

# 第三册

大青花鱼



# 目录

<b>第一章 无穷</b>	<b>5</b>
1.1 无穷集合的势 . . . . .	6
1.2 常见无穷集合的势 . . . . .	9
1.3 可数和不可数 . . . . .	12
<b>第二章 连续函数的变化</b>	<b>17</b>
2.1 函数在一点的变化 . . . . .	17
2.2 微变的运算法则 . . . . .	24
2.3 常见函数的微变 . . . . .	30
2.4 微变函数 . . . . .	36
2.5 多次微变 . . . . .	42
<b>第三章 研究可微函数</b>	<b>49</b>
3.1 增减与极值 . . . . .	49
3.2 局部性质 . . . . .	54

3.3 凹凸性质 . . . . .	61
<b>第四章 平直空间</b>	<b>67</b>
4.1 子空间与和空间 . . . . .	70
4.2 生成空间 . . . . .	73
4.3 基底和维数 . . . . .	75
<b>第五章 级数</b>	<b>81</b>
5.1 正项级数 . . . . .	82
5.2 正项级数的收敛 . . . . .	83
5.3 一般级数的收敛 . . . . .	89
<b>附录 A 无穷集合的势</b>	<b>93</b>
<b>附录 B 微变与函数</b>	<b>103</b>
2.1 函数的微变 . . . . .	103
2.2 常见函数的微变 . . . . .	109
2.3 研究可微函数 . . . . .	110
2.4 常见函数的局部展开 . . . . .	115
<b>附录 C 平直空间</b>	<b>117</b>

# 第一章 无穷

我们已经学习过无穷的概念。我们用数数原则来判定无穷。简单来说，数得尽的集合是有穷的，数不尽的集合是无穷的。现在，我们来进一步探讨无穷的性质。

数数原则把集合分成有穷的和无穷的两种。有穷的集合，集合元素的个数是可以知道的，因为按照定义，它是数得尽的。

我们把有穷集合的元素个数称为**集合的势**，用数字绝对值的符号来标记。比如，空集的势就是 0，记作  $|\emptyset| = 0$ 。集合  $\{2, 5, 1\}$  的势是 3，记作  $|\{2, 5, 1\}| = 3$ 。

比较两个集合的势，可以用最简单的“对消法”：每次从集合  $A$  中“拿掉”一个元素，对应地，也从集合  $B$  中“拿掉”一个元素。直到某个集合的元素被“拿光”。如果另一个集合里还有元素，就说明前者的势小于后者；如果另一个集合里也没有元素了，就说明两者的势相等。

容易验证，有穷集合的势有以下的基本性质：

1. 子集的势不多于母集的势；真子集的势小于真母集的势。
2. 集合  $A$  在集合  $B$  中补集的势，加上  $A$  的势，等于  $B$  的势。

## 1.1 无穷集合的势

有穷集合的基本性质符合我们日常生活中的经验。对于无穷集合，它们是否也成立呢？

来看以下的例子。

一个旅馆里有无穷个房间。房间的号码是正整数：1 号、2 号、3 号……

一天，有一个客人来旅馆里住宿。可所有的房间都住了人，怎么办呢？旅馆老板说：这样吧，我们把 1 号房的客人转到 2 号房，把 2 号房的客人转到 3 号房，把 3 号房的客人转到 4 号房……以此类推， $n$  号房的人转到  $n+1$  号房。

这样，1 号房就空出来了。客人顺利入住。

又一天，有无穷多个客人来旅馆里住宿。可所有的房间都住了人，怎么办呢？旅馆老板一看，每个客人有一个正整数号码，没有重复的也没有缺少的，恰好和房号一样。老板说：这样吧，我们把 1 号房的客人转到 2 号房，把 2 号房的客人转到 4 号房，把 3 号房的客人转到 6 号房……以此类推， $n$  号房的人转到  $2n$  号房。

这样，所有奇数号的房间就空出来了。于是，1 号客人住 1 号房，2 号客人住 3 号房，3 号客人住 5 号房，……以此类推， $n$  号客人入住  $2n-1$  号房。这样，所有的客人都顺利入住了。

以上的情况似乎有点违反我们的直觉。集合  $\{2, 3, 4, \dots\}$  和  $\{2, 4, 6, \dots\}$  是  $\{1, 2, 3, \dots\}$  的真子集，但它们的势与  $\{1, 2, 3, \dots\}$  相等。这说明，无穷集合的势，有着不同的性质。

为此，我们首先要定义无穷集合的势的关系。我们从“对消法”出发来构思。“对消法”中，我们实际上在给两个集合的元素建立一一对应的关系。每次从两个集合里分别“拿掉”的元素形成一一对应。

如果一个集合“拿光”的时候，另一个集合也“拿光”了，说明我们给两个集合的元素建立了一一对应的映射，也就是双射。

如果一个集合“拿光”的时候，另一个集合还有“剩余”，就说明我们无法在两个集合的元素之间建立双射。我们可以从前一个集合出发，建立到后一个集合的单射；但无法从后一个集合出发，建立到前一个集合的单射。这是因为后一个集合的元素太多了，总会有两个元素映射到同一个目标。

用这个思路，我们来定义无穷集合的势的关系。

**定义 1.1.1.** 设有集合  $A$  和  $B$ 。

- 如果存在  $A$  到  $B$  的双射，就说  $A$  和  $B$  的势相等，两者等势，记作  $|A| = |B|$ 。
- 如果存在从  $A$  到  $B$  的单射，就说  $A$  的势不大于  $B$  的势或  $A$  的势小于等于  $B$  的势，记作  $|A| \leq |B|$  或  $|B| \geq |A|$ 。
- 如果存在从  $A$  到  $B$  的单射，但不存在  $B$  到  $A$  的单射，就说  $A$  的势小于  $B$  的势，记作  $|A| < |B|$  或  $|B| > |A|$ 。

用这个定义，就可以理解无穷旅馆的例子。对于集合  $\{1, 2, 3, \dots\}$  和  $\{2, 3, 4, \dots\}$  来说，我们有双射  $n \mapsto n + 1$ ，因此两者等势。对于集合  $\{1, 2, 3, \dots\}$  和  $\{2, 4, 6, \dots\}$  来说，我们有双射  $n \mapsto 2n$ ，因此两者等势。

注意：

1. 以上定义对有穷集合、无穷集合都成立，也就是说，上面的  $A$ 、 $B$  分别可以是有穷、无穷集合。
2. 对于有穷集合  $A$ 、 $B$ ，这个定义与有穷集合的势的定义是兼容的。
3. **无穷集合的势不是数。**我们沿用有穷集合的记法，把无穷集合  $A$  的势记作  $|A|$ 。但  $|A|$  并不是任何数。因此，**无穷集合的势不能参与四则运算。**
4. 同样地，当我们写  $|A| \leq |B|$ 、 $|A| = |B|$  的时候，里面的等号和不等号也不能按数的相等、不等关系来理解。但在一定条件下，它们的性

质和数的相等、不等关系是一样的（具体参见附录 A）。

### 思考 1.1.1.

1. 在无穷旅馆的例子中，如果旅馆已经住满了，而有无穷多个旅客来住宿，旅客的编号用全体整数来编号。是否能腾出足够的房间呢？如何操作？如果旅客的编号用全体有理数呢？全体实数呢？
2. 请用满射代替单射，定义无穷集合的势的关系。
3. 能否给无穷集合的势定义四则运算？

### 习题 1.1.1.

1. 以下集合的势是多少？
  - 1.1.  $\{\text{一年中的月份}\}$
  - 1.2.  $\{\text{100以内的素数}\}$
  - 1.3.  $\{S\text{的所有子集}\}$ ，其中  $|S| = 9$ .
  - 1.4.  $\{(x, y, z) \mid 0 \leq x, y, z \leq 3, x + 2y + z \leq 8, x, y \in \mathbb{Z}\}$
2. 证明：把有穷集合的势定义为元素的个数，这样的定义，满足定义 1.1.1 中势的关系。
3. 证明：有穷集合的势总小于无穷集合的势。
4. 证明，如果集合  $A$  是  $B$  的子集，那么  $|A| \leq |B|$ 。
5. 证明：定义 1.1.1 中集合的势相等的关系，是一种等价关系，即满足：
  - 自反性：任意集合  $A$  的势等于自己： $|A| = |A|$ 。
  - 对称性：若集合  $A$  的势等于集合  $B$  的势，则集合  $B$  的势等于集合  $A$  的势。
  - 传递性：若集合  $A$  的势等于集合  $B$  的势，集合  $B$  的势等于集合  $C$  的势，则集合  $A$  的势等于集合  $C$  的势。
6. 证明：如果集合  $A, B$  满足  $|A| < |B|$ ，那么  $|A| = |B|$  不成立。



## 1.2 常见无穷集合的势

常见的无穷集合有：自然数集  $\mathbb{N}$ 、正整数集  $\mathbb{Z}^+$ 、整数集  $\mathbb{Z}$ 、有理数集  $\mathbb{Q}$ 、实数集  $\mathbb{R}$  等等。下面来看一些无穷集合的势的关系。

容易看出， $|\mathbb{N}| = |\mathbb{Z}^+|$ ，两者间的关系可以用双射  $n \mapsto n + 1$  确立。而双射  $n \mapsto -n$  可以说明  $|\mathbb{Z}^+| = |\mathbb{Z}^-|$ ，所以  $|\mathbb{N}| = |\mathbb{Z}^-|$ 。

考虑把整数映射到自然数的映射：

$$f, n \mapsto \begin{cases} 2n & \text{如果 } n \geq 0 \\ -2n - 1 & \text{如果 } n < 0 \end{cases}$$

也就是：

$$\begin{array}{cccccccc} 0 & -1 & 1 & -2 & 2 & -3 & 3 & \cdots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \cdots \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & \cdots \end{array}$$

它把所有自然数对应到自然数中的偶数，把所有负整数对应到自然数中的奇数。

不难验证，映射  $f$  是单射。这是因为，给定两个不相同的正数  $m \neq n$ ，如果两者一个小于零，一个不小于零，那么两者映射的结果奇偶性不同，因此不相等；如果两者同时小于零或同时不小于零，按定义  $2m \neq 2n$ 、 $-2m - 1 \neq -2n - 1$ ，也就是说两者映射的结果不相等。综上所述，如果  $m \neq n$ ，那么  $f(m) \neq f(n)$ 。

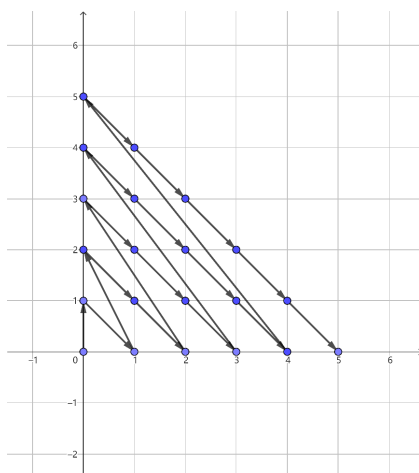
同时，映射  $f$  也是满射，因为任何偶数  $n$  都是整数  $\frac{n}{2}$  映射的结果，任何奇数  $n$  都是整数  $-\frac{n+1}{2}$  映射的结果。

$f$  既是单射也是满射，因此是双射。我们在自然数集  $\mathbb{N}$  和整数集  $\mathbb{Z}$  之间建立了双射  $f$ ，这说明  $|\mathbb{N}| = |\mathbb{Z}|$ 。

考虑平面所有坐标是自然数的点的集合： $\mathbb{N}^2 = \{(x, y) \mid x, y \in \mathbb{N}\}$ 。 $\mathbb{N}^2$  的势与  $\mathbb{N}$  的势关系如何呢？

我们希望找出一种按顺序数出  $\mathbb{N}^2$  中所有点的数数方法。这样，我们就可以把 1 映射到数数中的第一个点，把 2 映射到数数中的第二个点，等等，建立从  $\mathbb{Z}^+$  到  $\mathbb{N}^2$  的映射。这个映射如果是双射，那么  $\mathbb{N}^2$  的势就等于  $\mathbb{Z}^+$  的势，从而等于  $\mathbb{N}$  的势。

考虑这样的数数方法：把点的横坐标和纵坐标加起来。从和最小的点开始数起，逐步增大。对于和相等的点，则按照横坐标，从小数到大。也就是按照图中箭头的方法来数数。这样的数法，可以不重复不遗漏地数遍  $\mathbb{N}^2$  所有的点。



按照箭头方向，可以不重复不遗漏数遍  $\mathbb{N}^2$  中所有的点

相应地，把正整数  $n$  映射到这个数法的第  $n$  个点，这样的映射就是双射。因此，按前面的推理，我们得到结论：

$$|\mathbb{N}^2| = |\mathbb{N}|.$$

类似地，考虑正有理数集  $\mathbb{Q}^+$ 。对有理数  $r$ ，将它写成既约分数  $\frac{p}{q}$ ，考虑分子分母之和。从和最小的开始数起，逐步增加。如果某些有理数按以上方法得到的和相等，这样的有理数个数有限，把它们按分子从小到大排列，从分子最小的数起。这样，我们得到了一种不重复不遗漏数遍所有正有理

数的方法：

$$\frac{1}{1} \rightarrow \frac{1}{2} \rightarrow \frac{2}{1} \rightarrow \frac{1}{3} \rightarrow \frac{3}{1} \rightarrow \frac{1}{4} \rightarrow \frac{2}{3} \rightarrow \frac{3}{2} \rightarrow \frac{4}{1} \rightarrow \cdots$$

这表明  $\mathbb{Q}^+$  的势等于  $\mathbb{Z}^+$  的势，从而等于  $\mathbb{N}$  的势。

$$|\mathbb{Q}^+| = |\mathbb{N}|.$$

我们考虑把  $\mathbb{N}$  映射到  $\mathbb{Q}^+$  的双射  $f$ 。考虑以下映射：

$$g: r \mapsto \begin{cases} 0 & \text{如果 } r = 0 \\ f(r) & \text{如果 } r > 0 \\ -f(-r) & \text{如果 } r < 0 \end{cases}$$

$g$  把所有整数映射为有理数。由于  $f$  是双射，容易证明  $g$  也是双射。因此我们有  $\mathbb{Z}$  到  $\mathbb{Q}$  的双射，这说明  $\mathbb{Q}$  的势等于  $\mathbb{Z}$  的势，从而等于  $\mathbb{N}$  的势。

$$|\mathbb{Q}| = |\mathbb{N}|.$$

### 习题 1.2.1.

1. 给定数集  $A$  和正整数  $k \geq 2$ ，定义由  $k$  个  $A$  中元素构成的有序数组的集合为：

$$A^k = \{(a_1, a_2, \cdots, a_k) \mid a_1, a_2, \cdots, a_k \in A\}.$$

1.1. 考虑  $A = \mathbb{N}$  的情况。考虑  $\mathbb{N}^k$  中的有序数组  $(a_1, a_2, \cdots, a_k)$  中的  $k$  个数的和。记所有和等于自然数  $p$  的有序数组构成的集合为  $S_p$ 。证明：所有的  $S_p$  构成  $\mathbb{N}^k$  的分划。

1.2. 给定正整数  $k \geq 2$ ，考虑  $\mathbb{N}^{k+1}$  和  $\mathbb{N}^k$  按上一问构成的分划。如果  $|\mathbb{N}^k| = |\mathbb{N}|$ ，证明：可以给出一种不重复不遗漏数遍  $\mathbb{N}^{k+1}$  的方法。

1.3. 用归纳法证明：对任何正整数  $k \geq 2$ ， $|\mathbb{N}^k| = |\mathbb{N}|$ 。

1.4. 证明： $|\mathbb{Z}^k| = |\mathbb{N}|$ 。

2. 证明： $\mathbb{N}$  的子集要么有限，要么和  $\mathbb{N}$  等势。

3. 证明：如果  $|A| = |B| = |\mathbb{N}|$ ，那么  $|A \cup B| = |\mathbb{N}|$ 。

### 1.3 可数和不可数

上一节中，我们研究了一些常见数集的势。很多直观上似乎比自然数集“大得多”的数集，都和自然数集等势。那么，是否有比自然数集“大”的数集呢？

考虑数列的项只有 0 和 1 的数列。我们把所有这样的数列构成的集合记为  $2^{\mathbb{N}}$ ：

$$2^{\mathbb{N}} = \{ \{a_n\}_{n \in \mathbb{N}} \mid \forall n \in \mathbb{N}, a_n \in \{0, 1\} \}.$$

下面我们证明：

$$|\mathbb{N}| < |2^{\mathbb{N}}|$$

也就是说，存在  $\mathbb{N}$  到  $2^{\mathbb{N}}$  的单射，但不存在  $2^{\mathbb{N}}$  到  $\mathbb{N}$  的单射。

首先考虑映射：

$$\forall n \in \mathbb{N}, \quad n \mapsto \{a_k\}_{k \in \mathbb{N}}, \quad \text{其中 } a_k = 1 \text{ 当且仅当 } k = n.$$

这个映射显然是单射，因此存在  $\mathbb{N}$  到  $2^{\mathbb{N}}$  的单射。

如何证明不存在  $\mathbb{N}$  到  $2^{\mathbb{N}}$  的单射呢？我们用反证法证明。

假设存在  $2^{\mathbb{N}}$  到  $\mathbb{N}$  的单射  $f$ ，则  $2^{\mathbb{N}}$  所有元素经过映射得到的集合是  $\mathbb{N}$  的无穷子集，因而和  $\mathbb{N}$  等势。也就是说，我们可以假设  $f$  是满射，因而是双射。因此，我们考虑它的逆映射  $g$ 。 $g$  是把自然数映射到  $2^{\mathbb{N}}$  中元素的双

射。因此，我们可以把  $2^{\mathbb{N}}$  中的元素按数数的方式列出来：

$$\begin{array}{l}
 0 \mapsto a_{0,0}, a_{0,1}, a_{0,2}, \dots, a_{0,k}, \dots \\
 1 \mapsto a_{1,0}, a_{1,1}, a_{1,2}, \dots, a_{1,k}, \dots \\
 2 \mapsto a_{2,0}, a_{2,1}, a_{2,2}, \dots, a_{2,k}, \dots \\
 \vdots \\
 n \mapsto a_{n,0}, a_{n,1}, a_{n,2}, \dots, a_{n,k}, \dots \\
 \vdots
 \end{array}$$

其中  $a_{n,k}$  是  $n$  对应的数列中的第  $k$  项。现在考虑这么一个数列  $W$ ：它的第  $k$  项就是 1 减去上面排列中  $k$  对应的数列的第  $k$  项。也就是说，如果该项是 0， $W$  的第  $k$  项就是 1，如果该项是 1， $W$  的第  $k$  项就是 0。按定义， $W$  的第  $k$  项一定不等于上面排列中第  $k$  行数列的第  $k$  项。

按照  $g$  的定义， $W$  肯定是某个自然数  $m$  经过  $g$  得到的结果，因此，它排在上面排列中的第  $m$  行。但是，它的第  $m$  项就是  $a_{m,m}$ 。然而按照  $W$  的定义，它的第  $m$  项应该是  $1 - a_{m,m}$ ，不等于  $a_{m,m}$ 。这就构成了矛盾。

因此，不存在  $\mathbb{N}$  到  $2^{\mathbb{N}}$  的单射。

这样，我们得到了一个严格“大于”自然数集的集合。

以上结果说明：即便是无穷集合，也有“大小之分”。为此，我们要给无穷集合做更精细的划分。注意到以上证明里，我们通过把  $2^{\mathbb{N}}$  中的元素用数数的方式列出来，而导出了矛盾。这个矛盾说明了  $2^{\mathbb{N}}$  这样的集合的本质：它是“不可数”的。

我们把自然数集这样可以用数数的方式列出来的集合（也就是与  $\mathbb{N}$  等势的集合）称为**可数集合**或**可列集合**，而把  $2^{\mathbb{N}}$  这样的集合称为**不可数集合**或**不可列集合**。而以上的推理说明，**不可数集合总大于可数集合**。

不可数集合是否也有类似无限旅馆这样的现象呢？

对于可数集合，它的子集如果是无限的，就和它等势。换句话说，如果从可数集合中去掉有限个元素，得到的子集和它等势。用不严谨的话来说，这是由于有限集合相比可数集合是“非常小”的，是可以忽略的。

这个关系在可数集合与不可数集合之间也有体现。

**定理 1.3.1.** 设  $S$  是不可数集合，它的子集  $A$  是可数集合，则  $S$  去除  $A$  中元素得到的集合  $S \setminus A$  与  $S$  等势。

**证明：** 首先用反证法证明  $S \setminus A$  是不可数集合。 $S$  是  $A$  与  $S \setminus A$  的并集。如果  $A$  和  $S \setminus A$  都是可数集合，那么它们的并集  $S$  也是可数集合。矛盾！

接下来证明  $|S \setminus A| = |S|$ 。考虑  $S \setminus A$  的可数子集  $B$ ，记  $S \setminus A$  去掉  $B$  中元素得到的集合为  $C$ ，则  $A$ 、 $B$ 、 $C$  是  $S$  的分划， $S = A \cup B \cup C$ 。

注意到由于  $S \setminus A$  是不可数集合， $B$  是可数集合，所以  $C$  也是不可数集合。另外，由于  $A$ 、 $B$  是可数集合，所以  $|A| = |B| = |\mathbb{N}|$ ，于是  $|A \cup B| = |\mathbb{N}|$ 。因此，存在从  $B$  到  $A \cup B$  的双射  $g$ 。

考虑映射：

$$f: x \mapsto \begin{cases} g(x) & \text{如果 } x \in B \\ x & \text{如果 } x \in C \end{cases}$$

则  $f$  是  $S \setminus A$  到  $S$  的双射。因此  $|S \setminus A| = |S|$ 。

□

这说明，可数集合比起不可数集合，就和有限集合相比可数集合一样，是可以忽略的。

最后来看实数集  $\mathbb{R}$ 。它是否可数呢？我们可以用证明  $2^{\mathbb{N}}$  类似的想法。

给定实数  $x$ ，我们可以将它写成小数的形式。如果  $x$  是有穷小数，就在最后补上无穷多个 0。这样，每个实数都可以写成无穷小数。也就是说，我们可以把每个实数对应到类似 0 到 9 组成的无穷数列。

假设实数集可数，那么可以像前面的证明里那样，构造正整数集到实数集的双射，把所有实数依次排列出来。各个实数的小数部分就和  $2^{\mathbb{N}}$  里的数列一样。于是，我们构造这样的实数  $a$ ，它的小数部分第  $k$  位数字和排第  $k$  位的实数的小数部分第  $k$  位数字不一样。但另一方面， $a$  也是某个正整数  $m$  映射的结果。因此按照定义， $a$  的小数部分第  $m$  位不能等于自己。矛盾！因此我们可以得出结论：实数集不可数。

实际上，我们可以严谨证明：实数集  $\mathbb{R}$  和  $2^{\mathbb{N}}$  等势，具体参见附录 A。

研究了常见的无穷数集的势，我们发现，目前我们所知的无穷集合有两类。一类数集与自然数集  $\mathbb{N}$  等势，是为可数集合。另一类与实数集  $\mathbb{R}$  等势，等于  $2^{\mathbb{N}}$ 。那么，是否有既不等势于  $\mathbb{N}$ ，也不等势于  $2^{\mathbb{N}}$  的无穷集合呢？

数学研究者对无穷集合的研究发现：比  $2^{\mathbb{N}}$  “更大” 的无穷集合是存在的。我们把这些类别按从小到大的顺序，记为  $\aleph_0$ 、 $\aleph_1$ 、 $\aleph_2$ 、 $\aleph_3$  等等。 $\aleph_0$  就是自然数集  $\mathbb{N}$ ， $\aleph_1$  是  $2^{\mathbb{N}}$ 。 $\aleph_2$ 、 $\aleph_3$  等则是比  $2^{\mathbb{N}}$  更大的集合。

另一方面，是否存在介于  $\mathbb{N}$  和  $2^{\mathbb{N}}$  之间的无穷集合，则是困难得多的问题。由于实数集是连续的，我们把这个问题称为“连续统问题”或“连续统假设”。对这个问题的研究直接引发了对数学基本推理框架的质疑。

1963 年，数学家证明了：在我们常见的推理框架内，“连续统问题”是“独立的”，既不可能证明它成立，也不可能证明它不成立。从另一个角度来说，不论“存在介于  $\mathbb{N}$  和  $2^{\mathbb{N}}$  之间的无穷集合”还是“不存在介于  $\mathbb{N}$  和  $2^{\mathbb{N}}$  之间的无穷集合”，都不会导致矛盾。

### 习题 1.3.1.

1. 证明：可数集合去除有限个元素后仍然是可数集合。
2. 证明：无理数集是不可数集合。
3. 证明：可数个两两不相交的有限集合： $\{A_n\}_{n \in \mathbb{N}}$  的并集是可数集合。
4. 证明：有限个可数集合的并集是可数集合；可数个可数集合的并集是可数集合；不可数集合的并集是不可数集合。





## 第二章 连续函数的变化

日常生活、工程和科学研究中，我们关心事物的运动和变化。描述、衡量、分析事物的运动和变化，是人类了解世界、科学进步的关键。17 世纪，科学研究者发现了物体受力与运动变化的关系，建立了统一的力学理论，引发了工业革命。20 世纪初，科学研究者提出了“光速不变”的假设，在此基础上构建了相对论。这些革命性的进步，都离不开对运动本质的探索。而对运动的深入研究，也促使了数学的发展。

数学中，我们用实变映射描述事物的运动和变化。具体来说，我们需要研究的事物性质一些基本要素相关。这些基本要素，比如时间、物体的位置、温度等等，我们认为是连续变化的，称为变量，用实数表示。于是，事物的性质就是关于这些变量的映射。分析这些映射的性质，找到适合描述实验数据的映射，是科学研究的重要部分。

### 2.1 函数在一点的变化

我们已经学习过用数学描述运动和变化。对于物体的运动，我们定义了平均速度，描述运动的快慢。一般来说，我们使用变率来描述事物变化的快慢。

**定义 2.1.1. 实变映射的变率** 设  $f$  是定义在实数集上的映射, 给定实数  $t_1 < t_2$ , 则  $f$  在  $[t_1, t_2]$  上的变率是:

$$\frac{f(t_2) - f(t_1)}{t_2 - t_1}.$$

平均速度就是位移函数的变率。

使用平均速度, 我们可以近似描述物体位置在一段时间里的变化。如果把物体的运动看作关于时间  $t$  的函数, 设物体的位移为函数  $p(t)$ , 那么, 在  $(t_1, t_2)$  的时间段里, 物体的位移大致是:

$$p(t) \approx p(t_1) + \frac{p(t_2) - p(t_1)}{t_2 - t_1}(t - t_1).$$

这里我们用一次函数代替了真实的位移函数  $p$ 。直观上, 在区间  $(t_1, t_2)$  里, 我们用直线近似表示了函数  $p(t)$  的图像。

很多时候, 我们希望对事物的变化有更好的理解。比如, 我们可以测量一秒内物体的位移, 作为平均速度。但我们还想知道, 如果把一秒换成更短的 0.1 秒、0.01 秒, 得到的结果会不会完全不同。此外, 我们可以记录运动物体在某个时刻的位移, 但也希望能记录物体在该时刻的速度 (或者其他变化), 以更好地理解运动的性质。甚至, 我们发现物体受到的力与物体运动速度的变率相关, 也就是说, 和物体位移的变率的变率相关。这时候, 我们希望能更精确地知道物体运动速度的性质, 以研究它的变率。总之, 我们需要一个刻画物体在某个时刻 “附近” 变化快慢的量。

实践中, 我们发现, 对于大多数的运动物体, 如果在同样条件下重复测量平均速度, 那么随着取的时间间隔越来越短, 测得的平均速度会趋于某个固定的值。这个现象让我们想到函数 (在一点) 的极限。因此, 我们用极限的概念来描述物体某个时刻 “附近” 变化的快慢。

我们可以定义物体在某个时刻  $t_0$  的瞬时速度  $v(t_0)$ :

$$v(t_0) = \lim_{t \rightarrow t_0} \frac{p(t) - p(t_0)}{t - t_0}.$$

它是  $t$  附近的平均速度的极限。正如数列的极限不一定等于数列自身的项，瞬时速度作为变率的极限，描述了物体在该时刻运动的快慢，但它并非变率，也不是运动。

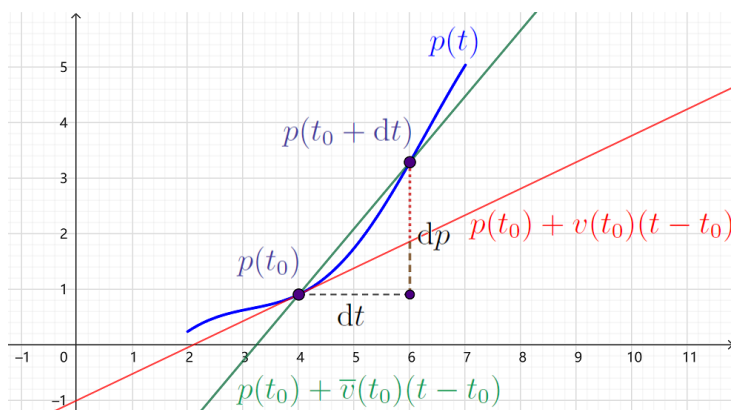
对于一般的函数，我们也可以用这个方法描述它在一点变化的快慢。

**定义 2.1.2.** 给定在某一点  $a$  附近有定义的函数  $f$ ，如果当  $x$  趋于  $a$  时， $f$  在  $a$  到  $x$  的变率收敛到某个极限，就说函数  $f$  在  $a$  处可微。我们把这个极限叫做函数  $f$  在点  $a$  处的微变率，简称微变，记作  $\partial f(a)$ 。<sup>1</sup>

$$\partial f(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}.$$

比如，运动物体在某时刻的瞬时速度，就是它的位移函数在该时刻的微变率。

$$v(t_0) = \partial p(t_0).$$



如何理解瞬时速度呢？上图是物体做直线运动时，位移关于时间的函数的图像。横坐标表示时间  $t$ ，纵坐标表示物体的位移  $p$ 。我们希望了解物体在  $t_0$  时刻“附近”的运动情况。

<sup>1</sup>函数  $f$  在点  $a$  处的微变率，一般记作  $\partial f(a)$  或  $f'(a)$ ，在物理书籍中也常记作  $\dot{f}(a)$ 。关于微变率的记法，可见附录 B。

从  $t_0$  时刻起, 经过固定时段  $dt$ , 物体的位移产生了变化:

$$dp = p(t_0 + dt) - p(t_0).$$

因此, 这段时间内的平均速度  $\bar{v}(t_0)$  就是  $dp$  与  $dt$  的比值:

$$\bar{v}(t_0) = \frac{dp}{dt}.$$

从图像来看, 它是过函数两点构成的绿色直线的斜率。

使用平均速度, 我们可以认为, 在  $t_0$  附近, 物体大致在做速度为  $\bar{v}(t_0)$  的匀速运动, 位移可以用一次函数近似表示:

$$p(t) \approx p(t_0) + \bar{v}(t_0)(t - t_0)$$

直观来说, 图中函数  $p(t)$  的图像曲线, 在  $t_0$  附近, 可以用绿色直线近似表示。

但是, 要注意的是, 平均速度  $\bar{v}(t_0)$  不仅仅与  $t_0$  相关。对不同的  $dt$ ,  $\bar{v}(t_0)$  是不同的。而瞬时速度  $v(t_0)$  的存在告诉我们,  $\bar{v}(t_0)$  会随着  $dt$  缩小而收敛。也就是说, 绿色直线会逐渐收拢到过点  $(t_0, p(t_0))$ 、以  $v(t_0)$  为斜率的红色直线。我们称这条直线为函数图像在  $t_0$  的**切线**。

使用瞬时速度  $v(t_0)$  来近似描述  $t_0$  附近的运动, 物体的位移可以用一次函数近似表示:

$$p(t) \approx p(t_0) + v(t_0)(t - t_0).$$

这样的表示和之前有什么不同呢?

首先, 我们注意到, 瞬时速度  $v(t_0)$  只与  $t_0$  相关, 不需要用别的时刻  $t$  来计算。其次, 我们可以证明, 如果要用一次函数 (也就是匀速运动) 来近似表示物体在  $t_0$  附近的运动, 瞬时速度  $v(t_0)$  是“最好”的系数。

直观来说, 如果用一条过  $(t_0, p(t_0))$  点的直线近似表示物体位移的曲线, 那么对于  $t_0$  附近的情况, 斜率为  $v(t_0)$  的直线是“最好”的。

为什么这么说呢？

假设我们用某个数  $u$  做系数，用这样的一次函数：

$$t \mapsto p(t_0) + u \cdot (t - t_0).$$

来近似表示物体的位移。考虑  $t_0$  附近的  $t$ ，实际的位移是  $p(t)$ ，近似的位移是  $p(t_0) + u(t - t_0)$ 。因此近似误差为：

$$d(t) = |p(t) - p(t_0) - u(t - t_0)|.$$

这是一个关于  $t$  的函数。 $t$  趋于 0 时， $d(t)$  趋于 0。

让我们来研究它趋于 0 有多快。我们这样来衡量：以  $t \mapsto t - t_0$  为参照物，考虑  $d(t)$  和  $|t - t_0|$  的比值：

$$\frac{d(t)}{|t - t_0|} = \left| \frac{p(t) - p(t_0) - u(t - t_0)}{t - t_0} \right| = \left| \frac{p(t) - p(t_0)}{t - t_0} - u \right|.$$

如果这个比值在  $t$  趋于  $t_0$  的时候的极限是 0，就说明只要  $t$  与  $t_0$  足够近，近似误差  $d(t)$  就可以比  $t - t_0$  小得多。这就是说  $d(t)$  比  $t - t_0$  收敛得更快。如果比值不趋于 0，就说明  $d(t)$  并不比  $t - t_0$  收敛得更快。

考虑这个比值在  $t$  趋于  $t_0$  时的极限：

$$\lim_{t \rightarrow t_0} \frac{d(t)}{|t - t_0|} = \left| \lim_{t \rightarrow t_0} \frac{p(t) - p(t_0)}{t - t_0} - u \right| = |v(t_0) - u|.$$

因此，这个比值趋于 0，当且仅当  $u$  等于瞬时速度  $v(t_0)$ 。

换句话说，当且仅当  $u = v(t_0)$  时，近似误差  $d(t)$  比  $t - t_0$  收敛得更快。只要  $t$  与  $t_0$  足够近，近似误差就可以比  $t - t_0$  小得多。而  $v$  取其他值的时候，就没有这样的效果，近似误差至多和  $t - t_0$  收敛得一样快。也就是说， $u = v(t_0)$  时，近似效果是最好的。在  $t_0$  附近，用微变率作为系数的一次函数和原来的函数最像。比如，我们取  $t_0$  处的瞬时速度  $v(t_0)$  作为  $t_0$  附近的“平均速度”，就比取其他的平均速度更能表现  $t_0$  附近的运动。

一般来说, 用微变率作为系数的一次函数:

$$t \mapsto f(t_0) + \partial f(t_0) \cdot (t - t_0)$$

称为函数  $f$  在  $t_0$  处的**微直观**。直观上, 微直观就是函数  $f$  在  $t_0$  处的切线, 是  $t_0$  附近模拟  $f$  图像曲线的最佳直线。

最后来看如何具体计算微变率。

从最简单的函数出发。常函数  $x \mapsto c$  在任意点的微变都是 0。恒等函数  $x \mapsto x$  在任意点的微变都是 1。正比例函数  $x \mapsto cx$  在任意点的微变都是系数  $c$ 。以上的结论就由读者来证明。

对于稍微复杂一点的函数, 我们一起来算一算。

### 例题 2.1.1.

1. 求函数  $f: x \mapsto x^2$  在点  $x = 3$  处的微变率。
2. 求函数  $f: x \mapsto \frac{1}{x}$  在点  $x = 2$  处的微变率。
3. 求函数  $f: x \mapsto (x - 1)^3$  在点  $x = 1$  处的微变率。

解答.

1. 按定义,  $f$  在点  $x = 3$  处的微变率为:

$$\begin{aligned} \partial f(3) &= \lim_{h \rightarrow 0} \frac{f(3+h) - f(3)}{h} \\ &= \lim_{h \rightarrow 0} \frac{(3+h)^2 - 3^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{6h + h^2}{h} \\ &= \lim_{h \rightarrow 0} (6 + h) \\ &= 6. \end{aligned}$$

2. 按定义,  $f$  在点  $x = 2$  处的微变率为:

$$\begin{aligned}
 \partial f(2) &= \lim_{h \rightarrow 0} \frac{f(2+h) - f(2)}{h} \\
 &= \lim_{h \rightarrow 0} \frac{\frac{1}{2+h} - \frac{1}{2}}{h} \\
 &= \lim_{h \rightarrow 0} \frac{\frac{2-(2+h)}{2(2+h)}}{h} \\
 &= \lim_{h \rightarrow 0} -\frac{1}{2(2+h)} \\
 &= -\frac{1}{2(2 + \lim_{h \rightarrow 0} h)} \\
 &= -\frac{1}{4}.
 \end{aligned}$$

3. 按定义,  $f$  在点  $x = 1$  处的微变率为:

$$\begin{aligned}
 \partial f(2) &= \lim_{h \rightarrow 0} \frac{f(1+h) - f(1)}{h} \\
 &= \lim_{h \rightarrow 0} \frac{h^3 - 0}{h} \\
 &= \lim_{h \rightarrow 0} h^2 \\
 &= 0.
 \end{aligned}$$

### 思考 2.1.1.

1. 在关于圆的章节中, 我们定义: 直线与圆恰有一个公共点, 是为相切, 公共点为切点。这个定义与本节中切线的定义相同吗? 是否有矛盾的地方?

2. 函数在一点可微, 是否需要在该点有定义? 是否需要在该点有极限? 是否需要在该点连续?

### 习题 2.1.1.

1. 证明本节提到的关于常函数、恒等函数、正比例函数的微变率的结论。

2. 求以下函数在给定点处的微变率。

2.1.  $f: x \mapsto x^4$  在点  $x = 2$  处的微变率。

2.2.  $f: x \mapsto x^2 - 3x + 1$  在点  $x = 3$  处的微变率。

2.3.  $f: x \mapsto \frac{1}{1-2x}$  在点  $x = -1$  处的微变率。

3. 已知函数  $f$  在  $a$  点可微, 证明: 函数  $f$  在  $a$  点连续。

4. 我们这样定义函数  $f$  在  $a$  点**左可微**: 函数  $f$  在  $a$  点左侧附近有定义。如果当  $x < a$  趋于  $a$  时,  $f$  从  $a$  到  $x$  的变率收敛到某个极限, 就说函数  $f$  在  $a$  点左可微, 称该极限为  $f$  在  $a$  点的**左微变率**或**左微变**, 记作:

$$\partial_- f(a) = \lim_{x \rightarrow a^-} \frac{f(x) - f(a)}{x - a} = \lim_{\substack{x \rightarrow a \\ x < a}} \frac{f(x) - f(a)}{x - a}.$$

4.1. 按照左可微的定义, 定义**右可微**。

4.2. 证明: 函数  $f$  在  $a$  点可微, 当且仅当它在  $a$  处左可微且右可微, 且左右微变相等。这时  $f$  在  $a$  点的微变就是相等的左微变和右微变。

4.3. 考虑绝对值函数:  $f: x \mapsto |x|$ 。它在  $0$  处是否可微?

5. 已知函数  $f$  在  $a$  附近有定义, 在  $a$  处可微, 微变率大于零。证明: 存在  $x_1 < a < x_2$ , 使得  $f(x_1) < f(a) < f(x_2)$ 。

6. 已知函数  $f$  在点  $a$  处可微, 证明以下函数在  $a$  处有极限:

$$x \mapsto \frac{xf(a) - af(x)}{x - a}.$$

7. 已知函数在包含  $0$  的开区间  $I$  上连续, 并且有

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

证明:  $f$  在  $0$  处可微。

## 2.2 微变的运算法则

已知函数的表达式, 如何具体计算它在一点的微变率呢? 与函数在一点的极限一样, 我们可以从研究最简单的函数的微变率开始, 通过四则运



算得到更复杂的函数的微变率。为此，我们先来了解函数的运算与它（在一点的）微变的关系。

首先来看加减法。给定函数  $f$ 、 $g$  和实数  $a$ 。设  $f$ 、 $g$  在  $a$  处可微，微变率为  $\partial f(a)$ 、 $\partial g(a)$ 。来看  $f + g$ 、 $f - g$  在  $a$  处是否可导，为此，研究  $f \pm g$  变率的极限：

$$\begin{aligned}\lim_{x \rightarrow a} \frac{(f \pm g)(x) - (f \pm g)(a)}{x - a} &= \lim_{x \rightarrow a} \left( \frac{f(x) - f(a)}{x - a} \pm \frac{g(x) - g(a)}{x - a} \right) \\ &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \pm \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\ &= \partial f(a) \pm \partial g(a)\end{aligned}$$

由此可见， $f + g$ 、 $f - g$  在  $a$  处也可微，微变率分别是  $\partial f(a)$ 、 $\partial g(a)$  的和与差。

$$\partial(f \pm g)(a) = \partial f(a) \pm \partial g(a)$$

接下来看乘法。同样地，研究  $f \cdot g$  变率的极限：

$$\begin{aligned}\lim_{x \rightarrow a} \frac{(f \cdot g)(x) - (f \cdot g)(a)}{x - a} &= \lim_{x \rightarrow a} \frac{f(x) \cdot g(x) - f(a) \cdot g(x) + f(a) \cdot g(x) - f(a) \cdot g(a)}{x - a} \\ &= \lim_{x \rightarrow a} \left( g(x) \cdot \frac{f(x) - f(a)}{x - a} + f(a) \cdot \frac{g(x) - g(a)}{x - a} \right) \\ &= \lim_{x \rightarrow a} g(x) \cdot \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} + f(a) \cdot \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\ &= g(a) \cdot \partial f(a) + f(a) \cdot \partial g(a)\end{aligned}$$

可以看到， $f \cdot g$  在  $a$  处可微。不过，它的微变率并不是  $\partial f(a)$ 、 $\partial g(a)$  的乘积。

$$\partial(f \cdot g)(a) = g(a) \cdot \partial f(a) + f(a) \cdot \partial g(a)$$

再来看除法。研究  $f \div g$  变率的极限<sup>2</sup>:

$$\begin{aligned}
 & \lim_{x \rightarrow a} \frac{(f \div g)(x) - (f \div g)(a)}{x - a} = \lim_{x \rightarrow a} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x - a} \\
 &= \lim_{x \rightarrow a} \frac{1}{g(x)g(a)} \cdot \frac{f(x)g(a) - f(a)g(x)}{x - a} \\
 &= \frac{1}{\lim_{x \rightarrow a} g(x)g(a)} \cdot \lim_{x \rightarrow a} \frac{f(x)g(a) - f(a)g(a) - (f(a)g(x) - f(a)g(a))}{x - a} \\
 &= \frac{1}{g(a)^2} \cdot \left( g(a) \cdot \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} - f(a) \cdot \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \right) \\
 &= \frac{g(a) \cdot \partial f(a) - f(a) \cdot \partial g(a)}{g(a)^2}
 \end{aligned}$$

可以看到,  $f \div g$  在  $a$  处可微。不过, 和函数乘法一样, 它的微变率并不是  $\partial f(a)$  除以  $\partial g(a)$  的商。

$$\partial(f \div g)(a) = \frac{g(a) \cdot \partial f(a) - f(a) \cdot \partial g(a)}{g(a)^2}$$

综上所述, 函数的四则运算的微变, 并不是简单地把函数的微变作四则运算。下面来看更复杂一点的, 复合函数的微变率。

设  $f$ 、 $g$  是定义在实数集上的函数,  $a$  为实数。 $g$  在  $a$  处可微, 微变率为  $\partial g(a)$ ;  $f$  在  $g(a)$  处可微, 微变率为  $\partial f(g(a))$ 。那么复合函数  $f \circ g$  是否在  $a$  处可微呢?

来看  $f \circ g$  在  $a$  处变率的极限:

$$\begin{aligned}
 & \lim_{x \rightarrow a} \frac{(f \circ g)(x) - (f \circ g)(a)}{x - a} \\
 &= \lim_{x \rightarrow a} \left( \frac{f(g(x)) - f(g(a))}{x - a} \cdot \frac{g(x) - g(a)}{g(x) - g(a)} \right) \\
 &= \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\
 &= \partial f(g(a)) \cdot \partial g(a)
 \end{aligned}$$

---

<sup>2</sup>和以往一样, 为了让  $f \div g$  在  $a$  处有定义, 这里需要假设  $g(a) \neq 0$ 。

上面推导中，我们从  $\lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)}$  算出  $\partial f(g(a))$ ，是因为  $g$  在  $a$  处可微，从而连续，因此  $x$  趋于  $a$  时， $g(x)$  趋于  $g(a)$ 。因此，把  $g(x)$  整体看作变化量，

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} &= \lim_{t \rightarrow g(a)} \frac{f(t) - f(g(a))}{t - g(a)} \\ &= \partial f(g(a)) \end{aligned}$$

综上，复合函数  $f \circ g$  在  $a$  处可微，微变率为：

$$\partial(f \circ g)(a) = \partial f(g(a)) \cdot \partial g(a)$$

如果是多个函数的复合，比如三个函数  $f$ 、 $g$ 、 $h$ ，那么可以算出，上式变为：

$$\partial(f \circ g \circ h)(a) = \partial f(g(h(a))) \cdot \partial g(h(a)) \cdot h(a).$$

多次复合的微变是各次复合的微变的乘积，仿佛链条一样。这个结果被形象地称为“链式法则”。

最后来看反函数的微变率。设  $f$  是定义在实数集上的函数， $a$  为实数， $f$  在  $a$  处可微，微变率为  $\partial g(a)$ 。 $f$  在  $a$  附近有反函数  $g$ ，即对  $a$  附近的  $x$ ，总有  $f(g(x)) = g(f(x)) = x$ 。 $g(f(a)) = a$ 。那么反函数  $g$  在  $f(a)$  处是否可微呢？

假设  $g$  在  $a$  处可微。考虑  $f \circ g$ ，它是恒等函数，恒等函数在任意点的微变率是 1。因此，根据链式法则，我们有：

$$1 = \partial(g \circ f)(a) = \partial g(f(a)) \cdot \partial f(a)$$

也就是说，

$$\partial g(f(a)) = \frac{1}{\partial f(a)}$$

$g$  在  $f(a)$  处的微变率是  $\partial f(a)$  的倒数。这要求  $\partial f(a)$  不能为零。

假设  $\partial f(a)$  不为零, 计算  $g$  在  $f(a)$  处变率的极限:

$$\begin{aligned}
 & \lim_{x \rightarrow f(a)} \frac{g(x) - g(f(a))}{x - f(a)} \\
 &= \lim_{x \rightarrow f(a)} \frac{g(x) - a}{x - f(a)} \\
 &= \lim_{g(x) \rightarrow a} \frac{g(x) - a}{f(g(x)) - f(a)} \\
 &= \frac{1}{\lim_{g(x) \rightarrow a} \frac{f(g(x)) - f(a)}{g(x) - a}} \\
 &= \frac{1}{\partial f(a)}
 \end{aligned}$$

上面推导中, 我们用到了反函数的连续性:  $f$  在  $a$  处连续, 因此  $g$  在  $f(a)$  处连续。因此  $x$  趋于  $f(a)$  时,  $g(x)$  趋于  $g(f(a))$ , 也就是  $a$ 。

### 例题 2.2.1.

1. 求以下函数在给定点处的微变率。

1.1.  $f: x \mapsto (x+1)(3x-4)$  在点  $x=2$  处的微变率。

1.2.  $f: x \mapsto \frac{x-1}{2x^2-x+1}$  在点  $x=1$  处的微变率。

1.3.  $f: x \mapsto (2x-1)^3$  在点  $x=-1$  处的微变率。

解答.

1. 应用函数乘法的求微法则:

$$\begin{aligned}
 \partial f(2) &= \partial(x+1)(2) \cdot (3 \cdot 2 - 4) + \partial(3x-4)(2) \cdot (2+1) \\
 &= 1 \cdot (3 \cdot 2 - 4) + 3 \cdot (2+1) \\
 &= 6 + 9 = 15
 \end{aligned}$$

其中  $\partial(x+1)(2)$  表示函数  $x \mapsto x+1$  在  $x=2$  处的微变率。 $\partial(3x-4)(2)$  同理。

2. 应用函数除法的求微法则:

$$\begin{aligned}\partial f(1) &= \frac{\partial(x-1)(1) \cdot (2 \cdot 1^2 - 1 + 1) - \partial(2x^2 - x + 1)(1) \cdot (1 - 1)}{(2 \cdot 1^2 - 1 + 1)^2} \\ &= \frac{1 \cdot 2 - 3 \cdot 0}{2^2} \\ &= \frac{1}{2}\end{aligned}$$

3. 应用复合函数的求微法则 (链式法则):

$$\begin{aligned}\partial f(-1) &= \partial(x^3)(2 \cdot (-1) - 1) \cdot \partial(2x - 1)(-1) \\ &= 3 \cdot (-3)^2 \cdot 2 \\ &= 54\end{aligned}$$

### 思考 2.2.1.

1. 在函数的除法和反函数的求微法则中, 都有要求取值不为零的情况。请对应函数图像, 给出这些要求的直观解释。

2. 对比函数在一点的极限, 函数在一点微变率的运算法则有什么不同? 你觉得为什么会有这样的不同。

3. 对于把实数变量映射到向量的映射, 能否定义在某个实数  $t$  的微变率? 如何在直观上解释你定义的微变率?

4. 对于把向量映射到向量的映射, 也就是点映射, 能否定义在某个点的微变率? 如何在直观上解释你定义的微变率?

### 习题 2.2.1.

1. 求以下函数在给定点处的微变率。

1.1.  $f: x \mapsto (2x - 1)(x - 4)(x^2 + 3)$  在点  $x = 2$  处的微变率。

1.2.  $f: x \mapsto \frac{(x+1)(x^2+1)}{x^3-2x+3}$  在点  $x = 1$  处的微变率。

1.3.  $f: x \mapsto (2x - \frac{1}{x+1})^5$  在点  $x = 0$  处的微变率。

2. 函数  $x \mapsto x^{\frac{1}{3}}$  在  $x = 0$  处是否可微?

3. 考虑以下函数:

$$f: x \mapsto \begin{cases} x & \text{如果 } x \text{ 为有理数} \\ -x & \text{如果 } x \text{ 为无理数} \end{cases}$$

- 3.1. 证明:  $f$  在  $x = 0$  处可微。  
 3.2. 证明:  $f$  在  $x = 1$  处不可微。  
 3.3. 找出  $f$  所有可微的点, 并给出证明。

## 2.3 常见函数的微变

使用上一节中的结论, 我们来研究一些较为复杂的常见函数的微变。

首先来看整式函数的微变。给定整式:  $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ , 其中  $a_0, a_1, \cdots, a_n$  是整式的系数。它是一系列单项式  $x^k$  乘以系数后相加的结果。因此, 我们可以先研究形如  $x^k$  的单项式。

给定自然数  $k$ ,  $x \mapsto x^k$  是  $k$  个  $x$  的乘积。对于  $k = 0$ 、 $k = 1$  的情况, 我们已经知道对应的微变:

$$\forall a, \quad \partial(x^0)(a) = 0, \quad \partial(x^1)(a) = 1$$

$k > 1$  时, 根据函数乘法的求微法则,

$$\partial(x^{k+1})(a) = \partial(x^k \cdot x)(a) = \partial(x^k)(a) \cdot a + 1 \cdot a^k.$$

用归纳法可以证明:

$$\forall k \in \mathbb{Z}^+, \quad a \in \mathbb{R}, \quad \partial(x^k)(a) = ka^{k-1}.$$

因此, 我们可以得出一般整式函数的微变:

$$\forall a, \quad \partial(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

接下来看幂函数  $x^r$  的微变。对自然数  $k$ , 函数  $x \mapsto x^{\frac{1}{k}}$  是单项式函数

$x \mapsto x^k$  的反函数, 所以根据反函数的求微法则:

$$\begin{aligned}\partial(x^{\frac{1}{k}})(a) &= \frac{1}{\partial(x^k)(a^{\frac{1}{k}})} \\ &= \frac{1}{ka^{\frac{k-1}{k}}} \\ &= \frac{1}{k}a^{\frac{1}{k}-1}\end{aligned}$$

要注意的是, 这里用到了反函数的求微法则, 所以, 根据对应的要求,  $a$  不能为 0。  $a = 0$  时, 函数不可微。

对于  $r = \frac{p}{q}$  为非零有理数时, 函数  $x \mapsto x^r$  可以看作函数  $x \mapsto x^{\frac{1}{q}}$  与函数  $x \mapsto x^p$  的复合函数, 因此, 根据复合函数的求微法则:

$$\begin{aligned}\partial(x^{\frac{p}{q}})(a) &= \partial(x^p)(a^{\frac{1}{q}}) \cdot \partial(x^{\frac{1}{q}})(a) \\ &= p \cdot a^{\frac{p-1}{q}} \cdot \frac{1}{q} \cdot a^{\frac{1}{q}-1} \\ &= \frac{p}{q}a^{\frac{p}{q}-1}\end{aligned}$$

综上可知, 对非零有理数  $r$ , 如果函数  $x \mapsto x^r$  在点  $a > 0$  的微变为:

$$\partial(x^r)(a) = ra^{r-1}.$$

再来看三角函数的微变。我们需要用到之前的结论:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

对于正弦函数  $\sin$ , 计算变率的极限:

$$\begin{aligned}\partial \sin(a) &= \lim_{h \rightarrow 0} \frac{\sin(a+h) - \sin a}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sin a \cos h + \cos a \sin h - \sin a}{h} \\ &= \sin a \cdot \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} + \cos a \cdot \lim_{h \rightarrow 0} \frac{\sin h}{h}\end{aligned}$$

这里我们要计算两个极限，第二个极限可以直接使用上面的结论，结果是1。对于第一个极限，我们把分子转化为关于  $\sin$  的表达式。

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{\cos h - 1}{h} &= \lim_{h \rightarrow 0} \frac{\cos^2 h - 1}{h(\cos h + 1)} \\ &= -\lim_{h \rightarrow 0} \frac{\sin h}{h} \cdot \lim_{h \rightarrow 0} \frac{\sin h}{\cos h + 1} \\ &= -1 \cdot \frac{0}{1 + 1} = 0.\end{aligned}$$

因此，正弦函数  $\sin$  在一点  $a$  的微变为：

$$\partial \sin(a) = \cos a$$

类似地，可以计算余弦函数  $\cos$  的微变。变率的极限：

$$\begin{aligned}\partial \cos(a) &= \lim_{h \rightarrow 0} \frac{\cos(a+h) - \cos a}{h} \\ &= \lim_{h \rightarrow 0} \frac{\cos a \cos h - \sin a \sin h - \cos a}{h} \\ &= \cos a \cdot \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} - \sin a \cdot \lim_{h \rightarrow 0} \frac{\sin h}{h} \\ &= \cos a \cdot 0 - \sin a \cdot 1 = -\sin a\end{aligned}$$

余弦函数  $\cos$  在一点  $a$  的微变为：

$$\partial \cos(a) = -\sin a$$

正切函数  $\tan$  可以看作正弦函数和余弦函数的比值，因此，它在一点  $a$  的微变为：

$$\begin{aligned}\partial \tan(a) &= \partial \left( \frac{\sin}{\cos} \right) (a) \\ &= \left( \frac{\partial \sin(a) \cdot \cos a - \partial \cos(a) \cdot \sin a}{\cos^2 a} \right) \\ &= \frac{\cos^2 a + \sin^2 a}{\cos^2 a} \\ &= \frac{1}{\cos^2 a}.\end{aligned}$$



正切函数  $\tan$  在一点  $a$  的微变为：

$$\partial \tan(a) = \frac{1}{\cos^2 a}$$

余切  $\cot$  是正切的倒数，使用求微法则可知，余切函数在一点  $a$  的微变为：

$$\partial \cot(a) = -\frac{1}{\sin^2 a}$$

最后来看指数函数。给定底数  $c > 1$ ，指数函数  $x \mapsto c^x$  是连续函数，计算它在一点  $a$  的微变：

$$\begin{aligned} \partial(c^x)(a) &= \lim_{h \rightarrow 0} \frac{c^{a+h} - c^a}{h} \\ &= c^a \lim_{h \rightarrow 0} \frac{c^h - 1}{h} \\ &= c^a \cdot \partial(c^x)(0) \end{aligned}$$

可以看到，如果指数函数在 0 处可微，那么它在任意点处可微，且微变率是函数值与函数在 0 处微变率的乘积。

于是，我们来研究指数函数在 0 处的微变。对于不同的底数  $c_1$ 、 $c_2$ ，有：

$$c_2^h = c_1^{\frac{\log c_2}{\log c_1} \cdot h}.$$

因此，

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{c_2^h - 1}{h} &= \lim_{h \rightarrow 0} \frac{c_1^{\frac{\log c_2}{\log c_1} \cdot h} - 1}{h} \\ &= \frac{\log c_2}{\log c_1} \cdot \lim_{h \rightarrow 0} \frac{c_1^{\frac{\log c_2}{\log c_1} \cdot h} - 1}{\frac{\log c_2}{\log c_1} \cdot h} \\ &= \frac{\log c_2}{\log c_1} \cdot \lim_{t \rightarrow 0} \frac{c_1^t - 1}{t} \end{aligned}$$

也就是说，如果底数  $c_1$  的指数函数在 0 处可微，那么底数  $c_2$  的指数函数在 0 处也可微，并且两个微变率只差一个乘法系数。

考虑通项公式如下的数列  $\{e_n\}$ :

$$\forall n \in \mathbb{Z}^+, \quad e_n = \left(1 + \frac{1}{n}\right)^n.$$

可以证明, 数列  $\{e_n\}$  有极限  $e$ , 而  $e$  满足 (见附录 B):

$$\lim_{t \rightarrow 0} \frac{e^t - 1}{t} = 1.$$

因此, 对任意底数  $c$ , 指数函数  $x \mapsto c^x$  在 0 处可微, 微变率是:

$$\lim_{h \rightarrow 0} \frac{c^h - 1}{h} = \log_e c.$$

从而, 指数函数  $x \mapsto c^x$  在  $a$  处的微变率为:

$$\partial(c^x)(a) = c^a \cdot \log_e c.$$

指数函数在任一点的微变率与它在该点的值成正比, 比值是以  $e$  为底数的对数  $\log_e c$ 。

特别来说, 以  $e$  为底数时, 指数函数  $f: x \mapsto e^x$  在任一点的微变率等于它在该点的值。

$$\partial f(a) = e^a = f(a).$$

对数函数是指数函数的反函数。因此, 可以用反函数的求微法则, 求出对数函数的微变。

$$\begin{aligned} \partial \log_c(a) &= \frac{1}{\partial(c^x)(\log_c a)} \\ &= \frac{1}{c^{\log_c a} \cdot \log_e c} = \frac{1}{a \cdot \log_e c} \end{aligned}$$

对数函数在定义域中任一点的微变率与它在该点的值成反比。

同样, 如果底数为  $e$ , 那么对数函数在任一点的微变率等于它在该点的值的倒数。

$$\partial \log_e(a) = \frac{1}{a}.$$

为此, 我们把以  $e$  为底数的对数函数称为**自然对数**, 记作  $\ln$ 。而  $e$  也叫做**自然对数的底数**。

**例题 2.3.1.**

1. 求以下函数在给定点处的微变率。
  - 1.1.  $f: x \mapsto \sin(2x - 7)$  在点  $x = 2$  处的微变率。
  - 1.2.  $f: x \mapsto \log_2\left(\frac{x+1}{x-1} + \sin(x)\right)$  在点  $x = 2$  处的微变率。
  - 1.3.  $f: x \mapsto 3^{2x - \frac{1}{\sin x + 2}}$  在点  $x = -1$  处的微变率。

**解答.**

1. 应用复合函数的求微法则:

$$\begin{aligned}\partial f(2) &= \partial \sin(2 \cdot 2 - 7) \cdot \partial(2x - 7)(2) \\ &= \cos(-3) \cdot 2 \\ &= 2 \cos 3\end{aligned}$$

2. 应用复合函数的求微法则:

$$\begin{aligned}\partial f(2) &= \partial \log_2\left(\frac{2+1}{2-1} + \sin 2\right) \cdot \left(\partial\left(\frac{x+1}{x-1}\right)(2) + \partial \sin(2)\right) \\ &= \frac{1}{\log_e 2 \left(\frac{2+1}{2-1} + \sin 2\right)} \cdot \left(-\frac{2}{(2-1)^2} + \cos 2\right) \\ &= \frac{\cos(2) - 2}{\ln 2 \cdot (3 + \sin 2)}\end{aligned}$$

3. 应用复合函数的求微法则:

$$\begin{aligned}\partial f(-1) &= \partial(3^x)(2 \cdot (-1) - \frac{1}{2 + \sin(-1)}) \cdot \left(\partial(2x)(-1) - \partial \frac{1}{\sin x + 2}(-1)\right) \\ &= \ln 3 \cdot 3^{-2 - \frac{1}{2 - \sin 1}} \cdot \left(2 - \frac{-1}{(\sin(-1) + 2)^2} \cdot \cos(-1)\right) \\ &= \frac{\ln 3}{3^{2 + \frac{1}{2 - \sin 1}}} \cdot \left(2 + \frac{\cos 1}{(2 - \sin 1)^2}\right)\end{aligned}$$

**习题 2.3.1.**

1. 求以下函数在给定点处的微变率。

1.1.  $f: x \mapsto \cos(3 - 5x)$  在点  $x = 2$  处的微变率。

1.2.  $f: x \mapsto 2^{(3-x)\log_2(x^2+1)+\frac{\sin(x)}{1+x^2}}$  在点  $x = 1$  处的微变率。

1.3.  $f: x \mapsto (\tan^2 x - e^{\frac{1}{1-x}})^{5-\ln(3-x)}$  在点  $x = 0$  处的微变率。

2. 求反正弦函数  $\arcsin$  和反余弦函数  $\arccos$  在一点  $a$  的微变。

3. 从结论

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1$$

出发, 证明:

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1.$$

4. 证明: 对任意非零实数  $r$ , 任意实数  $a > 0$ , 幂函数  $x \mapsto x^r$  在  $a > 0$  处可微, 微变率为  $ra^{r-1}$ 。

5. 设有函数

$$f: x \mapsto \begin{cases} x \sin \frac{1}{x} & \text{如果 } x \neq 0 \\ 0 & \text{如果 } x = 0 \end{cases}$$

5.1. 证明: 函数  $f$  在 0 处连续。

5.2. 证明: 函数  $f$  在 0 处不可微。

## 2.4 微变函数

从前面几节可以看出, 常见的显式函数, 在定义域中大致可微。因此, 我们可以引入微变函数的概念:

**定义 2.4.1. 微变函数**

如果函数  $f$  在集合  $D$  中各点都可微, 那么在  $D$  上可以定义它的微变函数:

$$\partial f: x \mapsto \partial f(x).$$

$f$  的微变函数把  $x$  映射到  $f$  在  $x$  处的微变率。

许多显式函数，其微变函数仍然是常见的显式函数。

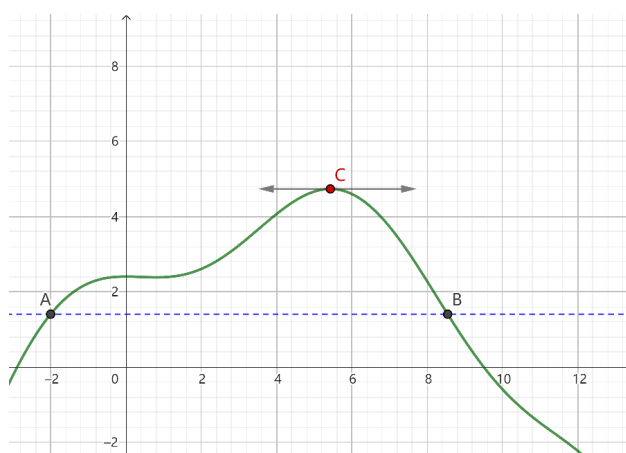
整式函数、分式函数的微变函数仍然是整式函数、分式函数。正弦函数、余弦函数的微变函数分别是余弦函数、正弦函数。正切、余切函数的微变函数仍然是三角函数。幂函数、指数函数的微变函数仍然是幂函数、指数函数。而对数函数的微变函数是反比例函数。

通过微变函数，我们可以讨论关于函数整体变化的一些性质。

**定理 2.4.1. 微变零值定理** 如果函数  $f$  在闭区间  $[a, b]$  上连续，在开区间  $(a, b)$  上可微，并且  $f(a) = f(b)$ ，那么，存在一点  $t \in (a, b)$ ，使得

$$\partial f(t) = 0.$$

这个定理让我们看到了“微变为零”的直观意义。函数在一点微变为 0，那么它在这一点的切线是水平的。也就是说，在这点附近的最佳近似直线是水平直线。函数曲线在这点附近几乎是水平的。



函数满足  $f(a) = f(b)$ ，于是会在某处附近变得水平。

那么，函数曲线怎样的情况下会变得水平呢？一种情况就是“掉头”，即函数从左到右逼近这一点时，函数值不断变大，但过了这一点，就不断变

小了。或者反过来，从左到右逼近这一点时，不断变小，但过了这一点，就不断变大了。

也就是说，函数在这一点附近达到局部的极大值或极小值时，会变得水平。

对于定理中提到的函数  $f$ ，我们知道  $f(a) = f(b)$ 。也就是说，把  $a$  看作起点，把  $b$  看作终点，那么函数经历了一段变化后，函数值又回到了起始的水平。那么，直觉上，要么函数值一直不变，要么总有“掉头”的一刻。比如说，如果函数值一开始不断变大，那么，为了在终点变回原值，一定要在某一点“掉头”变小。而该点就是微变为 0 的点。

要注意的是，这样的情形之所以存在，是因为函数满足“在闭区间  $[a, b]$  上连续，在开区间  $(a, b)$  上可微”的条件。否则，我们想象中的“不断变大”、“掉头”都有可能因为函数“跳跃”、“突转”而不存在。

下面给出定理的证明：

**证明：** 记  $f(a) = f(b) = v$ 。如果  $f$  在  $[a, b]$  上恒等于  $v$ ，那么任意挑选  $(a, b)$  中一点即可。

否则，由于  $f$  在  $[a, b]$  上连续，根据闭区间极值定理， $f$  在区间上有极大值  $v_{\text{大}}$  和极小值  $v_{\text{小}}$ 。而这两个值不全等于  $v$ 。不妨设极大值  $v_{\text{大}}$  不等于  $v$ 。设  $v_{\text{大}} = f(t)$ ，则  $t \in (a, b)$ 。

考虑  $t$  附近的变率：

$$\frac{f(x) - f(t)}{x - t}.$$

由于  $f$  在  $t$  处有极大值，所以总有  $f(x) \leq f(t)$ 。因此， $x < t$  时，变率大于等于 0； $x > t$  时，变率小于等于 0。 $\partial f(t)$  是  $x$  趋于  $t$  时，该变率的极限。根据极限传递不等关系的性质，变率在该点的左极限大于等于 0，右极限小于等于 0。而  $f$  在该点可微，因此左极限等于右极限，故都等于 0。

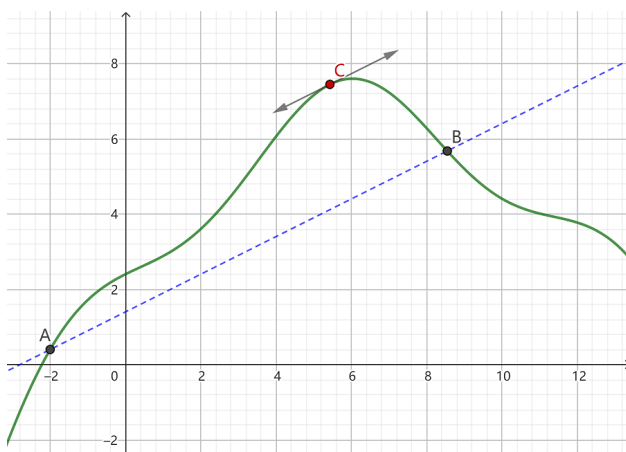
这就说明,  $\partial f(t) = 0$ . □

我们研究的函数, 不一定能满足  $f(a) = f(b)$  的条件。不过, 我们可以用一点技巧, 把这个结论拓展到更一般的情况。

**定理 2.4.2. 微变中值定理** 如果函数  $f$  在闭区间  $[a, b]$  上连续, 在开区间  $(a, b)$  上可微, 那么, 存在一点  $t \in (a, b)$ , 使得

$$\partial f(t) = \frac{f(b) - f(a)}{b - a}.$$

去掉了  $f(a) = f(b)$  的条件, 相当于允许函数在该段区间有一个整体的变化趋势。我们可以把这个整体的变化趋势用变率表示:  $k = \frac{f(b) - f(a)}{b - a}$ 。而中值定理告诉我们, 区间中有一点, 函数在其附近的变化趋势和整体趋势一样。



可以把  $f(a) \neq f(b)$  视为考虑了整体变化的趋势。

如果令  $f(a) = f(b)$ , 那么  $k = 0$ , 这时我们得到微变零值定理。如何把一般情况变为这种特殊情况呢? 我们可以用过  $f$  在区间两端点的直线, 大致描述函数的整体变化趋势:

$$y = f(a) + k(x - a).$$

$x = a$  时,  $y = f(a)$ ,  $x = b$  时,  $y = f(b)$ 。那么, 在函数中“剔除”这个趋势即可。

我们把函数“拆分”成两部分, 一部分描述整体趋势, 一部分描述函数在其中的“内部变化”。

$$f(x) = f(a) + k(x - a) + g(x).$$

其中  $g$  的定义就是  $g(x) = f(x) - f(a) - k(x - a)$ 。

那么,  $f(a) + k(x - a)$  就是  $f$  的趋势部分,  $g$  就是内部变化部分。 $g(a) = g(b) = 0$ 。对  $g$  运用微变零值定理可知, 存在  $t \in (a, b)$ , 使得  $\partial g(t) = 0$ 。然而,  $\partial g = \partial f - k$ 。也就是说,  $\partial f(t) - k = 0$ , 即:

$$\partial f(t) = k.$$

于是我们找到了满足要求的  $t$ , 微变中值定理成立。

以上推理过程中, 我们看到了函数变化和微变函数的关系。函数的单调性也能用微变描述。

**定理 2.4.3. 微变单调定理** 如果函数  $f$  在区间  $I$  上可微, 那么  $f$  在  $I$  上单调递增 (递减), 当且仅当它在  $I$  上的微变总大于等于 (小于等于) 零。

**证明:** 如果  $f$  在  $I$  上单调递增, 那么变率总大于等于 0, 因此微变作为变率的极限也总大于等于 0。

同理,  $f$  在  $I$  上单调递减, 那么微变总小于等于 0。

反之, 如果  $f$  在  $I$  上的微变总大于等于 0, 那么, 对于  $I$  中任意两点  $a < b$ , 根据微变中值定理, 总有  $c \in (a, b)$  使得:

$$f(b) - f(a) = \partial f(c) \cdot (b - a).$$

而等式右边的项都大于等于 0, 所以  $f(b) \geq f(a)$ 。这说明  $f$  在  $I$  上单调递增。



同理, 如果  $f$  在  $I$  上的微变总大于等于 0, 那么  $f$  在  $I$  上单调递减。

□

### 思考 2.4.1.

1. 微变零值定理中, 如果去掉 “在  $(a, b)$  上可微”, 结论是否还成立? 如果不成立, 举一个反例。

2. 微变单调定理中, 如果要使得  $f$  在  $I$  上严格单调递增 (递减), 需要什么条件?

3. 函数在一点微变率为 0, 是否说明该点是  $f$  在附近局部的极小值或极大值?

### 习题 2.4.1.

1. 证明: 每个整式函数都是某个函数的微变函数。

2. 已知函数  $f$  为:

$$f: x \mapsto \frac{\sin x + \cos x}{2 + \cos^2 x - \sin x}.$$

证明: 对任意实数  $a$ ,  $f$  在区间  $(a, a + 2\pi)$  上总有一处微变为 0。

3. 设函数  $f, g$  在区间  $[a, b]$  上连续, 在  $(a, b)$  上可微。  $g(a) \neq g(b)$ , 且  $f, g$  在任一点的微变率不同时为 0。

3.1. 设  $k_f$  为  $f$  从  $a$  到  $b$  的变率,  $k_g$  为  $g$  从  $a$  到  $b$  的变率, 作函数:

$$h: x \mapsto k_g \cdot (f(x) - f(a)) - k_f \cdot (g(x) - g(a))$$

证明:  $h$  在区间  $[a, b]$  上连续, 在  $(a, b)$  上可微, 且  $h(a) = h(b) = 0$ 。

3.2. 证明: 存在  $t \in (a, b)$ , 使得:

$$\frac{\partial f(t)}{\partial g(t)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

4.1. 证明:

$$\forall x > 0, \quad \frac{1}{x+1} < \ln(x+1) - \ln x < \frac{1}{x}.$$

4.2. 证明: 函数  $f: x \mapsto (1 + \frac{1}{x})^x$  在  $(0, \infty)$  上单调递增, 函数  $g: x \mapsto (1 + \frac{1}{x})^{x+1}$  在  $(0, \infty)$  上单调递减。(提示: 考虑复合函数  $\ln f$  和  $\ln g$ .)

4.3. 证明:  $x$  趋于正无穷时, 函数  $f, g$  有极限, 并给出这个极限的值。

4.4. 设数列  $\{v_n\}_{n \in \mathbb{N}}$  的通项为:

$$\forall n \in \mathbb{N}, v_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}.$$

证明:

$$\ln(2n+1) - \ln(n+1) < v_n < \ln(2n) - \ln(n).$$

从而证明数列  $\{v_n\}$  收敛, 并给出它的极限。

5. 设有界函数  $f$  在实数集上可微, 且它的微变函数  $\partial f$  在正无穷有极限  $c$ 。证明:  $c = 0$ 。(提示: 使用反证法。)

6. 设函数  $f$  在  $a$  处可微, 且微变  $\partial f(a) \neq 0$ 。

6.1. 证明: 存在  $h > 0$ , 使得对区间  $(a-h, a+h)$  中任意  $x$ ,  $f(x) = f(a)$  当且仅当  $x = a$ 。

6.2. 如果函数  $f$  还在  $a$  附近可微, 且微变函数  $\partial f$  在  $a$  处连续, 证明: 存在  $h > 0$ , 使得  $f$  在区间上是单射。

## 2.5 多次微变

既然对可微的函数, 可以定义它的微变函数, 那么, 如果微变函数仍然可微, 我们就可以继续定义它的微变函数。也就是说, 只要条件允许, 我们可以定义函数的多次微变函数。

举例来说, 正弦函数  $\sin$  的微变函数是余弦函数  $\cos$ 。而余弦函数  $\cos$  的微变函数是  $-\sin$ 。因此, 我们可以说, 正弦函数  $\sin$  的二次微变函数是  $-\sin$ 。对于一点  $a$ , 我们说正弦函数在  $a$  处二次可微, 二次微变率为  $\sin a$ 。

一般来说, 我们可以这样定义:

**定义 2.5.1. 多次微变**

设函数  $f$  在区间  $I$  上可微。如果它的微变函数  $\partial f$  在  $I$  上仍然可微，就说  $f$  二次可微。 $\partial f$  在  $I$  上的微变函数称为  $f$  的二次微变函数，简称二次微变，记作  $\partial^2 f$ 。

依此类推，对正整数  $k$ ，如果  $f$  的  $k$  次微变函数  $\partial^k f$  在  $I$  上仍然可微，就说  $f$  是  $k+1$  次可微的。 $\partial^k f$  的微变函数称为  $f$  的  $k+1$  次微变函数，简称  $k+1$  次微变，记作  $\partial^{k+1} f$ 。

对于  $I$  中一点  $a$  来说，如果  $f$  是  $k$  次可微的，且  $\partial^k f$  在  $a$  处可微，就说  $f$  在  $a$  处  $k+1$  次可微，微变率称为  $f$  在  $a$  处的  $k+1$  次微变率。

如果对任意正整数  $k$ ， $f$  都是  $k$  次可微的，就说  $k$  无穷可微。

可以验证，前几节提及的显式函数，都是无穷可微的。因为它们的微变总是由常见显式函数通过四则运算、复合、取反函数的方式组成。

比如，单项式函数  $x \mapsto x^n$  的  $k$  次微变为：

$$\partial^k(x^n) = \begin{cases} x \mapsto n(n-1) \cdots (n-k+1)x^{n-k} & \text{如果 } k \leq n \\ x \mapsto 0 & \text{如果 } k > n \end{cases}$$

指数函数  $x \mapsto a^x$  的  $k$  次微变为  $x \mapsto (\ln a)^k a^x$ 。特别地， $x \mapsto e^x$  的任意次微变都是自己。

正弦、余弦函数的多次微变和次数的奇偶性相关。

$$\begin{aligned} \forall n \in \mathbb{Z}^+, \\ \partial^{2n-1} \sin(x) &= (-1)^{n-1} \cos x, & \partial^{2n} \sin(x) &= (-1)^n \sin x, \\ \partial^{2n-1} \cos(x) &= (-1)^n \sin x, & \partial^{2n} \cos(x) &= (-1)^n \cos x, \end{aligned}$$

两个  $k$  次可微函数  $f$ 、 $g$  的和与差，仍然是  $k$  次可微函数，其  $k$  次微变函数为：

$$\partial^k(f \pm g) = \partial^k f \pm \partial^k g$$

而它们的乘积也依然是  $k$  次可微函数，其  $k$  次微变函数为：

$$\begin{aligned}\partial^k(f \cdot g) &= \partial^k f \cdot g + C_k^1 \partial^{k-1} f \cdot \partial g + \cdots + C_k^i \cdot \partial^{k-i} f \cdot \partial^i g + \cdots + \partial^k g \\ &= \sum_{i=0}^k C_k^i \partial^{k-i} f \cdot \partial^i g\end{aligned}$$

这个表达式和二项式的展开一样，因为函数乘积的微变就是（一次的）二项式。

对于多次可微的函数，使用多次微变，可以给出类似微变中值定理的结论。回顾微变中值定理，我们可以把它用来理解函数在一点附近的行为。对于符合条件的函数  $f$ ，我们可以写出：

$$f(a+h) = f(a) + \partial f(a+t)h.$$

其中  $0 < t < h$ 。我们把  $f$  在  $a$  附近的函数值  $f(a+h)$  表达为关于  $h$  的一次式，相比直接使用  $a$  处的微直观：

$$f(a+h) \approx f(a) + \partial f(a)h,$$

使用  $\partial f(a+t)$  可以得到精确的关系，但这时  $h$  的系数是随着  $h$  变化的。这与我们用平均速度来近似一样。我们希望  $h$  的系数只与  $a$  有关。比如，如果可以把  $f(a+h)$  表示成：

$$f(a+h) = f(a) + \partial f(a)h + w_2 \partial^2 f(a+t)h^2,$$

其中  $0 < t < h$ ， $w_2$  为常数系数。这样，我们就能具体地估计函数的微直观与精确值的误差。当  $h$  很接近 0 的时候， $h^2$  比  $h$  小得多，我们就可以忽略这一项。

一般来说，如果  $f$  多次可微，我们希望能有这样的近似表达式：

$$f(a+h) = f(a) + \partial f(a)h + c_2 \partial^2 f(a)h^2 + \cdots + c_n \partial^n f(a)h^n + w_{n+1} \partial^{n+1} f(a+t)h^{n+1}.$$

即用关于  $h$  的多项式来近似表示  $f(a+h)$  的值。其中  $c_2, \cdots, c_n$  是常数系数。而  $w_{n+1} \partial^{n+1} f(a+t)h^{n+1}$  和  $h$  的  $n+1$  次方成正比，当  $h$  接近 0 时，就变

成一个很小的误差项，比前面次数更低的项小得多。这样，我们就可以通过关于  $h$  的多项式来估计  $f(a+h)$  的值。

**定理 2.5.1. 微变展开定理** 设  $n$  为正整数。已知函数  $f$  在包含实数  $a$  的区间  $I$  中  $n+1$  次可微，那么，对于  $I$  中另一个数  $x$ ，存在介于  $a$  与  $x$  之间的数  $c$ ，使得下式成立：

$$\begin{aligned} f(x) &= f(a) + \sum_{k=1}^n \frac{\partial^k f(a)}{k!} (x-a)^k + \frac{\partial^{n+1} f(c)}{(n+1)!} (x-a)^{n+1}. \\ &= P_n(x-a) + Y_n(x-a). \end{aligned}$$

上式称为  $f$  在  $a$  处的  **$n$  次微变展开**。其中  $P_n$  称为  $f$  在  $a$  处的  **$n$  次微变展开多项式**， $Y_n$  称为  $f$  在  $a$  处展开的  **$n$  次余项**。

这个定理可以通过多次使用微变零值定理和中值定理得到（见附录 B），可以看作微分中值定理的增强版本。下面通过实际例子来看它的应用。

**例题 2.5.1.** 计算  $\sin \frac{\pi}{36}$  的值，精确到小数点后 5 位。

**解答.** 观察正弦函数的图像，它在 0 到  $\frac{\pi}{2}$  上单调递增。 $\frac{\pi}{36}$  相对于  $\frac{\pi}{2}$  很小。我们可以在 0 作多次微变展开，求  $\sin \frac{\pi}{36}$  的近似值。

记  $h = \frac{\pi}{36}$ 。由于正弦函数无穷可微，根据微变展开定理，把它在 0 处作  $n$  次展开，可得：

$$\sin h = \sin 0 + \sum_{k=1}^n \frac{\partial^k \sin(0)}{k!} h^k + Y_n.$$

其中余项  $Y_n$  为：

$$Y_n = \frac{\partial^{n+1} \sin(c)}{(n+1)!} h^{n+1}.$$

按题目要求，计算结果精确到小数点后 5 位。对此，我们希望使用前面的多项式来计算近似值，而余项  $Y_n$  就是近似误差。因此，我们要估计余项的大小。

正弦函数的多次微变要么是余弦函数, 要么是正弦函数, 至多相差一个正负号, 因此绝对值小于等于 1。\$h\$ 的值大约在 0.08 到 0.1 之间, 因此, \$h^5\$ 小于 \$10^{-5}\$。考虑到余项中还有作为分母的 \$(n+1)!\$, 我们可以尝试令 \$n=3\$。这时:

$$\begin{aligned} |Y_n| &= \left| \frac{\partial^{n+1} \sin(c)}{(n+1)!} h^{n+1} \right| \\ &= \frac{|\partial^4 \sin(c)|}{4!} |h|^4 \\ &\leq \frac{|h|^4}{4!} \\ &< \frac{10^{-4}}{24} < 5 \cdot 10^{-6}. \end{aligned}$$

具体计算可知, \$\frac{|h|^4}{4!} \approx 0.0000024\$。因此误差足够小。

接下来计算展开多项式 \$P\_3(h)\$ 的值。

$$\begin{aligned} P_3(h) &= \sin 0 + \frac{\cos(0)}{1!} h^1 - \frac{\sin(0)}{2!} h^2 - \frac{\cos(0)}{3!} h^3 \\ &= h - \frac{h^3}{6} \\ &\approx 0.08716. \end{aligned}$$

实际上, \$\sin \frac{\pi}{36}\$ 的值约为 0.08715574, 精确到小数点后 5 位为 0.08716。因此近似计算结果是正确的。

一般来说, 实践中, 我们使用显式函数作为经验公式来描述、总结某些实验现象。这时, 如果显式函数多次可微, 并且多次的微变大小可控 (比如上面的正弦函数), 那么, 通过合理估计余项的大小, 我们可以把显式函数的值用整式近似, 从而简化计算。

### 思考 2.5.1.

1. 两函数的商的多次微变是否有规律? 能否总结成统一的表达式?
2. 两函数的复合函数的多次微变是否有规律?

3. 函数可微次数是否越多越“好”？可微次数多代表什么？无穷可微的函数有什么“优点”？

**习题 2.5.1.**

1. 证明两函数乘积的多次微变公式。(提示：使用归纳法。)

2. 求以下函数的多次微变函数。

2.1. 求正切函数的 5 次微变函数。

2.2. 求幂函数  $x \mapsto x^r$  的  $k$  次微变函数。

2.3. 给定实数  $a$ ，求函数  $x \mapsto e^{ax}$  的  $k$  次微变函数。

2.4. 求函数  $x \mapsto (ax^2 + bx + c)e^x$  的  $k$  次微变函数。

2.5. 求函数  $x \mapsto x^2 \ln x$  的  $k$  次微变函数。

3. 证明：对任意正整数  $n$ ，函数  $x \mapsto x^n e^{\frac{1}{x}}$  的  $n+1$  次微变函数是

$$x \mapsto \frac{(-1)^{n+1}}{x^{n+2}} e^{\frac{1}{x}}.$$

4.  $n$  为正整数。 $f$  在区间  $[a, b]$  上  $n$  次可微。

4.1. 证明：如果  $f$  在  $[a, b]$  中有  $n+1$  个不同的零点，那么有  $c \in (a, b)$  使得  $\partial^n f(c) = 0$ 。

4.2. 函数  $f$  在  $[a, b]$  上  $n$  次可微。证明：如果

$$f(a) = \partial f(a) = \partial^2 f(a) = \cdots = \partial^{n-1} f(a) = f(b) = 0,$$

那么有  $c \in (a, b)$  使得  $\partial^n f(c) = 0$ 。





## 第三章 研究可微函数

我们很早就学习了用平面上的曲线表示函数的方法来直观地理解函数。函数的图像就是平面上的曲线。

不过，由于缺乏严谨描述实数和曲直质的语言，我们只用描点法来描述函数的图像。画出函数图像上足够多的点，然后用线段将相邻的点连起来，并假设这样能近似表示函数的图像。

通过这样的近似图像，我们可以大致了解函数的性质，但无法严谨地讨论。

对于简单的函数，我们可以直接利用函数的表达式分析函数的性质。对于比较复杂的函数，我们就需要更有力的工具。学习了函数的微变之后，使用微变的知识，我们可以用严谨的方式，讨论曲线的直观性质。

### 3.1 增减与极值

我们学习过单调递增和单调递减函数的概念。函数的增减是我们希望了解的基本性质。

我们已经了解过一些基本的常见函数的增减性，比如正比例函数、反比例函数、幂函数、三角函数等。对于较为复杂的函数，我们通过“增函数的复合仍然是增函数”、“增函数的反函数是增函数”等方法，可以推断一

些函数的单调性。

微变是讨论函数局部变化的有力工具。使用微变的概念，我们可以讨论更一般的情况下，函数的增减变化。

对于可微的函数，我们有以下结论：

**定理 3.1.1. 微变率与函数增减的关系** 函数  $f$  在区间  $I$  上可微。

- 函数  $f$  在  $I$  上为常数函数，当且仅当  $\partial f$  在  $I$  上恒等于 0。
- 函数  $f$  在  $I$  上单调递增（单调递减），当且仅当  $\partial f$  在  $I$  上总大于等于 0（小于等于 0）。
- 如果  $\partial f$  在  $I$  上总大于 0（小于 0），那么函数  $f$  在  $I$  上严格单调递增（严格单调递减）。

以上结论可以使用微变率的定义和微变中值定理得到。可以看到，函数在一段区间内的增减变化，与它在这段区间上的微变的正负性质是紧密相关的。这是判定函数增减的有力工具。

**例题 3.1.1.** 研究函数  $f: x \mapsto x + \frac{1}{x}$  的增减性。

**解答.**  $f$  在  $\mathbb{R}^*$  上有定义。易知  $f$  在定义域上可微。

函数的增减性质与微变函数的正负性质相关。因此，求  $f$  的微变函数，得到：

$$\partial f(x) = 1 - \frac{1}{x^2} = \frac{x^2 - 1}{x^2}.$$

分母  $x^2$  总大于 0，因此  $\partial f$  的正负性取决于二次函数  $x^2 - 1$  的正负性。

根据二次函数的性质， $x^2 - 1$  的正负性质为：

$$x^2 - 1 \begin{cases} < 0 & \text{如果 } x \in (-1, 1) \\ = 0 & \text{如果 } x = -1 \text{ 或 } x = 1 \\ > 0 & \text{如果 } x \in (-\infty, -1) \cup (1, +\infty) \end{cases}$$

因此,  $f$  在区间  $(-\infty, -1]$  上单调递增, 在  $x = -1$  处达到极大值; 然后在  $[-1, 0)$  上单调递减;  $[-1, 0)$  上单调递减; 在  $x = 1$  处达到极小值; 之后在  $[1, +\infty)$  上单调递增。

可以看出,  $f$  在  $-\infty$ 、 $-1$ 、 $0$ 、 $1$ 、 $+\infty$  的情况, 是我们了解  $f$  的几个“关键点”。容易算出在  $-1$ 、 $1$  处, 函数值分别为  $f(-1) = -2$ 、 $f(1) = 2$ 。

$x$  趋于负无穷时,  $x$  趋于负无穷,  $\frac{1}{x}$  趋于  $0$ , 因此  $f$  趋于负无穷。

$x$  趋于正无穷时,  $x$  趋于正无穷,  $\frac{1}{x}$  趋于  $0$ , 因此  $f$  趋于正无穷。

$x$  从左侧趋于  $0$  时,  $\frac{1}{x}$  趋于负无穷, 因此  $f$  趋于负无穷。

$x$  从右侧趋于  $0$  时,  $\frac{1}{x}$  趋于正无穷, 因此  $f$  趋于正无穷。

综上, 我们可以这样描述  $f$  的增减: 随着  $x$  增大,  $f$  在区间  $(-\infty, -1]$  上从负无穷单调递增, 在  $x = -1$  处达到极大值  $f(-1) = -2$ ; 然后在  $[-1, 0)$  上单调递减, 在  $x$  趋于  $0$  时趋于负无穷。 $f$  在  $(0, 1]$  上从正无穷单调递减, 在  $x = 1$  处达到极小值  $f(1) = 2$ ; 之后在  $[1, +\infty)$  上单调递增, 趋于正无穷。

**例题 3.1.2.** 研究函数  $f: x \mapsto x^2 e^{-2x}$  的增减性。

**解答.**  $f$  在  $\mathbb{R}$  上有定义。作为基本显式函数的复合函数, 易知  $f$  在  $\mathbb{R}$  上可微。

函数的增减性质与微变函数的正负性质相关。因此, 求  $f$  的微变函数, 得到:

$$\partial f(x) = (-2x^2 + 2x)e^{-2x}.$$

指数函数  $e^{-2x}$  总大于  $0$ , 因此  $\partial f$  的正负性取决于二次函数  $-2x^2 + 2x$  的正负性。

根据二次函数的性质,  $-2x^2 + 2x$  的正负性质为:

$$-2x^2 + 2x \begin{cases} > 0 & \text{如果 } x \in (0, 1) \\ = 0 & \text{如果 } x = 0 \text{ 或 } x = 1 \\ < 0 & \text{如果 } x \in (-\infty, 0) \cup (1, +\infty) \end{cases}$$

因此,  $f$  在区间  $(-\infty, 0]$  上单调递减, 在  $x = 0$  处达到极小值; 然后在  $(0, 1)$  上单调递增, 在  $x = 1$  处达到极大值; 之后在  $(1, +\infty)$  上单调递减。

可以看出,  $f$  在  $-\infty$ 、0、1、 $+\infty$  的情况, 是我们了解  $f$  的几个“关键点”。容易算出在 0、1 处, 函数值分别为  $f(0) = 0$ 、 $f(1) = e^{-2} \approx 0.135$ 。

$x$  趋于负无穷时,  $x^2$  趋于正无穷,  $e^{-2x}$  趋于正无穷, 因此  $f$  趋于正无穷。

$x$  趋于正无穷时,  $x^2$  趋于正无穷,  $e^{2x}$  趋于正无穷。需要研究两者的比值。 $x > 1$  时,  $e^{2x} > 2^{2x}$ , 于是

$$f(x) = \frac{x^2}{e^{2x}} < \frac{x^2}{2^x} \cdot \frac{1}{2^x},$$

而我们知道, 对于正整数  $n$ , 总有  $2^n \geq n^2$ 。因此,  $f(n) < 2^{-n}$ , 这说明  $n$  趋于无穷时,  $f(n)$  趋于 0。而  $f$  在  $(1, +\infty)$  上单调递减且大于 0, 因此有极限。这个极限就是  $f(n)$  的极限, 也就是 0。

综上, 我们可以这样描述  $f$  的增减: 随着  $x$  增大,  $f$  在区间  $(-\infty, 0]$  上从正无穷单调递减, 在  $x = 0$  处达到极小值  $f(0) = 0$ ; 然后在  $(0, 1)$  上单调递增, 在  $x = 1$  处达到极大值  $f(1) = e^{-2} \approx 0.135$ ; 之后在  $(1, +\infty)$  上单调递减, 趋于 0。

从例题可以知道, 研究可微函数的增减可以分为几个固定步骤:

1. 确定函数的定义域。
2. 确定函数可微, 求出函数的微变函数。
3. 求处微变函数的零点, 确定微变函数的正负性质。微变函数零点之中, 使得正负号改变的点, 就是函数的极值点。

4. 根据微变函数的正负性质, 确定函数的增减。
5. 研究函数在正负无穷的极限、极值点处的值。
6. 根据以上信息, 描述函数的增减。

**例题 3.1.3.** 研究函数  $f: x \mapsto \ln(1 + \sin^2 x)$  的增减性。

**解答.** 对任意实数  $x$ ,  $1 + \sin^2 x > 0$ , 所以  $f$  在  $\mathbb{R}$  上有定义。

作为基本显式函数的复合函数, 易知  $f$  在  $\mathbb{R}$  上可微。

函数的增减性质与微变函数的正负性质相关。因此, 求  $f$  的微变函数, 得到:

$$\partial f(x) = \frac{2 \sin x \cos x}{1 + \sin^2 x} = \frac{\sin 2x}{1 + \sin^2 x}.$$

从  $\partial f$  的表达式可知, 分母总大于 0。因此  $\partial f$  的正负性就是分子, 也就是  $\sin 2x$  的正负性。

根据正弦函数的性质,  $\sin 2x$  是以  $\pi$  为周期的周期函数, 它的正负性如下:

$$\sin 2x \begin{cases} > 0 & \text{如果 } x \in (k\pi, k\pi + \frac{\pi}{2}), k \in \mathbb{Z} \\ = 0 & \text{如果 } x = \frac{k\pi}{2}, k \in \mathbb{Z} \\ < 0 & \text{如果 } x \in (k\pi - \frac{\pi}{2}, k\pi), k \in \mathbb{Z} \end{cases}$$

因此,  $f$  的增减性质也是周期性的, 以  $\pi$  为周期。以一个周期  $[0, \pi]$  为例。随着  $x$  增大,  $f$  在  $[0, \frac{\pi}{2}]$  上单调递增, 在  $\frac{\pi}{2}$  处达到极大值, 然后在  $[\frac{\pi}{2}, \pi]$  上单调递减。在  $\pi$  处达到极小值。

**思考 3.1.1.**

1. 知道了函数的增减性质, 能否画出函数图像? 你认为还有哪些性质, 能帮我们把函数的图像画得更好一点?
2. 微变函数等于 0 的点, 是否一定是极值点? 如果不是的话, 它有什么性质?

**习题 3.1.1.**

1. 证明定理 3.1.1。

2. 研究以下函数的增减性质。

2.1.  $x \mapsto x^3 - x^5$ 。

2.2.  $x \mapsto \tan\left(x - \frac{\pi}{6}\right)$ 。

2.2.  $x \mapsto (1 - x^2)e^x$ 。

2.3.  $x \mapsto \frac{x^2-2}{(x+1)(x-3)}$ 。

2.4.  $x \mapsto e^{-\frac{1}{x}}$ 。

2.5.  $x \mapsto e^{-3x} \sin(2x)$ 。

3. 设  $I$  为实数区间,  $X$  是  $I$  的子集。如果对  $I$  中任意  $a < b$ , 都有  $x \in I$  使得  $a < x < b$ , 就说  $X$  在  $I$  中稠密。设  $f$  在  $I$  上可微, 微变函数  $\partial f$  在  $I$  上总大于等于 0。记  $A^+$  为  $I$  中微变函数  $\partial f$  大于 0 的点的集合。

3.1. 证明: 如果  $A^+$  在  $I$  中稠密, 那么  $f$  在  $I$  上严格单调递增。

3.2. 证明: 如果  $f$  在  $I$  上严格单调递增, 那么  $A^+$  在  $I$  中稠密。

## 3.2 局部性质

研究函数的增减性质时, 我们发现, 需要关注函数在一些“关键点”的行为。如果函数在“关键点”有定义, 那么我们只需要取函数值即可。如果“关键点”是正负无穷, 或者函数在“关键点”没有定义, 那么我们需要特别研究函数在这些地方附近的行为。我们把这样的性质称为函数的局部性质。

为此, 我们再次说明“附近”的严格定义: 一点“附近”指它的某个邻域。而邻域的概念我们已经学习过了, 这里再严谨地定义一次:

**定义 3.2.1. 邻域** 对于实数  $a$ , 我们把包含  $a$  的开区间  $I$  称为  $a$  的邻域。为了方便, 我们一般取对称的开区间  $(a - h, a + h)$ , 其中  $h > 0$  是足够小的实数, 以免涉及到别的区域。这样邻域记为  $I_h(a)$ 。

如果涉及函数  $f$ , 那么在提及  $a$  的邻域时, 一般考虑以上区间和  $f$  定义域的交集。但记法不变, 仍然是  $I_h(a)$ 。读者需要理解, 这里指的是关于  $f$  有定义的部分。

比如, 如果  $f$  在  $a$  附近有定义, 但在  $a$  处无定义, 那么交集就是在  $a$  的邻域中去掉  $a$  得到的子集, 称为  $a$  的**去心邻域**。

显然, 去心邻域就是  $a$  左右的两个开区间。比如,  $(a - h, a + h)$  对应的去心邻域就是  $(a - h, a) \cup (a, a + h)$ 。这种情况可以记为  $\overset{\circ}{I}_h(a)$ 。

如果我们只希望讨论  $a$  的一侧, 那么可以只取  $a$  的邻域小于等于  $a$  或大于等于  $a$  的部分, 称为  $a$  的**左邻域**、**右邻域**。

对于正负无穷, 我们把形如  $[H, +\infty)$ 、 $(H, +\infty)$  或  $(-\infty, H]$ 、 $(-\infty, H)$  的区间称为它们的邻域, 记为  $I_H(+\infty)$ 、 $I_H(-\infty)$ 。

研究函数的局部行为, 也就是研究函数在特定点或者无穷远点附近的行为。

**例子 3.2.1.** 1. 比较函数  $f_1: x \mapsto x$  和  $f_2: x \mapsto x^2$  在 0 附近的行为。

2. 比较函数  $f_1: x \mapsto \frac{1}{x}$  和  $f_2: x \mapsto \frac{1}{x^2}$  在 0 附近的行为。

对于第一个例子中的两个函数, 我们知道, 两者在  $x$  沿右侧趋于 0 时都趋于 0。但  $f_2$  比  $f_1$  要快得多。比如, 令  $x = 0.01$ , 则  $f_1(x) = 0.01$ , 而  $f_2(x) = 0.0001$ ; 令  $x = 0.001$ , 则  $f_1(x) = 0.001$ , 而  $f_2(x) = 0.000001$ 。

对于第二个例子中的两个函数, 我们知道, 两者在  $x$  沿右侧趋于 0 时都趋于正无穷。但  $f_2$  比  $f_1$  要快得多。比如, 令  $x = 0.01$ , 则  $f_1(x) = 100$ , 而  $f_2(x) = 10000$ ; 令  $x = 0.001$ , 则  $f_1(x) = 1000$ , 而  $f_2(x) = 1000000$ 。

为了比较不同函数的局部行为, 我们需要把“趋于无穷”、“趋于零”更细分。

**定义 3.2.2. 无穷的阶** 设  $f, g$  是从区间  $I$  到  $\mathbb{R}$  的函数,  $a \in I$ 。<sup>1</sup>

1. 如果对任意  $r > 0$ , 都有  $a$  的邻域  $I_a(h)$ , 使得

$$\forall x \in I_a(h), |f(x)| \leq r|g(x)|,$$

---

<sup>1</sup>注意: 这里的  $a$  也包括正负无穷。

就说相比于  $g$ ,  $f$  在  $a$  附近**可以忽略**。记为:

$$f \stackrel{a}{=} o(g)$$

如果  $x$  趋于  $a$  时,  $f$ 、 $g$  都趋于无穷, 就说  $g$  是  $f$  的**高阶无穷大**。

如果  $x$  趋于  $a$  时,  $f$ 、 $g$  都趋于零, 就说  $f$  是  $g$  的**高阶无穷小**。

2. 如果有  $a$  的邻域  $I_a(h)$  和  $M > 0$ , 使得

$$\forall x \in I_a(h), |f(x)| \leq M|g(x)|,$$

就说  $f$  在  $a$  附近**受制于** $g$ 。记为:

$$f \stackrel{a}{=} \mathcal{O}(g)$$

3. 如果  $f$  在  $a$  附近受制于  $g$ , 而  $g$  也受制于  $f$ , 就说两者**等阶**。

$f$  是  $g$  的**同阶无穷大**或**同阶无穷小**。记为:

$$f \stackrel{a}{=} (g)$$

4. 如果  $f - g$  相比  $g$  可以忽略, 就说  $f$  和  $g$  在  $a$  附近**等价**,

$f$  是  $g$  的**等价无穷大**或**等价无穷小**。

$$f \stackrel{a}{\sim} g$$

简单来说, 越高阶的无穷大(无穷小), 表示趋于无穷(零)的速度不是一个级别的。同阶的无穷, 趋于无穷(零)的速度虽然有快慢, 但仍然在同一个级别。等价的无穷, 趋于无穷(零)的速度相同。

在前面第一个例子中,  $f_1: x \mapsto x$  和  $f_2: x \mapsto x^2$  在 0 附近都是无穷小, 但对任意  $r > 0$ , 我们都可以设  $h = r$ , 这样, 只要  $|x| < r$ , 就有  $|x^2| < r|x|$ 。这说明  $f_2$  是  $f_1$  的高阶无穷小。

同理, 第二个例子中,  $f_1: x \mapsto \frac{1}{x}$  和  $f_2: x \mapsto \frac{1}{x^2}$  在 0 附近都是无穷大。但对任意  $r > 0$  我们都可以设  $h = r$ , 这样, 只要  $|x| < r$ , 就有  $|\frac{1}{x}| < r|\frac{1}{x^2}|$ 。这说明  $f_2$  是  $f_1$  的高阶无穷大。



从定义来看, 无穷的阶和极限很相似。两个函数在局部的无穷关系, 也可以用两者比值的极限来表示:

- $f$  在  $a$  附近相比  $g$  可以忽略, 当且仅当  $\lim_{x \rightarrow a} \frac{f}{g} = 0$ .
- $f$  在  $a$  附近与  $g$  等价, 当且仅当  $\lim_{x \rightarrow a} \frac{f}{g} = 1$ .
- 如果  $\lim_{x \rightarrow a} \frac{f}{g} > 0$ , 那么  $f$  在  $a$  附近是  $g$  的同阶无穷大。

我们可以把已知的说法转换成无穷的阶的“语言”。比如, 函数  $f \stackrel{a}{=} o(1)$ , 表示  $f$  在  $a$  处趋于 0。又如, 函数  $f \stackrel{a}{=} \mathcal{O}(1)$ , 表示  $f$  在  $a$  附近有界。

根据无穷的阶的定义, 我们还可以“开发”出一些“运算法则”。当然, 它们只是用无穷的阶的“语言”, 把按照极限定义得到的直接推论复述一遍。

1. 如果  $f_1 \stackrel{a}{=} o(g_1)$ ,  $f_2 \stackrel{a}{=} o(g_2)$ , 那么  $f_1 f_2 \stackrel{a}{=} o(g_1 g_2)$ 。
2. 如果  $f_1 \stackrel{a}{=} \mathcal{O}(g_1)$ ,  $f_2 \stackrel{a}{=} \mathcal{O}(g_2)$ , 那么  $f_1 f_2 \stackrel{a}{=} \mathcal{O}(g_1 g_2)$ 。
3. 如果  $f \stackrel{a}{=} \mathcal{O}(g)$ , 那么对任意正整数  $n$ ,  $f^n \stackrel{a}{=} \mathcal{O}(g^n)$ 。

我们可以用无穷的阶来更细致地刻画函数的局部性质。比如, 在上一章中, 我们研究了正弦函数在 0 附近的近似表示。在 0 附近使用微变展开定理, 正弦函数可以写成:

$$\sin h = \sin 0 + \sum_{k=1}^n \frac{\partial^k \sin(0)}{k!} h^k + Y_n.$$

其中余项  $Y_n$  为:

$$Y_n = \frac{\partial^{n+1} \sin(c)}{(n+1)!} h^{n+1}.$$

这说明正弦函数可以表示成多项式函数  $P_n : h \mapsto \sin 0 + \sum_{k=1}^n \frac{\partial^k \sin(0)}{k!} h^k$  与余项函数  $h \mapsto Y_n(h)$  的和。而我们已经证明了这个余项函数受制于  $h \mapsto h^{n+1}$ , 因此是  $P_n$  任一项的高阶无穷小。因此, 正弦函数在 0 附近可以用  $P_n$  来近似表示。

把函数用多项式近似表示, 有什么好处呢? 把不同的函数统一用多项式近似表示, 我们就能够(在局部)比较各种不同的函数了。

比如, 我们知道  $\sin x$  和  $\ln(1+x)$  在 0 附近都约等于 0。那么比值  $\frac{\sin x}{\ln(1+x)}$  在 0 附近是什么样子的呢?

首先利用微变展开定理把两者用多项式和余项表示。

$$\begin{aligned}\sin x &= \sin 0 + \sum_{k=1}^n \frac{\partial^k \sin(0)}{k!} x^k + o(x^n). \\ \ln(1+x) &= \ln 1 + \sum_{k=1}^n \frac{\partial^k \ln(1)}{k!} x^k + o(x^n).\end{aligned}$$

然后从最“低阶”的无穷小开始对比。

$$\begin{aligned}\sin x &= x - \frac{x^3}{6} + \frac{x^5}{120} + \cdots + o(x^n). \\ \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + o(x^n).\end{aligned}$$

$\sin x$  和  $\ln(1+x)$  的“最大项”都是  $x$ , 之后  $\ln(1+x)$  是  $-\frac{x^2}{2}$ , 而  $\sin x$  是  $-\frac{x^3}{6}$ 。因此, 我们可以写出它们的比值:

$$\begin{aligned}\frac{\sin x}{\ln(1+x)} &= \frac{x - \frac{x^3}{6} + o(x^3)}{x - \frac{x^2}{2} + o(x^2)} \\ &= \frac{1 - o(x)}{1 - \frac{x}{2} + o(x)} \\ &= \frac{(1 - o(x))(1 + \frac{x}{2} + o(x))}{(1 - \frac{x}{2} + o(x))(1 + \frac{x}{2} + o(x))} \\ &= \frac{1 + \frac{x}{2} + o(x)}{1 + o(x)} \\ &= \left(1 + \frac{x}{2} + o(x)\right)(1 + o(x)) \\ &= 1 + \frac{x}{2} + o(x)\end{aligned}$$

也就是说, 在 0 附近,  $\sin x$  和  $\ln(1+x)$  的比值大约是  $1 + \frac{x}{2}$ , 余项是关于  $x$  的高阶无穷小。

把函数在 0 附近用一次整式和高阶无穷小表示, 我们把这样的结果称为函数在 0 附近的一次展开。

我们把上面的方法称为**局部展开法**。简单来说，如果能够把函数在局部表示成多项式与高阶无穷小的和，那么，通过无穷的阶的运算，我们可以细致地分析函数的局部性质。从一些常见函数的局部展开（见附录 B）出发，我们可以比较准确地刻画较为复杂的函数的局部性质。

**例题 3.2.1.** 求极限：

$$\lim_{x \rightarrow +\infty} \frac{\left(1 + \frac{1}{x}\right)^{x^2}}{e^x}.$$

**解答.** 要求极限的函数由分子分母两部分组成。分子看起来很像  $\left(1 + \frac{1}{x}\right)^x$ ，也就是  $x$  趋于正无穷时极限为  $e$ 。不过指数从  $x$  换成了  $x^2$ 。

既然分子分母都和指数函数有关，我们可以把整个函数取对数，得到：

$$\ln \frac{\left(1 + \frac{1}{x}\right)^{x^2}}{e^x} = x^2 \ln \left(1 + \frac{1}{x}\right) - x.$$

显然，等式右边比较难估计的是对数函数的部分。 $x$  趋于正无穷时， $\frac{1}{x}$  趋于 0。我们把  $\ln \left(1 + \frac{1}{x}\right)$  展开，得到：

$$\ln \left(1 + \frac{1}{x}\right) \stackrel{+\infty}{=} \frac{1}{x} - \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right)$$

因此，

$$\begin{aligned} x^2 \ln \left(1 + \frac{1}{x}\right) - x &\stackrel{+\infty}{=} x^2 \left(\frac{1}{x} - \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right)\right) - x \\ &\stackrel{+\infty}{=} x - \frac{1}{2} + o(1) - x \\ &\stackrel{+\infty}{=} -\frac{1}{2} + o(1) \end{aligned}$$

也就是说，

$$\lim_{x \rightarrow +\infty} x^2 \ln \left(1 + \frac{1}{x}\right) - x = -\frac{1}{2}.$$

于是

$$\lim_{x \rightarrow +\infty} \frac{\left(1 + \frac{1}{x}\right)^{x^2}}{e^x} = e^{\lim_{x \rightarrow +\infty} x^2 \ln \left(1 + \frac{1}{x}\right) - x} = \frac{1}{\sqrt{e}}.$$

最后的问题：具有怎样性质的函数，才能在局部表示成多项式与高阶无穷小的和呢？我们有以下定理（证明见附录 B）：

**定理 3.2.1. 局部展开定理** 设  $n$  为正整数。函数  $f$  在区间  $I$  上  $n$  次可微，且  $n$  次微变函数在  $I$  上连续。那么，对于  $a \in I$ ， $f$  可以写成以下形式：

$$\forall x \in I, \quad f(x) = f(a) + \sum_{k=1}^n \frac{\partial^k f(a)}{k!} (x-a)^k + (x-a)^n h(x).$$

其中  $h$  是一个在  $I$  上连续，在  $a$  处趋于 0 的函数。

### 思考 3.2.1.

1. 说明：等价无穷大（无穷小）是一种等价关系。
2. 你认为，我们已知的关系中，什么关系和函数“可忽略”、“受制于”的关系最像？给出你的理由。
3. 函数在局部是否一定可以用多项式近似表示？如果是，说明理由。如果不是，给出反例。
4. 如果  $f_1 \stackrel{a}{\sim} g_1$ ,  $f_2 \stackrel{a}{\sim} g_2$ ，是否有  $f_1 + f_2 \stackrel{a}{\sim} g_1 + g_2$ ？

### 习题 3.2.1.

1. 证明：
  - 1.1. 如果  $f \stackrel{a}{\sim} g$ ,  $g \stackrel{a}{=} o(h)$ ，那么  $f \stackrel{a}{=} o(h)$ 。
  - 1.2. 如果  $f \stackrel{a}{\sim} g$ ,  $g \stackrel{a}{=} \mathcal{O}(h)$ ，那么  $f \stackrel{a}{=} \mathcal{O}(h)$ 。
  - 1.3. 如果  $f \stackrel{a}{=} o(g)$ ,  $g \stackrel{a}{=} \mathcal{O}(h)$ ，那么  $f \stackrel{a}{=} o(h)$ 。
  - 1.4. 如果  $f \stackrel{a}{\sim} o(g)$ ,  $g \stackrel{a}{=} \mathcal{O}(h)$ ，那么  $f \stackrel{a}{=} o(h)$ 。
2. 设  $a, b$  为正数。证明：
  - 2.1.  $x^a \stackrel{+\infty}{=} o(e^{bx})$ 。
  - 2.2.  $e^{-bx} \stackrel{+\infty}{=} o(|x|^{-a})$ 。
  - 2.3.  $\ln^a x \stackrel{+\infty}{=} o(x^b)$ 。
  - 2.4.  $|\ln x|^a \stackrel{0}{=} o(|x|^{-b})$ 。
3. 证明：
  - 3.1. 在 0 附近， $\frac{1}{1-x^2} = 1 + x^2 + o(x^2)$ 。

3.2. 在 0 附近,  $x^2 \cdot o(x^3) = o(x^5)$ 。

3.3. 在 0 附近,  $e^x - \sqrt{1+x} = \frac{x}{2} + \frac{5x^2}{8} + \frac{7x^3}{48} + o(x^3)$ 。

3.4. 在 0 附近, 偶函数的局部展开只有偶数次项, 奇函数的展开只有奇数次项。

4. 在局部展开以下函数:

4.1. 在 0 附近将  $x \mapsto \frac{\sqrt{1+x}}{1-x^2}$  作三次展开。

4.2. 在 0 附近将  $x \mapsto \frac{\cos x - 1}{\ln^2(1+x)}$  作三次展开。

4.3. 在 0 附近将  $x \mapsto \tan(x) - \frac{x}{1+x}$  作四次展开。

4.4. 在 0 附近将  $x \mapsto (1 - \ln(1+x))e^{-x}$  作四次展开。

4.5. 在 2 附近将  $x \mapsto x^2 \ln(x-1)$  作三次展开。

4.6. 在 1 附近将  $x \mapsto (1-x) \ln(x^2-1)$  作三次展开。

5. 使用局部展开法研究以下的极限:

$$1. \quad \lim_{x \rightarrow 0} \frac{\cos x - 1}{x^2}$$

$$2. \quad \lim_{x \rightarrow 0} \frac{1}{x(e^x - 1)} - \frac{1}{x^2}$$

$$3. \quad \lim_{x \rightarrow +\infty} \frac{1}{x} \ln\left(\frac{e^x - 1}{x}\right)$$

$$4. \quad \lim_{x \rightarrow 0} \frac{2x}{\ln(1+x) - \ln(1-x)}$$

6. 如果  $x \mapsto \frac{e^x + e^{ax} - 2}{x^2}$  在 0 处有极限, 求实数  $a$  可能的取值, 并计算相应的极限。

### 3.3 凹凸性质

要了解函数的性质, 仅靠增减性质是不够的。举例来说,  $x \mapsto \sqrt{x}$  和  $x \mapsto x^2$  在  $[0, 1]$  上都是增函数, 从 0 增长到 1。但两者的图像有很大差异。前者的图像是向上凸出的曲线, 后者则向下凸出。

我们把函数形状分为两类, 形如  $x \mapsto \sqrt{x}$  的称为上凸, 把形如  $x \mapsto x^2$  的称为下凸。更严谨的定义如下:

#### 定义 3.3.1. 函数的凸性

函数  $f$  区间  $I$  上有定义。

- 如果对  $I$  中任意两点  $x, y$ , 以及任意  $0 \leq c \leq 1$ , 都有

$$f(cx + (1 - c)y) \leq cf(x) + (1 - c)f(y),$$

就说  $f$  在区间  $I$  下凸。

- 如果对  $I$  中任意两点  $x, y$ , 以及任意  $0 < c < 1$ , 都有

$$f(cx + (1 - c)y) < cf(x) + (1 - c)f(y),$$

就说  $f$  在区间  $I$  严格下凸。

- 如果对  $I$  中任意两点  $x, y$ , 以及任意  $0 \leq c \leq 1$ , 都有

$$f(cx + (1 - c)y) \geq cf(x) + (1 - c)f(y),$$

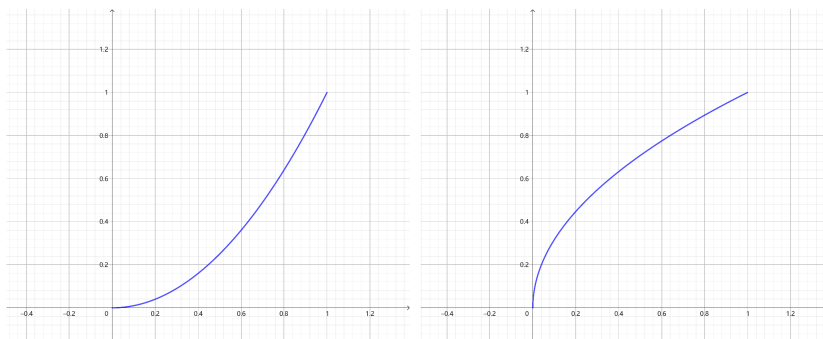
就说  $f$  在区间  $I$  上凸。

- 如果对  $I$  中任意两点  $x, y$ , 以及任意  $0 < c < 1$ , 都有

$$f(cx + (1 - c)y) > cf(x) + (1 - c)f(y),$$

就说  $f$  在区间  $I$  严格上凸。

不难验证, 前面提及的  $x \mapsto x^2$  和  $x \mapsto \sqrt{x}$  和  $x \mapsto x^2$  分别在  $[0, 1]$  严格上凸、严格下凸。

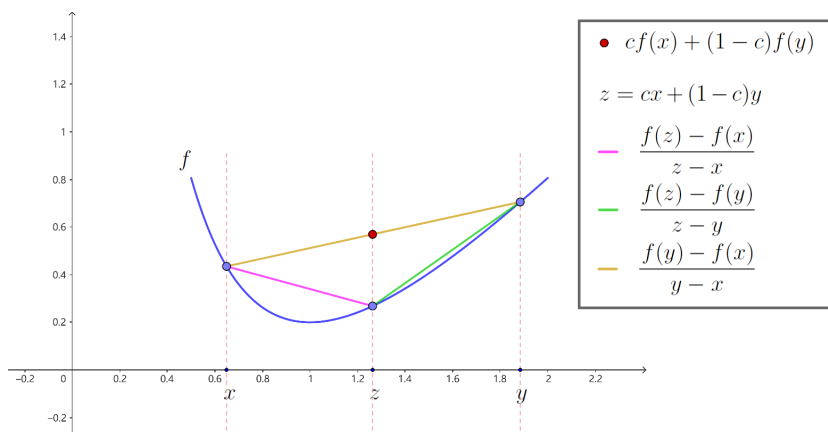


$x \mapsto x^2$  (左) 严格下凸,  $x \mapsto \sqrt{x}$  (右) 严格上凸。

设  $z = cx + (1 - c)y$ , 可得  $c = \frac{y-z}{y-x}$ 。带入凸性的定义, 就可以得到用变率描述的版本。比如, 下凸的定义变为: 对  $I$  中任意三点  $x < z < y$ , 都有

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{y - z}.$$

或者说, 从  $x$  到  $z$  的变率, 小于等于从  $z$  到  $y$  的变率。换句话说, 函数下凸, 就是变率不断变大。反之, 函数上凸, 就是变率不断变小。



直观来看, 设函数在  $[a, b]$  上有定义。考虑以函数在两端的点  $(a, f(a))$ 、 $(b, f(b))$  连成的线段。如果函数下凸, 那么在  $(a, b)$  上, 函数图像总在线段下方; 如果函数上凸, 那么在  $(a, b)$  上, 函数图像总在线段上方。

对于简单的函数, 可以按定义确定函数的凸性。比如,  $x \mapsto x^n$  在  $n > 1$  时是下凸函数,  $x \mapsto x^{\frac{1}{n}}$  在  $n > 1$  时是上凸函数。但对于比较复杂的函数, 直接按定义较难确定函数的凸性。为此, 我们需要更方便的判定方法。如果函数可微, 从凸性和变率的关系, 不难想到以下的判定方法 (证明见附录 B):

**定理 3.3.1. 凸性与微变** 设函数  $f$  在区间  $I$  上可微。

- $f$  在  $I$  下凸, 当且仅当  $\partial f$  在  $I$  上单调递增。
- $f$  在  $I$  上凸, 当且仅当  $\partial f$  在  $I$  上单调递减。
- $f$  在  $I$  严格下凸, 当且仅当  $\partial f$  在  $I$  上严格单调递增。

- $f$  在  $I$  严格下凸, 当且仅当  $\partial f$  在  $I$  上严格单调递减。

如果函数  $f$  二次可微, 那么我们有更简便的判定方法:

**定理 3.3.2. 凸性与二次微变** 设函数  $f$  在区间  $I$  上二次可微。

- $f$  在  $I$  下凸, 当且仅当  $\partial^2 f$  在  $I$  上总大于等于 0。
- $f$  在  $I$  上凸, 当且仅当  $\partial^2 f$  在  $I$  上总小于等于 0。

**证明:** 只需证明  $\partial f$  在  $I$  上单调递增当且仅当  $\partial^2 f$  在  $I$  上总大于等于 0。  
对可微函数  $\partial f$  使用定理 3.1.1 即可。  $\square$

用这个方法, 我们能够轻松判定大多数常见显式函数的凸性, 因为它们都二次可微。

幂函数  $x \mapsto x^a$  当  $a > 1$  时是下凸的, 当  $0 < a < 1$  时是上凸的;  $a < 0$  时, 在  $(0, +\infty)$  上是下凸的。

指数函数总是下凸的, 对数函数总是上凸的。

正弦函数在每个周期  $[2k\pi, 2(k+1)\pi)$  前半是上凸的, 后半是下凸的。正切函数也是如此, 在每个周期  $(k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2})$  前半是上凸的, 后半是下凸的。

如果函数的二次微变在某点  $a$  变号, 就说  $a$  是函数的**拐点**。比如,  $x \mapsto x^3$  在 0 处有拐点。它的二次微变在  $x < 0$  时小于 0, 在  $x > 0$  时大于 0。随着  $x$  增大, 函数在 0 处从上凸转为下凸。

在函数的增减性质分析基础上, 加入凸性的分析, 就大致完成了对函数性质的初步分析。有了这些分析结果, 我们就能大致画出函数的图像。

**例题 3.3.1.** 研究函数  $f: x \mapsto x^2 e^{-2x}$ 。

**解答.** 之前我们已经研究了  $f$  的增减性质。随着  $x$  增大,  $f$  在区间  $(-\infty, 0]$  上从正无穷单调递减, 在  $x = 0$  处达到极小值  $f(0) = 0$ ; 然后在  $(0, 1)$  上



单调递增，在  $x = 1$  处达到极大值  $f(1) = e^{-2} \approx 0.135$ ；之后在  $(1, +\infty)$  上单调递减，趋于 0。

在此基础上，我们加上对函数凸性的研究。

$f$  在  $\mathbb{R}$  上有定义。容易验证  $f$  在  $\mathbb{R}$  上二次可微。二次微变函数为：

$$\partial^2 f(x) = 2(2x^2 - 4x + 1)e^{-2x}.$$

$e^{-2x}$  总大于零，所以二次微变的零点为  $2x^2 - 4x + 1$  的根，即：

$$\begin{cases} x_1 = \frac{2+\sqrt{2}}{2} \\ x_2 = \frac{2-\sqrt{2}}{2} \end{cases}$$

根据二次函数的性质， $f$  的二次微变在  $(-\infty, x_1)$ 、 $(x_2, +\infty)$  上大于零，在  $(x_1, x_2)$  上小于零。因此， $x_1, x_2$  是函数  $f$  的拐点。 $f$  在  $(-\infty, x_1)$ 、 $(x_2, +\infty)$  下凸。在  $(x_1, x_2)$  上凸。

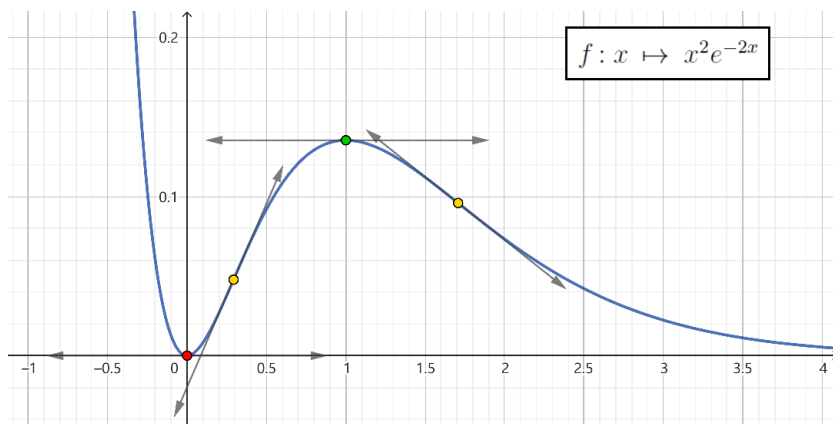
最后， $f$  在  $x_1, x_2$  处的值为：

$$f(x_1) = \frac{3+2\sqrt{2}}{2}e^{-2-\sqrt{2}} \approx 0.1, \quad f(x_2) = \frac{3-2\sqrt{2}}{2}e^{-2+\sqrt{2}} \approx 0.05.$$

微变率分别是：

$$\partial f(x_1) = -(\sqrt{2}+1)e^{-2-\sqrt{2}} \approx -0.08, \quad \partial f(x_2) = (\sqrt{2}-1)e^{-2+\sqrt{2}} \approx 0.23.$$

这样，我们可以大致画出  $f$  的图像。



**思考 3.3.1.**

1. 怎样用函数的二次微变判定它是否严格上凸 (下凸)?
2. 在区间  $I$  上不连续的函数能否在  $I$  上凸或下凸? 说明理由。

**习题 3.3.1.**

1. 判断以下函数的凸性:

$$1. \quad x \mapsto x^5 + 10x^2 - 1 \quad 2. \quad x \mapsto \tan x + 2 \ln \cos x$$

$$3. \quad x \mapsto (x+1) \ln \left(1 - \frac{1}{x}\right) \quad 4. \quad x \mapsto (1-x^2) e^{-\frac{x^2}{2}}$$

2. 证明: 如果函数下凸, 那么对  $I$  中任意三点  $x < z < y$ , 都有

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(z)}{y - z}.$$

3. 函数  $f$  在区间  $I$  上可微且下凸。证明:  $f$  在  $I$  中任一点的切线总在  $f$  图像的下方。

4. 证明**凸心不等式**: 如果函数  $f$  在区间  $I$  下凸, 那么对  $I$  中  $n$  个不同的数  $x_1, x_2, \dots, x_n$ , 以及大于等于零, 和为 1 的  $n$  个系数  $a_1, a_2, \dots, a_n$ , 总有:

$$f \left( \sum_{i=1}^n a_i x_i \right) \geq \sum_{i=1}^n a_i f(x_i).$$

5. 使用凸心不等式证明诸均值不等式。

6. 证明, 对任意  $x \in (0, \frac{\pi}{2})$ , 总有:

$$\frac{2}{\pi} x < \sin x < x.$$

7. 已知  $f, g$  是从  $\mathbb{R}$  到  $\mathbb{R}$  的下凸函数,  $g$  单调递增。证明: 复合函数  $g \circ f$  也是下凸函数。

8.  $f, g$  是区间  $I$  上的下凸函数。定义函数:

$$f \wedge g : x \mapsto \begin{cases} f(x) & \text{如果 } f(x) > g(x) \\ g(x) & \text{如果 } f(x) \leq g(x) \end{cases}$$

函数  $f \wedge g$  的值总是  $f, g$  中较大者。  $f \wedge g$  是否是下凸函数?

如果定义函数  $f \vee g$  为  $f, g$  中较小者,  $f \vee g$  是否是下凸函数?

## 第四章 平直空间

平面和立体空间是我们熟悉的概念。为了讨论其中的点、线等形状，我们引入了向量的概念。我们发现，除了实际中的空间，数学中还有别的概念，遵循和平面向量类似的法则。

**例子 4.0.1.**

1. 设  $\mathbb{K}$  是某个数域（比如有理数集或实数集），考虑所有次数不超过 3 次的整式的集合：

$$\mathbb{K}_3[x] = \{ax^3 + bx^2 + cx + d \mid a, b, c, d \in \mathbb{K}\}$$

2. 考虑  $S_2$  是所有满足以下方程的实变函数的集合。

$$\partial^2 f + f = 0$$

第一个例子里，集合  $\mathbb{K}_3[x]$  的每个元素都可以唯一表示成  $\{1, x, x^2, x^3\}$  的直组合。如果我们把  $\{1, x, x^2, x^3\}$  看作“基底”，那么  $\mathbb{K}_3[x]$  中的元素就可以看成向量。

第二个例子里，考虑  $S_2$  中元素  $u$ 、 $v$  的直组合  $au + bv$ ：

$$\partial^2(au + bv) = a\partial^2 u + b\partial^2 v,$$

所以

$$\begin{aligned}\partial^2(au + bv) + au + bv &= a\partial^2 u + b\partial^2 v + au + bv \\ &= a(\partial^2 u + u) + b(\partial^2 v + v) = 0.\end{aligned}$$

$au + bv$  也是  $S_2$  中元素。于是，我们也可以把  $S_2$  中元素看作向量。比如，可以验证正弦函数  $\sin$  和余弦函数  $\cos$  都是  $S_2$  中元素，因此，其任意直组合  $x \mapsto a \sin x + b \cos x$  也是  $S_2$  中元素。

就像日常计数中抽象出数和代数的概念一样，通过研究这些问题和平面、立体空间的共通之处，我们抽象出一种新概念，叫做平直空间。

**定义 4.0.1.** 给定作为系数的数域  $\mathbb{K}$  (如有理数或实数)， $\mathbb{K}$  上的平直空间  $\mathbb{V}$  是这样的一个集合：

首先， $\mathbb{V}$  上定义了两种二元运算： $\diamond$  和  $\bullet$ 。其中：

- $\diamond$  运算将两个  $\mathbb{V}$  中的元素映射为  $\mathbb{V}$  中的另一个元素。
- $\bullet$  运算将一个  $\mathbb{K}$  中的元素和一个  $\mathbb{V}$  中的元素映射为  $\mathbb{V}$  中另一个元素。

其次， $\mathbb{K}$ 、 $\mathbb{V}$  和以上两种运算具有以下七种性质：

1.  $\mathbb{V}$  中元素对于  $\diamond$  运算满足交换律：

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{V}, \mathbf{u} \diamond \mathbf{v} = \mathbf{v} \diamond \mathbf{u}.$$

2.  $\mathbb{V}$  中元素对于  $\diamond$  运算满足结合律：

$$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}, (\mathbf{u} \diamond \mathbf{v}) \diamond \mathbf{w} = \mathbf{u} \diamond (\mathbf{v} \diamond \mathbf{w}).$$

3.  $\mathbb{V}$  中元素对于  $\diamond$  运算存在零元素：

$$\exists \mathbf{0} \in \mathbb{V}, \text{ 使得 } \forall \mathbf{u} \in \mathbb{V}, \mathbf{u} \diamond \mathbf{0} = \mathbf{u}.$$

4.  $\mathbb{K}$  中有  $\bullet$  运算的单位元  $1_{\mathbb{V}}$ ：

$$\forall \mathbf{u} \in \mathbb{V}, 1_{\mathbb{V}} \bullet \mathbf{u} = \mathbf{u}.$$

5.  $\mathbb{K}$  中乘法  $\cdot$  和  $\bullet$  运算兼容：

$$\forall a, b \in \mathbb{K}, \forall \mathbf{u} \in \mathbb{V}, (a \cdot b) \bullet \mathbf{u} = a \bullet (b \bullet \mathbf{u}).$$

6.  $\bullet$  运算对  $\mathbb{K}$  中加法  $+$  满足分配律:

$$\forall a, b \in \mathbb{K}, \forall \mathbf{u} \in \mathbb{V}, (a + b) \bullet \mathbf{u} = (a \bullet \mathbf{u}) + (b \bullet \mathbf{u}).$$

7.  $\mathbb{K}$  中乘法  $\cdot$  对  $\diamond$  运算满足分配律:

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{V}, \forall a \in \mathbb{K}, a \cdot (\mathbf{u} \diamond \mathbf{v}) = (a \cdot \mathbf{u}) \diamond (a \cdot \mathbf{v}).$$

以上的定义可能比较难理解。我们可以比照平面、立体空间的向量来逐步分析。

以平面向量为例，作为系数的数域  $\mathbb{K}$  就是实数域  $\mathbb{R}$ 。两种二元运算： $\diamond$  和  $\bullet$  分别就是向量的加法和数乘。

这样看来，七种性质的前三种和平面向量是一样的。第五、六种性质就是平面向量定义中的“放缩和四则运算相容”。第七种性质就是平面向量定义中的“放缩和平移相容”。

总体来说，平直空间的定义是从平面、立体空间的向量衍生而来。新增的第四种性质是保证空间的尺度不变：乘以系数 1 不会改变向量的大小。

可以验证，前面的两个例子： $\mathbb{K}_3[x]$  和  $S_2$ ，都是平直空间。

一般来说，对任意正整数  $n$ ，由次数不超过  $n$  的整式构成的集合：

$$\mathbb{K}_n[x] = \{a_0 + a_1x + \cdots + a_nx^n \mid a_0, a_1, \cdots, a_n \in \mathbb{K}\}$$

是平直空间。所有整式的集合  $\mathbb{K}[x]$  也是平直空间。

给定整式  $P(x) = a_0 + a_1x + \cdots + a_nx^n$ ，它可以对应以下的方程：

$$a_0f + a_1\partial f + \cdots + a_n\partial^n f = 0.$$

这个方程所有解的集合  $S_P$  也是平直空间。

方便起见，平直空间的元素，也称为向量。对应的二元运算也称为加法、数乘。其他的概念也沿用平面、立体空间中的称呼。

**习题 4.0.1.**

1. 证明：零向量只有一个，任何向量乘 0 得到零向量。
2. 证明：零向量乘任何数得到零向量。
3. 证明：任何向量  $\mathbf{a}$  都有唯一的反向量  $\mathbf{b}$ ，满足  $\mathbf{a} + \mathbf{b} = \mathbf{0}$ 。
4. 实数集是否是平直空间？有理数集、整数集呢？数轴上的区间是否是平直空间？
5. 给定正整数  $n$ ，由次数为  $n$  的整式构成的集合是否是平直空间？
6. 所有有理数列的集合  $\mathbb{Q}^{\mathbb{N}}$  是否是平直空间？
7. 所有奇函数的集合是否是平直空间？所有偶函数的集合呢？
8. 所有有序数组  $(a, b, c, d)$ （其中元素为实数）的集合是否是平直空间？
9. 所有在区间  $[0, 1]$  上连续的函数的集合是否是平直空间？所有有界函数的集合是否是平直空间？
10. 所有在区间  $[0, 1]$  上 3 次可微的函数的集合是否是平直空间？
11. 平面中的直线是否是平直空间？立体空间中的平面是否是平直空间。

## 4.1 子空间与和空间

看了各种各样的平直空间，我们来了解它们之间的关系。

首先，我们想知道平直空间的子集是否是平直空间。如果看平面和立体空间的例子，我们会发现，平面、立体空间的子集中，有些满足平直空间的定义，另一些则不满足。比如，可以验证，平面中，过原点的直线是平直空间，不过原点的不是；立体空间中，过原点的平面和直线是平直空间，不过原点的平面和直线不是平直空间。

直觉上，平直空间的子集里的元素也是满足向量的性质的，唯一的问题是：子集中的向量运算后是否还在子集中。只要这一点能够满足，平直空间的子集应该就是平直空间了。

**定义 4.1.1. 子空间** 平直空间的非空子集，如果（关于空间的运算）仍然是平直空间，就称为该平直空间的**子空间**。其中，不等于平直空间自身的子空间称为该平直空间的**真子空间**。

**定理 4.1.1.** 平直空间的非空子集，如果关于空间的运算封闭（也就是说，子集的元素经过运算后仍在子集中），就是子空间。

显然，由于运算不变，平直空间子集的元素自动满足运算法则。唯一需要验证的是零向量也在子集里。这就要由运算封闭的性质来保证。对子集中元素  $\mathbf{u}$ ，零向量可以写成  $\mathbf{u} - \mathbf{u}$ ，因此根据封闭性质，还在子集中。

用这个方法，可以验证  $\mathbb{Q}_n[x]$  是  $\mathbb{R}[x]$  的子空间，如果正整数  $n < m$ ，那么  $\mathbb{Q}_n[x]$  是  $\mathbb{R}_m[x]$  的子空间。

对于立体空间来说，过原点的平面是它的子空间，而过原点的直线是它的子空间，也是包含它的平面的子空间。可以证明：

**定理 4.1.2.** 给定平直空间  $\mathbb{V}$ ， $\mathbb{V}$  的两个子空间的交集仍然是  $\mathbb{V}$  的子空间。

所有子空间的交集是  $\{\mathbf{0}\}$ ，它是任何平直空间都会有的子空间。我们也说它是**平凡**的。

子空间的并集是否是子空间呢？比如，两个过原点的平面，并集是否是子空间呢？一般来说并不是，并集无法对向量的运算法则封闭。可以发现，并集要成为子空间的话，需要“更多的”向量。为此，我们引入类似于并集的概念。

**定义 4.1.2. 和空间**

给定平直空间  $\mathbb{V}$  以及它的两个子空间  $\mathbb{U}_1$ 、 $\mathbb{U}_2$ ， $\mathbb{U}_1$ 、 $\mathbb{U}_2$  的**和空间**是：

$$\{\mathbf{u}_1 + \mathbf{u}_2 \mid \mathbf{u}_1 \in \mathbb{U}_1, \mathbf{u}_2 \in \mathbb{U}_2\}.$$

记作  $\mathbb{U}_1 + \mathbb{U}_2$ 。

可以验证, 这样定义的和空间也是平直空间  $\mathbb{V}$  的子空间。

子空间的和满足结合律和交换律。用归纳法不难得出, 平直空间的有限个子空间也可以定义和空间。

给定立体空间的两个子空间: 过原点的直线  $\mathbb{R}\mathbf{v}_1$ 、 $\mathbb{R}\mathbf{v}_2$ , 它们的和空间就是  $\mathbb{R}\mathbf{v}_1 + \mathbb{R}\mathbf{v}_2$ 。如果  $\mathbf{v}_1$ 、 $\mathbf{v}_2$  不共轴, 那么和空间是平面, 否则是直线。

给定立体空间的两个子空间: 过原点的直线  $\mathbb{R}\mathbf{v}_1$  与平面  $\gamma: \mathbb{R}\mathbf{v}_2 + \mathbb{R}\mathbf{v}_3$ , 它们的和空间就是  $\mathbb{R}\mathbf{v}_1 + \mathbb{R}\mathbf{v}_2 + \mathbb{R}\mathbf{v}_3$ 。如果  $\mathbf{v}_1$  在  $\mathbb{R}\mathbf{v}_2 + \mathbb{R}\mathbf{v}_3$  中, 那么和空间就是  $\gamma$ 。否则, 和空间就是整个空间。

给定立体空间的两个子空间: 过原点的平面  $\gamma_1: \mathbb{R}\mathbf{u}_1 + \mathbb{R}\mathbf{v}_1$ 、 $\gamma_2: \mathbb{R}\mathbf{u}_2 + \mathbb{R}\mathbf{v}_2$ , 它们的和空间是  $\mathbb{R}\mathbf{u}_1 + \mathbb{R}\mathbf{v}_1 + \mathbb{R}\mathbf{u}_2 + \mathbb{R}\mathbf{v}_2$ 。如果  $\gamma_2$  中有不在  $\gamma_1$  中的向量, 那么和空间是整个空间。否则, 和空间就是  $\gamma_1$ 。

如果平直空间  $\mathbb{V} = \mathbb{U}_1 + \mathbb{U}_2$ , 且  $\mathbb{U}_1 \cap \mathbb{U}_2 = \{\mathbf{0}\}$ , 就说  $\mathbb{V}$  是  $\mathbb{U}_1$ 、 $\mathbb{U}_2$  的直和, 记作。

$$\mathbb{V} = \mathbb{U}_1 \oplus \mathbb{U}_2.$$

要注意的是, 这里的  $\oplus$  只是记录两个子空间交集为  $\{\mathbf{0}\}$  的性质, 并不是一种运算。

#### 思考 4.1.1.

1. 子空间的交集、和空间, 与子集的交集、并集有什么相似之处? 有没有类似空集的概念?
2. 能否定义子空间“不相交”? 能否把平直空间用子空间划分?
3.  $\mathbb{Q}_3[x]$  是  $\mathbb{R}_3[x]$  的子空间吗?  $\mathbb{Q}^{\mathbb{N}}$  是  $\mathbb{R}^{\mathbb{N}}$  的子空间吗?

#### 习题 4.1.1.

1. 给出以下平直空间的 (非平凡的) 子空间的例子。



1.  $\mathbb{R}_3[x]$
2.  $\{f: \mathbb{R} \rightarrow \mathbb{R} \mid f(0) = 0, f(1) = 0\}$
3.  $\{f \mid \partial^3 f = f\}$
4.  $\{(a, b, c) \mid a, b, c \in \mathbb{R}\}$
5.  $\mathbb{R}^{\mathbb{N}}$
6.  $\{f: [0, 1] \rightarrow \mathbb{R} \mid f \text{ 可微, 且 } \partial f(0.5) = 0\}$

2. 设有立体空间中的向量  $(1, 0, 0)$ 、 $(0, 1, -1)$ 。求包含它们的最小子空间。

3. 证明：如果非常数整式  $Q$  是  $P$  的因式，那么  $S_Q$  是  $S_P$  的子空间。

4. 证明：如果  $\mathbb{U}_1$ 、 $\mathbb{U}_2$  是  $\mathbb{V}$  的子空间，那么  $\mathbb{U}_1 + \mathbb{U}_2$  是  $\mathbb{V}$  的子空间。

5. 证明： $\mathbb{V}$  平直空间能写成两个子空间  $\mathbb{U}_1$ 、 $\mathbb{U}_2$  的直和，当且仅当  $\mathbb{V}$  中任意元素都能以唯一的方式写成  $\mathbb{U}_1$  中元素与  $\mathbb{U}_2$  中元素的和。

## 4.2 生成空间

以上的章节中，我们已经认识到，平直空间是一种特殊的集合。给定集合，我们可以将它补充成平直空间。

**定义 4.2.1. 生成空间** 给定系数域为  $\mathbb{K}$  的平直空间  $\mathbb{V}$  的子集  $A$ ，它的生成空间<sup>1</sup>：

$$\langle A \rangle = \left\{ \sum_{i=1}^n a_i \mathbf{u}_i \mid n \in \mathbb{Z}^+, \forall 1 \leq i \leq n \ a_i \in \mathbb{K}, \mathbf{u}_i \in A \right\}$$

即考虑集合中有限个元素的直组合，所有这样的直组合的集合就是它的生成空间。

反之，如果某个集合  $A$  的生成空间是平直空间  $\mathbb{U}$ ，就说  $A$  是  $\mathbb{U}$  的生成集。

约定空集的生成空间是平凡空间  $\{\mathbf{0}\}$ 。

---

<sup>1</sup>如果没有强调  $A$  是哪个平直空间的子集，一般来说，要约定系数域  $\mathbb{K}$ 。我们约定：如果没有强调，系数域一般是实数集  $\mathbb{R}$ 。

显然, 集合  $A$  的生成空间是它的母集。

容易看出, 如果某个平直空间包含集合  $A$ , 那么肯定也包含  $A$  的生成空间。因此,  $A$  的生成空间是所有包含  $A$  的平直空间的子空间。换句话说,  $A$  的生成空间是包含  $A$  的最小的平直空间。

作为推论, 不难得出: 平直空间的生成空间就是自己。

根据定义可以给定集合  $A, B$ , 那么:

- 如果  $A \subset B$ , 那么  $\langle A \rangle$  是  $\langle B \rangle$  的子空间。
- $\langle A \cap B \rangle$  是  $\langle A \rangle \cap \langle B \rangle$  的子空间。
- $\langle A \cup B \rangle = \langle A \rangle + \langle B \rangle$ 。

可以注意到,  $A$  中元素是否重复, 不妨碍定义它的生成空间。因此, 我们可以定义“允许元素重复的集合”:

#### 定义 4.2.2. 族

设  $A$  为集合, 在  $A$  中可重复地抽取元素, 放在一起, 称为族。族可以是有序的。

我们可以把生成空间的概念扩大到向量族的生成空间。这样定义的生成空间也有以上提到的基本性质。

如果某个平直空间能通过有限个向量生成, 就说它是有限生成空间。比如,  $\mathbb{R}_3[x]$  就是有限生成空间, 因为每个向量都是  $\{1, x, x^2, x^3\}$  的直组合。可以验证

$$\mathbb{R}_3[x] = \langle \{1, x, x^2, x^3\} \rangle.$$

而  $\mathbb{R}[x]$  就不是有限生成空间。考虑  $\mathbb{R}[x]$  中  $n$  个元素, 设它们之中最高的次数是  $m$ , 那么它们的直组合的次数不会超过  $m$ , 因此无法生成次数大于  $m$  的整式。

平面和立体空间的基本定理说明, 它们都是有限生成空间。

**思考 4.2.1.**

1. 给定集合  $A, B$ ,  $\langle A \cap B \rangle = \langle A \rangle \cap \langle B \rangle$  成立吗?
2. 所有  $[0, 1]$  上的连续函数的集合, 是否是有限生成空间?

**习题 4.2.1.**

1. 求实系数多项式集合  $\{x, 4x^3\}$  的生成空间。
2. 设  $E$  是  $(1, 2, 0)$ 、 $(0, -1, 1)$  生成的空间。把  $E$  写成一次方程的解集的形式, 求该一次方程。
3. 求  $\{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + z = 0\}$  的生成集。

## 4.3 基底和维数

如果我们考虑  $1, x, x^2$  这三个向量, 它们生成的空间是  $\mathbb{R}_2[x]$ 。那么  $1, x, 2x - 1, x^2$  这四个向量呢? 它们生成的空间仍然是  $\mathbb{R}_2[x]$ 。

我们可以发现, 相比  $1, x, x^2$ , 加入  $2x - 1$  这个向量, 有点“多余”。为什么呢? 因为  $2x - 1$  本身就是  $1, x$  的直组合, 加入之后并不能“贡献”什么新的东西。

容易发现, 用足够多的向量来生成空间并不是难事。比如, 空间  $\mathbb{V}$  自己肯定可以生成  $\mathbb{V}$ 。所以, 我们比较关心的是最小生成集的问题, 即: 生成空间  $\mathbb{V}$ , 最少要用多少个向量?

平面和立体空间的基本定理说明, 生成平面至少要两个向量, 生成立体空间至少要三个向量。

直觉上来看, 向量个数增多会让它们“更相关”, 数量减少则使它们更“不相关”。对平直空间来说, 如果  $A$  是它的生成集,  $B$  是空间中一组直无关的向量, 那么前者的元素应该比后者多。

可以证明 (见附录 C):

**定理 4.3.1.** 有限生成空间中，任何一族直无关的向量，向量的个数都不多于任意一个生成族的向量个数。

这是一个重要的结论。它告诉我们，有限生成空间中直无关的向量，个数有上限。因此，平直空间的最小生成族也就是最大的直无关族。

**定义 4.3.1. 平直空间的基底** 平直空间的基底，就是它的直无关的生成族。基底的向量称为基向量。

这个定义和我们已经定义过的平面、立体空间的基底是一致的。

$\mathbb{R}_3[x]$  的一个基底是  $(1, x, x^2, x^3)$ 。一般来说， $(1, x, \dots, x^n)$  是  $\mathbb{R}_n[x]$  的基底。

和平面、立体空间一样，平直空间的基底也满足：

**定理 4.3.2. 向量的唯一分解**  $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  是平直空间  $\mathcal{V}$  的一组基，当且仅当空间中每个元素  $\mathbf{u}$  都能以唯一的方式写成

$$\mathbf{u} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_n \mathbf{e}_n, \quad a_1, a_2, \dots, a_n \in \mathbb{F}.$$

称为  $\mathbf{u}$  的  $\mathcal{B}$  分解。 $a_1, a_2, \dots, a_m$  称为  $\mathbf{u}$  (在该分解下) 的分量。

有限生成空间必然有基底吗？直觉上，我们可以从空间的生成集出发，移除一些“多余”的向量，直到剩下的向量直无关为止。

**定理 4.3.3.** 可以通过适当地移除元素，把有限生成空间  $\mathcal{V}$  的每个生成族变成基底。

用具体的例子来说明。

$((1, 0), (8, 2), (3, 1), (0, 4), (0, 0))$  是  $\mathbb{R}^2$  的生成集。

我们首先去除零向量，得到  $((1, 0), (8, 2), (3, 1), (0, 4))$ 。这组向量直相关：

$$(8, 2) = 2 \cdot (1, 0) + 2 \cdot (3, 1).$$

去除  $(8, 2)$ , 得到  $((1, 0), (3, 1), (0, 4))$ . 这组向量仍然直相关:

$$(3, 1) = 3 \cdot (1, 0) + 0.25 \cdot (0, 4).$$

去除  $(3, 1)$ , 得到  $((1, 0), (0, 4))$ . 现在我们得到了一组直无关的向量, 而且能生成  $\mathbb{R}^2$ , 即  $\mathbb{R}^2$  的一组基。

从上面的定理可以推出:

**定理 4.3.4.** 任意有限生成空间都有基底。

空间的基底就是最小生成集, 因此大小应该相同。我们可以用定理 4.3.1 来证明。选定同一空间的两个基底  $\mathcal{B}_1$ 、 $\mathcal{B}_2$ , 它们都是直无关的生成集。根据定理 4.3.1,  $\mathcal{B}_1$  直无关,  $\mathcal{B}_2$  是生成集, 所以  $|\mathcal{B}_1| \leq |\mathcal{B}_2|$ 。同理可证,  $|\mathcal{B}_2| \leq |\mathcal{B}_1|$ 。因此  $|\mathcal{B}_1| = |\mathcal{B}_2|$ 。

我们把有限生成空间  $\mathbb{V}$  基底的元素个数称为空间的**维数**, 记为  $\dim \mathbb{V}$ 。平直空间的维数是它的内禀性质, 不依赖于附带的假设。

平面的维数是 2, 所以也常称为二维空间。立体空间维数是 3, 所以也称为三维空间。过原点的直线作为子空间, 维数是 1, 也称为一维空间。

由于有限生成空间必然有维数, 我们也称其为**有限维空间**。

空间的任意生成集都可以缩减为它的基底。那么任意直无关的向量族能否扩充为基底呢? 答案是肯定的。

**定理 4.3.5.** 可以通过适当地添加元素, 把有限生成空间  $\mathbb{V}$  的每一组直无关的向量变成基底。

我们甚至有更具体的结论:

**定理 4.3.6.** 给定有限生成空间  $\mathbb{V}$  的子空间  $\mathbb{U}$ , 存在另一个子空间  $\mathbb{W}$ , 使得

$$\mathbb{V} = \mathbb{U} + \mathbb{W}.$$

且  $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$ 。 $\mathbb{W}$  称为子空间  $\mathbb{U}$  的**补空间**, 一般记作  $\mathbb{U}^c$ 。

这让我们联想到集合的补集。实际上，我们还可以证明：

**定理 4.3.7. 维数基本定理** 设  $\mathbb{U}$  和  $\mathbb{W}$  是有限维空间  $\mathbb{V}$  的两个子空间，则它们的维数满足以下关系：

$$\dim(\mathbb{U} + \mathbb{W}) = \dim \mathbb{U} + \dim \mathbb{W} - \dim(\mathbb{U} \cap \mathbb{W}).$$

如果我们把空间的和比作集合的并，把空间的交比作集合的交，把空间的维数比作集合的元素个数，则这个关系和有限集合的关系是一致的。

#### 思考 4.3.1.

1. 能否把  $\mathbb{R}$  看作系数域为  $\mathbb{Q}$  的平直空间？它是否是有限维空间？
2. 基底中，基向量的排列顺序重要吗？按不同顺序排列的基向量，能否看做不同的基底？给出你的理由。

#### 习题 4.3.1.

1. 说明以下结论：
  - 1.1. 有限维空间的生成集的元素个数不小于空间的维数。
  - 1.2. 有限维空间的直无关向量族的元素个数不小大于空间的维数。
  - 1.3. 有限维空间的生成集的元素个数等于空间的维数，则生成集是空间的基底。
  - 1.4. 有限维空间的直无关向量族的元素个数等于空间的维数，则直无关向量族是空间的基底。
2. 给出  $\{(x, y, z) \in \mathbb{R}^3 \mid x + 2y + z = 0\}$  的一个补空间。
3. 请把  $(1, 1, 1)$ 、 $(0, -1, 1)$  扩充为  $\mathbb{R}^3$  的基底。
4. 设二次可微函数  $f : \mathbb{R} \rightarrow \mathbb{R}$  是方程  $\partial^2 f + f = 0$  的解，并且  $f(0) = \partial f(0) = 0$ 。
  - 4.1. 对任意  $x$ ，证明：存在 0 与  $x$  之间的数  $c$ ，使得

$$f(x) = -\frac{f(c)x^2}{2}.$$

- 4.2. 设  $M$  是  $|f|$  在  $[0, 1]$  上的最大值。证明  $M \leq \frac{M}{2}$ ，从而  $M = 0$ 。
- 4.3. 证明  $f$  恒等于 0。

5. 设  $S_2$  为满足方程  $\partial^2 f + f = 0$  的实变函数的集合。

5.1. 证明：如果  $f, g \in S_2$  满足  $f(0) = g(0)$ ,  $\partial f(0) = \partial g(0)$ , 那么  $f = g$ 。

5.2. 分别找出  $S_2$  中满足  $f(0) = 0$ ,  $\partial f(0) = 1$  以及  $f(0) = 1$ ,  $\partial f(0) = 0$  的函数  $f_1$ 、 $f_2$ 。

5.3. 证明： $\{f_1, f_2\}$  是  $S_2$  的基底。给出  $S_2$  的维数。





## 第五章 级数

研究可微函数时，我们讨论过分析函数在某点附近的行为的问题。我们的研究方法是：把函数在该点附近表示成多项式的形式。换句话说，我们把函数表示成一系列简单函数的和。比如，指数函数在 0 附近可以写成：

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!} + o(x^n).$$

那么，我们能不能把  $e^x$  直接写成无穷多个简单函数的和呢？

具体来说，对于每个  $x$ ，我们能不能说， $e^x$  就是把“所有”  $\frac{x^n}{n!}$  这样的项加起来的和呢？

我们并未定义过“无穷多个数的和”。不过，这个问题可以转化成一个数列的问题。学习数列的时候，我们研究了趋于极限或趋于无穷的数列。上面的问题，换句话说，就是问数列：

$$1, 1 + x, 1 + x + \frac{x^2}{2}, \cdots, 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!}, \cdots$$

有没有极限。

传统上，我们把这样的数列用另一种方式表示，称为**级数**。

### 定义 5.0.1. 级数

给定数列  $\{a_n\}_{n \in \mathbb{N}}$ ，表达式：

$$a_0 + a_1 + a_2 + \cdots + a_n + \cdots$$

称为数列  $\{a_n\}$  的**级数**，简记为  $\sum_{n \in \mathbb{N}} a_n$  或  $\sum a_n$ 。  $a_n$  称为**级数的通项**， $\{a_n\}_{n \in \mathbb{N}}$  也称为级数的**通项数列**。

如果  $\{a_n\}$  的部分和序列：

$$a_0, a_0 + a_1, \dots, \sum_{i=0}^n a_i, \dots$$

有极限，就说**级数收敛**，其极限称为**级数和**，记作：

$$\sum_{i=0}^{\infty} a_i$$

否则说**级数发散**。

对于数列  $\{\frac{x^n}{n!}\}_{n \in \mathbb{N}}$  来说，它的级数就是

$$1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots$$

这个级数是否收敛，收敛的话，级数和是否是  $e^x$ ，就是我们希望研究的内容。

## 5.1 正项级数

如果级数的通项都是正数或非负数，那么就说级数是**正项级数**。

正项级数是最方便研究的级数。对一般级数的研究，经常会转化到正项级数。

正项级数的部分和数列是单调递增数列。我们知道，单调递增数列要么有极限，要么趋于正无穷。这两种情况分别对应级数收敛、级数发散。

举例来说，级数  $\sum 2^{-n}$  是正项级数。它的部分和  $H_n$  为：

$$H_n = 1 + \frac{1}{2} + \dots + \frac{1}{2^n}.$$

它的通项数列是无穷递缩等比数列, 因此  $H_n = 2 - 2^{-n}$ 。部分和数列  $\{H_n\}$  趋于 2。我们说级数  $\sum 2^{-n}$  收敛, 级数和是 2。

同理, 级数  $\sum r^{-n}$  的部分和  $H_n$  为

$$H_n = 1 + \frac{1}{r} + \cdots + \frac{1}{r^n} = \frac{1 - r^{-n-1}}{1 - r^{-1}}.$$

因此,  $r > 1$  时, 部分和数列  $\{H_n\}$  趋于  $\frac{r}{r-1}$ , 级数和是  $\frac{r}{r-1}$ 。

研究级数收敛, 和研究数列收敛, 有什么不同呢? 实际上, 级数是研究数列收敛性质的工具。

研究数列  $\{a_n\}$  是否收敛时, 我们发现, 将数列相邻项作差:

$$a_0, a_1 - a_0, a_2 - a_1, \cdots$$

这样得到的数列的性质, 对研究数列是否收敛更有用。我们把作差得到的数列称为原数列的**差分数列**。可以验证, 数列  $\{a_n\}$  的差分数列的部分和为:

$$H_n = a_0 + (a_1 - a_0) + (a_2 - a_1) + \cdots + (a_n - a_{n-1}) = a_n.$$

也就是说, **数列的差分数列的级数, 就是原数列**。

因此, 研究数列是否收敛, 就看它的差分数列, 研究差分数列的级数是否有级数和。如果差分数列有级数和, 那么级数和就是原数列的极限。

### 思考 5.1.1.

1. 通过差分数列来研究数列性质的方法, 与通过研究函数的微变率、微变函数来研究函数性质的方法, 有哪些共同之处? 有哪些不同之处?

## 5.2 正项级数的收敛

那么, 怎样的正项级数收敛呢?

首先来看级数  $\sum \frac{n}{n+1}$ 。这个级数的通项  $\frac{n}{n+1}$  趋于 1。也就是说,  $n$  足够大时, 每增加一项, 就等于给部分和加 1。这样看来, 级数不收敛。

下面严格来说明, 由于

$$\lim_{n \rightarrow +\infty} \frac{n}{n+1} = 1,$$

因此, 对于  $0 < r < 1$ , 总存在正整数  $N$ , 只要  $n > N$ , 就有  $\frac{n}{n+1} > 1 - r$ 。于是, 部分和  $H_n$  可以写成:

$$H_n > H_N + (n - N) \cdot (1 - r).$$

而数列  $\{(n - N) \cdot (1 - r)\}$  趋于正无穷, 所以部分和  $H_n$  也趋于正无穷。也就是说, 级数  $\sum \frac{n}{n+1}$  发散。

从这个证明可以看出, 只要正项级数的通项不收敛到 0, 就总会有比较大的项加到部分和里, 于是级数就会发散。我们将其总结为:

**定理 5.2.1.** 如果正项级数的通项没有极限或极限不为零, 那么级数发散。

**证明:** 设正项级数  $\sum a_n$  的通项数列不收敛到 0, 那么按照定义, 存在  $r > 0$ , 使得对任意正整数  $N$ , 总有  $n > N$ , 让  $a_n \geq r$ 。

因此, 我们可以找到第一个  $n_1$ , 使得  $a_{n_1} \geq r$ 。把  $n_1$  作为  $N$ , 又可以找到  $n_2 > n_1$ , 使得  $a_{n_2} \geq r$ 。以此类推, 我们可以找到  $\{a_n\}$  的子列, 其中每个元素都大于等于  $r$ 。

考虑级数  $\sum a_n$  的部分和  $H_n$ , 当  $n$  大于子列的第  $m$  个下标时,  $H_n$  就大于等于  $mr$ 。因此, 部分和趋于正无穷。这说明级数不收敛。  $\square$

如果通项数列  $\{a_n\}$  收敛到 0, 级数是否就收敛呢?

考虑级数  $\sum \frac{1}{n+1}$ , 它的通项是  $\frac{1}{n+1}$ , 因此趋于 0。但我们可以证明, 级数  $\sum \frac{1}{n+1}$  不收敛。

考虑数列  $\{\frac{1}{n+1}\}$  的前  $2^m$  项, 我们可以把它这样拆分:

$$\begin{aligned}
 1 &\geq 1 \\
 \frac{1}{2} &\geq \frac{1}{2} \\
 \frac{1}{3} + \frac{1}{4} &\geq 2 \cdot \frac{1}{4} = \frac{1}{2} \\
 \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} &\geq 4 \cdot \frac{1}{8} = \frac{1}{2} \\
 &\vdots \\
 \frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \cdots + \frac{1}{2^m} &\geq 2^{m-1} \cdot \frac{1}{2^m} = \frac{1}{2}
 \end{aligned}$$

因此, 级数前  $2^m$  项的部分和大于等于  $1 + \frac{1}{2} \cdot (m-1) = \frac{m+1}{2}$ 。这说明级数不收敛。

再看另一个级数:  $\sum \frac{1}{(n+1)^2}$ , 它的通项是  $\frac{1}{(n+1)^2}$ , 同样趋于 0。然而, 我们可以证明, 这个级数是收敛的。这是因为对任意正整数  $n$ ,

$$\frac{1}{(n+1)^2} < \frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}.$$

因此, 部分和  $H_n$  可以写为:

$$H_n = 1 + \sum_{i=1}^n \frac{1}{(n+1)^2} < 1 + \sum_{i=1}^n \left( \frac{1}{n} - \frac{1}{n+1} \right) = 1 + 1 - \frac{1}{n+1} < 2.$$

部分和数列有界, 因此收敛。

从上面的讨论可以顺便推出: 级数  $\sum \frac{1}{(n+1)^a}$  在  $a \geq 2$  时收敛, 在  $a \leq 1$  时发散。直觉上理解, 就是  $\{\frac{1}{n+1}\}$  收敛到 0 还不够快, 导致部分和增长得太多; 而  $\sum \frac{1}{(n+1)^2}$ 、 $\sum r^{-n}$  收敛得足够快, 因此部分和始终有限。

怎样把“收敛得足够快”这件事用严谨的方式说明呢? 比照无穷的阶的形式, 我们来考虑通项数列相邻两项的比值:

$$q_n = \frac{a_n}{a_{n+1}}$$

直觉来说, 如果  $q_n$  趋于某个小于 1 的数, 那么  $\{a_n\}$  大体来说是增长的, 不收敛到 0, 于是级数发散。如果  $q_n$  趋于某个大于 1 的数  $r$ , 那么  $a_n$  收敛到 0, 而且收敛得大致和  $r^{-n}$  一样快, 于是级数收敛。

不过, 对级数  $\sum \frac{1}{n+1}$ 、 $\sum \frac{1}{(n+1)^2}$ , 比值  $\{q_n\}$  都收敛到 1。因此, 需要更细致的分析。遵循研究函数局部性质的思路, 我们可以把它展开:

$$q_n = 1 + \frac{t_1}{n} + \frac{t_2}{n^2} + \cdots$$

级数  $\sum \frac{1}{n+1}$  对应  $t_1 = 1$ , 而  $\sum \frac{1}{(n+1)^2}$  对应  $t_1 = 2$ 。

我们把  $q_n$  写成

$$q_n = 1 + \frac{t_1}{n} + o\left(\frac{1}{n}\right).$$

也就是:

$$n(q_n - 1) = t_1 + o(1).$$

即  $n(q_n - 1)$  趋于  $t_1$ 。

如果  $t_1 < 1$ , 那么按照极限的定义, 总有  $N$  和  $r < 1$  使得  $n > N$  时总有:  $n(q_n - 1) < r$ , 即:

$$na_n - (n+1)a_{n+1} < 0.$$

即  $n > N$  后, 数列  $\{na_n\}$  单调递增。因此  $a_n > \frac{Na_N}{n}$ 。于是, 级数  $\sum a_n$  的部分和增长得比  $Na_N \sum \frac{1}{n}$  快, 因此发散。

如果  $t_1 > 1$ , 那么按照极限的定义, 总有  $N$  和  $r > 1$  使得  $n \geq N$  时总有:  $n(q_n - 1) > r$ , 即

$$na_n - (n+1)a_{n+1} \geq (r-1)a_{n+1}.$$

同理, 可以写出:

$$\begin{aligned}(n-1)a_{n-1} - na_n &\geq (r-1)a_n \\(n-2)a_{n-2} - (n-1)a_{n-1} &\geq (r-1)a_{n-1} \\&\vdots \quad \geq \quad \vdots \\Na_N - na_n &\geq (r-1)a_{N+1}\end{aligned}$$

以上等式分边相加, 就得到:

$$Na_N \geq Na_N - na_n \geq (r-1) \sum_{i=N+1}^n a_i.$$

于是部分和  $H_n$  可以写成:

$$H_n \leq H_N + Na_N.$$

这说明级数收敛。

$t_1 = 1$  时, 需要更精细的分析, 这里就不继续展开了。我们把以上的内容总结一下:

记正项级数  $\sum a_n$  的通项数列相邻两项的比值为:

$$q_n = \frac{a_n}{a_{n+1}}$$

如果在  $n$  足够大的时候, 总有以下情况, 则可以做出正项级数是否收敛的判断。

1. 如果  $q_n$  总大于某个大于 1 的数, 那么级数收敛;
2. 如果  $q_n$  总小于某个小于 1 的数, 那么级数发散;
3. 如果  $q_n$  可以写成  $1 + \frac{A}{n} + o(\frac{1}{n})$ , 且  $A > 1$ , 那么级数收敛;
4. 如果  $q_n$  可以写成  $1 + \frac{A}{n} + o(\frac{1}{n})$ , 且  $A < 1$ , 那么级数发散。

来看一些具体的例子。

级数  $\sum \frac{1}{n!}$  收敛, 因为

$$\forall n \geq 1, q_n = \frac{(n+1)!}{n!} = n+1 \geq 2.$$

$x > 0$  时, 级数  $\sum \frac{x^n}{n!}$  收敛, 因为

$$\forall n \geq x, q_n = \frac{x^n(n+1)!}{n!x^{n+1}} = \frac{n+1}{x} \geq 1 + \frac{1}{x}.$$

$x > 1$  时, 级数  $\sum \frac{n^k}{x^n}$  收敛, 因为

$$\forall n \geq 1, q_n = \frac{n^k x^{n+1}}{x^n(n+1)^k} = x \left(1 + \frac{1}{n}\right)^{-k}.$$

而数列  $\left\{\left(1 + \frac{1}{n}\right)^k\right\}$  趋于 1, 因此  $q_n$  趋于  $x > 1$ 。

级数  $\sum n^a$  的收敛情况跟  $a$  有关。比值  $q_n = \left(1 + \frac{1}{n}\right)^{-a}$ 。

- 如果  $a \leq 0$ , 那么级数发散。
- 如果  $a < 0$ , 那么  $q_n = \left(1 + \frac{1}{n}\right)^{-a} = 1 - \frac{a}{n} + o\left(\frac{1}{n}\right)$ 。
- 如果  $-1 < a < 0$ , 那么级数发散。
- 如果  $a = -1$ , 那么级数发散。
- 如果  $a < -1$ , 那么级数收敛。

### 思考 5.2.1.

1. 看了研究级数收敛性质的方法, 对你研究函数的性质, 有什么启发?
2. 如果数列从某一项开始都是正数, 这样的数列能否通过研究正项级数来讨论收敛性? 进一步来说, 数列的负项如果更多的话, 如何转为正项级数来研究?

### 习题 5.2.1.

1. 以下级数是否收敛? 说明原因。

$$\begin{array}{ll} 1. \sum \frac{2}{n^2-3n+3} & 2. \sum \frac{n!}{n^n} \\ 3. \sum \frac{1}{\sqrt{n+1}} & 4. \sum \frac{4^n(n!)^2}{(2n+1)!} \end{array}$$



2. 给定  $x > 0$ , 证明: 级数  $\sum \frac{x^n}{n!}$  收敛, 级数和为  $e^x$ 。

3. 考虑正项级数  $\sum a_n$ 。如果有正数列  $\{B_n\}$ , 正数  $c$ , 以及正整数  $N$ , 使得

$$\forall n > N, B_n \frac{a_n}{a_{n+1}} - B_{n+1} \geq c.$$

证明: 级数  $\sum a_n$  收敛。

4. 证明: 如果正项级数  $\sum a_n$  收敛, 那么可以找到正数列  $\{B_n\}$ , 正数  $c$ , 以及正整数  $N$ , 使得

$$\forall n > N, B_n \frac{a_n}{a_{n+1}} - B_{n+1} \geq c.$$

5. 如果正项级数  $\sum a_n$  从某一项起总有

$$\frac{a_{n+1}}{a_n} \geq 1 - \frac{1}{n},$$

那么级数  $\sum a_n$  发散。

## 5.3 一般级数的收敛

上一节我们讨论了正项级数的收敛性质。对于一般的级数, 怎么判断它是否收敛呢?

最简单的方法是考虑“绝对值”。我们把级数的绝对值定义为给通项取绝对值得到的级数。这样定义之下, 级数的绝对值肯定是正项级数。

**定理 5.3.1.** 如果级数的绝对值收敛, 那么级数收敛。

**证明:** 给定级数  $\sum a_n$ , 它的绝对值是正项级数  $\sum |a_n|$ 。如果后者收敛, 那么对应的部分和数列是自敛数列。因此, 对任意  $r > 0$ , 存在正整数  $N$ , 只要  $n > m > N$ , 就有

$$\sum_{i=m+1}^n |a_i| < r.$$

因此

$$\left| \sum_{i=m+1}^n a_i \right| \leq \sum_{i=m+1}^n |a_i| < r.$$

这说明级数  $\sum a_n$  的部分和数列是自敛数列。也就是说，级数  $\sum a_n$  收敛。

□

如果级数的绝对值收敛，就说它**绝对收敛**。正项级数显然都是绝对收敛的。给收敛的正项级数的通项随意加上正负号，得到的级数也绝对收敛。比如，从正项级数  $\sum \frac{1}{2^n}$  收敛，可以推出  $\sum \frac{(-1)^n}{2^n}$  绝对收敛。

**例题 5.3.1.** 证明：对任意实数  $x$ ， $\sum_{n \in \mathbb{N}} \frac{x^n}{n!}$  收敛。

**解答.**  $x \geq 0$  时，我们已经证明了正项级数  $\sum_{n \in \mathbb{N}} \frac{x^n}{n!}$  收敛。对于一般的  $x$ ，考虑级数的绝对值： $\sum_{n \in \mathbb{N}} \left| \frac{x^n}{n!} \right| = \sum_{n \in \mathbb{N}} \frac{|x|^n}{n!}$ ， $|x| \geq 0$ ，因此又回到了正项级数的情况，于是级数的绝对值收敛。也就是说， $\sum_{n \in \mathbb{N}} \frac{x^n}{n!}$  绝对收敛。

如果级数的绝对值发散，级数还能收敛吗？这样的级数也是存在的。我们称它们为**相对收敛**的级数。

举例来说，级数  $\sum \frac{(-1)^n}{(n+1)}$  不是绝对收敛的，因为它的绝对值  $\sum \frac{1}{(n+1)}$  发散。但我们可以证明  $\sum \frac{(-1)^n}{(n+1)}$  收敛。

考虑部分和数列的子列： $\{H_{2m}\}_{m \in \mathbb{N}}$  和  $\{H_{2m+1}\}_{m \in \mathbb{N}}$ 。我们证明两者都趋于同一个极限  $X$ ，因此部分和数列也趋于  $X$ 。

数列的偶数项大于等于 0，奇数项小于等于 0。

$$\begin{aligned} H_{2m+2} &= H_{2m} - \frac{1}{2m+2} + \frac{1}{2m+3} \leq H_{2m}, \\ H_{2m+3} &= H_{2m+1} + \frac{1}{2m+3} - \frac{1}{2m+4} \geq H_{2m+1}. \end{aligned}$$

即  $\{H_{2m}\}$  单调递减， $\{H_{2m+1}\}$  单调递增。即  $H_{2m} \leq H_2$ ， $H_{2m+1} \geq H_1$ 。

两者的通项之差为  $H_{2m} - H_{2m+1} = \frac{1}{2m+1} \geq 0$ ，所以

$$H_1 \leq H_{2m+1} \leq H_{2m} \leq H_2.$$

即两者都是有界单调数列, 因此都有极限。而它们的差  $\frac{1}{2m+1}$  趋于 0, 因此极限相等。

以上的情形可以稍做推广:

**定理 5.3.2.** 如果级数  $\sum a_n$  的通项正负交替, 且绝对值单调递减, 趋于 0, 那么级数收敛<sup>1</sup>。

**思考 5.3.1.**

1. 用级数的概念, 定义  $0.\dot{9}$ 。这样定义的  $0.\dot{9}$  是什么数? 它有多大?
2. 用级数的概念, 定义“无穷不循环小数”。这样定义的“无穷不循环小数”是什么数?
3. 如果级数的通项正负交替且趋于 0, 但绝对值不单调递减, 级数是否也收敛?
4. 如果级数  $\sum a_n$  的通项绝对值单调递减且趋于 0, 且有无穷多个正项和负项, 是否收敛? 正负项的排列满足什么条件时收敛? 给出你的想法。

**习题 5.3.1.**

1. 以下数列绝对收敛吗? 相对收敛吗?

$$\begin{array}{ll} 1. \sum \frac{(-1)^n}{(n+2)\sqrt{n+1}} & 2. \sum \left(1 - \cos\left(\frac{\pi}{n}\right)\right) \\ 3. \sum \frac{\sin \frac{n\pi}{6}}{n} & 4. \sum \frac{\sqrt{n} \sin n}{n^2+1} \end{array}$$

2. 考虑两个数列  $\{u_n\}$ 、 $\{v_n\}$ , 我们要研究级数  $\sum u_n v_n$  的收敛性质。记  $\{u_n\}$  的部分和数列为  $\{s_n\}$ 。

- 2.1. 设  $p \leq q$  为正整数, 证明以下恒等式。

$$\sum_{i=p}^q u_i v_i = s_q v_q - s_{p-1} v_p + \sum_{i=p}^{q-1} s_i (v_i - v_{i+1}).$$

- 2.2. 如果  $\{s_n\}$  是有界数列,  $\{v_n\}$  是单调递减且趋于 0 的正项数列, 证明: 级数  $\sum u_n v_n$  收敛。

---

<sup>1</sup>提醒: 这样的通项数列  $\{a_n\}$  就是我们见过的交替无穷小。

2.3. 使用以上结论, 研究级数:

$$\sum \frac{\sin(n\theta)}{\sqrt{n}}$$

的收敛性质。

3. 比较两个级数。

3.1. 证明: 级数  $\frac{(-1)^n}{\sqrt{n+1}}$  收敛。

3.2. 证明:

$$\frac{(-1)^n}{\sqrt{n+1} + (-1)^n} = \frac{(-1)^n}{\sqrt{n+1}} - \frac{1}{(n+1)} + \frac{(-1)^n}{(n+1)\sqrt{n+1}} + o\left(\frac{1}{n\sqrt{n}}\right)$$

3.3. 级数  $\frac{(-1)^n}{\sqrt{n+1} + (-1)^n}$  是否收敛? 从中可以得到什么结论?

## 附录 A 无穷集合的势

**定义 1.0.1. 像和原像** 给定映射  $f$ ，记  $f$  的定义域为  $D$ 。 $f$  将集合  $A \subseteq D$  映射得到的集合，称为  $A$  经过  $f$  映射的**像**，记作  $f(A)$ 。给定  $f$  值域的子集  $B$ ，将集合：

$$\{x \in D \mid f(x) \in B\}$$

称为  $B$ （关于  $f$ ）的**原像**。

**公理. 选择公理** 可以同时从任意多个非空集合里各选择一个元素，构成一个新集合。

从选择公理出发，我们可以得出一些有用的基本结论。

**定理 1.0.1. 构造双射** 设  $f$  是从集合  $A$  到  $B$  的单射，那么  $f$  是  $A$  到  $f(A)$  的双射。

设  $f$  是从集合  $A$  到  $B$  的满射，那么存在从  $A$  的子集到  $B$  的双射  $g$ ，使得  $g(x) = f(x)$  总成立。

这个定理其实说明了单射和满射相比双射“缺了什么”，以及如何修正，得到双射。

单射的“问题”在于它的值域“填不满”到达域，构不成满射。因此，把到达域中不在它值域里的部分去掉，只考虑它的值域作为到达域，那么它就是双射了。

满射的“问题”在于应变量可能是多个自变量映射的结果，也就是“多对一”，导致不是单射。因此，只要在里面选一个自变量就好了。而选择公理告诉我们，这么做是没问题的。

**证明：** 首先证明，可以从单射构造双射。设  $f$  是从集合  $A$  到  $B$  的单射，那么按照定义， $f$  是  $A$  到  $f(A) \subseteq B$  的满射。

$f$  既是单射又是满射，因此  $f$  是  $A$  到  $f(A)$  的双射。

再证明可以从满射构造双射。设  $f$  是从集合  $A$  到  $B$  的满射，那么按照定义， $B$  中每个元素  $y$  关于  $f$  的原像都是非空集合。

考虑  $B$  所有元素关于  $f$  的原像。根据选择公理，可以同时在这些原像中各选择一个元素。这样构成的集合是  $A$  的子集  $A'$ 。

我们定义新的映射

$$\begin{aligned} g: A' &\rightarrow B \\ x &\mapsto f(x) \end{aligned}$$

这样定义的  $g$ ，是从  $A$  的子集  $A'$  到  $B$  的满射。而且按照定义，总有  $g(x) = f(x)$ 。而由于不同的自变量  $x$  来自不同的原像，所以经过  $f$  映射的结果也不相同。这说明  $g$  是单射。因此  $g$  是双射。

□

**定义 1.0.2. 集合上的序** 给定一个集合，可以对其中的元素指定顺序。具体来说，序是一种二元关系。比如，整数集、有理数集或实数集上的大于、小于、大于等于、小于等于关系，都是序关系。

如果序关系  $\triangleleft$  满足以下条件，就称为**偏序**：

1. 自反性： $\forall x, x \triangleleft x$ 。
2. 反称性： $\forall x, y$ ，如果  $x \triangleleft y$  且  $y \triangleleft x$ ，那么  $x = y$ 。

3. 传递性:  $\forall x, y, z$ , 如果  $x \triangleleft y$  且  $y \triangleleft z$ , 那么  $x \triangleleft z$ 。

配备有偏序关系的集合称为**偏序集**。

偏序集不保证集合中任意两个元素都有序关系。如果集合中任意两个元素  $x, y$ , 总有  $x \triangleleft y$  或  $y \triangleleft x$  成立, 就说序关系是**全序**, 集合是**全序集**。

给定  $S$  上的序关系  $\triangleleft$ , 如果集合  $S$  中有元素  $x$ , 使得  $x \triangleleft y$  对  $S$  中任意元素  $y$  成立, 就说  $x$  是  $S$  的**最小元**。如果集合  $S$  中有元素  $x$ , 使得  $y \triangleleft x$  对  $S$  中任意元素  $y$  成立, 就说  $x$  是  $S$  的**最大元**。比如, 正整数集上的小于等于关系, 就有最小元 1, 通常称为最小值。

如果序关系使得集合的任意非空子集总有最小元, 就说它是**良序**, 集合是**良序集**。比如, 正整数集上的小于等于关系就是良序。

**定义 1.0.3.** 设有集合  $A$  和  $B$ 。

- 如果存在  $A$  到  $B$  的双射, 就说  $A$  和  $B$  的势相等, 两者等势, 记作  $|A| = |B|$ 。
- 如果存在从  $A$  到  $B$  的单射, 就说  $A$  的势不大于  $B$  的势或  $A$  的势小于等于  $B$  的势, 记作  $|A| \leq |B|$  或  $|B| \geq |A|$ 。
- 如果存在从  $A$  到  $B$  的单射, 但不存在  $B$  到  $A$  的单射, 就说  $A$  的势小于  $B$  的势, 记作  $|A| < |B|$  或  $|B| > |A|$ 。

**定理 1.0.2. 等势定理** 给定集合  $A, B$ , 如果  $|A| \leq |B|$ 、 $|B| \leq |A|$ , 那么  $|A| = |B|$ 。

这个定理是判定集合等势的重要方法。它的证明思路大致是这样的: 我们有从  $A$  到  $B$  的单射  $f$ , 也有从  $B$  到  $A$  的单射  $g$ 。根据前面的结论,  $f$  是  $A$  到  $B$  子集的双射,  $g$  又把  $B$  的这部分一一对应到  $A$  的一部分, 于是变成“更小”的双射。这个相互跳转的过程, 集合  $A, B$  会不断“剥离出”一部分。它们是通过双射得到的, 所以它们之间可以建立双射。而双方剩下的部分也可以建立双射。于是我们就在  $A, B$  之间建立了双射。

**证明：** 根据条件，存在从  $A$  到  $B$  的单射  $f$ ，以及从  $B$  到  $A$  的单射  $g$ 。记以下集合：

$$\begin{aligned} A_0 &= A, & B_0 &= B \\ A_1 &= g(B_0), & B_1 &= f(A_0) \\ A_2 &= g(B_1), & B_2 &= f(A_1) \\ &\vdots & &\vdots \\ A_{n+1} &= g(B_n), & B_{n+1} &= f(A_n) \\ &\vdots & &\vdots \end{aligned}$$

根据定理 1.0.1,  $f$  是  $A_0$  到  $B_1$  的双射,  $g$  是  $B_1$  到  $A_2$  的双射, 等等。因此, 我们可以写出以下两条等势关系“链”:

$$\begin{aligned} |A| &= |A_0| = |B_1| = |A_2| = |B_3| = |A_4| = |B_5| = |A_6| = \cdots \\ |B| &= |B_0| = |A_1| = |B_2| = |A_3| = |B_4| = |A_5| = |B_6| = \cdots \end{aligned}$$

其中集合的关系为:

$$\begin{aligned} A &= A_0 \supseteq A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \\ B &= B_0 \supseteq B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots \end{aligned}$$

以上集合关系中, 如果对某个正整数  $n$ , 有  $A_n = A_{n+1}$  或  $B_n = B_{n+1}$ , 那么两条等势“链”就连上了, 于是立刻得到  $|A| = |B|$ 。

如果所有的集合关系都是  $A_n \subset A_{n+1}$ 、 $B_n \subset B_{n+1}$ , 即真子集关系, 那么考虑  $A_n$  中去掉  $A_{n+1}$  的部分, 记为  $A_n^*$ 。同理定义  $B_n^*$ 。

$A_n$  到  $B_{n+1}$  有用  $f$  定义的双射,  $A_{n+1}$  到  $B_{n+2}$  有用  $f$  定义的双射, 因此,  $A_n^*$  到  $B_n^*$  也有用  $f$  定义的双射。我们可以得到以下两条等势关系“链”:

$$\begin{aligned} |A_0^*| &= |B_1^*| = |A_2^*| = |B_3^*| = |A_4^*| = \cdots \\ |B_0^*| &= |A_1^*| = |B_2^*| = |A_3^*| = |B_4^*| = \cdots \end{aligned}$$



$|A_0^*| = |B_1^*|$ ,  $|A_1^*| = |B_0^*|$ , 而  $A_0^*$  与  $A_1^*$  不相交, 而  $B_0^*$  与  $B_1^*$  不相交。因此, 这两对集合可以拼成  $A_0^* \cup A_1^*$  到  $B_0^* \cup B_1^*$  的“整齐”的双射。其中  $A_0^*$  的元素映射到  $B_1^*$ ,  $B_0^*$  的元素映射到  $A_1^*$ , 互不干扰。

对于一般的相邻奇偶数下标, 也可以这么做。因此, 我们可以构建这样的映射:

$$h: \bigcup_{n \in \mathbb{N}} A_n^* \rightarrow \bigcup_{n \in \mathbb{N}} B_n^*$$

$$x \mapsto \begin{cases} f(x) & \text{如果 } x \in A_{2m}^*, m \in \mathbb{N} \\ g(x) & \text{如果 } x \in A_{2m+1}^*, m \in \mathbb{N} \end{cases}$$

由于对所有的  $n$ ,  $A_n^*$  两两不相交,  $B_n^*$  两两不相交。所以以上定义的  $h$  成立, 而且是双射。

如果集合  $A^* = \bigcup_{n \in \mathbb{N}} A_n^*$  就是  $A$ , 同理  $B^* = \bigcup_{n \in \mathbb{N}} B_n^*$  就是  $B$ , 我们就在  $A$  和  $B$  之间建立了双射。不过很可惜,  $A^*$  还不是  $A$ , 遗漏了一部分元素。

遗漏了怎样的元素呢? 按照定义,  $\bigcup_{n \in \mathbb{N}} A_n^*$  是所有“在某个  $A_n$  中但不在  $A_{n+1}$  中”的元素的集合。因此, 只要  $A$  中一个元素同时在所有的  $A_n$  里, 它就不属于  $A^*$ 。反之,  $A$  中元素  $x$  只要不是同时属于所有的  $A_n$ , 那么总有一个最小的正整数  $m$ ,  $x$  不属于  $A_m$ 。也就是说,  $x$  属于  $A_{m-1}$  而不属于  $A_m$ 。这说明  $x \in A_{m-1}^*$ 。

综上所述, 设  $\overline{A}$  为所有  $A_n$  的交集, 那么  $A$  是  $A^*$  和  $\overline{A}$  的不交并。 $A^*$  和  $\overline{A}$  是  $A$  的分划。同理可以定义  $\overline{B}$  为所有  $B_n$  的交集, 且  $B^*$  和  $\overline{B}$  是  $B$  的分划。

只要有  $\overline{A}$  到  $\overline{B}$  的双射  $j$ , 那么用类似前面的思路, 我们可以建立  $A$

到  $B$  的双射:

$$H: A \rightarrow B$$

$$x \mapsto \begin{cases} h(x) & \text{如果 } x \in A^* \\ j(x) & \text{如果 } x \in \overline{A} \end{cases}$$

最后证明双射  $j$  存在。考虑  $y \in \overline{B}$ , 按照定义,  $y \in B_1 = f(A)$ , 所以它关于  $f$  的原像是非空集合。而  $f$  是单射, 所以其中恰有一个元素  $x \in A$ 。对任意正整数  $n$ ,  $y \in B_n = f(A_{n-1})$ , 所以  $y$  的原像  $x$  属于  $A_{n-1}$ 。因此,  $x \in \overline{A}$ 。这说明  $f$  是  $\overline{A}$  到  $\overline{B}$  的满射, 因此是双射, 即我们要找的双射  $j$ 。

□

**定理 1.0.3.** 集合的势之间的小于等于关系是偏序。

**证明:** 按照定义进行证明。

自反性: 集合总有到自己的单射 (恒等映射), 因此  $|A| \leq |A|$  总成立。

反称性: 根据等势定理, 反称性成立。

传递性: 给定集合  $A, B, C$ , 如果有  $A$  到  $B$  的单射  $f$ ,  $B$  到  $C$  的单射  $g$ , 那么两者的复合映射  $g \circ f$  就是  $A$  到  $C$  的单射。因此传递性成立。

□

**定理 1.0.4.** 实数集  $\mathbb{R}$  和  $2^{\mathbb{N}}$  等势。

**证明一:** 直接建立双射。

**证明:** 直接建立  $\mathbb{R}$  和  $2^{\mathbb{N}}$  之间的双射, 比较困难。我们使用区间  $(0, 1)$  作为“中介”。首先建立  $(0, 1)$  到  $\mathbb{R}$  的双射  $f_1$ , 然后建立  $2^{\mathbb{N}}$  到  $(0, 1)$  的双射  $f_2$ 。这样,  $f_1$  和  $f_2$  的复合就是  $2^{\mathbb{N}}$  到  $\mathbb{R}$  的双射。

首先建立  $(0, 1)$  到  $\mathbb{R}$  的双射。考虑反三角函数  $\arccos$  在  $(-1, 1)$  上的取值，它把  $(-1, 1)$  上的数映射到  $(0, \pi)$  上，是从  $(-1, 1)$  到  $(0, \pi)$  的双射。而余切函数  $\cot$  则是从  $(0, \pi)$  到  $\mathbb{R}$  的双射。因此映射

$$f_1: x \mapsto \cot(\arccos(2x - 1))$$

是从  $(0, 1)$  到  $\mathbb{R}$  的双射。

再来把区间  $(0, 1)$  的实数和  $2^{\mathbb{N}}$  联系起来。考虑  $2^{\mathbb{N}}$  中的数列  $\{a_n\}_{n \in \mathbb{N}}$ 。

取数轴上区间  $[0, 1]$  的中点  $0.5$ ，它把  $[0, 1]$  平分两个区间。如果  $a_0 = 0$ ，就取左边的区间  $[0, 0.5]$ ，否则取右边的区间  $[0.5, 1]$ 。然后在新的闭区间里，根据  $a_1$  的值重复上一步的操作。这样不断下去，得到一个闭区间套。根据闭区间套定理，它趋于某个  $[0, 1]$  中的实数。

如果数列从某一项后全是  $0$  或全是  $1$ ，那么从某一步操作后我们将总是取左边（右边）的区间。这样闭区间套会收敛到这一步对应的区间的左（右）端点，它对应着某个分母是  $2$  的乘方的有理数。除此以外，区间套收敛到某个  $(0, 1)$  中的实数。这个实数不是分母是  $2$  的乘方的有理数。

把  $2^{\mathbb{N}}$  中所有从某一项后全是  $0$  或全是  $1$  的数列的集合记为  $A$ ，把  $(0, 1)$  中所有分母是  $2$  的乘方的有理数的集合记为  $B$ ，那么以上的操作构造了从  $2^{\mathbb{N}} \setminus A$  到  $(0, 1) \setminus B$  的映射  $g_1$ 。

下面证明  $g_1$  是双射。

首先证明  $g_1$  是单射。给定  $2^{\mathbb{N}} \setminus A$  中两个不同的数列，设它们最早从第  $k$  项起不同，那么对应的第  $k$  步操作时就会选择同一区间  $[a, b]$  的左部分  $[a, \frac{a+b}{2}]$  和右部分  $[\frac{a+b}{2}, b]$ 。而由于数列不会从某一项后全是  $0$  或全是  $1$ ，所以最终闭区间套不会收敛到端点上，也就是说两者经过  $f$  映射的结果分别在  $(a, \frac{a+b}{2})$  和  $(\frac{a+b}{2}, b)$  中，因此不相等。

再证明  $g_1$  是满射。 $\forall x \in (0, 1) \setminus B$ ，用以下操作构建数列  $\{a_n\}$ ：取数轴上区间  $[0, 1]$  的中点  $0.5$ ，它把  $[0, 1]$  平分两个区间。如果  $x < 0.5$ ，则

$a_0 = 0$ , 取左边的区间; 否则  $a_0 = 1$ , 取右边的区间。然后在新的区间里重复上一步的操作, 决定  $a_1$  的值。以此类推, 得到数列  $\{a_n\} \in 2^{\mathbb{N}}$ 。

以上操作同样得到一个闭区间套, 且收敛到  $\{x\}$ 。由于  $x \notin B$ , 所以不会有从某次操作后总是取左边(右边)的区间的情况, 也就是说, 得到的数列不在  $A$  中。这就说明  $x$  必然是  $2^{\mathbb{N}} \setminus A$  中某个数列经过  $g_1$  映射的结果, 也就是说,  $g_1$  是满射。

$g_1$  既是单射又是满射, 因此是双射。

另一方面, 考虑集合  $A$  和  $B$ 。我们来建立  $A$  和  $B$  之间的双射。

$B$  是有理数集的无穷子集, 因此是可数集合。而集合  $A$  可以作以下分划:

设  $k$  为自然数, 记  $A_k$  为所有最早自  $a_k$  起全是 0 或全是 1 的数列的集合。所有  $A_k$  两两不相交, 且并集是  $A$ 。

计算它们的势。  $|A_0| = |A_1| = 2$ ,

$$\forall k > 1, \quad |A_k| = 2^{k-1}.$$

因此,  $A_0, A_1$  和  $\{0, 1, 2, 3\}$  一一对应;  $k > 1$  时, 每个  $A_k$  和  $\{2^{k-1}+3, 2^{k-1}+4, \dots, 2^k+2\}$  一一对应。于是  $A$  和自然数集一一对应。

于是, 存在  $A$  到  $B$  的双射  $g_2$ 。这样, 我们就可以构造  $2^{\mathbb{N}}$  到区间  $(0, 1)$  的双射  $f_2$ :

$$f_2: x \mapsto \begin{cases} g_1(x) & \text{如果 } x \in 2^{\mathbb{N}} \setminus A \\ g_2(x) & \text{如果 } x \in A \end{cases}$$

把  $f_1$  和  $f_2$  复合, 我们就完成了  $\mathbb{R}$  和  $2^{\mathbb{N}}$  之间双射关系的构造。结论:

$$|\mathbb{R}| = |2^{\mathbb{N}}|.$$

□

证明二: 使用等势定理, 建立双向的单射。

**证明：** 首先建立  $2^{\mathbb{N}}$  到  $\mathbb{R}$  的单射。

给定数列  $\{a_n\} \in 2^{\mathbb{N}}$ ，考虑以下数列  $\{u_n\}$ ：

$$u_0 = a_0, \quad \forall n \in \mathbb{Z}^+, \quad u_n = u_{n-1} + \frac{a_n}{3^n}.$$

不难证明，对任意  $n > m$ ，

$$|u_n - u_m| \leq \sum_{i=m+1}^n \frac{1}{3^i} < \frac{3^{-(m+1)}}{1 - \frac{1}{3}} = \frac{3^{-m}}{2}.$$

因此  $\{u_n\}$  是自敛数列，收敛到某个实数。因此，可以定义映射：

$$\begin{aligned} f : 2^{\mathbb{N}} &\rightarrow \mathbb{R} \\ \{a_n\} &\mapsto \lim_{n \rightarrow +\infty} u_n \end{aligned}$$

只需证明  $f$  是单射。

设有两个不相同的数列  $\{a_n\}$ 、 $\{b_n\}$ ，分别对应数列  $\{u_n\}$ 、 $\{v_n\}$ 。不妨设  $m$  是  $\{a_n\}$ 、 $\{b_n\}$  第一个不同之处的下标。即  $\forall i < m, a_i = b_i$ ，且  $a_m$ 、 $b_m$  一个是 1，一个是 0。

不妨设  $a_m = 1$ ， $b_m = 0$ 。那么对任意  $n > m$ ，

$$\begin{aligned} u_n - v_n &= \frac{a_m - b_m}{3^m} + \sum_{i=m+1}^n \frac{a_i - b_i}{3^i} \\ &= \frac{1}{3^m} + \sum_{i=m+1}^n \frac{a_i - b_i}{3^i} \\ &\geq \frac{1}{3^m} + \sum_{i=m+1}^n \frac{0 - 1}{3^i} \\ &= \frac{1}{3^m} - \frac{3^{-m-1} - 3^{-n-1}}{1 - \frac{1}{3}} \\ &> \frac{1}{3^m} - \frac{1}{2 \cdot 3^m} = \frac{1}{2 \cdot 3^m} \end{aligned}$$

这说明  $f(\{a_n\}) - f(\{b_n\}) \geq \frac{1}{2 \cdot 3^m} > 0$ 。于是  $f$  是单射。

其次建立  $\mathbb{R}$  到  $2^{\mathbb{N}}$  的单射。为此，仍然先建立  $\mathbb{R}$  与  $(0, 1)$  之间的双射。具体方法见证明一。然后只需建立  $(0, 1)$  到  $2^{\mathbb{N}}$  的单射即可。

由于只需要建立单射，我们可以用比较简单的二分法。 $\forall x \in (0, 1)$ ，用以下操作构建数列  $\{a_n\}$ ：

取数轴上区间  $[0, 1]$  的中点 0.5，它把  $[0, 1]$  平分成两个区间。如果  $x < 0.5$ ，则  $a_0 = 0$ ，取左边的区间；否则  $a_0 = 1$ ，取右边的区间。然后在新的区间里重复上一步的操作，决定  $a_1$  的值。以此类推，得到数列  $\{a_n\} \in 2^{\mathbb{N}}$ 。

最后证明这样定义的映射是单射。给定  $0 < x < y < 1$ 。由于无穷等比数列  $\{2^{-n}\}$  趋于 0，所以总有正整数  $n$  使得  $y - x > \frac{1}{2^n}$ 。也就是说，第  $n$  次操作时， $x$ 、 $y$  会落到不同的区间，因此它们分别对应的数列的第  $n$  项不相等。这说明映射是单射。

既然存在  $2^{\mathbb{N}}$  到  $\mathbb{R}$  的单射，也存在  $\mathbb{R}$  到  $2^{\mathbb{N}}$  的单射，根据等势定理，两者等势。

□

## 附录 B 微变与函数

### 2.1 函数的微变

**定义 2.1.1.** 给定在某一点  $a$  附近有定义的函数  $f$ ，如果当  $x$  趋于  $a$  时， $f$  在  $a$  到  $x$  的变率收敛到某个极限，就说函数  $f$  在  $a$  处**可微**。我们把这个极限叫做函数  $f$  在点  $a$  处的**微变率**，简称**微变**，记作  $\partial f(a)$ 。

$$\partial f(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}.$$

历史上微变率的记法有很多种。常见的记法分为两派。

一派采取直接对函数名  $f$  添加上标、角标等方法。比如  $\dot{f}(a)$  或  $f'(a)$ 。前者在物理方面的文献中常见，后者则广泛见于各个领域。这类记法的好处是简洁，方便手写，缺点是难以推广到多元函数的多次微变，容易与其他函数操作（比如反函数）混淆。

另一派则使用独立的微变符号标记。比如  $\partial f(a)$  或  $\frac{df}{dx}(a)$ 。这类记法的本质是将“微变”作为映射、变换考虑，即它把一个函数映射到它的微变函数。使用独立的微变符号标记，好处是方便推广到更复杂的情况，比如多元函数的多次微变，缺点是手写不方便，且可能会造成难以理解的问题。比如  $\frac{d}{dx}$  记法就容易让初学者感到困惑。

本书默认采用  $\partial$  记号记录微变，但也可能会用到其他的标记方法。

**定理 2.1.1. 直指不等式** 设有正整数  $n > 1$ , 实数  $x > -1$  且  $x \neq 0$ , 则  $(1+x)^n > 1+nx$ .

**证明:** 用归纳法证明。 $n=1$  时有  $(1+x)^n = 1+nx$ 。假设对正整数  $n$  有  $(1+x)^n \geq 1+nx$ , 下面证明  $(1+x)^{n+1} > 1+(n+1)x$ 。

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n \cdot (1+x) \\ &\geq (1+nx)(1+x) \\ &= 1+(n+1)x+nx^2 \\ &> 1+(n+1)x.\end{aligned}$$

因此, 由  $n=1$  的情况可以推出  $n=2$  时  $(1+x)^2 > 1+2x$ 。此后对所有  $n > 2$ , 总有  $(1+x)^n > 1+nx$ 。  $\square$

**定理 2.1.2.** 对任意正整数  $k$ , 数列  $\{(1+\frac{k}{n})^n\}_{n \in \mathbb{Z}^+}$  收敛。

**证明:** 记数列  $\{u_{k,n}\}_{n \in \mathbb{Z}^+}$ 、 $\{v_{k,n}\}_{n \in \mathbb{Z}^+}$  的通项分别是:

$$\forall n \in \mathbb{Z}^+, \quad \begin{aligned} u_{k,n} &= \left(1 + \frac{k}{n}\right)^n, \\ v_{k,n} &= \left(1 + \frac{k}{n}\right)^{n+1}. \end{aligned}$$

只要证明  $\{u_{k,n}\}$  是单调递增数列,  $\{v_{k,n}\}$  是单调递减数列。那么  $\{u_{k,n}\}$ 、 $\{v_{k,n}\}$  都是有界数列, 因而都有极限, 分别记为  $u_k, v_k$ 。

又注意到对任意正整数  $n$ ,

$$v_{k,n} - u_{k,n} = \frac{k}{n} \cdot u_{k,n},$$

因此差数列  $\{v_{k,n} - u_{k,n}\}$  收敛到 0。所以  $u_k = v_k$ 。我们定义  $e = u_1 = v_1$ 。

下面证明  $\{u_{k,n}\}$  是单调递增数列。



对任意正整数  $n$ :

$$\begin{aligned}
 \frac{u_{k,n+1}}{u_{k,n}} &= \frac{\left(1 + \frac{k}{n+1}\right)^{n+1}}{\left(1 + \frac{k}{n}\right)^n} \\
 &= \frac{\left(1 + \frac{k}{n+1}\right)^{n+1}}{\left(1 + \frac{k}{n}\right)^{n+1}} \cdot \left(1 + \frac{k}{n}\right) \\
 &= \left(\frac{n(n+k+1)}{(n+1)(n+k)}\right)^{n+1} \cdot \left(1 + \frac{k}{n}\right) \\
 &= \left(1 + \frac{-k}{(n+1)(n+k)}\right)^{n+1} \cdot \left(1 + \frac{k}{n}\right) \\
 &> \left(1 + \frac{-k}{n+k}\right) \cdot \left(1 + \frac{k}{n}\right) \\
 &= 1
 \end{aligned}$$

其中的不等号根据直指不等式 (2.1.1) 可得。

注意：由于直指不等式对任意  $x > -1$  且  $x \neq 0$  成立，所以只要  $k$  是非零实数，对于足够大的  $n$ ，不等号总成立。也就是说，对非零实数  $k$ ，数列  $\{(1 + \frac{k}{n})^n\}_{n \in \mathbb{Z}^+}$  在  $n$  足够大的时候总是单调递增数列。

再证明  $\{v_{k,n}\}$  是单调递减数列。对任意正整数  $n$ ，考虑均值不等式：

$$\frac{(n+1) \cdot \frac{1}{n+1} + k \cdot \frac{1}{n}}{n+k+1} > \left( \left(\frac{1}{n+1}\right)^{n+1} \cdot \left(\frac{1}{n}\right)^k \right)^{\frac{1}{n+k+1}}$$

据此可以得到：

$$\left(1 + \frac{k}{n}\right)^{n+k+1} > \frac{(n+k+1)^{n+k+1}}{(n+1)^{n+1} n^k}.$$

因此，

$$\begin{aligned}
 \left(1 + \frac{k}{n}\right)^{n+1} &> \frac{(n+k+1)^{n+k+1} \cdot n^k}{(n+1)^{n+1} \cdot n^k \cdot (n+k)^k} \\
 &= \left(1 + \frac{k}{n+1}\right)^{n+2} \cdot \frac{(n+1)(n+k+1)^{k-1}}{(n+k)^k}
 \end{aligned}$$

只需证明  $\frac{(n+1)(n+k+1)^{k-1}}{(n+k)^k} > 1$ , 就得到  $v_{k,n} > v_{k,n+1}$ 。

而这个不等式等价于:

$$\left(1 + \frac{1}{n+1}\right)^{k-1} > \frac{n+k}{n+1} = 1 + \frac{k-1}{n+1}.$$

根据直指不等式, 上式成立。于是我们证明了  $\{v_{k,n}\}$  是单调递减数列。□

对非零有理数  $r = \frac{p}{q}$ , 注意到

$$\left(1 + \frac{r}{n}\right)^n = \left(1 + \frac{p}{qn}\right)^{\frac{nq}{q}} = (u_{p,qn})^{\frac{1}{q}}$$

于是数列  $\{u_{r,n}\} = \{(1 + \frac{r}{n})^n\}_{n \in \mathbb{Z}^+}$  其实就是数列  $\{u_{p,n}^{\frac{1}{q}}\}$  的子列, 因而收敛。

**定理 2.1.3.** 对任意有理数  $r$ , 数列  $\{(1 + \frac{r}{n})^n\}_{n \in \mathbb{Z}^+}$  收敛。

考虑关于  $x$  的函数  $f_n: x \mapsto (1 + \frac{x}{n})^n$ 。对足够大的正整数  $n$ ,  $1 + \frac{x}{n} > 0$ , 于是  $f_n$  对于  $x$  总是单调递增函数。所以只要  $a < b$ , 对足够大的正整数  $n$  就有  $u_{a,n} < u_{b,n}$ 。因此它们对应的数列的极限  $u_a$ 、 $u_b$  就有  $u_a \leq u_b$ 。也就是说, 定义在有理数上的函数:

$$h: r \mapsto \lim_{n \rightarrow +\infty} \left(1 + \frac{r}{n}\right)^n$$

单调递增。

对于函数  $h$ , 我们还知道什么呢? 首先容易验证:

$$h(0) = 1, \quad h(1) = e.$$

从此不难做出这样的猜想:

$$h(x) = e^x.$$

对正整数  $k$ , 考虑  $u_{k,kn}$ :

$$u_{k,kn} = \left(1 + \frac{k}{kn}\right)^{kn} = \left(1 + \frac{1}{n}\right)^{kn},$$

因此, 当  $n$  趋于无穷时, 数列  $\{u_{k,kn}\}$  趋于  $h(1)^k$ 。而它又是  $\{u_{k,n}\}$  的子列, 所以趋于  $h(k)$ 。这说明  $h(k) = e^k$ 。

考虑  $u_{1,kn}$ , 它等于  $u_{\frac{1}{k},n}^k$ , 所以  $\{u_{\frac{1}{k},n}\}$  趋于  $h(1)^{\frac{1}{k}}$ , 即  $h(\frac{1}{k}) = e^{\frac{1}{k}}$ 。

同理, 对于有理数  $r = \frac{p}{q}$ , 考虑  $u_{p,qn}$ , 它等于  $u_{\frac{p}{q},n}^q$ , 所以  $\{u_{\frac{p}{q},n}\}$  趋于  $h(p)^{\frac{1}{q}}$ , 也就是  $h(1)^{\frac{p}{q}}$ 。这说明  $h(r) = e^{\frac{p}{q}}$ 。

综上可知, 对任何有理数  $r$ ,  $h(r) = e^r$ 。

给定实数  $x$ , 选择任意有理数  $r^\downarrow < x < r^\uparrow$ , 则  $n$  足够大时,  $u_{r^\downarrow,n} < u_{x,n} < u_{r^\uparrow,n}$ 。因此, 数列  $\{u_{x,n}\}$  有界且单调递增, 因此收敛。

我们把它的极限定为  $h(x)$ , 就把  $h$  的定义域扩延到了实数集上。

使用第一册中探索指数函数性质的方法, 我们可以证明, 对任何实数  $x$ ,  $h(x) = e^x$ 。

接下来证明:

#### 定理 2.1.4. 直指等极

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1.$$

**证明:** 在  $\frac{e^x - 1}{x}$  中, 用数列  $u_{x,n}$  替换  $e^x$ , 计算两者差别:

$$d(x, n) = \left| \frac{e^x - 1}{x} - \frac{u_{x,n} - 1}{x} \right| = \frac{|e^x - u_{x,n}|}{|x|}.$$

对给定的  $x$ , 数列  $\{u_{x,n}\}$  趋于  $e^x$ , 所以对任意正数  $r$ , 总有正整数  $N$ , 使得只要  $n > N$ , 就有:

$$|e^x - u_{x,n}| < \frac{|x|r}{2},$$

这样,  $d(x, n) < \frac{r}{2}$ 。

再考虑  $\frac{u_{x,n}-1}{x} - 1$ :

$$\begin{aligned}
 \frac{u_{x,n}-1}{x} - 1 &= \frac{\left(1 + \frac{x}{n}\right)^n - 1 - x}{x} \\
 &= \frac{\sum_{k=0}^n C_n^k \left(\frac{x}{n}\right)^k - 1 - x}{x} \\
 &= \frac{\sum_{k=2}^n C_n^k \left(\frac{x}{n}\right)^k}{x} \\
 &= \sum_{k=1}^{n-1} \frac{C_n^{k+1}}{n^{k+1}} x^k \\
 &= \sum_{k=1}^{n-1} \frac{n!}{(k+1)!(n-k-1)!n^{k+1}} x^k
 \end{aligned}$$

因此,  $|x| < 1$  且  $n$  足够大时,

$$\begin{aligned}
 \left| \frac{u_{x,n}-1}{x} - 1 \right| &\leq \sum_{k=1}^{n-1} \left| \frac{n!}{(k+1)!(n-k-1)!n^{k+1}} \right| |x|^k \\
 &< \sum_{k=1}^{n-1} \frac{|x|^k}{(k+1)!} \\
 &< |x| \cdot \sum_{k=1}^{n-1} \frac{1}{2^k} \\
 &< |x|.
 \end{aligned}$$

因此, 首先选择  $|x| < \frac{r}{2}$  且  $|x| < 1$  的  $x$ , 再选择使得  $d(x, n) < \frac{r}{2}$  的  $n$ 。这样就有:

$$\begin{aligned}
 \left| \frac{e^x - 1}{x} - 1 \right| &\leq \left| \frac{e^x - 1}{x} - \frac{u_{x,n} - 1}{x} \right| + \left| \frac{u_{x,n} - 1}{x} - 1 \right| \\
 &< \frac{r}{2} + \frac{r}{2} = r.
 \end{aligned}$$

这就说明:

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1.$$

□

## 2.2 常见函数的微变

下面给出一些常见的显式函数的微变函数。

函数名称	表达式	微变函数
常数函数	$x \mapsto a$	$x \mapsto 0$
整式函数	$x \mapsto \sum_{i=0}^n a_i x^i$	$x \mapsto \sum_{i=1}^n a_i x^{i-1}$
幂函数	$x \mapsto x^a, a \neq 0$	$x \mapsto ax^{a-1}$
正弦函数	$x \mapsto \sin x$	$x \mapsto \cos x$
正切函数	$x \mapsto \tan x$	$x \mapsto \frac{1}{\cos^2 x}$
正割函数	$x \mapsto \sec x$	$x \mapsto \frac{\sin x}{\cos^2 x}$
余弦函数	$x \mapsto \cos x$	$x \mapsto -\sin x$
余切函数	$x \mapsto \cot x$	$x \mapsto -\frac{1}{\sin^2 x}$
余割函数	$x \mapsto \csc x$	$x \mapsto -\frac{\cos x}{\sin^2 x}$
指数函数	$x \mapsto a^x$	$x \mapsto a^x \ln a$
对数函数	$x \mapsto \log_a x$	$x \mapsto \frac{1}{x \ln a}$
自然指数函数	$x \mapsto e^x$	$x \mapsto e^x$
自然对数函数	$x \mapsto \ln x$	$x \mapsto \frac{1}{x}$

反正弦函数	$x \mapsto \arcsin x$	$x \mapsto \frac{1}{\sqrt{1-x^2}}$
反正切函数	$x \mapsto \arctan x$	$x \mapsto \frac{1}{1+x^2}$
反余弦函数	$x \mapsto \arccos x$	$x \mapsto -\frac{1}{\sqrt{1-x^2}}$
反余切函数	$x \mapsto \operatorname{arccot} x$	$x \mapsto -\frac{1}{1+x^2}$

## 2.3 研究可微函数

**定理 2.3.1. 微变展开定理** 设  $n$  为正整数。已知函数  $f$  在包含实数  $a$  的区间  $I$  中  $n+1$  次可微，那么，对于  $I$  中另一个数  $x$ ，存在介于  $a$  与  $x$  之间的数  $c$ ，使得下式成立：

$$\begin{aligned}
 f(x) &= f(a) + \sum_{k=1}^n \frac{\partial^k f(a)}{k!} (x-a)^k + \frac{\partial^{n+1} f(c)}{(n+1)!} (x-a)^{n+1}. \\
 &= P_n(x) + Y_n(x).
 \end{aligned}$$

上式称为  $f$  在  $a$  处的  **$n$  次微变展开**。其中  $P_n$  称为  $f$  在  $a$  处的  **$n$  次微变展开多项式**， $Y_n$  称为  $f$  在  $a$  处展开的  **$n$  次余项**。

**证明：** 首先证明一个引理：

函数  $f$  在  $[a, b]$  上  $n+1$  次可微。如果

$$f(a) = \partial f(a) = \partial^2 f(a) = \cdots = \partial^n f(a) = f(b) = 0,$$

那么有  $c \in (a, b)$  使得  $\partial^{n+1} f(c) = 0$ 。

引理的证明：多次使用微变零值定理。 $f(a) = f(b) = 0$ ，因此，存在  $c_1 \in (a, b)$  使得  $\partial f(c_1) = 0$ 。接下来，由于  $\partial f(a) = \partial f(c_1) = 0$ ，因此，存在  $c_2 \in (a, c_1) \subset (a, b)$  使得  $\partial^2 f(c_2) = 0$ 。依此不断，最终得到，存在  $c_{n+1} \in (a, b)$  使得  $\partial^{n+1} f(c_{n+1}) = 0$ 。 $c_{n+1}$  就是我们要找的  $c$ 。

下面使用这个结论来证明微变展开定理。

可以验证，给定  $x$ ， $f(x)$  和  $P_n(x)$  在  $a$  处的值相等，不仅如此，它们在  $a$  处的 1 到  $n$  次微变率也相等。但  $f(x) \neq P_n(x)$ 。因此，构造函数：

$$F : t \mapsto f(t) - P_n(t) - q(t-a)^{n+1}.$$

可以验证， $F(a) = \partial F(a) = \cdots = \partial^n F(a) = 0$ ， $F(x) = f(x) - P_n(x) - q(x-a)^{n+1}$ 。因此，取

$$q = \frac{f(x) - P_n(x)}{(x-a)^{n+1}},$$

这样，就有  $F(x) = 0$ 。于是，根据上面的引理，存在  $c \in (a, x)$ ，使得  $\partial^{n+1} F(c) = 0$ 。将  $\partial^{n+1} F(c)$  具体计算出来：

$$\partial^{n+1} F(c) = \partial^{n+1} f(c) - (n+1)!q.$$

因此有：

$$\frac{f(x) - P_n(x)}{(x-a)^{n+1}} = q = \frac{\partial^{n+1} f(c)}{(n+1)!},$$

也就是：

$$f(x) = P_n(x) + \frac{\partial^{n+1} f(c)}{(n+1)!}(x-a)^{n+1}.$$

这就是我们需要的结论。 □

可以看出，这个证明有模仿微变零值定理和中值定理的痕迹。首先证明微变零值定理的加强版本，然后构造函数，把一般情况下转为零值定理成立的特殊情况。

如果我们细看函数的  $n$  次微变展开，多项式  $P_n$  中的每一项，都对应  $f$  的某次微变。如果单看函数  $p_k : x \mapsto \frac{\partial^k f(a)}{k!}(x-a)^k$ ，它在  $a$  处的值和前

$k-1$  次微变率都是 0, 而  $k$  次微变率等于  $f$  在  $a$  处的  $k$  次微变率, 但  $k+1$  次及之后的微变率又等于 0 了。 $P_n$  的各项互不相干扰, 而加起来, 就模拟了  $f$  在  $a$  处的各次微变率。

局部展开定理其实比微变展开定理要稍微“弱”一点。因为它不要求具体指出余项的形式, 只需要刻画余项函数的性质。

**定理 2.3.2. 局部展开定理** 设  $n$  为正整数。函数  $f$  在区间  $I$  上  $n$  次可微, 且  $n$  次微变函数在  $I$  上连续。那么, 对于  $a \in I$ ,  $f$  可以写成以下形式:

$$\forall x \in I, \quad f(x) = f(a) + \sum_{k=1}^n \frac{\partial^k f(a)}{k!} (x-a)^k + (x-a)^n h(x).$$

其中  $h$  是一个在  $I$  上连续, 在  $a$  处趋于 0 的函数。

**证明:** 对  $n$  使用归纳法证明。

$n=1$  时, 我们需要证明:

$$f(x) = f(a) + \partial f(a)(x-a) + (x-a)h(x).$$

其中  $h$  是一个在  $I$  上连续, 在  $a$  处趋于 0 的函数。

这个结论我们已经在讨论函数的微直观的时候证明了。具体来说, 按照微变率的定义, 函数

$$h: x \mapsto \frac{f(x) - f(a)}{x-a} - \partial f(a)$$

在  $I \setminus \{a\}$  上连续, 且在  $a$  处趋于 0。于是, 补上  $h(a) = 0$  即可。

假设对于正整数  $n$ , 局部展开定理成立。对于  $n+1$  的情形, 已知条件为: 函数  $f$  在区间  $I$  上  $n+1$  次可微, 且  $n+1$  次微变函数  $\partial^{n+1}f$  在  $I$  上连续。

写出  $h$  的具体形式:

$$h(x) = \frac{f(x) - P_{n+1}(x)}{(x-a)^{n+1}}$$



我们要证明  $h$  在  $I$  上连续, 且在  $a$  处趋于 0。按照条件,  $h$  在  $I \setminus \{a\}$  上连续, 因此, 只需证明它在  $a$  处趋于 0。

使用微变展开定理, 我们可以把  $f(x)$  写成:

$$f(x) = P_n(x) + \frac{\partial^{n+1}f(c)}{(n+1)!}(x-a)^{n+1}.$$

其中  $c$  介于  $a$  和  $x$  之间。于是

$$h(x) = \frac{\partial^{n+1}f(c) - \partial^{n+1}f(a)}{(n+1)!}.$$

$\partial^{n+1}f$  在  $a$  处连续。因此, 对于任意  $r > 0$ , 都有  $d > 0$ , 使得只要  $|x-a| < d$ , 就有

$$|\partial^{n+1}f(x) - \partial^{n+1}f(a)| < r.$$

因此, 只要  $|x-a| < d$ , 那么  $|c-a| < d$ , 于是

$$|h(x)| = \left| \frac{\partial^{n+1}f(c) - \partial^{n+1}f(a)}{(n+1)!} \right| < r.$$

这说明  $h$  在  $a$  处趋于 0。

综上所述, 对任意正整数  $n$ , 结论都成立。这就证明了局部展开定理。

□

### 思考 2.3.1.

1. 只讨论一点  $a$  处的局部展开的话, 是否可以把局部展开定理的条件: “函数  $f$  在区间  $I$  上  $n$  次可微, 且  $n$  次微变函数在  $I$  上连续” 变成更宽松的条件?

**定理 2.3.3. 凸性与微变** 设函数  $f$  在区间  $I$  上可微。

- $f$  在  $I$  下凸, 当且仅当  $\partial f$  在  $I$  上单调递增。
- $f$  在  $I$  上凸, 当且仅当  $\partial f$  在  $I$  上单调递减。
- $f$  在  $I$  严格下凸, 当且仅当  $\partial f$  在  $I$  上严格单调递增。
- $f$  在  $I$  严格上凸, 当且仅当  $\partial f$  在  $I$  上严格单调递减。

**证明：** 首先，如果函数在  $I$  下凸，那么任取  $I$  内四点  $a < t_1 < t_2 < b$ ，则变率满足：

$$\frac{f(t_1) - f(a)}{t_1 - a} \leq \frac{f(t_2) - f(t_1)}{t_2 - t_1} \leq \frac{f(b) - f(t_2)}{b - t_2}.$$

令  $t_1$  趋于  $a$ ， $t_2$  趋于  $b$ ，就得到：

$$\partial f(a) \leq \partial f(b).$$

于是  $\partial f$  在  $I$  上单调递增。

反之，如果  $\partial f$  在  $I$  上单调递增，那么任取  $I$  内三点  $a < t < b$ ，根据微变中值定理定理，存在  $a < c_1 < t < c_2 < b$ ，使得：

$$\frac{f(t) - f(a)}{t - a} = \partial f(c_1) \leq \partial f(c_2) = \frac{f(b) - f(t)}{b - t}.$$

这说明函数在  $I$  下凸。

同理可证， $f$  在  $I$  上凸，当且仅当  $\partial f$  在  $I$  上单调递减。

如果函数在  $I$  严格下凸，我们希望证明  $\partial f$  在  $I$  上严格单调递增。由于小于号取极限后会变成小于等于，我们需要给小于关系加一道“防线”。比如，取  $I$  内四点  $a < s_1 = \frac{2a+b}{3} < s_2 = \frac{a+2b}{3} < b$ ，

$$\frac{f(s_1) - f(a)}{s_1 - a} < \frac{f(s_2) - f(s_1)}{s_2 - s_1} < \frac{f(b) - f(s_2)}{b - s_2}.$$

然后再取  $a < t_1 < s_1$ ，则

$$\frac{f(t_1) - f(a)}{t_1 - a} < \frac{f(s_1) - f(t_1)}{s_1 - t_1}.$$

上式变形可得到：

$$\frac{f(t_1) - f(a)}{t_1 - a} < \frac{f(s_1) - f(a)}{s_1 - a}.$$

因此令  $t_1$  趋于  $a$ ，就有：

$$\partial f(a) \leq \frac{f(s_1) - f(a)}{s_1 - a}.$$

同理, 取  $s_2 < t_2 < b$ , 可得到:

$$\frac{f(b) - f(s_2)}{b - s_2} < \frac{f(b) - f(t_1)}{b - t_1}.$$

因此令  $t_2$  趋于  $b$ , 就有:

$$\frac{f(b) - f(s_2)}{b - s_2} \leq \partial f(b).$$

因此:

$$\partial f(a) \leq \frac{f(s_1) - f(a)}{s_1 - a} < \frac{f(b) - f(s_2)}{b - s_2} \leq \partial f(b).$$

这说明  $\partial f$  在  $I$  上严格单调递增。

反之, 如果  $\partial f$  在  $I$  上严格单调递增, 那么任取  $I$  内三点  $a < t < b$ , 根据微变中值定理定理, 存在  $a < c_1 < t < c_2 < b$ , 使得:

$$\frac{f(t) - f(a)}{t - a} = \partial f(c_1) < \partial f(c_2) = \frac{f(b) - f(t)}{b - t}.$$

这说明函数在  $I$  严格下凸。

同理可证,  $f$  在  $I$  下凸, 当且仅当  $\partial f$  在  $I$  上严格单调递减。

□

## 2.4 常见函数的局部展开

下面给出一些常见显式函数在  $x = 0$  处的局部展开。

函数表达式	局部展开
$\sin x$	$x - \frac{x^3}{3!} + \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + o(x^{2n+1})$

$\cos x$	$1 - \frac{x^2}{2!} + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} + o(x^{2n})$
$e^x$	$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + o(x^n)$
$\ln(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{(-1)^{n-1} x^n}{n} + o(x^n)$
$\ln(1-x)$	$-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots - \frac{x^n}{n} + o(x^n)$
$\frac{1}{1-x}$	$1 + x + x^2 + \cdots + x^n + o(x^n)$
$\frac{1}{1+x}$	$1 - x + x^2 + \cdots + (-1)^n x^n + o(x^n)$
$\arctan x$	$x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{(-1)^n x^{2n+1}}{2n+1} + o(x^{2n+1})$
$(1+x)^a$	$1 + ax + \frac{a(a-1)}{2}x^2 + \cdots + \frac{a(a-1)\cdots(a-n+1)}{n!}x^n + o(x^{2n+1})$
$\tan x$	$x + \frac{x^3}{3} + \frac{2x^5}{15} + o(x^5)$

## 附录 C 平直空间

**定理 3.0.1.** 设有不全为零的一组向量  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  直相关。不妨设  $\mathbf{v}_1 \neq \mathbf{0}$ , 那么存在下标  $j \in \{2, 3, \dots, m\}$  满足以下性质:

1.  $\mathbf{v}_j \in \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1} \rangle$
2. 如果从  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  中移除  $\mathbf{v}_j$ , 剩余向量生成的空间与原来  $m$  个向量生成的空间相同。

**证明:** 按直相关的定义, 存在不全为零的标量系数  $a_1, a_2, \dots, a_m$  使得:

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_m \mathbf{v}_m = \mathbf{0}.$$

由于  $\mathbf{v}_1 \neq \mathbf{0}$ , 所以  $a_2, \dots, a_m$  不全为 0. 否则可推出  $a_1 \mathbf{v}_1 = \mathbf{0}$ , 从而  $a_1$  也是 0, 矛盾。

假设  $j$  是  $a_2, \dots, a_m$  中不为 0 的数的下标中最大的那个, 也就是说  $\forall k > j, a_k = 0$ . 这表示  $a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_j \mathbf{v}_j = \mathbf{0}$ . 所以

$$\mathbf{v}_j = -\frac{a_1}{a_j} \mathbf{v}_1 - \frac{a_2}{a_j} \mathbf{v}_2 - \dots - \frac{a_{j-1}}{a_j} \mathbf{v}_{j-1} \in \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1} \rangle.$$

由于  $\mathbf{v}_j$  可以表示为  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}$  的直组合, 设某个向量是  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  的直组合:

$$\mathbf{u} = b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \dots + b_m \mathbf{v}_m$$

把右侧表达式中的  $\mathbf{v}_j$  替换为  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}$  的直组合, 就得到了不含  $\mathbf{v}_j$  的直组合表达式。

所以, 去除  $\mathbf{v}_j$ , 整组向量的生成空间不变<sup>1</sup>。

□

**定理 3.0.2.** 有限生成空间中, 任何一族直无关的向量, 向量的个数都不多于任意一个生成族的向量个数。

**证明:** 设  $\mathfrak{F}_u = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$  是平直空间  $\mathbb{V}$  中一组直无关的向量,  $\mathfrak{F}_v = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m)$  是  $\mathbb{V}$  的生成族。我们要证明  $n \leq m$ 。

我们把  $\mathfrak{F}_v$  中的元素逐一替换成  $\mathfrak{F}_u$  中元素, 一共替换  $n$  次。如果以上步骤能完成, 说明  $\mathfrak{F}_v$  的长度不少于  $n$ , 就完成了证明。

第一步替换具体操作为:

将  $\mathbf{u}_1$  添加到  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m)$  左侧, 形成

$$\mathfrak{F}_v^{(1)} = (\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m).$$

由于  $\mathfrak{F}_v$  是  $\mathbb{V}$  的生成族,  $\mathbf{u}_1$  可以表示为  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  的直组合。即  $\mathfrak{F}_v^{(1)}$  是直相关的。所以根据定理 3.0.1, 我们可以去掉某个  $\mathbf{v}$ , 使得其生成空间不变 (仍然为  $\mathbb{V}$ )。这样  $\mathfrak{F}_v^{(1)}$  变为一组元素个数为  $m$  的向量, 并且仍然是  $\mathbb{V}$  的生成族。

此后第  $j$  步, 具体操作为:

将  $\mathbf{u}_j$  添加到  $\mathfrak{F}_v^{(j-1)}$  中  $\mathbf{u}_{j-1}$  的右侧。注意到  $\mathfrak{F}_v^{(j-1)}$  是由左侧的  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{j-1}$  和右侧的  $\mathfrak{F}_v$  中元素拼接而成, 所以  $\mathbf{u}_j$  被放置在  $\mathfrak{F}_v^{(j-1)}$  中原属于  $\mathfrak{F}_u$  的元素和原属于  $\mathfrak{F}_v$  的元素的分界上。放置之后形成的  $\mathfrak{F}_v^{(j)}$ , “布局” 和  $\mathfrak{F}_v^{(j-1)}$  一样, 仍然是由左侧原属于  $\mathfrak{F}$  的元素和右侧原属于  $\mathfrak{F}_v$  的元素构成。

这时, 通过和第一步同样的逻辑, 我们知道  $\mathfrak{F}_v^{(j)}$  是直相关的, 所以根据定理 3.0.1, 可以从中找到某个元素: 它属于它之前的元素的生成空间, 并且去除这个元素后,  $\mathfrak{F}_v^{(j)}$  生成的空间不变, 仍然是  $\mathbb{V}$ 。

<sup>1</sup>其实我们只证明了一个方向的归属关系, 不过另一个方向比较显然, 留给读者思考。

注意由于  $\mathfrak{F}_u$  是直无关的，我们找到的这个元素，肯定不会是左侧原属于  $\mathfrak{F}_u$  的元素，因而只能是原属于  $\mathfrak{F}_v$  的元素。

这样，每一步操作，我们都加入一个原属于  $\mathfrak{F}_u$  的元素，去除一个原属于  $\mathfrak{F}_v$  的元素。元素个数不变，且最后形成的向量组仍然是  $\mathbb{V}$  的生成集。

那么，是否会在某一步操作时，发现已经没有原属于  $\mathfrak{F}_v$  的元素可以去除呢？

如果这种情况发生，说明加入一个原属于  $\mathfrak{F}_u$  的元素后，只有原属于  $\mathfrak{F}_u$  的元素了。这时候向量组一方面仍然是直相关的，但同时元素全是  $\mathfrak{F}_u$  中元素。说明  $\mathfrak{F}_u$  也是直相关的。矛盾！

所以每一步操作都可以完整进行，直到  $\mathfrak{F}_u$  中所有元素放入完毕。这样，我们就证明了  $\mathfrak{F}_v$  的长度不少于  $n$ ，即  $n \leq m$ 。  $\square$

**定理 3.0.3.** 可以通过适当地移除元素，把有限生成空间  $\mathbb{V}$  的每个生成族变成基底。

**证明：** 设  $\mathcal{F}_v = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m)$  是  $\mathbb{V}$  的生成族。

如果  $\mathcal{F}_v$  直无关，则按照定义， $\mathcal{F}_v$  是一组基。

如果  $\mathcal{F}_v$  直相关。我们进行如下操作：

首先移除  $\mathcal{F}_v$  中所有的零向量。如果  $\mathcal{F}_v$  完全由零向量构成，说明  $\mathbb{V} = \{\mathbf{0}\}$ 。这时， $\mathcal{F}_v$  移除全部元素后剩下空集，直无关且生成  $\mathbb{V}$ 。

如果  $\mathcal{F}_v$  中还有若干非零向量，且仍直相关，根据定理 3.0.1，可以从其中找出一个向量，把它从  $\mathcal{F}_v$  移除后，生成空间不变， $\mathcal{F}_v$  仍旧是生成族。

如果  $\mathcal{F}_v$  仍然直相关，则重复以上操作。每次操作后， $\mathcal{F}_v$  中元素数量减少 1。所以，有限次操作之后， $\mathcal{F}_v$  变为直无关的生成族。这时我们就得到了空间的基底。

□

从上面的定理可以推出：

**定理 3.0.4.** 任意有限生成空间都有基底。

空间的任意生成集都可以缩减为它的基底。那么任意直无关的向量能否扩充为基底呢？

**定理 3.0.5.** 可以通过适当地添加元素，把有限生成空间  $\mathbb{V}$  的每一组直无关的向量变成基底。

**证明：** 设  $\mathcal{F}_v = (v_1, v_2, \dots, v_m)$  是  $\mathbb{V}$  中一组直无关的向量。

如果  $\mathcal{F}_v$  生成  $\mathbb{V}$ ，则按照定义， $\mathcal{F}_v$  是一组基。

如果  $\mathcal{F}_v$  生成的空间不是  $\mathbb{V}$ ，则是  $\mathbb{V}$  的真子空间。

空间  $\mathbb{V}$  是有限生成空间，所以存在生成族： $\mathcal{F}_u = (u_1, u_2, \dots, u_n)$ 。我们进行以下操作：

逐个查看  $(u_1, u_2, \dots, u_n)$ ，其中必然有一个元素不在  $\langle \mathcal{F}_v \rangle$  中。

这是因为，反设  $\mathcal{F}_u$  中每个元素都能写成  $\mathcal{F}_v$  中元素的直组合，那么由于  $\mathbb{V}$  中任一元素都能写成  $\mathcal{F}_u$  中元素的直组合，从而也能写成  $\mathcal{F}_v$  中元素的直组合。这说明  $\mathcal{F}_v$  可以生成整个空间，矛盾！

所以，必然存在某个  $u_i$  不在  $\langle \mathcal{F}_v \rangle$  中。我们将它添加到  $\mathcal{F}_v$  里去，得到的仍然是一组直无关的向量。

每次操作， $\mathcal{F}_v$  中增加一个元素。所以，要么在有限次操作后， $\langle \mathcal{F}_v \rangle = \mathbb{V}$ ，我们得到  $\mathbb{V}$  的直无关的生成集（也就是基底）；要么  $\mathcal{F}_v$  中元素个数超过  $n$ ，而这和定理 3.0.2 矛盾。

□



我们甚至有更具体的结论：

**定理 3.0.6.** 给定有限生成空间  $\mathbb{V}$  的子空间  $\mathbb{U}$ ，存在另一个子空间  $\mathbb{W}$ ，使得

$$\mathbb{V} = \mathbb{U} + \mathbb{W}.$$

且  $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$   $\mathbb{W}$  称为子空间  $\mathbb{U}$  的补空间，一般记作  $\mathbb{U}^c$ 。

**证明：** 设  $\mathcal{B}_1$  是子空间  $\mathbb{U}$  的基底。 $\mathcal{B}_1$  是直无关的。所以可以通过适当添加元素，将  $\mathcal{B}_1$  扩充为  $\mathbb{V}$  的基底  $\mathcal{B}$ 。

设  $\mathcal{B}_2$  是  $\mathcal{B}_1$  扩充到  $\mathcal{B}$  过程中添加的所有元素。 $\mathcal{B}_2$  也是直无关的。考虑  $\mathbb{W} = \langle \mathcal{B}_2 \rangle$ ，可以验证：

$$\mathbb{V} = \mathbb{U} + \mathbb{W}.$$

下面证明  $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$ 。

设  $\mathcal{B}_1 = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m)$ ， $\mathcal{B}_2 = (\mathbf{v}_{m+1}, \dots, \mathbf{v}_n)$ 。设  $\mathbf{u} \in \mathbb{U} \cap \mathbb{W}$ ，则：

$$\mathbf{u} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_m \mathbf{v}_m = a_{m+1} \mathbf{v}_{m+1} + \dots + a_n \mathbf{v}_n.$$

即：

$$a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_m \mathbf{v}_m + (-a_{m+1}) \mathbf{v}_{m+1} + \dots + (-a_n) \mathbf{v}_n = \mathbf{0}.$$

然而  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$  是直无关的。所以，

$$a_1 = a_2 = \dots = a_m = -a_{m+1} = \dots = -a_n = 0.$$

所以  $\mathbf{u} = \mathbf{0}$ 。因此， $\mathbb{U} \cap \mathbb{W} = \{\mathbf{0}\}$ 。

□

关于子空间交集与和空间维数的定理：

**定理 3.0.7. 维数基本定理** 设  $\mathbb{U}$  和  $\mathbb{W}$  是有限维空间  $\mathbb{V}$  的两个子空间，则它们的维数满足以下关系：

$$\dim(\mathbb{U} + \mathbb{W}) = \dim \mathbb{U} + \dim \mathbb{W} - \dim(\mathbb{U} \cap \mathbb{W}).$$

**证明：** 记  $\mathbb{T} = \mathbb{U} \cap \mathbb{W}$ . 根据定理 3.0.6, 我们可以给  $\mathbb{T}$  在  $\mathbb{U}$  和  $\mathbb{W}$  里各自找一个补空间:

$$\mathbb{U} = \mathbb{T} \oplus \mathbb{U}', \quad \mathbb{W} = \mathbb{T} \oplus \mathbb{W}'.$$

记  $\dim \mathbb{T} = h$ ,  $\dim \mathbb{U}' = f$ ,  $\dim \mathbb{W}' = g$ . 我们考虑这三个空间的基:

$$\mathcal{B}_T = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_h), \quad \mathcal{B}_{U'} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_f), \quad \mathcal{B}_{W'} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_g).$$

其中  $\mathcal{B}_{U'}$  和  $\mathcal{B}_{W'}$  分别是  $\mathbb{T}$  按照定理 3.0.6 证明中的方式扩充为  $\mathbb{U}$  和  $\mathbb{W}$  时得到的向量族。这表示,  $\mathcal{B}_T \cup \mathcal{B}_{U'}$  是  $\mathbb{U}$  的基,  $\mathcal{B}_T \cup \mathcal{B}_{W'}$  是  $\mathbb{W}$  的基。所以,

$$\dim \mathbb{U} = h + f, \quad \dim \mathbb{W} = h + g.$$

只需证明:  $\dim(\mathbb{U} + \mathbb{W}) = h + f + g$ .

为此我们证明  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  是  $\mathbb{U} + \mathbb{W}$  的一组基。

首先,  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  可以生成  $\mathbb{U}$ , 也可以生成  $\mathbb{W}$ , 所以可以生成  $\mathbb{U} + \mathbb{W}$ 。

其次, 证明  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  直无关。

假设零向量可以分解为  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  的直组合, 则说明零向量可以表示为:

$$\mathbf{0} = \mathbf{t} + \mathbf{u} + \mathbf{w}. \quad \mathbf{t} \in \mathbb{T}, \quad \mathbf{u} \in \mathbb{U}', \quad \mathbf{w} \in \mathbb{W}'.$$

所以  $\mathbf{u} = -\mathbf{t} - \mathbf{w} \in \mathbb{W}$ . 也就是说,  $\mathbf{u} \in \mathbb{U}' \cap \mathbb{W}$ 。

然而  $\mathbb{U}' \cap \mathbb{W} \subset \mathbb{U} \cap \mathbb{W} = \mathbb{T}$ , 所以  $\mathbf{u} \in \mathbb{T} \cap \mathbb{U}' = \{\mathbf{0}\}$ 。

因此,  $\mathbf{u} = \mathbf{0}$ .

同理可得,  $\mathbf{w} = \mathbf{0}$ . 所以  $\mathbf{t} = \mathbf{0}$ .

而  $\mathcal{B}_T$ 、 $\mathcal{B}_{U'}$  和  $\mathcal{B}_{W'}$  分别是直无关的。所以零向量分解为  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  的直组合时, 每个分量系数都是 0. 这就证明  $\mathcal{B}_T \cup \mathcal{B}_{U'} \cup \mathcal{B}_{W'}$  直无关, 从而是  $\mathbb{U} + \mathbb{W}$  的一组基。



从这个定理可以看出子空间和集合的相似之处。