前　言

为适应充实高层科技人才培养的需要，必然地对工程及经济科学的许多专业的数学教育提出了更高的要求。由于泛函分析在现代数学科学中有着重大的理论和应用前景而成为基础课程之一。它是综合地运用几何、代数以及分析学的方法和观点来分析处理问题而形成的数学理论，因而内容极其丰富，体系更加系统、严谨，观点也尤为深刻。已成为探讨和解决现代科学技术中众多问题的一门重要的基础知识。因此，国内外许多大学（特别是著名大学）纷纷把泛函分析这门课作为理工科硕士生和博士生的必修课。

从 1996 年开始，作者在哈尔滨工业大学先后三次为硕士生和博士生编写应用泛函分析方面的教材，在校内曾试用多次。本书是在 2001 年 7 月编写的校内教材基础上经过详细修改而定稿的。此外，考虑到这门课程对博士生与硕士生的教学内容衔接问题，在原有内容的基础上，又增加了两章新内容，供博士生选用。

鉴于工科各专业研究生的大学数学基础存在差异，本书给出了大学数学方面的某些必要的基础知识，可起到承上启下的作用。对于工科学生来说，学习数学的目的是应用，但前提是要掌握基本理论。理论没有学到手，就无从去谈应用。本书编写过程中作者试图做到理论与实际相结合，特别与学生熟悉的数学知识相联系。为加强对基本理论和方法的理解和认识，书中每节都配有一定数量的习

题，供学生选做。

本书前四章的内容是供硕士生用的，大约 50 学时可讲授完毕；后两章的内容是供博士生选用的，大约 20 学时可讲授完毕。

在本书的编写和出版过程中得到吴从炘、李容录、张传义、付永强、邓廷权老师的关心、帮助，还得到了哈尔滨工业大学研究生院领导的支持和帮助，也得到了有关老专家的热情鼓励和支持。作者借此机会，向他们表示深切谢意。

写出一本高质量的有鲜明特色的教材是一个相当困难的工作。虽然我们历经数年艰苦的努力，但限于学识和经验，书中难免存在疏漏和不足，希望专家同行和同学们批评指正。

作　者

2002 年 6 月于哈尔滨工业大学

目　　录

第 1 章　预备知识

泛函分析是现代数学科学中重要分支之一，其内容涉及无穷维线性空间及其上的算子的基本理论，并且综合运用代数、几何与分析等经典学科中的观点和方法。为学好泛函分析，首先介绍实数空间及其上的函数的有关理论是十分必要的。

1.1　集合的一般知识

1.1.1　集及其运算

集合是现代数学的一个基本概念，依通常的观点，把具有一定性质或满足一定条件的对象的全体叫做集合，或简称为集，其中的每个对象叫做该集合的元或元素。本书常用大写字母表示集，用小写字母表示元素。

设 A,B 是集。元 a 属于集 A，记为 $a \in A$，而记号 $b \notin A$ 表示元素 b 不属于集 A。不包含任何元素的集合称为空集，记为 $\varnothing$。如果集 A 的每个元素都属于集 B，则说 A 为 B 的子集，记为 $A \subset B$ 或 $B \supset A$。规定空集 $\varnothing$ 是任何集的子集。如果 $A \subset B$，且 $B \subset A$，则称 A 与 B 相等，记成 $A = B$。若 $A \subset B$，但 $A \neq B$ 时，则称 A 为 B 的真子集。

若集 A 为具有某性质 π 的元素所构成时，我们常表示作 $A = \{x:x$ 具有性质 $\pi\}$，例如闭区间 $[a,b] = \{x:a \leqslant x \leqslant b\}$。

以下介绍集的运算。

设 A,B 是两个集，由集 A 及集 B 的全体元素构成的集叫做 A 与 B 的并，记为 $A\cup B$，即

$$A\cup B=\{x:x\in A\text{ 或 }x\in B\}。$$

所有同时属于 A 与 B 的元素组成的集叫做 A 与 B 的交，记为 $A\cap B$，即

$$A\cap B=\{x:x\in A\text{ 且 }x\in B\}。$$

并集与交集的概念可以推广到任意个集的情形。设 $\{A_i:i\in I\}$ 是一集族，其中 I 是指标集，i 在 I 中变化，则它们的并与交分别定义为

$$\bigcup_{i\in I}A_i=\{x:\exists\, i\in I,\text{有 }x\in A_i\},$$

$$\bigcap_{i\in I}A_i=\{x:\forall\, i\in I,\text{有 }x\in A_i\}。$$

不难证明并与交运算具有如下性质：

(1) $A\cup\left(\bigcap_{i\in I}A_i\right)=\bigcap_{i\in I}(A\cup A_i)$，

(2) $A\cap\left(\bigcup_{i\in I}A_i\right)=\bigcup_{i\in I}(A\cap A_i)$。

现在我们引入集合间的另一种运算：设 A,B 是两个集，由集 A 中不属于集 B 的那些元素组成的集，叫做 A 与 B 的差，记为 $A\setminus B$。特别，当 $B\subset A$ 时，常称差 $A\setminus B$ 为 B 关于 A 的余集或补集，记为 $\complement_A B$。如果我们仅考察某固定的集 A 的一些子集 B 时，则常记 $A\setminus B$ 为 B^C。以下的运算法则叫做 De Morgan 律：

(1) $\left(\bigcup_{i\in I}A_i\right)^C=\bigcap_{i\in I}A_i^C$，

(2) $\left(\bigcap_{i\in I}A_i\right)^C=\bigcup_{i\in I}A_i^C$。

它的证明留给读者去完成。这个法则为我们提供了一种对偶方法，可将集合某性质的证明转换到它的补集上去。

以下要介绍的直积概念（也叫 Cartesian 积）是集的重要运算之一。设 A,B 是给定的集，则全体有序对 (x,y) 所成之集叫做 A 与 B 的

有限个集合上去:对于集组$\{A_k: k = 1,2\cdots n\}$,它们的直积为

$$\prod_{k=1}^{n} A_k = \{(x_1, x_2, \cdots, x_n): x_k \in A_k, k = 1,2,\cdots,n\}。$$

1.1.2　映射

自然界各个对象间的相关性是科学关注的重要内容。抽象为数学问题就是要考察集合间元素的彼此联系。其中最重要的基本概念之一是关系(这里介绍的是二元关系)。设A,B是两个集,所谓A与B的一个关系R是直积$A \times B$的一个子集,即$R \subset A \times B$,此时我们用xRy表示$(x,y) \in R$,并且说在R意义下x与y相关。而集$\{x \in A: \exists y \in B$,使$(x,y) \in R\}$与集$\{y \in B: \exists x \in A$,使$(x,y) \in R\}$分别称为$R$的定义域或值域。例如,数学分析中讲过的函数$f$就确定了数集$A$与数集$B$之间的一个关系:$\{(x,f(x)): x \in A\} \subset A \times B$。

现在我们把函数概念一般化:设A,B是两个非空集,如果依某法则T,对每个$x \in A$,在B中有惟一确定的元素y与之对应,则说T是定义在A上且取值于B内的一个映射,记为$T: A \to B$,并将x与y的关系记成$y = T(x)$。此时A为T的定义域,而值域(或象)为$\{T(x): x \in A\}$,可记成$T(A)$,一般情形,它是B的一个子集。特别$B = T(A)$时,则称T是满射或映上的,即T为由A到B上的映射。又如果对每个$y \in T(A)$惟一存在$x \in A$,有$T(x) = y$,则说T是A上一个单映射或说T是由A到$T(A)$上的一对一映射,此时我们看到T存在逆映射$T^{-1}: T(A) \to A$,有$T^{-1}(y) = x, \forall y \in T(A)$。

若映射$I_A: A \to A$满足条件:$I_A(x) = x, \forall x \in A$,则称$I_A$为$A$上的恒等映射(或单位映射),在不致引起混淆时,可简记I_A为I。

设给定两个映射$T: A \to B$,$S: B \to C$,则记号ST表示由A到C内的映射,为$ST(x) = S(T(x))$,$x \in A$,并称为T与S的乘积(或复

合）。

设 T 为由 A 到 B 内的一个映射，若 $E \subset B$，引入记号 $T^{-1}(E) = \{x \in A : T(x) \in E\}$，它称为 E 在 T 映射下的原象（或逆象）。请注意这里并不表明 T 存在逆映射 T^{-1}。不难验证以下的关系式成立。

(1) $T(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} T(A_i)$，其中 $\{A_i : i \in I\}$ 是 A 的子集族；

(2) $T^{-1}(\bigcap_{i\in I} B_i) = \bigcap_{i\in I} T^{-1}(B_i)$，其中 $\{B_i : i \in I\}$ 是 B 的子集族。

以下引入另一个重要关系是等价关系：设 A 为给定的一个集，若关系 $R \subset A \times A$ 满足如下条件：

(1) 任 $x \in A$，有 $(x, x) \in R$，即 R 有自反性；

(2) 如果 $(x, y) \in R$，则 $(y, x) \in R$ 即 R 有对称性；

(3) 若 $(x, y) \in R$，$(y, z) \in R$，则 $(x, z) \in R$，也就是说 R 具有传递性。这时，称 R 为 A 上一个等价关系，此时，若 $(x, y) \in R$，可记为 $x \sim y$。

例如任何集上的相等概念就是一个等价关系。

现在我们给出重要概念：

【定义 1】 设 A，B 是两个集合，如果存在由 A 到 B 上的一对一映射 T，则说 A 与 B 成一一对应或称 A 与 B 对等，记为 $A \sim B$。

1.1.3 可列集

自 19 世纪 70 年代以来，由德国数学家 Cantor 所开创的无穷集合理论，已成为现代数学的基础。根据两个集合所包含元素的多少程度来决定其区别，他的基本方法就是利用对等概念。以后我们常用 **N** 表示自然数集，**Q** 表示有理数集，**Z** 表示整数集，**R** 表示实数集，**C** 表示复数集。它们都是常用的无限集。

当有人抬一筐苹果到某大教室内，会提出这样一个问题：教室中学生人数与筐中的苹果数哪个大？回答很简单，只要规定每个学生从

筐中只准拿走一个苹果，就可得到正确的判断。如果每个学生都拿到了苹果，但筐中仍剩有苹果时，则说明筐中苹果比教室中学生多；如果筐中苹果已拿光，而有学生没有拿到苹果，则自然是学生人数比筐中苹果数大；若每个学生都拿到了苹果，且筐中苹果都被拿光，明显得知学生数等于苹果数。请注意，以上的回答利用了对应的方法。由一一对应的概念可知，如果两个有限集互相对等时，则它们所含元素个数必然相同。然而对于无限集利用一一对应所导致的结果却是十分有趣的。例如全体正偶数集依方式 $2n \leftrightarrow n$，则与 $\mathbf{N}$ 成一一对应，即与 $\mathbf{N}$ 是对等的。明显前者是后者的真子集，因此对无限集，元素的个数已不能说明无限集所含的元素多少程度，而须引入更一般的概念。

【定义2】 设 A，B 是两个集，

(1) 若 A 与 B 对等，则说 A 与 B 有相同的势（或基数），记 A 的势为 $\bar{\bar{A}}$，B 的势为 $\bar{\bar{B}}$，则 $\bar{\bar{A}} = \bar{\bar{B}}$。

(2) 若 A 与 B 的某子集对等，则说 A 的势不大于 B 的势，记为 $\bar{\bar{A}} \leqslant \bar{\bar{B}}$（或 $\bar{\bar{B}} \geqslant \bar{\bar{A}}$），又若 B 不与 A 的任一子集对等时，则说 A 的势小于 B 的势，记为 $\bar{\bar{A}} < \bar{\bar{B}}$（或 $\bar{\bar{B}} > \bar{\bar{A}}$）。

注 (1) 势是集合元素个数的一般化，当 A 为有限集时，则 $\bar{\bar{A}}$ 就是 A 中元素的个数。

(2) 任有限集不能同自己的任真子集对等，即任真子集的势小于本身的势。

但无限集具有如下特征：

【定理1】 无限集必与它的某真子集对等。

证 设 A 为一个无限集，取出一个元素 $a_1 \in A$，因 A 为无限集，则 $A - \{a_1\} \neq \varnothing$，又从 $A - \{a_1\}$ 中取出一个元素 a_2，显然 $A - \{a_2, a_2\} \neq \varnothing$，此步骤可无限做下去，我们从 A 中取出一列元素 $\{a_1, a_2, a_3 \cdots\}$，记余集为 $B = A \setminus \{a_1, a_2, a_3 \cdots\}$，则集 $C = B \cup \{a_2, a_3, \cdots\}$

为 A 的真子集。定义映射 $T: A \to C$ 为 $T(a) = a$，当 $a \in B$ 时，$T(a_k) = a_{k+1}$　$(k = 1,2,3\cdots)$，则明显看出，T 为 A 到 C 上的一对一映射，证毕。

注　无限集可定义为可与自己的某真子集对等的集。

由于自然数集 **N** 是最简单的无限集，因此特别有如下概念：

【定义 3】　凡与 **N** 对等的集称为可列集(或可数集)。

由定义，若 A 是可列集，则 A 中全体元素可表示成无穷序列的形式：$A = \{x_1, x_2, x_3, \cdots, x_n, \cdots\}$。例如，整数集 $\mathbf{Z} = \{0, 1, -1, 2, -2, \cdots, n, -n, \cdots\}$ 及三角函数系 $\{1, \cos x, \sin x, \cos 2x, \cdots, \cos nx, \sin nx, \cdots\}$ 都是可列集。

为得到更多稍为复杂的可列集的实例，先考察其某些性质。

【定理 2】　任一无限集必含有可列子集。

可用类似于定理 1 的证明得到，留给读者练习。另外，也请读者证明如下事实：可列集的子集，若不是有限集，则一定为可列集。

【定理 3】　有限或可列个可列集的并集仍是可列集；可列个有限集的并(若为无限集)，也是可列集。

证　不失一般性，对可列集族(也称集列)$\{A_1, A_2, A_3, \cdots, A_n, \cdots\} = \{A_n : n \in N\}$，其中每个都是可列集且两两不相交，于是有

$$\begin{array}{l} A_1 = \{a_{11}, a_{12}, a_{13}, a_{14}, \cdots\} \\ A_2 = \{a_{21}, a_{22}, a_{23}, a_{24}, \cdots\} \\ A_3 = \{a_{31}, a_{32}, a_{33}, a_{34}, \cdots\} \\ A_4 = \{a_{41}, a_{42}, a_{43}, a_{44}, \cdots\} \\ \cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots \end{array}$$

我们依对角线原则(箭头所示)可把并 $\bigcup\limits_{n=1}^{\infty} A_n$ 中全部元素排成列：

$$a_{11}, a_{21}, a_{12}, a_{31}, a_{22}, a_{13}, a_{41}, a_{32}, a_{23}, a_{14}, \cdots$$

故它是可列集。

由此我们可以得到一个有趣的例子：有理数集 $\mathbf{Q}$ 是可列集。

事实上，因每个有理数 r 都可写成既约分数 p/q，其中 p 与 q 皆为整数，且规定 $q \in \mathbf{N}$。对每个固定的 $q \in \mathbf{N}$，$A_q = \{p/q : p \in \mathbf{Z}\}$ 是一可列集。明显 $\mathbf{Q} = \bigcup\limits_{q=1}^{\infty} A_q$ 为可列个可列集的并，故 $\mathbf{Q}$ 可列。

【定理 4】　设有一组可列集 $\{A_1, A_2, \cdots, A_n\}$，则直积 $\prod\limits_{k=1}^{n} A_k$ 也是可列集。

证　只须证明 $r = 2$ 情形，一般情况可由归纳法获得。设 $A_1 = \{x_1, x_2, x_3, \cdots\}$，$A_2 = \{y_1, y_2, y_3, \cdots\}$，则 $A_1 \times A_2 = \bigcup\limits_{k=1}^{\infty}\{(x_k, y_m) : m \in \mathbf{N}\}$，为可列个可列集之并，故可列。

例 1　有理系数多项式全体成为可列集。

事实上，容易知道零次有理系数多项式即是有理数，因而其全体是有理数集 $\mathbf{Q}$，故可列。对每个固定的 $n \in \mathbf{N}$，我们证明全体次数 $\leqslant n$ 的有理系数多项式的集合可列。任取次数 $\leqslant n$ 有理系数多项式 $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = P_n(x)$，则其与直积集 $\prod\limits_{k=1}^{n+1} \mathbf{Q}_k$ 中元素 $(a_0, a_1, a_2, \cdots, a_n)$ 一一对应，其中 $\mathbf{Q}_k = \mathbf{Q}$，$(k = 1, 2, \cdots, n+1)$，故推得次数 $\leqslant n$ 有理系数全体可列。从而有理系数多项式全体

$\bigcup\limits_{n=0}^{\infty}\{a_0x^n + a_1x^{n-1} + \cdots a_n : a_k \in \mathbf{Q}, k = 0, 1, 2, \cdots, n\}$ 是可列集。

但存在不可列无限集。

【定理 5】　点集 $(0,1) = \{x \in \mathbf{R} : 0 < x < 1\}$ 是不可列的无限集。

证　我们证明 $(0,1)$ 不可列。假设 $(0,1)$ 可列，则其中全体实数可排成一列 $x_1, x_2, x_3, \cdots$。将每个 x_k 用十进位无限小数表示，则有

$$x_1 = 0.t_{11}t_{12}t_{13}t_{14}\cdots$$
$$x_2 = 0.t_{21}t_{22}t_{23}t_{24}\cdots$$
$$x_3 = 0.t_{31}t_{32}t_{33}t_{34}\cdots$$
$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

其中所有的 t_{ij} 都是在 $0,1,2,\cdots,9$ 中取值，并且对每个 i，数列 $\{t_{ij}: j = 1,2,\cdots\}$ 应有无限个不为 0，即如小数 0.320 00… 应改写为 0.319 9…。

作十进位小数

$$x = 0.b_1b_2b_3\cdots,$$

当 $t_{ii} = 1$ 时，令 $b_i = 2$；而若 $t_{ii} \neq 1$，则令 $b_i = 1$。明显 $x \in (0,1)$，但因对每个 n，$b_n \neq t_{nn}$，故 $x \neq x_n$，这与假设矛盾，证毕。

注 集$(0,1)$的势称为连续点集的势，并且明显$\overline{\overline{(0,1)}}$大于可列集之势。

【定理 6】 设 A 为有限集或可列集，B 是任一无限集，则 $\overline{\overline{A \cup B}} = \overline{\overline{B}}$。

证 因 B 是无限集，由定理 2，存在可列子集 $\{x_1, x_2, \cdots\} = C \subset B$。不失一般性，设 $A \cap B = \varnothing$。由于 $A \cup C$ 仍是可列集，就是 $A \cup C \sim C$，再注意 $B - C \sim B - C$，可推得 $A \cup B = (A \cup C) \cup (B - C) \sim C \cup (B - C) = B$。即$\overline{\overline{A \cup B}} = \overline{\overline{B}}$。证毕。

由公式 $y = a + (b - a)x$，可建立点集$[0,1]$与点集$[a,b]$之间的一一对应，故任意区间$[a,b]$都是连续点集的势。而由公式 $y = \tan(\frac{2x-1}{2}\pi)$，又建立了$(0,1)$与实数集 $\mathbf{R}$ 之间的一一对应，因此 $\mathbf{R}$ 的势等同$(0,1)$的势。设无理数全体为 A，由定理 6 有

$$\overline{\overline{A}} = \overline{\overline{(A \cup \mathbf{Q})}} = \overline{\overline{\mathbf{R}}},$$

故无理数集的势为连续点集的势。据此，我们可得这样的认识：无理

数要比有理数多得多。

习题 1.1

1. 设映射 $T:A \to B$,试证明以下事实

(1) 若 A_1,A_2 都是 A 的子集,则

$T(A_1 \cap A_2) \subset T(A_1) \cap T(A_2)$,并举说明等号不一定成立;

(2) 若 $A_1 \subset A$,则 $T(A \setminus A_1) \supset T(A) \setminus T(A_1)$,并举例说明等号未必成立;

(3) 若 $A_1 \subset A$,则 $T^{-1}(T(A_1)) \supset A_1$,若 $B_1 \subset B$,则 $T(T^{-1}(B_1)) \subset B_1$,并举例说明等号不一定成立。

2. 对任何集 A,B,C,证明以下关系式成立

$$(A \cap B) \times C = (A \times C) \cap (B \times C)。$$

3. 设 A 为可列集,B 是由 A 的全体有限子集构成的集合,证明 B 是可列集。

4. 证明可列集 A 关于任意映射 T 的象至多为可列集。

5. 证明每个区间中的无理数集都不可列。

6. 证明以有理数为中心且以有理数为半径的区间的全体是可列集。

7. 建立闭区间 $[a,b]$ 与开区间 (a,b) 的一一对应。

8. 建立闭区间 $[a,b]$ 与全体实数集 $\mathbf{R}$ 的一一对应。

1.2　实数集的基本结构

关于数的加、减、乘、除等运算及其性质是代数学中的基本内容,

现引入概念。

【**定义1**】 设 P 是一些复数组成的集合，其中含有数0与1，若 P 中任两个数(也可相同)的和、差、积、商(除数不为0)仍属于 P，即 P 关于加、减、乘、除(有意义)运算是封闭的，则称 P 为一个数域。

显然，有理数集 $\mathbf{Q}$，实数集 $\mathbf{R}$ 及复数集 $\mathbf{C}$ 都是数域。关于实数域的这种代数结构，我们认为读者是熟悉的，以下研究实数集的另一种结构。

1.2.1 实数集的序关系

我们知道，实数集中两个数 x 与 y 的大小关系可用符号 $\leqslant$ 来说明，并且有如下的性质：

(1) 对任何 $x \in \mathbf{R}$，成立 $x \leqslant x$；

(2) 对 $x, y, z \in \mathbf{R}$，若 $x \leqslant y$，且 $y \leqslant z$，则 $x \leqslant z$；

(3) 对 $x, y \in \mathbf{R}$，若 $x \leqslant y$，且 $y \leqslant x$，则 $x = y$。

于是在 $\mathbf{R}$ 上由 $\leqslant$ 确定了一个序关系。因对 $\mathbf{R}$ 中任何两个数都可由关系 $\leqslant$ 联系，则 $\mathbf{R}$ 为全序数域。

在 $\mathbf{R}$ 的序结构基础上，引入以下概念。

对点集 $E \subset \mathbf{R}$，若存在 $y \in \mathbf{R}$，使得对任意 $x \in E$，恒成立 $x \leqslant y$，则称 y 为 E 的一个上界(此时可知 E 的全体上界组成的集合是一个无限集)。类似地，可以定义下有界及下界。设 $E \subset \mathbf{R}$，若它上有界且下有界，则称 E 为序有界，简称为有界，此时存在 $y, z \in \mathbf{R}$，有 $z \leqslant x \leqslant y$，$\forall x \in E$，即有 $E \subset [z, y]$。也不难得知 $E \subset R$ 有界的充要条件是存在 $y > 0$，成立 $E \subset (-y, y)$(亦可为闭区间)。如果点集 E 存在最大数，记为 $\max E$，立知它就是 E 的一个上界，且为 E 所有上界中最小者，但一般情况点集未必都存在最大数，于是有如下概念。

【**定义2**】 给定非空点集 E，若存在 $\beta \in \mathbf{R}$，满足以下条件：

(1) 对任一 $x \in E$,有 $x \leqslant \beta$;

(2) 对任给 $\varepsilon > 0$,至少存在一个 $x \in E$,有 $x > \beta - \varepsilon$,则称数 β 是点集 E 的上确界,记为 $\beta = \sup E$。

其中条件(1) 说明 β 是点集 E 的一个上界,而条件(2) 则说明它为 E 的最小的上界。

类似地,对于非空点集 E,若存在 $\alpha \in \mathbf{R}$,满足条件:

(1) 对任一 $x \in E$,有 $\alpha \leqslant x$;

(2) 对任给 $\varepsilon > 0$,至少存在一个 $x \in E$,有 $x < \alpha + \varepsilon$,则称 α 为点集 E 的下确界,记作 $\alpha = \inf E$。

我们规定,当 E 无上界时,记 $\sup E = \infty$;当 E 无下界时,则记 $\inf E = -\infty$。

关于上(下) 确界的存在性的回答是以下的基本定理,它刻画了实数域的重要特征。

【定理 1】 若非空点集 $E \subset \mathbf{R}$ 上有界,则必存在上确界 $\beta = \sup E$;同样,而下有界的点集 E 一定存在下确界 $\alpha = \inf E$。

它的证明可在较详细的数学分析教科书中找到。关于确界的惟一问题留作习题。

现在简要介绍一般的序概念,它是由实数集的大小关系抽象而得。

【定义 3】 对于非空集合 A 中某些元素对 x 与 y,定义了序的关系 $x \leqslant y$(比较关系),并且满足如下条件:

(1) 对任 $x \in A$,有 $x \leqslant x$;

(2) 若 $x \leqslant y$,且 $y \leqslant z$,则 $x \leqslant z$;

(3) 若 $x \leqslant y$,且 $y \leqslant x$,则 $x = y$,

则称 A 为半序集(关于序关系 $\leqslant$)。特别当 A 中任两个元素 x 与 y,关系式 $x \leqslant y$ 与 $y \leqslant x$ 至少有一个成立时,则称 A 为全序集。

例如，实数集 $\mathbf{R}$ 的所有子集构成的集合按包含关系 $A \subset B$ 定义序为 $A \leqslant B$，则成为半序集。

类似实数集，在半序集中可引入上(下)界概念及上(下)确界概念：设 B 为半序集 A 的子集，$a \in A$，若对任 $b \in B$，都有 $b \leqslant a$($a \leqslant b$)，则称 a 为 B 的一个上(下)界，且称 B 为 A 中上(下)有界的。如果 B 是 A 中上(下)有界子集，则其最小(大)的上(下)界(若存在)叫做 B 的上(下)确界，并记为 $\sup B$($\inf B$)。另外还有如下概念：

【定义 4】 设 A 为半序集，$a \in A$，若对任一 $x \in A$，都有 $x \leqslant a$，就称 a 为 A 中的最大元素。$a \in A$，若 $x \in A$，有 $a \leqslant x$，则必有 $x = a$，就称 a 为 A 的极大元素。我们指出，任何最大元素也是极大的，但反之一般不成立。当 a 为 A 的极大元素时，可能存在某些 $x \in A$ 与 a 不存在比较关系。一般情况半序集的极大元素不一定是惟一的，例如集 A 由两个元素 x 与 y 组成，规定序关系为 $x \leqslant x$，$y \leqslant y$，则 A 为半序集。此时 x 与 y 都是 A 的极大元素，但 A 无最大元素。

类似地有最小元素与极小元素的概念。

以下的引理在泛函分析的基本理论中常被用到。在集合论中是研究“无限过程”的一个逻辑工具(常叫做超限归纳法)，并且作为公理被承认。

引理(Zorn) 如果半序集 A 的每个全序子集都有上界，则在 A 中必存在极大元素。

1.2.2　实数集的度量结构

以微积分为核心的数学分析是建立在极限理论的基础之上的。而极限概念不仅使用了函数概念，还必然地引入几何学中的概念——距离。

在解析几何里，我们通过直线与实数集建立一一对应关系得到

数轴。设数轴上有两点 P 与 Q，它们的坐标分别是 x 与 y，则规定 P 与 Q 两点的距离为

$$\rho(P,Q)=|x-y|,$$

即由两个实数之差的绝对值给出了直线上两点距离概念，这具有直观上的意义，不难由绝对值的性质推知距离函数 $\rho(P,Q)$ 满足如下条件：

(1) 数轴上任两点 P 与 Q，有 $\rho(P,Q)\geqslant 0$，且 $\rho(P,Q)=0$，当且仅当 $P=Q$；

(2) 数轴上任两点 P 与 Q，有 $\rho(P,Q)=\rho(Q,P)$；

(3) 数轴上任三个点 P,Q 及 M，有 $\rho(P,Q)\leqslant\rho(P,M)+\rho(M,Q)$。

以距离概念为基础自然地引入了收敛点列的概念：点列 $x_n\in\mathbf{R}$ 收敛于点 $x\Leftrightarrow\forall\varepsilon>0,\exists N$，使当 $n\geqslant N$，有 $\rho(x_n,x)=|x_n-x|<\varepsilon$。此时记为 $x_n\to x(n\to\infty)$ 或 $\lim\limits_{n\to\infty}x_n=x$。

同实数集，也可在有理数集 $\mathbf{Q}$ 中引入距离及极限概念，但我们有例子：$r_1=1.4,r_2=1.41,r_3=1.414,\cdots$ 这样的有理数列 r_n，有 $\lim\limits_{n\to\infty}r_n=\sqrt{2}\notin\mathbf{Q}$，即 $\mathbf{Q}$ 关于极限运算是不封闭的。此时称 $\mathbf{Q}$ 是不完备的。那么实数集 $\mathbf{R}$ 是否完备呢？回答是肯定的。由上述的确界存在原理(定理 1)可证明实数集的完备性。以下列出的基本定理都同定理 1 等价，它们中的每一个都可以刻画实数集的完备性这个最基本的特征。

【定理 2】 (区间套引理)　设闭区间列 $[a_n,b_n]$ 满足条件：

(1) $[a_{n+1},b_{n+1}]\subset[a_n,b_n],n=1,2,3,\cdots$；

(2) $\lim\limits_{n\to\infty}(b_n-a_n)=0$，

则惟一存实数 $\xi\in[a_n,b_n],n=1,2,3,\cdots$。

注意引理中闭区间条件一般不能改为开区间。

【定理 3】 单调有界数列必有极限。

注意单调性概念是建立在实数集的序结构基础上：对于数列 $x_n \in \mathbf{R}$，若 $x_1 \leqslant x_2 \leqslant x_3 \cdots x_n \leqslant x_{n+1} \leqslant \cdots$（简记为 $x_n \leqslant x_{n+1}, n = 1, 2, \cdots$），则称 x_n 为单调增加数列，而若 $x_{n+1} \leqslant x_n, n \in \mathbf{N}$，则称 x_n 为单调减少数列。二者统称为单调数列。当序关系 $\leqslant$ 改为 $<$ 时，常称 x_n 为严格单调数列。

【定理 4】（Bolzano – Weierstrass 定理） 有界数列 $x_n \in \mathbf{R}$，必存在收敛子列。

由此定理可推出以下事实：若点列 x_n 无界，则存在子列 x_{n_k} 为无穷大量，即 $|x_{n_k}| \to \infty (k \to \infty)$。

【定理 5】（Heine – Borel 有限覆盖定理） 设 $[a, b]$ 为有界闭区间，对任一开区间族 $\{(a_i, b_i): i \in I\}$，只要 $[a, b] \subset \bigcup\limits_{i \in I} (a_i, b_i)$（此时，称此开区间族为 $[a, b]$ 的一个开覆盖），则从此开区间族中可选取出有限个开区间 $(a_{i_1}, b_{i_1}), (a_{i_2}, b_{i_2}), \cdots, (a_{i_n}, b_{i_n})$ 来覆盖 $[a, b]$，即 $[a, b] \subset \bigcup\limits_{k=1}^{n} (a_{i_k}, b_{i_k})$，且称这 n 个开区间为有限子覆盖。

定理中对闭区间所反映的这一重要属性引申出现代分析理论中一个重要概念，即紧集的概念，它在研究闭区间上的连续函数的许多重要的性质（如有界性，最大值及最小值的可达性及一致连续性等）时都有重要作用。

【定义 5】 若点集 E 具有如下性质：对 E 的任一个开覆盖，必存在一个有限子覆盖，则称 E 为实数集的一个紧集。

定理 5 告诉我们每个有界闭区间都是紧集。但却存在不是紧集的点集，例如每个开区间 (a, b) 都不是紧集。仅对 $a \in \mathbf{R}, b \in \mathbf{R}$ 时说明之，取自然数 N，使 $\dfrac{1}{N} < b - a$，则不难看到开区间族 $\left(a, b - \dfrac{1}{N}\right)$，

$\left(a,b-\frac{1}{N+1}\right),\left(a,b-\frac{1}{N+2}\right),\cdots\left(a,b-\frac{1}{n}\right),\cdots$ 是(a,b)的一个开覆盖，但它不存在有限子覆盖(其中 $n \geqslant N$)。

以下的定理给出了收敛点列的一个特征。

【定理6】 点列 x_n 收敛的充要条件是对任给 $\varepsilon>0$, $\exists N>0$, 当 $n,m \geqslant N$ 时，有 $|x_n - x_m| < \varepsilon$。

常称满足充分条件的点列 x_n 为Cauchy列(或称基本列)。一般称定理6为Cauchy收敛准则。它告诉了我们两个事实，一个是给出了判断点列是否收敛的方法，另一个则给出了实数集的一个重要特性，即实数集内每个Cauchy列一定收敛。实数集的这一特性称为完备性。它等价于实数集关于极限运算的封闭性。

关于以上定理的证明，建议读者找一本数学分析的教科书读一读，这对理解实数集的基本属性，加强分析理论的基础以及有效的思维方法的训练都是很有益处的。

介绍一个依赖于实数集结构的有用概念。

【定义6】 设 x_n 为实数序列，则定义 x_n 的上极限和下极限分别为

$$\overline{\lim_{n\to\infty}} x_n = \inf_{k\geqslant 1}(\sup_{n\geqslant k} x_n);$$
$$\underline{\lim_{n\to\infty}} x_n = \sup_{k\geqslant 1}(\inf_{n\geqslant k} x_n)。$$

注 (1) 令 $y_k = \sup_{n\geqslant k} x_n, k\in \mathbf{N}$，则明显 y_k 是单调减小序列，于是可记 $\overline{\lim_{n\to\infty}} x_n = \lim_{k\to\infty}(\sup_{n\geqslant k} x_n)$，同理有 $\underline{\lim_{n\to\infty}} x_n = \lim_{k\to\infty}(\inf_{n\geqslant k} x_n)$。

(2) 任给一点列 x_n，一般情况 $\lim_{n\to\infty} x_n$ 未必存在，但上(下)极限总有意义，上(下)极限可以是正或负无穷大(即 ∞ 或 $-\infty$)。

(3) $\lim_{n\to\infty} x_n$ 存在充要条件是 $\overline{\lim_{n\to\infty}} x_n = \underline{\lim_{n\to\infty}} x_n$。

习题 1.2

1. 试证明确界的惟一性。

2. 举例说明区间套引理中闭区间套$[a_n,b_n]$不能改为开区间套(a_n,b_n)。

3. 举例说明 Heine - Borel 有覆盖定理中不能把有界闭区间改为开区间(a,b)或半开半闭区间$(a,b]$及$[a,b)$,也不能改为无界闭区间$(-\infty,b]$及$[a,\infty)$。

4. 设 $E\subset\mathbf{R}$,求证 $\sup E=-\inf(-E)$,这里

$\inf E=-\sup(-E)$,其中 $-E=\{-x:x\in E\}$。

5. 设有两个函数列 $f_n(x)$ 与 $g_n(x)$,求证对任一固定的 x,有 $\inf\limits_n f_n(x)+\inf\limits_n g_n(x)\leqslant\inf\limits_n(f_n(x)+g_n(x))\leqslant\inf\limits_n f_n(x)+\sup\limits_n g_n(x)\leqslant\sup\limits_n(f_n(x)+g_n(x))\leqslant\sup\limits_n f_n(x)+\sup\limits_n g_n(x)$。

6. 设函数列 $f_n(x)$ 在定义域 E 上为有界函数且收敛于函数 $f(x)$。对每个 $n=1,2,\cdots$,令 $g_n(x)=\inf\limits_{k\geqslant n}f_k(x)=\inf\{f_n(x),f_{n+1}(x),\cdots\}$,求证 $g_n(x)$ 为单调增加函数列且有

$$f(x)=\lim_{n\to\infty}g_n(x)。$$

7. 对于点列 $x_n\in\mathbf{R}$,求证 $\lim\limits_{n\to\infty}x_n=a\Leftrightarrow\overline{\lim\limits_{n\to\infty}}x_n=\underline{\lim\limits_{n\to\infty}}x_n=a$。

8. 对于点列 $x_n\in\mathbf{R}$,若 $|x_{n+1}-x_n|\leqslant\frac{1}{2^n}$,$n=1,2,\cdots$,求证 x_n 是 Cauchy 列。又若 $\lim\limits_{n\to\infty}|x_{n+1}-x_n|=0$,问 x_n 是否一定为 Cauchy 列?

9. 对点列 $x_n\in\mathbf{R}$,设 $\overline{\lim\limits_{n\to\infty}}x_n=b\in\mathbf{R}$,求证对任 $\varepsilon>0$,有 x_n 的无穷多项含在邻域$(b-\varepsilon,b+\varepsilon)$内,而在$(b+\varepsilon,\infty)$内却仅有 x_n 的有限项。

10. 对点列 $x_n \in \mathbf{R}$，设 $\underline{\lim}_{n\to\infty} x_n = a \in \mathbf{R}$，请给出类似于 9 题的结论。

1.3 函数项级数的基本问题

由于实数集 $\mathbf{R}$ 的结构，对于其上的一个函数来说，我们不但可以研究它的代数运算，而且可以讨论其分析运算，比如极限、微分及积分运算等。函数项级数是在许多实际问题及理论中经常要碰到的一类函数，它的定义依赖于实数集的代数结构及度量结构。本节概要地研究这一类函数的某些分析性质。

1.3.1 函数项级数的基本概念

设 $u_n(x)$ 是定义在某点集 $E \subset \mathbf{R}$ 上的函数($n = 1,2,3,\cdots$)，则称

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \cdots$$

为函数项级数。并且称函数列

$$f_n(x) = \sum_{k=1}^{n} u_k(x)$$

为级数的前 n 项和(注意这里有代数运算)。

若某点 $x_0 \in E$，序列 $f_n(x_0) = \sum_{k=1}^{n} u_k(x_0)$ 收敛(注意此时为极限运算)，则称函数项级数在 x_0 收敛，$x_0 \in E$ 叫做级数的收敛点，否则称为级数的发散点。级数 $\sum_{n=1}^{\infty} u_n(x)$ 的收敛点全体称为级数的收敛域。特别当收敛域为 E 时，则 $\sum_{n=1}^{\infty} u_n(x)$ 就确定了 E 上一个新函数

$$f(x)=\sum_{n=1}^{\infty}u_n(x) \text{ 或 } f(x)=\lim_{n\to\infty}\sum_{k=1}^{n}u_k(x)。$$

一般情况下，函数的连续性是重要的。

【定义 1】 设有函数 $f:E\to\mathbf{R}$，其中 $E\subset\mathbf{R}$，$x_0\in E$，若对任一给定 $\varepsilon>0$，存在 $\delta>0$，使当 $x\in E\cap(x_0-\delta,x_0+\delta)$ 时，有

$$f(x)\in(f(x_0)-\varepsilon,f(x_0)+\varepsilon),$$

则称函数 f 在点 x_0 连续。

注 当 f 在每个点 $x\in E$ 都连续时（此时称 f 为 E 上连续函数），应特别注意，δ 的大小不但与 ε 的大小有关，而且同 $x\in E$ 也相关，即连续性是局部性概念。E 中含有无穷个点 x，对同一个 $\varepsilon>0$，相应就有无穷个 $\delta_x>0$，问题是不一定存在 $\delta=\inf\limits_{x\in E}\delta_x>0$（建议读者举出这样的例子）。而对于存在有对全体 E 中点都适用的 $\delta>0$ 的函数，我们给出以下概念。

【定义 2】 对于函数 $f:E\to\mathbf{R}$，$\forall\varepsilon>0$，$\exists\delta>0$，当 $x_1,x_2\in E$，只要 $|x_1-x_2|<\delta$，有 $|f(x_1)-f(x_2)|<\varepsilon$，则称 f 在 E 上一致连续。

注 显然，函数在 E 上一致连续，则必在 E 上连续，它是一个全局性概念。

一般情况是 E 上的连续函数未必一致连续，但对特别的定义域，有以下结果。

【定理 1】(Cantor) 若函数 f 在有界闭区间 $[a,b]$ 上连续，则 f 在 $[a,b]$ 上也一致连续。

证 用反证法，假设 f 在 $[a,b]$ 上不一致连续，于是必存在某 $\varepsilon_0>0$，使对 $\forall\delta>0$，$\exists x,y\in[a,b]$，$|x-y|<\delta$，而

$$|f(x)-f(y)|\geq\varepsilon_0。$$

取 $\delta = \frac{1}{n} > 0$,则存在 $x_n, y_n \in [a, b]$, $|x_n - y_n| < \frac{1}{n}$,而

$$|f(x_n) - f(y_n)| \geq \varepsilon_0。$$

由 Bolzano - Weierstrass 定理(2.2 节定理 4),有界数列 x_n 存在子列 $x_{n_k} \to x_0 \in [a, b]$。

因

$$\left| x_{n_k} - y_{n_k} \right| < \frac{1}{n_k} \quad (k = 1, 2, \cdots),$$

推知 $y_{n_k} \to x_0$,注意 f 在 x_0 处连续,则有

$$0 = |f(x_0) - f(x_0)| = \lim_{k \to \infty} \left| f(x_{n_k}) - f(y_{n_k}) \right| \geq \varepsilon_0,$$

矛盾,故 f 在 $[a, b]$ 上不一致连续的假设不真,f 在 $[a, b]$ 上一致连续性得证。证毕。

现在,关于函数项级数 $f(x) = \sum_{n=1}^{\infty} u_n(x) = \lim_{n \to \infty} \sum_{k=1}^{n} u_k(x)$ 提出问题:$f(x)$ 可否继承 $u_n(x)$ 或 $f_n(x) = \sum_{k=1}^{n} u_k(x)$ 的连续性、可积性以及可微性等性质,即有如下基本问题:

(1) 若 $u_n(x) \in C[a, b]$, $n = 1, 2\cdots$,其中 $C[a, b]$ 表示全体 $[a, b]$ 上连续函数组成的集合,问在什么条件下有 $f(x) \in C[a, b]$,即对任 $x_0 \in [a, b]$,有

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{x \to x_0} f_n(x) = f(x_0)$$

或

$$\lim_{x \to x_0} \sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \lim_{x \to x_0} u_n(x) = \sum_{n=1}^{\infty} u_n(x_0)。$$

此问题为两个极限交换顺序问题或逐项取极限问题。

(2) 设 $u_n(x)$ 在 $[a, b]$ 上 Riemann 可积,问在何种条件下有

$$\lim_{n\to\infty}\int_a^b f_n(x)\mathrm{d}x = \int_a^b \lim_{n\to\infty}f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$$

或

$$\int_a^b f(x)\mathrm{d}x = \int_a^b \sum_{n=1}^{\infty} u_n(x)\mathrm{d}x = \sum_{n=1}^{\infty}\int_a^b u_n(x)\mathrm{d}x。$$

此问题是极限与积分两种运算次序的可交换问题，或可否逐项积分的问题。

(3) 设 $u_n(x)$ 在 $[a,b]$ 上可微，问在什么条件下有

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}(\lim_{n\to\infty}f_n(x)) = \lim_{n\to\infty}\frac{\mathrm{d}f_n(x)}{\mathrm{d}x}$$

或

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sum_{n=1}^{\infty}u_n(x)) = \sum_{n=1}^{\infty}\frac{\mathrm{d}u_n(x)}{\mathrm{d}x}。$$

此问题为极限与微分两种运算的交换问题，或可否逐项求微分问题。

1.3.2 函数项级数的一致收敛性

函数列 $f_n(x) = \sum_{k=1}^{n} u_k(x) \to f(x) = \sum_{n=1}^{\infty} u_n(x), x \in E$ 就是对任给 $\epsilon > 0$，存在自然数 N，使当 $n > N$ 时，有 $|f_n(x) - f(x)| < \epsilon$(或 $\left|\sum_{k=n+1}^{\infty} u_k(x)\right| < \epsilon$)。一般情况下，上述中 N 不仅依赖于 ϵ，而且也与 x 有关，即对同一个 $\epsilon > 0$，对于 E 中不同 x，相应 N 通常是不同的。而对于存在对全体 E 中点都适用的 N(对同一个 $\epsilon > 0$) 的函数列 $f_n(x)$(或 $\sum_{n=1}^{\infty} u_n(x)$) 我们给出如下概念。

【定义 3】 对于函数序列 $f_n(x), x \in E$，若对任给 $\epsilon > 0$，存在只与 ϵ 相关的自然数 N，使当 $n > N$ 时，对一切 $x \in E$，都有

$$|f_n(x) - f(x)| < \epsilon,$$

则称 $f_n(x)$ 在 E 上一致收敛于 $f(x)$，记成 $f_n(x) \Rightarrow f(x)$（在 E 上）。

若 $f_n(x) = \sum_{k=1}^{n} u_k(x)$，则前面的不等式可写为

$$\left|\sum_{k=n+1}^{\infty} u_k(x)\right| < \varepsilon。$$

直接利用定义去检验函数项级数（或函数列）的一致收敛性，一般是困难的，以下介绍一个最简单且最常用的检验法，即 Weierstrass 判别法。

【定理 2】 设有 $f(x) = \sum_{n=1}^{\infty} u_n(x)$，$x \in E$，若存在 N，当 $n \geqslant N$ 时，有 $|u_n(x)| \leqslant a_n$，$\forall x \in E$，其中 $a_n > 0$，并且级数 $\sum_{n=1}^{\infty} a_n$ 收敛，则 $f(x)$ 在 E 上一致收敛。

证 $\forall \varepsilon > 0$，因 $\sum_{n=1}^{\infty} a_n$ 收敛，则存在 $N^* \geqslant N$，有

$$\sum_{n=N^*}^{\infty} a_n < \varepsilon，$$

于是由设，对任 $x \in E$ 有

$$\left|\sum_{n=N^*}^{\infty} u_n(x)\right| = \lim_{n\to\infty}\left|\sum_{k=N^*}^{n} u_k(x)\right| \leqslant \lim_{n\to\infty}\sum_{k=N^*}^{n} |u_k(x)| \leqslant \sum_{k=N^*}^{\infty} a_k < \varepsilon。$$

这就证明了 $\sum_{n=1}^{\infty} u_n(x)$ 在 E 上的一致收敛性。证毕。

以下是一致收敛的 Cauchy 收敛准则。

【定理 3】 函数序列 $f_n(x)$ 在 E 上一致收敛的充要条件是对任给 $\varepsilon > 0$，存在自然数 N，只要 $n, m > N$，对任 $x \in E$，有

$$|f_n(x) - f_m(x)| < \varepsilon。$$

对于级数形式，上式为

$$\left|\sum_{k=n+1}^{m} u_k(x)\right| < \varepsilon。$$

定理的证明类似于数列的 Cauchy 收敛准则。留给读者作练习。

1.3.3 一致收敛函数项级数的性质

【定理 4】 设函数列 $f_n(x)$ 在 E 上一致收敛于 $f(x)$，且每个 $f_n(x)$ 在 E 上连续，则 $f(x)$ 在 E 上连续。

证 我们要证 $f(x)$ 在 E 上的每个点连续。任意取定 $x_0 \in E$，对任给 $\varepsilon > 0$，因 $f_n \Rightarrow f$（在 E 上），则存在 N，有 $|f_N(x) - f(x)| < \frac{\varepsilon}{3}$，$\forall x \in E$。而 $f_N(x)$ 在 x_0 连续，故有 $\delta > 0$，当 $x \in E \cap (x_0 - \delta, x_0 + \delta)$ 时，有 $|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3}$。

于是当 $x \in E \cap (x_0 - \delta, x_0 + \delta)$ 时，有

$$|f(x) - f(x_0)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \varepsilon。$$

这就表明 $f(x)$ 在 x_0 连续，由 $x_0 \in E$ 有

$$\lim_{x \to x_0} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{x \to x_0} f_n(x)。$$

【定理 5】 设函数列 $f_n(x)$ 在 $[a, b]$ 上一致收敛于 $f(x)$，且每个 $f_n(x) \in C[a, b]$，则

$$\lim_{n \to \infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b \lim_{n \to \infty} f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x,$$

即极限与积分两运算可交换次序。

证 由本定理条件及定理 4 可知，$f_n(x)$ 及 $f(x)$ 都在 $[a, b]$ 上可积。现在证明 $\lim\limits_{n \to \infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$。

对任给 $\varepsilon > 0$，因 $f_n \Rightarrow f$（在 $[a, b]$ 上），存在 N，当 $n > N$ 时，

$$|f_n(x) - f(x)| < \varepsilon, \forall x \in [a,b]$$

从而只要 $n > N$ 时，有

$$\left|\int_a^b f_n(x)\mathrm{d}x - \int_a^b f(x)\mathrm{d}x\right| \leqslant \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x < (b-a)\varepsilon。$$

定理证毕。

【定理 6】 设函数列 $f_n(x)$ 在 $[a,b]$ 上收敛于 $f(x)$，每个 $f_n(x)$ 在 $[a,b]$ 上有连续导函数 $f_n'(x)$，并且 $f_n'(x) \Rightarrow g(x)$（在 $[a,b]$ 上），则 $f'(x) = g(x)$，即

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\lim_{n\to\infty} f_n(x)\right) = \lim_{n\to\infty} \frac{\mathrm{d}f_n(x)}{\mathrm{d}x}。$$

这表明极限与微分两运算在给定条件下可交换次序。

证明留给读者完成。

注　(1) 一般说来，定理 4、5、6 的逆命题不真。

(2) 关于定理 4、5、6，建议读者写出关于函数项级数相应的定理。

幂级数 $\sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ 作为一类特殊的函数项级数，它在很多实际和理论问题中都有重要应用。幂级数的前 $n+1$ 项和是一个多项式函数 $f_n(x) = \sum\limits_{k=0}^{n} a_k(x-x_0)^k$，根据函数项级数理论我们可以得到幂级数许多重要性质。例如 $f(x) = \sum\limits_{n=0}^{\infty} a_k(x-x_0)^n$ 在收敛区间内有各阶导函数，并且可逐项求导任意次。因此一般说来，$[a,b]$ 上的连续函数不一定可以表示为一个幂级数，但是却可以找到多项式 $P(x)$ 来逼近 f，使其误差一致地充分小，这就是著名的 Weierstrass 逼近定理。

【定理 7】 设 $f \in C[a,b]$，对任给 $\varepsilon > 0$，必存在多项式 $P(x)$，

对任一 $x \in [a,b]$，有

$$|f(x) - P(x)| < \varepsilon。$$

定理的证明较复杂，读者可在内容较详细的数学分析教科书中找到。

习题 1.3

1. 证明 $f(x) = \sin\dfrac{1}{x}$ 在 $(0,1)$ 内不一致连续。

2. 证明 $f(x) = x^{\frac{1}{3}}$ 在 $[0,\infty)$ 上一致连续。

3. 证明 $f(x) = x^2$ 在 R 上不一致连续。

4. 设有函数列 $f_n(x)$ 及函数 $f(x)$ 都在 E 上有定义，记

$$\| f_n - f \| = \sup_{x\in E} | f_n(x) - f(x) |,$$

证明 $f_n \rightrightarrows f$（在 E 上）$\Leftrightarrow \lim\limits_{n\to\infty} \| f_n - f \| = 0$。

5. 对于函数列 $f_n(x) = 2n^2 x e^{-n^2x^2}$，

(1) 证明它在 $[0,1]$ 上不是一致收敛的；

(2) 验证定理 4、5 的结论是否成立。对你得到的结果如何理解？

6. 证明定理 6。

7. 证明 $f(x) = \sum\limits_{n=1}^{\infty} \dfrac{1}{n^x}$ 在 $(1,\infty)$ 内连续，并且它有各阶连续导函数。

8. 设 $\sum\limits_{n=1}^{\infty} | u_n(x) |$ 在 E 上一致收敛，求证 $\sum\limits_{n=1}^{\infty} u_n(x)$ 在 E 上也一致收敛。并举例说明其逆不真。

9. 证明 $f(x) = \sum\limits_{n=1}^{\infty} n e^{-nx}$ 在 $(0,\infty)$ 内连续。

1.4　Lebesgue 积分

由 Newton 等科学家提出并最终为 Riemann 给出的严密定积分概念是分析数学中最重要的基本运算之一。以后人们又给出了广义积分的概念，但可积分的函数类仍不够广泛，并且这种积分的运算性质不很灵活。为此，19 世纪末以来，相继提出了许多积分概念，但直到 20 世纪初，法国人 Lebesgue 所提出的积分概念才部分的解决了积分学中的许多问题，而成为现代数学中最常用的基本运算。

关于 Lebesgue 积分理论（以后简称为 L 积分，而先前的 Riemann 积分则简记为 R 积分），原来要涉及集合的测度理论，且基本方法比较艰深。限于篇幅，我们采用较为直接的途径概要介绍 L 积分的基本理论。

1.4.1　零测度集

以下要介绍的零测度集的概念属于 Lebesgue，它在许多领域都有重要应用，并且可以避免使用一般测度的理论。

【定义 1】　对于点集 $E \subset \mathbf{R}$，若对任给 $\varepsilon > 0$，存在有限个或可列个（统称为至多可列个）开区间 (α_k, β_k)，使得

$$E \subset \bigcup_k (\alpha_k, \beta_k)，且 \sum_k (\beta_k - \alpha_k) < \varepsilon,$$

则称 E 为零测度集。

由于实数集 $\mathbf{R}$ 与直线上的点集的对等关系，则利用线段的长度概念定义对应的实数集的长度。亦即每个闭区间 $[a, b]$ 所对应的点集有长度 $b - a$。当闭区间蜕化为一点时，其长度应为 0，于是由长度的可加性（若一条线段等于另两条线段的和，则这条线段的长度等于

另两条线段长度的和),开区间或半开半闭的区间与对应的闭区间有相同的长度。若 $E \subset \mathbf{R}$ 为有限点集,则显然 E 的长度为 0。对于一般的无限点集 E,如何定义它的长度,则属于测度的一般理论,而零测度集是测度理论中最特殊情形。

由零测度集定义容易推知,零测度集的任子集都是零测度集。并且还知道,任何有限个或可列多个零测度集的并集也是零测度集,即设 E_n 为零测度集($n = 1,2,\cdots$),则并 $E = \bigcup_{n=1}^{\infty} E_n$ 是零测度集。事实上,对任 $\epsilon > 0$,由设对每个 E_n,存在至多可列个开区间,它们长度总和小于$\frac{\varepsilon}{2^n}$($n = 1,2,\cdots$),而所有这些开区间并起来仍是可列个,并且 E 属于它们的并集。再注意这些区间的总长度小于 $\sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \epsilon$。这就说明 E 是零测度集。由此可知,任何可列集都是零测度集。

现在介绍一个术语,如果某个命题或某个性质在集合 E 上除了某个零测度子集外处处成立,则常说为在 E 上几乎处处成立。例如我们说函数 f 与 g 在 E 上几乎处处相等,即指 $f(x)$ 与 $g(x)$ 不相等的点集是零测度集,常记成 $f(x) = g(x) a \cdot e$(almost everywhere)。不难验证几乎处处相等是一个等价关系。又例如函数列 $f_n(x)$ 在 E 上几乎处处收敛于函数 $f(x)$,即指集 $\{x \in E: \lim_{n\to\infty} f_n(x) \neq f(x)\}$ 是零测度集。

以下的结果也由 Lebesgue 给出,是分析数学中一个重要定理。我们仅叙述内容,证明较复杂,这里略去。

【定理 1】 设 $f(x)$ 为 $[a,b]$ 上单调函数,则它在此区间上几乎处处有有限的导数。

1.4.2 Riemann 积分

我们把正无穷 ∞(或记为 $+\infty$)及负无穷 $-\infty$ 视作两个"值"扩

充入实数域 $\mathbf{R}$ 中，记为 $\overline{\mathbf{R}} = [-\infty,\infty] = \mathbf{R}\cup\{-\infty,\infty\}$，称它为广义数轴。任一实数 $x\in\mathbf{R}$，有 $-\infty < x < \infty$，并对广义数 $-\infty,\infty$ 与实数 x 的运算有如下规定。

$$x+(\pm\infty) = (\pm\infty)+x = (\pm\infty)-x = x-(\mp\infty) = \pm\infty;$$

$$(\pm\infty)+(\pm\infty) = \pm\infty;$$

$$\frac{x}{\pm\infty} = 0\cdot(\pm\infty) = 0;$$

$x\cdot(\pm\infty) = \pm\infty$，当 $x>0$ 时；$x\cdot(\pm\infty) = \mp\infty$，当 $x<0$ 时。

又规定同号无穷大相乘得正无穷，而异号无穷相乘得负无穷。但 $(\pm\infty)-(\pm\infty),(\pm\infty)+(\mp\infty),\frac{\pm\infty}{\pm\infty},\frac{x}{0},\frac{\pm\infty}{0}$ 无意义。

一般情况下，称允许取无穷大的函数 $f:E\to\overline{\mathbf{R}}$ 为广义实函数，今后不加说明，凡说函数均指广义实函数，而对函数 $f:E\to\mathbf{R}$ 叫做有限函数。若称某函数 f 在 E 上几乎处处为有限函数，则指 f 可取无穷的是一个零测度集。注意有限函数未必是有界函数。

【定义 2】 设非负函数 $f:E\to\overline{\mathbf{R}}$，则 $E\times\overline{\mathbf{R}}$ 中集合

$$\{(x,y):x\in E,0\leqslant y\leqslant f(x)\}$$

称为 f 的下方图形。当 $E=[a,b]$，则有限函数 $f=[a,b]\to\mathbf{R}$ 的下方图形就是曲边梯形。下方图形记为 $G(E,f)$。

【定义 3】 设 $E\subset\mathbf{R}$，以下的函数称为 E 的特征函数。

$$\chi_E(x) = \begin{cases} 1, & \text{当 } x\in E; \\ 0, & \text{当 } x\in\mathbf{R}\setminus E。\end{cases}$$

对于有穷区间 (a,b)，作划分 $\pi: a=x_0<x_1<x_2<\cdots<x_n=b$，记 $\Delta x_k = (x_k,x_{k+1}),k=0,1,2,\cdots,n-1$，记 $\|\pi\| = \max\{|\Delta x_k| = x_{k+1}-x_k: 0\leqslant k\leqslant n-1\}$。

【定义 4】 常称以下函数 $\phi(x)$ 为阶梯函数：

$$\phi(x)=\sum_{k=0}^{n-1}c_k\chi_{\Delta x_k}(x)+\sum_{k=0}^{n-1}d_k\chi_{\{x_k\}}(x),x\in E=[a,b]$$

其中，$c_k,d_k\in\mathbf{R},0\leqslant k\leqslant n-1$。若 $a=-\infty$ 时，须取 $c_0=0$；若 $b=\infty$，则应取 $c_{n-1}=0$。

由几何的直观性，我们定义非负阶梯函数 $\phi(x)$ 在 (a,b) 上的积分值为它的下方图形的面积，即

$$\int_a^b\phi(x)\mathrm{d}(x)=\sum_{k=0}^{n-1}c_k\mid\Delta x_k\mid。$$

注 在讨论阶梯函数的积分问题时，由于它在每个小区间 Δx_k 的端点处的值的变化不影响其积分值，因此常可忽略阶梯函数在 Δx_k 端点处的定义。

依照这种方式得到如下一般定义。

【定义 5】 对于阶梯函数 $\phi:(a,b)\to\mathbf{R}$，它在 (a,b) 上的积分值定义为和数 $\sum\limits_{k=0}^{n-1}c_k\mid\Delta x_k\mid$，记号同非负情形。

借助如下无限逼近的方式可以得到 Riemann 积分概念。

设 f 在有穷区间 (a,b) 为有界函数。对于划分 π，令 $m_k=\inf\{f(x):x\in\Delta x_k\},M_k=\sup\{f(x):x\in\Delta x_k\},k=0,1,2,\cdots,n-1$，则相应确定两个阶梯函数分别为

$$\phi_\pi^{(1)}(x)=\sum_{k=0}^{n-1}m_k\chi_{\Delta x_k}(x)\text{ 及 }\phi_\pi^{(2)}(x)=\sum_{k=0}^{n-1}M_k\chi_{\Delta x_k}(x)。$$

将它们的积分值分别记作 $s_\pi=\sum\limits_{k=0}^{n-1}m_k\mid\Delta x_k\mid$ 及 $S_\pi=\sum\limits_{k=0}^{n-1}M_k\mid\Delta x_k\mid$。它们分别称为关于划分 π 的 Darboux 小和与 Darboux 大和。任取 $\xi_k\in\Delta x_k$，则有阶梯函数 $\phi_\pi(x)=\sum\limits_{k=0}^{n-1}f(\xi_k)\chi_{\Delta x_k}(x)$，它的积分叫做 Riemann 和，记作 $\sigma_\pi=\sum\limits_{k=0}^{n-1}f(\xi_k)\mid\Delta x_k\mid$。它不仅依赖于划分 π，且与 ξ_k

的选取有关。由定义立知

$$s_\pi \leqslant \sigma_\pi \leqslant S_\pi。$$

以下列出它们的一些基本性质，有关证明可在数学分析的教科书中查到。

【定理 2】 (1) 对任划分 π_1，再添入一些新分点后得到加细的划分 π_2，则 $s_{\pi_1} \leqslant s_{\pi_2}$；$S_{\pi_2} \leqslant S_{\pi_1}$；

(2) 若 π_1 与 π_2 是 (a,b) 的任意两个划分，则 $s_{\pi_1} \leqslant S_{\pi_2}$；

(3) 对 (a,b) 任划分 π，有 $m(b-a) \leqslant s_\pi \leqslant S_\pi \leqslant M(b-a)$。其中 $m = \inf\{f(x): x \in (a,b)\}$；$M = \sup\{f(x): x \in (a,b)\}$。

由定理推知小和有上界，大和有下界，则令

$$\underline{\int_a^b} f(x)\mathrm{d}x = \sup\{s_\pi : \pi \text{ 是} (a,b) \text{ 的划分}\};$$

$$\overline{\int_a^b} f(x)\mathrm{d}x = \inf\{S_\pi : \pi \text{ 是} (a,b) \text{ 的划分}\},$$

分别称它们为 f 在 (a,b) 的下 R 积分和上 R 积分。

一般情形易知有 $\underline{\int_a^b} (x)\mathrm{d}x \leqslant \overline{\int_a^b} f(x)\mathrm{d}x$。当特别成立 $\underline{\int_a^b} f(x)\mathrm{d}x = \overline{\int_a^b} f(x)\mathrm{d}x$ 时，则称 f 在 (a,b) 上 Riemann 可积，它们的公共值记为 $\int_a^b f(x)\mathrm{d}x$。

关于 R 积分可证明如下的判别准则。

【定理 3】 函数 f 在 (a,b) 上可积的充要条件是 $\lim\limits_{\|\pi\| \to 0}(S_\pi - s_\pi) = 0$，即任 $\varepsilon > 0$，$\exists\, \delta > 0$，对 (a,b) 任划分 π，只要 $\|\pi\| < \delta$，有 $S_\pi - s_\pi = \sum\limits_{k=0}^{n-1}(M_k - m_k)\Delta x_k < \varepsilon$。

常记 $\omega_k = M_k - m_k$，叫做 f 在 Δx_k 上的振幅。这个定理一定程度上给出了 R 可积函数的特征，而借助于零测度集概念得到的下述结

果给出了 R 可积函数的基本特征。

【定理 4】 函数 f 在 (a,b) 上 R 可积的充要条件是 f 在 (a,b) 内的间断点集为零测度集。

定理的证明要查阅实变函数的教科书。由定理看到 R 积分几乎是为连续函数设计的。我们来看函数

$$f(x)=\begin{cases}0, & \text{当 } x\in[0,1]\cap\mathbf{Q};\\ 1, & \text{当 } x \text{ 为}(0,1)\text{ 中无理数时。}\end{cases}$$

因为它在 $[0,1]$ 内处处不连续，故 f 在 $[0,1]$ 上不是 R 可积的，即在 Riemann 意义下，f 在 $[0,1]$ 上的下方图形无面积。但由直觉知，若函数 f 在零测度集 E 上为非负函数，它的下方图形的面积为 0。令 $E=[0,1]\cap\mathbf{Q}$，则 E 是零测度集，记 F 为 $[0,1]$ 中的无理数集，则 $E\cup F=[0,1]$，于是由面积的可加性公理，不难算出 f 在 $[0,1]$ 上的下方图形面积应是 1。这足以说明 R 积分适用范围的狭窄性。

1.4.3 Lebesgue 积分的概念

我们现在介绍以另一种方式把积分概念推广到更为广泛的函数上去。先给出两个基本引理。在不加说明情况下，我们所说的函数都在有穷或无穷区间 (a,b) 内有定义。

【引理 1】 设 $\phi_n(x)$ 为单调减小的阶梯函数列，即对每个 $x\in(a,b)$，有 $\phi_n(x)\geqslant\phi_{n+1}(x)$，$n=1,2,\cdots$，并且 $\phi_n(x)$ 在 (a,b) 内几乎处处收敛于零，则其积分序列 $\int_a^b\phi_n(x)\mathrm{d}x$ 趋于零。

证 由阶梯函数定义知每个 $\phi_n(x)$ 仅有有限个间断点。设 E_0 表示 $\phi_n(x)$ 不收敛于零的点及其间断点的全体组成的集合，则由引理条件推知 E_0 为零测度集。关于 $\phi_1(x)$，存在有限区间 $[a_1,b_1]\subset(a,b)$，有 $\phi_1(x)=0$，当 $x\in(a,b)-[a_1,b_1]$ 时。因 $\phi_n(x)$ 在 (a,b) 内

单调减趋于零，则每个 $\phi_n(x)$ 在 $(a,b)-[a_1,b_1]$ 内几乎处处等于零，于是有 $\int_a^b \phi_n(x)\mathrm{d}x = \int_{a_1}^{b_1} \phi_n(x)\mathrm{d}x$。

对任给 $\varepsilon > 0$，因 E_0 是零测度集，则存在至多可列开区间列 $\{(\alpha_k,\beta_k)\}$ 覆盖 E_0，且这些区间总长度小于 ε。现在任取 $x_0 \in [a_1, b_1] \setminus E_0$，则 $\phi_n(x_0)$ 单调减趋于零，于是存在 N^*。当 $n \geqslant N^*$ 时，有 $0 \leqslant \phi_n(x_0) < \varepsilon$。注意 x_0 是 ϕ_{N^*} 的连续点，由阶梯函数定义，$\phi_{N^*}(x)$ 在含 x_0 的某区间 $I_{N^*}(x_0)$ 内取定值 $\phi_{N^*}(x_0)$，因而在此区间内成立 $\phi_{N^*}(x) < \varepsilon$。让 x_0 在 $[a,b]-E_0$ 上变动，则得到开区族 $\{I_{N^*}(x_0)\}$，其中 N^* 依赖于 x_0，当 $x \in I_{N^*}(x_0)$ 时，有 $\phi_{N^*}(x) < \varepsilon$。显然，区间列 $\{(\alpha_k,\beta_k)\}$ 与区间族 $\{I_{N^*}(x_0)\}$ 合起来覆盖了 $[a_1,b_1]$，于是由 Heine－Borel 有限覆盖定理，可从开覆盖中选出有限个将 $[a_1,b_1]$ 覆盖住。记从 $\{(\alpha_k,\beta_k)\}$ 中选出的为 $(\alpha_1,\beta_1),(\alpha_2,\beta_2),\cdots,(\alpha_m,\beta_m)$，而从 $\{I_{N^*}(x_0)\}$ 中选出的记为 $I_{N_1}(x_1)=(t_1,s_1), I_{N_2}(x_2)=(t_2,s_2), \cdots, I_{N_l}(x_l)=(t_l,s_l)$。记 $M=\sup\{\phi_1(x): x\in(a,b)\}$，$N=\max\{N_1, N_2,\cdots,N_l\}$，只要 $n > N$，$\left|\int_a^b \phi_n(x)\mathrm{d}x\right| = \left|\int_{a_1}^{b_1}\phi_n(x)\mathrm{d}x\right| \leqslant \left|\sum_{k=1}^{m}\int_{\alpha_k}^{\beta_k}\phi_n(x)\mathrm{d}x\right| + \left|\sum_{k=1}^{l}\int_{t_k}^{s_k}\phi_n(x)\mathrm{d}x\right| \leqslant \varepsilon(M + b_1 - a_1)$，即 $\lim\limits_{n\to\infty}\int_a^b \phi_n(x)\mathrm{d}x = 0$，引理证毕。

【引理 2】　设 $\phi_n(x)$ 为单调增加的阶梯数列，且存在 $M > 0$，有 $\left|\int_a^b \phi_n(x)\mathrm{d}x\right| \leqslant M$，$n=1,2,\cdots$，则 $\phi_n(x)$ 在 (a,b) 内几乎处处存在有限的极限。

证　不妨设每个 $\phi_n(x)$ 都非负（否则只须对序列 $\phi_n-\phi_1$ 进行讨

论)。设 $E_0=\{x\in(a,b):\lim\limits_{n\to\infty}\phi_n(x)=\infty\}$,只须证明 E_0 为零测度集。

对任给 $\epsilon>0$,记 $E_\epsilon^{(n)}=\{x\in(a,b):\phi_n(x)>\frac{M}{\epsilon}\}$, $n=1,2,\cdots$。因 $\phi_n(x)$ 单调增,则 $E_\epsilon^{(n)}\subset E_\epsilon^{(n+1)}$,记 $E_\epsilon=\bigcup\limits_{n=1}^{\infty}E_\epsilon^{(n)}$,则明显有 $E_0\subset E_\epsilon$。对每个 n, $x\in E_\epsilon^{(n)}$,则 $\phi_n(x)>\frac{M}{\epsilon}$,因 $\phi_n(x)$ 是阶梯函数,则使上式成立的点集应是有限个不相交有界区间的并,故 E_ϵ 为至多可列个有界区间的并,它覆盖了 E_0。对 $E_\epsilon^{(n)}$ 我们有

$$M\geqslant\int_a^b\phi_n(x)\mathrm{d}x\geqslant\frac{M}{\epsilon}\left|E_\epsilon^{(n)}\right|,$$

故得 $\left|E_\epsilon^{(n)}\right|\leqslant\epsilon$。这里 $\left|E_\epsilon^{(n)}\right|$ 表示区间集的长度, $n=1,2,\cdots$。注意 $E_\epsilon^{(n)}\subset E_\epsilon^{(n+1)}$,则 E_ϵ 的总长度不大于 ϵ,这已说明 E_0 是零测度集。证毕。

我们用 C_0 表示全体阶梯函数组成的集合。设 $\phi_n(x)$ 为满足引理 2 条件的阶梯函数列,定义 $f(x)=\lim\limits_{n\to\infty}\phi_n(x)$,它在 (a,b) 内几乎处处有定义,且为有限值。全体这样的函数组成的集合用 C_1 表示。

对于 $f\in C_1$, $f(x)=\lim\limits_{n\to\infty}\int_a^b\phi_n(x)$,因 $\phi_n(x)$ 单调增加,则积分序列 $\int_a^b\phi_n(x)\mathrm{d}x$ 是单调上升且有界数列,故存在

$$\lim_{n\to\infty}\int_a^b\phi_n(x)\mathrm{d}x\in\mathbf{R}。$$

此时我们定义 $f(x)$ 在 (a,b) 上积分(Lebesgue 意义下)等于这个极限值,记为

$$(\mathrm{L})\int_a^b f(x)\mathrm{d}x=\lim_{n\to\infty}\int_a^b\phi_n(x)\mathrm{d}x。$$

注　(1) 须说明以上定义的合理性，即 $(L)\int_a^b f(x)\mathrm{d}x$ 不依赖于 $\phi_n(x)$ 的选择。事实上，若还有满足引理 2 条件的阶梯函数 $\Psi_n(x)$ 也几乎处处收敛于 $f(x)$，则亦存在有限极限 $\lim\limits_{n\to\infty}\int_a^b \Psi_n(x)\mathrm{d}x$。任取定 $\phi_m(x)$，则 $\phi_m(x)-\Psi_n(x)$ 是单调减小的阶梯函数列，其中 $n=1,2,\cdots$，则 (a,b) 分成两类区间集，第一类区间中的点 x，使 $\phi_m(x)-\Psi_n(x)>0, n=1,2,\cdots$；而第二类区间中点 x，存在相应的 N，使当 $n\geqslant N$ 时，有 $\phi_m(x)-\Psi_n(x)\leqslant 0$。记第一类区间全体为 I_1，第二类区间全体为 I_2，则

$$\lim_{n\to\infty}\int_a^b(\Psi_m(x)-\Psi_n(x))\mathrm{d}x=\lim_{n\to\infty}\left(\int_{I_1}(\phi_m(x)-\Psi_n(x))\mathrm{d}x+\right.$$
$$\left.\left(\int_{I_2}(\phi_m(x)-\Psi_n(x))\mathrm{d}x\right)\right)。$$

由设在 I_1 上几乎处处有

$$0\leqslant\lim_{n\to\infty}(\phi_m(x)-\Psi_n(x))=$$
$$\phi_m(x)-\lim_{n\to\infty}\Psi_n(x)=\phi_m(x)-f(x)\leqslant 0,$$

即当 $x\in I_1$ 时，有 $\phi_m(x)-\Psi_n(x)$ 单调减小的几乎处处收敛于 0，则由引理 1，$\lim\limits_{n\to\infty}\int_{I_1}(\phi_m(x)-\Psi_n(x))\mathrm{d}x=0$；因任 $x\in I_2$，有 $\phi_m(x)-\Psi_n(x)\leqslant 0$，对充分大的 n 成立，故

$$\lim_{n\to\infty}\int_{I_2}(\phi_m(x)-\Psi_n(x))\mathrm{d}x\leqslant 0。$$

综上得

$$\int_a^b\phi_m(x)\mathrm{d}x-\lim_{n\to\infty}\int_a^b\Psi_n(x)\mathrm{d}x=\lim_{n\to\infty}\int_a^b(\phi_m(x)-\Psi_n(x))\mathrm{d}x\leqslant 0。$$

对任 $m=1,2,\cdots$ 成立，于是有

$$\lim_{m\to\infty}\int_a^b \phi_m(x)\mathrm{d}x \leqslant \lim_{n\to\infty}\int_a^b \Psi_n(x)\mathrm{d}x。$$

对固定 m,考察 $\Psi_m(x) - \phi_n(x)$,类似上述过程又得

$$\lim_{m\to\infty}\int_a^b \Psi_m(x)\mathrm{d}x \leqslant \lim_{n\to\infty}\int_a^b \phi_n(x)\mathrm{d}x。$$

即恒有 $\lim\limits_{n\to\infty}\int_a^b \phi_n(x)\mathrm{d}x = \lim\limits_{n\to\infty}\int_a^b \Psi_n(x)\mathrm{d}x$。

现在我们引入 Lebesgue 可积函数类 $L(a,b)$,即 $f\in L(a,b)$,必存在 $f_1, f_2 \in C_1$,使 $f(x) = f_1(x) - f_2(x)$, $\forall x\in(a,b)$,此时定义 f 在 (a,b) 上的 Lebesgue 积分为

$$(\mathrm{L})\int_a^b f(x)\mathrm{d}x = \int_a^b f_1(x)\mathrm{d}x - \int_a^b f_2(x)\mathrm{d}x。$$

(2) 也要说明定义的合理性,若还存在 $g_1, g_2 \in C_1$,使 $f(x) = g_1(x) - g_2(x)$, $\forall x \in (a,b)$,则有

$$f_1(x) + g_2(x) = f_2(x) + g_1(x)$$

在 C_1 中积分的可加性是显然的,即明显有

$$\begin{aligned}\int_a^b f_1(x)\mathrm{d}x + \int_a^b g_2(x)\mathrm{d}x = &\int_a^b (f_1(x)\mathrm{d}x + g_2(x))\mathrm{d}x = \\ &\int_a^b (f_2(x)\mathrm{d}x + g_1(x))\mathrm{d}x = \\ &\int_a^b f_2(x)\mathrm{d}x + \int_a^b g_1(x)\mathrm{d}x。\end{aligned}$$

1.4.4 Lebesgue 积分的性质

Lebesgue 积分是 R 积分的推广,设 f 在有限区间 $[a,b]$ 上 R 可积,则 f 在 $[a,b]$ 上几乎处处连续,对每个 n,划分 π_n 是把 $[a,b]$ 分成 2^n 等份:$x_{k+1} - x_k = \dfrac{b-a}{2^n}$,则 π_{n+1} 是 π_n 的加细,得阶梯函数

$\phi_n(x) = \sum_{k=0}^{2^n-1} m_k \chi_{\Delta x_k}(x)$ 为单调增加阶梯函数列，其中 $m_k = \inf\{f(x): x \in \Delta x_k\}$，全体 π_n 的分点是可列集，因而是零测度集，记 E_0 为全体 f 的间断点及 π_n 的分点组成集合，它是零测度集。当 $x \in (a,b) \setminus E_0$ 时，不难得知 $\lim_{n\to\infty}\phi_n(x) = f(x)$，即 $f \in C_1$。分别由 R 积分及 L 积分定义有

$$(\mathrm{R})\int_a^b f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty}\sum_{k=0}^{2^n-1} m_k \mid \Delta x_k \mid =$$

$$\lim_{n\to\infty}\int_a^b \phi_n(x)\mathrm{d}x = (\mathrm{L})\int_a^b f(x)\mathrm{d}x。$$

这说明任 R 可积的函数一定 L 可积，且二者积分相等。另一方面，我们已知函数

$$f(x) = \begin{cases} 1; & \text{当 } x \text{ 取}[0,1] \text{ 中无理数时}, \\ 0; & \text{当 } x \text{ 取}[0,1] \text{ 中有理数时。} \end{cases}$$

在[0,1]上不是 R 可积的。记[0,1]中全体有理数为 $\{r_1, r_2, \cdots, r_n \cdots\} = Q_0$，$E$ 为[0,1]中全体无理数。对每个 $n = 1,2,\cdots$，定义阶梯函数

$$\phi_n(x) = \begin{cases} 0, & \text{当 } x \in \{r_1, r_2, \cdots, r_n\}; \\ -1, & \text{当 } x \in [0,1] \setminus \{r_1, r_2, \cdots, r_n\}。\end{cases}$$

明显 $\phi_n(x)$ 单调增加，$\int_0^1 \phi_n(x)\mathrm{d}x = -1$，且 $\lim_{n\to\infty}\phi_n(x) = f_1(x) = \begin{cases} 0, & \text{当 } x \in Q_0; \\ -1, & \text{当 } x \in E, \end{cases}$ 则 $f_1 \in C_1$，又 $f_2(x) = 0, \forall x \in [0,1]$，也属于 C_1，于是有

$$f(x) = f_2(x) - f_1(x) \text{ 是 L 可积函数。}$$

由此我们看到一个处处不连续的函数也可能 L 可积，并说明 L 积分

比 R 积分更广泛。

以下简单讨论一下这种新型积分的某些性质，在某些极限过程中较 R 积分的性质更强有力。

今后不加说明时，所述积分均指 L 积分，简记 $(\mathrm{L})\int_a^b f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x$，这里 $E = (a,b)$。

以下的几条性质，可直接由定义推得。

(1) $\forall f_1, f_2 \in \mathrm{L}(a,b)$，$\lambda_1, \lambda_2 \in \mathbf{R}$，则 $\lambda_1 f_1 + \lambda_2 f_2 \in \mathrm{L}(a,b)$，且 $\int_E (\lambda_1 f_1(x) + \lambda_2 f_2(x))\mathrm{d}x = \lambda_1 \int_E f_1(x)\mathrm{d}x + \lambda_2 \int_E f_2(x)\mathrm{d}x$，常称此性质为线性的。

(2) 积分关于积分域具有可加性，即 $a < c < b$，若 $f \in \mathrm{L}(a,b)$，则有 $\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x$。

(3) $f(x)$ 在 (a,b) 内几乎处处等于零 $\Leftrightarrow \int_0^1 |f(x)|\mathrm{d}x = 0$。

证　必要性显然。注意函数类 C_1 定义，取阶梯函数列

$$\phi_n(x) = 0, \forall x \in [a,b], n = 1,2,\cdots。$$

即知充分性成立。证毕。

(4) 对于 (a,b) 上函数 f，$f \in \mathrm{L}(a,b) \Leftrightarrow |f| \in \mathrm{L}(a,b)$，此时有 $\left|\int_a^b f(x)\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\mathrm{d}x$。

事实上，我们令 $f^+(x) = \sup\{f(x), 0\}$，$f^-(x) = \sup\{-f(x), 0\}$，不难知道 $f \in \mathrm{L}(a,b) \Leftrightarrow f^+ \in \mathrm{L}(a,b)$，$f^- \in \mathrm{L}(a,b)$。

再注意，$f(x) = f^+(x) - f^-(x)$，$|f(x)| = f^+(x) + f^-(x)$ 易得本性质成立。请注意此性质的充分性对 R 积分不成立。

(5) 设 $f, g \in \mathrm{L}(a,b)$，若 $f(x) \leqslant g(x)$ 在 (a,b) 几乎处处成立，

则有$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x$。

留给读者去验证。

(6) 任给定单调增加函数列 $f_n \in \mathrm{L}(a,b)$，若其积分一致有界，即$\int_a^b f_n(x)\mathrm{d}x \leqslant M, n = 1,2,\cdots$，则$f_n(x)$在$(a,b)$内几乎处处收敛于$f(x) \in \mathrm{L}(a,b)$，且有

$$\lim_{n\to\infty}\int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$$

注 性质(6)称为Levi引理，它告诉我们不能利用C_0类扩充到C_1类的同样方法把$\mathrm{L}(a,b)$扩充到更为广泛的函数类去，并且对于可积函数的单调增序列可积分号下取极限。证明较长，我们略过。

(7) 设$f_n \in \mathrm{L}(a,b)$，在(a,b)内几乎处处收敛于f，并且存在$g \in \mathrm{L}(a,b)$，成立$|f_n(x)| \leqslant g(x), \forall x \in (a,b), n = 1,2,\cdots$，则函数$f \in \mathrm{L}(a,b)$，并且有

$$\lim_{n\to\infty}\int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x。$$

证 对每个自然数n，令$g_n(x) = \sup\{f_n(x), f_{n+1}(x), \cdots\}$，我们要证$g_n(x) \in \mathrm{L}(a,b)$。事实上，因$\sup\{f_n(x), f_{n+1}(x)\} = (f_{n+1}(x) - f_n(x))^+ + f_n(x)$是两个可积函数之和，故其也可积。再由归纳法可证$\sup\{f_n(x), f_{n+1}(x), \cdots, f_{n+m}(x)\} \in \mathrm{L}(a,b)$，而$\sup\{f_n(x), f_{n+1}(x), \cdots, f_{n+m}(x)\}$成为单调增加可积函数列，且成立$\int_a^b \sup\{f_n(x), f_{n+1}(x), \cdots, f_{n+m}(x)\}\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x$，于是由Levi引理，$g_n(x) = \lim\limits_{m\to\infty}\sup\{f_n(x), f_{n+1}(x), \cdots, f_{n+m}(x)\} \in \mathrm{L}(a,b)$。

显然，$g_n(x)$为单调减小，且

$f(x) = \lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} g_n(x)$，在$(a,b)$上几乎处处成立。

令 $h_n(x) = \inf\{f_n(x), f_{n+1}(x), \cdots\} = -\sup\{-f_n(x), -f_{n+1}(x), \cdots\}$，则 $h_n(x) \in L(a,b)$，为单调增加序列且也几乎处处收敛于 $f(x)$。因而，由性质(6)得知 $f \in L(a,b)$。再注意

$$h_n(x) \leqslant f_n(x) \leqslant g_n(x)$$

及 $\lim\limits_{n\to\infty}\int_a^b h_n(x) = \lim\limits_{n\to\infty}\int_a^b g_n(x) = \lim\limits_{n\to\infty}\int_a^b f_n(x)$，立知

$$\lim_{n\to\infty}\int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x,$$

性质(7)得证。

注 性质(7)叫做Lebesgue控制收敛定理，在函数论、微分方程及概率统计理论中是极为常用的工具。

对于R积分，性质(6)与(7)一般不成立。前面举的例子

$$\phi_n(x) = \begin{cases} 0, & \text{当 } x \in \{r_1, r_2, \cdots, r_n\}; \\ -1, & \text{当 } x \in [0,1] \setminus \{r_1, r_2, \cdots, r_n\}。\end{cases}$$

因只有有限个间断点 $\{r_1, r_2, \cdots, r_n\}$，故 $\phi_n(x)$ 在 $[0,1]$ 上R可积。但

$$f(x) = \lim_{n\to\infty}\phi_n(x) = \begin{cases} 0, & \text{当 } x \text{ 为}[0,1]\text{中有理数}; \\ -1, & \text{当 } x \text{ 为}[0,1]\text{中无理数}, \end{cases}$$

在 $[0,1]$ 中不可积。而在L积分意义下，$\phi_n(x)$ 既满足性质(6)的条件，也满足性质(7)的条件，因而有

$$-1 = \int_0^1 f(x)\mathrm{d}x = \lim_{n\to\infty}\int_0^1 \phi_n(x)\mathrm{d}x。$$

这说明L积分的运算性质更灵活些。

(8) 若 $f \in L(a,b)$，则存在阶梯函数 ϕ_n，成立

$$\lim_{n\to\infty}\int_a^b |f(x) - \phi_n(x)| \mathrm{d}x = 0。$$

证 由定义，存在 $f_1, f_2 \in C_1$，使 $f = f_1 - f_2$，于是分别存在满足引理2条件的阶梯函数列 $\phi_n^{(1)}$ 与 $\phi_n^{(2)}$，使在 (a,b) 内分别几乎处处

收敛于 f_1 与 f_2。令 $\phi_n(x)=\phi_n^{(1)}(x)-\phi_n^{(2)}$，则有

$$\int_a^b|f(x)-\phi_n(x)|\mathrm{d}x\leqslant\int_a^b|f_1(x)-\phi_n^{(1)}(x)|\mathrm{d}x+$$

$$\int_a^b|f_2(x)-\phi_n^{(2)}(x)|\mathrm{d}x\to 0(n\to\infty)。$$

证毕。

(9) 若非负函数列 $f_n(x)\in\mathrm{L}(a,b)$，且在 (a,b) 几乎处处收敛于函数 $f(x)$，又若 $\int_a^b f_n(x)\mathrm{d}x\leqslant M,n=1,2,\cdots$，则 $f\in\mathrm{L}(a,b)$，且 $\int_a^b f(x)\mathrm{d}x\leqslant M$。

证　令 $g_n(x)=\inf\{f_n(x),f_{n+1}(x),\cdots\}$，则 $g_n(x)$ 为单调增加函数列，且在 (a,b) 内几乎处处收敛于 $f(x)$。又对每个 $m=1,2,\cdots$ 有 $g_n\leqslant f_{n+m}$，则有

$$\int_a^b g_n(x)\mathrm{d}x\leqslant\int_a^b f_{n+m}(x)\mathrm{d}x\leqslant M。$$

于是由性质(6)得证。证毕。

注　(1) 当 f 在有限区间 (a,b) 内是有界函数且 R 可积，那么 f 是 L 可积的。

(2) 当 f 在有限区间 (a,b) 内无界但广义 R 可积，不能推出 f 是 L 可积的，例如 $f(x)=\dfrac{1}{x}\sin\dfrac{1}{x},(a,b)=(0,1)$。

(3) 当 f 在无穷区间 (a,b) 内广义 R 可积，也不能推出 f 是 L 可积的，例如 $f(x)=\dfrac{\sin x}{x},(a,b)=(0,+\infty)$。

(4) 当 f 在 (a,b)（(a,b) 是有限或无穷区间）内不变号，则若 f 广义 R 可积，那么 f 也一定 L 可积。

习题 1.4

1. 证明定理 2。

2. 设 $\phi_1, \phi_2 \in C_0, \lambda_1, \lambda_2 \in \mathrm{R}$，求证 $\lambda_1\phi_1 + \lambda_2\phi_2 \in C_0$。

3. 设 $\phi_1, \phi_2 \in C_0$，求证 $\phi(x) = \sup\{\phi_1(x), \phi_2(x)\} \in C_0$。

4. 计算 $\lim\limits_{n\to\infty}\int_0^1 \frac{nx}{1+n^2x^2}\mathrm{d}x$。

5. 计算 $\lim\limits_{n\to\infty}\int_0^1 \mathrm{e}^{-nx^2}\mathrm{d}x$。

6. 证明 $\int_0^1 \frac{x^p}{1-x}\ln\frac{1}{x}\mathrm{d}x = \sum\limits_{n=1}^{\infty}\frac{1}{(n+p)^2}$，其中常数 $p > -1$。

7. 讨论函数 $f(x) = \frac{1}{\sqrt{x}}$ 是否可积，如可积，求出积分值。

(1) 在 $(0,1)$ 内是否 R 可积？

(2) 在 $(0,1)$ 内是否广义 R 可积？

(3) 在 $(0,1)$ 内是否 L 可积？

8. 讨论函数 $f(x) = \begin{cases} \sqrt{x}, & \text{当 } x \text{ 为无理数时；} \\ 1, & \text{当 } x \text{ 是有理数时，} \end{cases}$ 是否可积，如可积，求出积分值。

(1) 在 $[0,1]$ 上是否 R 可积？

(1) 在 $[0,1]$ 上是否广义 R 可积？

(3) 在 $[0,1]$ 上是否 L 可积？

1.5　函数类 $\mathrm{L}^p(E)$

$\mathrm{L}^p(a,b)$ 是分析数学中常见的一函数类，取定数值 $p \geq 1$，及有

穷或无穷区间 $E=(a,b)$(也可取作闭区间$[a,b]$),设有函数 f: $E\to\overline{\mathbf{R}}$,使 $|f(x)|^p$ 在E上L可积,则这种函数f的全体记作 $\mathrm{L}^p(E)$。常称 $\mathrm{L}^p(E)$ 为在 E 上 p 次幂可积的函数类。特别当 $p=1$ 时,即 $\mathrm{L}^1(E)$ 就是L可积函数 $\mathrm{L}(a,b)$。

设$f,g\in \mathrm{L}^p(E)$,则存在阶梯函数列 ϕ_n,Ψ_n,使在(a,b)内 ϕ_n 几乎处处收敛于 f, 而 Ψ_n 几乎处处收敛于 g, 于是阶梯函数列 $|\phi_n+\Psi_n|^p$ 几乎处处收敛于 $|f+g|^p$,再注意

$$\lim_{n\to\infty}\int_E|\phi_n+\Psi_n|^p\mathrm{d}x\leqslant\lim_{n\to\infty}\int_E 2^p(|\phi_n(x)|^p+|\Psi_n(x)|^p)\mathrm{d}x=$$

$$2^p(\int_E|f(x)|^p\mathrm{d}x+\int_E|g(x)|^p)\mathrm{d}x<\infty,$$

由性质(上节性质(7))得知$f+g\in \mathrm{L}^p(E)$。这说明 $\mathrm{L}^p(a,b)$ 关于加运算封闭。

设$f\in \mathrm{L}^p(E)$,任 $\lambda\in\mathbf{R}$,易知 $\lambda f\in \mathrm{L}^p(E)$,因此 $\mathrm{L}^p(E)$ 关于数乘运算也封闭。

设$f,g\in \mathrm{L}^p(E)$,若在(a,b)内$f(x)$与$g(x)$几乎处处相等,则把f,g看成同一个元素,可直接记成$f=g$。

设$f\in \mathrm{L}^p(E)$,确定了一个映射 $T:\mathrm{L}^p(E)\to[0,\infty)$ 为

$$T(f)=(\int_E|f(x)|^p_{\mathrm{d}x})^{\frac{1}{p}}。$$

人们常记 $T(f)=\|f\|_p$,称它为f的范数。

由范数的定义及积分的性质,立得以下两条性质。

(1) $\|f\|_p\geqslant 0,\forall f\in \mathrm{L}^p(E)$,且 $\|f\|_p=0\Leftrightarrow f$在$(a,b)$内为零函数 θ,即 $\theta(x)=0,\forall x\in(a,b)$。

(2) 设$f\in \mathrm{L}^p(E),\lambda\in\mathbf{R}$,有 $\|\lambda f\|_p=|\lambda|\,\|f\|_p$。

为获得范数另一条性质,先证明几个常用的不等式。

【引理 1】(Yong 不等式) 对任非负实数 a,b,成立不等式

$$a^{\frac{1}{p}} \cdot b^{\frac{1}{q}} \leqslant \frac{a}{p} + \frac{b}{q}$$

其中 $p > 1, q > 1$,且 $\frac{1}{p} + \frac{1}{q} = 1$。

证 令 $\alpha = \frac{1}{p}, \beta = \frac{1}{q}$,则 $\alpha + \beta = 1$。考虑函数 $y = x^{\alpha}$,因 $y'' = \alpha(\alpha - 1)x^{\alpha-2} < 0$(当 $x > 0$ 时),故 x^{α} 在 $x > 0$ 时为上凸函数,因而函数在点(1,1)的切线位于曲线的上方,则有不等式

$$x^{\alpha} \leqslant \alpha x + \beta \quad (x \geqslant 0),$$

令 $\frac{a}{b} = x$ 代入上式即得 Yong 不等式。

【定理 1】(Hölder 不等式) 设 p,q 如引理 1,则对任 $f \in L^p(E)$, $g \in L^q(E)$,有 $f \cdot g \in L(E)$,且成立

$$\int_E |f(x) \cdot g(x)| \mathrm{d}x \leqslant \|f\|_p \cdot \|g\|_q$$

证 若 f 或 g 是零函数,则定理显然真。现在假设 $\int_E |f(x)|^p \mathrm{d}x > 0, \int_E |g(x)|^q \mathrm{d}x > 0$。此时令

$$a = \frac{|f(x)|^p}{\int_E |f(x)|^p \mathrm{d}x}, \qquad b = \frac{|g(x)|^q}{\int_E |g(x)|^q \mathrm{d}x},$$

代入 Yong 不等式得

$$\frac{|f(x)| \cdot |g(x)|}{\|f\|_p \cdot \|g\|_q} \leqslant \frac{1}{p} \frac{|f(x)|^p}{\int_E |f(x)|^p \mathrm{d}x} + \frac{1}{q} \frac{|g(x)|^q}{\int_E |g(x)|^q \mathrm{d}x},$$

上式右侧为可积函数,则上式表明 $|f(x) \cdot g(x)|$ 有可积控制函数,故 $f \cdot g$ 可积,则对前不等式两侧进行积分得

$$\int_E |f(x) \cdot g(x)| \mathrm{d}x \leqslant \left(\frac{1}{p} + \frac{1}{q}\right) \|f\|_p \cdot \|g\|_q =$$

$$\|f\|_p \cdot \|g\|_q。$$

定理证毕。

注　常称定理1情况为连续型不等式,类似有离散型(或称级数形式)Hölder 不等式:记 $x=(t_1,t_2,\cdots)\in l^p$,则 $\sum_{n=1}^{\infty}|t_n|^p<\infty$,而 $l^q=\{y=(s_1,s_2\cdots):\sum_{n=1}^{\infty}|s_n|^q<\infty\}$,其中 $\frac{1}{p}+\frac{1}{q}=1,p>1,q>1$。成立

$$\sum_{n=1}^{\infty}|t_n s_n|\leqslant(\sum_{n=1}^{\infty}|t_n|^p)^{\frac{1}{p}}\cdot(\sum_{n=1}^{\infty}|s_n|^q)^{\frac{1}{q}}。$$

【定理 2】　取定 $p\geqslant1$,设 $f,g\in L^p(E)$,则有不等式

$$\|f+g\|_p\leqslant\|f\|_p+\|g\|_p。$$

证　当 $p=1$ 时,不等式明显成立。现在设 $p>1$,若 $\|f+g\|_p=0$,则定理真,现在假设 $\|f+g\|_p>0$。已知 $f+g\in L^p(E)$,取 $q>1$,使 $\frac{1}{p}+\frac{1}{q}=1$,则 $|f(x)+g(x)|^{\frac{p}{q}}\in L^p(E)$。应用 Hölder 不等式,得

$$\int_E|f(x)|\cdot|f(x)+g(x)|^{\frac{p}{q}}\mathrm{d}x\leqslant\|f\|_p\cdot$$
$$(\int_E|f(x)+g(x)|^p\mathrm{d}x)^{\frac{1}{q}}$$

及

$$\int_E|g(x)|\cdot|f(x)+g(x)|^{\frac{p}{q}}\mathrm{d}x\leqslant\|g\|_p\cdot$$
$$(\int_E|f(x)+g(x)|^p\mathrm{d}x)^{\frac{1}{q}}。$$

由此推得(注意 $\frac{p}{q}=p-1$)

$$\|f+g\|_p^p=\int_E|f(x)+g(x)|^p\mathrm{d}x\leqslant\int_E(|f(x)|+|g(x)|)\cdot$$

$$|f(x)+g(x)|^{p-1}\mathrm{d}x\leqslant$$

$$(\|f\|_p+\|g\|_p)\left(\int_E|f(x)+g(x)|^p\mathrm{d}x\right)^{\frac{1}{q}}=$$

$$(\|f\|_p+\|g\|_q)\left(\|f+g\|_p^{\frac{1}{q}}\right),$$

从而有

$$\|f+g\|_p\leqslant\|f\|_p+\|g\|_p。$$

证毕。

注　定理所述不等式叫做 Minkowski 不等式。它给出了范数的第 3 条性质:三角不等式。由此我们看到 f 的范数 $\|f\|_p$ 具有完全类似实数 x 的绝对值 $|x|$ 的基本性质。

另外,也成立离散型的 Minkowski 不等式:$p\geqslant 1$,任 $x=(t_1,t_2,\cdots),y=(s_1,s_2,\cdots,)\in l^p$,则有

$$\|x+y\|_p\leqslant\|x\|_p+\|y\|_p,$$

其中 $\|x\|_p=\left(\sum_{n=1}^{\infty}|t_n|^p\right)^{\frac{1}{p}}$。证明类似于定理 2。

关于集 $L^p(E)$ 或 l^p 今后我们还要讨论。

习题 1.5

1. 证明离散型 Hölder 不等式。

2. 证明离散型 Minkowski 不等式。

3. 设 $1\leqslant p<q<\infty$,问两关系式 $L^p(R)\subset L^q(R)$ 与 $L^q(R)\subset L^p(R)$ 是否必有一成立?

4. 若 $f(x)\in L(E)$,问是否有 $f^2(x)\in L(E)$?

5. 设$f(x),g(x)$在$(0,1)$内为非负函数,$f^{\frac{1}{2}}(x),g^{\frac{1}{2}}(x)\in L^2(0,1)$,且在$(0,1)$内几乎处处成立

$$x^{-1}\leqslant f(x)g(x),$$

求证:$4\leqslant\int_0^1 f(x)\mathrm{d}x\int_0^1 g(x)\mathrm{d}x$。

6. 证明$(\sum\limits_{n=1}^{\infty}|t_k|)^2\leqslant n\sum\limits_{n=1}^{\infty}|t_k|^2$。

第 2 章 度量空间与赋范线性空间

2.1 度量空间的基本定义

在第1章2.2节中我们简要介绍了实数集$\mathbf{R}$的度量结构。利用$\mathbf{R}$中两点的距离概念，可以定义$\mathbf{R}$中点列x_n的收敛性，于是才有可能研究函数的分析性质，如连续性、可微性及可积性等等。因此，距离在数学分析中是一基本的概念。现在，我们将实数集$\mathbf{R}$上的度量结构推广到一般集合上去，得到如下重要概念。

【定义 1】 设X是一个非空集合，$\rho(\cdot,\cdot):X\times X\to[0,\infty)$是一个定义在直积集$X\times X$上的二元函数，如果满足如下性质：

(1) $x,y\in X$，则$\rho(x,y)=0$当且仅当$x=y$；

(2) $x,y\in X$，则$\rho(x,y)=\rho(y,x)$；

(3) 任意$x,y,z\in X$，成立三角不等式

$$\rho(x,y)\leqslant\rho(x,z)+\rho(z,y),\tag{2.1}$$

则称$\rho(x,y)$是X中两个元素x与y的距离(或度量)。此时，称X按$\rho(\cdot,\cdot)$成为一个度量空间(或距离空间)，记为(X,ρ)。

注 当不致引起混淆时，(X,ρ)可简记成X，并且常称X中元素为点。在几何空间R^2中，任两点$M(x_1,y_1)$与$N(x_2,y_2)$的距离是

$$\rho(M,N)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}。$$

定义 1 中性质(3) 即是三角形中任一条边的长度不大于其余两条边长度之和，是三角形的一个公理。一般情况下，也称性质(3) 中不等式为三角不等式。

以下介绍几个度量空间的实例。

例 1　设 X 是任何一个非空集合，定义 $\rho(\cdot,\cdot)$ 为

$$\rho(x,y)=\begin{cases}1, & \text{当 } x\neq y;\\ 0, & \text{当 } x=y。\end{cases}$$

那么，$\rho(x,y)$ 满足度量的三个条件，因此是 X 集上的一个度量，称 $(X,\rho(\cdot,\cdot))$ 为离散度量空间。这种度量是最粗的，它只能区分 X 中任何两个元素是否相同，不能区分元素间的远近程度。这个例子说明任何非空集上都有度量存在。

例 2　n 维向量空间 $\mathbf{R}^n$ 表示 n 维向量 $x=(x_1,x_2,\cdots,x_n)$ 的全体组成的集合，也表示 n 个实数 $x_1,x_2,\cdots,x_n$ 组成的数组 $(x_1,x_2,\cdots,x_n)$ 的全体所形成的集合。每个数组 $x=(x_1,x_2,\cdots,x_n)$ 代表 $\mathbf{R}^n$ 中一个点。

对 $x=(x_1,x_2,\cdots,x_n)$，$y=(y_1,y_2,\cdots,y_n)\in\mathbf{R}^n$，定义

$$\rho(x,y)=\Big[\sum_{i=1}^{n}(x_i-y_i)^2\Big]^{\frac{1}{2}}。$$

下面来证 $\rho(\cdot,\cdot)$ 满足度量定义中的条件(1) ~ (3)。

由式(2.1) 不难验证 $\rho(x,y)$ 满足性质(1) 与(2)。来证(3)，需要利用 $P=2$ 时的离散型 Minkowski 不等式(见 1.5 节)。

取 $z=(z_1,z_2,\cdots,z_n)\in\mathbf{R}^n$，则有

$$\rho(x,y)=\Big[\sum_{i=1}^{n}(x_i-y_i)^2\Big]^{\frac{1}{2}}=\Big\{\sum_{i=1}^{n}[(x_i-z_i)+(z_i-y_i)]^2\Big\}^{\frac{1}{2}}\leqslant$$

$$\Big[\sum_{i=1}^{n}(x_i-z_i)^2\Big]^{\frac{1}{2}}+\Big(\sum_{i=1}^{n}(z_i-y_i)^2\Big)^{\frac{1}{2}}=$$

$\rho(x,y)+\rho(z,y)$。

通常 $\rho(x,y)$ 这个度量称为欧氏度量，而$(\mathbf{R}^n,\rho(\cdot,\cdot))$ 称为 n 维欧氏空间。

例 3 所有数列组成的集合 S，即 S 中的每个元素 $\xi=\{a_n\}$ 是实数列。对 $\xi=\{a_n\}$，$\eta=\{b_n\}$ 定义

$$\rho(\xi,\eta)=\sum_{n=1}^{\infty}\frac{1}{2^n}\frac{|a_n-b_n|}{1+|a_n-b_n|},\tag{2.2}$$

那么 $\rho(\xi,\eta)$ 是 S 上的度量。式(2.2)通常称为 Fréchet 组合。$\rho(\xi,\eta)$ 显然满足度量的条件(1) ~ (2)。我们来证也满足条件(3)。

对 ξ,η 及 $\gamma=\{c_n\}\in S$，由于函数 $\varphi(x)=\dfrac{x}{1+x}$ 是单调增函数 $\left(\varphi'(x)=\dfrac{1}{(1+x)^2}>0\right)$，因此由

$$|a_n-b_n|\leqslant|a_n-c_n|+|c_n-b_n|$$

得

$$\frac{|a_n-b_n|}{1+|a_n-b_n|}\leqslant\frac{|a_n-c_n|+|c_n-b_n|}{1+|a_n-c_n|+|c_n-b_n|}\leqslant$$
$$\frac{|a_n-c_n|}{1+|a_n-c_n|}+\frac{|c_n-b_n|}{1+|c_n-b_n|},$$

在上式中同乘$\dfrac{1}{2^n}$再求和便得 $\rho(\xi,\eta)\leqslant\rho(\xi,\gamma)+\rho(\gamma,\eta)$。

例 4 连续函数空间 $C[a,b]$。对两个连续函数 $f,g\in C[a,b]$ 定义 $\rho(f,g)=\max\{|f(t)-g(t)|:t\in[a,b]\}$，那么 $\rho(f,g)$ 是 $C[a,b]$ 上的一个度量。$\rho(f,g)$ 满足度量的(1) ~ (2) 条件。对另一连续函数 $h\in C[a,b]$ 由

$$|f(t)-g(t)|\leqslant|f(t)-h(t)|+|h(t)-g(t)|\leqslant$$
$$\max\{|f(t)-h(t)|:t\in[a,b]\}+\max\{|h(t)-g(t)|:t\in[a,b]\}\leqslant$$

$$\rho(f,h)+\rho(h,g)$$

得 $\rho(f,g)\leqslant\rho(f,h)+\rho(h,g)$，因此 $\rho(f,g)$ 也满足度量定义中的条件(3)。通常连续函数空间 $C[a,b]$ 就是指在此度量意义下的度量空间。如果在集合 $C[a,b]$ 上另外定义

$$\rho_1(f,g)=\int_a^b|f(t)-g(t)|\mathrm{d}t。$$

容易证明 $\rho_1(f,g)$ 也是 $C[a,b]$ 上的一个度量。两个度量空间 $(C[a,b],\rho)$ 与 $(C[a,b],\rho_1)$ 虽然其集合 $C[a,b]$ 是同一个，但度量 ρ 与 ρ_1 不同，因此，是两个完全不同的度量空间。在后面我们证明，这两个度量空间的性质有很大差异。

例 5　函数类 $L^p(E)(p\geqslant 1)$（参见 1.5 节）。对 $f,g\in L^p(E)$，定义

$$\rho(f,g)=\left(\int_E|f(t)-g(t)|^p\mathrm{d}t\right)^{\frac{1}{p}}。$$

由 $\rho(f,g)=0$，得 $\int_E|f(t)-g(t)|^p\mathrm{d}t=0$，根据 Lebesgne 积分的性质有 $f(t)=g(t)a\cdot e$，反之，若 $f(t)=g(t)a\cdot e$，那么 $\rho(f,g)=0$。所以 $\rho(f,g)$ 满足度量定义的条件(1)。条件(2) 显然满足。对另一函数 $h\in L^p(E)$，根据 1.5 节 Minkowski 不等式有

$$\rho(f,g)=\|f-g\|_p\leqslant\|f-h\|_p+\|h-g\|_p=$$
$$\rho(f,h)+\rho(h,g),$$

即 $\rho(f,g)$ 也满足度量定义中的(3)，$L^p(E)$ 在此度量意义下构成一个度量空间。

现在给出度量空间中点列的收敛性概念。

【定义 2】　设 X 是一个度量空间，$x_n,x\in X(n=1,2,\cdots)$，称点列 $\{x_n\}$ 收敛于 x，是指 $\rho(x_n,x)\to 0(n\to\infty)$，$x$ 叫做点列 $\{x_n\}$ 的极限，记作

$$\lim_{n\to\infty} x_n = x \text{ 或 } x_n \to x \quad (n \to \infty)。$$

度量空间中点列的收敛性质与数列的收敛性质有许多共同之处。

【定理1】 设度量空间中点列$\{x_n\}$满足$x_n \to x$，且$x_n \to y(n \to \infty)$，则$x = y$，即极限惟一。

证 由度量三角不等式

$$\rho(x,y) \leqslant \rho(x,x_n) + \rho(x_n,y) = \rho(x_n,x) + \rho(x_n,y),$$

令$n \to \infty$，得$\rho(x,y) \leqslant 0$，由$\rho(x,y) \geqslant 0$，知$\rho(x,y) = 0$。所以$x = y$。证毕。

在例2的欧氏空间$\mathbf{R}^n$中，点列$\{x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})\}$收敛于点$x(x = (x_1, x_2, \cdots, x_n))$等价于按坐标收敛，即对每个$i$，$(1 \leqslant i \leqslant n)$，数列$\{x_i^{(k)}\}$收敛于$x_i$。在例4的连续函数空间$C[a,b]$中，点列$\{f_n\}$收敛于$f$等价于$\{f_n(t)\}$在$[a,b]$上一致收敛于$f(t)$。

我们在学习数列收敛时，已经知道，判别数列收敛的准则是该数列是否是Cauchy数列，因为数列收敛的充要条件是数列是Cauchy列，这完全是由于实数集的完备性所致。在度量空间中这一结果未必总是成立。为此，我们引入一个重要的概念——度量的完备性。

【定义3】 度量空间X中点列$\{x_n\}$称为Cauchy列，是指对任意$\varepsilon > 0$，存在自然数N，当$n,m \geqslant N$时有$\rho(x_n,x_m) < \varepsilon$。度量空间$X$称为是完备的，是指$X$中任何Cauchy列都是收敛的。

例6 欧氏空间$\mathbf{R}^n$是完备度量空间。

证 设$\{x^{(k)}\}$是$\mathbf{R}^n$中任一Cauchy列，则对$\forall \varepsilon > 0$存在自然数N，当$k_1, k_2 \geqslant N$时有$\rho(x^{(k_1)}, x^{(k_2)}) < \varepsilon$。于是对每个坐标所形成的数列$\{x_i^{(k)}\}(x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}))(1 \leqslant i \leqslant n)$，

$$|x_i^{(k_1)} - x_i^{(k_2)}| \leqslant \rho(x^{(k_1)}, x^{(k_2)}) < \varepsilon。$$

这说明$\{x_i^{(k)}\}$是 Cauchy 数列，因此存在实数 x_i，满足 $x_i^{(k)} \to x_i$，记 $x=(x_1,x_2,\cdots,x_n)$，则 $x\in \mathbf{R}^n$。那么，有 $\rho(x^{(k)},x)\to 0(k\to\infty)$，即 $x^{(k)}\to x$。

例 7　连续函数空间 $C[a,b]$ 是完备的。

证　设$\{f_n\}$是 $C[a,b]$ 中任一 Cauchy 列，则对 $\forall\varepsilon>0$，存在自然数 N，当 $n,m\geqslant N$ 时，$\rho(f_n,f_m)<\varepsilon$。对每个固定 $t\in[a,b]$，由于

$$|f_n(t)-f_m(t)|\leqslant\rho(f_n,f_m)<\varepsilon, \tag{2.3}$$

所以数列$\{f_n(t)\}$是 Cauchy 数列，因此是收敛列。

记 $f(t)=\lim\limits_{n\to\infty}f_n(t)$，当 t 在$[a,b]$中变动时，$f(t)$ 便是$[a,b]$上定义的一个函数。在式(2.3)中令 $m\to\infty$，则当 $n\geqslant N$ 时对一切 $t\in[a,b]$ 均有

$$|f_n(t)-f(t)|\leqslant\varepsilon, \tag{2.4}$$

可见函数列$\{f_n(t)\}$一致收敛于 $f(t)$，因 $f_n\in C[a,b]$，所以 $f\in C[a,b]$。式(2.4)同时也说明 $\rho(f_n,f)\leqslant\varepsilon(n\geqslant N)$，即$\{f_n\}$收敛于 f。

例 8　连续函数组成的集合 $C[a,b]$ 上定义与例 7 不同的度量 ρ_1 为

$$\rho_1(f,g)=\int_a^b|f(t)-g(t)|\,\mathrm{d}t$$

时，则度量空间$(C[a,b],\rho_1)$不完备。

证　为了说明度量空间的不完备，仅需构造一个 Cauchy 列，它不是收敛列即可。为方便起见，设$[a,b]=[0,1]$。定义函数列

$$f_n(t)=\begin{cases}1, t\in\left[0,\dfrac{1}{2}-\dfrac{1}{n}\right);\\ -n\left(t-\dfrac{1}{2}\right), t\in\left[\dfrac{1}{2}-\dfrac{1}{n},\dfrac{1}{2}+\dfrac{1}{n}\right];(n>1)\\ -1, t\in\left(\dfrac{1}{2}+\dfrac{1}{n},1\right]。\end{cases}$$

那么,每个 $f_n(t)$ 都是连续函数。对任意 $\varepsilon > 0$,取自然数 N,使 $\frac{1}{N} < \frac{\varepsilon}{2}$,则当 $n,m \geqslant N$ 时,

$$\rho_1(f_n,f_m) = \int_0^1 |f_n(t) - f_m(t)| \, dt \leqslant \frac{1}{n} + \frac{1}{m} \leqslant \frac{2}{N} < \varepsilon,$$

所以 $\{f_n\}$ 是 Cauchy 列。显然,函数列 $\{f_n(t)\}$ 满足对每个 $t \in [0,1]$ 有 $\lim\limits_{n\to\infty} f_n(t) = f(t)$,这里

$$f(t) = \begin{cases} 1, t \in \left[0,\frac{1}{2}\right); \\ 0, t = \frac{1}{2}; \\ -1, t \in \left(\frac{1}{2},1\right]。 \end{cases}$$

而 $f(t)$ 在 $t = \frac{1}{2}$ 处不连续,所以 $f \notin C[0,1]$。因此找不到连续函数 g,使 $\rho_1(f_n,g) \to 0$,否则 $f(t) = g(t)$。这说明 $\{f_n\}$ 虽然是 Cauchy 列,但并不收敛。证毕。

例 7、例 8 表明,两个度量空间定义的集合相同,但度量不同所导致的性质也不同。

具有完备性的度量空间是在该空间中讨论求解各类方程问题中所必备的条件。

习题 2.1

1. 设 X 是度量空间,在 X 中若 $x_n \to x, y_n \to y$,证明 $\rho(x_n,y_n) \to \rho(x,y)$。

2. 证明度量空间 X 中每个收敛点列都是 Cauchy 列。

3. 设 x_n 是度量空间 X 中 Cauchy 列,若存在子列 $x_{n_k} \to x$,证明

$x_n \to x$。

4. 在实数集 $\mathbf{R}$ 中定义

$$\rho(x,y) = |\operatorname{arctg}x - \operatorname{arctg}y|, \forall x, y \in \mathbf{R}$$

证明 ρ 是 $\mathbf{R}$ 上一个度量，但$(\mathbf{R},\rho)$ 不完备。

5. 设非空集合 $X = \mathbf{R}^n$，定义 $\rho_1(x,y) = \max\limits_{1\leq i\leq n}|x_i - y_i|$ ($x = (x_1, x_2, \cdots, x_n)$, $y = (y_1, y_2, \cdots, y_n)$)，证明$(X,\rho_1)$是一个度量空间，并说明$(X,\rho_1)$是完备的。

6. 设 X 是度量空间，证明

$$|\rho(x,z) - \rho(y,z)| \leq \rho(x,y) \quad (x,y,z \in X)。$$

7. 设 X 是有理数集，定义 $\rho(x,y) = |x - y|$，证明(X,ρ)是不完备的度量空间。

8. 证明例 1 中所定义的离散度量空间是完备的。

9. 设 X 是度量空间，$x_n \to x$，证明：对$\{x_n\}$的任何子列$\{x_{n_k}\}$，均有 $x_{n_k} \to x$。

2.2　度量空间中的开、闭集与连续映射

2.2.1　开、闭集的性质

为了更深入研究度量空间中集合的内在结构，我们引入两类重要的集合 —— 开集与闭集。

设 X 是一个度量空间，$x_0 \in X$，$r > 0$ 是一个正数，记

$$B_r(x_0) = \{x \in X : \rho(x, x_0) < r\},$$

$$\bar{B}_r(x_0) = \{x \in X : \rho(x, x_0) \leq r\},$$

分别称为以 x_0 为中心，r 为半径的开球与闭球，也常记 $B_r(x_0)$ 为

$B(x_0,r)$。

【定义 1】 $x_0 \in G, G \subset X$,称 x_0 为 G 的内点,是指存在 $r > 0$,使 $B_r(x_0) \subset G$,又称 G 是开集,如果 G 中每一点都是 G 的内点,集 F 称为闭集,如果 F 的余集 $F^C = X \setminus F$ 是开集。

例 1 任何开球 $B_r(x_0)$ 是开集。

证 设 $x \in B_r(x_0)$,则 $\rho(x,x_0) < r$,令 $r_1 = r - \rho(x,x_0)$,那么 $B_{r_1}(x) \subset B_r(x_0)$。事实上,设 $y \in B_{r_1}(x)$,则 $\rho(y,x) < r_1$,由于

$$\rho(y,x_0) \leqslant \rho(y,x) + \rho(x,x_0) < r_1 + \rho(x,x_0) = r,$$

所以 $y \in B_r(x_0)$。

关于开集有如下性质。

【定理 1】 (1) $\varnothing$(空集),X 是开集;(2) 任意多个开集的并集是开集;(3) 有限多个开集的交集是开集。

证 (1) 显然。(2) 设一族开集 $\{G_\lambda\}_{\lambda\in\Gamma}$,来证 $\bigcup\limits_{\lambda\in\Gamma} G_\lambda$ 也是开集。设 $x \in \bigcup\limits_{\lambda\in\Gamma} G_\lambda$,那么存在 $\lambda_0 \in \Gamma$,使得 $x \in G_{\lambda_0}$。因 G_{λ_0} 是开集,所以有 $r > 0$,且 $B_r(x) \subset G_{\lambda_0}$,更有 $B_r(x) \subset \bigcup\limits_{\lambda\in\Gamma} G_\lambda$,即 $\bigcup\limits_{\lambda\in\Gamma} G_\lambda$ 是开集。(3) 设 $G_1, G_2, \cdots, G_n$ 是有限个开集,来证 $\bigcap\limits_{i=1}^{n} G_i$ 也是开集。设 $x \in \bigcap\limits_{i=1}^{n} G_i$,那么对每个 i,$x \in G_i$,又 G_i 是开集,那么存在 $r_i > 0$,使得 $B_{r_i}(x) \subset G_i$。取 $r = \min(r_1, r_2, \cdots, r_n)$,所以 $B_r(x_0) \subset \bigcap\limits_{i=1}^{n} G_i$,即 $\bigcap\limits_{i=1}^{n} G_i$ 是开集。证毕。

注意 定理 1 中的(3),对无限多个开集结论未必成立。

例 2 取 $G_n = \left(0, 1 + \dfrac{1}{n}\right)(n = 1,2,\cdots)$,则 $\bigcap\limits_{n=1}^{\infty} G_n = (0,1]$,而 $(0,1]$ 不是开集。

根据第一章 De Morgan 律对于闭集有

【定理 2】 (1) $\varnothing, X$ 是闭集;(2) 任意多个闭集的交是闭集;(3) 有限多个闭集的并是闭集。

闭集是通过开集定义的，这对闭集的结构了解有一定的障碍，为此，我们引入如下概念。

【定义 2】　设 X 是度量空间，$x_0 \in X$，$A \in X$，称 x_0 是 A 的聚点，是指对 $\forall \epsilon > 0$，都有 $(B_\epsilon(x_0) \setminus \{x_0\}) \cap A \neq \varnothing$。$A$ 的聚点全体所成集合称为 A 的导集，记为 A'，集合 $\bar{A} = A \cup A'$ 称为 A 的闭包。

例 3　设 $A = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\right\}$，则 $A' = \{0\}$，$\bar{A} = \left\{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\right\}$。

注　对于一个度量空间中的集 A，其闭包 $\bar{A}$ 中的点分成两类，一类是聚点，而另一类称为孤立点，即 $x \in A$，但 $x \notin A'$，如上例中 $1, \frac{1}{2}, \frac{1}{3}, \cdots$ 均是 A 的孤立点。

【定理 3】　A 为闭集的充要条件是 $A = \bar{A}$。

证　$A = \bar{A}$ 等价于 $A' \subset A$。设 A 为闭集，则 A^C 是开集。来证 $A^C \subset (A')^C$，这便是 $A' \subset A$。设 $x \in A^C$，有 $r > 0$，使 $B_r(x) \subset A^C$，故 $B_r(x) \cap A = \varnothing$，更有

$$(B_r(x) \setminus \{x\}) \cap A = \varnothing,$$

那么 $x \notin A'$，即 $x \in (A')^C$。反之，由 $A' \subset A$ 来证 A^C 是开集。设 $x \in A^C$，则 $x \notin A'$。于是有 $r > 0$，使 $(B_r(x) - \{x\}) \cap A = \varnothing$，因 $x \notin A$，所以 $B_r(x) \cap A = \varnothing$，即 $B_r(x) \subset A^C$，这说明 A^C 是开集。证毕。

注　A 为闭集的充要条件是对任意 A 中点列 x_n，若 $x_n \to x(n \to \infty)$，则 $x \in A$。X 是度量空间，$A \subset X$，A 的直径记为 $d(A)$，即 $d(A) = \sup\{\rho(x, y): x, y \in A\}$。如果 $d(A) < +\infty$，称 A 为有界集。

【定理 4】　设 X 是完备度量空间，$\{A_n\}$ 是一列非空有界闭集，若满足

(1) $A_1 \supset A_2 \supset A_3 \supset \cdots$;

(2) $d(A_n) \to 0(n \to \infty)$,

则 $\bigcap\limits_{n=1}^{\infty} A_n$ 为单点集,即有 $x \in X$,满足 $\bigcap\limits_{n=1}^{\infty} A_n = \{x\}$。

证 取 $x_n \in A_n(n = 1,2,\cdots)$。来证 $\{x_n\}$ 是 Cauchy 列。由(2) $d(A_n) \to 0$, 则对 $\forall \varepsilon > 0$, 存在自然数 N, 使当 $n \geq N$ 时, $d(A_n) < \varepsilon$,于是对自然数 $n, m \geq N$ 时,由(1), $A_n \subset A_N$, $A_m \subset A_N$,因此 $\rho(x_n, x_m) \leq d(A_N) < \varepsilon$。又度量空间 X 完备,所以存在 $x \in X$,使 $x_n \to x$。对任何自然数 k,由(1)及 x_n 的取法, $\{x_n\}_{n \geq k} \subset A_k$, A_k 是闭集,故 $x \in A_k$。即 $x \in \bigcap\limits_{k=1}^{\infty} A_k$。若有 $y \in \bigcap\limits_{k=1}^{\infty} A_k$,则 $\rho(x, y) \leq d(A_n) \to 0(n \to \infty)$, 即 $\rho(x, y) \leq 0$,故有 $x = y$。这便证得 $\bigcap\limits_{k=1}^{\infty} A_k = \{x\}$。证毕。

2.2.2 度量空间上的连续映射

【定义 3】 设 X 与 Y 是两个度量空间,其度量分别为 ρ 与 ρ_1。T: $X \to Y$ 是一个映射, $x_0 \in X$。称 T 在 x_0 点连续,是指对 $\forall \varepsilon > 0$,存在 $\delta > 0$,当 $\rho(x, x_0) < \delta$ 时,有

$$\rho_1(Tx, Tx_0) < \varepsilon。$$

若 T 在 X 中每一点连续,则称 T 为 X 上的连续映射。

度量空间之间的连续映射是数学分析中连续函数概念的推广。特别,当映射的值域空间 $Y = R$ 时,映射就是度量空间上的函数。

【定理 5】 设 X, Y 是两个度量空间, $T: X \to Y$, $x_0 \in X$,则下列命题是等价的。

(1) T 在 x_0 点连续;

(2) 对 $\forall \varepsilon > 0$,存在 $\delta > 0$,当 $x \in B(x_0, \delta)$ 时,有

$$Tx \subset B(Tx_0, \varepsilon);$$

(3) 对于 X 中任意点列 $\{x_n\}$，若 $x_n \to x_0 (n \to \infty)$，则 $Tx_n \to Tx_0 (n \to \infty)$。

证　(1)⇒(2)，由定义显然，(2)⇒(3)。由于 $x_n \to x_0$，存在自然数 N，当 $n \geqslant N$ 时，$\rho(x_n, x_0) < \delta$，即 $x_n \in B(x_0, \delta)$，因此 $Tx_n \in B(Tx_0, \varepsilon)$，即 $\rho_1(Tx_n, Tx_0) < \varepsilon$。(3)⇒(1)，用反证法。若 T 在 x_0 点不连续，则存在 $\varepsilon_0 > 0$，使对任意 $\delta > 0$，存在 $x_\delta \in X$，且 $\rho(x_\delta, x_0) < \delta$，但 $\rho_1(x_\delta, x_0) \geqslant \varepsilon_0$。特别，取 $\delta = \dfrac{1}{n} (n = 1, 2, \cdots)$，则有 $x_n \in X$，$\rho(x_n, x_0) < \dfrac{1}{n}$，但 $\rho_1(Tx_n, Tx_0) \geqslant \varepsilon_0$。这意味着，虽然 $\rho(x_n, x_0) < \dfrac{1}{n}$，但 $\rho_1(Tx_n, Tx_0) \geqslant \varepsilon_0$。也就是 $x_n \to x_0$，但 $Tx_n \not\to Tx_0$，矛盾。证毕。

下面定理是通过开集与闭集来刻画连续映射的。

【定理 6】　设 X, Y 是两个度量空间，$T: X \to Y$ 是一个映射，则下述命题等价：

(1) T 是连续映射；

(2) 对于 Y 中任何开集 G，$T^{-1}(G)$ 是 X 中开集；

(3) 对于 Y 中任何闭集 F，$T^{-1}(F)$ 是 X 中闭集。

证　(1)⇒(2)，设 $x_0 \in T^{-1}(G)$，则 $Tx_0 \in G$。因 G 是开集，所以存在 $\varepsilon > 0$，满足 $B(Tx_0, \varepsilon) \subset G$。对此 ε，由 T 在 x_0 点连续，存在 $\delta > 0$，当 $x \in B(x_0, \delta)$ 时，$Tx \in B(Tx_0, \varepsilon)$，即 $T(B(x_0, \delta)) \subset B(T(x_0, \varepsilon) \subset G$，故

$$B(x_0, \delta) \subset T^{-1}(G)。$$

(2)⇒(1)，对任意 $x_0 \in X$，及 $\varepsilon > 0$，取 $G = B(Tx_0, \varepsilon)$，那么 $T^{-1}(G)$ 是 X 中开集。而 $x_0 \in T^{-1}(G)$，所以有 $\delta > 0$，满足 $B(x_0, \delta) \subset T^{-1}(G)$，即 $T(B(x_0, \delta)) \subset G = B(Tx_0, \varepsilon)$。这说明 T 在 x_0 点连续。

再由 x_0 的任意性，T 在 X 的每一点都连续。

(2)⇒(3)。对任何闭集 F，F 的余集 F^C 是开集。根据映射象与原象的性质有

$$T^{-1}(F^C) = (T^{-1}(F))^C,$$

这便是(2)⇒(3)。对于任何开集 G，G^C 是闭集。同样有

$$T^{-1}(G^C) = (T^{-1}(G))^C,$$

这便是(3)⇒(2)。证毕。

注 关于映射的性质 $T^{-1}(A^C) = (T^{-1}(A))^C$ 留作习题。

【定义4】 设 X,Y 是两个度量空间，度量分别为 ρ 与 ρ_1，称 X 与 Y 是等距同构的，是指存在一一映射 $T:X\to Y$ 满足

$$\rho_1(Tx,Ty) = \rho(x,y) \quad (x,y\in X),$$

T 称为等距同构映射。

从定义4可见，$T:X\to Y$ 及 $T^{-1}:Y\to X$ 都是连续映射。两个等距同构的度量空间，从结构上来看，完全相同，可以视为同一个空间。

习题2.2

1. 设度量空间 $X=\mathbf{R}$，$\mathbf{Q}\subset X$ 是有理数集，求 $\mathbf{Q}'$，$\overline{\mathbf{Q}}$。

2. 设 X 是度量空间，$x_0\in X$，$r>0$，证明，闭球 $\overline{B}_r(x_0)$ 是闭集。

3. 设 X 是度量空间，$A\subset X$，A° 表示全体 A 的内点构成的集合，称为 A 的内部。证明 A° 是开集。

4. 设 X 是一度量空间，$x_0\in X$。证明：$f(x)=\rho(x,x_0)$ 是 X 上的连续函数。

5. 设 X 是度量空间，$F\subset X$ 是一个非空闭集，对 $x\in X$，记 $\inf\{\rho(x,y):y\in F\}=\rho(x,F)$。证明：对任何 $r>0$，集合

$$\{x\in X:\rho(x,F)<r\}$$

是开集。

6. 设 F_1 与 F_2 都是度量空间 X 中的闭集，且 $F_1 \cap F_2 = \varnothing$，证明存在开集 G_1, G_2，使 $F_1 \subset G_1, F_2 \subset G_2$，且 $G_1 \cap G_2 = \varnothing$。

7. 设 X 是度量空间，$A \subset X$，若 $x_0 \in A'$，证明对任意 $\varepsilon > 0$，集 $A \cap B(x_0, \varepsilon)$ 是无限集。

8. 设 X 是度量空间，$A, B \subset X$，证明

(1) 若 $A \subset B$，则 $\bar{A} \subset \bar{B}$；

(2) $\bar{\bar{A}} = \bar{A}$；

(3) $\overline{A \cup B} = \bar{A} \cup \bar{B}$；

(4) $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$，并举例说明等号未必成立。

9. 设 X 是度量空间，证明：(1) X 中每一非空闭集必为可列个开集的交；(2) X 中每一非空开集必为可列个闭集的并。

10. 设 X, Y 是两个非空集合。$T: X \to Y$ 是一个映射。$A \subset Y$。证明：

$$T^{-1}(A^C) = (T^{-1}(A))^C。$$

2.3　度量空间的可分性与紧性

有理数在实数中的稠密性及有界数列必有收敛子序列，是数学分析理论的源泉。本节的任务是把实数的这两个重要性质推广到一般度量空间中去，形成两个新的概念 —— 可分性与紧性。

2.3.1　度量空间的可分性

【定义 1】　设 X 是一个度量空间，A 与 B 都是 X 的子集，称 A 在 B 中稠密，是指 $B \subset \bar{A}$。

例 1　有理数在实数中稠密，有理数也在无理数中稠密。

从定义可见，A 在 B 中稠密，意指 B 中元素可由 A 中元素任意接近，即对 $x \in B$，必有 A 中一列元素 x_n，使 $\rho(x_n, x) \to 0$。

【定义 2】　称度量空间 X 是可分的，是指存在 X 中一可列集 A，使 A 在 X 中稠密。

例 2　欧氏空间 $\mathbf{R}^n$ 是可分的。

证　取 $A = \{(r_1, r_2, \cdots, r_n): r_i \text{ 是有理数}(1 \leqslant i \leqslant n)\}$，则 A 是可列集。对 $x \in \mathbf{R}^n$ 及 $\varepsilon > 0$，记 $x = (x_1, x_2, \cdots, x_n)$。每个 $x_i (1 \leqslant i \leqslant n)$ 是实数，取有理数 r_i 满足 $|x_i - r_i| < \dfrac{\varepsilon}{n}$。令 $a = (r_1, r_2, \cdots, r_n)$，则 $a \in A$，由于

$$\rho(x, a) = \sqrt{\sum_{i=1}^{n}(x_i - r_i)^2} < \sqrt{\sum_{i=1}^{n}\frac{\varepsilon^2}{n^2}} = \varepsilon,$$

所以 A 在 $\mathbf{R}^n$ 中稠密。

例 3　连续函数空间 $C[a, b]$ 是可分的

证　记 A 为系数是有理数的多项式组成的集合。由第一章的例子知道 A 是可数集。对任一连续函数 $f(t)$，根据 Weierstrass 定理，对 $\forall \varepsilon > 0$，存在多项式 $P(t)$ 满足

$$\rho(f, P) = \max\{|f(t) - P(t)|; t \in [a, b]\} < \frac{\varepsilon}{2}。$$

取多项式 $P_0 \in A$，满足

$$\rho(P, P_0) = \max\{|P(t) - P_0(t)|; t \in [a, b]\} < \frac{\varepsilon}{2}。$$

于是 $\rho(P_0, f) \leqslant \rho(P_0, P) + \rho(P, f) < \varepsilon$。从而 A 在 $C[a, b]$ 中稠密。

可以证明 $L^p(a, b)(p > 1)$ 也是可分的度量空间。

下面举一个不可分度量空间的例子。

例 4　有界数列 l^∞ 空间。l^∞ 是由所有有界数列组成的集合。在

l^∞ 上定义度量

$$\rho(\xi,\eta)=\sup_{i\geq 1}|a_i-b_i|,$$

这里 $\xi=\{a_i\},\eta=\{b_i\}$,则 l^∞ 在度量 ρ 下是不可分的。

证　我们用反证法。若 l^∞ 可分,则存在可列稠密集 A。取 l^∞ 的一个子集为

$$B=\{\xi=\{a_i\}:a_i=0\text{ 或 }1(i=1,2,\cdots)\}。$$

B 可以与 $[0,1]$ 中实数通过二进制小数建立如下对应:$\{a_i\}\leftrightarrow 0.a_1a_2a_3\cdots$,该对应是一一映射,因此 B 是不可数集。以 A 中的所有点为中心,$\frac{1}{3}$ 为半径的开球 $B(a,\frac{1}{3})(a\in A)$ 满足

$$\bigcup_{a\in A}B(a,\frac{1}{3})=l^\infty,$$

因此 $\bigcup_{a\in A}B(a,\frac{1}{3})\supset B$。由于 A 可数,B 不可数,所以至少存在 B 中两不同 ξ,η 落入某个开球 $B(a_0,\frac{1}{3})(a_0\in A)$。直接计算,显然 $\rho(\xi,\eta)=1$,但

$$\rho(\xi,\eta)\leq\rho(\xi,a_0)+\rho(\eta,a_0)<\frac{1}{3}+\frac{1}{3}=\frac{2}{3}$$

矛盾,故 l^∞ 不可数。

2.3.2　度量空间的紧性

【定义 3】　设 X 是度量空间,$A\subset X$,称 A 为列紧集,如果 A 中的每个点列都存在一个子列收敛于 X 中的某一点。称 A 是紧集,如果 A 中每个点列都存在一个子列收敛于 A 中的某一点。

由定义可见,一个集合是紧集必是列紧集,但反之不然。

例 5　$X=R,A=(0,1],\{\frac{1}{n}\}\subset A,\lim\limits_{n\to\infty}\frac{1}{n}=0$,但 $0\notin A$,因此

A 是列紧集，但不是紧集。

例 6 $X = \mathbf{R}^n, A \subset \mathbf{R}^n$ 是有界集，则 A 是列紧集。

证 $\{x^{(k)}\} \subset A$，记 $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$。由 A 有界知存在常数 $M > 0$，使 $|x_i^{(k)}| \leq M (1 \leq i \leq n)$，每个数列 $\{x_i^{(k)}\}$ 都是有界的。对 $\{x_1^{(k)}\}$，由第一章学过的 Bolzano - Weierstrass 定理，存在子序列收敛，记为 $\{x_1^{(k_1)}\}$，$\{k_1\}$ 为自然数集的子集。$\{x_2^{(k_1)}\}$ 仍是有界的，故又存在收敛子序列，记为 $\{x_2^{(k_2)}\}$，$\{k_2\}$ 是 $\{k_1\}$ 的子集。依次类推，得到自然数集的子列 $\{k_n\}$ 使 $\{x_1^{(k_n)}\}, \{x_2^{(k_n)}\}, \cdots, \{x_n^{(k_n)}\}$ 都收敛，因此 $x^{(k_n)}$ 在 $\mathbf{R}^n$ 中收敛，即 A 为列紧集。

一般地，度量空间中的有界集未必列紧。

例 7 $X = l^\infty, A = \{e_i\}$，这里 e_i 为如下数列

$$e_i = (0, 0, \cdots, 1, 0, 0 \cdots),$$

即该数列的第 i 项是 1，其余各项全是 0。由于 $d(A) = 1$，所以 A 是有界集。但因为

$$\rho(e_i, e_j) = 1 \quad (i \neq j)。$$

因此 $\{e_i\}$ 任何子列均不收敛，这说明 A 不是列紧集。

【定义 4】 设 X 是度量空间，$A \subset X$，称 A 是全有界集，如果对 $\forall \varepsilon > 0$，存在 A 中有限个点 $x_1, x_2, \cdots, x_n$ 满足

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

【定理 1】 如果 A 是度量空间 X 中的列紧集，则 A 是全有界集。

证 用反证法。若 A 不是全有界集，那么必存在 $\varepsilon_0 > 0$，使得 A 中任意有限个点为中心半径为 ε_0 的球的并不能盖住 A。取 $x_1 \in A$，球 $B(x_1, \varepsilon_0)$ 不能盖住 A，于是存在 $x_2 \in A$，且 $x_2 \notin B(x_1, \varepsilon_0)$，即有 $\rho(x_2, x_1) \geq \varepsilon_0$。同样，$B(x_1, \varepsilon_0) \cup B(x_2, \varepsilon_0)$ 也不能盖住 A，存在 $x_3 \in A$，且 $x_3 \notin B(x_1, \varepsilon_0) \cup B(x_2, \varepsilon_0)$，即 $\rho(x_1, x_3) \geq \varepsilon_0, \rho(x_2, x_3) \geq$

ϵ_0。如此继续下去，得到 A 中点到 $\{x_n\}$ 满足

$$\rho(x_n, x_m) \geqslant \epsilon_0 \quad (n \neq m)。$$

可见点列 x_n 的任何子列均不能收敛，这与 A 是列紧集矛盾。证毕。

进一步还可以证明：

【定理 2】 如果度量空间 X 完备，那么 A 全有界当且仅当 A 列紧。

证 仅需来证由全有界推出列紧。设 A 是全有界集，$\{x_n\} \subset A$。取 $\epsilon = \frac{1}{k}(k = 1,2,\cdots)$。对 $\epsilon = 1$，有 A 中有限个点为中心，半径为 1 的球的并盖住 A，所以必有某个球 $B(a_1, 1)$ 中含有 $\{x_n\}$ 的某个子列，记该子列为 $\{x_n^{(1)}\}$，取 $\epsilon = \frac{1}{2}$，同样有 A 中有限个点为中心，半径为 $\frac{1}{2}$ 的球盖住 A，所以必有某个球 $B(a_2, \frac{1}{2})$ 含有子列 $\{x_n^{(1)}\}$ 的子列，记为 $\{x_n^{(2)}\}$。如此进行下去，可得子列串为 $\{x_n^{(1)}\}, \{x_n^{(2)}\}, \cdots$，其中后一个总是前一个的子列，且 $\{x_n^{(k)}\} \in B(a_k, \frac{1}{k})$，从这一个子列串中重新选择一个子列 $\{x_n^{(n)}\}$ 即下面将子列串排成下表，选取对角线元素而得

$$\begin{array}{ccccc}
x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} & \cdots \\
x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} & \cdots \\
x_1^{(3)} & x_2^{(3)} & x_3^{(3)} & x_4^{(3)} & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
x_1^{(n)} & x_2^{(n)} & \cdots & \cdots & \cdots
\end{array}$$

我们来证 $\{x_n^{(n)}\}$ 是 Cauchy 列，事实上，对任意 $\epsilon > 0$，取自然数 N，使 $\frac{2}{N} < \epsilon$，则对任何 $n, m \geqslant N$ 时有 $\{x_n^{(n)}\}, \{x_m^{(m)}\} \in B(a_N, \frac{1}{N})$，所以

$$\rho(x_n^{(n)}, x_m^{(m)}) \leqslant \rho(x_n^{(n)}, a_N) + \rho(x_m^{(m)}, a_N) < \frac{2}{N} < \varepsilon。$$

即有$\{x_n^{(n)}\}$是 Cauchy 列，从而由 X 完备，$\{x_n^{(n)}\}$ 是收敛列。证毕。

现在我们将古典分析中闭区间上连续函数的某些性质推广到度量空间的紧集上。

【定理 3】 设 A 是度量空间 X 中一个紧集，$f(x)$ 是定义在 A 上的一个连续函数，那么 $f(x)$ 是有界的，且上、下确界可达。

证　先证 $f(x)$ 有界。若不然，则存在 $x_n \in A$，使 $\lim\limits_{n\to\infty} |f(x_n)| = \infty$。由于 A 紧，$\{x_n\}$ 有子列 $\{x_{n_k}\}$ 在 A 中收敛，即有 $x_0 \in A$，使 $x_{n_k} \to x_0(k \to \infty)$，再由于 $f(x)$ 在 x_0 点连续，有 $f(x_0) = \lim\limits_{k\to\infty} f(x_{n_k})$，这便是 $|f(x_0)| = +\infty$，这显然不可能。记 $\beta = \sup\limits_{x\in A} f(x)$，由上确界的定义，同样可找到 A 中点列 $\{x_n\}$，满足 $f(x_n) > \beta - \frac{1}{n}$，由 A 紧，存在子列 (x_{n_k}) 及 $x_0 \in A$，使 $x_{n_k} \to x_0(n \to \infty)$。由 $f(x)$ 在 x_0 点连续，得 $f(x_0) = \lim\limits_{k\to\infty} f(x_{n_k}) \geqslant \beta$。显然，$f(x_0) \leqslant \beta$，于是 $f(x_0) = \beta$。同理，可证下确界可达。证毕。

关于判别重要空间 $C[a,b]$ 中子集的列紧性有下述著名的 Arzela – Ascoli 定理。

【定理 4】 集合 $A \subset C[a,b]$ 是列紧的充分必要条件是下列两个条件成立：

(1) A 是一致有界的，即存在常数 $M > 0$，使对每个 $f \in A$，有 $\max\{|f(t)| : t \in [a,b]\} \leqslant M$；

(2) A 是等度连续的，即对 $\forall \varepsilon > 0$，存在 $\delta > 0$，使对任意 $t_1, t_2 \in [a,b]$，且 $|t_2 - t_1| < \delta$ 及 $f \in A$ 成立

$$|f(t_1) - f(t_2)| < \varepsilon。$$

该定理证明较为烦杂，故这里略去。

习题 2.3

1. 设 A 是列紧集且是闭集，证明：A 是紧集。

2. 设 A 是度量空间 X 中的稠密集，证明：对任何开球 $B_r(x_0)$，都有 $A \cap B_r(x_0) \neq \varnothing$。

3. 设 $\{F_n\}$ 是一列非空紧集，若满足

$$F_1 \supset F_2 \supset F_3 \supset \cdots,$$

则 $\bigcap\limits_{n=1}^{\infty} F_n \neq \varnothing$。

4. 设 $x \in X$，A 是 X 中紧集，记

$$\rho(x, A) = \inf\{\rho(x, y) : y \in A\},$$

证明：当 $\rho(x, A) = 0$，那么 $x \in A$。如果将 A 换为列紧集，结论是否成立。

5. 记 $l^1 = \{\xi = \{a_i\} : \sum\limits_{i=1}^{\infty} | a_i | < +\infty\}$，在 l^1 上定义度量为

$$\rho(\xi, \eta) = \sum_{i=1}^{\infty} | a_i - b_i |,$$

这里 $\xi = \{a_i\}$，$\eta = \{b_i\} \in l^1$。证明：l^1 可分。

6. 设 X 是度量空间，$A \subset X$，证明 A 在 X 中稠密 $\Leftrightarrow A^C$ 无内点。

7. 举例说明有界集未必是全有界集。

8. 设 X 是度量空间，$A \subset X$ 是紧集，$B \subset X$ 是闭集，记 $\rho(A, B) = \inf\{\rho(a, b) : a \in A, b \in B\}$，若 $\rho(A, B) = 0$，证明 $A \cap B \neq \varnothing$。

9. 设 M 是度量空间 X 中紧集，F_n 是 M 中非空闭子集列，且 $F_1 \supset F_2 \supset \cdots$，证明 $\bigcap\limits_{n=1}^{\infty} F_n = \varnothing$。

2.4　压缩映象原理及其应用

压缩映象原理是求解代数方程、微分方程、积分方程以及数值分析中迭代算法收敛性的理论依据，是数学和工程计算中最常用的方法之一。

2.4.1　Banach 压缩映象原理

在众多情况下，求解各种方程的问题可转化为求某一映射的不动点，压缩映象原理就是某一类映射不动点的存在和惟一性问题，特别，不动点可以通过迭代序列求出。

【定义 1】　设 X 是一个度量空间，$T: X \to X$ 是一个映射，称 x_* 是 T 的不动点，是指 $Tx_* = x_*$。

【定义 2】　设 X 是一个度量空间，$T: X \to X$ 称为压缩映象，是指存在常数 $a \in [0,1)$ 满足

$$\rho(Tx, Ty) \leqslant a\rho(x,y), \quad (x, y \in X)$$

从定义 2 可见，压缩映象一定是连续映射。因为若 $x_n \to x$，由 $\rho(Tx_n, Tx) \leqslant a\rho(x_n, x)$ 得 $Tx_n \to Tx$。

【定理 1】（Banach 压缩映象原理）　设 X 是完备度量空间，$T: X \to X$ 是压缩映象，那么 T 存在惟一的不动点。

证　任取 $x_0 \in X$，作迭代序列 $x_n = Tx_{n-1} (n \geqslant 1)$。为证明 $\{x_n\}$ 是收敛点列，仅需证明它是 Cauchy 点列，因为 X 是完备的。由于

$$\rho(x_2, x_1) = \rho(Tx_1, Tx_0) \leqslant a\rho(x_1, x_0),$$

$$\rho(x_3, x_2) = \rho(Tx_2, Tx_1) \leqslant a\rho(Tx_1, Tx_0) \leqslant a^2\rho(x_1, x_0)$$

…　…　…

$\rho(x_n,x_{n-1}) = \rho(Tx_{n-1},Tx_{n-2}) \leqslant a\rho(Tx_{n-2},Tx_{n-3}) \leqslant \cdots \leqslant a^{n-1}\rho(x_1,x_0)$，于是对任何自然数 n 及 P，有

$$\begin{aligned}\rho(x_{n+p},x_n) &\leqslant \rho(x_{n+p},x_{n+p-1}) + \rho(x_{n+p-1},x_{n+p-2}) + \cdots \\ &\quad + \rho(x_{n+1},x_n) \leqslant \\ &\quad (a^{n+p-1} + a^{n+p-2} + \cdots + a^n)\rho(x_1,x_0) = \\ &\quad \frac{a^n(1-a^p)}{1-a}\rho(x_1,x_0) \leqslant \frac{a^n}{1-a}\rho(x_1,x_0),\end{aligned}$$

可见，$\rho(x_{n+p},x_n) \to 0(n\to\infty)$，因此 $\{x_n\}$ 是 Cauchy 列。从而存在 X 中 x_* 使 $x_n \to x_*$。我们来证 x_* 便是 T 的不动点。事实上由

$$\begin{aligned}\rho(Tx_*,x_*) &\leqslant \rho(Tx_*,Tx_n) + \rho(Tx_n,x_*) \leqslant \\ &\quad a\rho(x_*,x_n) + \rho(x_{n+1},x_*) \to 0(n\to\infty)\end{aligned}$$

得 $\rho(Tx_*,x_*) \leqslant 0$，即 $\rho(Tx_*,x_*) = 0$，故 $Tx_* = x_*$。

最后来证惟一性。设另有一不动点 y_*，即有 $Ty_* = y_*$，因

$$\rho(x_*,y_*) = \rho(Tx_*,Ty_*) \leqslant a\rho(x_*,y_*),$$

$0 \leqslant a < 1$，所以 $\rho(x_*,y_*) = 0$，即 $x_* = y_*$。证毕。

注　(1) 从定理的证明过程中发现，迭代序列的初始值可任意选取，最终都能收敛到惟一不动点。

(2) 该定理提供了近似计算不动点的误差估计公式

$$\rho(x_*,x_n) \leqslant \frac{a^n}{1-a}\rho(Tx_0,x_0)。$$

(3) 因为完备度量空间的任何闭子集在原有度量下仍然是完备的，所以定理中的压缩映象不需要在整个空间 X 上有定义，只要在某个闭集上有定义，且象也在该闭集内，定理的结论依然成立。

在实际应用过程中，有时 T 本身未必是压缩映象，但 T 的若干次复合 T^n 是压缩映象，这时 T 仍然有惟一不动点，这就是如下所述的对压缩映象原理的改进定理。

【定理 2】 设 X 是完备度量空间，$T:X\to X$ 是一个映射。如果存在某个自然数 n_0，使 T^{n_0} 是压缩映射，那么 T 存在惟一的不动点（这里 T^{n_0} 是 T 的 n_0 次复合，即 $T^2x=T(Tx)$，$T^3x=T(T^2x)$，…，$T^{n_0}x=T(T^{n_0-1}x)$）。

证 T^{n_0} 是压缩映象，所以存在惟一不动点 x_* 即 $T^{n_0}x_*=x_*$，由于

$$T^{n_0}(Tx_*)=T^{n_0+1}x_*=T(T^{n_0}x_*)=Tx_*,$$

这说明 Tx_* 仍是 T^{n_0} 的不动点，而 T^{n_0} 的不动点惟一，所以 $Tx_*=x_*$，即 x_* 是 T 的不动点。若 T 另有不动点 y_* 即 $Ty_*=y_*$，则 $T^{n_0}y_*=T^{n_0-1}(Ty_*)=T^{n_0-1}y_*=\cdots=y_*$，那么 y_* 也是 T^{n_0} 的不动点，再据 T^{n_0} 不动点的惟一性有 $x_*=y_*$。证毕。

2.4.2 压缩映象原理的应用

本小节通过代数方程、微分方程及积分方程来说明定理 1 与定理 2 的具体应用。

例 1 线性代数方程组 $Ax=b$ 均可写为如下形式

$$x=Cx+D,\tag{2.5}$$

其中 $C=(c_{ij})_{n\times n}$，$D=(d_1,d_2,\cdots,d_n)^T$。如果矩阵 C 满足条件 $\sum\limits_{j=1}^{n}|c_{ij}|<1(i=1,2,\cdots,n)$，则式(2.5)存在惟一解，且此解可由迭代求得。

证 取 $X=\mathbf{R}^n$，定义度量为

$$\rho(\xi,\eta)=\max_{1\leqslant i\leqslant n}|a_i-b_i|,$$

$\xi=(a_1,a_2,\cdots,a_n)^T$，$\eta=(b_1,b_2,\cdots,b_n)^T$。构造映射 $T:X\to X$ 为 $Tx=Cx+D$，那么方程(2.5)的解等价于映射 T 的不动点。对于 $x=(x_1,x_2,\cdots,x_n)^T$，$y=(y_1,y_2,\cdots,y_n)^T$，由于

$$\rho(Tx,Ty)=\max_{1\leqslant i\leqslant n}\left|\left(\sum_{j=1}^{n}(c_{ij}x_j+d_j)\right)-\left(\sum_{j=1}^{n}(c_{ij}y_j+d_j)\right)\right|=$$

$$\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^{n}c_{ij}(x_j-y_j)\right|\leqslant\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|c_{ij}|\rho(x,y)。$$

记 $a=\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{n}|c_{ij}|$，由条件 $a<1$，因此 T 是压缩映象，于是 T 有惟一不动点，所以方程(2.5)有惟一解，且此解可由如下迭代序列

$$x^{(k)}=Cx^{(k-1)}+D$$

近似计算求得。

例 2　考察如下常微分方程的初值问题

$$\begin{cases}\dfrac{dy}{dx}=f(x,y)\\ y(x_0)=y_0\end{cases}\tag{2.6}$$

如果 $f(x,y)$ 在 $\mathbf{R}^2$ 连续，且关于第二变元 y 满足 Lipschitz 条件：

$$|f(x,y_1)-f(x,y_2)|\leqslant L|y_1-y_2|。$$

这里 $L>0$ 是常数，则方程(2.6)在 $[x_0-\delta,x_0+\delta]$ 上有惟一解 $(\delta<\dfrac{1}{L})$。

证　方程(2.6)的解等价于如下积分方程

$$y(x)=y_0+\int_{x_0}^{x}f(t,y(t))dt\tag{2.7}$$

的解。取连续函数空间 $\mathbf{C}[x_0-\delta,x_0+\delta]$，定义其上的映射 $T:\mathbf{C}[x_0-\delta,x_0+\delta]\to\mathbf{C}[x_0-\delta,x_0+\delta]$ 为

$$(Ty)(x)=y_0+\int_{x_0}^{x}f(t,y(t))dt。$$

则积分方程(2.7)的解等价于 T 的不动点。对任意两个连续函数 $y_1(x),y_2(x)\in\mathbf{C}[x_0-\delta,x_0+\delta]$，由于

$$\rho(Ty_1, Ty_2) = \max_{x\in[x_0-\delta, x_0+\delta]}\left|\int_{x_0}^{x}[f(t, y_1(t)) - f(t, y_2(t))]\mathrm{d}t\right| \leqslant$$

$$\max_{x\in[x_0-\delta, x_0+\delta]}\int_{x_0}^{x}|f(t, y_1(t)) - f(t, y_2(t))|\mathrm{d}t \leqslant$$

$$\max_{x\in[x_0-\delta, x_0+\delta]}L\int_{x_0}^{x}|y_1(t) - y_2(t)|\mathrm{d}t \leqslant \delta L\rho(y_1, y_2),$$

令 $a = L\delta$，则 $a < 1$，故 T 是压缩映射，从而 T 有惟一不动点，即积分方程(2.7) 有惟一解，从而微分方程(2.6) 在 $[x_0 - \delta, x_0 + \delta]$ 上有惟一解。

例 3 设 $K(s,t)$ 是定义在 $[a,b]\times[a,b]$ 上的二元连续函数，则对于任何常数 λ 及任何给定的连续函数 $f(t)\in\mathbf{C}[a,b]$，如下的 Volterra 型积分方程

$$x(t) = \lambda\int_a^t K(s,t)x(s)\mathrm{d}s + f(t) \tag{2.8}$$

存在惟一解。

证 取连续函数空间 $\mathbf{C}[a,b]$，其上定义映射 $T:\mathbf{C}[a,b]\to\mathbf{C}[a,b]$ 为

$$(Tx)(t) = \lambda\int_a^t K(s,t)x(s)\mathrm{d}s + f(t),$$

则方程(2.8) 的解等价于映射 T 的不动点。由于 $K(s,t)$ 在 $[a,b]\times[a,b]$ 上连续，于是 $|K(s,t)|$ 在 $[a,b]\times[a,b]$ 上有最大值，记为 M，即 $M = \max\{|K(s,t)|:(s,t)\in[a,b]\times[a,b]\}$。对任何两个连续函数 $x_1(t), x_2(t)$，由于

$$|(Tx_1)(t) - (Tx_2)(t)| = |\lambda|\left|\int_a^t K(s,t)[x_1(s) - x_2(s)]\mathrm{d}s\right| \leqslant$$

$$|\lambda|\int_a^t |K(s,t)||x_1(s) - x_2(s)|\mathrm{d}s \leqslant$$

$$|\lambda| M(t-a)\max_{a\leqslant s\leqslant b}|x_1(s)-x_2(s)|\leqslant$$
$$|\lambda| M(t-a)\rho(x_1,x_2),$$

$$|(T^2x_1)(t)-(T^2x_2)(t)|=|\lambda|\left|\int_a^t K(s,t)[(Tx_1)(s)-(Tx_2)(s)]\mathrm{d}s\right|\leqslant$$
$$|\lambda|^2M^2\rho(x_1,x_2)\int_a^t(s-a)\mathrm{d}s=$$
$$\frac{|\lambda|^2M^2(t-a)^2}{2}\rho(x_1,x_2),$$

一般地,对自然数 n,归纳可得

$$|(T^nx_1)(t)-(T^nx_2)(t)|\leqslant\frac{|\lambda|^nM^n(t-a)^n}{n!}\rho(x_1,x_2)。$$

因此 $\rho(T^nx_1,T^nx_2)=\max\limits_{a\leqslant t\leqslant b}|(T^nx_1)(t)-(T^nx_2)(t)|\leqslant$

$$\frac{|\lambda|^nM^n(b-a)^n}{n!}\rho(x_1,x_2)。$$

注意到 $\lim\limits_{n\to\infty}\dfrac{|\lambda|^nM^n(b-a)^n}{n!}=0$,因此存在自然数 n_0,满足

$$\frac{|\lambda|^{n_0}M^{n_0}(b-a)^{n_0}}{n_0!}=\alpha<1。$$

这说明 T^{n_0} 是压缩映象,由定理 2,T 有惟一不动点,亦即 Volterra 型积分方程(2.8)有惟一解。

压缩映象原理在数值分析、微分方程、代数方程等众多领域中有丰富的应用示例,希望读者能很好地体会此方法的优点。

习题 2.4

1. 用压缩映象原理证明方程 $x=a\sin x$ 只有惟一解 $x=0$,其中 $a\in(0,1)$。

2. 证明下述线性方程组

$$\begin{bmatrix} \frac{3}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{7} & \frac{5}{14} & -\frac{1}{5} \\ -\frac{1}{100} & -\frac{1}{120} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{7} \\ \frac{1}{8} \end{bmatrix}$$

有惟一解，并写出求方程近似解的迭代序列。

3. 用压缩映象原理构造迭代序列来求下述微分方程

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = 1 + x^2 \\ x(0) = 0 \end{cases}$$

的解。

4. 设 $K(s,t)$ 是 $[a,b]\times[a,b]$ 上的连续函数，$f(t)\in C[a,b]$，记 $M = \max\{|K(s,t)|:(s,t)\in[a,b]\times[a,b]\}$。证明下述 Fredholm 积分方程

$$x(t) = \lambda\int_a^b K(s,t)x(s)\mathrm{d}s + f(t)$$

当 $|\lambda| < \dfrac{1}{M(b-a)}$ 时有惟一解。

5. 设 X 是完备度量空间，$T:X\to X$，如果

$$a = \inf_n \sup_{x\neq y}\frac{\rho(T^n x, T^n y)}{\rho(x,y)} < 1,$$

证明 T 存在惟一不动点。

2.5 线性空间

在线性代数中，我们已经学过有限维向量空间 $\mathbf{R}^n$。对于向量，存

在两种基本运算即向量的加法与向量和数的乘积。本节的目的是将向量空间的概念一般化，建立一般线性空间。

2.5.1　线性空间的定义

【定义 1】　设 X 是一个非空集合，F 是实数域或复数域，称 X 为 F 上的线性空间，如果满足以下条件：

(1) 对任意两个元素 $x,y\in X$，存在 X 中惟一一个元素 u 与之对应，u 称为 x 与 y 的和，记为 $u=x+y$，且满足：

① 交换律：$x+y=y+x\quad(x,y\in X)$；

② 结合律：$x+(y+z)=(x+y)+z\quad(x,y,z\in X)$；

③ 在 X 中存在一个元素 θ，称为零元素使 $x+\theta=x(x\in X)$；

④ 对每个 $x\in X$，存在另一个元素，称为 x 的负元素，记为 $-x$，使 $x+(-x)=\theta$。

以上 4 条是对加法运算所要求的，下面 4 条是对数乘运算要求的。

(2) 对于任意数 α 及 $x\in X$，存在 X 中惟一元素 v 与之对应，记为 $v=\alpha x$，称为 α 与 x 的数乘，且满足：

① 结合律：$\alpha(\beta x)=(\alpha\beta)x$；$(\alpha,\beta\in F,x\in X)$

② $1x=x$；

③ 数乘对加法分配律：$(\alpha+\beta)x=\alpha x+\beta x$；

④ 加法对数乘分配律：$\alpha(x+y)=\alpha x+\alpha y$。

如果 $F=\mathbf{R}$(实数域)，称 X 为实线性空间。如果 $F=\mathbf{C}$(复数域)，称 X 为复线性空间。

例 1　欧氏空间 $\mathbf{R}^n$ 是一线性空间。加法为：$x=(a_1,a_2,\cdots,a_n)$，$y=(b_1,b_2,\cdots,b_n)$，$x+y=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$；数乘为 $\alpha x=(\alpha a_1,\alpha a_2,\cdots,\alpha a_n)$；零元素 $\theta=(0,0,\cdots,0)$；负元素为

$-x=(-a_1,-a_2,\cdots,-a_n)$。

例 2 所有系数为实数的多项式全体按多项式的加法和数乘组成一实线性空间。

例 3 $C[a,b]$ 按函数的加法与数乘运算组成一线性空间。

【定义 2】 设 $x_1,x_2,\cdots,x_n\in X$，如果存在不全为零的数 $\alpha_1,\alpha_2,\cdots,\alpha_n\in F$ 使得

$$\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n=\theta,$$

则称 $x_1,x_2,\cdots,x_n$ 是线性相关的。否则，称为线性无关。

例 4 线性空间 $X=C[a,b]$，那么 $\sin t,\sin 2t,\sin 3t,\sin 4t$ 是线性无关的，而 $\sin t,2\sin t,3\sin 2t,4\sin 2t$ 是线性相关的。

【定义 3】 如果线性空间 X 中可找到 n 个线性无关的向量且任意 $n+1$ 个向量均线性相关，则称 X 的维数为 n，记为 $\dim X=n$。对任何自然数 m，X 中都有 m 个线性无关的向量，则称 X 是无限维的，记为 $\dim X=\infty$。n 维线性空间中 n 个线性无关的向量称为空间的一组基。

例 5 $\mathbf{R}^n$ 空间是 n 维线性空间；$C[a,b]$ 是无限维线性空间。

X 是线性空间，$M\subset X$，称 M 是 X 的子空间，是指对任何 $x,y\in M$，及 $\alpha,\beta\in F$，有 $\alpha x+\beta y\in M$，亦即，M 在 X 的原有运算下仍构成一个线性空间。

例 6 例 2 中的线性空间是 $C[a,b]$ 的子空间。

设 A 是线性空间 X 的子集，记

$$\operatorname{span}A=\{\sum_{i=1}^{n}\alpha_ix_i:\alpha_i\in F,x_i\in A,i=1,2,\cdots,n;n\geq 1\},$$

称 $\operatorname{span}A$ 为 A 的线性包。易证 $\operatorname{span}A$ 是子空间，而且是包含 A 的最小子空间。

例 7 设 X 是线性空间，x_1,x_2 是线性无关的，则 $\operatorname{span}\{x_1,x_2\}$ 是

X 的 2 维子空间。

在线性空间中还有一类常用集合——凸集。一个集合 $A \subset X$ 称为是凸集，如果对任意 A 中两个元 x, y 及 $\lambda \in [0,1]$ 有 $\lambda x + (1-\lambda)y \in A$。特别，当 A 是子空间时，A 一定是凸集；相反，凸集未必是子空间。

2.4.2　线性算子与线性泛函

【定义 4】　设 X 与 Y 是两个线性空间，映射 $T: X \to Y$ 称为线性算子，如果对 $\forall x, y \in X$ 及 $\alpha, \beta \in F$，有 $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$，特别，当 $Y = F$ 时，线性算子称为线性泛函，F 是实数域时，称实线性泛函，F 是复数域时，称复线性泛函。

$T: X \to Y$ 是线性算子，记

$$\mathrm{Ker}(T) = \{x \in X : Tx = \theta\},$$

$$R(T) = \{y \in Y : y = Tx, x \in X\},$$

分别称为线性算子 T 的零空间和值域空间。容易证明，$\mathrm{Ker}(T)$ 是 X 的子空间，而 $R(T)$ 是 Y 的子空间。

【定义 5】　两个线性空间 X 与 Y 称为是同构的，是指存在一个线性算子 $T: X \to Y$ 是一一映射。

两个同构的线性空间维数相同。

例 8　设 X 与 Y 分别是 n 维与 m 维线性空间。X 中取一组基 $\{e_1, e_2, \cdots, e_n\}$，$Y$ 中取一组基为 $\{\delta_1, \delta_2, \cdots, \delta_m\}$，此时对任何一个线性算子 $T: X \to Y$ 存在相应的一个矩阵 $A = (a_{ij})_{m \times n}$ 使得若

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

则 $Tx = \sum_{j=1}^{m} \beta_j \delta_j$，其中 $\beta_j = \sum_{i=1}^{n} a_{ij} \alpha_i \ (j = 1, 2, \cdots, m)$。

证 对每个 e_i，Te_i 是 Y 中元素，存在 m 个数 $a_{i1},a_{i2},\cdots,a_{im}$ 使得

$$Te_i = \sum_{j=1}^{m} a_{ij}\delta_j,$$

于是有

$$Tx = T(\sum_{i=1}^{n}\alpha_i e_i) = \sum_{i=1}^{n}\alpha_i Te_i =$$

$$\sum_{i=1}^{n}\alpha_i(\sum_{j=1}^{m}a_{ij}\delta_j) = \sum_{j=1}^{m}(\sum_{i=1}^{n}a_{ij}\alpha_i)\delta_j = \sum_{j=1}^{m}(\beta_j\delta_j)$$

例 8 表明，两个有限维线性空间之间的线性算子均可在合适的基下通过矩阵表达。因此，线性代数所研究的矩阵本质上是有限维空间之间的线性算子。本书的主要目的是研究无限维空间上的线性算子。

例 9 连续函数空间 $C[a,b]$，其子空间 $C^1[a,b]$ 即 $[a,b]$ 上全体连续可微函数组成的线性空间，定义算子 $T=\dfrac{d}{dt}:C^1[a,b]\to C[a,b]$，则 T 是线性算子。

例 10 连续函数空间 $C[a,b]$，定义泛函为

$$f(x) = \int_a^b x(t)dt,$$

则 f 是线性泛函。

习题 2.5

1. 设 X 是线性空间，$A\subset X$ 是非空子集。证明：$\text{span}A$ 是子空间且满足对任一子空间 M，若 $M\supset A$，则 $M\supset\text{span}A$。

2. 设 M,N 是线性空间 X 的两个子空间，证明：$M+N=\{x+y: x\in M, y\in N\}$ 及 $M\cap N$ 均是子空间。$M\cup N$ 是否是子空间？

3. 设 X, Y 是两个线性空间，$T: X \to Y$ 是线性算子，证明：$\mathrm{Ker}(T)$ 和 $R(T)$ 分别是 X 及 Y 的子空间。

4. 证明：对欧氏空间 $\mathbf{R}^n$，任意线性泛函 $f: \mathbf{R}^n \to \mathbf{R}$ 都惟一存在 $a = (a_1, a_2, \cdots, a_n)$，是 n 个确定的实数，使对每个 $x = (x_1, x_2, \cdots, x_n) \in \mathbf{R}^n$ 都有

$$f(x) = \sum_{i=1}^{n} a_i x_i。$$

5. 下列函数集合按照函数的加法及数乘运算是否构成线性空间？

(1) $[a, b]$ 上所有次数 $\leqslant 3$ 的多项式全体。

(2) $[a, b]$ 上所有次数 $= 3$ 的多项式全体。

(3) $[a, b]$ 上满足 $x(a) = 0$ 的函数全体。

(4) $\mathbf{R}$ 上连续且周期为 2π 的函数全体。

(5) $\mathbf{R}$ 上一切单调函数全体。

6. 设 X 是 n 维实线性空间，证明 X 与 $\mathbf{R}^n$ 同构。

2.6　赋范线性空间

在前几节中我们在集合上引进了度量的概念，并且在度量意义下研究了点列的收敛及其映射的性质。本节我们在线性空间中引进一种与代数运算相联系的度量，即由向量范数诱导出的度量。

【定义 1】　设 X 是线性空间，函数 $\|\cdot\|: X \to [0, \infty)$ 称为 X 上定义的一个范数，如果满足：

(1) $\|x\| = 0$ 当且仅当 $x = \theta$；

(2) 对任意 $x \in X$ 及 $\alpha \in F$，$\|\lambda x\| = |\lambda| \|x\|$；

(3) 对任意 $x,y\in X$, $\|x+y\|\leqslant\|x\|+\|y\|$。

称二元体$(X,\|\cdot\|)$为赋范线性空间。

在 X 是赋范线性空间时,由范数导出的度量为

$$\rho(x,y)=\|x-y\|。$$

此时 X 在此度量意义成为度量空间。所以,赋范线性空间是一种特殊的度量空间。点列$\{x_n\}$收敛于 x,即 $\rho(x_n,x)=\|x_n-x\|\to 0$ $(n\to\infty)$,有时称这种收敛为依范数收敛。

【定义 2】 赋范线性空间称为 Banach 空间,是指由范数导出的度量是完备的。

例 1 欧氏空间 $\mathbf{R}^n$,连续函数空间 $C[a,b]$, $L^p[a,b]$ $(p\geqslant 1)$ 在下列范数下均是赋范线性空间,而且是 Banach 空间:

$$\|x\|=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}},x=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n;$$

$$\|x\|=\max\{|x(t)|:t\in[a,b]\},x\in C[a,b];$$

$$\|x\|=\left(\int_a^b|x(t)|^p\mathrm{d}t\right)^{\frac{1}{p}},x\in L^p[a,b]。$$

例 2 $\varphi=\{\xi=\{a_i\}$:存在某个数 i_0,使 $i>i_0$ 时,$a_i=0\}$,即 φ 是仅有有限项非零的所有实数列组成的集合。规定加法和数乘运算为

$$\xi+\eta=\{a_i+b_i\},\alpha\xi=\{\alpha a_i\},$$

$\xi=\{a_i\},\eta=\{b_i\},\alpha\in F$。此时 φ 是线性空间,定义范数为

$$\|\xi\|=\sup_{i\geqslant 1}|a_i|\ (\xi=\{a_i\}\in\varphi),$$

则$(\varphi,\|\cdot\|)$是赋范线性空间,但不是 Banach 空间。

证 仅需说明 φ 在范数导出的度量意义下不完备。取 $\xi^{(n)}=\{1,\frac{1}{2},\frac{1}{3},\cdots,\frac{1}{n},0,\cdots\}$。由于当 $n>m$ 时

$$\rho(\xi^{(n)},\xi^{(m)}) = \frac{1}{m+1} \to 0(m \to \infty),$$

所以 $\{\xi^{(n)}\}$ 是 Cauchy 点列，但它不收敛。若不然，存在 $\xi \in \varphi$ 使 $\|\xi^{(n)} - \xi\| \to 0(n \to \infty)$，于是求得数列 ξ 为 $\{1, \frac{1}{2}, \frac{1}{3}, \cdots\}$，这个数列的每一项均非零，因此 $\xi \notin \varphi$。矛盾。

2.6.1　赋范线性空间的性质

【性质 1】　设 X 是赋范线性空间，$\{x_n\} \subset X, x \in X$，若 $x_n \to x$，则 $\{x_n\}$ 是有界点列。

证　$\|x_n - x\| \to 0(n \to \infty)$，对 $\epsilon = 1$，存在自然数 N，当 $n > N$ 时，$\|x_n - x\| < 1$。于是

$$\|x_n\| \leqslant \|x\| + \|x_n - x\| < 1 + \|x\|。$$

令 $M = \max\{\|x_1\|, \|x_2\|, \cdots, \|x_N\|, 1 + \|x\|\}$，那么对一切自然数 n，均有 $\|x_n\| \leqslant M$，即 $\{x_n\}$ 有界。证毕。

【性质 2】　设 X 中点列 $\{x_n\}$，$\{y_n\}$ 及数域 F 中数列 $\{\alpha_n\}$ 满足 $x_n \to x, y_n \to y, \alpha_n \to \alpha(x, y \in X, \alpha \in F)$，则

(1) 加法连续：$x_n + y_n \to x + y$，即 $\lim\limits_{n\to\infty}(x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$；

(2) 数乘连续：$\alpha_n x_n \to \alpha x$，即 $\lim\limits_{n\to\infty}(\alpha_n x_n) = (\lim\limits_{n\to\infty}\alpha_n)\cdot(\lim\limits_{n\to\infty}x_n)$。

证　(1) 由 $\|x_n + y_n - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\|$，得 $x_n + y_n \to x + y$。

(2) 因 $x_n \to x$，由性质 1，$\{x_n\}$ 有界，所以存在常数 $M > 0$，满足 $\|x_n\| \leqslant M$。于是

$$\|\alpha_n x_n - \alpha x\| = \|\alpha_n x_n - \alpha x_n + \alpha x_n - \alpha x\| \leqslant$$
$$\|\alpha_n x_n - \alpha x_n\| + \|\alpha x_n - \alpha x\| =$$

$$|\alpha_n - \alpha| \, \|x_n\| + |\alpha| \, \|x_n - x\| \leqslant$$
$$M|\alpha_n - \alpha| + |\alpha| \, \|x_n - x\|$$

故 $\alpha_n x_n \to \alpha x$。证毕。

线性空间上两种范数的比较,我们有:

【定义 2】 设 X 是一线性空间,$\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 X 上定义的两个范数,称 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强,如果 $\|x_n\|_1 \to 0$,则 $\|x_n\|_2 \to 0$。又称 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价,如果 $\|x_n\|_1 \to 0$ 当且仅当 $\|x_n\|_2 \to 0$。

【性质 3】 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强当且仅当存在常数 $\lambda > 0$,使 $\|x\|_2 \leqslant \lambda \|x\|_1 (x \in X)$。

证 若上述不等式成立,显然 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强。若上述不等式不成立,则对任何自然数 n,存在 $x_n \in X$,使 $\|x_n\|_2 > n\|x_n\|_1$。令

$$y_n = \frac{x_n}{\|x_n\|_2}$$

于是 $\|y_n\|_2 = 1$,但 $\|y_n\|_1 < \dfrac{1}{n}$,这说明 $\|y_n\|_1 \to 0$,但 $\|y_n\|_2 \not\to 0$。这与 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强矛盾。证毕。

由性质 3,两个范数 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价当且仅当存在常数 $\lambda \geqslant \mu > 0$ 成立

$$\mu\|x\|_2 \leqslant \|x\|_1 \leqslant \lambda\|x\|_2。$$

例 3 在连续函数空间 $C[a,b]$ 上定义两种范数为

$$\|x\|_1 = \max\{|x(t)| : t \in [a,b]\}, \quad \|x\|_2 = \int_a^b |x(t)| \, dt,$$

则 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强,但两个范数不等价。

证 若 $\|x_n\|_1 \to 0$,则函数列 $\{x_n(t)\}$ 在 $[a,b]$ 上一致收敛于

恒等 0 的函数，于是

$$\lim_{n\to\infty}\int_a^b |x_n(t)|\,\mathrm{d}t = \int_a^b \lim_{n\to\infty}|x_n(t)|\,\mathrm{d}t = 0,$$

即 $\|x_n\|_2 \to 0$。若取连续函数列 $\{x_n(t)\}$ 为

$$x_n(t) = \begin{cases} 0, t \in [a + \dfrac{1}{n}, b]; \\ -n(t-a)+1, t \in [a, a+\dfrac{1}{n}]; \end{cases} \qquad (n = 1,2,\cdots)$$

则 $\|x_n\|_2 = \int_a^b |x_n(t)|\,\mathrm{d}t = \dfrac{1}{2n} \to 0(n\to\infty)$，但 $\|x_n\|_1 = 1$。

【定理 1】(Riesz 引理)　设 Y 是赋范线性空间 X 的一个真闭子空间，则对 $\forall \varepsilon > 0$，存在 $x_0 \in X$ 且 $\|x_0\| = 1$，满足

$$\rho(x_0, Y) = \inf\{\|y - x_0\| : y \in Y\} \geqslant 1 - \varepsilon。$$

证　因 $Y \neq X$，所以有 $x_1 \in X \setminus Y$。记

$$d = \inf\{\|y - x_1\| : y \in Y\},$$

则 $d > 0$。若不然，存在 $y_n \in Y$，使 $\|y_n - x_1\| \to d = 0$，由 Y 闭，得 $x_1 \in Y$，这与 x_1 的选取矛盾。设 $\varepsilon \in (0,1)$，那么 $\dfrac{d}{(1-\varepsilon)} > d$。根据下确界定义，存在 $y_1 \in Y$，使 $\|y_1 - x_1\| < \dfrac{d}{1-\varepsilon}$。令 $x_0 = \dfrac{x_1 - y_1}{\|x_1 - y_1\|}$，则 $\|x_0\| = 1$，且对任意 $y \in Y$，有

$$\|y - x_0\| = \left\| y - \frac{x_1 - y_1}{\|x_1 - y_1\|} \right\| = \frac{\| \|x_1 - y_1\| y + y_1 - x_1 \|}{\|x_1 - y_1\|}。 \tag{2.9}$$

又 $y, y_1 \in Y$，Y 是子空间，所以 $\|x_1 - y_1\| y + y_1 \in Y$，据 d 的定义及式(2.9)有

$$\|y - x_0\| \geqslant \frac{d}{\|x_1 - y_1\|} > 1 - \varepsilon,$$

即 $d(x_0,Y)\geqslant 1-\varepsilon$。证毕。

Riesz 引理是泛函分析的一条重要定理，它在有限维与无限维赋范线性空间的特征区别中起着关键作用。

2.6.2 有限维赋范线性空间

当线性空间的维数有限时，这类赋范线性空间具有许多良好性质。但当维数无限时，有限维空间所表现的某些特征已不复存在。

【定理 2】 设 X 是 n 维赋范线性空间，$\{e_1,e_2,\cdots,e_n\}$ 为 X 的一组基，则存在两个常数 $\lambda\geqslant\mu>0$，使对每个 $x\in X$，$x=\sum\limits_{i=1}^{n}\alpha_ie_i$ 时成立

$$\mu\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}}\leqslant\|x\|\leqslant\lambda\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}}\qquad(2.10)$$

证 由 Cauchy 不等式

$$\|x\|=\left\|\sum_{i=1}^{n}\alpha_ie_i\right\|\leqslant\sum_{i=1}^{n}|\alpha_i|\,\|e_i\|\leqslant\left(\sum_{i=1}^{n}\|e_i\|^2\right)^{\frac{1}{2}}\cdot\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}},$$

令 $\lambda=\left(\sum\limits_{i=1}^{n}\|e_i\|^2\right)^{\frac{1}{2}}$，则有 $\|x\|\leqslant\lambda\left(\sum\limits_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}}$，即式(2.10)右端不等式成立。

另一方面，考虑 $\mathbf{R}^n$ 的单位球面

$$S=\left\{\xi=(\alpha_1,\alpha_2,\cdots,\alpha_n):\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}}=\|\xi\|=1\right\},$$

在 S 上定义函数 $f(\xi)=\left\|\sum\limits_{i=1}^{n}\alpha_ie_i\right\|=\|x\|$，对于 $\xi,\eta\in S$，$\xi=(\alpha_1,\alpha_2,\cdots,\alpha_n)$，$\eta=(\beta_1,\beta_2,\cdots,\beta_n)$，由于

$$|f(\xi)-f(\eta)|=\big|\|x\|-\|y\|\big|\leqslant\|x-y\|\leqslant$$

$$\lambda\left(\sum_{i=1}^{n}|\alpha_i-\beta_i|^2\right)^{\frac{1}{2}}=\lambda\|\epsilon-\eta\|,$$

这里 $x=\sum_{i=1}^{n}\alpha_ie_i,y=\sum_{i=1}^{n}\beta_ie_i$,所以 $f(\xi)$ 在 S 上连续。而 S 是 $\mathbf{R}^n$ 中有界闭集,那么 S 是紧集。据本章 2.3 节定理 3,$f(\xi)$ 在 S 上达到下确界,即存在 $\xi_0\in S$,使 $f(\xi_0)=\inf\{f(\xi):\xi\in S\}$。显然 $f(\xi_0)\geqslant 0$。若 $f(\xi_0)=0$,则 $\left\|\sum_{i=1}^{n}\alpha_i^0e_i\right\|=0$,这里 $\xi_0=(\alpha_1^0,\alpha_2^0,\cdots,\alpha_n^0)$。从而有

$$\sum_{i=1}^{n}\alpha_i^0e_i=\theta。$$

因 $e_1,e_2,\cdots,e_n$ 线性无关,所以 $\alpha_1^0=\alpha_2^0=\cdots=\alpha_n^0=0$,那么 $\|\xi_0\|=0$,这与 $\xi_0\in S$ 矛盾。记 $\mu=f(\xi_0)$,则对每个 $\xi=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in\mathbf{R}^n$,如果 $\xi\neq 0$,那么 $\frac{\xi}{\|\xi\|}\in S$,于是 $f\left(\frac{\xi}{\|\xi\|}\right)\geqslant\mu$,即

$$\left\|\sum_{i=1}^{n}\frac{\alpha_i}{\|\xi\|}e_i\right\|\geqslant\mu,$$

亦即

$$\|x\|=\left\|\sum_{i=1}^{n}\alpha_ie_i\right\|\geqslant\mu\|\xi\|=\mu\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}},$$

这说明式(2.10)左端不等式成立。特别,当 $\|x\|=0$ 时,$\sum_{i=1}^{n}\alpha_ie_i=\theta$,即 $\alpha_1^0=\alpha_2^0=\cdots=\alpha_n^0=0$,这时式(2.10)不等式均变为等式。证毕。

【推论 1】　设 X 是 n 维线性空间,$\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 X 上定义的两个范数,则 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价。

证　由定理 2 存在常数 $\lambda_1\geqslant\mu_1>0$ 及 $\lambda_2\geqslant\mu_2>0$,满足

$$\mu_1\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}}\leqslant\|x\|_1\leqslant\lambda_1\left(\sum_{i=1}^{n}|\alpha_i|^2\right)^{\frac{1}{2}},$$

$$\mu_2(\sum_{i=1}^{n}|\alpha_i|^2)^{\frac{1}{2}} \leqslant \|x\|_2 \leqslant \lambda_2(\sum_{i=1}^{n}|\alpha_i|^2)^{\frac{1}{2}},$$

这里 $x=\sum_{i=1}^{n}\alpha_i e_i$。组合上面两个不等式有

$$\frac{\mu_1}{\lambda_2}\|x\|_2 \leqslant \|x\|_1 \leqslant \frac{\lambda_1}{\mu_2}\|x\|_2,$$

因此 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价。

【推论 2】 设 X 是 n 维赋范线性空间,则 X 是 Banach 空间。

证 设 $\{x^{(k)}\}$ 是 X 中 Cauchy 点列,则 $\{x^{(k)}\}$ 在基 $\{e_1,e_2,\cdots,e_n\}$ 下展开的系数对应 $\mathbf{R}^n$ 空间中的点列 $\{\xi^{(k)}\}$。由定理 2 不等式(2.10)的左端知 $\{\xi^{(k)}\}$ 是 $\mathbf{R}^n$ 空间的 Cauchy 列,而 $\mathbf{R}^n$ 完备,所以存在 $\xi\in\mathbf{R}^n$ 使得 $\xi^{(k)}\to\xi(k\to\infty)$。$\xi$ 在基下对应于 X 中的元素为 x,那么再由式(2.10)式的右端不等式得 $x^{(k)}\to x(k\to\infty)$。因此 X 是完备的,即为 Banach 空间。证毕。

当 X 为有限维空间时,X 中有界集是列紧集,有界闭集是紧集。这个性质在任何无限维空间中均不成立。

【定理 3】 设 X 是任一无限维赋范线性空间,则单位球 $\bar{B}_1(\theta)=\{x:\|x\|\leqslant 1\}$ 不是列紧集。

证 任取 $y_1\in\bar{B}_1(\theta)$, $\|y_1\|=1$,记 $Y_1=\operatorname{span}\{y_1\}$,则 Y_1 是一维闭子空间。由 Riesz 定理($\varepsilon=\frac{1}{2}$),存在 $y_2\in\bar{B}_1(\theta)$, $\|y_2\|=1$ 使

$$\|y_2-x\|\geqslant\frac{1}{2}\quad(x\in Y_1),$$

更有 $\|y_2-y_1\|\geqslant\frac{1}{2}$。记 $Y_2=\operatorname{span}\{y_1,y_2\}$,则 Y_2 是二维闭子空间,再由 Riesz 定理,存在 $y_3\in\bar{B}_1(\theta)$, $\|y_3\|=1$ 使

$$\| y_3 - x \| \geqslant \frac{1}{2} \quad (x \in Y_2),$$

更有 $\| y_3 - y_2 \| \geqslant \frac{1}{2}$，$\| y_3 - y_1 \| \geqslant \frac{1}{2}$。

由于 X 是无限维空间，这个过程可一直进行下去，得到 $\bar{B}_1(\theta)$ 中点列 $\{y_k\}$ 满足 $\| y_k - y_{k'} \| \geqslant \frac{1}{2}(k \neq k')$。从而 $\{y_k\}$ 无任何收敛子列。因此 $\bar{B}_1(\theta)$ 不是列紧集。证毕。

习题 2.6

1. 设 X 是赋范线性空间，$\{x_n\}$ 是 X 中 Cauchy 点列，证明：$\{x_n\}$ 是有界集。

2. 设 X 是赋范线性空间，证明：对任意 $x, y \in X$，有 $| \| x \| - \| y \| | \leqslant \| x - y \|$。

3. 在连续函数空间 $C[0,1]$ 上定义两个范数

$$\| f \|_1 = \left(\int_0^1 | f(x) |^2 \mathrm{d}x \right)^{\frac{1}{2}}, \quad \| f \|_2 = \left(\int_0^1 | f(x) |^2 (1 + x) \mathrm{d}x \right)^{\frac{1}{2}}$$

证明 $\| \cdot \|_1$ 与 $\| \cdot \|_2$ 等价。

4. 设 X 是赋范线性空间，Y 是 X 的子空间，证明：Y 的闭包 $\overline{Y}$ 仍是 X 的子空间。

5. 证明：赋范线性空间 X 是 Banach 空间的充要条件是：对任意 $\{x_n\} \subset X$，若 $\sum_{i=1}^{\infty} \| x_i \|$ 收敛，则 $\sum_{i=1}^{\infty} x_i$ 在 X 中收敛。

6. 设 X 是赋范线性空间，$x_i \in X$，且 $\sum_{i=1}^{\infty} x_i$ 在 X 中收敛，证明：$x_i \to \theta$。

7. 设 X 是赋范线性空间，$\{e_1, e_2, \cdots, e_n\} \subset X$，对任意 $x \in X$，证

明必存在数组 $\alpha_1,\alpha_2,\cdots,\alpha_n\in F$,使成立

$$\left\|x-\sum_{i=1}^{n}\alpha_i e_i\right\|=\min\left\{\left\|x-\sum_{i=1}^{n}t_i e_i\right\|:t_1,t_2,\cdots,t_n\in F\right\}。$$

8. 设 X 是赋范线性空间,Y 是 X 的有限维子空间,证明必存在 $x_0\in X$,且 $\|x_0\|=1$,使成立

$$\rho(x_0,Y)\geqslant 1。$$

9. 设 X 是赋范线性空间,若 X 中闭单位球 $\overline{B}_1(\theta)$ 是紧集时,证明 X 必是有限维的。

第3章 有界线性算子与有界线性泛函

本章所讲授的内容是泛函分析最重要也是最基本的。涉及到泛函分析的三大基本定理，即共鸣定理，逆算子定理以及 Hahn - Banach 定理。

3.1 有界线性算子

【定义1】 设 X 与 Y 是两个赋范线性空间，$T: X \to Y$ 是线性算子。称 T 是有界的，是指 T 将 X 中有界集映成 Y 中有界集。称 T 在 x 点是连续的，是指若 $x_n \to x$，则 $Tx_n \to Tx$，又称 T 在 X 上连续，是指 T 在 X 的每一点连续。

【性质1】 若线性算子 T 在 x_0 点连续，则 T 连续。

证 设 $x \in X$，$x_n \to x$，来证 $Tx_n \to Tx$。由于 T 在 x_0 点连续，而 $x_n - x + x_0 \to x_0$，于是有

$$T(x_n - x + x_0) = Tx_n - Tx + Tx_0 \to Tx_0$$

因此 $Tx_n - Tx \to \theta$（Y 中零元），所以 $Tx_n \to Tx$。证毕。

性质1说明对线性算子在一点连续可推出在空间的每点连续。特别，线性算子的连续性可由零元的连续性来刻画，即线性算子 T 连续等价于若 $x_n \to \theta$（X 中零元），则 $Tx_n \to \theta$（Y 中零元）。

【性质 2】 $T:X\to Y$ 是线性算子,则 T 有界的充分必要条件是存在常数 $M>0$,使不等式

$$\|Tx\|\leqslant M\|x\| \quad (x\in X) \tag{3.1}$$

成立。

证 必要性。因 T 有界,所以 T 将 X 的单位球 $\bar{B}_1(\theta)=\{x:\|x\|\leqslant 1\}$ 映成 Y 中有界集,即象集 $T\bar{B}_1(\theta)$ 是 Y 中有界集。记 $M=\sup\{\|Tx\|:x\in\bar{B}_1(\theta)\}$。此时,对每个 $x\in X, x\neq\theta$, $\dfrac{x}{\|x\|}\in\bar{B}_1(\theta)$,由 M 的定义有

$$\left\|T\left(\frac{x}{\|x\|}\right)\right\|\leqslant M,$$

即 $\|Tx\|\leqslant M\|x\|$。而当 $x=\theta$ 时,不等式(3.1) 变为等式,故式(3.1) 成立。

充分性 设 A 是 X 中有界集,即存在常数 $M_1>0$,使得 $\sup\{\|a\|:a\in A\}\leqslant M_1$。对每个 $y\in TA$,存在 $a\in A$,使 $y=Ta$,由不等式(3.1) 有

$$\|y\|=\|Ta\|\leqslant M\|a\|\leqslant M\cdot M_1,$$

故 $\sup\{\|y\|:y=Ta, a\in A\}\leqslant M\cdot M_1$,即 TA 是 Y 中有界集。

【定理 1】 线性算子 $T:X\to Y$ 连续的充要条件是 T 有界。

证 必要性。若 T 是无界的,则不等式(3.1) 不能成立。于是对任何自然数 n,存在 $x_n\in X$,使得 $\|Tx_n\|>n\|x_n\|$。令 $y_n=\dfrac{x_n}{n\|x_n\|}$,则 $\|Ty_n\|>1$,且

$$\|y_n\|=\frac{1}{n}\to 0(n\to\infty)。$$

可见,$y_n\to\theta$,但 $Ty_n\not\to\theta$,这与 T 连续矛盾。

充分性。由不等式(3.1),若 $x_n\to\theta$,则

$$\| Tx_n \| \leqslant M \| x_n \| \to 0 (n \to \infty),$$

故 $Tx_n \to \theta$，即 T 连续。证毕。

对于线性算子，连续性与有界性是两个等价的概念。

用 $L(X,Y)$ 表示所有从 X 到 Y 的有界线性算子组成的集合。

【定理 2】　$T \in L(X,Y)$，记

(1) $\alpha_1 = \sup\{\| Tx \| : \| x \| \leqslant 1, x \in X\}$，

(2) $\alpha_2 = \sup\{\| Tx \| : \| x \| = 1, x \in X\}$，

(3) $\alpha_3 = \sup\left\{\frac{\| Tx \|}{\| x \|} : x \neq \theta, x \in X\right\}$，

(4) $\alpha_4 = \inf\{M : \| Tx \| \leqslant M \| x \|, x \in X\}$。

那么，$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$。

证　显然 $\alpha_1 \geqslant \alpha_2 \geqslant \alpha_3$。对 $\forall \epsilon > 0$，由上确界的定义，有 $x_0 \in X$，$\| x_0 \| \leqslant 1$，满足 $\alpha_1 - \epsilon < \| Tx_0 \|$，于是有

$$\alpha_1 - \epsilon < \| Tx_0 \| \leqslant \frac{\| Tx_0 \|}{\| x_0 \|} \leqslant \alpha_3,$$

由 ϵ 的任意性得 $\alpha_1 \leqslant \alpha_3$，故 $\alpha_1 = \alpha_2 = \alpha_3$。对任意满足 $\| Tx \| \leqslant M \| x \|$ 的常数 M，当 $x \neq \theta$ 时，有

$$\frac{\| Tx \|}{\| x \|} \leqslant M,$$

因此 $\alpha_3 \leqslant M$，再由 α_4 的定义 $\alpha_3 \leqslant \alpha_4$。另一方面，据 α_3 的定义，对任意 $x \neq \theta$，有

$$\frac{\| Tx \|}{\| x \|} \leqslant \alpha_3,$$

即 $\| Tx \| \leqslant \alpha_3 \| x \|$，所以 $\alpha_4 \leqslant \alpha_3$。于是 $\alpha_3 = \alpha_4$。证毕。

【定义 2】　$T \in L(X,Y)$，定义 T 的范数为

$$\| T \| = \alpha_1 (\alpha_2 = \alpha_3 = \alpha_4)。$$

对于集合 $L(X,Y)$ 引进线性运算加法与数乘，使之成为一个线

性空间。设 $T_1, T_2 \in L(X,Y)$，定义 $(T_1+T_2): X \to Y$ 为 $(T_1+T_2)(x) = T_1x + T_2x$；$\alpha \in F$，$(\alpha T_1): X \to Y$ 为 $(\alpha T_1)(x) = \alpha T_1 x$。显然，$T_1 + T_2$ 及 αT_1 均是线性算子。又因

$$\|(T_1+T_2)x\| \leqslant \|T_1x\| + \|T_2x\| \leqslant (\|T_1\| + \|T_2\|)\|x\|,$$

$$\|\alpha(T_1)(x)\| \leqslant |\alpha|\,\|T_1\|\,\|x\|,$$

故 T_1+T_2，αT_1 也是有界线性算子，即 $T_1+T_2 \in L(X,Y)$，$\alpha T_1 \in L(X,Y)$。容易验证，有界线性算子在上述加法和数乘下构成一个线性空间。这个线性空间的零元素是零算子，记为 Θ。即 $\Theta: X \to Y$ 定义为 $\Theta x = \theta$（Y 中零元素），对一切 $x \in X$（在不致引起混淆时，可记 Θ 为 θ）。

$L(X,Y)$ 在定义 2 的意义下形成一个赋范线性空间。为此，需说明算子范数是 $L(X,Y)$ 上定义的一个范数。

(1) $\|T\| = 0$，那么对一切 $x \in X$，$\|x\| \leqslant 1$，有 $Tx = \theta$。对 $\|x\| > 0$，$T\left(\frac{x}{\|x\|}\right) = \theta$，因而有 $Tx = \theta$，故 $T = \Theta$。反之，$T = \Theta$，显然，$\|T\| = 0$。

(2) $$\begin{aligned}\|\alpha T\| &= \sup\{\|\alpha T(x)\| : \|x\| \leqslant 1, x \in X\} = \\ &|\alpha| \sup\{\|Tx\| : \|x\| \leqslant 1, x \in X\} = \\ &|\alpha|\,\|T\|。\end{aligned}$$

(3) $$\begin{aligned}\|T_1+T_2\| &= \sup\{\|(T_1+T_2)(x)\| : \|x\| \leqslant 1, x \in X\} \leqslant \\ &\sup\{\|T_1x\| : x \in X, \|x\| \leqslant 1\} + \\ &\sup\{\|T_2x\| : x \in X, \|x\| \leqslant 1\} = \\ &\|T_1\| + \|T_2\|。\end{aligned}$$

【定理 3】　当 Y 是 Banach 空间时，$L(X,Y)$ 也是 Banach 空间。

证　设 $\{T_n\}$ 是 $L(X,Y)$ 的一个 Cauchy 列，则对 $\forall \epsilon > 0$，存在自然数 N，当 $n, m \geqslant N$ 时

$$\| T_n - T_m \| < \varepsilon。\tag{3.2}$$

对每个 $x \in X$ 且 $x \neq \theta$ 时，由于

$$\| T_n x - T_m x \| \leqslant \| T_n - T_m \| \| x \|。$$

对 $\dfrac{\varepsilon}{\| x \|}$ 存在自然数 $N_1(\varepsilon, x)$，当 $n, m \geqslant N_1(\varepsilon, x)$ 时，

$$\| T_n - T_m \| < \frac{\varepsilon}{\| x \|},$$

所以 $\| T_n x - T_m x \| < \varepsilon$（当 $x = \theta$ 时亦成立）。这意味着，$\{T_n x\}$ 是 Y 中 Cauchy 列，从而 $\lim\limits_{n \to \infty} T_n x$ 存在。于是定义 $T: X \to Y$ 为

$$Tx = \lim_{n \to \infty} T_n x。$$

由于 T_n 是线性算子，那么 T 也是线性算子。又 $\{T_n\}$ 是 Cauchy 列，因而 $\| T_n \|$ 有界，即存在常数 $M > 0$，使 $\| T_n \| \leqslant M (n = 1, 2, \cdots)$。从而

$$\| Tx \| = \lim_{n \to \infty} \| T_n x \| \leqslant M \| x \|,$$

$T \in L(X, Y)$。

据式(3.2)，对 $x \in X$，$\| x \| \leqslant 1$，有 $\| T_n x - T_m x \| \leqslant \| T_n - T_m \| \| x \| < \varepsilon$，令 $m \to \infty$，则得 $\| T_n x - Tx \| \leqslant \varepsilon$ 对 $n \geqslant N$ 成立，这就是 $\| T_n - T \| \leqslant \varepsilon$，故 $\{T_n\}$ 收敛于 T。证毕。

例 1　X 是赋范线性空间，且 $X \neq \{\theta\}$，恒等算子 $I: X \to X$，为 $Ix = x (x \in X)$，则 $I \in L(X, Y)$，且 $\| I \| = 1$。

例 2　已知实矩阵 $\mathbf{A} = (a_{ij})_{n \times m}$，定义 $T: \mathbf{R}^m \to \mathbf{R}^n$ 为 $Tx = Ax$，则 $T \in L(\mathbf{R}^m, \mathbf{R}^n)$，且 $\| T \| \leqslant \left(\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} a_{ij}^2 \right)^{\frac{1}{2}}$。

证　$$\| Tx \| = \| Ax \| = \left[\sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} x_j \right)^2 \right]^{\frac{1}{2}} \leqslant$$

$$\left[\sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij}^2 \right) \left(\sum_{j=1}^{m} x_j^2 \right) \right]^{\frac{1}{2}} =$$

$$(\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^{2})^{\frac{1}{2}}\|x\|,$$

故 $\|T\| \leqslant (\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}^{2})^{\frac{1}{2}}$。

例2及第二章2.5节的例8表明,任何两个有限维赋范性空间之间的线性算子一定是有界线性算子。

例3 设 $K(s,t)$ 在 $[a,b]\times[a,b]$ 上连续,定义算子 $T:C[a,b]\to C[a,b]$ 为

$$Tx(s)=\int_a^b K(s,t)x(t)\mathrm{d}t,$$

则 $T\in L(C[a,b],C[a,b])$,且 $\|T\| \leqslant \max\{\int_a^b |K(s,t)|\mathrm{d}t: a\leqslant s\leqslant b\}$。

证 由于

$$\|Tx(s)\| = \max_{a\leqslant s\leqslant b}\left|\int_a^b K(s,t)x(t)\mathrm{d}t\right| \leqslant$$

$$\max_{a\leqslant s\leqslant b}\left|\int_a^b K(s,t)\right|\mathrm{d}t\cdot\max_{a\leqslant T\leqslant b}|x(t)| =$$

$$\max\{\int_a^b |K(s,t)|\mathrm{d}t: a\leqslant s\leqslant b\}\|x\|,$$

故结论成立。事实上,还可以进一步证明

$$\|T\| = \max\{\int_a^b |K(s,t)|\mathrm{d}t: a\leqslant s\leqslant b\}。$$

由于要用到实分析的知识,这里略去。

例4 $X=C^1[0,1]$ 即 $[0,1]$ 上连续可微函数全体,$Y=C[0,1]$。X 是 Y 的子空间,在 Y 的范数意义下也是赋范线性空间。定义微分算子 $T:X\to Y$ 为 $Tx(s)=\frac{\mathrm{d}}{\mathrm{d}s}x(s)=x'(s)$,则 T 是线性算子但不是有界的。

证　微分具有线性，因此 T 是线性算子。取 $x_n(s)=\sin(n\pi s)$，则 $\|x_n\|=\max\limits_{s\in[0,1]}|\sin(n\pi s)|=1$，但 $\|Tx_n\|=\max\limits_{s\in[0,1]}|n\pi\cos(n\pi s)|=n\pi\to\infty(n\to\infty)$，故 T 不是有界线性算子。

习题 3.1

1. 设 $T\in L(X,Y)$，证明：$\ker(T)=\{x:Tx=\theta\}$ 是 x 的闭子空间。

2. 设 $T\in L(X,Y)$，$S\in L(Y,Z)$，证明：复合算子 $ST:X\to Z$ 满足 $\|ST\|\leqslant\|S\|\|T\|$。

3. $X=C[0,1]$，定义 $T:X\to X$ 为

$$(Tx)(t)=t\int_0^1 x(s)\mathrm{d}s\qquad\forall t\in[0,1]$$

及 $S:X\to X$ 为 $(Sx)(t)=tx(t)\quad\forall t\in[0,1]$，

(1) 问 T 与 S 可交换吗？(即 $ST=TS$ 是否成立？)

(2) 求 $\|T\|$，$\|S\|$，$\|TS\|$ 及 $\|ST\|$。

4. 设 X 为所有有界数列组成的线性空间，范数为 $\|x\|=\sup\limits_{i\geqslant1}|a_i|(x=\{a_i\})$。给定无穷矩阵

$$T=(t_{ij})$$

满足 $\sup\limits_i\sum\limits_{j=1}^{\infty}|t_{ij}|<+\infty$，定义算子 $T:X\to X$ 为

$$Tx=y,$$

其中 $x=\{a_i\}$，$y=\{b_i\}$ 且

$$b_i=\sum_{j=1}^{\infty}t_{ij}a_j。$$

证明：$T\in L(X,X)$ 且 $\|T\|=\sup\limits_i\sum\limits_{j=1}^{\infty}|t_{ij}|$。

5. 设 $X = \mathbf{R}^n, Y = \mathbf{R}^m$，在 X, Y 上定义范数 $\|x\|_1 = \sum_{i=1}^{n} |x_i|$，$x = (x_1, x_2, \cdots, x_n) \in \mathbf{R}^n$；$\|y\|_1 = \sum_{i=1}^{m} |y_i|, y = (y_1, y_2, \cdots, y_m) \in \mathbf{R}^m$。矩阵 $\mathbf{A} = (a_{ij})_{m\times n}$，定义算子 $T: X \to Y$ 为

$$y = Tx = \mathbf{A}x。$$

证明：$\|T\| = \max\limits_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$。

6. 设 $f: \mathbf{R} \to \mathbf{R}$ 连续且可加，即对任 $x_1, x_2 \in \mathbf{R}$，有 $f(x_1 + x_2) = f(x_1) + f(x_2)$。证明 f 必为 $f(x) = \lambda x, \forall x \in \mathbf{R}$，其中 $\lambda \in \mathbf{R}$ 为常数。

7. 设 X, Y 都是 Banach 空间，$T \in L(X, Y)$ 且是满映射，证明对 X 中任意稠密子集 E，成立 $\overline{T(E)} = Y$。

8. 设 X 是 Banach 空间，$T \in L(X, X)$，且 $\|T\| < 1$，定义 $T^n = T \cdot T \cdots T$ 为 T 的 n 次复合，$T^0 = I$ 为单位算子，证明算子级数 $\sum_{n=0}^{\infty} T^n$ 在 $L(X, X)$ 中收敛，且 $T^n \to \theta$（零算子）。

3.2 共鸣定理

3.2.1 算子列的收敛

【定义 1】 设 $T, T_n \in L(X, Y)$，称 $\{T_n\}$ 一致收敛于 T，是指 $\|T_n - T\| \to 0 (n \to \infty)$，即在算子范数意义下收敛。称 $\{T_n\}$ 强收敛于 T，是指对每个 $x \in X$，$\|T_n x - Tx\| \to 0 (n \to \infty)$。

根据定义及 $\|T_n x - Tx\| \leq \|T_n - T\| \|x\| \to 0 (n \to \infty)$，知算子列 $\{T_n\}$ 一致收敛于 T 必强收敛于 T。反之不成立。

例 1　设 $X = Y = l^1 = \{x = \{a_i\}: \sum_{i=1}^{\infty} |a_i| < +\infty\}$，范数为 $\|x\| = \sum_{i=1}^{\infty} |a_i|$。定义算子列 T_n 为：

$$T_n x = (a_1, a_2, \cdots, a_{n-1}, 0, \cdots), x = \{a_i\}$$

及 $I: X \to X$ 为恒等算子，则 $\{T_n\}$ 强收敛于 I，但不一致收敛于 I。

证　对每个 $x \in X$，由于

$$\|T_n x - Ix\| = \sum_{i=n}^{\infty} |a_i| \to 0 (n \to \infty),$$

所以 $\{T_n\}$ 强收敛于 I。但取 $e_n = (\underbrace{0,0,\cdots,0}_{n-1个}, 1, 0, \cdots)$ 即第 n 个分量是 1，其余皆为 0，则 $\|e_n\| = 1$，于是

$$\|T_n - I\| \geqslant \|(T_n - I)e_n\| = 1,$$

故 $\{T_n\}$ 并不一致收敛于 I。

容易验证，算子列 $\{T_n\}$ 一致收敛于 T 的充要条件是 $\{T_n\}$ 在 X 的单位球上一致强收敛于 T。

3.2.2　共鸣定理及其应用

【定义 2】　设 X 是一个度量空间，$A \subset X$，称 A 是 X 中的稀疏集，是指 A 在 X 中的任何一个非空开集中均不稠密。又称 X 是第一纲的，是指 X 可表示成至多可列个稀疏集的并，不是第一纲的度量空间称为是第二纲的。

例 2　X = 有理数集，定义度量 $\rho(r_1, r_2) = |r_1 - r_2| \in X$，则 X 是第一纲的，因为 $X = \bigcup_{n=1}^{\infty} \{r_n\}$，而单点集 $\{r_n\}$ 是 X 的稀疏集。

下面是关于完备度量空间的一个重要定理，即 Baire 纲定理。它是证明共鸣定理的关键。

【定理 1】　设 X 是完备的度量空间，则 X 是第二纲的。

证 用反证法。若存在一列稀疏集$\{A_n\}$使$X=\bigcup\limits_{n=1}^{\infty}A_n$。任取一个闭球$\bar{B}_{r_0}(x_0)=\{x:\rho(x,x_0)\leqslant r_0\}$，由于$A_1$在开球$B_{r_0}(x_0)$中不稠密，从而可取一个闭球$\bar{B}_{r_1}(x_1)(0<r_1<1)$满足$A_1\cap\bar{B}_{r_1}(x_1)=\varnothing$，$\bar{B}_{r_1}(x_1)\subset\bar{B}_{r_0}(x_0)$。又$A_2$在开球$B_{r_1}(x_1)$中不稠密，同理，取闭球$\bar{B}_{r_2}(x_2)(0<r_2<\frac{1}{2})$满足$A_2\cap\bar{B}_{r_2}(x_2)=\varnothing$，$\bar{B}_{r_2}(x_2)\subset\bar{B}_{r_1}(x_1)$。按上述过程一直进行下去，可得出闭球列$\{\bar{B}_{r_n}(x_n)\}$满足

(1)$\bar{B}_{r_0}(x_0)\supset\bar{B}_{r_1}(x_1)\supset\bar{B}_{r_2}(x_2)\supset\cdots$；

(2)$\bar{B}_{r_n}(x_n)\cap A_n=\varnothing(n=1,2,\cdots)$；

(3)$0<r_n<\frac{1}{n}$。

由(3)知$\bar{B}_{r_n}(x_n)$的直径$d(\bar{B}_{r_n}(x_n))\leqslant\frac{2}{n}\to0(n\to\infty)$，据第二章2.2节中定理4，存在$x\in X$，且$\bigcap\limits_{n=1}^{\infty}\bar{B}_{r_n}(x_n)=\{x\}$。但从(2)中又有$x\notin\bigcup\limits_{n=1}^{\infty}A_n=X$，矛盾，故$X$是第二纲的。证毕。

应用上述定理来证明共鸣定理。

【定理2】 设X是Banach空间，Y是赋范线性空间。算子族$\{T_\lambda:\lambda\in\Lambda\}\subset L(X,Y)$。若对任意$x\in X$，满足$\sup\{\|T_\lambda x\|:\lambda\leqslant\Lambda\}<+\infty$，那么$\sup\limits_{\lambda\in\Lambda}\|T_\lambda\|<+\infty$。

证 定义X上的泛函$P(x)$为

$$P(x)=\sup\{\|T_\lambda x\|:\lambda\leqslant\Lambda\},$$

则$P:X\to[0,+\infty)$，且很容易验证$P(x)$满足

$$P(x+x')\leqslant P(x)+P(x'),P(\alpha x)=|\alpha|P(x)\quad(\alpha\in F),$$

记$A_n=\{x\in X:P(x)\leqslant n\}(n=1,2\cdots)$，则$X=\bigcup\limits_{n=1}^{\infty}A_n$。首先来证$A_n$

是闭集。设 $x_k \in A_n, x_k \to x$。对每个 $\lambda \in \Lambda$，因 T_λ 是连续的，所以 $\| T_\lambda x_k - T_\lambda x \| \to 0(k \to \infty)$，更有 $\| T_\lambda x_k \| \to \| T_\lambda x \|$。又 $\| T_\lambda x_k \| \leqslant P(x_k) \leqslant n$，故 $\| T_\lambda x \| \leqslant n$，即 $P(x) \leqslant n, x \in A_n$。Banach 空间 X 是完备的度量空间，由定理 1，必存在自然数 n_0，使 A_{n_0} 不是稀疏集。从而存在开球 $B_{r_0}(x_0)(r_0 > 0)$ 使 A_{n_0} 在 $\bar{B}_{r_0}(x_0)$ 中稠密，A_{n_0} 闭，所以 $A_{n_0} \supset \bar{B}_{r_0}(x_0)$。对任一 $x \in \bar{B}_1(\theta) = \{x : \| x \| \leqslant 1\}$，注意到 $x_0 + r_0 x, x_0 - r_0 x \in \bar{B}_{r_0}(x_0)$，则

$$P(x_0 + r_0 x) \leqslant n_0, P(x_0 - r_0 x) \leqslant n_0,$$

$$\begin{aligned} P(2r_0 x) \leqslant P(x_0 + r_0 x) + P(-x_0 + r_0 x) = \\ P(x_0 + r_0 x) + P(x_0 - r_0 x) \leqslant 2n_0, \end{aligned}$$

所以 $P(x) \leqslant \dfrac{n_0}{r_0}$。对每个 $\lambda \in \Lambda$，$\| T_\lambda x \| \leqslant P(x) \leqslant \dfrac{n_0}{r_0}$，即

$$\| T_\lambda \| = \sup\{ \| T_\lambda x \| : x \in \bar{B}_1(\theta)\} \leqslant \frac{n_0}{r_0}。$$

进一步有 $\sup\limits_{\lambda \in \Lambda} \| T_\lambda \| \leqslant \dfrac{n_0}{r_0} < +\infty$。证毕。

上述共鸣定理说明对每个 $x \in X$，$\{ \| T_\lambda x \| : \lambda \in \Lambda \}$ 有界，则 $\{ \| T_\lambda \| : \lambda \in \Lambda \}$ 有界。这蕴含算子族每点有界，可推出在单位球上一致有界。因此，共鸣定理又称一致有界原理。

共鸣定理的一个直接推论就是下面著名的 Banach – Steinhaus 定理。

【定理 3】 设 X 是 Banach 空间，Y 是赋范线性空间。$\{T_n\} \subset \mathrm{L}(X, Y)$，若满足对每个 $x \in X$，$\lim\limits_{n\to\infty} T_n x$ 在 Y 中存在，定义线性算子 $T: X \to Y$ 为 $Tx = \lim\limits_{n\to\infty} T_n x$，则 $T \in \mathrm{L}(X, Y)$。

证 由 $\lim\limits_{n\to\infty} T_n x$ 在 Y 中存在，知 $\sup\limits_n \| T_n x \| < +\infty$。据定理 2，存

在常数 $M > 0$,使 $\sup_n \| T_n \| \leqslant M$,故

$$\| Tx \| = \lim_{n\to\infty} \| T_n x \| \leqslant (\sup_n \| T_n \|) \| x \| \leqslant M \| x \|,$$

即 $T \in \mathrm{L}(X, Y)$。证毕。

下面列举几个共鸣定理的应用示例。

例 3 (Fourier 级数的发散问题)存在以 2π 为周期的连续函数,其 Fourier 级数在给定点发散。

证 用 $C_{2\pi}$ 表示定义在 $(-\infty, +\infty)$ 上以 2π 为周期的连续函数全体,赋予范数

$$\| x \| = \max\{ | x(t) | : t \in [0, 2\pi] \},$$

那么 $C_{2\pi}$ 是一个 Banach 空间。对每个 $x \in C_{2\pi}$ 其前 $n+1$ 项 Fourier 级数的部分和为

$$S_n(x, t) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt) = \frac{1}{\pi} \int_{-\pi}^{\pi} k_n(s, t) x(s) \mathrm{d}s,$$

这里 $k_n(s, t) = \dfrac{\sin[(n + \frac{1}{2})(s - t)]}{2\sin[\frac{1}{2}(s - t)]}$。令 $t = 0$,即

$$S_n(x, 0) = \frac{1}{\pi} \int_{-\pi}^{\pi} k_n(s, 0) x(s) \mathrm{d}s。$$

$S_n(x, 0)$ 是 $C_{2\pi}$ 到 R 的有界线性泛函,且可计算其范数为

$$\| S_n \| = \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \frac{\sin(n + \frac{1}{2}) s}{2\sin \frac{1}{2} s} \right| \mathrm{d}s。$$

注意到

$$\| S_n \| = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})s}{\sin\frac{s}{2}}\right|ds =$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})s}{\frac{s}{2}}\right|\cdot\left|\frac{\frac{s}{2}}{\sin\frac{s}{2}}\right|ds \geqslant$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(2n+1)\frac{s}{2}}{\frac{s}{2}}\right|ds =$$

$$\frac{1}{\pi}\int_{0}^{\frac{2n+1}{2}\pi}\left|\frac{\sin u}{u}\right|du \to \infty\,(n\to\infty),$$

所以$\sup\limits_n = \| S_n \| = +\infty$。从而由共鸣定理，必存在某个周期$2\pi$的连续函数 $x_0 \in C_{2\pi}$，使极限 $\lim\limits_{n\to\infty} S_n(x,0)$ 不存在，这意味着 $x_0(t)$ 的 Fourier 级数在 $t = 0$点发散。同理，对每一固定点 t_0，也必存在 $x_{t_0} \in C_{2\pi}$。其 Fourier 级数在 $t = t_0$ 发散。

例4 （Lagrange插值公式的发散性定理）　给定区间[0,1]内插入点，$(t_k^{(n)})(1 \leqslant k \leqslant n, n = 1,2,\cdots)$，构成无穷三角矩阵 $\boldsymbol{H}$ 为

$$\begin{bmatrix} t_1^{(1)} & 0 & 0 & 0 & \cdots & \cdots \\ t_1^{(2)} & t_2^{(2)} & 0 & 0 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ t_1^{(n)} & t_2^{(n)} & \cdots & t_n^{(n)} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix},$$

那么必存在 $x(t) \in C[0,1]$，使其与插入点相应的 n- 次插值 Lagrange 多项式

$$T(x)(t) = \sum_{k=1}^{n} x(t_k^{(n)})C_k^{(n)}(t),$$

其中 $C_k^{(n)}(t) = \dfrac{(t-t_1^{(n)})\cdots(t-t_{k-1}^{(n)})(t-t_{k+1}^{(n)})\cdots(t-t_n^{(n)})}{(t_k^{(n)}-t_1^{(n)})\cdots(t_k^{(n)}-t_{k-1}^{(n)})(t_k^{(n)}-t_{k+1}^{(n)})\cdots(t_k^{(n)}-t_n^{(n)})}$，当 $n\to\infty$ 时，不一致收敛于 $x(t)$。

证 在 C[0,1] 上定义算子序列 T_n：C[0,1] → C[0,1] 为

$$[T_n(x)](t) = \sum_{k=1}^{n} x(t_k^{(n)})C_k^{(n)}(t)。$$

通过计算得出

$$\| T_n \| = \max_{t\in[0,1]} \sum_{k=1}^{n} | C_k^{(n)}(t) |, n = 1,2,\cdots。$$

从而 $\{T_n\}$ 是有界线性算子序列。在函数逼近论中已知道

$$\| T_n \| > \frac{\lg n}{8\sqrt{\pi}} \quad (n = 1,2,\cdots),$$

因此 $\sup_n \| T_n \| = +\infty$。于是由共鸣定理必存在 $x_0 \in$ C[0,1]，使 $T_n(x_0)$ 不收敛于 x_0，即 $T_n(x_0)(t)$ 不一致收敛于 $x_0(t)$。

习题 3.2

1. 设 X 是 Banach 空间，Y 是赋范线性空间。$T_n \in$ L$(X,Y)(n = 1,2,\cdots)$。若 $\sup_n \| T_n \| = +\infty$。证明：存在 $x_0 \in X$，使得 $\sup_n \| T_n x_0 \| = +\infty$。

2. 设 X 为全体多项式构成的集合，按通常的函数的加法数乘运算成为一个线性空间。又对任意 $x(t) = a_0 + a_1 t + \cdots + a_n t^n \in X$，定义

$$\| x \| = \max\{ | a_0 |, | a_1 |, \cdots, | a_n | \}$$

(1) 证明 X 是一个赋范线性空间；

(2) 证明 X 不完备;

(3) 取 $Y = R$,定义算子列 $T_k: X \to R$ 为

$$T_k x = a_0 + a_1 + \cdots + a_k \quad (k = 1,2,\cdots),$$

证明 T_k 是有界线性算子,且对任意 $x \in X$,成立 $\sup\limits_k \| T_k x \| < \infty$,但 $\sup\limits_k \| T_k \| = \infty$。

3. 给定数列 $\{\beta_n\}$,若满足对任意收敛数列 $\{\alpha_n\}$,使级数 $\sum\limits_{n=1}^{\infty} \alpha_n \beta_n$ 收敛,证明级数 $\sum\limits_{n=1}^{\infty} |\beta_n| < +\infty$。

4. 给定数列 α_n,若对任意 $x = (t_1, t_2, \cdots) \in l^1$,使级数 $\sum\limits_{n=1}^{\infty} \alpha_n t_n$ 收敛,证明 $\sup\limits_n |\alpha_n| < \infty$。

5. 设 X, Y 都是 Banach 空间,$T_n \in L(X,Y)$,若 $\sup\limits_n \| T_n \| < \infty$,并且存在 X 中稠密子集 E,使 $\forall x \in E$,$T_n x$ 是 Cauchy 列,证明存在 $T \in L(X,Y)$,有 $T_n x \to Tx (x \in X)$。

6. 设 X 是度量空间,$A \subset X$,证明 A 是稀疏集 $\Leftrightarrow X \setminus \bar{A}$ 是稠密集。

3.3 Hahn – Banach 定理

X 是实赋范线性空间,$f: X \to \mathbf{R}$ 是线性泛函。由于泛函是算子的特例,因此,若 f 是有界的,则 f 的范数定义为

$$\| f \| = \sup\{ | f(x) | : x \in \bar{B}_1(\theta)\},$$

$\bar{B}_1(\theta) = \{x : \| x \| \leqslant 1\}$。所有 X 上有界线性泛函全体为 $L(X, \mathbf{R})$,特别记 $X^* = L(X, \mathbf{R})$。由于 $\mathbf{R}$ 是完备的,由本章 1.3 定理 3,X^* 是一

个 Banach 空间。

本节的目的是研究线性泛函的延拓问题。

【定义 1】 若 X 是一个实线性空间，称 $P:X\to\mathbf{R}$ 为次可加正齐次泛函，如果满足

(1) $P(x+y)\leqslant P(x)+P(y)$； (2) $P(\alpha x)=\alpha P(x)\quad(\alpha\geqslant 0)$。

例 1 当 X 是实赋线性空间时，令 $P(x)=\|x\|$，则 P 是次可加正齐次泛函。

【定理 1】 (Hahn－Banach) 设 X 是实线性空间，$P:X\to\mathbf{R}$ 是次可加正齐次泛函。$M\subset X$ 是一个子空间，f 是 M 上定义的一个实线性泛函且 $f(m)\leqslant P(m)(m\in M)$，那么存在 X 上的实线性泛函 F 满足

(1) $F(m)=f(m)\quad(m\in M)$；

(2) $F(x)\leqslant P(x)\quad(x\in X)$。

证 仅来证明一种特殊情况，当 X 比 M 仅多一维。此时 X 可表示为

$$X=\{x=m+\alpha x_0:m\in M,\alpha\in R\},$$

这里 $x_0\in X\setminus M$。定义 X 上的线性泛函 F 为

$$F(x)=F(m+\alpha x_0)=f(m)+\alpha C,$$

C 是一个待选择的常数。由于 f 是 M 上的线性泛函，那么 F 是 X 上的线性泛函，且显然满足(1)。为满足(2)，我们来确定常数 C，若满足(2)，则对一切 $m\in M$ 及 $\alpha\in\mathbf{R}$ 成立不等式

$$f(m)+\alpha C\leqslant P(m+\alpha x_0)。$$

这个不等式又等价于下面两个不等式：

$$(*)\begin{cases}f\left(\dfrac{m}{\alpha}\right)+C\leqslant P\left(x_0+\dfrac{m}{\alpha}\right),\alpha>0,\\[2ex] f\left(-\dfrac{m}{\alpha}\right)-C\leqslant P\left(-x_0-\dfrac{m}{\alpha}\right),\alpha<0。\end{cases}$$

因为对任何 $m',m''\in M$ 有

$$f(m')+f(m'')=f(m'+m'')\leqslant P(m'+m'')\leqslant p(m'-x_0)+P(m''+x_0),$$

即

$$(**)f(m')-P(m'-x_0)\leqslant P(m''+x_0)-f(m'')。$$

于是令

$$K_1=\sup\{f(m')-P(m'-x_0):m'\in M\},$$

$$K_2=\inf\{P(m''+x_0)-f(m''):m''\in M\}。$$

据$(**)$式，$-\infty<K_1\leqslant K_2\leqslant+\infty$，从而选取 $C\in[K_1,K_2]$，则这样的 C 满足$(*)$式，于是 $F(x)$ 在 X 上满足 $F(x)\leqslant P(x)$。证毕。

对于一般情况，由于涉及到超限归纳法，这里略去其证明。

下面我们把定理 1 应用到赋范空间的有界线性泛函的延拓上。

【定理 2】（Banach 保范延拓定理）　设 X 是实赋范线性空间，M 是 X 的子空间，f 是 M 上的有界线性泛函，则存在 X 上的有界线性泛函 F 满足

(1) $F(m)=f(m),m\in M$；

(2) $\|F\|=\|f\|_M$。

证　由于 f 是 M 上有界线性泛函，那么

$$|f(m)|\leqslant\|f\|_M\|m\|,\tag{3.3}$$

这里 $\|f\|_M=\sup\{|f(m)|:\|m\|\leqslant 1,m\in M\}$ 是 f 在 M 上的范数。令 $P(x)=\|f\|_M\|x\|$，则 P 是 X 上定义的次可加正齐次泛函，由式(3.3)对 $m\in M$，有 $f(m)\leqslant P(m)$。根据定理 1，存在 X 上泛函 F，满足结论(1)且 $F(x)\leqslant P(x)$。又 $F(-x)\leqslant P(-x)=\|f\|_M\|(-x)\|=\|f\|_M\|x\|$，所以

$$|F(x)|\leqslant\|f\|_M\|x\|。$$

可见 F 是 X 上有界线性泛函，且 $\|F\| \leqslant \|f\|_M$，又 F 是 f 的延拓（即满足(1)），所以

$$\|F\| = \sup\{|F(x)| : x \in X, \|x\| \leqslant 1\} \geqslant \sup\{|F(m)| : m \in M, \|m\| \leqslant 1\} = \sup\{|f(m)| : m \in M, \|m\| \leqslant 1\} = \|f\|_M,$$

即 $\|F\| = \|f\|_M$。证毕。

【推论 1】 设 X 是一个实赋范线性空间，M 是 X 的一个真闭子空间，$x_1 \in X \setminus M$。令

$$d = \inf\{\|x_1 - m\| : m \in M\},$$

则存在 X 上有界泛函 F 满足

$$\|F\| = \frac{1}{d} \text{ 且 } F(x) = \begin{cases} 1, \text{当 } x = x_1; \\ 0, \text{当 } x \in M。\end{cases}$$

证　首先来证 $d > 0$。若 $d = 0$，由下确界定义，存在 $m_n \in M$，满足 $\|x_1 - m_n\| \to d = 0$，即 $m_n \to x_1$，而 M 是闭的，所以 $x_1 \in M$，这与 $x_1 \in X \setminus M$ 矛盾。

记由 M 及 x_1 张成的子空间为 M_1，则 M_1 可表示成 $M_1 = \{m + \alpha x_1 : m \in M, \alpha \in \mathbf{R}\}$。在 M_1 上定义泛函 $f(m + \alpha x_1) = \alpha$，显然 f 是 M_1 上线性泛函，且 $f(m) = 0 (m \in M)$，$f(x_1) = 1$。下面来计算 f 在 M_1 上的范数。

对于 $\alpha \neq 0$，

$$|f(m + \alpha x_1)| = |\alpha| = \frac{|\alpha| \cdot \left\|\frac{m}{\alpha} + x_1\right\|}{\left\|\frac{m}{\alpha} + x_1\right\|} = \frac{\|m + \alpha x_1\|}{\left\|x_1 - \left(-\frac{m}{\alpha}\right)\right\|} \leqslant \frac{1}{d}\|m + \alpha x_1\|,$$

所以 $\|f\|_{M_1} \leqslant \frac{1}{d}$。另一方面，取 $m_n \in M$，使 $\|x_1 - m_n\| \to d$，而

$$\|f\|_{M_1} \geqslant \left|f\left(\frac{x_1 - m_n}{\|x_1 - m_n\|}\right)\right| = \frac{1}{\|x_1 - m_n\|}。$$

在上式中令 $n \to \infty$，得 $\|f\|_{M_1} \geqslant \frac{1}{d}$，故 $\|f\|_{M_1} = \frac{1}{d}$。最后，由定理 2，存在 X 上有界线性泛函 F 满足 $F(m) = f(m)(m \in M_1)$ 且 $\|F\| = \|f\|_{M_1}$，据 f 的构造，F 显然满足

$$F(x) = \begin{cases} 1, \text{当 } x = x_1; \\ 0, \text{当 } x \in M。\end{cases}$$

证毕。

注　推论 1 说明有界线性泛函可分离一点和一个闭子空间。

【推论 2】　设 X 是实赋范线性空间，$x_0 \in X$，且 $x_0 \neq \theta$，则存在 X 上有界线性泛函 F 满足 $F(x_0) = \|x_0\|$ 且 $\|F\| = 1$。

证　取 X 的一维子空间 $M = \{\alpha x_0 : \alpha \in R\}$。在 M 上定义泛函 f 为 $f(\alpha x_0) = \alpha\|x_0\|$，则 f 是 M 上线性泛函，且 $f(x_0) = \|x_0\|$。又

$$|f(\alpha x_0)| = |\alpha| \|x_0\| = \|\alpha x_0\|,$$

所以 $\|f\|_M = \sup\{|f(\alpha x_0)| : \|\alpha x_0\| \leqslant 1\} = 1$。由定理 2，存在 X 上有界线性泛函 F，它是 f 的保范延拓，因此 F 仍然满足 $F(x_0) = \|x_0\|$，$\|F\| = 1$。证毕。

当 X 是复线性空间时，Hahn - Banach 定理变形为：

【定理 3】　设 X 是复线性空间时，$P: X \to [0, +\infty)$ 满足：(1) $P(x+y) \leqslant P(x) + P(y)$；(2) $P(\alpha x) = |\alpha| P(x)(x, y \in X, \alpha \in \mathbf{C})$。$M$ 是 X 的一个复线性子空间，f 是 M 上定义的一个复线性泛函且 $|f(m)| \leqslant P(m)(m \in M)$，则存在 X 上定义的一个复线性泛函 F 满足

(1) $F(m)=f(m)$;$(m\in M)$,

(2) $|F(x)|\leqslant P(x)(x\in M)$。

由定理3,对于复线性空间上的复泛函,定理2及推论1和推论2仍然成立。

习题 3.3

1. 设 X 是实线性空间,M 是 X 的子空间,$x_0\in X\setminus M$。证明:$M_1=\{m+\alpha x_0: m\in M,\alpha\in R\}$ 是 X 的子空间。

2. 设 X 是赋范线性空间,$M\subset X$ 是非空闭集,$x_1\in X\setminus M$。记 $d=\inf\{\|x_1-m\|: m\in M\}$,证明:$d>0$。

3. 设 X 是赋范线性空间,$x,y\in X$,且 $x\neq y$。证明:存在有界线性泛函 F,满足 $F(x)\neq F(y)$。

4. 设 M 是赋范线性空间 X 的子集,$x_0\in X$ 且 $x_0\neq\theta$,证明 $x_0\in\overline{\mathrm{span}M}\Leftrightarrow$ 对每个 X 上有界线性泛函 f,只要 $f(x)=0,\forall x\in M$,则必 $f(x_0)=0$。

5. 设 X 是赋范线性空间,$x_0\in X$,若对任意 X 上有界线性泛函 f,且 $\|f\|=1$,使 $|f(x_0)|\leqslant\lambda$,证明 $\|x_0\|\leqslant\lambda$。

6. 在 $C[a,b]$ 上定义泛函

$$F(x)=\int_a^b x(t)\mathrm{d}t,\quad x\in[a,b]$$

证明:F 是有界线性泛函且 $\|F\|=b-a$。

7. 设 f 是线性空间 X 上的非零线性泛函,取 $x_0\in X\setminus \mathrm{Ker}(f)$。证明:$X=\{y+\alpha x_0: y\in \mathrm{Ker}(f),\alpha\in F\}$。

8. 设 f_1,f_2 是线性空间 X 上的两个线性泛函且 $\mathrm{Ker}(f_1)=\mathrm{Ker}(f_2)$,证明,存在常数 λ,使得 $f_1=\lambda f_2$。

9. 设 X,Y 都是赋范线性空间，且 $X\neq\{\theta\}$，若 $L(X,Y)$ 是 Banach 空间，证明 Y 必是 Banach 空间。

3.4　共轭空间与共轭算子

3.4.1　共轭空间

设 X 是赋范线性空间，称 X 上全体有界线性泛函组成的 Banach 空间 X^* 为 X 的共轭空间。为了寻求 X^* 的具体形式，我们引进下面等距同构的概念。

【定义 1】　设 X,Y 是两个赋范线性空间，称 X 是 Y 的嵌入子空间，如果存在线性算子 $T:X\to Y$ 满足 $\|Tx\|=\|x\|\ (x\in X)$；又称 X 与 Y 是等距同构的，如果存在线性算子 $T:X\to Y$ 是满映射且 $\|Tx\|=\|x\|\ (x\in X)$。

注　当 X 是 Y 的嵌入子空间时，X 与 Y 的子空间 TX 结构完全相同，因此可记为 $X\subset Y$。当 X 与 Y 等距同构时，这两个空间结构也完全相同，可记为 $X=Y$。

例 1　$c_0=\{\{a_i\}:\lim\limits_i a_i=0\}$，赋予范数

$$\|\{a_i\}\|=\sup_i|a_i|,$$

则 c_0 是赋线性空间。$l^1=\{\{b_i\}:\sum\limits_{i=1}^{\infty}|b_i|<+\infty\}$ 赋予范数

$$\|\{b_i\}\|=\sum_{i=1}^{\infty}|b_i|,$$

则 l^1 也是赋范线性空间。我们有 $(c_0)^*=l^1$。

证　对于任一 $\eta=\{b_i\}\in l^1$，定义 c_0 上线性泛函 F_η 为

$$F_{\eta}(\{a_i\}) = \sum_{i=1}^{\infty} a_i b_i,$$

于是 $|F_{\eta}(\{a_i\})| = \left|\sum_{i=1}^{\infty} a_i b_i\right| \leqslant \sup_i |a_i| \left(\sum_{i=1}^{\infty} |b_i|\right) = \|\eta\| \|\{a_i\}\|$，所以 $\|F_{\eta}\| \leqslant \|\eta\|$，即 $F_{\eta} \in (c_0)^*$。另一方面，对 $F \in (c_0)^*$，令 $b_i = F(e^{(i)})$，这里 $e^{(i)} = (0,0,\cdots,\underset{\text{第}i\text{个}}{\underset{\uparrow}{1}},0,\cdots)$，记 $\eta = \{b_i\}$。对每个 $\{a_i\} \in c_0$，由于 $\left\|\sum_{i=1}^{n} a_i e^{(i)} - \{a_i\}\right\| = \sup_{i \geqslant n+1} |a_i| \to 0 (n \to \infty)$，而 F 是连续性泛函，因此

$$F(\{a_i\}) = \lim_{n\to\infty} F\left(\sum_{i=1}^{n} a_i e^{(i)}\right) = \lim_{n\to\infty} \sum_{i=1}^{n} a_i F(e^{(i)}) =$$

$$\lim_{n\to\infty} \sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{\infty} a_i b_i。\tag{3.4}$$

接下来证明 $\eta \in l^1$。令 $\xi^{(N)} = \{\xi_n^{(N)}\}$，其中

$$\xi_n^{(N)} = \begin{cases} \operatorname{sgn} b_n, & n \leqslant N; \\ 0, & n > N。\end{cases}$$

这里 $\operatorname{sgn} x$ 是符号函数。则 $\lim_{n\to\infty} \xi_n^{(N)} = 0$，即 $\xi^{(N)} \in c_0$。且 $\|\xi^{(N)}\| \leqslant 1$。由式(3.4)

$$\sum_{n=1}^{N} |b_n| = \sum_{n=1}^{N} b_n \operatorname{sgn}(b_n) = F(\{\xi_n^{(N)}\}) \leqslant \|F\| \|\xi^{(N)}\| \leqslant \|F\|。$$

由 N 的任意性得 $\sum_{n=1}^{\infty} |b_n| < +\infty$，于是得到 $\|\eta\| \leqslant \|F\|$。根据上述两步，定义 $T: l^1 \to c_0^*$ 为 $T(\eta) = F_{\eta}$，$\eta \in l^1$，则 T 是线性算子，是满映射，而且 $\|T(\eta)\| = \|\eta\|$（因而是一一映射）。这说明 l^1 与 c_0^* 是等距同构，即 $c_0^* = l^1$。

例 2 设 $p > 1$，记

$$l^p = \{\{a_i\}: \sum_{i=1}^{\infty} |a_i|^p < +\infty\}。$$

在 l^p 上定义范数 $\| \{a_i\} \|_p = (\sum_{i=1}^{\infty} |a_i|^p)^{\frac{1}{p}}$,则 l^p 是赋范线性空间。我们有$(l^p)^* = l^q$,其中$\frac{1}{p} + \frac{1}{q} = 1, p, q > 1$。

证　对每个 $\eta = \{b_i\} \in l^q$,由级数形式的 Hölder 不等式对 $\{a_i\} \in l^p$,有

$$\sum_{i=1}^{\infty} |a_i| |b_i| \leqslant (\sum_{i=1}^{\infty} |a_i|^p)^{\frac{1}{p}} (\sum_{i=1}^{\infty} |b_i|^q)^{\frac{1}{q}}, \tag{3.5}$$

因此定义 l^p 上的线性泛函 F_η 为

$$F_\eta(\{a_i\}) = \sum_{i=1}^{\infty} a_i b_i,$$

那么由式(3.5) 得 $|F_\eta(\{a_i\})| \leqslant \|\eta\|_q \|\{a_i\}\|_p$,即 $\|F_\eta\| \leqslant \|\eta\|_q$。

另一方面,设 $F \in (l^p)^*$。记 $e^{(i)} = (0,0,\cdots, \underset{\substack{\uparrow \\ \text{第}i\text{个}}}{1}, 0,\cdots)$,对每个 $\{a_i\} \in l^p$,由于

$$\left\| \sum_{i=1}^{n} a_i e^{(i)} - \{a_i\} \right\|_p = (\sum_{i=n+1}^{\infty} |a_i|^p)^{\frac{1}{p}} \to 0(n \to \infty),$$

F 是连续线性泛函,所以

$$F(\{a_i\}) = \lim_{n\to\infty} F(\sum_{i=1}^{n} a_i e^{(i)}) = \lim_{n\to\infty} \sum_{i=1}^{n} a_i F(e^{(i)}) = \sum_{i=1}^{\infty} a_i F(e^{(i)})。 \tag{3.6}$$

记 $b_i = F(e^{(i)}), \eta = \{b_i\}$,来证 $\eta \in l^q$。对自然数 N,记

$$\xi_n^{(N)} = \begin{cases} \dfrac{|b_n|^q}{b_n}, & \text{当 } n \leqslant N \text{ 且 } b_n \neq 0; \\ 0, & \text{当 } n > N \text{ 或 } b_n = 0。 \end{cases}$$

则 $\xi^{(N)} = \{\xi_n^{(N)}\} \in l^p$。由式(3.6)得

$$\sum_{n=1}^{N} |b_n|^q = F(\xi^{(N)}) \leqslant \|F\| \|\xi^{(N)}\|_p =$$

$$\|F\| (\sum_{n=1}^{N} |b_n|^{p(q-1)})^{\frac{1}{p}} = \|F\| (\sum_{n=1}^{N} |b_n|^q)^{\frac{1}{p}},$$

于是$(\sum_{n=1}^{N} |b_n|^q)^{\frac{1}{q}} \leqslant \|F\|$,令 $N \to \infty$,则$(\sum_{n=1}^{\infty} |b_n|^q)^{\frac{1}{q}} \leqslant \|F\|$,即 $\|\eta\|_q \leqslant \|F\|$。根据上述两方面的证明定义线性算子 $T: l^q \to (l^p)^*$ 为 $T(\eta) = F_\eta$,则 T 是满映且 $\|T(\eta)\| = \|\eta\|_q$,故$(l^p)^* = l^q$。

记 $l^\infty = \{\{a_i\}: \sup\limits_i |a_i| < +\infty\}$,定义范数

$$\|\{a_i\}\|_\infty = \sup_i |a_i|,$$

可以证明$(l^1)^* = l^\infty$(留作习题)。

例3 $L^p[a,b] = \{f: |f|^p$ 是在$[a,b]$上Lebesgue可积函数$\}$($p > 1$),定义范数 $\|f\|_p = (\int_a^b |f|^p \mathrm{d}t)^{\frac{1}{p}}$,则 $F \in (L^p[a,b])^*$ 的充要条件是存在 $g \in L^q[a,b]$(其中$\frac{1}{p} + \frac{1}{q} = 1$)满足 $\|F\| = \|g\|_q$ 及

$$F(x) = \int_a^b x(t)g(t)\mathrm{d}t, x \in L^p[a,b],$$

即$(L^p[a,b])^* = L^q[a,b]$。

这里略去证明。

X 是赋范线性空间,我们已经知道 X 的共轭空间 X^* 是一个Banach 空间。同样,可定义 X^* 的共轭空间记为 X^{**},下面的定理阐明了 X 与 X^{**} 之间的关系。

【定理1】 X 是 X^{**} 的嵌入子空间即 $X \subset X^{**}$。

证　对每个 $x \in X$，定义 X^* 上的泛函 $J(x)$ 为

$$J(x)(x^*) = x^*(x) \quad (x^* \in X^*)。$$

不难验证算子 $J: X \to X^{**}$ 是线性的。由于 $|J(x)(x^*)| = |x^*(x)| \leqslant \|x^*\| \|x\|$，所以 $J(x)$ 是 X^* 上的有界线性泛函，即 $J(x) \in X^{**}$ 且 $\|J(x)\| \leqslant \|x\|$。另一方面，由本章 3.3 节推论 2，存在 $x_0^* \in X^*$ 且 $\|x_0^*\| = 1$，使 $x_0^*(x) = \|x\|$，那么

$$\|J(x)\| = \sup\{|J(x)(x^*)| : \|x^*\| \leqslant 1\} \geqslant |J(x)(x_0^*)| = |x_0^*(x)| = \|x\|,$$

故 $\|J(x)\| = \|x\|$。所以 X 等距嵌入 X^{**}，即 $X \subset X^{**}$。证毕。

【定义 2】　称 X 是自反的，如果 $X = X^{**}$，即 X 与 X^{**} 等距同构。

从上面的例子可见，$(c_0)^* = l^1$，$(l^1)^* = l^\infty$，因此有 $(c_0)^{**} = l^\infty$，且 $c_0 \neq l^\infty$，所以 c_0 不是自反空间。$(l^p)^* = (l^q)$，$(l^q)^* = (l^p)$，所以 l^p 是自反空间 $(p > 1)$。同样，$L^p[a,b]$ 也是自反空间 $(p > 1)$。

【定义 3】　设 X 是赋范线性空间，$\{x_n\} \subset X$，$x \in X$。称 $\{x_n\}$ 弱收敛于 x，是指对 $\forall f \in X^*$，有 $\lim\limits_{n \to \infty} f(x_n) = f(x)$。记为 $x_n \xrightarrow{w} x$。

关于弱收敛有如下性质：

(1) 若 $x_n \to x$，则 $x_n \xrightarrow{w} x$；

(2) 若 $x_n \xrightarrow{w} x$ 且 $x_n \xrightarrow{w} y$，则 $x = y$（弱收敛极限的惟一性）；

(3) 若 $x_n \xrightarrow{w} x$，则 $\{x_n\}$ 有界；

(4) 若 X 是有限维的，则由 $x_n \xrightarrow{w} x$ 可推出 $x_n \to x$。

证

(1) 对每个 $f \in X^*$，$|f(x_n) - f(x)| \leqslant \|f\| \|x_n - x\| \to 0$，

所以 $x_n \xrightarrow{w} x$。

(2) 对每个 $f \in X^*$，由 $x_n \xrightarrow{w} x$ 及 $x_n \xrightarrow{w} y$ 得

$$f(x) = \lim_{n\to\infty} f(x_n) = f(y),$$

于是 $f(x-y) = 0$。由本章 3.3 节推论 2，存在 $f_0 \in X^*$ 且 $\|f_0\| = 1$，$f_0(x-y) = \|x-y\|$，故 $\|x-y\| = 0$，即 $x = y$。

(3) 对每个 $f \in X^*$，由于 $\lim_{n\to\infty} f(x_n) = f(x)$，因此 $\{f(x_n)\}$ 有界，即 $\sup_n |f(x_n)| < +\infty$。定义 X^* 上有界线性泛函列 $J_n(f) = f(x_n)$，$(n = 1,2,\cdots)$，则由上面定理 1，$\|J_n\| = \|x_n\|$，再由共鸣定理知 $\sup\|J_n\| = \sup_n \|x_n\| < +\infty$，故 $\{x_n\}$ 有界。

(4) 取 X 的一组基为 $\{e_1, e_2, \cdots, e_d\}$。设 $x_n \to x$。$\{x_n\}$，x 在这组基下可展成

$$x_n = \sum_{j=1}^{d} a_j^{(n)} e_j, \quad (n = 1,2,\cdots)$$

$$x = \sum_{j=1}^{d} a_j e_j。$$

取特殊的坐标泛函 $f_i \in X^*$，使

$$f_i(e_j) = \delta_{ij} = \begin{cases} 1, i = j; \\ 0, i \neq j。\end{cases} \quad (i, j = 1,2,\cdots,d)$$

那么 $f_i(x_n) = a_i^{(n)}$，$f_i(x) = a_i$。由 $f_i(x_n) \to f(x)$ 得

$$a_i^{(n)} \to a_i (n \to \infty, i = 1,2,\cdots,d)。$$

因此

$$\|x_n - x\| = \left\|\sum_{i=1}^{d} (a_i^{(n)} - a_i) e_i\right\| \leqslant$$
$$\sum_{i=1}^{d} |a_i^{(n)} - a_i| \|e_i\| \to 0 (n \to \infty)。$$

即 $x_n \to x$。证毕。

注　在一般无穷维空间中 $x_n \xrightarrow{w} x$ 并不能推出 $x_n \to x$。

【定义 4】　设赋范线性空间 X 的共轭空间为 X^*，在 X^* 上有如下三种收敛：

(1) 按范数收敛，记为 $f_n \to f$，即 $\|f_n - f\| \to 0$；

(2) 弱收敛，记为 $f_n \xrightarrow{w} f$，即对每个 $x^{**} \in X^{**}$，$x^{**}(f_n) \to x^{**}(f)$；

(3) 弱 * 收敛，记为 $f_n \xrightarrow{w^*} f$，即对每个 $x \in X$，$f_n(x) \to f(x)$。

这三种收敛的关系是：若 $f_n \to f$，则 $f_n \xrightarrow{w} f$；若 $f_n \xrightarrow{w} f$，则 $f_n \xrightarrow{w^*} f$。当 X 是自反 Banach 空间时，$f_n \xrightarrow{w} f$ 等价于 $f_n \xrightarrow{w^*} f$。

3.4.2　共轭算子

大家知道，矩阵是有限维空间中算子的表示形式，矩阵的转置在矩阵理论中起着十分重要的作用。这种矩阵转置概念在无穷维空间的推广就是共轭算子。

设 X, Y 是两个赋范线性空间，共轭空间分别为 X^*, Y^*。设 $T \in L(X, Y)$，通过 T 诱导出另一个算子 $T^*: Y^* \to X^*$。具体做法如下：设 $y^* \in Y^*$。令 $f(x) = y^*(Tx)(x \in X)$，则

$$|f(x)| = |y^*(Tx)| \leqslant \|y^*\| \|Tx\| \leqslant \|y^*\| \|T\| \|x\|,$$

于是 f 是有界线性泛函且 $\|f\| \leqslant \|y^*\| \|T\|$，这样定义 $T^* y^* = f \in X^*$，便得到了 T^*。

【定义 5】　设 $T \in L(X, Y)$，称算子 $T^*: Y^* \to X^*$ 为 T 的共轭

算子,如果满足对任何 $x \in X$ 及 $y^* \in Y^*$ 有

$$T^* y^*(x) = y^*(Tx)。$$

【定理2】 设 $T \in L(X,Y)$,则 $T^* \in L(Y^*,X^*)$ 且 $\|T^*\| = \|T\|$。

证 由于

$$\begin{aligned}\|T^* y^*\| &= \sup\{|T^* y^*(x)| : \|x\| \leqslant 1\} = \\ &\sup\{|y^*(Tx)| : \|x\| \leqslant 1\} \leqslant \\ &\sup\{\|y^*\| \|T\| \|x\| : \|x\| \leqslant 1\} = \\ &\|T\| \|y^*\|,\end{aligned}$$

所以 $\|T^*\| \leqslant \|T\|$。另一方面,对每个 $x \in X$,由 Hahn - Banach 定理的推论 2,存在 $y_0^* \in Y^*$ 使 $\|y_0^*\| = 1$ 且 $y_0^*(Tx) = \|Tx\|$,于是

$$\begin{aligned}\|Tx\| &= y_0^*(Tx) = T^* y_0^*(x) \leqslant \\ &\|T^* y_0^*\| \cdot \|x\| \leqslant \|T^*\| \|y_0^*\| \|x\| = \\ &\|T^*\| \|x\|,\end{aligned}$$

故 $\|T\| \leqslant \|T^*\|$,即 $\|T\| = \|T^*\|$,而 T^* 的线性明显。证毕。

关于共轭算子,还有如下简单性质,读者根据定义可直接验证。

设 X,Y,Z 是赋范线性空间,若 $T_1 \in L(X,Y)$,$T_2 \in L(Y,Z)$,则

(1)$(\alpha T_1)^* = \alpha T_1^* (\alpha \in F)$;

(2)$(T_2 \cdot T_1)^* = T_1^* \cdot T_2^*$;

(3) 当 $T_2 \in L(X,Y)$ 时,$(T_1 + T_2)^* = T_1^* + T_2^*$;

(4) 若 $I: X \to X$ 是恒等算子,则 $I^*: X^* \to X^*$ 也是恒等算子。

习题 3.4

1. 证明:$(l^1)^* = l^\infty$。

2. 在 l^p 中作出一个点列 x_n 及 x，使 $x_n \xrightarrow{w} x$，但 $x_n \nrightarrow x$（即 x_n 依范数不收敛于 x），其中 $p > 1$。

3. 设 X 是赋范线性空间，$M \subset X$ 是闭线性子空间，若 $x_n \in M$，有 $x_n \xrightarrow{w} x_0$，证明 $x_0 \in M$ 且 $\varliminf\limits_{n\to\infty} \| x_n \| \geqslant \| x_0 \|$。

4. 设 X 是 Banach 空间，X^* 是共轭空间，若 $x_n \to x$，且 $f_n \xrightarrow{w^*} f$，证明：$f_n(x_n) \to f(x)$。

5. $x_n \xrightarrow{w} x$，$T \in \mathrm{L}(X, Y)$，证明：$Tx_n \xrightarrow{w} Tx$。

6. 设 $X = \mathrm{C}[a, b]$，且 $x_n \xrightarrow{w} x_0$。证明：$\lim\limits_{n\to\infty} x_n(t) = x_0(t)$。

7. 证明：若 X 是自反的，则 X^* 也是自反的。

8. 设矩阵 $A = (a_{ij})_{n\times m}$，定义算子 $T: \mathbf{R}^m \to \mathbf{R}^n$ 为 $Tx = Ax$，证明：$T^*: \mathbf{R}^n \to \mathbf{R}^m$ 为 $T^* x = A^* x$。这里 A^* 表示矩阵 A 的转置。

3.5　开映射、逆算子及闭图象定理

本节研究有界线性算子的最重要几个性质，开映射、逆算子及闭图象定理是泛函分析的支柱定理，不仅具有重要的理论意义，而且也有广泛的应用价值。

【定理 1】（Banach 开映射原理）设 X, Y 是两个 Banach 空间，$T \in \mathrm{L}(X, Y)$，若 T 满足 $TX = \{Tx : x \in X\} = Y$，那么存在常数 $\beta > 0$，使对一切 $y \in Y$ 有相应的 $x \in X$ 满足 $Tx = y$ 且 $\| x \| \leqslant \beta \| y \|$。

证　为方便起见，对 $r > 0$，记 $B(r) = \{x : \| x \| \leqslant r\}$，$U(r) = \{y \in Y : \| y \| \leqslant r\}$。由于 $X = \bigcup\limits_{n=1}^{\infty} B(n)$ 及 $Y = TX$，那么

$$Y = \bigcup_{n=1}^{\infty} TB(n)。$$

Y 是 Banach 空间，据 Baire 纲定理，Y 是第二纲的，因此存在某个自然数 n_0 使 $TB(n_0)$ 不是稀疏集。也就是说有 Y 中某闭球 $U_{r_0}(y_0) = \{y \in Y: \| y - y_0 \| \leqslant r_0\}(r_0 > 0)$，满足 $TB(n_0)$ 在 $U_{r_0}(y_0)$ 中稠密。令 $\delta_0 = \frac{r_0}{n_0}$。我们来证 $TB(1)$ 在 $U(\delta_0)$ 中稠密，即 $U(\delta_0) \subset \overline{TB(1)}$。设 $y \in U(\delta_0)$，则 $y_0 + n_0 y, y_0 - n_0 y \in U_{r_0}(y_0)$。据 $TB(n_0)$ 在 $U_{r_0}(y_0)$ 中稠密，于是存在点列 $\{x_k\}$ 及 $\{x'_k\} \subset B(n_0)$ 满足

$$\lim_{k \to \infty} Tx_k = y_0 + n_0 y, \lim_{k \to \infty} Tx'_k = y_0 - n_0 y。$$

进一步得

$$\lim_{k \to \infty} T\left(\frac{x_k - x'_k}{2n_0}\right) = y,$$

再注意到

$$\left\| \frac{x_k - x'_k}{2n_0} \right\| \leqslant \frac{\| x_k \| + \| x'_k \|}{2n_0} \leqslant \frac{n_0 + n_0}{2n_0} = 1,$$

所以 $\frac{x_k - x'_k}{2n_0} \in B(1)$。亦即 $TB(1)$ 在 $U(\delta_0)$ 中稠密。又因 $\frac{1}{2^n}TB(1) = TB(\frac{1}{2^n})$，$\frac{1}{2^n}U(\delta_0) = U(\frac{\delta_0}{2^n})$，故 $TB(\frac{1}{2^n})$ 也在 $U(\frac{\delta_0}{2^n})$ 中稠密。接下来，证明 $TB(1) \supset U(\frac{\delta_0}{2})$。因 $TB(\frac{1}{2})$ 在 $U(\frac{\delta_0}{2})$ 中稠密，任取 $y \in U(\frac{\delta_0}{2})$，从而存在 $x_1 \in B(\frac{1}{2})$，使

$$\| y - Tx_1 \| \leqslant \frac{\delta_0}{2^2},$$

即 $y - Tx_1 \in U(\frac{\delta_0}{2^2})$。又 $TB(\frac{1}{2^2})$ 在 $U(\frac{\delta_0}{2^2})$ 中稠密，从而存在 $x_2 \in$

$B(\frac{1}{2^2})$ 使

$$\| y - Tx_1 - Tx_2 \| = \| y - T(x_1 + x_2) \| \leqslant \frac{\delta_0}{2^3},$$

如此做下去可得点列 $\{x_n\}$ 满足：

(1) $x_n \in B(\frac{1}{2^n})(n = 1,2,\cdots)$;

(2) $\| y - T(x_1 + x_2 + \cdots + x_n) \| \leqslant \frac{\delta_0}{2^{n+1}}$。

令 $S_n = \sum\limits_{k=1}^{n} x_k$，由(1) $\| x_k \| \leqslant \frac{1}{2^k}$ 且当 $n > m$ 时有

$$\| S_n - S_m \| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leqslant \sum_{k=m+1}^{n} \frac{1}{2^k} \to 0(m \to \infty),$$

故 $\{S_n\}$ 是 Cauchy 列，X 是 Banaxh 空间，于是存在 $x \in X$，$S_n \to x$。据(2)及 T 的连续性，令 $n \to \infty$，得到 $y = Tx$。又 $\| x \| \leqslant \sum\limits_{k=1}^{\infty} \| x_k \| \leqslant \sum\limits_{k=1}^{\infty} \frac{1}{2^k} = 1$，所以 $x \in B(1)$。这便证明了 $TB(1) \supset U(\frac{\delta_0}{2})$。

最后，对任意 $y \in Y, y \neq 0$，$\frac{\delta_0 y}{2\| y \|} \in U(\frac{\delta_0}{2})$。由上面已证 $TB(1) \supset U(\frac{\delta_0}{2})$ 知，有 $x' \in B(1)$ 使

$$Tx' = \frac{\delta_0 y}{2\| y \|},$$

即 $T(\frac{2\| y \|}{\delta_0} x') = y$。令 $x = \frac{2\| y \|}{\delta_0} x'$，则

$$Tx = y \quad 且 \quad \| x \| \leqslant \frac{2}{\delta_0} \| y \|。$$

取 $\beta = \frac{2}{\delta_0}$，则 $\| x \| \leqslant \beta \| y \|$。证毕。

注　定理1是表征有界线性算子把开集映成开集，因此把它称为开映射原理。

【定理2】(逆算子定理)　设 X,Y 是两个 Banach 空间，$T:X\to Y$ 是一一映射，$TX=Y$，且 $T\in L(X,Y)$，则 T 的逆算子 $T^{-1}:Y\to X$ 也是有界线性算子，即 $T^{-1}\in L(Y,X)$。

证　因 T 是一一映射的线性算子，所以 T^{-1} 存在且也是线性算子。由定理1，存在 $\beta>0$，使 $\|T^{-1}y\|=\|x\|\leqslant\beta\|y\|$，所以 T^{-1} 是有界线性算子。证毕。

下面通过逆算子定理来说明常微分方程解的连续依赖性。

例1　给定 k 阶线性常微分方程及初值条件

$$\begin{cases}x^{(k)}(t)+p_1(t)x^{(k-1)}(t)+\cdots+p_{k-1}(t)x'(t)+p_k(t)x(t)=y(t),\\ x(0)=x'(0)=\cdots=x^{(k-1)}(0)=0。\end{cases}$$

其中 $p_i(t)(i=1,2,\cdots k)$ 是 $[0,1]$ 上的连续函数。根据常微分方程的理论，对每个连续函数 $y(t)\in C[0,1]$，上述微分方程均存在惟一解，这里的解连续相依性是指当左边函数 $y(t)$ 作微小变化时，相应的解也作微小变化。

证　定义 $C[0,1]$ 的一个线性子空间

$C_0^{(k)}[0,1]=\{x(t):x(0)=x'(0)=\cdots=x^{(k-1)}(0),x(t)$ 在 $[0,1]$ 上 k 次连续可微$\}$。

在 $C[0,1]$ 中定义通常范数 $\|y\|=\max\{|y(t)|:t\in[0,1]\}$。在空间 $C_0^{(k)}[0,1]$ 中定义范数

$$\|x\|_1=\sum_{i=0}^{k}\max\{|x^{(i)}(t)|:t\in[0,1]\},x^{(0)}(t)=x(t),$$

那么 $C[0,1]$ 及 $C_0^{(k)}[0,1]$ 在相应范数下均构成 Bananch 空间。定义线性算子 $T:C_0^{(k)}[0,1]\to C[0,1]$ 为

$$(Tx)(t)=x^{(k)}(t)+p_1(t)x^{(k-1)}(t)+\cdots+$$

$$p_{k-1}(t)x'(t) + p_k(t)x(t),$$

那么

$$\| Tx \| \leqslant (1 + \sum_{i=1}^{k} \| p_i \|) \| x \|_1,$$

这里 $\| p_i \| = \max\{ | p_i(t) | : t \in [0,1]\}$。所以 T 是有界线性算子，据 T 的定义及例 1 中的说明，T 是由 $C_0^{(k)}[0,1]$ 到 $C[0,1]$ 上的一一映射，因此由逆算子定理，T^{-1} 是有界线性算子。对任意 $\varepsilon > 0$，取 $\delta = \dfrac{\varepsilon}{\| T^{-1} \|}$，则只要 $\| y_1 - y_2 \| < \delta$ 时，$y_1(t), y_2(t)$ 相应的解为 $x_1(t), x_2(t)$，即 $y_1 = Tx_1, y_2 = Tx_2$，故 $x_1 = T^{-1}y_1, x_2 = T^{-1}y_2$，于是

$$\| x_1 - x_2 \|_1 = \| T^{-1}(y_1 - y_2) \| \leqslant \| T^{-1} \| \| y_1 - y_2 \| < \varepsilon。$$

这说明 $y(t)$ 作微小变动时，其相应的解 $x(t)$ 以及 $x(t)$ 的各阶导函数（直到 k 阶）也作微小的扰动。

判断一个线性算子是否有界，有时是十分困难的。我们下面介绍通过算子图象特征来判断算子有界的方法，这就是著名的闭图象定理。

在高等数学中，函数 $y = f(x)$ 的图象是平面上的一条曲线，也就是由平面上的点 $(x, f(x))$ 组成的集合，我们把这一概念推广到抽象空间。

设 X, Y 是两个赋范线性空间，其直积空间 $X \times Y$ 为

$$X \times Y = \{(x, y) : x \in X, y \in Y\}。$$

在 $X \times Y$ 中规定加法和数乘运算为

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2);$$

$$\lambda(x, y) = (\lambda x, \lambda y),$$

则 $X \times Y$ 为线性空间，引进范数 $\| (x, y) \| = \| x \| + \| y \|$，则

$X\times Y$ 在此范数下构成赋范线性空间。这里 $\|x\|$ 表示 X 中的范数，$\|y\|$ 表示 Y 中范数。此时称 $X\times Y$ 为赋范线性空间 X 与 Y 的直积赋范线性空间。特别，容易验证，当 X,Y 均是 Banach 空间时，$X\times Y$ 也是 Banach 空间。

【定义 1】 设 X,Y 是两个赋范线性空间，$T:A\to Y$ 是线性算子，其中 $A\subset X$，记 T 的图象为 $G(T)=\{(x,Tx)\in X\times Y:x\in A\}$，称 T 具有闭图象，如果 $G(T)$ 是 $X\times Y$ 的闭集。

由于 T 是线性算子，因此 $G(T)$ 是 $X\times Y$ 的线性子空间。

【定理 3】(闭图象定理) 设 X,Y 是两个 Banach 空间，$T:X\to Y$ 是线性算子。如果 T 具有闭图象，那么 T 是有界的。

证 因为 $G(T)\subset X\times Y$ 是闭线性子空间，因此在 $X\times Y$ 的范数下 $G(T)$ 本身也是一个 Banach 空间。定义算子 $H:G(T)\to X$ 为

$$H(x,Tx)=x。$$

显然 H 是映上的线性算子且是一一映射。又

$$\|H(x,Tx)\|=\|x\|\leqslant\|x\|+\|Tx\|=\|(x,Tx)\|,$$

所以 $\|H\|\leqslant 1$，即 H 是有界的。据逆算子定理，H^{-1} 也是有界的，于是

$$\|(x,Tx)\|=\|H^{-1}x\|\leqslant\|H^{-1}\|\,\|x\|,$$

从而有

$$\|x\|+\|Tx\|\leqslant\|H^{-1}\|\,\|x\|。$$

更有 $\|Tx\|\leqslant\|H^{-1}\|\,\|x\|$，故 T 是有界的。证毕。

习题 3.5

1. 设映射 $T:\mathbf{R}^2\to\mathbf{R}$ 为

$$Tx=t_1,\ \forall x=(t_1,t_2)\in\mathbf{R}^2$$

试问 T 是开映射吗?又令

$$Tx = (t_1, 0), \forall x = (t_1, t_2) \in \mathbf{R}^2$$

问 $T:\mathbf{R}^2 \to \mathbf{R}^2$ 是否为开映射?

3. 设 X, Y 是两个 Banach 空间,证明:$X \times Y$ 也是 Banach 空间。

4. 设 X, Y 是两个 Banach 空间,$T: X \to Y$ 是具有闭图象的线性算子。证明:$G(T)$ 也是 Banach 空间。

4. 设 X, Y 是赋范线性空间,$T \in \mathrm{L}(X, Y)$,证明:T 具有闭图象(即闭图象逆定理成立)。

5. 设 X, Y 都是赋范线性空间,对于线性算子 $T: A \to Y$,其中 A 是 X 的线性子空间,证明 T 具有闭图象 $\Leftrightarrow$ 若 $x_n \in A, y \in Y$,只要 $x_n \to x, Tx_n \to y$,则 $x \in A$ 且 $Tx = y$。

6. 设 X 是线性空间,$\|\cdot\|_1, \|\cdot\|_2$ 是 X 上两个范数,在这两个范数下,X 均是 Banach 空间。如果存在常数 $\mu > 0$,使 $\|x\|_1 \leqslant \mu\|x\|_2$,则一定也存在常数 $\lambda > 0$,使 $\|x\|_2 \leqslant \lambda\|x\|_1$(即两个范数等价)。

7. 设 X, Y 是两个 Banach 空间,$T \in \mathrm{L}(X, Y)$,证明:存在常数 $\lambda > 0$,使对 $x \in X$,$\|Tx\| \geqslant \lambda\|x\|$ 的充要条件是 $\mathrm{Ker}T = \{\theta\}$ 且 $R(T)$ 是闭集。

8. 在开映射原理的条件下,证明:T 将 X 中开集映成 Y 中开集。

3.6　算子谱理论简介

我们在线性代数中学过了矩阵的特征值与特征向量的基本理论,现在把这两个概念推广到 Banach 空间,建立算子的谱理论。

3.6.1 有界线性算子的谱

为了研究算子的谱,Banach 空间一般取复的。

【定义 1】 设 X 是复 Banach 空间,$T \in L(X,X)$,λ 为一复数。

(1) 称 λ 为 T 的正则值,如果 $\lambda I - T$ 有有界逆算子,即 $(\lambda I - T)^{-1} \in L(X,X)$。用 $\rho(T)$ 表示 T 的正则值组成的集合,称 $\rho(T)$ 为 T 的正则集。对 $\lambda \in \rho(T)$,称 $R_\lambda = (\lambda I - T)^{-1}$ 为 T 的预解式。

(2) 如果 λ 不是 T 的正则值,则称 λ 为 T 的谱点,全体谱点的集合记为 $\sigma(T)$,称为 T 的谱,对谱中的点又可分为以下三种类型:

① 算子方程 $(\lambda I - T)x = \theta$ 有非零解的 λ 称为 T 的特征值或点谱,相应的非零解称为 T 的特征值;

② 方程 $(\lambda I - T)x = \theta$ 仅有零解,但 $R(\lambda I - T) \neq X$,而 $R(\lambda I - T)$ 即 $(\lambda I - T)$ 的值域在 X 中稠密,则称这样的 λ 为 T 的连续谱;

③ 方程 $(\lambda I - T)x = \theta$ 仅有零解,但 $R(\lambda I - T)$ 在 X 中不稠密,则称 λ 是 T 的剩余谱。

对于 T 的特征值,有如下基本性质:

(1) 如果 λ 是 T 的特征值,则 λ 所对应的特征向量加上零元素 θ 正好组成 X 的一个闭子空间,称这个子空间为 λ 的特征子空间。

(2) 若 $\lambda_i (i = 1,2,\cdots,n)$ 是 T 的 n 个不同的特征值,λ_i 对应的特征向量为 x_i,则 $x_1, x_2, \cdots, x_n$ 线性无关。

证　(1) $\mathrm{Ker}(\lambda I - T) = \{x:(\lambda I - T)x = \theta\}$ 正好是 λ 的特征子空间,因此是闭子空间。

(2) 用数学归纳法来证明。当 $n = 1$ 时,特征向量 $x_1 \neq 0$,因此线性无关。假设 $n = k$ 时结论成立,即 $x_1, x_2, \cdots, x_k$ 线性无关。若 $x_1, x_2, \cdots, x_k, x_{k+1}$ 线性相关,于是存在不全为零的常数 $\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}$ 成立

$$\sum_{i=1}^{k+1} \alpha_i x_i = \theta, \tag{3.7}$$

从而有

$$T\left(\sum_{i=1}^{k+1} \alpha_i x_i\right) = \sum_{i=1}^{k+1} \alpha_i \lambda_i x_i = \theta。 \tag{3.8}$$

在式(3.7)两边乘 λ_{k+1} 得

$$\sum_{i=1}^{k+1} \alpha_i \lambda_{k+1} x_i = \theta。 \tag{3.9}$$

式(3.8)减式(3.9)式得

$$\sum_{i=1}^{k} \alpha_i (\lambda_i - \lambda_{k+1}) x_i = \theta,$$

而 $x_1, x_2, \cdots, x_k$ 线性无关，所以

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = \alpha_2(\lambda_2 - \lambda_{k+1}) = \cdots = \alpha_k(\lambda_k - \lambda_{k+1}) = 0。$$

又由式(3.7)，必有某个 $\alpha_{i_0}(1 \leqslant i_0 \leqslant k) \neq 0$，这样有

$$\lambda_{i_0} = \lambda_{k+1},$$

这与 $\lambda_1, \lambda_2, \cdots, \lambda_k, \lambda_{k+1}$ 互不相同，矛盾。因此 $x_1, x_2, \cdots, x_{k+1}$ 线性无关。证毕。

关于算子的谱有以下基本定理。

【定理 1】 设 X 是复 Banach 空间，$T \in L(X, X)$，则

(1) 当 $|\lambda| > \|T\|$ 时，λ 是 T 的正则值；

(2) $\rho(T)$ 是复平面 $\mathbf{C}$ 的开集，$\sigma(T)$ 是 $\mathbf{C}$ 的有界闭集。

证

(1) 令 $A = \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}}$，由于

$$\sum_{n=0}^{\infty} \left\| \frac{T^n}{\lambda^{n+1}} \right\| \leqslant \sum_{n=0}^{\infty} \frac{\|T\|^n}{\lambda^{n+1}} = \frac{1}{|\lambda|} \sum_{n=0}^{\infty} \left(\frac{\|T\|}{|\lambda|} \right)^n,$$

注意到 $|\lambda| > \|T\|$，所以级数 $\sum_{n=0}^{\infty}\left(\frac{\|T\|}{|\lambda|}\right)^n$ 收敛，而因 $L(X,X)$ 是 Banach 空间，所以 $A \in L(X,X)$。直接验证得

$$A(\lambda I - T) = (\lambda I - T)A = I,$$

故 $A = (\lambda I - T)^{-1}$，即 $\lambda \in \rho(T)$。

(2) 设 $\lambda \in \rho(T)$，则当 $|\mu - \lambda| < \frac{1}{\|(\lambda I - T)^{-1}\|}$ 时，我们来证 $\mu \in \rho(T)$。由于

$$\mu I - T = (\lambda I - T) + (\mu - \lambda)I = (\lambda I - T)[I + (\mu - \lambda)(\lambda I - T)^{-1}]$$

而 $|\mu - \lambda| < \frac{1}{\|(\lambda I - T)^{-1}\|}$，据(1)之证，$I + (\mu - \lambda)(\lambda I - T)^{-1}$ 有有界逆，即 $[I + (\mu - \lambda)(\lambda I - T)^{-1}]^{-1}$ 存在，因此 $(\mu I - T)^{-1} \in L(X,X)$，即 $\mu \in \rho(T)$。这说明 $\rho(T)$ 是开集且是无界的。又 $\sigma(T) = \mathbf{C} \setminus \rho(T) = (\rho(T))^C$，即 $\sigma(T)$ 是 $\rho(T)$ 的余集，所以 $\sigma(T)$ 是有界闭集。证毕。

注 当 $X \neq \{\theta\}$ 时，$\sigma(T) \neq \varnothing$。

【定义 2】 记

$$\gamma(T) = \max\{|\lambda| : \lambda \in \sigma(T)\},$$

称 $\gamma(T)$ 为算子 T 的谱半径。

【定理 2】 谱半径有公式

$$\gamma(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}。$$

这个公式的证明需用到算子值解析函数的基本定理，所以略去其证明。

对于有限维空间上的线性算子，谱理论十分简单，λ 要么是正则值，要么是特征值，即谱中仅含有特征值。但对于无限维空间，其谱集

的结构通常十分复杂。

例 1　取 $X = l^1 = \{x: x = \{\xi_n\}, \xi_n \in \mathbf{C}$ 且 $\sum_{n=1}^{\infty} |\xi_n| < +\infty\}$，定义范数

$$\|x\| = \sum_{n=1}^{\infty} |\xi_n|,$$

则 X 是一个复 Banach 空间。定义有界线性算子 $T: X \to X$ 为($x = \{\xi_n\}$)

$$y = Tx,$$

记 $y = \{\eta_n\}$，则 $\eta_1 = 0, \eta_n = -\xi_{n-1} (n \geqslant 2)$。证明：

$$\rho(T) = \{\lambda \in C: |\lambda| > 1\},$$
$$\sigma(T) = \{\lambda \in C: |\lambda| \leqslant 1\},$$

且 $\sigma(T)$ 中没有特征值。

证　由于

$$\|Tx\| = \sum_{n=1}^{\infty} |\xi_n| = \|x\|,$$

所以 $\|T\| = 1$。据定理 1，仅需证明 $\sigma(T) = \{\lambda \in C: |\lambda| \leqslant 1\}$。设算子方程 $(\lambda I - T)x = \theta$ 的解为 $x(x = \{\xi_n\})$，则

$$\{\lambda\xi_1, \lambda\xi_2 + \xi_1, \lambda\xi_3 + \xi_2, \cdots, \lambda\xi_n + \lambda\xi_{n-1}, \cdots\} = 0 \quad (3.10)$$

当 $\lambda = 0$ 时，由式(3.10)得

$$\xi_1 = \xi_2 = \cdots \xi_n = \cdots = 0,$$

于是 $x = \theta$ 即算子方程 $(\lambda I - T)x = 0$ 仅有零解。当 $\lambda \neq 0$，再由式(3.10)，得 $\lambda\xi_1 = 0$，则 $\xi_1 = 0$，由 $\lambda\xi_2 + \xi_1 = 0$ 又得 $\lambda\xi_2 = 0$，得 $\xi_2 = 0$，依次可得

$$\xi_1 = \xi_2 = \cdots \xi_n = \cdots = 0,$$

即 $x = \theta$，算子方程 $(\lambda I - T)x = 0$ 仅有零解。这说明 $\sigma(T)$ 中没有特

征值。下证当 $|\lambda| \leqslant 1$ 时，均有 $\lambda \in \sigma(T)$。

若 $\lambda = 0$，由算子 T 的定义，显然 $(\lambda I - T)(X) \neq X$，即 $\lambda \in \sigma(T)$。

若 $0 < |\lambda| \leqslant 1$。来证 $R(\lambda I - T) \neq X$。取 $y = \{1,0,0,\cdots\}$，则 $y \in X$。若存在 $x = \{\xi_n\}$ 满足

$$y = (\lambda I - T)x = \{\lambda\xi_1, \lambda\xi_2 + \xi_1, \cdots, \lambda\xi_n + \lambda\xi_{n-1}, \cdots\}$$

于是得

$$\lambda\xi_1 = 1, \lambda\xi_2 + \xi_1 = 0, \cdots, \lambda\xi_n + \xi_{n-1} = 0, \cdots$$

解得 $\xi_1 = \frac{1}{\lambda}, \xi_2 = -\frac{1}{\lambda^2}, \cdots, \xi_n = (-1)^{n-1}\frac{1}{\lambda^n}, \cdots$。但因

$$\sum_{n=1}^{\infty} |\xi_n| = \sum_{n=1}^{\infty} \frac{1}{|\lambda|^n} = +\infty (0 < |\lambda| \leqslant 1),$$

所以级数 $\sum_{n=1}^{\infty} |\xi_n|$ 不收敛，即 $x \notin X$。这说明不存在 $x \in X$ 使 $y = (\lambda I - T)x$，故 $R(\lambda I - T) \neq X$。于是

$$\sigma(T) = \{\lambda \in C: |\lambda| \leqslant 1\}。$$

3.6.2 紧算子的 Riesz – Schauder 定理

【定义 2】 设 X, Y 都是赋范线性空间，线性算子 $T: X \to Y$ 称为全连续算子(或紧算子)，是指 T 将 X 中有界集映成 Y 中列紧集。

注 全连续算子一定是有界线性算子。但一般逆不真，例如当 X 为无限维赋范线性空间时，则恒等算子 $I: X \to X$ 不是全连续算子。

关于全连续算子的谱理论十分简单，我们有下面著名的 Riesz – Schauder 定理。

【定理 3】 设 X 是复 Banach 空间，$T \in L(X, X)$，则

(1) $\sigma(T)$ 或是有限集，或是以 0 为聚点的可列集，且非零谱点均是特征值。

(2)$\sigma(T)=\sigma(T^*)$，这里 T^* 是 T 的共轭算子。

(3) 设 $\lambda,\mu(\lambda\neq\mu)$ 分别是 T,T^* 的特征值，则对应的特征子空间 L_λ 及 L_μ^* 直交，即若 $x\in L_\lambda$，$x^*\in L_\mu^*$，则 $x^*(x)=0$。

(4) 设 $\lambda\neq 0$ 是 T 与 T^* 的特征值，则相应的特征子空间 L_λ 及 L_λ^* 都是维数相同的有限维空间，且方程 $(\lambda I-T)x=y$ 有解的充要条件是 y 与 $\mathrm{Ker}(\lambda I^*-T^*)$ 直交；同样，方程 $(\lambda I^*-T^*)=y^*$ 有解的充要条件是 y^* 与 $\mathrm{Ker}(\lambda I-T)$ 直交。

习题 3.6

1. 设 $T_1,T_2\in \mathrm{L}(X,X)$ 且 T_1^{-1},T_2^{-1} 存在，$T_1^{-1},T_2^{-1}\in \mathrm{L}(X,X)$，证明：$(T_1T_2)^{-1}$ 存在且 $(T_1T_2)^{-1}\in \mathrm{L}(X,X)$，更有

$$(T_1T_2)^{-1}=T_2^{-1}T_1^{-1}。$$

2. 设 $T\in \mathrm{L}(X,X)$，$S\in \mathrm{L}(X,X)$，且 S 是全连续算子，证明：TS 与 ST 均是全连续算子。

3. 设 $T_n\in \mathrm{L}(X,X)$ 是全连续算子，$T\in \mathrm{L}(X,X)$ 且 $\|T_n-T\|\to 0(n\to\infty)$。证明：$T$ 也是全连续算子。

4. 设 $T\in \mathrm{L}(X,X)$，若 TX 是有限维子空间时，证明 T 是全连续算子。

5. 设 M 是赋范线性空间 X 的一个子空间，$T:X\to X$ 是有界线性算子，若 $TM\subset M$，则称 M 为 T 的不变子空间。

(1) 证明 T 的特征子空间是不变子空间；

(2) 设 X 是 n 维赋范线性空间，$\{e_1,e_2,\cdots,e_n\}$ 是 X 的一个基，对于线性算子 $T:X\to X$，若 $M=\mathrm{span}\{e_1,e_2,\cdots,e_m\}$ 是 T 的不变子空间，求 T 关于基 $\{e_1,e_2,\cdots,e_n\}$ 的矩阵表示式。

第 4 章　内积空间

在前一章中，我们把 n 维 Eulid 空间 $\mathbf{R}^n$ 中向量的模长推广到一般线性空间中去，得到了赋范线性空间的概念。但在 $\mathbf{R}^n$ 中可以通过两个向量的夹角讨论向量与方向有关的问题。这对仅有模长概念的赋范线性空间是做不到的。我们知道，$\mathbf{R}^n$ 中向量的夹角是通过向量的内积描述的，因此在本章中我们引入了一般的内积空间的概念。

4.1　内积空间的基本概念

首先回忆几何空间 $\mathbf{R}^3$ 中向量的内积的概念。设 $x = (t_1, t_2, t_3)$，$y = (s_1, s_2, s_3) \in \mathbf{R}^3$，设 x 与 y 夹角为 ϕ，由解析几何知识可得

$$\cos\phi = \frac{t_1 s_1 + t_2 s_2 + t_3 s_3}{\| x \| \cdot \| y \|},$$

其中 $\| x \| = (\sum\limits_{k=1}^{3} | t_k |^2)^{\frac{1}{2}}$，$\| y \| = (\sum\limits_{k=1}^{3} | s_k |^2)^{\frac{1}{2}}$。

令 $\langle x, y \rangle = \sum\limits_{k=1}^{3} t_k s_k$，称为 x 与 y 的内积，不难证明它有如下性质：

(1) $\langle x, y \rangle \geqslant 0, \forall x \in \mathbf{R}^3$，且 $\langle x, x \rangle = 0 \Leftrightarrow x = \theta$；

(2) $\langle x, y \rangle = < y, x >, \forall x, y \in \mathbf{R}^3$；

(3) $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle, \forall x_1, x_2, y \in \mathbf{R}^3$；

(4) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle, \forall \lambda \in \mathbf{R}, x \in \mathbf{R}^3$。

注　由定义可得 $\| x \| = \sqrt{\langle x, x \rangle}$，我们看到，两个向量的夹角

仅与向量的内积有关。利用内积我们可以讨论如向量的直交及投影等重要几何问题。

现在我们引入一般的内积空间的概念。

【定义 1】 设 X 为数域 F 上线性空间,若对任两个元素(也称为向量)$x,y\in X$,有惟一 F 中数与之对应,记为 $\langle x,y\rangle$,并且满足如下性质:

(1)$\langle x,x\rangle\geqslant 0,\forall x\in X$,且 $\langle x,x\rangle=0\Leftrightarrow x=\theta$;

(2)$\langle x,y\rangle=\overline{\langle y,x\rangle},\forall x,y\in X$;

(3)$\langle x_1+x_2,y\rangle=\langle x_1,y\rangle+\langle x_2,y\rangle,\forall x_1,x_2,y\in X$;

(4)$\langle\lambda x,y\rangle=\lambda\langle x,y\rangle,\forall\lambda\in F,x,y\in X$。

则称 $\langle x,y\rangle$ 为 x 与 y 的内积,有了内积的线性空间叫做内积空间,当 F 为实数域 $\mathbf{R}$(或复数域 $\mathbf{C}$)时,叫 X 为实(或复)内积空间。

注 由性质(3)与(4)知,内积运算关于第一个变元是线性的。由性质(2)与(4)可推知 $\langle x,\lambda y\rangle=\bar{\lambda}\langle x,y\rangle$。于是当 X 为实内积空间时,内积关于第二个变元也是线性的。而常称 $\langle x,\lambda y\rangle=\bar{\lambda}\langle x,y\rangle$ 为共轭齐次性,因此在 X 为复内积空间时,内积是共轭线性的。

今后讨论中不加注明时,恒设 X 为复内积空间。

引理(Schwarz 不等式) 设 X 为内积空间,对任意 $x,y\in X$,成立不等式:

$$|\langle x,y\rangle|\leqslant\sqrt{\langle x,x\rangle}\cdot\sqrt{\langle y,y\rangle}。$$

证 若 $y=\theta$,则任 $x\in X$,有 $\langle x,\theta\rangle=0$,则显然不等式成立。现在设 $y\neq\theta$,则 $\forall\lambda\in F$,有

$$0\leqslant\langle x+\lambda y,x+\lambda y\rangle=\langle x,x\rangle+\bar{\lambda}\langle x,y\rangle+\lambda\langle y,x\rangle+|\lambda|^2\langle y,y\rangle。$$

取 $\lambda=-\dfrac{\langle x,y\rangle}{\langle y,y\rangle}$ 代入上式可得 $\langle x,x\rangle-\dfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle}\geqslant 0$,由此立得 $\langle x,y\rangle\leqslant\sqrt{\langle x,x\rangle}\cdot\sqrt{\langle y,y\rangle}$。证毕。

【定理 1】 设 X 为内积空间,任 $x \in X$,令 $\|x\| = \sqrt{\langle x,x\rangle}$,则 $\|x\|$ 是 x 的范数。

证 因范数的前两条性质可直接由内积的性质推出,我们仅验证它满足第三条性质(即三角不等式)。事实上,

$$\|x+y\|^2 = \langle x+y,x+y\rangle = \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

故有 $\|x+y\| \leqslant \|x\| + \|y\|$。证毕。

注 常称 $\|x\| = \sqrt{\langle x,x\rangle}$ 为由内积导出的范数,于是内积空间按此范数成为一个赋范线性空间。在此意义下,第二章关于赋范线性空间的有关内容都适用于内积空间。特别当内积空间 X 按由内积导出的范数完备时,称 X 为 Hilbert 空间。

以下介绍几个常用的 Hilbert 空间的例子。

例 1 F^n 表示(实或复)Euclid 空间,对于 $x = (t_1,t_2,\cdots,t_n)$, $y = (s_1,s_2,\cdots,s_n) \in F^n$,类似于几何空间 $\mathbf{R}^3$ 中向量的内积定义,令

$$\langle x,y\rangle = \sum_{k=1}^{n} t_n \cdot \overline{s_n},$$

不难验证 F^n 成为一个 Hilbert 空间。

例 2 $l^2 = \{x = (t_1,t_2,\cdots): \sum\limits_{k=1}^{\infty} |t_n|^2 < \infty, t_n \in F, n = 1,2,\cdots\}$,当 $x = \{t_1,t_2,\cdots\}$, $y = \{s_1,s_2,\cdots\} \in l^2$ 时,令

$$\langle x,y\rangle = \sum_{k=1}^{\infty} t_n \cdot \overline{s_n},$$

容易验证 l^2 成为内积空间。以下证明 l^2 为 Hilbert 空间。任取 Cauchy 列 $x_n = (t_1^{(n)},t_2^{(n)},\cdots) \in l^2$,则对任 $\varepsilon > 0$, $\exists N$,当 $n,m \geqslant N$ 时,有

$$\|x_n - x_m\| = (\sum_{k=1}^{\infty} |t_k^{(n)} - t_k^{(m)}|^2)^{\frac{1}{2}} < \varepsilon。$$

因而有

$$|t_k^{(n)} - t_k^{(m)}| < \varepsilon, k = 1,2,\cdots,$$

故数列$\{t_k^{(n)}\}_{n=1}^{\infty} \subset F$ 是 Cauchy 列，因数域 F 完备，则存在 $s_k \in F(k=1,2,\cdots)$，使$\lim\limits_{n\to\infty} t_k^{(n)} = s_k$，令 $x = (s_1, s_2, \cdots)$，则任 $k = 1,2,\cdots\cdots$，当 $n, m \geqslant N$ 时，有

$$\sum_{i=1}^{k} |t_i^{(n)} - t_i^{(m)}|^2 \leqslant \| x_n - x_m \|^2 < \varepsilon^2,$$

则令 $m \to \infty$，对每个 $n \geqslant N$ 及任 $k = 1,2,\cdots$ 有

$$\sum_{i=1}^{k} |t_i^{(n)} - s_i|^2 \leqslant \varepsilon^2。$$

因而亦有

$$\left(\sum_{i=1}^{\infty} |t_i^{(n)} - s_i|^2\right)^{\frac{1}{2}} \leqslant \varepsilon, \text{只要 } n \geqslant N,$$

所以 $x_n - x \in l^2$，注意 l^2 是线性空间，则 $x = (x - x_n) + x_n \in l^2$，且 $\| x_n - x \| \leqslant \varepsilon, n \geqslant N$，这即表明 x_n 在l^2 中收敛，故 l^2 为 Hilbert 空间。

例 3 $L^2(E)$，E 为有限或无穷区间，对任 $x, \in L^2(E)$，定义内积

$$\langle x, y\rangle = \int_E x(t)\overline{y(t)}\mathrm{d}t。$$

这里 $L^2(E)$ 中元素是实值或复值二次可积函数，也不难验证 $L^2(E)$ 是内积空间。现在证明 $L^2(E)$ 是 Hilbert 空间。设 $x_n \in L^2(E)$ 为 Cauchy 列，则对每个 $k = 1,2,\cdots\cdots$，存在自然数 n_k，有

$$\| x_{n_k} - x_{n_{k+1}} \| \leqslant \frac{1}{2^k}, k = 1,2,\cdots。$$

对任有限区间 $e \subset E, me < \infty$，由 Hölder 不等式，有

$$\int_e |x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t \leqslant \left(\int_e |x_{n_k}(t)-x_{n_{k+1}}(t)|^2\mathrm{d}t\right)^{\frac{1}{2}}\cdot\left(\int_e 1^2\mathrm{d}t\right)^{\frac{1}{2}} =$$

$$(me)^{\frac{1}{2}}\|x_{n_k}-x_{n_{k+1}}\| \leqslant \frac{1}{2^k}\sqrt{me},k=1,2,\cdots,$$其中 me 为 e 的长度。

故级数 $\sum_{k=1}^{\infty}\int_e |x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t$ 收敛，于是由 Levi 引理(见第一章 1.4 节)关于 Lebesgue 积分的性质(6)我们有

$$\infty > \sum_{k=1}^{\infty}\int_e |x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t = \lim_{n\to\infty}\sum_{k=1}^{n}\int_e |x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t =$$

$$\lim_{n\to\infty}\int_e \sum_{k=1}^{n}|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t =$$

$$\int_e \lim_{n\to\infty}\sum_{k=1}^{n}|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t =$$

$$\int_e \sum_{k=1}^{\infty}|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t,$$

从而知 $\sum_{k=1}^{\infty}|x_{n_k}(t)-x_{n_{k+1}}(t)|$ 是集 e 上可积函数，则必在 e 上为几乎处处有限函数，即级数在 e 上几乎处处收敛，而 e 为 E 中任意有限区间，则级数 $\sum_{k=1}^{\infty}|x_{n_k}(t)-x_{n_{k+1}}(t)|$ 在 E 上几乎处处收敛，因而级数 $x_{n_1}(t)+(x_{n_2}(t)-x_{n_1}(t))+(x_{n_3}(t)-x_{n_2}(t))+\cdots$ 在 E 上也几乎处处收敛，亦即函数列 $x_{n_k}(t)$ 在 E 上几乎处处收敛于函数 $x(t)$。

现在证明 $x\in L^2(E)$，且 $\lim_{n\to\infty}\|x_n-x\|=0$。

对任意 $\epsilon>0$，因 x_n 为 $L^2(E)$ 中 Cauchy 列，则存在 N，当 $n,n_k>N$ 时，有 $\|x_n-x_{n_k}\|<\epsilon$，即

$$\int_E |x_n(t)-x_{n_k}(t)|^2\mathrm{d}t < \epsilon^2。$$

令 $k\to\infty$，利用第一章 1.4 节 Lebesgue 积分的性质(7)，得到

$$\int_E |x_n(t)-x(t)|^2\mathrm{d}t<\varepsilon^2\quad(\text{当 } n>N \text{ 时}),$$

即 $\|x_n-x\|\leqslant\varepsilon$，且 $x_n-x\in L^2(E)$，于是 $x=x_n-(x_n-x)\in L^2(E)$。于是 Cauchy 列 x_n 在 $L^2(E)$ 中收敛，故 $L^2(E)$ 是 Hilbert 空间。

本节最后介绍关于内积空间几条常用的性质。

(1) 内积的连续性。设 $\lim\limits_{n\to\infty}x_n=x$，$\lim\limits_{n\to\infty}y_n=y$，则有

$$\lim_{n\to\infty}\langle x_n,y_n\rangle=\langle\lim_{n\to\infty}x_n,\lim_{n\to\infty}y_n\rangle=\langle x,y\rangle。$$

证　由 Schwarz 不等式，得

$$|\langle x_n,y_n\rangle-\langle x,y\rangle|\leqslant|\langle x_n-x,y_n\rangle|+|\langle x,y_n-y\rangle|\leqslant \|y_n\|\,\|x_n-x\|+\|x\|\,\|y_n-y\|\to 0(n\to\infty),$$

因收敛列 y_n 有界。证毕。

(2) 极化恒等式。对内积空间 X 中元素 x 与 y，成立

$$\langle x,y\rangle=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2+i\|x+iy\|^2-i\|x-iy\|^2)。$$

证明可直接运用范数的定义和内积的性质得到。留给读者作为练习。

注　当 X 为实内积空间时，则极化恒等式为

$$\langle x,y\rangle=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2)。$$

(3) 中线公式。对内积空间 X 中元素 x 与 y，成立

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2)。$$

证

$$\begin{aligned}\|x+y\|^2+\|x-y\|^2&=\langle x+y,x+y\rangle+\langle x-y,x-y\rangle=\\&\langle x,x\rangle+\langle y,x\rangle+\langle x,y\rangle+\langle y,y\rangle+\langle x,x\rangle-\\&\langle y,x\rangle-\langle x,y\rangle+\langle y,y\rangle=2(\|x\|^2+\|y\|^2)。\end{aligned}$$

证毕。

注 也常称中线公式为平行四边形公式。因在平面 $\mathbf{R}^2$ 中,平行四边形的对角线长度的平方和等于四条边的长度平方和。另外,可以证明中线公式是内积空间中由内积导出的范数的特征性质,即当 X 为赋范线性空间时,若对其任何元素 x 与 y 关于范数成立中线公式,则必在 X 中可定义内积 $\langle x,y\rangle$,使范数可由此内积导出(此事实有兴趣读者可在泛函分析的教科书中查到)。也就是一个赋范线性空间成为内积空间的条件是其范数要满足中线公式。因此,内积空间是一类特殊的赋范线性空间。

例如当 $p\geqslant 1$ 且 $p\neq 2$ 时,l^p 不是内积空间。因为取 $x=(1,1,,0,0\cdots)$,$y=(1,-1,0,0,\cdots)\in l^p$,则 $\|x\|=\|y\|=2^{1/p}$ 且 $\|x+y\|=\|x-y\|=2$,显然不满足中线公式。

又例如 $C[a,b]$,按范数 $\|x\|=\max\limits_{a\leqslant t\leqslant b}|x(t)|$ 不是内积空间。这只要取 $x(t)=1,\forall t\in[a,b]$ 及 $y(t)=\dfrac{t-a}{b-a},\forall t\in[a,b]$,则 $\|x\|=\|y\|=1$,且 $\|x+y\|=2$,$\|x-y\|=1$,明显不满足中线公式。

再例如 $L^p[a,b]$ 当 $p\geqslant 1$ 且 $p\neq 2$ 时,也不是一个内积空间。

习题 4.1

1. 证明:Schwarz 不等式中等号成立 $\Leftrightarrow x$ 与 y 线性相关。

2. 设 X 为实内积空间,$x,y\in X$ 若 $\|x\|=\|y\|$,证明:$\langle x+y,x-y\rangle=0$。若 $X=\mathbf{R}^2$,所证事实有何几何意义。

3. 设 X 为内积空间,$u,v\in X$,若对任何 $x\in X$ 有 $\langle x,u\rangle=\langle x,v\rangle$,试证明 $u=v$。

4. 设 X 为 Hilbert 空间，$x_n, x \in X$，求证 $x_n \to x(n \to \infty)$ 的充要条件是 $\|x_n\| \to \|x\|$，且 $\langle x_n, x\rangle \to \langle x, x\rangle (n \to \infty)$。

5. 验证极化恒等式。

6. 设 $\{e_1, e_2, \cdots, e_n\}$ 是 n 维线性空间 X 的一组基，对于 $x, y \in X$ 有惟一表示 $x = \sum_{k=1}^{n} t_k e_k, y = \sum_{k=1}^{n} s_k e_k$，其中 $t_k, s_k \in F, k = 1, 2, \cdots, n$。求证 $\langle x, y\rangle$ 是 X 上一个内积的充要条件是存在正定矩阵 $A = (a_{ij})_{n\times n}$，使成立

$$\left\langle \sum_{i=1}^{n} t_i e_i, \sum_{j=1}^{n} s_j e_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} t_i \bar{s}_j。$$

4.2　内积空间中元素的直交与直交分解

4.2.1　直交及其性质

仿照 $\mathbf{R}^2$ 中两个向量的直交概念，我们有：

【定义 1】　设 X 是内积空间，$x, y \in X$，若 $\langle x, y\rangle = 0$，则称 x 与 y 直交，记为 $x \perp y$。设 $x \in X, M \subset X$，若 x 与 M 中每个元素直交时，则称 x 与 M 直交，记为 $x \perp M$。又 $N \subset X$，若 $x \in M, y \in N$，都有 $x \perp y$，则称 M 与 N 直交，记为 $M \perp N$。设 $M \subset X$，记 $M^{\perp} = \{x \in X: x \perp M\}$，则称 $M^{\perp}$ 为 M 的直交补。

由以上定义，立得如下简明事实：

(1) 零元素 θ 与 X 中每个元素 x 直交。

(2) 若 $x \perp y$，则 $y \perp x$。

(3) $x \perp X \Leftrightarrow x = \theta$。

(4) 若 $M \subset N \subset X$，则 $N^{\perp} \subset M^{\perp}$。

(5) 任 $M \subset X$,若 $\theta \notin M$,则 $M \cap M^{\perp} = \varnothing$;若 $\theta \in M$,则 $M \cap M^{\perp} = \{\theta\}$。

此外我们还有以下几条有用性质:

(6) 若 $x_n \to x(n \to \infty)$ 且 $x_n \perp y$,则 $x \perp y$。

这因 $\langle x, y\rangle = \lim\limits_{n\to\infty}\langle x_n, y\rangle = 0$。

(7) 若 $x, y \in X$,有 $x \perp y$,则成立勾股公式 $\|x + y\|^2 = \|x\|^2 + \|y\|^2$。

这留给读者验证。

(8) 对任 $M \subset X$,则 $M^{\perp}$ 是 X 的闭子空间。

事实上,任 $x_1, x_2 \in M^{\perp}$,则对每个 $y \in M$,有 $x_1 \perp y, x_2 \perp y$,于是有 $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle = 0$,故 $x_1 + x_2 \in M^{\perp}$;又任 $x \in M^{\perp}, \lambda \in F$,则任 $y \in M$ 有 $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle = 0$,故 $\lambda x \in M^{\perp}$,因此 $M^{\perp}$ 成为 X 的线性子空间。现在证明 $M^{\perp}$ 是闭集。若 $(M^{\perp})' = \varnothing$,则 $M^{\perp}$ 为闭集,当 $(M^{\perp})' \neq \varnothing$,任取 $x \in (M^{\perp})'$,则存在 $x_n \in M^{\perp}$,有 $\lim\limits_{n\to\infty} x_n = x$。对任 $y \in M$,应用性质(6),有

$$\langle x, y\rangle = \lim_{n\to\infty}\langle x_n, y\rangle = 0$$

则 $x \perp y$,于是推得 $x \perp M$,即 $x \in M^{\perp}$,因此 $M^{\perp}$ 为闭集。证毕。

(9) 设 $M \subset X$ 为非空集,则 $\overline{(\mathrm{span} M)}^{\perp} = M^{\perp}$。

事实上,因 $M \subset \overline{(\mathrm{span} M)}$,则 $\overline{(\mathrm{span} M)}^{\perp} \subset M^{\perp}$。另外,对任 $x \in M^{\perp}$,任意取 $y \in \overline{(\mathrm{span} M)} = (\mathrm{span} M) \cup (\mathrm{span} M)'$,若 $y \in (\mathrm{span} M)$,则 y 是 M 中有限个元素 $x_1, x_2, \cdots, x_n$ 的线性组合,即 $y = \sum\limits_{i=1}^{n} \lambda_i x_i$ $(\lambda_i \in F, 1 \leqslant i \leqslant n)$ 于是 $\langle y, x\rangle = \sum\limits_{i=1}^{n} \lambda_i \langle x_i, x\rangle = 0$,即 $x \perp y$。而当 $y \in (\mathrm{span} M)'$,则存在元素列 $y_n \in \mathrm{span} M$,有 $\lim\limits_{n\to\infty} y_n = y$,由以上证明

知 $y_n \perp x$，于是由性质(6)得知 $y \perp x$。综上所述，$x \in \overline{(\mathrm{span}M)}^{\perp}$，故 $M^{\perp} \subset \overline{(\mathrm{span}M)}^{\perp}$。证毕。

4.2.2　直交投影及变分引理

仿照 $\mathbf{R}^2$ 中向量在坐标轴上投影的概念引入以下定义。

【定义 2】　设 M 是内积空间 X 的一个线性子空间，$x \in X$，若存在 $x_0 \in M$，$z \in M^{\perp}$ 使成立 $x = x_0 + z$，则称 x_0 为 x 在 M 上的直交投影(可简称为投影)。

注　一般情况，某个元素 x 在 X 的某子空间 M 上不一定存在投影。但当投影存在时，则可证明投影的惟一性。因为若 x_0 及 x_1 都是 x 在 M 上的投影，则由定义有 $z = x - x_0$，$z_1 = x - x_1 \in M^{\perp}$，于是 $x_0 - x_1 = z_1 - z \in M \cap M^{\perp} = \{\theta\}$，故 $x_0 = x_1$。

对于 $\mathbf{R}^2$，任向量 $x = (t_1, t_2)$ 在 x 轴(即子空间 $M = \{(t, 0): t \in \mathbf{R}\}$)上有投影为 $x_0 = (t_1, 0)$，并且知道点 $x = (t_1, t_2)$ 到 x 轴上每个点的距离最小者为 $\| x - x_0 \| = | t_2 |$。这种现象如何在一般的(特别是无限维)内积空间中表现是个需要探讨的问题。为此，我们首先给出重要概念。

【定义 3】　设 X 是度量空间，M 是 X 中非空子集，$x \in X$，则称 $\inf\limits_{y \in M} \rho(x, y)$ 为 x 到集 M 的距离，记为 $\rho(x, M)$。若存在某 $x_0 \in M$，使 $\rho(x, x_0) = \rho(x, M)$，则称 x_0 为 x 在 M 中最佳逼近元。

注　一般情况下，某元 $x \in X$，在某集 $M \subset X$ 中不一定存在最佳逼近元。并且在最佳逼近元存在时也不一定惟一。因此，最佳逼近元的存在性及惟一性成为逼近理论中一个主要研究方向之一。

在此我们仅介绍一个在微分方程、现代控制论等学科都有重要应用的基本结果。

【定理 1】(极小化向量定理)　设 M 是 Hilbert 空间 X 中的凸闭

集,则任意 $x \in X$,必在 M 中惟一存在最佳逼近元。

证　令 $d = \inf\limits_{y \in M} \| x - y \| = \rho(x, M)$,则存在 $x_n \in M$,使 $\lim\limits_{n \to \infty} \| x_n - x \| = d$。因 M 是凸集,则 $\frac{1}{2}(x_n + x_m) \in M$,于是必有

$$\left\| x - \frac{1}{2}(x_n + x_m) \right\| \geq d。$$

在中线公式中以 $x_n - x$ 代换 x,以 $x - x_m$ 代换 y,则有

$$\begin{aligned}
\| x_n - x_m \|^2 = & \| (x_n - x) + (x - x_m) \|^2 = \\
& 2\| x_n - x \|^2 + 2\| x - x_m \|^2 - \| x_n + x_m - 2x \|^2 = \\
& 2\| x_n - x \|^2 + 2\| x_m - x \|^2 - \\
& 4\left\| x - \frac{1}{2}(x_n + x_m) \right\|^2 \leq \\
& 2\| x_n - x \|^2 + 2\| x_m - x \|^2 - 4d^2 \to 0 (n, m \to \infty)。
\end{aligned}$$

因此 x_n 是完备内积空间 X 中 Cauchy 列,则存在 $x_0 \in X$,使 $\lim\limits_{n \to \infty} x_n = x_0$。因 M 是闭集,则 $x_0 \in M$,并且有

$$d = \lim_{n \to \infty} \| x_n - x \| = \| x_0 - x \|,$$

这证明了最佳逼近元的存在性。

现在证明惟一性。设 $y_0 \in M$ 也是 x 的最佳逼近元。还由中线公式得

$$\begin{aligned}
0 \leq \| x_0 - y_0 \|^2 = & \| (x_0 - x) + (x - y_0) \|^2 = \\
& 2\| x_0 - x \|^2 + 2\| x - y_0 \|^2 - 4\left\| x - \frac{x_0 + y_0}{2} \right\| \leq \\
& 2d^2 + 2d^2 - 4d^2 = 0,
\end{aligned}$$

故 $\| x_0 - y_0 \| = 0$,即 $x_0 = y_0$。证毕。

也常称此定理为变分引理。由于子空间一定是凸集,并注意定理的证明过程,则定理条件改为 M 是内积空间 X 中完备的子空间时,

定理结论仍成立。

4.2.3 投影定理

【定理 2】(投影定理) 设 M 是内积空间 X 的完备线性子空间，则对任意 $x \in X$，必在 M 上惟一存在投影。即必惟一存在 $x_0 \in M$，$z \in M^{\perp}$，使 $x = x_0 + z$。

证 由设，依据极小化向量定理，x 在 M 中惟一存在最佳逼近元 x_0，记 $d = \| x - x_0 \| = \inf\limits_{y \in M} \| x - y \|$。任取复数 λ，$y \in M$，则 $x_0 + \lambda y \in M$，且有

$$d^2 \leqslant \| x - (x_0 + \lambda y) \|^2 = \langle (x - x_0) - \lambda y, (x - x_0) - \lambda y \rangle =$$
$$\| x - x_0 \|^2 - \bar{\lambda} \langle x - x_0, y \rangle - \lambda \langle y, x - x_0 \rangle + | \lambda |^2 \| y \|^2。$$

当 $y \neq \theta$ 时，取 $\lambda = \dfrac{\langle x - x_0, y \rangle}{\| y \|^2}$ 代入上式，得

$$d^2 \leqslant d^2 - \frac{| (x - x_0, y) |^2}{\| y \|^2}。$$

于是推得 $\langle x - x_0, y \rangle = 0$，再注意 $y = \theta$，此式也成立，因而 $x - x_0 \in M^{\perp}$。令 $z = x - x_0$，即有 $x = x_0 + z$。投影的存在性得证。

投影的惟一性已由定义 2 的注得证。证毕。

注 (1) X 为 Hilbert 空间时，则对任闭子空间 $M \subset X$ 投影定理成立。

(2) 表达式 $x = x_0 + z$ 也常称为元素 x 的直交分解，故投影定理也叫做直交分解定理，是 $\mathbf{R}^2$ 中向量的直交分解的推广。由于在一般赋范线性空间中没有直交概念，因此不能讨论直交分解的问题。

(3) 对于内积空间 X 及子空间 M，在投影定理条件下有

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

即 X 表示为两个直交子空间的直和，常叫 X 为 M 与 $M^{\perp}$ 的直交和，或

直交分解。

投影定理在内积空间理论中是极为重要的基本定理。由于投影 $x_0 \in M$，就是元素 $x \in X$ 在子空间 M 中的最佳逼近元，因此在现代逼近论、概率论以及控制论中许多问题都可以抽象为如下的数学问题。

设 X 是内积空间，$x, x_1, x_2, \cdots, x_n \in X$，问是否存在 n 个数 $\lambda_1, \lambda_2, \cdots, \lambda_n$，使得 $\left\| x - \sum_{i=1}^{n} \lambda_i x_i \right\| = \inf_{y \in M} \| x - y \|$，其中 $M = \operatorname{span}\{x_1, x_2, \cdots, x_n\}$。并且一般假设 $x_1, x_2, \cdots, x_n$ 线性无关。

由于 M 是一个 n 维赋范线性空间，故 M 完备，则由投影定理，对于 $x \in X$，必惟一存在 $x_0 = \sum_{i=1}^{n} \lambda_i x_i \in M$，使

$$\| x - x_0 \| = \inf_{y \in M} \| x - y \|。$$

现在我们给出求解 x_0 的方法，因 $x_i \in M, 1 \leqslant i \leqslant n$，则由投影定理，我们有

$$\langle x - \sum_{i=1}^{n} \lambda_i x_i, x_k \rangle = 0, 1 \leqslant k \leqslant n,$$

即得线性方程组

$$\sum_{i=1}^{n} \lambda_i \langle x_i, x_k \rangle = \langle x, x_k \rangle, 1 \leqslant k \leqslant n。$$

记其系数行列式为 Δ_n。因为方程组已知有惟一解，则由 Gram 法则，有 $\Delta_n \neq 0$。并且依据 Gram 法则可计算出 $\lambda_i, 1 \leqslant i \leqslant n$。

最后，我们再给出投影定理的两个推论。

【推论 1】 设 M 是 Hilbert 空间 X 的真闭线性子空间，则 $M^{\perp}$ 中必有非零元素。

证 由设 $M \neq X$，则存在 $x \in X \setminus M$。由投影定理，存在 $x_0 \in$

$M, z \in M^{\perp}$,使 $x = x_0 + z$,于是必 $z \neq \theta$,否则 $x = x_0 \in M$,矛盾。证毕。

【推论2】 设 M 是 Hilbert 空间 X 的线性子空间,则 $\overline{M} = (M^{\perp})^{\perp}$。特别当 $M^{\perp} = \{\theta\}$ 时,则 M 在 X 中稠密。

证 由性质(8),$(M^{\perp})^{\perp}$ 是 X 中闭线性子空间,因 X 完备,则 $(M^{\perp})^{\perp}$ 完备。显然,有 $M \subset (M^{\perp})^{\perp}$,于是 $\overline{M} \subset (M^{\perp})^{\perp}$。同样得知 $\overline{M}$ 也完备。如果 $\overline{M} \neq (M^{\perp})^{\perp}$,于是关于 $\overline{M} \subset (M^{\perp})^{\perp}$,应用推论1,存在非零元素 $x \in (M^{\perp})^{\perp}$,且 $x \in \overline{M}^{\perp} = M^{\perp}$,故 $< x, x > = 0$,从而 $x = \theta$,矛盾。从而必 $\overline{M} = (M^{\perp})^{\perp}$,证毕。

习题4.2

1. 设 X 为实内积空间,若 $\|x + y\|^2 = \|x\|^2 + \|y\|^2$,则 $x \perp y$。问 X 为复内积空间时,结论是否成立?

2. 证明内积空间 X 中两个元素 x, y 直交的充要条件是对任意数 $\lambda \in F$,成立 $\|x + \lambda y\| \geq \|x\|$。

3. 设 $x_1, x_2, \cdots, x_n$ 是内积空间 X 中两两直交的非零元素组,求证 $x_1, x_2, \cdots, x_n$ 线性无关。

4. 设 X 为内积空间,$x, y \in X$,则 $x \perp y \Leftrightarrow$ 对任 $\lambda \in F$,有 $\|x + \lambda y\| = \|x - \lambda y\|$。

5. 设 X 是 Hilbert 空间,M 是 X 的子集,求证 $(M^{\perp})^{\perp}$ 是包含 M 的最小闭子空间。

6. 设 M 是 Hilbert 空间 X 中非空子集,求证 $\overline{\mathrm{span}M} = X \Leftrightarrow M^{\perp} = \{\theta\}$。

7. 设 M 为 Hilbert 空间 $L^2[-1,1]$ 中全体偶函数的集合。

(1) 求证 $M^{\perp}$ 是 $L^2[-1,1]$ 中全体奇函数。

(2) 任 $x \in L^2[-1,1]$，求 x 在 M 上的投影。

8. 设 X 为 Hilbert 空间，元素列 $x_n \in X$，两两直交，求证级数 $\sum_{n=1}^{\infty} x_n$ 收敛 $\Leftrightarrow$ 数值级数 $\sum_{k=1}^{\infty} \| x_n \|^2$ 收敛。

9. 证明直交的性质(1) ~ (5)。

10. 设 $x_1, x_2, \cdots, x_n$ 是内积空间 X 中两两直交元素组，求证

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \| x_k \|^2。$$

4.3 直交系

仿照 $\mathbf{R}^2$ 中情况，在内积空间引入直角坐标概念。

【定义 1】 设 M 是内积空间 X 中一个不含零元的子集，若 M 中任两个不同元素都直交，则称 M 为 X 的一个直交系。又若 M 中每个元素的范数都是 1，则称 M 为标准直交系。

注 为简单起见，我们仅讨论至多含可列个元素的直交系，因为对不可列情况在方法上同可列情况并无本质的区别。

例 1 在(实或复)Euclid 空间 F^n 中

$$e_1 = (1,0,0,\cdots,0), e_2 = (0,1,0,\cdots,0), \cdots, e_n = (0,0,\cdots,0,1)$$

是一个标准直交系。

例 2 对内积空间 l^2，以下元素列是一个标准直交系。

$e_n = (0,\cdots,1,0\cdots)$，其第 n 个分量是 1，其余分量都是 0，$n = 1, 2\cdots$。

例 3 在实内积空间 $L^2[0,2\pi]$ 中，若定义内积为

$$\langle x, y \rangle = \int_0^{2\pi} x(t) y(t) \mathrm{d}t,$$

则三角函数系

$$\frac{1}{\sqrt{2\pi}},\frac{1}{\sqrt{\pi}}\cos t,\frac{1}{\sqrt{\pi}}\sin t,\cdots\frac{1}{\sqrt{\pi}}\cos nt,\frac{1}{\sqrt{\pi}}\sin nt,\cdots$$

是 $L^2[0,2\pi]$ 的一个标准直交系。

【定义 2】 设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准直交系。对任 $x\in X$,称 $c_n=\langle x,e_n\rangle$ 为元素 x 关于 e_n 的 Fourier 系数,常简称为 x 的 Fourier 系数。于是有形式级数 $\sum\limits_{n=1}^{\infty}c_ne_n$,称为元素 x 关于 $\{e_n\}_{n=1}^{\infty}$ 的 Fourier 级数。当特别 $x=\sum\limits_{n=1}^{\infty}c_ne_n$ 成立时,就说 x 关于 $\{e_n\}_{n=1}^{\infty}$ 可以展开为 Fourier 级数。

注　一般情况下,Fourier 级数不一定收敛。即或收敛,也不一定收敛于 x。在何条件下元素 x 可以展开为 Fourier 级数的问题自然是重要的。

【定理 1】 设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准直交系,记 $X_n=\mathrm{span}\{e_1,e_2,\cdots,e_n\}$,任意给定 $x\in X$,则 x 在 X_n 上的投影是 $s_n=\sum\limits_{k=1}^{n}c_ke_k$,即 s_n 是 x 在 X_n 内的最佳逼近元。

证　因 $x=s_n+(x-s_n)$,由于 $s_n\in X_n$ 则只须证 $x-s_n\in X_n^{\perp}$。由 4.2 性质(9),又仅须证 $(x-s_n)\perp e_k,k=1,2,\cdots,n$。于是由 $\langle x-s_n,e_k\rangle=\langle x,e_k\rangle-\langle\sum\limits_{i=1}^{n}c_ie_i,e_k\rangle=0$,知结论成立。证毕。

注　任意 $\sum\limits_{k=1}^{n}\lambda_ke_k\in X_n$,任 $x\in X$,成立

$$\|x-s_n\|\leqslant\left\|x-\sum_{k=1}^{n}\lambda_ke_k\right\|。$$

【定理 2】(Bessel 不等式)　设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准

直交系，则任 $x \in X$，成立 Bessel 不等式

$$\sum_{n=1}^{\infty} |c_n|^2 \leqslant \|x\|^2。$$

其中，$c_n = \langle x, e_n \rangle, n = 1,2,\cdots$。

证 已知 $s_n \perp (x - s_n)$，其中 $s_n = \sum_{k=1}^{n} c_k e_k$，则由勾股公式得

$$\|x\|^2 = \|s_n\|^2 + \|x - s_n\|^2 \geqslant \|s_n\|^2 = \left\|\sum_{k=1}^{n} c_k e_k\right\|^2 =$$

$$\sum_{k=1}^{n} \|c_k e_k\|^2 = \sum_{k=1}^{n} |c_k|^2,$$

令 $n \to \infty$，即得 Bessel 不等式。证毕。

注 (1)Bessel 不等式指元素 x 在每个 e_n 上投影 $c_n e_n$ 的范数的平方和不大于 x 的范数；由此知 $\sum_{n=1}^{\infty} |c_n|^2$ 为收敛级数，于是推得事实

$$\lim_{n\to\infty} c_n = 0。$$

特别对内积空间 $L^2[0,2\pi]$ 关于标准直交系三角函数系（见例 3），对任意 $x \in L^2[0,2\pi]$，其 Fourier 系数为

$$c_1 = \int_0^{2\pi} x(t) \cdot \frac{1}{\sqrt{2\pi}} \mathrm{d}t = \frac{\sqrt{\pi}}{\sqrt{2}} \cdot \frac{1}{\pi}\int_0^{2\pi} x(t)\mathrm{d}t = \frac{\sqrt{\pi}}{\sqrt{2}} a_0,$$

$$c_{2n} = \int_0^{2\pi} x(t) \cdot \frac{1}{\sqrt{\pi}} \cos nt \mathrm{d}t = \sqrt{\pi} \cdot \frac{1}{\pi}\int_0^{2\pi} x(t)\cos nt \mathrm{d}t = \sqrt{\pi} a_n, n = 1,$$
$2,\cdots,$

$$c_{2n+1} = \int_0^{2\pi} x(t) \cdot \frac{1}{\sqrt{\pi}} \sin nt \mathrm{d}t = \sqrt{\pi} \cdot \frac{1}{\pi}\int_0^{2\pi} x(t)\sin nt \mathrm{d}t = \sqrt{\pi} b_n, n = 1,$$
$2,\cdots,$

其中，a_n, b_n 即通常的 Fourier 系数，则由 Bessel 不等式，得

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \leqslant \frac{1}{\pi}\int_0^{2\pi} |x(t)|^2 \mathrm{d}t。$$

注意这里用了收敛正项级数的可交换性。

(2) 在内积空间 X 给定标准直交系 $\{e_n\}_{n=1}^{\infty}$ 情况下，$x \in X$，其对应的 Fourier 系数构成一个序列 $c = (c_1, c_2, \cdots) \in l^2$，并确定了由 X 到内积空间 l^2 内的一个映射 T 为

$$Tx = (c_1, c_2, \cdots), \forall x \in X,$$

其中 $c_n = < x, e_n >, n = 1,2,\cdots$。不难证明 T 是线性映射。

反之，任意 l^2 中元素 $c = (c_1, c_2, \cdots)$，一般情况下，不一定存在 X 中元素 x，使 $c_n = \langle x, e_n \rangle, n = 1,2,\cdots$ 但在 X 完备时，有以下定理。

【定理3】 (Riesz – Fisher) 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 的一个标准直交系，则对每个元素 $c = (c_1, c_2, \cdots) \in l^2$，惟一存在 $x \in X$，使 $c_n = \langle x, e_n \rangle, n = 1,2,\cdots$，且成立等式

$$\sum_{n=1}^{\infty} |c_n|^2 = \|x\|^2。$$

证　令 $s_k = \sum_{k=1}^{n} c_k e_k$，因 $\|s_n - s_m\|^2 = \sum_{k=n+1}^{m} |c_k|^2$，由于级数 $\sum_{n=1}^{\infty} |c_n|^2$ 收敛，则根据 Cauchy 收敛准则，有

$$\lim_{\substack{n\to\infty \\ m\to\infty}} \|s_n - s_m\| = \lim_{\substack{n\to\infty \\ m\to\infty}} \Big(\sum_{k=n+1}^{\infty} |c_n|^2\Big)^{\frac{1}{2}} = 0,$$

故 s_n 是完备空间 X 中一个 Cauchy 列，则存在 $x \in X$，有

$$x = \lim_{n\to\infty} s_n = \sum_{n=1}^{\infty} c_n e_n。$$

现在设 k 为任意自然数，则

$$\langle x, e_k \rangle = \langle \lim_{n\to\infty} s_n, e_k \rangle = \lim_{n\to\infty} \langle \sum_{i=1}^{n} c_i e_i, e_k \rangle = c_k。$$

再注意 $\|s_n\|^2 = \sum_{k=1}^{n} |c_n|^2$，令 $n \to \infty$ 即得等式 $\|x\|^2 =$

$\sum_{k=1}^{\infty} |c_n|^2$。

最后证明惟一性。若 $y \in X$，也满足定理结论：

$$c_n = \langle y, e_n \rangle \text{ 且 } \|y\|^2 = \sum_{k=1}^{n} |c_n|^2。$$

则因 $\|y\|^2 = \|s_n\|^2 + \|y - s_n\|^2$（由定理 1），令 $n \to \infty$ 推得 $s_n \to y$。由极限的惟一性，必 $y = x$。证毕。

注 在 X 为 Hilbert 空间时，可确定一个由 l^2 到 X 内的映射，但在一般情况下，不能断定映射是映上的。因此不一定为由 l^2 到 X 上的一一映射。

在 n 维 Euclid 空间中，标准直交基（直角坐标系）的极大性是至关重要的，对此我们有如下推广。

【定义 3】 设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准直角交系，若对任意 $x \in X$，有 $\langle x, e_n \rangle = 0, n = 1,2,\cdots$，则必 $x = \theta$，我们就称 $\{e_n\}_{n=1}^{\infty}$ 是完全的。

如例 2 中的标准直交系是 l^2 中一个完全的标准直交系。

【定理 4】 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 中一个标准直交系，则以下的命题等价：

(1) $\{e_n\}_{n=1}^{\infty}$ 是完全的；

(2) 对任意 $x \in X$，成立 Parseval 等式

$$\|x\|^2 = \sum_{n=1}^{\infty} |c_n|^2, \text{其中 } c_n = \langle x, e_n \rangle, n = 1,2,\cdots;$$

(3) 对任意 $x \in X$，有 $x = \sum_{n=1}^{\infty} c_n e_n$，其中 $c_n = \langle x, e_n \rangle, n = 1, 2,\cdots$;

(4) 对任意两个元素 $x, y \in X$ 有

$$\langle x,y\rangle=\sum_{n=1}^{\infty}\langle x,e_n\rangle\cdot\overline{\langle y,e_n\rangle}$$

证　(1)⇒(2) 设$\{e_n\}_{n=1}^{\infty}$是完全的，对任意$x\in X$，记$c_n=\langle x,e_n\rangle$，则由定理2知$c=(c_1,c_2,\cdots)\in l^2$，再由定理3，惟一存在$y\in X$，使$c_n=\langle y,e_n\rangle$且成立$\|y\|^2=\sum_{n=1}^{\infty}|c_n|^2$。因$\langle x,e_n\rangle=\langle y,e_n\rangle$，$n=1,2,\cdots$，则$\langle x-y,e_n\rangle=0$，$n=1,2,\cdots$。由于$\{e_n\}_{n=1}^{\infty}$完全，于是必$x=y$，因此有$\|x\|^2=\sum_{n=1}^{\infty}|c_n|^2$。命题(2)成立。

(2)⇒(3) 现在假设命题(2)真，任意取$x\in X$，令$c_n=\langle x,e_n\rangle$，$n=1,2,\cdots$。$s_n=\sum_{k=1}^{n}c_ke_k$，则有

$$\|x-s_n\|^2=\|x\|^2-\sum_{k=1}^{n}|c_n|^2\to 0(n\to\infty),$$

即得$\sum_{k=1}^{n}c_ne_n=\lim_{n\to\infty}\sum_{k=1}^{n}c_ke_k=x$。于是命题(3)真。

(3)⇒(4) 这里假设命题(3)真。任取$x,y\in X$，令$c_n=\langle x,e_n\rangle$，$d_n=\langle y,e_n\rangle$，则有$x=\lim_{n\to\infty}\sum_{k=1}^{n}c_ne_n$，$y=\lim_{n\to\infty}\sum_{k=1}^{n}d_ne_n$。于是可得

$$\langle x,y\rangle=\lim_{n\to\infty}\left\langle\sum_{k=1}^{n}c_ke_k,\sum_{k=1}^{n}d_ke_k\right\rangle=\lim_{n\to\infty}\sum_{k=1}^{n}c_k\bar{d}_k=\sum_{n=1}^{\infty}c_n\bar{d}_n。$$

即命题(4)成立。

(4)⇒(1) 现在假设命题(4)成立。取$x\in X$，若$x\perp e_n$，$n=1,2,\cdots$。此时任取$y\in X$，有$\langle x,y\rangle=\sum_{n=1}^{\infty}\langle x,e_n\rangle\cdot\overline{\langle y,e_n\rangle}=0$，即$x\perp X$，故$x=\theta$，因此命题(1)真。定理证毕。

注　若Hilbert空间X存在可列的完全的标准直交系$\{e_n\}_{n=1}^{\infty}$，则任意$x\in X$，有$c_n=\langle x,e_n\rangle$，$n=1,2,\cdots$。映射$Tx=(c_1,c_2,\cdots)$

$\in l^2$ 是由 X 到 l^2 上的一个等距同构映射，故 X 与 l^2 等距同构。

以下的定理在判别某标准直交系的完全性时是经常有用的。

【定理5】 设 $\{e_n\}_{n=1}^{\infty}$ 是Hilbert空间 X 中一个标准直交系，如果Parseval等式在 X 中某稠密子集 D 上成立，则 $\{e_n\}_{n=1}^{\infty}$ 是完全的。

证　$X_0 = \overline{\mathrm{span}\{e_n\}_{n=1}^{\infty}}$，则 X_0 是 X 闭线性子空间。任 $x \in D$，令 $c_n = \langle x, e_n \rangle, n = 1,2,\cdots$ 由设成立 $\|x\|^2 = \sum_{n=1}^{\infty} |c_n|^2$，同定理4(2)$\Rightarrow$(3)之证明得 $x = \sum_{n=1}^{\infty} c_n e_n$，故 $x \in X_0$。于是 $D \subset X_0$。因 X_0 是闭集，则 $X = \overline{D} \subset \overline{X}_0 = X_0$，即得 $X_0 = X$。由 X_0 定义，任 $x \in X_0 = X$，有 $x = \sum_{n=1}^{\infty} c_n e_n = \lim_{n\to\infty} \sum_{k=1}^{n} c_k e_k$，且 $c_n = \langle x, e_n \rangle, n = 1,2,\cdots$。因此由定理4命题(3)成立推得 $\{e_n\}_{n=1}^{\infty}$ 是完全的。证毕。

例4　$L^2[0,2\pi]$ 中三角函数系 $\left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos t, \frac{1}{\sqrt{\pi}}\sin t, \cdots, \frac{1}{\sqrt{\pi}}\cos nt, \frac{1}{\sqrt{\pi}}\sin nt, \cdots\right\}$ 是完全的。

因为取 D 为三角多项式全体组成的集合，依据数学分析知识及Lebesgue积分性质，可以证明 D 在 $L^2[0,2\pi]$ 中稠密。任三角多项式 $T_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos kt + b_k \sin kt)$ 不难验证成立Parseval等式。

根据定理4，对任意 $x \in L^2[0,2\pi]$，其中Fourier级数

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt)$$

依范数收敛于 x。但这并不能推知对每个 $t \in [0,2\pi]$，有

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt)。$$

由线性代数及解析几何的知识，我们知道直交组比一般的线性无关组的性质更为优越，若某向量可用标准直交组线性表示，其组合系数由内积容易求出，十分方便。

以下介绍一个得到标准直交系的常用方法。对内积空间 X 中已知的某线性无关序列 $\{e_n\}_{n=1}^{\infty}$，通过 Gram – Schmidt 标准直交化过程而获得一个标准直交系。其过程如下：

第一步，把 x_1 标准化，令

$$e_1 = \frac{x_1}{\|x_1\|}。$$

第二步，记 $X_1 = \text{span}\{e_1\} = \text{span}\{x_1\}$。由定理1，$x_2$ 在 X_1 上投影为 $\langle x_2, e_1\rangle e_1$，由投影定理，记 $x_2 = \langle x_2, e_1\rangle e_1 + y_2$，则 $y_2 \perp e_1$。因 x_2，e_1 线性无关，则 $y_2 \neq \theta$，此时令

$$e_2 = \frac{y_2}{\|y_2\|}。$$

不难看出有 $\text{span}\{e_1, e_2\} = \text{span}\{x_1, x_2\}$。

第三步，记 $X_2 = \text{span}\{e_1, e_2\}$，也由定理1，$x_3$ 在 X_2 上投影为 $\langle x_3, e_1\rangle e_1 + \langle x_3, e_2\rangle e_2$，依据投影定理，记 $x_3 = \sum\limits_{k=1}^{2}\langle x_3, e_k\rangle e_k + y_3$，则 $y_3 \perp e_k$，$k = 1,2$。因 x_3, e_1, e_2 线性无关，则 $y_3 \neq \theta$，此时令

$$e_3 = \frac{y_3}{\|y_3\|}。$$

且易知 $\text{span}\{e_1, e_2, e_3\} = \text{span}\{x_1, x_2, x_3\}$。

于是归纳有第 n 步，记 $X_{n-1} = \text{span}\{e_1, e_2, \cdots, e_{n-1}\}$，同样由定理1，$x_n$ 在 X_{n-1} 上投影为 $\sum\limits_{k=1}^{n-1}\langle x_n, e_k\rangle e_k$，并根据投影定理，记 $x_n = \sum\limits_{k=1}^{n-1}\langle x_n, e_k\rangle e_k + y_n$，则 $y_n \perp e_k$，$k = 1,2,\cdots,n-1$。又因 $x_n, e_1, e_2, \cdots$，

e_{n-1} 线性无关，则 $y_n \neq \theta$，令

$$e_n = \frac{y_n}{\| y_n \|},$$

则易知 $\mathrm{span}\{e_1, e_2, \cdots, e_n\} = \mathrm{span}\{x_1, x_2, \cdots, x_n\}$。

于是以上程序无限进行下去，即得一个标准直交系 $\{e_n\}_{n=1}^{\infty}$。

由定理 4 后面的注得知具有可列的完全的标准直交系的 Hilbert 空间与 l^2 等距同构。因 l^2 是可分的（即存在有限或可列稠密子集），则 X 也是可分的。相反地，我们有：

【定理 6】 设 X 是 Hilbert 空间，则

(1) 若 X 是可分的，则 X 必有至多可列的完全的标准直交系；

(2) 设 X 是无限维的可分空间，则 X 的每个完全标准直交系都是可列集。

证

(1) 由设 X 存在有限或可列（也称为至多可列）个元素 $\{x_k\}$，使 $\overline{\mathrm{span}\{x_k\}} = X$，且不妨设 $\{x_k\}$ 为 X 中线性无关集合。由 Gram - Schmidt 标准直交化程序，可构造出对应于 $\{x_k\}$ 的（等势的）标准直交系 $\{e_k\}$。当 X 为 n 维内积空间时，则有 $\mathrm{span}\{e_1, e_2, \cdots, e_n\} = \mathrm{span}\{x_1, x_2, \cdots, x_n\} = X$，于是 $\{e_1, e_2, \cdots, e_n\}$ 为 X 的一个完全标准直交系；而当 X 为无限维空间时，则对任自然数 n 有 $\mathrm{span}\{e_1, e_2, \cdots, e_n\} = \mathrm{span}\{x_1, x_2, \cdots, x_n\}$，故有

$$x_k \in \mathrm{span}\{e_1, e_2, \cdots\}, \quad k = 1, 2, \cdots,$$

从而有 $X = \overline{\mathrm{span}\{x_1, x_2, \cdots\}} \subset \overline{\mathrm{span}\{e_1, e_2, \cdots\}}$，于是必 $X = \overline{\mathrm{span}\{e_1, e_2, \cdots\}}$，故 $\{e_n\}_{n=1}^{\infty}$ 是完全的。(1) 证毕。

(2) 由设 X 存在可列稠密子集 D，任取 X 一个完全标准直交系 M，则 M 是一个无限集。任取 $e_i, e_j \in M$，且 $e_i \neq e_j$，都有

$$\| e_i - e_j \|^2 = 2。$$

记 $S_i = \{x \in X: \| x - e_i \| \leq \frac{\sqrt{2}}{3}\}$，$S_j = \{x \in X: \| x - e_j \| \leq \frac{\sqrt{2}}{3}\}$，则 $S_i \cap S_j = \varnothing$。由于 D 在 X 中稠密，则存在 $x_i \in D \cap S_i$，$x_j \in D \cap S_j$，有 $x_i \neq x_j$。于是 M 的势不大于 D 的势。因而 M 必是可列集。证毕。

习题 4.3

1. 在内积空间 l^2 中，试给出一个使 Bessel 不等式成为严格不等式的例子。

2. 设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准直交系，求证对任意 $x, y \in X$，有

$$\sum_{n=1}^{\infty} | \langle x, e_n \rangle \cdot \langle y, e_n \rangle | \leq \| x \| \cdot \| y \|。$$

3. 设 $\{e_n\}_{n=1}^{\infty}$ 是内积空间 X 中一个标准直交系，给定 $x \in X$，令 $c_n = < x, e_n >$，$n = 1, 2, \cdots$，则对任意 $\varepsilon > 0$，求证

(1) 使成立不等式 $| c_n | > \varepsilon$ 的 c_n 仅有有限个；

(2) 设 $\{n: | c_n | > \varepsilon\}$ 的个数为 m，则有

$$m < \frac{1}{\varepsilon^2} \| x \|^2。$$

4. 在 $L^2[-1,1]$ 中，试将 $x_1(t) = 1$，$x_2(t) = t$，$x_3(t) = t^2$，标准直交化。

5. 求 $(a_0, a_1, a_2) \in \mathbf{R}^3$，使 $\int_0^1 (e^t - a_0 - a_1 t - a_2 t^2)^2 \mathrm{d}t$ 取最小值。

6. 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 中一个标准直交系，若 $x, y \in X$，

有 $x = \sum_{n=1}^{\infty} t_n e_n, y = \sum_{n=1}^{\infty} s_n e_n$，求证

(1) $\langle x, y\rangle = \sum_{n=1}^{\infty} t_n \bar{s}_n$；

(2) 级数 $\sum_{n=1}^{\infty} t_n \bar{s}_n$ 绝对收敛。

7. 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 中一个标准直交系，给定 $x \in X$，若 $x = \sum_{n=1}^{\infty} t_n e_n$，求证 $t_n = \langle x, e_n\rangle, n = 1,2\cdots$，且有 $\|x\|^2 = \sum_{n=1}^{\infty} |t_n|^2$。

8. 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 中一个完全的标准直交系，试问是否每个 $x \in X$ 都可用 $\{e_n\}_{n=1}^{\infty}$ 线性表示？

9. 设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 中一个标准直交系，对任意 $x \in X$，求证 $y = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$ 在 X 中收敛，并且 $x - y$ 与每个 e_n 直交。

4.4 Hilbert 空间上有界线性泛函

在理论及应用中，对一个具体的赋范线性空间 X 来说，往往要和它的共轭空间 X^* 结合在一起来研究。为此，知道有界线性泛函 $f \in X^*$ 的一般形式，自然是十分重要的。对于一般的赋范线性空间，获得这种表示是相当困难的。但对于 Hilbert 空间，情况却非常简明。

4.4.1 Riesz 定理

【定理 1】(Riesz)　设 X 是 Hilbert 空间，对每个 $f \in X^*$，惟一存

在 $y \in X$，使任意 $x \in X$，有

$$f(x) = \langle x, y \rangle,$$

并且还有

$$\|f\| = \|y\|。$$

证　若 $f = \theta$ 为零泛函，则只要取 X 中零元素 $\theta = y$ 即可。

现在设 $f \neq \theta$。此时令 $M = \{x \in X : f(x) = 0\}$ 为 f 的零空间。因 f 是连续性线泛函，则 M 是 X 的闭子空间。因 $f \neq \theta$，则必 M 为 X 的真子空间。于是由投影定理的推论 1，必定有 $z \neq \theta$，且 $z \in M^{\perp}$。显然，$z \notin M$，所以 $f(z) \neq 0$。

任意取 $x \in X$，因为

$$f\left(x - \frac{f(x)}{f(z)} z\right) = 0,$$

则 $x - \frac{f(x)}{f(z)} z \in M$。于是必有

$$\left\langle x - \frac{f(x)}{f(z)} z, z \right\rangle = 0。$$

从而得 $f(x) = \frac{f(z)}{\|z\|^2} \langle x, z \rangle$。此时令 $y = \frac{\overline{f(z)}}{\|z\|^2}$，即有

$$f(x) = \langle x, y \rangle, \forall x \in X。$$

存在性得证。

现在证明 $y \in X$ 由 f 惟一确定。如果还有 $y_1 \in X$，使

$$f(x) = \langle x, y_1 \rangle, \forall x \in X。$$

于是有 $\langle x, y - y_1 \rangle = 0, \forall x \in X$。即 $y - y_1 \perp X$，因此必 $y_1 = y$。惟一性得证。

最后证明 $\|f\| = \|y\|$。当 $f = \theta$，事实明显。现在设 $f \neq \theta$，则 $y \neq \theta$。首先由 Schwarz 不等式有 $|f(x)| = |\langle x, y \rangle| \leqslant \|y\| \|x\|$，$\forall x \in X$，于是推得 $\|f\| \leqslant \|y\|$；另一方面，取 $x = y$，又有

$$f(y) = \langle y, y\rangle = \|y\|^2 \leqslant \|f\| \|y\|$$

于是又推知 $\|y\| \leqslant \|f\|$。因而必 $\|f\| = \|y\|$。定理证毕。

注 定理告诉我们产生了一个由 X^* 到 X 内的映射。现在要说明它是个一一映射。因为任意取定元素 $y \in X$，则确定 X 上一个泛函 f 为

$$f(x) = \langle x, y\rangle, \forall x \in X,$$

由内积的性质立知 f 是线性的。再由 Schwarz 不等式，有

$$|f(x)| = |\langle x, y\rangle| \leqslant \|y\| \|x\|, \forall x \in X。$$

因而 f 是有界泛函，且 $\|f\| \leqslant \|y\|$，故 $f \in X^*$。又类似于定理的证明，可推得 $\|f\| = \|y\|$。于是以下的由 X 到 X^* 上的映射 T 是个一一映射：

$$T(y) = f \in X^*, \forall y \in X, 使 f(x) = \langle x, y\rangle, \forall x \in X。$$

任取复数 λ_1, λ_2 及元素 $y_1, y_2 \in X$，令 $Ty_1 = f_1, Ty_2 = f_2, T(\lambda_1 y_1 + \lambda_2 y_2) = f$，则对任意 $x \in X$，有

$$\begin{aligned} f(x) = {} & \langle x, \lambda_1 y_1 + \lambda_2 y_2\rangle = \\ & \bar{\lambda}_1\langle x, y_1\rangle + \bar{\lambda}_2\langle x, y_2\rangle = \\ & \bar{\lambda}_1 f_1(x) + \bar{\lambda}_2 f_2(x) = (\bar{\lambda}_1 f_1 + \bar{\lambda}_2 f_2)(x) = \\ & (\bar{\lambda}_1 T(y_1) + \bar{\lambda}_2 T(y_2))(x), \end{aligned}$$

即有 $T(\lambda_1 y_1 + \lambda_2 y_2) = \bar{\lambda}_1 Ty_1 + \bar{\lambda}_2 Ty_2$。

因此称 T 为复共轭线性映射，并且有 $\|f\| = \|T(y)\| = \|y\|, \forall y \in X$。即 T 是一个等距映射(或称为保范映射)。故我们称映射 T 是 X 到 X^* 上的复共轭等距同构映射。在这种意义下，认为元素 $y \in X$ 与对应的泛函 $f \in X^*$ 是一致的，即 $X^* = X$。因此，称 X 为自共轭空间(必须注意是在复共轭等距同构意义下)。

4.4.2 Hilbert 空间上的共轭算子

我们曾在第三章讨论过赋范线性空间上的共轭算子问题。现在

我们利用 Hilbert 空间与共轭空间的一致化，引入所谓 Hilbert 空间上的共轭算子的概念。这类算子是在研究矩阵以及线性微分(或积分)方程的问题中提出来的，并且有着广泛的应用。

【定义 1】　设 X 和 Y 是两个内积空间，$T:X\to Y$ 是一个有界线性算子。又设 $T^*:Y\to X$ 是有界线性算子，若对任意的 $x\in X,y\in Y$，都有

$$\langle Tx,y\rangle=<x,T^*y>$$

就称 T^* 是 T 的共轭算子(或伴随算子)。

注　在复空间情况下，第三章关于赋范线性空间所引进的共轭算子与定义 1 所述的共轭算子并不完全一致，设 $T_1,T_2\in L(X,Y)$，及复数 λ_1,λ_2，按第三章所述定义，有 $(\lambda_1T_1+\lambda_2t_2)^*=\lambda_1T_1^*+\lambda_2T_2^*$。但依定义 1 的概念，却有 $(\lambda_1T_1+\lambda_2T_2)^*=\bar{\lambda}_1T_1^*+\bar{\lambda}_2T_2^*$。而在实空间情况下，则二者完全一致了。

例 1　设 $\mathbf{C}^n$ 及 $\mathbf{C}^m$ 为复 Euclid 空间，对于有界线性算子 $T:\mathbf{C}^n\to\mathbf{C}^m$，则 T 为一 m 行 n 列的矩阵

$$T=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}。$$

当 $x=(t_1,t_2,\cdots,t_n)\in\mathbf{C}^n$ 时，有

$$Tx=(\sum_{j=1}^n a_{1j}t_j,\sum_{j=1}^n a_{2j}t_j,\sum_{j=1}^n a_{mj}t_j)\in\mathbf{C}^m。$$

此时，任取 $y=(s_1,s_2,\cdots,s_m)\in\mathbf{C}^m$，有

$$\langle Tx,y\rangle=\sum_{i=1}^m(\sum_{j=1}^n a_{ij}t_j)\cdot\bar{s}_i=\sum_{j=1}^n t_j\cdot\overline{(\sum_{i=1}^m\bar{a}_{ij}s_i,)},$$

其中

$$T^* = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{m1} \\ \bar{a}_{12} & \bar{a}_{22} & \cdots & \bar{a}_{m2} \\ \vdots & \vdots & & \vdots \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{mn} \end{bmatrix}。$$

我们看到共轭算子 T^* 是 T 的转置共轭矩阵。

如果 X 是 n 维(实或复)内积空间时,取定 $\{e_1, e_2, \cdots, e_n\}$ 为其一个标准直交基,Y 是 m 维(实或复)内积空间,取定 $\{f_1, f_2, \cdots, f_m\}$ 为其一个标准直交基。设 $T: X \to Y$ 是线性算子(则 T 一定有界)。令

$$Te_j = \sum_{i=1}^{m} a_{ij} f_i, j = 1,2,\cdots,n。$$

则任意 $x \in X$,有惟一表示 $x = \sum_{j=1}^{n} t_j e_j$,于是有

$$Tx = \sum_{j=1}^{n} t_j Te_j = \sum_{j=1}^{n} t_j (\sum_{i=1}^{m} a_{ij} f_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij} t_j) f_i。$$

不难看出,线性算子 $T: X \to Y$ 由一个 m 行 n 列矩阵 $(a_{ij})_{m\times n}$ 所决定。类似于 Euclid 空间情形,可得 T 的共轭算子 $T^*: Y \to X$ 由 $(a_{ij})_{m\times n}$ 的转置共轭矩阵 $(\bar{a}_{ji})_{n\times m}$ 表示。

以下的定理说明了一般情况下共轭算子的存在性。

【定理 2】 设 X 是 Hilbert 空间,Y 是内积空间,则对任意有界线性算子 $T: X \to Y$,必惟一存在共轭算子 T^*。

证 对任取定 $y \in Y$,确定了 X 上线性泛函 $f(x) = \langle Tx, y\rangle$,其中 $x \in X$。因 $|f(x)| = |\langle Tx, y\rangle| \leq \|y\| \|Tx\| \leq \|y\| \|T\| \|x\|, \forall x \in X$,则 $f \in X^*$,且 $\|f\| \leq \|T\| \|y\|$。由 Riesz 定理,惟一存在 $z \in X$ 有

$$\langle Tx, y\rangle = \langle x, z \rangle, \forall x \in X。$$

我们得到了算子 $T^*: Y \to X$ 为

$$T^* y = z,\text{且} \| f \| = \| z \|。$$

使对任意 $x \in X, y \in Y$ 有 $\langle Tx, y \rangle = \langle x, T^* y \rangle$。

现在证明 T^* 是由 Y 到 X 的有界线性算子。任取复数 λ_1, λ_2 及元素 $y_1, y_2 \in Y$,因有

$$\begin{aligned}\langle Tx, \lambda_1 y_1 + \lambda_2 y_2 \rangle = {} & \bar{\lambda}_1 \langle Tx, y_1 \rangle + \bar{\lambda}_2 \langle Tx, y_2 \rangle = \\ & \bar{\lambda}_1 \langle x, T^* y_1 \rangle + \bar{\lambda}_2 \langle x, T^* y_2 \rangle = \\ & \langle x, \lambda_1 T^* y_1 + \lambda_2 T^* y_2 \rangle,\end{aligned}$$

因此 $T^*(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 T^* y_1 + \lambda_2 T^* y_2$。这说明 T^* 是线性的。再由 T^* 的定义,对任意 $y \in Y$,有 $\| T^* y \| = \| f \| \leqslant \| T \| \| y \|$。因此有 $\| T^* \| \leqslant \| T \|$,即 T^* 为有界线性算子,而 T^* 的惟一性是明显的。证毕。

再给出一个实例。设 $X = \mathrm{L}^2[a,b]$,$K(t,s)$ 是矩形区域 $D = [a,b] \times [a,b]$ 上平方可积函数,则由核 $K(t,s)$ 定义了空间 $\mathrm{L}^2[a,b]$ 上的有界线性算子 T 为

$$(Tx)(t) = \int_a^b K(t,s) x(s) \mathrm{d}s, x \in \mathrm{L}^2[a,b]。$$

T 是一个 Fredholm 型积分算子。现在求 T 的共轭算子。任取 $y \in \mathrm{L}^2[a,b]$,因为在给定条件下可交换积分次序,有

$$\begin{aligned}\langle x, T^* y \rangle = \langle Tx, y \rangle = {} & \int_a^b \left(\int_a^b K(t,s) x(s) \mathrm{d}s \right) \overline{y(t)} \mathrm{d}t = \\ & \int_a^b x(s) \cdot \left(\int_a^b K(t,s) \overline{y(t)} \mathrm{d}t \right) \mathrm{d}s = \\ & \int_a^b x(s) \cdot \overline{\left(\int_a^b \overline{K(t,s)} y(t) \mathrm{d}t \right)} \mathrm{d}s = \\ & \int_a^b x(t) \cdot \overline{\left(\int_a^b \overline{K(s,t)} y(s) \mathrm{d}s \right)} \mathrm{d}t =\end{aligned}$$

故有

$$(T^* y)(t) = \int_a^b \overline{K(s,t)} y(s) \mathrm{d}s$$

即 T^* 是以 $\overline{K(s,t)}$ 为核的 Fredholm 型积分算子。

由例 1，我们看到共轭算子是转置共轭矩阵概念的推广，因此，它必然地具有许多类似于转置共轭矩阵的性质。

【定理 3】(共轭算子的性质)　设 X,Z 是 Hilbert 空间，Y 是内积空间。$T,S\in L(X,Y)$，$Q\in L(Z,X)$，λ 是复数，则以下命题成立。

(1)$(\lambda T)^*=\bar{\lambda}T^*$；

(2)$(T+S)^*=T^*+S^*$；

(3)$(T^*)^*=T$；

(4)$\|T\|^2=\|T^*\|^2=\|T^*T\|$；

(5)$(TQ)^*=Q^*T^*$；

(6)T 存在有界线性逆算子的充要条件是 T^* 存在有界线性逆算子，且当 T 存在有界线性逆算子时，有 $(T^{-1})^*=(T^*)^{-1}$。

证

(1) 任取 $x\in X,y\in Y$，有

$\langle x,(\lambda T)^*y\rangle=\langle\lambda Tx,y\rangle=\lambda\langle Tx,y\rangle=\lambda\langle x,T^*y\rangle=<x,\bar{\lambda}T^*y>$

因此有 $(\lambda T)^*=\bar{\lambda}T^*$。(1) 得证。

(2) 的证明留给读者完成。

(3) 任取 $x\in X,y\in Y$，有 $\langle Tx,y\rangle=\langle x,T^*y\rangle$，因此有 $\langle T^*y,x\rangle=<y,Tx>$。于是由定义 1 得知 $(T^*)^*=T$。(3) 得证。

(4) 由定理 2 的证明已知 $\|T^*\|\leqslant\|T\|$。因此亦有 $\|(T^*)^*\|\leqslant\|T^*\|$，即 $\|T\|\leqslant\|T^*\|$。于是必 $\|T^*\|=\|T\|$。任取 $x\in X$，因

$$\|(T^*T)x\|=\|T^*(Tx)\|\leqslant\|T^*\|\|Tx\|\leqslant\|T^*\|\|T\|\|x\|=\|T\|^2\|x\|,$$

则得 $\|T^*T\|\leqslant\|T\|^2$。另一方面，任取 $x\in X$，且 $\|x\|=1$，有

$\|Tx\|^2=\langle Tx,Tx\rangle=\langle(T^*T)x,x\rangle\leqslant\|T^*T\|\|x\|^2=\|T^*T\|$，

则得 $\|T\| = \sup\limits_{\|x\|=1}\|Tx\| \leqslant (\|T^*T\|)^{\frac{1}{2}}$，即有 $\|T\|^2 \leqslant \|T^*T\|$。综上所证就得到了 $\|T\|^2 = \|T^*\|^2 = \|T^*T\|$。性质 (4) 得证。

(5) 由设知 $TQ \in L(Z,Y)$。任取 $z \in Z, y \in Y$，因

$\langle z,(TQ)^*y\rangle = \langle (TQ)z,y\rangle = \langle Qz,T^*y\rangle = \langle z,Q^*T^*y\rangle$，

于是有 $(TQ)^* = Q^*T^*$。(5) 得证。

(6) 设 T 存在有界线性逆算子 T^{-1}，则 $TT^{-1} = I_Y, T^{-1}T = I_X$，其中 I_X, I_Y 分别是 X 及 Y 上单位（恒等）算子。因明显有 $I_X^* = I_X, I_Y^* = I_Y$，则利用(5) 可得

$$(T^{-1})^*T^* = I_Y,\quad T^*(T^{-1})^* = I_X,$$

因此知 $(T^{-1})^*$ 是 T^* 的逆算子，即成立 $(T^*)^{-1} = (T^{-1})^*$。反之，设 T^* 存在有界线性逆算子，于是由前证有 $T = (T^*)^*$ 存在有界线性逆算子。(6) 得证。定理证毕。

习题 4.4

1. 设 X 是 Hilbert 空间，Y 是内积空间，若 $S_1, S_2 \in L(Y,X)$，有 $\langle x,S_1y\rangle = \langle x,S_2y\rangle, \forall x \in X, y \in Y$，求证 $S_1 = S_2$。

2. 设 X 是 Hilbert 空间，求证 X 是自反空间。

3. 证明 $I^* = I, \theta^* = \theta$，其中 I, θ 分别是 Hilbert 空间 X 上单位算子和零算子。

4. 试求作用于 l^2 上的算子的共轭算子：

(1) $T(t_1,t_2,\cdots) = (0,t_1,t_2,\cdots)$；

(2) $T(t_1,t_2,\cdots) = (t_2,t_3,\cdots)$。

5. 试求作用于 $L^2(-\infty,\infty)$ 上的算子 T 的共轭算子：

(1) $(Tx)(t) = x(t+h)$，其中 $x \in L^2(-\infty,\infty)$，$h$ 是实常数；

(2) $(Tx)(t) = \frac{1}{2}(x(t) + x(-t))$，其中 $x \in L^2(-\infty,\infty)$。

6. 设 X 是复 Hilbert 空间，$T \in L(X) = L(X,X)$。求证 $T = T^* \Leftrightarrow$ 对任何 $x \in X$，有 $\mathrm{Re}\langle Tx, x\rangle = 0$。

7. 设 X 是 Hilbert 空间，$T \in L(X)$ 且 $\|T\| \leqslant 1$，求证

$$\{x : Tx = x\} = \{x : T^* x = x\}。$$

8. 设 X, Y 都是 Hilbert 空间，$T \in L(X,Y)$。记 T 的零空间与值域分别为 $N(T) = \{x \in X : Tx = \theta\}$，$R(T) = \{Tx \in Y : x \in X\}$。

(1) 任 $A \subset X, B \subset Y$，若 $T(A) \subset B$，求证 $A^{\perp} \supset T^*(B^{\perp})$；

(2) 若(1)中 A, B 都是闭线性子空间，若 $A^{\perp} \supset T^*(B^{\perp})$，求证 $T(A) \subset B$；

(3) 求证：$\overline{R(T^*)} = (N(T))^{\perp}$；$\overline{R(T)} = (N(T^*))^{\perp}$；

$$N(T) = (R(T^*))^{\perp}；N(T^*) = (R(T))^{\perp}。$$

9. 设 X 是复 Hilbert 空间，M 是 X 的闭线性子空间，求证 M 是 X 上某个非零有界线性泛函 f 的零空间 $\Leftrightarrow M^{\perp}$ 是 X 的一维空间。

4.5 投影算子，自共轭算子，酉算子和正规算子

4.5.1 投影算子

利用投影定理我们引入投影算子的概念。

【定义 1】 设 M 为 Hilbert 空间 X 的一个给定的闭线性子空间，则任意 $x \in X$，由投影定理，存在惟一的直交分解 $x = x_0 + z$，其中 $x_0 \in M$，$z \in M^{\perp}$，则定义算子 $P : X \to M$ 为 $Px = x_0$，$\forall x \in X$，并称

P 为由 X 到 M 上投影算子。有时为清楚起见，记 P 为 P_M。

例 1　设 $\{e_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 X 的一个标准直交系，记 $M = \text{span}\{e_1, e_2, \cdots, e_n\}$，则任意 $x \in X$，有

$$Px = \sum_{k=1}^{n} \langle x, e_k \rangle e_k。$$

注　关于内积空间 X 的子空间 M，如果任意 $x \in X$，惟一存在 M 上的投影时，也可引入投影算子概念。

以下定理给出了投影算子的几条性质。

【定理 1】

(1) 投影算子一定是有界线性算子；

(2) 当 $M = \{\theta\}$ 时，则 $\|P\| = 0$；当 $M \neq \{\theta\}$ 时，$\|P\| = 1$；

(3) $Px = \theta \Leftrightarrow x \in M^{\perp}$；$Px = x \Leftrightarrow x \in M$；

(4) $P^2 = P$，即 P 有幂等性。

证

(1) 任意数 λ_1, λ_2 及任意元素 $x_1, x_2 \in X$，有

$x_1 = Px_1 + z_1, x_2 = Px_2 + z_2$，其中 $z_1, z_2 \in M^{\perp}$，

于是有 $\lambda_1 x_1 + \lambda_2 x_2 = (\lambda_1 Px_1 + \lambda_2 Px_2) + (\lambda_1 z_1 + \lambda_2 z_2)$，且

$$\lambda_1 Px_1 + \lambda_2 Px_2 \in M, \lambda_1 z_1 + \lambda_2 z_2 \in M^{\perp},$$

因此得 $P(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 Px_1 + \lambda_2 Px_2$。这说明 P 为线性算子。又任取 $x \in X$。由勾股公式有 $\|x\|^2 = \|Px\|^2 + \|z\|^2$，因此得 $\|Px\| \leqslant \|x\|$，故 P 是有界算子，且 $\|P\| \leqslant 1$。(1) 得证。

(2) 当 $M = \{\theta\}$ 时，则 P 就是 X 上零算子，故 $\|P\| = 0$。当 $M \neq \{\theta\}$ 时，则存在 $x_0 \in M$，且 $x_0 \neq \theta$，于是 $Px_0 = x_0$。此时有

$$\|P\| = \sup_{\substack{x \in X \\ x \neq \theta}} \frac{\|Px\|}{\|x\|} \geqslant \frac{\|Px_0\|}{\|x_0\|} = 1。$$

再由 (1) 知 $\|P\| \leqslant 1$。从而必有 $\|P\| = 1$。

(2) 得证。

(3) 事实明显。

(4) 任取 $x \in X$,有 $P^2x = P(Px) = Px$,故 $P^2 = P$。定理证毕。

4.5.2 自共轭算子

由 4.4 节中 Hilbert 空间 X 上共轭算子定义,对于有界线性算子 $T: X \to X$,共轭算子 T^* 也为由 X 到 X 的有界线性算子。因此对 T 与 T^* 可进行比较,特别有以下概念。

【定义 2】 设 T 是 Hilbert 空间 X 上的有界线性算子,若 $T = T^*$,则称 T 为自共轭算子(也常称为自伴算子)。

见 4.4.2 段中例 1,可以看出自共轭算子是矩阵理论中 Hermite 矩阵的一种推广。它是在理论和应用中一类有用的基本算子。也常有人称自共轭算子为 Hermite 算子。

以下的定理给出 T 为自共轭算子的一个重要且简明的判据。

【定理 2】 设 T 是 Hilbert 空间 X 上的有界线性算子,则 T 为自共轭算子的充要条件是对任意 $x \in X$,$\langle Tx, x\rangle$ 为实数。

证 先证充分性。设对任意 $x \in X$,$\langle Tx, x\rangle$ 为实数,则有 $\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^*x\rangle} = \langle T^*x, x\rangle$。因此对任意 $x \in X$,有

$$0 = \langle Tx, x\rangle - \langle T^*x, x\rangle = \langle (T - T^*)x, x\rangle。$$

令 $T - T^* = S$,即有 $\langle Sx, x\rangle = 0, \forall x \in X$。于是由

$$\langle S(x + y), x + y\rangle = 0,$$

推得 $\langle Sx, y\rangle + \langle Sy, x\rangle = 0$;又由

$$\langle S(x + iy), x + iy\rangle = 0,$$

推得 $\langle Sx, y\rangle - \langle Sy, x\rangle = 0$。二式相加可得

$$\langle Sx, y\rangle = 0,$$

上式对 X 中任两个元素 x, y 都成立。于是对任 $x \in X$,则 $Sx \perp X$,因

此必 $Sx=\theta,\forall x\in X$,故 S 是零算子,即得 $T=T^*$。充分性得证。

再证必要性。设 $T=T^*$,则对任意 $x\in X$,有

$$\langle Tx,x\rangle=\langle x,T^*x\rangle=\langle x,Tx\rangle=\overline{\langle x,Tx\rangle}$$

故$\langle Tx,x\rangle$是实数。必要性得证。定理证毕。

以下的定理说明了自共轭算子与投影算子的关系。

【定理 3】 设 T 是 Hilbert 空间 X 上的有界线性算子,则 T 为 X 中某闭线性子空间 M 上投影算子的充要条件是:

(1) T 是自共轭的;

(2) T 具有幂等性,即 $T^2=T$。

证　先证充分性。记 I 为单位算子,记算子 $I-T$ 的零空间为

$$N(I-T)=\{x\in X:(I-T)x=\theta\}。$$

由条件(2),任取 $x\in X$,有

$$(I-T)Tx=(T-T^2)x=\theta,$$

故 $Tx\in N(I-T),\forall x\in X$。这推得 $M=\{Tx\in X:x\in X\}\subset N(I-T)$。另一方面,任取 $x\in N(I-T)$,则有$(I-T)x=\theta$,即 $Tx=x$。这说明 $N(I-T)\subset M$。因而必有 $M=N(I-T)$。注意 T 为有界线性算子,所以 $N(I-T)$ 是 X 中闭线性子空间,即 M 是 X 中一个闭线性子空间。现在证明 T 为 M 上投影算子。任取 $x,y\in X$,有$\langle x-Tx,Ty\rangle=\langle Tx-T^2x,y\rangle=0$。因 y 取遍 X 时,则 Ty 取遍 M。因此推得 $x=Tx\perp M$,即 $x\in X$ 有直交分解

$$x=Tx+(x-Tx),$$

于是由定义 1 知 T 为 M 上投影算子。充分性得证。

再证必要性。任取 $x,y\in X$,由于 T 为闭线性子空间 M 上投影算子,则分别有直交分解

$$x=Tx+z_1;y=Ty+z_2,\text{其中 }Tx,Ty\in M,z_1,z_2\in M^{\perp}。$$

于是有$\langle Tx,y\rangle=\langle Tx,Ty+z_2\rangle=\langle Tx,Ty\rangle=\langle x-z_1,Ty\rangle=\langle x,$

$Ty\rangle$。因此 T 是自共轭算子，(1) 得证。而(2) 已由定理 1 得证。必要性得证。定理证毕。

以下的定理给出了自共轭算子的一条基本性质。

【定理 4】 设 T_n 是 Hilbert 空间 X 上的一个自共轭算子列。若存在 X 上算子 T，使 $\lim\limits_{n\to\infty}\|T_n - T\| = 0$，即 T_n 依照空间 $\mathrm{L}(X) = \mathrm{L}(X, X)$ 中范数收敛于 T，则 T 也是自共轭算子。

证 由定理条件可知 $T \in \mathrm{L}(X)$。再由 4.4 节定理 3 有

$$\|T^* - T_n^*\| = \|(T_n - T)^*\| = \|T_n - T\|,$$

从而有

$$\begin{aligned}\|T - T^*\| \leqslant{}& \|T - T_n\| + \|T_n - T_n^*\| + \|T_n^* - T^*\| \leqslant \\ & \|T_n - T\| + 0 + \|T_n - T\| = \\ & 2\|T_n - T\| \to 0 (n \to \infty)。\end{aligned}$$

由此立知 $T = T^*$。定理证毕。

4.5.3 酉算子和正规算子

酉矩阵和正规矩阵是矩阵理论中重要概念。它们的一般推广是以下的定义。

【定义 3】 设 X 是 Hilbert 空间，$T \in \mathrm{L}(X)$。

(1) 若 T 是可逆算子，且 $T^* = T^{-1}$，则称 T 为 X 上一个酉算子。

(2) 若 $TT^* = T^*T$，则称 T 为 X 上一个正规算子。

注 若 T 是自共轭算子或酉算子，则一定是正规的，但逆并不成立。例如设 I 为 X 上单位算子，则算子 $T = 2\mathrm{i}I$ 是一个正规算子。因 $T^* = -2\mathrm{i}I$，故 $TT^* = T^*T = 4I$。但 $T^* \neq T$，且 $T^* \neq T^{-1} = -2^{-1}\mathrm{i}I$。

先介绍酉算子一些基本性质。

【定理 5】 设 T 与 S 都是 Hilbert 空间 X 上酉算子，则以下命题

成立。

(1) T 为等距算子，即 $\|Tx\| = \|x\|, \forall x \in X$；

(2) 若 $X \neq \{\theta\}$，则 $\|T\| = 1$；

(3) T^{-1} 也是酉算子；

(4) TS 是酉算子。

证

(1) 因 $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle = \langle x, x\rangle = \|x\|^2$，因而(1)真。

(2) 这由(1)立即可得到。

(3) 由于 T 是酉算子，则 $T^{-1} = T^*$，于是有 $(T^{-1})^* = (T^*)^* = T = (T^{-1})^{-1}$，故(3)也真。

(4) 由共轭算子的性质，有 $(TS)^* = S^*T^* = S^{-1}\cdot T^{-1} = (TS)^{-1}$。这说明(4)真。定理证毕。

此外，再介绍酉算子的一个判据。

【定理 6】 设 X 为 Hilbert 空间，$T \in L(X)$，则 T 为酉算子 $\Leftrightarrow T$ 为等距的满映射。

证　$\Leftarrow$ 设 T 为等距的满映射，则由等距性 T 必为一对一映射，再由于是满射，所以 T 存在逆映射 T^{-1}。现证 $T^* = T^{-1}$。任取 $x \in X$，有

$$\langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 = \|x\|^2 = \langle x, x\rangle,$$

即得 $\langle (T^*T - I)x, x\rangle = 0$。同定理 2 的证明，可得 $T^*T = I$。由此又有

$$TT^* = TT^*(TT^{-1}) = T(T^*T)T^{-1} = TT^{-1} = I。$$

综上可得 $T^* = T^{-1}$。因此 T 是酉算子。充分性得证。

$\Rightarrow$ 设 T 为酉算子，则由定义立知 T 是一个满映射，而 T 的等距性由定理 5 得证。必要性得证。定理证毕。

以下简要介绍一些正规算子的基本性质。

借鉴复数的表示，任意给定一个 Hilbert 空间 X 上有界线性算子 T，令 $A = \frac{1}{2}(T + T^*)$，$B = \frac{1}{2i}(T - T^*)$，则有表示

$$T = A + \mathrm{i}B。$$

并且明显可知 A 和 B 都是自共轭算子。分别称 A，B 为 T 的实部与虚部。

【定理 7】 设 X 是 Hilbert 空间，$T \in L(X)$，则 $T = A + \mathrm{i}B$ 是正规算子 $\Leftrightarrow AB = BA$。

证 因为 $T^* = (A + \mathrm{i}B)^* = A^* - \mathrm{i}B^* = A - \mathrm{i}B$，则有

$$TT^* = (A + \mathrm{i}B)(A - \mathrm{i}B) = A^2 + B^2 + \mathrm{i}BA - \mathrm{i}AB$$

$$T^*T = (A - \mathrm{i}B)(A + \mathrm{i}B) = A^2 + B^2 - \mathrm{i}BA + \mathrm{i}AB$$

立即看到 $TT^* = T^*T \Leftrightarrow BA = AB$。定理证毕。

【定理 8】 设 X 是 Hilbert 空间。$T \in L(X)$，则 T 是正规算子 $\Leftrightarrow$ 对任意 $x \in X$，有 $\|Tx\| = \|T^*x\|$。

证 $\Leftarrow$ 因有 $\|Tx\| = \|T^*x\|$，$\forall x \in X$，则

$$\langle (T^*T - TT^*)x, x\rangle = \langle T^*Tx, x\rangle - \langle TT^*x, x\rangle = \|Tx\|^2 - \|T^*x\|^2 = 0。$$

于是依据定理 2 之证，必有 $T^*T = TT^*$，充分性得证。

$\Rightarrow$ 设 $TT^* = T^*T$，任取 $x \in X$，有

$$\|T^*x\|^2 = \langle T^*x, T^*x\rangle = \langle TT^*x, x\rangle = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2,$$

即有 $\|Tx\| = \|T^*x\|$。必要性得证。定理证毕。

4.5.4 自共轭算子的谱性质

类似于 Hermite 矩阵的特性值理论，自共轭算子有许多在实用上

很重要的一般性质。本段仅考察复 Hilbert 空间 X 上的自共轭算子。

【定理 9】　设 T 为 X 上一个自共轭算子，则

(1) 若 T 存在特征值时，则一定是实的；

(2) T 的不同特征值的特征元直交。

证

(1) 设 λ 是 T 的任一特征值，x 为对应的特征元，则 $x \neq \theta$，且 $Tx = \lambda x$。因为 T 自共轭，因此有

$$\lambda\langle x,x\rangle = \langle \lambda x,x\rangle = \langle Tx,x\rangle = \langle x,Tx\rangle = \langle x,\lambda x\rangle = \bar{\lambda}\langle x,x\rangle$$

因 $\langle x,x\rangle \neq 0$，于是必有 $\lambda = \bar{\lambda}$。所以 λ 是实数。(1) 得证。

(2) 设 λ,μ 都是 T 的特征值，且 $\lambda \neq \mu$，x 与 y 分别是对应的特征元，则 $Tx = \lambda x$，$Ty = \lambda y$，利用 T 的自共轭性及(1) 的结论，有

$$\lambda\langle x,y\rangle = \langle \lambda x,y\rangle = \langle Tx,y\rangle = \langle x,Ty\rangle = \langle x,\mu y\rangle = \mu\langle x,y\rangle,$$

因 $\lambda \neq \mu$，故一定有 $\langle x,y\rangle = 0$。(2) 得证。定理证毕。

不仅自轭算子的特征值只能是实数，而且它的整个谱集也只能是实数点集。为此，我们首先给出正则集的一个特征。

【定理 10】　设 $T \in \mathrm{L}(X)$ 且为自共轭的，则数 λ 属于正则集 $\rho(T) \Leftrightarrow$ 存在 $c > 0$，使有 $\|T_\lambda x\| \geqslant c\|x\|$，$\forall x \in X$，其中 $T_\lambda = T - \lambda I$。

证　$\Leftarrow$ 在充分性条件下，我们按以下步骤证明：

(1) $T_\lambda : X \to T_\lambda(X)$ 是个一对一映射；

(2) $T_\lambda(X)$ 在 X 中稠密；

(3) $T_\lambda(X)$ 在 X 中是闭的。

于是推得 $T_\lambda(X) = X$，则根据 Banach 逆算子定理，$T_\lambda^{-1} \in \mathrm{L}(X)$，即 $\lambda \in \rho(T)$。

首先证明(1)。任取 $x_1,x_2 \in X$，若 $T_\lambda x_1 = T_\lambda x_2$，则有

$$0 = \| T_\lambda x_1 - T_\lambda x_2 \| = \| T_\lambda(x_1 - x_2) \| \geqslant c \| x_1 - x_2 \|。$$

因 $c > 0$,故一定有 $\| x_1 - x_2 \| = 0$,即 $x_1 = x_2$。这说明 T_λ 是个一对一映射。(1) 获证。

再证明(2)。任取 $x_0 \in X$,且 $x_0 \perp \overline{T_\lambda(X)}$,则对任意 $x \in X$,有

$$0 = \langle T_\lambda x, x_0 \rangle = \langle (T - \lambda I)x, x_0 \rangle = \langle Tx, x_0 \rangle - \lambda \langle x, x_0 \rangle =$$
$$\langle x, Tx_0 \rangle - \langle x, \bar{\lambda} x_0 \rangle = \langle x, Tx_0 - \bar{\lambda} x_0 \rangle,$$

因此 $Tx_0 - \bar{\lambda} x_0 \perp X$,则一定有 $Tx_0 = \bar{\lambda} x_0$。而 $x_0 \neq 0$ 是不可能的,否则就推得 $\bar{\lambda}$ 是 T 的特征值。由定理 9,有 $\bar{\lambda} = \lambda$ 是实数,从而有 $T_\lambda x_0 = Tx_0 - \lambda x_0 = 0$。再由充分条件有

$$0 = \| T_\lambda x_0 \| \geqslant c \| x_0 \| > 0,$$

这就出现了矛盾。因此 $x_0 = \theta$,故推得 $\overline{(T_\lambda(X))}^{\perp} = \{\theta\}$。根据 4.2 节投影定理的推论 2 知 $T_\lambda(X)$ 在 X 中稠密。(2) 得证。

最后证明(3)。任取 $y \in \overline{T_\lambda(X)}$,则存在 $y_n \in T_\lambda(X)$,依 X 中范数收敛于 y。根据符号 $T_\lambda(X)$ 的意义,存在 $x_n \in X$,有 $y_n = T_\lambda x_n$。再利用充分条件有

$$\| x_n - x_m \| \leqslant \frac{1}{c} \| T_\lambda(x_n - x_m) \| = \frac{1}{c} \| y_n - y_m \|。$$

因 y_n 在 X 中收敛,则 y_n 是 X 中 Cauchy 列,故推得 x_n 是 X 中 Cauchy 列。于是有 $x_0 = \lim\limits_{n \to \infty} x_n \in X$。注意 T 在 X 上连续,则 T_λ 也在 X 上连续。所以 $\lim\limits_{n \to \infty} y_n = \lim\limits_{n \to \infty} T_\lambda x_n = T_\lambda x \in T_\lambda(X)$。由极限的惟一性,就有 $y = T_\lambda x \in T_\lambda(X)$。因此 $T_\lambda(x)$ 是闭的,(3) 也获证。充分性证毕。

$\Rightarrow$ 设 $\lambda \in \rho(T)$,则 $R_\lambda = T_\lambda^{-1} \in \mathrm{L}(X)$,且 $\| R_\lambda \| > 0$。因为 R_λ 不能为零算子。由于 $I = R_\lambda T_\lambda$,任取 $x \in X$,有

$$\| x \| = \| R_\lambda T_\lambda x \| \leqslant \| R_\lambda \| \| T_\lambda x \|,$$

即有

$$\|T_\lambda x\| \geqslant \frac{1}{\|R_\lambda\|}\|x\| = c\|x\|, \forall x \in X。$$

必要性成立。定理证毕。

【定理 11】 设 $T \in L(X)$ 且为自共轭算子，则谱集 $\sigma(T) \subset (-\infty, \infty)$。

证 任取复数 $\lambda = \alpha + i\beta$，且 $\beta \neq 0$。对任意 $x \in X$，且 $x \neq \theta$，有

$$\langle T_\lambda x, x\rangle = \langle Tx, x\rangle - \lambda(x, x)。$$

因 T 自共轭，则$\langle Tx, x\rangle$ 是实数，因此又有

$$\overline{\langle T_\lambda x, x\rangle} = \langle Tx, x\rangle - \bar{\lambda} < x, x >,$$

从而有

$$\overline{\langle T_\lambda x, x\rangle} - \langle T_\lambda, x\rangle = (\lambda - \bar{\lambda})\|x\|^2 = 2i\beta\|x\|^2$$

或

$$-2i\mathrm{Im}\langle T_\lambda x, x\rangle = 2i\beta\|x\|^2。$$

其中 Im 表示虚部。于是利用 Schwarz 不等式有

$$|\beta|\ \|x\|^2 = |\mathrm{Im}\langle T_\lambda x, x\rangle| \leqslant |\langle T_\lambda x, x\rangle| \leqslant \|T_\lambda x\|\ \|x\|,$$

注意 $\|x\| \neq 0$，则有

$$|\beta|\ \|x\| \leqslant \|T_\lambda x\|。$$

因 $\beta \neq 0$，由定理 10，有 $\lambda \in \rho(T)$，故必有 $\sigma(T) \in (-\infty, \infty)$。定理证毕。

习题 4.5

1. 分别设 $P_1: X \to M_1$ 及 $P_2: X \to M_2$ 为投影算子，其中 X 是 Hilbert 空间。

(1) 求证 $P = P_1 + P_2$ 是投影算子 $\Leftrightarrow M_1 \perp M_2$。

(2) 若 $P = P_1 + P_2$ 是一个投影算子,求证 $P(X) = M_1 \oplus M_2$。

2. 举例说明两个投影算子之和不一定是投影算子。

3. 分别设 $P_1: X \to M_1$ 及 $P_2: X \to M_2$ 皆为投影算子,其中 X 为 Hilbert 空间。

(1) 求证乘积算子 $P = P_1P_2$ 是 X 上一个投影算子 $\Leftrightarrow P_1$ 和 P_2 是可交换的,即 $P_1P_2 = P_2P_1$。并且 $P(X) = M_1 \cap M_2$。

(2) 求证 P_1P_2 为零算子 $\Leftrightarrow M_1 \perp M_2$。

4. 设 P_1, P_2 同题 3 意义,且可交换。求证算子 $P = P_1 + P_2 - P_1P_2$ 是一个投影算子,且 $P(X) = M_1 + M_2$。

5. 举例线性算子 $T: \mathbf{R}^2 \to \mathbf{R}^2$,它具有幂等性,但不是自共轭的。

6. 设 P_n 为 Hilbert 空间 X 上一个投影算子列。若 $\lim\limits_{n\to\infty} P_n = P$,求证 P 是 X 上投影算子。

7. 设 T 与 S 都是 Hilbert 空间 X 上自共轭算子,求证乘积算子 TS 为自共轭 $\Leftrightarrow T$ 与 S 可交换。

8. 设 X 为 $\mathbf{C}^2$(二维复 Euclid 空间),$x = (t_1, t_2) \in X$,定义算子 $T: X \to X$ 为

$$Tx = (t_1 + \mathrm{i}t_2, t_1 - \mathrm{i}t_2)。$$

求 T^*,并证明 $T^*T = TT^* = 2I$。

9. 设 T 为 Hilbert 空间 X 上自共轭算子。求证

$$\|T\| = \sup_{\|x\|=1} |\langle Tx, x\rangle|。$$

10. 设 T 为 Hilbert 空间 X 上自共轭算子,若对任意 $x \in X$ 有 $\langle Tx, x\rangle \geq 0$,则称 T 为 X 上一个正算子,记为 $T \geq \theta$(θ 为零算子)。

(1) 若 $T \geq \theta$,求证 $T^n \geq \theta$,任意 $n = 1, 2, \cdots$;

(2) 若 T 自共轭,求证 $T^{2n} \geq \theta$, $n = 1, 2, \cdots$;

(3) 若 $T \geqslant \theta$，求证惟一存在算子 $B \geqslant \theta$，使 $B^2 = T$(常称 B 为 T 的正平方根算子)。

11. 求证算子

$$Tx(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x(s)e^{its}\mathrm{d}s$$

是空间 $L^2(-\infty, \infty)$ 上一个酉算子。

12. 设 T_n 是 Hilbert 空间 X 上一个正规算子列。若 $\lim\limits_{n\to\infty} T_n = T$，求证算子 T 是正规算子。

13. 若 T 是 Hilbert 空间 X 上一个正规算子，求证 $\|T^2\| = \|T\|^2$。

14. 设 T 是 Hilbert 空间 X 上一个自共轭算子，令

$$m = \inf_{\|x\|=1}\langle Tx, x\rangle, M = \sup_{\|x\|=1}\langle Tx, x\rangle。$$

求证 $\sigma(T) \subset [m, M]$，且 $m, M \in \sigma(T)$。

第 5 章　非线性分析初步

本章主要介绍一些非线性分析的常用概念和基本方法。内容包括抽象函数的微积分，非线性算子的两种微分，反函数与隐函数定理，变分法及非线性最优化。

5.1　抽象函数的微分与积分

抽象函数是普通函数在 Banach 空间中的推广。设 X 是一个 Banach 空间，X^* 是 X 的对偶空间。称 $x(t):[a,b]\to X$ 为抽象函数。

首先介绍抽象函数的两种连续性。

【定义 1】　$x(t):[a,b]\to X$，$t_0\in[a,b]$，称 $x(t)$ 在 t_0 点是连续的，是指 $\lim\limits_{t\to t_0}\|x(t)-x(t_0)\|=0$；称 $x(t)$ 在 t_0 点是弱连续的，是指对每个 $f\in X^*$，有

$$\lim_{t\to t_0}f(x(t))=f(x(t_0))。$$

如果 $x(t)$ 在 $[a,b]$ 的每个点上连续，则称 $x(t)$ 在 $[a,b]$ 上连续；如果 $x(t)$ 在 $[a,b]$ 的每个点上弱连续，同样，称 $x(t)$ 在 $[a,b]$ 上弱连续。

注　若 $x(t)$ 在 t_0 点连续，则 $x(t)$ 在 t_0 点弱连续，这是因为对于 $f\in X^*$，

$$|f(x(t))-f(x(t_0))|\leqslant\|f\|\|x(t)-x(t_0)\|,$$

所以 $\lim\limits_{t\to t_0} f(x(t)) = f(x(t_0))$，反之不然。

类似于普通函数，有：

【定理1】 若 $x(t)$ 是 $[a,b]$ 上的连续函数，则 $x(t)$ 是一致连续的，即对于 $\forall \varepsilon > 0$，$\exists \sigma > 0$，当 $t, t' \in [a,b]$ 且 $|t - t'| < \sigma$ 时有

$$\| x(t) - x(t') \| < \varepsilon。$$

下面来介绍抽象函数的两种导数的概念。

【定义2】 设 $x(t):[a,b] \to X$ 是一抽象函数，$t_0 \in [a,b]$，若 $\exists x_0 \in X$，使

$$\lim_{\Delta t\to 0} \left\| \frac{x(t_0 + \Delta t) - x(t_0)}{\Delta t} - x_0 \right\| = 0,$$

则称 $x(t)$ 在 $t = t_0$ 点可微，而 x_0 称为 $x(t)$ 在 t_0 点的导数，记为 $x'(t_0)$，即

$$\lim_{\Delta t\to 0} \frac{x(t_0 + \Delta t) - x(t_0)}{\Delta t} = x'(t_0),$$

上述极限是在范数意义下取的。若对每个 $f \in X^*$，普通函数 $f(x(t))$ 满足 $\lim\limits_{\Delta t\to 0} \left| \dfrac{f(x(t_0 + \Delta t)) - f(x(t_0))}{\Delta t} - f(x_0) \right| = 0$，则称 $x(t)$ 在 t_0 点弱可微，x_0 称为 $x(t)$ 在 t_0 点的弱导数。

注　$x(t)$ 在 t_0 点可微，则 $x(t)$ 在 t_0 点弱可微，反之未然。$x(t)$ 在 t_0 点可微，则 $x(t)$ 在 t_0 点连续。

若 $x(t)$ 在 $[a,b]$ 中每一点均可微（a 点右可微，b 点左可微），则 $x(t)$ 在 $[a,b]$ 上可微，且导函数 $x'(t)$ 也是一个从 $[a,b]$ 到 X 的抽象函数。

例 1　$X = l^1$，$x(t):[a,b] \to l^1$，记为

$$x(t) = (x_1(t), x_2(t), \cdots, x_n(t), \cdots)。$$

若 $x(t)$ 在 t_0 点可微，那么

$$x'(t_0) = (x_1'(t_0), x_2'(t_0), \cdots, x_n'(t_0), \cdots)。$$

例 2　若 $x_0 \in X, x(t) \equiv x_0$，则 $x(t)$ 在 $[a,b]$ 上每一点可微，且 $x'(t) \equiv \theta$（零元）。反之，若 $x(t)$ 可微，且 $x'(t) \equiv \theta$，则 $x(t) \equiv x_0$。（X 中某一常元），事实上，对每个 $f \in X^*$，$f(x(t))$ 可微，且有

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x(t)) = f(x'(t)) = f(\theta) = 0,$$

故 $f(x(t)) \equiv$ 常数 $= f(x(a))$，于是 $x(t) \equiv x(a) = x_0$。

【定义 3】　设 $x(t):[a,b] \to X$ 是一抽象函数，对于分划 Δ：

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$$

作 Riemann 和

$$S(X,\Delta) = \sum_{i=1}^{n} x(\xi_i)\Delta t_i,$$

此处 $\Delta t_i = t_i - t_{i-1}, \xi_i \in [t_{i-1}, t_i]$ 可任取。再记分划 Δ 的范数为 $\|\Delta\| = \max\limits_i |t_i - t_{i-1}|$，仿照普遍函数的 Riemann 积分定义，抽象函数 $x(t)$ 的 Riemann 积分为：

若存在 $I \in X$ 使对 $\forall \varepsilon > 0, \exists \delta > 0$，使对任何分划 Δ 若满足 $\|\Delta\| < \delta$ 时，相应的任何 Riemann 和 $S(x,\Delta)$ 都成立

$$\| S(x,\Delta) - I \| < \varepsilon,$$

即 $\lim\limits_{\|\Delta\| \to 0} S(x,\Delta) = I$，则称 $x(t)$ 在 $[a,b]$ 上是 Riemann 可积的，并称 I 是 $x(t)$ 在 $[a,b]$ 上的 Riemann 积分，记为 $I = \int_a^b x(t)\mathrm{d}t$。

与通常函数有相同的性质，即若 $x(t)$ 在 $[a,b]$ 上连续，则 $x(t)$ 在 $[a,b]$ 上 Riemann 可积。

【定理 2】　若 $x(t)$ 在 $[a,b]$ 上连续可微，即 $x'(t)$ 存在且连续，则牛顿 - 莱布尼兹公式（Newton - Leibnitz）成立，即

$$\int_a^b x'(t)\mathrm{d}t = x(b) - x(a)。$$

证　对 $f\in X^*$，通常函数 $g(t)=f(x(t))$ 在 $[a,b]$ 上连续可微，且

$$\int_a^b g'(t)\mathrm{d}t=\int_a^b f(x'(t))\mathrm{d}t=g(b)-g(a)=$$
$$f(x(b))-f(x(a)),$$

所以有（因 $x'(t)$ Riemann 可积）

$$f\left(\int_a^b x'(t)\mathrm{d}t\right)=f(x(b)-x(a)),$$

从而由 Hahn – Banach 定理的推论得

$$\int_a^b x'(t)\mathrm{d}t=x(b)-x(a)。$$

【定理 3】　若 $x(t)$ 在 (a,b) 内可微，且 $x(t)$ 在 $[a,b]$ 上连续，则存在 $\xi\in(a,b)$ 使

$$\|x(b)-x(a)\|\leqslant(b-a)\|x'(\xi)\|。$$

证　对每个 $f\in X^*$，$g(t)=f(x(t))$ 满足通常函数的微分中值定理的条件。因此由 Haha – Bonach 定理的推论，取 $f_0\in X^*$ 且 $\|f_0\|=1$，那么存在 $\xi\in(a,b)$ 成立

$$\|x(b)-x(a)\|=f_0(x(b))-f_0(x(a))=$$
$$f_0(x'(\xi))(b-a)\leqslant\|f_0\|\,\|x'(\xi)\|(b-a)=$$
$$\|x'(\xi)\|(b-a)。$$

【定理 4】　设 $x(t):[a,b]\to X$ 连续，令

$$y(t)=\int_a^t x(s)\mathrm{d}s,$$

则 $y(t)$ 在 $[a,b]$ 上可微，且 $y'(t)=x(t)$。

证　对 $t_0\in[a,b]$，由于 $x(t)$ 在 t_0 点连续，因此对 $\forall\varepsilon>0$，$\exists\delta>0$，当 $|t-t_0|<\delta$ 时有

$$\|x(t)-x(t_0)\|<\varepsilon。$$

注意到当 $|\Delta t|<\delta$ 时有

$$\left\|\frac{y(t_0+\Delta t)-y(t_0)}{\Delta t}-x(t_0)\right\|=$$
$$\left\|\frac{1}{\Delta t}\int_{t_0}^{t_0+\Delta t}x(s)\mathrm{d}s-x(t_0)\right\|=$$
$$\left\|\frac{1}{\Delta t}\int_{t_0}^{t_0+\Delta t}[x(s)-x(t_0)]\mathrm{d}s\right\|\leqslant$$
$$\frac{1}{\Delta t}\int_{t_0}^{t_0+\Delta t}\|x(s)-x(t_0)\|\mathrm{d}s<\varepsilon,$$

所以 $\lim\limits_{\Delta t\to 0}\dfrac{y(t_0+\Delta t)-y(t_0)}{\Delta t}=x(t_0)$，即 $y'(t_0)=x(t_0)$。

注　定理 4 的证明中用到了抽象函数积分的如下公式：

若 $x(t),x_1(t),x_2(t)$ 均是 $[a,b]$ 上的抽象 Riemann 可积函数，α、β 是实数，那么

(1) $\int_a^b[\alpha x_1(t)+\beta x_2(t)]\mathrm{d}t=\alpha\int_a^b x_1(t)\mathrm{d}t+\beta\int_a^b x_2(t)\mathrm{d}t$；

(2) $\int_a^b x(t)\mathrm{d}t=\int_a^c x(t)\mathrm{d}t+\int_c^b x(t)\mathrm{d}t\quad(a<c<b)$；

(3) $\left\|\int_a^b x(t)\mathrm{d}t\right\|\leqslant\int_a^b\|x(t)\|\mathrm{d}t$。

这些性质的证明完全类似于普通函数，这里略去。

我们还可以定义抽象函数的高阶导数及其幂级数的展式等类似于通常函数的性质，这里就不再讨论了。

习题 5.1

1. 证明若 $x(t)$ 在 t_0 点可微，则 $x(t)$ 在 t_0 点连续。

2. 设 $x(t)$ 在 $[a,b]$ 上 Riemann 可积，则对每个 $f\in X^*$，有

$$\int_a^b f(x(t))\mathrm{d}t = f\left(\int_a^b x(t)\mathrm{d}t\right)。$$

3. 设 $x(t)$ 在 t_0 点可微，则对每个 $f \in X^*$，有 $g(t) = f(x(t))$ 在 t_0 点也可微，且

$$g'(t_0) = f(x'(t_0))。$$

4. 定义 $C[a,b]$ 内的抽象函数 $x(t):[a,b] \to C[a,b]$ 为

$$x(t) = \sin ts \quad (s \in [a,b])。$$

证明：$x(t)$ 在 $[a,b]$ 上可微，且 $x'(t) = s\cos ts (s \in [a,b])$。

5.2　非线性算子的微分

本节介绍非线性算子的两种常用微分 Fréchet 微分和 Gâteaux 微分，这是高等数学中多元函数的全微分与方向导数的概念在 Banach 空间中的推广。本节出现的 Banach 空间都是指实 Banach 空间。

【定义 1】　设 X, Y 是两个 Banach 空间，$\Omega \subset X$ 是开集，$F:\Omega \to Y$ 是算子，$x_0 \in \Omega$。

(1) 称 F 在 x_0 点连续，是指

$$\lim_{\|h\| \to 0} \| F(x_0 + h) - F(x_0) \| = 0 \quad (h \in X, x_0 + h \in \Omega);$$

(2) 称 F 在 x_0 点是 Fréchet 可微的，如果存在 $T \in L(X,Y)$ 满足

$$F(x_0 + h) = F(x_0) + Th + \omega(x_0, h) \quad (h \in X, x_0 + h \in \Omega),$$

其中 $\omega(x_0,h) = o(\|h\|)$，即

$$\lim_{\|h\| \to 0} \frac{\|\omega(x_0,h)\|}{\|h\|} = 0,$$

此时称 Th 为 F 在 x_0 点关于 h 的 Fréchet 微分，记为 $d[F(x_0)h]$，算子 T 称为 F 在 x_0 点的 Fréchet 导算子，并记为 $T = F'(x_0)$；

(3) 称 F 在 x_0 点是 Gâteaux 可微的，如果对任意 $h \in X$，极限

$$\lim_{t \to 0} \frac{F(x_0 + th) - F(x_0)}{t}$$

在 Y 中存在，记其极限为 $D[F(x_0)h]$，即

$$\lim_{t \to 0} \left\| \frac{F(x_0 + th) - F(x_0)}{t} - D[F(x_0)h] \right\| = 0,$$

此时称 $D[F(x_0)h]$ 为 F 在 x_0 点处沿方向 h 的 Gâteaux 微分，如果 Gâteaux 微分可以表示为 $D[F(x_0)h] = Th$，这里 $T \in L(X, Y)$，则称 F 在 x_0 点具有有界线性的 Gâteaux 微分，并称 T 为 Gâteaux 导算子，仍记为 $F'(x_0)$。

例 1 设 F 是由下列函数

$$\begin{cases} y_1 = f_1(x_1, x_2, \cdots, x_n) \\ y_2 = f_2(x_1, x_2, \cdots, x_n) \\ \cdots \quad \cdots \quad \cdots \\ y_m = f_m(x_1, x_2, \cdots, x_n) \end{cases}$$

所确定的由 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的算子。如果每个函数 $f_j(x_1, x_2, \cdots, x_n)$ 在 $x^{(0)} = (x_1^0, x_2^0, \cdots, x_n^0)$ 点附近是连续可微的，那么 F 在 x_0 点是 Fréchet 可微的，并且 Fréchet 导算子 $F'(x_0)$ 正好是 Jacobian 矩阵

$$F'(x^0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}) & \frac{\partial f_1}{\partial x_2}(x^{(0)}) & \cdots & \frac{\partial f_1}{\partial x_n}(x^{(0)}) \\ \frac{\partial f_2}{\partial x_1}(x^{(0)}) & \frac{\partial f_2}{\partial x_2}(x^{(0)}) & \cdots & \frac{\partial f_2}{\partial x_n}(x^{(0)}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x^{(0)}) & \frac{\partial f_m}{\partial x_2}(x^{(0)}) & \cdots & \frac{\partial f_m}{\partial x_n}(x^{(0)}) \end{bmatrix}。$$

证 取 $h = (h_1, h_2, \cdots, h_n) \in \mathbf{R}^n$，根据 Lagrange 中值定理，对

每个 $1 \leqslant j \leqslant m$ 存在 $0 < \theta_j < 1$ 满足

$$f_j(x^{(0)}+h)-f_j(x^0)=\sum_{i=1}^{n}\frac{\partial f_j(x^0+\theta_j h)h_i}{\partial x_i}。$$

(此处对辅助函数 $g_j(t)=f_j(x^{(0)}+th)$ 应用中值定理而得),又因为 $\dfrac{\partial f_i}{\partial x_i}$ 在 $x^{(0)}$ 点连续,所以有

$$A_{ij}(h)=\left|\frac{\partial f_j(x^{(0)}+\theta_j h)}{\partial x_i}-\frac{\partial f_j(x^{(0)})}{\partial x_i}\right|\to 0\quad(当\ \|h\|\to 0),$$

记 $\omega(x_0,h)=F(x^{(0)}+h)-F(x_0)-F'(x_0)h$,则

$$\|\omega(x_0,h)\| =\sqrt{\sum_{j=1}^{m}\Big(\sum_{i=1}^{n}A_{ij}(h)h_i\Big)^2}\leqslant$$

$$\sqrt{\sum_{j=1}^{m}\Big(\sum_{i=1}^{n}|A_{ij}(h)|^2\Big)\Big(\sum_{i=1}^{n}h_i^2\Big)}\ (\text{Cauchy 不等式})=$$

$$\sqrt{\sum_{j=1}^{m}\sum_{i=1}^{n}|A_{ij}(h)|^2}\,\|h\|,$$

因此

$$\lim_{\|h\|\to 0}\frac{\|\omega(x_0,h)\|}{\|h\|}=\lim_{\|h\|\to 0}\sqrt{\sum_{j=1}^{m}\sum_{i=1}^{n}|A_{ij}(h)|^2}=0。$$

例 2　设 $f(t,x)$ 在 $[a,b]\times(-\infty,+\infty)$ 上二元连续,且关于 x 可导,偏导数 $\dfrac{\partial f(t,x)}{\partial x}$ 在 $[a,b]\times(-\infty,+\infty)$ 上也二元连续。定义算子 $F:C[a,b]\to C[a,b]$ 为

$$(Fx)(t)=f(t,x(t))\quad(x=x(t)\in C[a,b]),$$

那么 F 在任意点 $x_0=x_0(t)\in C[a,b]$ 处 Fréchet 可微,且 Fréchet 导算子 $F'(x_0)$ 为 $(F'(x_0)h)(t)=f_x'(t,x_0(t))h(t)$。

证　取 $h\in C[a,b]$,令 $\varphi(\epsilon)=f(t,x_0(t)+\epsilon h(t))$,$\varphi(\epsilon)$ 作为 ϵ 的函数在[0,1] 上连续可微,且

$$\varphi'(\varepsilon) = f_x'(t, x_0(t) + \varepsilon h(t))h(t)。$$

对 $\varphi(\varepsilon)$ 应用 Lagrange 中值定理有 $\varepsilon_0 \in (0,1)$ 满足

$$\varphi(1) - \varphi(0) = \varphi'(\varepsilon_0) = f_x'(t, x_0(t) + \varepsilon_0 h(t))h(t) = F(x_0 + h) - F(x_0)。$$

由于 $f_x'(t,x)$ 连续，因此当 $\| h \| \to 0$ 时有

$$\max_{t\in[a,b]} |f_x'(t, x_0(t)) - f_x'(t, x_0(t) + \varepsilon_0 h(t))| \to 0。$$

令 $\omega(x_0, h) = F(x_0 + h) - F(x_0) - F'(x_0)h$，则

$$\begin{aligned}\| \omega(x_0, h) \| &= \max_{t\in[a,b]} |f_x'(t, x_0(t)) + \varepsilon_0 h(t))h(t) - \\ &\quad f_x'(t, x_0(t))h(t)| \leqslant \\ &\quad \max_{t\in[a,b]} |f_x'(t, x_0(t)) + \varepsilon_0 h(t)) - \\ &\quad f_x'(t, x_0(t))| \cdot \| h \|,\end{aligned}$$

所以

$$\lim_{\| h \| \to 0} \frac{\| \omega(x_0, h) \|}{\| h \|} = \lim_{\| h \| \to 0} \max_{t\in[a,b]} | f_x'(t, x_0(t)) + \varepsilon_0 h(t)) - f_x'(t, x_0(t)) | = 0。$$

从上面两个例子可见，计算一个具体的算子的 Fréchet 微分是比较困难的，它不同于求函数导数那样容易。

为了方便起见，Fréchet 可微，Gâteaux 可微分别简称 F - 可微，G - 可微。下面我们讨论这两种微分之间的关系。

【定理 1】 设 $F: \Omega \to Y$，$x_0 \in \Omega$，Ω 是 X 中开集。

(1) 若 F 在 x_0 点 F - 可微，则 F 在 x_0 点必有有界线性 G - 微分，并且 $D[F(x_0)h] = d[F(x_0)h]$，即 F 在 x_0 点的 F - 导算子与 G - 导算子相同；

(2) 若 F 在 Ω 的每一点都有有界线性 G - 微分，且 G - 导算子 $F'(x): \Omega \to L(X, Y)$ 在 x_0 点连续，那么 F 在 x_0 点 F - 可微。

证

(1) F 在 x_0 点附近可以表示成

$$F(x_0+h)=F(x_0)+F'(x_0)h+\omega(x_0,h),$$

于是当 t 充分小时，用 th 代替 h，有

$$F(x_0+th)=F(x_0)+F'(x_0)(th)+\omega(x_0,th),$$

即

$$\frac{F(x_0+th)-F(x_0)}{t}-F'(x_0)h=\frac{\omega(x_0,th)}{t}。$$

而

$$\lim_{t\to0}\left\|\frac{\omega(x_0,th)}{t}\right\|=\lim_{t\to0}\frac{\|\omega(x_0,th)\|}{\|th\|}\cdot\|h\|=0,$$

故

$$\lim_{t\to0}\frac{F(x_0+th)-F(x_0)}{t}=F'(x_0)h。$$

可见 F 在 x_0 点具有有界线性的 G – 微分，且两者导算子相同。

(2) 由于算子 F 是 G 可微的，且导算子 $F'(x)$ 在 x_0 点连续，因此对 $\forall\varepsilon>0,\exists\delta>0$，当 $\|h\|<\delta$ 时，

$$\|F'(x_0+h)-F'(x_0)\|<\varepsilon。$$

我们来证，当 $\|h\|<\delta$ 时有

$$\|F(x_0+h)-F(x_0)-F'(x_0)h\|\leqslant\varepsilon\|h\|。\quad(5.1)$$

事实上，根据 Hahn – Banach 定理的推论，存在 $y^*\in Y^*$，且 $\|y^*\|=1$ 满足

$$\|F(x_0+h)-F(x_0)-F'(x_0)h\|=$$
$$y^*[F(x_0+h)-F(x_0)-F'(x_0)h]。$$

定义辅助函数 $\varphi(t)=y^*[F(x_0+th)]$，根据 F 的 G – 可微，容易证明 $\varphi(t)$ 在 $[0,1]$ 上连续可微，且 $\varphi'(t)=y^*(F'(x_0+th))$，从而由

Lagrange 中值定理,存在 $\xi \in (0,1)$ 使

$$\varphi(1) - \varphi(0) = \varphi'(\xi),$$

即

$$y^*[F(x_0 + h) - F(x_0)] = y^*(F'(x_0 + \xi h)h)。$$

注意到 $\|\xi h\| \leqslant \|h\| < \delta$,那么

$$\|F'(x_0 + \xi h) - F'(x_0)\| < \varepsilon。$$

于是

$$\begin{aligned}
&\|F(x_0 + h) - F(x_0) - F'(x_0)h\| = \\
&y^*[F(x_0 + h) - F(x_0) - F'(x_0)h] = \\
&y^*(F'(x_0 + \xi h)h) - y^*(F'(x_0)h) = \\
&y^*([F'(x_0 + \xi h) - F'(x_0)]h) \leqslant \\
&\|y^*\| \|F'(x_0 + \xi h) - F'(x_0)\| \|h\| \leqslant \varepsilon \|h\|,
\end{aligned}$$

从而不等式(5.1) 成立。若令

$$\omega(x_0, h) = F(x_0 + h) - F(x_0) - F'(x_0)h,$$

那么由式(5.1) 有

$$\lim_{\|h\| \to 0} \frac{\|\omega(x_0, h)\|}{\|h\|} = 0,$$

即 F 在 x_0 点 Fréchet 可微。

注　若 F 在 x_0 点有有界线性的 G - 微分,一般并不能推出 F 在 x_0 点是 F - 可微的。

【定理 2】　设 X, Y 是两个 Banach 空间,$\Omega \subset X$ 是开集,$F: \Omega \to Y$。

(1) 若 $F(x) \equiv y_0$,则 $F'(x) = \theta$(θ 是零算子);

(2) 若 $F(x) = Tx, T \in L(X, Y)$,则 $F'(x) = T$;

(3) 若 $H: \Omega \to Y$,且对 $x_0 \in \Omega$, F, H 均在 x_0 点 F - 可微,则对

任何实数 α、$\beta \in \mathbf{R}$，算子 $\alpha F + \beta H$ 在 x_0 点亦 F – 可微，且 $(\alpha F + \beta H)'(x_0) = \alpha F'(x_0) + \beta H'(x_0)$。

(4) 若 Z 是一个 Banach 空间，$\Omega_1 \subset Y$ 是开集，$H:\Omega_1 \to Z$，$\Omega_1 \supset F(\Omega)$。如果 $x_0 \in \Omega$，F 在 x_0 点 F – 可微，H 在 $y_0 = F(x_0)$ 点 F – 可微，则复合算子 $HF:\Omega \to Z$ 仍在 x_0 点可微，且 $(HF)'(x_0) = H'(y_0)F'(x_0)$。

证　(1) ~ (3) 容易证明，留给读者，仅证(4)。

设 $h \in X$，$k \in Y$，那么

$$F(x_0 + h) - F(x_0) - F'(x_0)h = \omega(x_0, h) \tag{5.2}$$

$$H(y_0 + k) - H(y_0) - H'(y_0)k = \omega(y_0, k) \tag{5.3}$$

且

$$\lim_{\|h\|\to 0} \frac{\|\omega(x_0,h)\|}{\|h\|} = \lim_{\|k\|\to 0} \frac{\|\omega(y_0,k)\|}{\|k\|} = 0。$$

取 $k = F(x_0 + h) - F(x_0)$，则 $\|k\| \to 0$，且 $k = F'(x_0)h + \omega(x_0, h)$，代入式(5.3)得

$$(HF)(x_0 + h) - (HF)(x_0) - H'(y_0)F'(x_0)h = \omega(y_0, k) + H'(y_0)\omega(x_0, h),$$

注意到

$$\|k\| \leqslant \|F'(x_0)\| \|h\| + \|\omega(x_0,h)\|,$$

因此当 $\|h\| \to 0$ 时，$\|k\| \to 0$ 且

$$\lim_{\|h\|\to 0} \frac{\|\omega(y_0,k) + H'(y_0)\omega(x_0,h)\|}{\|h\|} \leqslant$$

$$\lim_{\|h\|\to 0}\left[\frac{\|\omega(y_0,k)\|}{\|k\|}\left(\|F'(x_0)\| + \frac{\|\omega(x_0,h)\|}{\|h\|}\right) + \frac{\|H'(y_0)\| \|\omega(x_0,h)\|}{\|h\|}\right] = 0,$$

故 HF 在 x_0 点 F - 可微，且 $(HF)'(x_0)=H'(y_0)F'(x_0)$。

【定义 2】 设 $F:\Omega\to Y$（$\Omega\subset X$ 是开集）在 Ω 中每一点都 F - 可微，那么导算子又决定一个算子 $F'(x):\Omega\to L(X,Y)$。若 $F'(x)$ 在点 $x_0\in\Omega$ 处 F - 可微，称 $F'(x)$ 在 x_0 点处的 F - 导算子 $(F')'(x_0)$ 为算子 F 在 x_0 点的二阶 F - 导算子，记为 $F''(x_0)$，可见 $F''(x_0)\in L(X,L(X,Y))$。类似地，可定义 n 阶导算子 $F^{(n)}(x_0)(n=3,4,\cdots)$。

关于高阶导算子的进一步讨论，我们在这里就不再赘述了。

习题 5.2

1. 证明定理 2 的(1) ~ (3)。

2. 若 X 是 Hilbert 空间，$F(x)=\|x\|$，若 $x_0\neq\theta$，证明：F 在 x_0 点处是 G - 可微的，并求出 $F'(x_0)$。

3. 在 $\mathbf{R}^2$ 上定义 $F:\mathbf{R}^2\to\mathbf{R}$ 如下：

$$F(x)=\begin{cases}x_1+x_2+\dfrac{x_1^3x^2}{x_1^4+x_2^2}, & \text{当}(x_1,x_2)\neq(0,0);\\ 0 & \text{当}(x_1,x_2)=(0,0)。\end{cases}$$

证明：F 在 $(0,0)$ 点 G - 可微，但不 F - 可微。

4. $X=C[0,1]$，定义 $F(x):X\to\mathbf{R}$ 为

$$F(x)=\int_0^1 f(x(t),t)\mathrm{d}t,$$

其中 $f(u,t)$ 是 $\mathbf{R}\times[0,1]$ 上的连续可微函数。证明：F 在任意点 $x_0\in C[0,1]$ 是 F - 可微的，并求出 $F'(x_0)$。

5. 设 X 是 Banach 空间，$\Omega\subset X$ 是开集，$x_0,h\in X$，记 $l=\{x:x_0+th,0\leqslant t\leqslant 1\}\subset\Omega$。

(1) 若 $F:\Omega\to\mathbf{R}$ 在 l 上是 F - 可微的，则存在 $\theta\in(0,1)$ 满足

$$F(x_0+h)-F(x_0)=F'(x_0+\theta h)h;$$

(2) 若 Y 是另一个 Banach 空间，$F:\Omega\to Y$ 在 l 上是 F - 可微的，则存在 $\theta\in(0,1)$ 满足

$$\|F(x_0+h)-F(x_0)\|\leqslant\|F'(x_0+\theta h)h\|\leqslant\|F'(x_0+\theta h)\|\,\|h\|$$

(这个习题就是非线性算子的微分中值定理)。

5.3　隐函数与反函数定理

本节将给出非线性分析中极其重要的两个基本定理——反函数定理与隐函数定理。

【定理 1】 (隐函数定理) 设 X,Y,Z 是三个 Banach 空间，$V\subset X\times Y$ 是开集，$(x_0,y_0)\in V$，$F(x,y):V\to Z$ 连续，且满足

(1) $F(x_0,y_0)=\theta$(Z 中零元素)，F 在 V 内 F - 可微；

(2) $F_y'(x_0,y_0)$(当 x 固定时，关于 y 的导算子) 有有界逆算子 $[F_y'(x_0,y_0)]^{-1}$；

(3) $F'_y(x,y)$ 在 (x_0,y_0) 处连续，

那么存在 x_0 点和 y_0 点的闭球($\delta>0,\lambda>0$)

$$\bar{B}_\delta(x_0)=\{x:\|x-x_0\leqslant\delta\|\},\bar{V}_\lambda(y_0)=\{y:\|y-y_0\|\leqslant\lambda\}$$

满足当 $x\in\bar{B}_\delta(x_0)$ 时，方程 $F(x,y)=\theta$ 在 $\bar{V}_\lambda(y_0)$ 内存在惟一连续解 $y=f(x)$，且 $y_0=f(x_0)$，即由 $F(x,y)=\theta$ 决定一个连续算子 $f:\bar{B}_\delta(x_0)\to\bar{V}_\lambda(y_0)$，且 $F(x,f(x))=\theta$。

证　记 $M=\|[F_y'(x_0,y_0)]^{-1}\|$，因 $F_y'(x,y)$ 在 (x_0,y_0) 处连

续,所以取 x_0,y_0 的恰当小闭球 $\bar{B}_\delta(x_0)$ 及 $\bar{V}_\lambda(y_0)$,而当 $(x,y)\in\bar{B}_\delta(x_0)\times\bar{V}_\lambda(y_0)$ 时成立

$$\| F_y'(x,y)-F_y'(x_0,y_0)\| < \frac{1}{2M}。\tag{5.4}$$

又 $F(x,y_0)$ 关于 x 连续,且 $F(x_0,y_0)=\theta$,因此可认为当 $x\in\bar{B}_\delta(x_0)$ 时有

$$\| F(x,y_0)\| < \frac{\lambda}{2M}。\tag{5.5}$$

对固定的 $x\in\bar{B}_\delta(x_0)$,作映射 $\varphi(x,y)=y-[F_y'(x_0,y_0)]^{-1}F(x,y)$,我们来证明 $\varphi(x,y)$ 关于 y 满足 $\varphi(x,y):\bar{V}_\lambda(y_0)\to\bar{V}_\lambda(y_0)$,且是压缩映射。事实上由式(5.4)知

$$\begin{aligned}\|\varphi_y'(x,y)\| &= \| I-[F_y'(x_0,y_0)]^{-1}F_y'(x,y)\| \leqslant \\ &\|[F_y'(x_0,y_0)]^{-1}\|\,\| F_y'(x_0,y_0)-F_y'(x,y)\| \leqslant \\ &M\cdot\frac{1}{2M}=\frac{1}{2}。\end{aligned}\tag{5.6}$$

再由中值定理(本章 5.2 节习题 5) 及式(5.5)、(5.6) 有

$$\begin{aligned}&\|\varphi(x,y)-y_0\| \leqslant \\ &\|\varphi(x,y)-\varphi(x,y_0)\| + \|\varphi(x,y_0)-y_0\| \leqslant \\ &\sup_{y\in\bar{V}_\lambda(y_0)}\|\varphi_y'(x,y)\|\,\| y-y_0\| + \\ &\|[F_y'(x_0,y_0)]^{-1}\|\,\| F(x,y_0)\| < \frac{\lambda}{2}+M\cdot\frac{\lambda}{2M}=\lambda,\end{aligned}$$

这里 I 是单位算子即 $I(y)=y$,这便证明了 $\varphi(x,y)$ 关于 y 为 $\bar{V}_\lambda(y_0)\to\bar{V}_\lambda(y_0)$。另一方面,对 $y_1,y_2\in\bar{V}_\lambda(y_0)$,再由中值定理,存在 $\theta\in(0,1)$ 满足

$$\begin{aligned}\|\varphi(x,y_1)-\varphi(x,y_2)\| &\leqslant \|\varphi_y'(x,y_1+\theta(y_2-y_1))\| \\ &\| y_1-y_2\| \leqslant \frac{1}{2}\| y_1-y_2\|,\end{aligned}$$

故 $\varphi(x,y)$ 是压缩映射，从而由压缩原理，存在惟一一个 $y=f(x)\in \overline{V}_\lambda(x_0)$ 使得

$$y=\varphi(x,y)$$

即

$$[F_y'(x_0,y_0)]^{-1}F(x,f(x))=\theta,$$

故 $F(x,f(x))=\theta$，特别当 $x=x_0$ 时，由 y 的惟一性，$y=y_0$，从而 $f(x_0)=y_0$，这意味着已确定一个映射 $f:\overline{B}_\delta(x_0)\to \overline{V}_\lambda(y_0)$ 满足 $F(x,f(x))=\theta$。

下面证明 f 是连续映射，对任意 $x_1\in \overline{B}_\delta(x_0)$，$x_2\in \overline{B}_\delta(x_0)$，记 $y_1=f(x_1)$，$y_2=f(x_2)$。利用中值定理及式(5.6)得

$$\|\varphi(x_2,y_1)-\varphi(x_2,y_2)\|\leqslant\frac{1}{2}\|y_1-y_2\|,$$

因 $y_1=\varphi(x_1,y_1)$，$y_2=\varphi(x_2,y_2)$ 故

$$\begin{aligned}\|y_1-y_2\| &= \|\varphi(x_1,y_1)-\varphi(x_2,y_2)\|\leqslant\\ &\|\varphi(x_1,y_1)-\varphi(x_2,y_1)\|+\frac{1}{2}\|y_1-y_2\|,\end{aligned}$$

即

$$\|f(x_1)-f(x_2)\|=\|y_1-y_2\|\leqslant 2\|\varphi(x_1,y_1)-\varphi(x_2,y_1)\|。$$

因 $F(x,y)$ 关于 (x,y) 连续，所以 $\varphi(x,y)$ 关于 (x,y) 连续。因此，当 $x_2\to x_1$ 时，$\|\varphi(x_1,y_1)-\varphi(x_2,y_1)\|\to 0$，更有

$$\|f(x_2)-f(x_1)\|\to 0,$$

这说明 $y=f(x)$ 是连续映射。

利用隐函数定理，我们可以推出下面的反函数定理。

【定理2】 （反函数定理）设 X，Y 是两个 Banach 空间，$\Omega\subset X$ 开集，$x_0\in\Omega$，$f:\Omega\to Y$ 连续，且 $y_0=f(x_0)$，若 f 满足：

(1) f 在 Ω 内 F – 可微；

(2)$f'(x_0)$具有有界逆算子$[f'(x_0)]^{-1}$,且$f'(x)$在x_0点连续,那么存在x_0点及y_0点的闭球$\bar{B}_\delta(x_0)$及$\bar{V}_\lambda(y_0)$使$f:\bar{B}_\delta(x_0)\to\bar{V}_\lambda(y_0)$的逆映射$f^{-1}:\bar{V}_\lambda(y_0)\to\bar{B}_\delta(x_0)$存在且连续。

证 令$F(y,x)=y-f(x)$,则$F:Y\times\Omega\to$满足

(1)$F(y_0,x_0)=\theta$;

(2)$F_x'(y_0,x_0)=-f'(x_0),[F_x'(y_0,x_0)]^{-1}=-[f'(x_0)]^{-1}$;

(3)$F_x'(y,x)=f'(x)$在(y_0,x_0)点连续。

对于F满足定理1的全部条件,从而有闭球$\bar{V}_\lambda(y_0)$及$\bar{B}_\delta(x_0)$使对任意$y\in\bar{V}_\lambda(y_0)$有惟一连续映射$g:\bar{V}_\lambda(y_0)\to\bar{B}_\delta(x_0)$成立$F(y,g(y))=\theta$,即

$$y=f(g(y))。$$

故$g=f^{-1}:\bar{B}_\delta(x_0)\to\bar{B}_\delta(x_0)$。

注 (1)若在定理1的条件下,再附加条件F在V内关于x和y的F-导算子$F_x'(x,y)$与$F_y'(x,y)$都存在并且连续,那么定理1中惟一确定的映射$y=f(x)$也具有F-导算子$f'(x)$,并且$f'(x)$连续,更可由下面公式得出$f'(x)$的表达式:

$$f'(x)=-[F_y'(x,f(x))]^{-1}F_x'(x,f(x))。$$

(2)若在定理2的条件下,再附加条件$f'(x)$在Ω内存在,则逆映射f^{-1}也是Fréchet可微的。

习题 5.3

1. 设$F:\mathbf{R}^5\to\mathbf{R}^5$为

$$F(x)=\begin{bmatrix} x_1+x_2^2 \\ x_2+x_3^2 \\ e^{x_3}\cos x_1 \\ x_4 \\ \sin x_5 \end{bmatrix}$$

证明 F 在 $\theta=(0,0,0,0,0)$ 点附近存在逆映射 F^{-1}。

2. 设 $F:\mathbf{R}^2\to\mathbf{R}$ 为

$$F(x_1,x_2)=x_2+x_1x_2^2+x_1^2,$$

证明：$\exists\,\delta>0,\lambda>0$ 及连续映射 $f:(-\delta,\delta)\to(-\lambda,\lambda)$ 满足

$$F(x,f(x))=0\quad,\quad\forall x\in(-\delta,\delta)。$$

即

$$f(x)+x[f(x)]^2+x^2=0,$$

并求出 $f(x)$ 的具体表达式。

5.4 变分法

上节讨论了非线性算子的微分，这节专门来研究非线性泛函的极值问题。类似于在高等数学中所学的用微分（导数）来求函数的极值，对非线性泛函通过微分求极值的方法称为变分法。变分法是泛函分析的起源，也是泛函分析的重要分支。变分法在力学、物理学、控制论等领域有广泛的应用。本节仅是对变分法这一基本原理作一简要介绍。

【定义 1】 设 X 是一个 Banach 空间，$\Omega\subset X$ 是开集，$f:\Omega\to\mathbf{R}$ 是 Ω 上定义的一个泛函。如果 f 在 Ω 中每一点都有有界线性的 G – 微

分，记 $F(x) = f'(x)(x \in \Omega)$，则称算子 $F:\Omega \to X^*$ 为泛函 f 的梯度，并记为 $F(x) = \mathrm{grad}f(x)$ 或简记为 $F = \mathrm{grad}f$，有时，又称泛函 f 为算子 F 的位势。

根据 G – 微分的定义，梯度算子 F 与其位势函数 f 之间成立

$$\lim_{t\to 0}\frac{1}{t}[f(x+th)-f(x)] = F(x)h(x \in \Omega, h \in X)。$$

【定义 2】 设 X 是一个空间，$\Omega \subset X$ 是开集，$f:\Omega \to \mathbf{R}$ 是泛函，$x_0 \in \Omega$。

(1) 若存在 x_0 点的开球 $B_r(x_0) = \{x: \|x - x_0\| < r\}(r > 0)$ 满足 $B_r(x_0) \subset \Omega$，且对一切 $x \in B_r(x_0)$ 有 $f(x) \geqslant f(x_0)(f(x) \leqslant f(x_0))$，则称泛函 $f(x)$ 在 x_0 点达到极小值(相应地，极大值)；极小值与极大值统称为极值。

(2) 设 $\varphi: X \to Y$(Y 是另一个 Banach 空间)，令 $M = \{x \in X: \varphi(x) = \theta\}$，设 $x_0 \in M$，若存在 x_0 点的开球 $B_r(x_0)$，使当 $x \in B_r(x_0) \cap M$ 时有

$$f(x) \geqslant f(x_0) \quad (f(x) \leqslant f(x_0)),$$

则称泛函 $f(x)$ 关于条件 $\varphi(x) = \theta$ 在 $x = x_0$ 达到条件极小值(相应地，条件极大值)；条件极小值与条件极大值统称为条件极值。

注 若 $x_0 \in \Omega$，$f'(x_0) = \theta$，常称 x_0 是泛函 $f(x)$ 的一个临界点，$c = f(x_0)$ 称为 $f(x)$ 的一个临界值。

【定理 1】 设 $f:\Omega \to X$ 具有有界线性的 Gâteaux 微分(G – 微分)，且在 $x_0 \in \Omega$ 达到极值，那么

$$f'(x) = \theta。$$

证 对于任意 $h \in X$，取常数 $a > 0$，使当 $|t| \leqslant a$ 时 $x_0 + th \in \Omega$。定义函数 $P(t):[-a,a] \to \mathbf{R}$ 为

$$P(t) = f(x_0 + th)。$$

由 G - 微分的性质，$P(t)$ 可微，且

$$P'(t) = f'(x_0 + th)h。$$

特别，由于 $P(t)$ 在 $t = 0$ 点达到极值，根据微分学的基本性质，$P'(0) = 0$，又 $P'(0) = f'(x_0)h$，于是

$$f'(x_0)h = 0。$$

而 h 是任意的，那么 $f'(x_0) = \theta$。

这个定理虽然十分简单，但为我们寻求泛函的极值提供了十分方便的条件。

例 1 设 $C^1[a,b]$ 为一切 $[a,b]$ 上连续可微函数组成的线性空间，对于 $x \in C^1[a,b]$ 定义范数为

$$\|x\| = \max\{|x(t)|:t \in [a,b]\} + \max\{|x'(t)|:t \in [a,b]\},$$

则 $C^1[a,b]$ 在此范数下是一个 Banach 空间。又设 $L(x,u,t)$ 是 $\mathbf{R}^3$ 上定义的一个连续可微函数，求泛函

$$J(x) = \int_a^b L(x(s),\dot{x}(s),s)\mathrm{d}s$$

达到极值的条件。

解 对于任意 $h \in C^1[a,b]$ 由于

$$\lim_{t\to 0}\frac{J(x+th) - J(x)}{t} =$$

$$\lim_{t\to 0}\int_a^b \frac{L(x(s)+th(s),\dot{x}(s)+t\dot{h}(s),s) - L(x(s),\dot{x}(s),s)}{t}\mathrm{d}s =$$

$$\int_a^b\left(\frac{\partial L}{\partial x}h + \frac{\partial L}{\partial \dot{x}}\dot{h}\right)\mathrm{d}s,$$

则 $J'(x)$ 满足

$$J'(x)h = \int_a^b\left(\frac{\partial L}{\partial x}h + \frac{\partial L}{\partial \dot{x}}\dot{h}\right)\mathrm{d}s。$$

若 J 在 x_0 点处达到极值,那么对任意 $h \in C^1[a,b]$,有

$$J'(x_0)h = 0,$$

即

$$\int_a^b \left(\frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h} \right) \mathrm{d}s = 0。 \tag{5.7}$$

这个条件十分不具体,为此,我们进一步假定 $L(x,u,t)$ 是一次连续可微的,且要求极值点满足固定条件 $x_0(a) = y_1, x_0(b) = y_2$,则由式(5.7)分部积分可得

$$\int_a^b \left(\frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h} \right) \mathrm{d}s =$$

$$\int_a^b \left[\frac{\partial L}{\partial x_0} - \frac{\mathrm{d}}{\mathrm{d}s} \left(\frac{\partial L}{\partial \dot{x}_0} \right) \right] h \mathrm{d}s + h \frac{\partial L}{\partial \dot{x}_0} \Big|_a^b = 0。 \tag{5.8}$$

在式(5.8)中根据 h 的任意性,我们得到 $x_0(t)$ 满足下面的方程

$$\begin{cases} \dfrac{\partial L}{\partial x_0} - \dfrac{\mathrm{d}}{\mathrm{d}t} \left(\dfrac{\partial L}{\partial \dot{x}_0} \right) = 0; \\ x_0(a) = y_1, x_0(b) = y_2, \end{cases}$$

这个方程通常称为 Euler – Lagrange 方程。

例 2 通过 Euler – Lagrange 方程求泛函

$$J(x) = \int_0^1 [(x(t)^2) + (x'(t))^2] \mathrm{d}t$$

满足条件 $x(0) = x_0, x(1) = x_1$ 的极值点函数。

解 由 Euler – Lagrange 方程得

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{x}} \right) = 2x(t) - \frac{\mathrm{d}}{\mathrm{d}t}(2x'(t)) = 0,$$

即

$$x''(t) - x(t) = 0。$$

这个方程的通解为

$$x(t) = c_1 e^t + c_2 e^{-t}。$$

由边界条件 $x(0) = x_0, x(1) = x_1$ 解得系数分别为

$$c_1 = \frac{x_1 - x_0 e^{-1}}{e - e^{-1}}, \quad c_2 = \frac{x_0 e - x_1}{e - e^{-1}}。$$

例 3　求泛函 $J(x,u) = \int_a^b L(x(t), u(t), t)\mathrm{d}$ 满足条件 $\dot{x} = f(x(t), u(t), t)$ 的极值条件。

解　这是一个条件极值问题，我们把它转化成无条件极值问题，为此引进一个辅助函数 $\lambda(t)$，将上述问题化成

$$v(x,u,\lambda) = \int_a^b \{L(x(t), u(t), t) + \lambda(t)[f(x(t), u(t), t) - \dot{x}(t)]\}\mathrm{d}t$$

的无条件极值。取 Banach 空间 $X = C^1[a,b] \times C^1[a,b] \times C^1[a,b]$，若 (x^*, u^*, λ^*) 是泛函 $v(x,u,\lambda)$ 的极值点，那么对任意 $(h, \eta, r) \in X$ 有

$$\lim_{s \to 0} \frac{v(x^* + sh, u^* + s\eta, \lambda^* + sr) - v(x^*, u^*, \lambda^*)}{s} =$$
$$\int_a^b \left\{ \frac{\partial L}{\partial x^*} h + \frac{\partial L}{\partial u^*}\eta + \lambda^* \frac{\partial f}{\partial x^*} h + \lambda^* \frac{\partial f}{\partial u^*}\eta - \lambda^* \dot{h} + (f - \dot{x}) r \right\}\mathrm{d}t, \tag{5.9}$$

选取 $h(a) = h(b) = 0$，那么由分部积分得

$$\int_a^b \lambda^* \dot{h}\mathrm{d}t = \lambda h \Big|_a^b - \int_a^b \dot{\lambda}^* h \mathrm{d}t = -\int_a^b \dot{\lambda} h \mathrm{d}t,$$

再由式(5.9)得

$$v'(x^*, u^*, \lambda^*)(h, \eta, r) = \int_a^b \left\{ \left[\frac{\partial L}{\partial x^*} + \lambda^* \frac{\partial f}{\partial x^*} + \dot{\lambda}^* \right] h + \left[\frac{\partial L}{\partial u^*} + \lambda^* \frac{\partial f}{\partial u^*} \right]\eta + (f - \dot{x}) r \right\}\mathrm{d}t \tag{5.10}$$

由无条件极值的必要条件 $v'(x^*,u^*,\lambda^*)=\theta$ 及 (h,η,r) 的任意性，(x^*,u^*,λ^*) 应满足下述微分方程

$$\left.\begin{aligned}&\frac{\partial L}{\partial x}+\lambda\frac{\partial f}{\partial x}+\dot{\lambda}=0\\&\frac{\partial L}{\partial u}+\lambda\frac{\partial f}{\partial u}=0\\&f=\dot{x}\end{aligned}\right\}\tag{5.11}$$

上述方程(5.11)也称为 Euler – Lagrange 方程。具体求解这个方程还需要一些别的定解条件。

例 4 通过 Euler – Lagrange 方程求泛函

$$J(x,u)=\int_0^1[(x^2(t)+u^2(t))]\mathrm{d}t$$

满足条件 $\dot{x}=x(t)+u(t)$，且 $x(0)=x_1,u(0)=u_1$ 的极值函数。

解 由 Euler – Lagrange 方程(5.11)得微分方程为

$$\begin{cases}2x(t)+\lambda(t)+\dot{\lambda}(t)=0\\2u(t)+\lambda(t)=0\\\dot{x}(t)=x(t)+u(t)\end{cases}$$

整理后得

$$\begin{cases}\dot{\lambda}(t)=-\lambda(t)-2x(t)\\\dot{x}(t)=x(t)-\dfrac{1}{2}\lambda(t)\\u(t)=-\dfrac{1}{2}\lambda(t)\end{cases}$$

通过初始条件 $x(0)=x_0,u(0)=u_1$ 来确定 $(x(t),u(t))$。记矩阵

$$\boldsymbol{A}=\begin{bmatrix}-1&-2\\-\dfrac{1}{2}&1\end{bmatrix}$$

则

$$\begin{bmatrix} \lambda(t) \\ x(t) \end{bmatrix} = e^{At}\begin{bmatrix} -2u_1 \\ x_0 \end{bmatrix}。$$

注　由定理 1 知，泛函的极值点一定是它的临界点，但临界点未必一定是极值点，因此临界点只是泛函极值点的必要条件，并非充分条件。

下面我们介绍求泛函条件极值的 Lagrange 乘子法。

【定义 3】　设 X, Y 是两个 Banach 空间，$\Omega \subset X$ 是开集，$T: \Omega \to Y$ 是 Fréchet 可微(即 F – 可微)的。$x_0 \in \Omega$，称 x_0 为 T 的一个正则点，如果 F – 导算子 $T'(x_0): X \to Y$ 是满射。

例 5　设 $T: \mathbf{R}^n \to \mathbf{R}^m$，$T(x) = (f_1(x), f_2(x), \cdots, f_m(x))$，$x = (x_1, x_2, \cdots, x_n)$，若 $x^{(0)} = (x_1^0, x_2^0, \cdots, x_n^0)$ 使矩阵

$$\begin{bmatrix} \dfrac{\partial f_1(x^{(0)})}{\partial x_1} & \cdots & \dfrac{\partial f_1(x^{(0)})}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m(x^{(0)})}{\partial x_1} & \cdots & \dfrac{\partial f_m(x^{(0)})}{\partial x_n} \end{bmatrix}$$

的秩为 m，则 $x^{(0)}$ 为 T 的正则点。

【定理 2】　(Lagrange 乘子) 设 X, Y 是两个 Banach 空间，$f: \Omega \to \mathbf{R}$，$T: \Omega \to Y$ 均是 Fréchet 可微的，$x_0 \in \Omega$，且 $f'(x)$，$T'(x)$ 在 x_0 点连续，x_0 点是 T 的正则点，那么，若 x_0 点是泛函 $f(x)$ 满足条件 $T(x) = \theta$ 的极值点，则存在 $y_0^* \in Y^*$ 使得

$$f'(x_0) + y_0^* T'(x_0) = \theta。$$

【定义 4】　称 $L(x, y^*) = f(x) + y^* T(x)$ 为 Lagrange 泛函，并记

$$\frac{\partial L}{\partial x} = f'(x) + y^* T'(x)。$$

若$\frac{\partial L}{\partial x}(x_0,y_0^*)=0$,则称$x_0$为$L$关于$y_0^*$的临界点,$y^*$称为Lagrange乘子。

例5 求函数$u\in L^2[0,1]$满足微分方程

$$\ddot{\varphi}(t)+\dot{\varphi}(t)=u(t),\varphi(0)=\dot{\varphi}(0)=0,$$

又同时满足$\varphi(1)=1,\dot{\varphi}(1)=0$,且使$J(u)=\int_0^1 u^2(t)\mathrm{d}t$达到最小。

解 满足微分方程初值问题的解可表示为

$$\begin{pmatrix}\varphi(t)\\ \dot{\varphi}(t)\end{pmatrix}=\begin{pmatrix}\int_0^t(1-\mathrm{e}^{-1}\mathrm{e}^s)u(s)\mathrm{d}s\\ \int_0^t \mathrm{e}^{-1}\mathrm{e}^s u(s)\mathrm{d}s\end{pmatrix}$$

于是取$H(u)$为

$$H(u)=\begin{pmatrix}H_1(u)\\ H_2(u)\end{pmatrix}=\begin{pmatrix}\int_0^1(1-\mathrm{e}^{-1}\mathrm{e}^s)u(s)\mathrm{d}s-\varphi(1)\\ \int_0^1 \mathrm{e}^{-1}\mathrm{e}^s u(s)\mathrm{d}s-\dot{\varphi}(1)\end{pmatrix}=$$

$$\begin{pmatrix}\int_0^1(1-\mathrm{e}^{-1}\mathrm{e}^s)u(s)\mathrm{d}s-1\\ \int_0^1 \mathrm{e}^{-1}\mathrm{e}^s u(s)\mathrm{d}s\end{pmatrix}$$

则上述问题化为如下条件极值问题

$$J:L^2[0,1]\to\mathbf{R}\quad H:L^2[0,1]\to\mathbf{R}^2$$

$$\min_u J(u)=\min_u\int_0^1 u^2(t)\mathrm{d}t,H(u)=0。$$

根据定理2,存在$\lambda=(\lambda_1,\lambda_2)\in\mathbf{R}^2$使得达到极值的$u$满足$J'(u)+\lambda_1 H_1'(u)+\lambda_2 H_2'(u)=0$,即对$\forall h\in L^2[0,1]$成立

$$J'(u)h+\lambda_1 H_1'(u)h+\lambda_2 H_2'(u)h=0。\tag{5.12}$$

计算 $J'(u)h, H_1'(u)h, H_2'(u)$ 为

$$J'(u)h = 2\int_0^1 u(s)h(s)\mathrm{d}s,$$

$$H_1'(u)h = \int_0^1 (1 - \mathrm{e}^{-1}\mathrm{e}^s)h(s)\mathrm{d}s,$$

$$H_2'(u)h = \int_0^1 \mathrm{e}^{-1}\mathrm{e}^s h(s)\mathrm{d}s。$$

根据式(5.12)可得

$$\int_0^1 [2u(s) + \lambda_1(1 - \mathrm{e}^{-1}\mathrm{e}^s) + \lambda_2 \mathrm{e}^{-1}\mathrm{e}^s] h(s)\mathrm{d}s = 0,$$

由于 h 的任意性得

$$2u(s) + \lambda_1(1 - \mathrm{e}^{-1}\mathrm{e}^s) + \lambda_2 \mathrm{e}^{-1}\mathrm{e}^s = 0,$$

即

$$u(s) = -\frac{1}{2}\lambda_1(1 - \mathrm{e}^{-1}\mathrm{e}^s) - \frac{1}{2}\lambda_2 \mathrm{e}^{-1}\mathrm{e}^s。$$

最后由约束条件 $H(u) = 0$ 计算出 λ_1, λ_2 为

$$\lambda_1 = -\frac{2(\mathrm{e}+1)}{3-\mathrm{e}}, \lambda_2 = -\frac{2(1-\mathrm{e})}{3-\mathrm{e}}。$$

例 6　考虑一般线性控制系统

$$\dot{x} = Ax + Bu, x(t_0) = x_0$$

其中 $A:\mathbf{R}^n \to \mathbf{R}^n, B:\mathbf{R}^m \to \mathbf{R}^n$,即 A 是 $n \times n$ 矩阵,B 是 $n \times m$ 矩阵,给定 $x_1 \in \mathbf{R}^n$,求控制 $u \in L^2([t_0, t_1], \mathbf{R}^m)$ 使

$$\min_u J(u) = \min_u \frac{1}{2}\int_{t_0}^{t_1} \| u(t) \|^2 \mathrm{d}t$$

且满足 $x(t_1) = x_1$。

解　记 $X = C^1([t_0, t_1], \mathbf{R}^n)$,则 $J: L^2([t_0, t_1], \mathbf{R}^m) \to \mathbf{R}$,约束条件为 $H(x, u) = \dot{x} - Ax - Bu = 0$,于是 $H: X \times L^2 \to L^2$。根定理 2

存在 $\lambda(t) \in L^2([t_0, t_1], \mathbf{R}^m)$ 使条件极值 u 满足

$$\frac{\partial L}{\partial x} = J'(u) + \lambda H'(x,u) = 0。 \tag{5.13}$$

对 $\forall h \in L^2$ 及 $k \in X$ 计算 $J'(u)h$ 及 $\lambda H'(x,u)(k,h)$ 为

$$J'(u)h = \int_{t_0}^{t_1} \langle u(t), h(t) \rangle \mathrm{d}t,$$

$$\lambda H'(x,u)(k,h) = \int_{t_0}^{t} [\langle \dot{k}(t), \lambda(t) \rangle - \langle Ak(t), \lambda(t) \rangle - \langle Bh(t), \lambda(t) \rangle] \mathrm{d}t,$$

此处 $\langle \cdot, \cdot \rangle$ 表示 $\mathbf{R}^m$ 中的内积运算，即若 $y = (y_1, y_2, \cdots, y_m), z = (z_1, z_2, \cdots, z_m) \in \mathbf{R}^m$，则

$$\langle y, z \rangle = \sum_{j=1}^{m} y_j z_j。$$

于是根据式(5.13)有

$$J'(u)h + \lambda H'(x,u)(k,h) = 0。 \tag{5.14}$$

对 $\forall h \in L^2$ 及 $\forall k \in X$ 成立，为了求解式(5.14)可假定 λ 连续可微且 $k(t_0) = k(t_1) = 0$，那么由式(5.14)及

$$\int_{t_0}^{t_1} \langle \dot{k}(t), \lambda(t) \rangle \mathrm{d}t = -\int_{t_0}^{t} \langle k(t), \dot{\lambda}(t) \rangle \mathrm{d}t$$

得

$$\int_{t_0}^{t_1} \langle u - B^T\lambda, h \rangle \mathrm{d}t - \int_{t_0}^{t_1} \langle \dot{\lambda}(t) + A^T\lambda(t), k(t) \rangle \mathrm{d}t = 0。$$

再由于 h 及 k 的任意性有

$$u(t) = B^T\lambda(t), \quad \dot{\lambda}(t) + A^T\lambda(t) = 0$$

这样求得 $\lambda(t) = C\mathrm{e}^{-A^T t}$，从而

$$u(t) = B^T C \mathrm{e}^{-A^T t}$$

其中矩阵 C 由条件 $x(t_1) = x_1$ 来确定。

习题 5.4

1. 设 $L(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_n)$ 是 $2n$ 个变元的连续可微函数，推导泛函

$$J(x_1, x_2, \cdots, x_n) = \int_a^b L(x_1, x_2, \cdots, x_n, \dot{x}_1, \dot{x}_2, \cdots, \dot{x}_n)\mathrm{d}t$$

达到极值时所满足的 Euler – Lagrange 方程。

2. 在 $C^1[t_1, t_2]$ 中确定一条通过点 $(t_1, x_1), (t_2, x_2)$ 的曲线，使泛函

$$J(x) = \int_{t_1}^{t_2} \sqrt{1 + [\dot{x}(t)]^2}\mathrm{d}t$$

取极小。

3. 在 $C^2[0,2]$ 中求一条曲线使泛函

$$J(x) = \int_0^2 \left(\frac{1}{2}\dot{x}^2 + x\dot{x} + \dot{x} + x\right)\mathrm{d}t$$

达到极小。

4. 求具有约束条件 $\dot{x} = x + u$ 及初始条件 $x(0) = x_0$ 下泛函 $J = \int_0^1 (x^2 + u^2)\mathrm{d}t$ 达到极值的 $x(t)$ 及 $u(t)$。

5. 求 $u \in L^2[0,1]$ 使 $J(u) = \min \dfrac{1}{2}\int_0^1 |u(t)|^2\mathrm{d}t$ 满足

$$\dot{x}(t) + x(t) + u(t) = 0$$

且 $x(0) = 0, x(1) = 1$。

5.5 凸集、凸泛函与最优化

凸集、凸泛函的理论，由于其在概率论、非线性规划以及现代控制等理论中的广泛应用而成为现代数学中一个极为重要的分支——凸分析。本节简要介绍关于凸集、凸泛函的一些最基本概念和性质。

5.5.1 凸集的分离性定理

我们曾在第三章3.3节中介绍过著名的Hahn－Banach定理，这里讨论的是关于凸集的Hahn－Banach定理的几何形式。

首先回顾一下曾在第二章线性空间这一节中定义过的凸集概念。设 X 是一个线性空间，$A \subset X$ 称为凸集，如果对任意 $x, y \in A$，$\alpha \in [0,1]$ 有

$$\alpha x + (1-\alpha)y \in A。$$

很容易验证任意多个凸集的交仍是凸集，但凸集对并运算不封闭，即 A_1, A_2 是两个凸集，而 $A_1 \cup A_2$ 未必是凸集。对于线性空间 X 中的任何子集 B，一定存在包含 B 的最小凸集，即所有包含 B 的凸集的交：

$$\bigcap \{A : A\text{ 是凸集且 } A \supset B\}。$$

记此集合为 $Co(B)$，称为 B 的凸包。

例1 设 $B = \{b_1, b_2, \cdots, b_n\} \subset X$，则

$$Co(B) = \left\{\sum_{i=1}^{n} \lambda_i b_i : \lambda_i \geqslant 0, \sum_{i=1}^{n} \lambda_i = 1\right\}。$$

我们将例1的证明留给读者。

【定义1】 设 X 是一个线性空间，X_1 是 X 的线性子空间，称

$$L = \{x : x = x_0 + y, y \in X_1\} = x_0 + X_1$$

为 X 中一个线性流形，又称线性流形 L 为 X 的一个超平面，如果 X_1 是比 X 低一维的线性子空间，即对任意 $x \in X \setminus X_1$，由 x 和 X_1 张成的线性空间就是 X，此时常称 X_1 为 X 的超子空间。

注　L 为闭线性流形当且仅当 X_1 为闭线性子空间。

【定理 1】　设 X 是赋范线性空间，一个闭的线性流形 L 是 X 的超平面的充要条件是存在非零线性泛函 $x^* \in X^*$ 及某个实数 $r \in \mathbf{R}$ 有

$$L = \{x \in X : x^*(x) = r\}。$$

证　必要性。设 x^* 及 r 如上式所要求。取 $x_0 \in L$，则 $x^*(x_0) = r$，那么对任意 $x \in L$ 有

$$x^*(x - x_0) = x^*(x) - x^*(x_0) = r - r = 0。$$

若记 $X_1 = Ker(x^*) = \{y : x^*(y) = 0\}$，那么 $x - x_0 \in X_1$，若令 $y = x - x_0$，则 $x = x_0 + y (y \in X_1)$，于是

$$L = x_0 + X_1。$$

由于 x^* 是连续线性泛函，所以 X_1 是闭子空间，即 L 是闭的。任取 $x_1 \in X \setminus X_1$，那么 $x^*(x_1) \neq 0$。对任意 $x \in X$，注意到

$$x^*\left(x - \frac{x^*(x)}{x^*(x_1)}x_1\right) = x^*(x) - \frac{x^*(x)}{x^*(x_1)}x^*(x_1) = 0,$$

所以 $x - \frac{x^*(x)}{x^*(x_1)}x_1 \in X_1$，即

$$x \in \frac{x^*(x)}{x^*(x_1)}x_1 + X_1。$$

这说明 X_1 与 x_1 张成的线性空间就是 X。

充分性。设 $L = x_0 + X_1$，X_1 是比 X 低一维的闭线性子空间。据第三章 3.3 节中 Hahn – Banach 定理的推论 1，存在 $x^* \in X^* (x^* \neq \theta)$

使对 $x \in X_1$ 有

$$x^*(x) = 0。$$

令 $r = x^*(x_0)$,那么

$$L = \{x : x^*(x) = r\}。$$

【定义 2】 设 X 是一个赋范线性空间,$A, B \subset X$。称超平面 $L = H_{x^*}^r = \{x : x^*(x) = r\}$ 分离(严格分离)A 与 B,如果满足

$$\begin{cases} x^*(x) \geqslant r (x \in A) \\ x^*(x) \leqslant r (x \in B) \end{cases} \left[\text{相应地} \begin{cases} x^*(x) > r \quad (x \in A) \\ x^*(x) < r \quad (x \in B) \end{cases} \right]。$$

【引理 1】 设 X 是实赋范线性空间,$B \subset X$ 是凸集且 $\exists \delta > 0$,使 $B_\delta(\theta) = \{x : \|x\| < \delta\} \subset B$,则可定义泛函 $P: X \to [0, +\infty)$ 为

$$P(x) = \inf\{\lambda : \lambda > 0, x \in \lambda B\},$$

且 P 满足:

(1) $P(\theta) = 0$;

(2) $P(\alpha x) = \alpha P(x) \quad (\alpha > 0, x \in X)$;

(3) $P(x + y) \leqslant P(x) + P(y) \quad (x, y \in X)$。

证 由对任意 $\lambda > 0, \theta \in \lambda B(\theta \in B)$ 得 $P(\theta) = 0$。对 $\forall \varepsilon > 0$,由于

$$x \in (P(x) + \varepsilon)B,$$

所以

$$\alpha x \in (P(x) + \varepsilon)\alpha B。$$

于是 $P(\alpha x) \leqslant \alpha P(x) + \varepsilon\alpha$,再由 ε 的任意性得 $P(\alpha x) \leqslant \alpha P(x)$。令 $\alpha' = \dfrac{1}{\alpha}, y = \alpha x$ 利用 $P(\alpha' y) \leqslant \alpha' P(y)$ 得 $P(x) \leqslant \dfrac{1}{\alpha} P(\alpha x)$,即 $P(\alpha x) \geqslant \alpha P(x)$,故 $P(\alpha x) = \alpha P(x) \quad (\alpha > 0)$。再注意到

$$x \in (P(x) + \varepsilon)B, \quad y \in (P(y) + \varepsilon)B。$$

又 B 是凸集，因此

$$\frac{P(x)+\varepsilon}{P(x)+P(y)+2\varepsilon}B+\frac{P(y)+\varepsilon}{P(x)+P(y)+2\varepsilon}B\subset B,$$

所以

$$x+y\in(P(x)+\varepsilon)B+(P(y)+\varepsilon)B=(P(x)+P(y)+2\varepsilon)$$
$$\left[\frac{P(x)+\varepsilon}{P(x)+P(y)+2\varepsilon}B+\frac{P(y)+\varepsilon}{P(x)+P(y)+2\varepsilon}B\right]\subset$$
$$(P(x)+P(y)+2\varepsilon)B,$$

即 $P(x+y)\leqslant P(x)+P(y)+2\varepsilon$，由 ε 的任意性有(3) 成立。

【定理 1】 设 X 是一个实赋范线性空间，$A\subset X$ 是凸集，且有 $a\in A$ 及相应的 $\delta>0$，使 $B_\delta(a)=\{x:\|x-a\|<\delta\}\subset A$。若 $x_0\notin A$，则存在超平面 $L=H_{x^*}^r$ 分离 x_0 与 A。

证　记 $B=A-a=\{x-a:x\in A\}$，则 B 仍是凸集，且

$$B_\delta(\theta)=\{x:\|x\|<\delta\}\subset B。$$

根据引理 1 定义泛函 $P(x)=\inf\{\lambda>0:x\in\lambda B\}$，那么对任意 $x\in B$ 有 $P(x)\leqslant 1$。又 $x_0\notin A$，则 $x_1=x_0-a\notin B$，故 $P(x_1)\geqslant 1$，记 $M=\{\lambda x_1:\lambda\in\mathbf{R}\}$，则 M 是 X 的一维子空间。在 M 上定义线性泛函 $\hat{x}^*(\cdot)$ 为

$$\hat{x}^*(\lambda x_1)=\lambda P(x_1)。$$

显然 $\hat{x}^*(\cdot)$ 是 M 上的连续线性泛函，且 $\hat{x}^*(x)\leqslant P(x)(x\in M)$。根据 Hahn – Banach 定理，$\hat{x}^*(\cdot)$ 可延拓成 X 上连续线性泛函 $x^*(\cdot)$ 且保持 $x^*(x)\leqslant P(x)(x\in X)$。对 $x\in B$，由于

$$x^*(x)\leqslant P(x)\leqslant 1,$$

而 $x^*(x_1)=\hat{x}^*(x_1)=P(x_1)\geqslant 1$。令 $x^*(a)=r-1$，则对 $x\in A$ 有

$$x^*(x)=x^*(x-a)+x^*(a)\leqslant 1+r-1=r,$$

但

$$x^*(x_0) = x^*(x_1) + x^*(a) \geq 1 + r - 1 = r。$$

故超平面 $H_{x^*}^r$ 分离点 x_0 与 A。

注　若 x_0 是 B 的内点，即 $x_0 \in \mathring{B}$，那么由引理 1 定义的泛函 P 满足 $P(x_0) < 1$。事实上，由于 x_0 是 B 的内点，可取 $\alpha_0 > 1$，使 $\alpha_0 x_0 \in B$，于是 $P(\alpha_0 x_0) \leq 1$，因此，$P(x_0) = \dfrac{1}{\alpha_0}P(\alpha_0 x_0) \leq \dfrac{1}{\alpha_0} < 1$。因此定理 1 的超平面 $H_{x^*}^r$ 可进一步满足若 $x \in \mathring{A}$ 有

$$x^*(x) < r。$$

【定理 2】　设 X 是实赋范线性空间，A 是闭凸集。若 $x_0 \notin A$，则存在超平面 $H_{x^*}^r$ 满足

$x^*(x_0) \geq r$，且对任意 $x \in A$，$x^*(x) < r$。

证　记 x_0 到 A 的距离 $d = \inf\{\| x_0 - a \| : a \in A\}$。因 A 是闭集，所以 $d > 0$，令集 G 为

$$G = \{x : d(x, A) < \frac{d}{2}\},$$

这里 $d(x, A)$ 表示 x 到 A 的距离即 $d(x, A) = \inf\{\| x - a \| : a \in A\}$。显然 G 是开集 $G \supset A$。我们来证 G 是凸集。设 $x_1, x_2 \in G$，则存在 $a_1, a_2 \in A$ 使

$$\| x_1 - a_1 \| < \frac{d}{2}, \quad \| x_2 - a_2 \| < \frac{d}{2}。$$

因 A 是凸集，那么 $\alpha a_1 + (1 - \alpha)a_2 \in A (\alpha \in [0,1])$。又

$$\| [\alpha x_1 + (1 - \alpha)x_2] - [\alpha a_1 + (1 - \alpha)]a_2 \| \leq$$

$$\alpha \| x_1 - a_1 \| + (1 - \alpha) \| x_2 - a_2 \| < \alpha \cdot \frac{d}{2} + (1 - \alpha)\frac{d}{2} = \frac{d}{2},$$

故 $d(\alpha x_1 + (1 - \alpha)x_2, A) < \dfrac{d}{2}$，即 $\alpha x_1 + (1 - \alpha)x_2 \in G$。由定理 1 及

注知存在超平面 $H_{x^*}^r$ 分离 x_0 点及 G，故

$x^*(x_0) \geqslant r$，且对任意 $x \in A$ 有 $x^*(x) < r$，这是因为 $x_0 \notin G$，且 $A \subset G$。

注　若在引理 1 中 $x_0 \notin B$，且 x_0 是 B^C 的内点，那么存在 $0 < \alpha_0 < 1$，使 $\alpha_0 x_0 \notin B$，故 $P(\alpha_0 x_0) \geqslant 1$，即 $P(x_0) = \frac{1}{\alpha_0}P(\alpha_0 x_0) \geqslant \frac{1}{\alpha_0} > 1$。因此，定理 2 的结论可加强为 $x^*(x_0) > r$，且对一切 $x \in A$ 有 $x^*(x) < r$，即超平面 $H_{x^*}^r$ 严格分离 $\{x_0\}$ 与 A。

【定理 3】　设 X 是实赋范线性空间，A，B 是 X 中的两个非空凸集，且 A 是紧集，B 是闭集。那么若 $A \cap B = \varnothing$，则存在超平面 $H_{x^*}^r$ 严格分离 A 与 B。

证　记 $C = A - B$，由于 $A \cap B = \varnothing$，那么 $\theta \notin C$。我们来证 C 是闭集。设 $x_n \in C$ 且 $x_n \to x$，于是有 $a_n \in A$，$b_n \in B$ 使 $x_n = a_n - b_n$。因 A 紧，从而 $\{a_n\}$ 有子列 $\{a_{n_k}\}$ 及 $a \in A$ 使 $a_{n_k} \to a$，又 $x_{n_k} \to x$，故 $b_{n_k} \to (a - x)$。又 B 是闭集，所以 $(a - x) \in B$，而 $x = a - (a - x)$，故 $x \in C$。据定理 2，存在超平面 $H_{x^*}^{r_1}$ 满足(见定理 2 注)：

$0 = x^*(\theta) > r_1$，对任意 $x \in C$，$x^*(x) < r_1$。

这样对任意 $a \in A$，$b \in B$ 有

$$x^*(a - b) < r_1 < 0。$$

即 $x^*(a) < r_1 + x^*(b)$；故

$$\sup_{a \in A} x^*(a) \leqslant r_1 + \inf_{b \in B} x^*(b),$$

更有

$$\sup_{a \in A} x^*(a) < \inf_{b \in B} x^*(b)。$$

取 r 满足 $\sup\limits_{a \in A} x^*(a) < r < \inf\limits_{b \in B} x^*(b)$，则超平面 $H_{x^*}^r$ 严格分离 A 与 B。

5.5.2 凸泛函

【定义 3】 设 X 是一个线性空间，$f: X \to (-\infty, +\infty]$ 称为凸泛函，如果对任意 $\{x_1, x_2, \cdots, x_n\} \subset X$，凸组合 $\sum\limits_{i=1}^{n} \lambda_i x_i (\lambda_i \geqslant 0, \sum\limits_{i=1}^{n} \lambda_i = 1)$ 满足

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)。$$

根据定义有：

【性质 1】 设 $f: X \to (-\infty, +\infty]$，则下述各条命题均等价。

(a) f 是凸泛函；

(b) 对任意 $x, y \in X, \alpha \in (0,1)$ 成立 $f(\alpha x - (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y)$；

(c) f 的上图 $E_p(f) = \{(x, \lambda) \in X \times \mathbf{R}: f(x) \leqslant \lambda\}$ 是 $X \times \mathbf{R}$ 中凸集。

证 (a)$\Rightarrow$(b) 显然。(b)$\Rightarrow$(c)，设 $(x, \lambda), (y, k) \in E_p(f)$，对 $\alpha \in (0,1)$，由于

$$\alpha(x, \lambda) + (1-\alpha)(y, k) = (\alpha x + (1-\alpha) y, \alpha\lambda + (1-\alpha) k)$$

及

$$f(\alpha x + (1-\alpha) y) \leqslant \alpha f(x) + (1-\alpha) f(y) \leqslant \alpha\lambda + (1-\alpha) k,$$

所以 $(\alpha x + (1-\alpha) y, \alpha\lambda + (1-\alpha) k) \in E_p(f)$，故 $E_p(f)$ 是凸集。

(c)$\Rightarrow$(a)，由于 $E_p(f)$ 是凸集，那么

$$\sum_{i=1}^{n} \lambda_i (x_i, f(x_i)) = \left(\sum_{i=1}^{n} \lambda_i x_i, \sum_{i=1}^{n} \lambda_i f(x_i)\right) \in E_p(f)$$

$$\left(\sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geqslant 0\right)。$$

因此 $f\left(\sum_{i=1}^{n}\lambda_i x_i\right)\leqslant\sum_{i=1}^{n}\lambda_i f(x_i)$。

【性质 2】　设 $f,g,f_i:X\to(-\infty,+\infty]$ 均是凸泛函,则

(a)$f+g$ 是凸泛函;

(b) 对 $\alpha>0,\alpha f$ 是凸泛函;

(c) 如果 $T:Y\to X$(Y 是另一线性空间)是线性算子,那么 $foT(x)=f(Tx)$ 是凸泛函;

(d) 如果 $\varphi:\mathbf{R}\to\mathbf{R}$ 单调增凸函数,那么 φof 是凸泛函;

(e)$\sup\{f_i(x):i\in I\}$ 是凸泛函,即一族凸泛函的上确界泛函仍是凸泛函;

(f) 对任意 $\lambda\in\mathbf{R}$,水平集

$$S(f,\lambda)=\{x\in X:f(x)\leqslant\lambda\}$$

是 X 中凸集。

我们把这些性质的证明留给读者。

【定义 4】　称 $f:X\to(-\infty,+\infty]$ 是严格凸的,如果对任意 f 的有效定义域 $\mathrm{Dom}(f)=\{x\in X:f(x)<+\infty\}$ 内的两点 x,y,满足 $f\left(\frac{x+y}{2}\right)<\frac{f(x)+f(y)}{2}$。

【性质 3】　$f:X\to(-\infty,+\infty]$ 是凸泛函,记

$$M=\{x\in X:f(x)=\inf\{f(y):y\in X\}\},$$

则 M 是凸集,即 f 达到最小值的点组成的集是凸的。特别,若 f 是严格凸的,则 M 最多含有一个点。

证　记 $\lambda=\inf\{f(y):y\in X\}$。设 $x,y\in M$,则 $f(x)=f(y)=\lambda$,对 $\alpha\in(0,1)$,有

$$\lambda\leqslant f(\alpha x+(1-\alpha)y)\leqslant\alpha f(x)+(1-\alpha)y=$$
$$\alpha\lambda+(1-\alpha)\lambda=\lambda,$$

故 $f(\alpha x+(1-\alpha)y)=\lambda$，即 $\alpha x+(1-\alpha)y\in M$，$M$ 为凸集。若 f 是严格凸的，且 $x\neq y$，$x,y\in M$，那么

$$\lambda\leqslant f\left(\frac{x+y}{2}\right)<\frac{f(x)+f(y)}{2}=\lambda,$$

这不可能，从而 M 最多含有一点。

【性质 4】 设 X,Y 是两个线性空间，$g:X\times Y\to(-\infty,+\infty]$ 是凸泛函，那么由下式定义的 $f:X\to(-\infty,+\infty]$

$$f(x)=\inf\{g(x,y):y\in Y\}$$

也是凸泛函。

证 设 $x,x'\in X$，若 $f(x)=+\infty$，则显然有

$$f(\alpha x+(1-\alpha)x')\leqslant\alpha f(x)+(1-\alpha)f(x')(\alpha\in(0,1)),$$

因此，不妨设 $x,x'\in\mathrm{Dom}(f)$。由定义，对任意 $\varepsilon>0$，存在 y 及 y' 满足

$$g(x,y)<f(x)+\varepsilon,\quad g(x',y')<f(x')+\varepsilon,$$

那么

$$g(\alpha x+(1-\alpha)x',\alpha y+(1-\alpha)y')\leqslant\alpha g(x,y)+(1-\alpha)g(x',y')<\alpha f(x)+(1-\alpha)f(x')+\varepsilon,$$

又

$$f(\alpha x+(1-\alpha)x')\leqslant g(\alpha x+(1-\alpha)x',\alpha y+(1-\alpha)y')。$$

所以

$$f(\alpha x+(1-\alpha)x')<\alpha f(x)+(1-\alpha)f(x')+\varepsilon,$$

由 ε 的任意性得 $f(\alpha x+(1-\alpha)x')\leqslant\alpha f(x)+(1-\alpha)f(x')$。

【定义 5】 设 X 是赋范线性空间，$\Omega\subset X$ 是开集。$f:\Omega\to R$ 是泛函，称 f 在 x_0 点是李普西兹的，是指 $\exists\delta>0$ 及常数 $k>0$ 满足当 x，$y\in B_\delta(x_0)$ 时有 $|f(x)-f(y)|\leqslant L\|x-y\|$，这里 $B_\delta(x_0)=\{y:\|y-x_0\|<\delta\}\subset\Omega$，又称 f 在 Ω 内是局部李普西兹的，是指 f 在 Ω

的每点都是李普西兹的。

【定理 4】　设 X 是赋范线性空间，$f: X \to (-\infty, +\infty]$ 是凸泛函，那么下面的两个命题是等价的。

(a) 存在 $\delta > 0, x_0 \in X$ 及常数 $M > 0$，使 $B_\delta(x_0) \subset \mathrm{Dom}(f)$，且对一切 $x \in B_\delta(x_0)$ 有 $f(x) \leqslant M$；

(b) f 在 $\mathrm{Dom}(f)$ 的内部是局部李普西兹的。

由于证明较复杂，这里略去证明。

5.5.3　Hilbert 空间的最优化

【定义 6】　设 X 是一个赋范线性空间，$f: X \to (-\infty, +\infty]$，称 f 在 $x_0(x_0 \in X)$ 点是下(上)半连续的，如果对任意的 $\{x_n\} \subset X$，且 $x_n \to x_0$ 有

$$f(x_0) \leqslant \varliminf_{n\to\infty} f(x_n)\text{（相应地 } f(x_0) \geqslant \varlimsup_{n\to\infty} f(x_n)\text{）。}$$

注　下(上)半连续的定义等价于：对任意 $\varepsilon > 0$，存在 $\delta > 0$，当满足 $x \in B_\delta(x_0)$ 时有

$$f(x) \geqslant f(x_0) - \varepsilon \quad \text{（相应地 } f(x) \leqslant f(x_0) + \varepsilon\text{）。}$$

$f(x)$ 在 x_0 点连续当且仅当 $f(x)$ 在 x_0 点既上半连续又下半连续。

关于下半连续的一些基本性质留作习题。

下面我们研究 Hilbert 空间中的下述最优化问题：

$$f_\lambda(x) = \inf_{y \in X}\left[f(y) + \frac{1}{2\lambda}\|y - x\|^2\right] \tag{5.15}$$

这里 f 是 Hilbert 空间 X 上定义的凸泛函，λ 是正参数。

【定理 5】　设 X 是 Hilbert 空间，$f: X \to (-\infty, +\infty]$ 是下半连续的凸泛函，那么式(5.15) 有惟一解，用 $J_\lambda x$ 表示，即

$$f_\lambda(x) = f(J_\lambda x) + \frac{1}{2\lambda}\|J_\lambda x - x\|^2,$$

并且满足下面的变分不等式

$$\frac{1}{\lambda}\langle J_\lambda x - x, J_\lambda x - y\rangle + f(J_\lambda x) - f(y) \leqslant 0\ (x, y \in X)。\qquad (5.16)$$

根据这个定理我们可直接推出 Hilbert 空间的最佳逼近定理。

【定理 6】 设 K 是 Hilbert 空间 X 中的一个闭凸子集，那么存在惟一的一个 Jx 满足：

(1) $Jx \in K$;

(2) $\| x - Jx \| = d(x, K) = \inf\{\| x - y \| : y \in K\}$,

且满足变分不等式

$$\langle Jx - x, Jx - y\rangle \leqslant 0 \quad (y \in K)。$$

证 取 $f(x)$ 为 K 的指示函数，即

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

则 $f(x)$ 是凸泛函，又 K 闭，则 $f(x)$ 是下半连续的，又 $f_\lambda(x) = \frac{1}{2\lambda}d^2(x, K)$，从而 $Jx \in K$，且

$$d(x, K) = \| x - Jx \|。$$

再由变分不等式(5.16)，对 $y \in K$ 成立

$$\langle Jx - x, Jx - y\rangle \leqslant 0。$$

【定理 7】 (Von Neumann 极小极大原理) 设 X, Y 是两个赋范空间，$A \subset X, B \subset Y$ 均是非空紧凸子集，$f: A \times B \to \mathbf{R}$ 满足：

(1) 对任意 $y \in B$ 固定，$f(\cdot, y): A \to \mathbf{R}$ 是上半连续的凹泛函；

(2) 对任意 $x \in A$ 固定，$f(x, \cdot): B \to \mathbf{R}$ 是下半连续的凸泛函，

则必存在 $(x_0, y_0) \in A \times B$ 满足

$$f(x, y_0) \leqslant f(x_0, y_0) \leqslant f(x_0, y) \quad (x \in A, y \in B)$$

注 称泛函 f 是凹的，如果 $-f$ 是凸的。通常把定理 7 也称为 Von

Neumann 鞍点存在定理，即定理中表述的(x_0, y_0)是鞍点。这个定理在最优化、对策论及经济数学中有广泛的应用。

习题 5.5

1. 设 X 是赋范线性空间，X^* 是 X 的共轭空间，对 $x^* \in X^*$ 及 $\lambda \in \mathbf{R}$，证明集合

$$\{x \in X : x^*(x) \leqslant \lambda\}$$

是 X 中闭凸集。

2. 设 X, X^* 如上题，$A \subset X$ 是非空有界集。定义 X^* 上泛函为 $\sigma_A(x^*) = \sup\limits_{a \in A} x^*(a)$，证明：存在常数 $M > 0$，使得

$$|\sigma_A(x^*) - \sigma_A(y^*)| \leqslant M \| x^* - y^* \| \quad (x^*, y^* \in X^*)。$$

3. X, X^* 如上题，$A \subset X$ 是闭凸集，$x_0 \in X$，证明：如果对每个 $x^* \in X^*$，成立 $\sigma_A(x^*) \geqslant x^*(x_0)$，则 $x_0 \in A$。

4. 设 X 是一个线性空间，$A \subset X$，证明：

$$Co(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : \lambda_i \geqslant 0, \sum_{i=1}^{n} \lambda_i = 1, x_i \in A, n = 1,2,\cdots \right\}。$$

5. 证明性质 2。

6. 设 X 是赋范线性空间，$f: X \to (-\infty, +\infty]$ 下半连续，证明：

(1) f 的上图 $E_p(f)$ 是 $X \times \mathbf{R}$ 中的闭集；

(2) 对 $\lambda \in \mathbf{R}$，水平集 $S(f, \lambda)$ 是 X 中闭集。

7. 设 X 是赋范线性空间，$f, g, f_i (i \in I)$ 均是 X 上下半连续泛函，证明：

(1) $f + g$ 是下半连续的；

(2) 对 $\alpha > 0$，αf 是下半连续的；

(3) $\inf(f, g)$ 是下半连续的；

(4) $\sup\limits_{i \in I} f_i$ 是下半连续的。

第 6 章　广义函数简介

函数是古典分析中的基本概念之一，然而这样的一个基本概念，在近代科学技术的发展中逐渐不够用了。下面用几个例子加以说明。

例 1(脉冲)　20 世纪初，Heaviside 在解电路方程时，提出了一种运算方法，称之为算子演算。这套算法要求对如下函数

$$h(x)=\begin{cases}1, & x\geqslant 0\\ 0, & x<0\end{cases}$$

求导数，并把导数记为 $\delta(x)$。但按照经典分析的理论，$h(x)$ 并不可导，因此 $\delta(x)$ 不可能是普通意义下的函数，它除了作为一个记号进行形式演算外，在数学上是没有意义的。但是，这个 $\delta(x)$ 在实际中却是有意义的，又代表一种理想化的"瞬时"单位脉冲。

例 2(Dirac 符号)　在微观世界中，把可观测到的物质的状态用波函数来描述，最简单的波函数具有形式 $e^{i\lambda x}(x\in(-\infty,+\infty))$，$\lambda$ 是实参数，并考虑如下形式的积分

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\lambda x}dx。$$

这种积分按 Cauchy 积分来定义，即

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\lambda x}dx=\lim_{n\to\infty}\frac{1}{2\pi}\int_{-n}^{n}e^{i\lambda x}=\lim_{n\to\infty}\frac{1}{\pi}\frac{\sin n\lambda}{\lambda}。$$

显然，这个极限在普通意义下不存在。然而，物理学家认为这个极限是前面所提到的 $\delta(\lambda)$，并认为是 Dirac 符号。特别，在量子力学中，进一步发展了不少关于 $\delta(\lambda)$ 的运算法则，并广泛地使用。

例 3(广义微商)　在数学本身的发展中,也时常要求冲破古典分析中对一些基本运算使用范围所加的限制。20 世纪 30 年代, Sobolev 为了确定微分方程解的存在性、惟一性问题,通过分部积分公式,推广了函数可微性的概念,建立了广义微商理论,形成了以他的名字命名的 Sobolev 空间理论。这标志着现代微分方程理论的诞生。

基于上述原因,扩充函数概念,为广义函数寻找坚实的数学基础,对数学家提出了新的挑战。20 世纪 40 年代, Schwartz 完成了这一艰巨的任务,创立了广义函数的系统理论,并因此而获得 1950 年数学最高奖 —— 菲尔兹奖。

6.1　基本函数空间与广义函数

6.1.1　基本函数空间

把普通函数视为某类函数空间上的线性泛函是推广函数概念的一条行之有效的途径。广义函数正是定义在一类性质很好的函数空间上的线性泛函。这类函数空间称为基本函数空间。

在引进基本函数空间之前,先介绍一些记号和术语。

对于欧氏空间 $\mathbf{R}^n$, $x=(x_1,x_2,\cdots,x_n)$ 表示 $\mathbf{R}^n$ 中的点,范数 $\|x\|=(x_1^2+x_2^2+\cdots+x_n^2)^{\frac{1}{2}}$。设 $p_1,p_2,\cdots,p_n$ 为 n 个非负整数,有序数组 $p=(p_1,p_2,\cdots,p_n)$ 称为多重指标。$|p|=p_1+p_2+\cdots+p_n$。对于多重指标 p,引进偏微分算子

$$D^p=\frac{\partial^{|p|}}{\partial x_1^{p_1}\partial x_2^{p_2}\cdots\partial x_n^{p_n}}。$$

$\Omega \subset \mathbf{R}^n$ 是非空开集，$\bar{\Omega}$ 是Ω 的闭包。$C(\bar{\Omega})$ 表示在 $\bar{\Omega}$ 上定义的连续函数全体组成的线性空间。对于任何非负整数 k，记 $C^k(\bar{\Omega})$ 表示全体在 Ω 内有 k 次连续可微的偏导数，且在 $\bar{\Omega}$ 上连续的函数组成的线性空间，特别 $C^0(\bar{\Omega}) = C(\bar{\Omega})$。设 $\varphi \in C(\bar{\Omega})$，$\varphi$ 的支集是集合

$$\{x \in \Omega : \varphi(x) \neq 0\}。$$

在 Ω 内的闭包，并记为 $\mathrm{supp}\varphi$，即

$$\mathrm{supp}\varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}。$$

$C_0^k(\Omega)$ 表示 $C^k(\bar{\Omega})$ 中满足支集是 Ω 内紧集（有界闭集）的所有函数组成的空间，记

$$C_0^\infty(\Omega) = \bigcap_{k=0}^{\infty} C_0^k(\Omega),$$

即表示支集是 Ω 内紧集的无穷次可微函数全体。显然，下面的包含关系成立：

$$C_0^\infty(\Omega) \subset \cdots \subset C_0^{k+1}(\Omega) \subset C_0^k(\Omega) \subset \cdots \subset C_0(\Omega)。$$

例 1 设 $\mathbf{R}^n$ 上定义的函数为

$$j(x) = \begin{cases} C_n e^{-\frac{1}{1-\|x\|^2}}, & \|x\| < 1; \\ 0, & \|x\| \geqslant 1。\end{cases}$$

这里 C_n 是依赖于维数 n 的常数，即

$$C_n = \left(\int_{\|x\| \leqslant 1} e^{-\frac{1}{1-\|x\|^2}} dx\right)^{-1}。$$

那么 $j(x)$ 是无穷次连续可微的，且 $\mathrm{supp} j \subset \{x \in \mathbf{R}^n : \|x\| \leqslant 1\}$，$\int_{\mathbf{R}^n} j(x) dx = 1$，因此 $j(x) \in C_0^\infty(\mathbf{R}^n)$。从 $j(x)$ 出发，我们可以构造出许多 $C_0^\infty(\mathbf{R}^n)$ 中的函数。下面我们来构造对任何非空开集 Ω，$C_0^\infty(\Omega)$ 中的函数。为此，对任意 $\delta > 0$，记

$$j_\delta(x) = \frac{1}{\delta^n} j\left(\frac{x}{\delta}\right),$$

那么 $j_\delta \in C_0^\infty(\mathbf{R}^n)$。

【定理 1】 设 $u(x)$ 是 Ω 上定义的一个可积函数，并且在 Ω 的一个紧集 K 外恒为零，则当 $\delta > 0$ 充分小时

$$u_\delta(x) = \int_\Omega u(y) j_\delta(x - y)\mathrm{d}y$$

是 $C_0^\infty(\mathbf{R}^n)$ 中的函数。

证　记 $K_\delta = \{x \in \mathbf{R}^n : \mathrm{dist}(x, K) \leqslant \delta\}$，这里 $\mathrm{dist}(x, K)$ 表示 x 到 K 的距离，当 δ 充分小时，$K_\delta \subset \Omega$，当 $x \notin K_\delta$ 时，对一切 $y \in K$ 均有 $\| x - y \| > \delta$，于是 $j_\delta(x - y) = 0$。

$$u_\delta(x) = \int_\Omega u(y) j_\delta(x - y)\mathrm{d}y = \int_K u(y) j_\delta(x - y)\mathrm{d}y = 0,$$

因此 $\mathrm{supp}\, u_\delta \subset K_\delta$，而

$$\frac{\partial u_\delta}{\partial x_1} = \lim_{h \to 0}\int_\Omega \frac{1}{h}[j_\delta(x + he_1 - y) - j_\delta(x - y)]u(y)\mathrm{d}y = \lim_{h \to 0}\int_\Omega \frac{\partial}{\partial x_1} j_\delta(x + \theta h e_1 - y)u(y)\mathrm{d}y。$$

上式利用了微分中值定理，$\theta \in (0,1)$，$e_1 = (1, 0, \cdots, 0) \in \mathbf{R}^n$，又 j_δ 是连续可微函数，因此存在 $M > 0$ 使

$$\left| \frac{\partial}{\partial x_1} j_\delta(x) \right| \leqslant M \quad (\forall x \in \mathbf{R}^n)。$$

应用 Lebesgue 控制收敛定理得

$$\frac{\partial u_\delta}{\partial x_1} = \lim_{h \to 0}\int_\Omega \frac{\partial}{\partial x_1} j_\delta(x + \theta h e_1 - y)u(y)\mathrm{d}y = \int_\Omega \frac{\partial}{\partial x_1} j_\delta(x - y)u(y)\mathrm{d}y。$$

注意到 $j_\delta \in C_0^\infty(\mathbf{R}^n)$，对任何多重指标 $p = (p_1, p_2, \cdots, p_n)$ 重复上述

过程可得到

$$D^p u_\delta = \int_\Omega D^p j_\delta(x-y)u(y)\mathrm{d}y,$$

于是 $u_\delta \in C_0^\infty(\Omega)$。

下面我们在 $C_0^\infty(\Omega)$ 上引进收敛的概念。

【定义 1】 设 $\{\varphi_i\} \subset C_0^\infty(\Omega), \varphi \in C_0^\infty(\Omega)$，如果满足下列条件：

(1) 存在一个紧集 $K \subset \Omega$ 使得

$$\operatorname{supp}(\varphi_j) \subset K(j=1,2,\cdots), \quad \operatorname{supp}(\varphi) \subset K;$$

(2) 对于任意多重指标 $p=(p_1,p_2,\cdots,p_n)$，函数列 $\{D^p\varphi_j\}$ 在 K 上一致收敛于 $D^p\varphi$，即

$$\max_{x\in K} |D^p\varphi_j(x) - D^p\varphi(x)| \to 0 (j\to\infty),$$

则称 $\{\varphi_i\}$ 收敛于 φ，记为 $\varphi_j \xrightarrow{D} \varphi$，而称 $C_0^\infty(\Omega)$ 按照收敛概念及线性运算为基本函数空间，并记为 $\mathscr{D}(\Omega)$，在 Ω 明确时，可简写为 $\mathscr{D}$。

根据 $\mathscr{D}$ 中收敛概念的定义，容易证明

① 设 $\{\varphi_j\}, \{\Psi_j\} \subset \mathscr{D}, \varphi, \Psi \in \mathscr{D}$，如果

$$\varphi_j \xrightarrow{\mathscr{D}} \varphi, \quad \Psi_j \xrightarrow{\mathscr{D}} \Psi,$$

则对任何数 α, β 有

$$\alpha\varphi_j + \beta\Psi_j \xrightarrow{\mathscr{D}} \alpha\varphi + \beta\Psi,$$

这说明 $\mathscr{D}$ 中的线性运算关于收敛概念是连续的。

② 对任一多重指标 p，$\mathscr{P}: \mathscr{D} \to \mathscr{D}$ 这一线性映射是连续的，即 $\{\varphi_j\} \subset \mathscr{D}, \varphi \in \mathscr{D}$，则若

$$\varphi_j \xrightarrow{\mathscr{D}} \varphi$$

那么 $D^p\varphi_j \xrightarrow{\mathscr{D}} D^p\varphi$。

【定义 2】　称$\{\varphi_j\}\subset\mathscr{D}$为 Cauchy 列，如果满足

(1) 存在紧集 K 使 $\mathrm{supp}(\varphi_j)\subset K(j=1,2,\cdots)$；

(2) 对 $\forall\varepsilon>0$，及多重指标 p，存在自然数 N，使当 $j_1,j_2\geqslant N$ 时有

$$\max_{x\in K}|D^p\varphi_{j_1}(x)-D^p\varphi_{j_2}(x)|<\varepsilon。$$

【定理 2】　$\mathscr{D}$是完备的，即若$\{\varphi_j\}\subset\mathscr{D}$是任意一个 Cauchy 列，则存在 $\varphi\in\mathscr{D}$，使得 $\varphi_j\xrightarrow{\mathscr{D}}\varphi$。

证明留作习题。

6.1.2　广义函数的基本概念

【定义 3】　$\mathscr{D}$上的一切线性连续泛函，称为广义函数，即$\mathscr{D}$上的广义函数 f 满足

(1) 线性：对任何数 α、β 及 $\varphi_1,\varphi_2\in\mathscr{D}$有

$$f(\alpha\varphi_1+\beta\varphi_2)=\alpha f(\varphi_1)+\beta f(\varphi_2);$$

(2) 连续：设$\{\varphi_j\}\subset\mathscr{D}$，$\varphi\in\mathscr{D}$，若 $\varphi_j\xrightarrow{\mathscr{D}}\varphi$，则有

$$f(\varphi_j)\to f(\varphi)(j\to\infty)。$$

一切广义函数所组成的集合，记作$\mathscr{D}'$。

例 2　δ - 函数，设 $\Omega\subset\mathbf{R}^n$ 是非空开集，$a\in\Omega$，对于任意 $\varphi\in\mathscr{D}$，定义

$$\delta_a(\varphi)=\varphi(a)$$

则 δ_a 是广义函数。

证　显然 δ_a 是$\mathscr{D}$上的线性泛函。设$\{\varphi_j\}\subset\mathscr{D}$，$\varphi\in\mathscr{D}$，若 $\varphi_j\xrightarrow{\mathscr{D}}\varphi$，则更有

$$|\varphi_j(a)-\varphi(a)|\to 0(j\to\infty)$$

从而

$$|\delta_a(\varphi_j)-\delta_a(\varphi)|=|\varphi_j(a)-\varphi(a)|\to 0(j\to\infty),$$

即 δ_a 在 $\mathscr{D}$ 上是连续的，所以 δ_a 是一个广义函数，称 δ_a 为集中在点 a 的 Dirac 广义函数，简称为 δ - 函数，特别，当 $a=\theta=(0,0,\cdots,0)$（$\mathbf{R}^n$ 中零元素）时，δ_θ 记为 δ。

例 3 设 $\Omega\subset\mathbf{R}^n$ 是非空开集，$f(x)$ 是 Ω 上定义的一个局部可积函数，即对于 Ω 的任何紧子集 K，积分

$$\int_K |f(x)|\,\mathrm{d}x<+\infty,$$

通过局部可积函数 f 定义 $\mathscr{D}$ 上的泛函

$$f^*(\varphi)=\int_\Omega f(x)\varphi(x)\mathrm{d}x, \tag{6.1}$$

则 f^* 是广义函数。

证 由于 $f(x)$ 是局部可积的，那么对任何 $\varphi\in\mathscr{D}$，$f(x)\varphi(x)$ 在 $\mathrm{supp}(\varphi)$ 上可积，从而由式(6.1)定义的积分有意义。根据式(6.1)，显然 f^* 是线性的。设 $\{\varphi_j\}\subset\mathscr{D}$，$\varphi\in\mathscr{D}$ 且 $\varphi_j\xrightarrow{\mathscr{D}}\varphi$，于是存在紧集 K 使 $\mathrm{supp}(\varphi_j)\subset K$，$\mathrm{supp}(\varphi)\subset K$ 且 φ_j 在 K 上一致收敛于 φ。取常数 $M>0$，使 $\sup\limits_{x\in K}|\varphi_j(x)|\leqslant M$，那么由 Lebesgue 控制收敛定理有

$$\begin{aligned}f^*(\varphi_j)&=\int_\Omega f(x)\varphi_j(x)\mathrm{d}x=\\&\int_K f(x)\varphi_j(x)\mathrm{d}x\to\int_K f(x)\varphi(x)\mathrm{d}x=\\&\int_\Omega f(x)\varphi(x)\mathrm{d}x\ (j\to\infty),\end{aligned}$$

即 $f^*(\varphi_j)\to f^*(\varphi)(j\to\infty)$，这说明 f^* 连续，因此 f^* 是 $\mathscr{D}$ 上的广义函数。

记 $L_{\mathrm{Loc}}(\Omega)$ 为 Ω 上全部局部可积函数组成的集合，通过例 3 知

道，每个 $f\in L_{\mathrm{Loc}}(\Omega)$ 都对应一个广义函数 f^*，称这样的 f^* 为函数型广义函数。

【定理3】 映射 $T:L_{\mathrm{Loc}}(\Omega)\to\mathscr{D}'$ 定义为

$$T(f)=f^*,$$

则 T 是一对一线性映射。

由于证明较繁烦，这里略去。

通过定理3，我们可以把局部可积函数 f 与由 f 定义的广义函数 f^* 视为同一，这样局部可积函数是广义函数。

例4 考察在 $\mathbf{R}$ 上的函数

$$h(x)=\begin{cases}0, & 当\ x\in(-\infty,0);\\ 1, & 当\ x\in(0,+\infty)。\end{cases}$$

通常 $h(x)$ 称为 Heaviside 函数。显然，$h\in L_{\mathrm{Loc}}(\mathbf{R})$，于是它定义 $\mathscr{D}(R)$ 上的广义函数 h^* 为

$$h^*(\varphi)=\int_{-\infty}^{+\infty}h(x)\varphi(x)\mathrm{d}x=\int_0^{+\infty}\varphi(x)\mathrm{d}x。$$

是否每个广义函数都是函数型的，即对广义函数 $g^*\in\mathscr{D}'$，是否存在局部可积函数 f 使

$$g^*(\varphi)=\int_\Omega f(x)\varphi(x)\mathrm{d}x,$$

回答是否定的。也就是说 $\mathscr{D}'$ 中确实存在非函数型广义函数。

例5 δ_a 不是函数型的广义函数。

证 用反证法。设 δ_a 是函数型的，则存在一个定义于 Ω 上的局部可积函数 f 使得对一切 $\varphi\in\mathscr{D}$ 成立

$$\int_\Omega f(x)\varphi(x)\mathrm{d}x=\varphi(a)。\tag{6.2}$$

取正数 r 充分小，使 $B_r(a)=\{x:\|x-a\|<r\}\subset\Omega$，定义函数

$$\varphi_{a,r}(x)=\begin{cases}e^{\frac{r^2}{\|x-a\|^2-r^2}}, & \text{当 } x\in B_r(a);\\ 0 & \text{,当 } x\in\Omega\setminus B_r(a)\text{。}\end{cases}$$

显然，$\varphi_{a,r}\in\mathscr{D}$，由式(6.2)得

$$\int_\Omega f(x)\varphi_{a,r}(x)\mathrm{d}x=\varphi_{a,r}(a)=e^{-1}\text{。}\tag{6.3}$$

另一方面，$\lim\limits_{r\to0^+}\varphi_{a,r}=\varphi(x)\equiv0\ a\cdot e$，由 Lebesgue 控制收敛定理有

$$\lim_{r\to0^+}\int_\Omega f(x)\varphi_{a,r}(x)=0\text{。}\tag{6.4}$$

这样式(6.3)与式(6.4)矛盾，故 δ_a 不是函数型的。

【定理 4】 $f\in\mathscr{D}$ 当且仅当对任意紧集 $K\subset\Omega$，存在常数 C 及非负整数 m，使得当 $\mathrm{supp}\varphi\subset K$ 时有

$$|f(\varphi)|\leqslant C\sum_{|p|\leqslant m}\sup_{x\in K}|D^p\varphi|\ (\varphi\in\mathscr{D})\tag{6.5}$$

证　充分性。由式(6.5)知 f 是 $\mathscr{D}$ 上定义的连续线性泛函，因此 $f\in\mathscr{D}$。

必要性。用反证法。若不然，有紧集 K，使式(6.5)不成立。于是对任何自然数 j，存在函数 $\varphi_j\in\mathscr{D}$，且 $\mathrm{supp}(\varphi_j)\subset K$ 使

$$|f(\varphi_j)|>j\sum_{|p|\leqslant j}\sup_{x\in K}|D^p\varphi_j(x)|,\tag{6.6}$$

于是令

$$\hat{\varphi}_j=\frac{\varphi_j(x)}{j\sum\limits_{|p|\leqslant j}\sup\limits_{x\in K}|D^p\varphi_j(x)|},$$

则 $\hat{\varphi}_j\in\mathscr{D}$，且 $\mathrm{supp}\hat{\varphi}_j\subset K$，再由式(6.6)可得

$$|f(\hat{\varphi}_j)|>1,\tag{6.7}$$

又

$$\sum_{|p|\leqslant j}\sup_{x\in K}|D^p\hat{\varphi}_j(x)|=\frac{1}{j},$$

因此，$\hat{\varphi}_j \xrightarrow{\mathscr{D}} \varphi = 0$，于是 $f(\hat{\varphi}_j) \to f(\varphi) = 0$，这与式(6.7)矛盾。

在广义函数空间 $\mathscr{D}'$ 上规定加法与数乘运算：设 $f_1, f_2 \in \mathscr{D}', \lambda \in$ **R** 定义

$$(f_1 + f_2)(\varphi) = f_1(\varphi) + f_2(\varphi),$$
$$(\lambda f_1)(\varphi) = \lambda f_1(\varphi),$$

则很容易证明 $f_1 + f_2 \in \mathscr{D}', \lambda f_1 \in \mathscr{D}'$，因此 $\mathscr{D}'$ 是一个线性空间。

【定义 4】　设 $\{f_j\} \subset \mathscr{D}', f \in \mathscr{D}'$，如果对一切 $\varphi \in \mathscr{D}$ 成立

$$\lim_{0\to\infty} f_j(\varphi) = f(\varphi),$$

则称 $\{f_j\}$ 在 $\mathscr{D}'$ 中收敛于 f，记为 $f_j \xrightarrow{\mathscr{D}'} f$。$\mathscr{D}'$ 按照这种收敛概念，称为广义函数空间。

容易证明，$\mathscr{D}'$ 中加法与数乘运算关于收敛是连续的，即如果 $f_j \xrightarrow{\mathscr{D}'} f, g_j \xrightarrow{\mathscr{D}'} g$，则对任何数 α, β 有

$$\alpha f_j + \beta g_j \xrightarrow{\mathscr{D}'} \alpha f + \beta g。$$

例 6　在 **R** 上，函数列

$$f_j(x) = \frac{1}{\pi} \frac{\sin jx}{x} \quad (j = 1,2,\cdots)$$

是 $L_{Loc}(\mathbf{R})$ 中的函数列，从而可视为广义函数列，那么 $f_j \xrightarrow{\mathscr{D}'} \delta$。

证　由于 $\int_{-\infty}^{+\infty} \frac{\sin x}{x} \mathrm{d}x = \pi$，因此

$$\lim_{T\to+\infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\sin jx}{x} \mathrm{d}x = 1$$

对任意 $\varphi \in \mathscr{D}(\mathbf{R})$，存在 $T_0 > 0$，使 $\operatorname{supp}\varphi \subset [-T_0, T_0]$。那么当 $T > T_0$ 时

$$\int_{-\infty}^{+\infty} f_j(x)\varphi(x)\mathrm{d}x = \int_{-T}^{T} f_j(x)\varphi(x)\mathrm{d}x。$$

另一方面,对 $\forall \varepsilon > 0$,取 T_1 足够大,使当 $T > T_1$ 时有

$$\left|\frac{1}{\pi}\int_{-T}^{T}\frac{\sin jx}{x}\mathrm{d}x - 1\right| < \frac{\varepsilon}{2},$$

从而当 $T > \max\{T_0, T_1\}$ 时有

$$\begin{aligned}|f_j(\varphi) - \varphi(0)| \leqslant & \frac{1}{\pi}\left|\int_{-T}^{T}\frac{\sin jx}{x}[\varphi(x) - \varphi(0)]\mathrm{d}x\right| + \frac{\varepsilon}{2}|\varphi(0)| = \\ & \frac{1}{\pi}\left|\int_{0}^{T}\sin jx\,\frac{\varphi(x) + \varphi(-x) - 2\varphi(0)}{x}\mathrm{d}x\right| + \\ & \frac{\varepsilon}{2}|\varphi(0)|。\end{aligned}$$

对于固定 T,由于函数 $\hat{\varphi}(x) = \dfrac{\varphi(x) + \varphi(-x) - 2\varphi(0)}{x}$ 是 Riemann 可积的,因此由 Riemann 引理

$$\lim_{j\to\infty}\int_0^T \sin jx\hat{\varphi}(x)\mathrm{d}x = 0,$$

于是存在自然数 n_0,当 $j > n_0$ 时

$$\frac{1}{\pi}\left|\int_0^T \sin jx\hat{\varphi}(x)\mathrm{d}x\right| < \frac{\varepsilon}{2}。$$

因此

$$|f_j(\varphi) - \varphi(0)| = |f_j(\varphi) - \delta(\varphi)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\varphi(0)。$$

由 ε 的任意性得

$$f_j(\varphi) \longrightarrow \delta(\varphi),$$

故 $f_j \xrightarrow{\mathscr{D}} \delta$。

这个例子给出了关于本章例 2 中的 Dirac 符号的合理数学解释。

习题 6.1

1. 在 $\mathscr{D}$ 中证明:

(1) $\dfrac{1}{\pi}\dfrac{1}{x^2+\varepsilon^2}\longrightarrow\delta(x)\quad(\varepsilon\to 0^+)$,

(2) $\dfrac{1}{2\sqrt{\pi t}}e^{-\frac{x^2}{4t}}\longrightarrow\delta(x)\quad(t\to 0^+)$。

2. 设 $f_j(x)=(1+\dfrac{x}{j})^j(j=1,2,\cdots)$,证明:$f_j(x)\xrightarrow{\mathscr{D}'}e^x$。

3. 设 $\Omega\subset\mathbf{R}^n$ 是一个开集,$f_j(x)$ 是 Ω 上的一列局部可积函数,并且对任意紧集 $K\subset\Omega$,存在常数 M_K 使得

$$|f_j(x)|\leqslant M_K,$$

又 $f_j(x)\to f_0(x)\quad a\cdot e$,证明:$f_j^*\xrightarrow{\mathscr{D}'}f_0^*$。

4. 设 $\Omega\subset\mathbf{R}^n$ 是一个非空开集,$K\subset\Omega$ 是紧集。证明:存在函数 $\varphi\in C_0^\infty(\Omega)$,使得 $0\leqslant\varphi(x)\leqslant 1$,且在 K 上恒有 $\varphi(x)=1$。

5. 证明定理 2。

6.2　广义函数的导数及其性质

广义函数求导的思想来源于经典分析学中的分部积分,为此,先回顾一下分部积分的基本思想。

设 $\Omega=(a,b)$,f 和 φ 是 Ω 上定义的两个连续可微函数,若 $\mathrm{supp}\varphi$ 是 Ω 内的紧集,则 $\varphi(a)=\varphi(b)=0$,于是有

$$\int_\Omega f'(x)\varphi(x)\mathrm{d}x=-\int_\Omega f(x)\varphi'(x)\mathrm{d}x。$$

可见,利用分部积分可将一个函数的求导运算转化为对另一个函数的求导,广义函数的导数引入,就是遵循这一法则而得来的。

对任一多重指标 p,D^p 是一个由 $\mathscr{D}$ 到 $\mathscr{D}$ 的连续线性映射,因此可有:

【性质 1】　设 $f\in\mathscr{D}'$，定义 $g(\varphi)=-f\left(\frac{\partial\varphi}{\partial x_1}\right)(\varphi\in\mathscr{D})$，则 $g\in\mathscr{D}'$。

证　显然 g 是 $\mathscr{D}$ 上线性泛函，设 $\{\varphi_j\}\subset\mathscr{D}$，$\varphi\in\mathscr{D}$，若 $\varphi_j\xrightarrow{\mathscr{D}}\varphi$，那么 $\frac{\partial\varphi_j}{\partial x_1}\xrightarrow{\mathscr{D}}\frac{\partial\varphi}{\partial x_1}$，从而

$$f\left(\frac{\partial\varphi_j}{\partial x_1}\right)\longrightarrow f\left(\frac{\partial\varphi}{\partial x_1}\right),$$

因此 $g(\varphi_j)\to g(\varphi)$，故 $g\in\mathscr{D}'$。

【定义 1】　$f\in\mathscr{D}'$ 定义 f 对 x_1 的一阶偏导数为 g，并记为 $g=\frac{\partial f}{\partial x_1}$，则 g 仍然是广义函数。一般地对任意多重指标 p，定义 D^pf 为如下广义函数：

$$g_p(\varphi)=(-1)^{|p|}f(D^p\varphi),$$

即 $D^pf=g_p$。

从定义 1 可以看出，广义函数可进行无限次求导运算。

例 1　由 Heaviside 函数 h 所定义的函数

$$h(x)=\begin{cases}1, & x\geqslant 0;\\ 0, & x<0。\end{cases}$$

的广义导数 $h'=\delta$。

证　对任意 $\varphi\in\mathscr{D}(R)$，有

$$h'(\varphi)=-h(\varphi')=-\int_0^\infty\varphi'\mathrm{d}x=\varphi(0)=\delta(\varphi),$$

因此 $h'=\delta$。

注　如果 f 是一元广义函数（即 f 为 $\mathscr{D}(R)$ 上的连续线性泛函），则 f 的一阶、二阶导数等分别用 f'，f'' 等来表示。

例 2　证明：$|x|''=2\delta$。

证　对任意 $\varphi \in \mathscr{D}(R)$，取 $a > 0$，使 $\mathrm{supp}\varphi \subset (-a, a)$。那么有

$$|x|''(\varphi) = |x|(\varphi'') = \int_{-\infty}^{+\infty} |x| \varphi''(x)\mathrm{d}x =$$
$$\int_0^a x\varphi''(x)\mathrm{d}x - \int_{-a}^0 x\varphi''(x)\mathrm{d}x =$$
$$-\int_0^a \varphi'(x)\mathrm{d}x + \int_{-a}^0 \varphi'(x)\mathrm{d}x =$$
$$\varphi(0) + \varphi(0) = 2\varphi(0) = 2\delta(\varphi),$$

所以 $|x|'' = 2\delta$。

【定义 2】　设 $\Psi \in C^\infty(\Omega)$，$f \in \mathscr{D}'$，定义 $\mathscr{D}$上的泛函为

$$(\Psi f)(\varphi) = f(\Psi\varphi) \quad (\varphi \in \mathscr{D}),$$

那么 Ψf 显然是 $\mathscr{D}$上连续线性泛函，即 $\Psi f \in \mathscr{D}'$，称 Ψ 是 $\mathscr{D}'$ 的一个乘子。

注　对于 $\Psi \in C^\infty(\Omega)$，$f \in D'$，可定义乘积 Ψf，但对于两个广义函数即不能定义乘积运算。

例 3　证明 $x^2\delta'' = 2\delta$。

证　$x^2\delta''(\varphi) = (-1)^2\delta[(x^2\varphi(x))'']$，又

$$(x^2\varphi(x))'' = (2x\varphi(x) + x^2\varphi'(x))' =$$
$$2\varphi(x) + 2x\varphi'(x) + 2x\varphi'(x) + x^2\varphi''(x),$$

故

$$\delta[(x^2\varphi(x))''] = 2\varphi(x) + 4x\varphi'(x) + x^2\varphi''(x)\Big|_{x=0} =$$
$$2\varphi(0) = 2\delta(\varphi),$$

即 $x^2\delta'' = 2\delta$。

注　对于任何正整数 k，$\delta^{(k)}$ 为按如下公式定义的广义函数

$$\delta^{(k)}(\varphi) = (-1)^k\delta(\varphi^{(k)}) = (-1)^k\varphi^{(k)}(0)。$$

【性质 2】 (1) 设 $f,g\in\mathscr{D}$，α、β 是数，则

$$\frac{\partial}{\partial x_j}(\alpha f+\beta g)=\alpha\frac{\partial f}{\partial x_j}+\beta\frac{\partial g}{\partial x_j}\quad(j=1,2,\cdots,n);$$

(2) 设 $\Psi\in C^\infty(\Omega)$，$f\in\mathscr{D}$，则

$$\frac{\partial}{\partial x_j}(\Psi f)=\Psi\frac{\partial f}{\partial x_j}+f\frac{\partial\Psi}{\partial x_j}\quad(j=1,2,\cdots,n)。$$

证　(1) 由定义显然。我们来证(2)，对于任意 $\varphi\in\mathscr{D}$，则

$$\begin{aligned}\frac{\partial(\Psi f)}{\partial x_j}(\varphi)&=-(\Psi f)\left(\frac{\partial\varphi}{\partial x_j}\right)=-f\left(\Psi\frac{\partial\varphi}{\partial x_j}\right)=\\&-\left[f\frac{\partial(\varphi\Psi)}{\partial x_j}\right]+f\left(\varphi\frac{\partial\Psi}{\partial x_j}\right)=\\&\frac{\partial f}{\partial x_j}(\Psi\varphi)+\left(f\frac{\partial\Psi}{\partial x_j}\right)(\varphi)=\\&\left(\Psi\frac{\partial f}{\partial x_j}\right)(\varphi)+\left(f\frac{\partial\Psi}{\partial x_j}\right)(\varphi),\end{aligned}$$

故(2) 成立。

【性质 3】 设 $f_j\in\mathscr{D}$，$f\in\mathscr{D}$，若 $f_j\xrightarrow{\mathscr{D}}f$，则

$$\frac{\partial f_j}{\partial x_i}\xrightarrow{\mathscr{D}}\frac{\partial f}{\partial x_i}\quad(i=1,2,\cdots,n)。$$

证　对任意 $\varphi\in\mathscr{D}$，由于

$$\begin{aligned}\frac{\partial f}{\partial x_i}(\varphi)&=-f\left(\frac{\partial\varphi}{\partial x_i}\right)=-\lim_{j\to\infty}f_j\left(\frac{\partial\varphi}{\partial x_i}\right)=\\&\lim_{j\to\infty}\frac{\partial f_j}{\partial x_i}(\varphi),\end{aligned}$$

因此 $\frac{\partial f_j}{\partial x_i}\xrightarrow{\mathscr{D}}\frac{\partial f}{\partial x_i}$。

【定义 3】 设 $\{f_j\}\subset\mathscr{D}$，$f\in\mathscr{D}$，称级数 $\sum\limits_{j=1}^{\infty}f_j$ 在 $\mathscr{D}$ 中收敛于 f，

是指前 m 项和 $S_m = \sum_{j=1}^{m} f_j$ 在 $\mathscr{D}'$ 中收敛于 f，记为 $\sum_{j=1}^{\infty} f_j = f$。

由性质 3 得：

【性质 4】　若级数 $\sum_{j=1}^{\infty} f_j$ 收敛于 f，则级数 $\sum_{j=1}^{\infty} \frac{\partial f_j}{\partial x_i}$ 收敛于 $\frac{\partial f}{\partial x_i}$ $(i = 1,2,\cdots,n)$。

注　对于广义函数级数可以逐项求导，然而普通函数级数即使每项是连续可导函数，且处处收敛于某个连续可导数，也不能逐项求导。

例 4　$f_j(x) = \frac{1}{j}\cos jx$，则 $f_j(x)$ 一致收敛于 $f(x) \equiv 0$，但其导函数 $f_j'(x) = -\sin jx$ 不收敛于 $f'(x) \equiv 0$。如果将 f_j、f 看成广义函数，则在 $\mathscr{D}'$ 中有 $f_j' \to f$。

【性质 5】　设 $\{f_j\} \subset \mathscr{D}'$，如果对每个 $\varphi \in \mathscr{D}$，极限 $\lim_{j\to\infty} f_j(\varphi)$ 存在且有限，则必存在 $f \in \mathscr{D}'$，使得 $f_j \xrightarrow{\mathscr{D}'} f$。

证明性质 5 需用到拓扑线性空间的专门知识，故此略去其证明。

通过广义函数上述的性质，我们看到，广义函数空间 $\mathscr{D}'$ 关于求导与极限运算是封闭的，因此，广义函数的求导与极限运算比普通微积分中函数的相应运算既灵活又方便。

习题 6.2

1. 计算 $|x|^{(3)}$。

2. 已知

$$x_+^{\lambda} = \begin{cases} x^{\lambda}; & x > 0 \\ 0 & ; x \leqslant 0 \end{cases} \quad (\lambda \neq -1),$$

计算$(x_+^\lambda)'$。

3. 求证$(\mathrm{Ln}\,|x|)' = P\cdot V(\frac{1}{x})$，即

$$(\mathrm{Ln}\,|x|)'(\varphi) = \lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon}\frac{\varphi(x)}{x}\mathrm{d}x,\ \forall\,\varphi\in\mathscr{D}(R)。$$

4. 设$\Omega=(\alpha,\beta)\subset R^1$，$a\in\Omega$，$f$在$(\alpha,\beta)$内除$a$点外是连续可微的，且在$a$点是第一类间断点，$f'$在$(\alpha,\beta)\setminus\{a\}$内有界。证明$f$的广义导数为

$$\frac{\mathrm{d}f}{\mathrm{d}x} = f' + [f(a+0) - f(a-0)]\delta_a。$$

参　考　文　献

1. 王声望,郑维行．实变函数与泛函分析概要．北京:高等教育出版社,1991

2. 郭大钧．非线性泛函分析．济南:山东科学技术出版社,1985

3. 王建举．泛函分析与最优化理论．厦门:厦门大学出版社,1991

4. 韩崇昭,胡保生．泛函分析及其在自动控制中的应用．西安:西安交通大学出版社,1991