第二次习题课讲义

于俊骜

2023 年 10 月 14 日

目录

1 复习回顾

数学的概念往往是抽象的，但并不意味着只能死记硬背。结合几何直观，将概念、结论与问题用形象化、通俗化的语言理解，便于我们更轻而易举地掌握知识，并从中发现数学之美。

1.1 连续函数

- 定义：$\forall\varepsilon>0$，$\exists\delta>0$，$|x-x_0|<\delta$ 时，$|f(x)-f(x_0)|<\varepsilon$
- 性质：函数和极限号可交换
- 理解：只要自变量离得够近，函数值一定不能差得太大
- 判定：左连续且右连续

1.2 等价替代

分析学的本质是用好的量逼近差的量。对于一般的函数，我们希望用更简单的函数去逼近它，从而更方便地去刻画其性质。比如后面要学的 Taylor 展开，就是用多项式逼近解析函数；B2 会接触到的 Fourier 级数，则是用三角函数逼近周期函数。

等价替代中，我们就希望用单项式 x^n 去描述一个复杂函数 $f(x)$ 趋于无穷大或无穷小的快慢，即 $f(x)$ 趋向于 0 或 ∞ 的速度和某一个 ax^n 相同。由于这里的“等价”是通过比值的极限定义的，在乘除运算中，可以直接进行替换。

$x\to0$ 时，我们有以下常用的等价无穷小替代：

- $\tan x\sim\sin x\sim x\sim\arcsin x\sim\arctan x$
- $1-\cos x\sim\frac{1}{2}x^2$
- $e^x-1\sim x\sim\ln(1+x)$

以及 $x\to+\infty$ 时常见的大小关系（设 $\alpha>\beta>0$）：

$$x^x\gg e^x\gg x^\alpha\gg x^\beta\gg\ln x$$

特别地，出现加减运算时，直接替换可能会出错。具体案例可见本讲义的 2.1。

另外，不少同学对 o 和 O 用法不够了解。事实上，设 $x \to x_0$ 时，$f(x)$ 和 $g(x)$ 都是无穷小量，则

1. $h(x) = o(1) \Longrightarrow h(x)$ 是无穷小量
2. $f(x) = o(g(x)) \Longrightarrow \frac{f(x)}{g(x)} = o(1)$，也即 $\frac{f(x)}{g(x)}$ 是无穷小量
3. $h(x) = O(1) \Longrightarrow h(x)$ 是有界量（可以是 0）
4. $f(x) = O(g(x)) \Longrightarrow \frac{f(x)}{g(x)} = O(1)$，即 $\frac{f(x)}{g(x)}$ 是有界量（可以是 0）

1.3 一致连续

“一致性”在第一次接触时，概念理解模糊是非常正常。

连续是局部性质，不过也可以推广到区间，即“$f(x)$ 在区间 I 连续 $\Longleftrightarrow f(x)$ 在 I 中每点连续”；一致连续是整体性质，只能说“$f(x)$ 在某个区间上是一致连续的”。

我们先从定义出发，类似第一次小测的格式，比较一下**区间上连续**和**区间上一致连续**：

c	$\forall\, \varepsilon > 0$	$\forall\, x_0$	$\exists\, \delta > 0$	$\lvert x - x_0\rvert < \delta \Rightarrow \lvert f(x) - f(x_0)\rvert < \varepsilon$
un.c.	$\forall\, \varepsilon > 0$	$\exists\, \delta > 0$	$\forall x_1, x_2$	$\lvert x_1 - x_2\rvert < \delta \Rightarrow \lvert f(x_1) - f(x_2)\rvert < \varepsilon$

不难发现，其中最关键的区别就是 δ 和 x 的前后顺序，即：“连续”中的 $\delta = \delta(\varepsilon, x_0)$ 后取，它不仅与 ε 有关，还与 x_0 有关；而“一致连续”中的 $\delta = \delta(\varepsilon)$ 先取，它只与 ε 有关。“一致”的含义即体现于此：这个连续性对于所有 $x \in I$ 是一致的，一个 δ 可以刻画 I 上所有 x 的连续性。

我们可以这样理解：取定 $\varepsilon > 0$，需要取的那个 δ 体现了这点的“连续程度”。如果 δ 可以取得很大，说明在一个更大的范围内，函数值相差不大，反之亦然。如果一个函数只是连续，那么它每个点的“连续程度”可以想多差糟糕有多糟糕。然而一旦一致连续，就要求整个区间上的点的“连续程度”都差不多，没有掉队的。

数学中的“一致”概念还很多，如**一致收敛**、**一致有界**、**一致可积**$\cdots$

以**一致收敛**为例，我们考虑一列函数 $\{f_n\}_{n=1}^{\infty}$，满足

$$f_n(x) = x^n,\ x \in I = [0,1)$$

不难注意到，$\forall \in I$，当 $n \to \infty$ 时，恒有 $f_n(x) \to 0$。但这个收敛对任意大的 n，都存在 $x \in (\frac{1}{\sqrt[n]{2}}, 1) \subset I$，使得 $f(x) > \frac{1}{2}$。因此，这个收敛对 I 上的所有 x 时不一致的，即“总有掉队的”。

2 作业选讲

2.1 习题 1.3.18

计算极限时，一旦某一项出现了两个极限相同的式子作差，例如

$$\lim_{x\to 0} \frac{\tan x - \sin x}{x^3}$$

则不能贸然等价替换。这是因为 $\sin x$ 和 $\tan x$ 趋于 0 的速度都与 x 差不多，但无法直接判断他们的差值 $\tan x - \sin x$ 趋于 0 的速度。

这就像下军棋，如果对手的子力只有“司令、师长、团长”，而你在此基础上多了 1 个“师长”，此时两边最大的子都是唯一的“司令”，你们的胜率差距一定程度上可以忽略；而万一两边“司令”对掉，此时你的最大棋子为 2 个“师长”，对方只有 1 个，这时候你的胜率就会明显大于对手。

在上述情形下，你们的子力都是“司令级”的，而子力差距是“师长级”的。正如上面那个极限式子，$\sin x$ 和 $\tan x$ 都是 x 级的，而它们的差值不是 0，而是 x^3 级的，所以最终结果是一个非零实数。

特别地，如果其中出现了根号，则可以利用平方差公式将其有理化，比如 (4)：

$$\begin{aligned}
\frac{\sqrt{2}-\sqrt{1+\cos x}}{\sin^2 x} &= \frac{(\sqrt{2}-\sqrt{1+\cos x})(\sqrt{2}+\sqrt{1+\cos x})}{\sin^2 x(\sqrt{2}+\sqrt{1+\cos x})} \\
&= \frac{1-\cos x}{\sin^2 x(\sqrt{2}+\sqrt{1+\cos x})} \\
&= \frac{2\sin^2 \frac{x}{2}}{\sin^2 x(\sqrt{2}+\sqrt{1+\cos x})} \\
&\sim \frac{1}{2(\sqrt{2}+\sqrt{1+\cos x})} \longrightarrow \frac{1}{4\sqrt{2}}
\end{aligned}$$

最后一步将 x 的极限值带入时，要当心计算错误，比如第 (5) 题有相当一部分同学将 $x \to 0$ 时 $1+\sqrt{x^2+x+1}$ 的极限值算成 1。

这种方法可以推广到开整数次根号，如第 (3) 题，可以利用恒等式

$$t^n - 1 = (t-1)(t^{n-1} + \cdots + 1)$$

并令 $1+\sin x = x^n$，则

$$\begin{aligned}
\frac{\sqrt[n]{1+\sin x}-1}{\arctan x} &= \frac{(1+\sin x)-1}{\arctan x((1+\sin x)^{\frac{n-1}{n}} + \cdots + 1)} \\
&\sim ((1+\sin x)^{\frac{n-1}{n}} + \cdots + 1)^{-1} \\
&\longrightarrow \frac{1}{n}
\end{aligned}$$

当然，该方法只是提供了一个可行的思路，未必是最简便的。

2.2 习题 2.1.4

(1)

直观上来看，判断函数是否连续，就是判断靠得近的点，函数值是不是差得不大。根据绝对值不等式 $||x|-|y|| \le |x-y|$，只要 $f(x)$ 的函数值差得不大，$|f(x)|$ 的函数值必然也离得很近。

证明. $\forall \varepsilon > 0, \exists \delta > 0$, 当 $|x-x_0| < \delta$ 时，有 $|f(x)-f(x_0)| < \varepsilon$。对上述 ε，注意到

$$||f(x)| - |f(x_0)|| \le |f(x) - f(x_0)| < \varepsilon$$

这说明 $|f(x)|$ 在 x_0 连续。 □

注 1. 这里也可以利用“连续函数的复合还是连续函数”，由于 $g(x)=|x|$ 连续、$f(x)$ 连续，可以推出 $g(f(x)) = |f(x)|$ 连续。

(2)

同一题的相邻两问，前一问往往给后一问提供现成结论或暗示。根据二元情形的最值函数的显式表达

$$\max\{x,y\} = \frac{x+y+|x-y|}{2} \qquad \max\{x,y\} = \frac{x+y-|x-y|}{2}$$

这里的第 (2) 问，就可以用第一问的结论一步出答案。

证明. 由 $f(x), g(x)$ 在 x_0 连续，根据 (1) 知

$$M(x)=\max\{f(x), g(x)\}=\frac{f(x)+g(x)+|f(x)-g(x)|}{2}$$

在 x_0 连续。同理，

$$m(x)=\min\{f(x), g(x)\}=\frac{f(x)+g(x)-|f(x)-g(x)|}{2}$$

在 x_0 连续。 □

注 2. 这里的 x_0 具有任意性，所以 $M(x)$ 和 $m(x)$ 都在整个定义域上连续。当我们需要证一个结论对于一个集合中的所有元素成立时，只需要取其中一个元素证明结论，并说明结论与选取无关。

如果对 $M(x)$ 和 $m(x)$ 的显式表达不熟悉，也可以发挥扎实的分析功底，直接分析函数局部性质。由于论证过程完全相同，只要证一个，另一个写“同理”。

另证：

证明. 只证 $M(x)$ 即可，$m(x)$ 同理。

①若 $f(x_0)=g(x_0)$，则 $\forall \varepsilon>0$，$\exists \delta_1>0, \delta_2>0$，使得 $|x-x_0|<\delta_1$ 时 $|f(x)-f(x_0)|<\varepsilon$；$|x-x_0|<\delta_2$ 时 $|g(x)-g(x_0)|<\varepsilon$。令 $\delta=\min\{\delta_1, \delta_2\}$，则

$$|M(x)-M(x_0)| \leq \max\{|f(x)-f(x_0)|,|g(x)-g(x_0)|\}<\varepsilon$$

故 $M(x)$ 在 x_0 连续。

②若 $f(x_0) \neq g(x_0)$，不妨设 $f(x_0)>g(x_0)$。对于 $\varepsilon=\frac{f(x_0)-g(x_0)}{2}>0$，$\exists \delta>0$，当 $|x-x_0|<\delta$ 时

$$\left.\begin{array}{r}|f(x)-f(x_0)|<\varepsilon \\ |g(x)-g(x_0)|<\varepsilon\end{array}\right\} \Longrightarrow f(x)-g(x)>(f(x_0)-\varepsilon)-(g(x_0)+\varepsilon)=0$$

即 $x \in(x_0-\delta, x_0+\delta)$ 时，恒有 $f(x)>g(x)$，于是

$$\lim_{x \rightarrow x_0} M(x)=\lim_{\substack{x \rightarrow x_0 \\ x_0-\delta<x<x_0+\delta}} M(x)=\lim_{\substack{x \rightarrow x_0 \\ x_0-\delta<x<x_0+\delta}} f(x)=f(x_0)=M(x_0)$$

故 $M(x)$ 在 x_0 连续。 □

注 3. 由于 $f(x)$ 和 $g(x)$ 地位相等，于是情形②可以“不妨设”。由 $f(x)-g(x)$ 的连续性，既然它在 x_0 取正值，就必然在 x_0 附近取正值。在这一个小区间上，$M(x)$ 就直接等于 $f(x)$ 了，大大简化。该想法体现了连续函数的本质，之后会多次出现在的证明中。从本题能看出来，“连续性”是一种很“刚性”的性质。

2.3 第十讲（一）(1)*

此题的背景是欧氏空间中的**范数**及其等价性。

范数是绝对值的推广，是用来评估向量长度的标尺。在欧氏空间

$$\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) | x_i \in \mathbb{R}, \forall i = 1, 2, \cdots, n\}$$

中，类似绝对值，我们可以定义范数：

定义 1 (欧氏空间中的范数). 若 $\mathbb{R}^n$ 到非负实数的函数 $\|\cdot\|$ 满足

1. 正定性：$\| x \| \geq 0$，等号成立当且仅当 $x = 0$
2. 齐次性：$\| \lambda x \| = |\lambda| \| x \|$，其中 λ 为常数
3. 三角不等式：$\| x \| + \| y \| \geq \| x + y \|$

则称它为一个范数

欧式空间的常用范数（以 2 维向量 $\vec{r} = (x, y) \in \mathbb{R}^2$ 为例）：

- 1 范数（直角距离）：$\| \vec{r} \|_1 = |x| + |y|$
- 2 范数（欧式距离）：$\| \vec{r} \|_2 = (|x|^2 + |y|^2)^{\frac{1}{2}}$
- p 范数（$p \geq 1$）：$\| \vec{r} \|_p = (|x|^p + |y|^p)^{\frac{1}{p}}$
- 无穷范数：$\| \vec{r} \|_\infty = \max\{|x|, |y|\}$

特别地，$p \to +\infty$ 是，p 范数的极限是无穷范数。

欧氏空间中的范数具有等价性，也就是说，对任意两个范数 $\|\cdot\|_{p_1}$ 和 $\|\cdot\|_{p_2}$，存在 $c_1, c_2 > 0$，使得

$$c_1 \| x \|_{p_1} \leq \| x \|_{p_2} \leq c_2 \| x \|_{p_1}, \ \forall x$$

该性质的一个直接推论就是同一个向量的不同范数，要么都有限，要么都无穷。

“向量”的概念很广，平面上的点组成 2 维欧式空间，那么这些点都可以是向量，可以对它们赋予以上的范数；$[0,1]$ 上的所有连续函数也可以组成线性空间 $C[0,1]$（这是一个无穷维空间），这些函数可以看作 $C[0,1]$ 中的向量，也可以给它们赋予范数，比如

$$\| f \|= \max_{x\in[0,1]} |f(x)|$$

矩阵也可以视为向量，它的最常用范数是最大特征值的模长，也就是用它乘一个向量，最多能把这个向量拉长多少倍。

扯了这么多不考的，现在回到原题。原式实际上就是一个 n 维欧氏空间 $\mathbb{R}$ 中的向量 $\vec{a}=(a_1,\cdots,a_n)$ 的 p 范数的极限，也就是无穷范数。显然，每一个分量都是有限数的有限维向量，到原点的欧式距离都有限，所以可以先验地判断这个式子的值不是 $+\infty$。事实上

$$\lim_{p\to+\infty}(|a_1|^p+\cdots+|a_n|^p)^{\frac{1}{p}}=\| \vec{a} \|_\infty=\max_n |a_n|$$

2.4 习题 2.2.8*

连续函数在闭区间上有界，无论这个闭区间长什么样，所以右端点可以取任意大。因为无穷远处极限存在，所以 x 很大时，$f(x)$ 只能在一个很小的范围内变化，必然有界。而当 x 没那么大时，就可以用连续函数在闭区间上有界了。将 $[a,+\infty)$ 分为“好+小”，两个区间上都有界，那么把两个区间并起来，肯定也有界。

证明. $\exists A\in\mathbb{R}$，$\forall\varepsilon>0$，$\exists b>a$，$x>b$ 时 $|f(x)-A|<\varepsilon$。

令 $\varepsilon=|A|+1$，再设 $|f(x)|<M,\forall x\in[a,b]$。则在 $[a,+\infty)$ 上恒有 $|f(x)|<\max\{M,2|A|+1\}$。 □

2.5 习题 2.2.13

函数在某点的极限与该点处的值毫无关系，所以本题的 a,b 完全不会出场。本题难度较大，需要对函数在边界附近的性质有着清晰的认识，才能判断其极限存在。不知道极限是多少的时候，Cauchy 列是一个很有效的工具。但 Cauchy 收敛准则只能应用在数列上，所以我们想要引入一个数列极限和函数极限关系的判别法，此即**Heine 归结原理**。

定理 1 (Heine 归结原理).

$$\lim_{x\to x_0} f(x) = A \Longleftrightarrow \lim_{n\to\infty} f(x_n) = A, \forall \{x_n\}\ s.t.\ x_n \longrightarrow x_0$$

应用该定理，我们可以给出以下正向证明。

证明. 只证 b 处左极限存在，a 处同理。

$\forall \varepsilon > 0$，$\exists \delta > 0$，$\forall x, y \in (a,b)$，只要 $|x-y| < \delta$，就有 $|f(x) - f(y)| < \varepsilon$。取定一列 $\{x_n\} \subset (a,b)$ 满足 $x_n \to b$。对上述 $\varepsilon > 0$，$\exists N > 0$，当 $n > N$ 时，恒有 $x_n \in (b-\delta, b)$。从而 $\forall m > n > N$，有 $|f(x_m) - f(x_n)| < \varepsilon$。这说明 $\{f(x_n)\}$ 是 Cauchy 列，从而收敛，有极限 B。

另一方面，任取一列 $\{y_n\} \subset (a,b)$ 满足 $y_n \to b$。$\forall \varepsilon > 0$，$\exists \delta > 0$，$\forall x, y \in (a,b)$，只要 $|x-y| < \delta$，就有 $|f(x) - f(y)| < \frac{\varepsilon}{2}$。同时，$\exists N_1 > 0$，$n > N_1$ 时，恒有 $|f(x_n) - B| < \frac{\varepsilon}{2}$；$\exists N_2 > 0$，$n > N_2$ 时，恒有 $y_n \in (b-\delta, b)$。因此 $n > \max\{N_1, N_2\}$ 时

$$|f(y_n) - B| \le |f(y_n) - f(x_n)| + |f(x_n) - B| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

说明 $f(y_n) \to B$。由 $\{y_n\}$ 的任意性知

$$\lim_{x\to b^-} f(x) = B$$

□

注 4. 先用一致连续找出一列趋向于边界的数列 $\{x_n\}$，使得 $f(x_n)$ 收敛，就可以设出极限 B，并证明它就是函数的极限。

3 专题：数学记号与解题步骤

3.1 常用缩写*

以下是一些可能在数学分析(B1)课程中遇到的缩写及其含义：

记号	thm	def	prop	i.e.	iff
原意	theorem	definition	proposition	id est	if and only if
含义	定理	定义	命题	（法）也即	当且仅当

记号	s.t	un.	C	d	i
原意	subject to/such that	uniformly	continuous	differential	index
含义	使得	一致	连续	微分	指标

记号	a.e.	sup	inf	e	log
原意	almost everywhere	supremum	infimum	exponent	logarithm
含义	几乎处处	上确界	下确界	指数	对数

记号	max	min	I	$P(x)$	L(R)HS
原意	maximum	minimum	interval	polynomial	left(right) hand side
含义	最大值	最小值	区间	多项式	左（右）式

接下来再补充一些数学分析中常用的希腊字母及其含义。

α/β	γ/Γ	δ/ε	ξ/ζ	θ	λ/μ	φ/ψ
alpha/beta	gamma	delta/epsilon	xi/zeta	theta	lambda/mu	phi/psi
指标/向量	曲线	极限定义	中值	余项	常系数	函数

3.2 解题规范

充要性 若题目要求证明“A 当且仅当 B”或“A 的充分必要条件是 B”，不建议写“充分性”“必要性”，而是用“$\Longrightarrow$”表示 A 推 B，用“$\Longleftarrow$”表示 B 推 A。这样写更便于理解。

另外，部分特殊的题目，证明的过程中每一步都可以是充要的，这时候只要推一个方向，并用“$\Longleftrightarrow$”从头连到尾。

任意性 若题目要求证明某性质对于整个区间成立，我们只需随便去一个 x，证明结论对这个 x 成立。最后补一句“由 x 的任意性，知结论对整个区间成立”就行。

先猜后证 部分题目过程较长，容易逻辑混乱，这是可以先提出结论，再对结论证明。常见的说法有：

1. 只需证$\cdots$
2. 先证一个引理/命题/结论：$\cdots$
3. 断言：$\cdots$

有时候，题目要求我们证结论 A，我们可以先断言结论 B 成立，然后用 B 把 A 给推出来，再去验证 B 的正确性。

3.3 简化步骤

高中时期，老师一再强调步骤的规范性。从作业中可以看出，很多同学步骤写得过繁琐。其实作业很多不必要的过程，完全可以省略。例如习题 2.1.7，解答完全可以一行写出：

$$1 = \lim_{x \to 0^-} e^x = \lim_{x \to 0^+} (a + x) = a \Longrightarrow a = 1$$

以下几个方面可以有效简化步骤，写着方便（助教批着也方便）。

省略不必要的说明 有些步骤一眼就能看出是为什么，就不需要文字说明。比如夹逼定理计算极限，若

$$0 < a_n < \frac{1}{n}$$

就不需要再赘述

$$\lim_{n \to \infty} 0 = 0 \qquad \lim_{n \to \infty} \frac{1}{n} = 0$$

而可以直接得到 $a_n \to 0$。

另外，我们可以把一些说明写在箭头或等号上。例如计算极限时，某一步需要用 Stolz 定理，就只需在那一步的等号上写一个 Stolz 之后接着算就行，不需要写“根据 Stolz 定理，可以得到$\cdots$”。

省略计算过程 计算过程中常规的四则运算与化简都是可以省略的。比如

$$\begin{aligned}
\lim_{x \to 0} \frac{\sqrt{1 + x + x^2} - 1}{\sin 2x} &= \lim_{x \to 0} \frac{(1 + x + x^2) - 1}{\sin 2x(\sqrt{1 + x + x^2} + 1)} \\
&= \lim_{x \to 0} \frac{x + x^2}{2x(\sqrt{1 + x + x^2} + 1)} \\
&= \lim_{x \to 0} \frac{1 + x}{2(\sqrt{1 + x + x^2} + 1)}
\end{aligned}$$

可以直接简写成

$$\lim_{x \to 0} \frac{\sqrt{1 + x + x^2} - 1}{\sin 2x} = \lim_{x \to 0} \frac{1 + x}{2(\sqrt{1 + x + x^2} + 1)}$$

使用箭头 中学阶段的推导基本被 $\because$ 和 $\therefore$ 给囊括了。但这种逻辑过于初等、繁琐，而用箭头推导则一目了然，如证明连续函数的连续函数次方仍连续时，这么写就行：

$$u(x), v(x) \in C(0, +\infty) \Longrightarrow u(x) \ln v(x) \in C(0, +\infty) \Longrightarrow u(x)^{v(x)} \in C(0, +\infty)$$

如果用 $\because\ \therefore$ 再加文字说明，没有个三行五行写不完。

适度化简 高考追求阅卷速率，往往要求“化到最简形式”。而你已经是个成熟的大学生了，可以化但没必要。比如 $\frac{1}{4\sqrt{2}}$ 不必化成 $\frac{\sqrt{2}}{8}$，前者反而更简洁；对于一个很麻烦函数的求导，结果不必通分，可以写成好几项的和；括号不是必须拆开，比如下面的式子就是一个合理的最终答案：

$$y' = \frac{(2x\ln x + \frac{1}{x} + x)(\sin x + \cos x) - (1+x^2)\ln x(\cos x - \sin x)}{(\sin x + \cos x)^2}$$

4 拓展：究竟什么是连续

4.1 开集与连续

第一次学到“连续”的概念时，会发现其定义比较繁琐。接下来，我们将用另一种方式定义连续函数。连续的原始定义中出现了两个开区间，即

$$(f(x_0)-\varepsilon, f(x_0)+\varepsilon) \qquad (x_0-\delta, x_0+\delta)$$

之后学多变量函数的连续性时，平面以及更高维的空间就没有“区间”这一概念了。事实上，多变量的连续性的定义中，开区间被推广成了“开球”(开区间是一维开球)。以二维为例，令 $\vec{r} = (x, y)$：

$$B_\delta(\vec{r}) = B_\delta(x_0, y_0) = \{(x,y) \in \mathbb{R}^2 | (x-x_0)^2 + (y-y_0)^2 < \delta^2\}$$

开区间的特点是，其中任何一个点，总能找到比它更左和更右的点。基于开区间，我们可以定义一类范围更广的集合：

定义 2 (开集). *若集合 O 满足：$\forall \vec{x} \in O \subset \mathbb{R}^n$，$\exists \varepsilon > 0$，使得 $B_\varepsilon(\vec{x}) \subset O$，则称 O 为开集。*

不难发现开集有两条重要性质:

- 任意多开集的并还是开集
- 两个开集的交还是开集

在开集的帮助下，我们可以给“连续函数”一个很精妙的等价定义：

定理 2 (连续函数的等价定义). *$f : D \subset \mathbb{R}^n \to \mathbb{R}$ 是连续函数 $\Leftrightarrow$ 任何开集在 f 下的原像是开集。*

证明. $\Longrightarrow$:

若 $f(x)$ 是连续函数，设 O 是 $f(x)$ 像集的任意一个开子集。

任取一个 $\vec{x}_0 \in f^{-1}(O)$，都有 $y_0 = f(\vec{x}_0) \in O$。结合开集的定义，对于 $y_0 \in O$，$\exists \varepsilon > 0$，使得 $(y_0 - \varepsilon, y_0 + \varepsilon) \subset O$。

接着，由 $f(x)$ 连续，知 $\exists \delta > 0$，只要 $\vec{x} \in B_\delta(\vec{x}_0)$，就有

$$f(\vec{x}) \in (y_0 - \varepsilon, y_0 + \varepsilon) \subset O$$

因此 $B_\delta(\vec{x}_0) \subset f^{-1}(O) \Rightarrow f^{-1}(O)$ 是开集。

$\Longleftarrow$:

若任意开集在 f 下的原像是开集，任取定义域中一点 $\vec{x}_0$。

$\forall \varepsilon > 0$，$f^{-1}(f(\vec{x}_0) - \varepsilon, f(\vec{x}_0) + \varepsilon)$ 是开集。

由 $\vec{x}_0 \in f^{-1}(f(\vec{x}_0) - \varepsilon, f(\vec{x}_0) + \varepsilon)$ 知，$\exists \delta > 0$，使得

$$B_\delta(\vec{x}_0) \subset f^{-1}(f(\vec{x}_0) - \varepsilon, f(\vec{x}_0) + \varepsilon)$$

因此只要 $|\vec{x} - \vec{x}_0| < \delta$，就有 $|f(x) - f(x_0)| < \varepsilon$，即 $f(x)$ 在 x_0 连续。 $\square$

这种定义方式的好处是，它不需要**度量结构**。也就是说，我们不需要考虑定义域或值域上的两个点的距离（体现为差的绝对值/模长）。事实上，有一些空间性质非常差，我们无法定义其上的“距离”，但可以定义什么是“开集”，从而也能定义连续函数。

事实上，开集的定义也可以推广。

定义 3 (拓扑). *设 X 是一个非空集合，$\mathcal{O}$ 是 X 的一些子集的集合，满足：*

1. *$X, \varnothing \in \mathcal{O}$*
2. *设 I 是一个指标集，$O_i \in \mathcal{O}, \forall i \in I$，则 $\bigcup_{i \in I} O_i \in \mathcal{O}$*
3. *$\forall O_1, O_2 \in \mathcal{O}$，有 $O_1 \cap O_2 \in \mathcal{O}$*

*则称 $\mathcal{O}$ 为 X 上的一个**拓扑**，$\mathcal{O}$ 中的元素称为该拓扑下的**开集**。*

拓扑有很多种，比如通常意义下的开集可以组成一个拓扑，X 的所有子集也可以组成一个拓扑。特别地，最小的拓扑（平凡拓扑）是 $\mathcal{O} = \{X, \varnothing\}$。

4.2 闭区间与紧性*

4.2.1 闭集

第二章中，我们学到了闭区间上连续函数的几大性质。但为什么偏偏是闭区间，它有什么特殊的？首先，我们类似地将闭区间推广为闭集：

定义 4 (闭集). *C 是闭集 $\Leftrightarrow$ C^c 是开集*

类似于开集，可以发现任意闭集的交仍是闭集、两个闭集的并仍是闭集。闭集有另一种等价描述：

定理 3 (闭集的等价定义). *C 是闭集 $\Leftrightarrow$ C 中任意收敛点列极限属于 C。*

证明. $\Longrightarrow$:

若 C 是闭集，则 C^c 是开集。

假设存在一个收敛点列 $\{x_n\} \subset C$，但其极限 $x \in C^c$，则存在 $\varepsilon_0 > 0$，使得 $B_{\varepsilon_0}(x) \subset C^c$。

此时 $|x_n - x| \geq \varepsilon_0, \forall n$，矛盾！

$\Longleftarrow$:

若 C 中任意收敛点列极限位于 C 中，假设 C 不是闭集，则 C^c 不是开集。

此时，$\exists x \in C^c$ 使得 $B_{\frac{1}{n}}(x) \not\subset C^c$，即 $B_{\frac{1}{n}}(x) \cap C \neq \varnothing, \forall n \in \mathbb{N}_0$。

于是可取 $x_n \in B_{\frac{1}{n}}(x) \cap C$，$\{x_n\}$ 收敛且极限是 x，与条件矛盾！ □

4.2.2 紧集

事实上，连续函数的这些性质来自于闭区间的**紧性**。

定义 5 (紧集). *若对于任何一族能覆盖 K 的开集，其中可以找到有限个，它们也能覆盖 K，则称 K 是一个***紧集***。*

顾名思义，紧集被紧紧框住，压得很“紧”。这个概念比较抽象，但我们有以下简单的描述：

定理 4 (Heine-Borel 定理（有限覆盖定理）). *有限维欧氏空间中，紧集与有界闭集等价。*

接下来，我们就可以研究连续函数在紧集上有什么性质。

定理 5. *设 $f(x)$ 是定义在紧集 K 上的连续函数，则*

1. $f(x)$ 有界且可以取到最值

2. $f(x)$ 在 K 上一致连续

证明. 我们将证明一个引理：$f(x)$ 将紧集映到紧集。

根据引理，$f(K)$，是 $\mathbb{R}$ 上的紧集，也就是有界闭集，所以 $f(x)$ 有界，且由闭集的等价定义，知可以 $f(K)$ 的最值存在。

下面只需证明引理。

事实上，设 $\{O_i\}_{i\in I}$ 是 $f(K)$ 的任意一个开覆盖，则 $\{f^{-1}(O_i)\}_{i\in I}$ 是 K 的一个开覆盖（开集的原像是开集）。

由 K 是紧集，知 K 存在有限子覆盖 $\{f^{-1}(O_i)\}_{i=1}^n$。

于是 $f(K)$ 存在有限子覆盖 $\{O_i\}_{i=1}^n$，即 $f(K)$ 是紧集。

引理得证!

最后证明一致连续。

任意固定 $\varepsilon > 0$，由 $f(x)$ 的连续性，$\forall x \in K$，$\exists \delta_x > 0$，只要 $y \in B_{\delta_x}(x)$，就有 $|f(y) - f(x)| < \varepsilon$。

注意到，K 有开覆盖

$$K \subset \bigcup_{x\in K} B_{\frac{\delta_x}{2}}(x)$$

由 K 的紧性，存在 $x_1, x_2, \cdots, x_m$，使得

$$K \subset \bigcup_{i=1}^{m} B_{\frac{\delta_{x_i}}{2}}(x_i)$$

令 $\delta_0 = \frac{1}{2}\min\{\delta_{x_1}, \delta_{x_2}, \cdots, \delta_{x_m}\} > 0$。

断言：对上述 $\varepsilon > 0$，存在 $\delta_0 > 0$，对于任意的 $x, y \in K$，只要 $|x - y| < \delta_0$，就有 $f(x) - f(y) < \varepsilon$。

事实上，不妨设 $x \in B_{\frac{\delta_{x_1}}{2}}(x_1)$，则

$$|y - x_1| \le |y - x| + |x - x_1| < \delta_0 + \frac{1}{2}\delta_{x_1} < \delta_{x_1}$$

于是 $x, y \in B_{\delta_{x_1}}(x_1)$，从而 $|f(x) - f(y)| < \varepsilon$。 □

从证明过程中可以看出，紧集的有限覆盖性保证了连续函数在其上的性质。而实数轴上的紧集，恰为闭区间或闭区间的有限并。这才促成了闭区间上的连续函数的性质。