泛函分析讲义

课堂笔记整理

编撰人： Clover

2026 年秋季学期

目录

第一章　度量空间速览 **3**

1.1 为什么要学泛函分析? 3

1.2 从 $\mathbb{R}^n$ 到抽象度量空间 5

1.3 序列的收敛、聚点与集合的闭包 10

1.4 连续映射 13

1.5 连续映射的进一步刻画 15

1.6 Cauchy 列与完备性 21

1.7 压缩映射与 Banach 不动点定理 27

1.8 紧性与列紧性 34

1.9 紧性的等价刻画 37

1.10 紧集上的连续映射 43

第二章　赋范空间和连续线性映射 **47**

2.1 Banach 空间 47

2.2 Banach 空间与级数 56

2.3 等价范数 61

2.4 有限维赋范空间 63

2.5 Riesz 引理与 Riesz 定理 66

2.6 连续线性映射（算子） 70

2.7 空间 $\mathcal{B}(E,F)$ 与对偶空间 E^* 78

2.8 延拓定理的应用；同构与 Neumann 级数 83

第三章　Hilbert 空间 **89**

3.1 内积空间 89

1

第一章

度量空间速览

1.1 为什么要学泛函分析?

三个问题

在有限维空间 $\mathbb{R}^n$ 里，很多直觉是”理所当然”的：有界闭集上的点列一定有收敛子列，线性映射一定连续，有限维子空间一定是闭集。本节用三个问题说明：这些直觉一旦搬到**无限维**的函数空间里，往往都会失效——这正是”为什么需要泛函分析”的出发点。

问题一：$C[0,1]$ 中有一致收敛的子列吗?

考虑 $C[0,1]$（$[0,1]$ 上全体连续函数构成的集合）中的函数列

$$f_n(x) = x^n, \qquad n = 1, 2, \dots.$$

问：能否从 $\{f_n\}$ 中选出一个**一致收敛**的子列?

答案是否定的。逐点来看，当 $x \in [0,1)$ 时 $f_n(x) \to 0$，而 $f_n(1) = 1$，所以 $\{f_n\}$ 的逐点极限是一个在 $x = 1$ 处跳跃的间断函数，它并不属于 $C[0,1]$。连续函数的一致极限仍然连续，于是 $\{f_n\}$ 的任何子列都不可能一致收敛（否则其一致极限会同时是这个间断的逐点极限，矛盾）。这与 $\mathbb{R}^n$ 中”有界闭集上任意点列都有收敛子列”（Bolzano–Weierstrass）的直觉完全不同：在函数空间里，”有界”不再自动导致这种紧性。

问题二：求导算子连续吗?

考虑**微分算子**

$$D : C^1[0,1] \longrightarrow C[0,1], \qquad Df = f',$$

两边都配备 $C[0,1]$ 上通常的度量（后面会看到就是 sup 范数）。问：D 是连续的吗？

答案是否定的。取 $f_n(x)=\dfrac{\sin(nx)}{n}$，则 $\sup_{x\in[0,1]}|f_n(x)|\le\dfrac{1}{n}\to 0$，即 $f_n\to 0$（一致收敛）；但 $f_n'(x)=\cos(nx)$ 根本不趋于 0。也就是说”输入趋于 0”不能推出”输出趋于 0”，D 不是连续（有界）算子。在有限维空间中，线性映射自动连续；这个例子说明无限维空间中的线性算子必须专门讨论**连续性**（或者说**有界性**）。

问题三：多项式全体是闭集吗？

多项式全体

$$P=\{\,[0,1]\text{ 上的多项式函数}\,\}\subset C[0,1]$$

显然是一个线性子空间（两个多项式的和、数乘仍是多项式）。问：P 是（sup 度量意义下的）闭集吗？

答案也是否定的。由 Weierstrass 逼近定理，P 在 $C[0,1]$ 中是**稠密**的——任何连续函数都可以被一列多项式一致逼近。但 $P\neq C[0,1]$（比如 e^x 或 $\sqrt{x}$ 不是多项式）。一个集合如果稠密却不等于全空间，就不可能是闭集。这说明：在无限维空间中，”有限维子空间是闭集”这条直觉，对无限维子空间并不成立。

泛函分析是什么

三个问题的共同点是：把 $\mathbb{R}^n$ 换成一个函数构成的**无限维**空间之后，线性代数、数学分析里许多”理所当然”的结论都需要重新证明，甚至根本不成立。粗略地说，

$$\text{泛函分析}\;=\;\text{无限维空间上的（线性代数}+\text{数学分析}+\text{几何）.}$$

”线性代数”提供了子空间、线性映射这些代数结构；”数学分析”提供了极限、收敛、连续这些分析结构；而当空间上还有”距离””长度””夹角”（几何结构）时，才能真正谈论收敛、连续、逼近这些分析概念。泛函分析要做的，就是在无限维空间上把这三者重新搭建起来。

研究对象与课程规划

研究对象

本课程研究的对象是**函数空间** E（E 中的元素本身是函数），以及定义在 E 上的两类基本映射：

- **泛函**：$J:E\to\mathbb{R}$，即把一个函数映成一个实数；
- **算子**：$T:E\to F$，其中 F 是另一个函数空间，即把一个函数映成另一个函数。

典型例子

- 能量泛函：$J(u)=\int_\Omega |\nabla u|^2\,dx$，把一个（定义在区域 Ω 上的）函数 u 映成一个实数，衡量 u 变化的”剧烈程度”；
- 微分算子：$T=\dfrac{d}{dx}$，把一个函数 u 映成另一个函数 u'，就是问题二里出现的那个算子。

这两个例子分别是”泛函”与”算子”最基本的原型，后面章节讨论的很多问题都是它们的推广或变体。

课程规划

要严格地讨论”泛函””算子”这些映射的连续性、有界性等分析性质，首先必须把它们的定义域和值域——函数空间 E——的结构说清楚。课程大致按下面的顺序展开：

(1) E 的结构：先是**线性**（向量空间）结构，再逐步配上 **距离** d、**范数** $\|\cdot\|$、或**内积** $\langle\cdot,\cdot\rangle$ 这些”能测量远近/长短/夹角”的附加结构，一层比一层丰富；

(2) 泛函：在配好结构的 E 上研究 $J:E\to\mathbb{R}$ 的性质，特别是连续性、有界性、极值等；

(3) 算子：在配好结构的 E,F 之间研究 $T:E\to F$ 的性质，尤其是（问题二里提出的）连续性/有界性问题；

(4) 横向比较：把**距离空间**、**赋范空间**、**内积空间** 等不同层次的结构放在一起比较，看清它们之间”谁包含谁””多了什么额外性质”。

下面就从最基础的”距离”结构开始。

1.2 从 $\mathbb{R}^n$ 到抽象度量空间

提纲

$\mathbb{R}^n$ 中两点 x,y 之间的距离是 $|x-y|$。上一节说过，泛函分析要讨论的是”距离””连续””收敛”这些分析概念在无限维函数空间中的类比。做法是：把 $\mathbb{R}^n$ 中具体的 $|x-y|$，换成一个抽象的函数 $d(x,y)$，只要求 d 具备 $|x-y|$ 所满足的那几条最基本的性质。$\mathbb{R}^n$ 中的点集拓扑（开集、闭集、收敛等概念）本质上只用到了这几条性质，所以一旦抽象出这几条公理，$\mathbb{R}^n$ 里关于开集、闭集、极限的全部结论就可以**原封不动地照搬**到任何满足这几条公理的抽象集合上——这正是本节要做的事。

度量的定义

设 X 是一个非空集合。如果存在一个函数

$$d:\ X\times X\longrightarrow \mathbb{R}_+=[0,\infty)$$

满足

(1) 正定性：$d(x,y)\geq 0$，且 $d(x,y)=0\iff x=y$，对一切 $x,y\in X$ 成立；

(2) 对称性：$d(x,y)=d(y,x)$，对一切 $x,y\in X$ 成立；

(3) 三角不等式：$d(x,y)+d(y,z)\geq d(x,z)$，对一切 $x,y,z\in X$ 成立，

则称 d 是 X 上的一个**度量**（或**距离**），称 (X,d) 为一个 **度量空间**（或**距离空间**）。

这三条公理其实就是把” 距离” 这个词日常意义上该有的样子写成了数学语言：距离不能是负的，而且” 距离为零” 才恰好对应” 是同一个点”（正定性）；从 x 到 y 的距离和从 y 到 x 的距离应该一样（对称性）；经过一个中间点 y 绕道而行，不会比直接从 x 走到 z 更近（三角不等式）。

常见的度量空间

基本例子

(1) $\mathbb{R}^n$ 配备通常的欧氏距离，是最基本的度量空间，本节讨论的所有抽象概念都以它为原型。

(2) 离散度量 在任意非空集合 X 上定义

$$d(x,y)=\begin{cases}1, & x\neq y,\\ 0, & x=y.\end{cases}$$

不难直接验证它满足度量的三条公理。这个例子看起来” 不自然”，但它是一个很方便的**构造反例**的工具：任何集合配上离散度量都能成为度量空间，可以用来检验某条一般性质是否真的对” 所有度量空间” 成立。

(3) 诱导度量 设 (X,d) 是度量空间，$A\subset X$ 是任意非空子集，将 d 限制在 $A\times A$ 上得到的 $d|_{A\times A}$ 仍然是 A 上的一个度量，称为 A 上由 X 诱导的度量。比如 $\mathbb{R}$ 中任意子集（区间、$\mathbb{Q}$ 等）都自动继承 $\mathbb{R}$ 上的度量。

(4) $C[0,1]$ **上的** sup **度量** 对 $f,g\in C[0,1]$，令

$$d_\infty(f,g)=\max_{x\in[0,1]}|f(x)-g(x)|.$$

（$[0,1]$ 是紧集，连续函数在其上必有最大值，所以这个 max 总有意义。）

(5) $L^p(\Omega)$ 上的度量 对 $f,g\in L^p(\Omega)$（$p\ge 1$），令

$$d_p(f,g)=\left(\int_\Omega |f-g|^p\,dx\right)^{1/p}.$$

三角不等式的证明：Minkowski 不等式

例子 (1)–(3) 的三条公理都可以直接验证。例子 (4)、(5) 中，正定性与对称性也是显然的；真正需要证明的是三角不等式，而它正是分析学中的 **Minkowski 不等式**：对 $p\ge 1$，

$$\left(\int_\Omega |f+g|^p\,dx\right)^{1/p}\le\left(\int_\Omega |f|^p\,dx\right)^{1/p}+\left(\int_\Omega |g|^p\,dx\right)^{1/p}.$$

把 f,g 换成 $f-h,h-g$ 即得 $d_p(f,g)\le d_p(f,h)+d_p(h,g)$。$d_\infty$ 的情形是它的极限（$p\to\infty$）版本，其实就是逐点绝对值不等式 $|f(x)-g(x)|\le|f(x)-h(x)|+|h(x)-g(x)|$ 两边对 x 取最大值。Minkowski 不等式的证明会在后续章节结合 Hölder 不等式给出，这里先记住结论、用它验证 (4)(5) 满足度量公理。

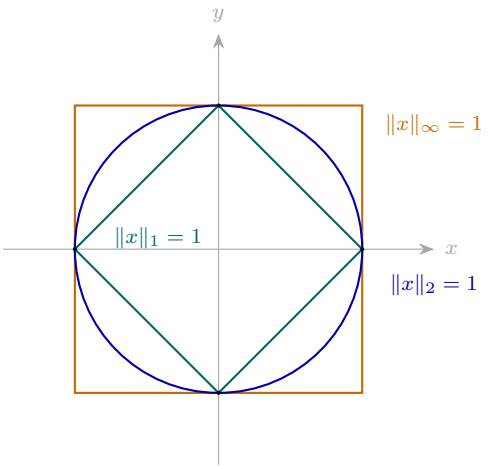
$\mathbb{R}^2$ 中三种度量的比较：d_1,d_2,d_∞

在 $\mathbb{R}^2$ 中，取 $x=(x_1,x_2)$，$y=(y_1,y_2)$，可以定义三种不同的度量：

$$d_1(x,y)=|x_1-y_1|+|x_2-y_2|,\qquad d_2(x,y)=\Big(\sum_{i=1}^{2}(x_i-y_i)^2\Big)^{1/2},$$

$$d_\infty(x,y)=\max_{1\le i\le 2}\big(|x_1-y_1|,\,|x_2-y_2|\big).$$

它们分别是例子 (4)（$p=\infty$）、(5)（$p=1,2$）在有限维、Ω 为两点集时的特殊情形。三者的”单位球” $\{x:d(x,0)=1\}$ 形状不同——d_1 是菱形，d_2 是圆，d_∞ 是正方形，菱形内切于圆、圆又内切于正方形：



从图上”菱形 $\subset$ 圆 $\subset$ 正方形”的包含关系，可以读出恒成立的不等式

$$d_\infty(x,0) \;\le\; d_2(x,0) \;\le\; d_1(x,0), \qquad \forall x \in \mathbb{R}^2.$$

更进一步还有 $d_1 \le \sqrt{2}\, d_2 \le 2\, d_\infty$，也就是说三种度量虽然数值不同，但可以互相”夹住”——这属于有限维空间中**度量（范数）等价**的现象：在 $\mathbb{R}^n$ 这样的有限维空间里，所有由范数导出的度量都彼此等价，诱导出**相同的**开集、闭集、收敛概念，下面讲拓扑时会看到这意味着什么。

开集与闭集

开球

设 (X,d) 是度量空间，$x \in X$，$r > 0$。称

$$B(x,r) = \{\, y \in X : d(x,y) < r \,\}$$

为以 x 为中心、r 为半径的**开球**。这是 $\mathbb{R}^n$ 中”以 x 为圆心、r 为半径的开圆盘/开球”这一概念在抽象度量空间中的直接类比，后面开集的定义完全依赖这个概念。

开集的定义

设 $A \subset X$。称 A 是 (X,d) 中的**开集**，如果

$$\forall x \in A,\ \exists r > 0,\ \text{s.t. } B(x,r) \subset A.$$

直观地说：A 中每一点都可以在 A 内部留出一点”活动空间”，不会紧贴在 A 的边界上。

开集的基本性质

(1) $\emptyset$ 与 X 都是开集（$\emptyset$ 是空真命题地满足定义；X 中任意一点取任意 $r>0$，$B(x,r) \subset X$ 显然成立）；

(2) 有限个开集的交仍是开集；

(3) 任意多个开集的并仍是开集；

(4) A 是开集，当且仅当 A 可以写成一族开球的并。

(2) 里的”有限”不能去掉。例如在 $\mathbb{R}$（通常度量）中，$A_n = \left(-\frac{1}{n}, \frac{1}{n}\right)$ 对每个 n 都是开集，但 $\bigcap_{n=1}^{\infty} A_n = \{0\}$，单点集 $\{0\}$ 不是开集（在 0 处放不下任何一个开球）。这说明”任意多个开集的交”未必开，(2) 特意限定为”有限个”。

闭集的定义

设 $B \subset X$。称 B 是 (X,d) 中的**闭集**，如果它的补集 $B^c = X \setminus B$ 是开集。

闭集是”用开集的语言定义出来的”，而不是独立的另一套概念；所以上面 (1)–(3) 中”开集”的性质，通过取补集，立刻可以翻译成”闭集”的对偶版本：$\emptyset, X$ 都是闭集，任意多个闭集的交仍是闭集，有限个闭集的并仍是闭集（并、交互换，正是德摩根律的体现）。

拓扑空间：抽象化的开集公理

定义

上面开集的三条性质 (1)(2)(3)——含 $\emptyset, X$、有限交封闭、任意并封闭——可以脱离”度量”单独提出来，作为一组公理。设 E 是一个集合，称 E 的子集族 τ（即 $\tau \subset \mathcal{P}(E)$）是 E 上的一个**拓扑**，如果 τ 满足以下条件：

(i) $E \in \tau$ 且 $\emptyset \in \tau$；

(ii) τ 中任意多个元素的并仍是 τ 中的元素；

(iii) τ 中有限个元素的交仍是 τ 中的元素。

此时称 τ 中的元素为**开集**，称 (E,τ) 为一个**拓扑空间**。

联系：度量诱导的拓扑

由上一小节的性质 (1)–(3) 可知：任给一个度量空间 (X,d)，把它所有开集收集起来组成的集族 τ_d，恰好满足拓扑空间的三条公理 (i)(ii)(iii)。也就是说，每一个度量空间 (X,d) 自动是一个拓扑空间 (X,τ_d)，这个 τ_d 称为由度量 d **诱导的拓扑**。

这正好解释了本节开头”提纲”里的那句话：$\mathbb{R}^n$ 中点集拓扑的结论之所以能”照搬”到一般度量空间，是因为拓扑空间的公理本来就是从”度量空间里开集应该具备的性质”中抽象出来的——拓扑空间是比度量空间更一般的框架，度量只是”生成”拓扑的一种（最常见的）手段。

反过来看，同一个集合 X 上可以定义很多形式不同的度量，只要它们诱导出同一个开集族 τ_d，就说这两个度量给出”相同的拓扑”（比如前面 $\mathbb{R}^2$ 上的 d_1, d_2, d_∞ 正是如此）。这就是后面会遇到的”度量等价”概念的拓扑本质：判断两个度量是否等价，归根结底是在问它们诱导出的开集族是否一致。

1.3 序列的收敛、聚点与集合的闭包

收敛序列

设 (X,d) 是一个度量空间。称序列 $\{x_n\}\subset X$ 是**收敛**的，如果

$$\exists\, x\in X,\quad \text{s.t.}\quad \lim_{n\to\infty} d(x_n,x)=0.$$

此时称 x 为 $\{x_n\}$ 的**极限**，也称 $\{x_n\}$ 收敛到 x，记作 $x_n \xrightarrow{d} x$。

这正是 $\mathbb{R}$ 中数列极限”当 $n\to\infty$ 时 $|x_n-x|\to 0$”这一定义，把绝对值换成抽象度量 d 之后在一般度量空间中的直接推广。

序列的聚点

称 $a\in X$ 是 (X,d) 中序列 $\{x_n\}$ 的一个**聚点**，如果存在子列 $\{x_{n_k}\}$ 收敛到 a。

收敛与聚点的关系

① 收敛序列只有唯一聚点。这是因为若 $x_n\to x$，则数学分析中的基本事实——收敛序列的任何子列都收敛到同一极限——在度量空间中依然成立，所以 x 是唯一的聚点。

② 只有唯一聚点的序列不一定收敛。例如取

$$x_n=\begin{cases} n, & n\text{ 为奇数},\\ \dfrac{1}{n}, & n\text{ 为偶数}. \end{cases}$$

偶数下标子列 $\dfrac{1}{2},\dfrac{1}{4},\dfrac{1}{6},\dots$ 收敛到 0；奇数下标子列 $1,3,5,\cdots\to\infty$，它以及它的任何子列都不可能收敛。于是 $\{x_n\}$ 唯一的聚点是 0，但整个序列 $\{x_n\}$ 并不收敛到 0（因为有无穷多项——奇数下标那些——离 0 很远，根本不有界）。这说明”聚点唯一”是比”收敛”更弱的条件。

接触点（粘着点）

设 (X,d) 是度量空间，$A\subset X$，$x\in X$。如果

$$\forall \varepsilon>0,\qquad B(x,\varepsilon)\cap A\neq\emptyset,$$

则称 x 为 A 的**接触点**（或**粘着点**）。

两个问题

① A 中的点一定是 A 的接触点吗？——**是**。因为 $x \in A$ 时，任意 $B(x,\varepsilon)$ 本身就含有 A 中的点 x。

② A 的接触点一定属于 A 吗？——**不是**。例如 $A=(a,b)\subset\mathbb{R}$ （$X=\mathbb{R}$），端点 a,b 都是 A 的接触点（任意以 a 或 b 为中心的球都会碰到 (a,b) 中的点），但 $a,b\notin A$。

集合的聚点（凝聚点）

设 (X,d) 是度量空间，$A\subset X$，$x\in X$。如果

$$\forall \varepsilon>0,\qquad \big(B(x,\varepsilon)\cap A\big)\setminus\{x\} \;=\; B(x,\varepsilon)\cap\big(A\setminus\{x\}\big)\;\neq\;\emptyset,$$

则称 x 为 A 的**聚点**（或**凝聚点**）。例如 $A=(a,b)\subset\mathbb{R}$ 中，端点 a,b 也是 A 的聚点。

注意，这里的”聚点”和上面”序列的聚点”用的是同一个名字，但含义不同——一个是相对于*序列*定义的（要求某个子列收敛到它），一个是相对于*集合*定义的（要求它附近总能找到 A 中*异于自身*的点）。二者的联系在于：x 是集合 A 的聚点，当且仅当存在 A 中互不相同的点组成的序列收敛到 x。读到”聚点”时要结合上下文判断说的是哪一种。

接触点与聚点的关系

- 聚点一定是接触点吗？——**是**：定义里 $B(x,\varepsilon)\cap(A\setminus\{x\})\neq\emptyset$ 显然蕴含 $B(x,\varepsilon)\cap A\neq\emptyset$。
- 接触点一定是聚点吗？——**不是**。例如 $A=[0,2]\cup\{3\}\subset\mathbb{R}$，点 3 是 A 的接触点（$3\in A$，由上面的问题①），但不是 A 的聚点：当 ε 充分小（比如 $\varepsilon<1$）时，$B(3,\varepsilon)$ 里除了 3 自己之外不再含有 A 中别的点，即 $B(3,\varepsilon)\cap(A\setminus\{3\})=\emptyset$。这种点称为 A 的**孤立点**。

闭包

设 (X,d) 是度量空间，$A\subset X$。A 的接触点的全体称为 A 的 **闭包**，记作 $\overline{A}$。

注：按定义取补集，$x\notin\overline{A}\iff\exists\, r>0,\ \text{s.t. } B(x,r)\cap A=\emptyset$。

闭包的基本定理

设 (X,d) 是度量空间，$A\subset X$，则以下结论成立：

(1) $\overline{A}$ 是一个闭集；

(2) $\overline{A}=\bigcap\{\,F\subset X\text{ 是闭集},\,F\supset A\,\}$，即 $\overline{A}$ 是 X 中包含 A 的**最小闭集**；

(3) A 是闭集 $\iff A=\overline{A} \iff A$ 包含了自身的所有接触点 $\iff \forall x\in(\overline{A})^c,\ \exists r>0,\ \text{s.t. } B(x,r)\subset(\overline{A})^c$。

证明

(1) $\overline{A}$ 是闭集。要证 $\overline{A}$ 是闭集，只需证 $(\overline{A})^c=X\setminus\overline{A}$ 是开集，即 $\forall x\in(\overline{A})^c,\ \exists r>0,\ \text{s.t. } B(x,r)\subset(\overline{A})^c$。

任取 $x\in(\overline{A})^c$，即 $x\notin\overline{A}$，由上面的”注”，存在 $r>0$ 使得 $B(x,r)\cap A=\emptyset$。对任意 $y\in B(x,r)$，令 $s=r-d(x,y)>0$，则 $B(y,s)\subset B(x,r)$（因为对任意 $z\in B(y,s)$，由三角不等式，$d(x,z)\le d(x,y)+d(y,z)<d(x,y)+s=r$）。于是

$$B(y,s)\cap A\subset B(x,r)\cap A=\emptyset \implies y\notin\overline{A} \implies y\in(\overline{A})^c.$$

由 $y\in B(x,r)$ 的任意性，$B(x,r)\subset(\overline{A})^c$。这说明 $(\overline{A})^c$ 是开集，从而 $\overline{A}$ 是闭集。

(2) $\overline{A}=\bigcap\{F\subset X$ 闭$, F\supset A\}$。先证 $\overline{A}\subset F$ 对任意闭集 $F\supset A$ 成立。设 F 是闭集且 $F\supset A$，任取 $x\in\overline{A}$，即 x 是 A 的接触点：$\forall\varepsilon>0$，$B(x,\varepsilon)\cap A\neq\emptyset$。由 $A\subset F$，自然 $B(x,\varepsilon)\cap F\neq\emptyset$，即 x 也是 F 的接触点。

下面说明 F 闭、x 是 F 的接触点，可以推出 $x\in F$（用反证法）：假设 $x\notin F$，则 $x\in F^c$，而 F^c 是开集，于是存在 $r>0$ 使得 $B(x,r)\subset F^c$，即 $B(x,r)\cap F=\emptyset$。这与”x 是 F 的接触点”（$B(x,r)\cap F\neq\emptyset$）矛盾，所以 $x\in F$。

因此 $\overline{A}\subset F$ 对每个满足 $F\supset A$ 的闭集 F 都成立，即 $\overline{A}\subset\bigcap\{F\subset X$ 闭$, F\supset A\}$。

反过来，由 (1)，$\overline{A}$ 本身就是一个包含 A 的闭集（A 中每一点显然都是自己的接触点，所以 $A\subset\overline{A}$），因此它本来就是右边这个交集所取的众多 F 之一，从而 $\bigcap\{F\subset X$ 闭$, F\supset A\}\subset\overline{A}$。两个包含关系合起来，就得到 $\overline{A}=\bigcap\{F\subset X$ 闭$, F\supset A\}$。

(3) 四条陈述互相等价。这一条要用到 (1)(2) 的结论。

先证 A 是闭集 $\iff A=\overline{A}$。

($\Longrightarrow$) 设 A 是闭集。A 本身就是” 包含 A 的闭集” 中的一个，由 (2) 的最小性，$\overline{A}\subset A$。又因为 A 中每一点显然都是自己的接触点，恒有 $A\subset\overline{A}$。两个包含关系合起来，$A=\overline{A}$。

($\Longleftarrow$) 设 $A=\overline{A}$。由 (1)，$\overline{A}$ 是闭集，从而 A 也是闭集。

再证 $A=\overline{A} \iff A$ 包含自身所有接触点。由闭包的定义，恒有 $A\subset\overline{A}$（A 中的点显然是自己的接触点），所以

$$A=\overline{A} \iff \overline{A}\subset A \iff A\text{ 的每一个接触点都属于 }A,$$

最后一句正是”A 包含自身所有接触点” 的意思。

最后看第四条陈述。”A 包含自身所有接触点” 即 $\overline{A}\subset A$，结合已经证得的 $A=\overline{A}$，此时 $(\overline{A})^c=A^c$。又由 (1)，$\overline{A}$ 恒是闭集，即 $(\overline{A})^c$ 恒是开集：

$$\forall x\in(\overline{A})^c,\ \exists r>0,\quad \text{s.t.}\quad B(x,r)\subset(\overline{A})^c,$$

这正是定理陈述的最后一条。把它和前面已证的 $A = \overline{A}$ 放在一起看，这一条其实就是把"A^c（$= (\overline{A})^c$）是开集"——也就是"A 闭" 本身的定义——重新展开了一遍，本质上不是一个独立于前三条的新事实，而是前三条串成一个逻辑闭环的最后一环：A 闭 $\Rightarrow$ $A = \overline{A} \Rightarrow \overline{A} \subset A \Rightarrow$（用 (1) 展开）第四条 $\Rightarrow$（因为此时 $(\overline{A})^c = A^c$）A^c 是开集 $\Rightarrow$ A 闭，转了一圈回到起点。

有界集

设 (X, d) 是度量空间。称 $A \subset X$ 是**有界**的，如果

$$\exists\, x_0 \in X,\ M > 0, \quad \text{s.t.} \quad A \subset B(x_0, M).$$

注：有界不是拓扑性质。取

$$\rho(x, y) = \frac{d(x, y)}{1 + d(x, y)},$$

可以验证 ρ 也是 X 上的一个度量，并且 ρ 与 d 有相同的收敛列（因为 $t \mapsto \dfrac{t}{1+t}$ 是 $[0, \infty)$ 到 $[0, 1)$ 上连续、严格递增的一一对应，逆映射 $s \mapsto \dfrac{s}{1-s}$ 在 0 处也连续，所以 $d(x_n, x) \to 0$ 当且仅当 $\rho(x_n, x) \to 0$）——也就是说 d 与 ρ 诱导出相同的拓扑。但 $\rho(x, y) < 1$ 对一切 x, y 恒成立，取 x_0 为任意一点、$M = 1$，就有 $X = B_\rho(x_0, 1)$，即 (X, ρ) 总是有界的。这说明" 有界" 这件事依赖于具体选哪个度量来衡量距离，同一个拓扑（同一族开集）可以由一个有界的度量给出，也可以由一个无界的度量给出，所以有界性不是像" 开集"" 收敛" 那样只依赖拓扑本身的性质。

注：闭球 $\overline{B}(x, r)$ 与开球 $B(x, r)$ 的闭包要区分。这里

$$\overline{B}(x, r) = \{\, y \in X : d(y, x) \le r \,\}$$

是把开球定义里的"$<$" 换成"$\le$" 直接给出的**闭球**，它总是一个闭集，但它不一定等于开球 $B(x, r)$ 的闭包（把 $B(x, r)$ 的所有接触点收集起来得到的集合）。例如取离散度量：$B(x, 1) = \{x\}$（因为 $d(x, y) < 1$ 迫使 $d(x, y) = 0$，即 $y = x$），而离散度量下任意集合都是闭集（其补集也总是可以写成一族单点开球的并，故为开集），所以 $B(x, 1) = \{x\}$ 的闭包就是 $\{x\}$ 本身；但 $\overline{B}(x, 1) = \{y : d(y, x) \le 1\}$ 却是整个 X（因为离散度量的值只有 0 或 1，恒 ≤ 1）。当 X 不止一个点时，$\{x\} \neq X$，两者并不相等。

1.4 连续映射

有了度量诱导的拓扑，就可以定义两个度量空间之间的**连续映射**了。设 (X, d_1) 和 (Y, d_2) 是两个度量空间，考虑映射 $f : X \to Y$。

一点处的连续性

称 f 在 $x \in X$ 点**连续**，是指

$$\begin{aligned}&\forall \varepsilon > 0,\ \exists \delta > 0, \quad \forall y \in X,\\ &\quad d_1(y,x) < \delta \ (\text{即} y \in B(x,\delta))\\ &\quad \Longrightarrow\ d_2\big(f(x),f(y)\big) < \varepsilon \ (\text{即} f(y) \in B(f(x),\varepsilon)).\end{aligned}$$

直观地说，x 附近（半径 δ 的球内）的点，经过 f 映射之后仍然落在 $f(x)$ 附近（半径 ε 的球内）；δ 的大小依赖于给定的 ε。这正是 $\mathbb{R}$ 上 ε-δ 语言在抽象度量空间里的直接翻版，把绝对值换成了各自空间上的度量 d_1, d_2。

等价刻画：极限语言

以上定义可以等价地刻画为

$$\lim_{d_1(y,x)\to 0} d_2\big(f(y),f(x)\big) = 0,$$

即当 $y \xrightarrow{d_1} x$ 时，有 $f(y) \xrightarrow{d_2} f(x)$。

整体连续性

称映射 f 在 X 上**连续**，若 f 在 X 中的每一点都连续。

连续的序列刻画

f 在 $x_0 \in X$ 连续 $\iff$ 对任意序列 $\{x_n\}$ 满足 $x_n \xrightarrow{d_1} x_0$，都有 $f(x_n) \xrightarrow{d_2} f(x_0)$。

证明。 ($\Longrightarrow$) 设 f 在 x_0 连续，$x_n \xrightarrow{d_1} x_0$。任给 $\varepsilon > 0$，由连续性存在 $\delta > 0$ 使得 $d_1(y,x_0) < \delta$ 就有 $d_2(f(y),f(x_0)) < \varepsilon$。又由 $x_n \to x_0$，存在 N 使得 $n \geq N$ 时 $d_1(x_n,x_0) < \delta$，从而 $d_2(f(x_n),f(x_0)) < \varepsilon$。这说明 $f(x_n) \xrightarrow{d_2} f(x_0)$。

($\Longleftarrow$) 用反证法。假设 f 在 x_0 不连续，即存在某个 $\varepsilon_0 > 0$，使得对每个 $\delta > 0$ 都能找到点 y 满足 $d_1(y,x_0) < \delta$ 但 $d_2(f(y),f(x_0)) \geq \varepsilon_0$。取 $\delta = \dfrac{1}{n}$ $(n = 1,2,\dots)$，依次找到点 x_n 满足 $d_1(x_n,x_0) < \dfrac{1}{n}$ 但 $d_2(f(x_n),f(x_0)) \geq \varepsilon_0$。于是 $x_n \xrightarrow{d_1} x_0$（因为 $d_1(x_n,x_0) < \frac{1}{n} \to 0$），但 $\{f(x_n)\}$ 不可能收敛到 $f(x_0)$（因为 $d_2(f(x_n),f(x_0)) \geq \varepsilon_0$ 对一切 n 成立，不趋于 0），这与假设”任意 $x_n \to x_0$ 都有 $f(x_n) \to f(x_0)$”矛盾。故 f 在 x_0 连续。

用开集刻画连续性

$$f:(X,d_1)\to(Y,d_2) \text{ 连续} \quad\Longleftrightarrow\quad Y \text{ 中任意开集的原像是 } X \text{ 中的开集},$$

即：$\forall O\subset Y$ 是 (Y,d_2) 中的开集，$f^{-1}(O)=\{x\in X: f(x)\in O\}$ 是 (X,d_1) 中的开集。

证明。($\Longrightarrow$) 设 f 连续，$O\subset Y$ 开集。任取 $x\in f^{-1}(O)$，即 $f(x)\in O$。因 O 开，存在 $\varepsilon>0$ 使得 $B(f(x),\varepsilon)\subset O$。由 f 在 x 处连续，对这个 ε 存在 $\delta>0$，使得 $d_1(y,x)<\delta\Rightarrow d_2(f(y),f(x))<\varepsilon$，即 $f\big(B(x,\delta)\big)\subset B(f(x),\varepsilon)\subset O$，从而 $B(x,\delta)\subset f^{-1}(O)$。这说明 $f^{-1}(O)$ 中每一点都有一个球包含在 $f^{-1}(O)$ 里，即 $f^{-1}(O)$ 是开集。

($\Longleftarrow$) 设 Y 中任意开集的原像都是 X 中的开集。任取 $x\in X$ 和 $\varepsilon>0$，$B(f(x),\varepsilon)$ 是 Y 中的开集，由假设 $f^{-1}\big(B(f(x),\varepsilon)\big)$ 是 X 中的开集，且显然 $x\in f^{-1}\big(B(f(x),\varepsilon)\big)$（因为 $f(x)\in B(f(x),\varepsilon)$）。由开集的定义，存在 $\delta>0$ 使得 $B(x,\delta)\subset f^{-1}\big(B(f(x),\varepsilon)\big)$，即对任意 y 满足 $d_1(y,x)<\delta$，有 $f(y)\in B(f(x),\varepsilon)$，即 $d_2(f(y),f(x))<\varepsilon$。这正是 f 在 x 处连续的定义。由 x 的任意性，f 在 X 上处处连续。

这个定理的意义在于：连续性原本是用度量（ε,δ）定义的一件”分析”的事情，但它完全可以用”开集的原像是开集”这句纯粹**拓扑**的话来刻画，不再直接提到 ε,δ 或者度量本身。这也是为什么”拓扑空间”上也能定义连续映射——只需要用这条开集刻画作为定义即可，不再需要度量。

1.5　连续映射的进一步刻画

连续的整体刻画

定理

设 $f:(X,d_1)\to(Y,d_2)$，则下列条件互相等价：

(i) f 在 X 上连续；

(ii) Y 中任意开集的原像是 X 中的开集；

(iii) Y 中任意闭集的原像是 X 中的闭集；

(iv) $\forall A\subset X$，$f(\overline{A})\subset\overline{f(A)}$。

证明

(i) $\Longleftrightarrow$ (ii) 就是上一讲已经证明过的”用开集刻画连续性”定理。下面依次证明 (ii) $\Longleftrightarrow$ (iii) 和 (iii) $\Longleftrightarrow$ (iv) 两步。

(ii) $\Longleftrightarrow$ (iii)。只需对开集、闭集的定义取补集，用到两个事实：$C\subset Y$ 闭 $\Longleftrightarrow$ $Y\setminus C$ 开；以及原像与补集可交换，$f^{-1}(Y\setminus C)=X\setminus f^{-1}(C)$。若 (ii) 成立，$C\subset Y$ 闭，则 $Y\setminus C$ 开，由 (ii)，$f^{-1}(Y\setminus C)=X\setminus f^{-1}(C)$ 是 X 中开集，从而 $f^{-1}(C)$ 是闭集，即 (iii) 成立；由 (iii) 推 (ii) 是完全对称的论证，把”开””闭”互换即可。

(iii)$\Rightarrow$(iv)。任取 $A\subset X$，令 $C=\overline{f(A)}$，它是 Y 中的闭集（闭包总是闭集）。因为 $f(A)\subset C$，所以 $A\subset f^{-1}(C)$。由 (iii)，$f^{-1}(C)$ 是 X 中的闭集；再由上一讲”闭包是包

含 A 的最小闭集”，得 $\overline{A}\subset f^{-1}(C)$。于是

$$f(\overline{A})\ \subset\ f\big(f^{-1}(C)\big)\ \subset\ C\ =\ \overline{f(A)},$$

即 (iv) 成立。

(iv)⇒(iii)。任取 Y 中闭集 C，令 $A=f^{-1}(C)$。要证 A 闭，即证 $A=\overline{A}$；恒有 $A\subset\overline{A}$，只需证 $\overline{A}\subset A$。由 (iv)，

$$f(\overline{A})\ \subset\ \overline{f(A)}\ =\ \overline{f\big(f^{-1}(C)\big)}\ \subset\ \overline{C}\ =\ C$$

（用到 $f(f^{-1}(C))\subset C$ 以及 C 闭）。所以对任意 $x\in\overline{A}$，都有 $f(x)\in C$，即 $x\in f^{-1}(C)=A$，说明 $\overline{A}\subset A$，故 A 闭，即 (iii) 成立。

综上，(i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (iv)，四条陈述互相等价。

注

① 连续是拓扑性质——由 (ii)(iii)，连续性完全可以用开集（或闭集）来描述，只依赖于 d_1,d_2 各自诱导出的拓扑，不需要直接用到度量的具体数值。

② 连续不保持开集、闭集——连续映射下，开集的像不一定是开集，闭集的像也不一定是闭集。要注意 (ii)(iii) 说的是*原像*，和这里说的”像”是两回事，不要混淆。

③ 可用连续映射分离闭集。

具体地，对于③，设 $A,B\subset X$ 是两个不相交的非空闭集。先回忆一个事实：点到集合的距离 $x\mapsto d(x,A):=\inf_{a\in A}d(x,a)$ 是 X 上的连续函数（事实上是 1-Lipschitz 的：对任意 $a\in A$，$d(x,a)\le d(x,y)+d(y,a)$，对 a 取下确界得 $d(x,A)\le d(x,y)+d(y,A)$，交换 x,y 再做一次同样的估计，就得到 $|d(x,A)-d(y,A)|\le d(x,y)$）。并且 $d(x,A)=0\iff x\in\overline{A}=A$（$A$ 闭）。

于是可以构造

$$f(x)=\frac{d(x,A)}{d(x,A)+d(x,B)}.$$

因为 $A\cap B=\emptyset$ 且两者都闭，分母 $d(x,A)+d(x,B)$ 对任意 x 都不为 0（否则 x 同时属于 $\overline{A}=A$ 和 $\overline{B}=B$，与不相交矛盾），所以 f 处处有定义。f 是两个连续函数的商（分母不为零），故连续；且 $f\equiv0$ 在 A 上（因为 $x\in A\Rightarrow d(x,A)=0$），$f\equiv1$ 在 B 上。这样就用一个连续映射把 A,B ”分开”了——A,B 分别落在 $f^{-1}(\{0\})$ 和 $f^{-1}(\{1\})$ 这两个不相交的开集（比如取 $f^{-1}((-\frac12,\frac12))$ 和 $f^{-1}((\frac12,\frac32))$）里。

一致连续

设 (X,d_1)，(Y,d_2) 是度量空间，$f:X\to Y$。称 f 在 X 上是 **一致连续**的，如果

$$\forall\varepsilon>0,\ \exists\delta=\delta(\varepsilon)>0,\quad \forall x,y\in X,\quad d_1(x,y)<\delta\ \Longrightarrow\ d_2(f(x),f(y))<\varepsilon.$$

把这个定义和上一讲” 一点处连续” 的定义（$\forall \varepsilon>0, \exists \delta=\delta(\varepsilon, x)>0, \ldots$）对比一下：两者形式几乎一样，唯一的区别是这里的 δ 只依赖于 ε，对 **整个** X 上所有点 x, y 统一有效，不能随点变化；而” 处处连续” 只要求在每一点各自能找到（可能依赖于该点的）δ 即可。所以一致连续 $\Rightarrow$ 处处连续，但反过来不成立（比如 $f(x)=x^2$ 在 $\mathbb{R}$ 上处处连续，但不一致连续：x 越大，同样的输出误差 ε 需要越小的 δ）。

命题：一致连续的序列刻画

f 一致连续 $\iff$ 对 X 中任意两个序列 $\{x_n\}, \{y_n\}$，只要 $d_1(x_n, y_n) \to 0$，就有 $d_2(f(x_n), f(y_n)) \to 0$。

证明： ($\Longrightarrow$) 由定义直接得到。设 f 一致连续，$d_1(x_n, y_n) \to 0$。任给 $\varepsilon>0$，由一致连续性存在 $\delta>0$（不依赖于点），使得 $d_1(x, y)<\delta \Rightarrow d_2(f(x), f(y))<\varepsilon$ 对一切 $x, y \in X$ 成立。又由 $d_1(x_n, y_n) \to 0$，存在 N 使得 $n \geq N$ 时 $d_1(x_n, y_n)<\delta$，从而 $d_2(f(x_n), f(y_n))<\varepsilon$。这说明 $d_2(f(x_n), f(y_n)) \to 0$。

($\Longleftarrow$) 用反证法。假设 f 不一致连续，即存在 $\varepsilon_0>0$，使得对每个 $\delta>0$，都能找到 $x, y \in X$ 满足 $d_1(x, y)<\delta$ 但 $d_2(f(x), f(y)) \geq \varepsilon_0$。取 $\delta=\dfrac{1}{n}(n=1,2, \ldots)$，依次找到 x_n, y_n 满足 $d_1(x_n, y_n)<\dfrac{1}{n}$ 但 $d_2(f(x_n), f(y_n)) \geq \varepsilon_0$。于是 $d_1(x_n, y_n) \to 0$，但 $\{d_2(f(x_n), f(y_n))\}$ 恒 $\geq \varepsilon_0$，不趋于 0，这与假设” 任意 $d_1(x_n, y_n) \to 0$ 都有 $d_2(f(x_n), f(y_n)) \to 0$” 矛盾。故 f 一致连续。

Lipschitz 连续

设 (X, d_1)，(Y, d_2) 是度量空间。称 $f: X \to Y$ 是 **M-Lipschitz 连续**的，如果

$$\exists\, M>0, \quad \text{s.t.} \quad d_2\big(f(x), f(y)\big) \leq M\, d_1(x, y), \qquad \forall x, y \in X.$$

注： *Lipschitz 连续 $\Rightarrow$ 一致连续，反之不对（例如 $f(x)=\sqrt{x}$）。*

证明： ($\Rightarrow$) 设 f 是 M-Lipschitz 的（不妨设 $M>0$）。任给 $\varepsilon>0$，取 $\delta=\varepsilon / M$，则 $d_1(x, y)<\delta$ 时 $d_2(f(x), f(y)) \leq M\, d_1(x, y)<M \delta=\varepsilon$，这就是一致连续（$\delta$ 的选取只依赖于 ε，与点无关）。

反例：$f(x)=\sqrt{x}$，$x \in[0, \infty)$，取通常的绝对值度量。先说明它一致连续但不 Lipschitz。

不妨设 $x \geq y \geq 0$，记 $a=\sqrt{x}$，$b=\sqrt{y}$ $(a \geq b \geq 0)$，则

$$|\sqrt{x}-\sqrt{y}|=a-b, \qquad \sqrt{|x-y|}=\sqrt{a^2-b^2}=\sqrt{(a-b)(a+b)}.$$

当 $a=b$ 时两边都是 0；当 $a>b$ 时，比较 $(a-b)^2$ 与 $(a-b)(a+b)$，等价于比较 $a-b$ 与 $a+b$，因为 $b \geq 0$ 显然 $a-b \leq a+b$，所以 $(a-b)^2 \leq(a-b)(a+b)$，两边开方即得

$$|\sqrt{x}-\sqrt{y}| \;\leq\; \sqrt{|x-y|}, \qquad \forall x, y \geq 0.$$

由此，任给 $\varepsilon>0$，取 $\delta=\varepsilon^2$，则 $|x-y|<\delta$ 时 $|\sqrt{x}-\sqrt{y}|\le\sqrt{|x-y|}<\sqrt{\delta}=\varepsilon$——这说明 $\sqrt{x}$ 在 $[0,\infty)$ 上一致连续。

但 $\sqrt{x}$ 不是 Lipschitz 的：若存在 M 使得 $|\sqrt{x}-\sqrt{0}|\le M|x-0|$ 对一切 $x>0$ 成立，即 $\sqrt{x}\le Mx$，即 $\dfrac{1}{\sqrt{x}}\le M$；但当 $x\to0^+$ 时 $\dfrac{1}{\sqrt{x}}\to\infty$，不可能被任何常数 M 控制，矛盾。这说明”一致连续”比”Lipschitz 连续”更弱。

等距映入与等距同构

设 (X,d_1)，(Y,d_2) 是度量空间。称 $f:X\to Y$ 为**等距映入**，如果

$$d_2\big(f(x),f(y)\big)=d_1(x,y),\qquad\forall x,y\in X.$$

若 f 还是满射，则称 f 为**等距同构**。

注：

① 等距映入 $\Rightarrow$ 一致连续，且是 1-Lipschitz 的——这是显然的，因为定义里的等号本身就给出 $d_2(f(x),f(y))=1\cdot d_1(x,y)$，取 $M=1$ 即满足 Lipschitz 的定义，再由上面证过的”Lipschitz $\Rightarrow$ 一致连续”即可。

② 等距映入 $\Rightarrow$ 单射——若 $f(x)=f(y)$，则 $d_1(x,y)=d_2(f(x),f(y))=0$，由度量的正定性 $x=y$。

映射复合的性质

设 $f:(X,d_1)\to(Y,d_2)$，$g:(Y,d_2)\to(Z,d_3)$。

1. 连续映射的复合仍连续；
2. 一致连续映射的复合仍一致连续；
3. 若 f 是 M_1-Lipschitz、g 是 M_2-Lipschitz 的，则 $g\circ f$ 是 M_1M_2-Lipschitz 的。

证明：1. 用开集刻画连续性最直接：对 Z 中任意开集 O，$(g\circ f)^{-1}(O)=f^{-1}\big(g^{-1}(O)\big)$。由 g 连续，$g^{-1}(O)$ 是 Y 中开集；再由 f 连续，$f^{-1}(g^{-1}(O))$ 是 X 中开集。所以 $(g\circ f)^{-1}(O)$ 总是开集，即 $g\circ f$ 连续。

2. 设 f,g 都一致连续。任给 $\varepsilon>0$，由 g 一致连续，存在 $\eta>0$ 使得 $d_2(y_1,y_2)<\eta\Rightarrow d_3(g(y_1),g(y_2))<\varepsilon$；再由 f 一致连续（对这个 η），存在 $\delta>0$ 使得 $d_1(x_1,x_2)<\delta\Rightarrow d_2(f(x_1),f(x_2))<\eta$。合起来，$d_1(x_1,x_2)<\delta\Rightarrow d_3\big(g(f(x_1)),g(f(x_2))\big)<\varepsilon$，且这里的 δ 只依赖于 ε（通过 η），与具体点无关，这正是 $g\circ f$ 一致连续。

3. 直接计算：

$$d_3\big(g(f(x)),g(f(y))\big)\le M_2\,d_2\big(f(x),f(y)\big)\le M_2M_1\,d_1(x,y),$$

所以 $g\circ f$ 是 M_1M_2-Lipschitz 的。

这三条都只是"⇒"，反过来一般不成立：复合映射有好的性质，不能反推出每一个"因子" 映射也有同样的性质。

拓扑等价与同胚

拓扑等价

设 X 是一个集合，d_1,d_2 是 X 上的两个度量。称 d_1,d_2（或它们各自诱导的拓扑）**拓扑等价**，如果它们诱导出*相同的开集族*。

等价地，我们称 d_1,d_2 **拓扑等价**，若对 X 中任意序列 $\{x_n\}$ 和任意点 $x\in X$，

$$x_n \xrightarrow{d_1} x \quad\Longleftrightarrow\quad x_n \xrightarrow{d_2} x,$$

即 $\{x_n\}$ 在 d_1 下收敛到 x 当且仅当它在 d_2 下收敛到 x。

定理

X 上 d_1,d_2 拓扑等价 $\Longleftrightarrow$ 恒同映射 $\mathrm{id}:(X,d_1)\to(X,d_2)$ 和它的逆 $\mathrm{id}^{-1}:(X,d_2)\to(X,d_1)$（也是恒同映射，只是把两个度量的角色互换）都连续。

下列条件都与"d_1,d_2 拓扑等价" 等价：

(i) d_1,d_2 诱导相同的收敛序列；

(ii) (X,d_1) 与 (X,d_2) 有相同的开集；

(iii) (X,d_1) 与 (X,d_2) 有相同的闭集；

(iv) $\forall x\in X,\ r>0,\ \exists s>0$，使得 $B_{d_2}(x,s)\subset B_{d_1}(x,r)$，并且把 d_1,d_2 互换后同样成立。

这几条其实都是"id 双向连续" 这件事的不同说法：(ii)(iii) 就是用开集/闭集刻画连续（上面刚证过的定理，应用在 $f=\mathrm{id}$ 上）；(iv) 是把" 开集" 这个抽象说法换成最原始的球的语言展开；(i) 则是连续的序列刻画的翻版。它们互相等价的证明方式和前面证连续的四种刻画完全一样，这里不重复。

同胚

设 (X,d_1)，(Y,d_2) 是（可以不同的）两个度量空间。称 $f:X\to Y$ 为 **同胚**，如果 f 是连续双射，且 f^{-1} 也连续。

注：

① 拓扑等价说的是*同一个集合* X 上两个度量 d_1,d_2 之间的关系；

② 同胚说的是（可能）*不同*的两个度量空间 (X,d_1) 与 (Y,d_2) 之间的关系；

③ 两者的联系：取 $Y=X$，$f=\mathrm{id}$，则 d_1 与 d_2 拓扑等价 $\iff$ $\mathrm{id}:(X,d_1)\to(X,d_2)$ 是同胚。也就是说，”拓扑等价”正是”同胚”这个更一般的概念，特殊化到”用恒同映射比较同一个集合上两个度量”的情形。

例

1. $f(x)=\dfrac{x}{1+|x|}$，$\mathbb{R}\to(-1,1)$，是同胚。
2. $\arctan:\mathbb{R}\to\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)$，是同胚。
3. 任意两个有界开区间都同胚（用线性映射构造：(a,b) 到 (c,d) 可取仿射映射 $x\mapsto c+\dfrac{d-c}{b-a}(x-a)$）。

先验证例 1：$f(x)=\dfrac{x}{1+|x|}$ 在 $x\ge 0$ 时化为 $\dfrac{x}{1+x}$，严格递增，从 0 增到（当 $x\to\infty$ 时）趋于 1；在 $x<0$ 时化为 $\dfrac{x}{1-x}$，同样严格递增，从（当 $x\to-\infty$ 时）趋于 -1 增到 0。两段在 $x=0$ 处平滑衔接（都取值 0），所以 f 在整个 $\mathbb{R}$ 上严格递增、连续，值域恰为 $(-1,1)$，从而是双射。可以解出反函数 $f^{-1}(y)=\dfrac{y}{1-|y|}$（$y\in(-1,1)$），同样是连续的有理函数（分母 $1-|y|$ 在 $(-1,1)$ 上恒不为零）。故 f 是同胚。

下面比较 $C[0,1]$ 上的两个度量。回忆 $d_\infty(f,g)=\max_{x\in[0,1]}|f(x)-g(x)|$，$d_1(f,g)=\int_0^1|f(x)-g(x)|\,dx$。因为

$$d_1(f,g)=\int_0^1|f(x)-g(x)|\,dx\ \le\ \int_0^1\max_{t\in[0,1]}|f(t)-g(t)|\,dx\ =\ d_\infty(f,g),$$

所以 $\mathrm{id}:(C[0,1],d_\infty)\to(C[0,1],d_1)$ 是（1-Lipschitz 从而）连续的：一致收敛（d_∞ 收敛）一定推出 d_1 收敛。但反过来不成立——存在在 d_1 意义下收敛、但不一致收敛的函数列（比如底边越来越窄、高度不变的”移动小山包”函数列，d_1 趋于 0 但 d_∞ 恒为常数）。这说明 $\mathrm{id}:(C[0,1],d_1)\to(C[0,1],d_\infty)$ 不连续，从而两个度量诱导出的拓扑 T_∞,T_1 不是拓扑等价的：T_∞ 比 T_1 ”更细/更强”（d_1-开集一定是 d_∞-开集，但反过来不然），记作 $T_1\subsetneq T_\infty$。

度量等价

设 X 是一个集合，d_1,d_2 是 X 上的两个度量。称 d_1,d_2 为 **等价度量**，如果存在常数 $c_1,c_2>0$，使得

$$c_1\,d_1(x,y)\ \le\ d_2(x,y)\ \le\ c_2\,d_1(x,y),\qquad \forall x,y\in X.$$

命题

X 上两个度量 d_1,d_2 度量等价 $\iff$ $\mathrm{id}:(X,d_1)\to(X,d_2)$ 与 $\mathrm{id}^{-1}:(X,d_2)\to(X,d_1)$ 都是 Lipschitz 的（Lipschitz 常数分别是 c_2 和 $\dfrac{1}{c_1}$）。

证明：若 $c_1 d_1 \le d_2 \le c_2 d_1$，则右边不等式 $d_2(x,y) \le c_2\, d_1(x,y)$ 正是说 $\mathrm{id} : (X,d_1) \to (X,d_2)$ 是 c_2-Lipschitz 的；左边不等式改写为 $d_1(x,y) \le \dfrac{1}{c_1} d_2(x,y)$，正是说 $\mathrm{id}^{-1} : (X,d_2) \to (X,d_1)$ 是 $\dfrac{1}{c_1}$-Lipschitz 的。反过来，若两个方向都是 Lipschitz 的（常数分别记为 c_2，$\dfrac{1}{c_1}$），逆推回去就得到两个不等式，即度量等价。

注：度量等价 $\Rightarrow$ 拓扑等价（*Lipschitz* 蕴含连续），但反之不然。

下面两个例子说明反之不然。

- d 与 $\rho = \dfrac{d}{1+d}$（上一讲”有界集”一节提到过）拓扑等价——诱导相同的收敛列。但当 X 在 d 下无界时，两者不是度量等价的：若存在 c_1 使 $c_1 d(x,y) \le \rho(x,y) = \dfrac{d(x,y)}{1+d(x,y)}$ 对一切 x, y 成立，取 $d(x,y) \to \infty$，右边趋于 1 而左边趋于 ∞（只要 $c_1 > 0$），不等式必然失效。
- $|\cdot|$（$\mathbb{R}$ 上通常的度量）与 arctan 型度量 $d_{\arctan}(x,y) = |\arctan x - \arctan y|$ 也是类似的情形：arctan 把整个 $\mathbb{R}$ ”压缩”进有界区间 $(-\pi/2, \pi/2)$，$d_{\arctan}$ 有界而 $|\cdot|$ 无界，同样不可能被一个常数上下夹住。

这说明”度量等价”是比”拓扑等价”更强的条件：它不仅要求两个度量有相同的收敛列，还要求它们对”距离大小”的度量方式在常数倍的意义下彼此可控。

1.6 Cauchy 列与完备性

Cauchy 列

设 (X,d) 是度量空间。称序列 $\{x_n\} \subset X$ 是 **Cauchy 列**，如果

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\quad \text{s.t.}\quad \forall m, n > N,\quad d(x_m, x_n) < \varepsilon.$$

直观地说，Cauchy 列是”自己跟自己越来越近”的序列——不需要预先知道极限是谁，只要求序列的任意两项（下标都足够大时）彼此靠近。这一点和”收敛”的定义不同：收敛需要预先给出一个候选极限 x，而 Cauchy 列的定义完全不提到 X 之外或之内的任何具体点。

命题

设 $\{x_n\}$ 是 (X,d) 中的序列，则：

① 收敛序列是 Cauchy 列；

② Cauchy 列不一定收敛；

③ 若 Cauchy 列有一个收敛子列（即有聚点），则它本身收敛到同一极限；

④ Cauchy 列有界；

⑤ 一致连续映射把 Cauchy 列映为 Cauchy 列，但连续映射不必。

下面逐条给出证明或例子。

①设 $x_n \to x$。任给 $\varepsilon > 0$，存在 N，使 $n > N$ 时 $d(x_n, x) < \varepsilon/2$。于是对 $m, n > N$，由三角不等式，

$$d(x_m, x_n) \le d(x_m, x) + d(x, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

即 $\{x_n\}$ 是 Cauchy 列。

②反例：在 $\mathbb{Q}$（有理数集，配通常的绝对值度量）中取一列有理数 $x_n \to \sqrt{2}$（比如 $\sqrt{2}$ 的小数截断：$1,\ 1.4,\ 1.41,\ 1.414, \dots$）。这列数在 $\mathbb{R}$ 中收敛（到 $\sqrt{2}$），由①它在 $\mathbb{R}$ 中是 Cauchy 列，而 Cauchy 列的定义只涉及数列自身两两之间的距离，与” 放在哪个空间里看” 无关，所以它在 $\mathbb{Q}$ 中同样是 Cauchy 列。但它在 $\mathbb{Q}$ 中**不收敛**——因为唯一可能的极限 $\sqrt{2} \notin \mathbb{Q}$。这说明” 是 Cauchy 列” 不能推出” 收敛”，除非所在的空间本身足够” 完整”（这正是下面要引入的” 完备性”）。

③设 $\{x_n\}$ 是 Cauchy 列，子列 $\{x_{n_k}\}$ 收敛到 a。任给 $\varepsilon > 0$：由 Cauchy 性，存在 N，使 $m, n > N$ 时 $d(x_m, x_n) < \varepsilon/2$；由子列收敛，存在 k_0，使 $n_{k_0} > N$ 且 $d(x_{n_{k_0}}, a) < \varepsilon/2$。于是对任意 $n > N$，

$$d(x_n, a) \le d(x_n, x_{n_{k_0}}) + d(x_{n_{k_0}}, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

（这里用到 $n, n_{k_0} > N$，故 $d(x_n, x_{n_{k_0}}) < \varepsilon/2$ 由 Cauchy 性给出）。所以 $x_n \to a$。

④取 $\varepsilon = 1$，由 Cauchy 性存在 N，使 $m, n > N$ 时 $d(x_m, x_n) < 1$。固定 $n_0 = N+1$，则对一切 $n > N$，$d(x_n, x_{n_0}) < 1$，即 $\{x_n : n > N\} \subset B(x_{n_0}, 1)$。前面有限项 $x_1, \dots, x_N$ 是有限集，取

$$M = \max\Big(1,\ \max_{1 \le k \le N} d(x_k, x_{n_0})\Big) + 1,$$

则整个序列 $\{x_n\} \subset B(x_{n_0}, M)$，即 $\{x_n\}$ 有界。

⑤设 f 一致连续，$\{x_n\}$ 是 Cauchy 列。任给 $\varepsilon > 0$，由一致连续性存在 $\delta > 0$（与点无关），使得 $d_1(x, y) < \delta \Rightarrow d_2(f(x), f(y)) < \varepsilon$。又因 $\{x_n\}$ 是 Cauchy 列，存在 N，使 $m, n > N$ 时 $d_1(x_m, x_n) < \delta$，从而 $d_2(f(x_m), f(x_n)) < \varepsilon$。这说明 $\{f(x_n)\}$ 是 Cauchy 列。

但仅有连续性是不够的：取 $f(x) = \dfrac{1}{x}$，$X = (0, 1)$（配通常的绝对值度量）。f 在 $(0, 1)$ 上处处连续，但不一致连续。取 $x_n = \dfrac{1}{n}$，这是 $\mathbb{R}$ 中的收敛列（收敛到 0），由① 它是 Cauchy 列，且各项都落在 $(0, 1)$ 中（$n \ge 2$ 时），所以 $\{x_n\}$ 是 $(0, 1)$ 中的 Cauchy 列。但 $f(x_n) = n \to \infty$，$\{f(x_n)\}$ 甚至无界，由④（逆否）它不可能是 Cauchy 列。这说明” 连续” 这个较弱的条件不足以保证 Cauchy 列被映成 Cauchy 列。

完备性

设 (X,d) 是度量空间，$A\subset X$。称 A 是**完备**的，如果 A 中任意 Cauchy 列都收敛到 A 中的点。若 X 本身是完备的，就称 (X,d) 为 **完备的度量空间**，也说 d 是完备的。

结合上面命题②③，完备性说的正是"Cauchy 列不会漏出这个空间之外"——$\mathbb{Q}$ 不完备，是因为它里面有 Cauchy 列（如上面那个逼近 $\sqrt{2}$ 的有理数列）"跑到"了 $\mathbb{Q}$ 之外（跑到了 $\mathbb{R}\setminus\mathbb{Q}$）；而完备的空间不会发生这种事。

例

1. $(\mathbb{R},|\cdot|)$ 完备；
2. $[0,1]$ 完备；
3. $\mathbb{R}^n$ 在 d_1,d_2,d_∞（任一种度量）下完备，其中 $d_\infty\le d_2\le d_1\le n\,d_\infty$；
4. 离散度量空间完备；
5. $(C[0,1],d_\infty)$ 完备；
6. $(C[0,1],d_1)$ 不完备（完备化之后就是 L^1 空间）；
7. $c_{00}(\mathbb{N})$（只有有限项非零的实数列全体，配 sup 度量）不完备（完备化之后是 c_0）。

下面逐条说明。

1.$\mathbb{R}$ 的完备性是实数理论的基本事实（比如可以从确界原理推出，这里作为已知结果使用，不再重新证明）。

2.$[0,1]$ 是 $\mathbb{R}$ 中的闭集，而完备空间的闭子集自身也完备：设 $\{x_n\}\subset[0,1]$ 是 Cauchy 列，由 $\mathbb{R}$ 完备，它在 $\mathbb{R}$ 中收敛到某个 $x\in\mathbb{R}$；又 $[0,1]$ 是闭集，收敛列的极限若各项都在闭集中，极限本身也在这个闭集中，所以 $x\in[0,1]$。故 $[0,1]$ 完备。

3. 关键在于：d_1,d_2,d_∞ 这三个度量两两度量等价（由给出的不等式链 $d_\infty\le d_2\le d_1\le n\,d_\infty$，任两个之间都能被彼此的常数倍夹住）。由上一节"度量等价 $\Rightarrow$ 两个方向都是 Lipschitz，从而都是一致连续"，再结合命题⑤，一个度量下的 Cauchy 列自动是另一个度量下的 Cauchy 列；同时度量等价也蕴含拓扑等价，极限点也是同一个。所以只需在一个度量（比如最熟悉的 d_2，欧氏度量）下验证完备性即可，而 $(\mathbb{R}^n,d_2)$ 的完备性可以逐坐标归结到 $\mathbb{R}$ 的完备性：$\{x^{(k)}\}\subset\mathbb{R}^n$ 是 d_2-Cauchy 列时，每个坐标分量数列 $\{x_i^{(k)}\}_k$（$i=1,\dots,n$）都是 $\mathbb{R}$ 中的 Cauchy 列（因为坐标差不超过整体的 d_2 距离），由 $\mathbb{R}$ 完备，各分量收敛，拼起来的极限向量就是 $\{x^{(k)}\}$ 在 $\mathbb{R}^n$ 中的极限。

4. 设 (X,d) 是离散度量空间，$\{x_n\}$ 是 Cauchy 列。取 $\varepsilon=\dfrac12$（小于离散度量唯一的非零取值 1），由 Cauchy 性存在 N，使 $m,n>N$ 时 $d(x_m,x_n)<\dfrac12$；但离散度量只取 0 或 1，所以只能是 $d(x_m,x_n)=0$，即 $x_m=x_n$。这说明 Cauchy 列在离散度量下必然 *最终是常值序列*，从而显然收敛（收敛到这个常值）。故离散空间完备。

5. 设 $\{f_n\}\subset C[0,1]$ 是 $(C[0,1],d_\infty)$ 中的 Cauchy 列。

第一步（构造极限候选，逐点极限）：任给 $\varepsilon>0$，存在 $N=N(\varepsilon)\in\mathbb{N}$，使得 $m,n>N$ 时

$$\max_{x\in[0,1]}|f_m(x)-f_n(x)|<\varepsilon. \tag{$*$}$$

对每个固定的 $x\in[0,1]$，由 $(*)$，$\{f_n(x)\}$ 是 $\mathbb{R}$ 中的 Cauchy 列（因为 $|f_m(x)-f_n(x)|\le\max_t|f_m(t)-f_n(t)|<\varepsilon$）。由 $\mathbb{R}$ 的完备性，令 $f(x):=\lim_{n\to\infty}f_n(x)$，即 f_n 在 $[0,1]$ 上逐点收敛到某个函数 f。

第二步（逐点收敛加强为一致收敛）：固定 $(*)$ 中的 N，对任意 $n>N$、任意 $x\in[0,1]$，$(*)$ 中令 $m\to\infty$（此时 $f_m(x)\to f(x)$，不等式两边取极限，$<$ 变为 $\le$）：

$$|f(x)-f_n(x)|\le\varepsilon,\qquad\forall x\in[0,1].$$

两边对 x 取上确界，得 $\sup_{x\in[0,1]}|f(x)-f_n(x)|\le\varepsilon$ （对一切 $n>N$）。这说明 f_n 一致收敛到 f，不只是逐点收敛。

第三步（极限函数 f 连续，即 $f\in C[0,1]$）：任取 $x_0\in[0,1]$，$\varepsilon>0$。由第二步，取 n 充分大使 $\sup_x|f(x)-f_n(x)|<\varepsilon/3$；又 f_n 连续，存在 $\delta>0$，使 $|x-x_0|<\delta$ 时 $|f_n(x)-f_n(x_0)|<\varepsilon/3$。于是对这样的 x，

$$\begin{aligned}|f(x)-f(x_0)|&\le|f(x)-f_n(x)|+|f_n(x)-f_n(x_0)|+|f_n(x_0)-f(x_0)|\\&<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

这说明 f 在 x_0 连续；由 x_0 的任意性，f 在 $[0,1]$ 上处处连续，即 $f\in C[0,1]$。

综合三步：$\{f_n\}$ 在 $(C[0,1],d_\infty)$ 中收敛到 $f\in C[0,1]$，故 $(C[0,1],d_\infty)$ 完备。（第三步正是”一致收敛的连续函数列，其极限函数仍连续”这一事实的证明。）

6. 考虑一列 $f_n\in C[0,1]$：

$$f_n(x)=\begin{cases}0, & 0\le x\le\frac12,\\ n\big(x-\frac12\big), & \frac12\le x\le\frac12+\frac1n,\\ 1, & \frac12+\frac1n\le x\le1\end{cases}$$

（在 $\left[\frac12,\frac12+\frac1n\right]$ 上用直线把 0 连到 1，其余地方是常数 0 或 1，f_n 显然连续）。

断言 1：$\{f_n\}$ 是 $(C[0,1],d_1)$ 中的 Cauchy 列。不妨设 $m>n$，这时两个函数只在 $\left[\frac12,\frac12+\frac1n\right]$ 上可能不同（在这个区间之外，只要 $x\ge\frac12+\frac1n\ge\frac12+\frac1m$ 就有 $f_m(x)=f_n(x)=1$，$x\le\frac12$ 时都是 0）。令 $t=x-\frac12\in[0,\frac1n]$，则 f_n 在这段是 nt，f_m 在这段是 $\min(1,mt)$（$t\ge\frac1m$ 之后封顶为 1）。分两段计算：

$$\int_0^{1/m}\big|(m-n)t\big|\,dt=\frac{m-n}{2m^2},\qquad\int_{1/m}^{1/n}\big|1-nt\big|\,dt=\frac{(m-n)^2}{2m^2n}$$

（第二段是从 $1-\frac{n}{m}$ 线性降到 0 的三角形面积）。两段相加，化简后得到

$$d_1(f_m,f_n)=\frac{m-n}{2m^2}+\frac{(m-n)^2}{2m^2n}=\frac{m-n}{2mn}=\frac{1}{2}\left(\frac{1}{n}-\frac{1}{m}\right)\le\frac{1}{2n}\xrightarrow[n\to\infty]{}0,$$

所以 $\{f_n\}$ 是 $(C[0,1],d_1)$ 中的 Cauchy 列。

断言 2：$\{f_n\}$ 在 $(C[0,1],d_1)$ 中不收敛。用反证法。假设存在 $f\in C[0,1]$ 使 $d_1(f_n,f)\to 0$。

先看右边：任取 $a\in\left(\frac{1}{2},1\right)$，只要 n 大到 $\frac{1}{n}<a-\frac{1}{2}$（即 $\frac{1}{2}+\frac{1}{n}<a$），就有 $f_n\equiv 1$ 在 $[a,1]$ 上。于是对这样的 n，

$$\int_a^1|1-f(x)|\,dx=\int_a^1|f_n(x)-f(x)|\,dx\le d_1(f_n,f)\xrightarrow[n\to\infty]{}0.$$

左边 $\int_a^1|1-f(x)|dx$ 是一个不依赖 n 的固定数，却被一列趋于 0 的数控制住，只能恒为 0。而 $|1-f(x)|$ 连续非负，积分为零又意味着它处处为零（如果某点 $|1-f(x_0)|>0$，由连续性它在 x_0 附近一个小区间上也保持远离 0，积分就不可能是 0）。所以 $f\equiv 1$ 在 $[a,1]$ 上；由 a 可以任意接近 $\frac{1}{2}$，得 $f\equiv 1$ 在 $\left(\frac{1}{2},1\right]$ 上。

再看左边：对一切 n，由定义 $f_n\equiv 0$ 在 $[0,\frac{1}{2}]$ 上，所以

$$\int_0^{1/2}|f(x)|\,dx=\int_0^{1/2}|f_n(x)-f(x)|\,dx\le d_1(f_n,f)\xrightarrow[n\to\infty]{}0,$$

同样的论证给出 $f\equiv 0$ 在 $\left[0,\frac{1}{2}\right]$ 上，特别地 $f(\frac{1}{2})=0$。

现在看 f 在 $x=\frac{1}{2}$ 处：一方面 $f(\frac{1}{2})=0$；另一方面，f 在 $\left(\frac{1}{2},1\right]$ 上恒为 1，由 f 连续，$\lim_{x\to\frac{1}{2}^+}f(x)=f(\frac{1}{2})$，但左边这个极限应该是 1（因为 $x>\frac{1}{2}$ 时 $f(x)\equiv 1$），右边却是 0，矛盾。所以假设不成立，$\{f_n\}$ 在 $(C[0,1],d_1)$ 中没有极限。

这说明 $(C[0,1],d_1)$ 不完备。这也是为什么后面课程要引入 L^1 空间——它正是把 $C[0,1]$ 按 d_1 这个度量" 补全" 之后得到的完备空间（$C[0,1]$ 的 **完备化**）。

7. $c_{00}(\mathbb{N})=\{x=(x_1,x_2,\dots,x_k,\dots):x_k\in\mathbb{R},\ x_k\text{ 中只有有限项不为0}\}$，配范数 $\|x\|_\infty=\max_k|x_k|$ 和度量 $d_\infty(x,y)=\max_k|x_k-y_k|$，这显然是一个度量空间。

取

$$x^{(n)}=\left(1,\ \frac{1}{2},\ \frac{1}{3},\ \dots,\ \frac{1}{n},\ 0,\ 0,\ \dots\right)\in c_{00}(\mathbb{N}).$$

对 $m>n$，$x^{(m)}$ 与 $x^{(n)}$ 只在第 $n+1,\dots,m$ 个坐标上不同（分别是 $\frac{1}{n+1},\dots,\frac{1}{m}$ 与 0），这些坐标里差的绝对值最大是 $\frac{1}{n+1}$（因为 $\frac{1}{k}$ 递减），所以 $d_\infty(x^{(n)},x^{(m)})=\frac{1}{n+1}\to 0$，即 $\{x^{(n)}\}$ 是 Cauchy 列。

但它在 $c_{00}(\mathbb{N})$ 中不收敛：假设某个 $y\in c_{00}(\mathbb{N})$ 是它的极限，y 只有有限项非零，设 $y_k=0$（$k>M$）。取固定坐标 $k=M+1$：只要 $n\ge M+1$，就有 $x^{(n)}_{M+1}=\frac{1}{M+1}$（不随 n 变化），而 $y_{M+1}=0$，所以 $d_\infty(x^{(n)},y)\ge\left|x^{(n)}_{M+1}-y_{M+1}\right|=\frac{1}{M+1}$ 对一切 $n\ge M+1$ 成立，不趋于 0，矛盾。所以 $\{x^{(n)}\}$ 没有 $c_{00}(\mathbb{N})$ 中的极限——它的" 极限候选" 其实是

$x^* = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right)$，这是一个无穷多项非零、但趋于 0 的数列，$x^* \notin c_{00}(\mathbb{N})$。这说明 $c_{00}(\mathbb{N})$ 不完备，它的完备化正是

$$c_0 := \left\{ x = (x_1, \dots, x_n, \dots) : x_n \in \mathbb{R},\ \lim_{n\to\infty} x_n = 0 \right\}.$$

完备与闭集

定理 1

设 (X, d) 是一个度量空间，$A \subset X$。若 (A, d) 完备，则 A 是 X 中的闭集。

注：这一条不需要假设 X 本身完备。

证明

只需证明 $A = \overline{A}$。任取 $x \in \overline{A}$，由闭包的定义（上一讲），存在 $\{x_n\} \subset A$，使 $x_n \xrightarrow{d} x$。由于 $\{x_n\}$ 收敛，由本节命题①，它是 A 中的 Cauchy 列；又 A 完备，所以存在 $y \in A$，使 $x_n \xrightarrow{d} y$。由极限的唯一性，$x = y$，从而 $x \in A$。这说明 $\overline{A} \subset A$；结合恒有的 $A \subset \overline{A}$，得 $A = \overline{A}$，即 A 是闭集。

定理 2

设 (X, d) 是一个完备的度量空间，$A \subset X$，则

$$(A, d) \text{ 完备} \quad \Longleftrightarrow \quad A \text{ 是} X \text{ 中的闭集.}$$

证明。（⇒）就是刚证过的定理 1（甚至不需要 X 完备）。下面证（⇐）：设 A 是闭集，$\{x_n\} \subset A$ 是 Cauchy 列。因为 X 完备，$\{x_n\}$ 在 X 中收敛到某个 $x \in X$。又 A 闭，$x_n \in A$ 且 $x_n \to x$ 意味着 x 是 A 的一个接触点，即 $x \in \overline{A} = A$（用到 A 闭）。所以 $\{x_n\}$ 收敛到 A 中的点 x，即 (A, d) 完备。这里 X 的完备性是关键——它保证了 $\{x_n\}$ 在”更大”的 X 里确实有极限可以收敛到，再靠 A 闭把这个极限拉回 A 内部。

注（判断是否完备的两个常用方法）：

(1) 要证明某个集合完备，往往只需证明它是一个已知完备空间的 闭子集（定理 2 的 ⇐ 方向）；

(2) 要证明某个集合不完备，往往可以把它看作一个已知完备空间里不闭的子集（定理 1 的逆否）。比如 $(0, 1] \subset \mathbb{R}$ 不闭，所以不完备；又比如 $c_{00} \subset \ell^\infty$（$\ell^\infty$ 是全体有界实数列配 sup 度量，完备性的证明和 $(C[0,1], d_\infty)$ 完备的证明是同一套路：逐点取极限、再证一致收敛、最后说明极限数列仍有界），c_{00} 在 ℓ^∞ 中不闭（上面 7. 的 $x^{(n)} \to x^* \notin c_{00}$ 就是见证），所以由定理 1 的逆否，c_{00} 不完备——这和上面直接验证的结论一致，从另一个角度重新说明了一遍。

1.7 压缩映射与 Banach 不动点定理

压缩映射

设 (X,d) 是度量空间。称 $f:X\to X$ 是**压缩映射**，如果它是 α-Lipschitz 的，且 Lipschitz 常数 $\alpha<1$，即

$$\exists\,\alpha\in(0,1),\quad \text{s.t.}\quad d\big(f(x),f(y)\big)\le \alpha\, d(x,y),\qquad \forall x,y\in X.$$

注意压缩映射要求 f 把 X 映到**自身**（$f:X\to X$），这是为了下面能不断迭代 $f(f(f(\cdots)))$；这一点在后面的反例里会看到它的重要性。

例

1. $\mathbb{R}$ 上 $f(x)=\alpha x$（$|\alpha|<1$）是压缩映射：$d(f(x),f(y))=|\alpha x-\alpha y|=|\alpha|\,|x-y|$，压缩常数就是 $|\alpha|<1$ 本身。
2. $\sin x$ 与 $|x|$ 是 1-Lipschitz 的，但不是压缩映射：$\lim\limits_{x\to0}\dfrac{|\sin x-\sin 0|}{|x-0|}=1$（即 $(\sin x)'|_{x=0}=1$）。具体地，sin 在 $\mathbb{R}$ 上确实 1-Lipschitz （因为 $|\sin'|=|\cos|\le1$，由中值定理 $|\sin x-\sin y|=|\cos c|\,|x-y|\le|x-y|$）。但要成为压缩映射，需要存在**某个固定的** $\alpha<1$ 对**所有**点对都成立；而比值 $\dfrac{|\sin x-\sin 0|}{|x-0|}$ 当 $x\to0$ 时趋于导数值 1，也就是说这个比值可以任意接近 1，找不到一个一致的 $\alpha<1$ 把它压下去。$|x|$ 也类似：当 x,y 同号时 $\big||x|-|y|\big|=|x-y|$，比值恒为 1，更直接地说明它不是压缩映射。
3. $f(x)=x^3$ 在 $\mathbb{R}$ 上不是 Lipschitz 的，但限制在 $\left[-\frac13,\frac13\right]$ 上是 $\frac13$-压缩映射。事实上，在 $\left[-\frac13,\frac13\right]$ 上 $|f'(x)|=3x^2\le3\cdot\left(\frac13\right)^2=\frac13$，由中值定理，$|f(x)-f(y)|\le\frac13|x-y|$，压缩常数 $\alpha=\frac13<1$；而在整个 $\mathbb{R}$ 上 $3x^2$ 无界（$x\to\infty$ 时），找不到统一的 Lipschitz 常数，所以 x^3 在 $\mathbb{R}$ 上不是（整体）Lipschitz 的。
4. 称 f 是**减小距离的**，如果 $d(f(x),f(y))<d(x,y)$ （$x\neq y$）。这**不等价于**压缩映射。例如 $h(x)=\sqrt{x^2+1}$，$|h'(x)|=\dfrac{|x|}{\sqrt{x^2+1}}<1$。具体地，由中值定理，$h(x)-h(y)=h'(c)(x-y)$（$c$ 在 x,y 之间），而 $|h'(c)|=\dfrac{|c|}{\sqrt{c^2+1}}<1$ 对**每一个有限的** c 都严格成立，所以 $|h(x)-h(y)|<|x-y|$（$x\neq y$），即 h 逐点”减小距离”。但当 $|c|\to\infty$ 时 $|h'(c)|\to1$（只是永远达不到 1），也就是说 $\sup_{c\in\mathbb{R}}|h'(c)|=1$，找不到一个**一致的** $\alpha<1$ 使 $|h(x)-h(y)|\le\alpha|x-y|$ 对一切 x,y（尤其是 x,y 都很大时）成立。这个例子说明”逐点减小距离”是比”压缩映射”更弱的条件——少了”用同一个 α 控制所有点对”这个一致性，下面 Banach 定理的证明里会看到这个一致性正是关键。

Banach 压缩映射原理

先引入一个基本概念。设 $g: X \to X$，称 $x_0 \in X$ 是 g 的**不动点**，如果 $g(x_0) = x_0$。

注：$x - g(x)$ 的零点对应 g 的不动点——把”求不动点”看成”求某个函数的零点”，是分析里常见的转化。

定理（Banach 压缩映射原理）

设 (X, d) 是一个完备的度量空间，$f: X \to X$ 是压缩映射，则

$$\exists!\, x \in X, \quad \text{s.t.} \quad f(x) = x$$

（即 f 在 X 中有且仅有一个不动点）。

证明

唯一性。因为 f 是压缩映射，存在 $\alpha \in (0,1)$，使 $d(f(x), f(y)) \le \alpha\, d(x,y)$ 对一切 $x, y \in X$ 成立。若有两个不动点 $x_1 \neq x_2$（即 $f(x_i) = x_i$，$i = 1, 2$），则

$$d(x_1, x_2) = d\big(f(x_1), f(x_2)\big) \le \alpha\, d(x_1, x_2).$$

因为 $x_1 \neq x_2$，$d(x_1, x_2) > 0$，两边同除以 $d(x_1, x_2)$ 得 $1 \le \alpha$，与 $\alpha < 1$ 矛盾。所以不动点至多一个。

存在性。任取 $x_0 \in X$，归纳地定义序列

$$x_{n+1} = f(x_n), \qquad n \ge 0.$$

断言：$\{x_n\}_{n \ge 0}$ 是 *Cauchy* 列。先估计相邻两项的距离：

$$\begin{aligned} d(x_2, x_1) &= d\big(f(x_1), f(x_0)\big) \le \alpha\, d(x_1, x_0), \\ d(x_3, x_2) &= d\big(f(x_2), f(x_1)\big) \le \alpha\, d(x_2, x_1) \le \alpha^2 d(x_1, x_0), \end{aligned}$$

依此类推（归纳法），一般地

$$d(x_{n+1}, x_n) \le \alpha^n\, d(x_1, x_0), \qquad \forall n \ge 0.$$

于是对任意 $m > n$，由三角不等式反复展开，

$$\begin{aligned} d(x_m, x_n) &\le d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\le \left(\alpha^{m-1} + \alpha^{m-2} + \cdots + \alpha^n\right) d(x_1, x_0), \end{aligned}$$

右边是等比数列求和：

$$\left(\alpha^{m-1} + \cdots + \alpha^n\right) d(x_1, x_0) = \frac{\alpha^n - \alpha^m}{1 - \alpha}\, d(x_1, x_0) \le \frac{\alpha^n}{1 - \alpha}\, d(x_1, x_0).$$

因为 $0<\alpha<1$，$\alpha^n\to 0$（$n\to\infty$），所以上式当 $n\to\infty$ 时趋于 0（对一切 $m>n$ 一致地趋于 0），这正是 Cauchy 列的定义。

由 (X,d) 的完备性，存在 $a\in X$，使 $\lim\limits_{n\to\infty}x_n=a$。又由 $x_{n+1}=f(x_n)$，两边令 $n\to\infty$ 取极限：左边 $\to a$；右边 $f(x_n)\to f(a)$（因为压缩映射本身就是连续的：压缩 $\Rightarrow$ Lipschitz$\Rightarrow$ 连续，前面讲过 Lipschitz 映射保持收敛列的极限）。所以 $a=f(a)$，即 a 就是 f 的不动点。

综合两步，f 在 X 中有且仅有一个不动点。

注

① 证明表明，从*任意* $x_0\in X$ 出发做迭代 $x_n=f(f(\cdots f(x_0)))=f^n(x_0)$，都收敛到*同一个*不动点——不动点不依赖于迭代的起点。

② 误差估计：在上面 $\forall m>n$ 的不等式里令 $m\to\infty$（用 $x_m\to a$），得到

$$d(x_n,a)\le\frac{\alpha^n}{1-\alpha}\,d(x_1,x_0),$$

这给出了用第 n 步迭代 x_n 逼近真正不动点 a 时的误差上界，在数值计算里很有用——不需要知道 a 是谁，只需要知道 α 和头两步的距离，就能估计迭代多少步能达到给定精度。

定理中的两个假设（X 完备；f 是 X 到自身的压缩映射）缺一不可，下面用两个反例说明。

反例一（X 不完备）：取 $X=\left(0,\frac14\right)$（配通常的绝对值度量），$f(x)=x^2$。首先 f 确实把 X 映入 X：$x\in\left(0,\frac14\right)$ 时 $x^2\in\left(0,\frac{1}{16}\right)\subset\left(0,\frac14\right)$。其次 f 是压缩映射：$|f'(x)|=2x<2\cdot\frac14=\frac12$ 对 $x\in\left(0,\frac14\right)$ 成立，所以 $\alpha=\frac12$ 即可。但方程 $x^2=x$ 的解只有 $x=0$ 或 $x=1$，都不在*开*区间 $\left(0,\frac14\right)$ 内——f 在这个 X 里没有不动点！问题出在 $X=\left(0,\frac14\right)$ 不是 $\mathbb{R}$ 中的闭集，由上一节的定理，它不完备（比如 $x_n=\frac{1}{n+10}\to0\notin X$ 就是一个跑出 X 的 Cauchy 列）。这说明 Banach 定理里"X 完备"这个假设不能去掉。

反例二（f 不是自映射）：取 $X=[0,1]$，但 $f:[0,1]\to\mathbb{R}$，$f(x)=\frac12x+1$。f 确实是压缩映射（$\frac12$-Lipschitz，$\frac12<1$），但 $f\big([0,1]\big)=\left[1,\frac32\right]\not\subset[0,1]$——$f$ 并没有把 X 映到*自身*，不满足 $f:X\to X$ 这个前提。方程 $x=\frac12x+1$ 的解是 $x=2$，不在 $[0,1]$ 里，f 在 $[0,1]$ 内没有不动点。这说明"$f:X\to X$"（压缩映射的定义本身要求的自映射性质）也不能少。

推广：某次迭代是压缩映射即可

定理

设 (X,d) 完备，$f:X\to X$，且存在 $k\ge1$，使 f^k（f 自己复合 k 次）是压缩映射，则 f 有唯一的不动点 x^*，且对任意 $x_0\in X$，$f^n(x_0)\to x^*$（$n\to\infty$）。

证明

记 $g = f^k$。因为 g 是完备空间 X 上的压缩映射，由 Banach 定理，g 有唯一不动点 x^*，且对任意 $y_0 \in X$，$g^n(y_0) \to x^*$。

x^* 也是 f 的不动点：f 与 $g = f^k$ 显然可交换（都是 f 的方幂）：$f \circ g = f \circ f^k = f^{k+1} = f^k \circ f = g \circ f$。于是

$$g\big(f(x^*)\big) = f\big(g(x^*)\big) = f(x^*),$$

即 $f(x^*)$ 也是 g 的一个不动点；而 g 的不动点唯一（就是 x^*），所以 $f(x^*) = x^*$。

f 的不动点唯一：若 $f(y) = y$，两边反复作用 f 得 $g(y) = f^k(y) = y$，即 y 也是 g 的不动点，由 g 不动点唯一，$y = x^*$。

从任意起点迭代收敛：这一步需要小心处理——定理只假设了 $g = f^k$ 是压缩映射（从而 g 自动连续），*并没有假设 f 本身连续*，所以不能像”取极限时把 f^r 直接作用到极限上”那样处理。下面用交换律绕开这个问题。

引理（f^r 与 g 可交换）：对一切整数 $r \geq 0$，$f^r \circ g = g \circ f^r$。

证明：对 r 归纳。$r = 0$ 时 $f^0 = \mathrm{id}_X$，显然与 g 交换。设 $f^{r-1} \circ g = g \circ f^{r-1}$ 成立，则用前面已证的 $f \circ g = g \circ f$，

$$f^r \circ g = f \circ (f^{r-1} \circ g) = f \circ (g \circ f^{r-1}) = (f \circ g) \circ f^{r-1} = (g \circ f) \circ f^{r-1} = g \circ f^r,$$

即 r 时也成立。故对一切 $r \geq 0$ 结论成立；进而把它反复作用 q 次，对一切 $q \geq 1$ 也有 $f^r \circ g^q = g^q \circ f^r$。

事实上，这个引理的证明纯粹是映射复合的代数运算（结合律 + 反复使用 $f \circ g = g \circ f$），*完全没有用到 f 连续*，对任意映射 $f : X \to X$ 都成立。

收敛性的证明：任取 $x_0 \in X$。对 $n = qk + r$（$0 \leq r < k$，$q \to \infty$），

$$f^n(x_0) = f^{qk+r}(x_0) = f^r\big(f^{qk}(x_0)\big) = f^r\big(g^q(x_0)\big) \overset{\text{引理}}{=} g^q\big(f^r(x_0)\big).$$

固定 r，记 $z_r := f^r(x_0) \in X$——这是 X 中一个*不随 q 变化*的固定点。因为 g 是完备空间 X 上的压缩映射，由 Banach 定理，从*任意*起点出发迭代 g 都收敛到 x^*；特别取起点为 z_r，得

$$f^{qk+r}(x_0) = g^q(z_r) \longrightarrow x^* \qquad (q \to \infty).$$

这对每一个固定的余数 $r \in \{0, 1, \ldots, k-1\}$（只有有限个）都成立，所以整个序列 $f^n(x_0) \to x^*$（$n \to \infty$）。

关键的技巧是先用交换律把 f^r 挪到*内层*，作用在固定点 $z_r = f^r(x_0)$ 上，这样”随 $q \to \infty$ 变化”的部分就完全由 g^q 承担——而 g 是压缩映射，Banach 定理本身就直接给出”从任意起点收敛”，不需要再对外层的 f^r 取极限（也就不需要 f^r 连续这个额外假设）。如果反过来想先对 $g^q(x_0) \to x^*$ 取极限、再把 f^r”穿过”极限号作用上去（即 $f^r(g^q(x_0)) \to f^r(x^*)$），那一步等价于假设 f^r 连续，而这并不是定理的已知条件，是不能这样用的。

应用

(a) 隐函数定理的证明；

(b) Newton 迭代 $g(x)=x-\dfrac{F(x)}{F'(x)}$ 在单根附近也是压缩映射（因为 $g'(x^*)=0$），故收敛；

(c) 一阶 ODE Cauchy 问题解的存在唯一性；

(d) 非线性方程组解的存在性。

对 (b)，设 x^* 是 F 的单根，即 $F(x^*)=0$ 且 $F'(x^*)\neq 0$。直接求导：

$$g'(x)=1-\frac{F'(x)F'(x)-F(x)F''(x)}{F'(x)^2}=\frac{F(x)F''(x)}{F'(x)^2},$$

代入 $x=x^*$：因为 $F(x^*)=0$，$g'(x^*)=0$。由 $F'(x^*)\neq 0$ 及 $F\in C^2$，g' 在 x^* 附近连续，故存在 x^* 的一个邻域，使 $|g'|$ 在其中小于 1（比如小于 $\frac{1}{2}$），由中值定理，g 在这个邻域内是压缩映射（把邻域取成一个以 x^* 为心的闭区间，并验证 g 把它映入自身，即可套用 Banach 定理），这就是 Newton 迭代在单根附近局部收敛的原因。(c)(d) 分别在下面两个例子里展开。

应用一：一阶 ODE 的存在唯一性（Picard–Lindelöf 定理）

定理

设 $f(t,x)$ 在矩形 $R=[-a,a]\times[\zeta-b,\zeta+b]$ 上连续，$M=\max_R|f|$，且 f 关于 x（在 R 上）一致 Lipschitz：

$$|f(t,x)-f(t,y)|\le L\,|x-y|,\qquad \forall\,|t|\le a,\ x,y\in[\zeta-b,\zeta+b].$$

令

$$h=\min\left\{a,\ \frac{b}{M},\ \frac{1}{2L}\right\}.$$

则 Cauchy 问题

$$\begin{cases}\dfrac{dx}{dt}=f(t,x),\\ x(0)=\zeta\end{cases}\tag{1}$$

在 $[-h,h]$ 上有唯一解。

证明

把微分方程改写成积分方程（不动点问题）。(1) 的解等价于

$$x(t)-x(0)=\int_0^t f\big(s,x(s)\big)\,ds$$

$$\Longleftrightarrow\quad x(t)=\zeta+\int_0^t f\big(s,x(s)\big)\,ds\ =:\ (Tx)(t).$$

这样一来，求解 (1) 就转化成寻找算子 T 的不动点。

构造完备的度量空间 X。取上面定义的 $h \in (0, a]$，令

$$X = \Big\{ x \in C[-h,h] \ : \ \max_{-h\le t\le h} |x(t) - \zeta| \le b \Big\}$$

（X 是以常值函数 ζ 为中心、半径 b 的闭球）。X 是 $\big(C[-h,h], d_\infty\big)$ 中的闭集（这是本讲前面" 完备与闭集" 那一节里" 闭集" 的直接验证：若 $x_n \in X$ 且 $x_n \to x$（d_∞ 意义下），则 $|x(t) - \zeta| = \lim_n |x_n(t) - \zeta| \le b$ 对每个 t 成立，所以极限 $x \in X$），而 $(C[-h,h], d_\infty)$ 完备（和 $(C[0,1], d_\infty)$ 完备的证明是同一套），所以由" 闭子集完备" 定理，X 完备。

定义算子 T。对 $x \in X$，令

$$(Tx)(t) = \zeta + \int_0^t f\big(s, x(s)\big)\, ds, \qquad t \in [-h, h].$$

因为 $s \mapsto f(s, x(s))$ 连续（f 连续、x 连续的复合），Tx 显然是 $[-h,h]$ 上的连续函数，即 $Tx \in C[-h,h]$。下面验证 $T : X \to X$ 是压缩映射。

① T 把 X 映入 X。对任意 $x \in X$，$t \in [-h,h]$，

$$\big|(Tx)(t) - \zeta\big| = \left| \int_0^t f\big(s, x(s)\big)\, ds \right| \le \int_0^{|t|} \big| f\big(s, x(s)\big) \big|\, ds \le |t|\, M \le hM \le b$$

（最后一步用 $h \le b/M$）。所以 $\max_{|t|\le h} |(Tx)(t) - \zeta| \le b$，即 $Tx \in X$。

② T 是压缩映射。对任意 $x, y \in X$，

$$\begin{aligned} \max_{|t|\le h} \big|(Tx)(t) - (Ty)(t)\big| &= \max_{|t|\le h} \left| \int_0^t \big[f(s, x(s)) - f(s, y(s)) \big]\, ds \right| \\ &\le \max_{|t|\le h} |t| \cdot L \cdot \max_{|s|\le h} |x(s) - y(s)| \\ &\le hL\, d_\infty(x, y). \end{aligned}$$

即 $d_\infty(Tx, Ty) \le hL\, d_\infty(x, y)$。由 $h \le \dfrac{1}{2L}$，压缩常数 $\alpha := hL \le \frac{1}{2} < 1$，所以 T 是压缩映射。

结论。由 Banach 压缩映射原理，T 在完备空间 X 上有唯一不动点 $x \in X$，即

$$x(t) = \zeta + \int_0^t f\big(s, x(s)\big)\, ds, \qquad t \in [-h, h].$$

两边对 t 求导（用微积分基本定理，被积函数 $f(s, x(s))$ 连续），得 $x'(t) = f(t, x(t))$，且显然 $x(0) = \zeta$，这正是 (1) 的解。

进一步，(1) 在 $[-h,h]$ 上的*所有*解都是唯一的，而不只是在 X 里唯一。设 $\tilde{x}$ 是 (1) 在 $[-h,h]$ 上的*任意*连续解，则 $\tilde{x}$ 也满足积分方程 $\tilde{x} = T\tilde{x}$，套用和" ①" 完全相同的估计（把 x 换成 $\tilde{x}$），得 $\max_{|t|\le h} |\tilde{x}(t) - \zeta| \le hM \le b$，即 $\tilde{x}$ 自动落在 X 里。所以 (1) 的任何解都是 T 在 X 中的不动点，由 T 在 X 中不动点唯一，(1) 在 $[-h,h]$ 上的解也唯一。

注

① 如果只假设 f 连续（不假设 Lipschitz），解仍然存在（Peano 存在定理），但可能失去唯一性。例如

$$\frac{dx}{dt} = 2\sqrt{|x|}, \qquad x(0) = 0$$

同时有解 $x \equiv 0$ 和 $x(t) = t|t|$（可以验证：$t \geq 0$ 时 $x(t) = t^2$，$x'(t) = 2t = 2\sqrt{|x(t)|}$；$t < 0$ 时 $x(t) = -t^2$，$x'(t) = -2t = 2|t| = 2\sqrt{|x(t)|}$，两段都满足方程，在 $t = 0$ 处也光滑衔接）——问题出在 $f(t,x) = 2\sqrt{|x|}$ 在 $x = 0$ 附近不是 Lipschitz 的（$\partial f/\partial x = 1/\sqrt{|x|} \to \infty$）。

② 证明本身是构造性的：可以用 T 反复迭代 $x_0(t) \equiv \zeta$，$x_{n+1} = Tx_n$ 来逐步逼近真正的解——这就是常微分方程里的**Picard 逐次逼近法**。

应用二：非线性方程组的存在唯一性

例

设 $a(t), b(t)$ 是给定的函数，证明对每个固定的 t，方程组

$$\begin{cases} x = \dfrac{1}{2}\cos(x+y) + a(t), \\ y = \dfrac{1}{3}\sin(x-y) + b(t) \end{cases}$$

都有唯一解 (x_t, y_t)。

证明

固定 t，在 $\mathbb{R}^2$ 上考虑映射

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\cos(x+y) + a \\ \frac{1}{3}\sin(x-y) + b \end{pmatrix},$$

则 $T : \mathbb{R}^2 \to \mathbb{R}^2$ 是连续映射。在 $\mathbb{R}^2$ 上取度量

$$d_1\big((x_1,y_1),(x_2,y_2)\big) = |x_1 - x_2| + |y_1 - y_2|$$

（$(\mathbb{R}^2, d_1)$ 是完备度量空间——有限维空间在任何等价的度量下都完备）。由 $|\cos u - \cos v| \leq |u - v|$，$|\sin u - \sin v| \leq |u - v|$（都是 1-Lipschitz，前面讲过），分别有

$$\begin{aligned} \left|\tfrac{1}{2}\cos(x_1{+}y_1) - \tfrac{1}{2}\cos(x_2{+}y_2)\right| &\leq \tfrac{1}{2}\big|(x_1{+}y_1) - (x_2{+}y_2)\big| \\ &\leq \tfrac{1}{2}\big(|x_1{-}x_2| + |y_1{-}y_2|\big), \end{aligned}$$

$$\begin{aligned} \left|\tfrac{1}{3}\sin(x_1{-}y_1) - \tfrac{1}{3}\sin(x_2{-}y_2)\right| &\leq \tfrac{1}{3}\big|(x_1{-}y_1) - (x_2{-}y_2)\big| \\ &\leq \tfrac{1}{3}\big(|x_1{-}x_2| + |y_1{-}y_2|\big), \end{aligned}$$

两式相加，得

$$\begin{aligned} d_1\big(T(x_1,y_1),T(x_2,y_2)\big) &\le \frac{1}{2}\big(|x_1-x_2|+|y_1-y_2|\big) \\ &\quad + \frac{1}{3}\big(|x_1-x_2|+|y_1-y_2|\big) \\ &= \frac{5}{6}\, d_1\big((x_1,y_1),(x_2,y_2)\big). \end{aligned}$$

故 T 是压缩映射（$\alpha=\frac{5}{6}<1$），存在唯一不动点，即方程组在 $\mathbb{R}^2$ 上有唯一解。

注（度量的选取有讲究；压缩映射不是拓扑性质）：如果换用 $d_\infty\big((x_1,y_1),(x_2,y_2)\big)=\max\{|x_1-x_2|,|y_1-y_2|\}$ 来重复上面的估计，会发现

$$\begin{aligned} \left|\tfrac{1}{2}\cos(x_1+y_1)-\tfrac{1}{2}\cos(x_2+y_2)\right| &\le \tfrac{1}{2}\big(|x_1-x_2|+|y_1-y_2|\big) \\ &\le \tfrac{1}{2}\cdot 2\, d_\infty\big((x_1,y_1),(x_2,y_2)\big) \\ &= d_\infty\big((x_1,y_1),(x_2,y_2)\big), \end{aligned}$$

即用 d_∞ 只能得到系数 1（不是压缩！），sin 那一项同理最多给出系数 $\frac{2}{3}$，两者取 max 后压缩常数依然是 1，达不到 <1。也就是说 **同一个映射 T**，换一个（哪怕是拓扑等价甚至度量等价的）度量，”是不是压缩映射”这个结论可能就不成立了——压缩映射的”压缩常数 α”是一个和具体度量的数值挂钩的概念，不像”连续””收敛”那样只依赖于拓扑本身。这就是为什么解决实际问题时，度量要选得”合适”：这里用 d_1 而不是 d_∞，才能把 Banach 定理用上。

1.8　紧性与列紧性

基本定义

设 (X,d) 是度量空间，$A\subset X$。

定义（紧集）：称 A 是**紧集**，如果 A 的任意开覆盖都有有限子覆盖：即对 A 的任意一族开集 $\{U_i\}_{i\in I}$，若 $A\subset\bigcup_{i\in I}U_i$，则存在有限个指标 $i_1,\dots,i_n\in I$，使得 $A\subset\bigcup_{k=1}^{n}U_{i_k}$。

定义（列紧集）：称 A 是**列紧集**，如果 A 中任意点列 $\{x_n\}$ 都存在子列 $\{x_{n_k}\}$ 在 X 中收敛（极限点未必在 A 内）。

定义（自列紧集）：称 A 是**自列紧集**，如果 A 是闭集且 A 是列紧集（等价地：A 中任意点列都存在子列收敛到 A 中一点）。

定义（预紧集 / 完全有界集）：称 A 是**完全有界集**（也叫 **预紧集**），如果对任意 $\varepsilon>0$，都存在有限个点 $x_1,\dots,x_n\in X$，使得

$$A\subset\bigcup_{i=1}^{n}B(x_i,\varepsilon).$$

称 $\{x_1,\dots,x_n\}$ 为 A 的一个 ε-网。

这四个概念是本节的核心对象，后面会逐步厘清它们之间的关系（在一般度量空间中" 紧" 与" 自列紧" 是等价的，这个等价性属于更深入的内容，这里先只讨论列紧、自列紧与完全有界之间的关系）。

几点注记

注①（列紧的遗传性）： 若 A 是列紧集，$B\subset A$，则 B 未必是列紧集；但若 A 本身列紧，A 的任意子集中的点列自然也是 A 中的点列，故在 A（作为全空间）中取子列收敛是自动成立的——需要小心的是极限点是否落在子集内。更常用的是相反方向的遗传性：若 X 列紧、$A\subset X$ 是闭集，则 A 也列紧（A 中的点列是 X 中的点列，取收敛子列后极限在 X 中，再由 A 闭得极限在 A 中）。

注②（列紧与自列紧的区别）： 以 $\mathbb{R}$ 中的 $(0,1]$ 为例。$(0,1]$ 作为 $\mathbb{R}$ 的子集是列紧的（Bolzano–Weierstrass 定理：$(0,1]$ 中任意点列都是 $\mathbb{R}$ 中的有界点列，故有收敛子列，极限点落在 $\mathbb{R}$ 中，即 $[0,1]$ 内），但 $(0,1]$ *不是*自列紧的：例如点列 $x_n=\frac{1}{n}\in(0,1]$，其任意子列都收敛到 $0\notin(0,1]$，即在 $(0,1]$ 内部找不到收敛子列的极限。

注③（全空间的情形）： 当 $A=X$（即把 A 看作是自己的全空间）时，" 列紧" 与" 自列紧" 是一回事：X 中点列的子列若收敛，极限自然也在 X 中。区别只在于 A 是 X 的*真子集*时才会出现——这也是注②中 $(0,1]\subsetneq\mathbb{R}$ 会出现区别、而把 $(0,1]$ 本身看作全空间时它是" 自列紧"（等价于说 $(0,1]$ 自身列紧）这一说法失效的原因。

完全有界的三点说明

说明 (i)（ε-网可以取在 A 内部）： 完全有界的定义中，ε-网的点 $x_1,\dots,x_n$ 原则上允许取在 X 中任意位置；但事实上总可以要求它们取自 A 本身。

具体地，设 $\{x_1,\dots,x_n\}$ 是 A 的一个 $\varepsilon/2$-网，即 $A\subset\bigcup_{i=1}^{n}B(x_i,\varepsilon/2)$。将其中使得 $A\cap B(x_i,\varepsilon/2)=\varnothing$ 的指标去掉（这些球对覆盖 A 没有贡献），不妨设剩下的指标为 $1,\dots,m$（$m\le n$），即对每个 $i=1,\dots,m$ 都有 $A\cap B(x_i,\varepsilon/2)\neq\varnothing$。在其中各取一点 $y_i\in A\cap B(x_i,\varepsilon/2)$。下证 $\{y_1,\dots,y_m\}$ 是 A 的一个 ε-网（且 $y_i\in A$）：对任意 $a\in A$，存在 $i\in\{1,\dots,m\}$ 使得 $a\in B(x_i,\varepsilon/2)$，即 $d(a,x_i)<\varepsilon/2$；又 $d(x_i,y_i)<\varepsilon/2$（因为 $y_i\in B(x_i,\varepsilon/2)$），故由三角不等式

$$d(a,y_i)\le d(a,x_i)+d(x_i,y_i)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

即 $a\in B(y_i,\varepsilon)$。故 $A\subset\bigcup_{i=1}^{m}B(y_i,\varepsilon)$，且每个 $y_i\in A$。证毕。

说明 (ii)（完全有界 $\Rightarrow$ 有界）： 若 A 完全有界，则 A 有界。这是显然的：取 $\varepsilon=1$ 的有限网 $\{x_1,\dots,x_n\}$，则 A 含于有限个半径为 1 的球的并，而有限个有界集的并是有界集。

说明 (iii)（$\mathbb{R}^n$ 中有界 $\Leftrightarrow$ 完全有界，但一般度量空间中不然）：在 $\mathbb{R}^n$（配备通常的欧氏度量）中，有界集与完全有界集是等价的概念：有界 $\Rightarrow$ 完全有界可以用有限维空间中” 用小方格网格覆盖有界区域” 的方式直接构造出 ε-网。但在一般的（特别是无穷维的）度量空间中，” 有界” 不再蕴含” 完全有界”——下面的例子说明了这一点。

例：$(C[0,1], d_\infty)$ 中的有界集未必完全有界

考虑 $(C[0,1], d_\infty)$，其中 $d_\infty(f,g) = \max_{x\in[0,1]} |f(x) - g(x)|$。构造函数列

$$f_n(x) = \begin{cases} 1 - nx, & 0 \le x < \dfrac{1}{n}, \\ 0, & \dfrac{1}{n} \le x \le 1. \end{cases} \qquad (n = 1, 2, 3, \dots)$$

即 f_n 是一个从 $(0,1)$ 出发、在 $x = \frac{1}{n}$ 降为 0 的” 三角形尖峰”，其后恒为 0。

首先，$\{f_n\}$ 是 $(C[0,1], d_\infty)$ 中的有界集：因为 $0 \le f_n(x) \le 1$ 对所有 n 和所有 $x \in [0,1]$ 成立，故 $d_\infty(f_n, \mathbf{0}) = \max_x |f_n(x)| \le 1$，即 $\{f_n\}$ 落在以零函数为心、半径 1 的球内。

下面计算 $d_\infty(f_m, f_n)$，不妨设 $m > n$（即 $\frac{1}{m} < \frac{1}{n}$）。将 $[0,1]$ 分成三段来看 $f_m - f_n$：

- 当 $x \in \left[0, \frac{1}{m}\right)$ 时，$f_m(x) = 1 - mx$，$f_n(x) = 1 - nx$（因为此时 $x < \frac{1}{m} < \frac{1}{n}$），故 $f_m(x) - f_n(x) = (n-m)x$，其绝对值 $(m-n)x$ 在此区间上单调递增，在右端点处（取极限）趋于 $(m-n)\cdot\frac{1}{m} = 1 - \frac{n}{m}$。
- 当 $x \in \left[\frac{1}{m}, \frac{1}{n}\right)$ 时，$f_m(x) = 0$，$f_n(x) = 1 - nx$，故 $|f_m(x) - f_n(x)| = 1 - nx$，在此区间上单调递减，左端点处取值 $1 - n\cdot\frac{1}{m} = 1 - \frac{n}{m}$。
- 当 $x \in \left[\frac{1}{n}, 1\right]$ 时，$f_m(x) = f_n(x) = 0$，差为 0。

两段在 $x = \frac{1}{m}$ 处的值都是 $1 - \frac{n}{m}$，且这正是 $|f_m - f_n|$ 在 $[0,1]$ 上的最大值。故

$$d_\infty(f_m, f_n) = 1 - \frac{n}{m}, \qquad m > n.$$

由此，对任意固定的 n，当 $m \to \infty$ 时 $d_\infty(f_m, f_n) \to 1$。这说明 $\{f_n\}$ 不是 Cauchy 列，且它的任意子列 $\{f_{n_k}\}$ 也不是 Cauchy 列：取 $\varepsilon = \frac{1}{2}$，对任意 N，取 k 充分大使 $n_k > N$，再取 $j > k$ 使 $n_j \ge 2n_k$，则 $d_\infty(f_{n_j}, f_{n_k}) = 1 - \frac{n_k}{n_j} \ge \frac{1}{2}$，且 $n_j > n_k > N$——即无论 N 多大，总能在其后找到两项距离 $\ge \frac{1}{2}$，故 $\{f_{n_k}\}$ 不满足 Cauchy 准则。从而 $\{f_n\}$ 没有 Cauchy 子列，也就没有收敛子列（收敛列必为 Cauchy 列）。

故 $\{f_n\}$ 没有收敛子列，即 $\{f_n\}$（作为 $C[0,1]$ 的子集）不是列紧集。下面进一步说明它也不是完全有界集。

完全有界与列紧之间有如下关系（在完备度量空间中两者等价，这是后续会遇到的更深刻的结果；这里只用到较易的一个方向）：完全有界的集合中的任意点列都存在

Cauchy 子列。因此，若 $\{f_n\}$ 完全有界，则它必有 Cauchy 子列，但上面已证明它没有 Cauchy 子列——矛盾。故 $\{f_n\}$ 不完全有界。

这个例子说明：说明 (iii) 中”有界 $\Rightarrow$ 完全有界”这一性质，在 $\mathbb{R}^n$ 之外的（无穷维）度量空间中一般不成立——$(C[0,1], d_\infty)$ 中存在有界但不完全有界（从而不列紧）的集合。

自列紧与闭集的关系

推论：若 $A \subset X$ 是自列紧集，则 A 是闭集。

证明：设 $x \in \overline{A}$，则存在 A 中点列 $\{x_n\}$ 使 $x_n \to x$。由 A 自列紧，存在子列 $\{x_{n_k}\}$ 收敛到某 $y \in A$。但 $\{x_{n_k}\}$ 是收敛列 $\{x_n\}$ 的子列，故 $x_{n_k} \to x$，由极限唯一性得 $y = x$，即 $x = y \in A$。故 $\overline{A} \subset A$，即 A 是闭集。证毕。

但反过来，闭集未必自列紧。例如 $[0, +\infty) \subset \mathbb{R}$ 是闭集，但取点列 $x_n = n$，它没有任何收敛子列（任意子列都趋于 $+\infty$），故 $[0, +\infty)$ 不是列紧集，从而也不是自列紧集。这说明”自列紧 $\Rightarrow$ 闭”不能逆推；自列紧比闭集要强（它还额外要求了一种”有界性”）。

1.9　紧性的等价刻画

本节承接上次课”紧性与列紧性”的定义，先补充两个”有界但不完全有界”的例子，再逐步建立预紧性、列紧性、完备性、紧性之间的完整等价关系，最后用一张关系图把这些概念之间的蕴含关系汇总起来。紧集上连续映射的基本性质放在下一节。

补充例子：更多”有界但不完全有界”的集合

例 1（无限集上的离散度量）：设 X 是一个无限集，配备离散度量 $d(x,y) = 1\ (x \neq y)$，$d(x,x) = 0$。则 X 有界：对任意 $x_0 \in X$，$X \subset B(x_0, 2)$。但 X 不完全有界：对 $\varepsilon \leq 1$，任意一个球 $B(x, \varepsilon)$ 只包含 x 自身一点（因为 $y \neq x$ 时 $d(x,y) = 1 \geq \varepsilon$），所以有限个这样的球至多覆盖有限个点，无法覆盖无限集 X。故 X 没有有限 ε-网（$\varepsilon \leq 1$ 时），X 不完全有界。

例 2（$(C[0,1], d_\infty)$ 中的 $g_n(x) = x^n$）：考虑 $g_n(x) = x^n$。因为 $|g_n(x)| \leq 1$ 对 $x \in [0,1]$ 恒成立，$d_\infty(g_n, \mathbf{0}) = 1$，故 $\{g_n\}$ 有界。

下面证明 $\{g_n\}$ 不完全有界。在 $x_n := 2^{-1/n}$ 处，$g_n(x_n) = \left(2^{-1/n}\right)^n = 2^{-1} = \frac{1}{2}$。而对 $m \geq 2n$，

$$g_m(x_n) = \left(2^{-1/n}\right)^m = 2^{-m/n} \leq 2^{-2} = \frac{1}{4} \qquad (\text{因为} m/n \geq 2),$$

所以

$$d_\infty(g_n,g_m)\ge\big|g_n(x_n)-g_m(x_n)\big|\ge\frac12-\frac14=\frac14,\qquad\forall\, m\ge 2n.$$

由此可以取出一个子列 $h_k:=g_{n_k}$（例如 $n_1=1$，$n_{k+1}\ge 2n_k$），使得任意两项之间的距离都 $\ge\frac14$：这是因为对 $k<j$，$n_j\ge 2^{j-k}n_k\ge 2n_k$，故 $d_\infty(h_k,h_j)\ge\frac14$ 恒成立。取 $\varepsilon=\frac18$，半径为 $\frac18$ 的球直径 $\le\frac14$，所以每个这样的球至多包含 $\{h_k\}$ 中的一项（两项若同在一个球内，距离就会 $<\frac14$，与上面矛盾）。于是 $\{h_k\}$（从而 $\{g_n\}$）没有有限的 $\frac18$-网，不完全有界。

事实上，$g_n(x)\xrightarrow{\text{逐点}}\begin{cases}0, & x\in[0,1),\\ 1, & x=1,\end{cases}$ 即 $\{g_n\}$ 只是逐点收敛而非一致收敛（不在 d_∞ 意义下收敛），这正是它在 $(C[0,1],d_\infty)$ 中不列紧、不完全有界的直观原因。

预紧性的遗传性

推论：设 $A\subset(X,d)$ 完全有界（预紧），则：

1. A 的任意子集 $B\subset A$ 也完全有界；
2. A 的闭包 $\overline{A}$ 也完全有界。

证明：1. 显然：若 $\{x_1,\dots,x_k\}$ 是 A 的一个 ε- 网，即 $A\subset\bigcup_{i=1}^k B(x_i,\varepsilon)$，而 $B\subset A$，则同一组点也是 B 的 ε-网。

2. 任取 $\varepsilon>0$，取 A 的有限 $\frac{\varepsilon}{2}$-网 $x_1,\dots,x_k$，即 $A\subset\bigcup_{i=1}^k B\big(x_i,\frac{\varepsilon}{2}\big)$。对任意 $y\in\overline{A}$，存在 $a\in A$ 使 $d(y,a)<\frac{\varepsilon}{2}$；又存在某个 i 使 $a\in B\big(x_i,\frac{\varepsilon}{2}\big)$，即 $d(a,x_i)<\frac{\varepsilon}{2}$。由三角不等式，

$$d(y,x_i)\le d(y,a)+d(a,x_i)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

即 $y\in B(x_i,\varepsilon)$。故 $\overline{A}\subset\bigcup_{i=1}^k B(x_i,\varepsilon)$，即 $\{x_1,\dots,x_k\}$ 也是 $\overline{A}$ 的一个 ε-网。证毕。

自列紧集的性质（续）

上次课已证明：自列紧集必是闭集。这里补充另外两条。

推论：若 $A\subset(X,d)$ 自列紧，则 A 有界，且 (A,d) 完备。

证明（有界性）：反证法。假设 A 无界，取定 $a_0\in A$，则 $\forall n\in\mathbb{N}^*$，$\exists a_n\in A$ 使 $d(a_n,a_0)>n$（否则对某个 n 有 $A\subset B(a_0,n)$，A 就有界了）。于是 $\{a_n\}$ 的任何子列 $\{a_{n_k}\}$ 都满足 $d(a_{n_k},a_0)>n_k\to\infty$，是无界数列；而收敛数列必有界，所以 $\{a_n\}$ 没有收敛子列，这与 A 自列紧（从而列紧）矛盾。故 A 有界。

证明（完备性）：设 $\{a_n\}\subset A$ 是 Cauchy 列。由 A 自列紧，存在子列 $\{a_{n_k}\}\to a\in A$。

下面说明整个数列收敛到同一极限（Cauchy 列有收敛子列，则整个数列收敛）。任给 $\varepsilon>0$，因 $\{a_n\}$ Cauchy，存在 N，使 $\forall m,n\ge N$，$d(a_m,a_n)<\frac{\varepsilon}{2}$。又 $a_{n_k}\to a$，可取

k 充分大，使 $n_k \geq N$ 且 $d(a_{n_k}, a) < \dfrac{\varepsilon}{2}$。于是对一切 $n \geq N$，

$$d(a_n, a) \leq d(a_n, a_{n_k}) + d(a_{n_k}, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

即 $a_n \to a$。

故 $\{a_n\}$（作为 A 中的 Cauchy 列）收敛到 $a \in A$，即 (A, d) 完备。证毕。

预紧性的等价刻画

命题 1（列紧 $\Rightarrow$ 预紧）：若 $A \subset (X, d)$ 列紧，则 A 预紧（完全有界）。

证明：反证法。假设 A 不预紧，则存在 $\varepsilon_0 > 0$，使任何有限个半径为 ε_0 的球都不能覆盖 A。于是可以依次取点：任取 $x_1 \in A$；因 $B(x_1, \varepsilon_0) \not\supset A$，取 $x_2 \in A \setminus B(x_1, \varepsilon_0)$；因 $B(x_1, \varepsilon_0) \cup B(x_2, \varepsilon_0) \not\supset A$，取 $x_3 \in A \setminus \big(B(x_1, \varepsilon_0) \cup B(x_2, \varepsilon_0)\big)$；依此类推，取 $x_{n+1} \in A \setminus \bigcup_{i=1}^{n} B(x_i, \varepsilon_0)$。这样得到的序列 $\{x_n\} \subset A$ 满足 $d(x_m, x_n) \geq \varepsilon_0$ $(\forall m \neq n)$，故没有 Cauchy 子列，更没有收敛子列，这与 A 列紧矛盾。故 A 预紧。

引理 1（预紧的序列刻画）：设 $A \subset (X, d)$，则 A 预紧 $\iff$ A 中任意序列都有 Cauchy 子列。

证明 $(\Leftarrow)$：即证：若 A 中任何序列都有 Cauchy 子列，则 A 预紧。反证法：假设 A 不预紧，同命题 1 证明中的构造，可得 $\varepsilon_0 > 0$ 及 $\{x_n\} \subset A$，使 $\forall m \neq n$，$d(x_m, x_n) \geq \varepsilon_0$。这样的 $\{x_n\}$ 没有 Cauchy 子列（任意两项距离都 $\geq \varepsilon_0$），与假设矛盾。故 A 预紧。

证明 $(\Rightarrow)$：已知 A 预紧，要证 A 中任意序列 $\{x_n\}$ 都有 Cauchy 子列。这是一个逐次取更细的网、再取对角线子列的构造。

Step 1：取 $\varepsilon_1 = 1$。由 A 预紧，存在有限 1-网 $z_1, \dots, z_p \in A$，使 $A \subset \bigcup_{i=1}^{p} B(z_i, 1)$。于是 $\{x_n\} \subset A \subset \bigcup_{i=1}^{p} B(z_i, 1)$：序列 $\{x_n\}$ 有无穷多项，而球只有有限个（p 个），由抽屉原理，必有一个球包含了 $\{x_n\}$ 中的无穷多项，即包含了 $\{x_n\}$ 的一个子列。把这个球的球心重新命名为 $y_1 \in A$，该子列记为 $\{x_n^{(1)}\}$，则 $\{x_n^{(1)}\} \subset \{x_n\}$ 且 $\{x_n^{(1)}\} \subset B(y_1, 1)$。

Step 2：类似地，取 $\varepsilon_2 = \frac{1}{2}$，对子列 $\{x_n^{(1)}\}$ 作同样的操作：A 有有限个 $\frac{1}{2}$-球覆盖，其中某个球包含 $\{x_n^{(1)}\}$ 中的无穷多项。于是可以找到 $y_2 \in A$ 以及 $\{x_n^{(1)}\}$ 的子列 $\{x_n^{(2)}\}$，使 $\{x_n^{(2)}\} \subset B\big(y_2, \frac{1}{2}\big)$。

Step k：由数学归纳法，$\forall k \geq 1$，取 $\varepsilon_k = \frac{1}{k}$，可以找到 $y_k \in A$ 以及 $\{x_n^{(k-1)}\}$ 的子列 $\{x_n^{(k)}\}$，使 $\{x_n^{(k)}\} \subset B\big(y_k, \frac{1}{k}\big)$，且 $\{x_n^{(k)}\} \subset \{x_n^{(k-1)}\} \subset \cdots \subset \{x_n^{(1)}\} \subset \{x_n\}$（第 k 行是第 $k-1$ 行的子列）。

对角线取法：把这些逐行嵌套的子列排成一张表（第 k 行落在球 $B(y_k, \frac{1}{k})$ 中）：

$$\begin{array}{cccccc}
x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} & \cdots & \subset B(y_1, 1) \\
x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} & \cdots & \subset B\big(y_2, \frac{1}{2}\big) \\
x_1^{(3)} & x_2^{(3)} & x_3^{(3)} & x_4^{(3)} & \cdots & \subset B\big(y_3, \frac{1}{3}\big) \\
\vdots & \vdots & \vdots & \vdots & &
\end{array}$$

取对角线 $z_k := x_k^{(k)}$（第 k 行的第 k 项）。

下面验证 $\{z_k\}$ 是 Cauchy 子列。首先 $\{z_k\}$ 确实是原序列 $\{x_n\}$ 的一个子列，因为每一行都是上一行（进而是 $\{x_n\}$）的子列，逐行往下选取，下标严格递增。其次，对固定的 k，当 $k' \geq k$ 时，$\{x_n^{(k')}\}$ 是 $\{x_n^{(k)}\}$ 的子列，故 $z_{k'} = x_{k'}^{(k')}$ 这一项也在第 k 行中出现，从而 $z_{k'} \in B\big(y_k, \frac{1}{k}\big)$。于是对一切 $k', k'' \geq k$，

$$d(z_{k'}, z_{k''}) \leq d(z_{k'}, y_k) + d(y_k, z_{k''}) < \frac{1}{k} + \frac{1}{k} = \frac{2}{k} \xrightarrow[k\to\infty]{} 0,$$

即 $\{z_k\}$ 是 Cauchy 列。这就找到了 $\{x_n\}$ 的一个 Cauchy 子列。证毕。

命题 2（完备空间中，预紧 $\Rightarrow$ 列紧）：设 (X, d) 完备，$A \subset X$ 预紧，则 A 列紧。

证明：由引理 1，A 预紧 $\Rightarrow$ A 中任何序列都有 Cauchy 子列；又 X 完备，该 Cauchy 子列在 X 中收敛。故 A 中任何序列都有子列在 X 中收敛，即 A 列紧。证毕。

结合命题 1（任意空间中列紧 $\Rightarrow$ 预紧）与命题 2（完备空间中预紧 $\Rightarrow$ 列紧），可知：*在完备度量空间中，预紧与列紧是等价的*。命题 1 不需要完备性，而命题 2 的逆方向恰恰需要用到完备性把 Cauchy 子列”接住”变成收敛子列。

Lebesgue 数引理

命题 3（Lebesgue 数引理）：设 $A \subset (X, d)$ 且 A 自列紧，$\{O_i\}_{i\in I}$ 是 A 的一个开覆盖，则存在 $\lambda > 0$（称为这个开覆盖的 **Lebesgue 数**），使得

$$\forall x \in A,\ \exists i_x \in I,\ \text{s.t. } B(x, \lambda) \subset O_{i_x}.$$

证明：反证法。假设不存在这样的 λ，则 $\forall n \in \mathbb{N}^*$，$\exists x_n \in A$，使 $B\big(x_n, \frac{1}{n}\big)$ 不包含在任何一个 O_i 内。由 A 列紧，$\{x_n\}$ 存在子列 $\{x_{n_k}\} \to y \in A \subset \bigcup_{i\in I} O_i$，故存在 $i_0 \in I$，使 $y \in O_{i_0}$。由 O_{i_0} 开，存在 $r > 0$，使 $B(y, r) \subset O_{i_0}$。当 k 足够大时，$d(x_{n_k}, y) < \frac{r}{2}$ 且 $\frac{1}{n_k} < \frac{r}{2}$。于是由三角不等式，

$$B\Big(x_{n_k}, \frac{1}{n_k}\Big) \subset B(y, r) \subset O_{i_0},$$

这与 x_{n_k} 的取法（$B(x_{n_k}, \frac{1}{n_k})$ 不包含于任何一个 O_i）矛盾。故这样的 λ 存在。证毕。

紧性的三个等价刻画

定理：设 (X, d) 是度量空间，$A \subset X$。则下列条件等价：

(1) A 是紧集（开覆盖意义下）；

(2) A 是自列紧集；

(3) A 完备且预紧。

证明（*(3) ⇒ (2)*）：设 A 完备且预紧。由命题 2（此时把 A 自身看作完备空间），A 预紧 $\Rightarrow$ A 列紧。又 A 完备 $\Rightarrow$ A 是 X 中的闭集（这是”Cauchy 列与完备性” 一节的定理 1）。故 A 闭且列紧，即 A 自列紧。

证明（*(2) ⇒ (3)*）：设 A 自列紧。一方面，A 自列紧 $\Rightarrow$ A 列紧 $\Rightarrow$（命题 1）A 预紧。另一方面，上一小节已证：A 自列紧 $\Rightarrow$ (A,d) 完备。合起来即 A 完备且预紧。

证明（*(1) ⇒ (2)*）：设 A 紧（开覆盖意义下），$\{x_n\}\subset A$ 是任意序列。反证法：假设 $\{x_n\}$ 没有子列收敛到 A 中的点。则对每个 $y\in A$，存在 $r_y>0$，使 $B(y,r_y)$ 只包含 $\{x_n\}$ 中有限多项。

事实上，若对某个 $y\in A$，任意 $r>0$ 时 $B(y,r)$ 都含有 $\{x_n\}$ 中无穷多项，则可取 $r=\frac{1}{m}$ $(m=1,2,\dots)$，每次都能选到指标 $n_m>n_{m-1}$ 使 $x_{n_m}\in B(y,\frac{1}{m})$，从而 $x_{n_m}\to y\in A$，即 $\{x_n\}$ 有子列收敛到 A 中的点 y，与假设矛盾。

于是 $\bigcup_{y\in A}B(y,r_y)$ 是 A 的一个开覆盖。由 A 紧，存在有限子覆盖 $y_1,\dots,y_k$，使 $\{x_n\}\subset A\subset\bigcup_{i=1}^{k}B(y_i,r_{y_i})$。但每个 $B(y_i,r_{y_i})$ 只含 $\{x_n\}$ 中有限多项，有限个这样的球的并也只含有限多项，这与 $\{x_n\}$ 是无穷序列矛盾。故 $\{x_n\}$ 必有子列收敛到 A 中的点，即 A 自列紧。

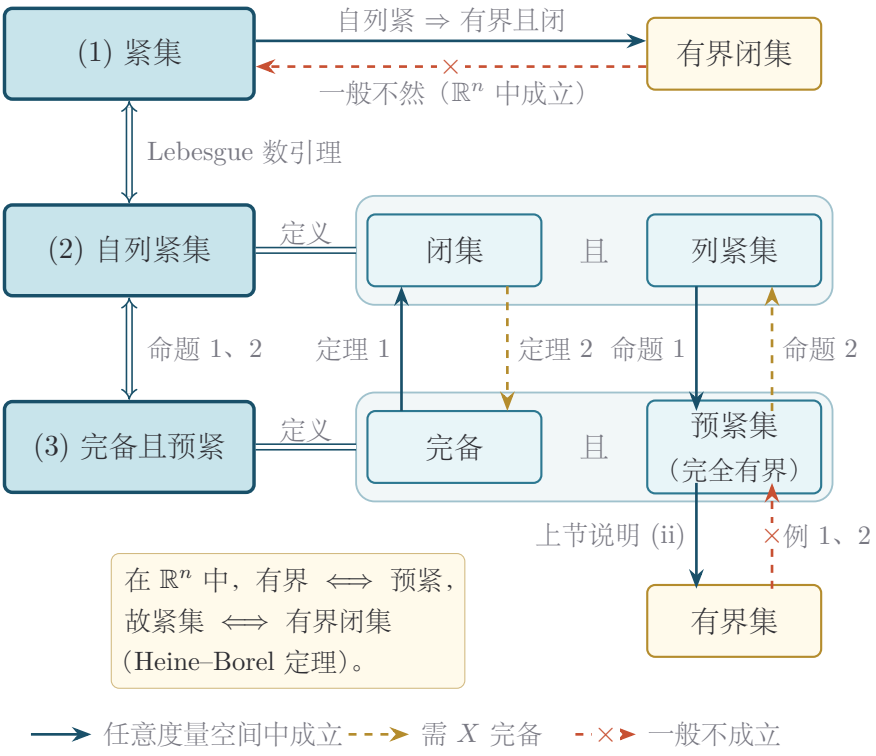
证明（*(2) ⇒ (1)*）：设 A 自列紧，$\{O_i\}_{i\in I}$ 是 A 的任意开覆盖。由命题 3（Lebesgue 数引理），取到 Lebesgue 数 $\lambda>0$。又 A 自列紧 $\Rightarrow$ A 列紧 $\Rightarrow$（命题 1）A 预紧，故可取 A 的有限 λ-网 $y_1,\dots,y_k$，使 $A\subset\bigcup_{p=1}^{k}B(y_p,\lambda)$。由 Lebesgue 数的定义，每个 $B(y_p,\lambda)$ 都包含于某个 O_{i_p}。故

$$A\subset\bigcup_{p=1}^{k}B(y_p,\lambda)\subset\bigcup_{p=1}^{k}O_{i_p},$$

即 $\{O_{i_1},\dots,O_{i_k}\}$ 是 $\{O_i\}_{i\in I}$ 的一个有限子覆盖。故 A 紧。证毕。

紧性相关概念的关系图

到这里，围绕” 紧性” 出现了好几个相互交织的概念：紧集、自列紧、列紧、预紧（完全有界）、完备、闭、有界。下面的图把它们之间的蕴含关系一次画清楚。实线箭头表示在任意度量空间中成立的蕴含，金色虚线箭头表示需要整个空间 X 完备才成立的蕴含，红色虚线加 $\times$ 的箭头表示一般不成立（有反例）的方向。双线箭头 $\Longleftrightarrow$ 与等号表示等价或定义。



图中” 定理 1、定理 2” 指”Cauchy 列与完备性” 一节中的定理 1（完备 ⇒ 闭）与定理 2（完备空间中闭 ⇔ 完备）；” 命题 1、2” 和” 例 1、2” 指本节；” 说明 (ii)” 指上一节中” 完全有界的三点说明” 的 (ii)。

读图提示：整张图可以这样理解。左边三个高亮的结点是本节的主要定理：紧集 ⟺ 自列紧 ⟺ 完备且预紧。中间的两个灰蓝色底板分别是自列紧与” 完备且预紧” 的定义拆开后的两个成分，竖直方向的箭头说明这两组成分一一对应：” 闭 ⇔ 完备” 和” 列紧 ⇔ 预紧” 都只有一个方向无条件成立，另一个方向要在完备空间里才成立。右侧的有界性始终是” 最弱” 的一环：它能由预紧、紧得到，但反过来一般不行；只有在 $\mathbb{R}^n$ 中，有界与预紧才重合。

Heine–Borel 定理：在 $\mathbb{R}^n$（欧氏度量）中，A 紧 ⟺ A 是有界闭集。

证明：($\Rightarrow$) 这就是图中最上面的实线箭头：紧即自列紧，而自列紧集有界且闭（本节前面的推论）。

($\Leftarrow$) 设 $A \subset \mathbb{R}^n$ 有界且闭。

第一步（有界 ⇒ 预紧）：这就是上一节说明 (iii) 中的” 网格” 想法。由 A 有界，存在 $M > 0$ 使 $A \subset [-M, M]^n$。任意给定 $\varepsilon > 0$，取正整数 N 满足 $\dfrac{2M\sqrt{n}}{N} < \varepsilon$，把每个坐标区间 $[-M, M]$ 等分成 N 段（每段长 $\dfrac{2M}{N}$），于是 $[-M, M]^n$ 被 N^n 个小闭方格覆盖，每个小方格的直径恰为 $\dfrac{2M\sqrt{n}}{N} < \varepsilon$。对每个与 A 相交的小方格，任取一点属于 A 与该方格的交集，得到 A 中的有限个点 $a_1, \dots, a_m$。对任意 $x \in A$，x 落在某个小方格中，而该方格与 A 相交，故取了点 a_j 属于同一个方格，于是 $d(x, a_j) \le$ 方格直径 $< \varepsilon$。所以 $\{a_1, \dots, a_m\}$ 是 A 的有限 ε-网，由 ε 任意，A 预紧。

第二步（闭 $\Rightarrow$ 完备）：$\mathbb{R}^n$ 完备，A 是完备空间的闭子集，由定理 2 知 A 完备。

于是 A 完备且预紧，由上面的三个等价刻画，A 紧。证毕。

这里再强调一次图中那条红色箭头：在一般的度量空间里，”有界闭集 $\Rightarrow$ 紧集”**不成立**，例如无限集上的离散度量（例 1）：整个空间有界且闭，却不是紧的。所以 Heine–Borel 定理依赖 $\mathbb{R}^n$ 的特殊结构（有限维），不是度量空间的一般事实。

1.10　紧集上的连续映射

注：拓扑性质

称 X（或其子集 A）的一条性质为**拓扑性质**，如果它的真假只由开集族 τ_d 决定。也就是说：对 X 上任意两个拓扑等价的度量 d, d'，这条性质在 (X, d) 中成立 $\iff$ 在 (X, d') 中成立。

回到上次课的”紧性的三个等价刻画”：

- (1) 中”紧”是拓扑性质；
- (2) 中”自列紧”是拓扑性质；
- (3) 中”完备”**不是**拓扑性质，”预紧”也**不是**拓扑性质。

例：$(\mathbb{R}, |\cdot|)$ 完备但不预紧；$(-1, 1)$ 预紧但不完备；而 $f(x) = \dfrac{x}{1+|x|}$ 给出这两个空间之间的同胚。

所以，”完备”与”预紧”各自都不是拓扑性质，但**合取之后**变成了拓扑性质——”完备且预紧”恰好等价于紧性。

具体地，设 d, d' 是 X 上拓扑等价的度量，即诱导相同的开集族（从而也有相同的闭集、相同的收敛序列，见”一致连续与同胚”一节）。”紧”的定义只涉及”开覆盖”与”有限子覆盖”，只用到开集，所以 A 在 d 下紧 $\iff$ 在 d' 下紧。”自列紧”的定义只涉及”闭”与”序列收敛到某个点”，而 d, d' 有相同的闭集、相同的收敛序列，所以 A 在 d 下自列紧 $\iff$ 在 d' 下自列紧。这与上次课证明的”紧 $\iff$ 自列紧”相互印证。

需要注意，”拓扑性质”比较的是**同一个集合**上的两个度量，所以上面的例子应这样理解。在 $X = \mathbb{R}$ 上取 $d(x, y) = |x - y|$ 与 $d'(x, y) = |f(x) - f(y)|$，其中 $f(x) = \dfrac{x}{1+|x|}$。因为 $f : \mathbb{R} \to (-1, 1)$ 是双射（见”一致连续与同胚”一节），d' 是 $\mathbb{R}$ 上的度量，并且 $f : (\mathbb{R}, d') \to (-1, 1)$ 是等距的双射。又 f 与 f^{-1} 都连续，所以恒同映射 $\mathrm{id} : (\mathbb{R}, d) \to (\mathbb{R}, d')$ 及其逆都连续，即 d, d' 拓扑等价。但是：

- $(\mathbb{R}, d)$ 完备，却不预紧（预紧蕴含有界，而 $\mathbb{R}$ 无界）；
- $(\mathbb{R}, d')$ 与 $(-1, 1)$ 等距，所以预紧（$(-1, 1)$ 是 $\mathbb{R}$ 中的有界集，而 $\mathbb{R}$ 中有界 $\iff$ 预

紧，见上一节），却不完备（$(-1,1)$ 不是完备空间 $\mathbb{R}$ 的闭子集，由”完备 $\Rightarrow$ 闭”知它不完备）。

这就说明完备性与预紧性都会随着度量的”拓扑等价替换”而改变。

紧集上连续映射的基本定理

设 (X,d_1)、(Y,d_2) 是度量空间。

定理：

(1) 紧集是有界闭集；紧集的闭子集是紧集。

(2) 若 $f:X\to Y$ 连续，$K\subset X$ 紧，则 $f(K)$ 紧。

(3) 若 X 紧，$f:X\to Y$ 连续，则 f 一致连续。

(4) 若 X 紧，$f:X\to\mathbb{R}$ 连续，则 f 有界，且取到最大值、最小值。

(5) 若 X 紧，$f:X\to Y$ 是连续双射，则 f 是同胚（即 f^{-1} 连续）。

证明 *(1)*：紧 $\Rightarrow$ 预紧 $\Rightarrow$ 有界；紧 $\Rightarrow$ 自列紧 $\Rightarrow$ 闭。故紧 $\Rightarrow$ 有界闭。

这里用到：紧 $\Longleftrightarrow$ 完备且预紧（所以紧 $\Rightarrow$ 预紧）；预紧 $\Rightarrow$ 有界（完全有界的说明(ii)）；紧 $\Longleftrightarrow$ 自列紧（所以紧 $\Rightarrow$ 自列紧），以及自列紧 $\Rightarrow$ 闭集（紧性与列紧性一节的推论）。

再设 K 紧，$F\subset K$ 是 X 中的闭集。K 紧 $\Rightarrow$ K 自列紧 $\Rightarrow$ F 列紧：事实上，任取 $\{x_n\}\subset F\subset K$，由 K 自列紧，存在子列 $x_{n_k}\to x\in K$，所以 F 中的任意序列在 X 中都有收敛子列（极限在 K 中）。又由于 F 闭，而 $\{x_{n_k}\}\subset F$，极限 x 属于 F 的闭包 $\overline{F}=F$。故 F 中任意序列都有子列收敛到 F 中的点，即 F 自列紧，从而 F 紧。

证明 *(2)*：设 $\{y_n\}\subset f(K)$，则存在 $\{x_n\}\subset K$ 使得 $y_n=f(x_n)$。由 K 紧 $\Rightarrow$ K 自列紧，存在子列 $x_{n_k}\to x\in K$。由 f 连续，$y_{n_k}=f(x_{n_k})\to f(x)\in f(K)$。所以 $f(K)$ 中任意序列都有子列收敛到 $f(K)$ 中的点，即 $f(K)$ 自列紧，从而紧。

证明 *(3)*：反设 f 不一致连续，则存在 $\varepsilon_0>0$，以及 $\{x_n\},\{y_n\}\subset X$，使得

$$d_1(x_n,y_n)<\frac{1}{n},\qquad \text{但}\qquad d_2\big(f(x_n),f(y_n)\big)\ge\varepsilon_0.$$

由 X 紧 $\Rightarrow$ X 自列紧，存在子列 $x_{n_k}\to x\in X$。则

$$d_1(y_{n_k},x)\le d_1(y_{n_k},x_{n_k})+d_1(x_{n_k},x)\to 0\quad(k\to\infty),$$

所以 $y_{n_k}\to x$。由 f 的连续性，$f(x_{n_k})\to f(x)$，$f(y_{n_k})\to f(x)$，故

$$d_2\big(f(x_{n_k}),f(y_{n_k})\big)\le d_2\big(f(x_{n_k}),f(x)\big)+d_2\big(f(x),f(y_{n_k})\big)\to 0,$$

这与 $d_2\big(f(x_{n_k}),f(y_{n_k})\big)\ge\varepsilon_0$ 矛盾。证毕。

此外也可以这样处理：在得到 $x_{n_k}\to x$ 之后，不用三角不等式，而是对 $\{y_{n_k}\}$ 再取收敛子列 $y_{n_{k_l}}\to y$。此时 $x_{n_{k_l}}\to x$，$y_{n_{k_l}}\to y$，而 $d_1\big(x_{n_{k_l}},y_{n_{k_l}}\big)<\dfrac{1}{n_{k_l}}\to 0$，由距离函数的连续性得 $d_1(x,y)=0$，即 $x=y$，之后的论证与上面完全相同。

证明 *(4)*：由 (2)，$f(X)$ 是 $\mathbb{R}$ 的紧子集，故由 (1) 有界，且是闭集。（设 X 非空。）由 $\mathbb{R}$ 中有界非空集合必有上、下确界，且上、下确界属于该集合的闭包，而 $f(X)$ 是闭集，所以 $f(X)$ 包含自己的上、下确界。即存在 $x_{\max},x_{\min}\in X$ 使得 $f(x_{\max})=\sup f(X)$，$f(x_{\min})=\inf f(X)$。故 $\min f$、$\max f$ 存在。

证明 *(5)*：连续映射 g 的”闭集刻画”是：g 连续 $\iff$ 闭集的原像是闭集。要证 $f^{-1}:Y\to X$ 连续，对任意 X 中闭集 F，$(f^{-1})^{-1}(F)=f(F)$（f 是双射），所以只需证明：f 把 X 中的闭集映为 Y 中的闭集。设 $F\subset X$ 闭，由 X 紧，据 (1)，F 紧；据 (2)，$f(F)$ 紧；再据 (1)，$f(F)$ 是 Y 中的闭集。证毕。

无限维中的新现象：“紧性丢失”

在有限维 $\mathbb{R}^n$ 中，紧 $\iff$ 有界闭。

在无限维中，这个等价不再成立。例如，在 $(C[0,1],d_\infty)$ 中，闭单位球不紧。

事实上，取”紧性与列紧性”一节中的函数列

$$f_n(x)=\begin{cases}1-nx, & 0\le x<\dfrac{1}{n},\\[2mm] 0, & \dfrac{1}{n}\le x\le 1,\end{cases}$$

它满足 $d_\infty(f_n,0)=\max|f_n|=1$，所以 $\{f_n\}\subset\overline{B}(0,1)$（闭单位球，0 是零函数）；并且当 $m>n$ 时 $d_\infty(f_m,f_n)=1-\dfrac{n}{m}$。若闭单位球是紧集，则它自列紧，$\{f_n\}$ 有收敛子列 $\{f_{n_k}\}$，收敛列必是 Cauchy 列。但对任意固定的 k，令 $l\to\infty$，$d_\infty(f_{n_l},f_{n_k})=1-\dfrac{n_k}{n_l}\to 1$，所以存在 $l>k$ 使 $d_\infty(f_{n_l},f_{n_k})>\dfrac{1}{2}$，这与 Cauchy 矛盾。故闭单位球不紧。

这正是”无限维空间中紧性丢失”的含义：有界闭集不再自动是紧集，紧集要比有界闭集苛刻得多。

相对紧集

定义（相对紧）：设 (X,d) 为度量空间，$A\subset X$。若 $\overline{A}$ 紧，则称 A **相对紧**。

推论：设 (X,d) 为完备度量空间，则

$$A\text{ 相对紧}\quad\Longleftrightarrow\quad A\text{ 预紧}.$$

证明：

($\Leftarrow$) A 预紧 $\Rightarrow$ $\overline{A}$ 预紧（预紧集的闭包预紧）；又 X 完备，所以 $\overline{A}$ 是完备空间的闭集，从而完备（完备空间中，闭 $\iff$ 完备）。完备且预紧，故 $\overline{A}$ 紧，即 A 相对紧。

($\Rightarrow$) A 相对紧 $\Rightarrow$ $\overline{A}$ 紧 $\Rightarrow$ $\overline{A}$ 预紧 $\Rightarrow$ A 预紧（预紧集的子集预紧，$A \subset \overline{A}$）。证毕。

这里用到：预紧性的遗传性（子集、闭包都保持预紧），" 完备空间中闭 $\iff$ 完备"（Cauchy 列与完备性一节的定理 2），以及" 紧 $\iff$ 完备且预紧"。($\Rightarrow$) 方向并不需要 X 完备；而 ($\Leftarrow$) 方向需要：例如在不完备的空间 $X = (0,1)$ 中，取 $A = X$，则 A 预紧（$(0,1) \subset \mathbb{R}$ 有界），但 $\overline{A} = (0,1)$ 在 X 中不是紧集（$\{1/n\}$ 没有在 $(0,1)$ 中收敛的子列），所以 A 不相对紧。

2

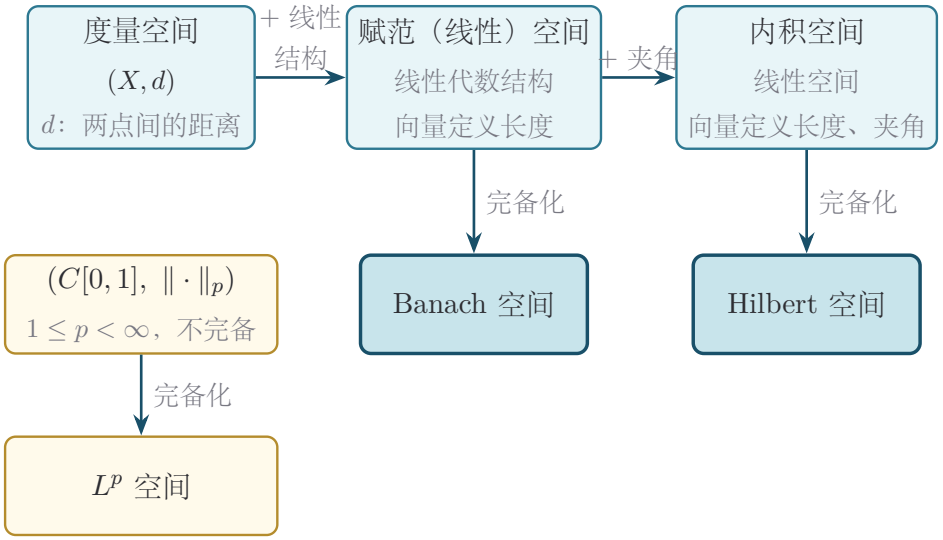
第二章

赋范空间和连续线性映射

2.1 Banach 空间

从度量空间到 Banach 空间与 Hilbert 空间

第一章里我们只用”两点间的距离” d 来研究空间。从这一章起，空间上多了 **线性代数结构**：可以做向量加法和数乘，于是可以用向量来定义”长度”，再进一步定义”夹角”。随着结构一层层增加，我们得到下面这张概念地图：



图中 $(C[0,1], \|\cdot\|_p)$ 指 $C[0,1]$ 配备 $\|f\|_p = \left(\int_0^1 |f(x)|^p\,dx\right)^{1/p}$ $(1 \le p < \infty)$。它不

完备，完备化之后得到 L^p 空间。不完备性可以这样直接验证：取

$$g_n(x)=\begin{cases}0, & 0\le x\le \dfrac{1}{2},\\ n\Big(x-\dfrac{1}{2}\Big), & \dfrac{1}{2}<x<\dfrac{1}{2}+\dfrac{1}{n},\\ 1, & \dfrac{1}{2}+\dfrac{1}{n}\le x\le 1,\end{cases}\qquad (n\ge 2).$$

当 $m>n$ 时，g_m-g_n 只在 $\left[\frac{1}{2},\frac{1}{2}+\frac{1}{n}\right]$ 上不为零，且 $|g_m-g_n|\le 1$，所以 $\|g_m-g_n\|_p^p\le \dfrac{1}{n}$，即 $\{g_n\}$ 是 Cauchy 列。若 $g_n\to g\in C[0,1]$（在 $\|\cdot\|_p$ 下），因为在 $\left[0,\frac{1}{2}\right]$ 上 $g_n\equiv 0$，

$$\int_0^{1/2}|g|^p\,dx=\int_0^{1/2}|g_n-g|^p\,dx\le \|g_n-g\|_p^p\to 0,$$

所以 $\int_0^{1/2}|g|^p=0$，由 g 连续知 $g\equiv 0$ 在 $\left[0,\frac{1}{2}\right]$ 上。又对任意 $\delta\in\left(0,\frac{1}{2}\right)$，当 $n\ge\dfrac{1}{\delta}$ 时 $g_n\equiv 1$ 在 $\left[\frac{1}{2}+\delta,1\right]$ 上，同理 $\displaystyle\int_{1/2+\delta}^{1}|g-1|^p=0$，所以 $g\equiv 1$ 在 $\left[\frac{1}{2}+\delta,1\right]$ 上。令 $\delta\to 0$ 得 $g\equiv 1$ 在 $\left(\frac{1}{2},1\right]$ 上，与 $g(1/2)=0$ 以及 g 在 $\frac{1}{2}$ 处连续矛盾。故 $\{g_n\}$ 在 $C[0,1]$ 中不收敛，$(C[0,1],\|\cdot\|_p)$ 不完备。

范数的定义

注：$\mathbb{K}=\mathbb{R}$ 或 $\mathbb{C}$，E 是数域 $\mathbb{K}$ 上的向量空间。

定义（赋范空间）：函数 $p:E\to[0,+\infty)$ 称为 E 上的 **范数**，若 $\forall x,y\in E$，$\lambda\in\mathbb{K}$，成立：

(1) 正定性：$p(x)\ge 0$，且 $p(x)=0\iff x=0_E$；

(2) 正齐性：$p(\lambda x)=|\lambda|\,p(x)$；

(3) 三角不等式：$p(x+y)\le p(x)+p(y)$。

范数通常记为 $\|\cdot\|$，$(E,\|\cdot\|)$ 或 (E,p) 称为**赋范空间**。只满足 (2)、(3) 的 p 称为**半范数**。

范数与度量的关系

注：范数给向量定义了”长度”。

1. 赋范空间是度量空间

在赋范空间 $(E,\|\cdot\|)$ 上令

$$d(x,y)=\|x-y\|\qquad (\text{验证}d\text{ 是度量}).$$

下面验证 d 是度量。**非负性与正定性**：$d(x,y)=\|x-y\|\geq 0$，且 $d(x,y)=0 \iff \|x-y\|=0 \iff x-y=0_E \iff x=y$（范数的 (1)）。**对称性**：$d(y,x)=\|y-x\|=\|(-1)(x-y)\|=|-1|\,\|x-y\|=d(x,y)$（范数的 (2)）。**三角不等式**：对任意 $x,y,z\in E$，

$$d(x,z)=\|x-z\|=\|(x-y)+(y-z)\|\leq\|x-y\|+\|y-z\|=d(x,y)+d(y,z)$$

（范数的 (3)）。故 d 是 E 上的度量，称为由范数**诱导**的度量。

于是赋范空间中有了开球、闭球、球面的概念：

$$\begin{aligned}\text{开球}\quad & B(x,r)=\{\,y\in E\mid\|y-x\|<r\,\},\\ \text{闭球}\quad & \overline{B}(x,r)=\{\,y\in E\mid\|y-x\|\leq r\,\},\\ \text{球面}\quad & S(x,r)=\{\,y\in E\mid\|y-x\|=r\,\}.\end{aligned}$$

2. 范数诱导的度量的两条额外性质

范数诱导的度量 $d(x,y)=\|x-y\|$ 有两条额外的性质：

(1) 平移不变性：$d(x,y)=d(x+z,y+z)$；

(2) 正齐次性：$d(\lambda x,\lambda y)=|\lambda|\,d(x,y)$。

命题：向量空间 E 上的度量 d 由某个范数诱导（即 $d(x,y)=\|x-y\|$）$\iff$ d 平移不变且齐次。

证明：

($\Rightarrow$) 若 $d(x,y)=\|x-y\|$，则 $d(x+z,y+z)=\|(x+z)-(y+z)\|=\|x-y\|=d(x,y)$，$d(\lambda x,\lambda y)=\|\lambda x-\lambda y\|=\|\lambda(x-y)\|=|\lambda|\,\|x-y\|=|\lambda|\,d(x,y)$。

($\Leftarrow$) 设 d 是 E 上平移不变且正齐次的度量。定义 $\|x\|:=d(x,0_E)$。验证它是范数：

(1) 因为 d 是度量，$\|x\|=d(x,0)\geq 0$，且 $\|x\|=0 \iff d(x,0)=0 \iff x=0_E$。

(2) $\|\lambda x\|=d(\lambda x,0)=d(\lambda x,\lambda\cdot 0)=|\lambda|\,d(x,0)=|\lambda|\,\|x\|$。

(3) 先由平移不变性（平移 $-y$）得 $\|x+y\|=d(x+y,0)=d(x,-y)$，再由度量的三角不等式，$d(x,-y)\leq d(x,0)+d(0,-y)$。而由正齐次性（取 $\lambda=-1$，注意 $-0=0$），$d(0,-y)=d\big((-1)\cdot 0,(-1)\cdot y\big)=|-1|\,d(0,y)=d(y,0)=\|y\|$。所以 $\|x+y\|\leq\|x\|+\|y\|$。

最后，这个范数诱导的度量就是 d：由平移不变性（平移 $-y$），$\|x-y\|=d(x-y,0)=d(x,y)$。证毕。

注：不来自范数的度量：

(i) $\mathbb{R}$ 上 $\rho(x,y)=\dfrac{|x-y|}{1+|x-y|}$（不正齐次）；

(ii) $\sigma(x,y)=|\arctan x-\arctan y|$。

事实上，由上面的命题，只需说明它们不满足” 平移不变且正齐次”。(i)：$\rho(0,1)=\dfrac{1}{2}$，而 $\rho(0,2)=\dfrac{2}{3}\neq 2\rho(0,1)=1$，所以不正齐次（它是平移不变的，因为只依赖 $|x-y|$）。(ii)：$\sigma(0,1)=\dfrac{\pi}{4}$，而平移 1 后 $\sigma(1,2)=\arctan 2-\dfrac{\pi}{4}\approx 0.322\neq\dfrac{\pi}{4}$，所以不平移不变（这里它也不正齐次：$\sigma(0,2)\approx 1.107\neq 2\sigma(0,1)$）。

依范数收敛

定义（依范数收敛）：$\{x_n\}\subset E$ 收敛到 x，记 $x_n\xrightarrow{\|\cdot\|}x$，指 $\|x_n-x\|\to 0$。
（也就是在由范数诱导的度量下收敛。）
E 中的单位球面和闭单位球分别为

$$S=\{x:\|x\|=1\},\qquad \overline{B}=\{x:\|x\|\le 1\}.$$

运算的连续性

命题（运算的连续性）：

(1) $\big|\,\|x\|-\|y\|\,\big|\le\|x-y\|$，即范数是 1-Lipschitz 的。特别地，$x_n\xrightarrow{\|\cdot\|}x\Rightarrow\|x_n\|\to\|x\|$。

(2) 加法 $(x,y)\mapsto x+y$ 与数乘 $(\lambda,x)\mapsto\lambda x$ 连续：当 $x_n\to x$，$y_n\to y$，$\lambda_n\to\lambda$ 时，有 $x_n+y_n\to x+y$，$\lambda_n x_n\to\lambda x$。

(3) E 的子空间 F 的闭包 $\overline{F}$ 仍是子空间；球 $B(x,r)$ 是凸集，且 $B(x,r)=x+rB(0,1)$。

证明 (1)：由三角不等式，$\|x\|=\|(x-y)+y\|\le\|x-y\|+\|y\|$，即 $\|x\|-\|y\|\le\|x-y\|$；交换 x,y 的位置得 $\|y\|-\|x\|\le\|y-x\|=\|x-y\|$。合起来 $\big|\|x\|-\|y\|\big|\le\|x-y\|$。于是 $x_n\to x$ 时 $\big|\|x_n\|-\|x\|\big|\le\|x_n-x\|\to 0$，即 $\|x_n\|\to\|x\|$。

证明 (2)：加法：$\|(x_n+y_n)-(x+y)\|\le\|x_n-x\|+\|y_n-y\|\to 0$。数乘：

$$\begin{aligned}\|\lambda_n x_n-\lambda x\|&\le\|\lambda_n x_n-\lambda x_n\|+\|\lambda x_n-\lambda x\|\\&=|\lambda_n-\lambda|\,\|x_n\|+|\lambda|\,\|x_n-x\|\to 0.\end{aligned}$$

最后一步需要 $\|x_n\|$ 有界，这由 (1) 保证：$\|x_n\|\to\|x\|$，收敛数列有界，所以 $|\lambda_n-\lambda|\,\|x_n\|\to 0$；另一项显然趋于 0。

证明 (3)：**子空间的闭包**：$0\in F\subset\overline{F}$，所以 $\overline{F}$ 非空。只需证：$\forall x,y\in\overline{F}$，$\lambda\in\mathbb{K}$，有 $x+\lambda y\in\overline{F}$。取 F 中序列 $x_n\to x$，$y_n\to y$。因为 F 是子空间，$x_n+\lambda y_n\in F$，且由 (2)，$x_n+\lambda y_n\to x+\lambda y$。所以 $x+\lambda y\in\overline{F}$。

球是凸集：设 $x_1,x_2\in B(x,r)$，$t\in[0,1]$，则

$$\begin{aligned}\|tx_1+(1-t)x_2-x\|&=\|t(x_1-x)+(1-t)(x_2-x)\|\\&\le t\|x_1-x\|+(1-t)\|x_2-x\|<r.\end{aligned}$$

这里最后的严格不等号成立，是因为 $\|x_1-x\|<r$，$\|x_2-x\|<r$，所以 $t\|x_1-x\|+(1-t)\|x_2-x\|\le\max\{\|x_1-x\|,\|x_2-x\|\}<r$，即 $tx_1+(1-t)x_2\in B(x,r)$。同样的计算把"$<r$" 换成"$\le r$"，就说明闭球 $\overline{B}(x,r)$ 也是凸集。

球的平移与伸缩：$y\in B(x,r)\iff\|y-x\|<r\iff\left\|\dfrac{y-x}{r}\right\|<1\iff\dfrac{y-x}{r}\in B(0,1)\iff y\in x+rB(0,1)$。故 $B(x,r)=x+rB(0,1)$。

赋范空间的例子

① $\mathbb{K}^n$ 上的欧氏范数、最大范数与和范数

设 $x=(x_1,\dots,x_n)\in\mathbb{K}^n$，

$$\|x\|_2=\Big(\sum_{k=1}^{n}|x_k|^2\Big)^{1/2},\qquad \|x\|_\infty=\max_{1\le k\le n}|x_k|,\qquad \|x\|_1=\sum_{k=1}^{n}|x_k|.$$

一般地，$\forall\,1\le p<\infty$，

$$\|x\|_p=\Big(\sum_{k=1}^{n}|x_k|^p\Big)^{1/p}.$$

三角不等式对 p-范数就是 **Minkowski 不等式**。

正定性与正齐性对这些范数都是显然的（例如 $\|\lambda x\|_p=\big(\sum|\lambda|^p|x_k|^p\big)^{1/p}=|\lambda|\,\|x\|_p$，且 $\|x\|_p=0$ 当且仅当每个 $x_k=0$）。三角不等式：对 $\|\cdot\|_1$ 与 $\|\cdot\|_\infty$，直接由 $\mathbb{K}$ 中的三角不等式 $|x_k+y_k|\le|x_k|+|y_k|$ 得到；对一般的 p 这正是 Minkowski 不等式 $\Big(\sum_k|x_k+y_k|^p\Big)^{1/p}\le\Big(\sum_k|x_k|^p\Big)^{1/p}+\Big(\sum_k|y_k|^p\Big)^{1/p}$，这里直接引用该不等式，不重复证明。

② $C[0,1]$ 上的一致范数与 L^p 范数

在 $C[0,1]$（或紧集 $\Omega\subset\mathbb{K}^n$ 上的 $C(\Omega)$）上有一致范数与 L^p 范数：

$$\|f\|_\infty=\max_{x\in[0,1]}|f(x)|,\qquad \|f\|_p=\Big(\int_0^1|f(x)|^p\,dx\Big)^{1/p}\quad(1\le p<\infty).$$

它们诱导的度量分别是 d_∞ 和（$p=1$ 时）d_1。

$\|f\|_\infty$ 中的"max" 之所以存在，是因为 $|f|$ 是紧集 $[0,1]$（或 Ω）上的连续函数，由上一节紧集上连续映射的定理 (4) 知它取得最大值。对 $\|f\|_p$：非负性与正齐性显然；三角不等式是积分形式的 Minkowski 不等式；正定性需要用到 f 的连续性：若 $\|f\|_p=0$ 但 $f(x_0)\neq0$，则由连续性存在 x_0 的邻域，使 $|f|>\dfrac{|f(x_0)|}{2}$ 在该邻域（与 $[0,1]$ 的交上）成立，于是 $\int_0^1|f|^p>0$，矛盾；所以 $f\equiv0$。（在更大的函数类上，例如允许在零测集上取不同的值，$\|f\|_p=0$ 不能推出 $f\equiv0$，这也是后面 L^p 空间要把几乎处处相等的函数视为同一个元素的原因。）

③ $M_n(\mathbb{K})$ 上的矩阵范数（算子范数）

$M_n(\mathbb{K})$ 是 $n\times n$ 矩阵全体，把矩阵看成 $\mathbb{K}^n$ 上的线性算子。取定 $\mathbb{K}^n$ 上的一个范数 $\|\cdot\|$，令

$$\|A\| := \sup_{\|x\|=1} \|Ax\|,$$

称为**矩阵范数**（**算子范数**）。

要说明它是范数，首先要说明 $\|A\| < +\infty$。这依赖于有限维空间上线性映射必连续（等价地，$\mathbb{K}^n$ 上的所有范数等价）这个事实，后续会证明，这里先认定。在 $\|A\| < \infty$ 的前提下：

(1) 非负性显然。若 $A \neq 0$，存在 x_0 使 $Ax_0 \neq 0$（此时 $x_0 \neq 0$），令 $x = \dfrac{x_0}{\|x_0\|}$，则 $\|x\| = 1$ 且 $\|Ax\| = \dfrac{\|Ax_0\|}{\|x_0\|} > 0$，所以 $\|A\| > 0$；反过来 $A = 0$ 时 $\|A\| = 0$。

(2) $\|\lambda A\| = \sup_{\|x\|=1} \|\lambda Ax\| = |\lambda| \sup_{\|x\|=1} \|Ax\| = |\lambda|\,\|A\|$。

(3) 对任意 $\|x\| = 1$，$\|(A+B)x\| \le \|Ax\| + \|Bx\| \le \|A\| + \|B\|$，对 x 取上确界得 $\|A+B\| \le \|A\| + \|B\|$。

注意矩阵范数依赖于 $\mathbb{K}^n$ 上取定的范数。

④序列空间

$\forall\, 1 \le p < \infty$，

$$\ell^p = \Big\{ x = (x_k)_{k\ge 1} : \sum_{k=1}^{\infty} |x_k|^p < \infty \Big\}, \qquad \|x\|_p = \Big(\sum_{k=1}^{\infty} |x_k|^p \Big)^{1/p}.$$

ℓ^∞ 是有界序列全体，赋范数 $\|x\|_\infty = \sup_{k\ge 1} |x_k|$。此外

$$c_0 = \{x : x_k \to 0\} \subset \ell^\infty, \qquad c_{00} = \text{只有有限项非零的序列全体},$$

并且（$1 \le p < \infty$）

$$c_{00} \subset \ell^p \subset c_0 \subset \ell^\infty.$$

这些包含关系可逐一验证：

- $c_{00} \subset \ell^p$：只有有限项非零，所以级数 $\sum |x_k|^p$ 是有限和。
- $\ell^p \subset c_0$：$\sum |x_k|^p < \infty$ 推出通项 $|x_k|^p \to 0$，即 $x_k \to 0$。
- $c_0 \subset \ell^\infty$：收敛数列必有界。

它们都是真包含：$(2^{-k})_k \in \ell^p \setminus c_{00}$；$(k^{-1/p})_k \in c_0$（趋于 0），但 $\sum (k^{-1/p})^p = \sum \dfrac{1}{k} = +\infty$，所以 $\notin \ell^p$；$(1,1,1,\dots) \in \ell^\infty \setminus c_0$。

$\|\cdot\|_p$ 确是 ℓ^p 上的范数：正定性、正齐性显然。对任意 N，由 $\mathbb{K}^N$ 上的 Minkowski 不等式，

$$\Big(\sum_{k=1}^{N}|x_k+y_k|^p\Big)^{1/p}\le\Big(\sum_{k=1}^{N}|x_k|^p\Big)^{1/p}+\Big(\sum_{k=1}^{N}|y_k|^p\Big)^{1/p}\le\|x\|_p+\|y\|_p.$$

令 $N\to\infty$，得 $x+y\in\ell^p$（所以 ℓ^p 是向量空间）以及 $\|x+y\|_p\le\|x\|_p+\|y\|_p$。$\ell^\infty, c_0, c_{00}$ 上的 sup 范数的三角不等式则由 $|x_k+y_k|\le|x_k|+|y_k|\le\|x\|_\infty+\|y\|_\infty$ 逐项取上确界得到。

⑤乘积空间

设 E,F 是赋范空间。在 $E\times F$ 上可以定义范数

$$\|(x,y)\|=\|x\|_E+\|y\|_F\quad\text{或}\quad\max\{\|x\|_E,\|y\|_F\}\quad\text{或}\quad\big(\|x\|_E^2+\|y\|_F^2\big)^{1/2},$$

这三个范数是等价的。

下面验证它们是范数并说明等价性。$E\times F$ 的向量运算是逐分量的。记 $N(a,b)$ 为 $\mathbb{R}^2$ 上的 ℓ^1、ℓ^∞ 或 ℓ^2 范数（$a,b\ge0$ 时分别是 $a+b$、$\max\{a,b\}$、$\sqrt{a^2+b^2}$），则上面三个函数都形如 $\|(x,y)\|=N(\|x\|_E,\|y\|_F)$。$N$ 是 $\mathbb{R}^2$ 上的范数，且对 $a,b\ge0$ 关于 a,b 单调不减。于是：正定性来自 $N(a,b)=0\iff a=b=0\iff x=0,y=0$；正齐性 $\|(\lambda x,\lambda y)\|=N(|\lambda|\|x\|_E,|\lambda|\|y\|_F)=|\lambda|\,\|(x,y)\|$；三角不等式：

$$\begin{aligned}\|(x+x',y+y')\|&=N(\|x+x'\|_E,\|y+y'\|_F)\\&\le N(\|x\|_E+\|x'\|_E,\ \|y\|_F+\|y'\|_F)\\&\le N(\|x\|_E,\|y\|_F)+N(\|x'\|_E,\|y'\|_F),\end{aligned}$$

其中第一个不等号用单调性，第二个用 N 自身的三角不等式。关于”等价”：记 $a=\|x\|_E$，$b=\|y\|_F$，则

$$\max\{a,b\}\le\sqrt{a^2+b^2}\le a+b\le2\max\{a,b\},$$

（中间一步由 $a^2+b^2\le(a+b)^2$ 得到），所以每个范数都被另一个的常数倍控制，这就是三个范数”等价”的含义。

序列空间与 L^p 空间的例子和范数一览

上面几个空间里，每个元素是什么样子、范数怎么算，这里集中举例说明。

各空间的范数

空间	元素	范数
c_{00}	只有有限项非零的序列	$\|x\|_\infty = \max_k \|x_k\|$（有限多项取最大）
c_0	趋于 0 的序列	$\|x\|_\infty = \sup_k \|x_k\|$
ℓ^p（$1 \le p < \infty$）	$\sum \|x_k\|^p < \infty$ 的序列	$\|x\|_p = \Big(\sum_{k=1}^{\infty} \|x_k\|^p\Big)^{1/p}$
ℓ^∞	有界序列	$\|x\|_\infty = \sup_k \|x_k\|$
$L^p[0,1]$（$1 \le p < \infty$）	$\int_0^1 \|f\|^p < \infty$ 的函数	$\|f\|_p = \Big(\int_0^1 \|f(x)\|^p\,dx\Big)^{1/p}$

关于上表，有以下几点说明。

- c_{00} 和 c_0 都是 ℓ^∞ 的子空间，所以带的是从 ℓ^∞ 继承来的 sup 范数（”Cauchy 列与完备性” 一节中 c_{00} 配的正是 sup 度量）。由于 $c_{00} \subset \ell^p$，c_{00} 也可以改配 $\|\cdot\|_p$，这是同一个向量空间上的另一个范数。
- c_{00} 在 $\|\cdot\|_\infty$ 下的完备化是 c_0，在 $\|\cdot\|_p$ 下的完备化是 ℓ^p。原因是 c_{00} 在两者中都是稠密的：对 $x \in c_0$，截断 $x^{(N)} = (x_1, \dots, x_N, 0, 0, \dots) \in c_{00}$ 满足 $\|x - x^{(N)}\|_\infty = \sup_{k>N} |x_k| \to 0$；对 $x \in \ell^p$，$\|x - x^{(N)}\|_p^p = \sum_{k>N} |x_k|^p \to 0$（收敛级数的尾项趋于 0）。再加上 c_0 完备（见下一节）、ℓ^p 完备（后证），就得到完备化。
- ℓ^p 里是*序列*、范数是*求和*；L^p 里是*函数*、范数是*积分*，两者是平行的（ℓ^p 可看成对计数测度的 L^p）。
- L^p 的精确定义（后续课程会严格建立）：$L^p[0,1]$ 由 $[0,1]$ 上满足 $\int_0^1 |f|^p\,dx < \infty$ 的（Lebesgue）可测函数组成，并且把几乎处处相等的两个函数视为同一个元素；这样 $\|f\|_p = 0 \iff f = 0$ 几乎处处，才满足范数的正定性。$L^\infty[0,1]$ 是几乎处处有界的可测函数（同样按几乎处处相等等同），范数是本性上确界 $\|f\|_\infty = \operatorname{ess\,sup} |f| = \inf\{M : |f| \le M \text{ 几乎处处}\}$。$L^p[0,1]$ 正是 $(C[0,1], \|\cdot\|_p)$ 的完备化。

c_{00} 的例子

- $e_k = (0, \dots, 0, \underbrace{1}_{\text{第}k\text{ 位}}, 0, \dots)$；
- $x = (1, 2, 3, 0, 0, \dots)$：$\|x\|_\infty = 3$（若看作 ℓ^p 中的元素，$\|x\|_1 = 6$，$\|x\|_2 = \sqrt{14}$）；
- $x^{(n)} = \big(1, \frac{1}{2}, \dots, \frac{1}{n}, 0, \dots\big)$（前面证明 c_{00} 不完备时用的 Cauchy 列）。

ℓ^p 的例子

- $x_k = 2^{-k}$：$\sum 2^{-k} = 1$，所以 $\|x\|_1 = 1$；$\sum 4^{-k} = \dfrac{1}{3}$，所以 $\|x\|_2 = \dfrac{1}{\sqrt{3}}$；$\|x\|_\infty = \dfrac{1}{2}$。它属于每个 ℓ^p，但不在 c_{00} 中。
- $x_k = \dfrac{1}{k}$：$\sum \dfrac{1}{k} = +\infty$，所以 $x \notin \ell^1$；但 $\sum \dfrac{1}{k^2} = \dfrac{\pi^2}{6}$，所以 $x \in \ell^2$，$\|x\|_2 = \dfrac{\pi}{\sqrt{6}}$。一般地，$x_k = k^{-\alpha}$（$\alpha > 0$）属于 $\ell^p \iff \alpha p > 1$（p-级数收敛的条件）。
- 这说明 $\ell^1 \subsetneq \ell^2$（更一般地 $\ell^p \subsetneq \ell^q$，$p < q$）。事实上，包含关系 $\ell^p \subset \ell^q$（$p < q$）成立：$x \in \ell^p$ 时 $x_k \to 0$，所以存在 K，使 $k \geq K$ 时 $|x_k| \leq 1$，从而 $|x_k|^q \leq |x_k|^p$，故 $\sum_{k \geq K} |x_k|^q \leq \sum_{k \geq K} |x_k|^p < \infty$；前面有限项的和当然有限，所以 $x \in \ell^q$。

c_0 的例子

- $x_k = \dfrac{1}{k}$，$\dfrac{1}{\sqrt{k}}$（趋于 0）；其中 $\dfrac{1}{\sqrt{k}} \in c_0$ 但 $\notin \ell^2$（$\sum \dfrac{1}{k} = \infty$），$\notin \ell^1$。
- $x_k = \dfrac{1}{\ln(k+1)} \to 0$：它在 c_0 中，但不属于任何 ℓ^p。事实上，对任意 p，当 k 充分大时 $(\ln(k+1))^p \leq k$，所以 $\dfrac{1}{(\ln(k+1))^p} \geq \dfrac{1}{k}$，$\sum \dfrac{1}{(\ln(k+1))^p}$ 发散。
- 不在 c_0 中的有界序列：$(1, 1, 1, \dots)$，$((-1)^k)_k$（它们属于 ℓ^∞，但不趋于 0）。

$L^p[0,1]$ 的例子

- 任何连续函数。因为 $|f| \leq \|f\|_\infty$，$\int_0^1 |f|^p \leq \|f\|_\infty^p < \infty$，所以 $C[0,1] \subset L^p[0,1]$，并且这时 L^p 的范数就是前面 $C[0,1]$ 上的 $\|\cdot\|_p$。例如 $f \equiv 1$：$\|f\|_p = 1$；$f(x) = x$：$\|f\|_p = \left(\dfrac{1}{p+1}\right)^{1/p}$。
- 不连续的函数：$f = \mathbf{1}_{[0,1/2]}$（$[0, \frac{1}{2}]$ 上取 1，其余取 0），$\|f\|_p = \left(\dfrac{1}{2}\right)^{1/p} = 2^{-1/p}$。
- 无界的函数：$f(x) = x^{-\alpha}$（$0 < x \leq 1$，$\alpha > 0$）。$\displaystyle\int_0^1 x^{-\alpha p}\,dx$ 在 $\alpha p < 1$ 时等于 $\dfrac{1}{1-\alpha p}$，在 $\alpha p \geq 1$ 时发散，所以 $f \in L^p \iff \alpha < \dfrac{1}{p}$，此时 $\|f\|_p = (1-\alpha p)^{-1/p}$。它不连续，甚至无界，所以不在 $C[0,1]$ 中，也不在 L^∞ 中。取 $p < q$ 与 $\dfrac{1}{q} \leq \alpha < \dfrac{1}{p}$，则 $f \in L^p \setminus L^q$。
- 几乎处处等于零的函数：$\mathbf{1}_{\mathbb{Q} \cap [0,1]}$（有理点上取 1）。它只在可数个点上非零，$\int |f|^p = 0$，所以在 L^p 里就是零元素——这正是”几乎处处相等视为同一元素”的意义。

2.2 Banach 空间与级数

Banach 空间的定义

定义（Banach 空间）：赋范空间 $(E, \|\cdot\|)$ 若在范数诱导下的度量完备，则称为 Banach 空间。

例子

（下面打 ✓ 的表示已证。）

完备的例子

① $\mathbb{K}^n$ 在任何范数下完备（后续会证明）；在 p-范数下完备（✓，由 $\mathbb{K}$ 的完备性）。

② $(C[0,1], \|\cdot\|_\infty)$ ✓。

③ ℓ^p、L^p（后证），$1 \le p \le \infty$；c_0 ✓。

先证①中 p-范数下的完备性。设 $\{x^{(m)}\}$ 是 $(\mathbb{K}^n, \|\cdot\|_p)$ 中的 Cauchy 列，$1 \le p \le \infty$。对每个分量 k，$|x_k^{(m)} - x_k^{(l)}| \le \|x^{(m)} - x^{(l)}\|_p$，所以 $\{x_k^{(m)}\}_m$ 是 $\mathbb{K}$ 中的 Cauchy 列，由 $\mathbb{K}$ 完备，收敛到某个 $x_k \in \mathbb{K}$。令 $x = (x_1, \dots, x_n)$，则

$$\|x^{(m)} - x\|_p \le n^{1/p} \max_{1 \le k \le n} |x_k^{(m)} - x_k| \to 0$$

（对 $p = \infty$ 取 $n^{1/p} = 1$），故 $x^{(m)} \to x$，$(\mathbb{K}^n, \|\cdot\|_p)$ 完备。这里用到了 $\|y\|_p \le n^{1/p}\|y\|_\infty$：因为 $\sum_k |y_k|^p \le n \max_k |y_k|^p$。

再看③中 c_0 的完备性，ℓ^∞ 同理。设 $\{x^{(n)}\}$ 是 c_0 中（sup 范数下）的 Cauchy 列。对每个 k，$|x_k^{(n)} - x_k^{(m)}| \le \|x^{(n)} - x^{(m)}\|_\infty$，所以 $\{x_k^{(n)}\}_n$ 是 $\mathbb{K}$ 中的 Cauchy 列，收敛到某个 x_k。令 $x = (x_k)_{k \ge 1}$。任给 $\varepsilon > 0$，取 N 使 $\forall n, m \ge N$ 有 $\|x^{(n)} - x^{(m)}\|_\infty < \varepsilon$，即 $\forall k$，$|x_k^{(n)} - x_k^{(m)}| < \varepsilon$。令 $m \to \infty$，得

$$|x_k^{(n)} - x_k| \le \varepsilon \qquad (\forall k,\ \forall n \ge N).$$

再因为 $x^{(N)} \in c_0$，存在 K 使 $k \ge K$ 时 $|x_k^{(N)}| < \varepsilon$，于是 $|x_k| \le |x_k^{(N)}| + |x_k - x_k^{(N)}| < 2\varepsilon$ $(k \ge K)$，故 $x_k \to 0$，$x \in c_0$；并且上式给出 $\|x^{(n)} - x\|_\infty \le \varepsilon$ $(n \ge N)$，即 $x^{(n)} \to x$。所以 c_0 完备。对 ℓ^∞：把"$x^{(N)} \in c_0$"换成"$x^{(N)}$ 有界"，同样得到 $|x_k| \le \|x^{(N)}\|_\infty + \varepsilon$，从而 x 有界，论证其余部分不变。

不完备的例子

① c_{00}。

② $(C[0,1],\|\cdot\|_p)$，其中 $p=1$ 时 ✓；$1<p<\infty$ 时同理。

③ $\mathbb{R}[x]$ 在 $\|\cdot\|_\infty$ 下，作为 $(C[0,1],\|\cdot\|_\infty)$ 的子空间。

（$\mathbb{R}[x]$：以 x 为变量的实多项式全体。）

其中 c_{00} 不完备在"Cauchy 列与完备性" 一节已证（取 $x^{(n)}=(1,\frac{1}{2},\dots,\frac{1}{n},0,\dots)$）；$(C[0,1],\|\cdot\|_1)$ 不完备也已在那里证明。对一般的 $1\le p<\infty$，本章开头" 概念地图" 之后给出了统一的证明（取阶梯式逼近的函数列 g_n），此处不再重复。

对③，把多项式看成 $[0,1]$ 上的函数（非零多项式在 $[0,1]$ 上只有有限个零点，所以这是 $\mathbb{R}[x]$ 到 $C[0,1]$ 的单射），$\|p\|_\infty=\max_{[0,1]}|p|$。取 $p_n(x)=\sum_{k=0}^{n}\dfrac{x^k}{k!}$，则 $p_n\in\mathbb{R}[x]$，并且对 $x\in[0,1]$，$m>n$：

$$|p_m(x)-p_n(x)|\le\sum_{k=n+1}^{m}\frac{1}{k!}\le\sum_{k>n}\frac{1}{k!}\xrightarrow{n\to\infty}0,$$

（因为 $\sum 1/k!$ 收敛），所以 $\{p_n\}$ 是 $(\mathbb{R}[x],\|\cdot\|_\infty)$ 中的 Cauchy 列；同时 $|e^x-p_n(x)|\le\sum_{k>n}\dfrac{1}{k!}$，所以 $p_n\to e^x$ 在 $C[0,1]$ 的 $\|\cdot\|_\infty$ 下成立。若 $\mathbb{R}[x]$ 完备，则存在多项式 q 使 $p_n\to q$（在 $\|\cdot\|_\infty$ 下）。极限唯一，故在 $[0,1]$ 上 $q(x)=e^x$。设 $\deg q=m$，则 $(\deg q+1)$ 阶导数 $q^{(m+1)}\equiv0$，而 $(e^x)^{(m+1)}=e^x\neq0$，矛盾。故 $(\mathbb{R}[x],\|\cdot\|_\infty)$ 不完备。

这也与" 完备 $\Rightarrow$ 闭" 一致：$\mathbb{R}[x]$ 不是 $C[0,1]$ 的闭子集（e^x 属于闭包却不在 $\mathbb{R}[x]$ 里）。

级数

定义（级数）：设 $\{x_n\}\subset E$。称

$$S_n=\sum_{i=1}^{n}x_i\qquad(n\in\mathbb{N}_+)$$

为级数 $\sum_{n=1}^{\infty}x_n$ 的**部分和**（partial sum）。

(1) 若 $S_n\xrightarrow{\|\cdot\|}S\in E$，则称级数**收敛**，$S$ 为其和，记 $S=\sum_{n=1}^{\infty}x_n$。

(2) 若 $\{S_n\}$ 是 Cauchy 列，称级数为 **Cauchy 级数**。

(3) 若数项级数 $\sum_{n=1}^{\infty}\|x_n\|$ 收敛，称级数 **绝对收敛**。

注：

(a) 若 $S_n\xrightarrow{\|\cdot\|}S$，则 $x_n=S_n-S_{n-1}\to0$（级数收敛的必要条件）。这是因为 $S_{n-1}\to S$ 也成立（子列的极限），所以 $\|x_n\|=\|S_n-S_{n-1}\|\le\|S_n-S\|+\|S-S_{n-1}\|\to0$。

(b) 绝对收敛 $\Rightarrow$ Cauchy 级数。事实上，对任意 $n, p \in \mathbb{N}_+$，

$$\|S_{n+p} - S_n\| = \Big\| \sum_{k=n+1}^{n+p} x_k \Big\| \le \sum_{k=n+1}^{n+p} \|x_k\| \xrightarrow{n\to\infty} 0.$$

这里右端的"$\to 0$" 与 p 无关：因为 $\sum \|x_k\|$ 收敛，它的尾项 $\sum_{k>n} \|x_k\| \to 0$，而对任意 p，$\sum_{k=n+1}^{n+p} \|x_k\| \le \sum_{k>n} \|x_k\|$。所以 $\sup_p \|S_{n+p} - S_n\| \to 0$，这正是 Cauchy 列的定义。

(c) 在 $\mathbb{R}$ 中，绝对收敛 $\Rightarrow$ 收敛。这依赖于 $\mathbb{R}$ 的完备性，它其实就是下面" 完备性的级数刻画" 定理 ($\Rightarrow$) 方向用在 $E = \mathbb{R}$ 上的特例。*在一般的赋范空间中不对。*

例：绝对收敛但不收敛的级数

在 $(C[0,1], \|\cdot\|_1)$ 中取

$$f_n(x) = \begin{cases} 0, & 0 \le x \le \dfrac{1}{2}, \\ 2^n\Big(x - \dfrac{1}{2}\Big), & \dfrac{1}{2} \le x < \dfrac{1}{2} + \dfrac{1}{2^n}, \\ 1, & \dfrac{1}{2} + \dfrac{1}{2^n} \le x \le 1. \end{cases}$$

令 $x_1 = f_1$，$x_n = f_n - f_{n-1}$（$n \ge 2$）。则级数 $\displaystyle\sum_{n=1}^{\infty} x_n$ 的部分和是（裂项相消）

$$\sum_{k=1}^{n} x_k = f_1 + (f_2 - f_1) + (f_3 - f_2) + \cdots + (f_n - f_{n-1}) = f_n.$$

Claim：$\displaystyle\sum_{n=1}^{\infty} x_n$ 绝对收敛，并且 $f_n \to \chi_{(1/2,1]}$。

即

$$\sum_{n=1}^{\infty} \|x_n\|_1 = \|f_1\|_1 + \sum_{n=1}^{\infty} \|f_{n+1} - f_n\|_1 = \|f_1\|_1 + \sum_{n=1}^{\infty} \frac{1}{2^{n+2}} < \infty,$$

其中（$f_{n+1} - f_n$ 的图像与横轴围成的三角形的面积）

$$\|f_{n+1} - f_n\|_1 = \frac{1}{2}\Big(\frac{1}{2^n} - \frac{1}{2^{n+1}}\Big) = \frac{1}{2^{n+2}}.$$

上面的计算验证如下。首先 f_n 连续：在 $x = \frac{1}{2}$ 处两段都等于 0，在 $x = \frac{1}{2} + \frac{1}{2^n}$ 处两段都等于 1。其次，$f_{n+1} \ge f_n$ 处处成立：在 $\big[\frac{1}{2}, \frac{1}{2} + \frac{1}{2^{n+1}}\big]$ 上 $f_{n+1} = 2^{n+1}(x - \frac{1}{2}) \ge 2^n(x - \frac{1}{2}) = f_n$；在 $\big[\frac{1}{2} + \frac{1}{2^{n+1}}, \frac{1}{2} + \frac{1}{2^n}\big]$ 上 $f_{n+1} = 1 \ge f_n$；其余地方两者相等（都是 0 或都是 1）。所以 $\|f_{n+1} - f_n\|_1 = \int_0^1 f_{n+1} - \int_0^1 f_n$。又

$$\int_0^1 f_n = \int_{1/2}^{1/2+2^{-n}} 2^n\Big(x - \frac{1}{2}\Big)dx + \Big(\frac{1}{2} - \frac{1}{2^n}\Big) = \frac{1}{2^{n+1}} + \frac{1}{2} - \frac{1}{2^n} = \frac{1}{2} - \frac{1}{2^{n+1}},$$

于是 $\|f_{n+1}-f_n\|_1=\left(\frac12-\frac{1}{2^{n+2}}\right)-\left(\frac12-\frac{1}{2^{n+1}}\right)=\frac{1}{2^{n+2}}$，与上面的三角形面积一致。同时 $\|f_1\|_1=\frac14$，所以 $\sum\|x_n\|_1=\frac14+\sum_{n\ge1}\frac{1}{2^{n+2}}=\frac14+\frac14=\frac12<\infty$。

下面说明这个级数在 $C[0,1]$ 中不收敛。若级数收敛，则部分和 f_n 在 $\|\cdot\|_1$ 下收敛到某个 $g\in C[0,1]$。因为 $f_n\equiv0$ 在 $\left[0,\frac12\right]$ 上，

$$\int_0^{1/2}|g|\,dx=\int_0^{1/2}|f_n-g|\,dx\le\|f_n-g\|_1\to0,$$

所以 $g\equiv0$ 在 $\left[0,\frac12\right]$ 上（g 连续、$|g|\ge0$）。又对任意 $\delta\in\left(0,\frac12\right)$，当 $2^{-n}\le\delta$ 时 $f_n\equiv1$ 在 $\left[\frac12+\delta,1\right]$ 上，同理 $\int_{1/2+\delta}^1|1-g|\to0$，所以 $g\equiv1$ 在 $\left[\frac12+\delta,1\right]$ 上。由 δ 的任意性，$g\equiv1$ 在 $\left(\frac12,1\right]$ 上，与 $g\left(\frac12\right)=0$ 以及 g 在 $\frac12$ 处连续矛盾。所以级数在 $(C[0,1],\|\cdot\|_1)$ 中不收敛（”极限”只能是不连续的 $\chi_{(1/2,1]}$）。这就说明在一般的赋范空间里，绝对收敛不蕴含收敛。

完备性的级数刻画

定理（完备性的级数刻画）：赋范空间 $(E,\|\cdot\|)$ 完备 $\iff$ E 中每个绝对收敛级数都收敛。

证明：

($\Rightarrow$) 设 $\{x_n\}\subset E$，且级数 $\sum_{n=1}^{\infty}x_n$ 绝对收敛。由定义，对 $n,p\in\mathbb{N}_+$，

$$\|S_{n+p}-S_n\|=\Big\|\sum_{k=n+1}^{n+p}x_k\Big\|\le\sum_{k=n+1}^{n+p}\|x_k\|\le\sum_{k=n+1}^{\infty}\|x_k\|\to0.$$

故 $\{S_n\}$ 是 Cauchy 列，由 $(E,\|\cdot\|)$ 完备，$\{S_n\}$ 收敛，即 $\sum x_n$ 收敛。

($\Leftarrow$) 设 $\{x_n\}\subset E$ 是 Cauchy 列。

第一步：抽子列使相邻项很近。由 Cauchy 性，$\forall\varepsilon=\frac{1}{2^k}$，$\exists N_k\in\mathbb{N}_+$，使得当 $m,n\ge N_k$ 时 $\|x_m-x_n\|<\frac{1}{2^k}$。令

$$n_1=N_1,\qquad n_{k+1}=\max\{N_{k+1},\,n_k+1\},$$

则 $\{n_k\}$ 严格递增，并且

$$\|x_{n_{k+1}}-x_{n_k}\|<\frac{1}{2^k}\qquad(k\ge1).$$

事实上，最后这个不等式由构造 $n_k\ge N_k$，且 $n_{k+1}>n_k\ge N_k$，所以 n_k,n_{k+1} 都 $\ge N_k$，于是由 N_k 的定义立即得到 $\|x_{n_{k+1}}-x_{n_k}\|<\frac{1}{2^k}$。

第二步：Claim：级数 $\sum_{k=1}^{\infty}(x_{n_{k+1}}-x_{n_k})$ 绝对收敛。事实上，

$$\sum_{k=1}^{\infty}\|x_{n_{k+1}}-x_{n_k}\|\le\sum_{k=1}^{\infty}\frac{1}{2^k}=1<\infty.$$

由假设，该级数收敛，而它的部分和

$$\sum_{k=1}^{K}(x_{n_{k+1}}-x_{n_k})=x_{n_{K+1}}-x_{n_1}$$

收敛，即 $\{x_{n_{K+1}}-x_{n_1}\}_K$ 收敛，从而 $\{x_{n_k}\}$ 收敛。又因为 $\{x_n\}$ 是 Cauchy 列，一个 Cauchy 列有收敛子列，则整个数列收敛，故 $\{x_n\}$ 收敛。这说明 E 完备。证毕。

推论：设 $(E,\|\cdot\|)$ 是 Banach 空间，$\{x_n\}\subset E$。若级数 $\sum\limits_{n=1}^{\infty}x_n$ 绝对收敛，则 $\sum\limits_{n=1}^{\infty}x_n$ 在 E 中收敛，并且

$$\Big\|\sum_{n=1}^{\infty}x_n\Big\|\le\sum_{n=1}^{\infty}\|x_n\|.$$

证明：前半部分已证（上面定理的 $(\Rightarrow)$）。因 $\sum\|x_n\|<\infty$，对 $S_n=\sum_{k=1}^{n}x_k$ $(n\in\mathbb{N}_+)$，

$$\|S_n\|\le\sum_{k=1}^{n}\|x_k\|\le\sum_{k=1}^{\infty}\|x_k\|<\infty.$$

令 $n\to\infty$：记 $S=\sum_{n=1}^{\infty}x_n$，由范数的连续性（上一节”运算的连续性”的命题 (1)），$S_n\to S$ 蕴含 $\|S_n\|\to\|S\|$，所以不等式 $\|S_n\|\le\sum_{k=1}^{\infty}\|x_k\|$ 取极限后保持，可得 $\Big\|\sum_{n=1}^{\infty}x_n\Big\|\le\sum_{n=1}^{\infty}\|x_n\|$。证毕。

补：一般 p 下 $C[-1,1]$ 不完备

例：$C[-1,1]$ 在 $\|\cdot\|_p$ 下不完备（$1\le p<\infty$）。$p=1$ 已证。

一般 $1\le p<\infty$，在 $[-1,1]$ 上取

$$f_n(x)=\begin{cases}-1, & x<-\dfrac{1}{n},\\ nx, & |x|\le\dfrac{1}{n},\\ 1, & x>\dfrac{1}{n}.\end{cases}$$

当 $m>n$ 时，$|f_m-f_n|\le 2$，但只在 $\left[-\frac{1}{n},\frac{1}{n}\right]$ 上不为零。于是

$$\|f_m-f_n\|_p^p=\int_{-1}^{1}|f_m-f_n|^p\,dx\le\frac{2^{p+1}}{n}\to 0.$$

所以 $\{f_n\}$ 是 Cauchy 列，它的极限只能是 $\operatorname{sgn}x$（不连续），故不收敛。

对于 $1\le p<\infty$，$(C[0,1],\|\cdot\|_p)\overset{\text{完备化}}{\rightsquigarrow}L^p(0,1)$。

上面的每一步验证如下。

- $|f_m-f_n|\le 2$：两个函数的取值都在 $[-1,1]$ 内。
- 只在 $\left[-\frac{1}{n},\frac{1}{n}\right]$ 上非零：当 $|x|>\dfrac{1}{n}$ 时，因为 $m>n$，也有 $|x|>\dfrac{1}{m}$，所以 $f_m(x)=f_n(x)=\operatorname{sgn}x$，差为 0。

- 因此 $\int_{-1}^{1}|f_m-f_n|^p \le 2^p\cdot\dfrac{2}{n}=\dfrac{2^{p+1}}{n}$。
- 不收敛：设有 $g\in C[-1,1]$ 使 $\|f_n-g\|_p\to 0$。对 $\delta\in(0,1)$，当 $n\ge\dfrac{1}{\delta}$ 时 $f_n\equiv -1$ 在 $[-1,-\delta]$ 上，所以 $\int_{-1}^{-\delta}|g+1|^p\le\|f_n-g\|_p^p\to 0$，由 g 连续知 $g\equiv -1$ 在 $[-1,-\delta]$ 上；同理 $g\equiv 1$ 在 $[\delta,1]$ 上。令 $\delta\to 0$，得 $g\equiv -1$ 在 $[-1,0)$ 上，$g\equiv 1$ 在 $(0,1]$ 上，与 g 在 0 处连续矛盾。
- 本章开头给出的不完备性证明是在 $C[0,1]$ 上用函数列 g_n 做的；这里的 $[-1,1]$ 版本思路完全一样，只是极限换成了符号函数 $\operatorname{sgn}x$。

2.3 等价范数

定义（等价范数）：E 上的两个范数 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 称为**等价**的，若存在常数 $C_1,C_2>0$，使得

$$C_1\|x\|_1\le\|x\|_2\le C_2\|x\|_1,\qquad\forall x\in E.$$

注：

① 即等价度量：$\mathrm{id}:(E,\|\cdot\|_1)\to(E,\|\cdot\|_2)$ 是双射、双 Lipschitz 的。

② 等价范数诱导的度量下，开集、收敛序列、Cauchy 列、有界集、紧集、完备性都一样（相同）。

具体地，记 $d_i(x,y)=\|x-y\|_i$。由 $d_2(x,y)\le C_2d_1(x,y)$ 知 $\mathrm{id}:(E,d_1)\to(E,d_2)$ 是 C_2-Lipschitz 的；由 $d_1(x,y)\le\dfrac{1}{C_1}d_2(x,y)$ 知 $\mathrm{id}:(E,d_2)\to(E,d_1)$ 是 $\dfrac{1}{C_1}$-Lipschitz 的，所以 id 是双射、双 Lipschitz 的，这就是①。对于②，逐项来看：

- 开集：$B_{d_1}(x,r)\subset B_{d_2}(x,C_2r)$，$B_{d_2}(x,s)\subset B_{d_1}\left(x,\frac{s}{C_1}\right)$，两个度量互相能把对方的球装进自己的球里，所以开集族相同（拓扑等价，见”一致连续与同胚”一节的判别条件 (iv)）。
- 收敛序列：$\|x_n-x\|_2\le C_2\|x_n-x\|_1$，反之 $\|x_n-x\|_1\le\dfrac{1}{C_1}\|x_n-x\|_2$，所以 $x_n\to x$ 在两个范数下同时成立或同时不成立。
- Cauchy 列：同理，用 $\|x_m-x_n\|$ 代替 $\|x_n-x\|$。
- 有界集：$\|x\|_2\le C_2\|x\|_1$ 与 $\|x\|_1\le\dfrac{1}{C_1}\|x\|_2$。
- 紧集：由开集相同（或由收敛序列相同、自列紧的序列刻画相同）。
- 完备性：Cauchy 列相同、收敛序列相同，所以一个空间在一个范数下完备当且仅当在另一个范数下完备。

需要注意，第一章里说过” 完备” 不是拓扑性质——两个拓扑等价的度量下完备性可以不同。这里并不矛盾：等价范数诱导的度量满足更强的” 双 Lipschitz 等价”，完备性就被保持了。

例

(1) $\mathbb{K}^n$ 上的 p-范数两两等价。设 $x=(x_1,\dots,x_n)$，$1\le p<\infty$，则

$$\|x\|_\infty \le \|x\|_p \le n^{1/p}\|x\|_\infty.$$

于是 $\|\cdot\|_p \sim \|\cdot\|_q$，$\forall 1\le p,q\le\infty$。最终将证明（见下一节” 有限维赋范空间”）：$\mathbb{K}^n$ 上（即任意有限维赋范空间上）任意两个范数都等价。

事实上，$\|x\|_\infty^p=\max_k|x_k|^p\le\sum_k|x_k|^p=\|x\|_p^p$，得 $\|x\|_\infty\le\|x\|_p$；又 $\sum_k|x_k|^p\le n\max_k|x_k|^p=n\|x\|_\infty^p$，得 $\|x\|_p\le n^{1/p}\|x\|_\infty$。对任意 $1\le p,q\le\infty$（$p=\infty$ 时 $n^{1/p}$ 理解为 1），

$$\|x\|_p\le n^{1/p}\|x\|_\infty\le n^{1/p}\|x\|_q,\qquad \|x\|_q\le n^{1/q}\|x\|_\infty\le n^{1/q}\|x\|_p,$$

即 $n^{-1/q}\|x\|_q\le\|x\|_p\le n^{1/p}\|x\|_q$，所以 $\|\cdot\|_p$ 与 $\|\cdot\|_q$ 等价。结合上面的注②，由 $\mathbb{K}^n$ 在 p-范数下完备，再加上有限维空间上任意两个范数等价（下一节），便得到 $\mathbb{K}^n$ 在任意范数下完备。

(2) $(C[0,1],\|\cdot\|_p)$ 两两不等价。要证两个范数不等价，只需找 $\{x_n\}$，使得两个范数的比值 $\dfrac{\|x_n\|_{(\text{一个})}}{\|x_n\|_{(\text{另一个})}}\to\infty$ 或 0。

在 $C[0,1]$ 取 $f_n(x)=x^n$，则 $\|f_n\|_\infty=1$，

$$\|f_n\|_p=\Big(\int_0^1 x^{np}\,dx\Big)^{1/p}=\Big(\frac{1}{np+1}\Big)^{1/p}\to 0.$$

故 $\|\cdot\|_p$ 与 $\|\cdot\|_\infty$ 不等价（$\dfrac{\|f_n\|_p}{\|f_n\|_\infty}\to0$）。同理，$p<q$ 时 $\dfrac{\|f_n\|_p}{\|f_n\|_q}\to0$，故不等价。

事实上，$\dfrac{\|f_n\|_p}{\|f_n\|_q}=\dfrac{(nq+1)^{1/q}}{(np+1)^{1/p}}$，分子 $\sim n^{1/q}$，分母 $\sim n^{1/p}$，因为 $p<q$ 时 $\dfrac1q<\dfrac1p$，比值 $\sim n^{1/q-1/p}\to0$。而如果两个范数等价，则 $\dfrac{\|x\|_p}{\|x\|_q}$ 在一切 $x\neq0$ 上同时有正的上下界，所以比值趋于 0 就足以说明不等价。

由 Hölder 不等式，$\|f\|_p\le\|f\|_\infty$，$\|f\|_p\le\|f\|_q$（$p<q$；这里 $\Omega=(0,1)$）。第一个是平凡的：$\int_0^1|f|^p\le\|f\|_\infty^p\cdot1$。第二个：对 $p<q$，用 Hölder 不等式（指数 $\dfrac{q}{p}$ 与它的共轭指数）：

$$\int_0^1|f|^p\cdot1\,dx\le\Big(\int_0^1|f|^{p\cdot\frac{q}{p}}\,dx\Big)^{p/q}\Big(\int_0^1 1\,dx\Big)^{1-p/q}=\|f\|_q^p,$$

即 $\|f\|_p\le\|f\|_q$。所以这些范数之间只有” 一个方向” 的控制关系：较小的 p 对应较弱的范数，而反方向的控制失败（见上面 $f_n=x^n$），这就是它们不等价的原因。

(3) 多项式空间 $\mathbb{R}[x] \subset C[0,1]$，赋予 $\|\cdot\|_\infty$，不完备。提示：$f_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$（多项式）在 $\|\cdot\|_\infty$ 下（一致收敛）收敛到 e^x，$\{f_n\}$ 是 Cauchy 列，但在 $\mathbb{R}[x]$ 中无极限。这正是前面" 不完备的例子③"，完整的证明见那里。

2.4 有限维赋范空间

有限维与坐标映射

定义（有限维）：设 E 是 $\mathbb{K}$ 上的向量空间。E **有限维**，指存在一个有限的基 $\{e_1, e_2, \dots, e_m\}$，使得对每个 $x \in E$，存在唯一的 $x_i \in \mathbb{K}$ $(1 \le i \le m)$，满足

$$x = \sum_{i=1}^{m} x_i e_i.$$

（下面总设 $m \ge 1$，即 $E \neq \{0\}$；$E = \{0\}$ 时一切结论平凡。）

定义（坐标映射）：

$$\varphi : E \to \mathbb{K}^m, \qquad x \mapsto (x_1, x_2, \dots, x_m),$$

φ 是线性（代数）同构。

事实上，由坐标的唯一性，φ 的定义合理。线性：$\alpha x + \beta y = \sum_i (\alpha x_i + \beta y_i) e_i$，由唯一性，$\alpha x + \beta y$ 的坐标就是 $\alpha x_i + \beta y_i$，即 $\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y)$。单射：坐标相同则 $x = \sum x_i e_i$ 相同。满射：对 $(c_1, \dots, c_m) \in \mathbb{K}^m$，$\varphi\big(\sum c_i e_i\big) = (c_1, \dots, c_m)$。

命题：令

$$p(x) := \sum_{i=1}^{m} |x_i|, \qquad \Big(x = \sum_{i=1}^{m} x_i e_i\Big),$$

则 p 是 E 上的范数，且 $\varphi : (E, p) \to (\mathbb{K}^m, \|\cdot\|_1)$ 是等距同构。

事实上，按定义 $p(x) = \|\varphi(x)\|_1$，而 φ 是线性的。于是 $p(\alpha x) = \|\alpha\varphi(x)\|_1 = |\alpha| p(x)$；$p(x+y) = \|\varphi(x) + \varphi(y)\|_1 \le p(x) + p(y)$；$p(x) = 0 \Rightarrow \varphi(x) = 0 \Rightarrow x = 0$（$\varphi$ 单射）。所以 p 是范数。等距在下面一小节直接给出。

(E, p) 具有 $(\mathbb{K}^m, \|\cdot\|_1)$ 的一切性质

对任意 $x, y \in E$，

$$\|\varphi(x) - \varphi(y)\|_1 = \sum_{i=1}^{m} |x_i - y_i| = p(x - y)$$

$$\Longrightarrow \quad \varphi \text{ 双射，且1-Lipschitz（实际上等距）.}$$

因此 $(\mathbb{K}^m, \|\cdot\|_1)$ 的性质 (E, p) 都有，特别是：

- (E,p) 完备；
- $A\subset(E,p)$ 紧 $\iff$ A 有界闭。

其中完备性的理由是：$(\mathbb{K}^m,\|\cdot\|_1)$ 完备（上一节” 完备的例子①”），φ 是等距满射，Cauchy 列、收敛列都被 φ、φ^{-1} 原样搬运，所以 (E,p) 完备。

证：($\Rightarrow$) 这是任何度量空间的一般事实（第一章” 紧性” 一节）：紧 $\Rightarrow$ 预紧（完全有界）$\Rightarrow$ 有界；紧 $\Rightarrow$ 列紧 $\Rightarrow$ 闭（列紧集上的收敛序列的极限仍在集内）。

($\Leftarrow$) A 有界闭 $\Rightarrow$ $\varphi(A)$ 有界闭 $\Rightarrow$ $\varphi(A)$ 是 $(\mathbb{K}^m,\|\cdot\|_1)$ 中的紧集，再由 φ^{-1} 连续，

$$\varphi^{-1}(\varphi(A))=A \text{ 紧，闭.}$$

具体地，$\varphi(A)$ 有界闭用到 φ 是等距满射（保持有界性，并把闭集映成闭集）；” 紧” 用 Heine–Borel：$\mathbb{K}^m$ 中 $\|\cdot\|_1$ 与欧氏范数 $\|\cdot\|_2$ 满足 $\|c\|_2\le\|c\|_1\le\sqrt{m}\,\|c\|_2$（后一个由 Cauchy–Schwarz），所以两者的有界闭集相同，欧氏范数下有界闭集紧（$\mathbb{K}=\mathbb{C}$ 时把 $\mathbb{C}^m$ 看成 $\mathbb{R}^{2m}$）；最后 φ^{-1} 是等距，当然连续，连续映射把紧集映成紧集。

有限维空间上的任意两个范数等价

定理：E 上所有的范数都与 p 范数等价。因此，E 上任意两个范数都等价。

证明：设 $\|\cdot\|$ 是 E 上的一个给定范数。

第一步：证明 $\|x\|\le M\cdot p(x)$，$\forall x\in E$，其中 $M:=\max\limits_{1\le i\le m}\|e_i\|$。

$\forall\, x=\sum\limits_{i=1}^{m}x_ie_i\in E$，

$$\|x\|=\Big\|\sum_{i=1}^{m}x_ie_i\Big\|\le\sum_{i=1}^{m}|x_i|\,\|e_i\|\le M\cdot\sum_{i=1}^{m}|x_i|=M\cdot p(x).$$

这里 $M>0$，因为基向量非零；第一个不等号是三角不等式与齐次性，第二个不等号是 $\|e_i\|\le M$。于是

$$\big|\,\|x\|-\|y\|\,\big|\le\|x-y\|\le M\cdot p(x-y),$$

其中第一个不等号是范数的反向三角不等式（上一节” 运算的连续性”）。也就是说

$$(E,p)\xrightarrow{\mathrm{id}}(E,\|\cdot\|)\xrightarrow{\|\cdot\|}\mathbb{R},\qquad x\longmapsto x\longmapsto\|x\|,$$

因此 $\|\cdot\|$ 是 M-Lipschitz 的，从而连续（作为 $(E,p)\to\mathbb{R}$ 的函数）。

第二步：在单位球面 $S_p:=\{x\in E\mid p(x)=1\}$ 上取最值。

S_p 非空：$p(e_1)=1$，所以 $e_1\in S_p$，最大、最小值才有意义。$S_p=p^{-1}(\{1\})$ 是连续函数的原像，所以在 (E,p) 中闭；有界显然（$p\equiv1$）。由命题（” 有界闭 $\Rightarrow$ 紧”）它是紧集。连续函数 $\|\cdot\|$ 在紧集 S_p 上取到最大值 β 与最小值 α，即

$$\alpha\le\|y\|\le\beta,\qquad\forall y\in S_p.$$

又由于 S_p 中的点非零（$0_E \notin S_p$，因为 $p(0_E)=0$），故 $\alpha>0$。（否则 $\alpha=0$，则存在 $x\in S_p$ 使得 $\|x\|=0 \iff x=0_E$，矛盾。）

第三步：齐次性。$\forall y\in E$ 且 $y\neq 0_E$，有 $\dfrac{y}{p(y)}\in S_p$，则由第二步，

$$\alpha \le \left\|\frac{y}{p(y)}\right\| \le \beta \quad\Longrightarrow\quad \alpha\, p(y) \le \|y\| \le \beta\, p(y).$$

$y=0$ 时显然成立。故 $\|\cdot\|$ 与 p 等价。而两个都与 p 等价的范数是相互等价的：设 $\alpha_i p \le \|\cdot\|_i \le \beta_i p$ $(i=1,2)$，则

$$\frac{\alpha_2}{\beta_1}\|x\|_1 \le \alpha_2 p(x) \le \|x\|_2 \le \beta_2 p(x) \le \frac{\beta_2}{\alpha_1}\|x\|_1,$$

这就是 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 的等价。（”等价”是对称、传递的：定义中交换两个范数，常数换成 $1/C_2, 1/C_1$ 即可。）　□

推论

推论：设 E 是有限维赋范空间，则：

(1) E 是 Banach 空间；

(2) 任何赋范空间（有限维或无限维）的有限维子空间都是闭空间；

(3) $A\subset E$ 紧 $\iff$ A 有界闭。特别地，闭单位球紧，有界序列有收敛子列；

(4) 同一数域 $\mathbb{K}$ 上维数相同的两个有限维赋范空间是拓扑同构的（存在双向连续的线性双射）。

证明：

(1) (E,p) 完备（已证），$(E,\|\cdot\|)$ 与 (E,p) 等价，等价范数保持完备性。

(2) 设 F 是 E（可以无限维）的有限维子空间。则 F 在诱导范数下是有限维赋范空间，由 (1) 完备，而完备子集是闭集：设 $x_n\in F$，$x_n\to x\in E$，则 $\{x_n\}$ 是 F 中的 Cauchy 列，由 F 完备，$x_n\to y\in F$。极限唯一，所以 $x=y\in F$。

(3) $A\subset(E,p)$ 时已证（上面的命题）。等价范数保持有界、紧、闭，故 A 对 $(E,\|\cdot\|)$ 下亦有界闭。特别地，有界序列落在某个闭球 $\bar B(0,R)$ 里，闭球有界闭，由前半部分是紧集，紧度量空间中的序列有收敛子列（紧 $\Rightarrow$ 列紧）。

(4) 设 E,F 是 $\mathbb{K}$ 上的 m 维赋范空间，范数分别记为 $\|\cdot\|_E,\|\cdot\|_F$，分别取基 $\{e_1,\dots,e_m\}$ 与 $\{f_1,\dots,f_m\}$，得坐标映射 $\varphi_E: E\to\mathbb{K}^m$，$\varphi_F: F\to\mathbb{K}^m$：

$$(E,\|\cdot\|_E) \xrightarrow{\varphi_E} (\mathbb{K}^m,\|\cdot\|_1) \xleftarrow{\varphi_F} (F,\|\cdot\|_F).$$

定义 $\varphi := \varphi_F^{-1} \circ \varphi_E$，则 φ 是 $(E, \|\cdot\|_E)$ 到 $(F, \|\cdot\|_F)$ 的拓扑同构。具体地，对 E 取 $p_E(x) = \sum |x_i|$，由定理 $\alpha_E p_E \le \|\cdot\|_E \le \beta_E p_E$，而 $\|\varphi_E(x)\|_1 = p_E(x)$，所以

$$\frac{1}{\beta_E}\|x\|_E \le \|\varphi_E(x)\|_1 \le \frac{1}{\alpha_E}\|x\|_E,$$

即 φ_E 是线性双射且双 Lipschitz 的；φ_F 同理。于是 $\varphi = \varphi_F^{-1} \circ \varphi_E$ 是线性双射。把两个估计合起来（用到 $\varphi_F \circ \varphi = \varphi_E$，因而 $\|\varphi_E(x)\|_1 = p_F(\varphi(x))$）：

$$\|\varphi(x)\|_F \le \beta_F\, p_F(\varphi(x)) = \beta_F \|\varphi_E(x)\|_1 \le \frac{\beta_F}{\alpha_E}\|x\|_E,$$

$$\|x\|_E \le \beta_E \|\varphi_E(x)\|_1 = \beta_E\, p_F(\varphi(x)) \le \frac{\beta_E}{\alpha_F}\|\varphi(x)\|_F,$$

所以 φ 与 φ^{-1} 都是 Lipschitz 连续的，φ 是拓扑同构。

注：以上性质在无穷维空间中都不成立

注：所有上述性质 (1)–(4) 在无穷维空间都是不对的。

(1) $\mathbb{R}[x] \subset C[0,1]$ 上任给一个范数都不是 Banach（作业；一个完整的证明放在下一节末尾，需要用到 Riesz 引理）。

$C[a,b]\,(\mathbb{K})$ 在 L^p 范数 $(1 \le p < \infty)$ 下不完备（已在前面证明，见” 赋范空间的例子” 与” 补：一般 p 下 $C[-1,1]$ 不完备”）。

(2) $\mathbb{R}[x]$ 在 $C[a,b]$ 中稠密（Weierstrass 多项式逼近定理）。于是 $\mathbb{R}[x]$ 是 $(C[a,b], \|\cdot\|_\infty)$ 的无穷维子空间，它的闭包是整个 $C[a,b]$，而 $e^x \in C[a,b] \setminus \mathbb{R}[x]$，所以 $\mathbb{R}[x]$ 不是闭子空间，推论 (2) 在无穷维时失败。

对 (3)、(4)，在无穷维下失败的原因分别是下一节的 Riesz 定理（闭单位球不紧），以及前一节已证的”$C[0,1]$ 上 $\|\cdot\|_p$ 两两不等价”（无穷维空间上的范数不一定等价）。

2.5 Riesz 引理与 Riesz 定理

点到集合的距离与最佳逼近

设 $(E, \|\cdot\|)$ 是赋范空间。$\forall F \subset E$（非空），$\forall z \in E$，记

$$d(z,F) = \inf_{w \in F} \|z - w\|.$$

$$d(z,F) = 0 \iff z \in \overline{F}. \qquad \text{（第 1 次作业）}$$

事实上，$d(z,F) = 0 \iff \forall \varepsilon > 0,\ \exists w \in F,\ \|z-w\| < \varepsilon \iff$ 存在 $w_n \in F$ 使 $w_n \to z \iff z \in \overline{F}$（第一章” 收敛、聚点与闭包”：闭包是 F 中序列的极限点集合）。

命题（到有限维子空间的最佳逼近存在）：设 F 是赋范空间 E 的有限维子空间，$z \in E$，则 $\exists x \in F$ s.t.

$$\|z - x\| = d(z, F) = \min_{w \in F} \|z - w\|.$$

证明：取 $\{x_n\} \subset F$ s.t.

$$\|z - x_n\| \le d(z, F) + \frac{1}{n} \qquad (n \in \mathbb{N}_+).$$

则

$$\|x_n\| \le \|z\| + \|z - x_n\| \le \|z\| + d(z, F) + \frac{1}{n} \implies \{x_n\} \text{ 是}F \text{ 中的有界序列.}$$

由紧性推论（上一节推论 (3)：有限维赋范空间中有界序列有收敛子列），存在子列 $\{x_{n_k}\} \to x \in F$（$\because F$ 闭）。这里 F 在诱导范数下是有限维赋范空间，上一节推论 (3) 用在 F 上：有界序列 $\{x_n\} \subset F$ 有 F 中的收敛子列，极限本来就落在 F 内（推论 (2) 也说明 F 是 E 中的闭集，所以不会有极限跑到 F 外面的问题）。由范数连续性，

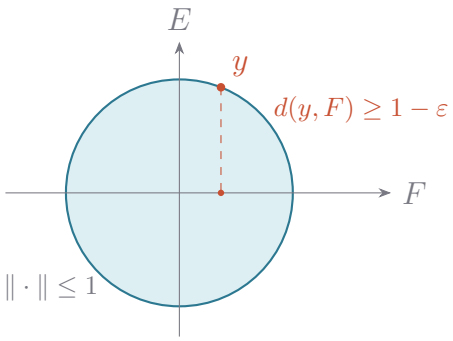
$$d(z, F) \le \|z - x\| = \lim_{k \to \infty} \|z - x_{n_k}\| \le d(z, F) \implies \|z - x\| = d(z, F). \qquad \square$$

最后一步中 $\lim_k \left(d(z, F) + \frac{1}{n_k}\right) = d(z, F)$，所以第二个不等号成立。

注：x 未必唯一。例如 $E = \mathbb{R}^2$, $\|\cdot\|_\infty$, $F = \{(t, 0) : t \in \mathbb{R}\}$, $z = (0, 1)$。则 $d(z, F) = 1$，且 $\|z - (t, 0)\|_\infty = \max(|t|, 1) = 1$ 对所有 $|t| \le 1$ 都成立，最佳逼近有无穷多个。

Riesz 引理

引理（Riesz 引理）：设 F 是赋范空间 E 的真闭子空间，则 $\forall \varepsilon \in (0, 1)$，$\exists y \in E$ s.t. $\|y\| = 1$，$d(y, F) \ge 1 - \varepsilon$。若 F 有限维，则 $\exists y$ s.t. $\|y\| = 1$，且 $d(y, F) = 1$。



证明：取 $y_0 \in E \setminus F$，因 F 闭，

$$\alpha := d(y_0, F) > 0 \qquad (\text{否则} y_0 \in \overline{F} = F).$$

$\forall \delta > 0$，由下确界的定义，$\exists x_\delta \in F$ s.t.

$$\|y_0 - x_\delta\| \le \alpha + \delta.$$

令

$$y=\frac{y_0-x_\delta}{\|y_0-x_\delta\|},\qquad \|y\|=1.$$

这里分母不为零：$y_0\notin F$ 而 $x_\delta\in F$，所以 $\|y_0-x_\delta\|>0$。$\forall x\in F$，因为 $x_\delta+\|y_0-x_\delta\|\,x\in F$，

$$\begin{aligned}\|y-x\|&=\frac{\big\|(y_0-x_\delta)-x\,\|y_0-x_\delta\|\big\|}{\|y_0-x_\delta\|}\\&=\frac{\big\|y_0-\overbrace{\big[x_\delta+\|y_0-x_\delta\|\cdot x\big]}^{\in F}\big\|}{\|y_0-x_\delta\|}\geq\frac{\alpha}{\alpha+\delta}.\end{aligned}$$

最后的不等号的来源是：分子 $\|y_0-[\cdots]\|\geq d(y_0,F)=\alpha$（方括号里的点属于 F），分母 $\|y_0-x_\delta\|\leq\alpha+\delta$。取 δ s.t. $\dfrac{\alpha}{\alpha+\delta}\geq 1-\varepsilon$，即得 $d(y,F)\geq 1-\varepsilon$。由于

$$\frac{\alpha}{\alpha+\delta}\geq 1-\varepsilon\iff\alpha\geq(1-\varepsilon)(\alpha+\delta)\iff\delta\leq\frac{\varepsilon\alpha}{1-\varepsilon},$$

取 $\delta=\dfrac{\varepsilon\alpha}{1-\varepsilon}>0$ 即可。

若 F 有限维，由上一命题，可取 $x_0\in F$ s.t. $\|y_0-x_0\|=\alpha$（即上式中 $\delta=0$）。这时 $y=\dfrac{y_0-x_0}{\alpha}$，$\|y\|=1$，上面的估计中分子 $\geq\alpha$、分母 $=\alpha$，此时 $d(y,F)\geq\dfrac{\alpha}{\alpha}=1$。又 $\because 0\in F$，$d(y,F)\leq\|y\|=1$，$\Rightarrow d(y,F)=1$。□

Riesz 定理

定理（Riesz 定理）：赋范空间 $(E,\|\cdot\|)$ 的单位球面 S（等价地，闭单位球 $\bar{B}$）紧 $\iff\dim E<\infty$。

证明：($\Leftarrow$) 已证（上一节推论 (3)：有限维空间中有界闭集紧，S 与 $\bar{B}$ 都有界闭）。

($\Rightarrow$) 证其逆否命题：设 $\dim E=\infty$，证 S 不紧。任取 $x_1\in S$，$F_1=\operatorname{span}\{x_1\}$，$F_1$ 有限维，故是闭子空间且 $F_1\neq E$。由 Riesz 引理（有限维的 case），$\exists x_2\in S$ s.t. $d(x_2,F_1)=1$。一般地，得到 $\{x_1,x_2,\dots,x_n\}$ 后，令 $F_n=\operatorname{span}\{x_1,x_2,\dots,x_n\}$，它是有限维真闭子空间，取 $x_{n+1}\in S$ s.t. $d(x_{n+1},F_n)=1$。于是 $\{x_n\}\subset S$。每一步都能继续：F_n 有限维，所以 $F_n\neq E$（$\dim E=\infty$），又由推论 (2) 它闭，所以 Riesz 引理（有限维情形）可用在 F_n 上。

$\forall m>n$，因为 $x_n\in F_{m-1}$（$x_n\in F_n\subset F_{m-1}$），

$$\|x_m-x_n\|\geq d(x_m,F_{m-1})=1.$$

$\{x_n\}\subset S$ 任意两项距离 ≥ 1，没有 Cauchy 列，更没有收敛子列。（序列 $\{x_n\}$ 的任何子列也满足任意两项距离 ≥ 1，所以不是 Cauchy 列；收敛序列必是 Cauchy 列，所以也没有收敛子列。）所以 S 不列紧 $\Rightarrow$ S 不紧。$\bar{B}\supset S$ 且 S 是 $\bar{B}$ 的闭子集，故 $\bar{B}$ 也不紧：$S=\{x:\|x\|=1\}$，$\|\cdot\|$ 连续，所以 S 在 E 中闭，若 $\bar{B}$ 紧，则它的闭子集 S 紧，与上面矛盾。□

推论：

(1) 无限维赋范空间中，有界闭集不一定紧，有界序列不一定有收敛子列。

事实上，由 Riesz 定理，无限维空间中闭单位球 $\bar{B}$ 是有界闭集，但不紧；上面构造的 $\{x_n\} \subset S \subset \bar{B}$ 就是一个有界却没有收敛子列的序列。具体例子：ℓ^p 中取单位向量 $e_n = (0, \dots, 0, 1, 0, \dots)$，则 $\|e_n\|_p = 1$，$\|e_n - e_m\|_p = 2^{1/p}$ $(m \neq n)$（ℓ^∞、c_0 中该距离为 1），所以没有收敛子列。

补：作业的一个证明（$\mathbb{R}[x]$ 上任何范数都不完备）

设 $\|\cdot\|$ 是 $V = \mathbb{R}[x]$（或 $\mathbb{C}[x]$）上的任意一个范数，假设 $(V, \|\cdot\|)$ 是 Banach 空间，导出矛盾。

记 V_n 为次数 $\le n$ 的多项式全体（$n \ge 0$），$V_{-1} = \{0\}$。V_n 是 $n+1$ 维子空间，由上一节推论 (2)，每个 V_n 都是 V 的闭子空间，且 $V_{n-1} \subsetneq V_n$。

第一步：取 y_n。对每个 $n \ge 0$，把 V_n 看作赋范空间，V_{n-1} 是它的有限维真闭子空间，由 Riesz 引理（有限维情形）取 $y_n \in V_n$ 满足

$$\|y_n\| = 1, \qquad d(y_n, V_{n-1}) = 1.$$

（$n = 0$ 时取 $y_0 = 1/\|1\|$。）

第二步：构造级数。$\sum_{n\ge 0} 3^{-n} y_n$ 绝对收敛（$\sum 3^{-n}\|y_n\| = \sum 3^{-n} < \infty$），由”完备 $\iff$ 绝对收敛级数收敛”，它在 V 中收敛到某个 $s \in V$。s 是多项式，设 $s \in V_N$。

第三步：导出矛盾。令 $r = s - \sum_{n=0}^{N} 3^{-n} y_n$。因为 $y_n \in V_n \subset V_N$ $(n \le N)$，所以 $r \in V_N$；另一方面

$$r = \sum_{n \ge N+1} 3^{-n} y_n = 3^{-(N+1)} y_{N+1} + t, \qquad t := \sum_{n \ge N+2} 3^{-n} y_n,$$
$$\|t\| \le \sum_{n \ge N+2} 3^{-n} = \frac{3^{-(N+1)}}{2}.$$

于是

$$\left\| y_{N+1} - 3^{N+1} r \right\| = 3^{N+1}\|t\| \le \frac{1}{2}.$$

但 $3^{N+1} r \in V_N$，所以 $d(y_{N+1}, V_N) \le \dfrac{1}{2}$，与 $d(y_{N+1}, V_N) = 1$ 矛盾。因此 $(\mathbb{R}[x], \|\cdot\|)$ 对任何范数都不是 Banach。 □

（用到的事实：有限维子空间是闭的、Riesz 引理（有限维情形）、Banach 空间中绝对收敛级数收敛。$\sum_{k \ge N+2} 3^{-k} = \dfrac{3^{-(N+2)}}{1 - 1/3} = \dfrac{3^{-(N+1)}}{2}$。）

2.6 连续线性映射（算子）

线性映射

定义（线性映射）：设 E, F 是 $\mathbb{K}$ 上的向量空间。映射 $u : E \to F$ 称为 **线性映射**（**线性算子**），若

$$u(\lambda x + y) = \lambda u(x) + u(y), \qquad \forall\, x, y \in E,\ \lambda \in \mathbb{K}.$$

这与通常的“可加 + 齐次”的定义等价：取 $\lambda = 1$ 得可加性 $u(x + y) = u(x) + u(y)$；取 $y = 0_E$ 得齐次性 $u(\lambda x) = \lambda u(x)$（这里要先知道 $u(0_E) = 0_F$，见下面注③，它只用到 $\lambda = 1, x = y = 0_E$ 这一特例）。反过来，可加与齐次显然给出 $u(\lambda x + y) = u(\lambda x) + u(y) = \lambda u(x) + u(y)$。

记号：

- $\mathcal{L}(E, F)$：$E \to F$ 的线性映射的全体；
- $\mathcal{L}(E, E) = \mathcal{L}(E)$；
- $\mathcal{L}(E, \mathbb{K})$ 中的元素称为 E 上的**线性泛函**。

注：

① $\mathcal{L}(E, F)$ 是一个向量空间：$\forall\, u, v \in \mathcal{L}(E, F)$，

$$(u + v)(x) = u(x) + v(x), \qquad (\lambda u)(x) = \lambda u(x),$$

则 $u + \lambda v \in \mathcal{L}(E, F)$，且 $0 \in \mathcal{L}(E, F)$（零映射）。

验证：对 $x, y \in E$，$\alpha \in \mathbb{K}$，

$$\begin{aligned}(u + \lambda v)(\alpha x + y) &= u(\alpha x + y) + \lambda v(\alpha x + y)\\ &= \alpha u(x) + u(y) + \lambda\alpha v(x) + \lambda v(y)\\ &= \alpha\,(u + \lambda v)(x) + (u + \lambda v)(y),\end{aligned}$$

所以 $u + \lambda v$ 仍是线性的；零映射 $x \mapsto 0_F$ 显然线性。向量空间的其余公理（交换律、结合律、分配律等）由 F 中的运算逐点继承。

② $\dim E = m$，$\dim F = n$。$E \cong \mathbb{K}^m$，$F \cong \mathbb{K}^n$，$u : E \to F$，$u \cong [u]$ 是 $n \times m$ 的矩阵。这里的“$\cong$”指线性（代数）同构，下一小节会把对应关系 $u \mapsto [u]$ 写出来。

线性映射的矩阵表示

取 E 的基 $\{e_1, e_2, \dots, e_m\}$，取 F 的基 $\{f_1, f_2, \dots, f_n\}$，$u \in \mathcal{L}(E, F)$。则

$$u(e_j) = \sum_{i=1}^{n} a_{ij} f_i = a_{1j} f_1 + a_{2j} f_2 + \cdots + a_{nj} f_n, \qquad j = 1, 2, \dots, m.$$

把 $u(e_j)$ 的坐标 $a_{1j}, a_{2j}, \dots, a_{nj}$ 竖着写成第 j 列，得

$$A = [a_{ij}] = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix},$$

其中第 j 列 $(a_{1j}, a_{2j}, \dots, a_{nj})^{\mathsf{T}}$ 就是 $u(e_j)$ 的坐标。或者写成

$$\big(u(e_1), u(e_2), \dots, u(e_m)\big) = (f_1, \dots, f_n)\, A.$$

最后一行是形式记号：右边是行向量 $(f_1, \dots, f_n)$ 乘以矩阵 A，第 j 个分量是 $\sum_i f_i a_{ij} = u(e_j)$。

对任意 $x = \sum_{j=1}^{m} x_j e_j \in E$（$x_j \in \mathbb{K}$ 唯一），

$$u(x) = \sum_{j=1}^{m} x_j u(e_j) = \sum_{j=1}^{m} x_j \Big(\sum_{i=1}^{n} a_{ij} f_i \Big) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{m} a_{ij} x_j \Big) f_i \overset{\text{记为}}{=} \sum_{i=1}^{n} y_i f_i,$$

即

$$y_i = \sum_{j=1}^{m} a_{ij} x_j \quad (1 \le i \le n), \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

坐标表示：$u(x) = Ax$。

例如：$\mathcal{L}(\mathbb{C}^2, \mathbb{C}^2) \cong M_2(\mathbb{C})$。一般地，映射 $\mathcal{L}(E, F) \to M_{n\times m}(\mathbb{K})$，$u \mapsto A$ 是线性同构：它是线性的，因为 $(u + \lambda v)(e_j) = u(e_j) + \lambda v(e_j)$，两边取坐标，第 j 列也是对应相加；单射，因为 u 由 $u(e_1), \dots, u(e_m)$ 唯一确定（上面的展开式）；满射，因为任给 $n \times m$ 矩阵 A，定义 $u\big(\sum_j x_j e_j\big) := \sum_i \big(\sum_j a_{ij} x_j\big) f_i$，由坐标的唯一性这是良好定义的线性映射，并且 $u(e_j) = \sum_i a_{ij} f_i$。所以 $\dim \mathcal{L}(E, F) = nm$。例子中 $m = n = 2$，$\dim_{\mathbb{C}} \mathcal{L}(\mathbb{C}^2, \mathbb{C}^2) = 4 = \dim_{\mathbb{C}} M_2(\mathbb{C})$。

注：

③ $u \in \mathcal{L}(E, F) \Rightarrow u(0_E) = 0_F$；$u$ 的像 $u(E)$（$= \mathrm{Range}(u)$）与核 $\ker u = u^{-1}(0)$ 都是子空间。

证明：在 $u(\lambda x + y) = \lambda u(x) + u(y)$ 中取 $\lambda = 1, x = y = 0_E$，得 $u(0_E) = u(0_E) + u(0_E)$，所以 $u(0_E) = 0_F$。像是 F 的子空间：$0_F = u(0_E) \in u(E)$；若 $u(x), u(y) \in u(E)$，则 $\lambda u(x) + u(y) = u(\lambda x + y) \in u(E)$。核是 E 的子空间：$0_E \in \ker u$；若 $x, y \in \ker u$，则 $u(\lambda x + y) = \lambda u(x) + u(y) = 0_F$。

线性映射的连续性

定理（线性映射的连续性）：设 E, F 是赋范空间，$u \in \mathcal{L}(E, F)$，则下列条件等价：

(1) u 在 E 上连续；

(2) u 在某一点 $x_0 \in E$ 处连续；

(3) u 在原点 0_E 处连续；

(4) $\exists C \geq 0$ s.t. $\|u(x)\| \leq C\|x\|$，$\forall x \in E$ 成立（u **有界**）；

(5) u 在 E 上 Lipschitz 连续。

证明：

(1) $\Rightarrow$ (2) 显然（在每一点连续蕴含在某一点连续）。

(2) $\Rightarrow$ (3)：由 u 在 x_0 连续，则 $\forall \varepsilon > 0$，$\exists \delta > 0$ s.t.

$$\|y - x_0\| < \delta \implies \|u(y) - u(x_0)\| < \varepsilon.$$

令 $z = y - x_0$，即 $y = x_0 + z$，当 z 遍历球 $B(0, \delta)$ 时 y 遍历 $B(x_0, \delta)$。于是只要 $\|z\| < \delta$，就有

$$\|u(z)\| = \|u(y - x_0)\| = \|u(y) - u(x_0)\| < \varepsilon \implies u \text{ 在} 0_E \text{ 连续}$$

（因为 $u(0_E) = 0_F$，“在 0_E 处连续”的条件 $\|u(z) - u(0_E)\| < \varepsilon$ 就是 $\|u(z)\| < \varepsilon$）。

(3) $\Rightarrow$ (4)：$\because u$ 在 0_E 连续，取 $\varepsilon = 1$，$\exists \delta > 0$ s.t. $\|z\| < \delta \Rightarrow \|u(z)\| < 1$。于是 $\|z\| \leq \dfrac{\delta}{2} < \delta$ 时，$\|u(z)\| < 1$。$\forall x \neq 0$，令 $z = \dfrac{\delta}{2\|x\|} \cdot x$，则 $\|z\| = \dfrac{\delta}{2}$，且 $\dfrac{2\|x\|}{\delta} z = x$。由线性（齐次性），

$$\|u(x)\| = \left\|u\Big(\frac{2\|x\|}{\delta}\, z\Big)\right\| = \frac{2\|x\|}{\delta} \underbrace{\|u(z)\|}_{\leq 1} \leq \frac{2}{\delta}\|x\|.$$

$x = 0$ 显然成立，取 $C = \dfrac{2}{\delta}$ 即可。

(4) $\Rightarrow$ (5)：

$$\|u(x) - u(y)\| = \|u(x - y)\| \leq C\|x - y\| \implies u \text{ 是} C\text{-Lipschitz}.$$

(5) $\Rightarrow$ (1) 显然（Lipschitz 蕴含一致连续，更蕴含连续，见第一章“一致连续与同胚”）。$\square$

注：线性映射的像 $u(E)$ 无界，除非 $u = 0$。（因 $\exists x \in E$ s.t. $u(x) \neq 0$；$u(\lambda x) = \lambda u(x)$，$\|u(\lambda x)\| = |\lambda| \cdot \|u(x)\|$，其中 $\|u(x)\| > 0$，当 $|\lambda| \to \infty$ 时趋于 ∞。）例如 $E = F = \mathbb{R}$，$u(x) = ax$：$a \neq 0$ 时像是整个 $\mathbb{R}$，无界；但它满足 $|u(x)| = |a||x|$，是有界算子（取 $C = |a|$）。所以 (4) 里的“u 有界”不是“像 $u(E)$ 有界”，下面算子范数一节末尾还会再强调。有限维定义域上的线性映射自动连续，以及无穷维时的反例，见下面“有界与连续；有限维定义域上的线性映射”一节。

算子范数

定义（算子范数）：若 $u \in \mathcal{L}(E,F)$ 有界（即上一定理中的 (4)），令

$$\|u\| := \sup_{x\neq 0} \frac{\|u(x)\|}{\|x\|},$$

称为 u 的**范数**（**算子范数**）。由上一定理，连续线性映射恰好就是有界的线性映射，故通常称为**有界算子**。

这里默认 $E \neq \{0\}$，否则 $x \neq 0$ 的上确界是对空集取的。由 (4)，$\dfrac{\|u(x)\|}{\|x\|} \leq C$ 对所有 $x \neq 0$ 成立，所以这个上确界是有限的。前面“赋范空间的例子③”里的矩阵范数就是这里的 $\|u\|$。

注：对于有界的 u，

$$\|u\| = \underbrace{\sup_{x\neq 0} \frac{\|u(x)\|}{\|x\|}}_{M_1} = \underbrace{\sup_{\|x\|\leq 1} \|u(x)\|}_{M_2}$$
$$= \underbrace{\sup_{\|x\|=1} \|u(x)\|}_{M_3} = \underbrace{\inf\{C \geq 0 :\ \|u(x)\| \leq C\|x\|,\ \forall x \in E\}}_{M_4},$$

并且

$$\|u(x)\| \leq \|u\|\,\|x\|, \qquad \forall x \in E \text{ 成立}$$

（即 $\|u\|$ 是最小的 Lipschitz 常数）。

证明：记四个量依次为 M_1, M_2, M_3, M_4。

显然 $M_3 \leq M_2$（单位球面含于闭单位球）。

下证 $M_2 \leq M_1$：当 $0 < \|x\| \leq 1$，

$$\frac{\|u(x)\|}{\|x\|} \leq M_1 \implies \|u(x)\| \leq M_1\|x\| \leq M_1,$$

而 $x = 0$ 时 $\|u(0)\| = 0 \leq M_1$，所以两边对 $0 \leq \|x\| \leq 1$ 取 sup，得 $M_2 \leq M_1$。

下证 $M_1 \leq M_3$ 即可：$x \neq 0$ 时 $\dfrac{x}{\|x\|} \in S$（单位球面），由齐次性（取 $\lambda = \dfrac{1}{\|x\|}$），

$$\frac{\|u(x)\|}{\|x\|} = \left\|u\left(\frac{x}{\|x\|}\right)\right\| \leq M_3,$$

两边取 sup 得 $M_1 \leq M_3$。

$\because M_3 \leq M_2 \leq M_1 \leq M_3 \Rightarrow M_1 = M_2 = M_3$。

最后，$\|u(x)\| \leq M_1\|x\|$（$x \neq 0$ 时由 M_1 的定义，$x = 0$ 时两边都是 0），即 M_1 属于集合 $\{C \geq 0 : \|u(x)\| \leq C\|x\|\ \forall x\}$，于是其下确界 $M_4 \leq M_1$。反之，该集合中每个 C 都满足 $M_1 \leq C$，所以 $M_4 \geq M_1$，$\Rightarrow M_1 = M_4$。

至于 $\|\cdot\|$ 是最小的 Lipschitz 常数：若 C 是 u 的 Lipschitz 常数，即 $\|u(x)-u(y)\| \le C\|x-y\|$ 对一切 x,y 成立，取 $y=0_E$ 得 $\|u(x)\| \le C\|x\|$，所以 $C \ge M_4 = \|u\|$；反过来 $\|u(x)-u(y)\| = \|u(x-y)\| \le \|u\|\|x-y\|$，$\|u\|$ 本身就是 Lipschitz 常数。□

注：“u 有界”指 u 把单位球（或任何有界集）映成有界集，不是 $u(E)$ 有界。非零映射在 E 上必无界（$\|u(nx)\| = n\|u(x)\|$）。事实上，若 $\|x\| \le R$，则 $\|u(x)\| \le \|u\|R$，所以有界的 u 把任何有界集映成有界集；反过来，若 u 把单位球映成有界集，即 $M_2 < \infty$，则对 $x \ne 0$ 有 $\dfrac{\|u(x)\|}{\|x\|} = \left\|u\left(\dfrac{x}{\|x\|}\right)\right\| \le M_2$，满足 (4)。所以“(4) u 有界”与“u 把单位球映成有界集”是等价的；而 $u(E)$ 作为整个空间的像，只要 $u \ne 0$ 就无界。

有界与连续；有限维定义域上的线性映射

推论：$u \in \mathcal{L}(E,F)$ 连续 $\iff$ u 有界 $\iff$ $\|u\| < \infty$。

前一个等价就是连续性定理的 (1) $\Leftrightarrow$ (4)。后一个等价：$\|u\| < \infty$ 当且仅当存在 $C \ge 0$ 使 $\|u(x)\| \le C\|x\|$ 对一切 x 成立。（$\Rightarrow$ 取 $C = \|u\|$，由 $\|u(x)\| \le \|u\|\|x\|$；$\Leftarrow$ 由 $\|u\| = M_4 \le C$。）

定理（有限维定义域上的线性映射自动连续）：假设 E,F 是 $\mathbb{K}$ 上的赋范空间，$\dim E < \infty$，则 $\forall\, u \in \mathcal{L}(E,F)$，$u$ 有界。

证明：$E=\{0\}$ 时 $u=0$，结论平凡。否则取 E 的基底 $\{e_1,e_2,\dots,e_m\}$，$x = \sum_{i=1}^{m} x_i e_i$。定义 E 上的 p 范数

$$p(x) = \sum_{i=1}^{m} |x_i|.$$

由上一节“有限维赋范空间”的定理，E 上任意范数都与 p 等价，即存在 $\alpha,\beta > 0$ 使 $\alpha p \le \|\cdot\| \le \beta p$；取左边这个不等式，$\exists C = 1/\alpha > 0$ s.t. $p(x) \le C\|x\|$。于是

$$\begin{aligned}
\|u(x)\| &= \Big\| \sum_{i=1}^{m} x_i u(e_i) \Big\| \le \sum_{i=1}^{m} |x_i|\,\|u(e_i)\| \\
&\le \Big(\max_{1\le i\le m} \|u(e_i)\| \Big) \cdot p(x) \\
&\le \Big(C \max_{1\le i\le m} \|u(e_i)\| \Big) \cdot \|x\| \implies u \text{ 有界.} \qquad \square
\end{aligned}$$

其中第一个等号用到 u 的线性，第一个不等号是三角不等式与齐次性，第二个不等号是把每个 $\|u(e_i)\|$ 放大成它们的最大值。

注：无穷维时存在反例。

① $E = (c_{00}, \|\cdot\|_\infty)$，$u \in \mathcal{L}(E)$：$\forall\, x = (x_1,x_2,\dots,x_n,\dots) \in c_{00}$，

$$u(x) = (x_1, 2x_2, \dots, nx_n, \dots) \in c_{00}.$$

这里 c_{00} 是只有有限项非零的序列全体；u 把有限支撑序列映成有限支撑序列（第 n 项为零则 nx_n 也为零），所以 u 确实是 $c_{00} \to c_{00}$ 的线性映射。取 $e_k =$

$(0,\dots,0,\underbrace{1}_{\text{第}k\text{位置}},0,\dots,0,\dots)$，则 $u(e_k)=ke_k$，$\|e_k\|_\infty=1$，

$$\frac{\|u(e_k)\|}{\|e_k\|}=\frac{\|ke_k\|}{\|e_k\|}=k\longrightarrow\infty\quad(k\to\infty)\implies u\text{ 无界.}$$

② $E=\mathbb{R}[x]$，赋 $\|\cdot\|_\infty$（在 $[0,1]$ 上），$\mathbb{R}[x]\subset(C[0,1],\|\cdot\|_\infty)$，$u\in\mathcal{L}(\mathbb{R}[x],\mathbb{R}[x])$，$u(f)=f'$。取 $f_n(x)=x^n$，$\|f_n\|_\infty=1$，$\|f_n'\|_\infty=\|nx^{n-1}\|_\infty=n$（在 $[0,1]$ 上最大值在 $x=1$ 取得；$n=1$ 时 $f_1'\equiv1$），无界。**微分算子不连续。**

这两个例子说明上面的定理中“$\dim E<\infty$”这个条件不能去掉：由前面的连续性定理，无界就等价于不连续。

有界线性算子空间 $\mathcal{B}(E,F)$

定义：

$$\mathcal{B}(E,F):=\{u\in\mathcal{L}(E,F):\ \|u\|<\infty\},$$

即 $E\to F$ 的连续（有界，bounded）线性映射的全体。对于算子范数 $\|u\|=\sup\limits_{\|x\|=1}\|u(x)\|$，$\|\cdot\|$ 的确是 $\mathcal{B}(E,F)$ 上的一个范数。

证明：设 $u,v\in\mathcal{B}(E,F)$，$\lambda\in\mathbb{K}$。

- 封闭性与三角不等式：$\|(u+v)(x)\|\le\|u(x)\|+\|v(x)\|\le(\|u\|+\|v\|)\|x\|$，所以 $u+v$ 有界，且 $\|u+v\|\le\|u\|+\|v\|$。
- 齐次性：$\|(\lambda u)(x)\|=|\lambda|\|u(x)\|$，两边在 $\|x\|=1$ 上取 sup，得 $\|\lambda u\|=|\lambda|\|u\|$（所以 λu 也有界）。零映射有界，所以 $\mathcal{B}(E,F)$ 是 $\mathcal{L}(E,F)$ 的子空间。
- 正定性：$\|u\|\ge0$；若 $\|u\|=0$，则 $\|u(x)\|\le0\cdot\|x\|=0$ 对一切 x 成立，所以 $u=0$。

特殊情况：

(1) $\mathcal{B}(E,\mathbb{K})$ 称为 E 的**对偶空间** E^*，即

$$E^*=\mathcal{B}(E,\mathbb{K})=E\text{ 上连续线性泛函的全体.}$$

这里强调“连续”，是因为 $\mathcal{L}(E,\mathbb{K})$ 里的线性泛函不一定连续，E^* 只取其中连续的那些。

(2) $\mathcal{B}(E)=\mathcal{B}(E,E)$。

(3) $\dim E<\infty$，$\mathcal{L}(E,F)=\mathcal{B}(E,F)$（上一定理）：$\mathcal{B}(E,F)\subset\mathcal{L}(E,F)$ 总成立，反过来在 $\dim E<\infty$ 时每个线性映射都有界。

算子范数的计算

计算 $\|u\|$ 的套路：

- **Step 1**：先证明不等式

$$\|u(x)\| \le C\|x\|, \qquad \forall\, x \in E,$$

得到 $\|u\| \le C$。

- **Step 2**：再找一个 x 或数列 $\{x_n\}$ s.t. $\|u(x)\| = C\|x\|$（可取 $\|x\| = 1$，即 $\|u(x)\| = C$）或 $\|x_n\| = 1$ 且 $\lim \|u(x_n)\| = C$（或者 $x_n \neq 0$、$\lim \|u(x_n)\|/\|x_n\| = C$）。

两步合起来给出 $\|u\| = C$：Step 1 是上界 $\|u\| \le C$；Step 2 中 $\|u\| \ge \|u(x_n)\|/\|x_n\|$ 对每个 n 成立，令 $n \to \infty$ 得 $\|u\| \ge C$。

若 $\dim E < \infty$，则 $\bar{B} = \{x : \|x\| \le 1\}$ 是紧集，而 u 自动连续（上一定理），$x \mapsto \|u(x)\|$ 连续，于是 $\|u(x)\|$ 在 $\bar{B}$ 上可取到最大值（$\bar{B}$ 紧见“有限维赋范空间”一节推论 (3)），即 $\|u\| = M_2 = \max_{\|x\|\le 1} \|u(x)\|$：此时 Step 2 里总能找到一个具体的 x。下面会看到无穷维时（Volterra 算子）上确界可以取不到。

例 1：$\varphi : \mathbb{R}^2 \to \mathbb{R}^2$，

$$\varphi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 4 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

①赋予 $(\mathbb{R}^2, \|\cdot\|_1)$，即 $\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_1 = |x| + |y|$，希望计算 $\|\varphi\|$。

$$\left\| \varphi \begin{pmatrix} x \\ y \end{pmatrix} \right\|_1 = \left\| \begin{pmatrix} 2x - 3y \\ 4x + 7y \end{pmatrix} \right\|_1 = |2x - 3y| + |4x + 7y|$$

$$\le 6|x| + 10|y| \ \le\ 10\,(|x| + |y|) = 10 \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|$$

$\Longrightarrow \|\varphi\| \le 10$。第一个“$\le$”是 $|2x - 3y| \le 2|x| + 3|y|$，$|4x + 7y| \le 4|x| + 7|y|$ 相加；第二个“$\le$”是 $6|x| + 10|y| \le 10|x| + 10|y|$，等号当且仅当 $x = 0$，所以上界 10 只可能在 $x = 0$ 处达到。

另一方面，取 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$，$\left\| \binom{0}{1} \right\|_1 = 1$，得

$$\left\| \varphi \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -3 \\ 7 \end{pmatrix} \right\| = 10 \implies \|\varphi\| \ge \frac{\left\| \varphi \binom{0}{1} \right\|}{\left\| \binom{0}{1} \right\|} = 10.$$

所以 $\|\varphi\| = 10$（即最大列和：各列绝对值之和为 $2 + 4 = 6$、$3 + 7 = 10$）。

②赋以 $(\mathbb{R}^2, \|\cdot\|_\infty)$，$\|\varphi\| = 11$。

每个分量估计：

$$|2x - 3y| \le 2|x| + 3|y| \le 5\|x\|_\infty, \qquad |4x + 7y| \le 4|x| + 7|y| \le 11\|x\|_\infty,$$

$\Longrightarrow \|\varphi x\|_\infty \le 11\|x\|_\infty \Longrightarrow \|\varphi\|_\infty \le 11$。当 $x=\begin{pmatrix}1\\1\end{pmatrix}$ 时，$\varphi(1,1)=(2-3,\ 4+7)=(-1,11)$，$\|\varphi x\|_\infty = 11 = 11\|x\|_\infty$，$\Longrightarrow \|\varphi\|_\infty = 11$（即最大行和：各行绝对值之和为 $2+3=5$、$4+7=11$）。

一般情形：矩阵的算子范数

命题：设 $A=(a_{ij})_{n\times n}$，把它看作 $(\mathbb{K}^n,\|\cdot\|_1)\to(\mathbb{K}^n,\|\cdot\|_1)$ 与 $(\mathbb{K}^n,\|\cdot\|_\infty)\to(\mathbb{K}^n,\|\cdot\|_\infty)$ 的算子，则分别有

$$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}| \quad \text{（最大列和；第}j\text{ 列绝对值之和）},$$

$$\|A\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| \quad \text{（最大行和；第}i\text{ 行绝对值之和）}.$$

证明（最大列和 $\|A\|_1$）：记 $C_j=\sum_{i=1}^{n}|a_{ij}|$（第 j 列绝对值之和），$C=\max_j C_j$。则

$$\begin{aligned}\|Ax\|_1&=\sum_{i=1}^{n}\Big|\sum_{j=1}^{n}a_{ij}x_j\Big|\le\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}x_j|\\&=\sum_{j=1}^{n}|x_j|\Big(\sum_{i=1}^{n}|a_{ij}|\Big)=\sum_{j=1}^{n}|x_j|C_j\le C\sum_{j=1}^{n}|x_j|=C\|x\|_1,\end{aligned}$$

$\Rightarrow \|A\|_1\le C$。反之，取 j^* s.t. $C_{j^*}=C$。令 $x=e_{j^*}=(0,\dots,0,\underset{j^*}{1},0,\dots,0)^{\mathsf T}$（第 j^* 位为 1），则 Ae_{j^*} 就是 A 的第 j^* 列，$\|Ae_{j^*}\|_1=C_{j^*}=C=C\|e_{j^*}\|_1$。所以 $\|A\|_1=C$。

证明（最大行和 $\|A\|_\infty$）：记 $r_i=\sum_{j=1}^{n}|a_{ij}|$（第 i 行绝对值之和），$r=\max_i r_i$（$i=i^*$ 时 $r_{i^*}=r$）。$\forall x\in\mathbb{K}^n$，

$$|(Ax)_i|=\Big|\sum_{j=1}^{n}a_{ij}x_j\Big|\le\sum_{j=1}^{n}|a_{ij}||x_j|\le\Big(\sum_{j=1}^{n}|a_{ij}|\Big)\|x\|_\infty=r_i\|x\|_\infty\le r\|x\|_\infty,$$

所以 $\|A\|_\infty\le r$。$r=0$ 时 $A=0$，结论平凡；否则令

$$x_j=\begin{cases}\dfrac{\overline{a_{i^*j}}}{|a_{i^*j}|}, & a_{i^*j}\neq 0,\\[2mm] 0, & a_{i^*j}=0\end{cases}\qquad(\mathbb{K}=\mathbb{R}\text{ 时即 }\operatorname{sgn}a_{i^*j}),$$

则 $a_{i^*j}x_j=|a_{i^*j}|$，某个 $a_{i^*j}\neq 0$ 故 $\|x\|_\infty=1$，$(Ax)_{i^*}=\sum_j|a_{i^*j}|=r$，于是 $\|Ax\|_\infty\ge|(Ax)_{i^*}|=r=r\|x\|_\infty$，再结合 $\le$ 得 $\|Ax\|_\infty=r\|x\|_\infty$，$\|A\|_\infty=r$。$\#$

对照例 1：$A=\begin{pmatrix}2&-3\\4&7\end{pmatrix}$，最大列和 $\max\{6,10\}=10$，最大行和 $\max\{5,11\}=11$，与①②相符；取到最大值的向量分别是 $e_2=(0,1)^{\mathsf T}$ 与第二行 $(4,7)$ 的符号向量 $x=(1,1)^{\mathsf T}$。

欧氏范数：$\|A\|_2=\sqrt{\lambda_{\max}(A^*A)}$（$A^*$ 指 $\bar{A}^{\mathsf{T}}$）。

证明：A^*A 是 Hermite 半正定矩阵，有标准正交特征基 $v_1,\dots,v_n$，特征值 $\lambda_k\ge 0$。设 $x=\sum_k c_k v_k$，则 $\|Ax\|_2^2=\langle Ax,Ax\rangle=\langle A^*Ax,x\rangle=\sum_k\lambda_k|c_k|^2\le\lambda_{\max}\sum_k|c_k|^2=\lambda_{\max}\|x\|_2^2$，取 x 为最大特征值对应的 v_k 时等号成立。所以 $\|A\|_2=\sqrt{\lambda_{\max}(A^*A)}$。

例：Volterra 算子

Volterra 算子：$E=L^1(a,b)$，

$$(\mathcal{V}f)(x)=\int_a^x f(t)\,dt,\qquad \mathcal{V}:L^1(a,b)\to L^1(a,b).$$

$$\|\mathcal{V}f\|_1=\int_a^b|\mathcal{V}f(x)|\,dx=\int_a^b\Big|\int_a^x f(t)\,dt\Big|\,dx\ \le\ \int_a^b\Big(\int_a^x|f(t)|\,dt\Big)dx$$
$$\le\ \int_a^b\|f\|_1\,dx=(b-a)\|f\|_1,$$

$\Rightarrow\|\mathcal{V}\|\le b-a$。（第一个“$\le$”在 $f\ge 0$ 时取等；第二个“$\le$”无法取到。）

取 $f_n(x)=n\cdot 1_{[a,a+\frac{1}{n}]}(x)\Rightarrow\|f_n\|_1=1$，

$$(\mathcal{V}f_n)(x)=\begin{cases}n(x-a), & a<x<a+\frac{1}{n},\\ 1, & x\ge a+\frac{1}{n},\end{cases}$$

$\Rightarrow\|\mathcal{V}f_n\|_1=(b-a)-\dfrac{1}{2n}$，$\dfrac{\|\mathcal{V}f_n\|_1}{\|f_n\|_1}\to b-a\ (n\to\infty)\Rightarrow\|\mathcal{V}\|=b-a$。

具体地，$\int_a^{a+1/n}n(x-a)\,dx=\dfrac{n}{2}\cdot\dfrac{1}{n^2}=\dfrac{1}{2n}$，$\int_{a+1/n}^b 1\,dx=(b-a)-\dfrac{1}{n}$，相加为 $(b-a)-\dfrac{1}{2n}$（要求 $n>\dfrac{1}{b-a}$）。

上确界取不到：$\|\mathcal{V}\|=b-a$，但对任何 $f\ne 0$ 都有 $\|\mathcal{V}f\|_1<(b-a)\|f\|_1$。因为 $x\mapsto\int_a^x|f|$ 连续、$\le\|f\|_1$，且在 x 靠近 a 时趋于 $0<\|f\|_1$，所以在一个正测度集上严格小于 $\|f\|_1$，第二个“$\le$”的积分是严格不等号。这说明在无穷维空间中，$\|u\|=\sup$ 不一定能取到，对照前面有限维时 max 总能取到。

2.7 空间 $\mathcal{B}(E,F)$ 与对偶空间 E^*

$\mathcal{B}(E,F)$ 是赋范空间

命题：$\mathcal{B}(E,F)$ 是 $\mathcal{L}(E,F)$ 的子空间，算子范数是 $\mathcal{B}(E,F)$ 上的范数。

证明：$\forall u,v\in\mathcal{B}(E,F)$，$\lambda\in\mathbb{K}$，$u+\lambda v\in\mathcal{B}(E,F)$。事实上，

$$\|(u+\lambda v)(x)\|\le\|u(x)\|+|\lambda|\|v(x)\|\le\big(\|u\|+|\lambda|\|v\|\big)\|x\|,$$

所以 $u+\lambda v$ 有界，且 $\|u+\lambda v\| \le \|u\| + |\lambda|\|v\|$；零映射有界，故 $\mathcal{B}(E,F)$ 是 $\mathcal{L}(E,F)$ 的子空间。范数三条公理（三角不等式、齐次性、正定性）的验证见上一节对算子范数的讨论。

$\mathcal{B}(E,F)$ 的完备性

定理：设 E 是赋范空间，F 是 Banach 空间，则 $(\mathcal{B}(E,F), \|\cdot\|)$ 是 Banach 空间。

证明：设 $\{u_n\}$ 是 $\mathcal{B}(E,F)$ 中的 Cauchy 列，即 $\forall \varepsilon > 0$，$\exists N \in \mathbb{N}_+$，当 $m,n > N$ 时，$\|u_n - u_m\| < \varepsilon$。

Step 1：**逐点收敛**。$\forall x \in E$，

$$\begin{aligned}\|u_m(x) - u_n(x)\|_F &= \|(u_m - u_n)(x)\|_F \\ &\le \|u_m - u_n\|_{\mathcal{B}(E,F)} \|x\|_E \to 0 \qquad (m,n \to \infty),\end{aligned}$$

具体地，任给 $\varepsilon' > 0$，取 $\varepsilon = \varepsilon'/(\|x\|+1)$ 对应的 N，则 $m,n > N$ 时 $\|u_m(x) - u_n(x)\| < \varepsilon\|x\| < \varepsilon'$。所以 $\{u_n(x)\}$ 是 F 中的 Cauchy 列，由 F 完备，令 $u(x) = \lim_{n\to\infty} u_n(x)$，这就定义了 $u : E \to F$。

Step 2：u **线性**。$\forall x, y \in E$，$\lambda \in \mathbb{K}$，

$$u(\lambda x + y) = \lim_{n\to\infty} u_n(\lambda x + y) = \lim_{n\to\infty} \big(\lambda u_n(x) + u_n(y)\big) = \lambda u(x) + u(y).$$

Step 3：u **有界**。Cauchy 列有界：取 $\varepsilon = 1$ 得对应的 N，则 $n > N$ 时 $\|u_n\| \le \|u_n - u_{N+1}\| + \|u_{N+1}\| < 1 + \|u_{N+1}\|$。令 $M = \max\{\|u_1\|, \dots, \|u_N\|, 1 + \|u_{N+1}\|\}$（若 $M = 0$ 就换成任意正数），则 $\|u_n\| \le M$，$\forall n$。于是

$$\|u_n(x)\| \le \|u_n\|\|x\| \le M\|x\|,$$

令 $n \to \infty$，由范数的连续性（$u_n(x) \to u(x)$ 蕴含 $\|u_n(x)\| \to \|u(x)\|$）得 $\|u(x)\| \le M\|x\|$，故 $\|u\| \le M$，即 $u \in \mathcal{B}(E,F)$。

Step 4：$\|u_n - u\| \to 0$。由 Step 3，$u_n - u \in \mathcal{B}(E,F)$，其范数有意义。$\forall \varepsilon > 0$，取 N 使 $m,n > N$ 时 $\|u_m - u_n\| < \varepsilon$，则对 $x \in E$，

$$\|u_m(x) - u_n(x)\| \le \varepsilon\|x\|.$$

固定 n, x，令 $m \to \infty$（同样用范数连续：$\|u_m(x) - u_n(x)\| \to \|u(x) - u_n(x)\|$），得

$$\|u(x) - u_n(x)\| \le \varepsilon\|x\| \qquad (\forall n > N,\ x \in E).$$

故 $\forall x \ne 0$，$\dfrac{\|u_n(x) - u(x)\|}{\|x\|} \le \varepsilon$，

$$\|u_n - u\| = \sup_{x\ne 0} \frac{\|u_n(x) - u(x)\|}{\|x\|} \le \varepsilon, \qquad \forall n > N.$$

注意这里的 N 与 x 无关，这正是从“逐点收敛”升级为“依算子范数收敛”的关键。从而 $\|u_n - u\| \to 0$。#

推论：

(1) $\dim F < \infty$ 时，$\mathcal{B}(E,F)$ 是 Banach 空间。特别地，$E^* = \mathcal{B}(E,\mathbb{K})$ 是 Banach 空间。

(2) 若 E 是 Banach 空间，则 $\mathcal{B}(E)$ 是 Banach 空间。

(3) $\dim E < \infty$ 时，$\mathcal{L}(E,F) = \mathcal{B}(E,F)$。

对 (1)：有限维赋范空间都是 Banach 空间，再用定理；$\mathbb{K}$ 完备，所以 E^* 完备，而且不要求 E 完备。(2) 取 $F = E$ 即可。(3) 即“有限维定义域上的线性映射自动连续”，前面已证。

复合算子

定理（复合算子）：设 E,F,G 是赋范空间，$u \in \mathcal{B}(E,F)$，$v \in \mathcal{B}(F,G)$，则 $v \circ u \in \mathcal{B}(E,G)$，且 $\|v \circ u\| \le \|v\|\,\|u\|$。

证明：

$$\|v \circ u(x)\| \le \|v\|\,\|u(x)\| \le \|v\|\,\|u\|\,\|x\|. \qquad \square$$

注：复合关于算子范数是 Lipschitz 的。固定 v，$u \mapsto v \circ u$ 是 $\|v\|$-Lipschitz 的；固定 u，$v \mapsto v \circ u$ 是 $\|u\|$-Lipschitz 的：

$$\begin{aligned}\|v \circ u_1 - v \circ u_2\| &= \|v \circ (u_1 - u_2)\| \le \|v\|\,\|u_1 - u_2\|,\\ \|v_1 \circ u - v_2 \circ u\| &= \|(v_1 - v_2) \circ u\| \le \|v_1 - v_2\|\,\|u\|.\end{aligned}$$

第一行的等号用 v 的线性；第二行的等号是因为 $(v_1 \circ u - v_2 \circ u)(x) = v_1(u(x)) - v_2(u(x)) = ((v_1 - v_2) \circ u)(x)$。

由此，复合运算 $(u,v) \mapsto v \circ u$ 关于算子范数是连续的：若 $u_n \to u$ in $\mathcal{B}(E,F)$，$v_n \to v$ in $\mathcal{B}(F,G)$，则

$$\|v_n u_n - vu\| = \|v_n(u_n - u) + (v_n - v)u\| \le \underbrace{\|v_n\|}_{\text{有界}} \cdot \underbrace{\|u_n - u\|}_{\to 0} + \|v_n - v\| \cdot \|u\| \to 0.$$

这里 $\|v_n\| \le \|v\| + \|v_n - v\|$ 且 $\|v_n - v\| \to 0$，所以 $\{\|v_n\|\}$ 有界；等号是加减 $v_n u$，不等号用三角不等式与复合定理。

推论：$u \in \mathcal{B}(E)$ 时，$u^k = u \circ u \circ \cdots \circ u \in \mathcal{B}(E)$，且 $\|u^k\| \le \|u\|^k$。

证明：对 k 归纳：$u^{k+1} = u \circ u^k$，由复合定理 $\|u^{k+1}\| \le \|u\|\|u^k\| \le \|u\|^{k+1}$。$\square$

注：不等号可以严格。对 $L^1(a,b)$ 上的 Volterra 算子 $\mathcal{V}$，$\|\mathcal{V}^n\| \le \dfrac{(b-a)^n}{n!}$，而 $\|\mathcal{V}\|^n = (b-a)^n$。

这是因为，先用归纳法证明

$$(\mathcal{V}^n f)(x)=\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}f(t)\,dt \qquad (n\geq 1).$$

$n=1$ 就是定义。若对 n 成立，则 $(\mathcal{V}^{n+1}f)(x)=\int_a^x\Big(\int_a^s \frac{(s-t)^{n-1}}{(n-1)!}f(t)\,dt\Big)ds$，交换积分次序（Fubini，被积函数可积）得 $\int_a^x f(t)\Big(\int_t^x \frac{(s-t)^{n-1}}{(n-1)!}\,ds\Big)dt=\int_a^x \frac{(x-t)^n}{n!}f(t)\,dt$，归纳完成。于是

$$\begin{aligned}\|\mathcal{V}^n f\|_1 &\leq \int_a^b\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}|f(t)|\,dt\,dx\\ &=\int_a^b |f(t)|\Big(\int_t^b \frac{(x-t)^{n-1}}{(n-1)!}\,dx\Big)dt\\ &=\int_a^b |f(t)|\frac{(b-t)^n}{n!}\,dt\leq \frac{(b-a)^n}{n!}\|f\|_1,\end{aligned}$$

所以 $\|\mathcal{V}^n\|\leq\dfrac{(b-a)^n}{n!}$。而 $\|\mathcal{V}\|^n=(b-a)^n$（上一节例），$n\geq 2$ 时 $n!>1$，$\dfrac{(b-a)^n}{n!}<(b-a)^n$ $(b>a)$，所以 $\|\mathcal{V}^n\|<\|\mathcal{V}\|^n$。

线性泛函的核

定义：$\varphi\in\mathcal{L}(E,F)$，

$$\operatorname{Ker}\varphi=\varphi^{-1}(0_F)=\{x\in E:\varphi(x)=0\},$$

它是 E 的子空间。

命题（线性泛函连续 $\Longleftrightarrow$ 核闭）：设 $\varphi\in\mathcal{L}(E,\mathbb{K})$，则 φ 连续 $\Longleftrightarrow$ $\operatorname{Ker}\varphi$ 在 E 中闭。

证明：$(\Rightarrow)$ $\{0\}$ 是 $\mathbb{K}$ 的闭集，$\operatorname{Ker}\varphi=\varphi^{-1}(\{0\})$ 是闭集 $\{0\}$ 在连续映射 φ 下的原像，故闭。

$(\Leftarrow)$ 已知 $\operatorname{Ker}\varphi$ 闭，反设 φ 不连续（无界）。无界意味着不存在 C 使 $|\varphi(x)|\leq C\|x\|$ 对一切 x 成立，所以对每个 n 存在 x_n' 使 $|\varphi(x_n')|>n\|x_n'\|$，此时 $x_n'\neq 0$，令 $x_n=x_n'/\|x_n'\|$，则 $\|x_n\|=1$，$|\varphi(x_n)|>n$；特别 $\varphi(x_n)\neq 0$。又因为 φ 无界，所以 $\varphi\neq 0$，取 $z\in E$ 使 $\varphi(z)=1$，令 $y_n=z-\dfrac{x_n}{\varphi(x_n)}$。则 $\varphi(y_n)=\varphi(z)-\dfrac{\varphi(x_n)}{\varphi(x_n)}=1-1=0\Rightarrow y_n\in\operatorname{Ker}\varphi$。

$$\left\|\frac{x_n}{\varphi(x_n)}\right\|=\frac{1}{|\varphi(x_n)|}\leq\frac{1}{n}\to 0,$$

故 $y_n\to z$。$\operatorname{Ker}\varphi$ 闭，所以 $z\in\operatorname{Ker}\varphi$，但 $\varphi(z)=1\neq 0$，矛盾。$\#$

注：

(1) 对于一般的线性算子，核闭 $\not\Rightarrow$ 连续。如 c_{00}（取 $\|\cdot\|_\infty$）上的 $u:(x_n)\mapsto(nx_n)$，$\operatorname{Ker}u=\{x:nx_n=0\ \forall n\}=\{0\}$ 是闭集，但 $u(e_k)=ke_k$，$\|u(e_k)\|/\|e_k\|=k$，所以 u 无界。因此上面的命题只对线性泛函（值域一维）成立。

(2) 不连续的线性泛函的核是稠密的真子空间（$\operatorname{Ker}\varphi$ 不闭，$\overline{\operatorname{Ker}\varphi}=E$）。

证明：φ 不连续，所以 $\varphi\neq0$，取 $\varphi(z)=1$。任意 $x=(x-\varphi(x)z)+\varphi(x)z$，前一项在 $\operatorname{Ker}\varphi$ 中，所以 $E=\operatorname{Ker}\varphi+\mathbb{K}z$，$\operatorname{Ker}\varphi$ 是真子空间（$z\notin\operatorname{Ker}\varphi$）且余维数为 1。闭包 $\overline{\operatorname{Ker}\varphi}$ 是包含它的子空间，而包含 $\operatorname{Ker}\varphi$ 的子空间只有 $\operatorname{Ker}\varphi$ 与 E（任何多出来的向量都与 $\operatorname{Ker}\varphi$ 张成 E）。$\operatorname{Ker}\varphi$ 不闭，所以闭包 $\neq\operatorname{Ker}\varphi$，只能是 E。

(3) 连续线性泛函 $\varphi\neq0$ 的核是闭超平面，且

$$|\varphi(x)|=\|\varphi\|\cdot d(x,\operatorname{Ker}\varphi).$$

证明：核是闭集由上面的命题。“$\le$”：对任意 $y\in\operatorname{Ker}\varphi$，$|\varphi(x)|=|\varphi(x-y)|\le\|\varphi\|\|x-y\|$，对 y 取下确界得 $|\varphi(x)|\le\|\varphi\|d(x,\operatorname{Ker}\varphi)$。“$\ge$”：任给 $\varepsilon\in(0,1)$，取 $\|w\|=1$ 使 $|\varphi(w)|\ge(1-\varepsilon)\|\varphi\|>0$。则 $x-\dfrac{\varphi(x)}{\varphi(w)}w\in\operatorname{Ker}\varphi$，所以 $d(x,\operatorname{Ker}\varphi)\le\left\|\dfrac{\varphi(x)}{\varphi(w)}w\right\|=\dfrac{|\varphi(x)|}{|\varphi(w)|}\le\dfrac{|\varphi(x)|}{(1-\varepsilon)\|\varphi\|}$，令 $\varepsilon\to0$ 得 $\|\varphi\|d(x,\operatorname{Ker}\varphi)\le|\varphi(x)|$。

延拓定理

定理（延拓定理）：设 E 是赋范空间，$G\subset E$ 是稠密子空间，F 是 Banach 空间。则任意 $u\in\mathcal{B}(G,F)$ 可以唯一地延拓为 $\tilde u\in\mathcal{B}(E,F)$，且 $\|\tilde u\|=\|u\|$。

证明：定义 $\tilde u$。$\forall x\in E$，由稠密性，取 $\{x_n\}\subset G$，$x_n\to x$，则 $\{x_n\}$ 是 Cauchy 列，且

$$\|u(x_m)-u(x_n)\|=\|u(x_m-x_n)\|\le\|u\|\,\|x_m-x_n\|\to0.$$

故 $\{u(x_n)\}$ 是 F 中的 Cauchy 列。这里需要 F 完备，才能保证它收敛：令 $\tilde u(x):=\lim_{n\to\infty}u(x_n)$。若 G 中另一列 $\tilde x_n\to x$，则

$$\|u(x_n)-u(\tilde x_n)\|\le\|u\|\underbrace{\|x_n-\tilde x_n\|}_{\to0}\to0,$$

所以 $\{u(x_n)\}$ 与 $\{u(\tilde x_n)\}$ 极限相同，$\tilde u$ 是良定义的。对 $x\in G$ 取常数列 $x_n=x$，得 $\tilde u(x)=u(x)$，即 $\tilde u$ 在 G 上与 u 相同。

$\tilde u$ **线性**。设 $x,y\in E$，$\lambda\in\mathbb{K}$，取 G 中 $x_n\to x$，$y_n\to y$，则 $\lambda x_n+y_n\in G$，$\lambda x_n+y_n\to\lambda x+y$，于是

$$\tilde u(\lambda x+y)=\lim_{n\to\infty}u(\lambda x_n+y_n)=\lim_{n\to\infty}\big(\lambda u(x_n)+u(y_n)\big)=\lambda\tilde u(x)+\tilde u(y).$$

$\tilde u$ **有界且** $\|\tilde u\|=\|u\|$。设 $x\in E$，取 G 中 $x_n\to x$，由范数的连续性，

$$\|\tilde u(x)\|=\Big\|\lim_{n\to\infty}u(x_n)\Big\|=\lim_{n\to\infty}\|u(x_n)\|\le\lim_{n\to\infty}\|u\|\,\|x_n\|=\|u\|\,\|x\|,$$

所以 $\tilde{u}\in\mathcal{B}(E,F)$ 且 $\|\tilde{u}\|\le\|u\|$。另一方面，$\tilde{u}$ 在 G 上与 u 相同，所以

$$\|\tilde{u}\|\ge\sup_{x\in G,x\neq0}\frac{\|\tilde{u}(x)\|}{\|x\|}=\sup_{x\in G,x\neq0}\frac{\|u(x)\|}{\|x\|}=\|u\|,$$

故 $\|\tilde{u}\|=\|u\|$。

唯一性。若 $\hat{u}\in\mathcal{B}(E,F)$ 也在 G 上与 u 相同，则对 $x\in E$ 取 G 中 $x_n\to x$，由 $\hat{u}$ 连续，$\hat{u}(x)=\lim\hat{u}(x_n)=\lim u(x_n)=\tilde{u}(x)$，所以 $\hat{u}=\tilde{u}$ （两个连续映射在稠密集上相同则处处相同）。#

2.8 延拓定理的应用；同构与 Neumann 级数

延拓定理的回顾与应用

回顾（延拓定理）：设 E 是赋范空间，$G\subset E$ 是稠密子空间，F 是 Banach 空间，则任意 $u\in\mathcal{B}(G,F)$ 可以唯一地延拓为 $\tilde{u}\in\mathcal{B}(E,F)$（即 $\tilde{u}|_G=u$），且 $\|\tilde{u}\|=\|u\|$。

证明：存在性与范数不变的构造同上一节：对 $x\in E$，取 G 中的 $x_n\to x$，则 $\|u(x_n)-u(x_m)\|\le\|u\|\,\|x_n-x_m\|\to0$，$\{u(x_n)\}$ 是 F 中的 Cauchy 列，由 F 完备，令 $\tilde{u}(x)=\lim u(x_n)$；这个极限与所取的列无关，$\tilde{u}$ 线性、有界，且 $\|\tilde{u}\|=\|u\|$（细节见下面注①与上一节）。

唯一性：设 $\tilde{u},\tilde{u}'$ 是 u 的两个连续延拓，它们在稠密集 G 上相等。由连续性，它们处处相等：

$$\{x\in E:\ \tilde{u}(x)=\tilde{u}'(x)\}=\mathrm{Ker}(\tilde{u}-\tilde{u}')$$

是闭集（连续映射 $\tilde{u}-\tilde{u}'$ 下闭集 $\{0\}$ 的原像），并且包含 G。而包含稠密集 G 的闭集必包含 $\overline{G}=E$，故它等于 E，即 $\tilde{u}=\tilde{u}'$。#

直观：u 只在 E 的一个稠密子空间 G 上已知。G 里的点可以任意逼近 E 中的点，而有界线性算子对逼近“不敏感”（$\|u(a)-u(b)\|\le\|u\|\|a-b\|$），所以 u 在 G 上的值已经把它在 E 上的值**完全决定了**：对 $x\in E$，取 G 中的 $x_n\to x$，则 $u(x_n)$ 必须收敛，且极限只能是 $\tilde{u}(x)$。F 完备保证这个极限存在，稠密性保证每个 x 都能这样逼近。下面三条是这个定理最常见的三种用法。

注（延拓定理的用法）：

①它是“一致连续映射的延拓”的线性版本

非线性的版本：设 X 是度量空间，$D\subset X$ 稠密，Y 是完备度量空间，$f:D\to Y$ 一致连续，则 f 可以唯一地延拓为连续映射 $\tilde{f}:X\to Y$。

证明的思路与上面完全一样：对 $x\in X$，取 D 中的 $x_n\to x$，则 $\{x_n\}$ 是 Cauchy 列，由 f 的一致连续性，$\{f(x_n)\}$ 是 Y 中的 Cauchy 列（$\forall\varepsilon>0$，$\exists\delta>0$ 使 $d(a,b)<\delta\Rightarrow$

$d(f(a),f(b))<\varepsilon$；当 m,n 充分大时 $d(x_n,x_m)<\delta$），所以收敛，令 $\tilde f(x)=\lim f(x_n)$。若 D 中另一列 $x_n'\to x$，则 $d(x_n,x_n')\to 0$，同样由一致连续得 $d(f(x_n),f(x_n'))\to 0$，所以极限与所取的列无关。

这里“一致连续”不能换成“连续”。例如 $X=[0,1]$，$D=(0,1]$ 在 X 中稠密，$f(x)=\dfrac{1}{x}$ 在 D 上连续（$Y=\mathbb{R}$ 完备），但无法连续地延拓到 $x=0$：$f(x)\to+\infty$。

线性的版本：有界线性算子 u 满足 $\|u(a)-u(b)\|\le\|u\|\|a-b\|$，是 Lipschitz 的，当然一致连续，所以由上面的版本，u 有唯一的连续延拓 $\tilde u$。延拓定理比它多出两条结论：

- $\tilde u$ **线性**：设 $x_n\to x$，$y_n\to y$（都在 G 中），则 $\lambda x_n+y_n\in G$ 且 $\lambda x_n+y_n\to\lambda x+y$，于是 $\tilde u(\lambda x+y)=\lim u(\lambda x_n+y_n)=\lim\big(\lambda u(x_n)+u(y_n)\big)=\lambda\tilde u(x)+\tilde u(y)$；
- **范数不变**：$\|\tilde u(x)\|=\lim\|u(x_n)\|\le\|u\|\lim\|x_n\|=\|u\|\|x\|$，所以 $\|\tilde u\|\le\|u\|$；反过来 $\tilde u$ 在 G 上等于 u，所以 $\|\tilde u\|\ge\|u\|$。

定理中的三个条件缺一不可：

- G **稠密**：否则延拓不唯一。例如 $E=\mathbb{R}^2$，$G=\{(x,0)\}$，$u(x,0)=x$，则对任意 $c\in\mathbb{R}$，$\tilde u_c(x,y)=x+cy$ 都是 u 的连续线性延拓。
- F **完备**：否则延拓可能不存在。例如 $E=(c_0,\|\cdot\|_\infty)$，$G=c_{00}$（在 E 中稠密），$F=(c_{00},\|\cdot\|_\infty)$（不完备），$u=\mathrm{Id}:G\to F$ 是有界线性算子（$\|u\|=1$）。若有连续延拓 $\tilde u:c_0\to c_{00}$，对 $x\in c_0\backslash c_{00}$ 取 $x_n\in c_{00}$，$x_n\to x$，则 $\tilde u(x)=\lim u(x_n)=\lim x_n=x$，但 $x\notin c_{00}$，矛盾。
- u **有界**：否则延拓可能不存在。例如 $E=(c_0,\|\cdot\|_\infty)$，$G=c_{00}$，$F=\mathbb{K}$，$u(x)=\sum_n nx_n$（G 中的 x 只有有限项非零，和有意义）是线性的，但 $u(e_k)=k$，$\|e_k\|_\infty=1$，u 无界。若有连续延拓 $\tilde u$，则 $|\tilde u(e_k)|\le\|\tilde u\|\,\|e_k\|=\|\tilde u\|$ 对一切 k 成立，与 $\tilde u(e_k)=u(e_k)=k$ 矛盾。

②“先在好函数上定义，再用稠密性延拓”

“好函数”（连续函数、有限和、紧支集函数、$L^1\cap L^2$ 中的函数等）上的算子常常容易定义、容易估计，但我们真正想要的是定义在整个（完备的）空间上的算子。通用的做法分四步：

(i) 找 E 的一个稠密子空间 G，使得在 G 上可以直接写出算子 u 的定义；

(ii) 验证 u 在 G 上线性，且值落在 Banach 空间 F 中；

(iii) 证明先验估计 $\|u(x)\|_F\le C\|x\|_E$，$\forall x\in G$——注意这里的范数 $\|\cdot\|_E$ 必须是使 G 稠密的那个范数；

(iv) 由延拓定理，u 唯一地延拓为 $\tilde u\in\mathcal{B}(E,F)$，且 $\|\tilde u\|\le C$。

例 1（积分）：$E=L^1(0,1)$，$G=C[0,1]$（在 $L^1(0,1)$ 中稠密），$F=\mathbb{K}$，$u(f)=\int_0^1 f(x)\,dx$。线性是显然的，且

$$|u(f)|=\left|\int_0^1 f\,dx\right|\le\int_0^1|f|\,dx=\|f\|_1,$$

所以 $u\in\mathcal{B}(G,\mathbb{K})$，$\|u\|\le 1$；取 $f\equiv 1$ 得 $\|u\|=1$。延拓定理给出 $L^1(0,1)$ 上唯一的连续线性泛函 $\tilde{u}$，$\|\tilde{u}\|=1$，对 $f\in L^1$ 它就是 $\lim\int_0^1 f_n\,dx$（$f_n\in C[0,1]$，$f_n\to f$ 于 L^1），即 f 的积分。

例 2（Fourier 变换）：对 $f\in L^1(\mathbb{R})$，定义

$$(\mathcal{F}f)(\xi)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}f(x)\,e^{-ix\xi}\,dx.$$

但对 $f\in L^2(\mathbb{R})$，这个积分未必有意义，因为 f 未必可积，如 $f(x)=\dfrac{1}{1+|x|}\in L^2\setminus L^1$。于是取 $G=L^1\cap L^2$，$E=F=L^2(\mathbb{R})$：

- G 在 L^2 中稠密：对 $f\in L^2$，$f_n=f\cdot 1_{[-n,n]}\in L^1\cap L^2$（由 Cauchy–Schwarz，$\int_{-n}^{n}|f|\le\sqrt{2n}\,\|f\|_2$），且 $f_n\to f$ 于 L^2（控制收敛定理）；
- 对 $f\in L^1\cap L^2$，Plancherel 等式 $\|\mathcal{F}f\|_2=\|f\|_2$ 成立，所以 $\mathcal{F}$ 在 G 上是 L^2 范数下的有界线性算子（等距）。

由延拓定理，$\mathcal{F}$ 可以唯一地延拓为 $L^2(\mathbb{R})\to L^2(\mathbb{R})$ 的有界线性算子，仍记为 $\mathcal{F}$，且 $\|\mathcal{F}\|=1$、$\|\mathcal{F}f\|_2=\|f\|_2$。具体地，$\mathcal{F}f=\lim_{n\to\infty}\mathcal{F}(f\cdot 1_{[-n,n]})$（极限在 L^2 中取）。

③赋范空间的完备化

定义：设 E 是赋范空间。若 $\hat{E}$ 是 Banach 空间，$J:E\to\hat{E}$ 是线性等距嵌入（$\|J(x)\|=\|x\|$），且 $J(E)$ 在 $\hat{E}$ 中稠密，则称 $(\hat{E},J)$ 是 E 的**完备化**。通常把 E 与 $J(E)$ 等同，说“E 在 $\hat{E}$ 中稠密”。

存在性（构造）：设 $\mathcal{C}$ 是 E 中全体 Cauchy 列构成的向量空间（按项加法、数乘），$\mathcal{N}\subset\mathcal{C}$ 是全体收敛到 0 的序列构成的子空间，令 $\hat{E}=\mathcal{C}/\mathcal{N}$，范数 $\big\|[\{x_n\}]\big\|:=\lim\limits_{n\to\infty}\|x_n\|$。这个极限存在（$\big|\|x_n\|-\|x_m\|\big|\le\|x_n-x_m\|$，所以 $\{\|x_n\|\}$ 是 $\mathbb{R}$ 中的 Cauchy 列），并且与代表元的选取无关（改变 $\{x_n\}$ 一个零序列，$\|x_n\|$ 的极限不变）。$J(x)=[\{x,x,x,\dots\}]$ 显然是线性等距；$J(E)$ 稠密，因为对 $[\{x_n\}]\in\hat{E}$ 有 $\|[\{x_n\}]-J(x_k)\|=\lim_n\|x_n-x_k\|\to 0$（$k\to\infty$，由 Cauchy 性）；完备性则由稠密性加上一个对角线论证得到。*这正是用 $\mathbb{Q}$ 的 Cauchy 列构造 $\mathbb{R}$ 的做法。*

唯一性（用延拓定理）：设 $(\hat{E}_1,J_1)$，$(\hat{E}_2,J_2)$ 是 E 的两个完备化。线性映射 $J_2\circ J_1^{-1}:J_1(E)\to\hat{E}_2$ 是等距的（特别地有界，范数 1），定义域 $J_1(E)$ 在 $\hat{E}_1$ 中稠密，$\hat{E}_2$ 是 Banach 空间，所以由延拓定理，它延拓为 $T_1\in\mathcal{B}(\hat{E}_1,\hat{E}_2)$，$\|T_1\|=1$；并且 T_1 仍是等距：对 $x\in\hat{E}_1$ 取 $J_1(E)$ 中的 $x_n\to x$，$\|T_1x\|=\lim\|T_1x_n\|=\lim\|x_n\|=\|x\|$。同理有等距的

$T_2 : \hat{E}_2 \to \hat{E}_1$。$T_2T_1$ 在稠密集 $J_1(E)$ 上是恒等映射，由延拓的唯一性，$T_2T_1 = \mathrm{Id}_{\hat{E}_1}$；同理 $T_1T_2 = \mathrm{Id}_{\hat{E}_2}$。所以 T_1 是等距同构，完备化在等距同构意义下唯一，因此可以说“*the*”完备化。

泛性质：设 F 是 Banach 空间，$u \in \mathcal{B}(E,F)$。则 $u \circ J^{-1} : J(E) \to F$ 是有界线性算子，定义域稠密于 $\hat{E}$，由延拓定理，u 唯一地延拓为 $\hat{u} \in \mathcal{B}(\hat{E},F)$（$\hat{u} \circ J = u$），且 $\|\hat{u}\| = \|u\|$。也就是说：从 E 出发的有界算子，到 *Banach* 空间里时，都可以无损地搬到 E 的完备化上。

例：

- $\mathbb{Q}$（带绝对值）的完备化是 $\mathbb{R}$。
- $(C[0,1], \|\cdot\|_p)$（$1 \le p < \infty$）的完备化是 $L^p(0,1)$：$(C[0,1], \|\cdot\|_p)$ 不完备（前面已证），$L^p(0,1)$ 是 Banach 空间（Riesz–Fischer），包含映射 $C[0,1] \hookrightarrow L^p(0,1)$ 是线性等距，并且 $C[0,1]$ 在 $L^p(0,1)$ 中稠密（$1 \le p < \infty$），所以 ($L^p(0,1)$, 包含映射) 满足完备化的定义。由唯一性，它就是 $(C[0,1], \|\cdot\|_p)$ 的完备化。这也解释了为什么 L^p 的元素是“函数的等价类”，而不是真正的函数：它们是 $C[0,1]$ 中 $\|\cdot\|_p$-Cauchy 列的“极限”。
- $p = \infty$ 时 $(C[0,1], \|\cdot\|_\infty)$ 本身就完备，完备化是它自己，所以上面的结论要求 $p < \infty$。

同构

定义（同构）：设 E, F 是赋范空间。若 $u \in \mathcal{B}(E,F)$ 是双射且 $u^{-1} \in \mathcal{B}(F,E)$，称 u 是 E 到 F 的（拓扑）**同构**；若还有 $\|u(x)\| = \|x\|$，称为**等距同构**。存在同构时称 E 与 F 同构，记 $E \cong F$。

注：

(1) u 是同构 $\iff$ u 是线性满射，且 $\exists c, C > 0$ s.t.

$$c\|x\| \le \|u(x)\| \le C\|x\|, \qquad \forall x \in E.$$

证明：($\Rightarrow$) 设 u 是同构（$E \neq \{0\}$）。u 是线性满射；取 $C = \|u\|$，则 $\|u(x)\| \le C\|x\|$。又

$$\|x\| = \|u^{-1}(u(x))\| \le \|u^{-1}\|\,\|u(x)\|,$$

取 $c = \dfrac{1}{\|u^{-1}\|}$，得 $c\|x\| \le \|u(x)\|$（$u^{-1} \neq 0$ 因为 $E \neq \{0\}$）。

($\Leftarrow$) 左边不等式给出 u 是单射（$u(x) = 0 \Rightarrow \|x\| = 0$），所以 u 是线性双射，从而 u^{-1} 也线性；右边不等式给出 $u \in \mathcal{B}(E,F)$。对 $y \in F$，令 $x = u^{-1}(y)$，则 $c\|u^{-1}(y)\| \le \|u(u^{-1}(y))\| = \|y\|$，即 $\|u^{-1}(y)\| \le \dfrac{1}{c}\|y\|$，所以 $u^{-1} \in \mathcal{B}(F,E)$，并且 $\|u^{-1}\| \le \dfrac{1}{c}$。□

(2) 同构保持完备性、可分性、紧性、闭子空间；同维数的有限维赋范空间都同构。

证明：设 $u:E\to F$ 是同构。由 (1) 或定理，u,u^{-1} 都是 Lipschitz 的，特别地 u 是同胚，而紧性、可分性是拓扑性质，被同胚保持。完备性：设 E 完备，$\{y_n\}$ 是 F 中的 Cauchy 列，则 $\|u^{-1}(y_n)-u^{-1}(y_m)\|\le\|u^{-1}\|\,\|y_n-y_m\|$，所以 $x_n=u^{-1}(y_n)$ 是 E 中的 Cauchy 列，$x_n\to x$，于是 $y_n=u(x_n)\to u(x)$，F 完备；F 完备推出 E 完备同理（对 u^{-1} 用同样的论证）。闭子空间：线性映射把子空间映成子空间，同胚把闭集映成闭集，所以 E 的闭子空间 M 的像 $u(M)$ 是 F 的闭子空间。最后，设 $\dim E=\dim F=m<\infty$，取基 $\{e_i\}$、$\{f_i\}$，线性映射 $u(e_i)=f_i$ 是线性双射；由有限维定义域上的线性映射自动连续，$u\in\mathcal{B}(E,F)$，u^{-1} 的定义域 F 也是有限维，所以 $u^{-1}\in\mathcal{B}(F,E)$，$u$ 是同构。□

(3) **Banach 逆算子定理**（后述）：Banach 空间之间的连续线性双射自动是同构（u^{-1} 的连续性可得）。

(4) $E=F$ 时，约定 $u^0=\mathrm{Id}_E$，$u^{n+1}=u\circ u^n$。

Neumann 级数与可逆算子

定理（Neumann 级数）：设 E 是 Banach 空间，$u\in\mathcal{B}(E)$，$\|u\|<1$。则 $\mathrm{Id}-u$ 在 $\mathcal{B}(E)$ 中可逆，且

$$(\mathrm{Id}-u)^{-1}=\sum_{n=0}^{\infty}u^n,\qquad\|(\mathrm{Id}-u)^{-1}\|\le\frac{1}{1-\|u\|}.$$

即 $(\mathrm{Id}-u)^{-1}\in\mathcal{B}(E)$，故 $\mathrm{Id}-u$ 是一个同构映射。

证明：第一步：级数收敛。由 $\|u^n\|\le\|u\|^n$，

$$\sum_{n=0}^{\infty}\|u^n\|\le\sum_{n=0}^{\infty}\|u\|^n=\frac{1}{1-\|u\|}<\infty,$$

所以级数 $\sum_{n=0}^{\infty}u^n$ 在 $\mathcal{B}(E)$ 中绝对收敛。E 完备 $\Rightarrow\mathcal{B}(E)$ 完备，因此 $\sum_{n=0}^{\infty}u^n$ 在 $\mathcal{B}(E)$ 中收敛。记 $v=\sum_{n=0}^{\infty}u^n$，则

$$\|v\|\le\sum_{n=0}^{\infty}\|u^n\|\le\frac{1}{1-\|u\|}$$

（部分和满足 $\|v_N\|\le\sum_{n=0}^{N}\|u^n\|$，令 $N\to\infty$，用范数的连续性）。

第二步：v 是 $\mathrm{Id}-u$ 的逆。对部分和 $v_N=\sum_{n=0}^{N}u^n$，直接展开计算得

$$(\mathrm{Id}-u)v_N=\sum_{n=0}^{N}u^n-\sum_{n=0}^{N}u^{n+1}=\mathrm{Id}-u^{N+1},$$

同样 $v_N(\mathrm{Id}-u)=\mathrm{Id}-u^{N+1}$。由复合关于算子范数连续（$w\mapsto(\mathrm{Id}-u)\circ w$ 与 $w\mapsto w\circ(\mathrm{Id}-u)$ 都是 Lipschitz 的），两端令 $N\to\infty$：左边 $v_N\to v$，右边因 $\|u^{N+1}\|\le\|u\|^{N+1}\to 0$ 而趋于 Id，得

$$(\mathrm{Id}-u)v=\mathrm{Id}=v(\mathrm{Id}-u).$$

所以 $\mathrm{Id}-u$ 可逆，且 $(\mathrm{Id}-u)^{-1}=v$，故

$$\|(\mathrm{Id}-u)^{-1}\|=\|v\|\le\frac{1}{1-\|u\|}.\qquad\#$$

注：

(1) 误差计算：用部分和 $\sum_{n=0}^{N}u^n$ 逼近 $(\mathrm{Id}-u)^{-1}$，

$$\Big\|(\mathrm{Id}-u)^{-1}-\sum_{n=0}^{N}u^n\Big\|=\Big\|\sum_{n=N+1}^{\infty}u^n\Big\|\le\sum_{n=N+1}^{\infty}\|u\|^n=\frac{\|u\|^{N+1}}{1-\|u\|}.$$

(2) 可逆元构成 $\mathcal{B}(E)$ 的开集：若 w 可逆，则 $\exists r>0$ s.t. $B(w,r)\subset\{\text{可逆算子}\}=\mathrm{GL}(E)$。
证明：对 $\tilde{w}\in\mathcal{B}(E)$，

$$\tilde{w}=w-(w-\tilde{w})=w\big[\mathrm{Id}-w^{-1}(w-\tilde{w})\big].$$

当 $\|w^{-1}(w-\tilde{w})\|<1$ 时，由 Neumann 级数定理，$\mathrm{Id}-w^{-1}(w-\tilde{w})$ 可逆，而 w 可逆，两个可逆元的乘积可逆（逆为 $\big[\mathrm{Id}-w^{-1}(w-\tilde{w})\big]^{-1}w^{-1}\in\mathcal{B}(E)$），故 $\tilde{w}$ 可逆。取 $r=\dfrac{1}{\|w^{-1}\|}$，则当 $\|w-\tilde{w}\|<r$ 时，

$$\|w^{-1}(w-\tilde{w})\|\le\|w^{-1}\|\cdot\|w-\tilde{w}\|<1,$$

所以 $B(w,r)\subset\mathrm{GL}(E)$。□

(3) 对 $|\lambda|>\|u\|$，有 $\|u/\lambda\|<1$，由 Neumann 级数定理 $\mathrm{Id}-\dfrac{u}{\lambda}$ 可逆，所以 $\lambda\cdot\mathrm{Id}-u=\lambda\Big(\mathrm{Id}-\dfrac{u}{\lambda}\Big)$ 可逆，且

$$(\lambda\cdot\mathrm{Id}-u)^{-1}=\frac{1}{\lambda}\sum_{n=0}^{\infty}\Big(\frac{u}{\lambda}\Big)^n=\sum_{n=0}^{\infty}\lambda^{-n-1}u^n.$$

结论：u 的谱（使 $\lambda\cdot\mathrm{Id}-u$ 不可逆的 λ 的全体）含于闭圆盘 $\{|\lambda|\le\|u\|\}$。

3

第三章

Hilbert 空间

3.1 内积空间

回顾：$\mathbb{R}^n$ 上的标准内积

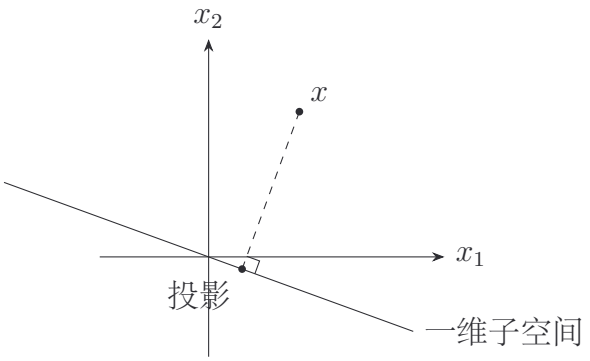
设 $x=(x_1,\dots,x_n)$，$y=(y_1,\dots,y_n)\in\mathbb{R}^n$。标准内积

$$\langle x,y\rangle=\sum_{i=1}^{n}x_iy_i$$

确定了 $\mathbb{R}^n$ 中向量的：

- **长度**：$\|x\|:=\sqrt{\langle x,x\rangle}=\sqrt{\sum_{i=1}^{n}x_i^2}$；
- **角度**：$\cos\theta=\dfrac{\langle x,y\rangle}{\|x\|\cdot\|y\|}\in[-1,1]$；
- **垂直**：$\langle x,y\rangle=0\iff x\perp y$。

有了垂直，就有了**投影**：把 x 沿垂直于子空间的方向落到一维子空间上，得到的就是 x 在该子空间上的投影。



Q：如果从有限维 $(\mathbb{R}^n)$ 推广到无穷维，会怎样？这一章就来回答这个问题：把内积的几条基本性质抽象成公理，在一般的向量空间上定义内积，从而在无穷维空间里也能谈长度、夹角、垂直与投影。

内积空间的定义

定义（内积）：设 H 是 $\mathbb{K}$ 上的向量空间（$\mathbb{K}=\mathbb{R}$ 或 $\mathbb{C}$），映射 $\varphi: H\times H\to\mathbb{K}$ 称为 H 上的**内积**，若

(1) **对第一个分量线性**：$\varphi(\lambda x_1+x_2,y)=\lambda\varphi(x_1,y)+\varphi(x_2,y)$，$\forall x_1,x_2,y\in H$，$\lambda\in\mathbb{K}$；

(2) **共轭对称**：$\varphi(x,y)=\overline{\varphi(y,x)}$，$\forall x,y\in H$；

(3) **正定**：$\varphi(x,x)\geq 0$，且 $\varphi(x,x)=0\iff x=0$。

此时称 (H,φ) 为**内积空间**。通常把内积 $\varphi(x,y)$ 记为 $\langle x,y\rangle$。

注：

① 由 (2)，$\varphi(x,x)=\overline{\varphi(x,x)}\Rightarrow\varphi(x,x)\in\mathbb{R}$，故 (3) 中“$\varphi(x,x)\geq 0$”有意义。

② 由 (1)(2) $\Rightarrow$ 内积对第二个分量**共轭线性**：

$$\langle x,\lambda y_1+y_2\rangle=\overline{\langle\lambda y_1+y_2,x\rangle}=\overline{\lambda\langle y_1,x\rangle+\langle y_2,x\rangle}=\bar\lambda\langle x,y_1\rangle+\langle x,y_2\rangle.$$

这样的映射 φ 是**共轭双线性型**（sesquilinear form）。$\mathbb{K}=\mathbb{R}$ 时共轭不起作用，内积就是双线性的。

③ $\langle x,0\rangle=\langle 0,y\rangle=0$。事实上，$\langle 0,y\rangle=\langle 0\cdot 0,y\rangle=0\cdot\langle 0,y\rangle=0$（左边的第一个 0 是零向量，$0\cdot 0$ 中第一个因子是数 0、第二个是零向量），而 $\langle x,0\rangle=\overline{\langle 0,x\rangle}=0$。

④ $\mathbb{K}=\mathbb{R}$ 时称为**实内积空间**（欧氏空间），$\mathbb{K}=\mathbb{C}$ 时称为**复内积空间**（酉空间）。

例：

(1) $\mathbb{K}^n$ 上的标准内积：

$$\langle x,y\rangle=\sum_{i=1}^{n}x_i\overline{y_i},\qquad x,y\in\mathbb{K}^n,$$

$\langle x,x\rangle=\sum_{i=1}^{n}x_i\overline{x_i}=\sum_{i=1}^{n}|x_i|^2=\|x\|^2$。对第一个分量线性、共轭对称是显然的（$\overline{y_i}$ 的共轭为 y_i，且数的乘法交换）；正定性：$\sum|x_i|^2\geq 0$，且为 0 当且仅当每个 $x_i=0$。

(2) $(C[0,1],\mathbb{K})$ 的内积：

$$\langle f,g\rangle=\int_0^1 f(x)\overline{g(x)}\,dx,$$

$\langle f,f\rangle=\int_0^1 f\bar f\,dx=\int_0^1|f|^2\,dx\geq 0$。积分的线性给出 (1)，$\overline{\int f\bar g}=\int\bar f g$ 给出 (2)；正定性中，$|f|^2$ 是连续的非负函数，积分为 0 当且仅当 $f\equiv 0$。这个内积可推广至 $L^2(0,1)$。

二次型

定义（二次型）：$q(x)=\langle x,x\rangle$ 称为内积对应的**二次型**。

由定义知：$q(x)\geq 0$ 且 $q(x)=0\iff x=0$。

命题（二次型的基本性质）：$\forall x,y\in H$，$\forall\lambda\in\mathbb{K}$，

(1) $q(\lambda x)=|\lambda|^2q(x)$；

(2) $q(x+y)=q(x)+q(y)+2\,\mathrm{Re}\langle x,y\rangle$；

(3) $q(x+\lambda y)=q(x)+|\lambda|^2q(y)+2\,\mathrm{Re}\left(\bar\lambda\langle x,y\rangle\right)$。

证明：(1) $\langle\lambda x,\lambda x\rangle=\lambda\bar\lambda\langle x,x\rangle=|\lambda|^2\langle x,x\rangle$。

(3) 由第一个分量线性、第二个分量共轭线性，

$$\langle x+\lambda y,x+\lambda y\rangle=\langle x,x\rangle+\bar\lambda\langle x,y\rangle+\lambda\langle y,x\rangle+|\lambda|^2\langle y,y\rangle.$$

又 $\lambda\langle y,x\rangle=\lambda\overline{\langle x,y\rangle}=\overline{\bar\lambda\langle x,y\rangle}$，所以中间两项之和是 $2\,\mathrm{Re}\left(\bar\lambda\langle x,y\rangle\right)$。

(2) 在 (3) 中取 $\lambda=1$ 即得。□

正交

记号：$\forall x,y\in H$，若 $\langle x,y\rangle=0$，称 x 与 y **正交**（垂直），记 $x\perp y$。

由共轭对称，$x\perp y\iff y\perp x$，且 $x\perp y$ 时（此时 $\mathrm{Re}\langle x,y\rangle=0$，由命题 (2)）

$$q(x+y)=q(x)+q(y).$$

（后文记 $\|x\|=\sqrt{q(x)}$，则上式即 $\|x+y\|^2=\|x\|^2+\|y\|^2$，这就是**勾股定理**。）

