数学分析讲义: 第五章 一元微分学

讲课教材: 《数学分析讲义》陈天权编著, 共三册, 北京大学出版社

参考书: 《数学分析》(第4版) 第一卷、第二卷, V.A. Zorich (卓里奇) 编著, 中译本, 高等教育出版社; 《数学分析》徐森林、薛春华编著, 共三册，清华大学出版社; 《数学分析教程》上下册, 常庚哲、史济怀编著, 中科大出版社

清华大学数学科学系 2016 级数学分析教学 卢旭光 Sept. 2016

数学分析讲义总目录

第五章　一元微分学

关于函数微分的概念、发展和微分的记号表示等见陈书第一册pp.161-168.

§5.1. 微分的概念, 可微函数

关键词: 高阶无穷小, 局部线性近似, 微分与切空间, 导数, 变化率.

【定义】设$E\subset\mathbb{R}$, $f:E\to\mathbb{C}$, $x_0\in E$ 是E的一个聚点. 若存在一个线性函数$h\mapsto A(x_0)h$ $(h\in\mathbb{R})$ 使得差(difference)

$$\alpha(x_0,h)=f(x_0+h)-f(x_0)-A(x_0)h\qquad(x_0+h\in E)$$

是关于h趋于零的高阶无穷小：$\alpha(x_0,h)=o(h)$ $(h\to0)$, 即

$$f(x_0+h)-f(x_0)=A(x_0)h+\alpha(x_0,h)\quad(x_0+h\in E)\text{ 且 }\lim_{h\to0}\frac{\alpha(x_0,h)}{h}=0$$

或等价地

$$f(x)-f(x_0)=A(x_0)(x-x_0)+\alpha(x_0,x-x_0)\quad(x\in E)\text{ 且 }\lim_{E\ni x\to x_0}\frac{\alpha(x_0,x-x_0)}{x-x_0}=0$$

则称f沿E在点x_0可微(differentiable), 同时称线性函数$h\mapsto A(x_0)h$ 是f 在x_0的微分(differentiation), 记之为$h\mapsto \mathrm{d}f(x_0)(h)$, 即

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

(**注意: 这个定义已经包含了x_0只是E的单侧聚点时的可微性, 例如若x_0是E的右聚点而非左聚点, 则高阶无穷小$\alpha(x_0,x-x_0)=o(x-x_0)$ 便只能是相对于右极限$E\ni x\to x_0+$而言的.**)

如果f 在E 的每个属于E的聚点处都可微, 则称f在E上处处可微, 也称f在E上可微. □

与函数的极限、连续等一样, 微分也是点点定义的, 属于函数的局部行为. 例如对任意(小的)$\delta>0$, 函数f 与函数$f|_{E\cap(x_0-\delta,x_0+\delta)}$在点$x_0$处有相同的微分.

导数与变化率. 由可微定义中的高阶无穷小$\alpha(x_0,x-x_0)=o(x-x_0)$ 易见当$E\ni x\to x_0$ 时差商$\frac{f(x)-f(x_0)}{x-x_0}$有极限且

$$\lim_{E\ni x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=A(x_0)+\lim_{E\ni x\to x_0}\frac{\alpha(x_0,x-x_0)}{x-x_0}=A(x_0).$$

这同时表明微分$h \mapsto A(x_0)h$ 中的$A(x_0)$ 由f 和x_0 唯一确定. 我们称$A(x_0)$为f沿E在x_0的导数(derivative), 记作$A(x_0) = f'(x_0)$, 即

$$f'(x_0) := \lim_{E \ni x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \in \mathbb{R}. \tag{$*$}$$

单独来看$(*)$, 它就是函数f 沿E在点x_0的导数的定义.

由于导数是由极限$(*)$定义的, 因此如果x_0 是E的双侧聚点, 则导数$f'(x_0)$存在当且仅当两个单侧导数

$$f'_-(x_0) := \lim_{E \ni x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}, \quad f'_+(x_0) := \lim_{E \ni x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}$$

皆存在有限且相等: $f'_-(x_0) = f'_+(x_0) \in \mathbb{R}$. 此时二者的公共值即为$f$沿$E$在$x_0$的导数:

$$f'(x_0) = f'_-(x_0) = f'_+(x_0).$$

如果x_0只是E的单侧聚点, 例如x_0是E的右聚点而非左聚点, 则当且仅当f在x_0的右导数$f'_+(x_0)$存在有限时,导数$f'(x_0)$存在. 此时$f'(x_0) = f'_+(x_0)$

导数$f'(x_0)$作为两个无穷小量之比的极限[1], 这一特征被莱布尼兹用优美且科学的记号记作

$$f'(x_0) = \frac{\mathrm{d}f}{\mathrm{d}x}(x_0).$$

从数值上看,导数$f'(x_0)$的绝对值$|f'(x_0)|$给出了函数$f(x)$在x_0 附近的变化率(放缩率):

$$\lim_{E \ni x \to x_0} \frac{|f(x) - f(x_0)|}{|x - x_0|} = |f'(x_0)|.$$

当$|f'(x_0)| > 1$时(相应地$|f'(x_0)| < 1$时), $|f(x) - f(x_0)|$ 在x_0附近是对于变量$|x - x_0|$的“放大镜”(“缩小镜”), 放大(缩小)倍数为$|f'(x_0)|$.

让我们把记号$A = f'(x_0)$ 带回到微分的定义中而将可微性可重写为

$$f(x_0 + h) - f(x_0) = f'(x_0)h + o(h) \qquad (x_0 + h \in E)$$

或等价地在x_0 附近

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + o(x - x_0) \qquad (x \in E).$$

[1]不难看出, f 在x_0 可微蕴含f首先在x_0 连续. 因此$f(x) - f(x_0)$ 是关于$x - x_0 \to 0$的无穷小量.

如果E的每一点都是E的聚点, 例如E 是区间, 而f 在E 上处处可微, 则$f'(x)$便是E 上的由f导出的一个新函数, 称之为f的导函数[导数derivative 原本的意思就是衍生物,派生物] 也记作

$$f'(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x), \quad x \in E \qquad \text{或} \qquad f' = \frac{\mathrm{d}f}{\mathrm{d}x}.$$

关于微分的表示. 上面已定义了f 沿E在点x_0的微分$\mathrm{d}f(x_0)(h) = A(x_0)h$,它是$h \in \mathbb{R}$的一个线性映射, 同时阐明了$A(x_0) = f'(x_0)$. 于是微分便可写成

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

如果不强调x_0是定点, 也即将x_0 用一般点的记号x表示, 那么微分便写成

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

这就是微分的现代主要表示法之一, 它的好处是你能直接看到它是h的线性函数. 另一方面注意到如果$f(x) \equiv x$, 则$f'(x) \equiv 1$, 因而有

$$\mathrm{d}x(h) = 1 \cdot h = h, \quad h \in \mathbb{R}.$$

代入上面即得

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

如果将其写为哑元形式, 即只保留映射关系而略去自变量h, 则写成

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

其中$\mathrm{d}f(x)$ 与$\mathrm{d}x$的共同点是二者都是h的线性映射！不同的是$\mathrm{d}f(x)$中的x 是E中的点, 可以是具体指定的点, 如x_0; 而$\mathrm{d}x$中的x 则只是提示我们: $\mathrm{d}x$ 是恒同映射$\mathrm{e}(x) \equiv x$ 的微分. 固然我们可以将$\mathrm{d}x$明确地写为$\mathrm{de}(x)$ 从而可以代入x的具体值x_0, 如$\mathrm{de}(x_0)$, 但一来是这样做费解, 二来是与自变量的一般记号x的意义相背：记号x不仅可以取具体值而且也提示读者它代表以x为自变量的恒同映射: $x \mapsto x$. 所以人们还是使用$\mathrm{d}x$ 而不是$\mathrm{de}(x)$. 这一点清楚后, 我们就知道下面的写法不妥

$$\text{“}\ \mathrm{d}f(a) = f'(a)\mathrm{d}a, \quad \mathrm{d}f(0.7) = f'(0.7)\mathrm{d}0.7\ \text{”}$$

正确的表达应为

$$\mathrm{d}f(a) = f'(a)\mathrm{d}x, \quad \mathrm{d}f(0.7) = f'(0.7)\mathrm{d}x, \quad \text{etc.}$$

它们与莱布尼兹记号是一致的(形式地把$\mathrm{d}x$除到左端):

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)=f'(x_0),\quad \frac{\mathrm{d}f}{\mathrm{d}x}(0.7)=f'(0.7),\quad \text{etc.}$$

在概念上还要注意: 作为“$\frac{0}{0}$”型极限的结果, $\frac{\mathrm{d}}{\mathrm{d}x}f(x)$中的“比值”$\frac{\mathrm{d}}{\mathrm{d}x}$是一个整体, 它是作用在$f$ 上的微商算子. 但是为了记住某些公式, 或在问题研究阶段做形式运算时, 把$\frac{\mathrm{d}f(x)}{\mathrm{d}x}$, $\frac{\mathrm{d}y(x)}{\mathrm{d}x}$等等当做分式, 也即让分子分母分别参与某些运算, 也是可以的, 只是在获得结果后别忘了做严格论证(有些结果确实只是形式上对, 实际上错).

一位法国著名数学家认为: 长期以来由于教学过于简化, 很多人只专注导数而没有重视微分, 特别是没有认识到微分是线性映射, 因此在后续的学习和研究中不得不补课. 另一位更著名的数学家威廉·瑟斯顿(美国, Fields 奖获得者) 则对导数和微分的概念给出六种解读, 包括速度,变化率和上面的观点. 我相信瑟斯顿的意思是, 这些不完全一样的认知各有各的侧重点和用途, 不必看轻那一方. 威廉·瑟斯顿本人是微分几何大专家, 他对微分的认识最具权威. 我们还是返回到那句话: 创造不是考逻辑; 靠谱的感知比仅是逻辑正确更好. 国际上不少好的教材既强调了微分的一般的现代观点, 又通过大量物理力学例子阐明了微分的实用性.

切空间. 回到函数在定点处的可微性, 按照现代数学的观点, 微分或可微的概念比导数或可导的概念更能反映函数的局部性质(高维空间尤其如此): 它是说函数的局部行为与线性函数几乎一样, 这里“几乎”就是那个高阶无穷小！所谓“切”、“贴切”指的就是这种局部线性近似. 因此我们把一维线性空间

$$\{f'(x_0)h \,|\, h\in\mathbb{R}\}$$

叫做曲线$\{(x,f(x))\,|\,x\in E\}$ 在 x_0 处的切空间(tangent space). 数学上的名词都是顾名思义的: 这空间实际就是把点$(x_0,f(x_0))$处的切线 $y-f(x_0)=f'(x_0)(x-x_0)$ 平移到过原点 $(0,0)$的直线 $\xi=f'(x_0)h$. 这两条直线只相差一个平移, 但直线$\{(h,f'(x_0)h)\,|\,h\in\mathbb{R}\}$ 含有零向量, 故它是线性空间. 由于两个线性空间

$$\{(h,f'(x_0)h)\,|\,h\in\mathbb{R}\}\quad 与\quad \{f'(x_0)h\,|\,h\in\mathbb{R}\}\quad 互相唯一决定$$

故人们习惯上称后者为 f 在 x_0 处的切空间.

继续: 如果f在E 上处处可微, 则导函数f' 在E 上就处处有定义了. 这时可以继续考察f' 的可微性. 如果f' 在点x_0 可微, 则称导数$(f')'(x_0)$为f 在点x_0 的二阶导数, 记作

$$f''(x_0)=(f')'(x_0)=\frac{\mathrm{d}f'}{\mathrm{d}x}(x_0).$$

注意, 根据微分的定义, 这意味着为了f在x_0有二阶导数, f 必须在x_0**附近** 处处有一阶导数. 如果f在E 上处处有二阶导数$f''(x)$, 则称f''为f在E上的二阶导函数, 记作

$$f''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\mathrm{d}f}{\mathrm{d}x}\Big)(x) = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x), \quad x \in E \qquad \text{或} \qquad f'' = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}.$$

归纳地我们可以定义f在x_0的n 阶导数

$$f^{(n)}(x_0) = \frac{\mathrm{d}^n f}{\mathrm{d}x^n}(x_0) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\mathrm{d}^{n-1} f}{\mathrm{d}x^{n-1}}\Big)(x_0)$$

只要f 在x_0附近的$n-1$阶导数$\frac{\mathrm{d}^{n-1}f}{\mathrm{d}x^{n-1}}(x)$处处存在.

关于高阶微分、高阶导数的一般观点　这方面可见陈书第一册pp.192-195. 这里我们只做一个直观但不失本质的解释. 以二阶微分、二阶导数为例.

设$f: E \to \mathbb{C}$. 定义一阶差分和二阶差分如下：　(总假设$x, x+h, x+k, x+h+k \in E$)

$$\begin{aligned}
&\Delta f(x,h) := f(x+h) - f(x),\\
&\Delta^2 f(x,h,k) := \Delta f(x+k,h) - \Delta f(x,h)\\
&\qquad\qquad\qquad = f(x+k+h) - f(x+k) - (f(x+h) - f(x)).
\end{aligned}$$

现在假定f在E上处处二次可微. 以下用α, β, γ 等表示关于$h \to 0, k \to 0$ 的“微分小”的量. 则至少形式地有

$$\begin{aligned}
\Delta f(x,h) &= f(x+h) - f(x) = f'(x)h + \alpha(x,h)h,\\
\Delta^2 f(x,h,k) &= \Delta f(x+k,h) - \Delta f(x,h)\\
&= f(x+k+h) - f(x+k) - (f(x+h) - f(x))\\
&= f'(x+k)h + \alpha(x+k,h)h - f'(x)h - \alpha(x,h)h\\
&= \Big(f'(x+k) - f'(x)\Big)h + \Big(\alpha(x+k,h) - \alpha(x,h)\Big)h\\
&= \Big(f''(x)k + \beta(x,k)k\Big)h + \widetilde{\alpha}(x,k,h)kh\\
&= f''(x)kh + \beta(x,k)kh + \widetilde{\alpha}(x,k,h)kh\\
&= f''(x)kh + \beta(x,k)kh + \widetilde{\alpha}(x,k,h)kh\\
&= f''(x)hk + \gamma(x,h,k)hk
\end{aligned}$$

其中

$$\gamma(x,h,k) \to 0 \ \ (\sqrt{h^2+k^2} \to 0) \quad \text{因此} \quad \gamma(x,h,k)hk = o(h^2+k^2) \quad (\sqrt{h^2+k^2} \to 0\,).$$

用高阶无穷小, 二阶差分可写成

$$\Delta^2 f(x,h,k) = f''(x)hk + o(h^2+k^2) \qquad (\sqrt{h^2+k^2}\to 0\,).$$

大家看到,

一阶差分的主部为线性函数: $\quad h \mapsto \mathrm{d}f(x)(h) := f'(x)h, \quad h\in\mathbb{R}$

二阶差分的主部为双线性函数: $\quad (h,k)\mapsto \mathrm{d}^2 f(x)(h,k) := f''(x)hk, \quad (h,k)\in\mathbb{R}^2.$

这就是一阶微分和二阶微分的主要概念, 其他就是(次要的)记号表示问题了.

复值、实值函数的微分的关系, 向量值函数的微分

对于实变复值函数$f = f_1 + \mathrm{i}f_2 : E\subset\mathbb{R}\to\mathbb{C}$, 其中$f_1, f_2$ 为实值函数, 由"复数极限存在当且仅当其实部和虚部的极限都存在"以及简单的加法关系可知: f 在$x\in E$可微当且仅当f 的实部虚部f_1, f_2 都在x 可微. 而当可微时有

$$f'(x) = f_1'(x) + \mathrm{i}f_2'(x).$$

同理, 对于向量值函数

$$\mathbf{f}(x) = (f_1(x), f_2(x), ..., f_m(x))$$

我们固然可以用$\mathbf{f}$的每个坐标函数$f_i(x)$的可微性来定义$\mathbf{f}$的可微性, 但这样做在观念上没有反映出微分的实质. 正确的做法还是用微分的本质— 局部线性近似— 来定义. 这一点在下学期学习多变量映射的微分以及在以后学习无穷维空间上一般映射的微分时会看得更清楚.

【定义】设$E\subset\mathbb{R}$, $\mathbf{f} = (f_1, f_2, ..., f_m) : E\to\mathbb{R}^m$, 并设$x\in E$ 为E的一个聚点. 若存在一个线性映射$h\mapsto \mathbf{A}(x)h$, $h\in\mathbb{R}$ (这里$\mathbf{A}(x) = (A_1(x), A_2(x), ..., A_m(x))\in\mathbb{R}^m$是向量) 使得

$$\mathbf{f}(x+h) - \mathbf{f}(x) = \mathbf{A}(x)h + \alpha(x,h) \qquad (x+h\in E)$$

其中$\alpha(x,h) = o(h)\in\mathbb{R}^m$ 是关于$h\to 0$的高阶无穷小[2], 或等价地

$$\mathbf{f}(y) - \mathbf{f}(x) = \mathbf{A}(x)(y-x) + \alpha(x, y-x), \quad y\in E$$

则称$\mathbf{f}$沿E在点x可微, 同时称线性映射$h\mapsto\mathbf{A}(x)h$ 为$\mathbf{f}$ 在点x的微分.

[2]即向量$\alpha(x,h)$的模长$|\alpha(x,h)|$ 是$h\,(\to 0)$的高阶无穷小.

如果$\mathbf{f}$ 在E 的每个属于E的聚点处都可微, 则称$\mathbf{f}$在E上处处可微, 也称$\mathbf{f}$在E 上可微. $\square$

与数值函数情形一样, 微分$h \mapsto \mathbf{A}(x)h$ 中的$\mathbf{A}(x)$ 由$\mathbf{f}$ 和x 唯一确定, 称它为$\mathbf{f}$ 在x的导数, 记作$\mathbf{f}'(x)$ 或$\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x}(x)$, 即

$$\mathbf{A}(x) = \mathbf{f}'(x) = \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x}(x) = \lim_{E \ni y \to x} \frac{\mathbf{f}(y) - \mathbf{f}(x)}{y - x}.$$

把记号$\mathbf{A}(x) = \mathbf{f}'(x)$ 带回到微分的定义中则上面可重写为

$$\mathbf{f}(x+h) - \mathbf{f}(x) = \mathbf{f}'(x)h + o(h) \qquad (x+h \in E)$$

或等价地

$$\mathbf{f}(y) - \mathbf{f}(x) = \mathbf{f}'(x)(y-x) + o(y-x) \qquad (y \in E).$$

如上, 我们称线性空间

$$\{\mathbf{f}'(x)h \,|\, h \in \mathbb{R}\}$$

为曲线$\{(x, \mathbf{f}(x)) \,|\, x \in E\}$ 在 x 处的切空间. 这空间仍是一维的因为$\mathbf{f}$的自变量x从而变量h是一维的.

回忆: 在$\mathbb{R}^m$中, 两个向量相等被定义为这两个向量的坐标或分量都对应相等; 一族向量趋于零向量当且仅当这族向量的所有坐标或分量都趋于零. 例如对于向量$\alpha(x,h) = (\alpha_1(x,h), \alpha_2(x,h), ..., \alpha_m(x,h))$, 由

$$|\alpha(x,h)| := \sqrt{\alpha_1(x,h)^2 + \alpha_2(x,h)^2 + \cdots + \alpha_m(x,h)^2}$$

和双向控制关系

$$|\alpha_i(x,h)| \le |\alpha(x,h)| \le \sum_{j=1}^{m} |\alpha_j(x,h)|, \quad i = 1, 2, ..., m$$

易见

$$\lim_{h \to 0} \frac{|\alpha(x,h)|}{|h|} = 0 \Longleftrightarrow \lim_{h \to 0} \frac{|\alpha_i(x,h)|}{|h|} = 0, \quad i = 1, 2, ..., m.$$

换言之,

$$\alpha(x,h) = o(h) \iff \alpha_i(x,h) = o(h), \quad i = 1, 2, ..., m.$$

据此我们看到: $\mathbf{f}(x+h) - \mathbf{f}(x) = \mathbf{A}(x)h + \alpha(x,h)$, $\alpha(x,h) = o(h)$, 等价于

$$f_i(x+h) - f_i(x) = A_i(x)h + \alpha_i(x,h), \quad \alpha_i(x,h) = o(h), \quad i = 1, 2, ..., m.$$

这说明$\mathbf{f}$ 在x 可微$\Longleftrightarrow$ $\mathbf{f}$的每个坐标函数f_i 都在x 可微.

此外根据微分和导数的唯一性可知$f_i'(x) = A_i(x)\,(i = 1,2,...,m)$. 把它代入$\mathbf{f}'(x) = \mathbf{A}(x) = (A_1(x), A_2(x), ..., A_m(x))$中便得到

$$\mathbf{f}'(x) = (f_1'(x), f_2'(x), ..., f_m'(x)).$$

它表明$\mathbf{f}$ 在x 可导$\Longleftrightarrow$ $\mathbf{f}$的每个坐标函数f_i 都在x 可导.

当然, $\mathbf{f}$与它的坐标f_i的导数的关系也可直接从向量的运算规则

$$\frac{\mathbf{f}(y)-\mathbf{f}(x)}{y-x} = \left(\frac{f_1(y)-f_1(x)}{y-x}, \frac{f_2(y)-f_2(x)}{y-x}, ..., \frac{f_m(y)-f_m(x)}{y-x}\right)$$

直接推出.

如上, 如果E的每一点都是E的聚点(例如E 是区间), 而$\mathbf{f}$ 在E 上处处可微, 则称函数$x \mapsto \mathbf{f}'(x) = \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}x}(x)$为$\mathbf{f}$的导函数. 进一步, 如果导函数$\mathbf{f}'$ 在点x 可微, 则称$\mathbf{f}'$的导数$(\mathbf{f}')'(x)$为$\mathbf{f}$ 在点x的二阶导数, 记作

$$\mathbf{f}''(x) = (\mathbf{f}')'(x) = \frac{\mathrm{d}\mathbf{f}'}{\mathrm{d}x}(x) = \frac{\mathrm{d}^2\mathbf{f}}{\mathrm{d}x^2}(x).$$

再次强调, 根据微分的定义, 这意味着为了$\mathbf{f}$在x有二阶导数, $\mathbf{f}$ 必须在x**附近** 处处有一阶导数.

将上面的结论用于导函数$\mathbf{f}'$ 可知

$$\mathbf{f}''(x) = (f_1''(x), f_2''(x), ..., f_m''(x)).$$

这同时说明: $\mathbf{f}$ 在x两次可导$\Longleftrightarrow$ $\mathbf{f}$ 的每个坐标函数f_i 都在x两次可导. 等等.

微分的几何意义. 设实值函数$f(x)$在x_0 可微, 令$y_0 = f(x_0)$. 则直线方程

$$y - y_0 = f'(x_0)(x - x_0), \quad x \in \mathbb{R}$$

是曲线$y = f(x)$ 在点(x_0, y_0) 处的切线或切线方程. See a graph.

按照解析几何, 即便$f'(x_0) = 0$, 切线方程的标准形式也是有意义(有定义)的:

$$\frac{x - x_0}{1} = \frac{y - y_0}{f'(x_0)}$$

由这个标准式可知曲线在点(x_0, y_0)处的切向量为$(1, f'(x_0))$, 其单位法化为$\frac{(1,f'(x_0))}{\sqrt{1+[f'(x_0)]^2}}$. 如果用参数来表达切线, 即令

$$\frac{x - x_0}{1} = \frac{y - y_0}{f'(x_0)} = t,$$

则得到该切线的参数表示:

$$x = x(t) = x_0 + t, \quad y = y(t) = y_0 + f'(x_0)t, \quad t \in \mathbb{R}.$$

它确是直线因为它关于t的导数是常值:

$$\frac{\mathrm{d}}{\mathrm{d}t}(x(t), y(t)) = (x'(t), y'(t)) \equiv (1, f'(x_0)).$$

曲线$y = f(x)$在点(x_0, y_0)处的法线方程: 如果把曲线在这点的切线方程 $y - y_0 = f'(x_0)(x - x_0)$用内积表示, 则有

$$\Big\langle (x - x_0, y - y_0), (-f'(x_0), 1) \Big\rangle = 0, \quad \text{它等于垂直关系} \quad (x - x_0, y - y_0) \perp (-f'(x_0), 1).$$

由此我们看到经过点(x_0, y_0)的切线的法线方程为

$$\frac{x - x_0}{-f'(x_0)} = \frac{y - y_0}{1} \qquad \text{即} \quad y - y_0 = -\frac{1}{f'(x_0)}(x - x_0)$$

它被定义为曲线$y = f(x)$在点(x_0, y_0)处的法线方程. 据此我们还得到

曲线在点(x_0, y_0)处的法向量为 $(-f'(x_0), 1)$, 单位法向量为 $\dfrac{(-f'(x_0), 1)}{\sqrt{1 + [f'(x_0)]^2}}$.

(图示)

【例】设$\alpha > 0$, $E \subset \mathbb{R}$, $x_0 \in E$ 为E 的聚点, $f(x) = |x - x_0|^\alpha$, $x \in E$.

(1) 若$0 < \alpha < 1$, 则

$$\lim_{E \ni x \to x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} \right| = \lim_{E \ni x \to x_0} \frac{1}{|x - x_0|^{1-\alpha}} = +\infty.$$

因此函数$f(x) = |x - x_0|^\alpha$在x_0不可导从而在x_0 不可微. (图示: 触及x-轴的尖点.)

(2) 若$\alpha = 1, E = [x_0 - 1, x_0 + 1]$, 则函数$f(x) = |x - x_0|$ 在x_0的左右导数

$$f'_-(x_0) = \lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0} = -1, \quad f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} = 1$$

不相等, 因此f在x_0不可导, 从而不可微. 事实上从图像上看, 曲线$y = |x - x_0|$ 是一条折线; 任何经过折点$(x_0, 0)$ 的直线都不可能在x_0 的左邻域$(x_0 - \delta, x_0]$和右邻域$[x_0, x_0 + \delta)$与此折线**同时**有高阶无穷小近似(图示).

(3) 若$\alpha=1, E=[x_0-1,x_0]$ 或$E=[x_0,x_0+1]$, 则函数$f(x)=|x-x_0|$沿E在x_0可微, 这是因为x_0只是E的单侧聚点. 事实上例如若$E=[x_0,x_0+1]$ 则$E\cap(x_0-1,x_0+1)=E\cap[x_0,x_0+1)$ 从而有$f(x)-f(x_0)=1\cdot(x-x_0)+0$ for all $x\in E\cap(x_0-1,x_0+1)$. 由微分的定义和微分的唯一性知f沿E在x_0 可微, 其中高阶无穷小$\alpha(x_0,x-x_0)\equiv 0$, $f'(x_0)=f'_+(x_0)=1$.

(4) 若$\alpha>1$, $E=[x_0-1,x_0+1]$, 则函数$f(x)=|x-x_0|^\alpha$本身便是$x-x_0$的高阶无穷小: $f(x)-f(x_0)=0\cdot(x-x_0)+|x-x_0|^\alpha$, 故$f$ 在x_0可微. 根据微分的唯一性或根据导数的定义知$f'(x_0)=0$. 几何上看, 曲线$y=|x-x_0|^\alpha$ 在点$(x_0,0)$的切线为x-轴. □

微分的运动学意义. 设一质点的运动轨迹 (即由质点的位移形成的空间曲线) 为

$$\mathbf{r}(t)=(x(t),y(t),z(t)),\quad t\in I.$$

如令$\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 是欧空间$\mathbb{R}^3$的标准正交基, 即

$$\mathbf{e}_1=(1,0,0),\quad \mathbf{e}_2=(0,1,0),\quad \mathbf{e}_3=(0,0,1)$$

则$\mathbf{r}(t)$ 可表示成三个垂直方向运动的叠加:

$$\mathbf{r}(t)=x(t)\mathbf{e}_1+y(t)\mathbf{e}_2+z(t)\mathbf{e}_3.$$

质点在时刻t和$t+h$之间的平均速度为

$$\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}=\frac{x(t+h)-x(t)}{h}\mathbf{e}_1+\frac{y(t+h)-y(t)}{h}\mathbf{e}_2+\frac{z(t+h)-z(t)}{h}\mathbf{e}_3.$$

假设质点的运动不会在某一时刻猛然转向, 即质点在任意时刻存在确定的速度, 那么令$h\to 0$便得到质点在时刻 t的速度为

$$\begin{aligned}\mathbf{v}(t)=\mathbf{r}'(t)&=\lim_{h\to0}\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}\\&=\lim_{h\to0}\frac{x(t+h)-x(t)}{h}\mathbf{e}_1+\lim_{h\to0}\frac{y(t+h)-y(t)}{h}\mathbf{e}_2+\lim_{h\to0}\frac{z(t+h)-z(t)}{h}\mathbf{e}_3\\&=x'(t)\mathbf{e}_1+y'(t)\mathbf{e}_2+z'(t)\mathbf{e}_3\\&=(x'(t),y'(t),z'(t)).\end{aligned}$$

如令

$$\alpha(t,h)=\mathbf{r}(t+h)-\mathbf{r}(t)-\mathbf{r}'(t)h$$

则由上述极限有

$$\frac{\alpha(t,h)}{h}=\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}-\mathbf{r}'(t)\to 0\quad(h\to 0)$$

因此$\alpha(t,h)=o(h)$ 是h的高阶无穷小. 于是得到

$$\mathbf{r}(t+h)-\mathbf{r}(t)=\mathbf{r}'(t)h+o(h).$$

这表明, 质点的运动轨迹在时刻t是可微的, 也即在时刻t的一个很小的邻域$(t-\delta,t+\delta)$, 内质点的运动与以$\mathbf{r}'(t)$为速度的匀速直线运动$h\mapsto\mathbf{r}(t)+\mathbf{r}'(t)h$几乎相同: 其差别是$h$的高阶无穷小. 这一观察很重要, 因为根据它我们就可以通过质点的速度信息确定质点的位移. 例如为了确定质点从$t=0$到$t=T\,(>0)$这段时间的位移$\mathbf{r}(T)-\mathbf{r}(0)$, 我们把总时间$[0,T]$分割成一些微小的时间段:

$$[0,T]=\bigcup_{k=0}^{n-1}\Big[\frac{k}{n}T,\ \frac{k+1}{n}T\Big],\qquad(n>>1)$$

质点在$t\in[\frac{k}{n}T,\frac{k+1}{n}T]$ 这段近似于匀速直线运动:

$$\mathbf{r}(t)=\mathbf{r}\Big(\frac{k}{n}T\Big)+\mathbf{r}'\Big(\frac{k}{n}T\Big)\Big(t-\frac{k}{n}T\Big)+\alpha\Big(\frac{k}{n}T,t-\frac{k}{n}T\Big)$$

其中$\alpha\Big(\frac{k}{n}T,h\Big)$ 是$h\to 0$的高阶无穷小. 特别有

$$\mathbf{r}\Big(\frac{k+1}{n}T\Big)-\mathbf{r}\Big(\frac{k}{n}T\Big)=\mathbf{r}'\Big(\frac{k}{n}T\Big)\frac{T}{n}+\alpha\Big(\frac{k}{n}T,\frac{T}{n}\Big).$$

因此质点从$t=0$到$t=T$的位移可计算为

$$\mathbf{r}(T)-\mathbf{r}(0)=\sum_{k=0}^{n-1}\Big(\mathbf{r}\Big(\frac{k+1}{n}T\Big)-\mathbf{r}\Big(\frac{k}{n}T\Big)\Big)=\sum_{k=0}^{n-1}\mathbf{r}'\Big(\frac{k}{n}T\Big)\frac{T}{n}+\sum_{k=0}^{n-1}\alpha\Big(\frac{k}{n}T,\frac{T}{n}\Big).$$

进一步假设质点的速度$t\mapsto\mathbf{r}'(t)$ 连续. 则从后面将要学习的积分理论可知

$$\lim_{n\to\infty}\sum_{k=0}^{n-1}\mathbf{r}'\Big(\frac{k}{n}T\Big)\frac{T}{n}=\int_0^T\mathbf{r}'(t)\mathrm{d}t.$$

另一方面可以证明$\max\limits_{0\le k\le n-1}|\alpha\Big(\frac{k}{n}T,\frac{T}{n}\Big)|\le o(\frac{T}{n})$. 因此

$$\left|\sum_{k=0}^{n-1}\alpha\Big(\frac{k}{n}T,\frac{T}{n}\Big)\right|\le n\cdot o(\frac{T}{n})=T\frac{o(\frac{T}{n})}{\frac{T}{n}}\to 0\quad(n\to\infty)$$

换言之, 上面众多小时间段中产生的误差, 加起来还是无穷小. 于是令$n\to\infty$ 得到

$$\mathbf{r}(T)-\mathbf{r}(0)=\int_0^T\mathbf{r}'(t)\mathrm{d}t.$$

这等式表明, 知道了速度和时间段即可求出位移. 这就是著名的牛顿-莱布尼茨公式.

如用向量的叠加表示, 上述积分等式也可写成

$$(x(T)-x(0))\mathbf{e}_1+(y(T)-y(0))\mathbf{e}_2+(z(T)-z(0))\mathbf{e}_3$$
$$=\int_0^T x'(t)\mathrm{d}t\mathbf{e}_1+\int_0^T y'(t)\mathrm{d}t\mathbf{e}_2+\int_0^T z'(t)\mathrm{d}t\mathbf{e}_3$$

它等价于对应的分量相等:

$$x(T)-x(0)=\int_0^T x'(t)\mathrm{d}t,\quad y(T)-y(0)=\int_0^T y'(t)\mathrm{d}t,\quad z(T)-z(0)=\int_0^T z'(t)\mathrm{d}t.$$

类似地, 质点在时刻t的加速度为

$$\mathbf{a}(t)=\mathbf{v}'(t)=\mathbf{r}''(t)=x''(t)\mathbf{e}_1+y''(t)\mathbf{e}_2+z''(t)\mathbf{e}_3=(x''(t),y''(t),z''(t)).$$

设质点在时刻 t 受到的外力的合力为 $\mathbf{F}(t)$, 则由牛顿第二定律有

$$\mathbf{F}(t)=m\mathbf{a}(t)=m\mathbf{r}''(t)$$

其中 m 为质点的质量. 如上, 力$\mathbf{F}(t)$可以被表示成三个垂直分力的叠加(合成):

$$\mathbf{F}(t)=F_1(t)\mathbf{e}_1+F_2(t)\mathbf{e}_2+F_3(t)\mathbf{e}_2=(F_1(t),F_2(t),F_3(t)),$$

$$F_1(t)=mx''(t),\quad F_2(t)=my''(t),\quad F_3(t)=mz''(t).$$

可微与连续的关系.

从微分的定义和上面例子看到, 可微性是一个比连续性要求高得多的性质. 我们把这个事实列为一个命题.

【命题5.1.】 设$E\subset\mathbb{R}$, $f:E\to\mathbb{C}$, 并设$x_0\in E$ 是E的一个聚点. 若f在x_0可微, 则存在$\delta=\delta_{x_0}>0$ 和$L=L_{x_0}>0$ 使得

$$|f(x)-f(x_0)|\le L|x-x_0|\qquad \text{for all }\ x\in E\cap(x_0-\delta,x_0+\delta)$$

特别可知f在x_0 连续. 反之, 若f在x_0 连续, 则f远不必在x_0 可微.

【证】设f在x_0可微. 由定义知

$$\alpha(x_0, x-x_0) = f(x) - f(x_0) - f'(x_0)(x-x_0)$$

是$x-x_0$的高阶无穷小, 即$|\alpha(x_0,x-x_0)|/|x-x_0| \to 0$ as $E \ni x \to x_0$. 因此存在$\delta > 0$使得当$x \in E\cap(x_0-\delta, x_0+\delta)$时$|\alpha(x_0,x-x_0)| \le |x-x_0|$. 令$L = |f'(x_0)|+1$. 则当$x \in E\cap(x_0-\delta, x_0+\delta)$ 时有

$$|f(x)-f(x_0)| \le |f'(x_0)(x-x_0)| + |\alpha(x_0,x-x_0)| \le L|x-x_0|.$$

反之, 设$0<\alpha<1$, $f(x) = |x-x_0|^\alpha$, 则f在x_0 连续但不可微. □

一元函数的可微等价于可导.

在后面学习多元函数微分学时将会继续看到, 函数的可微性与否是一个局部的整体性质, 即它描述的是在由多元函数给出的曲面上的一点处是否存在切平面的问题; 大量例子表明多元函数关于自变量的分量的导数即便存在, 也未必能保证切平面存在. **但是在一元函数的情形, 也即对于切线问题而言, 可微与可导是等价的:**

【命题5.2.】设$E\subset\mathbb{R}$, $f: E\to\mathbb{C}$, 并设$x_0\in E$ 是E的一个聚点. 则f在x_0可微$\Longleftrightarrow$ f在x_0可导.

【证】上面已证明了若f在x_0可微, 则f在x_0可导.

反之设f在x_0可导. 令$\alpha(x_0,x-x_0) = f(x)-f(x_0)-f'(x_0)(x-x_0)$, 则由$f$在$x_0$可导知

$$\lim_{E\ni x\to x_0}\frac{|\alpha(x_0,x-x_0)|}{|x-x_0|} = \lim_{E\ni x\to x_0}\left|\frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)\right| = 0$$

所以$\alpha(x_0,x-x_0)$ 是$x-x_0\to 0$的高阶无穷小. 于是再由表达式$f(x)-f(x_0) = f'(x_0)(x-x_0)+\alpha(x_0,x-x_0)$ 即知f在x_0可微. □

【例】设函数f在区间I 上有定义且在点$x_0\in I$ 可微, 设数列$a_n, b_n\in I$ 满足$a_n\le x_0\le b_n$, $n=1,2,3,\dots$, $0<b_n-a_n\to 0\,(n\to\infty)$. 则有

$$\lim_{n\to\infty}\frac{f(b_n)-f(a_n)}{b_n-a_n} = f'(x_0).$$

【证】[草稿分析: 如果用导数=差商极限来证, 则必要讨论分母可能为零的情形, 即涉及忽而$a_n = x_0$, 忽而$b_n = x_0$ 等琐碎的情形. 能够避免这些麻烦的简单做法就是直接用可微的定义, 即不是从除法出发而是从乘法和高阶无穷小出发.]

由f在x_0可微和$b_n \to x_0, a_n \to x_0$ $(n \to \infty)$ 知

$$f(b_n) = f(x_0) + f'(x_0)(b_n - x_0) + \beta_n |b_n - x_0|,$$

$$f(a_n) = f(x_0) + f'(x_0)(a_n - x_0) + \gamma_n |a_n - x_0|$$

其中$\beta_n \to 0, \gamma_n \to 0$ $(n \to \infty)$. 两式相减并除以$b_n - a_n$得到

$$\frac{f(b_n) - f(a_n)}{b_n - a_n} = f'(x_0) + \frac{\beta_n |b_n - x_0| - \gamma_n |a_n - x_0|}{b_n - a_n}.$$

因$a_n \le x_0 \le b_n$, 故

$$\begin{aligned}\Big|\frac{\beta_n |b_n - x_0| - \gamma_n |a_n - x_0|}{b_n - a_n}\Big| &\le \frac{|\beta_n|(b_n - x_0) + |\gamma_n|(x_0 - a_n)}{b_n - a_n} \\ &\le \max\{|\beta_n|, |\gamma_n|\} \frac{b_n - x_0 + x_0 - a_n}{b_n - a_n} = \max\{|\beta_n|, |\gamma_n|\} \to 0 \quad (n \to \infty).\end{aligned}$$

所以$\lim\limits_{n\to\infty} \frac{f(b_n)-f(a_n)}{b_n-a_n} = f'(x_0)$. $\square$

§5.2. 导数与微分的运算规则

1. 四则运算

【定理 5.3. 微分的四则运算】设 $f, g : E \to \mathbb{C}$ 在 $x \in E$ 可微. 则 $\alpha f + \beta g, fg$ 也在 x 可微, 其中α, β 为常数. 又若$g(x) \neq 0$, 则f/g 也在x 可微. 此外有

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x),$$
$$(fg)'(x) = f'(x)g(x) + f(x)g'(x),$$

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \quad \text{分母 } g(x) \neq 0.$$

用莱布尼兹符号,上面写成

$$\frac{\mathrm{d}}{\mathrm{d}x}(\alpha f + \beta g)(x) = \alpha \frac{\mathrm{d}f(x)}{\mathrm{d}x} + \beta \frac{\mathrm{d}g(x)}{\mathrm{d}x},$$

$$\frac{\mathrm{d}}{\mathrm{d}x}(fg)(x) = g(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x} + f(x)\frac{\mathrm{d}g(x)}{\mathrm{d}x},$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f}{g}\right)(x) = \frac{g(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x} - f(x)\frac{\mathrm{d}g(x)}{\mathrm{d}x}}{g(x)^2}$$

做分母的$g(x) \neq 0$.

[注]: 以上表明, 线性组合的求导运算就是通常的线性运算, 但关于乘除法的求导运算则是微分或导数运算的**特殊性质**!

【证】我们将使用微分的定义. 首先回忆高阶无穷小的定义:

$$\psi(h) = o(h) \ \ (h \to 0) \iff \lim_{h \to 0} \left| \frac{\psi(h)}{h} \right| = 0.$$

由假设知对任意$h \neq 0$满足$x + h \in E$ 有

$$f(x+h) - f(x) = f'(x)h + o(h), \quad g(x+h) - g(x) = g'(x)h + o(h) \quad (h \to 0).$$

因此

$$\begin{aligned}
&(\alpha f + \beta g)(x+h) - (\alpha f + \beta g)(x) \\
&= \alpha f(x+h) + \beta g(x+h) - \alpha f(x) - \beta g(x) \\
&= (\alpha f'(x) + \beta g'(x))h + \alpha o(h) + \beta o(h) \\
&= (\alpha f'(x) + \beta g'(x))h + o(h) \qquad (h \to 0).
\end{aligned}$$

所以由可微的定义知$\alpha f+\beta g$ 在点x可微且

$$(\alpha f+\beta g)'(x)=\alpha f'(x)+\beta g'(x).$$

类似地对于乘法有

$$\begin{aligned}&(fg)(x+h)-(fg)(x)=f(x+h)g(x+h)-f(x)g(x)\\&=\Big(f(x)+f'(x)h+o(h)\Big)\Big(g(x)+g'(x)h+o(h)\Big)-f(x)g(x)\\&=\Big(f'(x)g(x)+f(x)g'(x)\Big)h+f(x)o(h)+g(x)o(h)+o(h)o(h)\\&=\Big(f'(x)g(x)+f(x)g'(x)\Big)h+o(h)\qquad (h\to 0).\end{aligned}$$

所以由可微的定义知fg 在点x可微且

$$(fg)'(x)=f'(x)g(x)+f(x)g'(x).$$

对于除法, 先看$1/g$的可微性. 由g在点x可微知g在点x 连续, 因此由$g(x)\neq 0$ 有

$$\text{当}|h|<<1\text{ 时 }|g(x+h)|\geq\frac{|g(x)|}{2}>0;\quad \frac{1}{g(x+h)}-\frac{1}{g(x)}=o(1)\quad (h\to 0).$$

于是有

$$\begin{aligned}&\Big(\frac{1}{g}\Big)(x+h)-\Big(\frac{1}{g}\Big)(x)=\frac{1}{g(x+h)}-\frac{1}{g(x)}=\frac{g(x)-g(x+h)}{g(x+h)g(x)}\\&=-\frac{g'(x)h+o(h)}{g(x)g(x+h)}=-\frac{g'(x)h}{g(x)g(x+h)}+o(h).\end{aligned}$$

而

$$\frac{g'(x)h}{g(x)g(x+h)}=\frac{g'(x)}{(g(x))^2}h+\frac{g'(x)}{g(x)}\Big(\frac{1}{g(x+h)}-\frac{1}{g(x)}\Big)h=\frac{g'(x)}{(g(x))^2}h+o(h)$$

所以

$$\Big(\frac{1}{g}\Big)(x+h)-\Big(\frac{1}{g}\Big)(x)=-\frac{g'(x)}{(g(x))^2}h+o(h)\quad (h\to 0).$$

所以$1/g$ 在x可微且

$$\Big(\frac{1}{g}\Big)'(x)=-\frac{g'(x)}{(g(x))^2}.$$

最后应用乘法的结果即得

$$\Big(\frac{f}{g}\Big)'(x)=\Big(f\frac{1}{g}\Big)'(x)=f'(x)\frac{1}{g(x)}+f(x)\Big(\frac{1}{g}\Big)'(x)=\frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}.$$

□

对函数个数用归纳法易证下面

【定理 5.3的推论1】设$f_1, f_2, ..., f_n$ 都在其定义域中的点x处可微, $c_1, c_2, ..., c_n$ 为常数. 则$\sum\limits_{k=1}^{n} c_k f_k$, $\prod_{k=1}^{n} f_k$ 都在x可微且

$$\Big(\sum_{k=1}^{n} c_k f_k\Big)'(x) = \sum_{k=1}^{n} c_k f_k'(x),$$

$$\Big(\prod_{k=1}^{n} f_k\Big)'(x) = f_1'(x)f_2(x)\cdots f_n(x) + f_1(x)f_2'(x)f_3(x)\cdots f_n(x) + \cdots + f_1(x)\cdots f_{n-1}(x)f_n'(x)$$

$$= \sum_{j=1}^{n} f_j'(x) \prod_{1\le k\le n, k\ne j} f_k(x). \qquad \square$$

如果采用代数学中的做法, 上面这个乘积求和关系可以(至少形式地)写成具有结构意义的等式:

$$\Big(\prod_{k=1}^{n} f_k\Big)'(x) = \Big(\prod_{k=1}^{n} f_k(x)\Big)\sum_{k=1}^{n}\frac{f_k'(x)}{f_k(x)}$$

当然要注意哪些分母可能为零.

【例】设$P(x)$ 为一个n次多项式, $x_1, x_2, ..., x_n$ 是它的全部n个根,包括重根. 则(当x 不属于这些根时) 有

$$\frac{P'(x)}{P(x)} = \sum_{k=1}^{n}\frac{1}{x-x_k}.$$

这个关系式为研究多项式的性质提供可了方便.

【证】写$P(x) = a(x-x_1)(x-x_2)\cdots(x-x_n)$. 因$(x-x_k)' = 1$, 故由上面给出的函数乘积的导数计算公式知当$x$ 不属于$\{x_1, x_2, ..., x_n\}$ 也即当$P(x)\neq 0$ 时有

$$P'(x) = P(x)\sum_{k=1}^{n}\frac{1}{x-x_k}. \qquad \square$$

【定理5.3的推论2: 微分运算】设 $f, g: E\to\mathbb{C}$ 在 $x\in E$ 可微, α, β 为常数. 则

$$\mathrm{d}(\alpha f+\beta g)(x) = \alpha\mathrm{d}f(x) + \beta\mathrm{d}g(x),$$

$$\mathrm{d}(fg)(x) = g(x)\mathrm{d}f(x) + f(x)\mathrm{d}g(x).$$

又若$g(x)\neq 0$,

$$\mathrm{d}\Big(\frac{f}{g}\Big)(x) = \frac{g(x)\mathrm{d}f(x) - f(x)\mathrm{d}g(x)}{g(x)^2}.$$

【证】以除法为例. 注意在证明中x 是固定的. 由微分的定义: $\mathrm{d}\varphi(x)(h)=\varphi'(x)h$ 和导数运算的结果有

$$\begin{aligned}&\mathrm{d}\Big(\frac{f}{g}\Big)(x)(h)=(\frac{f}{g}\Big)(x)'h=\Big(\frac{g(x)f'(x)-f(x)g'(x)}{g(x)^2}\Big)h\\&=\frac{g(x)f'(x)h-f(x)g'(x)h}{g(x)^2}=\frac{g(x)\mathrm{d}f(x)(h)-f(x)\mathrm{d}g(x)(h)}{g(x)^2}\\&=\frac{g(x)\mathrm{d}f(x)-f(x)\mathrm{d}g(x)}{g(x)^2}(h)\qquad\forall\, h\in\mathbb{R}.\end{aligned}$$

根据两个函数相等的定义即知所证成立. □

【例】我们来看物体速度的定义在静态仿射坐标系下的不变性.

设在$\mathbb{R}^3$ 中有两个参考系, 二者都是$\mathbb{R}^3$的与时间无关的仿射坐标系. 不妨设其中一个是$\mathbb{R}^3$的标准直角坐标系(x,y,z), 另一个坐标系为$(\widetilde{x},\widetilde{y},\widetilde{z})$. 则由线代数知$(\widetilde{x},\widetilde{y},\widetilde{z})$与$(x,y,z)$ 的关系是仿线性关系(即线性变换加平移):

$$\widetilde{x}=a_{11}x+a_{12}y+a_{13}z+b_1,$$

$$\widetilde{y}=a_{21}x+a_{22}y+a_{23}z+b_2,$$

$$\widetilde{z}=a_{31}x+a_{32}y+a_{33}z+b_3.$$

为书写紧凑, 让我们使用向量和矩阵. 记$\mathbf{r}=(x,y,z)^T,\widetilde{\mathbf{r}}=(\widetilde{x},\widetilde{y},\widetilde{z})^T$, $\mathbf{b}=(b_1,b_2,b_3)^T$, $A=(a_{ij})_{3\times3}$, 这里自然要求A非奇异. 则两个坐标系的关系写为

$$\widetilde{\mathbf{r}}=A\mathbf{r}+\mathbf{b},\qquad \mathbf{r}=A^{-1}\widetilde{\mathbf{r}}+\widetilde{\mathbf{b}}$$

其中$\widetilde{\mathbf{b}}=-A^{-1}\mathbf{b}$. 设一物体(或质点)在$\mathbb{R}^3$的标准坐标系下的运动轨迹为$\mathbf{r}(t),t\in I$. 则这一轨迹在上述静态仿射坐标系下的表现为$\widetilde{\mathbf{r}}(t)=A\mathbf{r}(t)+\widetilde{\mathbf{b}}$. 根据这个线性关系可见二者在同一时间区间$[t,t+h]$上的平均速度的关系为

$$\frac{\widetilde{\mathbf{r}}(t+h)-\widetilde{\mathbf{r}}(t)}{h}=A\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}.$$

令$h\to 0$ 得到

$$\widetilde{\mathbf{r}}'(t)=A\mathbf{r}'(t),\quad \mathbf{r}'(t)=A^{-1}\widetilde{\mathbf{r}}'(t).$$

换言之, 因为$A,\mathbf{b}$ 与时间无关, 故

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{r}}(t)=\frac{\mathrm{d}}{\mathrm{d}t}\Big(A\mathbf{r}(t)+\mathbf{b}\Big)=A\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{r}(t).$$

从这一关系看到, 若物体(或质点)在某一静态坐标系下的速度确定了, 则它在任何其它静态坐标系下的速度也就确定了, 并且无论这速度是在新坐标系下主动定义(测量)的还是经变换被动定义(测量)的, 其速度定义的结构(或速度的测量方式)都是相同的, 只是比例系数矩阵A可能不同. 这相当于说在时刻t 附近(此时物体近于直线运动), 我们的三维眼睛从不同的出发点和不同的角度看物体运动, 会得出不同的速度值, 但这不同的值(即速度向量)之间有着确定的“比例”关系或换算关系, 除此之外, 没有别的不同. $\square$

【定理5.4. 基本初等函数的导数】设x分别属于下面的函数和对应的等式右端的函数的公共定义域. 则有

$$\frac{\mathrm{d}x^\alpha}{\mathrm{d}x}=\alpha x^{\alpha-1},\qquad \text{即}\quad (x^\alpha)'=\alpha x^{\alpha-1}$$
$$\frac{\mathrm{d}e^x}{\mathrm{d}x}=e^x,\qquad \text{即}\quad (e^x)'=e^x,$$
$$\frac{\mathrm{d}\log x}{\mathrm{d}x}=\frac{1}{x},\qquad \text{即}\quad (\log x)'=\frac{1}{x},$$
$$\frac{\mathrm{d}\sin x}{\mathrm{d}x}=\cos x,\qquad \text{即}\quad (\sin x)'=\cos x,$$
$$\frac{\mathrm{d}\cos x}{\mathrm{d}x}=-\sin x,\qquad \text{即}\quad (\cos x)'=-\sin x.$$

【证】应用第三章后面例题中证明的几个特殊极限我们有:

对于幂函数,

$$(x^\alpha)'=\lim_{h\to0}\frac{(x+h)^\alpha-x^\alpha}{h}=x^{\alpha-1}\lim_{h\to0}\frac{(1+h/x)^\alpha-1}{h/x}=x^{\alpha-1}\lim_{y\to0}\frac{(1+y)^\alpha-1}{y}=\alpha x^{\alpha-1}.$$

易见当$x=0$ 时(这时要求$\alpha\ge1$), 结果也成立.

对于对数函数,

$$(\log x)'=\lim_{h\to0}\frac{\log(x+h)-\log x}{h}=\frac{1}{x}\lim_{h\to0}\frac{\log(1+h/x)}{h/x}=\frac{1}{x}\lim_{y\to0}\frac{\log(1+y)}{y}=\frac{1}{x}.$$

对于指数函数, 由幂级数展开式$e^z=\sum_{n=0}^{\infty}\frac{z^n}{n!}$ 易见

$$|e^z-1-z|\le\sum_{n=2}^{\infty}\frac{|z|^n}{n!}=|z|^2\sum_{n=0}^{\infty}\frac{|z|^n}{(n+2)!}\le\frac{|z|^2}{2}e^{|z|},\quad z\in\mathbb{C}$$

从而有

$$\lim_{z\to0}\frac{e^z-1}{z}=1.\tag{$*$}$$

因此
$$(e^z)'=\frac{\mathrm{d}}{\mathrm{d}z}e^z=\lim_{h\to 0}\frac{e^{z+h}-e^z}{h}=e^z\lim_{h\to 0}\frac{e^h-1}{h}=e^z,\quad z\in\mathbb{C}.$$
由$(*)$还有
$$\begin{aligned}(e^{\mathrm{i}x})'&=\frac{\mathrm{d}}{\mathrm{d}x}e^{\mathrm{i}x}=\lim_{h\to 0}\frac{e^{\mathrm{i}(x+h)}-e^{\mathrm{i}x}}{h}=\mathrm{i}e^{\mathrm{i}x}\lim_{h\to 0}\frac{e^{\mathrm{i}h}-1}{\mathrm{i}h}=\mathrm{i}e^{\mathrm{i}x},\\(e^{-\mathrm{i}x})'&=\frac{\mathrm{d}}{\mathrm{d}x}e^{-\mathrm{i}x}=\lim_{h\to 0}\frac{e^{-\mathrm{i}(x+h)}-e^{-\mathrm{i}x}}{h}=-\mathrm{i}e^{-\mathrm{i}x}\lim_{h\to 0}\frac{e^{-\mathrm{i}h}-1}{-\mathrm{i}h}=-\mathrm{i}e^{-\mathrm{i}x}.\end{aligned}$$
对于三角函数, 由$\sin x=\frac{e^{\mathrm{i}x}-e^{-\mathrm{i}x}}{2\mathrm{i}}$, $\cos x=\frac{e^{\mathrm{i}x}+e^{-\mathrm{i}x}}{2}$ 有
$$\begin{aligned}(\sin x)'&=\frac{(e^{\mathrm{i}x})'-(e^{-\mathrm{i}x})'}{2\mathrm{i}}=\frac{\mathrm{i}e^{\mathrm{i}x}-(-\mathrm{i})e^{-\mathrm{i}x}}{2\mathrm{i}}=\frac{e^{\mathrm{i}x}+e^{-\mathrm{i}x}}{2}=\cos x,\\(\cos x)'&=\frac{(e^{\mathrm{i}x})'+(e^{-\mathrm{i}x})'}{2}=\frac{\mathrm{i}e^{\mathrm{i}x}-\mathrm{i}e^{-\mathrm{i}x}}{2}=-\frac{e^{\mathrm{i}x}-e^{-\mathrm{i}x}}{2\mathrm{i}}=-\sin x.\end{aligned}$$
□

有了以上基本结果再结合导数的四则运算我们就可以计算$\tan x=\sin x/\cos x$, $x^\alpha\log x$等初等函数的导数. 但是要计算形如$\log(\sin x+3e^{\cos(x^2)})$的初等函数的导数, 我们还需要更重要的工具——

2. 复合函数的微分.

先看一个典型例子. 设A,B 为常数,
$$y=y(x)=Ax,\quad z=z(y)=By.$$
则$y'(x)=A$, $z'(y)=B$, $z(y(x))=BAx$, 从而有
$$(z\circ y)'(x)=BA=z'(y)\Big|_{y=y(x)}y'(x)=z'(y(x))y'(x).$$
如果把$y=y(x)$, $z=z(y)$ 看成放大镜(即$A>1,B>1$), 那么这就是说$y(x)$先把$x\ (\geq 0)$放大了A 倍, 然后$z(y)$ 又把$y=y(x)$ 放大了B倍, 二者合成的结果是$z(y(x))$把x 放大了BA 倍.

由于微分是函数的局部线性近似并且加减任何常数(即平移)对微分无影响, 故上述例子本质上具有一般性.

【定理5.5. 复合函数的微分 — 链锁规则】设 $X, Y \subset \mathbb{R}$, $g : X \to Y$ 在点 $x \in X$ 可微, 而函数 $f : Y \to \mathbb{C}$ 在点 $y = g(x)$ 可微. 则复合函数 $f \circ g : X \to \mathbb{C}$ 在 x 可微并且

$$(f \circ g)'(x) = f'(y)\Big|_{y=g(x)} g'(x) = f'(g(x))g'(x),$$

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

【证】由可微和高阶无穷小的定义有

$$g(x+h) - g(x) = g'(x)h + \alpha(h)h, \quad \alpha(h) \to 0 \ \ (h \to 0), \quad \alpha(0) = 0,$$

$$f(y+k) - f(y) = f'(y)k + \beta(k)k, \quad \beta(k) \to 0 \ \ (k \to 0), \quad \beta(0) = 0.$$

将$y = g(x)$ 代入g的微分式并取$k = k_h = g(x+h) - g(x)$, 则有

$$\begin{aligned}
& k_h = g'(x)h + \alpha(h)h, \qquad g(x+h) = g(x) + k_h, \\
& f(g(x+h)) - f(g(x)) = f(g(x) + k_h) - f(g(x)) = f'(g(x))k_h + \beta(k_h)k_h \\
& = f'(g(x))\Big(g'(x)h + \alpha(h)h\Big) + \beta(k_h)\Big(g'(x)h + \alpha(h)h\Big) \\
& = f'(g(x))g'(x)h + f'(g(x))\alpha(h)h + \beta(k_h)(g'(x) + \alpha(h))h \\
& = f'(g(x))g'(x)h + \gamma(h)h
\end{aligned}$$

其中

$$\gamma(h) = f'(g(x))\alpha(h) + \beta(k_h)(g'(x) + \alpha(h)).$$

因可微蕴含连续: $k_h = g(x+h) - g(x) \to 0$ $(h \to 0)$, 故$\beta(k_h) \to 0$ $(h \to 0)$. 因此$\gamma(h) \to 0$ $(h \to 0)$. 于是有$\gamma(h)h = o(h)$ $(h \to 0)$ 从而我们得到微分关系:

$$f(g(x+h)) - f(g(x)) = f'(g(x))g'(x)h + o(h).$$

由可微的定义知$f \circ g$ 在点x 可微且

$$(f \circ g)'(x) = f'(g(x))g'(x) \qquad \text{即} \qquad \mathrm{d}(f \circ g)(x)(h) = f'(g(x))g'(x)h.$$

最后再由$g'(x)h = g'(x)\mathrm{d}x(h)$ 即得

$$\mathrm{d}(f \circ g)(x)(h) = f'(g(x))g'(x)\mathrm{d}x(h) \qquad \forall\, h \in \mathbb{R}.$$

所以$\mathrm{d}(f \circ g)(x) = f'(g(x))g'(x)\mathrm{d}x$. $\qquad\square$

【定理5.5的推论1:一阶微分形式的不变性)】设f 在区间I内处处可微. 则无论x 是否为自变量都有$\mathrm{d}f(x) = f'(x)\mathrm{d}x$.

【证】当x为自变量时$\mathrm{d}f(x) = f'(x)\mathrm{d}x$ 自然成立. 当$x = \varphi(t)$ 时其中$\varphi(t) \in I$在$t \in J$内处处可微, 则由链锁规则和$\mathrm{d}x = \mathrm{d}\varphi(t) = \varphi'(t)\mathrm{d}t$有

$$\mathrm{d}f(x) = \mathrm{d}f(\varphi(t)) = f'(\varphi(t))\varphi'(t)\mathrm{d}t = f'(\varphi(t))\mathrm{d}\varphi(t) = f'(x)\mathrm{d}x. \qquad \square$$

【定理5.5的推论2: 多个映射复合的链锁规则】设有可微函数$y_1 = f_1(x), y_2 = f_2(y_1), ..., y_n = f_{n-1}(y_{n-1})$ 的复合$(f_n \circ f_{n-1} \circ \cdots \circ f_1)(x)$, $n \ge 2$. 则有

$$(f_n \circ f_{n-1} \circ \cdots \circ f_1)'(x) = f_n'(y_{n-1})f_{n-1}'(y_{n-2}) \cdots f_1'(x).$$

【证】当$n = 2$ 时, 这是定理5.5的结论. 设在$n = k$ 时成立, 则当$n = k + 1$时, 利用复合运算的结合律和两个函数$x \mapsto y_k = (f_k \circ f_{k-1} \circ \cdots \circ f_1)(x), y_k \mapsto f_{k+1}(y_k)$ 复合的结果以及归纳假设有

$$\begin{aligned}&(f_{k+1} \circ f_k \circ f_{k-1} \circ \cdots \circ f_1)'(x) = \Big(f_{k+1} \circ (f_k \circ f_{k-1} \circ \cdots \circ f_1)\Big)'(x)\\ &= f_{k+1}'(y_k)(f_k \circ f_{k-1} \circ \cdots \circ f_1)'(x) = f_{k+1}'(y_k)f_k'(y_{k-1})f_{k-1}'(y_{k-2}) \cdots f_1'(x).\end{aligned}$$

据归纳法原理, 链锁规则对任意有限多个可微映射的复合都成立. $\square$

【例】指数函数的导数. 设$a > 0$. 则

$$(a^x)' = a^x \log a, \quad x \in \mathbb{R}.$$

事实上将a 写成$a = e^{\log a}$ 有$a^x = e^{x \log a}$. 根据复合函数求导法则便知a^x 可导且

$$(a^x)' = (e^{x \log a})' = (e^y)_y'\Big|_{y = x\log a}(x \log a)_x' = e^y\Big|_{y = x\log a}\log a = a^x \log a, \quad x \in \mathbb{R}.$$

$\square$

【例】求初等函数$\log(\sin x + 3e^{\cos(x^2)})$的导函数.

【解】由$\sin x + 3e^{\cos(x^2)} \ge -1 + 3e^{-1} > 0$ for all $x \in \mathbb{R}$ 可知函数$x \mapsto \log(\sin x + 3e^{\cos(x^2)})$的定义域为$\mathbb{R}$. 根据复合函数求导法则, 我们计算

$$\begin{aligned}&(\log(\sin x + 3e^{\cos(x^2)}))_x' = (\log y)_y'\Big|_{y = \sin x + 3e^{\cos(x^2)}}(\sin x + 3e^{\cos(x^2)})_x'\\ &= \frac{1}{y}\Big|_{y = \sin x + 3e^{\cos(x^2)}}(\cos x + 3e^{\cos(x^2)}(\cos(x^2))_x')\\ &= \frac{1}{\sin x + 3e^{\cos(x^2)}}(\cos x + 3e^{\cos(x^2)}(-\sin(x^2)2x) = \frac{\cos x - 6xe^{\cos(x^2)}\sin(x^2)}{\sin x + 3e^{\cos(x^2)}}\end{aligned}$$

$x \in \mathbb{R}$. $\square$

4. 反函数的微分法

应用复合函数的微分法我们可以得到计算反函数的微分的公式.

【定理5.6.反函数的微分法】设 $X \subset \mathbb{R}$, $f : X \to \mathbb{R}$ 为单射. 假设函数 f 在点$x_0 \in X$ 可微且 $f'(x_0) \neq 0$, 又设反函数$f^{-1} : f(X) \to X$ 在对应点 $y_0 = f(x_0)$ 连续. 则f^{-1} 在 y_0 可微并有

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

【证】由可微的定义知$x_0 \in X$ 是X的聚点. 又由可微蕴含连续以及$f : X \to \mathbb{R}$ 为单射可知$y_0 = f(x_0)$ 是$f(X)$的聚点. 我们有

$$y \neq y_0 \Longleftrightarrow f^{-1}(y) \neq f^{-1}(y_0); \quad y \to y_0 \Longleftrightarrow f^{-1}(y) \to f^{-1}(y_0).$$

结合$f'(x_0) \neq 0$ 即得

$$\begin{aligned} &\lim_{Y \ni y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{Y \ni y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{f(f^{-1}(y)) - f(f^{-1}(y_0))} \\ &= \lim_{Y \ni y \to y_0} \frac{1}{\frac{f(f^{-1}(y)) - f(f^{-1}(y_0))}{f^{-1}(y) - f^{-1}(y_0)}} = \lim_{X \ni x = f^{-1}(y) \to x_0} \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}} = \frac{1}{f'(x_0)}. \end{aligned}$$

因此导数$(f^{-1})'(y_0)$ 存在且$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}$. $\square$

【注】一般来说, “$f'(x_0) \neq 0$”不蕴含“f^{-1} 在y_0 连续”. 例如设$f : X = (-\infty, -1) \cup [0, +\infty) \to \mathbb{R}$ 定义为

$$f(x) = x, \quad x \in [0, +\infty); \quad f(x) = x + 1, \quad x \in (-\infty, -1).$$

则$f : X \to \mathbb{R}$ 是单满射; 它的反函数$f^{-1} : \mathbb{R} \to X$的表达式为

$$f^{-1}(y) = y, \quad y \in [0, +\infty); \quad f^{-1}(y) = y - 1, \quad y \in (-\infty, 0).$$

根据导数的定义我们有

$$\lim_{X \ni x \to 0} \frac{f(x) - f(0)}{x} = \lim_{x \to 0+} \frac{x}{x} = 1.$$

因此$f'(0) = 1 \neq 0$. 但反函数$f^{-1}(y)$ 显然在对应点$f(0) = 0$处不连续.

然而若X是区间, 则上述定理可以改进为

【定理 5.7. 区间上严格单调函数的反函数的微分】 设 $I\subset\mathbb{R}$ 是区间, $f: I\to\mathbb{R}$ 严格单调. 如果f 在 $x\in I$ 可微且 $f'(x)\neq 0$, 则反函数$f^{-1}: f(I)\to I$ 在 $y=f(x)$ 可微且

$$(f^{-1})'(y)=\frac{1}{f'(x)}\Big|_{x=f^{-1}(y)}=\frac{1}{f'(f^{-1}(y))}.$$

【证明】因反函数$f^{-1}: f(I)\to I$ 也是严格单调的且其值域$f^{-1}(f(I))=I$ 是区间, 故由第四章单调函数的性质(见本讲义定理4.20)可知 f^{-1} 在 $f(I)$ 上连续, 特别 f^{-1} 在 $y=f(x)$ 连续. 因此据上一定理即知本定理成立. $\Box$

【函数单调性与导数的关系】 设$I\subset\mathbb{R}$ 为区间, 函数$f: I\to\mathbb{R}$ 在I 上可微. 则

(1) f 在I 上单调不减 $\iff$ $f'(x)\ge 0\quad\forall x\in I$.

(2) 若$f'(x)>0\quad\forall x\in I$, 则$f$在$I$ 上严格单调增加.

【证】(1): "$\Longrightarrow$": 设f 在I 上单调不减. 对任意$x\in I$, 若x 不是I的右端点, 则当$y\in I$ 且$y>x$时有$f(y)-f(x)\ge 0$. 因此

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

若x 不是I的左端点, 则当$y\in I$ 且$y<x$时有$f(y)-f(x)\le 0$. 因此仍有

$$f'(x)=\lim_{y\to x-}\frac{f(y)-f(x)}{y-x}\ge 0.$$

"$\Longleftarrow$": 设$f'(x)\ge 0\quad\forall x\in I$. 为证明$f$ 在I 上单调不减, 只需证明下列性质成立:

对任意 $a,b\in I$ 满足 $a<b$, 都存在 $c\in[a,b]$ 使得 $f(b)-f(a)\ge f'(c)(b-a)$. $\quad(*)$

设$a,b\in I$ 且$a<b$. 由

$$f(b)-f(\frac{a+b}{2})+f(\frac{a+b}{2})-f(a)=f(b)-f(a)$$

知$[a,b]$的二等分区间$[a_1,b_1]=[a,\frac{a+b}{2}]$ 或$[a_1,b_1]=[\frac{a+b}{2},b]$ 满足

$$f(b_1)-f(a_1)\le\frac{f(b)-f(a)}{2}.$$

同理存在$[a_1,b_1]$的一个二等分区间$[a_2,b_2]$ 满足

$$f(b_2)-f(a_2)\le\frac{f(b)-f(a)}{4}.$$

运用归纳法我们得到闭区间套$[a,b]\supset[a_1,b_1]\supset[a_2,b_2]\supset\cdots\supset[a_n,b_n]\supset\cdots$ 满足

$$f(b_n)-f(a_n)\le\frac{f(b)-f(a)}{2^n},\quad b_n-a_n=\frac{b-a}{2^n}.\quad n=1,2,3,\dots.$$

这给出差商不等式

$$\frac{f(b_n)-f(a_n)}{b_n-a_n}=\frac{f(b_n)-f(a_n)}{b-a}\cdot 2^n\le\frac{f(b)-f(a)}{b-a},\quad n=1,2,3,\dots.$$

设$c\in[a_n,b_n]$ 是所有$[a_n,b_n]$的公共点(闭区间套原理). 则由前面例题的结果知

$$\lim_{n\to\infty}\frac{f(b_n)-f(a_n)}{b_n-a_n}=f'(c).$$

于是有$f'(c)\le\frac{f(b)-f(a)}{b-a}$, 即$f(b)-f(a)\ge f'(c)(b-a)$.

(2): 设$f'(x)>0\ \ \forall x\in I$. 则由上面的(*) 知对任意$x,y\in I, x<y$, 存在$\xi\in(x,y)$ 使得$f(y)-f(x)\ge f'(\xi)(y-x)>0$. 所以$f$在$I$ 上严格递增. □

• 一些初等函数的反函数及其导数

(1). 正弦余弦函数.

我们承认一个事实: $\sin x>0$ for all $x\in(0,\pi)$. 这给出$\cos'(x)=-\sin x<0$ for all $x\in(0,\pi)$. 因此$\cos x$ 在$[0,\pi]$ 上严格单调减少. 由此可知$\sin(x)=\cos(\pi/2-x)$ 在$[-\pi/2,\pi/2]$ 上严格单调增加且$\sin'(x)=\cos x>0$ for all $x\in(-\pi/2,\pi/2)$ (这里用到$\cos x$ 是偶函数和$\cos x>\cos(\pi/2)=0$ for all $x\in[0,\pi/2)$). 令$\arcsin(y)=\sin^{-1}(y)$是$\sin(x)$ 在$[-\pi/2,\pi/2]$ 上的反函数, $\arccos(y)=\cos^{-1}(y)$是$\cos(x)$ 在$[0,\pi]$ 上的反函数, 则有

$$\arcsin'(y)=\frac{1}{\sin' x}=\frac{1}{\cos x}=\frac{1}{\sqrt{1-\sin^2 x}}=\frac{1}{\sqrt{1-y^2}},\quad y\in(-1,1).$$

$$\arccos'(y)=\frac{1}{\cos' x}=\frac{-1}{\sin x}=\frac{-1}{\sqrt{1-\cos^2 x}}=\frac{-1}{\sqrt{1-y^2}},\quad y\in(-1,1).$$

(2). 正切余切函数.

正切$\tan x=\frac{\sin x}{\cos x}, x\in(-\pi/2,\pi/2)$. $\tan'(x)=\frac{1}{\cos^2 x}>0$ for all $x\in(-\pi/2,\pi/2)$. 所以$\tan x$ 在$(-\pi/2,\pi/2)$ 上严格递增. 令$\arctan y=\tan^{-1}(y)$ 为$\tan x$在$(-\pi/2,\pi/2)$ 上的

反函数, 则

$$\arctan'(y) = \frac{1}{\tan' x}\Big|_{x=\arctan y} = \cos x\Big|_{x=\arctan y} = \frac{1}{1+\tan^2 x}\Big|_{x=\arctan y}.$$

因此

$$\arctan'(y) = \frac{1}{1+y^2} \qquad \forall\, y \in \mathbb{R}.$$

余切$\cot x = \tan(\frac{\pi}{2} - x) = \frac{\cos x}{\sin x}, x \in (0,\pi)$, $\cot'(x) = \frac{-1}{\sin^2 x} < 0$ for all $x \in (0,\pi)$. $\Longrightarrow$ $\cot x$ 在$(0,\pi)$ 上严格递减, 其反函数的导数为

$$\operatorname{arccot}'(y) = \frac{-1}{1+y^2}, \quad y \in \mathbb{R}.$$

(3). 双曲函数, 定义域为$\mathbb{R}$:

$$\text{双曲正弦：}\quad \operatorname{sh}(x) = \frac{e^x - e^{-x}}{2}; \qquad \text{双曲余弦：}\quad \operatorname{ch}(x) = \frac{e^x + e^{-x}}{2}.$$

“双曲”的由来:

$$\operatorname{ch}^2(x) - \operatorname{sh}^2(x) \equiv 1, \quad x \in \mathbb{R}.$$

双曲正弦的反函数: 首先易见双曲正弦$\operatorname{sh}(x)$ 的值域也是$\mathbb{R}$, 这是因为$\operatorname{sh}(+\infty) = +\infty, \operatorname{sh}(-\infty) = -\infty$. 其次由

$$\operatorname{sh}'(x) = \operatorname{ch}(x) > 0, \quad x \in \mathbb{R}$$

知$\operatorname{sh}(x)$ 严格单调增加. 为求其反函数的表达式, 令$y = \frac{e^x - e^{-x}}{2}$, 则$(e^x)^2 - 2ye^x - 1 = 0$. 因$e^x > 0$, 故得到$e^x = y + \sqrt{1+y^2}$. 令$\operatorname{arsh}(y) = \operatorname{sh}^{-1}(y)$ 为$\operatorname{sh}(x)$的反函数, 则有

$$\begin{aligned} &\operatorname{arsh}(y) = \log(y + \sqrt{1+y^2}) \qquad \text{也称作面积正弦}, \\ &\operatorname{arsh}'(y) = \frac{1}{\sqrt{1+y^2}}, \qquad y \in \mathbb{R}. \end{aligned}$$

双曲余弦$\operatorname{ch}(x) = \frac{e^x + e^{-x}}{2} : \mathbb{R} \to [1,+\infty)$是偶函数, 因此我们只能考虑它在半轴例如$\mathbb{R}_+ = [0,+\infty)$ 上的反函数. 由$\operatorname{ch}'(x) = \operatorname{sh}(x) = \frac{e^x - e^{-x}}{2} > 0$ for all $x > 0$ 和$\operatorname{ch}(0) = 1, \operatorname{ch}(+\infty) = +\infty$ 知$\operatorname{ch}(x)$ 在$\mathbb{R}_+$ 上严格增加, 其反函数的定义域为$[1,+\infty)$. 为求反函数的表达式, 令$y = \frac{e^x + e^{-x}}{2}$, 则$(e^x)^2 - 2ye^x + 1 = 0$. 因$e^x \ge 1$且$y \ge 1$, 故解得$e^x = y + \sqrt{y^2 - 1}$. 令$\operatorname{arch}_+(y) = \operatorname{ch}_+^{-1}(y)$ 为$\operatorname{ch}(x)$在$\mathbb{R}_+ = [0,+\infty)$ 上的反函数, 则有

$$\begin{aligned} &\operatorname{arch}_+(y) = \log(y + \sqrt{y^2-1}), \qquad y \ge 1, \\ &\operatorname{arch}_+'(y) = \frac{1}{\sqrt{y^2-1}}, \qquad y > 1. \end{aligned}$$

类似地, 令$\operatorname{arch}_-(y)=\operatorname{ch}_-^{-1}(y)$ 为$\operatorname{ch}(x)$在$\mathbb{R}_-=(-\infty,0]$ 上的反函数, 则有

$$\begin{aligned}
\operatorname{arch}_-(y) &= \log(y-\sqrt{y^2-1}), \qquad y\geq 1,\\
\operatorname{arch}_-'(y) &= \frac{-1}{\sqrt{y^2-1}}, \qquad y>1.
\end{aligned}$$

- **一些常用的初等函数及其反函数的导数表**

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}x^m = mx^{m-1}, \qquad x\in\mathbb{R},\ m\in\mathbb{N},\\
&\frac{\mathrm{d}}{\mathrm{d}x}x^m = mx^{m-1}, \qquad x\in\mathbb{R}\setminus\{0\},\ m\in\mathbb{Z}_{<0},\\
&\frac{\mathrm{d}}{\mathrm{d}x}x^\alpha = \alpha x^{\alpha-1}, \qquad x>0,\ \alpha\in\mathbb{R}\\
&\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x, \qquad x\in\mathbb{R},\\
&\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x\log a, \qquad x\in\mathbb{R},\ a>0,\\
&\frac{\mathrm{d}}{\mathrm{d}x}\log(x) = \frac{1}{x}, \qquad x>0,\\
&\frac{\mathrm{d}}{\mathrm{d}x}\log_a(x) = \frac{1}{x\log a}, \qquad x>0,\ a>0,\ a\neq 1,\\
&\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x), \qquad x\in\mathbb{R},\\
&\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x), \qquad x\in\mathbb{R},\\
&\tan(x)=\frac{\sin(x)}{\cos(x)},\quad \frac{\mathrm{d}}{\mathrm{d}x}\tan(x) = \frac{1}{\cos^2 x}, \qquad x\neq k\pi+\frac{\pi}{2},\, k\in\mathbb{Z},\\
&\cot(x)=\frac{\cos(x)}{\sin(x)},\quad \frac{\mathrm{d}}{\mathrm{d}x}\cot(x) = -\frac{1}{\sin^2 x}, \qquad x\neq k\pi,\, k\in\mathbb{Z},\\
&\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}, \qquad x\in(-1,1),\\
&\frac{\mathrm{d}}{\mathrm{d}x}\arccos(x) = -\frac{1}{\sqrt{1-x^2}}, \qquad x\in(-1,1),\\
&\frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{1+x^2}, \qquad x\in\mathbb{R},\\
&\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arccot}(x) = -\frac{1}{1+x^2}, \qquad x\in\mathbb{R},\\
&\operatorname{sh}(x)=\frac{e^x-e^{-x}}{2},\quad \operatorname{ch}(x)=\frac{e^x+e^{-x}}{2},\\
&\frac{\mathrm{d}}{\mathrm{d}x}\operatorname{sh}(x)=\operatorname{ch}(x),\quad \frac{\mathrm{d}}{\mathrm{d}x}\operatorname{ch}(x)=\operatorname{sh}(x), \qquad x\in\mathbb{R},\\
&\operatorname{arsh}(x)=\log(x+\sqrt{1+x^2}),\quad \frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arsh}(x)=\frac{1}{\sqrt{1+x^2}}, \qquad x\in\mathbb{R},\\
&\operatorname{arch}_\pm(y)=\log(x\pm\sqrt{x^2-1}),\ x\geq 1,\quad \frac{\mathrm{d}}{\mathrm{d}x}\operatorname{arch}_\pm(x)=\pm\frac{1}{\sqrt{x^2-1}}, \quad x>1.
\end{aligned}$$

5. 高阶导数.

【高阶导数的归纳定义】设$E \subset \mathbb{R}$, $f: E \to \mathbb{C}$. 定义$f^{(0)}(x) \equiv f(x)$, 并称$f^{(0)} = f$为f零阶导函数. 设n 为自然数. 若f 在E 上处处有$n-1$阶导数$f^{(n-1)}(x) \equiv \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} f(x)$, 则$f$ 在$x \in E$的第n 阶导数记作

$$f^{(n)}(x) = (f^{(n-1)})'(x) \qquad \text{即} \qquad \frac{\mathrm{d}^n f(x)}{\mathrm{d}x^{n-1}} = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\mathrm{d}^{n-1} f(x)}{\mathrm{d}x^{n-1}}\Big).$$

若第n次导数$f^{(n)}(x)$ 处处存在, $x \in E$, 则称f在E 上n次可导, 并称$f^{(n)}$ 是f 在E 上的第n 阶导函数.

若对每个$n \in \mathbb{N}$, f 在E 上都有n 阶导数, 则称f在E 上无穷次可导.

其他记号: 对于$k \in \mathbb{N}$, 我们用$C^k(E) = C^k(E, \mathbb{C})$ 表示在E 上具有k 次导数且导函数$f^{(k)}$ 在E 上连续的复值函数的全体, 即

$$C^k(E) = \{f : E \to \mathbb{C} \mid \ f\text{的第}k\text{ 阶的导数}f^{(k)}(x)\text{在}E\text{上处处存在且在}E\text{ 上连续}\}$$

用$C^\infty(E) = C^\infty(E, \mathbb{C})$ 表示在E 上具有任意次导数且所有导函数在E 上连续的复值函数的全体, 即

$$C^\infty(E) = \bigcap_{k=1}^{\infty} C^k(E).$$

把值域$\mathbb{C}$ 换成$\mathbb{R}$, 同样可定义实值的连续可微函数类$C^k(E, \mathbb{R})$ 和$C^\infty(E, \mathbb{R})$ □

【注】因为可微蕴含连续, 故易见对任意$k \in \mathbb{N}$ 有$C^k(E) = \bigcap_{j=0}^{k} C^j(E)$. 这里定义$C^0(E) = C(E)$.

基本初等函数的高阶导数: 对每个自然数n 有

$$(x^\alpha)' = \alpha x^{\alpha-1},\ (x^\alpha)'' = \alpha(\alpha-1)x^{\alpha-2},\ ......,\ (x^\alpha)^{(n)} = \alpha(\alpha-1)\cdots(\alpha-n+1)x^{\alpha-n}.$$

$$(a^x)' = a^x \log a, \quad (a^x)'' = a^x(\log a)^2, \quad ......, \quad (a^x)^{(n)} = a^x(\log a)^n.$$

$$(e^x)' = e^x, \quad (e^x)'' = e^x, \quad ......, \quad (e^x)^{(n)} = e^x.$$

$$(\log x)' = x^{-1},\ (\log x)'' = -x^{-2},\ ......,\ (\log x)^{(n)} = (-1)^n(n-1)!x^{-n}.$$

$$(\log|x|)' = \Big(\frac{1}{2}\log x^2\Big)' = x^{-1},\ (\log|x|)'' = -x^{-2},\ ......,\ (\log|x|)^{(n)} = (-1)^n(n-1)!x^{-n}.$$

$$(\sin x)' = \cos x, \quad (\sin x)'' = -\sin x, \quad ......, \quad (\sin x)^{(n)} = \sin(x + \frac{n\pi}{2}).$$

$$(\cos x)' = -\sin x, \quad (\cos x)'' = -\cos x, \quad ......, \quad (\cos x)^{(n)} = \cos(x + \frac{n\pi}{2}).$$

以上可用归纳法验证.

【定理5.8. 函数乘积的导数公式——莱布尼兹公式】设两个函数u, v 在E 上有直到n 阶的导数. 则它们的乘积uv的导数等于它们各自导数的乘积的“二项式”和:

$$(uv)^{(n)}(x) = \sum_{k=0}^{n} C_n^k u^{(n-k)}(x) v^{(k)}(x).$$

其中$u^{(0)} = u, v^{(0)} = v$, $C_n^k = \frac{n!}{k!(n-k)!}$ 是二项组合数.

【证】我们对导数的阶数n 用归纳法. 当$n = 1$ 时, $(uv)^{(1)}(x) = (uv)'(x) = u'(x)v(x) + u(x)v'(x)$. 当$n = 2$ 时

$$(uv)^{(2)}(x) = (uv)''(x) = (u'(x)v(x))' + (u(x)v'(x))'$$

$$= u''(x)v(x)+u'(x)v'(x)+u'(x)v'(x)+u(x)v''(x) = u''(x)v(x)+2u'(x)v'(x)+u(x)v''(x).$$

假设乘积求导的阶数等于$n(\geq 2)$ 时莱布尼兹公式成立. 看$n+1$时：计算

$$\begin{aligned}
&(uv)^{(n+1)}(x) = ((uv)^{(n)}(x))' = \Big(\sum_{k=0}^{n} C_n^k u^{(n-k)}(x) v^{(k)}(x)\Big)' \\
&= \sum_{k=0}^{n} C_n^k \Big(u^{(n-k)}(x) v^{(k)}(x)\Big)' = \sum_{k=0}^{n} C_n^k \Big(u^{(n+1-k)}(x) v^{(k)}(x) + u^{(n-k)}(x) v^{(k+1)}(x)\Big) \\
&= \sum_{k=0}^{n} C_n^k u^{(n+1-k)}(x) v^{(k)}(x) + \sum_{k=0}^{n} C_n^k u^{(n-k)}(x) v^{(k+1)}(x) \\
&= \sum_{k=0}^{n} C_n^k u^{(n+1-k)}(x) v^{(k)}(x) + \sum_{k=1}^{n+1} C_n^{k-1} u^{(n+1-k)}(x) v^{(k)}(x) \qquad (\text{上标}k\text{已被平移}) \\
&= u^{(n+1)}(x) v^{(0)}(x) + \sum_{k=1}^{n} (C_n^k + C_n^{k-1}) u^{(n+1-k)}(x) v^{(k)}(x) + u^{(0)}(x) v^{(n+1)}(x) \\
&= \sum_{k=0}^{n+1} C_{n+1}^k u^{(n+1-k)}(x) v^{(k)}(x)
\end{aligned}$$

其中用到组合恒等式

$$C_n^k + C_n^{k-1} = C_{n+1}^k, \qquad C_{n+1}^0 = C_{n+1}^{n+1} = 1.$$

这证明了当乘积求导的阶数等于$n+1$ 时莱布尼兹公式仍成立. 据归纳法原理, 莱布尼兹公式普遍成立. $\square$

导数的另一常用记号为：

$$D=\frac{\mathrm{d}}{\mathrm{d}x},\quad D^k=D(D^{k-1})=\frac{\mathrm{d}^k}{\mathrm{d}x^k}.$$

用此记号, 莱布尼兹公式写为

$$D^n(uv)=\sum_{k=0}^{n}C_n^kD^{n-k}uD^kv.$$

由此看出, 莱布尼兹公式的右端几乎就是二项式展开. 但莱布尼兹公式的左端是乘积uv的n阶导数而不是和$u+v$的n阶导数！原因是: 乘积uv的一阶导数是两项之和:

$$\frac{u(x+h)v(x+h)-u(x)v(x)}{h}=\frac{u(x+h)-u(x)}{h}v(x+h)+u(x)\frac{v(x+h)-v(x)}{h}$$
$$\Longrightarrow\qquad (uv)'(x)=u'(x)v(x)+u(x)v'(x),$$

【例题】设$n\ge 2$. 求$x^2\sin(x)$的n阶导函数.

【解】x^2 和$\sin(x)$ 都是任意次可导的. 由莱布尼兹公式和$\sin^{(m)}(x)=\sin(x+\frac{m\pi}{2})$ 有

$$\begin{aligned}&(x^2\sin(x))^{(n)}=\sum_{k=0}^{n}C_n^k(x^2)^{(k)}\sin^{(n-k)}(x)=\sum_{k=0}^{2}C_n^k(x^2)^{(k)}\sin^{(n-k)}(x)\\&=x^2\sin^{(n)}(x)+n2x\sin^{(n-1)}(x)+\frac{n(n-1)}{2!}2\sin^{(n-2)}(x)\\&=x^2\sin(x+\frac{n\pi}{2})+2nx\sin(x+\frac{(n-1)\pi}{2})+n(n-1)\sin(x+\frac{(n-2)\pi}{2})\\&=\Big(x^2-n(n-1)\Big)\sin\Big(x+\frac{n\pi}{2}\Big)-2nx\cos\Big(x+\frac{n\pi}{2}\Big).\qquad\square\end{aligned}$$

【例题(一个常用的无穷次可微的函数)】令

$$f(x)=\begin{cases}e^{-1/x} & \text{当 } x>0,\\ 0 & \text{当 } x\le 0.\end{cases}$$

则f 在$\mathbb{R}$ 上无穷次可导即$f\in C^\infty(\mathbb{R})$, 且对任意$n\in\mathbb{N}$有

$$f^{(n)}(x)=\begin{cases}e^{-1/x}\dfrac{P_{n-1}(x)}{x^{2n}} & \text{当 } x>0,\\ 0 & \text{当 } x\le 0.\end{cases}\qquad (*)$$

其中当$n=1$ 时$P_0(x)\equiv 1$, 当$n\ge 2$ 时, $P_{n-1}(x)$是次数$\le n-1$的多项式.

【证】由f的定义知f在$(0,+\infty)$ 和$(-\infty,0)$ 内分别是初等函数因而在$(0,+\infty)$ 和$(-\infty,0)$ 内分别无穷次可导. 证明过程集中在证明f在交接点$x=0$ 处也是无穷次可导的(它表明交接得无限光滑!). 我们将使用事实

$$\lim_{x\to 0+}\frac{e^{-1/x}}{x^\alpha}=0\qquad \forall\,\alpha\in\mathbb{R}.$$

为证$(*)$, 我们对导数的次数n 用归纳法.

设$n=1$. 计算: 当$x>0$ 时

$$f'(x)=e^{-1/x}(-1)(-1)x^{-2}=e^{-1/x}x^{-2},$$

当$x<0$ 时, $f'(x)=0$. 当$x=0$时

$$\lim_{x\to 0+}\frac{f(x)-f(0)}{x}=\lim_{x\to 0+}\frac{e^{-1/x}}{x}=0,\quad \lim_{x\to 0-}\frac{f(x)-f(0)}{x}=\lim_{x\to 0-}\frac{0}{x}=0.$$

这表明$f'(0)$ 存在且$f'(0)=0$. 这证明了当$n=1$ 时$(*)$ 成立.

假设在导数次数为$n\in\mathbb{N}$ 时$(*)$ 成立, 来看导数次数为$n+1$ 时. 由归纳假设易见当$x>0$时

$$\begin{aligned}
&f^{(n+1)}(x)\\
&=e^{-1/x}x^{-2}x^{-2n}P_{n-1}(x)+e^{-1/x}(-2n)x^{-2n-1}P_{n-1}(x)+e^{-1/x}x^{-2n}P'_{n-1}(x)\\
&=e^{-1/x}x^{-2(n+1)}\Big(P_{n-1}(x)-2nxP_{n-1}(x)+x^2P'_{n-1}(x)\Big)\\
&=e^{-1/x}\frac{P_n(x)}{x^{2(n+1)}},
\end{aligned}$$

当$x<0$ 时$f^{(n+1)}(x)=0$, 当$x=0$时

$$\begin{aligned}
\lim_{x\to 0+}\frac{f^{(n)}(x)-f^{(n)}(0)}{x}&=\lim_{x\to 0+}e^{-1/x}\frac{P_{n-1}(x)}{x^{2n+1}}=0,\\
\lim_{x\to 0-}\frac{f^{(n)}(x)-f^{(n)}(0)}{x}&=\lim_{x\to 0-}\frac{0}{x}=0
\end{aligned}$$

所以$f^{(n)}(x)$在$x=0$ 可导且$f^{(n+1)}(0)=0$. 这证明了在导数次数为$n+1$ 时$(*)$也成立. 据归纳法原理知$(*)$ 对任意$n\in\mathbb{N}$ 成立. $\square$

【作业题】设f为上面例题定义的非负函数, 即$f(x)=\begin{cases}e^{-1/x} & \text{当 } x>0,\\ 0 & \text{当 } x\le 0.\end{cases}$

(1) 设$P(x)$为一个多项式. 证明复合函数$x\mapsto f(P(x))$ 属于$C^\infty(\mathbb{R})$.

(2) (制作光滑的"特征函数") 对于$\varepsilon>0$, 令

$$g(x)=\frac{f((1+\varepsilon)^2-x^2)}{f((1+\varepsilon)^2-x^2)+f(x^2-1)},\quad x\in\mathbb{R}.$$

证明$g(x)$ 在$\mathbb{R}$上是良好定义的(即分母处处非零) 且属于$C^\infty(\mathbb{R})$, 且满足

$$0\le g(x)\le 1\quad \forall\, x\in\mathbb{R};\quad g(x)=1\quad \forall\,|x|\le 1;\quad g(x)=0\quad \forall\,|x|\ge 1+\varepsilon.$$

6. 参数曲线的导数, 速度, 加速度

以平面曲线为例. 设有平面参数曲线$t \mapsto \mathbf{r}(t) = (x(t), y(t)), t \in I$ 这里$I \subset \mathbb{R}$ 为区间. 可以把t看做时间, 把$\mathbf{r}(t)$看做一个质点的位移. 设$t \in I$. 若

$$\lim_{I \ni s \to t} |\mathbf{r}(s) - \mathbf{r}(t)| = 0$$

则称$\mathbf{r}(\cdot)$ 在t连续。若存在$\mathbf{v}(t) = (v_1(t), v_2(t)) \in \mathbb{R}^2$ 使得

$$\lim_{I \ni s \to t} \Big| \frac{\mathbf{r}(s) - \mathbf{r}(t)}{s - t} - \mathbf{v}(t) \Big| = 0$$

则称$\mathbf{r}(\cdot)$ 在t可导并记

$$\mathbf{r}'(t) = \mathbf{v}(t).$$

假设$\mathbf{r}(\cdot)$ 在I 上处处可导并且导函数即速度向量$\mathbf{v}(\cdot) = \mathbf{r}'(\cdot)$ 在时刻t 可导, 则自然有

$$\mathbf{r}''(t) = \mathbf{v}'(t).$$

称其为位移$\mathbf{r}(\cdot)$ 在t 时刻的加速度.

注意: 由

$$|\mathbf{r}(s) - \mathbf{r}(t)| = \sqrt{(x(s) - x(t))^2 + (y(s) - y(t))^2}$$

和

$$\Big| \frac{\mathbf{r}(s) - \mathbf{r}(t)}{s - t} - \mathbf{v}(t) \Big| = \sqrt{\Big(\frac{x(s) - x(t)}{s - t} - v_1(t) \Big)^2 + \Big(\frac{y(s) - y(t)}{s - t} - v_2(t) \Big)^2}$$

等等, 可知

$\mathbf{r}(\cdot)$ 在时刻t连续 $\iff$ $x(\cdot), y(\cdot)$ 都在时刻t连续.

$\mathbf{r}(\cdot)$ 在时刻t可导 $\iff$ $x(\cdot), y(\cdot)$ 都在时刻t可导且 $\mathbf{r}'(t) = (x'(t), y'(t))$.

$\mathbf{r}(\cdot)$ 在时刻t两次可导 $\iff$ $x(\cdot), y(\cdot)$ 都在时刻t两次可导且 $\mathbf{r}''(t) = (x''(t), y''(t))$.

等等.

与实值函数一样我们可以定义光滑的向量值函数类:

$$C^k(I, \mathbb{R}^2) = \{\mathbf{r} : I \to \mathbb{R}^2 \,|\, \mathbf{r}\text{的第}k\text{ 阶的导数}\mathbf{r}^{(k)}(t)\text{在}I\text{上处处存在且在}I\text{上连续}\}.$$

易见

$$C(I, \mathbb{R}^2) \supset C^1(I, \mathbb{R}^2) \supset C^2(I, \mathbb{R}^2) \supset \cdots.$$

【奇点与正则点】设$\mapsto \mathbf{r}(t)=(x(t),y(t))$ 属于$C^1(I,\mathbb{R}^2)$. 若$t_0\in I$使得$\mathbf{r}'(t_0)\neq 0$ 即$x'(t_0)\neq 0$ 或者$y'(t_0))\neq 0$, 则称t_0 或$\mathbf{r}(t_0)$为曲线$\mathbf{r}(t)$ 的一个正则点. 若$t_0\in I$使得$\mathbf{r}'(t_0)=0$ 即$x'(t_0)=0$ 且$y'(t_0))=0$, 则称t_0 或$\mathbf{r}(t_0)$ 为曲线$\mathbf{r}(t)$ 的一个奇点. 奇点的物理意义是: 它表示运动着的质点在时刻t_0有停顿. 对奇点的研究涉及深刻的代数、几何和非线性分析学, 是本科数学分析无法包揽的. 而实际问题中多数情况下奇点是很少见的. 因此本科数学分析中我们主要研究没有奇点的情形.

【参数导数与坐标导数的关系】设曲线$\mathbf{r}(t)=(x(t),y(t))$ 属于$C^1(I,\mathbb{R}^2)$ 且在I 上无奇点, 即$\mathbf{r}'(t)=(x'(t),y'(t))\neq 0\ \forall t\in I$. 因为导函数是连续的, 故例如若$x'(t)$ 在某时刻t_0不等于零, 则在t_0的一个邻域内$x'(t)$ 都不等于零. 因此将区间I 缩小后, 我们可以假定$x'(t)\neq 0$ for all $t\in I$. 于是$t\mapsto x(t)$ 在I上严格单调从而其反函数$x\mapsto t=t(x)$也是严格单调、可微、且导数不等于零. 根据反函数求导定理有

$$t'(x)=\frac{\mathrm{d}t}{\mathrm{d}x}=\frac{1}{x'(t)}\Big|_{t=t(x)}.$$

把$t=t(x)$ 代入$y(x(t))$ 可知

$$x\mapsto y=f(x):=y(t(x))$$

也是x的函数, 且根据复合函数求导法则便得到$y=f(x)$ 关于x的导数:

$$f'(x)=y'(t(x))t'(x)=\frac{y'(t)}{x'(t)}\Big|_{t=t(x)}.$$

若进一步假设$t\mapsto (x(t),y(t))$ 在区间I上处处两次可微, 则由$x'(t)\neq 0$ for all $t\in I$, $t\mapsto x'(t)$可导、$x\mapsto t(x)$ 可导以及复合函数求导法可知$x\mapsto t'(x)=\frac{1}{x'(t(x)}$ 对仍x 可导且

$$t''(x)=\Big(\frac{1}{x'(t(x))}\Big)'_x=-\frac{x''(t(x))t'(x)}{(x'(t(x)))^2}=-\frac{x''(t(x))}{(x'(t(x)))^3}=-\frac{x''(t)}{(x'(t))^3}\Big|_{t=t(x)}.$$

再次应用复合函数求导法则可知$x\mapsto f'(x)=y'(t(x))t'(x)$对$x$ 仍可导, 并将以上关系式代入, 便得到$y=f(x)$对x的二阶导数为

$$f''(x)=(y'(t(x))t'(x))'_x=y''(t(x))(t'(x))^2+y'(t(x))t''(x)=\frac{y''(t)}{(x'(t))^2}-y'(t)\frac{x''(t)}{(x'(t))^3}\Big|_{t=t(x)}$$

即

$$f''(x)=\frac{y''(t)x'(t)-y'(t)x''(t)}{(x'(t))^3}\Big|_{t=t(x)}.$$

参数曲线的一阶导数二阶导数等有直接的物理意义, 如速度、加速度等.

下面的作业是这方面的例题与练习.

【作业题】一个物体, 可以把它看作是一个质点, 在重力的作用下从一个光滑的小山上滑下来, 这里山坡是一个函数$y=f(x)$的图像：$\{(x,f(x))\,|\,x\in I\}$, I为一个区间, $f\in C^2(I)$.

(1) 求物体在山坡上的点(x_0,y_0) (即$y_0=f(x_0)$) 处的加速度的水平分量和竖直分量.

(2) 当 $y=x^2$ 时且物体从很高的地方滑下来时, 求抛物线 $y=x^2$ 上使加速度的水平分量最大的点. 这里假定物体在下滑之前处于静态, 即下滑的初速度为零.

【说明】本题稍复杂, 难点在于列出正确的计算方案. 题目的叙述暗含这样的假定: 这个质点在三维空间中滑动的轨迹位于一张平面内, 并假设这平面就是xy-平面. 设

$$[a,b]\subset I,\quad 0=f(a)\le f(x)\le f(b),\qquad x\in[a,b]$$

质点开始下滑的位置为$(b,f(b))$, 其中高度$H:=f(b)>0$. 注意: 质点在点$(b,f(b))$处可以有两种状态: 静止状态(即初速度为零)和非静止状态. 本题假定是前一种即静止状态, 因此可设开始下滑的时刻(即初始时刻)为$t=0$, 而在质点滑到山坡上的点(x_0,y_0)处的时刻为$t=t_0$.

在求解时可以考虑机械能守恒律(动能与势能之和等于常数)：

$$\frac{1}{2}m|\mathbf{v}(t)|^2+m\mathrm{g}\,y(t)\equiv C\ (\text{常数}),\qquad t\in[0,T]\,.$$

其中m为质点的质量, $\mathrm{g}=9.8...$为重力加速度, $\mathbf{v}(t)=(x'(t),y'(t))$为质点速度.

本题(2)的答案是: 质点从抛物型山坡$y=x^2$上很高的地方(b,b^2)滑下来时, 在点$(\frac{1}{\sqrt{12}},\frac{1}{12})$ 处其加速度的水平分量达到最大值.

【解答参考(作业布置一周后公布)】首先由题意知, 本题假定质点在三维空间中的滑动轨迹位于一张平面内, 并假设这平面就是xy-平面, 而山形函数f 在区间I 上是光滑的, $[a,b]\subset I$, $0=f(a)\le f(x)\le f(b)$ for all $x\in[a,b]$. 不妨设质点开始下滑的位置为$(b,f(b))$, 其中高度$H:=f(b)>0$.

(a) 设$t\mapsto(x(t),y(t)),t\in[0,T]$, 是质点沿着山坡$(x,f(x))$ 下滑时的运动轨迹, 也即

$$y(t)=f(x(t)),\quad x(0)=b,\quad y(0)=f(x(0))=f(b)=H.$$

在以下分析中我们略去量纲. 设m 为质点的质量, $\mathrm{g}=9.8...$为重力加速度. 设质点在时刻t的速度为$\mathbf{v}(t)$. 则$\mathbf{v}(t)=(x'(t),y'(t))$. 根据机械能守恒律(动能与势能之和等于常数)我们有

$$\frac{1}{2}m|\mathbf{v}(t)|^2+m\mathrm{g}y(t)\equiv C\ (\text{常数}),\qquad t\in[0,T]\,.$$

由复合函数求导法则有

$$y(t)=f(x(t))\quad\Longrightarrow\quad y'(t)=\frac{d}{dt}f(x(t))=f'(x(t))x'(t)\,,\quad t\in[0,T]$$

因此

$$\mathbf{v}(t)=x'(t)(1,f'(x(t))\,,\quad t\in[0,T]\,. \tag{$*$}$$

将这一关系代入守恒律中得到

$$\frac{m}{2}[x'(t)]^2(1+[f'(x(t))]^2)+m\mathrm{g}f(x(t))\equiv C,\quad t\in[0,T].$$

即

$$[x'(t)]^2(1+[f'(x(t))]^2)+2\mathrm{g}f(x(t))\equiv\frac{2C}{m}=:C_1,\quad t\in[0,T]. \tag{$**$}$$

两边对t 求导得到

$$\begin{aligned}&0\equiv 2x'(t)x''(t)(1+[f'(x(t))]^2)+[x'(t)]^2 2f'(x(t))f''(x(t))x'(t)+2\mathrm{g}f'(x(t))x'(t)\\&=2x'(t)\Big(x''(t)(1+[f'(x(t))]^2)+[x'(t)]^2f'(x(t))f''(x(t))+\mathrm{g}f'(x(t))\Big)\,.\end{aligned}$$

因当$0<t<T$ 时质点处于正在下滑中, 故其速度$\mathbf{v}(t)\neq\mathbf{0}$. 这等价于$x'(t)\neq 0$ (见($*$)). 于是得到

$$x''(t)(1+[f'(x(t))]^2)+[x'(t)]^2f'(x(t))f''(x(t))+\mathrm{g}f'(x(t))=0$$

即

$$x''(t)=-\frac{f'(x(t))}{1+[f'(x(t))]^2}\Big(\mathrm{g}+[x'(t)]^2f''(x(t))\Big)\,.$$

再次运用复合函数求导法得到$y''(t)$的表达式:

$$\begin{aligned}y''(t)&=\frac{d}{dt}(f'(x(t))x'(t)))=f''(x(t))[x'(t)]^2+f'(x(t))x''(t)\\&=f''(x(t))[x'(t)]^2-\frac{[f'(x(t))]^2}{1+[f'(x(t))]^2}\Big(\mathrm{g}+[x'(t)]^2f''(x(t))\Big)\\&=-\frac{\mathrm{g}[f'(x(t))]^2}{1+[f'(x(t))]^2}+\frac{f''(x(t))[x'(t)]^2}{1+[f'(x(t))]^2}.\end{aligned}$$

而由$(**)$ 知

$$[x'(t)]^2 = \frac{C_1 - 2\mathrm{g}f(x(t))}{1+[f'(x(t))]^2}$$

故将其代入$x''(t), y''(t)$的表达式得到

$$x''(t) = -\frac{\mathrm{g}f'(x(t))}{1+[f'(x(t))]^2} - \frac{f'(x(t))f''(x(t))(C_1 - 2\mathrm{g}f(x(t))}{(1+[f'(x(t))]^2)^2},$$

$$y''(t) - \frac{\mathrm{g}[f'(x(t))]^2}{1+[f'(x(t))]^2} + \frac{f''(x(t))(C_1 - 2\mathrm{g}f(x(t)))}{(1+[f'(x(t))]^2)^2}.$$

现在设$t_0 \in (0,T)$ 使得$(x(t_0), y(t_0))$ 为指定点, 即$(x(t_0), y(t_0)) = (x_0, y_0) = (x_0, f(x_0))$. 则物体在点 (x_0, y_0) 处的加速度的水平分量和竖直分量分别为

$$x''(t_0) = -\frac{\mathrm{g}f'(x_0)}{1+[f'(x_0)]^2} - \frac{f'(x_0)f''(x_0)(C_1 - 2\mathrm{g}y_0)}{(1+[f'(x_0)]^2)^2},$$

$$y''(t_0) = -\frac{\mathrm{g}[f'(x_0)]^2}{1+[f'(x_0)]^2} + \frac{f''(x_0)(C_1 - 2\mathrm{g}y_0))}{(1+[f'(x_0)]^2)^2}.$$

其中常数$C_1 = \frac{2C}{m}$ 可以通过在守恒律中取$t=0$确定为(注意$y(0) = f(b) = H$)

$$C_1 = \frac{2C}{m} = |\mathbf{v}(0)|^2 + 2\mathrm{g}H.$$

若质点从静止开始下滑, 即质点开始下滑时的初速度为零, 则$C_1 = 2\mathrm{g}H$ 从而得

$$x''(t_0) = -\frac{\mathrm{g}f'(x_0)}{1+[f'(x_0)]^2} - \frac{2\mathrm{g}f'(x_0)f''(x_0)(H - y_0)}{(1+[f'(x_0)]^2)^2},$$

$$y''(t_0) = -\frac{\mathrm{g}[f'(x_0)]^2}{1+[f'(x_0)]^2} + \frac{2\mathrm{g}f''(x_0)(H - y_0)}{(1+[f'(x_0)]^2)^2}.$$

(b). 此时有$f(x) = x^2$ 其中$0 \le x \le b$. 这对应于(a)中的$a=0, b>0$ 和$f(b) = b^2 = H$. 根据假设还知质点下滑的初速度为零: $\mathbf{v}(0) = \mathbf{0}$. 计算: $f'(x) = 2x$, $f''(x) = 2$. 代入上面所得公式有

$$x''(t) = \left(-\frac{\mathrm{g}2x}{1+4x^2} - \frac{2\mathrm{g}2x2(b^2 - x^2)}{(1+4x^2)^2}\right)_{x=x(t)} = -2\mathrm{g}(1+4b^2)\frac{x}{(1+4x^2)^2}\bigg|_{x=x(t)}.$$

$$\Longrightarrow \quad |x''(t)| = 2\mathrm{g}(1+4b^2)\frac{x}{(1+4x^2)^2}\bigg|_{x=x(t)}$$

为了确定$\max\limits_{0\le t\le T} |x''(t)|$, 只需求出$x_0 \in (0,b)$ 使得

$$\frac{x_0}{(1+4x_0^2)^2} = \max_{x\in[0,b]} \frac{x}{(1+4x^2)^2}\,.$$

令$\varphi(x)=\frac{x}{(1+4x^2)^2}$, $x\in[0,+\infty)$. 则$\varphi(0)=0,\varphi(+\infty)=0$, $\varphi(x)\ge 0$ for all $x\in[0,+\infty)$. 因此最大值$\max\limits_{x\in[0,+\infty)}\varphi(x)$存在且显然$>0$. 设$x_0\in(0,+\infty)$ s.t. $\varphi(x_0)=\max\limits_{x\in[0,+\infty)}\varphi(x)$. 因$\varphi(x_0)>0$, 故$x_0\in(0,+\infty)$, 也即$\varphi$ 的最大值点x_0 属于φ 的定义域的**内部**. 根据极值原理有$\varphi'(x_0)=0$. 计算:

$$\varphi'(x)=\frac{(1+4x^2)^2-x2(1+4x^2)8x}{(1+4x^2)^4}=\frac{1+4x^2-16x^2}{(1+4x^2)^3}=\frac{1-12x^2}{(1+4x^2)^3},\quad x>0.$$

$$\text{因此}\quad \varphi'(x_0)=0\quad\Longleftrightarrow\quad x_0=\frac{1}{\sqrt{12}}\,.$$

由此看出, 为使最大值能够在$(0,b)$ 内达到, 应使$b>\frac{1}{\sqrt{12}}$. 此时有

$$\max_{x\in[0,b]}\frac{x}{(1+4x^2)^2}=\max_{x\in[0,+\infty)}\frac{x}{(1+4x^2)^2}=\frac{x}{(1+4x^2)^2}\Big|_{x=x_0=\frac{1}{\sqrt{12}}}=\frac{3\sqrt{3}}{32}.$$

这就是说, 为了质点的加速度的水平分量达到最大值, 质点开始下滑的位置(b,b^2) 应满足$b>\frac{1}{\sqrt{12}}$. 例如取$b>\frac{1}{2}$ 则$x_0=\frac{1}{\sqrt{12}}\in(0,b)$.

结论: 质点从抛物型山坡$y=x^2$上很高的地方(b,b^2)滑下来时, 在点$(\frac{1}{\sqrt{12}},\frac{1}{12})$ 处其加速度的水平分量达到最大值。 [最后观察：上述最大值为

$$\max_{0\le t\le T}|x''(t)|=2\mathrm{g}(1+4b^2)\frac{3\sqrt{3}}{32}=\mathrm{g}(1+4b^2)\frac{3\sqrt{3}}{16}\,.$$

它表明质点水平方向加速度的最大值与质点最初的高度b^2 成正比. 它显然与事实相符.] $\square$

§5.3. 可微函数的整体性质

我们将以极值原理为中心展开分析. **大家将看到, 极值的观点是富有成效的!**

1. 局部极值和关于极值的费马定理.

【极值的定义】 设$E \subset \mathbb{R}$, $f : E \to \mathbb{R}$, $x_0 \in E$. 若存在x_0的一个(相对于E的)邻域$U_E(x_0) = E \cap (x_0 - \delta, x_0 + \delta)\,(\delta > 0)$ 使得

$$f(x) \le f(x_0) \quad (\ f(x) \ge f(x_0)\ ) \qquad \forall\, x \in U_E(x_0)$$

则称$f(x_0)$ 是f在E 上的一个局部极大(极小)值并称x_0 是f在E 上的一个局部极大(极小)值点. 而若

$$f(x) < f(x_0) \quad (\ f(x) > f(x_0)\ ) \qquad \forall\, x \in U_E(x_0) \setminus \{x_0\}$$

则称$f(x_0)$ 是f在E 上的一个局部严格极大(严格极小)值并称x_0 是f在E 上的一个局部严格极大(严格极小)值点. (图形)

f的局部极大(极小)值和局部极大(极小)值点统称为f的局部极值和局部极值点. 也常简称为f的极值和极值点. □

【注】整体最大(小)值点当然也是局部极大(小)值点, 反之不然. 例如设
$f(x) = x^2(1-x), x \in [-1, 3]$. 则$x = 0$ 是f 在$[-1, 3]$上的一个局部极小值点: $f(x) \ge f(0) = 0$ for all $x \in (-1, 1)$. 但$f(2) = -4 < 0 = f(0)$ 所以$x = 0$ 不是f 在$[-1, 3]$ 上的最小值.

【临界点的定义】 设$E \subset \mathbb{R}$, $f : E \to \mathbb{R}$, $x_0 \in E$. 若f 在x_0 可微且$f'(x_0) = 0$, 则称x_0 是f在E上的一个**临界点** (critical point). (两类图形) □

【定理5.9. Fermat 极值原理[3]】 设$E \subset \mathbb{R}$, $f : E \to \mathbb{R}$, 并设$x_0 \in E$ 是E 的双侧聚点(即x_0 既是E的左聚点也是E的右聚点. **例如当E 是区间时, x_0属于E的内部**). 设f在x_0 可微. 则x_0是f 在E 上的一个局部极值点的一个**必要条件**是: x_0 是f的一个临界点, 即$f'(x_0) = 0$. (图形)

[3]Pierre de Fermat, 1601-1665, 法国律师、业余数学家; 对概率论、数论(提出著名的Fermat 大定理) 和光学(提出最小作用原理) 都有很大贡献.

【证】不妨设x_0 是f 在E 上的一个局部极小值点(否则考虑$-f$). 则存在$\delta > 0$ 使得

$$f(x) \ge f(x_0) \quad \forall\, x \in E \cap (x_0 - \delta, x_0 + \delta).$$

又由x_0是E的双侧聚点知

$$E \cap (x_0 - \varepsilon, x_0) \ne \emptyset, \quad E \cap (x_0, x_0 + \varepsilon) \ne \emptyset \qquad \forall\, 0 < \varepsilon < \delta.$$

这蕴含对任意$0 < \varepsilon < \delta$ 有

$$\frac{f(x) - f(x_0)}{x - x_0} \le 0 \qquad \forall\, x \in E \cap (x_0 - \varepsilon, x_0),$$
$$\frac{f(x) - f(x_0)}{x - x_0} \ge 0 \qquad \forall\, x \in E \cap (x_0, x_0 + \varepsilon).$$

因f 在x_0 可微, 故

$$f'(x_0) = f'_-(x_0) = \lim_{E \ni x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0} \le 0,$$
$$f'(x_0) = f'_+(x_0) = \lim_{E \ni x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} \ge 0.$$

所以$f'(x_0) = 0$. $\square$

【例】(1) 设$f(x) = x^3, x \in \mathbb{R}$. 则$f'(x) = 3x^2 \ge 0$ 且$f'(x) = 0 \iff x = 0$. 因此f在$\mathbb{R}$上严格单调增加. 这说明$x = 0$ 是f的临界点但不是f的极值点. 因此Fermat 极值原理只是函数取得极值的一个必要条件而(一般来说)非充分条件.

(2) 设$f : [a, b] \to \mathbb{R}$ 严格单调增加且$f'(a) > 0, f'(b) > 0$. (例如$f(x) \equiv x$). 则$x = a, x = b$ 分别是f在$[a, b]$ 上的上的最小值点和最大值点从而也是f在$[a, b]$上的极小值点和极大值点. 虽然$f'(a) \ne 0, f'(b) \ne 0$，但这与Fermat 极值原理无矛盾因为Fermat 极值原理要求极值点x_0 是函数定义域中的双侧聚点, 而这里的a, b 都是$[a, b]$ 的单侧聚点. (图象解释)

【命题5.10.】设$I \subset \mathbb{R}$ 为区间, $f : I \to \mathbb{R}$ 在I 内处处可微. 假设最值$\min\limits_{x\in I} f(x)$或$\max\limits_{x\in I} f(x)$ 存在且最值点可以在I的**内部**达到, 那么这些位于I内部的最值点都是f 的临界点. 特别若f在I的内部只有一个临界点, 则这个临界点就是f的位于I内部的唯一最值点.

【证】以最小值为例. 设$\min\limits_{x\in I} f(x)$ 存在并设$x_0 \in I^\circ$ (即x_0 属于I的内部) 满足$f(x_0) = \min\limits_{x\in I} f(x)$. 则存在$\delta > 0$ 使得$(x_0 - \delta, x_0 + \delta) \subset I$ 且x_0 也是f 在I 的一个局部极小值点.

显然x_0 是I的双侧聚点. 因f 在I 内处处可微, 故Fermat 极值原理知x_0 是f的一个临界点, 即$f'(x_0)=0$. 进一步假设f在I° 内只有一个临界点, 即方程$f'(x)=0$ 在I° 内的解是唯一的, 则必有$\min\limits_{x\in I} f(x)=f(x_0)$ 其中$x_0\in I^\circ$ 是f在I°内的唯一临界点.

【**例**】设$-\infty\le a<b\le+\infty$, $f:(a,b)\to\mathbb{R}$ 可微且f在(a,b) 内只有有限多个临界点: $x_1,x_2,...,x_n$. 不失一般性, 考虑f的最小值问题. 设存在$c\in(a,b)$ 使得

$$\liminf_{x\to a+} f(x)>f(c),\quad \liminf_{x\to b-} f(x)>f(c).$$

则f 在(a,b) 上有最小值且最小值点属于$\{x_1,x_2,...,x_n\}$. 因此

$$\min_{x\in(a,b)} f(x)=\min\{f(x_1),f(x_2),...,f(x_n)\}.$$

特别当$n=1$ 时 $\min\limits_{x\in(a,b)} f(x)=f(x_1)$.

【证】由假设条件知存在A,B 满足$a<A<c<B<b$ 使得

$$f(x)>f(c)\quad \forall\, x\in(a,A)\cup(B,b).$$

因f在$[A,B]$ 上连续, 故存在$x_0\in[A,B]$ 使得$f(x_0)=\min\limits_{x\in[A,B]} f(x)$. 来证明

$$f(x_0)=\min_{x\in(a,b)} f(x)$$

即$f(x_0)$ 是f在(a,b) 上的整体最小值. 首先由$c\in[A,B]$ 知$f(c)\ge f(x_0)$. 对任意$x\in(a,b)$, 若$x\in(a,A)\cup(B,b)$, 则$f(x)>f(c)\ge f(x_0)$; 若$x\in[A,B]$ 则$f(x)\ge \min\limits_{x\in[A,B]} f(x)=f(x_0)$. 所以$f(x)\ge f(x_0)$ for all $x\in(a,b)$. 因此$f(x_0)$ 是f在(a,b) 上的整体最小值. 由**命题5.10** 知$x_0\in\{x_1,x_2,...,x_n\}$. 这蕴含

$$f(x_0)\ge\min\{f(x_1),f(x_2),...,f(x_n)\}.$$

而由$\{x_1,x_2,...,x_n\}\subset(a,b)$ 和$f(x_0)=\min\limits_{x\in(a,b)} f(x)$ 又有

$$f(x_0)\le\min\{f(x_1),f(x_2),...,f(x_n)\}.$$

所以$f(x_0)=\min\{f(x_1),f(x_2),...,f(x_n)\}$. $\square$

2. **微分中值定理**.

作为极值原理的重要应用, 我们来证明下面的Rolle 定理并由它证明其它中值定理.

【**定理5.11. Rolle 定理**】设$f:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续, 在(a,b) 内可导且$f(a)=f(b)$. 则存在$\xi\in(a,b)$ 使得$f'(\xi)=0$. (图形)

【证】由假设, 存在$\alpha,\beta\in[a,b]$ 使得$f(\alpha)=\min\limits_{x\in[a,b]}f(x)$, $f(\beta)=\max\limits_{x\in[a,b]}f(x)$. 若$\alpha,\beta\{a,b\}$, 则由$f(a)=f(b)$ 知$f(\alpha)=f(\beta)$ 从而f 在$[a,b]$ 恒为常数: $f(x)\equiv f(\alpha),x\in[a,b]$. 此时对任意$\xi\in(a,b)$ 都有$f'(\xi)=0$. 设例如$\beta\not\in\{a,b\}$ 即$\beta\in(a,b)$. 则存在$\delta>0$ 使得$(\beta-\delta,\beta+\delta)\subset(a,b)$ 因而$f(\beta)$ 也是f的一个局部极大值. 据Fermate 极值原理知$f'(\beta)=0$. 于是可取$\xi=\beta$. □

在函数的大范围分析中, 下面的中值定理最常用.

【**定理5.12. Lagrange 中值定理**[4]】设$f:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续, 在(a,b)内可微. 则

$$\exists\,\xi\in(a,b)\quad\text{s.t.}\quad f(b)-f(a)=f'(\xi)(b-a).$$

【注1】将等式两边乘以-1 有

$$f(a)-f(b)=f'(\xi)(a-b).$$

这表明Lagrange 中值定理对$a<b$ 和$a>b$ 都成立. (图象)

【注2】Lagrange 中值定理也常写成顾名思义的形式:

$$\exists\,\xi\in(a,b)\quad\text{s.t.}\quad f'(\xi)=\frac{f(b)-f(a)}{b-a}.$$

这等式可以理解为, $f'(\xi)$ 等于变量$y=f(x)$ 在$x\in[a,b]$上的平均变化率$\frac{f(b)-f(a)}{b-a}$. 也可理解为, $f'(\xi)$等于质点在时间段$[a,b]$上的平均速度$\frac{f(b)-f(a)}{b-a}$.

【Lagrange 中值定理的证明】方法是考虑辅助函数, 把Lagrange 中值定理归结为Rolle 定理. 令 (曲线减直线)

$$F(x)=f(x)-f(a)-\frac{f(b)-f(a)}{b-a}(x-a),\quad x\in[a,b].$$

则

$$F(a)=F(b)=0,\quad F'(x)=f'(x)-\frac{f(b)-f(a)}{b-a},\quad x\in(a,b).$$

[4]Joseph-Louis Lagrange, 1735-1813, 法国力学家数学家, 是分析力学的奠基人, 对抽象代数也有贡献.

据Rolle 定理, 存在$\xi \in (a,b)$ 使得$F'(\xi)=0$, 即$f'(\xi)=\frac{f(b)-f(a)}{a-b}$. 这等价于Lagrange 中值公式. $\square$

利用Lagrange 中值定理我们对导函数会有进一步认识：导函数的间断点(如果有的话)都是第二类的!

【命题5.13.】设$f: I \to \mathbb{R}$ 在区间I 上可微. 若$x_0 \in I$ 是$f'(x)$的一个间断点, 则x_0 必是$f'(x)$的第二类间断点, 即单侧极限$\lim\limits_{I\ni x\to x_0-} f'(x)$, $\lim\limits_{I\ni x\to x_0+} f'(x)$ 不都存在有限. 换言之, 若单侧极限$\lim\limits_{I\ni x\to x_0-} f'(x)$, $\lim\limits_{I\ni x\to x_0+} f'(x)$ 都存在有限, 则$f'(x)$ 在x_0 连续.
[注: 若$x_0 \in I$ 是I的一个端点, 例如$x_0 \in I$ 是I的左端点, 则上面的记号应被理解为只有$\lim\limits_{I\ni x\to x_0+} f'(x)$ 出现而$\lim\limits_{I\ni x\to x_0-} f'(x)$不出现.]
【证】我们只对$x_0 \in I^\circ$ (I的内部)的情形给出证明, 其论证过程已包含了x_0是I 的端点的情形的证明. 如命题中所言, 我们来证明若$A := \lim\limits_{I\ni x\to x_0-} f'(x), B := \lim\limits_{I\ni x\to x_0+} f'(x)$ 都存在有限, 则$A = B = f'(x_0)$ 因而$f'(x)$ 在x_0 连续.

对任意$x \in I$ 满足$x < x_0$, 由Lagrange 中值定理存在$x < \xi_x < x_0$ 使得

$$\frac{f(x)-f(x_0)}{x-x_0} = f'(\xi_x).$$

因$x \to x_0-$ 蕴含$\xi_x \to x_0-$, 故由左极限$A = \lim\limits_{I\ni x\to x_0-} f'(x)$ 存在有限和导数的定义可知

$$f'(x_0) = \lim_{x\to x_0-} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0-} f'(\xi_x) = A.$$

同理有

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

所以$A = B = f'(x_0)$. 因此$f'(x)$ 在x_0 连续. $\square$
配合上述命题, 看一个例子:

【例】令

$$f(x) = \begin{cases} x^2 \sin(\dfrac{1}{x}) & \text{当 } x \in \mathbb{R}\setminus\{0\}, \\ 0 & \text{当 } x = 0. \end{cases}$$

则f在$\mathbb{R}$ 上处处可导且导函数为

$$f'(x) = \begin{cases} 2x\sin(\dfrac{1}{x}) - \cos(\dfrac{1}{x}) & \text{当 } x \in \mathbb{R}\setminus\{0\}, \\ 0 & \text{当 } x = 0. \end{cases}$$

注意导数是点点定义的! 在注意$\mathbb{R}\setminus\{0\}=(-\infty,0)\cup(0,+\infty)$ 是开集, 容易验证当$x\in\mathbb{R}\setminus\{0\}$时$f'(x)=2x\sin(\frac{1}{x})-\cos(\frac{1}{x})$. 而当$x=0$ 时, 由

$$\lim_{x\to 0}\frac{f(x)-f(0)}{x}=\lim_{x\to 0}x\sin(\frac{1}{x})=0$$

知f在$x=0$ 可导且$f'(0)=0$. 这证明了$f'(x)$的上述表达式.

从$f'(x)$的表达式可知$\mathbb{R}\setminus\{0\}$中的每一点都是$f'(x)$的连续点. 看$f'(x)$在$x=0$处是否连续: 由

$$f'(\frac{1}{2n\pi})=0,\quad f'(\frac{1}{(2n+1/2)\pi})=1,\quad n=1,2,3,...$$

可知极限$\lim\limits_{x\to 0}f'(x)$ 不存在. 因此导函数$f'(x)$ 在$x=0$ 不连续且$x=0$ 是$f'(x)$的第二类不连续点(即第二类间断点). $\square$

Lagrange 中值定理在处理大范围分析时很有用, 因为它建立了函数的整体性质与局部性质(即导数)的基本关系.

【定理5.14. 关于单调函数】设$f:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续, 在(a,b)内可微. 则

- f 在$[a,b]$ 上单调不减(单调不增) $\iff$ $f'(x)\ge 0\,(\le 0\,)$ for all $x\in(a,b)$.
- f 在$[a,b]$ 上严格单调增加(严格单调减少) $\iff$ $f'(x)\ge 0\,(\le 0\,)$ for all $x\in(a,b)$ 并且$f'(x)$ 在(a,b) 内的任何子区间上不恒为零.

【证】我们以严格单调增加为例进行证明, 其它情形的证明要素已含在其中.

"$\Longrightarrow$": 设f 在$[a,b]$ 上严格单调增加. 对任意$x\in(a,b)$, 存在$\delta>0$ 使得$(x,x+\delta)\subset(a,b)$. 因此当$0<h<\delta$ 时$x+h\in(x,x+\delta)$ 从而有$\frac{f(x+h)-f(x)}{h}>0$. 于是得到

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

所以$f'(x)\ge 0$ for all $x\in(a,b)$. 假设存在$(\alpha,\beta)\subset(a,b)$ 使得$f'(x)\equiv 0, x\in(\alpha,\beta)$, 则由Lagrange 中值定理, 存在$\eta\in(\alpha,\beta)$ 使得$f(\beta)-f(\alpha)=f'(\eta)(\beta-\alpha)=0\cdot(\beta-\alpha)=0$. 这与$f$ 在$[a,b]$上严格单调增加矛盾.

"$\Longleftarrow$": 设$f'(x)\ge 0$ for all $x\in(a,b)$ 并且$f'(x)$ 在(a,b) 内的任何子区间上不恒为零.

对任意$x,y\in[a,b]$ with $x<y$. 由Lagrange 中值定理, 存在$z\in(x,y)$ 使得$f(y)-f(x)=f'(z)(y-x)\ge 0\cdot(y-x)=0$. 这证明了$f$ 在$[a,b]$ 上单调不减. 假设f 不是严格增加的, 则存在$a\le\alpha<\beta\le b$ 使得$f(\alpha)=f(\beta)$. 因f 在$[a,b]$ 上单调不减, 故这蕴含f 在$[\alpha,\beta]$

上等于常数: $f(x) \equiv f(\alpha), x \in [\alpha, \beta]$, 因而有$f'(x) \equiv 0, x \in [\alpha, \beta]$. 这与假设条件矛盾. 所以$f$ 必在$[a,b]$ 上严格增加. $\square$

【例】函数$f(x) = x^3$ 满足$f'(x) = 3x^2 \geq 0$ for all $x \in \mathbb{R}$ 且临界点方程$f'(x) = 0$只有一个解$x = 0$. 由定理5.12 知$f(x) = x^3$ 在$\mathbb{R}$ 上严格增加. 其反函数为$f^{-1}(y) = y^{1/3}$, $y \in \mathbb{R}$. 注意$f^{-1}(y) = y^{1/3}$ 在$y = 0$ 不可导, 或者你可以认为它在$y = 0$ 的变化率是无穷大, 这与$f'(0) = 0$差不多是一回事. $\square$

Rolle 定理和Lagrange 中值定理对于向量值函数或复值函数不成立, 例如对于复值函数$f(x) = e^{\mathrm{i}x}$ 有$f(2\pi) - f(0) = 0$. 但$f'(x) = \mathrm{i}f(x) \neq 0$ for all $x \in \mathbb{R}$.

但是情况不会很差, 因为至少成立Lagrange 中值不等式(也称微分中值不等式).

首先我们回忆向量的欧几里德范数$|\cdot|$:

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}, \quad \mathbf{x} = (x_1, x_2, ..., x_m) \in \mathbb{R}^m.$$

【定理5.15. 向量值函数的微分中值不等式】设$\mathbf{r} : [a,b] \to \mathbb{R}^m$ 在$[a,b]$ 上连续, 在(a,b) 内可微. 则

$$\exists \xi \in (a,b) \quad \text{s.t.} \quad |\mathbf{r}(b) - \mathbf{r}(a)| \leq |\mathbf{r}'(\xi)|(b-a).$$

【证】首先易见当$\mathbf{r}(b) - \mathbf{r}(a) = \mathbf{0}$ 时上式对任何$\xi \in (a,b)$ 成立. 以下设$\mathbf{r}(b) - \mathbf{r}(a) \neq \mathbf{0}$, 即$|\mathbf{r}(b) - \mathbf{r}(a)| > 0$. 让我们写

$$\mathbf{r}(t) = (x_1(t), x_2(t), ..., x_m(t)), \quad t \in [a,b].$$

由假设知$\mathbf{r}(t)$的每个坐标函数$x_i(t)$ 都在$[a,b]$ 上连续、在(a,b) 内可微. 任取向量$\mathbf{c} = (c_1, c_2, ..., c_m) \in \mathbb{R}^m$. 考虑由内积定义的实值函数

$$f(t) = \langle \mathbf{c}, \mathbf{r}(t) \rangle = \sum_{i=1}^{m} c_i x_i(t), \quad t \in [a,b].$$

我们有

$$f(b) - f(a) = \langle \mathbf{c}, \mathbf{r}(b) - \mathbf{r}(a) \rangle, \quad f'(t) = \sum_{i=1}^{m} c_i x_i'(t) = \langle \mathbf{c}, \mathbf{r}'(t) \rangle, \quad t \in (a,b).$$

据Lagrange 中值定理, 存在$\xi \in (a,b)$ (ξ与$\mathbf{c}$ 有关) 使得$f(b)-f(a)=f'(\xi)(b-a)$ i.e.

$$\langle \mathbf{c}, \mathbf{r}(b)-\mathbf{r}(a)\rangle = \langle \mathbf{c}, \mathbf{r}'(\xi)\rangle(b-a).$$

由Cauchy 不等式有$|\langle \mathbf{c}, \mathbf{r}'(\xi)\rangle| \le |\mathbf{c}||\mathbf{r}'(\xi)|$ 从而有

$$|\langle \mathbf{c}, \mathbf{r}(b)-\mathbf{r}(a)\rangle| = |\langle \mathbf{c}, \mathbf{r}'(\xi)\rangle|(b-a) \le |\mathbf{c}||\mathbf{r}'(\xi)|(b-a).$$

现在取$\mathbf{c}=\mathbf{r}(b)-\mathbf{r}(a)$. 则存在相应的$\xi \in (a,b)$ 使得

$$|\mathbf{r}(b)-\mathbf{r}(a)|^2 \le |\mathbf{r}(b)-\mathbf{r}(a)||\mathbf{r}'(\xi)|(b-a).$$

消去$|\mathbf{r}(b)-\mathbf{r}(a)|>0$ 即得

$$|\mathbf{r}(b)-\mathbf{r}(a)| \le |\mathbf{r}'(\xi)|(b-a). \qquad \square$$

下面的Cauchy 中值定理是Lagrange 中值定理的自然推广. 我们在后面将看到, 在研究两个函数的商f/g的极限时Cauchy 中值定理起重要作用. 而对于变量的精密估计, Cauchy 中值公式有时也比Lagrange 中值公式好.

【定理5.16. Cauchy 中值定理】设$f,g:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续, 在(a,b) 内可导. 则

(a) 若$g(a)\ne g(b)$. 则

$$\exists \xi \in (a,b) \quad \text{s.t.} \quad f'(\xi)=\frac{f(b)-f(a)}{g(b)-g(a)}g'(\xi).$$

(b) 若$g'(x)\ne 0$ for all $x\in(a,b)$, 则

$$\exists \xi \in (a,b) \quad \text{s.t.} \quad \frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}.$$

【证】如果$g'(x)\ne 0$ for all $x\in(a,b)$, 则由Rolle定理或Lagrange 中值定理易见$g(a)\ne g(b)$. 因此以下我们假设$g(a)\ne g(b)$.

如上, 考虑辅助函数, 把Cauchy 中值定理归结为Rolle 定理. 令

$$F(x)=f(x)-f(a)-\frac{f(b)-f(a)}{g(b)-g(a)}(g(x)-g(a)), \quad x\in[a,b].$$

则

$$F(a)=F(b)=0, \quad F'(x)=f'(x)-\frac{f(b)-f(a)}{g(b)-g(a)}g'(x), \quad x\in(a,b).$$

据Rolle 定理, 存在$\xi \in (a,b)$ 使得$F'(\xi)=0$, 即$f'(\xi)=\frac{f(b)-f(a)}{g(b)-g(a)}g'(\xi)$. $\square$

作为极值原理的另一重要应用, 我们有关于导函数的介值定理, 它不要求导函数连续.

【**定理5.17. 导函数的介值定理(Darboux 定理)**】设$I\subset\mathbb{R}$ 为区间, $f:I\to\mathbb{R}$ 在I 上可微. 则其导函数$f'(x)$ 具有介值性质: 对任意$a,b\in I$, 若$f'(a)\neq f'(b)$, 则对任意p 满足$f'(a)<p<f'(b)$ 或$f'(b)<p<f'(a)$, 存在$c\in(a,b)$ 或$c\in(b,a)$ 使得$f'(c)=p$. 换言之, $f'(I)$ 是一个区间.

[注意: 对于$I=[a,b]$ $(a<b)$ 的情形, $f'(a),f'(b)$ 自然是f在$[a,b]$上的**单侧**导数.]

【证】设$a,b\in I$ 且$f'(a)\neq f'(b)$. 调整记号可设$a<b$.

先假设0 介于$f'(a)$ 与$f'(b)$之间, 也即$f'(a)$ 与$f'(b)$ 反号. 不妨设$f'(a)<0<f'(b)$ (否则考虑$-f$ 便有$-f'(a)<0<-f'(b)$.] 来证明存在$c\in(a,b)$ 使得$f'(c)=0$.

方法1: 利用极值原理. 由导数的定义有

$$\lim_{x\to a+}\frac{f(x)-f(a)}{x-a}=f'(a)<0,\quad \lim_{x\to b-}\frac{f(x)-f(b)}{x-b}=f'(b)>0.$$

因此存在x_1,x_2 满足$a<x_1<x_2<b$ 使得

$$\frac{f(x_1)-f(a)}{x_1-a}<0,\quad \frac{f(x_2)-f(b)}{x_2-b}>0.$$

这蕴含 (图形)

$$f(x_1)<f(a),\quad f(x_2)<f(b).$$

因f 在I 上可微, 故f 在I 上从而在$[a,b]$ 上连续. 因此存在$c\in[a,b]$ 使得

$$f(c)=\min_{x\in[a,b]}f(x).$$

显然$f(c)\le\min\{f(x_1),f(x_2)\}<\min\{f(a),f(b)\}$. 因此$c\notin\{a,b\}$ 即$a<c<b$. 取$0<\delta<\min\{c-a,b-c\}$, 则$(c-\delta,c+\delta)\subset(a,b)$. 由$f(c)=\min\limits_{x\in[a,b]}f(x)$ 有$f(x)\ge f(c)$ for all $x\in(c-\delta,c+\delta)$. 因此c 也是f 在(a,b) 内的局部极小值点. 据Fermat 极值原理便有$f'(c)=0$.

方法2: 利用单调函数的特性和Rolle定理. 先证明f 在$[a,b]$ 上不是单射. 否则, f 在$[a,b]$ 上是单射, 则由可微蕴含连续和第四章**定理4.23** 知f 在$[a,b]$ 上严格单调. 由导数的定义可知: 若f 在$[a,b]$上单调增加, 则$f'(a)\ge 0$, 这与$f'(a)<0$ 矛盾; 若$f(x)$

在$[a,b]$ 上单调减少, 则$f'(b)\le 0$, 这与$0<f'(b)$ 矛盾. 这矛盾证明了f 在$[a,b]$ 上不是单射. 因此存在$x_1,x_2\in[a,b]$, $x_1\ne x_2$, 使得$f(x_1)=f(x_2)$. 不妨设$x_1<x_2$. 则由Rolle 定理知存在$c\in(x_1,x_2)$ 使得$f'(c)=0$.

对于一般情形, 对任意p 满足$f'(a)<p<f'(b)$, 令

$$F(x)=f(x)-px,\quad x\in I.$$

则F 在I 上可微, $F'(x)=f'(x)-p$, 且

$$F'(a)=f'(a)-p<0<f'(b)-p=F'(b).$$

因此由上面已证的结论知存在$c\in(a,b)$ 使得$F'(c)=0$, 即$f'(c)=p$.

最后证明$f'(I)$ 是一个区间. 根据第四章**命题4.14**(区间的刻画), 只需证明对任意$\alpha,\beta\in f'(I)$ with $\alpha<\beta$, 都有$[\alpha,\beta]\subset f'(I)$. 由$\alpha,\beta\in f'(I)$ 知存在$a,b\in I$ 使得$\alpha=f'(a),\beta=f'(b)$. 对任意$p\in[\alpha,\beta]$, 若$p=\alpha$ 或$p=\beta$, 则$p\in f'(I)$; 若$\alpha<p<\beta$, 则由上面已证的结论知存在$c\in(a,b)$ 或$c\in(b,a)$ 使得$p=f'(c)$ 从而$p\in f'(I)$. 所以$[\alpha,\beta]\subset f'(I)$. 所以$f'(I)$ 是一个区间. $\Box$

【Darboux 定理的推论1】设$f:I\to\mathbb{R}$ 在区间I 上可微且$f'(x)\ne 0$ for all $x\in I$. 则$f'(x)$ 在I 上不变号, 即要么$f'(x)>0$ for all $x\in I$ 从而f在I 上严格单调增加; 要么$f'(x)<0$ for all $x\in I$ 从而f 在I 上严格单调减少. $\Box$

【Darboux 定理的推论2】设$f:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上可微且导函数$f'(x)$在开区间(a,b)上单调(严格单调). 则$f'(x)$ 在闭区间$[a,b]$ 上连续且单调(严格单调).

【证】我们只对$f'(x)$在开区间(a,b)上严格单调增加的情形给予证明, 单调不减的情形以及其他情形的证明已含在其中了. 为证明$f'(x)$ 在$[a,b]$ 上严格单调增加, 只需证明$f'(a)<f'(x)<f'(b)$ for all $a<x<b$.

任取$a<x<b$. 取$a<c<x$. 则对任意$a<t<c$, 由Lagrange 中值定理存在$a<\xi_t<t$ 使得$\frac{f(t)-f(a)}{t-a}=f'(\xi_t)$. 再由导数的定义和$f'$ 在(a,b) 内严格单调增加即得

$$f'(a)=\lim_{t\to a+}\frac{f(t)-f(a)}{t-a}=\lim_{t\to a+}f'(\xi_t)\le f'(c)<f'(x).$$

同理可证$f'(x)<f'(b)$. 所以f' 在$[a,b]$ 上严格单调增加.

另一方面由Darboux 定理知$f'([a,b])$是一个区间. 因$f'(x)$在$[a,b]$上单调, 故由第四章**定理4.20** 知$f'(x)$ 在$[a,b]$ 上连续. $\square$

• 函数本身的Lagrange中值定理和Cauchy 中值定理可以推广到导函数, 其中区别只在于在导函数的情形, 只要求导函数在所考虑的闭区间上存在而不要求其连续。利用这个性质. 利用这个性质我们可以把Taylor公式中的条件减弱一点。迄今为止几乎所有《数学分析》教材在引进和证明具有Lagrange 余项的Taylor 公式和具有Cauchy 余项的Taylor 公式时, 都是要求函数在闭区间$[a,b]$ 上有**连续**的$n\,(\ge 1)$阶导函数, 在开区间(a,b) 有$n+1$ 阶导数. 2014年我在课上对此向同学们提问: 可否将上面的**"连续"**条件去掉？当时班上的**徐凯**同学在提问的第二天就证明了这是可以做到的. 徐凯的证明除了应用Darboux介值定理外, 重要的是概括出了一个函数类$S([a,b])$, 它差不多是使Cauchy中值定理(包括Lagrange 中值定理)成立的最大的类：

$$S([a,b])=\Big\{f:[a,b]\to\mathbb{R}\,\Big|\, f \text{ 在开区间 } (a,b) \text{ 内可微, 且}$$
$$\liminf_{x\to a+} f(x)\le f(a)\le \limsup_{x\to a+} f(x),\ \liminf_{x\to b-} f(x)\le f(b)\le \limsup_{x\to b-} f(x)\Big\}.$$

【命题5.A】设$f\in S([a,b])$, $g:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续, 在(a,b) 内可微. 则对任何实常数α,β 有线性组合$\beta f+\alpha g\in S([a,b])$.

【证】因αg 与g的性质相同故只需证明

$$\beta f\in S([a,b]),\quad f+g\in S([a,b]).$$

当$\beta=0$ 时显然$\beta f=0\in S([a,b])$. 设$\beta>0$. 则由上下极限运算法则知

$$\liminf_{x\to a+}\beta f(x)=\beta\liminf_{x\to a+} f(x),\quad \limsup_{x\to a+}\beta f(x)=\beta\limsup_{x\to a+} f(x),$$

此式对趋于b的极限也成立. 因此$\beta f\in S([a,b])$.

设$\beta<0$. 则由上下极限运算法则知

$$\begin{aligned}&\liminf_{x\to a+}\beta f(x)=\liminf_{x\to a+}\Big(-(-\beta f(x)\Big)=-\limsup_{x\to a+}(-\beta f(x))\\&=(-1)(-\beta)\limsup_{x\to a+} f(x)=\beta\limsup_{x\to a+} f(x),\\&\limsup_{x\to a+}\beta f(x)=\limsup_{x\to a+}\Big(-(-\beta f(x)\Big)=-\liminf_{x\to a+}(-\beta f(x))\\&=(-1)(-\beta)\liminf_{x\to a+} f(x)=\beta\liminf_{x\to a+} f(x).\end{aligned}$$

于是由$f\in S([a,b])$和$\beta<0$ 得到

$$\liminf_{x\to a+}\beta f(x)=\beta\limsup_{x\to a+}f(x)\le\beta f(a)\le\beta\liminf_{x\to a+}f(x)=\limsup_{x\to a+}\beta f(x).$$

此式对趋于b的极限也成立. 因此$\beta f\in S([a,b])$.

对于加法$f+g$, 根据上下极限运算法则和g在a,b 连续知

$$\liminf_{x\to a+}(f(x)+g(x))=\liminf_{x\to a+}f(x)+\lim_{x\to a}g(x)\le f(a)+g(a),$$

$$\limsup_{x\to b-}(f(x)+g(x))=\limsup_{x\to b-}f(x)+\lim_{x\to b-}g(x)\ge f(b)+g(b).$$

所以$f+g\in S([a,b])$.　　□

【**命题5.B**】设$f:[a,b]\to\mathbb{R}$ 在闭区间$[a,b]$ 上可导, 在开区间(a,b) 内两次可导. 又设$g:[a,b]\to\mathbb{R}$在$[a,b]$ 上连续, 在(a,b) 内可微. 则$f'g\in S([a,b])$. 特别取$g\equiv 1$ 知$f'\in S([a,b])$.

【证】由假设知f 在$[a,b]$ 上连续. 因此由Lagrange中值定理对任意$x\in(a,b)$ 存在$\xi_x\in(a,x)$ 使得

$$\frac{f(x)-f(a)}{x-a}=f'(\xi_x).$$

因此由导数的定义有

$$f'(a)=\lim_{x\to a+}\frac{f(x)-f(a)}{x-a}=\lim_{x\to a+}f'(\xi_x).$$

看

$$f'(\xi_x)g(\xi_x)=f'(\xi_x)(g(\xi_x)-g(x))+f'(\xi_x)g(x).$$

因$\lim\limits_{x\to a+}f'(\xi_x)=f'(a)$ 故存在$0<\delta<b-a$ 使得当$x\in(a,a+\delta)$ 时$|f'(\xi_x)|<M:=1+|f'(a)|$. 因此当$x\in(a,a+\delta)$ 时

$$|f'(\xi_x)(g(\xi_x)-g(x))|\le M|g(\xi_x)-g(a)|.$$

于是由g的连续性有$\lim\limits_{x\to a+}f'(\xi_x)(g(\xi_x)-g(x))=0$ 从而有

$$\lim_{x\to a+}f'(\xi_x)g(\xi_x)=\lim_{x\to a+}f'(\xi_x)g(x)=f'(a)g(a).$$

因$a<\xi_x\to a\,(x\to a+)$, 故易见

$$\liminf_{x\to a+}f'(x)g(x)\le\lim_{x\to a+}f'(\xi_x)g(\xi_x)\le\limsup_{x\to a+}f'(x)g(x)$$

即

$$\liminf_{x\to a+} f'(x)g(x) \le f'(a)g(a) \le \limsup_{x\to a+} f'(x)g(x).$$

同理有

$$\liminf_{x\to b-} f'(x)g(x) \le f'(b)g(b) \le \limsup_{x\to b-} f'(x)g(x).$$

所以$f'g \in S([a,b])$. $\square$

【**命题5.C**】若$f \in S([a,b])$ 且$f'(x) \ne 0$ for all $x \in (a,b)$, 则f 在$[a,b]$ 上严格单调.

【证】由【**Darboux 定理的推论1**】知f 在(a,b)内严格单调. 不妨设f 在(a,b)内严格单调增加. 为证明f 在$[a,b]$ 上严格单调增加, 只需证明$f(a) < f(x) < f(b)$ for all $a < x < b$.

任取$a < x < b$. 取x_1, x_2 满足$a < x_1 < x < x_2 < b$. 则对任意$a < t < x_1, x_2 < \tau < b$ 有

$$f(t) < f(x_1) < f(x) < f(x_2) < f(\tau).$$

因此由$f \in S([a,b])$ 知

$$f(a) \le \limsup_{t\to a+} f(t) \le f(x_1) < f(x) < f(x_2) \le \liminf_{\tau\to b-} f(\tau) \le f(b).$$

所以f 在$[a,b]$ 上严格单调. $\square$

【**定理5.18(较弱条件下的Cauchy 中值定理)**】

设$f, g \in S([a,b])$, f, g, 在(a,b) 内可微, 且$g'(x) \ne 0$ for all $x \in (a,b)$. 则

$$\exists \xi \in (a,b) \quad \text{s.t.} \quad \frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

【证】由假设和**命题5.C**知g在$[a,b]$ 上严格单调. 特别有$g(a) \ne g(b)$. 令

$$k = \frac{f(b)-f(a)}{g(b)-g(a)}, \quad F(x) = f(x) - f(a) - k(g(x) - g(a)), \quad x \in [a,b].$$

则

$$F(a) = F(b)\,(=0), \quad F'(x) = f'(x) - kg'(x), \quad x \in (a,b).$$

由**命题5.A**知$F \in S([a,b])$. 但F 在$[a,b]$ 上不是单调的, 故由**命题5.C**知存在$\xi \in (a,b)$使得$F'(\xi) = 0$, 即$f'(\xi) = kg'(\xi)$. 最后由$g'(\xi) \ne 0$ 即得$f'(\xi)/g'(\xi) = k$. $\square$

【**例1 (导函数的Cauchy 中值定理)**】设函数$f,g:[a,b]\to\mathbb{R}$ 在闭区间$[a,b]$上可微, 在开区间(a,b) 内两次可微且$g''(x)\neq 0$ for all $x\in(a,b)$. 则

$$\exists\xi\in(a,b)\qquad \text{s.t.}\qquad \frac{f'(b)-f'(a)}{g'(b)-g'(a)}=\frac{f''(\xi)}{g''(\xi)}.$$

【证】由假设和**命题5.B**知$f',g'\in S([a,b])$. 于是将**定理5.18(较弱条件下的Cauchy 中值定理)**应用于函数$f'(x),g'(x)$ 即得上述导函数的Cauchy 中值公式. □

【**例2**】设n为正整数, $F,G:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上有n 阶导数, 在(a,b)内有$n+1$ 阶导数且$G^{(k)}(x)\neq 0$ for all $x\in(a,b)$ and all $k=1,2,...,n+1$. 则有:

(1) 若$F(a)=F'(a)=\cdots=F^{(n)}(a)=0$, $G(a)=G'(a)=\cdots=G^{(n)}(a)=0$, 则

$$\exists\xi\in(a,b)\qquad \text{s.t.}\qquad \frac{F(b)}{G(b)}=\frac{F^{(n+1)}(\xi)}{G^{(n+1)}(\xi)}.$$

(2) 若$F(b)=F'(b)=\cdots=F^{(n)}(b)=0$, $G(b)=G'(b)=\cdots=G^{(n)}(b)=0$, 则

$$\exists\xi\in(a,b)\qquad \text{s.t.}\qquad \frac{F(a)}{G(a)}=\frac{F^{(n+1)}(\xi)}{G^{(n+1)}(\xi)}.$$

【证】(1) 我们对n 用归纳法. 当$n=1$ 时, 由Cauchy 中值定理和导函数的Cauchy 中值定理(见上面的**例1**)以及$F(a)=F'(a)=0, G(a)=G'(a)=0$ 可知存在$a<\xi<\xi_1<b$ 使得

$$\frac{F(b)}{G(b)}=\frac{F(b)-F(a)}{G(b)-G(a)}=\frac{F'(\xi_1)}{G'(\xi_1)}=\frac{F'(\xi_1)-F'(a)}{G'(\xi_1)-G'(a)}=\frac{F''(\xi)}{G''(\xi)}.$$

假设所证关系式对于$n=k$ 成立. 看$n=k+1$. 此时由$(G')^{(j)}(x)=G^{(j+1)}(x)$ 可知对任意$\beta\in(a,b]$, 一阶导函数F',G' 在$[a,\beta]$ 上满足$n=k$时的条件. 于是由Cauchy 中值定理和归纳假设可知依次存在$a<\beta<b$ 和$a<\xi<\beta$ 使得

$$\frac{F(b)}{G(b)}=\frac{F(b)-F(a)}{G(b)-G(a)}=\frac{F'(\beta)}{G'(\beta)}=\frac{(F')^{(k+1)}(\xi)}{(G')^{(k+1)}(\xi)}=\frac{F^{(k+2)}(\xi)}{G^{(k+2)}(\xi)}.$$

据归纳法原理, 这证明了(1) 成立。

(2) 令$\widetilde{F}(x)=F(a+b-x), \widetilde{G}(x)=G(a+b-x)$, $x\in[a,b]$. 则由复合函数微分的性质易见$\widetilde{F},\widetilde{G}$ 在$[a,b]$ 上有n 阶导数, 在(a,b)内有$n+1$ 阶导数且

$$\widetilde{G}^{(k)}(x)=G^{(k)}(a+b-x)(-1)^k\neq 0,\quad x\in(a,b),\quad k=1,2,...,n+1,$$

$$\widetilde{F}^{(k)}(a)=F^{(k)}(b)(-1)^k=0,\quad \widetilde{G}^{(k)}(a)=G^{(k)}(b)(-1)^k=0,\quad k=0,1,...,n.$$

这说明$\widetilde{F},\widetilde{G}$ 满足(1)中的条件. 于是由(1)即知(2) 成立. □

§5.4. Taylor 公式和Taylor 展开

函数$f(x)$在x_0的一次可微性是用局部线性近似定义的. 假如$f(x)$在x_0附近是n次可微的, 则相信$f(x)$在x_0附近与一个次数不超过n的多项式的近似程度是$(x-x_0)^n$的高阶无穷小. 这些属于局部行为.

另一方面, 对高阶可微函数的大范围性质而言, 例如设$f(x)$ 有$n+1$次导数, 我们也将有"$n+1$ 阶Lagrange 中值定理", 也即$f(x)$与一个次数不超过n的多项式$P(x)$的大范围"误差"可以浓缩成$\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$.

下面我们先讲整体(或大范围)性质, 然后讲局部性质.

1. 整体或大范围Taylor 公式.

对任一区间$I\subset\mathbb{R}$ (开, 闭, 半开半闭, 有界无界均可), 我们用I° 表示区间I 的内部, 即I° 是与I 有相同端点的**开区间**. 例如若$I=[a,b]$ 则$I^\circ=(a,b)$; 若$I=[a,b)$ 则$I^\circ=(a,b)$; 若$I=(a,b)$ 则$I^\circ=I$, etc.

【定理 5.19.具有Lagrange 余项的Taylor公式】设n为正整数, $f:I\to\mathbb{R}$ 在区间I上n次可导, 在开区间I°内$n+1$次可导. 则对任意$x_0,x\in I$, 存在$\xi\in(x_0,x)$ 或$\xi\in(x,x_0)$ 使得[5]

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

这一定理的推广形式是下列

【定理 5.20.具有Cauchy 余项的Taylor公式】设n 为正整数, 函数$f:I\to\mathbb{R}$ 在区间I上n次可导数, 在开区间I°内$n+1$次可导. 设$\varphi:I\to\mathbb{R}$ 在I 上连续, 在I°内可导且$\varphi'(x)\neq 0$ for all $x\in I^\circ$. 则对任意$x_0,x\in I$, 存在$\xi\in(x_0,x)$ 或$\xi\in(x,x_0)$ 使得

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{\varphi(x)-\varphi(x_0)}{\varphi'(\xi)n!}f^{(n+1)}(\xi)(x-\xi)^n.$$

【两个定理的证明】先证Cauchy 余项的Taylor公式.

任取$x_0,x\in I$. 当$x=x_0$ 时, 定理中的Taylor 公式自动成立且此时ξ 可取为I°内的任意一点. 设$x\neq x_0$. 令$I_{x_0,x},I^\circ_{x_0,x}$ 分别表示为以x_0,x 为端点的闭区间和开区间. 我们让f的展开点变动, 即考虑函数

$$F(t)=f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(t)}{k!}(x-t)^k,\quad t\in I_{x_0,x}.$$

[5] **当$x=x_0$ 时ξ 可取为I°内的任意一点. 在下一定理即具有Cauchy 余项的Taylor公式中也如此处理.**

由对f的假设知F 在开区间$I^\circ_{x_0,x}$内可微. 计算(对于$t\in I^\circ_{x_0,x}$)

$$\begin{aligned}F'(t)&=-\sum_{k=0}^{n}\Big(\frac{f^{(k)}(t)}{k!}(x-t)^k\Big)'_t\\&=-f'(t)-\sum_{k=1}^{n}\Big(\frac{f^{(k+1)}(t)}{k!}(x-t)^k-\frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1}\Big)\\&=-\frac{f^{(n+1)}(t)}{n!}(x-t)^n,\quad t\in I^\circ_{x_0,x}.\end{aligned}$$

另一方面令

$$h(t)=f(x)-\sum_{k=0}^{n-1}\frac{f^{(k)}(t)}{k!}(x-t)^k,\quad g(t)=\frac{(x-t)^n}{n!},\quad t\in I_{x_0,x}.$$

则

$$F(t)=h(t)-f^{(n)}(t)g(t),\quad t\in I_{x_0,x}.$$

注意$I_{x_0,x}=[x_0,x]$或$[x,x_0]$是闭区间. 设$S(I_{x_0,x})$ 是上一节定义的函数类. 由假设知$f^{(n-1)}(t)$ 在闭区间$I_{x_0,x}$上可微, 在开区间$I^\circ_{x_0,x}$内两次可微, $g(t)$ 在闭区间$I_{x_0,x}$ 上任意次可导. 据**命题5.B** 知$f^{(n)}(t)g(t)$ 属于函数类$S(I_{x_0,x})$. 而$h(t)$ 在闭区间$I_{x_0,x}$上可微。因此由**命题5.A** 知$F(t)$ 属于函数类$S(I_{x_0,x})$.

在$I_{x_0,x}$ 上对函数$F(t),\varphi(t)$ 应用**定理5.18(较弱条件下的Cauchy 中值定理)**, 则存在$\xi\in I^\circ_{x_0,x}$ 使得

$$\frac{F(x)-F(x_0)}{\varphi(x)-\varphi(x_0)}=\frac{F'(\xi)}{\varphi'(\xi)}=-\frac{f^{(n+1)}(\xi)}{\varphi'(\xi)n!}(x-\xi)^n.$$

但$F(x)=0$, 故

$$F(x_0)=(\varphi(x)-\varphi(x_0))\frac{f^{(n+1)}(\xi)}{\varphi'(\xi)n!}(x-\xi)^n.$$

将$F(x_0)=f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ 代入上式即得Cauchy 余项的Taylor公式.

最后取$\varphi(t)=(x-t)^{n+1}$, 则$\varphi'(\xi)=-(n+1)(x-\xi)^n$, $\varphi(x)-\varphi(x_0)=-(x-x_0)^{n+1}$,

$$F(x_0)=(\varphi(x)-\varphi(x_0))\frac{f^{(n+1)}(\xi)}{\varphi'(\xi)n!}(x-\xi)^n=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$$

也即得到了Lagrange 余项的Taylor公式. □

• 在应用中, Cauchy 余项中的函数φ 将根据问题的特点选取. 例如取$\varphi(t)=x-t$, 则得到Cauchy早先给出的Taylor公式:

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-x_0)$$

它在研究Taylor级数敛散性的某些临界问题时将起重要作用(见后面).

• 虽然Cauchy余项很灵活, 但在多数情况下, 常用的还是Lagrange 余项的Taylor公式, 而其中最常用的是二阶Taylor公式:

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(\xi)}{2}(x-x_0)^2.$$

这个简单优美的公式密切联系着几何(曲线的凹凸性, 曲率)、力学(速度与加速度) 和优化问题.

回忆连续可微函数类的记号. 设I 为区间.

$C(I)=C(I,\mathbb{R})$ （或$C(I,\mathbb{C})$）表示在I 上连续的实值(或复值)函数的集合.

$C^n(I)=C^n(I,\mathbb{R})$ （或$C^n(I,\mathbb{C})$）表示在I 上具有连续的n阶导数的实值(或复值)函数的集合($n\in\mathbb{N}$).

$C^\infty(I)$ 表示在I 上具有任意阶导数的实值(或复值)函数的集合. 注意可微蕴含连续, 因此$C^\infty(I)=\bigcap_{n=0}^\infty C^n(I)$.

Taylor 多项式与Taylor 级数. 若f在x_0 附近n次可微, 则称形如

$$P_n(x)=\sum_{k=0}^n\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k,\quad x\in\mathbb{R}$$

的多项式叫做f 的n 次(或次数不超过n)的Taylor 多项式.

设f在x_0 附近属于C^∞, 则称形如

$$\sum_{n=0}^\infty\frac{f^{(n)}(x_0)}{k!}(x-x_0)^n$$

的**形式幂级数** 叫做f 在点x_0 展开的Taylor 级数 或 f在点x_0的Taylor展开.

两个自然的问题是: Taylor 级数在多大范围内收敛？(显然当$x=x_0$时该级数总收敛); 若收敛, 是否收敛到$f(x)$?

利用具有Lagrange 余项或Cauchy余项的Taylor公式, 这问题很容易解答:

【定理5.21．函数能展开成其Taylor级数的充分必要条件】设$f\in C^\infty(I), x_0, x\in I$. 则

$$f(x)=\sum_{n=0}^\infty\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\quad(\text{收敛且相等})\quad\Longleftrightarrow\quad\lim_{n\to\infty}R_n(f,x)=0$$

其中

$$R_n(f,x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}\qquad \text{或}\qquad R_n(f,x)=\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-x_0)$$

是f的Taylor公式中的Lagrange 余项或Cauchy余项. (注意这里$\xi=\xi_{x_0,x,n}$ 介于x_0,x 之间且还依赖于n).

【证】由具有Lagrange 余项或Cauchy余项的Taylor公式有

$$f(x)=\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+R_N(f,x).$$

因此根据级数收敛的定义即知

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{k!}(x-x_0)^n\quad(\text{收敛})\quad\Longleftrightarrow\quad f(x)=\lim_{N\to\infty}\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

$$\Longleftrightarrow\quad\lim_{N\to\infty}\Big(f(x)-\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\Big)=\lim_{N\to\infty}R_N(f,x)=0\,.\qquad\Box$$

下面的推论给出两个常见的情形:

【定理5.21.的推论】设$f\in C^{\infty}(I), x_0\in I$.

(a) 若存在常数$A,B>0$ 使得

$$\sup_{x\in I}|f^{(n)}(x)|\le AB^n,\quad n=0,1,2,...,.$$

则

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{k!}(x-x_0)^n\qquad\forall\,x\in I.$$

(b)若存在常数$M>0$ 使得

$$\sup_{x\in I}|f^{(n)}(x)|\le M^n n!,\quad n=1,2,3,...,.$$

则对于$\delta=1/M$ 有

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\qquad\forall\,x\in I\cap(x_0-\delta,x_0+\delta).$$

【证】(a)留为作业(简单). 下证(b). 根据**定理5.21**, 只需证明余项$R_n(f,x)\to 0\,(n\to\infty)$ for all $x\in I\cap(x_0-\delta,x_0+\delta)$. 考虑Lagrange 余项: 由假设条件知当$x\in I$且$|x-x_0|<\delta=1/M$ 时有$0\le M|x-x_0|<1$ 从而有

$$|R_n(f,x)|=\frac{|f^{(n+1)}(\xi)|}{(n+1)!}|x-x_0|^{n+1}\le M^{n+1}|x-x_0|^{n+1}=(M|x-x_0|)^{n+1}\to 0$$

$(n\to\infty)$. $\square$

• **幂级数收敛性和连续性的Abel定理.**

【**Abel 定理**】对于$R>0$ 和形式幂级数

$$S(x)=\sum_{n=0}^{\infty}a_nx^n$$

若$S(R)$ 收敛, 则$S(x)$ 在$(-R,R]$ 上收敛, 在$[0,R]$ 上有界, 且$\lim\limits_{x\to R-}S(x)=S(R)$.

若$S(-R)$ 收敛, 则$S(x)$ 在$[-R,R)$ 上收敛, 在$[-R,0]$ 上有界, 且$\lim\limits_{x\to -R+}S(x)=S(-R)$.

因此若$S(\pm R)$均收敛, 则$S(x)$在$[-R,R]$ 上收敛且有界.

【证】设$S(R)$收敛. 则通项$a_nR^n\to 0\,(n\to\infty)$ 从而有界$C=\sup\limits_{n\ge 0}|a_nR^n|<+\infty$. 由此可见对任意$|x|<R$ 有

$$\sum_{n=0}^{\infty}|a_nx^n|=\sum_{n=0}^{\infty}|a_nR^n|(|x|/R)^n\le C=\sum_{n=0}^{\infty}(|x|/R)^n=C\frac{1}{1-|x|/R}<+\infty.$$

这表明$S(x)$ 在$(-R,R)$绝对收敛从而收敛. 因$S(R)$已收敛, 故$S(x)$ 在$(-R,R]$上收敛.

下证$S(x)$在$[0,R]$ 上有界且$\lim\limits_{x\to R-}S(x)=S(R)$. 由$S(R)$收敛和Cauchy收敛准则知, 对任意$\varepsilon>0$ 存在$N=N_\varepsilon\in\mathbb{N}$ 使得

$$\Big|\sum_{n=N+1}^{m}a_nR^n|<\varepsilon\qquad \forall\, m>N\,.$$

应用Abel 分部求和公式, 对任意$x\in[0,R]$, 令$t=x/R$, 则对任意$m>N+1$有

$$\sum_{n=N+1}^{m}a_nx^n=\sum_{n=n+1}^{m}a_nR^nt^n=\sum_{n=N+1}^{m-1}(\sum_{n=N+1}^{n}a_kR^k)(t^n-t^{n+1})+(\sum_{n=N+1}^{m}a_nR^n)t^m$$

$\Longrightarrow$

$$\Big|\sum_{n=N+1}^{m}a_nx^n\Big|\le\sum_{n=N+1}^{m-1}\Big|\sum_{k=N+1}^{n}a_kR^k\Big|(t^n-t^{n+1})+\Big|\sum_{n=N+1}^{m}a_nR^n\Big|t^m$$
$$\le\varepsilon\Big(\sum_{n=N+1}^{m-1}(t^n-t^{n+1})+t^m\Big)=\varepsilon t^{N+1}\le\varepsilon\,.$$

以上用到$0\le t\le 1$. 令$m\to\infty$ 得到

$$\Big|\sum_{n=N+1}^{\infty}a_nx^n\Big|\le\varepsilon\qquad\forall\,x\in[0,R].$$

由此和分解式

$$S(x)-S(R)=\sum_{n=0}^{N}a_nx^n-\sum_{n=0}^{N}a_nR^n+\sum_{n=N+1}^{\infty}a_nx^n-\sum_{n=N+1}^{\infty}a_nR^n,\quad x\in[0,R]$$

得到

$$|S(x)-S(R)|\le\Big|\sum_{n=0}^{N}a_nx^n-\sum_{n=0}^{N}a_nR^n\Big|+2\varepsilon,\quad \forall\, x\in[0,R].$$

取$\varepsilon=1$ 则有

$$|S(x)|\le|S(R)|+2\sum_{k=0}^{N}|a_k|R^k+2,\quad \forall\, x\in[0,R]$$

这证明了$S(x)$在$[0,R]$ 上有界. 现在保持$\varepsilon>0$的任意性. 因多项式$x\mapsto\sum\limits_{n=0}^{N}a_nx^n$ 连续, 故存在$0<\delta<R$ 使得

$$|\sum_{n=0}^{N}a_nx^n-\sum_{n=0}^{N}a_nR^n|<\varepsilon\qquad \forall\, x\in(R-\delta,R)\,.$$

于是由面的估计式便有

$$|S(x)-S(R)|<\varepsilon+2\varepsilon=3\varepsilon\qquad \forall\, x\in(R-\delta,R)\,.$$

这就证明了$\lim\limits_{x\to R-}S(x)=S(R)$.

最后设$S(-R)$收敛. 则幂级数$\widetilde{S}(x)=S(-x)$ 在$x=R$收敛. 将已证结论用于$\widetilde{S}(x)$ 即知$S(x)=\widetilde{S}(-x)$ 在$[-R,R)$ 上收敛, 在$[-R,0]$ 上有界, 且$\lim\limits_{x\to-R+}S(x)=S(-R)$. □

【命题5.22. 幂级数的唯一性】设J是一个包含x_0的区间[6] ,

$$\sum_{n=0}^{\infty}a_n(x-x_0)^n=\sum_{n=0}^{\infty}b_n(x-x_0)^n\quad(\text{收敛且相等})\qquad\forall\, x\in J.$$

则$a_n=b_n$, $n=0,1,2,\dots$.

【证】令$c_n=a_n-b_n$. 要证一切$c_n=0$. 令$t=x-x_0$, $I=\{x-x_0\,|\,x\in J\}$. 则I 是包含0的区间. 由假设知$\sum_{n=0}^{\infty}c_nt^n$ 在I 上收敛且

$$\sum_{n=0}^{\infty}c_nt^n=0\qquad\forall\, t\in I.\tag{$*$}$$

取$t=0$ 得到$c_0=0$. 设对于非负整数k 有$c_0=c_1=\cdots=c_k=0$. 看c_{k+1}. 此时$(*)$ 变成

$$\sum_{n=k+1}^{\infty}c_nt^n=0\qquad\forall\, t\in I.$$

[6]按惯例, 除非特别说明, 区间都是指非退化区间.

即$t^{k+1}\sum_{n=0}^{\infty}c_{n+k+1}t^n=0\ \forall\, t\in I$. 这蕴含$\sum_{n=0}^{\infty}c_{n+k+1}t^n=0\ \ \forall\, t\in I\setminus\{0\}$, 也即

$$c_{k+1}=-t\sum_{n=0}^{\infty}c_{n+k+2}t^n\qquad \forall\, t\in I\setminus\{0\}.$$

易见级数$\sum_{n=0}^{\infty}c_{n+k+2}t^n$ 对每个$t\in I\setminus\{0\}$ 从而对每个$t\in I$ 收敛. 取$t_1\in I\setminus\{0\}$ 使$\sum_{n=0}^{\infty}c_{n+k+2}t_1^n$ 收敛. 由Abel 定理知, 幂级数$\sum_{n=0}^{\infty}c_{n+k+2}t^n$ 在$(-|t_1|,|t_1|)$内收敛从而它在$[-|t_1|/2,|t_1|/2]$ 上有界. 于是令t 沿$I\setminus\{0\}$ 趋于0 得到

$$c_{k+1}=-\lim_{I\ni t\to 0}t\sum_{n=0}^{\infty}c_{n+k+2}t^n=0.$$

据归纳法原理, 这证明了一切$c_n=0, n=0,1,2,\dots$. □

【定理5.23. Taylor级数的唯一性】设$f\in C^{\infty}(I), x_0\in I$. 假设$f(x)$在$x_0$附近可以展开成它的Taylor级数(即$f(x)$的Taylor级数在$x_0$附近收敛于$f(x)$). 又设存在包含$x_0$的子区间$I_0\subset I$ 使得

$$f(x)=\sum_{n=0}^{\infty}a_n(x-x_0)^n,\quad x\in I_0.$$

则必有

$$a_n=\frac{f^{(n)}(x_0)}{n!},\quad n=0,1,2,\dots.$$

因此无论用什么方法, 只要将$f(x)$表示成$x-x_0$的幂级数, 则这幂级数就是f在x_0展开的Taylor级数.

【证】由假设知存在包含x_0的子区间$I_1\subset I$ 使得

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\qquad \forall\, x\in I_1.$$

于是对于包含x_0的区间$J=I_0\cap I_1$ 有

$$\sum_{n=0}^{\infty}a_n(x-x_0)^n=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\qquad \forall\, x\in J.$$

根据**命题5.22**(幂级数的唯一性) 即知$a_n=\frac{f^{(n)}(x_0)}{n!},\ n=0,1,2,\dots$. □

基本初等函数的Taylor公式和Taylor 展开.

(1) 指数函数$f(x)=e^x$, Taylor公式Lagrange 余项:

$f^{(n)}(x) = f(x) = e^x$, $f^{(n)}(0) = 1 \Longrightarrow$

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{\theta x}}{(n+1)!} x^{n+1}, \quad 0 < \theta = \theta_{x,n} < 1, \qquad x \in \mathbb{R}.$$

根据典型极限知

$$\lim_{n\to\infty} \frac{|x|^n}{n!} = 0 \qquad \forall\, x \in \mathbb{R}.$$

因此余项

$$\Big|\frac{e^{\theta x}}{(n+1)!} x^{n+1}\Big| \le e^{|x|} \frac{|x|^{n+1}}{(n+1)!} \to 0 \quad (n \to \infty).$$

所以

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \forall\, x \in \mathbb{R}.$$

在Taylor公式中取$x = 1$ 得到e的有限表示

$$e = \sum_{k=0}^{n} \frac{1}{k!} + \frac{e^{\theta_n}}{(n+1)!}, \qquad 0 < \theta_n < 1, \quad n = 1, 2, 3, \dots.$$

(2) 正弦余弦函数: $f(x) = \sin x, g(x) = \cos x$, Taylor 公式Lagrange 余项:

$f^{(n)}(x) = \sin(x + n\pi/2), f^{(2k-1))}(0) = \sin((2k-1)\pi/2) = (-1)^{k-1}, f^{(2k))}(0) = 0;$

$g^{(n)}(x) = \cos(x + n\pi/2), g^{(2k)}(0) = \cos(k\pi) = (-1)^k, g^{(2k+1)}(0) = 0 \;\Longrightarrow$

$$\sin x = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} + \frac{\sin(\theta x + n\pi)}{(2n)!} x^{2n},$$
$$\cos x = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + \frac{\cos(\theta x + (2n+1)\pi/2)}{(2n+1)!} x^{2n+1}.$$

如同(1)中的证明有

$$\sin x = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!}, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \qquad \forall\; x \in \mathbb{R}.$$

(3) 对数函数$f(x) = \log(1 + x)$,

$$f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}, \quad f^{(n)}(0) = (-1)^{n-1}(n-1)!.$$

Taylor公式Lagrange 余项:

$$\log(1 + x) = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^k}{k} + (-1)^n \frac{x^{n+1}}{(n+1)(1+\xi)^{n+1}}, \quad x > -1.$$

Taylor公式Cauchy 余项:

$$\log(1+x)=\sum_{k=1}^{n}(-1)^{k-1}\frac{x^k}{k}+(-1)^n\Big(\frac{x-\xi}{1+\xi}\Big)^n\frac{x}{1+\xi},\quad x>-1.$$

ξ 介于$0, x$ 之间, 因此ξ, x 同号!

对于Cauchy 余项我们有(因ξ, x 同号且$|\xi|<|x|$)

$$\begin{aligned}|x|<1 \implies &\Big|(-1)^n\Big(\frac{x-\xi}{1+\xi}\Big)^n\frac{x}{1+\xi}\Big|=\Big(\frac{|x|-|\xi|}{1+\xi}\Big)^n\frac{|x|}{1+\xi}\\ &\le\Big(\frac{|x|-|\xi|}{1-|\xi|}\Big)^n\frac{|x|}{1-|x|}\le\frac{|x|^{n+1}}{1-|x|}\to 0\ (n\to\infty)\end{aligned}$$

$\Longrightarrow$

$$\log(1+x)=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n},\quad \forall\, x\in(-1,1).$$

对于Lagrange 余项我们有(因ξ, x同号)

$$0\le x\le 1\quad\Longrightarrow\quad\Big|(-1)^n\frac{x^{n+1}}{(n+1)(1+\xi)^{n+1}}\Big|\le\frac{1}{n+1}\to 0\ (n\to\infty).$$

因此总结果是:

$$\log(1+x)=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n},\quad \forall\, x\in(-1,1].$$

特别取$x=1$ 得到

$$\log 2=\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}.$$

注意当$x>1$ 时通项$|(-1)^{n-1}\frac{x^n}{n}|=\frac{x^n}{n}\to+\infty\,(n\to\infty)$. 因此对数函数$\log(1+x)$在$x=0$ 展开的Taylor级数在且只在$(-1,1]$上收敛.

(3) 幂函数$f(x)=(1+x)^\alpha$. 设$\alpha\in\mathbb{R}\setminus\{0,1,2,...\}$. 照顾到一般情形, 我们先考虑$x\ge -1$ 或$x>-1$. 此时

$$f^{(n)}(x)=\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)(1+x)^{\alpha-n},\quad x>-1.$$

Taylor公式:

$$(1+x)^\alpha=1+\sum_{k=1}^{n}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k+R_n(x),$$

$$\text{Lagrange 余项}:\quad R_n(x)=\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)(1+\xi)^{\alpha-n-1}}{(n+1)!}x^{n+1},$$

$$\text{Cauchy 余项：}\quad R_n(x) = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)}{n!} x(1+\xi)^{\alpha-n-1}(x-\xi)^n.$$

利用Cauchy 余项我们有

$$|x|<1 \implies |\xi|<|x|<1 \implies |(1+\xi)^{\alpha-n-1}(x-\xi)^n| = |(1+\xi)^{\alpha-1}|\Big(\frac{|x-\xi|}{|1+\xi|}\Big)^n$$

$$\le C(x)\Big(\frac{|x|-|\xi||}{1-|\xi|}\Big)^n \le C(x)|x|^n$$

其中$C(x) = \left(\frac{1}{1-|x|}\right)^{1-\alpha}$ if $\alpha<1, C(x) = 2^{\alpha-1}$ if $\alpha>1$. 因此

$$|R_n(x)| \le C(x)\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)||x|^{n+1}}{n!}.$$

令

$$a_n = \frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)|}{n!}.$$

则有

$$\frac{a_{n+1}}{a_n} = \left|1-\frac{\alpha}{n+1}\right| \to 1\ (n\to\infty) \quad\implies\quad \lim_{n\to\infty}(a_n)^{1/n} = \lim_{n\to\infty}\frac{a_{n+1}}{a_n} = 1$$

(见第三章典型极限). 因

$$|R_n(x)| \le C(x)a_n|x|^{n+1} = C(x)|x|a_n|x|^n, \quad |x|<1$$

故当$|x|<1$ 时

$$\limsup_{n\to\infty}|R_n(x)|^{1/n} \le \limsup_{n\to\infty}\Big((C(x)|x|)^{1/n}(a_n)^{1/n}\Big)|x| = |x| < 1$$

从而有

$$\limsup_{n\to\infty}|R_n(x)| = 0.$$

因此

$$(1+x)^\alpha = 1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}x^n \quad (\text{收敛}) \qquad \forall\,|x|<1.$$

而当$|x|>1$ 时, 对余项$R_n(x)$的估计不方便, 但是从上面的分析易见此时$(1+x)^\alpha$的Taylor级数的通项不趋于零！事实上从这个Taylor级数中x^n的系数的绝对值与a_{n-1}的关系式

$$\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!} = \frac{a_{n-1}}{n} \qquad (\,n>1\,)$$

和

$$(a_{n-1})^{\frac{1}{n}} = \left((a_{n-1})^{\frac{1}{n-1}}\right)^{\frac{n-1}{n}} \to 1 \quad (n\to\infty)$$

可知

$$\lim_{n\to\infty}\left(\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!}\right)^{1/n} = \lim_{n\to\infty}\left(\frac{1}{n}\right)^{1/n}(a_{n-1})^{1/n} = 1.$$

因此当$|x|>1$ 时

$$\lim_{n\to\infty}\left(\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!}|x|^n\right)^{1/n} = |x|>1$$

从而有

$$\lim_{n\to\infty}\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!}|x|^n = +\infty.$$

所以当$|x|>1$ 时$(1+x)^\alpha$ 的Taylor级数发散.

而当$|x|=1$ 时, $(1+x)^\alpha$的Taylor级数的敛散性依赖于α 的取值范围(见下面).

牛顿最早研究了$(1+x)^\alpha$的展开式, 所以人们把$(1+x)^\alpha$ (在$x=0$展开的)Taylor 级数叫做$(1+x)^\alpha$ 的**牛顿二项式展开**. 易见当$\alpha=m$ 为正整数时, 牛顿二项式展开中x^n的系数当$n\ge m+1$ 时均为零. 此时牛顿二项式展开式就是多项式$(1+x)^m$的二项式公式.

• 下面我们利用Abel定理证明: 若$\alpha>0$, 则$(1+x)^\alpha$的牛顿二项式展开在整个闭区间$[-1,1]$ 上都收敛于$(1+x)^\alpha$.

设$\alpha>0$. 令$S(x)$ 为$(1+x)^\alpha$的牛顿二项式展开, 即

$$S(x) = 1+\sum_{n=1}^\infty \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}x^n.$$

来证明$S(\pm1)$ 绝对收敛, 即

$$\sum_{n=1}^\infty \frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!} < +\infty.$$

为此我们充分地来证明通项的衰减估计式:

$$\frac{|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)|}{n!} \le C_\alpha\frac{1}{n^{1+\alpha}}, \qquad n=1,2,3,... \tag{$\spadesuit$}$$

其中$C_\alpha>0$ 是只与$\alpha>0$ 有关的常数.

先考虑$n\ge 3+[\alpha]$. 利用不等式

$$1+t\le e^t,\quad 1+\frac{1}{2}+\cdots+\frac{1}{n-1}\ge \log n$$

有

$$
\begin{aligned}
&\frac{|\alpha(\alpha-1)\cdots(\alpha-n+1)|}{n!}=|\alpha||1-\alpha||1-\frac{\alpha}{2}||1-\frac{\alpha}{3}|\cdots|1-\frac{\alpha}{[\alpha]+1}|\cdots|1-\frac{\alpha}{n-1}|\frac{1}{n}\\
&=|\alpha(1-\alpha)(1-\frac{\alpha}{2})|\cdots(1-\frac{\alpha}{[\alpha]+1})(1-\frac{\alpha}{[\alpha]+2})\cdots(1-\frac{\alpha}{n-1})\frac{1}{n}\\
&\le|\alpha(1-\alpha)(1-\frac{\alpha}{2})|\cdots(1-\frac{\alpha}{[\alpha]+1})e^{-\frac{\alpha}{[\alpha]+2}}\cdots e^{-\frac{\alpha}{n-1}}\frac{1}{n}\\
&=|\alpha(1-\alpha)(1-\frac{\alpha}{2})|\cdots(1-\frac{\alpha}{[\alpha]+1})e^{-\alpha(\frac{1}{[\alpha]+2}+\cdots+\frac{1}{n-1})}\frac{1}{n}\\
&=C_\alpha\frac{1}{n}e^{-\alpha(1+\frac{1}{2}+\cdots+\frac{1}{n-1})}\le C_\alpha\frac{1}{n}e^{-\alpha\log n}=C_\alpha\frac{1}{n^{1+\alpha}}.
\end{aligned}
$$

然后将常数C_α适当放大使得不等式(♠)对$1\le n\le 2+[\alpha]$自动成立.

由这一衰减估计便有

$$
\sum_{n=1}^{\infty}\frac{|\alpha(\alpha-1)\cdots(\alpha-n+1)|}{n!}\le C_\alpha\sum_{n=1}^{\infty}\frac{1}{n^{1+\alpha}}<\infty\,.
$$

所以$S(\pm1)$ 绝对收敛.

因当$x\in(-1,1)$ 时$(1+x)^\alpha=S(x)$, 故由Abel 定理和$(1+x)^\alpha$在$[-1,1]$ 上的连续性(因$\alpha>0$)知

$$
\begin{aligned}
(1+x)^\alpha\Big|_{x=1}&=\lim_{x\to1-}(1+x)^\alpha=\lim_{x\to1-}S(x)=S(1),\\
(1+x)^\alpha\Big|_{x=-1}&=\lim_{x\to-1+}(1+x)^\alpha=\lim_{x\to-1+}S(x)=S(-1).
\end{aligned}
$$

这证明了$(1+x)^\alpha=S(x)\ \ \forall x\in[-1,1]$.

由上式还得到: 当$\alpha>0$ 时

$$
2^\alpha=1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}.
$$

$$
0=1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}(-1)^n.
$$

• 从幂级数理论中大家将会看到, 幂级数在其收敛区间的内部是无穷次可微的. 因此像折线函数$|x-x_0|, a+b|x-x_0|\,(b\ne0)$ 在折点x_0 附近就不可能表示成$x-x_0$的幂级数. 但是一个有趣且具有重要意义的现象是, (例如)折线函数$|x|$ 在$x=0$ 附近可以表示成x^2-1的幂级数. 事实上对$(1+t)^{1/2}$做牛顿二项式展开($|t|\le1$) 然后取$t=x^2-1$ 就得到

$$
|x|=1+\sum_{n=1}^{\infty}\frac{\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-n+1)}{n!}(x^2-1)^n,\quad x\in[-1,1].
$$

进一步, 令上式右端的部分和为

$$P_n(x)=1+\sum_{k=1}^{n}\frac{\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-k+1)}{k!}(x^2-1)^k.$$

则$P_n(x)$是$2n$次多项式, 它们对连续函数$|x|$在$[-1,1]$上有一致逼近:

$$\max_{x\in[-1,1]}\Big|P_n(x)-|x|\Big|\le C\frac{1}{\sqrt{n}},\quad n=1,2,3,\ldots.$$

其中$C>0$ 为常数.

事实上在不等式(♠)中取$\alpha=1/2$ 就得到

$$\max_{x\in[-1,1]}\Big|P_n(x)-|x|\Big|\le\sum_{k=n+1}^{\infty}\frac{|\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-k+1)|}{k!}\le C_{1/2}\sum_{k=n+1}^{\infty}\frac{1}{k^{3/2}}$$
$$\le 2C_{1/2}\sum_{k=n+1}^{\infty}\Big(\frac{1}{\sqrt{k-1}}-\frac{1}{\sqrt{k}}\Big)=C\frac{1}{\sqrt{n}}$$

这里我们用到了Lagrange中值定理: 对于$k>1$ 存在$k-1<\xi_k<k$ 使得

$$(k-1)^{-1/2}-k^{-1/2}=(-\frac{1}{2})\xi_k^{-3/2}(-1)=\frac{1}{2}\xi_k^{-3/2}\ge\frac{1}{2}k^{-3/2}.$$

【注】上面这个特殊的多项式逼近性质将来在研究生阶段学习和证明著名的Stone-Weierstrass 逼近定理时会用到.

函数的刚性和柔性. 一个函数$f(x)$ 若能在某区间上展开成$x-x_0$的幂级数, 也即在该区间上, $f(x)$的Taylor级数收敛到$f(x)$, 则称f在这区间上是解析的. 若某区间I上的两个解析函数f,g在点$x_0\in I^\circ$ 的各阶导数都相同, 或等价地, $f(x)-g(x)=0$ 在x_0的某个小邻域内有无穷多个根, 则不难证明$f(x)\equiv g(x)$ 于I. 这个性质说明, 与多项式相似, 解析函数如果局部地确定了, 则它就整体被确定了, 没有其他自由度了. 这就是解析函数的刚性. 注意C^∞ 光滑性远非刚性的本质！大量C^∞函数不是解析的; 他们可以在不相交的闭区间上互不干扰地随意调整而仍保持整体的C^∞性. 这就是函数的柔性. 我们将在习题课上制作一些C^∞函数, 他们在指定区间$[a,b]$上不为零, 例如等于1, 而在$[a-\delta,b+\delta]$之外恒等于零($\delta>0$), 等等. 用于构造这种柔性函数的一个常用函数就是下面这个著名的C^∞函数:

$$f(x)=\begin{cases}e^{-1/x} & \text{if}\quad x>0,\\ 0 & \text{if}\quad x\le 0.\end{cases}$$

不难验证$f\in C^{\infty}(\mathbb{R})$ 且$f^{(n)}(0)=0, n=0,1,2,...,$. 这后一条性质使得函数$f(x)$在$x=0$ 展开的Taylor级数恒等于零. 于是由$f(x)>0$ for all $x>0$ 可知

$$f(x)\neq\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n\qquad\forall\, x>0.$$

细节上看, 根据**定理5.21**(函数能展开成其Taylor级数的充分必要条件), $f(x)$的Taylor公式中的余项$R_n(f,x)$在$x=0$的任一邻域内的变化幅度都不小, 因而$R_n(f,x)\not\longrightarrow 0\,(n\to\infty)$. 利用这个$f(x)$, 根据复合函数微分法易见函数$g(x)=f(x^2)$ 属于$C^{\infty}(\mathbb{R})$ 且$g^{(n)}(0)=0, n=0,1,2,...,$. 但$g(x)\neq 0$ for all $x\neq 0$, 故g在0附近也不是解析的. 又若令$\varphi(x)=f(1-x^2)$, 则φ 属于$C^{\infty}(\mathbb{R})$ 且$\varphi(x)>0$ for all $x\in(-1,1)$; $\varphi(x)=0$ for all $|x|\geq 1$. φ 在$x=\pm 1$ 附近不是解析的.

2. 函数的局部Taylor公式.

函数的局部Taylor公式完全是函数在一点处的一阶微分向高阶微分的直接推广:

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+o((x-x_0)^n).$$

由于我们只关心在$f(x)$在x靠近x_0的行为, 即局部的多项式行为, 因此对$(x-x_0)^n$的高阶无穷小$o((x-x_0)^n)$不做细致要求, 意即我们用同一个记号$o((x-x_0)^n)$表示所有$(x-x_0)^n$的高阶无穷小. 但这就要求在运算中保持

$$o((x-x_0)^n)\pm o((x-x_0)^n)=o((x-x_0)^n),\quad \text{等等}$$

而不要误算成"$o((x-x_0)^n)-o((x-x_0)^n)=0$"!

提醒注意: 函数在一点处有n阶导数意味着这函数在此点附近已有$n-1$阶导数. 同时注意: 单独一个高阶无穷小$o((x-x_0)^n)$不能保证函数在x_0 有n阶导数. 例如$f(x)=x^9 1_{\mathbb{Q}}(x)=o(x^8)$, 但$f(x)$只在$x=0$这一点可微因此谈不上高阶可微.

【引理5.24】设n 为正整数, 函数$\varphi(x)$在点x_0 附近有$n-1$阶导数, 在点x_0处有n阶导数. 又设

$$\varphi(x_0)=\varphi'(x_0)=\cdots=\varphi^{(n)}(x_0)=0.$$

则

$$\varphi(x)=o((x-x_0)^n).$$

【证】我们对引理中的$n \in \mathbb{N}$ 用归纳法. 当$n = 1$时, 由可微的定义有

$$\varphi(x) = \varphi(x_0) + \varphi'(x_0)(x - x_0) + o(x - x_0) = o(x - x_0).$$

假设引理对于某个$n \in \mathbb{N}$ 成立. 则当引理中的正整数为$n + 1$ 时, 导函数$\varphi'(x)$ 满足引理中的正整数为n时的条件. 因此$\varphi'(x) = o((x - x_0)^n) = \alpha(x)(x - x_0)^n$ 其中$\alpha(x)$满足$\alpha(x) \to 0$ as $x \to x_0$. 此外由于$\varphi'(x_0) = 0$我们可以定义$\alpha(x_0) = 0$因而$\alpha(x)$ 在x_0连续. 因$n \ge 1$, 故由可微蕴含连续和Lagrnge中值定理知, 对x_0 附近的任意x, 存在ξ 介于x_0, x 之间使得

$$\varphi(x) = \varphi'(\xi)(x - x_0) = \alpha(\xi)(\xi - x_0)^n(x - x_0).$$

因ξ 介于x, x_0 之间, 故$|\xi - x_0| \le |x - x_0|$ 从而有

$$|\varphi(x)| \le |\alpha(\xi)||x - x_0|^{n+1}.$$

因$x \to x_0 \implies \xi \to x_0 \implies \alpha(\xi) \to 0$, 故$\varphi(x) = o((x - x_0)^{n+1})$. 据归纳法原理知引理任意$n \in \mathbb{N}$ 成立. $\square$

【引理5.25】设n 为正整数, $P_n(x) = \sum_{k=0}^{n} c_k(x - x_0)^k$ 为一多项式. 则

$$P_n(x) = o((x - x_0)^n)\ (x \to x_0) \quad \implies \quad c_k = 0, \quad k = 0, 1, ..., n.$$

【证】易见$P(x_0) = 0$ 即$c_0 = 0$. 设对于$m \le n - 1$已有$c_k = 0, k = 0, 1, ..., m$. 则

$$P_n(x) = \sum_{k=0}^{n} c_k(x - x_0)^k = c_{m+1}(x - x_0)^{m+1} + c_{m+2}(x - x_0)^{m+2} + \cdots + c_n(x - x_0)^n.$$

从而由$P_n(x) = o((x - x_0)^n)$ 和$m + 1 \le n$ 有

$$c_{m+1} = \lim_{x \to x_0} \frac{P_n(x)}{(x - x_0)^{m+1}} = \lim_{x \to x_0} \frac{o(x - x_0)^n}{(x - x_0)^n} \cdot (x - x_0)^{n-m-1} = 0.$$

据归纳法原理, 这证明了$c_k = 0$, $k = 0, 1, ..., n$. $\square$

【定理5.26.】设n 为正整数, $f(x)$ 在x_0 附近有$n - 1$阶导数, 在x_0 有n阶导数. 则在x_0附近有

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + o((x - x_0)^n). \tag{$*$}$$

此外若

$$f(x)=\sum_{k=0}^{n}a_k(x-x_0)^k+o((x-x_0)^n)$$

则

$$a_k=\frac{f^{(k)}(x_0)}{k!},\quad k=0,1,2,...,n.$$

【证】令

$$\varphi(x)=f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

则$\varphi^{(k)}(x_0)=0,k=0,1,...,n$. 由**引理5.24** 知$\varphi(x)=o((x-x_0)^n)$. 所以$(*)$成立.

设又有$f(x)=\sum\limits_{k=0}^{n}a_k(x-x_0)^k+o((x-x_0)^n)$, 则由$f$的两个表达式得到

$$0=f(x)-f(x)=\sum_{k=0}^{n}\Big(a_k-\frac{f^{(k)}(x_0)}{k!}\Big)(x-x_0)^k+o((x-x_0)^n)-o((x-x_0)^n)$$

从而有

$$\sum_{k=0}^{n}\Big(a_k-\frac{f^{(k)}(x_0)}{k!}\Big)(x-x_0)^k=o((x-x_0)^n)\qquad(\,x\to x_0\,).$$

于是由**引理5.25** 即得 $a_k-\frac{f^{(k)}(x_0)}{k!}=0,k=0,1,...,n.$ □

对于较好的函数, 例如初等函数, 常用的是具有$O((x-x_0)^{n+1})$余项的局部Taylor 公式:

【定理5.27】设n 为正整数, $f(x)$ 在x_0 附近有$n+1$阶导数且$f^{(n+1)}(x)$ 在x_0 附近有界. 则在x_0 附近有

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+O((x-x_0)^{n+1}).$$

【证】由假设知存在$\delta>0$ 使得$f(x)$在$I:=[x_0-\delta,x_0+\delta]$上有$n+1$阶导数且$f^{(n+1)}(x)$在$I$上有界: $M=\sup\limits_{x\in I}|f^{(n+1)}(x)|<+\infty$. 据具有Lagrange余项的Taylor 公式, 对任意$x\in I$有

$$\Big|f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\Big|=\frac{|f^{(n+1)}(\xi)|}{(n+1)!}|x-x_0|^{n+1}\le\frac{M}{(n+1)!}|x-x_0|^{n+1}$$

其中ξ 介于x_0,x 之间因而属于I. 由$O((x-x_0)^{n+1})$的定义即知
$f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k=O((x-x_0)^{n+1}).$ □

- **Taylor展开的唯一性的应用**

A. 关于复合函数的Taylor展开/Taylor公式的计算方式.

设$g(u), f(x)$ 是光滑或解析函数且复合$g(f(x))$ 有意义. 为了求$x \mapsto g(f(x))$ 的局部Taylor 公式或Taylor展开, 在没有其它好办法的时候可以用定义直接计算在定点x_0的各阶导数. 而有些时候可以应用Taylor 公式或Taylor展开的唯一性, 采取代入的方法: 先将$g(u)$ 对u 做Taylor 展开, 然后再将$u = f(x)$ 对x的Taylor展开代入, 最后按x的升幂求和即得到$g(f(x))$ 对x的Taylor 展开. 这种方法特别适合于内层函数$f(x)$是多项式的情形.

【例】研究函数的Taylor 展开.

(1) 设$g(u)$在$u = 0$展开的Tayor级数在$\mathbb{R}$上收敛于$g(u)$. 求$g(-x^2)$的Taylor展开：

$$g(-x^2) = \sum_{n=0}^{\infty} \frac{g^{(n)}(0)}{n!} u^n \Big|_{u=-x^2} = \sum_{n=0}^{\infty} \frac{(-1)^n g^{(n)}(0)(2n)!/n!}{(2n)!} x^{2n}, \quad x \in \mathbb{R}.$$

根据Tayor级数的唯一性知这就是$g(-x^2)$的Taylor展开. 此外还得到

$$(g(-x^2))^{(2n-1)}\Big|_{x=0} = 0, \quad (g(-x^2))^{(2n)}\Big|_{x=0} = (-1)^n \frac{g^{(n)}(0)}{n!}(2n)!$$

例如

$$e^{-x^2} = \sum_{n=0}^{\infty} \frac{(-x^2)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{n!} = \sum_{n=0}^{\infty} \frac{\frac{(-1)^n (2n)!}{n!}}{(2n)!} x^{2n}.$$

$$(e^{-x^2})^{(2n)}\Big|_{x=0} = \frac{(-1)^n (2n)!}{n!}, \quad (e^{-x^2})^{(2n+1)}\Big|_{x=0} = 0, \quad n = 0, 1, 2, \dots.$$

(2) 若函数$f(x)$ 是解析的, 则$f(\sqrt{x})$ 一般不是解析的. 但$\frac{1}{2}(f(\sqrt{x}) + f(-\sqrt{x}))$ 可以延拓成解析函数:

$$\frac{1}{2}(f(\sqrt{x}) + f(-\sqrt{x})) = \frac{1}{2} \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \Big((\sqrt{x})^n + (-\sqrt{x})^n\Big) = \sum_{n=0}^{\infty} \frac{f^{(2n)}(0)}{(2n)!} x^n.$$

由此可知当$f(x)$是偶函数时, $f(\sqrt{x})$可以有解析延拓. 例如

$$\cos(\sqrt{x}) = \sum_{n=0}^{\infty} (-1)^n \frac{(\sqrt{x})^{2n}}{(2n!)} = \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{(2n!)}.$$

因此可以把$\cos(\sqrt{x})$ 解析延拓到$\mathbb{R}$ 上, 即用上面的幂级数定义$\cos(\sqrt{x}), x \in \mathbb{R}$. 等等.

(3) $f(x) = \log(\cos x)$.

由$\log(1+u)=u-\frac{u^2}{2}+\frac{u^3}{3}+O(u^4)$ 和$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)$ 有

$$\begin{aligned}
&\log(\cos x)=\cos x-\frac{(\cos x)^2}{2}+\frac{(\cos x)^3}{3}+O((\cos x)^4)\\
&=\Big(-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)\Big)-\frac{1}{2}\Big(-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)\Big)^2\\
&+\frac{1}{3}\Big(-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)\Big)^3+O\left(\Big(-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)\Big)^4\right)\\
&=\cdots=-\frac{1}{2}x^2-\frac{1}{12}x^4-\frac{1}{45}x^6+O(x^8).
\end{aligned}$$

或者由$\log(1-u)=-u-\frac{u^2}{2}-\frac{u^3}{3}+O(u^4)$ 有

$$\begin{aligned}
&2\log(\cos x)=\log(\cos^2 x)=\log(1-\sin^2 x)\\
&=-\sin^2 x-\frac{1}{2}\sin^4 x-\frac{1}{3}\sin^6 x+O(\sin^8 x)
\end{aligned}$$

然后将$\sin x = x-\frac{1}{3!}x^3+\frac{1}{5!}x^5+O(x^7)$ 代入. 但需要计算$\sin x$的平方..., 也即这里把原来的两次复合升高为三次复合, 所以计算量较大. 如果直接展开$\cos x = 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+O(x^8)$, 则不涉及新的复合运算.

(4) 求极限

$$\lim_{x\to 0}\frac{\arctan x-\sin x}{\tan x-\arcsin x}.$$

一个方法是用罗比达法则(后面讲): $\lim\limits_{x\to 0}\frac{f(x)}{g(x)}=\lim\limits_{x\to 0}\frac{f'(x)}{g'(x)}$. 但题目放在这儿的意思是对分子分母直接做Taypor 展开(利用已知的展开式). 具体计算留作练习. □

B. 从一阶导数的Taylor展开确定原函数的Taylor展开.

有些时候函数$f(x)$的导数$f'(x)$比$f(x)$ 结构简单. 例如$f(x)=\log(1+x),\arctan x$ 就是典型的这种情形. 这时可以考虑先求$f'(x)$的Taylor展开然后通过两个级数的系数关系推出$f(x)$的Taylor展开. 当然我们也常用$f(x)$的展开确定$f'(x)$的展开.

让我们以在$x=0$展开的Taylor级数为例做具体说明. 假定$f(x)$ 在$x=0$ 附近解析. 设已有

$$f'(x)=\sum_{n=0}^{\infty}a_n x^n,\quad x\in(-R,R).$$

则

$$f(x)=f(0)+\sum_{n=0}^{\infty}\frac{a_n}{n+1}x^{n+1},\quad x\in(-R,R).$$

事实上由Taylor 级数的唯一性知

$$a_n = \frac{(f')^{(n)}(0)}{n!} = \frac{f^{(n+1)}(0)}{n!}.$$

于是有

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = f(0) + \sum_{n=0}^{\infty} \frac{f^{(n+1)}(0)}{(n+1)!} x^{n+1} = f(0) + \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

【实际上这属于简单的逐项微分、逐项积分定理. 只是现在还没讲这些定理.】

【例】试将$\arctan x$展开成(在$x=0$展开的)Taylor 级数. 并证明

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}.$$

【解】当$|x|<1$ 时

$$(\arctan x)' = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-x^2)^n = \sum_{n=0}^{\infty} (-1)^n x^{2n} = \sum_{n=0}^{\infty} a_n x^n.$$

其中$a_{2n} = (-1)^n,\ a_{2n+1} \equiv 0.\ \Longrightarrow$ (注意$\arctan 0 = 0$)

$$\arctan x = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1} = \sum_{n=0}^{\infty} a_{2n} \frac{x^{2n+1}}{2n+1} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}, \quad x \in (-1,1).$$

在区间的右端点$x=\pm1$处, 因由交错级数收敛性判别法知$\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}$收敛, 故根据**Abel定理** 和$\arctan x$的连续性得到

$$\frac{\pi}{4} = \arctan 1 = \lim_{x\to 1-} \arctan x = \lim_{x\to 1-} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}.$$

同理(或由奇函数性质)有

$$-\frac{\pi}{4} = \arctan(-1) = \lim_{x\to -1+} \arctan x = -\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}.$$

当$|x|>1$ 时级数的通项不趋于零因而级数发散. 结论是: $\arctan x$在$x=0$展开的Taylor 级数的收敛区间等于$[-1,1]$ 且

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}, \quad x \in [-1,1]. \qquad \square$$

§5.5. 用微分学的方法研究函数.

我们将学习用微分的方法研究函数的**局部极值、整体最(大,小)值、单调性、凹凸性、不等式、一些实际问题中的优化问题**等.

1. 用导数研究极值、最值、不等式. 函数的局部极值、整体最值、单调性的基本概念已在前面讲过且不难理解。下面从微分的角度再讲些进一步的性质.

首先回忆Fermat 极值原理: 设$f: I \to \mathbb{R}$ 在内点$x_0 \in I^\circ$ 附近可微. 则当x_0 是f的一个局部极值点时, x_0 必是f_0的一个临界点, 即必有$f'(x_0) = 0$. 因此若$f'(x_0) \neq 0$, 则x_0 不是f的局部极值点!

【定理5.28(取得极值和严格极值的充分条件1)】设$I \subset \mathbb{R}$ 为区间, $f: I \to \mathbb{R}$ 在$(x_0-\delta, x_0+\delta) \subset I$ 内可微. 则有:

(1) 若$f'(x) \geq 0$ for all $x \in (x_0-\delta, x_0)$; $f'(x) \leq 0$ for all $x \in (x_0, x_0+\delta)$. 则x_0 是f的一个局部极大值点, 即$f(x_0)$ 是f的一个局部极大值.

(2) 若$f'(x) \leq 0$ for all $x \in (x_0-\delta, x_0)$; $f'(x) \geq 0$ for all $x \in (x_0, x_0+\delta)$. 则x_0 是f的一个局部极小值点, 即$f(x_0)$ 是f的一个局部极小值.

(3) 若(1),(2) 中的不等号"$\geq, \leq$"可以换成严格不等号"$>, <$", 则x_0分别是f的局部严格极大、严格小值点.

【证】(1) 由假设知f 在$(x_0-\delta, x_0]$ 上单调不减, 在$[x_0, x_0+\delta)$ 上单调不增. 因此

$f(x) \leq f(x_0)$ for all $x \in (x_0-\delta, x_0]$;

$f(x) \leq f(x_0)$ for all $x \in [x_0, x_0+\delta)$.

所以$f(x) \leq f(x_0)$ for all $x \in (x_0-\delta, x_0+\delta)$.

所以x_0 是f的一个局部极大值点. 同理可证(2),(3). $\square$

【定理5.29(取得严格极值的充分条件2)】设$I \subset \mathbb{R}$ 为区间, 函数$f: I \to \mathbb{R}$ 在内点$x_0 \in I^\circ$附近可微且在x_0 二次可微. 则有:

(1) 若$f'(x_0) = 0$, $f''(x_0) < 0$, 则x_0是f的一个局部严格极大值点.

(2) 若$f'(x_0) = 0$, $f''(x_0) > 0$, 则x_0是f的一个局部严格极小值点.

这一定理的一般形式如下:

定理5.30(取得严格极值的充分条件3)】 设$I \subset \mathbb{R}$ 为区间, 函数$f : I \to \mathbb{R}$ 在内点$x_0 \in I^\circ$附近有$n-1$ 阶导数而在x_0处有n 阶导数, 其中$n \geq 2$. 假设

$$f'(x_0) = f''(x_0) = \cdots f^{(n-1)}(x_0) = 0, \quad f^{(n)}(x_0) \neq 0.$$

则

(1) 若n 为**奇数**, 则x_0不是f的局部极值点.

(2) 若n 为**偶数**, 则x_0是f的一个局部极值点, 并且当$f^{(n)}(x_0) < 0$ 时, x_0是f 的一个局部极大值点; 当$f^{(n)}(x_0) > 0$ 时, x_0是f 的一个局部极小值点.

【两个定理的证明】 只需证第二个定理. 由假设条件和Taylor 公式有

$$f(x) - f(x_0) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n).$$

因$f^{(n)}(x_0) \neq 0$, 故由高阶无穷小的定义知

$$\alpha(x) := \frac{o((x - x_0)^n)}{\frac{f^{(n)}(x_0)}{n!}(x - x_0)^n} \to 0 \quad (x \to x_0).$$

因此存在$\delta > 0$ 使得$(x_0 - \delta, x_0 + \delta) \subset I$ 且$|\alpha(x)| < 1/2$ for all $x \in (x_0 - \delta, x_0 + \delta)$. 于是有

$$f(x) - f(x_0) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n(1 + \alpha(x)), \quad x \in (x_0 - \delta, x_0 + \delta).$$

由此易见$f(x) - f(x_0) \neq 0$ for all $x \in (x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$.

并且当n 为奇数时, $f(x) - f(x_0)$ 在$(x_0 - \delta, x_0)$ 和$(x_0, x_0 + \delta)$ 上有相反的符号. 因此x_0不是f的极值点.

而当n为偶数时, $f(x) - f(x_0)$ 在$(x_0 - \delta, x_0)$ 和$(x_0, x_0 + \delta)$ 上有相同的符号. 因此x_0是f的一个严格极值点. 进一步, 若$f^{(n)}(x_0) > 0$ 则x_0 是f的一个严格极小值点; 若$f^{(n)}(x_0) < 0$ 则x_0 是f 的一个严格极大值点. $\square$

【定理5.31(取得最值的充分条件1)】 设$I \subset \mathbb{R}$ 为区间, 函数$f : I \to \mathbb{R}$ 在I 上连续, 在I° 内可微. 设$x_0 \in I$ (包括x_0是区间端点的情形[7]).

(1) 若$f'(x) \geq 0$ for all $x \in I^\circ \cap (-\infty, x_0)$; $f'(x) \leq 0$ for all $x \in I^\circ \cap (x_0, +\infty)$. 则$f(x_0)$是$f$在$I$ 上的最大值, 即$f(x_0) = \max_{x \in I} f(x)$.

[7] **当$I = [x_0, b), [x_0, b]$ 时或$I = (a, x_0], [a, x_0]$ 时定理被理解为只涉及x_0的某一侧, 而另一侧不出现.**

(2) 若$f'(x) \le 0$ for all $x \in I^\circ \cap (-\infty, x_0)$; $f'(x) \ge 0$ for all $x \in I^\circ \cap (x_0, +\infty)$. 则$f(x_0)$是$f$在$I$ 上的最小值, 即$f(x_0) = \min\limits_{x \in I} f(x)$.

(3) 若(1),(2) 中的不等号“$\ge, \le$”可以换成严格不等号“$>, <$”, 则$f(x_0)$分别是f在I上的严格最大值、严格最小值.

【证】(1) 有假设知f 在$I \cap (-\infty, x_0]$ 上单调不减, 在$I \cap [x_0, +\infty)$ 上单调不增. 所以不等式$f(x) \le f(x_0)$对所有$x \in I$ 成立. 因此$f(x_0) = \max\limits_{x \in I} f(x)$.

同理可证(2),(3). $\square$

【定理5.32(取得最值的充分条件2)】设$I \subset \mathbb{R}$ 为区间, 函数$f : I \to \mathbb{R}$ 在I 上连续, 在I° 内两次可微. 设$x_0 \in I^\circ$ 且$f'(x_0) = 0$.

(1) 若$f''(x) \le 0$ for all $x \in I^\circ$ 则$f(x_0) = \max\limits_{x \in I} f(x)$.

(2) 若$f''(x) \ge 0$ for all $x \in I^\circ$ 则$f(x_0) = \min\limits_{x \in I} f(x)$.

(3) 若(1),(2) 中的不等号“$\le, \ge$”可以换成严格不等号“$<, >$”, 则$f(x_0)$分别是f在I上的严格最大值、严格最小值.

【证】(1) 由假设知$f'(x)$ 在I° 上单调不增. 因此

$f'(x) \ge f'(x_0) = 0$ for all $x \in I^\circ \cap (-\infty, x_0)$,

$f'(x) \le f'(x_0) = 0$ for all $x \in I^\circ \cap (x_0, +\infty)$.

故由上面定理知$f(x_0) = \max\limits_{x \in I} f(x)$. 同理可证(2),(3). $\square$

【注】以上定理提供了证明不等式$f(x) \le g(x)$ 或$f(x) \ge g(x)$ 的方法：只要将以上定理应用于函数$f - g$.

【例1】证明贝努力不等式:

$$
\begin{aligned}
&(1+x)^\alpha \le 1 + \alpha x \qquad \text{if} \quad 0 < \alpha \le 1, \quad x > -1; \\
&(1+x)^\alpha \ge 1 + \alpha x \qquad \text{if} \quad \alpha \ge 1, \quad x > -1.
\end{aligned}
$$

【证】先设$0 < \alpha \le 1$. 考虑减法:

$$f(x) = (1+x)^\alpha - 1 - \alpha x, \quad x > -1.$$

则$f(0)=0$,

$$f'(x)=\alpha(1+x)^{\alpha-1}-\alpha=\alpha((1+x)^{\alpha-1}-1)\begin{cases}>0 & \text{if}\quad -1<x<0,\\ <0 & \text{if}\quad x>0.\end{cases}$$

因此f 在$(-1,0]$ 上单调增加, 在$[0,+\infty)$ 上单调减少. 故$f(x)\le f(0)$ for all $x\in(-1,0]$, $f(x)\le f(0)$ for all $x\in[0,+\infty)$. 所以$f(x)\le 0$ for all $x\in(-1,+\infty)$, 即$(1+x)^{\alpha}\le 1+\alpha x$ for all $x>-1$.

同法可证$\alpha\ge 1$的情形. $\square$

两个应用题:

【例2】设一圆柱形蒸锅的表面积被限定. 问怎样设计才可使蒸锅的容量最大? 此时蒸锅是什么样子?

【解】设S,V分别为圆柱形蒸锅的表面积(有底无顶)和体积, 设r,h 分别为此圆柱的半径和高. 则

$$\begin{aligned}&S=2\pi rh+\pi r^2,\quad V=\pi r^2h,\\ &S=2\frac{V}{r}+\pi r^2,\quad Sr=2V+\pi r^3,\\ &V=V(r)=\frac{1}{2}(Sr-\pi r^3)=\frac{1}{2}r(S-\pi r^2),\quad r\in(0,\sqrt{S/\pi}).\end{aligned}$$

计算导数:

$$\begin{aligned}&V'(r)=\frac{1}{2}(S-3\pi r^2)=\frac{3\pi}{2}\Big(\frac{S}{3\pi}-r^2\Big),\\ &0<r<\sqrt{\frac{S}{3\pi}}\implies V'(r)>0;\quad \sqrt{\frac{S}{3\pi}}<r<\sqrt{\frac{S}{\pi}}\implies V'(r)<0.\end{aligned}$$

由此知对于$r_0=\sqrt{\frac{S}{3\pi}}$ 有

$$\max_{0<r\sqrt{S/\pi}}V(r)\ (\text{存在且})=V(r_0)=\frac{1}{2}\sqrt{\frac{S}{3\pi}}(S-\pi\frac{S}{3\pi})=\frac{S}{3}\sqrt{\frac{S}{3\pi}}.$$

或者由

$$V'(r_0)=0,\quad V''(r)=-\frac{3\pi}{2}\cdot 2r<0\ \text{ for all }\ r\in(0,\sqrt{S/\pi})$$

也得知$\max\limits_{0<r\sqrt{S/\pi}}V(r)=V(r_0)$.

此时

$$h = \frac{V}{\pi r^2} = \frac{S}{3}\sqrt{\frac{S}{3\pi}} \cdot \frac{1}{\pi \frac{S}{3\pi}} = \sqrt{\frac{S}{3\pi}} = r_0.$$

反之假设$h = r$, 则

$$S = 2\pi r^2 + \pi r^2 = 3\pi r^2 \quad \Longrightarrow \quad r = \sqrt{\frac{S}{3\pi}}.$$

结论: 当且仅当蒸锅底的直径等于蒸锅高的二倍时, 蒸锅的容量最大. 此时从侧面看, 此蒸锅是一个长方形, 其底等于高的二倍. $\square$

【例3】试根据Fermart原理—— 光线行走的实际路程是耗时最少的路程—— 证明几何光学中**光的折射定律:**

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$$

其中v_1, v_2 是光在介质1 和介质2 中的传播速度, $\theta_1, \theta_2 \in (0, \pi/2)$ 是光线与介质公共界面折射点处的法线的夹角 (图示).

【证】设光从介质1 中的点A_1 传播到介质2 中的点A_2. 根据上述Fermart原理我们可以假定光的这个传播路线位于一张平面内(此处略去证明). 于是我们可以以两界面的交线为x-轴、以平行于界面的法线的直线为y-轴建立坐标系. 设A_1, A_2 到x-轴的距离分别为h_1, h_2. 不妨设A_1 到x-轴的垂足位于A_2 到x-轴的垂足的左侧并取该点为坐标原点$(0,0)$. 于是A_2 到x-轴的垂足为$(a,0)$ 其中$a > 0$. 设$(x,0)$ 为光的折射点. 则光从A_1 传到A_2时理论上所有可能的时间为

$$t(x) = \frac{\text{路程1}}{v_1} + \frac{\text{路程2}}{v_2} = \frac{\sqrt{x^2 + h_1^2}}{v_1} + \frac{\sqrt{(x-a)^2 + h_2^2}}{v_2}, \quad x \in \mathbb{R}.$$

但由**Fermart原理(光线行走的实际路程是耗时最少的路程)** 可知光从A_1 传到A_2所用的实际时间为

$$\min_{x\in\mathbb{R}} t(x).$$

利用光的这个极值性质我们来导出上述折射定律. 计算导数:

$$\begin{aligned}
&t'(x) = \frac{1}{v_2} \cdot \frac{x}{\sqrt{x^2 + h_1^2}} + \frac{1}{v_2} \cdot \frac{x-a}{\sqrt{(x-a)^2 + h_2^2}}, \\
&t''(x) = \frac{1}{v_2} \cdot \frac{h_1^2}{(x^2 + h_1^2)^{3/2}} + \frac{1}{v_2} \cdot \frac{h_2^2}{((x-a)^2 + h_2^2)^{3/2}} > 0 \quad \forall x \in \mathbb{R}, \\
&t'(0) < 0 < t'(a).
\end{aligned}$$

因此存在唯一的$x_0\in(0,a)$ 使得$t'(x_0)=0$. 根据**定理5.32** 知光所用的实际时间为

$$\min_{x\in\mathbb{R}} t(x)=t(x_0).$$

而由$t'(x_0)=0$ 和$0<x_0<a$ 得到

$$\frac{1}{v_2}\cdot\frac{x_0}{\sqrt{x_0^2+h_1^2}}=\frac{1}{v_2}\cdot\frac{a-x_0}{\sqrt{x_0-a)^2+h_2^2}}>0.$$

这正是

$$\frac{\sin\theta_1}{v_1}=\frac{\sin\theta_2}{v_2}.$$

反之若$\theta_1,\theta_2\in(0,\pi/2)$ 满足这一关系式, 则由θ_1,θ_2 为锐角可以确定一点$x_0\in(0,a)$ 使得

$$\frac{x_0}{\sqrt{x_0^2+h_1^2}}=\sin\theta_1,\quad \frac{a-x_0}{\sqrt{x_0-a)^2+h_2^2}}=\sin\theta_2.$$

将其代入$t'(x)$的表达式中得知$t'(x_0)=0$. 于是由$t(x)$的最小值点的唯一性即知$t(x_0)=\min_{x\in\mathbb{R}} t(x)$. 这说明, 在光的折射定律中, 折射角$\theta_1,\theta_2$ 与光的实际传播时间是相互确定的. $\square$

2 凸函数.

凸函数有很好的几何性质和分析性质. 它们广泛联系着几何问题, 优化问题和各种各样的不等式. 首先引进**凸集**和**凸组合**的概念.

凸集: 设集合$E\subset\mathbb{R}$ 或$E\subset\mathbb{R}^m$. 若蕴含关系$x,y\in E, 0<t<1 \implies (1-t)x+ty\in E$ 恒成立, 则称E 是一个凸集.

根据$\mathbb{R}$中区间的定义和实数的全序性, 易见I 是一个区间当且仅当I 是一个凸集.

凸组合: 设$x_1,x_2,...,x_n$ 是$\mathbb{R}$ 中或$\mathbb{R}$上的任一向量空间中的元素. 若实数$t_1,t_2,...,t_n$ 满足

$$t_k\ge 0,\ \ k=0,1,...,n;\quad t_1+t_2+\cdots+t_n=1$$

则称

$$t_1x_1+t_2x_2+\cdots+t_nx_n$$

为$x_1,x_2,...,x_n$的一个凸组合. 这个凸组合实际上就是$x_1,x_2,...,x_n$的一种等权或非等权的算术平均.

下面我们着重介绍一元凸函数.

【**凸(凹)函数的定义**】设$I\subset\mathbb{R}$ 为区间, $f:I\to\mathbb{R}$.

(1) 若

$$f((1-t)x+ty)\le(1-t)f(x)+tf(y)\quad \forall\, x,y\in I,\quad \forall\, 0<t<1$$

则称f 是I 上的一个凸函数(convex function).　进一步, 若

$$f((1-t)x+ty)<(1-t)f(x)+tf(y)\quad \forall\, x,y\in I\quad \text{with}\ \ x\ne y,\quad \forall\, 0<t<1$$

则称f 是I 上的一个严格凸函数.

(2) 若

$$f((1-t)x+ty)\ge(1-t)f(x)+tf(y)\quad \forall\, x,y\in I,\quad \forall\, 0<t<1$$

则称f 是I 上的一个凹函数(concave function).　进一步, 若

$$f((1-t)x+ty)>(1-t)f(x)+tf(y)\quad \forall\, x,y\in I\quad \text{with}\ \ x\ne y,\quad \forall\, 0<t<1$$

则称f 是I 上的一个严格凹函数.　□

显然 f **是(严格) 凸函数** $\Longleftrightarrow$ $-f$ **是(严格) 凹函数.**

因此不失一般性只需研究凸函数. 以下总设I 为$\mathbb{R}$ 中的区间.

【**凸函数的几何特征**】

• 曲线上任意两点间的直线段总是位于对应的曲线段的上方 (图示).

• 曲线上任意三点构成的三角形的三条边的斜率关系(图示):

左边斜率$\le$ 上边斜率$\le$ 右边斜率

即

$$x_1<x_2<x_3\implies \frac{f(x_1)-f(x_2)}{x_1-x_2}\le\frac{f(x_1)-f(x_3)}{x_1-x_3}\le\frac{f(x_2)-f(x_3)}{x_2-x_3}.\qquad(*)$$

【证】经简单运算知: 位于x_1,x_3 之间的x_2可以写成x_1,x_3的如下凸组合:

$$x_2=(1-\frac{x_2-x_1}{x_3-x_1})x_1+\frac{x_2-x_1}{x_3-x_1}x_3.$$

于是由凸函数的定义有

$$f(x_2)\le(1-\frac{x_2-x_1}{x_3-x_1})f(x_1)+\frac{x_2-x_1}{x_3-x_1}f(x_3)=f(x_1)+\frac{x_2-x_1}{x_3-x_1}(f(x_3)-f(x_1)).$$

它也可写成

$$f(x_2) \le f(x_3) + \frac{x_3 - x_2}{x_3 - x_1}(f(x_1) - f(x_3)).$$

由此可分别得到$(*)$ 中的第一个和第二个不等式. $\square$

【**Jensen 不等式**】设$f : I \to \mathbb{R}$ 为凸函数. 则对任意$x_1, x_2, ..., x_n \in I$ 有

$$f(x_1 + x_2 + \cdots + x_n) \le \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n}.$$

一般地, 若$\lambda_k \ge 0, k = 1, 2, ..., n$ 且$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$, 则

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \le \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n)$$

即

$$f\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) \le \sum_{k=1}^{n} \lambda_k f(x_k).$$

换言之, 凸组合的函数小于等于函数的凸组合.

此外当f为严格凸函数时, 若所有$\lambda_k > 0, k = 1, 2, ..., n; \sum_{k=1}^{n} \lambda_k = 1$, 则

$$f\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) = \sum_{k=1}^{n} \lambda_k f(x_k) \iff x_1 = x_2 = \cdots = x_n.$$

【证】对x_k的个数n 用归纳法. 当$n = 2$时, 这是凸函数的定义. 设在个数$= n - 1 (\ge 1)$ 时Jensen 不等式成立. 看个数$= n$ 时. 设$x_1, x_2, ..., x_n \in I$, $\lambda_k \ge 0, k = 1, 2, ..., n$ 且$\lambda_1 + \lambda_2 + \cdots + \lambda_n = 1$. 若$\lambda_n = 0$ 则由归纳假设知Jensen 不等式成立; 若$\lambda_n = 1$, 则$\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = 0$, Jensen 不等式显然成立. 设$0 < \lambda_n < 1$. 此时有

$$\sum_{k=1}^{n} \lambda_k x_k = (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} x_k + \lambda_n x_n,$$

$$\frac{\lambda_k}{1 - \lambda_n} \ge 0, \quad k = 1, 2, ..., n - 1; \quad \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} = 1 \quad \Longrightarrow \quad \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} x_k \in I$$

$$\Longrightarrow \quad f\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) = f\Big((1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} x_k + \lambda_n x_n\Big)$$

$$\le (1 - \lambda_n) f\Big(\sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} x_k\Big) + \lambda_n f(x_n)$$

$$\le (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} f(x_k) + \lambda_n f(x_n) = \sum_{k=1}^{n} \lambda_k f(x_k).$$

所以当x_k的个数$= n$ 时Jensen 不等式仍成立.

最后设f为严格凸函数. 设$n \ge 2, \lambda_k > 0, k = 1, 2, ..., n; \sum_{k=1}^{n} \lambda_k = 1$ 并设$x_1, x_2, ..., x_n \in I$ 使等式$f\Big(\sum_{k=1}^{n} \lambda_k x_k\Big) = \sum_{k=1}^{n} \lambda_k f(x_k)$ 成立. 因有限多个实数可以按大小排列, 故调整记号后可设$x_1 \ge x_2 \ge \cdots \ge x_n$. 由上面的推导和$f$的严格凸性知

$$\sum_{k=1}^{n-1} \frac{\lambda_k}{1-\lambda_n} x_k = x_n \quad 即 \quad \sum_{k=1}^{n-1} \frac{\lambda_k}{1-\lambda_n}(x_k - x_n) = 0.$$

因$\frac{\lambda_k}{1-\lambda_n} > 0, k = 1, 2, ..., n-1$, 故得到$x_k - x_n = 0, k = 1, 2, ..., n-1$. 所以$x_k$ 全相等. □

【**命题5.33**】设$f : I \to \mathbb{R}$ 为凸函数. 则f 在每一点$x \in I^\circ$ 都有左右导数, 即$f'_-(x), f'_+(x)$ 皆存在有限. 此外f'_-, f'_+ 在I° 内单调不减并有

$$f'_-(x) \le f'_+(x) \le f'_-(y) \le f'_+(y) \qquad \forall\, x, y \in I^\circ \quad \text{s.t.} \quad x < y.$$

特别若f 在I°内可微, 则导函数f' 在I°内单调不减.

【证】对于I中的任意七点$x_1 < x_2 < x_3 < x_4 < x_5 < x_6 < x_7$ 应用三边斜率不等式$(*)$ 得到(图示)

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \le \frac{f(x_3)-f(x_4)}{x_3-x_4} \le \frac{f(x_3)-f(x_5)}{x_3-x_5} \le \frac{f(x_4)-f(x_5)}{x_4-x_5} \le \frac{f(x_6)-f(x_7)}{x_6-x_7}. \tag{$*1$}$$

从第一个不等号、第二个不等号和最后一个不等号我们看到函数$x \mapsto \frac{f(x)-f(x_3)}{x-x_3}$ 在$x \in (x_3, x_6)$ 上单调不减且有界. 因此

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \le f'_+(x_3) := \lim_{x \to x_3+} \frac{f(x)-f(x_3)}{x-x_3} \le \frac{f(x_6)-f(x_7)}{x_6-x_7}. \tag{$*2$}$$

再从$(*1)$中的第一个不等号、第三个不等号和最后一个不等号我们看到函数$x \mapsto \frac{f(x)-f(x_5)}{x-x_5}$ 在$x \in (x_2, x_5)$ 上单调不减且有界. 因此

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \le f'_-(x_5) := \lim_{x \to x_5-} \frac{f(x)-f(x_5)}{x-x_5} \le \frac{f(x_6)-f(x_7)}{x_6-x_7}. \tag{$*3$}$$

因x_3, x_5可分别取为I°中的任意一点, 故这证明了极限$f'_-(x), f'_+(x)$ 对所有$x \in I^\circ$ 存在有限.

在$(*1)$中让$x_3 \to x_4-,\ x_5 \to x_4+$ 得到$f'_-(x_4) \le f'_+(x_4)$. 因x_4可以取为I°中任意点, 这证明了

$$f'_-(x) \le f'_+(x), \quad x \in I^\circ.$$

在$(*2)$中让$x_2\to x_1+$ 得到$f'_+(x_1)\le f'_+(x_3)$. 因$x_1<x_3$可以取为I°中任意两点, 这证明了$x\mapsto f'_+(x)$ 在I° 上单调不减.

在$(*3)$中让$x_1\to x_2-$ 得到$f'_-(x_2)\le f'_-(x_5)$. 因$x_2<x_5$可以取为I°中任意两点, 这证明了$x\mapsto f'_-(x)$ 在I° 上单调不减.

最后在$(*2)$中让$x_6\to x_7-$ 得到$f'_+(x_3)\le f'_-(x_7)$. 因$x_3<x_7$可以取为I°中任意两点, 故得到

$$f'_+(x)\le f'_-(y),\quad x,y\in I^\circ,\ \ x<y.$$

□

【命题5.34】设$f:I\to\mathbb{R}$ 为凸函数. 则

$$f(x)\ge f(x_0)+f'_-(x_0)(x-x_0)\qquad\forall\,x_0\in I^\circ,\quad\forall\,x\in I;$$

$$f(x)\ge f(x_0)+f'_+(x_0)(x-x_0)\qquad\forall\,x_0\in I^\circ,\quad\forall\,x\in I.$$

特别若f在I°内可微, 则

$$f(x)\ge f(x_0)+f'(x_0)(x-x_0)\qquad\forall\,x_0\in I^\circ,\quad\forall\,x\in I.$$

这个不等式的几何意义是: 曲线$y=f(x)$ 在其上任意点处的切线的上方(图示).

【证】任取$x_0\in I^\circ$. 从上一命题的证明可知$x\mapsto\frac{f(x)-f(x_0)}{x-x_0}$ 分别在$I\cap(-\infty,x_0)$和$I\cap(x_0,+\infty)$ 都是单调不减的. 因此

$$\frac{f(x)-f(x_0)}{x-x_0}\le\lim_{y\to x_0-}\frac{f(y)-f(x_0)}{y-x_0}=f'_-(x_0)\le f'_+(x_0)\qquad\forall\,x\in I\quad\text{with}\quad x<x_0,$$

$$\frac{f(x)-f(x_0)}{x-x_0}\ge\lim_{y\to x_0+}\frac{f(y)-f(x_0)}{y-x_0}=f'_+(x_0)\ge f'_-(x_0)\qquad\forall\,x\in I\quad\text{with}\quad x>x_0.$$

由这两个不等式得到

$$f(x)\ge f(x_0)+f'_-(x_0)(x-x_0)\ge f(x_0)+f'_+(x_0)(x-x_0)\qquad\forall\,x\in I\quad\text{with}\quad x\le x_0,$$

$$f(x)\ge f(x_0)+f'_+(x_0)(x-x_0)\ge f(x_0)+f'_-(x_0)(x-x_0)\qquad\forall\,x\in I\quad\text{with}\quad x\ge x_0.$$

所以命题成立. □

下面这个命题在以后学习和研究积分不等式时是比较常用的:

【命题5.35】设$f:I\to\mathbb{R}$. 则

f 在I 上是凸函数$\iff$ 对任意$x_0 \in I^\circ$ 存在常数$A_{x_0} \in \mathbb{R}$ 使得

$$f(x) \ge f(x_0) + A_{x_0}(x - x_0) \qquad \forall\, x \in I.$$

【证】“$\Longrightarrow$”: 设f 在I 上是凸函数. 则由上一命题知$A_{x_0} = f'_-(x_0)$ 或$A_{x_0} = f'_+(x_0)$ 满足所证不等式.

“$\Longleftarrow$”: 设f 满足上述不等式.

对任意$x, y \in I, 0 < t < 1$. 不妨设$x < y$. 令$z = (1-t)x + ty$. 则$x < z < y$ 从而$z \in I^\circ$. 由假设知存在$A_z \in \mathbb{R}$ 使得

$$f(x) \ge f(z) + A_z(x - z), \quad f(y) \ge f(z) + A_z(y - z).$$

对这两个不等式分别乘以$1 - t$ 和t 有

$$\begin{aligned}(1-t)f(x) &\ge (1-t)f(z) + A_z\Big((1-t)x - (1-t)z\Big),\\ tf(y) &\ge tf(z) + A_z\Big(ty - tz\Big).\end{aligned}$$

两边对应相加并注意$z = (1-t)x + ty$ 得到

$$(1-t)f(x) + tf(y) \ge f(z) + A_z\Big((1-t)x + ty - z\Big) = f(z) = f((1-t)x + ty).$$

所以f 在I 上是凸函数. $\Box$

下面我们进一步用导数刻画函数的凸性.

【命题5.36】设$f: I \to \mathbb{R}$ 在I上连续, 在I°内可导. 则以下$1^\circ, 2^\circ, 3^\circ$ 彼此等价:

1° f 在I上为凸函数.

2° 导函数f'在I°内单调不减.

3° $f(x) \ge f(x_0) + f'(x_0)(x - x_0)$ for all $x \in I$ and all $x_0 \in I^\circ$.

第3°条说明: 凸函数曲线位于曲线任一点处的切线的上方.

【证】

$1^\circ \Longrightarrow 2^\circ$: 设$f$ 在I上为凸函数. 由**命题5.33** 知f' 在I°内单调不减.

$2^\circ \Longrightarrow 3^\circ$: 设$f'$在$I^\circ$内单调不减. 对任意$x \in I, x_0 \in I^\circ$. 不妨设$x \ne x_0$. 由中值定理, 存在$\xi \in (x, x_0)$ 或$\xi \in (x_0, x)$ 使得$f(x) - f(x_0) = f'(\xi)(x - x_0)$. 由$f'$在$I^\circ$内单调不减知

若$x < x_0$ 则$f'(\xi) \le f'(x_0)$ 于是由$x - x_0 < 0$ 有

$$f(x) - f(x_0) = f'(\xi)(x - x_0) \ge f'(x_0)(x - x_0),$$

若$x > x_0$ 则$f'(\xi) \ge f'(x_0)$ 从而仍有

$$f(x) - f(x_0) = f'(\xi)(x - x_0) \ge f'(x_0)(x - x_0).$$

$3^\circ \Longrightarrow 1^\circ$: 此时由**命题5.35** 知$f$ 在I 是凸函数. □

下面我们用导数刻画函数的严格凸性.

【命题5.37】设$f : I \to \mathbb{R}$ 在I上连续, 在I°内可导. 则以下$1^\circ, 2^\circ, 3^\circ$ 彼此等价:

1° f 在I上为严格凸函数.

2° 导函数f'在I°内严格单调增加.

3° $f(x) > f(x_0) + f'(x_0)(x - x_0)$ for all $x \in I$ and all $x_0 \in I^\circ$ with $x \neq x_0$.

【证】

$1^\circ \Longrightarrow 2^\circ$: 设$f$ 在I上为严格凸函数. 由**命题5.33** 知f' 在I°内单调不减. 若存在$x, y \in I^\circ$, $x < y$ 使得$f'(x) = f'(y)$, 则$f'(t) = f'(x)$ for all $t \in [x, y]$. 这蕴含

$$f(t) = f(x) + f'(x)(t - x) \quad \forall\, t \in [x, y]$$

即f 在$[x, y]$ 上为仿线性函数, 即f 可表为$f(t) = a + bt$, $t \in [x, y]$ 其中a, b 为常数. 于是有

$$f(\frac{x + y}{2}) = \frac{1}{2}f(x) + \frac{1}{2}f(y).$$

这与f 为严格凸函数矛盾. 所以f'必在I°内严格单调增加.

$2^\circ \Longrightarrow 3^\circ$: 设$f'$在$I^\circ$内严格单调增加. 对任意$x \in I, x_0 \in I^\circ$ with $x \neq x_0$. 由中值定理, 存在$\xi \in (x, x_0)$ 或$\xi \in (x_0, x)$ 使得$f(x) - f(x_0) = f'(\xi)(x - x_0)$. 由$f'$在$I^\circ$内严格单调增加知若$x < x_0$ 则$f'(\xi) < f'(x_0)$ 于是由$x - x_0 < 0$ 有

$$f(x) - f(x_0) = f'(\xi)(x - x_0) > f'(x_0)(x - x_0),$$

若$x > x_0$ 则$f'(\xi) > f'(x_0)$ 从而仍有

$$f(x) - f(x_0) = f'(\xi)(x - x_0) > f'(x_0)(x - x_0).$$

$3^\circ \Longrightarrow 1^\circ$: 这个证明与**命题5.35**的“$\Longleftarrow$”的证明相同: 只需在那里取$A_z = f'(z)$ 并注意那里的不等号现在可以取为严格不等号. 因此f 在I上为严格凸函数. $\Box$

根据单调性与导数的关系, 立即得到下列常用命题:

【**命题5.38**】设$f : I \to \mathbb{R}$ 在I上连续, 在I°内两次可导. 则

f 在I上为凸函数 $\Longleftrightarrow$ $f''(x) \ge 0$ for all $x \in I^\circ$. $\Box$

• **凸函数的最大值与最小值**.

【**命题5.39**】设$f : [a, b] \to \mathbb{R}$为凸函数.

$$f(x) \le \max\{f(a), f(b)\} \qquad \forall\, x \in [a, b].$$

换言之凸函数f 在$[a, b]$上必有最大值且最大值一定在端点a, b 处达到, 即

$$\max_{a \le x \le b} f(x) = \max\{f(a), f(b)\}.$$

【证明】对任意$x \in [a, b]$, 只需考虑$a < x < b$的情形. 此时有

$$x = \frac{b-x}{b-a}a + \frac{x-a}{b-a}b.$$

因$\frac{b-x}{b-a} > 0, \frac{x-a}{b-a} > 0$ 且$\frac{b-x}{b-a} + \frac{x-a}{b-a} = 1$, 故由凸函数的定义便有

$$f(x) \le \frac{b-x}{b-a}f(a) + \frac{x-a}{b-a}f(b) \le \max\{f(a), f(b)\}.$$

$\Box$

【**命题5.40**】设$f : I \to \mathbb{R}$为凸函数. 若f在某点$x_0 \in I^\circ$可微且$f'(x_0) = 0$, 则x_0 便是f 在I 上的最小值点, 即

$$f(x) \ge f(x_0) \qquad \forall\, x \in I.$$

换言之, 凸函数的临界点必是整体最小值点.

【证】由**命题5.36** 和$f'(x_0) = 0$ 有

$$f(x) \ge f(x_0) + f'(x_0)(x - x_0) = f(x_0) \qquad \forall\, x \in I. \qquad \Box$$

• **一些例题和著名不等式**.

【**例1**】设$f:[a,+\infty)\to\mathbb{R}$ 为凸函数且$\lim\limits_{x\to+\infty}f(x)=A\in\mathbb{R}$. 则$f$ 在$[a,+\infty)$ 上单调不增. 从而有$f(a)\ge f(x)\ge A$ for all $x\in[a,+\infty)$.

【证】对任意$y\in(a,+\infty)$, 由**命题5.34** 知

$$f(x)\ge f(y)+f'_-(y)(x-y)\qquad \forall\, x\in[a,+\infty).$$

若$f'_-(y)>0$ 则 $\lim\limits_{x\to+\infty}f(x)=+\infty$, 与假设矛盾. 因此必有$f'_-(y)\le 0$. 于是有

$$f(x)\ge f(y)+f'_-(y)(x-y)\ge f(y)\qquad \forall\, a\le x<y.$$

这证明了f 在$[a,+\infty)$ 上单调不增. $\Box$

【**例2**】设$f:[a,b]\to\mathbb{R}$ 在$[a,b]$ 上连续、在(a,b)内两次可微, 且满足

$$f(a)=f(b)=0,\quad M:=\sup_{a<x<b}|f''(x)|<\infty.$$

则

$$|f(x)|\le\frac{M}{2}(x-a)(b-x)\qquad \forall\, x\in[a,b].$$

【证】本题如用中值定理或其它方法, 则将陷入较繁的运算和估计. 此处的证明是化成凸函数的最值问题: 凸函数的最大值必在区间端点达到. 为此令

$$F(x)=f(x)-\frac{M}{2}(x-a)(b-x),$$

$$G(x)=-f(x)-\frac{M}{2}(x-a)(b-x).$$

则有$F(a)=0,F(b)=0,G(a)=0,G(b)=0,$

$$F''(x)=f''(x)+M\ge 0,\quad G''(x)=-f''(x)+M\ge 0,\qquad \forall\, x\in(a,b).$$

由**命题5.38**知$F(x),G(x)$都是$[a,b]$上的凸函数. 因此

$$F(x)\le\max\{F(a),F(b)\}=0,\quad G(x)\le\max\{G(a),G(b)\}=0\qquad \forall\, x\in[a,b]$$

即

$$f(x)\le\frac{M}{2}(x-a)(b-x),\quad -f(x)\le\frac{M}{2}(x-a)(b-x)\quad \forall\, x\in[a,b].$$

所以$|f(x)|\le\frac{M}{2}(x-a)(b-x)$ for all $x\in[a,b]$. $\Box$

【**Young 不等式**】设$p>1,q>1$ 满足$1/p+1/q=1$. 则

$$a^{1/p}b^{1/q}\le \frac{a}{p}+\frac{b}{q}\qquad \text{for all }\ a\ge 0, b\ge 0.$$

等号成立当且仅当$a=b$.

一般地, 若$\alpha_k>0, k=1,2,...,n; \alpha_1+\alpha_2+\cdots\alpha_n=1$, 则对任意$x_k\ge 0\,(k=1,2,...,n)$ 有“几何平均不超过算术平均”:

$$x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}\le \alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n.$$

并且等号成立当且仅当$x_1=x_2=\cdots=x_n$.

【证】易见我们只需考虑所有$x_k>0\,(k=1,2,...,n)$的情形. 此时我们可以使用对数函数. 由$(\log x)''=-x^{-2}<0\ (x>0)$ 知$\log x$ 为严格凹函数. 因此

$$\log\Big(\sum_{k=1}^n\alpha_kx_k\Big)\ge\sum_{k=1}^n\alpha_k\log x_k=\log\Big(\prod_{k=1}^n x_k^{\alpha_k}\Big)$$

即

$$\sum_{k=1}^n\alpha_kx_k\ge\prod_{k=1}^n x_k^{\alpha_k}.$$

再由$\log x$ 为严格凹函数知: 等号成立当且仅当$x_1=x_2=\cdots=x_n$. $\Box$

【**Hölder 不等式**】设$p>1,q>1$ 满足$1/p+1/q=1$. 则对任意$x_k\ge 0, y_k\ge 0, k=1,2,...,n$ 有

$$\sum_{i=1}^n x_ky_k\le\Big(\sum_{k=1}^n x_k^p\Big)^{1/p}\Big(\sum_{k=1}^n y_k^q\Big)^{1/q}.$$

此外例如当y_k 不全为零时, Hölder 不等式中的等号成立当且仅当存在$\lambda\ge 0$ 使得$x_k^p=\lambda y_k^q$, $k=1,2,...,n$.

【证】只需考虑x_k 不全为零且y_k 不全为零的情形. 令

$$X=\sum_{k=1}^n x_k^p,\quad Y=\sum_{k=1}^n y_k^q.$$

则$X>0, Y>0$. 据Young 不等式有

$$\frac{x_k}{X^{1/p}}\cdot\frac{y_k}{Y^{1/q}}\le\frac{1}{p}\frac{x_k^p}{X}+\frac{1}{q}\frac{y_k^q}{Y},\quad k=1,2,...,n$$

两边对k 求和得到

$$\frac{1}{X^{1/p}Y^{1/q}}\sum_{i=1}^n x_ky_k\le\frac{1}{p}\frac{1}{X}\sum_{k=1}^n x_k^p+\frac{1}{q}\frac{1}{Y}\sum_{k=1}^n y_k^q=\frac{1}{p}+\frac{1}{q}=1.$$

即

$$\sum_{i=1}^{n} x_k y_k \le X^{1/p} Y^{1/q}$$

即Hölder 不等式成立.   易见当Hölder 不等式的等号成立时, 上面的n个Young 不等式中的等号都成立. 因此由Young 不等式中等号成立的充要条件知$\frac{x_k^p}{X} = \frac{y_k^q}{Y}, k = 1,2,...,n$. 反之当存在$\lambda \ge 0$ 使得$x_k^p = \lambda y_k^q\, k = 1,2,...,n$ 时, 易见Hölder 不等式中的等号成立. $\Box$

【**Minkowski 不等式**】设$1 \le p < +\infty$. 则对任意$x_k \ge 0, y_k \ge 0, k = 1,2,...,n$ 有

$$\Big(\sum_{k=1}^{n}(x_k + y_k)^p\Big)^{1/p} \le \Big(\sum_{k=1}^{n} x_k^p\Big)^{1/p} + \Big(\sum_{k=1}^{n} y_k^p\Big)^{1/p}.$$

此外当$1 < p < +\infty$时, 若(例如) y_k 不全为零, 则Minkowski 不等式中的等号成立当且仅当存在$\lambda \ge 0$ 使得$x_k = \lambda y_k$, $k = 1,2,...,n$.

【证】只需考虑$p > 1$ 和x_k 不全为零且y_k 不全为零的情形. 令$q = p/(p-1)$. 则$q(p-1) = p, 1/p + 1/q = 1$. 考虑分解$(x_k + y_k)^p = x_k(x_k + y_k)^{p-1} + y_k(x_k + y_k)^{p-1}$. 由此分解和Hölder 不等式有

$$\begin{aligned}
&\sum_{k=1}^{n}(x_k + y_k)^p = \sum_{k=1}^{n} x_k(x_k + y_k)^{p-1} + \sum_{k=1}^{n} y_k(x_k + y_k)^{p-1} \\
&\le \Big(\sum_{k=1}^{n} x_k^p\Big)^{1/p}\Big(\sum_{k=1}^{n}(x_k + y_k)^{(p-1)q}\Big)^{1/q} + \Big(\sum_{k=1}^{n} y_k^p\Big)^{1/p}\Big(\sum_{k=1}^{n}(x_k + y_k)^{(p-1)q}\Big)^{1/q} \\
&= \left\{\Big(\sum_{k=1}^{n} x_k^p\Big)^{1/p} + \Big(\sum_{k=1}^{n} y_k^p\Big)^{1/p}\right\}\Big(\sum_{k=1}^{n}(x_k + y_k)^p\Big)^{1/q}.
\end{aligned}$$

两边消去$\Big(\sum_{k=1}^{n}(x_k + y_k)^p\Big)^{1/q} > 0$ 即得Minkowski 不等式. 当Minkowski 不等式中的等号成立时, 上面推导中出现的两个Hölder 不等式中的等号成立. 因此存在$\lambda_1 > 0, \lambda_2 > 0$ 使得$x_k^p = \lambda_1(x_k + y_k)^p, y_k^p = \lambda_2(x_k + y_k)^p, k = 1,2,...,n$. 这蕴含对某个$\lambda > 0$ 有$x_k = \lambda y_k, k = 1,2,...,n$. 反之当这一共线条件满足时, 易见当Minkowski不等式中的等号成立. $\Box$

【**例3**】设$f : [0,+\infty) \to [0,+\infty)$ 单调不减且为凹函数. 又设$f(0) = 0$. 则

$$f(x + y) \le f(x) + f(y), \qquad x, y \ge 0.$$

【证】对任意$x, y > 0$ 有

$$x = \frac{x}{x+y}\cdot(x+y) + \frac{y}{x+y}\cdot 0, \quad y = \frac{y}{x+y}\cdot(x+y) + \frac{x}{x+y}\cdot 0$$

$\Longrightarrow$

$$f(x)=f(\frac{x}{x+y}\cdot(x+y)+\frac{y}{x+y}\cdot 0)\geq \frac{x}{x+y}f(x+y)+\frac{y}{x+y}f(0)=\frac{x}{x+y}f(x+y),$$

$$f(y)=f(\frac{y}{x+y}\cdot(x+y)+\frac{x}{x+y}\cdot 0)\geq \frac{y}{x+y}f(x+y)+\frac{x}{x+y}f(0)=\frac{y}{x+y}f(x+y).$$

两式相加即得

$$f(x)+f(y)\geq \frac{x}{x+y}f(x+y)+\frac{y}{x+y}f(x+y)=f(x+y). \qquad \square$$

【例4】设$f:[0,T]\to\mathbb{R}$ 为凹函数且$f(0)=0$. 则$x\mapsto f(x)/x$ 单调不增, 即

$$\frac{f(x)}{x}\geq\frac{f(y)}{y}\qquad \forall\, 0<x<y\leq T.$$

【证】对任意$0<x<y\leq T$, 可写

$$x=(1-\frac{x}{y})0+\frac{x}{y}\cdot y.$$

由凹函数的定义和$f(0)=0$ 有

$$f(x)=f\Big((1-\frac{x}{y})0+\frac{x}{y}\cdot y\Big)\geq(1-\frac{x}{y})f(0)+\frac{x}{y}f(y)=\frac{x}{y}f(y).$$

所以$\frac{f(x)}{x}\geq\frac{f(y)}{y}$. $\square$

【例5】证明Jordan 不等式

$$\sin x\geq\frac{2}{\pi}x,\qquad x\in[0,\pi/2].$$

【证】由$\sin''(x)=-\sin x\leq 0$ for all $x\in[0,\pi]$ 和**命题5.38** 可知$\sin x$ 在$[0,\pi]$ 上为凹函数. 又$\sin 0=0$, 故由例4 知

$$\frac{\sin x}{x}\geq\frac{\sin(\pi/2)}{\pi/2}=\frac{2}{\pi}\qquad\forall\, 0<x\leq\pi/2. \qquad \square$$

【例6】证明对任意$x_1,x_2,...,x_n\in[0,\pi]$ 有

$$\Big(\prod_{k=1}^{n}\sin x_k\Big)^{1/n}\leq\sin\bar{x},\quad 其中\ \ \bar{x}=\frac{x_1+x_2+\cdots+x_n}{n}.$$

【证】可以假设$0<x_k<\pi$, $k=1,2,...,n$. 令$f(x)=\log(\sin x)$, $x\in(0,\pi)$. 则

$$f'(x)=\frac{\cos x}{\sin x},\quad f''(x)=\frac{-1}{\sin^2 x}<0,\quad x\in(0,\pi).$$

所以$f(x)$在$(0,\pi)$ 内为凹函数. 因此$f(\bar{x}) \geq \frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}$ 即

$$\log(\sin\bar{x}) \geq \frac{\log(\sin x_1)+\log(\sin x_2)+\cdots+\log(\sin x_n)}{n} = \log\Big(\prod_{k=1}^{n}\sin x_k\Big)^{1/n}$$

即

$$\sin\bar{x} \geq \Big(\prod_{k=1}^{n}\sin x_k\Big)^{1/n}.$$

另法: 由几何平均不超过算术平均有

$$\Big(\prod_{k=1}^{n}\sin x_k\Big)^{1/n} \leq \frac{\sin x_1+\sin x_2+\cdots+\sin x_n}{n}.$$

而由$\sin x$ 在$[0,\pi]$是凹函数有

$$\frac{\sin x_1+\sin x_2+\cdots+\sin x_n}{n} \leq \sin\bar{x}. \qquad \square$$

3. 罗比达法则(L'Hospital rule)

这一著名法则是: 在自变量的同一极限过程中, 函数比值$\frac{f(x}{g(x)}$的"$\frac{0}{0}$"或"$\frac{*}{\infty}$"型极限等于导数比值$\frac{f'(x)}{g'(x)}$的极限, 只要后者存在:

$$\lim_{x}\frac{f(x)}{g(x)} = \lim_{x}\frac{f'(x)}{g'(x)}.$$

这一法则显然与Cauchy 中值定理有紧密联系. 因此我们先从推广的Rolle 定理和Cauchy 中值定理开始.

【命题5.41 (广义Rolle定理)】(1) 设$f:[a,+\infty)\to\mathbb{R}$ 在$[a,+\infty)$上连续, 在$(a,+\infty)$内可微且

$$f(a) = f(+\infty) := \lim_{x\to+\infty} f(x).$$

则

$$\exists\xi\in(a,+\infty) \quad \text{s.t.} \qquad f'(\xi)=0.$$

(2)设$f:\mathbb{R}\to\mathbb{R}$ 在$\mathbb{R}$上可微且

$$f(-\infty) := \lim_{x\to+\infty} f(x) = f(+\infty) := \lim_{x\to+\infty} f(x) \qquad (\text{有限或无限}).$$

则

$$\exists\xi\in\mathbb{R} \qquad \text{s.t.} \qquad f'(\xi)=0.$$

【证】(1) 假设$f'(x) \neq 0$ for all $x \in (a,+\infty)$. 则由Darboux 定理的推论知$f(x)$在$[a,+\infty)$上严格单调. 不妨设为严增. 取$c \in (a,+\infty)$. 则由单调性有$f(a) < f(c) \leq f(+\infty)$. 这与假设条件$f(a) = f(+\infty)$ 矛盾. 因此必存在$\xi \in (a,+\infty)$ 使得$f'(\xi) = 0$. 同理可证(2).
□

【命题5.42 (广义Cauchy 中值定理)】(1) 设$f, g : [a,+\infty) \to \mathbb{R}$ 在$[a,+\infty)$ 上连续, 在$(a,+\infty)$ 内可微且$f(+\infty), g(+\infty)$皆存在有限. 又设$g'(x) \neq 0$ for all $x \in (a,+\infty)$. 则

$$\exists \xi \in (a,+\infty) \qquad \text{s.t.} \qquad \frac{f(a)-f(+\infty)}{g(a)-g(+\infty)} = \frac{f'(\xi)}{g'(\xi)}.$$

(2) 设$f, g : \mathbb{R} \to \mathbb{R}$ 在$\mathbb{R}$ 上可微且$f(+\infty), f(-\infty), g(+\infty), g(-\infty)$皆存在有限. 又设$g'(x) \neq 0$ for all $x \in \mathbb{R}$. 则

$$\exists \xi \in \mathbb{R} \qquad \text{s.t.} \qquad \frac{f(+\infty)-f(-\infty)}{g(+\infty)-g(-\infty)} = \frac{f'(\xi)}{g'(\xi)}.$$

【证】(1) 由Darboux 定理的推论知$g(x)$在$[a,+\infty)$上严格单调. 令

$$F(x) = f(x) - f(a) - \frac{f(+\infty)-f(a)}{g(+\infty)-g(a)}(g(x)-g(a)), \quad x \in [a,+\infty).$$

则$F(a) = F(+\infty)$ $(= 0)$. 据广义Rolle 定理, 存在$\xi \in (a,+\infty)$ 使得$F'(\xi) = 0$. 即$f'(\xi) = \frac{f(+\infty)-f(a)}{g(+\infty)-g(a)} g'(\xi)$. 再由$g'(\xi) \neq 0$ 即得所证. 同理可证(2). □

【注】利用反射变换, 以上命题对于$(-\infty, b]$ 也成立.

【命题5.43 (“$\frac{0}{0}$” 型极限)】

(1) 设$f, g : (x_0, b) \to \mathbb{R}$ 在(x_0, b) 内可微且$g'(x) \neq 0$ for all $x \in (x_0, b)$. 又设$\lim\limits_{x \to x_0+} f(x) = \lim\limits_{x \to x_0+} g(x) = 0$ 并且$\lim\limits_{x \to x_0+} \frac{f'(x)}{g'(x)}$存在(有限或无限). 则

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

(2) 设$f, g : (a,+\infty) \to \mathbb{R}$ 在$(a,+\infty)$ 内可微且$g'(x) \neq 0$ for all $x \in (a,+\infty)$. 又设$\lim\limits_{x \to +\infty} f(x) = \lim\limits_{x \to +\infty} g(x) = 0$并且$\lim\limits_{x \to +\infty} \frac{f'(x)}{g'(x)}$存在(有限或无限). 则

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)}.$$

【证】(1) 定义$f(x_0)=g(x_0)=0$, 则由假设知f,g 在$[x_0,b)$ 上连续. 对任意$x\in(x_0,b)$, 由Cauchy 中值定理, 存在$x_0<\xi_x<x$ 使得

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(x_0)}{g(x)-g(x_0)}=\frac{f'(\xi_x)}{g'(\xi_x)}.$$

因当$x\to x_0+$ 时$\xi_x\to x_0+$, 故由假设条件知

$$\lim_{x\to x_0+}\frac{f(x)}{g(x)}=\lim_{x\to x_0+}\frac{f'(\xi_x)}{g'(\xi_x)}=\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}.$$

(2) 对任意$x\in(a,+\infty)$, 由广义Cauchy 中值定理, 存在$x<\xi_x<+\infty$ 使得

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(+\infty)}{g(x)-g(+\infty)}=\frac{f'(\xi_x)}{g'(\xi_x)}.$$

因当$x\to+\infty$ 时$\xi_x\to+\infty$, 故由假设条件知

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{f'(\xi_x)}{g'(\xi_x)}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.\qquad\square$$

【命题5.43 (“$\frac{*}{\infty}$”型极限)】(1) 设$f,g:(x_0,b)\to\mathbb{R}$ 在(x_0,b) 内可微且$g'(x)\neq0$ for all $x\in(x_0,b)$. 又设$\lim\limits_{x\to x_0+}|g(x)|=+\infty$ 并且$\lim\limits_{x\to x_0+}\frac{f'(x)}{g'(x)}$存在(有限或无限). 则

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

(2) 设$f,g:(a,+\infty)\to\mathbb{R}$ 在$(a,+\infty)$ 内可微且$g'(x)\neq0$ for all $x\in(a,+\infty)$. 又设$\lim\limits_{x\to+\infty}|g(x)|=+\infty$ 并且$\lim\limits_{x\to+\infty}\frac{f'(x)}{g'(x)}$存在(有限或无限). 则

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}.$$

【证】先证(2). 取$R_0>a$ 使得$g(x)\neq0$ for all $x\ge R_0$. 则对任意$x>R>R_0$有

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(R)}{g(x)-g(R)}\Big(1-\frac{g(R)}{g(x)}\Big)+\frac{f(R)}{g(x)}=\frac{f'(\xi_x)}{g'(\xi_x)}\Big(1-\frac{g(R)}{g(x)}\Big)+\frac{f(R)}{g(x)}\qquad(*)$$

其中$\xi_x\in(R,x)$.

先设$A:=\lim\limits_{x\to+\infty}\frac{f'(x)}{g'(x)}$有限. 则由$(*)$ 有

$$\left|\frac{f(x)}{g(x)}-\frac{f'(\xi_x)}{g'(\xi_x)}\right|\le\left|\frac{f'(\xi_x)}{g'(\xi_x)}\right|\frac{|g(R)|}{|g(x)|}+\frac{|f(R)|}{|g(x)|},$$

$$\left|\frac{f(x)}{g(x)}-A\right|\le\left|\frac{f'(\xi_x)}{g'(\xi_x)}\right|\frac{|g(R)|}{|g(x)|}+\frac{|f(R)|}{|g(x)|}+\left|\frac{f'(\xi_x)}{g'(\xi_x)}-A\right|.$$

对任意$0<\varepsilon<1$, 存在$R>R_0$ 使得

$$\left|\frac{f'(x)}{g'(x)}-A\right|<\varepsilon/2 \qquad \forall\, x\ge R.$$

由此有:

$$\left|\frac{f(x)}{g(x)}-A\right|\le(|A|+1)\frac{|g(R)|}{|g(x)|}+\frac{|f(R)|}{|g(x)|}+\varepsilon/2 \qquad \forall\, x>R.$$

因

$$(|A|+1)\frac{|g(R)|}{|g(x)|}+\frac{|f(R)|}{|g(x)|}\to 0 \quad \text{as} \quad x\to+\infty$$

故存在$R_1>R$ 使得

$$(|A|+1)\frac{|g(R)|}{|g(x)|}+\frac{|f(R)|}{|g(x)|}<\varepsilon/2 \qquad \forall\, x>R_1.$$

于是有

$$\left|\frac{f(x)}{g(x)}-A\right|<\varepsilon/2+\varepsilon/2=\varepsilon \qquad \forall\, x>R_1.$$

这证明了$\lim\limits_{x\to+\infty}\frac{f(x)}{g(x)}=A$.

其次设$\lim\limits_{x\to+\infty}\frac{f'(x)}{g'(x)}=+\infty$. 对任意$M>1$, 存在$R>R_0$ 使得

$$\frac{f'(x)}{g'(x)}>3M \qquad \forall\, x>R.$$

取$R_1>R$ 使得

$$\frac{|f(R)|}{|g(x)|}<\frac{1}{2} \qquad \forall\, x>R_1.$$

则由$(*)$ 有

$$\frac{f(x)}{g(x)}>3M\Big(1-\frac{1}{2}\Big)+\frac{1}{2}>M \qquad \forall\, x>R_1.$$

这证明了$\lim\limits_{x\to+\infty}\frac{f(x)}{g(x)}=+\infty$.　对于$\lim\limits_{x\to+\infty}\frac{f'(x)}{g'(x)}=-\infty$ 的情形, 证明相同.

(1)的证明: 我们做变量替换使之化为(2)的情形. 取$a>0$ 充分大使得$x_0+1/t<b$ for all $t>a$. 令

$$f_1(t)=f(x_0+1/t),\quad g_1(t)=g(x_0+1/t),\quad t\in(a,+\infty).$$

则

$$f_1'(t)=f'(x_0+1/t)(-1/t^2),\quad g_1'(t)=g'(x_0+1/t)(-1/t^2),\quad t\in(a,+\infty),$$

$$\lim_{t\to+\infty}\frac{f_1'(t)}{g_1'(t)}=\lim_{t\to+\infty}\frac{f'(x_0+1/t)}{g'(x_0+1/t)}=\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}.$$

这表明f_1, g_1 满足(2) 的条件. 因此有

$$\lim_{x\to x_0+}\frac{f(x)}{g(x)}=\lim_{t\to+\infty}\frac{f_1(t)}{g_1(t)}=\lim_{t\to+\infty}\frac{f_1'(t)}{g_1'(t)}=\lim_{x\to x_0+}\frac{f'(x)}{g'(x)}. \qquad \square$$

【注1】利用反射变换, 以上命题对于(a, x_0), $\lim\limits_{x\to x_0-}$ 和 $(-\infty, b)$, $\lim\limits_{x\to-\infty}$ 也成立. 将这些命题同时用于(a, x_0) 和(x_0, b) 即知这些命题对于$\lim\limits_{x\to x_0}$ 也成立.

【注2】对于不定型"1^∞"和"$\frac{1}{0}-\frac{1}{0}$", 可以通过初等变化将其化为"$\frac{0}{0}$"或"$\frac{*}{\infty}$"型.

例如设$\lim\limits_{x} u(x)=1$, $\lim\limits_{x} v(x)=+\infty$. 则由

$$(u(x))^{v(x)}=\exp(v(x)\log u(x))$$

和$v(x)\log u(x)=\frac{\log u(x)}{\frac{1}{v(x)}}$ 可知这化为"$\frac{0}{0}$"型; 或从$v(x)\log u(x)=\frac{v(x)}{\frac{1}{\log(x)}}$ 将其化为"$\frac{\infty}{\infty}$"型.

"$\frac{1}{0}-\frac{1}{0}$"化为"$\frac{0}{0}$"的一般操作: "$\frac{1}{0}-\frac{1}{0}=\frac{0-0}{0\cdot 0}=\frac{0}{0}$".

【例1】求极限

$$\lim_{x\to 0}\Big(\frac{1}{\sin^2 x}-\frac{1}{x^2}\Big).$$

【解】首先定性来看: $0<|x|<1 \Longrightarrow 0<|\sin x|<|x| \Longrightarrow$ 上述极限(如果存在的话)必定非负! 因此如果你的结果是负的, 则一定错.

方法1: 用L'Hospital rule: 先做整理(变形)

$$\frac{1}{\sin^2 x}-\frac{1}{x^2}=\frac{x^2-\sin^2 x}{x^2\sin^2 x}=\frac{x^2}{\sin^2 x}\cdot\frac{x+\sin x}{x}\cdot\frac{x-\sin x}{x^3}.$$

因

$$\lim_{x\to 0}\frac{\sin x}{x}=1$$

故

$$\lim_{x\to 0}\Big(\frac{1}{\sin^2 x}-\frac{1}{x^2}\Big)=2\lim_{x\to 0}\frac{x-\sin x}{x^3}=2\lim_{x\to 0}\frac{1-\cos x}{3x^2}=2\lim_{x\to 0}\frac{\sin x}{6x}=\frac{1}{3}.$$

方法2: 用Taylor 公式: 仍不是蛮干! 在上述变形后, 由Taylor 公式有

$$\frac{x-\sin x}{x^3}=\frac{x-\left(x-\frac{x^3}{3!}+O(x^5)\right)}{x^3}=\frac{1}{3!}+O(x^2)$$
$$\Longrightarrow\quad \lim_{x\to 0}\left(\frac{1}{\sin^2 x}-\frac{1}{x^2}\right)=2\lim_{x\to 0}\frac{x-\sin x}{x^3}=\frac{2}{3!}+\lim_{x\to 0}O(x^2)=\frac{1}{3}.\qquad\Box$$

【例2】设$f:(a,+\infty)\to\mathbb{R}$ 可微且对于常数$\lambda>0$ 有

$$\lim_{x\to+\infty}(f(x)+\lambda f'(x))=A\qquad\text{(有限或无限)}.$$

证明

$$\lim_{x\to+\infty}f(x)=A.$$

因此当A 有限时还有$\lim\limits_{x\to+\infty}f'(x)=0$.

此外举例说明当$\lambda<0$ 时, 上述结论不成立. 因此当λ 与x的极限的符号相同时, 上述结论成立; 符号相反时, 结论不成立.

【证】考虑

$$f(x)=\frac{f(x)e^{x/\lambda}}{e^{x/\lambda}}.$$

计算导数:

$$(f(x)e^{x/\lambda})'=f'(x)e^{x/\lambda}+f(x)e^{x/\lambda}\frac{1}{\lambda}=e^{x/\lambda}\frac{1}{\lambda}(f(x)+\lambda f'(x)).$$

应用"$\frac{*}{\infty}$"型的L'Hospital rule 有

$$\lim_{x\to+\infty}f(x)=\lim_{x\to+\infty}\frac{f(x)e^{x/\lambda}}{e^{x/\lambda}}=\lim_{x\to+\infty}\frac{e^{x/\lambda}\frac{1}{\lambda}(f(x)+\lambda f'(x))}{e^{x/\lambda}\frac{1}{\lambda}}=\lim_{x\to+\infty}(f(x)+\lambda f'(x))=A.$$

设$\alpha>0$, $f(x)=e^{\alpha x}$. 则$f'(x)=\alpha f(x)$ 即$f(x)-\frac{1}{\alpha}f'(x)\equiv 0$. 而$\lim\limits_{x\to+\infty}f(x)=+\infty$, 故取$\lambda=-\frac{1}{\alpha}<0$ 即知上述结论不成立. $\Box$

§5.6. 复变量的幂级数; 代数基本定理; 有理函数.

关于复数域$\mathbb{C} = \{z = x + \mathrm{i}y \mid x, y \in \mathbb{R}\}$ $(\mathrm{i} = \sqrt{-1})$ 的一些基本性质(距离、极限、级数等) 已在前面章节讲过. 在那里大家看到, 实数域中凡是不涉及序的性质, 对于复数集合、复数序列、聚点、复数函数等等也都成立(当然这里不是指这些性质的证明方法). 事实上这个结论对于一般欧几里德空间$\mathbb{R}^m$ 也成立.

下面我们以常用的Weierstrass 极限点定理以及函数的连续、可微的概念为例再来阐明这一点.

【命题5.44】任何有界的复数序列至少有一个收敛的子序列.

【证】设$\{z_n\}_{n=1}^{\infty}$ 为一个有界的复数序列. 回忆: 有界指存在$R > 0$ 使得$|z_n| \le R$ for all $n = 1, 2, 3, \dots$. 写$z_n = x_n + \mathrm{i}y_n$, $x_n, y_n \in \mathbb{R}$. 则易见$|x_n|, |y_n| \le R, n = 1, 2, 3, \dots$. 由$\mathbb{R}$上的Weierstrass 极限点定理, 存在子列$\{x_{m_j}\}_{j=1}^{\infty}$ 和$x_0 \in \mathbb{R}$ 使得$x_{m_j} \to x_0\,(j \to \infty)$. 对有界实数列$\{y_{m_j}\}_{j=1}^{\infty}$ 再次应用Weierstrass 极限点定理, 存在子列$\{y_{m_{j_k}}\}_{k=1}^{\infty}$ 和$y_0 \in \mathbb{R}$ 使得$y_{m_{j_k}} \to y_0\,(k \to \infty)$. 于是令$z_0 = x_0 + \mathrm{i}y_0$ 则有

$$|z_{m_{j_k}} - z_0| = \sqrt{(x_{m_{j_k}} - x_0)^2 + (y_{m_{j_k}} - y_0)^2} \le |x_{m_{j_k}} - x_0| + |y_{m_{j_k}} - y_0| \to 0 \quad (k \to \infty).$$

如记$n_k = m_{j_k}$, 则我们得到了$\{z_n\}_{n=1}^{\infty}$的一个收敛子列$\{z_{n_k}\}_{k=1}^{\infty}$: $\lim\limits_{k\to\infty} z_{n_k} = z_0 \in \mathbb{C}$.
□

【定义】设$E \subset \mathbb{C}$ 且E中每一点都是E 的聚点. 设$f : E \to \mathbb{C}$.

(1) 若对任意$z_0 \in E$ 都有$\lim\limits_{E\ni z\to z_0} f(z) = f(z_0)$, 则称$f$在$E$上连续.

(2) 若对任意$z_0 \in E$ 都存在差商极限

$$f'(z_0) := \lim_{E\ni z\to z_0} \frac{f(z) - f(z_0)}{z - z_0} \in \mathbb{C}$$

则称f在E上可微或可导, 并称$z \mapsto f'(z)$ 为f的导函数.

若$f'(z)$ 仍在E上可导, 则记$f''(z) = (f')'(z)$. 一般地定义

$$f^{(k)}(z) = (f^{(k-1)})'(z) \qquad \square$$

【例】对任意$n \in \mathbb{N}$, 单项式z^n 在$\mathbb{C}$ 上可导且

$$(z^n)' = nz^{n-1}, \quad (z^n)^{(k)} = \frac{n!}{(n-k)!} z^{n-k}, \quad z \in \mathbb{C}.$$

而当$k > n$ 时, $(z^n)^{(k)} \equiv 0$. 此外有如下估计(它在后面有用):

$$\left| \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right| \le n(n-1)(|z| + |h|)^{n-2}|h|.$$

【证】$n = 1$ 的情形是显然的. 设$n \ge 2$. 由二项式公式和中值定理有

$$\frac{(z+h)^n - z^n}{h} - nz^{n-1} = \sum_{k=2}^{n} C_n^k z^{n-k} h^{k-1},$$

$$\left| \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right| \le \sum_{k=2}^{n} C_n^k |z|^{n-k} |h|^{k-1} = \frac{(|z| + |h|)^n - |z|^n}{|h|} - n|z|^{n-1}$$

$$= n(\xi_1^{n-1} - |z|^{n-1}) = n(n-1)\xi_2^{n-2}(\xi_1 - |z|) \le n(n-1)(|z| + |h|)^{n-2}|h|$$

其中$|z| < \xi_2 < \xi_1 < |z| + |h|$. 由此可知

$$(z^n)' = \lim_{h \to 0} \frac{(z+h)^n - z^n}{h} = nz^{n-1}.$$

即$(z^n)' = nz^{n-1}$ for all $z \in \mathbb{C}$. 应用归纳法可导出k 阶导数的关系式. $\square$

【定理5.45】设$\{c_n\}_{n=0}^{\infty} \subset \mathbb{C}, z_0 \in \mathbb{C}$. 令

$$R = \frac{1}{\limsup\limits_{n \to \infty} |c_n|^{1/n}}.$$

则幂级数

$$\sum_{n=0}^{\infty} c_n (z - z_0)^n$$

在开圆盘$|z - z_0| < R$内处处绝对收敛, 而在开圆盘之外$|z - z_0| > R$内处处发散.

故此我们称R为这个幂级数的收敛半径. 这里定义$\frac{1}{0+} = +\infty$, $\frac{1}{\infty} = 0$.

当$R = 0$ 时, 幂级数只在$z = z_0$ 收敛.

【证】先看发散: 设$z \in \mathbb{C}$ 满足$|z - z_0| > R$. 则

$$\limsup_{n \to \infty} |c_n (z - z_0)^n|^{1/n} = (\limsup_{n \to \infty} |c_n|^{1/n})|z - z_0| = \frac{|z - z_0|}{R} > 1$$

$\Longrightarrow$ 存在子列$\{n_k\}_{k=1}^{\infty}$ 使得$\limsup\limits_{k \to \infty} |c_{n_k}(z - z_0)^{n_k}|^{1/n_k} = \frac{|z - z_0|}{R} > 1$. 取$1 < \rho < \frac{|z - z_0|}{R}$, 则存在$N \in \mathbb{N}$ 使得当$k \ge N$ 时$|c_{n_k}(z - z_0)^{n_k}|^{1/n_k} > \rho$. 于是有

$$|c_{n_k}(z - z_0)^{n_k}| > \rho^{n_k} \to +\infty \quad (k \to \infty).$$

这表明上述级数的通项不趋于零, 因此级数发散.

下证绝对收敛: 对任意$z \in \mathbb{C}$ 满足$|z-z_0|<R$, 有

$$\limsup_{n\to\infty} |c_n(z-z_0)^n|^{1/n} = (\limsup_{n\to\infty} |c_n|^{1/n})|z-z_0| = \frac{|z-z_0|}{R} < 1.$$

取ρ 满足$\frac{|z-z_0|}{R} < \rho < 1$, 则由上极限的定义知存在$N \in \mathbb{N}$使得当$n > N$ 时 $|c_n(z-z_0)^n|^{1/n} < \rho$. 由此便有

$$\sum_{n=0}^{\infty} |c_n(z-z_0)^n| \le \sum_{n=0}^{N} |c_n(z-z_0)^n| + \sum_{n=N+1}^{\infty} \rho^n < +\infty. \qquad \square$$

【定理5.46】设幂级数

$$f(z) := \sum_{n=0}^{\infty} c_n(z-z_0)^n$$

的收敛半径为$R>0$. 则和函数$f(z)$ 在开圆盘$|z-z_0|<R$内无穷次可导且

$$f^{(k)}(z) = \sum_{n=k}^{\infty} \frac{n!c_n}{(n-k)!}(z-z_0)^{n-k} = \sum_{n=0}^{\infty} \frac{(n+k)!c_{n+k}}{n!}(z-z_0)^n \tag{$*$}$$

$k = 0,1,2,\ldots$. 特别有

$$f^{(k)}(z_0) = k!c_k, \qquad k = 0,1,2,...$$

因此幂级数的系数与幂级数的和函数互相唯一确定.

【证】设$k \in \mathbb{N}$. 由

$$\Big(\frac{(n+k)!|c_{n+k}|}{n!}\Big)^{1/n} \le \Big((n+k)^k|c_{n+k}|\Big)^{1/n} = (n+k)^{\frac{k}{n}}\Big((|c_{n+k}|)^{\frac{1}{n+k}}\Big)^{\frac{n+k}{n}}$$

知

$$\limsup_{n\to\infty}\Big(\frac{(n+k)!|c_{n+k}|}{n!}\Big)^{1/n} \le 1\cdot\Big(\limsup_{n\to\infty}(|c_{n+k}|)^{\frac{1}{n+k}}\Big)^1 = \limsup_{n\to\infty}(|c_n|)^{\frac{1}{n}} = \frac{1}{R}.$$

同时由

$$\Big(\frac{(n+k)!|c_{n+k}|}{n!}\Big)^{1/n} \ge |c_{n+k}|^{1/n} = \Big((|c_{n+k}|)^{\frac{1}{n+k}}\Big)^{\frac{n+k}{n}}$$

有

$$\limsup_{n\to\infty}\Big(\frac{(n+k)!|c_{n+k}|}{n!}\Big)^{1/n} \ge \frac{1}{R}.$$

所以

$$\limsup_{n\to\infty}\Big(\frac{(n+k)!|c_{n+k}|}{n!}\Big)^{1/n} = \frac{1}{R}.$$

这表明幂级数(*) 与原幂级数有相同的收敛半径.

下证f 在收敛圆盘$|z-z_0|<R$内处处可导且上述导数关系式成立. 为记号简便, 我们以$f(z_0+z)$ 代替$f(z)$. 或者不失一般性可设$z_0=0$ 并令

$$f(z) := \sum_{n=0}^{\infty} c_n z^n, \quad |z|<R.$$

对任意z 满足$|z|<R$, 取r满足$0<r<R-|z|$. 则对任意h 满足$0<|h|\le r$, 我们有$|z|+|h|\le |z|+r<R$,

$$\begin{aligned}
&\left|\frac{f(z+h)-f(z)}{h}-\sum_{n=1}^{\infty} c_n n z^{n-1}\right| = \left|\sum_{n=2}^{\infty} c_n\Big(\frac{(z+h)^n-z^n}{h}-nz^{n-1}\Big)\right| \\
&\le \sum_{n=2}^{\infty} |c_n|\left|\frac{(z+h)^n-z^n}{h}-nz^{n-1}\right| \le |h|\sum_{n=2}^{\infty}|c_n| n(n-1)(|z|+r)^{n-2} \\
&\to 0 \quad \text{as} \quad h\to 0
\end{aligned}$$

这里用到绝对收敛: $\sum_{n=2}^{\infty}|c_n|n(n-1)(|z|+r)^{n-2}<+\infty$. 所以$f$ 在z 可导且

$$f'(z)=\sum_{n=1}^{\infty} c_n n z^{n-1}, \qquad |z|<R.$$

因导函数对应的幂级数与原级数有相同的收敛半径, 故又可知二阶导数$f''(z)$ 存在且

$$f''(z)=\sum_{n=2}^{\infty} c_n n(n-1) z^{n-2}, \qquad |z|<R.$$

应用归纳法即得$z_0=0$ 时的一般关系式(*). $\square$

【定理5.46(幂级数的乘法)】 设幂级数$\sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} b_n z^n$ 的收敛半径分别为$R_1>0, R_2>0$. 令$R=\min\{R_1,R_2\}$. 则由乘法$(\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} b_n z^n)$ 定义的幂级数的收敛半径至少是R 且

$$\Big(\sum_{n=0}^{\infty} a_n z^n\Big)\Big(\sum_{n=0}^{\infty} b_n z^n\Big)=\sum_{n=0}^{\infty} c_n z^n, \qquad |z|<R$$

其中

$$c_n=\sum_{k=0}^{n} C_n^k a_{n-k} b_k, \quad n=0,1,2,\dots.$$

【证】对任意$|z|<R$ 有

$$\sum_{n=0}^{\infty}|a_n||z|^n<+\infty, \quad \sum_{n=0}^{\infty}|b_n||z|^n<+\infty.$$

据正项级数的乘法公式有

$$\sum_{n=0}^{\infty}|c_n||z|^n \le \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{n}C_n^k|a_{n-k}||b_k|\Big)|z|^n = \Big(\sum_{n=0}^{\infty}|a_n||z|^n\Big)\Big(\sum_{n=0}^{\infty}|b_n||z|^n\Big) < +\infty.$$

所以$\sum_{n=0}^{\infty}c_nz^n$ 的收敛半径至少是R. 再由两个绝对收敛的级数的乘法关系式即得上述关系式. □

【例】幂级数的乘法和幂级数系数的唯一性在发现组合恒等式方面的应用.

设n, m, p 为非负整数且$n \ge m+1$. 证明

$$\sum_{k=0}^{p}(-1)^k C_n^{p-k}C_{m+k}^m = C_{n-m-1}^p, \qquad 0 \le p \le n-m-1.$$

【证】考虑同一函数的两种表示:

$$(1+t)^{n-m-1} = (1+t)^n(1+t)^{-m-1}, \qquad |t| < 1.$$

一方面由二项式公式有

$$(1+t)^{n-m-1} = \sum_{p=0}^{n-m-1}C_{n-m-1}^p t^p, \qquad (1+t)^n = \sum_{k=0}^{n}C_n^k t^k.$$

另一方面由$|t| < 1$ 知

$$(1+t)^{-1} = \sum_{k=0}^{\infty}(-1)^k t^k,$$

$$(-1)^m m!(1+t)^{-m-1} = ((1+t)^{-1})^{(m)} = \sum_{k=0}^{\infty}(-1)^k(t^k)^{(m)} = \sum_{k=m}^{\infty}(-1)^k\frac{k!}{(k-m)!}t^{k-m}$$

$\Longrightarrow$

$$(1+t)^{-m-1} = \sum_{k=m}^{\infty}(-1)^{k-m}\frac{k!}{(k-m)!m!}t^{k-m} = \sum_{k=m}^{\infty}(-1)^{k-m}C_k^m t^{k-m} = \sum_{k=0}^{\infty}(-1)^k C_{m+k}^m t^k.$$

按惯例我们对组合数C_n^k 做补充规定:

$$C_n^k = 0 \quad \text{if} \quad k < 0 \ \text{ or } \quad k > n.$$

则可写

$$(1+t)^n = \sum_{k=0}^{\infty}C_n^k t^k.$$

于是由级数乘法有

$$(1+t)^n(1+t)^{-m-1} = \Big(\sum_{k=0}^{\infty}C_n^k t^k\Big)\Big(\sum_{k=0}^{\infty}(-1)^k C_{m+k}^m t^k\Big) = \sum_{p=0}^{\infty}\Big(\sum_{k=0}^{p}C_n^k(-1)^{p-k}C_{m+p-k}^m\Big)t^p.$$

最后根据幂级数系数的唯一性(因收敛半径>0) 即知

$$\sum_{k=0}^{p}(-1)^{p-k}C_n^k C_{m+p-k}^m = C_{n-m-1}^p, \quad p=0,1,...,n-m-1.$$

即

$$\sum_{k=0}^{p}(-1)^{k}C_n^{p-k} C_{m+k}^m = C_{n-m-1}^p, \quad p=0,1,...,n-m-1. \qquad \square$$

• **复数的幅角表示.**

【命题5.47】

(1) 函数$\theta\mapsto e^{\mathrm{i}\theta}$ 是$[0,2\pi)$ 到单位圆周$\mathbf{S}^1=\{z\in\mathbb{C}\,|\,|z|=1\}$的一一对应(单满射).

(2) 对任意$z\in\mathbb{C}\setminus\{0\}$, 存在唯一的$\theta\in[0,2\pi)$ 使得$z=|z|e^{\mathrm{i}\theta}$.

我们称$\theta\in[0,2\pi)$ 为z 的幅角, 即θ 是z与实数轴的正向沿逆时针方向旋转的夹角.

【证】(1) 先证$\theta\mapsto e^{\mathrm{i}\theta}$ 是$[0,2\pi)$ 到$\mathbf{S}^1$的满射. 任取$z=x+\mathrm{i}y\in\mathbf{S}^1$, 其中$x,y\in\mathbb{R}$. 则$x^2+y^2=1$. 由连续函数介值定理, 存在$\theta_1\in[0,\pi],\theta_2\in[\pi,2\pi]$ 使得$x=\cos\theta_1=\cos\theta_2$. 当$y\ge 0$ 时有

$$y=\sqrt{1-x^2}=\sqrt{1-\cos^2\theta_1}=|\sin\theta_1|=\sin\theta_1.$$

当$y<0$ 时有

$$y=-\sqrt{1-x^2}=-\sqrt{1-\cos^2\theta_2}=-|\sin\theta_2|=\sin\theta_2.$$

注意: 由$0\neq y=\sin\theta_2$知$\theta_2\neq 2\pi$. 这证明了存在$\theta\in[0,2\pi)$ 使得$(x,y)=(\cos\theta,\sin\theta)$ 即$z=e^{\mathrm{i}\theta}$. 所以$\theta\mapsto e^{\mathrm{i}\theta}$ 是$[0,2\pi)$ 到$\mathbf{S}^1$的满射.

其次证明$\theta\mapsto e^{\mathrm{i}\theta}$在$[0,2\pi)$ 上是单射. 设$\theta_1,\theta_2\in[0,2\pi)$ 满足$e^{\mathrm{i}\theta_1}=e^{\mathrm{i}\theta_2}$. 不妨设$\theta_1\le\theta_2$. 则$0\le\theta_2-\theta_1<2\pi$ 且$e^{\mathrm{i}(\theta_2-\theta_1)}=e^{\mathrm{i}\theta_2}e^{-\mathrm{i}\theta_1}=1$ 即$\cos(\theta_2-\theta_1)=1,\sin(\theta_2-\theta_1)=0$. 这蕴含$\theta_2-\theta_1=0$ 或π. 但$\cos(\pi)=-1$ 故只能有$\theta_2-\theta_1=0$ 即$\theta_1=\theta_2$. 所以$\theta\mapsto e^{\mathrm{i}\theta}$ 在$[0,2\pi)$ 上是单射.

(2) 设$z\in\mathbb{C}\setminus\{0\}$. 则$z/|z|\in\mathbf{S}^1$. 因此存在唯一的$\theta\in[0,2\pi)$ 使得$z/|z|=e^{\mathrm{i}\theta}$ 即$z=|z|e^{\mathrm{i}\theta}$. $\square$

关于代数基本定理

这定理说: 复数域$\mathbb{C}$是代数封闭的, 即任何次数≥ 1的复系数多项式

$$P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

必有一复根. 进而任何n次复系数多项式必有n个复根(含重数).

实数域$\mathbb{R}$不是代数封闭的, 例如实系数多项式$1+x^2$ 便没有实根, 即方程$1+x^2=0$ 无实解. 数学大师欧拉(或许人们永远不知道他最初是怎么想的), 愣是添加了一个记号$\mathrm{i}=\sqrt{-1}$ 迫使方程$1+x^2$ 有"解": $1+\mathrm{i}^2=0$. 这一添加打开了全新局面, 而且在自然科学中起重要作用, 例如量子力学中实质上用到了复数(相位) 而不是仅仅让复数简化公式.

为证明代数基本定理, 我们先证两个引理.

【**引理5.48**】设连续函数$f:\mathbb{C}\to\mathbb{C}$ 具有下列性质:

(i) $\lim\limits_{|z|\to+\infty}|f(z)|=+\infty$.

(ii) 对任意$z_0\in\mathbb{C}$ 满足$f(z_0)\neq 0$, 存在$z_0'\in\mathbb{C}$ 使得$|f(z_0')|<|f(z_0)|$.

则f 必有一根, 即存在$z_0\in\mathbb{C}$ 使得$f(z_0)=0$.

【证】令

$$\mu = \inf_{z\in\mathbb{C}}|f(z)|.$$

由确界的定义, 存在$\{z_n\}_{n=1}^{\infty}\subset\mathbb{C}$ 使得

$$\mu \leq |f(z_n)| < \mu + 1/n, \quad n = 1, 2, 3, \ldots.$$

由假设(i) 易见$\{z_n\}_{n=1}^{\infty}$ 有界. 事实上对于$\mu+1$ 存在$R>0$ 使得当$|z|>R$时$|f(z)|>\mu+1$. 因此由$|f(z_n)|<\mu+1$ 知$|z_n|\leq R$, $n=1,2,3,\ldots$. 所以$\{z_n\}_{n=1}^{\infty}$ 有界. 据Weierstrass极限点定理, 存在子列$\{z_{n_k}\}_{k=1}^{\infty}$ 和$z_0\in\mathbb{C}$ 使得$\lim\limits_{k\to\infty}z_{n_k}=z_0$. 再由$f(z)$ 连续从而$|f(z)|$连续知

$$\mu = \lim_{k\to\infty}|f(z_{n_k})| = |f(z_0)|. \qquad \text{因此} \quad |f(z_0)| = \min_{z\in\mathbb{C}}|f(z)|.$$

由假设(ii) 易见必有$f(z_0)=0$. 否则, $f(z_0)\neq 0$, 则由假设(ii) 存在$z_0'\in\mathbb{C}$ 使得$|f(z_0')|<|f(z_0)|$. 这矛盾于$|f(z_0)|$ 是$|f|$的最小值. 所以必有$f(z_0)=0$. $\square$

【引理5.48】设$Q(z)$为一个次数≥ 1的多项式且$Q(0)=1$. 则存在$z_1\in\mathbb{C}$使得$|Q(z_1)|<1$.

【证】由假设知存在$k\in\mathbb{N}$使得

$$Q(z)=1+q_kz^k+q_{k+1}z^{k+1}+\cdots+q_nz^n,\quad q_k\neq 0,\quad q_n\neq 0.$$

往下我们只就$n\geq k+1$的情形进行证明. 对于$n=k$的情形, 证明更简单.

证明的想法是: 让z在一个适当半径的圆周上旋转, 旋转到一个合适的角度时, $|Q(z)|$便达到"最小". 这当中将充分使用复数的性质—— 可以做360度旋转！(实数没有此性质,因此这引理对于实变量多项式显然不成立). 根据复数的幅角表示, 可写$q_k=|q_k|e^{\mathrm{i}\theta}$. 取$z_1=re^{\mathrm{i}\frac{\pi-\theta}{k}}$, $0<r<1$, 则有

$$\begin{aligned}
&q_kz_1^k=q_kr^ke^{\mathrm{i}(\pi-\theta)}=r^kq_ke^{-\mathrm{i}\theta}e^{\mathrm{i}\pi}=r^k|q_k|(-1)=-|q_k|r^k,\\
&Q(z_1)=1+q_kz_1^k+\sum_{m=k+1}^n q_mz_1^m=1-|q_k|r^k+\sum_{m=k+1}^n q_mr^me^{\mathrm{i}\frac{m(\pi-\theta)}{k}}\\
&\Longrightarrow\\
&|Q(z_1)|\leq|1-|q_k|r^k|+\sum_{m=k+1}^n|q_m|r^m\leq|1-|q_k|r^k|+Cr^{k+1}
\end{aligned}$$

其中$C=\sum_{m=k+1}^n|q_m|>0$. 现在取

$$0<r<\min\left\{1,\ |q_k|^{-1/k},\ \frac{|q_k|}{C}\right\}.$$

则$|q_k|r^k<1$ 且$|q_k|>Cr$ 从而有

$$|Q(z_1)|\leq 1-|q_k|r^k+Cr^{k+1}=1-r^k(|q_k|-Cr)<1.\qquad\square$$

【代数基本定理】设P为一个次数≥ 1的多项式. 则P 在$\mathbb{C}$ 中有根, 即存在$z_0\in\mathbb{C}$使得$P(z_0)=0$. 因此若$P(z)=a_0+a_1z+\cdots+a_nz^n$ $(a_n\neq 0)$ 为n 次多项式$(n\geq 1)$, 则P可分解成

$$P(z)=a_n(z-z_1)(z-z_2)\cdots(z-z_n)$$

其中z_j 不必互不相同.

【证】设$P(z)=\sum_{k=0}^n a_kz^k$, $a_n\neq 0, n\geq 1$. 以$\frac{1}{a_n}P(z)$ 代替$P(z)$ 可以假设$a_n=1$, 即

$$P(z)=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0.$$

易见

$$|P(z)| \ge |z|^n - \sum_{k=0}^{n-1} |a_k||z|^k = |z|^n \left(1 - \sum_{k=0}^{n-1} |a_k| \frac{1}{|z|^{n-k}}\right) \to +\infty \quad (|z| \to +\infty).$$

因此 $\lim\limits_{|z|\to+\infty} |P(z)| = +\infty$.

下证: 对任意$z_0 \in \mathbb{C}$, 若$P(z_0) \ne 0$, 则存在$z_0' \in \mathbb{C}$ 使得$|P(z_0')| < |P(z_0)|$. 若此事成立, 则由**引理5.48** 知P在$\mathbb{C}$中有根.

设$z_0 \in \mathbb{C}$满足$P(z_0) \ne 0$. 将每个单项式$(z_0+z)^k$ 做二项式展开

$$(z_0+z)^k = \sum_{j=0}^{k} C_k^j z_0^{k-j} z^j = \sum_{j=0}^{n} C_k^j z_0^{k-j} z^j$$

(这里用到对组合数的约定: 当$j > k$时$C_k^j = 0$) 我们有

$$P(z_0+z) = \sum_{k=0}^{n} a_k (z_0+z)^k = \sum_{k=0}^{n} a_k \sum_{j=0}^{n} C_k^j z_0^{k-j} z^j = \sum_{j=0}^{n} \Big(\sum_{k=0}^{n} a_k C_k^j z_0^{k-j}\Big) z^j,$$

$$\sum_{k=0}^{n} a_k C_k^j z_0^{k-j}\Big|_{j=n} = \sum_{k=0}^{n} a_k C_k^n z_0^{k-n} = a_n = 1.$$

这表明$z \mapsto P(z_0+z)$ 仍是n次多项式. 令

$$Q(z) = \frac{P(z_0+z)}{P(z_0)}.$$

则Q是n次多项式且$Q(0) = 1$. 由**引理5.48**知存在$z_1 \in \mathbb{C}$ 使得$|Q(z_1)| < 1$. 于是对于$z_0' = z_0 + z_1 \in \mathbb{C}$ 有$|P(z_0')| = |P(z_0+z_1)| < |P(z_0)|$.

至此我们已证明了P 有根. 设$z_1 \in \mathbb{C}$ 为P的一个根. 则

$$P(z) = P(z) - P(z_1) = z^n - z_1^n + a_{n-1}(z^{n-1} - z_1^{n-1}) + \cdots + a_1(z - z_1) = (z - z_1)P_1(z).$$

其中P_1 是$n-1$ 次首1 多项式. 这里用了到因式分解:

$$a^k - b^k = (a-b)(a^{k-1} + a^{k-2}b + \cdots + ab^{k-2} + b^{k-1}).$$

如果$n-1 \ge 1$, 则对P_1 用上述结果知存在$z_2 \in \mathbb{C}$ 使得$P_1(z_2) = 0$ 从而有$P_1(z) = (z-z_2)P_2(z)$, 其中P_2 是$n-2$ 次首1 多项式. 于是得到$P(z) = (z-z_1)(z-z_2)P_2(z)$. 应用归纳法操作程序可知这一分解进行至第$n-1$步即得到定理中的完全分解. $\square$

设$n(\geq 1)$次多项式$P(z) = a_0 + a_1 z + \cdots + a_n z^n\,(a_n \neq 0)$的所有互不相同的根为$z_1, z_2, ..., z_m\ (m \leq n)$. 则$P(z)$ 可写成

$$P(z) = a_n(z - z_1)^{k_1}(z - z_2)^{k_2}\cdots(z - z_m)^{k_m}$$

其中$k_i \geq 1$ 且$k_1 + k_2 + \cdots + k_m = n$.

下面着重学习实系数多项式. 首先看实系数多项式的一个简单刻画:

$$P(z)\ \text{是实系数多项式} \quad \Longleftrightarrow \quad \overline{P(\overline{z})} \equiv P(z), \quad z \in \mathbb{C}.$$

事实上对任一多项式$P(z) = \sum_{k=0}^{n} a_k z^k$ 我们有

$$\overline{P(z)} = \sum_{k=0}^{n} \overline{a_k}\ \overline{z}^k, \qquad \overline{P(\overline{z})} = \sum_{k=0}^{n} \overline{a_k}\, z^k.$$

由此可知当所有系数a_k均为实数时有$\overline{P(\overline{z})} \equiv P(z)$. 反之若$\overline{P(\overline{z})} \equiv P(z)$, 则有$\sum_{k=0}^{n}(a_k - \overline{a_k})\, z^k \equiv 0$, $z \in \mathbb{C}$, 从而推知$a_k - \overline{a_k} = 0$, $k = 0, 1, ..., n$. 因此所有a_k均为实数.

【引理5.50】设$P(z), P_1(z), P_2(z)$为多项式且$P(z) = P_1(z)P_2(z)$. 若$P(z), P_1(z)$为实系数多项式, 则$P_2(z)$也是实系数多项式.

【证】根据复数的共轭运算和$P(z), P_1(z)$为实系数多项式有

$$P(z) = \overline{P(\overline{z})} = \overline{P_1(\overline{z})}\,\overline{P_2(\overline{z})} = P_1(z)\overline{P_2(\overline{z})}.$$

设$P_1(z)$的根的集合为$\{z_1, z_2, ..., z_r\}$. 则由上式有

$$\overline{P_2(\overline{z})} = P_2(z) \qquad \forall\, z \in \mathbb{C} \setminus \{z_1, z_2, ..., z_r\}.$$

即$\overline{P_2(\overline{z})} - P_2(z) = 0\ \ \forall\, z \in \mathbb{C} \setminus \{z_1, z_2, ..., z_r\}$. 根据多项式性质知$\overline{P_2(\overline{z})} - P_2(z) \equiv 0, z \in \mathbb{C}$. 所以是实系数多项式. $\square$

【实系数多项式的分解定理】设$n \geq 1$, $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ 是一个实系数多项式, 即所有系数a_k均为实数且$a_n \neq 0$. 则在实数域$\mathbb{R}$ 上P 具有以下三种分解

(1) 若P 只有实根, 则

$$P(z) = a_n(z - \alpha_1)^{k_1}(z - \alpha_2)^{k_2}\cdots(z - \alpha_r)^{k_r}$$

其中$\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{R}$ 互不相同, $k_i\geq 1$ with $k_1+k_2+\cdots+k_r=n$.

(2) 若P只有虚根没有实根, 则n 为偶数且

$$P(z)=a_n(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}\cdots(z^2+p_sz+q_s)^{m_s}$$

其中$(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$互不相同, 满足$p_j^2<4q_j\, j=1,2,...,s;\ m_j\geq 1$ 且$2(m_1+m_2+\cdots+m_s)=n$

(3) 若P 既有实根又有虚根, 则

$$P(z)=a_n(z-\alpha_1)^{k_1}(z-\alpha_2)^{k_2}\cdots(z-\alpha_r)^{k_r}(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}\cdots(z^2+p_sz+q_s)^{m_s}$$

其中$\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{R}$ 互不相同, $(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$互不相同, 满足
$p_j^2<4q_j,\ j=1,2,...,s;\ \ k_i,m_j$ 为正整数且$k_1+k_2+\cdots+k_2+2(m_1+m_2+\cdots+m_s)=n$.

【证】如前, 可以假设$a_n=1$, 否则考虑$\frac{1}{a_n}P(z)$.

(1) 设$P(z)$ 只有实根, 则由代数基本定理有

$$P(z)=(z-\alpha_1)^{k_1}(z-\alpha_2)^{k_2}\cdots(z-\alpha_r)^{k_r}$$

其中$\alpha_1,\alpha_2,...,\alpha_r$ 为互不相同的实数, $k_i\geq 1$, $k_1+k_2+\cdots+k_r=n$.

(2) 设P只有虚根没有实根, 设z_1 为$P(z)$的一个根. 因$P(z)$为实系数多项式, 故
$0=\overline{P(z_1)}=P(\overline{z_1})$. 因此$\overline{z}_1$ 也是$P(z)$的根. 但P只有虚根, 故$z_1\neq\overline{z}_1$. 于是有分解

$$P(z)=(z-z_1)^{m_1}(z-\overline{z}_1)^{l_1}P_1(z)$$

其中$P_1(z)$ 是多项式且$P_1(z_1)\neq 0, P_1(\overline{z}_1)\neq 0$. 再由$P(z)=\overline{P(\overline{z})}$ 得到

$$(z-z_1)^{m_1}(z-\overline{z}_1)^{l_1}P_1(z)=(z-\overline{z}_1)^{m_1}(z-z_1)^{l_1}\overline{P_1(\overline{z})}.$$

由此易见$m_1=l_1$. 否则, $m_1\neq l_1$, 不妨设$m_1>l_1$. 则当$z\neq z_1$ 时, 消去$(z-z_1)^{l_1}$

$$\Longrightarrow\qquad (z-z_1)^{m_1-l_1}(z-\overline{z}_1)^{l_1}P_1(z)=(z-\overline{z}_1)^{m_1}\overline{P_1(\overline{z})},\quad z\neq z_1.$$

因多项式连续, 故令$z\to z_1$ 得到矛盾:

$$0=(z_1-\overline{z}_1)^{m_1}\overline{P_1(\overline{z_1})}\neq 0.$$

这矛盾证明了$m_1=l_1$. 于是得到分解

$$P(z)=\Big((z-z_1)(z-\overline{z}_1)\Big)^{m_1}P_1(z)=(z^2+p_1z+q_1)^{m_1}P_1(z)$$

其中$p_1=-(z_1+\overline{z}_1), q_1=|z_1|^2$, $p_1^2<4q_1$.

因$P(z),(z^2+p_1z+q_1)^{m_1}$ 都是实系数多项式, 故由**引理5.50** 知$P_1(z)$ 也是实系数多项式. 又因$P(z)$只有虚根, 故$P_1(z)$也只有虚根. 将已证的结论应用于$P_1(z)$ 知若$P_1(z)$的次数>0, 则存在$z_2\in\mathbb{C}\setminus\{z_1,\overline{z_1}\}$ 和正整数m_2使得$P_1(z)=(z^2+p_2z+q_2)^{m_2}P_2(z)$, 其中$p_2=-(z_2+\overline{z}_2), q_2=|z_2|^2$, $p_2^2<4q_2$, 而$P_2(z)$ 为实系数多项式. 于是得到

$$P(z)=(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}P_2(z).$$

应用有界归纳法原理知经有限步后得到完全分解:

$$P(z)=\prod_{j=1}^{s}\Big((z-z_j)(z-\overline{z}_j)\Big)^{m_j}=\prod_{i=1}^{s}\Big(z^2+p_jz+q_j\Big)^{m_i}$$

其中$p_j=-(z_j+\overline{z}_j), q_j=|z_j|^2$, $p_j^2<4q_j$.

(3) 设P 既有实根又有虚根. 设$\alpha_1,\alpha_2,...,\alpha_r$ 是P 的所有互不相同的实根. 则有分解

$$P(z)=(z-\alpha_1)^{k_1}(z-\alpha_2)^{k_2}\cdots(z-\alpha_r)^{k_r}Q(z)$$

其中$k_i\geq 1$ with $k_1+k_2+\cdots+k_r<n$, 而多项式$Q(z)$ 只有虚根. 因$P(z)$ 和$(z-\alpha_1)^{k_1}(z-\alpha_2)^{k_2}\cdots(z-\alpha_r)^{k_r}$ 都是实系数多项式, 故**引理5.50** 知$Q(z)$也是实系数多项式. , 其次数为$n-(k_1+k_2+\cdots+k_r)$. 由(2)的结论知$Q(z)$可被分解成

$$Q(z)=(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}\cdots(z^2+p_sz+q_s)^{m_s}$$

其中$(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$ 互不相同, $p_j^2<4q_j$, $m_i\geq 1$ 且$2(m_1+m_2+\cdots+m_s)=n-(k_1+k_2+\cdots+k_r)$. 因此

$$P(z)=(z-\alpha_1)^{k_1}(z-\alpha_2)^{k_2}\cdots(z-\alpha_r)^{k_r}(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}\cdots(z^2+p_sz+q_s)^{m_s}.$$

这就完成了证明. $\Box$

【例(单位根)】设$n\in\mathbb{N}$. 试给出多项式z^n-1, z^n+1, $z^n-\mathrm{i}$, $z^n+\mathrm{i}$的分解.

【解】根据代数基本定理, 只需确定这四个多项式的所有根.

由$e^{\mathrm{i}k2\pi}=1$ 易见

$$z_k=e^{\mathrm{i}\frac{k}{n}2\pi},\quad k=0,1,...,n-1$$

是方程$z^n=1$ (即方程$z^n-1=0$) 的n个互不相同的根(这是因为$\frac{k}{n}2\pi\in[0,2\pi)$ ($k=0,1,...,n-1$) 互不相同). 因此

$$z^n-1=\prod_{k=0}^{n-1}(z-e^{\mathrm{i}\frac{k}{n}2\pi}).$$

类似地由$e^{\mathrm{i}(2k-1)\pi}=-1$ 易见

$$z_k=e^{\mathrm{i}\frac{2k-1}{n}\pi},\quad k=1,2,...,n$$

是方程$z^n=-1$ (即方程$z^n+1=0$) 的n个互不相同的根(这是因为$\frac{2k-1}{n}\pi\in(0,2\pi)$ ($k=1,2,...,n$) 互不相同). 因此

$$z^n+1=\prod_{k=1}^{n}(z-e^{\mathrm{i}\frac{2k-1}{n}\pi}).$$

再由$e^{\mathrm{i}\frac{4k+1}{2}\pi}=e^{\mathrm{i}\frac{\pi}{2}}=\mathrm{i}$ 易见

$$z_k=e^{\mathrm{i}\frac{4k+1}{2n}\pi},\quad k=0,1,...,n-1$$

是方程$z^n=\mathrm{i}$ 的n个互不相同的根(这是因为$\frac{4k+1}{2n}\pi\in[0,2\pi)$ ($k=0,1,...,n-1$) 互不相同). 因此

$$z^n-\mathrm{i}=\prod_{k=0}^{n-1}(z-e^{\mathrm{i}\frac{4k+1}{2n}\pi}).$$

最后由$e^{\mathrm{i}\frac{4k-1}{2}\pi}=e^{-\mathrm{i}\frac{\pi}{2}}=-\mathrm{i}$ 易见

$$z_k=e^{\mathrm{i}\frac{4k-1}{2n}\pi},\quad k=1,2,...,n$$

是方程$z^n=-\mathrm{i}$ 的n个互不相同的根(这是因为$\frac{4k-1}{2n}\pi\in[0,2\pi)$ ($k=1,2,...,n$) 互不相同). 因此

$$z^n+\mathrm{i}=\prod_{k=1}^{n}(z-e^{\mathrm{i}\frac{4k-1}{2n}\pi}).$$

□

关于实系数有理函数的分解定理

以下我们学习实系数有理函数的基本知识, 为后面学习有理函数的不定积分做准备. 设 $P(x),Q(x)$ 是两个实系数多项式. 假如 $\deg P\ge\deg Q$, 则由多项式的带余除法知

$$P(x)=P_0(x)Q(x)+P_1(x),\quad \deg P_1<\deg Q.$$

(这里定义 $\deg 0 = -1$, 即恒等于 0 的多项式的次数定义为 -1.) 于是

$$\frac{P(x)}{Q(x)} = P_0(x) + \frac{P_1(x)}{Q(x)}.$$

因此为研究有理函数的结构, 我们只需研究真分式.

【定义(真分式和最简分式)】若有理函数

$$\frac{P(x)}{Q(x)} \qquad \text{满足} \quad \deg P < \deg Q \ \ (\geq 1)$$

则称此有理函数是一个**真分式**.

形如

$$\frac{A}{(x-a)^k}, \qquad \frac{Bx+C}{(x^2+px+q)^m}, \quad p^2 < 4q$$

称为（实）最简分式, 其中 a, A, B, C, p, q 皆为实数; k, m 为正整数.

【真分式分解定理】任何实的真分式可被唯一地表示为有限个最简分式的线性组合. 详细来说, 设 $P(x), Q(x)$ 为实系数多项式, $\deg P < \deg Q \ \ (\geq 1)$,

$$Q(x) = (x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\cdots(x-\alpha_r)^{k_r}(x^2+p_1x+q_1)^{m_1}(x^2+p_2x+q_2)^{m_2}\cdots(x^2+p_sx+q_s)^{m_s}$$

其中 $\alpha_1, \alpha_2, ..., \alpha_r \in \mathbb{R}$ 互不相同, $(p_1, q_1), (p_2, q_2), ..., (p_s, q_s) \in \mathbb{R}^2$ 互不相同满足 $p_j^2 < 4q_j \ (j = 1, 2, ..., s)$, k_i, m_j 为正整数 $(i = 1, 2, ..., r, j = 1, 2, ..., s)$. 则存在唯一的三组实数

$$A_k^{(i)}, \quad k = 1, 2, ..., k_i; \quad i = 1, 2, ..., r;$$

$$B_m^{(j)},\ C_m^{(j)}, \quad m = 1, 2, ..., m_j; \quad j = 1, 2, ..., s$$

使得

$$\frac{P(x)}{Q(x)} \equiv \sum_{i=1}^{r}\sum_{k=1}^{k_i}\frac{A_k^{(i)}}{(x-\alpha_i)^k} + \sum_{j=1}^{s}\sum_{m=1}^{m_j}\frac{B_m^{(j)}x + C_m^{(j)}}{(x^2+p_jx+q_j)^m}. \qquad (*1)$$

注：若 $Q(x)$ 只有实根, 则等式 $(*1)$ 右端第二项不出现; 若 $Q(x)$ 只有虚根, 则等式 $(*1)$ 右端第一项不出现.

【证】根据代数中的约定, 若 $Q_1(x)$ 是 $Q(x)$ 的一个多项式因子, 即 $Q(x) = Q_1(x)P_1(x)$, 则我们用 $\frac{Q(x)}{Q_1(x)}$ 表示多项式 $P_1(x)$ 即 $\frac{Q(x)}{Q_1(x)} := P_1(x)$.

我们直接考虑一般情形, 即 $Q(x)$ 既有实根又有虚根, 其它情形的证明已包含在一般情形的证明中. 为证明 $(*1)$, 只需证明存在唯一的三组实数 $A_k^{(i)}$,$B_m^{(j)}$,$C_m^{(j)}$ 使得

$$P(x)\equiv\sum_{i=1}^{r}\sum_{k=1}^{k_i}A_k^{(i)}\frac{Q(x)}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}x+C_m^{(j)})\frac{Q(x)}{(x^2+p_jx+q_j)^m}. \qquad (*2)$$

记 $n=\deg Q$. 则由 $Q(x)$ 的分解式知 $k_1+k_2+\cdots+k_r+2(m_1+m_2+\cdots+m_s)=n$. 为证明 $(*2)$, 我们只需证明下列 n 个多项式

$$\frac{Q(x)}{(x-\alpha_i)^k},\quad k=1,2,...,k_i;\quad i=1,2,...,r;$$

$$\frac{xQ(x)}{(x^2+p_jx+q_j)^m},\quad \frac{Q(x)}{(x^2+p_jx+q_j)^m},\quad m=1,2,...,m_j;\quad j=1,2,...,s$$

线性无关. 事实上如果此事成立, 则因这 n 个多项式都属于 $\mathbf{P}_{n-1}$(= 次数 $\le n-1$ 的多项式的全体构成的 n 维线性空间), 故这 n 个多项式便是 $\mathbf{P}_{n-1}$ 的一组基. 因而任何 $P\in\mathbf{P}_{n-1}$ 都可以唯一地表示成 $(*2)$ 的形式.

为证上述 n 个多项式线性无关, 我们令它们的线性组合等于零 (即等于零多项式), 也即

$$\sum_{i=1}^{r}\sum_{k=1}^{k_i}A_k^{(i)}\frac{Q(x)}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}x+C_m^{(j)})\frac{Q(x)}{(x^2+p_jx+q_j)^m}\equiv 0,\quad x\in\mathbb{R}. \qquad (*3)$$

则我们要证明所有系数 $A_k^{(i)},B_m^{(j)},C_m^{(j)}$ 都等于零. 首先我们注意, 上式左边是一个多项式, 因此它在 $\mathbb{R}$ 上恒等于零蕴含它在 $\mathbb{C}$ 上也恒等于零.

由 $Q(x)$ 的分解式可写 $x^2+p_jx+q_j=(x-a_j)^2+b_j^2$, 其中 $a_j,b_j\in\mathbb{R}$ 且 $b_j\neq 0$. 于是 $Q(z)$ 的根 $\alpha_1,\alpha_2,...,\alpha_r,a_1+\mathrm{i}b_1,a_1-\mathrm{i}b_1,a_2+\mathrm{i}b_2,a_2-\mathrm{i}b_2,...,a_s+\mathrm{i}b_s,a_s-\mathrm{i}b_s$ 互不相同. 在恒等式 $(*3)$ 中取 $x=\alpha_1$ 得到

$$\sum_{k=1}^{k_1}A_k^{(1)}\frac{Q(x)}{(x-\alpha_1)^k}\Big|_{x=\alpha_1}=0.$$

因

$$\frac{Q(x)}{(x-\alpha_1)^k}\Big|_{x=\alpha_1}=0,\quad k=1,2,...,k_1-1;\qquad \frac{Q(x)}{(x-\alpha_1)^{k_1}}\Big|_{x=\alpha_1}\neq 0$$

故得到 $A_{k_1}^{(1)}=0$. 同理依次取 $x=\alpha_2,...,\alpha_r$ 得到 $A_{k_i}^{(i)}=0$, $i=1,2,...,r$. 令

$$\begin{aligned}\widetilde{Q}(x)&=\frac{Q(x)}{(x-\alpha_1)(x-\alpha_1)\cdots(x-\alpha_r)}\\&=(x-\alpha_1)^{k_1-1}(x-\alpha_2)^{k_2-1}\cdots(x-\alpha_r)^{k_r-1}\\&\times(x^2+p_1x+q_1)^{m_1}(x^2+p_2x+q_2)^{m_2}\cdots(x^2+p_sx+q_s)^{m_s}\,.\end{aligned}$$

则 $(*3)$ 式变成

$$\sum_{i=1}^{r}\sum_{k=1}^{k_i-1}A_k^{(i)}\frac{\widetilde{Q}(x)}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}x+C_m^{(j)})\frac{\widetilde{Q}(x)}{(x^2+p_jx+q_j)^m}\equiv 0,\quad x\in\mathbb{R}.$$

注意：若对某个 i 有 $k_i=1$, 则已完成了 $A_k^{(i)}=0\,(k=1,...,k_i)$ 的证明, 从而上面第一个和式中的第 i 项不出现. 使用相同的方法若干次 (用有限归纳法原理), 即可证明所有系数 $A_k^{(i)}$ 均为零 ($k=1,...,k_i; i=1,2,...,r.$)

这样一来, $(*3)$ 式变成 (通过消去因子 $(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\cdots(x-\alpha_r)^{k_r}$)

$$\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}x+C_m^{(j)})\frac{Q_*(x)}{(x^2+p_jx+q_j)^m}\equiv 0,\quad x\in\mathbb{R}.$$

如上所说, 这等价于

$$\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}z+C_m^{(j)})\frac{Q_*(z)}{(z^2+p_jz+q_j)^m}\equiv 0,\quad z\in\mathbb{C}$$

也即

$$\sum_{j=1}^{s}\sum_{m=1}^{m_j}(B_m^{(j)}z+C_m^{(j)})\frac{Q_*(z)}{[(z-z_1)(z-\overline{z}_1)]^m}\equiv 0,\quad z\in\mathbb{C}\qquad (*4)$$

其中

$$\begin{aligned}&Q_*(z)=(z^2+p_1z+q_1)^{m_1}(z^2+p_2z+q_2)^{m_2}\cdots(z^2+p_sz+q_s)^{m_s}\\&=[(z-z_1)(z-\overline{z}_1)]^{m_1}[(z-z_2)(z-\overline{z}_2)]^{m_2}\cdots[(z-z_s)(z-\overline{z}_s)]^{m_s},\\&z_j=a_j+\mathrm{i}b_j.\end{aligned}$$

取 $z=z_1$ 得到

$$\sum_{m=1}^{m_1}(B_m^{(1)}z+C_m^{(1)})\frac{Q_*(z)}{[(z-z_1)(z-\overline{z}_1)]^m}\Big|_{z=z_1}=0.$$

因

$$\frac{Q_*(z)}{[(z-z_1)(z-\overline{z}_1)]^m}\Big|_{z=z_1}=0,\quad m=1,2,...,m_1-1;\qquad \frac{Q_*(z)}{[(z-z_1)(z-\overline{z}_1)]^{m_1}}\Big|_{z=z_1}\neq 0$$

故得到

$$B_{m_1}^{(1)}z_1+C_{m_1}^{(1)}=0\quad\text{i.e.}\quad B_{m_1}^{(1)}(a_1+\mathrm{i}b_1)+C_{m_1}^{(1)}=0.$$

同理取 $z=\overline{z}_1$ 得到

$$B_{m_1}^{(1)}\overline{z}_1+C_{m_1}^{(1)}=0\quad\text{i.e.}\quad B_{m_1}^{(1)}(a_1-\mathrm{i}b_1)+C_{m_1}^{(1)}=0.$$

两式相减得到 $B^{(1)}_{m_1}=0$ 其中用到 z_1 的虚部 $b_1\neq 0$. 进而得到 $C^{(1)}_{m_1}=0$.

运用同样方法, 依次取 $z=z_2,\overline{z}_2,z_3,\overline{z}_3,...,z_s,\overline{z}_s$ 得到 $B^{(j)}_{m_j}=C^{(j)}_{m_j}=0, j=1,2,...,s$.

令

$$\begin{aligned}\widetilde{Q_*}(x)&=\frac{Q_*(x)}{(z-z_1)(z-\overline{z}_1)(z-z_2)(z-\overline{z}_2)\cdots(z-z_s)(z-\overline{z}_s)}\\&=[(z-z_1)(z-\overline{z}_1)]^{m_1-1}[(z-z_2)(z-\overline{z}_2)]^{m_2-1}\cdots[(z-z_s)(z-\overline{z}_s)]^{m_s-1}\end{aligned}$$

则 $(*4)$ 式变为

$$\sum_{j=1}^{s}\sum_{m=1}^{m_j-1}(B^{(j)}_m z+C^{(j)}_m)\frac{\widetilde{Q_*}(z)}{[(z-z_1)(z-\overline{z}_1)]^m}\equiv 0,\quad z\in\mathbb{C}.$$

这里, 如上所说, 若对某个 j 有 $m_j=1$, 则已完成了 $B^{(j)}_m=C^{(j)}_m=0(m=1,...,m_j)$ 的证明, 从而上面和式中的第 j 项不出现. 使用相同的方法若干次 (用有限归纳法原理), 即可证明所有系数 $B^{(j)}_m,C^{(j)}_m$ 均为零 ($m=1,...,m_j;j=1,2,...,s.$) 这就证明了上述 n 个多项式线性无关, 从而完成了定理的证明. $\square$

【注】上述证明只是保证了最简分式中的系数$A^{(i)}_k$, $B^{(j)}_m$, $C^{(j)}_m$的存在唯一性, 并没有给出计算这些系数的具体方法. 由于多项式$P(x),Q(x)$是已知的, 在实际问题中, 一般是通过比较恒等式$(*2)$ 两边的多项式的系数来确定$A^{(i)}_k$, $B^{(j)}_m$, $C^{(j)}_m$. 当然这需要解一个线性方程组. 在特殊情况下, 例如当$Q(x)$ 无重根时, 最简分式中系数可以通过计算$P(x),Q(x)$ 在$Q(x)$的根处的函数值和导数来确定. 这就是下面定理的内容.

【当$Q(x)$ 无重根时真分式$P(x)/Q(x)$分解中系数的确定】设 $P(x),Q(x)$ 为实系数多项式, $\deg P<\deg Q\ (\geq 1)$, 且$Q(x)$ 无重根:

$$Q(x)=\prod_{j=1}^{r}(x-\alpha_j)\cdot\prod_{k=1}^{s}(x-z_k)(x-\overline{z_k})=\prod_{j=1}^{r}(x-\alpha_j)\cdot\prod_{k=1}^{s}(x^2+p_kx+q_k)$$

其中 $\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{R}$ 互不相同, $(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$ 互不相同满足 $p_k^2<4q_k$ $(k=1,2,...,s)$. 则在最简分解中

$$\frac{P(x)}{Q(x)}=\sum_{j=1}^{r}\frac{A_j}{x-\alpha_j}+\sum_{k=1}^{s}\frac{B_kx+C_k}{x^2+p_kx+q_k}\qquad(*5)$$

系数A_j, B_k, C_k 由P, Q 在Q的根处的函数值和导数确定:

$$\begin{aligned}
&A_j = \frac{P(\alpha_i)}{Q'(\alpha_j)}, \quad j = 1, 2, ..., r, \\
&B_k = 2\mathrm{Re}\left(\frac{P(z_k)}{Q'(z_k)}\right), \quad C_k = -2\mathrm{Re}\left(\frac{P(z_k)}{Q'(z_k)}\overline{z_k}\right), \quad k = 1, 2, ..., s\,.
\end{aligned}$$

注：若 $Q(x)$ 只有实根, 则等式 $(*5)$ 右端第二项不出现; 若 $Q(x)$ 只有虚根, 则等式 $(*5)$ 右端第一项不出现.

【证】如前, 我们直接考虑一般情形, 即 $Q(x)$ 既有实根又有虚根, 其它情形的证明已包含在一般情形的证明中. 易见恒等式$(*5)$ 对所有复数也成立. 因此有

$$P(z) = \sum_{j=1}^{r} A_j \frac{Q(z)}{z-\alpha_j} + \sum_{k=1}^{s}(B_k z + C_k)\frac{Q(z)}{(z-z_k)(z-\overline{z_k})}, \quad z \in \mathbb{C}. \qquad (*6)$$

计算:

$$\frac{Q(z)}{z-\alpha_j}\Big|_{z=\alpha_j} = \lim_{z\to\alpha_j}\frac{Q(z)}{z-\alpha_j} = \lim_{z\to\alpha_j}\frac{Q(z)-Q(\alpha_j)}{z-\alpha_j} = Q'(\alpha_j),$$

$$\frac{Q(z)}{(z-z_k)(z-\overline{z_k})}\Big|_{z=\alpha_j} = \frac{Q(\alpha_j)}{(\alpha_j-z_k)(\alpha_j-\overline{z_k})} = 0,$$

$$\frac{Q(z)}{z-z_k}\Big|_{z=z_k} = \lim_{z\to z_k}\frac{Q(z)}{z-z_k} = \lim_{z\to z_k}\frac{Q(z)-Q(z_k)}{z-z_k} = Q'(z_k),$$

$$\frac{Q(z)}{z-\overline{z_k}}\Big|_{z=\overline{z_k}} = \lim_{z\to\overline{z_k}}\frac{Q(z)}{z-\overline{z_k}} = \lim_{z\to z_k}\frac{Q(z)-Q(\overline{z_k})}{z-\overline{z_k}} = Q'(\overline{z_k}),$$

$$\frac{Q(z)}{z-\alpha_j}\Big|_{z=z_k} = \frac{Q(z_k)}{z_k-\alpha_j} = 0, \quad \frac{Q(z)}{z-\alpha_j}\Big|_{z=\overline{z_k}} = \frac{Q(\overline{z_k})}{\overline{z_k}-\alpha_j} = 0.$$

于是在$(*6)$中分别取$z=\alpha_j$, $z=z_k$, $z=\overline{z_k}$ 得到

$$P(\alpha_j) = A_j Q'(\alpha_j), \quad \text{i.e.} \quad A_j = \frac{P(\alpha_i)}{Q'(\alpha_j)}, \quad j = 1, 2, ..., r;$$

$$P(z_k) = (B_k z_k + C_k)\frac{Q'(z_k)}{z_k-\overline{z_k}},$$

$$P(\overline{z_k}) = (B_k\overline{z_k} + C_k)\frac{Q'(\overline{z_k})}{\overline{z_k}-z_k}.$$

对后两个等式先做除法然后相减得到

$$B_k(z_k-\overline{z_k})=\left(\frac{P(z_k)}{Q'(z_k)}+\frac{P(\overline{z_k})}{Q'(\overline{z_k})}\right)(z_k-\overline{z_k})$$

即

$$B_k=\frac{P(z_k)}{Q'(z_k)}+\frac{P(\overline{z_k})}{Q'(\overline{z_k})}=2\mathrm{Re}\left(\frac{P(z_k)}{Q'(z_k)}\right).$$

这里用到P,Q 是实系数多项式, 它蕴含

$$\overline{\left(\frac{P(z_k)}{Q'(z_k)}\right)}=\frac{\overline{P(z_k)}}{\overline{Q'(z_k)}}=\frac{P(\overline{z_k})}{Q'(\overline{z_k})}.$$

将B_k的表达式代入前面方程即得到C_k的表达式:

$$B_kz_k+C_k=\frac{P(z_k)}{Q'(z_k)}(z_k-\overline{z_k}),$$

$$C_k=\frac{P(z_k)}{Q'(z_k)}(z_k-\overline{z_k})-B_kz_k=\frac{P(z_k)}{Q'(z_k)}(z_k-\overline{z_k})-\left(\frac{P(z_k)}{Q'(z_k)}+\frac{P(\overline{z_k})}{Q'(\overline{z_k})}\right)z_k$$

$$=-\left(\frac{P(z_k)}{Q'(z_k)}\overline{z_k}+\frac{P(\overline{z_k})}{Q'(\overline{z_k})}z_k\right)=-2\mathrm{Re}\left(\frac{P(z_k)}{Q'(z_k)}\overline{z_k}\right). \qquad \square$$

【例】设$n\ge 2$. 试给出$\frac{1}{x^{2n-1}+1}$, $\frac{1}{x^{2n}+1}$ 的最简分式分解, 并具体计算$\frac{1}{x^3+1}$, $\frac{1}{x^5+1}$, $\frac{1}{x^4+1}$, $\frac{1}{x^6+1}$的最简分式分解.

【解】(1) 奇数阶的情形: 由上面关于单位根的例题知

$$z^{2n-1}+1=\prod_{k=1}^{2n-1}(z-e^{\mathrm{i}\frac{2k-1}{2n-1}\pi})=(z+1)\prod_{k=1}^{n-1}(z-e^{\mathrm{i}\frac{2k-1}{2n-1}\pi})\prod_{k=n+1}^{2n-1}(z-e^{\mathrm{i}\frac{2k-1}{2n-1}\pi}).$$

因

$$\prod_{k=n+1}^{2n-1}(z-e^{\mathrm{i}\frac{2k-1}{2n-1}\pi})=\prod_{k=1}^{n-1}(z-e^{\mathrm{i}\frac{2(2n-k)-1}{2n-1}\pi})=\prod_{k=1}^{n-1}(z-e^{-\mathrm{i}\frac{2k-1}{2n-1}\pi})$$

故

$$z^{2n-1}+1=(z+1)\prod_{k=1}^{n-1}(z-e^{\mathrm{i}\frac{2k-1}{2n-1}})(z-e^{-\mathrm{i}\frac{2k-1}{2n-1}\pi})z=(z+1)\prod_{k=1}^{n-1}(z^2-2\cos(\frac{2k-1}{2n-1}\pi)z+1).$$

即

$$x^{2n-1}+1=(x+1)\prod_{k=1}^{n-1}(x^2-2\cos(\frac{2k-1}{2n-1}\pi)x+1).$$

因无重根, 故

$$\frac{1}{x^{2n-1}+1}=\frac{A_1}{x+1}+\sum_{k=1}^{n-1}\frac{B_kx+C_k}{x^2-2\cos(\frac{2k-1}{2n-1}\pi)x+1}.$$

根据上述定理, 实常数A_1, B_k, C_k 由$Q(z)=z^{2n-1}+1$ 在其根

$$\alpha_1=-1,\quad z_k=e^{\mathrm{i}\frac{2k-1}{2n-1}\pi},\quad k=1,2,...,n-1$$

处的导数确定:

$$A_1=\frac{1}{Q'(-1)},\quad B_k=2\mathrm{Re}\left(\frac{1}{Q'(z_k)}\right),\quad C_k=-2\mathrm{Re}\left(\frac{1}{Q'(z_k)}\overline{z_k}\right),\quad k=1,2,...,n-1.$$

计算:

$$\begin{aligned}A_1&=\frac{1}{Q'(-1)}=\frac{1}{(2n-1)z^{2n-2}}\Big|_{z=-1}=\frac{1}{2n-1},\\ \frac{1}{Q'(z_k)}&=\frac{1}{(2n-1)z_k^{2n-2}}=\frac{z_k}{(2n-1)z_k^{2n-1}}=-\frac{z_k}{2n-1},\\ \frac{1}{Q'(z_k)}\overline{z_k}&=-\frac{z_k\overline{z_k}}{2n-1}=-\frac{1}{2n-1}\end{aligned}$$

其中用到$z_k^{2n-1}=-1$, $z_k\overline{z_k}=|z_k|^2=1$. 所以

$$B_k=-\frac{2}{2n-1}\mathrm{Re}(z_k)=-\frac{2}{2n-1}\cos(\frac{2k-1}{2n-1}\pi),\quad C_k=\frac{2}{2n-1},\quad k=1,2,...,n-1.$$

结果是:

$$\frac{1}{x^{2n-1}+1}=\frac{1}{2n-1}\cdot\frac{1}{x+1}+\frac{2}{2n-1}\sum_{k=1}^{n-1}\frac{-\cos(\frac{2k-1}{2n-1}\pi)x+1}{x^2-2\cos(\frac{2k-1}{2n-1}\pi)x+1}.$$

当$n=2$ 时, $\cos(\frac{1}{3}\pi)=\frac{1}{2}$,

$$\frac{1}{x^3+1}=\frac{1}{3}\cdot\frac{1}{x+1}+\frac{2}{3}\cdot\frac{-\frac{1}{2}x+1}{x^2-x+1}=\frac{1}{3}\cdot\frac{1}{x+1}+\frac{1}{3}\cdot\frac{-x+2}{x^2-x+1}.$$

当$n=3$ 时,

$$\frac{1}{x^5+1}=\frac{1}{5}\cdot\frac{1}{x+1}+\frac{2}{5}\cdot\frac{-\cos(\frac{1}{5}\pi)x+1}{x^2-2\cos(\frac{1}{5}\pi)x+1}+\frac{2}{5}\cdot\frac{-\cos(\frac{3}{5}\pi)x+1}{x^2-2\cos(\frac{3}{5}\pi)x+1}.$$

或写成

$$\frac{1}{x^5+1}=\frac{1}{5}\cdot\frac{1}{x+1}+\frac{2}{5}\cdot\frac{-\cos(\frac{1}{5}\pi)x+1}{x^2-2\cos(\frac{1}{5}\pi)x+1}+\frac{2}{5}\cdot\frac{\cos(\frac{2}{5}\pi)x+1}{x^2+2\cos(\frac{2}{5}\pi)x+1}.$$

(2) 偶数阶的情形:

$$z^{2n}+1=\prod_{k=1}^{2n}(z-e^{\mathrm{i}\frac{2k-1}{2n}\pi})=\prod_{k=1}^{n}(z-e^{\mathrm{i}\frac{2k-1}{2n}\pi})\cdot\prod_{k=n+1}^{2n}(z-e^{\mathrm{i}\frac{2k-1}{2n}\pi}).$$

因

$$\prod_{k=n+1}^{2n}(z-e^{\mathrm{i}\frac{2k-1}{2n}\pi})=\prod_{k=1}^{n}(z-e^{\mathrm{i}\frac{2(2n+1-k)-1}{2n}\pi})=\prod_{k=1}^{n}(z-e^{-\mathrm{i}\frac{2k-1}{2n}\pi})$$

故

$$z^{2n}+1=\prod_{k=1}^{n}(z-e^{\mathrm{i}\frac{2k-1}{2n}\pi})(z-e^{-\mathrm{i}\frac{2k-1}{2n}\pi})=\prod_{k=1}^{n}(z^2-2\cos(\frac{2k-1}{2n}\pi)z+1).$$

即

$$x^{2n}+1=\prod_{k=1}^{n}(x^2-2\cos(\frac{2k-1}{2n}\pi)x+1).$$

因无重根, 故

$$\frac{1}{x^{2n}+1}=\sum_{k=1}^{n}\frac{B_kx+C_k}{x^2-2\cos(\frac{2k-1}{2n}\pi)x+1}.$$

根据上述定理, 实常数B_k, C_k 由$Q(z)=z^{2n}+1$ 在其根

$$z_k=e^{\mathrm{i}\frac{2k-1}{2n}\pi},\quad k=1,2,...,n$$

处的导数确定:

$$B_k=2\mathrm{Re}\left(\frac{1}{Q'(z_k)}\right),\quad C_k=-2\mathrm{Re}\left(\frac{1}{Q'(z_k)}\overline{z_k}\right),\quad k=1,2,...,n\,.$$

计算:

$$\frac{1}{Q'(z_k)}=\frac{1}{2nz_k^{2n-1}}=\frac{z_k}{2nz_k^{2n}}=-\frac{z_k}{2n},\qquad \frac{1}{Q'(z_k)}\overline{z_k}=-\frac{z_k\overline{z_k}}{2n}=-\frac{1}{2n}$$

其中用到$z_k^{2n}=-1,\ z_k\overline{z_k}=|z_k|^2=1$. 所以

$$B_k=-\frac{1}{n}\mathrm{Re}(z_k)=-\frac{1}{n}\cos(\frac{2k-1}{2n}\pi),\quad C_k=\frac{1}{n},\quad k=1,2,...,n.$$

结果是:

$$\frac{1}{x^{2n}+1}=\frac{1}{n}\sum_{k=1}^{n}\frac{-\cos(\frac{2k-1}{2n}\pi)x+1}{x^2-2\cos(\frac{2k-1}{2n}\pi)x+1}.$$

当$n=2$ 时,

$$\cos(\frac{1}{4}\pi)=\frac{\sqrt{2}}{2},\qquad \cos(\frac{3}{4}\pi)=-\frac{\sqrt{2}}{2},$$

$$\frac{1}{x^4+1} = \frac{1}{2} \cdot \frac{-\frac{\sqrt{2}}{2}x+1}{x^2-\sqrt{2}\,x+1} + \frac{1}{2} \cdot \frac{\frac{\sqrt{2}}{2}x+1}{x^2+\sqrt{2}\,x+1}, \qquad x \in \mathbb{R}.$$

当$n=3$ 时,

$$\cos(\frac{1}{6}\pi) = \frac{\sqrt{3}}{2}, \quad \cos(\frac{3}{6}\pi) = 0, \quad \cos(\frac{5}{6}\pi) = -\frac{\sqrt{3}}{2},$$

$$\frac{1}{x^6+1} = \frac{1}{3} \cdot \frac{-\frac{\sqrt{3}}{2}x+1}{x^2-\sqrt{3}\,x+1} + \frac{1}{3} \cdot \frac{1}{x^2+1} + \frac{1}{3} \cdot \frac{\frac{\sqrt{3}}{2}x+1}{x^2+\sqrt{3}\,x+1}, \qquad x \in \mathbb{R}. \qquad \square$$

§5.7. 多项式与微分算子, 一些简单的微分方程

设$P(z)$为n次多项式(不失一般性假设$a_n = 1$):

$$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0.$$

因为多项式运算只涉及加法和乘法, 不涉及除法, 故如大家在线代数中学到的那样,可以将z 换成方阵$A \in \mathbb{C}^{m\times m}$, 这是因为

$$P(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 \mathrm{I}$$

有意义且仍为同类方阵. 同理对于微分算子

$$D := \frac{\mathrm{d}}{\mathrm{d}t}, \quad D^k := D^{k-1}(D) = \frac{\mathrm{d}^k}{\mathrm{d}t^k}, \quad k = 1, 2, ...$$

其相应的多项式

$$P(D) = D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0 \mathrm{I}$$

也有意义: $P(D)$ 是作用在函数类$C^n(I, \mathbb{C})$ 上的微分算子:

$$P(D)f = D^n f + a_{n-1}D^{n-1}f + \cdots + a_1 Df + a_0 f$$

即

$$P(D)f(t) = D^n f(t) + a_{n-1}D^{n-1}f(t) + \cdots + a_1 Df(t) + a_0 f(t), \quad t \in \mathbb{R}.$$

本节我们利用多项式的性质研究如下形式的n阶常微分方程[8]

$$f^{(n)}(t) + a_{n-1}f^{(n-1)}(t) + \cdots + a_1 f'(t) + a_0 f(t) = 0, \quad t \in \mathbb{R}$$

如用微分算子$D^k f = f^{(k)}$, 则这个微分方程写为

$$P(D)f(t) = 0, \quad t \in \mathbb{R}.$$

设$\lambda \in \mathbb{C}$ 为$P(z)$的任一根. 则有分解$P(z) = P_1(z)(z - \lambda)$. 如果$f$ 是一阶微分方程

$$f'(t) = \lambda f(t) \qquad \text{即} \qquad (D - \lambda)f(t) = 0, \qquad t \in \mathbb{R}$$

的解, 则由

$$P(D)f(t) = P_1(D)(D - \lambda\mathrm{I})f(t) = P_1(D)\Big((D - \lambda\mathrm{I})f(t)\Big) = P_1(D)0 = 0, \quad t \in \mathbb{R}$$

[8]**注意：微分方程是关于未知函数f的方程, 而不是关于函数自变量t的方程, 因此微分方程中的等号是关于自变量t的恒等式!**

知f 也是n阶微分方程$P(D)f=0$ 的解.

这就引导我们先研究上述类型的一阶微分方程的解. 很多数学模型都涉及这种形式的一阶微分方程, 而其解都是指数型函数.

【命题5.51: 指数函数的作用】设$I\subset\mathbb{R}$ 为一区间, 不失一般性, 假设$0\in I$.

(1) 设实值或复值函数$a(t)$在I 上连续. 则一个实值或复值的可微函数f满足微分方程

$$f'(t)=a(t)f(t),\quad t\in I \tag{$*1$}$$

当且仅当

$$f(t)=Ce^{A(t)},\quad t\in I$$

其中C是常数, $A(t)$ 是$a(t)$的一个原函数, 即$A(t)$ 满足$A'(t)=a(t)$.

解的表达式$f(t)=Ce^{A(t)}$ 中的常数C 可由$t=0$ 确定: $C=f(0)e^{-A(0)}$. 于是方程$(*1)$ 的解都可表示成

$$f(t)=f(0)e^{A(t)-A(0)},\quad t\in I$$

此外在在满足相同初值的条件下, 方程$(*1)$的解还是唯一的: 若$g(t)$ 也满足方程$(*)$ 且$g(0)=f(0)$, 则$g(t)\equiv f(t)$.

(2) 特殊情形: $a(t)\equiv a$ 为常数. 此时$A(t)=at$. 则一个实值或复值的可微函数f满足微分方程

$$f'(t)=af(t),\quad t\in\mathbb{R} \tag{$*2$}$$

当且仅当

$$f(t)=Ce^{at},\quad t\in\mathbb{R}$$

其中C为常数. 此外, 在满足相同初值的条件下, 方程$(*2)$ 的解还是唯一的.

【证】只需证(1). 我们设法把方程$(*1)$ 变形成$F'(t)=0$ 的形式, 从而容易求得方程$(*1)$的解并给出解的结构. 经观察, 这个$F(t)$ 可取为$f(t)e^{-A(t)}$. 事实上我们有

$$(f(t)e^{-A(t)})'=f'(t)e^{-A(t)}+f(t)e^{-A(t)}(-A'(t))=e^{-A(t)}\Big(f'(t)-f(t)a(t)\Big).$$

由此即知

$$f\ \text{是方程}(*1)\text{的解}\iff (f(t)e^{-A(t)})'\equiv 0\iff f(t)e^{-A(t)}=\text{常数 }C,\quad t\in I.$$

最后这一条当然等价于$f(t)=Ce^{A(t)}$, 而常数$C=f(0)e^{-A(0)}$.

用同样的论证易见在满足相同初值的条件下方程$(*1)$的解还是唯一的. 事实上如令$h(t) = g(t) - f(t)$, 则由$h'(t) = g'(t) - f'(t)$ 知$h(t)$ 也满足方程$(*1)$. 于是根据上述推导并注意$h(0) = g(0) - f(0) = 0$ 得到$h(t) = h(0)e^{A(t)-A(0)} \equiv 0$. 因此$g(t) \equiv f(t)$. $\square$

【命题5.52: 指数函数的作用(续)】设$I \subset \mathbb{R}$ 为一区间, 不失一般性, 假设$0 \in I$.

(1) 设实值或复值函数$a(t), b(t)$在I 上连续. 则一个实值或复值的可微函数f满足微分方程

$$f'(t) = a(t)f(t) + b(t), \quad t \in I \tag{$*3$}$$

当且仅当

$$f(t) = (c + B(t))e^{A(t)}, \quad t \in I$$

其中c 是一个常数, $A(t)$ 是$a(t)$的一个原函数, $B(t)$ 是$b(t)e^{-A(t)}$ 的一个原函数, 即$A'(t) = a(t)$, $B'(t) = b(t)e^{-A(t)}$. 解的表达式$f(t) = (c + B(t))e^{A(t)}$ 中的常数c 可由$t = 0$ 确定: $c = f(0)e^{-A(0)} - B(0)$.

此外在满足相同初值的条件下, 方程$(*3)$的解还是唯一的: 若$g(t)$ 也满足方程$(*3)$且$g(0) = f(0)$, 则$g(t) \equiv f(t)$.

(2) 特殊情形: 设$a(t) \equiv a \neq 0, b(t) \equiv b$均为常数. 此时$A(t) = at, B(t) = -\frac{b}{a}e^{-at}$. 则一个实值或复值的可微函数$f$满足微分方程

$$f'(t) = af(t) + b, \quad t \in \mathbb{R} \tag{$*4$}$$

当且仅当

$$f(t) = (C - \frac{b}{a}e^{-at})e^{at} = Ce^{at} - \frac{b}{a}, \quad t \in \mathbb{R}$$

其中C为常数. 此常数C 可由$f(0)$ 确定: $C = f(0) + \frac{b}{a}$. 因此方程$(*4)$的任意解可表为

$$f(t) = f(0)e^{at} - \frac{b}{a}\Big(1 - e^{at}\Big), \quad t \in \mathbb{R}.$$

此外, 在满足相同初值的条件下, 方程$(*4)$ 的解还是唯一的.

【证】只需证(1). 为求解此方程, 我们设法将其化成$F'(t) = G(t)$ 的形式. 如上, $F(t)$ 可取为$f(t)e^{-A(t)}$. 我们有

$$(f(t)e^{-A(t)})' = f'(t)e^{-A(t)} + f(t)e^{-A(t)}(-A'(t)) = e^{-A(t)}\Big(f'(t) - f(t)a(t)\Big).$$

于是由原函数的定义得知

$$f \text{ 是方程}(*3)\text{的解} \iff (f(t)e^{-A(t)})' \equiv b(t)e^{-A(t)} \iff f(t)e^{-A(t)} = c+B(t), \quad t \in I.$$

最后这一条当然等价于$f(t) = c + B(t))e^{A(t)}$, 而常数$c = f(0)e^{-A(0)} - B(0)$.

设$g(t)$也是方程$(*3)$的解且$g(0) = f(0)$. 令$h(t) = g(t) - f(t)$, 则$h(t)$ 满足微分方程$h'(t) = a(t)h(t)$. 据**命题5.51** 和$h(0) = 0$ 即知$h(t) \equiv 0$. 因此$g(t) \equiv f(t)$. □

【命题5.53】对任意一组复数$a_0, a_1, ..., a_{n-1}$ $(n \geq 1)$, 考察微分方程

$$f^{(n)}(t) + a_{n-1}f^{(n-1)}(t) + \cdots + a_1 f'(t) + a_0 f(t) = 0, \quad t \in \mathbb{R}. \tag{$*5$}$$

我们有下列性质:

(1)对于相应的多项式$P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ 的任一根λ 和任意常数$C \in \mathbb{C}$, 函数$f(t) = Ce^{\lambda t}$ 都是方程$(*5)$的解.

(2) 方程$(*5)$的解的全体是复数域上的一个线性空间, 即若f, g 是方程$(*5)$的解, α 为常数, 则$f \pm g$ 和αf 也都是方程$(*5)$的解.

(3) 若$g(t)$ 是方程$(*5)$的解且$g(0) = f(0), g'(0) = f'(0), ..., g^{(n-1)}(0) = f^{(n-1)}(0)$, 则$g(t) \equiv f(t)$.

(4) 若上述多项式$P(z)$的所有根$\lambda_1, \lambda_2, ..., \lambda_n$ 互不相同, 方程$(*5)$的解都可表成$f(t) = \sum_{k=1}^{n} C_k e^{\lambda_k t}$, 其中$C_k$ 为常数.

此外对任意给定的初值$y_0, y_1, ..., y_{n-1} \in \mathbb{C}$, 方程$(*5)$存在唯一的解$f(t) = \sum_{k=1}^{n} C_k e^{\lambda_k t}$ 满足$f(0) = y_0, f'(0) = y_1, ..., f^{(n-1)}(0) = y_{n-1}$.

【证】前面已说明, 方程$(*5)$可写成$P(D)f(t) = 0$, $t \in \mathbb{R}$.

(1) 设$\lambda = \lambda$为$P(z)$的任一根, $f(t) = Ce^{\lambda t}$. 则由前面分析的结果知f 也是方程$(*5)$的一个解.

(2) 因导数运算是线性的, 即$D^k(f \pm g) = D^k f \pm D^k g$, $D^k(\alpha f) = \alpha D^k f$, 故易见性质(2)成立.

(3) 为证明方程$(*5)$的解关于初值条件的唯一性, 根据性质(2), 只需证明若$f(t)$是方程$(*5)$的解且$f(0) = 0, f'(0) = 0, ..., f^{(n-1)}(0) = 0$, 则必有$f(t) \equiv 0$. 我们对方程的阶数$n$ 用归纳法. 当$n = 1$ 时, 这是**命题5.51**的结论. 假设所证性质对

于n 阶方程成立. 则对$n+1$阶方程, 设f满足$P(D)f(t) := \sum_{k=0}^{n+1} a_k D^k f(t) \equiv 0$ (其中$a_{n+1}=1$) 且$f(0)=0, f'(0)=0, ..., f^{(n)}(0)=0$. 设$\lambda$ 为多项式$P(z)$的一个根. 作分解$P(z)=(z-\lambda)P_1(z)$. 令

$$F(t)=P_1(D)f(t).$$

则

$$(D-\lambda \mathrm{I})F(t)=(D-\lambda \mathrm{I})P_1(D)f(t)=P(D)f(t)=0, \quad t\in\mathbb{R}.$$

这表明F是一阶方程$(D-\lambda \mathrm{I})F(t)=0$ 的解. 而由f的初条件有

$$F(0)=P_1(D)f(0)=\sum_{k=0}^{n} b_k D^k f(0)=\sum_{k=0}^{n} b_k f^{(k)}(0)=0 \qquad (\text{其中 } b_n=1)$$

因此$F(t)\equiv 0$, 也即$P_1(D)f(t)=0$ for all $t\in\mathbb{R}$. 因P_1的次数为n, 故归纳假设和$f^{(k)}(0)=0, k=0,1,2,...,n$ 知$f(t)\equiv 0$. 据归纳法原理, 所证性质对任意阶方程成立.

(4) 首先由性质(3) 知对任意常数$C_1,C_2,...,C_n$, 函数$f(t)=\sum_{k=1}^{n} C_k e^{\lambda_k t}$ 是方程($*5$)的解. 此外易见

$$f^{(j)}(t)=C_1\lambda_1^j e^{\lambda_1 t}+C_2\lambda_2^j e^{\lambda_2 t}+\cdots+C_n\lambda_n^j e^{\lambda_n t}, \quad j=0,1,2,\dots.$$

设g是方程($*5$)的任一解. 考虑关于系数$(C_1,C_2,...,C_n)$的线性方程组

$$f(0)=C_1+C_2+\cdots+C_n=g(0),$$

$$f'(0)=C_1\lambda_1+C_2\lambda_2+\cdots+C_n\lambda_n=g'(0),$$

$$f''(0)=C_1\lambda_1^2+C_2\lambda_2^2+\cdots+C_n\lambda_n^2=g''(0),$$

....

$$f^{(n-1)}(0)=C_1\lambda_1^{n-1}+C_2\lambda_2^{n-1}+\cdots+C_n\lambda_n^{n-1}=g^{(n-1)}(0).$$

因$\lambda_1,\lambda_2,...,\lambda_n$ 互不相同, 故由线代数知这个线性方程组的系数矩阵可逆（即范德蒙矩阵可逆), 因此这方程组存在唯一解$(C_1,C_2,...,C_n)$. 由于相应的函数$f(t)=\sum_{k=1}^{n} C_k e^{\lambda_k t}$ 满足$f^{(k)}(0)=g^{(k)}(0), k=0,1,...,n-1$, 故根据性质(3) 知$g(t)\equiv f(t)$. 这证明了方程($*5$)的解都可表成$\sum_{k=1}^{n} C_k e^{\lambda_k t}$ 的形式. 用同样的论证可知(4)中的后一结论也成立. $\square$

【注】性质(3) 也称为解的叠加原理. 这个性质常用于寻求解的其它性质(见下面).

【二阶微分方程】在物理、力学中, 一阶和二阶微分方程最常见. 一阶方程已讨论过, 现在着重考察实系数二阶微分方程

$$f''(t) + bf'(t) + cf(t) = 0, \quad t \in \mathbb{R} \tag{$*6$}$$

和它的实解, 其中b, c 为实常数. 它对应的多项式为$P(z) = z^2 + bz + c$. 此时方程$P(z) = 0$的两个根为

$$\lambda_1 = \frac{-b + \sqrt{b^2 - 4c}}{2}, \quad \lambda_2 = \frac{-b - \sqrt{b^2 - 4c}}{2}.$$

分以下几种情形.

Case 1: $b^2 > 4c$. 此时λ_1, λ_2为不相等的实数. 方程$(*6)$的两个实解为

$$f_1(t) = C_1 e^{\lambda_1 t}, \quad f_2(t) = C_2 e^{\lambda_2 t}.$$

根据叠加原理, 方程$(*6)$的一般实解为

$$f(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}.$$

Case 2: $b^2 = 4c$ 即$\lambda_1 = \lambda_2 =: \lambda \in \mathbb{R}$. 此时不难验证方程$(*6)$的两个实解为

$$f_1(t) = C_1 e^{\lambda t}, \quad f_2(t) = C_2 t e^{\lambda t}.$$

方程$(*6)$的一般实解为

$$f(t) = e^{\lambda t}(C_1 + C_2 t).$$

Case 3: $b^2 < 4c$, 即λ_1, λ_2为共轭复数:

$$\lambda_1 = \alpha + \mathrm{i}\beta, \quad \lambda_2 = \alpha - \mathrm{i}\beta$$

其中α, β 为非零实数. 此时方程$(*6)$的两个特解为

$$f_1(t) = e^{(\alpha+\mathrm{i}\beta)t} = e^{\alpha t}(\cos\beta t + \mathrm{i}\sin\beta t), \quad f_2(t) = e^{(\alpha-\mathrm{i}\beta)t} = e^{\alpha t}(\cos\beta t - \mathrm{i}\sin\beta t).$$

根据解的叠加原理知

$$C_1 \frac{f_1(t) + f_2(t)}{2} = C_1 e^{\alpha t}\cos\beta t,$$

$$C_2 \frac{f_1(t) - f_2(t)}{2\mathrm{i}} = C_2 e^{\alpha t} \sin \beta t$$

是方程$(*6)$的两个实解. 再由叠加原理, 方程$(*6)$的一般实解为

$$f(t) = e^{\alpha t}\Big(C_1 \cos \beta t + C_2 \sin \beta t\Big).$$

Case 4: $b = 0, c > 0$. 此时$\lambda_1 = \mathrm{i}\beta, \lambda_2 = -\mathrm{i}\beta$ 其中$\beta = \sqrt{c} > 0$. 方程写为

$$f''(t) + \beta^2 f(t) = 0, \quad t \in \mathbb{R}. \tag{$*7$}$$

方程$(*7)$的两个特解为

$$f_1(t) = e^{\mathrm{i}\beta t} = \cos \beta t + \mathrm{i} \sin \beta t, \quad f_2(t) = e^{-\mathrm{i}\beta t} = \cos \beta t - \mathrm{i} \sin \beta t.$$

根据解的叠加原理知

$$C_1 \frac{f_1(t) + f_2(t)}{2} = C_1 \cos \beta t,$$
$$C_2 \frac{f_1(t) - f_2(t)}{2\mathrm{i}} = C_2 \sin \beta t$$

是方程$(*7)$的两个实解. 再由叠加原理知方程$(*7)$的一般实解为

$$f(t) = C_1 \cos \beta t + C_2 \sin \beta t.$$

§5.8. 自然科学中应用微分学的一些例子.

自从微积分建立以来, 人们对自然科学中的问题(如物理、力学、经济学中的) 就有了科学的方法. 这就是说, 一方面对早先只是靠直观、经验得到的粗糙结果现在有了理论解释和精确化, 包括不同程度的修正; 为了精确, 人们用导数代替局部平均, 用积分代替离散求和。另一方面现代微积分方法使我们能对更复杂的现象给予描述和解释并能预测. 这里所谓描述就是建立相关的数学模型, 这些数学模型的计算结果和预言与具体事物的实际测量和发展状态十分吻合. 微积分在建立数学模型中的主要作用是建立微分方程(或微分-积分方程). 为什么是微分方程? 原因很多：一个主要原因是人们在建模时经常要使用物理基本定律和力学基本原理, 这些定律和原理本身就是用微积分的语言描述的; 另一个主要原因是, 微分方程能够准确地描述大部分自然规律—— 事物的发展规律不仅与事物的属性有关，还与事物的变化率有关; 把事物自身和它的变化率(导数)通过某个合适的微分方程联系起来, 人们就获得了该事物自身的发展轨道. 进一步, 再把微分方程的解(具有给定初值或边界值)数值地计算出来, 人们就知道了事物发展的所有细节, 包括它的历史和未来状态.

下面三个例子取自卓里奇《数学分析》第一卷, pp.260-272.

【例：齐奥尔柯夫斯基公式】考察一个在开阔的宇宙中直线运动的火箭. 假设火箭远离任何具有引力的物体. 我们把火箭和助推燃料作为一个整体, 称之为火箭系统. 设$M(t), V(t)$分别是火箭系统在时刻t的质量和速度, ω (正常数) 是燃料在从火箭尾部管口爆燃喷出时的相对于火箭的速度. 我们将从动量守恒律来导出$M(t), V(t), \omega$ 三者的关系, 并做些设计上的分析.

对任一时刻t, 火箭系统在此刻的动量为$M(t)V(t)$. 让我们以t作为新阶段的初始时刻. 则在临近的下一时刻$t+h$ $(h>0)$, 火箭系统的动量等于两个动量之和: 第一个动量是$M(t+h)V(t+h)$, 它等于火箭自己在时刻$t+h$的动量, 即它不包含喷射的燃料的动量. 第二个动量是$|M(t+h)-M(t)|(V(t+h)+(-\omega)) = (M(t)-M(t+h))(V(t+h)-\omega)$, 它等于在新时刻$t+h$, 由火箭喷射时消耗的燃料$|M(t+h)-M(t)|$和这些消耗的燃料的速度$V(t+h)-\omega$构成的动量. 这里考虑$-\omega$是因为喷射速度的方向与火箭运行的方向相反. 根据动量守恒定律, 火箭系统在时刻$t+h$的动量等于火箭系统在新的初始时刻t的动量, 即

$$M(t+h)V(t+h) + (M(t)-M(t+h))(V(t+h)-\omega) = M(t)V(t).$$

化简后得到

$$(M(t+h)-M(t))\omega = -M(t)(V(t+h)-V(t)).$$

两边除以h 并令$h \to 0$ 得到

$$M'(t)\omega = -M(t)V'(t).$$

因为t是任意时刻, 于是得到微分方程

$$M'(t) = -\frac{1}{\omega}V'(t)M(t) \qquad \forall\, t > 0.$$

由**命题5.51** 知

$$M(t) = M(0)e^{-\frac{1}{\omega}(V(t)-V(0))}, \quad t \geq 0.$$

两边取对数并记$V_0 = V(0), M_0 = M(0)$ 得到

$$V(t) = V_0 + \omega \log\Big(\frac{M_0}{M(t)}\Big), \quad t \geq 0.$$

【**注**: 上面分析中说到动量守恒定律, 容易使人认为$M(t+h)V(t+h) \equiv M(t)V(t)$ 从而有$(M(t)V(t))' \equiv 0$. 这蕴含$M(t)V(t) \equiv M_0V_0$ 即$V(t) = \frac{M_0}{M(t)}V_0$. 当然这个结果明显不妥, 因为第一它没有用到喷射速度ω这一重要因素; 第二, 因(一般来说)初速度$V_0 = 0$, 故导致荒谬结果: $V(t) \equiv 0$. 产生这个问题的原因在于: 在上面的分析中, 两个动量$M(t)V(t)$ 与$M(t+h)V(t+h)$的局部意义不完全相同. 动量$M(t)V(t)$ 表示在时间段$[t, t+h]$上火箭系统在初始时刻t的总动量, 它是一个已经合成的结果. 而动量$M(t+h)V(t+h)$ 则是上面所说的意思, 即它只是火箭系统在新刻$t+h$的一部分动量, 另一部分动量, 也是最重要的动量, 是来自于喷射推进时产生的新动量$|M(t+h)-M(t)|(V(t+h)-\omega)$. 所以本题是根据这个分析运用动量守恒律的. 当然, 对于本题的这种分析方法, 在理解上确实比其它题目难一些. 但其实建模过程就是一个基本原理与实际情况相结合的过程, 其中专业经验和对物理力学基本定律的理解起很大作用, 这一点与纯粹的理论证明有些不同. **欢迎同学们继续讨论这个例题, 特别希望给出更加自圆其说的分析.**】

现在设m_k 是火箭自身的质量, m_T 是燃料的总质量, 而$V_\infty = V(t_\infty)$ 是燃料完全烧尽时火箭达到的终极速度. 那么把$M_0 = m_k + m_T$ 和$M(t_\infty) = m_k$ 代入上式就求得

$$V_\infty = V_0 + \omega \log\Big(1 + \frac{m_T}{m_k}\Big).$$

特别对于实际情形$V_0 = 0$ 有

$$V_\infty = \omega \log\Big(1 + \frac{m_T}{m_k}\Big).$$

从这一关系式还可知, 为使火箭的速度在有限时间内达到预订速度V, 也即为使$V_\infty \ge V$, 那么燃料的总质量m_T和喷射速度ω就应满足不等式

$$\omega \log\Big(1 + \frac{m_T}{m_k}\Big) \ge V.$$

但对数函数$\log(\cdot)$对自变量的放大作用很小, 因此设计的主要任务在于提高喷射速度ω. 当ω确定后, 燃料的总质量m_T 就需满足

$$m_T \ge m_k(e^{V/\omega} - 1)$$

换言之, 火箭在发射时必须至少装满数量为$m_k(e^{V/\omega} - 1)$的燃料. $\square$

【例：放射衰变、连锁反应及原子反应堆】本例详见卓里奇pp. 264-266, 其中遇到微分方程

$$N'(t) = \Big(\alpha - \frac{\beta}{r}\Big)N(t) + n, \qquad t \ge 0.$$

当$\alpha - \frac{\beta}{r} = 0$ 时, 解得$N(t) = N_0 + nt$. 当$\alpha - \frac{\beta}{r} \ne 0$ 时, 由**命题5.52**之(2) 知

$$N(t) = N_0 e^{(\alpha - \frac{\beta}{r})t} - \frac{n}{\alpha - \frac{\beta}{r}}\Big(1 - e^{(\alpha - \frac{\beta}{r})t}\Big), \quad t \ge 0.$$

$\square$

【例：振动】参见卓里奇pp. 270-273. 依据是牛顿定律

$$x''(t) = \frac{1}{m}F(t).$$

(1) 设物体的受力$F(t)$ 为弹力$F(t) = -kx(t)$, 这里$k > 0$ 是常数(即k 是弹性系数). 注意到当物体的位置$x(t) = 0$ 时, 其受力为零. 因此$x = 0$是弹性物体的平衡位置. 描述弹性物体运动的牛顿定律即为下列二阶常微分方程

$$x''(t) + \frac{k}{m}x(t) = 0, \quad t \in \mathbb{R}.$$

如令$a = \sqrt{k/m}$, 则这方程化为

$$x''(t) + a^2 x(t) = 0, \quad t \in \mathbb{R}.$$

根据上一节的分析结果知这方程的一般实解为

$$x(t)=C_1\cos at+C_2\sin at,\quad t\in\mathbb{R}$$

其中常数C_1,C_2 可由初值$x_0=x(0),v_0=x'(0)$ 确定：

$$x_0=x(0)=C_1,\quad v_0=x'(0)=C_2a,\quad \text{i.e.}\quad C_2=\frac{v_0}{a}.$$

因此

$$x(t)=x_0\cos at+\frac{v_0}{a}\sin at=\sqrt{x_0^2+\frac{v_0^2}{a^2}}\sin(at+\alpha)$$

即

$$x(t)=\sqrt{x_0^2+v_0^2\frac{m}{k}}\,\sin\Big(\sqrt{\frac{k}{m}}\,t+\alpha\Big),\quad t\in\mathbb{R}$$

其中

$$\alpha=\arcsin\Bigg(\frac{x_0}{\sqrt{x_0^2+v_0^2\frac{m}{k}}}\Bigg)\in[-\pi/2,\pi/2],\quad \text{它是简谐振动的初始相位角.}$$

(2) 设物体的受力$F(t)$与物体位移同向: $F(t)=kx(t)$, $k>0$ 是常数. 则物体的运动由二阶常微分方程描述:

$$x''(t)-\frac{k}{m}x(t)=0,\quad t\in\mathbb{R}.$$

当物体的位置$x(t)=0$ 时, 其受力为零. 故$x=0$是该物体的平衡位置.

令$a=\sqrt{k/m}$, 则这方程化为

$$x''(t)-a^2x(t)=0,\quad t\in\mathbb{R}.$$

因二次多项式方程$z-a^2=0$ 有两个不同实根$z=\pm a$, 故由上一节的分析结果知这方程的一般实解为

$$x(t)=C_1e^{at}+C_2e^{-at},\quad t\in\mathbb{R}.$$

其中常数C_1,C_2 可由初值$x_0=x(0),v_0=x'(0)$ 确定. 如果C_1,C_2 不全为零, 则易见

$$|x(t)|\to+\infty\quad \text{当 } t\to-\infty \quad \text{或}\ \ t\to+\infty\ \text{ 时.}$$

这表示: 要么物体的早期历史是远离平衡位置的, 要么物体的未来的发展是远离平衡位置的. 现在让我们只考虑物体的未来演化行为即只考虑$t\in[0,+\infty)$. 则易见当且仅当$C_1\neq 0$ 时, $|x(t)|\to+\infty\,(t\to+\infty)$. 即物体随时间趋于$+\infty$ 而远离平衡位置.

设$C_1=0$, 则$C_2\neq 0$. 此时$x(t)=C_2e^{-at}$. 因此$x(t)\to 0\,(t\to+\infty)$, 即物体随时间趋于$+\infty$而趋于平衡位置.

(3) 设物体的受力$F(t)$不仅与物体的位置$x(t)$有关还与物体的运动速度$x'(t)$有关, 例如设$F(t)=-\alpha x'(t)-kx(t)$, 其中α,k 为非零常数. 此时牛顿定律化为二阶微分方程:

$$x''(t)+\frac{\alpha}{m}x'(t)+\frac{k}{m}x(t)=0,\quad t\in\mathbb{R}. \tag{$*8$}$$

它对应的二次多项式$P(z)=z^2+\frac{\alpha}{m}z+\frac{k}{m}$的两根为

$$\lambda_1=-\frac{\alpha}{2m}+\frac{1}{2m}\sqrt{\alpha^2-4mk},\quad \lambda_2=-\frac{\alpha}{2m}-\frac{1}{2m}\sqrt{\alpha^2-4mk}.$$

以下讨论三种情形, 其中我们将使用上节的结果.

Case 1: $\alpha^2>4mk$. 此时λ_1,λ_2为不相等的实数. 此时方程$(*8)$的一般实解为

$$x(t)=C_1e^{\lambda_1 t}+C_2e^{\lambda_2 t}.$$

根据λ_1,λ_2 的正负, 不难给出当$t\to+\infty$时$x(t)$的渐近行为.

Case 2: $\alpha^2=4mk$ 即$\lambda_1=\lambda_2=-\frac{\alpha}{2m}$. 此时不难验证方程$(*8)$的一般实解为

$$x(t)=e^{-\frac{\alpha}{2m}t}(C_1+C_2t).$$

根据α 的正负, 不难给出当$t\to+\infty$时$x(t)$的渐近行为.

Case 3: $\alpha^2<4mk$. 此时λ_1,λ_2为共轭复数:

$$\lambda_1=-\frac{\alpha}{2m}+\mathrm{i}\frac{1}{2m}\sqrt{4mk-\alpha^2},\quad \lambda_2=-\frac{\alpha}{2m}-\mathrm{i}\frac{1}{2m}\sqrt{4mk-\alpha^2}.$$

此时方程$(*8)$的一般实解为

$$x(t)=e^{-\frac{\alpha}{2m}t}\Big(C_1\cos\omega t+C_2\sin\omega t\Big)$$

其中$\omega=\frac{1}{2m}\sqrt{4mk-\alpha^2}$. 如上, 这个一般解也可表示成

$$x(t)=Ae^{-\frac{\alpha}{2m}t}\sin(\omega t+\varphi)$$

其中常数$\varphi\in[-\pi/2,\pi/2], A>0$ 可由初值$x(0),x'(0)$ 确定.

因$\alpha\neq 0$, 故易见当且仅当$\alpha>0$ 时$\lim\limits_{t\to+\infty}x(t)=0$. 此时物体随时间趋于正无穷而趋于平衡位置.

但当$\alpha < 0$ 时, 与上面其它所有情况不同的是, 此时物体在时间趋于正无穷时是无界振动的:

$$\limsup_{t\to+\infty} x(t) = +\infty, \quad \liminf_{t\to+\infty} x(t) = -\infty.$$

□

§5.9. 原函数(不定积分)

【**定义**】设函数$f, F: I\to\mathbb{R}$. 若$F'(x)=f(x)$ for all $x\in I$, 也即$\mathrm{d}F(x)=f(x)\mathrm{d}x$ for all $x\in I$, 则称F 是f的一个原函数或f的一个不定积分, 记作

$$F(x)=\int f(x)\mathrm{d}x,\quad x\in I.\qquad\square$$

从定义看出, 若F是f的一个原函数, 则f就是F的导函数. 因此求原函数的运算与求导函数的运算几乎是互为逆运算, 需要注意的只是: 如果函数f有原函数, 则它有无穷多个原函数, 这是因为若F 是f的一个原函数, 则对任意常数C, 函数$F(t)+C$ 也是f的原函数. 当然若限死(例如) $F(0)=0$, 则常数C 就是唯一的, 从而可知求原函数与求导函数就确为互逆运算了.

【**记号$\int f(x)\mathrm{d}x$的来源和原函数的存在性**】

我们以$I=[a,b]$ 和f为连续函数为例来说明原函数$\int f(x)\mathrm{d}x$ 的存在性.

先设f 在$[a,b]$ 上恒大于零. 对任意$x\in[a,b]$, 考察子区间$[a,x]$. 从几何上看, 集合$\{(t,s)\,|\,t\in[a,x],\ 0\le s\le f(t)\}$ 是一个曲边梯形. 我们相信这个曲边梯形有面积. 记这面积为(图示)

$$\int_a^x f(t)\mathrm{d}t,\quad x\in(a,b].$$

常识告诉我们, 面积具有可加性, 即若$x<y$ 则

$$\int_a^x f(t)\mathrm{d}t+\int_x^y f(t)\mathrm{d}t=\int_a^y f(t)\mathrm{d}t.$$

如用可减性, 则上式等价地写为

$$\int_a^y f(t)\mathrm{d}t-\int_a^x f(t)\mathrm{d}t=\int_x^y f(t)\mathrm{d}t.$$

如将面积函数记为

$$F(x)=\int_a^x f(t)\mathrm{d}t,\quad x\in(a,b]\tag{$*$}$$

则上面的减法即为

$$F(y)-F(x)=\int_x^y f(t)\mathrm{d}t.\tag{$**$}$$

现在固定x. 根据f的连续性, 对任意$\varepsilon>0$, 存在$\delta>0$ 使得当$t\in[x,x+\delta]$ 时$f(x)-\varepsilon<f(t)<f(x)+\varepsilon$. 从集合的角度看, 这蕴含当$y\in(x,x+\delta]$ 时

$$[x,y]\times[0,f(x)-\varepsilon]\subset\{(t,s)\,|\,t\in[x,y],\ 0\le s\le f(t)\}\subset[x,y]\times[0,f(x)+\varepsilon].$$

常识告诉我们: 小的集合其面积也小, 故得到

$$(y-x)(f(x)-\varepsilon) \le \int_x^y f(t)\mathrm{d}t \le (y-x)(f(x)+\varepsilon).$$

两边除以$y-x$ 并注意关系式$(**)$ 得到

$$f(x)-\varepsilon \le \frac{1}{y-x}\int_x^y f(t)\mathrm{d}t = \frac{F(y)-F(x)}{y-x} \le f(x)+\varepsilon.$$

也即我们得到

$$\Big|\frac{F(y)-F(x)}{y-x} - f(x)\Big| \le \varepsilon.$$

根据导数的定义, 这证明了

$$F'_+(x) = f(x).$$

同理对于$x \in (a,b]$, 让$y \in [x-\delta, x)$ 也可证得

$$F'_-(x) = f(x).$$

所以F 在任一点$x \in [a,b]$ 可微且$F'(x) = f(x)$. 因此F 是f的一个原函数.

对于f不恒大于零的情形, 取$M > 0$ 充分大使得$f(x)+M > 0$ for all $x \in [a,b]$. 则$f(x)+M$ 有原函数. 设$F_1(x)$ 是$f(x)+M$的一个原函数. 令$F(x) = F_1(x) - Mx$, 则F 可微且

$$F'(x) = F_1'(x) - M = f(x), \quad x \in [a,b]$$

这说明此时f仍有原函数.

回到记号$(*)$: 我们已说明了

$$F(x) = \int_a^x f(t)\mathrm{d}t, \quad x \in (a,b]$$

是f的一个原函数. 由于这个原函数通过变动积分上限x定义的, 故我们也称$F(x)$是f的一个不定积分(即不限定积分上限的积分). 由于这个不定积分的导数与积分下限a 无关, 故为了简单且突出实质, 人们就将上述不定积分记作

$$F(x) = \int f(x)\mathrm{d}x.$$

这个记号当然不是很好因为从上面的分析可知左边的x 与右边的x 意义不同, 但因没有更合适的简单记号(简明与详细总是对立的!), 人们就接受了它. 不仅接受, 而且在进一步的运算中, 例如在做变量替换时, 大家发现, 记号$\int f(x)\mathrm{d}x$ 竟有很多优点. 见下

面.

【不定积分的基本性质和计算方法】

【性质1：唯一性(不计常数)】$F_1(x), F_2(x)$ 都是$f(x)$的不定积分(原函数) 当且仅当$F_2(x) = F_1(x) + C$ 其中C 为常数.

【性质2: 线性性】设α, β 为常数. 则

$$\int(\alpha f(x) + \beta g(x))\mathrm{d}x = \alpha\int f(x)\mathrm{d}x + \beta\int g(x)\mathrm{d}x.$$

【证】由求导运算的线性性和原函数(不定积分)的定义有

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha\int f(x)\mathrm{d}x + \beta\int g(x)\mathrm{d}x\right) = \alpha\frac{\mathrm{d}}{\mathrm{d}x}\int f(x)\mathrm{d}x + \beta\frac{\mathrm{d}}{\mathrm{d}x}\int g(x)\mathrm{d}x = \alpha f(x) + \beta g(x).$$

这表明$\alpha\int f(x)\mathrm{d}x + \beta\int g(x)\mathrm{d}x$ 是$\alpha f(x) + \beta g(x)$的原函数. 再根据原函数(不定积分)的定义即知所证成立. $\square$

【性质3: 分部积分】

$$\int u(x)v'(x)\mathrm{d}x = u(x)v(x) - \int v(x)u'(x)\mathrm{d}x.$$

或写成便于理解的形式:

$$\int u(x)\mathrm{d}v(x) = u(x)v(x) - \int v(x)\mathrm{d}u(x).$$

【证】由原函数的定义和不定积分的线性性有

$$u(x)v(x) = \int(u(x)v(x))'\mathrm{d}x = \int(u'(x)v(x)+u(x)v'(x))\mathrm{d}x = \int u'(x)v(x)\mathrm{d}x+\int u(x)v'(x)\mathrm{d}x.$$

移项即得分部积分公式. $\square$

【性质4: 凑微分法】设

$$F(x) = \int f(x)\mathrm{d}x$$

并设$\varphi(t)$ 是区间I上的可微函数且$\varphi(t)$ 属于f的定义域. 则

$$\int f(\varphi(t))\varphi'(t)\mathrm{d}t = F(\varphi(t)).$$

【证】由复合函数求导有

$$(F(\varphi(t)))' = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

所以$F(\varphi(t))$ 是$f(\varphi(t))\varphi'(t)$ 的一个原函数. $\square$

【性质5: 换元公式】设$x = \varphi(t) \in J$ 是区间I的严格单调的可微函数且反函数$t = \varphi^{-1}(x)$ 在J 上可微. 设$f(x)$ 在J 上有原函数. 设$F(t)$ 是$f(\varphi(t))\varphi'(t)$的原函数, 即

$$F(t) = \int f(\varphi(t))\varphi'(t)\mathrm{d}t.$$

则

$$\int f(x)\mathrm{d}x = F(\varphi^{-1}(x)).$$

【证】证法1: 直接验证:

$$(F(\varphi^{-1}(x)))' = F'(t)\Big|_{t=\varphi(x)}(\varphi^{-1})'(x) = f(\varphi(t))\varphi'(t)\Big|_{t=\varphi^{-1}(x)}\frac{1}{\varphi'(t)}\Big|_{t=\varphi^{-1}(x)} = f(x).$$

证法2: **体现换元的意义**: 做换元$x = \varphi(t)$, 则$\mathrm{d}x = \varphi'(t)\mathrm{d}t$ 从而有

$$\int f(x)\mathrm{d}x = \int f(\varphi(t))\mathrm{d}\varphi(t) = \int f(\varphi(t))\varphi'(t)\mathrm{d}t = F(t) = F(\varphi^{-1}(x)). \qquad \square$$

♠ 一些例子

1. 凑微分法和换元法

$$\int te^{t^2}\mathrm{d}t = \int e^{t^2}\frac{1}{2}\mathrm{d}t^2 = \frac{1}{2}\int e^x\mathrm{d}x\Big|_{x=t^2} = C + \frac{1}{2}e^x\Big|_{x=t^2} = C + \frac{1}{2}e^{t^2}.$$

$$\int \frac{t\mathrm{d}t}{1+t^2}\mathrm{d}t = \frac{1}{2}\int \frac{\mathrm{d}t^2}{1+t^2} = \frac{1}{2}\int \frac{\mathrm{d}x}{1+x}\Big|_{x=t^2} = C + \frac{1}{2}\log(1+x)\Big|_{x=t^2} = C + \frac{1}{2}\log(1+t^2).$$

$$\begin{aligned}
&\int \tan^4 x\mathrm{d}x = \int \tan^4(\arctan t)\mathrm{d}(\arctan t) = \int \frac{t^4}{1+t^2}\mathrm{d}t \\
&= \int \frac{t^4-1}{1+t^2}\mathrm{d}t + \int \frac{1}{1+t^2}\mathrm{d}t = \int (t^2-1)\mathrm{d}t + \arctan t \\
&= C + \frac{1}{3}t^3 - t + \arctan t = C + \frac{1}{3}\tan^3 x - \tan x + x.
\end{aligned}$$

2. 指数函数的不定积分和相关的不定积分.

周知若$0 \neq \lambda \in \mathbb{C}$ 为复常数, 则

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{\lambda}e^{\lambda x}\Big) = e^{\lambda x}$$

意即

$$\int e^{\lambda x}\mathrm{d}x = C + \frac{1}{\lambda}e^{\lambda x}.$$

现在取$\lambda = \alpha + \mathrm{i}\beta$ 其中α, β 为非零实数. 则由

$$e^{(\alpha+\mathrm{i}\beta)x} = e^{\alpha x}(\cos\beta x + \mathrm{i}\sin\beta x) = e^{\alpha x}\cos\beta x + \mathrm{i}e^{\alpha x}\sin\beta x$$

和不定积分的线性性得到(推导中不考虑常数)

$$\begin{aligned}
&\int e^{\alpha x}\cos\beta x\mathrm{d}x + \mathrm{i}\int e^{\alpha x}\sin\beta x\mathrm{d}x\\
&= \int e^{(\alpha+\mathrm{i}\beta)x}\mathrm{d}x = \frac{1}{\alpha+\mathrm{i}\beta}e^{(\alpha+\mathrm{i}\beta)x} = \frac{1}{\alpha+\mathrm{i}\beta}\Big(e^{\alpha x}\cos\beta x + \mathrm{i}e^{\alpha x}\sin\beta x\Big)\\
&= \frac{\alpha-\mathrm{i}\beta}{\alpha^2+\beta^2}\Big(e^{\alpha x}\cos\beta x + \mathrm{i}e^{\alpha x}\sin\beta x\Big) = \frac{e^{\alpha x}}{\alpha^2+\beta^2}(\alpha-\mathrm{i}\beta)(\cos\beta x + \mathrm{i}\sin\beta x)\\
&= \frac{e^{\alpha x}}{\alpha^2+\beta^2}\Big(\alpha\cos\beta x + \beta\sin\beta x + \mathrm{i}(-\beta\cos\beta x + \alpha\sin\beta x)\Big).
\end{aligned}$$

比较实部、虚部得到

$$\int e^{\alpha x}\cos\beta x\mathrm{d}x = \frac{e^{\alpha x}(\alpha\cos\beta x + \beta\sin\beta x)}{\alpha^2+\beta^2} + C,$$
$$\int e^{\alpha x}\sin\beta x\mathrm{d}x = \frac{e^{\alpha x}(-\beta\cos\beta x + \alpha\sin\beta x)}{\alpha^2+\beta^2} + C.$$

2. 分部积分法

$$\int \log x\mathrm{d}x = (\log x)x - \int x\mathrm{d}\log x = x\log x - \int\frac{x}{x}\mathrm{d}x = x\log x - \int\mathrm{d}x = x\log x - x + C.$$

下面用分部积分再次计算$\int e^{\alpha x}\cos\beta x\mathrm{d}x$.

$$\begin{aligned}
I(x) &:= \int e^{\alpha x}\cos\beta x\mathrm{d}x = \int e^{\alpha x}\frac{1}{\beta}\mathrm{d}(\sin\beta x) = \frac{1}{\beta}e^{\alpha x}\sin\beta x - \frac{1}{\beta}\int\sin\beta x\mathrm{d}e^{\alpha x}\\
&= \frac{1}{\beta}e^{\alpha x}\sin\beta x - \frac{1}{\beta}\int(\sin\beta x)\alpha e^{\alpha x}\mathrm{d}x = \frac{1}{\beta}e^{\alpha x}\sin\beta x - \frac{\alpha}{\beta}\int e^{\alpha x}\sin\beta x\mathrm{d}x.
\end{aligned}$$

类似地,

$$\int e^{\alpha x}\sin\beta x\mathrm{d}x=\int e^{\alpha x}\frac{-1}{\beta}\mathrm{d}(\cos\beta x)=-\frac{1}{\beta}e^{\alpha x}\cos\beta x+\frac{\alpha}{\beta}\int e^{\alpha x}\cos\beta x\mathrm{d}x.$$

所以$\Longrightarrow$

$$\begin{aligned}
&I(x)=\frac{1}{\beta}e^{\alpha x}\sin\beta x+\frac{\alpha}{\beta^2}e^{\alpha x}\cos\beta x-\frac{\alpha^2}{\beta^2}I(x),\\
&\Big(1+\frac{\alpha^2}{\beta^2}\Big)I(x)=e^{\alpha x}\Big(\frac{1}{\beta}\sin\beta x+\frac{\alpha}{\beta^2}\cos\beta x\Big),\\
&I(x)=\frac{e^{\alpha x}}{\alpha^2+\beta^2}\beta^2\Big(\frac{1}{\beta}\sin\beta x+\frac{\alpha}{\beta^2}\cos\beta x\Big)=\frac{e^{\alpha x}}{\alpha^2+\beta^2}\Big(\beta\sin\beta x+\alpha\cos\beta x\Big).
\end{aligned}$$

3. 最简真分式的不定积分. 这里给出基本递推关系. 以下设n为正整数.

(1)

$$\int\frac{\mathrm{d}x}{x^2+1}=\arctan x+C,$$

$$\int\frac{\mathrm{d}x}{(x^2+1)^n}=\frac{1}{2(n-1)}\cdot\frac{x}{(x^2+1)^{n-1}}+\frac{2n-3}{2(n-1)}\int\frac{\mathrm{d}x}{(x^2+1)^{n-1}},\quad n\ge 2.$$

【证】由分部积分有

$$\begin{aligned}
&I_{n-1}(x):=\int(x^2+1)^{1-n}\mathrm{d}x=(x^2+1)^{1-n}x-\int(1-n)(x^2+1)^{-n}2x^2\mathrm{d}x\\
&=\frac{x}{(x^2+1)^{n-1}}+2(n-1)\int\frac{x^2}{(x^2+1)^n}\mathrm{d}x\\
&=\frac{x}{(x^2+1)^{n-1}}+2(n-1)\int\frac{x^2+1}{(x^2+1)^n}\mathrm{d}x-2(n-1)\int\frac{1}{(x^2+1)^n}\mathrm{d}x\\
&=\frac{x}{(x^2+1)^{n-1}}+2(n-1)I_{n-1}(x)-2(n-1)I_n(x).
\end{aligned}$$

所以

$$2(n-1)I_n(x)=\frac{x}{(x^2+1)^{n-1}}+(2n-3)I_{n-1}(x).$$

(2) 设$p^2<4q$. 则

$$\int\frac{\mathrm{d}x}{x^2+px+q}=\frac{2}{\sqrt{4q-p^2}}\arctan\Big(\frac{2x+p}{\sqrt{4q-p^2}}\Big)+C,$$

$$\begin{aligned}
&\int \frac{\mathrm{d}x}{(x^2+px+q)^n}\\
&= \frac{1}{(n-1)(4q-p^2)}\cdot\frac{2x+p}{(x^2+px+q)^{n-1}}+\frac{2(2n-3)}{(n-1)(4q-p^2)}\int\frac{\mathrm{d}x}{(x^2+px+q)^{n-1}},\quad n\geq 2.
\end{aligned}$$

从而有

$$\int \frac{\mathrm{d}x}{(x^2+px+q)^n}=\frac{P(x)}{(x^2+px+q)^{n-1}}+A\arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right)+C,\quad n\geq 2$$

其中$\deg P\leq 2n-3$; A, C 是常数.

【证】令$a=-\frac{1}{2}p, b=\frac{1}{2}\sqrt{4q-p^2}$. 则$-2a=p, a^2+b^2=q$,

$$x^2+px+q=(x-a)^2+b^2.$$

做变量替换$x=a+bt$, $t=\frac{x-a}{b}$, 我们计算

$$\begin{aligned}
&I_n(x):=\int\frac{\mathrm{d}x}{(x^2+px+q)^n}=\int\frac{\mathrm{d}x}{[(x-a)^2+b^2]^n}\\
&=\frac{1}{b^{2n}}\int\frac{\mathrm{d}x}{[(\frac{x-a}{b})^2+1]^n}=\frac{1}{b^{2n}}\int\frac{b\mathrm{d}t}{(t^2+1)^n}\\
&=\frac{1}{b^{2n-1}}\int\frac{\mathrm{d}t}{(t^2+1)^n}.
\end{aligned}$$

若$n=1$, 则

$$I_1(x)=\frac{1}{b}\int\frac{\mathrm{d}t}{t^2+1}=\frac{1}{b}\arctan t=\frac{1}{b}\arctan(\frac{x-a}{b})=\frac{2}{\sqrt{4q-p^2}}\arctan(\frac{2x+p}{\sqrt{4q-p^2}})+C.$$

设$n\geq 2$. 则

$$\begin{aligned}
&I_n(x)=\frac{1}{b^{2n-1}}\int\frac{\mathrm{d}t}{(t^2+1)^n}\\
&=\frac{1}{b^{2n-1}}\frac{1}{2(n-1)}\cdot\frac{t}{(t^2+1)^{n-1}}+\frac{1}{b^{2n-1}}\frac{2n-3}{2(n-1)}\int\frac{\mathrm{d}t}{(t^2+1)^{n-1}}\\
&=\frac{1}{b^{2n-1}}\frac{1}{2(n-1)}\cdot\frac{\frac{x-a}{b}}{[(\frac{x-a}{b})^2+1]^{n-1}}+\frac{1}{b^{2n-1}}\frac{2n-3}{2(n-1)}b^{2(n-1)-1}I_{n-1}(x)\\
&=\frac{1}{2(n-1)b^2}\cdot\frac{x-a}{[(x-a)^2+b^2]^{n-1}}+\frac{1}{b^2}\cdot\frac{2n-3}{2(n-1)}I_{n-1}(x)\\
&=\frac{2}{(n-1)(4q-p^2)}\cdot\frac{x+p/2}{(x^2+px+q)^{n-1}}+\frac{1}{b^2}\cdot\frac{2n-3}{2(n-1)}I_{n-1}(x)\\
&=\frac{1}{(n-1)(4q-p^2)}\cdot\frac{2x+p}{(x^2+px+q)^{n-1}}+\frac{2(2n-3)}{(n-1)(4q-p^2)}I_{n-1}(x).
\end{aligned}$$

(3) 一般情形: 设$a,b,p,q\in\mathbb{R}, p^2<4q$. 则

$$\int\frac{ax+b}{x^2+px+q}\mathrm{d}x=\frac{a}{2}\log(x^2+px+q)+\frac{2b-ap}{\sqrt{4q-p^2}}\arctan\Big(\frac{2x+p}{\sqrt{4q-p^2}}\Big)+C,$$

$$\int\frac{ax+b}{(x^2+px+q)^n}\mathrm{d}x=-\frac{a}{2(n-1)(x^2+px+q)^{n-1}}+(b-ap/2)\int\frac{1}{(x^2+px+q)^n}\mathrm{d}x,\quad n\ge 2$$

从而有

$$\int\frac{ax+b}{(x^2+px+q)^n}\mathrm{d}x=\frac{P(x)}{(x^2+px+q)^{n-1}}+A\arctan\Big(\frac{2x+p}{\sqrt{4q-p^2}}\Big)+C,\quad n\ge 2$$

其中$\mathrm{deg}P\le 2n-3$, A,C 是常数.

【证】

$$\begin{aligned}&\int\frac{ax+b}{x^2+px+q}\mathrm{d}x=\int\frac{\frac{a}{2}(2x+p)+b-ap/2}{x^2+px+q}\mathrm{d}x\\&=\frac{a}{2}\int\frac{2x+p}{x^2+px+q}\mathrm{d}x+(b-ap/2)\int\frac{1}{x^2+px+q}\mathrm{d}x\\&=\frac{a}{2}\log(x^2+px+q)+\frac{2b-ap}{\sqrt{4q-p^2}}\arctan\Big(\frac{2x+p}{\sqrt{4q-p^2}}\Big)+C,\end{aligned}$$

$$\begin{aligned}&\int\frac{ax+b}{(x^2+px+q)^n}\mathrm{d}x=\int\frac{\frac{a}{2}(2x+p)+b-ap/2}{(x^2+px+q)^n}\mathrm{d}x\\&=\frac{a}{2}\int\frac{2x+p}{(x^2+px+q)^n}\mathrm{d}x+(b-ap/2)\int\frac{1}{(x^2+px+q)^n}\mathrm{d}x\\&=-\frac{a}{2(n-1)(x^2+px+q)^{n-1}}+(b-ap/2)\int\frac{1}{(x^2+px+q)^n}\mathrm{d}x.\qquad\Box\end{aligned}$$

4. 真分式的不定积分. 真分式的原函数等于一个低阶真分式与一些可能的对数函数和反正切函数之和. 具体来说, 设 $P(x)/Q(x)$ 为实真分式. 不失一般性设

$$Q(x)=\prod_{i=1}^{r}(x-\alpha_i)^{k_i}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{m_j}$$

其中 $\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{R}$ 互不相同, $(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$ 互不相同, 满足 $p_j^2<4q_j$ $(j=1,2,...,s)$, k_i,m_j 为正整数 $(i=1,2,...,r, j=1,2,...,s)$. 则

$$\int\frac{P(x)}{Q(x)}\mathrm{d}x=\frac{P_1(x)}{Q_1(x)}+H(x)$$

其中

$$Q_1(x)=\prod_{i=1}^{r}(x-\alpha_i)^{k_i-1}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{m_j-1},\quad \deg P_1<\deg Q_1,$$

$$H(x)=\sum_{i=1}^{r}A_i\log|x-\alpha_i|+\sum_{j=1}^{s}B_j\log(x^2+p_jx+q_j)+\sum_{j=1}^{s}C_j\arctan\Big(\frac{2x+p_j}{\sqrt{4q_j-p_j^2}}\Big)+常数,$$

A_i,B_j,C_j 为实常数. 函数 $H(x)$ 也称为超越函数 (即指 $H(x)$ 经自变量 x的任何有理变换都不能化成有理函数).

【证】由真分式分解定理有

$$\begin{aligned}\frac{P(x)}{Q(x)}&=\sum_{i=1}^{r}\sum_{k=1}^{k_i}\frac{A_k^{(i)}}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{m=1}^{m_j}\frac{B_m^{(j)}x+C_m^{(j)}}{(x^2+p_jx+q_j)^m}\\&=\sum_{i=1}^{r}\frac{A_1^{(i)}}{x-\alpha_i}+\sum_{j=1}^{s}\frac{B_1^{(j)}x+C_1^{(j)}}{x^2+p_jx+q_j}\\&+\sum_{i=1}^{r}\sum_{2\le k\le k_i}\frac{A_k^{(i)}}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{2\le m\le m_j}\frac{B_m^{(j)}x+C_m^{(j)}}{(x^2+p_jx+q_j)^m}.\end{aligned}$$

两边取不定积分, 利用不定积分的线性性和上面计算结果便有

$$\begin{aligned}&\int\frac{P(x)}{Q(x)}\mathrm{d}x\\&=\sum_{i=1}^{r}A_1^{(i)}\int\frac{1}{x-\alpha_i}\mathrm{d}x+\sum_{j=1}^{s}\int\frac{B_1^{(j)}x+C_1^{(j)}}{x^2+p_jx+q_j}\mathrm{d}x\\&+\sum_{i=1}^{r}\sum_{2\le k\le k_i}A_k^{(i)}\int\frac{1}{(x-\alpha_i)^k}\mathrm{d}x+\sum_{j=1}^{s}\sum_{2\le m\le m_j}\int\frac{B_m^{(j)}x+C_m^{(j)}}{(x^2+p_jx+q_j)^m}\mathrm{d}x\\&=\sum_{i=1}^{r}A_1^{(i)}\log|x-\alpha_i|+\sum_{j=1}^{s}\left(b_j\log(x^2+p_jx+q_j)+c_j\arctan\Big(\frac{2x+p_j}{\sqrt{4q_j-p_j^2}}\Big)\right)\\&+\sum_{i=1}^{r}\sum_{2\le k\le k_i}\frac{A_k^{(i)}}{1-k}\cdot\frac{1}{(x-\alpha_i)^{k-1}}+\sum_{j=1}^{s}\sum_{2\le m\le m_j}\left(\frac{P_{m,j}(x)}{(x^2+p_jx+q_j)^{m-1}}+c_{m,j}\arctan\Big(\frac{2x+p_j}{\sqrt{4q_j-p_j^2}}\Big)\right)\\&+常数\\&=\frac{P_1(x)}{Q_1(x)}+H(x).\qquad\square\end{aligned}$$

<u>**奥斯特洛格拉德斯基方法**</u>. 对真分式的不定积分公式两边求导得到

$$\frac{P(x)}{Q(x)}=\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{P_1(x)}{Q_1(x)}\Big)+\frac{\mathrm{d}}{\mathrm{d}x}H(x).$$

而根据 $H(x)$ 的表达式和前面已证明的最简分式的不定积分有

$$\frac{\mathrm{d}}{\mathrm{d}x}H(x) = \frac{P_2(x)}{Q_2(x)}$$

其中

$$Q_2(x) = \prod_{i=1}^{r}(x-\alpha_i)\prod_{j=1}^{s}(x^2+p_jx+q_j), \qquad \deg P_2 < \deg Q_2. \tag{q}$$

于是真分式 $P(x)/Q(x)$ 可分解为

$$\frac{P(x)}{Q(x)} = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{P_1(x)}{Q_1(x)}\Big) + \frac{P_2(x)}{Q_2(x)}. \tag{$*$}$$

注意：由 $Q(x), Q_1(x), Q_2(x)$ 的根式表达式易见 $Q_1(x)$ 是 $Q(x)$ 与 $Q'(x)$ 的最大公因式, 且

$$Q_1(x)Q_2(x) = Q(x).$$

公式 $(*)$ 也称为奥斯特洛格拉德斯基公式. 可以想象, 如果不用不定积分方法, 则很难证明分解式 $(*)$. 在实际应用中, 假如 $Q(x)$ 的上述根式分解式已经知道, 则多项式 $P_1(x), P_2(x)$ 可以通过 $(*)$ 和待定系数法确定, 这是多数教科书中的做法. 但有些教科书如卓里奇的《数学分析》则指出了应用奥斯特洛格拉德斯基公式 $(*)$ 的一个永远可行的方法：虽然一般来说很难给出 $Q(x)$ 的根式分解从而无法直接得到 $Q_1(x)$, 但因为 $Q_1(x)$ 是 $Q(x)$ 与 $Q'(x)$ 的最大公因式, 故可以应用欧几里德算法得到 $Q_1(x)$. 然后对恒等式 $Q_1(x)Q_2(x) = Q(x)$ 用待定系数法可以得到 $Q_2(x)$. 这样一来, $Q_1(x), Q_2(x)$ 就都确定下来了. 余下的工作同上, 即通过 $(*)$ 和待定系数法确定出 $P_1(x), P_2(x)$. 这种方法虽然麻烦, 但与高次多项式求根这个几乎不可行的方法相比, 毕竟是可行的！公式 $(*)$ 与上述待定系数法等统称为奥斯特洛格拉德斯基方法.

回到真分式的不定积分, 我们将用公式 $(*)$ 再次证明真分式的不定积分定理: 一般情况下, 超越部分 $H(x)$ 不可避免地要出现, 并且为了得到 $H(x)$ 的表达式, 我们也不得不使用多项式基本定理. 注意到 $Q_1(x)$ 是 $Q(x)$ 与 $Q'(x)$ 的最大公因式, 易见多项式 $Q_2(x)$ 只有单根而无重根, 也即 $Q_2(x)$ 具有 (q) 中的分解式并且 $Q_2(x)$ 的所有根就是 $Q(x)$ 的所有根. **但 $Q_2(x)$ 的主要优点在于它的次数比 $Q(x)$ 的低**—— 除非 $Q(x)$ 无重根 (此时 $Q_2(x) \equiv Q(x), P(x) \equiv P_2(x), Q_1(x) \equiv 1, P_1(x) \equiv 0$). 例如若 $Q_2(x)$ 的次

数 ≤ 4, 则我们可以具体写出 $Q_2(x)$ 的根式分解 (q). 因 $P_2(x)/Q_2(x)$ 是真分式, 故由真分式按最简分式的分解定理有

$$\frac{P_2(x)}{Q_2(x)} \equiv \sum_{i=1}^{r} \frac{A^{(i)}}{x-\alpha_i} + \sum_{j=1}^{s} \frac{B^{(j)}x + C^{(j)}}{x^2 + p_j x + q_j}.$$

两边取不定积分并应用上面最简分式的不定积分的结果我们再次得到 $\int \frac{P(x)}{Q(x)}\mathrm{d}x$ 的超越部分 $H(x)$ 的表达式:

$$\begin{aligned} H(x) &= \int \frac{P_2(x)}{Q_2(x)}\mathrm{d}x \\ &= \sum_{i=1}^{r} A_i \log|x-\alpha_i| + \sum_{j=1}^{s} B_j \log(x^2 + p_j x + q_j) + \sum_{j=1}^{s} C_j \arctan\left(\frac{2x+p_j}{\sqrt{4q_j - p_j^2}}\right) + \text{常数}. \end{aligned}$$

5. 二元有理函数和相关的不定积分.

称形如

$$P(x,y) = \sum_{k=0}^{n} \sum_{i+j=k} a_{i,j} x^i y^j, \quad (x,y) \in \mathbb{R}^2$$

为一个次数不超过n的二元多项式. 当最高项系数a_{i+j} $(i+j=n)$ 不全为零时, 称$P(x,y)$ 为一个n 次二元多项式.

称两个二元多项式$P(x,y), Q(x,y)$ 的商

$$R(x,y) := \frac{P(x,y)}{Q(x,y)} = \frac{\sum_{k=0}^{n} \sum_{i+j=k} a_{i,j} x^i y^j}{\sum_{k=0}^{m} \sum_{i+j=k} b_{i,j} x^i y^j}, \quad (x,y) \in \mathbb{R}^2$$

为一个二元有理函数.

例如

$$R(x,y) = \frac{x^2y^3 + xy^2 + y + 1}{x^4y^2 + x^3y^2 + x - 2}$$

是一个二元有理函数.

对于一个二元有理函数$R(x,y) = P(x,y)/Q(x,y)$, 若令$x = p(t), y = q(t)$, 其中$p(t), q(t)$ 都是t的多项式, 则易见$P(p(t),q(t)), Q(p(t),q(t))$ 都是t的多项式从而$R(p(t),q(t))$ 是t的有理函数.

例如

$$R(t^2, 1-t^3) = \frac{x^2y^3 + xy^2 + y + 1}{x^4y^2 + x^3y^2 + x - 2}\Big|_{x=t^2, y=1-t^3} = \frac{t^4(1-t^3)^3 + t^2(1-t^3)^2 - t^3}{t^8(1-t^3)^2 + t^6(1-t^3)^2 + t^2 - 2}$$

是t的有理函数.

我们此处引进二元有理函数的目的是为了算出如下形式的原函数(不定积分):

三角函数类

$$\int R(\cos x, \sin x)\mathrm{d}x, \quad \int R(\cos^2 x, \sin^2 x)\mathrm{d}x, \quad \int R(\cos x, \sin^2 x)\sin x\mathrm{d}x, \quad \text{etc.}$$

无理函数类:

$$\int R\Big(x, \Big(\frac{ax+b}{cx+d}\Big)^{1/n}\Big)\mathrm{d}x, \quad \int R(x, \sqrt{ax^2+bx+c}\,)\mathrm{d}x, \quad \int R(x, \sqrt{P(x)}\,)\mathrm{d}x$$

其中$P(x)$ 是次数≥ 3的多项式.

这些不定积分的计算通常需要寻求合适的变量替换. 具体的分类操作和算例参见陈书第6章, 卓里奇书第一卷pp. 287-292; 徐森林、薛春华第一册pp.295-325. (徐书很nice!)

通过不定积分的计算, 我们要学习和培养

<u>观察、变形的能力</u>, <u>变量替换的观念</u>, <u>多步计算能力</u> (特别是在运用分部积分导出某些关系时). 同时须熟记一些最基本的不定积分的结果.

我们用一个综合算列结束本节:

【**例**】求不定积分

$$I_{p,q}(x) = \int \frac{\mathrm{d}x}{x\sqrt{x^2+px+q}} \qquad \text{其中} \;\; p^2 > 4q.$$

【**解**】做反射变换$x=-y$ 有

$$I_{p,q}(x) = I_{p,q}(-y) = \int \frac{\mathrm{d}(-y)}{-y\sqrt{y^2-py+q}} = \int \frac{\mathrm{d}y}{y\sqrt{y^2-py+q}} = I_{-p,q}(y).$$

再以$y=-x$ 带回得到关系式

$$I_{p,q}(x) = I_{-p,q}(-x).$$

这说明只要$I_{p,q}(x), I_{-p,q}(x)$有一个求出, 则另一个也就求出. 详细来说, 若$p<0$, 则先计算$I_{-p,q}(x)$. 得到结果后再把x 换成$-x$ 就得到$I_{p,q}(x) = I_{-p,q}(-x)$的表达式.

因此我们只需对$p>0$情形计算即可. 但是最后的结果必须满足关系式$I_{p,q}(x) = I_{-p,q}(-x)$. 这就为什么在下面的计算中我们要保持因子$px$ 或x/p. 事实上关系式$I_{p,q}(x) = I_{-p,q}(-x)$ 得以成立就是由于px, x/p具有不变性:

$$(-p)(-x) \equiv px, \qquad \frac{-x}{-p} \equiv \frac{x}{p}.$$

为减少记号, 在计算时, 不定积分中的常数有时省略. 记

$$a = p/2\ (>0), \qquad b = \sqrt{p^2/4 - q}.$$

则

$$\begin{aligned}
I_{p,q}(x) &= \int \frac{\mathrm{d}x}{x\sqrt{x^2+px+q}} = \int \frac{\mathrm{d}x}{x\sqrt{(x+p/2)^2-(p^2/4-q)}} \\
&= \int \frac{\mathrm{d}x}{x\sqrt{(x+a)^2-b^2}} \qquad (x+a = bu \quad \text{i.e.} \quad x = bu-a,\ \mathrm{d}x = b\mathrm{d}u\ ) \\
&= \int \frac{b\mathrm{d}u}{(bu-a)b\sqrt{u^2-1}} = \int \frac{\mathrm{d}u}{(bu-a)\sqrt{u^2-1}} \qquad (u = \frac{1}{\cos\theta}, \quad \mathrm{d}u = \frac{\sin\theta}{\cos^2\theta}\mathrm{d}\theta\ ) \\
&= \int \frac{\frac{\sin\theta}{\cos^2\theta}\mathrm{d}\theta}{(\frac{b}{\cos\theta}-a)\sqrt{\frac{1}{\cos^2\theta}-1}} = \int \frac{\tan\theta\mathrm{d}\theta}{(b-a\cos\theta)\tan\theta} \qquad (\text{这里假定}\ \tan\theta>0) \\
&= \int \frac{\mathrm{d}\theta}{b-a\cos\theta} \qquad (t=\tan(\theta/2)\ \text{即}\ \theta = 2\arctan t, \quad \mathrm{d}\theta = \frac{2}{1+t^2}\mathrm{d}t,\ \cos\theta = \frac{1-t^2}{1+t^2}\ ) \\
&= \int \frac{\frac{2}{1+t^2}\mathrm{d}t}{b-a\frac{1-t^2}{1+t^2}} = \int \frac{2\mathrm{d}t}{b(1+t^2)-a(1-t^2)} = \int \frac{2\mathrm{d}t}{(b+a)t^2+b-a} \\
&= \frac{2}{b+a}\int \frac{\mathrm{d}t}{t^2+\frac{b-a}{b+a}} =: J(t).
\end{aligned}$$

在进一步计算之前, 我们先将变量逐一返回:

$$t = \tan(\theta/2) = \sqrt{\frac{1-\cos\theta}{1+\cos\theta}} = \sqrt{\frac{u-1}{u+1}} = \sqrt{\frac{x+a-b}{x+a+b}}.$$

讨论:

(1) 设$q=0$. 则由$p>0$ 知$b=a=p/2$, $b+a=p$,

$$I_{p,0}(x) = J(t) = \frac{2}{b+a}\int\frac{\mathrm{d}t}{t^2} = \frac{-2}{p}t^{-1} = -\frac{2}{p}\sqrt{\frac{x+a+b}{x+a-b}} = -\frac{2}{p}\sqrt{\frac{x+p}{x}} = -\frac{2\sqrt{x^2+px}}{px}.$$

(2) 设$q\neq 0$. 则

$$I_{p,q}(x) = J(t) = \frac{2}{b+a}\int\frac{\mathrm{d}t}{t^2+\frac{b-a}{b+a}}.$$

当$q>0$ 时, $b-a<0$. 令$c=\sqrt{\frac{a-b}{a+b}}$. 则

$$\int\frac{\mathrm{d}t}{t^2+\frac{b-a}{b+a}} = \int\frac{\mathrm{d}t}{t^2-c^2} = \int\frac{1}{2c}\Big(\frac{1}{t-c}-\frac{1}{t+c}\Big)\mathrm{d}t$$

$$
\begin{aligned}
&= \frac{1}{2c}\Big(\log(t-c)-\log(t+c)\Big) = \frac{1}{2c}\log\left(\frac{t-c}{t+c}\right) = \frac{1}{2c}\log\left(\frac{t-c}{t+c}\right) \\
&= \frac{1}{2c}\log\left(\frac{\sqrt{\frac{x+a-b}{x+a+b}}-c}{\sqrt{\frac{x+a-b}{x+a+b}}+c}\right) = \frac{1}{2c}\log\left(\frac{\sqrt{x+a-b}-c\sqrt{x+a+b}}{\sqrt{x+a-b}+c\sqrt{x+a+b}}\right) \\
&= \frac{1}{2c}\log\left(\frac{x+a-b-c^2(x+a+b)}{(\sqrt{x+a-b}+c\sqrt{x+a+b})^2}\right) \\
&= \frac{1}{2c}\log\left(\frac{(1-c^2)x}{(x+a-b+c^2(x+a+b)+2c\sqrt{x+a-b}\sqrt{x+a+b}}\right) \\
&= \frac{1}{2c}\log\left(\frac{(1-c^2)x}{(1+c^2)x+2(a-b)+2c\sqrt{x^2+px+q}}\right) \\
&= \frac{-1}{2c}\log\left(\frac{(1+c^2)x+2(a-b)+2c\sqrt{x^2+px+q}}{(1-c^2)x}\right) \\
&= \frac{-1}{2c}\log\left(\frac{a}{b}+\frac{2\frac{a-b}{1-c^2}+2\frac{c}{1-c^2}\sqrt{x^2+px+q}}{x}\right) \\
&= \frac{-1}{2c}\log\left(\frac{p}{2\sqrt{q}}+\frac{\sqrt{q}+\sqrt{x^2+px+q}}{x}\right)+C \\
&= \frac{-1}{2c}\log\left(\frac{1}{2\sqrt{q}}+\frac{\sqrt{q}+\sqrt{x^2+px+q}}{px}\right)+C
\end{aligned}
$$

其中用到

$$\frac{a}{b}=\frac{p}{2b},\quad 2\frac{a-b}{1-c^2}=\frac{q}{b},\quad 2\frac{c}{1-c^2}=\frac{\sqrt{q}}{b}.$$

代入上式并注意

$$(a+b)c=\sqrt{a^2-b^2}=\sqrt{q}$$

得到

$$I_{p,q}(x)=\frac{-1}{\sqrt{q}}\log\left(\frac{1}{2\sqrt{q}}+\frac{\sqrt{q}+\sqrt{x^2+px+q}}{px}\right)+C.$$

当$q<0$ 时, $b-a>0$. 令$c=\sqrt{\frac{b-a}{a+b}}$. 则

$$
\begin{aligned}
&\int\frac{\mathrm{d}t}{t^2+\frac{b-a}{b+a}} = \int\frac{\mathrm{d}t}{t^2+c^2} = \frac{1}{c}\int\frac{\mathrm{d}(t/c)}{(t/c)^2+1} = \frac{1}{c}\arctan(t/c) \\
&= \frac{1}{c}\arctan\left(\frac{1}{c}\sqrt{\frac{x+a-b}{x+a+b}}\right) = \frac{1}{c}\arctan\left(\frac{\sqrt{x^2+px+q}}{c(x+a+b)}\right) \\
&= \frac{a+b}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\frac{\sqrt{-q}}{a+b}x+\sqrt{-q}}\right)
\end{aligned}
$$

$\Longrightarrow$

$$
\begin{aligned}
I_{p,q}(x) &= \frac{2}{b+a}\int\frac{\mathrm{d}t}{t^2+\frac{b-a}{b+a}} = \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\frac{\sqrt{-q}}{a+b}x+\sqrt{-q}}\right)\\
&= \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\sqrt{-q}\,(\frac{x}{p/2+\sqrt{p^2/4-q}}+1)}\right)\\
&= \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\sqrt{-q}\,(\frac{2x}{p+\sqrt{p^2-4q}}+1)}\right).
\end{aligned}
$$

当$p>0$ 时,

$$
I_{p,q}(x) = \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\sqrt{-q}\,(\frac{2x}{p(1+\sqrt{1-4q/p^2})}+1)}\right).
$$

当$p=0$ 时,

$$
I_{0,q}(x) = \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+q}}{x+\sqrt{-q}}\right).
$$

结论(包括$p<0$的情形):

当$q=0$ 时

$$
I_{p,0}(x) = -\frac{2\sqrt{x^2+px}}{px}+C.
$$

当$q>0$ 时

$$
I_{p,q}(x) = \frac{-1}{\sqrt{q}}\log\left(\frac{1}{2\sqrt{q}}+\frac{\sqrt{q}+\sqrt{x^2+px+q}}{px}\right)+C.
$$

当$q<0$ 时, 若$p\neq 0$, 则

$$
I_{p,q}(x) = \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+px+q}}{\sqrt{-q}\,(\frac{2x}{p(1+\sqrt{1-4q/p^2})}+1)}\right)+C\,;
$$

若$p=0$,则

$$
I_{0,q}(x) = \frac{2}{\sqrt{-q}}\arctan\left(\frac{\sqrt{x^2+q}}{x+\sqrt{-q}}\right)+C.
$$

最后验证结果: 对所得结果求导得知上述计算是正确的. $\square$

附: 关于无理函数的不定积分 $\int R(x,\sqrt{ax^2+bx+c})\mathrm{d}x$ 的结构.

设$R(x,y)$ 是一个实系数二元有理函数, 即

$$
R(x,y) = \frac{P(x,y)}{Q(x,y)}
$$

其中P,Q 为实系数二元多项式, 即

$$P(x,y)=\sum_{i,j\geq 0,\, i+j\leq n} a_{i,j}x^iy^j,\quad Q(x,y)=\sum_{i,j\geq 0,\, i+j\leq m} b_{i,j}x^iy^j;\quad (x,y)\in\mathbb{R}^2.$$

按y的奇、偶次幂分解有

$$P(x,y)=\sum_{i,j\geq 0,\, i+2j\leq n} a_{i,2j}x^iy^{2j}+\sum_{i,j\geq 0,\, i+2j+1\leq n} a_{i,2j+1}x^iy^{2j+1}=P_1(x,y^2)+yP_2(x,y^2).$$

同样有

$$Q(x,y)=Q_1(x,y^2)+yQ_2(x,y^2).$$

由此得(对不同的多项式给予编号)

$$R(x,y)=\frac{P_1(x,y^2)+yP_2(x,y^2)}{Q_1(x,y^2)+yQ_2(x,y^2)}=\frac{(P_1(x,y^2)+yP_2(x,y^2))(\,Q_1(x,y^2)-yQ_2(x,y^2)\,)}{(Q_1(x,y^2))^2-y^2(Q_2(x,y^2))^2}$$

$$=\frac{P_3(x,y^2)+yP_4(x,y^2)}{Q_3(x,y^2)}=\frac{P_3(x,y^2)}{Q_3(x,y^2)}+\frac{y^2P_4(x,y^2)}{yQ_3(x,y^2)}=\frac{P_3(x,y^2)}{Q_3(x,y^2)}+\frac{P_5(x,y^2)}{Q_3(x,y^2)}\cdot\frac{1}{y}$$

$\Longrightarrow$

$$R(x,\sqrt{ax^2+bx+c})=\frac{P_3(x,ax^2+bx+c)}{Q_3(x,ax^2+bx+c)}+\frac{P_5(x,ax^2+bx+c)}{Q_3(x,ax^2+bx+c)}\cdot\frac{1}{\sqrt{ax^2+bx+c}}.$$

如令

$$P_1(x)=P_3(x,ax^2{+}bx{+}c),\quad P_2(x)=P_5(x,ax^2{+}bx{+}c),\qquad Q(x)=Q_3(x,ax^2{+}bx{+}c)$$

则P_1,P_2,Q 都是x的实系数多项式且

$$R(x,\sqrt{ax^2+bx+c})=\frac{P_1(x)}{Q(x)}+\frac{P_2(x)}{Q(x)}\cdot\frac{1}{\sqrt{ax^2+bx+c}}.$$

再由多项式的带余除法有分解

$$\frac{P_1(x)}{Q(x)}=P_3(x)+\frac{P_4(x)}{Q(x)},\qquad \frac{P_2(x)}{Q(x)}=P_5(x)+\frac{P_6(x)}{Q(x)}$$

其中$\deg P_4<\deg Q$, $\deg P_6<\deg Q$, 即$\frac{P_4(x)}{Q(x)},\frac{P_6(x)}{Q(x)}$ 是真分式. 于是得到

$$R(x,\sqrt{ax^2+bx+c})=P_3(x)+\frac{P_4(x)}{Q(x)}+\frac{P_5(x)}{\sqrt{ax^2+bx+c}}+\frac{P_6(x)}{Q(x)}\cdot\frac{1}{\sqrt{ax^2+bx+c}}.$$

$\Longrightarrow$

$$\begin{aligned}&\int R(x,\sqrt{ax^2+bx+c})\mathrm{d}x\\ &=\int P_3(x)\mathrm{d}x+\int\frac{P_4(x)}{Q(x)}\mathrm{d}x+\int\frac{P_5(x)}{\sqrt{ax^2+bx+c}}\mathrm{d}x+\int\frac{P_6(x)}{Q(x)}\cdot\frac{1}{\sqrt{ax^2+bx+c}}\mathrm{d}x.\end{aligned}$$

考虑最一般的情形, 可以假定分母Q 的次数≥ 1. 此外可以假定Q的最高幂的系数为1 (即Q 是首1 多项式) 否则把Q的最高幂的系数吸收到分子上去.

设

$$Q(x)=\prod_{i=1}^{r}(x-\alpha_i)^{k_i}\prod_{j=1}^{s}(x^2+p_jx+q_j)^{m_j}$$

其中 $\alpha_1,\alpha_2,...,\alpha_r\in\mathbb{R}$ 互不相同, $(p_1,q_1),(p_2,q_2),...,(p_s,q_s)\in\mathbb{R}^2$ 互不相同, 满足 $p_j^2<4q_j$ $(j=1,2,...,s)$, k_i,m_j 为正整数 $(i=1,2,...,r,j=1,2,...,s)$. 则由真分式分解定理知

$$\frac{P_4(x)}{Q(x)},\ \frac{P_6(x)}{Q(x)}\quad \text{可表示成}\quad \sum_{i=1}^{r}\sum_{k=1}^{k_i}\frac{A_k^{(i)}}{(x-\alpha_i)^k}+\sum_{j=1}^{s}\sum_{m=1}^{m_j}\frac{B_m^{(j)}x+C_m^{(j)}}{(x^2+p_jx+q_j)^m}$$

的形式, 其中若 $Q(x)$ 只有实根, 则等式右端第二项不出现; 若 $Q(x)$ 只有虚根, 则等式右端第一项不出现.

这样一来, 不定积分$\int R(x,\sqrt{ax^2+bx+c})\mathrm{d}x$ 就等于下面六种类型的**线性组合**:

$$(1)\qquad \int P(x)\mathrm{d}x,\qquad P(x)\ \text{是多项式(下同)};$$

$$(2)\qquad \int\frac{\mathrm{d}x}{(x-x_0)^k},\qquad x_0\in\mathbb{R},\qquad k\in\mathbb{N};$$

$$(3)\int\frac{Bx+C}{(x^2+px+q)^m}\mathrm{d}x\qquad \text{其中}\quad B,C,p,q\ \text{为常数且}\ p^2<4q\ \text{(下同)},\quad m\in\mathbb{N};$$

$$(4)\qquad \int\frac{P(x)}{\sqrt{ax^2+bx+c}}\mathrm{d}x\,;$$

$$(5)\qquad \int\frac{\mathrm{d}x}{(x-x_0)^k\sqrt{ax^2+bx+c}},\qquad x_0\in\mathbb{R},\qquad k\in\mathbb{N};$$

$$(6)\qquad \int\frac{Bx+C}{(x^2+px+q)^m\sqrt{ax^2+bx+c}}\mathrm{d}x,\qquad m\in\mathbb{N}.$$

为证明这一点, 我们先将(4),(5),(6) 的线性组合写成统一形式:

$$\int\frac{P(x)}{Q(x)}\cdot\frac{1}{\sqrt{ax^2+bx+c}}\mathrm{d}x$$

其中P, Q 为多项式. 如果不定积分$\int R(x, \sqrt{ax^2+bx+c})\mathrm{d}x$ 等于上面(4),(5),(6) 三种类型的线性组合, 则有

$$\int R(x, \sqrt{ax^2+bx+c})\mathrm{d}x = \int \frac{P(x)}{Q(x)} \cdot \frac{1}{\sqrt{ax^2+bx+c}}\mathrm{d}x.$$

根据原函数或不定积分的定义(或对上式两端对x 求导) 可知它等价于

$$R(x, \sqrt{ax^2+bx+c}) = \frac{P(x)}{Q(x)} \cdot \frac{1}{\sqrt{ax^2+bx+c}}. \tag{$*1$}$$

现将实系数二次多项式ax^2+bx+c 写为(考虑$a \neq 0$和单根的情形)

$$ax^2+bx+c = a(x-\lambda_1)(x-\lambda_2), \qquad a \neq 0, \quad \lambda_1, \lambda_2 \in \mathbb{C}, \ \ \lambda_1 \neq \lambda_2.$$

让我们取有理函数为

$$R(x, y) = 1 + \frac{P_0(x)}{y}$$

其中P_0 为一个实系数多项式. 则由式($*1$) 有

$$1 + \frac{P_0(x)}{\sqrt{ax^2+bx+c}} = \frac{P(x)}{Q(x)} \cdot \frac{1}{\sqrt{ax^2+bx+c}}, \quad x \in \mathbb{R}. \tag{$*2$}$$

我们来导出矛盾. 可以假设P, Q 无次数≥ 1的公因式(否则约去公因式).

由($*2$) 有

$$\sqrt{ax^2+bx+c} + P_0(x) = \frac{P(x)}{Q(x)}, \quad \sqrt{ax^2+bx+c} = \frac{P(x) - P_0(x)Q(x)}{Q(x)} =: \frac{P_1(x)}{Q(x)},$$

$$ax^2+bx+c = \frac{(P_1(x))^2}{(Q(x))^2}$$

$\Longrightarrow$

$$(ax^2+bx+c)(Q(x))^2 = (P_1(x))^2 \tag{$*3$}$$

$\Longrightarrow$

$$P_1(\lambda_1) = P_1(\lambda_2) = 0 \quad \Longrightarrow \quad P_1(x) = (x-\lambda_1)(x-\lambda_2)P_2(x) = (ax^2+bx+c)P_3(x).$$

代入($*3$) 得

$$(ax^2+bx+c)(Q(x))^2 = (ax^2+bx+c)^2(P_3(x))^2 \quad \Longrightarrow \quad (Q(x))^2 = (ax^2+bx+c)^2(P_3(x))^2$$

$\Longrightarrow$ (如上)

$$Q(\lambda_1) = Q(\lambda_2) = 0 \quad \Longrightarrow \quad Q(x) = (ax^2+bx+c)Q_1(x).$$

$\Longrightarrow$

$$P(x) = P_1(x) + P_0(x)Q(x) = (ax^2+bx+c)(\,P_3(x) + P_0(x)Q_1(x)\,)$$

$\Longrightarrow$ $P(x), Q(x)$ 有公因式ax^2+bx+c, 这与P, Q 无次数≥ 1的公因式矛盾. 这矛盾说明仅有(4),(5),(6) 这三种类型是不够的.

最后, 就这个例子而言, 我们有

$$\int R(x, \sqrt{ax^2+bx+c})\mathrm{d}x = \int \Big(1 + \frac{P_0(x)}{\sqrt{ax^2+bx+c}}\Big)\mathrm{d}x = x + \int \frac{P_0(x)}{\sqrt{ax^2+bx+c}}\mathrm{d}x$$

它是(1),(4) 两种类型的线性组合. $\square$

本章一些复习练习题:

1. 求下列不定积分:

$$(1)\quad \int 5^x \mathrm{d}x, \qquad (2)\quad \int \tan^2 x \mathrm{d}x, \qquad (3)\quad \int \frac{x^2+x^3}{1+x^2}\mathrm{d}x,$$
$$(4)\quad \int \frac{\mathrm{d}x}{1-\sin x}, \qquad (5)\quad \int \frac{\mathrm{d}x}{\sqrt{7x^2-8}}, \qquad (6)\quad \int \frac{\mathrm{d}x}{x\sqrt{x^2+1}}$$

2. 设非负函数$F(x)$ 是$f(x)$ 在$[0,+\infty)$ 上的一个原函数且$F(0)=1$. 假设

$$F(x)f(x) = \sin^2 x, \quad x \in [0,+\infty). \qquad 求 \quad F(x).$$

3. 设$Q(x) = (x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_{n+1})$ 其中$\alpha_1, \alpha_2, ..., \alpha_{n+1} \in \mathbb{R}$ 互不相同 ($n \in \mathbb{N}$). 证明对任意次数$\leq n$的多项式$P(x)$ 有

$$\frac{P(x)}{Q(x)} = \sum_{k=1}^{n+1} \frac{P(\alpha_k)}{Q'(\alpha_k)(x-\alpha_k)}$$

或等价地

$$P(x) = \sum_{k=1}^{n+1} \frac{P(\alpha_k)}{Q'(\alpha_k)} \cdot \frac{Q(x)}{(x-\alpha_k)}, \qquad x \in \mathbb{R}.$$

4. 设$f(x)$在点a 可导且$f(a) \neq 0$. 证明

$$\lim_{n\to\infty} \left(\frac{f(a+\frac{1}{n})}{f(a)}\right)^n = \exp(\frac{f'(a)}{f(a)}).$$

5. 一个静止质量为m_0的物体以速度v运动, 按相对论公式, 其动能为

$$K = m_0 c^2 \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1 \right) \qquad (c\text{ 为光速}).$$

证明当比值$v/c \to 0$ 时有$K \sim \frac{1}{2} m_0 v^2$. (后者表示两个量之比趋于1.)

6.(复合函数的微分) 设函数f 在$\mathbb{R}$ 上无穷次可导, 函数g 在(a,b) 内可导且满足$g'(x) = f(g(x))$, $x \in (a,b)$. 证明g 在(a,b) 内无穷次可导.

7.(复合函数的微分) 设$0 < \varepsilon < 1$. 设$y = y(x)$ 是由方程$y - \varepsilon \sin y = x$ 定义的x的连续函数. (即$y(x)$满足恒等式$y(x) - \varepsilon \sin(y(x)) \equiv x, x \in \mathbb{R}$.) 证明$y(x)$处处可微, 并求$y'(x)$的表达式(当然这个表达式仍然是隐式的, 即与$y(x)$有关).

8. (高阶导数Leibniz) 试求$(x^2 \cos(2x))^{(n)}$. (化简后结果只有很少几项).

9. (Leibniz). 定义

$$\|f\| = \max_{x\in[a,b]} \sum_{k=0}^{n} \frac{1}{k!} |f^{(k)}(x)|, \quad f \in C^n([a,b]).$$

证明对任意$f, g \in C^n([a,b])$

$$\|fg\| \le \|f\| \|g\|.$$

10. (1) 设$f \in C^2([-a,a])$, $f(0) = 0$. 证明

$$\lim_{n\to\infty} \sum_{k=1}^{n} f\left(\frac{k}{n^2}\right) = \frac{f'(0)}{2}.$$

(2) 设$a > 0$. 求极限

$$\lim_{n\to\infty} \sum_{k=1}^{n} (a^{k/n^2} - 1).$$

11. 证明方程$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} = 0$ 在$\mathbb{R}$ 上有实根但不超过两个. (提示: 考察方程$f(x) - f'(x) = 0$.)

12. (1) 设f 在$[a,b]$ 上连续, 在(a,b) 内可微且$f(a) = f(b) = 0$. 证明存在$\xi \in (a,b)$ 所得$f(\xi) - f'(\xi) = 0$.

(2) 试由(1) 证明方程$e^x - x^n = 0$ 至多有三个不同实根. 这里n 为自然数.

13.(指数函数的作用) 设$f(x)$ 在$[a,b]$ 上可微且存在常数A 所得

$$f(a) < A, \quad f(x)+f'(x) < A \quad \forall x\in(a,b).$$

证明$f(x) < A$ for all $x\in(a,b)$.

14. (从结论出发寻找辅助函数) 设f,g 在$[a,b]$ 上可导且$g'(x)\neq 0$ for all $x\in(a,b)$. 证明存在$\xi\in(a,b)$ 使得

$$\frac{f(a)-f(\xi)}{g(\xi)-g(b)}=\frac{f'(\xi)}{g'(\xi)}.$$

15. 设$f(x)$ 在有界开区间(a,b) 内可导且导函数$f'(x)$ 在(a,b) 内有界. 证明f 在(a,b) 内一致连续从而可以定义$f(a):=\lim_{x\to a+}f(x)$, $f(b):=\lim_{x\to b-}f(x)$.

16. 设f 在区间I 上二次可微且$f'(x)f''(x)\equiv 0, x\in I$. 问$f$ 是那类函数?

17. 设f,g 在$[0,+\infty)$ 上二次可微且有$f(0)=g(0), f'(0)=g'(0), f''(x)>g''(x)$ for all $x>0$. 证明$f(x)>g(x)$ for all $x>0$.

18. 设f在$[0,+\infty)$上连续**非负**, $f(0)=0$, 且f 在$(0,+\infty)$ 内可导. 假设$f(x)\geq f'(x)$ for all $x\in(0,+\infty)$. 证明$f(x)\equiv 0$, $x\in[0,+\infty)$.

19. 证明

$$2\arctan x+\arcsin\left(\frac{2x}{1+x^2}\right)=\pi, \qquad x\geq 1.$$

20. 证明下列不等式:

$$\frac{2}{2x+1}<\log\left(1+\frac{1}{x}\right), \quad x>0.$$

$$1+x\log(x+\sqrt{1+x^2})\geq\sqrt{1+x^2}, \quad x\in\mathbb{R}.$$

$$\frac{1}{x}+\arctan x>\frac{\pi}{2}, \quad x>0.$$

21. (Darboux). 设f 在$[a,b]$ 上两次可导且$f(a)=f(b), f'(a)>0, f'(b)>0$. 证明存在$x_0\in(a,b)$ 使得$f''(x_0)=0$.

此外在初等函数范围内举例说明这样的函数存在.

22. 墙上有一幅画高a 尺, 底边距观察者眼睛高b尺. 当观察者离墙多远时, 观察画的视角最大?

23. (凹凸性) 设$x_k \in (0, \pi), k = 1, 2, ..., n$. 证明

$$\prod_{k=1}^{n} \frac{\sin x_k}{x_k} \le \Big(\frac{\sin \bar{x}}{\bar{x}}\Big)^n, \qquad \bar{x} := \frac{1}{n}\sum_{k=1}^{n} x_k.$$

24. 设f 在$[a, b]$ 是凸(convex)函数, A 为常数. 则为使$f(x) \le A$ for all $x \in [a, b]$, 只需$f(a) \vee f(b) \le A$.

25. 设$f(x)$ 在(a, b) 内是凸函数且处处可微. 证明$f \in C^1((a, b))$.

26. 设$f(x)$ 在$(0, +\infty)$ 内是凸函数且处处可微. 又设$\lim\limits_{x\to+\infty} f(x) = A \in \mathbb{R}$. 证明$\lim\limits_{x\to+\infty} f'(x) = 0$.

27.(L'Hospital) 设二阶导数$f''(x_0)$ 存在, $f'(x_0) \ne 0$. 求

$$\lim_{x\to x_0}\left(\frac{1}{f(x) - f(x_0)} - \frac{1}{f'(x_0)(x - x_0)}\right) = ?$$

28.(Taylor) 熟知$\log 2 = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$. 但有人说实际上有很快的收敛:

$$\log 2 = \sum_{n=1}^{\infty} \frac{1}{n}\Big(\frac{1}{2}\Big)^n. \qquad \text{这对吗?}$$

【$\log 2 = 0.693147...$ 你可以用简易计算器或手算验证一下, 以加深印象.】

29. 设$f, g \in C^{\infty}(\mathbb{R})$ 满足: 函数在任一点x_0处展开的Taylor级数在$(x_0 - 1, x_0 + 1)$内处处收敛于函数自己. 假设f, g 在一个有界无限集上相等. 证明$f(x) \equiv g(x)$, $x \in \mathbb{R}$.

30. 设$f \in C^{\infty}(\mathbb{R})$ 满足

$$f^{(n)}(x) \ge 0 \qquad \forall\, x \in \mathbb{R}, \quad \forall\, n = 0, 1, 2, \dots.$$

证明f在任一点处展开的Taylor级数处处收敛且收敛于f. 不失一般性只需考虑在$x = 0$ 展开的Taylor级数, 并证明

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \qquad \forall\, x \in \mathbb{R}.$$

习题课

1. 设$f,g:[0,1]\to[0,1]$ 可交换, 即$f\circ g=g\circ f$. 又设f,g 在$[0,1]$ 上两次可导且

$$f'(x)f''(x)g'(x)g''(x)\neq 0 \qquad \forall\, x\in[0,1].$$

令A_f为f的不动点的集合, 即$A_f=\{x\in[0,1]\,|\,f(x)=x\}$. 证明$A_f=A_g$. 由此可知$f,g$有公共不动点. 此外证明$A_f$至多有两个元素.

例子:

$$f(x)=\frac{2}{7}x^2+\frac{4}{7}x+\frac{1}{7},\quad g(x)=f(f(x)).$$

此时f的两个不动点为$x=\frac{1}{2},x=1$.

【证】由复合函数求导法则和$f(g(x))\equiv g(f(x))$ 有

$$f'(g(x))g'(x)=(f(g(x))'=(g(f(x))'=g'(f(x))f'(x),\quad x\in[0,1].$$

对任意$c\in A_f$, 代入$x=c$于上式并注意$f(c)=c$ 有

$$f'(g(c))g'(c)=g'(f(c))f'(c)=g'(c)f'(c).$$

由假设知$g'(c)\neq 0$ 因此$f'(g(c))=f'(c)$. 再由$f'(x)$ 严格单调(因$f''(x)\neq 0\,\forall\, x\in[0,1]$) 知$g(c)=c$, 即$c\in A_g$. 这证明了$A_f\subset A_g$. 因$f,g$地位对称, 故同理有$A_g\subset A_f$. 所以$A_f=A_g$.

最后证明A_f至多有两个元素. 令$h(x)=f(x)-x$. 则

$$h''(x)=f''(x)\neq 0\quad \forall\, x\in[0,1].$$

因此方程$h(x)=0$在$[0,1]$中至多只有两个根. 否则设$0\le x_1<x_2<x_3\le 1$是$h(x)=0$的三个根: 则由$h(x_1)=h(x_2)=h(x_3)\,(=0)$ 知存在$x_1<\xi_1<x_2<\xi_2<x_3$使得$h'(\xi_1)=0=h'(\xi_2)$. 进而存在$\xi_1<\xi<\xi_2$使得$h''(\xi)=0$, 与$h''(x)$ 无零点矛盾. 所以A_f 至多有两个元素. $\square$

2. 设$[a,b]$ 为有界闭区间, $f:[a,b]\to\mathbb{R}, g:(a,b)\to\mathbb{R}$, f 在$[a,b]$ 上连续, 在(a,b) 内两次可微且满足微分不等式

$$f''(x)+g(x)f'(x)-f(x)\ge 0 \qquad \forall\, x\in(a,b).$$

假设$\max\{f(a),f(b)\}>0$. 证明f在$[a,b]$ 上的最大值只能在区间端点达到, 即

$$f(x)<\max\{f(a),f(b)\}\qquad\forall\, x\in(a,b).$$

【证】由假设知f 在$[a,b]$ 上连续因而有最大值. 假设存在$x_0\in(a,b)$ 使得$f(x_0)=\max\limits_{x\in[a,b]}f(x)$. 则$f(x_0)\ge\max\{f(a),f(b)\}>0$ 且x_0 也是f在开区间(a,b)内的一个局部极大值点. 由Fermat 极值原理知$f'(x_0)=0$ 从而由上述微分不等式有

$$f''(x_0)-f(x_0)\ge 0\qquad\text{即}\qquad f''(x_0)\ge f(x_0)>0.$$

根据极值定理知x_0 是f 在(a,b)内的一个局部严格极小值点, 这与x_0是f在(a,b)内的一个局部极大值点矛盾. 这矛盾证明了f在$[a,b]$ 上的最大值只能在区间端点达到. 因此$\max\limits_{x\in[a,b]}f(x)=\max\{f(a),f(b)\}$.

又假若存在$x_0\in(a,b)$ 使得$f(x_0)=\max\{f(a),f(b)\}$. 则x_0 是f的最大值点. 由上面证明知这是不可能的. 因此必是对所有$x\in(a,b)$ 都有$f(x)<\max\{f(a),f(b)\}$. $\Box$

3. 设$f:\mathbb{R}\to\mathbb{R}$ 处处可微且满足$f\circ f=f$, $f(x)\not\equiv$ 常数. 证明$f(x)\equiv x, x\in\mathbb{R}$.

【证】先证明

$$f(x)=x\qquad\forall\, x\in f(\mathbb{R}).\tag{1}$$

事实上对任意$x\in f(\mathbb{R})$, 写$x=f(y), y\in\mathbb{R}$. 则有$f(x)=f(f(y))=f(y)=x$. 所以(1)成立.

由(1) 知为证本题, 只需证明$f(\mathbb{R})=\mathbb{R}$.

因连续函数把区间映为区间, 故$f(\mathbb{R})$ 是区间. 又f不是常数, 所以$f(\mathbb{R})$ 不是退化的区间.

由假设$f(x)\equiv f(f(x))$ 和链锁规则有

$$f'(x)=f'(f(x))f'(x)\qquad\forall\, x\in\mathbb{R}.\tag{2}$$

由此和(1)得到

$$f'(x)=(f'(x))^2,\quad x\in f(\mathbb{R}).$$

这说明: 对任意$x\in\mathbb{R}$, 要么$f'(x)=0$, 要么$f'(x)=1$. 于是由导函数介值定理知

要么有$f'(x)\equiv 0$, $x\in f(\mathbb{R})$; 要么有$f'(x)\equiv 1$, $x\in f(\mathbb{R})$.

但因f 不是常值函数, 故存在$\xi \in \mathbb{R}$ 使得$f'(\xi) \neq 0$. 由(2) 得到

$$0 \neq f'(\xi) = f'(f(\xi))f'(\xi) \quad \Longrightarrow \quad f'(f(\xi)) = 1.$$

因$f(\xi) \in f(\mathbb{R})$, 于是有

$$f'(x) \equiv 1, \quad x \in f(\mathbb{R}). \tag{3}$$

为证明区间$f(\mathbb{R}) = \mathbb{R}$, 只需证明

$$\inf f(\mathbb{R}) = -\infty, \quad \sup f(\mathbb{R}) = +\infty.$$

反证法: 假设$\inf f(\mathbb{R}) > -\infty$, 则$\inf f(\mathbb{R}) \in \mathbb{R}$. 令$a = \inf f(\mathbb{R})$. 由下确界的定义存在$\{y_n\}_{n=1}^{\infty} \subset f(\mathbb{R})$ 使得$y_n \to a\,(n \to \infty)$. 由(1) 知$f(y_n) = y_n$ 因此再由f连续得到

$$a = \lim_{n\to\infty} y_n = \lim_{n\to\infty} f(y_n) = f(a).$$

于是由$f(a) = a = \inf f(\mathbb{R})$ 得知

$$f(x) \geq f(a) \qquad \forall\, x \in \mathbb{R}.$$

这表明a是f在$\mathbb{R}$上的最小值点从而也是极值点. 因此必有$f'(a) = 0$. 但由$a = f(a) \in f(\mathbb{R})$ 和(3)又知应有$f'(a) = 1$, 矛盾. 这矛盾证明了必有$\inf f(\mathbb{R}) = -\infty$. 同理可证$\sup f(\mathbb{R}) = +\infty$. 所以$f(\mathbb{R}) = \mathbb{R}$. $\square$

4. Let $P(x)$ be a real polynomial on $\mathbb{R}$ satisfying

$$P'''(x) - P''(x) - P'(x) + P(x) \geq 0 \qquad \forall\, x \in \mathbb{R}.$$

Prove that $P(x) \geq 0$ for all $x \in \mathbb{R}$.

Proof. Let

$$f(x) = P''(x) - P(x).$$

Then

$$f'(x) - f(x) = P'''(x) - P'(x) - P''(x) + P(x) \geq 0 \qquad \forall\, x \in \mathbb{R}.$$

So

$$(e^{-x}f(x))' = e^{-x}\Big(f'(x) - f(x)) \geq 0 \qquad \forall\, x \in \mathbb{R}.$$

So the function $x \mapsto e^{-x} f(x)$ is non-decreasing in $\mathbb{R}$. So for any $x \in \mathbb{R}$ we have

$$e^{-y} f(y) \geq e^{-x} f(x) \qquad \forall\, y > x.$$

Letting $y \to +\infty$ gives

$$0 = \lim_{y \to +\infty} e^{-y} f(y) \geq e^{-x} f(x)$$

and so $f(x) \leq 0$ for all $x \in \mathbb{R}$ i.e.

$$P''(x) \leq P(x) \qquad \forall\, x \in \mathbb{R}.$$

Here we used the fact that $f(x)$ is a polynomial and that

$$\lim_{y \to +\infty} \frac{y^n}{e^y} = 0 \qquad \forall\, n \in \{0, 1, 2, ...\}. \tag{$*$}$$

From this we have

$$\Big(e^{-x}(P'(x)+P(x))\Big)' = e^{-x}\Big(P''(x)+P'(x)-P'(x)-P(x)\Big) = e^{-x}\Big(P''(x)-P(x)\Big) \leq 0 \qquad \forall\, x \in \mathbb{R}.$$

So the function $x \mapsto e^{-x}(P'(x) + P(x))$ is non-increasing on $\mathbb{R}$. So for any $x \in \mathbb{R}$ we have

$$e^{-y}(P'(y) + P(y)) \leq e^{-x}(P'(x) + P(x)) \qquad \forall\, y > x.$$

Letting $y \to +\infty$ and using $(*)$ gives

$$0 = \lim_{y \to +\infty} e^{-y}(P'(y) + P(y)) \leq e^{-x}(P'(x) + P(x)).$$

So

$$P'(x) + P(x) \geq 0 \qquad \forall\, x \in \mathbb{R}.$$

From this we have

$$\Big(e^{x} P(x)\Big)' = e^{x}\Big(P'(x) + P(x)\Big) \geq 0 \qquad \forall\, x \in \mathbb{R}.$$

So the function $x \mapsto e^{x} P(x)$ is non-decreasing on $\mathbb{R}$. So for any $x \in \mathbb{R}$ we have

$$e^{y} P(y) \leq e^{x} P(x) \qquad \forall\, y < x.$$

Letting $y \to -\infty$ and using $(*)$ gives

$$0 = \lim_{y \to -\infty} e^{y} P(y) \leq e^{x} P(x).$$

So $P(x) \geq 0$ for all $x \in \mathbb{R}$. $\square$

5. Let $I \subset \mathbb{R}$ be an interval (bounded or unbounded) and let $f : I \to \mathbb{C}$ be differentiable on I satisfying for some constant $0 < A < \infty$

$$|f'(x)| \leq A|f(x)|, \quad x \in I.$$

Then for any $a \in I$ we have

$$|f(x)| \leq e^{A|x-a|}|f(a)|, \quad x \in I \qquad (**)$$

Proof. Let

$$F(x) = e^{-2A(x-a)}|f(x)|^2, \quad G(x) = e^{2A(x-a)}|f(x)|^2, \quad x \in I.$$

Compute

$$|f(x)|^2 = f(x)\overline{f(x)},$$

$$(|f(x)|^2)' = f'(x)\overline{f(x)} + f(x)(\overline{f(x)})' = f'(x)\overline{f(x)} + f(x)\overline{f'(x)} = 2\mathrm{Re}(f(x)\overline{f'(x)}),$$

$$\begin{aligned}
&F'(x) = e^{-2A(x-a)}(-2A)|f(x)|^2 + e^{-2A(x-a)}(|f(x)|^2)' \\
&= e^{-2A(x-a)}(-2A)|f(x)|^2 + e^{-2A(x-a)}2\mathrm{Re}(f(x)\overline{f'(x)}) \\
&= 2e^{-2A(x-a)}(-A|f(x)|^2 + \mathrm{Re}(f(x)\overline{f'(x)})) \leq 2e^{-2A(x-a)}(-A|f(x)|^2 + |f(x)||f'(x)|) \\
&= 2e^{-2A(x-a)}|f(x)|(|f'(x)| - A|f(x)|) \leq 0, \quad x \in I;
\end{aligned}$$

$$\begin{aligned}
&G'(x) = e^{2A(x-a)}2A[f(x)]^2 + e^{2A(x-a)}2\mathrm{Re}(f(x)\overline{f'(x)}) \\
&= 2e^{2A(x-a)}(A|f(x)|^2 + \mathrm{Re}(f(x)\overline{f'(x)}) \geq 2e^{2A(x-a)}(A[f(x)|^2 - |f(x)||f'(x)|) \\
&= 2e^{2A(x-a)}|f(x)|(A|f(x)| - |f'(x)|) \geq 0, \quad x \in I.
\end{aligned}$$

So $F(x)$ is non-increasing on I and $G(x)$ is non-decreasing on I. So for any $x \in I$, if $x \geq a$, then

$$F(x) \leq F(a) \quad \text{and so} \quad |f(x)| \leq e^{A(x-a)}|f(a)| = e^{A|x-a|}|f(a)|;$$

if $x \leq a$, then

$$G(x) \leq G(a) \quad \text{and so} \quad |f(x)| \leq e^{A(a-x)}|f(a)| = e^{A|x-a|}|f(a)|.$$

This proves $(**)$. $\square$

6. Let $[a, b]$ be a bounded closed interval and let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable in (a, b). Suppose that f is not linear on $[a, b]$. Prove that there are $\xi, \eta \in (a, b)$ such that

$$f'(\xi) < \frac{f(b) - f(a)}{b - a} < f'(\eta).$$

Proof. Let

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a), \quad x \in [a, b].$$

Then $F(a) = F(b) = 0$ and

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}, \quad x \in (a, b).$$

If $F'(x) \geq 0$ for all $x \in (a, b)$, then by the Lagrange mean-value theorem that $F(x)$ is non-decreasing on $[a, b]$ so that

$$0 = F(a) \leq F(x) \leq F(b) = 0 \qquad \forall x \in [a, b]$$

so that

$$F(x) \equiv 0, \quad x \in [a, b]$$

i.e.

$$f(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \qquad \forall x \in [a, b]$$

which contradicts the assumption that f is not linear on $[a, b]$. So there must be a $\xi \in (a, b)$ such that $F'(\xi) < 0$ i.e.

$$f'(\xi) < \frac{f(b) - f(a)}{b - a}.$$

Similarly if $F'(x) \leq 0$ for all $x \in (a, b)$, then by the Lagrange mean-value theorem that $F(x)$ is non-increasing on $[a, b]$ so that

$$0 = F(a) \geq F(x) \geq F(b) = 0 \qquad \forall x \in [a, b]$$

so that

$$F(x) \equiv 0, \quad x \in [a, b]$$

i.e.

$$f(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \qquad \forall\, x \in [a, b]$$

which contradicts the assumption that f is not linear on $[a, b]$. So there must be an $\eta \in (a, b)$ such that $F'(\eta) > 0$ i.e.

$$f'(\eta) > \frac{f(b) - f(a)}{b - a}.$$

□

7. Let $f : (0, b) \to \mathbb{R}$ be differentiable in $(0, b)$ and satisfy $\sup\limits_{0<x<b} x^{1-\alpha}|f'(x)| < +\infty$ for some constant $0 < \alpha < 1$. Prove that f is Hölder continuous in $(0, b)$ with the Hölder index α, i.e. there is a constant $0 < L < \infty$ such that

$$|f(y) - f(x)| \le L|y - x|^{\alpha} \qquad \forall\, x, y \in (0, b).$$

Proof. Let $C = \sup\limits_{0<x<b} x^{1-\alpha}|f'(x)|$. Using Cauchy's mean-value theorem we have that for any $0 < x < y < b$ there is ξ satisfying $x < \xi < y$ such that

$$\frac{f(y) - f(x)}{y^{\alpha} - x^{\alpha}} = \frac{f'(\xi)}{\alpha \xi^{\alpha - 1}} = \frac{1}{\alpha}\xi^{1-\alpha} f'(\xi)$$

so that

$$\frac{|f(y) - f(x)|}{y^{\alpha} - x^{\alpha}} = \frac{1}{\alpha}\xi^{1-\alpha}|f'(\xi)| \le \frac{C}{\alpha} := L.$$

So

$$|f(y) - f(x)| \le L(y^{\alpha} - x^{\alpha}) \le L(y - x)^{\alpha} = L|y - x|^{\alpha}$$

where we used the inequality

$$y^{\alpha} = (x + y - x)^{\alpha} \le x^{\alpha} + (y - x)^{\alpha}, \quad 0 < x < y < \infty. \qquad \square$$