

线性代数：从有限维到无穷维

Shun Li

Jul 3, 2026

整理一下线性代数从有限维到无穷维的一些差异，实际就是从线性代数到最基本的泛函分析的内容。

目录

第一部分 基础篇	4
1 基的概念不再唯一	4
1.1 线性空间的 Hamel 基	4
1.2 可分 Banach 空间的 Schauder 基	5
1.3 Hilbert 空间的标准正交基	5
2 范数/拓扑不再唯一	6
3 对偶空间与双对偶空间	8
3.1 代数对偶与连续对偶空间	9
3.2 连续对偶空间与原空间的关系	10
3.3 双对偶空间、典范嵌入与自反性	12
4 范数拓扑下紧性的丢失	16
5 弱收敛、弱拓扑与弱紧性	17
5.1 强收敛与弱收敛	17
5.2 弱拓扑	19
5.3 弱紧性	19
5.4 弱星拓扑	21
6 子空间的闭性与投影定理	21
6.1 子空间的闭性	22
6.2 投影定理	22
第二部分 进阶篇	25

目录	2
7 线性算子的连续性和收敛性	25
7.1 线性算子的连续性	25
7.2 线性算子的收敛性	27
8 线性方程的可解性、闭值域与适定性	28
9 有限秩算子与紧算子	30
10 特征值与谱理论	33
11 酉算子、正规算子与谱定理	34
11.1 酉算子与正规算子	35
11.2 谱定理	37
12 Fourier 理论	39
12.1 交换群上平移的谱表示	39
12.2 离散 Fourier 变换	41
12.3 Fourier 级数	41
12.4 离散时间 Fourier 变换	41
12.5 Fourier 变换	42
13 Fredholm 算子与 Fredholm 二择一定理	42
14 行列式与迹	45
14.1 迹的推广	45
14.2 行列式的推广	46
15 从有限维到无穷维的总览	46
 第三部分 基本定理篇	 48
16 Baire 纲定理	48
17 开映射定理与有界逆定理	49
18 闭图像定理	51
19 一致有界原理（共鸣定理）	53
20 Hahn–Banach 定理	55
21 基本定理之间的联系	57
 第四部分 数值篇	 59
22 Lax 等价定理	59

23 变分问题	65
23.1 连续变分问题	65
23.2 离散变分问题	69
24 混合问题	72
24.1 连续混合问题	72
24.2 离散混合问题	74
24.3 鞍点矩阵	75

第一部分 基础篇

1 基的概念不再唯一

1.1 线性空间的 Hamel 基

在有限维线性空间中，基通常指代数意义下的 Hamel 基。

Definition 1.1 (有限维线性空间的基). 设 V 是数域 $\mathbb{F}$ 上的线性空间，其中 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 。一组向量

$$\{e_1, \dots, e_n\} \subset V$$

称为空间 V 的一组基，是指它线性无关，并且

$$V = \text{span}\{e_1, \dots, e_n\}.$$

等价地，对任意 $v \in V$ ，存在唯一的一组标量 $\{c_1, \dots, c_n\}$ ，使得

$$v = \sum_{j=1}^n c_j e_j.$$

有限维情形的一个基本事实是：任意两组基所含元素个数相同，这个公共的个数称为空间的维数。不同的有限维线性空间可以用维数进行分类：任意的 n 维的线性空间都线性同构于 $\mathbb{F}^n$ 。对于一个固定的有限维线性空间，虽然基本本身并不唯一，但不同基之间总可以通过可逆线性变换相互转换。因此，在有限维线性代数中，“基”主要是坐标表示的工具，不同基之间不会导致本质上不同的分析框架。

在无穷维线性空间中，情况则存在本质区别：从代数结构出发，虽然仍然可以定义 Hamel 基，但是它不再像有限维空间的基那么好用了。

Definition 1.2 (Hamel 基). 设 V 是线性空间。一组向量 $\mathcal{B} \subset V$ 称为 V 的 Hamel 基，如果 $\mathcal{B}$ 线性无关，并且任意 $v \in V$ 都可以表示为有限线性组合

$$v = \sum_{j=1}^m c_j b_j, \quad b_j \in \mathcal{B},$$

且这种表示在合并相同基向量后唯一。

这个定义保留了有限维线性代数中“有限线性组合”的要求，因为无限求和（包括可数以及不可数）需要依赖收敛概念，无法从纯线性结构中得到。但是，在许多常见的无穷维 Banach 空间中，Hamel 基的基数非常大，通常不可数。

借助选择公理，任意线性空间都存在 Hamel 基。但是在无穷维赋范空间中，Hamel 基通常并不适合分析问题：它往往具有非常大的基数，在许多常见空间中是不可数的，并且一般不与空间的范数、极限、收敛等拓扑结构相容。因此，Hamel 基虽然在代数上完备，却很少用于具体的泛函分析和数值分析。

1.2 可分 Banach 空间的 Schauder 基

为了反映无穷维空间中的收敛结构，需要在纯线性结构之外引入拓扑结构。最常见的情形是完备的赋范线性空间（即 Banach 空间），此时空间中存在一个范数 $\|\cdot\|$ ，允许讨论收敛和级数，也可以定义 Schauder 基。

Definition 1.3 (Schauder 基). 设 X 是 Banach 空间。一系列向量

$$\{e_j\}_{j=1}^{\infty} \subset X$$

称为 X 的 Schauder 基，如果对任意 $x \in X$ ，存在唯一的一系列标量 $\{c_j\}_{j=1}^{\infty}$ ，使得

$$x = \sum_{j=1}^{\infty} c_j e_j,$$

其中级数在范数意义下收敛，即

$$\lim_{N \rightarrow \infty} \left\| x - \sum_{j=1}^N c_j e_j \right\| = 0.$$

Schauder 基与 Hamel 基存在明显区别：Schauder 基的向量个数是至多可数的，Hamel 基的向量个数通常是不可数的；Schauder 基允许无穷级数表示（依赖空间中的范数和收敛概念）；Hamel 基只允许有限求和（纯代数概念）。

Schauder 基的存在性直接蕴含了 Banach 空间的可分性，但是可分 Banach 空间并不是必然存在 Schauder 基。可分性只说明空间中存在一个可数稠密子集，而 Schauder 基要求每个元素都能以唯一的收敛级数形式表示出来，这比可分性强得多：可分性提供的是“可以用可数集合逼近”，而 Schauder 基提供的是“可以用固定的一系列向量作唯一的级数展开”。因此

$$\text{存在 Schauder 基} \implies \text{空间可分},$$

反之一般不成立。

1.3 Hilbert 空间的标准正交基

对于 Hilbert 空间，内积结构允许进一步讨论 Hilbert 基（标准正交基），甚至还可以不依赖可分性的条件。

Definition 1.4 (Hilbert 基). 设 H 是 Hilbert 空间。一组标准正交向量族

$$\{e_{\alpha}\}_{\alpha \in A} \subset H$$

称为 H 的 Hilbert 基（或完备标准正交系），如果

$$\langle e_{\alpha}, e_{\beta} \rangle = \delta_{\alpha\beta},$$

并且满足

$$\overline{\text{span}\{e_{\alpha} : \alpha \in A\}} = H.$$

对于可分 Hilbert 空间，标准正交基包括可数多个向量，并且可以据此给出任意向量的正交展开。

Proposition 1.1 (可分 Hilbert 空间中的正交展开). 设 H 是无穷维可分 Hilbert 空间。在 H 中存在一组可数标准正交基

$$\{e_j\}_{j=1}^{\infty}.$$

此时任意 $x \in H$ 都可以展开为

$$x = \sum_{j=1}^{\infty} \langle x, e_j \rangle e_j,$$

并且级数在 H 的范数意义下收敛。同时成立 Parseval 等式：

$$\|x\|_H^2 = \sum_{j=1}^{\infty} |\langle x, e_j \rangle|^2.$$

这正是有限维欧氏空间中标准正交基展开的无穷维推广。可分 Hilbert 空间在结构上与有限维内积空间最为接近：虽然维数可以是无穷的，但仍然可以用可数正交基和平方可和的系数序列来描述空间中的元素。

Remark 1.1 (不可分 Hilbert 空间). 若 Hilbert 空间不可分，其标准正交基通常是不可数的。此时向量的展开不能简单理解为普通的可数级数。对于任意 $x \in H$ ，集合

$$\{\alpha \in A : \langle x, e_\alpha \rangle \neq 0\}$$

至多可数，并且满足 Parseval 恒等式

$$\|x\|_H^2 = \sum_{\alpha \in A} |\langle x, e_\alpha \rangle|^2,$$

其中对不可数指标集的求和按平方可和族的意义理解，而不是普通的可数级数求和。

综上，从有限维线性代数过渡到泛函分析时，“基”的概念发生了分化。Hamel 基保留代数上的有限线性组合，但通常不适合分析；Schauder 基适合 Banach 空间中的级数展开，但不是所有可分 Banach 空间都有；Hilbert 基利用内积结构给出正交展开，在可分 Hilbert 空间中具有最接近有限维正交基的形式，此时也是一个特殊的 Schauder 基。这里的核心区别在于：有限维空间中线性结构已经足够支撑基和坐标表示，而无穷维空间中的基必须同时考虑范数、极限、完备性、可分性和内积结构。

2 范数/拓扑不再唯一

有限维线性空间与无穷维线性空间的一个重要差异在于：有限维中范数虽然形式不唯一，但诱导的拓扑唯一。在无穷维情形下才会发生真正的分化。

Theorem 2.1 (有限维空间中的范数等价). 设 V 是数域 $\mathbb{F}$ 上的有限维线性空间，其中 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 。若 $\|\cdot\|_a$ 和 $\|\cdot\|_b$ 是 V 上的两个范数，则存在常数 $C_1, C_2 > 0$ ，使得对任意 $v \in V$ ，都有

$$C_1 \|v\|_a \leq \|v\|_b \leq C_2 \|v\|_a.$$

因此，有限维空间上的所有范数诱导出相同的收敛概念。换言之，对任意向量列 $\{v_m\}_{m=1}^{\infty} \subset V$ ，

有

$$v_m \rightarrow v_0 \text{ in } \|\cdot\|_a \iff v_m \rightarrow v_0 \text{ in } \|\cdot\|_b.$$

Proof. 取 V 的一组基 $e_1, \dots, e_n$, 并定义坐标范数

$$\left\| \sum_{j=1}^n x_j e_j \right\|_1 = \sum_{j=1}^n |x_j|.$$

对任意范数 $\|\cdot\|$, 由三角不等式,

$$\left\| \sum_{j=1}^n x_j e_j \right\| \leq \left(\max_{1 \leq j \leq n} \|e_j\| \right) \sum_{j=1}^n |x_j|.$$

因此 $x \mapsto \|x\|$ 关于坐标范数 $\|\cdot\|_1$ 连续。集合

$$S = \{x \in V : \|x\|_1 = 1\}$$

在坐标映射下对应于 $\mathbb{F}^n$ 中的紧集, 所以 S 是紧的。连续函数 $x \mapsto \|x\|$ 在 S 上取得最小值 m , 并且 $m > 0$, 因为 S 不含零向量。于是存在常数 $m, M > 0$, 使得

$$m\|x\|_1 \leq \|x\| \leq M\|x\|_1, \quad x \in V.$$

任意两个范数都与 $\|\cdot\|_1$ 等价, 因而彼此等价。 $\square$

在有限维赋范线性空间中, 虽然可以写下许多不同形式的范数, 例如 $\|x\|_1, \|x\|_2, \|x\|_\infty$, 但它们不会导致不同的拓扑结构。在这种意义下, 有限维空间中由范数诱导的拓扑是唯一的。

Proposition 2.2 (有限维赋范空间自动完备). 有限维赋范线性空间 V 配备任意范数 $\|\cdot\|$ 得到的赋范空间 X 都是完备的。即若 $\{x_m\}_{m=1}^\infty \subset X$ 是 *Cauchy* 列, 则存在 $x_0 \in X$, 使得

$$\lim_{m \rightarrow \infty} \|x_m - x_0\| = 0.$$

Proof. 证明思路是先证明 $\mathbb{F}^n$ 配备 ℓ^1 范数后的完备性 (因为 $\mathbb{F}$ 本身是完备的), 然后利用有限维范数等价定理完成证明。 $\square$

因此, 所有有限维赋范线性空间都是 Banach 空间。更具体地说, 若 $\dim V = n$, 则 V 作为线性空间同构于 $\mathbb{F}^n$ 。对于有限维情形, 这个线性同构同时也是拓扑意义下的同构。也就是说, 任意有限维赋范线性空间不仅在线性代数意义下等价于 $\mathbb{F}^n$, 而且在由范数诱导的拓扑意义下也等价于 $\mathbb{F}^n$ 。这说明有限维空间中的线性结构与拓扑结构高度相容: 一旦维数固定, 范数的具体选择通常不会改变空间的基本分析性质。

无穷维情形则完全不同。设 X 是无穷维线性空间, 定义在 X 上的两个范数未必等价。不同范数可能诱导出不同的收敛概念和不同的拓扑结构。某个序列在一种范数下收敛, 并不意味着它在另一种范数下也收敛。

Example 2.1 (多项式空间上的不等价范数). 考虑多项式空间 $\mathcal{P}([0, 1])$, 在这个空间上可以定义

$$\|p\|_\infty = \max_{x \in [0, 1]} |p(x)|, \quad \|p\|_1 = \int_0^1 |p(x)| dx.$$

这两个范数在有限维多项式子空间 $\mathcal{P}_n([0, 1])$ 上等价，但在整个无穷维空间 $\mathcal{P}([0, 1])$ 上不等价。

Proof. 在 $\mathcal{P}([0, 1])$ ，显然有 $\|p\|_1 \leq \|p\|_\infty$ ，下面证明不存在常数 $C > 0$ 使得

$$\|p\|_\infty \leq C\|p\|_1, \quad \forall p \in \mathcal{P}([0, 1])$$

只需要构造多项式序列 $p_n(x) = x^n$ ，易得

$$\|p_n\|_\infty = 1, \quad \|p_n\|_1 = \frac{1}{n+1}$$

显然不存在 $C > 0$ 对所有 n 保证上述不等式成立。因此，两个范数不等价。 $\square$

无穷维空间中存在大量的 Banach 空间例子，例如

$$C[0, 1], \quad L^p(\Omega), \quad \ell^p, \quad 1 \leq p \leq \infty.$$

这些空间在各自的标准范数下是完备的，但它们显然不再同构于某个有限维空间 $\mathbb{F}^n$ 。因此，无穷维 Banach 空间不是有限维线性代数的简单延伸，而是具有新的拓扑和分析结构。例如配备上述两个范数后的多项式空间 $\mathcal{P}([0, 1])$ 都不是完备的，并且进行完备化会得到不同的 Banach 空间。

Example 2.2 (多项式空间的完备化). 对于多项式空间 $\mathcal{P}([0, 1])$:

1. 在配备一致范数

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

时不是完备空间。由 Weierstrass 逼近定理， $\mathcal{P}([0, 1])$ 在 $C([0, 1])$ 中稠密，因此它在一致范数下的完备化是 $C[0, 1]$ 。

2. 在配备 L^1 范数

$$\|p\|_1 = \int_0^1 |p(x)| dx.$$

时不是完备空间。它的完备化是 $L^1[0, 1]$ 。

综上，在有限维线性空间中，所有范数等价，所有范数诱导的拓扑相同，并且所有有限维赋范线性空间都是 Banach 空间。因此，有限维线性代数中通常不需要特别区分不同范数带来的拓扑差异。而在无穷维线性空间中，不同范数可能不等价，诱导的拓扑可能不同，赋范空间也未必完备。于是，范数的选择、拓扑结构和完备性本身都成为泛函分析中的核心内容。

3 对偶空间与双对偶空间

对于有限维空间，对偶空间以及双对偶空间的结构都是相对平凡的，因为维数还是和原本的空间一样。但是对于无穷维空间，尤其在无穷维赋范空间中，会产生非常复杂的结果。

3.1 代数对偶与连续对偶空间

Definition 3.1 (代数对偶空间). 设 V 是数域 $\mathbb{F}$ 上的线性空间, 其中 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$. V 的代数对偶空间定义

$$V^\# = \{f: V \rightarrow \mathbb{F} : f \text{ 是线性映射}\}.$$

对于赋范空间, 范数的引入允许我们考虑更小的连续对偶空间。这里用记号 $V^\#$ 以区分代数对偶 (所有线性泛函) 和在泛函分析中更常见的连续对偶 X^* (所有连续线性泛函), 下文中对偶默认指的是连续对偶, 除非强调代数对偶。

Definition 3.2 (连续对偶空间). 若 X 是赋范线性空间, 定义连续对偶空间:

$$X^* = \{f: X \rightarrow \mathbb{F} : f \text{ 是连续线性泛函}\}.$$

等价地, $f \in X^*$ 当且仅当存在常数 $C > 0$, 使得

$$|f(x)| \leq C\|x\|_X, \quad \forall x \in X.$$

此时可以定义对偶范数

$$\|f\|_{X^*} = \sup_{\|x\|_X \leq 1} |f(x)|.$$

在有限维赋范线性空间中, 代数对偶空间 $V^\#$ 与连续对偶空间 V^* 没有区别。

Proposition 3.1 (有限维中的代数对偶与连续对偶). 对于有限维线性空间 V , 配备任意范数得到的赋范空间 X , 代数对偶空间 $V^\#$ 与连续对偶空间 V^* 相同。

Proof. 先考虑 ℓ^1 范数, 此时任意线性泛函 f 的连续性是显然的: 对 x 展开

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

易得

$$|f(x)| \leq \left(\max_i |f(e_i)| \right) \sum_{i=1}^n |x_i| = \left(\max_i |f(e_i)| \right) \|x\|_1$$

然后用范数等价定理转换即可。 □

但是在无穷维赋范线性空间中, 代数对偶与连续对偶通常差别很大。由于无穷维赋范空间中存在大量不连续线性泛函, 连续对偶空间 X^* 只是代数对偶空间 $X^\#$ 的一部分:

$$X^* \subset X^\#.$$

这个包含关系一般是严格成立的。

Example 3.1. 考虑多项式空间 $\mathcal{P}([0, 1])$, 配备 $\|p\|_\infty$ 范数, 那么在 $x = 1$ 处求导的线性泛函

$$f: p \mapsto p'(1)$$

是不连续的。

Proof. 取 $p_n(x) = x^n$ 计算可得

$$f(p_n) = p'_n(1) = n$$

易得 f 不连续。 $\square$

代数对偶空间只关心线性结构，不关心范数和连续性，因此它通常过于庞大，不适合泛函分析中的极限、收敛和估计问题。泛函分析中主要研究的是连续对偶空间，因为连续线性泛函与空间的拓扑结构相容。

3.2 连续对偶空间与原空间的关系

对于有限维线性空间 V ，连续对偶空间 V^* 与原空间具有相同维数：

$$\dim V^\# = \dim V.$$

选定 V 的一组基

$$\{e_1, \dots, e_n\} \subset V,$$

可以自然地定义 $V^\#$ 的一组基（称为对偶基）

$$\{e^1, \dots, e^n\} \subset V^\#,$$

使得

$$e^i(e_j) = \delta_{ij}.$$

因此，在有限维线性代数中，对偶空间可以看作原空间的自然镜像。虽然 V 与 $V^\#$ 之间没有不依赖于额外结构的典范同构，但选定基以后，二者可以自然地一一对应。

对于无穷维 Banach 空间，连续对偶空间与原空间则可能有显著差异。

Example 3.2.

$$(\ell^1)^* \cong \ell^\infty.$$

Proof. 对任意 $a = (a_n)_{n=1}^\infty \in \ell^\infty$ ，定义

$$f_a(x) = \sum_{n=1}^{\infty} a_n x_n, \quad x = (x_n)_{n=1}^\infty \in \ell^1.$$

由

$$|f_a(x)| \leq \|a\|_{\ell^\infty} \|x\|_{\ell^1}$$

可知 f_a 是 ℓ^1 上的连续线性泛函。

反过来，设 $f \in (\ell^1)^*$ 。令 e_n 表示第 n 个标准基向量，并定义 $a_n = f(e_n)$ 。由于

$$|a_n| = |f(e_n)| \leq \|f\|_{(\ell^1)^*} \|e_n\|_{\ell^1} = \|f\|_{(\ell^1)^*},$$

所以 $a = (a_n)_{n=1}^\infty \in \ell^\infty$ 。对任意有限支撑序列 x ，有

$$f(x) = \sum_{n=1}^{\infty} a_n x_n.$$

由于有限支撑序列在 ℓ^1 中稠密，且两边都是 ℓ^1 上的连续线性泛函，因此上式对任意 $x \in \ell^1$ 成立。因此

$$(\ell^1)^* \cong \ell^\infty$$

在等距同构意义下成立。 $\square$

Example 3.3.

$$(\ell^p)^* \cong \ell^q.$$

其中 $1 < p < \infty$, q 为 p 的共轭指数, 即

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Proof. 对任意 $a = (a_n)_{n=1}^\infty \in \ell^q$, 定义

$$f_a(x) = \sum_{n=1}^{\infty} a_n x_n, \quad x \in \ell^p.$$

由 Hölder 不等式,

$$|f_a(x)| \leq \|a\|_{\ell^q} \|x\|_{\ell^p}.$$

因此 $f_a \in (\ell^p)^*$.

反过来, 设 $f \in (\ell^p)^*$. 令

$$a_n = f(e_n).$$

对任意有限支撑序列 x , 有

$$f(x) = \sum_{n=1}^{\infty} a_n x_n.$$

下面验证 $a \in \ell^q$. 对每个 N , 定义有限支撑序列

$$x_n^{(N)} = \begin{cases} \operatorname{sgn}(a_n) |a_n|^{q-1}, & 1 \leq n \leq N, \\ 0, & n > N, \end{cases}$$

其中实数情形取通常的符号函数, 复数情形取 $\operatorname{sgn}(a_n) = \overline{a_n}/|a_n|$, 并约定 $\operatorname{sgn}(0) = 0$. 于是

$$\sum_{n=1}^N |a_n|^q = f(x^{(N)}) \leq \|f\|_{(\ell^p)^*} \left(\sum_{n=1}^N |a_n|^q \right)^{1/p}.$$

因此

$$\left(\sum_{n=1}^N |a_n|^q \right)^{1/q} \leq \|f\|_{(\ell^p)^*}.$$

令 $N \rightarrow \infty$, 得到 $a \in \ell^q$, 并且 $\|a\|_{\ell^q} \leq \|f\|_{(\ell^p)^*}$. 由于有限支撑序列在 ℓ^p 中稠密, 所以对任意 $x \in \ell^p$, 仍有

$$f(x) = \sum_{n=1}^{\infty} a_n x_n.$$

因此每个 ℓ^p 上的连续线性泛函都唯一地由某个 ℓ^q 序列给出, 从而

$$(\ell^p)^* \cong \ell^q.$$

□

Example 3.4. ℓ^∞ 的对偶空间最为复杂, 虽然有自然嵌入

$$\ell^1 \hookrightarrow (\ell^\infty)^*,$$

但是

$$\ell^1 \subsetneq (\ell^\infty)^*.$$

Proof. 对任意 $a = (a_n)_{n=1}^\infty \in \ell^1$, 定义

$$f_a(x) = \sum_{n=1}^{\infty} a_n x_n, \quad x \in \ell^\infty.$$

由于

$$|f_a(x)| \leq \|a\|_{\ell^1} \|x\|_{\ell^\infty},$$

所以 $f_a \in (\ell^\infty)^*$. 因此 ℓ^1 可以自然地嵌入到 $(\ell^\infty)^*$ 中。

但这个嵌入不是满射。令 c 表示所有收敛序列构成的空间, 并定义

$$L(x) = \lim_{n \rightarrow \infty} x_n, \quad x \in c.$$

则 L 是 c 上的连续线性泛函。由 Hahn-Banach 定理, L 可以延拓为 ℓ^∞ 上的连续线性泛函

$$\tilde{L} \in (\ell^\infty)^*.$$

如果 $\tilde{L}$ 可以由某个 $a \in \ell^1$ 表示, 即

$$\tilde{L}(x) = \sum_{n=1}^{\infty} a_n x_n,$$

那么对每个标准基向量 e_n , 由于 $e_n \in c$ 且 $\lim_{k \rightarrow \infty} (e_n)_k = 0$, 应有

$$a_n = \tilde{L}(e_n) = 0.$$

因此 $a = 0$ 。但对常数序列 $\mathbf{1} = (1, 1, \dots)$, 有

$$\tilde{L}(\mathbf{1}) = 1,$$

这与 $a = 0$ 矛盾。因此 $\tilde{L}$ 不是由任何 ℓ^1 序列表示的泛函。所以

$$\ell^1 \subsetneq (\ell^\infty)^*.$$

□

在 Hilbert 空间中, 连续对偶空间与原空间仍然保持非常好的对应关系。

Theorem 3.2 (Riesz 表示定理). 设 H 是 Hilbert 空间。对任意连续线性泛函 $f \in H^*$, 存在唯一的 $y \in H$, 使得

$$f(x) = \langle x, y \rangle, \quad \forall x \in H.$$

因此, Hilbert 空间的连续对偶空间可以通过内积与其自身建立等距同构。从这个意义上说, Hilbert 空间在对偶结构上仍然非常接近有限维内积空间。

3.3 双对偶空间、典范嵌入与自反性

Definition 3.3 (典范映射). 设 X 是赋范线性空间. 其双对偶空间定义为

$$X^{**} = (X^*)^*.$$

典范映射定义为如下线性映射

$$J: X \rightarrow X^{**}, \quad (Jx)(f) = f(x), \quad f \in X^*.$$

由 Hahn-Banach 定理可知, 典范映射 J 满足

$$\|Jx\|_{X^{**}} = \|x\|_X.$$

因此, J 是等距嵌入.

在有限维情形中, 典范映射

$$J: V \rightarrow V^{##}$$

总是双射, 因此有限维线性空间与其双对偶空间存在不依赖于基的典范同构. 这一点与 V 和 $V^\#$ 的关系不同: 虽然 V 、 $V^\#$ 和 $V^{##}$ 的维数都一样, 但是 V 与 $V^\#$ 之间的同构通常需要借助基或内积定义, 而 V 与 $V^{##}$ 的同构是典范的. 这里的典范映射也称为自然映射, 因为它的定义不依赖于选取的基或内积, 也不会随基或内积的选择而改变.

Example 3.5. 设 V 是 n 维线性空间, 取一组基 $\mathcal{E} = (e_1, \dots, e_n)$, 记其对偶基为

$$\mathcal{E}^\# = (\varepsilon^1, \dots, \varepsilon^n), \quad \varepsilon^i(e_j) = \delta_{ij}.$$

利用基和对偶基可定义线性同构 $I_{\mathcal{E}}: V \rightarrow V^\#$:

$$I_{\mathcal{E}}: v = \sum_{i=1}^n x_i e_i \mapsto I_{\mathcal{E}}(v) = \sum_{i=1}^n x_i \varepsilon^i$$

再取另一组基 $\mathcal{E}' = (e'_1, \dots, e'_n)$, 相应的对偶基记作

$$\mathcal{E}'^\# = (\eta^1, \dots, \eta^n), \quad \eta^i(e'_j) = \delta_{ij}.$$

同理可定义线性同构 $I_{\mathcal{E}'}: V \rightarrow V^\#$:

$$I_{\mathcal{E}'}: v = \sum_{i=1}^n y_i e'_i \mapsto I_{\mathcal{E}'}(v) = \sum_{i=1}^n y_i \eta^i$$

下面说明 $I_{\mathcal{E}}$ 和 $I_{\mathcal{E}'}$ 一般并不相同. 记两组基之间的过渡矩阵为 P :

$$\begin{bmatrix} e'_1 & \dots & e'_n \end{bmatrix} = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} P, \quad e'_j = \sum_i e_i P_{ij}$$

两组对偶基之间也由线性变换联系. 设对应的变换矩阵为 A , 即

$$\eta^j = \sum_k A_{jk} \varepsilon^k$$

那么

$$\delta_{ji} = \eta^j(e'_i) = \sum_k A_{jk} \varepsilon^k \left(\sum_\ell e_\ell P_{\ell i} \right) = \sum_{k, \ell} A_{jk} P_{\ell i} \delta_{k\ell} = \sum_k A_{jk} P_{ki}$$

因此 $AP = I_n$, 得到 $A = P^{-1}$, 即

$$\begin{bmatrix} \eta'_1 \\ \vdots \\ \eta'_n \end{bmatrix} = P^{-1} \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{bmatrix}, \quad \eta^j = \sum_k (P^{-1})_{jk} \varepsilon^k$$

对于一个固定的向量 v , 分别在两组基下展开

$$v = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} e'_1 & \dots & e'_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

那么

$$I_{\mathcal{E}}(v) = \begin{bmatrix} \varepsilon^1 & \dots & \varepsilon^n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad I_{\mathcal{E}'}(v) = \begin{bmatrix} \eta^1 & \dots & \eta^n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

代入上述变换关系可知

$$I_{\mathcal{E}'}(v) = \begin{bmatrix} \eta^1 & \dots & \eta^n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \varepsilon^1 & \dots & \varepsilon^n \end{bmatrix} P^{-\top} P^{-1} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

当且仅当基变换满足 $P^{-\top} P^{-1} = I_n$ 时, 对应的两个线性变换相同, 但是这个条件显然并不是恒成立的。因此, 基于对偶基定义的从 V 到 $V^\#$ 的变换不是典范的。

Example 3.6. 设 V 是 n 维线性空间, 取一组基 $\mathcal{E} = (e_1, \dots, e_n)$, 记其对偶基为

$$\mathcal{E}^\# = (\varepsilon^1, \dots, \varepsilon^n), \quad \varepsilon^i(e_j) = \delta_{ij}.$$

考虑 $V^{\#\#}$ 关于 $\mathcal{E}^\#$ 的对偶基, 记作

$$\mathcal{E}^{\#\#} = (\mathbf{e}_1, \dots, \mathbf{e}_n), \quad \mathbf{e}_i(\varepsilon^j) = \delta_{ij}.$$

利用这组基可定义线性同构 $J_{\mathcal{E}}: V \rightarrow V^{\#\#}$:

$$J_{\mathcal{E}} \left(\sum_{i=1}^n x_i e_i \right) = \sum_{i=1}^n x_i \mathbf{e}_i.$$

再取 V 的另一组基 $\mathcal{E}' = (e'_1, \dots, e'_n)$, 两组基之间的过渡矩阵记为 P :

$$\begin{bmatrix} e'_1 & \dots & e'_n \end{bmatrix} = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} P, \quad e'_j = \sum_i e_i P_{ij}.$$

相应的对偶基 $\mathcal{E}'^\# = (\eta^1, \dots, \eta^n)$ 满足

$$\begin{bmatrix} \eta'_1 \\ \vdots \\ \eta'_n \end{bmatrix} = P^{-1} \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{bmatrix}, \quad \eta^j = \sum_k (P^{-1})_{jk} \varepsilon^k$$

$V^{##}$ 关于 $\mathcal{E}'^\#$ 的对偶基记作 $\mathcal{E}'^{##} = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$, 其中 $\mathbf{e}'_i(\eta^j) = \delta_{ij}$ 。类似地定义同构 $J_{\mathcal{E}'}: V \rightarrow V^{##}$:

$$J_{\mathcal{E}'}\left(\sum_{i=1}^n y_i \mathbf{e}'_i\right) = \sum_{i=1}^n y_i \mathbf{e}'_i.$$

下面说明 $J_{\mathcal{E}}$ 与 $J_{\mathcal{E}'}$ 实际上是同一个映射。只需要两次应用从基变换到对偶基的变换关系: 首先是从 V 到 $V^\#$ 的转换

$$\begin{bmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} P \Rightarrow \begin{bmatrix} \eta^1 \\ \vdots \\ \eta^n \end{bmatrix} = P^{-1} \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{bmatrix}.$$

然后是从 $V^\#$ 到 $V^{##}$ 的转换

$$\begin{bmatrix} \eta^1 \\ \vdots \\ \eta^n \end{bmatrix} = P^{-1} \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^n \end{bmatrix}, \Rightarrow \begin{bmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} P$$

对于一个固定的向量 v , 分别在两组基下展开

$$v = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

那么

$$J_{\mathcal{E}}(v) = \sum_i x_i \mathbf{e}_i = \begin{bmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad J_{\mathcal{E}'}(v) = \sum_i y_i \mathbf{e}'_i = \begin{bmatrix} \mathbf{e}'_1 & \dots & \mathbf{e}'_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

因此

$$J_{\mathcal{E}'}(v) = J_{\mathcal{E}}(v), \quad \forall v.$$

这说明两个映射完全相同, 与基的选取无关, 这是一个典范的自然同构。

对于 Banach 空间 X , 典范嵌入

$$J: X \rightarrow X^{**}$$

未必是满射, 即 X^{**} 中可能存在不能写成

$$f \mapsto f(x)$$

形式的连续线性泛函。

Definition 3.4 (自反 Banach 空间). 对于 Banach 空间 X , 称其为自反 Banach 空间, 如果典范嵌入 $J: X \rightarrow X^{**}$ 是满射。

需要注意, 自反性要求的不是任意抽象同构, 而是要求原空间通过典范嵌入恰好填满双对偶空间。

许多常见空间是自反的, 例如

$$\ell^p, \quad 1 < p < \infty,$$

以及更一般的

$$L^p(\Omega), \quad 1 < p < \infty,$$

在适当条件下都是自反 Banach 空间。但是 ℓ^1 和 ℓ^∞ 都不是自反空间。由 Riesz 表示定理易知, 任意 Hilbert 空间都是自反的。

在有限维线性代数中, 对偶空间与原空间具有相同维数, 代数对偶与连续对偶没有区别, 并且原空间与双对偶空间通过典范映射自然同构。而在无穷维泛函分析中, 代数对偶和连续对偶发生分离, 连续对偶空间可能与原空间结构差异很大, 典范嵌入 $X \rightarrow X^{**}$ 也未必满射, 还需要自反性来刻画这种嵌入何时没有遗漏。

4 范数拓扑下紧性的丢失

紧性控制极限过程。在有限维赋范空间中, 闭有界集就是紧集; 许多存在性和收敛性论证都隐含地依赖这一事实。

Theorem 4.1 (有限维空间中的 Heine–Borel 性质). 设 X 是有限维赋范线性空间。则集合 $K \subset X$ 是紧集, 当且仅当 K 是有界闭集。特别地, 单位闭球

$$\overline{B}_1(0) = \{x \in X : \|x\| \leq 1\}$$

是紧集。

因此, 在有限维空间中, 只要一个序列是有界的, 就可以把它放进某个闭球中, 而闭球是紧的。于是由紧性可知, 有界序列必然存在收敛子列。这正是 Bolzano–Weierstrass 性质, 也是有限维中许多极限论证能够直接成立的原因。但是这个性质在无穷维赋范线性空间中不再成立。

Theorem 4.2 (Riesz 紧性定理). 设 X 是赋范线性空间。则 X 的单位闭球在范数拓扑下紧, 当且仅当 X 是有限维空间。

换句话说, 若 X 是无穷维赋范线性空间, 则单位闭球

$$\overline{B}_1(0) = \{x \in X : \|x\| \leq 1\}$$

在范数拓扑下必定不是紧集。因此, 在无穷维赋范线性空间中, 有界闭集未必紧, 有界序列也未必存在范数收敛子列。在无穷维中, 单纯的有界性通常不足以提供强收敛子列。下面是一个标准反例。

Example 4.1 (ℓ^2 中无强收敛子列的有界序列). 在 Hilbert 空间 ℓ^2 中, 设

$$e_n = (0, \dots, 0, 1, 0, \dots),$$

则

$$\|e_n\|_{\ell^2} = 1,$$

所以 $\{e_n\}_{n=1}^\infty$ 是单位球中的有界序列。但是当 $m \neq n$ 时,

$$\|e_m - e_n\|_{\ell^2} = \sqrt{2}.$$

因此, $\{e_n\}_{n=1}^{\infty}$ 不存在范数收敛子列。

从有限维线性代数过渡到泛函分析时, 紧性不再是一个自动性质。在有限维空间中, 有界闭集就是紧集, 有界序列总有收敛子列; 而在无穷维空间中, 单位闭球在范数拓扑下已经不再紧, 有界序列也可能没有范数收敛子列。范数拓扑下的紧性在无穷维中很稀缺。弱拓扑、弱紧性和紧算子等理论, 都是不同层面上弥补强紧性缺失的工具。

5 弱收敛、弱拓扑与弱紧性

弱收敛用连续线性泛函来测试收敛。它比范数收敛弱, 但常常能保留足够的极限信息, 并在自反空间、对偶空间等情形中提供可用的紧性结论。

5.1 强收敛与弱收敛

强收敛就是范数收敛; 弱收敛则只要求所有连续线性泛函下的观测量收敛。二者在有限维中等价, 但在无穷维空间中需要分开处理。

Definition 5.1 (强收敛与弱收敛). 设 X 是赋范线性空间。

- 若

$$\|x_n - x\|_X \rightarrow 0,$$

则称 x_n 强收敛到 x , 也称为范数收敛, 记作

$$x_n \rightarrow x.$$

- 若对任意连续线性泛函 $f \in X^*$, 都有

$$f(x_n) \rightarrow f(x),$$

则称 x_n 弱收敛到 x , 记作

$$x_n \rightharpoonup x.$$

强收敛要求向量本身在范数意义下接近; 弱收敛只要求所有连续线性观测量的取值收敛。因此, 弱收敛是比强收敛更弱的收敛概念。

Proposition 5.1 (强收敛蕴含弱收敛). 设 X 是赋范线性空间。若

$$x_n \rightarrow x,$$

则

$$x_n \rightharpoonup x.$$

Proof. 对任意 $f \in X^*$, 由连续性可得

$$|f(x_n) - f(x)| = |f(x_n - x)| \leq \|f\|_{X^*} \|x_n - x\|_X \rightarrow 0.$$

因此 $x_n \rightharpoonup x$ 。 □

在有限维空间中，弱收敛和强收敛实际上没有区别。

Proposition 5.2 (有限维中强弱收敛等价). 设 X 是有限维赋范线性空间。则

$$x_n \rightarrow x \iff x_n \rightharpoonup x.$$

Proof. 只需要从弱收敛推出强收敛。设 $x_n \rightharpoonup x$ 。取 X 的一组基

$$e_1, \dots, e_d.$$

将

$$x_n - x = \sum_{j=1}^d \alpha_{n,j} e_j.$$

有限维空间中的坐标泛函是连续线性泛函，因此弱收敛推出

$$\alpha_{n,j} \rightarrow 0, \quad j = 1, \dots, d.$$

由于只有有限多个坐标，得到

$$\|x_n - x\|_X \rightarrow 0.$$

因此 $x_n \rightarrow x$ 。 □

无穷维空间中的情况不同。弱收敛一般不能推出强收敛，下面是一个例子。

Example 5.1 (ℓ^2 中弱收敛但不强收敛). 在 Hilbert 空间 ℓ^2 中，令

$$e_n = (0, \dots, 0, 1, 0, \dots)$$

为第 n 个标准基向量。对任意

$$y = (y_j)_{j=1}^\infty \in \ell^2,$$

有

$$\langle e_n, y \rangle_{\ell^2} = \overline{y_n} \rightarrow 0,$$

因为 $y \in \ell^2$ 蕴含 $y_n \rightarrow 0$ 。由 Riesz 表示定理， ℓ^2 上的连续线性泛函都可以由内积表示，因此

$$e_n \rightharpoonup 0.$$

但是

$$\|e_n\|_{\ell^2} = 1, \quad n = 1, 2, \dots,$$

所以

$$e_n \not\rightarrow 0$$

在范数意义下成立。

这个例子说明，在无穷维空间中，弱收敛序列仍然可能在范数意义下不靠近极限。不过，弱收敛虽然比强收敛弱，但它仍然蕴含有界性，这一点在无穷维分析中很重要。

Proposition 5.3 (弱收敛序列有界). 设 X 是 Banach 空间。若

$$x_n \rightharpoonup x,$$

则 $\{x_n\}_{n=1}^{\infty}$ 在 X 中有界。

Proof. 对每个 n , 定义 X^* 上的线性泛函

$$T_n(f) = f(x_n), \quad f \in X^*.$$

则 $T_n \in X^{**}$, 且

$$\|T_n\|_{X^{**}} = \|x_n\|_X.$$

由于 $x_n \rightharpoonup x$, 对任意固定的 $f \in X^*$, 数列

$$T_n(f) = f(x_n)$$

收敛, 因此有界。由一致有界原理,

$$\sup_n \|T_n\|_{X^{**}} < \infty.$$

于是

$$\sup_n \|x_n\|_X < \infty.$$

□

5.2 弱拓扑

弱收敛可以通过弱拓扑来理解。

Definition 5.2 (弱拓扑). 设 X 是赋范线性空间。 X 上的弱拓扑是使所有连续线性泛函

$$f \in X^*$$

都连续的最弱拓扑, 记作

$$\sigma(X, X^*).$$

在这个拓扑下, 序列 x_n 收敛到 x , 正是指

$$f(x_n) \rightarrow f(x), \quad \forall f \in X^*.$$

因此, 弱收敛就是弱拓扑下的收敛。

弱拓扑比范数拓扑更弱。也就是说, 范数拓扑中的开集通常比弱拓扑中的开集更多。因此, 范数收敛一定推出弱收敛, 但弱收敛一般不能推出范数收敛。引入弱拓扑的目的不是为了得到更强的极限, 而是为了在无穷维空间中恢复一部分紧性。范数拓扑太强, 使得无穷维空间中的单位闭球不再紧; 弱拓扑较弱, 因此单位闭球有可能在弱拓扑下变成紧集。

5.3 弱紧性

在有限维空间中, 闭有界集是紧的。因此, 有界序列总能抽取强收敛子列。但在无穷维赋范空间中, 单位闭球在范数拓扑下一般不紧, 所以有界序列通常没有强收敛子列。弱拓扑提供了一种替代: 虽然不能总是抽取强收敛子列, 但在某些空间中可以抽取弱收敛子列。

Theorem 5.4 (自反空间中的弱紧性). 设 X 是 *Banach* 空间. 则 X 是自反空间, 当且仅当单位闭球

$$\overline{B}_X(0) = \{x \in X : \|x\|_X \leq 1\}$$

在弱拓扑 $\sigma(X, X^*)$ 下是紧的。

这个定理给出的是拓扑意义下的弱紧性。但在无穷维空间中，弱拓扑一般不是度量拓扑，因此不能直接由紧性推出序列紧性。为了得到“有界序列存在弱收敛子列”的结论，需要使用 Eberlein–Šmulian 定理。

Theorem 5.5 (Eberlein–Šmulian 定理). 设 X 是 *Banach* 空间, $K \subset X$. 则 K 在弱拓扑下紧, 当且仅当 K 在弱拓扑下序列紧。也就是说, K 弱紧当且仅当 K 中任意序列都存在一个弱收敛子列, 并且其弱极限仍属于 K 。

因此可以得到如下序列形式的结论。

Corollary 5.6 (自反空间中的弱收敛子列). 若 X 是自反 *Banach* 空间, 则 X 中每个有界序列都存在弱收敛子列。

也就是说, 在自反空间中, 有限维中的 Bolzano–Weierstrass 性质在无穷维中可以以弱收敛的形式部分恢复:

$$\text{有界序列} \implies \text{存在弱收敛子列}.$$

Hilbert 空间都是自反空间, 因此, 在 Hilbert 空间中, 每个有界序列都存在弱收敛子列。

Example 5.2. ℓ^2 是 Hilbert 空间, 其有界序列 $\{e_n\}$ 没有任何强收敛子列, 因为

$$\|e_n - e_m\|_{\ell^2} = \sqrt{2}, \quad n \neq m.$$

但是它本身弱收敛到 0:

$$e_n \rightharpoonup 0.$$

这里的自反性是一个关键条件, 如果缺少自反性, 弱紧性就无法保证。

Example 5.3. ℓ^1 不是自反空间, 其有界序列 $\{e_n\}$ 没有弱收敛子列。

Proof. 事实上, 若某个子列 $\{e_{n_k}\}$ 弱收敛到 $x \in \ell^1$, 则对每个坐标泛函可得

$$x_j = 0, \quad j = 1, 2, \dots,$$

因而 $x = 0$ 。但取

$$a = (1, 1, \dots) \in \ell^\infty = (\ell^1)^*,$$

则

$$a(e_{n_k}) = 1, \quad k = 1, 2, \dots,$$

不可能收敛到

$$a(0) = 0.$$

因此 $\{e_n\}$ 没有弱收敛子列。 $\square$

这个例子说明，弱拓扑并不是自动解决无穷维中所有紧性问题。只有在自反空间等合适结构下，弱紧性才能提供有效的替代。

5.4 弱星拓扑

除了弱拓扑，还经常需要考虑弱星拓扑。弱拓扑定义在原空间中，而弱星拓扑则定义在对偶空间上。

Definition 5.3 (弱星拓扑). 设 X 是赋范线性空间。对偶空间 X^* 上的弱星拓扑是由 X 中元素诱导的评价映射给出的拓扑，记作

$$\sigma(X^*, X).$$

也就是说，若 $\{f_\alpha\} \subset X^*$ ，则

$$f_\alpha \xrightarrow{*} f$$

当且仅当

$$f_\alpha(x) \rightarrow f(x), \quad \forall x \in X.$$

我们当然也可以在对偶空间中考虑弱拓扑，即

$$x(f_\alpha) \rightarrow x(f), \quad \forall x \in X^{**}.$$

弱收敛测试的是 X^{**} 中的元素，而弱星收敛只测试 X 中的元素

$$x(f_\alpha) \rightarrow x(f), \quad \forall x \in X \subset X^{**}.$$

X^{**} 通常远大于 X ，因此作为 X^* 上的拓扑，弱星拓扑通常比弱拓扑更弱（更粗）。但是若 X 自反，显然两者一致。

Theorem 5.7 (Banach–Alaoglu 定理). 设 X 是赋范线性空间。则 X^* 的单位闭球

$$\overline{B}_{X^*}(0) = \{f \in X^* : \|f\|_{X^*} \leq 1\}$$

在弱星拓扑 $\sigma(X^*, X)$ 下是紧的。

这个定理不要求 X 自反。它说明：即使原空间没有弱紧性，对偶空间的单位闭球仍然总是在弱星拓扑下紧。

不过需要注意，Banach–Alaoglu 定理给出的是拓扑意义下的紧性。在一般情形下，它不能直接理解为“任意有界序列都有弱星收敛子列”。若 X 可分，则 X^* 的单位闭球在弱星拓扑下可度量化，此时可以得到相应的弱星收敛子列结论。

6 子空间的闭性与投影定理

子空间的闭性决定了极限是否仍留在子空间内，也决定了投影和最佳逼近问题是否有解。在线性代数中这通常不需要单独强调；在赋范空间和 Hilbert 空间中，它是一个真正的条件。

6.1 子空间的闭性

有限维子空间总是闭的。无穷维子空间则可能不是闭集，即使它的代数描述很简单。

Proposition 6.1 (有限维子空间的闭性). 设 X 是赋范线性空间, $M \subset X$ 是有限维线性子空间, 则 M 是闭子空间。

在有限维空间中, 任意线性子空间都是有限维的, 因此自动是闭的。因此, 在有限维情形下讨论子空间时, 通常不需要额外区分“子空间”和“闭子空间”。但是在无穷维空间中, 这一区分是必要的, 无穷维子空间可能不是闭的。

Example 6.1 (ℓ^2 中的非闭子空间). 记

$$c_{00} = \{x = (x_n)_{n=1}^{\infty} : x \text{ 只有有限多个非零分量}\}.$$

则 c_{00} 是 ℓ^2 的线性子空间, 但它在 ℓ^2 范数下不是闭的。

Proof. 事实上, 取

$$x = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right) \in \ell^2.$$

令

$$x^{(N)} = \left(1, \frac{1}{2}, \dots, \frac{1}{N}, 0, 0, \dots\right).$$

则 $x^{(N)} \in c_{00}$, 并且

$$\|x^{(N)} - x\|_{\ell^2}^2 = \sum_{n=N+1}^{\infty} \frac{1}{n^2} \rightarrow 0.$$

因此

$$x^{(N)} \rightarrow x \quad \text{in } \ell^2.$$

但是 $x \notin c_{00}$, 所以 c_{00} 不是 ℓ^2 中的闭子空间。 □

这个例子说明, 在无穷维空间中, 有限线性组合形成的子空间可能通过取极限逼近出原本不属于该子空间的新元素。事实上, c_{00} 在 ℓ^2 中是稠密的, 但不是闭的。这也说明, “稠密”和“闭”是两种不同的性质。一个子空间可以非常大, 大到它的闭包等于整个空间; 但它本身仍然可能不是闭子空间。

6.2 投影定理

子空间闭性的差异会直接影响投影理论。在有限维内积空间中, 任意子空间都可以作正交投影; 但在无穷维 Hilbert 空间中, 正交投影定理要求子空间必须是闭的。

Proposition 6.2 (有限维中的正交分解). 设 V 是有限维内积空间, $M \subset V$ 是线性子空间。则

$$V = M \oplus M^{\perp},$$

其中

$$M^{\perp} = \{v \in V : \langle v, m \rangle = 0, \forall m \in M\}.$$

因此任意 $v \in V$ 都可以唯一分解为

$$v = v_M + v_{M^\perp}, \quad v_M \in M, \quad v_{M^\perp} \in M^\perp.$$

其中 v_M 是 v 在 M 上的正交投影, 并且满足最佳逼近性质

$$\|v - v_M\| = \min_{m \in M} \|v - m\|.$$

有限维情形中不需要额外假设 M 闭, 是因为 M 自动闭。无穷维 Hilbert 空间中的对应版本则需要条件中强调闭子空间。

Theorem 6.3 (Hilbert 空间投影定理). 设 H 是 Hilbert 空间, $M \subset H$ 是闭线性子空间。则任意 $x \in H$ 都存在唯一分解

$$x = m + n, \quad m \in M, \quad n \in M^\perp.$$

等价地,

$$H = M \oplus M^\perp.$$

进一步, 存在唯一的有界线性算子

$$P_M : H \rightarrow M,$$

使得

$$x - P_M x \in M^\perp.$$

这个算子称为 H 到 M 上的正交投影。并且 $P_M x$ 是 x 在 M 中的最佳逼近:

$$\|x - P_M x\| = \min_{m \in M} \|x - m\|.$$

闭性条件在投影定理中不可省略。若 M 不是闭子空间, 则一般不能把任意 $x \in H$ 正交投影到 M 上。此时虽然仍然可以定义正交补

$$M^\perp = \{x \in H : \langle x, m \rangle = 0, \forall m \in M\},$$

但真正成立的是

$$H = \overline{M} \oplus M^\perp,$$

也就是说, Hilbert 空间中的正交分解对应的是闭子空间 $\overline{M}$, 而不是非闭子空间 M 本身。

Example 6.2 (非闭子空间上最佳逼近可能不存在). 仍然考虑 $c_{00} \subset \ell^2$ 。取

$$x = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right) \in \ell^2.$$

如上所述, $x \in \overline{c_{00}}$, 但 $x \notin c_{00}$ 。因此

$$\inf_{m \in c_{00}} \|x - m\|_{\ell^2} = 0.$$

但是不存在 $m \in c_{00}$ 使得

$$\|x - m\|_{\ell^2} = 0.$$

因为若 $\|x - m\|_{\ell^2} = 0$, 则 $x = m$, 这与 $x \notin c_{00}$ 矛盾。因此, x 到 c_{00} 的范数下确界等于 0, 但这个下确界不能在 c_{00} 中达到。也就是说, x 在 c_{00} 中没有最佳逼近。

这个现象在有限维空间中不会发生。有限维中, 子空间闭、有界闭集紧, 因此最小化问题通常可以通过正交投影或紧性得到解; 而在无穷维空间中, 若子空间不是闭的, 最佳逼近可能只有下确界而没有最小值。

在一般 Banach 空间中, 情况还要更复杂。由于 Banach 空间没有内积结构, 通常不能直接谈论正交补。此时若希望得到类似

$$X = M \oplus N$$

的分解, 需要寻找某个闭子空间 N 作为 M 的补空间。等价地, 需要存在一个有界投影算子

$$P : X \rightarrow M, \quad P^2 = P.$$

若这样的有界投影存在, 则称 M 是 X 中的补子空间。但是, 在一般 Banach 空间中, 即使 M 是闭子空间, 也未必存在这样的有界投影。因此需要区分三个层次:

$$\text{子空间} \supset \text{闭子空间} \supset \text{有补闭子空间}.$$

闭性和可补性是不同性质。Hilbert 空间中的闭子空间有正交补; 一般 Banach 空间中的闭子空间则未必有有界补投影。

第二部分 进阶篇

7 线性算子的连续性和收敛性

7.1 线性算子的连续性

线性算子是否连续，取决于定义域和值域上的范数。有限维中这一点常被范数等价性掩盖；在无穷维中，同一个公式放在不同空间之间，可能得到完全不同的连续性结论。

Definition 7.1 (线性算子的连续性). 设 X, Y 是赋范线性空间, $T: X \rightarrow Y$ 是线性算子。若对任意 $x \in X$ 和任意满足

$$x_n \rightarrow x \quad \text{in } X$$

的序列 $\{x_n\}_{n=1}^{\infty} \subset X$, 都有

$$Tx_n \rightarrow Tx \quad \text{in } Y,$$

则称 T 是连续线性算子。

对线性算子来说，连续性等价于有界性，可以用一个全局范数估计刻画。

Definition 7.2 (线性算子的有界性). 设 X, Y 是赋范线性空间, $T: X \rightarrow Y$ 是线性算子。若存在常数 $C > 0$, 使得

$$\|Tx\|_Y \leq C\|x\|_X, \quad \forall x \in X.$$

那么称 T 是有界线性算子。此时 T 把 X 中的有界集映成 Y 中的有界集。定义其算子范数为

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y.$$

Proposition 7.1 (线性算子连续性的等价刻画). 设 X, Y 是赋范线性空间, $T: X \rightarrow Y$ 是线性算子。则下列性质等价：

$$T \text{ 连续} \iff T \text{ 在 } 0 \text{ 点连续} \iff T \text{ 有界}.$$

在有限维空间中，线性算子的连续性自动成立。

Proposition 7.2 (有限维线性算子必然连续). 设 V 是有限维赋范线性空间, Y 是赋范线性空间, $T: V \rightarrow Y$ 是线性算子。则 T 是连续线性算子。

Proof. 取 V 的一组基

$$e_1, \dots, e_d.$$

任意 $v \in V$ 可写为

$$v = \sum_{j=1}^d \alpha_j e_j.$$

由于有限维空间中所有范数等价, 存在常数 $C_1 > 0$, 使得

$$\sum_{j=1}^d |\alpha_j| \leq C_1 \|v\|_V.$$

因此

$$\|Tv\|_Y = \left\| \sum_{j=1}^d \alpha_j T e_j \right\|_Y \leq \sum_{j=1}^d |\alpha_j| \|T e_j\|_Y \leq C_1 \left(\max_j \|T e_j\|_Y \right) \|v\|_V.$$

所以 T 有界, 从而连续。 $\square$

因此, 在有限维线性代数中, 通常不需要区分线性算子和连续线性算子, 也不需要额外检查算子范数是否有限。但是这个性质依赖有限维, 在无穷维空间中, 线性算子并不是必然连续的。

Example 7.1 (微分算子的连续性依赖于范数). 考虑空间 $C^1[0, 1]$, 并先赋予一致范数

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

定义微分算子

$$Df = f'.$$

若把 D 看作

$$D : (C^1[0, 1], \|\cdot\|_\infty) \rightarrow (C[0, 1], \|\cdot\|_\infty),$$

则 D 不是有界线性算子。

取

$$f_n(x) = \frac{1}{n} \sin(n^2 x). \quad \|f_n\|_\infty \leq \frac{1}{n} \rightarrow 0.$$

但是

$$f'_n(x) = n \cos(n^2 x), \quad \|Df_n\|_\infty = \|f'_n\|_\infty = n \rightarrow \infty.$$

所以 D 不可能满足

$$\|Df\|_\infty \leq C \|f\|_\infty.$$

因而 D 不是有界线性算子, 也不是连续线性算子。

但如果给 $C^1[0, 1]$ 配备 C^1 范数

$$\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty,$$

那么

$$D : (C^1[0, 1], \|\cdot\|_{C^1}) \rightarrow (C[0, 1], \|\cdot\|_\infty)$$

就是有界线性算子, 因为

$$\|Df\|_\infty = \|f'\|_\infty \leq \|f\|_{C^1}.$$

这个例子说明, 在无穷维空间中, 一个算子是否连续不能只看它的公式, 还必须说明它作用在哪两个赋范空间之间, 以及这些空间分别配备什么范数。

从这个角度看, 无穷维算子理论不再只是矩阵理论的直接推广。在有限维空间中, 线性算子自动连续; 而在无穷维空间中, 线性算子严格分为有界算子和无界算子。对于有界线性算子, 可以发展与矩阵理论非常类似的算子范数、谱理论、紧算子理论等内容; 但是 PDE 中关注的微分算子、偏微分算子等往往是无界算子, 为了研究这些算子, 还需要额外讨论其定义域、闭性、稠密性和谱性

质。因此，从有限维线性代数过渡到泛函分析时，一个基本变化是：线性算子不再自动具有良好的拓扑性质。连续性、有界性以及定义域的选择，都必须成为讨论的一部分。

7.2 线性算子的收敛性

前面主要讨论的收敛性是向量的收敛。对于有界线性算子组成的序列，还需要区分：是在单位球上统一逼近，还是对每个固定输入分别逼近。

Definition 7.3 (算子范数收敛与强算子收敛). 设 X, Y 是赋范空间, $T_n, T \in \mathcal{L}(X, Y)$ 。

1. 若

$$\|T_n - T\|_{\mathcal{L}(X, Y)} = \sup_{\|x\|_X \leq 1} \|T_n x - T x\|_Y \rightarrow 0,$$

则称 T_n 在算子范数下收敛到 T 。

2. 若对每个固定的 $x \in X$ ，都有

$$\|T_n x - T x\|_Y \rightarrow 0,$$

则称 T_n 强算子收敛到 T ，记作

$$T_n \xrightarrow{\text{strong}} T.$$

这也称为算子列的逐点收敛；等价地，每个向量序列 $T_n x$ 都在 Y 中强收敛到 $T x$ 。

这里的“强算子收敛”不等于把 $\mathcal{L}(X, Y)$ 本身作为赋范空间时的范数收敛。强算子收敛允许逼近速度依赖于输入 x ；算子范数收敛则要求对单位球中的所有输入统一控制误差。

Proposition 7.3. 算子范数收敛蕴含强算子收敛。

Proof. 对每个固定的 $x \in X$ ，有

$$\|T_n x - T x\|_Y \leq \|T_n - T\|_{\mathcal{L}(X, Y)} \|x\|_X \rightarrow 0.$$

□

Example 7.2 (有限维截断强算子收敛，但不按算子范数收敛). 在 ℓ^2 上定义前 n 个坐标的截断投影

$$P_n(x_1, x_2, \dots) = (x_1, \dots, x_n, 0, 0, \dots).$$

对每个固定的 $x \in \ell^2$ ，有

$$\|P_n x - x\|_{\ell^2}^2 = \sum_{j>n} |x_j|^2 \rightarrow 0,$$

所以 $P_n \xrightarrow{\text{strong}} I$ 。但是取标准基向量 e_{n+1} ，有

$$\|(I - P_n)e_{n+1}\|_{\ell^2} = 1, \quad \|I - P_n\|_{\mathcal{L}(\ell^2)} = 1.$$

因此 P_n 不在算子范数下收敛到 I 。

若定义域 X 是有限维的, 强算子收敛也蕴含算子范数收敛: 只需控制有限个基向量的像, 再利用坐标范数等价性即可。上面的截断例子说明, 二者在无穷维中会分离。

8 线性方程的可解性、闭值域与适定性

线性方程 $Ax = b$ 的代数可解性由线性算子 A 的核和值域决定:

1. 若 $b \in \text{Ran}(A)$, 则解存在; 反之, 若 $b \notin \text{Ran}(A)$, 则解不存在;
2. 若 A 单射, 则解唯一;
3. 若 A 双射, 则对任意右端 b 都存在唯一解。

在有限维线性代数中, 上述结果已经足够完整。在无穷维 Banach 空间中, 虽然上述结论仍然成立, 但是无穷维会带来一些新问题, 例如: 值域 $\text{Ran}(A)$ 可能不是闭的, 逆映射也可能不是有界的 (有界逆定理表明, 这两个条件是等价的)。此时会出现一种在有限维中不会出现的现象:

$$b \in \overline{\text{Ran}(A)} \quad \text{and} \quad b \notin \text{Ran}(A).$$

这表示

$$\inf_{x \in X} \|Ax - b\|_Y = 0,$$

也就是说, 可以找到一列 $x_n \in X$, 使得

$$Ax_n \rightarrow b.$$

但是并不存在真正的 $x \in X$ 使得

$$Ax = b.$$

注意这里的“可逼近”并不意味着问题在数值意义下是良好的。相反, 它对于数值求解是一个糟糕的现象: 这说明即使得到了一组残差任意小的所谓逼近解 $\{x_n\}$, 但实际上它们并不会收敛到所谓的解, 甚至 $\{x_n\}$ 本身并不会收敛。

Example 8.1. 定义

$$A: \ell^2 \rightarrow \ell^2, \quad (Ax)_n = \frac{x_n}{n}.$$

则 A 是有界线性单射。事实上,

$$\|Ax\|_{\ell^2}^2 = \sum_{n=1}^{\infty} \frac{|x_n|^2}{n^2} \leq \sum_{n=1}^{\infty} |x_n|^2 = \|x\|_{\ell^2}^2.$$

它的值域为

$$\text{Ran}(A) = \{y = (y_n)_{n=1}^{\infty} \in \ell^2 : (ny_n)_{n=1}^{\infty} \in \ell^2\}.$$

由于有限支撑序列都属于 $\text{Ran}(A)$, 而有限支撑序列在 ℓ^2 中稠密, 所以

$$\overline{\text{Ran}(A)} = \ell^2.$$

但是 $\text{Ran}(A) \neq \ell^2$ 。例如

$$y = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right) \in \ell^2.$$

若 $y = Ax$, 则必须有

$$x_n = ny_n = 1, \quad n = 1, 2, \dots,$$

由于 $x = (1, 1, \dots) \notin \ell^2$, 因此 $y \notin \text{Ran}(A)$ 。所以值域虽然是稠密的, 但是仍然不是闭的

$$\text{Ran}(A) \subsetneq \overline{\text{Ran}(A)} = \ell^2.$$

取 $x^{(N)} = e_N$, 则

$$\|x^{(N)}\|_{\ell^2} = 1, \quad Ax^{(N)} = \frac{1}{N}e_N, \quad \|Ax^{(N)}\|_{\ell^2} = \frac{1}{N} \rightarrow 0.$$

因此不可能存在常数 $C > 0$, 使得

$$\|x\|_{\ell^2} \leq C\|Ax\|_{\ell^2}, \quad \forall x \in \ell^2.$$

这表明, 对线性方程 $Ax = b$ 的右端施加小扰动, 即使问题始终可解, 相应的解的扰动也可能是不可控的。

对于无穷维空间中的线性方程组, 显然只讨论解的存在唯一性是不够的, 还需要新的要求——稳定性: 要求解连续依赖于右端数据。换言之, 若 $Ax = b$ 的右端产生小扰动, 则对应的解 x 的扰动必须很小。

Definition 8.1 (稳定性). 设 X, Y 是赋范空间, $A : X \rightarrow Y$ 是单射线性算子。称方程 $Ax = b$ 在数据空间 $\text{Ran}(A)$ 上是稳定的, 如果存在常数 $C > 0$, 使得

$$\|x_1 - x_2\|_X \leq C\|Ax_1 - Ax_2\|_Y, \quad x_1, x_2 \in X.$$

等价地, $A^{-1} : \text{Ran}(A) \rightarrow X$ 是有界线性算子。

取 $x' = x_1 - x_2$, 上述条件也可以等价改写为

$$\|x'\|_X \leq C\|Ax'\|_Y, \quad x' \in X.$$

Proposition 8.1 (闭值域与有界逆). 设 X, Y 是 Banach 空间, $A : X \rightarrow Y$ 是有界线性算子, 并且是单射。则以下两个条件等价:

$$\text{Ran}(A) \text{ 是 } Y \text{ 中的闭子空间,} \iff A^{-1} : \text{Ran}(A) \rightarrow X \text{ 是有界线性算子.}$$

Proof. 若 $\text{Ran}(A)$ 是闭的, 则 $\text{Ran}(A)$ 按照 Y 中继承的范数本身是 Banach 空间。由于

$$A : X \rightarrow \text{Ran}(A)$$

是有界线性双射, 由有界逆定理可知

$$A^{-1} : \text{Ran}(A) \rightarrow X$$

有界。

反过来, 假设 A^{-1} 有界。取任意收敛序列

$$y_n \in \text{Ran}(A), \quad y_n \rightarrow y \quad \text{in } Y.$$

对每个 n , 令 $y_n = Ax_n$ 。由于 A^{-1} 有界,

$$\|x_n - x_m\|_X = \|A^{-1}(y_n - y_m)\|_X \leq C\|y_n - y_m\|_Y.$$

因此 $\{x_n\}$ 是 X 中的 Cauchy 列。由 X 的完备性, 存在 $x \in X$, 使得

$$x_n \rightarrow x.$$

又因为 A 连续,

$$Ax_n \rightarrow Ax.$$

另一方面, $Ax_n = y_n \rightarrow y$, 故 $y = Ax \in \text{Ran}(A)$ 。因此 $\text{Ran}(A)$ 是闭的。 $\square$

Definition 8.2 (适定性). 设 X, Y 是赋范空间, $A : X \rightarrow Y$ 是线性算子。称线性方程

$$Ax = b$$

在数据空间 Y 上适定, 如果对每个 $b \in Y$ 都存在唯一解 $x \in X$, 并且是稳定的。

对于 Banach 空间之间的有界线性算子 $A : X \rightarrow Y$, 若只把 b 限制在 $\text{Ran}(A)$ 中, 则 A 单射保证解唯一, 闭值域条件进一步保证问题在 $\text{Ran}(A)$ 上稳定。若要求问题在整个 Y 上适定, 则还必须要求 A 满射; 此时 A 是有界线性双射, 有界逆定理自动给出 $A^{-1} : Y \rightarrow X$ 的有界性。因此, 适定性实际就是要求 A 是双射且有界。

9 有限秩算子与紧算子

在有限维空间之间的线性算子存在很多好的性质: 始终把有限维空间映射到有限维空间, 始终把有界集映射到相对紧集。在无穷维情形下, 这些性质不再自动成立, 但是自然诱导出仍然满足这些性质的算子概念。

Definition 9.1 (有限秩算子). 设 X, Y 是线性空间, $T : X \rightarrow Y$ 是线性算子。称 T 是有限秩算子, 若

$$\dim \text{Ran}(T) < \infty,$$

有限秩算子保证输出始终落在一个有限维子空间中, 因此它是最接近矩阵的概念, 保留了有限维矩阵的主要特征。显然, 对于有限维情形, 所有的线性算子都是有限秩算子。

Example 9.1. 典型的有限秩算子形如

$$Tx = \sum_{j=1}^m f_j(x)y_j, \quad f_j \in X^\#, y_j \in Y$$

此时

$$\text{Ran}(T) \subset \text{span}\{y_1, \dots, y_m\}.$$

有限秩算子的定义是纯代数的, 因此并不能保证算子的有界性, 通常明确讨论有界有限秩算子。

Example 9.2 (有限秩算子不一定有界). 设 $X = c_{00}$, 即只有有限项非零的实数列全体, 配备上确界范数 $\|x\|_{\infty} = \sup_n |x_n|$. 定义 $T: X \rightarrow X$ 为

$$T(x_1, x_2, x_3, \dots) = \left(\sum_{k=1}^{\infty} kx_k, 0, 0, \dots \right).$$

由于输入序列只有有限项非零, 该求和总是有限和, 因此定义合理. T 的值域包含在第一个坐标的一维子空间中, 故 $\dim \text{Ran}(T) = 1$, 是有限秩算子. 取 e_n 为第 n 个标准单位向量, 则 $\|e_n\|_{\infty} = 1$, 而

$$Te_n = (n, 0, 0, \dots) \Rightarrow \|Te_n\|_{\infty} = n \rightarrow \infty,$$

故 T 无界.

我们还要引入一类具有某种有限维紧性特征的算子, 称为紧算子.

Definition 9.2 (紧算子). 设 X, Y 是赋范线性空间, $T: X \rightarrow Y$ 是线性算子. 若 T 把 X 中的有界集映成 Y 中的相对紧集, 则称 T 是紧算子. 等价地, 若对任意有界序列 $\{x_n\}_{n=1}^{\infty} \subset X$, 序列 $\{Tx_n\}_{n=1}^{\infty} \subset Y$ 都存在范数收敛子列, 则 T 是紧算子. 紧算子空间记作 $\mathcal{C}(X, Y)$.

下面的命题会保证紧算子全体组成的空间是一个闭空间, 即一系列紧算子的极限也是紧算子.

Proposition 9.1 (紧算子的算子范数闭性). 设 X 是赋范线性空间, Y 是 Banach 空间. 则紧算子空间 $\mathcal{C}(X, Y)$ 是 $\mathcal{L}(X, Y)$ 在算子范数下的闭子空间. 换言之, 若一系列的紧算子 $\{T_m\} \subset \mathcal{C}(X, Y)$ 在算子范数下收敛到 $T \in \mathcal{L}(X, Y)$,

$$\|T_m - T\|_{\mathcal{L}(X, Y)} \rightarrow 0,$$

则 T 也是紧算子.

Proof. 下面只证明闭性. 设 $\{x_k\}_{k=1}^{\infty} \subset X$ 是任意有界序列, 由于 T_1 是紧算子, 可以从 x_k 中抽取一个子列, 使得该子列在 T_1 作用下收敛. 再从这个子列中抽取一个更小的子列, 使其在 T_2 作用下收敛. 如此继续, 可以得到一族逐层嵌套的子列. 取其对角子列 $\{z_k\}$, 则对每个固定的 m , 序列 $\{T_m z_k\}_{k=1}^{\infty}$ 都收敛, 因而是 Cauchy 序列. 用 $\varepsilon - \delta$ 语言容易证明 $\{T z_k\}_{k=1}^{\infty}$ 是 Banach 空间 Y 的 Cauchy 序列, 由于 Y 完备, $\{T z_k\}_{k=1}^{\infty}$ 在 Y 中收敛. 因此, 任意有界序列经过 T 作用后都存在范数收敛子列, 所以 T 是紧算子. $\square$

下面的两个命题表明, 有界有限秩算子、紧算子和有界算子存在包含关系.

Proposition 9.2 (紧算子的有界性). 紧算子一定是有界线性算子.

Proof. 设 X, Y 是赋范线性空间, $T: X \rightarrow Y$ 是紧算子. 若 T 无界, 则存在序列 $\{x_n\}_{n=1}^{\infty} \subset X$ 满足 $\|x_n\| = 1$, 但 $\|Tx_n\| \rightarrow \infty$. 此时 $\{x_n\}$ 是 X 中的有界序列, 但 $\{Tx_n\}$ 的任何子列都不可能范数收敛 (因为范数趋于无穷), 这与 T 是紧算子矛盾. 因此 T 必有界. $\square$

Proposition 9.3 (有限秩算子的紧性). 有界有限秩算子一定是紧算子.

Proof. 设 $T: X \rightarrow Y$ 是有界线性算子且 $\dim \text{Ran}(T) < \infty$ 。任取 X 中的有界集 B ，由于 T 有界， $T(B)$ 是 Y 中的有界集。又因为 $\text{Ran}(T)$ 是有限维子空间且 $T(B) \subset \text{Ran}(T)$ ，故 $T(B)$ 包含在有限维赋范空间 $\text{Ran}(T)$ 中。有限维赋范空间中的有界闭集是紧集，从而有界集相对紧，因此 $T(B)$ 在 $\text{Ran}(T)$ 中相对紧，进而在 Y 中也相对紧。由紧算子的定义， T 是紧算子。 $\square$

因此，存在如下包含关系

有界有限秩算子 $\subset$ 紧算子 $\subset$ 有界线性算子。

在有限维空间中，这些概念是统一的：所有线性算子 T 都是有界的，都是有限秩算子，都是紧算子。无穷维空间是导致这些概念发生分裂的原因，并且此时的包含关系通常是严格的。

Example 9.3 (有界算子不一定是紧算子). 若 X 是无穷维 Banach 空间，则恒等算子

$$I: X \rightarrow X$$

是有界线性算子，但不是紧算子。否则， I 会把单位闭球映成相对紧集，也就是说单位闭球本身相对紧。这与 Riesz 紧性定理矛盾。

Example 9.4 (紧算子不一定是有限秩算子). 考虑

$$T: \ell^2 \rightarrow \ell^2, \quad (Tx)_n = \frac{x_n}{n}.$$

首先， T 是有界线性算子，因为

$$\|Tx\|_{\ell^2}^2 = \sum_{n=1}^{\infty} \frac{|x_n|^2}{n^2} \leq \sum_{n=1}^{\infty} |x_n|^2 = \|x\|_{\ell^2}^2.$$

因而

$$\|T\|_{\mathcal{L}(\ell^2, \ell^2)} \leq 1.$$

其次， T 不是有限秩算子，因为

$$Te_n = \frac{1}{n}e_n, \quad n = 1, 2, \dots,$$

这些像向量线性无关，可以张成无穷维子空间，因此 $\text{Ran}(T)$ 是无穷维的。

我们用紧算子全体的闭性来证明 T 是紧算子：取

$$T_N x = \left(x_1, \frac{x_2}{2}, \dots, \frac{x_N}{N}, 0, \dots \right),$$

则每个 T_N 的值域包含在 $\text{span}\{e_1, \dots, e_N\}$ 中，因此 T_N 是有界有限秩算子，自然也是紧算子。注意到

$$\|T - T_N\|_{\mathcal{L}(\ell^2, \ell^2)} = \sup_{n > N} \frac{1}{n} = \frac{1}{N+1} \rightarrow 0 \quad (N \rightarrow \infty).$$

像空间 ℓ^2 显然完备，因此 T_N 的极限 T 也是紧算子。

在无穷维空间中，有界算子的概念太大，有限秩算子的概念太小，紧算子则构成一个介于二者之间的重要类别。它们不再是有限维矩阵本身，但保留了有限维紧性的某些影子：有界序列经过紧算子作用后，可以抽取强收敛子列。紧性本身不是代数性质，而是拓扑性质；它依赖于范数拓扑下的相对紧性。从谱理论中也可以体现，紧算子在谱结构上也比一般有界算子更接近有限维矩阵。

10 特征值与谱理论

在有限维线性代数中, 算子不可逆性完全由特征值刻画, 无穷维情形则不同, 我们关注有界逆的存在性, 需要进一步引入谱的概念。下面始终设 X 是复 Banach 空间, $A: X \rightarrow X$ 是有界线性算子。这里只讨论复 Banach 空间到自身的算子, 因为要使用到恒等算子。

Definition 10.1 (特征值). 对于线性空间 V , 若存在非零向量 $v \in V$ 及系数 $\lambda \in \mathbb{F}$, 使得

$$Av = \lambda v,$$

则称 λ 为 A 的特征值, v 为对应的特征向量。

在有限维空间中, $A - \lambda I$ 不可逆当且仅当它有非零核。无穷维空间中, 我们关注有界逆算子的存在性, $A - \lambda I$ 不存在有界逆算子有很多原因, 因此引入更广的谱概念。

Definition 10.2 (谱). 设 X 是复 Banach 空间, $A: X \rightarrow X$ 是有界线性算子。称 $\lambda \in \mathbb{C}$ 属于 A 的谱, 如果算子

$$A - \lambda I: X \rightarrow X$$

不是有界可逆的 (即不存在定义在整个 X 上的有界逆算子)。 A 的谱记作 $\sigma(A)$ 。

可以按照 $A - \lambda I$ 失去有界逆算子的原因进行进一步分类。

Definition 10.3 (点谱、连续谱与剩余谱). 设 X 是复 Banach 空间, $A: X \rightarrow X$ 是有界线性算子。

1. 若 $\ker(A - \lambda I) = \{0\}$, 对值域进行区分
 - (a) 若 $\text{Ran}(A - \lambda I) = X$, 那么 $A - \lambda I$ 存在有界逆算子。
 - (b) 若 $\text{Ran}(A - \lambda I) \neq X$, 但是在 X 中稠密, 则称 λ 属于连续谱。
 - (c) 若 $\text{Ran}(A - \lambda I) \neq X$, 且在 X 中不稠密, 则称 λ 属于剩余谱。
2. 若 $\ker(A - \lambda I) \neq \{0\}$, 即 $A - \lambda I$ 存在非零核, 则称 λ 属于点谱, 记作 $\lambda \in \sigma_p(A)$ 。这正是通常意义下的特征值。

对于有限维空间, 由于 $\ker(A - \lambda I) = \{0\}$ 直接等价于 $\text{Ran}(A - \lambda I) = X$, 因此只存在两种情况: 要么存在有界逆算子, 要么是点谱 (即特征值)。

Definition 10.4 (预解集和预解算子). 设 X 是复 Banach 空间, $A: X \rightarrow X$ 是有界线性算子。谱 $\sigma(A)$ 的补集称为预解集, 记作

$$\rho(A) = \mathbb{C} \setminus \sigma(A).$$

根据谱的定义可知, 若 $\lambda \in \rho(A)$, 则可以定义有界算子

$$R(\lambda, A) = (A - \lambda I)^{-1}$$

称其为 A 在 λ 处的预解算子。这里采用的是 $(A - \lambda I)^{-1}$ 约定; 有些教材也使用 $(\lambda I - A)^{-1}$,

二者只相差一个负号。

Example 10.1 (一个谱严格大于特征值集合的例子). 考虑 ℓ^2 上的对角算子

$$A : \ell^2 \rightarrow \ell^2, \quad (Ax)_n = \frac{x_n}{n}.$$

对每个 $k \geq 1$, 有

$$Ae_k = \frac{1}{k}e_k,$$

因此

$$\frac{1}{k} \in \sigma_p(A).$$

另一方面, 0 不是特征值, 因为 A 是单射。但是 0 仍然属于谱。事实上, $A^{-1} : \text{Ran}(A) \rightarrow \ell^2$ 无界; 这可由

$$x^{(N)} = e_N, \quad \|x^{(N)}\|_{\ell^2} = 1, \quad \|Ax^{(N)}\|_{\ell^2} = \frac{1}{N} \rightarrow 0$$

说明。因此 A 没有有界逆, 故

$$0 \in \sigma(A).$$

但是

$$0 \notin \sigma_p(A).$$

实际上,

$$\sigma(A) = \left\{ \frac{1}{n} : n = 1, 2, \dots \right\} \cup \{0\}.$$

这个例子说明, 在无穷维空间中, 谱可以包含不是特征值的点。

这个例子也展示了无穷维谱理论的一个基本现象: 特征值可能有聚点, 而聚点本身不一定是特征值。例如非零特征值序列

$$1, \frac{1}{2}, \frac{1}{3}, \dots$$

趋于 0, 而 0 是谱点但不是特征值。这是有限维中不存在的结构, 因为有限维算子只有有限个特征值。

对于一般有界算子, 谱结构可能远比有限维矩阵复杂。但对紧算子, 仍然保留了较多有限维特征值理论的影子。

Theorem 10.1 (紧算子的非零谱). 设 X 是复无穷维 Banach 空间, $A : X \rightarrow X$ 是紧线性算子。则 A 的每个非零谱点都是特征值。并且每个非零特征值对应的特征子空间有限维; 非零特征值至多可数, 只能以 0 为可能的聚点。

这个定理说明, 紧算子虽然作用在无穷维空间上, 但其非零谱仍然很接近有限维矩阵的特征值结构。真正体现无穷维差异的是 0: 它往往是谱点, 并且可能不是特征值。

11 酉算子、正规算子与谱定理

有限维复内积空间中, 正规矩阵可以酉对角化。把这一结论推广到 Hilbert 空间时, 需要区分两种情形: 紧正规算子仍然有由特征向量组成的正交分解; 一般有界正规算子则未必有特征向量,

此时对角矩阵要由投影值测度取代。

以下均设 H 是复 Hilbert 空间, 并简记 $\mathcal{L}(H) := \mathcal{L}(H, H)$ 。对 $T \in \mathcal{L}(H)$, 其伴随算子记为 T^* 。

11.1 酉算子与正规算子

Definition 11.1 (线性等距映射). 设 H_1, H_2 是 Hilbert 空间, $V : H_1 \rightarrow H_2$ 是线性映射。如果

$$\|Vx\|_{H_2} = \|x\|_{H_1}, \quad \forall x \in H_1,$$

则称 V 是线性等距映射。由极化恒等式, 这等价于 V 保持内积:

$$\langle Vx, Vy \rangle_{H_2} = \langle x, y \rangle_{H_1}, \quad \forall x, y \in H_1.$$

Definition 11.2 (酉算子与正规算子). 算子 $U \in \mathcal{L}(H)$ 称为酉算子, 如果

$$U^*U = UU^* = I.$$

算子 $T \in \mathcal{L}(H)$ 称为正规算子, 如果

$$T^*T = TT^*.$$

酉算子的定义与酉矩阵完全相同: 条件 $U^*U = I$ 说明 U 保持内积和范数, 而 $UU^* = I$ 保证 U 是满射。因此

$$U \text{ 酉} \iff U \text{ 是满射的线性等距映射},$$

并且 $U^{-1} = U^*$ 。根据定义可知, 每个酉算子都是正规算子。自伴算子 $T = T^*$ 也是正规算子, 但酉性与自伴性互不包含。

在有限维中, 线性等距映射自动是满射, 所以仅验证 $U^*U = I$ 即可。在无穷维中这一点不再成立。

Example 11.1 (不是酉算子的单边移位). 在 $\ell^2(\mathbb{N})$ 上定义

$$S(x_1, x_2, \dots) = (0, x_1, x_2, \dots).$$

对任意 $x, y \in \ell^2(\mathbb{N})$, 按标准内积计算可得

$$\langle Sx, y \rangle = \sum_{n=2}^{\infty} x_{n-1} \overline{y_n} = \sum_{n=1}^{\infty} x_n \overline{y_{n+1}} = \langle x, (y_2, y_3, \dots) \rangle.$$

因此

$$S^*(y_1, y_2, y_3, \dots) = (y_2, y_3, \dots).$$

于是

$$S^*S = I, \quad SS^*(y_1, y_2, \dots) = (0, y_2, y_3, \dots),$$

所以 $SS^* \neq I$ 。算子 S 是线性等距映射, 但不是满射, 因而不是酉算子; 又因为 $S^*S \neq SS^*$, 它也不是正规算子。

在无穷维情况下, 酉算子甚至无法保证特征值和特征向量的存在性。

Example 11.2 (没有特征值的双边移位). 在 $\ell^2(\mathbb{Z})$ 上定义

$$(Ux)_n = x_{n-1}, \quad n \in \mathbb{Z}.$$

对任意 $x, y \in \ell^2(\mathbb{Z})$, 有

$$\langle Ux, y \rangle = \sum_{n \in \mathbb{Z}} x_{n-1} \overline{y_n} = \sum_{m \in \mathbb{Z}} x_m \overline{y_{m+1}}.$$

因此

$$(U^*y)_m = y_{m+1}, \quad m \in \mathbb{Z}.$$

这正是 U 的逆向移位, 所以 $U^* = U^{-1}$, 并且

$$U^*U = UU^* = I.$$

故 U 是酉算子。若 $Ux = \lambda x$ 且 $x \neq 0$, 则酉性给出 $|\lambda| = 1$; 递推关系 $x_{n-1} = \lambda x_n$ 又说明 $|x_n|$ 与 n 无关, 这与 $x \in \ell^2(\mathbb{Z})$ 矛盾。因此 U 没有特征值, 也不存在由特征向量组成的标准正交基。

Proposition 11.1 (正规、自伴与酉算子的谱). 设 $T \in \mathcal{L}(H)$ 是正规算子, 则

$$\max_{\lambda \in \sigma(T)} |\lambda| = \|T\|.$$

正规算子没有剩余谱; 若 $\lambda \neq \mu$ 是 T 的两个特征值, 则相应特征空间相互正交。进一步地,

$$T = T^* \implies \sigma(T) \subset \mathbb{R}, \quad T \text{ 酉} \implies \sigma(T) \subset \{z \in \mathbb{C} : |z| = 1\}.$$

这里的谱点未必是特征值; 酉算子的谱可以是单位圆的无限子集, 甚至没有任何特征值。

Proof. 若 N 正规, 则 $\|Nx\| = \|N^*x\|$ 。对 $N = T - \lambda I$ 使用这一等式可得

$$\ker(T - \lambda I) = \ker(T^* - \bar{\lambda} I).$$

因此 $T - \lambda I$ 单射时, 其值域稠密, 故正规算子没有剩余谱。同一等式还说明 $Tx = \lambda x$ 时 $T^*x = \bar{\lambda}x$, 由此可得不同特征值对应的特征向量正交。

正规性和恒等式 $\|T^*T\| = \|T\|^2$ 给出 $\|T^{2^n}\| = \|T\|^{2^n}$; 结合谱半径公式即得谱的最大模等于 $\|T\|$ 。若 T 自伴, 则对 $\lambda \notin \mathbb{R}$ 有

$$\|(T - \lambda I)x\| \geq |\operatorname{Im} \lambda| \|x\|,$$

对伴随算子使用同一估计可知 $T - \lambda I$ 可逆。若 T 酉, 则当 $|\lambda| > 1$ 或 $|\lambda| < 1$ 时, 分别对 $T - \lambda I$ 或 $T(I - \lambda T^*)$ 使用 Neumann 级数, 均可得到有界逆算子。因此酉算子的谱只能位于单位圆上。 $\square$

11.2 谱定理

Theorem 11.2 (有限维谱定理). 设 $A \in \mathbb{C}^{n \times n}$. 则以下条件等价:

1. A 正规, 即 $A^*A = AA^*$;
2. $\mathbb{C}^n$ 存在一组由 A 的特征向量组成的标准正交基;
3. 存在酉矩阵 Q 和对角矩阵 D , 使得 $A = QDQ^*$.

Remark 11.1. 若 $Ae_j = \lambda_j e_j$, 其中 $e_1, \dots, e_n$ 是一组标准正交基. 可以在这组基下展开向量, 将 A 的作用的坐标表示为逐坐标乘法

$$(c_1, \dots, c_n) \mapsto (\lambda_1 c_1, \dots, \lambda_n c_n).$$

这组基下的坐标称为谱坐标。

Remark 11.2. 如果从复数域改到实数域, 对应的结论是: 实矩阵 A 可以正交相似对角化当且仅当 A 是对称方阵. 注意这里的要求不是正规矩阵。

在无穷维情形, 紧算子的性质非常接近有限维, 因为在紧算子的谱中: 每个非零谱点都是特征值, 并且每个非零特征值对应的特征子空间有限维; 非零特征值至多可数, 并且只能以 0 为可能的聚点. 因此对紧正规算子的结论仍然接近于正规矩阵。

Theorem 11.3 (紧正规算子的谱定理). 设 $T \in \mathcal{L}(H)$ 是紧正规算子. 那么, 在各个非零特征值 (重数有限, 且个数至多可数) 对应的特征子空间中选取标准正交基, 合并后得到特征系统 $\{(\lambda_j, e_j)\}_{j \in J}$ (同一个特征值按其重数重复出现). 可以将整个空间 H 分解为

$$H = \ker(T) \oplus \overline{\text{span}\{e_j : j \in J\}},$$

并且将算子 T 表示为

$$Tx = \sum_{j \in J} \lambda_j \langle x, e_j \rangle e_j, \quad \forall x \in H.$$

其中级数在 H 的范数意义下收敛。

当 H 可分时, 再取 $\ker(T)$ 的一组标准正交基, 就得到一组由 T 的特征向量组成的标准正交基; 核空间中的向量对应特征值 0. 这与有限维的酉对角化最接近, 但“对角矩阵”成为一列趋于 0 的特征值。

在一般情形下, 直接按照特征值求和的形式要改写为关于谱的积分, 需要用到投影值测度。

Definition 11.3 (投影值测度). 设 $T \in \mathcal{L}(H)$. 定义在 $\sigma(T)$ 的 Borel 集 $\mathcal{B}(\sigma(T))$ 上的映射 E 称为投影值测度, 如果满足: 对于任意 $\Delta \in \mathcal{B}(\sigma(T))$, $E(\Delta)$ 都是 H 上的正交投影, 并且

$$E(\emptyset) = 0, \quad E(\sigma(T)) = I, \quad E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2).$$

此外, 对任意两两不交的 Borel 集列 $\{\Delta_k\}$

$$E\left(\bigcup_{k=1}^{\infty} \Delta_k\right) = \sum_{k=1}^{\infty} E(\Delta_k).$$

Theorem 11.4 (有界正规算子的谱定理). 设 $T \in \mathcal{L}(H)$ 。那么 T 正规当且仅当在 $\sigma(T)$ 上存在唯一的投影值测度 E , 使得

$$T = \int_{\sigma(T)} z \, dE(z).$$

Example 11.3 (有限维谱积分). 对正规矩阵 A , 记 $\lambda_1, \dots, \lambda_m$ 是 A 的不同特征值, P_j 是到相应特征子空间的正交投影, 则对角化也可以写成

$$A = \sum_{j=1}^m \lambda_j P_j, \quad P_i P_j = 0 \ (i \neq j), \quad \sum_{j=1}^m P_j = I.$$

可以用投影值测度的概念重写: 投影值测度集中在有限个特征值上

$$E(\{\lambda_j\}) = P_j$$

因此谱积分变成了有限和

$$A = \int_{\sigma(A)} z \, dE(z) = \sum_{j=1}^m \lambda_j P_j.$$

Example 11.4 (乘法算子的谱积分). 在 $L^2([0, 1])$ 上定义乘法算子

$$(Mf)(t) = tf(t).$$

对任意 $f, g \in L^2([0, 1])$, 有

$$\|Mf\|_2^2 = \int_0^1 t^2 |f(t)|^2 \, dt \leq \|f\|_2^2, \quad \langle Mf, g \rangle = \langle f, Mg \rangle.$$

因此 M 有界且自伴, 并且 $\|M\| = 1$; 后一个等式可由取支撑逐渐集中到 $t = 1$ 附近的单位函数得到。

下面验证 $\sigma(M) = [0, 1]$ 。若 $\lambda \notin [0, 1]$, 则

$$(M - \lambda I)^{-1} f(t) = \frac{f(t)}{t - \lambda}, \quad \|(M - \lambda I)^{-1}\| = \frac{1}{\text{dist}(\lambda, [0, 1])},$$

所以 λ 属于预解集。反之, 对任意 $\lambda \in [0, 1]$, 令

$$I_n = [0, 1] \cap (\lambda - 1/n, \lambda + 1/n), \quad f_n = \frac{\mathbf{1}_{I_n}}{\sqrt{|I_n|}}.$$

则 $\|f_n\|_2 = 1$, 而

$$\|(M - \lambda I)f_n\|_2 \leq \frac{1}{n} \rightarrow 0.$$

因此 $M - \lambda I$ 不存在有界逆, 故 $\lambda \in \sigma(M)$ 。综上, $\sigma(M) = [0, 1]$ 。

相应的谱投影为

$$(E(\Delta)f)(t) = \mathbf{1}_\Delta(t)f(t).$$

因此

$$M = \int_{[0, 1]} t \, dE(t).$$

对每个 $\lambda \in [0, 1]$, 都有 $E(\{\lambda\}) = 0$ 。这里的谱分解不来自特征向量的可数求和, 而由投影值测度描述。

本节的概念关系可以概括如下。线性等距映射保持范数；当定义域和值域都是 H 时，满射的线性等距映射就是酉算子。酉算子和自伴算子都属于正规算子，正规性则是谱定理的假设。相应的谱分解依次为

$$\begin{aligned} \text{有限维正规矩阵} &\implies \text{有限个正交谱投影之和 / 正交特征向量展开,} \\ \text{紧正规算子} &\implies \text{离散的至多可数的正交特征向量展开,} \\ \text{一般有界正规算子} &\implies \text{投影值测度给出的谱积分.} \end{aligned}$$

前两种分解以特征值求和为主；一般情形还允许连续谱，因此必须使用谱积分。

12 Fourier 理论

12.1 交换群上平移的谱表示

设 G 是局部紧交换群。对 $a \in G$ ，在 $L^2(G)$ 上定义平移算子

$$\tau_a : L^2(G) \rightarrow L^2(G), \quad (\tau_a f)(x) = f(x - a),$$

则有

$$\|\tau_a f\|_2 = \|f\|_2, \quad \tau_a \tau_b = \tau_{a+b} = \tau_b \tau_a.$$

所以每个 τ_a 都是酉算子，并且

$$G \longrightarrow \mathcal{U}(L^2(G)), \quad a \longmapsto \tau_a$$

给出 G 的平移酉表示。

Remark 12.1. 局部紧性是为了保证 G 上存在合适的平移不变测度。平移不变是指集合 E 与 $E + a$ 具有相同测度，这种测度称为 Haar 测度。离散群上的计数测度和实轴上的 Lebesgue 测度都是具体例子。

Remark 12.2. 一般地，设 H 是 Hilbert 空间， $\mathcal{U}(H)$ 是 H 上的酉算子群。群同态

$$\pi : G \longrightarrow \mathcal{U}(H), \quad \pi(a + b) = \pi(a)\pi(b)$$

称为 G 在 H 上的酉表示。由同态性质，

$$\pi(0) = I, \quad \pi(-a) = \pi(a)^{-1} = \pi(a)^*.$$

若 G 是拓扑群，通常还要求对每个 $h \in H$ ，映射 $a \mapsto \pi(a)h$ 连续，这称为强连续性；离散群上的连续性条件自动成立。

记单位圆乘法群为

$$\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

从 G 到 $\mathbb{S}^1$ 的群同态

$$\chi : G \longrightarrow \mathbb{S}^1, \quad \chi(x + y) = \chi(x)\chi(y),$$

称为 G 的特征标。如果 G 还带有拓扑，则只考虑连续特征标；对离散群，连续性自动成立。全体特征标在逐点乘法下构成交换群

$$\widehat{G} = \text{Hom}_{\text{cont}}(G, \mathbb{S}^1),$$

称为 G 的对偶群；对离散群可省略下标 cont 。

用 $\xi \in \widehat{G}$ 标记特征标 χ_ξ 。由同态性质，

$$\tau_a \chi_\xi = \overline{\chi_\xi(a)} \chi_\xi.$$

因此特征标 χ_ξ 是所有平移算子的共同特征函数；对于平移算子 τ_a ，相应的特征值为 $\overline{\chi_\xi(a)}$ 。如果 $\chi_\xi \notin L^2(G)$ ，这个等式仍有形式意义，此时称其为广义特征函数。这些共同广义特征函数恰好由 $\xi \in \widehat{G}$ 参数化，因此可以把 $\widehat{G}$ 看作平移表示的谱参数空间。

对 $f \in L^1(G) \cap L^2(G)$ 定义

$$(\mathcal{C}_G f)(\xi) = c_G \int_G f(x) \overline{\chi_\xi(x)} dx, \quad \xi \in \widehat{G}.$$

这里离散群上的积分就是求和， c_G 用来调整归一化。一般的 Plancherel 定理说明，在 G 与 $\widehat{G}$ 上选取相容的测度后， $\mathcal{C}_G$ 可延拓为酉映射

$$\mathcal{C}_G : L^2(G) \longrightarrow L^2(\widehat{G}), \quad \|\mathcal{C}_G f\|_2 = \|f\|_2.$$

对 $\widehat{G}$ 上的有界可测函数 m ，记 $\mathcal{D}_m$ 为相应的乘法算子：

$$(\mathcal{D}_m h)(\xi) = m(\xi) h(\xi).$$

固定 $a \in G$ ，令 $m_a(\xi) = \overline{\chi_\xi(a)}$ ，则平移在谱坐标中变成乘法算子：

$$\mathcal{C}_G \tau_a \mathcal{C}_G^{-1} = \mathcal{D}_{m_a}.$$

更一般地，与所有平移交换的有界算子在这些谱坐标下都是乘法算子：

$$A \tau_a = \tau_a A \quad (\forall a \in G) \quad \implies \quad \mathcal{C}_G A \mathcal{C}_G^{-1} = \mathcal{D}_m.$$

这给出了平移酉表示的谱分解。这些一般结论属于调和分析，本文不作证明。

Fourier 理论正是这一抽象结构在四个交换群上的具体化。记 $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ ，则

G	特征标	$\widehat{G}$
$\mathbb{Z}/N\mathbb{Z}$	$\chi_k(n) = e^{2\pi i k n / N}$	$\mathbb{Z}/N\mathbb{Z}$
$\mathbb{T}$	$\chi_k(x) = e^{i k x}$	$\mathbb{Z}$
$\mathbb{Z}$	$\chi_\omega(n) = e^{i n \omega}$	$\mathbb{T}$
$\mathbb{R}$	$\chi_\xi(x) = e^{i x \xi}$	$\mathbb{R}$

例如，在 $\mathbb{T}$ 上要求 $e^{i\xi x}$ 具有 2π 周期，便迫使 $\xi \in \mathbb{Z}$ ；在 $\mathbb{Z}$ 上， ω 与 $\omega + 2\pi$ 给出同一序列，所以频率参数属于 $\mathbb{T}$ 。因此，周期性使频率离散，离散性使频率周期。在这四种情形中，抽象的谱坐标映射 $\mathcal{C}_G$ 分别就是 DFT、Fourier 级数、DTFT 和 Fourier 变换。

Fourier 表示	原空间 $L^2(G)$	自然算子	Fourier 空间 $L^2(\widehat{G})$	乘子	谱类型
DFT	$\mathbb{C}^N$	循环移位 S_N	$\mathbb{C}^N$	$e^{-2\pi i k / N}$	有限谱
Fourier 级数	$L^2(\mathbb{T})$	$D = -i\partial_x$	$\ell^2(\mathbb{Z})$	k	离散谱 $\mathbb{Z}$
DTFT	$\ell^2(\mathbb{Z})$	双边移位 S	$L^2(\mathbb{T})$	$e^{-i\omega}$	单位圆上的连续谱
Fourier 变换	$L^2(\mathbb{R})$	$D = -i\partial_x$	$L^2(\mathbb{R})$	ξ	连续谱 $\mathbb{R}$

Remark 12.3. Fourier 理论在各个领域中有不同的系数习惯，只有对称的系数可以保证其作为酉映射，但是差一个常系数不影响主要结论。

12.2 离散 Fourier 变换

Definition 12.1 (离散 Fourier 变换, DFT). 对 $x = (x_0, \dots, x_{N-1}) \in \mathbb{C}^N$, 定义

$$\hat{x}_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n e^{-2\pi i k n / N}, \quad k = 0, \dots, N-1.$$

逆变换为

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \hat{x}_k e^{2\pi i k n / N}.$$

满足

$$\sum_{n=0}^{N-1} |x_n|^2 = \sum_{k=0}^{N-1} |\hat{x}_k|^2.$$

Remark 12.4. 著名的 FFT 只是计算 DFT 的快速算法, 并不是另一种变换。

12.3 Fourier 级数

Definition 12.2 (Fourier 级数). 对 $f \in L^2(\mathbb{T})$, 定义 Fourier 系数

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx, \quad k \in \mathbb{Z}.$$

得到 Fourier 级数展开

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ikx},$$

这里的等式和级数收敛均按 $L^2(\mathbb{T})$ 意义理解, 满足 Parseval 等式

$$\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{k \in \mathbb{Z}} |\hat{f}(k)|^2.$$

12.4 离散时间 Fourier 变换

Definition 12.3 (离散时间 Fourier 变换, DTFT). 对 $x \in \ell^1(\mathbb{Z})$, 定义函数 $X \in L^2(\mathbb{T})$

$$X(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} x_n e^{-in\omega}, \quad \omega \in \mathbb{R}.$$

逆变换为

$$x_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} X(\omega) e^{in\omega} d\omega.$$

由 Fourier 级数的完备性, 可以唯一延拓到 $\ell^2(\mathbb{Z})$, 并且满足

$$\sum_{n \in \mathbb{Z}} |x_n|^2 = \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega.$$

Remark 12.5. DTFT 与 Fourier 级数本质上是同一个西对应的两个方向。

12.5 Fourier 变换

Definition 12.4 (Fourier 变换). 对 $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, 定义

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix\xi} dx.$$

逆变换

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{f}(\xi) e^{ix\xi} d\xi.$$

由 Plancherel 定理可以唯一延拓到 $L^2(\mathbb{R})$, 并且满足

$$\int_{\mathbb{R}} |f(x)|^2 dx = \int_{\mathbb{R}} |\widehat{f}(\xi)|^2 d\xi$$

对一般 L^2 函数, 正逆变换积分都按 L^2 极限理解; 若 $f, \widehat{f} \in L^1(\mathbb{R})$, 则反演公式在 f 的 Lebesgue 点成立。

Remark 12.6 (Riemann–Lebesgue 引理). 若 $f \in L^1(\mathbb{R})$, 则 $\widehat{f}$ 连续, 且

$$\widehat{f}(\xi) \rightarrow 0, \quad |\xi| \rightarrow \infty.$$

换言之, Fourier 变换把 $L^1(\mathbb{R})$ 映入 $C_0(\mathbb{R})$ 。

13 Fredholm 算子与 Fredholm 二择一定理

在有限维线性代数中, 线性方程 $Tx = y$ 的可解性可以被维数公式很好地刻画。设

$$T : X \rightarrow Y$$

是有限维线性空间之间的线性映射, 那么显然有

$$X / \ker T \cong \operatorname{Ran}(T).$$

对应的维数关系为

$$\dim X - \dim \ker T = \dim \operatorname{Ran}(T) = \dim Y - \dim \operatorname{codim} \operatorname{Ran}(T).$$

其中

$$\operatorname{codim} \operatorname{Ran}(T) := \dim(Y / \operatorname{Ran}(T)).$$

特别地, 若 $X = Y$, 则

$$T \text{ 单射} \iff T \text{ 满射} \iff T \text{ 可逆}.$$

这正是有限维线性方程理论中最核心的结构。

在无穷维情形下, 单射和满射并不是等价的。

Example 13.1. 考虑 $X = Y = \ell^2$, 左移算子 (第一位丢弃) 是满射但不是单射, 右移算子 (第一位补零) 是单射但不是满射。

在无穷维空间中, 我们关注的不是从映射角度的可逆性——如果一个映射是单射, 那么总是可以在值域上自然定义逆映射, 这对于无穷维情形仍然成立, 但是没什么用。我们实际关注的是有界线性算子是否存在有界逆算子, 这才对应线性方程问题的适定性。不存在有界逆的原因通常有两类:

1. 核可能非平凡；（不是单射显然不可逆）
2. 值域可能不是整个空间，值域也可能不闭。（保证稳定性，有界逆定理）

Fredholm 理论试图在无穷维空间中延续有限维中的维数数量关系，以及单射满射和可逆的等价关系，处理方法是只考虑一类特定的算子：它们可以不可逆，但不可逆的部分只发生在有限维方向上。

Definition 13.1 (Fredholm 算子). 设 X, Y 是 Banach 空间, $T : X \rightarrow Y$ 是有界线性算子。若

$$\dim \ker T < \infty, \quad \operatorname{codim} \operatorname{Ran}(T) < \infty,$$

并且

$$\operatorname{Ran}(T) \text{ 是 } Y \text{ 中的闭子空间,}$$

则称 T 是 Fredholm 算子。

三个条件分别控制线性方程的非唯一性、不可解性障碍和稳定性：核有限维表示齐次方程的自由方向有限；余维有限表示不可解的相容性条件只有有限多个；值域闭表示可解性对扰动稳定。

Definition 13.2 (Fredholm 指标). 若 $T : X \rightarrow Y$ 是 Fredholm 算子，则定义

$$\operatorname{ind}(T) = \dim \ker T - \operatorname{codim} \operatorname{Ran}(T).$$

Proposition 13.1 (Fredholm 指标为零时的可逆性判别). 设 X 是 Banach 空间, $T : X \rightarrow X$ 是 Fredholm 算子，并且 $\operatorname{ind}(T) = 0$ ，则以下三者等价：

$$T \text{ 单射} \iff T \text{ 满射} \iff T \text{ 存在有界逆}.$$

Proof. 若 T 单射，则 $\dim \ker T = 0$ 。由 $\operatorname{ind}(T) = 0$ 得

$$\operatorname{codim} \operatorname{Ran}(T) = 0.$$

又因为 $\operatorname{Ran}(T)$ 是闭子空间，所以 $\operatorname{Ran}(T) = X$ 。因此 T 是双射，再由有界逆定理可知 T^{-1} 有界。反过来，满射且指标为零推出 $\dim \ker T = 0$ ，所以 T 单射。有有界逆显然等价于双射且逆连续。□

这个结果在有限维中是平凡的：在有限维情形下，有限维线性算子始终是 Fredholm 算子，在 $X = Y$ 时的 Fredholm 指标总是 0，上述定理的条件始终满足，结论自然成立。

Fredholm 算子的一个基本来源是单位算子的紧扰动：若 $K : X \rightarrow X$ 是紧算子，考虑 $I - K$ 的可逆性分析，这在很多偏微分方程问题中是自然出现的。下面的定理表明，对于单位算子的紧扰动，不可逆性只能来自有限维特征空间。

Theorem 13.2 (Fredholm 二择一定理). 设 X 是 Banach 空间, $K : X \rightarrow X$ 是紧算子。则 $I - K$ 是 Fredholm 算子，并且 $\operatorname{ind}(I - K) = 0$ ，进而有如下等价关系

$$I - K \text{ 单射} \iff I - K \text{ 满射} \iff I - K \text{ 存在有界逆}.$$

这个结论常写成如下二择一形式。对方程

$$(I - K)x = y,$$

恰有以下两种情况:

1. 齐次方程 $(I - K)x = 0$ 只有零解。此时对任意 $y \in X$, 非齐次方程存在唯一解。
2. 齐次方程 $(I - K)x = 0$ 有非零解。此时 $\ker(I - K)$ 是有限维的, $\text{Ran}(I - K)$ 是闭子空间, 并且

$$\text{codim Ran}(I - K) = \dim \ker(I - K).$$

从谱理论角度看, Fredholm 二择一定理也解释了为什么紧算子的非零谱点都像有限维矩阵的特征值。由于

$$(I - K)x = 0 \iff Kx = x,$$

1 是 K 的特征值当且仅当 $I - K$ 不是单射。另一方面, 由 Fredholm 二择一定理, $I - K$ 单射、满射和可逆三者等价。因此

$$I - K \text{ 不可逆} \iff 1 \text{ 是 } K \text{ 的特征值}.$$

更一般地, 若 $\lambda \neq 0$, 则

$$\lambda I - K = \lambda \left(I - \frac{1}{\lambda} K \right).$$

由于 $\frac{1}{\lambda}K$ 仍然是紧算子, Fredholm 二择一定理表明 $\lambda I - K$ 不可逆当且仅当 K 有特征值 λ 。因此, 对紧算子而言, 非零谱点都是特征值。这正是紧算子谱理论中最接近有限维矩阵谱理论的一点。

在 Hilbert 空间中, 值域的余维可以通过伴随算子的核表示。若 $T: H_1 \rightarrow H_2$ 是有界线性算子且 $\text{Ran}(T)$ 闭, 则

$$\text{Ran}(T)^\perp = \ker T^*.$$

因此, 当 T 是 Fredholm 算子时,

$$\text{ind}(T) = \dim \ker T - \dim \ker T^*.$$

并且方程

$$Tx = y$$

可解当且仅当

$$y \perp \ker T^*.$$

Example 13.2 (积分方程). 在 $L^2(a, b)$ 中考虑

$$u(x) - \int_a^b k(x, s)u(s) ds = f(x).$$

若 $k \in L^2([a, b] \times [a, b])$, 则

$$K: L^2(a, b) \rightarrow L^2(a, b), \quad (Ku)(x) = \int_a^b k(x, s)u(s) ds,$$

是 Hilbert-Schmidt 算子, 因而是紧算子。原方程写作

$$(I - K)u = f.$$

因此该方程满足 Fredholm 二择一定理：若齐次方程只有零解，则对任意 $f \in L^2(a, b)$ 存在唯一解；若齐次方程有非零解，则可解性由有限多个相容性条件决定。

14 行列式与迹

迹和行列式是有限维矩阵的基本不变量。到了无穷维空间，普通的对角元求和和特征值乘积通常不再收敛，也不一定与基的选择无关。因此这些概念只能在特殊算子类上重新定义。

14.1 迹的推广

以下关于特征值的公式均在复数域上讨论。对于有限维复线性空间 V ，选定一组基以后，线性算子 A 自然得到一个 $n \times n$ 的矩阵表示，其迹定义为

$$\operatorname{tr}(A) = \sum_{j=1}^n a_{jj} = \sum_{j=1}^n \lambda_j.$$

其中 λ_j 是 A 的特征值（按代数重数重复出现）。当空间变成无穷维时，算子通常不能由有限矩阵表示；即使选取某种基，也往往会得到无穷矩阵。此时对角元之和

$$\sum_{j=1}^{\infty} \langle T e_j, e_j \rangle$$

未必收敛；即使收敛，也未必与所选基无关。因此，对于一般的无穷维有界线性算子，迹并没有自动定义。

Definition 14.1 (迹类算子). 设 H 是复可分 Hilbert 空间, $T \in \mathcal{L}(H)$ 是紧线性算子。记 T 的奇异值为

$$s_j(T), \quad j = 1, 2, \dots,$$

也就是非负算子 $(T^*T)^{1/2}$ 的特征值，按重数排列。若

$$\sum_{j=1}^{\infty} s_j(T) < \infty,$$

则称 T 是迹类算子。

Definition 14.2 (迹类算子的迹). 设 T 是复可分 Hilbert 空间 H 上的迹类算子, $\{e_j\}_{j=1}^{\infty}$ 是 H 的一组标准正交基。定义

$$\operatorname{tr}(T) = \sum_{j=1}^{\infty} \langle T e_j, e_j \rangle.$$

该级数绝对收敛，并且其值与所选标准正交基无关。

在迹类算子上，有限维迹的一些基本性质仍然成立。例如，迹是线性的：

$$\operatorname{tr}(\alpha S + \beta T) = \alpha \operatorname{tr}(S) + \beta \operatorname{tr}(T),$$

并且若 T 是迹类算子, $B \in \mathcal{L}(H)$ ，则满足循环性质：

$$\operatorname{tr}(BT) = \operatorname{tr}(TB).$$

14.2 行列式的推广

行列式的推广比迹更加困难。对于有限维复线性空间 V ，选定一组基以后，线性算子 A 自然得到一个 $n \times n$ 的矩阵表示，其行列式定义为

$$\det(A) = \prod_{j=1}^n \lambda_j.$$

其中 λ_j 是 A 的特征值（按代数重数重复出现）。在无穷维中，特征值可能有无穷多个，也可能存在连续谱。即使存在一系列特征值，形式上的无穷乘积

$$\prod_{j=1}^{\infty} \lambda_j$$

也通常没有良好的收敛性。因此，一般无穷维算子没有自然的行列式。

Definition 14.3 (Fredholm 行列式). 设 T 是复可分 Hilbert 空间上的迹类算子。将 T 的非零特征值按照代数重数重复计数，记为 $\{\lambda_j\}$ 。定义 $I + T$ 的 Fredholm 行列式为

$$\det(I + T) = \prod_j (1 + \lambda_j).$$

迹类算子保证

$$\sum_j |\lambda_j| < \infty,$$

从而上述无穷乘积具有良好的收敛性。因此，Fredholm 行列式可以看作有限维公式

$$\det(I + A) = \prod_{j=1}^n (1 + \lambda_j)$$

在迹类算子上的无穷维延伸。Fredholm 行列式不是对任意无穷维算子 A 定义 $\det(A)$ ，而是对恒等算子的迹类扰动算子

$$I + T$$

进行定义，其中 T 是迹类算子。

一般有界算子没有自然的迹和行列式。迹类算子提供了一个足够小的算子类别，使无穷级数和无穷乘积具有良好意义，并保留一部分有限维公式。

15 从有限维到无穷维的总览

从有限维线性代数进入无穷维泛函分析，核心变化不是“矩阵变成无穷矩阵”，而是有限维中自动统一的结构在无穷维中发生分裂。有限维里，线性结构几乎决定一切：范数等价、子空间闭合、线性算子连续、单位球紧、代数可逆通常也意味着稳定可逆。进入无穷维后，这些性质不再由线性结构自动保证，必须分别由拓扑、完备性、对偶性、紧性和算子估计来控制。这种分裂也解释了泛函分析中大量新概念的来源。Banach 空间处理完备性，连续对偶和弱拓扑处理极限过程，自反性和 Banach–Alaoglu 定理补偿强紧性的丢失，闭子空间和投影定理控制最佳逼近，有界逆、闭值域和适定性刻画线性方程的稳定性，紧算子、迹类算子和 Fredholm 行列式则是在特殊算子类上恢复一部分有限维矩阵理论。

有限维与一般无穷维之间的主要差异可以压缩为下表：

有限维线性代数	无穷维泛函分析
基是有限线性组合	Hamel 基、Schauder 基、Hilbert 基分化
所有范数等价	不同范数诱导不同拓扑
空间自动完备	赋范空间未必完备，需要 Banach 空间
代数对偶等于连续对偶	代数对偶与连续对偶分离
有界闭集紧	单位球通常不紧
强收敛与弱收敛等价	强收敛与弱收敛分离
子空间自动闭合	子空间可能稠密但不闭
任意子空间可投影	正交投影需要闭子空间
线性算子自动连续	线性算子可能无界
可逆性自动稳定	逆算子可能无界，值域可能不闭
谱基本就是特征值	谱包含点谱、连续谱、剩余谱
所有矩阵有迹和行列式	迹和行列式只在特殊算子类上恢复

在实际分析和数值问题中，经常不需要面对最一般的 Banach 空间，而是工作在可分 Hilbert 空间中（例如 ℓ^2 、 $L^2(\Omega)$ 、和 $H^k(\Omega)$ ）。此时内积和可数标准正交基恢复了很多有限维欧氏空间的几何图像：对偶可以由 Riesz 表示定理识别为自身，闭子空间有正交投影，Hilbert 空间自反，弱紧性也更容易使用。但可分 Hilbert 空间只是“最接近有限维”的无穷维空间，不是有限维空间。只要维数仍然无穷，单位球的范数紧性、强弱收敛的分离、非闭子空间、非紧有界算子、连续谱、无界算子和不适定逆问题仍然存在。它简化的是一般 Banach 空间的粗糙性，而不是消除无穷维本身的困难。

可分 Hilbert 空间在这两方面的位置可以概括为：

一般无穷维空间中的复杂性	可分 Hilbert 空间中的情况
可能没有可数基	有可数标准正交基
连续对偶可能很复杂	由 Riesz 定理识别为自身
双对偶可能严格更大	Hilbert 空间自反
一般没有正交结构	有内积和正交投影
闭子空间未必有补	闭子空间有正交补
单位球不紧	仍然不紧
强弱收敛分离	仍然分离
子空间可能不闭	仍然可能不闭
有界算子不一定紧	仍然不一定紧
谱不等于特征值集合	仍然不等于
无界算子需要定义域理论	仍然需要
逆算子可能无界	仍然可能无界

换言之，如果只在可分 Hilbert 空间中工作，那么很多结构会重新变得接近有限维：可以用正交基展开，可以用内积表示对偶，可以对闭子空间做正交投影，也可以利用自反性得到弱紧性。但是，只要空间仍然是无穷维的，范数紧性的丢失、强弱收敛的分离、非闭子空间、非紧有界算子、连续谱、无界算子和不适定逆问题仍然是不可避免的。可分 Hilbert 空间是无穷维分析中最接近有限维欧氏空间的情形，但它仍然不是有限维线性代数。

第三部分 基本定理篇

16 Baire 纲定理

Baire 纲定理是从完备性提取全局信息的基本工具。它说明，完备度量空间不可能由可数个“处处稀薄”的闭集拼成；这个看似拓扑的结论将直接用于证明一致有界原理和开映射定理。

Theorem 16.1 (Baire 纲定理). 完备度量空间不能表示为可数个无处稠密闭集的并。

Proof. 反设完备度量空间 X 可以写成

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

其中每个 F_n 都是无处稠密闭集。任取一个非空开球 B_0 。由于 F_1 无处稠密，可以在 $B_0 \setminus F_1$ 中取一个闭球 $\bar{B}_1$ ，使得 $\bar{B}_1 \subset B_0 \setminus F_1$ 。

递归地，若已经取到闭球 $\bar{B}_{n-1}$ ，则 F_n 在 $\bar{B}_{n-1}$ 的内部不可能稠密。因此可以选取闭球 $\bar{B}_n$ ，满足

$$\bar{B}_n \subset \text{int}(\bar{B}_{n-1}) \setminus F_n, \quad \text{diam}(\bar{B}_n) < 2^{-n}.$$

这些闭球彼此嵌套，且直径趋于零。取各球心 x_n ，则 $\{x_n\}$ 是 Cauchy 列。由完备性， $x_n \rightarrow x \in X$ 。闭球的嵌套性给出

$$x \in \bigcap_{n=1}^{\infty} \bar{B}_n.$$

但 $\bar{B}_n \cap F_n = \emptyset$ ，所以 $x \notin F_n$ 对所有 n 都成立，这与 $X = \bigcup_n F_n$ 矛盾。□

直观地说，一个完备空间不能被可数个“没有内部”的薄集合完全填满。这个定理本身看起来不像线性代数，但它是 Banach 空间理论中许多结论的基础。有限维中，很多结论可以通过紧性、坐标分解或矩阵计算直接得到；而在无穷维 Banach 空间中，单位球不再紧，坐标也未必可用，于是完备性开始发挥核心作用。可以把 Baire 纲定理理解为：完备性虽然不能恢复有限维的紧性，但它仍然排除了很多病态情况。

Corollary 16.2 (无穷维 Banach 空间没有可数 Hamel 基). 无穷维 Banach 空间不可能有可数 Hamel 基。

Proof. 若 $\{e_n\}_{n=1}^{\infty}$ 是 Banach 空间 X 的 Hamel 基，令

$$X_n = \text{span}\{e_1, \dots, e_n\}.$$

Hamel 基只允许有限线性组合，因此 $X = \bigcup_{n=1}^{\infty} X_n$ 。每个 X_n 都是有限维闭子空间。又因为 X 无穷维， X_n 是 X 的真子空间，因而没有内点。于是 X 是可数个无处稠密闭集的并，与 Baire 纲定理矛盾。□

Example 16.1. 考虑空间

$$c_{00} = \{x = (x_j)_{j \geq 1} : x_j \text{ 只有有限个非零项}\},$$

并赋予 ℓ^2 范数。显然 c_{00} 不完备，它的完备化空间是 ℓ^2 。

注意到 c_{00} 是无穷维空间, 而且

$$c_{00} = \bigcup_{n=1}^{\infty} \text{span}\{e_1, \dots, e_n\}.$$

右端每个子空间在 c_{00} 中都是无处稠密闭集。由 Baire 纲定理及其推论也可以得到 c_{00} 不完备。

在上面这个例子中, 尤其要注意区分基的概念: c_{00} 存在可数 Hamel 基 $\{e_n\}$, 不完备; 它的完备化空间是 ℓ^2 , 由上述命题可知, ℓ^2 不存在可数 Hamel 基; 在配置 ℓ^2 内积后, ℓ^2 作为 Hilbert 空间, 存在可数的标准正交基 $\{e_n\}$ 。

17 开映射定理与有界逆定理

Theorem 17.1 (开映射定理与有界逆定理). 若 X, Y 是 Banach 空间, $T: X \rightarrow Y$ 是有界线性算子, 并且 T 是满射, 则 T 把 X 中的开集映成 Y 中的开集。特别地, 若 $T: X \rightarrow Y$ 是有界线性双射, 则

$$T^{-1}: Y \rightarrow X$$

也是有界线性算子。

Proof. 先证明开映射定理。记 $B_X = B_X(0, 1)$ 。由 T 满射,

$$Y = \bigcup_{n=1}^{\infty} T(nB_X) \subset \bigcup_{n=1}^{\infty} \overline{T(nB_X)}.$$

由 Baire 纲定理, 存在 n , 使得 $\overline{T(nB_X)}$ 含有一个开球 $B(y_0, r)$ 。将该球中关于 y_0 对称的两点相减, 可得

$$B_Y(0, 2r) \subset \overline{T(2nB_X)}.$$

缩放后存在 $\delta > 0$ 使得

$$B_Y(0, \delta) \subset \overline{T(B_X)}.$$

还需要把右端的闭包去掉。若 $\|y\|_Y < \delta/2$, 由上式及缩放, 可以依次选取 $x_j \in X$, 使得

$$\|x_j\|_X < 2^{-j}, \quad \|y - T(x_1 + \dots + x_j)\|_Y < \frac{\delta}{2^{j+1}}.$$

级数 $\sum_j x_j$ 在 Banach 空间 X 中收敛到某个 x , 且 $\|x\|_X < 1$ 。由 T 连续,

$$Tx = y.$$

因此

$$B_Y(0, \delta/2) \subset T(B_X).$$

通过平移和缩放可知, T 把任意开集映成开集。

若 T 还是双射, 则上面的球包含关系说明: 对任意 $y \in Y$,

$$\|T^{-1}y\|_X \leq \frac{2}{\delta} \|y\|_Y.$$

所以 T^{-1} 有界, 这就是有界逆定理。 □

Example 17.1 (微分算子的开映射性质). 考虑 $C([0, 1])$ 配备范数

$$\|u\|_C = \sup_{x \in [0, 1]} |u(x)|.$$

以及 $C^1([0, 1])$ 配备范数

$$\|u\|_{C^1} = \|u\|_C + \|u'\|_C.$$

易得这两个空间都是 Banach 空间。考虑微分算子

$$D : C^1([0, 1]) \rightarrow C([0, 1]), \quad Du = u',$$

可以证明 D 是有界线性算子：对任意 $u \in C^1([0, 1])$

$$\|Du\|_C = \|u'\|_C \leq \|u\|_{C^1}.$$

并且 D 是满射：对任意 $f \in C([0, 1])$ ，取

$$u(x) = \int_0^x f(t) dt,$$

易得 $Du = f$ 以及

$$\|u\|_{C^1} = \left\| \int_0^x f(t) dt \right\|_C + \|f\|_C \leq 2\|f\|_C$$

应用开映射定理可得到： D 把 $C^1([0, 1])$ 中的开集映成 $C([0, 1])$ 中的开集。

注：显然微分算子 D 不是单射，因为常数函数构成它的核。

有界逆定理在有限维中上述结论是平凡的：矩阵 A 可逆时，逆矩阵 A^{-1} 也自动存在并且连续。因此，有限维中通常不会区分“代数上有唯一解”和“解对数据稳定依赖”。但是在无穷维中，这两件事必须分开：一个线性方程

$$Tx = y$$

即使对每个 y 都存在唯一解，也不必然说明求解过程是稳定的。如果逆算子无界，那么右端 y 的一个很小扰动可能导致解 x 的巨大变化。这正是不适定问题的典型来源。有界逆定理表明：在 Banach 空间之间，如果 T 本身有界、双射，并且两个空间都完备，那么代数可逆性可以提升为稳定可逆性。这可以看作有限维中“可逆矩阵必有连续逆”的无穷维版本。

有界逆定理可以用来证明两个范数的等价关系：只需要一个方向的范数控制不等式就足够了。

Corollary 17.2 (完备范数的等价性判别). 若同一个线性空间上赋予两个范数 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 都是 Banach 空间，且存在常数 C 使得 $\|x\|_2 \leq C\|x\|_1$ 对所有 x 成立，则这两个范数等价（即存在 C' 使得 $\|x\|_1 \leq C'\|x\|_2$ ）。

Proof. 由条件可知恒等映射

$$I : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2)$$

是双射且有界。由于两个空间都完备，有界逆定理说明 I^{-1} 有界，即存在 $C' > 0$ 使得

$$\|x\|_1 \leq C'\|x\|_2.$$

□

完备性在这里还是发挥了重要作用, 缺乏完备性时, 单向的范数控制不等式无法推出范数等价。

Example 17.2. 考虑空间

$$c_{00} = \{x = (x_j)_{j \geq 1} : x_j \text{ 只有有限个非零项}\},$$

并分别配备 ℓ^1 范数和 ℓ^∞ 范数。注意 $(c_{00}, \|\cdot\|_{\ell^1})$ 和 $(c_{00}, \|\cdot\|_{\ell^\infty})$ 都不完备。显然 $\|\cdot\|_{\ell^1}$ 可以控制 $\|\cdot\|_{\ell^\infty}$

$$\|x\|_{\ell^\infty} \leq \|x\|_{\ell^1}$$

但是 $\|\cdot\|_{\ell^\infty}$ 无法控制 $\|\cdot\|_{\ell^1}$, 两个范数不等价。

18 闭图像定理

Theorem 18.1 (闭图像定理). 若 X, Y 是 Banach 空间, $T: X \rightarrow Y$ 是定义在整个 X 上的线性算子, 则 T 有界当且仅当它的图像

$$\mathcal{G}(T) = \{(x, Tx) : x \in X\} \subset X \times Y$$

是闭的。

Proof. 若 T 有界, 设 $x_n \rightarrow x$ 于 X 且 $Tx_n \rightarrow y$ 于 Y 。由连续性, $Tx_n \rightarrow Tx$, 所以 $y = Tx$ 。因此 $\mathcal{G}(T)$ 是闭集。

反之, 设 $\mathcal{G}(T)$ 闭。因为 $X \times Y$ 是 Banach 空间, $\mathcal{G}(T)$ 也是 Banach 空间。考虑投影

$$P: \mathcal{G}(T) \rightarrow X, \quad P(x, Tx) = x.$$

这是有界线性双射。由有界逆定理, 逆映射

$$P^{-1}: X \rightarrow \mathcal{G}(T), \quad P^{-1}x = (x, Tx),$$

有界。因此存在 $C > 0$ 使得

$$\|x\|_X + \|Tx\|_Y \leq C\|x\|_X,$$

从而 $\|Tx\|_Y \leq C\|x\|_X$ 。所以 T 有界。 $\square$

闭图像定理在有限维中是平凡的结论: 在有限维中, 任何线性映射都连续, 它的图像总是闭的。有限维中的线性已经足够推出连续性, 但是在无穷维中的线性映射可以不连续。

闭图像定理的直观意义是: 在 Banach 空间的完备性前提下, 一个线性算子 T 只要保证“输入收敛到 x , 输出也收敛到 Tx ”, 那么这个算子就是有界的。并且反过来也成立, 一个无界线性算子一定会存在这样的异常情况: 输入收敛到 x , 输出却不收敛到 Tx 。

从实践的角度来说, 直接估计

$$\|Tx\|_Y \leq C\|x\|_X$$

有时并不容易, 闭图像定理提供了另一条更容易的证明路线: 只要能够证明

$$(x_n \rightarrow x, \quad Tx_n \rightarrow y) \rightarrow y = Tx$$

就可以推出 T 连续。偏微分方程和算子理论中的很多连续性证明都是基于这条路线进行的。

Example 18.1 (处处定义但无界的线性算子). 记 $\{e_n\}_{n \geq 1}$ 为 ℓ^2 的标准正交序列。它是线性无关的, 因而可以扩充为 ℓ^2 的一组 Hamel 基 $\mathcal{B}$ 。在 $\mathcal{B}$ 上指定

$$Te_n = ne_1, \quad n \geq 1,$$

并令

$$Tb = 0, \quad b \in \mathcal{B} \setminus \{e_n : n \geq 1\}.$$

由于每个 $x \in \ell^2$ 都能唯一表示为有限个 Hamel 基向量的线性组合, 上述规定唯一确定了一个处处定义的线性算子

$$T : \ell^2 \rightarrow \ell^2.$$

由

$$\|e_n\|_{\ell^2} = 1, \quad \|Te_n\|_{\ell^2} = n$$

可知 T 是无界线性算子。下面说明其图像也不是闭的: 取

$$x_n = \frac{1}{n}e_n.$$

则

$$x_n \rightarrow 0, \quad Tx_n = e_1 \rightarrow e_1.$$

因此在图像 $\mathcal{G}(T)$ 中

$$\mathcal{G}(T) \ni (x_n, Tx_n) \rightarrow (0, e_1) \notin \mathcal{G}(T)$$

这里 $(0, e_1) \notin \mathcal{G}(T)$ 是因为 $T0 = 0 \neq e_1$ 。

Theorem 18.2 (Hellinger–Toeplitz 定理). 设 H 是 Hilbert 空间, $T : H \rightarrow H$ 是定义在整个 H 上的线性算子。若 T 是对称的, 即

$$\langle Tx, z \rangle = \langle x, Tz \rangle, \quad x, z \in H,$$

则 T 是有界线性算子。

Proof. 由闭图像定理, 只需证明 $\mathcal{G}(T)$ 是闭集。设

$$x_n \rightarrow x, \quad Tx_n \rightarrow y \quad \text{in } H.$$

对任意 $z \in H$, 由内积的连续性和 T 的对称性,

$$\langle y, z \rangle = \lim_{n \rightarrow \infty} \langle Tx_n, z \rangle = \lim_{n \rightarrow \infty} \langle x_n, Tz \rangle = \langle x, Tz \rangle = \langle Tx, z \rangle.$$

因此 $\langle y - Tx, z \rangle = 0$ 对所有 $z \in H$ 成立。取 $z = y - Tx$, 得到 $y = Tx$ 。所以 $(x, y) \in \mathcal{G}(T)$, 即 $\mathcal{G}(T)$ 是闭集。由闭图像定理, T 有界。□

这里“定义在整个 H 上”不可省略。一般的无界对称算子只能定义在 H 的真子空间上; Hellinger–Toeplitz 定理说明, 处处定义的对称算子不可能无界。

Remark 18.1 (从无界微分算子到 Sobolev 空间与弱解). Hellinger–Toeplitz 定理在现代 PDE 理论中扮演的最核心角色是一个“否定性定理”(No-Go Theorem) 的警示牌: 它彻底堵死了在古典分析框架内处理微分算子的老路, 从而迫使人们开辟无界算子理论与弱解理论的全新范式。以 PDE

中最基本的拉普拉斯算子 Δ 为例。我们自然地希望在 $L^2(\mathbb{R}^d)$ 中研究它。通过经典的分部积分，只要边界项消失，例如考虑光滑紧支函数，就有

$$\langle -\Delta u, v \rangle_{L^2} = \int_{\mathbb{R}^d} \nabla u \cdot \nabla v \, dx = \langle u, -\Delta v \rangle_{L^2},$$

这说明 $-\Delta$ 在 L^2 内积下是对称的。然而它却是无界的。固定非零实值函数 $\varphi \in C_c^\infty(\mathbb{R}^d)$ ，令

$$u_k(x) = k^{d/2} \varphi(kx).$$

此时 $\|u_k\|_{L^2} = \|\varphi\|_{L^2}$ ，而

$$\Delta u_k(x) = k^{d/2+2} (\Delta \varphi)(kx).$$

因此

$$\|\Delta u_k\|_{L^2} = k^2 \|\Delta \varphi\|_{L^2}.$$

由于非零紧支光滑函数不可能满足 $\Delta \varphi = 0$ ，故

$$\frac{\|\Delta u_k\|_{L^2}}{\|u_k\|_{L^2}} = k^2 \frac{\|\Delta \varphi\|_{L^2}}{\|\varphi\|_{L^2}} \rightarrow \infty.$$

所以 $-\Delta$ 不是 $L^2(\mathbb{R}^d)$ 上的有界算子。此时由 Hellinger–Toeplitz 定理可知：既然 Δ 对称且无界，它就不可能定义在整个 L^2 上。这意味着，如果要在 Hilbert 空间框架下研究微分方程，就必须放弃“算子处处有定义”的要求，转而接受算子带有一个稠密定义域。为了给微分运算提供一个同时足够大又能保证完备性的空间，需要引入 Sobolev 空间。例如， Δ 在 $L^2(\mathbb{R}^d)$ 上的自然定义域是 $H^2(\mathbb{R}^d)$ ；它是 L^2 的稠密子空间，并且在 H^2 范数下成为 Hilbert 空间。无界自伴算子理论以及 PDE 中的弱解与强解理论都建立在这一框架之上。

19 一致有界原理（共鸣定理）

Theorem 19.1 (一致有界原理，也称 Banach–Steinhaus 定理). 设 X 是 Banach 空间， Y 是赋范空间， $\mathcal{F} \subset \mathcal{L}(X, Y)$ 是一族有界线性算子。如果对每个固定的 $x \in X$ ，都有

$$\sup_{T \in \mathcal{F}} \|Tx\|_Y < \infty,$$

那么事实上

$$\sup_{T \in \mathcal{F}} \|T\|_{\mathcal{L}(X, Y)} < \infty.$$

Proof. 对每个正整数 n ，令

$$E_n = \left\{ x \in X : \sup_{T \in \mathcal{F}} \|Tx\|_Y \leq n \right\}.$$

每个 E_n 都是闭集。事实上，若 $x_k \in E_n$ 且 $x_k \rightarrow x$ ，则对任意 $T \in \mathcal{F}$ ，连续性给出

$$\|Tx\|_Y = \lim_{k \rightarrow \infty} \|Tx_k\|_Y \leq n.$$

逐点有界性说明 $X = \bigcup_{n=1}^{\infty} E_n$ 。由 Baire 纲定理，某个 E_N 含有开球 $B(x_0, r)$ 。若 $\|x\|_X \leq 1$ ，则

$$x_0 + \frac{r}{2}x \in E_N, \quad x_0 - \frac{r}{2}x \in E_N$$

因此对所有 $T \in \mathcal{F}$,

$$r\|Tx\|_Y \leq \|T(x_0 + \frac{r}{2}x)\|_Y + \|T(x_0 - \frac{r}{2}x)\|_Y \leq 2N.$$

于是

$$\|T\|_{\mathcal{L}(X,Y)} \leq \frac{2N}{r}, \quad T \in \mathcal{F},$$

结论成立。 $\square$

这条定理的有限维版本非常简单: 若 X 是有限维的, 取一组基 $e_1, \dots, e_n$ 。如果每个 Te_j 关于有界线性算子族 $T \in \mathcal{F}$ 都一致有界, 那么任意 $x = \sum_{j=1}^n c_j e_j$ 的像也可以被这些有限个方向上的界控制。因此, 在有限维中, “逐点有界” 转化为 “算子范数有界” 是显然的。无穷维中问题在于: 方向有无穷多个, 而且不能只检查有限组基向量。一个算子族可能在每个固定向量上看起来都没有发散, 但不同算子可能在越来越多的新方向上变得很大。一致有界原理表明, 在 Banach 空间的完备性假设下, 如果一族连续线性算子在每个具体输入上都稳定, 那么它们作为算子整体也是稳定的。

Corollary 19.2 (逐点收敛算子列的一致有界性). 设 X 是 Banach 空间, Y 是赋范空间, 且 $T_n \in \mathcal{L}(X, Y)$ 。若对每个 $x \in X$, 序列 $\{T_n x\}$ 都在 Y 中收敛, 则

$$\sup_n \|T_n\| < \infty.$$

此时可以自然定义极限算子 $T: X \rightarrow Y$

$$Tx := \lim_{n \rightarrow \infty} T_n x$$

可以保证 T 是有界线性算子, 并且

$$\|T\| \leq \liminf_{n \rightarrow \infty} \|T_n\|.$$

Proof. 对每个 $x \in X$, 收敛列 $\{T_n x\}$ 必然有界, 所以一致有界原理给出

$$\sup_n \|T_n\| < \infty.$$

极限算子 T 的线性性由逐点取极限得到。由于

$$\|Tx\|_Y = \lim_{n \rightarrow \infty} \|T_n x\|_Y \leq (\sup_n \|T_n\|) \|x\|_X,$$

因此 T 有界。对任意 $x \in X$ 还有

$$\|Tx\|_Y \leq \left(\liminf_{n \rightarrow \infty} \|T_n\| \right) \|x\|_X.$$

对 $\|x\|_X \leq 1$ 取上确界, 即得

$$\|T\| \leq \liminf_n \|T_n\|.$$

$\square$

一致有界原理经常被反向地使用: 在 Banach 空间中, 构造一系列有界线性算子 $\{T_n\}$, 使得算子范数 $\|T_n\| \rightarrow \infty$ 发散, 那么必然可以存在某个输入 x , 使得对应输出不可能保持有界

$$\sup_n \|T_n x\| = \infty$$

一致有界原理依赖于空间的完备性; 空间不完备时, 结论一般不成立。

Example 19.1. 考虑空间

$$c_{00} = \{x = (x_j)_{j \geq 1} : x_j \text{ 只有有限个非零项}\},$$

并赋予 ℓ^2 范数。显然 c_{00} 不完备，它的完备化空间是 ℓ^2 。定义 $T_n : c_{00} \rightarrow \mathbb{R}$ 如下

$$T_n x = nx_n, \quad x = (x_j) \in c_{00}.$$

显然 T_n 是有界线性算子，并且 $\|T_n\| = n$ ，因此

$$\sup_n \|T_n\| = \infty$$

对每个固定的 $x \in c_{00}$ ，只有有限个 x_n 非零，所以

$$\sup_n |T_n x| < \infty.$$

因此逐点有界。

20 Hahn-Banach 定理

前面几条定理依赖 Banach 空间的完备性，Hahn-Banach 定理关注的则是连续线性泛函的延拓与分离，并不要求空间完备。

Theorem 20.1 (Hahn-Banach 定理). 设 X 是实或复赋范空间。若 $Y \subset X$ 是线性子空间， $f \in Y^*$ ，则存在 $F \in X^*$ ，使得

$$F|_Y = f, \quad \|F\|_{X^*} = \|f\|_{Y^*}.$$

Proof. 先考虑实数情形。记 $C = \|f\|_{Y^*}$ ，并令 $p(x) = C\|x\|$ 。先说明如何把满足 $g \leq p$ 的线性泛函 $g : M \rightarrow \mathbb{R}$ 从子空间 M 延拓到多一个方向的空间。取 $x_0 \notin M$ ，考虑

$$G(m + tx_0) = g(m) + ta.$$

为了使 $G \leq p$ ，只需选择 a 满足

$$\sup_{m \in M} (g(m) - p(m - x_0)) \leq a \leq \inf_{m \in M} (p(m + x_0) - g(m)).$$

这个区间非空。事实上，对任意 $m, n \in M$ ，由 $g \leq p$ 和 p 的次可加性，

$$g(m) - p(m - x_0) \leq p(n + x_0) - g(n).$$

因此一步延拓总能完成，并且仍保持 $G \leq p$ 。

现在考虑所有满足 $G|_Y = f$ 且 $G \leq p$ 的线性延拓，按定义域包含关系排序。任意全序链的并仍是一个满足相同控制的线性延拓。由 Zorn 引理，存在极大延拓 F 。若其定义域不等于 X ，上面的一步延拓还可以增加一个方向，与极大性矛盾。因此 F 定义在整个 X 上，并满足

$$|F(x)| \leq C\|x\|.$$

所以 $\|F\|_{X^*} \leq C$ 。另一方面， $F|_Y = f$ 给出 $\|F\|_{X^*} \geq \|f\|_{Y^*} = C$ ，故两者相等。

若 X 是复赋范空间, 把 X 和 Y 暂时看成实赋范空间, 并令

$$u(y) = \operatorname{Re} f(y), \quad y \in Y.$$

则 u 是实线性泛函。显然 $\|u\| \leq \|f\|$; 反过来, 对任意 $y \in Y$, 可以选择模为 1 的标量 α , 使 $\alpha f(y) = |f(y)|$, 从而

$$|f(y)| = u(\alpha y) \leq \|u\| \|y\|.$$

因此 $\|u\| = \|f\|$ 。由实数情形, u 可以等范数延拓为实线性泛函 $U: X \rightarrow \mathbb{R}$ 。定义

$$F(x) = U(x) - iU(ix).$$

直接计算可知 F 是复线性的。对 $y \in Y$, 有

$$F(y) = \operatorname{Re} f(y) - i \operatorname{Re} f(iy) = f(y),$$

所以 $F|_Y = f$ 。又因为 $\operatorname{Re} F = U$, 对任意 $x \in X$ 选择模为 1 的标量 α , 使得 $\alpha F(x) = |F(x)|$, 便有

$$|F(x)| = U(\alpha x) \leq \|U\| \|x\| = \|f\| \|x\|.$$

因而 $\|F\| \leq \|f\|$; 结合 $F|_Y = f$, 得到 $\|F\| = \|f\|$ 。□

在有限维空间中, 延拓线性泛函本身并不神秘。给定子空间 Y 的一组基, 把它扩充成整个空间 X 的一组基, 然后在新增加的基向量上任意指定泛函值, 就得到一个线性延拓。有限维中真正容易被忽略的是: 由于所有范数等价、所有线性泛函连续, 延拓后的连续性通常不是问题。无穷维中困难恰恰在这里。代数上仍然可以延拓, 但延拓之后是否仍然连续、范数是否可控, 是一个非平凡问题。Hahn-Banach 定理说明: 连续线性泛函具有足够多的延拓, 而且可以保持范数不变。

Hahn-Banach 定理的直观意义是: 无穷维空间中的点和子空间仍然可以被连续线性泛函有效地区分。有限维中经常用法向量、超平面、线性方程来分离点和子空间; Hahn-Banach 定理就是这种几何分离思想在无穷维空间中的延续。因此它也是对偶空间理论、弱拓扑、凸分析和分离定理的基础。

Corollary 20.2 (范数可由连续泛函检测). 设 X 是实或复赋范空间。若 $x \in X$ 且 $x \neq 0$, 则存在 $f \in X^*$, 使得

$$\|f\|_{X^*} = 1, \quad f(x) = \|x\|.$$

Proof. 在一维子空间 $\operatorname{span}\{x\}$ 上定义

$$f_0(tx) = t\|x\|.$$

则 $\|f_0\| = 1$ 。由 Hahn-Banach 定理, f_0 可以等范数延拓为 X 上的连续线性泛函 f 。于是

$$\|f\|_{X^*} = 1, \quad f(x) = \|x\|.$$

□

Example 20.1. 在实 Banach 空间 ℓ^1 中考虑任意 $x = (x_j) \in \ell^1$, 定义

$$a_j = \begin{cases} 1, & x_j \geq 0, \\ -1, & x_j < 0. \end{cases}$$

则 $a = (a_j) \in \ell^\infty$, 并且泛函

$$f_a(y) = \sum_{j=1}^{\infty} a_j y_j$$

满足 $\|f_a\| = 1$ 以及 $f_a(x) = \|x\|_{\ell^1}$.

下面这个结论则是商空间、弱拓扑和凸分离理论中反复使用的基本工具。

Corollary 20.3 (闭子空间与外点可分离). 设 X 是实或复赋范空间, $M \subset X$ 是闭线性子空间. 若 $x \in X \setminus M$, 则存在 $f \in X^*$, 使得

$$f(M) = \{0\}, \quad f(x) \neq 0.$$

Proof. 考虑商空间 X/M 及商映射 $Q: X \rightarrow X/M$. 因为 M 闭且 $x \notin M$, 所以

$$\|Qx\|_{X/M} = \text{dist}(x, M) > 0.$$

由上一个推论, 存在 $\varphi \in (X/M)^*$ 使得 $\varphi(Qx) \neq 0$. 令 $f = \varphi \circ Q$, 则 $f \in X^*$, $f|_M = 0$ 且 $f(x) \neq 0$. $\square$

Example 20.2. 在 ℓ^2 中取闭子空间

$$M = \{y = (y_j) \in \ell^2 : y_1 = 0\},$$

并令 $x = e_1 = (1, 0, 0, \dots)$. 坐标泛函

$$f(y) = y_1$$

属于 $(\ell^2)^*$, 且 $\|f\| = 1$. 显然 $f(y) = 0$ 对所有 $y \in M$ 成立, 而 $f(e_1) = 1$. 因此, f 将闭子空间 M 与外点 e_1 分离开来。

21 基本定理之间的联系

这一篇的定理可以分成两条主线。第一条主线来自完备性: Baire 纲定理直接推出一致有界原理和开映射定理; 有界逆定理是开映射定理的推论; 闭图像定理又可以由有界逆定理得到。它们的逻辑关系可以概括为

$$\text{Baire 纲定理} \implies \begin{cases} \text{开映射定理} \implies \text{有界逆定理} \implies \text{闭图像定理}, \\ \text{一致有界原理}. \end{cases}$$

开映射定理把满射这一代数性质提升为开性和定量可解性; 有界逆定理保证双射有界算子的逆仍然有界; 闭图像定理则把极限过程中的闭性提升为算子的连续性。这三者从不同角度表达了同一现象:

Banach 空间之间的线性问题一旦具有适当的代数性质或闭性，就会自动获得稳定性。一致有界原理把算子族的逐点有界提升为统一的范数有界。

第二条主线是 Hahn–Banach 定理。它不依赖完备性，而是通过线性延拓和极大性原理保证连续线性泛函足够丰富。借助这些泛函，可以刻画范数、分离点与闭子空间，并把向量问题转化为数值问题。因此，Baire 纲定理提供的是完备空间中的整体控制，Hahn–Banach 定理提供的是连续对偶中的观测与分离；二者共同构成 Banach 空间理论的基础工具。

有限维中，上述结论常可借助坐标、矩阵和紧性直接得到。在无穷维中，单位球通常不紧，坐标也未必可用，因此完备性与连续对偶承担了这些结构性作用。前面的反例也说明，若缺少完备性，一致有界、开映射、有界逆和闭图像等结论都可能失效。

第四部分 数值篇

22 Lax 等价定理

Lax 等价定理把数值方法的相容性、稳定性和收敛性联系起来。其中，“稳定性推出收敛性”的证明较为直接；“收敛性推出稳定性”则需要使用一致有界原理。下面先引入相关概念，再给出允许离散空间随离散参数变化的抽象形式。

设 X, Y 是 Banach 空间, $D \subset X$ 是稠密线性子空间, 考虑连续问题

$$Lu = f,$$

其中 $L \in \mathcal{L}(X, Y)$ 是双射。由有界逆定理,

$$L^{-1} \in \mathcal{L}(Y, X).$$

对每个 n , 设 X_n, Y_n 是 Banach 空间, $L_n \in \mathcal{L}(X_n, Y_n)$ 是双射, 并引入有界线性算子

$$\Pi_n : X \rightarrow X_n, \quad E_n : X_n \rightarrow X, \quad R_n : Y \rightarrow Y_n, \quad J_n : Y_n \rightarrow Y.$$

称 Π_n 为插值算子, R_n 为限制算子, E_n 为重构算子 (也称延拓算子), J_n 为提升算子。

需要如下假设:

1. 假设提升算子是限制算子的右逆, 满足:

$$R_n J_n = I_{Y_n},$$

其中 $I_{Y_n} : Y_n \rightarrow Y_n$ 是 Y_n 上的恒等算子。

2. 假设限制算子和提升算子都满足

$$\sup_n \|R_n\| < \infty, \quad \sup_n \|J_n\| < \infty,$$

3. 假设插值算子和重构算子有足够的逼近性:

$$\|E_n \Pi_n u - u\|_X \rightarrow 0 \quad u \in D.$$

4. 假设重构算子 E_n 满足: 存在与 n 无关的常数 $0 < c_E \leq C_E < \infty$, 使得对任意 $v_n \in X_n$,

$$c_E \|v_n\|_{X_n} \leq \|E_n v_n\|_X \leq C_E \|v_n\|_{X_n}.$$

引入如下概念:

1. 称离散问题与连续问题相容, 如果

$$\|(L_n \Pi_n - R_n L)u\|_{Y_n} = \|L_n \Pi_n u - R_n f\|_{Y_n} \rightarrow 0, \quad u \in D.$$

2. 称离散问题稳定, 如果

$$\sup_n \|L_n^{-1}\|_{\mathcal{L}(Y_n, X_n)} < \infty.$$

3. 称离散问题收敛到连续问题, 如果对每个 $f \in Y$, 都有

$$\|(E_n L_n^{-1} R_n - L^{-1})f\|_X \rightarrow 0.$$

也就是说

$$E_n L_n^{-1} R_n \xrightarrow{\text{strong}} L^{-1}.$$

这里使用第 7.2 节的强算子收敛, 即对每个固定右端项收敛, 不要求算子范数收敛。

注意: 对离散问题的稳定性要求都是指关于离散参数 (或者对应的下标 n) 一致的, 而不是固定离散参数后的离散算子的稳定性。具体的离散算子通常都是有限维的, 稳定性是平凡的, 我们关注的是对应的界是否关于离散参数一致。

Theorem 22.1 (Lax 等价定理). 在上述假设下, 若离散问题与连续问题相容, 则

$$\text{稳定性} \iff \text{收敛性}.$$

Proof. 稳定性 $\implies$ 收敛性: 存在 $M > 0$, 使得

$$\|L_n^{-1}\|_{\mathcal{L}(Y_n, X_n)} \leq M, \quad n \geq 1.$$

定义重构后的离散解算子以及连续解算子

$$T_n = E_n L_n^{-1} R_n \in \mathcal{L}(Y, X), \quad T = L^{-1}.$$

任取 $u \in D$, 令 $f = Lu$. 由 $L_n^{-1} L_n \Pi_n u = \Pi_n u$, 有

$$T_n f - T f = E_n L_n^{-1} R_n L u - u = E_n L_n^{-1} (R_n L u - L_n \Pi_n u) + (E_n \Pi_n u - u).$$

因此

$$\|T_n f - T f\|_X \leq \|E_n\| \|L_n^{-1}\| \|R_n L u - L_n \Pi_n u\|_{Y_n} + \|E_n \Pi_n u - u\|_X.$$

由假设 $\|E_n\| \leq C_E$ 、稳定性、相容性和逼近性质可得

$$T_n f \rightarrow T f, \quad \forall f \in L(D).$$

由于

$$\|T_n\| \leq \|E_n\| \|L_n^{-1}\| \|R_n\| \leq C_E M \sup_n \|R_n\|,$$

所以 $\{T_n\}$ 一致有界。由于 $L: X \rightarrow Y$ 是同胚且 D 在 X 中稠密, $L(D)$ 在 Y 中稠密。任取 $f \in Y$ 和 $g \in L(D)$, 有

$$\begin{aligned} \|T_n f - T f\|_X &\leq \|T_n(f - g)\|_X + \|T_n g - T g\|_X + \|T(g - f)\|_X \\ &\leq (\sup_n \|T_n\| + \|T\|) \|f - g\|_Y + \|T_n g - T g\|_X. \end{aligned}$$

先令 $n \rightarrow \infty$, 再利用 $L(D)$ 在 Y 中的稠密性令 $g \rightarrow f$, 得到

$$T_n f \rightarrow T f, \quad \forall f \in Y.$$

因此离散问题收敛到连续问题。

收敛性 $\implies$ 稳定性: 已知

$$T_n f \rightarrow T f, \quad f \in Y.$$

因此, 对每个固定的 $f \in Y$, 序列 $\{T_n f\}$ 有界. 由共鸣定理可得,

$$\sup_n \|T_n\|_{\mathcal{L}(Y, X)} < \infty.$$

固定 n , 任取 $y_n \in Y_n$, 利用 $R_n J_n = I_{Y_n}$ 可得

$$E_n L_n^{-1} y_n = E_n L_n^{-1} R_n J_n y_n = T_n J_n y_n.$$

再利用 E_n 的下界假设得到

$$c_E \|L_n^{-1} y_n\|_{X_n} \leq \|E_n L_n^{-1} y_n\|_X = \|T_n J_n y_n\|_X \leq \|T_n\| \|J_n\| \|y_n\|_{Y_n}.$$

因此

$$\|L_n^{-1}\|_{\mathcal{L}(Y_n, X_n)} \leq \frac{1}{c_E} \|T_n\| \|J_n\|.$$

由 $\{T_n\}$ 和 $\{J_n\}$ 的一致有界性可得

$$\sup_n \|L_n^{-1}\|_{\mathcal{L}(Y_n, X_n)} < \infty.$$

因此离散问题稳定。 $\square$

Remark 22.1. 在有限差分法中, Lax 等价定理是核心理论基础, 尤其对于含时间演化的偏微分方程. 该定理的反方向 (收敛推出稳定) 可用来严格排除那些相容但不稳定的格式; 对于条件稳定的格式, 它同时给出了时间步长等参数必须满足的限制.

与此相对, 有限元方法通常用于求解稳态椭圆方程, 其理论基础主要由 Lax-Milgram 定理与 Céa 引理构成. 由于在构造有限元空间与双线性形式时, 就已将一致稳定性 (如强制性或 inf-sup 条件) 作为优先保证的性质, 实际分析中只需使用正方向 (稳定推出收敛) 即可推导出拟最优误差估计, 而无须借道收敛反推稳定.

Example 22.1. 考虑第二类 Volterra 积分方程

$$u(x) - \int_0^x u(t) dt = f(x), \quad 0 \leq x \leq 1.$$

例如 $f \equiv 1$ 时, 对应的解为 $u(x) = e^x$.

在 $X = Y = C([0, 1])$ 上定义算子

$$(Lu)(x) = u(x) - \int_0^x u(t) dt.$$

对任意 $u \in C([0, 1])$, 有

$$\|Lu\|_{C([0, 1])} \leq \|u\|_{C([0, 1])} + \sup_{0 \leq x \leq 1} \int_0^x |u(t)| dt \leq 2\|u\|_{C([0, 1])},$$

因此 L 有界. 它的逆算子可以显式写成

$$(L^{-1}f)(x) = f(x) + \int_0^x e^{x-t} f(t) dt.$$

直接代入并交换积分次序可验证 $LL^{-1} = L^{-1}L = I$, 因此 L 是双射. 此外,

$$\begin{aligned} |(L^{-1}f)(x)| &\leq \|f\|_{C([0, 1])} \left(1 + \int_0^x e^{x-t} dt\right) \\ &= e^x \|f\|_{C([0, 1])} \leq e \|f\|_{C([0, 1])}, \end{aligned}$$

故 $\|L^{-1}\|_{C([0,1]) \rightarrow C([0,1])} \leq e$ 。

取均匀网格 $x_i = ih$, 其中 $h = 1/n$, $0 \leq i \leq n$, 并令

$$X_n = Y_n = \mathbb{R}^{n+1}, \quad \|v\|_\infty = \max_{0 \leq i \leq n} |v_i|.$$

插值算子 Π_n 和限制算子 R_n 都取为网格点取值算子, 用于离散未知量和右端项

$$\Pi_n u = (u(x_0), \dots, u(x_n)), \quad R_n f = (f(x_0), \dots, f(x_n)).$$

对 $v = (v_0, \dots, v_n)$, 记 $p(v)$ 为通过 (x_i, v_i) 的分段线性函数, 重构算子 E_n 和提升算子 J_n 都取为 p 。易得

$$\|E_n\|_{\infty \rightarrow C([0,1])} = \|J_n\|_{\infty \rightarrow C([0,1])} = \|R_n\|_{C([0,1]) \rightarrow \infty} = 1, \quad R_n J_n = I_{Y_n}.$$

对积分用梯形公式离散, 得到数值格式如下:

$$(L_n v)_i = v_i - h \left(\frac{1}{2} v_0 + v_1 + \dots + v_{i-1} + \frac{1}{2} v_i \right), \quad 1 \leq i \leq n,$$

以及 $(L_n v)_0 = v_0$ 。由于对应的矩阵为下三角矩阵, 对角元为 $1, 1 - h/2, \dots, 1 - h/2$, 故 L_n 可逆。

因为梯形公式对分段线性函数是精确的, 得到

$$L_n = R_n L E_n.$$

因此

$$L_n \Pi_n u - R_n L u = R_n L (E_n \Pi_n u - u).$$

因为连续函数的分段线性插值一致收敛, 右端在最大范数下趋于零, 这保证了格式的相容性。

下面具体验证稳定性。任取 $g = (g_0, \dots, g_n) \in Y_n$, 令 $v = L_n^{-1} g$ 。具体即 $v_0 = g_0$

$$\left(1 - \frac{h}{2}\right) v_i = g_i + \frac{h}{2} v_0 + h \sum_{j=1}^{i-1} v_j, \quad 1 \leq i \leq n.$$

因为 $0 < h \leq 1$, 所以 $(1 - h/2)^{-1} \leq 2$, 从而

$$|v_i| \leq 2\|g\|_\infty + h|v_0| + 2h \sum_{j=1}^{i-1} |v_j| \leq 3\|g\|_\infty + 2h \sum_{j=0}^{i-1} |v_j| =: S_i.$$

补充定义 $S_0 = 3\|g\|_\infty$ 。于是 $|v_i| \leq S_i$, 并且 S_i 满足递推关系

$$S_{i+1} = S_i + 2h|v_i| \leq (1 + 2h)S_i.$$

因此

$$S_i \leq (1 + 2h)^i S_0 = 3\|g\|_\infty (1 + 2h)^i.$$

再由 $1 + 2h \leq e^{2h}$ 以及 $ih \leq nh = 1$, 可得

$$|v_i| \leq S_i \leq 3\|g\|_\infty e^{2ih} \leq 3e^2 \|g\|_\infty, \quad 0 \leq i \leq n.$$

这实际就是离散版本的 Gronwall 不等式。因而

$$\|L_n^{-1} g\|_\infty \leq 3e^2 \|g\|_\infty, \quad \|L_n^{-1}\|_{\infty \rightarrow \infty} \leq 3e^2.$$

右端常数与 n 无关, 所以该格式稳定。

前文已经验证了格式的相容性和稳定性, 由 Lax 等价定理可得收敛性

$$E_n L_n^{-1} R_n f \longrightarrow L^{-1} f \quad \text{in } C([0, 1]).$$

Example 22.2 (输运方程的一阶迎风格式). 考虑周期初值问题

$$u_t + u_x = 0, \quad u(x, 0) = u_0(x), \quad u(0, t) = u(1, t).$$

下面只说明它如何放入前述 Lax 等价定理的记号中。记 $Q = [0, 1] \times [0, T]$, 空间变量按周期 1 理解, 并定义

$$X = C_{\text{per}}(Q), \quad Y = C_{\text{per}}([0, 1]) \times C_{\text{per},0}(Q),$$

其中

$$C_{\text{per},0}(Q) = \{p \in C_{\text{per}}(Q) : p(x, 0) = 0\}.$$

在 X 和 Y 上分别取范数

$$\|u\|_X = \|u\|_{C(Q)}, \quad \|(g, r)\|_Y = \max\{\|g\|_{C([0,1])}, \|r\|_{C(Q)}\}.$$

连续算子 $L : X \rightarrow Y$ 定义为

$$L : u \mapsto Lu = (g, r),$$

其中

$$g(x) = u(x, 0), \quad r(x, t) = u(x, t) - u(x - t, 0).$$

因而 $Lu = (u_0, 0)$ 恰好表示

$$u(x, t) = u_0(x - t),$$

也就是输运方程的解。这个写法把微分方程改写成了沿特征线的有界算子方程。其逆算子为

$$(L^{-1}(g, r))(x, t) = g(x - t) + r(x, t),$$

所以 L 是有界双射。

取 $x_j = jh$, $h = 1/N$ 和 $t_m = mk$, $k = T/M$, 并令

$$X_n = \{U = (U_j^m) : 0 \leq j < N, 0 \leq m \leq M\}.$$

Y_n 中的元素写成 (G, R) , 其中

$$G = (G_j)_{j=0}^{N-1}, \quad R = (R_j^m)_{\substack{0 \leq j < N \\ 1 \leq m \leq M}}.$$

两个离散空间都取相应的最大范数。

令 $\lambda = k/h$, 并定义一步迎风算子

$$(A_n V)_j = (1 - \lambda)V_j + \lambda V_{j-1},$$

其中下标 j 周期延拓。离散算子 $L_n : X_n \rightarrow Y_n$ 定义为

$$L_n U = (G, R), \quad G_j = U_j^0, \quad R_j^m = U_j^m - (A_n^m U^0)_j.$$

因此

$$L_n U = R_n(u_0, 0),$$

等价于

$$U^m = A_n^m U^0,$$

也就是逐步使用迎风格式

$$U_j^{m+1} = (1 - \lambda)U_j^m + \lambda U_{j-1}^m.$$

这里把所有时间层放进一个向量 U ，于是 L_n^{-1} 正是从离散初值和离散残量生成整段离散解的算子。

现在定义前述定理中的四个辅助算子。插值算子 $\Pi_n : X \rightarrow X_n$ 取时空网格值：

$$(\Pi_n u)_j^m = u(x_j, t_m).$$

限制算子 $R_n : Y \rightarrow Y_n$ 分别对 g 和 r 取网格值：

$$R_n(g, r) = ((g(x_j))_j, (r(x_j, t_m))_{j,m}).$$

重构算子 $E_n : X_n \rightarrow X$ 对 U 作周期双线性插值；提升算子 $J_n : Y_n \rightarrow Y$ 分别对 G 和 R 作相同的分段线性重构。于是

$$R_n J_n = I_{Y_n}, \quad \|E_n\| = \|R_n\| = \|J_n\| = 1,$$

并且对连续函数 u 有

$$\|E_n \Pi_n u - u\|_{C(Q)} \rightarrow 0.$$

在这些记号下，相容性写成

$$\|L_n \Pi_n u - R_n L u\|_{Y_n} \rightarrow 0.$$

对 $u \in C_{\text{per}}^2(Q)$ ，这个量衡量 $A_n^m(u(x_j, 0))_j$ 与特征线平移 $u(x_j - t_m, 0)$ 之间的误差；它由迎风差分的局部截断误差趋于零得到。

稳定性也正好是前述定义中的

$$\sup_n \|L_n^{-1}\|_{\mathcal{L}(Y_n, X_n)} < \infty.$$

当 $0 \leq \lambda \leq 1$ 时， A_n 是两个移位值的凸组合，故 $\|A_n^m\| \leq 1$ 。由

$$(L_n^{-1}(G, R))^m = A_n^m G + R^m,$$

立即得到

$$\|L_n^{-1}(G, R)\|_{X_n} \leq 2\|(G, R)\|_{Y_n}.$$

因此 $\|L_n^{-1}\| \leq 2$ ，且上界与网格无关。

由 Lax 等价定理，

$$E_n L_n^{-1} R_n(u_0, 0) \longrightarrow L^{-1}(u_0, 0) \quad \text{in } X.$$

右端就是 $u(x, t) = u_0(x - t)$ 。

23 变分问题

23.1 连续变分问题

有限元方法通常不直接离散微分算子，而是先把偏微分方程写成变分问题。设 X, Y 是 Banach 空间， Y 自反， $a : X \times Y \rightarrow \mathbb{R}$ 是连续双线性型。给定 $f \in Y^*$ ，考虑如下变分问题：求 $u \in X$ ，使得

$$a(u, v) = f(v), \quad \forall v \in Y.$$

定义线性算子 $A : X \rightarrow Y^*$ 如下

$$\langle Au, v \rangle_{Y^*, Y} = a(u, v).$$

因为 a 连续，存在 $M > 0$ ，使得

$$|a(u, v)| \leq M \|u\|_X \|v\|_Y.$$

因此

$$\|Au\|_{Y^*} = \sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|v\|_Y} \leq M \|u\|_X,$$

所以 $A \in \mathcal{L}(X, Y^*)$ 。也就是说， A 的有界性已经由双线性型的连续性直接保证。

变分问题等价于算子方程

$$Au = f.$$

由于 A 已经是有界线性算子，问题的适定性归结为 A 是否为双射。若 A 是双射，则有界逆定理保证 $A^{-1} : Y^* \rightarrow X$ 有界。

Theorem 23.1 (Banach–Nečas–Babuška 定理，广义 Lax–Milgram 定理). 上述变分问题对每个 $f \in Y^*$ 都适定，当且仅当以下两个条件成立：

1. *inf-sup* 条件：存在常数 $\beta > 0$ ，使得

$$\inf_{0 \neq u \in X} \sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|u\|_X \|v\|_Y} \geq \beta;$$

2. 伴随非退化条件：若 $v \in Y$ 满足

$$a(u, v) = 0, \quad \forall u \in X,$$

则 $v = 0$ 。

此时解满足

$$\|u\|_X \leq \frac{1}{\beta} \|f\|_{Y^*}.$$

Proof. 记 $A : X \rightarrow Y^*$ 为双线性型诱导的算子。第一个条件等价于

$$\|Au\|_{Y^*} = \sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|v\|_Y} \geq \beta \|u\|_X, \quad u \in X.$$

因此 A 是单射。

这个下界还保证了 $\text{Ran}(A)$ 是闭的：取 $\text{Ran}(A)$ 中任意一个收敛序列 $\{Au_n\}$ ，设

$$Au_n \rightarrow f \quad \text{in } Y^*.$$

只需要证明 $f \in \text{Ran}(A)$ 。因为收敛列必为 Cauchy 列, $\{Au_n\}$ 是 Y^* 中的 Cauchy 列。由于

$$\|u_n - u_m\|_X \leq \frac{1}{\beta} \|Au_n - Au_m\|_{Y^*},$$

因此 $\{u_n\}$ 是 Banach 空间 X 中的 Cauchy 列。由 X 的完备性, 存在 $u \in X$, 使得

$$u_n \rightarrow u \quad \text{in } X.$$

又因为 A 连续,

$$Au_n \rightarrow Au \quad \text{in } Y^*.$$

由 Y^* 中极限的唯一性,

$$f = Au \in \text{Ran}(A).$$

下面证明 A 满射。若 $\text{Ran}(A)$ 是 Y^* 的真闭子空间, Hahn-Banach 定理给出非零的 $\Phi \in Y^{**}$, 使得

$$\Phi(Au) = 0, \quad \forall u \in X.$$

因为 Y 自反, 存在非零的 $v \in Y$, 使得 $\Phi(f) = f(v)$ 对所有 $f \in Y^*$ 成立。于是

$$a(u, v) = \langle Au, v \rangle_{Y^*, Y} = 0, \quad \forall u \in X,$$

这与第二个条件矛盾。因此 A 满射, 从而是双射。

由于 A 是双射, A^{-1} 存在。再由有界逆定理, A^{-1} 有界, 因此问题适定。由第一个条件易得, $Au = f$ 的解满足

$$\|u\|_X \leq \frac{1}{\beta} \|f\|_{Y^*}.$$

反过来, 若问题适定, 则 A 是有界双射。有界逆定理给出

$$\|u\|_X \leq \|A^{-1}\| \|Au\|_{Y^*},$$

所以第一个条件成立, 直接取 $\beta = \|A^{-1}\|^{-1}$ 即可。若 $a(u, v) = 0$ 对所有 $u \in X$ 成立, 由 A 满射可知 $f(v) = 0$ 对所有 $f \in Y^*$ 成立。连续对偶可以分离点, 故 $v = 0$, 第二个条件也成立。□

inf-sup 条件的含义为: 对固定的 $u \in X$, 内层上确界

$$\sup_{0 \neq v \in Y} \frac{|a(u, v)|}{\|v\|_Y}$$

表示检验空间 Y 能够观测到 u 的最大强度; 外层下确界则要求这种观测对所有非零 u 都有统一下界。换言之, 对每个 $u \neq 0$, 都能在 Y 中找到合适的检验方向 v , 使 $a(u, v)$ 不会相对于 $\|u\|_X \|v\|_Y$ 任意小, 并且常数 β 与 u 无关。利用 A 的定义可知, 双线性形式满足 inf-sup 条件等价于 A 满足

$$\|Au\|_{Y^*} \geq \beta \|u\|_X, \quad u \in X.$$

因此它给出 A 的有下界性, 保证 A 单射、值域闭, 并且逆算子 $A^{-1} : \text{Ran}(A) \rightarrow X$ 有界。第二个条件等价于伴随算子 $A^* : Y \rightarrow X^*$ 单射, 它排除了值域的余维, 从而保证 A 满射。

Example 23.1 (BNB 定理的有限维矩阵版本). 取 $X = \mathbb{R}^n$ 、 $Y = \mathbb{R}^m$, 并令

$$a(x, y) = y^\top Ax, \quad A \in \mathbb{R}^{m \times n}.$$

对固定的 x 易得

$$\sup_{0 \neq y \in \mathbb{R}^m} \frac{|y^\top Ax|}{\|y\|_2} = \|Ax\|_2.$$

因此, BNB 定理的第一个条件即

$$\|Ax\|_2 \geq \beta \|x\|_2, \quad x \in \mathbb{R}^n.$$

这等价于 A 列满秩, 最佳常数 β 是 A 的最小奇异值. BNB 定理的第二个条件即

$$y^\top Ax = 0 \quad \forall x \in \mathbb{R}^n \implies y = 0,$$

也就是 $A^\top$ 单射, 等价于 A 行满秩.

BNB 定理此时的表述变为: A 可逆, 当且仅当列满秩和行满秩这两个条件同时成立. 在非方阵情形下 ($m \neq n$), 两个条件和 A 可逆都不能成立. 在方阵情形下 ($m = n$), 第一个条件要求行满秩, 第二个条件要求列满秩, 这两个条件实际是重复的, 因为方阵的行满秩与列满秩等价.

上述例子表明, 在有限维中, 这两个条件相互等价, 实际只相当于一个条件. 但是在无穷维情形下, 两个条件并不等价, 单射不等于满射.

当 $X = Y = H$ 是 Hilbert 空间时, 可以用更容易验证的强制性条件代替一般的 inf-sup 条件. 称双线性型 $a : H \times H \rightarrow \mathbb{R}$ 是强制的, 如果存在 $\alpha > 0$, 使得

$$a(v, v) \geq \alpha \|v\|_H^2, \quad \forall v \in H.$$

Theorem 23.2 (Lax–Milgram 定理). 设 H 是实 Hilbert 空间, $a : H \times H \rightarrow \mathbb{R}$ 是连续双线性型. 若存在常数 $M, \alpha > 0$, 使得

$$|a(u, v)| \leq M \|u\|_H \|v\|_H, \quad a(v, v) \geq \alpha \|v\|_H^2,$$

则对任意 $f \in H^*$, 存在唯一的 $u \in H$, 使得

$$a(u, v) = f(v), \quad \forall v \in H.$$

并且

$$\|u\|_H \leq \frac{1}{\alpha} \|f\|_{H^*}.$$

Proof. 强制性条件可以同时推出 BNB 定理中需要的 in-sup 条件和伴随非退化条件. 对任意 $0 \neq u \in H$, 在上确界中取 $v = u$, 便有

$$\sup_{0 \neq v \in H} \frac{|a(u, v)|}{\|v\|_H} \geq \frac{a(u, u)}{\|u\|_H} \geq \alpha \|u\|_H.$$

若 $v \in H$ 满足 $a(u, v) = 0$ 对所有 $u \in H$ 成立, 取 $u = v$, 由强制性得到 $v = 0$. 因此 BNB 定理的两个条件都成立, 并给出存在唯一性和稳定估计. $\square$

基于 Riesz 表示定理的证明. 由 Riesz 表示定理, 存在唯一的 $Tu \in H$, 使得

$$a(u, v) = \langle Tu, v \rangle_H, \quad \forall v \in H.$$

双线性型 a 的线性和连续性说明 $T: H \rightarrow H$ 是有界线性算子, 并且 $\|T\| \leq M$ 。同理, 存在唯一的 $g \in H$, 使得

$$f(v) = \langle g, v \rangle_H, \quad \forall v \in H.$$

因此原变分问题等价于算子方程

$$Tu = g.$$

由强制性和 Cauchy-Schwarz 不等式,

$$\alpha \|u\|_H^2 \leq a(u, u) = \langle Tu, u \rangle_H \leq \|Tu\|_H \|u\|_H.$$

因而

$$\|Tu\|_H \geq \alpha \|u\|_H, \quad u \in H.$$

所以 T 是单射, 并且与 BNB 证明中相同的 Cauchy 列论证表明 $\text{Ran}(T)$ 是闭子空间。

下面证明 T 满射。若 $w \in \text{Ran}(T)^\perp$, 则

$$0 = \langle Tu, w \rangle_H = a(u, w), \quad \forall u \in H.$$

取 $u = w$, 由强制性得到

$$0 = a(w, w) \geq \alpha \|w\|_H^2,$$

所以 $w = 0$ 。因此

$$\text{Ran}(T)^\perp = \{0\},$$

从而 $\text{Ran}(T)$ 在 H 中稠密。它又是闭子空间, 故 $\text{Ran}(T) = H$ 。因此 T 是双射, 方程 $Tu = g$ 存在唯一解。

最后, 在变分方程中取 $v = u$, 得到

$$\alpha \|u\|_H^2 \leq a(u, u) = f(u) \leq \|f\|_{H^*} \|u\|_H,$$

因而

$$\|u\|_H \leq \frac{1}{\alpha} \|f\|_{H^*}.$$

□

Remark 23.1. 对比这两种证明过程, 基于 Riesz 表示定理的证明体现了 Hilbert 空间相对于一般 Banach 空间的便利。在 BNB 定理中, 双线性型诱导的算子具有形式

$$A: X \rightarrow Y^*.$$

为了证明 A 满射, 需要先由 inf-sup 条件证明 $\text{Ran}(A)$ 闭, 再用 Hahn-Banach 定理在 Y^{**} 中分离这个值域, 最后利用 Y 的自反性把所得泛函表示成 Y 中的元素。因此, BNB 定理还需要单独的伴随非退化条件。

在 Hilbert 空间中, Riesz 表示定理使得双线性型可以写成

$$a(u, v) = \langle Tu, v \rangle_H, \quad T: H \rightarrow H.$$

此时, Banach 空间证明中值域的零化子变成了 $\text{Ran}(T)^\perp$ 。若 w 属于这个正交补, 则 $a(u, w) = 0$ 对所有 $u \in H$ 成立; 直接取 $u = w$, 强制性便给出 $w = 0$ 。因而不需要显式使用 Hahn-Banach 定理和双对偶空间。

23.2 离散变分问题

前一小节讨论了连续变分问题的适定性。下面转向相应的离散问题：先考虑试探空间和检验空间不同的 Petrov–Galerkin 方法，再考虑二者相同的 Galerkin 方法。

沿用连续变分问题的记号。设

$$a(u, v) = f(v), \quad \forall v \in Y$$

适定，其中 $a : X \times Y \rightarrow \mathbb{R}$ 连续。取有限维试探空间 $X_h \subset X$ 和检验空间 $Y_h \subset Y$ 。Petrov–Galerkin 离散问题是求 $u_h \in X_h$ ，使得

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in Y_h.$$

定义 $A_h : X_h \rightarrow Y_h^*$ ：

$$\langle A_h u_h, v_h \rangle_{Y_h^*, Y_h} = a(u_h, v_h).$$

于是离散问题为

$$A_h u_h = f|_{Y_h}.$$

下面的讨论始终要求 $\dim X_h = \dim Y_h$ ，使得离散矩阵为方阵。

Theorem 23.3 (Petrov–Galerkin 离散问题的适定性). 设 X, Y 是 Banach 空间, $a : X \times Y \rightarrow \mathbb{R}$ 是连续双线性型, $X_h \subset X$ 与 $Y_h \subset Y$ 是有限维子空间, 并且 $\dim X_h = \dim Y_h$ 。定义 $A_h : X_h \rightarrow Y_h^*$ 为

$$\langle A_h w_h, v_h \rangle_{Y_h^*, Y_h} = a(w_h, v_h).$$

则以下两个条件等价：

1. 每个 A_h 都是双射，并且离散逆算子族一致有界；
2. 存在与 h 无关的常数 $\beta > 0$ ，使得

$$\inf_{0 \neq w_h \in X_h} \sup_{0 \neq v_h \in Y_h} \frac{|a(w_h, v_h)|}{\|w_h\|_X \|v_h\|_Y} \geq \beta, \quad \forall h.$$

若上述条件成立，则对任意 $f \in Y^*$ ，每个 Petrov–Galerkin 离散问题都有唯一解 $u_h \in X_h$ ，并且

$$\|u_h\|_X \leq \frac{1}{\beta} \|f|_{Y_h}\|_{Y_h^*} \leq \frac{1}{\beta} \|f\|_{Y^*}, \quad \forall h.$$

Proof. 离散 inf-sup 条件等价于

$$\|A_h w_h\|_{Y_h^*} \geq \beta \|w_h\|_X, \quad \forall h.$$

对每个固定的 h ，这说明 A_h 单射。因为 X_h 与 Y_h^* 维数相同，有限维中的单射等价于满射，所以 A_h 是双射；而且

$$\|A_h^{-1}\|_{\mathcal{L}(Y_h^*, X_h)} \leq \frac{1}{\beta}, \quad \forall h.$$

因而离散逆算子族一致有界。

反过来，若每个 A_h 都是双射，并且存在与 h 无关的常数 C ，使得

$$\|A_h^{-1}\|_{\mathcal{L}(Y_h^*, X_h)} \leq C, \quad \forall h,$$

则

$$\|A_h w_h\|_{Y_h^*} \geq \frac{1}{C} \|w_h\|_X, \quad \forall w_h \in X_h, \quad \forall h.$$

取 $\beta = 1/C$, 即得一致离散 inf-sup 条件。

对任意 $f \in Y^*$, 离散解为 $u_h = A_h^{-1}(f|_{Y_h})$, 所以

$$\|u_h\|_X \leq \frac{1}{\beta} \|f|_{Y_h}\|_{Y_h^*}.$$

又因为 $Y_h \subset Y$, 所以

$$\|f|_{Y_h}\|_{Y_h^*} = \sup_{0 \neq v_h \in Y_h} \frac{|f(v_h)|}{\|v_h\|_Y} \leq \|f\|_{Y^*}.$$

□

Remark 23.2. 对每个固定的 h , 上述论证是 BNB 定理的有限维版本。由于 $\dim X_h = \dim Y_h$, 第一个 inf-sup 条件已经保证 A_h 单射, 进而保证它满射, 因此不需要另行验证伴随非退化条件。不过, 逐个应用这个结论只能得到可能依赖于 h 的常数 $\beta_h > 0$, 也只能说明每个离散矩阵分别可逆。这里的适定性实际是对离散问题族的一致适定性: 要求同一个 $\beta > 0$ 对所有 h 成立, 从而保证网格加密时离散逆算子的范数不会发散。

Proposition 23.4 (Petrov–Galerkin 正交性). 连续解 u 与离散解 u_h 满足

$$a(u - u_h, v_h) = 0, \quad \forall v_h \in Y_h.$$

Proof. 因为 $Y_h \subset Y$, 连续问题和离散问题在任意 $v_h \in Y_h$ 上的右端都等于 $f(v_h)$, 相减得证。□

Theorem 23.5 (Petrov–Galerkin 拟最优误差估计). 设 a 的连续性常数为 M , 并且上述一致离散 inf-sup 条件成立。则

$$\|u - u_h\|_X \leq \left(1 + \frac{M}{\beta}\right) \inf_{w_h \in X_h} \|u - w_h\|_X.$$

Proof. 任取 $w_h \in X_h$ 。由离散 inf-sup 条件和 $a(u - u_h, v_h) = 0$ 可得

$$\beta \|u_h - w_h\|_X \leq \sup_{0 \neq v_h \in Y_h} \frac{|a(u_h - w_h, v_h)|}{\|v_h\|_Y} = \sup_{0 \neq v_h \in Y_h} \frac{|a(u - w_h, v_h)|}{\|v_h\|_Y} \leq M \|u - w_h\|_X.$$

因此, 由三角不等式,

$$\|u - u_h\|_X \leq \|u - w_h\|_X + \|w_h - u_h\|_X \leq \left(1 + \frac{M}{\beta}\right) \|u - w_h\|_X.$$

对 w_h 取下确界即得结论。□

Petrov–Galerkin 收敛包含稳定性和逼近性两个方面。一致离散 inf-sup 条件控制稳定性, 而

$$\inf_{w_h \in X_h} \|u - w_h\|_X$$

衡量 X_h 对 u 的最佳逼近误差。如果存在插值算子 $I_h : X \rightarrow X_h$ 满足

$$\|u - I_h u\|_X \rightarrow 0,$$

就有

$$\|u - u_h\|_X \leq \left(1 + \frac{M}{\beta}\right) \inf_{w_h \in X_h} \|u - w_h\|_X \leq \left(1 + \frac{M}{\beta}\right) \|u - I_h u\|_X \rightarrow 0.$$

从而得到收敛性

$$u_h \rightarrow u.$$

更精细的定量插值估计可以给出相应阶的有限元误差估计。

接下来考虑较容易处理的情形。设 V 是 Hilbert 空间，并令

$$X = Y = V, \quad X_h = Y_h = V_h \subset V.$$

设 $a : V \times V \rightarrow \mathbb{R}$ 连续且强制，即

$$|a(u, v)| \leq M \|u\|_V \|v\|_V, \quad a(v, v) \geq \alpha \|v\|_V^2.$$

Galerkin 离散问题是求 $u_h \in V_h$ ，使得

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in V_h.$$

选取 V_h 的基 $\{\phi_i\}_{i=1}^N$ ，离散问题成为

$$A_h U = F, \quad (A_h)_{ij} = a(\phi_j, \phi_i), \quad F_i = f(\phi_i).$$

Theorem 23.6 (Galerkin 离散问题的适定性). 设 V 是 Hilbert 空间， $a : V \times V \rightarrow \mathbb{R}$ 是连续双线性型，并且存在 $\alpha > 0$ ，使得

$$a(v, v) \geq \alpha \|v\|_V^2, \quad \forall v \in V.$$

则对任意 $f \in V^*$ 和任意有限维子空间 $V_h \subset V$ ，Galerkin 离散问题都有唯一解 $u_h \in V_h$ ，并且

$$\|u_h\|_V \leq \frac{1}{\alpha} \|f|_{V_h}\|_{V_h^*} \leq \frac{1}{\alpha} \|f\|_{V^*}.$$

Proof. 强制性限制到 V_h 后仍成立。对任意 $0 \neq w_h \in V_h$ ，在离散 inf-sup 条件中取 $v_h = w_h$ ，得到

$$\sup_{0 \neq v_h \in V_h} \frac{|a(w_h, v_h)|}{\|v_h\|_V} \geq \frac{a(w_h, w_h)}{\|w_h\|_V} \geq \alpha \|w_h\|_V.$$

因而离散 inf-sup 条件以与 h 无关的常数 $\beta = \alpha$ 成立。由 Petrov–Galerkin 离散问题的适定性定理，离散问题有唯一解，且

$$\|u_h\|_V \leq \frac{1}{\alpha} \|f|_{V_h}\|_{V_h^*}.$$

最后，由 $V_h \subset V$ ，

$$\|f|_{V_h}\|_{V_h^*} = \sup_{0 \neq v_h \in V_h} \frac{|f(v_h)|}{\|v_h\|_V} \leq \|f\|_{V^*}.$$

□

Theorem 23.7 (Céa 引理). Galerkin 解满足

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V.$$

Proof. 任取 $v_h \in V_h$ 。由强制性和 Galerkin 正交性,

$$\begin{aligned}\alpha \|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\ &= a(u - u_h, u - v_h) \\ &\leq M \|u - u_h\|_V \|u - v_h\|_V.\end{aligned}$$

若 $u = u_h$, 结论显然; 否则约去误差范数, 再对 v_h 取下确界。 $\square$

一般 Petrov–Galerkin 估计需要利用三角不等式, 对应的常数是 $1 + M/\beta$ 。在强制 Galerkin 情形中, 试探空间与检验空间相同, 可以把误差本身代入强制性估计, 并利用正交性消去离散分量, 从而得到更好的常数 M/α 。

24 混合问题

前一节讨论了单个变分方程的连续与离散理论。引入约束条件和 Lagrange 乘子后, 未知量分属两个空间, 适定性需要同时控制约束算子和它的核空间。本节先讨论连续混合问题, 再讨论离散混合问题及其鞍点矩阵形式。

24.1 连续混合问题

许多约束问题和混合有限元方法产生如下鞍点问题: 给定 $f \in V^*$ 和 $g \in Q^*$, 求 $(u, p) \in V \times Q$, 使得

$$\begin{aligned}a(u, v) + b(v, p) &= f(v), \quad \forall v \in V, \\ b(u, q) &= g(q), \quad \forall q \in Q.\end{aligned}$$

这里 V, Q 是 Hilbert 空间, $a : V \times V \rightarrow \mathbb{R}$ 和 $b : V \times Q \rightarrow \mathbb{R}$ 是连续双线性型。记

$$Z = \{v \in V : b(v, q) = 0, \forall q \in Q\}.$$

定义 $A \in \mathcal{L}(V, V^*)$ 、 $B \in \mathcal{L}(V, Q^*)$ 和 $B^* \in \mathcal{L}(Q, V^*)$:

$$\langle Au, v \rangle = a(u, v), \quad \langle Bu, q \rangle = \langle B^*q, u \rangle = b(u, q).$$

则混合问题可以写成算子方程

$$\begin{bmatrix} A & B^* \\ B & 0 \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}.$$

Remark 24.1. 这里讨论的是经典的鞍点问题: 主变量的试探空间和检验空间同为 V , 乘子的试探空间和检验空间同为 Q , 并且两个约束项由同一个双线性型 b 给出。因此算子矩阵中的两个非对角块互为伴随。更一般地, 可以分别取主变量和乘子的试探空间 U, P 以及检验空间 V, Q , 并考虑

$$\begin{aligned}a(u, v) + b(v, p) &= f(v), \quad \forall v \in V, \\ c(u, q) &= g(q), \quad \forall q \in Q,\end{aligned}$$

其中 $a : U \times V \rightarrow \mathbb{R}$ 、 $b : V \times P \rightarrow \mathbb{R}$ 、 $c : U \times Q \rightarrow \mathbb{R}$ 。这个问题可以直接看成乘积空间 $U \times P$ 与 $V \times Q$ 之间的 Petrov–Galerkin 变分问题, 并用 BNB 定理讨论适定性。此时两个约束算子不必互为伴随, 适定性条件也不能只用单个核空间 Z 表示。

Theorem 24.1 (混合问题的一般适定性条件). 上述混合问题适定, 当且仅当以下三个条件成立:

1. 存在 $\alpha > 0$, 使得

$$\inf_{0 \neq u \in Z} \sup_{0 \neq v \in Z} \frac{|a(u, v)|}{\|u\|_V \|v\|_V} \geq \alpha;$$

2. 若 $v \in Z$ 满足 $a(u, v) = 0$ 对所有 $u \in Z$ 成立, 则 $v = 0$;

3. 存在 $\beta > 0$, 使得

$$\inf_{0 \neq q \in Q} \sup_{0 \neq v \in V} \frac{|b(v, q)|}{\|v\|_V \|q\|_Q} \geq \beta.$$

此时存在与 f, g 无关的常数 $C > 0$, 使得唯一解 (u, p) 满足

$$\|u\|_V + \|p\|_Q \leq C(\|f\|_{V^*} + \|g\|_{Q^*}).$$

前两个条件正是 $a|_{Z \times Z}$ 的 BNB 条件. 第三个条件称为 Ladyzhenskaya–Babuška–Brezzi 条件, 简称 LBB 条件; 它要求所有 $q \in Q$ 都能被 V 中的检验函数一致地检测.

Proof. LBB 条件等价于 $B : V \rightarrow Q^*$ 满射, 并给出 $B^* : Q \rightarrow V^*$ 的有下界估计. 因而对任意 $g \in Q^*$, 可以找到 $u_g \in Z^\perp$, 使得

$$Bu_g = g, \quad \|u_g\|_V \leq \frac{1}{\beta} \|g\|_{Q^*}.$$

前两个条件正是双线性型 $a|_{Z \times Z}$ 的 BNB 条件. 因此存在唯一的 $z \in Z$, 使得

$$a(z, w) = f(w) - a(u_g, w), \quad \forall w \in Z.$$

令 $u = z + u_g$, 并定义 $\ell = f - Au \in V^*$. 由上式, ℓ 在 Z 上为零. LBB 条件给出

$$\text{Ran}(B^*) = Z^\circ := \{\ell \in V^* : \ell|_Z = 0\},$$

所以存在唯一的 $p \in Q$ 使得 $B^*p = \ell$. 于是 (u, p) 是混合问题的唯一解.

下面写出稳定性估计. 记 M_a 为 a 的连续性常数. 由 Z 上的 BNB 条件和 a 的连续性,

$$\begin{aligned} \|z\|_V &\leq \frac{1}{\alpha} \|f|_Z - a(u_g, \cdot)|_Z\|_{Z^*} \\ &\leq \frac{1}{\alpha} (\|f\|_{V^*} + M_a \|u_g\|_V) \\ &\leq \frac{1}{\alpha} \|f\|_{V^*} + \frac{M_a}{\alpha\beta} \|g\|_{Q^*}. \end{aligned}$$

因为 $u = z + u_g$, 所以

$$\|u\|_V \leq \frac{1}{\alpha} \|f\|_{V^*} + \left(\frac{1}{\beta} + \frac{M_a}{\alpha\beta}\right) \|g\|_{Q^*}.$$

再由 $B^*p = f - Au$ 和 LBB 条件,

$$\beta \|p\|_Q \leq \|B^*p\|_{V^*} \leq \|f\|_{V^*} + M_a \|u\|_V.$$

合并以上两个估计, 便得到定理中的结论.

反过来, 假设混合问题适定。对任意 $g \in Q^*$, 方程的解满足 $Bu = g$, 因此 B 满射; 由闭值域理论得到第三个条件。现在任取 $\varphi \in Z^*$, 由 Hahn-Banach 定理将它延拓为 $f \in V^*$, 并取 $g = 0$ 。对应的解满足 $u \in Z$, 而对所有 $v \in Z$ 有

$$a(u, v) = f(v) = \varphi(v).$$

所以 a 在 Z 上诱导的算子 $A_Z : Z \rightarrow Z^*$ 满射。若 $A_Z u = 0$, 则 Au 在 Z 上为零, 故 $Au \in Z^\circ = \text{Ran}(B^*)$ 。可取 $p \in Q$ 使 $Au + B^*p = 0$; 又因为 $u \in Z$, 有 $Bu = 0$ 。混合问题的唯一性推出 $u = p = 0$, 所以 A_Z 也是单射。对 A_Z 应用 BNB 定理, 便得到前两个条件。□

一般适定性定理要求 $a|_{Z \times Z}$ 满足两个 BNB 条件。由于 $Z = \ker(B)$ 是 V 的闭子空间, 它本身也是 Hilbert 空间。我们还可以像从 BNB 定理过渡到 Lax-Milgram 定理一样, 用核空间上的强制性同时保证这两个条件。若存在 $\alpha > 0$, 使得

$$a(v, v) \geq \alpha \|v\|_V^2, \quad \forall v \in Z,$$

则对任意 $0 \neq u \in Z$, 在上确界中取 $v = u$, 得到

$$\sup_{0 \neq v \in Z} \frac{|a(u, v)|}{\|v\|_V} \geq \frac{a(u, u)}{\|u\|_V} \geq \alpha \|u\|_V.$$

这给出第一个 BNB 条件。若 $v \in Z$ 满足 $a(u, v) = 0$ 对所有 $u \in Z$ 成立, 则取 $u = v$, 由强制性得到 $v = 0$, 所以伴随非退化条件也成立。这里仍不要求 a 对称。由此得到下面较容易验证的充分条件。

Theorem 24.2 (Brezzi 适定性条件). 假设存在 $\alpha, \beta > 0$, 使得

$$a(v, v) \geq \alpha \|v\|_V^2, \quad \forall v \in Z,$$

以及

$$\inf_{0 \neq q \in Q} \sup_{0 \neq v \in V} \frac{|b(v, q)|}{\|v\|_V \|q\|_Q} \geq \beta.$$

则鞍点问题存在唯一解, 并且存在只依赖于连续性常数、 α 和 β 的常数 C , 使得

$$\|u\|_V + \|p\|_Q \leq C(\|f\|_{V^*} + \|g\|_{Q^*}).$$

Proof. 核空间上的强制性保证 $a|_{Z \times Z}$ 满足一般适定性定理中的两个 BNB 条件, 而第二个假设正是 LBB 条件。因此, 一般混合问题的适定性定理直接给出解的存在唯一性和稳定性估计。□

24.2 离散混合问题

取有限维子空间 $V_h \subset V$ 和 $Q_h \subset Q$, 并记

$$Z_h = \{v_h \in V_h : b(v_h, q_h) = 0, \forall q_h \in Q_h\}.$$

离散混合问题是求 $(u_h, p_h) \in V_h \times Q_h$, 使得

$$a(u_h, v_h) + b(v_h, p_h) = f(v_h), \quad \forall v_h \in V_h,$$

$$b(u_h, q_h) = g(q_h), \quad \forall q_h \in Q_h.$$

Theorem 24.3 (离散混合问题的适定性). 设 a 和 b 的连续性常数与 h 无关. 假设存在与 h 无关的常数 $\alpha, \beta > 0$, 使得

$$a(v_h, v_h) \geq \alpha \|v_h\|_V^2, \quad \forall v_h \in Z_h, \quad \forall h,$$

以及

$$\inf_{0 \neq q_h \in Q_h} \sup_{0 \neq v_h \in V_h} \frac{|b(v_h, q_h)|}{\|v_h\|_V \|q_h\|_Q} \geq \beta, \quad \forall h.$$

则每个离散混合问题都有唯一解, 并且存在与 h 无关的常数 C , 使得

$$\|u_h\|_V + \|p_h\|_Q \leq C (\|f\|_{V_h^*} + \|g\|_{Q_h^*}).$$

Proof. 对每个固定的 h , 把 Brezzi 定理应用于 $V_h \times Q_h$. 两个假设分别给出 a 在离散核空间 Z_h 上的强制性和 b 的离散 LBB 条件, 因而离散问题存在唯一解. Brezzi 定理中的稳定常数只依赖于 a, b 的连续性常数以及 α, β . 这些常数均与 h 无关, 所以所得估计对整个离散问题族一致成立. $\square$

这里必须使用离散核空间 Z_h , 而不是连续核空间 Z . 一般只有 $Z \cap V_h \subset Z_h$; 离散检验空间 Q_h 不能检测 Q_h 之外的约束, 所以 Z_h 可能比 $Z \cap V_h$ 更大. 连续 LBB 条件也不会自动推出离散 LBB 条件, 这正是混合有限元空间需要单独验证的稳定性条件.

24.3 鞍点矩阵

混合问题对应的有限维版本就是常见的鞍点矩阵问题. 设 $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times n}$, 其中 $m < n$. 考虑鞍点矩阵方程

$$\mathcal{A} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} F \\ G \end{bmatrix}, \quad \mathcal{A} = \begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix}.$$

它的齐次方程为

$$Ax + B^\top p = 0, \quad Bx = 0.$$

第二个方程把 x 限制在 $\ker(B)$ 中; 第一个方程中的 $B^\top p$ 则用于补足 A 在这个核空间之外的分量.

根据 LBB 条件, 定义

$$\beta_B := \inf_{0 \neq p \in \mathbb{R}^m} \sup_{0 \neq x \in \mathbb{R}^n} \frac{|p^\top Bx|}{\|x\|_2 \|p\|_2}.$$

固定 $p \neq 0$, 有

$$\sup_{0 \neq x \in \mathbb{R}^n} \frac{|p^\top Bx|}{\|x\|_2} = \sup_{0 \neq x \in \mathbb{R}^n} \frac{|(B^\top p)^\top x|}{\|x\|_2} = \|B^\top p\|_2.$$

因此

$$\beta_B = \inf_{0 \neq p \in \mathbb{R}^m} \frac{\|B^\top p\|_2}{\|p\|_2} = \sigma_{\min}(B).$$

所以 $\beta_B > 0$ 当且仅当 $B^\top$ 单射, 也就是 B 行满秩. 因此 LBB 条件在有限维情形下实质就是要求 B 行满秩.

Theorem 24.4 (鞍点矩阵可逆性的核空间判别). 上述鞍点矩阵 $\mathcal{A}$ 可逆, 当且仅当以下两个条件成立:

1. B 行满秩;

2. 任取列满秩矩阵 $Z \in \mathbb{R}^{n \times (n-m)}$ 满足

$$\text{Ran}(Z) = \ker(B),$$

约化矩阵 $Z^\top A Z$ 可逆。

Proof. 先设两个条件成立，并考虑齐次方程

$$Ax + B^\top p = 0, \quad Bx = 0.$$

由 $Bx = 0$ ，存在 $\xi \in \mathbb{R}^{n-m}$ 使得 $x = Z\xi$ 。第一个方程左乘 $Z^\top$ ，并利用 $Z^\top B^\top = 0$ ，得到

$$Z^\top A Z \xi = 0.$$

由 $Z^\top A Z$ 可逆可知 $\xi = 0$ ，所以 $x = 0$ 。此时 $B^\top p = 0$ 。因为 B 行满秩， $B^\top$ 单射，故 $p = 0$ 。因此 $\ker(\mathcal{A}) = \{0\}$ ，有限维方阵 $\mathcal{A}$ 可逆。

反过来，设 $\mathcal{A}$ 可逆。若 B 不满行秩，则存在 $0 \neq p \in \ker(B^\top)$ ，于是

$$\mathcal{A} \begin{bmatrix} 0 \\ p \end{bmatrix} = 0,$$

与 $\mathcal{A}$ 可逆矛盾。因此 B 行满秩。

若 $Z^\top A Z$ 不可逆，则存在 $0 \neq \xi \in \ker(Z^\top A Z)$ 。令 $x = Z\xi$ ，则 $x \neq 0$ 、 $Bx = 0$ ，并且

$$Z^\top A x = 0.$$

因为 $\text{Ran}(Z) = \ker(B)$ ，上式说明

$$Ax \in (\ker B)^\perp = \text{Ran}(B^\top).$$

因此存在 $p \in \mathbb{R}^m$ ，使得 $Ax + B^\top p = 0$ 。于是

$$\mathcal{A} \begin{bmatrix} x \\ p \end{bmatrix} = 0.$$

由于 $x \neq 0$ ，这又与 $\mathcal{A}$ 可逆矛盾。所以 $Z^\top A Z$ 必须可逆。 $\square$

这里的第一个条件对应 LBB 条件，第二个条件对应双线性型限制到核空间后满足 BNB 条件。在有限维中，二者分别化为矩阵的行满秩和约化矩阵可逆。同样可以给出 Brezzi 条件对应的推论。

Corollary 24.5 (Brezzi 条件的有限维形式). 设 B 行满秩，并且存在 $\alpha > 0$ ，使得

$$x^\top A x \geq \alpha \|x\|_2^2, \quad \forall x \in \ker(B).$$

则鞍点矩阵 $\mathcal{A}$ 可逆。

Proof. 取列向量为 $\ker(B)$ 的一组标准正交基的矩阵 Z 。对任意 $0 \neq \xi \in \mathbb{R}^{n-m}$ ，有

$$\xi^\top Z^\top A Z \xi = (Z\xi)^\top A (Z\xi) \geq \alpha \|Z\xi\|_2^2 = \alpha \|\xi\|_2^2.$$

因此 $Z^\top A Z$ 可逆，由核空间判别可知 $\mathcal{A}$ 可逆。 $\square$

Corollary 24.6 (对称正定情形与 Schur 补). 在上述记号下, 若 A 对称正定, 则鞍点矩阵 $\mathcal{A}$ 可逆, 当且仅当 B 行满秩。

Proof. 若 $\mathcal{A}$ 可逆而 B 不满行秩, 则存在 $0 \neq p \in \ker(B^\top)$, 从而 $(0, p)$ 是 $\mathcal{A}$ 的非零核向量, 矛盾。因此 B 必须行满秩。

反过来, 设 B 行满秩。由于 A 对称正定, A^{-1} 也对称正定, 消去第一个方程中的 u 后得到关于 p 的 Schur 补方程, 其系数矩阵为

$$S = BA^{-1}B^\top.$$

相应的消元过程可以写成 Schur 分解

$$\mathcal{A} = \begin{bmatrix} I & 0 \\ BA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & -S \end{bmatrix} \begin{bmatrix} I & A^{-1}B^\top \\ 0 & I \end{bmatrix}.$$

两个三角因子都可逆, 因此 $\mathcal{A}$ 可逆当且仅当 S 可逆。对任意 $0 \neq p \in \mathbb{R}^m$, $B^\top p \neq 0$, 并且

$$p^\top Sp = (B^\top p)^\top A^{-1}(B^\top p) > 0.$$

所以 S 对称正定, 进而可逆。先解出 p , 再由 $u = A^{-1}(f - B^\top p)$ 解出 u , 故 $\mathcal{A}$ 可逆。 □