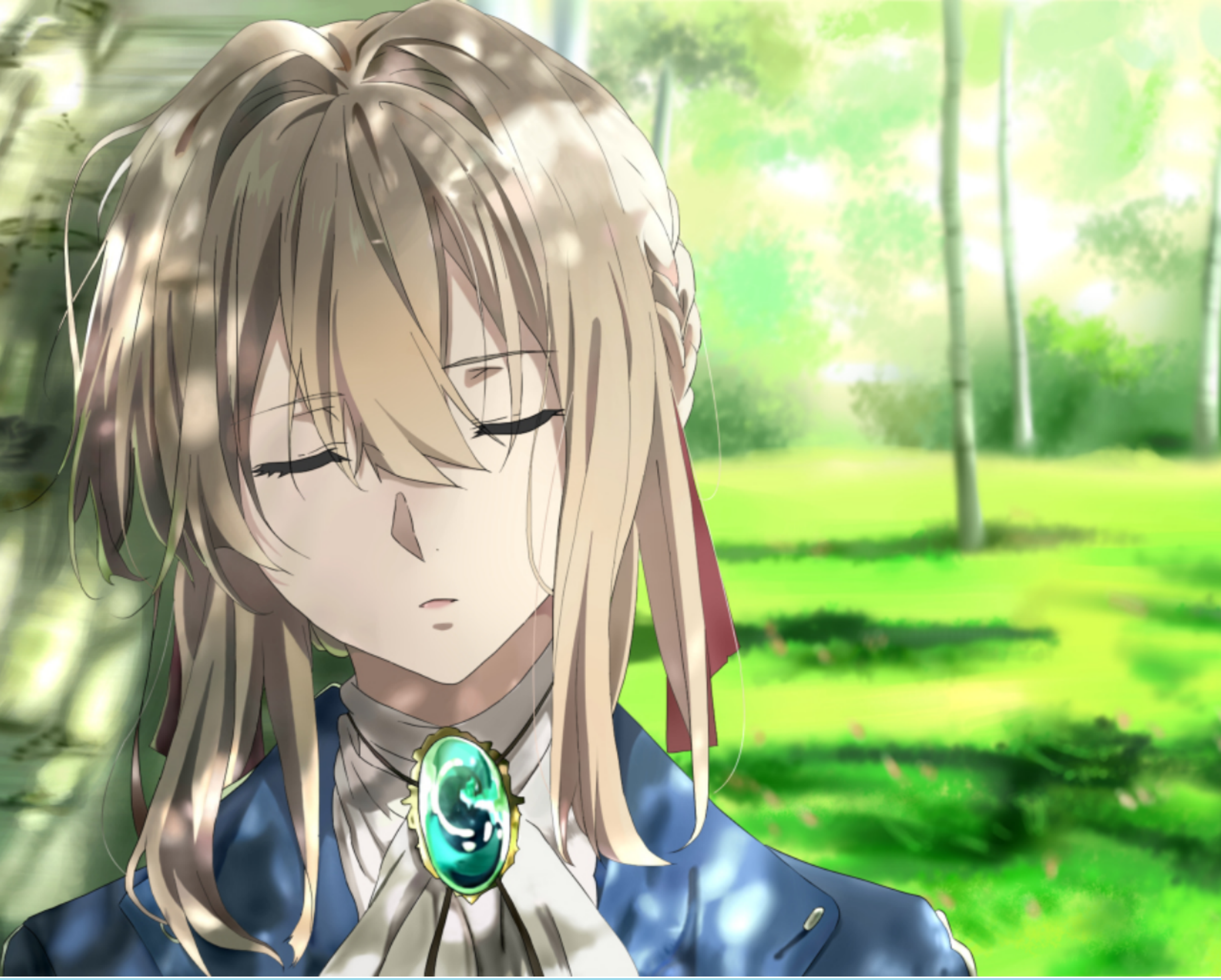

更进一步的泛函分析学习笔记

算子谱论, 广义函数论和非线性泛函分析 (等)

作者：UN

组织：北京师范大学数学科学学院

时间：2024.4.8 起著

模板来源：https://github.com/ElegantLaTeX/

封面来源：https://www.pixiv.net/artworks/67787485

道虽远，不行不至；事虽难，不为不成。

目录

第一部分

从拓扑空间出发的泛函分析

第一章 拓扑空间简介

本章选自 [**XMY**].

1.1 基本概念

1.1.1 开集, 闭集, 邻域系

定义 1.1 (拓扑, 开集, 拓扑空间)

设 E 是一个集合, 称 E 的子集族 τ(即 $\tau \subset \mathcal{P}(E)$) 是一个拓扑, 如果 τ 满足下述条件:

(i) $E \in \tau, \varnothing \in \tau$.

(ii) τ 中任意多个元素的并仍然是 τ 中的元素 (即满足任意并性质).

(iii) τ 中有限个元素的交仍然是 τ 中的元素 (即满足有限交性质).

称 τ 中的元素为开集, 称 (E, τ)(或简记为 E) 为拓扑空间. ♣

例 1.1

(1) $\tau = \{\varnothing, E\}$ 是 E 上的拓扑, 称为**平凡拓扑**.

(2) $\tau = \mathcal{P}(E)$ 是 E 上的拓扑, 称为**离散拓扑**.

例 1.2(度量空间) 设 E 是一个非空集合, d 是定义在 $E \times E$ 上的实值函数. 若任取 $x, y, z \in E$, 有

(i) **非负性**: $d(x, y) \geq 0$;

(ii) **正定性**: $d(x, y) = 0 \Leftrightarrow x = y$;

(iii) **对称性**: $d(x, y) = d(y, x)$;

(iv) **三角不等式**: $d(x, y) \leq d(x, z) + d(y, z)$,

就称 d 是集合 E 上的一个**度量** (或称**距离**), 称 (E, d) 为**度量空间**.

设 (E, d) 是度量空间, E 上的子集族 τ 定义为: 设 $U \subset E$, 且

$$U \in \tau \Leftrightarrow \forall x \in U \exists r > 0 (B(x, r) \subset U)$$

其中 $B(x, r) = \{y \in E : d(y, x) < r\}$. 则 τ 是 E 上的一个拓扑, 称 τ 为**由距离 d 诱导的拓扑**.

定义 1.2 (可度量化空间)

若 E 上的拓扑 τ 可由某个距离 d 诱导, 则称 (E, τ) 是一个可度量化空间. ♣

注 同一个拓扑可以被不同的距离 (乃至无限个距离) 诱导. 如 d 是 E 上的距离, 则 $d, \min\{1, d\}, rd (r > 0)$ 在 E 上诱导的拓扑是一样的.

定义 1.3 (闭集)

设 (E, τ) 是一个拓扑空间, $A \subset E$. 若 A 的补集 A^c 是关于拓扑 τ 的开集, 则称 A 是关于拓扑 τ 的闭集. ♣

由开集和闭集的互补关系立得闭集的下述性质:

定理 1.1 (闭集的基本性质)

设 (E, τ) 是一个拓扑空间, 则

(i) $E, \varnothing$ 都是闭集;

(ii) 任意多个闭集的交仍然是闭集 (即闭集满足任意交性质).

(iii) 有限多个闭集的并仍然是闭集 (即闭集满足有限并性质). ♡

注 可以用定理1.1中的性质 (i)-(iii) 来定义 E 上的拓扑 $\mathcal{F}$, 此时称 $\mathcal{F}$ 中的元素为闭集, 称它们的补集为开集.

特别注意: 无限多个开集的交未必是开集, 无限多个闭集的并未必是闭集.

定义 1.4 (邻域, 邻域系, 邻域基)

设 (E,τ) 是拓扑空间, $x \in E$.

(i) 若 $V \subset E$ 满足: 存在开集 $U \in \tau$, 使得 $x \in U, U \subset V$, 则称集合 V 为 x 的邻域. 用 $\mathcal{N}(x)$ 表示 x 点处所有邻域构成的集族, 称之为 x 的邻域系.

(ii) 若 $\mathcal{B}(x)$ 是 $\mathcal{N}(x)$ 的子集族, 且满足对任意 $V \in \mathcal{N}(x)$, 总存在 $U \in \mathcal{B}(x)$ 使得 $U \subset V$, 则称 $\mathcal{B}(x)$ 为 x 的邻域基 (或称基础邻域系).

♣

例 1.3

(i) $\mathcal{N}(x)$ 是 x 处的邻域基.

(ii) 含有 x 的所有开集族是 x 的一个邻域基, 但由 x 的所有闭邻域构成的集族未必是 x 的一个邻域基.

(iii) 设 (E,d) 是一个度量空间, 其拓扑由 d 诱导, $x \in E$, $\mathcal{B}(x)$ 表示所有中心在 x, 半径为 $\frac{1}{n}(n \in \mathbb{N}^{+})$ 的开球构成的集族, 则 $\mathcal{B}(x)$ 是 x 的邻域基.

邻域基有下述运算性质:

定理 1.2

设 E 是一个拓扑空间, $x \in E$, 则:

(i) $V \in \mathcal{N}(x) \Rightarrow x \in V$.

(ii) $V \in \mathcal{N}(x), U \supset V \Rightarrow U \in \mathcal{N}(x)$.

(iii) $\mathcal{N}(x)$ 中有限个元素的交仍然属于 $\mathcal{N}(x)$.

(iv) 对任意 $V \in \mathcal{N}(x)$, 存在 $U \in \mathcal{N}(x)$, 满足 $U \subset V$ 且任取 $y \in U$ 都有 $U \in \mathcal{N}(y)$.

♡

上面的 (iv) 实际上刻画了开集的性质, 即:

定理 1.3

设 (E,τ) 是一个拓扑空间, 且对任意 $x \in E$, 记 $\mathcal{B}(x)$ 为其邻域基, $U \subset E$, 则

$$U \text{ 是开集} \Leftrightarrow \forall x \in U \exists V \in \mathcal{B}(x)(V \subset U).$$

因此, U 是开集当且仅当它是它其中任意一点的邻域.

♡

证明 当 U 是开集, 任取 $x \in U$, 因为本身就存在开集 $U \subset U$, 故 U 是 x 的邻域, 亦即 $U \in \mathcal{B}(x)$.

当 $\forall x \in U \exists V \in \mathcal{B}(x)(V \subset U)$, 根据 $\mathcal{B}(x)$ 的定义知存在开集 O_x 使得

$$x \in O_x \subset V \subset U.$$

因此 $U = \bigcup_{x \in U} O_x$, 又因为任意开集的并还是开集, 故 U 是开集.

事实上, 集合 E 上的拓扑可由其上的邻域基唯一确定, 而邻域基可以事先由公理约定.

定理 1.4

设 E 是非空集合. 若对任意 $x \in E$, $\mathcal{B}(x)$ 表示关于 x 的由 E 的子集构成的非空集族, 且其满足:

(i) $\forall V \in \mathcal{B}(x)(x \in V)$.

(ii) $\forall (U,V) \in \mathcal{B}(x) \times \mathcal{B}(x) \exists W \in \mathcal{B}(x)(W \subset U \cap V)$.

(iii) 对任意 $V \in \mathcal{B}(x)$, 存在 $U \subset V$ 满足 $x \in U$, 且对任意 $y \in U$ 均存在 $W \in \mathcal{B}(y)$ 使得 $W \subset U$.

则存在 E 上唯一的拓扑 τ 使得对每一个 $x \in E$ 而言, $\mathcal{B}(x)$ 都是 x 关于拓扑 τ 的邻域基.

♡

1.1.2 粘着集, 闭包, 内部

定义 1.5 (凝聚点, 粘着点, 粘着集)

设 (E,τ) 是拓扑空间, $A \subset E$.

(i) 设 $x \in E$, 若任取 $V \in \mathcal{N}(x)$ 均有

$$(V \cap A) \backslash \{x\} \neq \varnothing,$$

则称 x 为 A 的凝聚点.

(ii) 若 x 是 A 中的元素或 A 的凝聚点, 则称 x 是 A 的粘着点. 也就是说当 x 是 A 的粘着点时有

$$V \cap A \neq \varnothing, \quad \forall V \in \mathcal{N}(x).$$

(iii) 称 A 的所有粘着点构成的集合为 A 的粘着集, 记作 $\overline{A}$.

♣

注 $x \notin \overline{A} \Leftrightarrow \exists V \in \mathcal{N}(x)(V \cap A = \varnothing)$.

定理 1.5

设 E 是一个拓扑空间, $A \subset E$.

(i) $\overline{A}$ 是 E 中包含 A 的最小闭集, 等价地, $\overline{A}$ 是 E 中包含 A 的所有闭集之交.

(ii) A 是闭集当且仅当 $\overline{A} = A$, 等价地, A 是闭集当且仅当 A 含有它自身所有的粘着点.

♡

注 E 中包含 A 的最小闭集称为 A 的**闭包**, 因此 A 的闭包和粘着集相同.

定义 1.6 (稠密)

设 $A \subset E$, 若 $\overline{A} = E$, 则称 A 在 E 中稠密.

♣

定义 1.7 (内部)

设 E 是拓扑空间, $x \in E, A \subset E$. 若 A 是 x 的一个邻域, 则称 x 为 A 的内点, 并称由 A 的所有内点构成的集合为 A 的内部, 记作 A°.

♣

定理 1.6

设 E 是拓扑空间, $A \subset E$.

(i) A° 是包含于 A 的所有开集的并, 等价地, A° 是包含于 A 的最大开集.

(ii) A 是开集当且仅当 $A^\circ = A$.

♡

定义 1.8 (边界)

设 E 是拓扑空间, $A \subset E$. 令 $\partial A = \overline{A} \backslash A^\circ$, 称 ∂A 为 A 的边界.

♣

定理 1.7

设 E 是拓扑空间, A, B 都是 E 的子集, 则

(i) $\overline{\overline{A}} = \overline{A}, (A^\circ)^\circ = A^\circ$.

(ii) $E \backslash A^\circ = \overline{E \backslash A}$.

(iii) $\overline{A \cup B} = \overline{A} \cup \overline{B}, (A \cap B)^\circ = A^\circ \cap B^\circ$.

♡

1.1.3 拓扑空间的比较和拓扑子空间

在同一个集合 E 上可以定义多个不同的拓扑, 这些拓扑之间可以建立偏序关系:

定义 1.9 (拓扑的强弱)

设 τ, τ' 是集合 E 上的两个拓扑. 若 $\tau' \subset \tau$, 则称 τ 是 τ' 的强拓扑 (也就是说, 每个 τ' 中的开集也是 τ 中的开集).

对集合 E 而言, 其上的离散拓扑 $\tau = \mathcal{P}(E)$ 是最强的拓扑, 平凡拓扑 $\tau = \{\varnothing, E\}$ 是最弱的拓扑.

下面我们讨论拓扑子空间的概念. 设 (E, τ) 是一个拓扑空间, $F \subset E$, 定义

$$\tau_F = \{U \cap F : U \in \tau\}.$$

容易验证 τ_F 是子集 F 上的拓扑, 称该拓扑为 F 上**由 τ 诱导的拓扑**. 相应地, 我们称 (F, τ_F) 是 (E, τ) 的**拓扑子空间**.

定理 1.8

设 F 是 E 的拓扑子空间, $A \subset F$, 那么 A 是 F 中的闭集当且仅当存在 E 中的闭集 B 使得 $A = B \cap F$.

1.1.4 Hausdorff 空间

定义 1.10 (Hausdorff 空间)

设 (E, τ) 是拓扑空间, 若任取 $x, y \in E, x \neq y$, 均存在 $U \in \mathcal{N}(x), V \in \mathcal{N}(y)$ 使得 $U \cap V = \varnothing$, 则称拓扑空间 (E, τ) 是 Hausdorff 空间 (或称为分离空间).

显见离散拓扑空间和度量空间是 Hausdorff 空间, 而至少有两个元素的平凡拓扑空间不是 Hausdorff 空间.

定理 1.9

设 E 是拓扑空间, 则 E 是 Hausdorff 空间当且仅当对任意一点 $x \in E$, 其所有闭邻域的交为单点集 $\{x\}$.

证明 任取 $x \in E$, 设 F 是 x 全体闭邻域的交集, 则 F 首先是闭集. 若 E 是 Hausdorff 空间, 任取 E 中不同于 x 的元素 y, 知存在开集 $V \in \mathcal{N}(y)$ 与 $U \in \mathcal{N}(x)$ 满足 $V \cap U = \varnothing$, 因此 $U \subset V^c$, 故 V^c 是 x 的闭邻域. 由此知 $F \subset V^c$, 故 $y \notin F$, 这说明 $F = \{x\}$.

反之, 若 $F = \{x\}$ 且 y 是 E 中不同于 x 的元素, 则总存在 x 的闭邻域 U 使得 $y \notin U$, 由此知 U^c 为开集且 $y \in U^c$, 故 E 是 Hausdorff 空间. □

推论 1.1

若拓扑空间 E 是 Hausdorff 空间, 则对任意 $x \in E$ 有

$$\bigcap_{V \in \mathcal{N}(x)} V = \{x\}.$$

上述推论的逆命题一般不正确. 在分析学中一般不使用非 Hausdorff 空间, 故以后我们默认提到的拓扑空间都是 Hausdorff 空间. 显见任何 Hausdorff 空间的子空间依旧是 Hausdorff 空间.

1.2 收敛序列和连续映射

定义 1.11 (极限)

设 (E, τ) 是拓扑空间, $\{x_n\}$ 是 E 中的序列, $x \in E$. 若任取 $V \in \mathcal{N}(x)$, 总存在 $n_0 \in \mathbb{N}$, 使得当 $n \geq n_0$ 时有 $x_n \in V$, 则称 $\{x_n\}$ 收敛于 x, 并称 x 为序列 $\{x_n\}$ 的极限, 记作 $\tau - \lim_{n\to\infty} x_n = x$, 在不混淆的情形下记作 $\lim_{n\to\infty} x_n = x$.

注

1. 上述定义中的 $\mathcal{N}(x)$ 可以换成任意邻域基 $\mathcal{B}(x)$.
2. 若 E 是 Hausdorff 空间, 则 E 中任一序列至多只有一个极限.

定义 1.12 (粘着值)

设 $\{x_n\}$ 是拓扑空间 E 中的序列, $x \in E$. 若任取 $V \in \mathcal{N}(x)$, 对任意 $n_0 \in \mathbb{N}$, 总存在 $n \geq n_0$ 使得 $x_n \in V$, 则称 x 是序列 $\{x_n\}$ 的粘着值.

从定义中可见粘着值之于极限就相当于数学分析中的子列极限之于极限. 关于粘着值有下述定理成立:

定理 1.10

设 $\{x_n\}$ 是拓扑空间 E 上的序列, $x \in E$. 令

$$A_n = \{x_m : m \geq n\},$$

则 x 是 $\{x_n\}$ 的粘着值当且仅当 $x \in \bigcap_n \overline{A_n}$.

注

(i) 设 $A \subset E$, 上述定理意味着 $\overline{A}$ 含有 A 中全体收敛序列的极限值 (也含有 A 中所有序列的粘着值). 当 E 是度量空间时, $\overline{A}$ 正是由 A 中所有收敛序列的极限值构成. 然而在一般拓扑空间下, 该结论未必成立, 此时 $\overline{A}$ 中可能含有 A 中收敛序列极限值以外的其他元素.

(ii) 度量空间和一般拓扑空间的根本区别在于: 度量拓扑可以由收敛序列刻画, 但一般拓扑不能这样.

下面讨论连续映射.

定义 1.13 (连续)

设 E, F 是两个拓扑空间, $f : E \to F$ 是一个映射, 且 $x \in E$. 若任取 $V \in \mathcal{N}(f(x))$, 总存在 $U \in \mathcal{N}(x)$, 使得 $f(U) \subset V$, 则称 f 在 x 点连续. 也就是说, 只要 f 在 x 点连续, 则对任意 $V \in \mathcal{N}(f(x))$, 都有 $f^{-1}(V) \in \mathcal{N}(x)$. 若映射 f 在 E 上每一点都连续, 则称 f 在 E 上连续.

注 显见若 f 在 x 点连续且 $\{x_n\}$ 收敛到 x, 那么 $\{f(x_n)\}$ 收敛到 $f(x)$. 在度量空间下, 这一命题的逆命题也成立, 亦即若 E, F 是度量空间, 那么由 $x_n \to x \Rightarrow \lim_{n\to\infty} f(x_n) = f(x)$ 可推知 f 在 x 点连续. 但对一般拓扑空间而言, 该命题不成立.

定理 1.11

设 E, F 是两个拓扑空间, $f : E \to F$ 是一个映射, 那么下面的命题等价:

(i) f 在 E 上连续.
(ii) 任取 F 中的开集 O, $f^{-1}(O)$ 都是 E 中的开集.
(iii) 任取 F 中的闭集 A, $f^{-1}(A)$ 都是 E 中的闭集.

推论 1.2

两个连续映射的复合映射依旧连续.

注 设 τ, τ' 是 E 上的两个拓扑, 则 τ 是 τ' 的强拓扑当且仅当 E 上的恒等映射 $\mathrm{Id}_E : (E, \tau) \to (E, \tau')$ 是连续的.

特别注意连续映射未必把开集映成开集.

定义 1.14 (开映射, 同胚映射)

设 E, F 是拓扑空间, $f : E \to F$ 是映射.

(i) 若任取开集 $O \subset E$, $f(O)$ 都是 F 中的开集, 则称 $f : E \to F$ 是开映射.
(ii) 若 f 是双射, f, f^{-1} 均连续 (也就是说 f 同时是连续双射和开映射), 则称 $f : E \to F$ 为同胚映射. 若 E, F 之间存在同胚映射, 则称 E, F 同胚.

1.3 紧性

定义 1.15 (开覆盖, 紧集)

设 E 是拓扑空间.

(i) 若 E 上的开集族 $\{O_i\}_{i\in\mathcal{I}}$ 满足 $E=\bigcup_{i\in\mathcal{I}}O_i$, 则称 $\{O_i\}_{i\in\mathcal{I}}$ 是 E 的开覆盖.

(ii) 若 E 的任意一个开覆盖 $\{O_i\}_{i\in\mathcal{I}}$ 中都可取出有限子覆盖, 则称 E 是紧的 (也就是说, 存在 $\mathcal{I}$ 的有限子集 $\mathcal{J}$ 使得 $\bigcup_{i\in\mathcal{J}}O_i=E$).

♣

由开集和闭集的互补性立得下述定理.

定理 1.12

拓扑空间 E 是紧的当且仅当任取 E 的一族闭集 $\{F_i\}_{i\in\mathcal{I}}$ 均有

$$\left(\forall 有限集\mathcal{J}\subset\mathcal{I}\Rightarrow\bigcap_{i\in\mathcal{J}}F_j\neq\varnothing\right)\Rightarrow\bigcap_{i\in\mathcal{I}}F_i\neq\varnothing.$$

♡

注 我们称上述定理中提到的闭集族 $\{F_i\}_{i\in\mathcal{I}}$ 具有的性质为**有限交性质**, 即

$$\forall 有限集\mathcal{J}\subset\mathcal{I},\bigcap_{i\in\mathcal{J}}F_i\neq\varnothing.$$

因此上述定理可表述为: 拓扑空间 E 是紧的当且仅当任意具有有限交性质的闭集族中所有元素的交集非空.

大多数情况下, 我们使用的命题是下述特殊情形:

推论 1.3

设 E 是紧拓扑空间, $\{F_n\}_{n\geq 1}$ 是 E 中单调递减的非空闭集族, 则

$$\bigcap_{n\geq 1}F_n\neq\varnothing.$$

♡

下面的定理介绍拓扑子空间紧性的判定方法:

定理 1.13

设 E 是拓扑空间, $F\subset E$ 是子空间, 则 F 是紧的当且仅当在 E 中的开集族 $\{O_i\}_{i\in\mathcal{I}}$ 满足 $F\subset\bigcup_{i\in\mathcal{I}}O_i$ 时, 存在有限集 $\mathcal{J}\subset\mathcal{I}$ 使得 $F\subset\bigcup_{i\in\mathcal{J}}O_i$(也就是说, F 的任一在 E 中的开覆盖 $\{O_i\}_{i\in\mathcal{I}}$ 都存在有限子覆盖).

♡

证明 设 $F\subset E$, E 中的开集族 $\{O_i\}_{i\in\mathcal{I}}$ 满足 $F\subset\bigcup_{i\in\mathcal{I}}O_i$, 令

$$U_i=O_i\cap F,\quad i\in\mathcal{I}.$$

则 $\{U_i\}_{i\in\mathcal{I}}$ 是空间 F 上的开覆盖. 反之, F 的任一开覆盖 $\{U_i\}_{i\in\mathcal{I}}$ 均可表示为上述形式.

现若 F 是紧子空间, 则存在有限集 $\mathcal{J}\subset\mathcal{I}$ 使得 $F=\bigcup_{i\in\mathcal{J}}U_i$, 因此 $F\subset\bigcup_{i\in\mathcal{J}}O_i$. 反之, 若存在有限集 $\mathcal{J}\subset\mathcal{I}$ 使得 $F\subset\bigcup_{i\in\mathcal{J}}O_i$, 则

$$F=F\cap\left(\bigcup_{i\in\mathcal{J}}O_i\right)=\bigcup_{i\in\mathcal{J}}U_i,$$

因此 F 是紧子空间. □

定理 1.14

设 E 是拓扑空间, $F\subset E$ 是子空间.

(i) 若 E 是 Hausdorff 空间且 F 是紧的, 那么 F 是 E 的闭子集.

(ii) 若 E 是紧 Hausdorff 空间, 则 F 是紧的当且仅当 F 是 E 的闭子集.

♡

证明 (i) 设 $x \in F^c$. 因为 E 是 Hausdorff 空间, 故对任意 $y \in F$, 均存在 E 中的开集 U_y, V_y 使得 $x \in U_y, y \in V_y, U_y \cap V_y = \varnothing$, 因此

$$F \subset \bigcup_{y \in F} V_y.$$

因为 F 是紧的, 故存在有限个点 $y_1, \cdots, y_n$ 使得

$$F \subset \bigcup_{i=1}^{n} V_{y_i}.$$

相应地, 令 $U = \bigcap_{i=1}^{n} U_{y_i}$, 则 U 是 E 中含有 x 的开集, 且

$$U \cap \left(\bigcup_{i=1}^{n} V_{y_i} \right) = \varnothing.$$

故 $U \cap F = \varnothing$. 这说明对任意 $x \in F^c$, 总能找到完全包含在 F^c 中的开集 U 作为 x 的邻域, 从而 F^c 是开集, 故 F 是闭集.

(ii) 设 E 是紧空间, F 是 E 中的闭集. 设 $\{F_i\}_{i \in \mathcal{I}}$ 是 F 中一族具有有限交性质的闭集, 因为 F 本身是 E 中的闭集, 故 F_i 同样是 E 中的闭集. 根据定理1.12与 E 的紧性知 $\bigcap_{i \in \mathcal{I}} F_i \neq \varnothing$, 进而由定理1.12可得 F 的紧性. □

在任一 Hausdorff 空间上, 如果一个具有有限交性质的集族是紧集族, 那么该紧集族中所有元素的交集非空. 换言之, 我们有下述结论:

定理 1.15

设 E 是 Hausdorff 空间, $\{K_i\}_{i \in \mathcal{I}}$ 是 E 中的紧集族, 若 $\bigcap_{i \in \mathcal{I}} K_i = \varnothing$, 则存在有限个 K_i 使得它们的交集是空集.

证明 在集族 $\{K_i\}_{i \in \mathcal{I}}$ 中任取一个元素, 记为 K_0. 因为 $\bigcap_{i \in \mathcal{I}} K_i = \varnothing$, 故 $\{K_i^c\}_{i \in \mathcal{I}}$ 是 K_0 的开覆盖[1]. 由 K_0 的紧性知存在有限集 $\mathcal{J} \subset \mathcal{I}$ 使得 $K_0 \subset \bigcup_{i \in \mathcal{J}} K_i^c$, 因此

$$K_0 \cap \left(\bigcap_{i \in \mathcal{J}} K_i \right) = \varnothing.$$

此即欲证. □

推论 1.4

设 E 是紧 Hausdorff 空间, 则 E 中每一点均有一个紧邻域基.

证明 取 $x \in E$, V 是 x 的一个开邻域, 则 V^c 是闭集, 进而 V^c 是紧集. 根据定理1.9, 任取 x 的闭邻域 U, 知

$$\bigcap_{U \in \mathcal{N}(x)} (U \cap V^c) = \left(\bigcap_{U \in \mathcal{N}(x)} U \right) \cap V^c = \{x\} \cap V^c = \varnothing;$$

进而由定理1.15知存在有限个紧集 $U_1, \cdots, U_n \in \mathcal{N}(x)$ 使得 $\bigcap_{i=1}^{n} (U_i \cap V^c) = \varnothing$, 因此

$$U_1 \cap \cdots \cap U_n \subset V.$$

因为上左式中有限紧邻域的交依旧是 x 的紧邻域, 故 x 存在紧邻域基. □

对于一个非紧的空间, 上述结论也可能正确, 例如 $\mathbb{R}$. 这便启发我们给出下述定义:

定义 1.16 (局部紧空间)

设 E 是拓扑空间, 若 E 上的每一点均有一个紧邻域, 则称 E 为局部紧空间.

此时推论1.4可修改为下述定理:

定理 1.16

若 E 是局部紧 Hausdorff 空间, 则 E 中每一点均有一个紧邻域基.

---

[1] Hausdorff 空间的条件用在这里了: Hausdorff 空间中的紧集必为闭集.

紧性在连续映射下不变, 即:

定理 1.17

设 E 是紧空间, F 是拓扑空间, $f: E \to F$ 是连续映射, 则 $f(E)$ 是 F 的紧子空间.

证明 设 $\{U_i\}_{i\in\mathcal{I}}$ 是 $f(E)$ 的开覆盖, 即 $f(E) \subset \bigcup_{i\in\mathcal{I}} U_i$, 则 $E = \bigcup_{i\in\mathcal{I}} f^{-1}(U_i)$. 因为 f 连续, 故 $f^{-1}(U_i)$ 是 E 中的开集. 因为 E 本身是紧集, 故存在有限个元素 $i_1, \cdots, i_n$ 使得

$$E = \bigcup_{k=1}^{n} f^{-1}(U_{i_k}),$$

因此 $f(E) \subset \bigcup_{k=1}^{n} U_{i_k}$, 从而 $f(E)$ 是 F 的紧子空间. □

推论 1.5

设 E, F 是拓扑空间, $f: E \to F$ 是连续映射, 则只要 $K \subset E$ 是紧集, $f(K)$ 就也是紧集.

推论 1.6

设 E 是紧空间, F 是 Hausdorff 空间, $f: E \to F$ 是连续单射, 则 f 是 $E \to f(E)$ 的同胚映射.

证明 只需说明 $f^{-1}: f(E) \to E$ 连续即可. 设 A 是 E 中的闭集, 因为 E 是紧空间, 故 A 是紧集. 又因为 f 连续, 故 $f(A)$ 是紧集, 因此 $f(A) = (f^{-1})^{-1}(A)$ 是 $f(E)$ 中的闭集, 从而 f^{-1} 是连续映射. □

推论 1.7

设 E 是紧空间, $f: E \to \mathbb{R}$ 是连续函数, 则 f 在 E 上有界, 且能取到上确界和下确界.

证明 由前述推论知 $f(E)$ 是 $\mathbb{R}$ 上的紧集, 进而其为有界闭集, 自然存在上下确界. □

本节末尾我们给出著名的 Urysohn 引理 (证明就略过了):

定理 1.18 (Urysohn 引理)

设 E 是局部紧 Hausdorff 空间, A, B 是 E 中不相交的非空闭集, 且其中一个是紧集, 则存在连续函数 $f: E \to [0,1]$ 使得 $f|_A \equiv 0, f|_B \equiv 1$.

注 当 (E, d) 是度量空间时, 取

$$f(x) = \frac{d(x,A)}{d(x,A) + d(x,B)}, \quad d(x,A) = \inf\{d(x,y) : y \in A\}$$

即可说明结论成立. 我们称具有 Urysohn 引理提到的性质的空间为**正规空间**, 因此局部紧 Hausdorff 空间和可度量化空间都是正规空间.

1.4 乘积拓扑

定义 1.17 (直积, 基础开集, 乘积拓扑)

设 (E_i, τ_i) 是一族拓扑空间, $E = \prod_{i\in\mathcal{I}} E_i$ 称为集族 $\{E_i\}_{i\in\mathcal{I}}$ 的直积, 其中的元素记为 $x = (x_i)_{i\in\mathcal{I}} \in \prod_{i\in\mathcal{I}} E_i (x_i \in E_i, \forall i \in \mathcal{I})$, 称 x_i 为 x 的坐标.

(i) 若 $\mathcal{J}$ 是 $\mathcal{I}$ 的有限子集, $U_i (i \in \mathcal{J})$ 是拓扑空间 (E_i, τ_i) 上的开集, 则称

$$O = \prod_{i\in\mathcal{J}} U_i \times \prod_{i\in\mathcal{I}\setminus\mathcal{J}} E_i$$

为 E 的基础开集.

(ii) 称由基础开集的并构成的集合为 E 中的开集. 事实上, 所有这种开集构成了 E 上的拓扑, 称该拓扑为

$E = \prod_{i\in\mathcal{I}} E_i$ 上的乘积拓扑.

注

(i) 如果 $\mathcal{I}$ 是有限集, 则 O 是基础开集当且仅当 $O = \prod_{i\in\mathcal{I}} U_i$, 其中 U_i 是 E_i 中的开集. 特别注意如果 $\mathcal{I}$ 是无限集, 则前文提到的这种开集的直积通常不是乘积拓扑空间中的开集.

(ii) $\mathbb{R}^n$ 上的自然拓扑与 $\mathbb{R}\times\cdots\times\mathbb{R}$ 上的乘积拓扑一致.

(iii) 若对每个 $i\in\mathcal{I}$ 而言, F_i 都是 E_i 中的闭集, 则 $\prod_{i\in\mathcal{I}} F_i$ 是 E 中的闭集.

(iv) 设 $x = (x_i)_{i\in\mathcal{I}} \in E$, 且对每个 $i\in\mathcal{I}$ 而言, 记 $\mathcal{N}(x_i)$ 为 x_i 在 E_i 中的邻域系. 则 x 在 E 中的一个邻域基可由下述集合构成:

$$\prod_{i\in\mathcal{J}} V_i \times \prod_{i\in\mathcal{I}\setminus\mathcal{J}} E_i,$$

其中 $\mathcal{J}$ 是 $\mathcal{I}$ 的有限子集, 且对每个 $i\in\mathcal{J}$ 均有 $V_i\in\mathcal{N}(x_i)$.

(v) 拓扑空间的乘积满足结合律.

下面的定理解释了前面定义的乘积拓扑是自然的, 也就是说它是使得 $\prod_{i\in\mathcal{I}} E_i$ 上每个正规投影都连续的最弱拓扑.

定理 1.19 (乘积空间拓扑的最弱拓扑刻画)

设 $\{E_i\}_{i\in\mathcal{I}}$ 是一族拓扑空间, $E = \prod_{i\in\mathcal{I}} E_i$, 记 τ 为 E 上的乘积拓扑, 另对每个 $i\in\mathcal{I}$, 记 p_i 为 $E\to E_i$ 的正规投影, 即 $p_i(x) = x_i (x = (x_i)_{i\in\mathcal{I}})$. 那么 τ 是 E 上使得每个 $p_i (i\in\mathcal{I})$ 都连续的最弱拓扑, 且每个 $p_i (i\in\mathcal{I})$ 都是开映射.

证明 对每个 $i\in\mathcal{I}$, 设 U_i 是 E_i 中的开集, 则

$$p_i^{-1}(U_i) = U_i \times \prod_{j\neq i} E_j.$$

根据乘积拓扑的定义即知 $p_i^{-1}(U_i)$ 是 E 中的开集, 因此 p_i 是连续映射.

下面说明 τ 的最弱性. 设 σ 是 E 上另一个使每个 $p_i (i\in\mathcal{I})$ 都连续的拓扑, 则对每个 $E_i (i\in\mathcal{I})$ 上的开集 U_i 而言, $U_i\times\prod_{j\neq i} E_j$ 作为某开集在连续映射下的原像集是关于 σ 的开集. 对这样的开集 $U_i\times\prod_{j\neq i} E_j$ 求有限交可知拓扑 τ 中的基础开集是拓扑 σ 中的开集, 因此 $\tau\leq\sigma$, 亦即 τ 是使得 E 上所有正规投影均连续的最弱拓扑.

最后说明在拓扑 τ 下每个 p_i 都是开映射. 设 O 是 E 中的开集, 并设 $x\in O$. 因为 O 本身是 E 中基础开集之并, 故 x 必属于某个基础开集 $U = \prod_{i\in\mathcal{I}} U_i$(其中 U_i 是 E_i 中的开集, 且除去有限个情形后总有 $U_i = E_i$), 且 $U\subset O$. 这说明 $U_i\subset p_i(O)$, 即

$$x_i\in U_i\subset p_i(O).$$

因此 $p_i(O)$ 是其中每个点的邻域, 故 $p_i(O)$ 是 E_i 中的开集, 从而 p_i 是开映射. □

下面的推论刻画了乘积拓扑空间中序列的收敛性, 即序列在乘积拓扑空间中的收敛等价于依坐标收敛, 这意味着乘积拓扑正是我们期望刻画的拓扑.

推论 1.8

设 $\{E_i\}_{i\in\mathcal{I}}$ 是一族拓扑空间, $E = \prod_{i\in\mathcal{I}} E_i$ 是乘积拓扑空间.

(i) 设 $\{x^n\}_{n\in\mathbb{N}}$ 是 E 中的序列, 其中 $x^n = (x_i^n)_{i\in\mathcal{I}}$, 另记 $x = (x_i)_{i\in\mathcal{I}}\in E$, 则 $\{x^n\}_{n\in\mathbb{N}}$ 在乘积拓扑空间 E 中收敛到 x 当且仅当对每个 $i\in\mathcal{I}$ 而言, $\{x_i^n\}_{n\in\mathbb{N}}$ 在空间 E_i 中收敛到 x_i.

(ii) 设 F 是拓扑空间, $f: F\to E$ 是映射, 则 f 是 $F\to(E,\tau)$ 的连续映射当且仅当对每个 $i\in\mathcal{I}$ 而言, $p_i\circ f$ 是 $F\to E_i$ 的连续映射.

证明 (i) 设 $\{x^n\}_{n\in\mathbb{N}}$ 收敛到 x, 因为 p_i 连续, 故对每个 $i\in\mathcal{I}$ 都有 $\{x_i^n\}_{n\in\mathbb{N}}$ 收敛到 x_i.

反之若对每个 $i \in \mathcal{I}$ 都有 $\{x_i^n\}_{n\in\mathbb{N}}$ 收敛到 x_i, 设 O 是 τ 中含有点 x 的任意开集, 则存在基础开集

$$U = \prod_{i\in\mathcal{J}} U_i \times \prod_{i\in\mathcal{I}\setminus\mathcal{J}} E_i$$

满足 $x \in U \subset O$, 其中 $\mathcal{J}$ 是有限集, U_i 是 E_i 中的开集. 现在因为 $\mathcal{J}$ 是有限集, 故存在 n_0 使得 $n \geq n_0$ 时, 对每个 $i \in \mathcal{J}$ 总有 $x_i^n \in U_i$. 也就是说, 当 $n \geq n_0$ 时有 $x^n \in O$, 因此 $\{x^n\}_{n\in\mathbb{N}}$ 收敛到 x.

(ii) 的证明与 (i) 是类似的, 这里就省略了. □

注 如果所有的 $E_i(i \in \mathcal{I})$ 都是同一个集合 F, 那么这种情况下

$$\prod_{i\in\mathcal{I}} E_i = \prod_{i\in\mathcal{I}} F = \{f : \mathcal{I} \to F, i \mapsto f(i) \in F\}.$$

也就是说, $\prod_{i\in\mathcal{I}} F$ 中的元素是指标集 $\mathcal{I}$ 到集合 F 的映射, 这种直积集通常记为 $F^{\mathcal{I}}$. 若取 $F = \mathbb{R}$, 则 $\prod_{i\in\mathcal{I}} = \mathbb{R}^{\mathcal{I}}$ 中的元素就是从指标集 $\mathcal{I}$ 到实数集 $\mathbb{R}$ 上的函数. 例如令 $\mathcal{I} = \mathbb{N}$, 则 $\mathbb{R}^{\mathbb{N}}$ 中的元素就是实无穷序列 $(x_1, x_2, \cdots)$.

当我们在 $\mathbb{R}^{\mathcal{I}}$ 上赋乘积拓扑后, 由推论1.8可知, $\mathbb{R}^{\mathcal{I}}$ 中的序列 $\{f_n\}$ 依乘积拓扑收敛到 $f \in \mathcal{R}^{\mathcal{I}}$ 等价于 f_n 作为 $\mathcal{I}$ 上的函数列逐点收敛到函数 f, 这说明乘积拓扑是刻画函数逐点收敛的拓扑.

定理 1.20

如果对每个 $i \in \mathcal{I}$, E_i 都是 Hausdorff 空间, 则乘积拓扑空间 $E = \prod_{i\in\mathcal{I}} E_i$ 也是 Hausdorff 空间.

证明 设 $x = (x_i)_{i\in\mathcal{I}}, y = (y_i)_{i\in\mathcal{I}}$ 是 $E = \prod_{i\in\mathcal{I}} E_i$ 中的两个不同元素, 则存在 $i_0 \in \mathcal{I}$ 使得 $x_{i_0} \neq y_{i_0}$. 因为 E_{i_0} 是 Hausdorff 空间, 故存在 E_{i_0} 中的开集 U_{i_0}, V_{i_0}, 使得 $x_{i_0} \in U_{i_0}, y_{i_0} \in V_{i_0}, U_{i_0} \cap V_{i_0} = \varnothing$. 因此

$$x \in U = U_{i_0} \times \prod_{i\neq i_0} E_i, \quad y \in V = V_{i_0} \times \prod_{i\neq i_0} E_i.$$

这里 U, V 是乘积拓扑空间 E 中的两个基础开集, 且满足 $U \cap V = \varnothing$, 因此 $E = \prod_{i\in\mathcal{I}}$ 同样是 Hausdorff 空间. □

下面讨论反映乘积拓扑空间稳定性的重要性质: 紧性. 我们首先给出两个成员情形下的结论:

定理 1.21

设 E_1, E_2 是两个紧拓扑空间, 则乘积拓扑空间 $E = E_1 \times E_2$ 依旧是紧的.

证明 设 $\{U_i\}_{i\in\mathcal{I}}$ 是 E 上的开覆盖, 则任取 $x = (x_1, x_2) \in E$, 必存在 $i_x \in \mathcal{I}$ 使得 $x \in U_{i_x}$. 根据乘积拓扑的定义知存在 V_x, W_x 分别是 x_1, x_2 在相应拓扑上的开邻域, 且 $V_x \times W_x \subset U_{i_x}$. 现在固定 $x_2 \in E_2$, 设 $F = E_1 \times \{x_2\}$, 则 $\{V_x\}_{x\in F}$ 是 E_1 上的一个开覆盖. 因为 E_1 是紧的, 故存在 F 的有限子集 $\mathcal{J}(x_2)$ 使得 $\{V_x\}_{x\in\mathcal{J}(x_2)}$ 覆盖 E_1. 下面令

$$A_{x_2} = \bigcap_{x\in\mathcal{J}(x_2)} W_x,$$

则 A_{x_2} 是含有 x_2 的开集. 当 x_2 在 E_2 中变动时, 可知 $\{A_{x_2}\}_{x_2\in E_2}$ 是 E_2 的开覆盖, 因此由 E_2 的紧性知存在有限集 $K \subset E_2$ 使得相应的有限集族 $\{A_{x_2}\}_{x_2\in K}$ 覆盖 E_2. 由此知

$$E = \bigcup_{x\in\mathcal{J}} V_x \times W_x, \quad \mathcal{J} = \bigcup_{x_2\in K} \mathcal{J}(x_2).$$

现由上述讨论, 任取 $a = (a_1, a_2) \in E$ 时, 必存在某个 $x_2 \in K$ 使得 $a_2 \in A_{x_2}$, 进而存在 $x \in \mathcal{J}(x_2)$ 使得 $a_1 \in V_x$, 这里确定的 $x \in \mathcal{J}$ 且 $a \in V_x \times W_x$. 因此 E 能被 $\{U_{i_x}\}_{x\in\mathcal{J}}$ 覆盖. 又因为 $\mathcal{J}$ 时有限集, 故 E 是紧的. □

一般情形下有下述结论成立 (证明就省略了):

定理 1.22 (Tychonoff)

设 $\{E_i\}_{i\in\mathcal{I}}$ 是一族紧拓扑空间, 则乘积拓扑空间 $\prod_{i\in\mathcal{I}} E_i$ 依旧是紧的.

最后, 可以说明可度量化在可数乘积运算下是稳定的 (证明就省略了):

定理 1.23

设 $\{E_n\}_{n\geq 1}$ 是一族可度量化的拓扑空间, 则乘积拓扑空间 $E=\prod_{n\in\mathbb{N}} E_n$ 也可度量化. ♡

第二章 四大定理的回顾

本章选自 [**HB**].

2.1 Hahn-Banach 定理

2.1.1 Hahn-Banach 定理的解析形式: 线性泛函的延拓

Hahn-Banach 定理解析形式的证明需要用到 Zorn 引理, 为此需要准备一些定义. 设 P 是具有 (偏) 序关系 $\leq$ 的集合, 称 $Q \subset P$ 是**全序集 (totally ordered)**, 如果对 Q 中的任意两个元素 a, b, 要么 $a \leq b$, 要么 $b \leq a$. 称 P 是**归纳集 (inductive)**, 如果其中的任意全序子集都有上界.

引理 2.1 (Zorn)

任意非空归纳序集都有最大元.

定理 2.1 (Hahn-Banach 定理的解析形式)

设 E 是 $\mathbb{R}$ 上的向量空间, $p : E \to \mathbb{R}$ 作为函数满足[a]

$$p(\lambda x) = \lambda p(x), \quad \forall x \in E, \forall \lambda > 0, \tag{2.1}$$

$$p(x + y) \leq p(x) + p(y), \quad \forall x, y \in E. \tag{2.2}$$

设 $G \subset E$ 是线性子空间, $g : G \to \mathbb{R}$ 作为线性泛函满足

$$g(x) \leq p(x), \quad \forall x \in G. \tag{2.3}$$

在上述条件下, 存在定义在 E 上的线性泛函 f, 使得对任意 $x \in G$ 均有 $g(x) = f(x)$, 且

$$f(x) \leq p(x), \quad \forall x \in E. \tag{2.4}$$

[a] 满足(2.1),(2.2)的函数 p 有时又称为 Minkowski 泛函.

证明 考虑集合

$$P = \left\{ h : D(h) \subset E \to \mathbb{R} \;\middle|\; \begin{array}{l} D(h)\text{是 } E \text{ 的线性子空间} \\ h\text{是线性的}, G \subset D(h), \\ h \text{ 是 } g \text{ 的延拓}, h(x) \leq p(x)(\forall x \in D(h)) \end{array} \right\}$$

在 P 上定义序关系:

$$(h_1 \leq h_2) \Leftrightarrow (D(h_1) \subset D(h_2) \wedge h_2 \text{ 是 } h_1 \text{ 的延拓}).$$

显见 P 非空, 因为 $g \in P$. 下面说明 P 是归纳集. 设 $Q \subset P$ 是全序子集, 记 $Q = \{h_i\}_{i \in \mathcal{I}}$, 另记

$$D(h) = \bigcup_{i \in \mathcal{I}} D(h_i), \quad h(x) = h_i(x), x \in D(h_i), i \in \mathcal{I}.$$

显见上式中的 h 良定义, $h \in P$, 且 h 是 Q 的一个上界. 现在对 P 应用 Zorn 引理可得其最大元 f.

下面说明 $D(f) = E$. 若 $D(f) \neq E$, 则取 $x_0 \notin D(f)$, 记 $D(h) = D(f) + \mathbb{R}x_0$, 且对任意 $x \in D(f)$ 记

$$h(x + tx_0) = f(x) + t\alpha, \quad t \in \mathbb{R}$$

其中常数 $\alpha \in \mathbb{R}$ 是为满足 $h \in P$ 而选取的. 现在需要说明

$$f(x) + t\alpha \leq p(x + t\alpha), \quad \forall x \in D(f), \forall t \in \mathbb{R}.$$

由齐次性(2.1)知只需验证

$$\begin{cases} f(x)+\alpha \le p(x+x_0), & \forall x \in D(f) \\ f(x)-\alpha \le p(x-x_0), & \forall x \in D(f) \end{cases}$$

即可. 也就是说, α 需要满足

$$\sup_{y\in D(f)} \{f(y)-p(y-x_0)\} \le \alpha \le \inf_{x\in D(f)} \{p(x+x_0)-f(x)\}.$$

因为由三角不等式(2.2)知

$$f(x)+f(y) \le p(x+y) \le p(x+x_0)+p(y-x_0)$$

故

$$f(y)-p(y-x_0) \le p(x+x_0)-f(x), \quad \forall x \in D(f), \forall y \in D(f),$$

因此前面要求的 α 肯定是存在的, 由此可知 $f \le h$. 但因为 $f \ne h$, 故这与 f 是 P 的最大元矛盾! □

解析形式的 Hahn-Banach 定理多用于对偶空间的研究. 回忆 E 的**对偶空间 (dual space)**E^* 是 E 上全体线性泛函构成的空间, 其上范数定义为

$$\|f\|_{E^*} = \sup_{\substack{\|x\|\le 1\\ x\in E}} |\langle f,x\rangle|.$$

本科泛函分析中已经说明了就算 E 不完备的赋范空间, E^* 也是 Banach 空间, 这是因为 $\mathbb{R}$ 是完备的.

解析形式的 Hahn-Banach 定理有下述重要推论:

推论 2.1

设 $G \subset E$ 是线性子空间, 如果 $g: G \to \mathbb{R}$ 是连续线性泛函, 则它在 E 上存在延拓 $f \in E^*$, 且

$$\|f\|_{E^*} = \|g\|_{G^*}.$$

♡

证明 令 $p(x) = \|g\|_{G^*}\|x\|_E$ 即可. □

推论 2.2

对任意 $x_0 \in E$, 总存在 $f_0 \in E^*$ 使得

$$\|f_0\|_{E^*} = \|x_0\|_E, \quad \langle f_0, x_0\rangle = \|x_0\|_E^2.$$

♡

证明 在推论2.1中令 $G = \mathbb{R}x_0, g(tx_0) = t\|x_0\|^2$ 即可.

注 推论2.2中给出的 f_0 一般来说不是唯一的. 但如果 E^* 是严格凸空间[1], 例如 Hilbert 空间或 L^p 空间 $(1 < p < \infty)$, 则 f_0 唯一. 在一般情况下, 对任意 $x_0 \in E$ 记

$$F(x_0) = \{f_0 \in E^* : \|f_0\|_{E^*} = \|x_0\|_E, \langle f_0, x_0\rangle = \|x_0\|^2\}.$$

称 (多值) 映射 $x_0 \mapsto F(x_0)$ 为 $E \to E^*$ 的**对偶映射 (duality map)**.

推论 2.3

对任意 $x \in E$ 均有

$$\|x\|_E = \sup_{\substack{f\in E^*\\ \|f\|\le 1}} |\langle f,x\rangle| = \max_{\substack{f\in E^*\\ \|f\|\le 1}} |\langle f,x\rangle|.$$

♡

证明 不妨设 $x \ne 0$, 显见

$$\sup_{\substack{f\in E^*\\ \|f\|\le 1}} |\langle f,x\rangle| \le \|x\|.$$

另一方面, 由推论2.2知存在 $f_0 \in E^*$ 使得 $\|f_0\|_{E^*} = \|x\|_E, \langle f_0, x\rangle = \|x\|_E^2$, 取 $f_1 = f_0/\|x\|$ 即可. □

[1]称赋范空间 E 严格凸, 如果对任意 $t \in (0,1)$ 与任意 $x, y \in E$, 由 $\|x\| = \|y\| = 1, x \ne y$ 可推知 $\|tx + (1-t)y\| < 1$.

2.1.2 Hahn-Banach 定理的几何形式: 凸集的分离

我们从超平面的一些预备知识出发. 下面默认 E 是赋范线性空间.

定义 2.1 (仿射超平面 (affine hyperplane))

E 上的仿射超平面 H 是 E 的一个子集, 其形如

$$H = \{x \in E : f(x) = \alpha\},$$

其中 f 是不恒为零的线性泛函[a], $\alpha \in \mathbb{R}$ 是给定的常数. 我们将上式简记为 $H = [f = \alpha]$, 称 $f = \alpha$ 是 H 的方程.

[a]注意这里我们没要求连续性.

命题 2.1

超平面 $H = [f = \alpha]$ 是闭集当且仅当 f 连续.

证明 从连续性的定义出发显见 f 连续 $\Rightarrow H$ 是闭集. 反之设 H 是闭集, 因为 f 不恒为零, 故 H^c 是非空开集. 取 $x_0 \in H^c$, 知 $f(x_0) \neq \alpha$, 不妨设 $f(x_0) < \alpha$.

取定 $r > 0$ 满足 $B(x_0, r) \subset H^c$, 下面说明

$$f(x) < \alpha, \quad \forall x \in B(x_0, r). \tag{2.5}$$

这是因为若存在 $x_1 \in B(x_0, r)$ 使得 $f(x_1) > \alpha$, 根据球 $B(x_0, r)$ 的凸性知线段

$$\{x_t = (1-t)x_0 + tx_1 : t \in [0,1]\}$$

被囊括在 $B(x_0, r)$ 内, 因此对任意 $t \in [0,1]$ 均有 $f(x_t) \neq \alpha$. 另一方面, 如果取 $t = \frac{f(x_1)-\alpha}{f(x_1)-f(x_0)}$, 可得 $f(x_t) = \alpha$, 这便导出了矛盾. 因此(2.5)式成立. 由(2.5)式知

$$f(x_0 + rz) < \alpha, \quad \forall z \in B(0,1).$$

因此 f 连续, 且 $\|f\|_{E^*} \leq \frac{1}{r}(\alpha - f(x_0))$. □

定义 2.2 (分离 (seperate), 严格分离 (strictly seperate), 凸集 (convex set))

设 A, B 是 E 的子集. 称超平面 $H = [f = \alpha]$ 分离 A, B, 如果

$$f(x) \leq \alpha (\forall x \in A), \quad f(x) \geq \alpha (\forall x \in B).$$

称 H 严格分离 A, B, 如果存在 $\varepsilon > 0$ 使得

$$f(x) \leq \alpha - \varepsilon (\forall x \in A), \quad f(x) \geq \alpha + \varepsilon (\forall x \in B).$$

从几何上看, A, B 能够被某超平面 H 分离意味着 A 在由 H 确定的其中一个半空间中, 而 B 在另一半空间中.

称 $A \subset E$ 是凸集, 如果

$$tx + (1-t)y \in A, \quad \forall x, y \in A, \forall t \in [0,1].$$

定理 2.2 (Hahn-Banach 定理的第一几何形式)

设 $A \subset E, B \subset E$ 是满足 $A \cap B = \varnothing$ 的非空凸集, 另设其一是开集, 则存在闭超平面分离 A, B.

定理2.2的证明需要下述两个引理作为铺垫:

引理 2.2

设 $C \subset E$ 是开凸集, $0 \in C$. 对任意 $x \in E$, 记[a]

$$p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in C\},$$

则 p 满足齐次性与三角不等式, 存在常数 M 使得 $0 \le p(x) \le M\|x\|(\forall x \in E)$, 且 $C = \{x \in E : p(x) < 1\}$.

[a]称 p 为 C 的度规或 Minkowski 泛函.

证明 p 的齐次性是显见的. 对于 p 的有界性, 取 $r > 0$ 满足 $B(0, r) \subset C$, 则

$$p(x) \le \frac{1}{r}\|x\|, \quad \forall x \in E.$$

取 $M = \frac{1}{r}$ 即可.

对于 C 的等价刻画, 设 $x \in C$, 因为 C 是开集, 故存在足够小的 $\varepsilon > 0$ 使得 $(1+\varepsilon)x \in C$, 因此 $p(x) \le \frac{1}{1+\varepsilon} < 1$, 这说明 $C \subset \{x \in E : p(x) < 1\}$. 反之, 若 $p(x) < 1$, 则存在 $\alpha \in (0, 1)$ 使得 $\alpha^{-1}x \in C$, 因此 $x = \alpha(\alpha^{-1}x) + (1-\alpha)0 \in C$, 故 $\{x \in E : p(x) < 1\} \subset C$.

对于三角不等式, 设 $x, y \in E, \varepsilon > 0$. 由齐次性与 C 的等价刻画知 $\frac{x}{p(x)+\varepsilon} \in C$, $\frac{y}{p(y)+\varepsilon} \in C$, 因此对任意 $t \in [0, 1]$ 均有

$$\frac{tx}{p(x)+\varepsilon} + \frac{(1-t)y}{p(y)+\varepsilon} \in C.$$

取

$$t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$$

可得

$$\frac{x+y}{p(x)+p(y)+2\varepsilon} \in C$$

再次由齐次性与 C 的等价刻画可得 $p(x+y) < p(x)+p(y)+2\varepsilon$ 对任意 $\varepsilon > 0$ 成立, 令 $\varepsilon \to 0$ 即得欲证. □

引理 2.3

设 $C \subset E$ 是非空开凸集, $x_0 \in E \backslash C$, 则存在 $f \in E^*$ 使得 $f(x) < f(x_0)(\forall x \in C)$. 特别地, 超平面 $[f = f(x_0)]$ 分离 $\{x_0\}$ 和 C.

证明 不妨设 $0 \in C$. 记 C 的度规为 p, 另记 $G = \mathbb{R}x_0$, 取线性泛函 $g : G \to \mathbb{R}$ 为

$$g(tx_0) = t, \quad t \in \mathbb{R}.$$

显见

$$g(x) \le p(x), \quad \forall x \in G.$$

于是由 Hahn-Banach 定理的解析形式2.1知在 E 上存在 g 的延拓 f, 其满足

$$f(x) \le p(x), \quad \forall x \in E.$$

特别地, 根据 g 的构造知 $f(x_0) = 1$, 且由 p 的有界性知 f 是连续线性泛函. 由 p 的三角不等式即得对任意 $x \in C$ 均有 $f(x) < 1$. □

下面来证明定理2.2.

证明 设 $C = A - B$, 可以说明 C 同样是凸集, 且因为 $C = \bigcup_{y \in B}(A - y)$, 故 C 也是开集, 另因 $A \cap B \neq \varnothing$, 故 $0 \notin C$. 由引理2.3知存在 $f \in E^*$ 使得

$$f(z) < 0, \quad \forall z \in C,$$

亦即

$$f(x) < f(y), \quad \forall x \in A, \forall y \in B.$$

取定常数 α 满足

$$\sup_{x\in A} f(x) \le \alpha \le \inf_{y\in B} f(y),$$

可见超平面 $[f=\alpha]$ 分离 A, B, 由 f 的连续性即得该超平面的闭性. □

定理 2.3 (Hahn-Banach 定理的第二几何形式)

设 $A \subset E, B \subset E$ 是满足 $A \cap B = \varnothing$ 的两个非空凸集, 另设 A 是闭集, B 是紧集, 则存在闭超平面严格分离 A, B.

证明 设 $C = A - B$, 可以说明 C 依旧是闭凸集, 且 $0 \notin C$. 因此存在 $r > 0$ 使得 $B(0,r) \cap C = \varnothing$, 由 Hahn-Banach 定理的第一几何形式2.2知存在闭超平面分离 $B(0,r)$ 和 C, 亦即存在 $f \in E^*$ 满足 $f \not\equiv 0$, 且

$$f(x-y) \le f(rz), \quad \forall x \in A, \forall y \in B, \forall z \in B(0,1).$$

这说明

$$f(x-y) \le -r\|f\|_{E^*}, \quad \forall x \in A, \forall y \in B.$$

记 $\varepsilon = \frac{1}{2} r\|f\| > 0$, 可得

$$f(x) + \varepsilon \le f(y) - \varepsilon, \quad \forall x \in A, \forall y \in B.$$

选取 α 满足

$$\sup_{x\in A} f(x) + \varepsilon \le \alpha \le \inf_{y\in B} f(y) - \varepsilon,$$

此即超平面 $[f=\alpha]$ 严格分离 A, B. □

注 如果只设 $A \subset E, B \subset E$ 是不交的非空凸集, 那么一般来说是没法用闭超平面分离 A, B 的. 甚至于就算 A, B 都是闭集, 也可以构造不能用闭超平面分离的反例. 然而, 如果 E 是有限维空间, 我们就总可以用闭超平面分离两个不交非空凸集.

从 Hahn-Banach 定理的几何形式出发, 可以得到下述推论:

推论 2.4

设 $F \subset E$ 是线性子空间, 且 $\overline{F} \neq E$, 则存在 $f \in E^*$ 使得 $f \not\equiv 0$ 的同时有

$$\langle f, x\rangle = 0, \quad \forall x \in F.$$

证明 取 $x_0 \in E \backslash \overline{F}$, 在 Hahn-Banach 定理的第二几何形式2.3中代入 $A = \overline{F}, B = \{x_0\}$ 知存在闭超平面 $[f=\alpha]$ 严格分离 $\overline{F}$ 和 $\{x_0\}$. 于是

$$\langle f, x\rangle < \alpha < \langle f, x_0\rangle, \quad \forall x \in F.$$

因为 F 本身关于乘法封闭, 故对任意 $\lambda \in \mathbb{R}$ 而言均有 $\lambda\langle f, x\rangle < \alpha$, 这说明只能有 $\langle f, x\rangle = 0 (\forall x \in F)$. □

注 推论2.4经常被用来证明线性子空间 $F \subset E$ 的稠密性. 也就是说, 要说明 $F \subset E$ 在 E 中稠, 只需说明在 F 上恒为零的连续线性泛函在 E 上也恒为零.

2.1.3 双重对偶空间 E^{**}, 正交关系

设 E 是赋范线性空间, E^* 是它的对偶空间, 其上范数定义为

$$\|f\|_{E^*} = \sup_{\substack{x\in E \\ \|x\|\le 1}} |\langle f, x\rangle|.$$

双重对偶空间 E^{**} 定义为 E^* 的对偶空间, 其上范数定义为

$$\|\xi\|_{E^{**}} = \sup_{\substack{f\in E^* \\ \|f\|\le 1}} |\langle \xi, f\rangle|, \quad \xi \in E^{**}.$$

从 E 到 E^{**} 存在一个**典则单射 (canonical injection)**J: 取定 $x\in E$, 知映射 $f\mapsto\langle f,x\rangle$ 是 E^* 上的连续线性泛函, 因此这个映射是 E^{**} 中的元素, 记该映射为 Jx, 则有:

$$\langle Jx,f\rangle_{(E^{**},E^*)}=\langle f,x\rangle_{(E^*,E)},\quad \forall x\in E,\forall f\in E^*. \tag{2.6}$$

显见 J 是线性的, 同时 J 是等距映射, 亦即 $\|Jx\|_{E^{**}}=\|x\|_E$. 这是因为

$$\|Jx\|_{E^{**}}=\sup_{\substack{f\in E^*\\ \|f\|\le 1}}|\langle Jx,f\rangle|=\sup_{\substack{f\in E^*\\ \|f\|\le 1}}|\langle f,x\rangle|=\|x\|_E.$$

一般来说, 典则映射 J 并不是 $E\to E^{**}$ 的满射, 不过通过这一映射, 我们可以把 E 等同于 E^{**} 的一个子空间. 如果 J 能够成为满射, 就称 E 是**自反空间**, 此时把 E^{**} 与 E 视作同一空间.

定义 2.3 (空间的正交)

若 $M\subset E$ 是线性子空间, 则记

$$M^\perp=\{f\in E^*:\langle f,x\rangle=0,\forall x\in M\}.$$

若 $N\subset E^*$ 是线性子空间, 则记

$$N^\perp=\{x\in E:\langle f,x\rangle=0,\forall f\in N\}.$$

根据定义, $N^\perp$ 是 E(而非 E^{**}) 的子集. 显见 $M^\perp$(或 $N^\perp$) 是 E^*(或 E) 的闭线性子空间. 称 $M^\perp$(或 $N^\perp$) 是正交于 M(或 N) 的空间.

♣

命题 2.2

设 $M\subset E$ 是线性子空间, 则

$$(M^\perp)^\perp=\overline{M}.$$

设 $N\subset E^*$ 是线性子空间, 则

$$(N^\perp)^\perp\supset\overline{N}.$$

♠

证明 显见 $M\subset(M^\perp)^\perp$, 又因为 $(M^\perp)^\perp$ 是闭集, 故 $\overline{M}\subset(M^\perp)^\perp$. 下面说明 $(M^\perp)^\perp\subset\overline{M}$. 若存在 $x_0\in(M^\perp)^\perp\backslash\overline{M}$, 由 Hahn-Banach 定理的第二几何形式2.3知存在闭超平面严格分离 $\{x_0\}$ 和 $\overline{M}$, 根据严格分离的定义知存在 $f\in E^*$ 与 $\alpha\in\mathbb{R}$ 使得

$$\langle f,x\rangle<\alpha<\langle f,x_0\rangle,\quad \forall x\in M.$$

因为 M 是线性空间, 故只能有 $\langle f,x\rangle=0(\forall x\in M)$, 且 $\langle f,x_0\rangle>0$. 因此 $f\in M^\perp$, 又由 $x_0\in(M^\perp)^\perp$ 知 $\langle f,x_0\rangle=0$, 矛盾!

$N\subset(N^\perp)^\perp$ 是显然的, 因此 $\overline{N}\subset(N^\perp)^\perp$. □

注 不同于 E 的情形, 对于 E^* 而言, $(N^\perp)^\perp$ 是可能严格大于 $\overline{N}$ 的. 我们不妨尝试证明 $(N^\perp)^\perp=\overline{N}$, 看看证明在哪个地方过不去. 设 $f_0\in E^*$ 满足 $f_0\in(N^\perp)^\perp\backslash\overline{N}$, 对 E^* 用 Hahn-Banach 定理知可以严格分离 $\{f_0\}$ 和 $\overline{N}$, 因此存在 $\xi\in E^{**}$ 使得 $\langle\xi,f_0\rangle>0$. 为了导出矛盾, 我们需要说明 $\langle\xi,f_0\rangle=0$, 但这一式子的成立需要 $\xi\in E$, 而 E^{**} 未必等于 E, 证明至此就过不去了. 也就是说, 只要 E 是自反空间, 那么就确实有 $(N^\perp)^\perp=\overline{N}$. 在一般情况下, 可以证明 $(N^\perp)^\perp$ 在弱 * 拓扑 $\sigma(E^*,E)$ 下与 N 的闭包重合.

2.1.4 共轭凸函数理论速览

共轭凸函数与凸分析理论在单调算子理论, 非线性半群理论与 PDE 中都有应用, 因此我们在这里讨论一下共轭凸函数. 我们从下半连续函数和凸函数的一些简单事实开始. 设 φ 是定义在 E 上, 在 $(-\infty,+\infty]$ 中取值的函数. 记 $D(\varphi)$ 是 φ 的定义域, 即

$$D(\varphi)=\{x\in E:\varphi(x)<+\infty\}.$$

定义 2.4 (上图 (epigraph))

函数 φ 的上图是集合[a]

$$\operatorname{epi}\varphi = \{(x,\lambda) \in E \times \mathbb{R} : \varphi(x) \le \lambda\}.$$

[a] 这里设 $\mathbb{R} = (-\infty, \infty)$, 因此 λ 不会取到 ∞.

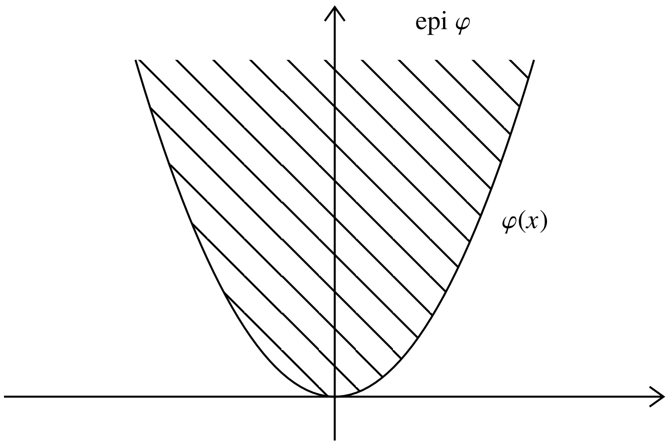

图 2.1: 上图示意.

设 E 是拓扑空间, 回忆下半连续的定义:

定义 2.5 (下半连续 (lower semicontinuous, l.s.c))

称函数 $\varphi : E \to (-\infty, +\infty]$ 下半连续, 如果对任意 $\lambda \in \mathbb{R}$ 而言, 集合

$$[\varphi \le \lambda] = \{x \in E : \varphi(x) \le \lambda\}$$

均为闭集.

下半连续函数有下述基本性质:

(i) 若 φ 下半连续, 则 $\operatorname{epi}\varphi$ 是 $E \times \mathbb{R}$ 中的闭集, 反之亦然.

(ii) 若 φ 下半连续, 则对任意 $x \in E$ 与任意 $\varepsilon > 0$, 存在 x 的某邻域使得

$$\varphi(y) \ge \varphi(x) - \varepsilon, \quad \forall y \in V$$

反之亦然. 特别地, 若 φ 下半连续, 则对 E 中任意趋向 x 的序列 $\{x_n\}$ 均有

$$\varliminf_{n\to\infty} \varphi(x_n) \ge \varphi(x)$$

若 E 是度量空间, 则反之亦然.

(iii) 若 φ_1, φ_2 都下半连续, 则 $\varphi_1 + \varphi_2$ 下半连续.

(iv) 若 $\{\varphi_i\}_{i\in I}$ 是下半连续函数族, 则其**上包络**是下半连续函数, 亦即

$$\varphi(x) := \sup_{i\in I} \varphi_i(x)$$

下半连续.

(v) 若 E 是紧集, φ 下半连续, 则 $\inf_E \varphi$ 能够达到.

现设 E 是线性空间, 回忆凸函数的定义:

定义 2.6 (凸函数 (convex function))

函数 $\varphi: E \to (-\infty, +\infty]$ 称为凸函数, 如果

$$\varphi(tx + (1-t)y) \le t\varphi(x) + (1-t)\varphi(y), \quad \forall x, y \in E, \forall t \in (0,1).$$

凸函数有下述基本性质:

(i) 若 φ 是凸函数, 则 $\operatorname{epi}\varphi$ 是 $E \times \mathbb{R}$ 中的凸集, 反之亦然.

(ii) 若 φ 是凸函数, 则对任意 $\lambda \in \mathbb{R}$, 集合 $[\varphi \le \lambda]$ 都是凸集, 但反之不成立.

(iii) 若 φ_1, φ_2 都是凸函数, 则 $\varphi_1 + \varphi_2$ 是凸函数.

(iv) 若 $\{\varphi_i\}_{i\in I}$ 是凸函数族, 则其上包络 $\sup_i \varphi_i$ 是凸函数.

下面默认 E 是赋范线性空间.

定义 2.7 (共轭函数 (conjugate function))

设 $\varphi: E \to (-\infty, +\infty]$ 是满足 $\varphi \not\equiv \infty$(亦即 $D(\varphi) \ne \varnothing$) 的函数, 定义其共轭函数 $\varphi^*: E^* \to (-\infty, +\infty]$ 为[a]

$$\varphi^*(f) = \sup_{x\in E}\{\langle f, x\rangle - \varphi(x)\}, \quad f \in E^*.$$

[a]有时称 φ^* 为 φ 的 Legendre 变换.

注意到 φ^* 总是 E^* 上的下半连续凸函数. 这是因为任取 $x \in E$, 函数 $f \mapsto \langle f, x\rangle - \varphi(x)$ 都是 E^* 上的连续凸函数 (进而是下半连续凸函数), 从而 x 跑遍 E 时形成的函数族的上包络同样是下半连续凸函数.

注 显见

$$\langle f, x\rangle \le \varphi(x) + \varphi^*(f), \quad \forall x \in E, \forall f \in E^*,$$

该不等式称为 Young 不等式. 如果令 $E = E^* = \mathbb{R}, \varphi(t) = \frac{1}{p}|t|^p, \varphi^*(s) = \frac{1}{p'}|s|^{p'}$(其中 $1 < p < \infty, \frac{1}{p} + \frac{1}{p'} = 1$), 那么上述不等式就成为了经典的 Young 不等式:

$$ab \le \frac{1}{p}a^p + \frac{1}{p'}b^{p'}, \quad \forall a, b \ge 0.$$

命题 2.3

设 $\varphi: E \to (-\infty, +\infty]$ 是下半连续凸函数, 且 $\varphi \not\equiv +\infty$, 则 $\varphi^* \not\equiv +\infty$. 特别地, φ 总在某仿射连续函数之上.

证明 设 $x_0 \in D(\varphi), \lambda_0 < \varphi(x_0)$. 在空间 $E \times \mathbb{R}$ 中对 $A = \operatorname{epi}\varphi, B = \{(x_0, \lambda_0)\}$ 用 Hahn-Banach 定理的第二几何形式2.3知存在 $E \times \mathbb{R}$ 中的闭超平面 $H = [\Phi = \alpha]$ 严格分离 A, B. 注意到函数 $E \ni x \mapsto \Phi((x, 0))$ 是 E 上的连续线性泛函, 故存在 $f \in E^*$ 使得 $\Phi((x,0)) = \langle f, x\rangle$. 取 $k = \Phi((0,1))$, 知

$$\Phi((x, \lambda)) = \langle f, x\rangle + k\lambda, \quad \forall (x, \lambda) \in E \times \mathbb{R}.$$

因为闭超平面 $[\Phi = \alpha]$ 严格分离 A, B, 可设在 A 上 $\Phi > \alpha$, 在 B 上 $\Phi < \alpha$, 则

$$\langle f, x\rangle + k\lambda > \alpha, \quad \forall (x, \lambda) \in \operatorname{epi}\varphi,$$

且

$$\langle f, x_0\rangle + k\lambda_0 < \alpha.$$

特别有

$$\langle f, x\rangle + k\varphi(x) > \alpha, \quad \forall x \in D(\varphi), \tag{2.7}$$

因此

$$\langle f, x_0\rangle + k\varphi(x_0) > \alpha > \langle f, x_0\rangle + k\lambda_0$$

这说明 $k>0$, 进而由(2.7)式知

$$\left\langle -\frac{1}{k}f,x\right\rangle - \varphi(x) < -\frac{\alpha}{k}, \quad \forall x\in D(\varphi),$$

因此 $\varphi^*\left(-\frac{1}{k}f\right)<+\infty$. $\left\langle -\frac{1}{k}f,x\right\rangle+\frac{\alpha}{k}$ 正是欲求的仿射连续函数. □

如果我们对 φ^* 再用一次算子 $*$, 就能得到定义在 E^{**} 上的函数 φ^{**}, 但 E^{**} 不如 E 好研究. 因为 $E\subset E^{**}$, 故我们谈 φ^{**} 时会把它限制在 E 上, 亦即

$$\varphi^{**}(x):=\sup_{f\in E^*}[\langle f,x\rangle-\varphi^*(f)], \quad x\in E.$$

定理 2.4 (Fenchel-Moreau)

若 $\varphi:E\to(-\infty,+\infty]$ 是下半连续凸函数, 且 $\varphi\not\equiv+\infty$, 则 $\varphi^{**}=\varphi$. ♡

证明 我们分两步来证明.

第一步: 另设 $\varphi\geq 0$, 往证 $\varphi^{**}=\varphi$.

由 φ^* 的定义知对 $\forall x\in E$ 与 $\forall f\in E^*$ 均有 $\langle f,x\rangle-\varphi^*(f)\leq\varphi(x)$, 因此 $\varphi^{**}\leq\varphi$. 为证 $\varphi^{**}=\varphi$, 考虑反证法, 设 $\exists x_0\in E$ 使得 $\varphi^{**}(x_0)<\varphi(x_0)$. 考虑证明 $\varphi(x_0)=+\infty$, 而 $\varphi^{**}(x_0)$ 始终有限. 现在在空间 $E\times\mathbb{R}$ 中对 $A=\operatorname{epi}\varphi, B=(x_0,\varphi^{**}(x_0))$ 应用 Hahn-Banach 定理的第二几何形式, 由命题2.3知 $\exists f\in E^*, k\in\mathbb{R}, \alpha\in\mathbb{R}$ 使得

$$\langle f,x\rangle+k\lambda>\alpha, \quad \forall (x,\lambda)\in\operatorname{epi}\varphi, \tag{2.8}$$

$$\langle f,x_0\rangle+k\varphi^{**}(x_0)<\alpha. \tag{2.9}$$

在(2.8)式中取定 $x\in D(\varphi)$ 并令 $\lambda\to+\infty$ 可知只能有 $k\geq 0$(这里我们没法像命题2.3的证明一样断言 $k>0$; $k=0$ 是可能出现的, 这种情况对应 $E\times\mathbb{R}$ 中的一个" 铅直" 超平面). 现取 $\varepsilon>0$, 因为 $\varphi\geq 0$, 故由(2.8)式知

$$\langle f,x\rangle+(k+\varepsilon)\varphi(x)\geq\alpha, \quad \forall x\in D(\varphi).$$

因此

$$\varphi^*\left(-\frac{f}{k+\varepsilon}\right)\leq-\frac{\alpha}{k+\varepsilon}.$$

由 $\varphi^{**}(x_0)$ 的定义可知

$$\varphi^{**}(x_0)\geq\left\langle-\frac{f}{k+\varepsilon},x_0\right\rangle-\varphi^*\left(-\frac{f}{k+\varepsilon}\right)\geq\left\langle-\frac{f}{k+\varepsilon},x_0\right\rangle+\frac{\alpha}{k+\varepsilon}.$$

因此

$$\langle f,x_0\rangle+(k+\varepsilon)\varphi^{**}(x_0)\geq\alpha, \quad \forall\varepsilon>0,$$

这便与(2.9)式矛盾了. $\varphi\geq 0$ 的情况得证.

第二步: 一般情况.

取定 $f_0\in D(\varphi^*)$(由命题2.3可知 $D(\varphi^*)\neq\varnothing$), 定义

$$\overline{\varphi}(x)=\varphi(x)-\langle f_0,x\rangle+\varphi^*(f_0)$$

知 $\overline{\varphi}$ 同样是下半连续凸函数, 且 $\overline{\varphi}\geq 0$. 根据 φ^* 的定义知 $\overline{\varphi}\geq 0$, 因此由第一步的结论知 $(\overline{\varphi})^{**}=\overline{\varphi}$, 下面计算 $(\overline{\varphi})^*$ 与 $(\overline{\varphi})^{**}$, 知

$$(\overline{\varphi})^*(f)=\varphi^*(f+f_0)-\varphi^*(f_0),$$

$$(\overline{\varphi})^{**}(x)=\varphi^{**}(x)-\langle f_0,x\rangle+\varphi^*(f_0).$$

因为 $(\overline{\varphi})^{**}=\overline{\varphi}$, 故 $\varphi^{**}=\varphi$. □

下面给出一些适用于 Fenchel-Moreau 定理的例子.

例 2.1 考察 $\varphi(x)=\|x\|$, 显见

$$\varphi^*(f)=\begin{cases}0, & \|f\|\leq 1,\\ +\infty, & \|f\|>1.\end{cases}$$

因此

$$\varphi^{**}(x) = \sup_{\substack{f\in E^* \\ \|f\|\le 1}} \langle f, x\rangle.$$

此时由 Fenchel-Moreau 定理得到的等式

$$\varphi^{**} = \varphi$$

实际上就是推论2.3.

例 2.2 取定非空集合 $K \subset E$, 设

$$I_K(x) = \begin{cases} 0, & x \in K, \\ +\infty, & x \notin K. \end{cases}$$

函数 I_K 称为 K 的**指示函数 (indicator function)**(这与示性函数是不同的). 显见 I_K 是凸函数当且仅当 K 是凸集, I_K 是下半连续函数当且仅当 K 是闭集. I_K 的共轭函数 $(I_K)^*$ 称为 K 的**支撑函数 (supporting function)**.

显见如果 $K = M$ 是线性子空间, 那么 $(I_M)^* = I_{M^\perp}, (I_M)^{**} = I_{(M^\perp)^\perp}$. 若 M 是闭子空间, 由 I_M 的下半连续凸性可得 $(I_M)^{**} = I_M$, 这实际上就是 $(M^\perp)^\perp = M$. 在这种意义下, Fenchel-Moreau 定理和命题2.2说的其实是一回事.

我们用共轭函数的另一有用性质结束本节.

定理 2.5 (Fenchel-Rockafellar)

设 $\varphi, \psi : E \to (-\infty, +\infty]$ 是凸函数, 且存在 $x_0 \in D(\varphi) \cap D(\psi)$ 使得 φ 在 x_0 处连续, 则

$$\begin{aligned}\inf_{x\in E}(\varphi(x)+\psi(x)) &= \sup_{f\in E^*}(-\varphi^*(-f)-\psi^*(f)) \\ &= \max_{f\in E^*}(-\varphi^*(-f)-\psi^*(f)) = -\min_{f\in E^*}(\varphi^*(-f)+\psi^*(f)).\end{aligned}$$

♡

Fenchel-Rockafellar 定理的证明需要下述引理:

引理 2.4

若 $C \subset E$ 是凸集, 则 C° 也是凸集. 另若 $C^\circ \neq \varnothing$, 则

$$\overline{C} = \overline{(C^\circ)}.$$

♡

引理2.4的证明就略过了, 下面证明 Fenchel-Rockafellar 定理.

证明 记

$$a = \inf_{x\in E}(\varphi(x)+\psi(x)),$$

$$b = \sup_{f\in E^*}(-\varphi^*(-f)-\psi^*(f)).$$

易证 $b \le a$. 若 $a = -\infty$, 则定理结论是显见的, 下设 $a \in \mathbb{R}$. 令 $C = \operatorname{epi}\varphi$, 因为 φ 至少在 x_0 处连续, 故 $C^\circ \neq \varnothing$. 现在对于 $A = C^\circ$ 与

$$B = \{(x,\lambda) \in E\times\mathbb{R} : \lambda \le a - \psi(x)\},$$

知 A, B 均为非空凸集. 另有 $A\cap B = \varnothing$, 这是因为若 $(x,\lambda) \in A$, 则 $\lambda > \varphi(x)$, 而另一方面由 a 的定义知 $\varphi(x) \ge a-\psi(x)$, 故 $(x,\lambda) \notin B$.

对 A, B 应用 Hahn-Banach 定理的第二几何形式可得闭超平面 H 分离 A, B, 进而 H 同样分离 $\overline{A}, B$. 但由引理2.4知 $\overline{A} = \overline{C}$, 故存在 $f \in E^*, k \in \mathbb{R}, \alpha \in \mathbb{R}$ 使得 $E\times\mathbb{R}$ 中的超平面 $H = [\Phi = \alpha]$ 分离 C, B, 其中

$$\Phi((x,\lambda)) = \langle f, x\rangle + k\lambda, \quad \forall (x,\lambda) \in E\times\mathbb{R}.$$

因此可设

$$\langle f, x\rangle + k\lambda \ge \alpha, \quad \forall (x,\lambda) \in C, \tag{2.10}$$

$$\langle f, x\rangle + k\lambda \le \alpha, \quad \forall (x,\lambda) \in B. \tag{2.11}$$

取 $x = x_0$, 在(2.10)式中令 $\lambda \to +\infty$ 可得 $k \geq 0$, 下面说明进一步有

$$k > 0. \tag{2.12}$$

考虑反证法, 设 $k = 0$, 则由 $\Phi \not\equiv 0$ 知 $\|f\| \neq 0$. 由(2.10),(2.11)两式可得

$$\langle f, x\rangle \geq \alpha, \quad \forall x \in D(\varphi),$$

$$\langle f, x\rangle \leq \alpha, \quad \forall x \in D(\psi).$$

又因为存在足够小的 $\varepsilon_0 > 0$ 使得 $B(x_0, \varepsilon_0) \subset D(\varphi)$, 故

$$\langle f, x_0 + \varepsilon_0 z\rangle \geq \alpha, \quad \forall z \in B(0,1),$$

这说明 $\langle f, x_0\rangle \geq \alpha + \varepsilon_0\|f\|$. 另一方面, 因为 $x_0 \in D(\psi)$, 故 $\langle f, x_0\rangle \leq \alpha$, 结合前述结果知只能有 $\|f\| = 0$, 矛盾! (2.12)式由此得证.

由(2.10),(2.11)两式可知

$$\varphi^*\left(-\frac{f}{k}\right) \leq -\frac{\alpha}{k}$$

且

$$\psi^*\left(\frac{f}{k}\right) \leq \frac{\alpha}{k} - a,$$

因此

$$-\varphi^*\left(-\frac{f}{k}\right) - \psi^*\left(\frac{f}{k}\right) \geq \alpha.$$

另一方面, 由 b 的定义可得

$$-\varphi^*\left(-\frac{f}{k}\right) - \psi^*\left(\frac{f}{k}\right) \leq b,$$

结合 $b \leq a$ 即得

$$a = b = -\varphi^*\left(-\frac{f}{k}\right) - \psi^*\left(\frac{f}{k}\right).$$

□

例 2.3 设 K 是非空凸集, 利用 Fenchel-Rockafellar 定理可以说明对任意 $x_0 \in E$ 均有

$$\operatorname{dist}(x_0, K) = \inf_{x\in K}\|x - x_0\| = \max_{\substack{f\in E^* \\ \|f\|\leq 1}}(\langle f, x_0\rangle - I_K^*(f)). \tag{2.13}$$

这是因为取 $\varphi(x) = \|x - x_0\|, \psi(x) = I_K(x)$ 可得

$$\inf_{x\in K}\|x - x_0\| = \inf_{x\in E}(\varphi(x) + \psi(x)).$$

由 Fenchel-Rockafellar 定理即得(2.13)式. 特别在 $K = M$ 为线性子空间时有

$$\operatorname{dist}(x_0, M) = \inf_{x\in M}\|x - x_0\| = \max_{\substack{f\in M^\perp \\ \|f\|\leq 1}}\langle f, x_0\rangle.$$

注 (2.13)式在 $\inf_{x\in K}\|x - x_0\|$ 没法达到时可以提供一些有用的信息. 这种情况多见于极小曲面理论: 此时原始问题 (即 $\inf_{x\in E}(\varphi(x) + \psi(x))$) 一般没有解, 但其对偶问题 (即 $\max_{f\in E^*}(-\varphi^*(-f) - \psi^*(f))$) 有解.

例 2.4 若 $\varphi : E \to \mathbb{R}$ 是连续凸函数, $M \subset E$ 是线性子空间, 则

$$\inf_{x\in M}\varphi(x) = -\min_{f\in M^\perp}\varphi^*(f).$$

在 Fenchel-Rockafellar 定理中令 $\psi = I_M$ 即得上述结论.

2.2 一致有界原理与闭图像定理

2.2.1 Baire 纲定理

下面的结果在本节的证明中至关重要.

定理 2.6 (Baire)

设 X 是完备距离空间, $\{X_n\}_{n\geq 1}$ 是 X 的闭子集列. 若

$$X_n^\circ = \varnothing, \quad \forall n \geq 1$$

则

$$\left(\bigcup_{n=1}^{\infty} X_n\right)^\circ = \varnothing.$$

注 通常还会用到 Baire 纲定理的下述形式. 设 X 是非空完备距离空间, $\{X_n\}_{n\geq 1}$ 是其中的闭集列, 且

$$\bigcup_{n=1}^{\infty} X_n = X,$$

则存在 n_0 使得 $X_{n_0}^\circ \neq \varnothing$.

证明 设 $O_n = X_n^c$, 则 O_n 是开集, 且对每个 $n \geq 1$ 而言, O_n 都在 X 中稠密, 往证 $G = \bigcap_{n=1}^{\infty} O_n$ 在 X 中稠密. 设 ω 是 X 中的非空开集, 我们考虑证明 $\omega \cap G \neq \varnothing$.

记

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

任取 $x_0 \in \omega, r_0 > 0$ 满足

$$\overline{B(x_0,r_0)} \subset \omega$$

因为 O_1 是开集且在 X 中稠密, 故可以进一步取 $x_1 \in B(x_0,r_0) \cap O_1, r_1 > 0$ 满足

$$\begin{cases} \overline{B(x_1,r_1)} \subset B(x_0,r_0) \cap O_1, \\ 0 < r_1 < \frac{r_0}{2}, \end{cases}$$

归纳可知存在序列 $\{x_n\}, \{r_n\}$ 使得

$$\begin{cases} \overline{B(x_{n+1},r_{n+1})} \subset B(x_n,r_n) \cap O_{n+1}, \ \forall n \geq 0, \\ 0 < r_{n+1} < \frac{r_n}{2}. \end{cases}$$

显见 $\{x_n\}$ 是 X 中的基本列, 设 $x_n \to l$.

因为 $x_{n+p} \in B(x_n,r_n)$ 对任意 $n \geq 0$ 与任意 $p \geq 0$ 均成立, 故令 $p \to \infty$ 可得

$$l \in \overline{B(x_n,r_n)}, \quad \forall n \geq 0.$$

特别有 $l \in \omega \cap G$, 此即欲证. □

2.2.2 一致有界原理

定理 2.7 (Banach-Steinhaus 一致有界原理 (共鸣定理))

设 E, F 是 Banach 空间, $\{T_i\}_{i\in I}$ 是 $E \to F$ 的连续线性算子族 (该算子族未必可数). 若

$$\sup_{i\in I} \|T_i x\| < \infty, \quad \forall x \in E. \tag{2.14}$$

则

$$\sup_{i\in I} \|T_i\|_{E\to F} < \infty. \tag{2.15}$$

也就是说, 存在常数 c 使得

$$\|T_i x\| \leq c\|x\|, \quad \forall x \in E, \forall i \in I.$$

注 一致有界原理的结论是相当出人意料的, 它将点态估计升级成了全局估计 (一致估计).

证明 对 $n \geq 1$, 记

$$X_n = \{x \in E : \forall i \in I, \|T_i x\| \leq n\},$$

由 T_i 的连续性知 X_n 是闭集, 由(2.14)式知

$$\bigcup_{n=1}^{\infty} X_n = E.$$

因为 $E \neq \varnothing$, 由 Baire 纲定理知存在 $n_0 \geq 1$ 使得 $X_{n_0}^{\circ} \neq \varnothing$. 取 $x_0 \in E, r > 0$ 满足 $B(x_0, r) \subset X_{n_0}$, 由 X_n 的构造可得

$$\|T_i(x_0 + rz)\| \leq n_0, \quad \forall i \in I, \forall z \in B(0,1).$$

由 T_i 的线性性与范数的三角不等式进而可得

$$r\|T_i\|_{E \to F} \leq n_0 + \|T_i x_0\|,$$

此即(2.15)式. □

从一致有界原理能推知下述结果.

推论 2.5

设 E, F 是 Banach 空间, $\{T_n\}$ 是 $E \to F$ 的连续线性算子列, 且对 $\forall x \in E$ 而言, $T_n x$ 在 $n \to \infty$ 时均收敛到 Tx, 则

(i) $\sup_n \|T_n\|_{E \to F} < \infty$,

(ii) $T \in L(E, F)$,

(iii) $\|T\|_{E \to F} \leq \liminf_{n \to \infty} \|T_n\|_{E \to F}$.

证明 由极限 Tx 的存在性知 (i) 是一致有界原理的直接推论, 进而存在常数 c 使得

$$\|T_n x\| \leq c\|x\|, \quad \forall n, \forall x \in E.$$

令 $n \to \infty$ 可得

$$\|Tx\| \leq c\|x\|, \quad \forall x \in E.$$

又因为 T 显然是线性的, 故 (ii) 得证. 最后因为

$$\|T_n x\| \leq \|T_n\|_{E \to F}\|x\|, \quad \forall x \in E,$$

令 $n \to \infty$ 即得 (iii).

推论 2.6

设 G 是 Banach 空间, $B \subset G$. 若对任意 $f \in G^*$ 而言, 集合 $f(B) = \{\langle f, x\rangle : x \in B\}$ 都是 $\mathbb{R}$ 中的有界集, 则 B 在 G 中有界.

证明 在一致有界原理中令 $E = G^*, F = \mathbb{R}, I = B$, 对任意 $b \in B$ 记

$$T_b(f) = \langle f, b\rangle, \quad f \in E = G^*,$$

于是由 $f(B)$ 的有界性知

$$\sup_{b \in B} |T_b(f)| < \infty, \quad \forall f \in E.$$

由一致有界原理进而知存在常数 c 使得

$$|\langle f, b\rangle| \leq c\|f\|, \quad \forall f \in G^*, \forall b \in B.$$

根据 Hhan-Banach 定理, 可以取到某个 $f \in G^*$ 使得 $|\langle f, b\rangle| = \|b\|, \|f\| = 1$, 因此

$$\|b\| \leq c, \quad \forall b \in B.$$

□

注 推论2.6表明要想证明某个集合 B 有界, 只需关注有界线性泛函在 B 上的作用即可. 这件事其实在有限维空间

我们干过很多次: 要说明某个集合有界, 只需说明该集合内的元素和全体线性基作内积均有界即可, 典例即各分量均有界的集合是有界的. 上述结果在某种意义下展示了无穷维空间中" 内积" 的运用. 有时我们也把推论2.6概括为" 弱有界"⇔" 强有界".

下面阐述推论2.6的对偶形式:

推论 2.7

设 G 是 Banach 空间, $B^* \subset G^*$. 若对任意 $x \in G$ 而言, 集合 $\langle B^*, x\rangle := \{\langle f, x\rangle : f \in B^*\}$ 都是 $\mathbb{R}$ 中的有界集, 则 B^* 在 G^* 中有界.

证明 在一致有界原理中令 $E = G, F = \mathbb{R}, I = B^*$, 对任意 $b \in B^8$ 记

$$T_b(x) = \langle b, x\rangle, \quad x \in G = E.$$

由条件可知存在常数 c 使得

$$|\langle b, x\rangle| \leq c\|x\|, \quad \forall b \in B^*, \forall x \in G.$$

进而由 $\|b\|$ 的定义即知

$$\|b\| \leq c, \quad \forall b \in B^*.$$

2.2.3 开映射定理与闭图像定理

本节介绍 Banach 空间的两个基本结果.

定理 2.8 (开映射定理)

设 E, F 是 Banach 空间, T 是 $E \to F$ 的连续线性满射, 则存在常数 $c > 0$ 使得

$$T(B_E(0,1)) \supset B_F(0,c). \tag{2.16}$$

注 (2.16)式表明 E 中的任意开集在作用 T 之后的象依旧是 F 中的开集. 这是因为设 U 是 E 中的开集, 任取 $y_0 \in T(U)$, 知存在 $x_0 \in U$ 使得 $y_0 = Tx_0$. 取 $r > 0$ 满足 $B(x_0, r) \subset U$, 知这可以转化为 $x_0 + B(0, r) \subset U$, 两边作用 T 可得

$$y_0 + T(B(0,r)) \subset T(U)$$

由(2.16)式得

$$T(B(0,r)) \supset B(0, rc)$$

因此

$$B(y_0, rc) \subset T(U).$$

从开映射定理出发可以得到下述用于证明算子逆连续性的重要结论.

推论 2.8

设 E, F 是 Banach 空间, T 是 $E \to F$ 的连续线性双射, 则 T^{-1} 是 $F \to E$ 的连续线性算子.

我们先证明推论2.8.

证明 由(2.16)式与 T 是单射知只要 $x \in E$ 满足 $\|Tx\| \leq c$, 就必有 $\|x\| < 1$, 由范数的齐次性可得

$$\|x\| \leq \frac{1}{c}\|Tx\|, \quad \forall x \in E$$

此即 T^{-1} 的连续性. □

另外, 开映射定理还能导出下面的范数等价定理:

推论 2.9 (范数等价定理)

设 E 是线性空间, 其上赋有范数 $\|\cdot\|_1$ 与 $\|\cdot\|_2$. 若 E 在这两种范数下都是 Banach 空间, 且存在常数 $C \geq 0$ 使得

$$\|x\|_2 \leq C\|x\|_1, \quad \forall x \in E.$$

则这两个范数等价, 亦即存在常数 $c > 0$ 使得

$$\|x\|_1 \leq c\|x\|_2, \quad \forall x \in E.$$

同样地, 我们先证明范数等价定理2.9.

证明 在推论2.8中代入

$$E = (E, \|\cdot\|_1), F = (E, \|\cdot\|_2), T = \mathrm{Id}$$

即可. □

下面我们着手证明开映射定理.

证明 我们把证明分成两部分进行.

$\boxed{\text{第一部分: 从 } T \text{ 的满射性推导它在单位球上的非退化性}}$

这部分设 T 是 $E \to F$ 的线性满射, 往证存在常数 $c > 0$ 使得

$$\overline{T(B(0,1))} \supset B(0,2c). \tag{2.17}$$

记 $X_n = n\overline{T(B(0,1))}$. 因为 T 是满射, 故 $\bigcup_{n=1}^{\infty} X_n = F$, 又因为 $F \neq \varnothing$, 故由 Baire 纲定理知存在 n_0 使得 $X_{n_0}^{\circ} \neq \varnothing$, 因此

$$(\overline{T(B(0,1))})^{\circ} \neq \varnothing.$$

取 $c > 0, y_0 \in F$ 满足

$$B(y_0, 4c) \subset \overline{T(B(0,1))}, \tag{2.18}$$

特别地, $y_0 \in \overline{T(B(0,1))}$, 再根据 T 的线性性与 $B(0,1)$ 关于原点的对称性可知

$$-y_0 \in \overline{T(B(0,1))}. \tag{2.19}$$

将(2.18),(2.19)两式相加可得

$$B(0,4c) \subset \overline{T(B(0,1))} + \overline{T(B(0,1))}.$$

另一方面, 因为 $\overline{T(B(0,1))}$ 是凸集, 故

$$\overline{T(B(0,1))} + \overline{T(B(0,1))} = 2\overline{T(B(0,1))}$$

此即(2.17)式.

$\boxed{\text{第二部分: 从 } T \text{ 的连续性与(2.17)式推导欲证结论.}}$

这部分考虑 T 是 $E \to F$ 的连续线性算子, 且其满足(2.17)式, 往证

$$T(B(0,1)) \supset B(0,c). \tag{2.20}$$

任取 $y \in F$ 满足 $\|y\| < c$, 要证明(2.20)式, 就是要找到 $x \in E$ 使得

$$\|x\| < 1, Tx = y. \tag{2.21}$$

由(2.17)式知

$$\forall \varepsilon > 0 \exists z \in E \left(\|z\| < \frac{1}{2} \wedge \|y - Tz\| < \varepsilon \right). \tag{2.22}$$

取 $\varepsilon = c/2$, 根据上式知存在 $z_1 \in E$ 使得

$$\|z_1\| < \frac{1}{2}, \|y - Tz_1\| < \frac{c}{2}.$$

把 y 换成 $y - Tz_1$, ε 换成 $c/4$, 重复这一过程知存在 $z_2 \in E$ 使得

$$\|z_2\| < \frac{1}{4}, \|(y - Tz_1) - Tz_2\| < \frac{c}{4}.$$

进一步重复该过程, 归纳可得序列 $\{z_n\}$ 满足

$$\|z_n\| < \frac{1}{2^n}, \|y - T(z_1 + z_2 + \cdots + z_n)\| < \frac{c}{2^n}, \quad \forall n.$$

显见序列 $x_n = z_1 + z_2 + \cdots + z_n$ 是 E 中的基本列, 设 $x_n \to x$, 由 T 的连续性即得 $\|x\| < 1, y = Tx$. □

定理 2.9 (闭图像定理)

设 E, F 是 Banach 空间, T 是 $E \to F$ 的线性算子. 若 T 的图 $G(T) = \{(x, Tx) : x \in E\}$ 是 $E \times F$ 中的闭集, 则 T 连续.

注 闭图像定理的逆命题是显然成立的, 因为任意连续映射 (不管线性与否) 的图都是闭集.

证明 在 E 上考虑两种范数:

$$\|x\|_1 = \|x\|_E + \|Tx\|_F, \|x\|_2 = \|x\|_E$$

(其中 $\|\cdot\|_1$ 称为图范数). 根据 $G(T)$ 的闭性知 E 在 $\|\cdot\|_1$ 下同样是 Banach 空间, 而 E 本身在 $\|\cdot\|_2$ 下是 Banach 空间, 且显见 $\|\cdot\|_2 \leq \|\cdot\|_1$. 故由范数等价定理知这两个范数等价, 因此存在常数 $c > 0$ 使得 $\|x\|_1 \leq c\|x\|_2$, 此即 $\|Tx\|_F \leq c\|x\|_E$, 因此 T 连续. □

2.2.4 拓扑余子空间, 右 (左) 可逆算子

我们先讨论由开映射定理出发得到的 Banach 空间闭子空间的一些几何性质.

定理 2.10

设 E 是 Banach 空间, 若 G, L 是两个闭子空间, 且 $G + L$ 是闭集, 则存在常数 $C \geq 0$ 使得

$$\begin{cases} \text{每个 } z \in G + L \text{ 均可分解为 } z = x + y, \text{ 其中 } x \in G, y \in L. \\ \text{且 } \|x\| \leq C\|z\|, \|y\| \leq C\|z\|. \end{cases} \tag{2.23}$$

证明 考虑乘积空间 $G \times L$, 在其上赋范

$$\|[x, y]\| = \|x\| + \|y\|.$$

另对空间 $G+L$ 赋 E 上的范数. 现在定义映射 $T : G \times L \to G + L$ 为 $T[x, y] = x + y$, 显见 T 是连续线性满射, 进而由开映射定理知存在常数 $c > 0$ 使得只要 $z \in G + L$ 满足 $\|z\| < c$, 它就能写成 $z = x + y$, 其中 $x \in G, y \in L, \|x\| + \|y\| < 1$. 根据范数的齐次性知每个 $z \in G + L$ 都能表为

$$z = x + y, \quad x \in G, y \in L, \|x\| + \|y\| \leq \frac{1}{c}\|z\|.$$

□

推论 2.10

在定理2.10的假设下, 存在常数 C 使得

$$\operatorname{dist}(x, G \cap L) \leq C(\operatorname{dist}(x, G) + \operatorname{dist}(x, L)), \quad \forall x \in E. \tag{2.24}$$

证明 取定 $x \in E, \varepsilon > 0$, 根据 dist 的定义知存在 $a \in G, b \in L$ 使得

$$\|x - a\| \leq \operatorname{dist}(x, G) + \varepsilon, \|x - b\| \leq \operatorname{dist}(x, L) + \varepsilon.$$

对 $z = a - b$ 应用定理2.10知存在 $a' \in G, b' \in L$ 使得

$$a - b = a' + b', \|a'\| \leq C\|a - b\|, \|b'\| \leq C\|a - b\|.$$

故 $a-a' \in G\cap L$, 且

$$\begin{aligned}\operatorname{dist}(x, G\cap L) &\le \|x-(a-a')\| \le \|x-a\|+\|a'\|\\ &\le \|x-a\|+C\|a-b\| \le \|x-a\|+C(\|x-a\|+\|x-b\|)\\ &\le (1+C)\operatorname{dist}(x,G)+\operatorname{dist}(x,L)+(1+2C)\varepsilon.\end{aligned}$$

令 $\varepsilon\to 0$ 即得欲证.

注 推论2.10的逆命题同样成立: 如果 G, L 是满足(2.24)式的闭子空间, 那么 $G+L$ 同样是闭子空间.

定义 2.8 (拓扑余子空间 (topological complement))

设 $G\subset E$ 是 Banach 空间 E 的闭子空间. 称 E 的子空间 L 为 G 的拓扑余子空间 (简称为余子空间), 如果

(i) L 是闭集,

(ii) $G\cap L=\{0\}, G+L=E$.

♣

上述情形成立时, 称 G, L 在 E 中互为余子空间, 此时每个 $z\in E$ 都能唯一地表为 $z=x+y, x\in G, y\in L$. 根据定理2.10, 投影算子 $z\mapsto x, z\mapsto y$ 是连续线性算子 (这一性质同样可以作为余子空间的定义).

例 2.5 每个有限维子空间 G 都具有余子空间. 这是因为设 $e_1,\cdots,e_n$ 为 G 的一组基, 显见每个 $x\in G$ 都能表为 $x=\sum_{i=1}^n x_ie_i$. 记 $\varphi_i(x)=x_i$, 由 Hahn-Banach 定理的解析形式知每个 φ_i 都能延拓为定义在 E 上的连续线性泛函 $\widetilde{\varphi}_i$, 可以验证 $L=\bigcap_{i=1}^n(\widetilde{\varphi}_i)^{-1}(0)$ 正是 G 的余子空间.

例 2.6 任意余维[2]有限的闭子空间 G 都具有余子空间. 只需找到使得 $G\cap L=\{0\}, G+L=E$ 的任意有限维空间 L(亦即 G 的有限维正交补空间) 即可, 下面介绍这种情形的典例.

设 $N\subset E^*$ 是 p 维子空间, 则

$$G=\{x\in E:\langle f,x\rangle=0,\forall f\in N\}=N^{\perp}$$

是余维为 p 的闭集. 这是因为取 $f_1,\cdots,f_p$ 为 N 的基, 则由 Hahn-Banach 定理知存在 $e_1,\cdots,e_p\in E$ 使得

$$\langle f_i,e_j\rangle=\delta_{ij},\quad \forall i,j=1,2,\cdots,p.$$

(这是因为对映射

$$x\in E\mapsto \mathbf{\Phi}(x)=(\langle f_1,x\rangle,\cdots,\langle f_p,x\rangle)$$

而言, 显见 $\mathbf{\Phi}$ 是满射, 否则由 Hahn-Banach 定理的第二几何形式知存在 $\alpha=(\alpha_1,\cdots,\alpha_p)\neq 0$ 使得

$$\alpha\cdot\mathbf{\Phi}(x)=\left\langle\sum_{i=1}^p\alpha_if_i,x\right\rangle=0,\quad \forall x\in E,$$

这显然是不可能的.)

显见 $\{e_i\}_{1\le i\le p}$ 线性独立, 因此该向量组张成的空间就是 G 的余子空间.

例 2.7 Hilbert 空间中的每个闭子空间都有余子空间.

注 需要注意就算在自反空间中, 我们也能构造不具有余子空间的闭子空间. Lindenstrauss-Tzafriri 表明: 任何不同构于 Hilbert 空间的 Banach 空间都存在不具有余子空间的闭子空间.

定义 2.9 (右逆算子 (right inverse), 左逆算子 (left inverse))

设 $T\in L(E,F)$, T 的右逆算子为算子 $S\in L(F,E)$, 其满足 $T\circ S=\mathrm{Id}_F$. T 的左逆算子为算子 $S\in L(F,E)$, 其满足 $S\circ T=\mathrm{Id}_E$.

♣

下面的结果给出了逆算子存在的充要条件.

[2]若存在 $X\subset E$ 使得 $G+X=E$, 就称 $\dim X$ 为 G 的余维数, 记作 $\operatorname{codim}G$.

定理 2.11 (右逆算子的存在性)

设 $T \in L(E.F)$ 是满射, 则下述命题等价:

(i) T 具有右逆算子.

(ii) $N(T) = T^{-1}(0)$ 在 E 中具有余子空间.

证明 (i)⇒(ii). 设 S 是 T 的右逆, 可以验证 $R(S) = S(F)$ 正是 $N(T)$ 在 E 中的余子空间.

(ii)⇒(i). 设 L 是 $N(T)$ 的余子空间, 令 P 为 $E \to L$ 的 (连续) 投影算子. 任取 $f \in F$, 记 x 为 $Tx = f$ 的解 (因为 T 是满射, 所以解必定存在). 现定义 $Sf = Px$, 注意到 S 的定义实际上与 x 的选择无关 (可能存在的 x 之间的差异被投影抹去了), 易证 $S \in L(F, E)$ 且 $T \circ S = \mathrm{Id}_F$. □

注 存在自反空间 E, F, 它们之间存在不具有右逆算子的连续线性满射. 例如对没有拓扑余子空间的闭子空间 $G \subset E$, 取 $F = E/G$, T 是 $E \to F$ 的典则投影.

定理 2.12 (左逆算子的存在性)

设 $T \in L(E, F)$ 是单射, 则下述命题等价:

(i) T 具有左逆算子.

(ii) $R(T) = T(E)$ 是闭子空间, 且它在 F 中具有余子空间.

证明 (i)⇒(ii). 记 $f = TS(f) + (f - TSf)$, 显见 $R(T)$ 是闭子空间, 且 $N(S)$ 是 $R(T)$ 的余子空间.

(ii)⇒(i). 设 P 是 $F \to R(T)$ 的连续投影算子. 取 $f \in F$, 因为 $Pf \in R(T)$, 故存在唯一的 $x \in E$ 使得 $Tx = Pf$. 令 $Sf = x$, 可见 $S \circ T = \mathrm{Id}_E$. 另由推论2.8可知 S 连续. □

2.2.5 再论正交关系

我们首先给出空间的和运算, 交运算与正交运算之间的关系.

命题 2.4

设 G, L 是 E 的闭子空间, 则

$$G \cap L = (G^\perp + L^\perp)^\perp, \tag{2.25}$$

$$G^\perp \cap L^\perp = (G + L)^\perp. \tag{2.26}$$

证明 (2.25): 首先证明 $G \cap L \subset (G^\perp + L^\perp)^\perp$. 取 $x \in G \cap L$, $f \in G^\perp + L^\perp$, 显见 $\langle f, x\rangle = 0$, 此即 $G \cap L \subset (G^\perp + L^\perp)^\perp$. 再说明 $(G^\perp + L^\perp)^\perp \subset G \cap L$. 显见 $G^\perp \subset G^\perp + L^\perp$, 因此 $(G^\perp + L^\perp)^\perp \subset G^{\perp\perp} = G$, 类似可得 $(G^\perp + L^\perp)^\perp \subset L$, 因此 $(G^\perp + L^\perp)^\perp \subset G \cap L$.

(2.26)的证明与上述证明是类似的. □

通过命题2.2, 容易得到下述结果:

推论 2.11

设 G, L 是 E 的闭子空间, 则

$$(G \cap L)^\perp \supset \overline{G^\perp + L^\perp}, \tag{2.27}$$

$$(G^\perp \cap L^\perp)^\perp = \overline{G + L}. \tag{2.28}$$

下面给出一些更深刻的结果.

定理 2.13

设 G, L 是 Banach 空间 E 的闭子空间, 则下述命题等价:

(i) $G + L$ 在 E 中闭,

(ii) $G^\perp + L^\perp$ 在 E^* 中闭,

(iii) $G + L = (G^\perp \cap L^\perp)^\perp$,

(iv) $G^\perp + L^\perp = (G \cap L)^\perp$.

证明 (i)⇔(iii) 由(2.28)式立得, (iv)⇒(ii) 显见, 下面证明 (i)⇒(iv),(ii)⇒(i).

(i)⇒(iv). 由(2.27)式知只需证明 $(G \cap L)^\perp \subset G^\perp + L^\perp$ 即可. 任取 $f \in (G \cap L)^\perp$, 对每个 $x \in G + L$, 记 $x = a + b, a \in G, b \in L$, 定义泛函 $\varphi : G + L \to \mathbb{R}$ 为

$$\varphi(x) = \langle f, a\rangle.$$

显见 φ 与 x 的分解方式无关, 且 φ 是线性映射. 由定理2.10知可选取 a, b 使得 $\|a\| \le C\|x\|$, 因此

$$|\varphi(x)| \le C\|x\|, \quad \forall x \in G + L.$$

利用 Hahn-Banach 定理的解析形式将 φ 延拓为 E 上的连续线性泛函 $\widetilde{\varphi}$, 可得:

$$f = (f - \widetilde{\varphi}) + \widetilde{\varphi}, \quad f - \widetilde{\varphi} \in G^\perp, \widetilde{\varphi} \in L^\perp.$$

此即 $(G \cap L)^\perp \subset G^\perp + L^\perp$.

(ii)⇒(i). 由推论2.10知存在常数 C 使得

$$\operatorname{dist}(f, G^\perp \cap L^\perp) \le C(\operatorname{dist}(f, G^\perp) + \operatorname{dist}(f, L^\perp)), \quad \forall f \in E^*. \tag{2.29}$$

另一方面, 在 Fenchel-Rockafellar 定理2.5中令 $\varphi(x) = I_{B_E}(x) - \langle f, x\rangle, \psi(x) = I_G(x)$(其中 $B_E = \{x \in E : \|x\| \le 1\}$) 可得

$$\operatorname{dist}(f, G^\perp) = \sup_{\substack{x \in G \\ \|x\| \le 1}} \langle f, x\rangle, \quad \forall f \in E^*. \tag{2.30}$$

类似可得

$$\operatorname{dist}(f, L^\perp) = \sup_{\substack{x \in L \\ \|x\| \le 1}} \langle f, x\rangle, \quad \forall f \in E^* \tag{2.31}$$

结合(2.26)式, 利用与上述结果相同的方法可得

$$\operatorname{dist}(f, G^\perp \cap L^\perp) = \operatorname{dist}(f, (G + L)^\perp) = \sup_{\substack{x \in \overline{G+L} \\ \|x\| \le 1}} \langle f, x\rangle, \quad \forall f \in E^*. \tag{2.32}$$

现结合(2.29)-(2.32)式可得

$$\sup_{\substack{x \in \overline{G+L} \\ \|x\| \le 1}} \langle f, x\rangle \le C\Big(\sup_{\substack{x \in G \\ \|x\| \le 1}} \langle f, x\rangle + \sup_{\substack{x \in L \\ \|x\| \le 1}} \langle f, x\rangle\Big), \quad \forall f \in E^*. \tag{2.33}$$

下面说明由(2.33)式可得

$$\overline{B_G + B_L} \supset \frac{1}{C} B_{\overline{G+L}}. \tag{2.34}$$

考虑反证法, 若存在 $x_0 \in \overline{G + L}$ 满足 $\|x_0\| \le \frac{1}{C}, x_0 \notin \overline{B_G + B_L}$, 则存在 E 中的闭超平面严格分离 $\{x_0\}$ 与 $\overline{B_G + B_L}$, 因此存在 $f_0 \in E^*$ 与 $\alpha \in \mathbb{R}$ 使得

$$\langle f_0, x\rangle < \alpha < \langle f_0, x_0\rangle, \quad \forall x \in B_G + B_L.$$

由此可得

$$\sup_{\substack{x \in G \\ \|x\| \le 1}} \langle f_0, x\rangle + \sup_{\substack{x \in L \\ \|x\| \le 1}} \langle f_0, x\rangle \le \alpha < \langle f_0, x_0\rangle,$$

这与(2.33)式矛盾! (2.34)式即证.

最后, 考察空间 $X = G \times L$, 在其上赋范

$$\|(x, y)\| = \max(\|x\|, \|y\|),$$

以及 $Y=\overline{G+L}$, 在其上赋 E 的范数. 显见映射 $T:X\to Y,(x,y)\mapsto x+y$ 是连续线性的, 进而由(2.34)式知

$$\overline{T(B_X)}\supset\frac{1}{C}B_Y.$$

利用开映射定理证明过程的第二部分可得

$$T(B_X)\supset\frac{1}{2C}B_Y$$

这说明 T 实际上是 $X\to Y$ 的满射, 故 $G+L=\overline{G+L}$.

2.2.6 无界线性算子引论, 共轭算子的定义

定义 2.10 (无界线性算子 (unbounded linear operator))

设 E,F 是 Banach 空间, 称线性映射 $A:D(A)\subset E\to F$ 为 $E\to F$ 的无界线性算子, 如果它定义在线性子空间 $D(A)\subset E$ 上且在 F 中取值. $D(A)$ 称为 A 的定义域.

♣

回忆有界算子 A 指的是 $D(A)=E$ 且存在常数 $c\geq 0$ 使得

$$\|Au\|\leq c\|u\|,\quad \forall u\in E.$$

注 无界线性算子显然可能有界, 因此" 无界" 这个词其实不太妥当, 但通常都是这么用的.

关于无界线性算子需要注意下述定义:

$$\begin{aligned}
A\text{ 的图 (graph)}&=G(A)=\{(u,Au):u\in D(A)\}\subset E\times F,\\
A\text{ 的像集 (range)}&=R(A)=\{Au:u\in D(A)\}\subset F,\\
A\text{ 的核 (kernal)}&=N(A)=\{u\in D(A):Au=0\}\subset E.
\end{aligned}$$

回忆如果 $G(A)$ 在 $E\times F$ 中是闭集, 就称 A 为闭算子.

注 一般来说我们用下述方法证明 A 是闭算子: 取 $D(A)$ 中的序列 $\{u_n\}$ 使得在 E 中 $u_n\to u$, 在 F 中 $Au_n\to f$, 往证 $u\in D(A), f=Au$. 因为只谈论闭算子时没有涉及线性性, 故这套证明中不能只考虑 $u_n\to 0, Au_n\to F$ 的情况.

注 在实际应用中, 我们碰到的无界算子大多都是闭算子, 且其定义域 $D(A)$ 在 E 中稠密.

定义 2.11 (无界算子的共轭算子 (adjoint))

设 $A:D(A)\subset E\to F$ 是定义域在 E 中稠密的无界线性算子, 我们依照下述方法定义无界算子 $A^*:D(A^*)\subset F^*\to E^*$. 令

$$D(A^*)=\{v\in F^*:\exists c>0(|\langle v,Au\rangle|\leq c\|u\|,\forall u\in D(A))\}.$$

显见 $D(A^*)$ 是 F^* 的一个线性子空间, 下面定义 A^*v. 取定 $v\in D(A^*)$, 定义映射 $g:D(A)\to\mathbb{R}$ 为

$$g(u)=\langle v,Au\rangle,\quad \forall u\in D(A).$$

知

$$|g(u)|\leq c\|u\|,\quad \forall u\in D(A).$$

于是由 Hahn-Banach 定理的解析形式知存在线性映射 $f:E\to\mathbb{R}$ 作为 g 的延拓满足

$$|f(u)|\leq c\|u\|,\quad \forall u\in E.$$

因此 $f\in E^*$. 又因为 $D(A)$ 在 E 中稠密, 故该延拓唯一. 定义

$$A^*v=f.$$

无界线性算子 $A^*:D(A^*)\subset F^*\to E^*$ 就称为 A 的共轭算子. 简单来说, A 与 A^* 之间的关系为:

$$\langle v,Au\rangle_{F^*,F}=\langle A^*v,u\rangle_{E^*,E},\quad \forall u\in D(A),\forall v\in D(A^*).$$

♣

注 其实前面在延拓 g 的时候不需要用到 Hahn-Banach 定理. 因为 $D(A)$ 在 E 中稠密, g 在 $D(A)$ 上一致连续, 且 $\mathbb{R}$ 完备, 故只需要用连续性即可完成延拓.

注 就算 A 是闭算子, $D(A^*)$ 也可能在 F^* 中不稠密, 但这种情况相当病态. 事实上, 如果 A 是闭算子, 那么在弱 * 拓扑 $\sigma(F^*, F)$ 下 $D(A^*)$ 确实在 F^* 中稠密. 特别当 F 是自反空间时, $D(A^*)$ 在通常的范数拓扑下在 F^* 中稠密.

注 如果 A 是有界算子, 那么 A^* 同样是 $F^* \to E^*$ 的有界算子, 且

$$\|A^*\|_{F^* \to E^*} = \|A\|_{E \to F}.$$

这是因为 A 有界时显见 $D(A^*) = F^*$. 根据 A^* 的定义知

$$|\langle A^*v, u\rangle| = |\langle v, Au\rangle| \le \|A\|\|u\|\|v\|, \quad \forall u \in E, \forall v \in F^*.$$

这说明 $\|A^*v\| \le \|A\|\|v\|$, 进而 $\|A^*\| \le \|A\|$.

另一方面, 知

$$|\langle v, Au\rangle| = |\langle A^*v, u\rangle| \le \|A^*\|\|u\|\|v\|, \quad \forall u \in E, \forall v \in F^*,$$

由 Hahn-Banach 定理的解析形式, 通过选取恰当的 v 即知 $\|Au\| \le \|A^*\|\|u\|$, 因此 $\|A\| \le \|A^*\|$.

命题 2.5

设 $A : D(A) \subset E \to F$ 是定义在 E 的某稠密子集上的无界线性算子, 则 A^* 是闭算子 (即 $G(A^*)$ 在 $F^* \times E^*$ 中闭).

证明 取 $v_n \in D(A^*)$ 满足 $v_n \to v$ 在 F^* 中成立, $A^*v_n \to f$ 在 E^* 中成立, 往证 $v \in D(A^*), A^*v = f$.

根据共轭算子的定义知

$$\langle v_n, Au\rangle = \langle A^*v_n, u\rangle, \quad \forall u \in D(A).$$

在上式两端令 $n \to \infty$ 有

$$\langle v, Au\rangle = \langle f, u\rangle, \quad \forall u \in D(A).$$

于是对 $\forall u \in D(A)$ 有 $|\langle v, Au\rangle| \le \|f\|\|u\|$, 此即 $v \in D(A^*)$, 根据 A^* 的定义立得 $A^*v = f$. □

A 的图与 A^* 的图之间有一个非常简单的正交关系: 考察同构:

$$I : F^* \times E^* \to E^* \times F^*, \quad I((v, f)) = (-f, v). \tag{2.35}$$

设 $A : D(A) \subset E \to F$ 是定义在 E 的某稠密子集上的无界线性算子, 则

$$I(G(A^*)) = G(A)^{\perp}.$$

这是因为取 $(v, f) \in F^* \times E^*$, 可得

$$\begin{aligned}(v, f) \in G(A^*) &\Leftrightarrow \langle f, u\rangle = \langle v, Au\rangle, \quad \forall u \in D(A)\\ &\Leftrightarrow -\langle f, u\rangle + \langle v, Au\rangle = 0, \quad \forall u \in D(A)\\ &\Leftrightarrow (-f, v) \in G(A)^{\perp}.\end{aligned}$$

下面再探讨共轭算子与原算子的像集与核之间的正交关系.

推论 2.12

设 $A : D(A) \subset E \to F$ 是定义在 E 的某稠密子集上的无界闭线性算子, 则

(i) $N(A) = R(A^*)^{\perp}$,

(ii) $N(A^*) = R(A)^{\perp}$,

(iii) $N(A)^{\perp} \supset \overline{R(A^*)}$,

(iv) $N(A^*)^{\perp} = \overline{R(A)}$.

证明 由命题2.2可直接由 (i),(ii) 推知 (iii),(iv), 故只需证明 (i),(ii). 记 $X = E \times F, X^* = E^* \times F^*$, 另记

$$G = G(A), L = E \times \{0\}$$

为 X 的子空间. 易证

$$N(A)\times\{0\}=G\cap L, \tag{2.36}$$

$$E\times R(A)=G+L. \tag{2.37}$$

显见 $L^{\perp}=\{0\}\times E^{*}$, 又由(2.35)式知 $G^{\perp}=I(G(A^{*}))$, 于是

$$\{0\}\times N(A^{*})=G^{\perp}\cap L^{\perp}. \tag{2.38}$$

根据 A 的线性性, $D(A)$ 至少是对加法封闭的, 故

$$R(A^{*})\times F^{*}=G^{\perp}+L^{\perp}. \tag{2.39}$$

下面来证明推论2.12. 对于 (i), 由(2.39)式知

$$\begin{aligned} R(A^{*})^{\perp}\times\{0\}=(G^{\perp}+L^{\perp})^{\perp}&\overset{(A)}{=}G\cap L\\ &\overset{(B)}{=}N(A)\times\{0\}. \end{aligned}$$

其中 (A) 是(2.25)式; (B) 是(2.36)式. 对于 (ii), 由(2.37)式知

$$\begin{aligned} \{0\}\times R(A)^{\perp}=(G+L)^{\perp}&\overset{(C)}{=}G^{\perp}\cap L^{\perp}\\ &\overset{(D)}{=}N(A)\times\{0\} \end{aligned}$$

其中 (C) 是(2.26)式; (D) 是(2.38)式.

注 对于 (iii) 需要特别注意: 就算 A 是有界线性算子, $N(A)^{\perp}$ 也可能不等于 $\overline{R(A^{*})}$. 但在弱 * 拓扑 $\sigma(E^{*},E)$ 下 $N(A)^{\perp}$ 总是 $R(A^{*})$ 的闭包, 特别若 E 自反, 则 $N(A)^{\perp}=\overline{R(A^{*})}$.

2.2.7 闭像集算子的刻画, 满射算子的刻画

定理 2.14 (闭像集算子的刻画)

设 $A:D(A)\subset E\to F$ 是定义在 E 的某稠密子集上的无界闭线性算子, 则下述命题等价:

(i) $R(A)$ 是闭集,

(ii) $R(A^{*})$ 是闭集,

(iii) $R(A)=N(A^{*})^{\perp}$,

(iv) $R(A^{*})=N(A)^{\perp}$.

♡

证明 采用与推论2.12相同的记号知:

$$\begin{aligned} &\text{(i)}\Leftrightarrow G+L\text{ 在 }X\text{ 中闭 (见(2.37)式)},\\ &\text{(ii)}\Leftrightarrow G^{\perp}+L^{\perp}\text{ 在 }X^{*}\text{ 中闭 (见(2.39)式)},\\ &\text{(iii)}\Leftrightarrow G+L=(G^{\perp}\cap L^{\perp})^{\perp}\text{(见(2.37),(2.38)式)},\\ &\text{(iv)}\Leftrightarrow(G\cap L)^{\perp}=G^{\perp}+L^{\perp}\text{(见(2.36),(2.39)式)}. \end{aligned}$$

由定理2.13即得欲证. □

注 设 $A:D(A)\subset E\to F$ 是无界闭线性算子, 则 $R(A)$ 是闭集当且仅当存在常数 C 使得

$$\mathrm{dist}(u,N(A))\le C\|Au\|,\quad \forall u\in D(A).$$

下面的结果给出了满射算子的一个常用刻画.

定理 2.15 (满射算子的刻画)

设 $A:D(A)\subset E\to F$ 是定义在 E 的某稠密子集上的无界闭线性算子, 则下述命题等价:

(i) A 是满射 (即 $R(A)=F$),

(ii) 存在常数 C 使得

$$\|v\|\le C\|A^{*}v\|,\quad \forall v\in D(A^{*}),$$

(iii) $N(A^*) = \{0\}$, 且 $R(A^*)$ 是闭集.

注 (ii)⇒(i) 在实际应用中常用于证明算子 A 是满射 (在 PDE 中对应的正是方程的可解性), 过程如下: 设 v 满足 $A^*v = f$, 往证 $\|v\| \le C\|f\|$(其中 C 与 f 无关). 该方法称为**先验估计 (a priori estimates)**, 在这一方法中我们并不需要关心方程 $A^*v = f$ 是否有解, 而是**先验地**假定 v 就是方程的解, 并尝试估计 v 的范数.

证明 (i)⇒(ii). 记

$$B^* = \{v \in D(A^*) : \|A^*v\| \le 1\}.$$

下面先证明 B^* 有界. 因为 $B^* \subset F^*$, 由一致有界原理知只需证明对 $\forall f_0 \in F$ 而言, 集合 $\langle B^*, f_0\rangle$ 在 $\mathbb{R}$ 中有界. 因为 A 是满射, 故对这样的 $f_0 \in F$ 总会存在 $u_0 \in D(A)$ 使得 $Au_0 = f_0$. 因此对任意 $v \in B^*$ 有

$$\langle v, f_0\rangle = \langle v, Au_0\rangle = \langle A^*v, u_0\rangle$$

于是

$$|\langle v, f_0\rangle| \le \|u_0\|.$$

在 B^* 的有界性得证后, 任取 $v \in D(A^*)$, 由范数的齐次性知

$$\left\|A^*\left(\frac{v}{2\|A^*v\|}\right)\right\| \le 1$$

因此 $v/2\|A^*v\| \in B^*$, 亦即存在 $C > 0$ 使得 $\|v\| \le C\|A^*v\|$.

(ii)⇒(iii). 只要 $A^*v = 0$, 由 (ii) 立得 $v = 0$, 此即 $N(A^*) = 0$. 再设 $f_n = A^*v_n \to f$, 将 $v_n - v_m$ 代入 (ii) 可知 $\{v_n\}$ 自身也是基本列, 因此可设 $v_n \to v$. 由命题2.5知 A^* 是闭算子, 故 $A^*v = f \in R(A)$.

(iii)⇒(i). 因为 $R(A^*)$ 是闭集, 由闭像集算子的刻画2.14知 $R(A) = N(A^*)^\perp$, 又由 $N(A^*) = \{0\}$ 知 $N(A^*)^\perp = F$, 此即 $R(A) = F$. □

上述结果具有下述对偶表述:

定理 2.16

设 $A : D(A) \subset E \to F$ 是定义在 E 的某稠密子集上的无界闭线性算子, 则下述命题等价:

(i) A^* 是满射 (即 $R(A^*) = E^*$),

(ii) 存在常数 C 使得

$$\|u\| \le C\|Au\|, \quad \forall u \in D(A),$$

(iii) $N(A) = \{0\}$, 且 $R(A)$ 是闭集.

证明 (i)⇒(ii). 记

$$B = \{u \in D(A) : \|Au\| \le 1\},$$

往证 B 有界, 由一致有界原理的推论2.6知只需证明对 $\forall f \in E^*$ 而言, 集合 $\langle f, B\rangle$ 在 $\mathbb{R}$ 中有界即可. 因为 A^* 是满射, 故总能找到 $v \in F^*$ 使得 $A^*v = f$. 因此对任意 $u \in B$ 有

$$\langle f, u\rangle = \langle A^*v, u\rangle = \langle v, Au\rangle$$

于是

$$|\langle f, u\rangle| \le \|v\|.$$

(ii) 进而由范数的齐次性得证.

(ii)⇒(iii). 显见 $N(A) = \{0\}$. 设 $f_n = Au_n \to f$, 由 (ii) 可知 $\{u_n\}$ 自身也是基本列, 因此可设 $u_n \to u$, 由 A 的闭性即得 $Au = f \in R(A)$.

(iii)⇒(i). 由闭像集算子的刻画2.14(iv) 立得结论. □

注 如果另外知道要么 $\dim E < \infty$, 要么 $\dim F < \infty$, 则有下述等价成立:

$$A \text{ 是满射} \Leftrightarrow A^* \text{ 是单射},$$

$$A^* \text{ 是满射} \Leftrightarrow A \text{ 是单射}.$$

这是因为此时 $R(A)$ 和 $R(A^*)$ 为有限维空间, 进而它们是闭空间. 但在一般情形下, 只有下述关系成立:

$$A \text{ 是满射} \Rightarrow A^* \text{ 是单射},$$

$$A^* \text{ 是满射} \Rightarrow A \text{ 是单射}.$$

反命题不再成立的例子如下: 取 $E = F = l^2$, 对每个 $x \in l^2$ 记 $x = \{x_n\}_{n\geq 1}$, 取 $Ax = \left\{\frac{1}{n}x_n\right\}_{n\geq 1}$. 显见 A 是有界算子, 且 $A^* = A$, A^* 和 A 此时都是单射, 但它们不是满射. 同时 $R(A)$ 和 $R(A^*)$ 都在对应空间稠密, 但它们不是闭集.

第三章 弱拓扑, 自反空间, 可分空间, 一致凸性

本章选自 [**HB**].

3.1 使一族映射连续的最粗糙拓扑

我们从拓扑中最基本的一个概念开始. 设 X 是 (没有赋任何结构的) 一个集合, $\{Y_i\}_{i\in I}$ 是一族拓扑空间. 现在有一族映射 $\{\varphi_i\}_{i\in I}$ 满足对每个 $i\in I$ 而言, φ_i 都是 $X\to Y_i$ 的. 考虑下述问题:

$\boxed{\text{问题一.}}$ 在 X 上构造一个拓扑使得所有映射 $\{\varphi_i\}_{i\in I}$ 均连续. 如果可能的话找出满足条件的最经济的拓扑 $\mathscr{T}$(即具有最少开集的拓扑).

注意到如果我们给 X 装备离散拓扑 (即 X 的每个子集都是开集), 自然会有每个 φ_i 都连续, 但这个拓扑肯定不是最经济的, 实际上它是最不经济的! 之后我们会发现, 实际上在 X 上总是存在 (唯一) 的最经济拓扑 $\mathscr{T}$ 使得每个 φ_i 都连续. 该拓扑称为 $\{\varphi_i\}_{i\in I}$ 生成的最粗糙拓扑 (或最弱拓扑).

若 $\omega_i\subset Y_i$ 是开集, 根据连续性的要求, $\varphi_i^{-1}(\omega_i)$ 也需要是 $\mathscr{T}$ 中的开集. 随着 ω_i 跑遍 Y_i 中所有开集, i 跑遍 I 中所有指标, 我们可以得到由 $\varphi_i(\omega_i)$ 形成的 X 的一个子集族, 其中的每个元素都需要在拓扑 $\mathscr{T}$ 中. 记该子集族为 $\{U_\lambda\}_{\lambda\in\Lambda}$, 显见 $\{U_\lambda\}_{\lambda\in\Lambda}$ 本身不一定是拓扑, 故我们可以将问题归结为:

$\boxed{\text{问题二.}}$ 对给定的集合 X 与 X 中的子集族 $\{U_\lambda\}_{\lambda\in\Lambda}$, 构造最经济的拓扑 $\mathscr{T}$ 使得对全体 $\lambda\in\Lambda$ 而言, U_λ 都是开集.

也就是说, 我们需要找到 X 的某个最经济的子集族 $\mathscr{F}$, 使得该子集族对有限交与任意并封闭, 且对 $\forall\lambda\in\Lambda$ 均有 $U_\lambda\in\mathscr{F}$. 我们可以从 $\{U_\lambda\}_{\lambda\in\Lambda}$ 中的有限交开始构造, 亦即形如 $\bigcap_{\lambda\in\Gamma}U_\lambda(\Gamma\subset\Lambda$ 是有限集) 的集合. 将所有这些有限交收集起来可以得到一个新集族, 记之为 Φ, 可见 Φ 是包含 $\{U_\lambda\}_{\lambda\in\Lambda}$ 且对有限交封闭的 X 的子集族, 但它依旧未必对任意并封闭. 因此, 我们下一步考察由 Φ 中全体元素的任意并构成的集族 $\mathscr{F}$. 显见 $\mathscr{F}$ 对任意并封闭, 但经过这一改动后没法确定 $\mathscr{F}$ 是否还对有限交封闭, 于是我们给出下述引理:

引理 3.1

集族 $\mathscr{F}$ 确对有限交封闭. ♡

引理的证明就省略了.

注 **特别注意前面构造 $\mathscr{F}$ 的顺序是不能颠倒的.** 虽说先取任意并再取有限交也很自然, 但这样生成的集族只对有限交封闭, 而不会再对任意并封闭. 要想得到对任意并的封闭性, 还需要再取一次任意并.

总结一下前述讨论, 对拓扑 $\mathscr{F}$ 而言, 其中的开集都是由形如 $\varphi_i^{-1}(\omega_i)$ 的集合的有限交的任意并构成的. 因此对每个取定的 $x\in X$, 设 V_i 是 $\varphi_i(x)$ 在 Y_i 中的邻域, x 在拓扑 $\mathscr{F}$ 下的邻域基就由形如 $\bigcap_{\text{finite}}\varphi_i^{-1}(V_i)$ 的集合组成[1].

之后我们就可以在 X 上赋使得映射族 $\{\varphi_i\}_{i\in I}$ 连续的最弱拓扑 $\mathscr{F}$ 了. 下面给出拓扑 $\mathscr{F}$ 的两条基本简单性质.

命题 3.1 (最弱拓扑的收敛性)

设 $\{x_n\}$ 是 X 中的序列, 则 $x_n\to x$ 在 $\mathscr{F}$ 下成立当且仅当 $\varphi_i(x_n)\to\varphi_i(x)$ 对任意 $i\in I$ 成立. ♠

证明 若 $x_n\to x$, 因为每个 φ_i 都在 $\mathscr{F}$ 下连续, 故 $\varphi_i(x_n)\to\varphi_i(x)$. 反之, 设 U 是 x 的某邻域, 根据前述邻域基的构造可设 $U=\bigcap_{i\in J}\varphi_i^{-1}(V_i)$(其中 $J\subset I$ 是有限集). 现任取 $i\in J$, 根据 $\varphi_i(x_n)\to\varphi_i(x)$ 知存在正整数 N_i 使得 $n\geq N_i$ 时有 $\varphi_i(x_n)\in V_i$, 于是 $n\geq N=\max_{i\in J}N_i$ 时 $x_n\in U$, 此即 $x_n\to x$. □

命题 3.2 (映射在最弱拓扑下的连续性)

设 Z 是拓扑空间, ψ 是 $Z\to X$ 的映射, 则 ψ 连续当且仅当对每个 $i\in I$ 而言, 映射 $\varphi_i\circ\psi:Z\to Y_i$ 均连续. ♠

---

[1]回忆在拓扑空间下, 点 x 的邻域基本质上是 x 的一族邻域, 其满足 x 的每个邻域都会包含该集族中的某元素.

证明 若 ψ 连续, 则显见对每个 $i \in I$ 而言 $\varphi_i \circ \psi$ 均连续. 反之, 往证对 X 中的每个开集 U 而言 $\psi^{-1}(U)$ 都是 Z 中的开集. 根据 $\mathscr{F}$ 的构造知 U 形如 $\bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} \varphi_i^{-1}(\omega_i)$, 其中 ω_i 是 Y 中的开集. 故

$$\psi^{-1}(U) = \bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} \psi^{-1}(\varphi_i^{-1}(\omega_i)) = \bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} (\varphi_i \circ \psi)^{-1}(\omega_i),$$

因为 $\varphi_i \circ \psi$ 连续, 故 $\psi^{-1}(U)$ 是 Z 中的开集. □

3.2 弱拓扑 $\sigma(E,E^*)$ 的定义与基本性质

设 E 是 Banach 空间, $f \in E^*$. 记 $\varphi_f : E \to \mathbb{R}$ 为线性泛函 $\varphi_f(x) = \langle f, x\rangle$. 随着 f 跑遍 E^*, 可以得到 $E \to \mathbb{R}$ 的泛函族 $\{\varphi_f\}_{f\in E^*}$. 下面我们在 E(由 $\|\cdot\|$ 诱导的) 的通常拓扑之外再定义 E 上的一个新拓扑:

定义 3.1 (弱拓扑 (weak topology))

E 上的弱拓扑 $\sigma(E,E^*)$ 指的是泛函族 $\{\varphi_f\}_{f\in E^*}$ 生成的最弱拓扑. ♣

注意到每个映射 φ_f 在通常拓扑下已经连续了, 故弱拓扑只会弱于通常拓扑.

命题 3.3

弱拓扑 $\sigma(E,E^*)$ 是 Hausdorff 拓扑 (即弱拓扑是分离的). ♠

证明 取 $x_1, x_2 \in E, x_1 \neq x_2$, 往证弱拓扑 $\sigma(E,E^*)$ 中存在开集 O_1, O_2 使得 $x_1 \in O_1, x_2 \in O_2, O_1 \cap O_2 = \varnothing$. 根据 Hahn-Banach 定理的第二几何形式, 存在闭超平面严格分离 $\{x_1\}, \{x_2\}$, 因此存在 $f \in E^*$ 与 $\alpha \in \mathbb{R}$ 使得

$$\langle f, x_1\rangle < \alpha < \langle f, x_2\rangle.$$

取

$$O_1 = \{x \in E : \langle f, x\rangle < \alpha\} = \varphi_f^{-1}((-\infty, \alpha)),$$
$$O_2 = \{x \in E : \langle f, x\rangle > \alpha\} = \varphi_f^{-1}((\alpha, +\infty)).$$

可见 O_1, O_2 是 $\sigma(E,E^*)$ 下的开集, 且它们满足要求的性质. □

命题 3.4 (弱拓扑的邻域基)

设 $x_0 \in E$, 取定 $\varepsilon > 0$ 与 E^* 中的有限集 $\{f_1, f_2, \cdots, f_k\}$, 记

$$V = V(f_1, f_2, \cdots, f_k; \varepsilon) = \{x \in E : |\langle f_i, x - x_0\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\}.$$

则 V 是 x_0 在弱拓扑 $\sigma(E,E^*)$ 下的邻域, 且跑遍 $\varepsilon > 0, k \in \mathbb{N}, f_i \in E^*$ 后得到的集族正是 x_0 在 $\sigma(E,E^*)$ 下的一族邻域基. ♠

证明 把 V 写成

$$V = \bigcap_{i=1}^{k} \varphi_{f_i}^{-1}((a_i - \varepsilon, a_i + \varepsilon)), \quad a_i = \langle f_i, x_0\rangle.$$

由 φ_{f_i} 的连续性可知 V 是弱拓扑 $\sigma(E,E^*)$ 下的开集, 且 $x_0 \in V$, 这说明 V 是 x_0 在 $\sigma(E,E^*)$ 下的邻域. 反之, 设 U 是 x_0 在 $\sigma(E,E^*)$ 下的一个邻域, 根据 $\sigma(E,E^*)$ 的构造知存在开集 $W \ni x_0$ 满足 $W \subset U$, 且 W 形如 $\varphi_{f_i}^{-1}(\omega_i)$ 的有限交, 其中 ω_i 是 $a_i = \langle f_i, x_0\rangle$ 在 $\mathbb{R}$ 中的邻域. 由 $\mathbb{R}$ 中的开集构造可知总存在 $\varepsilon > 0$ 使得 $(a_i - \varepsilon, a_i + \varepsilon) \subset \omega_i$ 对 $\forall i$ 成立. 故 $x_0 \in V \subset W \subset U$. □

如果 E 中的序列 $\{x_n\}$ 在弱拓扑 $\sigma(E,E^*)$ 下收敛到 x, 就将这一收敛记为

$$x_n \rightharpoonup x.$$

为避免歧义, 有时将这一收敛称为 $x_n \rightharpoonup x$ 在 $\sigma(E,E^*)$ 中是弱收敛. 另称 $x_n \to x$ 为强收敛, 意为 $\|x_n - x\| \to 0$.

命题 3.5 (弱收敛的性质)

设 $\{x_n\}$ 是 E 中的序列, 则

(i) $x_n \rightharpoonup x \Leftrightarrow \langle f, x_n\rangle \to \langle f, x\rangle, \forall f \in E^*$.

(ii) 若 $x_n \to x$ 是强收敛, 则 $x_n \rightharpoonup x$.

(iii) 若在 $\sigma(E,E^*)$ 中 $x_n \rightharpoonup x$, 则 $\{\|x_n\|\}$ 有界, 且 $\|x\| \le \liminf \|x_n\|$.

(iv) 若在 $\sigma(E,E^*)$ 中 $x_n \rightharpoonup x$, 且在 E^* 中 $f_n \to f$(即 $\|f_n - f\|_{E^*} \to 0$), 则 $\langle f_n, x_n\rangle \to \langle f, x\rangle$.

证明 (i) 由最弱拓扑的收敛性3.1与弱拓扑 $\sigma(E,E^*)$ 的定义即得欲证.

(ii) 因为 $|\langle f, x_n\rangle - \langle f, x\rangle| \le \|f\|\|x_n - x\|$, 由 (i) 即得欲证. 这条性质说明弱拓扑比强拓扑 (即范数拓扑) 弱.

(iii) 由 (i) 知对每个 $f \in E^*$ 而言, 序列 $\{\langle f, x_n\rangle\}_n$ 都是有界的, 根据一致有界原理的推论2.6可知 $\{x_n\}$ 本身是 E 中的有界集, 亦即 $\{\|x_n\|\}$ 是 $\mathbb{R}$ 中的有界集. 在不等式

$$|\langle f, x_n\rangle| \le \|f\|\|x_n\|$$

两端令 $n \to \infty$, 可得

$$|\langle f, x\rangle| \le \|f\| \varliminf_{n\to\infty} \|x_n\|,$$

再由 Hahn-Banach 定理可得

$$\|x\| = \sup_{\|f\|\le 1} |\langle f, x\rangle| \le \varliminf_{n\to\infty} \|x_n\|.$$

(iv) 考虑不等式

$$|\langle f_n, x_n\rangle - \langle f, x\rangle| \le |\langle f_n - f, x_n\rangle| + |\langle f, x_n - x\rangle| \le \|f_n - f\|\|x_n\| + |\langle f, x_n - x\rangle|,$$

结合 (i),(iii) 即得欲证.

命题 3.6 (有限维空间中弱拓扑与通常拓扑的等价性)

若 E 是有限维空间, 则弱拓扑 $\sigma(E,E^*)$ 与通常拓扑是同一拓扑. 特别地, 序列 $\{x_n\}$ 弱收敛当且仅当其强收敛.

证明 因为弱拓扑总比通常拓扑包含更少开集, 故只需说明每个通常拓扑下的开集在弱拓扑下也是开集即可. 设 $x_0 \in E$, U 是 x_0 在通常拓扑下的邻域, 往证存在 x_0 在弱拓扑 $\sigma(E,E^*)$ 下的邻域 V 满足 $V \subset U$, 从邻域基的角度出发知只需说明存在 $f_1, \cdots, f_k \in E^*$ 与 $\varepsilon > 0$ 满足

$$V = \{x \in E : |\langle f_i, x - x_0\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\} \subset U$$

即可. 现取 $r > 0$ 满足 $B(x_0, r) \subset U$, 取 E 中的一组标准基 $e_1, \cdots, e_k$, 知每个 $x \in E$ 都能表成 $x = \sum_{i=1}^k x_i e_i$, 且投影映射 $x \mapsto x_i$ 显然是 E 上的连续线性泛函, 记之为 f_i, 则有

$$\|x - x_0\| \le \sum_{i=1}^k |\langle f_i, x - x_0\rangle| < k\varepsilon, \quad \forall x \in V.$$

令 $\varepsilon = r/k$ 即得 $V \subset U$. $\square$

注 根据弱拓扑的最弱性可知弱拓扑 $\sigma(E,E^*)$ 中的开集 (或闭集) 总是强拓扑中的开集 (或闭集). 命题3.6表明有限维空间中这两个拓扑等价, 但在无限维空间中弱拓扑总是严格粗糙于强拓扑的. 也就是说, 总存在集合是强拓扑中的开集 (或闭集), 但不是弱拓扑中的开集 (或闭集). 下面给出两个例子:

例 3.1 设 E 是无穷维空间, $S = \{x \in E : \|x\| = 1\}$ 是其中的单位球面. 下面我们说明 S 不可能在弱拓扑 $\sigma(E,E^*)$ 中闭, 进一步还有

$$\overline{S}^{\sigma(E,E^*)} = B_E. \tag{3.1}$$

其中 $\overline{S}^{\sigma(E,E^*)}$ 表示 S 在弱拓扑 $\sigma(E,E^*)$ 下的闭包, B_E 表示 E 中的闭单位球 $\{x \in E : \|x\| \le 1\}$.

首先, 我们说明只要 $x_0 \in E$ 满足 $\|x_0\| < 1$, 就有 $x_0 \in \overline{S}^{\sigma(E,E^*)}$. 设 V 是 x_0 在 $\sigma(E,E^*)$ 中的一个邻域, 往证

$V \cap S \neq \varnothing$. 从邻域基的角度出发可设

$$V = \{x \in E : |\langle f_i, x - x_0\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\}, \quad \varepsilon > 0, f_1, \cdots, f_k \in E^*.$$

取 $y_0 \in E$ 满足 $y_0 \neq 0$, 且

$$\langle f_i, y_0\rangle = 0, \quad \forall i = 1, 2, \cdots, k.$$

这样的 y_0 是能取到的, 否则映射 $\varphi : E \to \mathbb{R}^k, x \mapsto (\langle f_i, x\rangle)_{1\le i\le k}$ 就是单射了, 此时 φ 成为 E 到 $\varphi(E)$ 的同构, 从而 $\dim E \le k$, 这与 E 的无限维假设矛盾[2]. 现取函数 $g(t) = \|x_0 + ty_0\|$, 显见 g 在 $[0, \infty)$ 上连续, $g(0) < 1, g(t) \to +\infty(t \to \infty)$, 故存在 $t_0 > 0$ 使得 $\|x_0 + t_0y_0\| = 1$, 也就是说 $x_0 + t_0y_0 \in V \cap S$, 因此

$$S \subset B_E \subset \overline{S}^{\sigma(E,E^*)}.$$

为完成(3.1)式的证明, 只需再说明 B_E 是 $\sigma(E, E^*)$ 下的闭集即可, 但由 Hahn-Banach 定理知

$$B_E = \bigcap_{\substack{f\in E^*\\ \|f\|\le 1}} \{x \in E : |\langle f, x\rangle| \le 1\},$$

上右式是弱拓扑中闭集的交集 (见定理3.1), 故 B_E 在 $\sigma(E, E^*)$ 中闭. □

例 3.2 无限维空间 E 中的单位球 $U = \{x \in E : \|x\| < 1\}$ 不可能是弱拓扑 $\sigma(E, E^*)$ 中的开集. 用反证法, 如果 U 是弱拓扑下的开集, 则其补集 $U^c = \{x \in E : \|x\| \ge 1\}$ 是弱拓扑下的闭集, 于是 $S = B_E \cap U^c$ 也应该是弱拓扑下的闭集, 这与前例矛盾了. □

注 无限维空间中的弱拓扑不可度量化. 也就是说, 在 E 上不存在能诱导出弱拓扑的度量 (也不存在这样的范数). 但是我们可以说明, 如果 E^* 是可分空间, 就可以在 E 上定义一个范数, 使其在 E 的有界子集上能诱导出弱拓扑 $\sigma(E, E^*)$.

注 一般来说, 在无限维空间中, 存在弱收敛但不强收敛的序列. 例如, 若 E^* 可分或 E 自反, 就能构造序列 $\{x_n\} \subset E$ 使得 $\|x_n\| = 1, x_n \rightharpoonup 0$. 但同样存在每个弱收敛序列都强收敛的无限维空间, 如 l^1. 这样的空间通常都很罕见, 性质也比较病态. 这一事实并不与前面提到的弱拓扑与强拓扑不同相冲突 (回忆两个具有相同收敛列的**度量**空间有相同拓扑, 但两个具有相同收敛列的**拓扑**空间未必有相同拓扑).

3.3 弱拓扑, 凸集与线性算子

前面说明了弱拓扑下的闭集都是强拓扑下的闭集, 但在无限维空间中反之不然. 下面说明对于凸集而言, 在弱拓扑下闭就等价于在强拓扑下闭.

定理 3.1 (凸集的闭性等价)

若 $C \subset E$ 是凸集, 则 C 在弱拓扑 $\sigma(E, E^*)$ 下闭当且仅当它在强拓扑下闭. ♡

证明 只需设 C 是强拓扑中的闭集, 往证 C 是弱拓扑中的闭集即可. 考虑说明 C 的补集 C^c 是弱拓扑中的开集, 为此设 $x_0 \notin C$. 因为 C 是凸集, 根据 Hahn-Banach 定理知存在闭超平面严格分离 $\{x_0\}$ 与 C, 因此存在 $f \in E^*$ 与 $\alpha \in \mathbb{R}$ 使得

$$\langle f, x_0\rangle < \alpha < \langle f, y\rangle, \quad \forall y \in C.$$

取

$$V = \{x \in E : \langle f, x\rangle < \alpha\},$$

知 $x_0 \in V, V \cap C = \varnothing$(即 $V \subset C^c$) 且 V 是 $\sigma(E, E^*)$ 中的开集, 故 C^c 是 $\sigma(E, E^*)$ 中的开集. □

注 这一证明说明闭凸集等于所有包含该集合的闭半空间之交.

[2] 从几何角度看, 当 E 是无限维空间的时候, x_0 在拓扑 $\sigma(E, E^*)$ 下的每个邻域都至少包含一条过 x_0 延伸到无穷远的直线.

推论 3.1 (Mazur)

设 $\{x_n\}$ 弱收敛到 x, 则存在由 x_n 的凸线性组合构成的序列 $\{y_n\}$ 强收敛到 x.

证明 记 $C = \operatorname{conv}\left(\bigcup_{p=1}^{\infty}\{x_p\}\right)$ 是 $\{x_n\}$ 的凸包. 因为 x 从属于 $\bigcup_{p=1}^{\infty}\{x_p\}$ 在弱拓扑下的闭包, 故它显然从属于 C 在弱拓扑下的闭包. 根据定理3.1知 $x \in \overline{C}$, 此即欲证. □

推论 3.2

设 $\varphi : E \to (-\infty, +\infty]$ 是强拓扑下的下半连续凸函数, 则它在弱拓扑 $\sigma(E, E^*)$ 中下半连续.

证明 任取 $\lambda \in \mathbb{R}$, 由 φ 的性质知知集合

$$A = \{x \in E : \varphi(x) \leq \lambda\}$$

是凸集且在强拓扑下闭. 根据定理3.1知 A 同样是弱拓扑下的闭集, 因此 φ 在弱拓扑下下半连续. □

注 一般来说很难直接证明函数在弱拓扑下下半连续, 推论3.2给出了一条路径:

$$\varphi\text{作为凸函数在强拓扑下连续} \Rightarrow \varphi\text{在弱拓扑下下半连续}.$$

例如, 函数 $\varphi(x) = \|x\|$ 作为凸函数是强拓扑下的连续函数, 因此它在弱拓扑下下半连续, 由此可知只要 $x_n \rightharpoonup x$, 则 $\|x\| \leq \liminf \|x_n\|$.

定理 3.2

设 E, F 是 Banach 空间, $T : E \to F$ 是线性算子, 则 T 在强拓扑下连续当且仅当其在弱拓扑下连续.

证明 设 T 在强拓扑下连续. 由映射在最弱拓扑下的连续性3.2知只需证明对 $\forall f \in F^*$ 而言, 映射 $x \mapsto \langle f, Tx\rangle$ 都是在弱拓扑 $\sigma(E, E^*)$ 下 $E \to \mathbb{R}$ 的连续映射, 但显见映射 $x \mapsto \langle f, Tx\rangle$ 是 E 上的连续线性泛函, 故 T 在弱拓扑 $\sigma(E, E^*)$ 下连续.

反之, 设 T 是 $(E, \sigma(E, E^*)) \to (F, \sigma(F, F^*))$ 的连续算子, 则 $G(T)$ 是 $(E \times F, \sigma(E, E^*) \times \sigma(F, F^*))$ 中的闭集. 因为 $\sigma(E, E^*) \times \sigma(F, F^*) = \sigma(E \times F, (E \times F)^*)$, 故 $G(T)$ 作为弱拓扑中的闭集必定是强拓扑中的闭集, 由闭图像定理即知 $T : E \to F$ 在强拓扑下连续. □

注 上面的断言可以进一步细化: 如果在 E 上赋强拓扑, F 上赋弱拓扑, 且 $T : E \to F$ 在这种拓扑下连续, 则 T 在 E, F 同时赋强拓扑时也连续. 因此对线性算子而言, 强 $\to$ 弱, 弱 $\to$ 弱, 强 $\to$ 弱时的连续性质是一样的. 但另一方面, 只有很少的线性算子能保证弱 $\to$ 强时依旧连续, 弱 $\to$ 强时线性算子保持连续性当且仅当 T 强 $\to$ 强连续, 且 $\dim R(T) < \infty$.

另外, 一般情况下, 在 E, F 同时赋强拓扑时连续的非线性映射在它们同时赋弱拓扑下并不连续, 这是非线性问题与线性问题主要的不同点之一.

3.4 弱 * 拓扑 $\sigma(E^*, E)$

目前为止, 我们已经研究了 E^* 上的两种拓扑:

(i) E^* 上范数诱导的通常拓扑 (强拓扑),

(ii) 依照弱拓扑构造方式生成的弱拓扑 $\sigma(E^*, E^{**})$.

下面我们定义 E^* 上称为**弱 * 拓扑**的第三种拓扑, 记之为 $\sigma(E^*, E)$(这里的 * 意在提示这一拓扑只定义在对偶空间上). 对每个 $x \in E$, 考察线性泛函 $\varphi_x : E^* \to \mathbb{R}, f \mapsto \varphi_x(f) = \langle f, x\rangle$. 随着 x 跑遍 E, 我们可以得到 $E^* \to \mathbb{R}$ 的泛函族 $\{\varphi_x\}_{x\in E}$.

定义 3.2 (弱 * 拓扑 (weak* topology))

弱 * 拓扑 $\sigma(E^*, E)$ 定义为 E^* 上由泛函族 $\{\varphi_x\}_{x\in E}$ 生成的最弱拓扑.

因为 $E \subset E^{**}$, 故拓扑 $\sigma(E^*, E)$ 显然要弱于拓扑 $\sigma(E^*, E^{**})$, 即 $\sigma(E, E^*)$ 包含的开集 (或闭集) 比 $\sigma(E^*, E^{**})$ 更少, 后者包含的开集 (或闭集) 又比强拓扑更少.

注 读者可能会疑惑为什么我们这么执着于弱拓扑. 这是因为**一个拓扑越粗糙, 其中的紧集就越多**. 例如 E^* 中的闭单位球 B_{E^*} 不可能是强拓扑下的紧集 (除非 E^* 是有限维的, 这将在后面进行证明), 但它总是弱 * 拓扑下的紧集. 既然紧集在诸多学科中 (例如力学的最小作用量存在性) 都是基本概念, 我们也就不难理解弱 * 拓扑的重要性了.

命题 3.7

弱 * 拓扑是 Hausdorff 拓扑.

证明 取 $f_1, f_2 \in E^*$ 满足 $f_1 \neq f_2$, 知存在 $x \in E$ 使得 $\langle f_1, x\rangle \neq \langle f_2, x\rangle$, 不妨设 $\langle f_1, x\rangle < \langle f_2, x\rangle$, 知存在 α 使得

$$\langle f_1, x\rangle < \alpha < \langle f_2, x\rangle.$$

记

$$O_1 = \{f \in E^* : \langle f, x\rangle < \alpha\} = \varphi_x^{-1}((-\infty, \alpha)),$$
$$O_2 = \{f \in E^* : \langle f, x\rangle > \alpha\} = \varphi_x^{-1}((\alpha, +\infty)).$$

由 φ_x 的连续性知 O_1, O_2 是 $\sigma(E^*, E)$ 中的开集, 显见 $f_1 \in O_1, f_2 \in O_2$ 且 $O_1 \cap O_2 = \varnothing$. □

命题 3.8 (弱 * 拓扑的邻域基)

设 $f_0 \in E^*$, 对给定的有限集 $\{x_1, x_2, \cdots, x_k\} \subset E$ 与 $\varepsilon > 0$, 记

$$V = V(x_1, x_2, \cdots, x_k; \varepsilon) = \{f \in E^* : |\langle f - f_0, x_i\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\}.$$

则 V 是 f 在拓扑 $\sigma(E^*, E)$ 下的一个邻域. 当 ε 跑遍全体正数, k 跑遍全体正整数, x_i 跑遍 E 时, V 形成的集族构成 f_0 在 $\sigma(E^*, E)$ 中的邻域基.

证明 这与弱拓扑的邻域基3.4证明是一样的. □

如果 E^* 中的序列 $\{f_n\}$ 在弱 * 拓扑下收敛到 f, 就记之为

$$f_n \overset{*}{\rightharpoonup} f.$$

为免歧义, 有时我们会强调 $f_n \overset{*}{\rightharpoonup} f$ 在 $\sigma(E^*, E)$ 中, $f_n \rightharpoonup f$ 在 $\sigma(E^*, E^{**})$ 中, $f_n \to f$ 在强拓扑中.

命题 3.9 (弱 * 收敛的性质)

设 $\{f_n\}$ 是 E^* 中的序列, 则

(i) $f_n \overset{*}{\rightharpoonup} f \Leftrightarrow \forall x \in E(\langle f_n, x\rangle \to \langle f, x\rangle)$.

(ii) 若 $f_n \to f$, 则 $f_n \rightharpoonup f$. 若 $f_n \rightharpoonup f$, 则 $f_n \overset{*}{\rightharpoonup} f$.

(iii) 若 $f_n \overset{*}{\rightharpoonup} f$, 则 $\{\|f_n\|\}$ 有界, 且 $\|f\| \leq \liminf \|f_n\|$.

(iv) 若 $f_n \overset{*}{\rightharpoonup} f, x_n \to x$, 则 $\langle f_n, x_n\rangle \to \langle f, x\rangle$.

证明 这与弱收敛的性质3.5证明是一样的. □

注 如果在 $f_n \overset{*}{\rightharpoonup} f$(乃至 $f_n \rightharpoonup f$) 的同时只知道 $x_n \rightharpoonup x$, 一般来说是没法推知 $\langle f_n, x_n\rangle \to \langle f, x\rangle$ 的. 例如在 l^2 中设 $f_n = x_n = T_n(\{\frac{1}{n^2}\}_n)$(其中 T_n 是右推移算子), 显见 $f_n \rightharpoonup 0, x_n \rightharpoonup 0$, 但 $\langle f_n, x_n\rangle$ 恒为某常值.

注 当 E 是有限维空间时, 前面提到的在 E^* 上的三种拓扑 (强拓扑, 弱拓扑, 弱 * 拓扑) 均重合. 这是因为由 $\dim E = \dim E^{**}$ 知典则映射 $J : E \to E^{**}$ 是满射, 因此 $\sigma(E^*, E) = \sigma(E^*, E^{**})$.

命题 3.10

若 $\varphi : E^* \to \mathbb{R}$ 是在弱 * 拓扑下连续的线性泛函, 则存在 $x_0 \in E$ 使得

$$\langle \varphi, f\rangle = \langle f, x_0\rangle, \quad \forall f \in E^*.$$

命题3.10的证明需要下述代数引理:

引理 3.2

设 X 是线性空间, $\varphi, \varphi_1, \cdots, \varphi_k$ 是 X 上的 $k+1$ 个线性泛函, 它们满足

$$(\langle \varphi_i, v\rangle = 0, \forall i = 1, \cdots, k) \Rightarrow (\langle \varphi, v\rangle = 0) \tag{3.2}$$

则存在常数 $\lambda_1, \cdots, \lambda_k \in \mathbb{R}$ 使得 $\varphi = \sum_{i=1}^k \lambda_i \varphi_i$.

下面证明引理3.2.

证明 考察依下式定义的映射:

$$F : X \to \mathbb{R}^{k+1},\ u \mapsto (\langle \varphi, u\rangle, \langle \varphi_1, u\rangle, \cdots, \langle \varphi_k, u\rangle).$$

由(3.2)式知 $a = (1, 0, \cdots, 0) \notin R(F)$, 又由 φ 的连续性与线性性知 $R(F)$ 是闭凸集, 因此在 $\mathbb{R}^{k+1}$ 中存在严格分离 $\{a\}$ 与 $R(F)$ 的超平面. 也就是说, 存在常数 $\lambda, \lambda_1, \cdots, \lambda_k$ 与 α 满足

$$\lambda < \alpha < \lambda\langle \varphi, u\rangle + \sum_{i=1}^k \lambda_i \langle \varphi_i, u\rangle, \quad \forall u \in X.$$

如果存在 $u \in X$ 使得 $\lambda\langle \varphi, u\rangle + \sum\limits_{i=1}^k \lambda_i \langle \varphi_i, u\rangle \neq 0$, 就总能通过选取某个 v 使得 $\lambda\langle \varphi, v\rangle + \sum\limits_{i=1}^k \lambda_i \langle \varphi_i, v\rangle$ 任意小, 因此只能有

$$\lambda\langle \varphi, u\rangle + \sum_{i=1}^k \lambda_i \langle \varphi_i, u\rangle = 0, \quad \forall u \in X.$$

由此可知 $\lambda < 0$, 因此 $\lambda \neq 0$, 此即欲证. □

下面证明命题3.10.

证明 因为 φ 在弱 * 拓扑下连续, 故在 $\sigma(E^*, E)$ 中存在 0 的邻域 V 使得

$$|\langle \varphi, f\rangle| < 1, \quad \forall f \in V.$$

从邻域基的角度出发可设

$$V = \{f \in E^* : |\langle f, x_i\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\},$$

其中 $x_i \in E, \varepsilon > 0$. 下面说明

$$(\langle f, x_i\rangle = 0, \forall i = 1, 2, \cdots, k) \Rightarrow (\langle \varphi, f\rangle = 0). \tag{3.3}$$

用反证法, 若 $\langle \varphi, f\rangle \neq 0$ 时 $\langle f, x_i\rangle = 0$ 对 $\forall i = 1, 2, \cdots, k$ 成立, 则对任意 $M \in \mathbb{R}$ 而言, Mf 总会在 V 中, 通过调整 M 的大小可使得 $|\langle \varphi, Mf\rangle| \geq 1$, 矛盾! 故(3.3)式成立. 现在对 $\varphi, x_1, \cdots, x_k$ 应用引理3.2可得

$$\langle \varphi, f\rangle = \sum_{i=1}^k \lambda_i \langle f, x_i\rangle = \left\langle f, \sum_{i=1}^k \lambda_i x_i \right\rangle, \quad \forall f \in E^*.$$

□

推论 3.3

设 E^* 中的超平面 H 是 $\sigma(E^*, E)$ 中的闭集, 则 H 具有形式:

$$H = \{f \in E^* : \langle f, x_0\rangle = \alpha\},$$

其中 $x_0 \in E, x_0 \neq 0, \alpha \in \mathbb{R}$.

证明 由 H 是 E^* 中的超平面知可设

$$H = \{f \in E^* : \langle \varphi, f\rangle = \alpha\},$$

其中 φ 是 E^* 上的线性泛函, $\varphi \neq 0$. 下面说明 φ 在 $\sigma(E^*, E)$ 下的连续性, 由线性性知只需说明 φ 在 0 处连续. 取 $f_0 \notin H$, 设 V 是 f_0 在 $\sigma(E^*, E)$ 下的一个邻域, 且 $V \subset H^c$(这样的 V 能取到正是因为 H^c 是开集), 从邻域基的角度出发可设

$$V = \{f \in E^* : |\langle f - f_0, x_i\rangle| < \varepsilon, \forall i = 1, 2, \cdots, k\}.$$

显见 V 是凸集, 故要么

$$\langle\varphi, f\rangle < \alpha, \quad \forall f \in V \tag{3.4}$$

要么

$$\langle\varphi, f\rangle > \alpha, \quad \forall f \in V. \tag{3.5}$$

不妨设(3.4)式成立, 可得

$$\langle\varphi, g\rangle < \alpha - \langle\varphi, f_0\rangle, \quad \forall g \in W = V - f_0.$$

因为 $-W = W$, 故

$$|\langle\varphi, g\rangle| < |\alpha - \langle\varphi, f_0\rangle|, \quad \forall g \in W. \tag{3.6}$$

因为 $W \subset E^*$ 是 0 的邻域, 故(3.6)式表明在 $\sigma(E^*, E)$ 下 φ 在 0 处连续. 由命题3.10即知存在 $x_0 \in E$ 使得

$$\langle\varphi, f\rangle = \langle f, x_0\rangle, \quad \forall f \in E^*.$$

□

注 第一次推到这里的时候可能会采用下述方法: 因为 $\sigma(E^*, E)$ 弱于通常拓扑, 故 H 作为 $\sigma(E^*, E)$ 中的闭集一定是通常拓扑下的闭集, 利用通常拓扑下的闭超平面形式即得 $H = \{f \in E^* : \langle f, x_0\rangle = \alpha\}$. 但这么做实际上是有问题的: 典则映射 $J: E \to E^{**}$ 未必是满射, 因此这里出现的 x_0 未必落在 E 内. 实际上, 在典则映射不满的情况下, $\sigma(E^*, E)$ 要严格弱于 $\sigma(E^*, E^{**})$. 例如, 取 $\xi \in E^{**} \backslash J(E)$, 显见

$$H = \{f \in E^* : \langle\xi, f\rangle = 0\}$$

是 $\sigma(E^*, E^{**})$ 中的闭集, 但由推论3.3知 H 并不是 $\sigma(E^*, E)$ 中的闭集. 通过这个例子我们也可以发现在强拓扑下闭的凸集未必在弱 * 拓扑下闭. 需记住在 E^* 中有两类闭凸集:

(a) 在强拓扑下闭的凸集 (由凸集的闭性等价3.1知这也就是在 $\sigma(E^*, E^{**})$ 下闭的凸集),

(b) 在 $\sigma(E^*, E)$ 下闭的凸集.

定理 3.3 (Banach-Alaoglu-Bourbaki)

对于 Banach 空间 E 而言, 闭单位球

$$B_{E^*} = \{f \in E^* : \|f\| \le 1\}$$

*在弱 * 拓扑 $\sigma(E^*, E)$ 中紧.*

♡

注 B_{E^*} 的紧性是弱 * 拓扑的最本质性质. 这是因为有结果表明: E^* 的闭单位球在强拓扑下紧当且仅当 E 有限维. 因此只要 E^* 是无限维空间, B_{E^*} 就不会是强拓扑下的紧集, 但它总能成为弱 * 拓扑下的紧集.

证明 证明的主要思路在于把 E^* 拉入乘积空间 $\mathbb{R}^E$, 再在乘积空间中用 Tychonoff 定理1.22证明命题. 现考虑由全体 $E \to \mathbb{R}$ 的映射构成的直积 $Y = \mathbb{R}^E$, 记 Y 的元素为 $\omega = \{\omega_x\}_{x\in E}, \omega_x \in \mathbb{R}$. 可以说明 Y 上能赋乘积拓扑, 即 Y 上全体映射 $\omega \mapsto \omega_x$(在 x 跑遍 E 时) 诱导出的最弱拓扑, 而这又是点态收敛对应的拓扑 (见推论1.8).

下面默认 E^* 上赋弱 * 拓扑 $\sigma(E^*, E)$. 因为 E^* 本质上是由 $E \to \mathbb{R}$ 的一些特殊的映射 (即连续线性映射) 组成的, 故可以将 E^* 看成 Y 的一个子集. 准确来说, 设 $\Phi: E^* \to Y$ 是 E^* 到 Y 的典则映射, 可知

$$\Phi(f) = \{\omega_x\}_{x\in E}, \quad \text{其中}\omega_x = \langle f, x\rangle,$$

往证 $\Phi: E^* \to \Phi(E^*)$ 是同胚. 因为对每个固定的 $x \in E$ 而言, 映射 $E^* \ni f \mapsto (\Phi(f))_x = \langle f, x\rangle$ 作为作用在 f 上的泛函依照弱 * 拓扑的构造连续, 根据该连续性可知对 $\forall x \in E$ 而言 $\varphi_x \circ \Phi: E^* \to \mathbb{R}$ 均连续, 故由映射在最弱拓扑下的连续性3.2知 $\Phi: E^* \to Y$ 连续. 另外, 逆映射 $\Phi^{-1}: \Phi(E^*) \to E^*$ 在 $\Phi(E^*)$ 上赋 Y 的子空间拓扑时也连续, 这是因为再次应用映射在最弱拓扑下的连续性3.2知只需说明对每个取定的 $x \in E$, 映射 $Y \ni \omega \mapsto \langle\Phi^{-1}(\omega), x\rangle \in \mathbb{R}$ 在 $\Phi(E^*)$ 上关于 Y 的子空间拓扑连续, 而由 $\omega \in \Phi(E^*)$ 知必存在 $f \in E^*$ 使得 $\omega = \Phi(f)$, 于是

$$\langle\Phi^{-1}(\omega), x\rangle = \langle f, x\rangle = \omega_x$$

这正是乘积空间拓扑所要保证连续的正规投影 (见定理1.19), 因此映射 $\omega \mapsto \omega_x$ 必连续. 至此, 我们便证明了 Φ 是

E^* 到 $\Phi(E^*)$ 上的同胚.

现在把闭单位球 B_{E^*} 同胚到乘积空间 Y 中, 设 $\Phi(B_{E^*}) = K$, 其中

$$K = \{\omega \in Y : |\omega_x| \le \|x\|, \omega_{x+y} = \omega_x + \omega_y, \omega_{\lambda x} = \lambda\omega_x, \forall\lambda \in \mathbb{R}, \forall x, y \in E\}.$$

因为同胚保持紧性, 故只需证明 K 在 Y 的乘积空间拓扑下成为紧集即可. 把 K 写成 $K = K_1 \cap K_2$, 其中

$$K_1 = \{\omega \in Y : |\omega_x| \le \|x\|, \forall x \in E\},$$

$$K_2 = \{\omega \in Y : \omega_{x+y} = \omega_x + \omega_y, \omega_{\lambda x} = \lambda\omega_x, \forall\lambda \in \mathbb{R}, \forall x, y \in E\}.$$

根据乘积拓扑的构造, K_1 可以进一步写成一族紧区间的乘积:

$$K_1 = \prod_{x\in E}[-\|x\|, +\|x\|].$$

因为每个 $[-\|x\|, +\|x\|]$ 都是 $\mathbb{R}$ 中的紧集, 故由 Tychonoff 定理1.22知 K_1 是 Y 中的紧集. 下面再说明 K_2 是 Y 中的闭集. 对任意取定的 $\lambda \in \mathbb{R}, x, y \in E$, 因为映射 $\omega \mapsto \omega_{x+y} - \omega_x - \omega_y, \omega \mapsto \omega_{\lambda,x} - \lambda\omega_x$ 依照乘积拓扑的构造必定连续, 故集合

$$A_{x,y} = \{\omega \in Y : \omega_{x+y} - \omega_x - \omega_y = 0\}$$

$$B_{\lambda,x} = \{\omega \in Y : \omega_{\lambda x} - \lambda\omega_x = 0\}$$

是 Y 中的闭集, 因此

$$K_2 = \left[\bigcap_{x,y\in E} A_{x,y}\right] \cap \left[\bigcap_{\substack{x\in E\\ \lambda\in\mathbb{R}}} B_{\lambda,x}\right]$$

作为闭集的任意交是 Y 中的闭集. 最后, 因为 $K_1 \cap K_2$ 可以视作紧空间 K_1 的闭子空间, 故 $K = K_1 \cap K_2$ 是 Y 中的紧集, 此即欲证. □

3.5 自反空间

定义 3.3 (自反空间 (reflexive space))

设 E 是 Banach 空间, $J : E \to E^{**}$ 是 E 到 E^{**} 中依(2.6)式定义的典则映射. 若 J 是满射 (即 $J(E) = E^{**}$), 就称 E 是自反的. ♣

当 E 是自反空间时, 一般把 E^{**} 就看成 E, 此时 J 是 $E \to E^{**}$ 的保范同构,

注 分析中的很多重要空间都是自反空间. 显见有限维空间是自反空间 (因为 $\dim E = \dim E^* = \dim E^{**}$)[3]. 另外在 $1 < p < \infty$ 时, L^p(与 l^p) 也是自反空间, 同时 Hilbert 空间都是自反空间. 然而, 也有一些重要空间并不自反, 例如:

- L^1, L^∞(与 l^1, l^∞) 不是自反空间.
- 无限维紧度量空间 K 上的连续函数构成的空间 $C(K)$ 不是自反空间.

注 自反空间的定义必须通过典则映射 J. 事实上, 存在非自反空间 E 满足 E 与 E^{**} 等距同构.

下述结果给出了自反空间的一条基本性质:

定理 3.4 (Kakutani(角谷静夫))

设 E 是 Banach 空间, 则 E 自反当且仅当

$$B_E = \{x \in E : \|x\| \le 1\}$$

在弱拓扑 $\sigma(E, E^*)$ 下紧. ♡

证明 设 E 自反, 知 $J(B_E) = B_{E^{**}}$. 由 Banach-Alaoglu-Bourbaki 定理3.3知 $B_{E^{**}}$ 是弱 * 拓扑 $\sigma(E^{**}, E^*)$ 下的紧集, 因此只要能说明 J^{-1} 是从赋 $\sigma(E^{**}, E^*)$ 的 E^{**} 到赋 $\sigma(E, E^*)$ 的 E 的连续映射, 就能得到 J 是同胚, 从而利用同胚的保紧性即得 B_E 的紧性. 根据映射在最弱拓扑下的连续性3.2, 只需证明对每个取定的 $f \in E^*$ 而言, 映射

[3] 然而需要特别注意 E 与 E^* 并不等距同构, 它们之间可能只存在线性同构.

$\varphi_f \circ J^{-1}: \xi \mapsto \langle f, J^{-1}\xi\rangle$ 在赋 $\sigma(E^{**}, E^*)$ 的 E^{**} 上连续即可. 又因为 $\langle f, J^{-1}\xi\rangle = \langle \xi, f\rangle$, 而映射 $\xi \mapsto \langle \xi, f\rangle$ 依照 $\sigma(E^{**}, E^*)$ 的构造必定连续, 故 J^{-1} 连续, 因此 B_E 在 $\sigma(E, E^*)$ 中紧.

逆命题的证明更为精细, 它需要下述两条引理:

引理 3.3 (Helly)

设 E 是 Banach 空间, $f_1, \cdots, f_k \in E^*, \gamma_1, \cdots, \gamma_k \in \mathbb{R}$ 是给定的元素, 则下述两条性质等价:

(i) 对 $\forall \varepsilon > 0$ 均存在 $x_\varepsilon \in E$ 使得 $\|x_\varepsilon\| \leq 1$ 的同时有

$$|\langle f_i, x_\varepsilon\rangle - \gamma_i| < \varepsilon, \quad \forall i = 1, \cdots, k.$$

(ii) $\left|\sum_{i=1}^k \beta_i \gamma_i\right| \leq \left\|\sum_{i=1}^k \beta_i f_i\right\|, \forall \beta_1, \cdots, \beta_k \in \mathbb{R}$.

下面证明 Helly 引理3.3. 对于 (i)$\Rightarrow$(ii), 取定 $\beta_1, \cdots, \beta_k \in \mathbb{R}$, 令 $S = \sum_{i=1}^k |\beta_i|$. 由 (i) 知

$$\left|\sum_{i=1}^k \beta_i \langle f_i, x_\varepsilon\rangle - \sum_{i=1}^k \beta_i \gamma_i\right| \leq \varepsilon S,$$

因此

$$\left|\sum_{i=1}^k \beta_i \gamma_i\right| \leq \left\|\sum_{i=1}^k \beta_i f_i\right\| \|x_\varepsilon\| + \varepsilon S \leq \left\|\sum_{i=1}^k \beta_i f_i\right\| + \varepsilon S.$$

因为 $\varepsilon > 0$ 是任意的, 故 (ii) 成立.

对于 (ii)$\Rightarrow$(i), 取 $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_k) \in \mathbb{R}^k$, 考虑下述映射:

$$\varphi: E \to \mathbb{R}^k, \quad x \mapsto (\langle f_1, x\rangle, \cdots, \langle f_k, x\rangle).$$

现在性质 (i) 就是在说 $\gamma \in \overline{\varphi(B_E)}$. 现在考虑反证, 若 (i) 不成立, 则 $\gamma \notin \overline{\varphi(B_E)}$, 因此 $\{\gamma\}$ 与 $\overline{\varphi(B_E)}$ 必能在 $\mathbb{R}^k$ 中被某超平面严格分离, 亦即存在 $\beta = (\beta_1, \beta_2, \cdots, \beta_k) \in \mathbb{R}^k$ 与 $\alpha \in \mathbb{R}$ 使得

$$\beta \cdot \varphi(x) < \alpha < \beta \cdot \gamma, \quad \forall x \in B_E.$$

把上式两端写开即

$$\left\langle \sum_{i=1}^k \beta_i f_i, x \right\rangle < \alpha < \sum_{i=1}^k \beta_i \gamma_i, \quad \forall x \in B_E,$$

因此

$$\left\|\sum_{i=1}^k \beta_i f_i\right\| \leq \alpha < \sum_{i=1}^k \beta_i \gamma_i$$

这与 (ii) 矛盾! □

引理 3.4 (Goldstine)

设 E 是任意 Banach 空间, 则 $J(B_E)$ 在拓扑 $\sigma(E^{**}, E^*)$ 下是 $B_{E^{**}}$ 的稠子集, 因此 $J(E)$ 在拓扑 $\sigma(E^{**}, E^*)$ 下是 E^{**} 的稠子集.

下面证明 Goldstine 引理3.4. 取 $\xi \in B_{E^{**}}$, 设 V 是 ξ 在拓扑 $\sigma(E^{**}, E^*)$ 下的邻域, 往证 $V \cap J(B_E) \neq \varnothing$. 利用弱 * 拓扑的邻域基刻画3.8, 可设

$$V = \{\eta \in E^{**} : |\langle \eta - \xi, f_i\rangle| < \varepsilon, \forall i = 1, \cdots, k\},$$

其中 $f_1, \cdots, f_k \in E^*$ 取定, $\varepsilon > 0$. 现需找到 $x \in B_E$ 使得 $J(x) \in V$, 即

$$|\langle f_i, x\rangle - \langle \xi, f_i\rangle| < \varepsilon, \quad \forall i = 1, \cdots, k.$$

记 $\gamma_i = \langle \xi, f_i\rangle$, 根据 Helly 引理3.3, 只需验证

$$\left|\sum_{i=1}^k \beta_i \gamma_i\right| \leq \left\|\sum_{i=1}^k \beta_i f_i\right\|,$$

而

$$\left|\sum_{i=1}^{k}\beta_i\gamma_i\right| = \left|\left\langle \xi, \sum_{i=1}^{k}\beta_i f_i\right\rangle\right| \le \|\xi\|\left\|\sum_{i=1}^{k}\beta_i f_i\right\|$$

由 $\|\xi\| \le 1$ 即得欲证. □

现在着手证明逆命题. 根据映射在最弱拓扑下的连续性3.2, 因为对每个取定的 $f \in E^*$ 而言, 映射 $x \mapsto \langle Jx, f\rangle = \langle f, x\rangle$ 总关于 $\sigma(E, E^*)$ 连续, 故典则映射 $J : E \to E^{**}$ 实际上总是从 $\sigma(E, E^*)$ 到 $\sigma(E^{**}, E^*)$ 的连续映射. 因为连续映射把紧集映成紧集, 故由 B_E 在 $\sigma(E, E^*)$ 下是紧集知 $J(B_E)$ 是 $\sigma(E^{**}, E^*)$ 下的紧集. 由 $\sigma(E^{**}, E^*)$ 的 Hausdorff 性可知 $J(B_E)$ 同样是 $\sigma(E^{**}, E)$ 中的闭集, 又因为 Goldstine 引理3.4表明 $J(B_E)$ 在 $B_{E^{**}}$ 中稠, 故只能有 $J(B_E) = B_{E^{**}}$, 因此 $J(E) = E^{**}$, 此即欲证. □

关于自反空间的紧性性质, 我们还有下述两条结果:

定理 3.5

设 E 是自反 Banach 空间, $\{x_n\}$ 是 E 中的有界序列, 则 $\{x_n\}$ 存在在弱拓扑 $\sigma(E, E^*)$ 下收敛的子列 $\{x_{n_k}\}$. ♡

定理 3.6 (Eberlein-Smulian)

设 E 作为 Banach 空间满足 E 中的每个有界列都有在 $\sigma(E, E^*)$ 下收敛的子列, 则 E 自反. ♡

这两条定理的证明先省略了.

注 定理3.4, 定理3.5与定理3.6并不是同一件事的重申, 我们可以回忆:

(i) 若 X 是度量空间, 则

$$X \text{ 紧} \Leftrightarrow X \text{ 中的任意序列都有收敛子列.}$$

(ii) 存在紧拓扑空间 X 使得其中的序列不存在任何收敛子列. 这种空间的典例为 $X = B_{E^*}$, 根据 Banach-Alaoglu-Bourbaki 定理可知 B_{E^*} 是 $\sigma(E^*, E)$ 中的紧集, 但在 $E = l^\infty$ 时可以构造 X 中不存在收敛子列的序列.

(iii) 就算 X 是拓扑空间且 X 中的任意序列都有收敛子列, X 也未必紧.

下面给出自反空间的更多性质.

命题 3.11 (闭线性子空间对自反性的继承)

设 E 是自反 Banach 空间, $M \subset E$ 是 E 的闭线性子空间, 则 M 自反. ♠

证明 当在 M 上赋 E 的范数后, M 上可以自动出现两种弱拓扑:

(a) 由 $\sigma(E, E^*)$ 诱导出的拓扑.

(b) 它自身的弱拓扑 $\sigma(M, M^*)$.

事实上, 这两种拓扑是相同的, 这是因为由 Hahn-Banach 定理知 M 上的每个连续线性泛函都是 E 上连续线性泛函在 M 上的限制, 于是 $\sigma(M, M^*)$ 需要确保连续的每个线性泛函 $f \in M^*$ 均具有在 E 中的延拓 $f' \in E^*$. 现在根据 Kakutani 定理, 要证明 M 自反, 只需证明 B_M 在 $\sigma(M, M^*)$(或等价地在 $\sigma(E, E^*)$) 中紧. 然而, 因为 B_E 已经在 $\sigma(E, E^*)$ 中紧, 且由凸集的闭性等价3.1知 M 在 $\sigma(E, E^*)$ 中闭, 故 B_M 在 $\sigma(E, E^*)$ 中紧. □

推论 3.4 (对偶保持自反性)

Banach 空间 E 自反当且仅当其对偶空间 E^* 自反. ♡

证明 $\Rightarrow$: 因为 E 自反, 故 $E^{**} = E$. 设 $J : E \to E^{**}$ 是典则映射, 任取 $\varphi \in E^{***}$, 知映射 $x \mapsto \langle \varphi, Jx\rangle$ 是 E 上的连续线性泛函, 记该映射为 f, 则

$$\langle \varphi, Jx\rangle = \langle f, x\rangle, \quad \forall x \in E.$$

但 $f \in E^*$, 故 $\langle f, x\rangle = \langle Jx, f\rangle$, 于是

$$\langle \varphi, Jx\rangle = \langle Jx, f\rangle, \quad \forall x \in E.$$

因为 J 是满射, 所以当 x 跑遍 E 时 Jx 跑遍 E^{**}, 因此

$$\langle \varphi, \xi \rangle = \langle \xi, f \rangle, \quad \forall \xi \in E^{**}.$$

这说明 $E^* \to E^{***}$ 的典则映射是满射.

$\Leftarrow$: 根据前述结论, 由 E^* 自反可知 E^{**} 自反. 因为 $J(E)$ 是 E^{**} 在强拓扑下的闭子空间, 故由命题3.11知 $J(E)$ 自反. 若能说明线性等距同构保持自反性, E 自反就自然成立了.

下面说明线性等距同构保持自反. 若 E, F 是 Banach 空间, $T : E \to F$ 是线性等距同构, 则考虑映射 $T^* : E^* \to F^*, f \mapsto f \circ T^{-1}$, 可以验证 T^* 是 $E^* \to F^*$ 的线性等距同构, 类似基于 T^* 可以构造 $E^{**} \to F^{**}$ 的线性等距同构 $T^{**} : \xi \mapsto \xi \circ (T^*)^{-1}$. 现设 E 自反, 任取 $\xi \in F^{**}$, 知存在 $\tau \in E^{**}$ 使得 $T^{**}\tau = \xi$, 由 E 的自反性可知

$$\langle \xi, g \rangle = \langle T^* \tau, g \rangle = \langle \tau, (T^*)^{-1} g \rangle = \langle (T^*)^{-1} g, J_E^{-1} \tau \rangle = \langle g, T(J_E^{-1} \tau) \rangle, \quad \forall \xi \in F^{**}, \forall g \in F^*.$$

其中 $J_E : E \to E^{**}$ 是典则映射. 容易验证 $T \circ J_E^{-1} : E^{**} \to F$ 是满射, 下面说明 $T^{**} \circ (T \circ J_E^{-1})^{-1} = T^{**} \circ J_E \circ T^{-1}$ 实际上就是 $F \to F^{**}$ 的典则映射. 任取 $x \in F, f \in F^*$, 重复上述过程知

$$\langle T^{**} \circ J_E \circ T^{-1} x, f \rangle = \langle f, x \rangle.$$

这正是典则映射的定义. 现在 $T^{**} \circ J_E \circ T^{-1}$ 显然是 $F \to F^{**}$ 的满射, 故 F 自反. □

推论 3.5 (自反空间弱拓扑下的紧集判定)

设 E 是自反 Banach 空间, $K \subset E$ 是强拓扑下的有界闭凸子集, 则 K 在 $\sigma(E, E^*)$ 下紧. ♡

证明 因为 K 是凸集, 故由凸集的闭性等价3.1知 K 在强拓扑下闭等价于 K 在 $\sigma(E, E^*)$ 下闭. 另一方面, 根据 K 有界知存在常数 m 使得 $K \subset mB_E$, 而 Kakutani 定理3.4表明 mB_E 在 $\sigma(E, E^*)$ 下紧, 因此 K 作为紧集的闭子集自然在 $\sigma(E, E^*)$ 下紧. □

推论 3.6 (自反空间下半连续凸函数的最小值)

设 E 是自反 Banach 空间, $A \subset E$ 是强拓扑下的非空闭凸子集, 设 $\varphi : A \to (-\infty, +\infty]$ 是下半连续凸函数, $\varphi \not\equiv +\infty$, 且

$$\lim_{\substack{x \in A \\ \|x\| \to \infty}} \varphi(x) = +\infty \quad (A \text{ 有界时忽略该假设.}) \tag{3.7}$$

则 φ 可在 A 上达到最小值, 亦即存在 $x_0 \in A$ 使得

$$\varphi(x_0) = \min_A \varphi.$$

♡

证明 任取 $a \in A$ 满足 $\varphi(a) < +\infty$, 考察集合

$$\widetilde{A} = \{x \in A : \varphi(x) \le \varphi(a)\}.$$

根据(3.7)式知 $\widetilde{A}$ 是强拓扑下的有界闭凸集, 进而由自反空间弱拓扑下的紧集判定3.5知 A 是 $\sigma(E, E^*)$ 下的紧集. 另一方面, 推论3.2表明 φ 在 $\sigma(E, E^*)$ 下依旧下半连续, 因此根据下半连续函数的基本性质 (v) 可知 φ 在 $\widetilde{A}$ 上达到最小值, 即存在 $x_0 \in \widetilde{A}$ 使得

$$\varphi(x_0) \le \varphi(x), \quad \forall x \in \widetilde{A}.$$

当 $x \in A \ni \widetilde{A}$ 时, 知 $\varphi(x_0) \le \varphi(a) < \varphi(x)$, 故

$$\varphi(x_0) \le \varphi(x), \quad \forall x \in A.$$

□

注 推论3.6正是自反空间与凸函数在变分法与优化论的许多问题中重要地位的基石.

定理 3.7 (自反空间之间无界线性算子的延拓)

设 E, F 是自反 Banach 空间, $A : D(A) \subset E \to F$ 是无界闭线性算子, 且 $D(A)$ 在 E 中稠密. 则 $D(A^*)$ 在 F^* 中稠密, 因此 $A^{**} : D(A^{**}) \subset E^{**} \to F^{**}$ 良定义, 它也能视作 $E \to F$ 的无界算子, 于是

$$A^{**} = A.$$

注 回忆

$$D(A^*) = \{v \in F^* : \exists c > 0(|\langle v, Au\rangle| \le c\|u\|, \forall u \in D(A))\}$$

$$\langle v, Au\rangle_{F^*,F} = \langle A^*v, u\rangle_{E^*,E}, \quad \forall u \in D(A), \forall v \in D(A^*).$$

证明 首先证明 $D(A^*)$ 在 F^* 中稠密. 设 φ 是 F^* 上的连续线性泛函, 且 φ 在 $D(A^*)$ 上为零, 由推论2.4知只需说明在 F^* 上 $\varphi \equiv 0$ 即可. 因为 F 自反, 故 $\varphi \in F$, 且

$$\langle w, \varphi\rangle = 0, \quad \forall w \in D(A^*). \tag{3.8}$$

若 $\varphi \ne 0 = A(0)$, 则 $(0, \varphi) \notin G(A) \subset E \times F$, 而由 A 的闭性知 $G(A)$ 是闭集, 故由 Hahn-Banach 定理的第二几何形式知 $E \times F$ 中存在严格分离 $\{(0, \varphi)\}$ 与 $G(A)$ 的闭超平面, 亦即存在 $(f, v) \in E^* \times F^*$ 与 $\alpha \in \mathbb{R}$ 使得

$$\langle f, u\rangle + \langle v, Au\rangle < \alpha < \langle v, \varphi\rangle, \quad \forall u \in D(A).$$

因为 $u \mapsto \langle f, u\rangle + \langle v, Au\rangle$ 依旧是 $D(A)$ 上的连续线性泛函, 故只能有

$$\langle f, u\rangle + \langle v, Au\rangle = 0, \quad \forall u \in D(A). \tag{3.9}$$

且

$$\langle v, \varphi\rangle \ne 0.$$

由(3.9)式知 $v \in D(A^*)$, 在(3.8)式中取 $w = v$ 即得矛盾.

下面说明 $A^{**} = A$. 回忆(2.35)式, 知

$$I(G(A^*)) = G(A)^{\perp}, \quad I(G(A^{**})) = G(A^*)^{\perp},$$

又因为由 A 的闭性知 $G(A)$ 闭, 于是 $G(A)^{\perp\perp} = G(A)$, 因此

$$G(A^{**}) = G(A)^{\perp\perp} = G(A).$$

□

3.6 可分空间

定义 3.4 (可分空间 (seperable space))

若度量空间 E 存在可数稠密子集, 就称 E 可分.

分析中的许多重要空间都是可分空间, 例如有限维空间, $L^p(l^p)(1 \le p < \infty)$ 与紧度量空间 K 上的连续函数空间 $C(K)$. 然而, $L^\infty(l^\infty)$ 并不可分.

命题 3.12 (可分性的传递性)

设 E 是可分度量空间, $F \subset E$ 是任意子集, 则 F 可分.

证明 设 $\{u_n\}$ 是 E 的可数稠密子集, $\{r_m\}$ 是满足 $r_m \to 0$ 的任意正数列. 现在对每个 m, n 而言, 在 $B(u_n, r_m) \cap F$ 非空时选取一个点 $a_{m,n} \in B(u_n, r_m)$, 可见 $\{a_{m,n}\}$ 是 F 的可数稠密子集. □

定理 3.8 (Banach)

设 E 是 Banach 空间, 若 E^* 可分, 则 E 可分.

注 上述定理的逆命题并不正确. 考虑可分空间 $E = L^1$, 其对偶空间 $E^* = L^\infty$ 并不可分.

证明 设 $\{f_n\}_{n\geq 1}$ 是 E^* 的可数稠密子集. 因为

$$\|f_n\| = \sup_{\substack{x\in E\\ \|x\|\leq 1}} \langle f_n, x\rangle,$$

故存在 $x_n \in E$ 使得

$$\|x_n\| = 1, \quad \langle f_n, x_n\rangle \geq \frac{1}{2}\|f_n\|.$$

记 L_0 为 $\{x_n\}_{n\geq 1}$ 以 $\mathbb{Q}$ 为背景所张成的线性空间; 也就是说, L_0 包含系数在 $\mathbb{Q}$ 中的 $\{x_n\}_{n\geq 1}$ 的全体有限线性组合. 这样构造的 L_0 是可数的, 这是因为对每个整数 n, 记 Λ_n 为 $\{x_n\}_{1\leq k\leq n}$ 以 $\mathbb{Q}$ 为背景张成的空间. 显见 Λ_n 可数, 且 $L_0 = \bigcup_{n\geq 1}\Lambda_n$, 因此 L_0 可数.

现在记 L 为 $\{x_n\}_{n\geq 1}$ 以 $\mathbb{R}$ 为背景张成的线性空间. 显见 L_0 是 L 的稠密子集, 如果能说明 L 是 E 的稠密子空间, 证明就结束了 (因为这样就说明了 L_0 是 E 的可数稠密子集). 设 $f \in E^*$ 是在 L 上取值为零的连续线性泛函, 我们考虑用推论2.4证明 L 的稠密性, 这就是要说明 $f = 0$. 根据 $\{f_n\}_{n\geq 1}$ 的稠密性, 任取 $\varepsilon > 0$, 总存在整数 N 使得 $\|f - f_N\| < \varepsilon$, 因此

$$\frac{1}{2}\|f_N\| \leq \langle f_N, x_N\rangle \overset{(A)}{=} \langle f_N - f, x_N\rangle < \varepsilon$$

(其中 (A) 是因为 $\langle f, x_N\rangle = 0$). 现有 $\|f\| \leq \|f - f_N\| + \|f_N\| < 3\varepsilon$, 因此 $f = 0$. □

推论 3.7

设 E 是 Banach 空间, 则

$$E \text{ 自反且可分} \Leftrightarrow E^* \text{ 自反且可分}.$$

证明 根据对偶对自反性的保持3.4与前述 Banach 定理3.8已经有

$$E^* \text{ 自反且可分} \Rightarrow E \text{ 自反且可分},$$

$$E \text{ 自反且可分} \Rightarrow E^* \text{ 自反}.$$

现若 E 自反且可分, 根据自反性知 $E^{**} = J(E)$, 而典则嵌入保持子空间的稠密性, 故 E^{**} 可分, 由 Banach 定理3.8即知 E^* 可分. □

可分性和弱拓扑下的可度量性关系密切. 回忆拓扑空间 X 可度量指的是 X 上存在能诱导 X 上拓扑的度量.

定理 3.9 (可分与单位闭球可度量的等价性)

若 E 是可分 Banach 空间, 则 B_{E^*} 在弱 * 拓扑 $\sigma(E^*, E)$ 下可度量. 反之, 若 B_{E^*} 在 $\sigma(E^*, E)$ 下可度量, 则 E 可分.

上述命题还有下述" 对偶" 表述:

定理 3.10

若 E 是满足 E^* 可分的 Banach 空间, 则 B_E 在弱拓扑 $\sigma(E, E^*)$ 下可度量. 反之, 若 B_E 在 $\sigma(E, E^*)$ 下可度量, 则 E^* 可分.

下面证明定理3.9.

证明 设 $\{x_n\}_{n\geq 1}$ 是 B_E 的可数稠密子集. 对任意 $f \in E^*$, 记

$$[f] = \sum_{n=1}^{\infty} \frac{1}{2^n}|\langle f, x_n\rangle|.$$

容易说明 $[\cdot]$ 是 E^* 上的范数, 且 $[f] \leq \|f\|$. 现在记 $d(f,g) = [f-g]$ 为该范数对应的度量, 下面说明度量 d 在 B_{E^*} 上诱导的拓扑与 $\sigma(E^*, E)$ 在 B_{E^*} 上的限制 (为叙述方便, 依旧把该拓扑记为 $\sigma(E^*, E)$) 是相同的.

首先说明 $\sigma(E^*, E)$ 中的开集是 d 诱导拓扑中的开集. 根据定理1.3, 只需说明 B_{E^*} 中全体点在 $\sigma(E^*, E)$ 下的全体邻域都是 d 诱导拓扑中的开集即可. 现取 $f_0 \in B_{E^*}$, 记 V 是 f_0 在 $\sigma(E^*, E)$ 下的邻域. 为证 V 是 d 诱导拓扑中

的开集, 就是要找到 $r>0$ 使得

$$U=\{f\in B_{E^*}:d(f,f_0)<r\}\subset V.$$

根据弱 * 拓扑的邻域基3.8, 不妨设

$$V=\{f\in B_{E^*}:|\langle f-f_0,y_i\rangle|<\varepsilon,\forall i=1,\cdots,k\},$$

其中 $\varepsilon>0,y_1,\cdots,y_k\in E$. 根据 $f-f_0$ 的线性性与 ε 的任意性, 不失一般性可以再设 $\|y_i\|\le 1(i=1,\cdots,k)$, 于是 $\{y_i\}_{i=1}^k\subset B_E$. 因为 $\{x_n\}_{n\ge1}$ 是 B_E 的稠子集, 故对每个 i 而言均存在整数 n_i 使得

$$\|y_i-x_{n_i}\|<\frac{\varepsilon}{4}.$$

取 $r>0$ 满足

$$2^{n_i}r<\frac{\varepsilon}{2},\quad\forall i=1,\cdots,k.$$

下面说明对这样的 r 即有 $U\subset V$. 这是因为若 $d(f,f_0)<r$, 根据 $d(f,g)$ 的构造知

$$\sum_{n=1}^{\infty}\frac{1}{2^n}|\langle f-f_0,x_n\rangle|<r,$$

因此

$$\frac{1}{2^{n_i}}|\langle f-f_0,x_{n_i}\rangle|<r,\quad\forall i=1,\cdots,k.$$

于是

$$|\langle f-f_0,y_i\rangle|=|\langle f-f_0,y_i-x_{n_i}\rangle+\langle f-f_0,x_{n_i}\rangle|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}.$$

这说明 $f\in V$.

下面再说明 d 诱导拓扑中的开集是 $\sigma(E^*,E)$ 中的开集. 任取 $f_0\in B_{E^*}$, 取定 $r>0$, 现在需要找到 f_0 在 $\sigma(E^*,E)$ 中的某邻域 V 使得

$$V\subset U=\{f\in B_{E^*}:d(f,f_0)<r\}.$$

与前文同理, 不妨设

$$V=\{f\in B_{E^*}:\langle f-f_0,x_i\rangle<\varepsilon,\forall i=1,\cdots,k\},$$

其中 ε,k 就是我们用来调整以使得 $V\subset U$ 的量. 现在对 $f\in V$ 有

$$\begin{aligned}d(f,f_0)&=\sum_{n=1}^{k}\frac{1}{2^n}|\langle f-f_0,x_n\rangle|+\sum_{n=k+1}^{\infty}\frac{1}{2^n}|\langle f-f_0,x_n\rangle|\\&<\varepsilon+\sum_{n=k+1}^{\infty}\frac{1}{2^n}((\|f\|+\|f_0\|)\|x_n\|)<\varepsilon+2\sum_{n=k+1}^{\infty}\frac{1}{2^n}=\varepsilon+\frac{1}{2^{k-1}}.\end{aligned}$$

取 $\varepsilon=\frac{r}{2}$, 令 k 充分大以满足 $\frac{1}{2^{k-1}}<\frac{r}{2}$ 即可.

反之, 设 B_{E^*} 在 $\sigma(E^*,E)$ 下可度量, 往证 E 可分. 记

$$U_n=\left\{f\in B_{E^*}:d(f,0)<\frac{1}{n}\right\}$$

另记 V_n 是 0 在 $\sigma(E^*,E)$ 中的邻域, 且其满足 $V_n\subset U_n$. 不妨设

$$V_n=\{f\in B_{E^*}:|\langle f,x\rangle|<\varepsilon_n,\forall x\in\Phi_n\},$$

其中 $\varepsilon_n>0,\Phi_n$ 是 E 的有限子集. 记

$$D=\bigcup_{n=1}^{\infty}\Phi_n,$$

显见 D 可数.

显见 D 以 $\mathbb{Q}$ 为背景张成的线性空间在 D 以 $\mathbb{R}$ 为背景张成的线性空间中稠密, 往证 D 以 $\mathbb{R}$ 为背景张成的线性空间 $D_{\mathbb{R}}$ 在 E 中稠密, 为此考虑推论2.4. 设 $f\in E^*$ 满足 $\langle f,x\rangle=0(\forall x\in D_{\mathbb{R}})$, 根据 f 的线性性知 $\langle f,x\rangle=0(\forall x\in D)$, 由 D 的构造可知 $f\in V_n(\forall n)$, 因此 $f\in U_n(\forall n)$, 这说明 $f=0$. □

对于定理3.10, 证明 E^* 可分 $\Rightarrow B_E$ 在 $\sigma(E,E^*)$ 中可度量的过程与前面完全一样, 只需把 E 和 E^* 的位置对调即可. 但是它的逆命题不能再用前面的方法了, 这是因为 E 未必是自反空间, 当 $E^{**} \supsetneq J(E)$ 时, 就算构造了与前面类似的 $D \subset E^*$, 通过推论2.4我们只能设 $x \in E^{**}$ 满足 $\langle x, f\rangle = 0(\forall f \in D)$, 而这没法说明 x 在 $V_n \subset E$ 中. 这一方向的证明我们就暂时省略了.

注 我们需要再次强调: 若 E 是无穷维空间, 则全空间 E(或 E^*) 在弱拓扑 $\sigma(E,E^*)$(或弱 * 拓扑 $\sigma(E^*,E)$) 下并不可度量化. 特别地, E^* 上由范数 $[\cdot]$ 诱导的拓扑与弱 * 拓扑并不相同.

推论 3.8

设 E 是可分 Banach 空间, $\{f_n\}$ 是 E^* 中的有界序列, 则该序列存在在弱 * 拓扑 $\sigma(E^*,E)$ 下收敛的子列 $\{f_{n_k}\}$.

证明 不失一般性, 可设 $\|f_n\| \leq 1(\forall n)$. 根据 Banach-Alaoglu-Bourbaki 定理3.3, 集合 B_{E^*} 在弱 * 拓扑 $\sigma(E^*,E)$ 中紧, 由可分与单位闭球可度量的等价性3.9知 B_{E^*} 在 $\sigma(E^*,E)$ 下可度量. 回忆在度量空间下, 紧集等价于自列紧集, 故 B_{E^*} 在 $\sigma(E^*,E)$ 的意义下是自列紧集. 利用有界的定义, 设存在 $M > 0$ 使得 $\{f_n\} \subset MB_{E^*}$, 容易说明 MB_{E^*} 同样是 $\sigma(E^*,E)$ 下的可度量紧集, 由其列紧性即得欲证. □

最后我们给出定理3.5的证明, 为方便阅读这里重申内容:

定理 3.11 (自反空间的弱列紧性)

设 E 是自反 Banach 空间, $\{x_n\}$ 是 E 中的有界序列, 则 $\{x_n\}$ 存在在弱拓扑 $\sigma(E,E^*)$ 下收敛的子列 $\{x_{n_k}\}$.

证明 设 M_0 是由 $\{x_n\}$ 以 $\mathbb{R}$ 为背景张成的线性空间, 记 $M = \overline{M_0}$, 仿照 Banach 定理3.8证明中的过程可以说明 M 可分. 又因为 M 是 E 的闭线性子空间, 由闭线性子空间对自反性的继承3.11可知 M 也是自反空间. 现在因为 M 自反且可分, 由推论3.7可知 M^* 也自反且可分, 于是由 M^* 的可分性与定理3.10可知 B_M 是弱拓扑 $\sigma(M,M^*)$ 下的可度量紧集. 根据度量空间中紧性与自列紧性的等价, 可知 M 中的有界列 $\{x_n\}$ 在 $\sigma(M,M^*)$ 的意义下必有收敛子列 $\{x_{n_k}\}$, 又因为命题3.11的证明中已经指出在 M 上 $\sigma(E,E^*)$ 所诱导的拓扑与 $\sigma(M,M^*)$ 实际上是同一种拓扑, 故 $\{x_{n_k}\}$ 在 $\sigma(E,E^*)$ 的意义下收敛. □

3.7 一致凸空间

定义 3.5 (一致凸 (uniformly convex))

若 Banach 空间 E 满足

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in E \left(\|x\| \leq 1, \|y\| \leq 1, \|x-y\| > \varepsilon \Rightarrow \left\|\frac{x+y}{2}\right\| < 1-\delta\right),$$

就称 E 一致凸.

一致凸性可以看成单位球的几何性质: 如果我们在单位球中牵一条长度为 ε 的弦, 则这条弦的中点必须落在半径为 $1-\delta$ 的球内 (其中 $\delta > 0$). 特别地, 这说明单位球面必须是" 圆" 的, 它不能包含任何线段.

例 3.3 设 $E = \mathbb{R}^2$, 则范数 $\|x\|_2 = [|x_1|^2 + |x_2|^2]^{\frac{1}{2}}$ 一致凸, 而范数 $\|x\|_1 = |x_1| + |x_2|, \|x\|_\infty = \max(|x_1|, |x_2|)$ 并不一致凸. 画出这些范数下单位球的形状就能知道原因了.

例 3.4 之后会说明在 $1 < p < \infty$ 时, L^p 都是一致凸空间, 且 Hilbert 空间也是一致凸空间.

定理 3.12 (Milman-Pettis)

每个一致凸 Banach 空间均自反.

注 一致凸性本身是范数的几何性质, 范数的等价并不保持其一致凸性. 另一方面, 自反性是一个拓扑性质: 范数的等价保持空间的自反性. 因此, 上述定理的结果非常惊人: 它表明一个几何性质能导出一个拓扑性质. 通过上述定

理, 我们会经常用一致凸性来证明自反性, 但一致凸性并不是干这件事的最佳工具: 存在一些诡异的自反空间, 在它们上面不存在任何一致凸的等价范数!

证明 设 $\xi \in E^{**}$ 满足 $\|\xi\| = 1$, 往证 $\xi \in J(B_E)$. 因为 $J(B_E)$ 在强拓扑下是 E^{**} 中的闭集, 故只需说明

$$\forall \varepsilon > 0 \exists x \in B_E(\|\xi - J(x)\| \le \varepsilon). \tag{3.10}$$

取定 $\varepsilon > 0$, 设 $\delta > 0$ 是一致凸性产生的数. 因为 $\|\xi\| = 1$, 故总能取到 $f \in E^*$ 使得 $\|f\| = 1$ 的同时有

$$\langle \xi, f \rangle > 1 - \frac{\delta}{2}. \tag{3.11}$$

记

$$V = \left\{ \eta \in E^{**} : |\langle \eta - \xi, f \rangle| < \frac{\delta}{2} \right\}$$

根据弱 * 拓扑的邻域基3.8可知 V 是 ξ 在弱 * 拓扑 $\sigma(E^{**}, R^*)$ 下的邻域. 根据 Goldstine 引理3.4, $J(B_E)$ 在 $\sigma(E^{**}, E^*)$ 下是 $B_{E^{**}}$ 的稠子集, 因此 $V \cap J(B_E) \ne \varnothing$, 故可设 $x \in B_E$ 满足 $J(x) \in V$. 下面说明这个 x 就满足(3.10)式.

考虑反证法, 设 $\|\xi - Jx\| > \varepsilon$, 即 $\xi \in (Jx + \varepsilon B_{E^{**}})^c = W$. 因为 $B_{E^{**}}$ 在 $\sigma(E^{**}, E^*)$ 下闭[4], 而闭性是平移与缩放不变的, 故 $Jx + \varepsilon B_{E^{**}}$ 是 $\sigma(E^{**}, E^*)$ 中的闭集, 因此 W 是 ξ 在拓扑 $\sigma(E^{**}, E^*)$ 下的邻域. 再次应用 Goldstine 引理3.4, 知 $V \cap W \cap J(B_E) \ne \varnothing$, 这说明存在 $y \in B_E$ 使得 $J(y) \in V \cap W$. 因为 $J(x), J(y) \in V$, 故

$$|\langle f, x \rangle - \langle \xi, f \rangle| < \frac{\delta}{2},$$
$$|\langle f, y \rangle - \langle \xi, f \rangle| < \frac{\delta}{2}.$$

将上述两式相加可得

$$2\langle \xi, f \rangle < \langle f, x + y \rangle + \delta \le \|x + y\| + \delta.$$

与(3.11)式结合可知

$$\left\| \frac{x + y}{2} \right\| > 1 - \delta.$$

根据一致凸性, 上式成立时只能有 $\|x - y\| \le \varepsilon$, 但这与 $J(y) \in W$ 所蕴含的 $\|x - y\| > \varepsilon$ 矛盾! 至此即证命题. □

我们最后以一致凸性的一个有用性质来收尾.

命题 3.13

设 E 是一致凸 Banach 空间, $\{x_n\} \subset E$ 满足 $x_n \rightharpoonup x$ 在 $\sigma(E, E^*)$ 中成立, 且

$$\limsup \|x_n\| \le \|x\|,$$

则 $x_n \to x$ 在强拓扑意义下成立.

证明 不妨设 $x \ne 0$(否则结论显然成立). 记

$$\lambda_n = \max(\|x_n\|, \|x\|), y_n = \lambda_n^{-1} x_n, y = \|x\|^{-1} x,$$

可知 $\lambda_n \to \|x\|$, 且 $y_n \hookrightarrow y$ 在 $\sigma(E, E^*)$ 中成立. 根据弱收敛的性质3.5(iii) 可知

$$\|y\| \le \liminf \|(y_n + y)/2\|.$$

另一方面, 注意到 $\|y\| = 1, \|y_n\| \le 1$, 故 $\|(y_n + y)/2\| \to 1$, 根据一致凸性知只能有 $\|y_n - y\| \to 0$, 因此 $x_n \to x$ 在强拓扑意义下成立. □

[4]前者显然, 对后者而言, 由 $\sigma(E^{**}, E^*)$ 是 Hausdorff 拓扑知 $\sigma(E^{**}, E^*)$ 中的紧集一定是闭集, 根据 Banach-Alaoglu-Bourbaki 定理3.3可得 $B_{E^{**}}$ 在 $\sigma(E^{**}, E^*)$ 中紧.

第四章 L^p 空间

本章选自 [**HB**].

记 $(\Omega, \mathcal{M}, \mu)$ 为测度空间, 此时 Ω 是集合, 且

(i) $\mathcal{M}$ 是 Ω 中的 σ-代数. 也就是说, $\mathcal{M}$ 是 Ω 的某些子集构成的集族, 其中

  (a) $\varnothing \in \mathcal{M}$,

  (b) $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$,

  (c) 当 $\forall n (A_n \in \mathcal{M})$ 时 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

(ii) μ 是一个**测度 (measure)**, 即 μ 作为 $\mathcal{M} \to [0, \infty]$ 的函数满足

  (a) $\mu(\varnothing) = 0$,

  (b) 当 $\{A_n\}$ 是 $\mathcal{M}$ 中的不交可数集族时有 $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

  称 $\mathcal{M}$ 中的元素为**可测集 (measurable set)**. 有时我们把 $\mu(A)$ 写成 $|A|$. 另外, 尽管下述假设并不是本质上的, 我们还是考虑:

(iii) Ω 是 σ-有限的, 即存在 $\mathcal{M}$ 中的可数集族 $\{\Omega_n\}$ 使得 $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$, 且对 $\forall n$ 均有 $\mu(\Omega_n) < \infty$.

如果集合 $E \in \mathcal{M}$ 满足 $\mu(E) = 0$, 就称 E 为**零测集 (null set)**. 如果一个性质在 Ω 上除去某零测集后处处成立, 就称该性质 a.e. 成立.

在阅读本章时, **可测函数 (measurable function)**, **可积函数 (integrable function)** 与 L^p 空间的一些基本性质已经应该是老朋友了, 因此我们不再给出它们的具体说明. 一般我们把 $\Omega \to \mathbb{R}$ 的全体可积函数构成的空间记为 $L^1(\Omega, \mu)$, 简记之为 $L^1(\Omega)$(或 L^1). 另外简便起见, 我们把积分写成 $\int f$ 而非 $\int_\Omega f d\mu$, 且在 L^1 范数上作下述符号约定:

$$\|f\|_{L^1} = \|f\|_1 = \int_\Omega |f| d\mu = \int |f|.$$

同时, L^1 中两元素的相等是在 a.e. 意义下的. 接下来我们回忆 L^1 上的一些最基本性质.

4.1 关于积分人人须知的结果

定理 4.1 (Beppo Levi 单调收敛定理)

设 $\{f_n\} \subset L^1$ 满足

(i) $f_1 \le f_2 \le \cdots \le f_n \le f_{n+1} \le \cdots$ 在 Ω 上 a.e.

(ii) $\sup_n \int f_n < \infty$.

则 $f_n(x)$ 在 Ω 上 a.e. 收敛到某有限极限. 若记该极限为 $f(x)$, 则 $f \in L^1$, 且 $\|f_n - f\|_1 \to 0$.

定理 4.2 (Lebesgue 控制收敛定理)

设 $\{f_n\} \subset L^1$ 满足

(i) $f_n(x) \to f(x)$ 在 Ω 上 a.e.

(ii) 存在 $g \in L^1$ 使得对 $\forall n$ 而言, $|f_n(x)| \le g(x)$ 在 Ω 上 a.e.

则 $f \in L^1$, 且 $\|f_n - f\|_1 \to 0$.

引理 4.1 (Fatou 引理)

设 $\{f_n\} \subset L^1$ 满足

(i) 对 $\forall n$ 而言, $f_n \ge 0$ a.e.

(ii) $\sup_n \int f_n < \infty$.

对 a.e. $x \in \Omega$ 记 $f(x) = \liminf_{n\to\infty} f_n(x) \le +\infty$, 则 $f \in L^1$, 且

$$\int f \le \liminf_{n\to\infty} \int f_n.$$

在 $\Omega = \mathbb{R}^N$ 这一最基础的情况中, $\mathcal{M}$ 由 Lebesgue 可测集组成, μ 是 $\mathbb{R}^N$ 上的 Lebesgue 测度. 记 $C_c(\mathbb{R}^N)$ 为 $\mathbb{R}^N$ 中全体紧支连续函数构成的空间, 即

$$C_c(\mathbb{R}^N) = \{f \in C(\mathbb{R}^N) : \exists K \subset \mathbb{R}^N \forall x \in \mathbb{R}^N \setminus K (f(x) = 0), \text{ 其中 } K \text{ 是紧集}\}$$

定理 4.3 ($C_c(\mathbb{R}^N)$ 在 $L^1(\mathbb{R}^N)$ 中的稠密性)

空间 $C_c(\mathbb{R}^N)$ 在 $L^1(\mathbb{R}^N)$ 中稠密, 即

$$\forall f \in L^1(\mathbb{R}^N) \forall \varepsilon > 0 \exists f_1 \in C_c(\mathbb{R}^N)(\|f - f_1\|_1 \le \varepsilon).$$

设 $(\Omega_1, \mathcal{M}_1, \mu_1), (\Omega_2, \mathcal{M}_2, \mu_2)$ 是两个 σ-有限测度空间, 通过乘积测度可以在直积 $\Omega = \Omega_1 \times \Omega_2$ 上定义测度空间结构 $(\Omega, \mathcal{M}, \mu)$.

定理 4.4 (Tonelli)

设 $F(x, y) : \Omega_1 \times \Omega_2 \to \mathbb{R}$ 是可测函数, 且

(i) $\int_{\Omega_2} |F(x, y)| d\mu_2 < \infty$ 对 a.e. $x \in \Omega_1$ 成立,

(ii) $\int_{\Omega_1} d\mu_1 \int_{\Omega_2} |F(x, y)| d\mu_2 < \infty$,

则 $F \in L^1(\Omega_1 \times \Omega_2)$.

定理 4.5 (Fubini)

设 $F \in L^1(\Omega_1 \times \Omega_2)$, 则对 a.e. $x \in \Omega_1$ 而言, $F(x, y) \in L^1_y(\Omega_2)$, $\int_{\Omega_2} F(x, y) d\mu_2 \in L^1_x(\Omega_1)$. 类似地, 对 a.e. $y \in \Omega_2$ 而言, $F(x, y) \in L^1_x(\Omega_1)$, $\int_{\Omega_1} F(x, y) d\mu_1 \in L^1_y(\Omega_2)$. 另有下述积分换序成立:

$$\int_{\Omega_1} d\mu_1 \int_{\Omega_2} F(x, y) d\mu_2 = \int_{\Omega_2} d\mu_2 \int_{\Omega_1} F(x, y) d\mu_1 = \iint_{\Omega_1 \times \Omega_2} F(x, y) d\mu_1 d\mu_2.$$

4.2 L^p 空间的定义与初等性质

定义 4.1 (L^p)

设 $p \in \mathbb{R}$ 满足 $1 < p < \infty$, 则记

$$L^p(\Omega) = \{f : \Omega \to \mathbb{R} : f \text{ 可测, 且 } |f|^p \in L^1(\Omega)\}$$

并记

$$\|f\|_{L^p} = \|f\|_p = \left[\int_\Omega |f(x)|^p d\mu\right]^{\frac{1}{p}}.$$

容易验证 $\|\cdot\|_p$ 是范数.

定义 4.2 (L^∞)

记

$$L^\infty(\Omega) = \{f : \Omega \to \mathbb{R} : f \text{ 可测, 且存在常数 } C \text{ 使得 } |f(x)| \le C \text{ 在 } \Omega \text{ 上 a.e.}\}$$

并记

$$\|f\|_{L^\infty} = \|f\|_\infty = \inf\{C : |f(x)| \le C \text{在 } \Omega \text{ 上 a.e.}\}$$

同样可以验证 $\|\cdot\|_\infty$ 是范数. ♣

对 $1 \le p \le \infty$, 约定 p' 为其**共轭指标 (conjugate exponent)**:

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

定理 4.6 (Hölder 不等式)

设 $1 \le p \le \infty, f \in L^p, g \in L^{p'}$, 则 $fg \in L^1$, 且

$$\int |fg| \le \|f\|_p \|g\|_{p'}.$$

♡

注 Hölder 不等式有下述拓展: 设 $f_1, \cdots, f_k$ 满足

$$f_i \in L^{p_i}, \quad 1 \le i \le k, \frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k} \le 1$$

则 $f = f_1 \cdots f_k \in L^p$, 且

$$\|f\|_p \le \|f_1\|_{p_1} \cdots \|f_k\|_{p_k}.$$

特别若 $f \in L^p \cap L^q (1 \le p \le q \le \infty)$, 则对 $\forall r$ 满足 $p \le r \le q$, 总有 $f \in L^r$, 且成立下述**插值不等式 (interpolation inequality)**:

$$\|f\|_r \le \|f\|_p^\alpha \|f\|_q^{1-\alpha}, \quad \frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}, 0 \le \alpha \le 1.$$

定理 4.7

L^p 是线性空间, $\|\cdots\|_p$ 对任意 $1 \le p \le \infty$ 均为范数. ♡

定理 4.8 (Fischer-Riesz)

L^p 在 $1 \le p \le \infty$ 时是 Banach 空间. ♡

定理 4.9 (Riesz)

设 $\{f_n\} \subset L^p, f \in L^p$ 满足 $\|f_n - f\|_p \to 0$, 则存在子列 $\{f_{n_k}\}$ 与函数 $h \in L^p$ 使得

(i) $f_{n_k}(x) \to f(x)$ 在 Ω 上 a.e.

(ii) $|f_{n_k}(x)| \le h(x) (\forall k)$ 在 Ω 上 a.e.

♡

4.3 L^p 的自反性, 可分性与对偶

我们分下述三种情况考虑:

(A) $1 < p < \infty$,

(B) $p = 1$,

(C) $p = \infty$.

4.3.1 $1 < p < \infty$ 的情况

在 $1 < p < \infty$ 时, L^p 的性质是最令人舒适的: 它自反, 可分, 且 L^p 的对偶正是 $L^{p'}$.

定理 4.10

在 $1 < p < \infty$ 时 L^p 自反. ♡

证明 证明分为下述三步:

$\boxed{\text{第一步: Clarkson 第一不等式.}}$ 设 $2 \le p < \infty$, 我们断言

$$\left\|\frac{f+g}{2}\right\|_p^p + \left\|\frac{f-g}{2}\right\|_p^p \le \frac{1}{2}(\|f\|_p^p + \|g\|_p^p), \quad \forall f, g \in L^p. \tag{4.1}$$

下面证明(4.1)式. 事实上, 只要能证明

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \le \frac{1}{2}(|a|^p + |b|^p), \quad \forall a, b \in \mathbb{R} \tag{4.2}$$

令 $a = f(x), b = g(x)$ 并在两端对 x 积分即得(4.1)式, 故我们转而考虑(4.2)式. 注意到函数

$$(x^2+1)^{\frac{p}{2}} - x^p - 1$$

在 $[0, \infty)$ 上递增, 由此可得

$$\alpha^p + \beta^p \le (\alpha^2+\beta^2)^{\frac{p}{2}}, \quad \forall \alpha, \beta \ge 0.$$

取 $\alpha = \left|\frac{a+b}{2}\right|, b = \left|\frac{a-b}{2}\right|$ 即得

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \le \left(\left|\frac{a+b}{2}\right|^2 + \left|\frac{a-b}{2}\right|^2\right)^{\frac{p}{2}} = \left(\frac{a^2}{2} + \frac{b^2}{2}\right)^{\frac{p}{2}} \le \frac{1}{2}(|a|^p + |b|^p).$$

至此即得(4.1)式.

$\boxed{\text{第二步: } 2 \le p < \infty \text{ 时 } L^p \text{ 一致凸, 此时它自反.}}$ 事实上, 取定 $\varepsilon > 0$, 设 $f, g \in L^p$ 满足 $\|f\|_p \le 1, \|g\|_p \le 1, \|f-g\|_p > \varepsilon$, 由 Clarkson 第一不等式(4.1)知

$$\left\|\frac{f+g}{2}\right\|_p^p < 1 - \left(\frac{\varepsilon}{2}\right)^p,$$

因此 $\left\|\frac{f+g}{2}\right\|_p < 1-\delta$, 其中 $\delta = 1 - [1-(\varepsilon/2)^p]^{\frac{1}{p}} > 0$. 这说明 L^p 是一致凸的, 由 Milman-Pettis 定理3.12即知 L^p 自反.

$\boxed{\text{第三步: } 1 < p \le 2 \text{ 时 } L^p \text{ 自反.}}$ 设 $1 < p < \infty$, 按照下述方法定义算子 $T : L^p \to (L^{p'})^*$: 固定 $u \in L^p$, 知映射 $L^{p'} \ni f \mapsto \int uf$ 定义了 $L^{p'}$ 上的一个连续线性泛函, 因此这一映射是 $(L^{p'})^*$ 中的元素, 记之为 Tu. 即:

$$\langle Tu, f\rangle = \int uf, \quad \forall f \in L^{p'}.$$

现在断言

$$\|Tu\|_{(L^{p'})^*} = \|u\|_p, \quad \forall u \in L^p. \tag{4.3}$$

这是因为一方面根据 Hölder 不等式知

$$|\langle Tu, f\rangle| \le \|u\|_p \|f\|_{p'}, \quad \forall f \in L^{p'}$$

因此 $\|Tu\|_{(L^{p'})^*} \le \|u\|_p$. 另一方面, 记

$$f_0(x) = |u(x)|^{p-2}u(x) \quad (u(x) = 0\text{时}\, f_0(x) = 0)$$

显见

$$f_0 \in L^{p'}, \|f_0\|_{p'} = \|u\|_p^{p-1}, \langle Tu, f_0\rangle = \|u\|_p^p;$$

因此

$$\|Tu\|_{(L^{p'})^*} \ge \frac{\langle Tu, f_0\rangle}{\|f_0\|_{p'}} = \|u\|_p. \tag{4.4}$$

故 T 实际上是 $L^p \to (L^{p'})^*$ 的等距同构, 又因为 L^p 是 Banach 空间, 故 $T(L^p)$ 是 $(L^{p'})^*$ 的闭线性子空间.

现设 $1 < p \le 2$, 第二步说明了 $L^{p'}$ 自反, 由对偶保持自反性3.4知 $(L^{p'})^*$ 也自反. 因为闭线性子空间继承自反性 (命题3.11), 故 $T(L^p)$ 自反, 而推论3.4的证明中说明了线性等距同构保持自反性, 故 L^p 自反. □

注 事实上, $1 < p \le 2$ 时 L^p 也是一致凸空间, 这需要用到 $1 < p \le 2$ 时成立的 Clarkson 第二不等式:

$$\left\|\frac{f+g}{2}\right\|_p^{p'} + \left\|\frac{f-g}{2}\right\|_p^{p'} \le \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{\frac{1}{p-1}}, \quad \forall f, g \in L^p.$$

这一不等式的具体证明我们就省略了.

定理 4.11 (Riesz 表示定理)

设 $1 < p < \infty, \phi \in (L^p)^*$, 则存在唯一函数 $u \in L^{p'}$ 使得

$$\langle \phi, f \rangle = \int uf, \quad \forall f \in L^p$$

且

$$\|u\|_{p'} = \|\phi\|_{(L^p)^*}.$$

注 Riesz 表示定理4.11在 L^p 空间理论中非常重要. 它表明 $1 < p < \infty$ 时 L^p 上的每个连续线性泛函都能" 具体地" 表为某个积分. 映射 $\phi \mapsto u$ 作为线性等距同构允许我们把" 抽象的" 空间 $(L^p)^*$ 与 $L^{p'}$ 等同起来.

在 Riesz 表示定理得证后, 我们将统一作下述恒同约定:

$$(L^p)^* = L^{p'}.$$

证明 考察算子

$$T : L^{p'} \to (L^p)^*, \langle Tu, f \rangle = \int uf, \quad \forall u \in L^{p'}, \forall f \in L^p.$$

定理4.10证明中的第三步表明

$$\|Tu\|_{(L^p)^*} = \|u\|_{p'}, \quad \forall u \in L^{p'}.$$

我们断言 T 是满射. 事实上, 记 $E = T(L^{p'})$, 因为 E 已经是 $(L^p)^*$ 的闭子空间了, 故只需证明 E 在 $(L^p)^*$ 中稠密, 为此考虑推论2.4. 设 $h \in (L^p)^{**}$ 满足 $\langle h, Tu \rangle = 0 (\forall u \in L^{p'})$, 因为 L^p 自反, 故 $h \in L^p$, 因此 $\int uh = 0 (\forall u \in L^{p'})$. 取 $u = |h|^{p-2}h$ 即知 $h = 0$. □

定理 4.12 ($C_c(\mathbb{R}^N)$ 在 $L^p(\mathbb{R}^N)$ 中的稠密性)

当 $1 \le p < \infty$ 时, 空间 $C_c(\mathbb{R}^N)$ 在 $L^p(\mathbb{R}^N)$ 中稠密.

在证明定理4.12前, 我们先作下述符号约定: **截断算子 (truncation operation)**$T_n : \mathbb{R} \to \mathbb{R}$ 定义为

$$T_n r = \begin{cases} r, & |r| \le n \\ \frac{nr}{|r|}, & |r| > n. \end{cases}$$

给定集合 $E \subset \Omega$, 定义**示性函数 (characteristic function)**χ_E 为

$$\chi_E(x) = \begin{cases} 1, & x \in E \\ 0, & x \in \Omega \backslash E. \end{cases}$$

证明 首先, 我们断言对给定的 $f \in L^p(\mathbb{R}^N)$ 与 $\varepsilon > 0$ 而言, 存在函数 $g \in L^\infty(\mathbb{R}^N)$ 与紧集 $K \subset \mathbb{R}^N$ 使得 $\operatorname{supp} g \subset K$, 且

$$\|f - g\|_p < \varepsilon. \tag{4.5}$$

事实上, 设 χ_n 是 $B(0, n)$ 的示性函数, 记 $f_n = \chi_n T_n f$. 根据控制收敛定理可知 $\|f_n - f\|_p \to 0$, 因此可在 n 足够大时取 $g = f_n$. 下一步, 对于已经取出的 g, 根据 $C_c(\mathbb{R}^N)$ 在 $L^1(\mathbb{R}^N)$ 中的稠密性知任意给定 $\delta > 0$, 总存在 $g_1 \in C_c(\mathbb{R}^N)$ 使得

$$\|g - g_1\|_1 < \delta.$$

不妨设 $\|g_1\|_\infty \le \|g\|_\infty$, 否则可用 $T_n g_1$ 代替 g_1(其中 $n = \|g\|_\infty$). 知

$$\|g - g_1\|_p \le \|g - g_1\|_1^{\frac{1}{p}} \|g - g_1\|_\infty^{1-\frac{1}{p}} \le \delta^{\frac{1}{p}} (2\|g\|_\infty)^{1-\frac{1}{p}},$$

取 δ 足够小使得

$$\delta^{\frac{1}{p}} (2\|g\|_\infty)^{1-\frac{1}{p}} < \varepsilon$$

即可. □

定义 4.3 (可分测度空间 (separable measure space))

称测度空间 Ω 可分, 如果存在 $\mathcal{M}$ 中的可数集族 $\{E_n\}$ 使得由 $\{E_n\}$ 生成的 σ-代数与 $\mathcal{M}$ 重合 (即 $\mathcal{M}$ 是包含所有 E_n 的最小 σ-代数.)

♣

例 4.1 测度空间 $\Omega = \mathbb{R}^N$ 是可分的. 这是因为可取 $\{E_n\}$ 为以有理点为球心, 以有理数为半径的开球族, 显见 $\mathbb{R}^N$ 中的每个开集都可以表为某些 E_n 的并. 更一般地, 只要 Ω 是可分度量空间, $\mathcal{M}$ 由 Borel 集组成 (即 $\mathcal{M}$ 是 Ω 中的开集生成的 σ-代数), 则 Ω 是可分测度空间.

定理 4.13 (L^p 的可分性)

设 Ω 是可分测度空间, 则 $1 \le p < \infty$ 时 $L^p(\Omega)$ 可分.

♡

这里我们只考虑 $\Omega = \mathbb{R}^N$ 的情况, 一般情况的证明需要更多技巧. 注意从该定理出发, 可以得到只要可测集 $\Omega \subset \mathbb{R}^N$, 那么 $L^p(\Omega)$ 就也可分. 这是因为存在 $L^p(\Omega) \to L^p(\mathbb{R}^N)$ 的典则同构 (即在 Ω 外令函数值为 0 的延拓映射); 因此 $L^p(\Omega)$ 与 $L^p(\mathbb{R}^N)$ 的某闭线性子空间等距同构, 根据闭线性子空间对可分性的继承3.11与线性等距同构保持可分性即知 $L^p(\Omega)$ 可分.

证明 当 $\Omega = \mathbb{R}^N$ 时, 记 $\mathcal{R}$ 为 $\mathbb{R}^N$ 中形如 $R = \prod_{k=1}^N (a_k, b_k)(a_k, b_k \in \mathbb{Q})$ 的集合构成的可数集族, 记 $\mathcal{E}$ 为函数族 $\{\chi_R\}_{R\in\mathcal{R}}$ 以 $\mathbb{Q}$ 为背景张成的线性空间, 即 $\mathcal{E}$ 由 χ_R 的全体有理系数有限线性组合构成, 显见 $\mathcal{E}$ 可数.

下面说明 $\mathcal{E}$ 在 $L^p(\mathbb{R}^N)$ 中稠密. 取定 $f \in L^p(\mathbb{R}^N), \varepsilon > 0$, 根据定理4.12知存在 $f_1 \in C_c(\mathbb{R}^N)$ 使得 $\|f - f_1\|_p < \varepsilon$, 设 $R \in \mathcal{R}$ 是包含 $\operatorname{supp} f_1$ 的方体. 对取定的 $\delta > 0$, 容易构造函数 $f_2 \in \mathcal{E}$ 满足 $\|f_1 - f_2\|_\infty < \delta, \operatorname{supp} f_2 \subset R$. 由此可知

$$\|f_1 - f_2\|_p \le \|f_1 - f_2\|_\infty |R|^{\frac{1}{p}} < \delta |R|^{\frac{1}{p}}$$

现在取 $\delta > 0$ 满足 $\delta |R|^{\frac{1}{p}} < \varepsilon$, 可知 $\|f - f_2\|_p < 2\varepsilon$, 因此 $\mathcal{E}$ 在 $L^p(\mathbb{R}^N)$ 中稠密. □

4.3.2 $p = 1$ 的情况

我们从 $L^1(\Omega)$ 对偶空间的描述开始介绍.

定理 4.14 (Riesz 表示定理)

设 $\phi \in (L^1)^*$, 则存在唯一函数 $u \in L^\infty$ 使得

$$\langle \phi, f \rangle = \int uf, \quad \forall f \in L^1$$

且

$$\|u\|_\infty = \|\phi\|_{(L^1)^*}.$$

♡

注 定理4.14表明 L^1 上的每个连续线性泛函都能" 具体地" 表为某个积分. 映射 $\phi \mapsto u$ 作为线性等距同构允许我们把" 抽象的" 空间 $(L^1)^*$ 与 L^∞ 等同起来. 在上述定理得证后, 我们将统一作下述恒同约定:

$$(L^1)^* = L^\infty.$$

证明 设 $\{\Omega_n\}$ 为 Ω 中的可测集列, 其满足 $\Omega = \bigcup_{n=1}^\infty \Omega_n, |\Omega_n| < \infty (\forall n)$. 记 $\chi_n = \chi_{\Omega_n}$.

u 的唯一性首先是显然的, 这是因为若 $u \in L^\infty$ 满足

$$\int uf = 0, \quad \forall f \in L^1$$

取 $f = \chi_n \operatorname{sgn} u$(约定 $\operatorname{sgn} 0 = 0$) 可知 $u = 0$ 在 Ω_n 上 a.e., 令 n 跑遍 $\mathbb{N}$ 即知 $u = 0$ 在 Ω 上 a.e.

下面说明 u 的存在性. 首先构造函数 $\theta \in L^2(\Omega)$ 使得

$$\theta(x) \ge \varepsilon_n > 0, \quad \forall x \in \Omega_n.$$

这样的 θ 肯定存在. 事实上, 只要定义 θ 在 Ω_1 上取值为 α_1, 在 $\Omega_2 \backslash \Omega_1$ 上取值为 α_2,..., 在 $\Omega_n \backslash \Omega_{n-1}$ 上取值为 α_n 等等, 调整 $\alpha_n > 0$ 的值使得 $\theta \in L^2$ 即可.

根据 Hölder 不等式, $\theta f \in L^1$, 于是映射 $L^2(\Omega) \ni f \mapsto \langle \phi, \theta f \rangle$ 是 $L^2(\Omega)$ 上的连续线性泛函. 根据 $p = 2$ 时的 Riesz 表示定理4.11知存在函数 $v \in L^2(\Omega)$ 使得

$$\langle \phi, \theta f \rangle = \int v f, \quad \forall f \in L^2(\Omega). \tag{4.6}$$

记 $u(x) = v(x)/\theta(x)$. 因为在 Ω 上 $\theta > 0$, 故 u 首先是良定义的. 另外显见 u 可测, 且 $u\chi_n \in L^2(\Omega)$, 下面说明 u 就是欲求的函数. 对 $g \in L^\infty(\Omega)$, 令 $f = \chi_n g/\theta$, 因为 f 在 Ω_n 上有界且在 Ω_n 外为零, 故 $f \in L^2(\Omega)$, 将其代入(4.6)式可得

$$\langle \phi, \chi_n g \rangle = \int u \chi_n g, \quad \forall g \in L^\infty(\Omega), \forall n. \tag{4.7}$$

下面我们说明 $u \in L^\infty(\Omega)$, 且

$$\|u\|_\infty \le \|\phi\|_{(L^1)^*}. \tag{4.8}$$

取定常数 $C > \|\phi\|_{(L^1)^*}$, 记

$$A = \{x \in \Omega : |u(x)| > C\}.$$

下面说明 A 是零测集. 在(4.7)式中取 $g = \chi_A \operatorname{sgn} u$ 可得

$$\int_{A\cap\Omega_n} |u| \le \|\phi\|_{(L^1)^*} |A \cap \Omega_n|$$

因此

$$C|A \cap \Omega_n| \le \|\phi\|_{(L^1)^*} |A \cap \Omega_n|.$$

这说明只能有 $|A \cap \Omega_n| = 0 (\forall n)$, 因此 A 是零测集, 这便证明了(4.8)式.

最后, 我们说明

$$\langle \phi, h \rangle = \int u h, \quad \forall h \in L^1(\Omega). \tag{4.9}$$

这只需要在(4.7)式中取 $g = T_n h$, 观察到 $\chi_n T_n h \to h$ 在 $L^1(\Omega)$ 意义下成立, 利用 Hölder 不等式与 Lebesgue 控制收敛定理即得上式.

为完成定理4.14的证明, 我们还需验证 $\|u\|_\infty = \|\phi\|_{(L^1)^*}$. 由(4.9)式知

$$|\langle \phi, h \rangle| \le \|u\|_\infty \|h\|_1, \quad \forall h \in L^1(\Omega)$$

因此 $\|\phi\|_{(L^1)^*} \le \|u\|_\infty$, 结合(4.8)式即得欲证. □

注 除非 Ω 仅包含有限个原子[1](此时 $L^1(\Omega)$ 有限维), 不然空间 $L^1(\Omega)$ 不可能自反. 这是因为如若 $L^1(\Omega)$ 自反, 则只可能有下述两种情况出现:

(i) 对任意 $\varepsilon > 0$, 存在 $\omega \subset \Omega$ 可测且 $0 < \mu(\omega) < \varepsilon$.

(ii) 存在 $\varepsilon > 0$ 使得对每个可测集 $\omega \subset \Omega$ 而言, 只要 $\mu(\omega) > 0$, 就有 $\mu(\omega) \ge \varepsilon$.

对于情形 (i), 可以构造递减可测集列 $\{\omega_n\}$ 使得 $\mu(\omega_n) > 0 (\forall n)$, 且 $\mu(\omega_n) \to 0$. (其中一个构造方法为任取集列 $\{\omega'_k\}$ 满足 $0 < \mu(\omega'_k) < 1/2^k$, 令 $\omega_n = \bigcup_{k=n}^\infty \omega'_k$.) 记 $\chi_n = \chi_{\omega_n}$, 定义 $u_n = \chi_n / \|\chi_n\|_1$. 因为 $\|u_n\|_1 = 1$ 且 $L^1(\Omega)$ 自反, 根据定理3.5与 $\{u_n\}$ 有界知存在子列 (依旧记为 $\{u_n\}$) 与 $u \in L^1$ 使得 $u_n \rightharpoonup u$ 在弱拓扑 $\sigma(L^1, L^\infty)$ 意义下成立, 亦即

$$\int u_n \phi \to \int u \phi, \quad \forall \phi \in L^\infty.$$

另一方面, 取定 j, 对 $n > j$ 可知 $\int u_n \chi_j = 1$. 在等式两端令 $n \to \infty$ 并由控制收敛定理可得 $\int u \chi_j = 1 (\forall j)$. 然而如果对 $\int u \chi_j$ 取 $j \to \infty$ 并应用控制收敛定理, 应该得到 $\int u \chi_j \to 0$, 矛盾!

对于情形 (ii), 此时空间 Ω 必为原子测度空间[2], 且 Ω 可表为不交原子 $\{a_n\}$ 的可数并 (除非它仅包含有限多个

[1]若测度空间 (Ω, μ) 的可测子集 A 满足 $\mu(A) > 0$, 且对任意可测集 $B \subset A$ 而言, 要么 $\mu(B) = 0$, 要么 $\mu(B) = \mu(A)$, 就称 A 是 X 的一个原子.

[2]即包含原子的测度空间.

原子). 此时 $L^1(\Omega)$ 与 l^1 同构, 因此只需证明 l^1 不自反即可. 考虑标准基底:

$$e_n = (\underbrace{0, 0, \cdots, 1}_{n}, 0, 0, \cdots).$$

若 l^1 自反, 根据定理3.5知存在子列 $\{e_{n_k}\}$ 与 $x \in l^1$ 使得 $e_{n_k} \rightharpoonup x$ 在弱拓扑 $\sigma(l^1, l^\infty)$ 下成立, 即

$$\langle \varphi, e_{n_k} \rangle \to \langle \varphi, x \rangle (k \to \infty), \quad \forall \varphi \in l^\infty.$$

现取

$$\varphi = \varphi_j = (\underbrace{0, 0, \cdots, 1}_{j}, 1, 1, \cdots)$$

知 $\langle \varphi_j, x \rangle = 1 (\forall j)$, 而由 $x \in l^1$ 知本应有 $\langle \varphi_j, x \rangle \to 0 (j \to \infty)$, 矛盾!

4.3.3 $p = \infty$ 的情况

表示定理4.14已经表明 $L^\infty = (L^1)^*$. 在作为对偶空间时, L^∞ 具有一些良好性质, 例如:

(i) 根据 Banach-Alaoglu-Bourbaki 定理3.3, 闭单位球 B_{L^∞} 在弱 * 拓扑 $(\sigma(L^\infty, L^1))$ 下紧.

(ii) 若 Ω 是 $\mathbb{R}^N$ 中的可测集, $\{f_n\} \subset L^\infty(\Omega)$, 则存在子列 $\{f_{n_k}\}$ 与 $f \in L^\infty(\Omega)$ 使得 $f_{n_k} \rightharpoonup f$ 在弱 * 拓扑 $\sigma(L^\infty, L^1)$ 下成立. (这由 L^1 可分与推论3.8即得.)

然而, $L^\infty(\Omega)$ 并不自反, 除非 Ω 仅包含有限个原子. 这是因为如果 $L^\infty(\Omega)$ 自反, 根据对偶对自反性的保持3.4可知 $L^1(\Omega)$ 也应该自反, 而前一条注已经说明了 $L^1(\Omega)$ 不自反. 从这一事实出发可知, L^∞ 的对偶空间 $(L^\infty)^*$ 包含 L^1(因为 $L^\infty = (L^1)^*$), 且 $(L^\infty)^*$ 严格大于 L^1. 也就是说, 存在 L^∞ 上的连续线性泛函 ϕ, 它不能表为

$$\langle \phi, f \rangle = \int uf, \quad u \in L^1, \forall f \in L^\infty.$$

下面我们给出这种泛函的一个" 具体" 例子. 记

$$\phi_0 : C_c(\mathbb{R}^N) \to \mathbb{R}, \quad f \mapsto f(0).$$

显见 ϕ_0 是赋范数 $\|\cdot\|_\infty$ 的 $C_c(\mathbb{R}^N)$ 上的连续线性泛函. 根据 Hahn-Banach 定理, 我们可以把 ϕ_0 延拓为 $L^\infty(\mathbb{R}^N)$ 上的连续线性泛函 ϕ, 且

$$\langle \phi, f \rangle = f(0), \quad \forall f \in C_c(\mathbb{R}^N). \tag{4.10}$$

现在说明不存在函数 $u \in L^1(\mathbb{R}^N)$ 使得

$$\langle \phi, f \rangle = \int uf, \quad \forall f \in L^\infty(\mathbb{R}^N). \tag{4.11}$$

用反证法, 设存在这样的函数 u, 由(4.10),(4.11)两式可知

$$\int_u f = 0, \quad \forall f \in C_c(\mathbb{R}^N), f(0) = 0.$$

这说明只能让 $u = 0$ 在 $\mathbb{R}^N \backslash \{0\}$ 上 a.e., 进而 $u = 0$ 在 $\mathbb{R}^N$ 上 a.e. 由(4.11)式进而可得

$$\langle \phi, f \rangle = 0, \quad \forall f \in L^\infty(\mathbb{R}^N),$$

但在 $f(0) \neq 0$ 时这显然与(4.10)式矛盾!

注 除非 Ω 由有限多个原子构成, 否则空间 $L^\infty(\Omega)$ 不自反. 为说明此事, 我们先说明下述引理.

引理 4.2

设 E 是 Banach 空间, 若存在集族 $\{O_i\}_{i\in I}$ 使得

(i) 对每个 $i \in I$, O_i 都是 E 的非空开子集,

(ii) $i \neq j$ 时 $O_i \cap O_j = \varnothing$,

(iii) I 不可数,

则 E 不可分.

证明 用反证法, 设 E 可分, 记 $\{u_n\}_{n\in\mathbb{N}}$ 是 E 的可数稠密子集. 对每个 $i \in I$, 根据 $\{u_n\}_{n\in\mathbb{N}}$ 的稠密性知 $I_i \cap \{u_n\}_{n\in\mathbb{N}} \neq$

$\emptyset$, 设 $u_{n(i)} \in O_i$. 映射 $i \mapsto n(i)$ 是单射, 这是因为若 $n(i) = n(j)$, 则 $u_{n(i)} = u_{n(j)} \in O_i \cap O_j$, 由条件 (ii) 知只能有 $i = j$. 因此 I 可数, 这与条件 (iii) 矛盾! □

下面说明 $L^\infty(\Omega)$ 不可分, 为此我们利用引理4.2, 断言存在 Ω 的两两不交的不可数可测集族 $\{\omega_i\}_{i\in I}$, 此时对称和 $\omega_i \Delta \omega_j$ 在 $i \neq j$ 时有正测度. 现在在引理4.2中取

$$O_i = \{f \in L^\infty(\Omega) : \|f - \chi_{\omega_i}\|_\infty < 1/2\}$$

(注意到 ω, ω' 不交时 $\|\chi_\omega - \chi_{\omega'}\|_\infty = 1$) 即可. 当 Ω 为 $\mathbb{R}^N$ 中的开集时, 这样的不可数集族 $\{\omega_i\}$ 显然存在 (考虑球 $B(x_0, r), x_0 \in \Omega, r > 0$ 充分小即可).

当 Ω 是一般的测度空间时, 我们把 Ω 分成其原子部分 Ω_a 与非原子部分 Ω_d; 下面讨论两种情况:

(i) Ω_d 不是零测集.

(ii) Ω_d 是零测集.

对于情形 (i), 对每个实数 $t \in (0, \mu(\Omega_d))$, 总能找到可测集 ω 使得 $\mu(\omega) = t$, 通过这种方法即可得到不交的不可数可测集族.

对于情形 (ii), 此时 Ω 由不交原子 $\{a_n\}$ 的可数并组成 (除非它仅包含有限多个原子). 对任意正整数族 $\mathbb{A} \subset \mathbb{N}$, 定义 $\omega_{\mathbb{A}} = \bigcup_{n\in\mathbb{A}} a_n$, 显见 $\{\omega_{\mathbb{A}}\}$ 就是我们要找的不交不可数可测集族.

下表总结了 Ω 为 $\mathbb{R}^N$ 中可测集时 $L^p(\Omega)$ 的主要性质:

	自反	可分	对偶空间
$L^p(1 < p < \infty)$	✓	✓	$L^{p'}$
L^1	×	✓	L^∞
L^∞	×	×	比 L^1 严格大

4.4 卷积与正则化

我们首先定义函数 $f \in L^1(\mathbb{R}^N)$ 与函数 $g \in L^p(\mathbb{R}^N)$ 的卷积:

定理 4.15 (Young)

设 $f \in L^1(\mathbb{R}^N), g \in L^p(\mathbb{R}^N)(1 \le p \le \infty)$, 则对 a.e. $x \in \mathbb{R}^N$ 而言函数 $y \mapsto f(x-y)g(y)$ 在 $\mathbb{R}^N$ 上可积. 若记

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x-y)g(y)dy$$

则 $f * g \in L^p(\mathbb{R}^N)$, 且

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

对 $\mathbb{R}^N$ 上给定的函数 f, 记 $\widetilde{f}(x) = f(-x)$, 则:

命题 4.1

设 $f \in L^1(\mathbb{R}^N), g \in L^p(\mathbb{R}^N), h \in L^{p'}(\mathbb{R}^N)$, 则

$$\int_{\mathbb{R}^N} (f * g)h = \int_{\mathbb{R}^N} g(\widetilde{f} * h).$$

为了介绍卷积的支集, 我们回忆函数 f 的支集这一概念是在标准意义下谈的: supp f 为使得 f 为零的最大开集的补集; 也就是说 supp f 是集合 $\{x : f(x) \neq 0\}$ 的闭包. 这个定义在函数变为函数等价类的时候其实不再适用, 而 L^p 空间中的元素不幸的是后者. 因此我们需要一个内蕴的定义, 亦即 supp f_1, supp f_2 在 $f_1 = f_2$ a.e. 时应该相同 (或相差一个零测集). 读者很容易发现之前的概念在 $f = \chi_Q$ 时其实没有意义, 因此我们通过下述命题引入一个更合适的概念:

命题 4.2

设 $f:\mathbb{R}^N\to\mathbb{R}$ 是任意函数, 考虑 $\mathbb{R}^N$ 中的开集族 $\{\omega_i\}_{i\in I}$, 它满足对每个 $i\in I$ 而言, $f=0$ 都在 ω_i 上 a.e. 记 $\omega=\bigcup_{i\in I}\omega_i$, 则 $f=0$ 在 ω 上 a.e., 记 $\operatorname{supp} f$ 为 ω 在 $\mathbb{R}^N$ 中的补集.

注

(a) 设 $f_1=f_2$ 在 $\mathbb{R}^N$ 上 a.e., 根据上述定义显见 $\operatorname{supp} f_1=\operatorname{supp} f_2$, 因此我们可以直接谈论函数 $f\in L^p$ 的支集 $\operatorname{supp} f$, 同时避开约定它所在的等价类.

(b) 若 f 是 $\mathbb{R}^N$ 上的连续函数, 则容易验证 $\operatorname{supp} f$ 的新定义与老定义得到的结果是相同的.

命题 4.3

设 $f\in L^1(\mathbb{R}^N), g\in L^p(\mathbb{R}^N)(1\le p\le\infty)$, 则

$$\operatorname{supp}(f*g)\subset\overline{\operatorname{supp} f+\operatorname{supp} g}.$$

证明 取定 $x\in\mathbb{R}^N$ 使得函数 $y\mapsto f(x-y)g(y)$ 可积[3], 知

$$(f*g)(x)=\int f(x-y)g(y)dy=\int_{(x-\operatorname{supp} f)\cap\operatorname{supp} g} f(x-y)g(y)dy.$$

若 $x\notin\operatorname{supp} f+\operatorname{supp} g$, 则 $(x-\operatorname{supp} f)\cap\operatorname{supp} g=\varnothing$, 于是 $(f*g)(x)=0$, 因此

$$(f*g)(x)=0\quad 在\ (\operatorname{supp} f+\operatorname{supp} g)^c\ 上\ \text{a.e.}$$

特别有

$$(f*g)(x)=0\quad 在\ \operatorname{Int}[(\operatorname{supp} f+\operatorname{supp} g)^c]\ 上\ \text{a.e.}$$

因此

$$\operatorname{supp}(f*g)\subset[\operatorname{Int}(\operatorname{supp} f+\operatorname{supp} g)^c]^c=\overline{\operatorname{supp} f+\operatorname{supp} g}.$$

注 若 f,g 均紧支, 则 $f*g$ 也紧支. 然而, 如果 f,g 只有一个紧支, $f*g$ 就未必紧支.

定义 4.4 (L^p_{loc})

设 $\Omega\subset\mathbb{R}^N$ 是开集, $1\le p\le\infty$. 若函数 $f:\Omega\to\mathbb{R}$ 满足对 Ω 中的每个紧集 K 均有 $f\chi_K\in L^p(\Omega)$, 就称 f 是局部 p 次可积的, 记 $f\in L^p_{loc}(\Omega)$.

显见 $f\in L^p_{loc}(\Omega)\Rightarrow f\in L^1_{loc}(\Omega)$.

命题 4.4

设 $f\in C_c(\mathbb{R}^N), g\in L^1_{loc}(\mathbb{R}^N)$, 则 $(f*g)(x)$ 对每个 $x\in\mathbb{R}^N$ 均良定义, 且 $f*g\in C(\mathbb{R}^N)$.

证明 注意到对每个 $x\in\mathbb{R}^N$ 而言, 函数 $y\mapsto f(x-y)g(y)$ 在 $\mathbb{R}^N$ 上可积, 因此 $(f*g)(x)$ 对每个 $x\in\mathbb{R}^N$ 均有定义.

现设 $x_n\to x$, 取定 $\mathbb{R}^N$ 中的紧集 K 使得 $(x_n-\operatorname{supp} f)\subset K(\forall n)$, 由此可知 $f(x_n-y)=0(\forall n,\forall y\notin K)$. 由 f 的一致连续性知

$$|f(x_n-y)-f(x-y)|\le\varepsilon_n\chi_K(y),\quad\forall n,\forall y\in\mathbb{R}^N,\varepsilon_n\to 0.$$

由此即知

$$|(f*g)(x_n)-(f*g)(x)|\le\varepsilon_n\int_K|g(y)|dy\to 0.$$

□

另外, 卷积可以提升原有函数的正则性. 为此我们先约定一些记号. 设 $\Omega\subset\mathbb{R}^N$ 是开集, 则:

- $C(\Omega)$ 表示 Ω 上全体连续函数构成的空间.

[3]需要强调取定 x 是因为仅仅有 a.e. $x\in\mathbb{R}^N$ 使得 $f*g$ 可积.

- $C^k(\Omega)$ 表示 Ω 上全体 k 阶连续可微函数构成的空间, 其中 $k \geq 1$ 是整数.
- $C^\infty(\Omega) = \bigcap_k C^k(\Omega)$.
- $C_c(\Omega)$ 表示 Ω 上全体在 Ω 中紧支的连续函数构成的空间, 亦即该空间中的函数都能对应存在某个紧集 $K \subset \Omega$ 使得函数在 K 之外为零.
- $C_c^k(\Omega) = C^k(\Omega) \cap C_c(\Omega)$.
- $C_c^\infty(\Omega) = C^\infty(\Omega) \cap C_c(\Omega)$(一些作者会把 $C_c^\infty(\Omega)$ 写成 $\mathcal{D}(\Omega)$ 或 $C_0^\infty(\Omega)$).

若 $f \in C^1(\Omega)$, 则其梯度定义为

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_N}\right).$$

若 $f \in C^k(\Omega)$, $\alpha = (\alpha_1, \cdots, \alpha_N)$ 是长为 $|\alpha| = \alpha_1 + \cdots + \alpha_N \leq k$ 的多重指标, 则

$$D^\alpha f := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} f.$$

命题 4.5 (卷积对正则性的提升)

若 $f \in C_c^k(\mathbb{R}^N)(k \geq 1), g \in L_{loc}^1(\mathbb{R}^N)$, 则 $f * g \in C^k(\mathbb{R}^N)$, 且

$$D^\alpha(f * g) = (D^\alpha f) * g, \quad \forall |\alpha| \leq k.$$

特别若 $f \in C_c^\infty(\mathbb{R}^N), g \in L_{loc}^1(\mathbb{R}^N)$, 则 $f * g \in C^\infty(\mathbb{R}^N)$.

♠

4.5 磨光子

定义 4.5 (磨光子 (mollifier))

磨光子序列 $\{\rho_n\}_{n\geq 1}$ 指的是 $\mathbb{R}^N$ 上的一个函数列, 其满足

$$\rho_n \in C_c^\infty(\mathbb{R}^N), \operatorname{supp}\rho_n \subset \overline{B(0, 1/n)}, \int \rho_n = 1, \text{在 } \mathbb{R}^N \text{ 上} \rho_n \geq 0.$$

♣

之后我们统一用 $\{\rho_n\}$ 来表示磨光子序列.

从单个函数 $\rho \in C_c^\infty(\mathbb{R}^N)$ 出发, 只要它满足 $\operatorname{supp}\rho \subset \overline{B(0,1)}$, 在 $\mathbb{R}^N$ 上 $\rho \geq 0$, 且 ρ 不恒为 0, 就可以基于 ρ 构造一个磨光子序列出来. 例如取

$$\rho(x) = \begin{cases} e^{1/(|x|^2-1)}, & |x| < 1 \\ 0, & |x| > 1 \end{cases}$$

令 $\rho_n(x) = Cn^N\rho(nx), C = 1/\int\rho$ 即得磨光子序列 $\{\rho_n\}$.

磨光子最重要的用途在于其逼近性质:

命题 4.6 (磨光子的一致逼近)

设 $f \in C(\mathbb{R}^N)$, 则 $(\rho_n * f) \to f(n \to \infty)$ 在 $\mathbb{R}^N$ 的紧集上一致成立.

♠

证明 设 $K \subset \mathbb{R}^N$ 是取定的紧集, 给定 $\varepsilon > 0$, 根据 f 的连续性知存在 (依赖于 K 和 ε 的)$\delta > 0$ 使得

$$|f(x-y) - f(x)| < \varepsilon, \quad \forall x \in K, \forall y \in B(0,\delta).$$

于是对 $x \in \mathbb{R}^N$ 有

$$\begin{aligned}(\rho_n * f)(x) - f(x) &= \int [f(x-y) - f(x)]\rho_n(y)dy \\ &= \int_{B(0,1/n)} [f(x-y) - f(x)]\rho_n(y)dy.\end{aligned}$$

当 $n > 1/\delta, x \in K$ 时有

$$|(\rho_n * f)(x) - f(x)| \le \varepsilon \int \rho_n = \varepsilon.$$

□

定理 4.16 (磨光子的 L^p 逼近)

设 $f \in L^p(\mathbb{R}^N)(1 \le p < \infty)$, 则 $(\rho_n * f) \to f(n \to \infty)$ 在 $L^p(\mathbb{R}^N)$ 意义下成立.

推论 4.1 ($C_c^\infty(\Omega)$ 在 $L^p(\Omega)$ 中的稠密性)

设 $\Omega \subset \mathbb{R}^N$ 是开集, 则 $C_c^\infty(\Omega)$ 在 $L^p(\Omega)(1 \le p < \infty)$ 中稠密.

推论 4.2

设 $\Omega \subset \mathbb{R}^N$ 是开集, $u \in L^1_{loc}(\Omega)$ 满足

$$\int uf = 0, \quad \forall f \in C_c^\infty(\Omega).$$

则 $u = 0$ 在 Ω 上 a.e.

4.6 L^p 中的强紧性判别法

确定 $L^p(\Omega)$ 中怎样的函数类具有 (强拓扑下的) 紧闭包这件事非常重要. 对于 $C(K)$(其中 K 是紧度量空间) 而言, 这个问题已经由 Ascoli-Arzela 定理给出了解答:

定理 4.17 (Ascoli-Arzela)

设 K 是紧度量空间, $\mathcal{H}$ 是 $C(K)$ 中的有界集. 若 $\mathcal{H}$ 是等度连续的, 即

$$\forall \varepsilon > 0 \exists \delta > 0 \forall f \in \mathcal{H}(d(x_1, x_2) < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon), \tag{4.12}$$

则 $\mathcal{H}$ 在 $C(K)$ 中的闭包是紧的.

现在记 $(\tau_h f)(x) = f(x+h)(x \in \mathbb{R}^N, h \in \mathbb{R}^N)$ 是函数 f 的平移, Ascoli-Arzela 定理在 L^p 空间中有下述版本:

定理 4.18 (Kolmogorov-Riesz-Frechet)

设 $\mathcal{F}$ 是 $L^p(\mathbb{R}^N)(1 \le p < \infty)$ 中的有界集, 设[a]

$$\lim_{|h| \to 0} \|\tau_h f - f\|_p = 0 \quad \text{关于 } f \in \mathcal{F} \text{ 一致} \tag{4.13}$$

亦即对 $\forall \varepsilon > 0$ 均 $\exists \delta > 0$ 使得 $\|\tau_h f - f\|_p < \varepsilon$ 对 $\forall f \in \mathcal{F}$ 与满足 $|h| < \delta$ 的 $\forall h \in \mathbb{R}^N$ 成立, 则只要 $\Omega \subset \mathbb{R}^N$ 是任意有限测度集, $F|_\Omega$ 在 $L^p(\Omega)$ 中的闭包就是紧的. (其中 $F|_\Omega$ 表示 $\mathcal{F}$ 中函数在 Ω 上的限制构成的空间.)

[a]这个条件对应(4.12)式, 它可以看成" 积分的" 等度连续条件.

证明 证明分为下述四步.

$\boxed{\text{第一步: 将 } f \text{ 磨光.}}$ 我们断言

$$\|(\rho_n * f) - f\|_{L^p(\mathbb{R}^N)} \le \varepsilon, \quad \forall f \in \mathcal{F}, \forall n > \frac{1}{\delta}. \tag{4.14}$$

这是因为根据 Hölder 不等式有

$$\begin{aligned} |(\rho_n * f)(x) - f(x)| &\le \int |f(x-y) - f(x)| \rho_n(y) dy \\ &\le \left[\int |f(x-y) - f(x)|^p \rho_n(y) dy \right]^{\frac{1}{p}}. \end{aligned}$$

于是只要 $1/n < \delta$, 就有

$$\int |(\rho_n * f)(x) - f(x)|^p dx \le \int\int |f(x-y) - f(x)|^p \rho_n(y) dx dy$$
$$= \int_{B(0,1/n)} \rho_n(y) dy \int |f(x-y) - f(x)|^p dx \le \varepsilon^p.$$

$\boxed{\text{第二步: } f \text{ 磨光后所得函数的一致有界与等度连续性}}$ 我们断言

$$\|\rho_n * f\|_{L^\infty(\mathbb{R}^N)} \le C_n \|f\|_{L^p(\mathbb{R}^N)}, \quad \forall f \in \mathcal{F} \tag{4.15}$$

且

$$|(\rho_n * f)(x_1) - (\rho_n * f)(x_2)| \le C_n \|f\|_p |x_1 - x_2|, \quad \forall f \in \mathcal{F}, \forall x_1, x_2 \in \mathbb{R}^N, \tag{4.16}$$

其中 C_n 仅依赖于 n.

注意到 $\rho_n * f$ 本质上是一个积分, 利用 Hölder 不等式并令 $C_n = \|\rho_n\|_{p'}$ 即证(4.15)式. 对于(4.16)式, 注意到 $\nabla(\rho_n * f) = (\nabla \rho_n) * f$, 于是

$$\|\nabla(\rho_n * f)\|_{L^\infty(\mathbb{R}^N)} \le \|\nabla \rho_n\|_{L^{p'}(\mathbb{R}^N)} \|f\|_{L^p(\mathbb{R}^N)},$$

令 $C_n = \|\nabla \rho_n\|_{L^{p'}(\mathbb{R}^N)}$, 由 Lagrange 中值定理即得(4.16)式.

$\boxed{\text{第三步: 取出 } \Omega \text{ 中适用于 Ascoli-Arzela 定理的紧子集.}}$ 取定 $\varepsilon > 0, \Omega \subset \mathbb{R}^N$ 测度有限, 我们断言存在 Ω 的有界可测子集使得

$$\|f\|_{L^p(\Omega \backslash \omega)} < \varepsilon, \quad \forall f \in \mathcal{F} \tag{4.17}$$

这是因为

$$\|f\|_{L^p(\Omega \backslash \omega)} \le \|f - (\rho_n * f)\|_{L^p(\mathbb{R}^N)} + \|\rho_n * f\|_{L^p(\Omega \backslash \omega)}.$$

上右式前一项由第一步已经得证是充分小的了. 对于上右式第二项, 由(4.15)式知只需取 ω 使得 $|\Omega \backslash \omega|$ 足够小即可.

$\boxed{\text{第四步: 证明结论.}}$ $L^p(\Omega)$ 首先可度量, 因此 $\overline{\mathcal{F}|_\Omega}$ 紧当且仅当其自列紧. 因为度量空间的子集列紧当且仅当其闭包自列紧, 故要说明 $\overline{F|_\Omega}$ 紧, 只需说明 $\mathcal{F}|_\Omega$ 列紧. 因为 $L^p(\Omega)$ 完备, 根据 Hausdorff 定理可知 $\mathcal{F}|_\Omega$ 列紧当且仅当它是完全有界集, 亦即对任意 $\varepsilon > 0$ 而言均可以用半径为 ε 的球有限覆盖 $\mathcal{F}|_\Omega$. 现在取定 $\varepsilon > 0$, 根据第三步知存在有界可测子集 ω 使得(4.17)式成立. 再取定 $n > 1/\delta$, 根据第二步可知函数族 $\mathcal{H} = (\rho_n * \mathcal{F})|_{\overline{\omega}}$ 满足 Ascoli-Arzela 定理的条件, 因此 $\mathcal{H}$ 在 $C(\overline{\omega})$ 中有紧闭包. 因为 $C(\overline{\omega})$ 上自然的拓扑肯定强于 $L^p(\omega)$ 上的拓扑, 而 $C(\overline{\omega})$ 本身是 $L^p(\omega)$ 中的闭线性子空间, 故 $\mathcal{H}$ 在 $L^p(\omega)$ 中同样具有紧闭包. 根据 $\overline{\mathcal{H}}$ 在 $L^p(\omega)$ 中的自列紧性与 Hausdorff 定理, 知 $\overline{\mathcal{H}}$ 应在 $L^p(\omega)$ 中完全有界, 于是存在半径为 ε 的有限个球 $B(g_i, \varepsilon)$ 覆盖 $\overline{\mathcal{H}}$. 特别这些球也覆盖 $\mathcal{H}$, 即

$$\mathcal{H} \subset \bigcup_i B(g_i, \varepsilon), \quad g_i \in L^p(\omega).$$

现在定义

$$\overline{g_i} : \Omega \to \mathbb{R}, \quad x \mapsto \begin{cases} g_i(x), & x \in \omega \\ 0, & x \in \Omega \backslash \omega \end{cases}$$

考虑 $L^p(\Omega)$ 中的球 $B(\overline{g_i}, 3\varepsilon)$. 我们接下来就说明这族球覆盖 $\mathcal{F}|_\Omega$. 事实上, 任取 $f \in \mathcal{F}$, 根据 $\{B(g_i, \varepsilon)\}$ 覆盖 $\mathcal{H}$ 知存在 i 使得

$$\|(\rho_n * f) - g_i\|_{L^p(\omega)} < \varepsilon.$$

因为

$$\|f - \overline{g_i}\|^p_{L^p(\Omega)} = \int_{\Omega \backslash \omega} |f|^p + \int_\omega |f - g_i|^p$$

由(4.17)式知

$$\|f - \overline{g_i}\|_{L^p(\Omega)} \le \varepsilon + \|f - g_i\|_{L^p(\omega)}$$
$$\le \varepsilon + \|f - (\rho_n * f)\|_{L^p(\mathbb{R}^N)} + \|(\rho_n * f) - g_i\|_{L^p(\omega)} < 3\varepsilon$$

因此 $\mathcal{F}|_\Omega$ 被 $\{B(\overline{g_i}, 3\varepsilon)\}$ 覆盖, 故 $\mathcal{F}|_\Omega$ 在 $L^p(\Omega)$ 中具有紧闭包. □

如果只有定理4.18的条件, 我们一般没法断言 $\mathcal{F}$ 自身在 $L^p(\mathbb{R}^N)$ 中有紧闭包. 要说明这件事, 我们需要添一个条件, 此即下述推论:

推论 4.3

设 $\mathcal{F}$ 是 $L^p(\mathbb{R}^N)(1 \le p < \infty)$ 中的有界集, (4.13)式成立, 且

$$\forall \varepsilon > 0 \exists \Omega \subset \mathbb{R}^N (\Omega \text{ 有界, 可测, 且} \|f\|_{L^p(\mathbb{R}^N\backslash\Omega)} < \varepsilon (\forall f \in \mathcal{F})), \tag{4.18}$$

则 $\mathcal{F}$ 在 $L^p(\mathbb{R}^N)$ 中有紧闭包.

证明 取定 $\varepsilon > 0$, 设 $\Omega \subset \mathbb{R}^N$ 是满足(4.18)式的有界可测集. 根据 Kolmogorov-Riesz-Frechet 定理4.18已知 $\mathcal{F}|_\Omega$ 在 $L^p(\Omega)$ 中有紧闭包, 因此我们可以在 $L^p(\Omega)$ 中用有限个半径为 ε 的球覆盖 $\mathcal{F}|_\Omega$. 比方说

$$\mathcal{F}|_\Omega \subset \bigcup_i B(g_i, \varepsilon), \quad g_i \in L^p(\Omega).$$

记

$$\overline{g_i}(x) = \begin{cases} g_i(x), & x \in \Omega \\ 0, & x \in \mathbb{R}^N \backslash \Omega \end{cases}$$

则可以证明 $\mathcal{F}$ 在 $L^p(\mathbb{R}^N)$ 中被 $\{B(\overline{g_i}, 2\varepsilon)\}$ 覆盖. □

注 上述推论的逆命题也成立, 因此上述推论实际上给出了 $L^p(\mathbb{R}^N)$ 中紧集的一个完全刻画.

下面给出 Kolmogorov-Riesz-Frechet 定理的一个有用应用:

推论 4.4

设 G 是 $L^1(\mathbb{R}^N)$ 中取定的函数, 记

$$\mathcal{F} = G * \mathcal{B},$$

其中 $\mathcal{B}$ 是 $L^p(\mathbb{R}^N)(1 \le p < \infty)$ 中的有界集, 则对任意有限测度集 Ω 而言, $F|_\Omega$ 在 $L^p(\Omega)$ 中均有紧闭包.

证明 显见 $\mathcal{F}$ 在 $L^p(\mathbb{R}^N)$ 中有界. 另一方面, 若记 $f = G * u (u \in \mathcal{B})$, 根据 $\mathcal{B}$ 的有界性知

$$\|\tau_h f - f\|_p = \|(\tau_h G - G) * u\|_p \le C\|\tau_h G - G\|_1,$$

现在只要能证明下述引理:

引理 4.3

若 $G \in L^q(\mathbb{R}^N)(1 \le q < \infty)$, 则 $\lim_{j\to\infty} \|\tau_h G - G\|_q = 0$.

根据 Kolmogorov-Riesz-Frechet 定理4.18就立得结论. 下面说明上述引理. 取定 $\varepsilon > 0$, 根据 $C_c(\mathbb{R}^N)$ 在 $L^q(\mathbb{R}^N)$ 中的稠密性4.3知存在 $G_1 \in C_c(\mathbb{R}^N)$ 使得 $\|G - G_1\|_q < \varepsilon$. 于是

$$\begin{aligned} \|\tau_h G - G\|_q &\le \|\tau_h G - \tau_h G_1\|_q + \|\tau_h G_1 - G_1\|_q + \|G_1 - G\|_q \\ &\le 2\varepsilon + \|\tau_h G_1 - G_1\|_q. \end{aligned}$$

因为 $\lim_{h\to 0} \|\tau_h G_1 - G_1\|_q = 0$, 故

$$\limsup_{h\to 0} \|\tau_h G - G\|_q \le 2\varepsilon, \quad \forall \varepsilon > 0.$$

引理至此即证. □

第五章 Hilbert 空间

本章选自 [**HB**].

5.1 定义, 初等性质, 到闭凸集上的投影

定义 5.1 (内积 (scalar product))

设 H 是线性空间, 则内积 (u,v) 是 $H\times H\to\mathbb{R}$ 的双线性型 (即关于两个变量都线性的映射), 它满足

$$\begin{aligned}&\text{对称性:}\quad (u,v)=(v,u),\forall u,v\in H\\&\text{非负性:}\quad (u,u)\geq 0,\forall u\in H\\&\text{正定性:}\quad (u,u)\neq 0,\forall u\neq 0\end{aligned}$$

♣

回忆内积满足 Cauchy-Schwarz 不等式:

$$|(u,v)|\leq (u,u)^{\frac{1}{2}}(v,v)^{\frac{1}{2}},\quad \forall u,v\in H.$$

(注意 Cauchy-Schwarz 不等式的证明不需要用到 $(u,u)\neq 0(\forall u\neq 0)$.) 由 Cauchy-Schwarz 不等式立得

$$|u|=(u,u)^{\frac{1}{2}}$$

是一个范数. 一般用 $|\cdot|$(而非 $\|\cdot\|$) 来记内积诱导的范数.

回忆经典的平行四边形法则:

$$\left|\frac{a+b}{2}\right|^2+\left|\frac{a-b}{2}\right|^2=\frac{1}{2}(|a|^2+|b|^2),\quad \forall a,b\in H. \tag{5.1}$$

定义 5.2 (Hilbert 空间 (Hilbert space))

Hilbert 空间指的是装备了内积的线性空间 H, 且 H 关于范数 $|\cdot|$ 完备.

♣

之后统一用 H 表示 Hilbert 空间.

例 5.1 在 $L^2(\Omega)$ 上装备内积

$$(u,v)=\int_\Omega u(x)v(x)d\mu$$

后, $L^2(\omega)$ 成为 Hilbert 空间. 特别地, l^2 也是 Hilbert 空间. 之后研究的 Sobolev 空间 H^1 也是 Hilbert 空间.

命题 5.1

H 是一致凸的, 因此它自反.

♠

证明 取 $\varepsilon>0, u,v\in H$ 满足 $|u|\leq 1,|v|\leq 1,|u-v|>\varepsilon$, 根据平行四边形法则可知

$$\left|\frac{u+v}{2}\right|^2<1-\frac{\varepsilon^2}{4}\Rightarrow\left|\frac{u+v}{2}\right|<1-\delta,\delta=1-\left(1-\frac{\varepsilon^2}{4}\right)^{\frac{1}{2}}>0.$$

□

定理 5.1 (到闭凸集上的投影)

设 $K\subset H$ 是非空闭凸集, 则对每个 $f\in H$ 而言均存在唯一 $u\in K$ 使得

$$|f-u|=\min_{v\in K}|f-v|=\operatorname{dist}(f,K). \tag{5.2}$$

且 u 由下述性质刻画:

$$u\in K,\quad (f-u,v-u)\leq 0,\forall v\in K. \tag{5.3}$$

♡

注 上述定理中断言存在的 u 称为 f 到 K 上的**投影 (projection)**, 记之为

$$u = P_K f.$$

不等式(5.3)表明向量 $\overrightarrow{uf}$(以 u 为起点, 以 f 为终点) 与任意向量 $\overrightarrow{uv}(v \in K)$ 之间的内积均 ≤ 0, 也就是说这两个向量之间的夹角 $\theta \geq \pi/2$(见图5.1).

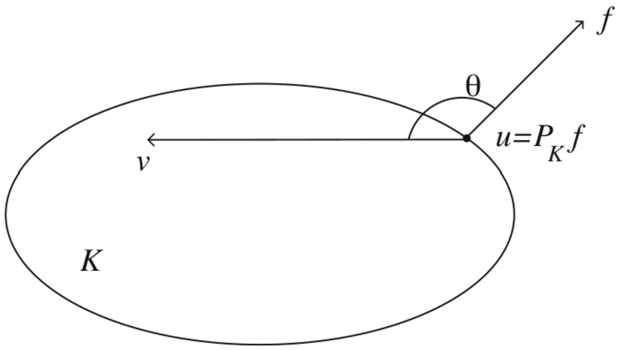

图 5.1

证明 $\boxed{\text{存在性.}}$ 我们对存在性给出下述两种证明:

1. 函数 $\varphi(v) = |f - v|$ 是凸函数, φ 连续, 且 $\lim_{|v|\to\infty} \varphi(v) = +\infty$. 因此由 H 的自反性与自反空间下半连续凸函数的最小值3.6知 φ 在 K 上可以达到最小值.
2. 第二种证明不依赖于自反理论与一致凸空间, 该证明是直接进行的. 设 $\{v_n\}$ 是(5.2)式的最小化序列, 即 $v_n \in K$ 且

$$d_n = |f - v_n| \to d = \inf_{v\in K} |f - v|.$$

下面说明 $\{v_n\}$ 是基本列. 事实上, 对 $a = f - v_n, b = f - v_m$ 应用平行四边形法则可得

$$\left|f - \frac{v_n + v_m}{2}\right|^2 + \left|\frac{v_n - v_m}{2}\right|^2 = \frac{1}{2}(d_n^2 + d_m^2).$$

因为 K 是凸的, 故 $(v_n + v_m)/2 \in K$, 因此 $|f - (v_n + v_m)/2| \geq d$, 于是

$$\left|\frac{v_n - v_m}{2}\right|^2 \leq \frac{1}{2}(d_n^2 + d_m^2) - d^2 \Rightarrow \lim_{m,n\to\infty} |v_n - v_m| = 0.$$

因此序列 $\{v_n\}$ 收敛到 $u \in K$, 且 $d = |f - u|$.

$\boxed{\text{(5.2),(5.3)的等价.}}$ 设 $u \in K$ 满足(5.2)式, 记 $w \in K$, 根据 K 的凸性知

$$v = (1 - t)u + tw \in K, \quad \forall t \in [0, 1]$$

因此

$$|f - u| \leq |f - [(1 - t)u + tw]| = |(f - u) - t(w - u)|.$$

于是

$$|f - u|^2 \leq |f - u|^2 - 2t(f - u, w - u) + t^2|w - u|^2,$$

这说明 $2(f - u, w - u) \leq t|w - u|^2(\forall t \in (0, 1])$, 令 $t \to 0$ 即得(5.3)式.

反之, 设 u 满足(5.3)式, 知

$$|u - f|^2 - |v - f|^2 = 2(f - u, v - u) - |u - v|^2 \leq 0, \quad \forall v \in K$$

此即(5.2)式.

$\boxed{\text{唯一性.}}$ 设 u_1, u_2 满足(5.3)式, 则

$$(f - u_1, v - u_1) \leq 0, \quad \forall v \in K \tag{5.4}$$

$$(f - u_2, v - u_2) \leq 0, \quad \forall v \in K \tag{5.5}$$

在(5.4)式中取 $v = u_2$, 在(5.5)式中取 $v = u_1$ 并将得到的不等式相加即得 $|u_1 - u_2|^2 \leq 0$. □

注 不难发现最小化问题总是会和不等式组联系起来. 这里回忆一个广为人知的例子. 设 $F:\mathbb{R}\to\mathbb{R}$ 是可微函数, $u\in[0,1]$ 是 F 在 $[0,1]$ 上的最小值点, 则要么 $u\in(0,1)$ 且 $F'(u)=0$, 要么 $u=0, F'(u)\geq 0$, 要么 $u=1, F'(u)\leq 0$, 这三种情况都能划归为 $u\in[0,1]$ 满足 $F'(u)(v-u)\geq 0(\forall v\in[0,1])$.

注 定理5.1的结论对一致凸 Banach 空间也成立. 也就是说, 设 $K\subset E$ 是非空闭凸集, 其中 E 是一致凸 Banach 空间, 则对每个 $f\in E$ 而言存在唯一 $u\in E$ 使得

$$\|f-u\|=\min_{v\in K}\|f-v\|=\operatorname{dist}(f,K).$$

命题 5.2

设 $K\subset H$ 是非空闭凸集, 则 P_K 不可能增加距离, 即

$$|P_K f_1-P_K f_2|\leq|f_1-f_2|,\quad \forall f_1,f_2\in H.$$

证明 设 $u_1=P_K f_1, u_2=P_K f_2$, 则

$$(f_1-u_1,v-u_1)\leq 0,\quad \forall v\in K \tag{5.6}$$

$$(f_2-u_2,v-u_2)\leq 0,\quad \forall v\in K \tag{5.7}$$

在(5.6)式中代入 $v=u_2$, 在(5.7)式中代入 $v=u_1$, 将得到的不等式相加可知

$$|u_1-u_2|^2\leq(f_1-f_2,u_1-u_2).$$

故 $|u_1-u_2|\leq|f_1-f_2|$. □

推论 5.1 (正交投影定理)

设 $M\subset H$ 是闭线性子空间, 若 $f\in H$, 则 $u=P_M f$ 能被刻画为

$$u\in M,\quad (f-u,v)=0,\forall v\in M. \tag{5.8}$$

进一步, P_M 是线性算子, 称之为正交投影.

证明 由(5.3)式知

$$(f-u,v-u)\leq 0,\quad \forall v\in M$$

因此

$$(f-u,tv-u)\leq 0,\quad \forall v\in M,\forall t\in\mathbb{R}$$

由此即得(5.8)式.

反之, 若 u 满足(5.8)式, 则

$$(f-u,v-u)=0,\quad \forall v\in M$$

由(5.3)式即知 $u=P_M f$, 另外由此显见 P_M 是线性的. □

5.2 Hilbert 空间的对偶空间

在 Hilbert 空间中, 连续线性泛函的形式非常简单. 任取 $f\in H$, 则映射 $u\mapsto(f,u)$ 就是 H 上的连续线性泛函. Hilbert 空间中的一个重要结论就在于其上的全体连续线性泛函都是这种形式:

定理 5.2 (Riesz-Frechet 表示定理)

对任意给定的 $\varphi\in H^*$, 总存在唯一 $f\in H$ 使得

$$\langle\varphi,u\rangle=(f,u),\quad \forall u\in H$$

且

$$|f| = \|\varphi\|_{H^*}.$$

♡

证明 我们给出两种证明:

1. 第一个证明和 Riesz 表示定理4.11的证明基本完全一样. 考虑依照下述方法定义的映射 $T: H \to H^*$: 对任意给定的 $f \in H$, 映射 $u \mapsto (f, u)$ 是 H 上的连续线性泛函, 因此该映射是 H^* 中的元素, 记之为 Tf, 即

$$\langle Tf, u\rangle = (f, u), \quad \forall u \in H.$$

显见 $\|Tf\|_{H^*} = |f|$, 故 T 是 H 到 $T(H)$ 上的线性等距同构, 且 $T(H)$ 是 H^* 的一个闭子空间. 为说明 $T(H) = H^*$, 只需说明 $T(H)$ 在 H^* 中稠密, 为此考虑推论2.4, 设 h 是 H^* 上的连续线性泛函, 且 h 在 $T(H)$ 上为零. 因为 H 自反, 故 $h \in H$, 于是 $\langle h, Tf\rangle = \langle Tf, h\rangle = 0 (\forall f \in H)$. 这说明 $(f, h) = 0 (\forall f \in H)$, 因此只能有 $h = 0$.

2. 第二个证明更直接, 它绕开了 H 的自反性. 设 $M = \varphi^{-1}(\{0\})$, 则 M 是 H 的闭子空间, 不妨设 $M \neq H$(否则 $\varphi \equiv 0$, 定理5.2的结论此时显然成立). 我们断言存在 $g \in H$ 使得

$$|g| = 1, \quad (g, v) = 0, \forall v \in M (\text{因此 } g \notin M).$$

事实上, 取 $g_0 \in H$ 满足 $g_0 \notin M$, 令 $g_1 = P_M g_0$, 则

$$g = \frac{g_0 - g_1}{|g_0 - g_1|}$$

就满足我们需要的性质.

现在任意取定 $u \in H$, 记

$$v = u - \lambda g, \quad \lambda = \frac{\langle \varphi, u\rangle}{\langle \varphi, g\rangle}.$$

因为 $g \notin M$, 故 $\langle \varphi, g\rangle \neq 0$, 因此 v 良定义. 又因为 $\langle \varphi, v\rangle = 0$, 故 $v \in M$, 从而 $(g, v) = 0$, 亦即

$$\langle \varphi, u\rangle = \langle \varphi, g\rangle (g, u), \quad \forall u \in H$$

取 $f = \langle \varphi, g\rangle g$ 即可. □

注 **H 与 H^* 是否可以等同?** Riesz-Frechet 表示定理5.2断言存在 H 到 H^* 上的典则等距同构, 因此把 H 和 H^* 等同的这个想法看似是" 符合逻辑" 的. 我们经常会这么干, 但不总是如此. 下面我们给出一个典型例子. 设 H 是赋内积 $(\cdot, \cdot)$ 与对应范数 $|\cdot|$ 的 Hilbert 空间, $V \subset H$ 是在 H 中稠密的线性子空间. 设 V 自身有范数 $\|\cdot\|$, 且 V 关于范数 $\|\cdot\|$ 成为 Banach 空间. 设包含映射 $V \subset H$ 是连续的, 即

$$|v| \le C\|v\|, \quad \forall v \in V.$$

(例如取 $H = L^2(0,1), V = L^p(0,1)(p > 2)$ 或 $V = C[0,1]$.)

通过把 H 上的连续线性泛函 φ 限制在 V 上即可得到典则映射 $T: H^* \to V^*$, 即

$$\langle T\varphi, v\rangle_{V^*, V} = \langle \varphi, v\rangle_{H^*, H}.$$

显见 T 有下述性质:

(i) $\|T\varphi\|_{V^*} \le C|\varphi|_{H^*} (\forall \varphi \in H^*)$,

(ii) T 是单射,

(iii) 若 V 自反, 则 $R(T)$ 在 V^* 中稠密[1].

现在如果把 H^* 和 H 等同, 把 T 看成 H^* 到 V^* 的典则嵌入, 则

$$V \subset H \simeq H^* \subset V^*, \tag{5.9}$$

其中出现的所有包含映射都连续, 且若 V 自反, 则这些包含映射还稠密. 此时称 H 为**基准空间 (pivot space)**. 注意到此时只要序偶 $\langle \cdot, \cdot\rangle_{V^*, V}$ 与内积 $(\cdot, \cdot)$ 均有意义, 那它们就是相同的, 即

$$\langle f, v\rangle_{V^*, V} = (f, v), \quad \forall f \in H, \forall v \in V.$$

[1]不过 T 一般不是满射.

然而, 当 V 成为 Hilbert 空间, 其上赋有它本身的内积 $((\cdot,\cdot))$ 与该内积诱导的范数 $\|\cdot\|$ 时, 情况就更复杂了. 此时我们当然可以通过 $((\cdot,\cdot))$ 把 V 和 V^* 等同起来, 但因为一般来说 $V \neq H$, 故此时(5.9)式不能再成立了. 这说明我们不能同时把 V, H 与它们的对偶空间等同: 我们必须做一个取舍. 通常的习惯是等同 H^* 与 H, 让(5.9)式成立, 而不把 V^* 和 V 等同起来. (此时从 V 到 V^* 上当然依旧还有等距映射, 但此时我们不把它看成恒同映射.) 参见下例:

记

$$H = l^2 = \left\{ u = \{u_n\}_{n\geq 1} : \sum_{n=1}^{\infty} u_n^2 < \infty \right\}$$

其上赋内积 $(u, v) = \sum_{n=1}^{\infty} u_n v_n$. 记

$$V = \left\{ u = \{u_n\}_{n\geq 1} : \sum_{n=1}^{\infty} n^2 u_n^2 < \infty \right\}$$

其上赋内积 $((u, v)) = \sum_{n=1}^{\infty} n^2 u_n v_n$.

显见包含映射 $V \subset H$ 是连续单射, 且 V 在 H 中稠密. 当我们等同 H^* 和 H 时, V^* 就等同于下述空间

$$V^* = \left\{ f = \{f_n\}_{n\geq 1} : \sum_{n=1}^{\infty} \frac{1}{n^2} f_n^2 < \infty \right\},$$

该空间严格大于 H. 序偶 $\langle \cdot,\cdot \rangle_{V^*,V}$ 此时为

$$\langle f, v \rangle_{V^*,V} = \sum_{n=1}^{\infty} f_n v_n,$$

而 Riesz-Frechet 同构 $T : V \to V^*$ 此时为

$$u = \{u_n\}_{n\geq 1} \mapsto Tu = \{n^2 u_n\}_{n\geq 1}.$$

这显然不再是恒同算子.

注 绕过一致凸空间理论也可以很容易地说明 Hilbert 空间自反, 用两次 Riesz-Frechet 同构即可 (从 H 到 H^* 上, 再从 H^* 到 H^{**} 上).

注 设 H 是与其对偶空间 H^* 等同的 Hilbert 空间, M 是 H 的子空间. 前面我们定义过作为 H^* 子空间的空间 $M^\perp$. 因为 H^* 和 H 恒同, 故此时 $M^\perp$ 可以看成 H 的子空间, 即

$$M^\perp = \{u \in H : (u, v) = 0, \forall v \in M\}.$$

显见 $M \cap M^\perp = \{0\}$. 另若 M 是闭子空间, 则 $M + M^\perp = H$. 这是因为每个 $f \in H$ 都能写成

$$f = (P_M f) + (f - P_M f)$$

可以验证 $f - P_M f \in M^\perp$, 且 $f - P_M f = P_{M^\perp} f$. 由此可知 Hilbert 空间中的任何闭子空间都有余子空间.

5.3 Stampacchia 定理与 Lax-Milgram 定理

定义 5.3

对双线性型 $a : H \times H \to \mathbb{R}$ 而言:

(i) 若存在常数 C 使得

$$|a(u, v)| \leq C|u||v|, \quad \forall u, v \in H$$

就称 a 连续.

(ii) 若存在常数 $\alpha > 0$ 使得

$$a(v, v) \geq \alpha |v|^2, \quad \forall v \in H$$

就称 a 强正 (coercive). ♣

定理 5.3 (Stampacchia)

设 $a(u,v)$ 是 H 上的连续强正双线性型, $K \subset H$ 是非空闭凸子集, 则对任意给定的 $\varphi \in H^*$, 均存在唯一 $u \in K$ 使得

$$a(u, v-u) \geq \langle \varphi, v-u \rangle, \quad \forall v \in K. \tag{5.10}$$

另若 a 对称, 则 u 能被下述性质刻画:

$$u \in K, \text{ 且 } \frac{1}{2}a(u,u) - \langle \varphi, u \rangle = \min_{v \in K} \left\{ \frac{1}{2}a(v,v) - \langle \varphi, v \rangle \right\}. \tag{5.11}$$

♡

Stampacchia 定理的证明依赖于下述非常经典的结果:

定理 5.4 (Banach 不动点定理-压缩映射原理)

设 X 是非空完备度量空间, $S: X \to X$ 是严格压缩映射, 即

$$d(Sv_1, Sv_2) \leq k d(v_1, v_2), \quad k < 1, \forall v_1, v_2 \in X$$

则 S 有唯一不动点 u(即 $u = Su$).

♡

下面证明 Stampacchia 定理.

证明 对于(5.10)右式而言, 根据 Riesz-Frechet 表示定理5.2知存在唯一 $f \in H$ 使得

$$\langle \varphi, v \rangle = (f, v), \quad \forall v \in H.$$

对于(5.10)左式, 如果固定 $u \in H$, 则映射 $v \mapsto a(u,v)$ 是 H 上的连续线性泛函, 再次应用 Riesz-Frechet 表示定理知在 H 中存在唯一元素 (记之为 Au) 满足 $a(u,v) = (Au, v) (\forall v \in H)$. 显见 A 是 $H \to H$ 的线性算子, 且其满足:

$$|Au| \leq C|u|, \quad \forall u \in H, \tag{5.12}$$

$$(Au, u) \geq \alpha |u|^2, \quad \forall u \in H. \tag{5.13}$$

因此(5.10)式归结为找到 $u \in K$ 使得

$$(Au, v-u) \geq (f, v-u), \quad \forall v \in K. \tag{5.14}$$

设 $\rho > 0$ 是待定常数, 注意到(5.14)式等价于

$$(\rho f - \rho Au + u - u, v - u) \leq 0, \quad \forall v \in K, \tag{5.15}$$

这说明

$$u = P_K(\rho f - \rho Au + u).$$

现对每个 $v \in K$, 记 $Sv = P_K(\rho f - \rho Au + v)$. 我们下面说明当 $\rho > 0$ 取合适值时 S 是严格压缩映射. 事实上, 因为 P_K 不增加距离 (命题5.2), 故

$$|Sv_1 - Sv_2| \leq |(v_1 - v_2) - \rho(Av_1 - Av_2)|$$

因此

$$\begin{aligned} |Sv_1 - Sv_2|^2 &= |v_1 - v_2|^2 - 2\rho(Av_1 - Av_2, v_1 - v_2) + \rho^2 |Av_1 - Av_2|^2 \\ &\leq |v_1 - v_2|^2 (1 - 2\rho\alpha + \rho^2 C^2). \end{aligned}$$

取 $\rho > 0$ 满足 $k^2 = 1 - 2\rho\alpha + \rho^2 C^2 < 1$(即 $0 < \rho < 2\alpha/C^2$) 即知 S 有唯一不动点.

现设 $a(u,v)$ 对称, 则 $a(u,v)$ 定义了 H 上的一个新内积, 由其连续性与强正性知该内积对应的范数 $a(u,u)^{\frac{1}{2}}$ 与原来的范数 $|u|$ 等价, 因此 H 在这个新内积下依旧是 Hilbert 空间. 根据 Riesz-Frechet 表示定理, 我们可以在新内积下表示泛函 φ, 即存在唯一 $g \in H$ 使得

$$\langle \varphi, v \rangle = a(g, v), \quad \forall v \in H.$$

此时(5.10)式归结为找到 $u \in K$ 使得

$$a(g-u, v-u) \le 0, \quad \forall v \in K. \tag{5.16}$$

然而(5.16)式的解已经是个老朋友了: u 就是 g 在新内积 a 下到 K 上的投影. 根据定理5.1, u 是 K 中唯一一个让 $a(g-v, g-v)^{\frac{1}{2}}(v \in K)$ 达到最小值的元素. 也就是说, 它是函数

$$v \mapsto a(g-v, g-v) = a(v,v) - 2a(g,v) + a(g,g) = a(v,v) - 2\langle\varphi, v\rangle + a(g,g)$$

在 K 上的最小值点, 该最小值点等价于函数

$$v \mapsto \frac{1}{2}a(v,v) - \langle\varphi, v\rangle$$

的最小值点. □

注 容易说明若 $a(u,v)$ 作为双线性型同时还满足

$$a(v,v) \ge 0, \quad \forall v \in H$$

那么函数 $v \mapsto a(v,v)$ 就是凸的.

推论 5.2 (Lax-Milgram)

设 $a(u,v)$ 是 H 上的连续强正双线性型, 则对任意给定的 $\varphi \in H^*$, 均存在唯一 $u \in H$ 使得

$$a(u,v) = \langle\varphi, v\rangle, \quad \forall v \in H. \tag{5.17}$$

另若 a 对称, 则 u 能被下述性质刻画:

$$u \in H, \text{ 且 } \frac{1}{2}a(u,u) - \langle\varphi, u\rangle = \min_{v \in H}\left\{\frac{1}{2}a(v,v) - \langle\varphi, v\rangle\right\}. \tag{5.18}$$

♡

证明 在 Stampacchia 定理5.3中代入 $K = H$ 并用与正交投影定理5.1相同的证明方法即可. □

注 Lax-Milgram 定理是一个解线性椭圆 PDEs 的非常简单且有效的方法. 特别注意方程(5.17)与最小化问题(5.18)之间的联系. 当这类问题出现在力学或物理学中时, 它们通常都有一个自然的对应量: 最小作用量, 能量最小化等等. 在变分法中, 我们通常称(5.17)式是与最小化问题(5.18)关联的 Euler 方程. 简单来说, (5.17)式表明"$F'(u) = 0$", 其中 F 是函数 $F(v) = \frac{1}{2}a(v,v) - \langle\varphi, v\rangle$.

注 方程(5.17)解的存在唯一性有一个直接的初等证法. 此即说明

$$\forall f \in H \exists! u \in H (Au = f). \tag{5.19}$$

也就是说 A 是 H 到 H 上的双射. 这通过下述步骤即可实现:

(a) A 是单射 (因为 A 强正),

(b) $R(A)$ 是闭集, 因为 $\alpha|v| \le |Av| (\forall v \in H)$(这是强正性的一个结论),

(c) $R(A)$ 在 H 中稠密, 这是因为若 $v \in H$ 满足

$$(Au, v) = 0, \quad \forall u \in H$$

则 $v = 0$.

5.4 Hilbert 和, 正交基

定义 5.4 (Hilbert 和 (Hilbert sum))

设 $\{E_n\}_{n \ge 1}$ 是 H 中的闭子空间列. 若 E_n 满足

(i) E_n 两两正交, 即

$$(u,v) = 0, \quad \forall u \in E_n, \forall v \in E_m, m \ne n,$$

(ii) $\bigcup_{n=1}^{\infty} E_n$ 张成的线性空间[a]在 H 中稠密,

则称 H 是 E_n 的 Hilbert 和, 记之为 $H = \oplus_n E_n$.

[a]这里" 张成" 理解成代数意义下的, 即由 $\bigcup_{n=1}^{\infty} E_n$ 中元素的有限线性和构成的空间.

定理 5.5

设 H 是 E_n 的 Hilbert 和, 给定 $u \in H$, 记

$$u_n = P_{E_n} u, \quad S_n = \sum_{k=1}^{n} u_k,$$

则

$$\lim_{n\to\infty} S_n = u, \tag{5.20}$$

且有 Bessel-Parseval 等式成立:

$$\sum_{k=1}^{\infty} |u_k|^2 = |u|^2. \tag{5.21}$$

为证上述定理, 我们首先需要下述引理:

引理 5.1

设 $\{v_n\} \subset H$ 满足

$$(v_m, v_n) = 0, \quad \forall m \neq n \tag{5.22}$$

$$\sum_{k=1}^{\infty} |v_k|^2 < \infty. \tag{5.23}$$

记

$$S_n = \sum_{k=1}^{n} v_k,$$

则

$$S = \lim_{n\to\infty} S_n \text{ 存在}$$

且

$$|S|^2 = \sum_{k=1}^{\infty} |v_k|^2. \tag{5.24}$$

下面证明引理5.1.

证明 当 $m > n$ 时有

$$|S_m - S_n|^2 = \sum_{k=n+1}^{m} |v_k|^2.$$

由(5.23)式知 S_n 是基本列, 因此 $S = \lim_{n\to\infty} S_n$ 存在. 进一步, 根据(5.22)式知

$$|S_n|^2 = \sum_{k=1}^{n} |v_k|^2$$

令 $n \to \infty$ 即得(5.24)式. □

下面证明定理5.5.

证明 因为 $u_n = P_{E_n} u$, 由正交投影定理5.1可知

$$(u - u_n, v) = 0, \quad \forall v \in E_n \tag{5.25}$$

特别有

$$(u, u_n) = |u_n|^2.$$

将上式对 n 求和可得

$$(u, S_n) = \sum_{k=1}^{n} |u_k|^2.$$

但另一方面有

$$\sum_{k=1}^{n} |u_k|^2 = |S_n|^2, \tag{5.26}$$

因此

$$(u, S_n) = |S_n|^2.$$

故 $|S_n| \le |u|$, 从而 $\sum_{k=1}^{n} |u_k|^2 \le |u|^2$.

现在应用引理5.1即知 $S = \lim_{n\to\infty} S_n$ 存在. 实际上, 就算我们没有 Hilbert 和的条件 (ii), 在这里也可以确定 S. 如果设 F 是 $\{E_n\}$ 张成的线性空间, 我们断言

$$S = P_{\overline{F}} u. \tag{5.27}$$

事实上, 根据正交性条件 (i) 与正交投影定理知

$$(u - S_n, v) = 0, \quad \forall v \in E_m, m \le n$$

令 $n \to \infty$ 知

$$(u - S, v) = 0, \quad \forall v \in E_m, \forall m$$

因此

$$(u - S, v) = 0, \quad \forall v \in F,$$

根据内积的连续性知

$$(u - S, v) = 0, \quad \forall v \in \overline{F}.$$

另一方面显见 $S_n \in F(\forall n)$, 因此 $S \in \overline{F}$, 利用正交投影定理即得(5.27)式. 当然, 如果加上条件 (ii), 则 $\overline{F} = H$, 因此 $S = u$, 在(5.26)式中令 $n \to \infty$ 即得(5.21)式. □

定义 5.5 (标准正交基 (orthonormal basis))

若 H 中的序列 $\{e_n\}_{n\ge1}$ 满足:

(i) $|e_n| = 1(\forall n)$ 且 $(e_m, e_n) = 0(\forall m \ne n)$,

(ii) $\{e_n\}$ 张成的线性空间在 H 中稠密,

就称 $\{e_n\}_{n\ge1}$ 是 H 中的一组标准正交基 (或 Hilbert 基[a], 在无歧义时简称为基)[b].

[a]注意不要把这个概念和代数基 (即 Hamel 基) 弄混了, 后者指的是 H 中的元素族 $\{e_i\}_{i\in I}$, 其中每个 $u \in H$ 都能唯一表成 e_i 的有限线性组合.

[b]有的作者会把 $\{e_n\}$ 称为完备正交基.

♣

推论 5.3

设 $\{e_n\}$ 是标准正交基, 则对每个 $u \in H$ 有

$$u = \sum_{k=1}^{\infty} (u, e_k) e_k, \text{ 即 } u = \lim_{n\to\infty} \sum_{k=1}^{n} (u, e_k) e_k$$

且

$$|u|^2 = \sum_{k=1}^{\infty} |(u, e_k)|^2.$$

反之, 对任意给定的 $\{\alpha_n\} \in l^2$, 级数 $\sum_{k=1}^{\infty} \alpha_k e_k$ 总能收敛到 H 中的某元素 u, 使得 $(u, e_k) = \alpha_k(\forall k)$, 且 $|u|^2 = \sum_{k=1}^{\infty} \alpha_k^2$.

♡

证明 注意到 H 是空间 $E_n = \mathbb{R}e_n$ 的 Hilbert 和, $P_{E_n}u = (u, e_n)e_n$, 由定理5.5与引理5.1即得结论. □

注 一般来说, 定理5.5中的级数 $\sum u_k$ 与推论5.3中的级数 $\sum(u, e_k)e_k$ 并不绝对收敛. 也就是说, 可能出现 $\sum_{k=1}^{\infty}|u_k| = \infty$ 或 $\sum_{k=1}^{\infty}|(u, e_k)| = \infty$ 的情况.

定理 5.6 (标准正交基的存在性)

每个可分 Hilbert 空间都有标准正交基. ♡

证明 设 $\{v_n\}$ 是 H 的可数稠密子集, 记 F_k 为 $\{v_1, \cdots, v_k\}$ 张成的线性空间, 则集列 $\{F_k\}$ 是有限维空间的不降列, 且 $\bigcup_{k=1}^{\infty} F_k$ 在 H 中稠密. 任取 F_1 中的单位向量 e_1, 若 $F_2 \neq F_1$, 则在 F_2 中存在向量 e_2 使得 $\{e_1, e_2\}$ 是 F_2 的标准正交基. 重复这一构造即得 H 上的标准正交基. □

注 将定理5.6与推论5.3结合可知每个可分 Hilbert 空间都与 l^2 等距同构. 尽管这一结果非常惊人, 其它 Hilbert 空间 (比如 $L^2(\Omega)$, Sobolev 空间 $H^1(\Omega)$ 等等) 也是很重要的. 这是因为很多漂亮的线性 (或非线性) 算子在用基表示的时候可能看起来很恐怖.

注 如果 H 是不可分 Hilbert 空间 (这个情况不那么常见), 则可以 (用 Zorn 引理) 证明其上存在不可数标准正交基 $\{e_i\}_{i\in I}$.

第二部分

算子谱论

第六章 紧算子的谱理论

本章选自 **[YDC]**, 作为杨大春老师开设的研究生课程泛函分析的课程笔记.

课堂笔记 (2024.9.10)

泛函分析就是无穷维空间上的分析 (或称为无穷维空间上的数学分析与高等代数).

(i) 起源 (背景): 数学与物理中的具体分析问题.

(ii) 研究方法: 具体的分析问题 $\overset{\text{抽象}}{\Longrightarrow}$(具有拓扑 (度量) 和代数结构的) 空间上的分析问题.

(iii) 现状: 内容丰富, 应用广泛, 数学工作者必须的.

6.1 有界线性算子的谱

约定 $\mathbb{N}_+ = \{1, 2, \cdots\}$, $\mathbb{C}$ 表示全体复数. 设 X, Y 为赋范线性空间, 用 $L(X, Y)$ 表示 $X \to Y$ 的有界线性算子全体. 另对每个 $T \in L(X, Y)$, 用 $\|T\|_{X\to Y}$ 表示 T 在 $L(X, Y)$ 中的范数 (亦即 T 的算子范数). 特别地, 若 $X = Y$, 则记 $L(X) := L(X, X)$, 用 $\|T\|_{X\to X}$ 表示 $T \in L(X)$ 的算子范数.

定义 6.1 (闭算子 (closed operator))

设 $A : X \to Y$ 是线性算子, $D(A) \subset X$ 是 A 的定义域. 称 A 是闭算子, 如果

$$\begin{cases} \{x_n\}_{n\in\mathbb{N}} \subset D(A) \\ n \to \infty \text{时有} x_n \to x, Ax_n \to y \end{cases} \Rightarrow \begin{cases} x \in D(A) \\ y = Ax \end{cases}$$

由闭算子的定义易知, 若算子 A 是闭算子, 则对任意 $\lambda \in \mathbb{C}$ 而言, 算子 $\lambda I - A$ 依旧是闭算子, 其中 I 是恒同算子. 另外若[1] $A \in L(X)$, 那么 A 是闭算子[2]. 这是因为任取 $\{x_n\}_{n\in\mathbb{N}_+} \subset X$ 与 $x, y \in X$, 设 $x_n \to x, Ax_n \to y (n \to \infty)$, 因为 $D(A) = X$ 且 A 连续, 故 $x \in X = D(A)$, $Ax = y$, 因此 A 是闭算子.

关于闭算子, 需要回忆一下闭图像定理:

定理 6.1 (闭图像定理)

设 X, Y 为 Banach 空间, 若 $A : D(A) \subset X \to Y$ 是闭线性算子, 且 $D(A)$ 是 X 中的闭集, 那么 A 是有界线性算子.

注 关于闭算子还需要强调下述注记:

(i) 若 X, Y 是 Banach 空间, $A : D(A) \subset X \to Y$ 是线性算子, 则在 $D(A)$ 闭的时候由闭图像定理可知 A 是闭算子 $\Leftrightarrow$ A 是有界算子. 具体来说, 若 A 是闭算子, 由闭图像定理知 A 是有界算子. 若 A 是有界算子, 设 $x, y \in X$, $\{x_n\}_{n\in\mathbb{N}_+} \subset D(A)$ 满足 $x_n \to x, Ax_n \to y (n \to \infty)$, 根据闭性知 $x \in D(A)$, 于是由有界性所等价的连续性知 $y = \lim_{n\to\infty} Ax_n = A(\lim_{n\to\infty} x_n) = Ax$, 因此 A 是闭算子.

(ii) 令 $X = C[0, 1]$, 且对任意 $u \in C[0, 1]$ 令

$$\|u\|_{C[0,1]} := \max_{t\in[0,1]} |u(t)|.$$

再令 $A : D(A) \subset X \to X, u \mapsto \frac{du}{dt}$, 其中 $D(A) = C^1[0, 1]$, 则 A 是闭算子, 但 A 不是有界算子. 对于 A 是闭算子, 任取 $\{x_n\}_{n\in\mathbb{N}_+} \in C^1[0, 1]$ 与 $x, y \in C[0, 1]$ 满足 $x_n \to x (n \to \infty)$ 在 $C[0, 1]$ 意义下成立, $Ax_n \to y (n \to \infty)$ 在 $C[0, 1]$ 意义下成立, 则 $n \to \infty$ 时对任意 $t \in [0, 1]$ 有

$$x_n(t) - x_n(0) = \int_0^t x_n'(\tau)d\tau \to \int_0^t y(\tau)d\tau,$$

---

[1] 注意: $A \in L(X, Y)$ 这句话默认了 $D(A) = X$.

[2] 注意这里其实不需要 X 是 Banach 空间, 这是因为在取点的时候只能在 X 内取, 而 $D(A) = X$, 故 $x \in D(A)$ 这一条其实是自动成立了.

且 $x_n(t)-x_n(0)\to x(t)-x(0)$. 故对任意 $t\in[0,1]$ 有

$$x(t)-x(0)=\int_0^t y(\tau)d\tau.$$

因此 $x\in C^1[0,1], x'=y$, 即 $x\in D(A), Ax=y$, 故 A 是闭算子.

下面说明 A 不是有界算子, 为此对 $\forall n\in\mathbb{N}_+$ 与 $\forall t\in[0,1]$, 令 $u_n(t)=\sin(n\pi t)$, 则 $u_n\in C^1[0,1], \|u_n\|_{C[0,1]}=1$. 但

$$\|Au_n\|_{C[0,1]}=\max_{t\in[0,1]}|n\pi\cos(n\pi t)|=n\pi.$$

因此

$$\begin{aligned}\|A\|_{C^1[0,1]\to C[0,1]}&=\sup_{\|u\|_{C[0,1]}=1}\|Au\|_{C[0,1]}\geq\sup_{n\in\mathbb{N}_+}\|Au_n\|_{C[0,1]}\\&=\sup_{n\in\mathbb{N}_+}(n\pi)=\infty.\end{aligned}$$

故 A 不是有界算子. □

注 需要注意: 上面证明了存在不是有界算子的闭算子, 但不会与闭图像定理矛盾, 这是因为上面提到的 $D(A)=C^1[0,1]$ 并非 $C[0,1]$ 的闭子集. 事实上, $\overline{D(A)}^{C[0,1]}=C[0,1]$. 显见 $\overline{D(A)}^{C[0,1]}\subset C[0,1]$, 下证 $\overline{D(A)}^{C[0,1]}\supset C[0,1]$. 任取 $f\in C[0,1]$, 由 Weierstrass 逼近定理知存在 $\mathbb{R}$ 上的多项式列 $\{P_n\}_{n\in\mathbb{N}_+}$ 在 $[0,1]$ 上一致收敛到 f, 即 $\max_{t\in[0,1]}|P_n(t)-f(t)|\to 0(n\to\infty)$. 此即 $\|P_n-f\|_{C[0,1]}\to 0(n\to\infty)$, 又 $\{P_n\}_{n\in\mathbb{N}_+}\subset D(A)$, 故 $f\in\overline{D(A)}^{C[0,1]}$, 即 $\overline{D(A)}^{C[0,1]}\supset C[0,1]$.

(iii) 如果把 (ii) 的设置稍作改动, 对 $\forall u\in C^1[0,1]$ 令 $\|u\|_{C^1[0,1]}:=\max\{\|u\|_{C[0,1]},\|u\|_{C[0,1]}\}$. 若视 $T:=\frac{d}{dt}:C^1[0,1]\to C[0,1]$, 则 T 成为有界算子. 这是因为此时对 $\forall u\in C^1[0,1]$ 有 $\|Tu\|_{C[0,1]}=\|u'\|_{C[0,1]}\leq\|u\|_{C^1[0,1]}$.

现在考虑非平凡[3]复 Banach 空间 X 与闭线性算子 $A:D(A)\subset X\to X$. 若存在 $\lambda\in\mathbb{C}$ 和 $x_0\in D(A)\backslash\{0_X\}$ 使得

$$Ax_0=\lambda x_0,$$

则称 λ 为 A 的**本征值 (eigenvalue)**, 称 x_0 为 λ 的**本征元 (eigenvector)**.

定义 6.2 (预解集, 正则值 (resolvent set, regular value))

设 X 是复 Banach 空间, $A:D(A)\subset X\to X$ 是闭线性算子, I 是 X 上的恒同算子. 称集合

$$\rho(A):=\{\lambda\in\mathbb{C}:(\lambda I-A)^{-1}\in L(X)\}$$

为 A 的预解集, $\rho(A)$ 中的元素 λ 称为 A 的正则值.

♣

当 $\dim X<\infty$($\dim X$ 表示 X 的维数) 时, $A\in L(X)$ 必是矩阵, 因而由线性代数理论知 $\lambda\in\mathbb{C}$ 的选取只有下述两种可能:

(i) 要么 λ 是 A 的本征值;

(ii) 要么 $(\lambda I-A)^{-1}$ 是一个矩阵, 亦即 $(\lambda I-A)^{-1}\in L(X)$, 此时 λ 是 A 的正则值.

课堂笔记 (2024.9.10)

- 只会存在两种可能是因为在 $\dim X<\infty, \lambda\in\mathbb{C}$ 时, 要么 $|\lambda I-A|=0$, 要么 $|\lambda I-A|\neq 0$. 若 $|\lambda I-A|=0$, 则 $\exists x_0\in X\backslash\{0_X\}$ s.t. $(\lambda I-A)x_0=0_X$, 因此 λ 是本征值. 而若 $|\lambda I-A|\neq 0$, 则 $(\lambda I-A)^{-1}$ 存在, 因此 $(\lambda I-A)^{-1}\in L(X)$, 故 λ 是正则值.
- 有限维线性赋范空间上的线性算子完全由一个矩阵确定且必有界. 简便起见, 设 X 是 $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$ 上的线性赋范空间, $\dim X=2$, $\{e_1,e_2\}$ 为 X 的基. 设 $A:X\to X$ 是线性算子, 则 $Ae_1\in X, Ae_2\in X$, 因

[3]这里的非平凡指的是 $X\supsetneq\{0_X\}$, 其中 0_X 是 X 的零元.

此存在 $\{a_{11}, a_{12}, a_{21}, a_{22}\} \subset \mathbb{K}$ 使得

$$\begin{cases} Ae_1 = a_{11}e_1 + a_{12}e_2 \\ Ae_2 = a_{21}e_1 + a_{22}e_2. \end{cases}$$

现任取 $x \in X$, 则存在 $\{x_1, x_2\} \subset \mathbb{K}$ 使得 $x = x_1e_1 + x_2e_2$, 因此由 A 的线性性知

$$\begin{aligned} Ax &= x_1Ae_1 + x_2Ae_2 \\ &= x_1a_{11}e_1 + x_1a_{12}e_2 + x_2a_{21}e_1 + x_2a_{22}e_2 \\ &= (x_1a_{11} + x_2a_{21})e_1 + (x_1a_{12} + x_2a_{22})e_2 \\ &= (x_1a_{11} + x_2a_{21}, x_1a_{12} + x_2a_{22}) \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \\ &= (x_1, x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}. \end{aligned}$$

因此, A 完全由矩阵

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

确定.

下面说明 A 有界. 为此, 利用范数等价定理2.9知对 $\forall x \in X$ 有

$$\|x\| \sim \sqrt{|x_1|^2 + |x_2|^2}.$$

由 Cauchy-Schwarz 不等式可知对 $\forall x \in X$ 有

$$\begin{aligned} \|Ax\| &\sim \sqrt{|x_1a_{11} + x_2a_{21}|^2 + |x_1a_{12} + x_2a_{22}|^2} \\ &\le \sqrt{(|x_1|^2 + |x_2|^2)(|a_{11}|^2 + |a_{21}|^2 + |a_{12}|^2 + |a_{22}|^2)} \end{aligned}$$

由此, 进一步有

$$\begin{aligned} \|A\|_{X\to X} &:= \sup_{x\neq 0_X} \frac{\|Ax\|}{\|x\|} \\ &\sim \sup_{x\neq 0_X} \frac{\sqrt{|x_1a_{11} + x_2a_{21}|^2 + |x_1a_{12} + x_2a_{22}|^2}}{\sqrt{|x_1|^2 + |x_2|^2}} \\ &\le \sqrt{|a_{11}|^2 + |a_{21}|^2 + |a_{12}|^2 + |a_{22}|^2} < \infty. \end{aligned}$$

故 A 有界.

- 特别地, 若 A 为 $\mathbb{R}$ 上的线性算子, 则对 $\forall x \in \mathbb{R}$ 有

$$Ax = A(x1) = xA(1).$$

因此 A 完全由 $A(1)$(这是一个 1×1 方阵) 决定, 且

$$\|A\|_{\mathbb{R}\to\mathbb{R}} := \sup_{x\neq 0} \frac{|Ax|}{|x|} = |A(1)| < \infty.$$

故 A 有界. 注意 A 为 $\mathbb{R}$ 上的有界线性算子不等价于 A 为 $\mathbb{R}$ 上的有界函数. 事实上, 设 A 为 $\mathbb{R}$ 上的有界线性算子, 则 A 为 $\mathbb{R}$ 上的有界函数 $\Leftrightarrow A = 0_{L(\mathbb{R})}$.

♠

对 $\dim X = \infty$ 的情况, $\lambda \in \mathbb{C}$ 的选取会且仅会导致下述四个结果其一:

(i) $(\lambda I - A)^{-1}$ 不存在, 亦即 λ 是 A 的本征值. 记这样的 $\lambda \in \mathbb{C}$ 构成的全体为[4]$\sigma_p(A)$, 称这样的 λ 为 A 的**点谱 (point spectrum)**;

[4]其中角标 p 意为 proper value.

(ii) $(\lambda I-A)^{-1}$ 存在, 且 $\lambda I-A$ 的值域

$$R(\lambda I-A)^{-1}:=(\lambda I-A)(D(A))=X$$

此时 λ 是 A 的正则值;

(iii) $(\lambda I-A)^{-1}$ 存在, $R(\lambda I-A)\neq X$, 但 $\overline{R(\lambda I-A)}=X$. 此时称 λ 为 A 的**连续谱 (continuous spectrum)**, 记这样的 $\lambda\in\mathbb{C}$ 构成的全体为 $\sigma_c(A)$;

(iv) $(\lambda I-A)^{-1}$ 存在, 且 $\overline{R(\lambda I-A)}\neq X$. 此时称 λ 为 A 的**剩余谱 (residual spectrum)**, 记这样的 $\lambda\in\mathbb{C}$ 构成的全体为 $\sigma_r(A)$.

注 设 X 为非平凡复 Banach 空间, $A:D(A)\subset X\to X$ 为闭算子, 则

$$\begin{aligned}(\lambda I-A)^{-1}\text{不存在}&\Leftrightarrow \lambda I-A\text{不是单射}\\&\Leftrightarrow \exists x_0\in D(A)\backslash\{0_X\}((\lambda I-A)x_0=0_X)\\&\Leftrightarrow \lambda\text{ 为 }A\text{ 的本征值.}\end{aligned}$$

现称 $\sigma(A):=\mathbb{C}\backslash\rho(A)$ 为 A 的**谱集 (spectrum set)**, $\lambda\in\sigma(A)$ 称为 A 的**谱点 (spectrum)**. 可知

$$\sigma(A)=\sigma_p(A)\cup\sigma_c(A)\cup\sigma_r(A).$$

根据定义可知, $\rho(A),\sigma_p(A),\sigma_c(A),\sigma_r(A)$ 互不相交.

下述命题表明 $\lambda\in\rho(A)$ 当且仅当 $(\lambda I-A)^{-1}$ 存在且 $R(\lambda I-A)=X$, 即前述正则值的定义合理.

命题 6.1

设 X 为 Banach 空间, $A:D(A)\subset X\to X$ 为闭线性算子. 若 $(\lambda I-A)^{-1}$ 存在, 且 $\lambda I-A$ 的值域

$$R(\lambda I-A):=(\lambda I-A)(D(A))=X,$$

则 $(\lambda I-A)^{-1}\in L(X)$.

证明 因为 $D((\lambda I-A)^{-1})=R(\lambda I-A)=X$ 是 X 中的闭集, 故根据闭图像定理6.1, 要想说明 $(\lambda I-A)^{-1}\in L(X)$, 只需说明 $(\lambda I-A)^{-1}$ 是闭算子.

首先说明 $\lambda I-A$ 是闭算子. 设 $\{x_n\}_{n\in\mathbb{N}}\subset D(A)$ 满足

$$x_n\to x\text{ 且 }(\lambda I-A)x_n\to y,\quad n\to\infty \tag{6.1}$$

往证 $x\in D(A), y=(\lambda I-A)x$. 注意到(6.1)第二式等价于 $\lambda x-\lim_{n\to\infty}Ax_n=y$, 因而(6.1)式即

$$Ax_n\to\lambda x-y\text{ 且 }x_n\to x,\quad n\to\infty$$

现由 A 是闭算子知 $x\in D(A), Ax=\lambda x-y$, 于是 $(\lambda I-A)x=y$, 这说明 $\lambda I-A$ 是闭算子.

下证 $(\lambda I-A)^{-1}$ 是闭算子. 设 $\{y_n\}_{n\in\mathbb{N}}\subset D((\lambda I-A)^{-1})=X$ 满足

$$y_n\to y\text{ 且 }x_n:=(\lambda I-A)^{-1}y_n\to x,\quad n\to\infty \tag{6.2}$$

往证 $y\in D((\lambda I-A)^{-1}), x=(\lambda I-A)^{-1}y$. 由(6.2)式知

$$(\lambda I-A)x_n=y_n\to y\text{ 且 }x_n\to x,\quad n\to\infty$$

现由 $\lambda I-A$ 是闭算子与 $\{x_n\}_{n\in\mathbb{N}}\subset D(\lambda I-A)$ 知 $x\in D(\lambda I-A)=D(A), (\lambda I-A)x=y$, 于是

$$y\in R(\lambda I-A)=D((\lambda I-A)^{-1})$$

且 $(\lambda I-A)^{-1}y=x$, 这说明 $(\lambda I-A)^{-1}$ 为闭算子, 命题因而得证. □

下面的一些例子表明当 $\dim X=\infty$ 时, 点谱、连续谱与剩余谱都是有可能出现的.

例 6.1 设 $X=C[0,1]$, $A:u(t)\mapsto -\frac{d^2}{dt^2}u(t)$, 其定义域为

$$D(A):=\{u\in C^2[0,1]:u(0)=u(1),u'(0)=u'(1)\}.$$

则 A 为 $D(A)\subset C[0,1]\to C[0,1]$ 的闭线性算子, $\sigma(A)=\sigma_p(A)=\{(2n\pi)^2:n\in\mathbb{N}\}$, 且 $\rho(A)=\mathbb{C}\backslash\{(2n\pi)^2:n\in\mathbb{N}\}$.

证明 首先说明 A 是闭算子, 亦即说明当 $\{u_n\}_{n\in\mathbb{N}} \subset D(A)$ 且

$$\begin{cases} u_n \to u, & n\to\infty, \\ -u_n'' \to v, & n\to\infty \end{cases}$$

在 $C[0,1]$ 意义下成立时, 有 $u\in D(A)$ 且 $-u''=v$. 为此, 对 $\forall t\in[0,1]$ 令[5]

$$y(t) := -\int_0^t v(s)ds$$

注意到对 $\forall t\in[0,1]$, 当 $n\to\infty$ 时有

$$|u_n'(t)-u_n'(0)-y(t)| = \left|\int_0^t [u_n''(s)+v(s)]ds\right| \le \|v-(-u_n'')\|_{C[0,1]} \to 0, \quad n\to\infty \tag{6.3}$$

从而

$$\|u_n'-u_n'(0)-y\|_{C[0,1]} \to 0, \quad n\to\infty.$$

特别在(6.3)式中令 $t=1$ 有

$$\int_0^1 v(s)ds = 0.$$

这说明 $y(0)=y(1)=0$.

下面证明存在 $\alpha\in\mathbb{C}$ 使得 $\lim_{n\to\infty} u_n'(0)=\alpha$, 由此与前述已证结论可得

$$u_n' \to \alpha + y, \quad n\to\infty$$

在 $C[0,1]$ 意义下成立. 事实上, 对 $\forall n\in\mathbb{N}$ 由 $u_n(0)=u_n(1)$ 知

$$\int_0^1 u_n'(t)dt = 0$$

从而当 $n,m\to\infty$ 时

$$\begin{aligned} |u_n'(0)-u_m'(0)| &= \left|\int_0^1 [-u_n'(t)+u_n'(0)+u_m'(t)-u_m'(0)]dt\right| \\ &\le \|[u_m'-u_m'(0)]-[u_n'-u_n'(0)]\|_{C[0,1]} \to 0, \quad n,m\to\infty \end{aligned}$$

由此可知 $\{u_n'(0)\}_{n\in\mathbb{N}}$ 是 $\mathbb{C}$ 上的 Cauchy 列, 因而存在 $\alpha\in\mathbb{C}$ 使得 $n\to\infty$ 时有

$$u_n'(0) \to \alpha$$

对 $\forall t\in[0,1]$, 定义

$$\Omega(t) := \int_0^t [\alpha+y(s)]ds$$

则 $\Omega(0)=0$ 显然成立. 下面说明 $\{u_n-u_n(0)\}_{n\in\mathbb{N}}$ 在 $C[0,1]$ 意义下收敛到 Ω. 事实上, 注意到对 $\forall t\in[0,1]$ 与 $\forall n\in\mathbb{N}$ 有

$$u_n(t)-u_n(0) = \int_0^t u_n'(s)ds$$

故对 $\forall t\in[0,1]$, 当 $n\to\infty$ 时有

$$\left|u_n(t)-u_n(0)-\int_0^t [\alpha+y(s)]ds\right| = \left|\int_0^t [u_n'(s)-\alpha-y(s)]ds\right| \le \|u_n'-(\alpha+y)\|_{C[0,1]} \to 0$$

因此由 $u_n(1)-u_n(0)=0(\forall n\in\mathbb{N})$ 知 $\Omega(1)=0$, 且 $n\to\infty$ 时

$$\|u_n-u_n(0)-\Omega\|_{C[0,1]} \le \|u_n'-(\alpha+y)\|_{C[0,1]} \to 0.$$

又由 $u_n\to u(n\to\infty)$ 在 $C[0,1]$ 意义下成立知

$$\lim_{n\to\infty} u_n(0) = u(0).$$

[5]整个证明的思路在于把 u 用 v 的二重积分表示出来, 因此先说明 u_n' 的极限能用 v 的一重积分表示, 再说明 u_n 的仅限能用 v 的二重积分表示.

令 $\beta := u(0)$, 则当 $n \to \infty$ 时

$$u_n \to \beta + \Omega$$

在 $C[0,1]$ 意义下成立. 注意到对 $\forall n \in \mathbb{N}$ 有

$$\|u-(\beta+\Omega)\|_{C[0,1]} \le \|u-u_n\|_{C[0,1]} + \|u_n-(\beta+\Omega)\|_{C[0,1]}$$

令 $n \to \infty$ 易得

$$\|u-(\beta+\Omega)\|_{C[0,1]} = 0.$$

由此知对 $\forall t \in [0,1]$ 有

$$u(t) = \beta + \Omega(t) = \beta + \alpha t - \int_0^t \int_0^s v(\tau) d\tau ds.$$

显见 $u \in C^2[0,1]$ 且 $u'' = -v$. 由 $\Omega(1) = \Omega(0) = 0$ 知 $u(0) = \beta = u(1)$. 又因为对 $\forall t \in [0,1]$ 有

$$u'(t) = \alpha - \int_0^t v(\tau) d\tau,$$

故 $u'(0) = \alpha$ 且

$$u'(1) = \alpha - \int_0^1 v(\tau) d\tau = \alpha = u'(0).$$

故 $u \in D(A)$, 从而 A 为闭算子.

但 A 不是有界算子. 为证此事, 对 $\forall n \in \mathbb{N}$ 与 $\forall t \in [0,1]$ 令

$$f_n(t) := \sin(2n\pi t),$$

则 $f_n \in D(A)$ 且 $\|f_n\|_{C[0,1]} = 1$. 又对 $\forall n \in \mathbb{N}$ 及 $\forall t \in [0,1]$ 有

$$Af_n(t) = (2n\pi)^2 \sin(2n\pi t),$$

故 $\|Af_n\|_{C[0,1]} = (2n\pi)^2$. 因此对 $\forall n \in \mathbb{N}$ 均有

$$\|A\|_{C[0,1]\to C[0,1]} \ge (2n\pi)^2.$$

令 $n \to \infty$ 即知 A 无界.

下面证明 $\sigma(A) = \sigma_p(A) = \{(2n\pi)^2 : n \in \mathbb{N}\}$. 注意到对任意 $n \in \mathbb{N}$ 均有

$$A\sin(2n\pi t) = (2n\pi)^2 \sin(2n\pi t), \quad A\cos(2n\pi t) = (2n\pi)^2 \cos *2n\pi t.$$

因此 $(2n\pi)^2 \in \sigma_p(A)$. 若令

$$E := \{(2n\pi)^2 : n \in \mathbb{N}\},$$

则 $E \subset \sigma_p(A)$. 下面说明 $\sigma_p(A) \subset E$, 为此只需说明 $(\mathbb{C}\backslash E) \subset \rho(A)$, 这是因为若该事实成立, 则有

$$\mathbb{C} = E \cup (\mathbb{C}\backslash E) \subset \sigma_p(A) \cup \rho(A) \subset \mathbb{C}$$

因而 $\mathbb{C} = \sigma_p(A) \cup \rho(A)$, 故 $\sigma_p(A) = \rho(A)^c \subset E$, 结论进而成立.

现在说明 $\lambda \in \mathbb{C}\backslash E$ 时, $(\lambda I - A)^{-1}$ 存在, 这只需说明 $\lambda I - A$ 是单射. 设 $u_1, u_2 \in D(A)$ 满足 $u_1 \ne u_2$, 若 $(\lambda I - A)u_1 = (\lambda I - A)u_2$, 代入 $A = -\frac{d^2}{dt^2}$ 知

$$-\frac{d^2}{dt^2}(u_1 - u_2) = \lambda(u_1 - u_2)$$

由 ODE 理论解得

$$u_1 - u_2 = C_1 e^{\gamma_0 t} + C_2 e^{-\gamma_0 t} \tag{6.4}$$

其中 $C_1, C_2 \in \mathbb{C}$, γ_0 满足 $\lambda + \gamma_0^2 = 0$ 且 $\gamma_0 = i\sqrt{|\lambda|}e^{i\frac{\arg\lambda}{2}}$. 因为 $\lambda \in \mathbb{C}\backslash E, 0 \in E$, 故 $\lambda \ne 0$, 从而 $\gamma_0 \ne 0$. 下面断言当 $\lambda \in \mathbb{C}\backslash E$ 时有

$$e^{\pm\gamma_0} \ne 1. \tag{6.5}$$

这是因为 $e^{-\gamma_0}=1 \Leftrightarrow e^{\gamma_0}=1$, 而

$$
\begin{aligned}
e^{\gamma_0}=1 &\Leftrightarrow e^{i\sqrt{|\lambda|}e^{i\frac{\arg\lambda}{2}}}=1\\
&\Leftrightarrow 1=e^{i\sqrt{|\lambda|}[\cos\frac{\arg\lambda}{2}+i\sin\frac{\arg\lambda}{2}]}=e^{-\sqrt{|\lambda|}\sin\frac{\arg\lambda}{2}}e^{i\sqrt{|\lambda|}\cos\frac{\arg\lambda}{2}}\\
&\Leftrightarrow e^{-\sqrt{|\lambda|}\sin\frac{\arg\lambda}{2}}=1\text{且}e^{i\sqrt{|\lambda|}\cos\frac{\arg\lambda}{2}}=1\\
&\Leftrightarrow \sqrt{|\lambda|}\sin\frac{\arg\lambda}{2}=0, \cos\left(\sqrt{|\lambda|}\cos\frac{\arg\lambda}{2}\right)=1, \sin\left(\sqrt{|\lambda|}\cos\frac{\arg\lambda}{2}\right)=0\\
&\Leftrightarrow \begin{cases}\sin\frac{\arg\lambda}{2}=0,\\ \sqrt{|\lambda|}\cos\frac{\arg\lambda}{2}=2k\pi, k\in\mathbb{Z}\end{cases}\\
&\Leftrightarrow \begin{cases}\frac{\arg\lambda}{2}=n\pi, & n\in\mathbb{Z}\\ \sqrt{|\lambda|}=2k\pi, & k\in\mathbb{Z}\backslash\{0\}\end{cases}\\
&\Leftrightarrow \begin{cases}\arg\lambda=2n\pi, & n\in\mathbb{Z}\\ |\lambda|=(2k\pi)^2, & k\in\mathbb{Z}\backslash\{0\}\end{cases}\\
&\Leftrightarrow \lambda=(2k\pi)^2, k\in\mathbb{Z}\backslash\{0\}.
\end{aligned}
$$

注意到

$$\{(2k\pi)^2 : k\in\mathbb{Z}\}=\{(2k\pi)^2 : k\in\mathbb{N}\}$$

且 $\lambda=0$ 时 $\gamma_0=0$, 因此 $e^{\gamma_0}=1$. 故 $e^{\gamma_0}=1 \Leftrightarrow \lambda\in E$, 从而 $\lambda\in C\backslash E$ 时 $e^{\pm\gamma_0}\neq 1$. 所证断言因而成立.

又因为对 $\forall i\in\{1,2\}$ 有

$$u_i(0)=u_i(1), u_i'(0)=u_i'(1),$$

代入(6.4)式可得

$$C_1=C_2=0.$$

因此 $u_1=u_2$, 矛盾! 故 $(\lambda I-A)u_1\neq(\lambda I-A)u_2$, 从而 $\lambda I-A$ 为单射, 亦即 $(\lambda I-A)^{-1}$ 存在.

下面还需证明 $\lambda\in\mathbb{C}\backslash E$ 时有

$$R(\lambda I-A)=C[0,1].$$

亦即对任意 $v\in C[0,1]$, 需证明存在 $u\in D(A)$ 使得 $v=(\lambda I-A)u$. 这等价于

$$\frac{d^2}{dt^2}u+\lambda u=\left(\frac{d}{dt}+\gamma_0\right)\left(\frac{d}{dt}-\gamma_0\right)u=v, \tag{6.6}$$

其中 $\gamma_0=i\sqrt{|\lambda|}e^{i\frac{\arg\lambda}{2}}, v\in C[0,1]$. 令

$$\widetilde{u}:=\left(\frac{d}{dt}-\gamma_0\right)u, \tag{6.7}$$

则(6.6)式即

$$\left(\frac{d}{dt}+\gamma_0\right)\widetilde{u}=v. \tag{6.8}$$

在(6.7)式两端同乘 $e^{-\gamma_0\tau}$ 并对 τ 在 $[0,t]$ 上积分, 则对 $\forall t\in[0,1]$ 有

$$\int_0^t\frac{d}{d\tau}(e^{-\gamma_0\cdot}u(\cdot))(\tau)=\int_0^t e^{-\gamma_0\tau}\widetilde{u}(\tau)d\tau.$$

因此对 $\forall t\in[0,1]$ 有

$$u(t)=e^{\gamma_0 t}\left[\int_0^t e^{-\gamma_0\tau}\widetilde{u}(\tau)d\tau+u(0)\right]. \tag{6.9}$$

类似地, 在(6.8)式两端同乘 $e^{\gamma_0\tau}$ 并重复上述步骤可得对 $\forall t\in[0,1]$ 有

$$\widetilde{u}(t)=e^{-\gamma_0 t}\left[\int_0^t e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]. \tag{6.10}$$

下面求解 $u(0)$ 与 $\widetilde{u}(0)$. 结合(6.9),(6.10)两式知

$$u(t)=e^{\gamma_0 t}\int_0^t e^{-2\gamma_0 t}\left[\int_0^s e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]ds+e^{\gamma_0 t}u(0)$$

且

$$u'(t)=\gamma_0 e^{\gamma_0 t}\int_0^t e^{-2\gamma_0 t}\left[\int_0^s e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]ds+e^{-\gamma_0 t}\left[\int_0^t e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]+\gamma_0 e^{\gamma_0 t}u(0).$$

显见 $u'(0)=\gamma_0 u(0)+\widetilde{u}(0)$, 由此及 $u(0)=u(1), u'(0)=u'(1)$ 知

$$\begin{cases} u(0)=e^{\gamma_0}\int_0^1 e^{-2\gamma_0 t}\left[\int_0^s e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]ds+e^{\gamma_0}u(0) \\ \gamma_0 u(0)+\widetilde{u}(0)=\gamma_0 e^{\gamma_0}\int_0^1 e^{-2\gamma_0 s}\left[\int_0^s e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]ds+e^{-\gamma_0}\left[\int_0^t e^{\gamma_0\tau}v(\tau)d\tau+\widetilde{u}(0)\right]+\gamma_0 e^{\gamma_0}u(0).\end{cases}$$

可以验证上述方程组有唯一解 $u(0),\widetilde{u}(0)$, 因此存在 $u\in D(A)$ 使得 $v=(\lambda I-A)u$, 故 $R(\lambda I-A)=C[0,1]$, 从而 $\lambda\in\rho(A)$, 即 $(\mathbb{C}\backslash E)\subset\rho(A)$. □

例 6.2 设 $X=C[0,1]$, 则 $A:u(t)\mapsto tu(t)$ 是 X 上的有界线性算子, 且

$$\sigma(A)=\sigma_r(A)=[0,1].$$

证明 首先说明 A 是有界线性算子. 对任意 $u\in C[0,1]$, 注意到 $\|u\|_X:=\max_{t\in[0,1]}|u(t)|$, 因而

$$\|Au\|_X=\max_{t\in[0,1]}|tu(t)|\le\max_{t\in[0,1]}|u(t)|=\|u\|_X,$$

故 $\|A\|_{L(X)}\le 1$. 进一步取 $u\equiv 1$ 可知 $\|A\|_{L(X)}=1$, 这说明 $A\in L(X)$, 又因为 $D(A)=X=C[0,1]$ 是 X 中的闭集, 故 A 是闭算子.

下面计算 $\sigma(A)$. 根据 A 的构造, 对任意 $\lambda\in\mathbb{C}$ 与 $u\in C[0,1]$ 有

$$(\lambda I-A)u=(\lambda-t)u(t)$$

于是 $(\lambda I-A)u=0$ 意味着对任意 $t\in[0,1]$ 均有 $(\lambda-t)u(t)=0$. 当 $\lambda\neq t$ 时, 显见 $u(t)=0$; 当 $\lambda=t$ 时, 由 u 的连续性可知 $u(\lambda)=0$, 因此 $u\equiv 0$, 这说明 $\lambda I-A$ 是单射, 故 $(\lambda I-A)^{-1}$ 存在.

下面说明 $\lambda\notin[0,1]\Rightarrow\lambda\in\rho(A)$. 事实上, 若 $\lambda\notin[0,1]$, 则 $\min_{t\in[0,1]}|\lambda-t|>0$, 于是对 $v\in C[0,1]$ 与 $t\in[0,1]$, 令 $u(t):=(\lambda-t)^{-1}v(t)$ 可得

$$\|u\|_X=\max_{t\in[0,1]}\frac{|v(t)|}{\lambda-t|}\le\|v\|_X\frac{1}{\min\limits_{t\in[0,1]}|\lambda-t|}<\infty.$$

故 $u\in C[0,1]$ 且 $(\lambda I-A)u=v$, 由此可知 $R(\lambda I-A)=C[0,1]$, 故 $\lambda\in\rho(A)$.

再说明 $\lambda\in[0,1]\Rightarrow\lambda\in\sigma_r(A)$, 由剩余谱的定义知只需证明 $\overline{R(\lambda I-A)}\subsetneq C[0,1]$. 事实上, 对任意 $v\in R(\lambda I-A)$, 根据值域的定义知存在 $u\in C[0,1]$ 使得

$$v(t)=(\lambda I-A)u(t)=(\lambda-t)u(t).$$

故 $v(\lambda)=0$, 这说明 $v\in R(\lambda I-A)\to v(\lambda)=0$. 现注意到 $1\in C[0,1]$, 但 $1\notin\overline{R(\lambda I-A)}$. 若不然, 设 $1\in\overline{R(\lambda I-A)}$, 根据闭包的定义知存在 $\{v_n\}_{n\in\mathbb{N}}\subset R(\lambda I-A)$ 使得

$$\|v_n-1\|_X=\max_{t\in[0,1]}|v_n(t)-1|\to 0,\quad n\to\infty.$$

故对任意 $t\in[0,1]$ 均有 $|v_n(t)-1|\to 0(n\to\infty)$. 但另一方面, 根据前述推导可知对任意 $n\in\mathbb{N}$ 均有 $v_n(\lambda)=0$, 矛盾! 故 $1\notin\overline{R(\lambda I-A)}$, 从而 $\overline{R(\lambda I-A)}\subsetneq C[0,1]$, 亦即 $\lambda\in\sigma_r(A)$. □

例 6.3 设 $X=L^2[0,1]$, 则 $A:u(t)\mapsto tu(t)$ 是 X 上的有界线性算子, 且

$$\sigma(A)=\sigma_c(A)=[0,1].$$

证明 首先说明 A 是有界线性算子. 对任意 $u\in L^2[0,1]$, 因为

$$\|Au\|_{L^2[0,1]}:=\left(\int_0^1 t^2|u(t)|^2dt\right)^{\frac{1}{2}}\le\left(\int_0^1|u(t)|^2dt\right)^{\frac{1}{2}}=\|u\|_{L^2[0,1]},$$

故 $\|A\|_{L(L^2[0,1])}\le 1$, 因此 A 为 $L^2[0,1]$ 上的有界线性算子. 另因 $D(A)=L^2[0,1]$ 在 X 中闭, 故 A 是 X 上的闭算子.

再说明 $(\lambda I-A)^{-1}$ 存在, 为此只需说明 $\lambda I-A$ 是单射. 对任意 $\lambda\in\mathbb{C}$ 与 $u\in L^2[0,1]$, 若

$$(\lambda I-A)u(t)=(\lambda-t)u(t)=0_{L^2[0,1]},$$

在 $L^2[0,1]$ 意义下成立, 则对几乎处处 $t\in[0,1]$ 有 $(\lambda-t)u(t)=0$, 进而对几乎处处 $t\in[0,1]$ 有 $u(t)=0$, 亦即 $u=0_{L^2[0,1]}$, 这便说明了 $\lambda I-A$ 是单射.

下证 $\lambda\notin[0,1]\Rightarrow\lambda\in\rho(A)$, 这只需证明 $R(\lambda I-A)=L^2[0,1]$. 注意到 $\lambda\notin[0,1]$ 时有

$$\frac{1}{\min\limits_{t\in[0,1]}|\lambda-t|}<\infty,$$

故对任意 $v\in L^2[0,1]$, 令

$$u(t):=\frac{v(t)}{\lambda-t},\quad t\in[0,1],$$

则

$$\|u\|_{L^2[0,1]}\le\frac{1}{\min\limits_{t\in[0,1]}|\lambda-t|}\|v\|_{L^2[0,1]}<\infty,$$

故 $u\in L^2[0,1]$ 且 $(\lambda I-A)u=v$, 因此 $R(\lambda I-A)=L^2[0,1]$. 由此可知 $\lambda\in\rho(A)$.

下证 $\lambda\in[0,1]\Rightarrow\sigma_c(A)$, 根据连续谱的定义知只需证明 $R(\lambda I-A)\subsetneq L^2[0,1]$ 且 $\overline{R(\lambda I-A)}=L^2[0,1]$. 首先由

$$\int_0^1\frac{dt}{(\lambda-t)^2}\ge\max\left(\int_\lambda^1\frac{dt}{(\lambda-t)^2},\int_0^\lambda\frac{dt}{(\lambda-t)^2}\right)=\max\left(\int_0^{1-\lambda}\frac{dt}{t^2},\int_0^\lambda\frac{dt}{t^2}\right)=\infty$$

可知 $(\lambda-t)^{-1}\notin L^2[0,1]$, 因而 $1\notin R(\lambda I-A)$. 又因为 $1\in L^2[0,1]$, 故 $R(\lambda I-A)\subsetneq L^2[0,1]$.

再说明 $\overline{R(\lambda I-A)}=L^2[0,1]$, 为此设 $v\in L^2[0,1],\lambda\in[0,1]$, 定义

$$v_n:=\begin{cases}v(t), & |t-\lambda|\ge\frac{1}{n},t\in[0,1],\\ 0, & |t-\lambda|<\frac{1}{n},t\in[0,1],\end{cases}$$

则 $\|v_n\|_{L^2[0,1]}\le\|v\|_{L^2[0,1]}<\infty$, 且

$$\|v_n-v\|_{L^2[0,1]}=\left(\int_{\{t\in[0,1]:|t-\lambda|<\frac{1}{n}\}}|v(t)|^2dt\right)^{\frac{1}{2}}\to 0,\quad n\to\infty.$$

现定义

$$u_n(t):=\begin{cases}\frac{v_n}{\lambda-t}, & |t-\lambda|\ge\frac{1}{n},t\in[0,1],\\ 0, & |t-\lambda|<\frac{1}{n},t\in[0,1],\end{cases}$$

则

$$\|u_n\|_{L^2[0,1]}\le n\|v\|_{L^2[0,1]}<\infty,$$

即 $u_n\in L^2[0,1]$, 且 $(\lambda I-A)u_n=v_n$. 这说明 $v_n\in R(\lambda I-A), v\in\overline{R(\lambda I-A)}$, 结论即证. □

在说明清楚谱集 $\sigma(A)$ 所蕴含的三种可能性后, 下面来研究 $\sigma(A)$ 的性质. 在 $\dim X<\infty$ 时, 由线性代数理论可知对任意 $A\in L(X)$ 均有 $\sigma(A)\neq\varnothing$, 这是因为特征多项式 $|\lambda I-A|=0$ 根据代数基本定理是必定存在根的. 但在 $\dim X=\infty$ 时, 我们并没有类似于代数基本定理的结论, 故这套方法在无穷维时行不通了. 这个问题由 Gelfand 利用算子值函数的解析性得以解决, 下面介绍其方法.

定义 6.3 (预解式 (resolvent))

算子值函数 $R_\lambda(A):\rho(A)\to L(X)$ 定义为

$$\lambda\mapsto(\lambda I-A)^{-1},\quad\lambda\in\rho(A),$$

称 $R_\lambda(A)$ 为 A 的预解式. ♣

首先, 说明 $\rho(A)$ 是开集, 为此需要下述引理:

引理 6.1

设 $A \in L(X)$ 满足 $\|A\|_{L(X)} < 1$, 则 $(I-A)^{-1} \in L(X)$, 且

$$\|(I-A)^{-1}\|_{L(X)} \leq \frac{1}{1-\|A\|_{L(X)}}.$$

证明 首先证明 $I-A$ 是双射, 为此只需证明对任意 $y \in X$, 存在唯一 $x_y \in X$ 使得 $(I-A)x_y = y$. 任取 $x \in X$, 定义算子 $Sx := y + Ax$, 注意到对任意 $x_1, x_2 \in X$ 有

$$\|Sx_1 - Sx_2\|_X = \|Ax_1 - Ax_2\|_X \leq \|A\|_{L(X)}\|x_1 - x_2\|_X,$$

由 $\|A\|_{L(X)} < 1$ 知 S 为一个压缩映射, 故由压缩映射原理知存在唯一 $x_y \in X$ 使得 $Sx_y = x_y$, 即 $y + Ax_y = x_y$, 亦即 $y = (I-A)x_y$. 至此可知 $I-A$ 是双射, 从而 $(I-A)^{-1}$ 存在.

再说明 $(I-A)^{-1} \in L(X)$. 由 $\|A\|_{L(X)} < 1$ 知对 $\forall N, p \in \mathbb{N}$, 当 $N \to \infty$ 时有

$$\left\|\sum_{n=N}^{N+p} A^n\right\|_{L(X)} \leq \sum_{n=N}^{N+p} \|A\|_{L(X)}^n \to 0.$$

因此 $\left\{\sum_{n=0}^{N} A^n\right\}_{N\in\mathbb{N}}$ 是基本列. 因为 $L(X)$ 完备, 故 $\left\{\sum_{n=0}^{N} A^n\right\}_{N\in\mathbb{N}}$ 在 $L(X)$ 中收敛, 记 $\sum_{n=0}^{N} A^n \to \sum_{n=0}^{\infty} A^n$, 由此可知

$$\begin{aligned}\left\|\sum_{n=0}^{\infty} A^n\right\|_{L(X)} &\leq \left\|\sum_{n=0}^{\infty} A^n - \sum_{n=0}^{N} A^n\right\|_{L(X)} + \left\|\sum_{n=0}^{N} A^n\right\|_{L(X)} \\ &\leq \left\|\sum_{n=0}^{\infty} A^n - \sum_{n=0}^{N} A^n\right\|_{L(X)} + \sum_{n=0}^{N} \|A\|_{L(X)}^n.\end{aligned}$$

令 $N \to \infty$ 即得

$$\left\|\sum_{n=0}^{\infty} A^n\right\|_{L(X)} \leq \sum_{n=0}^{\infty} \|A\|_{L(X)}^n < \infty. \tag{6.11}$$

下面说明

$$(I-A)\left(\sum_{n=0}^{\infty} A^n\right) = I = \left(\sum_{n=0}^{\infty} A^n\right)(I-A).$$

为此, 设 $x \in X$, 由 $\sum_{n=0}^{\infty} A^n \in L(X)$ 可知 $k \to \infty$ 时:

$$\left\|\sum_{n=0}^{\infty} A^n x - \sum_{n=0}^{k} A^n x\right\|_X \leq \left\|\sum_{n=0}^{\infty} A^n - \sum_{n=0}^{k} A^n\right\|_{L(X)} \|x\|_X \to 0.$$

从而 $k \to \infty$ 时有

$$\sum_{n=0}^{k} A^n x =: y_k \to y := \sum_{n=0}^{\infty} A^n x, \quad k \to \infty \tag{6.12}$$

又由 $\|A\|_{L(X)} < 1$ 知当 $k \to \infty$ 时

$$\|A^{k+1}x\|_X \leq \|A\|_{L(X)}^{k+1}\|x\|_X \to 0,$$

由此及(6.12)式知 $k \to \infty$ 时有

$$(I-A)y_k = (I-A)\left(\sum_{n=0}^{k} A^n\right)x = x - A^{k+1}x \to x, \quad k \to \infty.$$

又由 $I-A \in L(X)$ 知 $(I-A)y = x$, 亦即 $(I-A)\left(\sum_{n=0}^{\infty} A^n\right) = I$. 类似可证

$$\left(\sum_{n=0}^{\infty} A^n\right)(I-A) = I.$$

因此

$$(I-A)^{-1} = \sum_{n=0}^{\infty} A^n. \tag{6.13}$$

从而

$$\|(I-A)^{-1}\|_{L(X)} \le \sum_{n=0}^{\infty} \|A\|_{L(X)}^n = \frac{1}{1-\|A\|_{L(X)}}.$$

至此引理即证. □

注

(i) 设 $A \in L(X), \lambda \in \mathbb{C}$ 且 $|\lambda| > \|A\|_{L(X)}$. 由引理6.1知 $(I-\lambda^{-1}A)^{-1} \in L(X)$, 且

$$\|(I-\lambda^{-1}A)^{-1}\|_{L(X)} \le \frac{1}{1-|\lambda|^{-1}\|A\|_{L(X)}},$$

此即 $(\lambda I - A)^{-1} \in L(X)$, 且

$$\|(\lambda I-A)^{-1}\|_{L(X)} \le \frac{1}{|\lambda|-\|A\|_{L(X)}}.$$

由此可知 $\lambda \in \rho(A)$, 即

$$\mathbb{C} \backslash \overline{B(0, \|A\|_{L(X)})} \subset \rho(A),$$

故 $\sigma(A) = \rho(A)^c \subset \overline{B(0, \|A\|_{L(X)})}$.

(ii) 在引理6.1的证明中, 因为 $Sx = y + Ax (\forall x \in X)$, 故 $Sy = y + Ay$, 进而

$$S^2 y = S(Sy) = y + A(Sy) = y + Ay + A^2 y.$$

更一般地, 对任意 $n \in \mathbb{N}$ 有

$$S^n y = \sum_{k=0}^{n} A^k y.$$

由此与压缩映射原理可知

$$x_y = \lim_{n\to\infty} S^n y = \sum_{k=0}^{\infty} A^k y,$$

这说明当 $\|A\|_{L(X)} < 1$ 时总有(6.13)式成立. 级数(6.13)称为 Neumann 级数.

(iii) Neumann 级数的形式非常像 $(1-x)^{-1}$ 的 Taylor 展开, 实际上后者就是前者的特殊形式. 令 $X := \mathbb{R}$, 对任意取定的 $x \in \mathbb{R}$, 令

$$T_x : \mathbb{R} \to \mathbb{R}, \quad y \mapsto xy.$$

显见 $\|T_x\|_{L(\mathbb{R})} = |x|$. 现在当 $\|T_x\|_{L(\mathbb{R})} = |x| < 1$ 时, 对 $\forall y \in \mathbb{R}$, 知 $(I - T_x)y = (1-x)y$, 从而

$$(I-T_x)^{-1} y = \frac{1}{1-x} y, \quad \|(I-T_x)^{-1}\|_{L(\mathbb{R})} = \frac{1}{1-x}.$$

此时 Neumann 级数(6.13)成为

$$\frac{1}{1-x} = (I-T_x)^{-1} = \sum_{n=0}^{\infty} T_x^n = \sum_{n=0}^{\infty} x^n, \quad |x| < 1$$

这正是 $(1-x)^{-1}$ 的 Taylor 展开.

推论 6.1 (预解集的开性)

设 X 是 Banach 空间, A 是闭线性算子, 则 $\rho(A)$ 是开集. ♡

证明 设 $\lambda_0 \in \rho(A)$, 注意到对任意 $\lambda \in \mathbb{C}$ 均有

$$\begin{aligned} \lambda I - A &= (\lambda - \lambda_0) I + (\lambda_0 I - A) \\ &= (\lambda_0 I - A)[I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]. \end{aligned} \tag{6.14}$$

又因为对任何具有有界逆的线性算子而言, 其自身与其逆的算子范数均不为零[6], 故当 $|\lambda - \lambda_0| < \|(\lambda_0 I - A)^{-1}\|_{L(X)}^{-1}$

[6]这是因为若 $\|(\lambda_0 I - A)^{-1}\|_{L(X)} = 0$, 则 $(\lambda_0 I - A)^{-1} = 0_{L(X)}$, 从而 $I = (\lambda_0 I - A)^{-1}(\lambda_0 I - A) = 0_{L(X)}$, 这与 X 的非平凡性 (即 $I \ne 0_{L(X)}$) 矛盾.

时, 有

$$\|(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}\|_{L(X)}<1,$$

进而由引理6.1知

$$B:=[I+(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}]^{-1}\in L(X).$$

由此可得

$$(\lambda I-A)^{-1}=B(\lambda_0 I-A)^{-1}\in L(X),$$

故 $\lambda\in\rho(A)$, 即 $B(\lambda_0,\|(\lambda_0 I-A)^{-1}\|_{L(X)}^{-1})\subset\rho(A)$, 推论得证. □

接下来说明 $R_\lambda(A)$ 是算子值解析函数[7], 为此首先考虑对 $R_\lambda(A)$ 求导, 引入下述第一预解公式:

引理 6.2 (第一预解公式)

设 $\lambda,\mu\in\rho(A)$, 则

$$R_\lambda(A)-R_\mu(A)=(\mu-\lambda)R_\lambda(A)R_\mu(A).$$

进一步有

$$R_\lambda(A)R_\mu(A)=R_\mu(A)R_\lambda(A).$$

证明 注意到对 $\lambda,\mu\in\rho(A)$ 有:

$$\begin{aligned}
R_\lambda(A)&=(\lambda I-A)^{-1}\\
&=(\lambda I-A)^{-1}(\mu I-A)(\mu I-A)^{-1}\\
&=(\lambda I-A)^{-1}[(\mu-\lambda)I+\lambda I-A](\mu I-A)^{-1}\\
&=(\mu-\lambda)(\lambda I-A)^{-1}(\mu I-A)^{-1}+(\mu I-A)^{-1}\\
&=(\mu-\lambda)R_\lambda(A)R_\mu(A)+R_\mu(A).
\end{aligned}$$

引理第一式即证. 将上式中的 λ,μ 互换并将两式相加即得引理的第二式. □

定理 6.2 (预解式的解析性)

若 X 是 Banach 空间, A 是闭线性算子, 则 A 的预解式 $R_\lambda(A)$ 在 $\rho(A)$ 内是算子值解析函数.

证明 首先说明 $R_\lambda(A)$ 关于 $\lambda\in\rho(A)$ 连续. 设 $\lambda_0\in\rho(A)$, 则当

$$|\lambda-\lambda_0|<\frac{1}{2\|R_{\lambda_0}(A)\|_{L(X)}}$$

时, 有

$$\|(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}\|_{L(X)}<\frac{1}{2}<1$$

此时由(6.14)式知

$$R_\lambda(A)=(\lambda I-A)^{-1}=[I+(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}]^{-1}R_{\lambda_0}(A)$$

进而由引理6.1知

$$\begin{aligned}
\|R_\lambda(A)\|_{L(X)}&\le\|R_{\lambda_0}(A)\|_{L(X)}\|[I+(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}]^{-1}\|_{L(X)}\\
&\le\frac{\|R_{\lambda_0}(A)\|_{L(X)}}{1-\|(\lambda-\lambda_0)(\lambda_0 I-A)^{-1}\|_{L(X)}}\\
&<2\|R_{\lambda_0}(A)\|_{L(X)}<\infty
\end{aligned}$$

[7] 设 X 是 Banach 空间, $G\subset\mathbb{C}$, 称抽象函数 $x(z):G\to X$ 解析, 如果对任意 $z_0\in G$, 在 X 中均存在极限 $\lim\limits_{z\to z_0}\frac{x(z)-x(z_0)}{z-z_0}$.

于是由第一预解公式6.2知, 当 $\lambda \to \lambda_0$ 时有:

$$\begin{aligned}\|R_\lambda(A) - R_{\lambda_0}(A)\|_{L(X)} &\leq |\lambda - \lambda_0| \|R_\lambda(A)\|_{L(X)} \|R_{\lambda_0}(A)\|_{L(X)} \\ &\leq 2|\lambda - \lambda_0| \|R_{\lambda_0}(A)\|_{L(X)}^2 \to 0\end{aligned}$$

因此 $R_\lambda(A)$ 关于 $\lambda \in \rho(A)$ 连续.

再说明 $R_\lambda(A)$ 的可微性. 由第一预解公式6.2可得

$$\frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} = -R_\lambda(A) R_{\lambda_0}(A)$$

在等式两边令 $\lambda \to \lambda_0$ 可得

$$\lim_{\lambda \to \lambda_0} \frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} = -\lim_{\lambda \to \lambda_0} R_\lambda(A) R_{\lambda_0}(A) = -R_{\lambda_0}(A)^2 \in L(X)$$

定理即证. □

下面考虑谱点的存在性定理:

定理 6.3 (Gelfand-Mazur)

若 X 是 Banach 空间, $A \in L(X)$, 则 $\sigma(A) \neq \varnothing$.

证明 用反证法. 若 $\rho(A) = \mathbb{C}$, 则由预解式的解析性6.2知 $R_\lambda(A)$ 在 $\mathbb{C}$ 上解析. 又当 $|\lambda| \leq 2\|A\|_{L(X)}$ 时, 由 $R_\lambda(A)$ 在 $\rho(A)$ 中的连续性与

$$\left| \|R_\lambda(A)\|_{L(X)} - \|R_{\lambda_0}(A)\|_{L(X)} \right| \leq \|R_\lambda(A) - R_{\lambda_0}(A)\|_{L(X)}$$

可知 $\|R_\lambda(A)\|$ 在 $\mathbb{C}$ 上连续, 进而它必在 $\overline{B(0, 2\|A\|_{L(X)})}$ 上连续. 又因为 $\overline{B(0, 2\|A\|_{L(X)})}$ 是欧氏空间中的紧集, 故 $\|R_\lambda(A)\|_{L(X)}$ 在 $\overline{B(0, 2\|A\|_{L(X)})}$ 上有界. 当 $|\lambda| > 2\|A\|_{L(X)}$ 时, 由级数(6.13)可得

$$R_\lambda(A) = (\lambda I - A)^{-1} = \frac{1}{\lambda}\left(I - \frac{A}{\lambda}\right)^{-1} = \frac{1}{\lambda} \sum_{n=0}^{\infty} \frac{A^n}{\lambda^n} = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n$$

另由引理6.1的注 (i) 知

$$\|R_\lambda(A)\|_{L(X)} \leq \frac{1}{|\lambda| - \|A\|_{L(X)}} < \frac{1}{\|A\|_{L(X)}}$$

故 $\|R_\lambda(A)\|_{L(X)}$ 在复平面上有上界, 记之为 M.

对任意 $f \in (L(X))^*$, 定义

$$u_f(\lambda) : \mathbb{C} \to \mathbb{C},\ \lambda \mapsto \langle f, R_\lambda(A) \rangle$$

其中 $\langle\ ,\ \rangle$ 表示 $((L(X))^*, L(X))$ 上的序偶. 下面说明 u_f 是 $\mathbb{C}$ 上的解析函数. 为此, 任取 $\lambda_0 \in \mathbb{C}$, 对任意 $\lambda \in \mathbb{C}$, 在 $\lambda \neq \lambda_0$ 时有:

$$\frac{u_f(\lambda) - u_f(\lambda_0)}{\lambda - \lambda_0} = \frac{\langle f, R_\lambda(A) \rangle - \langle f, R_{\lambda_0}(A) \rangle}{\lambda - \lambda_0} = \left\langle f, \frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} \right\rangle.$$

根据 $f \in (L(X))^*$ 所蕴含的 f 的连续性, 可知

$$\lim_{\lambda \to \lambda_0} \frac{u_f(\lambda) - u_f(\lambda_0)}{\lambda - \lambda_0} = \left\langle f, \lim_{\lambda \to \lambda_0} \frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} \right\rangle.$$

再由预解式的解析性6.2即知 u_f 在 $\mathbb{C}$ 上解析. 又注意到

$$|u_f(\lambda)| \leq \|f\|_{(L(X))^*} \|R_\lambda(A)\|_{L(X)} \leq M \|f\|_{(L(X))^*}$$

这说明 u_f 是 $\mathbb{C}$ 上的整函数, 根据 Liouville 定理可知存在与 λ 无关的常数 $C_f \in \mathbb{C}$ 使得

$$u_f(\lambda) = \langle f, R_\lambda(A) \rangle \equiv C_f$$

根据这件事可得 $R_\lambda(A)$ 是与 λ 无关的算子, 这是因为若设 $R_{\lambda_1}(A) \neq R_{\lambda_2}(A)$, 则由 Hahn-Banach 定理的推论知存在 $f \in (L(X))^*$ 使得 $\|f\|_{(L(X))^*} = 1$, 且

$$0 = u_f(\lambda_1) - u_f(\lambda_2) = \langle f, R_{\lambda_1}(A) - R_{\lambda_2}(A) \rangle \leq \|R_{\lambda_1}(A) - R_{\lambda_2}(A)\|_{L(X)}$$

因而 $R_{\lambda_1}(A)=R_{\lambda_2}(A)$, 矛盾! 因此 $R_\lambda(A)$ 是与 λ 无关的算子. 由此与第一预解公式6.2知当 $\lambda_1\neq\lambda_2$ 时有

$$(\lambda_2-\lambda_1)R_{\lambda_1}(A)R_{\lambda_2}(A)=R_{\lambda_1}(A)-R_{\lambda_2}(A)=0_{L(X)}$$

于是

$$I=\frac{1}{\lambda_2-\lambda_1}(\lambda_2 I-A)(\lambda_1 I-A)(\lambda_2-\lambda_1)R_{\lambda_1}(A)R_{\lambda_2}(A)=\frac{1}{\lambda_2-\lambda_1}(\lambda_2 I-A)(\lambda_1 I-A)0_{L(X)}=0_{L(X)}$$

矛盾! 因此 $\sigma(A)\neq\varnothing$, 定理得证. □

下面考虑有界线性算子谱集的范围. 由引理6.1的注 (i) 已经知道 $\sigma(A)\subset\overline{B(0,\|A\|_{L(X)})}$, 另由预解集的开性6.1与 Gelfand-Mazur 定理6.3知 $\sigma(A)$ 是非空紧集.

定义 6.4 (谱半径 (spectrum radius))

设 X 是 Banach 空间, A 是 X 上的有界线性算子, 称

$$r_\sigma(A):=\sup\{|\lambda|:\lambda\in\sigma(A)\}$$

为 A 的谱半径. ♣

注 从谱半径的定义立即可以得到下述结论:

(i) $\sigma(A)\subset\overline{B(0,r_\sigma(A))}$;

(ii) $r_\sigma(A)\le\|A\|_{L(X)}$.

其中 (ii) 是因为 $\sigma(A)\subset\overline{B(0,\|A\|_{L(X)})}$.

Gelfand 对谱半径给出了下述更精确的结论:

定理 6.4 (Gelfand)

设 X 是 Banach 空间, A 是 X 上的有界线性算子, 则

$$r_\sigma(A)=\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}.$$

♡

证明 首先说明极限 $\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$ 存在. 令

$$a:=\inf_{n\in\mathbb{N}}\|A^n\|_{L(X)}^{\frac{1}{n}},$$

知

$$a\le\liminf_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}=\lim_{k\to\infty}\inf_{n\ge k}\|A^n\|_{L(X)}^{\frac{1}{n}}. \tag{6.15}$$

由下确界的定义知, 对任意 $\varepsilon\in(0,\infty)$, 总存在 $n_0\in\mathbb{N}$ 使得

$$a\le\|A^{n_0}\|_{L(X)}^{\frac{1}{n_0}}<a+\varepsilon.$$

对任意 $n\ge n_0$, 总存在 $k\in\mathbb{N}^+$ 使得 $n=kn_0+s$, 其中 $s\in\mathbb{N}$ 且 $0\le s<n_0$, 由此可知

$$\|A^n\|_{L(X)}=\|A^{kn_0+s}\|_{L(X)}\le\|A^{n_0}\|_{L(X)}^k\|A\|_{L(X)}^s.$$

因此

$$\|A^n\|_{L(X)}^{\frac{1}{n}}\le\|A^{n_0}\|_{L(X)}^{\frac{k}{kn_0+s}}\|A\|_{L(X)}^{\frac{s}{n}}.$$

令 $n\to\infty$, 此时 $k\to\infty$, 于是

$$\begin{aligned}\limsup_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}&\le\limsup_{n\to\infty}\|A^{n_0}\|_{L(X)}^{\frac{k}{kn_0+s}}\|A\|_{L(X)}^{\frac{s}{n}}\\&=\lim_{n\to\infty}\|A^{n_0}\|_{L(X)}^{\frac{k}{kn_0+s}}\|A\|_{L(X)}^{\frac{s}{n}}\\&=\|A^{n_0}\|_{L(X)}^{\frac{1}{n_0}}<a+\varepsilon.\end{aligned}$$

由 ε 的任意性知

$$\limsup_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}\le a.$$

结合上式与(6.15)式即知

$$\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}=a.$$

下面说明 $r_\sigma(A)\le\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$. 为了证明逻辑的明晰, 首先介绍一下证明这件事的思路. 回顾 Cauchy-Hadamard 公式: 对于函数项级数

$$f(x):=\sum_{n=0}^{\infty}a_nx^n,\quad \forall x\in\mathbb{R},$$

令

$$R:=\frac{1}{\limsup\limits_{n\to\infty}|a_n|^{\frac{1}{n}}},$$

则当 $|x|<R$ 时 $f(x)$ 绝对收敛, $|x|>R$ 时 $f(x)$ 发散. 观察到 $\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$ 正是 $\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}$ 的形式, 而级数(6.13)也恰好有上面提到的函数项级数的形式, 通过变量代换可以猜测

$$(\lambda I-A)^{-1}=\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}},\tag{6.16}$$

现在要证明 $r_\sigma(A)\le\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$, 其实就是证明 $|\lambda|>\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$ 时 $(\lambda I-A)^{-1}\in L(X)$. 如果(6.16)式成立, 这就是要说明 $\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}\in L(X)$, 而让这件事成立的范数估计恰好是 Cauchy-Hadamard 公式的结果.

现在开始正式证明. 当 $|\lambda|>\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}$ 时, 有

$$\sum_{n=0}^{\infty}\frac{\|A^n\|_{L(X)}}{|\lambda|^n}<\infty \text{ 且 } \sum_{n=0}^{\infty}\frac{\|A^n\|_{L(X)}}{|\lambda|^{n+1}}<\infty.$$

因此

$$\sum_{n=0}^{\infty}\frac{A^n}{\lambda^n},\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}\in L(X).$$

故

$$\sum_{n=0}^{\infty}\frac{A^{n+1}}{\lambda^{n+1}}=\lim_{l\to\infty}\sum_{n=0}^{l}\frac{A^{n+1}}{\lambda^{n+1}}=\lim_{l\to\infty}A\sum_{n=0}^{l}\frac{A^n}{\lambda^{n+1}}.\tag{6.17}$$

注意到, 当 $l\to\infty$ 时,

$$\left\|A\sum_{n=0}^{l}\frac{A^n}{\lambda^{n+1}}-A\lim_{k\to\infty}\sum_{n=0}^{k}\frac{A^n}{\lambda^{n+1}}\right\|_{L(X)}\le\|A\|_{L(X)}\left\|\sum_{n=0}^{l}\frac{A^n}{\lambda^{n+1}}-\lim_{k\to\infty}\sum_{n=0}^{k}\frac{A^n}{\lambda^{n+1}}\right\|_{L(X)}\to 0.$$

由上式与(6.17)式进一步有

$$\sum_{n=0}^{\infty}\frac{A^{n+1}}{\lambda^{n+1}}=A\lim_{k\to\infty}\sum_{n=0}^{k}\frac{A^n}{\lambda^{n+1}}=A\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}\in L(X).$$

从而

$$\begin{aligned}(\lambda I-A)\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}&=\lambda\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}-A\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}\\&=\lambda\lim_{N\to\infty}\sum_{n=0}^{N}\frac{A^n}{\lambda^{n+1}}-A\lim_{N\to\infty}\sum_{n=0}^{N}\frac{A^n}{\lambda^{n+1}}\\&=\lim_{N\to\infty}\sum_{n=0}^{N}\frac{A^n}{\lambda^n}-\lim_{N\to\infty}\sum_{n=0}^{N}\frac{A^{n+1}}{\lambda^{n+1}}\\&=\sum_{n=0}^{\infty}\frac{A^n}{\lambda^n}-\sum_{n=0}^{\infty}\frac{A^{n+1}}{\lambda^{n+1}}=I,\end{aligned}$$

其中 I 表示 X 上的恒同算子. 类似有

$$\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}(\lambda I-A)=I,$$

因此

$$R_\lambda(A):=(\lambda I-A)^{-1}=\sum_{n=0}^{\infty}\frac{A^n}{\lambda^{n+1}}\in L(X).$$

故 $\lambda\in\rho(A)$, 从而

$$r_\sigma(A)\le\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}. \tag{6.18}$$

下面来证明

$$r_\sigma(A)=\lim_{n\to\infty}\|A^n\|_{L(X)}^{\frac{1}{n}}.$$

根据 $r_\sigma(A)$ 的定义, 当 $|\lambda|>r_\sigma(A)$ 时, 必有 $\lambda\in\rho(A)$, 即 $R_\lambda(A)\in L(X)$. 根据预解式的解析性6.2可知 $R_\lambda(A)$ 在 $|\lambda|>r_\sigma(A)$ 时解析. 下面证明 $R_\lambda(A)$ 在 $\lambda=\infty$ 处解析. 事实上, 若令 $\omega:=\frac{1}{\lambda}$, 则

$$\begin{aligned}R_\lambda(A)=(\lambda I-A)^{-1}&=\frac{1}{\lambda}\left(I-\frac{A}{\lambda}\right)^{-1}\\&=\omega(I-\omega A)^{-1}=:\widetilde{R}_\omega(A).\end{aligned}$$

现在要说明 $R_\lambda(A)$ 在 $\lambda=\infty$ 处可微, 就是要说明 $\widetilde{R}_\omega(A)$ 在 $\omega=0$ 处可微. 注意到 $\widetilde{R}_0(A)=0_{L(X)}$, 故

$$\frac{\widetilde{R}_\omega(A)-\widetilde{R}_0(A)}{\omega-0}=(I-\omega A)^{-1}.$$

又当 $|\omega|\|A\|_{L(X)}<\frac{1}{2}$ 时, 由引理6.1与级数(6.13)知

$$(I-\omega A)^{-1}=\sum_{n=0}^{\infty}(\omega A)^n=I+\omega A\sum_{n=0}^{\infty}(\omega A)^n,$$

且

$$\left\|\sum_{n=0}^{\infty}(\omega A)^n\right\|_{L(X)}\le\frac{1}{1-|\omega|\|A\|_{L(X)}}<2.$$

故当 $\omega\to0$ 时

$$0\le\|(I-\omega A)^{-1}-I\|_{L(X)}\le2|\omega|\|A\|_{L(X)}\to0.$$

这说明 $\omega\to0$ 时有

$$\frac{\widetilde{R}_\omega(A)-\widetilde{R}_0(A)}{\omega-0}\to I.$$

因此

$$\left.\frac{dR_\lambda(A)}{d\lambda}\right|_{\lambda=\infty}=\left.\frac{d\widetilde{R}_\omega(A)}{d\omega}\right|_{\omega=0}=I.$$

由此即知 $R_\lambda(A)$ 在 $\lambda=\infty$ 处解析, 再结合 $R_\lambda(A)$ 在 $|\lambda|>r_\sigma(A)$ 时解析即知 $\widetilde{R}_\omega(A)$ 在 $|\omega|<\frac{1}{r_\sigma(A)}$ 时关于 ω 解析.

任取 $f\in(L(X))^*$, 记

$$u_f(\lambda)=\langle f,R_\lambda(A)\rangle.$$

知

$$u_f(\lambda)=\langle f,R_\lambda(A)\rangle=\langle f,\omega(I-\omega A)^{-1}\rangle=:\widetilde{u}_f(\omega)$$

是开圆盘 $\{\omega\in\mathbb{C}:|\omega|<\frac{1}{r_\sigma(A)}\}$ 内关于 ω 的解析函数. 根据解析函数的 Taylor 定理可知

$$\widetilde{u}_f(\omega)=\sum_{n=0}^{\infty}c_n\omega^n,$$

其中

$$c_n := \frac{\widetilde{u}_f^{(n)}(0)}{n!}, \quad \forall n \in \mathbb{N}$$

且该展式唯一. 注意到

$$c_0 = \widetilde{u}_f^{(0)}(0) = \widetilde{u}_f(0) = 0,$$

故

$$\widetilde{u}_f(\omega) = \sum_{n=1}^{\infty} c_n \omega^n.$$

这说明当 $|\lambda| > r_\sigma(A)$ 时有

$$\begin{aligned} u_f(\lambda) = \widetilde{u}_f(\omega) &= \sum_{n=1}^{\infty} c_n \omega^n = \sum_{n=1}^{\infty} c_n \frac{1}{\lambda^n} \\ &= \sum_{n=0}^{\infty} \frac{c_{n+1}}{\lambda^{n+1}} =: \sum_{n=0}^{\infty} \frac{a_n}{\lambda_{n+1}}. \end{aligned}$$

因为 $|\lambda| > \lim_{n\to\infty} \|A^n\|_{L(X)}^{\frac{1}{n}}$ 时有

$$R_\lambda(A) = \sum_{n=0}^{\infty} \frac{A^n}{\lambda^{n+1}} \in L(X),$$

故由 $f \in (L(X))^*$ 所蕴含的连续性可知

$$u_f(\lambda) = \langle f, R_\lambda(A) \rangle = \sum_{n=0}^{\infty} \frac{\langle f, A^n \rangle}{\lambda^{n+1}}.$$

根据 Laurent 展式的唯一性, 对 $\forall n \in \mathbb{N}$ 均有 $a_n = f(A^n)$, 因此当 $|\lambda| > r_\sigma(A)$ 时也应该有

$$u_f(\lambda) = \sum_{n=0}^{\infty} \frac{\langle f, A^n \rangle}{\lambda^{n+1}}.$$

因为 Laurent 展式在其收敛区域内必绝对收敛, 故对任意 $\varepsilon \in (0, \infty)$ 与任意 $f \in (L(X))^*$, 由 $r_\sigma(A) + \varepsilon > r_\sigma(A)$ 知

$$\sum_{n=0}^{\infty} \frac{|\langle f, A^n \rangle|}{(r_\sigma(A) + \varepsilon)^{n+1}} = \sum_{n=0}^{\infty} \left| \left\langle f, \frac{A^n}{(r_\sigma(A) + \varepsilon)^{n+1}} \right\rangle \right| < \infty.$$

现在对 $\forall n \in \mathbb{N}$, 把 $\frac{A^n}{(r_\sigma(A)+\varepsilon)^{n+1}}$ 看成 $(L(X))^{**}$ 中的元素, 则对 $\forall f \in (L(X))^*$ 有

$$\left| \left\langle f, \frac{A^n}{(r_\sigma(A) + \varepsilon)^{n+1}} \right\rangle \right| \le \sum_{n=0}^{\infty} \left| \left\langle f, \frac{A^n}{(r_\sigma(A) + \varepsilon)^{n+1}} \right\rangle \right| < \infty,$$

这说明 $\left\langle f, \frac{A^n}{(r_\sigma(A)+\varepsilon)^{n+1}} \right\rangle$ 关于 $n \in \mathbb{N}$ 一致有界, 根据典则嵌入的等距性与共鸣定理即知存在 $M > 0$ 使得对任意 $n \in \mathbb{N}$ 均有

$$\frac{\|A^n\|_{L(X)}}{(r_\sigma(A) + \varepsilon)^{n+1}} = \left\| \frac{A^n}{(r_\sigma(A) + \varepsilon)^{n+1}} \right\|_{(L(X))^{**}} \le M.$$

因此对任意 $n \in \mathbb{N}$ 有

$$\|A^n\|_{L(X)}^{\frac{1}{n}} \le M^{\frac{1}{n}} (r_\sigma(A) + \varepsilon)^{\frac{n+1}{n}},$$

令 $n \to \infty$ 可得

$$\varlimsup_{n\to\infty} \|A^n\|_{L(X)}^{\frac{1}{n}} \le r_\sigma(A) + \varepsilon,$$

由 ε 的任意性可得

$$\varlimsup_{n\to\infty} \|A^n\|_{L(X)}^{\frac{1}{n}} \le r_\sigma(A).$$

结合(6.18)式, 这便得到了

$$r_\sigma(A) = \lim_{n\to\infty} \|A^n\|_{L(X)}^{\frac{1}{n}}.$$

此即欲证. □

6.1.1 习题

✍ **练习 6.1** 设 X 是有限维 Banach 空间, $A: X \to X$ 是有界线性算子, 则对 $\forall \lambda \in \mathbb{C}$, λ 要么是 A 的正则值, 要么是 A 的特征值.

证明 若 $\dim X < \infty, A \in L(X)$, 则 A 完全由一个矩阵确定. 现在对于 $\lambda \in \mathbb{C}$, 要么 $|\lambda I - A| = 0$, 要么 $|\lambda I - A| \neq 0$, 其中 $|\cdot|$ 是行列式函数, I 是单位矩阵. 对于 $|\lambda I - A| = 0$ 的情况, 知存在 $x_0 \in X \backslash \{0_X\}$ 使得 $(\lambda I - A)x_0 = 0_X$, 因此 λ 是 A 的特征值. 对于 $|\lambda I - A| \neq 0$ 的情况, 知此时 $(\lambda I - A)^{-1}$ 存在, 故 λ 是 A 的正则值. □

下面的习题旨在研究微分算子的谱. 我们首先回忆弱导数与 Sobolev 空间的相关概念. 设 $\Omega \subset \mathbb{R}^n$ 是连通开集. 对于定义在 Ω 上的函数 u 而言, 若对任意 $x \in \Omega$ 均存在 $r \in (0, \infty)$ 使得 $u\chi_{B(x,r)} \in L^1(\Omega)$, 就称 u 在 Ω 上局部可积, 记 Ω 上的全体局部可积函数构成的集合为 $L^1_{loc}(\Omega)$. 对于 $L^1_{loc}(\Omega)$ 中的元素, 我们可以定义弱导数: 对于 $u \in L^1_{loc}(\Omega)$ 与多重指标 $\alpha \in \mathbb{N}^n$, 若存在 $g \in L^1_{loc}(\Omega)$ 满足对任意 $\varphi \in C_c^\infty(\Omega)$ 有

$$\int_\Omega g(x)\varphi(x)dx = (-1)^{|\alpha|} \int_\Omega u(x)D^\alpha \varphi(x)dx,$$

就称 g 是 u 的 α 阶弱导数, 记作 $g = \widetilde{D}^\alpha u$, 其中

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}.$$

对任意 $m \in \mathbb{N}$, 令

$$H^m(\Omega) = \{u \in L^2(\Omega) : \widetilde{D}^\alpha u \in L^2(\Omega), \forall \alpha \in \mathbb{N}^n, |\alpha| \leq m\}.$$

称 $H^m(\Omega)$ 为 Sobolev 空间. 对 $u \in H^m(\Omega)$, 记

$$\|u\|_{H^m(\Omega)} = \left[\sum_{|\alpha| \leq m} \|\widetilde{D}^\alpha u\|^2_{L^2(\Omega)}\right]^{\frac{1}{2}}.$$

方便起见, 我们再回忆一下 Fourier 变换与 Sobolev 空间的 Fourier 刻画. 任取 $f \in L^1(\mathbb{R}^n), \xi \in \mathbb{R}^n$, 记

$$\mathcal{F}[f](\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)e^{-i2\pi x \cdot \xi}dx$$

称 $\widehat{f}$ 为 f 的 Fourier 变换. 通过 Hahn-Banach 定理可以说明 $\mathcal{F}$ 可以延拓为 $L^2 \to L^2$ 的有界线性算子, 且有 Plancherel 公式成立:

$$\|\widehat{f}\|_{L^2(\mathbb{R}^n)} = \|f\|_{L^2(\mathbb{R}^n)}.$$

✍ **练习 6.2** 本题说明 Fourier 变换的乘法公式与 Sobolev 空间的 Fourier 刻画.

(i) 对任意 $f, g \in L^2(\mathbb{R}^n)$ 有

$$\int_{\mathbb{R}^n} \widehat{f}(x)g(x)dx = \int_{\mathbb{R}^n} f(x)\widehat{g}(x)dx.$$

(ii) 设 $m \in \mathbb{N}^*$, 则 $f \in H^m(\mathbb{R}^n)$ 当且仅当 $f \in L^2(\mathbb{R}^n)$ 且 $(1 + |\cdot|^2)^{\frac{m}{2}}\widehat{f} \in L^2(\mathbb{R}^n)$. 另外, 存在正常数 C_1, C_2 使得

$$C_1\|f\|_{H^m(\mathbb{R}^n)} \leq \|(1 + |\cdot|^2)^{\frac{m}{2}}\widehat{f}\|_{L^2(\mathbb{R}^n)} \leq C_2\|f\|_{H^m(\mathbb{R}^n)}.$$

证明 (i) 根据 Plancherel 公式首先可以说明 $\widehat{f} \cdot g \in L^1, f \cdot \widehat{g} \in L^1$, 因此根据 Fubini 定理知

$$\begin{aligned}
\int_{\mathbb{R}^n} \widehat{f}(x)g(x)dx &= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} f(y)e^{-i2\pi y \cdot x}dy\right]g(x)dx \\
&= \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} g(x)e^{-i2\pi y \cdot x}dx\right]f(y)dy \\
&= \int_{\mathbb{R}^n} f(x)\widehat{g}(x)dx
\end{aligned}$$

✍ **练习 6.3** 考虑 $L^2(0, \infty)$ 上的微分算子

$$A : x(t) \mapsto x'(t), \quad D(A) := H^1(0, \infty).$$

则 $\sigma_p(A)=\{\lambda\in\mathbb{C}:\operatorname{Re}\lambda<0\},\sigma_c(A)=\{\lambda\in\mathbb{C}:\operatorname{Re}\lambda=0\},\sigma_r(A)=\varnothing$.

证明 我们首先说明 $\{\lambda\in\mathbb{C}:\operatorname{Re}\lambda<0\}\subset\sigma_p(A)$. 当 $\operatorname{Re}\lambda<0$ 时, 取 $x(t)=e^{\lambda t}$, 显见 $x\in L^2(0,\infty)$, 且

$$(\lambda I-A)x(t)=\lambda e^{\lambda t}-\lambda e^{\lambda t}=0$$

因此 $(\lambda I-A)^{-1}$ 不存在, 即 $\lambda\in\sigma_p(A)$.

6.2 紧算子

定义 6.5 (紧算子 (compact operator))

设 X,Y 是赋范线性空间, 称线性算子 $A:X\to Y$ 为紧算子, 如果下述条件之一成立:

(i) $\overline{A(B_1)}$ 是 Y 中 (强拓扑下) 的紧集, 其中 $B_1:=\{x\in X:\|x\|_X<1\}$ 为 X 中的单位球.

(ii) 对 X 中任意有界集 B 而言, $\overline{A(B)}$ 是 Y 中 (强拓扑下) 的紧集.

(iii) 对 X 中任意有界点列 $\{x_n\}_{n\in\mathbb{N}}$ 而言, $\{Ax_n\}_{n\in\mathbb{N}}$ 在 Y 中 (强拓扑意义下) 有收敛子列.

记 $X\to Y$ 的紧算子全体为 $\mathfrak{C}(X,Y)$, 特别当 $X=Y$ 时记之为 $\mathfrak{C}(X)$.

下面说明上述定义中的 (i),(ii),(iii) 彼此等价, 为此我们回忆度量空间中的下述结论:

引理 6.3

度量空间 X 的子集 M 是紧集当且仅当它是自列紧集, 亦即 M 中的任意点列在 X 中都有一个收敛子列, 且该子列收敛到 M 中的某点.

引理 6.4

若度量空间 X 的子集 M 是列紧集, 则 $\overline{M}$ 是自列紧集.

证明 记 X 的度量为 d. 设 $\{x_n\}_{n\in\mathbb{N}}\subset\overline{M}$, 则对 $\forall n\in\mathbb{N}$ 而言存在 $y_n\in M$ 使得

$$d(x_n,y_n)<\frac{1}{n}. \tag{6.19}$$

因为 M 是列紧集且 $\{y_n\}_{n\in\mathbb{N}}\subset M$, 故 $\{y_n\}_{n\in\mathbb{N}}$ 有收敛子列, 记之为 $\{y_{n_k}\}_{k\in\mathbb{N}}$. 不妨设

$$\lim_{k\to\infty}y_{n_k}=y_0\in X.$$

根据极限定义可知

$$\forall\varepsilon\in(0,\infty)\exists k_0\in\mathbb{N}\forall k>k_0\left(d(y_{n_k},y_0)<\frac{\varepsilon}{2}\right).$$

取 $k_1\in\mathbb{N}$ 使得 $k>k_1$ 时 $\frac{1}{n_k}<\frac{\varepsilon}{2}$, 则当 $k>\max\{k_0,k_1\}$ 时由(6.19)式知

$$d(x_{n_k},y_0)\le d(x_{n_k},y_{n_k})+d(y_{n_k},y_0)<\frac{1}{n_k}+\frac{\varepsilon}{2}<\varepsilon.$$

因此 $\lim_{k\to\infty}x_{n_k}=y_0$. 又因为 $\overline{M}$ 是闭集, 故 $y_0\in\overline{M}$, 这说明 $\overline{M}$ 是自列紧集. □

另外, 还需回忆定理1.14:

引理 6.5

紧 Hausdorff 空间 X 中子集的闭性与紧性等价.

现在开始正式说明定义6.5中的 (i),(ii),(iii) 彼此等价.

(i)⟹(ii): 对 X 中的任意有界集 B, 根据有界的定义知存在 $M>0$ 使得

$$B\subset MB_1:=\{Mx:x\in B_1\},$$

其中 B_1 表示 X 中的单位球. 由 A 的线性性可知

$$\begin{aligned} x \in \overline{MA(B_1)} &\Leftrightarrow \exists \{x_n\}_{n\in\mathbb{N}} \subset MA(B_1), x_n \to x (n\to\infty) \\ &\Leftrightarrow \exists \left\{\frac{x_n}{M}\right\}_{n\in\mathbb{N}} \subset A(B_1), \frac{x_n}{M} \to \frac{x}{M} (n\to\infty) \\ &\Leftrightarrow \frac{x}{M} \in \overline{A(B_1)} \Leftrightarrow x \in M\overline{A(B_1)}, \end{aligned}$$

这说明 $\overline{A(MB_1)} = M\overline{A(B_1)}$, 从而由 $\overline{A(B_1)}$ 紧与引理6.3知 $\overline{A(MB_1)}$ 紧. 又因为

$$A(B) \subset A(MB_1) \Rightarrow \overline{A(B)} \subset \overline{A(MB_1)},$$

故由引理6.5知 $\overline{A(B)}$ 紧, 此即定义6.5(ii).

(ii)⇒(iii): 设 $\{x_n\}_{n\in\mathbb{N}}$ 是 X 中的有界点列, 由 (ii) 知 $\overline{\{Ax_n\}_{n\in\mathbb{N}}}$ 是 Y 中的紧集, 从而由引理6.3知 $\overline{\{Ax_n\}_{n\in\mathbb{N}}}$ 是自列紧集, 因此 $\{Ax_n\}_{n\in\mathbb{N}}$ 在 Y 中有收敛子列.

(iii)⇒(i): 根据引理6.3, 要说明 $\overline{A(B_1)}$ 是 Y 中的紧集, 只需说明 $\overline{A(B_1)}$ 是自列紧集. 现任取 $A(B_1)$ 中的点列 $\{y_n\}_{n\in\mathbb{N}}$, 根据 $A(B_1)$ 的定义知存在 $\{x_n\}_{n\in\mathbb{N}} \subset B_1$ 使得对 $\forall n\in\mathbb{N}$ 有

$$y_n = Ax_n.$$

因为 $\{x_n\}_{n\in\mathbb{N}} \subset B_1$, 故 $\{x_n\}_{n\in\mathbb{N}}$ 至少是有界集, 进而由 (iii) 知 $\{y_n\}_{n\in\mathbb{N}}$ 有收敛子列, 因此 $A(B_1)$ 是列紧的, 进一步由引理6.4知 $\overline{A(B_1)}$ 自列紧.

至此即知定义6.5的 (i),(ii),(iii) 彼此等价. □

注 对 B^* 空间 Y 而言, 若集合 $E\subset Y$ 满足 $\overline{E}$ 是 Y 中的紧集, 就称 E 在 Y 中预紧. 回忆

$$\text{线性算子 } T \text{ 是有界算子} \Leftrightarrow T \text{ 把有界集映成有界集}^{8},$$

而定义6.5表明

$$\begin{aligned} \text{线性算子 } T \text{ 为紧算子} &\Leftrightarrow T \text{ 把有界集映成预紧集} \\ &\Leftrightarrow T \text{ 把有界集映成列紧集}. \end{aligned}$$

在范数诱导的拓扑下, 显见预紧集首先是有界集, 因此紧算子必为有界算子. 下面说明有界线性算子成为紧算子的一个充分条件.

命题 6.2

设 X, Y 是赋范线性空间, 若 $A\in L(X,Y)$ 且 $\dim Y<\infty$, 则 $A\in\mathfrak{C}(X,Y)$.

证明 设 B_1 是 X 中的开单位球, 由 $A\in L(X,Y)$ 知 $\overline{A(B_1)}$ 是 Y 中的有界闭集. 因为 $\dim Y<\infty$, 而有限维空间中的有界闭集就是紧集, 故 $\overline{A(B_1)}$ 是 Y 中的紧集, 由定义6.5(i) 即得欲证. □

另外, 紧算子定义中的开单位球可以改为单位球面, 此即下述命题:

命题 6.3 (紧算子的等价定义)

设 X, Y 是赋范线性空间, 则 $A: X\to Y$ 为紧算子等价于下述任一命题:

(i) $\overline{A(S_1)}$ 是 Y 中的紧集, 其中 $S_1 := \{x\in X : \|x\|_X = 1\}$ 是 X 中的单位球面;

(ii) 对 X 中单位球面 S_1 中的任意点列 $\{x_n\}_{n\in\mathbb{N}}$ 而言, $\{Ax_n\}_{n\in\mathbb{N}}$ 在 Y 中均有收敛子列.

证明 我们用 A 是紧算子 ⇒(i)⇒(ii)⇒ A 是紧算子这一思路进行证明.

当 A 是紧算子时, 因为 S_1 是有界集, 根据定义6.5(ii) 即知 $\overline{A(S_1)}$ 是紧集, 此即 (i).

当 (i) 成立时, 任取 S_1 中的点列 $\{x_n\}_{n\in\mathbb{N}}$. 因为 $\overline{A(S_1)}$ 是紧集, 由引理6.3可知 $\overline{A(S_1)}$ 是自列紧集, 因此 $\{Ax_n\}_{n\in\mathbb{N}}$ 在 Y 中有收敛子列, 此即 (ii).

当 (ii) 成立时, 任取 X 中的有界点列 $\{x_n\}_{n\in\mathbb{N}}$, 考虑用定义6.5(iii) 证明 A 是紧算子, 这就希望把 $\{x_n\}_{n\in\mathbb{N}}$ 拉到

[8] ⇒ 显然, 对于 ⇐, 记 $T: X\to Y$, 且存在 $b>0$ 使得 $TB_X(0_X,1)\subset B_Y(0_Y,b)$, 则任取 $0_X\neq x\in X$, 由 $x/2\|x\|_X\in B_X(0_X,1)$ 即得 $\|Tx\|_Y\le 2b\|x\|_X$.

S_1 上. 为此首先排除 $x_n = 0_X$ 的情况. 若存在 $\mathbb{N}$ 的子列 $\{n_i\}_{i\in\mathbb{N}}$ 使得对 $\forall i \in \mathbb{N}$ 有 $x_{n_i} = 0_X$, 则由 A 的线性性知

$$Ax_{n_i} = 0_Y, \quad \forall i \in \mathbb{N}.$$

此时 $\{Ax_{n_i}\}_{i\in\mathbb{N}}$ 正是 $\{Ax_n\}_{n\in\mathbb{N}}$ 的收敛子列. 现设对每个 $n \in \mathbb{N}$ 而言均有 $x_n \neq 0_X$, 因为 $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ 是 $(0,\infty)$ 中的有界点列, 根据实数的极限点引理知该序列必有收敛子列, 不妨设 $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ 自身收敛到 $a \in [0,\infty)$. 因为 $\{x_n/\|x_n\|_X\}_{n\in\mathbb{N}}$ 是 S_1 中的点列, 故根据假设知 $\{A(x_n/\|x_n\|_X)\}_{n\in\mathbb{N}}$ 在 Y 中有收敛子列 $\{A(x_{n_i}/\|x_{n_i}\|_X)\}_{i\in\mathbb{N}}$, 不妨设

$$A\left(\frac{x_{n_i}}{\|x_{n_i}\|_X}\right) \to y, \quad i \to \infty.$$

注意到

$$Ax_{n_i} - ay = \|x_{n_i}\|_X \left[A\left(\frac{x_{n_i}}{\|x_{n_i}\|_X}\right) - y\right] + (\|x_{n_i}\|_X - a)y$$

因为 $\{\|x_{n_i}\|_X\}_{i\in\mathbb{N}}$ 有界, 且 $\|x_{n_i}\|_X \to a(i\to\infty)$, 故 $Ax_{n_i} \to ay(i\to\infty)$, 这说明 $\{Ax_n\}_{n\in\mathbb{N}}$ 此时也有收敛子列, 根据定义6.5即知 A 是紧算子. □

紧算子有下述基本性质 (其中空间上的拓扑默认为强拓扑):

命题 6.4 (紧算子的基本性质)

若 X, Y, Z 为赋范线性空间, 则下述命题成立:

(i) $\mathfrak{C}(X,Y) \subset L(X,Y)$.

(ii) 设 $A, B \in \mathfrak{C}(X,Y), \alpha, \beta \in \mathbb{C}$, 则 $\alpha A + \beta B \in \mathfrak{C}(X,Y)$.

(iii) 若 Y 是 Banach 空间, 则 $\mathfrak{C}(X,Y)$ 是 $L(X,Y)$ 的闭线性子空间.

(iv) 设 $A \in \mathfrak{C}(X,Y)$, X_0 是 X 的一个线性子空间, 则 $A|_{X_0} \in \mathfrak{C}(X_0,Y)$.

(v) 若 $A \in \mathfrak{C}(X,Y)$, 则 $R(A) := \{Ax : x \in X\}$ 可分.

(vi) 若 $A \in L(X,Y), B \in L(Y,Z)$ 其中之一是紧算子, 则

$$B \circ A \in \mathfrak{C}(X,Z).$$

下面逐条进行证明.

证明 (i) 要证明 $\mathfrak{C}(X,Y) \subset L(X,Y)$, 设 $A \in \mathfrak{C}(X,Y)$, B_1 为 X 中的单位球, 则

$$\|A\|_{L(X,Y)} = \sup_{x\in B_1}\|Ax\|_Y = \sup_{y\in A(B_1)}\|y\|_Y.$$

下面断言

$$\sup_{y\in A(B_1)}\|y\|_Y = \sup_{y\in\overline{A(B_1)}}\|y\|_Y. \tag{6.20}$$

事实上, 显见 $\sup_{y\in A(B_1)}\|y\|_Y \leq \sup_{y\in\overline{A(B_1)}}\|y\|_Y$. 现在任取 $y \in \overline{A(B_1)}$, 根据闭包的定义知存在 $\{y_n\}_{n\in\mathbb{N}} \subset A(B_1)$ 使得 $n\to\infty$ 时 $\|y_n - y\|_Y \to 0$. 故当 $n \to \infty$ 时有

$$|\|y_n\|_Y - \|y\|_Y| \leq \|y_n - y\|_Y \to 0.$$

因此

$$\|y\|_Y = \lim_{n\to\infty}\|y_n\|_Y \leq \sup_{y\in A(B_1)}\|y\|_Y.$$

这说明

$$\sup_{y\in\overline{A(B_1)}}\|y\|_Y \leq \sup_{y\in A(B_1)}\|y\|_Y.$$

此即(6.20)式. 因为 $A \in \mathfrak{C}(X,Y)$, 根据定义6.3(i) 知 $\overline{A(B_1)}$ 是 Y 中的紧集, 而 $y \mapsto \|y\|_Y$ 是 Y 上的连续函数, 故该函数必在 $\overline{A(B_1)}$ 上至少有界, 亦即

$$\|A\|_{L(X,Y)} = \sup_{y\in A(B_1)}\|y\|_Y = \sup_{y\in\overline{A(B_1)}}\|y\|_Y < \infty.$$

此即 (i). 这条结论说明 $\mathfrak{C}(X,Y)$ 上可赋 $L(X,Y)$ 的子空间拓扑, (iii) 提到的闭性是该拓扑下的.

(ii) 因为收敛列的线性组合依旧收敛, 故由定义6.3(iii) 立得 (ii).

注 由命题6.2与基本性质6.4(i) 的证明知, 若 X 为线性赋范空间, Y 为有限维线性赋范空间, 则

$$A \in L(X,Y) \Leftrightarrow A \in \mathfrak{C}(X,Y).$$

为证明 (iii), 我们先回忆下述概念与引理:

定义 6.6 (ε 网, 完全有界 (ε net, totally bounded))

设 M 是度量空间 X 的子集.

(a) 设 $\varepsilon \in (0,\infty)$, $N \subset M$. 若对 $\forall x \in M$ 均 $\exists y \in N$ 使得 $d(x,y) < \varepsilon$, 就称 N 是 M 的一个 ε 网. 等价地, N 为 M 的一个 ε 网 $\Leftrightarrow N \subset M \subset \bigcup_{y\in N} B(y,\varepsilon)$.

(b) 若对 $\forall \varepsilon \in (0,\infty)$, 均存在 M 的有限 ε 网, 就称 M 完全有界.

♣

引理 6.6 (Hausdorff)

设 X 是度量空间, $M \subset X$. 若 M 列紧, 则 M 完全有界. 另若 X 完备, 则 M 列紧 $\Leftrightarrow M$ 完全有界.

♡

下面证明 (iii). 因为 $L(X,Y)$ 完备, 故由 (i) 知

$$\overline{\mathfrak{C}(X,Y)} \subset L(X,Y).$$

现设 $T \in \overline{\mathfrak{C}(X,Y)}$, 往证 $T \in \mathfrak{C}(X,Y)$, 由定义6.5(iii) 知只需证 $\overline{T(B_1)}$ 在 Y 中紧, 根据引理6.3知这等价于说明 $\overline{T(B_1)}$ 在 Y 中自列紧. 注意 $\overline{T(B_1)}$ 显然是 Y 中 (在强拓扑意义下) 的闭集, 故只需说明 $\overline{T(B_1)}$ 列紧, 根据 Hausdorff 定理 (引理6.6) 与 Y 的完备性知这等价于说明 $\overline{T(B_1)}$ 完全有界. 为说明 $\overline{T(B_1)}$ 完全有界, 对 $\forall \varepsilon \in (0,\infty)$ 我们就需要找到 $\overline{T(B_1)}$ 的一个有限 ε 网. 根据闭包的定义, 知存在紧算子列 $\{T_n\}_{n\in\mathbb{N}} \subset \mathfrak{C}(X,Y)$ 使得

$$\lim_{n\to\infty} \|T_n - T\|_{L(X,Y)} = 0. \tag{6.21}$$

由此进一步知对 $\forall \varepsilon \in (0,\infty)$ 均 $\exists n \in \mathbb{N}$ 使得

$$\|T_n - T\|_{L(X,Y)} < \frac{\varepsilon}{8}. \tag{6.22}$$

因为 $\{T_n\}_{n\in\mathbb{N}} \subset \mathfrak{C}(X,Y)$, 故 $\overline{T_n(B_1)}$ 在 Y 中紧 (其中 B_1 为 X 中的单位球), 根据引理6.3知 $\overline{T_n(B_1)}$ 自列紧, 进而由 Hausdorff 定理6.6知 $\overline{T_n(B_1)}$ 完全有界. 现在考虑 $\overline{T_n(B_1)}$ 的 $\varepsilon/8$ 网, 知存在

$$\{y_1,\cdots,y_m\} \subset \overline{T_n(B_1)}$$

使得

$$\overline{T_n(B_1)} \subset \bigcup_{i=1}^m B\left(y_i, \frac{\varepsilon}{8}\right). \tag{6.23}$$

注意到对 $\forall y \in \overline{T(B_1)}$, 总 $\exists x \in B_1$ 使得 $\|y - Tx\|_Y < \frac{\varepsilon}{8}$, 于是由(6.22)式知

$$\|y - T_n x\|_Y \le \|y - Tx\|_Y + \|Tx - T_n x\|_Y < \frac{\varepsilon}{4}.$$

又由(6.23)式知 $\exists i \in \{1,\cdots,m\}$ 使得

$$\|T_n x - y_i\|_Y < \frac{\varepsilon}{8},$$

因此

$$\|y - y_i\|_Y \le \|y - T_n x\|_Y + \|T_n x - y_i\|_Y < \frac{\varepsilon}{2},$$

这说明

$$\overline{T(B_1)} \subset \bigcup_{i=1}^m B\left(y_i, \frac{\varepsilon}{2}\right).$$

但这还不能导出 $y_i \in \overline{T(B_1)}$, 下面分两种情况进行讨论:

(a) 若 $B(y_i, \frac{\varepsilon}{2}) \cap \overline{T(B_1)} = \varnothing$, 就在 $\{y_1,\cdots,y_m\}$ 中去掉 y_i.

(b) 若 $B(y_i, \frac{\varepsilon}{2}) \cap \overline{T(B_1)} \ne \varnothing$, 则取 $z_i \in B(y_i, \frac{\varepsilon}{2}) \cap \overline{T(B_1)}$. 现对任意 $z \in B(y_i, \frac{\varepsilon}{2}) \cap \overline{T(B_1)}$ 知总有 $z \in B(z_i, \varepsilon)$. 此时

用 z_i 替代 y_i 可得

$$\{z_1, \cdots, z_l\} \subset \overline{T(B_1)}$$

其中 $1 \le l \le m$ 满足

$$\overline{T(B_1)} \subset \bigcup_{i=1}^{l} B(z_i, \varepsilon).$$

根据上述讨论知 $\overline{T(B_1)}$ 有有限 ε 网, 这便完成了 (iii) 的证明.

对于 (iv), 因为 X_0 中的有界点列 $\{x_n\}_{n\in\mathbb{N}}$ 同时是 X 中的有界点列, 且对 $\forall n \in \mathbb{N}$ 有

$$A|_{X_0} x_n = A x_n,$$

故根据定义6.5(iii) 知 $\{A|_{X_0} x_n\}_{n\in\mathbb{N}} = \{A x_n\}_{n\in\mathbb{N}}$ 有收敛子列, 由 $\{x_n\}_{n\in\mathbb{N}}$ 的任意性即知

$$A|_{X_0} \in \mathfrak{C}(X_0, Y).$$

此即 (iv).

对于 (v), 证明的思路在于用 $\overline{A}(B_1)$ 的 ε 网构造 $R(A)$ 的可数稠密子集. 注意到 $X = \bigcup_{n\in\mathbb{N}} nB_1$(其中 B_1 为 X 中的单位球), 由 A 的线性性知

$$R(A) = \bigcup_{n\in\mathbb{N}} nA(B_1). \tag{6.24}$$

根据定义6.5(i) 与引理6.3知 $\overline{A(B_1)}$ 自列紧, 因而 $A(B_1)$ 列紧, 由 Hausdorff 定理知 $A(B_1)$ 完全有界, 故对任意 $k \in \mathbb{N}$ 而言, 存在 $A(B_1)$ 的有限 $\frac{1}{k}$ 网 $N_k \subset A(B_1)$ 使得

$$A(B_1) \subset \bigcup_{y\in N_k} B\left(y, \frac{1}{k}\right).$$

现在令

$$N := \bigcup_{k\in\mathbb{N}} N_k.$$

注意到对 $\forall k \in \mathbb{N}$ 有

$$N_k \subset A(B_1) \subset \bigcup_{y\in N_k} B\left(y, \frac{1}{k}\right)$$

故

$$N \subset A(B_1) \subset \bigcap_{k\in\mathbb{N}} \bigcup_{y\in N_k} B\left(y, \frac{1}{k}\right).$$

对上述集合同时缩放知对 $\forall n \in \mathbb{N}$ 有

$$nN \subset nA(B_1) \subset n\left\{\bigcap_{k\in\mathbb{N}} \bigcup_{y\in N_k} B\left(y, \frac{1}{k}\right)\right\}.$$

显见 $\bigcup_{n\in\mathbb{N}} nN$ 可数, 下面验证 $\bigcup_{n\in\mathbb{N}} nN$ 在 $R(A)$ 中稠密. 任取 $z \in R(A)$, 根据上式与分解(6.24)知存在 $n_0 \in \mathbb{N}$ 使得

$$z \in n_0 A(B_1) \subset n_0\left\{\bigcap_{k\in\mathbb{N}} \bigcup_{y\in N_k} B\left(y, \frac{1}{k}\right)\right\}.$$

对 $\forall \varepsilon \in (0, \infty)$, 取 $k_0 \in \mathbb{N}$ 使得 $\frac{n_0}{k_0} < \varepsilon$. 现在因为 $\frac{z}{n_0} \in \bigcup_{y\in N_{k_0}} B(y, \frac{1}{k_0})$, 故存在 $y_0 \in N_{k_0}$ 使得

$$\frac{z}{n_0} \in B\left(y_0, \frac{1}{k_0}\right).$$

此即 $\|\frac{z}{n_0} - y_0\|_Y < \frac{1}{k_0}$, 于是

$$\|z - n_0 y_0\|_Y < \frac{n_0}{k_0} < \varepsilon.$$

又因为 $n_0 y_0 \in n_0 N \subset \bigcup_{n\in\mathbb{N}} nN$, 故 $\bigcup_{n\in\mathbb{N}} nN$ 在 $R(A)$ 中稠密, 因此 $R(A)$ 可分, 此即 (v).

对于 (vi), 若 A 是紧算子, 由定义6.5(iii) 知对 X 中的任意有界点列 $\{x_n\}_{n\in\mathbb{N}}$ 而言, 序列 $\{Ax_n\}_{n\in\mathbb{N}}$ 均有收敛子列 $\{Ax_{n_k}\}_{k\in\mathbb{N}}$. 又因为 B 有界, 故 $\{(B \circ A)x_{n_k}\}_{k\in\mathbb{N}}$ 收敛, 因此 $B \circ A \in \mathfrak{C}(X, Y)$.

若 B 是紧算子, 对 X 中的任意有界点列 $\{x_n\}_{n\in\mathbb{N}}$ 而言, 由 A 有界知 $\{Ax_n\}_{n\in\mathbb{N}}$ 是有界点列, 因此由 $B\in\mathfrak{C}(X,Y)$ 与定义6.5知 $\{(B\circ A)x_n\}_{n\in\mathbb{N}}$ 有收敛子列, 故 $B\circ A\in\mathfrak{C}(X,Y)$. 此即 (vi).

注

- 关于有界集与完全有界集我们需要特别回忆一下下述事实:
  (i) M 有界 $\Leftrightarrow$ 存在 $\varepsilon_0\in(0,\infty)$ 使得 M 有有限 ε_0 网. 事实上, 若 M 有界, 则存在 $x_0\in M$ 与 $\varepsilon_0\in(0,\infty)$ 使得 $M\subset B(x_0,\varepsilon_0)$, 这说明 M 有一个有限 ε_0 网. 反之, 若存在 $\varepsilon_0\in(0,\infty)$ 使得 M 有有限 ε_0 网 $\{y_i\}_{i=1}^m$, 设背景空间 X 上的度量为 d, 并记

  $$r_0:=\varepsilon_0+\max\{d(y_i,y_1):i\in\{2,3,\cdots,m\}\},$$

  则 $M\subset B(y_1,r_0)$, 这是因为对 $\forall x\in M$, 根据 ε_0 网的定义知存在 $i_0\in\{1,\cdots,m\}$ 使得 $x\in B(y_{i_0},\varepsilon_0)$, 于是

  $$d(x,y_1)\le d(x,y_{i_0})+d(y_{i_0},y_1)<\varepsilon_0+d(y_{i_0},y_1)\le r_0.$$

  此即 $M\subset B(y_1,r_0)$, 因此 M 有界.

  (ii) 完全有界集必有界, 但反之未必. 例如在 l^2 中对 $\forall n\in\mathbb{N}$, 记

  $$e_n:=\{\underbrace{0,\cdots,0,1}_{n},0\cdots\}$$

  知 $n\neq m\in\mathbb{N}$ 时有 $\|e_n-e_m\|_{l^2}=\sqrt{2}$, 且对 $\forall n\in\mathbb{N}$ 而言 $\|e_n\|_{l^2}=1$. 下面说明 $\varepsilon\in(0,\frac{\sqrt{2}}{2}]$ 时 $\{e_n\}_{n\in\mathbb{N}}$ 没有有限 ε 网. 考虑反证法, 设 $\{e_n\}_{n\in\mathbb{N}}$ 有有限 ε 网 $\{B(e_{i_k},\varepsilon)\}_{k=1}^N$, 根据鸽笼原理知必有一个球 (设之为 $B(e_{i_{k_0}},\varepsilon)$) 含有不同的 $e_n\neq e_m$. 但

  $$\sqrt{2}=\|e_n-e_m\|_{l^2}\le\|e_n-e_{i_{k_0}}\|_{l^2}+\|e_{i_{k_0}}-e_m\|_{l^2},$$

  这与

  $$\|e_n-e_{i_{k_0}}\|_{l^2}+\|e_{i_{k_0}}-e_m\|_{l^2}<2\varepsilon\le\sqrt{2}$$

  矛盾! 故 $\{e_n\}_{n\in\mathbb{N}}$ 没有有限 ε 网, 亦即 $\{e_n\}_{n\in\mathbb{N}}$ 作为有界集并不完全有界.

  (iii) M(完全) 有界 $\Leftrightarrow$ $\overline{M}$(完全) 有界. 这里我们只说明完全有界集的情形. 当 $\overline{M}$ 是完全有界集时, 根据定义知对 $\forall\varepsilon\in(0,\infty)$ 而言, 存在 $N\in\mathbb{N}$ 与 $\{\widetilde{x}_n\}_{n=1}^N\subset\overline{M}$ 使得 $\overline{M}\subset\bigcup_{n=1}^N B(\widetilde{x}_n,\frac{\varepsilon}{2})$. 现对 $\forall n\in\{1,\cdots,n\}$, 取 $x_n\in M$ 使得 $d(\widetilde{x}_n,x_n)<\frac{\varepsilon}{2}$, 则

  $$M\subset\overline{M}\subset\bigcup_{n=1}^N B(x_n,\varepsilon).$$

  故 M 完全有界. 反之, 若 M 完全有界, 则对 $\forall\varepsilon\in(0,\infty)$ 存在 $N\in\mathbb{N}$ 与 $\{x_n\}_{n=1}^N\subset M$ 使得 $M\subset\bigcup_{n=1}^N B(x_n,\frac{\varepsilon}{2})$. 任取 $x\in\overline{M}$, 知必存在 $y\in M$ 使得 $d(x,y)<\frac{\varepsilon}{2}$. 又因为 $y\in M\subset\bigcup_{n=1}^N B(x_n,\frac{\varepsilon}{2})$, 故必存在 $n_0\in\{1,\cdots,N\}$ 使得 $d(y,x_{n_0})<\frac{\varepsilon}{2}$, 从而

  $$d(x,x_{n_0})\le d(x,y)+d(y,x_{n_0})<\varepsilon,$$

  这说明 $\overline{M}\subset\bigcup_{n=1}^N B(x_n,\varepsilon)$, 故 $\overline{M}$ 完全有界.
- 紧算子的基本性质6.4(iii) 还有另一种 (更直接的) 证法. 和原先的证明一样, 设 $\{T_n\}_{n\in\mathbb{N}}$ 是紧算子列, $T\in\mathfrak{C}(X,Y)$. 我们从定义6.5(iii) 入手证明 T 是紧算子. 任取 X 中的有界集 $\{x_n\}_{n\in\mathbb{N}}$, 记其界为 M. 由(6.21)式知对 $\forall\varepsilon\in(0,\infty)$ 而言存在 $N\in\mathbb{N}$ 使得

  $$\|T_N-T\|_{L(X,Y)}<\frac{\varepsilon}{4(M+1)}.$$

  因为 $\{T_n\}_{n\in\mathbb{N}}$ 是紧算子列, 故通过定义6.5(iii) 与对角线法则可以找到 $\{x_n\}_{n\in\mathbb{N}}$ 的子列 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 使得对 $\forall n\in\mathbb{N}$ 而言, 序列 $\{T_nx_{n_k}\}$ 在 Y 中均收敛, 故存在 $K\in\mathbb{N}$ 使得 $k>K$ 时对 $\forall l\in\mathbb{N}$ 有

  $$\|T_Nx_{n_{k+l}}-T_Nx_{n_k}\|_Y<\frac{\varepsilon}{2}.$$

因此

$$\begin{aligned}\|Tx_{n_{k+l}} - Tx_{n_k}\|_Y &\le \|Tx_{n_{k+l}} - T_N x_{n_{k+l}}\|_Y + \|T_N x_{n_{k+l}} - T_N x_{n_k}\|_Y + \|T_N x_{n_k} - Tx_{n_k}\|_Y \\ &\le 2M\|T_N - T\|_{L(X,Y)} + \frac{\varepsilon}{2} < \varepsilon.\end{aligned}$$

这说明 $\{Tx_{n_k}\}_{k\in\mathbb{N}}$ 是 Y 中的基本列, 又因为 Y 完备, 故 $\{Tx_{n_k}\}_{k\in\mathbb{N}}$ 在 Y 中收敛, 这说明 T 是紧算子, 此即 (iii).

- 若 Y 不是 Banach 空间, 则紧算子的基本性质6.4(iii) 未必成立, 亦即 $\mathfrak{C}(X,Y)$ 未必在 $L(X,Y)$ 中闭.

根据紧算子的定义即知, 紧算子大多是在强拓扑下谈的. 对于弱拓扑而言, 我们同样能给出与紧算子类似定义的算子, 此即全连续算子:

定义 6.7 (全连续算子 (completely comtinuous operator))

设 X,Y 均为 Banach 空间. 若线性算子 $A: X\to Y$ 满足对 X 中的任意弱收敛列 $x_n \rightharpoonup x(n\to\infty)$ 而言, 均有 $Ax_n \to Ax(n\to\infty)$, 就称 A 为全连续算子. ♣

注 若线性算子 A 全连续, 则其必连续. 根据 A 的线性性, 只需说明 A 在 0_X 处连续. 设 $x_n\to 0_X(n\to\infty)$, 则 $x_n \rightharpoonup 0_X(n\to\infty)$, 由 A 的全连续性可得 $Ax_n\to A0_X = 0_Y(n\to\infty)$, 此即欲证.

全连续算子与紧算子之间的关系由下述命题给出:

命题 6.5

若 $A\in\mathfrak{C}(X,Y)$, 则 A 全连续. 反之, 若 X 自反, 则 $A\in\mathfrak{C}(X,Y)\Leftrightarrow A$ 全连续. ♠

证明 设 $A\in\mathfrak{C}(X,Y)$, $\{x_n\}_{n\in\mathbb{N}}\subset X$, 记 $x_n\rightharpoonup x\in X(n\to\infty)$, 往证 $Ax_n\to Ax(n\to\infty)$. 考虑反证法, 假设存在 $\varepsilon_0>0$ 与子列 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 使得

$$\forall k\in\mathbb{N}(\|Ax_{n_k} - Ax\|_Y \ge \varepsilon_0). \tag{6.25}$$

因为 $\{x_n\}_{n\in\mathbb{N}}$ 弱收敛, 故它作为 X^{**} 中的元素时满足 $\forall f\in X^*(\langle x_n, f\rangle < \infty)$, 根据 Banach-Steinhaus 定理2.7知 $\{x_n\}_{n\in\mathbb{N}}$ 是 X^{**} 中的有界集, 故它同样是 X 中的有界集. 因此 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 是 X 中的有界集, 根据 $A\in\mathfrak{C}(X,Y)$ 知 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 存在子列 (不妨仍将其记为 $\{x_{n_k}\}$) 与 $z\in Y$ 使得

$$Ax_{n_k}\to z,\quad k\to\infty.$$

另一方面, 对任意 $y^*\in Y^*$, 由 $x_n\rightharpoonup x\in X(n\to\infty)$ 与 $A^*y^*\in X^*$ 知 $k\to\infty$ 时

$$\langle y^*, Ax_{n_k} - Ax\rangle = \langle A^*y^*, x_{n_k} - x\rangle \to 0.$$

因此

$$Ax_{n_k}\rightharpoonup Ax,\quad k\to\infty.$$

由弱极限的唯一性知 $Ax = z$ 且 $Ax_{n_k}\to Ax(k\to\infty)$, 这与(6.25)式矛盾! 故 A 全连续.

反之, 若 X 自反且 A 全连续, 由 Eberlein-Smulian 定理3.6知 X 中的有界集必弱列紧, 这说明对 X 中的任意有界点列 $\{x_n\}_{n\in\mathbb{N}}$, 必存在子列 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 与 $x\in X$ 使得 $x_{n_k}\rightharpoonup x(k\to\infty)$. 因为 A 全连续, 故 $Ax_{n_k}\to Ax(k\to\infty)$, 由定义6.5(iii) 即知 $A\in\mathfrak{C}(X,Y)$. □

定理 6.5 (Schauder)

设 X 是 B^* 空间, Y 是 Banach 空间, 则 $A\in\mathfrak{C}(X,Y)\Leftrightarrow A^*\in\mathfrak{C}(Y^*,X^*)$. ♡

证明 $\Rightarrow$: 设 $A\in\mathfrak{C}(X,Y)$, 往证 $A^*\in\mathfrak{C}(Y^*,X^*)$, 此即证明当 $\{y_n^*\}_{n\in\mathbb{N}}$ 为 Y^* 中的有界点列时, $\{A^*y_n^*\}_{n\in\mathbb{N}}$ 在 X^* 中有收敛子列. 为此不妨设 $\|y_n^*\|_{Y^*}\le M$(其中 $n\in\mathbb{N}$, M 是正常数). 对 $\forall n\in\mathbb{N}$ 与 $y\in\overline{A(B_1)}$, 令

$$\varphi_n(y) := \langle y_n^*, y\rangle.$$

则 φ_n 是 $\overline{A(B_1)}$ 上的连续函数. 记 $\overline{A(B_1)}$ 上连续函数的全体为 $C(\overline{A(B_1)})$, 注意到

$$\begin{aligned}\|A^*y_n^*\|_{X^*} &= \sup_{x\in B_1}|\langle A^*y_n^*,x\rangle| = \sup_{y\in A(B_1)}|\langle y_n^*,y\rangle|\\ &= \sup_{y\in\overline{A(B_1)}}|\langle y_n^*,y\rangle| = \sup_{y\in\overline{A(B_1)}}|\varphi_n(y)|\\ &=: \|\varphi_n\|_{C(\overline{A(B_1)})},\end{aligned}$$

因此 $\{A^*y_n^*\}_{n\in\mathbb{N}}$ 在 X^* 中有收敛子列当且仅当 $\{\varphi_n\}_{n\in\mathbb{N}}$ 在 $C(\overline{A(B_1)})$ 中有收敛子列, 因此只需证明 $\{\varphi_n\}_{n\in\mathbb{N}}$ 是 $C(\overline{A(B_1)})$ 中的列紧集. 根据 Arzela-Ascoli 定理与 $\overline{A(B_1)}$ 的紧性, 这便只需证明 $\{\varphi_n\}_{n\in\mathbb{N}}$ 一致有界且等度连续.

对于一致有界性, 任取 $y\in\overline{A(B_1)}$, 知存在 $\{x_n\}_{n\in\mathbb{N}}\subset B_1$ 使得

$$\|y-Ax_n\|_Y\to 0,\quad n\to\infty.$$

于是

$$\begin{aligned}\|y\|_Y &\le \|y-Ax_n\|_Y+\|Ax_n\|_Y\\ &\le \|y-Ax_n\|_Y+\|A\|_{L(X,Y)}\\ &\to \|A\|_{L(X,Y)},\quad n\to\infty,\end{aligned}$$

此即 $\|y\|_Y\le\|A\|_{L(X,Y)}$, 因此对 $\forall n\in\mathbb{N}$ 与 $y\in\overline{A(B_1)}$ 有

$$|\varphi_n(y)| = |\langle y_n^*,y\rangle| \le \|y_n^*\|_{Y^*}\|y\|_Y \le M\|A\|_{L(X,Y)}.$$

故 $\{\varphi_n\}_{n\in\mathbb{N}}$ 在 $\overline{A(B_1)}$ 上一致有界.

对于等度连续性, 对 $\forall n\in\mathbb{N}$ 与 $y,z\in\overline{A(B_1)}$ 有

$$|\varphi_n(y)-\varphi_n(z)| = |\langle y_n^*,y-z\rangle| \le M\|y-z\|_Y.$$

因此 $\{\varphi_n\}_{n\in\mathbb{N}}$ 在 $\overline{A(B_1)}$ 上等度连续, 至此 $\Rightarrow$ 得证.

$\Leftarrow$: 设 $A^*\in\mathfrak{C}(Y^*,X^*)$, 由 $\Rightarrow$ 的结论知 $A^{**}\in\mathfrak{C}(X^{**},Y^{**})$. 现设 $U:X\to X^{**},V:Y\to Y^{**}$ 为典则嵌入, 可见 U,V 等距. 下面说明对 $\forall x\in X^*$ 有

$$A^{**}Ux = VAx.$$

这是因为对 $\forall y^*\in Y^*$ 有

$$\begin{aligned}\langle A^{**}Ux,y^*\rangle &= \langle Ux,A^*y^*\rangle = \langle A^*,y^*,x\rangle\\ &= \langle y^*,Ax\rangle = \langle VAx,y^*\rangle.\end{aligned}$$

现设 $\{x_n\}_{n\in\mathbb{N}}$ 是 X 中的有界点列, 由 U 的等距性知 $\{Ux_n\}_{n\in\mathbb{N}}$ 是 X^{**} 中的有界点列, 于是根据 $A^{**}\in\mathfrak{C}(X^{**},Y^{**})$ 知 $\{A^{**}Ux_n\}_{n\in\mathbb{N}}$ 在 Y^{**} 中有收敛子列, 此即 $\{VAx_n\}_{n\in\mathbb{N}}$ 在 Y^{**} 中有收敛子列, 记该子列为 $\{VAx_{n_k}\}_{k\in\mathbb{N}}$, 于是

$$\forall\varepsilon>0\exists K>0\forall p,q>K(\|VAx_{n_p}-VAx_{n_q}\|_{Y^{**}}<\varepsilon)$$

根据 V 的等距性可得

$$\forall\varepsilon>0\exists K>0\forall p,q>K(\|Ax_{n_p}-Ax_{n_q}\|_Y<\varepsilon)$$

由 Y 是 Banach 空间即知 $\{Ax_{n_k}\}_{k\in\mathbb{N}}$ 作为 Y 中的基本列在 Y 中收敛, 因此 $A\in\mathfrak{C}(X,Y)$. □

下面给出一个紧算子的例子.

例 6.4 设有界闭集 $\Omega\subset\mathbb{R}^n, K\in C(\Omega\times\Omega), X=Y=C(\Omega)$, 记

$$T: u\mapsto\int_\Omega K(\cdot,y)u(y)dy,\quad \forall u\in C(\Omega),$$

则 $T\in\mathfrak{C}(X)$.

证明 对 $\forall u\in C(\Omega)$, 首先说明 $Tu\in C(\Omega)$. 事实上, 因为 $K\in C(\Omega\times\Omega)$ 且 Ω 紧, 故 K 在 $\Omega\times\Omega$ 上一致连续, 根据一致连续的定义知

$$\forall\varepsilon\in(0,\infty)\exists\delta\in(0,\infty)\forall x_1,x_2\in\Omega\forall y\in\Omega(|x_1-x_2|<\delta\Rightarrow|K(x_1,y)-K(x_2,y)|<\varepsilon).$$

因此

$$\begin{aligned}|Tu(x_1)-Tu(x_2)| &= \left|\int_\Omega [K(x_1,y)-K(x_2,y)]u(y)dy\right| \\ &< \varepsilon\int_\Omega |u(y)|du \le \varepsilon\|u\|_{C(\Omega)}|\Omega|.\end{aligned}$$

故 Tu 在 Ω 上连续, 进而 $Tu\in C(\Omega)$, 这说明 T 首先是 $X\to X$ 的.

下面说明 $T\in\mathfrak{C}(X)$, 我们从定义6.5(iii) 入手, 任取 $C(\Omega)$ 中的有界点列 $\{u_m\}_{m\in\mathbb{N}}$, 往证 $\{Tu_m\}_{m\in\mathbb{N}}$ 有收敛子列, 这也就是证明 $\{Tu_m\}_{m\in\mathbb{N}}$ 列紧, 根据 Arzela-Ascoli 定理知只需证明 $\{Tu_m\}_{m\in\mathbb{N}}$ 一致有界且等度连续即可. 不妨设对 $\forall m\in\mathbb{N}$ 有 $\|u_m\|_{C(\Omega)}\le L$(其中 L 是与 m 无关的非负常数), 根据 T 的连续性与 Ω 的紧性知

$$\begin{aligned}\|Tu_m\|_{C(\Omega)} &= \max_{x\in\Omega}\left|\int_\Omega K(x,y)u_m(y)dy\right| \\ &\le L|\Omega|\max_{(x,y)\in\Omega\times\Omega}|K(x,y)|<\infty.\end{aligned}$$

因此 $\{Tu_m\}_{m\in\mathbb{N}}$ 在 Ω 上一致有界.

对于等度连续性, 由 Ω 紧与 $K\in C(\Omega\times\Omega)$ 知 K 在 $\Omega\times\Omega$ 上一致连续, 即

$$\forall\varepsilon\in(0,\infty)\exists\delta\in(0,\infty)\forall x,\widetilde{x}\in\Omega\forall y\in\Omega(|x-\widetilde{x}|<\delta\Rightarrow|K(x,y)-K(\widetilde{x},y)|<\varepsilon).$$

于是对 $\forall n\in\mathbb{N}$ 有

$$|Tu_m(x)-Tu_m(\widetilde{x})|\le\int_\Omega|K(x,y)-K(\widetilde{x},y)||u_m(y)|dy<\varepsilon L|\Omega|.$$

因此 $\{Tu_m\}_{m\in\mathbb{N}}$ 在 Ω 上等度连续, 至此即得 $T\in\mathfrak{C}(X)$.

定义 6.8 (有穷秩算子 (finite rank operator))

设 $T\in L(X,Y)$, 若 $\dim R(T)<\infty$, 就称 T 是有穷秩算子. 有穷秩算子的全体记作 $F(X,Y)$, 特别当 $X=Y$ 时记 $F(X,Y)=F(X)$.

♣

因为有限维线性空间中紧集与有界闭集等价, 故有穷秩算子必为紧算子.

定义 6.9 (秩 1 算子 (rank 1 operator))

设 $f\in X^*, y\in Y$, 定义算子 $y\otimes f$ 为

$$y\otimes f: X\to Y,\quad x\mapsto\langle f,x\rangle y.$$

称之为秩 1 算子.

♣

注 我们特别提醒关于秩 1 算子的下述两点:

(i) 对秩 1 算子 $y\otimes f$ 有:

$$\dim R(y\otimes f)=\begin{cases}1, & y\ne 0_Y\wedge f\ne 0_{X^*}\\ 0, & y=0_Y\vee f=0_{X^*}\end{cases}$$

(ii) 因为

$$\begin{aligned}\|y\otimes f\|_{L(X,Y)} &= \sup_{\|x\|_X=1}\|\langle f,x\rangle y\|_Y \\ &= \sup_{\|x\|_X=1}|\langle f,x\rangle|\|y\|_Y=\|f\|_{X^*}\|y\|_Y,\end{aligned}$$

故秩 1 算子必为有界线性算子.

利用秩 1 算子可以刻画有穷秩算子, 此即下述定理:

定理 6.6

算子 $T\in F(X,Y)$ 当且仅当存在 $m\in\mathbb{N}$ 使得对 $\forall i\in\{1,\cdots,m\}$ 而言, 存在 $y_i\in Y$ 与 $f_i\in X^*$ 使得

$$T=\sum_{i=1}^{m}y_i\otimes f_i.$$

♡

证明 $\Leftarrow$: 记 $\operatorname{span}\{y_1,\cdots,y_m\}$ 为 $\{y_1,\cdots,y_m\}$ 张成的线性子空间, 即

$$\operatorname{span}\{y_1,\cdots,y_m\}:=\left\{y\in Y:y=\sum_{i=1}^{m}\lambda_iy_i,\lambda_i\in\mathbb{C},i\in\{1,\cdots,m\}\right\}.$$

由 $R(T)\subset\operatorname{span}\{y_1,\cdots,y_m\}$ 即知 $T\in F(X,Y)$.

$\Rightarrow$: 若 $T\in F(X,Y)$, 则 $m:=\dim R(T)<\infty$. 设 $\{y_1,\cdots,y_m\}$ 为 $R(T)$ 的基, 则 $\{y_1,\cdots,y_m\}$ 线性无关, 因此对 $\forall x\in X$ 而言存在唯一序列 $\{l_i(x)\}_{i=1}^m\subset\mathbb{C}$ 使得

$$Tx=\sum_{i=1}^{m}l_i(x)y_i. \tag{6.26}$$

下面说明 $l_i\in X^*(\forall i\in\{1,\cdots,m\})$. 为此首先说明 $\{l_i\}_{i=1}^m$ 是线性的. 事实上, 对任意 $\alpha_1,\alpha_2\in\mathbb{C}$ 与任意 $x_1,x_2\in X$ 有

$$\begin{aligned}&T(\alpha_1x_1+\alpha_2x_2)=\alpha_1Tx_1+\alpha_2Tx_2\\ \Leftrightarrow&\sum_{i=1}^{m}l_i(\alpha_1x_1+\alpha_2x_2)=\alpha_1\sum_{i=1}^{m}l_i(x_1)y_i+\alpha_2\sum_{i=1}^{m}l_i(x_2)y_i\\ \Leftrightarrow&\sum_{i=1}^{m}[l_i(\alpha_1x_1+\alpha_2x_2)-\alpha_1l_i(x_1)-\alpha_2l_i(x_2)]y_i=0.\end{aligned}$$

由此与 $\{y_i\}_{i=1}^m$ 的线性无关性知对 $\forall i\in\{1,\cdots,m\}$ 有

$$l_i(\alpha_1x_1+\alpha_2x_2)=\alpha_1l_i(x_1)+\alpha_2l_i(x_2),$$

因此 $l_i(x)$ 是线性的.

下面说明对 $\forall i\in\{1,\cdots,m\}$ 而言 l_i 有界, 为此首先断言对 $\forall x\in X$ 而言, $\|Tx\|_Y$ 与 $\sum_{i=1}^m|l_i(x)|$ 均为 $R(T)$ 上的范数[9]. 事实上, $\|Tx\|_Y$ 显然是 $R(T)\subset Y$ 上的范数. 对于 $\sum_{i=1}^m|l_i(x)|$, 首先它非负, 其次

$$\begin{aligned}\sum_{i=1}^{m}|l_i(x)|=0&\Leftrightarrow l_i(x)=0,\forall i\in\{1,\cdots,m\}\\ &\Leftrightarrow Tx=0_Y.\end{aligned}$$

而 $\sum_{i=1}^m|l_i(x)|$ 的齐次性与三角不等式由 $\{l_i\}_{i=1}^m$ 的线性性立得, 故 $\sum_{i=1}^m|l_i(x)|$ 也为 $R(T)$ 上的范数. 现在因为有限维空间上的诸范数互相等价, 可知存在 $M>0$ 使得对 $\forall x\in X$ 均有

$$\sum_{i=1}^{m}|l_i(x)|\le M\|Tx\|_Y\le M\|T\|_{L(X,Y)}\|x\|_X.$$

由此可知对 $\forall i\in\{1,2,\cdots,m\}$ 均有 $l_i\in X^*$, 因此对每个 $i\in\{1,2,\cdots,m\}$ 而言, 存在 $f_i\in X^*$ 使得对 $\forall x\in X$ 有

$$l_i(x)=\langle f_i,x\rangle.$$

由此可得

$$Tx=\sum_{i=1}^{m}\langle f_i,x\rangle y_i=\sum_{i=1}^{m}y_i\otimes f_i(x).$$

这说明

$$T=\sum_{i=1}^{m}y_i\otimes f_i,$$

[9] 这里 $\sum_{i=1}^m|l_i(x)|$ 作为 $R(T)$ 上的范数可以理解为: 对每个 $y\in R(T)$, 总存在 $x\in X$ 使得 $y=Tx$. 根据 $R(T)$ 的基 $\{y_1,\cdots,y_m\}$ 的线性无关性, 这个 $x\in X$ 对应的序列 $\{l_i(x)\}_{i=1}^m$ 唯一, 因此 $\sum_{i=1}^m|l_i(x)|$ 良定义, 且它唯一对应 $R(T)$ 中的元素 y.

此即欲证. □

回忆紧算子的基本性质6.4(iii), 当 Y 是 Banach 空间时 $\mathfrak{C}(X,Y)$ 本身是 $L(X,Y)$ 中 (在强拓扑下) 的闭集, 而前面已经提过 $F(X,Y)\subset\mathfrak{C}(X,Y)$, 故 $\overline{F(X,Y)}^{L(X,Y)}\subset\mathfrak{C}(X,Y)$, 这也就是下述推论:

推论 6.2

设 X,Y 是 Banach 空间, $\{T_n\}_{n\in\mathbb{N}}\subset F(X,Y), T\in L(X,Y)$. 若 $\|T_n-T\|_{L(X,Y)}\to 0(n\to\infty)$, 则 $T\in\mathfrak{C}(X,Y)$.

注 由 Banach 与 Grothendieck 提出的著名" 逼近问题" 实际上就是推论6.2的反问题: 对给定的紧算子 T 而言, 是否总存在有穷秩算子列 $\{T_n\}_{n\in\mathbb{N}}$ 在 $L(X,Y)$ 范数下收敛到 T? 这个问题公开了很久, 直到 1972 年 P. Enflo 才给出第一个反例. 最初的反例构造相当复杂, 不过之后陆续发现了更简单的反例: 比方说当 Y 是 $l^p(1<p<\infty,p\neq 2)$ 的某些闭子空间时这件事就不成立. 读者可以在 J. Lindenstrauss-L. Tzafriri 的文章中找到对逼近问题的详尽讨论. 注意到逼近问题在某些情况下是成立的, 例如下述命题讨论的 $X=Y$ 为 Hilbert 空间的情形:

命题 6.6

设 H 是 Hilbert 空间, 则 $\overline{F(H)}^{L(H)}=\mathfrak{C}(H)$.

证明 已知 $\overline{F(H)}^{L(H)}\subset\mathfrak{C}(H)$, 故只需再说明 $\mathfrak{C}(H)\subset\overline{F(H)}^{L(H)}$ 即可. 设 $T\in\mathfrak{C}(H)$, 根据紧算子的定义知 $\overline{T(B_1)}$ 紧, 其中 B_1 是 H 中的单位球. 现在 H 作为 Hilbert 空间首先是赋范的, 因此它是可度量的, 而在度量空间中根据引理6.3知紧性和自列紧性等价, 故 $\overline{T(B_1)}$ 自列紧, 又由 Hausdorff 定理6.6可知 $\overline{T(B_1)}$ 完全有界, 因此任取 $\varepsilon\in(0,\infty)$, 总存在 $\overline{T(B_1)}$ 的有限 $\varepsilon/2$ 网 $\{y_1,y_2,\cdots,y_m\}$ 使得

$$\overline{T(B_1)}\subset\bigcup_{i=1}^m B\left(y_i,\frac{\varepsilon}{2}\right).$$

令 $E_\varepsilon:=\mathrm{span}\{y_1,y_2,\cdots,y_m\}$. 因为 E_ε 是有限维线性空间, 故它是 H 的闭子空间, 根据 Hilbert 空间的正交分解定理知对 $\forall x\in H$, 存在唯一正交分解 $x=y+z$, 其中 $y\in E_\varepsilon$, 而

$$z\in E_\varepsilon^\perp=\{z\in H:(z,y)=0,\forall y\in E_\varepsilon\}.$$

记 $P_\varepsilon x=y$, 则 P_ε 是 H 到 E_ε 的正交投影. 由

$$\|x\|_H^2=\|y\|_H^2+\|z\|_H^2$$

进一步可得

$$\|P_\varepsilon\|_{L(H)}=\sup_{\|x\|_H=1}\|P_\varepsilon x\|_H=\sup_{\|x\|_X=1}\|y\|_H\leq 1$$

因此 $\|P_\varepsilon\|_{L(H)}\leq 1$, 这说明 $P_\varepsilon\in L(H)$, 且显见 $P_\varepsilon T\in F(H)$.

下面再用 $P_\varepsilon T$ 尝试逼近 T. 回到 $\overline{T(B_1)}$, 因为 $\{y_1,\cdots,y_m\}$ 是 $\overline{T(B_1)}$ 的有限 $\varepsilon/2$ 网, 故对 $\forall x\in B_1$, 总存在 $y_i(i\in\{1,\cdots,m\})$ 使得

$$\|Tx-y_i\|_H<\frac{\varepsilon}{2}.$$

由此及 $P_\varepsilon y_i=y_i$ 可知

$$\begin{aligned}\|P_\varepsilon Tx-y_i\|_H&=\|P_\varepsilon(Tx-y_i)\|_H\\&\leq\|P_\varepsilon\|_H\|Tx-y_i\|_H<\frac{\varepsilon}{2}.\end{aligned}$$

因此对 $\forall x\in B_1$ 有

$$\|Tx-P_\varepsilon Tx\|_H\leq\|Tx-y_i\|_H+\|y_i-P_\varepsilon Tx\|_H<\varepsilon.$$

由此可知 $\|T-P_\varepsilon T\|_{L(H)}\leq\varepsilon$, 此即 $\mathfrak{C}(H)\subset\overline{F(H)}^{L(H)}$. □

注 注意在上述证明中, 我们对定义域只用到了其上可赋范 (即可定义单位球 B_1), 而对到达域才用到了 Hilbert 空间的性质. 因此上述命题可以改进为下述版本:

命题 6.7

设 X 是 B^* 空间, H 是 Hilbert 空间, 则 $\overline{F(X,H)}^{L(X,H)} = \mathfrak{C}(X,H)$.

下面我们再考察可分空间上的逼近问题, 为此我们先引入 Schauder 基的概念:

定义 6.10 (Schauder 基 (Schauder basis))

设 X 为可分 Banach 空间, 称序列 $\{e_n\}_{n\in\mathbb{N}} \subset X$ 为 X 的一组 Schauder 基, 如果对 $\forall x \in X$, 均存在唯一序列 $\{C_n(x)\}_{n\in\mathbb{N}} \subset \mathbb{C}$ 使得

$$x = \lim_{N\to\infty} \sum_{n=1}^{N} C_n(x)e_n$$

在 X 中成立, 亦即

$$\left\| x - \sum_{n=1}^{N} C_n(x)e_n \right\|_X \to 0,\ N \to \infty.$$

注

(i) 若 $\{e_n\}_{n\in\mathbb{N}}$ 为 X 的一组 Schauder 基, 则由 $\{C_n(x)\}_{n\in\mathbb{N}}$ 的唯一性可知对 $\forall n \in \mathbb{N}$ 均有 $e_n \neq 0_X$, 故对 $\forall n \in \mathbb{N}$ 均有 $\|e_n\|_X \neq 0$.

(ii) 若 X 有 Schauder 基, 则 X 可分. 事实上, 设 $\{e_n\}_{n\in\mathbb{N}}$ 为 X 的 Schauder 基, 记

$$M := \left\{ \sum_{j=1}^{N} (a_j + ib_j)e_j : N \in \mathbb{N}, a_j, b_j \in \mathbb{Q}, j \in \{1,2,\cdots,N\} \right\},$$

显见 M 可数. 要证明 X 可分, 现在就只需证明 M 在 X 中稠密. 为此取 $x \in X$, 根据 Schauder 基的定义知 $N \to \infty$ 时有

$$\left\| x - \sum_{n=1}^{N} C_n(x)e_n \right\|_X \to 0,$$

因此对 $\forall \varepsilon \in (0,\infty)$, 总存在 $N \in \mathbb{N}$ 使得

$$\left\| x - \sum_{n=1}^{N} C_n(x)e_n \right\|_X < \frac{\varepsilon}{2}.$$

记 $L := \max\{\|e_n\|_X : n \in \{1,2,\cdots,N\}\}$, 由本注记 (i) 知 $L > 0$. 对 $\forall n \in \{1,2,\cdots,N\}$, 取 $a_n, b_n \in \mathbb{Q}$ 使得 $|C_n(x) - (a_n + ib_n)| < \frac{\varepsilon}{2LN}$, 则

$$\begin{aligned}\left\| x - \sum_{n=1}^{N} (a_n + ib_n)e_n \right\|_X &\leq \left\| x - \sum_{n=1}^{N} C_n(x)e_n \right\|_X + \sum_{n=1}^{N} |C_n(x) - (a_n + ib_n)| \|e_n\|_X \\ &< \frac{\varepsilon}{2} + LN\frac{\varepsilon}{2LN} = \varepsilon.\end{aligned}$$

因此 M 在 X 中稠密, 故 X 可分.

引理 6.7

若 X 是可分 Banach 空间, 则定义6.10中的 $C_n(x)$ 是 X 上的有界线性泛函.

证明 首先说明 C_n 是 X 上的线性函数. 任取 $\alpha, \beta \in \mathbb{C}, x, y \in X$ 知

$$\begin{aligned}\lim_{N\to\infty}\sum_{n=1}^{N} C_n(\alpha x+\beta y)e_n &= \alpha x+\beta y\\ &= \alpha\lim_{N\to\infty}\sum_{n=1}^{N} C_n(x)e_n+\beta\lim_{N\to\infty}\sum_{n=1}^{N} C_n(x)e_n\\ &= \lim_{N\to\infty}\sum_{n=1}^{N}[\alpha C_n(x)+\beta C_n(y)]e_n,\end{aligned}$$

由 C_n 的唯一性可知

$$C_n(\alpha x+\beta y)=\alpha C_n(x)+\beta C_n(y).$$

因此 C_n 首先是 X 上的线性函数.

下面说明连续性. 为此对 $\forall x \in X$ 定义

$$|||x||| := \sup_{n\in\mathbb{N}} \|S_n(x)\|_X,$$

其中 $S_n(x) := \sum_{i=1}^{n} C_i(x)e_i$. 易证 $|||\cdot|||$ 是 X 上的范数, 又因为当 $N\to\infty$ 时有

$$0 \le |\|x\|_X - \|S_n(x)\|_X| \le \|x-S_n(x)\|_X \to 0,$$

故

$$\|x\|_X = \lim_{n\to\infty}\|S_n(x)\|_X \le |||x|||.$$

下面说明 $(X, |||\cdot|||)$ 完备, 为此设 $\{x_m\}_{m\in\mathbb{N}}$ 为 $(X, |||\cdot|||)$ 中的基本列. 因为对任意 $m\in\mathbb{N}$ 均有

$$x_m = \lim_{n\to\infty}\sum_{i=1}^{n} C_i(x_m)e_i$$

在 $\|\cdot\|_X$ 范数下成立, 故对 $\forall m,k,n\in\mathbb{N}$ 有

$$\begin{aligned}|C_n(x_m)-C_n(x_k)|\|e_n\|_X &= \left\|\sum_{i=1}^{n}[C_i(x_m)-C_i(x_k)]e_i-\sum_{i=1}^{n-1}[C_i(x_m)-C_i(x_k)]e_i\right\|_X\\ &= \|S_n(x_m-x_k)-S_{n-1}(x_m-x_k)\|_X\\ &\le 2|||x_m-x_k|||,\end{aligned}$$

因此

$$|C_n(x_m)-C_n(x_k)| \le 2\|e_n\|_X^{-1}|||x_m-x_k|||.$$

于是 $\{C_n(x_m)\}_{m\in\mathbb{N}}$ 是 $\mathbb{C}$ 中的基本列, 故存在 $\alpha_n\in\mathbb{C}$ 使得

$$\lim_{m\to\infty} C_n(x_m)=\alpha_n.$$

进一步利用三角不等式可知对任意取定的 $n\in\mathbb{N}$, 当 $m\to\infty$ 时有

$$\left\|S_n(x_m)-\sum_{i=1}^{n}\alpha_i e_i\right\|_X \to 0. \tag{6.27}$$

下面说明 $\sum_{i=1}^{\infty}\alpha_i e_i\in X$, 且当 $m\to\infty$ 时有

$$\left|\left|\left|x_m-\sum_{i=1}^{\infty}\alpha_i e_i\right|\right|\right| \to 0.$$

事实上, 因为 $\{x_m\}_{m\in\mathbb{N}}$ 是 $(X, |||\cdot|||)$ 中的基本列, 故对 $\forall\varepsilon\in(0,\infty)$ 存在 $N_\varepsilon\in\mathbb{N}$ 使得对 $\forall m,k\ge N_\varepsilon$ 与 $n\in\mathbb{N}$ 有

$$\|S_n(x_m)-S_n(x_k)\|_X \le |||x_m-x_k||| < \frac{\varepsilon}{3},$$

从而

$$\begin{aligned}\left\|S_n(x_m)-\sum_{i=1}^n\alpha_ie_i\right\|_X&\leq\|S_n(x_m)-S_n(x_k)\|_X+\left\|S_n(x_k)-\sum_{i=1}^n\alpha_ie_i\right\|_X\\&<\frac{\varepsilon}{3}+\left\|S_n(x_k)-\sum_{i=1}^n\alpha_ie_i\right\|_X.\end{aligned}$$

在上式中令 $k\to\infty$ 并利用(6.27)式可知

$$\left\|S_n(x_m)-\sum_{i=1}^n\alpha_ie_i\right\|_X\leq\frac{\varepsilon}{3}.\tag{6.28}$$

由此知对任意的 $l\geq k\geq 2$ 有

$$\left\|\sum_{i=k}^lC_i(x_m)e_i-\sum_{i=k}^l\alpha_ie_i\right\|_X\leq\left\|S_l(x_m)-\sum_{i=1}^l\alpha_ie_i\right\|_X+\left\|S_{k-1}(x_m)-\sum_{i=1}^{k-1}\alpha_ie_i\right\|_X\leq\frac{2\varepsilon}{3}.\tag{6.29}$$

固定 $m\geq N_\varepsilon$, 因为 $x_m=\lim_{n\to\infty}S_n(x_m)$ 在 $\|\cdot\|_X$ 意义下成立, 故存在 $N\in\mathbb{N}$ 使得对 $\forall l\geq k\geq N$ 有

$$\left\|\sum_{i=k}^lC_i(x_m)e_i\right\|_X=\|S_l(x_m)-S_{k-1}(x_m)\|_X<\frac{\varepsilon}{3}.$$

由此及(6.29)式知

$$\left\|\sum_{i=k}^l\alpha_ie_i\right\|_X\leq\left\|\sum_{i=k}^l\alpha_ie_i-\sum_{i=k}^lC_i(x_m)e_i\right\|_X+\left\|\sum_{i=k}^lC_i(x_m)e_i\right\|_X<\varepsilon.$$

因此 $\left\{\sum_{i=1}^n\alpha_ie_i\right\}_{n\in\mathbb{N}}$ 为 $(X,\|\cdot\|_X)$ 中的基本列, 由 X 的完备性可知 $\sum_{i=1}^\infty\alpha_ie_i\in X$.

现在根据 Schauder 基分解的唯一性, 当 $m\geq N_\varepsilon$ 时, 因为

$$x:=\sum_{i=1}^\infty\alpha_ie_i=\sum_{i=1}^\infty C_i(x)e_i,$$

$$S_n(x)=\sum_{i=1}^nC_i(x)e_i=\sum_{i=1}^n\alpha_ie_i,$$

由此及(6.28)式有

$$\begin{aligned}|||x_m-x|||&=\sup_{n\in\mathbb{N}}\|S_n(x_m)-S_n(x)\|_X\\&=\sup_{n\in\mathbb{N}}\left\|S_n(x_m)-\sum_{i=1}^n\alpha_ie_i\right\|\leq\frac{\varepsilon}{3},\end{aligned}$$

这说明 x 是 $\{x_m\}_{m\in\mathbb{N}}$ 在 $|||\cdot|||$ 下的极限, 因此 $(X,|||\cdot|||)$ 完备. 根据等价范数定理2.9知存在 $M>0$ 使得对任意 $x\in X$ 均有

$$|||x|||\leq M\|x\|_X.$$

在上式中代入 $|||\cdot|||$ 的定义即知对 $\forall n\in\mathbb{N}$ 均有 $\|S_n\|_{L(X)}\leq M$, 且对 $\forall x\in X$ 有

$$|C_n(x)|\|e_n\|_X=\|S_n(x)-S_{n-1}(x)\|_X\leq 2M\|x\|_X,$$

因此

$$\|C_n\|_{X^*}=\sup_{x\in X\setminus\{0_X\}}\frac{|C_n(x)|}{\|x\|_X}\leq 2M\|e_n\|_X^{-1},$$

故 $C_n\in X^*$, 引理至此得证. □

定理 6.7

若 Banach 空间 X 上有 Schauder 基, 则 $\overline{F(X)}^{L(X)}=\mathfrak{C}(X)$.

证明 因为已经有 $\overline{F(X)}^{L(X)}\subset\mathfrak{C}(X)$, 故只需证明 $\mathfrak{C}(X)\subset\overline{F(X)}^{L(X)}$. 设 S_n 与引理6.7证明中的设置相同, 根据引

理6.7的证明可知存在 $M>0$ 使得对 $\forall n\in\mathbb{N}$ 有

$$\|S_n\|_{L(X)}\le M.$$

设 $T\in\mathfrak{C}(X)$, 则 $\overline{T(B_1)}$ 是紧集, 因此根据引理6.3知 $\overline{T(B_1)}$ 是自列紧集, 再由 Hausdorff 定理6.6知对 $\forall\varepsilon\in(0,\infty)$ 均存在 $\overline{T(B_1)}$ 的有穷 $\varepsilon/3(M+1)$ 网 $\{y_1,y_2,\cdots,y_m\}$ 使得

$$\overline{T(B_1)}\subset\bigcup_{i=1}^m B\Big(y_i,\frac{\varepsilon}{3(M+1)}\Big).$$

因此对 $\forall x\in B_1$, 总存在 y_i 使得

$$\|Tx-y_i\|_X<\frac{\varepsilon}{3(M+1)}.$$

故对 $\forall N\in\mathbb{N}$ 有

$$\|S_N(Tx)-S_N(y_i)\|_X<\frac{M\varepsilon}{3(M+1)}<\frac{\varepsilon}{3}.$$

另由 Schauder 基的定义知存在 $N\in\mathbb{N}$ 使得对 $\forall i\in\{1,2,\cdots,m\}$ 有

$$\|S_N(y_i)-y_i\|_X<\frac{\varepsilon}{3}.$$

于是

$$\|Tx-S_N(Tx)\|_X\le\|Tx-y_i\|_X+\|y_i-S_N(y_i)\|_X+\|S_N(y_i)-S_N(Tx)\|_X<\varepsilon.$$

记 $T_\varepsilon:=S_NT$, 显见 $T_\varepsilon\in F(X)$, 且 $\|T-T_\varepsilon\|_{L(X)}\le\varepsilon$. 因此 $\varepsilon\to0^+$ 时 T_ε 在 $L(X)$ 中收敛到 T, 至此定理即证. □

注 设 X,Y 均为 B^* 空间, $T:X\to Y$ 是线性算子.

(i) 如果 T 把有界集映成完全有界集, 就称 T 是完全有界算子. 根据引理6.3与 Hausdorff 定理6.6进而可知

$$T\text{ 是紧算子}\Rightarrow T\text{ 是完全有界算子}\Rightarrow T\text{ 是有界算子}.$$

另若 Y 是 Banach 空间, 则由 Hausdorff 定理6.6进一步知

$$T\text{ 是紧算子}\Rightarrow T\text{ 是完全有界算子}.$$

(ii) 回忆紧算子必全连续 (命题6.5), 全连续算子必有界线性, 可知当 Y 为 Banach 空间时有:

$$\begin{aligned}\overline{F(X,Y)}^{L(X,Y)}&\subset\mathfrak{C}(X,Y)\\&\subset\{X\to Y\text{ 的全体全连续算子}\}\\&\subset L(X,Y).\end{aligned}$$

因为定理6.7中对定义域只要求可赋范, 而对到达域才用到了完备性与 Schauder 基的存在性, 结合命题6.7可知当 Y 为 Hilbert 空间或具 Schauder 基的 Banach 空间时, 上式中第一个包含关系取等; 当 X 为自反空间时, 由命题6.5知上式中第二个包含关系取等. 特别若 Y 为有限维 B^* 空间, 则 Y 为 Banach 空间的同时还具有 Schauder 基, 于是 $\overline{F}(X,Y)^{L(X,Y)}=\mathfrak{C}(X,Y)$, 又因为 Y 本身是有限维的, 故 $\mathfrak{C}(X,Y)=L(X,Y)$, 因此

$$\overline{F(X,Y)}^{L(X,Y)}=\mathfrak{C}(X,Y)=L(X,Y).$$

注 回到前面注记提起过的逼近问题, 下面介绍一个在非线性分析中用有穷秩非线性映射逼近 (线性或非线性) 连续映射的重要方法. 设 X 是拓扑空间, Y 是 Banach 空间, $T:X\to Y$ 是连续映射, 其满足 $T(X)$ 在 Y 中预紧 (即 $\overline{T(X)}$ 在 Y 中紧). 下面断言存在有穷秩连续映射 $T_\varepsilon:X\to Y$ 使得

$$\|T_\varepsilon(x)-T(x)\|_Y<\varepsilon,\quad\forall x\in X.\tag{6.30}$$

事实上, 因为 $K=\overline{T(X)}$ 是 Y 中的紧集, 故存在 K 的有限覆盖 $\bigcup_{i\in I}B(f_i,\varepsilon/2)$. 记

$$T_\varepsilon(x)=\frac{\sum\limits_{i\in I}q_i(x)f_i}{\sum\limits_{i\in I}q_i(x)},\quad q_i(x)=\max\{\varepsilon-\|Tx-f_i\|_Y,0\};$$

直接验证即知 T_ε 满足(6.30)式.

这类逼近非常有用, 比方说它可以用在从 Brouwer 不动点定理推导 Schauder 不动点定理中. Lomonosov 另外

还把上式的一个类似构造与 Schauder 不动点定理结合起来, 给出了线性算子的一个大类中非平凡不变子空间存在性的精彩证明. 通过 Schauder 不动点定理能高度简化证明的另一个线性结果为 Krein-Rutman 定理, 这个定理我们后面也会具体介绍.

注 关于 Schauder 基需要特别注意: 不可分空间一定没有 Schauder 基, 而可分空间也未必有 Schauder 基.

6.3 Riesz-Fredholm 定理

我们首先预备一个引理:

引理 6.8 (Riesz)

设 X 是 B^* 空间, $M \subset X$ 是真闭线性子空间, 则

$$\forall \varepsilon \in (0,1) \exists u \in X (\|u\| = 1 \text{ 且 } \operatorname{dist}(u, M) \geq 1 - \varepsilon).$$

证明 取 $v \in X, v \notin M$. 因为 M 是闭的, 故

$$d = \operatorname{dist}(v, M) > 0.$$

任取 $m_0 \in M$ 使得

$$d \leq \|v - m_0\| \leq \frac{d}{1-\varepsilon}.$$

则

$$u = \frac{v - m_0}{\|v - m_0\|}$$

即为欲求. □

根据 Riesz 引理可以得到下述重要结果:

定理 6.8 (Riesz)

若 E 是 B^* 空间, B_E 作为 E 中的闭单位球是紧集, 则 E 是有限维的.

证明 用反证法, 设 E 无穷维, 则存在 E 的有限维子空间列 $\{E_n\}$ 使得 $E_{n-1} \subsetneq E_n$. 根据 Riesz 引理6.8, 存在序列 $\{u_n\}$ 满足 $u_n \in E_n, \|u_n\| = 1, \operatorname{dist}(u_n, E_{n-1}) \geq 1/2$. 特别地, 这说明 $\|u_n - u_m\| \geq 1/2$ 对 $m < n$ 均成立, 因此 $\{u_n\}$ 不收敛 (它的任意子列也不收敛). 但因为 B_E 是紧集且 E 赋范, 故 B_E 是自列紧集, 矛盾! 此即欲证. □

定理 6.9 (Fredholm 二择一定理)

设 E 是 B^* 空间, 若 $T \in \mathfrak{C}(E)$, 则

(i) $N(I-T)$ 是有限维的,

(ii) $R(I-T)$ 是闭集, 且 $R(I-T) = N(I-T^*)^{\perp}$,

(iii) $N(I-T) = \{0_E\} \Leftrightarrow R(I-T) = E$,

(iv) $\dim N(I-T) = \dim N(I-T^*)$.

注 之所以称定理6.9为二择一定理, 是因为它解决了方程 $u - Tu = f$ 的可解性. 该定理表明:

- 要么对每个 $f \in E$ 而言, 方程 $u - Tu = f$ 都存在唯一解 (这对应着结论 (iii)),
- 要么齐次方程 $u - Tu = 0$ 具有 n 个线性独立解, 此时非齐次方程 $u - Tu = f$ 有解当且仅当 f 满足 n-正交条件, 即

$$f \in N(I-T^*)^{\perp}.$$

(这对应着结论 (i),(ii),(iv).)

注 注意二择一定理6.9的结论 (iii) 在有限维空间中其实是老朋友了. 当 $\dim E < \infty$ 时, E 到它自身的线性算子是单射当且仅当它是满射. 然而, 无穷维空间中有界算子成为单射时不必为满射, 反之亦然, 例如 l^2 中的右推移算子是

不满的单射, 而左推移算子是不单的满射. 因此, 结论 (iii) 实际上是形如 $I-T(T\in\mathfrak{C}(E))$ 的算子的一个惊人结论.

证明 (i) 设 $E_1=N(I-T)$, 则 E_1 中的每个元素 x 均满足 $x=Tx$, 因此 $\|x\|_E\le 1\Rightarrow\|Tx\|_E\le 1$, 故 $B_{E_1}\subset T(B_E)\subset\overline{T(B_E)}$. 又因为 $\overline{T(B_E)}$ 是紧集, 故 B_{E_1} 作为紧空间的闭子集也应该是紧集. 根据 Riesz 定理6.8, E_1 只能是有限维的.

(ii) 设 $f_n=u_n-Tu_n\to f$, 往证 $f\in R(I-T)$. 记 $d_n=\operatorname{dist}(u_n,N(I-T))$, 因为 $N(I-T)$ 是有限维的, 故对每个 u_n 都存在 $v_n\in N(I-T)$ 使得 $d_n=\|u_n-v_n\|$. 于是

$$f_n=(u_n-v_n)-T(u_n-v_n). \tag{6.31}$$

下面说明 $\|u_n-v_n\|$ 关于 n 一致有界. 若不然, 则存在子列使得 $\|u_{n_k}-v_{n_k}\|\to\infty$. 取

$$w_n=\frac{u_n-v_n}{\|u_n-v_n\|}.$$

因为 f_n 本身是强拓扑下的收敛列, 故 $\|f_n\|$ 必关于 n 一致有界, 进而由(6.31)式可知 $w_{n_k}-Tw_{n_k}\to 0$. 因为 $w_n\in B_E$, 由 $T\in\mathfrak{C}(E)$ 与引理6.3知 $\{Tw_{n_k}\}_k$ 列紧, 取其收敛子列 (不妨仍记为 w_{n_k}) 并设 $Tw_{n_k}\to z$. 根据极限的唯一性知 $w_{n_k}\to z$, 从而 $z-Tz=0$, 亦即 $z\in N(I-T)$. 这说明 $\operatorname{dist}(w_{n_k},N(I-T))\to 0$. 另一方面, 因为 $v_n\in N(I-T)$, 故由 w_n 本身的构造知:

$$\operatorname{dist}(w_n,N(I-T))=\frac{\operatorname{dist}(u_n,N(I-T))}{\|u_n-v_n\|}=1$$

矛盾! 因此 $\|u_n-v_n\|$ 关于 n 一致有界, 从而由 $T\in\mathfrak{C}(E)$ 知存在子列使得 $T(u_{n_k}-v_{n_k})$ 收敛到某个极限 l, 进而由(6.31)式知 $u_{n_k}-v_{n_k}\to f+l$. 记 $g=f+l$, 则 $g-Tg=f$, 此即 $f\in R(I-T)$, 因此 $R(I-T)$ 是闭集.

现在对 $I-T$ 应用闭像集算子的刻画2.14立得

$$R(I-T)=N(I^*-T^*)^\perp=N(I-T)^\perp,$$
$$R(I-T^*)=N(I-T)^\perp.$$

(iii) 首先证明 $\Rightarrow$. 用反证法, 设

$$E_1=R(I-T)\ne E.$$

根据 $R(I-T)$ 的闭性知 E_1 是 Banach 空间, 又因为显见 $T(E)\subset E$, 故 $T(E_1)\subset E_1$. 由紧算子的基本性质6.4(iii) 可知 $T|_{E_1}\in\mathfrak{C}(E_1)$, 重复上述步骤可知 $E_2=(I-T)(E_1)$ 是 E_1 的一个闭子空间. 因为 $N(I-T)=\{0_E\}$, 故 $(I-T)$ 是单射, 从而由反证的假设知依旧有 $E_2\ne E_1$. 现设 $E_n=(I-T)^n(E)$, 我们便得到了一个 (严格) 递减的闭子空间列. 对其中的每个闭子空间 E_n, 根据 Riesz 引理6.8我们都可以对应得到 $u_n\in E_n$ 使得 $\|u_n\|=1,\operatorname{dist}(u_n,E_{n+1})\ge 1/2$. 现有

$$Tu_n-Tu_m=-(u_n-Tu_n)+(u_m-Tu_m)+(u_n-u_m).$$

注意到若 $n>m$, 则 $E_{n+1}\subset E_n\subset E_{m+1}\subset E_m$, 而 $u_n-Tu_n\in E_{n+1},u_m-Tu_m\in E_{m+1},u_n\in E_n,u_m\in E_m$, 因此

$$-(u_n-Tu_n)+(u_m-Tu_m)+(u_n-u_m)\in E_{m+1}.$$

故

$$\|Tu_n-Tu_m\|\ge\operatorname{dist}(u_m,E_{m+1})\ge\frac{1}{2},\quad\forall n>m$$

但这说明 $\{Tu_n\}$ 作为 $T(B_{E_{m+1}})$ 中的序列不存在任何收敛子列, 这与 $T\in\mathfrak{C}(E_{m+1})$ 矛盾! 故只能有 $R(I-T)=E$.

反之, 设 $R(I-T)=E$, 根据满射算子的刻画2.15与推论2.12知 $N(I-T^*)=R(I-T)^\perp=\{0_{E^*}\}$. 因为 $T\in\mathfrak{C}(E)\Rightarrow T^*\in\mathfrak{C}(E^*)$, 利用 $\Rightarrow$ 的结论可知 $R(I-T^*)=E^*$, 再用一次满射算子的刻画2.15与推论2.12即知 $N(I-T)=R(I-T^*)^\perp=\{0_E\}$.

(iv) 记 $d=\dim N(I-T),d^*=\dim N(I-T^*)$, 我们首先证明 $d^*\le d$. 如若不然, 设 $d<d^*$, 因为 $N(I-T)$ 是有限维的, 故由例2.5可知 $N(I-T)$ 在 E 中有余子空间, 因此存在从 E 到 $N(I-T)$ 上的连续投影 P. 另一方面, 由例2.6可知 $R(I-T)=N(I-T^*)^\perp$ 余维为 d^*, 因此它作为余维有限的闭子空间同样有余子空间, 记它的余子空间为 F, 显见 $\dim F=d^*$. 因为 $d<d^*$, 故存在单但不满的线性映射 $\Lambda:N(I-T)\to F$. 取 $S=T+\Lambda\circ P$, 因为 $\Lambda\circ P$ 是有穷秩算子, 故 $\Lambda\circ P\in\mathfrak{C}(E)$, 因此 $S\in\mathfrak{C}(E)$. 下面说明 $N(I-S)=\{0_E\}$, 这是因为若

$$0_E=u-Su=(u-Tu)-(\Lambda\circ Pu),$$

因为 $u-Tu, \Lambda\circ Pu$ 活在两个互为余子空间的空间中, 故只能有

$$u-Tu=0_E, \quad \Lambda\circ Pu=0_E$$

也就是说 $u\in N(I-T), \Lambda u=0$, 故只能有 $u=0_E$.

现在把结论 (iii) 应用到 S 上, 可知 $R(I-S)=E$, 但这并不可能! 因为 Λ 不是满射, 故总会存在 $f\in F$ 使得 $f\notin R(\Lambda)$, 此时对于方程 $u-Su=f$, 注意到 $u-Tu\in R(I-T)$, 而 $F\ni f$ 是 $R(I-T)$ 的余子空间, 因此等式成立仅当 $u-Tu=0_E$, 也就是说方程其实是 $-\Lambda\circ Pu=f$, 但后者不在前者的值域内, 因此方程无解.

至此我们已经说明了 $d^*\le d$, 把这一结论再应用到 T^* 上可知

$$\dim N(I-T^{**})\le \dim N(I-T^*)\le \dim N(I-T)$$

又因为 $N(I-T^{**})\supset N(I-T)$, 故只能有 $d=d^*$. □

6.4 紧算子的谱

对紧算子的谱, 我们有下述结论:

定理 6.10 (紧算子的谱性质)

设 X 是 Banach 空间, 若 $A\in\mathfrak{C}(X)$, 则:

(i) 若 $\dim X=\infty$, 则 $0\in\sigma(A)$;

(ii) $\sigma(A)\backslash\{0\}=\sigma_p(A)\backslash\{0\}$;

(iii) $\sigma_p(A)$ 至多以 0 为聚点.

证明 (i) 设 $\dim X=\infty$, 用反证法. 若 $0\notin\sigma(A)$, 则 $0\in\rho(A)$, 这意味着

$$A^{-1}\in L(X).$$

根据紧算子的基本性质6.4(vi) 与 $A\in\mathfrak{C}(X)$ 知 $I=A\circ A^{-1}\in\mathfrak{C}(X)$, 因此对 X 中的任意有界集 M 而言, 其闭包 $\overline{M}$ 均为 X 中的紧集, 特别 X 的闭单位球也是紧集, 因此根据 Riesz 定理6.8知 X 有限维, 矛盾!

(ii) 设 $\lambda\in\sigma(A), \lambda\neq 0$, 往证 λ 必为本征值. 用反证法, 若 λ 不是本征值, 则 $(\lambda I-A)^{-1}$ 存在, 因此必须有 $N(\lambda I-A)=\{0_X\}$. 根据二择一定理6.9(iii), 这表明 $R(\lambda I-A)=X$, 因此 λ 也不可能是连续谱或剩余谱, 这说明 λ 只能在 $\rho(A)$ 内, 矛盾!

(iii) 我们需要下述引理:

引理 6.9

设 $A\in\mathfrak{C}(X)$, $\{\lambda_n\}_{n\in\mathbb{N}}\subset\sigma(A)\backslash\{0\}$ 是两两不同的复数列, 且 $\lambda_n\to\lambda(n\to\infty)$, 则必有 $\lambda=0$.

上述引理也就是在说 $\sigma(A)\backslash\{0_X\}$ 中的所有点都只能是孤立点. 下面证明上述引理. 根据 (ii) 的结果, 已经知道 $\lambda_n\in\sigma_p(A)(\forall n\in\mathbb{N})$, 于是对每个 $n\in\mathbb{N}$ 均有 $\{0_X\}\subsetneq N(A-\lambda_n I)$. 设 $e_n\neq 0_X$ 满足 $(A-\lambda_n I)e_n=0_X$, 记 $E_n=\operatorname{span}\{e_1,e_2,\cdots,e_n\}$, 下面说明 $E_n\subsetneq E_{n+1}$ 对所有 n 均成立, 为此只需说明对所有 n 而言向量 $e_1,e_2,\cdots,e_n$ 均线性无关即可. 考虑归纳法, 设 $e_1,e_2,\cdots,e_n$ 线性无关, 若 $e_{n+1}=\sum_{i=1}^n\alpha_ie_i$, 则

$$Ae_{n+1}=\sum_{i=1}^n\alpha_i\lambda_ie_i=\lambda_{n+1}e_{n+1}=\sum_{i=1}^n\alpha_i\lambda_{n+1}e_i.$$

因此对 $i=1,2,\cdots,n$ 均有 $\alpha_i(\lambda_i-\lambda_{n+1})=0$. 因为 λ_i 两两不同, 故只能有 $\alpha_i=0(i=1,2,\cdots,n)$, 这便说明了 e_{n+1} 与 $e_1,\cdots,e_n$ 线性无关, 于是对全体 n 均有 $E_n\subsetneq E_{n+1}$.

根据 Riesz 引理6.8, 当 $n\ge 2$ 时对每个 E_n 都可以构造 u_n 使得 $\|u_n\|=1, \operatorname{dist}(u_n,E_{n-1})\ge 1/2$. 现在对 $2\le m<n$ 有

$$E_{m-1}\subset E_m\subset E_{n-1}\subset E_n.$$

另一方面, 根据 e_n 的构造显见 $(A-\lambda_n I)E_n \subset E_{n-1}$, 于是

$$\begin{aligned}\left\|\frac{Au_n}{\lambda_n}-\frac{Au_m}{\lambda_m}\right\| &= \left\|\frac{(Au_n-\lambda_n u_n)}{\lambda_n}-\frac{(Au_m-\lambda_m u_m)}{\lambda-m}+u_n-u_m\right\| \\ &\geq \operatorname{dist}(u_n, E_{n-1}) \geq \frac{1}{2}.\end{aligned}$$

若 $\lambda_n \to \lambda, \lambda \neq 0$, 这就说明总能找到某严格正数 β 使得 $\|Au_n - Au_m\| \geq \beta > 0$, 因此 $\{Au_n\}$ 不存在任何收敛子列, 但这与 $A \in \mathfrak{C}(X)$ 矛盾! 引理至此得证.

回到原定理的证明, 对每个整数 $n \geq 1$, 集合

$$\sigma(A) \cap \{\lambda \in \mathbb{C} : |\lambda| \geq 1/n\}$$

要么是空集, 要么是有限集. 这是因为如果上述集合是无限集, 因为 $\sigma(A)$ 本身是有界集, 故上述集合作为 $\mathbb{C}$ 中的有界无限子集, 必存在某收敛到它自身的子列, 但这意味着 $\sigma_p(A)$ 中将出现收敛到 $\lambda(|\lambda| \geq 1/n)$ 的子列, 这与引理6.9矛盾! 因此只要 $\sigma(A)\backslash\{0\}$ 包含了无穷多个值, 这无穷多个值必定以 0 为聚点. □

注 定理6.10表明对于无穷维 Banach 空间上的紧算子 A, 其谱集只可能是下述三种情况:

(i) $\sigma(A) = \{0\}$;

(ii) $\sigma(A) = \{0, \lambda_1, \cdots, \lambda_m\}$, 其中 $\{\lambda_i\}_{i=1}^m$ 非零, 它们均为特征值, 且

$$|\lambda_1| \geq \cdots \geq |\lambda_m|;$$

(iii) $\sigma(A) = \{0, \lambda_1, \lambda_2, \cdots, \lambda_m, \cdots\}$, 其中 $\{\lambda_i\}_{i=1}^\infty$ 非零, 它们均为特征值, 满足模的递增关系

$$|\lambda_1| \geq \cdots \geq |\lambda_m| \geq \cdots$$

且当 $m \to \infty$ 时 $\lambda_m > 0$.

注 对上注 (iii) 其实还有某种意义下的逆命题成立: 任意给定趋零的序列 $\{\lambda_n\}_{n\in\mathbb{N}}$, 总会存在紧算子 A 使得 $\sigma(A) = \{\lambda_n\}_{n\in\mathbb{N}} \cup \{0\}$. 在 l^2 中, 只需把该算子取成乘积算子:

$$A : l^2 \to l^2, \quad u = (u_1, \cdots, u_n, \cdots) \mapsto (\alpha_1 u_1, \cdots, \alpha_n u_n, \cdots)$$

即可. 因为 T 可以视作有穷秩算子 $T_n u = (\alpha_1 u_1, \cdots, \alpha_n u_n, 0, \cdots)$ 的极限, 故 $T \in \mathfrak{C}(l^2)$. 通过这个例子同样可以看出, 定理6.10没有断定 0 是否一定在 $\sigma_p(A)$ 内. 就算 $0 \in \sigma_p(A)$, 也不能依此断定 $N(A)$ 是有限维还是无限维的.

我们下面举出几个例子以体现前述三种情况:

例 6.5 取 $X = l^2, A = 0_{L(l^2)}$, 则对任意 $\lambda \in \mathbb{C}\backslash\{0_{l^2}\}$ 均有 $(\lambda I - A)^{-1} = \lambda^{-1} I \in L(l^2)$, 从而 $\lambda \in \rho(A)$, 因此 $\sigma(A) = \{0\} = \sigma_p(A)$.

例 6.6 取 $X = l^2$, 取定 $\{\lambda_1, \cdots, \lambda_m\} \subset \mathbb{C}$ 满足 $\lambda_i \neq 0 (i \in \{1, \cdots, m\})$, 定义算子

$$A : l^2 \to l^2, \quad \{a_k\}_{k\in\mathbb{N}} \mapsto \{\lambda_1 a_1, \cdots, \lambda_m a_m, 0, \cdots\},$$

此时 $\dim R(A) < \infty$, 下面说明 $A \in L(l^2)$, 且

$$\|A\|_{L(l^2)} = \max\{|\lambda_1|, |\lambda_2|, \cdots, |\lambda_m|\}.$$

一方面显见

$$\|A\|_{L(l^2)} = \sup_{\|\{a_k\}_{k\in\mathbb{N}}\|_{l^2}=1} \left\{\sum_{i=1}^m |\lambda_i a_i|^2\right\}^{\frac{1}{2}} \leq \max\{|\lambda_i| : i \in \{1, \cdots, m\}\}.$$

另一方面, 对任意 $i \in \{1, \cdots, m\}$, 对应令 e_i 是第 i 个分量为 1, 其余分量均为 0 的向量, 则

$$\|A\|_{L(l^2)} \geq \|Ae_1\|_{l^2} = |\lambda_i|$$

因此

$$\|A\|_{L(l^2)} \geq \max\{|\lambda_1|, |\lambda_2|, \cdots, |\lambda_m|\}.$$

从而 $A \in L(l^2)$ 且

$$\|A\|_{L(l^2)} = \max\{|\lambda_1|, |\lambda_2|, \cdots, |\lambda_m|\}.$$

另外 A 作为有穷秩算子显然是紧算子.

接下来考察 $\sigma(A)$. 由 $\dim l^2=\infty$ 与谱性质6.10(i) 知 $0\in\sigma(A)$. 下面我们验证实际上 $0\in\sigma_p(A)$, 且

$$\sigma(A)=\sigma_p(A)=\{0,\lambda_1,\cdots,\lambda_m\}.$$

注意到对前面定义的 e_i 有

$$Ae_i=\lambda_i e_i,\quad i\in\{1,2,\cdots,m\}$$

且

$$Ae_{m+1}=0_{l^2}=0e_{m+1},$$

因此 $\{0,\lambda_1,\cdots,\lambda_m\}\subset\sigma_p(A)$. 现在对 $\lambda\notin\{0,\lambda_1,\cdots,\lambda_m\}$, 要想

$$(\lambda I-A)(\{a_k\}_{k\in\mathbb{N}})=0_{l^2}$$

即

$$\{(\lambda-\lambda_1)a_1,\cdots,(\lambda-\lambda_m)a_m,\lambda a_{m+1},\cdots\}=\{0,\cdots,0,0,\cdots\}$$

就只能有 $\{a_k\}_{k\in\mathbb{N}}=0_{l^2}$, 这说明 $\lambda I-A$ 是单射, 从而 $(\lambda I-A)^{-1}$ 至少存在, 进一步可知

$$(\lambda I-A)^{-1}:l^2\to l^2,\quad \{a_k\}_{k\in\mathbb{N}}\mapsto\left\{\frac{a_1}{\lambda-\lambda_1},\cdots,\frac{a_m}{\lambda-\lambda_m},\frac{a_{m+1}}{\lambda},\cdots\right\},$$

且

$$\begin{aligned}\|(\lambda I-A)^{-1}\|_{L(l^2)}&=\sup_{\|\{a_k\}_{k\in\mathbb{N}}\|_{l^2}=1}\left\{\sum_{i=1}^{m}\left|\frac{a_i}{\lambda-\lambda_i}\right|^2+\sum_{i=m+1}^{\infty}\left|\frac{a_i}{\lambda}\right|^2\right\}^{\frac{1}{2}}\\&\le\max\left\{\frac{1}{|\lambda-\lambda_1|},\cdots,\frac{1}{|\lambda-\lambda_m|},\frac{1}{|\lambda|}\right\}.\end{aligned}$$

又因为对 $\forall i\in\{1,\cdots,m\}$ 有

$$\|(\lambda I-A)^{-1}\|_{L(l^2)}\ge\|(\lambda I-A)^{-1}e_i\|_{l^2}=\frac{1}{|\lambda-\lambda_i|}$$

且

$$\|(\lambda I-A)^{-1}\|_{L(l^2)}\ge\|(\lambda I-A)^{-1}e_{m+1}\|_{l^2}=\frac{1}{|\lambda|},$$

故

$$\|(\lambda I-A)^{-1}\|_{L(l^2)}\|\ge\max\left\{\frac{1}{|\lambda-\lambda_1|},\cdots,\frac{1}{|\lambda-\lambda_m|},\frac{1}{|\lambda|}\right\}.$$

因此 $(\lambda I-A)^{-1}\in L(l^2)$, 这说明 $\lambda\in\rho(A)$, 从而

$$\sigma(A)=\sigma_p(A)=\{0,\lambda_1,\cdots,\lambda_m\}.$$

□

例 6.7 取 $X=l^2$, 设 $\{\lambda_m\}_{m\in\mathbb{N}}\subset\mathbb{C}$ 满足 $\lambda_m\ne0(\forall m\in\mathbb{N})$ 且 $\lambda_m\to0(m\to\infty)$. 定义算子

$$A:l^2\to l^2,\quad \{a_k\}_{k\in\mathbb{N}}\mapsto\{\lambda_k a_k\}_{k\in\mathbb{N}}.$$

因为 $m\to\infty$ 时 $\lambda_m\to0$, 故 $\{|\lambda_m|\}_{m\in\mathbb{N}}$ 有界, 于是

$$\|A\|_{L(l^2)}=\sup_{\|\{a_k\}_{k\in\mathbb{N}}\|_{l^2}=1}\|\{\lambda_k a_k\}_{k\in\mathbb{N}}\|_{l^2}\le\sup_{m\in\mathbb{N}}|\lambda_m|<\infty.$$

这说明 $A\in L(l^2)$. 又因为对 $\forall m\in\mathbb{N}$ 有

$$\|A\|_{L(l^2)}\ge\|Ae_m\|_{l^2}=|\lambda_m|$$

故 $\|A\|_{L(l^2)}\ge\sup_{m\in\mathbb{N}}|\lambda_m|$, 因此 $A\in L(l^2)$ 且

$$\|A\|_{L(l^2)}=\sup_{m\in\mathbb{N}}|\lambda_m|. \tag{6.32}$$

下面说明 $A\in\mathfrak{C}(l^2)$. 因为 l^2 是 Hilbert 空间, 根据命题6.7知只需证明存在有穷秩算子 A_m 使得

$$\|A_m-A\|_{L(l^2)}\to0,\quad m\to\infty$$

即可. 为此对 $\forall m \in \mathbb{N}$ 定义 A_m 为前例中的算子:

$$A_m : \{a_k\}_{k\in\mathbb{N}} \mapsto \{\lambda_1 a_1, \cdots, \lambda_m a_m, 0, \cdots\}$$

由 $\lambda_m \to 0(m \to 0)$ 立得 $\|A_m - A\|_{L(l^2)} \to 0(m \to \infty)$, 因此 $A \in \mathfrak{C}(l^2)$.

下面研究 A 的谱. 显见对任意 $i \in \mathbb{N}$ 均有 $Ae_i = \lambda_i e_i$, 故

$$\{\lambda_1, \cdots, \lambda_m, \cdots\} \subset \sigma_p(A).$$

又有 $\dim l^2 = \infty$ 与谱性质6.10(i) 知 $0 \in \sigma(A)$.

下面说明 $0 \in \sigma_c(A)$. 因为对 $\forall i \in \mathbb{N}$ 而言均有 $\lambda_i \neq 0$, 故 A 是单射, 从而 A^{-1} 至少存在, 于是 $0 \notin \sigma_p(A)$. 又因为 $0 \in \sigma(A)$, 故要么 $0 \in \sigma_c(A)$, 要么 $0 \in \sigma_r(A)$, 这两种情况均指向 $R(A) \subsetneq l^2$.

现在说明 $\overline{R(A)} = l^2$. 事实上, 任取 $b = \{b_1, \cdots, b_m, \cdots\} \in l^2$, 对 $\forall m \in \mathbb{N}$ 令

$$a_m := \left\{\frac{b_1}{\lambda_1}, \cdots, \frac{b_m}{\lambda_m}, 0, \cdots\right\}$$

则 $a_m \in l^2, Aa_m = \{b_1, \cdots, b_m, 0, \cdots\}$, 且当 $m \to \infty$ 时有

$$\|b - Aa_m\|_{l^2} = \left\{\sum_{i=m+1}^{\infty} |b_i|^2\right\}^{\frac{1}{2}} \to 0.$$

因此 $b \in \overline{R(A)}$, 从而 $\overline{R(A)} = l^2$, 故 $0 \in \sigma_c(A)$.

现在对于

$$\lambda \notin \{0, \lambda_1, \cdots, \lambda_m, \cdots\}$$

由 $\lambda_k \to 0(k \to \infty)$ 知 $|\lambda - \lambda_k| \to |\lambda| > 0(k \to \infty)$, 故存在 $k_0 \in \mathbb{N}$ 使得 $k \geq k_0$ 时 $|\lambda - \lambda_k| \geq \frac{|\lambda|}{2}$, 此时对任意 $k \in \mathbb{N}$ 有

$$|\lambda - \lambda_k| \geq \min\left\{\frac{|\lambda|}{2}, |\lambda - \lambda_1|, \cdots, |\lambda - \lambda_{k_0-1}|\right\} =: C_\lambda > 0.$$

于是

$$(\lambda I - A)^{-1} : l^2 \to l^2, \quad \{a_k\}_{k\in\mathbb{N}} \mapsto \left\{\frac{a_1}{\lambda - \lambda_1}, \cdots, \frac{a_m}{\lambda - \lambda_m}, \cdots\right\}$$

是有界算子, 且容易说明

$$\|(\lambda I - A)^{-1}\|_{L(l^2)} = \sup\left\{\frac{1}{|\lambda - \lambda_1|}, \cdots, \frac{1}{|\lambda - \lambda_m|}, \cdots\right\}$$

故 $\lambda \in \rho(A)$, 因此

$$\sigma(A) = \sigma_c(A) \cup \sigma_p(A) = \{0, \lambda_1, \cdots, \lambda_m, \cdots\}.$$

□

例 6.8 取 $X = l^2$, $\{\lambda_k\}_{k\in\mathbb{N}}$ 的设置与上例相同, 定义

$$T : l^2 \to l^2, \quad \{a_k\}_{k\in\mathbb{N}} \mapsto \{0, \lambda_1 a_1, \cdots, \lambda_m a_m, \cdots\},$$

则 $T \in \mathfrak{C}(l^2)$ 且 $\sigma(T) = \sigma_r(T) = \{0\}$.

事实上, 记

$$\widetilde{A} : l^2 \to l^2, \quad \{a_k\}_{k\in\mathbb{N}} \mapsto \{0, a_1, \cdots, a_m, \cdots\}$$

则 $\widetilde{A}$ 作为右推移算子显然有界, 记 A 是上例中定义的算子, 由 $A \in \mathfrak{C}(l^2)$, 紧算子的基本性质6.4(vi) 与 $T = \widetilde{A}A$ 即知 $T \in \mathfrak{C}(l^2)$.

下面研究 $\sigma(T)$. 因为 $\dim l^2 = \infty$, 故 $0 \in \sigma(T)$. 若 $\lambda \in \mathbb{C}$ 为 T 的特征值, 则必存在 $\{a_k\}_{k\in\mathbb{N}} \in l^2 \backslash \{0_{l^2}\}$ 使得

$$(\lambda I - T)(\{a_k\}_{k\in\mathbb{N}}) = 0_{l^2}.$$

代入 T 的定义可知

$$(\lambda a_1, \lambda a_2 - \lambda_1 a_1, \cdots, \lambda a_{k+1} - \lambda_k a_k, \cdots) = 0_{l^2}.$$

这说明

$$\begin{cases} \lambda a_1 = 0, \\ \lambda a_{k+1} = \lambda_k a_k, \quad \forall k \in \mathbb{N}. \end{cases}$$

若 $\lambda = 0$, 则对 $\forall k \in \mathbb{N}$ 均有 $\lambda_k a_k = 0$, 又因为 $\lambda_k \neq 0 (\forall k \in \mathbb{N})$, 故只能有 $a_k = 0$, 因此

$$\{a_k\}_{k\in\mathbb{N}} = 0_{l^2},$$

矛盾! 而若 $\lambda \neq 0$, 则由 $\lambda a_1 = 0$ 可知 $a_1 = 0$, 进而推知 $a_k = 0 (\forall k \in \mathbb{N})$, 同样矛盾, 故 T 没有特征值, 即 $\sigma_p(T) = \varnothing$. 又显见 $\overline{R(T)} \neq l^2$, 故 $\{0\} \in \sigma_r(T)$. 由谱性质6.10(ii) 即知 $\sigma(T) = \{0\} = \sigma_r(T)$. □

6.5 不变子空间

定义 6.11 (不变子空间 (invariant subspace))

设 X 是 Banach 空间, $A \in L(X)$, $M \subset X$ 是线性子空间. 若 $A(M) \subset M$, 就称 M 为算子 A 的不变子空间. ♣

从定义出发可得下述结论:

命题 6.8 (不变子空间的基本性质)

设 X 是 Banach 空间, $A \in L(X)$, 则

(i) $\{0_X\}$ 与 X 都是 A 的不变子空间, 称这两个不变子空间为 A 的平凡不变子空间.

(ii) 若 M 是 A 的不变子空间, 则 $\overline{M}$ 是 A 的闭不变子空间.

(iii) 若 $\lambda \in \sigma_p(A)$, 则 $N(\lambda I - A)$ 是 A 的闭不变子空间.

(iv) 对 $\forall y \in X$, 若记

$$L_y := \{P(A)y : P \text{ 是数域 } \mathbb{K} \text{ 上的任意一元多项式}\}$$

则 L_y 是 A 的不变子空间. ♠

证明 (i) 显然.

(ii) 只需说明对 $\forall x \in \overline{M}$ 均有 $Ax \in \overline{M}$. 事实上, 因为 $x \in \overline{M}$, 故存在 $\{x_n\}_{n\in\mathbb{N}} \subset M$ 使得 $x_n \to x (n \to \infty)$. 因为 $A(M) \subset M$, 故 $\{Ax_n\}_{n\in\mathbb{N}} \subset M$. 又因为 $A \in L(X)$, 故当 $n \to \infty$ 时有

$$Ax_n \to Ax.$$

因此 $Ax \in \overline{M}$, 从而 $\overline{M}$ 是 A 的 (闭) 不变子空间.

(iii) 若 $\lambda \in \sigma_p(A)$, 则 $(\lambda I - A)^{-1}$ 不存在, 因此 $N(\lambda I - A) \neq \{0_X\}$, 且 $N(\lambda I - A)$ 为 X 的一个线性子空间. 现任取 $x \in N(\lambda I - A)$, 知 $(\lambda I - A)x = 0_X$, 即

$$Ax = \lambda x \in N(\lambda I - A),$$

这说明 $N(\lambda I - A)$ 是不变子空间.

下面说明 $N(\lambda I - A)$ 是闭的. 任取 $x \in \overline{N(\lambda I - A)}$, 知存在 $\{x_n\}_{n\in\mathbb{N}} \subset N(\lambda I - A)$ 使得 $x_n \to x (n \to \infty)$. 又对 $\forall n \in \mathbb{N}$ 有

$$x_n \in N(\lambda I - A) \Leftrightarrow (\lambda I - A)x_n = 0_X \Leftrightarrow Ax_n = \lambda x_n.$$

在上右式两端对 n 取极限, 由 A 的连续性即知

$$\lambda x = \lim_{n\to\infty} Ax_n = Ax.$$

因此 $x \in N(\lambda I - A)$, 故 $N(\lambda I - A)$ 是闭的.

(iv) 任取 $x \in L_y$, 知存在多项式 P 使得 $x = P(A)y$. 注意若 P 是 $\mathbb{K}$ 上的多项式, 则 $xP(x)$ 依旧是 $\mathbb{K}$ 上的多项式,

因此

$$Ax = AP(A)y \in L_y,$$

故 L_y 是 A 的不变子空间. □

定理 6.11 (紧算子不变子空间的存在性)

设 X 是 Banach 空间, $\dim X \geq 2$, 则对任意 $A \in \mathfrak{C}(X)$ 而言, A 均有非平凡的闭不变子空间.

证明 若 $A = 0_{L(X)}$, 任取 $y \in X \backslash \{0_X\}$. 因为 $\dim X \geq 2$, 故

$$\{0_X\} \subsetneq \operatorname{span}\{y\} \subsetneq X$$

且

$$A(\operatorname{span}\{y\}) = \{0_X\} \subset \operatorname{span}\{y\}$$

这说明 $\operatorname{span}\{y\}$ 就是 A 的一个非平凡闭不变子空间.

下设 $A \neq 0_{L(X)}, \|A\|_{L(X)} = 1$. 若 $\dim X = n < \infty$, 则 A 成为一个 $n \times n$ 矩阵. 根据代数基本定理知必存在 $\lambda_0 \in \mathbb{C}$ 使得 $|A - \lambda_0 I| = 0$, 因此 $(A - \lambda_0 I)x = 0_X$ 必存在非零解 x. 由 $\dim X \geq 2$ 知 $\{\lambda x : \lambda \in \mathbb{C}\}$ 为 A 的一个非平凡闭不变子空间.

若 $\dim X = \infty$, 我们首先说明可设 $\sigma_p(A) \backslash \{0\} = \varnothing$. 这是因为当 $\sigma_p(A) \backslash \{0\} \neq \varnothing$ 时, 由 $A \in \mathfrak{C}(X) \subset L(X)$ 与不变子空间的基本性质6.8(iii) 知, 对 $\forall \lambda \in \sigma_p(A) \backslash \{0\}$ 而言, $N(\lambda I - A)$ 都是 A 的闭不变子空间. 根据 $\sigma_p(A)$ 的定义显见 $\{0_X\} \subsetneq N(\lambda I - A)$. 若 $N(\lambda I - A) = X$, 则

$$X = N(\lambda I - A) = \{x \in X : Ax = \lambda x\}$$

设 S_1 是 X 的单位球面, 则对任意 $\{x_n\}_{n\in\mathbb{N}} \subset S_1 \subset X = N(\lambda I - A)$ 均有

$$Ax_n = \lambda x_n, \quad \forall n \in \mathbb{N}. \tag{6.33}$$

又因为 $\{x_n\}_{n\in\mathbb{N}}$ 本身有界, 而 $A \in \mathfrak{C}(X)$, 故 $\{Ax_n\}_{n\in\mathbb{N}}$ 存在收敛子列 $\{Ax_{n_k}\}_{k\in\mathbb{N}}$, 由(6.33)式与 $\lambda \neq 0$ 可知 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 也收敛. 因为 $\{x_n\}_{n\in\mathbb{N}}$ 是在 S_1 上任取的, 这便说明 S_1 是 X 中的列紧集. 但从 Riesz 定理6.8出发可以说明此时 X 是有限维空间, 矛盾! 故必有 $N(\lambda I - A) \subsetneq X$, 因此 $N(\lambda I - A)$ 为 A 的非平凡闭不变子空间. 至此我们只需要讨论 $\sigma_p(A) \backslash \{0\} = \varnothing$ 的情况就可以了.

根据紧算子的谱性质6.10(i),(ii), 此时只能有 $\sigma(A) = \{0\}$, 因此 $r_\sigma(A) = 0$.

下面用反证法证明 A 有非平凡闭不变子空间. 若 A 不存在非平凡闭不变子空间, 则由基本性质6.8(ii),(iv) 知对任意 $y \in X \backslash \{0_X\}$ 而言均有 $\overline{L_y} = X$. 因为 $\|A\|_{L(X)} = 1$, 故存在 $\tilde{x}_0 \in X$ 使得 $\|A\tilde{x}_0\|_X \neq 0$, 从而当 $n \to \infty$ 时有

$$\|A(n\tilde{x}_0)\|_X = n\|A\tilde{x}_0\|_X \to \infty.$$

由此可知必存在 $x_0 \in X$ 使得 $\|Ax_0\|_X > 1$, 进而必有 $\|x_0\|_X > 1$, 否则

$$\|Ax_0\|_X \leq \|A\|_{L(X)} \|x_0\|_X \leq 1,$$

这与 $\|Ax_0\|_X > 1$ 矛盾. 现令 $C := \overline{A(B(x_0, 1))}$, 根据 $A \in \mathfrak{C}(X)$ 知 C 是紧集. 下面说明 $0_X \notin C$. 如若不然, 则存在 $\{x_n\}_{n\in\mathbb{N}}$ 使得 $Ax_n \to 0_X (n \to \infty)$, 但另一方面

$$\|Ax_n - Ax_0\|_X \leq \|A\|_{L(X)} \|x_n - x_0\|_X < 1$$

而 $\|Ax_0\|_X > 1$, 故必有 $\|Ax_n\|_X > 0 (\forall n \in \mathbb{N})$, 矛盾! 故 $0_X \notin C$.

现在对 $\forall y_0 \in C$, 因为 $y_0 \neq 0_X$, 根据反证法的假设知 $\overline{L_{y_0}} = X$, 因此 $L_{y_0} \cap B(x_0, 1) \neq \varnothing$. 设 L_{y_0} 中的元素 $T_{y_0} y_0$ 满足 $\|T_{y_0} y_0 - x_0\|_X < 1$(其中 T_{y_0} 是关于 A 的多项式, $T_{y_0} \neq 0_{L(X)}$(否则 $\|x_0\|_X < 1$, 矛盾!)), 记

$$\varepsilon := 1 - \|T_{y_0} y_0 - x_0\|_X, \quad \delta_{y_0} := \frac{\varepsilon}{\|T_{y_0}\|_{L(X)}}.$$

则 $\delta_{y_0} > 0$, 且当 $\|y - y_0\|_X < \delta_{y_0}$ 时有

$$\|T_{y_0} y - T_{y_0} y_0\|_X \leq \|T_{y_0}\|_{L(X)} \|y - y_0\| < \varepsilon.$$

于是对 $\forall y\in B(y_0,\delta_{y_0})$ 有

$$\begin{aligned}\|T_{y_0}y-x_0\|_X&\le\|T_{y_0}y-T_{y_0}y_0\|_X+\|T_{y_0}y_0-x_0\|_X\\&<\varepsilon+\|T_{y_0}y_0-x_0\|_X=1.\end{aligned}\tag{6.34}$$

注意到

$$C\subset\bigcup_{y\in C}B(y,\delta_y),$$

由 C 的紧性知存在 $\{y_1,\cdots,y_n\}\subset C$ 与 $\delta_i:=\delta_{y_i}(i\in\{1,\cdots,n\})$ 使得

$$C\subset\bigcup_{i=1}^{n}B(y_i,\delta_{y_i}).\tag{6.35}$$

因此对 $\forall y\in C$, 总会存在 $i_1\in\{1,\cdots,n\}$ 使得

$$y\in B(y_{i_1},\delta_{i_1}).$$

于是由(6.34)式知

$$\|T_{y_{i_1}}y-x_0\|_X<1.$$

即 $T_{y_{i_1}}y\in B(x_0,1)$. 又因为 $T_{y_{i_1}}$ 是 A 的多项式, 故

$$T_{y_{i_1}}Ay=AT_{y_{i_1}}y\in C,$$

又因为 $0_X\notin C$, 故 $T_{y_{i_1}}\neq 0_X$. 由此及(6.35)式知存在 $i_2\in\{1,\cdots,n\}$ 使得

$$T_{y_{i_1}}Ay\in B(y_{i_2},\delta_{i_2}).$$

重复上述过程可得

$$\|T_{y_{i_2}}T_{y_{i_1}}Ay-x_0\|_X<1,$$

此即 $T_{y_{i_2}}T_{y_{i_1}}Ay\in B(x_0,1)$. 又因为 $T_{y_{i_2}}$ 同样是 A 的多项式, 故

$$T_{y_{i_2}}T_{y_{i_1}}A^2y=AT_{y_{i_2}}T_{y_{i_1}}Ay\in C,$$

又因为 $0_X\notin C$, 故 $T_{y_{i_2}}\neq 0_X$. 依此类推, 存在 $\{i_k\}_{k\in\mathbb{N}}\subset\{1,\cdots,n\}$ 使得对 $\forall k\in\mathbb{N}$ 均有

$$\left\|\prod_{j=1}^{k}T_{y_{i_j}}(A^{k-1}y)-x_0\right\|_X<1.$$

由此可得对 $\forall k\in\mathbb{N}$ 有

$$\|x_0\|_X-1<\left\|\prod_{j=1}^{k}T_{y_{i_j}}(A^{k-1}y)\right\|_X.\tag{6.36}$$

定义

$$\mu:=\max_{j\in\mathbb{N}}\|T_{y_{i_j}}\|_{L(X)}.\tag{6.37}$$

由 $i_j\in\{1,\cdots,n\}$ 可知 $\mu<\infty$, 又因为 $0_X\notin C$, 故 $\mu>0$. 因此对 $\forall k\in\mathbb{N}$, 由(6.36)式可得

$$\|x_0\|_X-1<\mu^k\|A^{k-1}y\|_X\le\mu^k\|A^{k-1}\|_{L(X)}\|y\|_X.$$

整理可得

$$\frac{1}{\mu}\left(\frac{\|x_0\|_X-1}{\mu\|y\|_X}\right)^{\frac{1}{k-1}}<\|A^{k-1}\|_{L(X)}^{\frac{1}{k-1}},\quad\forall k\in\mathbb{N}.$$

令 $k\to\infty$, 由 Gelfand 定理6.4知

$$\frac{1}{\mu}\leftarrow\frac{1}{\mu}\left(\frac{\|x_0\|_X-1}{\mu\|y\|_X}\right)^{\frac{1}{k-1}}<\|A^{k-1}\|_{L(X)}^{\frac{1}{k-1}}\to r_\sigma(A)=0$$

这与 $\mu<\infty$ 矛盾! 故 A 有非平凡闭不变子空间. □

6.6 Hilbert-Schmidt 定理

本节主要讨论对称紧算子的特征值表现. 首先我们给出对称算子的定义:

定义 6.12 (对称算子 (symmetric operator))

设 H 是 Hilbert 空间, $A \in L(H)$. 若

$$\langle Ax, y\rangle = \langle x, Ay\rangle, \quad \forall x, y \in H$$

则称 A 是对称算子.

注 回忆共轭算子 A^* 指的是对任意 $x, y \in X$ 均有

$$\langle Ax, y\rangle = \langle x, A^* y\rangle,$$

因此 Hilbert 空间上的有界线性算子 A 对称当且仅当 $A = A^*$, 因此又称 A 为**自共轭算子 (self-conjugate operator)** 或**自伴算子 (self-adjoint operator)**.

例 6.9 设 $X = L^2(\Omega, \mu)$ 以 $\mathbb{R}$ 为背景数域, $K \in L^2(\Omega\times\Omega, \mu\otimes\mu)$ 满足对 $\forall x, y \in \Omega$ 均有 $K(x, y) = K(y, x)$, 则算子

$$A : u(x) \mapsto \int_\Omega K(x, y)u(y)d\mu(y)$$

是 $L^2(\Omega, \mu)$ 上的对称算子. 这是因为对任取 $u \in L^2(\Omega, \mu)$ 满足 $\|u\|_{L^2(\Omega,\mu)} \le 1$, 由 Hölder 不等式知

$$\begin{aligned}\|Au\|_{L^2(\Omega,\mu)} &= \left[\int_\Omega\left|\int_\Omega K(x, y)u(y)d\mu(y)\right|^2 d\mu(x)\right]^{\frac{1}{2}}\\ &\le \left[\int_\Omega\int_\Omega |K(x, y)|d\mu(y)d\mu(x)\right]^{\frac{1}{2}}\|u\|_{L^2(\Omega,\mu)}\\ &\le \|K\|_{L^2(\Omega\times\Omega,\mu\otimes\mu)} < \infty.\end{aligned}$$

这便说明了 $A \in L^2(\Omega, \mu)$.

下面说明 A 对称. 为此任取 $u, v \in L^2(\Omega, \mu)$, 首先说明 $K(x, y)u(y)v(x)$ 在 $\Omega\times\Omega$ 上可测, 这只需说明 $u(y)v(x)$ 在 $\Omega\times\Omega$ 上可测即可. 注意到对任意 $\lambda \in \mathbb{R}$ 有

$$\{(x, y) \in \Omega\times\Omega : u(y) > \lambda\} = \Omega\times\{y \in \Omega : u(y) > \lambda\}$$

而上右式可测, 故 $u(y)$ 是 $\Omega\times\Omega$ 上的可测函数, 同理 $v(x)$ 也是 $\Omega\times\Omega$ 上的可测函数, 因此 $u(y)v(x)$ 在 $\Omega\times\Omega$ 上可测. 现在根据 Hölder 不等式知

$$\begin{aligned}\iint_{\Omega\times\Omega}|K(x, y)u(y)v(x)|d\mu(x)d\mu(y) &\le \left[\iint_{\Omega\times\Omega}|K(x, y)|^2 d\mu(x)d\mu(y)\right]^{\frac{1}{2}}\left[\iint_{\Omega\times\Omega}|u(y)v(x)|^2 d\mu(x)d\mu(y)\right]^{\frac{1}{2}}\\ &= \|K\|_{L^2(\Omega\times\Omega,\mu\otimes\mu)}\|u\|_{L^2(\Omega,\mu)}\|v\|_{L^2(\Omega,\mu)} < \infty.\end{aligned}$$

因此 $K(x, y)u(y)v(x) \in L^1(\Omega\times\Omega, \mu\otimes\mu)$, 由此可知对 $\forall u, v \in L^2(\Omega, \mu)$, 根据 Fubini 定理有

$$\begin{aligned}\langle Au, v\rangle &= \int_\Omega\left[\int_\Omega K(x, y)u(y)d\mu(y)\right]v(x)d\mu(x)\\ &= \int_\Omega u(y)\left[\int_\Omega K(x, y)v(x)d\mu(x)\right]d\mu(y)\\ &= \langle u, Av\rangle.\end{aligned}$$

由此可知 A 对称. □

例 6.10 设 H 是 Hilbert 空间, M 是 H 的闭线性子空间, 则 $H \to M$ 的投影算子 P_M 是对称的. 这是因为根据正交投影定理5.1知对 $\forall x, y \in H$ 存在唯一正交分解:

$$x = x_M + x_{M^\perp}, \quad x_M \in M, x_{M^\perp} \in M^\perp$$
$$y = y_M + y_{M^\perp}, \quad y_M \in M, y_{M^\perp} \in M^\perp$$

其中 $M^{\perp}$ 表示 M 的正交补:

$$M^{\perp} = \{x \in X : (x, y) = 0, \forall y \in M\}.$$

现在 $P_M : H \to M, x \mapsto x_M$, 故 $x_M = P_M x, y_M = P_M y$, 从而

$$\begin{aligned}(P_M x, y) &= (x_M, y) = (x_M, y_M + y_{M^{\perp}}) \\ &= (x_M, y_M) = (x_M + x_{M^{\perp}}, y_M) \\ &= (x, P_M y),\end{aligned}$$

此即 P_M 是对称算子. □

命题 6.9 (对称算子的基本性质)

设 H 是 Hilbert 空间, $A \in L(H)$, 则:

(i) A 对称当且仅当对任意 $x \in H$ 有 $(Ax, x) \in \mathbb{R}$.

(ii) 若 A 对称, 则 $\sigma(A) \subset \mathbb{R}$, 且对任意 $x \in H$ 与满足 $\operatorname{Im}\lambda \neq 0$ 的 $\lambda \in \mathbb{C}$ 有

$$\|(\lambda I - A)^{-1} x\|_X \leq \frac{1}{|\operatorname{Im}\lambda|}\|x\|_X.$$

(iii) 设 A 是 H 上的对称算子, H_1 是 A 的一个闭不变线性子空间, 则 $A|_{H_1}$ 也是 H_1 上的对称算子.

(iv) 若 A 对称, $\lambda, \widetilde{\lambda} \in \sigma_p(A)$ 且 $\lambda \neq \widetilde{\lambda}$, 则

$$N(\lambda I - A) \perp N(\widetilde{\lambda} I - A).$$

(v) 若 A 对称, 则

$$\|A\|_{L(H)} = \sup_{\|x\|_H = 1} |(Ax, x)|.$$

♠

证明 (i) 任取 $x, y \in H$, 令

$$a(x, y) := (Ax, y),$$

显见 $a(\cdot, \cdot)$ 是 H 上的共轭双线性函数. 根据 a 的构造可知对任意对称算子 A 与任意 $x, y \in H$ 有

$$\begin{aligned}&(Ax, y) = (x, Ay) \\ \Leftrightarrow\, &a(x, y) = (Ax, y) = (x, Ay) = \overline{(Ay, x)} = \overline{a(y, x)}.\end{aligned}$$

令 $x = y$ 知

$$a(x, x) = (Ax, x) = \overline{(Ax, x)} = \overline{a(x, x)},$$

这说明 $(Ax, x) \in \mathbb{R}$.

反之, 因为 H 是 Hilbert 空间, 故 $(Ax, x) \in \mathbb{R}$ 意味着对任意 $x, y \in H$ 有

$$(A(x + y), x + y) = \overline{(A(x + y), x + y)},$$

且

$$(A(x + iy), x + iy) = \overline{(A(x + iy), x + iy)},$$

将上述两式展开有

$$\begin{aligned}a(x, y) + a(y, x) &= \overline{a(x, y)} + \overline{a(y, x)} \\ -a(x, y) + a(y, x) &= \overline{a(x, y)} - \overline{a(y, x)}\end{aligned}$$

两式相减可知对任意 $x, y \in H$ 有

$$a(x, y) = \overline{a(y, x)},$$

此即 A 对称.

(ii) 设 $\lambda=\mu+i\nu,\nu\neq 0,\mu,\nu\in\mathbb{R}$, 下面证明 $\lambda\in\rho(A)$, 为此首先说明 $\lambda I-A$ 是单射, 事实上:

$$\begin{aligned}\|(\lambda I-A)x\|_H^2&=((\lambda I-A)x,(\lambda I-A)x)\\&=((\mu I-A)x+i\nu x,(\mu I-A)x+i\nu x)\\&=\|(\mu I-A)x\|_H^2+((\mu I-A)x,i\nu x)+(i\nu x,(\mu I-A)x)+|\nu|^2\|x\|_H^2\\&=\|(\mu I-A)x\|_H^2+|\nu|^2\|x\|_H^2\geq|\nu|^2\|x\|_H^2.\end{aligned}$$

亦即

$$\|(\lambda I-A)x\|_H\geq|\nu|\|x\|_H. \tag{6.38}$$

因此 $\lambda I-A$ 是单射, 从而 $(\lambda I-A)^{-1}$ 至少存在.

通过(6.38)式还可得到 $R(\lambda I-A)$ 闭. 这是因为设 $\{y_n\}_{n\in\mathbb{N}}$ 是 $R(\lambda I-A)$ 中的基本列, 则存在 $\{x_n\}_{n\in\mathbb{N}}\subset H$ 使得

$$y_n=(\lambda I-A)x_n,$$

由(6.38)式知 $\{x_n\}_{n\in\mathbb{N}}$ 为 H 中的基本列, 由 H 的完备性可知存在 $x\in H$ 使得 $x_n\to x(n\to\infty)$. 进一步由 $\lambda I-A\in L(H)$ 可知

$$y_n\to(\lambda I-A)x\in R(\lambda I-A),\quad n\to\infty.$$

因此 $R(\lambda I-A)$ 是 H 中的闭集.

下面说明 $R(\lambda I-A)=H$, 为此考虑说明 $(R(\lambda I-A))^{\perp}=\{0_H\}$. 根据 A 的对称性知

$$\begin{aligned}y\in(R(\lambda I-A))^{\perp}\Leftrightarrow&((\lambda I-A)x,y)=0,\forall x\in H\\\Leftrightarrow&\lambda(x,y)=(Ax,y),\forall x\in H\\\Leftrightarrow&\lambda(x,y)=(x,Ay),\forall x\in H\\\Leftrightarrow&(x,(\bar{\lambda}I-A)y)=0,\forall x\in H\\\Leftrightarrow&(\bar{\lambda}I-A)y=0_H.\end{aligned}$$

这说明 $y\in N(\bar{\lambda}I-A)$. 又由(6.38)式知对任意 $y\in H$ 均有

$$\|(\bar{\lambda}I-A)y\|_H\geq|\nu|\|y\|_H$$

由 $\Im\lambda=\nu\neq 0$ 可知 $y\in N(\bar{\lambda}I-A)\Leftrightarrow y=0_H$, 故

$$(R(\lambda I-A))^{\perp}=\{0_H\}. \tag{6.39}$$

现在通过(6.39)式说明 $R(\lambda I-A)=H$. 事实上若 $R(\lambda I-A)\subsetneq H$, 则存在 $x\in H$ 满足 $x\notin R(\lambda I-A)$. 因为 $R(\lambda I-A)$ 是 H 的闭线性子空间, 故 x 关于 $R(\lambda I-A)$ 的正交分解投影 $x_{(R(\lambda I-A))^{\perp}}$ 不能是 0_H, 否则 $x\in R(\lambda I-A)$. 但另一方面, 这说明

$$(R(\lambda I-A))^{\perp}\neq\{0_H\}$$

矛盾! 故 $R(\lambda I-A)=H$. 由此可得 $\lambda\in\rho(A)$, 这说明 $\sigma(A)\subset\mathbb{R}$. 由(6.38)式即知对任意满足 $\operatorname{Im}\lambda\neq 0$ 的 $\lambda\in\mathbb{C}$ 与 $x\in H$ 有

$$\|(\lambda I-A)^{-1}x\|_H\leq\frac{1}{|\operatorname{Im}\lambda|}\|x\|_H.$$

注 本注给出(6.38),(6.39)两式的另外一种证法.

对于(6.38)式, 根据 A 的对称性与 (i) 知 $(Ax,x)\in\mathbb{R}$, 于是

$$|((\lambda I-A)x,x)|=|\mu(x,x)-(Ax,x)+i\nu(x,x)|\geq|\nu|\|x\|_H^2. \tag{6.40}$$

又由 Cauchy-Schwarz 不等式知

$$|((\lambda I-A)x,x)|\leq\|(\lambda I-A)x\|_H\|x\|_H$$

于是

$$|\nu|\|x\|_H^2 \le \|(\lambda I - A)x\|_H\|x\|_H$$

此即(6.38)式.

对于(6.39)式, 任取 $x \in (R(\lambda I - A))^{\perp}$, 根据定义知

$$((\lambda I - A)x, x) = 0.$$

由上式与(6.40)式知只能有 $x = 0_H$, 于是

$$(R(\lambda I - A))^{\perp} \subset \{0_H\}.$$

又显见 $0_H \in (R(\lambda I - A))^{\perp}$, 故 $(R(\lambda I - A))^{\perp} = \{0_H\}$.

(iii) 因为 H_1 是 A 的不变子空间, 故对任意 $x, y \in H_1$ 知 $Ax, Ay \in H_1$, 又由 A 的对称性可得 $A|_{H_1}$ 的对称性.

(iv) 由 (ii) 知 $\lambda, \widetilde{\lambda} \in \mathbb{R}$. 现若 $x \in N(\lambda I - A), \widetilde{x} \in N(\widetilde{\lambda} I - A)$, 则

$$\lambda(x, \widetilde{x}) = (Ax, \widetilde{x}) = (x, A\widetilde{x}) = \widetilde{\lambda}(x, \widetilde{x}),$$

由 $\lambda \ne \widetilde{\lambda}$ 知只能有 $(x, \widetilde{x}) = 0$, 因此 $N(\lambda I - A) \perp N(\widetilde{\lambda} I - A)$.

(v) 记

$$\widetilde{C} := \sup_{\|x\|_H = 1} |(Ax, x)|.$$

由 Cauchy-Schwarz 不等式知 $\widetilde{C} \le \|A\|_{L(H)}$, 下面说明 $\|A\|_{L(H)} \le \widetilde{C}$. 因为 A 对称, 故对 $\forall x, y \in H$ 有

$$\begin{aligned}(A(x+y), x+y) - (A(x-y), x-y) &= 2[(Ax, y) + (Ay, x)] \\ &= 2[(Ax, y) + \overline{(Ax, y)}] \\ &= 4\,\mathrm{Re}(Ax, y).\end{aligned}$$

由上式与平行四边形法则知对任意 $x, y \in H$, 若 $\|x\|_H = \|y\|_H = 1$ 且 $x \ne \pm y$, 则

$$\begin{aligned}\mathrm{Re}(Ax, y) &= \frac{1}{4}[(A(x+y), x+y) - (A(x-y), x-y)] \\ &= \frac{1}{4}\left[\|x+y\|_H^2\left(A\left(\frac{x+y}{\|x+y\|_H}\right), \frac{x+y}{\|x+y\|_H}\right) - \|x-y\|_H^2\left(A\left(\frac{x-y}{\|x-y\|_H}\right), \frac{x-y}{\|x-y\|_H}\right)\right] \\ &\le \frac{\widetilde{C}}{4}[\|x+y\|_H^2 + \|x-y\|_H^2] \\ &= \frac{\widetilde{C}}{2}[\|x\|_H^2 + \|y\|_H^2] = \widetilde{C}.\end{aligned}$$

而若 $x = \pm y$, 则上式显然成立. 现取 $\alpha \in \mathbb{C}, |\alpha| = 1$ 使得

$$\alpha(Ax, y) = |(Ax, y)|,$$

则

$$|(Ax, y)| = \alpha(Ax, y) = (Ax, \overline{\alpha}y) = \mathrm{Re}(Ax, \overline{\alpha}y) \le \widetilde{C}.$$

因此

$$\|A\|_{L(H)} = \sup_{\|x\|_H = \|y\|_H = 1} |(Ax, y)| \le \widetilde{C},$$

故 $\|A\|_{L(H)} = \widetilde{C}$, 由此即得

$$\|A\|_{L(H)} = \sup_{\|x\|_H = 1} |(Ax, x)|.$$

□

对有限维空间的情形, 基本性质6.9(v) 还可以从下述论述中自然得到. 设 $\boldsymbol{A}$ 是 $n \times n$ 实对称阵, 通过代数上的

处理可知存在正交阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{A}=\boldsymbol{U}^T\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}\boldsymbol{U},$$

其中 $\{\lambda_1,\cdots,\lambda_n\}$ 是 $\boldsymbol{A}$ 的特征值. 现设 $\boldsymbol{U}^T=(\xi_1,\cdots,\xi_n)$, 则

$$\boldsymbol{A}\boldsymbol{U}^T=\boldsymbol{U}^T\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix},$$

也就是说对 $\forall i\in\{1,\cdots,n\}$ 有

$$\boldsymbol{A}\xi_i=\lambda_i\xi_i.$$

下面说明特征值 $\{\lambda_1,\cdots,\lambda_n\}$ 实际上是 $\boldsymbol{A}$ 所对应的二次型 $(\boldsymbol{A}x,x)$ 在单位球面 $\|x\|_{\mathbb{R}^n}=1$ 上的各临界值. 为此不妨设

$$|\lambda_1|\geq\cdots\geq|\lambda_n|,$$

根据正交阵的定义, 对 $\forall i\in\{1,\cdots,n\}$ 而言均有 $\|\xi_i\|_{\mathbb{R}^n}=1$, 且 $\{\xi_i\}_{i=1}^n$ 构成 $\mathbb{R}^n$ 的一组规范正交基. 因此对任意 $x\in\mathbb{R}^n,\|x\|_{\mathbb{R}^n}=1$, 均存在 $\{a_i\}_{i=1}^n\subset[-1,1]$ 使得

$$\left[\sum_{i=1}^n|a_i|^2\right]^{\frac{1}{2}}=1,\quad x=\sum_{i=1}^n a_i\xi_i. \tag{6.41}$$

因此

$$\begin{aligned}(\boldsymbol{A}x,x)&=\left(\sum_{i=1}^n a_i\boldsymbol{A}\xi_i,\sum_{i=1}^n a_i\xi_i\right)\\ &=\left(\sum_{i=1}^n a_i\lambda_i\xi_i,\sum_{i=1}^n a_i\xi_i\right)=\sum_{i=1}^n\lambda_i a_i^2.\end{aligned}$$

接下来我们说明 $x=\xi_i(i\in\{1,\cdots,n\})$ 是 $(\boldsymbol{A}x,x)$ 的临界点, 为此需要说明

$$\nabla(\boldsymbol{A}x,x)=(2\lambda_1a_1,\cdots,2\lambda_na_n)$$

在 ξ_i 处退化. 现在已知

$$(\boldsymbol{A}\cdot,\cdot):\mathbb{S}^{n-1}\to\mathbb{R}$$

显然是流形间的光滑映射, 于是考虑该映射在 x 点的切映射有

$$T_x(\boldsymbol{A}\cdot,\cdot):T_x\mathbb{S}^{n-1}\to\mathbb{R},$$

其中 $T_x\mathbb{S}^{n-1}$ 表示 $\mathbb{S}^{n-1}$ 在 x 处的切平面. 几何上不难发现 $T_x\mathbb{S}^{n-1}=(\mathrm{span}\{x\})^\perp$, 于是对任意 $i\in\{1,\cdots,n\}$ 有

$$y\in T_{\xi_i}\mathbb{S}^{n-1}=(\mathrm{span}\{\xi_i\})^\perp\Leftrightarrow y=\sum_{j\neq i}b_j\xi_j.$$

从而

$$[\nabla(\boldsymbol{A}\xi_i,\xi_i)]\cdot y=(0,\cdots,0,2\lambda_i,0,\cdots,0)\begin{pmatrix}b_1\\ \vdots\\ b_{i-1}\\ 0\\ b_{i+1}\\ \vdots\\ b_n\end{pmatrix}=0$$

因此 $\nabla(\boldsymbol{A}x,x)$ 在 ξ_i 处退化, 类似可证 $\nabla(\boldsymbol{A}x,x)$ 在 $-\xi_i$ 处也退化, 故 $x=\xi_i$ 确为 $(\boldsymbol{A}x,x)$ 的临界点. 下面说明对任意 $i\in\{1,\cdots,n\}$ 均有

$$|\lambda_i|=\sup\{|\boldsymbol{A}x,x|:\|x\|_{\mathbb{R}^n}=1,x\perp\operatorname{span}\{\xi_1,\cdots,\xi_{i-1}\}\}.\tag{6.42}$$

记(6.42)右式为 d, 取 $x=\xi_i$ 知

$$d\geq|(\boldsymbol{A}\xi_i,\xi_i)|=|\lambda_i|.$$

另一方面, 对任意满足 $\|x\|_{\mathbb{R}^n}=1$ 的 $x\perp\operatorname{span}\{\xi_1,\cdots,\xi_{i-1}\}$, 由(6.41)式与 Cauchy-Schwarz 不等式知

$$\begin{aligned}|(\boldsymbol{A}x,x)|&=\left|\left(\sum_{j=1}^n a_j\boldsymbol{A}\xi_j,x\right)\right|=\left|\left(\sum_{j=1}^n a_j\lambda_j\xi_j,x\right)\right|\\&=\left|\left(\sum_{j=i}^n a_j\lambda_j\xi_j,x\right)\right|\leq\left|\sum_{j=i}^n a_j\lambda_j\xi_j\right|\\&=\left(\sum_{j=i}^n a_j^2|\lambda_j|^2\right)^{\frac{1}{2}}\leq|\lambda_i|\end{aligned}$$

因此 $d\leq|\lambda_i|$, 故 $\lambda=d$. □

下面的定理表明对于对称紧算子而言, 基本性质6.9(v) 中的上确界是可以达到的.

定理 6.12

若 A 是 Hilbert 空间 H 上的对称紧算子, 则存在 $x_0\in H$ 满足 $\|x_0\|_H=1$, 且

$$|(Ax_0,x_0)|=\sup_{\|x\|_H=1}|(Ax,x)|=\|A\|_{L(H)},$$

另有

$$Ax_0=(Ax_0,x_0)x_0.$$

证明 当 $A=0_{L(H)}$ 时, 定理结论自动成立, 下设 $A\neq0_{L(H)}$. 记 S_1 为 H 的单位球面, 因为 A 对称, 故由基本性质6.9(i) 知对任意 $x\in S_1$ 有 $(Ax,x)\in\mathbb{R}$, 因此可设

$$\sup_{x\in S_1}|(Ax,x)|=\sup_{x\in S_1}(Ax,x).$$

令

根据上确界的定义知存在 $\{x_n\}_{n\in\mathbb{N}}\subset S_1$ 使得 $f(x_n)\to\lambda(n\to\infty)$. 又因为 Hilbert 自反, 而 Eberlein-Smulian 定理3.6表明自反空间的单位闭球弱自列紧, 故 $\{x_n\}_{n\in\mathbb{N}}$ 自身存在弱收敛子列, 不妨依旧记该子列为 $\{x_n\}_{n\in\mathbb{N}}$, 并设 $x_0\in H$ 满足 $\|x_0\|_H\leq1,x_n\rightharpoonup x_0(n\to\infty)$. 因为 A 是紧算子, 故由命题6.5知 A 全连续, 进而 $n\to\infty$ 时有

$$Ax_n\to Ax_0.$$

于是 $n\to\infty$ 时有

$$\begin{aligned}|f(x_n)-(Ax_0,x_0)| &= |(Ax_n,x_n)-(Ax_0,x_0)|\\ &\le |(Ax_n-Ax_0,x_n)|+|(Ax_0,x_n-x_0)|\\ &\le \|Ax_n-Ax_0\|_H\|x_n\|_H+|(Ax_0,x_n-x_0)|\\ &\le \|Ax_0-Ax_0\|_H+|(Ax_0,x_n-x_0)|\to 0,\end{aligned}$$

因此 $\lambda=(Ax_0,x_0)$.

下面说明 $\|x_0\|_H=1$. 用反证法, 若 $\|x_0\|_H<1$, 注意到本身根据基本性质6.9(v) 有

$$\lambda=\sup_{x\in S_1}(Ax,x)=\sup_{x\in S_1}|(Ax,x)|=\|A\|_{L(H)}>0,$$

另一方面根据 Cauchy-Schwarz 不等式知

$$\begin{aligned}\lambda=(Ax_0,x_0)&\le\|A\|_{L(H)}\|x_0\|_H^2\\ &=\sup_{x\in S_1}(Ax,x)\|x_0\|_H^2<\sup_{x\in S_1}(Ax,x),\end{aligned}$$

这与 $\sup_{x\in S_1}(Ax,x)=\lambda$ 矛盾! 故 $\|x_0\|_H=1$, 且

$$(Ax_0,x_0)=\sup_{x\in S_1}(Ax,x).$$

下面说明 $Ax_0=\lambda x_0$. 为此任取 $y\in H$ 与 $y\in\mathbb{R}$ 满足 $|t|$ 充分小, 且

$$\|x_0+ty\|_H>0,$$

对 y,t 定义:

$$\varphi_y(t):=\frac{(A(x_0+ty),x_0+ty)}{(x_0+ty,x_0+ty)}=\left(A\left(\frac{x_0+ty}{\|x_0+ty\|_H}\right),\frac{x_0+ty}{\|x_0+ty\|_H}\right).$$

因为 $(x_0+ty)/\|x_0+ty\|_H\in S_1$, 故 $\varphi_y(t)$ 在 $t=0$ 时达到极大值, 这说明 $\varphi_y'(0)=0$. 又因为

$$\varphi_y'(0)=\lim_{t\to0}\frac{\varphi_y(t)-\varphi_y(0)}{t}=\lim_{t\to0}\frac{\varphi_y(t)-\lambda}{t}=\lim_{t\to0}\frac{(A(x_0+ty),x_0+ty)-\lambda\|x_0+ty\|_H^2}{t\|x_0+ty\|_H^2}\tag{6.43}$$

而对 $\forall t\in\mathbb{R}$ 与 $\forall y\in H$ 有

$$\begin{aligned}(A(x_0+ty),x_0+ty)-\lambda\|x_0+ty\|_H^2&=(Ax_0,x_0)+t(Ax_0,y)+t(Ay,x_0)+t^2(Ay,y)-\lambda[(x_0,x_0)+t(x_0,y)+t(y,x_0)+t^2(y,y)]\\ &=t[(Ax_0,x_0)+(Ay,x_0)+t(Ay,y)-\lambda[(x_0,y)+(y,x_0)+t(y,y)]]\end{aligned}$$

将上式代入(6.43)式知

$$\begin{aligned}\varphi_y'(0)&=\lim_{t\to0}\frac{(Ax_0,y)+(Ay,x_0)+t(Ay,y)-\lambda[(x_0,y)+(y,x_0)+t(y,y)]}{\|x_0+ty\|_H^2}\\ &=(Ax_0,y)+(Ay,x_0)-\lambda(x_0,y)-\lambda(y,x_0),\end{aligned}$$

由 A 对称可知对 $\forall y\in H$ 有

$$(Ax_0,y)+(Ay,x_0)-\lambda(x_0,y)-\lambda(y,x_0)=(Ax_0-\lambda x_0,y)+(y,Ax_0-\lambda x_0)=2\,\mathrm{Re}(Ax_0-\lambda x_0,y)=0.$$

将 y 换成 iy 即得 $\mathrm{Im}(Ax_0-\lambda x_0,y)=0$, 因此 $(Ax_0,\lambda x_0,y)=0$, 又由 y 的任意性可知 $Ax_0=\lambda x_0$, 此即欲证. □

注 本注给出 $Ax_0=\lambda x_0$ 的另一种证法. 根据 $\lambda=(Ax_0,x_0)$ 与 $\|x_0\|_H=1$ 可知

$$\begin{aligned}\|Ax_0-\lambda x_0\|_H^2&=(Ax_0-\lambda x_0,Ax_0-\lambda x_0)\\ &=(Ax_0,Ax_0)-\lambda^2=\|Ax_0\|_H^2-\lambda^2\\ &\le\|A\|_{L(H)}^2-\lambda^2=0,\end{aligned}$$

因此 $Ax_0=\lambda x_0$.

注 现在整理一下关于对称紧算子我们已经知道的事: 若 A 是 Hilbert 空间 H 上的对称紧算子, 则:

(i) A 必有实特征值, 因为定理6.12表明 (Ax_0,x_0) 是 A 的特征值, 而基本性质6.9(i) 表明 $(Ax_0,x_0)\in\mathbb{R}$. 另若 A 非

零, 则它必有非零实特征值.

(ii) 若 x_0 是定理6.12给出的, 且 $\widetilde{\lambda}$ 是 A 的任意实特征值, 则 $|\widetilde{\lambda}| \le |(Ax_0, x_0)|$.

(iii) 根据 (ii), A 的紧性与谱性质6.10可知 A 的算子范数是其特征值绝对值的最大值, 也是其所有谱的绝对值的最大值, 即 $\|A\|_{L(H)} = r_\sigma(A)$. 事实上, 因为 (Ax_0, x_0) 是 A 的一个特征值, 故

$$|(Ax_0, x_0)| \le \sup\{|\lambda| : \lambda \in \sigma_p(A)\}.$$

根据谱性质6.10(ii) 可知

$$\sup\{|\lambda| : \lambda \in \sigma_p(A)\} = \sup\{|\lambda| : \lambda \in \sigma(A)\} = r_\sigma(A).$$

根据谱性质6.10的注, 可知集合 $\{|\lambda| : \lambda \in \sigma(A)\}$ 能取到最大值, 于是

$$r_\sigma(A) = \max\{|\lambda| : \lambda \in \sigma_p(A)\}.$$

但谱半径本身满足

$$r_\sigma(A) \le \|A\|_{L(H)},$$

且由基本性质6.9(v) 知

$$\|A\|_{L(H)} = (Ax_0, x_0),$$

综合上述诸式可知

$$r_\sigma(A) = \max\{|\lambda| : \lambda \in \sigma_p(A)\} = |(Ax_0, x_0)| = \|A\|_{L(H)}.$$

□

对于 H 上的紧算子 A, 我们已经根据谱性质6.10知道

$$\sigma(A)\backslash\{0\} = \sigma_p(A)\backslash\{0\} = \{\lambda_1, \cdots, \lambda_m, \cdots\}$$

若 $\{\lambda_1, \cdots, \lambda_m, \cdots\}$ 中有无穷多个 $\lambda_i (i = 1, 2, \cdots)$ 不相等, 则 $\lambda_n \to 0 (n \to \infty)$. 当 A 还是自伴算子 (即对称算子) 时, 基本性质6.9(ii) 表明 $\{\lambda_1, \cdots, \lambda_m, \cdots\} \subset \mathbb{R}$. 下面的 Hilbert-Schmidt 定理表明实际上可以通过这列实数完全刻画 A.

定理 6.13 (Hilbert-Schmidt)

若 A 是 Hilbert 空间 H 上的对称紧算子, 则存在可数个非零且以 0 为聚点的实数 $\{\lambda_i\}_i$(相同数依照重数计算), 这列实数是算子 A 的本征值, 且它们对应了一组规范正交集 $\{e_i\}_i$, 它满足对 $\forall x \in H$ 有

$$x = \sum_i (x, e_i) e_i + \sum_\alpha (x, e_\alpha^{(0)}) e_\alpha^{(0)} \tag{6.44}$$

且

$$Ax = \sum_i \lambda_i (x, e_i) e_i, \tag{6.45}$$

其中 $\{e_\alpha^{(0)}\}_\alpha$ 是 $N(A)$ 的规范正交基 (若 0 不是 A 的本征值, 则 $\{e_\alpha^{(0)}\}_\alpha = \varnothing$), 且 $\{e_i\}_i \cup \{e_\alpha^{(0)}\}_\alpha$ 是 H 的规范正交基.

♡

注 在上述定理中, 规范正交基 $\{e_\alpha^{(0)}\}_\alpha$ 未必可数, 但对取定的 x 而言, 只有可数个 $e_\alpha^{(0)}$ 能使得 $(x, e_\alpha^{(0)}) \ne 0$.

为证明定理6.13, 我们需要下述引理.

引理 6.10

设 X 是 Banach 空间, $A \in \mathfrak{C}(X)$, 若

$$\sigma_p(A)\backslash\{0\} \ne \varnothing,$$

则对 $\forall \lambda \in \sigma_p(A)\backslash\{0\}$ 均有 $\dim N(\lambda I - A) < \infty$.

♡

证明 既然 $\lambda \ne 0$, 故 $\lambda^{-1} A \in \mathfrak{C}(X)$, 于是根据二择一定理6.9即知 $N(I - \lambda^{-1} A) = N(\lambda I - A)$ 是有限维空间. □

下面证明定理6.13.

证明 任取 $\lambda\in\sigma_p(A)\backslash\{0\}$, 根据引理6.10知

$$m(\lambda):=\dim N(\lambda I-A)<\infty,$$

回忆非退化内积空间必存在完备正交集, 而 $N(\lambda I-A)$ 显然非退化, 且作为内积空间的闭线性子空间依旧是内积空间, 故 $N(\lambda I-A)$ 中存在完备正交集, 再由 $N(\lambda I-A)$ 有限维知该正交集必定是有限集, 于是可设 $N(\lambda I-A)$ 存在规范正交基 $\{e_i^{(\lambda)}\}_{i=1}^{m(\lambda)}$. 因为 A 是紧算子, 根据谱性质6.10可知 $\sigma_p(A)\backslash\{0\}$ 至多由以 0 为聚点的一列复数构成, 又因为 A 对称, 所以该复数列实际上是实数列, 记之为 $\{\lambda_i\}_i$. 现在对每个 $\lambda\in\sigma_p(A)\backslash\{0\}$ 重复前述操作, 得到向量集

$$\bigcup_{\lambda\in\sigma_p(A)\backslash\{0\}}\{e_i\}_{i=1}^{m(\lambda)}$$

由 $\sigma_p(A)\backslash\{0\}$ 的至多可列性知上述集合至多可列, 因此可以把它记作 $\{\widetilde{e_i}\}_i$. 根据基本性质6.9(iv), 不同的 λ 诱导出的 $e_i^{(\lambda)}$ 彼此正交, 故 $\{\widetilde{e_i}\}_i$ 是规范正交集.

现设 $0\in\sigma_p(A)$, 则 $N(A)$ 是 H 的非退化闭子空间, 从而 $N(A)$ 存在规范正交基 $\{e_\alpha^{(0)}\}_\alpha$. 注意这样得到的 $\{e_\alpha^{(0)}\}_\alpha$ 未必可数. 现在令

$$\{e_\beta\}_\beta:=\begin{cases}\{\widetilde{e_i}\}_i, & 0\notin\sigma_p(A)\\ \{\widetilde{e_i}\}_i\cup\{e_\alpha^{(0)}\}_\alpha, & 0\in\sigma_p(A),\end{cases}$$

记 $M=\operatorname{span}\{e_\beta\}$(即 $\{e_\beta\}_\beta$ 中任意有限个元素的线性组合构成的集合), 下面验证在 M 上成立(6.44),(6.45)式. 事实上, 任取 $x\in M$, 根据 M 的构造知存在 $e_1^x,\cdots,e_m^x\in\{e_i\}_i$ 使得

$$x=\sum_{i=1}^m(x,e_i^x)e_i^x.$$

因为 e_i^x 要么是 λ_i 诱导出来的, 要么是 0 诱导出来的, 故每个 e_i^x 都能对应某个特征值. 设 e_i^x 对应的特征值为 λ_i^x, 则

$$Ax=\sum_{i=1}^m(x,e_i^x)Ae_i^x=\sum_{i=1}^m\lambda_i^x(x,e_i^x)e_i^x,$$

根据 e_i 的取法, 若 $e_i\notin\{e_1^x,\cdots,e_m^x\}$, 就必有 $(x,e_i)=0$, 因此(6.44),(6.45)两式在 M 上成立.

下面再说明 $M=H$, 为此考虑说明 $\overline{M}^\perp=\{0_H\}$. 用反证法, 若 $\overline{M}^\perp\neq\{0_H\}$, 则 $\overline{M}^\perp$ 是 H 的闭子空间, 下面说明

$$A(\overline{M}^\perp)\subset\overline{M}^\perp.$$

为说明上式, 首先说明 $A(\overline{M})\subset\overline{M}$. 任取 $x\in\overline{M}$, 知存在 $\{x_n\}_{n\in\mathbb{N}}\subset M$ 满足 $x_n\to x(n\to\infty)$. 由 $A\in L(H)$ 知

$$Ax=A\left(\lim_{n\to\infty}x_n\right)=\lim_{n\to\infty}Ax_n.$$

又因为对任意 $n\in\mathbb{N}$ 而言, x_n 都是 A 的特征向量的有限线性组合, 于是 Ax_n 依旧是这些特征向量的有限线性组合, 从而 $Ax_n\in M$, 故 $Ax\in\overline{M}$, 此即 $A(\overline{M})\subset\overline{M}$. 现取 $\xi\in\overline{M}^\perp,\eta\in\overline{M}$, 则由 A 的对称性与 $A\eta\in\overline{M}$ 知

$$(A\xi,\eta)=(\xi,A\eta)=0,$$

因此 $A\xi\in\overline{M}^\perp$, 故 $\overline{M}^\perp$ 是 A 的闭不变子空间. 现在令 $\widetilde{A}:=A|_{\overline{M}^\perp}$, 我们说明 $\widetilde{A}$ 无特征值. 这是因为若存在 λ 与 $0_H\neq\xi\in\overline{M}^\perp$ 使得 $\widetilde{A}\xi=\lambda\xi$, 则 $A\xi=\lambda\xi$ 且 $\lambda\in\mathbb{R}$, 于是 ξ 作为 A 的特征向量理应在 $\overline{M}$ 中[10], 矛盾. 因此 $\widetilde{A}$ 无特征值. 又因为 $A\in\mathfrak{C}(H)$, 根据紧算子的基本性质6.4(iv) 知 $\widetilde{A}$ 是 $\overline{M}^\perp$ 上的紧算子, 再由 A 的对称性与对称算子的基本性质6.9(iii) 知 $\widetilde{A}$ 是 $\overline{M}^\perp$ 上的对称算子, 因此定理6.12表明存在 $x_0\in\overline{M}^\perp$ 满足 $\|x_0\|_H=1$, 且

$$\widetilde{A}x_0=(\widetilde{A}x_0,x_0)x_0$$

这说明 $(\widetilde{A}x_0,x_0)$ 是 $\widetilde{A}$ 的特征值, 矛盾! 故只能有 $\overline{M}^\perp=\{0_H\}$, 于是 $M^\perp=\overline{M}^\perp=\{0_H\}$, 又因为

$$(\operatorname{span}\{e_i\})^\perp=M^\perp=\{0_H\},$$

故 $\{e_i\}_i$ 是 H 中的完备集, 因此 $\{e_i\}_i$ 是 H 的一组规范正交基, 于是(6.44)式在 H 上成立. 又因为 A 连续, 故(6.45)式

[10] 这里必须是 $\xi\in\overline{M}$ 而非 $\xi\in M$, 这是因为当 $\lambda=0$ 为 A 的特征值时, $\dim N(A)$ 是可能到达 ∞ 的, 此时 ξ 是 $\{e_i^{(0)}\}_i$ 中可数个元素的集合, 从而不再有 $\xi\in M$.

也在 H 上成立. □

注 回忆秩 1 算子引入的定义, 如果把 A 的非零特征值重排为

$$|\lambda_1| \geq \cdots \geq |\lambda_m| \geq \cdots$$

则可将 A 记为

$$A = \sum_{j\geq 1} \lambda_j e_j \otimes e_j,$$

其中对 $\forall j, \forall x \in H$ 有

$$(e_j \otimes e_j)(x) = (x, e_j)e_j.$$

我们下面说明这一等式是在 $L(H)$ 的强拓扑下成立的. 只需考虑计重数时 A 有无穷个非零特征值的情况即可, 往证

$$\left\| A - \sum_{j=1}^{n} \lambda_j e_j \otimes e_j \right\|_{L(H)} \leq |\lambda_{n+1}| \to 0. \tag{6.46}$$

为此任取 $x \in H$ 满足 $\|x\|_H = 1$, 由(6.44)式与引理6.10知

$$\|x\|_H = \left[\sum_{i=1}^{\infty} |(x, e_i)|^2 + \sum_j |(x, e_j^{(0)})|^2 \right]^{\frac{1}{2}} \geq \left[\sum_{i=1}^{\infty} |(x, e_i)|^2 \right]^{\frac{1}{2}}.$$

于是

$$\begin{aligned}
\left\| \left(A - \sum_{i=1}^{n} \lambda_i e_i \otimes e_i \right)(x) \right\|_H &= \left\| Ax - \sum_{i=1}^{n} \lambda_i (x, e_i) e_i \right\|_H \\
&= \left\| \sum_{i=n+1}^{\infty} \lambda_i (x, e_i) e_i \right\|_H \\
&= \left[\sum_{i=n+1}^{\infty} \lambda_i^2 |(x, e_i)|^2 \right]^{\frac{1}{2}} \\
&\leq |\lambda_{n+1}| \left[\sum_{i=n+1}^{\infty} |(x, e_i)|^2 \right]^{\frac{1}{2}} \\
&\leq |\lambda_{n+1}| \|x\|_H,
\end{aligned}$$

这便得到了(6.46)式. □

注 本注说明 A 的特征值具有极值性质. 若 A 的非零特征值满足

$$|\lambda_1| \geq \cdots \geq |\lambda_n| \geq \cdots$$

则对任意 $n \in \mathbb{N}$ 有

$$|\lambda_n| = \sup\{|(Ax, x)| : x \perp \operatorname{span}\{e_1, \cdots, e_{n-1}\}, \|x\|_H = 1\}, \tag{6.47}$$

其中 $e_1, \cdots, e_{n-1}$ 是分别对应 $\lambda_1, \cdots, \lambda_{n-1}$ 的本征元.

事实上, 记(6.47)右式为 μ_n, 取 $x = e_n$, 则

$$x \perp \operatorname{span}\{e_1, \cdots, e_[n-1]\},$$

且 $\|x\|_H = 1$, 于是

$$|\lambda_n| = |(Ae_n, e_n)| \leq \mu_n.$$

反之, 对 $\forall x \perp \operatorname{span}\{e_1, \cdots, e_{n-1}\}$ 且 $\|x\|_H = 1$, 由 Parseval 等式与定理6.13知

$$\|x\|_H = \left[\sum_{i\geq 1} |(x, e_i)|^2 + \sum_j |(x, e_j^{(0)})|^2 \right]^{\frac{1}{2}} \geq \left[\sum_{i\geq 1} |(x, e_i)|^2 \right]^{\frac{1}{2}},$$

且

$$|(Ax,x)| = \left|\left(\sum_{i\geq 1}\lambda_i(x,e_i)e_i, x\right)\right| = \left|\sum_{i\geq n}\lambda_i(x,e_i)(e_i,x)\right| \leq |\lambda_n|\|x\|_H^2 = |\lambda_n|$$

于是 $\mu_n \leq |\lambda_n|$, 因此 $|\lambda_n| = \mu_n$, 此即(6.47)式. □

如果 A 的特征值能进一步排成

$$\lambda_1^+ \geq \lambda_2^+ \geq \cdots > 0,$$
$$\lambda_1^- \leq \lambda_2^- \leq \cdots < 0,$$

其中除 0 以外允许依重数重复, 则有下述极大极小刻画:

定理 6.14 (对称紧算子特征值的极大极小刻画) 设 A 是对称紧算子, 其非零特征值为 $\{\lambda_i^+\}_{i\geq 1} \cup \{\lambda_i^-\}_{i\geq 1}$, 则对 $\forall n \in \mathbb{N}$ 有

$$\lambda_n^+ = \inf_{E_{n-1}} \sup_{\substack{x\in E_{n-1}^\perp \\ x\neq 0_H}} \frac{(Ax,x)}{(x,x)}, \tag{6.48}$$

且

$$\lambda_n^- = \sup_{E_{n-1}} \sup_{\substack{x\in E_{n-1}^\perp \\ x\neq 0_H}} \frac{(Ax,x)}{(x,x)}, \tag{6.49}$$

其中 E_{n-1} 是 H 的任意 $n-1$ 维闭线性子空间. ♡

证明 我们只给出(6.48)式的证明, 在(6.48)式中用 $-A$ 代替 A 即得(6.49)式. 现设 e_i^+, e_i^- 分别为 λ_i^+, λ_i^- 对应的单位特征向量, 则由 Hilbert-Schmidt 定理6.13知

$$\{e_i^+\}_{i\geq 1} \cup \{e_i^-\}_{i\geq 1} \cup \{e_j^{(0)}\}_j$$

构成 H 的一组规范正交基, 其中 $\{e_j^{(0)}\}_j$ 的设置与 Hilbert-Schmidt 定理相同. 若 $x \in H$, 则根据 Hilbert-Schmidt 定理6.13有

$$x = \sum_{i\geq 1}(a_i^+e_i^+ + a_i^-e_i^-) + \sum_j a_je_j^{(0)},$$

于是

$$\frac{(Ax,x)}{(x,x)} = \frac{\sum\limits_{i\geq 1}\lambda_i^+|a_i^+|^2 + \sum\limits_{i\geq 1}\lambda_i^-|a_i^-|^2}{\sum\limits_{i\geq 1}|a_i^+|^2 + \sum\limits_{i\geq 1}|a_i^-|^2 + \sum\limits_j |a_j|^2}, \tag{6.50}$$

对 $\forall n \in \mathbb{N}$, 记(6.48)右式为 μ_n, 下面说明 $\lambda_n^+ \leq \mu_n$. 任取 E_{n-1} 为 H 的 $n-1$ 维闭线性子空间, 则 E_{n-1} 有基 $\{\widetilde{e}_1, \cdots, \widetilde{e}_{n-1}\}$, 而 $\{e_1^+, \cdots, e_n^+\}$ 是 n 个线性无关的向量, 故根据鸽笼原理必定存在

$$0_H \neq x_n \in \operatorname{span}\{e_1^+, \cdots, e_n^+\}$$

满足 $x_n \perp E_{n-1}$, 设 $x_n = \sum_{i=1}^n y_ie_i^+$, 代入(6.50)式有

$$\sup_{\substack{x\in E_{n-1}^\perp \\ x\neq 0_H}} \frac{(Ax,x)}{(x,x)} \geq \frac{(Ax_n,x_n)}{(x_n,x_n)} = \frac{\sum\limits_{i=1}^n \lambda_i^+|y_i|^2}{\sum\limits_{i=1}^n |y_i|^2} \geq \lambda_n^+.$$

上左式是只关于 E_{n-1} 的, 因此上式两端对 E_{n-1} 取下确界即得 $\lambda_n^+ \leq \mu_n$.

下面说明 $\lambda_n^+ \geq \mu_n$. 为此令

$$E_{n-1} := \operatorname{span}\{e_1^+, \cdots, e_{n-1}^+\},$$

则对任意 $x \in H$, 若 $x \in E_{n-1}^\perp$ 且 $x \neq 0_H$, 则由 Hilbert-Schmidt 定理知

$$x = \sum_{i\geq n}a_i^+e_i^+ + \sum_{i\geq 1}a_i^-e_i^- + \sum_j a_je_j^{(0)},$$

于是由(6.50)式进一步知

$$\frac{(Ax,x)}{(x,x)}=\frac{\sum\limits_{i\geq n}\lambda_i^+|a_i^+|^2+\sum\limits_{i\geq 1}\lambda_i^-|a_i^-|^2}{\sum\limits_{i\geq n}|a_i^+|^2+\sum\limits_{i\geq 1}|a_i^-|^2+\sum\limits_{j}|a_j|^2}\leq\frac{\sum\limits_{i\leq n}\lambda_i^+|a_i^+|^2}{\sum\limits_{i\geq n}|a_i^+|^2}\leq\lambda_n^+\frac{\sum\limits_{i\geq n}|a_i^+|^2}{\sum\limits_{i\geq n}|a_i^+|^2}=\lambda_n^+.$$

在上式两端对 x 取上确界即得 $\lambda_n^+\geq\mu_n$, 因此 $\lambda_n^+=\mu_n$, 此即(6.48)式. □

推论 6.3

若 Hilbert 空间 H 上的两个对称紧算子 A,B 满足 $A\leq B$(即对 $\forall x\in H$ 均有 $(Ax,x)\leq(Bx,x)$), 则对 $\forall j\in\mathbb{N}$ 有

$$0<\lambda_j^+(A)\leq\lambda_j^+(B),\quad \lambda_j^-(A)\leq\lambda_j^-(B)<0.$$

♡

第七章 Hille-Yosida 定理

本章选自 [**HB**], 统一用 H 表示 Hilbert 空间.

7.1 极大单调算子的定义与初等性质

定义 7.1 (单调算子 (monotone operator), 极大单调算子 (maximal monotone operator))

若无界线性算子 $A : D(A) \subset H \to H$ 满足

$$(Av, v) \geq 0, \quad \forall v \in D(A)$$

就称 A 是单调算子[a]. 若另外还有 $R(I + A) = H$, 即

$$\forall f \in H \exists u \in D(A)\ s.t.\ u + Au = f,$$

就称 A 是极大单调算子.

[a]有的作者会把这样的 A 称为增强算子或把 $-A$ 称为耗散算子.

命题 7.1 (极大单调算子的性质)

设 A 是极大单调算子, 则

(i) $D(A)$ 在 H 中稠密,

(ii) A 是闭算子,

(iii) 对 $\forall \lambda > 0$ 而言, $I + \lambda A$ 是 $D(A) \to H$ 的双射, $(I + \lambda A)^{-1}$ 是有界算子, 且 $\|(I + \lambda A)^{-1}\|_{L(H)} \leq 1$.

证明 (i) 我们考虑用推论2.4证明 $D(A)$ 的稠密性, 为此设 $f \in H$ 满足 $\forall v \in D(A)((f, v) = 0)$, 往证 $f = 0_H$. 根据极大单调算子的定义知对这个选定的 f 存在 $v_0 \in D(A)$ 使得

$$v_0 + Av_0 = f.$$

两边与 v_0 作内积可得

$$0 = (f, v_0) = \|v_0\|_H^2 + (Av_0, v_0) \geq \|v_0\|_H^2$$

因此只能有 $v_0 = 0_H$, 于是 $f = v_0 + Av_0 = 0_H$.

(ii) 从闭算子的定义出发, 设 $\{u_n\} \subset D(A)$ 满足 $u_n \to u, Au_n \to f$, 往证 $u \in D(A), f = Au$. 现在已知

$$u_n + Au_n = (I + A)u_n \to u + f,$$

如果现在能说明 $(I + A)^{-1} \in L(H)$, 就有

$$u_n = (I + A)^{-1}(I + A)u_n \to (I + A)^{-1}(u + f) = u,$$

于是一方面 $u \in D(I + A) = D(A)$, 另一方面通过 $u + f = Au + u$ 也能说明 $f = Au$.

下面说明确有 $(I+A)^{-1} \in L(H)$. 因为 $R(I+A) = H$, 故 $(I+A)^{-1}$ 的存在性不成问题, 但我们还需要验证 $(I+A)^{-1}$ 是单值映射, 亦即对任意 $f \in H$ 而言, 均存在唯一的 $u \in D(A)$ 使得 $u + Au = f$. 现设 $\tilde{u}$ 同样满足 $\tilde{u} + A\tilde{u} = f$, 则

$$u - \tilde{u} + A(u - \tilde{u}) = 0.$$

上式两端与 $u - \tilde{u}$ 作内积可得

$$\|u - \tilde{u}\|_H^2 + (A(u - \tilde{u}), u - \tilde{u}) = 0$$

但上左式的两项都是非负的, 故只能有

$$\|u - \tilde{u}\|_H^2 = (A(u - \tilde{u}), u - \tilde{u}) = 0$$

因此 $u=\widetilde{u}$, 于是 $(I+A)^{-1}$ 确为映射, 显见 $(I+A)^{-1}$ 是线性的. 接下来再说明 $(I+A)^{-1}$ 的有界性. 因为

$$\|u\|_H^2 \le \|u\|_H^2 + (Au,u) = (f,u) \le \|f\|_H\|u\|_H.$$

因此 $\|u\|_H = \|(I+A)^{-1}f\|_H \le \|f\|_H$, 于是 $(I+A)^{-1}\in L(H)$, (ii) 至此即证.

(iii) (ii) 中已经说明了结论在 $\lambda=1$ 时成立, 因此我们只需说明由 A 是极大单调算子能推知对 $\forall\lambda>0$ 而言 λA 都是极大单调算子即可, 后者实际上就是求解方程

$$u+\lambda Au = f,\quad \lambda>0.$$

现在把方程重写为

$$u+Au = \frac{1}{\lambda}f + \left(1-\frac{1}{\lambda}\right)u,$$

进一步写成能应用压缩映射原理的形式:

$$u = (I+A)^{-1}\left(\frac{1}{\lambda}f + \left(1-\frac{1}{\lambda}\right)u\right),$$

对于算子

$$T: u\mapsto (I+A)^{-1}\left(\frac{1}{\lambda}f + \left(1-\frac{1}{\lambda}\right)u\right),$$

知

$$\|Tu_1 - Tu_2\|_H = \left\|(I+A)^{-1}\left(1-\frac{1}{\lambda}\right)(u_1-u_2)\right\|_H.$$

现若 $|1-\frac{1}{\lambda}|\le\frac{1}{2}$, 即 $\frac{2}{3}\le\lambda\le 2$, 则利用压缩映射原理即可说明方程 $u+\lambda Au=f$ 确实有解.

为了将结论推广到 $\mathbb{R}^+$ 上, 我们需要对更一般的情形进行说明. 假设现在已知 $R(I+\lambda_0 A)=H$(其中 $\lambda_0>0$), 我们考虑当 λ 在怎样的范围内时依旧有 $R(I+\lambda A)=H$. 通过与前文相同的方法, 不难发现要想 $R(I+\lambda A)=H$, 只需 $|1-\frac{\lambda_0}{\lambda}|\le\frac{1}{2}$, 亦即 $\frac{2}{3}\lambda_0\le\lambda\le 2\lambda_0$. 现在分别取 $\lambda_0=\frac{2}{3}$ 与 $\lambda=2$, 可知 $R(I+\lambda A)=H$ 在 $\lambda\in[(\frac{2}{3})^2, 2^2]$ 时也成立. 重复该操作即知对任意 $n\in\mathbb{N}$ 而言, $R(I+\lambda A)=H$ 都在 $\lambda\in[(\frac{2}{3})^n, 2^n]$ 时成立, 因此结论在 $\lambda\in(0,\infty)$ 时成立. □

注 根据前述证明, 只要 A 是极大单调算子, 那么对任意 $\lambda>0$ 而言 λA 就也是极大单调算子. 但是当 A,B 都是极大单调算子时, 定义在 $D(A)\cap D(B)$ 上的算子 $A+B$ 未必是极大单调算子.

定义 7.2 (Yosida 逼近 (Yosida approximation))

设 A 是极大单调算子, 定义

$$J_\lambda = (I+\lambda A)^{-1},\ A_\lambda = \frac{1}{\lambda}(I-J_\lambda),$$

称 A_λ 为 A 的 Yosida 逼近 (或 Yosida 正则化).

注 根据命题7.1的证明显见 $\|J_\lambda\|_{L(H)}\le 1$.

命题 7.2 (Yosida 逼近的性质)

设 A 是极大单调算子, 则:

(i) 对 $\forall v\in H$ 与 $\forall\lambda>0$ 均有 $A_\lambda v = A(J_\lambda v)$.

(ii) 对 $\forall v\in D(A)$ 与 $\forall\lambda>0$ 均有 $A_\lambda v = J_\lambda(Av)$.

(iii) 对 $\forall v\in D(A)$ 与 $\forall\lambda>0$ 均有 $\|A_\lambda v\|_H\le\|Av\|_H$.

(iv) 对 $\forall v\in H$ 有 $J_\lambda v\to v(\lambda\to 0)$.

(v) 对 $\forall v\in D(A)$ 有 $A_\lambda v\to Av(\lambda\to 0)$.

(vi) 对 $\forall v\in H$ 与 $\forall\lambda>0$ 有 $(A_\lambda v, v)\ge 0$.

(vii) 对 $\forall v\in H$ 与 $\forall\lambda>0$ 有 $\|A_\lambda v\|_H\le(1/\lambda)\|v\|_H$.

证明 (i) 欲证结论即

$$\frac{1}{\lambda}(v-J_\lambda v) = A(J_\lambda v)\Rightarrow v = (J_\lambda v)+\lambda A(J_\lambda v),$$

代入 J_λ 的定义立得结论.

(ii) 由 (i) 知

$$A_\lambda v + A(v - J_\lambda v) = Av,$$

此即

$$A_\lambda v + \lambda A(A_\lambda v) = Av,$$

这表明 $A_\lambda v = (I_\lambda A)^{-1} Av = J_\lambda(Av)$.

(iii) 由 (ii) 与 $\|J_\lambda\|_{L(H)} \le 1$ 立得结论.

(iv) 首先设 $v \in D(A)$, 则由 (ii),(iii) 知

$$\|v - J_\lambda v\|_H = \lambda\|A_\lambda v\|_H \le \lambda\|Av\|_H$$

因此 $J_\lambda v \to v(\lambda \to 0)$.

现在设 $v \in H$, 根据 $D(A)$ 在 H 中的稠密性 (命题7.1(i)) 知对任意 $\varepsilon > 0$ 总存在 $v_1 \in D(A)$ 使得 $|v - v_1| \le \varepsilon$, 于是

$$\begin{aligned}\|J_\lambda v - v\|_H &\le \|J_\lambda v - J_\lambda v_1\|_H + \|J_\lambda v_1 - v_1\|_H + \|v_1 - v\|_H \\ &\le 2\|v_1 - v\|_H + \|J_\lambda v_1 - v_1\|_H \le 2\varepsilon + \|J_\lambda v_1 - v_1\|_H.\end{aligned}$$

结合 $v \in D(A)$ 的情况可知

$$\limsup_{\lambda\to 0} \|J_\lambda v - v\|_H \le 2\varepsilon, \quad \forall \varepsilon > 0$$

因此

$$\lim_{\lambda\to 0} \|J_\lambda v - v\|_H = 0.$$

(v) 由 (ii),(iv) 可知

$$A_\lambda v = J_\lambda(Av) \to Av, \quad \lambda \to 0.$$

(vi) 知

$$(A_\lambda v, v) = (A_\lambda v, v - J_\lambda v) + (A_\lambda v, J_\lambda v) = \lambda\|A_\lambda v\|_H^2 + (A(J_\lambda v), J_\lambda v),$$

由 A 的单调性可知

$$(A_\lambda v, v) \ge \lambda\|A_\lambda v\|_H^2. \tag{7.1}$$

(vii) 由(7.1)式与 Cahuchy-Schwarz 不等式知

$$\|A_\lambda v\|_H^2 \le \frac{1}{\lambda}(A_\lambda v, v) \le \frac{1}{\lambda}\|A_\lambda v\|_H\|v\|_H$$

两边约去 $\|A_\lambda v\|_H$ 即可. □

注 命题7.2表明 $\{A_\lambda\}_{\lambda>0}$ 实际上是在 $\lambda \to 0$ 时" 逼近" 无界算子 A 的一族有界算子. 这一逼近在 PDE 中应用得非常广. 当然, 一般情况下 $\|A_\lambda\|_L(H)$ 在 $\lambda \to 0$ 时都会爆破.

7.2 演化方程 $\frac{du}{dt}+Au=0, u(0)=u_0$ 在 $[0,+\infty)$ 上解的存在性与唯一性

我们首先回顾下述经典结果:

定理 7.1 (Cauchy,Lipschitz,Picard)

设 E 是 Banach 空间, $F: E \to E$ 是 Lipschitz 映射, 即存在常数 L 使得

$$\|Fu - Fv\|_E \le L\|u - v\|_E, \quad \forall u, v \in E.$$

则对任意给定的 $u_0\in E$, 下述初值问题总存在唯一解 $u\in C^1([0,+\infty);E)$:

$$\begin{cases}\frac{du}{dt}(t)=Fu(t),\ t\in[0,+\infty)\\ u(0)=u_0.\end{cases}\tag{7.2}$$

此时称 u_0 为初值.

证明 首先说明解的存在性. 显见要求解方程(7.2), 首先可以考虑在 $C([0,+\infty);E)$ 中求解下述积分方程:

$$u(t)=u_0+\int_0^t F(u(s))ds.\tag{7.3}$$

为了应用压缩映射原理5.4, 我们就需要找到一个合适的工作空间 X. X 作为 $C([0,+\infty),E)$ 的子空间应该非空完备可度量, 且它满足

$$u\in X,\ \Phi(u):=u_0+\int_0^t F(u(s))ds\in X,\ d_X(\Phi(u),\Phi(v))\le\alpha d_X(u,v),$$

其中 d_X 是 X 上的度量. 不妨就设 X 可赋范, 从形式上看, 如果 F 是常值映射, 那么方程的解应该是

$$u(t)=e^{tF}u_0.$$

亦即 $u_0=e^{-tF}u(t)$, 这暗示看似与 t 有关的 $e^{-tF}u(t)$ 实际上与 t 无关, 因此我们可以尝试考虑形如 $e^{-kt}u(t)$ 的量构造范数. 又因为 $X\subset C([0,+\infty);E)$, 故自然可以对 $e^{-kt}u(t)$ 应用 $C([0,+\infty);E)$ 的范数, 即 $\sup_{t\ge0}e^{-kt}\|u(t)\|_E$. 因此我们考虑空间

$$X=\left\{u\in C([0,+\infty);E);\|u\|_X:=\sup_{t\ge0}e^{-kt}\|u(t)\|_E<\infty\right\},$$

其中 $k>0$ 待定. 因为 $u\in C([0,+\infty);E)\Rightarrow e^{-k\cdot}u\in C([0,+\infty);E)$, 故 X 首先是赋范且非空的. 下面说明 X 完备, 设 $\{u_n\}_{n\in\mathbb{N}}$ 是 X 中的基本列, 根据定义知

$$\forall\varepsilon>0\exists N>0\forall m,n>N\left(\sup_{t\ge0}e^{-kt}\|u_m(t)-u_n(t)\|_E<\varepsilon\right)$$

于是对每个取定的 $t'\in[0,+\infty)$ 均有

$$e^{-kt'}\|u_m(t')-u_n(t')\|_E<\varepsilon$$

这说明对每个取定的 $t'\in[0,+\infty)$ 而言, $\{u_n(t')\}_{n\in\mathbb{N}}$ 都是 E 中的基本列, 于是对每个固定的 t' 均有

$$\|u_m(t')-u_n(t')\|_E\to\|u_m(t')-u(t')\|_E,\quad n\to\infty$$

下面说明 $\{u_n\}_{n\in\mathbb{N}}$ 在 $\|\cdot\|_X$ 的意义下收敛到 u, 事实上对任意 $m,n>N$ 有:

$$\begin{aligned}\sup_{t\ge0}e^{-kt}\|u_m(t)-u(t)\|_E&=\sup_{t\ge0}\lim_{n\to\infty}e^{-kt}\|u_m(t)-u_n(t)\|_E\\&\le\lim_{n\to\infty}\sup_{t\ge0}e^{-kt}\|u_m(t)-u_n(t)\|_E<\varepsilon.\end{aligned}$$

现在只需说明 $u\in X$, 这通过 $\|u\|_X\le\|u-u_n\|_X+\|u_n\|_X<\infty$ 立得, 因此 $(X,\|\cdot\|_X)$ 完备.

下面说明当 $u\in X$ 时, Φu 也在 X 内. 由 $u\in X$ 知可设存在 $M>0$ 使得

$$\sup_{t\ge0}e^{-kt}\|u(t)\|_E\le M<\infty,$$

另外显见 $F(0_E)=0_E$, 于是

$$\begin{aligned}\|\Phi u\|_X &\le \|u_0\|_X+\left\|\int_0^t F(u(s))ds\right\|_X\\ &\le \|u_0\|_X+\sup_{t\ge 0}e^{-kt}\int_0^t \|F(u(s))\|_E ds\\ &\le \|u_0\|_X+L\sup_{t\ge 0}e^{-kt}\int_0^t \|u(s)\|_E ds\\ &\le \|u_0\|_X+L\sup_{t\ge 0}e^{-kt}\int_0^t e^{ks}\sup_{s\ge 0}e^{-ks}\|u(s)\|_E ds\\ &\le \|u_0\|_X+ML\sup_{t\ge 0}e^{-kt}\int_0^t e^{ks}ds<\infty.\end{aligned}$$

因此 $\Phi u\in X$. 现在

$$\begin{aligned}\|\Phi u-\Phi v\|_X &= \left\|\int_0^t (F(u(s))-F(v(s)))ds\right\|_X\\ &\le \sup_{t\ge 0}e^{-kt}\int_0^t \|F(u(s))-F(v(s))\|_E ds\\ &\le L\sup_{t\ge 0}e^{-kt}\int_0^t \|u(s)-v(s)\|_E ds\\ &\le L\sup_{t\ge 0}e^{-kt}\int_0^t e^{ks}\sup_{s\ge 0}e^{-ks}\|u(s)-v(s)\|_E ds\\ &= \frac{L}{k}\sup_{t\ge 0}(1-e^{-kt})\|u-v\|_X=\frac{L}{k}\|u-v\|_X.\end{aligned}$$

取 $k>L$, 则 Φ 是 X 中的压缩映射, 根据压缩映射原理5.4即知 Φ 在 X 中存在不动点 u, 而由 $X\subset C([0,+\infty);E)$ 可知 $u\in C([0,+\infty);E)$, 因此 u 是积分方程(7.3)的解, 而 u 能表成 Bochner 积分这件事自然意味着它对 t 一阶可微, 因此 u 也是方程(7.2)的解.

下面说明解的唯一性. 设 $u,\overline{u}$ 是方程(7.2)的两个解, 设

$$\varphi(t)=\|u(t)-\overline{u}(t)\|_E.$$

由(7.3)式知

$$\varphi(t)\le L\int_0^t \varphi(s)ds,\quad \forall t\ge 0,$$

于是 Gronwall 不等式表明此时只能有 $\varphi\equiv 0$. □

前述定理在 ODE 的研究中非常有用, 但它在 PDE 的研究中用的很少, 这是因为微分算子往往不是 Lipschitz 映射. 因此我们给出求解演化方程的下述强力工具:

定理 7.2 (Hille-Yosida)

设 A 是极大单调算子, 则对任意给定的 $u_0\in D(A)$ 而言, 存在唯一函数[a]

$$u\in C^1([0,+\infty);H)\cap C([0,+\infty);D(A))$$

满足下述方程:

$$\begin{cases}\frac{du}{dt}+Au=0 & t\in[0,+\infty)\\ u(0)=u_0.\end{cases}\tag{7.4}$$

且

$$|u(t)|\le|u_0|,\ \left|\frac{du}{dt}(t)\right|=|Au(t)|\le|Au_0|, \forall t\ge 0.$$

[a]其中 $D(A)$ 赋的是图范数 $\|v\|_H+\|Av\|_H$ 或与图范数等价的 Hilbert 范数 $(\|v\|_H^2+\|Av\|_H^2)^{\frac{1}{2}}$.

♡

注 Hille-Yosida 定理的主要威力在于只要 A 是单调算子 (这在实践中很容易验证, 尤其是对微分算子), 它就可以

把”演化问题”的研究转化为”驻定问题”$u+Au=f$ 的研究.

证明 定理的证明主要分为六步.

$\boxed{\text{第一步: 唯一性.}}$ 设 $u,\overline{u}$ 是方程(7.4)的两个解, 则

$$\left(\frac{d}{dt}(u-\overline{u}),(u-\overline{u})\right)=-(A(u-\overline{u}),u-\overline{u})\le 0,$$

但另一方面有[1]:

$$\frac{1}{2}\frac{d}{dt}\|u(t)-\overline{u}(t)\|_H^2=\left(\frac{d}{dt}(u(t)-\overline{u}(t)),u(t)-\overline{u}(t)\right).$$

因此函数 $t\mapsto|u(t)-\overline{u}(t)|$ 在 $[0,+\infty)$ 上不增. 又因为 $|u(0)-\overline{u}(0)|=0$, 故

$$|u(t)-\overline{u}(t)|=0,\quad \forall t\ge 0.$$

对于存在性的证明, 我们的思路是在方程(7.4)中用 A_λ 替换 A, 对替换后的方程用定理7.1, 再用与 λ 无关的估计得到 $\lambda\to 0$ 时的极限. 现设 u_λ 时下述初值问题的解:

$$\begin{cases}\frac{du_\lambda}{dt}+A_\lambda u_\lambda=0,\ t\in[0,+\infty),\\ u_\lambda(0)=u_0\in D(A).\end{cases}\tag{7.5}$$

$\boxed{\text{第二步: 估计 } |u_\lambda(t)|,|\frac{du_\lambda}{dt}(t)|}$. 往证估计

$$|u_\lambda(t)|\le|u_0|,\quad \forall t\ge 0,\forall\lambda>0\tag{7.6}$$

$$\left|\frac{du_\lambda}{dt}(t)\right|=|A_\lambda u_\lambda(t)|\le|Au_0|,\quad \forall t\ge 0,\forall\lambda>0.\tag{7.7}$$

为此我们考虑下述引理:

引理 7.1

设 $w\in C^1([0,+\infty);H)$ 是满足下述方程的函数:

$$\frac{dw}{dt}+A_\lambda w=0,\ t\in[0,+\infty)\tag{7.8}$$

则函数 $t\mapsto|w(t)|, t\mapsto|\frac{dw}{dt}(t)|=|A_\lambda w(t)|$ 在 $[0,+\infty)$ 上不增.

♡

现在证明引理7.1. 知

$$\left(\frac{dw}{dt},w\right)+(A_\lambda w,w)=0.$$

根据 Yosida 逼近的性质7.2(vi) 知 $(A_\lambda w,w)\ge 0$, 因此只能有 $(\frac{dw}{dt},w)\le 0$, 亦即 $\frac{1}{2}\frac{d}{dt}\|w\|_H^2\le 0$, 从而 $\|w(t)\|_H$ 不增. 另一方面, 因为 A_λ 是有界线性算子, 故

$$\begin{aligned}\frac{d^2w}{dt^2}(t)&=\lim_{h\to 0}\frac{\frac{dw}{dt}(t+h)-\frac{dw}{dt}(t)}{h}\\&=-\lim_{h\to 0}\frac{A_\lambda w(t+h)-A_\lambda w(t)}{t}\\&=-A_\lambda\left(\frac{dw}{dt}(t)\right)\end{aligned}$$

因此 $w\in C^1([0,+\infty);H)\Rightarrow w\in C^2([0,+\infty);H)$, 归纳可得 $w\in C^\infty([0,+\infty);H)$. 且上式表明

$$\frac{d}{dt}\left(\frac{dw}{dt}\right)+A_\lambda\left(\frac{dw}{dt}\right)=0.$$

将上式两端与 $\frac{dw}{dt}$ 作内积, 与前文同理可得 $\|\frac{dw}{dt}\|_H$ 不增. 事实上对任意 $k\in\mathbb{N}$ 而言, 函数 $\|\frac{d^kw}{dt^k}(t)\|_H$ 都是不增的. 引理证毕.

根据上述引理与 Yosida 逼近的性质7.2(iii) 立得(7.6),(7.7)式.

$\boxed{\text{第三步: } u_\lambda(t) \text{ 的收敛性.}}$ 往证对任意 $t\ge 0$ 而言, $u_\lambda(t)$ 在 $\lambda\to 0$ 时均收敛, 记该极限为 $u(t)$. 另外这一收敛在

[1]注意当 $\varphi\in C^1([0,+\infty);H)$ 时必有 $\|\varphi\|_H^2\in C^1([0,+\infty);\mathbb{R})$, 且 $\frac{d}{dt}\|\varphi\|_H^2=2(\frac{d\varphi}{dt},\varphi)$.

任意有界区间 $[0,T]$ 上都是一致的.

对任意 $\lambda,\mu>0$, 知

$$\frac{du_\lambda}{dt}-\frac{du_\mu}{dt}+A_\lambda u_\lambda-A_\mu u_\mu=0,$$

因此

$$\frac{1}{2}\frac{d}{dt}\|u_\lambda(t)-u_\mu(t)\|_H^2+(A_\lambda u_\lambda(t)-A_\mu u_\mu(t),u_\lambda(t)-u_\mu(t))=0. \tag{7.9}$$

简便起见我们把 $u(t)$ 就写成 u, 对上左式第二项利用 Yosida 逼近的性质7.2(i) 有

$$\begin{aligned}(A_\lambda u_\lambda-A_\mu u_\mu,u_\lambda-u_\mu)&=(A_\lambda u_\lambda-A_\mu u_\mu,u_\lambda-J_\lambda u_\lambda+J_\lambda u_\lambda-J_\mu u_\mu+J_\mu u_\mu-u_\mu)\\&=(A_\lambda u_\lambda-A_\mu u_\mu,\lambda A_\lambda u_\lambda-\mu A_\mu u_\mu)+(A(J_\lambda u_\lambda-J_\mu u_\mu),J_\lambda u_\lambda-J_\mu u_\mu)\\&\geq(A_\lambda u_\lambda-A_\mu u_\mu,\lambda A_\lambda u_\lambda-\mu A_\mu u_\mu).\end{aligned} \tag{7.10}$$

于是由(7.7),(7.9),(7.10)式知

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u_\lambda-u_\mu\|^2&=-(A_\lambda u_\lambda(t)-A_\mu u_\mu(t),u_\lambda(t)-u_\mu(t))\\&\leq-(A_\lambda u_\lambda-A_\mu u_\mu,\lambda A_\lambda u_\lambda-\mu A_\mu u_\mu)\\&\leq\lambda\|A_\lambda u_\lambda\|_H^2+(\lambda+\mu)\|A_\mu u_\mu\|_H\|A_\lambda u_\lambda\|_H+\mu\|A_\mu u_\mu\|_H^2\\&\leq 2(\lambda+\mu)\|Au_0\|_H^2.\end{aligned}$$

对上述不等式两端积分有

$$\|u_\lambda(t)-u_\mu(t)\|_H^2\leq 4(\lambda+\mu)t\|Au_0\|_H^2,$$

即

$$\|u_\lambda(t)-u_\mu(t)\|_H\leq 2\sqrt{(\lambda+\mu)t}\|Au_0\|_H. \tag{7.11}$$

因此对每个取定的 $t\geq 0$ 而言, $\{u_\lambda(t)\}_{\lambda>0}$ 都是 $\lambda\to 0$ 时的 Cauchy 列, 由 H 的完备性知 $\lim_{\lambda\to 0}u_\lambda(t)$ 存在, 记之为 $u(t)$. 现在在(7.11)式两端令 $\mu\to 0$ 可得

$$\|u_\lambda(t)-u(t)\|_H\leq 2\sqrt{\lambda t}\|Au_0\|_H.$$

故在每个有界区间 $[0,T]$ 上, $u_\lambda(t)\to u(t)$ 都是关于 t 一致的, 于是根据 $C[0,T]$ 的完备性知 $u\in C([0,T];H)$ 对任意 $T>0$ 均成立, 因此 $u\in C([0,+\infty);H)$.

$\boxed{\text{第四步: } \frac{du_\lambda}{dt} \text{ 的收敛性.}}$ 设 $u_0\in D(A^2)$(即 $u_0\in D(A^2), Au_0\in D(A)$), 往证 $\lambda\to 0$ 时 $\frac{du_\lambda}{dt}(t)$ 能收敛到某极限, 且该收敛在每个有界区间 $[0,T]$ 上均一致.

设 $v_\lambda=\frac{du_\lambda}{dt}$, 则由引理7.1的证明可知 $\frac{dv_\lambda}{dt}+A_\lambda v_\lambda=0$, 于是第三步的结果表明

$$\frac{1}{2}\frac{d}{dt}\|v_\lambda-v_\mu\|_H^2\leq(\|A_\lambda v_\lambda\|_H+\|A_\mu v_\mu\|_H)(\lambda\|A_\lambda v_\lambda\|_H+\mu\|A_\mu v_\mu\|_H). \tag{7.12}$$

由引理7.1知

$$\|A_\lambda v_\lambda(t)\|_H\leq\|A_\lambda v_\lambda(0)\|_H=\|A_\lambda A_\lambda u_0\|_H. \tag{7.13}$$

类似有

$$\|A_\mu v_\mu(t)\|_H\leq\|A_\mu v_\mu(0)\|_H=\|A_\mu A_\mu u_0\|_H. \tag{7.14}$$

又因为 $Au_0\in D(A)$, 故根据 Yosida 逼近的性质7.2(i),(ii) 知

$$A_\lambda A_\lambda u_0=J_\lambda AJ_\lambda Au_0=J_\lambda J_\lambda AAu_0=J_\lambda^2A^2u_0,$$

由 $\|J_\lambda\|_{L(H)}\leq 1$ 知

$$\|A_\lambda A_\lambda u_0\|_H\leq\|A^2u_0\|_H,\quad \|A_\mu A_\mu u_0\|_H\leq\|A^2u_0\|_H. \tag{7.15}$$

结合(7.12)-(7.15)式可知

$$\frac{1}{2}\frac{d}{dt}\|v_\lambda-v_\mu\|_H^2\leq 2(\lambda+\mu)\|A^2u_0\|_H^2.$$

于是重复第三步的过程可知 $v_\lambda(t)=\frac{du_\lambda}{dt}(t)$ 在 $\lambda\to 0$ 时收敛到某极限, 且这一收敛在任意有界区间 $[0,T]$ 上关于 t 都是一致的.

第五步: 说明 u 是方程(7.4)的解. 设 $u_0\in D(A^2)$, 第三步与第四步说明了对任意 $T<\infty$ 均有

$$\begin{cases} u_\lambda(t)\to u(t)(\lambda\to 0)\text{在 } [0,T] \text{ 上一致,}\\ \lim\limits_{\lambda\to 0}\frac{du_\lambda}{dt}(t)\text{存在且在 } [0,T] \text{ 上一致.}\end{cases}$$

因此 $u\in C^1([0,+\infty);H)$, 且在 $[0,T]$ 上总有 $\frac{du_\lambda}{dt}(t)\to\frac{du}{dt}(t)(\lambda\to 0)$ 关于 t 一致成立. 现在把(7.5)式重写为

$$\frac{du_\lambda}{dt}(t)+A(J_\lambda u_\lambda(t))=0. \tag{7.16}$$

根据极大单调算子的性质7.1(ii) 知 A 的图是闭集, 因此对任意 $t\geq 0$ 而言, $u(t)$ 作为 $u_\lambda(t)$ 在 $D(A)$ 中的聚点也应在 $D(A)$ 中. 现在对每个 $t\geq 0$, 在(7.16)式中令 $\lambda\to 0$ 即得

$$\frac{du}{dt}(t)+Au(t)=0.$$

最后, 因为 $u\in C^1([0,+\infty);H)$, 故函数 $t\mapsto Au(t)=-\frac{du}{dt}(t)$ 是 $[0,+\infty)\to H$ 的连续函数, 这说明函数 $t\mapsto u(t)$ 在范数 $\|\cdot\|_H+\|A\cdot\|_H$ 下连续, 因此 $u\in C([0,+\infty);D(A))$. 至此我们便得到了方程(7.4)的一个解, 且该解满足

$$\|u(t)\|_H\leq\|u_0\|_H,\quad \left|\frac{du}{dt}(t)\right|_H=\|Au(t)\|_H\leq\|Au_0\|_H,\quad \forall t\geq 0.$$

第六步: 减弱 $u_0\in D(A^2)$ 的条件. 回到定理本身, 注意我们并没有限定 $u_0\in D(A^2)$, 而是只知道 $u_0\in D(A)$. 因此我们需要通过下述引理把条件减弱:

引理 7.2

设 $u_0\in D(A)$, 则对 $\forall\varepsilon>0$ 均存在 $\overline{u}_0\in D(A^2)$ 使得 $\|u_0-\overline{u}_0\|_H<\varepsilon, \|Au_0-A\overline{u}_0\|_H<\varepsilon$, 亦即 $D(A^2)$ 在图范数意义下在 $D(A)$ 中稠密.

现在证明引理7.2. 取 $\overline{u}_0=J_\lambda u_0$, 其中 $\lambda>0$ 待定. 知

$$\overline{u}_0\in D(A),\quad \overline{u}_0+\lambda A\overline{u}_0=u_0.$$

于是 $A\overline{u}_0\in D(A)\Rightarrow\overline{u}_0\in D(A^2)$. 另一方面, 根据 Yosida 逼近的性质7.2(v),(i),(ii) 知

$$\lim_{\lambda\to 0}\|J_\lambda u_0-u_0\|_H=0,\ \lim_{\lambda\to 0}\|J_\lambda Au_0-Au_0\|_H=0,\ J_\lambda Au_0=AJ_\lambda u_0.$$

因此只需取 $\lambda>0$ 即得欲证. 引理证毕.

回到定理7.2的证明. 现在对每个给定的 $u_0\in D(A)$, 根据引理7.2知总可以找到 $D(A^2)$ 中的序列 $\{u_{0,n}\}_{n\in\mathbb{N}}$ 满足 $u_{0,n}\to u_0, Au_{0,n}\to Au_0$. 根据第五步的结果可知对每个 $u_{0,n}$ 而言, 都存在 u_n 满足下述初值问题:

$$\begin{cases}\frac{du_n}{dt}+Au_n=0,\ t\in[0,+\infty)\\ u_n(0)=u_{0,n}.\end{cases} \tag{7.17}$$

于是对任意 $t\geq 0$ 有

$$\|u_n(t)-u_m(t)\|_H\leq\|u_{0,n}-u_{0,m}\|_H\to 0,\quad m,n\to\infty$$
$$\left\|\frac{du_n}{dt}(t)-\frac{du_m}{dt}(t)\right\|_H\leq\|Au_{0,n}-Au_{0,m}\|_H\to 0,\quad m,n\to\infty$$

这说明存在 $u\in C^1([0,+\infty);H)$ 使得

$$u_n(t)\to u(t)\text{ 在 }[0,+\infty)\text{ 上一致,}$$
$$\frac{du_n}{dt}(t)\to\frac{du}{dt}(t)\text{ 在 }[0,+\infty)\text{ 上一致.}$$

因为 A 是闭算子, 故 $u\in D(A)$, 因此在方程(7.17)中令 $n\to\infty$ 即知 u 满足方程(7.4), 再由方程(7.4)的成立即知 $u\in C([0,+\infty);D(A))$. □

注 设 u_λ 是方程(7.5)的解:

(a) 若假设 $u_0 \in D(A)$, 则根据第三步可知对每个 $t \geq 0$, $u_\lambda(t)$ 在 $\lambda \to 0$ 时都将收敛到某极限 $u(t)$. 我们实际上可以从这点出发直接证明 $u \in C^1([0,+\infty);H) \cap C([0,+\infty);D(A))$, 且 u 是方程(7.4)的解.

(b) 若只假设 $u_0 \in H$, 则我们依旧可以说明对每个 $t \geq 0$, $u_\lambda(t)$ 在 $\lambda \to 0$ 时都收敛到某极限 $u(t)$, 但此时极限函数 $u(t)$ 可能对 $\forall t > 0$ 都不在 $D(A)$ 内, 且 $u(t)$ 在 $[0,+\infty)$ 上无处可微. 因此 $u(t)$ 此时并不是方程(7.4)的" 经典" 解. 事实上, 对这样的 u_0 而言, 方程(7.4)是没有经典解的, 不过我们可以把 u 看成方程(7.4)的" 广义" 解. 后面我们会说明这种情况在 A 是自伴算子时不会发生: 此时对每个 $u_0 \in H$ 而言 $u(t)$ 都是方程(7.4)的" 经典" 解, 就算 $u_0 \notin D(A)$ 也是如此.

注 对每个 $t \geq 0$, 考虑线性映射 $D(A) \ni u_0 \mapsto u(t) \in D(A)$, 其中 $u(t)$ 是 Hille-Yosida 定理断言存在的方程(7.4)的解. 因为 $\|u(t)\|_H \leq \|u_0\|_H$, 而 $D(A)$ 在 H 中稠密, 故该连续线性映射可以延拓为 $H \to H$ 的有界算子, 记之为[2] $S_A(t)$. 显见 $S_A(t)$ 具有下述性质:

(a) 对每个 $t \geq 0$ 而言均有 $S_A(t) \in L(H)$, 且 $\|S_A(t)\|_{L(H)} \leq 1$,

(b) 对 $\forall t_1, t_2 \geq 0$ 均有 $S_A(t_1+t_2) = S_A(t_1) \circ S_A(t_2)$, 且 $S_A(0) = I$,

(c) $\lim_{t\to 0^+} \|S_A(t)u_0 - u_0\|_H = 0 (\forall u_0 \in H)$.

依赖于参数 $t \geq 0$ 且满足性质 (a)-(c) 的 $H \to H$ 算子族 $\{S(t)\}_{t\geq 0}$ 就称为**压缩连续半群 (continuous semigroup of contractions)**.

Hille 和 Yosida 给出的一个惊人结果断言了压缩连续半群均可以用 $S_A(t)$ 刻画. 也就是说对给定的在 H 上的压缩连续半群 $S(t)$, 总存在唯一极大单调算子 A 使得对 $\forall t \geq 0$ 均有 $S(t) = S_A(t)$. 这表明极大单调算子与压缩连续半群之间存在双射.

注 设 A 是极大单调算子, $\lambda \in \mathbb{R}$, 则初值问题

$$\begin{cases} \frac{du}{dt} + Au + \lambda u = 0, \ t \in [0,+\infty) \\ u(0) = u_0, \end{cases}$$

可以通过下述过程转化为初值问题(7.5): 取

$$v(t) = e^{\lambda t} u(t),$$

则 v 满足

$$\begin{cases} \frac{dv}{dt} + Av = 0, \ t \in [0,+\infty) \\ v(0) = u_0. \end{cases}$$

7.3 正则性

本节说明只要对初值 u_0 加一点额外条件, 就能让 Hille-Yosida 定理7.2中给出的方程(7.4)的解 u 比 $C^1([0,+\infty);H) \cap C([0,+\infty);D(A))$ 更加正则. 为此我们先归纳地定义空间

$$D(A^k) = \{v \in D(A^{k-1}) : Av \in D(A^{k-1})\},$$

其中 k 是整数, 且 $k \geq 2$. 显见 $D(A^k)$ 在赋下述内积后成为 Hilbert 空间:

$$(u,v)_{D(A^k)} = \sum_{j=0}^{k} (A^j u, A^j v)_H,$$

该内积对应的范数为

$$\|u\|_{D(A^k)} = \left(\sum_{j=0}^{k} \|A^j u\|_H^2 \right)^{\frac{1}{2}}.$$

[2] 当然如果我们采用上一注 (b) 的想法, 允许映射的输出是方程(7.4)的广义解的话, 就可以直接在 H 上定义 $S_A(t)$ 为 $H \ni u_0 \mapsto u(t) \in H$.

定理 7.3

设 $u_0 \in D(A^k)$, 其中 k 是不小于 2 的整数, 则通过定理7.2得到的初值问题(7.4)的解 u 满足

$$u \in C^{k-j}([0,+\infty); D(A^j)), \quad \forall j = 0, 1, \cdots, k.$$

证明 首先设 $k = 2$, 考察装备内积 $(u, v)_{D(A)}$ 的 Hilbert 空间 $H_1 = D(A)$. 根据 A 的极大单调性, 显见算子

$$A_1 : H_1 \supset D(A_1) = D(A^2) \to H_1, \quad u \mapsto A_1 u = Au$$

是 H_1 中的极大单调算子. 因此在空间 H_1 中对 A_1 应用 Hille-Yosida 定理7.2知存在函数

$$u \in C^1([0,+\infty); H_1) \cap C([0,+\infty); D(A_1))$$

满足

$$\begin{cases} \frac{du}{dt} + A_1 u = 0, \ t \in [0,+\infty) \\ u(0) = u_0. \end{cases}$$

特别地, u 也满足方程(7.4), 因此根据解的唯一性知 u 实际上就是方程(7.4)的解. 现在只需说明 $u \in C^2([0,+\infty); H)$. 因为

$$A \in L(H_1, H), \ u \in C^1([0,+\infty); H_1),$$

故 $Au \in C^1([0,+\infty); H)$, 且

$$\frac{d}{dt}(Au) = A\left(\frac{du}{dt}\right) \tag{7.18}$$

回到方程(7.4), 由 $Au \in C^1([0,+\infty); H)$ 知 $\frac{du}{dt} \in C^1([0,+\infty); H)$, 因此 $u \in C^2([0,+\infty); H)$ 且

$$\frac{d}{dt}\left(\frac{du}{dt}\right) + A\left(\frac{du}{dt}\right) = 0, \ t \in [0,+\infty). \tag{7.19}$$

下面讨论 $k \geq 3$ 的一般情况, 我们对 k 使用归纳法: 设结果对直到 $k-1$ 的整数都成立, 取 $u_0 \in D(A^k)$. 根据前述分析可知方程(7.4)的解 $u \in C^2([0,+\infty); H) \cap C^1([0,+\infty); D(A))$, 且 u 满足(7.19)式. 设

$$v = \frac{du}{dt},$$

则

$$v \in C^1([0,+\infty); H) \cap C([0,+\infty); D(A)),$$

$$\begin{cases} \frac{dv}{dt} + Av = 0, \ t \in [0,+\infty) \\ v(0) = -Au_0. \end{cases}$$

也就是说, v 是方程(7.4)对应初值 $v_0 = -Au_0$ 的解. 由 $u_0 \in D(A^k)$ 知 $v_0 \in D(A^{k-1})$, 于是根据归纳假设知

$$v \in C^{k-1-j}([0,+\infty); D(A^j)), \quad \forall j = 0, 1, \cdots, k-1, \tag{7.20}$$

亦即

$$u \in C^{k-j}([0,+\infty); D(A^j)), \quad \forall j = 0, 1, \cdots, k-1.$$

现在只需说明

$$u \in C([0,+\infty); D(A^k)). \tag{7.21}$$

在(7.20)式中取 $j = k-1$ 知

$$\frac{du}{dt} \in C([0,+\infty); D(A^{k-1})), \tag{7.22}$$

于是由(7.22)式与方程(7.4)知

$$Au \in C([0,+\infty); D(A^{k-1})),$$

这便是(7.21)式. $\square$

7.4 自伴算子的情形

设 $A : H \supset D(A) \to H$ 是满足 $\overline{D(A)} = H$ 的无界线性算子, 在将 H^* 与 H 等同后, 我们就可以把 A^* 看成 H 中的无界线性算子.

定义 7.3 (对称算子 (symmetric operator), 自伴算子 (self-adjoint operator))

- 若 A 满足对 $\forall u, v \in D(A)$ 均有 $(Au, v) = (u, Av)$, 就称 A 是对称算子.
- 若 A 满足 $D(A^*) = D(A)$, $A^* = A$, 就称 A 是自伴算子.

注 对有界算子而言, 对称与自伴这两个概念是相同的. 但如果 A 无界, 则对称算子与自伴算子之间还是有些细微的差距. 显见任意自伴算子都是对称算子, 但反之不然: 算子 A 对称当且仅当 $A \subset A^*$(即 $D(A) \subset D(A^*)$, 且在 $D(A)$ 上 $A^* = A$, 实际上 A 作为对称算子使得 $D(A) \neq D(A^*)$ 的情况是存在的). 我们的下一条结果表明只要 A 是极大单调算子, 那么

$$A \text{ 对称} \Leftrightarrow A \text{ 自伴}.$$

命题 7.3 (极大单调算子自伴与对称的等价性)

设 A 是极大单调对称算子, 则 A 自伴.

证明 记 $J_1 = (I + A)^{-1}$, 我们首先说明 J_1 自伴. 因为 $J_1 \in L(H)$, 故只需说明

$$(J_1 u, v) = (u, J_1 v), \quad \forall u, v \in H. \tag{7.23}$$

记 $u_1 = J_1 u, v_1 = J_1 v$, 则

$$u_1 + Au_1 = u,$$
$$v_1 + Av_1 = v.$$

因为 A 对称, 故 $(u_1, Av_1) = (Au_1, v_1)$, 因此 $(u_1, v) = (u, v_1)$, 这便是(7.23)式.

现设 $u \in D(A^*)$, 记 $f = u + A^* u$, 知

$$(f, v) = (u, v + Av), \quad \forall v \in D(A),$$

即

$$(f, J_1 w) = (u, w), \quad \forall w \in H.$$

其中 $w \in H$ 是因为 $R(I + A) = H$. 由上式可知 $u = J_1 f$, 因此 $u \in D(A)$, 故 $D(A^*) = D(A)$, 从而 A 自伴. □

注 特别注意当 A 是单调算子 (乃至对称单调算子) 时, A^* 未必是单调算子. 不过我们可以证明下述等价:

$$\begin{aligned} A \text{ 是极大单调算子} &\Leftrightarrow A^* \text{ 是极大单调算子} \\ &\Leftrightarrow A \text{ 是闭算子}, D(A) \text{ 在 } H \text{ 中稠密}, A, A^* \text{ 同时是单调算子}. \end{aligned}$$

定理 7.4 (自伴算子的 Hille-Yosida 定理)

设 A 是自伴极大单调算子, 则对每个 $u_0 \in H$ 均存在唯一函数

$$u \in C([0, +\infty); H) \cap C^1((0, +\infty); H) \cap C((0, +\infty); D(A))$$

满足

$$\begin{cases} \frac{du}{dt} + Au = 0, \ t \in (0, +\infty) \\ u(0) = u_0. \end{cases}$$

且有

$$\|u(t)\|_H \le \|u_0\|_H, \quad \left\|\frac{du}{dt}(t)\right\|_H = \|Au(t)\|_H \le \frac{1}{t}\|u_0\|_H, \quad \forall t > 0,$$

$$u \in C^k((0,+\infty); D(A^l)), \quad \forall k, l \in \mathbb{N}. \tag{7.24}$$

注 我们这里强调一下定理7.4与原本的 Hille-Yosida 定理7.2有何区别. 定理7.4减弱了 u_0 的条件 (这里只需要 $u_0 \in H$, 而定理7.2要求 $u_0 \in D(A)$), 且增强了方程(7.4)解在 $t=0$ 之外的正则性. 但 $|\frac{du}{dt}(t)|$ 此时可能在 $t \to 0$ 时爆破.

证明 对于唯一性, 设 $u, \overline{u}$ 是初值问题的两个解, 由 A 的单调性可知 $\varphi(t) = \|u(t) - \overline{u}(t)\|_H^2$ 在 $(0,+\infty)$ 上不增. 另一方面, φ 在 $[0,+\infty)$ 上连续且 $\varphi(0) = 0$, 因此 $\varphi \equiv 0$.

对于存在性, 我们分两步进行证明.

$\boxed{\text{第一步: 通过增强 } u_0 \text{ 的条件得到结论.}}$ 设 $u_0 \in D(A^2)$, 则由 Hille-Yosida 定理7.2知方程(7.4)存在解 u. 下面说明

$$\left\|\frac{du}{dt}(t)\right\| \le \frac{1}{t}\|u_0\|_H, \quad \forall t > 0. \tag{7.25}$$

根据极大单调算子的性质7.1, 对 $\forall \lambda > 0$ 而言 λA 是极大单调算子, 而由 A 的自伴性显然可得 λA 的自伴性, 于是仿照命题7.3的证明可知

$$J_\lambda^* = J_\lambda, \quad \forall \lambda > 0$$

根据 A_λ 的构造进一步可知

$$A_\lambda^* = A_\lambda, \quad \forall \lambda > 0$$

回到证明 Hille-Yosida 定理7.2时给出的逼近初值问题:

$$\begin{cases} \frac{du_\lambda}{dt} + A_\lambda u_\lambda = 0, \ t \in [0,+\infty) \\ u_\lambda(0) = u_0. \end{cases} \tag{7.26}$$

我们希望先对 u_λ 给出形如(7.25)式的估计, 再用极限过渡到 u 上. 为此将(7.26)式两端与 u_λ 作内积, 再在 $[0,T]$ 上积分, 可得

$$\frac{1}{2}\|u_\lambda(T)\|_H^2 + \int_0^T (A_\lambda u_\lambda, u_\lambda)dt = \frac{1}{2}\|u_0\|_H^2. \tag{7.27}$$

再将(7.26)式两端与 $t\frac{du_\lambda}{dt}$ 作内积并在 $[0,T]$ 上积分, 可得

$$\int_0^T \left\|\frac{du_\lambda}{dt}(t)\right\|_H^2 tdt + \int_0^T \left(A_\lambda u_\lambda(t), \frac{du_\lambda}{dt}(t)\right) tdt = 0. \tag{7.28}$$

因为 $A_\lambda^* = A_\lambda$, 故

$$\frac{d}{dt}(A_\lambda u_\lambda, u_\lambda) = \left(A_\lambda \frac{du_\lambda}{dt}, u_\lambda\right) + \left(A_\lambda u_\lambda, \frac{du_\lambda}{dt}\right) = 2\left(A_\lambda u_\lambda, \frac{du_\lambda}{dt}\right).$$

现在对(7.28)左式第二项分部积分有

$$\int_0^T \left(A_\lambda u_\lambda(t), \frac{du_\lambda}{dt}\right) tdt = \frac{1}{2}\int_0^T \frac{d}{dt}[(A_\lambda u_\lambda, u_\lambda)]tdt = \frac{1}{2}(A_\lambda u_\lambda(T), u_\lambda(T))T - \frac{1}{2}\int_0^T (A_\lambda u_\lambda, u_\lambda)dt. \tag{7.29}$$

另一方面, 根据引理7.1, 函数 $t \mapsto |\frac{du_\lambda}{dt}(t)|$ 不增, 因此

$$\int_0^T \left\|\frac{du_\lambda}{dt}(t)\right\|_H^2 tdt \ge \left|\frac{du_\lambda}{dt}(T)\right|^2 \frac{T^2}{2}. \tag{7.30}$$

综合(7.27)-(7.30)式知

$$\frac{1}{2}\|u_\lambda(T)\|_H^2 + T(A_\lambda u_\lambda(T), u_\lambda(T)) + T^2\left\|\frac{du_\lambda}{dt}(T)\right\|^2 \le \frac{1}{2}\|u_0\|_H^2,$$

根据 A_λ 的单调性知上式表明

$$\left\|\frac{du_\lambda}{dt}(t)\right\|_H \le \frac{1}{T}\|u_0\|_H, \quad \forall T > 0. \tag{7.31}$$

在(7.31)式中令 $\lambda \to 0$, 由 Hille-Yosida 定理7.2证明第五步给出的 $\frac{du_\lambda}{dt} \to \frac{du}{dt}$ 即得(7.25)式.

$\boxed{\text{第二步: 减弱 } u_0 \text{ 的条件.}}$ 现在设 $u_0 \in H$. 引理7.2表明 $D(A^2)$ 在图范数意义下在 $D(A)$ 中稠密, 进而 $D(A^2)$ 在 $\|\cdot\|_h$ 意义下在 $D(A)$ 中稠密, 又因为 $D(A)$ 在 $\|\cdot\|_H$ 意义下在 H 中稠密, 故 $D(A^2)$ 在 $\|\cdot\|_H$ 意义下在 H 中稠密, 这说明存在 $\{u_{0,n}\}_{n\in\mathbb{N}} \subset D(A^2)$ 使得 $u_{0,n} \to u_0$. 设 u_n 是下述初值问题的解:

$$\begin{cases} \frac{du_n}{dt} + Au_n = 0, \ t \in [0, +\infty) \\ u_n(0) = u_{0,n}. \end{cases}$$

由 Hille-Yosida 定理7.2证明的第六步知

$$\|u_n(t) - u_m(t)\|_H \le \|u_{0,n} - u_{0,m}\|_H, \quad \forall m, n, \forall t \ge 0.$$

再由前述第一步的结果知

$$\left\|\frac{du_n}{dt}(t) - \frac{du_m}{dt}(t)\right\|_H \le \frac{1}{t}\|u_{0,n} - u_{0,m}\|_H, \quad \forall m, n, \forall t > 0.$$

这说明 $u_n(t)$ 在 $[0, +\infty)$ 上关于 t 一致收敛到某极限 $u(t)$, $\frac{du_n}{dt}(t)$ 在任意形如 $[\delta, +\infty)(\delta > 0)$ 的区间上关于 t 一致收敛到 $\frac{du}{dt}(t)$. 因此极限函数 u 满足

$$u \in C([0, +\infty); H) \cap C^1((0, +\infty); H).$$

又因为 A 是闭算子, 故

$$\begin{cases} u_n(t) \to u(t), \forall t > 0 \\ \frac{du_n}{dt}(t) = -Au_n(t) \to \frac{du}{dt}(t), \forall t > 0 \end{cases} \Rightarrow \begin{cases} u(t) \in D(A), \forall t > 0 \\ \frac{du}{dt}(t) = -Au(t), \forall t > 0 \end{cases}$$

因此 u 是方程(7.4)的解. 下面我们证明正则性结果(7.24), 考虑对 $k \ge 2$ 用归纳法, 说明

$$u \in C^{k-j}((0, +\infty); D(A^j)), \quad \forall j = 0, 1, \cdots, k. \tag{7.32}$$

设(7.32)式对直到 $k-1$ 的整数均成立, 于是特别对 $k-1$ 有

$$u \in C((0, +\infty); D(A^{k-1})). \tag{7.33}$$

现在为了说明(7.32)式在 k 时也成立, 仿照定理7.3的证明知只需证明

$$u \in C((0, +\infty); D(A^k)) \tag{7.34}$$

即可. 考虑 Hilbert 空间 $\widetilde{H} = D(A^{k-1})$ 与算子

$$\widetilde{A} : \widetilde{H} \supset D(\widetilde{A}) = D(A^k) \to \widetilde{H}, \quad u \mapsto \widetilde{A}u = Au,$$

显见 $\widetilde{A}$ 是 $\widetilde{H}$ 中的极大单调对称算子, 因此它是自伴的. 对 $\widetilde{A}$ 在 $\widetilde{H}$ 中应用已经证明的定理7.4第一条断言知对任意给定的 $v_0 \in \widetilde{H}$ 而言, 下述初值问题存在唯一解 v:

$$\begin{cases} \frac{dv}{dt} + Av = 0, \ t \in (0, +\infty) \\ v(0) = v_0. \end{cases} \tag{7.35}$$

且

$$v \in C([0, +\infty); \widetilde{H}) \cap C^1((0, +\infty); \widetilde{H}) \cap C((0, +\infty); D(\widetilde{A})).$$

根据(7.33)式, 不妨取 $v_0 = u(\varepsilon)(\varepsilon > 0)$, 这便得到 $u \in C((\varepsilon, +\infty); D(A^k))$, 令 $\varepsilon \to 0$ 即得欲证. □

参考文献

[HB] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*[M]. New York: Springer, 2010.

[HB2] Haim Brezis. 泛函分析-理论和应用 [M]. 北京: 清华大学出版社, 2009.

[LPD] 刘培德. 拓扑线性空间与算子谱理论 [M]. 北京: 高等教育出版社, 2013.

[PGC2] Philippe G. Ciarlet. 线性与非线性泛函分析及其应用 (下册)[M]. 北京: 高等教育出版社, 2017.

[SW] L. Schwartz. 广义函数论 [M]. 北京: 高等教育出版社, 2010.

[XMY] 许全华, 马涛, 尹智. 泛函分析讲义 [M]. 北京: 高等教育出版社, 2017.

[YDC] 杨大春, 袁文. 泛函分析选讲 [M]. 北京: 北京师范大学出版社, 2016.

[ZGQ] 张恭庆, 林源渠. 泛函分析讲义 (第二版)(上)[M]. 北京: 北京大学出版社, 2021.

[ZMQ] 周民强. 实变函数论 (第三版)[M]. 北京: 北京大学出版社, 2016.

[Zo] B. A. 卓里奇. 数学分析 (第二卷)(第四版)[M]. 北京: 高等教育出版社, 2019.