泛函分析与偏微分方程学习笔记

作者：UN

组织：北京师范大学数学科学学院

时间：2023.6.22 起著

模板来源：https://github.com/ElegantLaTeX/

指穷于为薪，火传也，不知其尽也。

目录

为了突出课程重点, 笔记与课后习题中杨大春老师的部分会用 $\star$ 标记, 高志强老师的部分会用 $\heartsuit$ 标记, 徐桂香老师的部分会用 $\clubsuit$ 标记. 自己暂未想清楚的题目用 $\bullet$ 标记.

第一部分

泛函分析

第一章 预备知识: L^p 空间

本章内容选自 [**ZMQ**], 旨在回顾 L^p 空间相关内容, 该类空间在泛函分析的诸多例子中扮演着重要角色. 特别说明在实变函数理论之后的课程中, 一般默认积分为 Lebesgue 积分, 测度为 Lebesgue 测度.

1.1 L^p 空间的定义与不等式

定义 1.1.1 (L^p 空间, 本性有界, 本性上确界)

1. 设 $f(\boldsymbol{x})$ 是 $E \subset \mathbb{R}^n$ 上的可测函数, 记
$$\|f\|_p = \left(\int_E |f(\boldsymbol{x})|^p d\boldsymbol{x}\right)^{\frac{1}{p}}, \ 0 < p < +\infty,$$
用 $L^p(E)$ 表示使 $\|f\|_p < +\infty$ 的 f 的全体, 称其为 L^p 空间.
2. 设 $f(\boldsymbol{x})$ 是 $E \subset \mathbb{R}^n$ 上的可测函数, $m(E) > 0$. 若存在 M, 使得 $|f(\boldsymbol{x})| \le M$, a.e. $\boldsymbol{x} \in E$, 则称 $f(\boldsymbol{x})$ 在 E 上本性有界, M 称为 $f(\boldsymbol{x})$ 的本性上界. 再对一切本性上界取下确界, 记为 $\|f\|_\infty$, 称它为 $f(\boldsymbol{x})$ 在 E 上的本性上确界. 此时用 $L^\infty(E)$ 表示在 E 上本性有界的函数的全体.

♣

命题 1.1.1

若 $0 < m(E) < +\infty$, 则
$$\lim_{p\to\infty} \|f\|_p = \|f\|_\infty.$$

♠

证明 设 $M = \|f\|_\infty$, 则根据本性上确界的定义知对任一满足 $M' < M$ 的 M', 点集
$$A = \{\boldsymbol{x} \in E : |f(\boldsymbol{x})| > M'\}$$
有正测度. 故由
$$\|f\|_p = \left(\int_E |f(\boldsymbol{x})|^p d\boldsymbol{x}\right)^{\frac{1}{p}} \ge \left(\int_E M'^p d\boldsymbol{x}\right)^{\frac{1}{p}} = M'(m(E))^{\frac{1}{p}}$$
知 $\liminf_{p\to\infty} \|f\|_p \ge M'$. 再令 $M' \to M$ 得 $\liminf_{p\to\infty} \|f\|_p \ge M$.

另一方面, 知
$$\|f\|_p \le \left(\int_E M^p dx\right)^{\frac{1}{p}} = M(m(E))^{\frac{1}{p}},$$
得 $\limsup_{p\to\infty} \|f\|_p \le M$, 夹逼即得欲证. □

定理 1.1.1 (L^p 是线性空间)

若 $f, g \in L^p(E), 0 < p \le +\infty, \alpha, \beta$ 是实数, 则
$$\alpha f + \beta g \in L^p(E).$$

♡

证明 当 $0 < p < \infty$, 注意首先有
$$\left|\frac{\alpha f(x) + \beta g(x)}{2}\right|^p \le |\max\{\alpha f(x), \beta g(x)\}|^p \le |\alpha f(x)|^p + |\beta g(x)|^p,$$
得到
$$|\alpha f(x) + \beta g(x)|^p \le 2^p(|\alpha|^p |f(x)|^p + \beta^p |g(x)|^p).$$
两边积分得到
$$\int_E |\alpha f(x) + \beta g(x)|^p dx \le \int_E 2^p(|\alpha|^p |f(x)|^p + \beta^p |g(x)|^p) dx < +\infty.$$

因而 $\alpha f+\beta g\in L^p(E)$.

当 $p=\infty$, 有

$$|\alpha f(x)+\beta g(x)|\le|\alpha|\cdot\|f\|_\infty+|\beta|\cdot\|g\|_\infty,\ \text{a.e. } x\in E,$$

根据本性上确界的定义知

$$\|\alpha f(x)+\beta g(x)\|_\infty\le|\alpha|\cdot\|f\|_\infty+|\beta|\cdot\|g\|_\infty,\ \text{a.e. } x\in E.$$

因而 $\alpha f+\beta g\in L^p(E)$. □

注意研究 $L^p(E)$ 时, 主要的兴趣都在 $p\ge1$ 的情形, 故下面若无特殊说明, 均默认 $p\ge1$.

定义 1.1.2 (共轭指标)

若 $p,p'>1$, 且 $\frac{1}{p}+\frac{1}{p'}=1$, 则称 p 与 p' 为共轭指标. 若 $p=1$, 则规定共轭指标 $p'=\infty$; 若 $p=\infty$, 则规定共轭指标 $p'=1$. ♣

定理 1.1.2 (Hölder 不等式)

设 p 与 p' 为共轭指标, 若 $f\in L^p(E), g\in L^{p'}(E)$, 则

$$\|fg\|_1\le\|f\|_p\cdot\|g\|_{p'},\ 1\le p\le\infty,$$

也即

$$\int_E|f(x)g(x)|dx\le\left(\int_E|f(x)|^pdx\right)^{\frac{1}{p}}\left(\int_E|g(x)|^{p'}dx\right)^{\frac{1}{p'}},\ 1<p<+\infty.$$

以及

$$\int_E|f(x)g(x)|dx\le\|g\|_\infty\int_E|f(x)|dx,\ p=1,$$

与

$$\int_E|f(x)g(x)|dx\le\|f\|_\infty\int_E|g(x)|dx,\ p=\infty.$$

♡

证明 当 $p=1$ 或 ∞ 时, 欲证式显然成立.

当 $\|f\|_p=0$ 或 $\|g\|_{p'}=0$, 知 $f(x)g(x)=0$, a.e. $x\in E$, 欲证式进而成立.

当 $\|f\|_p>0,\|g\|_{p'}>0,p,p'<+\infty$, 注意 Young 不等式

$$a^{\frac{1}{p}}b^{\frac{1}{p'}}\le\frac{a}{p}+\frac{b}{p'},\ a>0,b>0,$$

代入 $a=\frac{|f(x)|^p}{\|f\|_p},b=\frac{|g(x)|^{p'}}{\|g\|_{p'}}$ 得到

$$\frac{|f(x)g(x)|}{\|f\|_p\cdot\|g\|_{p'}}\le\frac{1}{p}\cdot\frac{|f(x)|^p}{\|f\|_p^p}+\frac{1}{p'}\cdot\frac{|g(x)|^{p'}}{\|g\|_{p'}^{p'}},$$

两边积分有

$$\frac{\|fg\|_1}{\|f\|_p\cdot\|g\|_{p'}}\le\frac{1}{p}\cdot\frac{\|f\|_p^p}{\|f\|_p^p}+\frac{1}{p'}\cdot\frac{\|g\|_{p'}^{p'}}{\|g\|_{p'}^{p'}}=\frac{1}{p}+\frac{1}{p'}=1.$$

因而

$$\|fg\|_1\le\|f\|_p\cdot\|g\|_{p'}.$$

□

推论 1.1.1 (Cauthy-Bunyakowsky-Schwarz 不等式)

若 $f,g\in L^2(E)$, 则

$$\int_E|f(x)g(x)|dx\le\left(\int_E|f(x)|^2dx\right)^{\frac{1}{2}}\left(\int_E|g(x)|^2dx\right)^{\frac{1}{2}}.$$

♡

证明 在 Hölder 不等式中代入 $p = p' = 2$ 即可. □

下面的定理在泛函分析中有广泛用途.

定理 1.1.3

若 $m(E) < +\infty$, 且 $0 < p_1 < p_2 \le \infty$, 则 $L^{p_2}(E) \subset L^{p_1}(E)$, 且有

$$\|f\|_{p_1} \le (m(E))^{\frac{1}{p_1}-\frac{1}{p_2}} \|f\|_{p_2}.$$

证明 当 $p_2 = +\infty$, 知 $\frac{1}{p_2} = 0$, 有

$$\left(\int_E |f(x)|^{p_1} dx\right)^{\frac{1}{p_1}} \le \left(\int_E \|f\|_\infty^{p_1} dx\right)^{\frac{1}{p_1}} = (m(E))^{\frac{1}{p_1}} \|f\|_\infty.$$

当 $p_2 < +\infty$, 针对 $\int_E |f(x)|^{p_1} dx$, 考虑 $r = \frac{p_2}{p_1}, r' = \frac{p_2}{p_2-p_1}, \frac{1}{r} + \frac{1}{r'} = 1$, 由 Hölder 不等式得

$$\int_E |f(x)|^{p_1} dx \le \left(\int_E |f(x)|^{p_1 r} dx\right)^{\frac{1}{r}} \left(\int_E 1 dx\right)^{\frac{1}{r'}} = \left(\int_E |f(x)|^{p_2} dx\right)^{\frac{p_1}{p_2}} \cdot (m(E))^{\frac{p_2-p_1}{p_2}}.$$

两边同取 $\frac{1}{p_1}$ 次幂得到

$$\|f\|_{p_1} \le (m(E))^{\frac{1}{p_1}-\frac{1}{p_2}} \|f\|_{p_2}.$$

□

定理 1.1.4

若 $f \in L^r(E) \cap L^s(E)$, 且令 $0 < r < p < s \le +\infty$ 以及

$$0 < \lambda < 1, \frac{1}{p} = \frac{\lambda}{r} + \frac{1-\lambda}{s},$$

则

$$\|f\|_p \le \|f\|_r^\lambda \cdot \|f\|_s^{1-\lambda}.$$

证明 当 $r < s < \infty$, 注意 $1 = \frac{\lambda p}{r} + \frac{(1-\lambda)p}{s}$, 由 Hölder 不等式, 针对 $\int_E |f(x)|^p dx$ 有

$$\begin{aligned}\int_E |f(x)|^p dx &= \int_E |f(x)|^{\lambda p} \cdot |f(x)|^{(1-\lambda)p} dx \\ &\le \left(\int_E (|f(x)|^{\lambda p})^{\frac{r}{\lambda p}} dx\right)^{\frac{\lambda p}{r}} \left(\int_E (|f(x)|^{(1-\lambda)p})^{\frac{s}{(1-\lambda)p}} dx\right)^{\frac{(1-\lambda)p}{s}} \\ &= \left(\int_E |f(x)|^r dx\right)^{\frac{\lambda p}{r}} \left(\int_E |f(x)|^s dx\right)^{\frac{(1-\lambda)p}{s}},\end{aligned}$$

两边同取 $\frac{1}{p}$ 次幂得到

$$\|f\|_p \le \|f\|_r^\lambda \cdot \|f\|_s^{1-\lambda}.$$

当 $r < s = +\infty$, 知此时 $p = \frac{r}{\lambda}$, 故

$$\begin{aligned}\int_E |f(x)|^p dx &= \int_E |f(x)|^{p-r} \cdot |f(x)|^r \\ &\le \|f^{p-r}\|_\infty \cdot \|f^r\|_1 \\ &= \|f^{p-p\lambda}\|_\infty \cdot \int_E |f(x)|^r dx \\ &= \|f\|_\infty^{p(1-\lambda)} \cdot \|f\|_r^{p\lambda}.\end{aligned}$$

□

注 上述结果还说明了当 $r < p < s \le +\infty$ 时有

$$\|f\|_p \le \max\{\|f\|_r, \|f\|_s\}.$$

定理 1.1.5 (反向 Hölder 不等式)

设 $0<p<1, q<0, \frac{1}{p}+\frac{1}{q}=1$, 则对 $f\in L^p(E), g\in L^q(E)$ 有

$$\int_E |f(x)g(x)|dx \geq \|f\|_p\cdot\|g\|_q.$$

证明 不妨设 $fg\in L(E)$, 记 $\overline{p}=\frac{1}{p}>1, \overline{q}=\frac{1}{1-p}>1$, 则有

$$\begin{aligned}\|f\|_p &= \left(\int_E |f(x)g(x)|^p\cdot|g(x)|^{-p}dx\right)^{\frac{1}{p}} \leq \left(\int_E |f(x)g(x)|^{p\cdot\overline{p}}dx\right)^{\frac{1}{p}\cdot\frac{1}{\overline{p}}}\cdot\left(\int_E |g(x)|^{-p\cdot\overline{q}}dx\right)^{\frac{1}{p}\cdot\frac{1}{\overline{q}}}\\ &= \left(\int_E |f(x)g(x)|dx\right)\cdot\left(\int_E |g(x)|^{-\frac{p}{1-p}}dx\right)^{\frac{1-p}{p}} = \int_E \frac{|f(x)g(x)|}{\|g\|_q}dx.\end{aligned}$$

两边同乘 $\|g\|_q$ 即得欲证. □

定理 1.1.6 (Minkowski 不等式)

若 $f,g\in L^p(E)(1\leq p\leq\infty)$, 则

$$\|f+g\|_p\leq\|f\|_p+\|g\|_p.$$

证明 当 $p=1$ 时, 欲证式即三角不等式. 当 $p=\infty$ 时, 根据本性上确界的定义有

$$|f(x)|\leq\|f\|_\infty, |g(x)|\leq\|g\|_\infty,\ \text{a.e. } x\in E.$$

故

$$|f(x)+g(x)|\leq|f(x)|+|g(x)|\leq\|f\|_\infty+\|g\|_\infty,\ \text{a.e. } x\in E.$$

从而 $\|f+g\|_\infty\leq\|f\|_\infty+\|g\|_\infty$, 欲证式成立.

当 $1<p<+\infty$, 有

$$\begin{aligned}\int_E |f(x)+g(x)|^p dx &\leq \int_E |f(x)+g(x)|^{p-1}\cdot|f(x)+g(x)|dx\\ &\leq \int_E |f(x)+g(x)|^{p-1}\cdot|f(x)|dx+\int_E |f(x)+g(x)|^{p-1}\cdot|g(x)|dx\\ &\leq (\int_E |f(x)+g(x)|^{(p-1)\cdot\frac{p}{p-1}}dx)^{\frac{p-1}{p}}\cdot(\int_E |f(x)|^p dx)^{\frac{1}{p}}+(\int_E |f(x)+g(x)|^{(p-1)\cdot\frac{p}{p-1}}dx)^{\frac{p-1}{p}}\cdot(\int_E |g(x)|^p dx)^{\frac{1}{p}}\\ &= \|f+g\|_p^{p-1}\cdot\|f\|_p+\|f+g\|_p^{p-1}\cdot\|g\|_p,\end{aligned}$$

当 $\|f+g\|_p\neq 0$ 时, 这便是

$$\|f+g\|_p^p\leq\|f+g\|_p^{p-1}(\|f\|_p+\|g\|_p)\Rightarrow\|f+g\|_p\leq\|f\|_p+\|g\|_p.$$

而当 $\|f+g\|=0$, 原式显然成立. □

定理 1.1.7 (反 Minkowski 不等式)

设 $0<p<1$, 则对 $f,g\in L^p(E)$ 有

$$\||f|+|g|\|_p\geq\|f\|_p+\|g\|_p.$$

证明 不妨设 $\||f|+|g|\|>0$, 则由反 Hölder 不等式知

$$\begin{aligned}\||f|+|g|\|_p^p &= \int_E (|f(x)|+|g(x)|)^{p-1}\cdot|f(x)|dx+\int_E (|f(x)|+|g(x)|)^{p-1}\cdot|g(x)|dx\\ &\geq \left(\int_E (|f(x)|+|g(x)|)^{(p-1)\cdot\frac{p}{p-1}}dx\right)^{\frac{p-1}{p}}\cdot\int_E |f(x)|^p dx^{\frac{1}{p}}+\left(\int_E (|f(x)|+|g(x)|)^{(p-1)\cdot\frac{p}{p-1}}dx\right)^{\frac{p-1}{p}}\cdot\int_E |g(x)|^p dx^{\frac{1}{p}}\\ &= \||f|+|g|\|_p^{p-1}\cdot\|f\|_p+\||f|+|g|\|_p^{p-1}\cdot\|g\|_p.\end{aligned}$$

这便得到

$$\||f|+|g|\|_p\geq\|f\|_p+\|g\|_p.$$

□

定理 1.1.8

设 $f \in L^{p_1}(E), g \in L^{p_2}(E)(0 < p_1, p_2 < +\infty)$, 则

$$fg \in L^{\tilde{p}}(E),\ \tilde{p} = \frac{p_1 p_2}{p_1 + p_2}.$$

证明 取 $\lambda = \frac{1}{p_1} + \frac{1}{p_2}, p = \lambda p_1, q = \lambda p_2$, 则

$$0 < p, q < +\infty,\ \frac{1}{p} + \frac{1}{q} = 1.$$

知

$$f \in L^{p_1}(E) \Rightarrow \||f|^{p_1}\|_1 = \||f|^{\frac{p}{\lambda}}\|_1 < \infty \Rightarrow \||f|^{\frac{1}{\lambda}}\|_p < \infty \Rightarrow |f|^{\frac{1}{\lambda}} \in L^p(E).$$

同理可得 $|g|^{\frac{1}{\lambda}} \in L^q(E)$, 从而 $|fg|^{\frac{1}{\lambda}} \in L^1(E)$, 因而

$$\begin{aligned}\int_E |f(x)g(x)|^{\tilde{p}} dx &= \int_E |f(x)|^{\frac{1}{\lambda}} \cdot |g(x)|^{\frac{1}{\lambda}} dx \\ &\le \left(\int_E |f(x)|^{\frac{1}{\lambda}\cdot p_1 \lambda} dx\right)^{\frac{1}{p_1 \lambda}} \cdot \left(\int_E |g(x)|^{\frac{1}{\lambda}\cdot 2\lambda} dx\right)^{\frac{1}{p_2 \lambda}} \\ &= \|f\|_{p_1}^{\frac{p_1}{p}} \cdot \|g\|_{p_2}^{\frac{p_2}{p}} < \infty.\end{aligned}$$

故 $fg \in L^{\tilde{p}}(E)$. □

1.2 L^p 空间的结构

本节主要证明 L^p 空间的完备性和可分性, 会用到后面的概念, 整理于此仅供补充 L^p 空间的重要性质. 这里采用与实变函数论相同的函数等价类, 即用几乎处处相等作为等价关系把 $L^p(E)$ 中的元划分为等价类.

1.2.1 $L^p(E)$ 的完备性

定理 1.2.1 ($L^p(E)$ 是距离空间)

对于 $f, g \in L^p(E)$, 定义

$$d(f,g) = \|f - g\|_p (= (\int_E |f(x) - g(x)|^p dx)^{\frac{1}{p}}),\ 1 \le p \le +\infty,$$

则 $(L^p(E), d)$ 是距离空间, 简记为 $L^p(E)$. 称 d 为 L^p 距离.

证明 对于正定性, 显见 $d(f,g) \ge 0$. 因为 $\|f - g\|_p = 0 \Leftrightarrow f(x) = g(x)$ a.e., 故 $d(f,g) = 0 \Leftrightarrow f = g$.

显见 $d(f,g) = d(g,f)$, 对称性得证.

对于三角不等式, 由 Minkowski 不等式知

$$\|f - g\|_p = \|f - h + h - g\|_p \le \|f - h\|_p + \|h - g\|_p,$$

也即

$$d(f,g) \le d(f,h) + d(h,g).$$

三角不等式得证, 故 d 为距离, 命题得证. □

定义 1.2.1 (L^p 意义下的极限)

设 $f_k \in L^p(E)(k = 1, 2, \cdots)$. 若存在 $f \in L^p(E)$, 使得

$$\lim_{k\to\infty} d(f_k, f) = \lim_{k\to\infty} \|f_k - f\|_p = 0,$$

则称 $\{f_k\}$ 依 $L^p(E)$ 的意义收敛到 f, $\{f_k\}$ 为 $L^p(E)$ 中的收敛列, f 为 $\{f_k\}$ 在 $L^p(E)$ 中的极限. ♣

容易证明 L^p 极限的两条重要性质:

命题 1.2.1 (唯一性)

若

$$\lim_{k\to\infty}\|f_k-f\|_p=0,\ \lim_{k\to\infty}\|f_k-g\|_p=0,$$

则 $f=g$(也即 $f(x)=g(x)$ a.e.). ♠

命题 1.2.2

若 $\lim_{k\to\infty}\|f_k-f\|_p=0$, 则 $\lim_{k\to\infty}\|f_k\|_p=\|f\|_p$. ♠

定义 1.2.2 ($L^p(E)$ 中的基本列)

设 $\{f_k\}\subset L^p(E)$. 若 $\lim_{k,j\to\infty}\|f_k-f_j\|_p=0$, 则称 $\{f_k\}$ 是 $L^p(E)$ 中的基本列 (或 Cauchy 列). ♣

在介绍 $L^p(E)$ 的完备性之前, 先来回顾一下实变函数论中的三个重要定理.

定理 1.2.2 (Fatou)

设 $g\in L(E)$, $f_n\in L(E)(n\in\mathbb{N})$. 若 $f_n(x)\geq g(x)$, a.e. $x\in E$, 则

$$\int_E \liminf_{n\to\infty} f_n(x)dx\leq \liminf_{n\to\infty}\int_E f_n(x)dx.$$

♡

定义 1.2.3 (依测度收敛)

设 $f(x), f_1(x),\cdots, f_k(x),\cdots$ 是 E 上几乎处处有限的可测函数. 若对任给的 $\varepsilon>0$, 有

$$\lim_{k\to\infty} m(\{x\in E: |f_k(x)-f(x)|>\varepsilon\})=0,$$

就称 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$. ♣

定义 1.2.4 (依测度基本列)

设 $\{f_k(x)\}$ 是 E 上几乎处处有限的可测函数列. 若对任给的 $\varepsilon>0$, 有

$$\lim_{\substack{k\to\infty\\ j\to\infty}} m(\{x\in E: |f_k(x)-f_j(x)|>\varepsilon\})=0,$$

则称 $\{f_k(x)\}$ 为 E 上的依测度基本 (Cauchy) 列. ♣

定理 1.2.3

若 $\{f_k(x)\}$ 是 E 上的依测度基本列, 则在 E 上存在几乎处处有限的可测函数 $f(x)$, 使得 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$. ♡

定理 1.2.4 (Riesz)

若 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$, 则存在子列 $\{f_{k_i}(x)\}$, 使得

$$\lim_{i\to\infty} f_{k_i}(x)=f(x),\ \text{a.e. } x\in E.$$

♡

定理 1.2.5 (Riesz-Fisher 定理)

$L^p(E)$ 是完备的距离空间. ♡

证明 设 $1 \le p < +\infty$, 取基本列 $\{f_k\} \subset L^p(E)$ 满足 $\lim_{k,j\to\infty} \|f_j - f_k\|_p = 0$, 则对任给的 $\sigma > 0$, 取

$$E_{j,k}(\sigma) = \{x \in E : |f_j(x) - f_k(x)| \ge \sigma\},$$

则根据 $E_{j,k}(\sigma)$ 的定义有

$$\int_{E_{j,k}(\sigma)} |f_j(x) - f_k(x)|^p dx \ge \int_{E_{j,k}(\sigma)} \sigma^p dx = \sigma^p m(E_{j,k}(\sigma)),$$

得到

$$\sigma(m(E_{j,k}(\sigma)))^{\frac{1}{p}} \le (\int_{E_{j,k}(\sigma)} |f_j(x) - f_k(x)|^p dx)^{\frac{1}{p}} \le (\int_E |f_j(x) - f_k(x)|^p dx)^{\frac{1}{p}} = \|f_j - f_k\|_p.$$

两边同取 $j,k \to \infty$ 得

$$\lim_{j,k\to\infty} m(E_{j,k}(\sigma)) = 0.$$

这说明 $\{f_k(x)\}$ 在 E 上是依测度收敛列, 从而根据定理1.2.3知存在 E 上几乎处处有限的可测函数 $f(x)$, 使得 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$, 进而由 Riesz 定理1.2.4知可选出 $\{f_k(x)\}$ 的子列 $\{f_{k_i}(x)\}$, 使得

$$\lim_{i\to\infty} f_{k_i}(x) = f(x), \text{ a.e. } x \in E.$$

下面说明 $f_k \to f$ 在 L^p 意义下成立, 由 Fatou 引理知

$$\int_E |f_k(x) - f(x)|^p dx = \int_E \lim_{i\to\infty} |f_k(x) - f_{k_i}(x)|^p dx \le \liminf_{i\to\infty} \int_E |f_k(x) - f_{k_i}(x)|^p dx,$$

故

$$\lim_{k\to\infty} \int_E |f_k(x) - f(x)|^p dx = 0.$$

也即

$$\lim_{k\to\infty} \|f_k - f\|_p = 0.$$

最后根据

$$\|f\|_p \le \|f - f_k\|_p + \|f_k\|_p < \infty$$

得到 $f \in L^p(E)$.

再考虑 $p = \infty$ 的情况. 取基本列 $\{f_k\} \subset L^\infty(E)$ 满足 $\lim_{k,j\to\infty} \|f_k - f_j\|_\infty = 0$. 根据本性上确界的定义, 对任意一对自然数 k, j 都有:

$$|f_k(x) - f_j(x)| \le \|f_k - f_j\|_\infty, \text{ a.e. } x \in E,$$

故存在零测集 Z, 使得对一切自然数 k, j 都有

$$|f_k(x) - f_j(x)| \le \|f_k - f_j\|_\infty, \ x \notin Z.$$

现在在点态的意义下便存在 $f(x)$, 使得

$$\lim_{k\to\infty} f_k(x) = f(x), \ x \in E \backslash Z.$$

根据基本列定义中极限的定义, 任取 $\varepsilon > 0$ 都存在自然数 N 使得

$$\|f_k - f_j\|_\infty < \varepsilon, \ j, k > N.$$

在上式中令 $j \to \infty$ 得

$$|f_k(x) - f(x)| \le \varepsilon, \ k > N, x \in E \backslash Z.$$

故当 $k > N$ 时 $\|f_k - f\|_\infty \le \varepsilon$. 这一方面说明 $f \in L^\infty(E)$, 另一方面说明

$$\|f_k - f\|_\infty \to 0, \ k \to \infty.$$

□

1.2.2 $L^p(E)(1 \le p < +\infty)$ 的可分性

定义 1.2.5 (稠密, 可分)

设 Γ 是 $L^p(E)$ 的子集. 若对任意的 $f \in L^p(E)$ 与 $\varepsilon > 0$, 总存在 $g \in \Gamma$ 使得 $\|f - g\|_p < \varepsilon$, 则称 Γ 在 $L^p(E)$ 中稠密; 若 $L^p(E)$ 中存在稠密的可数子集, 就称 $L^p(E)$ 是可分的. ♣

引理 1.2.1

设 $f \in L^p(E)(1 \le p < +\infty)$, 则对任意的 $\varepsilon > 0$, 有

(i) 存在 $\mathbb{R}^n$ 上具有紧支集的连续函数 $g(x)$, 使得

$$\int_E |f(x) - g(x)|^p dx < \varepsilon.$$

(ii) 存在 $\mathbb{R}^n$ 上具有紧支集的阶梯函数

$$\varphi(x) = \sum_{i=1}^{k} c_i \chi_{I_i}(x),$$

其中每个 I_i 都是二进方体, 使得

$$\int_E |f(x) - \varphi(x)|^p dx < \varepsilon.$$

♡

定理 1.2.6

$L^p(E)(1 \le p < +\infty)$ 是可分空间. ♡

证明 首先设 $E = \mathbb{R}^n$, $f \in L^p(\mathbb{R}^n)$, 则对任意的 $\varepsilon > 0$, 由引理1.2.1(ii) 知存在 $\mathbb{R}^n$ 上的阶梯函数 $\varphi(x)$, 使得 $\|f - \varphi\|_p < \varepsilon$, 其中

$$\varphi(x) = \sum_{i=1}^{k} c_i \chi_{I_i}(x).$$

不妨设 $|c_i| < M, m(I_i) < M^p(i = 1, 2, \cdots, k)$. 根据有理数在 $\mathbb{R}$ 中的稠密性, 知对每个 i, 都存在有理数 r_i 使得

$$|r_i| < M, |c_i - r_i| < \varepsilon,\ i = 1, 2, \cdots, k.$$

另取

$$\psi(x) = \sum_{i=1}^{k} r_i \chi_{I_i}(x),$$

知

$$\begin{aligned}\|\varphi - \psi\|_p = \|\sum_{i=1}^{k} c_i \chi_{I_i} - \sum_{i=1}^{k} r_i \chi_{I_i}\|_p &\le \sum_{i=1}^{n} \|c_i \chi_{I_i} - r_i \chi_{I_i}\|_p \\ &\le \sum_{i=1}^{k} |c_i - r_i| \cdot \|\chi_{I_i}\|_p \le \varepsilon \sum_{i=1}^{k} (m(I_i))^{\frac{1}{p}} \le kM\varepsilon.\end{aligned}$$

从而

$$\|f - \psi\|_p \le \|f - \varphi\|_p + \|\varphi - \psi\|_p < (1 + kM)\varepsilon.$$

由 ε 的任意性与形如 ψ 的阶梯函数全体 Γ 是可数集, 得到 Γ 正是 $L^p(\mathbb{R}^n)$ 的一个可数稠密子集.

再考虑一般可测集 E, 设 $f \in L^p(E)$, 作

$$f_1(x) = \begin{cases} f(x), & x \in E, \\ 0, & x \notin E. \end{cases}$$

显见 $f_1 \in L^p(\mathbb{R}^n)$, 从而由上述结论知存在 $g \in \Gamma$ 使得

$$\left(\int_{\mathbb{R}^n} |f_1(x) - g(x)|^p dx\right)^{\frac{1}{p}} < \varepsilon,$$

进而

$$\left(\int_E |f(x) - g(x)|^p dx\right)^{\frac{1}{p}} < \varepsilon.$$

若将 Γ 中每个函数的定义域限定在 E 上, 并记这样的函数全体为 Γ', 则 Γ' 正是 $L^p(E)$ 的一个可数稠密子集. □

推论 1.2.1

若 $1 \leq p < +\infty, 1 \leq r \leq +\infty$, 则 $L^p(E) \cap L^r(E)$ 在 $L^p(E)$ 中稠密.

第二章 度量空间

2.1 压缩映射原理

2.1.1 知识梳理

定义 2.1.1 (距离)

设 $\mathscr{X}$ 是一个非空集. $\mathscr{X}$ 叫作度量空间, 是指在 $\mathscr{X}$ 上定义了一个双变量的实值函数 $\rho(x,y)$, 满足下列三个条件:

1. $\rho(x,y) \geq 0$, 且 $\rho(x,y)=0$ 当且仅当 $x=y$;
2. $\rho(x,y)=\rho(y,x)$;
3. $\rho(x,z) \leq \rho(x,y)+\rho(y,z)(\forall x,y,z \in \mathscr{X})$.

这里 ρ 叫作 $\mathscr{X}$ 上的一个距离, 以 ρ 为度量的度量空间 $\mathscr{X}$ 记作 $(\mathscr{X},\rho)$.

例 2.1(连续函数空间) 区间 $[a,b]$ 上的连续函数全体记为 $C[a,b]$, 按距离[1]

$$\rho(x,y) := \max_{a\leq t\leq b}|x(t)-y(t)| \tag{2.1}$$

形成度量空间 $(C[a,b],\rho)$, 简记作 $C[a,b]$.

注 一般称下述度量为欧氏度量:

$$\rho(x,y)=[(x_1-y_1)^2+\cdots+(x_n-y_n)^2]^{\frac{1}{2}}. \tag{2.2}$$

定义 2.1.2 (收敛性)

度量空间 $(\mathscr{X},\rho)$ 上的点列 $\{x_n\}$ 叫作收敛到 x_0 是指: $\rho(x_n,x_0)\to 0(n\to\infty)$. 这时记作 $\lim_{n\to\infty} x_n = x_0$, 或简单记作 $x_n \to x_0$.

注 在 $C[a,b]$ 中点列 $\{x_n\}$ 收敛到 x_0 是指: $\{x_n(t)\}$ 一致收敛到 $x_0(t)$.

定义 2.1.3 (闭集)

度量空间 $(\mathscr{X},\rho)$ 中的一个子集 E 称为闭集, 是指: $\forall\{x_n\}\subset E$, 若 $x_n\to x_0\in\mathscr{X}$, 则 $x_0\in E$.

定义 2.1.4 (基本列与完备空间)

度量空间 $(\mathscr{X},\rho)$ 上的点列 $\{x_n\}$ 叫作基本列, 是指: $\rho(x_n,x_m)\to 0(n,m\to\infty)$. 这也就是说: $\forall\varepsilon>0,\exists N(\varepsilon)$, 使得 $m,n\geq N(\varepsilon)\Rightarrow\rho(x_n,x_m)<\varepsilon$. 如果空间中所有基本列都收敛, 就称该空间完备.

例 2.2 $(\mathbb{R}^n,\rho)$ 完备, 其中 ρ 是欧氏度量(2.2).

命题 2.1.1

$(C[a,b],\rho)$ 是完备的.

证明 设 $\{x_n\}$ 是 $(C[a,b],\rho)$ 中的一串基本列, 则根据基本列定义知

$$\forall\varepsilon>0\exists N=N(\varepsilon)\forall m,n\geq N(\varepsilon)(\rho(x_m,x_n)=\max_{a\leq t\leq b}|x_m(t)-x_n(t)|<\varepsilon).$$

故 $\forall t\in[a,b]$ 有

$$|x_m(t)-x_n(t)|\leq\max_{a\leq t\leq b}|x_m(t)-x_n(t)|<\varepsilon,\ \forall m,n\geq N(\varepsilon).$$

---

[1] 该度量称为一致收敛性度量, 在 **[Zo]** 中有提及.

固定 $t \in [a,b]$, 知关于 n 的数列 $\{x_n(t)\}$ 是基本列, 进而极限 $\lim_{n\to\infty} x_n(t)$ 存在, 记之为 $x_0(t)$, 再在上式中令 $m \to \infty$ 得到 $|x_0(t) - x_n(t)| \le \varepsilon (\forall n \ge N(\varepsilon))$, 这说明 $x_n(t) \rightrightarrows x_0(t)$, 再由一致收敛保连续知 $x_0(t)$ 本身连续, 且在 $C[a,b]$ 有 $x_n \to x_0, n \to \infty$. □

课堂笔记 (♡)

- (基本列的性质)
  1. 收敛列必为基本列.
  2. 若基本列 $\{x_n\}$ 的某个子列 $\{x_{n_k}\}$ 收敛到 x, 那么 $\{x_n\}$ 也收敛, 且它收敛到 x. 这条性质在习题2.11中会用到.
- (一个不完备空间的例子) $(C[a,b], L^1)$ 不是完备的, 其中 L^1 表示 L^1 距离.
- 完备距离空间的任意闭子空间都完备.
- 仿照在欧氏空间的情形, 在一般的距离空间 (X,ρ) 上也可以定义开球

$$B(x_0, r) := \{x \in X : \rho(x, x_0) < r\},\ x_0 \in X, r > 0$$

与闭球:

$$\overline{B}(x_0, r) := \{x \in X : \rho(x, x_0) \le r\},\ x_0 \in X, r > 0.$$

下面考察度量空间之间的映射, 给定度量空间 $(\mathscr{X}, \rho), (\mathscr{Y}, r)$, 考察映射 $T : \mathscr{X} \to \mathscr{Y}$.

定义 2.1.5 (连续映射)

设 $T : (\mathscr{X}, \rho) \to (\mathscr{Y}, r)$ 是一个映射, 称它是连续的, 如果对于 $\mathscr{X}$ 中任意点列 $\{x_n\}$ 和点 x_0,

$$\rho(x_n, x_0) \to 0 \Rightarrow r(Tx_n, Tx_0) \to 0,\ n \to \infty.$$

命题 2.1.2

要使得 $T : (\mathscr{X}, \rho) \to (\mathscr{Y}, r)$ 是连续的, 必须且只需

$$\forall \varepsilon > 0 \forall x_0 \in \mathscr{X} \exists \delta = \delta(x_0, \varepsilon) > 0 \forall x \in \mathscr{X} (\rho(x, x_0) < \delta \Rightarrow r(Tx, Tx_0) < \varepsilon). \tag{2.3}$$

证明 如果(2.3)不成立, 则

$$\exists \varepsilon_0 > 0 \exists x_0 \in \mathscr{X} \forall n \in \mathbb{N} \exists x_n \in \mathscr{X} \left(\rho(x_n, x_0) < \frac{1}{n} \wedge r(Tx_n, Tx_0) \ge \varepsilon_0 \right),$$

进而 $\lim_{n\to\infty} \rho(x_n, x_0) = 0, \lim_{n\to\infty} r(Tx_n, Tx_0) \ne 0$, 矛盾!

当(2.3)成立, 则对任意满足 $\lim_{n\to\infty} \rho(x_n, x_0) = 0$ 的点列, 根据该极限的定义本身有

$$\forall \varepsilon > 0 \exists N = N(\delta(x_0, \varepsilon)) \forall n > N(\rho(x_n, x_0) < \delta),$$

进而由(2.3)知 $r(Tx_n, Tx_0) < \varepsilon$, 也即 $\lim_{n\to\infty} r(Tx_n, Tx_0) = 0$. □

在拓扑的意义下, 连续映射还能按照下述方式定义:

补充定义 2.1.1 ((拓扑意义下的) 连续映射 ♡)

设 $T : (\mathscr{X}, \rho) \to (\mathscr{Y}, r)$ 是一个映射, 称它是连续的, 如果对任意的 $x_0 \in \mathscr{X}$ 与任意的 $\varepsilon > 0$, 存在 $\delta > 0$ 使得

$$TB_{\mathscr{X}}(x_0, \delta) \subset B_{\mathscr{Y}}(Tx_0, \varepsilon).$$

注 [You] 中给出了下述命题:

命题 2.1.3

设 $f : X \to Y$ 是映射, 则下述各条件互相等价:

(i) f 是连续映射;

(ii) Y 中任一开集在 f 下的原像是 X 的开集;

(iii) Y 中任一闭集在 f 下的原像是 X 的闭集.

课堂笔记 (♡)

- 如果一个映射满足: 开集的原像都是开集, 那么该映射是连续映射.
- 下面举一个度量空间中连续映射的例子, 设 $\mathscr{X} = C[a,b], T : \mathscr{X} \to \mathbb{R}, Tx := \int_a^b x(t)dt$, 现在证明 T 连续. 任取 $\{x_n\} \subset \mathscr{X}$, 设 $x_n \to x_0 (n \to \infty)$, 则
$$|Tx_n - Tx_0| = |\int_a^b x_n(t)dt - \int_a^b x_0(t)dt| \le \int_a^b |x_n(t) - x_0(t)|dt \le \rho(x_n, x_0)|b-a| \to 0,\ n \to \infty,$$
故 T 连续.

对压缩映射原理, 首先介绍其引入动机: 不动点问题. 设 φ 是 $\mathbb{R}$ 上定义的实函数, 则求解下述方程根的问题:
$$\varphi(x) = 0$$
可以看成求解 $\mathbb{R} \to \mathbb{R}$ 的下述映射的不动点问题:
$$f(x) = x - \varphi(x),$$
也即求 $x \in \mathbb{R}$ 满足
$$f(x) = x.$$

同样, 下述常微分方程初值问题
$$\begin{cases} \dfrac{dx}{dt} = F(t,x), \\ x(0) = \xi, \end{cases}$$
或其积分方程等价形式
$$x(t) = \xi + \int_0^t F(\tau, x(\tau))d\tau$$
也可以看成一个不动点问题, 这是因为在以 $t = 0$ 为中心的某区间 $[-h, h]$ 上考察度量空间 $C[-h, h]$, 引入映射
$$(Tx)(t) = \xi + \int_0^t F(\tau, x(\tau))d\tau,$$
则原问题等价于求 $C[-h, h]$ 上的一个点 x, 使得 $x = Tx$, 也即求 T 的不动点.

定义 2.1.6 (不动点[WL])

设 X 是集合, 映射 $f : X \to X$. 若存在 $x_0 \in X$ 使得 $f(x_0) = x_0$, 就称 x_0 为 f 的一个不动点.

定义 2.1.7 (压缩映射)

称 $T : (\mathscr{X}, \rho) \to (\mathscr{X}, \rho)$ 是一个压缩映射, 如果存在 $0 < \alpha < 1$, 使得 $\rho(Tx, Ty) \le \alpha\rho(x, y) (\forall x, y \in \mathscr{X})$.

课堂笔记 (♡)

连续映射不一定是压缩映射, 但压缩映射一定是连续映射.

注 [WL] 中给出了另外两种形似压缩映射的定义作为拓展, 设 $\mathscr{X} = (\mathscr{X}, \rho)$ 是度量空间, 映射 $f : \mathscr{X} \to \mathscr{X}$:

定义 2.1.8 (弱压缩映射)

称 f 为弱压缩映射, 如果
$$d(f(x), f(y)) < d(x, y),\ \forall x, y \in X, x \ne y.$$

定义 2.1.9 (非扩张映射)

称 f 为非扩张映射, 如果

$$d(f(x), f(y)) \le d(x, y),\ \forall x, y \in X.$$

♣

显见压缩映射必为弱压缩映射, 弱压缩映射必为非扩张映射.

例 2.3 设 $\mathscr{X} = [0, 1]$, $T(x)$ 是 [0,1] 上的一个可微函数, 满足:

$$T(x) \in [0, 1], |T'(x)| \le \alpha < 1,\ \forall x \in [0, 1],$$

则映射 $T : \mathscr{X} \to \mathscr{X}$ 是一个压缩映射.

证明

$$\rho(Tx, Ty) = |T(x) - T(y)| = |T'(\theta x + (1-\theta)y)(x - y)| \le \alpha|x - y| = \alpha\rho(x, y),\ \forall x, y \in \mathscr{X}, 0 < \theta < 1.$$

□

定理 2.1.1 (Banach 不动点定理-压缩映射原理)

设 $(\mathscr{X}, \rho)$ 是一个完备的度量空间, T 是 $(\mathscr{X}, \rho)$ 到其自身的一个压缩映射, 则 T 在 $\mathscr{X}$ 上存在唯一的不动点.

♡

证明 设 $T : (\mathscr{X}, \rho) \to (\mathscr{X}, \rho)$ 是压缩映射, 满足:

$$\rho(Tx, Ty) \le \alpha\rho(x, y),\ \forall x, y \in \mathscr{X}, 0 < \alpha < 1,$$

任取初始点 $x_0 \in \mathscr{X}$, 考察迭代产生的序列

$$x_{n+1} = Tx_n,$$

知

$$\rho(x_{n+1}, x_n) = \rho(Tx_n, Tx_{n-1}) \le \alpha\rho(x_n, x_{n-1}) \le \cdots \le \alpha^n\rho(x_1, x_0),$$

这说明

$$\begin{aligned}
\rho(x_{n+p}, x_n) &\le \sum_{i=1}^{p} \rho(x_{n+i}, x_{n+i-1}) \\
&\le \sum_{i=1}^{\infty} \rho(x_{n+i}, x_{n+i-1}) \\
&\le \sum_{k=n}^{\infty} \alpha^k\rho(x_1, x_0) \\
&= \frac{\alpha^n}{1-\alpha}\rho(x_1, x_0) \to 0,\ \forall p \in \mathbb{N}, n \to \infty.
\end{aligned}$$

这说明 $\{x_n\}$ 是基本列, 又因为 $(\mathscr{X}, \rho)$ 完备, 故 $\{x_n\}$ 在 $(\mathscr{X}, \rho)$ 中有 (唯一的) 极限. □

★ 只要空间完备, 对于非压缩映射也有下述结论, 这在 **[WL]** 中也出现了:

补充定理 2.1.1 (♡)

设度量空间 $\mathscr{X}$ 是完备的, 映射 $T : \mathscr{X} \to \mathscr{X}$, 如果存在正整数 n, 使得 T^n 是 $\mathscr{X}$ 上的一个压缩映射, 那么映射 T 在 X 中必有唯一的不动点.

♡

证明 因为 T^n 是完备度量空间 $(\mathscr{X}, \rho)$ 上的压缩映射, 故根据压缩映射原理2.1.1知 T^n 存在唯一的不动点 $x^* \in \mathscr{X}$, 且根据压缩映射的定义设

$$\exists\alpha \in (0, 1)\forall x, y \in \mathscr{X}(\rho(T^n x, T^n y) \le \alpha\rho(x, y)),$$

根据不动点的定义知

$$T^n x^* = x^* \Rightarrow \rho(Tx^*, x^*) = \rho(T(T^n x^*), T^n x^*) = \rho(T^{n+1}x^*, T^n x^*) \le \alpha\rho(Tx^*, x^*),\ \alpha < 1.$$

这说明 $\rho(Tx^*, x^*) = 0$, 也即 x^* 也是 T 的不动点. 若 T 另有不动点 x^{**}, 知

$$Tx^{**} = x^{**} \Rightarrow T^n x^{**} = T^{n-1}x^{**} = \cdots = x^{**},$$

这说明 x^{**} 也是 T^n 的不动点, 从而由 T^n 不动点的唯一性知 $x^{**} = x^*$, 故 T 的不动点是唯一的. □

补充定理 2.1.2 (♡)

设 $(\mathscr{X}, \rho)$ 完备, $T : \mathscr{X} \to \mathscr{X}$, 若 T 在 $\mathscr{X}$ 的闭球 $Y = \{x \in \mathscr{X} : \rho(x_0, x) \le r\}$ 上是压缩的, 且满足条件:

$$\rho(x_0, Tx_0) \le (1-\alpha)r, \rho(Tx, Ty) \le \alpha\rho(x, y),\ \alpha \in (0, 1), \forall x, y \in Y,$$

则 T 在 Y 内有唯一不动点.

证明 因为 Y 是 $\mathscr{X}$ 的闭子集, 故 (Y, ρ) 为完备度量空间, 且 T 在 Y 内确为压缩映射, 故只需证明 $T(Y) \subset Y$. 任取 $x \in Y$, 设 $y := Tx$, 知

$$\rho(x_0, y) = \rho(x_0, Tx) \le \rho(x_0, Tx_0) + \rho(Tx_0, Tx) \le (1-\alpha)r + \alpha\rho(x_0, x) \le (1-\alpha)r + \alpha r = r,$$

故 $\rho(x_0, y) < r$, 也即 $y \in Y$, 命题即证. □

注 [WL] 中给出了另外两个与不动点相关的重要定理, 这些定理在后面也会介绍.

定理 2.1.2 (Schauder)

设 U 是 Banach 空间 $\mathscr{X}$ 中的紧凸集, $f : U \to U$ 是连续映射, 则 f 必有不动点.

定理 2.1.3 (Schauder-Tychonoff)

设 $\mathscr{X}$ 是一个局部凸的拓扑线性空间, U 是 $\mathscr{X}$ 中的紧凸集. 又设 $f : U \to U$ 是连续映射, 则存在 $x \in U$ 使得 $f(x) = x$.

现在再回头考虑常微分方程的初值问题

$$\begin{cases} \dfrac{dx}{dt} = F(t, x), \\ x(0) = \xi, \end{cases}$$

将其视作积分方程

$$x(t) = \xi + \int_0^t F(\tau, x(\tau))d\tau,$$

取映射

$$C[-h, h] \ni (Tx)(t) = \xi + \int_0^t F(\tau, x(\tau))d\tau,$$

考虑一致收敛性度量:

$$\rho(Tx, Ty) = \max_{|t|\le h}\left|\int_0^t F(\tau, x(\tau))d\tau - \int_0^t F(\tau, y(\tau))d\tau\right| \le h\max_{|t|\le h}|F(t, x(t)) - F(t, y(t))|.$$

现在希望找到某个 α 并令 F 有足够好的性质, 使得

$$h\max_{|t|\le h}|F(t, x(t)) - F(t, y(t))| \le \alpha\rho(x, y).$$

其中一种情况是 $F(t, x)$ 对变元 x 关于 t 一致地满足局部 Lipschitz 条件, 即

$$\exists\delta > 0 \exists L > 0 \forall |t| \le h \forall x_1, x_2 \in \overline{B}_\delta(\xi)(|F(t, x_1) - F(t, x_2)| \le L|x_1 - x_2|),$$

其中

$$\overline{B}_\delta(\xi) := \{x(t) \in C[-h, h] : \max_{|t|\le h}|x(t) - \xi| \le \delta\}.$$

此时

$$\rho(Tx, Ty) \le h\max_{|t|\le h}|F(t, x(t)) - F(t, y(t))| \le Lh\rho(x, y),\ \forall x, y \in \overline{B}(\xi, \delta).$$

但现在还不能将 $C[-h,h]$ 作为压缩映射原理2.1.1中的度量空间 $\mathscr{X}$, 这是因为当 $Lh<1$ 时, T 仅仅在 $C[-h,h]$ 的子集 $\overline{B}(\xi,\delta)$ 上才是压缩映射, 这里 ξ 是 $[-h,h]$ 上恒等于 ξ 的常值函数. 故为了使 T 是压缩映射, 同时其在 $\overline{B}(\xi,\delta)$ 内, 考虑令 $\mathscr{X}=\overline{B}(\xi,\delta)$, 并设

$$M:=\max\{|F(t,x)|:(t,x)\in[-h,h]\times[\xi-\delta,\xi+\delta]\},$$

取 $h>0$ 足够小, 使得

$$\max|(Tx)(t)-\xi|=\max\left|\int_0^t F(\tau,x(\tau))d\tau\right|\leq Mh\leq\delta.$$

可以证明 $(\mathscr{X},\rho)$ 依旧是完备度量空间, 此时 $T:\mathscr{X}\to\mathscr{X}$ 是压缩映射, 得到结果. 这便是常微分方程解的存在唯一性定理.

定理 2.1.4 (解的存在唯一性定理)

设函数 $F(t,x)$ 在 $[-h,h]\times[\xi-\delta,\xi+\delta]$ 上定义, 连续且满足局部 Lipschitz 条件:

$$\exists\delta>0\exists L>0\forall|t|\leq h\forall x_1,x_2\in\overline{B}_\delta(\xi)(|F(t,x_1)-F(t,x_2)|\leq L|x_1-x_2|),$$

则当 $h<\min\{\frac{\delta}{M},\frac{1}{L}\}$ 时, 初值问题

$$\begin{cases}\dfrac{dx}{dt}=F(t,x),\\ x(0)=\xi,\end{cases}$$

在 $[-h,h]$ 上存在唯一解.

下面是压缩映射原理的另一个推论: 隐函数存在定理.

定理 2.1.5 (隐函数存在定理)

设 $f(x,y)=(f_1(x,y),\cdots,f_m(x,y)):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m, U\times V\subset\mathbb{R}^n\times\mathbb{R}^m$ 是 (x_0,y_0) 在 $\mathbb{R}^n\times\mathbb{R}^m$ 中的一个邻域, 设 $f(x,y)$ 与

$$\frac{\partial f(x,y)}{\partial y}=\begin{pmatrix}\dfrac{\partial f_1}{\partial y_1}(x,y) & \cdots & \dfrac{\partial f_1}{\partial y_m}(x,y)\\ \vdots & & \vdots\\ \dfrac{\partial f_m}{\partial y_1}(x,y) & \cdots & \dfrac{\partial f_m}{\partial y_m}(x,y)\end{pmatrix},$$

在 $U\times V$ 内连续, 又设

$$f(x_0,y_0)\neq 0,\ \left(\det\left(\frac{\partial f}{\partial y}(x,y)\right)\right)(x_0,y_0)\neq 0,$$

则存在 (x_0,y_0) 的一个邻域 $U_0\times V_0\subset U\times V$ 以及唯一的连续函数 $\varphi:U_0\to V_0$, 满足

$$\begin{cases}f(x,\varphi(x))=0,\ x\in U_0,\\ \varphi(x_0)=y_0.\end{cases}$$

证明 考虑映射 $T:C(\overline{B}(x_0,r),\mathbb{R}^m)\to C(\overline{B}(x_0,r),\mathbb{R}^m),\varphi\mapsto T\varphi$ 满足

$$(T\varphi)(x)=\varphi(x)-f(x,\varphi(x))\left(\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^{-1}\right)^T,$$

其中 $r>0, C(\overline{B}(x_0,r),\mathbb{R}^m)$ 表示定义在闭球 $\overline{B}(x_0,r)$ 上, 取值在 $\mathbb{R}^m$ 中的向量值连续函数空间, 其上的距离定义为:

$$\rho(\varphi,\psi):=\max_{\substack{x\in\overline{B}(x_0,r)\\1\leq i\leq m}}|\varphi_i(x)-\psi_i(x)|,\ \varphi=(\varphi_1,\cdots,\varphi_m),\psi=(\psi_1,\cdots,\psi_m).$$

根据假设, $\frac{\partial f}{\partial y}(x,y)$ 在 $U\times V$ 上连续, 故

$$\exists\delta>0\left(\left|\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)_{ij}-\left(\frac{\partial f}{\partial y}(x,y)\right)_{ij}\right|<\frac{1}{2m}\right),\ i,j=1,2,\cdots,m,x\in\overline{B}(x_0,\delta),y\in\overline{B}(y_0,\delta),$$

其中 $(\cdot)_{ij}$ 表示括号内矩阵的第 i 行, 第 j 列元素. 现在记

$$d_i(x)=\varphi_i(x)-\psi_i(x),\ i=1,2,\cdots,m,$$

根据微分中值定理有:

$$\begin{aligned}\rho(T\varphi,T\psi)&=\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}\left|\varphi_i(x)-\left(f(x,\varphi(x))\left(\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^{-1}\right)^T\right)_i-\psi_i(x)+\left(f(x,\psi(x))\left(\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^{-1}\right)^T\right)_i\right|\\&=\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}\left|d_i(x)-\left((f(x,\varphi(x))-f(x,\psi(x)))\left(\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^{-1}\right)^T\right)_i\right|\\&=\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}\left|d_i(x)-\left((\varphi(x)-\psi(x))\left(\frac{\partial f}{\partial y}(x,\hat{y})\right)^T\left(\left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^{-1}\right)^T\right)_i\right|\\&\le\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}|d_i(x)-(\varphi_i(x)-\psi_i(x))(1+\tau_i)|\le\frac{1}{2}\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}|d_i(x)|=\frac{1}{2}\rho(\varphi,\psi),\end{aligned}\tag{2.4}$$

其中 $0<\tau_i<1, i=1,2,\cdots,m$, 且

$$\varphi,\psi\in\mathscr{X},\hat{y}(x)=(\hat{y}_1(x),\hat{y}_2(x),\cdots,\hat{y}_m(x)),$$
$$\mathscr{X}:=\{\varphi\in C(\overline{B}(x_0,r),\mathbb{R}^m):\varphi(x_0)=y_0,\varphi(x)\in\overline{B}(y_0,\delta),x\in\overline{B}(x_0,r)\},$$
$$\hat{y}_i(x)=\theta_i(x)\varphi(x)+(1-\theta_i(x))\psi(x),0<\theta_i(x)<1,x\in\overline{B}(x_0,r).$$

可以证明 $\mathscr{X}$ 在 $C(\overline{B}(x_0,r),\mathbb{R}^m)$ 中是闭的, 进而是完备度量空间. 现在式(2.4)表明 $T:\mathscr{X}\to C(\overline{B}(x_0,r),\mathbb{R}^m)$ 是压缩映射, 故只需证 $T:\mathscr{X}\to\mathscr{X}$ 即可. 注意

$$\rho(T\varphi,y_0)\le\rho(T\varphi,Ty_0)+\rho(Ty_0,y_0)\le\frac{1}{2}\rho(\varphi,y_0)+\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}|f_i(x,y_0)|,$$

其中在讨论 Ty_0 时, 是将 y_0 视作映射: $x\mapsto y_0,x\in\overline{B}(x_0,r)$, 此时显见 $y_0\in\mathscr{X}$. 根据 f 的连续性知:

$$\exists\eta>0\left(\max_{\substack{x\in\overline{B}(x_0,r)\\1\le i\le m}}|f_i(x,y_0)|<\frac{1}{2}\delta\right),$$

因此, 当 $0<r<\min\{\eta,\delta\}$ 时, 有

$$\rho(Ty_0,y_0)<a\delta\Rightarrow\rho(T\varphi,y_0)\le\frac{1}{2}\rho(\varphi,y_0)+\frac{1}{2}\delta\le\frac{1}{2}\delta+\frac{1}{2}\delta=\delta,\ \varphi\in\mathscr{X}.$$

此外

$$(T\varphi)(x_0)=\varphi(x_0)-f(x_0,\varphi(x_0))=y_0-f(x_0,y_0),$$

这说明 $T:\mathscr{X}\to\mathscr{X}$, 进而根据压缩映射原理知 T 有唯一不动点, 命题即证. □

补充命题 2.1.1 (♡)

设 $F(x,y)$ 是 $[a,b]\times(-\infty,\infty)$ 上的二元连续函数, 且 $F(x,y)$ 对 y 的一阶偏导数 $\frac{\partial F}{\partial y}(x,y)$ 处处连续, 满足条件:

$$0<m\le\frac{\partial F}{\partial y}(x,y)\le M,\ a\le x\le b,-\infty<y<\infty,$$

则存在连续函数 $y=u(x),a\le x\le b$ 满足

$$F(x,u(x))=0,\ x\in[a,b].$$

证明 令 $\mathscr{X}=C[a,b]$, 定义算子 T:

$$(Tf)(x):=f(x)-\frac{2}{M+m}F(x,f(x)),\ x\in[a,b],f\in C[a,b],$$

下面证明 T 是压缩映射. 任取 $f_1,f_2\in C[a,b]$, 由中值定理知:

$$F(x,f_1(x))-F(x,f_2(x))=\frac{\partial F}{\partial y}(x,\theta(x))(f_1(x)-f_2(x)).$$

故

$$
\begin{aligned}
|Tf_1(x)-Tf_2(x)| &= \left|f_1(x)-f_2(x)-\frac{2}{M+m}(F(x,f_1(x))-F(x,f_2(x)))\right| \\
&= |f_1(x)-f_2(x)|\cdot\left|1-\frac{2}{M+m}\frac{\partial F}{\partial y}(x,\theta(x))\right| \le \frac{M-m}{M+m}\rho(f_1,f_2).
\end{aligned}
$$

这说明 $\rho(Tf_1,Tf_2)\le\frac{M-m}{M+m}\rho(f_1,f_2)$, 从而根据 Banach 不动点定理知 T 有不动点 $y=u(x)\in C[a,b]$, 得到

$$
Ty=y\Rightarrow u(x)-\frac{2}{M+m}F(x,u(x))=u(x)\Rightarrow F(x,u(x))=0.
$$

□

补充命题 2.1.2 (♡)

给定 $Ax=c$, 其中 $A=(a_{ij})_{n\times n}, x^T=(x_1,\cdots,x_n), c=(c_1,\cdots,c_n)$, 若 A 满足

$$
\sum_{\substack{1\le j\le 1\\ j\ne i}}|a_{ij}|<|a_{ii}|,\ i=1,2,\cdots,n
$$

则 A 可逆, 且 $x=A^{-1}c$.

证明 令 $A=B+G$, 其中 $B=\operatorname{diag}(a_{11},\cdots,a_{nn})$, 则根据条件知 $|a_{ii}|>0(i=1,2,\cdots,n)$, 从而 B 可逆, 得到:

$$
\begin{aligned}
Ax=c &\Leftrightarrow Bx=-Gx+c \\
&\Leftrightarrow x=-B^{-1}Gx+B^{-1}c.
\end{aligned}
$$

将 $\mathbb{R}^n$ 视作 l^∞ 的子空间 $((x_1,\cdots,x_n)\mapsto(x_1,\cdots,x_n,0,\cdots))$, 在 $\mathbb{R}^n$ 上赋 l^∞ 距离:

$$
d_\infty(x,y)=\max_{1\le i\le n}|x_i-y_i|,
$$

则 $\mathbb{R}^n_\infty=(\mathbb{R}^n,d_\infty)$ 为完备距离空间. 定义:

$$
T:\mathbb{R}^n_\infty\to\mathbb{R}^n_\infty,\ x\mapsto Tx=(-B^{-1}G)x+B^{-1}c.
$$

下面说明 T 是压缩映射. 注意

$$
B=\begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn}\end{pmatrix}\Rightarrow B^{-1}=\begin{pmatrix} \frac{1}{a_{11}} & & \\ & \ddots & \\ & & \frac{1}{a_{nn}}\end{pmatrix},
$$

而

$$
G=\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix},
$$

故

$$
B^{-1}G=\begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \frac{a_{13}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \frac{a_{23}}{a_{22}} & \cdots & \frac{a_{2n}}{a_{22}} \\ \frac{a_{31}}{a_{33}} & \frac{a_{32}}{a_{33}} & 0 & \cdots & \frac{a_{3n}}{a_{33}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \frac{a_{n3}}{a_{nn}} & \cdots & 0\end{pmatrix}=:(d_{ij})_{n\times n}.
$$

现对任意的 $x, y \in \mathbb{R}^n_\infty$, 知 $Tx - Ty = -B^{-1}G(x-y)$, 从而

$$d_\infty(Tx - Ty) = \max_{1\le j\le n} |(Tx)_j - (Ty)_j| = \max_{1\le j\le n}\left|\sum_{i=1}^n d_{ji}(x_i - y_i)\right|$$
$$\le \max_{1\le j\le n}\sum_{i=1}^n |d_{ji}|\cdot|x_i - y_i| \le \max_{1\le j\le n}\left(\sum_{i=1}^n |d_{ji}|\right)\cdot\max_{1\le i\le n}|x_i - y_i|,$$

根据命题条件:

$$\sum_{i=1}^n |d_{ji}| = \sum_{\substack{1\le i\le n\\ i\ne j}}\left|\frac{a_{ji}}{a_{jj}}\right| = \left|\frac{1}{a_{jj}}\right|\sum_{\substack{1\le i\le n\\ i\ne j}}|a_{ji}| < 1,\ \forall 1\le j\le n.$$

这说明 $\max_{1\le j\le n}\sum_{i=1}^n |d_{ji}| < 1$, 从而 T 是压缩映射, 其有唯一不动点, 命题即证. □

证明

2.1.2 一些例子

本节例子若无特别说明均选自 **[WL]**.

2.1.2.1 度量空间

例 2.4(距离空间中两个点列的极限存在, 但它们的和不等于它们和的极限) 对 $x, y \in \mathbb{R}^1$, 令

$$d(x,y) = \begin{cases} 0, & x = y, \\ \max\{|x|, |y|\}, & x \ne y, \end{cases}$$

则容易证明 d 是 $\mathbb{R}^1$ 上的一个距离. 现取 $x_n = 1, y_n = -\frac{1}{n}, n = 1, 2, \cdots$, 则

$$\lim_{n\to\infty} x_n = 1,\ \lim_{n\to\infty} y_n = 0.$$

但对任意正整数 n, 总有 $d(1-\frac{1}{n}, 1) = 1$, 这说明至少

$$\lim_{n\to\infty}(x_n + y_n) \ne 1,$$

故

$$\lim_{n\to\infty}(x_n + y_n) \ne \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n.$$

注 上述反例表明: 在度量空间中两个点列的和的极限不一定是它们极限的和.

例 2.5(一个集合上的两个距离, 使得这个集合上没法再定义同时小于它们的距离) 首先, 对任意非空集 X 上的任意两个距离 d_1, d_2, 当然可以在 X 上定义一个新的距离 d_3, 使得对一切的 $x, y \in X$ 都有:

$$d_3(x,y) \ge d_1(x,y),\ d_3(x,y) \ge d_2(x,y),$$

也即可以定义一个同时大于这两个距离的距离. 令 $d_3 = d_1 + d_2$ 即证命题. 但至于说能不能定义一个同时小于这两个距离的距离, 答案是未必的.

譬如, 设 $X = \{x : 0 \le x \le 1\}$, 在 X 上取通常距离 $d_1(x,y) = |x-y|, x, y \in X$. 令

$$f(x) = \begin{cases} 1, & x = 0, \\ 0, & x = 1, \\ x, & 0 < x < 1, \end{cases}$$

并定义 $d_2(x,y) = d_1(f(x), f(y))$, 则当 $0 < x < 1$ 时有

$$d_1(x, 0) = |x| = x,$$
$$d_2(x, 1) = d_1(f(x), f(1)) = d_1(x, 0) = x.$$

下面证明不存在对一切 $x, y \in X$ 都同时小于等于 d_1, d_2 的距离 d. 用反证法, 假设存在这样的距离 d, 则

$$\forall x \in (0,1)(d(x,0) \le d_1(x,0) = x, d(x,1) \le d_2(x,1) = x),$$

进而根据三角不等式有

$$d(0,1) \le d(0,x) + d(x,1) \le 2x,$$

但上不等式左端是常数, 右端是任意的量, 矛盾! 命题即证.

例 2.6(两个不相交的有界闭集, 它们之间的距离为零) 取 $X = (0,1) \cup (1,2)$, 在 X 上取通常距离 $d(x,y) = |x-y|, x, y \in X$, 则 X 是一个度量空间. 显见 $(0,1), (1,2)$ 均为 X 中的两个不相交的有界闭集, 而它们之间的距离为零.

注 回忆闭集的概念: 如果度量空间 $(\mathscr{X}, \rho)$ 的子集 E 满足:

$$\forall \{x_n\} \subset E(x_n \to x_0 \in \mathscr{X} \Rightarrow x_0 \in E),$$

则 E 是该空间中的闭集. 所以在这个例子中, (0,1) 能作为给出的空间的闭集正是因为 0,1 这两个点本身都不在这个空间内.

后面的习题中会证明, 度量空间中任意两个不相交的非空闭集, 只要其中有一个是紧集, 那么它们之间就有正值的距离, 所以这个例子之所以成立, 是因为给出的这两个闭集都不是紧集 (譬如取开集族 $\{(0, 1 - \frac{1}{n}) : n \ge 1\}$).

例 2.7(一个度量空间中的某个开球是闭集但不是闭球, 某个闭球是开集但不是开球) 设 $X = \{x = (\xi, \eta) : -\frac{1}{2} \le \xi \le \frac{1}{2}, \eta = 0\} \cup \{(0,1)\} \cup \{(0,-1)\}$, 在 X 上取通常距离, 则 $\{x = (\xi, \eta) : x \in X, d(x,0) < 1\}$ 是 X 中的开球也是闭集, 但它不是闭球; $\{x = (\xi, \eta) : x \in X, d(x,0) \le \frac{1}{2}\}$ 是 X 中的闭球也是开集, 但它不是开球.

注 上述例子说明在一些度量空间中, 虽然开球与闭球的球心和半径都相同, 但开球的闭包未必是一个闭球.

在任意赋范线性空间中这种例子都是不存在的.

例 2.8(某个度量空间中开球的闭包都是闭球, 但它的某个子空间没有这个性质) 显见 $\mathbb{R}^1$ 中每个开球的闭包都是闭球, 取 $\mathbb{R}^1$ 的度量子空间 $X = [0,1] \cup [2,3]$, 则 X 中开球 $\{x \in X : d(1,x) < 1\}$ 的闭包是 $[0,1]$, 但 $\{x \in X : d(1,x) \le 1\} = [0,1] \cup \{2\}$, 故开球 $\{x \in X : d(1,x) < 1\}$ 的闭包不是闭球.

例 2.9(某个度量空间中半径分别为 r_1, r_2 的闭球 B_1, B_2, 尽管 $r_1 > r_2$, 但 $B_1 \subset B_2$) 设 X 为 $\mathbb{R}^2$ 上的闭圆盘: $X = \{(x,y) : x^2 + y^2 \le 9\}$, 在 X 上取通常距离, 设 $B_2 = X$, 设 B_1 为:

$$B_1 = \{(x,y) : (x,y) \in X, d((2,0),(x,y)) \le 4\} = B_2 \cap \{(x,y) : (x-2)^2 + y^2 \le 16\}.$$

故 $B_1 \subset B_2$, 但 $r_1 = 4 > 3 = r_2$.

注 上述例子表明在度量空间中半径为 4 的球可能是半径为 3 的球的真子集, 特别注意上述例子里 B_1 的半径 $r_1 = 4$, 而直径则为 6.

在任意赋范线性空间中这种例子都是不存在的.

课堂笔记 (♡)

两个是度量空间但不是赋范空间的例子.

- 设 E 是 $\mathbb{R}^n$ 中的有界 Lebesgue 可测集, $m(E) < \infty$, 记 $S(E)$ 为 E 上几乎处处有限的可测函数全体, 其中几乎处处相等的函数视作同一函数. 定义

$$\rho(f,g) = \int_E \frac{|f(t) - g(t)|}{1 + |f(t) - g(t)|} dt,$$

则容易验证 ρ 是良定义的二元函数, 进一步可以验证 ρ 是距离, 从而 $(S(E), \rho)$ 是距离空间, 下面证明

$(S(E), \rho)$ 是完备的. 取 $\{f_k\} \subset S(E), f \in S(E)$, 首先说明 $\rho(f_k, f) \to 0$ 当且仅当 f_k 依 (Lebesgue) 测度收敛到 f, 亦即

$$\forall \sigma > 0(m\{t \in E : |f_k(t) - f(t)| \geq \sigma\} \to 0, k \to \infty).$$

当 $\rho(f_k, f) \to 0$, 按照定义此即

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall k \geq N \left(\int_E \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt < \varepsilon \right).$$

记

$$A_k(\sigma) := \{t \in E : |f_k(t) - f(t)| \geq \sigma\},\ B_k(\sigma) = E \backslash A_k(\sigma),$$

则根据函数 $\frac{x}{1+x}$ 的递增性知:

$$\forall t \in A_k(\sigma) \left(\frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} \geq \frac{\sigma}{1+\sigma} \right).$$

现在将 E 上的积分分为 $A_k(\sigma), B_k(\sigma)$ 两部分得:

$$\begin{aligned} \varepsilon > \int_E \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt > \int_{A_k(\sigma)} \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt + \int_{B_k(\sigma)} \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt \\ \geq m(A_k(\sigma)) \cdot \frac{\sigma}{1+\sigma} + \int_{B_k(\sigma)} \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt + \geq \frac{\sigma}{1+\sigma} m(A_k(\sigma)), \end{aligned}$$

故

$$m(A_k(\sigma)) \to 0, k \to \infty.$$

这便是 f_k 依测度收敛到 f 的定义.
而当 f_k 依测度收敛到 f, 根据定义即

$$\forall \sigma > 0(m\{t \in E : |f_k(t) - f(t)| \geq \sigma\} \to 0, k \to \infty),$$

而根据函数 $\frac{u}{1+u}$ 的单调性:

$$|f_k(t) - f(t)| \geq \sigma \Leftrightarrow \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} \geq \frac{\sigma}{1+\sigma}.$$

这说明

$$m\left(t \in E : \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} > \frac{\sigma}{1+\sigma}\right) = m(t \in E : |f_k(t) - f(t)| > \sigma) \to 0,\ k \to \infty.$$

再根据 σ 的任意性知

$$\frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} \xrightarrow{m} 0,\ k \to \infty.$$

根据依测度收敛型的 Lebesgue 控制定理:

$$\int_E \frac{|f_k(t) - f(t)|}{1 + |f_k(t) - f(t)|} dt \to 0,\ k \to \infty.$$

此即 $\rho(f_k, f) \to 0 (k \to \infty)$[a].
最后来说明 $S(E)$ 的完备性, 设 $\{f_n\}$ 是 $S(E)$ 中的基本列, 根据基本列的定义有:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall m, n > N(\rho(f_n, f_m) < \varepsilon),$$

根据前述结论:

$$\rho(f_n, f_m) \to 0 \Leftrightarrow \forall \sigma > 0(m(t \in E : |f_n(t) - f_m(t)| > \sigma) \to 0, n, m \to \infty),$$

也即 $\{f_n\}$ 是依测度的基本列. 根据 Riesz 定理, 存在子列 $\{f_{n_k}\}$ 几乎处处收敛, 进而在点态情况下可以确定可测函数 $f(t)$ 使得

$$f_{n_k} \to f \text{ a.e..}$$

根据 $S(E)$ 的定义已经可知 $f\in S(E)$, 最后验证 $f_m\xrightarrow{\rho}f$. 根据基本列的性质知当 m,k 足够大时:

$$\rho(f_m,f_{m_k})<\varepsilon,$$

现令 $k\to\infty$, 代入 $\rho(\cdot,\cdot)$ 的定义并根据 Lebesgue 控制收敛定理得

$$\varepsilon\geq\lim_{k\to\infty}\rho(f_m,f_{m_k})=\lim_{k\to\infty}\int_E\frac{|f_m(t)-f_{m_k}(t)|}{1+|f_m(t)-f_{m_k}(t)|}dt=\int_E\frac{|f_m(t)-f(t)|}{1+|f_m(t)-f(t)|}dt=\rho(f_m,f).$$

故 $f_m\xrightarrow{\rho}f\in S(E)$, 因而 $S(E)$ 完备.

- 空间 S 表示一切实 (或复) 数列的全体, 对任意的 $x,y\in S$, 记

$$x=(x_1,x_2,\cdots),\ \ y=(y_1,y_2,\cdots),$$

记

$$\rho_n(x,y)=\frac{|x_n-y_n|}{1+|x_n-y_n|}(n=1,2,\cdots),\ \ \rho(x,y):=\sum_{n=1}^{\infty}\alpha_n\rho_n(x,y)<\infty,$$

其中正数列 $\{\alpha_n\}$ 满足

$$\sum_{n=1}^{\infty}\alpha_n<\infty,$$

则可以验证 ρ 是 S 上的距离函数, 下面证明 (S,ρ) 内点列的收敛等价于依坐标收敛, 也即若设 $x^{(k)}=\{x_n^{(k)}\}$, 则欲证 $\rho(x^{(k)},x)\to0\Leftrightarrow x_n^{(k)}\to x_n(\forall n\in\mathbb{N})$.

当 $\rho(x^{(k)},x)=\sum_{n=1}^{\infty}\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}\to0(k\to\infty)$, 则对每个 n 都有

$$\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}\leq\rho(x^{(k)},x)\to0,\ \ k\to\infty.$$

根据定义此即

$$\forall\varepsilon>0\exists N\in\mathbb{N}\forall k>N\left(\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}<\varepsilon\right).$$

而

$$\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}<\varepsilon\Leftrightarrow|x_n^{(k)}-x_n|<\frac{\varepsilon}{1-\varepsilon},$$

根据 ε 的任意性即得 $x_n^{(k)}\to x_n(k\to\infty)(\forall n\in\mathbb{N})$.

当 $\lim_{k\to\infty}x_n^{(k)}=x_n(n=1,2,\cdots)$, 则根据 $\sum_{n=1}^{\infty}\alpha_n<\infty$ 的性质有

$$\forall\varepsilon>0\exists N\in\mathbb{N}\forall m>N\left(\sum_{n=m+1}^{\infty}\alpha_n<\varepsilon\right)$$

而对 $i=1,2,\cdots,m$, 均存在对应的 $N_i\in\mathbb{N}$, 使得当 $k>N_i$ 时

$$|x_i^{(k)}-x_i|<\varepsilon,$$

取 $N'=\max\{N,N_1,\cdots,N_m\}$, 当 $k>N'$ 时有

$$\begin{aligned}\rho(x^{(k)},x)=\sum_{n=1}^{\infty}\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}&\leq\sum_{n=1}^{m}\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}+\sum_{n=m+1}^{\infty}\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}\\&\leq\sum_{n=1}^{m}\alpha_n\frac{|x_n^{(k)}-x_n|}{1+|x_n^{(k)}-x_n|}+\sum_{n=m+1}^{\infty}\alpha_n<\varepsilon\sum_{n=1}^{m}\alpha_n+\varepsilon=\left(\sum_{n=1}^{m}\alpha_n+1\right)\varepsilon,\end{aligned}$$

根据 ε 的任意性即得 $\rho(x^{(k)},x)\to0(k\to\infty)$, 命题即证.

[a] 老师上课的笔记提出另一种证法, 如果记 $g_k(t)=\frac{|f_k(t)-f(t)|}{1+|f_k(t)-f(t)|}$, 显见 $|f_k(t)-f(t)|\geq g_k(t)$, 所以 $m(t\in E:|f_k(t)-f(t)|>\sigma)\geq m(t\in E:g_k(t)>\sigma)$, 从而由前者趋零推知后者趋零.

注 实变函数论中对于依测度收敛的函数列一般有下面几条重要的处理方法与结论.

定义 2.1.10 (依测度基本列)

设 $\{f_k(x)\}$ 是 E 上几乎处处有限的可测函数列, 若对任给的 $\varepsilon > 0$, 都有

$$\lim_{\substack{k\to\infty \\ j\to\infty}} m(\{x \in E : |f_k(x) - f_k(x)| > \varepsilon\}) = 0,$$

则称 $\{f_k(x)\}$ 为 E 上的依测度基本列

下面这个定理一般用于确定依测度基本列的某个依测度极限函数.

定理 2.1.6

若 $\{f_k(x)\}$ 是 E 上的依测度基本列, 则在 E 上存在几乎处处有限的可测函数 $f(x)$, 使得 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$.

但依测度收敛无法点态地确定一个函数, 此时可以用下述定理来 (几乎处处) 点态地确定一个函数.

定理 2.1.7 (Riesz)

若 $\{f_k(x)\}$ 在 E 上依测度收敛于 $f(x)$, 则存在子列 $\{f_{k_i}(x)\}$, 使得

$$\lim_{i\to\infty} f_{k_i}(x) = f(x), \text{ a.e. } x \in E.$$

关于依测度收敛列的积分有两个重要结论:

定理 2.1.8 (依测度收敛型的 Fatou 引理)

设 $\{f_k(x)\}$ 是 E 上依测度收敛于 $f(x)$ 的非负可测函数列, 则:

$$\int_E f(x)dx \le \liminf_{k\to\infty} \int_E f_k(x)dx.$$

定理 2.1.9 (依测度收敛型的 Lebesgue 控制收敛定理)

设 $f_k \in L^1(\mathbb{R}^n)(k = 1, 2, \cdots)$, 且 $f_k(x)$ 在 $\mathbb{R}^n$ 上依测度收敛于 $f(x)$. 若存在 $F \in L(\mathbb{R}^n)$, 使得

$$|f_k(x)| \le F(X), \ k = 1, 2, \cdots, \text{a.e. } x \in \mathbb{R}^n,$$

则 $f \in L^1(\mathbb{R}^n)$, 且

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k(x)dx = \int_{\mathbb{R}^n} f(x)dx.$$

2.1.2.2 压缩映射

课堂笔记 (♡)

压缩映射定理的条件不能改成 $\rho(Tx, Ty) < \rho(x, y)(\forall x, y \in \mathscr{X})$, 也即弱压缩映射不满足压缩映射原理. 例如设 $\mathscr{X} = \mathbb{R}, Tx = \frac{\pi}{2} - \arctan x + x$, 则 T 是弱压缩映射, 但它没有不动点. 这是因为任取 $x, y \in \mathscr{X}$ 有

$$|Tx - Ty| = |x - y + \arctan y - \arctan x| = \left| x - y + \frac{y - x}{1 + \xi^2} \right| = \frac{\xi^2}{1 + \xi^2} |y - x| = \alpha(x, y)|y - x|, \ \xi \text{在 } x, y \text{ 之间}.$$

所以本例同时说明 α 不能与 x, y 有关.

例 2.10(T^2 是压缩映射, 但 T 本身不是压缩映射) 本例见习题2.5.

2.1.3 习题

✍ **练习 2.1** 证明: 完备空间的闭子集是一个完备的子空间, 而任一度量空间中的完备子空间必是闭子集.

证明 届于完备空间 $\mathscr{X}$ 的闭子集 E 继承了 $\mathscr{X}$ 的距离, 只需证明 E 是完备空间. 任取 $\{x_n\}$ 是 E 中的基本列, 注意 $\{x_n\} \subset E \subset \mathscr{X}$, 故由 $\mathscr{X}$ 的完备性推知 $\{x_n\}$ 在 $\mathscr{X}$ 中有极限 x, 根据 E 的闭性知 $x \in E$, 故 $\{x_n\}$ 在 E 中收敛, 也即 E 为完备空间.

另一方面, 如若 $E \subset \mathscr{X}$ 作为完备子空间并非闭子集, 考虑 $E \ni x_n \to x \notin E$, 既然 $\mathscr{X}$ 是完备空间, 知 $\{x_n\}$ 是基本列, 这与 E 的完备性矛盾! 命题即证.

练习 2.2(Newton 法)★ 设 f 是定义在 $[a,b]$ 上的二次连续可微的实值函数, $\hat{x} \in (a,b)$ 使得 $f(\hat{x}) = 0, f'(\hat{x}) \neq 0$. 求证: 存在 $\hat{x}$ 的邻域 $U(\hat{x})$, 使得 $\forall x_0 \in U(\hat{x})$, 迭代序列

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},\ n = 0, 1, 2, \cdots$$

是收敛的, 并且

$$\lim_{n\to\infty} x_n = \hat{x}.$$

证明 取 $T(x) = x - \frac{f(x)}{f'(x)}$, 则既然 $f \in C^2[a,b]$, 有

$$T'(x) = 1 - \frac{f'^2(x) - f(x)f''(x)}{f'^2(x)} = \frac{f(x)f''(x)}{f'^2(x)},\ x \in (a,b).$$

既然 $f(\hat{x}) = 0$, 知 $T'(\hat{x}) = 0$, 根据 T' 的连续性知

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in U_\delta(\hat{x})(|T'(x)| < \varepsilon).$$

现在在 $B(\hat{x},\delta)$ 中任取两点 x_1, x_2, 有

$$\left|\frac{T(x_1) - T(x_2)}{x_1 - x_2}\right| = |T'(\xi)| < \varepsilon,\ \xi \in B(\hat{x},\delta)$$

当然可以取 $\varepsilon < 1$, 从而 T 是压缩映射, 进而 $\forall x_0 \in B(\hat{x},\delta)$, 迭代序列

$$x_{n+1} = T(x_n),\ n = 0, 1, 2, \cdots$$

是收敛的, 且注意到 $T(\hat{x}) = \hat{x}$, 根据压缩映射不动点的唯一性知 $\lim_{n\to\infty} x_n = \hat{x}$.

练习 2.3 设 $(\mathscr{X},\rho)$ 是度量空间, 映射 $T: \mathscr{X} \to \mathscr{X}$ 满足

$$\rho(Tx, Ty) < \rho(x, y),\ \forall x \neq y,$$

并已知 T 有不动点, 求证: 此不动点是唯一的.

证明 如若存在 $x^* \neq x^{**}$ 使得

$$Tx^* = x^*, Tx^{**} = x^{**},$$

则

$$\rho(Tx^*, Tx^{**}) = \rho(x^*, x^{**}) < \rho(x^*, x^{**}),$$

矛盾! 命题即证.

练习 2.4★ 设 T 是度量空间上的压缩映射, 求证: T 是连续的.

证明 任取 $(\mathscr{X},\rho)$ 中的点列 $\{x_n\}$ 与 x_0, 注意

$$\rho(Tx_n, Tx_0) \le a\rho(x_n, x_0),\ 0 < a < 1, n = 1, 2, \cdots,$$

取极限有

$$0 \le \lim_{n\to\infty} \rho(Tx_n, Tx_0) \le a \cdot 0 = 0,$$

得到 $\lim_{n\to\infty} \rho(Tx_n, Tx_0) = 0$, 命题即证.

练习 2.5★ 设 T 是压缩映射, 求证: $T^n (n \in \mathbb{N})$ 也是压缩映射, 并说明逆命题不一定成立.

证明 当 T 是压缩映射, 任取 $x_1, x_2 \in (\mathscr{X},\rho)$, 知存在 $0 < a < 1$ 使得:

$$\rho(T^n x_1, T^n x_2) = \rho(T(T^{n-1}x_1), T(T^{n-1}x_2)) \le a\rho(T^{n-1}x_1, T^{n-1}x_2) \le \cdots \le a^n \rho(x_1, x_2).$$

这说明 $T^n (n \in \mathbb{N})$ 也是压缩映射.

下面给出逆命题的两个来自 **[WL]** 的反例.

其一: T 是连续映射, 但不是压缩映射, 而 T^2 是压缩映射. 设 $T: C[0,1] \to C[0,1]$ 定义为: 对 $x \in C[0,1]$, 取

$$Tx = y := \int_0^t x(u)du,\ 0 \le t \le 1.$$

现取 $x_1(t) \equiv C_1, x_2(t) \equiv C_2$, 则

$$d(Tx_1, Tx_2) = \max_{0\le t\le 1}\left|\int_0^t C_1 du - \int_0^t C_2 du\right| = |C_1 - C_2| = d(x_1, x_2),$$

这并不满足压缩映射的定义! 故 T 本身连续, 但不是压缩映射. 再研究 $T^2: C[0,1] \to C[0,1]$, 任取 $x_1, x_2 \in C[0,1]$, 有

$$d(T^2x_1, T^2x_2) = \max_{0\le t\le 1}\left|\int_0^t\int_0^u (x_1(v) - x_2(v))dvdu\right| \le \max_{0\le t\le 1}\left|\int_0^t ud(x_1,x_2)du\right| = \frac{1}{2}d(x_1,x_2).$$

故 T^2 是压缩映射.

其二: T 是不连续映射, 但 T^2 是压缩映射[2]. 设 $f: [0,2] \to [0,2]$ 定义为:

$$f(x) = \begin{cases} 0, & x \in [0,1], \\ 1, & x \in (1,2], \end{cases}$$

则显见任取 $x \in [0,2]$ 总有 $f[f(x)] = 0$, 故 f^2 是 $[0,2]$ 到 $[0,2]$ 的压缩映射, 同时 f 本身不连续. □

✍ **练习 2.6** 设 M 是 $(\mathbb{R}^n, \rho)$ 中的有界闭集, 映射 $T: M \to M$ 满足: $\rho(Tx, Ty) < \rho(x,y)(\forall x, y \in M, x \ne y)$. 求证: T 在 M 中存在唯一的不动点.

证明 考虑说明存在常数 a 使得

$$\rho(Tx, Ty) < a\rho(x,y),\ \forall x, y \in M, x \ne y.$$

如若不然, 则知

$$\forall \varepsilon > 0 \exists x_\varepsilon, y_\varepsilon \in M(\rho(Tx_\varepsilon, Ty_\varepsilon) \ge (1-\varepsilon)\rho(x_\varepsilon, y_\varepsilon)),$$

取 $\varepsilon = \frac{1}{n}, n = 1, 2, \cdots$ 即得序列 $\{x_n\}, \{y_n\} \in M$ 满足

$$\lim_{n\to\infty}\rho(Tx_n, Ty_n) = \lim_{n\to\infty}\rho(x_n, y_n),$$

既然 M 是闭集, 知 $\{x_n\}, \{y_n\}$ 在 M 中有极限, 记之为 x_0, y_0, 此时

$$\rho(Tx_0, Ty_0) = \rho(x_0, y_0),\ x_0, y_0 \in M,$$

矛盾! 命题即证.

✍ **练习 2.7**♡ 对于积分方程

$$x(t) - \lambda\int_0^1 e^{t-s}x(s)ds = y(t)$$

其中 $y(t) \in C[0,1]$ 为一给定函数, λ 为常数, $|\lambda| < 1$, 求证: 存在唯一解 $x(t) \in C[0,1]$.

证明 注意

$$x(t) - \lambda\int_0^1 e^{t-s}x(s)ds = y(t) \Leftrightarrow e^{-t}x(t) - \lambda\int_0^1 e^{-s}x(s)ds = e^{-t}y(t),$$

现在记

$$e^{-t}x(t) = z(t), e^{-t}y(t) = \zeta(t),\ t \in [0,1],$$

得到

$$z(t) = \lambda\int_0^1 z(s)ds + \zeta(t).$$

[2]注意: 压缩映射必连续.

令 $(Tz)(t)=\lambda\int_0^1 z(s)ds+\zeta(t)$, 现在证明 T 是 $C[0,1]$ 上的压缩映射:

$$\begin{aligned}\rho(Tz_1,Tz_2)&=\max_{t\in[0,1]}|Tz_1-Tz_2|=\max_{t\in[0,1]}\left|\lambda\int_0^1(z_1(s)-z_2(s))ds\right|\\&\le|\lambda|\max_{t\in[0,1]}\int_0^1|z_1(s)-z_2(s)|ds\le|\lambda|\max|z_1(t)-z_2(t)|\\&=|\lambda|\rho(z_1,z_2),\end{aligned}$$

这说明 $T:C[0,1]\to C[0,1]$ 是压缩映射, 进而由压缩映射原理即得欲证.

2.2 完备化

2.2.1 知识梳理

首先注意一个压缩映射原理在非完备空间中不成立的例子:

例 2.11 函数 $T(x):=\frac{1}{2}\sqrt{x+1}$ 在 $[0,1]$ 上有定义, 有唯一不动点 $x_0=\frac{\sqrt{17}+1}{8}$. 若取 $\mathscr{X}=[0,1]\backslash\{x_0\}$, 则 T 依旧是 $\mathscr{X}\to\mathscr{X}$ 的压缩映射, 但却不再有不动点.

空间是否完备是与距离紧密相关的, 譬如 $C[a,b]$ 在一致收敛性度量下完备, 但若赋以 L^1 距离:

$$\rho_1(x,y):=\int_a^b|x(t)-y(t)|dt,$$

则 $(C[a,b],\rho_1)$ 不再完备. 以 $C[0,1]$ 为例, 这是因为考虑函数列

$$f_n(x)=\begin{cases}1-nx, & x\in[0,\frac{1}{n}),\\0, & x\in[\frac{1}{n},1],\end{cases}$$

显见其在 L^1 距离下是基本列, 但其极限函数并不在 $C[a,b]$ 中. 下面仿照实数理论中从有理数域出发定义无理数的方法, 对空间 $(\mathscr{X},\rho)$ 增添元素使之完备化.

定义 2.2.1 (等距同构)

设 $(\mathscr{X},\rho_1),(\mathscr{X}_1,\rho_1)$ 是两个度量空间, 如果存在映射 $\varphi:\mathscr{X}\to\mathscr{X}_1$ 满足

(i) φ 是满射,

(ii) $\rho(x,y)=\rho_1(\varphi x,\varphi y),\forall x,y\in\mathscr{X}$

则称 $(\mathscr{X},\rho)$ 和 $(\mathscr{X}_1,\rho_1)$ 是等距同构的, 并称 φ 为等距同构映射, 简称为等距同构. ♣

定义 2.2.2 (嵌入)

如果度量空间 $(\mathscr{X}_1,\rho_1)$ 与另一个度量空间 $(\mathscr{X}_2,\rho_2)$ 的子空间 $(\mathscr{X}_0,\rho_2)$ 是等距同构的, 我们就说 $(\mathscr{X}_1,\rho_1)$ 可以嵌入 $(\mathscr{X}_2,\rho_2)$. 在嵌入的意义下, 也称 $(\mathscr{X}_1,\rho_1)$ 是 $(\mathscr{X}_2,\rho_2)$ 的一个子空间, 简记作 $(\mathscr{X}_1,\rho_1)\subset(\mathscr{X}_2,\rho_2)$. ♣

定义 2.2.3 (稠密子集)

设 $(\mathscr{X},\rho)$ 是度量空间. 集合 $E\subset\mathscr{X}$ 叫作在 $\mathscr{X}$ 中的稠密子集, 如果

$$\forall x\in\mathscr{X}\forall\varepsilon>0\exists z\in E(\rho(x,z)<\varepsilon),$$

也即

$$\forall x\in\mathscr{X}\exists\{x_n\}\subset E(x_n\to x,n\to\infty).$$

♣

课堂笔记 (⋆)

- 根据等距同构定义 (ii) 知等距同构映射 φ 是单射, 这是因为若 $\varphi(x)=\varphi(y)$, 则 $\rho(x,y)=\rho_1(\varphi(x),\varphi(y))=0$, 进而 $x=y$, φ 是单射. 进一步知 $\varphi^{-1}:\mathscr{X}_1\to\mathscr{X}$ 也是等距同构映射.

- 从定义显见, 对于等距同构的两个度量空间, 它们的一切与距离相联系的性质都是一样的.
- 存在无穷维距离空间与其自身真子空间等距同构. 如:
$$T: l^2 \to \tilde{l}^2 \subsetneq l^2, (x_1, \cdots, x_n, \cdots) \mapsto (0, x_1, \cdots, x_n, \cdots),$$
则 T 是满射且等距, 故 l^2 与 $\tilde{l}^2$ 等距同构.
- 稠密子集有等价定义: 设 $(\mathscr{X}, \rho)$ 是度量空间, 称 $E \subset \mathscr{X}$ 是 $\mathscr{X}$ 中的稠密子集, 如果
$$\forall x \in \mathscr{X} \forall \varepsilon > 0 \exists z \in E(\rho(x,z) < \varepsilon).$$
两者等价的思想类似于实数理论中极限点引理的两种表述.

例 2.12 $[a,b]$ 上的多项式全体记为 $P[a,b]$, 根据 Weierstrass 定理知 $P[a,b]$ 在 $C[a,b]$ 中稠密.

课堂笔记 (⋆)

例(2.12)中 $[a,b]$ 作为有界闭区间 (进而是紧集) 的条件是至关重要的, 否则结论不一定成立, 即存在 $\mathbb{R}$ 上的 (有界) 连续函数 f 使得不存在多项式列 $\{P_n\}_{n\in\mathbb{N}}$ 满足
$$\sup_{x\in\mathbb{R}} |f(x) - P_n(x)| \to 0,\ n \to \infty. \tag{2.5}$$
事实上, 容易构造在 $(-\infty, 0)$ 无界, 但在 $[0, +\infty)$ 有界且非常数的连续函数 $f(x)$, 此时若存在 $\mathbb{R}$ 上的多项式列 $\{P_n\}_{n\in\mathbb{N}}$ 满足(2.5)式, 则
$$\sup_{[0,\infty)} |P_n(x)| \le \sup_{[0,\infty)} |P_n(x) - f(x)| + \sup_{[0,\infty)} |f(x)| \le \sup_{\mathbb{R}} |P_n(x) - f(x)| + \sup_{[0,\infty)} |f(x)|.$$
根据(2.5)式定义知对 $\varepsilon_0 = 1$ 而言:
$$\exists N_0 \in \mathbb{N} \forall n \ge N_0 (\sup_{\mathbb{R}} |P_n(x) - f(x)| < \varepsilon_0 = 1) \Rightarrow \sup_{[0,\infty)} |P_n(x)| \le 1 + \sup_{[0,\infty)} |f(x)| < \infty.$$
这说明 $n \ge N_0$ 时 $P_n(x)$ 只能为常值函数, 得到矛盾. 进一步, 设 f^+ 是前面构造的函数 $f(x)$ 在 $x \in \mathbb{R}^+$ 的部分, 对 f^+ 作偶延拓得到 $\tilde{f}$, 将上述 f 换为 $\tilde{f}$ 知证明依旧成立, 从而至少在 $\mathbb{R}$ 上 Weierstrass 定理不再成立, 进而结论不再成立, $P(\mathbb{R})$ 并不在 $C(\mathbb{R})$ 中稠密[a].

[a]直观的联想诸如 Runge 现象.

定义 2.2.4 (完备化空间)

包含给定度量空间 $(\mathscr{X}, \rho)$ 的最小的完备度量空间称为 $\mathscr{X}$ 的完备化空间, 其中最小的含义是: 任何一个以 $(\mathscr{X}, \rho)$ 为子空间的完备度量空间都以此空间为子空间.

命题 2.2.1

如果 $(\mathscr{X}_1, \rho_1)$ 是一个以 $(\mathscr{X}, \rho)$ 为子空间的完备度量空间, $\rho_1|_{\mathscr{X}\times\mathscr{X}} = \rho$, 并且 $\mathscr{X}$ 在 $\mathscr{X}_1$ 中稠密, 则 $\mathscr{X}_1$ 是 $\mathscr{X}$ 的完备化空间.

证明 事实上, 根据 $\mathscr{X}$ 在 $\mathscr{X}_1$ 中稠密知
$$\forall \xi \in \mathscr{X}_1 \exists x_n \in \mathscr{X} (\rho_1(x_n, \xi) \to 0, n \to \infty).$$
如果存在 $(\mathscr{X}_2, \rho_2)$ 以 $(\mathscr{X}, \rho)$ 为子空间, 因为
$$\rho_2(x_n, x_m) = \rho_1(x_n, x_m) \to 0,\ n, m \to \infty,$$
故 $\{x_n\}$ 同样是 $\mathscr{X}_2$ 中的基本列, 因为 $(\mathscr{X}_2, \rho_2)$ 是完备的, 故其在 $\mathscr{X}_2$ 中有极限:
$$\exists \hat{\xi} \in \mathscr{X}_2, \rho_2(x_n, \hat{\xi}) \to 0.$$
作映射:
$$T: \mathscr{X}_1 \to \mathscr{X}_2,\ \xi \mapsto \hat{\xi}.$$

下面证明 T 是等距的, 这是因为再取 $\eta \in \mathscr{X}_1$, 知 $\exists y_n \in \mathscr{X}$ 使得 $\rho_1(y_n, \eta) \to 0$, 故

$$\rho_1(\xi, \eta) = \lim_{n\to\infty} \rho_1(x_n, y_n) = \lim_{n\to\infty} \rho_2(x_n, y_n) = \rho_2(\hat{\xi}, \hat{\eta}).$$

这说明 $(\mathscr{X}_1, \rho_1)$ 能嵌入到 $(\mathscr{X}_2, \rho_2)$ 中, 进而是其一个子空间. □

补充命题 2.2.1 (完备化空间的唯一性 ⋆)

设 $(\mathscr{X}, \rho)$ 是度量空间, $(\mathscr{X}', \rho')$ 是其完备化空间. 若 $(\mathscr{X}, \rho)$ 在等距同构的意义下在 $(\mathscr{X}', \rho')$ 中稠密 (即 $(\mathscr{X}, \rho)$ 与 $(\mathscr{X}', \rho')$ 的某个稠密子集等距同构), 则具有这种性质的完备化空间 $(\mathscr{R}', \rho')$ 是唯一的.

证明 设 $(\widetilde{\mathscr{X}}, \widetilde{\rho})$ 是另一个满足命题条件的完备化空间, 也即 $\mathscr{X}$ 同时等距同构于 $\widetilde{\mathscr{X}}$ 与 $\mathscr{X}'$ 的稠密子空间, 设相应的等距同构分别为 $\widetilde{\varphi}$ 和 φ', 下面证明 $\widetilde{\mathscr{X}}$ 与 $\mathscr{X}'$ 等距同构.

任取 $\xi \in \widetilde{\mathscr{X}}$, 因为 $\mathscr{X}$ 与 $\widetilde{\mathscr{X}}$ 的某个稠密子空间等距同构, 根据稠密的定义知存在 $\{x_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ 使得 $\widetilde{\varphi}(x_n) \to \xi \in \widetilde{\mathscr{X}} (n \to \infty)$. 现在因为 $\widetilde{\varphi}, \varphi'$ 都是等距同构, 故因为 $\{\widetilde{\varphi}(x_n)\}_{n\in\mathbb{N}}$ 是 $\widetilde{\mathscr{X}}$ 中的基本列, 有:

$$\widetilde{\rho}(\widetilde{\varphi}(x_n), \widetilde{\varphi}(x_m)) = \rho(x_n, x_m) = \rho'(\varphi'(x_n), \varphi'(x_m)) \to 0, \ n, m \to \infty.$$

这说明 $\{\varphi'(x_n)\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}'$ 中的基本列, 由 $(\mathscr{X}', \rho')$ 完备知存在 $x' \in \mathscr{X}'$ 使得 $\varphi'(x_n) \to x' \in \mathscr{X}'(n \to \infty)$. 现在定义

$$\varphi : \widetilde{\mathscr{X}} \to \mathscr{X}', \ \xi \mapsto x'.$$

下面证明 φ 首先是良定义的, 此即证明 φ 的定义与 $\{x_n\}_{n\in\mathbb{N}}$ 的选取无关. 另设 $\{\overline{x}_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ 满足 $\widetilde{\overline{x}_n} \to \xi \in \widetilde{\mathscr{X}}(n \to \infty)$, 则由 $\widetilde{\varphi}$ 的等距性与三角不等式知

$$0 \le \rho(x_n, \overline{x}_n) = \widetilde{\rho}(\widetilde{\varphi}(x_n), \widetilde{\varphi}(\overline{x}_n)) \le \widetilde{\rho}(\widetilde{\varphi}(x_n), \xi) + \widetilde{\rho}(\xi, \widetilde{\varphi}(\overline{x}_n)) \to 0, \ n \to \infty,$$

这说明

$$\rho(x_n, \overline{x}_n) \to 0(n \to \infty) \Rightarrow \rho'(\varphi'(x_n), \varphi'(\overline{x}_n)) \to 0(n \to \infty).$$

也即 $\varphi'(\overline{x}_n) \to x'(n \to \infty)$, 从而 φ 是良定义的.

再说明 φ 是等距同构, 为此先说明 φ 是满射. 任取 $y' \in \mathscr{X}'$, 因为 $\mathscr{X}$ 等距同构于 $\mathscr{X}'$ 的一个稠密子空间, 故存在 $\{y_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ 使得 $\varphi'(y_n) \to y' \in \mathscr{X}'(n \to \infty)$. 根据 φ' 的等距性知 $\{y_n\}_{n\in\mathbb{N}}$ 也是 $\mathscr{X}$ 中的基本列, 再由 $\widetilde{\varphi}$ 的等距性知 $\{\widetilde{\varphi}(y_n)\}_{n\in\mathbb{N}}$ 在 $\widetilde{\mathscr{X}}$ 中也是基本列. 因为 $\widetilde{\mathscr{X}}$ 本身完备, 故存在 $\eta \in \widetilde{\mathscr{X}}$ 使得 $\widetilde{\varphi}(y_n) \to \eta(n \to \infty)$, 根据 φ 的定义知 $\varphi(\eta) = y'$, 故 φ 是满射.

最后证明 φ 是等距的. 任取 $\xi, \eta \in \widetilde{\mathscr{X}}$, 因为 $\mathscr{X}$ 等距同构于 $\widetilde{\mathscr{X}}$ 的一个稠密子空间, 故存在 $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \subset \mathscr{X}$ 使得

$$\widetilde{\rho}(\xi, \widetilde{\varphi}(x_n)) \to 0, \widetilde{\rho}(\eta, \widetilde{\varphi}(y_n)) \to 0 \ n \to \infty,$$

记 $x' := \varphi(\xi), y' := \varphi(\eta)$, 则由 φ 的定义知

$$\lim_{m\to\infty} \rho'(\varphi'(x_n), x') = 0 = \lim_{n\to\infty} \rho'(\varphi'(y_n), y'). \tag{2.6}$$

下面证明 $\widetilde{\rho}(\xi, \eta) = \rho'(x', y')$, 事实上

$$\begin{aligned} |\widetilde{\rho}(\xi, \eta) - \widetilde{\rho}(\widetilde{\varphi}(x_n), \widetilde{\varphi}(y_n))| &\le |\widetilde{\rho}(\xi, \eta) - \widetilde{\rho}(\xi, \widetilde{\varphi}(y_n))| + |\widetilde{\rho}(\xi, \widetilde{\varphi}(y_n)) - \widetilde{\rho}(\widetilde{\varphi}(x_n), \widetilde{\varphi}(y_n))| \\ &\overset{\text{(i)}}{\le} \widetilde{\rho}(\eta, \widetilde{\varphi}(y_n)) + \widetilde{\rho}(\xi, \widetilde{\varphi}(x_n)) \to 0, \ n \to \infty, \end{aligned} \tag{2.7}$$

其中 (i) 是三角不等式的直接推论:

推论 2.2.1

对任意的 $x, y, z \in (\mathscr{X}, \rho)$ 有:

$$|\rho(x, z) - \rho(y, z)| \le \rho(x, y)$$

这从极限的定义上说明:

$$\widetilde{\rho}(\xi, \eta) = \lim_{n\to\infty} \widetilde{\rho}(\widetilde{\varphi}(x_n), \widetilde{\varphi}(y_n)). \tag{2.8}$$

根据 $\widetilde{\varphi}$ 的等距性:

$$\widetilde{\rho}(\widetilde{\varphi}(x_n),\widetilde{\varphi}(y_n))=\rho(x_n,y_n),\ \forall n\in\mathbb{N}, \tag{2.9}$$

再根据 φ' 的等距性:

$$\rho(x_n,y_n)=\rho'(\varphi'(x_n),\varphi'(y_n)),\ \forall n\in\mathbb{N}, \tag{2.10}$$

类似(2.7)式并利用(2.6)式可证

$$\lim_{n\to\infty}\rho'(\varphi'(x_n),\varphi'(y_n))=\rho'(x',y')=\rho'(\varphi(\xi),\varphi(\eta)). \tag{2.11}$$

联立(2.8)-(2.10)式即知 $\widetilde{\rho}(\xi,\eta)=\rho'(\varphi(\xi),\varphi(\eta))$, 从而 φ 是 $\widetilde{\mathscr{X}}$ 到 $\mathscr{X}'$ 的等距同构. □

课堂笔记 (⋆)

定义(2.2.4)中的完备化空间在等距同构的意义下不一定唯一, 下为反例:

设 $\mathscr{X}_1:=\mathbb{N}$, $\mathscr{X}_2=\mathbb{N}\backslash\{2\}$, 赋度量 $\rho(x,y)=|x-y|$, 则 $\mathscr{X}_1,\mathscr{X}_2$ 均完备[a]. $\mathscr{X}$ 为其自身的一个完备化空间, 且显见 $\mathscr{X}_2$ 可以等距嵌入到 $\mathscr{X}_1$, 下面说明 $\mathscr{X}_2$ 是 $\mathscr{X}_1$ 的一个完备化空间, 取

$$\varphi:\mathscr{X}_1\to\mathscr{X}_2,\ n\mapsto n+2,$$

显见 φ 是 $\mathscr{X}_1$ 到 $\mathscr{X}_2$ 的等距映射, 故 $\mathscr{X}_1$ 可以等距嵌入到 $\mathscr{X}_2$ 中, 从而 $\mathscr{X}_2$ 也是 $\mathscr{X}_1$ 的一个完备化空间. 但 $\mathscr{X}_1$ 与 $\mathscr{X}_2$ 并不等距同构, 这是因为若设 ψ 是 $\mathscr{X}_1$ 到 $\mathscr{X}_2$ 的等距同构映射, 则 $\psi^{-1}(1)$ 与 $\psi^{-1}(1)+1$ 在 $\mathscr{X}_1$ 中的距离为 1, 但在 $\mathscr{X}_2$ 中不存在到 1 距离为 1 的点, 也即 $\psi^{-1}(1)+1\in\mathscr{X}_1$, 但 $\psi(\psi^{-1}(1)+1)\notin\mathscr{X}_2$, 矛盾! 故 $\mathscr{X}_1$ 与 $\mathscr{X}_2$ 不等距同构.

[a]基本列为从某项开始往后均为同一常数的数列.

注

- [3]关于上述反例仍有疑惑: 定义(2.2.4)中的"包含"是在什么意义下的? 如果 $A\subset B$ 指的是

$$x\in A\Rightarrow x\in B$$

  那么显见上述反例中的 $\mathscr{X}_2$ 在定义(2.2.4)下并非 $\mathscr{X}_1$ 的完备化空间. 而依照嵌入的定义(2.2.2)中所说的等距同构意义下的包含 $\subset$, 已经说明了 $\mathscr{X}_1\subset\mathscr{X}_2,\mathscr{X}_2\subset\mathscr{X}_1$, 而 $\mathscr{X}_1,\mathscr{X}_2$ 本身并不等距同构, 这说明至少在现在这种情况下, 关系 $\subset$ 作为偏序, 其所对应的等价关系 $=$ 并非等距同构. 回到定义(2.2.4), 要讨论 $\mathscr{X}_1$ 的完备化空间, 必须要说明满足"包含 $(\mathscr{X}_1,\rho)$ 的最小的完备度量空间"这一条件的度量空间的存在性, 最好也能说明这个度量空间 (暂时设为 $\mathscr{A}$) 是包含于 $\mathbb{N}$ 的. 首先, 为了保证嵌入映射的存在性, $\mathscr{A}$ 不能是有限集. 其次, 为了保证等距嵌入的等距性, $\mathscr{A}$ 中必须至少有可数个元素之间的距离 ρ 为 1, 进一步必须至少有一条可数长度的连续数串[4], 这是因为对 $\mathscr{X}_1$ 中任意长度的数串, 根据等距嵌入的等距性, 它都应该对应 $\mathscr{A}$ 中相等长度的数串. 这说明 $\mathscr{A}$ 中存在连续数串 $\{a_1,a_2,\cdots\}$, 取等距嵌入 $\varphi:n\mapsto n+a_1$ 即可将 $\mathscr{X}_1$ 等距嵌入 $\mathscr{A}$ 内, 又显见 $\mathscr{A}$ 可以等距嵌入 φ, 这说明如果一个度量空间满足前述条件且包含于 $\mathbb{N}$, 那么它与 $\mathscr{X}_1$ 之间的包含关系 $\subset$ 就至少不是全序, 因而至少没法用 Zorn 引理断言这样的度量空间中"最小者"的存在性, 得到疑问: $\mathscr{X}_2$ 为什么是 $\mathscr{X}_1$ 在定义(2.2.4)下的完备化空间?
- [5]上述反例确疑有问题, 它所说明的是"最小"一词定义模糊, 即可能并不存在所谓"最小的"完备化空间. 一般来说, 谈论完备化空间时就按下面给出的新定义即可.

在诸多泛函分析教材中, 命题2.2.1才是完备化空间的定义:

[3]2023.9.6 记.

[4]这里是自己定义的一个概念: 如果 $\mathbb{N}\ni a_1,a_2,\cdots$ 是互不相同的元素, $a_1<a_2<\cdots$, 满足 $1=|a_1-a_2|=|a_2-a_3|=\cdots$, 就称集合 $\{a_1,a_2,\cdots\}$ 是一条可数长度的连续数串. 而若 $\mathbb{N}\ni a_1,a_2,\cdots,a_n$ 是互不相同的元素, $a_1<\cdots<a_n$, 满足 $1=|a_1-a_2|=\cdots=|a_{n-1}-a_n|$, 则称集合 $\{a_1,\cdots,a_n\}$ 为长度为 n 的连续数串, n 称为该数串的长度.

[5]2023.9.6 答疑后记.

定理 2.2.1 (度量空间的完备化[PGC])

(i) 设 (X,d) 是度量空间, 则存在完备度量空间 $(\widetilde{X},\widetilde{d})$ 与等距同构 $\sigma: X\to\widetilde{X}$, 满足 $\sigma(X)$ 在 $\widetilde{X}$ 中稠密.

(ii) 若 X 可分, 则 $\widetilde{X}$ 可分.

(iii) 若 $(\widehat{X},\widehat{d})$ 满足: 存在从 X 到 $\widehat{X}$ 的某稠密子集的等距同构, 则存在从 $(\widetilde{X},\widetilde{d})$ 到 $(\widehat{X},\widehat{d})$ 上的等距同构.

定义 2.2.5 (完备化空间[PGC])

满足定理2.2.1的空间 $(\widetilde{X},\widetilde{d})$ 称为度量空间 (X,d) 的完备化空间. 注意从条件 (iii) 可以推知这样的空间在双射同构的意义下唯一.

定理 2.2.2

每一个度量空间都有一个 (定义(2.2.5)意义下的) 完备化空间.

注 ★ 首先对可能的证明思路进行分析. 度量空间 $\mathscr{X}$ 不完备, 根本原因就在于其基本列不一定收敛, 也就是其元素不足以囊括其基本列的极限. 为了克服这个问题, 可以把 $\mathscr{X}$ 中的基本列依照某种等价关系进行适当的分类, 每个等价类看做一个新元素, 这些元素的全体构成一个新集合 $\mathscr{X}_1$. 在这个集合 $\mathscr{X}_1$ 中适当引进距离 ρ_1, 使得对于 $\mathscr{X}$ 中关于 ρ 的基本列 $\{x_n\}_{n\in\mathbb{N}}$, $\mathscr{X}_1$ 中的序列 $\{[\{x_n\}_{m\in\mathbb{N}}]\}_{n\in\mathbb{N}}$ 也是关于 ρ_1 的基本类, 且其收敛到 $\mathscr{X}_1$ 中的元素 $[\{x_n\}_{n\in\mathbb{N}}]$. 又因为可以证明映射

$$T:\mathscr{X}\to\mathscr{X}_1,\ x\mapsto[\{x\}_{n\in\mathbb{N}}]$$

将 $\mathscr{X}$ 等距同构于 $\mathscr{X}_1$ 的稠子集 $\mathscr{X}':=T(\mathscr{X})$, 故对 $\mathscr{X}_1$ 中的任意基本列 $\{\xi^{(n)}\}_{n\in\mathbb{N}}$, 可以证明存在 $T(\mathscr{X})$ 中的基本列 $\{\overline{\xi}^{(n)}\}_{n\in\mathbb{N}}\subset\mathscr{X}_1$ 使得 $\{\overline{\xi}^{(n)}\}_{n\in\mathbb{N}}$ 与 $\{\xi^{(n)}\}_{n\in\mathbb{N}}$ 在 $\mathscr{X}_1$ 中有相同的极限, 同时可证 $\{\overline{\xi}^{(n)}\}_{n\in\mathbb{N}}$ 对应于 $\mathscr{X}$ 中的一个基本列 $\{x_n\}_{n\in\mathbb{N}}$, 进而存在 $\xi\in\mathscr{X}_1$ 使得 $\{x_n\}_{n\in\mathbb{N}}\in\xi$, 可以证明 ξ 正是 $\{\xi^{(n)}\}_{n\in\mathbb{N}}$ 的极限, 从而 $\mathscr{X}_1$ 完备.

证明 设 $(\mathscr{X},\rho)$ 是一个度量空间, 下面分三步证明它有一个完备化空间.

(1) 将 $\mathscr{X}$ 中的基本列分类. 凡是满足

$$\lim_{n\to\infty}\rho(x_n,y_n)=0$$

的两个基本列 $\{x_n\},\{y_n\}$ 称为等价的. 彼此等价的基本列归于同一类且只归于一类, 称为等价类. 我们把一个等价类看成是一个元素, 并用 $\mathscr{X}_1$ 表示一切这种元素 (等价类) 组成的集合. 在 $\mathscr{X}_1$ 上定义距离: $\forall\xi,\eta\in\mathscr{X}_1$, 任取 $\{x_n\}\in\xi,\{y_n\}\in\eta$, 令

$$\rho_1(\xi,\eta)=\lim_{n\to\infty}\rho(x_n,y_n),$$

根据前述等价类的定义容易验证上式右端极限确实存在, 且与 $\{x_n\},\{y_n\}$ 的选取无关. 为了验证 ρ_1 确实是距离, 注意到正则性与对称性显然, 而三角不等式可由

$$\rho(x_n,y_n)\le\rho(x_n,z_n)+\rho(z_n,y_n)$$

两边取极限得到, 其中 $\{x_n\},\{y_n\},\{z_n\}$ 是分别属于等价类 ξ,η,ζ 的基本列. 故 $(\mathscr{X}_1,\rho_1)$ 是一个度量空间.

(2) $\forall x\in\mathscr{X}$, 用 $\xi_x\in\mathscr{X}_1$ 表示包含序列 $(x,x,\cdots,x,\cdots)$ 的等价类, 这样的 ξ_x 的全体记作 $\mathscr{X}'$. 显见 $\mathscr{X}'\subset\mathscr{X}_1$, 且映射 $T: x\mapsto\xi_x$ 作为 $(\mathscr{X},\rho)\to(\mathscr{X}',\rho_1)$ 的映射满足等距同构的定义, 故 $(\mathscr{X},\rho)$ 和 $(\mathscr{X}',\rho_1)$ 等距同构, 即 $(\mathscr{X},\rho)\subset(\mathscr{X}_1,\rho_1)$, 进一步可以验证 $\mathscr{X}$ 在 $\mathscr{X}_1$ 中稠密.

(3) 证明 $(\mathscr{X}_1,\rho_1)$ 是完备的, 也即证明 $(\mathscr{X}_1,\rho_1)$ 中的基本列必收敛. 设 $\{\xi^{(n)}\}$ 是 $\mathscr{X}_1$ 中的基本列. 要证 $\exists\xi\in\mathscr{X}_1$, 使得

$$\rho_1(\xi^{(n)},\xi)\to0,\ n\to\infty.$$

首先证明特殊情形, 假定 $\{\xi^{(n)}\}\subset\mathscr{X}'$, 则令 $x_n=T^{-1}(\xi^{(n)})$, 知 $\{x_n\}$ 是 $\mathscr{X}$ 中的基本列. 设 $\{x_n\}\in\xi$, 则知 $\xi^{(n)}\to\xi$.

再证一般情形, 由于 $\mathscr{X}'$ 在 $\mathscr{X}_1$ 中稠密, 知

$$\forall \xi^{(n)} \in \mathscr{X}_1 \exists \overline{\xi}^{(n)} \in \mathscr{X}' \left(\rho_1(\overline{\xi}^{(n)}, \xi^{(n)}) < \frac{1}{n} \right).$$

进而根据特殊情形可设 $\overline{\xi}^{(n)} \to \xi \in \mathscr{X}_1$, 得到 $\xi^{(n)} \to \xi$.

最后, 综合 (1)-(3) 并结合命题2.2.1即得命题. □

例 2.13 $P[a,b]$($[a,b]$ 上的多项式全体) 按一致收敛度量

$$\rho(x,y) = \max_{a \le t \le b} |x(t) - y(t)|$$

的完备化空间是 $C[a,b]$.

例 2.14 $C[a,b]$ 按照 L^1 度量:

$$\rho_1(x,y) := \int_a^b |x(t) - y(t)| dt$$

完备化, 完备化空间是 $L^1[a,b]$.

2.2.2 一些例子

例 2.15(某个完备度量空间中存在渐缩非空闭球列 $\{B_n\}$, 使得 $\bigcap_{n=1}^{\infty} B_n = \varnothing$) 设 X 是一切正整数所成的集合, 对 $m,n \in X$, 令

$$d(m,n) = \begin{cases} 1 + \dfrac{1}{m+n}, & m \neq n \\ 0, & m = n \end{cases}$$

易证 d 是 X 上的距离. 任取 X 中的基本列 $\{x_n\}$, 根据距离 d 的定义知存在正整数 n_0, 当 $n > n_0$ 时全体 x_n 均相同, 这说明 $\{x_n\}$ 为一收敛点列, 进而 X 是完备度量空间.

设 $B_n = \{m : d(n,m) \le 1 + \frac{1}{2n}\} = \{n, n+1, \cdots\}$, 知各个 B_n 都是非空闭球, 且 $\{B_n\}$ 为渐缩列, 但 $\bigcap_{n=1}^{\infty} B_n = \varnothing$.

注 考虑下述命题:

命题 2.2.2

设 X 是不空的完备度量空间, 若 $B_n = \{x : d(x_n, x) \le \varepsilon_n\}$ 是 X 中渐缩的闭球列, 且 $\varepsilon_n \to 0$, 则必有唯一的点 $x \in \bigcap_{n=1}^{\infty} B_n$.

上述反例说明命题中 $\varepsilon_n \to 0$ 的条件不能去掉.

例 2.16(任何子集都既开又闭的完备度量空间) 设 X 为非空集合, 在 X 上取距离:

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

任取基本列 $\{x_n\} \subset X$, 根据基本列的定义知当 n 充分大后 x_n 均相同, 从而 $\{x_n\}$ 是 X 中的收敛点列, 亦即 X 是完备度量空间.

现任取 $x \in X$, 因为以 x 为球心, $\frac{1}{2}$ 为半径的开球只含有点 x, 故 X 中每一单元素集均为开集, 又因为开集的任意并依旧是开集, 故 X 的任一子集均为开集, 取补集即知任一子集也为闭集.

例 2.17(任何子集都既开又闭的不完备度量空间) 设 X 为全体正整数组成的集合, 令 $d(m,n) = \frac{|m-n|}{mn}$, 可以验证 (X,d) 为度量空间, 又因为 $\{n\}_{n=1}^{\infty}$ 是 (X,d) 中的一个基本列, 且其并不收敛, 故 (X,d) 不完备.

下面证明 X 中任一单元素集 $\{n\}$ 都是开集. 若 $m > n$, 则

$$f(m) = d(m,n) = \frac{m-n}{mn}$$

是关于 m 的递增函数, 因而

$$d(m,n) \ge \frac{1}{n(n+1)}, \ m > n.$$

若 $0<m<n$, 则

$$f(m)=d(m,n)=\frac{n-m}{mn}$$

是关于 m 的递减函数, 这说明

$$d(m,n)\geq\frac{1}{n(n-1)}>\frac{1}{n(n+1)},\ n>m>0,$$

进而对任意正整数 $m(m\neq n)$, 总有 $d(m,n)\geq\frac{1}{n(n+1)}$, 取 $r=\frac{1}{n(n+1)}$, 则以 n 为球心, r 为半径的开球正是单点集 $\{n\}$, 故 $\{n\}$ 是开集, 从而 X 的全体子集均为开集, 进而全体子集均为闭集.

注 上述情况在欧氏空间中不可能出现. 对欧氏空间 (乃至任意的连通拓扑空间), 其中既开又闭的集合只有空间本身与 $\varnothing$[6].

定义 2.2.6 (离散度量空间)

如果度量空间中的单点集均为开集, 则称该度量空间为离散度量空间. 上两例表明离散度量空间与完备性没有必然联系.

2.2.3 习题

练习 2.8(空间 S)★ 令 S 为一切实 (或复) 数列

$$x=(\xi_1,\xi_2,\cdots,\xi_n,\cdots)$$

组成的集合, 在 S 中定义距离为

$$\rho(x,y)=\sum_{k=1}^{\infty}\frac{1}{2^k}\cdot\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|},$$

其中 $x=(\xi_1,\xi_2,\cdots,\xi_k,\cdots),y=(\eta_1,\eta_2,\cdots,\eta_k,\cdots)$. 求证: S 为一个完备的度量空间.

证明 首先验证 $\rho(x,y)$ 是距离, 出于 $0\leq\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|}\leq 1$ 知 $\rho(x,y)$ 良定义. 显见 $\rho(x,y)\geq 0$, 且 $\rho(x,y)=0\Leftrightarrow x=y$, $\rho(x,y)=\rho(y,x)$. 最后考虑 $x=(\xi_1,\cdots,\xi_k,\cdots),y=(\eta_1,\cdots,\eta_k,\cdots),z=(\zeta_1,\cdots,\zeta_k,\cdots)$, 有

$$\begin{aligned}\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|}+\frac{|\eta_k-\zeta_k|}{1+|\eta_k-\zeta_k|}&=\frac{|\xi_k-\eta_k|+|\eta_k-\zeta_k|+2|\xi_k-\eta_k|\cdot|\eta_k-\zeta_k|}{1+|\xi_k-\eta_k|+|\eta_k-\zeta_k|+|\xi_k-\eta_k|\cdot|\eta_k-\zeta_k|}\\&\geq\frac{|\xi_k-\eta_k|+|\eta_k-\zeta_k|+|\xi_k-\eta_k|\cdot|\eta_k-\zeta_k|}{1+|\xi_k-\eta_k|+|\eta_k-\zeta_k|+|\xi_k-\eta_k|\cdot|\eta_k-\zeta_k|}\geq\frac{|\xi_k-\zeta_k|}{1+|\xi_k-\zeta_k|},\end{aligned}$$

其中最后一步是出于 $|\xi_k-\eta_k|+|\eta_k-\zeta_k|+|\xi_k-\eta_k|\cdot|\eta_k-\zeta_k|\geq|\xi_k-\zeta_k|$ 与 $\frac{x}{1+x}$ 的增性. 综上 $\rho(x,y)$ 确为距离, 进而 S 是度量空间.

再考虑完备性, 考虑 S 中的基本列 $\{x_n\}=(\xi_n^1,\cdots,\xi_n^k,\cdots)$, 知:

$$\rho(x_n,x_m)=\sum_{k=1}^{\infty}\frac{1}{2^k}\cdot\frac{|\xi_k^n-\xi_k^m|}{1+|\xi_k^n-\xi_k^m|}\to 0,\ n,m\to\infty.$$

注意若存在 $k_0\in\mathbb{N}$ 使得 $|\xi_n^{k_0}-\xi_m^{k_0}|\nrightarrow 0$, 则必有 $\frac{1}{2^{k_0}}\cdot\frac{|\xi_n^{k_0}-\xi_m^{k_0}|}{1+|\xi_n^{k_0}-\xi_m^{k_0}|}\nrightarrow 0$, 进而导出矛盾. 这说明

$$\forall k\in\mathbb{N}(|\xi_n^k-\xi_m^k|\to 0),\ n,m\to\infty.$$

因而每个 $\{\xi_n^k\}$ 均为 $\mathbb{R}$(或 $\mathbb{C}$) 中的基本列, 由 $\mathbb{R}$(或 $\mathbb{C}$) 的完备性知 $\{\xi_n^k\}$ 有极限 $\{\xi^k\}$, 得到极限序列

$$x=(\xi^1,\xi^2,\cdots,\xi^k,\cdots)\in S.$$

下面说明序列 $x=(\xi^1,\cdots,\xi^k,\cdots)\in S$ 确实是 $\{x_n\}$ 的极限. 注意 $\sum_{k=1}^{\infty}\frac{1}{2^k}=1<\infty$, 故

$$\forall\varepsilon>0\exists N>0\forall n>N\left(\sum_{k=n+1}^{\infty}\frac{1}{2^k}<\varepsilon\right),$$

而对每个 $k\leq N$, 都存在 $N_k>0$, 使得

$$n>N_k\Rightarrow|\xi_n^k-\xi_n|<\varepsilon.$$

[6] 这个断言的理解参见 PDE 部分 Laplace 方程调和函数强极值原理证明处的笔记.

故取 $N' = \max\{N, N_1, \cdots, N_N\}$, 有:

$$\rho(x_n, x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{|\xi_n^k - \xi^k|}{1 + |\xi_n^k - \xi^k|} = \sum_{k=1}^{N'} \frac{1}{2^k} \cdot \frac{|\xi_n^k - \xi^k|}{1 + |\xi_n^k - \xi^k|} + \sum_{k=N'+1}^{\infty} \frac{1}{2^k} \cdot \frac{|\xi_n^k - \xi^k|}{1 + |\xi_n^k - \xi^k|}$$
$$< \sum_{k=1}^{N'} \frac{1}{2^k} \cdot \frac{\varepsilon}{1+\varepsilon} + \sum_{k=N'+1}^{\infty} \frac{1}{2^k} < \sum_{k=1}^{N'} \frac{\varepsilon}{2^k} + \varepsilon < 2\varepsilon,$$

由 ε 的任意性即得命题.

练习 2.9★ 在一个度量空间 $(\mathscr{X}, \rho)$ 上, 求证: 基本列是收敛列, 当且仅当其中存在一串收敛子列.

证明 当基本列是收敛列, 命题显见, 进而现在考虑基本列 $\{x_n\}$ 中存在一串收敛子列 $\{x_{n_k}\}$, 记 $x_{n_k} \to x \in \mathscr{X}, k \to \infty$, 根据定义有

$$\forall \varepsilon > 0 \exists N > 0 \forall k > N(\rho(x_{n_k}, x) < \varepsilon),$$

根据基本列的定义有

$$\exists N' > 0 \forall n, m > N'(\rho(x_n, x_m) < \varepsilon),$$

故取 $N'' = \max\{N, N'\}$, 则当 $n, k > N''$ 时:

$$\rho(x_n, x) \le \rho(x_n, x_{n_k}) + \rho(x_{n_k}, x) < \varepsilon + \varepsilon = 2\varepsilon.$$

故 $x_n \to x, n \to \infty$, 也即 $\{x_n\}$ 在 $(\mathscr{X}, \rho)$ 中是收敛列, 命题得证.

练习 2.10★♡ 设 F 是只有有限项不为 0 的实数列全体, 在 F 上引进距离

$$\rho(x, y) = \sup_{k \ge 1} |\xi_k - \eta_k|$$

其中 $x = \{\xi_k\} \in F, y = \{\eta_k\} \in F$, 求证: (F, ρ) 不完备, 并指出它的完备化空间.

证明 考虑实数列:

$$x_n = \left(1, \frac{1}{2}, \cdots, \frac{1}{2^n}, 0, \cdots, 0\right), \ n \in \mathbb{N},$$

显见对固定的 n, x_n 都是有限项不为零的实数列, 进而 $x_n \in F \Rightarrow \{x_n\}_{n\in\mathbb{N}} \subset F$. 下面证明 $\{x_n\}$ 是 (F, ρ) 中的基本列: 对任意的 $1 > \varepsilon > 0$, 取 $N = \log_2(\frac{1}{\varepsilon}) - 1$, 则对任意的 $m > n > N$, 都有:

$$\rho(x_m, x_n) = \frac{1}{2^{n+1}} < \frac{1}{2^{N+1}} = \varepsilon,$$

这说明 $\{x_n\}_{n\in\mathbb{N}}$ 是 (F, ρ) 中的基本列, 但

$$\lim_{n\to\infty} x_n = (1, \frac{1}{2}, \cdots, \frac{1}{2^n}, \cdots) \notin F.$$

故 (F, ρ) 不完备. 现设 F' 为满足

$$|x_k| \to 0 (k \to \infty) \text{ 且 } \sup_{k\ge1} |x_k| < \infty$$

的实数列 $x = (x_1, \cdots, x_n, \cdots)$ 的全体, 下面验证 ρ 确为 F' 上的距离, 任取 $x = (x_k)_{k\ge1}, y = (y_k)_{k\ge1}, z = (z_k)_{k\ge1} \in F'$:

- 正定性: $|x_k - y_k| \ge 0 (\forall k \in \mathbb{N}) \Rightarrow \rho(x, y) \ge 0$, 同时 $\rho(x, y) = 0 \Rightarrow \sup_{k\ge1} |x_k - y_k| = 0 \Rightarrow 0 \le |x_k - y_k| \le 0 (\forall k \in \mathbb{N})$, 故 $\forall k \in \mathbb{N}(x_k = y_k)$, 亦即 $x = y$.
- 交换律: $\rho(x, y) = \sup_{k\ge1} |x_k - y_k| = \sup_{k\ge1} |y_k - x_k| = \rho(y, x)$;
- 三角不等式: 根据实数的三角不等式有: $|x_k - y_k| \le |x_k - z_k| + |z_k - y_k| (\forall k \in \mathbb{N})$, 故

$$\sup_{k\ge1} |x_k - y_k| \le \sup_{k\ge1} (|x_k - z_k| + |z_k - y_k|) \le \sup_{k\ge1} |x_k - z_k| + \sup_{k\ge1} |z_k - y_k| \Rightarrow \rho(x, y) \le \rho(x, z) + \rho(z, y).$$

综上, (F', ρ) 为度量空间, 下面简记为 F'.

再证明 F' 的完备性, 任取 $x_m = (x_m^k)_{k\ge1} \in F'$ 是基本列, 根据基本列的定义知:

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N(\rho(x_m^k, x_n^k) = \sup_{k\ge1} |x_m^k - x_n^k| < \varepsilon),$$

根据上确界的定义知:

$$\forall k \geq 1(|x_m^k - x_n^k| \leq \sup_{k\geq 1}|x_m^k - x_n^k| < \varepsilon),$$

这说明对每个固定的 k 而言, $\{x_m^k\}$ 是 $\mathbb{R}$ 中的基本列, 进而根据 $\mathbb{R}$ 的完备性知其有极限 x^k, 记:

$$x = (x^1, x^2, \cdots, x^k, \cdots).$$

下面说明 $x \in F'$. 对任意的 $k \geq 1$, 前述 $\mathbb{R}$ 中的极限定义为:

$$\exists N_k > 0 \forall n > N_k(|x_n^k - x^k| < \varepsilon). \tag{2.12}$$

故若 $x \notin F'$, 根据 F' 的定义知

$$\forall M > 0 \exists k_M \geq 1(|x^{k_M}| > M) \text{ 或 } \exists \varepsilon_0 > 0 \forall K > 0 \exists k_K > K(|x^{k_K}| > \varepsilon_0).$$

对第一种情况, 取 $n > N_k$ 有

$$\varepsilon > |x_n^{k_M} - x^{k_M}| \geq |x^{k_M}| - |x_n^{k_M}| > M - |x_n^{k_M}| \Rightarrow |x_n^{k_M}| > M - \varepsilon,$$

不妨设 $\varepsilon < \frac{M}{2}$, 这便说明当 n 足够大时, 对任意 $M > 0$, 总能取到 $k_M \geq 1$, 使得 $|x_n^{k_M}| > \frac{M}{2}$, 根据 M 的任意性知 $x_n \notin l^\infty$, 矛盾! 而对第二种情况, 取 $\varepsilon = \frac{\varepsilon_0}{2}$ 并取 $n > N_k$, 有:

$$\varepsilon > |x_n^{k_K} - x^{k_K}| \geq |x^{k_K}| - |x_n^{k_K}| > \varepsilon_0 - |x_n^{k_K}| \Rightarrow |x_n^{k_K}| > \frac{\varepsilon_0}{2}.$$

这说明当 n 足够大时, 总存在足够大的 k_K 使得 $|x_n^{k_K}| > \frac{\varepsilon_0}{2}$, 但这与 x_n 的定义不符, 矛盾! 综上知 $x \in F'$.

下面说明 x 确为 $\{x_n\}$ 在 F' 中的极限, 用反证法. 如若存在 ε_0, 使得对任意的 $N > 0$, 总存在 $n_N > N$, 使得

$$\rho(x, x_{n_N}) = \sup_{k\geq 1}|x^k - x_{n_N}^k| \geq \varepsilon_0,$$

根据上确界的定义, 这表明对任意的 $\varepsilon_1 > 0$, 总存在 $k_{\varepsilon_1} \geq 1$, 使得

$$|x^{k_{\varepsilon_1}} - x_{n_N}^{k_{\varepsilon_1}}| \geq \varepsilon_0 - \varepsilon_1.$$

取 $\varepsilon_1 = \frac{\varepsilon_0}{2}$, 这便与(2.61)式矛盾, 故 $\lim_{n\to\infty}\rho(x, x_n) = 0$, 确有 $x_n \to x \in F'(n \to \infty)$.

最后说明 F 与 F' 的某个稠密子集等距同构. 取恒等映射:

$$\mathrm{id} : F \to F', \ (x_1, \cdots, x_n, 0, \cdots) \mapsto (x_1, \cdots, x_n, 0, \cdots),$$

显见 id 是等距映射, 下面再说明 F 在 F' 中稠密. 任取 $x = (x^1, \cdots, x^k, \cdots) \in F'$, 取:

$$x_n = (x^1, \cdots, x^n, 0, \cdots),$$

根据 F' 的定义, 对 x 有:

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N(|x^n| < \varepsilon \Rightarrow \sup_{n\geq N+1}|x^n| \leq \varepsilon).$$

而易知

$$\rho(x_n, x) = \sup_{k\geq n+1}|x^k|,$$

取 $n > N$ 即知 $\rho(x_n, x) \leq \varepsilon$, 进而由 ε 的任意性即得 $x_n \to x \in F'(n \to \infty)$, 故 F 在完备度量空间 F' 中稠密, 进而 F' 为 F 的完备化空间. □

✍ **练习 2.11**★♡ 求证: $[0,1]$ 上的多项式全体按距离

$$\rho(p, q) = \int_0^1 |p(x) - q(x)|dx \ (p, q\text{是多项式})$$

是不完备的, 并指出它的完备化空间.

证明 首先简单说明距离 d 确为距离函数. d 的正定性与对称性由绝对值的性质立得, 而 d 满足三角不等式是因为绝对值函数 $|\cdot|$ 满足三角不等式.

再说明题设距离空间不完备. 考虑多项式列 $\{p_n\}_{n=1}^\infty$, 其定义为

$$p_n(x) = \sum_{k=1}^n \frac{x^k}{k!},$$

下面说明 $\{p_n\}$ 是题设空间在距离 d 下的基本列, 设 $m > n$, 则

$$d(p_m, p_n) = \int_0^1 \left|\sum_{k=n+1}^{m} \frac{x^k}{k!}\right| dx = \sum_{k=n+1}^{m} \int_0^1 \frac{x^k}{k!} dx = \sum_{k=n+1}^{m} \frac{1}{(k+1)!}. \tag{2.13}$$

容易证明当 $n \in \mathbb{N}$ 时 $(n+1)! > (n+1)n > n^2$, 故由比较审敛法知级数 $\sum_{n=1}^{\infty} \frac{1}{(n+1)!}$ 收敛, 从而由 Cauchy 收敛准则知

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N \left(\left| \sum_{k=n+1}^{m} \frac{1}{(n+1)!} \right| < \varepsilon \right). \tag{2.14}$$

将(2.56)式代入(2.57)的不等式中得到

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N (d(p_m, p_n) < \varepsilon),$$

故 $\{p_n\}$ 是题设空间在距离 d 下的基本列, 但极限 $\lim_{n\to\infty} p_n(x) = e^x$ 在 $[0,1]$ 上由 Dini 定理是一致成立的, 而函数 e^x 并非多项式, 这便说明题设空间在距离 d 下不完备.

由 Weierstrass 逼近定理知, [0,1] 上的连续函数总能由 [0,1] 上的多项式一致逼近, 这说明题设距离空间在 $(C[0,1], d)$ 中稠密, 又根据实变函数论的相关定理知: [0,1] 上的任意可积函数可由 $C[0,1]$ 中的函数按 L^1 距离逼近, 这说明 $(C[0,1], d)$ 在 $L^1[0,1]$ 中稠密. 最后来证明 $L^1[0,1]$ 的完备性[7]:

设 $\{f_n\}$ 是 L^1 中的一个基本列, 即

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N \left(\int_0^1 |f_m - f_n| dx < \varepsilon \right), \tag{2.15}$$

现取 $\varepsilon_k = \frac{1}{2^k} (k \in \mathbb{N})$, 则存在子列 $\{f_{n_k}\}$ 满足

$$\int_0^1 |f_{n_{k+1}} - f_{n_k}| dx < \frac{1}{2^k}. \tag{2.16}$$

现说明 $\lim_{k\to\infty} f_{n_k} \in L^1[0,1]$, 考虑级数

$$F(x) := |f_{n_1}(x)| + \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|,$$

根据逐项积分定理与(2.59)式, 因为

$$\sum_{k=1}^{\infty} \int_0^1 |f_{n_{k+1}} - f_{n_k}| dx \le \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 < \infty,$$

故

$$\begin{aligned} \int_0^1 F dx &= \int_0^1 |f_{n_1}| dx + \int_0^1 \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}| dx \\ &= \int_0^1 |f_{n_1}| dx + \sum_{k=1}^{\infty} \int_0^1 |f_{n_{k+1}} - f_{n_k}| dx < \int_0^1 |f_{n_1}| dx + 1 < \infty. \end{aligned}$$

其中最后一步是因为 $f_{n_1} \in L^1[0,1]$. 这个可积性一方面说明级数 $F(x)$ 在 $[0,1]$ 上几乎处处收敛, 另一方面通过设

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

可由 $|f| \le F$ 得到 $f \in L^1[0,1]$, 从而级数 $f(x)$ 同样在 [0,1] 上几乎处处收敛. 从定义中可见

$$f(x) = \lim_{k\to\infty} f_{n_k}(x).$$

同时对任意固定的 k, 显见 $|f - f_{n_k}| \le F$, 故由 Lebesgue 控制收敛定理知

$$\lim_{k\to\infty} \int_0^1 |f - f_{n_k}| dx = \int_0^1 |f - \lim_{k\to\infty} f_{n_k}| dx = 0.$$

这说明 f_{n_k} 在 L^1 距离 (也即距离 d) 的意义下收敛到 f. 沿用(2.58)式的记号并将上式按定义写开即

$$\exists K > 0 \forall k > K \left(\int_0^1 |f - f_{n_k}| dx < \varepsilon \right). \tag{2.17}$$

[7]这个完备性又叫做 Riesz-Fisher 定理[ST2].

最后,考虑说明 f_n 在 L^1 距离的意义下收敛到 f. 沿用(2.58),(2.60)式的记号,取 $n>N,k>K$ 满足 $n_k>N$, 则

$$\int_0^1 |f-f_n|dx \le \int_0^1 |f-f_{n_k}|dx + \int_0^1 |f_{n_k}-f_n|dx < 2\varepsilon,$$

由 ε 的任意性即得 f_n 在 L^1 距离的意义下收敛到 $f \in L^1[0,1]$.

综上, $L^1[0,1]$ 是完备空间, 且根据 [0,1] 上的多项式函数在 $L^1[0,1]$ 中的稠密性知 $L^1[0,1]$ 正是题设空间的完备化空间.

练习 2.12 在完备的度量空间 $(\mathscr{X},\rho)$ 中给定点列 $\{x_n\}$, 如果 $\forall\varepsilon>0$, 存在基本列 $\{y_n\}$, 使得

$$\rho(x_n,y_n)<\varepsilon\ (n\in\mathbb{N}),$$

求证: $\{x_n\}$ 收敛.

证明 只需说明 $\{x_n\}$ 是基本列, 注意 $\{y_n\}$ 是基本列, 根据定义知

$$\forall\varepsilon>0\exists N>0\forall n,m>N(\rho(y_m,y_n)<\varepsilon),$$

此时

$$\rho(x_m,x_n)=\rho(x_m,y_m)+\rho(y_m,y_n)+\rho(y_n,x_n)<3\varepsilon.$$

这说明 $\{x_n\}$ 也是 $(\mathscr{X},\rho)$ 中的基本列, 又因为 $(\mathscr{X},\rho)$ 完备, 故 $\{x_n\}$ 收敛.

2.3 列紧集

2.3.1 知识梳理

定义 2.3.1 (有界)

设 $(\mathscr{X},\rho)$ 是一个度量空间, A 是 $\mathscr{X}$ 的一个子集, A 称为是有界的, 如果 $\exists x_0\in\mathscr{X}$ 及 $r>0$, 使得 $A\subset B(x_0,r)$, 其中

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

♣

注意有穷维欧氏空间中, 有界无穷集必含有一个收敛子列, 但这个性质不能推广到任意度量空间.

例 2.18(有界序列不含收敛子列) 在 $C[0,1]$ 上, 考察点列

$$x_n(t)=\begin{cases}0, & t\ge\frac{1}{n},\\ 1-nt, & t<\frac{1}{n}\end{cases}\quad n=1,2,\cdots$$

显见 $\{x_n\}\subset B(\theta,1)$, 其中 θ 表示恒等于 0 的函数, 但 $\{x_n\}$ 不含有收敛子列.

定义 2.3.2 (列紧, 自列紧, 列紧空间)

设 $(\mathscr{X},\rho)$ 是一个度量空间, A 为其一子集. 称 A 是列紧的[a], 如果 A 中的任意点列在 $\mathscr{X}$ 中有一个收敛子列. 若这个子列还收敛到 A 中的点, 则称 A 是自列紧的. 如果空间 $\mathscr{X}$ 是列紧的, 那么称 $\mathscr{X}$ 为列紧空间.

[a] 列紧可以保证收敛点落在 $\mathscr{X}$ 中.

♣

注 列紧空间总是自列紧的.

命题 2.3.1

在 $\mathbb{R}^n$ 中任意有界集都是列紧集, 任意有界闭集都是自列紧集.

♠

补充命题 2.3.1 (⋆)

在 $\mathbb{R}^n$ 中, 列紧集 $\Leftrightarrow$ 有界集, 自列紧集 $\Leftrightarrow$ 有界闭集.

证明

根据命题2.3.1, $\mathbb{R}^n$ 中的有界集必为列紧集, 下面证明 $\mathbb{R}$ 中的列紧集必有界. 用反证法, 设 $A \subset \mathbb{R}^n$ 是列紧集且无界, 根据定义知存在子列 $\{x_n\}_{n\in\mathbb{N}} \subset A$ 使得 $x_n + 1 < x_{n+1}(\forall n \in \mathbb{N})$ 或 $x_n - 1 > x_{n+1}(\forall n \in \mathbb{N})$, 但这两种情形均说明 $\{x_n\}_{n\in\mathbb{N}}$ 不收敛, 矛盾! 故 A 必有界. $\mathbb{R}^n$ 中的情况与有界闭集情况是类似的. □

课堂笔记 (⋆)

在紧性的讨论中, 不可避免的一个例子正是实数连续性六大等价形式中的 Heine-Borel 引理. 下面先回顾实数集的 Dedekind 分割: $\subset \mathbb{R}$ 的一个切割是指将 $\mathbb{R}$ 分成子集 S, T, 它们满足:

(i) $S \neq \varnothing, T \neq \varnothing$;
(ii) $\mathbb{R} = S \cup T$;
(iii) $\forall x \in S \forall y \in T(x < y)$, 进而称 S 为左集, T 为右集.

实数集的连续性进而可表述为: 对于 $\mathbb{R}$ 的任何一个切割, 要么左集 S 具有最大实数, 要么右集 T 具有最小实数. 该命题有下述六种等价形式:

(i) 确界原理: 有上 (下) 界的非空数集有上 (下) 确界;
(ii) 单调收敛定理: 单调有界数列必收敛;
(iii) Cantor 闭区间套定理: " 非增" 的闭区间序列, 如果其长度趋于 0, 则该区间序列有唯一的公共点;
(iv) Heine-Borel 有限覆盖引理: 直线上有限闭区间的任一开覆盖必有有限子覆盖;
(v) Bolzano-Weierstrass 定理: 有界数列必有收敛子列;
(vi) Cauthy 收敛准则: 数列 $\{a_n\}_{n\in\mathbb{N}}$ 收敛当且仅当 $\{a_n\}_{n\in\mathbb{N}}$ 为基本列.

补充定义 2.3.1 (闭包 ⋆)

设 $(\mathscr{X}, \rho)$ 为度量空间, $E \subset \mathscr{X}$, 则称 $\overline{E}$ 为 E 的闭包, 如果 $x \in \overline{E} \Leftrightarrow \forall \delta > 0(B(x,\delta) \subset E \neq \varnothing)$.

补充命题 2.3.2 (闭包的等价定义 ⋆)

设 $(\mathscr{X}, \rho)$ 为度量空间, $E \subset \mathscr{X}$, 则 $x \in \overline{E} \Leftrightarrow \exists \{x_n\}_{n\in\mathbb{N}} \subset E(x_n \to x, n \to \infty)$.

证明

当 $x \in \overline{E}$, 根据定义知 $\forall n \in \mathbb{N}(B(x, \frac{1}{n}) \cap E \neq \varnothing)$, 取 $x_n \in B(x, \frac{1}{n}) \cap E$, 可得 $\{x_n\}_{n\in\mathbb{N}} \subset E$, 且 $0 \leq \rho(x, x_n) < \frac{1}{n} \to 0(n \to \infty)$, 亦即 $x_n \to x(n \to \infty)$.

当 $\exists \{x_n\}_{n\in\mathbb{N}} \subset E(x_n \to x, n \to \infty)$, 根据极限的定义知任取 $\delta > 0$, 存在 $N_0 \in \mathbb{N}$ 使得 $n \geq N_0$ 时有 $\rho(x_n, x) < \delta$, 故 $B(x,\delta) \cap E \supset \{x_n\}_{n=N_0}^{\infty} \neq \varnothing$, 从而 $x \in \overline{E}$. □

补充命题 2.3.3 (闭包的极小性 ⋆)

设 $(\mathscr{X}, \rho)$ 为度量空间, 则 $\overline{E}$ 为 $\mathscr{X}$ 中包含 E 的最小闭集.

证明

首先证明 $\overline{E}$ 是闭集. 任取 $\{x_n\}_{n\in\mathbb{N}} \subset \overline{E}$ 满足 $x_n \to x \in \mathscr{X}(n \to \infty)$, 欲证 $x \in \overline{E}$. 根据闭包的等价定义(2.3.2), 对任意的 $n \in \mathbb{N}$, 因为 $x_n \in \overline{E}$, 故 $B(x_n, \frac{1}{n}) \cap E \neq \varnothing$. 取[8]$\widetilde{x}_n \in B(x_n, \frac{1}{n}) \cap E$, 则对任意的 $n \in \mathbb{N}$ 有:

$$\rho(\widetilde{x}_n, x) \leq \rho(\widetilde{x}_n, x_n) + \rho(x_n, x) < \frac{1}{n} + \rho(x_n, x) \to 0, \ n \to \infty$$

[8]选 $\widetilde{x}_n$ 是为了确保序列在 E 中, 进而套用前述定理.

故 $E \ni \widetilde{x}_n \to x \in \mathscr{X}(n \to \infty)$, 从而根据命题2.3.2知 $x \in \overline{E}$, 因此 $\overline{E}$ 是闭集.

再设 $F \supset E$ 是闭集, 任取 $x \in \overline{E}$, 由命题2.3.2知 $\exists\{x_n\}_{n\in\mathbb{N}} \subset E(\rho(x_n, x) \to 0(n \to \infty))$. 因为 $\{x_n\}_{n\in\mathbb{N}} \subset E \subset F$, 且 $x_n \to x(n \to \infty)$, 故由 F 的闭性知 $x \in F$, 因而 $\overline{E} \subset F$, 这便说明 $\overline{E}$ 是 $\mathscr{X}$ 中包含 E 的最小闭集. □

补充命题 2.3.4 (⋆)

设 $(\mathscr{X}, \rho)$ 为度量空间, 则 F 为 E 的闭包当且仅当 F 为 $\mathscr{X}$ 中包含 E 的最小闭集.

证明

只需证明 $\mathscr{X}$ 中包含 E 的最小闭集正是 $\overline{E}$ 即可. 设 F 为 $\mathscr{X}$ 中包含 E 的最小闭集, 因为 $\overline{E} \supset E$ 也是闭集, 故 $F \subset \overline{E}$. 又因为 $E \subset F$, 且 F 是闭集, 故 $\overline{E} \subset F$, 从而只能有 $\overline{E} = F$. □

命题 2.3.2

列紧空间内任意 (闭) 子集都是 (自) 列紧集.

补充命题 2.3.5 (⋆)

度量空间 $(\mathscr{X}, \rho)$ 的子集 A 为自列紧集当且仅当 A 为列紧闭集.

证明

当 A 是自列紧集, 任取 $\{x_n\}_{n\in\mathbb{N}} \subset A$, 根据自列紧集的定义知 $\{x_n\}_{n\in\mathbb{N}}$ 中必有收敛到 A 中点的子列, 故 A 是列紧集. 又因为任取 $x \in \overline{A}$, 根据闭包的等价定义知存在点列 $\{x_n\}_{n\in\mathbb{N}} \subset A$ 使得 $x_n \to x(n \to \infty)$, 而届于 A 是自列紧的, 这说明 $\{x_n\}_{n\in\mathbb{N}}$ 应收敛到 A 中的点, 故 $x \in A$, 从而 $\overline{A} \subset A$, A 是闭集.

当 A 是列紧闭集, 任取 $\{x_n\}_{n\in\mathbb{N}} \subset A$, 因为 A 是列紧的, 故 $\{x_n\}_{n\in\mathbb{N}}$ 在 $\mathscr{X}$ 中有收敛子列. 又因为 A 是闭集, 故该收敛子列的极限必属于 A, 这说明 $\{x_n\}_{n\in\mathbb{N}}$ 中有收敛到 A 中的子列, 从而 A 是自列紧集. □

注 $(0,1]$ 是 $\mathbb{R}$ 的列紧集, 但不是 $\mathbb{R}$ 的自列紧集, 故 $(0,1]$ 本身不是列紧空间. 列紧空间必定自列紧.

补充命题 2.3.6 (⋆)

度量空间 $(\mathscr{X}, \rho)$ 的子集 A 列紧 $\Leftrightarrow \overline{A}$ 列紧 $\Leftrightarrow \overline{A}$ 自列紧.

证明

第二个等价关系藉由 $\overline{A}$ 的闭性与前述命题已经得证, 下面证明第一个等价关系.

当 $\overline{A}$ 列紧, 任取 A 的一个无穷点列 $\{x_n\}_{n\in\mathbb{N}}$, 则由 $A \subset \overline{A}$ 知 $\{x_n\}_{n\in\mathbb{N}}$ 也是 $\overline{A}$ 的一个无穷点列, 进而由 $\overline{A}$ 列紧知 $\{x_n\}_{n\in\mathbb{N}}$ 在 $\mathscr{X}$ 中有收敛子列, 从而 A 是列紧集.

当 A 是列紧集, 任取 $\{x_n\}_{n\in\mathbb{N}} \subset \overline{A}$, 由闭包的定义知对任意的 $n \in \mathbb{N}$, 存在 $y_n \in A$ 使得 $\rho(x_n, y_n) < \frac{1}{n}$. 因为 A 是列紧集, $\{y_n\}_{n\in\mathbb{N}} \subset A$, 故 $\{y_n\}_{n\in\mathbb{N}}$ 在 $\mathscr{X}$ 中有收敛子列, 设为 $\{y_{n_k}\}_{k\in\mathbb{N}}$. 不妨设 $\lim_{k\to\infty} y_{n_k} = y_0 \in \mathscr{X}$, 下面证明 $x_{n_k} \to y_0(k \to \infty)$. 根据极限的定义知:

$$\forall \varepsilon > 0 \exists k_0 \in \mathbb{N} \forall k > k_0(\rho(y_{n_k}, y_0) < \varepsilon)$$

因为 $n_K \to \infty(k \to \infty)$, 故取 $k_1 \in \mathbb{N}$ 使得 $k > k_1$ 时有 $\frac{1}{n_k} < \varepsilon$. 从而当 $k > \max\{k_0, k_1\}$ 时:

$$\rho(x_{n_k}, y_0) \leq \rho(x_{n_k}, y_{n_k}) + \rho(y_{n_k}, y_0) < \frac{1}{n_k} + \varepsilon < 2\varepsilon$$

由 ε 的任意性知 $\lim_{k\to\infty} x_{n_k} = y_0$, 这说明 $\overline{A}$ 列紧. 进一步, 因为 $\overline{A}$ 是闭集, 故 $y_0 \in \overline{A}$, 从而 $\overline{A}$ 是自列紧的. □

注 上面两个命题表明, 列紧集与自列紧集之间差一个闭性.

命题 2.3.3

列紧空间必是完备空间.

证明

设 $(\mathscr{X},\rho)$ 是一个列紧空间, $\{x_n\}$ 是其中的一串基本列, 不妨设其是一无穷点列. 由列紧性知, 存在 $\{x_n\}$ 的子列 $\{y_n\}$ 收敛到 $x_0\in\mathscr{X}$. 根据练习(2.9)即得 $x_n\to x_0(n\to\infty)$. □

在有限维空间中, 有界集与列紧集是等价的, 但在无穷维空间中未必如此. 要想让无穷维空间有类似的性质, 只能从" 有界" 与" 紧" 两方面下手: 要么给出更强的有界性, 要么给出更弱的紧性 (也即收敛性). 对收敛性的讨论会放在弱收敛一阶, 这里就讨论前一种方法: 引入更强的有界性.

定义 2.3.3 (ε– 网)

设 M 是 $(\mathscr{X},\rho)$ 中的一个子集, $\varepsilon>0$, $N\subset M$. 如果

$$\forall x\in M\exists y\in N(\rho(x,y)<\varepsilon)$$

则称 N 是 M 的一个 ε– 网. 如果 N 还是一个有穷集[a](个数依赖于 ε), 那么称 N 为 M 的一个有穷 ε– 网.

[a]即有限集.

注 根据定义显然有

$$M\subset\bigcup_{y\in N}B(y,\varepsilon).$$

补充命题 2.3.7 (⋆)

M 有界当且仅当存在 $\varepsilon_0>0$, 使得 M 有有穷 ε_0 网.

证明

当 M 有界, 根据定义知存在 $x_0\in\mathscr{X}, r>0$ 使得 $M\subset B(x_0,r)$. 另任取 $y_0\in M$, 知 $M\subset B(y_0,2r)$, 取 $\varepsilon_0=2r$ 即可.

当 M 有有穷 ε_0 网, 记 $M\subset\bigcup_{i=1}^N B(y_i,\varepsilon_0)$, 令

$$L:=\varepsilon_0+\max\{\rho(y_i,y_j):i,j\in\{1,\cdots,n\}\}$$

则 $B(y_i,\varepsilon_0)\subset B(y_1,L)$, 故 $M\subset B(y_1,L)$, 从而 M 有界. □

定义 2.3.4 (完全有界)

集合 M 称为是完全有界的, 如果 $\forall\varepsilon>0$, 都存在着 M 的一个有穷 ε– 网.

课堂笔记 (♡)

有界集不一定是完全有界集, 下面给出两个反例.

- 对任意的 $n\in\mathbb{N}$, 记

$$e_n:=\{\underbrace{0,\cdots,0}_{n-1},1,0,\cdots,\}\in l^2$$

  则 $\rho(e_n,\theta)=1$, 从而 $\{e_n\}_{n\in\mathbb{N}}\subset B(\theta,1+\varepsilon_0)(\varepsilon_0>0)$, 这说明 $\{e_n\}_{n\in\mathbb{N}}$ 是 l^2 中的有界集. 但对任意的 $\varepsilon\in(0,\frac{\sqrt{2}}{2}]$, $\{e_n\}_{n\in\mathbb{N}}$ 不存在有穷 ε 网. 这是因为若 $\{e_n\}_{n\in\mathbb{N}}$ 有有穷 ε 网 $\{B(_i,\varepsilon)\}_{i=1}^I$, 注意当 $n\neq m$ 时, $\rho(e_n,e_m)=\sqrt{2}\geq2\varepsilon$, 故对任意的 $i\in\{1,\cdots,I\}$, $B(y_i,\varepsilon)$ 中不能同时含有两个不同的 e_n,e_m. 又因为 $\{e_n\}_{n\in\mathbb{N}}$ 是无穷集, 故 $I=\infty$, 矛盾! 从而 $\{e_n\}_{n\in\mathbb{N}}$ 是 l^2 中的有界集, 但不是完全有界集.
- 设 $\mathscr{X}$ 为无穷集, 对任意的 $x,y\in\mathscr{X}$, 令

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

故 $\mathscr{X}$ 是有界集, 但 $\mathscr{X}$ 并非完全有界集, 这是因为对任意的 $\varepsilon \in (0, \frac{1}{2}]$, $\mathscr{X}$ 都没有有穷的 ε 网.

定理 2.3.1 (Hausdorff)

为了 (完备) 度量空间 $(\mathscr{X}, \rho)$ 中的集合 M 是列紧的, 必须 (且仅须)M 是完全有界集.

证明

如果 M 并非完全有界, 也即存在 $\varepsilon_0 > 0$ 使得 M 中没有有穷的 ε_0 网, 则任取 $x_1 \in M$, 知存在 $x_2 \in M \backslash B(x_1, \varepsilon_0)$. 对 $\{x_1, x_2\} \subset M$, 知存在 $x_3 \in M \backslash (B(x_1, \varepsilon_0) \cup B(x_2, \varepsilon_0))$, 这样一直下去, 知对 $\{x_1, \cdots, x_n\} \subset M$, 既然 M 没有有穷 ε_0 网, 总会存在 $x_{n+1} \in M \backslash \bigcup_{k=1}^{n} B(x_k, \varepsilon_0)$. 这样产生的点列 $\{x_n\} \subset M$ 显然满足 $\rho(x_n, x_m) \geq \varepsilon_0 (n \neq m)$, 这说明它没有收敛子列, 这与 M 的列紧性矛盾!

如果 M 完全有界且 $(\mathscr{X}, \rho)$ 完备, 对 M 中的无穷点列 $\{x_n\}$, 现在想找出它的一个收敛子列. 既然 M 完全有界, 知任取 $\varepsilon > 0$, 总是能找到 M 的一个有穷 ε 网, 这进而说明存在 $N(\varepsilon) = \{y_1, y_2, \cdots, y_{n(\varepsilon)}\}$ 使得 $M \subset \bigcup_{y \in N(\varepsilon)} B(y, \varepsilon)$, 从而根据鸽笼原理, 既然 $\{x_n\} \subset M$ 是无穷列, 知总会有某个 $B(y_i, \varepsilon)$ 包含了 $\{x_n\}$ 中的无穷多项.

现在, 对 1 网, 根据上面的讨论知存在 $y_1 \in M$, 使得 $\{x_1\}$ 的一个子列 $\{x_n^{(1)}\} \subset B(y_1, 1)$;

对 $\frac{1}{2}$ 网, 存在 $y_2 \in M$, 使得 $\{x_n^{(1)}\}$ 的一个子列 $\{x_n^{(2)}\} \subset B(y_2, \frac{1}{2})$;

这样一直下去, 对 $\frac{1}{k}$ 网, 存在 $y_k \in M$, 使得 $x_n^{(k-1)}$ 的一个子列 $\{x_n^{(k)}\} \subset B(y_k, \frac{1}{k})$.

最后, 抽出对角线子列 $\{x_k^{(k)}\}$, 下面说明它是一个基本列. 这是因为 $\forall \varepsilon > 0$, 当 $n > \frac{2}{\varepsilon}$ 时, 对任意的 $p \in \mathbb{N}$ 有

$$\rho(x_{(n+p)}^{n+p}, x_n^{(n)}) \leq \rho(x_{n+p}^{(n+p)}, y_n) + \rho(x_n^{(n)}, y_n) \leq \frac{2}{n} < \varepsilon.$$

从而 $\{x_k^{(k)}\}$ 确为基本列, 进而由 $\mathscr{X}$ 的完备性知 $\{x_k^{(k)}\}$ 存在极限, 因而 $\{x_n\}$ 有收敛子列, 从而 M 是列紧集. □

课堂笔记 (♡)

为免歧义, Hausdorff 定理还可以表述成: 给定 (X, ρ), $M \subset X$, 则

(i) M 在 X 内列紧 $\Rightarrow$ M 完全有界;

(ii) (X, ρ) 完备, M 完全有界 $\Rightarrow$ M 在 X 内列紧;

(iii) M 在 X 内自列紧 $\Leftrightarrow$ M 完全有界且 (M, ρ) 完备.

定义 2.3.5 (可分)

一个度量空间若有可数的稠密子集, 就称这个度量空间是可分的.

定理 2.3.2

完全有界的度量空间是可分的.

证明

设 $(\mathscr{X}, \rho)$ 完全有界, 根据定义知对任意的 ε, 其均有有穷 ε 网. 现取 N_n 为有穷的 $\frac{1}{n}$ 网, 则 $\bigcup_{n=1}^{\infty} N_n$ 首先可数, 其次任取 $x \in (\mathscr{X}, \rho)$, 总可依次在 $N_1, N_2, \cdots$ 中找到元素 $x_1, x_2, \cdots$ 满足

$$\rho(x_1, x) < 1, \rho(x_2, x) < \frac{1}{2}, \cdots, \rho(x_m, x) < \frac{1}{m}, \cdots$$

这说明 $\bigcup_{n=1}^{\infty} N_n$ 在 $(\mathscr{X}, \rho)$ 中稠密, 因而是可数稠密子集. □

定义 2.3.6 (紧集)

在拓扑空间 $\mathscr{X}$ 中, 集合 M 称为是紧的, 如果 $\mathscr{X}$ 中每个覆盖 M 的开集族中有有穷个开集覆盖集合 M.

注 集合还有一种称为预紧的紧性:

补充定义 2.3.2 (预紧集[RAJF])

对赋范空间 $\mathscr{X}$ 而言, 称 A 在 $\mathscr{X}$ 中预紧, 如果 A 的闭包 $\overline{A}$ 在 $\mathscr{X}$ 中是紧集. ♣

可以证明列紧集必是预紧集: 对任意列紧集 A, 知 $\overline{A}$ 是列紧集, 又因为 $\overline{A}$ 是闭集, 故 $\overline{A}$ 是自列紧集, 因而 $\overline{A}$ 是紧集, 从而 A 是预紧集.

在 Banach 空间中, Hausdorff 定理实际上对预紧集成立:

补充定理 2.3.1 (Hausdorff[RAJF])

Banach 空间 $\mathscr{X}$ 的子集 A 是预紧集当且仅当其完全有界. ♡

集合的预紧性在讨论算子的紧性时起着决定性的作用.

定理 2.3.3

设 $(\mathscr{X},\rho)$ 是一个度量空间, 为了 $M\subset\mathscr{X}$ 是紧的必须且仅须 M 是自列紧集. ♡

证明

当 M 是紧集, 首先证明 M 是闭集, 进而只需证明 M 的余集是开集. 对任意的 $x_0\in\mathscr{X}\backslash M$, 注意到显见的开覆盖:

$$M\subset\bigcup_{x\in M}B(x,\frac{1}{2}\rho(x,x_0))$$

既然 M 是紧的, 知存在 $x_k\in M(k=1,2,\cdots,n)$, 使得

$$M\subset\bigcup_{k=1}^{n}B(x_k,\frac{1}{2}\rho(x_k,x_0))$$

取 $\delta=\min\limits_{1\le k\le n}\frac{1}{2}\rho(x_k,x_0)$, 则显见 $\delta>0$, 且

$$\forall x\in B(x_0,\delta),\rho(x,x_k)\ge\rho(x_k,x_0)-\rho(x_0,x)>2\delta-\delta=\delta,\ k=1,2,\cdots,n$$

这说明 $x\notin M$, 也即 $B(x_0,\delta)\cap M=\varnothing$, 故 M 的余集是开集, M 是闭集. 现在只需要说明 M 列紧即可, 这是因为 $M'\subset M$, 也即 M 的全体极限点就在 M 中, 故只要 M 列紧, 其必自列紧.

下面证明 M 是列紧集, 用反证法. 如果有 M 中的点列 $\{x_n\}$ 不含有收敛子列, 不妨假定 x_n 互异. 对每个 $n\in\mathbb{N}$, 作集合

$$S_n:=\{x_1,x_2,\cdots,x_{n-1},x_{n+1},x_{n+2},\cdots\}$$

因为 S_n 无收敛子列, 故 $S_n'=\varnothing$, 进而 S_n 是闭集, 从而每个 $\mathscr{X}\backslash S_n$ 是开集, 但

$$\bigcup_{n=1}^{\infty}(\mathscr{X}\backslash S_n)=\mathscr{X}\backslash\bigcap_{n=1}^{\infty}S_n=\mathscr{X}\backslash\varnothing=X\supset M$$

由 M 的紧性, 既然 $\bigcup_{n=1}^{\infty}(\mathscr{X}\backslash S_n)$ 是 M 的开覆盖, 就必存在子覆盖 $\bigcup_{n=1}^{N}(\mathscr{X}\backslash S_n)$, 即

$$\mathscr{X}\backslash\{x_n\}_{n=N+1}^{\infty}\supset M$$

但这是不可能的, 因为 x_{N+1} 在上式右端中, 却不在上式左端中, 矛盾. 进而 M 是列紧集.

当 M 是自列紧集, 要在 M 的任一开覆盖中取出有限子覆盖. 用反证法, 如果某个开覆盖 $\bigcup_{\lambda\in\Lambda}G_\lambda\supset M$ 不能取出 M 的有限子覆盖, 届于 M 是自列紧的, 知对任意的 $n\in\mathbb{N}$, 存在有穷的 $\frac{1}{n}$ 网

$$N_n=\{x_1^{(n)},x_2^{(n)},\cdots,x_{k(n)}^{(n)}\}$$

显见 $\bigcup_{y\in N_n}B(y,\frac{1}{n})\supset M$, 进而对任意的 $n\in\mathbb{N}$, 存在 $y_n\in N_n$, 使得 $B(y_n,\frac{1}{n})$ 不能被有限个 G_λ 所覆盖. 因为 M 是自列紧集, 故这样的序列 $\{y_n\}$ 必存在收敛子列 y_{n_k}, 记其收敛到 $y_0\in G_{\lambda_0}$. 又因为 G_{λ_0} 是开集, 故存在 $\delta>0$, 使得 $B(y_0,\delta)\subset G_{\lambda_0}$. 对这个 δ, 取足够大的 k 使得 $n_k>\frac{2}{\delta}$, 且 $\rho(y_{n_k},y_0)<\frac{\delta}{2}$, 则对任意的 $x\in B(y_{n_k},\frac{1}{n_k})$ 有

$$\rho(x,y_0)\le\rho(x,y_{n_k})+\rho(y_{n_k},y_0)\le\frac{1}{n_k}+\frac{\delta}{2}<\delta$$

这说明 $x\in B(y_0,\delta)$, 故 $B(y_{n_k},\frac{1}{n_k})\subset B(y_0,\delta)\subset G_{\lambda_0}$, 这与每个 $B(y_n,\frac{1}{n})$ 不能被有限个 G_λ 所覆盖矛盾! □

补充命题 2.3.8 (♡)

连续映射把紧集映为紧集, 也就是说, 紧集的连续像依旧是紧集.

证明

给定度量空间 $(X,\rho),(Y,\tilde{\rho})$, 设 $f:X\to Y$ 是连续映射, 设 $M\subset X$ 是紧集, 希望证明 $f(M)\subset Y$ 是紧集. 若 $A\subset Y$, 约定:

$$f^{-1}(A)=\{x\in X:f(x)\in A\}$$

现任取 $f(M)$ 在 Y 中的一族开覆盖 $\{G_\lambda:\lambda\in\Lambda\}$, 知 $\{f^{-1}(G_\lambda):\lambda\in\Lambda\}$ 是 X 中的开集族, 且显见 $M\subset\bigcup_{\lambda\in\Lambda}f^{-1}(G_\lambda)$, 故 $\{f^{-1}(G_\lambda)\}$ 是 M 在 X 中的开覆盖, 由 M 的紧性知其具有有限子覆盖 $\{f^{-1}(G_k)\}_{k=1}^n$, 因而 $\{G_k\}_{k=1}^n$ 是 $f(M)$ 的有限子覆盖, 故 $f(M)$ 是紧集, 命题得证. □

如果知道了集合的紧性, 就说明可以对集合进行某种" 剖分", 从而可取用有限个" 代表" 进行研究, 得到某些一致性的结论. 应用该方法的一个典型结果为:

补充命题 2.3.9 (♡)

紧集上的连续映射必定一致连续.

证明

设 M 是紧集, 任取 $\varepsilon>0,x\in M,\varphi\in C(M)$, 由 φ 在 x 处连续知:

$$\exists\delta=\delta(\varepsilon,x)\forall y\in B(x,\delta)(|\varphi(y)-\varphi(x)|<\varepsilon)$$

因为 $M\subset\bigcup_{x\in M}B(x,\frac{\delta(\varepsilon,x)}{2})$, 故根据 M 的紧性知存在 $x_1,\cdots,x_n\in M$ 使得 $M\subset\bigcup_{i=1}^n B(x_i,\frac{\delta(\varepsilon,x_i)}{2})$. 取 $\delta_0=\min_{1\le i\le n}\delta(\varepsilon,x_i)$, 知 $\delta_0>0$. 现任取 $x,y\in M$, 当 $\rho(x,y)<\frac{\delta_0}{2}$ 时, 知必存在 $i_0\in\{1,\cdots,n\}$ 使得 $x\in B(x_{i_0},\frac{\delta(\varepsilon,x_{i_0})}{2})$, 进而

$$\rho(y,x_{i_0})\le\rho(y,x)+\rho(x,x_{i_0})<\frac{\delta_0}{2}+\frac{\delta(\varepsilon,x_{i_0})}{2}\le\delta(\varepsilon,x_{i_0})$$

因而

$$|\varphi(y)-\varphi(x)|\le|\varphi(y)-\varphi(x_{i_0})|+|\varphi(x_{i_0})-\varphi(x)|<2\varepsilon$$

故 φ 一致连续. □

课堂笔记 (⋆)

由上述命题知当 M 是紧集, 则 $C(M)$ 中任意有限个函数构成的函数族必定等度连续. 事实上, 设 $\{\varphi_i\}_{i=1}^n\subset C(M)$, 对任意取定的 $i\in\{1,\cdots,n\}$, 知 φ_i 在 M 上一致连续, 进而

$$\forall\varepsilon>0\exists\delta_i(\varepsilon)>0\forall x_1,x_2\in M(\rho(x_1,x_2)<\delta_i(\varepsilon)\Rightarrow|\varphi_i(x_1)-\varphi_i(x_2)|<\varepsilon)$$

现令 $\delta=\min_{1\le i\le n}\{\delta_i(\varepsilon)\}$, 即得

$$\forall i\in\{1,\cdots,n\}(\rho(x_1,x_2)<\delta\Rightarrow|\varphi_i(x_1)-\varphi_i(x_2)|<\varepsilon)$$

故 $\{\varphi_i\}_{i=1}^n$ 等度连续.

现在对 $C[a,b]$ 作推广: 设 M 是一个紧的度量空间, 带有距离 ρ, 用 $C(M)$ 表示 $M\to\mathbb{R}$ 的一切连续映射全体, 定义

$$d(u,v)=\max_{x\in M}|u(x)-v(x)| \tag{2.18}$$

命题 2.3.4

$(C(M), d)$ 是一个度量空间.

证明

只需验证 d 定义的合理性, 也即对任意的 $u \in C(M)$, 存在最大值 $\max\limits_{x\in M}|u(x)|$.

首先说明 $u(M)$ 是紧集. 对任意点列 $y_n \in u(M)$, 知存在 $x_n \in M$ 使得 $u(x_n) = y_n$. 因为 M 是紧的, 故其是自列紧的, 进而存在子列 $x_{n_k} \to x_0, k \to \infty$. 又因为 u 连续, 故 $u(x_{n_k}) \to u(x_0) \in u(M), k \to \infty$. 令 $y_0 := u(x_0)$, 即得 $y_{n_k} = u(x_{n_k}) \to y_0$, 故 $u(M)$ 是紧集. 又因为 $u(M) \subset \mathbb{R}$, 故其为有界闭集. 设

$$\min u(M) = \alpha,\ \max u(M) = \beta$$

由闭性知 $\alpha, \beta \in u(M)$, 故 $\max\limits_{x\in M}|u(x)|$ 存在. □

命题 2.3.5

$(C(M), d)$ 是完备的.

证明

任取 $C(M)$ 中的基本列 $\{u_n\}_{n\in\mathbb{N}}$, 根据定义有:

$$\forall\varepsilon > 0\exists N = N(\varepsilon) > 0\forall m > n > N(\max_{x\in M}|u_m(x) - u_n(x)| < \varepsilon)$$

亦即

$$\forall\varepsilon > 0\exists N = N(\varepsilon) > 0\forall m > n > N\forall t \in M(|u_m(x) - u_n(x)| < \varepsilon) \tag{2.19}$$

这说明对每个固定的 x_0, $\{u_n(x_0)\}_{n\in\mathbb{N}}$ 都是 $\mathbb{R}$ 上的基本列, 因而由 $\mathbb{R}$ 的完备性知其存在极限, 设之为 $u(x_0)$. 现在在(2.19)式中令 $m \to \infty$ 知:

$$\forall\varepsilon > 0\exists N = N(\varepsilon) > 0\forall n > N\forall x \in M(|u(x) - u_n(x)| \le \varepsilon)$$

这说明

$$u(x) = \lim_{n\to\infty} u_n(x),\ \forall x \in M$$

下面说明 $u \in C(M)$. 因为 $\{u_n\} \subset C(M)$, 故取定 $n_0 \in \mathbb{N}$, 有:

$$\forall s \in M\forall\varepsilon > 0\exists\delta > 0\forall s' \in B(s, \delta)(|u_{n_0}(s) - u_{n_0}(s')| < \varepsilon)$$

不妨令 $n_0 > N$, 此时有:

$$\begin{aligned}|u(s) - u(s')| &= |u(s) - u_{n_0}(s) + u_{n_0}(s) - u_{n_0}(s') + u_{n_0}(s') - u(s')|\\ &\le |u(s) - u_{n_0}(s)| + |u_{n_0}(s) - u_{n_0}(s')| + |u_{n_0}(s') - u(s')| < 3\varepsilon\end{aligned}$$

由 ε 的任意性即得 $u \in C(M)$. □

下面讨论连续函数空间上列紧集的刻画.

定义 2.3.7 (一致有界, 等度连续)

设 F 是 $C(M)$ 的一个子集. 称 F 是一致有界的, 如果

$$\exists M_1 > 0\forall x \in M\forall\varphi \in F(|\varphi(x)| \le M_1)$$

称 F 是等度连续的, 如果

$$\forall\varepsilon > 0\exists\delta(\varepsilon) > 0\forall x_1, x_2 \in M\forall\varphi \in F(\rho(x_1, x_2) < \delta \Rightarrow |\varphi(x_1) - \varphi(x_2)| < \varepsilon)$$

课堂笔记 (♡)

回忆 F 一致有界的定义, 这相当于在说:

$$\exists M_1 > 0 \forall \varphi \in F(\rho(\varphi, 0) \leq M_1)$$

其中 $\rho(\cdot) = \max_M(\cdot)$ 是一致收敛性度量, 从而 F 一致有界意味着 F 是 $C(M)$ 中的有界集.

定理 2.3.4 (Arzela-Ascoli)

为了 $F \subset C(M)$ 是一个列紧集, 必须且仅须 F 是一致有界且等度连续的函数族.

证明

因为 $C(M)$ 是完备的, 故由 Hausdorff 定理2.3.1知, F 是列紧的当且仅当它是完全有界的. 故下面考虑证明 F 完全有界当且仅当 F 是一致有界且等度连续的函数族.

当 F 是完全有界集, 知其首先是有界集, 又因为 M 是紧集, 故 F 是一致有界函数族. 现在要证明 F 等度连续, 即证

$$\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon) \forall \varphi \in F \forall x_1, x_2 \in M(\rho(x_1, x_2) < \delta \Rightarrow |\varphi(x_1) - \varphi(x_2)| < \varepsilon)$$

因为 F 是完全有界集, 故其 $\frac{\varepsilon}{3}$ 网是一个有穷集: $N(\frac{\varepsilon}{3}) = \{\varphi_1, \varphi_2, \cdots, \varphi_n\}$, 对这有穷个函数, 因为它们均连续, 故

$$\exists \delta = \delta(\frac{\varepsilon}{3})(\rho(x_1, x_2) < \delta \Rightarrow |\varphi_i(x_1) - \varphi_i(x_2)| < \frac{\varepsilon}{3}, i = 1, 2, \cdots, n)$$

又根据 $\frac{\varepsilon}{3}$ 网的定义知

$$\forall \varphi \in F \exists \varphi_i \in N(\frac{\varepsilon}{3})(d(\varphi, \varphi_i) < \frac{\varepsilon}{3})$$

故当 $\rho(x, x') < \delta$ 时有

$$\begin{aligned} |\varphi(x) - \varphi(x')| &\leq |\varphi(x) - \varphi_i(x)| + |\varphi_i(x) - \varphi_i(x')| + |\varphi_i(x') - \varphi(x')| \\ &\leq 2d(\varphi, \varphi_i) + |\varphi_i(x) - \varphi_i(x')| < \varepsilon \end{aligned}$$

故 F 等度连续.

当 F 一致有界且等度连续, 只需找有穷的 ε 网, 因为这样一来就能说明 F 完全有界, 而完备空间中的完全有界集必是列紧集. 由 F 等度连续知

$$\exists \delta = \delta(\frac{\varepsilon}{3}) > 0 \forall x, x' \in M \forall \varphi \in F(\rho(x, x') < \delta \Rightarrow |\varphi(x) - \varphi(x')| < \frac{\varepsilon}{3})$$

因为 M 本身是紧的, 故其自然是自列紧的, 进而有 Hausdorff 定理知 M 完全有界, 这说明对这个 δ, 可以选取 M 上的有穷 δ 网 $N(\delta) = \{x_1, x_2, \cdots, x_n\}$. 做映射 $T : F \to \mathbb{R}^n$:

$$T\varphi := (\varphi(x_1), \varphi(x_2), \cdots, \varphi(x_n)),\ \forall \varphi \in F$$

记 $\tilde{F} = T(F)$, 下面说明 $\tilde{F}$ 在 $\mathbb{R}^n$ 中有界, 这是因为设 $|\varphi| \leq M_1 (\forall \varphi \in F)$, 则

$$(\sum_{i=1}^{n} |\varphi(x_i)|^2)^{\frac{1}{2}} \leq \sqrt{n} \max_{x \in M} |\varphi(x)| \leq \sqrt{n} M_1,\ \forall \varphi \in F$$

故 $\tilde{F}$ 以 $\sqrt{n} M_1$ 为上界. 这说明 $\tilde{F}$ 是列紧集[9], 故根据 Hausdorff 定理2.3.1, $\tilde{F}$ 是完全有界集, 进而有有穷的 $\frac{\varepsilon}{3}$ 网:

$$\tilde{N}(\frac{\varepsilon}{3}) = \{T\varphi_1, T\varphi_2, \cdots, T\varphi_m\}.$$

因为对任意的 $\varphi \in F$, 总存在 φ_i, 使得 $\rho_n(T\varphi, T\varphi_i) < \frac{\varepsilon}{3}$, 故取定 $x_r \in N(\delta)$ 满足 $\rho(x, x_r) < \delta$ 时, 有:

$$|\varphi(x) - \varphi_i(x)| \leq |\varphi(x) - \varphi(x_r)| + |\varphi(x_r) - \varphi_i(x_r)| + |\varphi_i(x_r) - \varphi_i(x)| < \frac{2}{3}\varepsilon + \rho_n(T\varphi, T\varphi_i) < \varepsilon$$

这说明 $\{\varphi_1, \varphi_2, \cdots, \varphi_m\}$ 是 F 的 ε 网, 命题即证. □

注 事实上 Arzela-Ascoli 定理也对预紧集成立:

[9]这是因为 $\tilde{F}$ 活在欧氏空间 $\mathbb{R}^n$ 中, 此时有界可推出任意序列有收敛子列.

补充定理 2.3.2 (Arzela-Ascoli[RAJF])

设 Ω 是 $\mathbb{R}^n$ 中的区域, 则 $K \subset C(\overline{\Omega})$ 是 $C(\overline{\Omega})$ 中的预紧集当且仅当 K 一致有界且等度连续.

例 2.19 设 $\Omega \subset \mathbb{R}^n$ 是有界开凸集. 若 M_1, M_2 是两个给定的正数, 则集合

$$F := \{\varphi \in C^{(1)}(\overline{\Omega}) || \varphi(x)| \le M_1, |\operatorname{grad}\varphi(x)| \le M_2, \forall x \in \Omega\}$$

是 $C(\overline{\Omega})$ 上的一个列紧集, 其中 $C^{(1)}(\overline{\Omega})$ 表示 $\overline{\Omega}$ 上的连续可微函数全体.

证明

因为

$$\forall \varphi \in F \forall x_1, x_2 \in \overline{\Omega} \exists \theta \in (0,1)(\varphi(x_1) - \varphi(x_2) = \operatorname{grad}\varphi(\theta x_1 + (1-\theta)x_2) \cdot (x_1 - x_2))$$

故

$$|\varphi(x_1) - \varphi(x_2)| \le M_2 \rho_n(x_1, x_2), \ \forall \varphi \in F$$

这说明 F 等度连续, 而其根据构造已经是一致有界的了, 由 Arzela-Ascoli 定理即证命题. □

补充定理 2.3.3 (♡)

l^p 中的子集 A 列紧当且仅当下述两个断言成立:

(i) 一致有界: $\exists K > 0 \forall x = \{x_n\}_{n=1}^\infty \in A(\sum_{n=1}^\infty |x_n|^p < K)$;

(ii) 等度收敛[a]: $\forall \varepsilon > 0 \exists N = N(\varepsilon) > 0 \forall m > N \forall x = \{x_n\}_{n=1}^\infty \in A(\sum_{n=m}^\infty |x_n|^p < \varepsilon)$.

[a]注意这个概念比一致收敛更强: 一致收敛指的对象是单个序列 x, 而等度收敛指的对象是序列族 $A \supset \{x\}$.

证明[WKY]

当 A 是列紧集, 显见其有界, 亦即一致有界性成立. 对于等度收敛性, 由 Hausdorff 定理知 M 是完全有界集, 因而其对任意 ε 有有限 ε 网 $\{x_r : x_r = (x_r^1, x_r^2, \cdots) \in l^p, r = 1, 2, \cdots, r_\varepsilon\}$. 现在该有限 ε 网本身是等度收敛的, 亦即对任意的 $\varepsilon' > 0$ 存在 $N > 0$ 使得:

$$(\sum_{n=N+1}^\infty |x_r^n|^p)^{\frac{1}{p}} < \varepsilon', \ \forall r = 1, 2, \cdots, r_\varepsilon$$

现在任取 $x = (x^1, x^2, \cdots) \in A$, 根据 ε 网的定义知存在 x_{r_x} 使得:

$$\rho(x, x_{r_x}) = (\sum_{k=1}^\infty |x^k - x_{r_x}^k|^p)^{\frac{1}{p}} < \varepsilon$$

取 $x(N) = (x^1, x^2, \cdots, x^N, 0, \cdots)$, 则

$$(\sum_{i=N+1}^\infty |x^i|^p)^{\frac{1}{p}} = \rho(x, x(N)) \le \rho(x, x_{r_x}) + \rho(x_{r_x}, x(N)) < \varepsilon + \rho(x_{r_x}, x(N))$$

而

$$\begin{aligned}\rho(x_{r_x}, x(N)) &= (\sum_{j=1}^N |x_{r_x}^j - x^j|^p + \sum_{j=N+1}^\infty |x_{r_x}^j|^p)^{\frac{1}{p}} \\ &\le (\sum_{j=1}^N |x_{r_x}^j - x^j|^p)^{\frac{1}{p}} + (\sum_{j=N+1}^\infty |x_{r_x}^j|^p)^{\frac{1}{p}} \\ &\le \varepsilon + \varepsilon'\end{aligned}$$

故

$$(\sum_{i=N+1}^\infty |x^i|^p)^{\frac{1}{p}} < 2\varepsilon + \varepsilon'$$

令 $\varepsilon' = \varepsilon$, 并由 ε 的任意性即得等度收敛性.

当 $A\subset l^p$ 一致有界且等度收敛, 根据等度收敛的定义:

$$\forall\varepsilon>0\exists N=N(\varepsilon)>0\forall m>N\forall x=\{x_n\}_{n=1}^{\infty}\in A(\sum_{n=m}^{\infty}|x_n|^p<\varepsilon)$$

取

$$S=\{(x_1,\cdots,x_N,0,\cdots):x=(x_1,\cdots,x_n,\cdots)\in A\}$$

知 S 为 A 的 ε 网, 由 A 在 l^p 内有界知 S 在 l^p 内有界, 因而 S 可视作 $\mathbb{R}^N$ 中的有界集, 进而其为列紧集. 另一方面, $\mathbb{R}^N$ 中的列紧性可导出 l^p 中的列紧性, 故 S 在 A 中列紧, 因而其为 A 中的列紧 ε 网. 最后利用 ε 网的定义即知 A 列紧. □

下面讨论定理2.3.3的连续版本: Riesz 定理. 设 u 是在 $\Omega\subset\mathbb{R}^n$ 上几乎处处有定义的函数, 设

$$\widetilde{u}(x)=\begin{cases}u(x), & x\in\Omega\\ 0, & x\in\mathbb{R}^n\backslash\Omega\end{cases}$$

补充定理 2.3.4 (Riesz♡RAJF)

设 $1\le p<\infty$, 则有界子集 $K\subset L^p(\Omega)$ 是 $L^p(\Omega)$ 中的预紧集, 当且仅当对任意 $\varepsilon>0$, 存在 $\delta>0$ 与子集 $G\subset\Omega$, 满足 $\overline{G}\subset\Omega$ 是紧集, 且对任意 $u\in K, h\in\mathbb{R}^n$, 若 $|h|<\delta$, 则:

(i) $\int_\Omega|\widetilde{u}(x+h)-\widetilde{u}(x)|^p dx<\varepsilon^p$;

(ii) $\int_{\Omega\backslash\overline{G}}|u(x)|^p dx<\varepsilon^p$.

♡

证明

记

$$T_h(u):=u(x+h)$$

当 K 是 $L^p(\Omega)$ 中的预紧集, 取定 $\varepsilon>0$, 由 Hausdorff 定理知 K 完全有界, 因而其有有穷 ε 网 $\{u_i\}_{i=1}^{n_1}$. 已知紧支连续函数空间 $C_0(\Omega)$ 在 $L^p(\Omega)$ 中稠密, 故存在有限个紧支连续函数 $\{\phi_j\}_{j=1}^{n_2}$ 使得

$$\forall u\in K\exists j\in\{1,\cdots,n_2\}(\|u-\phi_j\|_{L^p(\Omega)}\le\|u-u_i\|_{L^p(\Omega)}+\|u_i-\phi_j\|_{L^p(\Omega)}<2\varepsilon)$$

设 $G=\bigcup_{j=1}^{n_2}\operatorname{supp}(\phi_j)$, 则显见 $\overline{G}\subset\Omega$ 是紧集, 且:

$$\int_{\Omega\backslash\overline{G}}|u(x)|^p dx\le\int_{\Omega\backslash\overline{G}}(|u(x)-\phi_j(x)|^p+|\phi_j(x)|^p)dx<2^p\varepsilon^p m(\Omega\backslash\overline{G})$$

进而 (ii) 得证. 要证明 (i), 选取 $\overline{B}(0,r)\supset G$, 知 $|h|<1$ 时 $(T_h\phi_j-\phi_j)(x)=\phi_j(x+h)-\phi_j(x)$ 一致连续且在 $B(0,r+1)$ 外为零 $(1\le j\le n_2)$, 进而

$$\lim_{|h|\to0}\int_{\mathbb{R}^n}|T_h\phi_j(x)-\phi_j(x)|^p dx=0,\ 1\le j\le n_2$$

是一致的. 当 $|h|$ 充分小, 可设 $\|T_h\phi_j-\phi\|_{L^p(\Omega)}<\varepsilon$, 则当 $\|u-\phi_j\|_{L^p(\Omega)}<\varepsilon$ 时, 依照 Lebesgue 积分的平移不变性得 $\|T_h\widetilde{u}-T_h\phi_j\|_{L^p(\Omega)}<\varepsilon$, 因而对充分小的 $|h|$ 有

$$\|T_h\widetilde{u}-\widetilde{u}\|_{L^p(\Omega)}\le\|T_h\widetilde{u}-T_h\phi_j\|_{L^p(\Omega)}+\|T_h\phi_j-\phi_j\|_{L^p(\Omega)}+\|\phi_j-u\|_{L^p(\Omega)}<3\varepsilon$$

(i) 进而成立.

当对任意 $\varepsilon>0$, 存在 $\delta>0$ 与子集 $G\subset\Omega$, 满足 $\overline{G}\subset\Omega$ 是紧集, 且 (i),(ii) 式对 $|h|<\delta$ 成立, 固定该 ε, 记 $\widetilde{K}=\{\widetilde{u}:u\in K\}$, 由 (ii) 知可取 $G\subset\mathbb{R}^n$ 使得

$$\int_{\Omega\backslash\overline{G}}|u(x)|^p dx<\varepsilon$$

设 $\xi > 0, \eta_\xi$ 是磨光核, 显见 $\eta_\xi * u \in C^\infty(\Omega) \supset C^\infty(\overline{G})$. 现对 $\phi \in C_0(\Omega)$, 由 Hölder 不等式知:

$$\begin{aligned}
|\eta_\xi * \phi(x) - \phi(x)|^p &= |\int_\Omega \eta_\xi(y)(\phi(x-y) - \phi(x))dy|^p \\
&\leq (\int_\Omega (\eta_\xi^{\frac{1}{p}}(y)|\phi(x-y) - \phi(x)|)\eta_\xi^{1-\frac{1}{p}}(y)dy)^p \\
&\leq ((\int_\Omega \eta_\xi(y)|\phi(x-y) - \phi(x)|^p dy)^{\frac{1}{p}} (\int_\Omega \eta_\xi^{\frac{p-1}{p}\frac{p}{p-1}}(y)dy)^{1-\frac{1}{p}})^p \\
&= \int_{\mathrm{supp}(\eta_\xi)} \eta_\xi(y)|T_{-y}\phi(x) - \phi(x)|^p dx
\end{aligned} \tag{2.20}$$

这说明

$$\|\eta_\xi * \phi - \phi\|_{L^p(\Omega)} \leq \sup_{h \in \mathrm{supp}(\eta_\xi)} \|T_h \phi - \phi\|_{L^p(\Omega)}$$

当 $u \in L^p(\Omega)$, 设 $\{\phi_j\}$ 是依 L^p 范数趋于 u 的 $C_0(\Omega)$ 函数列, 进而 $\{\eta_\xi * \phi\}$ 是依 L^p 范数趋于 $\eta_\xi * u$ 的序列, 进一步知

$$\|\eta_\xi * u - u\|_{L^p(\Omega)} \leq \sup_{h \in \mathrm{supp}(\eta_\xi)} \|T_h u - u\|_p$$

由 (i) 知 $\lim_{|h|\to 0} \|T_h u - u\|_{L^p(\Omega)} = 0$ 对 $u \in K$ 一致成立, 因而 $\lim_{\xi \to 0} \|\eta_\xi * u - u\|_{L^p(\Omega)} = 0$ 对 $u \in K$ 一致成立, 进而可选定 $\xi_0 > 0$ 使得

$$\int_G |\eta_{\xi_0} * u(x) - u(x)|^p dx < \varepsilon, \ \forall u \in K$$

下面说明 $\{\eta_{\xi_0} * u : u \in K\}$ 是 $C(\overline{G})$ 上的预紧集, 根据 Arzela-Ascoli 定理只需说明其一致有界且等度连续. 由(2.20)式知

$$|\eta_{\xi_0} * u(x)| \leq (\sup_{y \in \mathbb{R}^n} \eta_{\xi_0}(y))^{\frac{1}{p}} \|u\|_{L^p(\Omega)}$$

因而该函数族对 $u \in K$ 一致有界. 类似有

$$|\eta_{\xi_0} * u(x+h) - \eta_{\xi_0} * u(x)| \leq (\sup_{y \in \mathbb{R}^n} \eta_{\xi_0}(y))^{\frac{1}{p}} \|T_h u - u\|_{L^p(\Omega)}$$

进而该函数族对 $u \in K$ 等度连续, 从而其为 $C(\overline{G})$ 中的预紧集, 因而由 Hausdorff 定理知其具有穷 ε 网, 最后易证该有穷 ε 网亦为 K 在 $L^p(\Omega)$ 中的 ε 网, 从而 K 作为完全有界集, 由 Hausdorff 定理可推知其预紧性. □

补充定理 2.3.5 (Montel♡)

设 $\Omega \subset \mathbb{C}$ 是区域, 并设 $\{f_n\} \subset H(\Omega)$ 是 Ω 上的解析函数, 若 $\{f_n\}$ 在 Ω 上一致有界, 则对任意区域 $D \subset\subset \Omega$[a], 总存在子列 $\{f_{n_j}\}$ 在 D 上一致收敛.

[a]即 $D \subset \overline{D} \subset \Omega$, 这个关系叫做紧包含, 在之后的 PDE 课程中也会遇到.

♡

注 [**Zo**] 中给出了更一般的 Arzela-Ascoli 定理, 为此先介绍一般情况下的等度连续, 一致有界与完全有界.

定义 2.3.8 (一致有界)

设 X 是集合, Y 是度量空间, $\mathcal{F}$ 是函数 $f: X \to Y$ 构成的函数族. 如果函数值的集合 $V = \{y \in Y : \exists f \in \mathcal{F} \exists x \in X (y = f(x))\}$ 在 Y 中有界, 就称 $\mathcal{F}$ 为集合 X 上的一致有界函数族. ♣

定义 2.3.9 (完全有界)

如果函数族 $\mathcal{F}$ 的函数值集合 $V \subset Y$ 完全有界 (亦即对任意 $\varepsilon > 0$ 存在有限 ε 网), 就称 $\mathcal{F}$ 为完全有界函数族. ♣

定义 2.3.10 (等度连续)

设 X, Y 是度量空间, $\mathcal{F}$ 是函数 $f : X \to Y$ 构成的函数族. 如果

$$\forall \varepsilon > 0 \exists \delta > 0 \forall f \in \mathcal{F} \forall x_1, x_2 \in X (d_X(x_1, x_2) < \delta \Rightarrow d_Y(f(x_1), f(x_2)) < \varepsilon)$$

就称 $\mathcal{F}$ 为集合 X 上的等度连续函数族.

♣

定理 2.3.5 (Arzela-Ascoli)

设 K 是度量紧空间[a], Y 是完备度量空间, $\mathcal{F}$ 是函数 $f : K \to Y$ 构成的函数族. 则 $\mathcal{F}$ 中的任意序列 $\{f_n\}_{n\in\mathbb{N}}$ 都包含一致收敛子序列, 当且仅当 $\mathcal{F}$ 完全有界且等度连续.

[a]即自列紧空间.

♡

证明

当任意序列 $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$ 均有一致收敛子列, 用反证法.

若 $\mathcal{F}$ 不是完全有界函数族, 根据定义知存在 ε_0 使得 $\mathcal{F}$ 中没有有穷 ε_0 网. 任取 $f_1 \in \mathcal{F}$, 知存在 $f_2 \in \mathcal{F} \backslash B(f_1, \varepsilon_0)$; 对于 $\{f_1, f_2\}$, 知存在 $f_3 \in \mathcal{F} \backslash \bigcup_{k=1}^{2} B(f_k, \varepsilon_0) \cdots$, 这样选取得到的序列 $\{f_n\}_{n\in\mathbb{N}}$ 显然没有收敛子列, 矛盾.

若 $\mathcal{F}$ 不是等度连续函数族, 根据定义知

$$\exists \varepsilon_0 > 0 \forall \delta > 0 \exists f_\delta \in \mathcal{F} \exists x_{1\delta}, x_{2\delta} \in X (d_X(x_{1\delta}, x_{2\delta}) < \delta \wedge d_Y(f_\delta(x_{1\delta}), f_\delta(x_{2\delta})) \geq \varepsilon_0)$$

从序列的角度说, 此即

$$\exists \varepsilon_0 > 0 \forall N > 0 \exists n > N \exists f_n \in \mathcal{F} \exists x_{1n}, x_{2n} \in X (d_X(x_{1n}, x_{2n}) \to 0 (n \to \infty) \wedge d_Y(f_n(x_{1n}), f_n(x_{2n})) \geq \varepsilon_0)$$

但此时 $\{f_n\}_{n\in\mathbb{N}}$ 不包含一致收敛子序列, 矛盾! 故 $\mathcal{F}$ 只能是完全有界且等度连续的.

当 $\mathcal{F}$ 完全有界且等度连续, 不妨设 K 是无限集, 根据紧性的定义知 K 是完全有界的, 因而其为可分空间, 设 $K \supset E = \{x_n\}_{n\in\mathbb{N}}$ 是 K 的稠密子集, 并任取 $\mathcal{F}$ 中的函数列 $\{f_k\}_{k\in\mathbb{N}}$.

对于 Y 中的序列 $\{f_k(x_1)\}_{k\in\mathbb{N}}$, 因为 $\mathcal{F}$ 是完全有界的, 故 $\{f_k(x_1)\}_{k\in\mathbb{N}}$ 本身作为 Y 的子集是完全有界的, 又因为 Y 完备, 故由 Hausdorff 定理知 $\{f_k(x_1)\}_{k\in\mathbb{N}}$ 有收敛子列 $\{f_{k_p}(x_1)\}_{p\in\mathbb{N}}$. 为方便表示, 记 $f_{k_p}(x_1)$ 对应的函数为 $f_{k_p}^1$, 其中上标 1 表示这是基于 x_1 得到的.

现在在序列 $\{f_{k_p}^1(x_2)\}_{p\in\mathbb{N}}$ 中按同样方式选取收敛子列 $\{f_{k_p}^1(x_2)\}_{p\in\mathbb{N}}$, 其对应函数列 $\{f_{k_p}^2\}_{p\in\mathbb{N}}$. 如此下去, 可得函数列:

$$\begin{aligned}
&f_{k_1}^1, f_{k_2}^1, \cdots, f_{k_n}^1, \cdots \text{ s.t.} \{f_{k_n}^1(x_1)\}_{n\in\mathbb{N}} \text{收敛} \\
&f_{k_1}^2, f_{k_2}^2, \cdots, f_{k_n}^2, \cdots \text{ s.t.} \{f_{k_n}^2(x_1)\}_{n\in\mathbb{N}}, \{f_{k_n}^2(x_2)\}_{n\in\mathbb{N}} \text{收敛} \\
&\qquad\qquad\vdots \\
&f_{k_1}^q, f_{k_2}^q, \cdots, f_{k_n}^q, \cdots \text{ s.t.} \{f_{k_n}^q(x_1)\}_{n\in\mathbb{N}}, \{f_{k_n}^q(x_2)\}_{n\in\mathbb{N}}, \cdots, \{f_{k_n}^q(x_q)\}_{n\in\mathbb{N}} \text{收敛}
\end{aligned}$$

现取对角线序列 $\{g_n\}_{n\in\mathbb{N}}, g_n = f_{k_n}^n$, 知对任意的 $x_p (p \in \mathbb{N})$, $\{g_n(x_p)\}_{n\in\mathbb{N}}$ 均收敛, 也即 $\{g_n(x)\}_{n\in\mathbb{N}}$ 在 E 上收敛, 下面证明 $\{g_n(x)\}_{n\in\mathbb{N}}$ 在 K 上收敛. 取定 $\varepsilon > 0$, 根据等度连续函数族的定义对应选取 $\delta > 0$, 因为 E 在 K 中稠密, 故可取 $E_k = \{x_1, \cdots, x_k\}$ 是 K 的有限 δ 网. 根据 g_n 的构造, 知序列 $\{g_n(x_i)\}_{n\in\mathbb{N}} (k = 1, \cdots, n)$ 收敛, 根据收敛的定义知

$$\exists N > 0 \forall m > n > N \forall 1 \leq i \leq k (d_Y(g_m(x_i), g_n(x_i)) < \varepsilon)$$

现任取 $x \in K$, 根据 δ 网的定义知存在 $x_j \in K$, 使得 $d_X(x, x_j) < \delta$, 又因为 $\mathcal{F}$ 是等度连续函数族, 故 $d_Y(g_n(x), g_n(x_j)) < \varepsilon$ 对任意 $n \in \mathbb{N}$ 均成立, 故

$$\begin{aligned}
d_Y(g_m(x), g_n(x)) &\leq d_Y(g_n(x), g_n(x_j)) + d_Y(g_n(x_j), g_m(x_j)) + d_Y(g_m(x_j), g_m(x)) \\
&< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon
\end{aligned}$$

又因为 x 是紧空间 K 上的任意一点, 且前述 N 与 x 无关, 故 $\{g_n(x)\}_{n\in\mathbb{N}}$ 在 K 中一致收敛, 亦即序列 $\{g_n\}_{n\in\mathbb{N}}$ 在 $\mathcal{F}$

中按一致收敛度量收敛. □

2.3.2 一些例子

例 2.20(完备度量空间中的紧集 F_1, 闭集 F_2, 使对任意 $x \in F_1, y \in F_2$ 恒有 $d(x,y) > d(F_1,F_2)$) 根据习题2.16, 当 F_1, F_2 都是度量空间中的紧集时, 必存在 $x_0 \in F_1, y_0 \in F_2$ 使得 $d(x_0,y_0) = d(F_1,F_2)$, 但 F_1, F_2 不全为紧集时该结论不真.

对空间 l^2 取 l^2 距离 $d(x,y)$, 则 l^2 为完备度量空间. 在 l^2 中取 $F_1 = \{0\}, F_2 = \{x_n\}_{n=1}^{\infty}$, 其中

$$0 = (0,0,\cdots),\ x_n = (0,\cdots,0,\frac{n+1}{n},0,\cdots)$$

显见 F_1 是紧集, 又对任意的 $x_m, x_n \in F_2 (m \neq n)$, 有

$$d(x_m,x_n) = \sqrt{\frac{(m+1)^2}{m} + \frac{(n+1)^2}{n^2}} > \sqrt{2}$$

这说明 F_2 中无收敛列, 进而其也可视作闭集, 但对任意的 $x \in F_1, y \in F_2$, 根据定义知 $x = 0, y = x_k$, 其中 k 给定. 故

$$d(x,y) = \sqrt{(1+\frac{1}{k})^2} > 1 = d(F_1,F_2)$$

这说明不存在 $x_0 \in F_1, y_0 \in F_2$ 使得 $d(x_0,y_0) = d(F_1,F_2)$.

例 2.21(在同一个集合 X 上存在两个距离 d_1, d_2, 使得 (X,d_1) 可分, (X,d_2) 不可分) 考虑元素为全体有界数列 $\{\xi_n\}$ 的空间, 对 $x = \{\xi_n\}, y = \{\eta_n\} \in X$, 令

$$d_1(x,y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdots \frac{|\xi_n - \eta_n|}{1 + |\xi_n - \eta_n|}$$

易证 (X,d_1) 是度量空间, 再设 M 为形如 $(r_1, r_2, \cdots, r_n, 0, 0\cdots)$ 的元素 x 构成的集合, 其中 $r_i \in \mathbb{Q} (i = 1,\cdots,n)$, $n \in \mathbb{N}$, 显见 M 为可数集, 且 M 在 (X,d_1) 中稠密, 故 (X,d_1) 可分.

再考虑 l^∞ 距离 $d_2(x,y) = \sup_n |\xi_n - \eta_n|$, 显见 $(X,d_2) = l^\infty$ 同样是度量空间 (称为有界数列空间), 考虑 l^∞ 中坐标仅取 0,1 的元素的集合, 知它们中不同者之间的距离为 1, 且每个元素都能 (二进制地) 与 [0,1) 上的元素一一对应, 因而该集合不可数, 从而该集合不可分, 进而 l^∞ 也不可分.

例 2.22(完全有界但不列紧的集合) 设 X 为 $\mathbb{Q} \cap [0,1]$, 在 X 上取通常距离, 知此时 X 不完备. 任取 $\varepsilon > 0$, 取 $n \in \mathbb{N}$ 满足 $n\varepsilon > 1$, 则此时 $\{0, \frac{1}{n}, \cdots, \frac{n-1}{n}, 1\}$ 正是 X 的一个有限 ε 网, 从而 X 是完全有界的. 但 X 不列紧, 因为若取 $x_n = \frac{1}{3}(1+\frac{1}{n})^n (n \in \mathbb{N})$, 则序列 $\{x_n\}_{n\in\mathbb{N}}$ 在 X 中无极限.

注 Hausdorff 定理保证了完备度量空间内列紧与完全有界等价, 故完备度量空间中不存在上述例子.

2.3.3 习题

练习 2.13★ 在完备的度量空间中求证: 子集 A 列紧的充要条件是对 $\forall \varepsilon > 0$, 存在 A 的列紧的 ε 网.

证明

根据 Hausdorff 定理, $(\mathscr{X},\rho) \supset A$ 列紧当且仅当 A 完全有界, 此时其对任意的 $\varepsilon > 0$ 有有穷 ε 网 N_ε. 既然 N_ε 有穷, 知 $N'_\varepsilon = \varnothing \subset (\mathscr{X},\rho)$, 其自然是列紧的.

现在当 A 由列紧的 ε 网 N, 知 N 作为列紧集完全有界, 设其 ε 网为 N', 则任取 $x \in A$, 设 $y \in N$ 满足 $\rho(x,y) < \varepsilon$, 根据 N' 的定义知另存在 $y' \in N'$ 满足 $\rho(y,y') < \varepsilon$, 故 $\rho(x,y') < 2\varepsilon$, 由 ε 的任意性知可通过这种方式构造 A 的有限 ε 网, 因而 A 是列紧的.

练习 2.14♡ 在度量空间中求证: 紧集上的连续函数必是有界的, 并且达到它的上, 下确界.

证明

设紧集 A 上有连续函数 $f: A \to \mathbb{R}$, 根据连续性的定义知:

$$\forall x \in A \forall \varepsilon > 0 \exists \delta = \delta(x,\varepsilon) \forall y \in A (d(x,y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)$$

现在固定 ε, 显见 $A \subset \bigcup_{x\in A} B(x, \delta(x,\varepsilon))$, 这说明 $\bigcup_{x\in A} B(x,\varepsilon)$ 可代表 A 的一个开覆盖. 因为 A 是紧集, 故其任意开覆盖必存在有限子覆盖, 进而存在 $x_1, \cdots, x_n \in A$ 使得:

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

从而

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

现在对于 $f(B(x_i, \delta(x_i,\varepsilon)))$ 知:

$$\forall 1 \le i \le n \forall y \in B(x_i, \delta(x_i,\varepsilon))(|f(y) - f(x_i)| < \varepsilon \Rightarrow |f(y)| < |f(x_i)| + \varepsilon)$$

现取 $M = \max\{|f(x_1)| + \varepsilon, \cdots, |f(x_n)| + \varepsilon\}$, 则

$$\forall x \in A(|f(x)| < M)$$

这正是 f 在 A 上有界的定义.

再证明 f 在 A 上可达到上下确界. 设 $f(A)$ 的上确界为 u, 依照上确界的定义知:

$$\forall n \in \mathbb{N} \exists y_n \in f(A)(u - \frac{1}{n} \le y_n \le u)$$

因为 $f(A)$ 有界, 故 $\{y_n\} \subset f(A)$ 是有界序列, 从而由 $\mathbb{R}$ 的完备性知其至少在 $\mathbb{R}$ 中有收敛子列 $\{y_{n_q}\}$, 设 $y_{n_q} \to y_0 \in \mathbb{R}(q \to \infty)$. 因为 $n_q \to \infty(q \to \infty)$, 故由夹逼定理知 $y_0 = u$. 另一方面, 根据 $f(A)$ 的定义, 设对每个 $y_{n_q} \in f(A)$, 存在 $x_{n_q} \in A$ 使得 $y_{n_q} = f(x_{n_q})$. 对这样得到的 $\{x_{n_q}\} \subset A$, 因为 A 作为度量空间中的紧集必是自列紧集, 而其至少作为列紧集是有界的, 故 $\{x_{n_q}\}$ 是 A 中的有界序列, 因而由 A 的自列紧性知其具有收敛到 A 中的子列 $\{x_{n_{q_r}}\}$, 设 $x_{n_{q_r}} \to x_0 \in A(r \to \infty)$, 并设 $f(x_{n_{q_r}}) = y_{n_{q_r}} \in \{y_{n_q}\}$, 根据 f 的连续性有:

$$u = y_0 = \lim_{r\to\infty} y_{n_{q_r}} = \lim_{r\to\infty} f(x_{n_{q_r}}) = f(\lim_{r\to\infty} x_{n_{q_r}}) = f(x_0)$$

这说明 $u = f(x_0), x_0 \in A$, 故 f 可在 A 上达到上确界, 下确界类似可证. □

✍ **练习 2.15★** 在度量空间中求证: 完全有界的集合是有界的, 并通过考虑 l^2 的子集 $E = \{e_k\}_{k=1}^{\infty}$, 其中

$$e_k = \{\underbrace{0, 0, \cdots, 0, 1}_{k}, 0, \cdots\},$$

来说明一个集合可以是有界但不完全有界的.

证明

当 $A \subset (\mathscr{X}, \rho)$ 完全有界, 知任取 $\varepsilon > 0$, 总存在 A 的有限 ε 网 $\bigcup_{k=1}^{n} B(x_k, \varepsilon)$. 现取 x_1 为球心, $\max\{\rho(x_1,x_2), \cdots, \rho(x_1,x_n)\} + \varepsilon$ 为半径, 易证

$$\bigcup_{k=1}^{n} B(x_k, \varepsilon) \subset B(x_1, \max\{\rho(x_1,x_2), \cdots, \rho(x_1,x_n)\} + \varepsilon)$$

因而 A 是有界的.

另一方面, 依题意考虑 $E = \{e_k\}_{k=1}^{\infty} \subset l^2$, 显见 $E \subset B(0,2)$, 进而其为有界集. 但 E 中任意两元素之间的距离都至少大于 1, 这说明 E 中至少不存在有穷 1 网, 进而 E 不是完全有界的.

✍ **练习 2.16★** 设 $(\mathscr{X}, \rho)$ 是度量空间, F_1, F_2 是它的两个紧子集, 求证: $\exists x_i \in F_i (i = 1,2)$, 使得 $\rho(F_1,F_2) = \rho(x_1,x_2)$, 其中

$$\rho(F_1,F_2) := \inf\{\rho(x,y) : x \in F_1, y \in F_2\}$$

证明[10]

既然 $F_1, F_2 \subset (\mathscr{X}, \rho)$ 是紧集, 对它们的开覆盖作 Cartesian 积易证 $F_1 \times F_2$ 是 $(\mathscr{X} \times \mathscr{X}, \rho')$ 中的紧集, 其中 ρ' 是与 $\mathscr{X} \times \mathscr{X}$ 上乘积拓扑相容的距离. 易证 $\rho(\cdot,\cdot): \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ 是连续映射, 根据练习(2.14)知其在 $F_1 \times F_2$ 上可取到下确界, 命题即证.

练习 2.17★ 设 S 是 $C[a,b]$ 中的有界集, 求证: 集合

$$\{F(x) = \int_a^x f(t)dt : f \in S\}$$

是列紧集.

证明

要证明 $S \subset C[a,b]$ 是列紧集, 只需证明其一致有界且等度连续.

对一致有界性, 既然 $M \subset C[a,b]$ 有界, 设存在 $C > 0$ 使得:

$$\forall f \in M(\|f\| = \max_{x\in[a,b]} |f(x)| \le C)$$

从而任取 $F \in S$, 知:

$$|F(x)| = |\int_a^x f(t)dt| \le \int_a^x |f(t)|dt \le \int_a^b \max_{t\in[a,b]} |f(t)|dt \le C(b-a),\ \forall x \in [a,b]$$

这说明 $\max_{x\in[a,b]} |F(x)| = \|F\| \le C(b-a)$, 而这对 S 中的任意 F 均成立, 故 S 在 $C[a,b]$ 中有界, 亦即对 x 一致有界.

对等度连续性, 对任意的 $\varepsilon > 0$ 取 $\delta = \frac{\varepsilon}{C}$, 则对任意的 $F \in S$, 当任意 $x_1, x_2 \in [a,b]$ 满足 $|x_1 - x_2| < \delta$ 时, 有:

$$|F(x_1) - F(x_2)| \le |\int_{x_2}^{x_1} |f(t)|dt| \le C|x_1 - x_2| < \varepsilon$$

从而 S 等度连续, 因而由 Arzela-Ascoli 定理知 S 是 $C[a,b]$ 中的列紧集. □

练习 2.18 设 $E = \{\sin nt\}_{n=1}^{\infty}$, 求证: E 在 $C[0,\pi]$ 中不是列紧的.

证明

根据 Arzela-Ascoli 定理, 要证明 E 在 $C[0,\pi]$ 中不是列紧的, 只需证明 E 并非等度连续, 根据定义即证

$$\exists \varepsilon_0 > 0 \forall \delta(\varepsilon) > 0 \exists x_1, x_2 \in [0,\pi] \exists n \in \mathbb{N}(|x_1 - x_2| < \delta \wedge |\sin nx_1 - \sin nx_2| \ge \varepsilon_0)$$

注意

$$|\sin nx_1 - \sin nx_2| = 2|\cos \frac{n}{2}(x_1 + x_2)| \cdot |\sin \frac{n}{2}(x_1 - x_2)|$$

现考虑

$$\begin{cases} 0 < \alpha_1 < \frac{n}{2}(x_1 + x_2) < \alpha_2 < \frac{\pi}{2} \\ 0 < \beta_1 < \frac{n}{2}(x_1 - x_2) < \beta_2 < \frac{\pi}{2} \end{cases}$$

得到

$$x_1 \in (\frac{\alpha_1 + \beta_1}{n}, \frac{\alpha_2 + \beta_2}{n}),\ x_2 \in (\frac{\alpha_1 - \beta_2}{n}, \frac{\alpha_2 - \beta_1}{n})$$

只需

$$\frac{\alpha_2 + \beta_2}{n} > \frac{\alpha_1 - \beta_2}{n}$$

即得对任意小的 $|x_1 - x_2|$, 总可找到 $\varepsilon_0 = 2\cos\alpha_2 \sin\beta_1$, 使得

$$|\sin nx_1 - \sin nx_2| > \varepsilon_0$$

这说明 $\{\sin nt\}_{n=1}^{\infty}$ 不是等度连续的, 命题即证.

[10]该证明涉及拓扑空间的内容, 参见补充部分.

✍ **练习 2.19★** 求证: S 空间[11]的子集 A 列紧的充要条件是: $\forall n\in\mathbb{N},\exists C_n>0$, 使得对 $\forall x=(\xi_1,\xi_2,\cdots,\xi_n,\cdots)\in A$, 有 $|\xi_n|\le C_n(n=1,2,\cdots)$.

证明

若

$$\exists n\in\mathbb{N}\forall C_n>0\exists x_{C_n}=(\xi_1,\cdots,\xi_n,\cdots)\in A(|\xi_n|>C_n)$$

取用 $C_n=k$, 考虑序列 $x_k=(\xi_k^1,\cdots,\xi_k^n,\cdots)$, 知 $\{x_k\}$ 无收敛子列, 因而 A 不是列紧集.

当

$$\forall n\in\mathbb{N}\exists C_n>0\forall x=(\xi_1,\cdots,\xi_n,\cdots)\in A(|\xi_n|\le C_n, n=1,2,\cdots)$$

注意 $\xi_k\in\mathbb{R}$(或 $\mathbb{C}$), $k=1,2,\cdots$, 根据 $\mathbb{R}$(或 $\mathbb{C}$) 的完备性知: 对任意的序列 $x_p=(\xi_p^1,\cdots,\xi_p^n,\cdots)$, 其中的第一个分量 $\{\xi_p^1\}$ 必有收敛子列 $\{\xi_{p_{k_1}}^1\}$, 设 $\xi_{p_{k_1}}^1\to\xi^1$, 再在 $x_{p_{k_1}}=(\xi_{p_{k_1}}^1,\cdots,\xi_{p_{k_1}}^n,\cdots)$ 中取第二个分量, 知其亦有收敛子列, 设其收敛到 ξ^2. 这样一直下去, 可找到子列 $\{x_q\}=(\xi_q^1,\cdots,\xi_q^n,\cdots)$, 其收敛到 $x=(\xi^1,\cdots,\xi^n,\cdots)$, 命题即证.

✍ **练习 2.20★** 设 $(\mathscr{X},\rho)$ 是度量空间, M 是 $\mathscr{X}$ 中的列紧集, 映射 $f:\mathscr{X}\to M$ 满足

$$\rho(f(x_1),f(x_2))<\rho(x_1,x_2)\ (\forall x_1,x_2\in\mathscr{X},x_1\ne x_2)$$

求证: f 在 $\mathscr{X}$ 中存在唯一的不动点.

证明

根据条件显见 $f:\mathscr{X}\to M$ 是连续映射, 为了应用压缩映射原理, 考虑证明存在 $0<a<1$, 使得

$$\rho(f(x_1),f(x_2))<a\rho(x_1,x_2),\ \forall x_1,x_2\in M,x_1\ne x_2$$

如若不然, 知对任意的 $n\in\mathbb{N}$, 总存在 $x_{1n},x_{2n}\in M,x_{1n}\ne x_{2n}$, 使得

$$\rho(f(x_{1n}),f(x_{2n}))\ge(1-\frac{1}{n})\rho(x_{1n},x_{2n})$$

既然 M 是列紧集, 知 $f(x_{1n}),f(x_{2n}),x_{1n},x_{2n}$ 均有收敛子列, 进而不妨就设它们分别收敛到 $f(x_1),f(x_2),x_1,x_2\in\mathscr{X}$, 两边取极限有

$$\lim_{n\to\infty}\rho(f(x_{1n}),f(x_{2n}))=\rho(f(x_1),f(x_2))=\lim_{n\to\infty}\rho(x_{1n},x_{2n})=\rho(x_1,x_2)$$

这与条件矛盾! 故存在 $0<a<1$ 使得

$$\rho(f(x_1),f(x_2))<a\rho(x_1,x_2),\ \forall x_1,x_2\in M,x_1\ne x_2$$

根据压缩映射原理即知 f 在 M 中存在唯一的不动点. 最后对那些不在 M 中的点, 根据 $f:\mathscr{X}\to M$ 知这些点不可能成为不动点, 故 f 在 $\mathscr{X}$ 中存在唯一不动点.

✍ **练习 2.21** 设 (M,ρ) 是一个紧度量空间, 又 $E\subset C(M)$, E 中的函数一致有界并满足下列 Hölder 条件:

$$|x(t_1)-x(t_2)|\le C\rho(t_1,t_2)^\alpha\ (\forall x\in E,\forall t_1,t_2\in M)$$

其中 $0<\alpha\le 1,C>0$. 求证: E 在 $C(M)$ 中是列紧集.

证明

根据 Arzela-Ascoli 定理与 (M,ρ) 的紧性 (进而是完备性), 只需证明 E 一致有界且等度连续. 已知 E 一致有界, 根据 Hölder 条件知

$$\forall\varepsilon>0\exists\delta=(\frac{\varepsilon}{C})^{\frac{1}{\alpha}}\forall t_1,t_2\in M(\rho(t_1,t_2)<\delta\Rightarrow|x(t_1)-x(t_2)|\le C\rho(t_1,t_2)^\alpha<\varepsilon)$$

故 E 等度连续, 命题得证.

[11] 具体见练习(2.8), 其为一切实 (或复) 数列 $x=(\xi_1,\cdots,\xi_n,\cdots)$ 组成的集合, 距离为 $\rho(x,y)=\sum_{k=1}^{\infty}\frac{1}{2^k}\cdot\frac{|\xi_k-\eta_k|}{1+|\xi_k-\eta_k|}$.

2.4 赋范线性空间

2.4.1 知识梳理

度量空间讨论的是拓扑结构, 其可以给出空间中的收敛. 为了考虑元素间的代数运算, 现在开始讨论代数结构.

定义 2.4.1 (线性空间)

设 $\mathscr{X}$ 是一个非空集, $\mathbb{K}$ 是复 (或实) 数域. 如果下列条件满足, 便称 $\mathscr{X}$ 为一复 (或实) 线性空间:

1. $\mathscr{X}$ 是一加法交换群, 即对 $\forall x, y \in \mathscr{X}, \exists u \in \mathscr{X}$, 记作 $u = x + y$, 称 u 是 x, y 之和, 如果其满足
   (a). $x + y = y + x$;
   (b). $(x + y) + z = x + (y + z)$;
   (c). 存在唯一的 $\theta \in \mathscr{X}$, 对 $\forall x \in \mathscr{X}$, 有 $x + \theta = \theta + x$;
   (d). 对任意的 $x \in \mathscr{X}$, 存在唯一的 $x' \in \mathscr{X}$, 使得 $x + x' = \theta$, 记这个 x' 为 $-x$.
2. 定义了数域 $\mathbb{K}$ 中的数 α 与 $x \in \mathscr{X}$ 的数乘运算, 即 $\forall(\alpha, x) \in \mathbb{K} \times \mathscr{X}$, 存在 $u \in \mathscr{X}$, 记作 $u = \alpha x$, 称 u 为 x 对 α 的数乘, 适合
   (a). $\alpha(\beta x) = (\alpha\beta)x, \forall \alpha, \beta \in \mathbb{K}, \forall x \in \mathscr{X}$;
   (b). $1 \cdot x = x$;
   (c). $(\alpha + \beta)x = \alpha x + \beta x, \forall \alpha, \beta \in \mathbb{K}, \forall x \in \mathscr{X}, \alpha(x + y) = \alpha x + \alpha y, \forall x, y \in \mathscr{X}, \forall \alpha \in \mathbb{K}$.

♣

定义 2.4.2 (线性同构)

设 $\mathscr{X}, \mathscr{X}_1$ 都是线性空间, $T : \mathscr{X} \to \mathscr{X}_1$ 称为一个线性同构, 如果

1. 它既是单射又是满射, 即它是一对一且在上的;
2. $T(\alpha x + \beta y) = \alpha Tx + \beta Ty, \forall x, y \in \mathscr{X}, \forall \alpha, \beta \in \mathbb{K}$.

♣

定义 2.4.3 (线性子空间)

设 $E \subset \mathscr{X}$, 若 E 依 $\mathscr{X}$ 上的加法与数乘还构成一个线性空间, 则称 E 是 $\mathscr{X}$ 的一个线性子空间.

♣

注 $\mathscr{X}$ 以及 $\{\theta\}$ 都是 $\mathscr{X}$ 的线性子空间, 称它们为平凡的子空间, 而称其他的子空间为真子空间.

定义 2.4.4 (线性流形)

设 $E \subset \mathscr{X}$, 若存在 $x_0 \in \mathscr{X}$ 及线性子空间 $E_0 \subset \mathscr{X}$, 使得 $E = E_0 + x_0 := \{x + x_0 | x \in E_0\}$, 则称 E 为线性流形. 简单来说, 线性流形就是子空间对某个向量的平移.

♣

定义 2.4.5 (线性相关)

一组向量 $x_1, x_2, \cdots, x_n \in \mathscr{X}$ 称为是线性相关的, 如果存在 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{K}$ 不全为 0, 使得

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = 0;$$

否则称为是线性无关的.

♣

定义 2.4.6 (线性基)

若 A 是 $\mathscr{X}$ 中的一个极大线性无关向量组, 即 A 中的向量是线性无关的, 而且任意的 $x \in \mathscr{X}$ 都是 A 中的向量的线性组合, 则称 A 是 $\mathscr{X}$ 的一组线性基.

♣

课堂笔记 (♡)

这里引入的线性基正是 Hamel 基.

♠

补充定义 2.4.1 (Hamel 基[PGC])

设 $\mathscr{X}$ 是线性空间, 所谓 $\mathscr{X}$ 的一个 Hamel 基, 指的是任何一个由向量 $e_i \in \mathscr{X}$ 组成的向量族 $\{e_i\}_{i\in I}$, 其满足:

(i) 向量族中的元线性无关, 即对族 $\{e_i\}_{i\in I}$ 的任何有限子集 $\{e_j\}_{j\in J}$ 与任何一组数 $\alpha_j \in \mathbb{K}(j \in J)$, 当 $\sum_{j\in J}\alpha_j e_j = 0$ 时, 必有 $\alpha_j = 0(j \in J)$.

(ii) $\operatorname{span}\{e_i\}_{i\in I} = \mathscr{X}$, 也就是说对任意给定的向量 $x \in \mathscr{X}$, 存在族 $\{e_i\}_{i\in I}$ 的有限子族 $\{e_j\}_{j\in J(x)}$ 和数 $x_j \in \mathbb{K}(j \in J(x))$, 使得 $x = \sum_{j\in J(x)} x_j e_j$.

♣

[**PGC**] 给出: Hamel 基在任意线性空间中均存在, 且其基数是确定的.

补充定理 2.4.1

设 $\mathscr{X} \neq \{0\}$ 是线性空间, 则

(i) $\mathscr{X}$ 的 Hamel 基 E 存在.

(ii) 若 E, F 是 $\mathscr{X}$ 的两个 Hamel 基, 则 $\operatorname{card} E = \operatorname{card} F$.

♡

证明

考虑到 Hamel 基的本质是极大无关组, 故可以考虑用 Zorn 引理断言极大元的存在性.

(i) 用 $\mathcal{F}$ 表示 $\mathscr{X}$ 中向量的所有线性无关族组成的集合. 因为 $\{e\} \in \mathcal{F}$, 其中 $\mathscr{X} \ni e \neq 0$, 故 $\mathcal{F}$ 首先非空. 当 $\bigcup_{i\in I}\{e_i\} \subset \bigcup_{j\in J}\{e_j\}$ 时, 规定 $E = \{e_i\}_{i\in I} \preccurlyeq F = \{e_j\}_{j\in J}$, 进而在 $\mathcal{F}$ 上可以定义偏序关系 $\preccurlyeq$. 因为线性无关族的元素互不相同, 故 $E = \{e_i\}_{i\in I}$ 恒同于 $\bigcup_{i\in I}\{e_i\}$, 显见关系 $E \preccurlyeq F$ 正是通常的包含关系 $E \subset F$.

现设 $\mathcal{E}$ 是 $\mathcal{F}$ 的全序子集, 下面说明族 $G := \bigcup_{E\in\mathcal{E}} E$ 是 $\mathcal{F}$ 中的元. 因为 $\mathcal{E}$ 全序, 故 G 的任何有限子集 $\{e_i\}_{i=1}^m$ 都是某个族 $E \in \mathcal{E}$ 的子族[12], 而 $E \in \mathcal{E} \subset F$, 故 $\{e_i\}_{i=1}^m$ 线性无关, 且 $G \in \mathcal{F}$. 根据 G 的构造, 只要 $E \in \mathcal{E}$, 就必有 $E \subset G$, 这说明 G 是 $\mathcal{E}$ 在序关系 $\preccurlyeq$ 下的一个上界.

上面说明了 $\mathcal{F}$ 的每个全序子集在 $\mathcal{F}$ 中均有上界, 故由 Zorn 引理知非空偏序集 $\mathcal{F}$ 在序关系 $\preccurlyeq$ 下有极大元 M, 下面证明 M 就是 $\mathscr{X}$ 的一个 Hamel 基. 如若不然, 根据 Hamel 基的定义知必存在非零向量 $e \in \mathscr{X}$, 使得 e 不能表为 M 中元素的线性组合. 此时 $M \cup \{e\} \in \mathcal{F}$, 且 $M \prec M \cup \{e\}$(即 $M \preccurlyeq M \cup \{e\}$ 且 $M \neq M \cup \{e\}$), 这与 M 的极大性矛盾! 故 M 是 $\mathscr{X}$ 的 Hamel 基, (i) 得证.

(ii) 设 $E = \bigcup_{i\in I}\{e_i\}, F = \bigcup_{j\in J}\{f_j\}$ 分别为 $\mathscr{X}$ 的两个 Hamel 基, 特别设 E 中的每个元 e_i 可以表为 $f_j(j \in J(i))$ 的有限线性组合, 其中 $J(i)$ 是 J 的有限子集.

下面证明 $F = \bigcup_{i\in I} F_i$, 其中 $F_i := \bigcup_{k\in J(i)}$, 进一步只需证明 $F \subset \bigcup_{i\in I} F_i$. 用反证法, 设存在 $j_0 \in J$, 使得 $f_{j_0} \notin \bigcup_{i\in I} F_i$, 也即在表示 E 的全体元素时均不会用到 f_{j_0}. 因为 E 是基, 故用于表示 E 的 $F\backslash\{f_{j_0}\}$ 也是基, 但因为 $F \supset F\backslash\{f_{j_0}\}$ 也是基, 故这与前述 Hamel 基的极大性矛盾! 从而 $F = \bigcup_{i\in I} F_i$.

现在先设这两个基之一 (比如 E) 是有限集, 从而 $\bigcup_{i\in I}\{f_j\}_{j\in J(i)}$ 是有限集, 因而 F 是有限集, 此时容易构造双射说明 $\operatorname{card} E = \operatorname{card} F$.

再设 E 是无限集, 等价地有 I 是无限集. 现对每个 $i \in I$, 都存在满射 $f_i : \mathbb{N} \to F_i, i \mapsto f_i$(满射是因为根据前面的设定, F_i 是有限集), 从而映射 $(i, n) : I \times \mathbb{N} \to F = \bigcup_{i\in I} F_i, (i, n) \mapsto f_n \in F_i$ 是满射, 进而由 Schröder-Bernstein 定理知 $\operatorname{card} F \preccurlyeq \operatorname{card}(I \times \mathbb{N})$. 因为 I 是无限集, 故 $\operatorname{card}\mathbb{N} \preccurlyeq \operatorname{card} I$, 因而

$$\operatorname{card}(I \times \mathbb{N}) \preccurlyeq \operatorname{card}(I \times I) = \operatorname{card} I$$

从而 $\operatorname{card} F \preccurlyeq \operatorname{card} I = \operatorname{card} E$, 同理可证 $\operatorname{card} E \preccurlyeq \operatorname{card} F$, 故 $\operatorname{card} E = \operatorname{card} F$. □

定义 2.4.7 (维数)

线性空间中的线性基的元素个数 (势), 称为维数.

♣

[12] 全序保证的是集合中的两个元素总是具有序关系, 现在序关系对应于包含关系, 这说明 $\mathcal{E}$ 中的线性无关族 E 两两之间都有包含 (或被包含) 关系. 现在在 G 中随便取有限个 $\{e_i\}_{i=1}^m$, 这些 e_i 可能来自不同的 E, 但因为前面的全序性, 总是可以找到一个新的 E 把所有这些 e_i 都给包含住.

定义 2.4.8 (线性包)

设 Λ 是一个指标集, $\{x_\lambda|\lambda\in\Lambda\}$ 是 $\mathscr{X}$ 中的向量族, 一切由 $\{x_\lambda|\lambda\in\Lambda\}$ 的有穷线性组合组成的集合

$$\{y=\alpha_1 x_{\lambda_1}+\cdots+\alpha_n x_{\lambda_n}|\lambda_i\in\Lambda,\alpha_i\in\mathbb{K},i=1,2,\cdots,n\}$$

称为 $\{x_\lambda|\lambda\in\Lambda\}$ 的线性包. 线性包是一个线性子空间, 不难证明其为包含 $\{x_\lambda|\lambda\in\Lambda\}$ 的一切线性子空间的交, 因此称线性包为 $\{x_\lambda|\lambda\in\Lambda\}$ 张成的线性子空间, 记为 $\mathrm{span}\{x_\lambda|\lambda\in\Lambda\}$. ♣

定义 2.4.9 (线性和与直和)

设 E_1,E_2 是 $\mathscr{X}$ 的子空间, 称集合 $\{x+y|x\in E_1,y\in E_2\}$ 为 E_1,E_2 的线性和, 记为 E_1+E_2. 对于任意有限个子空间, 定义以此类推. 又若 (E_1,E_2) 中的任意一对非零向量都是线性无关的, 则称线性和 E_1+E_2 为直和, 记作 $E_1\oplus E_2$, 这时 $E_1\cap E_2=\{\theta\}$, 对 $\forall x\in E_1\oplus E_2$, 有唯一的分解:

$$x=x_1+x_2,\ x_i\in E_i,i=1,2.$$

♣

现在把前述空间 $\mathscr{X}$ 的代数结构: 线性空间与拓扑结构: 距离 ρ 结合起来导出范数的概念. 我们要求:

(1) **距离有平移不变性**

$$\rho(x+z,y+z)=\rho(x,y),\ \forall x,y,z\in\mathscr{X}$$

进而知 ρ 对加法是连续的, 即

$$(\rho(x_n,x)\to 0\wedge\rho(y_n,y)\to 0)\Rightarrow\rho(x_n+y_n,x+y)\to 0,\ n\to\infty$$

这本身是因为

$$\begin{aligned}\rho(x_n+y_n,x+y)&=\rho(x_n+y_n-x-y,\theta)=\rho(x_n-x,y-y_n)\\&\le\rho(x_n-x,\theta)+\rho(y-y_n,\theta)=\rho(x_n,x)+\rho(y,y_n)\to 0,n\to\infty\end{aligned}$$

反之, 如果距离 ρ 对加法连续, 则满足平移不变性.

(2) **数乘的连续性**

(a). $\rho(x_n,x)\to 0\Rightarrow\rho(\alpha x_n,\alpha x)\to 0,n\to\infty,\forall\alpha\in\mathbb{K}$;

(b). $\alpha_n\to\alpha\in\mathbb{K}\Rightarrow\rho(\alpha_n x,\alpha x)\to 0,n\to\infty,\forall x\in\mathscr{X}$.

若令 $p:\mathscr{X}\to\mathbb{R},p(x):=\rho(x,\theta),\forall x\in\mathscr{X}$, 则由距离的平移不变性知

$$p(x-y)=\rho(x-y,\theta)=\rho(x,y).$$

注 事实上, 从距离空间到 F^* 空间并没有这么容易, 这主要体现在上文标蓝部分实际上是不平凡的:

补充定理 2.4.2 (Klee[LPD]♡)

若拓扑线性空间 $\mathscr{X}$ 具有可数局部基, 则 $\mathscr{X}$ 上存在度量 ρ 使得:

(i) 由 ρ 诱导的拓扑与 $\mathscr{X}$ 上原有拓扑相同;

(ii) ρ 具有平移不变性;

(iii) 在度量 ρ 之下, 每个球 $B(0,r)=\{x:d(x,0)<r,r>0\}$ 都是平衡集[a].

[a] 若 $\forall\alpha\in\mathbb{K}(|\alpha|\le 1\Rightarrow\alpha W\subset W)$, 就称 $W\subset X$ 是平衡集.

♡

推论 2.4.1

拓扑线性空间可度量化当且仅当其有可数局部基. ♡

补充定义 2.4.2 (等价距离[LPD]♡)

称 $\mathscr{X}$ 上的距离 ρ,ρ' 等价, 如果对 $\mathscr{X}$ 中的每个序列 $\{x_n\}_{n\ge 1}$, 有 $\rho(x_m,x_n)\to 0\Leftrightarrow\rho'(x_m,x_n)\to 0(m,n\to\infty)$. ♣

推论 2.4.2 (Klee[LPD]♡)

若 $(\mathscr{X},\rho)$ 是线性距离空间, ρ 不是平移不变的, 则 $\mathscr{X}$ 上可以改赋一个等价距离 ρ', 使得 $(\mathscr{X},\rho')$ 成为平移不变线性距离空间. ♡

现在根据距离公理, 可以列出 p 需要的条件:

$$\rho(x,y)\geq 0 \Leftrightarrow p(x)\geq 0, \forall x,y\in\mathscr{X}$$
$$\rho(x,y)=0\text{当且仅当}x=y \Leftrightarrow p(x)=0\text{当且仅当}x=\theta;$$
$$\rho(x,y)\leq\rho(x,z)+\rho(z,y) \Leftrightarrow \rho(x+y)\leq p(x)+p(y);$$
$$\rho(x,y)=\rho(y,x) \Leftrightarrow p(-x)=p(x)$$

此外还有

$$(a) \Leftrightarrow p(\alpha x_n)\to 0, p(x_n)\to 0;$$
$$(b) \Leftrightarrow p(\alpha_n x)\to 0, \alpha_n\to 0.$$

根据 p 有下述定义:

定义 2.4.10 (准范数)

线性空间 $\mathscr{X}$ 上的准范数 (准模) 定义为这个空间上的一个函数 $\|\cdot\|:\mathscr{X}\to\mathbb{R}$, 满足条件:

1. $\|x\|\geq 0, \forall x\in\mathscr{X}; \|x\|=0\Leftrightarrow x=\theta$;
2. $\|x+y\|\leq\|x\|+\|y\|, \forall x,y\in\mathscr{X}$;
3. $\|-x\|=\|x\|, \forall x\in\mathscr{X}$;
4. $\lim_{\alpha_n\to 0}\|\alpha_n x\|=0, \lim_{\|x_n\|\to 0}\|\alpha x_n\|=0, \forall x\in\mathscr{X}, \forall\alpha\in\mathbb{K}$.

♣

定义 2.4.11 (F^* 空间)

一个赋准范数的线性空间 $\mathscr{X}$, 如果按照

$$\|x_n-x\|\to 0,\ n\to\infty$$

来定义 $x_n\to x(n\to\infty)$, 就称其为 F^* 空间. ♣

定义 2.4.12 (Frechet 空间)

完备的 F^* 空间称为 Frechet 空间, 简称 F 空间. ♣

例 2.23 对空间 $C(M)(M$ 是一个紧度量空间), 显见

$$\|u\|=\max_{x\in M}|u(x)|$$

是一个准范数, 进而 $C(M)$ 是一个 F 空间.

例 2.24 对 Euclid 空间 $\mathbb{R}^n$, 设 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$, 定义

$$\|x\|=(\sum_{i=1}^{n}|x_i|^2)^{\frac{1}{2}}$$

显见其为准范数, 进而 $\mathbb{R}^n$ 是一个 F 空间.

例 2.25 对空间 S, 其中 S 表示一切序列 $x=(x_1,x_2,\cdots,x_n)$ 组成的线性空间, 加法和数乘按自然方式定义:

$$x+y=(x_1+y_1,x_2+y_2,\cdots,x_n+y_n,\cdots),$$
$$\alpha x=(\alpha x_1,\alpha x_2,\cdots,\alpha x_n,\cdots),\alpha\in\mathbb{K}$$

其中 $x=(x_1,x_2,\cdots,x_n,\cdots), y=(y_1,y_2,\cdots,y_n,\cdots)$. 对任意的 $x\in S$ 定义其范数为

$$\|x\|=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n|}{1+|x_n|}$$

现在说明其为准范数, 显见正定性与偶性是满足的, 下面验证三角不等式, 注意到本身有初等不等式:

$$\frac{\alpha+\beta}{1+\alpha+\beta}=\frac{\alpha}{1+\alpha+\beta}+\frac{\beta}{1+\alpha+\beta}\leq\frac{\alpha}{1+\alpha}+\frac{\beta}{1+\beta}, \forall\alpha,\beta>0$$

得

$$\|x+y\|=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n+y_n|}{1+|x_n+y_n|}\leq\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n|+|y_n|}{1+|x_n|+|y_n|}\leq\sum_{n=1}^{\infty}\frac{1}{2^n}(\frac{|x_n|}{1+|x_n|}+\frac{|y_n|}{1+|y_n|})=\|x\|+\|y\|$$

最后验证连续性, 同样注意到初等不等式

$$\frac{\alpha\beta}{1+\alpha\beta}\leq\begin{cases}\frac{\alpha\beta}{1+\beta}, & \alpha\geq 1,\beta\geq 0\\ \frac{\beta}{1+\beta}, & 0<\alpha<1,\beta\geq 0\end{cases}$$

故对任意的 $\alpha\in\mathbb{K}$, 有

$$\|\alpha x_n\|\leq\max(|\alpha|,1)\|x_n\|\to 0,\ \|x_n\|\to 0$$

又如果 $|\alpha_m|\to 0$, 则对任意的 $\varepsilon>0$, 取 $n_0\in\mathbb{N}$ 满足 $\frac{1}{2^{n_0}}<\varepsilon$. 固定 n_0, 再取 $N=N(\varepsilon)$, 使得当 $m>N$ 时有

$$|\alpha_m|\max_{1\leq i\leq n_0}|x_i|<\varepsilon$$

得到:

$$\|\alpha_m x\|=\sum_{n=1}^{n_0}\frac{1}{2^n}\cdot\frac{|\alpha_m x_n|}{1+|\alpha_m x_n|}+\sum_{n=n_0+1}^{\infty}\frac{1}{2^n}\cdot\frac{|\alpha_m x_n|}{1+|\alpha_m x_n|}\leq\varepsilon\sum_{n=1}^{n_0}\frac{1}{2^n}+\sum_{n=n_0+1}^{\infty}\frac{1}{2^n}<\varepsilon+\varepsilon=2\varepsilon$$

这便证明了 S 是 F^* 空间, 最后来说明完备性. 若:

$$\|x^{(m+p)}-x^{(m)}\|=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n^{(m+p)}-x_n^{(m)}|}{1+|x_n^{(m+p)}-x_n^{(m)}|}\to 0,\ m\to\infty,\forall p\in\mathbb{N}$$

因为

$$\frac{1}{2^n}\frac{|x_n^{m+p}-x_n^{(m)}|}{1+|x_n^{(m+p)}-x_n^{(m)}|}\leq\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n^{(m+p)}-x_n^{(m)}|}{1+|x_n^{(m+p)}-x_n^{(m)}|}, \ \forall n\in\mathbb{N}$$

故对任意的 $n\in\mathbb{N}$, 有 $|x_n^{(m+p)}-x_n^{(m)}|\to 0, m\to\infty,\forall p\in\mathbb{N}$. 这说明存在 x_n^*, 使得 $x_n^{(m)}\to x_n^*, m\to\infty$, 故对任意的 $\varepsilon>0$, 取 n_0 使得 $\frac{1}{2^{n_0}}<\varepsilon$, 再取 N 使得当 $m>N$ 时有

$$|x_n^{(m)},x_n^*|<\varepsilon,\ n=1,2,\cdots,n_0$$

即得

$$\|x^{(m)}-x^*\|=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{|x_n^{(m)}-x_n^*|}{1+|x_n^{(m)}-x_n^*|}\leq\sum_{n=1}^{n_0}\frac{1}{2^n}|x_n^{(m)}-x_n^*|+\sum_{n=n_0+1}^{\infty}\frac{1}{2^n}<\varepsilon+\varepsilon=2\varepsilon, m>N$$

其中 $x^*=(x_1^*,x_2^*,\cdots,x_n^*,\cdots)$, 故 S 是一个 F 空间.

注 从上面的过程可以看出, 点列

$$x^{(m)}=(x_1^{(m)},x_2^{(m)},\cdots,x_n^{(m)},\cdots),\ m=1,2,\cdots$$

收敛到 $\theta=(0,0,\cdots,0,\cdots), m\to\infty$, 当且仅当对每个 $n\in\mathbb{N}$ 都有 $x_n^{(m)}\to 0, m\to\infty$, 这说明按 S 范数收敛与按坐标收敛等价.

例 2.26 $C(\mathbb{R}^n)$ 空间表示 $\mathbb{R}^n$ 上一切连续函数全体, 并令

$$\|u\|=\sum_{k=1}^{\infty}\frac{1}{2^k}\cdot\frac{\max\limits_{|x|\leq k}|u(x)|}{1+\max\limits_{|x|\leq k}|u(x)|}$$

其中 $u(x) \in C(\mathbb{R}^n), x = (x_1, x_2, \cdots, x_n), |x| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. 此时 $\|\cdot\|$ 是一个准范数, 且 $C(\mathbb{R}^n)$ 构成一个 F 空间.

证明

$\|\cdot\|$ 作为准范数的性质容易验证, 因而 $C(\mathbb{R}^n)$ 是 F^* 空间, 下面说明其完备. 任取基本列 $\{u_n\}_{n\in\mathbb{N}} \subset C(\mathbb{R}^n)$, 根据基本列的定义知

$$\forall \varepsilon > 0 \exists N > 0 \forall m > \nu > N(\|u_m - u_\nu\| = \sum_{k=1}^{\infty} \frac{1}{2^k} \cdot \frac{\max_{|x|\le k} |u_m(x) - u_\nu(x)|}{1 + \max_{|x|\le k} |u_m(x) - u_\nu(x)|} < \varepsilon)$$

用反证法可以证明

$$\|u_m - u_\nu\| \to 0(m, \nu \to \infty) \Leftrightarrow \forall k \in \mathbb{N}(\max_{|x|\le k} |u_m(x) - u_\nu(x)| \to 0(m, \nu \to \infty))$$

现在说明 $\{u_\nu\}$ 存在点态收敛的极限. 任取 $x_0 \in \mathbb{R}^n$, 知必存在 $K > 0$ 使得 $|x_0| \le K$. 因而由

$$|u_m(x_0) - u_\nu(x_0)| \le \max_{|x|\le K} |u_m(x) - u_\nu(x)| \to 0, \ m, \nu \to \infty \tag{2.21}$$

知 $\{u_\nu(x_0)\}_{\nu\in\mathbb{N}}$ 有极限 $u(x_0)$, 进而由 $x_0 \in \mathbb{R}^n$ 的任意性可得 $\mathbb{R}^n$ 上的函数 $u(x)$, 且在(2.21)式中先令 $m \to \infty$ 即得:

$$|u(x) - u_\nu(x)| \to 0, \ \nu \to \infty, \forall x \in \mathbb{R}^n$$

从而 $u_\nu(x)$ 点态收敛到 $u(x)$, 因而 $\|u - u_\nu\| \to 0(\nu \to \infty)$. 下面说明 $u \in C(\mathbb{R}^n)$, 将前述极限按定义写开:

$$\forall \varepsilon > 0 \exists N > 0 \forall \nu > N(\|u - u_\nu\| < \varepsilon)$$

由 $\|\cdot\|$ 与一致范数族的等价性可取

$$\forall \varepsilon > 0 \forall k \in \mathbb{N} \exists N = N(\varepsilon, k) > 0 \forall \nu > N(\max_{|x|\le k} |u(x) - u_\nu(x)| < \varepsilon)$$

根据 $u_\nu \in C(\mathbb{R}^n)$ 知

$$\forall x_0 \in \mathbb{R}^d \exists \delta = \delta(x_0, \varepsilon) > 0 \forall x \in B(x_0, \delta)(|u_\nu(x) - u_\nu(x_0)| < \varepsilon)$$

对该固定的 x_0, 必存在 $K > 0$ 使得 $\overline{B(x_0, \delta)} \subset B(0, K)$, 从而

$$\begin{aligned} |u(x) - u(x_0)| &\le |u(x) - u_\nu(x)| + |u_\nu(x) - u_\nu(x_0)| + |u_\nu(x_0) - u(x_0)| \\ &\le \max_{|x|\le K} |u(x) - u_\nu(x)| + \varepsilon + \max_{|x|\le K} |u(x) - u_\nu(x)| < 3\varepsilon \end{aligned}$$

命题即证. □

注 $C(\mathbb{R}^n)$ 的例子给出了非紧集上连续函数空间准范数的刻画.

2.4.1.1 范数与 Banach 空间

在上述诸例中, 各类准范数存在齐次性上的差异:

$$\|\alpha x\| = |\alpha| \cdot \|x\|, \ \forall \alpha \in \mathbb{K}, \forall x \in \mathscr{X}$$

具有齐次性的准范数称为范数, 有时称为模.

定义 2.4.13 (范数)

线性空间 $\mathscr{X}$ 上的范数 $\|\cdot\|$ 是一个非负值函数:$\mathscr{X} \to \mathbb{R}$, 满足:

1. $\|x\| \ge 0, \forall x \in \mathscr{X}$. $\|x\| = 0 \Leftrightarrow x = \theta$(正定性);
2. $\|x + y\| \le \|x\| + \|y\|, \forall x, y \in \mathscr{X}$(三角不等式);
3. $\|\alpha x\| = |\alpha| \cdot \|x\|, \forall \alpha \in \mathbb{K}, \forall x \in \mathscr{X}$(齐次性).

♣

显见, 范数必是准范数.

定义 2.4.14 (赋范线性空间)

当赋准范数的线性空间中的准范数是范数时, 称该空间为赋范线性空间, 或称为 B^* 空间. 完备的 B^* 空间称作 B 空间或 Banach 空间. ♣

下面给出一个 Banach 空间的例子.

例 2.27(p 次可积函数空间) 设 $(\Omega, \mathscr{B}, \mu)$ 是一个测度空间, μ 是 Ω 上的可测函数, 且 $|u(x)|^p$ 在 Ω 上可积. 这种函数 u 的全体记作 $L^p(\Omega, \mu)$, 称作 $(\Omega, \mathscr{B}, \mu)$ 上的 p 次可积函数空间. $L^p(\Omega, \mu)$ 中的加法和数乘是自然意义, 并把几乎处处 (a.e.) 相等的函数看成同一个向量. 经这样处理过的空间 $L^p(\Omega, \mu)$ 仍然是一个线性空间, 且定义

$$\|u\| = (\int_\Omega |u(x)|^p dx)^{\frac{1}{p}}$$

则 $\|\cdot\|$ 是一个范数. 这是因为正定性与齐次性显然, 而三角不等式正是 Minkowski 不等式:

$$(\int_\Omega |u(x)+v(x)|^p d\mu)^{\frac{1}{p}} \le (\int_\Omega |u(x)|^p d\mu)^{\frac{1}{p}} + (\int_\Omega |v(x)|^p d\mu)^{\frac{1}{p}}$$

又根据 Riesz-Fisher 定理[13], $L^p(\Omega, \mu)$ 还是一个 Banach 空间.

注 这个例子中有两个重要特例:

1. Ω 是 $\mathbb{R}^n$ 中的一个可测集, $d\mu$ 是普通的 Lebesgue 测度, 此时对应的空间记作 $L^p(\Omega)$.
2. $\Omega = \mathbb{N}$, 测度 μ 等分布: $\mu(\{n\}) = 1, \forall n \in \mathbb{N}$, 此时空间 $L^p(\Omega, \mu)$ 由满足 $\sum_{n=1}^\infty |u_n|^p < \infty$ 的序列 $u = \{u_n\}_{n=1}^\infty$ 组成, 其对应的空间记作 l^p, 范数为

$$\|u\| = (\sum_{n=1}^\infty |u_n|^p)^{\frac{1}{p}}$$

例 2.28(本性有界可测函数空间) 设 $(\Omega, \mathscr{B}, \mu)$ 是一个测度空间, μ 对于 Ω 是 $\sigma-$ 有限的, $u(x)$ 是 Ω 上的可测函数. 如果 $u(x)$ 与 Ω 上的一个有界函数几乎处处相等, 则称 $u(x)$ 是 Ω 上的一个本性有界可测函数. Ω 上的一切本性有界可测函数 (把 a.e. 相等的两个函数视作同一个向量) 的全体记作 $L^\infty(\Omega, \mu)$, 在其上规定

$$\|u\| = \inf_{\substack{\mu(E_0)=0 \\ E_0 \subset \Omega}} (\sup_{x \in \Omega \setminus E_0} |u(x)|)$$

该式右端有时也记作 $\operatorname{ess\,sup}_{x\in\Omega} |u(x)|$ 或 $\operatorname{l.u.b}_{x\in\Omega} |u(x)|$. 显见 $L^\infty(\Omega, \mu)$ 首先是线性空间, 下面说明 $\|\cdot\|$ 是范数. 注意 $\|u\| \ge 0$ 和齐次性都显然, 下面逐一验证:

1. $\|u\| = 0 \Leftrightarrow u = \theta$. 只需证明必要性. 若 $\|u\| = 0$, 则根据下确界的定义知

$$\forall n \in \mathbb{N} \exists E_n \subset \Omega (\mu(E_n) = 0 \wedge \sup_{x \in \Omega \setminus E_n} |u(x)| < \frac{1}{n})$$

令

$$\Omega_1 := \bigcap_{n=1}^\infty (\Omega \setminus E_n) = \Omega \setminus \bigcup_{n=1}^\infty E_n$$

因为 $u(x) = 0, x \in \Omega_1, \mu(\bigcup_{n=1}^\infty E_n) = 0$, 故

$$u(x) = 0, \text{a.e. } x \in \Omega, \text{ 即} u = \theta$$

2. $\|u+v\| \le \|u\| + \|v\|$. 根据 $\|\cdot\|$ 的下确界定义知

$$\forall \varepsilon > -\exists E_0, E_1 \subset \Omega (\mu(E_0) = \mu(E_1) = 0 \wedge \sup_{x \in \Omega \setminus E_0} |u(x)| \le \|u\| + \varepsilon \wedge \sup_{x \in \Omega \setminus E_1} |v(x)| \le \|v\| + \varepsilon)$$

进而

$$\|u+v\| \le \sup_{x \in \Omega \setminus (E_0 \cup E_1)} |u(x)+v(x)| \le \sup_{x \in \Omega \setminus E_0} |u(x)| + \sup_{x \in \Omega \setminus E_1} |v(x)| \le \|u\| + \varepsilon + \|v\| + \varepsilon$$

[13]空间 $L^p(\Omega, \mu)(1 \le p \le \infty)$ 是完备的.

由 ε 的任意性即得结论.

最后说明 $L^\infty(\Omega,\mu)$ 完备, 设

$$\|u_{n+p}-u_n\|\to 0,\ n\to\infty,\forall p\in\mathbb{N}$$

根据 $L^\infty(\Omega,\mu)$ 中范数作为下确界定义:

$$\exists Z_{n,p}\subset\Omega\forall x\in\Omega\backslash Z_{n,p}(\mu(Z_{n,p})=0\wedge|u_{n+p}(x)-u_n(x)|\le\|u_{n+p}-u_n\|+\frac{1}{2^{n+p}})$$

令 $Z=\bigcup_{n,p}Z_{n,p}$, 则 $\mu(Z)=0$, 且

$$\forall x\in\Omega\backslash Z\forall n,p\in\mathbb{N}(|u_{n+p}(x)-u_n(x)|\le\|u_{n+p}-u_n\|+\frac{1}{2^{n+p}})$$

进而从极限的定义知

$$\forall\varepsilon>0\exists N\in\mathbb{N}\forall n\ge N,p\in\mathbb{N}(|u_{n+p}(x)-u_n(x)|<\varepsilon+\frac{1}{2^{n+p}},x\in\Omega\backslash Z).$$

这说明对任意的 $x\in\Omega\backslash Z$, $\lim_{n\to\infty}u_n(x)=u(x)$ 存在. 再在上式中令 $p\to\infty$ 得到

$$|u(x)-u_n(x)|\le\varepsilon,\ x\in\Omega\backslash Z$$

根据上确界性质知

$$\sup_{x\in\Omega\backslash Z}|u(x)-u_n(x)|\le\varepsilon$$

因为 $\mu(Z)=0$, 故 $u\in L^\infty(\Omega,\mu)$, 且

$$\|u_n-u\|\le\varepsilon,\ n\ge N$$

也即

$$\|u_n-u\|\to 0,\ n\to\infty.$$

注 当 Ω 是 $\mathbb{R}^n$ 中的一个可测集时, 对应的空间记作 $L^\infty(\Omega)$; 当 $\Omega=\mathbb{N}$ 时, 对应的空间记作 l^∞, 其为一切有界序列 $u=\{u_n\}_{n=1}^\infty$ 组成的空间, 范数为 $\|u\|=\sup\limits_{n\ge1}|u_n|$.

例 2.29(k 阶连续可微函数空间) 设 Ω 是 $\mathbb{R}^n$ 中的一个有界连通开区域, $k\in\mathbb{N}$, 用 $C^k(\overline{\Omega})$ 表示在 $\overline{\Omega}$ 上具有直到 k 阶连续偏导数的函数 $u(x)=u(x_1,x_2,\cdots,x_n)$ 的全体, 加法与数乘取自然定义, 再规定范数为

$$\|u\|=\max_{|\alpha|\le k}\max_{x\in\overline{\Omega}}|\partial^\alpha u(x)|$$

其中 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n),|\alpha|=\alpha_1+\alpha_2+\cdots+\alpha_n$, 以及

$$\partial^\alpha u(x)=\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}}u(x)$$

容易验证这样定义的 $\|\cdot\|$ 是范数, 进而 $C^k(\overline{\Omega})$ 是赋范线性空间.

再说明其完备, 事实上, 若 $\{u_n\}$ 是一个基本列, 根据定义知

$$\forall\varepsilon>0\exists N>0\forall n,m>N(\|u_n-u_m\|=\max_{|\alpha|\le k}\max_{x\in\overline{\Omega}}|\partial^\alpha u_n(x)-\partial^\alpha u_m(x)|<\varepsilon)$$

这说明全体 $\partial^\alpha u$ 按范数 $\|u\|=\max\limits_{x\in\overline{\Omega}}|u(x)|$ 在 $C(\overline{\Omega})$ 中是基本列. 根据 $C(\overline{\Omega})$ 的完备性知存在连续函数 v_α, 使得[14]

$$\partial^\alpha u_n(x)\rightrightarrows v_\alpha(x),\ n\to\infty,|\alpha|\le k,\text{关于 } x\in\overline{\Omega}\text{ 一致}$$

只需证明 $v_\alpha=\partial^\alpha v_\theta$ 即可, 先证

$$v_{(1,0,\cdots,0)}=\frac{\partial}{\partial x_1}v_{(0,0,\cdots,0)}$$

这是因为

$$\frac{\partial}{\partial x_1}u_n(x)\rightrightarrows v_{(1,0,\cdots,0)}(x),u_n(x)\rightrightarrows v_{(0,0,\cdots,0)}(x),\ n\to\infty,\text{对 } x\in\overline{\Omega}\text{ 一致}$$

[14] 一致收敛是因为该范数诱导的是一致收敛性度量.

且

$$u_n(x)=\int_{x_1^0}^{x_1}\frac{\partial}{\partial\xi}u_n(\xi,x_2,\cdots,x_n)d\xi+u_n(x_1^0,x_2,\cdots,x_n)$$

两边令 $n\to\infty$ 有

$$v_{(0,0,\cdots,0)}(x)=\int_{x_1^0}^{x_1}v_{(1,0,\cdots,0)}(\xi,x_2,\cdots,x_n)d\xi+v_{(0,0,\cdots,0)}(x_1^0,x_2,\cdots,x_n)$$

两边对 x_1 求偏导即得 $v_{(1,0,\cdots,0)}=\frac{\partial}{\partial x_1}v_{(0,0,\cdots,0)}$, 同理知

$$v_{(\underbrace{0,0,\cdots,0,1}_{j},0,\cdots,0)}=\frac{\partial}{\partial x_j}v_{(0,0,\cdots,0)},\ j=2,3,\cdots,n$$

剩下的部分用归纳法即得.

例 2.30(Sobolev 空间) 设 Ω 是 $\mathbb{R}^n$ 中的一个有界连通开区域, m 是一个非负整数, $1\le p<\infty$, 对于 $C^m(\overline{\Omega})$ 中的任意 u, 定义范数

$$\|u\|_{m,p}=(\sum_{|\alpha|\le m}\int_\Omega|\partial^\alpha u(x)|^p dx)^{\frac{1}{p}}$$

可以验证 $\|\cdot\|_{m,p}$ 是范数, 但 $C^m(\overline{\Omega})$ 依 $\|\cdot\|_{m,p}$ 并不完备, 进而考虑把它完备化, 即确定一个完备的赋范线性空间 S, 使得 $C^m(\overline{\Omega})$ 可以连续地嵌入 S 成为其稠密的子空间. 现在将 $C^m(\Omega)$ 的子集

$$S:=\{u\in C^m(\Omega)|\|u\|_{m,p}<\infty\}$$

按照上述范数完备化, 得到的完备化空间称为 Sobolev 空间, 记作 $H^{m,p}(\Omega)$. 特别当 $p=2$ 时, $H^{m,2}(\Omega)$ 简记作 $H^m(\Omega)$.

补充定义 2.4.3 (级数依范数收敛, 依范数绝对收敛 ♡)

设 $\{x_n\}\subset(\mathscr{X},\|\cdot\|)$, $S_n=x_1+\cdots+x_n$. 若存在 $S\in\mathscr{X}$ 使得 $\lim_{n\to\infty}\|S_n-S\|=0$, 就称 $\mathscr{X}$ 内的无穷级数 $\sum_{n=1}^\infty x_n$ 依范数收敛, $S=\sum_{n=1}^\infty x_n$ 为级数和. 若 $\sum_{n=1}^\infty\|x_n\|<\infty$, 则称 $\sum_{n=1}^\infty x_n$ 依范数绝对收敛. ♣

补充命题 2.4.1 (♡)

$(\mathscr{X},\|\cdot\|)$ 是 Banach 空间当且仅当对其中的任意序列, 由依范数绝对收敛可推知依范数收敛. ♠

补充定义 2.4.4 (Schauder 基 ♡)

给定 Banach 空间 $(\mathscr{X},\|\cdot\|)$, 可列集 $\{e_n\}\subset\mathscr{X}$. 若 $\forall x\in\mathscr{X}$, 均有唯一确定的数列 $\{\alpha_n\}\subset K$ 使得

$$\lim_{n\to\infty}\|x-\sum_{k=1}^\infty\alpha_k e_k\|=0$$

则称 $\{e_n\}$ 为 $\mathscr{X}$ 的 Schauder 基, 并记 $x=\sum_{k=1}^\infty\alpha_k e_k$. ♣

课堂笔记 (♡)

一些 Schauder 基的例子.

- l^2 中有 Schauder 基. 设 $e_n=(0,\cdots,0,1,0,\cdots)$, 则 $\{e_n\}$ 为 l^2 中的 Schauder 基. 这是因为对任意 $x=(x^1,\cdots,x^n,\cdots)\in l^2$, 知

$$\|x-\sum_{k=1}^n x^k e_k\|=(\sum_{k=n+1}^\infty|x_k|^2)^{\frac{1}{2}}\to0,\ n\to\infty$$

故 $x=\sum_{n=1}^\infty x^n e_n$, 亦即 $\{e_n\}$ 为 l^2 的一个 Schauder 基.

- l^∞ 中 $\{e_n\}$ 不再是 Schauder 基, 这是因为取 $\mathbf{1}=(1,\cdots,1,\cdots)\in l^\infty$, 知

$$\|\mathbf{1}-\sum_{k=1}^{n}e_k\|=1\neq 0,\ n\to\infty$$

这说明 $\mathbf{1}\neq\sum_{n=1}^{\infty}e_n$, 亦即 $\sum_{n=1}^{\infty}e_n$ 依 l^∞ 范数是不收敛的.

- 取 c_0 是 l^∞ 中收敛到零的序列的全体, 亦即 $c_0=\{(x_1,\cdots,x_n,\cdots):\lim_{n\to\infty}x_n=0\}$, 取 c_0 范数为 l^∞ 范数, 则 c_0 以 $\{e_n\}$ 为 Schauder 基.
- 取 c 为 l^∞ 中一切收敛序列的全体, 亦即 $c=\{(x_1,\cdots,x_n,\cdots):\lim_{n\to\infty}x_n\exists\}$, 则 c 的一个 Schauder 基为 $\{e_n\}_{n=1}^{\infty}\cup\{e_0\}$, 其中 $e_0=(1,\cdots,1,\cdots)$. 这是因为任取 $x=(x_1,\cdots,x_n,\cdots)\in c$, 设 $x_n\to x_0(n\to\infty)$, 则 $x-x_0e_0\in c_0$. 因为 $\{e_n\}_{n=1}^{\infty}$ 是 c_0 的 Schauder 基, 故存在 $x_k(k=1,2,\cdots)$ 使得 $x-x_0e_0=\sum_{n=1}^{\infty}x_ne_n$, 因而 $x=\sum_{n=0}^{\infty}x_ne_n$.

补充命题 2.4.2 (♡)

有 Schauder 基的 Banach 空间是可分的, 但可分 Banach 空间未必有 Schauder 基.

2.4.1.2 赋范线性空间上的范数等价

定义 2.4.15 (范数的强弱与等价)

设在线性空间 $\mathscr{X}$ 上给定了两个范数 $\|\cdot\|_1$ 与 $\|\cdot\|_2$, 我们说 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 是指

$$\|x_n\|_2\to 0\Rightarrow\|x_n\|_1\to 0,\ n\to\infty$$

如果 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 且 $\|\cdot\|_1$ 比 $\|\cdot\|_2$ 强, 则称 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价.

命题 2.4.1

要使得 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 必须且仅须存在常数 $C>0$, 使得

$$\|x\|_1\leq C\|x\|_2,\ \forall x\in\mathscr{X}$$

证明

充分性显见, 下证必要性, 用反证法. 如果 $\|x\|_1\leq C\|x\|_2$ 不成立, 这说明

$$\forall n\in\mathbb{N}\exists x_n\in\mathscr{X}(\|x\|_1>n\|x_n\|_2)$$

令[15] $y_n:=\frac{x_n}{\|x_n\|_1}$, 一方面知道 $\|y_n\|_1=1$, 另一方面注意到

$$0\leq\|y_n\|_2<\frac{1}{n},\ \forall n\in\mathbb{N}$$

这说明 $\|y_n\|_2\to 0, n\to\infty$, 这与 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强的条件矛盾! □

课堂笔记 (♡)

范数等价定义中的" 等价" 本质上是拓扑的等价. 在 **[ZL]** 的脉络中, 是先定义了极限, 随后依照极限定义了闭集. 从而如果两者的极限行为一致, 就说明它们所定义的闭集是一致的, 因而它们的拓扑结构等价.

推论 2.4.3

要使得 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价, 必须且只需存在常数 $C_1,C_2>0$, 使得

$$C_1\|x\|_1\leq\|x\|_2\leq C_2\|x\|_1,\ \forall x\in\mathscr{X}$$

[15] 可以额外留意这里对 y_n 做的处理, 这是一个非常常用的技巧.

设 $\mathscr{X}$ 是一个赋范线性空间, $\dim \mathscr{X} = n$, 此时 $\mathscr{X}$ 存在一组基 $e_1, e_2, \cdots, e_n$, 任意一个元素 $x \in \mathscr{X}$ 有唯一表示:

$$x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$$

代数同构意义下, 这说明两个有穷维线性空间等价的充要条件是它们有相同的维数, 现在如果两个有穷维赋范线性空间的维数相同, 想讨论它们的拓扑之间的关系. 知道每个 $x \in \mathscr{X}$ 唯一对应 $\mathbb{K}^n$ 空间中的一个点 $\xi = Tx := (\xi_1, \xi_2, \cdots, \xi_n)$. 现在希望建立 x 在 $\mathscr{X}$ 中的范数 $\|x\|$ 与 Tx 在 $\mathbb{K}^n$ 中的范数 (欧氏范数)

$$|Tx| = |\xi| := (\sum_{j=1}^{n} |\xi_j|^2)^{\frac{1}{2}}$$

之间的关系. 为此考察函数 $p(\xi) := \|\sum_{j=1}^{n} \xi_j e_j\|, \forall \xi \in \mathbb{K}^n$.

首先, p 对 ξ 一致连续, 这是因为对任意的 $\xi = (\xi_1, \xi_2, \cdots, \xi_n), \eta = (\eta_1, \eta_2, \cdots, \eta_n) \in \mathbb{K}^n$, 根据三角形不等式与 Schwarz 不等式:

$$|p(\xi) - p(\eta)| \leq p(\xi - \eta) \leq \sum_{i=1}^{n} |\xi_i - \eta_i| \|e_i\| \leq (\sum_{i=1}^{n} |\xi_i - \eta_i|^2)^{\frac{1}{2}} (\sum_{i=1}^{n} \|e_i\|^2)^{\frac{1}{2}} \leq |\xi - \eta| (\sum_{i=1}^{n} \|e_i\|^2)^{\frac{1}{2}}$$

其次, 根据范数的齐次性, 对任意的 $\xi \in \mathbb{K}^n \backslash \{\theta\}$, 有:

$$p(\xi) = |\xi| \| \sum_{j=1}^{n} \frac{|\xi_j|}{|\xi|} e_j \| = |\xi| p(\frac{\xi}{|\xi|})$$

因为 $\mathbb{K}^n$ 上的单位球面 $\mathbb{S}_1 := \{\xi \in \mathbb{K}^n | |\xi| = 1\}$ 是一个紧集, 故 $p(\xi)$ 在 $\mathbb{S}_1$ 上必有非负的最小值 C_1 与最大值 C_2, 即

$$C_1 \leq p(\xi) \leq C_2, \ \forall \xi \in \mathbb{S}_1$$

代入前式有

$$C_1 |\xi| \leq p(\xi) \leq C_2 |\xi|, \ \forall \xi \in \mathbb{K}^n$$

前面说 C_1 非负, 现在证明 $C_1 > 0$, 用反证法. 如若 $C_1 = 0$, 则因为其为最小值, 知存在 $\xi^* \in \mathbb{S}_1$ 满足 $p(\xi^*) = 0$. 设 $\xi^* = (\xi_1^*, \xi_2^*, \cdots, \xi_n^*)$, 此即

$$\xi_1^* e_1 + \xi_2^* e_2 + \cdots + \xi_n^* e_n = 0$$

因为本身 $\{e_1, e_2, \cdots, e_n\}$ 是基, 故上式蕴含 $\xi^* = \theta$, 但这与 $\xi^* \in \mathbb{S}_1$ 矛盾!

现在知道 $|\xi| = |Tx|, p(\xi) = \|x\|$, 得到

$$C_1 |Tx| \leq \|x\| \leq C_2 |Tx|, \ \forall x \in \mathscr{X}$$

如果把 $|Tx|$ 看作是在 $\mathscr{X}$ 空间中引入的另一个范数 $\|x\|_T$, 也即 $\|x\|_T := |Tx| (\forall x \in \mathscr{X})$, 则上式表明 $\|\cdot\|$ 与 $\|\cdot\|_T$ 等价. 以后简称 $\|\cdot\|_T$ 为 $\mathscr{X}$ 的 $\mathbb{K}^n$ 范数, 进而这说明 n 维赋范线性空间的范数与其 $\mathbb{K}^n$ 范数等价.

定理 2.4.1

设 $\mathscr{X}$ 是有穷维线性空间, 若 $\|\cdot\|_1, \|\cdot\|_2$ 都是 $\mathscr{X}$ 上的范数, 则必有正常数 C_1, C_2 使得

$$C_1 \|x\|_1 \leq \|x\|_2 \leq C_2 \|x\|_1, \ \forall x \in \mathscr{X}$$

♡

证明

设 $\dim \mathscr{X} = n$, 因为 $\|\cdot\|_1, \|\cdot\|_2$ 都与 $\mathbb{K}^n$ 范数等价, 故它们之间也等价. □

课堂笔记 (上述定理的动机整理 ♡)

对于给定的 n 维空间 $(\mathscr{X}, \|\cdot\|)$, 任取 $x \in \mathscr{X}$, 设

$$x = \xi_1 e_1 + \cdots + \xi_n e_n, \ (\xi_1, \cdots, \xi_n) \in \mathbb{K}^n$$

定义坐标映射:

$$T : \mathscr{X} \to \mathbb{K}^n, \ x = \xi_1 e_1 + \cdots + \xi_n e_n \mapsto (\xi_1, \cdots, \xi_n) = \xi$$

如果只从拓扑上考虑, 想要说明 $\mathscr{X}$ 与 $\mathbb{K}^n$ 是同胚的, 就希望说明 T 是连续单射. 对于连续性, 需要证明的是 T 是有界算子, 亦即存在常数 $C_1 > 0$ 使得:

$$\|Tx\|_{\mathbb{K}^n} \le C_1\|x\|, \ \forall x \in \mathscr{X}$$

对于单射, 只需说明 T 把非 0 映成非 0, 亦即存在常数 $C_2 > 0$ 使得

$$\|Tx\|_{\mathbb{K}^n} \ge C_2\|x\|, \ \forall x \in \mathscr{X}\backslash\{0\}$$

要证明上述两件事, 实际上就是说明 T 在 $\mathscr{X}$ 的单位球面上取到非负最值, 这就希望 $\mathscr{X}$ 的单位球面是紧集, 但现在并不知道这件事, 而只知道 $\mathbb{K}^n$ 的单位球面确为紧集. 所以需要重新构造映射:

$$p : \mathbb{K}^n \to \mathscr{X}, \ \xi = (\xi_1, \cdots, \xi_n) \mapsto \xi_1 e_1 + \cdots + \xi_n e_n = x$$

需要证明的事是同样的, 即存在 $C_1, C_2 > 0$ 使得

$$C_1\|\xi\|_{\mathbb{K}^n} \le \|p\xi\|_{\mathscr{X}} \le C_2\|\xi\|_{\mathbb{K}^n}$$

现在说明 p 在 $\mathbb{K}^n$ 的单位球面上取到正最值. 已经知道 $\mathbb{K}^n$ 的单位球面是紧集, 因而只需说明 p 的连续性, 任取 $\xi = (\xi_1, \cdots, \xi_n), \eta = (\eta_1, \cdots, \eta_n) \in \mathbb{K}^n$, 有:

$$\begin{aligned}
\|p\xi - p\eta\| &\le \|\sum_{i=1}^{n} |\xi_i - \eta_i| e_i\| \\
&\le \sum_{i=1}^{n} |\xi_i - \eta_i| \|e_i\| \\
&\le (\sum_{i=1}^{n} |\xi_i - \eta_i|^2)^{\frac{1}{2}} (\sum_{i=1}^{n} \|e_i\|^2)^{\frac{1}{2}} \\
&= \|\xi - \eta\|_{\mathbb{K}^n} \cdot (\sum_{i=1}^{n} \|e_i\|^2)^{\frac{1}{2}}
\end{aligned}$$

记 $C = (\sum_{i=1}^{n} \|e_i\|^2)^{\frac{1}{2}} < \infty$ 即得 p 的连续性, 因而 p 在 $\mathbb{K}^n$ 的单位球面上取到最值. 非负性是因为 $\mathbb{K}^n$ 是连通的, 根据 p 的连续性可得介值原理, 利用反证可知 p 不可能有非正最小值, 欲证进而成立.

注 本定理表明: 具有相同维数的两个有穷维赋范线性空间在代数上同构, 在拓扑上同胚.

要进一步说明拓扑等价 (亦即同胚), 首先说明同胚的定义:

补充定义 2.4.5 (同胚[You])

若 $f : X \to Y$ 是一一对应, 并且 f 与其逆 $f^{-1} : Y \to X$ 均连续, 就称 f 是一个同胚映射, 或称拓扑变换, 简称为同胚. 若存在 X 到 Y 的同胚映射, 就称 X 与 Y 同胚.

现在单独来证明上述断言.

证明 ($\star$)

$$\begin{tikzcd}
x \in \mathscr{X}_1 \arrow[r, "S"] \arrow[d, "T_1"'] & x \in \mathscr{X}_2 \arrow[d, "T_2"] \\
\mathbb{K}^n \arrow[r, equal] & \mathbb{K}^n
\end{tikzcd}$$

图 2.1: S 与 T_j 关系示意图

设 $\dim \mathscr{X}_1 = \dim \mathscr{X}_2 = n < \infty$, 记 $\{e_i^j\}_{i=1}^n$ 分别为 $\mathscr{X}_j$ 的基 $(j = 1, 2)$. 若令 $Se_i^{(1)} = e_i^{(2)} (i = 1, \cdots, n)$, 作线性扩张

有:

$$S:\mathscr{X}_1\to\mathscr{X}_2,\ x=\sum_{i=1}^n\xi_ie_i^{(1)}\mapsto\sum_{i=1}^n\xi_ie_i^{(2)}=:S(x)$$

易证 $S:\mathscr{X}_1\to\mathscr{X}_2$ 是代数同构, 下面再验证 $S:\mathscr{X}_1\to\mathscr{X}_2$ 是同胚. 记

$$T_j:\mathscr{X}_i\to\mathbb{K}^N,x=\sum_{i=1}^n\xi_ie_i^{(j)}\mapsto(\xi_1,\cdots,\xi_n),\ j=1,2$$

定义

$$\|x\|_{T_1}:=\|T_1x\|_{\mathbb{K}^n}=\|T_2(S(x))\|_{\mathbb{K}^n}=:\|S(x)\|_{T_2}$$

已经知道 $\|\cdot\|_{\mathscr{X}_j}$ 与 $\|\cdot\|_{T_j}$ 等价 $(j=1,2)$, 又因为 $\|x\|_{T_1}=\|T_1x\|_{\mathbb{K}^n}=\|T_2S(x)\|_{\mathbb{K}^n}=\|S(x)\|_{T_2}$, 因而 S 是连续映射. 同理可证 $\|S^{-1}(x)\|_{T_1}=\|x\|_{T_2}$, 进一步由 $\|\cdot\|_{\mathscr{X}_j}$ 与 $\|\cdot\|_{T_j}$ 的等价性可知 S^{-1} 也连续, 故 S 为拓扑同胚.

推论 2.4.4

有穷维 B^* 空间必为 B 空间.

推论 2.4.5

B^* 空间上的任意有穷维子空间必是闭子空间.

例 2.31★(B^* 空间的无穷维子空间不一定闭) 对任意 $n\in\mathbb{N}$, 令

$$e_n:=(\underbrace{0,\cdots,0}_{n},1,0,\cdots)$$

知 $e_n\in l^2$, 令 $E:=\operatorname{span}\{e_n\}_{n\in\mathbb{N}}$, 则 $E\subset l^2$ 是赋范线性空间, 但其并非闭集, 这是因为取 $x_n:=(1,\cdots,\frac{1}{n},0,\cdots)$, 则 $x_n\in E$, 但 $x_n\to(1,\cdots,\frac{1}{n},\cdots)\notin E$, 故 E 不闭.

推论(2.4.4),(2.4.5)可以进一步衍生得到下述推论.

推论 2.4.6

在有限维 B^* 空间 $\mathscr{X}$ 中, 集合 M 完全有界 $\Leftrightarrow$ 集合 M 有界.

证明

已经知道完全有界必定有界. 当 M 有界时, 根据推论(2.4.4)知 $\mathscr{X}$ 完备, 因而 M 必是列紧集, 从而由 Hausdorff 定理即得命题. □

推论 2.4.7

B^* 空间中, 紧集一定是完全有界闭集, 反之不一定成立.

证明

设 $\mathscr{X}$ 是 B^* 空间, 由定理2.3.3知 $A\subset\mathscr{X}$ 紧当且仅当 A 自列紧, 因而 A 一定是闭集. 又由 Hausdorff 定理, 从 A 的列紧性可导出其完全有界性, 故 A 必是完全有界闭集.

反之, 设 $\mathscr{X}:=(C[a,b],\|\cdot\|_{L^1})$, 记

$$x_n=\begin{cases}1-nx, & x\in[0,\frac{1}{n}]\\ 0, & x\in(\frac{1}{n},1]\end{cases}$$

显见 $\{x_n\}_{n\in\mathbb{N}}$ 依 $\|\cdot\|_{L^1}$ 有极限 $x\notin C[a,b]$, 故 $\{x_n\}_{n\in\mathbb{N}}$ 为 $L^1[a,b]$ 中的列紧集, 进而由 Hausdorff 定理知其为 $L^1[a,b]$ 中的完全有界集. 根据定义知

$$\forall\mathcal{E}>0\exists N\in\mathbb{N}\exists\{f_k\}_{k=1}^N\subset\{x_n\}_{n\in\mathbb{N}}(\{x_n\}_{n\in\mathbb{N}}\subset\bigcup_{k=1}^N B_{L^1[a,b]}(f_k,\varepsilon))$$

其中 $B_{L^1[a,b]}(f_k,\varepsilon) := \{x \in L^1[a,b] : \|x - f_k\|_{L^1} < \varepsilon\}$. 从而必有

$$\{x_n\}_{n\in\mathbb{N}} \subset \bigcup_{k=1}^{N} B_{C[a,b]}(f_k,\varepsilon) := \{x \in C[a,b] : \|x - f_k\|_{L^1} < \varepsilon\}$$

这说明 $\{x_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}$ 中的完全有界集, 又因为 $\{x_n\}_{n\in\mathbb{N}}$ 在 $\mathscr{X}$ 内的导集为 $\varnothing$, 故 $\{x_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}$ 中的闭集, 因而是完全有界闭集, 但 $\{x_n\}_{n\in\mathbb{N}}$ 显然不是紧集.

推论 2.4.8

B 空间的子集为紧集当且仅当其为完全有界闭集.

证明

已经知道紧集必是完全有界闭集, 从而只需证明完全有界闭集为紧集. 知 B 空间中的完全有界闭集必定自列紧, 而自列紧集必为紧集, 故完全有界闭集必为紧集. □

定义 2.4.16 (次线性泛函)

设 $P : \mathscr{X} \to \mathbb{R}$ 是线性空间 $\mathscr{X}$ 上的一个函数, 若其满足

1. $P(x+y) \le P(x) + P(y), \forall x, y \in \mathscr{X}$(次可加性),
2. $P(\lambda x) = \lambda P(x), \forall \lambda > 0, \forall x \in \mathscr{X}$(正齐次性)

则称 P 为 $\mathscr{X}$ 上的一个次线性泛函.

注 如果 P 还满足 $P(x) \ge 0 (\forall x \in \mathscr{X})$, 且代替条件 2 的是齐次性: $P(\alpha x) = |\alpha| P(x), \forall \alpha \in \mathbb{K}, \forall x \in \mathscr{X}$, 则称 P 是一个半范数或半模.

定理 2.4.2 (次线性泛函被范数控制的条件)

设 P 是有穷维 B^* 空间 $\mathscr{X}$ 上的一个次线性泛函, 如果 $P(x) \ge 0 (\forall x \in \mathscr{X})$, 且 $P(x) = 0 \Leftrightarrow x = \theta$, 则存在正常数 C_1, C_2, 使得

$$C_1\|x\| \le P(x) \le C_2\|x\|, \ \forall x \in \mathscr{X}.$$

证明 (★)

设 $\mathscr{X}$ 有一组基 $\{e_1, \cdots, e_n\}$, 定义

$$T : \mathscr{X} \to \mathbb{K}^n, x = \sum_{i=1}^{n} \xi_i e_i \mapsto Tx = \xi = (\xi_1, \cdots, \xi_n)$$

定义 $\|x\|_T = |Tx| = (\sum_{i=1}^{n} |\xi_j|^2)^{\frac{1}{2}}$. 根据定理2.4.1的笔记部分, 已经知道存在常数 $\widetilde{C}_1 > 0, \widetilde{C}_2 > 0$ 使得对任意 $x \in \mathscr{X}$ 均有:

$$\widetilde{C}_1\|x\|_T \le \|x\| \le \widetilde{C}_2\|x\|_T$$

此即证明

$$\widetilde{C}_1 \le \frac{\|x\|}{\|x\|_T} \le \widetilde{C}_2$$

亦即证明 P 在 $\{x \in \mathscr{X} : \|x\|_T = 1\}$ 上有界. 如若 $\mathbb{K} = \mathbb{R}$, 设 $M := \max\{P(\pm e_j) : j = 1, \cdots, n\}$, 则至少 $M < \infty$. 进一步因为 $P(x) = 0 \Leftrightarrow x = \theta$, 知 $M > 0$, 否则 $P(x) \equiv 0$, 得到 $\forall x \in \mathscr{X} (x = \theta)$ 矛盾. 现取 $x \in \mathscr{X}$ 满足 $\|x\|_T = 1$, 由 P 的次线性性知:

$$\begin{aligned}
P(x) &= P(\sum_{i=1}^{n} \xi_i e_i) \le |\xi_i| P(e_i \cdot \operatorname{sgn} \xi_i) \\
&\le M \sum_{i=1}^{n} |\xi_i| \le M(1^2 + \cdots + 1^2)^{\frac{1}{2}} \cdot (\xi_1^2 + \cdots + \xi_n^2)^{\frac{1}{2}} \\
&= \sqrt{n} M \|x\|_T = \sqrt{n} M =: M_0
\end{aligned}$$

而若 $\mathbb{K}=\mathbb{C}$, 则令

$$M := \max\{P(\pm e_j), P(\pm ie_j) : j=1,\cdots,n\},\ M_0 := \sqrt{2n}M$$

亦即将 $\mathbb{C}$ 视作 $\mathbb{R}^{2n}$, 类似上述证明可证 P 的有界性, 从而 P 在 $\{x\in\mathscr{X} : \|x\|_T=1\}$ 上有界 M_0. 对于 $\xi\in\mathbb{K}^n$, 设 $\widetilde{P}(\xi)=P(T^{-1}\xi)$, 则当 $\xi\neq\eta$ 时有:

$$\begin{aligned}
|\widetilde{P}(\xi)-\widetilde{P}(\eta)| &= |P(T^{-1}\xi)-P(T^{-1}\eta)| \\
&\overset{\text{(i)}}{\le} \max\{P(T^{-1}(\xi-\eta)), P(T^{-1}(\eta-\xi))\} \\
&\overset{\text{(ii)}}{=} |\xi-\eta|\cdot\max\{P(T^{-1}(\frac{\xi-\eta}{|\xi-\eta|})), P(T^{-1}(\frac{\eta-\xi}{|\eta-\xi|}))\} \\
&\le M_0|\xi-\eta|
\end{aligned}$$

其中 (i) 是因为对任意 $\xi,\eta\in\mathscr{X}$, 根据次可加性知:

$$P(T^{-1}\xi)\le P(T^{-1}\xi-T^{-1}\eta)+P(T^{-1}\eta)=P(T^{-1}(\xi-\eta))+P(T^{-1}\eta)$$

得到

$$P(T^{-1}\xi)-P(T^{-1}\eta)\le P(T^{-1}(\xi-\eta))$$

交换 ξ,η 位置并结合二式即得 (i), 而 (ii) 是出于 P 的齐次性.

上式说明 $\widetilde{P}(\xi)$ 在 $\{\xi\in\mathbb{K}^n : |\xi|=1\}$ 上连续, 因而由 $\{\xi\in\mathbb{K}^n : |\xi|=1\}$ 的紧性知其有最大最小值. 设 $\widetilde{P}(\xi_0)=\min_{\substack{\xi\in\mathbb{K}^n\\|\xi|=1}}\widetilde{P}(\xi)$, 若 $\widetilde{P}(\xi_0)=0$, 则

$$P(T^{-1}\xi_0)=0\Leftrightarrow T^{-1}(\xi_0)=0\Leftrightarrow \xi_0=0$$

这与 $|\xi_0|=1$ 矛盾! 故存在常数 $C_1',C_2'>0$ 使得对任意 ξ, 若 $|\xi|=1$, 则

$$C_1'\le\widetilde{P}(\xi)\le C_2'$$

从而 $\xi\neq0$ 时有:

$$\begin{aligned}
&C_1'\le\widetilde{P}(\frac{\xi}{|\xi|})\le C_2' \\
\Rightarrow& C_1'|\xi|\le\widetilde{P}(\xi)\le C_2'|\xi| \\
\Rightarrow& C_1'|\xi|\le P(T^{-1}\xi)\le C_2'|\xi|
\end{aligned}$$

设 $T^{-1}\xi=x$, 则 $\xi=Tx$, 因而

$$\frac{C_1'}{\widetilde{C}_2}\|x\|\le C_1'\|x\|_T\le P(x)\le C_2'|Tx|=C_2'\|x\|_T\le\frac{C_2'}{\widetilde{C}_1}\|x\|$$

而当 $\xi=0$ 时 $x=0$, 上述不等式依旧成立. 取 $C_1=\frac{C_1'}{\widetilde{C}_2}, C_2=\frac{C_2'}{\widetilde{C}_1}$ 即得命题. □

注 半范数与范数之间只相差原点的确定. 亦即如果其余条件均满足, 而 $P(x)=0\Rightarrow x=\theta$, 就说明 P 是范数, 而这件事不一定成立就说明 P 是半范数. 一般可以设置一族半范数, 使得元素的这族半范数全为 0 可推出元素本身为 0, 进而这族半范数就扮演了范数的角色. 半范数在分析中有非常重要的地位, 事实上在更加现代的教材中, Frechet 空间的定义其实并非范数收敛, 而是半范数收敛. **[FL]** 中给出了半范数定义拓扑的方式, 进而给出了 Frechet 空间的定义.

补充定理 2.4.3 (半范数确定的拓扑)

设 $\{p_\alpha\}_{\alpha\in A}$ 是线性空间 $\mathscr{X}$ 上的一族半范数. 设 $x\in\mathscr{X},\alpha\in A,\varepsilon>0$, 令

$$U_{x\alpha\varepsilon}=\{y\in\mathscr{X} : p_\alpha(y-x)<\varepsilon\}$$

并设 $\mathscr{T}$ 是集族 $U_{x\alpha\varepsilon}$ 生成的拓扑, 则:

(i) 对任意 $x\in\mathscr{X}$, 集合 $U_{x\alpha\varepsilon}(\alpha\in A,\varepsilon>0)$ 有限并构成 x 的一个邻域.

(ii) 若 $\langle x_i\rangle_{i\in I}$ 是 $\mathscr{X}$ 中的网[a], 则 $x_i\to x$ 当且仅当对任意 $\alpha\in A$ 都有 $p_\alpha(x_i-x)\to0$.

(iii) $(\mathscr{X},\mathscr{T})$ 是局部紧拓扑空间.

[a]与滤子的概念类似, 都表示极限过程.

半范数定义的拓扑也可以阐述连续性:

补充命题 2.4.3

设 $\mathscr{X},\mathscr{Y}$ 是线性拓扑空间, 其上的拓扑分别由半范数族 $\{p_\alpha\}_{\alpha\in A},\{q_\beta\}_{\beta\in B}$ 定义, $T:\mathscr{X}\to\mathscr{Y}$ 是线性映射, 则 T 连续当且仅当对任意 $\beta\in B$, 存在 $\alpha_1,\cdots,\alpha_k\in A$ 与 $C>0$ 使得 $q_\beta(Tx)\le C\sum_{j=1}^k p_{\alpha_j}(x)$.

前面定义半范数拓扑是通过邻域定义的, 自然也可以借此说明拓扑空间是否为 Hausdorff 空间:

补充命题 2.4.4

设 $\mathscr{X}$ 是装备了由半范数族 $\{p_\alpha\}_{\alpha\in A}$ 定义的拓扑的线性空间, 则:

(i) $\mathscr{X}$ 是 Hausdorff 空间当且仅当对任意 $x\neq 0$, 都存在 $\alpha\in A$ 使得 $p_\alpha(x)\neq 0$.

(ii) 若 $\mathscr{X}$ 是 Hausdorff 空间, A 是可数集, 则 $\mathscr{X}$ 可赋平移不变度量.

补充定义 2.4.6 (Frechet 空间)

拓扑由可数半范数族定义的紧 Hausdorff 拓扑线性空间称为 Frechet 空间.

2.4.1.3 最佳逼近问题

逼近论的一个基本问题是, 给定一组函数 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 和一个函数 f, 要求 (按某种尺度) 用 $\varphi_1,\varphi_2,\cdots,\varphi_n$ 去逼近 f, 问是否有最佳的逼近存在? 例如 f 是 $[0,2\pi]$ 上的一个周期函数, $\varphi_i(x)=\cos ix(i=1,2,\cdots,n)$, 用 $\sum_{i=1}^n\lambda_i\varphi_i$ 去逼近 f, 求在 $L^p[0,2\pi]$ 意义下的最佳逼近.

现在在 B^* 空间中重述问题, 给定一个 B^* 空间 $\mathscr{X}$, 并给定 $\mathscr{X}$ 中的有穷个向量 $e_1,e_2,\cdots,e_n$. 对给定的向量 $x\in\mathscr{X}$, 求一组数 $(\lambda_1,\lambda_2,\cdots,\lambda_n)\in\mathscr{K}^n$, 使得

$$\|x-\sum_{i=1}^n\lambda_ie_i\|=\min_{a\in\mathbb{K}^n}\|x-\sum_{i=1}^n a_ie_i\|,\ a=(a_1,a_2,\cdots,a_n).$$

首先需要探究这组数 $(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 是否存在, 不妨设 $e_1,e_2,\cdots,e_n$ 线性无关, 要求函数

$$F(a)=\|x-\sum_{i=1}^n a_ie_i\|,\ a\in\mathbb{K}^n$$

的最小值. 显见 F 是 $\mathbb{K}^n$ 上的连续函数, 又注意到

$$F(a)\ge\|\sum_{i=1}^n a_ie_i\|-\|x\|,\ \forall a\in\mathbb{K}^n$$

令 $P(a):=\|\sum_{i=1}^n a_ie_i\|$, 显见 $P(\cdot)$ 是 $\mathbb{K}^n$ 上的一个范数, 又因为 $\mathbb{K}^n$ 是有穷维空间, 故应用定理2.4.1, 知存在 $c_1>0$ 使得

$$P(a)\ge c_1|a|,\ \forall a\in\mathbb{K}^n$$

其中

$$|a|:=(|a_1|^2+|a_2|^2+\cdots+|a_n|^2)^{\frac{1}{2}},\ \forall a=(a_1,a_2,\cdots,a_n)\in\mathscr{K}^n$$

这说明 $|a|\to\infty\Rightarrow F(a)\to\infty$, 这说明 F 确有最小值存在. 事实上, 如果 F 没有最小值, 因为 F 首先是连续函数, 故 $-\infty$ 必在 $|a|=\infty$ 时取, 但这与前面的论述矛盾! 进而得到下面的定理.

定理 2.4.3

设 $\mathscr{X}$ 是一个 B^* 空间, 若 $e_1, e_2, \cdots, e_n$ 是 $\mathscr{X}$ 中给定的向量组, 则 $\forall x \in \mathscr{X}$, 存在最佳逼近系数 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 满足

$$\|x - \sum_{i=1}^{n} \lambda_i e_i\| = \min_{a \in \mathbb{K}^n} \|x - \sum_{i=1}^{n} a_i e_i\|, \ a = (a_1, a_2, \cdots, a_n)$$

注 若记 $M := \operatorname{span}\{e_1, e_2, \cdots, e_n\}, \rho(x, M) := \inf\limits_{y \in M} \|x - y\|, x_0 := \sum_{i=1}^{n} \lambda_i e_i$, 则可将上式改写为

$$\rho(x, x_0) = \rho(x, M)$$

适合该式子的 $x_0 \in M$ 称作 x 在 M 上的最佳逼近元. 该定理说明: 在 B^* 空间中, 任一指定元素在给定的有限维子空间上的最佳逼近元总是存在的.

下面再来讨论最佳逼近元的唯一性, 首先需假设给定的向量组 $e_1, e_2, \cdots, e_n$ 线性无关, 同时唯一性依赖于 B^* 空间 $\mathscr{X}$ 的范数的性质.

定义 2.4.17 (严格凸)

B^* 空间 $(\mathscr{X}, \|\cdot\|)$ 称为严格凸的, 如果对任意的 $x, y \in \mathscr{X}, x \neq y$, 必有

$$\|x\| = \|y\| = 1 \Rightarrow \|\alpha x + \beta y\| < 1, \ \forall \alpha, \beta > 0, \alpha + \beta = 1$$

课堂笔记 (★)

严格凸的条件有下述等价形式:

- $(\mathscr{X}, \|\cdot\|)$ 严格凸 $\Leftrightarrow \forall x, y \in \mathscr{X}((x \neq y) \wedge (\|x\| = \|y\| = 1) \Rightarrow \|x + y\| < 2)$.

  证明:

  当 $(\mathscr{X}, \|\cdot\|)$ 严格凸时, 取 $\alpha = \beta = \frac{1}{2}$ 即可. 当 $\forall x, y \in \mathscr{X}((x \neq y) \wedge (\|x\| = \|y\| = 1) \Rightarrow \|x + y\| < 2)$, 不妨设 $\beta > \alpha$, 则:

$$\begin{aligned} \|\alpha x + \beta y\| &= \|\alpha x + \alpha y + (\beta - \alpha) y\| \leq \|\alpha x + \alpha y\| + (\beta - \alpha)\|y\| \\ &= 2\alpha \|\frac{x + y}{2}\| + \beta - \alpha < 2\alpha + \beta - \alpha = \beta + \alpha = 1 \end{aligned}$$

  此即严格凸的定义. □

- $(\mathscr{X}, \|\cdot\|)$ 严格凸 $\Leftrightarrow \exists \alpha_0, \beta_0 \in (0,1) \forall x, y \in \mathscr{X}((\alpha_0 + \beta_0 = 1) \wedge (x \neq y) \wedge (\|x\| = \|y\| = 1) \Rightarrow \|\alpha_0 x + \beta_0 y\| < 1)$.

  证明:

  当 $(\mathscr{X}, \|\cdot\|)$ 严格凸, 令 $\alpha = \alpha_0, \beta = \beta_0$ 即可. 当 $\exists \alpha_0, \beta_0 \in (0,1) \forall x, y \in \mathscr{X}((\alpha_0 + \beta_0 = 1) \wedge (x \neq y) \wedge (\|x\| = \|y\| = 1) \Rightarrow \|\alpha_0 x + \beta_0 y\| < 1)$, 任取 $\alpha, \beta \in (0,1)$ 满足 $\alpha + \beta = 1$, 若 $\frac{\beta}{\alpha} \geq \frac{\beta_0}{\alpha_0}$, 则 $\beta \geq \alpha \frac{\beta_0}{\alpha_0}$, 从而

$$\begin{aligned} \|\alpha x + \beta y\| &= \|\alpha x + \frac{\alpha}{\alpha_0} \beta_0 y + (\beta - \frac{\alpha}{\alpha_0} \beta_0) y\| \\ &\leq \frac{\alpha}{\alpha_0} \|\alpha_0 x + \beta_0 y\| + (\beta - \frac{\alpha}{\alpha_0} \beta_0)\|y\| \\ &< \frac{\alpha}{\alpha_0} + \beta - \frac{\alpha}{\alpha_0} \beta_0 = \beta + \frac{\alpha}{\alpha_0}(1 - \beta_0) = \beta + \alpha = 1 \end{aligned}$$

  若 $\frac{\beta}{\alpha} \leq \frac{\beta_0}{\alpha_0}$, 则 $\alpha \geq \frac{\alpha_0}{\beta_0} \beta$, 从而

$$\begin{aligned} \|\alpha x + \beta y\| &= \|(\alpha - \frac{\alpha_0}{\beta_0} \beta) x + \frac{\alpha_0}{\beta_0} \beta x + \beta y\| \\ &\leq (\alpha - \frac{\alpha_0}{\beta_0} \beta)\|x\| + \frac{\beta}{\beta_0} \|\alpha_0 x + \beta_0 y\| \\ &< (\alpha - \frac{\alpha_0}{\beta_0} \beta) + \frac{\beta}{\beta_0} = \alpha + \frac{\beta}{\beta_0}(1 - \alpha_0) = \alpha + \beta = 1 \end{aligned}$$

  □

如果 $\mathscr{X}$ 是严格凸的 B^* 空间, 就不能有两个不同的最佳逼近元. 这是因为如果 $d := \rho(x, M) > 0$, 且有 $\|x-y\| = \|x-z\| = d$, 则对任意的 $\alpha, \beta > 0, \alpha+\beta = 1$, 由严格凸性:

$$\frac{1}{d}\|x-(\alpha y+\beta z)\| = \frac{1}{d}\|\alpha(x-y)+\beta(x-z)\| = \|\alpha(\frac{x-y}{d})+\beta(\frac{x-z}{d})\| < 1$$

这说明 $\|x-(\alpha y+\beta z)\| < d$, 进而 d 对应的元并非最佳逼近元, 矛盾! 而如果 $d = 0$, y 是相应的最佳逼近元, 必有 $\|x-y\| = 0$, 也即 $y = x$, 进而最佳逼近元必唯一.

定理 2.4.4 (最佳逼近)

设 $\mathscr{X}$ 是严格凸的 B^* 空间, $\{e_1, e_2, \cdots, e_n\}$ 是 $\mathscr{X}$ 上给定的一组线性无关向量, 则对任意的 $x \in \mathscr{X}$, 存在唯一的一组最佳逼近系数 $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ 满足

$$\|x-\sum_{i=1}^{n}\lambda_i e_i\| = \min_{a\in\mathbb{K}^n}\|x-\sum_{i=1}^{n}a_i e_i\|, \ a = (a_1, a_2, \cdots, a_n).$$

例 2.32 空间 $L^p(\Omega, \mu)(1 < p < \infty)$ 是严格凸的.

证明

这是因为 Minkowski 不等式 $\|u+v\| \le \|u\|+\|v\|$ 等号成立的充要条件是, $\exists k_1, k_2 \ge 0(k_1+k_2 > 0)$ s.t. $k_1 u = k_2 v$(a.e.). 这意味着 $\forall u, v \in L^p(\Omega, \mu)$, 当 $\|u\| = \|v\| = 1, u \ne v$ 时:

$$\|tu+(1-t)v\| < t\|u\|+(1-t)\|v\| = 1, \ \forall 0 < t < 1$$

这说明 $L^p(\Omega, \mu)$ 是严格凸的. □

例 2.33 $C(M)$ 与 $L^1(\Omega, \mu)$ 均不是严格凸的.

证明

对 $C[0,1]$, 取 $x(t) \equiv 1, y(t) = t$, 知 $\|x\| = \|y\| = 1$, 而 $\|\frac{1}{2}(x+y)\| = 1$.

对 $L^1[0,1]$, 取 $x(t) \equiv 1, y(t) = 2t$, 知 $\|x\| = \|y\| = 1$, 而 $\|\frac{1}{2}(x+y)\| = 1$. □

例 2.34★ $(\mathbb{R}^2, \|\cdot\|_3)$ 不是严格凸的, 其中 $\|(x,y)\|_3 = \max\{|x|, |y|\}$.

证明

取 $(1,1), (1,-1)$, 知

$$\|\alpha(1,1)+\beta(1,-1)\|_3 = \|(1, \alpha-\beta)\|_3 = 1$$

其并不满足严格凸的定义! □

注 上例说明空间的严格凸性依赖于取用的范数.

2.4.1.4 有穷维 B^* 空间的刻画

首先说明有穷维 B^* 空间 $\mathscr{X}$ 上的单位球面

$$\mathbb{S}_1 := \{x \in \mathscr{X} \mid \|x\| = 1\}$$

是列紧的 (事实上用下述方法类似可证有穷维 B^* 空间中任意有界集均为列紧集):

课堂笔记 (★)

记有穷维 B^* 空间 $\mathscr{X}$ 的维数为 n, 取

$$T: \mathscr{X} \to \mathbb{K}^n, \ x = \xi_1 e_1+\cdots+\xi_n e_n \mapsto (\xi_1, \cdots, \xi_n)$$

记 $|Tx| = (\sum_{i=1}^n |\xi_i|^2)^{\frac{1}{2}}$, 则回忆推论(2.4.3)中的证明, 已知存在 $C_1, C_2 > 0$ 使得

$$C_1|Tx| \le \|x\| \le C_2|Tx|$$

故对任意 $x \in \mathbb{S}_1$ 有 $|Tx| \leq \frac{1}{C_1}$, 从而 $T\mathbb{S}_1 := \{Tx : x \in \mathbb{S}_1\}$ 是 $\mathbb{K}^n$ 中的有界集. 再说明 $\mathbb{S}_1$ 的闭性, 若 $x \in \overline{\mathbb{S}_1}$, 则存在 $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{S}_1$ 使得 $\|x_n - x\| \to 0(n \to \infty)$, 从而

$$||\|x\| - 1| = |\|x\| - \|x_n\|| \leq \|x - x_n\| \to 0,\ n \to \infty$$

因而 $\|x\| = 1$, 从而 $x \in \mathbb{S}_1$, 故 $\mathbb{S}_1$ 是闭集, 根据 T 的连续性知 $T\mathbb{S}_1$ 是 $\mathbb{K}^n$ 中的闭集, 因而其为 $\mathbb{K}^n$ 中的有界闭集, 从而 $T\mathbb{S}_1$ 为 $\mathbb{K}^n$ 中的自列紧集, 根据 T 的连续性知 $\mathbb{S}_1$ 为 $\mathscr{X}$ 中的自列紧集, 亦即紧集.

现在反过来证明, 如果一个 B^* 空间 $\mathscr{X}$ 的单位球面是列紧的, 那么该空间必是有穷维. 事实上, 如果在 $\mathbb{S}_1$ 上给定了有穷个线性无关的向量 $\{x_1, x_2, \cdots, x_n\}$, 如果它们的线性包 M_n 张不满 $\mathscr{X}$, 则存在 $x_{n+1} \in \mathbb{S}_1$ 使得

$$\|x_{n+1} - x_i\| \geq 1,\ i = 1, 2, \cdots, n$$

这是因为任取 $y \notin M_n$, 根据定理2.4.3[16]知存在 $x \in M_n$, 使得

$$\|y - x\| = d := \rho(y, M_n)$$

注意必定有 $d > 0$, 这是因为如果 $d = 0$, 根据距离的下确界定义知存在 M_n 中的序列 $\{y_n\}$ 使得 $\lim_{n\to\infty} \rho(y, y_n) = 0$. 又因为 M_n 作为有穷个向量的线性包必定是有限维 B^* 空间, 进而其是 $\mathscr{X}$ 中的闭集, 得到 $y \in M_n$, 矛盾! 现在令 $x_{n+1} := \frac{y-x}{d}$, 则 $x_{n+1} \in \mathbb{S}_1$, 且

$$\|x_{n+1} - x_i\| = \frac{1}{d}\|y - (x + dx_i)\| \geq \frac{1}{d} \cdot d = 1,\ i = 1, 2, \cdots, n$$

由此, 若 $\mathscr{X}$ 是无穷维的, 就可以逐次在 $\mathbb{S}_1$ 上选出一串 $\{x_n\}_{n=1}^{\infty}$ 适合 $\|x_n - x_m\| \geq 1(n, m \in \mathbb{N}, n \neq m)$, 进而 $\mathbb{S}_1$ 不是列紧的. 这便得到下述定理.

定理 2.4.5 (Riesz)

要使得 B^* 空间 $\mathscr{X}$ 是有穷维的, 必须且只需 $\mathscr{X}$ 的单位球面是列紧的.

定义 2.4.18 (有界)

B^* 空间 $\mathscr{X}$ 上的一个子集 A 称为是有界的, 如果存在常数 $c > 0$, 使得 $\|x\| \leq c, \forall x \in A$.

推论 2.4.9

要使得 B^* 空间 $\mathscr{X}$ 是有穷维的, 必须且只需其任意有界集均列紧.

证明 (★)

只需证明有穷维 B^* 空间 $\mathscr{X}$ 的任意有界集列紧. 已知 $\mathscr{X}$ 的单位球面列紧, 设 $M := \{x \in \mathscr{X} : \|x\| \leq C\}$ 为有界集, 任取 $\{x_n\} \subset M$, 因为 $\|x_n\| \leq C$, 故 $\{\|x_n\|\}_{n\in\mathbb{N}}$ 为 $\mathbb{R}$ 中的有界序列, 因而存在收敛子列 $\{\|x_{n_i}\|\}$, 对应 $\{x_n\}$ 的子列 $\{x_{n_i}\}$. 设

$$\|x_{n_i}\| \to a,\ i \to \infty$$

进一步不妨设 $a \neq 0$, 这是因为若 $a = 0$, 则知 $\|x_{n_i}\| \to 0(i \to \infty)$, 从而根据范数的正定性知 $\mathscr{X} \ni x_{n_i} \to \theta(i \to \infty)$, $\{x_n\}$ 自然有收敛子列, 命题已然成立. 当 $a \neq 0$ 时, 可设 $\forall i \in \mathbb{N}(\|x_{n_i}\| \neq 0)$, 知 $\{\frac{x_{n_i}}{\|x_{n_i}\|}\} \subset \mathbb{S}_1$, 又因为 $\mathbb{S}_1$ 列紧, 故有收敛子列 $\{\frac{x_{n_{i_k}}}{\|x_{n_{i_k}}\|}\}_{k\in\mathbb{N}}$. 现设

$$\frac{x_{n_{i_k}}}{\|x_{n_{i_k}}\|} \to x_0,\ k \to \infty$$

则由

$$x_{n_{i_k}} = \|x_{n_{i_k}}\|(\frac{x_{n_{i_k}}}{\|x_{n_{i_k}}\|} - x_0) + \|x_{n_{i_k}}\|x_0$$

与 $\|x_{n_{i_k}}\| \leq C$, 两边同时令 $k \to \infty$ 得 $x_{n_{i_k}} \to ax_0$, 因而 $\{x_n\}$ 有收敛子列, 故 M 列紧. □

[16] 不选取最佳逼近定理2.4.4是因为现在不知道 M_n 是否严格凸.

引理 2.4.1 (Riesz)

如果 $\mathscr{X}_0$ 是 B^* 空间 $\mathscr{X}$ 的一个真闭子空间, 则对任意的 $0<\varepsilon<1$, 存在 $y\in\mathscr{X}$, 使得 $\|y\|=1$, 且

$$\|y-x\|\geq 1-\varepsilon,\ \forall x\in\mathscr{X}_0$$

证明

任取 $y_0\in\mathscr{X}\backslash\mathscr{X}_0$, 因为 $\mathscr{X}_0$ 是闭的, 故

$$d:=\inf_{x\in\mathscr{X}_0}\|y_0-x\|>0$$

根据下确界性质知

$$\forall\eta>0\exists x_0\in\mathscr{X}_0(d\leq\|y_0-x_0\|<d+\eta)$$

现取 $y:=\frac{y_0-x_0}{\|y_0-x_0\|}$, 则 $\|y\|=1$, 且对任意 $x\in\mathscr{X}_0$ 有:

$$\|y-x\|=\frac{\|y_0-x'\|}{\|y_0-x_0\|}>\frac{d}{d+\eta}=1-\frac{\eta}{d+\eta}$$

其中 $x_0'=x_0+\|y_0-x_0\|x\in\mathscr{X}_0$. 故对于任意的 $0<\varepsilon<1$, 令 $\eta=\frac{d\varepsilon}{1-\varepsilon}$, 即得 $\|y-x\|>1-\varepsilon$. □

注

- Riesz 引理表明, 总能在赋范空间的单位球面中, 找到与其真闭子空间" 足够远" 的点.
- WL Riesz 引理的表述不能改成存在 $y\in\mathscr{X}$, 使得 $\|y\|=1$ 且 $\|y-x\|\geq 1(\forall x\in\mathscr{X}_0)$. 这是因为譬如 $\mathscr{X}$ 是定义在 $[0,1]$ 上满足 $x(0)=0$ 的全体连续函数 $x(t)$ 构成的空间, 线性运算赋通常运算. 取用一致范数 $\|x\|=\max_{t\in[0,1]}|x(t)|$, 知 $(\mathscr{X},\|\cdot\|)$ 是 Banach 空间. 令

$$\mathscr{X}_0=\{x:x\in\mathscr{X}\text{ 且 }\int_0^1 x(t)dt=0\}$$

显见 $\mathscr{X}_0$ 是 $\mathscr{X}$ 的真闭子空间, 下面说明不存在 $x_0\in\mathscr{X}$ 同时满足 $\|x_0\|=1$ 与 $\|x_0-x\|\geq 1(\forall x\in\mathscr{X}_0)$. 用反证法, 若存在 $x_0\in\mathscr{X}$ 同时满足这两件事, 则对任意 $y\in\mathscr{X}\backslash\mathscr{X}_0$, 取

$$b_y=\frac{\int_0^1 x_0(t)dt}{\int_0^1 y(t)dt}$$

显见 $\int_0^1[x_0(t)-b_y y(t)]dt=0$, 因而 $x_0-b_y y\in\mathscr{X}_0$, 从而

$$1\leq\|x_0-(x_0-b_y y)\|=\|b_y y\|$$

亦即

$$\frac{1}{|b_y|}\leq\|y\|$$

将 b_y 的构造代入可得

$$|\int_0^1 y(t)dy|=\frac{|\int_0^1 x_0(t)dt|}{|b_y|}\leq\|y\||\int_0^1 x_0(t)dt|,\ \forall y\in\mathscr{X}\backslash\mathscr{X}_0$$

现在任取 $n\in\mathbb{N}$, 记 $g_n(t)=t^{\frac{1}{n}}\in\mathscr{X}\backslash\mathscr{X}_0$, 将 g_n 代入上式得到

$$\frac{n}{n+1}=|\int_0^1 g_n(t)dt|\leq\|g_n\||\int_0^1 x_0(t)dt|=|\int_0^1 x_0(t)dt|$$

令 $n\to\infty$ 知 $|\int_0^1 x_0(t)dt|\geq 1$. 又因为 $x_0\in\mathscr{X},\|x_0\|=1,x_0(0)=0$ 能推出 $|\int_0^1 x_0(t)dt|<1$, 矛盾! 故欲证成立.
□

2.4.1.5 商空间

设 $(\mathscr{X},\|\cdot\|)$ 是 B^* 空间, $\mathscr{X}_0\subset\mathscr{X}$ 是一个闭线性子空间, $x',x''\in\mathscr{X}$. 若 $x'-x''\in\mathscr{X}_0$, 则记 $x'\sim x''$. 可知这是一个等价关系, 用 $[x]$ 表 x 所在的等价类, 全体等价类组成的空间称为关于 $\mathscr{X}_0$ 的商空间, 记为 $\mathscr{X}/\mathscr{X}_0$, 其中加法和数乘定义为:

1. $[x]+[y]=[x+y], [x],[y]\in\mathscr{X}/\mathscr{X}_0$;
2. $\lambda[x]=[\lambda x], \lambda\in\mathbb{K}, [x]\in\mathscr{X}/\mathscr{X}_0$.

可以验证 $\mathscr{X}/\mathscr{X}_0$ 关于上述运算组成一个线性空间.

定理 2.4.6

设 $[x]\in\mathscr{X}/\mathscr{X}_0$, 定义

$$\|[x]\|_0=\inf_{y\in[x]}\|y\|$$

则 $(\mathscr{X}/\mathscr{X}_0,\|\cdot\|_0)$ 是 B^* 空间, 又当 $(\mathscr{X},\|\cdot\|)$ 是 Banach 空间时, $(\mathscr{X}/\mathscr{X}_0,\|\cdot\|_0)$ 也是 Banach 空间.

证明

根据商空间自身的定义, $\mathscr{X}/\mathscr{X}_0$ 中零元素为 $[x], x\in\mathscr{X}_0$. 又根据下确界性质知

$$\|[x]\|_0\geq 0, \|\lambda[x]\|=|\lambda|\|[x]\|_0,\ \forall[x]\in\mathscr{X}/\mathscr{X}_0, \lambda\in\mathbb{R}$$

成立. 下面验证三角不等式:

$$\|[x]+[y]\|_0=\|[x+y]\|_0=\inf_{x\in[x+y]}\|z\|$$

这是因为根据下确界性质, 可取 $x_n\in[x], y_n\in[y]$, 使得

$$\|x_n\|\to\|[x]\|_0, \|y_n\|\to\|[y]\|_0,\ n\to\infty$$

此时 $x_n+y_n\in[x+y]$, 进而

$$\|[x]+[y]\|_0\leq\|x_n+y_n\|\leq\|x_n\|+\|y_n\|\to\|[x]\|_0+\|[y]\|_0,\ n\to\infty$$

也即三角不等式

$$\|[x]+[y]\|_0\leq\|[x]\|_0+\|[y]\|_0$$

最后说明连续性, 设 $\|[x]\|_0=0$, 根据下确界性质知存在 $x_n\in[x]$ 使得 $\|x_n\|\to 0, n\to\infty$. 因为 $\mathscr{X}_0$ 是闭子空间, 其中 Cauchy 列必有极限, 进而由 $x-x_n\in\mathscr{X}_0$ 得到

$$x=\lim_{n\to\infty}(x-x_n)\in\mathscr{X}_0$$

也即 $[x]=\theta$ 是 $\mathscr{X}/\mathscr{X}_0$ 中的零元素, 进而 $(\mathscr{X}/\mathscr{X}_0,\|\cdot\|_0)$ 是 B^* 空间.

当 $(\mathscr{X},\|\cdot\|)$ 是 Banach 空间, 下面证明 $(\mathscr{X}\backslash\mathscr{X}_0,\|\cdot\|_0)$ 完备. 令 $\{[x_n]\}$ 是 Cauchy 列, 进而

$$\|[x_n]-[x_m]\|_0=\|[x_n-x_m]\|_0\to 0,\ n,m\to\infty$$

取子列, 为简便仍然记作 $\{[x_n]\}$, 使得

$$\|[x_n-x_{n+1}]\|_0\leq\frac{1}{2^{n+1}}$$

根据下确界性质, 存在 $y_{n,n+1}\in\mathscr{X}, y_{n,n+1}\in[x_n-x_{n+1}]$, 满足

$$\|y_{n,n+1}\|\leq\|[x_n-x_{n+1}]\|_0+\frac{1}{2^{n+1}}\leq\frac{1}{2^n}$$

现在记 $y_1=x_1, y_{n+1}=y_n-y_{n,n+1}, n\geq 1$, 知存在 $z_{n,n+1}\in\mathscr{X}_0$, 使得

$$y_{n+1}=y_n-(x_n-x_{n+1}+z_{n,n+1})$$

也即

$$y_{n+1}-x_{n+1}=y_n-x_n-z_{n,n+1}\in\mathscr{X}_0$$

因为 $y_1=x_1$, 故归纳知 $[y_{n+1}]=[x_{n+1}]$, 此时

$$\|y_n-y_{n+p}\|\leq\|y_n-y_{n+1}\|+\cdots+\|y_{n+p+1}-y_{n+p}\|=\|y_{n,n+1}\|+\cdots+\|y_{n+p-1,n+p}\|\leq\frac{1}{2^{n+1}}+\cdots+\frac{1}{2^{n+p}}\leq\frac{1}{2^n}$$

这说明 $\{y_n\}$ 是 $\mathscr{X}$ 中的 Cauchy 列, 进而因为 $\mathscr{X}$ 是 Banach 空间, 有 $y\in\mathscr{X}$ 使得

$$\|y_n-y\|\to 0,\ n\to\infty$$

故
$$\|[x_n]-[y]\|_0=\|[y_n]-[y]\|_0=\|[y_n-y]\|_0\le\|y_n-y\|\to 0,\ n\to\infty$$
也即 $(\mathscr{X}/\mathscr{X}_0,\|\cdot\|_0)$ 完备.

2.4.2 一些例子

例 2.35(任给线性空间 X, 都可在 X 上赋范, 得到赋范线性空间) 因为线性空间必有 Hamel 基, 可设 $\{h_\alpha\}_{\alpha\in A}$ 是 X 的一个 Hamel 基, 根据定义知
$$\forall x\in X(x=\sum_{\alpha\in A(x)}t_\alpha h_\alpha)$$
其中 x 的表示唯一, $A(x)\subset A$ 是有限集. 令 $\|x\|=\max_{\alpha\in A(x)}|t_\alpha|$, 下面验证 $\|\cdot\|$ 是范数:

- $\forall x\in X(\|x\|\ge 0)$, 且 $\|x\|=0\Leftrightarrow x=0$. 这点是显然的.
- $\forall\alpha\in\mathbb{K}(\|\alpha x\|=|\alpha|\|x\|)$, 这点是显然的.
- $\forall x,y\in X(\|x+y\|\le\|x\|+\|y\|)$. 这是因为根据 x 表示的唯一性, 可以说明 $A(x+y)=A(x)\cup A(y)$, 再利用通常的三角不等式即得结论.

综上, $\|\cdot\|$ 是范数, $(X,\|\cdot\|)$ 是赋范线性空间.

例 2.36(不能赋予完备范数的线性空间) 前例告诉我们任意线性空间都能成为赋范线性空间, 下面问任意线性空间是否都能成为 B 空间? 答案是否定的, 例如取 E^∞ 为只有有限个非零坐标的实数列 $x=(a_1,\cdots,a_n,\cdots)$ 的全体, 定义通常的加法与数乘运算, 显见 E^∞ 是线性空间.

下面证明 E^∞ 上不能赋完备范数, 如若不然, 设 $(E^\infty,\|\cdot\|)$ 为 Banach 空间, 则其作为完备赋范空间应为第二纲集 (这是 Baire 纲定理). 另一方面, 取 $e_n=(0,\cdots,0,1,0,\cdots)$, 则 $\{e_n\}_{n=1}^\infty$ 所张成的线性空间正是 E^∞, 而由有限向量组 $\{e_1,\cdots,e_n\}$ 张成的线性子空间 $E_n=\mathrm{span}\{e_1,\cdots,e_n\}$ 是 $(E^\infty,\|\cdot\|)$ 的闭真子空间, 且
$$E^\infty=\bigcup_{n=1}^\infty E_n$$
下面证明各 E_n 在 $(E^\infty,\|\cdot\|)$ 内均无处稠密, 也即 $\overline{(E_n^\circ)}=E_n^\circ=\varnothing$. 用反证法, 若存在 n_0 使得 $E_{n_0}^\circ\ne\varnothing$, 则可取 $x_0\in E_{n_0}^\circ$ 与 x_0 的某开球 $B(x_0,r)=\{x:\|x_0-x\|<r\}$, 使得
$$B(x_0,r)\subset E_{n_0}^\circ\subset E_{n_0}$$
任取 $x\in E^\infty$, 显见存在 $\mathbb{R}\ni\beta>0$, 使得 $\beta x+x_0\in B(x_0,r)\subset E_{n_0}$. 因为 E_{n_0} 是线性空间, 故 $x\in E_{n_0}$, 从而 $E_{n_0}=E^\infty$, 但这与 E_{n_0} 真包含于 $(E^\infty,\|\cdot\|)$ 矛盾! 故 E_{n_0} 在 $(E^\infty,\|\cdot\|)$ 中无处稠密, 因而由 Baire 定理知 E^∞ 是第一纲集, 矛盾! 故 E^∞ 上不能赋完备范数.

注 上例出现的关键定理与定义如下.

定义 2.4.19 (第一纲集, 第二纲集)

在度量空间 $(\mathscr{X},\rho)$ 上, 集合 E 称为第一纲集, 如果 $E=\bigcup_{n=1}^\infty E_n$, 其中 E_n 是疏集[a]. 不是第一纲集的集合称作第二纲集.

[a]闭包无内点的集.

定理 2.4.7 (Baire 纲定理)

完备度量空间均为第二纲集.

定理 2.4.8 (Baire[ZMQ])

设 $E=\bigcup_{k=1}^{\infty}F_k$, 其中 $F_k(k=1,2,\cdots)$ 是闭集. 若每个 F_k 皆无内点, 则 E 也无内点.

同时可以证明, 具有无穷 Hamel 维数 a 的线性空间可赋予完备范数的充要条件为 $a\aleph^0=a$, 其中 $\aleph^0$ 表示可数集的势. 因为 $\{e_n\}_{n=1}^{\infty}$ 是线性空间 E^{∞} 的一个 Hamel 基, 故 E^{∞} 的 Hamel 维数为 $\aleph^0$, 而 $\aleph^0\aleph_0\neq\aleph_0$, 故不能在 E^{∞} 上赋完备范数.

例 2.37(存在某个线性空间上的两个不可比较的完备范数) 记 l^{∞} 中所有收敛到 0 的序列组成的线性空间为 c_0, 可以证明 c_0 与 l^2 均为 Banach 空间, 且有相同的线性维数, 从而存在由 c_0 到 l^2 上的一对一线性映射 f. 借助该线性映射可在 c_0 上取新范数:

$$\|x\|=\|f(x)\|_{l^2},\ x\in c_0$$

可以验证 $(c_0,\|\cdot\|)$ 依旧是 Banach 空间.

下面证明 c_0 上的原范数 $\|x\|_{c_0}=\|x\|_{l^{\infty}}$ 与新范数 $\|x\|$ 不可比较. 用反证法, 如若它们可以比较, 则由等价范数定理知必它们必等价, 即存在正常数 m,M 使得

$$m\|x\|_{c_0}\leq\|f(x)\|_{l^2}\|\leq M\|x\|_{c_0},\ \forall x\in c_0$$

故 f 是 c_0 到 l^2 上的一个线性同胚映射, 故 Banach 空间 c_0 与 l^2 要么均自反, 要么均不自反, 而 l^2 自反, c_0 不自反, 矛盾!

注 上例出现的关键定理与定义如下.

定理 2.4.9 (等价范数定理)

设线性空间 $\mathscr{X}$ 上有两个范数 $\|\cdot\|_1,\|\cdot\|_2$, 若 $\mathscr{X}$ 关于这两个范数均构成 Banach 空间, 且 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 则 $\|\cdot\|_2$ 必与 $\|\cdot\|_1$ 等价.

例 2.38(存在某线性空间上的强, 弱两个范数, 使得强范数完备而弱范数不完备) 在 l^1 上取范数 $\|x\|_1=\sum_{n=1}^{\infty}|\xi_n|$, $\|x\|_{\infty}=\sup_n|\xi_n|, x=\{\xi_n\}\in l^1$, 则 $(l^1,\|\cdot\|_1)$ 完备, 而 $(l^1,\|\cdot\|_{\infty})$ 不完备, 且

$$\|x\|_{\infty}=\sup_n|\xi_n|\leq\sum_{n=1}^{\infty}|\xi_n|=\|x\|_1$$

且取 $x_n=(\frac{1}{n},\cdots,\frac{1}{n},0,\cdots)$ 可知 $\|x_n\|=1(n=1,2,\cdots),\|x_n\|_{\infty}=\frac{1}{n}\to 0(n\to\infty)$, 故 $\|\cdot\|_1$ 严格强于 $\|\cdot\|_{\infty}$.

例 2.39(存在某线性空间上的强, 弱两个范数, 使得弱范数完备而强范数不完备) 回忆例(2.37)中出现的 $c_0,\|\cdot\|_{c_0},\|\cdot\|$, 知 $\|\cdot\|_{c_0},\|\cdot\|$ 是不等价的完备范数, 令

$$\|x\|'=\|x\|_{c_0}+\|x\|$$

显见 $\|\cdot\|'$ 同样是 c_0 上的一个范数, 且同时强于 $\|\cdot\|_{c_0},\|\cdot\|$ 这两个范数.

如果 $\|\cdot\|'$ 是完备范数, 则由等价范数定理知 $\|\cdot\|'$ 与 $\|\cdot\|_{c_0},\|\cdot\|$ 均等价, 进而 $\|\cdot\|_{c_0},\|\cdot\|$ 等价, 矛盾!

例 2.40(存在某线性空间上的两个完备范数, 它们的和不完备) 回忆例(2.39)中的 $c_0,\|\cdot\|_{c_0},\|\cdot\|,\|\cdot\|'$ 即可.

2.4.3 习题

练习 2.22 在二维空间 $\mathbb{R}^2$ 中, 对每一点 $z=(x,y)$, 令

$$\|z\|_1=|x|+|y|;\|z\|_2=\sqrt{x^2+y^2};\|z\|_3=\max(|x|,|y|);\|z\|_4=(x^4+y^4)^{\frac{1}{4}}$$

(1) 求证 $\|\cdot\|_i(i=1,2,3,4)$ 都是 $\mathbb{R}^2$ 的范数.

证明

统一设 $z_1=(x_1,y_1),z_2=(x_2,y_2),\alpha\in\mathbb{R}$, 易知题设的四个函数正定性与齐次性均显然, 故只需研究三角不等式即可.

对 $\|z\|_1$, 知

$$\|z_1+z_2\|_1=|x_1+x_2|+|y_1+y_2|\le|x_1|+|x_2|+|y_1|+|y_2|=\|z_1\|_1+\|z_2\|_1$$

故 $\|z\|_1$ 是范数.

对 $\|z\|_2$, 知

$$\begin{aligned}\|z_1+z_2\|_2&=\sqrt{(x_1+x_2)^2+(y_1+y_2)^2}=\sqrt{x_1^2+x_2^2+y_1^2+y_2^2+2x_1x_2+2y_1y_2}\\&\le\sqrt{x_1^2+x_2^2+y_1^2+y_2^2+2\sqrt{(x_1^2+y_1^2)(x_2^2+y_2^2)}}=\sqrt{x_1^2+y_1^2}+\sqrt{x_2^2+y_2^2}=\|z_1\|_2+\|z_2\|_2\end{aligned}$$

故 $\|z\|_2$ 是范数.

对 $\|z\|_3$, 知

$$\|z_1+z_2\|_3=\max(|x_1+x_2|,|y_1+y_2|)\le\max(|x_1|+|x_2|,|y_1|+|y_2|)\le\max(|x_1|,|y_1|)+\max(|x_2|+|y_2|)=\|z_1\|_3+\|z_2\|_3$$

故 $\|z\|_3$ 是范数.

对 $\|z\|_4$, 要证明

$$((x_1+x_2)^4+(y_1+y_2)^4)^{\frac14}\le(x_1^4+y_1^4)^{\frac14}+(x_2^4+y_2^4)^{\frac14}$$

两边 4 次方有

$$(x_1+x_2)^4+(y_1+y_2)^4\le x_1^4+x_2^4+x_2^4+y_2^4+4(x_1^4+y_1^4)^{\frac14}(x_2^4+y_2^4)^{\frac34}+6(x_1^4+y_1^4)^{\frac12}(x_2^4+y_2^4)^{\frac12}+4(x_1^4+y_1^4)^{\frac34}(x_2^4+y_2^4)^{\frac14}$$

展开左式, 只需说明

$$4(x_1x_2^3+x_1^3x_2+y_1y_2^3+y_1^3y_2)+6(x_1^2x_2^2+y_1^2y_2^2)\le4(x_1^4+y_1^4)^{\frac14}(x_2^4+y_2^4)^{\frac14}((x_2^4+y_2^4)^{\frac12}+(x_1^4+y_1^4)^{\frac12})+6(x_1^4+y_1^4)^{\frac12}(x_2^4+y_2^4)^{\frac12}$$

注意

$$\begin{aligned}x_1x_2(x_1^2+x_2^2)+y_1y_2(y_1^2+y_2^2)&\le\sqrt{(x_1^2x_2^2+y_1^2y_2^2)((x_1^2+x_2^2)^2+(y_1^2+y_2^2)^2)}=\sqrt{x_1^2x_2^2+y_1^2y_2^2}\cdot\sqrt{(x_1^2+x_2^2)^2+(y_1^2+y_2^2)^2}\\&\le(x_1^4+y_1^4)^{\frac14}(x_2^4+y_2^4)^{\frac14}\sqrt{x_1^4+x_2^4+y_1^4+y_2^4+2x_1^2x_2^2+2y_1^2y_2^2}\\&\le(x_1^4+y_1^4)^{\frac14}(x_2^4+y_2^4)^{\frac14}\sqrt{x_1^4+x_2^4+y_1^4+y_2^4+2\sqrt{(x_1^4+y_1^4)(x_2^4+y_2^4)}}\\&=(x_1^4+y_1^4)^{\frac14}(x_2^4+y_2^4)^{\frac14}((x_2^4+y_2^4)^{\frac12}+(x_1^4+y_1^4)^{\frac12})\end{aligned}$$

而 $x_1^2x_2^2+y_1^2y_2^2\le(x_1^4+y_1^4)^{\frac12}(x_2^4+y_2^4)^{\frac12}$ 是 Cauchy 不等式. 进而原不等式得证, $\|z\|_4$ 是范数.

(2) 画出 $(\mathbb{R}^2,\|\cdot\|_i)(i=1,2,3,4)$ 各空间中的单位球面图形.

解

(3) 在 $\mathbb{R}^2$ 中取定三点 $O=(0,0),A=(1,0),B=(0,1)$, 试在上述四种不同范数下求出 $\triangle OAB$ 三边的长度.

解

在 $\|z\|_1$ 下: $\overline{OA}=1,\overline{OB}=1,\overline{AB}=2$.

在 $\|z\|_2$ 下: $\overline{OA}=1,\overline{OB}=1,\overline{AB}=\sqrt{2}$.

在 $\|z\|_3$ 下: $\overline{OA}=1,\overline{OB}=1,\overline{AB}=1$.

在 $\|z\|_4$ 下: $\overline{OA}=1,\overline{OB}=1,\overline{AB}=2^{\frac14}$.

✍ **练习 2.23★** 设 $C(0,1]$ 表示 $(0,1]$ 上连续且有界的函数 $x(t)$ 全体. 对 $\forall x\in C(0,1]$, 令 $\|x\|=\sup\limits_{0<t\le1}|x(t)|$. 求证:

(1) $\|\cdot\|$ 是 $C(0,1]$ 上的范数.

证明

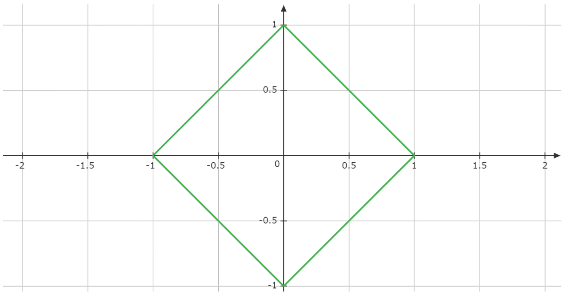


图 2.2: $\|z\|_1$

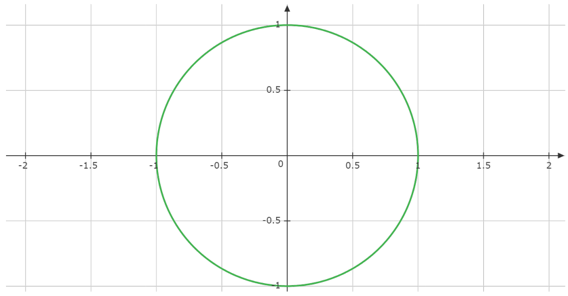


图 2.3: $\|z\|_2$

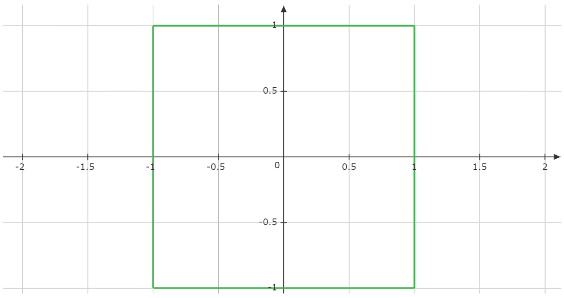


图 2.4: $\|z\|_3$

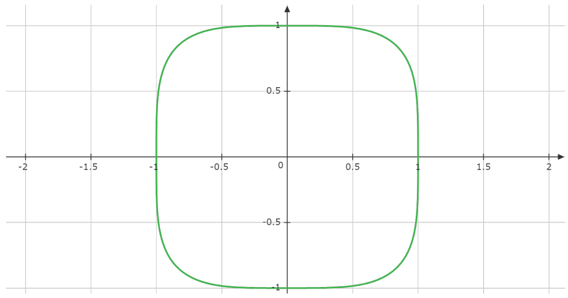


图 2.5: $\|z\|_4$

正定性与齐次性显然, 对三角不等式, 知

$$\|x+y\| = \sup_{0<t\le 1}|x(t)+y(t)| \le \sup_{0<t\le 1}|x(t)| + \sup_{0<t\le 1}|y(t)| = \|x\|+\|y\|, \forall x,y \in C(0,1]$$

故 $\|\cdot\|$ 是 $C(0,1]$ 上的范数.

(2) l^∞ 与 $C(0,1]$ 的一个子空间是等距同构的.

证明

考虑映射

$$\varphi : a = (\xi_1, \cdots, \xi_n, \cdots) \mapsto x \in C(0,1]$$

其中 x 是将 $(1,\xi_1), \cdots, (\frac{1}{n}, \xi_n), \cdots$ 线性联结的函数. 显见这样的 x 的全体构成 $C(0,1]$ 的一个子空间 $\mathscr{X}$, 且 φ 是线性映射, 下面说明 φ 是等距同构.

首先说明 φ 是双射, 这是因为对任意不同的 $a \in l^\infty$, 按前述方法找到的 $x \in C(0,1]$ 必不同, 因而 φ 是单射. 对任意的 $x \in \mathscr{X}$, 总可以取 $a_x = (x(1), \cdots, x(\frac{1}{n}), \cdots) \in l^\infty$, 故 φ 是满射, 因而其为双射.

再说明 φ 等距, 此即证明 $\|x\| = \|a\|$. 注意届于 $x \in \mathscr{X}$ 均为折线, 有

$$\|x\| = \sup_{n\in\mathbb{N}}\{x(1), \cdots, x(\frac{1}{n}), \cdots\} = \|a_x\|$$

故 φ 等距, 因而其为等距同构, 命题即证.

练习 2.24 在 $C^1[a,b]$ 中, 令

$$\|f\|_1 = (\int_a^b (|f|^2 + |f'|^2)dx)^{\frac{1}{2}} \ (\forall f \in C^1[a,b])$$

(1) 求证 $\|\cdot\|_1$ 是 $C^1[a,b]$ 上的范数.

证明

正定性和齐次性显然, 对三角不等式, 希望证明

$$(\int_a^b (|f_1+f_2|^2 + |f_1'+f_2'|^2)dx)^{\frac{1}{2}} \le (\int_a^b (|f_1|^2+|f_1'|^2)dx)^{\frac{1}{2}} + (\int_a^b (|f_2|^2+|f_2'|^2)dx)^{\frac{1}{2}}$$

两边平方即

$$\int_a^b(|f_1+f_2|^2+|f_1'+f_2'|^2)dx \le \int_a^b(|f_1|^2+|f_1'|^2)dx+\int_a^b(|f_2|^2+|f_2'|^2)dx+2(\int_a^b(|f_1|^2+|f_1'|^2)dx)^{\frac{1}{2}}(\int_a^b(|f_2|^2+|f_2'|^2)dx)^{\frac{1}{2}}$$

展开左式即

$$\int_a^b(|f_1f_2|+|f_1'f_2'|)dx \le (\int_a^b(|f_1|^2+|f_1'|^2)dx)^{\frac{1}{2}}(\int_a^b(|f_2|^2+|f_2'|^2)dx)^{\frac{1}{2}}$$

此即 Cauchy-Schwarz 不等式, 命题即证.

(2) 问 $(C^1[a,b],\|\cdot\|_1)$ 是否完备?

解

不完备, 比如对 $(C^1[0,1],\|\cdot\|_1)$, 考虑序列

$$C^1[0,1]\ni f_n(x)=\frac{\pi}{2}-\arctan(nx),\ x\in[0,1]$$

知

$$\|f_n-f_m\|_1=(\int_0^1(|\arctan(mx)-\arctan(nx)|^2+|\frac{m}{1+(mx)^2}-\frac{n}{1+(nx)^2}|^2)dx)^{\frac{1}{2}},\ n>m>0$$

当 $x\neq 0, m,n\to\infty$, 知

$$\arctan(mx)-\arctan(nx)=\frac{1}{1+\xi^2}\to 0,\ (mx,nx)\ni\xi\to\infty$$

同时

$$\frac{m}{1+(mx)^2}\to 0,\ \frac{n}{1+(nx)^2}\to 0$$

而 $x=0$ 时 $\|f_n-f_m\|_1\equiv 0$, 故 $\{f_n\}$ 是基本列. 但

$$\lim_{n\to\infty}f_n(x)=\begin{cases}\frac{\pi}{2}, & x=0\\ 0, & x\in(0,1]\end{cases}$$

其并不属于 $C^1[0,1]$, 故 $(C^1[a,b],\|\cdot\|_1)$ 不完备.

✍ **练习 2.25** 在 $C[0,1]$ 中, 对每一个 $f\in C[0,1]$, 令

$$\|f\|_1=(\int_0^1|f(x)|^2dx)^{\frac{1}{2}}$$

$$\|f\|_2=(\int_0^1(1+x)|f(x)|^2dx)^{\frac{1}{2}}$$

求证: $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 是 $C[0,1]$ 中的两个等价范数.

证明

显见 $\|f\|_2\ge\|f\|_1$, 又

$$\|f\|_2\le(\int_0^1 2|f(x)|^2dx)^{\frac{1}{2}}=\sqrt{2}\|f\|_1$$

故 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 是 $C[0,1]$ 中的两个等价范数.

✍ **练习 2.26★** 设 $BC[0,\infty)$ 表示 $[0,\infty)$ 上连续且有界的函数 $f(x)$ 全体, 对于每个 $f\in BC[0,\infty)$ 及 $a>0$, 定义

$$\|f\|_a=(\int_0^\infty e^{-ax}|f(x)|^2dx)^{\frac{1}{2}}$$

(1) 求证 $\|\cdot\|_a$ 是 $BC[0,\infty)$ 上的范数.

证明

根据 $BC[0,\infty)$ 的定义知 $\|\cdot\|_a$ 良定义. 正定性和齐次性显然, 对于三角不等式, 欲证

$$(\int_0^\infty e^{-ax}|f_1(x)+f_2(x)|^2dx)^{\frac{1}{2}}\le(\int_0^\infty e^{-ax}|f_1(x)|^2dx)^{\frac{1}{2}}+(\int_0^\infty e^{-ax}|f_2(x)|^2dx)^{\frac{1}{2}}$$

两边平方即

$$\int_0^{\infty} e^{-ax}|f_1(x)+f_2(x)|^2dx \le \int_0^{\infty} e^{-ax}|f_1(x)|^2dx + \int_0^{\infty} e^{-ax}|f_2(x)|^2dx + 2(\int_0^{\infty} e^{-ax}|f_1(x)|^2dx)^{\frac{1}{2}}(\int_0^{\infty} e^{-ax}|f_2(x)|^2dx)^{\frac{1}{2}}$$

展开左式整理即

$$\int_0^{\infty} e^{-ax}|f_1(x)f_2(x)|dx = \int_0^{\infty} |e^{-\frac{a}{2}x}f_1(x)|\cdots|e^{-\frac{a}{2}x}f_2(x)|dx \le (\int_0^{\infty} e^{-ax}|f_1(x)|^2dx)^{\frac{1}{2}}(\int_0^{\infty} e^{-ax}|f_2(x)|^2dx)^{\frac{1}{2}}$$

此即 Cauchy-Schwarz 不等式, 命题即证.

(2) 若 $a,b>0, a\neq b$, 求证 $\|\cdot\|_a$ 与 $\|\cdot\|_b$ 作为 $BC[0,\infty)$ 上的范数是不等价的.

证明

不妨设 $a<b$, 取 $f(x)\equiv 1$ 知

$$\|1\|_a = \frac{1}{\sqrt{a}} > \frac{1}{\sqrt{b}} = \|1\|_b$$

现需找到函数列 $\{f_n\}\subset BC[0,\infty)$ 使得 $\frac{\|f_n\|_a^2}{\|f_n\|_b^2}\to\infty$, 取

$$|f_n(x)|^2 = \begin{cases} e^{ax}, & 0\le x\le n \\ e^{ax}(n+1-x), & n\le x\le n+1 \\ 0, & x\ge n+1 \end{cases}$$

知

$$\|f_n\|_a^2 \ge \int_0^n e^{-ax}\cdot e^{ax}dx = n, \|f_n\|_b^2 \le \int_0^{+\infty} e^{-bx}\cdot e^{ax}dx = \frac{1}{b-a}$$

因而

$$\frac{\|f_n\|_a^2}{\|f_n\|_b^2} \ge \frac{n}{b-a} \to \infty,\ n\to\infty$$

命题即证.

✍ **练习 2.27** 设 $\mathscr{X}_1,\mathscr{X}_2$ 是两个 B^* 空间, $x_1\in\mathscr{X}_1$ 和 $x_2\in\mathscr{X}_2$ 的序对 (x_1,x_2) 全体构成空间 $\mathscr{X}=\mathscr{X}_1\times\mathscr{X}_2$, 并赋以范数

$$\|x\| = \max(\|x_1\|_1, \|x_2\|_2)$$

其中 $x=(x_1,x_2), x_1\in\mathscr{X}_1, x_2\in\mathscr{X}_2$, $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 分别是 $\mathscr{X}_1$ 和 $\mathscr{X}_2$ 的范数. 求证: 如果 $\mathscr{X}_1,\mathscr{X}_2$ 是 B 空间, 那么 $\mathscr{X}$ 也是 B 空间.

证明

届于 $\mathscr{X}_1,\mathscr{X}_2$ 均为 B 空间, 知 $\mathscr{X}$ 至少是 B^* 空间, 容易验证 $\|\cdot\|$ 确为范数, 下面说明完备性. 任取 $\mathscr{X}$ 中的基本列 $\{x_n\}=\{(x_n^1,x_n^2)\}, n\in\mathbb{N}, x_n^1\in\mathscr{X}_1, x_n^2\in\mathscr{X}_2$, 根据定义有

$$\|x_n-x_m\| = \max(\|x_n^1-x_m^1\|_1, \|x_n^2-x_m^2\|_2)\to 0,\ n,m\to\infty$$

这说明

$$\|x_n^1-x_m^1\|_1\to 0, \|x_n^2-x_m^2\|_2\to 0, n,m\to\infty$$

即 $\{x_n^1\},\{x_n^2\}$ 分别是 $\mathscr{X}_1,\mathscr{X}_2$ 中的基本列. 因为 $\mathscr{X}_1,\mathscr{X}_2$ 均为 B 空间, 故 $\{x_n^1\},\{x_n^2\}$ 分别在 $\mathscr{X}_1,\mathscr{X}_2$ 中收敛, 设 $x_n^1\to x^1\in\mathscr{X}_1, x_n^2\to x^2\in\mathscr{X}_2, n\to\infty$. 知 $(x^1,x^2)\in\mathscr{X}$, 故 $\{x_n\}$ 在 $\mathscr{X}$ 中收敛到 (x^1,x^2), 也即 $\mathscr{X}$ 完备, 命题即证.

✍ **练习 2.28**★♡ 设 $\mathscr{X}$ 是 B^* 空间. 求证: $\mathscr{X}$ 是 B 空间, 必须且仅须对 $\forall\{x_n\}_{n=1}^{\infty}\subset\mathscr{X}, \sum_{n=1}^{\infty}\|x_n\|<\infty \Rightarrow \sum_{n=1}^{\infty}x_n$ 收敛.

证明

当 $\mathscr{X}$ 是 B 空间, 任取 $\{x_n\}_{n=1}^{\infty}\subset\mathscr{X}$, 记 $S_n=\sum_{k=1}^{n}x_n$, 根据 $\sum_{n=1}^{\infty}\|x_n\|<\infty$, 由 Cauchy 准则知

$$\sum_{n=p}^{q}\|x_n\|\to 0,\ p,q\to\infty$$

进而知

$$0\le\|S_q-S_p\|\le\sum_{n=p}^{q}\|x_n\|\to 0,\ p,q\to\infty$$

这说明 $\{S_n\}\subset\mathscr{X}$ 是基本列, 又根据 $\mathscr{X}$ 的完备性知 $\{S_n\}$ 收敛, 也即 $\sum_{n=1}^{\infty}x_n$ 收敛.

当

$$\forall\{x_n\}_{n=1}^{\infty}\subset\mathscr{X},\sum_{n=1}^{\infty}\|x_n\|<\infty\Rightarrow\sum_{n=1}^{\infty}x_n\text{收敛}$$

任取基本列 $\{y_n\}_{n=1}^{\infty}\subset\mathscr{X}$, 根据基本列的定义有

$$\|y_m-y_n\|\to 0,\ m,n\to\infty$$

可以选取 $\{y_n\}$ 的一个子列 $\{y_{n_k}\}$ 满足

$$\|y_{n_{k+1}}-y_{n_k}\|\le\frac{1}{p^2},\ p\in\mathbb{N}$$

考虑 $\zeta_k=y_{n_k}-y_{n_{k-1}},k\ge 2,\zeta_1=y_{n_1}$, 则 $y_{n_k}=\sum_{i=1}^{k}\zeta_i$, 知

$$\sum_{k=1}^{\infty}\|\zeta_k\|=\|y_{n_1}\|+\sum_{k=2}^{\infty}\|y_{n_k}-y_{n_{k-1}}\|\le\|y_{n_1}\|+\sum_{p=1}^{\infty}\frac{1}{p^2}<\infty$$

根据条件知 $\sum_{k=1}^{\infty}\zeta_k$ 收敛, 也即 $\{y_{n_k}\}$ 收敛, 根据练习(2.9)知 $\{y_n\}$ 收敛, 因而 $\mathscr{X}$ 是 B 空间.

✍ **练习 2.29**♡ 记 $[a,b]$ 上次数不超过 n 的多项式全体为 $\mathbb{P}_n$. 求证: $\forall f(x)\in C[a,b],\exists P_0(x)\in\mathbb{P}_n$, 使得

$$\max_{a\le x\le b}|f(x)-P_0(x)|=\min_{P\in\mathbb{P}_n}\max_{a\le x\le b}|f(x)-P(x)|$$

也就是说, 如果用所有次数不超过 n 的多项式去对 $f(x)$ 一致逼近, 那么 $P_0(x)$ 是最佳的.

证明

根据定理2.4.3, 显见 $C[a,b]$ 是 B^* 空间, 因而对给定的向量组 $1,x,\cdots,x^n$, 对任意的 $f\in C[a,b]$ 总存在最佳逼近系数 $\lambda_0,\cdots,\lambda_n$ 满足

$$\max_{a\le x\le b}|f(x)-\sum_{k=0}^{n}\lambda_kx^k|=\|f-P_\lambda\|=\min_{a\in\mathbb{R}^n}\|f-P_a\|,P_\lambda=\sum_{k=0}^{n}\lambda_kx^k,P_a=\sum_{k=0}^{n}a_kx^k$$

命题即证.

✍ **练习 2.30** 在 $\mathbb{R}^2$ 中, 对 $\forall x=(x_1,x_2)\in\mathbb{R}^2$, 定义范数

$$\|x\|=\max(|x_1|,|x_2|)$$

并设 $e_1=(1,0),x_0=(0,1)$. 求 $a\in\mathbb{R}$ 适合

$$\|x_0-ae_1\|=\min_{\lambda\in\mathbb{R}}\|x_0-\lambda e_1\|$$

并问这样的 a 是否唯一? 请对结果做出几何解释.

解

此即要求

$$\max(|a|,1)=\min_{\lambda\in\mathbb{R}}\max(|\lambda|,1)$$

取用 $a\in[-1,1]$ 即可. 几何上, 结合练习(2.22)的图知当 $a\in[-1,1]$ 时, x_0-ae_1 落在单位圆上, 且全体 $x_0-\lambda e_1$ 的模长均不少于单位圆上元素的模长.

✍ **练习 2.31★** 求证: 范数的严格凸性等价于下列条件:

$$\|x+y\| = \|x\| + \|y\| (\forall x \neq \theta, y \neq \theta) \Rightarrow x = cy \ (c > 0)$$

证明

当

$$\|x+y\| = \|x\| + \|y\| (\forall x \neq \theta, y \neq \theta) \Rightarrow x = cy \ (c > 0)$$

取 $x \neq y, \|x\| = \|y\| = 1$, 并取 $\alpha, \beta > 0, \alpha + \beta = 1$, 知

$$\|\alpha x + \beta y\| \leq \alpha\|x\| + \beta\|y\| = 1$$

现在说明等号不成立, 这是因为如若等号成立, 即

$$\|\alpha x + \beta y\| \leq \alpha\|x\| + \beta\|y\|$$

此时应有 $\alpha x = c\beta y, c > 0$, 取范数知 $\alpha = c\beta \neq 0$, 也即 $x = y$, 矛盾! 故

$$\|\alpha x + \beta y\| < 1$$

当 $\|\cdot\|$ 严格凸, 即

$$\forall x, y \in \mathscr{X}, x \neq y (\|x\| = \|y\| = 1 \Rightarrow \|\alpha x + \beta y\| < 1, \forall \alpha, \beta > 0, \alpha + \beta = 1)$$

将 x 换为 $\frac{x}{\|x\|}$, y 换为 $\frac{y}{\|y\|}$ 知

$$\forall \alpha, \beta > 0, \alpha + \beta = 1 (\|\frac{\alpha x}{\|x\|} + \frac{\beta y}{\|y\|}\| < 1)$$

取 $\alpha = \frac{\|x\|}{\|x\|+\|y\|}, \beta = \frac{\|y\|}{\|x\|+\|y\|}$ 知

$$\|x+y\| < \|x\| + \|y\|, \ \forall x, y \in \mathscr{X}, \frac{x}{\|x\|} \neq \frac{y}{\|y\|}$$

这说明

$$\|x+y\| = \|x\| + \|y\| (\forall x \neq \theta, y \neq \theta) \Rightarrow \frac{x}{\|x\|} = \frac{y}{\|y\|}$$

取 $c = \frac{\|x\|}{\|y\|}$ 即得命题.

✍ **练习 2.32★** 设 $\mathscr{X}$ 是赋范线性空间, 函数 $\varphi : \mathscr{X} \to \mathbb{R}$ 称为凸的, 如果不等式

$$\varphi(\lambda x + (1-\lambda)y) \leq \lambda\varphi(x) + (1-\lambda)\varphi(y) \ (\forall 0 \leq \lambda \leq 1)$$

成立. 求证: 凸函数的局部极小值必然是全空间最小值.

证明

用反证法, 若 $\varphi(x_0)$ 作为局部极小值并非全空间最小值, 另设全空间最小值为 $\varphi(x_1), x_1 \neq x_0$, 根据局部极小的定义知

$$\exists \delta > 0 \forall x \in B(x_0, \delta)(\varphi(x) \geq \varphi(x_0))$$

特别设 $x_2 \in B(x_0, \delta)$ 满足 $\exists \lambda \in [0,1](x_2 = \lambda x_0 + (1-\lambda)x_1)$, 根据 $\mathscr{X}$ 是线性空间知这样的 x_2 必可取到, 现在根据凸性:

$$\varphi(x_2) = \varphi(\lambda x_0 + (1-\lambda)x_1) \leq \lambda\varphi(x_0) + (1-\lambda)\varphi(x_1)$$

另一方面 $\varphi(x_2) \geq \varphi(x_0)$, 有

$$(1-\lambda)\varphi(x_0) \leq (1-\lambda)\varphi(x_1) \Rightarrow \varphi(x_0) \leq \varphi(x_1)$$

但根据假设应有 $\varphi(x_0) > \varphi(x_1)$, 矛盾! 命题即证.

✍ **练习 2.33** 设 $(\mathscr{X}, \|\cdot\|)$ 是一赋范线性空间, M 是 $\mathscr{X}$ 的有限维子空间, $\{e_1, e_2, \cdots, e_n\}$ 是 M 的一组基, 给定 $g \in \mathscr{X}$,

引进函数 $F:\mathbb{K}^n\to\mathbb{R}$, 对 $\forall c=(c_1,c_2,\cdots,c_n)\in\mathbb{K}^n$ 规定

$$F(c)=F(c_1,c_2,\cdots,c_n)=\|\sum_{i=1}^n c_ie_i-g\|$$

(1) 求证: F 是一个凸函数.

证明

此即证明对任意的 $\lambda\in[0,1]$ 有:

$$F(\lambda c+(1-\lambda)\zeta)=\|\sum_{i=1}^n(\lambda c_i+(1-\lambda)\zeta_i)e_i-g\|\le\lambda\|\sum_{i=1}^n c_ie_i-g\|+(1-\lambda)\|\sum_{i=1}^n\zeta_ie_i-g\|=\lambda F(c)+(1-\lambda)F(\zeta)$$

注意

$$\begin{aligned}\|\sum_{i=1}^n(\lambda c_i+(1-\lambda)\zeta_i)e_i-g\|&=\|(\sum_{i=1}^n\lambda c_i-\lambda g)+(\sum_{i=1}^n(1-\lambda)\zeta_i-(1-\lambda)g)\|\\&\le\lambda\|\sum_{i=1}^n c_ie_i-g\|+(1-\lambda)\|\sum_{i=1}^n\zeta_ie_i-g\|=\lambda F(c)+(1-\lambda)F(\zeta)\end{aligned}$$

命题即证.

(2) 若 $F(c)$ 的最小值点是 $c=(c_1,c_2,\cdots,c_n)$, 求证:

$$f:=\sum_{i=1}^n c_ie_i$$

给出 g 在 M 中的最佳逼近元.

证明

根据练习(2.32)知

$$F(c)=\min_{\zeta\in\mathbb{K}^n}F(\zeta)=\min_{\zeta\in\mathbb{K}^n}\|\sum_{i=1}^n\zeta_ie_i-g\|$$

这恰是最佳逼近元的定义.

练习 2.34★♡ 设 $\mathscr{X}$ 是 B^* 空间, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间, 假定 $\exists c\in(0,1)$, 使得

$$\inf_{x\in\mathscr{X}_0}\|y-x\|\le c\|y\|\ (\forall y\in\mathscr{X})$$

求证: $\mathscr{X}_0$ 在 $\mathscr{X}$ 中稠密.

证明

用反证法, 如若 $\overline{\mathscr{X}_0}\subsetneq\mathscr{X}$, 根据 Riesz 引理知

$$\forall\varepsilon\in(0,1)\exists y\in\mathscr{X}(\|y\|=1\wedge\|y-x\|\ge1-\varepsilon,\forall x\in\overline{\mathscr{X}_0})$$

在题目条件中用 $\frac{y}{\|y\|}$ 替换 y 有

$$\inf_{x\in\mathscr{X}_0}\|y-x\|\le c,\forall y\in\{y:\|y\|=1\}$$

现取 $\varepsilon_n=\frac{1}{n}$, 可产生序列 $\{y_n\}\subset\{y:\|y\|=1\}$ 满足

$$\liminf_{n\to\infty}\|y_n-x\|\ge1-\liminf_{n\to\infty}\frac{1}{n}=1,\ \forall x\in\mathscr{X}_0$$

这与条件断言 c 的存在相悖! 命题即证.

另法

根据下确界的定义, 任意取定 $y\in\mathscr{X}$, 有:

$$\forall\varepsilon>0\exists x_1^y\in\mathscr{X}_0(\|y-x_1^y\|\le\varepsilon+\inf_{x\in\mathscr{X}_0}\|y-x\|)$$

因为 $\mathscr{X}$ 是 B^* 空间, 故 $y-x_1^y\in\mathscr{X}$, 进而沿用上述记号, 根据下确界的定义有:

$$\exists x_2^y\in\mathscr{X}_0(\|(y-x_1^y)-x_2^y\|\le\varepsilon+\inf_{x\in\mathscr{X}_0}\|(y-x_1^y)-x\|)$$

依题知

$$\inf_{x\in\mathscr{X}_0}\|(y-x_1^y)-x\|\le c\|y-x_1^y\|\le c\varepsilon+c\inf_{x\in\mathscr{X}^0}\|y-x\|$$

得到

$$\|y-x_1^y-x_2^y\|\le\varepsilon+c\varepsilon+c\inf_{x\in\mathscr{X}^0}\|y-x\|$$

重复上述过程, 在第 n 次知:

$$\exists x_n^y\in\mathscr{X}_0(\|y-\sum_{k=1}^{n}x_k^y\|\le\varepsilon\sum_{k=1}^{n-1}c^k+c^{n-1}\inf_{x\in\mathscr{X}^0}\|y-x\|)$$

当 $n\to\infty$ 时, 由 $c\in(0,1)$ 是常数知 $c^{n-1}\to 0,\sum_{k=1}^{n-1}c^k<\infty$, 故由 ε 的任意性即知 $\mathscr{X}_0$ 中的序列 $\{\sum_{k=1}^{n}x_k^y\}_{n=1}^{\infty}$ 趋向 y, 进而由 y 的任意性即得 $\mathscr{X}_0$ 在 $\mathscr{X}$ 中稠密.

✍ **练习 2.35**★ 设 C_0 表示以 0 为极限的实数列全体, 并在 C_0 中赋以范数

$$\|x\|=\max_{n\ge1}|\xi_n|\ (\forall x=(\xi_1,\xi_2,\cdots,\xi_n)\in C_0)$$

又设 $M:=\{x=\{\xi_n\}_{n=1}^{\infty}\in C_0:\sum_{n=1}^{\infty}\frac{\xi_n}{2^n}=0\}$.

(1) 求证: M 是 C_0 的闭线性子空间.

证明

M 的线性性显然, 现设 $x_0=(\xi_0^1,\cdots,\xi_0^n,\cdots)\in M'$, 根据定义知存在序列 $\{x_k\}=\{(\xi_k^1,\cdots,\xi_k^n,\cdots)\}\subset M$ 满足

$$\|x_k-x_0\|\to 0,\ n\to\infty$$

根据 $\|\cdot\|$ 的定义知此时

$$\xi_k^i\to\xi_0^i,\ i=1,\cdots,n,\cdots,k\to\infty$$

考虑 $\sum_{n=1}^{p}\frac{\xi_0^n}{2^n}$, 既然 p 有限, 知

$$\forall\varepsilon>0\exists K>0\forall k>K(|\xi_k^i-\xi_0^i|<\varepsilon,i=1,\cdots,p)$$

得到

$$\xi_k^i-\varepsilon<\xi_0^i<\xi_k^i+\varepsilon,\ i=1,\cdots,p$$

因而

$$\sum_{n=1}^{p}\frac{\xi_k^n}{2^n}-(1-\frac{1}{2^p})\varepsilon<\sum_{n=1}^{p}\frac{\xi_0^n}{2^n}<\sum_{n=1}^{p}\frac{\xi_k^n}{2^n}+(1-\frac{1}{2^p})\varepsilon$$

令 $p\to\infty$ 知

$$\sum_{n=1}^{\infty}\frac{\xi_k^n}{2^n}-\varepsilon=-\varepsilon\le\sum_{n=1}^{\infty}\frac{\xi_0^n}{2^n}\le\varepsilon=\sum_{n=1}^{\infty}\frac{\xi_k^n}{2^n}+\varepsilon$$

由 ε 的任意性即得

$$\sum_{n=1}^{\infty}\frac{\xi_0^n}{2^n}=0\Rightarrow x_0\in M$$

故 $M'\subset M$, 因而 M 是 C_0 的闭线性子空间.

(2) 设 $x_0=(2,0,\cdots,0,\cdots)$, 求证:

$$\inf_{z\in M}\|x_0-z\|=1,$$

但 $\forall y\in M$ 有 $\|x_0-y\|>1$.

证明

取

$$z_n = (1 - \frac{1}{2^n}, \underbrace{-1, \cdots, -1}_{n}, 0, \cdots) \in M, \ n \in \mathbb{N}$$

知

$$\|x_0 - z_n\| = 1 + \frac{1}{2^n} \geq 1$$

且

$$\lim_{n\to\infty} \|x_0 - z_n\| = 1$$

因而

$$\inf_{z\in M} \|x_0 - z\| = 1$$

但若存在 $y \in M$ 使得 $\|x_0 - y\| \leq 1$, 设 $y = (y^1, \cdots, y^n, \cdots)$, 知

$$\max(|2 - y^1|, |y^2|, \cdots, |y^n|, \cdots) \leq 1$$

得到 $y^1 \geq 1$, 因而

$$|\sum_{k=1}^{\infty} \frac{y^k}{2^k}| \geq \frac{y^1}{2} - |\sum_{k=2}^{\infty} \frac{y^k}{2^k}| \geq \frac{1}{2} - \sum_{k=2}^{\infty} \frac{|y^k|}{2^k} \geq 0$$

等号全部成立当且仅当 $y^1 = 1, y^k = -1, k = 2, 3, \cdots$, 但这与 $y^k \to 0, k \to \infty$ 相悖! 命题即证.

注 该练习说明对于无穷维闭线性子空间来说, 给定其外一点 x_0, 未必能在其上找到一点 y 满足

$$\|x_0 - y\| = \inf_{z\in M} \|x_0 - z\|$$

也即给定 $x \notin M$, 未必能在 M 上找到最佳逼近元.

练习 2.36★ 设 $\mathscr{X}$ 是 B^* 空间, M 是 $\mathscr{X}$ 的有限维真子空间. 求证: $\exists y \in \mathscr{X}, \|y\| = 1$, 使得

$$\|y - x\| \geq 1 \ (\forall x \in M)$$

证明

取 $X \subset \mathscr{X}$ 是有限维真子空间, 满足 $M \subset X$, 知 X 中的单位球 $\mathbb{S}$ 列紧, 根据 Riesz 引理知对任意的 $\varepsilon > 0$, 存在 $y \in \mathbb{S}$ 使得

$$\|y - x\| \geq 1 - \varepsilon, \ \forall x \in M$$

取 $\varepsilon_n = \frac{1}{n}$, 产生序列 $\{y_n\} \subset \mathbb{S}$, 取极限有

$$\lim_{n\to\infty} \|y_n - x\| \geq 1, \ \forall x \in M$$

既然 $\mathbb{S}$ 是有限维 B^* 空间的单位球, 由 Riesz 定理2.4.5知其列紧, 因而 $\{y_n\}$ 有收敛子列 $\{y_{n_k}\}$, 设 $y_{n_k} \to y, k \to \infty$, y 即欲求, 命题得证.

练习 2.37 若 f 是定义在区间 $[0, 1]$ 上的复值函数, 定义

$$\omega_\delta(f) = \sup\{|f(x) - f(y)| : \forall x, y \in [0, 1], |x - y| \leq \delta\}.$$

如果 $0 < \alpha \leq 1$ 对应的 Lipschitz 空间 $\mathrm{Lip}\,\alpha$, 由满足

$$\|f\| := |f(0)| + \sup_{\delta>0}\{\delta^{-\alpha}\omega_\delta(f)\} < \infty$$

的一切 f 组成, 并以 $\|f\|$ 为范数. 又设

$$\mathrm{lip}\,\alpha := \{f \in \mathrm{Lip}\,\alpha : \lim_{\delta\to 0} \delta^{-\alpha}\omega_\delta(f) = 0\}.$$

求证: $\mathrm{Lip}\,\alpha$ 是 B 空间, 而且 $\mathrm{lip}\,\alpha$ 是 $\mathrm{Lip}\,\alpha$ 的闭子空间.

证明[17]

易证 $\|\cdot\|$ 确为范数, 且 $\operatorname{Lip}\alpha$ 是线性空间, 下面证明 $\operatorname{Lip}\alpha$ 完备. 设 $\{f_n\}$ 是 $\operatorname{Lip}\alpha$ 中的基本列, 根据定义知

$$\|f_m - f_n\| \to 0,\ m, n \to \infty$$

代入 $\|\cdot\|$ 的定义得到:

$$|f_m(0) - f_n(0)| + \sup_{\delta>0}\{\delta^{-\alpha}\omega_\delta(f_m - f_n)\} \to 0,\ m, n \to \infty$$

根据极限定义, 此即

$$\forall\varepsilon > 0\exists N > 0\forall m > n > N(|f_m(0) - f_n(0)| + \sup_{\delta>0}\{\delta^{-\alpha}\omega_\delta(f_m - f_n)\} < \varepsilon)$$

故分别有

$$|f_m(0) - f_n(0)| < \varepsilon,\ \sup_{\delta>0}\{\delta^{-\alpha}\omega_\delta(f_m - f_n)\} < \varepsilon \tag{2.22}$$

上第二式说明

$$\forall\delta > 0(\delta^{-\alpha}\omega_\delta(f_m - f_n) < \varepsilon)$$

再代入 $\omega_\delta(f)$ 的定义有

$$\forall\delta > 0(\delta^{-\alpha}\sup\{|f_m(x) - f_n(x) - f_m(y) + f_n(y)| : \forall x, y \in [0, 1], |x - y| \le \delta\} < \varepsilon)$$

写开得到

$$\forall\delta > 0\forall x, y(|x - y| \le \delta \Rightarrow \delta^{-\alpha}|f_m(x) - f_n(x) - f_m(y) + f_n(y)| < \varepsilon) \tag{2.23}$$

特别取 $y = 0$, 则有

$$|f_m(x)-f_n(x)|-|f_m(0)-f_n(0)| \le |f_m(x)-f_n(x)-(f_m(0)-f_n(0))| < \varepsilon\delta^\alpha \Rightarrow |f_m(x)-f_n(x)| \le \varepsilon(1+\delta^\alpha),\ \forall\delta > 0, |x| \le \delta$$

注意前面 N 本身的选取与 x 没有任何关系, 这说明至少在 $|x| \le \delta$ 上, $\{f_n\} \subset \operatorname{Lip}\alpha \subset C[-\delta, \delta]$ 是一致收敛度量意义下的基本列, 从而根据 $C[-\delta, \delta]$ 在一致收敛度量下的完备性知存在 $f \in C[-\delta, \delta]$, 使得 $f_n \to f(n \to \infty)$.

再来证明 $f \in \operatorname{Lip}\alpha$. 根据前述 $\{f_n\}$ 的一致收敛性, 在(2.23)式中令 $m \to \infty$ 得到

$$|f_n(x) - f(x) + f(y) - f_n(y)| < \varepsilon\delta^\alpha,\ \forall\delta > 0, |x - y| \le \delta \tag{2.24}$$

同样在(2.22)第一式中令 $m \to \infty$ 得

$$|f_n(0) - f(0)| < \varepsilon \tag{2.25}$$

(2.24),(2.25)式说明 $\|f - f_n\| < \infty$, 从而

$$\|f\| \le \|f - f_n\| + \|f_n\| < \infty$$

故 $f \in \operatorname{Lip}\alpha$, 因而 $\operatorname{Lip}\alpha$ 是 B 空间.

显见 $\operatorname{lip}\alpha$ 也是线性空间, 且 $\operatorname{lip}\alpha \subset \operatorname{Lip}\alpha$, 故只需说明 $\operatorname{lip}\alpha$ 是闭的. 任取 $f \in \overline{\operatorname{lip}\alpha}$, 根据闭包的定义知存在 $\{f_n\} \subset \operatorname{lip}\alpha$ 使得 $\|f - f_n\| \to 0$, 与(2.22)式的导出同理, 知此时

$$\delta^{-\alpha}\omega_\delta(f - f_n) \to 0,\ n \to \infty$$

又因为 $\{f_n\} \subset \operatorname{lip}\alpha$, 故

$$\lim_{\delta\to 0}\delta^{-\alpha}\omega_\delta(f_n) = 0$$

从而

$$\delta^{-\alpha}\omega_\delta(f) \le \delta^{-\alpha}\omega_\delta(f - f_n) + \delta^{-\alpha}\omega_\delta(f_n) \to 0,\ n \to \infty, \delta \to 0$$

这便说明 $f \in \operatorname{lip}\alpha$, 也即 $\overline{\operatorname{lip}\alpha} = \operatorname{lip}\alpha$, 进而 $\operatorname{lip}\alpha$ 是闭的, 命题即证.

[17] 23.6.25 想法: $\operatorname{Lip}\alpha$ 是 $C[0, 1]$ 的子空间, 是否可以考虑先断言 $(C[0, 1], \|\cdot\|)$ 完备, 在 $C[0, 1]$ 中求出基本列的极限函数, 再证明它在 $\operatorname{Lip}\alpha$ 中? 个人感觉不能从列紧空间入手.
23.8.31 评论: 前面的想法在思想上是对的, 就是证明 $(C[0, 1], \|\cdot\|)$ 完备这点太复杂了.

✍ **练习 2.38★** 设有商空间 $\mathscr{X}/\mathscr{X}_0$.

(1) 设 $[x]\in\mathscr{X}/\mathscr{X}_0$, 求证: 对 $\forall x\in[x]$, 有

$$\inf_{z\in\mathscr{X}_0}\|x-z\|=\|[x]\|_0.$$

证明

根据定义知

$$\|[x]\|_0=\inf_{x\in[x]}\|x\|$$

而当 $x\in[x], z\in\mathscr{X}_0$, 自然有 $x-z\in[x]$, 故

$$\inf_{z\in\mathscr{X}_0}\|x-z\|=\inf_{x\in[x]}\|x\|=\|[x]\|_0$$

(2) 定义映射 $\varphi:\mathscr{X}\to\mathscr{X}/\mathscr{X}_0$ 为

$$\varphi(x)=[x]:=x+\mathscr{X}_0\ (\forall x\in\mathscr{X})$$

求证: φ 是连续线性映射.

证明

线性性显见, 下证连续性. 知

$$\|\varphi(y)-\varphi(x)\|=\|[y-x]\|_0=\inf_{y-x\in[y-x]}\|y-x\|\le\|y-x\|$$

因而 φ 是 Lipschitz 映射, 进而是连续映射.

(3) $\forall[x]\in\mathscr{X}/\mathscr{X}_0$, 求证: $\exists x\in\mathscr{X}$, 使得

$$\varphi(x)=[x],\ \text{且}\|x\|\le2\|[x]\|_0$$

证明

根据 $\|\cdot\|_0$ 作为下确界的定义知

$$\forall\varepsilon>0\exists x\in\mathscr{X}(\|x\|\le(1+\varepsilon)\|[x]\|_0)$$

取 $\varepsilon=1$ 即得欲证.

(4) 设 $\mathscr{X}=C[0,1], \mathscr{X}_0=\{f\in\mathscr{X}:f(0)=0\}$, 求证:

$$\mathscr{X}/\mathscr{X}_0\cong\mathbb{K}$$

其中记号"$\cong$" 表示等距同构.

证明

考虑映射

$$\varphi:\mathscr{X}/\mathscr{X}_0\to\mathbb{K},[f]\mapsto f(0)$$

显见 φ 是满的, 再证明 φ 是单的, 这是因为任取 $f_1,f_2\in\mathscr{X}$ 满足 $f_1(0)=f_2(0)$, 知 $f_1-f_2\in\mathscr{X}_0$, 因而 $[f_1-f_2]=\theta$. 故 φ 是双射.

再证明 φ 等距, 知

$$\|[f]\|_0=\inf_{f\in[f]}|f(0)|=|f(0)|$$

最后一个等号成立是因为 $[f]$ 代表所有 $f(0)$ 等于某一特定值的函数类. 故 φ 是等距同构, $\mathscr{X}/\mathscr{X}_0\cong\mathbb{K}$.

2.4.4 补充: 拓扑空间

本节选自 [**PGC**], 在 [**ZL**] 中仅仅给出了范数诱导拓扑下的讨论, 但并未给出一般拓扑的情况, 而后者在抽象分析中弥足轻重, 故在此补充. 更详细的教材可参见 [**You**].

2.4.4.1 基本定义

定义 2.4.20 (拓扑空间)

拓扑空间是指一对 $(X, \mathcal{O})$, 其中 X 是一个集合, $\mathcal{O}$ 是 $\mathcal{P}(X)$[a]的满足下述条件的一个子集:

1. 对任何一族 $\{O_i\}_{i\in I}$, 其中 $O_i \in \mathcal{O}$, 满足 $\bigcup_{i\in I} O_i \in \mathcal{O}$(其中指标集 I 可以是有限集, 也可为可数或不可数的无限集);
2. 对任意有限个 $\{O_j\}_{j=1}^n$, 其中 $O_j \in \mathcal{O}$, 满足 $\bigcap_{j=1}^n O_j \in \mathcal{O}$;
3. $X \in \mathcal{O}, \varnothing \in \mathcal{O}$.

[a]即 X 的幂集.

♣

定义 2.4.21 (开集, 闭集)

如果 $(X, \mathcal{O})$ 是一个拓扑空间, 称集合 X 被赋予一个拓扑 (对应于 $\mathcal{P}(X)$ 的子集 $\mathcal{O}$), X 中属于 $\mathcal{O}$ 的子集称为关于这个拓扑的开集; 设 F 是 X 的子集, 若 $X\backslash F$ 是开集, 则称 F 是关于这个拓扑的闭集.

♣

根据 de Morgan 律, 任意闭集族 $\{F_i\}_{i\in I}$ 的交集 $\bigcap_{i\in I} F_i$ 是闭集, 任意有限个闭集 $\{F_j\}_{j=1}^\infty$ 的并集 $\bigcup_{j=1}^n F_j$ 是闭集. X 和 $\varnothing$ 都是闭集.

定义 2.4.22 (邻域)

在拓扑空间 $(X, \mathcal{O})$ 中, 点 $x \in X$ 的邻域是 X 的包含点 x 的某开子集的任意子集. 点 $x \in X$ 的所有邻域组成的集合记作 $\mathcal{V}(x)$.

♣

例 2.41($\mathbb{R}$ 上的通常拓扑) 给定点 $a, b \in \mathbb{R}, a < b$, 记 $(a, b) = \{y \in \mathbb{R} : a < y < b\}$, 考虑拓扑空间 $(\mathbb{R}, \mathcal{O})$, 其中

$$\mathbb{R} \supset O \in \mathcal{O} \Leftrightarrow \forall x \in O \exists a < b(x \in (a, b) \wedge (a, b) \in O).$$

称该拓扑为 $\mathbb{R}$ 上的通常拓扑, (a, b) 作为通常拓扑下的开集称为开区间.

在通常拓扑下, 可以得到下述已在实变函数课程中证明的开集构造定理.

定理 2.4.10 (开集构造)

设 $\mathbb{R}$ 赋予通常拓扑, 则 $\mathbb{R}$ 中任何非空开子集可表示为可数个互不相交的有界或无界的开区间的并集.

♡

定义 2.4.23 (内部, 闭包, 边界)

设 $(X, \mathcal{O})$ 为一个拓扑空间, A 是 X 的一个子集. 称包含于 A 的所有开集的并集为 A 的内部, 记作 A°, 即:

$$A^\circ := \{x \in X : A \in \mathcal{V}(x)\}.$$

A 的闭包是包含 A 的所有闭子集的交集, 记作 $\overline{A}$, 即

$$\overline{A} := \{x \in X : \forall V \in \mathcal{V}(x)(V \cap A \neq \varnothing)\} = X\backslash(X\backslash A)^\circ.$$

A 的边界定义为 $\overline{A}$ 和 $\overline{X\backslash A}$ 的交集, 记作 ∂A, 即

$$\partial A := \{x \in X : \forall V \in \mathcal{V}(x)(V \cap A \neq \varnothing \wedge V \cap (X\backslash A) \neq \varnothing)\}.$$

♣

注意 $\partial A = \overline{A}\backslash A^\circ$.

定义 2.4.24 (支集)

设 (X, O) 为拓扑空间, 实值或复值函数 $f : X \to \mathbb{K}$ 的支集是指集合

$$\operatorname{supp} f := \overline{\{x \in X : f(x) \neq 0\}}.$$

♣

定义 2.4.25 (稠密)

设 A 为拓扑空间 (X, O) 的子集, 若 $\overline{A} = X$, 则称 A 在 X 中稠密. ♣

定义 2.4.26 (可分)

拓扑空间 (X, O) 称为可分的, 是指它包含一个可数的稠密子集, 即存在 $x_n \in X, n \geq 0$, 使得 $\overline{\bigcup_{n=0}^{\infty}\{x_n\}} = X$. ♣

定义 2.4.27 (Hausdorff 空间与 Hausdorff 拓扑)

设 (X, O) 为拓扑空间, 如果对任意两个不同的点 $x \in X$ 和 $y \in X$, 存在 x 的邻域 V 和 y 的邻域 W, 使得 $V \cap W = \emptyset$, 那么称 X 为 Hausdorff 空间或具有 Hausdorff 拓扑. ♣

定义 2.4.28 (正规)

拓扑空间 (X, O) 称为正规的, 是指对 X 中任意给定的两个互不相交的闭子集 F_1, F_2, 存在互不相交的开子集 O_1, O_2, 使得 $F_1 \subset O_1, F_2 \subset O_2$. 显见, 正规空间是 Hausdorff 空间. ♣

定义 2.4.29 (极限)

设 X 为拓扑空间, $x_n \in X, n \geq 0$. 若存在 $x \in X$, 对 x 的任一邻域 V, 存在整数 $n_0 = n_0(V) \geq 0$, 使得当 $n \geq n_0$ 时 $x_n \in V$, 则称 x 为序列 $\{x_n\}_{n=0}^{\infty}$ 的极限. 用记号

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

表示 x 为收敛序列 $\{x_n\}_{n=0}^{\infty}$ 的极限. ♣

定义 2.4.30 (点态收敛)

设 X 为一集合, (Y, O) 为 Hausdorff 拓扑空间. 称映射 $f_n : X \to Y$ 的序列 $\{f_n\}_{n=0}^{\infty}$ 点态收敛于映射 $f : X \to Y$, 如果对每个 $x \in X$, 当 $n \to \infty$ 时, 都有 $f_n(x) \to f(x)$. ♣

定义 2.4.31 (诱导拓扑)

设 (X, O) 为拓扑空间, A 为 X 的一个子集, 用 O_A 表示 $\mathcal{P}(A)$ 的一个子集, 其元素为 A 的所有形如 $O_A = O \cap A$ 的子集, 其中 $O \in O$. 这样一来, (A, O_A) 也是一个拓扑空间, 称这个拓扑是由 (X, O) 的拓扑在 A 上诱导的拓扑, 在无歧义的情况下简称为诱导拓扑. ♣

需要注意 (X, O) 和 (A, O_A) 的拓扑性质不一定相同.

定义 2.4.32 ((有限) 乘积拓扑)

设 $(X_j, O_j)(1 \leq j \leq n)$ 均为拓扑空间, $X := \prod_{j=1}^{n} X_j$ 表示集合 $X_j(1 \leq j \leq n)$ 的 (有限) 乘积, 记

$$O := \{O \in \mathcal{P}(X) : \forall x \in O \exists O_j \in O_j (1 \leq j \leq n)(x \in O_1 \times \cdots \times O_n \wedge O_1 \times \cdots \times O_n \subset O)\}.$$

这样, (X, O) 也是一个拓扑空间, 称 X 上的拓扑是相应于 $\mathcal{P}(X_j)$ 的子集 $O_j(1 \leq j \leq n)$ 的乘积拓扑. ♣

定义 2.4.33 ((一般) 乘积拓扑)

对任意一族拓扑空间 $(X_i, \mathcal{O}_i), i \in I$, 乘积集合 $X = \prod_{i\in I} X_i$ 上的乘积拓扑定义如下: 在该拓扑下子集 $O \subset X$ 称为开集是指, 对任何 $x \in O$, 存在开集 $O_i \in \mathcal{O}_i$ 的有限族 $\{O_i\}_{i\in J(x)}$, 使得

$$x \in (\prod_{i\in J(x)} O_i) \times (\prod_{i\in I(x)} X_i) \text{ 且 } (\prod_{i\in J(x)} O_i) \times (\prod_{i\in I(x)} X_i) \subset O$$

其中 $I(x) = I - J(x)$.

2.4.4.2 拓扑空间中的连续性

定义 2.4.34 (在某点处连续)

设 X, Y 均为拓扑空间, $f : X \to Y$ 为从 X 到 Y 的映射, $x \in X$. 若对 Y 中 $f(x)$ 的任何邻域 V, 存在 X 中 x 的一个邻域 U, 使得 U 在 f 下的像 $f(U) \subset V$, 则称 f 在点 x 处连续.

Hausdorff 空间之间连续映射的一个基本性质是: 连续映射把收敛序列映成收敛序列.

定理 2.4.11

设 X, Y 为两个 Hausdorff 空间, 映射 $f : X \to Y$ 在点 $x \in X$ 处连续, 则对 X 中收敛于点 x 的任何序列 $\{x_n\}_{n=0}^{\infty}$, 序列 $\{f(x_n)\}_{n=0}^{\infty}$ 在 Y 中收敛于 $f(x)$.

注 该定理的逆命题在 X 为距离空间时是成立的.

定理 2.4.12 (复合映射的连续性)

设 X, Y, Z 为三个拓扑空间, 映射 $f : X \to Y$ 在点 $x \in X$ 处连续, 映射 $g : Y \to Z$ 在点 $f(x) \in Y$ 处连续, 则复合映射 $g \circ f : X \to Z$ 在 x 处连续.

定义 2.4.35 (连续映射)

若映射 $f : X \to Y$ 在 X 上每一点处均连续, 则称 f 是 X 上的连续映射, X 到 Y 的所有连续映射组成的集合记作 $C(X, Y)$, 当 $Y = \mathbb{R}$ 时简记为 $C(X)$.

定理 2.4.13

设 X, Y 为两个拓扑空间, 映射 $f : X \to Y$ 是连续的充要条件为 Y 中任何开集在 f 下的逆像是 X 中的开集, 也等价于 Y 中任何闭集在 f 下的逆像是 X 中的闭集.

定义 2.4.36 (同胚)

设 X 和 Y 为两个拓扑空间, 称映射 $f : X \to Y$ 是 X 到 Y 上的同胚, 如果 f 是双射, 且 $f \in C(X, Y), f^{-1} \in C(Y, X)$. 如果存在拓扑空间 X 到 Y 上的同胚映射, 则称 X 到 Y 是同胚的.

定理 2.4.14

设 X, Y 为两个拓扑空间, $f \in C(X, Y)$ 是一个双射. 则 f 是 X 到 Y 上的同胚当且仅当 X 中任何开子集或闭子集在 f 下的直接像是 Y 中的开子集或相应的闭子集.

在乘积空间中有下述定理.

定理 2.4.15

设 $X_j(1 \le j \le n)$ 和 Y 均为拓扑空间, 乘积空间 $X := \prod_{j=1}^n X_j$ 上装备乘积拓扑, 映射 $f: X \to Y$ 在点 $a = (a_1, \cdots, a_n) \in X$ 连续, 则对每个 $1 \le j \le n$, 映射

$$X_j \ni x_j \mapsto f(a_1, \cdots, a_{j-1}, x_j, a_{j+1}, \cdots, a_n) \in Y$$

在点 a_j 处连续.

定理 2.4.16

设 X 和 $Y_i(1 \le i \le m)$ 均为拓扑空间, 乘积空间 $Y := \prod_{i=1}^m Y_i$ 上装备乘积拓扑, 则映射 $f = (f_1, \cdots, f_m) : X \to Y$ 在点 $a \in X$ 处连续, 当且仅当每个映射 $f_i : X \to Y_i(1 \le i \le m)$ 在点 $a \in X$ 处连续.

特别注意下面关于连续函数的扩张定理, 它在定义 Brouwer 拓扑度, 建立 Hairy 球定理和 Borsuk-Ulam 定理等方面应用广泛.

定理 2.4.17 (Tietze-Urysohn 扩张定理)

设 X 为正规的拓扑空间, F 为 X 的闭子集, $f : F \to \mathbb{R}$ 为连续函数, 则存在连续函数 $\tilde{f} : X \to \mathbb{R}$, 使得对所有的 $x \in F$,

$$\tilde{f}(x) = f(x).$$

现在介绍利用一个集合到拓扑空间上给定的映射, 构造该集合上一类特殊拓扑的基本方法.

定理 2.4.18 (最弱拓扑的存在性)

设给定集合 X 和从 X 到拓扑空间 Y_i 上映射 φ_i 的族 $\{\varphi_i\}_{i\in I}$, 则存在 X 上具有下述两个性质的拓扑:

1. 所有的映射 $\varphi_i : X \to Y_i, i \in I$ 在该拓扑下均连续.
2. X 上关于该拓扑的开子集, 关于 X 上使所有映射 $\varphi_i : X \to Y_i(i \in I)$ 均连续的拓扑也是开的.

根据上述两个性质, 在定理2.4.18中断言存在 (且显然唯一) 的拓扑称为 X 上使映射 $\varphi_i : X \to Y_i(i \in I)$ 均连续的最弱拓扑.

2.4.4.3 拓扑空间中的紧性

定义 2.4.37 ((具有 Heine-Borel-Lebesgue 性质的) 紧集)

设 (X, O) 为拓扑空间, K 为 X 的一个子集, 如果对 X 中任何一族开集 $\{O_i\}_{i\in I}$, 当 $K \subset \bigcup_{i\in I} O_i$ 时, 必存在 $\{O_i\}_{i\in I}$ 的有限子族 $\{O_j\}_{j\in J}$, 使得 $K \subset \bigcup_{j\in J} O_j$, 则称 K 为紧集.

注 上述定义即: 若 X 的子集 K 的任何开覆盖均有有限子覆盖, 则 K 为紧集.

命题 2.4.2

拓扑空间 (X, O) 的子集 K 为紧集当且仅当装备诱导拓扑的拓扑空间 (K, O_K) 是紧的.

下列定理给出了有关紧性的基本性质.

定理 2.4.19

拓扑空间 X 是紧的充要条件是对 X 中闭子集 F_i 组成的任何闭集族 $\{F_i\}_{i\in I}$, 如果对 $\{F_i\}_{i\in I}$ 的任何有限子族 $\{F_j\}_{j\in J}$ 均有 $\bigcap_{j\in J} F_j \ne \emptyset$, 那么必有 $\bigcap_{i\in I} F_i \ne \emptyset$.

定理 2.4.20

1. Hausdorff 空间的任何紧子集均为闭集.
2. 紧拓扑空间的闭子集为紧集.

定理 2.4.21 (连续映射把紧集映到紧集)

设 X, Y 为两个拓扑空间, $f: X \to Y$ 为连续映射, 则 X 中任何紧集 K 的像 $f(K)$ 是 Y 中的紧集.

定理 2.4.22 (同胚保持紧性)

紧拓扑空间 X 到拓扑空间 Y 上的连续双射是 X 到 Y 上的同胚, 因而 Y 是紧的.

定理 2.4.23 (乘积空间的紧性)

设 $X_j (1 \le j \le n)$ 均为紧拓扑空间, 则装备了乘积拓扑的乘积空间 $\prod_{j=1}^{n} X_j$ 是紧空间.

注 定理2.4.23是 Tychonoff 定理的一个特殊情况, 该定理断言对任何一族紧的拓扑空间 $\{X_i\}_{i\in I}$, 装备乘积拓扑的乘积空间 $\prod_{i\in I} X_i$ 是紧的.

定义 2.4.38 (相对紧)

设 A 为拓扑空间 X 的子集, 如果闭包 $\overline{A}$ 是 X 的紧子集, 则称 A 为相对紧的.

2.4.4.4 拓扑空间中的连通和单连通性

定义 2.4.39 (连通)

拓扑空间 $(X, \mathcal{O})$ 称为连通的, 是指 X 中既是开集又是闭集的子集只有 X 和 $\emptyset$. X 的子集 A 称为连通的, 是指 A 关于 X 在 A 上的诱导拓扑, 形成的拓扑空间是连通的.

下列定理罗列了一些关于连通性的基本性质.

定理 2.4.24

设 A 为拓扑空间 X 的连通子集, 则 X 中满足 $A \subset B \subset \overline{A}$ 的子集 B 也是连通的, 特别地, $B = \overline{A}$ 是连通的.

定理 2.4.25

设 X 为连通的拓扑空间, Y 为拓扑空间, 则任何局部常值函数 $f: X \to Y$ 必为常值函数, 其中局部常值函数是指对每个点 $x \in X$, 存在邻域 V_x, 使得 $f|_{V_x}$ 为取常值的函数.

定理 2.4.26 (连续映射保连通性)

设 X, Y 为两个拓扑空间, $f: X \to Y$ 为连续映射, 则 X 中连通子集 A 的像 $f(A)$ 是 Y 中的连通子集.

定理 2.4.27 ($\mathbb{R}$ 上连通子集构造)

关于 $\mathbb{R}$ 上的通常拓扑, $\mathbb{R}$ 的子集为连通的当且仅当它是区间 (有界或无界).

定理 2.4.28 (Bolzano 中值定理)

设 X 为连通的拓扑空间, $f: X \to \mathbb{R}$ 为连续函数, $a, b \in X$ 且 $f(a) < f(b)$, 则对任意给定的 $y \in (f(a), f(b))$, 存在 $x \in X$ 使得 $f(x) = y$.

下述三个定理给出了关于连通性的常用充分条件.

定理 2.4.29

设 X 为拓扑空间, $\{A_i\}_{i\in I}$ 为 X 的一族连通子集. 若交集 $\bigcap_{i\in I} A_i$ 非空, 则其并集 $\bigcup \Im_{i\in I} A_i$ 是连通的.

定理 2.4.30

设 $X_j (1 \le j \le n)$ 均为连通的拓扑空间, 则其装备乘积拓扑的乘积空间 $X = \prod_{j=1}^{n} X_j$ 是连通的.

定义 2.4.40 (连通分支)

设 X 为拓扑空间, 关系"存在 X 中包含 x 与 y 的连通子集"是 X 中的一个等价关系, 模这个等价关系的等价类为 X 的子集, 称为 X 的连通分支. 给定 $x \in X$, X 中包含 x 的连通分支称为 x 的连通分支.

定理 2.4.31

设 X 为拓扑空间, $x \in X$, 则 x 的连通分支即 X 包含 x 的所有连通子集的并.

定义 2.4.41 (道路, 弧连通)

设 x 和 y 是拓扑空间 X 中两个点. 所谓连接 x 到 y 的道路是指连续映射 $\gamma : [0,1] \to X$, 满足 $\gamma(0) = x, \gamma(1) = y$. 如果对拓扑空间 X 中任何两个不同的点, 均存在连接 x 到 y 的道路, 那么称 X 为弧连通的.

定理 2.4.32

1. 弧连通的拓扑空间是连通的.
2. 赋范向量空间中的任何连通开集均是弧连通的.

定义 2.4.42 (道路同胚)

设 x 和 y 是拓扑空间 X 中的两个点, 连接 x 到 y 的两条道路 $\gamma_0 : [0,1] \to X$ 和 $\gamma_1 : [0,1] \to X$ 称为同胚的, 是指存在连续映射 $H : [0,1] \times [0,1] \to X$, 满足 $H(\cdot, 0) = \gamma_0, H(\cdot, 1) = \gamma_1, H(0, \cdot) = x, H(1, \cdot) = y$, 称 H 为连接 γ_0 到 γ_1 的一个同胚.

定义 2.4.43 (单连通)

拓扑空间 X 称为单连通的, 是指它是弧连通的, 且对任何两条道路 $\gamma_0 : [0,1] \to X$ 和 $\gamma_1 : [0,1] \to X$, 当 $\gamma_0(0) = \gamma_1(0), \gamma_0(1) = \gamma_1(1)$ 时, 它们必是同胚的.

2.4.5 补充: Tychonoff 定理及其证明

该部分选自 **[SL]**, 旨在完善前面介绍的乘积拓扑内容. 在有限个拓扑空间的乘积空间中, 讨论紧性一般是比较自然的. 比方说, 设 E, F 是赋范线性空间, S, T 分别是它们的紧子集. 设 $\{z_n\}$ 是 $S \times T$ 中的序列, 记 $z_n = (x_n, y_n), x_n \in E, y_n \in F$. 显见可以在 S 中找到收敛到 a 的子列 $\{x_{n_i}\}$, 进而在 T 中找到收敛到 b 的子列 $\{y_{n_{i_k}}\}$, 则 $\{z_{n_{i_k}}\}$ 就是欲求的收敛子列.

上述证明的思想在于先对坐标作投影, 再在坐标收敛的基础上得到乘积空间中的收敛. 但这个思路在无穷乘积中失效了, 因为可能没法穷尽坐标. 不过, 这其中蕴含的基本思想还是适用于一般情况. 上述过程在无穷乘积中的阻碍在于多重坐标空间中点的表示可能不唯一, 进而必须找到一个给出在所有坐标空间中同时表示点的系统方法. 下面的证明属于 Bourbaki 学派.

定理 2.4.33 (Tychonoff)

设 $\{X_\alpha\}_{\alpha \in A}$ 是紧空间族, 则其乘积空间

$$X = \prod_{\alpha \in A} X_\alpha$$

是紧的.

证明

设 $\mathscr{F}=\{F_i\}_{i\in I}$ 是一组闭子集的乘积, 其中任意有限个集合的交均非空 (称该性质为有限交性). 可以说明包含上述 $\mathscr{F}$ 且满足有限交性的 X 的子集族关于 $\subset$ 构成序关系, 且进一步可以证明该序集是归纳序集[18], 故存在具有有限交性的最大集族 $\mathscr{F}^*$. 记

$$\pi_\alpha : X \to X_\alpha$$

是对第 α 个向量的投影. 对每个 α, 闭集族 $\{\overline{\pi_\alpha(F)}\}_{F\in\mathscr{F}^*}$ 有有限交性, 进而存在 x_α 在全体 $\overline{\pi_\alpha(F)}$ 中. 记 $x=\{x_\alpha\}$, 则只需说明 x 在全体 $F\in\mathscr{F}^*$ 中.

观察到既然 $\mathscr{F}^*$ 是最大集族, $\mathscr{F}^*$ 中有限个集合的交依旧落在 $\mathscr{F}^*$ 中. 设 $U\subset X$ 是包含 x 的开集, 形如

$$U=U_{\alpha_1}\times\cdots\times U_{\alpha_n}\times\prod_{\alpha\neq\alpha_1}X_\alpha$$

其中每个 U_{α_i} 在 X_{α_i} 中是开集. 进而对任意的 i, $x_{\alpha_i}\in U_{\alpha_i}$, 因而对全体 $F\in\mathscr{F}^*$, U_{α_i} 均包含 $\pi_{\alpha_i}(F)$ 中的一个点. 故

$$\pi_{\alpha_i}^{-1}(U_{\alpha_i})=U_{\alpha_i}\times\prod_{\alpha\neq\alpha_i}X_\alpha$$

对每个 $F\in\mathscr{F}^*$ 都包含了其中一个点. 因为 $\mathscr{F}^*$ 同时具有最大性和有限交性, 知 $\pi_{\alpha_i}^{-1}(U_{\alpha_i})\in\mathscr{F}^*$, 进而这些集合的有限交 $\bigcap_{i=1}^n\pi_{\alpha_i}^{-1}(U_{\alpha_i})$ 也在 $\mathscr{F}^*$ 中. 但这个有限交正是 U, 故 U 与每个 $F\in\mathscr{F}^*$ 均有交, 因而也与每个 $F\in\mathscr{F}$ 有交. 故 x 在每个 $F\in\mathscr{F}$ 的闭包中, 进而根据它们的闭性知 $x\in F, \forall F\in\mathscr{F}$. □

推论 2.4.10

$\mathbb{R}^n$ 的子集是紧集当且仅当其为有界闭集.

推论 2.4.11

$\mathbb{R}^n$ 上的全体函数均等价.

推论 2.4.12 (Riesz)

赋范线性空间局部紧[a]当且仅当其有限维.

[a]每个点具有紧邻域.

2.5 凸集与不动点

2.5.1 知识梳理

2.5.1.1 定义与基本性质

一般线性空间的凸集可以借由平面凸集来理解: 若 E 是一个平面凸集, 则对于 E 中任意两点 x,y, 联结这两点的线段也在 E 内, 即

$$\lambda x+(1-\lambda)y\in E\ (\forall x,y\in E,\forall 0\le\lambda\le 1)$$

凸性不要求空间具有拓扑结构, 因而这个概念可以推广到一般线性空间.

[18]每个非空全序子集均有上界, S 全序意味着在序关系 $\le$ 下, S 中的元素 x,y 要么满足 $x\le y$, 要么满足 $y\le x$.

定义 2.5.1 (凸集)

设 $\mathscr{X}$ 是线性空间, $E \subset \mathscr{X}$, 称 E 为一个凸集, 如果

$$\lambda x + (1-\lambda)y \in E, \ \forall x, y \in E, \forall 0 \le \lambda \le 1$$

补充命题 2.5.1 (⋆)

上述定义中的凸集条件

$$\lambda x + (1-\lambda)y \in E, \ \forall x, y \in E, \forall \lambda \in [0,1] \tag{2.26}$$

不能减弱为

$$\lambda_1 x + (1-\lambda_1)y \in E, \ \forall x, y \in E, \forall \lambda_1 \in [0,1] \cap \mathbb{Q} \tag{2.27}$$

但可以减弱为

$$\lambda_2 x + (1-\lambda_2)y \in E, \ \forall x, y \in E, \forall \lambda_2 \in (0,1) \backslash \mathbb{Q} \tag{2.28}$$

证明

显见(2.26)⇒(2.27), (2.26)⇒(2.28).

现在说明(2.27)⇏(2.26). 取 $\mathscr{X} = \mathbb{R}, E = \mathbb{Q}$, 显见 E 满足(2.27), 但若取 $x = 1, y = 0 \in E, \lambda = \frac{\sqrt{2}}{2} \in [0,1]$, 则 $\lambda x + (1-\lambda)y = \frac{\sqrt{2}}{2} \notin E$, 因而 E 不满足(2.26).

再说明(2.28)⇒(2.26). 当 $\lambda \in (0,1)$ 是无理数时, 由(2.28)成立知(2.26)成立. 而当 λ 是有理数时, 不妨设 $\lambda \in (0,1)$(亦即设 $\lambda \neq 0, 1$), 注意到任意有理数均可写成两个无理数之和, 这是因为

$$\forall \lambda \in (0,1) \exists n \in \mathbb{N}(\lambda + \frac{\sqrt{2}}{n} \in (0,1) \backslash \mathbb{Q} \wedge \lambda - \frac{1}{(1-\frac{\sqrt{2}}{2})n} \in (0,1) \backslash \mathbb{Q})$$

由(2.28), 任取 $x, y \in E$ 均有

$$m_1 := (\lambda + \frac{\sqrt{2}}{n})x + (1 - \lambda - \frac{\sqrt{2}}{n})y \in E$$
$$m_2 := (\lambda - \frac{1}{(1-\frac{\sqrt{2}}{2})n})x + (1 - \lambda + \frac{1}{(1-\frac{\sqrt{2}}{2})n})y \in E$$

再次应用(2.28)得

$$\lambda x + (1 - \lambda y) = \frac{\sqrt{2}}{2} m_1 + (1 - \frac{\sqrt{2}}{2}) m_2 \in E$$

从而(2.26)成立. □

注 事实上, (2.26)也不能减弱为

$$\exists \lambda_0 \in (0,1) \backslash \mathbb{Q} \forall x, y \in E(\lambda_0 x + (1-\lambda_0)y \in E) \tag{2.29}$$

显见(2.26)⇒(2.29), 欲证明(2.29)⇏(2.26), 取

$$E := \{a_0 + a_1\lambda_0 + \cdots + a_k\lambda_0^k : k \in \mathbb{Z}^+, a_0, \cdots, a_k \in \mathbb{Q}\}$$

则 E 可数, 且 $\forall x, y \in E(\lambda_0 x + (1-\lambda_0)y \in E)$. 又因为 (0,1) 不可数, 故存在 $z \in (0,1) \backslash E$, 使得当 $x = 1 \in E, y = 0 \in E, \lambda = z$ 时 $\lambda x + (1-\lambda)y = z \notin E$, 从而 E 并非凸集.

根据定义可得下述命题.

命题 2.5.1

若 $\{E_\lambda : \lambda \in \Lambda\}$ 是线性空间 $\mathscr{X}$ 中的一族凸集, 则 $\bigcap_{\lambda \in \Lambda} E_\lambda$ 也是凸集.

定义 2.5.2 (凸包, 凸组合)

设 $\mathscr{X}$ 是线性空间, $A \subset \mathscr{X}$. 若 $\{E_\lambda : \lambda \in \Lambda\}$ 为 $\mathscr{X}$ 中包含 A 的一切凸集, 则称 $\bigcap_{\lambda\in\Lambda} E_\lambda$ 为 A 的凸包, 记作 $\mathrm{co}(A)$. 又对任意的 $n \in \mathbb{N}, x_1, x_2, \cdots, x_n \in A$, 称 $\sum_{i=1}^n \lambda_i x_i$ 为 $x_1, x_2, \cdots, x_n$ 的凸组合, 如果其系数满足 $\lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1$.

♣

命题 2.5.2

设 $\mathscr{X}$ 是线性空间, $A \subset \mathscr{X}$, 那么 A 的凸包是 A 中元素任意凸组合的全体, 即

$$\mathrm{co}(A) = \{\sum_{i=1}^n \lambda_i x_i : \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, x_i \in A, i = 1, 2, \cdots, n, \forall n \in \mathbb{N}\} \tag{2.30}$$

♠

证明

记(2.30)式右端为 S, 则显见 $A \subset S$, 下说明 S 是凸集. 这是因为任取 $x, y \in S$, 根据 S 的定义记

$$x = \sum_{i=1}^n \lambda_i x_i, y = \sum_{j=1}^m \eta_j y_j,\ x_i, y_j \in A, i = 1, \cdots, n, j = 1, \cdots, m, \sum_{i=1}^n \lambda_i = \sum_{j=1}^m \eta_j = 1$$

任取 $\lambda \in [0, 1]$ 有

$$\lambda x + (1-\lambda) y = \lambda \sum_{i=1}^n \lambda_i x_i + (1-\lambda) \sum_{j=1}^m \eta_j y_j$$

知

$$\lambda \sum_{i=1}^n \lambda_i + (1-\lambda) \sum_{j=1}^m \eta_j = \lambda + 1 - \lambda = 1$$

这说明 $\lambda x + (1-\lambda) y \in S$, 因而 S 是凸集, 进而 $\mathrm{co}(A) \subset S$.

反之, 设 F 是包含 A 的任一凸集, 则出于 $A \subset F$ 知 $x_i \in F (i = 1, 2, \cdots, n)$, 进而根据凸集定义 $\sum_{i=1}^n \lambda_i x_i \in F$, 得到 $S \subset F$, 故令 F 跑遍全体包含 A 的凸集, 根据凸包作为全体凸集交的定义即得 $S \subset \mathrm{co}(A)$. □

定义 2.5.3 (Minkowski 泛函)

设 $\mathscr{X}$ 是线性空间, C 是 $\mathscr{X}$ 上含有 θ 的凸子集, 在 $\mathscr{X}$ 上规定一个取值于 $[0, \infty]$ 的函数

$$P(x) = \inf\{\lambda > 0 : \frac{x}{\lambda} \in C\},\ \forall x \in \mathscr{X}$$

与 C 对应, 称函数 P 为 C 的 Minkowski 泛函.

♣

注 从几何上看, Minkowski 泛函表示了 $\mathscr{X}$ 中的元素最少需要缩放多少才能进入 C. 特别注意根据 C 的凸性与 $\theta \in C$, 只要确定了 $P(x)$, 则 $\lambda \in (P(x), \infty]$ 均可使 $\frac{x}{\lambda} \in C$.

命题 2.5.3

设 $\mathscr{X}$ 是线性空间, C 是 $\mathscr{X}$ 上含有 θ 的凸子集, 若 P 为 C 的 Minkowski 泛函, 则 P 具有下列性质:

1. $P(x) \in [0, \infty], P(\theta) = 0$;
2. $P(\lambda x) = \lambda P(x), \forall x \in \mathscr{X}, \forall \lambda > 0$(正齐次性);
3. $P(x + y) \leq P(x) + P(y), \forall x, y \in \mathscr{X}$(次可加性).

♠

证明

屈于下确界定义只需验证 3. 不妨设 $P(x), P(y)$ 有穷, 对任意的 $\varepsilon > 0$, 取 $\lambda_1 = P(x) + \frac{\varepsilon}{2}, \lambda_2 = P(y) + \frac{\varepsilon}{2}$, 则

$$\frac{x}{\lambda_1} \in C, \frac{y}{\lambda_2} \in C$$

因为 C 是凸的, 所以

$$\frac{x+y}{\lambda_1+\lambda_2} = \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{x}{\lambda_1} + \frac{\lambda_2}{\lambda_1+\lambda_2}\frac{y}{\lambda_2} \in C$$

这说明

$$P(x+y) \leq \lambda_1 + \lambda_2 = P(x) + P(y) + \varepsilon$$

根据 ε 的任意性得到 3.

为了让 $P(x)$ 刨除成为 ∞ 的可能, 且让正齐次性转变为齐次性, 引入下述概念.

定义 2.5.4 (吸收, 对称)

线性空间中, 含有 θ 的凸集 C 称作是吸收的, 如果 $\forall x \in \mathscr{X}, \exists \lambda > 0$, 使得 $\frac{x}{\lambda} \in C$; 称 C 是对称的, 如果 $x \in C \Rightarrow -x \in C$. ♣

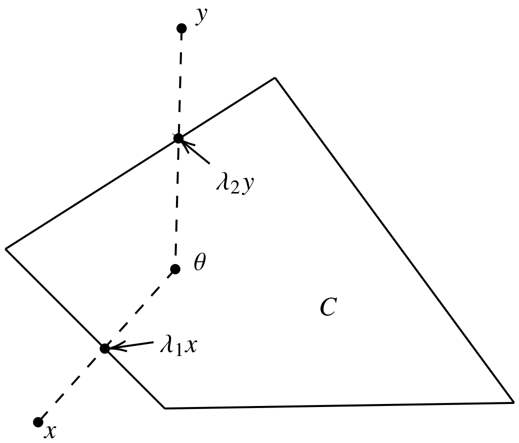


图 2.6: 吸收凸集示意: $\mathscr{X}$ 中的任意点总能在缩放某一比例后进入 C.

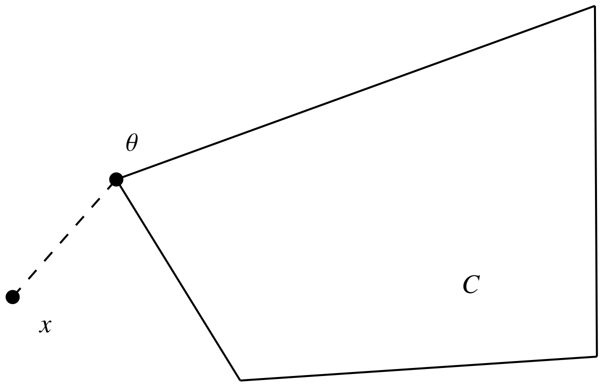


图 2.7: 不吸收凸集示意: 存在 $\mathscr{X}$ 中的点, 无论怎样缩放, 都不会进入 C.

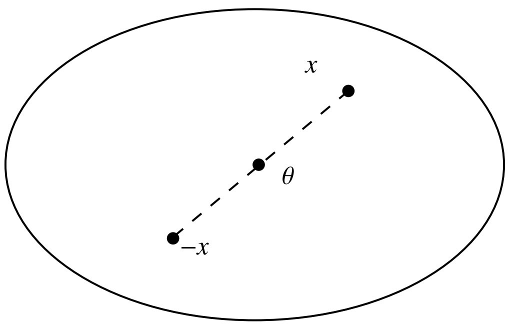


图 2.8: 对称凸集示意: 只要 $x \in C$, 则 $-x \in C$.

注意如果在吸收集的定义中将凸集的要求取出, 则就算 C 是吸收的, θ 也不一定是 C 的内点. 参见下例:

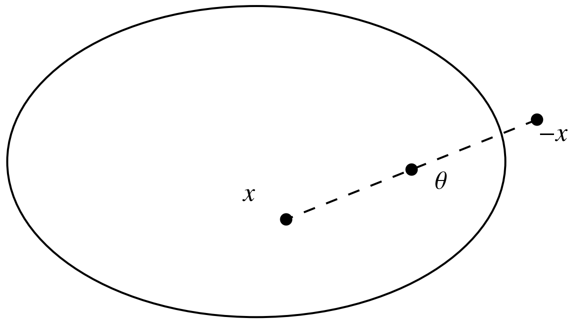


图 2.9: 不对称凸集示意: 存在 $x \in C$, 但 $-x \notin C$.

例 2.42★ 设 C 为 $\mathbb{R}^2$ 上的点集 $B((0,0),1)\backslash\{(x,x^2) : x \in \mathbb{R}, x \neq 0\}$, 则 C 不是凸集, 但确为吸收集, 且 $(0,0)$ 并非 C 的内点.

如果在有限维空间中设 C 是凸集, $\theta \in C$, 且 $\forall x \in \mathscr{X} \exists \lambda > 0 (\frac{x}{\lambda} \in C)$, 问 θ 是否一定是 C 的内点, 答案是肯定的:

补充命题 2.5.2 (★)

设 $\mathscr{X}$ 为有限维 B^* 空间, 若 $C \subset \mathscr{X}$ 为含有 θ 的吸收凸集, 则 θ 必为 C 的内点.

证明

设 $\{e_i\}_{i=1}^n$ 为 $\mathscr{X}$ 的基, 对任意 $i \in \{1, \cdots, n\}$, 由 C 的吸收性知存在 $\lambda_i > 0, \mu_i > 0$ 使得 $\lambda_i e_i \in C, \mu_i(-e_i) \in C$. 令 $\lambda_0 := \min_{1 \le i \le n}\{\lambda_i, \mu_i\}$, 则 $\lambda_0 > 0$. 依此与 $\theta \in C$ 和 C 为凸集可知: 对任意 $i \in \{1, \cdots, n\}$ 与任意 $0 \le \lambda \le \lambda_0$, 均有

$$\begin{cases} \lambda e_i = \frac{\lambda}{\lambda_i}(\lambda_i e_i) + (1 - \frac{\lambda}{\lambda_i})\theta \in C \\ -\lambda e_i = \frac{\lambda}{\mu_i}(-\mu_i e_i) + (1 - \frac{\lambda}{\mu_i})\theta \in C \end{cases} \tag{2.31}$$

取

$$A := \{x = \sum_{i=1}^n x_i e_i : \sum_{i=1}^n |x_i| \le \lambda_0\}$$

下面证明 $A \subset C$. 这是因为对任意 $x \in A$, 有:

$$x = \sum_{i=1}^n [\frac{|x_i|}{\sum_{j=1}^n |x_j|}(\sum_{j=1}^n |x_j|) \operatorname{sgn}(x_i) e_i]$$

有 $\sum_{j=1}^n |x_j| \le \lambda_0$ 与(2.31)式知 $(\sum_{j=1}^n |x_j|) \operatorname{sgn}(x_i) e_i \in C (\forall i \in \{1, \cdots, n\})$. 据此与 C 的凸性知 $x \in \operatorname{co}(\{(\sum_{j=1}^n |x_j|) \operatorname{sgn}(x_i) e_i\}_{i=1}^n) \subset C$, 从而 $A \subset C$. 因为 $\mathscr{X}$ 本身是有限维, 故根据有穷维空间的范数等价定理2.4.1知存在 $C_1 > 0$ 使得:

$$\forall x := \sum_{i=1}^n x_i e_i (|x| := (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} \le \frac{1}{C_1}\|x\|_{\mathscr{X}}) \tag{2.32}$$

现证明 $B(\theta, \frac{C_1 \lambda_0}{\sqrt{n}}) \subset C$. 任取 $x := \sum_{i=1}^n x_i e_i \in B(\theta, \frac{C_1 \lambda_0}{\sqrt{n}})$, 由(2.32)式与 Hölder 不等式知

$$\sum_{i=1}^n |x_i| \le \sqrt{n}|x| \le \frac{\sqrt{n}}{C_1}\|x\|_{\mathscr{X}} < \lambda_0$$

从而 $x \in A \subset C$. 这说明 $B(\theta, \frac{C_1 \lambda}{\sqrt{n}}) \subset C$, 因而 θ 是 C 的内点, 命题得证. □

注 在无穷维空间中, 上述结论不一定成立, 反例见 **[WL]**.

命题 2.5.4

要使得 C 是吸收凸集, 必须且仅须其 Minkowski 泛函 $P(x)$ 是实值函数[a]; 要使得 C 是对称凸集, 必须 $P(x)$ 是实齐次的, 即

$$P(\alpha x) = |\alpha| P(x),\ \forall \alpha \in \mathbb{R}$$

[a]这里说的实值函数是相对于广义实数系说的, 也就是不存在 $x \in \mathscr{X}$ 使得 $P(x) = \infty$.

证明 ($\star$)

当 $P(x)$ 是实值函数, 知 $\forall x \in \mathscr{X}(P(x) \in [0, +\infty))$, 进而任取 $\lambda > P(x)$, 根据 Minkowski 泛函的定义知 $\frac{x}{\lambda} \in C$, 因而 C 是吸收的.

当 C 是吸收凸集, 根据吸收的定义知

$$\forall x \in \mathscr{X} \exists \lambda_x > 0 (\frac{x}{\lambda_x} \in C)$$

则由 Minkowski 泛函的定义知:

$$P(x) := \inf\{\lambda > 0 : \frac{x}{\lambda} \in C\}$$

可得 $P(x) \le \lambda_x < \infty$, 亦即 $P(x)$ 是实值函数.

当 C 是对称凸集, 任取 $\mathbb{R}^1 \ni \alpha \ne 0$ 知

$$\begin{aligned} P(\alpha x) &= \inf\{\lambda > 0 : \frac{\alpha x}{\lambda} \in C\} = |\alpha| \inf\{\frac{\lambda}{|\alpha|} > 0 : \operatorname{sgn}(\alpha)\frac{x}{\frac{\lambda}{|\alpha|}} \in C\} \\ &\overset{\text{(i)}}{=} |\alpha| \inf\{\frac{\lambda}{|\alpha|} > 0 : \frac{x}{\frac{\lambda}{|\alpha|}} \in C\} = |\alpha| P(x) \end{aligned}$$

其中 (i) 是因为由 C 的对称性知 $\operatorname{sgn}(\alpha)\frac{x}{\frac{\lambda}{|\alpha|}} \in C \Leftrightarrow \frac{x}{\frac{\lambda}{|\alpha|}} \in C$. 又当 $\alpha = 0$ 时, 根据 C 的对称凸性知 $\theta \in C$, 因而 $P(\theta) = 0$, 欲证式依旧成立. 综上即知 $P(x)$ 是实齐次的, 命题即证. □

定义 2.5.5 (均衡)

复线性空间 $\mathscr{X}$ 的一个子集 C 称作是均衡的, 如果

$$x \in C \Rightarrow \alpha x \in C, \forall \alpha \in \mathbb{C}, |\alpha| = 1.$$

在给定了半范数的定义[19]后, 有下述命题成立.

命题 2.5.5

复线性空间 $\mathscr{X}$ 上的任一个均衡吸收凸集 C, 决定了该空间上的一个半范数.

对于赋范线性空间 $\mathscr{X}$, 有下述更强的结果.

命题 2.5.6

设 $\mathscr{X}$ 是一个 B^* 空间, C 是一个含有 θ 点的闭凸集. 如果 $P(x)$ 是 C 的 Minkowski 泛函, 那么 $P(x)$ 下半连续, 且有

$$C = \{x \in \mathscr{X} : P(x) \le 1\}. \tag{2.33}$$

此外, 如果 C 有界, 那么 $P(x)$ 满足

$$P(x) = 0 \Leftrightarrow x = \theta.$$

又如果 C 以 θ 为一个内点, 那么 C 是吸收的, 并且 $P(x)$ 一致连续.

证明

[19]满足次可加性, 正定, 齐次的函数.

首先证明

$$C=\{x\in\mathscr{X}:P(x)\le 1\}$$

为此考虑证明

$$\alpha C=\{x\in\mathscr{X}:P(x)\le\alpha\}\ (\forall\alpha>0)$$

这是因为对任意的 $\alpha>0$, 若 $x\in\alpha C$, 也即 $\frac{x}{\alpha}\in C$, 根据 $P(x)$ 的定义知 $P(x)\le\alpha$. 反之, 当 $P(x)\le\alpha$, 因为 $P(x)$ 本质上是下确界, 有

$$\frac{x}{\alpha+\frac{1}{n}}\in C,\frac{x}{\alpha+\frac{1}{n}}\to\frac{x}{\alpha}\ (n\to\infty)$$

又因为 C 是闭的, 故 $\frac{x}{\alpha}\in C\Rightarrow x\in\alpha C$. 故

$$\alpha C=\{x\in\mathscr{X}:P(x)\le\alpha\}\ (\forall\alpha>0)$$

令 $\alpha=1$ 即得

$$C=\{x\in\mathscr{X}:P(x)\le 1\}$$

又因为对任意的 $\alpha>0$, αC 都是闭集, 故 P 是下半连续的.

因为 $\theta\in C$, 当然有 $P(\theta)=0$. 当 C 有界, 根据定义知

$$\exists r>0(C\subset B(\theta,r))$$

进而

$$\forall x\in\mathscr{X}\backslash\{\theta\}(\frac{rx}{\|x\|}\notin C)\Rightarrow P(x)\ge\frac{\|x\|}{r}$$

这说明 $P(x)=0\Rightarrow x=\theta$.

当 C 以 θ 为内点, 根据定义知

$$\exists r>0(B(\theta,r)\subset C)$$

此时

$$\forall x\in\mathscr{X}\backslash\{\theta\}(\frac{rx}{2\|x\|}\in C)$$

这说明 C 是吸收的, 且有

$$\forall x\in\mathscr{X}(P(x)\le\frac{2\|x\|}{r})$$

于是

$$|P(x)-P(y)|\le\max(P(x-y),P(y-x))\le\frac{2}{r}\|x-y\|,\ \forall x,y\in\mathscr{X}$$

因而 $P(x)$ 一致连续. □

推论 2.5.1

若 C 是 $\mathbb{R}^n$ 中的一个紧凸子集, 则必存在正整数 $m\le n$, 使得 C 同胚于 $\mathbb{R}^m$ 中的单位球. ♡

证明

用 E 表示包含 C 的最小闭线性子流形, 设其维数为 $m(\le n)$, 于是根据线性流形的定义, 在 C 上必有 $m+1$ 个向量 $e_1,\cdots,e_m,e_{m+1}$, 使得 $e_i-e_{m+1}(i=1,2,\cdots,m)$ 是线性无关的. 现在找一个向量 e_0 在 C 的内部用以确定原点, 令

$$e_0=\frac{1}{m+1}\sum_{i=1}^{m+1}e_i$$

根据 C 的凸性知 $e_0 \in C \subset E$, 这说明 $E - e_0$ 是一个 m 维线性子空间, 因而对任意的 $y \in E$, 存在唯一的表示[20]:

$$y = \sum_{i=1}^{m} \mu_i (e_i - e_0) + e_0 \tag{2.34}$$

在 $E - e_0$ 上可引进等价范数

$$\|z\| = (\sum_{i=1}^{m} |\mu_i|^2)^{\frac{1}{2}} \; (z = y - e_0, y \in E). \tag{2.35}$$

下面证明当(2.35)式中的 $\|z\|$ 足够小时, 对应由(2.34)式给出的 y 是在 C 中的. 这是因为

$$y = \sum_{i=1}^{m} \mu_i e_i + (1 - \sum_{j=1}^{m} \mu_j) e_0 = \sum_{i=1}^{m} (\mu_i + \frac{1}{m+1}(1 - \sum_{j=1}^{m} \mu_j)) e_i + \frac{1}{m+1}(1 - \sum_{j=1}^{m} \mu_j) e_{m+1} \tag{2.36}$$

当 $|\mu_j| (j = 1, \cdots, m)$ 足够小时, (2.36)式右端各项系数都是正的, 且各项系数总和满足

$$\sum_{i=1}^{m} (\mu_i + \frac{1}{m+1}(1 - \sum_{j=1}^{m} \mu_j)) + \frac{1}{m+1}(1 - \sum_{j=1}^{m} \mu_j) = 1$$

故 $y \in \mathrm{co}\{e_1, \cdots, e_m, e_{m+1}\} \subset C$, 这件事说明了在 $E - e_0$ 中, θ 是 $C - e_0$ 的内点.

现在在 $E - e_0$ 上, C 是一个以 θ 为内点的有界闭凸集, 根据命题2.5.6, $C - e_0$ 的 Minkowski 泛函 $P(z)$ 是 $E - e_0$ 上的一个一致连续, 正齐次, 次可加的泛函, 且 $P(z) = 0 \Leftrightarrow z = \theta$. 根据定理2.4.2知, 存在常数 $C_1, C_2 > 0$, 使得

$$C_1 \|z\| \le P(z) \le C_2 \|z\| \; (\forall z \in E - e_0)$$

设 $B^m(\theta, 1)$ 是 $E - e_0$ 中的单位球, 若令

$$\varphi(z) = \begin{cases} e_0 + \frac{\|z\| z}{P(z)}, & z \neq \theta, \\ e_0, & z = \theta \end{cases}$$

则 $\varphi : B^m(\theta, 1) \to C$ 是一个在上同胚. 事实上, φ 的连续性显见, 现只需说明其既单又满. 对于单射, 设 $\varphi(x) = \varphi(y)$, 得到 $\frac{\|x\| x}{P(x)} = \frac{\|y\| y}{P(y)}$. 因为 $\frac{\|x\|}{P(x)}, \frac{\|y\|}{P(y)}$ 均为数量, 故 $\frac{\|x\|}{P(x)} = \frac{\|y\|}{P(y)}$, 因而 $x = y$. 对于满射, 任取 $y \in C$, 对应取 $z = \frac{P(y-e_0)}{\|y-e_0\|}(y - e_0) \in B^m(\theta, 1)$, 知

$$\varphi(z) = e_0 + P(y - e_0) \| \frac{y - e_0}{\|y - e_0\|} \| \cdot \frac{(y - e_0) P(y - e_0)}{\|y - e - 0\|} \cdot \frac{\|y - e - 0\|}{P(y - e_0) P(y - e_0)} = y$$

因而 φ 是满射. □

2.5.1.2 Brouwer 与 Schauder 不动点定理

定理 2.5.1 (Brouwer)

设 B 是 $\mathbb{R}^n$ 中的闭单位球, 又设 $T : B \to B$ 是一个连续映射, 那么 T 必有一个不动点 $x \in B$.

结合推论(2.5.1)与 Brouwer 不动点定理可得

推论 2.5.2

设 C 是 $\mathbb{R}^n$ 中的一个紧凸子集, $T : C \to C$ 是连续的, 则 T 必有一个在 C 上的不动点.

证明

由于 C 与 $\mathbb{R}^m (m \le n)$ 中的一个单位球同胚, 记该同胚为 $\varphi : B^m(\theta, 1) \to C$. 考察映射

$$T_\varphi = \varphi^{-1} \circ T \circ \varphi$$

显见 $T_\varphi : B^m(\theta, 1) \to B^m(\theta, 1)$ 是连续的, 对其应用 Brouwer 不动点定理2.5.1知, 存在 $x \in B^m(\theta, 1)$ 使得 $T_\varphi x = x$, 这说明 $y = \varphi x \in C$ 是 T 的不动点. □

[20]这里相当于在 $E - e_0$ 中取 y 的坐标表示, 再在两边加上 e_0 将其拉回到 E.

上面讨论的是有穷维空间的不动点定理, 下面在无穷维的情况讨论这件事.

定理 2.5.2 (Schauder)

设 C 是 B^* 空间 $\mathscr{X}$ 中的一个闭凸子集, $T:C\to C$ 连续且 $T(C)$ 列紧, 则 T 在 C 上必有一个不动点.

证明

根据 Hausdorff 定理2.3.1, 因为 T(C) 是列紧集, 故对任意的 $n\in\mathbb{N}$, 存在 $\frac{1}{n}$ 网 $N_n=\{y_1,\cdots,y_{r_n}\}$, 即

$$T(C)\subset\bigcup_{i=1}^{r_n}B(y_i,\frac{1}{n})\ (y_i\in T(C),i=1,\cdots,r_n)$$

记 $E_n=\operatorname{span}N_n$, 即 E_n 为由 N_n 张成的有穷维线性子空间, 作 $T(C)\to\operatorname{co}(N_n)$ 的映射 I_n:

$$I_n(y)=\sum_{i=1}^{r_n}y_i\lambda_i(y)\ (\forall y\in T(C)) \tag{2.37}$$

其中

$$\lambda_i(y)=\frac{m_i(y)}{\sum_{i=1}^{r_n}m_i(y)},\ m_i(y)=\begin{cases}1-n\|y-y_i\|, & y\in B(y_i,\frac{1}{n}),\\ 0, & y\notin B(y_i,\frac{1}{n})\end{cases}$$

下面说明 I_n 定义的合理性. 因为 $m_i(y)\geq 0$, 且对任意的 $y\in T(C)$, 都存在 $i_0(1\leq i_0\leq r_n)$ 使得

$$y\in B(y_{i_0},\frac{1}{n})\Rightarrow m_{i_0}(y)>0$$

故

$$\forall y\in T(C)(\sum_{i=1}^{r_n}m_i(y)>0)$$

因而 $\lambda_i(y)(1\leq i\leq r_n)$ 有定义且满足

$$\lambda_i(y)\geq 0(1\leq i\leq r_n),\ \sum_{i=1}^{r_n}\lambda_i(y)=1 \tag{2.38}$$

故 I_n 在 $T(C)$ 上有定义, 且(2.37)式与(2.38)式说明 $I_n(y)$ 是 N_n 中元素的凸组合, 故 $I_n(y)\in\operatorname{co}(N_n)$. 此外:

$$\begin{aligned}\|I_ny-y\|&=\|\sum_{i=1}^{r_n}y_i\lambda_i(y)-\sum_{i=1}^{r_n}y\lambda_i(y)\|=\|\sum_{i=1}^{r_n}(y_i-y)\lambda_i(y)\|\\&\leq\sum_{i=1}^{r_n}\|y_i-y\|\lambda_i(y)=\sum_{\substack{1\leq i\leq n\\ y\in B(y_i,\frac{1}{n})}}\|y_i-y\|\lambda_i(y)+\sum_{\substack{1\leq i\leq n\\ y\notin B(y_i,\frac{1}{n})}}\|y_i-y\|\lambda_i(y)\\&<\frac{1}{n}\sum_{\substack{1\leq i\leq n\\ y\in B(y_i,\frac{1}{n})}}\lambda_i(y)+0\cdot\sum_{\substack{1\leq i\leq n\\ y\notin B(y_i,\frac{1}{n})}}\|y_i-y\|\leq\frac{1}{n}\sum_{i=1}^{r_n}\lambda_i(y)+0=\frac{1}{n}\end{aligned} \tag{2.39}$$

注意 $T:C\to C,N_n\subset T(C)\subset C$, 而 C 是凸的, 这说明 $\operatorname{co}(N_n)\subset C$. 令 $T_n:=I_n\circ T$, 知 $T_n:\operatorname{co}(N_n)\to\operatorname{co}(N_n)$, 又因为 $\operatorname{co}(N_n)$ 是 E_n 中的一个有界闭凸子集, 根据推论(2.5.2)知存在 $x_n\in\operatorname{co}(N_n)\subset C$, 使得

$$T_nx_n=x_n \tag{2.40}$$

又因为 $T(C)\subset C$ 列紧, C 是闭集, 故存在子列 n_k 与 $x\in C$ 使得

$$Tx_{n_k}\to x,\ k\to\infty \tag{2.41}$$

结合(2.39)式与(2.40)式有:

$$\begin{aligned}\|x_n-x\|&=\|T_nx_n-x\|=\|I_nTx_n-Tx_n+Tx_n-x\|\\&\leq\|I_nTx_n-Tx_n\|+\|Tx_n-x\|<\frac{1}{n}+\|Tx_n-x\|\ (\forall n\in\mathbb{N})\end{aligned} \tag{2.42}$$

结合(2.41)式与(2.42)式可得 $x_{n_k}\to x(k\to\infty)$, 再根据 T 的连续性与(2.41)可得

$$Tx=x.$$

□

注 事实上上述证明中的 $m_i(y)$ 构造成 $\chi_{B(y_i,\frac{1}{n})}$ 也可以, 这里的 $\{m_i(y)\}$ 只需要满足两点: 其一是可以记录 y 是否在 $B(y_i,\frac{1}{n})$ 内, 其二是保证至少存在一个为正的 $m_{i_0}(y)$.

定义 2.5.6 (紧映射)

设 $\mathscr{X}$ 是 B^* 空间, E 是 $\mathscr{X}$ 的一个子集, 称映射 $T: E\to\mathscr{X}$ 是紧的, 如果它是连续的并且把 E 中的任意有界集映为 $\mathscr{X}$ 中的列紧集.

推论 2.5.3

设 C 为 B^* 空间 $\mathscr{X}$ 中的一个有界闭凸子集, $T: C\to C$ 是紧的, 则 T 在 C 上必有不动点.

2.5.1.3 应用

现在再来讨论常微分方程初值问题的存在性定理, 只假设函数

$$f(t,x):\mathbb{R}\times\mathbb{R}\to\mathbb{R}$$

在 $[-h,h]\times[\xi-b,\xi+b]$ 上二元连续 (进而有常数 $M>0$ 使得 $|f(t,x)|\le M$). 考察 $C[-h,h]$ 中的球 $\overline{B}(\xi,b)$ 上的映射:

$$(Tx)(t)=\xi+\int_0^t f(\tau,x(\tau))d\tau.$$

下面证明: 对足够小的 h, T 把 $\overline{B}(\xi,b)$ 映到 $\overline{B}(\xi,b)$, 并且 T 是紧的. 这是因为

$$\|Tx-\xi\|_{C[-h,h]}=\|\int_0^t f(\tau,x(\tau))d\tau\|=\max_{[-h,h]}(|\int_0^t f(\tau,x(\tau))d\tau|)\le M\cdot h,\ \forall x\in\overline{B}(\xi,b)$$

故当 $h\le\frac{b}{M}$ 时, T 把 $\overline{B}(\xi,b)$ 映到 $\overline{B}(\xi,b)$. 又因为

$$|(Tx)(t)-(Tx)(t')|=|\int_{t'}^t f(\tau,x(\tau))d\tau|\le M|t-t'|\ (\forall t,t'\in[-h,h])$$

$$|(Tx)(t)|\le|\xi|+Mh\ (\forall t\in[-h,h])$$

故 T 连续 (进一步是 Lipschitz 的), 可知 T 把 x 映成一致有界且等度连续的函数, 根据 Arzela-Ascoli 定理2.3.4知 $T(\overline{B}(\xi,b))$ 是列紧集, 故由 Schauder 定理2.5.2知 T 在 $\overline{B}(\xi,b)$ 上有不动点. 这便证明了下述定理.

定理 2.5.3 (Caratheodory)

假设函数 $f(t,x)$ 在 $[-h,h]\times[\xi-b,\xi+b]$ 上二元连续, $|f(t,x)|\le M$, 那么当 $h<\frac{b}{M}$ 时, 方程的初值问题

$$\begin{cases}\dot{x}(t)=f(t,x(t)),\\ x(0)=\xi\end{cases}$$

在 $[--h,h]$ 上存在解 $x(t)$.

2.5.2 习题

练习 2.39★ 设 $\mathscr{X}$ 是 B^* 空间, E 是以 θ 为内点的真凸子集, P 是由 E 产生的 Minkowski 泛函, 求证:

(1) $x\in E^\circ\Leftrightarrow P(x)<1$;

证明

当 $x\in E^\circ$, 根据定义知

$$\exists r>0(B(x,r)\subset E)$$

因而至少存在 $\delta>0$, 使得 $(1+\delta)x\in B(x,r)\subset E$, 这说明

$$P(x)\le\frac{1}{1+\delta}<1$$

当 $P(x) < 1$, 根据实数的稠密性知存在 $a > 0$ 使得 $P(x) + a < 1$. 现欲证明存在 $r > 0$, 使得对任意 $y \in B(x, r)$, 总有 $y \in E$. 注意若 $P(y) := \lambda_y < 1$, 则必存在 $\varepsilon > 0$ 使得 $\lambda_y + \varepsilon < 1$, 且 $\frac{y}{\lambda_y+\varepsilon} \in E$, 因而根据凸性知 $y \in E$, 故只需证明 $P(y) < 1$ 即可. 注意

$$P(y) = P(y - x + x) \leq P(y - x) + P(x)$$

$$P(x) = P(x - y + y) \leq P(x - y) + P(y)$$

得到

$$|P(y) - P(x)| \leq \max\{P(x - y), P(y - x)\}$$

因为 θ 是 E 的内点, 故存在 $r_\theta > 0$ 使得 $B(\theta, r_\theta) \subset E$. 任取 $x \in E \backslash \{\theta\}$, 知 $\frac{r_\theta}{2}\frac{x}{\|x\|} \in B(\theta, r)$, 因而根据 Minkowski 范数的定义有 $\frac{2\|x\|}{r_\theta} \geq P(x)$. 这便得到

$$\max\{P(x - y), P(y - x)\} \leq \frac{2\|x - y\|}{r_\theta}$$

现只需取 $r < \frac{r_\theta}{2}a$, 即得

$$\frac{2\|x - y\|}{r_\theta} \leq \frac{2r}{r_\theta} < a \Rightarrow |P(y) - P(x)| < a \Rightarrow P(y) < P(x) + a < 1$$

欲证成立, 命题得证.

(2) $\overline{E^\circ} = \overline{E}$.

证明

已知 $\overline{E^\circ} \subset \overline{E}$, 只需证明 $\overline{E} \subset \overline{E^\circ}$. 显见 $\overline{E}$ 是 $\mathscr{X}$ 的一个真闭凸子集, 因而

$$\overline{E} = \{x \in \mathscr{X} : P(x) \leq 1\}$$

而由 (1) 已经有 $x \in E^\circ \Leftrightarrow P(x) < 1$, 往证 $P(x) = 1 \Rightarrow x \in \overline{E^\circ}$. 根据定义有

$$\inf\{\lambda > 0 : \frac{x}{\lambda} \in E\} = 1$$

根据下确界的定义与凸性知:

$$\frac{x}{1 + \frac{1}{n}} \in E,\ n \in \mathbb{N}$$

这说明 $P(\frac{x}{1+\frac{1}{n}}) \leq \frac{1}{1+\frac{1}{n}} < 1$, 因而由 (1) 知 $\frac{x}{1+\frac{1}{n}} \in E^\circ$, 令 $n \to \infty$ 即知 $x \in \overline{E}^\circ$. 综上知

$$P(x) \leq 1 \Rightarrow x \in \overline{E^\circ}$$

故 $\overline{E} \subset \overline{E^\circ}$, 命题得证.

练习 2.40★ 求证: 在 B 空间中, 列紧集的凸包是列紧集.

证明

设 $N \subset B$ 是列紧集, 根据凸包的定义知

$$\mathrm{co}(N) = \{\sum_{i=1}^n \lambda_i x_i : \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, x_i \in N, i = 1, \cdots, n, \forall n \in \mathbb{N}\}$$

根据 Hausdorff 定理, 只需证明 $\mathrm{co}(N)$ 对任意的 ε 都有有限 ε 网, 进一步根据练习(2.13)只需证明 $\mathrm{co}(N)$ 对任意的 ε 都有列紧 ε 网. 因为 N 列紧, 故可取 N 的 ε 网 $N_\varepsilon = \{y_1, \cdots, y_n\}$, 下证 $\mathrm{co}(N_\varepsilon)$ 列紧.

取 $y_p = \sum_{k=1}^n \lambda_k^p y_k^p, \sum_{k=1}^n \lambda_k^p = 1, \lambda_k \geq 0 (1 \leq k \leq n), p \in \mathbb{N}$. 因为 $y_1, \cdots, y_n$ 只有有限个, 而 $y_k^p \in \{y_1, \cdots, y_n\}$, 故依据鸽笼原理, 在序列 $\{y_p\}$ 中, 总有某个 $y_{i_1}, \cdots, y_{i_q}$ 的加和组合出现过无数次, 不妨就设形如 $\lambda_1 y_1 + \cdots + \lambda_q y_q$ 的项出现过无数次, 为简化符号不妨就设 $y_p = \lambda_1^p y_1 + \cdots + \lambda_q^p y_q, p \in \mathbb{N}$. 现在在 $\{\lambda_1^p\}$ 中选出收敛子列 $\{\lambda_1^{p_{1k}}\}$, 在 $\{\lambda_2^{p_{k1}}\}$ 中选出收敛子列 $\{\lambda_2^{p_{k2}}\}$, 一直如此, 直到在 $\{\lambda_n^{p_{k,n-1}}\}$ 中选出收敛子列 $\{\lambda_n^{p_{kn}}\}$, 知 $\{y_{p_{kn}}\}$ 正是一个收敛子列, 故对任意的 $\varepsilon > 0$, $\mathrm{co}(N_\varepsilon)$ 列紧.

最后证明 $\mathrm{co}(N_\varepsilon)$ 是 $\mathrm{co}(N)$ 的 ε 网, 这是因为任取 $y = \sum_{i=1}^n \lambda_i x_i$, 既然 N_ε 是 N 的 ε 网, 对每个 x_i 都能在 N 中

找到 $\zeta_i \in \{y_1, \cdots, y_n\}$ 使得 $\|x_i - \zeta_i\| < \varepsilon$, 进而取 $\zeta = \sum_{i=1}^n \lambda_i \zeta_i$, 有

$$\|\sum_{i=1}^n \lambda_i x_i - \sum_{i=1}^n \lambda_i \zeta_i\| \le \sum_{i=1}^n \lambda_i \|x_i - \zeta_i\| < \varepsilon \cdot \sum_{i=1}^n \lambda_i = \varepsilon$$

综上, 命题得证.

练习 2.41 设 C 是 B^* 空间 $\mathscr{X}$ 中的一个紧凸集, 映射 $T: C \to C$ 连续, 求证: T 在 C 上有一个不动点.

证明

既然 C 是紧凸集, 而连续映射把紧集映成紧集, 故 $T(C)$ 是紧集, 进而是自列紧集, 因而 T 在 C 上必有不动点.

练习 2.42 设 C 是 B 空间 $\mathscr{X}$ 中的一个有界闭凸集, 映射 $T_i: C \to \mathscr{X}(i = 1, 2)$ 适合

1. $\forall x, y \in C \Rightarrow T_1 x + T_2 y \in C$;
2. T_1 是一个压缩映射, T_2 是一个紧映射.

求证: $T_1 + T_2$ 在 C 上至少有一个不动点.

证明

希望找到 x 满足 $x = (T_1 + T_2)x$, 只需 $(I - T_1)x = T_2 x$, 进而只需 $x = (I - T_1)^{-1} \circ T_2 x$, 记 $T = (I - T_1)^{-1} \circ T_2 : C \to C$, 下证 T 是紧映射. 因为连续映射把列紧集映到列紧集, 又只需证明 $(I - T_1)^{-1}$ 是连续映射.

既然 T_1 是压缩映射, 知存在 $0 < \alpha < 1$ 使得

$$\|T(y_2 - y_1)\| \le \alpha \|y_2 - y_2\|$$

根据条件 1, 对任意的 $z \in T_2(C)$, 有 $T_1 y + z : C \to C$ 是关于 y 的映射, 进而根据压缩映射原理知其有唯一不动点 y:

$$T_1 y + z = y$$

再对任意的 $z_0 \in \overline{T_2(C)}$, 根据闭性知存在 $\{z_n\} \subset T_2(C)$ 满足 $z_n \to z_0 (n \to \infty)$. 现在对每个 z_n, 根据前述讨论都存在相应的 y_n 使得

$$T_1 y_n + z_n = y_n$$

为了令 $n \to \infty$ 时 $\{y_n\}$ 收敛, 下证其为基本列. 知:

$$\|y_n - y_m\| \le \|T_1 y_n - T_1 y_m\| + \|z_n - z_m\| \le \alpha \|y_n - y_m\| + \|z_n - z_m\|$$

故

$$\|y_n - y_m\| \le \frac{1}{1 - \alpha} \|z_n - z_m\|$$

故 $\{y_n\}$ 是基本列, 又 C 是闭集, 故存在 $y_0 \in C$ 满足 $y_n \to y_0 (n \to \infty)$, 得到

$$T_1 y_0 + z_0 = y_0$$

现已知

$$\forall z \in \overline{T}_2(C) \exists y \in C (T_1 y + z = y)$$

此即 $z = (I - T_1) y$ 或 $y = (I - T_1)^{-1} z$ 在 $\overline{T_2(C)}$ 上有定义. 知

$$\|z_2 - z_1\| = \|y_2 - y_1 - T_1(y_2 - y_1)\| \ge \|y_2 - y_1\| - \|T_1(y_2 - y_1)\| \ge (1 - \alpha) \|y_2 - y_1\|$$

这说明

$$\|y_2 - y_1\| \le \frac{1}{1 - \alpha} \|z_2 - z_1\|$$

代入 $y = (I - T_1)^{-1} z$ 知

$$\|(I - T_1)^{-1} z_2 - (I - T_1)^{-1} z_1\| \le \frac{1}{1 - \alpha} \|z_2 - z_1\|$$

这说明 $(I - T_1)^{-1}$ 连续, 令

$$T = (I - T_1)^{-1} \circ T_2$$

已知 T_2 是紧映射, 又因为连续映射把列紧集映到列紧集, 故 $T : C \to C$ 也是紧映射, 又因为 C 是有界闭凸集, 故 T 在 C 上有不动点, 依照最开始的分析知此即 $T_1 + T_2$ 在 C 上有不动点.

✍ **练习 2.43** 设 A 是 $n \times n$ 矩阵, 其元素 $a_{ij} > 0(1 \le i, j \le n)$, 求证: 存在 $\lambda > 0$ 及各分量非负但不全为零的向量 $x \in \mathbb{R}^n$, 使得

$$Ax = \lambda x.$$

证明

取

$$\mathbb{R}^n \supset C := \{x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1, x_i \ge 0(i = 1, \cdots, n)\}$$

考虑映射 $f(x) = \frac{Ax}{\sum_{j=1}^n (Ax)_j}$, 其中 v_j 是向量 v 的第 j 个分量. 显见 $f : C \to C$, 且 C 是有界凸集, 故只需证明 f 是紧映射. 任取 C 中的有界集 E, 显见 $f(E) \subset \mathbb{R}^n$ 依旧有界, 因而是列紧集, 故 f 是紧映射, 故其存在不动点:

$$\frac{Ax}{\sum_{j=1}^n (Ax)_j} = x \Rightarrow Ax = \sum_{j=1}^n (Ax)_j x, \ x \in C, \sum_{j=1}^n (Ax)_j > 0$$

令 $\lambda = \sum_{j=1}^n (Ax)_j$ 即证命题.

•

✍ **练习 2.44** 设 $K(x, y)$ 是 $[0, 1] \times [0, 1]$ 上的正值连续函数, 定义映射

$$(Tu)(x) = \int_0^1 K(x, y)u(y)dy \ (\forall u \in C[0, 1])$$

求证: 存在 $\lambda > 0$ 及非负但不恒为零的连续函数 u, 满足

$$Tu = \lambda u.$$

证明

取

$$C[0, 1] \supset C := \{u : \int_0^1 u(x)dx = 1, u(x) \ge 0\}$$

考虑映射

$$(T'u)(x) = \frac{\int_0^1 K(x, y)u(y)dy}{\int_0^1 dt \int_0^1 K(t, y)u(y)dy}, \ u \in C$$

显见 $T' : C \to C$ 连续[21], 且 C 是有界闭凸集, 故只需证明 T' 是紧映射.

2.6 内积空间

2.6.1 知识梳理

前面引进的是 B^* 上的代数与拓扑结构, 但并未定义角度, 即还未确定几何结构, 下面在无穷维空间上通过类似于内积的概念引入角度.

[21] 为什么?

2.6.1.1 定义与基本性质

定义 2.6.1 (共轭双线性函数, 二次型)

线性空间 $\mathscr{X}$ 上的一个二元函数 $a(\cdot,\cdot):\mathscr{X}\times\mathscr{X}\to\mathbb{K}$, 称为是共轭双线性函数, 如果

1. $a(x,\alpha_1 y_1+\alpha_2 y_2)=\overline{\alpha}_1 a(x,y_1)+\overline{\alpha}_2 a(x,y_2)$,
2. $a(\alpha_1 x_1+\alpha_2 x_2,y)=\alpha_1 a(x_1,y)+\alpha_2 a(x_2,y)$,

其中 $\forall x,y,x_1,x_2,y_1,y_2\in\mathscr{X},\forall\alpha_1,\alpha_2\in\mathbb{K}$. 还称由

$$q(x):=a(x,x),\ \forall x\in\mathscr{X}$$

定义的函数为 $\mathscr{X}$ 上由 a 诱导的二次型.

命题 2.6.1

设 a 是 $\mathscr{X}$ 上的共轭双线性函数, q 是由 a 诱导的二次型, 那么

$$q(x)\in\mathbb{R}(\forall x\in\mathscr{X})\Leftrightarrow a(x,y)=\overline{a(y,x)},\ \forall x,y\in\mathscr{X}$$

证明

当 $a(x,y)=\overline{a(y,x)}$, 知 $a(x,x)=\overline{a(x,x)}(\forall x\in\mathscr{X})$, 因而只能有 $q(x)=a(x,x)\in\mathbb{R}(\forall x\in\mathscr{X})$.

当 $q(x)\in\mathbb{R}(\forall x\in\mathscr{X})$, 知

$$q(x+y)=\overline{q(x+y)}\ (\forall x,y\in\mathscr{X})$$

拆分可知

$$a(x,y)+a(y,x)=\overline{a(x,y)}+\overline{a(y,x)}$$

将 y 换成 iy 有

$$-ia(x,y)+ia(y,x)=\overline{ia(x,y)}-\overline{ia(y,x)}\Rightarrow -a(x,y)+a(y,x)=\overline{a(x,y)}-\overline{a(y,x)}$$

联合两式可得 $a(x,y)=\overline{a(y,x)}$. □

注 ⋆ 上述命题只对 $a(\cdot,\cdot):\mathscr{X}\times\mathscr{X}\to\mathbb{C}$ 成立, 而对 $a(\cdot,\cdot):\mathscr{X}\times\mathscr{X}\to\mathbb{R}$ 不成立. 因为此时 $q(x)=a(x,x)$ 显然是实数, 但未必有 $a(x,y)=\overline{a(y,x)}(=a(y,x))$.

定义 2.6.2

线性空间 $\mathscr{X}$ 上的一个共轭双线性函数

$$(\cdot,\cdot):\mathscr{X}\times\mathscr{X}\to\mathbb{K}$$

称为是一个内积, 如果它满足:

(i) $(x,y)=\overline{(y,x)},\forall x,y\in\mathscr{X}$(共轭对称性)

(ii) $(x,x)\geq 0,\forall x\in\mathscr{X},(x,x)=0\Leftrightarrow x=\theta$(正定性)

具有内积的线性空间称为内积空间, 记作 $(\mathscr{X},(\cdot,\cdot))$.

注 若在 (ii) 中只保留非负定条件, 则称 $(\cdot,\cdot)$ 是一个半内积, 对应的空间称为半内积空间.

例 2.43($\mathbb{R}^n,\mathbb{C}^n$) $\mathbb{R}^n,\mathbb{C}^n$ 均为内积空间, 它们的内积分别为

$$(x,y)=\sum_{i=1}^n x_i y_i,\forall x,y\in\mathbb{R}^n$$

$$(x,y)=\sum_{i=1}^n x_i\overline{y_i},\forall x,y\in\mathbb{C}^n$$

其中 $x=(x_1,\cdots,x_n), y=(y_1,\cdots,y_n)$. 特别称定义了上述内积的 $\mathbb{C}$ 为酉空间.

例 2.44(l^2 空间) l^2 空间是内积空间, 规定内积

$$(x,y)=\sum_{i=1}^{\infty}x_i\overline{y}_i$$

其中 $x=(x_1,\cdots,x_n), y=(y_1,\cdots,y_n)\in l^2$.

例 2.45($L^2(\Omega,\mu)$) $L^2(\Omega,\mu)$ 是内积空间, 规定内积

$$(u,v)=\int_{\Omega}u(x)\cdot\overline{v(x)}d\mu$$

其中 $u,v\in L^2(\Omega,\mu)$.

例 2.46($C^k(\overline{\Omega})$) 在空间 $C^k(\overline{\Omega})$ 中, 规定内积

$$(u,v)=\sum_{|\alpha|\le k}\int_{\Omega}\partial^{\alpha}u(x)\cdot\overline{\partial^{\alpha}v(x)dx}, \forall u,v\in C^k(\overline{\Omega})$$

则 $(C^k(\overline{\Omega}),(\cdot,\cdot))$ 是一个内积空间. 这个空间与 Sobolev 空间有很强的联系.

从内积导出范数需要用到下述不等式.

命题 2.6.2 (Cauchy-Schwarz 不等式)

设 $(\mathscr{X},(\cdot,\cdot))$ 是内积空间, 若令

$$\|x\|=(x,x)^{\frac{1}{2}}\ (\forall x\in\mathscr{X})$$

则有

$$|(x,y)|\le\|x\|\cdot\|y\|\ (\forall x,y\in\mathscr{X}) \tag{2.43}$$

上式等号当且仅当 x,y 线性相关时成立.

(2.43)式有下述更一般的形式.

命题 2.6.3

设 a 是线性空间 $\mathscr{X}$ 上的共轭双线性函数, $q(x)$ 是由 a 诱导的二次型. 如果

$$q(x)\ge 0(\forall x\in\mathscr{X})\ \text{且}\ q(x)=0\Leftrightarrow x=\theta$$

则

$$|a(x,y)|\le[q(x)q(y)]^{\frac{1}{2}}, \forall x,y\in\mathscr{X} \tag{2.44}$$

上式等号当且仅当 x,y 线性相关时成立.

注 ⋆ 上述命题需要额外补充条件 $a\in\mathbb{C}$. 在 $a\in\mathbb{R}$ 时, 需要补充条件: $a(x,y)=a(y,x)$.

证明

不妨设 $y\neq\theta$, 任取 $\lambda\in\mathbb{K}$, 考察

$$q(x+\lambda y)=q(x)+\overline{\lambda}a(x,y)+\lambda a(y,x)+|\lambda|^2q(y)\ge 0 \tag{2.45}$$

现取 $\lambda=-\frac{a(x,y)}{q(y)}$, 因为 $a(x,y)=\overline{a}(y,x)$, 故

$$q(x)-\frac{2|a(x,y)|^2}{q(y)}+\frac{|a(x,y)|^2}{q(y)}\ge 0\Rightarrow q(x)q(y)\ge|a(x,y)|^2$$

此即(2.44)式. 当 $x=-\lambda y(\lambda\in\mathbb{K})$ 时, 上式等号成立, 反之上式等号成立时, 知(2.45)式中的等号成立, 因而 $x+\lambda y=$

$0, x = -\lambda y(\lambda \in \mathbb{K})$, 亦即 x 与 y 线性相关. □

注★ 特别对于 L^2 空间而言, Cauchy-Schwarz 不等式与 Hölder 不等式等价. 下面先回忆 Hölder 不等式:

补充命题 2.6.1 (Hölder 不等式 (指标为 (2,2)))

对任意 $f, g \in L^2(\mathbb{R}^n)$ 有

$$\int_{\mathbb{R}^n} |f(x)g(x)|dx \leq \|f\|_{L^2}\|g\|_{L^2}$$

现在说明等价性. 当 Hölder 不等式成立, 知

$$|\int_{\mathbb{R}^n} f(x)g(x)dx| \leq \int_{\mathbb{R}^n} |f(x)g(x)|dx \leq \|f\|_{L^2}\|g\|_{L^2}$$

此即 Cauchy-Schwarz 不等式. 而当 Cauchy-Schwarz 不等式成立, 知

$$\int_{\mathbb{R}^n} |f(x)g(x)|dx = |\int_{\mathbb{R}^n} |f(x)g(x)|dx| \leq \||f|\|_{L^2}\||g|\|_{L^2} = \|f\|_{L^2}\|g\|_{L^2}$$

此即 Hölder 不等式. □

命题 2.6.4

若内积空间 $(\mathscr{X}, (\cdot,\cdot))$ 按 $\|x\| = (x, x)^{\frac{1}{2}}$ 定义范数, 则其为 B^* 空间.

证明

只需验证 $\|\cdot\|$ 确为范数. 正定性与齐次性显然, 下证三角不等式:

$$\begin{aligned}\|x+y\|^2 &= (x+y, x+y) = (x,x) + (x,y) + (y,x) + (y,y) \\ &\leq \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \ (\forall x, y \in \mathscr{X})\end{aligned}$$

□

命题 2.6.5

在内积空间 $(\mathscr{X}, (\cdot,\cdot))$ 中, 内积 (x, y) 是 $\mathscr{X} \times \mathscr{X}$ 上关于范数 $\|\cdot\|$ 的连续函数.

证明

设 $x_n \to x, y_n \to y$, 显见 $\|x_n\|, \|y_n\|$ 均有界, 设它们的一个上界为 M, 知

$$\begin{aligned}|(x_n, y_n) - (x, y)| &\leq |(x_n, y_n) - (x, y_n)| + |(x, y_n) - (x, y)| = |(x_n - x, y_n)| + |(x, y_n - y)| \\ &\leq \|x_n - x\|\cdot\|y_n\| + \|x\|\cdot\|y_n - y\| \leq M\|x_n - x\| + \|x\|\cdot\|y_n - y\| \to 0,\ n \to \infty\end{aligned}$$

□

命题 2.6.6

内积空间 $(\mathscr{X}, (\cdot,\cdot))$ 是严格凸的 B^* 空间.

证明

任取 $0 < \lambda < 1$, 当 $\|x\| = \|y\| = 1, x \neq y$ 时有

$$\|\lambda x + (1-\lambda)y\|^2 = \lambda^2\|x\|^2 + 2\lambda(1-\lambda)\operatorname{Re}(x,y) + (1-\lambda)^2\|y\|^2$$

现希望 $\operatorname{Re}(x, y) < \|x\|\|y\|$, 依照 Cauchy-Schwarz 不等式的正常程序只能有

$$\operatorname{Re}(x,y) \leq |(x,y)| \leq \|x\|\|y\|$$

当不存在 $a \in \mathbb{K}$ 使得 $x = ay$, 知 x, y 不可能线性相关, 从而 Cauchy-Schwarz 不等式的等号无法取到. 而若存在 $\alpha \in \mathbb{K}$ 使得 $x = \alpha y$, 则根据 $\|x\| = \|y\| = 1$ 知

$$1 = \|x\| = |\alpha|\|y\| = |\alpha| \Rightarrow \alpha = e^{i\theta}(0 < \theta \leq 2\pi)$$

又因为 $x \neq y$, 故 $\theta \in (0, 2\pi)$, 从而

$$\operatorname{Re}(x, y) = \operatorname{Re}(\alpha y, y) = \operatorname{Re}\alpha = \cos\theta < 1 = \|x\|\|y\|$$

综上有 $\operatorname{Re}(x, y) < \|x\|\|y\| = 1$, 因而

$$\lambda^2\|x\|^2 + 2\lambda(1-\lambda)\operatorname{Re}(x, y) + (1-\lambda)^2\|y\|^2 < (\lambda + (1-\lambda))^2 = 1$$

□

前面考虑的是在内积空间引入范数达到 B^* 空间, 现在考虑在 B^* 空间中引入内积满足

$$(x, x)^{\frac{1}{2}} = \|x\|, \forall x \in \mathscr{X}.$$

命题 2.6.7

在 B^* 空间 $(\mathscr{X}, \|\cdot\|)$ 中, 为了在 $\mathscr{X}$ 上可引入一个内积 $(\cdot,\cdot)$ 满足

$$(x, x)^{\frac{1}{2}} = \|x\|, \ \forall x \in \mathscr{X}$$

必须且仅须范数 $\|\cdot\|$ 满足平行四边形等式:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \forall x, y \in \mathscr{X}.$$

证明

引入内积后容易验证其满足平行四边形等式, 下面当范数满足平行四边形等式, 可设

$$(x, y) = \begin{cases} \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2), & \mathbb{K} = \mathbb{R}, \\ \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2), & \mathbb{K} = \mathbb{C}. \end{cases}$$

下证 $(\cdot,\cdot)$ 确为内积, 这需要验证共轭双线性函数的线性性, 齐次性与内积的共轭对称性, 正定性. 当 $\mathbb{K} = \mathbb{R}$ 时, 共轭对称性与正定性显然, 且易知 $(0, z) = 0$. 下面验证线性性, 注意任取 $x, y, z \in \mathscr{X}$ 有

$$\begin{aligned}
(x, z) + (y, z) &= \frac{1}{4}(\|x+z\|^2 - \|x-z\|^2 + \|y+z\|^2 - \|y-z\|^2) \\
&= \frac{1}{4}((\|\frac{x+y}{2} + z + \frac{x-y}{2}\|^2 + \|\frac{x+y}{2} + z - \frac{x-y}{2}\|^2) - (\|\frac{x+y}{2} - z + \frac{x-y}{2}\|^2 + \|\frac{x+y}{2} - z - \frac{x-y}{2}\|^2)) \\
&\overset{(i)}{=} \frac{1}{4}(2(\|\frac{x+y}{2} + z\|^2 + \|\frac{x-y}{2}\|^2) - 2(\|\frac{x+y}{2} - z\|^2 + \|\frac{x-y}{2}\|^2)) \\
&= \frac{1}{2}(\|\frac{x+y}{2} + z\|^2 - \|\frac{x+y}{2} - z\|^2) = 2(\frac{x+y}{2}, z)
\end{aligned}$$

其中 (i) 是平行四边形等式. 在上式中取 $y = 0$ 得

$$(x, z) = 2(\frac{x}{2}, z)$$

再将 x 换为 $x + y$ 即得

$$2(\frac{x+y}{2}, z) = (x+y, z)$$

这说明

$$(x, z) + (y, z) = (x+y, z)$$

线性性得证. 下面验证齐次性, 对给定的 $x, z \in \mathscr{X}$ 与任意 $t \in \mathbb{R}$, 记

$$f(t) = (tx, z)$$

根据线性性易知 $f(t)$ 满足:

$$f(t_1 + t_2) = f(t_1) + f(t_2), \ t_1, t_2 \in \mathbb{R}$$

另外, 当 $t_n \to t (n \to \infty)$ 时, 有

$$|\|t_n x \pm z\| - \|tx \pm z\|| \leq \|t_n x - tx\| = |t_n - t|\|x\| \to 0, \ n \to \infty$$

从而 $|f(t) - f(t_n)| = |f(t - t_n)| \to 0 (n \to \infty)$, 因而 f 连续. 现在断言: 若连续函数 $f(t)$ 满足 $f(t_1 + t_2) = f(t_1) + f(t_2)$,

则 $f(t)=f(1)t$. 这是因为对线性性归纳可得 $f(nt)=nf(t)(\forall n\in\mathbb{N},\forall t\in\mathbb{R})$, 进一步可得

$$f(\frac{n}{m})=nf(\frac{1}{m})=\frac{n}{m}f(1),\ \forall n,m\in\mathbb{N}$$

另外, 取 $t_1=t_2=0$ 易得 $f(0)=0$, 进一步可得 $f(-t)=-f(t)$, 从而

$$f(t)=tf(1),\ \forall t\in\mathbb{Q}$$

最后根据 f 的连续性与 $\mathbb{Q}$ 在 $\mathbb{R}$ 中的稠性即知

$$f(t)=tf(1),\ \forall t\in\mathbb{R}$$

回到原命题, 上述断言说明

$$(tx,z)=t(x,z),\ \forall x,z\in\mathscr{X},\forall t\in\mathbb{R}$$

齐次性得证, 因而 $(\cdot,\cdot)$ 是内积, 且显见 $(x,x)=\|x\|^2$, 这说明 $\mathscr{X}$ 是实赋范线性空间时, 仅依据范数 $\|\cdot\|$ 与平行四边形等式即可诱导内积 $(\cdot,\cdot)$.

当 $\mathscr{X}$ 是复赋范线性空间, 回忆构造

$$\begin{aligned}(x,y)_2&=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2+i\|x+iy\|^2-i\|x-iy\|^2)\\&=(x,y)+i(x,iy)\end{aligned}$$

其中 $(\cdot,\cdot)$ 是 $\mathscr{X}$ 为实赋范线性空间时对应的内积构造, 下面说明 $(\cdot,\cdot)_2$ 是 $\mathscr{X}$ 上的内积. 根据 $(\cdot,\cdot)$ 已获证的线性性, 知

$$(x,z)_2+(y,z)_2=(x+y,z)_2$$

根据 $(\cdot,\cdot)$ 已获证的齐次性, 知

$$(\alpha x,y)_2=\alpha(x,y)_2,\ \forall\alpha\in\mathbb{R}$$

另外容易验证

$$(ix,y)_2=i(x,y)_2$$

因而

$$(\alpha x,y)_2=\alpha(x,y)_2,\ \forall\alpha\in\mathbb{C}$$

从而 $(\cdot,\cdot)_2$ 具齐次性. 对于共轭对称性知:

$$\overline{(y,x)_2}=(y,x)-i(y,ix)=(x,y)-i(ix,y)\overset{\text{(i)}}{=}(x,y)+i(x,iy)$$

其中 (i) 是直接将 x,iy 代入 $(\cdot,\cdot)$ 的定义得到的, 共轭对称性进而得证. 综上知 $(\cdot,\cdot)_2$ 是 $\mathscr{X}$ 上的内积, 且显见 $(x,x)=\|x\|^2$, 命题得证. □

定义 2.6.3 (Hilbert 空间)

完备的内积空间称为 Hilbert 空间. ♣

下面介绍一个在偏微分方程边值问题理论中特别有用的内积空间: $H_0^m(\Omega)$, 为此需要先证明下述不等式.

引理 2.6.1 (Poincaré 不等式)

设 $C_0^m(\Omega)$ 表示有界开区域 $\Omega\subset\mathbb{R}^n$ 上一切 m 次连续可微, 并在边界 $\partial\Omega$ 的某邻域内为 0 的函数集合, 即:

$$C_0^m(\Omega)=\{u\in C^m(\overline{\Omega}):u(x)=0,x\in\partial\Omega\text{的某邻域}\}$$

则 $\forall u\in C_0^m(\Omega)$ 有

$$\sum_{|\alpha|<m}\int_\Omega|\partial^\alpha u(x)|^2dx\le C\sum_{|\alpha|=m}\int_\Omega|\partial^\alpha u(x)|^2dx \tag{2.46}$$

其中 C 是仅依赖于区域 Ω 与 m 的常数. ♡

证明

因为 Ω 有界, 可将其放在某个边长为 a 的立方体 Ω_1 内, 进一步在适当选取坐标系后记

$$\Omega_1 = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n | 0 \le x_i \le a (i = 1, 2, \cdots, n)\}.$$

在 $\Omega_1 \backslash \Omega$ 上补充定义 $u = 0$, 补充定义后, $u(x)$ 在 Ω_1 上 m 次连续可微, 且在边界上等于 0. 对任意的 $x = (x_1, \cdots, x_n) \in \Omega_1$ 有:

$$u(x) = \int_0^{x_1} \frac{\partial u}{\partial t}(t, x_2, \cdots, x_n) dt$$

再根据 Cauchy-Schwarz 不等式有

$$|u(x)| \le \int_0^a |\frac{\partial u}{\partial x_1}| dx_1 = \int_0^a |\frac{\partial u}{\partial x_1}| \cdot 1 dx_1 =\le (\int_0^a |\frac{\partial u}{\partial x_1}|^2 dx_1)^{\frac{1}{2}} (\int_0^a 1^2 dx_1)^{\frac{1}{2}} \Rightarrow |u(x)|^2 \le a \int_0^a |\frac{\partial u}{\partial x_1}|^2 dx_1$$

在 Ω_1 上对上述不等式积分有

$$\begin{aligned} \int_{\Omega_1} |u(x)|^2 dx &\le \int_0^a dx_1 \int_0^a dx_2 \cdots \int_0^a dx_n \int_0^a a |\frac{\partial u}{\partial x_1}|^2 dx_1 = \int_0^a dx_1 \int_{\Omega_1} a |\frac{\partial u}{\partial x_1}|^2 dx \\ &= a^2 \int_{\Omega_1} |\frac{\partial u}{\partial x_1}|^2 dx \le a^2 \int_{\Omega_1} \sum_{k=1}^n |\frac{\partial u}{\partial x_k}|^2 dx = a^2 \int_{\Omega_1} |\nabla u(x)|^2 dx \end{aligned}$$

将上式逐次应用在 $\partial^\alpha u(x) (|\alpha| < m)$ 上, 因为(2.46)左式均满足 $|\alpha| < m$, 故总能通过对相应指标不断取偏微达到 $|\alpha| = m$, 进而得到命题. □

Poincaré 不等式(2.46)表明在 $C_0^m(\Omega)$ 上

$$\|u\|_m := (\sum_{|\alpha|=m} \int_\Omega |\partial^\alpha u(x)|^2 dx)^{\frac{1}{2}} \tag{2.47}$$

与

$$\|u\|_m := (\sum_{|\alpha|\le m} \int_\Omega |\partial^\alpha u(x)|^2 dx)^{\frac{1}{2}} \tag{2.48}$$

是一对等价范数. 记 C_0^m 按(2.47)式完备化后的空间为 $H_0^m(\Omega)$, 其为 $H^m(\Omega)$ 的一个闭子空间.

例 2.47($H_0^m(\Omega)$) $H_0^m(\Omega)$ 是一个 Hilbert 空间, 其内积定义为

$$(u, v)_m = \sum_{|\alpha|=m} \int_\Omega \partial^\alpha u(x) \cdot \overline{\partial^\alpha v(x)} dx, \forall u, v \in C_0^m(\Omega)$$

注 当 Ω 的边界 $\partial\Omega$ 具有光滑的法向导数时, $H_0^m(\Omega)$ 中的元素实际上是满足边界条件

$$u|_{\partial\Omega} = \frac{\partial u}{\partial n}|_{\partial\Omega} = \cdots = (\frac{\partial}{\partial n})^{m-1} u|_{\partial\Omega} = 0$$

的 $C^m(\Omega)$ 函数 u 的一种推广, 其中 $\frac{\partial}{\partial n}$ 是 $\partial\Omega$ 上的法向导数.

2.6.1.2 正交与正交基

对内积空间中的两个向量 x, y, 用

$$\theta := \arccos \frac{|(x, y)|}{\|x\| \cdot \|y\|}$$

表示它们之间的夹角.

注 * 注意当背景域为 $\mathbb{C}$ 时, $\theta \in [0, \frac{\pi}{2}]$; 而当背景域为 $\mathbb{R}$ 时, $\theta \in [0, \pi]$.

定义 2.6.4 (正交, 正交补)

内积空间 $\mathscr{X}$ 上的两个元素 x 与 y 称作是正交的, 如果

$$(x, y) = 0$$

记作 $x \perp y$. 又设 M 是 $\mathscr{X}$ 的一个非空子集, $x \in \mathscr{X}$. 若对 $\forall y \in M$ 都有 $x \perp y$, 则称 x 与 M 正交, 记作 $x \perp M$. 此外称集合

$$\{x \in \mathscr{X} : x \perp M\}$$

为 M 的正交补, 记作 $M^{\perp}$. ♣

通过定义可以证明下述命题.

命题 2.6.8

设 $\mathscr{X}$ 是内积空间, M 是 $\mathscr{X}$ 的一个非空子集.

(i) 若 $x \perp y_i (i = 1, 2)$, 则

$$x \perp \lambda_1 y_1 + \lambda_2 y_2, \forall \lambda_1, \lambda_2 \in \mathbb{K}.$$

(ii) 若 $x = y + z, y \perp z$, 则

$$\|x\|^2 = \|y\|^2 + \|z\|^2.$$

(iii) 若 $x \perp y_n (n \in \mathbb{N}), y_n \to y$, 则 $x \perp y$.

(iv) 若 $x \perp M$, 则 $x \perp \operatorname{span} M$.

(v) $M^{\perp}$ 是 $\mathscr{X}$ 的一个闭线性子空间.

(vi) $\star \forall M \in \mathscr{X} (\overline{M}^{\perp} = M^{\perp})$. ♠

性质 (vi) 的证明

因为 $M \subset \overline{M}$, 根据正交补的定义显见 $\overline{M}^{\perp} \subset M^{\perp}$, 下证 $M^{\perp} \subset \overline{M}^{\perp}$. 任取 $y \in M^{\perp}$, 根据闭包的定义知对任意 $x \in \overline{M}$ 均存在序列 $\{x_n\}_{n \in \mathbb{N}} \subset M$ 使得 $x_n \to x (n \to \infty)$. 注意 $x_n \in M \Rightarrow (x_n, y) = 0$, 因而根据内积的连续性知 $(x, y) = 0$, 亦即 $y \in \overline{M}^{\perp}$, 命题得证. □

下面推广欧氏空间中的直角坐标系.

定义 2.6.5 (正交集, 正交规范集, 完备)

设 $\mathscr{X}$ 是一个内积空间, 集合 $S = \{e_\alpha : \alpha \in A\}$ 是 $\mathscr{X}$ 的一个子集. 称 S 为正交集, 如果

$$e_\alpha \perp e_\beta \ (\alpha \neq \beta, \forall \alpha, \beta \in A)$$

如果还有 $\|e_\alpha\| = 1 (\forall \alpha \in A)$, 则称 S 为正交规范集. 又如果在 $\mathscr{X}$ 中不存在非零元与 S 正交, 即 $S^{\perp} = \{\theta\}$, 则称 S 是完备的. ♣

为了研究一个内积空间是否一定有完备正交集, 首先引入 Zorn 引理.

补充定义 2.6.1 (半序 (偏序), 可比较 ⋆)

赋等价关系 $=$ 的[a]集合 S 上的一个二元关系称为一个偏序, 记作 $\leq$, 如果:

(i) 反身性: $\forall a \in S (a \leq a)$;

(ii) 反对称性: 若 $a \leq b$ 且 $b \leq a$, 则 $a = b$.

(iii) 传递性: 若 $a \leq bb \leq c$, 则 $a \leq c$.

若 $x, y \in S$ 有关系 $x \leq y$ 或 $y \leq x$, 就称 x, y 可比较.

[a] 是否一定需要? 偏序不能诱导等价关系吧? ♣

补充定义 2.6.2 (全序子集 ⋆)

$\mathscr{X}$ 的子集 A 称为是全序子集, 如果 A 中的任意两个元素均可比较.

引理 2.6.2 (Zorn)

设 $\mathscr{X}$ 是一个半序集. 如果它的每一个全序子集有一个上界, 那么 $\mathscr{X}$ 有一个极大元.

命题 2.6.9

非 $\{\theta\}$ 内积空间 $\mathscr{X}$ 中必存在完备正交集.

证明

因为 $\mathscr{X} = \{\theta\}$, 故 $\mathscr{X}$ 中的正交集根据包含关系 $\subset$ 构成一个半序集类, 并且每个全序子集类都有一个上界, 即它们的并集. 根据 Zorn 引理2.6.2, 这个半序集类有极大元, 下面证明该极大元 S 正是完备正交集. 如若不然, 则存在 $x_0 \in S^{\perp}, x_0 \neq \theta$, 取 $S_1 = \{x_0\} \cup S$, 知 S_1 依旧是正交集, 且 $S \subsetneq S_1$, 这与 S 的极大性矛盾! □

定义 2.6.6 (基, 封闭, Fourier 系数)

内积空间 $\mathscr{X}$ 中的正交规范集 $S = \{e_\alpha : \alpha \in A\}$ 称为一个基 (或封闭的) 是指 $\forall x \in \mathscr{X}$, 有下列表示:

$$x = \sum_{\alpha \in A} (x, e_\alpha) e_\alpha$$

其中 $\{(x, e_\alpha) : \alpha \in A\}$ 称为 x 关于基 $\{e_\alpha : \alpha \in A\}$ 的 Fourier 系数.

定理 2.6.1 (Bessel 不等式)

设 $\mathscr{X}$ 是一个内积空间. 如果 $S = \{e_\alpha : \alpha \in A\}$ 是 $\mathscr{X}$ 中的正交规范集, 那么 $\forall x \in \mathscr{X}$, 有

$$\sum_{\alpha \in A} |(x, e_\alpha)|^2 \leq \|x\|^2.$$

证明

首先考虑 A 中的任意有限子集, 不妨设 $\alpha \in \{1, \cdots, n\}$, 下面证明

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2 \tag{2.49}$$

这是因为

$$0 \leq \|x - \sum_{i=1}^{n} (x, e_i) e_i\|^2 = (x - \sum_{i=1}^{n} (x, e_i) e_i, x - \sum_{j=1}^{n} (x, e_j) e_j) = \|x\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2$$

因而 $\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2$, 这说明任取 $n \in \mathbb{N}$, 满足 $|(x, e_\alpha)| > \frac{1}{n}$ 的 $\alpha \in A$ 至多只有有限个. 更具体地, 对任意 $n \in \mathbb{N}$, 满足 $|(x, e_\alpha)| > \frac{1}{n}$ 的指标 $\alpha \in A$ 至多只有 $[n^2\|x\|^2] =: N_0$, 其中 $[x](x \in \mathbb{R})$ 表示不超过 x 的最大整数. 这是因为若存在指标列 $\{\alpha_i\}_{i=1}^{N_0+1}$ 使得 $|(x, e_{\alpha_i})| > \frac{1}{n}(1 \leq i \leq N_0 + 1)$, 则根据(2.49)式知 $\sum_{i=1}^{N_0+1} |(x, e_{\alpha_i})|^2 \leq \|x\|^2$, 从而

$$\frac{N_0 + 1}{n^2} < \sum_{i=1}^{N_0+1} |(x, e_{\alpha_i})|^2 \leq \|x\|^2 \Rightarrow N_0 + 1 < n^2\|x\|^2$$

但根据构造有 $N_0 + 1 > n^2\|x\|^2$, 矛盾! 从而令 n 遍历 $\mathbb{N}$, 知满足 $(x, e_\alpha) \neq 0$ 的 $\alpha \in A$ 至多只有可数个. 故 $\sum_{\alpha \in A} |(x, e_\alpha)|^2$ 实际上是至多可数项求和的级数, 又因为对 A 的任意有限子集 A_f 都有

$$\sum_{\alpha \in A_f} |(x, e_\alpha)|^2 \leq \|x\|^2$$

由极限的保号性即得欲证. □

推论 2.6.1 (勾股定理)

设 $\mathscr{X}$ 是 Hilbert 空间, 且 $\{e_\alpha : \alpha \in A\}$ 是 $\mathscr{X}$ 中的正交规范集, 则对 $\forall x \in \mathscr{X}$, 有

$$\sum_{\alpha \in A}(x, e_\alpha)e_\alpha \in \mathscr{X}, \tag{2.50}$$

且

$$\|x - \sum_{\alpha \in A}(x, e_\alpha)e_\alpha\|^2 = \|x\|^2 - \sum_{\alpha \in A}|(x, e_\alpha)|^2.$$

注 ★ 注意从上述定理的证明过程中仅仅可以得知 $\sum_{\alpha \in A}(x, e_\alpha)e_\alpha$ 中至多有可列个非零项, 这并不意味着 $\{e_\alpha : \alpha \in A\}$ 一定是可列的, 亦即 $\{e_\alpha : \alpha \in A\}$ 可以不可列.

证明

不妨设使得 $(x, e_\alpha) \neq 0$ 的可数多个 $\alpha \in A$ 为 $1, 2, \cdots, n, \cdots$, 则

$$\sum_{\alpha \in A}(x, e_\alpha)e_\alpha = \sum_{n=1}^{\infty}(x, e_n)e_n$$

根据 Bessel 不等式(2.6.1)知 $\sum_{n=1}^{\infty}|(x, e_n)|^2$ 收敛, 故

$$\|\sum_{n=m}^{m+p}(x, e_n)e_n\|^2 = \sum_{n=m}^{m+p}|(x, e_n)|^2 \to 0\ (m \to \infty, \forall p \in \mathbb{N})$$

记 $x_m = \sum_{n=1}^{m}(x, e_n)e_n$, 知 $\{x_m\}$ 是基本列, 从而

$$\sum_{\alpha \in A}(x, e_\alpha)e_\alpha = \sum_{n=1}^{\infty}(x, e_n)e_n = \lim_{m \to \infty} x_m \in \mathscr{X}$$

下面说明 $x - \sum_{n=1}^{\infty}(x, e_n)e_n \perp \sum_{n=1}^{\infty}(x, e_n)e_n$, 这是因为:

$$\begin{aligned}
&(x - \sum_{m=1}^{N}(x, e_n)e_n, \sum_{n=1}^{N}(x, e_n)e_n)\\
=&(x, \sum_{n=1}^{N}(x, e_n)e_n) - \sum_{n=1}^{N}(x, e_n)(e_n, \sum_{n=1}^{N}(x, e_n)e_n)\\
=&\sum_{n=1}^{N}\overline{(x, e_n)}(x, e_n) - \sum_{n=1}^{N}(x, e_n)\overline{(x, e_n)} = 0
\end{aligned}$$

根据正交性与内积的连续性得到

$$\begin{aligned}
\|x - \sum_{n=1}^{\infty}(x, e_n)e_n\|^2 &= (x - \sum_{n=1}^{\infty}(x, e_n)e_n, x - \sum_{n=1}^{\infty}(x, e_n)e_n)\\
&= (x - \sum_{n=1}^{\infty}(x, e_n)e_n, x) - (x - \sum_{n=1}^{\infty}(x, e_n)e_n, \sum_{n=1}^{\infty}(x, e_n)e_n)\\
&= (x, x) - (\sum_{n=1}^{\infty}(x, e_n)e_n, x) = (x, x) - \sum_{n=1}^{\infty}(x, e_n)(e_n, x)\\
&= \|x\|^2 - \sum_{n=1}^{\infty}(x, e_n)\overline{(e_n, x)} = \|x\|^2 - \sum_{n=1}^{\infty}|(x, e_n)|^2
\end{aligned}$$

此即欲证. □

注 ★ 结合 Bessel 不等式(2.6.1)与其推论勾股定理2.6.1, 可以证明正交规范基的定义(2.6.6)中级数的收敛是无条件收敛, 亦即任意重排后, $\sum_{\alpha \in A}(x, e_\alpha)e_\alpha$ 仍然收敛到 x.

证明

由 Bessel 不等式的证明知, 任取 $x \in \mathscr{X}$, 满足 $(x, e_\alpha) \neq 0$ 的 α 只有可数个, 不妨设这样的 e_α 为 $\{e_i\}_{i=1}^{N}$, 其中

$N \in \mathbb{N}$ 或 $N = \infty$. 当 $N < \infty$ 时, 欲证显然成立, 下设 $N = \infty$, 亦即已知

$$x = \sum_{\alpha \in A} (x, e_\alpha) e_\alpha = \sum_{i=1}^{\infty} (x, e_i) e_i$$

根据勾股定理知

$$0 = \|\theta\|^2 = \|x - \sum_{i=1}^{\infty} (x, e_i) e_i\|^2 = \|x\|^2 - \sum_{i=1}^{\infty} |(x, e_i)|^2$$

因而

$$\sum_{i=1}^{\infty} |(x, e_i)|^2 = \|x\|^2 < \infty$$

这说明级数 $\sum_{i=1}^{n} |(x, e_i)|^2$ 绝对收敛, 而数项级数的绝对收敛等价于无条件收敛, 故对 $\mathbb{N}$ 的任意重排 $\{a(i)\}_{i=1}^{\infty}$ 均有:

$$\sum_{i=1}^{\infty} |(x, e_{a(i)})|^2 = \sum_{i=1}^{\infty} |(x, e_i)|^2$$

故根据勾股定理知:

$$\|x - \sum_{i=1}^{\infty} (x, e_{a(i)}) e_{a(i)}\|^2 = \|x\|^2 - \sum_{i=1}^{\infty} |(x, e_{a(i)})|^2 = \|x\|^2 - \sum_{i=1}^{\infty} |(x, e_i)|^2 = 0$$

故 $x = \sum_{i=1}^{\infty} (x, e_{a(i)}) e_{a(i)}$, 欲证成立. □

下面说明 Bessel 不等式(2.6.1)取等号与 $\sum_{\alpha \in A} (x, e_\alpha) e_\alpha = x$ 的条件.

定理 2.6.2

设 $\mathscr{X}$ 是一个 Hilbert 空间, 若 $S = \{e_\alpha | \alpha \in A\}$ 是 $\mathscr{X}$ 中的正交规范集, 则下述三条等价:

1. S 是封闭的;
2. S 是完备的;
3. Parseval 等式:

$$\|x\|^2 = \sum_{\alpha \in A} |(x, e_\alpha)|^2, \ \forall x \in \mathscr{X} \tag{2.51}$$

成立.

证明

1. ⇒ 2. 若 S 不完备, 则存在 $x \in \mathscr{X} \backslash \{\theta\}$, 使得

$$(x, e_\alpha) = 0 \ (\forall \alpha \in A)$$

但根据封闭性, 知

$$x = \sum_{\alpha \in A} (x, e_\alpha) e_\alpha = \theta$$

矛盾!

2. ⇒ 3. 若存在 $x \in \mathscr{X}$ 不满足 Parseval 等式, 则由推论(2.6.1)知

$$\|x - \sum_{\alpha \in A} (x, e_\alpha) e_\alpha\|^2 = \|x\|^2 - \sum_{\alpha \in A} |(x, e_\alpha)|^2 > 0$$

这说明 $y := x - \sum_{\alpha \in A} (x, e_\alpha) e_\alpha \neq \theta$, 但任取 $e_\alpha \in S$ 知 $(y, e_\alpha) = 0$, 故 $y \in S^\perp$, 这说明 S 不完备, 矛盾!

3. ⇒ 1. 根据 Parseval 等式(2.51)与推论(2.6.1)知

$$\|x - \sum_{\alpha \in A} (x, e_\alpha) e_\alpha\|^2 = \|x\|^2 - \sum_{\alpha \in A} |(x, e_\alpha)|^2 = 0, \ \forall x \in \mathscr{X}$$

故

$$x = \sum_{\alpha \in A} (x, e_\alpha) e_\alpha, \ \forall x \in \mathscr{X}$$

此即 S 封闭. $\square$

例 2.48 在 $L^2[0, 2\pi]$ 上,

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \ n = 0, \pm 1, \pm 2, \cdots$$

是一组正交规范基. $\forall u \in L^2[0, 2\pi]$, 其对应的 Fourier 系数是

$$(u, e_n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} u(t) e^{-int} dt, \ n = 0, \pm 1, \pm 2, \cdots$$

注 ★[JD] 中给出三角多项式在 $L^2[0, 2\pi]$ 中稠密, 进一步可以推出 $\{e^{int}\}_{n\in\mathbb{Z}}$ 在 $L^2[0, 2\pi]$ 中完备. 这是因为若设 $f \in L^2[0, 2\pi]$ 且 $f \perp \{e^{int}\}_{n\in\mathbb{Z}}$, 显见 $f \perp \mathrm{span}\{e^{int}\}_{n\in\mathbb{Z}}$, 进一步由内积的连续性知 $f \perp \overline{\mathrm{span}\{e^{int}\}_{n\in\mathbb{Z}}}$. 又根据 Fourier 级数的相关理论知 $\overline{\mathrm{span}\{e^{int}\}_{n\in\mathbb{Z}}} = L^2[0, 2\pi]$, 从而 $f \perp f$, 进而只能有 $f = \theta$. 这说明 $\{e^{int}\}_{n\in\mathbb{Z}}$ 在 $L^2[0, 2\pi]$ 中完备, 从而 $\{\frac{e^{int}}{\sqrt{2\pi}}\}_{n\in\mathbb{Z}}$ 是 $L^2[0, 2\pi]$ 的正交规范基.

例 2.49 在 l^2 空间上,

$$e_n = (\underbrace{0, 0, \cdots, 0, 1}_{n}, 0, \cdots), \ n = 1, 2, 3, \cdots$$

是一组正交规范基.

证明

显见 $\{e_n\}_{n\in\mathbb{N}}$ 规范正交. 任取 $x = (x_1, \cdots, x_n, \cdots) \in l^2$, 知

$$\|x\|_{l^2}^2 = \sum_{i=1}^{\infty} |x_i|^2 = \sum_{i=1}^{\infty} |(x, e_i)|^2$$

这说明 Parseval 等式成立, 因而 $\{e_n\}_{n\in\mathbb{N}}$ 是 l^2 的一组正交规范基. $\square$

例 2.50 设 D 是 $\mathbb{C}$ 中的单位开圆域, $H^2(D)$ 表示在 D 内满足

$$\iint_D |u(z)|^2 dxdy < \infty (z = x + iy)$$

的解析函数全体组成的空间. 规定内积为

$$(u, v) = \iint_D u(z)\overline{v(z)} dxdy$$

此时函数组

$$\varphi_n(z) = \sqrt{\frac{n}{\pi}} z^{n-1}, \ n = 1, 2, 3, \cdots$$

是一组正交规范基. 设 $u(z) = \sum_{k=0}^{\infty} b_k z^k \in H^2(D)$, 其对应的 Fourier 系数是

$$(u, \varphi_n) = \iint_D \sum_{k=0}^{\infty} b_k z^k \sqrt{\frac{n}{\pi}} \bar{z}^{n-1} dxdy = \sum_{k=0}^{\infty} b_k \sqrt{\frac{\pi}{k+1}} (\varphi_{k+1}, \varphi_n) = b_{n-1} \sqrt{\frac{\pi}{n}}, \ n = 1, 2, 3, \cdots$$

2.6.1.3 正交化与 Hilbert 空间的同构

线性代数中的 Gram-Schmidt 正交化方法可以套用到内积空间中. 设 $\{x_1, x_2, \cdots\}$ 是内积空间 $\mathscr{X}$ 中的一列线性无关的元素, 希望构造一组正交规范集

$$S = \{e_i : i = 1, 2, \cdots\}$$

使得对 $\forall x \in \mathbb{N}$, e_n 是 $x_1, \cdots, x_n$ 的线性组合, 且 x_n 也是 $e_1, \cdots, e_n$ 的线性组合. 其构造过程为:

$$y_1 = x_1,\ e_1 = \frac{y_1}{\|y_1\|}$$

$$y_2 = x_2 - (x_2, e_1)e_1,\ e_2 = \frac{y_2}{\|y_2\|}$$

$$\vdots, \vdots$$

$$y_n = x_n - \sum_{i=1}^{n-1}(x_n, e_i)e_i,\ e_n = \frac{y_n}{\|y_n\|}$$

$$\vdots, \vdots$$

定义 2.6.7

设 $(\mathscr{X}_1, (\cdot,\cdot)_1)$ 与 $(\mathscr{X}_2, (\cdot,\cdot)_2)$ 是两个内积空间, 如果存在 $\mathscr{X}_1 \to \mathscr{X}_2$ 的一个线性同构 T, 满足

$$(Tx, Ty)_2 = (x, y)_1,\ \forall x, y \in \mathscr{X}_1$$

则称内积空间 $\mathscr{X}_1$ 与 $\mathscr{X}_2$ 是同构的. ♣

定理 2.6.3

为了 Hilbert 空间 $\mathscr{X}$ 是可分的, 必须且仅须它有至多可列的正交规范基 S. 又若 S 的元素个数 $N < \infty$, 则 $\mathscr{X}$ 同构于 $\mathbb{K}^N$; 若 $N = \infty$, 则 $\mathscr{X}$ 同构于 l^2. ♡

证明

当 $\mathscr{X}$ 可分, 设其可列稠密子集为 $\{x_n\}_{n=1}^{\infty}$, 取线性包知必存在一个线性无关的子集 $\{y_n\}_{n=1}^{N}(N \le \infty)$ 满足

$$\operatorname{span}\{y_n\}_{n=1}^{N} = \operatorname{span}\{x_n\}_{n=1}^{\infty}$$

再对 $\{y_n\}_{n=1}^{N}$ 应用 Gram-Schmidt 过程得到正交规范集 $\{e_n\}_{n=1}^{N}$. 又因为

$$\overline{\operatorname{span}\{e_n\}_{n=1}^{N}} = \overline{\operatorname{span}\{y_n\}_{n=1}^{N}} = \overline{\operatorname{span}\{x_n\}_{n=1}^{\infty}} = \mathscr{X}$$

故 $\overline{\operatorname{span}\{e_n\}_{n=1}^{N}}^{\perp} = \{\theta\}$, 依照命题2.6.8(vi) 知 $\overline{\operatorname{span}\{e_n\}_{n=1}^{N}}^{\perp} = \operatorname{span}\{e_n\}_{n=1}^{N} = \{\theta\}$, 因而 $\{e_n\}_{n=1}^{N}$ 完备, 从而其为至多可列的正交规范基.

当 $\mathscr{X}$ 有至多可数的正交规范基, 设之为 $\{e_n\}_{n=1}^{N}(N \le \infty)$, 则集合

$$\{x = \sum_{n=1}^{N} a_n e_n : \operatorname{Re} a_n, \operatorname{Im} a_n \in \mathbb{Q}\}$$

是 $\mathscr{X}$ 中的稠密子集, 因而 $\mathscr{X}$ 可分.

现在讨论正交规范基 $\{e_n\}_{n=1}^{N}(N \le \infty)$, 作对应

$$T : x \mapsto \{(x, e_n)\}_{n=1}^{N}\ (\forall x \in \mathscr{X})$$

由 Parseval 等式知

$$\|x\| = \sum_{n=1}^{N} |(x, e_n)|^2\ (\forall x \in \mathscr{X})$$

现在说明 $T : \mathscr{X} \to \mathbb{K}^N(N < \infty)$ 或 $T : \mathscr{X} \to l^2(N = \infty)$ 是双射. 针对单射, 只需证明 $Tx = \theta \Rightarrow x = \theta$. 当 $Tx = \theta$, 知 $\sum_{n=1}^{N} |(x, e_n)|^2 = 0$, 因而由 Parseval 等式知 $\|x\| = 0$, 亦即 $x = \theta$, 单射得证. 针对满射, 设 $N < \infty$ 时 $\{x_i\}_{i=1}^{N} \subset \mathbb{K}^N$, 而 $N = \infty$ 时 $\{x_i\}_{i=1}^{\infty} \in l^2$. 令 $x := \sum_{i=1}^{N} x_i e_i$, 则显见 $x \in \mathscr{X}$, 且 $Tx = \{x_i\}_{i=1}^{N}$. 当 $N < \infty$ 时, 显见 T 为满射; 当 $N = \infty$ 时, 任取 $M \in \mathbb{N}$, 记 $x^{(M)} = \sum_{i=1}^{M} x_i e_i$, 则由 Parseval 等式知

$$\|x^{(M+p)} - x^{(M)}\|^2 = \sum_{i=M+1}^{M+p} |x_i|^2 \to 0,\ M \to \infty, \forall p \in \mathbb{N}$$

这说明 $\{x^{(M)}\}_{M\in\mathbb{N}}$ 是 $\mathscr{X}$ 中的基本列, 又因为 $\mathscr{X}$ 作为 Hilbert 空间是完备的, 故 $x = \lim_{M\to\infty} x^{(M)} = \sum_{i=1}^{\infty} x_i e_i \in \mathscr{X}$,

从而 $Tx=\{x_i\}_{i=1}^{\infty}$, T 是满射.

综上, 对应 T 是 $\mathscr{X}\to\mathbb{K}^N(N<\infty)$ 或 $\mathscr{X}\to l^2(N=\infty)$ 的一对一在上线性同构, 此外因为 (在 $N=\infty$ 时由内积的连续性)

$$(x,y)=(\sum_{i=1}^{N}(x,e_i)e_i,\sum_{j=1}^{N}(y,e_j)e_j)=\sum_{i=1}^{N}(x,e_i)\overline{(y,e_j)}\ (\forall x,y\in\mathscr{X})$$

故 T 保持内积. 从而当 $N<\infty$ 时, $\mathscr{X}$ 与 $\mathbb{K}^N$ 同构; 当 $N=\infty$ 时, $\mathscr{X}$ 与 l^2 同构. □

课堂笔记 (★)

上述证明中标蓝处有致命错误: 集合 $\{x=\sum_{n=1}^{N}a_ne_n:\operatorname{Re}a_n,\operatorname{Im}a_n\in\mathbb{Q}\}$ 未必可数, 因而仅通过这个过程未必能说明 $\mathscr{X}$ 可分. 下面证明: 若 $\{e_n\}_{n=1}^{\infty}$ 为 $\mathscr{X}=l^2$ 的正交规范基, 则

$$\mathscr{A}:=\{\sum_{n=1}^{\infty}a_ne_n:a_n\in\mathbb{Q},\sum_{n=1}^{\infty}a_n^2<\infty\}$$

不可数. 注意 $\sum_{n=1}^{\infty}\frac{1}{n^2}<\infty$. 若令

$$\mathscr{A}_0:=\{\sum_{n=1}^{\infty}a_ne_n:a_n\in\{0,\frac{1}{n}\},\forall n\in\mathbb{N}\}$$

则 $\mathscr{A}_0\subset\mathscr{A}\subset\mathscr{X}$. 要证明 $\mathscr{A}$ 不可数, 只需证明 $\mathscr{A}_0$ 不可数. 考虑映射:

$$i:\prod_{n=1}^{\infty}\{0,\frac{1}{n}\}\to\mathscr{A}_0,\ \{a_n\}_{n=1}^{\infty}\mapsto\sum_{n=1}^{\infty}a_ne_n$$

其中 $\prod_{n=1}^{\infty}\{0,\frac{1}{n}\}$ 表示 $\{0,1\}\times\{0,\frac{1}{2}\}\times\cdots\times\{0,\frac{1}{n}\}\times\cdots$. 注意该无穷 Cartesian 乘积中的每个元素 $(a_1,a_2,\cdots,a_n,\cdots)$ 都可以与 $[0,1]$ 上数的二进制表示一一对应 (a_i 为 0 时映成 0, 非零时映成 1), 故 $\prod_{n=1}^{\infty}\{0,\frac{1}{n}\}$ 的势不可数. 现在只需说明 i 良定义且是双射, 其中良定义与满射显见. 对于单射, 这是因为 $\{e_n\}_{n=1}^{\infty}$ 是 $\mathscr{X}$ 的正交基, 故 $\sum_{n=1}^{\infty}a_ne_n=0\Rightarrow a_n=0(n\in\mathbb{N})$, 从而 i 是单射. 故 $\mathscr{A}_0$ 与 $\prod_{n=1}^{\infty}\{0,\frac{1}{n}\}$ 等势, 因而 $\mathscr{A}_0$ 不可数, 从而 $\mathscr{A}\supset\mathscr{A}_0$ 不可数.

为了补救证明, 在 $\mathbb{K}=\mathbb{R}$ 时另设

$$\mathscr{B}:=\{x=\sum_{n=1}^{N}a_ne_n:a_n\in\mathbb{Q}(\forall n\in\mathbb{N}),\text{且仅有有限个 }a_n\text{ 不为 }0\}$$

此时 $\mathscr{B}$ 是可数集, 下面证明 $\mathscr{B}$ 在 $\mathscr{X}$ 中稠密.

若 $N<\infty$, 则任取 $x\in\mathscr{X}$, 根据 $\mathbb{Q}$ 在 $\mathbb{R}$ 中的稠密性可知对任意 $\varepsilon>0$ 存在序列 $\{a_n\}_{n=1}^{N}\subset\mathbb{Q}$ 使得

$$|(x,e_n)-a_n|^2<\varepsilon$$

取 $\tilde{x}=\sum_{n=1}^{N}a_ne_n\in\mathscr{B}$, 因为 $\{e_n\}_{n=1}^{N}$ 已经是 $\mathscr{X}$ 的正交规范基了, 故由 Parseval 等式知:

$$\|x-\tilde{x}\|^2=\sum_{n=1}^{N}|(x,e_n)-a_n|^2<N\varepsilon$$

由 ε 的任意性即得 $\mathscr{B}$ 在 $\mathscr{X}$ 中稠密.

若 $N=\infty$, 任取 $x\in\mathscr{X},\varepsilon>0$, 因为 $\{e_n\}_{n=1}^{\infty}$ 是 $\mathscr{X}$ 的正交规范基, 故由 Parseval 等式知

$$\|x\|^2=\sum_{n=1}^{\infty}|(x,e_n)|^2<\infty$$

这说明存在 $M\in\mathbb{N}$ 充分大使得

$$\sum_{n=M+1}^{\infty}|(x,e_n)|^2<\varepsilon$$

再对 $n=1,\cdots,M$ 应用前述过程, 对任意 ε' 均可取 $a_n\in\mathbb{Q}$ 使得

$$|(x,e_n)-a_n|^2<\varepsilon'$$

令 $\widetilde{x}=\sum_{n=1}^{M}a_ne_n$, 知 $\widetilde{x}\in\mathscr{B}$, 且由 Parseval 等式知:

$$\|x-\widetilde{x}\|^2=\sum_{n=1}^{M}|(x,e_n)-a_n|^2+\sum_{n=M+1}^{\infty}|(x,e_n)|^2<M\varepsilon'+\varepsilon$$

最后由 ε,ε' 的任意性即得 $\mathscr{B}$ 在 $\mathscr{X}$ 中稠密.

而在 $\mathbb{K}=\mathbb{C}$ 时, 类似可以证明

$$\mathscr{B}:=\{x=\sum_{n=1}^{N}a_ne_n:\operatorname{Re}a_n\in\mathbb{Q},\operatorname{Im}a_n\in\mathbb{Q},\text{仅有有限个 }a_n\text{ 不为 }0\}$$

在 $\mathscr{X}$ 中稠密, 综上即知 $\mathscr{X}$ 可分.

2.6.1.4 再论最佳逼近

先前讨论的最佳逼近问题是求空间上一点到它的一个线性子空间的距离问题, 当时给定的子空间是存在无穷维的. 根据练习(2.35), 对于无穷维闭线性子空间 M, 给定 $x\in\mathscr{X}$, 未必能找到 $y\in M$ 满足

$$\inf_{z\in M}\|x-z\|=\|x-y\|$$

但在 Hilbert 空间中, 这样的 y 是存在的, 进一步还可以用更一般的闭凸子集 C 代替闭线性子空间 M. 下面从 $x=\theta$ 开始讨论.

定理 2.6.4

如果 C 是 Hilbert 空间 $\mathscr{X}$ 中的一个闭凸子集, 那么在 C 上存在唯一元素 x_0 取到最小范数.

证明

先证明存在性. 若 $\theta\in C$, 则取 $x_0=\theta$ 即可. 若 $\theta\notin C$, 知

$$d:=\inf_{z\in C}\|z\|>0$$

根据下确界的定义, 对任意的 $n\in\mathbb{N}$, 存在 $x_n\in C$, 使得

$$d\le\|x_n\|\le d+\frac{1}{n}$$

当 $\{x_n\}$ 有极限 x_0, 则根据 C 的闭性知 $x_0\in C$, 且夹逼有 $\|x_0\|=d$, 这说明 x_0 取到 C 上元素的最小范数, 故只需证明 $\{x_n\}$ 有极限, 这进一步只需说明 $\{x_n\}$ 是基本列. 考虑平行四边形等式:

$$\|x_m-x_n\|^2=2(\|x_m\|^2+\|x_n\|^2)-4\|\frac{x_m+x_n}{2}\|^2\le2((d+\frac{1}{n})^2+(d+\frac{1}{m})^2)-4d^2\to0\ (n,m\to\infty)$$

存在性得证.

再证明唯一性, 若 $x_0,\widehat{x_0}\in C$ 满足 $\|x_0\|=\|\widehat{x_0}\|=d$, 则

$$\|x_0-\widehat{x_0}\|^2=2(\|x_0\|^2+\|\widehat{x_0}\|^2)-4\|\frac{x_0+\widehat{x_0}}{2}\|^2\le4d^2-4d^2=0$$

故 $x_0=\widehat{x_0}$, 唯一性得证. □

注 上述证明中, 特别在平行四边形等式中把 $\|x_m+x_n\|^2$ 写成 $4\|\frac{x_m+x_n}{2}\|^2$ 是出于放缩需要, 根据 C 的凸性可知 $\frac{x_m+x_n}{2}\in C$, 因而 $\|\frac{x_m+x_n}{2}\|\ge d$, 代入放缩即可.

推论 2.6.2

若 C 是 Hilbert 空间 $\mathscr{X}$ 中的一个闭凸子集, 则对 $\forall y\in\mathscr{X},\exists|x_0\in C$, 使得

$$\|y-x_0\|=\inf_{x\in C}\|x-y\|.$$

证明

考察集合 $C-\{y\}:=\{x-y: x\in C\}$, 知该集合依旧是 $\mathscr{X}$ 上的闭凸子集. 根据定理2.6.4知存在唯一的 $z_0\in C-\{y\}$ 使得

$$\|z_0\|=\inf_{z\in C-\{y\}}\|z\|$$

令 $x_0:=z_0+y$ 即得欲证. □

例 2.51 若 M 是 Hilbert 空间空间 $\mathscr{X}$ 上的一个闭线性子空间, 则对 $\forall y\in\mathscr{X},\exists|x_0\in M$, 使得

$$\|y-x_0\|=\inf_{x\in M}\|x-y\|.$$

下面给出最佳逼近元的刻画.

定理 2.6.5

设 C 是内积空间 $\mathscr{X}$ 中的一个闭凸子集, $\forall y\in\mathscr{X}$, 要让 x_0 是 y 在 C 上的最佳逼近元, 必须且仅须其满足

$$\operatorname{Re}(y-x_0,x_0-x)\geq 0\ (\forall x\in C)$$

证明

对任意的 $x\in C$, 考察函数

$$\varphi_x(t)=\|y-tx-(1-t)x_0\|^2,\ t\in[0,1]$$

知 x_0 是 y 在 C 上的最佳逼近元当且仅当

$$\varphi_x(t)\geq\varphi_x(0)\ (\forall x\in C,\forall t\in[0,1])$$

下面证明

$$\forall x\in C\forall t\in[0,1](\varphi_x(t)\geq\varphi_x(0))\Leftrightarrow\forall x\in C(\operatorname{Re}(y-x_0,x_0-x)\geq 0)\tag{2.52}$$

这是因为

$$\varphi_x(t)=\|(y-x_0)+t(x_0-x)\|^2=\|y-x_0\|^2+2t\operatorname{Re}(y-x_0,x_0-x)+t^2\|x_0-x\|^2$$

故

$$\varphi_x'(0)=2\operatorname{Re}(y-x_0,x_0-x)$$

从而

$$\forall x\in C(\operatorname{Re}(y-x_0,x_0-x)\geq 0\Leftrightarrow\varphi_x'(0)\geq 0)$$

又因为

$$\varphi_x(t)-\varphi_x(0)=\varphi_x'(0)t+\|x_0-x\|^2t^2\geq 0,\ t\in[0,1]$$

故

$$\varphi_x'(0)\geq 0\Leftrightarrow\varphi_x(t)\geq\varphi_x(0)$$

因而(2.52)成立, 命题得证. □

推论 2.6.3

设 M 是 Hilbert 空间 $\mathscr{X}$ 的一个闭线性子流形. $\forall x\in\mathscr{X}$, 要使得 y 是 x 在 M 上的最佳逼近元, 必须且仅须它满足

$$x-y\perp M-\{y\}:=\{z-y:z\in M\}$$

证明

根据定理2.6.5, 要使得 y 是 x 在 M 上的最佳逼近元, 必须且仅须

$$\operatorname{Re}(x-y,y-z)\geq 0\ (\forall z\in M)\tag{2.53}$$

因为 M 是线性流形, 根据定义知任意 $z \in M$ 都可以表示为

$$z = y + w,\ \ w \in M - \{y\} \tag{2.54}$$

因为 $M - \{y\}$ 是线性子空间, 且 z 跑遍 M 时 w 跑遍 $M - \{y\}$. 将(2.54)式代入(2.53)式有:

$$\mathrm{Re}(x - y, w) \leq 0\ (\forall w \in M - \{y\})$$

用 $-w$ 代替 w 得到

$$\mathrm{Re}(x - y, w) \geq 0\ (\forall w \in M - \{y\})$$

故

$$\mathrm{Re}(x - y, w) = 0\ (\forall w \in M - \{y\})$$

再用 iw 代替 w 有

$$\mathrm{Im}(x - y, w) = 0\ (\forall w \in M - \{y\})$$

故

$$(x - y, w) = 0\ (\forall w \in M - \{y\})$$

此即欲证. □

例 2.52 当 M 是线性子空间时, 要使得 y 是 x 在 M 上的最佳逼近元, 必须且仅须 $x - y \perp M$.

证明

y 是 x 在 M 上的最佳逼近元等价于

$$\forall z \in M \forall \lambda \in \mathbb{K}(\|x - y - \lambda z\| \geq \|x - y\|)$$

两边平方有

$$\|x - y - \lambda z\|^2 = |\lambda|^2\|z\|^2 - \overline{\lambda}(x - y, z) - \lambda(z, x - y) + \|x - y\|^2 \geq \|x - y\|^2$$

亦即

$$|\lambda|^2\|z\|^2 - \overline{\lambda}(x - y, z) - \lambda(z, x - y) \geq 0,\ \forall \lambda \in \mathbb{K}$$

取 $\lambda = \frac{1}{\|z\|}(x - y, z)$ 代入即得

$$|(x - y, z)|^2\|z\|^2 \leq 0$$

从而只能有 $(x - y, z) = 0$, 命题得证. □

注 该例有形象的几何解释, 见图(2.10).

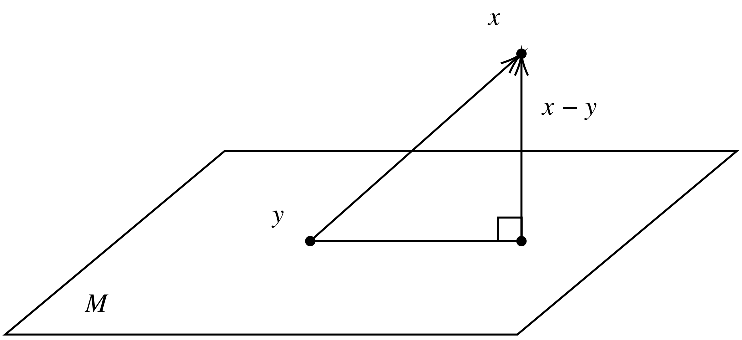


图 2.10: 最佳逼近在几何上表现为正交投影.

推论 2.6.4 (正交分解)

设 M 是 Hilbert 空间 $\mathscr{X}$ 上的一个闭线性子空间, 那么 $\forall x \in \mathscr{X}$, 存在下述唯一的正交分解:

$$x = y + z\ (y \in M, z \in M^{\perp})$$

证明

取 y 为 x 在 M 上的最佳逼近元, $z = x - y$, 知此即满足条件的一个正交分解. 若还存在另一种分解:

$$x = y' + z'\ (y' \in M, z' \in M^{\perp})$$

则

$$y - y' = z - z'$$

这说明 $y - y', z - z' \in M \cap M^{\perp}$, 故它们均为 θ, 唯一性得证. □

2.6.1.5 应用: 最小二乘法

实际问题中, 诸多情况需要考虑求解

$$y = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

其中 $y, x_1, \cdots, x_n$ 是测量数据, 但 $\lambda_1, \cdots, \lambda_n$ 未知. 理论上观测 n 次产生 n 组 $(y, x_1, \cdots, x_n)$ 即可求解 $\lambda_1, \cdots, \lambda_n$, 但观测往往会产生误差, 因而需要进行额外的观测, 导致方程个数大于未知数个数. 现在希望按下述意义确定系数: 求 $\lambda_1, \cdots, \lambda_n$, 使得

$$\min_{\alpha_1, \cdots, \alpha_n} \sum_{j=1}^{m} |y^{(j)} - \sum_{i=1}^{n} \alpha_i x_i^{(j)}|^2 = \sum_{j=1}^{m} |y^{(j)} - \sum_{i=1}^{n} \lambda_i x_i^{(i)}|^2$$

这进一步可以视作在 $\text{span}\{x_i^{(j)}\}_{1 \le i \le n, 1 \le j \le m}$ 中, 求解 $(y^{(1)}, \cdots, y^{(m)})$ 的最佳逼近.

在函数逼近论中, 对于给定的一般函数 $f \in L^2[a, b]$, 希望用给定的 n 个 $L^2[a, b]$ 函数 $\varphi_1, \cdots, \varphi_n$ 的线性组合按平方平均意义最佳逼近. 也即求解系数 $\lambda_1, \cdots, \lambda_n$ 使得

$$\min_{\alpha_1, \cdots, \alpha_n} \int_a^b |f(x) - \sum_{i=1}^{n} \alpha_i \varphi_i(x)|^2 dx = \int_a^b |f(x) - \sum_{i=1}^{n} \lambda_i \varphi_i(x)|^2 dx$$

随机数学中也有最佳估计问题. 设 $(\Omega, \mathscr{B}, \mathbb{P})$ 是一个概率空间 (即测度空间), 满足 $\mathbb{P}(\Omega) = 1$. 随机变量 X 是 $(\Omega, \mathscr{B}, \mathbb{P})$ 上的一个可测函数. 现在希望用随机变量 $X_1(\omega), \cdots, X_n(\omega)$ 的线性组合来估计 $X(\omega)$. 设 $X, X_1, \cdots, X_n \in L^2(\Omega, \mathscr{B}, \mathbb{P})$, 求解系数 $\lambda_1, \cdots, \lambda_n$, 使得

$$\min_{(\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n} \int_{\Omega} |X(\omega) - \sum_{i=1}^{n} \alpha_i X_i(\omega)|^2 d\mathbb{P}(\omega) = \int_{\Omega} |X(\omega) - \sum_{i=1}^{n} \lambda_i X_i(\omega)|^2 d\mathbb{P}(\omega)$$

这三个问题的本质均为: 在 Hilbert 空间 $\mathscr{X}$ 上给定 x 及 $x_1, x_2, \cdots, x_n$, 希望求出 $(\lambda_1, \cdots, \lambda_n) \in \mathbb{R}^n$, 使得

$$\min_{(\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n} \|x - \sum_{i=1}^{n} \alpha_i x_i\| = \|x - \sum_{i=1}^{n} \lambda_i x_i\|$$

从几何上看这件事, 就是求 x 在由 $x_1, \cdots, x_n$ 张成的子空间 M 上的正交投影. 不妨设 $\{x_1, \cdots, x_n\}$ 是线性无关的, 根据例(2.52), 要使得 $x_0 = \sum_{i=1}^{n} \lambda_i x_i$ 是欲求的解, 必须且仅须

$$(x - x_0, x_j) = 0\ (j = 1, 2, \cdots, n)$$

代入 $x_0 = \sum_{i=1}^{n} \lambda_i x_i$ 即

$$\sum_{i=1} \lambda_i (x_i, x_j) = (x, x_j)\ (j = 1, 2, \cdots, n) \tag{2.55}$$

因为解 $\{\lambda_i\}_{i=1}^n$ 存在唯一, 故线性方程组(2.55)的系数行列式非零, 因而

$$\lambda_i = \frac{\begin{vmatrix} (x_1,x_1) & \cdots & (x,x_1) & \cdots & (x_n,x_1) \\ (x_1,x_2) & \cdots & (x,x_2) & \cdots & (x_n,x_2) \\ \vdots & & \vdots & & \vdots \\ (x_1,x_n) & \cdots & (x,x_n) & \cdots & (x_n,x_n) \end{vmatrix}}{\begin{vmatrix} (x_1,x_1) & \cdots & (x_i,x_1) & \cdots & (x_n,x_1) \\ (x_1,x_2) & \cdots & (x_i,x_2) & \cdots & (x_n,x_2) \\ \vdots & & \vdots & & \vdots \\ (x_1,x_n) & \cdots & (x_i,x_n) & \cdots & (x_n,x_n) \end{vmatrix}} \ (i = 1, 2, \cdots, n)$$

2.6.2 一些例子

例 2.53(存在某 Banach 空间的非空闭凸子集不含最小范数元)

2.6.3 习题

✍ **练习 2.45(极化恒等式)** 设 a 是复线性空间 $\mathscr{X}$ 上的共轭双线性函数, q 是由 a 诱导的二次型, 求证: 对 $\forall x, y \in \mathscr{X}$ 有

$$a(x,y) = \frac{1}{4}[q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)].$$

证明

知

$$\begin{aligned}
&\frac{1}{4}[q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy)] \\
=&\frac{1}{4}(a(x+y,x+y) - a(x-y,x-y) + ia(x+iy,x+iy) - ia(x-iy,x-iy)) \\
=&\frac{1}{4}(a(x,x) + a(y,x) + a(x,y) + a(y,y) - a(x,x) + a(y,x) + a(x,y) - a(y,y) \\
&+ ia(x,x) - a(y,x) + a(x,y) + ia(y,y) - ia(x,x) - a(y,x) + a(x,y) - ia(y,y)) \\
=&a(x,y)
\end{aligned}$$

✍ **练习 2.46★♡** 求证: 在 $C[a,b]$ 中不可能引进一种内积 $(\cdot,\cdot)$, 使其满足

$$(f,f)^{\frac{1}{2}} = \max_{a\le x\le b} |f(x)| \ (\forall f \in C[a,b])$$

证明

用反证法, 如果存在这样的内积满足题式, 根据平行四边形等式应有

$$\|f+g\|^2 + \|f-g\|^2 = 2(\|f\|^2 + \|g\|^2) \ (\forall f, g \in C[a,b])$$

现取 $f(x) = 1, g(x) = \frac{x-a}{b-a}$, 知 $\|f\| = \|g\| = 1$, 易知 $\|f+g\| = 2, \|f-g\| = 1$, 故

$$\|f+g\|^2 + \|f-g\|^2 = 5 \neq 4 = 2(\|f\|^2 + \|g\|^2)$$

矛盾!

✍ **练习 2.47♡** 在 $L^2[0,T]$ 中, 求证: 函数

$$x \mapsto \left|\int_0^T e^{-(T-\tau)} x(\tau) d\tau\right| \ (\forall x \in L^2[0,T])$$

在单位球面上达到最大值, 并求出此最大值和达到最大值的元素 x.

证明

注意在单位球面上 $\int_0^T |x(\tau)|^2 d\tau = 1$, 因而由 Cauchy-Schwarz 不等式有

$$|\int_0^T e^{-(T-\tau)} x(\tau) d\tau| \le \int_0^T e^{-(T-\tau)} |x(\tau)| d\tau \le (\int_0^T e^{-2(T-\tau)} d\tau)^{\frac{1}{2}} \cdot (\int_0^T |x(\tau)|^2 d\tau)^{\frac{1}{2}} = (\int_0^T e^{-2(T-\tau)} d\tau)^{\frac{1}{2}} = \frac{1}{\sqrt{2}}(1-e^{-2T})^{\frac{1}{2}}$$

等号成立当且仅当 $x(\tau)$ 与 $e^{-(T-\tau)}$ 线性相关, 且 $x(\tau)$ 不变号. 设 $x(\tau) = ke^{-(T-\tau)}$, 要求

$$\|x\| = (\int_0^T k^2 e^{-2(T-\tau)} d\tau)^{\frac{1}{2}} = 1$$

解得

$$k = \pm \frac{\sqrt{2}}{(1-e^{-2T})^{\frac{1}{2}}}$$

故

$$x(t) = \pm \frac{\sqrt{2} e^{t-T}}{(1-e^{-2T})^{\frac{1}{2}}}, \ t \in [0, T]$$

练习 2.48★♡ 设 M, N 是内积空间中的两个子集, 求证:

$$M \subset N \Rightarrow N^{\perp} \subset M^{\perp}$$

证明

任取 $x \in N^{\perp}$, 根据定义知

$$\forall n \in N((n, x) = 0)$$

又因为 $M \subset N$, 知

$$\forall n \in M((n, x) = 0)$$

这说明 $n \in M^{\perp}$, 即 $N^{\perp} \subset M^{\perp}$.

练习 2.49★ 设 M 是 Hilbert 空间 $\mathscr{X}$ 的子集, 求证:

$$(M^{\perp})^{\perp} = \overline{\operatorname{span} M}$$

证明

当 $x \in M^{\perp}$, 根据内积的双线性性知 $x \in (\operatorname{span} M)^{\perp}$, 再根据内积的连续性知 $x \in (\overline{\operatorname{span} M})^{\perp}$, 故 $M^{\perp} \subset (\overline{\operatorname{span} M})^{\perp}$, 又根据练习(2.49)与 $M \subset \overline{\operatorname{span} M}$ 知 $(\overline{\operatorname{span} M})^{\perp} \subset M^{\perp}$, 故 $(\overline{\operatorname{span} M})^{\perp} = M^{\perp}$, 因而题即证明

$$((\overline{\operatorname{span} M})^{\perp})^{\perp} = \overline{\operatorname{span} M}$$

现在说明对任意的闭线性子空间 M, 总有 $(M^{\perp})^{\perp} = M$. 显见 $M \subset (M^{\perp})^{\perp}$, 若存在 $x \neq \theta$ 满足 $x \in (M^{\perp})^{\perp} \backslash M$, 根据正交分解定理, 设 $x = y + z, y \in M, z \in M^{\perp}$, 因为 $x \perp M^{\perp}$, 故 $z = \theta$, 这说明 $x = y \in M$, 矛盾! 故 $(M^{\perp})^{\perp} = M$, 把 M 换成 $\overline{\operatorname{span} M}$ 即得欲证.

练习 2.50 在 $L^2[-1, 1]$ 中, 问偶函数集的正交补是什么? 证明你的结论.

解

(在几乎处处意义下的) 奇函数集. 设 $L^2[-1, 1]$ 中偶函数的全体为 E, (几乎处处) 奇函数[22]的全体为 O, 另设 $f \in E, g \in O$, 知

$$(f, g) = \int_{-1}^{1} f(x)\overline{g(x)} dx = 0$$

[22]即满足 $f(x) + f(-x) \neq 0$ 的 $x \in [0, 1]$ 构成零测集.

故 $g \in E^{\perp}$, 这说明 $O \subset E^{\perp}$. 现考虑 $g \notin O$, 根据定义知存在正测集 $U \subset [0,1]$ 使得 $g(x)+g(-x) \neq 0, x \in U$, 得

$$(g,f) = \int_{-1}^{1} g(x)\overline{f(x)}dx = \int_{-U} g(x)\overline{f(x)}dx + \int_{U} g(x)\overline{f(x)}dx = \int_{U} (g(x)+g(-x))\overline{f(x)}dx$$

取 $\overline{f(x)} = \chi_V(x)$, 其中 V 是任意关于原点对称的区域, 知

$$\int_{U} (g(x)+g(-x))\overline{f(x)}dx = 0 \Rightarrow m(\{x \in U : g(x)+g(-x) \neq 0\}) = 0$$

其中 m 是 Lebesgue 测度, 这与假设矛盾! 故 $E^{\perp} \subset O$, 即 $O = E^{\perp}$.

练习 2.51 在 $L^2[a,b]$ 中, 考察函数集 $S = \{e^{2\pi inx}\}$

(1) 若 $|b-a| \leq 1$, 求证: $S^{\perp} = \{\theta\}$;

证明

当 $|b-a| = 1$ 时, 根据 $e^{2\pi inx}$ 的周期性, 只需考虑 $L^2[0,1]$. 根据 Fourier 级数的均方收敛性, 对任意的 $f \in L^2[0,1]$, 总存在 Fourier 级数 $\sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}$ 使得

$$\int_0^1 |f(x) - \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx}|^2 dx = 0$$

若 $f \in S^{\perp}$, 则

$$\forall n \in \mathbb{Z}(\int_0^1 f(x)e^{2\pi inx}dx = 0) \Rightarrow \forall n \in \mathbb{Z}(c_n = 0)$$

故

$$\int_0^1 |f(x)|^2 dx = 0 \Rightarrow f = \theta$$

即 $S^{\perp} = \{\theta\}$.

当 $|b-a| < 1$, 根据周期性只需考虑 $L^2[0,b](0 < b < 1)$. 对任意的 $f \in L^2[0,b]$, 设

$$\hat{f}(x) = \begin{cases} f, & x \in [0,b] \\ 0, & x \in (b,1] \end{cases}$$

则当 $f \in S^{\perp}$, 知对任意的 $n \in \mathbb{Z}$ 有

$$\int_0^b f(x)e^{2\pi inx}dx = 0 \Rightarrow \int_0^1 \hat{f}(x)e^{2\pi inx}dx = 0 \Rightarrow \hat{f} = \theta \Rightarrow f = \theta$$

故 $S^{\perp} = \{\theta\}$.

•[23](2) 若 $|b-a| > 1$, 求证: $S^{\perp} \neq \{\theta\}$.

证明

根据周期性只需考虑 $L^2[0,b](b > 1)$. 当 $f \in S^{\perp}$, 知对任意的 $n \in \mathbb{Z}$ 有

$$\int_0^b f(x)e^{2\pi inx}dx = \int_0^1 f(x)e^{2\pi inx}dx + \int_1^b f(x)e^{2\pi inx}dx = 0$$

只需设计 f 使其满足

$$\int_0^1 f(x)e^{2\pi inx}dx = -\int_1^b f(x)e^{2\pi inx}dx,\ \forall n \in \mathbb{N}$$

即可, 当 $b \notin \mathbb{Q}$, 这样的一个例子为

$$f(x) = \begin{cases} x, & x \in [0,1] \\ g(x), & x \in (1,b] \end{cases}$$

[23]这题最后不知道怎么设计了.

练习 2.52 设 $\mathscr{X}$ 表示闭单位圆上的解析函数全体, 内积定义为

$$(f,g)=\frac{1}{i}\oint_{|z|=1}\frac{f(z)\overline{g(z)}}{z}dz\ (\forall f,g\in\mathscr{X})$$

求证: $\{\frac{z^n}{\sqrt{2\pi}}\}$ 是一组正交规范集.

证明

任取 $n>m\Rightarrow n-m\geq 1$, 设 $f(z)=\frac{z^n}{\sqrt{2\pi}},g(z)=\frac{z^m}{\sqrt{2\pi}}$, 有

$$(f,g)=\frac{1}{i}\oint_{|z|=1}\frac{|z|^m z^{n-m}}{2\pi z}dz=0$$

$n<m$ 的情况同理. 当 $n=m$, 知

$$(f,f)=\frac{1}{i}\oint_{|z|=1}\frac{|z|^{2n}}{2\pi z}dz=\frac{1}{i}\cdot\frac{1}{2\pi}\cdot 2\pi i=1$$

故 $\{\frac{z^n}{\sqrt{2\pi}}\}$ 是一组正交规范集.

练习 2.53★♡ 设 $\{e_n\}_{n=1}^{\infty},\{f_n\}_{n=1}^{\infty}$ 是 Hilbert 空间 $\mathscr{X}$ 中的两个正交规范集, 满足条件

$$\sum_{n=1}^{\infty}\|e_n-f_n\|^2<1$$

求证: $\{e_n\}$ 和 $\{f_n\}$ 两者中一个完备蕴含另一个完备.

证明

当 $\{e_n\}$ 完备, 知任取 $x\in\mathscr{X}$, 总有

$$\forall n\in\mathbb{N}((x,e_n)=0)\Rightarrow x=\theta$$

如果 $\{f_n\}_{n=1}^{\infty}$ 不完备, 则存在 $\theta\neq x\in\mathscr{X}$ 使得 $(x,f_n)\neq 0,\forall n\in\mathbb{N}$. 进而

$$\|x\|^2=\sum_{n=1}^{\infty}|(x,e_n)|^2=\sum_{n=1}^{\infty}|(x,e_n-f_n)|^2\leq\sum_{n=1}^{\infty}\|x\|^2\cdot\|e_n-f_n\|^2<\|x\|^2$$

矛盾! 命题即证.

注 ♡ 该题的条件可以弱化为 $\sum_{n=1}^{\infty}\|e_n-f_n\|^2<\infty$, 下面证明该条件下的结论.

证明

设 $\{e_n\}$ 完备, 下面说明 $\{f_n\}$ 完备, 根据完备的定义只需说明对任意的 $f_0\in\mathscr{X}$, 只要 $\forall n\in\mathbb{N}(f_0\perp f_n)$, 则 $f_0=\theta$ 即可. 用反证法, 若 $f_0\neq\theta$, 则因为 $\sum_{n=1}^{\infty}\|e_n-f_n\|^2<\infty$, 知必存在 $N\in\mathbb{N}$ 使得 $\sum_{n=N+1}^{\infty}\|e_n-f_n\|^2<1$. 现令 $g_k=\sum_{j=1}^{N}(f_k,e_j)e_j(k=0,1,\cdots,N)$, 注意 $\{g_0,g_1,\cdots,g_N\}$ 的基为 $\{e_1,\cdots,e_N\}$, 故 $\dim\{g_0,g_1,\cdots,g_N\}\leq N$, 从而 $\{g_0,g_1,\cdots,g_N\}$ 线性相关, 因而存在不全为 0 的 $a_0,a_1,\cdots,a_N$ 使得 $\sum_{k=0}^{N}a_kg_k=0$. 代入 $g_k(k=0,1,\cdots,N)$ 的定义有

$$\sum_{k=0}^{N}a_k(\sum_{j=1}^{N}(f_k,e_j))e_j=0\Rightarrow\sum_{j=1}^{N}(\sum_{k=0}^{N}a_kf_k,e_j)e_j=0$$

不妨设 $h=\sum_{k=0}^{N}a_kf_k$, 得到 $\sum_{j=1}^{N}(h,e_j)e_j=0$, 且根据 $\{f_k\}$ 的正交性知 $\forall k\geq n+1(h\perp f_k)$, 现在只要得到 $h=0$ 即得矛盾. 由 $\{e_1,\cdots,e_N\}$ 的正交性知它们线性无关, 因而由 $\sum_{j=1}^{N}(h,e_j)e_j=0$ 知 $(h,e_j)=0(j=1,\cdots,n)$.

因为 $\{e_k\}$ 完备, 故由 Parseval 等式知

$$\|h\|^2=\sum_{j=1}^{\infty}|(h,e_j)|^2=\sum_{j=N+1}^{\infty}|(h,e_j)|^2=\sum_{j=N+1}^{\infty}|(h,e_j)-(h,f_j)|^2=\sum_{j=N+1}^{\infty}|(h,e_j-f_j)|^2$$

进一步由 Cauchy-Schwartz 不等式知

$$\|h\|^2\leq\|h\|^2\sum_{j=N+1}^{\infty}\|e_j-f_j\|^2$$

又因为 $\sum_{j=N+1}^{\infty}\|e_j-f_j\|^2<1$, 故 $\|h\|=0$, 因而 $h=\sum_{k=0}^{N}a_kf_k=0$. 又因为 $\{f_0,f_1,\cdots,f_N\}$ 正交, 故它们线性无关, 因而 $a_0=a_1=\cdots=a_N=0$, 矛盾! 命题即证. □

✍ **练习 2.54★** 设 $\mathscr{X}$ 是 Hilbert 空间, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的闭线性子空间, $\{e_n\},\{f_n\}$ 分别是 $\mathscr{X}_0$ 和 $\mathscr{X}_0^{\perp}$ 的正交规范基. 求证: $\{e_n\}\cup\{f_n\}$ 是 $\mathscr{X}$ 的正交规范基.

证明

根据 $\{e_n\}\subset\mathscr{X}_0,\{f_n\}\subset\mathscr{X}_0^{\perp}$ 知 $(e_n,f_m)=0,\forall n,m$, 这说明 $\{e_n\}\cup\{f_n\}$ 是 $\mathscr{X}$ 的正交规范集, 再证明其为基底. 因为 $\mathscr{X}_0$ 是闭线性子空间, 故根据正交分解定理知任取 $x\in\mathscr{X}$, 有表示 $x=y+z,y\in\mathscr{X}_0,z\in\mathscr{X}_0^{\perp}$, 又因为 $\{e_n\},\{f_n\}$ 分别是 $\mathscr{X}_0$ 和 $\mathscr{X}_0^{\perp}$ 的正交规范基, 知 $y=\sum_n y_ne_n, z=\sum_m z_mf_m$, 因而 $x=\sum_n y_ne_n+\sum_m z_mf_m$, 故 $\{e_n\}\cup\{f_n\}$ 是 $\mathscr{X}$ 的基底, 命题即证.

✍ **练习 2.55** 设 $H^2(D)$ 是按例(2.50)定义的内积空间.

(1) 如果 $u(z)$ 的 Taylor 展开式是 $u(z)=\sum_{k=0}^{\infty}b_kz^k$, 求证:

$$\sum_{k=0}^{\infty}\frac{|b_k|^2}{1+k}<\infty.$$

证明

根据 Bessel 不等式, 令 $\varphi_n(z)=\sqrt{\frac{n}{\pi}}z^{n-1}$, 有

$$\sum_{n=1}^{\infty}|(u,\varphi_n)|^2=\sum_{n=1}^{\infty}\frac{\pi b_{n-1}^2}{n}\le\|u\|^2<\infty$$

得到

$$\sum_{k=0}^{\infty}\frac{|b_k|^2}{1+k}<\infty$$

(2) 设 $u(z),v(z)\in H^2(D)$, 并且

$$u(z)=\sum_{k=0}^{\infty}a_kz^k,\ \ v(z)=\sum_{k=0}^{\infty}b_kz^k$$

求证:

$$(u,v)=\pi\sum_{k=0}^{\infty}\frac{a_k\overline{b_k}}{k+1}$$

证明

考虑

$$u(z)=\sum_{k=1}^{\infty}a_{k-1}\sqrt{\frac{\pi}{k}}\varphi_k(z),\ \ v(z)=\sum_{k=1}^{\infty}b_{k-1}\sqrt{\frac{\pi}{k}}\varphi_k(z)$$

得到

$$(u,v)=(\sum_{k=1}^{\infty}a_{k-1}\sqrt{\frac{\pi}{k}}\varphi_k(z),\sum_{k=1}^{\infty}b_{k-1}\sqrt{\frac{\pi}{k}}\varphi_k(z))=\sum_{p=1}^{\infty}\sum_{q=1}^{\infty}\frac{\pi a_{p-1}\overline{b_{q-1}}}{\sqrt{pq}}(\varphi_p,\varphi_q)=\pi\sum_{k=0}^{\infty}\frac{a_k\overline{b_k}}{k+1}$$

•(3) 设 $u(z)\in H^2(D)$, 求证:

$$|u(z)|\le\frac{\|u\|}{\sqrt{\pi}(1-|z|)}\ (\forall|z|<1)$$

证明

既然 $u \in H^2(D)$, 可设

$$u(z)=\sum_{k=0}^{\infty}a_kz^k=\sum_{k=0}^{\infty}\frac{a_k}{\sqrt{1+k}}\sqrt{1+k}z^k$$

由 Cauchy-Schwartz 不等式知

$$|u(z)|\leq\sum$$

•[24](4) 验证 $H^2(D)$ 是 Hilbert 空间.

✍ **练习 2.56★♡** 设 $\mathscr{X}$ 是内积空间, $\{e_n\}$ 是 $\mathscr{X}$ 中的正交规范集, 求证:

$$|\sum_{n=1}^{\infty}(x,e_n)\overline{(y,e_n)}|\leq\|x\|\cdot\|y\|\ (\forall x,y\in\mathscr{X}).$$

证明

根据 Cauchy-Schwarz 不等式有

$$|\sum_{n=1}^{\infty}(x,e_n)\overline{(y,e_n)}|\leq(\sum_{n=1}^{\infty}|(x,e_n)^2|)^{\frac{1}{2}}(\sum_{n=1}^{\infty}|(y,e_n)^2|)^{\frac{1}{2}}$$

根据 Bessel 不等式有

$$(\sum_{n=1}^{\infty}|(x,e_n)^2|)^{\frac{1}{2}}\leq\|x\|,\ (\sum_{n=1}^{\infty}|(y,e_n)^2|)^{\frac{1}{2}}\leq\|y\|$$

两式联合即为欲证.

✍ **练习 2.57★** 设 $\mathscr{X}$ 是一个内积空间, $\forall x_0\in\mathscr{X},\forall r>0$, 令

$$C=\{x\in\mathscr{X}:\|x-x_0\|\leq r\}.$$

(1) 求证: C 是 $\mathscr{X}$ 中的闭凸集.

证明

要验证 C 是 $\mathscr{X}$ 中的闭凸集, 只需验证 $C-\{x_0\}:=\{x-x_0:\|x-x_0\|\leq r\}$ 是 $\mathscr{X}$ 中的闭凸集即可, 这是因为闭性和凸性不随平移而改变.

对于 $D:=C-\{x_0\}$ 的闭性, 任取 $y\in D'$, 根据定义知存在 $\{y_n\}\subset D$ 使得

$$\forall\varepsilon>0\exists N\in\mathbb{N}\forall n>N(\|y-y_n\|<\varepsilon)$$

注意

$$\|y-y_n\|\geq\|y-x_0\|-\|y_n-x_0\|>\|y-x_0\|-r$$

其中最后的不等号是 $y_n\in D$ 的定义. 故

$$\|y-x_0\|<\varepsilon+r,\ \forall\varepsilon>0$$

因而 $\|y-x_0\|\leq r$, 亦即 $y\in D$, 故 D 是闭集.

对于 D 的凸性, 任取 $x,y\in D,\lambda\in[0,1]$, 知

$$\begin{aligned}\|\lambda x+(1-\lambda)y-x_0\|&=\|\lambda(x-x_0)+(1-\lambda)(y-x_0)\|\\&\leq\lambda\|x-x_0\|+(1-\lambda)\|y-x-0\|\leq r\end{aligned}$$

故 $\lambda x+(1-\lambda)y\in D$, 从而 D 是凸集. 综上即得欲证.

[24]不清楚一般验证完备空间的方法.

(2) $\forall x \in \mathscr{X}$, 令

$$y = \begin{cases} x_0 + \frac{r(x-x_0)}{\|x-x_0\|}, & x \notin C \\ x, & x \in C \end{cases}$$

求证: y 是 x 在 C 中的最佳逼近元.

证明

根据 y 的构造知 x_0, y, x 共线. 任取 $z \in C$, 知

$$\|z - x\| \geq \|x - x_0\| - \|x_0 - z\| \geq \|x - x_0\| - r$$

而

$$\|x - x_0\| = \|x - y\| + \|y - x_0\| = \|x - y\| + r$$

故

$$\|z - x\| \geq \|y - x\|$$

此即欲证.

练习 2.58 求 $(a_0, a_1, a_2) \in \mathbb{R}^3$, 使得 $\int_0^1 |e^t - a_0 - a_1 t - a_2 t^2|^2 dt$ 取最小值.

解

考虑

$$(e^t, 1) = e - 1, (e^t, t) = 1, (e^t, t^2) = e - 2,$$
$$(1, 1) = 1, (1, t) = \frac{1}{2}, (1, t^2) = \frac{1}{3}$$
$$(t, t) = \frac{1}{3}, (t, t^2) = \frac{1}{4}, (t^2, t^2) = \frac{1}{5}$$

得到

$$a_0 = \frac{\begin{vmatrix} (e^t, 1) & (t, 1) & (t^2, 1) \\ (e^t, t) & (t, t) & (t^2, t) \\ (e^t, t^2) & (t, t^2) & (t^2, t^2) \end{vmatrix}}{\begin{vmatrix} (1, 1) & (t, 1) & (t^2, 1) \\ (1, t) & (t, t) & (t^2, t) \\ (1, t^2) & (t, t^2) & (t^2, t^2) \end{vmatrix}} = 3(13e - 35)$$

$$a_1 = \frac{\begin{vmatrix} (1, 1) & (e^t, 1) & (t^2, 1) \\ (1, t) & (e^t, t) & (t^2, t) \\ (1, t^2) & (e^t, t^2) & (t^2, t^2) \end{vmatrix}}{\begin{vmatrix} (1, 1) & (t, 1) & (t^2, 1) \\ (1, t) & (t, t) & (t^2, t) \\ (1, t^2) & (t, t^2) & (t^2, t^2) \end{vmatrix}} = 12(49 - 18e)$$

$$a_2 = \frac{\begin{vmatrix} (1, 1) & (t, 1) & (e^t, 1) \\ (1, t) & (t, t) & (e^t, t) \\ (1, t^2) & (t, t^2) & (e^t, t^2) \end{vmatrix}}{\begin{vmatrix} (1, 1) & (t, 1) & (t^2, 1) \\ (1, t) & (t, t) & (t^2, t) \\ (1, t^2) & (t, t^2) & (t^2, t^2) \end{vmatrix}} = 30(7e - 19)$$

故欲求即 $(3(13e - 35), 12(49 - 18e), 30(7e - 19))$.

•

✍ **练习 2.59** 设 $f(x) \in C^2[a,b]$, 满足边界条件

$$f(a) = f(b) = 0, f'(a) = 1, f'(b) = 0$$

求证:

$$\int_a^b |f''(x)|^2 dx \geq \frac{4}{b-a}$$

证明

•

✍ **练习 2.60(变分不等式)** 设 $\mathscr{X}$ 是一个 Hilbert 空间, $a(x,y)$ 是 $\mathscr{X}$ 上的共轭对称的双线性函数, $\exists M > 0, \delta > 0$, 使得

$$\delta\|x\|^2 \leq a(x,x) \leq M\|x\|^2 \ (\forall x \in \mathscr{X})$$

又设 $u_0 \in \mathscr{X}$, C 是 $\mathscr{X}$ 上的一个闭凸子集. 求证: 函数

$$x \mapsto a(x,x) - \mathrm{Re}(u_0, x)$$

在 C 上达到最小值, 并且达到最小值的点 x_0 唯一, 且满足

$$\mathrm{Re}[2a(x_0, x-x_0) - (u_0, x-x_0)] \geq 0 \ (\forall x \in C).$$

证明

2.7 泛函分析 (高志强老师班) 第一章作业与解答合集

2.7.1 第一次作业

1. 设 (X,d) 是距离空间, 令 $\rho(x,y) = \frac{d(x,y)}{1+d(x,y)}$. 求证 (X,ρ) 也是距离空间.

证明

只需证明 ρ 也是距离即可.

对 ρ 的正定性, 由 d 是距离知 $d \geq 0$, 故 $\rho(x,y) = \frac{d(x,y)}{1+d(x,y)} \geq 0$. 且 $\rho(x,y) = 0 \Leftrightarrow d(x,y) = 0 \Leftrightarrow x = y$. 正定性得证.

对 ρ 的对称性, 由 d 是距离知 $d(x,y) = d(y,x)$, 故

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

对称性得证.

对三角不等式, 此即证明

$$\rho(x,y) + \rho(y,z) \geq \rho(x,z), \ \forall x,y,z \in X$$

代入定义即证

$$\begin{aligned}
&\frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)} \geq \frac{d(x,z)}{1+d(x,z)} \\
\Leftarrow &\frac{d(x,y)(1+d(y,z)) + d(y,z)(1+d(x,y))}{(1+d(x,y))(1+d(y,z))} \geq \frac{d(x,z)}{1+d(x,z)} \\
\Leftarrow &\frac{d(x,y)+d(y,z)+2d(x,y)d(y,z)}{1+d(x,y)+d(y,z)+d(x,y)d(y,z)} \geq \frac{d(x,z)}{1+d(x,z)} \\
\Leftarrow &d(x,y)+d(y,z)+2d(x,y)d(y,z)+d(x,y)d(x,z)+d(y,z)d(x,z)+2d(x,y)d(y,z)d(x,z) \\
&\geq d(x,z)+d(x,y)d(x,z)+d(y,z)d(x,z)+d(x,y)d(y,z)d(x,z) \\
\Leftarrow &d(x,y)+d(y,z)+2d(x,y)d(y,z)+d(x,y)d(y,z)d(x,z) \geq 0
\end{aligned}$$

最后的不等式由 d 的正定性显然成立, 故三角不等式得证.

综上, ρ 是 X 上的距离, 因而 (X,ρ) 是距离空间. □

2. 求证 [0,1] 上的全体多项式, 定义距离

$$d(p,q)=\int_0^1 |p(x)-q(x)|dx\ (p,q\text{ 是多项式})$$

形成的距离空间是不完备的, 并指出它的完备化空间.

证明

首先简单说明距离 d 确为距离函数. d 的正定性与对称性由绝对值的性质立得, 而 d 满足三角不等式是因为绝对值函数 $|\cdot|$ 满足三角不等式.

再说明题设距离空间不完备. 考虑多项式列 $\{p_n\}_{n=1}^{\infty}$, 其定义为

$$p_n(x)=\sum_{k=1}^{n}\frac{x^k}{k!}$$

下面说明 $\{p_n\}$ 是题设空间在距离 d 下的基本列, 设 $m>n$, 则

$$d(p_m,p_n)=\int_0^1 |\sum_{k=n+1}^{m}\frac{x^k}{k!}|dx=\sum_{k=n+1}^{m}\int_0^1\frac{x^k}{k!}dx=\sum_{k=n+1}^{m}\frac{1}{(k+1)!} \tag{2.56}$$

容易证明当 $n\in\mathbb{N}$ 时 $(n+1)!>(n+1)n>n^2$, 故由比较审敛法知级数 $\sum_{n=1}^{\infty}\frac{1}{(n+1)!}$ 收敛, 从而由 Cauchy 收敛准则知

$$\forall\varepsilon>0\exists N>0\forall m>n>N(|\sum_{k=n+1}^{m}\frac{1}{(n+1)!}|<\varepsilon) \tag{2.57}$$

将(2.56)式代入(2.57)的不等式中得到

$$\forall\varepsilon>0\exists N>0\forall m>n>N(d(p_m,p_n)<\varepsilon)$$

故 $\{p_n\}$ 是题设空间在距离 d 下的基本列, 但极限 $\lim_{n\to\infty}p_n(x)=e^x$ 在 $[0,1]$ 上由 Dini 定理是一致成立的, 而函数 e^x 并非多项式, 这便说明题设空间在距离 d 下不完备.

由 Weierstrass 逼近定理知, [0,1] 上的连续函数总能由 [0,1] 上的多项式一致逼近, 这说明题设距离空间在 $(C[0,1],d)$ 中稠密, 又根据实变函数论的相关定理知: [0,1] 上的任意可积函数可由 $C[0,1]$ 中的函数按 L^1 距离逼近, 这说明 $(C[0,1],d)$ 在 $L^1[0,1]$ 中稠密. 最后来证明 $L^1[0,1]$ 的完备性:

设 $\{f_n\}$ 是 L^1 中的一个基本列, 即

$$\forall\varepsilon>0\exists N>0\forall m>n>N(\int_0^1|f_m-f_n|dx<\varepsilon) \tag{2.58}$$

现取 $\varepsilon_k=\frac{1}{2^k}(k\in\mathbb{N})$, 则存在子列 $\{f_{n_k}\}$ 满足

$$\int_0^1|f_{n_{k+1}}-f_{n_k}|dx<\frac{1}{2^k} \tag{2.59}$$

现说明 $\lim_{k\to\infty}f_{n_k}\in L^1[0,1]$, 考虑级数

$$F(x):=|f_{n_1}(x)|+\sum_{k=1}^{\infty}|f_{n_{k+1}}(x)-f_{n_k}(x)|$$

根据逐项积分定理与(2.59)式, 因为

$$\sum_{k=1}^{\infty}\int_0^1|f_{n_{k+1}}-f_{n_k}|dx\le\sum_{k=1}^{\infty}\frac{1}{2^k}=1<\infty$$

故

$$\int_0^1 F dx = \int_0^1 |f_{n_1}|dx + \int_0^1 \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|dx$$
$$= \int_0^1 |f_{n_1}|dx + \sum_{k=1}^{\infty} \int_0^1 |f_{n_{k+1}} - f_{n_k}|dx < \int_0^1 |f_{n_1}|dx + 1 < \infty$$

其中最后一步是因为 $f_{n_1} \in L^1[0,1]$. 这个可积性一方面说明级数 $F(x)$ 在 $[0,1]$ 上几乎处处收敛, 另一方面通过设

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$$

可由 $|f| \le F$ 得到 $f \in L^1[0,1]$, 从而级数 $f(x)$ 同样在 [0,1] 上几乎处处收敛. 从定义中可见

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

同时对任意固定的 k, 显见 $|f - f_{n_k}| \le F$, 故由 Lebesgue 控制收敛定理知

$$\lim_{k\to\infty} \int_0^1 |f - f_{n_k}|dx = \int_0^1 |f - \lim_{k\to\infty} f_{n_k}|dx = 0$$

这说明 f_{n_k} 在 L^1 距离 (也即距离 d) 的意义下收敛到 f. 沿用(2.58)式的记号并将上式按定义写开即

$$\exists K > 0 \forall k > K (\int_0^1 |f - f_{n_k}|dx < \varepsilon) \tag{2.60}$$

最后, 考虑说明 f_n 在 L^1 距离的意义下收敛到 f. 沿用(2.58),(2.60)式的记号, 取 $n > N, k > K$ 满足 $n_k > N$, 则

$$\int_0^1 |f - f_n|dx \le \int_0^1 |f - f_{n_k}|dx + \int_0^1 |f_{n_k} - f_n|dx < 2\varepsilon$$

由 ε 的任意性即得 f_n 在 L^1 距离的意义下收敛到 $f \in L^1[0,1]$.

综上, $L^1[0,1]$ 是完备空间, 且根据 [0,1] 上的多项式函数在 $L^1[0,1]$ 中的稠密性知 $L^1[0,1]$ 正是题设空间的完备化空间. □

2.7.2 第二次作业

1. 证明存在闭区间 [0,1] 上的连续函数 $x(t)$, 使得

$$x(t) = \frac{1}{2}\sin x(t) - a(t)$$

其中 $a(t)$ 是给定的 [0,1] 上的连续函数.

证明

设 $Tx(t) = \frac{1}{2}\sin x(t) - a(t)$, 只需证明 $T : C[0,1] \to C[0,1]$ 是压缩映射即可. 任取 $x_1(t), x_2(t) \in C[0,1]$, 设 ρ 为 C 上的一致收敛性度量, 有:

$$\rho(Tx_1, Tx_2) = \max_{t\in[0,1]} |\frac{1}{2}\sin x_1(t) - \frac{1}{2}\sin x_2(t)| = \frac{1}{2}\max_{t\in[0,1]} |2\cos\left(\frac{x_1(t)+x_2(t)}{2}\right)\sin\left(\frac{x_1(t)-x_2(t)}{2}\right)|$$
$$\le \max_{t\in[0,1]} |\sin\left(\frac{x_1(t)-x_2(t)}{2}\right)| \le \frac{1}{2}\max_{t\in[0,1]} |x_1(t) - x_2(t)| = \frac{1}{2}\rho(x_1, x_2)$$

这说明 $T : C[0,1] \to C[0,1]$ 是完备度量空间 $C[0,1]$ 上的压缩映射, 进而存在唯一不动点 $x_0 \in C[0,1]$, 亦即

$$x_0(t) = \frac{1}{2}\sin x_0(t) - a(t)$$

□

2. 考虑积分方程:

$$x(t) - \lambda \int_0^1 e^{t-s}x(s)ds = y(t)$$

其中 $y(t) \in C[0,1]$, λ 为常数, $|\lambda| < 1$, 证明存在唯一解 $x(t) \in C[0,1]$.

证明

知

$$x(t)-\lambda\int_0^1 e^{t-s}x(s)ds=y(t)\Leftrightarrow e^{-t}x(t)-\lambda\int_0^1 e^{-s}x(s)ds=e^{-t}y(t)$$

记 $z(t)=e^{-t}x(t)$, 设 $Tz(t)=\lambda\int_0^1 z(s)ds+e^{-t}y(t)$, 下面证明 $T:C[0,1]\to C[0,1]$ 是压缩映射. 任取 $z_1(t),z_2(t)\in C[0,1]$, 设 ρ 为 C 上的一致收敛性度量, 有:

$$\begin{aligned}\rho(Tz_1,Tz_2)&=\max_{t\in[0,1]}|\lambda\int_0^1 z_1(s)ds-\lambda\int_0^1 z_2(s)ds|\\&\le|\lambda|\max_{t\in[0,1]}|\int_0^1(z_1(s)-z_2(s))ds|\le|\lambda|\max_{t\in[0,1]}|z_1(t)-z_2(t)|=|\lambda|\rho(z_1,z_2)\end{aligned}$$

因为 $|\lambda|<1$ 是常数, 故 $T:C[0,1]\to C[0,1]$ 是完备度量空间 $C[0,1]$ 上的压缩映射, 进而存在唯一不动点 $z_0\in C[0,1]$, 亦即积分方程

$$z(t)-\lambda\int_0^1 z(s)ds=e^{-t}y(t)$$

存在唯一解 $z_0(t)$, 从而 $x_0(t)=e^tz_0(t)$ 是积分方程

$$x(t)-\lambda\int_0^1 e^{t-s}x(s)ds=y(t)$$

的唯一解. □

2.7.3 第三次作业

1. 设 F 是只有有限项不为零的实数列全体, 在 F 上引入距离

$$\rho(x,y)=\sup_{k\ge1}|\xi_k-\eta_k|,\ \forall x=(\xi_k)_{k\ge1},y=(\zeta_k)_{k\ge1}$$

求证 (F,ρ) 不完备, 并求其完备化空间.

证明

考虑实数列:

$$x_n=(1,\frac{1}{2},\cdots,\frac{1}{2^n},0,\cdots,0),\ n\in\mathbb{N}$$

显见对固定的 n, x_n 都是有限项不为零的实数列, 进而 $x_n\in F\Rightarrow\{x_n\}_{n\in\mathbb{N}}\subset F$. 下面证明 $\{x_n\}$ 是 (F,ρ) 中的基本列: 对任意的 $1>\varepsilon>0$, 取 $N=\log_2(\frac{1}{\varepsilon})-1$, 则对任意的 $m>n>N$, 都有:

$$\rho(x_m,x_n)=\frac{1}{2^{n+1}}<\frac{1}{2^{N+1}}=\varepsilon$$

这说明 $\{x_n\}_{n\in\mathbb{N}}$ 是 (F,ρ) 中的基本列, 但

$$\lim_{n\to\infty}x_n=(1,\frac{1}{2},\cdots,\frac{1}{2^n},\cdots)\notin F$$

故 (F,ρ) 不完备. 现设 F' 为满足

$$|x_k|\to0(k\to\infty)\text{ 且 }\sup_{k\ge1}|x_k|<\infty$$

的实数列 $x=(x_1,\cdots,x_n,\cdots)$ 的全体, 下面验证 ρ 确为 F' 上的距离, 任取 $x=(x_k)_{k\ge1},y=(y_k)_{k\ge1},z=(z_k)_{k\ge1}\in F'$:

- 正定性: $|x_k-y_k|\ge0(\forall k\in\mathbb{N})\Rightarrow\rho(x,y)\ge0$, 同时 $\rho(x,y)=0\Rightarrow\sup_{k\ge1}|x_k-y_k|=0\Rightarrow0\le|x_k-y_k|\le0(\forall k\in\mathbb{N})$, 故 $\forall k\in\mathbb{N}(x_k=y_k)$, 亦即 $x=y$.
- 交换律: $\rho(x,y)=\sup_{k\ge1}|x_k-y_k|=\sup_{k\ge1}|y_k-x_k|=\rho(y,x)$;
- 三角不等式: 根据实数的三角不等式有: $|x_k-y_k|\le|x_k-z_k|+|z_k-y_k|(\forall k\in\mathbb{N})$, 故

$$\sup_{k\ge1}|x_k-y_k|\le\sup_{k\ge1}(|x_k-z_k|+|z_k-y_k|)\le\sup_{k\ge1}|x_k-z_k|+\sup_{k\ge1}|z_k-y_k|\Rightarrow\rho(x,y)\le\rho(x,z)+\rho(z,y)$$

综上, (F',ρ) 为度量空间, 下面简记为 F'.

再证明 F' 的完备性, 任取 $x_m = (x_m^k)_{k\geq 1} \in F'$ 是基本列, 根据基本列的定义知:

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N (\rho(x_m^k, x_n^k) = \sup_{k\geq 1} |x_m^k - x_n^k| < \varepsilon)$$

根据上确界的定义知:

$$\forall k \geq 1 (|x_m^k - x_n^k| \leq \sup_{k\geq 1} |x_m^k - x_n^k| < \varepsilon)$$

这说明对每个固定的 k 而言, $\{x_m^k\}$ 是 $\mathbb{R}$ 中的基本列, 进而根据 $\mathbb{R}$ 的完备性知其有极限 x^k, 记:

$$x = (x^1, x^2, \cdots, x^k, \cdots)$$

下面说明 $x \in F'$. 对任意的 $k \geq 1$, 前述 $\mathbb{R}$ 中的极限定义为:

$$\exists N_k > 0 \forall n > N_k (|x_n^k - x^k| < \varepsilon) \tag{2.61}$$

故若 $x \notin F'$, 根据 F' 的定义知

$$\forall M > 0 \exists k_M \geq 1 (|x^{k_M}| > M) \text{ 或 } \exists \varepsilon_0 > 0 \forall K > 0 \exists k_K > K (|x^{k_K}| > \varepsilon_0)$$

对第一种情况, 取 $n > N_k$ 有

$$\varepsilon > |x_n^{k_M} - x^{k_M}| \geq |x^{k_M}| - |x_n^{k_M}| > M - |x_n^{k_M}| \Rightarrow |x_n^{k_M}| > M - \varepsilon$$

不妨设 $\varepsilon < \frac{M}{2}$, 这便说明当 n 足够大时, 对任意 $M > 0$, 总能取到 $k_M \geq 1$, 使得 $|x_n^{k_M}| > \frac{M}{2}$, 根据 M 的任意性知 $x_n \notin l^\infty$, 矛盾! 而对第二种情况, 取 $\varepsilon = \frac{\varepsilon_0}{2}$ 并取 $n > N_k$, 有:

$$\varepsilon > |x_n^{k_K} - x^{k_K}| \geq |x^{k_K}| - |x_n^{k_K}| > \varepsilon_0 - |x_n^{k_K}| \Rightarrow |x_n^{k_K}| > \frac{\varepsilon_0}{2}$$

这说明当 n 足够大时, 总存在足够大的 k_K 使得 $|x_n^{k_K}| > \frac{\varepsilon_0}{2}$, 但这与 x_n 的定义不符, 矛盾! 综上知 $x \in F'$.

下面说明 x 确为 $\{x_n\}$ 在 F' 中的极限, 用反证法. 如若存在 ε_0, 使得对任意的 $N > 0$, 总存在 $n_N > N$, 使得

$$\rho(x, x_{n_N}) = \sup_{k\geq 1} |x^k - x_{n_N}^k| \geq \varepsilon_0$$

根据上确界的定义, 这表明对任意的 $\varepsilon_1 > 0$, 总存在 $k_{\varepsilon_1} \geq 1$, 使得

$$|x^{k_{\varepsilon_1}} - x_{n_N}^{k_{\varepsilon_1}}| \geq \varepsilon_0 - \varepsilon_1$$

取 $\varepsilon_1 = \frac{\varepsilon_0}{2}$, 这便与(2.61)式矛盾, 故 $\lim_{n\to\infty} \rho(x, x_n) = 0$, 确有 $x_n \to x \in F'(n \to \infty)$.

最后说明 F 与 F' 的某个稠密子集等距同构. 取恒等映射:

$$\mathrm{id} : F \to F', \ (x_1, \cdots, x_n, 0, \cdots) \mapsto (x_1, \cdots, x_n, 0, \cdots)$$

显见 id 是等距映射, 下面再说明 F 在 F' 中稠密. 任取 $x = (x^1, \cdots, x^k, \cdots) \in F'$, 取:

$$x_n = (x^1, \cdots, x^n, 0, \cdots)$$

根据 F' 的定义, 对 x 有:

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N (|x^n| < \varepsilon \Rightarrow \sup_{n\geq N+1} |x^n| \leq \varepsilon)$$

而易知

$$\rho(x_n, x) = \sup_{k\geq n+1} |x^k|$$

取 $n > N$ 即知 $\rho(x_n, x) \leq \varepsilon$, 进而由 ε 的任意性即得 $x_n \to x \in F'(n \to \infty)$, 故 F 在完备度量空间 F' 中稠密, 进而 F' 为 F 的完备化空间. □

2. 设 (X, ρ) 完备, $\{F_n\}$ 是 X 内的单调下降非空闭集序列, 即

$$F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots, \ F_n \neq \varnothing$$

且 F_n 的直径 $d_n = d(F_n) \to 0$, 证明: $\bigcap_{n\geq 1} F_n \neq \varnothing$. 如果没有条件 $d_n \to 0$, 结论成立吗? 如果 X 为列紧空间时, 关于直径的条件是否还需要?

证明

取 $x_1 \in F_1, \cdots, x_n \in F_n, \cdots$, 由 $d(F_n) \to 0$ 的定义知:

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N (d(F_n) < \varepsilon)$$

现对 $m > n > N$, 由 $\{F_n\}$ 的递降性知 $x_m, x_n \in F_n$, 故

$$\rho(x_m, x_n) \le d(F_n) < \varepsilon$$

这说明点列 $\{x_n\}$ 是 (X, ρ) 中的基本列, 根据 (X, ρ) 的完备性知 $\{x_n\}$ 有极限, 记之为 x, 下面证明 $\forall n \in \mathbb{N}(x \in F_n)$. 用反证法: 若

$$\exists n_0 \in \mathbb{N}(x \notin F_{n_0} \Leftrightarrow x \in F_{n_0}^c)$$

因为 F_{n_0} 是闭集, 故 $F_{n_0}^c$ 是开集, 从而根据开集的定义知:

$$\exists \delta > 0 (B(x, \delta) \cap F_{n_0} = \varnothing)$$

现对 $F_{n_0} \ni x_{n_0}, x_{n_0+1}, \cdots$ 有

$$\rho(x, x_k) \ge \rho(x, F_{n_0}) \ge \delta > 0,\ k \ge n_0 \tag{2.62}$$

但另一方面, 根据 $x_n \to x (n \to \infty)$ 的定义:

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N (\rho(x, x_n) < \varepsilon)$$

这与(2.62)式矛盾! 故 $\forall n \in \mathbb{N}(x \in F_n)$, 也即 $x \in \bigcap_{n \ge 1} F_n$, 故 $\bigcap_{n \ge 1} F_n \ne \varnothing$.

若没有 $d_n \to 0$, 结论不成立. 下面给出来自*泛函分析中的反例 (汪林著)* 中的例子:

任取 $m, n \in \mathbb{N}$, 取

$$d(m, n) = \begin{cases} 1 + \frac{1}{m+n}, & m \ne n \\ 0, & m = n \end{cases}$$

下面说明 d 是 $\mathbb{N}$ 上的距离. 正定性与交换律显见, 设 $p, q, r \in \mathbb{N}$, 有:

$$d(p, r) = 1 + \frac{1}{p+r} < 2 < (1 + \frac{1}{p+q}) + (1 + \frac{1}{q+r}) = d(p, q) + d(q, r)$$

故 d 为 $\mathbb{N}$ 上的距离, 任取 $(\mathbb{N}, d)$ 中的基本列, 根据定义知:

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N (d(x_m, x_n) < \varepsilon)$$

若 $x_m \ne x_n$, 知 $d(x_m, x_n) = 1 + \frac{1}{x_m + x_n} > 1$, 矛盾! 故只能有 $x_m = x_n$, 也即 $\{x_n\}$ 在 $n > N$ 时为常数列 (设该常数为 k_x), 进而其收敛到 $k_x \in \mathbb{N}$, 故 $(\mathbb{N}, d)$ 完备. 现设 $B_n = \{m : d(n, m) \le 1 + \frac{1}{2n}\}$, 知 $B_n = \{n, n+1, n+2, \cdots\}$, 显见 $\{B_n\}$ 是单调下降非空闭集列, 但 $\bigcap_{n=1}^{\infty} B_n = \varnothing$.

最后对于列紧空间, 结论不再需要直径趋零. 此时因为 F_1 是列紧空间 (X, ρ) 中的闭集, 故其为自列紧集. 取 $x_1 \in F_1, x_2 \in F_2 \subset F_1, \cdots, x_n \in F_n \subset F_1, \cdots$, 知 $\{x_n\} \subset F_1$, 进而由 F_1 的自列紧性知其有收敛到 F_1 中的子列 $\{x_{n_k}\}$, 记 $x_{n_k} \to x (k \to \infty)$, 下面证明 $\forall k \in \mathbb{N}(x \in F_{n_k})$. 用反证法: 若

$$\exists k_0 \in \mathbb{N}(x \notin F_{n_{k_0}} \Leftrightarrow x \in F_{n_{k_0}}^c)$$

因为 $F_{n_{k_0}}$ 是闭集, 故 $F_{n_{k_0}}^c$ 是开集, 从而根据开集的定义知

$$\exists \delta > 0 (B(x, \delta) \cap F_{n_{k_0}} = \varnothing)$$

现对 $F_{n_{k_0}} \ni x_{n_{k_0}}, x_{n_{k_0+1}}, \cdots$ 有

$$\rho(x, x_{n_k}) \ge \rho(x, F_{n_{k_0}}) \ge \delta > 0,\ k \ge k_0$$

但这与 $x_{n_k} \to x (k \to \infty)$ 矛盾! 故 $\forall k \in \mathbb{N}(x \in F_{n_k})$. 又因为 $\{F_n\}$ 是递降列, 且 $n_k \to \infty (k \to \infty)$, 故 $\forall n \in \mathbb{N}(x \in F_n)$, 也即 $x \in \bigcap_{n \ge 1} F_n$, 故 $\bigcap_{n \ge 1} F_n \ne \varnothing$. □

2.7.4 第四次作业

1. 给定距离空间 (X,d), 设 $A\subset X$ 是自列紧集, 求证: A 上连续函数必有界, 亦达到它的上下确界.

证明

设 A 上有连续函数 $f:A\to\mathbb{R}$, 根据连续性的定义知:

$$\forall x\in A\forall\varepsilon>0\exists\delta=\delta(x,\varepsilon)\forall y\in A(d(x,y)<\delta\Rightarrow|f(x)-f(y)|<\varepsilon)$$

现在固定 ε, 显见 $A\subset\bigcup_{x\in A}B(x,\delta(x,\varepsilon))$, 这说明 $\bigcup_{x\in A}B(x,\varepsilon)$ 可代表 A 的一个开覆盖. 因为 A 是距离空间 X 中的自列紧集, 故其为紧集, 从而其任意开覆盖必存在有限子覆盖, 进而存在 $x_1,\cdots,x_n\in A$ 使得:

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

从而

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

现在对于 $f(B(x_i,\delta(x_i,\varepsilon)))$ 知:

$$\forall 1\le i\le n\forall y\in B(x_i,\delta(x_i,\varepsilon))(|f(y)-f(x_i)|<\varepsilon\Rightarrow|f(y)|<|f(x_i)|+\varepsilon)$$

现取 $M=\max\{|f(x_1)|+\varepsilon,\cdots,|f(x_n)|+\varepsilon\}$, 则

$$\forall x\in A(|f(x)|<M)$$

这正是 f 在 A 上有界的定义.

再证明 f 在 A 上可达到上下确界. 设 $f(A)$ 的上确界为 u, 依照上确界的定义知:

$$\forall n\in\mathbb{N}\exists y_n\in f(A)(u-\frac{1}{n}\le y_n\le u)$$

因为 $f(A)$ 有界, 故 $\{y_n\}\subset f(A)$ 是有界序列, 从而由 $\mathbb{R}$ 的完备性知其至少在 $\mathbb{R}$ 中有收敛子列 $\{y_{n_q}\}$, 设 $y_{n_q}\to y_0\in\mathbb{R}(q\to\infty)$. 因为 $n_q\to\infty(q\to\infty)$, 故由夹逼定理知 $y_0=u$. 另一方面, 根据 $f(A)$ 的定义, 设对每个 $y_{n_q}\in f(A)$, 存在 $x_{n_q}\in A$ 使得 $y_{n_q}=f(x_{n_q})$. 对这样得到的 $\{x_{n_q}\}\subset A$, 因为 A 至少作为列紧集是有界的, 故 $\{x_{n_q}\}$ 是 A 中的有界序列, 因而由 A 的自列紧性知其具有收敛到 A 中的子列 $\{x_{n_{q_r}}\}$, 设 $x_{n_{q_r}}\to x_0\in A(r\to\infty)$, 并设 $f(x_{n_{q_r}})=y_{n_{q_r}}\in\{y_{n_q}\}$, 根据 f 的连续性有:

$$u=y_0=\lim_{r\to\infty}y_{n_{q_r}}=\lim_{r\to\infty}f(x_{n_{q_r}})=f(\lim_{r\to\infty}x_{n_{q_r}})=f(x_0)$$

这说明 $u=f(x_0),x_0\in A$, 故 f 可在 A 上达到上确界, 下确界类似可证. □

2. 设 X 是距离空间, $M\subset X$ 是自列紧集, $f:M\to\mathbb{R}^1$ 是连续函数, 则 $f(x)$ 在 M 上一致连续.

证明

设 X 上的距离为 d, 根据 f 的连续性知:

$$\forall x\in X\forall\varepsilon>0\exists\delta=\delta(x,\varepsilon)\forall y\in B(x,\delta)(|f(x)-f(y)|<\varepsilon)$$

显见 $M\subset\bigcup_{x\in M}B(x,\frac{1}{2}\delta(x,\varepsilon))$, 故 $\bigcup_{x\in M}B(x,\frac{1}{2}\delta(x,\varepsilon))$ 可代表 M 的一个开覆盖. 因为 $M\subset X$ 是自列紧集, X 是距离空间, 故 M 是紧集, 从而其任意开覆盖均存在有限子覆盖, 设 $M\subset\bigcup_{i=1}^{n}B(x_i,\frac{1}{2}\delta(x_i,\varepsilon))$. 取 $\delta_0:=\min\{\delta(x_1,\varepsilon),\cdots,\delta(x_n,\varepsilon)\}$, 现对任意的 $x,y\in M$, 若 $d(x,y)<\frac{1}{2}\delta_0$, 对于 x 知必存在 $i_x\in\{1,\cdots,n\}$ 使得 $x\in B(x_{i_x},\frac{1}{2}\delta(x_{i_x},\varepsilon))$, 进而

$$d(y,x_{i_x})\le d(y,x)+d(x,x_{i_x})<\frac{1}{2}\delta_0+\frac{1}{2}\delta(x_{i_x},\varepsilon)\le\delta(x_{i_x},\varepsilon)$$

这说明 $y\in B(x_{i_x},\delta(x_{i_x},\varepsilon))$, 因而 $x,y\in B(x_{i_x},\varepsilon)$, 有:

$$|f(x)-f(y)|\le|f(x)-f(x_{i_x})|+|f(x_{i_x})-f(y)|<2\varepsilon$$

最后由 ε 的任意性即知 $f(x)$ 在 M 上一致连续. □

3. 设 M 是 $C[a,b]$ 中的有界集, 证明集合

$$S=\{F(x)=\int_a^x f(t)dt : f\in M\}$$

是列紧集.

证明

要证明 $S\subset C[a,b]$ 是列紧集, 只需证明其一致有界且等度连续.

对一致有界性, 既然 $M\subset C[a,b]$ 有界, 设存在 $C>0$ 使得:

$$\forall f\in M(\|f\|=\max_{x\in[a,b]}|f(x)|\le C)$$

从而任取 $F\in S$, 知:

$$|F(x)|=|\int_a^x f(t)dt|\le\int_a^x|f(t)|dt\le\int_a^b\max_{t\in[a,b]}|f(t)|dt\le C(b-a),\ \forall x\in[a,b]$$

这说明 $\max_{x\in[a,b]}|F(x)|=\|F\|\le C(b-a)$, 而这对 S 中的任意 F 均成立, 故 S 在 $C[a,b]$ 中有界, 亦即对 x 一致有界.

对等度连续性, 对任意的 $\varepsilon>0$ 取 $\delta=\frac{\varepsilon}{C}$, 则对任意的 $F\in S$, 当任意 $x_1,x_2\in[a,b]$ 满足 $|x_1-x_2|<\delta$ 时, 有:

$$|F(x_1)-F(x_2)|\le|\int_{x_2}^{x_1}|f(t)|dt|\le C|x_1-x_2|<\varepsilon$$

从而 S 等度连续, 因而由 Arzela-Ascoli 定理知 S 是 $C[a,b]$ 中的列紧集. □

4. 在 $\mathbb{R}^n$ 中, 对于 $x=(x_1,\cdots,x_n)$, 规定

$$\varphi(x)=(\sqrt{|x_1|}+\cdots+\sqrt{|x_n|})^2$$

φ 是范数吗?

证明

不是, 取 $x=(0,1,0,\cdots,0)$, $y=(1,0,\cdots,0)$, 则 $\varphi(x)=\varphi(y)=1$, 而 $\varphi(x+y)=4>\varphi(x)+\varphi(y)$, 故 φ 不满足三角不等式, 因而不是范数. □

5. 设 X 表示复序列 $x=(x_1,x_2,\cdots)$ 的全体, 定义

$$\|x\|=\sum_{n=1}^{\infty}2^{-n}\min\{1,|x_n|\}$$

(1) $\|\cdot\|$ 是否是 X 上的范数?

(2) $\|x-y\|$ 是否可定义为 X 上的距离? 假若可以, 说明 $\|x^{(n)}-x\|\to 0(n\to\infty)$ 的意义.

证明

(1) 不是. 取 $x=(\frac{1}{3},0,0,\cdots)$, 知 $\|x\|=\frac{1}{2}\times\min\{1,\frac{1}{3}\}=\frac{1}{6}$, 而对 $4x=(\frac{4}{3},0,0,\cdots)$ 有 $\|4x\|=\frac{1}{2}\times\min\{1,\frac{4}{3}\}=\frac{1}{2}\ne 4\|x\|$, 故 $\|\cdot\|$ 不满足齐次性, 因而不是范数.

(2) 可以. 任取 $x=(x_1,x_2,\cdots)$, $y=(y_1,y_2,\cdots)$, $z=(z_1,z_2,\cdots)$, 下面验证距离定义.

- (正定性) 因为 $|x_n-y_n|\ge 0(n=1,2,\cdots)$, 故 $\|x-y\|=\sum_{n=1}^{\infty}2^{-n}\min\{1,|x_n-y_n|\}\ge 0$. 显见 $x=y\Rightarrow\|x-y\|=0$. 当 $\|x-y\|=0$ 时, 若存在 $|x_k-y_k|\ne 0$, 则 $2^{-k}\min\{1,|x_k-y_k|\}>0$, 从而

$$\sum_{n=1}^{\infty}2^{-n}\min\{1,|x_n-y_n|\}\ge 2^{-k}\min\{1,|x_k-y_k|\}>0$$

  矛盾! 故只能有 $x_k=y_k(k=1,2,\cdots)$, 亦即 $x=y$, 正定性至此得证.

- (对称性) 因为 $|x_k-y_k|=|y_k-x_k|(k=1,2,\cdots)$, 故

$$\sum_{n=1}^{\infty}2^{-n}\min\{1,|x_n-y_n|\}=\sum_{n=1}^{\infty}2^{-n}\min\{1,|y_n-x_n|\}$$

  亦即 $\|x-y\|=\|y-x\|$, 对称性至此得证.

- (三角不等式) 因为 $|x_k - y_k| \le |x_k - z_k| + |z_k - y_k| (k = 1, 2, \cdots)$, 故对 $k \in \mathbb{N}$ 有:

$$\min\{1, |x_k - y_k|\} \le \min\{1, |x_k - z_k| + |z_k - y_k|\} \le \min\{1, |x_k - z_k|\} + \min\{1, |z_k - y_k|\}$$

因而

$$\sum_{n=1}^{\infty} 2^{-n} \min\{1, |x_n - y_n|\} \le \sum_{n=1}^{\infty} 2^{-n} \min\{1, |x_n - z_n|\} + \sum_{n=1}^{\infty} 2^{-n} \min\{1, |z_n - y_n|\}$$

亦即 $\|x - y\| \le \|x - z\| + \|z - y\|$, 三角不等式至此得证.

综上, $\|x - y\|$ 是 X 上的距离, 下面说明 $\|x^{(n)} - x\| \to 0 (n \to \infty)$ 和 $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, \cdots)$ 依坐标趋于 $x = (x_1, x_2, \cdots)$ 等价, 即:

$$\|x^{(n)} - x\| \to 0 (n \to \infty) \Leftrightarrow \forall k \in \mathbb{N} \forall \varepsilon > 0 \exists N = N(k, \varepsilon) > 0 \forall n > N(|x_k^{(n)} - x_k| < \varepsilon)$$

当 $x^{(n)}$ 依坐标趋于 x 时, 对上式中已取定的 ε, 知存在 $K > 0$ 使得

$$\sum_{k=K}^{\infty} 2^{-k} < \varepsilon \Rightarrow \sum_{k=K}^{\infty} 2^{-k} \min\{1, |x_k^{(n)} - x_k|\} \le \sum_{k=K}^{\infty} 2^{-k} < \varepsilon$$

而对每个满足 $k < K$ 的 k, 总对应有 $N_k = N(k, \varepsilon)$ 使得 $n > N_k$ 时 $|x_k^{(n)} - x_k| < \varepsilon$, 现取 $N_0 = \max_{1 \le k < K} N_k$, 则知当 $n > N_0$ 时对 $1 \le k < K$ 都有 $|x_k^{(n)} - x_k| < \varepsilon$, 进而 $n > N_0$ 时有:

$$\begin{aligned}
\sum_{k=1}^{\infty} 2^{-k} \min\{1, |x_k^{(n)} - x_k|\} &= \sum_{k=1}^{K-1} 2^{-k} \min\{1, |x_k^{(n)} - x_k|\} + \sum_{k=K}^{\infty} 2^{-k} \min\{1, |x_k^{(n)} - x_k|\} \\
&\le \sum_{k=1}^{K-1} 2^{-k} \varepsilon + \sum_{k=K}^{\infty} 2^{-k} \le \varepsilon \sum_{k=1}^{\infty} 2^{-k} = \varepsilon
\end{aligned}$$

故根据定义知 $\|x^{(n)} - x\| \to 0 (n \to \infty)$.

如若 $x^{(n)}$ 并非依坐标趋于 x, 即:

$$\exists k_0 \in \mathbb{N} \exists \varepsilon_0 > 0 \forall N > 0 \exists n > N(|x_{k_0}^{(n)} - x_{k_0}| \ge \varepsilon_0)$$

则

$$\|x^{(n)} - x\| = \sum_{k=1}^{\infty} 2^{-k} \min\{1, |x_k^{(n)} - x_k|\} \ge 2^{-k_0} \min\{1, |x_{k_0}^{(n)} - x_{k_0}|\}$$

令 $\varepsilon_0' = 2^{-k_0} \min\{1, |x_{k_0}^{(n)} - x_{k_0}|\}$, 这便说明 $\|x^{(n)} - x\| \nrightarrow 0 (n \to \infty)$. 取逆否命题即知当 $\|x^{(n)} - x\| \to 0 (n \to \infty)$ 时, 必有 $x^{(n)}$ 依坐标趋于 x.

综上, $\|x^{(n)} - x\| \to 0 (n \to \infty)$ 与 $x^{(n)}$ 依坐标趋于 x 等价. □

2.7.5 第五次作业

1. 设 $\{x_n\}$ 是 Banach 空间 X 中的点列. 若存在 $(0, \infty)$ 上的非负递减的可积函数 $g(t)$, 使得 $\|x_n\| \le g(n) (n = 1, 2, 3, \cdots)$. 证明 $\sum_{n=1}^{\infty} x_n$ 收敛.

证明

根据 X 是 Banach 空间, 知只需证明 $\|\sum_{n=1}^{\infty} x_n\| < \infty$. 根据 $g(t)$ 是 $(0, \infty)$ 上的非负递减可积函数知:

$$\sum_{n=0}^{\infty} g(n) = \sum_{n=0}^{\infty} \int_n^{n+1} g(n) dt \le \int_0^{\infty} g(t) dt < \infty$$

因而

$$\sum_{n=0}^{\infty} \|x_n\| \le \sum_{n=0}^{\infty} g(n) < \infty$$

故

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

亦即 $\sum_{n=1}^{\infty} x_n$ 收敛到 X 中, 命题得证. □

2. 设 $(X_k, \|\cdot\|_k)$ 是一列赋范空间, $x = \{x_k\}$, 其中 $x_k \in X_k (k = 1, 2, \cdots)$ 且满足条件 $\sum_{k=1}^{\infty} \|x_k\|_k^p < \infty$, 用 X 表示所有 x 的全体, 按坐标定义线性运算构成的线性空间, 在 X 中定义

$$\|x\| = (\sum_{k=1}^{\infty} \|x_k\|_k^p)^{\frac{1}{p}} \quad p \geq 1$$

证明 $(X, \|\cdot\|)$ 是一个赋范空间.

证明

只需验证 $\|\cdot\|$ 确为范数.

- 正定性: 已知 $\|\cdot\|_k (k = 1, 2, \cdots)$ 均为范数, 由它们的正定性易得 $\|\cdot\|$ 的正定性.
- 齐次性: 任取 $\alpha \in \mathbb{K}$, 因为 X 中的线性运算按坐标定义, 知

$$\alpha x = \{\alpha x_k\}_{k=1}^{\infty}$$

因而

$$\|\alpha x\| = (\sum_{k=1}^{\infty} \|\alpha x_k\|_k^p)^{\frac{1}{p}} = (\sum_{k=1}^{\infty} |\alpha| \|x_k\|_k^p)^{\frac{1}{p}} = |\alpha| (\sum_{k=1}^{\infty} \|x_k\|_k^p)^{\frac{1}{p}} = |\alpha| \|x\|$$

齐次性得证.

- 三角不等式: 任取 $x, y \in X$, 设

$$x = \{x_k\}_{k=1}^{\infty}, \ y = \{y_k\}_{k=1}^{\infty}$$

因为 X 中的线性运算按坐标定义, 故

$$x + y = \{x_k + y_k\}_{k=1}^{\infty}$$

知

$$\begin{aligned}\|x + y\| &= (\sum_{k=1}^{\infty} \|x_k + y_k\|_k^p)^{\frac{1}{p}} \leq (\sum_{k=1}^{\infty} (\|x_k\|_k + \|y_k\|_k)^p)^{\frac{1}{p}} \\ &\overset{\text{(i)}}{\leq} (\sum_{k=1}^{\infty} \|x_k\|_k^p)^{\frac{1}{p}} + (\sum_{k=1}^{\infty} \|y_k\|_k^p)^{\frac{1}{p}} = \|x\| + \|y\|\end{aligned}$$

其中 (i) 是 $\mathbb{R}$ 上的 Minkowski 不等式, 三角不等式得证.

综上, $\|\cdot\|$ 确为 X 上的范数, $(X, \|\cdot\|)$ 是赋范空间, 命题得证. □

2.7.6 第六次作业

1. 设 $\mathscr{X}$ 是 B^* 空间. 求证: $\mathscr{X}$ 是 B 空间, 必须且仅须对 $\forall \{x_n\}_{n=1}^{\infty} \subset \mathscr{X}, \sum_{n=1}^{\infty} \|x_n\| < \infty \Rightarrow \sum_{n=1}^{\infty} x_n$ 收敛.

证明

当 $\mathscr{X}$ 是 B 空间, 任取 $\{x_n\}_{n=1}^{\infty} \subset \mathscr{X}$, 若 $\sum_{n=1}^{\infty} \|x_n\| < \infty$, 则知 $\{\sum_{n=1}^{k} \|x_n\|\}_{k=1}^{\infty}$ 是 $\mathbb{R}$ 中的基本列, 因而由定义知:

$$\forall \varepsilon > 0 \exists K > 0 \forall p > q > K (|\sum_{n=1}^{p} \|x_n\| - \sum_{n=1}^{q} \|x_n\|| < \varepsilon)$$

即

$$\sum_{n=q+1}^{p} \|x_n\| < \varepsilon$$

现记 $S_k = \sum_{n=1}^{k} x_n$, 要说明 $\sum_{n=1}^{\infty} x_n$ 收敛, 就是说明 $\{S_k\}$ 收敛, 进而由 $\mathscr{X}$ 是 B 空间知只需说明 $\{S_k\}$ 是基本列. 沿用前述记号, 知:

$$\|S_p - S_q\| = \|\sum_{n=q+1}^{p} x_n\| \leq \sum_{n=q+1}^{p} \|x_n\| < \varepsilon$$

故 $\{S_k\}$ 是基本列, 因而收敛.

当 $\forall\{x_n\}_{n=1}^{\infty}\subset\mathscr{X},\sum_{n=1}^{\infty}\|x_n\|<\infty\Rightarrow\sum_{n=1}^{\infty}x_n$ 收敛, 用反证法. 若 $\mathscr{X}$ 并非 B 空间, 亦即存在 $\mathscr{X}$ 中的基本列 $\{y_n\}$ 在 $\mathscr{X}$ 中不收敛, 根据基本列的定义知:

$$\forall\varepsilon>0\exists N>0\forall p>q>N(\|y_p-y_q\|<\varepsilon)$$

取 $\varepsilon=\frac{1}{2^k}$, 并设 $\{y_n\}$ 的子列 $\{y_{n_k}\}$ 满足:

$$\|y_{n_{k+1}}-y_{n_k}\|\le\frac{1}{2^k}$$

现定义

$$x_1=y_{n_1},\ x_k=y_{n_{k+1}}-y_{n_k}(k\ge 1)$$

显见 $\{x_k\}_{k=1}^{\infty}\subset\mathscr{X}$, 知

$$\sum_{k=1}^{\infty}\|x_k\|\le\sum_{k=1}^{\infty}\frac{1}{2^k}=\frac{1}{2}<\infty$$

这说明 $\sum_{k=1}^{\infty}x_k$ 收敛, 亦即 $\{y_{n_k}\}_{k=1}^{\infty}$ 收敛, 又因为其为基本列的子列, 故基本列 $\{y_n\}$ 也收敛, 矛盾!

综上, 命题即证. □

2. 记 $[a,b]$ 上次数不超过 n 的多项式全体为 $\mathbb{P}_n$. 求证: $\forall f(x)\in C[a,b],\exists P_0(x)\in\mathbb{P}_n$, 使得

$$\max_{a\le x\le b}|f(x)-P_0(x)|=\min_{P\in\mathbb{P}_n}\max_{a\le x\le b}|f(x)-P(x)|.$$

证明

设 $g_k(x)=x^k(1\le k\le n,x\in[a,b])$, 显见 $g_k\in C[a,b](k=1,\cdots,n)$. 根据 $\mathbb{P}_n$ 的定义易证

$$\mathbb{P}_n=\operatorname{span}\{1,x,\cdots,x^n\}=\{\sum_{k=1}^{n}\alpha_k x^k:(\alpha_1,\cdots,\alpha_n)\in\mathbb{K}^n\}$$

届于 $C[a,b]$ 是 B^* 空间, 故对其中固定的向量组 $\{g_k\}_{k=1}^{n}$ 与任意元素 f 而言, 总存在系数 $\lambda_1,\cdots,\lambda_n\in\mathbb{K}$ 使得:

$$\|f-\sum_{k=1}^{n}\lambda_k g_k\|=\min_{\{\alpha_k\}_{k=1}^{n}\in\mathbb{K}^n}\|f-\sum_{k=1}^{n}\alpha_k g_k\|$$

代入一致收敛范数的定义:

$$\max_{x\in[a,b]}|f(x)-\sum_{k=1}^{n}\lambda_k g_k(x)|=\min_{\{\alpha_k\}_{k=1}^{n}\in\mathbb{K}^n}\max_{x\in[a,b]}|f(x)-\sum_{k=1}^{n}\alpha_k g_k(x)|$$

命题即证. □

3. 设 $\mathscr{X}$ 是 B^* 空间, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间, 假定 $\exists c\in(0,1)$, 使得

$$\inf_{x\in\mathscr{X}_0}\|y-x\|\le c\|y\|,\ \forall y\in\mathscr{X}$$

求证: $\mathscr{X}_0$ 在 $\mathscr{X}$ 中稠密.

证明

根据下确界的定义, 任意取定 $y\in\mathscr{X}$, 有:

$$\forall\varepsilon>0\exists x_1^y\in\mathscr{X}_0(\|y-x_1^y\|\le\varepsilon+\inf_{x\in\mathscr{X}_0}\|y-x\|)$$

因为 $\mathscr{X}$ 是 B^* 空间, 故 $y-x_1^y\in\mathscr{X}$, 进而沿用上述记号, 根据下确界的定义有:

$$\exists x_2^y\in\mathscr{X}_0(\|(y-x_1^y)-x_2^y\|\le\varepsilon+\inf_{x\in\mathscr{X}_0}\|(y-x_1^y)-x\|)$$

依题知

$$\inf_{x\in\mathscr{X}_0}\|(y-x_1^y)-x\|\le c\|y-x_1^y\|\le c\varepsilon+c\inf_{x\in\mathscr{X}^0}\|y-x\|$$

得到

$$\|y-x_1^y-x_2^y\| \le \varepsilon + c\varepsilon + c \inf_{x\in\mathscr{X}^0} \|y-x\|$$

重复上述过程, 在第 n 次知:

$$\exists x_n^y \in \mathscr{X}_0(\|y-\sum_{k=1}^{n} x_k^y\| \le \varepsilon \sum_{k=1}^{n-1} c^k + c^{n-1} \inf_{x\in\mathscr{X}^0} \|y-x\|)$$

当 $n\to\infty$ 时, 由 $c\in(0,1)$ 是常数知 $c^{n-1}\to 0, \sum_{k=1}^{n-1} c^k < \infty$, 故由 ε 的任意性即知 $\mathscr{X}_0$ 中的序列 $\{\sum_{k=1}^{n} x_k^y\}_{n=1}^{\infty}$ 趋向 y, 进而由 y 的任意性即得 $\mathscr{X}_0$ 在 $\mathscr{X}$ 中稠密. □

2.7.7 第七次作业

1. 求证: 在 $C[a,b]$ 中不可能引进一种内积 $(\cdot,\cdot)$, 使其满足

$$(f,f)^{\frac{1}{2}} = \max_{a\le x\le b} |f(x)|, \ \forall f\in C[a,b].$$

证明

只需验证一致范数不满足平行四边形等式. 用反证法, 如若:

$$\max_{[a,b]} |f+g| + \max_{[a,b]} |f-g| = 2(\max_{[a,b]} |f| + \max_{[a,b]} |g|), \ \forall f,g\in C[a,b]$$

取

$$f(x)=1, \ g(x)=\begin{cases} \frac{2}{b-a}(x-\frac{a+b}{2}), & x\in[a,\frac{a+b}{2}] \\ \frac{2}{a-b}(x-\frac{a+b}{2}), & x\in[\frac{a+b}{2},b] \end{cases}$$

计算知

$$\max_{[a,b]} |f+g| = 1, \max_{[a,b]} |f-g| = 2, \max_{[a,b]} |f| = 1, \max_{[a,b]} |g| = 1$$

从而

$$\max_{[a,b]} |f+g| + \max_{[a,b]} |f-g| = 3 \ne 4 = 2(\max_{[a,b]} |f| + \max_{[a,b]} |g|)$$

命题进而得证. □

2. 在 $L^2[0,T]$ 中, 求证: 函数

$$x\mapsto |\int_0^T e^{-(T-\tau)} x(\tau)d\tau|, \ \forall x\in L^2[0,T]$$

在单位球面上达到最大值, 并求出此最大值和达到最大值的元素 x.

证明

当 x 在单位球面上时, 知 $\int_0^T x^2(\tau)d\tau = 1$. 由 Cauchy-Schwartz 不等式知

$$|\int_0^T e^{-(T-\tau)}x(\tau)d\tau| \le (\int_0^T e^{-2(T-\tau)}d\tau)^{\frac{1}{2}} \cdot (\int_0^T x^2(\tau)d\tau)^{\frac{1}{2}} = \frac{1}{\sqrt{2}}(1-e^{-2T})^{\frac{1}{2}}$$

等号成立当且仅当 $x(t) = ke^{-(T-t)}(t\in[0,T])$, 其中 k 为常数. 代入 $\|x\|_{L^2}=1$ 知:

$$k^2\int_0^T e^{-2(T-\tau)}d\tau = 1 \Rightarrow k = \pm\frac{\sqrt{2}}{(1-e^{-2T})^{\frac{1}{2}}}$$

因而

$$x(t) = \sqrt{2}\frac{e^{-(T-t)}}{(1-e^{-2T})^{\frac{1}{2}}} \text{ 或 } x(t) = -\sqrt{2}\frac{e^{-(T-t)}}{(1-e^{-2T})^{\frac{1}{2}}}, \ t\in[0,T].$$

□

3. 设 M,N 是内积空间中的两个子集, 求证:

$$M\subset N \Rightarrow N^{\perp}\subset M^{\perp}.$$

证明

当 $M \subset N$ 时, 任取 $x \in N^{\perp}$, 根据定义知 $\forall n \in N((x,n)=0)$. 因为 $M \subset N$, 故 $\forall m \in M \subset N((x,m)=0)$, 这说明 $x \in M^{\perp}$. 根据 x 的任意性即得 $N^{\perp} \subset M^{\perp}$. □

2.7.8 第八次作业

1. 设 $A=\{e_k\}$ 是内积空间 X 中的标准正交系, 证明对 $\forall x, y \in X$, 有

$$\sum_{k=1}^{\infty} |(x,e_k)(y,e_k)| \le \|x\|\|y\|.$$

证明

根据 Cauchy-Schwarz 不等式知

$$\sum_{k=1}^{\infty} |(x,e_k)(y,e_k)| \le (\sum_{k=1}^{\infty} |(x,e_k)|^2)^{\frac{1}{2}} (\sum_{k=1}^{\infty} |(y,e_k)|^2)^{\frac{1}{2}}, \ \forall x, y \in X$$

根据 A 的正交规范性与 Bessel 不等式知

$$(\sum_{k=1}^{\infty} |(x,e_k)|^2)^{\frac{1}{2}} \le \|x\|$$

$$(\sum_{k=1}^{\infty} |(y,e_k)|^2)^{\frac{1}{2}} \le \|y\|$$

综上即得

$$\sum_{k=1}^{\infty} |(x,e_k)(y,e_k)| \le \|x\|\|y\|, \ \forall x, y \in X$$

□

2. 设 H 为 Hilbert 空间, $\{e_k\}, \{e_k'\}$ 是 H 中的两个标准正交系, 并且 $\sum_{k=1}^{\infty} \|e_k - e_k'\|^2 < 1$. 证明如果 $\{e_k\}, \{e_k'\}$ 中之一是完备的, 则另一个也是完备的.

证明

不妨设 $\{e_k\}$ 完备, 则 $\{e_k\}^{\perp} = \{\theta\}$. 现若 $\{e_k'\}$ 不完备, 则存在 $h \in \{e_k'\}^{\perp} \backslash \{\theta\}$, 根据正交补空间的定义知 $(h, e_k') = 0 (\forall k \in \mathbb{N})$. 对 h 和 $\{e_k\}$ 应用 Parseval 等式有:

$$\|h\|^2 = \sum_{k=1}^{\infty} |(h,e_k)|^2 = \sum_{k=1}^{\infty} |(h, e_k - e_k')|^2 \le \sum_{k=1}^{\infty} (\|h\|^2 \|e_k - e_k'\|^2) < \|h\|^2$$

其中第一个不等式是 Cauchy-Schwarz 不等式, 得到矛盾! 故 $\{e_k'\}$ 完备. □

3. 对于内积空间 H, 下述条件等价:

(1) $x \perp y$;

(2) $\|x + \alpha y\| \ge \|x\|, \forall \alpha \in \mathbb{C}$;

(3) $\|x + \alpha y\| = \|x - \alpha y\|, \forall \alpha \in \mathbb{C}$.

证明

(1) $\Rightarrow$ (2): 当 $x \perp y$, 知

$$\begin{aligned}(x+\alpha y, x+\alpha y) &= (x,x) + (x,\alpha y) + (\alpha y, x) + (\alpha y, \alpha y) \\ &= \|x\|^2 + \alpha(x,y) + \overline{\alpha}(x,y) + |\alpha|^2\|y\|^2 \\ &= \|x\|^2 + |\alpha|^2\|y\|^2 \ge \|x\|, \ \forall \alpha \in \mathbb{C}.\end{aligned}$$

(2) $\Rightarrow$ (3): 当 $\|x + \alpha y\| \ge \|x\|$, 知一方面有

$$\|x + \alpha y\| = \|x - \alpha y + 2\alpha y\| \ge \|x - \alpha y\|$$

另一方面

$$\|x-\alpha y\| = \|x+\alpha y-2\alpha y\| \geq \|x+\alpha y\|$$

故只能有

$$\|x+\alpha y\| = \|x-\alpha y\|,\ \forall \alpha \in \mathbb{C}$$

(3) ⇒ (1): 由 (3) 易知

$$\|x+\alpha y\|^2 = \|x-\alpha y\|^2,\ \forall \alpha \in \mathbb{C}$$

写成内积形式即:

$$(x+\alpha y, x+\alpha y) = (x-\alpha y, x-\alpha y)$$

其中对左式有

$$(x+\alpha y, x+\alpha y) = \|x\|^2 + (\alpha+\overline{\alpha})(x,y) + |\alpha|^2\|y\|^2$$

对右式有

$$(x-\alpha y, x-\alpha y) = \|x\|^2 - (\alpha+\overline{\alpha})(x,y) + |\alpha|^2\|y\|^2$$

取用 α 满足 $\alpha+\overline{\alpha} \neq 0$ 即得 $(x,y)=0$, 从而 $x \perp y$. □

第三章 线性算子与线性泛函

3.1 线性算子的概念

3.1.1 知识梳理

3.1.1.1 线性算子和线性泛函的定义

算子是运算的抽象表述, 首先给出一些运算的例子:

(1) 代数运算:

$$x \mapsto Ax, \quad \forall x \in \mathbb{R}^n$$

其中 A 是 $n \times n$ 矩阵.

(2) 求导运算:

$$u(x) \mapsto P(\partial_x)u(x), \quad \forall u \in C^\infty(\overline{\Omega})$$

其中 $P(\cdot)$ 是一个多项式, 而 ∂_x 是偏导数运算.

(3) 积分变换:

$$u(x) \mapsto \int_\Omega K(x,y)u(y)dy \quad \forall u \in C(\overline{\Omega})$$

其中 $K(x,y)$ 是 $\Omega \times \Omega$ 上的可积函数.

线性算子的概念来自线性代数中的线性变换.

定义 3.1.1 (定义域, 值域, 线性算子)

设 $\mathscr{X}, \mathscr{Y}$ 是两个线性空间, D 是 $\mathscr{X}$ 的一个线性子空间, $T: D \to \mathscr{Y}$ 是一种映射, D 称为 T 的定义域, 有时记作 $D(T)$. $R(T) = \{Tx : \forall x \in D\}$ 称为 T 的值域. 如果

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty, \quad \forall x, y \in D, \forall \alpha, \beta \in \mathbb{K}$$

那么称 T 是一个线性算子.

♣

例 3.1 设 $\mathscr{X} = \mathbb{R}^n, \mathscr{Y} = \mathrm{R}^m, T = (t_{ij})_{m\times n}$. 如果

$$x \mapsto Tx = (\sum_{j=1}^n t_{ij}x_j)_{i=1}^m, \quad \forall x = (x_1, \cdots, x_n) \in \mathbb{R}^n$$

那么 T 是一个线性算子.

例 3.2 设 $\mathscr{X} = \mathscr{Y} = C^\infty(\overline{\Omega})$, 又设微分多项式

$$P(\partial_x) = \sum_{|\alpha|\le m} a_\alpha(x)\partial_x^\alpha, \quad a_\alpha(x) \in C^{(\infty)}(\overline{\Omega})$$

如果 $T: u(x) \mapsto P(\partial_x)u(x)(\forall u \in \mathscr{X})$, 那么 T 便是一个 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性算子.

若 $\mathscr{X} = \mathscr{Y} = L^2(\Omega), D(T) = C^m(\overline{\Omega})$, 则上面定义的算子 T 也是线性的.

例 3.3 设 $\mathscr{X} = L^1(-\infty,\infty), \mathscr{Y} = L^\infty(-\infty,\infty)$, 若规定

$$T: u(x) \mapsto \int_{-\infty}^\infty e^{i\xi x}u(x)dx, \quad \forall u \in \mathscr{X}$$

那么 T 是一个 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性算子.

定义 3.1.2 (线性泛函)

取值为实数 (复数) 的线性算子称为实 (复) 线性泛函, 记作 $f(x)$ 或 $\langle f, x\rangle$(即线性函数).

例 3.4 设 $\mathscr{X} = C(\overline{\Omega})$, 若规定

$$f(x) \triangleq \int_{\Omega} x(\xi)d\xi, \quad \forall x \in \mathscr{X}$$

则 f 是一个线性泛函, 但 $x(\xi) \mapsto \int_{\Omega} x^2(\xi)d\xi$ 不是线性泛函.

例 3.5 设 $\mathscr{X} = C^{\infty}(\Omega)$, 若对某个指标 α 及 $\xi_0 \in \Omega$ 规定

$$f(u) = \partial^{\alpha} u(\xi_0), \quad \forall u \in \mathscr{X},$$

则 f 是 $C^{\infty}(\Omega)$ 上的一个线性泛函.

3.1.1.2 线性算子的连续性和有界性

算子的连续性即映射的连续性.

定义 3.1.3 (连续)

设 $\mathscr{X}, \mathscr{Y}$ 是 F^* 空间, $D(T) \subset \mathscr{X}$, 称线性算子 $T : D(T) \to \mathscr{Y}$ 在 $x_0 \in D(T)$ 是连续的, 如果

$$x_n \in D(T), \quad x_n \to x_0 \Rightarrow Tx_n \to Tx_0.$$

注 设 $\mathscr{X}$ 有度量 d, $\mathscr{Y}$ 有度量 $\tilde{d}$, 则上述定义的形式逻辑表述为:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in \mathscr{X} (d(x, x_0) < \delta \Rightarrow \tilde{d}(Tx, Tx_0) < \varepsilon).$$

命题 3.1.1

对于线性算子 T, 为了它在 $D(T)$ 内处处连续, 必须且仅须它在 $x = \theta$ 处连续.

证明

若 T 在 θ 处连续, 则对 $\forall x_n, x_0 \in D(T), x_n \to x_0$ 有:

$$x_n - x_0 \to \theta \Rightarrow Tx_n - Tx_0 = T(x_n - x_0) \to T\theta = \theta.$$

□

定义 3.1.4 (有界)

设 $\mathscr{X}, \mathscr{Y}$ 都是 B^* 空间, 称线性算子 $T : \mathscr{X} \to \mathscr{Y}$ 是有界的, 如果有常数 $M \geq 0$, 使得

$$\|Tx\|_{\mathscr{Y}} \leq M\|x\|_{\mathscr{X}}, \quad \forall x \in \mathscr{X}.$$

命题 3.1.2

设 $\mathscr{X}, \mathscr{Y}$ 都是 B^* 空间, 为了线性算子 T 连续, 必须且仅须 T 有界.

证明

当 T 有界, 只需验证 T 在 θ 处连续. 设 $x_n \to \theta, n \to \infty$, 知

$$\|Tx_n\|_{\mathscr{Y}} \leq M\|x_n\|_{\mathscr{X}} \to 0, \quad n \to \infty$$

故 T 处处连续.

当 T 连续, 若存在 $x_n \in \mathscr{X}$ 使得

$$\|Tx_n\| > n\|x_n\|$$

令 $y_n = \frac{x_n}{n\|x_n\|}$, 知 $\|Ty_n\| > 1$, 但 $y_n \to \theta (n \to \infty)$, 这与 T 的连续性矛盾! □

补充定义 3.1.1 (♡ 有界算子)

设 $(\mathscr{X}, \|\cdot\|_{\mathscr{X}}), (\mathscr{Y}, \|\cdot\|_{\mathscr{Y}})$ 为 B^* 空间, $T : \mathscr{X} \supset D(T) \to \mathscr{Y}$ 是线性算子. 如果 T 把其定义域 $D(T)$ 中的任一有界子集映到 $\mathscr{Y}$ 中的有界子集, 则称 T 为有界算子.

补充命题 3.1.1 (♡ 有界算子定义的等价性)

$T : \mathscr{X} \to \mathscr{Y}$ 是定义(3.1.1)中的有界线性算子当且仅当存在常数 $M \geq 0$ 使得

$$\|Tx\|_{\mathscr{Y}} \leq M\|x\|_{\mathscr{X}}, \quad \forall x \in \mathscr{X}.$$

证明

当 T 是定义(3.1.1)中的有界线性算子, 取

$$S = \{x \in \mathscr{X} : \|x\|_{\mathscr{X}} = 1\}$$

知 $S \subset \mathscr{X}$ 是有界集. 根据有界定义知 $T(S)$ 是 $\mathscr{Y}$ 中的有界集, 从而存在常数 $M > 0$ 使得

$$\|Tx\|_{\mathscr{Y}} \leq M, \quad \forall x \in S$$

现对任意 $x \in \mathscr{X}$, 只要 $x \neq \theta$, 则有 $x' = \frac{x}{\|x\|_{\mathscr{X}}} \in S$, 得到

$$\|Tx'\|_{\mathscr{Y}} = \|T\frac{x}{\|x\|_{\mathscr{X}}}\|_{\mathscr{Y}} = \frac{\|Tx\|_{\mathscr{Y}}}{\|x\|_{\mathscr{X}}} \leq M$$

亦即 $\|Tx\|_{\mathscr{Y}} \leq M\|x\|_{\mathscr{X}}$, 此即欲证.

当存在常数 $M > 0$ 使得

$$\|Tx\|_{\mathscr{Y}} \leq M\|x\|_{\mathscr{X}}, \quad \forall x \in \mathscr{X}.$$

则对 $\mathscr{X}$ 中的任意有界子集 A, 设存在 $r > 0$ 使得 $A \subset B(\theta, r)$, 则

$$\|Tx\|_{\mathscr{Y}} \leq M\|x\|_{\mathscr{X}} \leq Mr, \quad \forall x \in A$$

这说明 $T(A) \subset B(\theta, Mr)$, 此即欲证. □

定义 3.1.5 (有界线性算子空间, 有界线性泛函空间)

用 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 表示一切由 $\mathscr{X}$ 到 $\mathscr{Y}$ 的有界线性算子的全体, 并规定

$$\|T\| = \sup_{x \in \mathscr{X} \setminus \theta} \|Tx\|/\|x\| = \sup_{\|x\|=1} \|Tx\|$$

为 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 的范数. 特别用 $\mathscr{L}(\mathscr{X})$ 表示 $\mathscr{L}(\mathscr{X}, \mathscr{X})$, 用 $\mathscr{X}^*$ 表示 $\mathscr{L}(\mathscr{X}, \mathbb{K})$, 即 $\mathscr{X}^*$ 表示 $\mathscr{X}$ 上的有界线性泛函全体.

定理 3.1.1

设 $\mathscr{X}$ 是 B^* 空间, $\mathscr{Y}$ 是 B 空间, 若在 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 上规定线性运算:

$$(\alpha_1 T_1 + \alpha_2 T_2)(x) = \alpha_1 T_1 x + \alpha_2 T_2 x, \quad \forall x \in \mathscr{X}$$

其中 $\alpha_1, \alpha_2 \in \mathscr{K}, T_1, T_2 \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 则 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 按 $\|T\|$ 构成一个 Banach 空间.

证明

显见 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 是一个线性空间, 下证 $\|T\|$ 是范数.

$$\|T\| \geq 0, \|T\| = 0 \Leftrightarrow (Tx = \theta(\forall x \in \mathscr{X}) \Leftrightarrow T = \theta)$$

$$\|T_1 + T_1\| = \sup_{\|x\|=1} \|T_1 x + T_2 x\| \leq \sup_{\|x\|=1} \|T_1 x\| + \sup_{\|x\|=1} \|T_2 x\| = \|T_1\| + \|T_2\|$$

$$\|\alpha T\| = \sup_{\|x\|=1} \|\alpha T x\| = |\alpha| \sup_{\|x\|=1} \|Tx\| = |\alpha|\|T\|.$$

再证完备性, 设 $\{T_n\}_1^\infty$ 是基本列, 则

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) \forall x \in \mathscr{X} (\|T_{n+p}x - T_n x\| \leq \varepsilon \|x\|, \quad \forall n < N, \forall p \in \mathbb{N}).$$

故根据 $\mathscr{Y}$ 作为 Banach 空间的完备性知 $T_n x \to y \in \mathscr{Y} (n \to \infty)$, 记 $y = Tx$, 要证明 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. 显见 T 线性, 进而只需证明其有界. 事实上, $\exists n \in \mathbb{N}$ 使得

$$\|Tx\| = \|y\| \leq \|T_n x\| + 1 \leq (\|T_n\| + 1)\|x\|, \forall x \in \mathscr{X}, \|x\| = 1$$

可得 $\|T\| \leq \|T_n\| + 1$. □

注 当 $\mathscr{X}$ 是赋范线性空间, $A, B \in \mathscr{L}(\mathscr{X})$ 时, AB 也是有界线性算子, 且 $\|AB\| \leq \|A\| \cdot \|B\|$. 这是因为任取 $x \in \mathscr{X}$ 有:

$$\|ABx\| \leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\|$$

上述不等式作为一个核心条件被用于 Banach 代数的定义中, 为此先回顾代数的定义:

补充定义 3.1.2 (AHCB 代数, 含幺代数, 交换代数)

称线性空间 $\mathscr{X}$ 为一个代数, 如果在 $\mathscr{X}$ 中引进乘法运算 $\cdot$, 其满足:

(1) $\forall x, y, z \in \mathscr{X} ((x \cdot y) \cdot z = x \cdot (y \cdot z))$;

(2) $\forall x, y, z \in \mathscr{X} ((x \cdot (y + z) = x \cdot y + x \cdot z) \wedge ((y + z) \cdot x = y \cdot x + z \cdot x))$;

(3) $\forall \alpha \in \mathbb{K} \forall x, y \in \mathscr{X} (\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y))$.

如果存在元素 $e \in \mathscr{X}$, 使得对任意 $x \in \mathscr{X}$ 均有

$$e \cdot x = x \cdot e = x$$

则称 e 为代数 $\mathscr{X}$ 的幺元, $\mathscr{X}$ 本身称为一个含幺代数.

如果 $\mathscr{X}$ 上的乘法运算 $\cdot$ 是交换的, 亦即

$$\forall x, y \in \mathscr{X} (x \cdot y = y \cdot x)$$

就称代数 $\mathscr{X}$ 为交换代数. ♣

补充定义 3.1.3 (AHCB 赋范代数, Banach 代数)

设 $\mathscr{X}$ 是赋范线性空间, 如果它是具有单位元的代数, 同时满足:

(1) $\|e\| = 1$;

(2) $\|x \cdot y\| \leq \|x\| \cdot \|y\|$,

就称 $\mathscr{X}$ 是赋范代数. 如果赋范代数 $\mathscr{X}$ 完备 (亦即 $\mathscr{X}$ 还是 Banach 空间), 就称 $\mathscr{X}$ 为 Banach 代数. ♣

补充命题 3.1.2 (⋆)

若 $T : \mathbb{R}^n \to \mathbb{R}^m$ 是线性算子, 则 T 可用矩阵表示. ♠

证明

设 $\{e_1, \cdots, e_n\}$ 是 $\mathbb{R}^n$ 的一组基, 现任取 $j \in \{1, \cdots, n\}$, 设 $Te_j = (t_{1j}, \cdots, t_{mj}) = \sum\limits_{i=1}^m t_{ij}\tilde{e}_i$. 知对任意的 $x \in \mathbb{R}^n$, 设 $x = \sum\limits_{j=1}^n x_j e_j$, 则有

$$Tx = \sum_{j=1}^n x_j T e_j = \sum_{j=1}^n x_j (\sum_{i=1}^m t_{ij}\tilde{e}_i) = \sum_{i=1}^m (\sum_{j=1}^n t_{ij} x_j)\tilde{e}_j$$

写成坐标形式即

$$Tx = (\sum_{j=1}^n t_{ij} x_j)_{i=1}^m$$

亦即 T 完全由矩阵 $(t_{ij})_{m \times n}$ 确定. □

例 3.6(矩阵算子) 设 T 是有穷维 B^* 空间 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性映射, 则 T 必连续.

证明

有穷维 B^* 空间之间的线性算子可通过矩阵表示, 设 $T=(t_{ij})$. 因为有穷维空间之间, 维数相等则任意范数等价, 故不妨设 $\mathscr{X}=\mathbb{K}^n, \mathscr{Y}=\mathbb{K}^m$, 得到

$$\|Tx\|=(\sum_{i=1}^{m}|\sum_{j=1}^{n}t_{ij}x_j|^2)^{\frac{1}{2}}\leq(\sum_{i=1}^{m}\sum_{j=1}^{n}|t_{ij}|^2\cdot\sum_{j=1}^{n}|x_j|^2)^{\frac{1}{2}}=(\sum_{i=1}^{m}\sum_{j=1}^{n}|t_{ij}|^2)^{\frac{1}{2}}\|x\|$$

得到

$$\|Tx\|\leq(\sum_{i=1}^{m}\sum_{j=1}^{n}|t_{ij}|^2)^{\frac{1}{2}}\|x\|$$

因而 T 有界, 进而其连续. □

补充定理 3.1.1 (♡ 基本引理)

给定数域 $\mathbb{K}$ 上的线性赋范空间 $(\mathscr{X},\|\cdot\|)$. 若 $x_1,\cdots,x_n\in\mathscr{X}$ 线性无关, 则存在常数 $\gamma>0$ 使得对一切 $(\alpha_1,\cdots,\alpha_n)\in\mathbb{K}^n$ 有

$$\|\sum_{j=1}^{n}\alpha_jx_j\|\geq\gamma(\sum_{j=1}^{n}|\alpha_j|^2)^{\frac{1}{2}}$$

♡

基本引理恰是有限维 B^* 空间 $\mathrm{span}\{x_1,\cdots,x_n\}$ 与 $\mathbb{K}^n$ 之间范数等价的体现.

例 3.7(♡ 有限维线性赋范空间上的线性算子) 给定 $(\mathscr{X},\|\cdot\|_{\mathscr{X}}),(\mathscr{Y},\|\cdot\|_{\mathscr{Y}}),x_1,\cdots,x_n$ 是 $\mathscr{X}$ 的一组基, $T:\mathscr{X}\to\mathscr{Y}$ 是线性算子. 设 $D(T)=\mathscr{X}$, 令 $y_i=Tx_i(i=1,\cdots,n)$, 则 $R(T)$ 是由 $\{y_1,\cdots,y_n\}$ 生成的 $\mathscr{Y}$ 的线性子空间.

下面来说明 T 是有界算子, 任取 $x\in\mathscr{X}$, 设 $x=\sum\limits_{i=1}^{n}\alpha_ix_i$, 因而 $Tx=\sum\limits_{i=1}^{n}\alpha_iTx_i=\sum\limits_{i=1}^{n}\alpha_iy_i$. 若记 $|\alpha|=(\sum\limits_{i=1}^{n}|\alpha_i|^2)^{\frac{1}{2}},(\sum\limits_{i=1}^{n}\|y_i\|^2)^{\frac{1}{2}}=M$[1], 则根据 Cauchy-Schwarz 不等式有

$$\|Tx\|_{\mathscr{Y}}\leq\sum_{i=1}^{n}|\alpha_i|\|y_i\|_{\mathscr{Y}}\leq(\sum_{i=1}^{n}\|y_i\|^2)^{\frac{1}{2}}(\sum_{i=1}^{n}|\alpha_i|^2)^{\frac{1}{2}}=M|\alpha|$$

设 $\|x\|_{\mathscr{X}}=\|\sum\limits_{i=1}^{n}\alpha_ix_i\|=1$, 由基本引理知 $|\alpha|\leq\gamma^{-1}$, 其中 γ 只依赖于 $x_1,\cdots,x_n$, 从而

$$\sup_{\|x\|_{\mathscr{X}}=1}\|Tx\|_{\mathscr{Y}}\leq M\gamma^{-1}$$

亦即 T 是有界算子.

下面主要考察 $\mathscr{X}=\mathscr{Y}=\mathbb{R}^n$ 时 T 的范数. 设 $l^p(1\leq p\leq\infty)$ 范数:

$$\|x\|_{l^p}=\begin{cases}(\sum\limits_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}, & 1\leq p<\infty\\ \max\limits_{1\leq i\leq n}|x_i|, & p=\infty\end{cases}$$

并设 T 是 $(\mathscr{X},l^p)$ 到 $(\mathscr{Y},l^q)$ 的线性算子. 当 $p=q=2$ 时, 设 T 可表为矩阵 $A=(a_{ij})_{1\leq j,k\leq n}$, 下面来求解 $\|T\|$.

(a) 当 A 是对称阵时, 考虑

$$\det(A-\lambda E)=0$$

设该方程有 n 个实根 $\lambda_1,\cdots,\lambda_n$, 并记

$$\Lambda=\max_{1\leq i\leq n}|\lambda_i|=|\lambda_{i_0}|$$

任取 $x=(x_1,\cdots,x_n),y=(y_1,\cdots,y_n)$, 令 $\langle x,y\rangle=x_1y_1+\cdots+x_ny_n$, 则 $|x|^2=\langle x,x\rangle$, 下面证明 $\|T\|=\sup\limits_{\|x\|=1}|Ax|=$

[1]注意这里 y_i 只与 x_i 有关, 而 x_i 是一开始选定的, 所以 y_i 也是固定的.

Λ. 注意对任意 λ_i, 总存在 $\xi^{(i)} \in \mathbb{R}^n$, 使得 $\|\xi^{(i)}\| = 1$ 且 $A\xi^{(i)} = \lambda_i \xi^{(i)}$. 根据矩阵特征向量之间的正交性, 知

$$\langle \xi^{(i)}, \xi^{(j)} \rangle = 0, \quad \forall i \neq j$$

现有

$$\|T\| \geq (\langle A\xi^{(i_0)}, A\xi^{(i_0)} \rangle)^{\frac{1}{2}} = (\langle \lambda_{i_0}\xi^{(i_0)}, \lambda_{i_0}\xi^{(i_0)} \rangle)^{\frac{1}{2}} = |\lambda_{i_0}||\xi^{(i_0)}| = |\lambda_{i_0}| = \Lambda$$

另一方面, 任取 $x \in \mathbb{R}^n, \|x\| = 1$, 知

$$x = \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle \xi^{(i)}$$

根据 Parseval 等式知

$$\sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle^2 = |x|^2 = 1$$

且根据 $\xi^{(i)}$ 的构造有

$$Ax = \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle A\xi^{(i)} = \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle \lambda_i \xi^{(i)}$$

进而由 $|\lambda_i| \leq \Lambda (i = 1, \cdots, n)$ 知:

$$\begin{aligned}|Ax|^2 = \langle Ax, Ax \rangle &= \langle \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle \lambda_i \xi^{(i)}, \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle \lambda_i \xi^{(i)} \rangle \\ &= \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle^2 \lambda_i^2 \leq \Lambda^2 \sum_{i=1}^{n} \langle x, \xi^{(i)} \rangle^2 = \Lambda^2 |x|^2 = \Lambda^2\end{aligned}$$

得到 $\|T\| = \sup\limits_{|x|=1} |Ax| \leq \Lambda$, 综合即得 $\|T\| = \Lambda$.

(b) 一般情况下, 设 T 可由矩阵 $A = (a_{jk})_{n\times n}$ 表示, 记 A 的转置为 A^T, 显见 A^TA 是 $n \times n$ 对称阵. 设 λ_1 是 A^TA 的最大特征根, 下面说明 $\|T\| = \sup\limits_{|x|=1} \langle Ax, Ax \rangle^{\frac{1}{2}} = \sqrt{\lambda_1}$.

设 $\lambda_1, \cdots, \lambda_n$ 是 A^TA 的特征根, $\eta_i (1 \leq i \leq n)$ 是对应于 λ_i 的单位特征向量, 亦即

$$|\eta_i| = 1, \langle \eta_i, \eta_j \rangle = 0, \quad \forall i \neq j$$

则

$$\|T\| \geq \langle A\eta_1, A\eta_1 \rangle^{\frac{1}{2}} = \langle A^TA\eta_1, \eta_1 \rangle^{\frac{1}{2}} = \langle \lambda_1\eta_1, \eta_1 \rangle^{\frac{1}{2}} = \sqrt{\lambda_1}$$

另一方面, 任取 $x \in \mathbb{R}^n, |x| = 1$, 知

$$x = \sum_{i=1}^{n} \langle x, \eta_i \rangle \eta_i, \quad |x|^2 = \sum_{i=1}^{n} \langle x, \eta_i \rangle^2 = 1$$

那么

$$\begin{aligned}|Ax|^2 = \langle Ax, Ax \rangle = \langle A^TAx, x \rangle &= \langle A^TA \sum_{i=1}^{n} \langle x, \eta_i \rangle \eta_i, \sum_{i=1}^{n} \langle x, \eta_i \rangle \eta_i \rangle \\ &= \langle \sum_{i=1}^{n} \langle x, \eta_i \rangle A^TA\eta_i, \sum_{i=1}^{n} \langle x, \eta_i \rangle \eta_i \rangle = \langle \sum_{i=1}^{n} \langle x, \eta_i \rangle \lambda_i \eta_i, \sum_{i=1}^{n} \langle x, \eta_i \rangle \eta_i \rangle \\ &= \sum_{i=1}^{n} \lambda_i |\eta_i|^2 \leq \lambda_1 \sum_{i=1}^{n} |\eta_i|^2 = \lambda_1\end{aligned}$$

这说明

$$\|T\| \leq \sup_{|x|=1} |Ax| \leq \sqrt{\lambda_1}$$

综上即得 $\|T\| = \sqrt{\lambda_1}$.

例 3.8(正交投影算子) 设 M 是 $\mathscr{X}$ 的一个闭线性子空间, 根据正交分解定理(2.6.4)知对任意的 $x \in \mathscr{X}$, 存在唯一的

分解

$$x = y + z$$

其中 $y \in M, z \in M^{\perp}$, 对应 $x \mapsto y$ 称作由 $\mathscr{X}$ 到 M 的正交投影算子, 记作 P_M. 在不强调子空间 M 时, 省略 M 简记为 P, 下面证明 P 是一个连续线性算子, 且若 $M \neq \{\theta\}$, 则 $\|P\| = 1$.

证明

首先证明线性性. 设 $x_i = Px_i + z_i (i = 1, 2), z_i \in M^{\perp}$, 则

$$\alpha_1 x_1 + \alpha_2 x_2 = (\alpha_1 Px_1 + \alpha_2 P(x_2)) + (\alpha_1 z_1 + \alpha_2 z_2), \quad \forall \alpha_1, \alpha_2 \in \mathbb{K}$$

因为 $\alpha_1 z_1 + \alpha_2 z_2 \in M^{\perp}, \alpha_1 Px_1 + \alpha_2 Px_2 \in M$, 故

$$P(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Px_1 + \alpha_2 Px_2$$

线性性得证.

再证连续性, 注意

$$\|Px\|^2 = \|x\|^2 - \|z\|^2 \leq \|x\|^2$$

故 $\|Px\| \leq \|x\|$, 得到 $\|P\| \leq 1$, 故 P 有界, 因而连续.

最后, 当 $M \neq \{\theta\}$, 任取 $x \in M \backslash \{\theta\}$, 知 $\|Px\| = \|x\|$, 因而 $\|P\| = 1$.

例 3.9(♡ 积分核算子) 设 $K \in C[a, b]^2$, 亦即 $K(s, t)$ 在 $[a, b] \times [a, b]$ 上连续, 设

$$T : C[a, b] \to C[a, b], x(t) \mapsto y(s) = \int_a^b K(s, t)x(t)dt, \quad s \in [a, b]$$

则显见 T 是线性算子, 下面计算 T 的范数.

任取 $x \in C[a, b], \|x\| = \sup\limits_{[a,b]} |x(t)| = 1$, 则知

$$\begin{aligned} \|Tx\| &= \sup_{s \in [a,b]} \left| \int_a^b K(s, t)x(t)dt \right| \\ &\leq \sup_{s \in [a,b]} \int_a^b |K(s, t)||x(t)|dt \\ &\leq \sup_{s \in [a,b]} \int_a^b |K(s, t)|dt =: M \end{aligned}$$

因为 $K(s, t)$ 作为 s 的函数在 $[a, b]$ 上连续, 故 $\int_a^b |K(s, t)|dt$ 也是关于 s 在 $[a, b]$ 上的连续函数. 由闭区间上连续函数的性质可知存在 $s_0 \in [a, b]$ 使得

$$\int_a^b |K(s_0, t)|dt = M$$

现在希望找到 $C[a, b]$ 中的函数序列 $\{x_n\}$ 使得 $\|x_n\| = 1 (n = 1, 2, \cdots)$, 且 $\|Tx_n\| \to M (n \to \infty)$. 记 $\operatorname{sgn} K(s_0, t) = x_0(t)$, 令

$$y_0(s) = \int_a^b K(s, t)x_0(t)dt$$

则

$$y_0(s_0) = \int_a^b |K(s_0, t)|dt = M$$

显见 $y_0 \in C[a, b]$, 这便只需要说明存在 $\{x_n\} \subset C[a, b]$ 使得 $Tx_n \to y_0 (n \to \infty)$ 即可. 事实上, 根据 Luzin 定理, 对任意 $\varepsilon > 0$ 均存在 $x_\varepsilon \in C[a, b]$ 满足

(a) $\forall t \in [a, b] (|x_\varepsilon(t)| \leq 1)$;

(b) $m\{t \in [a, b] : x_\varepsilon(t) \neq x_0(t)\} < \varepsilon$

记

$$y_\varepsilon(s) = Tx_\varepsilon(s) = \int_a^b K(s,t)x_\varepsilon(t)dt$$

则

$$|y_\varepsilon(s_0) - y_0(s_0)| = |\int_a^b K(s_0,t)(x_\varepsilon(t) - x_0(t))dt| \le \int_a^b |K(s_0,t)||x_\varepsilon(t) - x_0(t)|dt$$

其中根据 x_ε 与 x_0 的构造有

$$|x_\varepsilon(t) - x_0(t)| \le \begin{cases} 2, & x_\varepsilon(t) \ne x_0(t) \\ 0, & x_\varepsilon(t) = x_0(t) \end{cases}$$

记 $C = \max\limits_{(s,t)\in[a,b]^2} |K(s,t)|$, 则

$$\int_a^b |K(s_0,t)||x_\varepsilon(t) - x_0(t)|dt \le C \cdot 2\{t \in [a,b] : x_\varepsilon(t) \ne x_0(t)\} < 2C\varepsilon$$

根据 ε 的任意性即知 $\|Tx_\varepsilon\| \to M(\varepsilon \to 0^+)$, 故 $\|T\| = M$.

另外, 可以考虑 $L^1[a,b]$ 上的积分核算子, 亦即设 $x \in L[a,b]$, 取

$$T : L[a,b] \to L[a,b], x(s) \mapsto y(s) = \int_a^b K(s,t)x(t)dt, \quad s \in [a,b]$$

则若设 $y = Tx$, 有

$$\begin{aligned} \|y\|_{L^1} &= \int_a^b |\int_a^b K(s,t)x(t)dt|ds \\ &\le \int_a^b \int_a^b |K(s,t)||x(t)|dtds \\ &= \int_a^b |x(t)|(\int_a^b |K(s,t)|ds)dt \end{aligned}$$

现记 $M' = \max\limits_{t\in[a,b]} \int_a^b |K(s,t)|ds$, 则

$$\int_a^b |x(t)|(\int_a^b |K(s,t)|ds)dt \le M' \int_a^b |x(t)|dt = M'\|x\|_{L^1}$$

从而

$$\|Tx\|_{L^1} \le M'\|x\|_{L^1} \Rightarrow \|T\| \le M'$$

下面再说明 $\|T\| \ge M'$, 这需要用到 Δ 型函数族的思想, 关于这个函数族的具体介绍可参见 **[Zo]** 或 **[ST1]**. 根据 $\int_a^b |K(s,t)|ds$ 关于 t 的连续性, 知存在 $t_0 \in [a,b]$ 使得 $\int_a^b |K(s,t_0)|ds = M' = \max\limits_{t\in[a,b]} \int_a^b |K(s,t)|ds$. 目前的动机在于找到一族函数 $x_\delta(t) \in L[a,b]$, 使得 $\|x_\delta\|_{L^1} = 1$ 的同时有

$$\int_a^b |x_\delta(t)|(\int_a^b |K(s,t)|ds)dt \to \int_a^b |K(s,t_0)|ds, \quad \delta \to 0$$

根据 $K(s,t)$ 关于 t 的连续性知:

$$\forall\varepsilon > 0 \exists\delta > 0 \forall t \in (t_0 - \delta, t_0 + \delta)(|K(s,t) - K(s,t_0)| < \varepsilon)$$

现取

$$x_\delta(t) = \begin{cases} \frac{1}{2\delta}, & t \in (t_0 - \delta, t_0 + \delta) \\ 0, & t \notin (t_0 - \delta, t_0 + \delta) \end{cases}$$

有

$$\begin{aligned}
\int_a^b |x_\delta(t)|(\int_a^b |K(s,t)|ds)dt &= \frac{1}{2\delta}\int_{t_0-\delta}^{t_0+\delta}(\int_a^b |K(s,t)|ds)dt \\
&= \frac{1}{2\delta}\int_{t_0-\delta}^{t_0+\delta}(\int_a^b |K(s,t)-K(s,t_0)+K(s,t_0)|ds)dt \\
&\geq \frac{1}{2\delta}\int_{t_0-\delta}^{t_0+\delta}(\int_a^b |K(s,t_0)|ds - \int_a^b |K(s,t)-K(s,t_0)|ds)dt \\
&= M' - (b-a)\varepsilon
\end{aligned}$$

由 ε 的任意性即知 $\|T\| \geq M'$, 因而 $\|T\| = M'$.

例 3.10(♡ 有界序列乘积算子) 设 $\mathscr{X} = \mathscr{Y} = l^\infty$, $a = \{a_n\} \in l^\infty$ 是固定元素, 设

$$T_a : \mathscr{X} \to \mathscr{Y}, \quad x = (x_1, x_2, \cdots) \mapsto (a_1x_1, a_2x_2, \cdots)$$

则 T_a 是有界线性算子, 且

$$\|T_a\| = \|a\|_\infty.$$

证明

显见 T_a 是线性算子, 下面证明有界性. 任取 $x = (x_1, x_2, \cdots)$, 知 $|a_i| \leq \|a\|_\infty (i = 1, 2, \cdots)$, 有

$$\|T_a x\|_\infty = \sup_{i\geq 1} |a_i x_i| \leq \|a\|_\infty \sup_{i\geq 1} |x_i| = \|a\|_\infty \|x\|_\infty$$

这说明 T_a 是有界线性算子, 且 $\|T_a\| \leq \|a\|_\infty$. 若存在 i_0 使得 $|a_{i_0}| = \|a\|_\infty$, 取用 e_{i_0} 即得命题. 现若不存在 i 使得 $|a_i| = \|a\|_\infty$, 则根据上确界的定义有

$$\forall \varepsilon > 0 \exists N \in \mathbb{N}(|a_N| \geq \|a\|_\infty - \varepsilon)$$

设 e_N 为 l^∞ 中第 N 个坐标为 1, 其余坐标为 0 的向量, 显见 $\|e_N\|_\infty = 1$, 有

$$\|Ta\| \geq \|Tae_N\|_\infty = |a_N| \geq \|a\|_\infty - \varepsilon$$

根据 ε 的任意性有 $\|Ta\| \geq \|a\|_\infty$, 因而 $\|Ta\| = \|a\|_\infty$.

例 3.11(♡ 无界线性算子) 取 $\mathscr{X} = P[0,1]$ 是 $[0,1]$ 上的多项式全体, $\|\cdot\|_{\mathscr{X}} = \max\limits_{[0,1]} |\cdot|$, 取

$$T : \mathscr{X} \to \mathscr{X}, x \mapsto (Tx)(t) = x'(t), \quad \forall t \in [0,1], x \in \mathscr{X}$$

则显见 T 是线性算子, 但 T 无界, 这是因为取 $x_n(t) = t^n (n = 1, 2, \cdots)$, 知 $\|x_n\| = 1$, 而 $Tx_n(t) = nt^{n-1}$, 这说明 $\|Tx_n\| = n \Rightarrow \|T\| \geq n (\forall n \in \mathbb{N})$, 故 T 无界.

注 [WL] 与高志强老师的课堂上都介绍了几种不同的算子收敛性, 补充于此. 设 $(\mathscr{X}, \|\cdot\|_{\mathscr{X}}), (\mathscr{Y}, \|\cdot\|_{\mathscr{Y}})$ 均为赋范线性空间, $T_n, T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$.

补充定义 3.1.4 (♡ 一致收敛 (依范数收敛)[WL])

若 $\|T_n - T\| \to 0 (n \to \infty)$, 则称 $\{T_n\}$ 一致收敛于 T. ♣

补充定义 3.1.5 (♡ 点态收敛 (强收敛)[WL])

若 $\forall x \in \mathscr{X} (\|T_n x - Tx\| \to 0 (n \to \infty))$, 则称 $\{T_n\}$ 点态收敛于 T. ♣

补充定义 3.1.6 (♡ 弱收敛[WL])

若 $\forall x \in \mathscr{X} \forall f \in \mathscr{X}^* (f(T_n x) \to f(Tx))$, 则称 $\{T_n\}$ 弱收敛于 T. ♣

补充命题 3.1.3 (收敛强弱关系WL)

一致收敛 $\Rightarrow$ 点态收敛 $\Rightarrow$ 弱收敛.

例 3.12(♡ 点态收敛但不一致收敛的例子) 设 $\mathscr{X}=\mathscr{Y}=l^p(1\le p<\infty)$, 任取 $x=(\xi_1,\xi_2,\cdots)\in l^p$, 记 $x_n=(\xi_n,\xi_{n+1},\cdots)$, 令 $T_nx=x_n(n=1,2,\cdots)$, 则 T_n 是 l^p 到 l^p 的线性算子. 显见 $\|T_nx\|=\|x_n\|\le\|x\|$, 故 $\|T_n\|\le 1$, 亦即 $\{T_n\}\subset\mathscr{L}(l^p)$. 任取 $x\in l^p$, 知

$$\|T_nx\|=\|x_n\|=(\sum_{k=n}^{\infty}|\xi_k|^p)^{\frac{1}{p}}\to 0,\quad n\to\infty$$

这说明 T_n 点态收敛到零算子 $\theta(\forall x\in\mathscr{X}(\theta(x)=0))$. 但 T_n 并不一致收敛到零算子. 事实上, 知 $T_ne_n=e_1$, 这说明 $\|T_n\|\ge\|T_ne_n\|=1$, 从而 $\|T_n\|=1\ne 0$.

3.1.2 习题

★ 本节各题中, $\mathscr{X},\mathscr{Y}$ 均为 Banach 空间.

练习 3.1 求证: $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$ 的充要条件是 T 为线性算子, 并将 $\mathscr{X}$ 中的有界集映为 $\mathscr{Y}$ 中的有界集.

证明

当 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 线性性显见. 根据有界的定义知存在 $M>0$, 使得

$$\|Tx\|_{\mathscr{Y}}\le M\|x\|_{\mathscr{X}}$$

故若 $\|x\|_{\mathscr{X}}$ 有界, 必有 $\|Tx\|_{\mathscr{Y}}$ 有界, 也即 T 将有界集映成有界集.

当 T 为线性算子, 且将有界集映为有界集, 知

$$\|T\|=\sup_{\|x\|=1}\|Tx\|<\infty$$

这说明 T 有界, 故 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$.

练习 3.2 设 $A\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 求证:

(1) $\|A\|=\sup\limits_{\|x\|\le 1}\|Ax\|$;

证明

因为 $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|$, 而显见

$$\sup_{\|x\|=1}\|Ax\|\le\sup_{\|x\|\le 1}\|Ax\|$$

故 $\|A\|\le\sup\limits_{\|x\|\le 1}\|Ax\|$. 又由 Cauchy-Schwarz 不等式有

$$\|Ax\|\le\|A\|\cdot\|x\|\le\|A\|,\quad\|x\|\le 1$$

故

$$\sup_{\|x\|\le 1}\|Ax\|\le\|A\|$$

综上 $\|A\|=\sup\limits_{\|x\|\le 1}\|Ax\|$.

(2) $\|A\|=\sup\limits_{\|x\|<1}\|Ax\|$.

证明

既然 $\{x\in\mathscr{X}:\|x\|=1\}$ 是 $\mathscr{X}$ 的闭凸子集, 而 $\mathscr{X},\mathscr{Y}$ 均为 Banach 空间, 故 $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|$ 是可取到的, 不

妨设 $\|A\| = \|Ax_0\|$. 取点列 $\{x_n\}, x_n = (1-\frac{1}{n})x_0$, 知 $\{x_n\} \subset \{x \in \mathscr{X} : \|x\| < 1\}, x_n \to x_0, n \to \infty$. 知

$$\sup_{\|x\|<1} \|Ax\| \geq (1-\frac{1}{n})\|Ax_0\| = (1-\frac{1}{n})\|A\|, \quad \forall n \in \mathbb{N}$$

令 $n \to \infty$ 知

$$\sup_{\|x\|<1} \|Ax\| \geq \|A\|$$

又根据 $\|A\| = \sup\limits_{\|x\|\leq 1} \|Ax\|$ 知

$$\sup_{\|x\|<1} \|Ax\| \leq \|A\|$$

即得

$$\sup_{\|x\|<1} \|Ax\| = \|A\|$$

练习 3.3 设 $f \in \mathscr{L}(\mathscr{X}, \mathbb{R})$, 求证:

(1) $\|f\| = \sup\limits_{\|x\|=1} f(x)$;

证明

知

$$\|f\| = \sup_{\|x\|_{\mathscr{X}}=1} \|f(x)\|_{\mathbb{R}} = \sup_{\|x\|=1} |f(x)| \geq \sup_{\|x\|=1} f(x)$$

又根据 f 的线性性:

$$|f(x)| = \operatorname{sgn}(f(x)) \cdot f(x) = f(\operatorname{sgn}(f(x)) \cdot x) \leq \sup_{\|\operatorname{sgn} f(x) \cdot x\|=1} f(x) = \sup_{\|x\|=1} f(x)$$

得到

$$\sup_{\|x\|=1} |f(x)| \leq \sup_{\|x\|=1} f(x)$$

综上有

$$\|f\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|=1} f(x)$$

(2) $\sup\limits_{\|x\|<\delta} f(x) = \delta\|f\| (\forall \delta > 0)$.

证明

根据练习(3.2)与前述结论知

$$\sup_{\|x\|=1} f(x) = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|<1} |f(x)|$$

而对任意的 $\delta > 0$, 知

$$\sup_{\|x\|<\delta} f(x) = \sup_{\|x\|<1} f(\delta x) = \delta \sup_{\|x\|<1} f(x) = \delta\|f\|$$

联立即得

$$\sup_{\|x\|<\delta} f(x) = \delta\|f\|.$$

练习 3.4♡ 设 $y(t) \in C[0,1]$, 定义 $C[0,1]$ 上的泛函

$$f(x) = \int_0^1 x(t)y(t)dt \quad (\forall x \in C[0,1])$$

求 $\|f\|$.

解

取 $f(t) = \operatorname{sgn}\phi(t)$, 因为 $[0,1]$ 是 $\mathbb{R}$ 上的紧集, 故根据 Luzin 定理知存在连续函数列 $f_\varepsilon(t)$ 使得

$$f_\varepsilon(t) \to f(t)(\forall t \in [0,1]), m(\{t : f_\varepsilon(t) \neq f(t)\}) \to 0, \quad \varepsilon \to 0^+$$

因此

$$\begin{aligned}|\Phi(f_\varepsilon)| &= |\int_0^1 \phi(t)f_\varepsilon(t)dt| \\ &\geq \int_0^1 |\phi(t)|\cdot|f(t)|dt - \int_0^1 |\phi(t)|\cdot|f_\varepsilon(t)-f(t)|dt \\ &\geq \int_0^1 |\phi(t)|dt - m(\{t : f_\varepsilon(t)\neq f(t)\})\max_{t\in[0,1]}|\phi(t)| \to \|\phi\|_{L^1}, \quad \varepsilon \to 0^+\end{aligned}$$

另一方面容易说明 $|\Phi(f)| \leq \|\phi\|_{L^1}\|f\|$, 因而 $\|\Phi\| = \|\phi\|_{L^1}$. □

练习 3.5♡ 设 f 是 $\mathscr{X}$ 上的非零有界线性泛函, 令

$$d = \inf\{\|x\| : f(x) = 1, x \in \mathscr{X}\}$$

求证: $\|f\| = \frac{1}{d}$.

证明

先说明 $\|f\| \geq \frac{1}{d}$, 这是因为

$$\|f\| = \sup_{x\in\mathscr{X}\setminus\{\theta\}} \frac{\|f(x)\|}{\|x\|} \geq \sup_{\substack{x\in\mathscr{X}\setminus\{\theta\}\\ |f(x)|=1}} \frac{|f(x)|}{\|x\|} = \frac{1}{\inf\{\|x\| : f(x)=1, x\in\mathscr{X}\}} = \frac{1}{d}$$

再考虑 $\|f\| \leq \frac{1}{d}$, 根据 $\|f\|$ 作为上确界的定义知, 对任意的 $\varepsilon > 0$, 存在 $x_\varepsilon \in \mathscr{X}$ 使得

$$\frac{|f(x_\varepsilon)|}{\|x_\varepsilon\|} \geq \|f\| - \varepsilon$$

既然 f 非零, 得到

$$\|\frac{x_\varepsilon}{f(x_\varepsilon)}\| = \frac{\|x_\varepsilon\|}{|f(x_\varepsilon)|} \leq \frac{1}{\|f\|-\varepsilon}$$

又因为

$$f(\frac{x_\varepsilon}{f(x_\varepsilon)}) = 1$$

故根据 d 的定义有

$$\|\frac{x_\varepsilon}{f(x_\varepsilon)}\| \geq d \Rightarrow d \leq \frac{1}{\|f\|-\varepsilon}, \forall\varepsilon > 0 \Rightarrow \|f\| \leq \frac{1}{d}$$

综上即得 $\|f\| = \frac{1}{d}$.

练习 3.6♡ 设 $f \in \mathscr{X}^*$, 求证: $\forall\varepsilon > 0, \exists x_0 \in \mathscr{X}$, 使得 $f(x_0) = \|f\|$, 且 $\|x_0\| < 1+\varepsilon$.

证明

根据 f 的线性性, 对任意的 $x_0 \in \mathscr{X}$, 总能找到 $x_1 = \frac{x_0}{f(x_0)}\cdot\|f\|$, 使得 $f(x_1) = \|f\|$. 故要想 $\|x_1\| < 1+\varepsilon$, 只需验证存在 $x_0 \in \mathscr{X}$ 使得 $\|\frac{x_0}{f(x_0)}\|\cdot\|f\| < 1+\varepsilon$ 即可. 注意此即

$$\frac{\|x_1\|}{|f(x_1)|} > \frac{\|f\|}{1+\varepsilon}$$

现根据 $\|f\|$ 作为上确界的定义知, 对任意的 $\varepsilon' > 0$, 存在 $x_{\varepsilon'} \in \mathscr{X}$, 使得

$$\frac{|f(x_{\varepsilon'})|}{\|x_{\varepsilon'}\|} \geq \|f\| - \varepsilon'$$

可定义 ε' 满足 $\|f\| - \varepsilon' = \frac{\|f\|}{1+\varepsilon}$, 因而 x_1 存在, 命题即证.

✍ **练习 3.7**♡ 设 $T : \mathscr{X} \to \mathscr{Y}$ 是线性的, 令

$$N(T) := \{x \in \mathscr{X} : Tx = \theta\}.$$

(1) 若 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 求证: $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间.

证明

显见 $T\theta = \theta$, 故 $\theta \in N(T)$, 亦即 $N(T) \neq \varnothing$. 若 $x_1, x_2 \in N(T)$, 则任取 $a, b \in \mathbb{K}$ 知

$$T(ax_1 + bx_2) = aTx_1 + bTx_2 = \theta \Rightarrow ax_1 + bx_2 \in N(T)$$

故 $N(T)$ 是 $\mathscr{X}$ 的线性子空间. 最后, 只要 $x \in \overline{N(T)}$, 根据极限点的定义知存在 $\{x_n\} \subset N(T)$ 使得 $\|x_n - x\|_{\mathscr{X}} \to 0 (n \to \infty)$, 根据 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 知 $\|Tx_n - Tx\|_{\mathscr{Y}} \to 0 (n \to \infty)$, 又因为 $Tx_n = \theta (\forall n \in \mathbb{N})$, 故 $Tx = \theta$, 亦即 $x \in N(T)$, 故 $\overline{N}(T) \subset N(T)$, 从而 $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间. □

(2) 问 $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间能否推出 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$?

解

未必如此. 设 $\mathscr{X} = (l^1, \|\cdot\|_{l^\infty})$, 另取 $a = (1, -1, 0, \cdots) \in \mathscr{X}$ 与线性泛函:

$$f(x) = \sum_{n=1}^{\infty} \xi_n, \quad \forall x = (\xi_1, \cdots, \xi_n, \cdots) \in \mathscr{X}$$

知 $\|a\|_{l^\infty} = 1, f(a) = 0$. 取用:

$$Tx := x - af(x), \quad \forall x = (\xi_1, \cdots, \xi_n, \cdots) \in \mathscr{X}$$

现讨论 $N(T)$, 令 $Tx = x - af(x) = \theta$, 得到 $x = af(x)$, 两边作用 f 得 $f(x) = f(a)f(x) = 0$, 因而 $f(x) = 0$, 从而 $Tx = x = \theta$, 故 $N(T) \subset \{\theta\}$. 这便得到 $N(T) = \{\theta\}$.

下面说明 T 的无界性, 首先说明线性泛函 f 的无界性. 取 $e_k = (\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots)$, 并设 $x_n = \sum_{k=1}^{n} e_k \in \mathscr{X}$, 知

$$\|x_n\| = 1, \quad f(x_n) = n, \quad \forall n \in \mathbb{N}$$

故 $\|f\| \geq n (\forall n \in \mathbb{N})$, 因而 f 无界.

再证明 T 无界, 用反证法. 如若 T 有界, 知

$$\exists M > 0 \forall x \in \mathscr{X} (\|Tx\| \leq M\|x\|)$$

则

$$\begin{aligned} & Tx = x - af(x) \\ \Rightarrow & af(x) = x - Tx \\ \Rightarrow & \|a\||f(x)| \leq \|x\| + \|Tx\| \leq (1 + M)\|x\| \\ \Rightarrow & |f(x)| \leq (1 + M)\|x\| \end{aligned}$$

其中最后一步是因为 $\|a\| = 1$. 故 f 有界, 但这是矛盾的! 从而 $T \notin \mathscr{L}(\mathscr{X}, \mathscr{Y})$. □

(3) 若 f 是线性泛函, 求证:

$$f \in \mathscr{X}^* \Leftrightarrow N(f) \text{是闭线性子空间.}$$

证明

当 $f \in \mathscr{X}^*$, 这便是 (1) 所证, 下设 $N(f)$ 是闭线性子空间, 往证 $f \in \mathscr{X}^*$, 根据 f 的线性性知只需说明 f 有界. 用反证法, 若 f 无界, 亦即

$$\forall n \in \mathbb{N} \exists x_n \in \mathscr{X} (\|x_n\| = 1 \wedge |f(x_n)| > n)$$

现取 $y_n = \frac{x_n}{f(x_n)} - \frac{x_1}{f(x_1)}$, 知 $f(y_n) = 1 - 1 = 0 \Rightarrow y_n \in N(f)$. 同时 $\|\frac{x_n}{|f(x_n)|}\| < \frac{\|x_n\|}{n} \to 0 (n \to \infty)$, 故 $\|y_n + \frac{x_1}{f(x_1)}\| \to 0 (n \to \infty)$, 亦即 $y_n \to -\frac{x_1}{f(x_1)} (n \to \infty)$. 根据 $N(f)$ 的闭性知 $-\frac{x_1}{f(x_1)} \in N(f)$, 但 $f(-\frac{x_1}{f(x_1)}) = -1 \neq 0$, 矛盾! 故 f 有界, 因而 $f \in \mathscr{X}^*$. □

练习 3.8 设 f 是 $\mathscr{X}$ 上的线性泛函, 记

$$H_f^\lambda := \{x \in \mathscr{X} : f(x) = \lambda\} \quad (\forall \lambda \in \mathbb{K})$$

如果 $f \in \mathscr{X}^*$, 并且 $\|f\| = 1$, 求证:

(1) $|f(x)| = \inf\{\|x - z\| : \forall z \in H_f^0\}(\forall x \in \mathscr{X})$.

证明

既然 $\|f\| = 1$, 知

$$\frac{|f(x-z)|}{\|x-z\|} = \frac{|f(x)|}{\|x-z\|} \leq \|f\| = 1, \quad \forall x \in \mathscr{X}, \forall z \in H_f^0$$

也即对任意的 $x \in \mathscr{X}$ 都有

$$|f(x)| \leq \|x - z\|, \forall z \in H_f^0 \Rightarrow |f(x)| \leq \inf\{\|x - z\| : \forall z \in H_f^0\}$$

再设 $y = x - \frac{f(x)}{f(z)}z, z \in \mathscr{X}$, 知 $f(y) = 0 \Rightarrow y \in H_f^0$, 故

$$\inf\{\|x - z\| : \forall z \in H_f^0\} \leq \|x - y\| = \|\frac{f(x)}{f(z)}z\| = |f(x)|\|\frac{z}{f(z)}\|$$

因为

$$\|f\| = 1 = \sup_{z \in \mathscr{X} \setminus \{\theta\}} \|\frac{f(z)}{z}\|$$

故

$$1 \geq \|\frac{z}{f(z)}\| \Rightarrow \inf\{\|x - z\| : \forall z \in H_f^0\} \leq \|x - y\| \leq |f(x)|$$

综上即得

$$|f(x)| = \inf\{\|x - z\| : \forall z \in H_f^0\}, \quad \forall x \in \mathscr{X}.$$

(2) $\forall \lambda \in \mathbb{K}$, H_f^λ 上任一点 x 到 H_f^0 的距离都等于 $|\lambda|$. 并对 $\mathscr{X} = \mathbb{R}^2, \mathbb{K} = \mathbb{R}$ 情形解释 (1) 和 (2) 的几何意义.

证明

知 H_f^λ 上任一点 x 到 H_f^0 的距离即

$$\rho(x) := \inf\{\|x - z\| : z \in H_f^0\}$$

根据 (1), $\rho(x) = |f(x)|$. 又根据 H_f^λ 的定义知 $|f(x)| = |\lambda|$, 故 $\rho(x) = |\lambda|$.

对 (1), 题设情形下的几何意义即函数值 $f(x, y)$ 的绝对值可视作 (x, y, z) 到 xOy 平面的距离.

对 (2), 题设情形下的几何意义即在 $\mathbb{R}^3$ 中代表 $f(x, y) = \lambda$ 的等值面上的点, 函数取值的绝对值恰为该等值面到 xOy 平面的距离.

练习 3.9 设 $\mathscr{X}$ 是实 B^* 空间, f 是 $\mathscr{X}$ 上的非零实值线性泛函, 求证: 不存在开球 $B(x_0, \delta)$, 使得 $f(x_0)$ 是 $f(x)$ 在 $B(x_0, \delta)$ 中的极大值或极小值.

证明[2]

用反证法, 若存在满足题设的开球 $B(x_0, \delta)$, 不妨设 $f(x_0)$ 是 $f(x)$ 在 $B(x_0, \delta)$ 中的极大值, 进一步不妨设其恰为最大值, 进而对任意的 $x \in B(x_0, \delta)$, 都有 $f(x) \leq f(x_0)$. 取 $x_1, x_2 \in B(x_0, \delta), \mathbb{K} \ni \lambda \in (0, 1)$ 满足

$$x_0 = \lambda x_1 + (1 - \lambda)x_2$$

根据线性性有

$$f(x_0) = \lambda f(x_1) + (1 - \lambda)f(x_2) \leq f(x_0)$$

[2]此处极大值允许 f 是常值.

等号取得当且仅当 $f(x_1)=f(x_2)$. 根据 $B(x_0,\delta)$ 的凸性, 该结论可推广为 $f(x)=f(y),\forall x,y\in B(x_0,\delta)$, 这说明 f 为常值函数, 否则矛盾.

3.2 补充: 拓扑空间的更多内容

先前选自 [**PGC**] 的拓扑空间内容实际上更注重的是元素空间, 选取的最终动机是为了介绍乘积拓扑与 Tychonoff 定理. 这里再选用 [**XDX**] 的拓扑空间内容作一些复习与补充, 动机在于为算子空间收敛性作出拓扑意义上的解释, 以及为更后面的弱收敛, 弱 * 收敛等等作准备.

首先还是回忆一下拓扑的一些基本定义:

定义 3.2.1 (拓扑, 拓扑空间, 开集, 邻域, 闭集, Hausdorff 空间)

设 S 是一个不空的集合, $\mathcal{O}$ 是 S 的一些子集所组成的集类. 如果 $\mathcal{O}$ 满足:

(1) 空集 $\varnothing$ 与全空间 S 在 $\mathcal{O}$ 中;

(2) $\mathcal{O}$ 中任意个集合的并集在 $\mathcal{O}$ 中;

(3) $\mathcal{O}$ 中任意两个集合的交集在 $\mathcal{O}$ 中,

就称 $\mathcal{O}$ 是空间 S 中的一个拓扑, 称 $(S,\mathcal{O})$ 为一个拓扑空间, 简记作 S. $\mathcal{O}$ 中的集合称为 S 的开集, S 中的元素称为点. 如果开集 U 含有点 x, 就称 U 为点 x 的邻域. 任何开集 $O\in\mathcal{O}$ 的补集 $S\backslash O$ 称为闭集. 如果 S 还满足: 对任意 $x,y\in S$, 当 $x\neq y$ 时必有 x,y 的邻域 U_x,V_y 满足

$$U_x\cap V_y=\varnothing$$

就称 S 是 Hausdorff 空间.

♣

但是直接定义拓扑有时会比较费劲. 比如在度量空间中, 一般是先对每个点 x 给出一个特殊的邻域 (即球 $B(x,r)$), 再定义一般邻域与开集. 下面介绍的邻域基就是常见的一种" 利用在一点的一族特殊邻域来定义开集" 的方法.

定义 3.2.2 (邻域基)

设 $(S,\mathcal{O})$ 是拓扑空间, $x\in S$, 又设 $\mathscr{U}(x)$ 是 x 点的某些邻域组成的邻域族. 如果对 x 点的任何邻域 V, 都存在 $U\in\mathscr{U}(x)$ 使得 $U\subset V$, 就称 $\mathscr{U}(x)$ 是拓扑 $\mathcal{O}$ 在 x 点的邻域基.

♣

例如当 R 是度量空间时, 取 $\mathscr{U}(x)=\{B(x,r):r>0\}$, 则 $\mathscr{U}(x)$ 就是 x 点的邻域基. 显然, $\mathscr{U}(x)=\{B(x,r):r\in\mathbb{Q}^+\}$ 也是 x 点的邻域基.

引理 3.2.1

设 $(S,\mathcal{O})$ 是拓扑空间, $\mathscr{U}(x)$ 是在 x 点的邻域基, 则其必然满足条件:

(1) 每个 $U\in\mathscr{U}(x)$ 均含有 x;

(2) 对任意 $U_1,U_2\in\mathscr{U}(x)$, 必存在 $U\in\mathscr{U}(x)$, 使得 $U\subset U_1\cap U_2$;

(3) 设 $U\in\mathscr{U}(x)$, 且 $y\in U$, 则必存在 $V\in\mathscr{U}(y)$ 使得 $V\subset U$.

♡

引理依照邻域的定义即可得证. 需要注意的一件不平凡的结果是: 如果对每点 x 指定了符合引理3.2.1的邻域基, 就相当于给定了拓扑.

引理 3.2.2

设 S 是集合, 如果对每点 $x\in S$ 均指定 S 的子集族 $\mathscr{U}(x)$ 满足引理3.2.1的条件 (1)-(3), 那么 S 上存在唯一的拓扑 $\mathcal{O}$ 使得 $\mathscr{U}(x)$ 成为 $\mathcal{O}$ 在 x 点的邻域基.

♡

证明

利用每个点 x 处的 $\mathscr{U}(x)$ 定义 $\mathcal{O}$ 为: 任意取 S 中的 (可重复的) 若干点 $\{x_\lambda:\lambda\in\Lambda\}$, 且对每点 x_λ 任取 $\mathscr{U}(x_\lambda)$

中的一个集合 U_λ, 构造 S 的子集:

$$U = \bigcup_{\lambda\in\Lambda} U_\lambda$$

现在由 $\mathscr{U}(x)(x\in S)$ 作出的这种类型的集合 U 的全体再加上空集构成的集族记为 $\mathcal{O}$, 下面说明 $\mathcal{O}$ 就是 S 的一个拓扑.

对拓扑的条件 (1): 针对每个 $x\in S$, 取 $U_x\in\mathscr{U}(x)$, 根据 $\mathscr{U}(x)$ 满足的引理3.2.1中的条件 (1) 知 $S=\bigcup\limits_{x\in S} U_x\in\mathcal{O}$, 故 $S\in\mathcal{O},\varnothing\in\mathcal{O}$.

对拓扑的条件 (2): 这由 $\mathcal{O}$ 的构造即得.

对拓扑的条件 (3): 设 $W_1,W_2\in\mathcal{O}$, 任取 $y\in W:=W_1\cap W_2$, 根据 $\mathcal{O}$ 的构造知 $W_\nu=\bigcup\limits_{\lambda\in\Lambda} U_\lambda^\nu(\nu=1,2)$, 故必存在 x_ν 的邻域 $U_\lambda^\nu(x_\nu)(\nu=1,2)$ 使得 $y\in U_\lambda^\nu(x_\nu)(\nu=1,2)$, 故取 $U_\nu=U_\lambda^\nu(x_\nu)$ 即知存在 $U_\nu\in\mathscr{U}(x_\nu)$ 使得 $y\in U_\nu\subset W_\nu(\nu=1,2)$. 现在根据引理3.2.1条件 (3) 知存在 $V_\nu\in\mathscr{U}(y)$ 使得 $V_\nu\subset U_\nu(\nu=1,2)$. 因为 $V_1,V_2\in\mathscr{U}(y)$, 故由引理3.2.1条件 (2) 知存在 $V_y\in\mathscr{U}(y)$ 使得

$$V_y\subset V_1\cap V_2\subset U_1\cap U_2\subset W$$

而这件事对任意 $y\in W$ 均成立, 这说明 $W=\bigcup\limits_{y\in W} V_y$, 从而 $W\in\mathcal{O}$, 因而 $\mathcal{O}$ 是一个拓扑.

再证明 $\mathscr{U}(x)(x\in\mathscr{S})$ 是 $\mathcal{O}$ 所对应的邻域基, 首先说明每个 $U\in\mathscr{U}(y)$ 是开集, 这是因为在 $\mathcal{O}$ 的构造中取单点 y 即可. 再任取 $\mathcal{O}$ 在 y 点的一个邻域 O, 因为 $O\in\mathcal{O}$, 根据 $\mathcal{O}$ 的构造知 $O=\bigcup\limits_{\lambda\in\Lambda} U_\lambda$, 从而存在 $x\in S, U\in\mathscr{U}(x)$ 使得 $y\in U\subset O$, 根据 $\mathscr{U}(x)$ 满足的引理3.2.1条件 (3) 知存在 $V\in\mathscr{U}(y)$ 使得 $V\subset U\subset O$, 故 $\mathscr{U}(y)(y\in S)$ 是邻域基.

最后说明拓扑的唯一性. 如果有另一个拓扑 $\mathcal{O}'$ 满足要求, 因为对任意 $x\in S$ 均有 $\mathscr{U}(x)\subset\mathcal{O}'$, 故 $\mathcal{O}$ 中的任意 $U=\bigcup\limits_{\lambda\in\Lambda} U_\lambda\in\mathcal{O}'$, 这说明 $\mathcal{O}\subset\mathcal{O}'$. 相反地, 如果 $U'\in\mathcal{O}'$, 则对任意 $x\in U'$, 必存在 $\mathscr{U}(x)$ 中的 $O(x)\subset U'$, 于是 $U'=\bigcup\limits_{O(x)=U'} O(x)\in\mathcal{O}$, 故 $\mathcal{O}'\subset\mathcal{O}$, 这便说明 $\mathcal{O}=\mathcal{O}'$. □

定义 3.2.3 (邻域基诱导的拓扑)

称引理3.2.2中的拓扑为邻域基 $\mathscr{U}(x)(x\in S)$ 诱导的拓扑. ♣

从上述证明可以看出, 要想确定拓扑, 可以选择去确定每一点的邻域基. 从邻域基出发有下述定义:

定义 3.2.4

设 $(S,\mathcal{O})$ 是拓扑空间. 如果 S 中的每个点 x 都存在一个邻域基 $\mathscr{U}(x)$ 是可列集, 那么称 $(S,\mathcal{O})$ 是满足第一可列公理的. ♣

回忆从拓扑出发定义的极限:

定义 3.2.5

设 $(S,\mathcal{O})$ 是拓扑空间, $\{x_n\}_{n\in\mathbb{N}}$ 是 S 中的点列, $x\in S$. 如果对 x 的任何邻域 O, 均存在 $N\in\mathbb{N}$, 使得当 $n\geq N$ 时 $x_n\in O$, 那么就称 x_n 按拓扑 $\mathcal{O}$ 收敛到 x, 记作 $x_n\to x(\mathcal{O})$. ♣

曾经在谈论闭集的时候, 一个思想是通过极限去定义闭集. 但这里需要阐明从拓扑出发的闭集和从极限出发的闭集事实上并不等价, 他们的等价性由下述引理保证.

引理 3.2.3

设 $(S,\mathcal{O})$ 是拓扑空间, $A\subset S$. 如果 A 是闭集, 那么 A 中任何收敛点列必收敛于 A 中一点. 又设 $(S,\mathcal{O})$ 满足第一可列公理, 那么当 A 中任何收敛点列必收敛于 A 中一点时, A 是闭集. ♡

例 3.13(极限收敛于自身的集合, 但它不是闭集) 设 $\mathbb{R}$ 是实轴, 但其中规定拓扑为: 令

$$\mathcal{O}=\{\mathbb{R}\backslash B: B\text{是 }\mathbb{R}\text{ 中任意有限子集, 或是可列子集, 或是 }\varnothing\text{, 或是 }\mathbb{R}\}$$

此时 $(\mathbb{R},\mathcal{O})$ 中没有收敛点列, 这是因为若存在点列 $\{x_n\}_{n\in\mathbb{N}}$ 满足 $x_n\to x(\mathcal{O})$, 则考虑 x 的邻域 $O:=\mathbb{R}\backslash\{x_n\}_{n\in\mathbb{N}}$, 知 $\{x_n\}_{n\in\mathbb{N}}\cap O=\varnothing$, 显然与 $x_n\to x(\mathcal{O})$ 的定义矛盾! 现在任取 $(\mathbb{R},\mathcal{O})$ 中的一个不闭的集合 A, 譬如 $A=[0,\infty)$(其补集并非开集), 它满足其中任何收敛点列必收敛到其中一点 (因为它没有收敛点列), 但它不是闭集.

定义 3.2.6 (拓扑线性空间)

设 R 是 $\mathbb{K}$ 上的线性空间, 又设 $(R,\mathcal{O})$ 是一个拓扑空间, 满足:

(i) $(R,\mathcal{O})$ 是 Hausdorff 空间 (即任意两个不同点有不相交的邻域);

(ii) R 中的线性运算是连续的:

(a) 加法运算

$$+:(R,\mathcal{O})\times(R,\mathcal{O})\to(R,\mathcal{O}),\quad (x,y)\mapsto x+y$$

作为映射是连续的;

(b) 数乘运算

$$\cdot:(\mathbb{K},\mathcal{O}_{\mathbb{K}})\times(R,\mathcal{O})\to(R,\mathcal{O}),\quad (\alpha,x)\mapsto \alpha x$$

作为映射是连续的.

那么就称 $(R,\mathcal{O})$ 是拓扑线性空间. ♣

在分析与方程的相关学科中, 最常用的是下述拟赋范线性空间:

定义 3.2.7 (拟赋范线性空间)

设 R 是 $\mathbb{K}$ 上的线性空间, $\{p_\alpha(x):\alpha\in\Lambda\}$ 是 R 上的一族半范数. 若对任意 $x\in R$, 只要 $x\neq 0$, 就存在 $\alpha\in\Lambda$ 使得 $p_\alpha(x)\neq 0$, 那么就称 R 按半范族 $\{p_\alpha:\alpha\in\Lambda\}$ 构成拟赋范线性空间. ♣

对于拟赋范线性空间 R, 若 $\{p_\alpha:\alpha\in\Lambda\}$ 是其半范族, 引进邻域基为: 当 $x\in R$ 时, 任取有限个 $\alpha_1,\cdots,\alpha_n\in\Lambda$, 任取正数 ε 作

$$U(x;\alpha_1,\cdots,\alpha_n,\varepsilon)=\{y:p_{\alpha_\nu}(y-x)<\varepsilon,\nu=1,2,\cdots,n\}$$

对任意 $x\in R$, 记

$$\mathscr{U}(x)=\{U(x;\alpha_1,\cdots,\alpha_n,\varepsilon):n\in\mathbb{N},\alpha_1,\cdots,\alpha_n\in\Lambda,\varepsilon>0\}$$

可以验证它满足引理3.2.1的三个条件, 于是由引理3.2.2知它导出唯一的拓扑, 称该拓扑为由半范族 $\{p_\alpha:\alpha\in\Lambda\}$ 导出的拓扑, 可以证明 R 按这个拓扑称为拓扑线性空间. 一般来说, 拟赋范线性空间总按这个方式引入拓扑, 这样得到的拓扑线性空间又叫做局部凸拓扑线性空间.

3.3 Riesz 表示定理及其应用

3.3.1 知识梳理

在前面的讨论中, 可以证明内积本身可以视作一个线性泛函. 这是因为设 $\mathscr{X}$ 是一个 Hilbert 空间, $\forall y\in\mathscr{X}$, 若定义

$$f_y:x\mapsto(x,y),\quad \forall x\in\mathscr{X},$$

则根据 Cauchy-Schwarz 不等式有

$$|f_y(x)|\leq\|y\|\|x\|,\quad \forall x\in\mathscr{X},$$

因而 $\|f_y\|\leq\|y\|$. 特别若 $y\neq\theta$, 取 $x=y$ 有

$$|f_y(y)|=(y,y)=\|y\|^2$$

得到 $\|f_y\| = \|y\|$. 事实上这个结论反过来也成立, 即: 线性泛函都可以看做内积.

定理 3.3.1 (Riesz 表示定理 (Hilbert 空间))

设 f 是 Hilbert 空间 $\mathscr{X}$ 上的一个连续线性泛函, 则必存在唯一的 $y_f \in \mathscr{X}$, 使得

$$f(x) = (x, y_f), \quad \forall x \in \mathscr{X}. \tag{3.1}$$

且若设 $\mathscr{X}$ 上由内积诱导的范数为 $\|\cdot\|_{\mathscr{X}}$, 则有

$$\|f\|_{\mathscr{X}^*} = \|y_f\|_{\mathscr{X}}.$$

♡

启发

在三维空间中, (3.1)式即

$$f(\boldsymbol{x}) = ax + by + cz = \boldsymbol{n} \cdot \boldsymbol{x}, \quad \forall \boldsymbol{x} \in \mathbb{R}^3$$

其中 $\boldsymbol{x} = (x, y, z), \boldsymbol{n} = (a, b, c)$, 要找的 y_f 就是 $\boldsymbol{n}$, 这是平面 $f(\boldsymbol{x}) = 0$ 的法线.

证明

不妨设 f 不是零泛函, 考察集合 $M := \{x \in \mathscr{X} | f(x) = 0\}$. 因为 f 连续线性, 根据练习(3.7)知 M 是一个真闭线性子空间. 既然 M 是闭线性子空间, 根据正交分解定理(2.6.4), 可取 $x_0 \perp M$, 不妨设 $\|x_0\| = 1$, 则 $\mathscr{X}$ 中任意元素 x 可分解为:

$$x = \alpha x_0 + y \tag{3.2}$$

其中 $y \in M, \alpha = f(x)/f(x_0)$. $y \in M$ 是因为:

$$f(y) = f(x - \alpha x_0) = f(x) - \alpha f(x_0) = 0$$

在(3.2)式两边同时与 x_0 作内积有

$$(x, x_0) = (\alpha x_0 + y, x_0) = \alpha(x_0, x_0) + (y, x_0) = \alpha$$

故

$$f(x) = f(\alpha x_0 + y) = \alpha f(x_0) + f(y) = \alpha f(x_0) = (x, \overline{f(x_0)} x_0)$$

取 $y_f = \overline{f(x_0)} x_0$ 即为欲求.

再证唯一性, 若存在 $y, y' \in \mathscr{X}$ 满足

$$f(x) = (x, y) = (x, y'), \quad \forall x \in \mathscr{X}$$

则

$$(x, y - y') = 0, \quad \forall x \in \mathscr{X}$$

取 $x = y - y'$ 即得 $y = y'$.

最后证明函数关系, 由 Cauchy-Schwarz 不等式知

$$|f(x)| = |(x, y_f)| \le \|x\|\|y_f\| \Rightarrow \|f\|_{\mathscr{X}^*} \le \|y_f\|_{\mathscr{X}}$$

反之, 取 $x = y_f$ 知

$$|f(y_f)| = \|y_f\|^2 \Rightarrow |f(\frac{y_f}{\|y_f\|})| = \|y_f\|$$

从而 $\|f\|_{\mathscr{X}^*} \ge |f(\frac{y_f}{\|y_f\|_{\mathscr{X}}})| = \|y_f\|_{\mathscr{X}}$, 欲证进而成立. □

注 Riesz 表示定理(3.3.1)的几何意义在于: 连续线性泛函 $f(x)$ 的等值面都是互相平行的超平面, 故每个向量 x 的泛函值 $f(x)$ 应该由 x 的垂直于这些等值面的分量决定.

定理 3.3.2

设 $\mathscr{X}$ 是一个 Hilbert 空间, $a(x,y)$ 是 $\mathscr{X}$ 上的共轭双线性函数, 并存在 $M>0$, 使得

$$|a(x,y)| \le M\|x\|\cdot\|y\|, \quad \forall x,y\in\mathscr{X}$$

则存在唯一的 $A\in\mathscr{L}(\mathscr{X})$, 使得

$$a(x,y)=(x,Ay), \quad \forall x,y\in\mathscr{X}$$

且

$$\|A\| = \sup_{\substack{(x,y)\in\mathscr{X}\times\mathscr{X}\\ x\neq\theta,y\neq\theta}} \frac{|a(x,y)|}{\|x\|\cdot\|y\|}.$$

证明

固定 $y\in\mathscr{X}, x\mapsto a(x,y)$ 是一个连续线性泛函, 进而由 Riesz 表示定理(3.3.1)知存在 $z=z(y)\in\mathscr{X}$ 使得

$$a(x,y)=(x,z), \quad \forall x\in\mathscr{X}$$

定义映射 $A:y\mapsto z(y)$, 则有 $a(x,y)=(x,Ay)(\forall x,y\in\mathscr{X})$. 又因为对任意的 $x,y_1,y_2\in\mathscr{X},\alpha_1,\alpha_2\in\mathbb{K}$ 有

$$(x,A(\alpha_1y_1+\alpha_2y_2))=a(x,\alpha_1y_1+\alpha_2y_2)=\overline{\alpha_1}a(x,y_1)+\overline{\alpha_2}a(x,y_2)=(x,\alpha_1Ay_1+\alpha_2Ay_2)$$

故 A 是线性的. 再讨论 $Ay:x\mapsto a(x,y)$ 的范数, 由 $|\alpha(x,y)|\le M\|x\|\cdot\|y\|$ 知:

$$\|Ay\| = \sup_{x\in\mathscr{X}\ni\{\theta\}} \frac{|a(x,y)|}{\|x\|} \le M\|y\|,$$

故 A 有界, 因而 $A\in\mathscr{L}(\mathscr{X})$. 最后知

$$\|A\| = \sup_{y\in\mathscr{X}\setminus\{\theta\}} \frac{\|Ay\|}{\|y\|} = \sup_{\substack{(x,y)\in\mathscr{X}\times\mathscr{X}\\ x\neq\theta,y\neq\theta}} \frac{|a(x,y)|}{\|x\|\cdot\|y\|}.$$

□

注 该定理与后续将介绍的 Lax-Milgram 定理有密切联系, 后者是 PDE 理论中证明解的存在与适定性的强力工具.

3.3.1.1 应用: Laplace 方程 $-\Delta u=f$ 的 Dirichlet 边值问题的弱解

设 $\Omega\subset\mathbb{R}^n$ 是一个有界开区域, $f\in L^2(\Omega)$, 称实函数 u 是

$$\begin{cases} -\Delta u=f & \text{在 } \Omega \text{ 内} \\ u|_{\partial\Omega}=0 \end{cases} \tag{3.3}$$

的一个弱解是指 $u\in H_0^1(\Omega)$(参见例(2.47)), 满足

$$\int_\Omega \nabla u\cdot\nabla v dx=\int_\Omega fvdx, \quad \forall v\in H_0^1(\Omega)$$

这是因为如果 $u\in C^2(\overline{\Omega})$, 且是方程(3.3)的解, 那么对任意的 $v\in C^2(\overline{\Omega}), v|_{\partial\Omega}=0$, 有

$$-\Delta u\cdot v=f\cdot v \Rightarrow \int_\Omega(-\Delta\cdot v)dx=\int_\Omega fvdx$$

回忆 Green 公式:

$$\int_{\partial\Omega} v\frac{\partial u}{\partial n}d\sigma=\int_\Omega\nabla u\cdot\nabla vdx+\int_\Omega v\Delta udx$$

应用到 $\int_\Omega(-\Delta\cdot v)dx$ 得:

$$\int_\Omega(-\Delta u\cdot v)dx=\int_\Omega\nabla u\cdot\nabla vdx-\int_{\partial\Omega}\frac{\partial u}{\partial n}vd\sigma=\int_\Omega\nabla u\cdot\nabla vdx$$

即得

$$\int_\Omega\nabla u\cdot\nabla vdx=\int_\Omega f\cdot vdx, \quad \forall v\in C^2(\Omega), v|_{\partial\Omega}=0$$

因为 $H_0^1(\Omega)$ 是 $C^2(\Omega)$ 的完备化, 故集合 $\{v\in C^2(\Omega): v|_{\partial\Omega}=0\}$ 在 $H_0^1(\Omega)$ 中稠密, 由 $u\in H_0^1(\Omega)$ 与 $f\in L^2(\Omega)$ 可得

$$\int_\Omega \nabla u\cdot\nabla v dx=\int_\Omega f\cdot v dx,\quad \forall v\in H_0^1(\Omega).$$

这说明解必是弱解.

一般来说, 直接求解 PDE 方程的解是很困难的, 所以一般研究 PDE 解的思路都是先求弱解, 再去证明弱解的光滑性. 下面断言方程(3.3)的弱解存在唯一.

定理 3.3.3

$\forall f\in L^2(\Omega)$, 方程

$$-\Delta u=f,\quad x\in\Omega$$

的 0-Dirichlet 问题 (即以 $u|_{\partial\Omega}=0$ 为边值条件) 弱解存在唯一.

证明

先证明存在性, 根据 Poincaré 不等式(2.6.1), 可以定义

$$(u,v)_1:=\int_\Omega \nabla u\cdot\nabla v dx,\quad \forall u,v\in H_0^1(\Omega)$$

是 $H_0^1(\Omega)$ 上的一个内积. 而根据 Cauchy-Schwarz 不等式与 Poincaré 不等式知

$$|\int_\Omega f\cdot v dx|\le(\int_\Omega |f|^2dx)^{\frac12}(\int_\Omega |v|^2dx)^{\frac12}\le\|f\|\cdot C\|v\|_1,\quad \forall v\in H_0^1(\Omega) \tag{3.4}$$

其中 $\|\cdot\|,\|\cdot\|_1$ 分别表示 $L^2(\Omega)$ 与 $H_0^1(\Omega)$ 上的范数. (3.4)式表明对给定的 f:

$$v\mapsto\int_\Omega f\cdot v dx,\quad \forall v\in H_0^1(\Omega)$$

是 $H_0^1(\Omega)$ 上的一个连续线性泛函. 应用 Riesz 表示定理(3.3.1)知存在 $u_0\in H_0^1(\Omega)$ 使得

$$(u_0,v)_1=\int_\Omega \nabla u_0\cdot\nabla v dx=\int_\Omega f v dx,\quad \forall v\in H_0^1(\Omega)$$

故 u_0 是一个弱解.

再证唯一性, 如果 u_0,u_0' 都是弱解, 则

$$(u_0-u_0',v)=0,\forall v\in H_0^1(\Omega)\Rightarrow u_0=u_0'.$$

□

对于非 0-Dirichlet 问题, 总是化成 0-Dirichlet 问题来解决. 给定 $\partial\Omega$ 上的函数 g, 若存在 $u_0\in C^2(\overline{\Omega})$, 使得 $u_0|_{\partial\Omega}=g$, 则考虑定义 $f_0:=-\Delta u_0, v:=u-u_0$, 如果 v 是

$$\begin{cases}-\Delta v=f-f_0\\ v|_{\partial\Omega}=0\end{cases} \tag{3.5}$$

的弱解, 则 u 就是

$$\begin{cases}-\Delta u=f\\ u|_{\partial\Omega}=g\end{cases}$$

的弱解, 而方程(3.5)正是 0-Dirichlet 问题.

3.3.1.2 应用: 变分不等式

定理 3.3.4

设 C 是 $H_0^1(\Omega)$ 中的闭凸子集, 若 $f \in L^2(\Omega)$, 则下列不等式存在唯一解 $u_0^* \in C$:

$$\int_\Omega \nabla u_0^* \cdot \nabla(v - u_0^*)dx \geq \int_\Omega f \cdot (v - u_0^*)dx, \quad \forall v \in C$$

证明

利用 Riesz 表示定理(3.3.1)知对于线性泛函

$$w \mapsto \int_\Omega f \cdot w dx$$

存在唯一 $u_0 \in H_0^1(\Omega)$, 使得

$$(u_0, w)_1 = \int_\Omega \nabla u_0 \cdot \nabla w dx = \int_\Omega f \cdot w dx, \quad \forall w \in H_0^1(\Omega)$$

故题式可化为

$$\int_\Omega \nabla u_0^* \cdot \nabla(v - u_0^*)dx \geq \int_\Omega \nabla u_0 \cdot \nabla(v - u_0^*)dx, \quad \forall v \in C$$

此即

$$\int_\Omega (\nabla u_0^* - u_0) \cdot \nabla(v - u_0^*)dx \geq 0 \Leftrightarrow (u_0^* - y_0, v - u_0^*)_1 \geq 0, \quad \forall v \in C \tag{3.6}$$

根据最佳逼近元的刻画(2.6.5), (3.6)式说明 u_0^* 是 u_0 在 C 上的最佳逼近元, 而最佳逼近元是存在唯一的. □

注

1. 本定理可以换成更一般的结果. 设 $A = (a_{ij}(x))$ 是 $n \times n$ 正定阵, 满足
$$\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \delta \sum_{i=1}^n |\xi_i|^2 \quad (\delta > 0)$$
其中 $a_{ij}(x) \in C(\overline{\Omega})$, 则对任意的 $f \in L^2(\Omega)$, 存在唯一的 $u^* \in C$, 使得
$$\int_\Omega \sum_{i,j=1}^n a_{ij}(x)\partial_j u^*(x)\partial_i(v(x) - u^*(x))dx \geq \int_\Omega f(x)(v(x) - u^*(x))dx \quad (\forall v \in C)$$
2. 若 C 由一个连续函数 $\psi(x) \in C(\overline{\Omega})$ 给定:
$$C := \{v \in H_0^1(\Omega) : v(x) \leq \psi(x)\}$$
则上述变分不等式问题称为障碍问题, 此时 u 表示薄膜位移, f 表示外力, $\psi(x)$ 是一个障碍.

3.3.1.3 Radon-Nikodym 定理

根据 [**ZMQ**] 所言, Radon-Nikodym 定理是”经典测度论的最高成就”.

定理 3.3.5 (Radon-Nikodym)

设 $(\Omega, \mathcal{B}, \mu), (\Omega, \mathcal{B}, \nu)$ 是两个 $\sigma-$ 有限测度[a], 且 ν 关于 μ 绝对连续, 即

$$E \in \mathcal{B}, \mu(E) = 0 \Rightarrow \nu(E) = 0$$

则存在关于 μ 的可测函数 g, 且 $g(x) \geq 0$ a.e. μ 使得

$$\nu(E) = \int_E g(x)d\mu, \quad \forall E \in \mathcal{B}.$$

[a] 对测度空间 $(\Omega, \mathcal{B}, \mu)$, 任取 $E \in \mathcal{B}$, 存在一列 $\{E_n\} \subset \mathcal{B}$ 满足 $\forall n \in \mathbb{N}(\mu(E_i) < \infty)$ 且 $E \subset \bigcup\limits_{i=1}^{\infty} E_i$, 则称测度 μ 是 $\sigma-$ 有限的.

下面介绍 Von Neumann 的证明.

证明

先设 $\mu(\Omega)<\infty$. 考虑实 Hilbert 空间 $L^2(\Omega,(\mu+\nu))$, 其范数为

$$\|u\|^2=\int_\Omega u^2(x)d(\mu+\nu)$$

令 $l(u)=\int_\Omega ud\mu$ 是 $L^2(\Omega,(\mu+\nu))$ 到 $\mathbb{R}$ 的泛函, 显见 $l(u)$ 关于 u 是线性的. 由 Cauchy-Schwarz 不等式知[3]

$$\|l(u)\|\le\mu(\Omega)^{\frac{1}{2}}(\int_\Omega u^2d\mu)^{\frac{1}{2}}\le\mu(\Omega)^{\frac{1}{2}}\|u\|,\quad\forall u\in L^2(\Omega,(\mu+\nu))$$

这说明 $l(u)$ 有界, 因而是连续线性泛函. 根据 Riesz 表示定理(3.3.1), 存在函数 $v\in L^2(\Omega,(\mu+\nu))$, 使得

$$\int_\Omega ud\mu=\int_\Omega u\cdot vd(\mu+\nu)=\int_\Omega u\cdot vd\mu+\int_\Omega u\cdot vd\nu$$

也即

$$\int_\Omega u\cdot(1-v)d\mu=\int_\Omega u\cdot vd\nu,\quad\forall u\in L^2(\Omega,(\mu+\nu))\tag{3.7}$$

现在证明

$$0<v(x)\le1,\quad\text{a.e. }\mu.$$

这是因为设 $F=\{x\in\Omega:v(x)\le0\}$, 取 $u(x)=\chi_F(x)$ 代入(3.7)式得到

$$\int_F(1-v)d\mu=\int_F vd\nu$$

故

$$\mu(F)=\int_F d\mu=\int_F vd(\mu+\nu)\le0$$

从而 $\mu(F)=0$. 同样, 设 $G=\{x\in\Omega:v(x)>1\}$, 取 $u(x)=\chi_F(x)$ 代入(3.7)式得到

$$0\ge\int_F(1-v)d\mu=\int_G vd\nu\ge\nu(G)\ge0$$

故

$$\int_G(1-v)d\mu=0$$

又因为 $1-v(x)<0,x\in G$, 故 $\mu(G)=0$. 因而 $0<v(x)\le1,x\in\Omega$, a.e. μ. 令 $g(x)=\frac{1-v(x)}{v(x)}$, 则 $g(x)\ge0$, a.e. μ, 且 $g(x)$ 关于 μ 可测. 对任意的 $E\in\mathcal{B}$, 取 $u(x)=\frac{\chi_E(x)}{v(x)+\frac{1}{n}}$ 代入(3.7)式得到

$$\int_\Omega\chi_E(x)\frac{1-v(x)}{v(x)+\frac{1}{n}}d\mu=\int_\Omega\chi_E(x)\frac{v(x)}{v(x)+\frac{1}{n}}d\nu$$

因为 ν 关于 μ 绝对连续, 且 $v>0$, a.e. μ, 故 $v>0$, a.e. ν. 令 $n\to\infty$, 根据单调收敛定理有

$$\int_E g(x)d\mu=\nu(E),\quad\forall E\in\mathcal{B}$$

至此 $\mu(\Omega)<\infty$ 的情形得证, 下面考虑情形 $\mu(\omega)=\infty$. 根据 σ 有限性, 取集列 $\{\Omega_n\}$ 满足 $\Omega_n\subset\Omega_{n+1},\Omega=\bigcup\limits_{n\ge1}\Omega_n,\mu(\Omega_n)<\infty,n\ge1$. 根据前述结论, 任取 $E\subset\Omega$, 知

$$\nu(E\cap\Omega_n)=\int_{E\cap\Omega_n}g_nd\mu$$

根据集列的增性显见 $g_n(x)=g_{n+1}(x),x\in\Omega_n$. 令 $g(x)=\lim\limits_{n\to\infty}g_n(x)$, 由单调收敛定理知

$$\nu(E)=\nu(E\cap\Omega)=\nu(E\cap\lim_{n\to\infty}\Omega_n)=\int_{E\cap\lim\limits_{n\to\infty}\Omega_n}d\nu=\lim_{n\to\infty}\int_{\Omega_n\cap E}d\nu=\lim_{n\to\infty}\int_{\Omega_n\cap E}g_nd\mu=\int_E gd\mu,\quad E\in\mathcal{B}$$

定理得证. □

[3]注意 $\|l(u)\|$ 是指 $|\int_\Omega ud\mu|$, 此时的范数是在 $\mathbb{R}$ 中的范数.

3.3.2 习题

★ 本节各题中, H 均为 Hilbert 空间.

练习 3.10★ 设 $f_1, f_2, \cdots, f_n$ 是 H 上的一组有界线性泛函,

$$M := \bigcap_{k=1}^{n} N(f_k), N(f_k) := \{x \in H : f_k(x) = 0\} \quad (k = 1, 2, \cdots, n)$$

$\forall x_0 \in H$, 记 y_0 为 x_0 在 M 上的正交投影, 求证: $\exists y_k \in N(f_k)^{\perp}(k = 1, 2, \cdots, n)$ 及 $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$, 使得

$$y_0 = x_0 - \sum_{k=1}^{n} \alpha_k y_k.$$

证明

由 Riesz 表示定理知存在 $\{y_k\}_{k=1}^n \subset H$ 使得

$$f_k(x) = (x, y_k), \quad k = 1, 2, \cdots, n$$

则 M 可被刻画为:

$$\forall x \in M = \bigcap_{k=1}^{n} N(f_k) \Rightarrow \forall k \in \{1, \cdots, n\} \forall x \in M((x, y_k) = 0)$$

故 $y_k \perp M(k = 1, 2, \cdots, n)$, 因而 $\operatorname{span}\{y_k\}_{k=1}^n \subset M^{\perp}$. 根据 Gram-Schmidt 过程可构造一组正交规范集 $\{z_k\}_{k=1}^m (1 \leq m \leq n)$ 使得

$$\operatorname{span}\{z_k\}_{k=1}^m = \operatorname{span}\{y_k\}_{k=1}^n$$

于是

$$z_k \perp M, \quad k = 1, 2, \cdots, m$$

现取 $z_0 = x_0 - y_0$, 则 $z_0 \perp M$. 任取 $x \in H$ 有:

$$(x, z_0) = (\sum_{k=1}^{m} (x, z_k) z_k + x - \sum_{k=1}^{m} (x, z_k) z_k, z_0)$$

注意对任意 $1 \leq k \leq n$ 均有

$$f_k(x - \sum_{k=1}^{m} (x, z_k) z_k) = (x - \sum_{k=1}^{m} (x, z_k) z_k, z_k) = (x, z_k) - (x, z_k) = 0$$

故 $\forall k \in \{1, \cdots, n\}(x - \sum_{k=1}^{m} (x, z_k) z_k \in N(f_k)) \Rightarrow x - \sum_{k=1}^{m} (x, z_k) z_k \in \bigcap_{k=1}^{n} N(f_k) = M$. 又因为 $z_0 \perp M$, 故

$$(\sum_{k=1}^{m} (x, z_k) z_k + x - \sum_{k=1}^{m} (x, z_k) z_k, z_0) = (\sum_{k=1}^{m} (x, z_k) z_k, z_0) = (x, \sum_{k=1}^{m} \overline{(z_k, z_0)} z_k)$$

根据 x 的任意性即知 $z_0 = \sum_{k=1}^{m} \overline{(z_k, z_0)} z_k$, 故

$$x_0 - y_0 = z_0 \in \operatorname{span}\{z_k\}_{k=1}^m = \operatorname{span}\{y_k\}_{k=1}^n$$

命题即证.

练习 3.11★(能量泛函) 设 l 是 H 上的实值有界线性泛函, C 是 H 中的一个闭凸子集. 又设

$$f(v) = \frac{1}{2}\|v\|^2 - l(v) \quad (\forall v \in C)$$

(1) 求证: $\exists u^* \in H$, 使得

$$f(v) = \frac{1}{2}\|u^* - v\|^2 - \frac{1}{2}\|u^*\|^2 \quad (\forall v \in C).$$

证明

根据 Riesz 表示定理知存在 $u^* \in H$ 使得 $l(v) = (u^*, v)$, 进而

$$\begin{aligned} f(v) &= \frac{1}{2}(v,v) - (u^*,v) = \frac{1}{2}(v-u^*,v) - \frac{1}{2}(u^*,v) \\ &= \frac{1}{2}(v-u^*,v-u^*) - \frac{1}{2}(u^*,u^*) = \frac{1}{2}\|u^*-v\|^2 - \frac{1}{2}\|u^*\|^2, \quad \forall v \in C \end{aligned}$$

(2) 求证: $\exists| u_0 \in C$, 使得 $f(u_0) = \inf\limits_{v\in C} f(v)$.

证明

题即证明存在唯一的 $u_0 \in C$ 满足

$$\|u^* - u_0\| = \inf_{v\in C}\|u^* - v\|$$

在 (1) 中将 v 视作 u^* 在 C 上的逼近元, 既然 C 是闭凸子集, 知 u^* 在 C 上的最佳逼近元存在唯一, 取 u_0 为该元素即可.

练习 3.12 设 H 的元素是定义在集合 S 上的复值函数. 又若 $\forall x \in S$, 由

$$J_x(f) = f(x) \quad (\forall f \in H)$$

定义的映射 $J_x : H \to \mathbb{C}$ 是 H 上的连续线性泛函. 求证: 存在 $S \times S$ 上的复值函数 $K(x,y)$, 适合条件:

(1) 对任意固定的 $y \in S$, 作为 x 的函数有 $K(x,y) \in H$;

(2) $f(y) = (f, K(\cdot, y)), \forall f \in H, \forall y \in S$.

证明

既然 J_y 是连续线性泛函, 根据 Riesz 表示定理知存在 $K_y \in H$ 使得

$$J_y(f) = (f, K_y)$$

当 y 固定, $K_y(x)$ 即为 H 中一个固定的函数. 而对任意的 $y \in S$, 总可以找到这样的 K_y 使得 $f(y) = J_y(x) = (f, K_y)$. 故另记 $K(x,y) = K_y(x)$ 即为欲求.

注 满足条件 (1) 和 (2) 的函数 $K(x,y)$ 称为 H 的再生核.

练习 3.13 求证: $H^2(D)$(定义见例(2.50)) 的再生核为

$$K(z,w) = \frac{1}{\pi(1-z\overline{w})^2} \quad (z, w \in D).$$

证明[4]

显见对任意固定的 w, $K(z,w) \in H^2(D)$, 故只需验证 $f(w) = (f, K(\cdot, w))$ 即可, 此即

$$f(w) = \iint_D f(z) \cdot \overline{\frac{1}{\pi(1-z\overline{w})^2}} dz$$

练习 3.14★ 设 L, M 是 H 上的闭线性子空间, 求证:

(1) $L \perp M \Leftrightarrow P_L P_M = 0$.

证明

当 $L \perp M$, 知任取 $x \in H$, 有 $P_M x \in M$, 故 $P_M x \perp L \Rightarrow P_L P_M x = 0$.

当 $P_L P_M = 0$, 知对任意的 $x \in H$, $P_M x \perp L$, 故 $P_M(H) \perp L$. 又因为 $M \subset P_M(H)$, 故 $M \perp L$.

(2) $L = M^\perp \Leftrightarrow P_L + P_M = I$(恒同算子).

证明

当 $L = M^\perp$, 由正交分解定理知对任意的 $x \in H$, $x = P_M x + P_{M^\perp} x = (P_M + P_L)x$. 故 $P_L + P_M = I$.

[4]对于 $H^2(D)$ 内积的性质还有些没弄清楚.

当 $P_L+P_M=I$, 知此时对任意的 $x\in H$, 都有 $x=P_Mx+P_Lx$. 又根据正交分解定理知本应有 $x=P_Mx+P_{M^\perp}x$, 由唯一性得 $P_L=P_{M^\perp}$. 既然 L,M 都是闭线性子空间, 这便说明 $L=M^\perp$.

(3) $P_LP_M=P_{L\cap M}\Leftrightarrow P_LP_M=P_MP_L$.

说明

该命题也出现在了 **[XDX]** 上, 这里整理该定理作为解答.

补充定义 3.3.1 (投影算子)

设 $\mathscr{X}$ 是线性空间, P 是定义在 $\mathscr{X}$ 上的线性算子. 若 $P^2=P$, 那么称 P 是 $\mathscr{X}$ 上的投影算子. 当 $\mathscr{X}$ 是赋范线性空间时, 满足 $P^2=P$ 的有界线性算子称为投影算子.

补充定理 3.3.1

设 P_L,P_M 是两个在 Hilbert 空间上定义的投影算子, 那么 P_LP_M 成为投影算子的充要条件是 $P_LP_M=P_MP_L$. 而且在 P_LP_M 是投影算子时, 它就是在 $L\cap M$ 上的投影算子.

证明

首先说明 Hilbert 空间 H 上的投影算子 P 必是自共轭算子, 这是因为设 P 是从 H 到 $L\subset H$ 的投影算子, 则对任意 $x_1,x_2\in H$, 设

$$\begin{cases} x_1=Px_1+z_1, z_1\perp L \\ x_2=Px_2+z_2, z_2\perp L \end{cases}$$

故:

$$(Px_1,x_2)=(Px_1,Px_2+z_2)=(Px_1,Px_2)=(Px_1+z_1,Px_2)=(x_1,Px_2)$$

另一方面从共轭算子的定理出发即得

$$(Px_1,x_2)=(x_1,P^*x_2)$$

比对即得 $P=P^*$, 于是 P 是自共轭算子.

现在若 P_LP_M 是投影算子, 那么它是自共轭算子, 于是

$$P_MP_L=(P_MP_L)^*=P_L^*P_M^*=P_LP_M$$

而若 $P_LP_M=P_MP_L$, 则

$$(P_LP_M)^2=P_LP_MP_LP_M=P_LP_LP_MP_M=P^2LP^2M=P_LP_M$$

于是 P_LP_M 是投影算子.

最后, 当 P_LP_M 是投影算子, 若 $x\in L\cap M$, 则 $P_LP_Mx=P_Lx=x$. 另一边, 若 x 满足 $x=P_LP_Mx$, 则至少有 $x\in L$. 又因为 $x=P_MP_Lx$, 故 $x\in M$, 从而 $x\in L\cap M$, 这说明 P_LP_M 是 $L\cap M$ 上的投影算子, 命题即证. □

3.4 纲与开映射定理

3.4.1 知识梳理

从泛函分析的视角来看, 大部分解方程的问题就是对给定的算子 $T:\mathscr{X}\to\mathscr{Y}$ 和 $y\in\mathscr{Y}$, 求 $x\in\mathscr{X}$ 使得

$$Tx=y \tag{3.8}$$

要说明解的存在性, 就是要说明算子 T 有右逆 T_r^{-1}:

$$TT_r^{-1}=I\quad (I\text{表示恒同算子})$$

这是因为若令 $x=T_r^{-1}y$, 则 $Tx=TT_r^{-1}y=y$.

要说明解的唯一性, 就是要说明算子 T 有左逆 T_l^{-1}:

$$T_l^{-1}T = I$$

这是因为 $Tx = y \Rightarrow x = T_l^{-1}Ty$ 唯一被 y 确定.

通过这些讨论知道, 要想说明解的存在唯一性, 只需线性算子 T 既有左逆又有右逆. 注意当 T 的左右逆同时存在, 则它们相等:

$$T_l^{-1} = T_l^{-1}I = T_l^{-1}TT_r^{-1} = IT_r^{-1} = T_r^{-1}$$

故此时称算子 T 有逆, 记之为 T^{-1}.

当 $\mathscr{X},\mathscr{Y}$ 都具有拓扑结构, 且方程(3.8)的解已经存在唯一, 下一步就希望讨论这个解是否稳定, 即当 y 有微小变化的时候, x 也只有微小变化, 也就是说 T^{-1} 是连续的. 已经知道映射 T 连续, 是指开集 U 在 T 作用下的原像 $T^{-1}(U)$ 是开的, 故要使得 T^{-1} 连续, 就需要 T 把开集映成开集. 为了绕过对 T^{-1} 存在性的讨论, 称 $T : \mathscr{X} \to \mathscr{Y}$ 是开映射, 如果它把开集映成开集.

3.4.1.1 纲与纲推理

在介绍 Baire 纲定理之前, 首先回忆 Cantor 闭集套定理:

补充定理 3.4.1 (♡Cantor 闭集套定理)

设 $(\mathscr{X}, d)$ 是完备距离空间, 如果 $\{F_n\}_{n=1}^{\infty}$ 是 $\mathscr{X}$ 中的一列闭集套: $F_n \supset F_{n+1}(n = 1, 2, \cdots)$, 且 $\lim\limits_{n\to\infty} d(F_n) = 0$, 则存在唯一 $x_0 \in \mathscr{X}$, 使得 $\bigcap\limits_{n=1}^{\infty} F_n = \{x_0\}$.

♡

[PGC] 中给出了 Baire 定理的两种表述:

补充定理 3.4.2 (♡Baire[PGC])

设 $(\mathscr{X}, d)$ 是完备度量空间, 则下列两条等价性质成立:

(a) 设 $\{F_n\}_{n=0}^{\infty}$ 是 $\mathscr{X}$ 中的闭子集序列, 且对任意 $n \geq 0$, 均有 $F_n^{\circ} = \varnothing$, 则 $(\bigcup\limits_{n=0}^{\infty} F_n)^{\circ} = \varnothing$.

(b) 设 $\{O_n\}_{n=0}^{\infty}$ 是 $\mathscr{X}$ 中的开子集序列, 且对任意 $n \geq 0$, 均有 $\overline{O_n} = \mathscr{X}$, 则 $\overline{\bigcap\limits_{n=0}^{\infty} O_n} = \mathscr{X}$.

♡

证明

首先说明 (a) 与 (b) 的等价性. 注意对任意的 $A \subset \mathscr{X}$, 都有 $\overline{A} = \mathscr{X}\backslash(\mathscr{X}\backslash A)^{\circ}$. 这是因为任取 $x \in \overline{A}$, 知必有 $x \notin \mathscr{X}\backslash\overline{A}$, 而 $\mathscr{X}\backslash\overline{A} = (\mathscr{X}\backslash A)^{\circ}$, 故 $x \in \mathscr{X}\backslash(\mathscr{X}\backslash A)^{\circ}$, 亦即 $\overline{A} \subset \mathscr{X}\backslash(\mathscr{X}\backslash A)^{\circ}$. 另一方面, 任取 $x \in \mathscr{X}\backslash(\mathscr{X}\backslash A)^{\circ}$, 知 $x \notin (\mathscr{X}\backslash A)^{\circ} = \mathbb{X}\backslash\overline{A}$, 故 $x \in \overline{A}$, 亦即 $\mathscr{X}\backslash(\mathscr{X}\backslash A)^{\circ} \subset \overline{A}$, 即得断言. 该断言说明

$$\overline{A} = \mathscr{X} \Leftrightarrow (\mathscr{X}\backslash A)^{\circ} = \varnothing$$

若 (a) 成立, 任取开子集 $O_n \subset \mathscr{X}$ 且 $\overline{O_n} = \mathscr{X}(n = 1, 2, \cdots)$. 记 $F_n = \mathscr{X}\backslash O_n$, 则 F_n 是闭集, 且

$$F_n^{\circ} = (\mathscr{X}\backslash O_n)^{\circ} = \mathscr{X}\backslash\overline{O_n} = \varnothing, \quad \forall n \geq 0$$

由 (a) 知 $(\bigcup\limits_{n=0}^{\infty} F_n)^{\circ} = \varnothing$. 根据 de Morgan 律知

$$\bigcup_{n=0}^{\infty} F_n = \bigcup_{n=0}^{\infty}(\mathscr{X}\backslash O_n) = \mathscr{X}\backslash(\bigcap_{n=0}^{\infty} O_n)$$

这说明

$$(\mathscr{X}\backslash(\bigcup_{n=0}^{\infty} O_n))^{\circ} = \varnothing = \mathscr{X}\backslash\overline{(\bigcup_{n=0}^{\infty} O_n)}$$

因而

$$\overline{(\bigcup_{n=0}^{\infty} O_n)} = \mathscr{X}.$$

故 (b) 成立. 类似可证 (b)⇒(a).

现在着手来证明 (a). 注意 $\mathscr{X}$ 子集 A 的内部非空当且仅当存在 $\mathscr{X}$ 的非空开子集 O 包含于 A, 该断言由开集定义即可得证. 注意 $O \subset A \subset \mathscr{X}$ 等价于 $O \cap (\mathscr{X} \backslash A) = \varnothing$, 因而有下述逻辑表达式:

$$A^\circ \neq \varnothing \Leftrightarrow \exists \text{非空开集} O \subset \mathscr{X} (O \cap (\mathscr{X} \backslash A) = \varnothing)$$

该逻辑表达式等价于

$$A^\circ = \varnothing \Leftrightarrow \forall \text{非空开集} O \subset \mathscr{X} (O \cap (\mathscr{X} \backslash A) \neq \varnothing)$$

现给定 $\mathscr{X}$ 中的闭集列 $\{F_n\}_{n=0}^{\infty}$, 满足 $\forall n \geq 0 (F_n^\circ = \varnothing)$, 希望证明 $(\bigcup\limits_{n=0}^{\infty} F_n)^\circ = \varnothing$. 根据前述等价, 只需证明对任意非空开集 $O \subset \mathscr{X}$, 均有 $O \cap (\mathscr{X} \backslash (\bigcup\limits_{n=0}^{\infty} F_n)) \neq \varnothing$. 为此任取非空开集 $O_0 \subset \mathscr{X}$. 因为 $F_0 \in \{F_n\}_{n=0}^{\infty}$, 故 $F_0^\circ = \varnothing$, 这等价于对任意非空开集 $O \subset \mathscr{X}$ 均有 $O \cap (\mathscr{X} \backslash F_0) \neq \varnothing$, 自然也有 $O_0 \cap (\mathscr{X} \backslash F_0) \neq \varnothing$. 因为 F_0 本身是闭集, 故 $\mathscr{X} \backslash F_0$ 是开集, 从而 $O_0 \cap (\mathscr{X} \backslash F_0)$ 是非空开集, 因而可取闭球 $\overline{B}(a_1, r_1)$ 使得

$$\overline{B}(a_1, r_1) \subset O_0 \cap (\mathscr{X} \backslash F_0), \quad r_1 < 1$$

记 O_1 为开球 $B(a_1, r_1)$, 因为 $F_1 \in \{F_n\}_{n=0}^{\infty}$, 故 $F_1^\circ = \varnothing$, 因而对任意非空开集 O 均有 $O \cap (\mathscr{X} \backslash F_1) \neq \varnothing$, 同理可说明 $O_1 \cap (\mathscr{X} \backslash F_1)$ 是非空开集, 因而又存在闭球 $\overline{B}(a_2, r_2)$ 使得

$$\overline{B}(a_2, r_2) \subset O_1 \cap (\mathscr{X} \backslash F_1), \quad r_2 < \frac{1}{2}$$

如此下去得到非空闭球列

$$\overline{B}(a_1, r_1) \supset \overline{B}(a_2, r_2) \supset \cdots \supset \overline{B}(a_n, r_n) \supset \cdots, \quad r_n < \frac{1}{n}$$

其中

$$\overline{B}(a_{n+1}, r_{n+1}) \subset B(a_n, r_n) \cap (\mathscr{X} \backslash F_n), \quad n \geq 1$$

故由 Cantor 闭集套原理知存在唯一 $x \in \mathscr{X}$ 使得

$$\{x\} = \bigcap_{n=1}^{\infty} B(a_n, r_n)$$

一方面, 知 $x \in \overline{B}(a_1, r_1) \subset O_0$. 另一方面, 对任意 $n \geq 1$ 均有 $x \in \overline{B}(a_{n+1}, r_{n+1}) \subset \mathscr{X} \backslash F_n$, 这说明 $x \in \bigcap\limits_{n=0}^{\infty} (\mathscr{X} \backslash F_n) = \mathscr{X} \backslash (\bigcup\limits_{n=0}^{\infty} F_n)$. 故 $x \in O_0 \cap (\mathscr{X} \backslash (\bigcup\limits_{n=0}^{\infty} F_n))$, 亦即 $O_0 \cap (\mathscr{X} \backslash (\bigcup\limits_{n=0}^{\infty} F_n)) \neq \varnothing$. 根据 O_0 的任意性即得

$$(\bigcup_{n=0}^{\infty} F_n)^\circ = \varnothing$$

欲证进而成立. □

[ZL] 中的表述将内部为空的集合新下了一个定义:

定义 3.4.1 (疏集)

设 $(\mathscr{X}, \rho)$ 是一个度量空间, 集合 $E \subset \mathscr{X}$, 称 E 是疏的, 如果 $\overline{E}$ 无内点. ♣

例 3.14 在 $\mathbb{R}^n$ 上, 有穷点集是疏集, Cantor 集是疏集.

命题 3.4.1

设 $(\mathscr{X},\rho)$ 是一度量空间. 为了 $E\subset\mathscr{X}$ 是疏集, 必须且仅须对任意的球 $B(x_0,r_0)$, 存在 $B(x_1,r_1)\subset B(x_0,r_0)$, 使得

$$\overline{E}\cap\overline{B}(x_1,r_1)=\varnothing$$

证明

当 E 是疏集, 根据定义知 $\overline{E}$ 无内点, 因而 $\overline{E}$ 不能包含任意一个球 $B(x_0,r_0)$, 也即对任意球 $B(x_0,r_0)$, 总存在 $x_1\in B(x_0,r_0)$ 使得 $x_1\notin\overline{E}$. 因为 $\overline{E}$ 是闭集, 故 $\rho(x_1,\overline{E})>0$, 进而可取 $r_1=\min\{\frac{1}{2}\rho(x_1,\overline{E}),r_0-\rho(x_0,x_1)\}$, 得到 $B(x_1,r_1)\subset B(x_0,r_0)$, 且 $\overline{B}(x_1,r_1)\cap\overline{E}=\varnothing$.

如果 E 并非疏集, 也即 $\overline{E}$ 有内点, 设其内点为 x_0, 根据内点的定义知存在 $\delta>0$ 使得 $B(x_0,\delta)\subset\overline{E}$, 进而对任意的 $B(x_1,r_1)\subset B(x_0,\delta)$, 总有 $B(x_1,r_1)\subset\overline{E}\Rightarrow\overline{B}(x_1,r_1)\cap\overline{E}\neq\varnothing$, 矛盾! □

补充命题 3.4.1 (⋆)

设度量空间 $\mathscr{X}$ 由子空间 $\mathscr{X}_1$, 且有集合 $E\subset\mathscr{X}_1$. 若 E 为 $\mathscr{X}_1$ 的疏集, 则 E 也是 $\mathscr{X}$ 的疏集.

证明

用反证法. 若 E 在 $\mathscr{X}$ 中并非疏集, 亦即存在 $x_0\in\mathscr{X},r_0>0$ 使得 $B(x_0,r_0)\subset\overline{E}|_{\mathscr{X}}$, 其中 $\overline{E}|_{\mathscr{X}}$ 表示 E 在 $\mathscr{X}$ 中的闭包. 若 x_0 是 E 的孤立点, 则

$$\{x_0\}\subset E\cap B(x_0,r_0)\neq\varnothing$$

而若 x_0 为 E 的极限点, 根据极限点的定义亦有 $E\cap B(x_0,r_0)\neq\varnothing$, 从而 $E\cap B(x_0,r_0)$ 非空. 现取 $x_1\in E\cap B(x_0,r_0)$, 令 $r_1:=r_0-\rho(x_0,x_1)$, 记 $\overline{E}|_{\mathscr{X}_1}$ 为 E 在 $\mathscr{X}_1$ 中的闭包, 则:

$$\overline{E}|_{\mathscr{X}_1}=\overline{E}|_{\mathscr{X}}\cap\mathscr{X}_1 \tag{3.9}$$

这是因为一方面显见 $\overline{E}|_{\mathscr{X}_1}\subset\overline{E}|_{\mathscr{X}},\overline{E}|_{\mathscr{X}_1}\subset\mathscr{X}_1$, 故 $\overline{E}|_{\mathscr{X}_1}\subset\overline{E}|_{\mathscr{X}}\cap\mathscr{X}_1$. 另一方面, 任取 $x\in\overline{E}|_{\mathscr{X}}\cap\mathscr{X}_1$, 根据 $\overline{E}|_{\mathscr{X}}$ 的定义知存在 $\{x_n\}_{n\in\mathbb{N}}\subset E$ 使得 $x_n\to x(n\to\infty)$ 在 $\mathscr{X}$ 中成立. 又因为 $\mathscr{X}_1\subset\mathscr{X}$, 故 $x_n\to x(n\to\infty)$ 在 $\mathscr{X}_1$ 中也成立. 从而由 $x\in\mathscr{X}_1$ 知 $x\in\overline{E}|_{\mathscr{X}_1}$, 因而 $\overline{E}|_{\mathscr{X}_1}=\overline{E}|_{\mathscr{X}}\cap\mathscr{X}_1$.

现在由(3.9)式知

$$(B(x_1,r_1)\cap\mathscr{X}_1)\subset(B(x_0,r_0)\cap\mathscr{X}_1)\subset\overline{E}|_{\mathscr{X}}\cap\mathscr{X}_1=\overline{E}|_{\mathscr{X}_1}$$

从而 x_1 为 $\overline{E}|_{\mathscr{X}_1}$ 在 $\mathscr{X}_1$ 中的内点, 故 E 在 $\mathscr{X}_1$ 中并非疏集, 矛盾! 命题即证. □

定义 3.4.2 (第一纲集, 第二纲集)

在度量空间 $(\mathscr{X},\rho)$ 上, 集合 E 称为第一纲集, 如果 $E=\bigcup\limits_{n=1}^{\infty}E_n$, 其中 E_n 是疏集. 不是第一纲集的集合称作第二纲集.

例 3.15 在 $\mathbb{R}$ 上, 有理点集是第一纲集. 更一般地, 可数点集总是第一纲集.

定理 3.4.1 (Baire)

完备度量空间 $(\mathscr{X},\rho)$ 是第二纲集.

证明

用反证法. 若 $\mathscr{X}$ 是第一纲集, 根据定义知存在疏集 $\{E_n\}$, 使得

$$\mathscr{X}=\bigcup_{n=1}^{\infty}E_n$$

由命题(3.4.1)知, 对任意的球 $B(x_0,r_0)$, 存在 $B(x_1,r_1)\subset B(x_0,r_0)(r_1<1)$, 使得

$$\overline{B}(x_1,r_1)\cap\overline{E_1}=\varnothing$$

对球 $B(x_1,r_1)$, 存在 $B(x_2,r_2)\subset B(x_1,r_1)(r_2<\frac{1}{2})$, 使得

$$\overline{B}(x_2,r_2)\cap(\overline{E_1}\cup\overline{E_2})=\varnothing$$

这样一直下去, 对球 $B(x_{n-1},r_{n-1})$, 存在 $B(x_n,r_n)\subset B(x_{n-1},r_{n-1})(r_n<\frac{1}{n})$, 使得 $\overline{B}(x_n,r_n)\cap\overline{E_n}=\varnothing$, 故

$$\overline{B}(x_n,r_n)\cap(\bigcup_{i=1}^{n}\overline{E_i})=\varnothing\quad(\forall n\in\mathbb{N})$$

通过 $B(x_n,r_n)$ 的构造方式知

$$\overline{B}(x_0,r_0)\supset\overline{B}(x_1,r_1)\supset\cdots\supset\overline{B}(x_n,y_n)\supset\cdots$$

而

$$\rho(x_{n+p},x_n)\le r_n<\frac{1}{n}\quad(\forall n,p\in\mathbb{N})$$

这说明 $\{x_n\}$ 是基本列, 故存在 $x\in\mathscr{X}$ 使得 $\lim\limits_{n\to\infty}x_n=x$. 另一方面, 令 $p\to\infty$ 得到 $\rho(x,x_n)\le r_n$, 故

$$x\in\overline{B}(x_n,r_n),\quad\forall n\in\mathbb{N}$$

进而 $x\notin\bigcup\limits_{n=1}^{\infty}E_n$, 但这与 $\mathscr{X}=\bigcup\limits_{n=1}^{\infty}E_n$ 矛盾! 命题即证. □

特别注意第二纲不能推知连续统势, 也就是说存在可数的第二纲集.

例 3.16★(可数的第二纲集) 取 $(\mathbb{Z},|\cdot|)$, 其中 $\mathbb{Z}$ 是整数集. 可以证明 $(\mathbb{Z},|\cdot|)$ 是完备度量空间, 因而由 Baire 定理知 $(\mathbb{Z},|\cdot|)$ 是第二纲集, 但 $(\mathbb{Z},|\cdot|)$ 是可数的.

注 在上面的例子中需要注意单点集 $\{n\}$ 此时不是疏集. 这是因为一方面确实存在开球 $(n-\frac{1}{2},n+\frac{1}{2})=\{n\}\subset\{n\}$, 另一方面它显然与自己的闭包相等. 于是 $(\overline{\{n\}})^\circ=\{n\}$ 非空, 进而 $\{n\}$ 不是疏集. 集是否为第一纲集与集势的关系依赖于下述命题.

补充命题 3.4.2 (★)

若度量空间 $\mathscr{X}$ 的单点集 $\{x\}$ 不是开集, 则 $\mathscr{X}$ 中的可数集必为第一纲集.

证明

注意单点集永远是闭集. 若 $E=\{x\}$ 不是开集, 则 $\overline{E}=E$ 也不是开集, 亦即 $\overline{E}$ 无内点, 故 E 是疏集. 因为可数集为可数个单点集之并, 而第一纲集为可数个疏集之并, 故可数集是第一纲集. □

例 3.17★(不完备的第二纲空间) 设 $\mathscr{X}$ 为全体无理数, 任取 $x,y\in\mathscr{X}$, 赋距离 $d(x,y):=|x-y|$, 则 $(\mathscr{X},d)$ 为不完备的第二纲集. 这是因为 $\mathscr{X}^c$ 为全体有理数, 其作为可数集的同时单点集不是开集, 从而 $\mathscr{X}^c$ 是第一纲集. 若 $\mathscr{X}$ 也是第一纲集, 则 $\mathbb{R}=\mathscr{X}\cup\mathscr{X}^c$ 是第一纲集. 但 Baire 定理断言 $(\mathbb{R},d)$ 作为完备距离空间是第二纲集, 矛盾! 故 $\mathscr{X}$ 是第二纲集.

例 3.18★(利用 Hamel 基在任意无穷维 Banach 空间内作出不完备第二纲线性子空间)WL 设 X 是无穷维 Banach 空间, H 是其 Hamel 基, 从 H 中取可数真子集 $\{h_n\}_{n\in\mathbb{N}}\subset H$, 令

$$X_0=\operatorname{span}\{H\backslash\{h_n\}_{n\in\mathbb{N}}\},X_n=\operatorname{span}\{H\backslash\{h_k:k\ge n+1\}\},\quad n=1,2,\cdots$$

则由 Hamel 基的定义知 $X=\bigcup\limits_{n=0}^{\infty}X_n$, 且由 X 的完备性知 X 是第二纲的, 这说明 $X_n(n\ge 0)$ 不能全为第一纲集. 设 X_{n_0} 为第二纲的线性子空间, 进而 $\overline{X_{n_0}}$ 在 X 中必有内点, 亦即存在 X 中的球 $B(x_0,\delta)\subset\overline{X_{n_0}}$, 于是由 $\overline{X_{n_0}}$ 的线性性可推知 $\overline{X_{n_0}}=X$. 另一方面, 显见 $X_{n_0}\subset X,X_{n_0}\ne X$, 这说明 X_n 不是闭的, 因而不是完备空间, 进而其为欲求.

下面的定理说明处处连续但处处不可微的函数占大多数.

定理 3.4.2

在 $C[0,1]$ 中处处不可微的函数集合 E 是非空的, 更确切的, E 的余集是第一纲集.

证明

取 $\mathscr{X}=C[0,1]$, 设 A_n 表示 $\mathscr{X}$ 中这样一些元素 f 所成的集合: 对 f, 存在 $s\in[0,1]$, 使得对满足 $0\le s+h\le 1$ 与 $|h|\le\frac{1}{n}$ 的任何 h, 成立:

$$|\frac{f(s+h)-f(s)}{h}|\le n$$

如果 f 在某个点 s 处可微, 则必有正整数 n, 使得 $f\in A_n$, 于是

$$\mathscr{X}\backslash E\subset\bigcup_{n=1}^{\infty}A_n$$

下面证明每个 A_n 都是疏集, 首先证明 A_n 是闭的. 这是因为若 $f\in\mathscr{X}\backslash A_n$ 则对任意的 $s\in[0,1]$, 都存在 h_s 满足

$$|h_s|\le\frac{1}{n}\quad 且\quad |f(s+h_s)-f(s)|>n|h_s|$$

根据 f 的连续性, 存在 $\varepsilon_s>0$ 及 s 的某个邻域 J_s, 使得对任意的 $\sigma\in J_s$, 都有

$$|f(\sigma+h_s)-f(\sigma)|>n|h_s|+2\varepsilon_s \tag{3.10}$$

根据有限覆盖定理, 设 $J_{s_1},\cdots,J_{s_m}$ 覆盖 $[0,1]$, 并设

$$\varepsilon=\min\{\varepsilon_{s_1},\varepsilon_{s_2},\cdots,\varepsilon_{s_m}\}$$

现在若 $g\in\mathscr{X}$ 满足 $\|g-f\|<\varepsilon$, 则由(3.10)式知对任意的 $\sigma\in J_{s_k}(k=1,\cdots,m)$ 有

$$|g(\sigma+h_{s_k})-g(\sigma)|\ge|f(\sigma+h_{s_k})-f(\sigma)|-2\varepsilon>n|h_{s_k}|$$

故 $g\in\mathscr{X}\backslash A_n\Rightarrow\mathscr{X}\backslash A_n$ 是开集, 故 A_n 是闭集.

再证明 A_n 无内点. 任取 $f\in A_n,\varepsilon>0$, 根据 Weierstrass 逼近定理, 存在多项式 p 使得

$$\|f-p\|<\frac{\varepsilon}{2}$$

因为 p 的导数在 $[0,1]$ 上有界, 故根据中值定理, 存在 $M>0$, 使得对任意的 $s\in[0,1],|h|<\frac{1}{n}$ 成立

$$|p(s+h)-p(s)|\le M|h|$$

另设 $g\in C[0,1]$ 是一个分段线性函数, 满足 $\|g\|<\frac{\varepsilon}{2}$, 且各条线段斜率绝对值均大于 $M+n$, 得到

$$p+g\in B(f,\varepsilon),\quad p+g\notin A_n$$

故 A_n 是闭集, 因而其为疏集, 得到 $\bigcup\limits_{n=1}^{\infty}A_n$ 是第一纲集. 又因为 $\mathscr{X}$ 是完备的, 由 Baire 定理(3.4.1)知 $\mathscr{X}$ 是第二纲集, 故根据 $\mathscr{X}\backslash E\subset\bigcup\limits_{n=1}^{\infty}A_n$ 知 E 是第二纲集.

3.4.1.2 补充: Baire 定理的更多应用

[PGC] 直言: 大部分线性泛函分析主要定理的共同特点是, 它们有赖于下述两个基本结果或其中的某一个: Baire 定理和赋范向量空间中的 Hahn-Banach 定理. 其中 Baire 定理断言: Banach 空间 (或更一般的完备距离空间) 中可列无穷多个稠密开子集的交仍然稠密. 从 Baire 定理出发可以直接推知多项式空间的不完备性与处处不可微连续函数的" 多" 性 (这个性质在前面已经介绍过了). 同样, Baire 定理可以导出共鸣定理 (即 Banach-Steinhaus 定理, 又称一致有界原理), 从共鸣定理可以推知存在连续函数使得其 Lagrange 多项式插值不一定收敛, 也存在连续函数使得其 Fourier 级数不一定收敛. 开映射定理与闭图像定理同样是 Baire 定理的推论. 下面介绍一些 Baire 定理的直观应用.

例 3.19(♡ 有理数域的不完备性) 考察所有有理数组成的可列无限集 $\mathbb{Q}=\bigcup\limits_{n=0}^{\infty}\{q_n\}$, 其中 $\{q_n\}$ 以任意方式排序, 赋通常距离 $d(p,q)=|p-q|$. 因为 $\mathbb{Q}$ 的每一子集 $\{q_n\}$ 都是闭集且内部为空, 故 $(\mathbb{Q},d)$ 内部也应为空, 从而其并不完备.

例 3.20♡[PGC] 平面 $\mathbb{R}^2$ 不能表示为可列条直线的并集, 因为后者依 Baire 定理内部为空, 这与前者显见不符. 更一般地, 空间 $\mathbb{R}^n$ 不能表示为可列个超平面的并集.

例 3.21[PGC] 无穷维的 Banach 空间不可能具有可列无穷的 Hamel 基. 特别地, 由单变量或多变量的多项式组成的空间不可能装备某范数成为 Banach 空间.

证明

用反证法, 在赋范向量空间 $(\mathscr{X}, \|\cdot\|)$ 中给定一组可列无穷 Hamel 基 $\{e_j\}_{j=0}^{\infty}$, 定义 $\mathscr{X}$ 的子集 F_n 为:

$$F_n := \operatorname{span}\{e_j\}_{j=0}^{\infty}, \quad n = 1, 2, \cdots$$

显见 $\mathscr{X} = \bigcup\limits_{n=0}^{\infty} F_n$, 且每一个子集 F_n 在 $\mathscr{X}$ 中作为有限维线性子空间均是闭集. 下面证明对全体 $n \geq 0$ 均有 $F_n^{\circ} = \varnothing$. 实际上, 如若存在 $n' \geq 0$ 使得 $F_{n'}^{\circ} \neq \varnothing$, 则必存在 $x = \sum\limits_{j=0}^{n'} x_j e_j \in F_{n'}$ 与 $r > 0$, 使得 $\overline{B}(x, r) \subset F_{n'}$, 于是点

$$y := \frac{r}{\|e_{n'+1}\|} e_{n'+1} + x \in \overline{B}(x, r) \subset F_{n'}$$

但因为 $\{e_j\}_{j=0}^{n+1}$ 线性无关, 故 $y \notin \operatorname{span}\{e_j\}_{j=0}^{n}$, 矛盾! 从而全体 F_n 的内部均空, 因而 $\mathscr{X}^{\circ} = \varnothing$, $\mathscr{X}$ 不可能完备.

特别对于 n 变量多项式空间 $\mathbb{P}$ 而言, 多项式

$$x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \mapsto x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}, \quad k_i \in \mathbb{N}, 1 \leq i \leq n$$

恰是 $\mathbb{P}$ 的一个可列无穷 Hamel 基. □

3.4.1.3 开映射定理

定义 3.4.3 (单射, 满射)

设 $\mathscr{X}, \mathscr{Y}$ 都是 B 空间, $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 算子 T 称为单射, 如果 T 是一一的; 算子 T 称为满射, 如果 $T(\mathscr{X}) = \mathscr{Y}$. ♣

当 T 是单射时, 可以定义 T^{-1}, 它是线性算子, 但定义域不一定是整个 $\mathscr{Y}$. 只有当 T 是满射时, T^{-1} 可以成为 $\mathscr{Y}$ 到 $\mathscr{X}$ 的线性算子, 这时希望确定 T^{-1} 的连续性, 进而有 Banach 定理.

定理 3.4.3 (Banach 逆算子定理)

设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间. 若 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 它既是单射又是满射, 那么 $T^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$. ♡

Banach 定理更一般的形式即开映射定理. 回忆开映射的定义:

定义 3.4.4 (开映射)

称 $T : \mathscr{X} \to \mathscr{Y}$ 是开映射, 如果 T 把 $\mathscr{X}$ 中的开集映成 $\mathscr{Y}$ 中的开集. ♣

定理 3.4.4 (开映射定理)

设 $\mathscr{X}, \mathscr{Y}$ 都是 B 空间, 若 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 是一个满射, 则 T 是开映射. ♡

证明

下用 $B(x_0, a), U(y_0, b)$ 分别表示 $\mathscr{X}, \mathscr{Y}$ 中的开球. 要证明 T 是开映射, 也即其把开集映成开集, 必须且只需证明: 存在 $\delta > 0$, 使得

$$TB(\theta, 1) \supset U(\theta, \delta) \tag{3.11}$$

现在来证明这个等价.

当 T 是开映射, $TB(\theta, 1)$ 自然是开集, 且 θ 是 $TB(\theta, 1)$ 的内点, 故根据内点的定义即得 $U(\theta, \delta) \subset TB(\theta, 1)$.

当存在 $\delta > 0$ 使得 $U(\theta, \delta) \subset TB(\theta, 1)$, 由于 T 的线性性, 知条件等价于

$$TB(x_0, r) \supset U(Tx_0, r\delta) \quad (\forall x_0 \in \mathscr{X}, \forall r > 0)$$

对任意的 $y_0 \in T(W)$, 根据 $T(W)$ 的定义知存在 $x_0 \in W$ 使得 $y_0 = Tx_0$. 因为 W 是开集, 故存在 $B(x_0, r) \subset W$, 再取 $\varepsilon = r\delta$ 得到

$$U(Tx_0, \varepsilon) \subset TB(x_0, r) \subset T(W)$$

这说明 $y_0 = Tx_0$ 是 $T(W)$ 的内点.

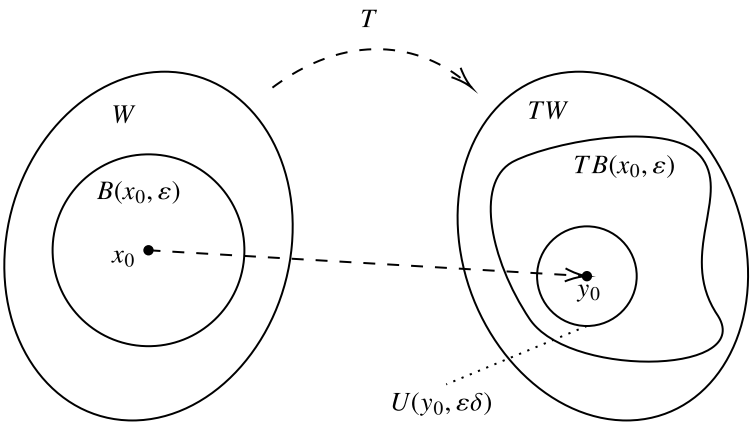


图 3.1: T 把内点映成内点, 因而把开集映成开集.

回到原命题, 现在只需证明(3.11)式, 首先证明存在 $\delta > 0$ 满足 $\overline{TB(\theta, 1)} \supset U(\theta, 3\delta)$. 这是因为由 $\mathscr{X} = \bigcup\limits_{n=1}^{\infty} B(\theta, n)$ 可得

$$\mathscr{Y} = T\mathscr{X} = T(\bigcup_{n=1}^{\infty} B(\theta, n)) \overset{\text{(i)}}{=} \bigcup_{n=1}^{\infty} TB(\theta, n)$$

其中 (i) 是因为 T 是满射. 由于 $\mathscr{Y}$ 是完备的, 故根据 Baire 定理(3.4.1)知 $\mathscr{Y}$ 是第二纲集, 因而 $\bigcup\limits_{n=1}^{\infty} TB(\theta, n)$ 是第二纲集, 所以至少存在一个 $n \in \mathbb{N}$ 使得 $TB(\theta, n)$ 不是疏集, 也即 $\overline{TB(\theta, n)}$ 至少含有一个内点, 亦即存在 $U(y_0, r) \subset \overline{TB(\theta, n)}$. 因为 T 把 θ 映成 θ, 故 $TB(\theta, n)$ 是一个对称凸集, 因而 $U(-y_0, r) \subset \overline{TB(\theta, n)}$, 从而

$$U(\theta, r) \subset \frac{1}{2}(U(y_0, r) + U(-y_0, r)) \subset \overline{TB(\theta, n)}$$

其中

$$\frac{1}{2}(U(y_0, r) + U(-y_0, r)) = \{\frac{z}{2} : z = x + y, x \in U(y_0, r), y \in U(-y_0, r)\}$$

对于这个断言的证明, 任取 $y \in U(\theta, r)$, 考虑 $y_0 + y \in U(y_0, r), -y_0 + y \in U(-y_0, r)$ 即可. 根据 T 的齐次性, 取 $\delta = \frac{r}{3n}$ 即得 $U(\theta, 3\delta) \subset \overline{TB(\theta, 1)}$.

再证明 $U(\theta, \delta) \subset TB(\theta, 1)$ 亦即证明任取 $y_0 \in U(\theta, \delta)$, 都存在 $x_0 \in B(\theta, 1)$ 满足 $Tx_0 = y_0$, 此即求方程 $Tx = y_0$ 在 $B(\theta, 1)$ 内的一个解 x_0. 下面用逐次逼近法.

对于 $y_0 \in U(\theta, \delta) \subset \overline{TB(\theta, \frac{1}{3})}$, 根据闭包的定义知存在 $x_1 \in B(\theta, \frac{1}{3})$ 满足

$$\|y_0 - Tx_1\| < \frac{\delta}{3}$$

对于 $y_1 = y_0 - Tx_1 \in U(\theta, \frac{\delta}{3}) \subset \overline{TB(\theta, \frac{1}{3^2})}$, 根据闭包的定义知存在 $x_2 \in B(\theta, \frac{1}{3^2})$ 满足

$$\|y_1 - Tx_2\| < \frac{\delta}{3^2}$$

这样一直下去, 对 $y_n = y_{n-1} - Tx_n \in U(\theta, \frac{\delta}{3^n}) \subset \overline{TB(\theta, \frac{1}{3^{n+1}})}$, 知存在 $x_{n+1} \in B(\theta, \frac{1}{3^{n+1}})$ 满足

$$\|y_n - Tx_{n+1}\| < \frac{\delta}{3^{n+1}}$$

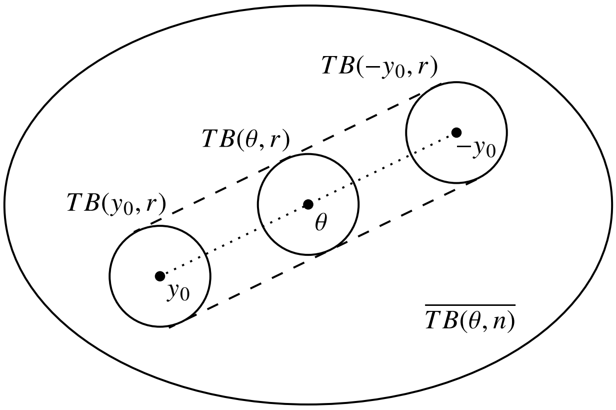


图 3.2: $\frac{1}{2}(U(y_0, r) + U(-y_0, r))$ 可以看作这两个球与它们的元素的中点的集合, 而根据线性性 $U(\theta, r)$ 恰在这个集合中.

于是 $\sum\limits_{n=1}^{\infty} \|x_n\| \leq \frac{1}{2}$, 取 $x_0 := \sum\limits_{n=1}^{\infty} x_n$, 则知 $x_0 \in B(\theta, 1)$. 又因为

$$\|y_n\| = \|y_{n-1} - Tx_n\| = \cdots = \|y_0 - T(x_1 + \cdots + x_n)\| < \frac{\delta}{3^n}, \quad \forall n \in \mathbb{N}$$

得到

$$S_n := \sum_{i=1}^{n} x_i \to x_0 \quad \text{且} \quad TS_n \to y_0, \quad n \to \infty$$

又因为 T 连续, 故

$$Tx_0 = y_0 \Rightarrow y_0 \in TB(\theta, 1)$$

这说明 $U(\theta, \delta) \subset TB(\theta, 1)$. □

现在来补全 Banach 定理的证明.

证明

根据 T 的线性性, 已知存在 $\delta > 0$ 使得

$$U(\theta, 1) \subset TB(\theta, \frac{1}{\delta})$$

即

$$T^{-1}U(\theta, 1) \subset B(\theta, \frac{1}{\delta}) \Rightarrow \|T^{-1}y\| < \frac{1}{\delta}, \quad \forall y \in \mathscr{Y}, \|y\| < 1$$

故

$$\|T^{-1}\| = \sup_{\|y\|<1} \|T^{-1}y\| \leq \frac{1}{\delta}$$

也即 $T^{-1} \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. □

注

1. 在定理(3.4.3)中, $T\mathscr{X}$ 是第二纲集的假设是必不可少的 (此处是由满射与 $\mathscr{Y}$ 的完备性保证).
2. Banach 逆算子定理(3.4.3)与开映射定理(3.4.4)中的 Banach 空间可以换成 F 空间, 后面将会补充相关定理.

例 3.22(Banach 逆算子定理在 $T\mathscr{X}$ 为第一纲集时不再成立) 取 $\mathscr{X} = \mathscr{Y} = C[0, 1]$, 设

$$(Tx)(t) = \int_0^t x(\tau)d\tau \quad (\forall x \in \mathscr{X})$$

T 显然是连续线性的, 但

$$T\mathscr{X} = \mathscr{Y}_0 = \{y \in C^1[0, 1] : y(0) = 0\}$$

在一致收敛范数下不是 $C[0,1]$ 的第二纲集, 这是因为依据定理(3.4.2), $C^1[0,1]$ 本身是 $C[0,1]$ 的第一纲集, 因而 $\mathscr{Y}_0$ 作为 $C^1[0,1]$ 的子空间也是 $C[0,1]$ 的第一纲集. 此时 $T^{-1}=\frac{d}{dt}$ 在 $C[0,1]$ 中并不连续 (即使以 $C[0,1]$ 的子集 $\mathscr{Y}_0$ 作为 T^{-1} 定义域, 它也不是连续的). 这是因为取 $x_n:=\sin n\pi t$, 显见 $\|x_n\|=1$, 但

$$\|\frac{d}{dt}x_n(t)\|=n\pi\|\cos n\pi t\|=n\pi\to\infty,\quad n\to\infty$$

其中 $\|\cdot\|$ 表示 $C[0,1]$ 空间中的范数. 但如果更改范数, 在 $\mathscr{Y}_0$ 上取 $C^1[0,1]$ 的范数 $\|\cdot\|_1$:

$$\|y\|_1=\max_{t\in[0,1]}|y(t)|+\max_{t\in[0,1]}|y'(t)|$$

则 $\mathscr{Y}_0$ 是 B 空间, 此时 $T^{-1}=\frac{d}{dt}$ 有界, 这是因为

$$\|T^{-1}y\|=\|\frac{d}{dt}y(t)\|\le\|y\|_1\quad(\forall y\in\mathscr{Y}_0)$$

□

反思前述对 Banach 逆算子定理(3.4.3)与开映射定理(3.4.4)的证明, 注意 T 的连续性仅仅用在由 $S_n\to x_0$ 且 $TS_n\to y_0$ 推出 $Tx_0=y_0$ 这一步, 也就是说只要 T 能使得这个过程成立, 开映射定理对 T 就成立, 这诱导出下述闭算子的定义:

定义 3.4.5 (闭算子)

设 T 是 $\mathscr{X}\to\mathscr{Y}$ 的线性算子, $D(T)$ 是其定义域. 称 T 是闭的, 是指由 $x_n\in D(T), x_n\to x$ 与 $Tx_n\to y$ 就能推出 $x\in D(T), y=Tx$.

♣

例 3.23 在 $C[0,1]$ 上, $D(T)=C^1[0,1], T=\frac{d}{dt}$ 是一个闭线性算子.

证明

若 $x_n(t)\in C^1[0,1]$, 且有

$$x_n\to x\in C[0,1],\quad \frac{dx_n}{dt}\to y\in C[0,1]$$

则

$$x_n(t)-x_n(t)\to\int_0^t y(\tau)d\tau\quad(\forall t\in[0,1])$$

$$x_n(t)-x_n(0)\to x(t)-x(0)\quad(\forall t\in[0,1])$$

故

$$x(t)-x(0)=\int_0^t y(\tau)d\tau\quad(\forall t\in[0,1])$$

这说明 $x\in C^1[0,1]$, 且 $\frac{dx}{dt}=y(t)$. □

下面给出 Banach 逆算子定理的一个弱化版本.

定理 3.4.5

若 $\mathscr{X},\mathscr{Y}$ 是 B 空间, T 是 $\mathscr{X}\to\mathscr{Y}$ 是一个闭线性算子, 满足 $R(T)$[a]是 $\mathscr{Y}$ 中的第二纲集, 则 $R(T)=\mathscr{Y}$, 且 $\forall\varepsilon>0,\exists\delta=\delta(\varepsilon)>0$, 使得 $\forall y\in\mathscr{Y}$, 若它满足 $\|y\|<\delta$, 则存在 $x\in D(T)$ 满足 $\|x\|<\varepsilon$ 且 $y=Tx$.

[a] T 的值域.

♡

证明

只需证明 $R(T)=\mathscr{Y}$. 根据定理(3.4.4)的证明过程, 已知对 $\varepsilon=1$, 存在 $\delta>0$ 满足

$$U(\theta,\delta)\subset T(B(\theta,1)\cap D(T))$$

现在对任意的 $y\in\mathscr{Y}$, 不妨设 $y\neq\theta$(显见 $\theta\in R(T)$), 则对任意的 $\delta_1\in(0,\delta)$ 有

$$\frac{\delta_1 y}{\|y\|}\in U(\theta,\delta)\Rightarrow\frac{\delta_1 y}{\|y\|}\in T(B(\theta,1)\cap D(T))$$

故存在 $x \in B(\theta,1)\cap D(T)$ 使得

$$\frac{\delta_1 y}{\|y\|} = Tx \Rightarrow y = T(\frac{\|y\|}{\delta_1}x) \Rightarrow y \in R(T)$$

□

注 该定理中

$$\forall \varepsilon > 0 \exists \delta > 0 \forall y \in \mathscr{Y}(\|y\| < \delta \Rightarrow \exists x \in D(T)(\|x\| < \varepsilon \wedge y = Tx))$$

这一结果事实上说明的是 T^{-1}(如果存在的话) 的连续性.

最后, [**WR**] 中给出了最一般的开映射定理:

补充定理 3.4.3 (开映射定理^{WR})

若

(a) $\mathscr{X}$ 是 Frechet 空间,

(b) $\mathscr{Y}$ 是拓扑向量空间,

(c) $T \in \mathscr{L}(\mathscr{X},\mathscr{Y})$,

(d) $T\mathscr{X}$ 是 $\mathscr{Y}$ 中的第二纲集,

则:

(i) $T\mathscr{X} = \mathscr{Y}$,

(ii) T 是开映射,

(iii) $\mathscr{Y}$ 是 Frechet 空间.

利用 Banach 逆算子定理可以推知下述极其有用的等价范数定理:

推论 3.4.1 (等价范数定理)

设线性空间 $\mathscr{X}$ 上有两个范数 $\|\cdot\|_1$ 与 $\|\cdot\|_2$. 如果 $\mathscr{X}$ 关于这两个范数都构成 Banach 空间, 而且 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 则 $\|\cdot\|_2$ 与 $\|\cdot\|_1$ 必等价.

证明

考察恒同映射 $I : (\mathscr{X},\|\cdot\|_2) \to (\mathscr{X},\|\cdot\|_1)$, 显见它是线性算子. 由假设 $\|\cdot\|_2$ 比 $\|\cdot\|_1$ 强, 从而根据范数强弱的等价定义知存在 $C > 0$, 使得

$$\|Ix\|_1 \le C\|x\|_2 \quad (\forall x \in \mathscr{X})$$

其中 $Ix = x$, 故 I 是连续的, 且显见它既是单射又是满射, 故由 Banach 逆算子定理(3.4.3)知 I 可逆且 I^{-1} 连续, 因而 I^{-1} 有界, 即存在 $M > 0$ 使得

$$\|I^{-1}x\|_2 \le M\|x\|_1 \quad (\forall x \in \mathscr{X})$$

又因为 $I^{-1}x = x$, 故 $\|\cdot\|_1$ 与 $\|\cdot\|_2$ 等价. □

3.4.1.4 闭图像定理

回忆闭算子的定义:

定义 3.4.6 (闭算子)

设 T 是 $\mathscr{X} \to \mathscr{Y}$ 的线性算子, $D(T)$ 是其定义域. 称 T 是闭的, 是指由 $x_n \in D(T), x_n \to x$ 与 $Tx_n \to y$ 就能推出 $x \in D(T), y = Tx$.

现在讨论线性算子的连续性和闭性之间的关系. 有结论: 连续线性算子 $T : D(T) \to \mathscr{Y}$ 总可以延拓到 $\overline{D(T)}$ 上.

定理 3.4.6

设 T 是 B^* 空间 $\mathscr{X}$ 到 B 空间 $\mathscr{Y}$ 的连续线性算子, 那么 T 能唯一地延拓到 $\overline{D(T)}$ 上成为连续线性算子 T_1, 使得 $T_1|_{D(T)} = T$, 且 $\|T_1\| = \|T\|$.

证明

对任意的 $x \in \overline{D(T)}$, 根据闭包的定义知均存在 $x_n \in D(T)$ 使得 $\lim\limits_{n\to\infty} x_n = x$. 因为 T 是连续线性算子, 故其有界, 即存在 $M > 0$ 满足

$$\|Tx\|_{\mathscr{Y}} \le M\|x\|_{\mathscr{X}}, \quad \forall x \in D(T)$$

故

$$\|T(x_{n+p} - x_n)\|_{\mathscr{Y}} = \|Tx_{n+p} - Tx_n\|_{\mathscr{X}} \le M\|x_{n+p} - x_n\|$$

这说明 $\{Tx_n\}$ 是 $\mathscr{Y}$ 中的基本列, 因为 $\mathscr{Y}$ 完备, 故存在 $y \in \mathscr{Y}$, 使得 $Tx_n \to y$. 可知 y 仅依赖于 x, 而与 x_n 的选择无关, 故可定义 $T_1 : x \mapsto y$. 显见 T_1 是线性的, $T_1|_{D(T)} = T$, 且 $\|T_1x\| \le M\|x\|(\forall x \in \overline{D(T)})$. □

至此, 可以把每个连续线性算子 T 的定义域都看成闭的, 于是有下述命题:

命题 3.4.2

设 T 是 B^* 空间 $\mathscr{X}$ 到 B^* 空间 $\mathscr{Y}$ 的有界线性算子, 且 $D(T)$ 是闭集, 那么 T 是闭算子.

但一般闭线性算子未必能延拓到 $\overline{D(T)}$ 上, 使其仍然是闭算子, 故只要当结论需要用到 $D(T)$ 中的收敛时, 都需要说明 $D(T)$ 自身的闭性. 在注意到这一点后, 闭图像定理介绍的就是闭线性算子怎样成为连续线性算子.

定理 3.4.7 (闭图像定理)

设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间. 若 T 是 $\mathscr{X} \to \mathscr{Y}$ 的闭线性算子, 并且 $D(T)$ 是闭的, 则 T 是连续的.

证明

因为 $D(T)$ 是闭的, 故 $D(T)$ 作为 $\mathscr{X}$ 的线性子空间可以看作 B 空间. 在 $D(T)$ 上引进范数 $\|\cdot\|_G$:

$$\|x\|_G = \|x\| + \|Tx\| \quad (\forall x \in D(T))$$

下面证明 $(D(T), \|\cdot\|_G)$ 也是 B 空间, 这是因为根据

$$\|x_n - x_m\|_G = \|x_n - x_m\| + \|Tx_n - Tx_m\| \to 0 \quad (n, m \to \infty)$$

知存在 $x^* \in \mathscr{X}, y^* \in \mathscr{Y}$, 使得 $x_n \to x^*, Tx_n \to y^*$. 根据 T 的闭性可得 $y^* = Tx^*$, 故 $Tx_n \to Tx^*$, 这说明 $\|x_n - x^*\|_G \to 0$, 得到 $(D(T), \|\cdot\|_G)$ 是 B 空间. 又因为 $\|\cdot\|_G$ 比 $\|\cdot\|$ 强, 故根据等价范数定理(3.4.1)知 $\|\cdot\|_G$ 与 $\|\cdot\|$ 等价, 故存在 $M > 0$, 使得

$$\|Tx\| \le \|x\|_G \le M\|x\| \quad (\forall x \in D(T))$$

□

注 集合 $G(T) := \{(x, Tx) : x \in D(T)\}$ 称为算子 T 的图像, $\|x\|_G$ 是 (x, Tx) 在乘积空间 $\mathscr{X} \times \mathscr{Y}$ 的范数, 故 $\|\cdot\|_G$ 称为图模. 算子 T 是闭的, 也即 $G(T)$ 按图模是闭的.

补充定理 3.4.4 (♡ 闭算子的等价定义)

设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间, $T : \mathscr{X} \to \mathscr{Y}$ 是线性算子, 则 T 是闭算子当且仅当 $G(T)$ 是闭集.

证明

当 T 是闭算子, 欲证即闭图像定理; 当 $G(T)$ 是闭集, 知只要 $G(T) \ni (x_n, Tx_n) \to (x, y)$, 则 $(x, y) \in G(T)$, 从而 $y = Tx$, 亦即 T 是闭算子, 命题得证. □

闭图像定理的作用在于证明算子连续性时, 不必再按照有界性去进行不等式的证明, 而是可以先验证闭算子这一借助极限的定义, 再由闭图像定理得到连续性, 亦即把证明从不等式转换成了极限. 下面给出闭图像定理的一些应用.

补充定理 3.4.5 (♡Hellinger-Toeplitz)

设 $(\mathscr{X},(\cdot,\cdot))$ 是 Hilbert 空间, $A:\mathscr{X}\to\mathscr{X}$ 是自伴线性算子, 亦即其满足

$$(Ax,y)=(x,Ay),\quad \forall x,y\in\mathscr{X}$$

则 A 是连续的.

证明

根据闭图像定理, 只需证明 A 是闭算子即可. 设 $\{x_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}$ 中的点列, 且满足

$$x_n\to x\in\mathscr{X}, Ax_n\to y\in\mathscr{X},\quad n\to\infty$$

往证 $y=Ax$. 现任取 $z\in\mathscr{X}$, 由内积的连续性知:

$$(Ax_n,z)\to(y,z),(Ax_n,z)=(x_n,Az)\to(x,Az),\quad n\to\infty$$

故

$$(y,z)=(x,Az)=(Ax,z),\quad \forall z\in\mathscr{X}$$

取 $z=y-Ax$ 即得

$$(y-Ax,y-Ax)=0\Rightarrow y=Ax$$

从而 A 是闭算子, 因而由闭图像定理知 A 连续. □

补充定理 3.4.6 (♡)

$\mathscr{X},\mathscr{Y},\mathscr{Z}$ 是 B 空间, $A\in\mathscr{L}(\mathscr{X},\mathscr{Z}),B\in\mathscr{L}(\mathscr{Y},\mathscr{Z})$, 且对任意 $x\in\mathscr{X}$, 方程 $Ax=By$ 都有唯一解 $y\in\mathscr{Y}$, 由此定义算子 $T:\mathscr{X}\to\mathscr{Y},x\mapsto y$, 则 T 是连续线性算子.

证明

由 A,B 的线性性显见 T 是线性算子, 且 $D(T)=\mathscr{X}$ 是闭的, 故只需验证 T 是闭算子, 亦即若取 $\{x_n\}_{n\in\mathbb{N}}\subset\mathscr{X}$ 满足 $x_n\to x,Tx_n\to y$, 往证 $Tx=y$. 现由 T 的构造知:

$$Ax_n=B(Tx_n)$$

两边令 $n\to\infty$ 知 $Ax=By$, 因而由 T 的构造知 $Tx=y$, 故 T 是闭算子, 因而由闭图像定理知其为连续线性算子. □

3.4.1.5 共鸣定理

定理 3.4.8 (共鸣定理 (一致有界定理))

设 $\mathscr{X}$ 是 B 空间, $\mathscr{Y}$ 是 B^* 空间, 如果 $W\subset\mathscr{L}(\mathscr{X},\mathscr{Y})$, 使得

$$\sup_{A\in W}\|Ax\|<\infty\quad(\forall x\in\mathscr{X})$$

那么存在常数 M, 使得 $\|A\|\le M(\forall A\in W)$.

证明

未免混淆, 过程中用 $\|\cdot\|_{\mathscr{A}}$ 表示空间 $\mathscr{A}$ 中的范数. 对任意的 $x\in\mathscr{X}$, 定义

$$\|x\|_W=\|x\|_{\mathscr{X}}+\sup_{A\in W}\|Ax\|_{\mathscr{Y}}$$

显见 $\|\cdot\|_W$ 是 $\mathscr{X}$ 上的范数, 且强于 $\|\cdot\|_{\mathscr{X}}$. 为了应用等价范数定理(3.4.1), 下面说明 $(\mathscr{X},\|\cdot\|_W)$ 完备. 这是因为当

$$\|x_m-x_n\|_{\mathscr{X}}+\sup_{A\in W}\|A(x_m-x_n)\|_{\mathscr{Y}}\to 0\quad(m,n\to\infty)$$

出于 $\mathscr{X}$ 对 $\|\cdot\|_{\mathscr{X}}$ 的完备性, 存在 $x\in\mathscr{X}$ 使得 $\|x_n-x\|_{\mathscr{X}}\to 0(n\to\infty)$. 针对 $\sup\limits_{A\in W}\|A(x_m-x_n)\|_{\mathscr{Y}}\to 0(m,n\to\infty)$, 根据极限的定义知

$$\forall\varepsilon>0\exists N=N(\varepsilon)\forall m,n\geq N(\sup_{A\in W}\|Ax_m-Ax_n\|_{\mathscr{Y}}<\varepsilon)$$

再根据上确界的定义[5]并在两边令 $m\to\infty$ 可知对任意的 $A\in W$ 有 $\|Ax_n-Ax\|_{\mathscr{Y}}\leq\varepsilon(\forall n\geq N)$. 故

$$\|x_n-x\|_{\mathscr{X}}+\sup_{A\in W}\|A(x_n-x)\|_{\mathscr{Y}}\to 0\quad(n\to\infty)$$

也即 $\|x_n-x\|_W\to 0$, 故 $(\mathscr{X},\|\cdot\|_W)$ 完备, 进而根据等价范数定理(3.4.1)知 $\|\cdot\|_W$ 与 $\|\cdot\|_{\mathscr{X}}$ 等价, 从而存在常数 M 使得

$$\sup_{A\in W}\|Ax\|_{\mathscr{Y}}\leq M\|x\|_{\mathscr{X}}\quad(\forall x\in\mathscr{X})$$

由此可知 $\|A\|\leq M(\forall A\in W)$. □

注 注意条件中 $\forall x\in\mathscr{X}(\sup\limits_{A\in W}\|Ax\|<\infty)$ 说明

$$\forall x\in\mathscr{X}\exists M=M(x)>0\forall A\in W(\|Ax\|\leq M(x)\|x\|)$$

而结论则说明 $M(x)$ 实际上与 x 无关. 从而一致有界定理的精髓在于从点态的有界推出了一致有界. 另一方面, 取该定理的逆否命题即得如果 $\sup\limits_{A\in W}\|A\|=\infty$, 则必存在 $x_0\in\mathscr{X}$ 使得 $\sup\limits_{A\in W}\|Ax_0\|=\infty$, 故该定理又称为共鸣定理.

课堂笔记 (♡)

[**ZL**] 上的证明本质上依赖的是等价范数定理, 而 [**PGC**] 中给出了共鸣定理仅依赖 Baire 纲定理的另一种证法, 整理如下.

证明

对每个整数 $n\geq 0$, 定义集合

$$F_n:=\{x\in\mathscr{X}:\sup_{A\in W}\|Ax\|_{\mathscr{Y}}\leq n\}$$

任取 $x\in\mathscr{X}$, 根据条件有 $\sup\limits_{A\in W}\|Ax\|_{\mathscr{Y}}<\infty$, 故必存在关于 x 的整数 n_x 使得 $\sup\limits_{A\in W}\|Ax\|_{\mathscr{Y}}\leq n_x$, 亦即 $x\in F_{n_x}$, 故

$$\mathscr{X}=\bigcup_{n=0}^{\infty}F_n$$

根据 F_n 的构造知, $x\in F_n$ 的充要条件是对全体 $A\in W$ 均有 $\|Ax\|_{\mathscr{Y}}\leq n$, 亦即对任意 $A\in W$ 均有 $x\in\{z\in\mathscr{X}:\|Az\|_{\mathscr{Y}}\leq n\}$. 这说明 F_n 可表为

$$F_n=\bigcap_{A\in W}\{z\in\mathscr{X}:\|Az\|_{\mathscr{Y}}\leq n\}$$

因为 $A\in W\subset\mathscr{L}(\mathscr{X},\mathscr{Y})$, 故 $\{z:\mathscr{X}:\|Az\|_{\mathscr{Y}}\leq n\}$ 是 $\mathscr{X}$ 中的闭集, 因而 F_n 作为闭集之交是闭集. 因为 $\mathscr{X}$ 是完备的, 故由 Baire 纲定理知 $\mathscr{X}$ 是第二纲集, 亦即存在整数 $n_0\geq 0$ 使得 $F_{n_0}^{\circ}\neq\varnothing$. 根据内部的定义知存在 $x_0\in F_{n_0}, r>0$ 使得 $\overline{B}(x_0;r)\subset F_{n_0}$. 根据 F_{n_0} 的构造知

$$\forall z\in\overline{B}(x_0;r)\forall A\in W(\|Az\|\leq n_0)$$

[5] 在这里把 sup 写开是因为不确定在 $\sup\limits_{A\in W}\|Ax_m-Ax_n\|_{\mathscr{Y}}<\varepsilon$ 两边令 $m\to\infty$ 时, 极限能否与 sup 交换.

注意任何一个非零向量 $x \in \mathscr{X}$ 均可写成

$$x = \frac{\|x\|_{\mathscr{X}}}{r}(z - x_0), \quad \text{其中} z := (x_0 + \frac{r}{\|x\|_{\mathscr{X}}}x) \in \overline{B}(x_0; r)$$

故

$$\begin{aligned}\|Ax\|_{\mathscr{Y}} &= \|A(\frac{\|x\|_{\mathscr{X}}}{r}(z - x_0))\|_{\mathscr{Y}} \leq \frac{\|x\|_{\mathscr{X}}}{r}(\|Az\|_{\mathscr{Y}} + \|Ax_0\|_{\mathscr{Y}}) \\ &\leq \frac{\|x\|_{\mathscr{X}}}{r}(n_0 + \|Ax_0\|_{\mathscr{Y}}) \leq \frac{1}{r}(n_0 + \sup_{A \in W} \|Ax_0\|_{\mathscr{Y}})\|x\|_{\mathscr{X}}, \quad \forall A \in W, \forall x \in \mathscr{X}\end{aligned}$$

根据假设知 $\sup\limits_{A \in W} \|Ax_0\|_{\mathscr{Y}} < \infty$, 故

$$\sup_{A \in W} \|A\| \leq \frac{1}{r}(n_0 + \sup_{A \in W} \|Ax_0\|_{\mathscr{Y}}) < \infty$$

此即欲证. □

定理 3.4.9 (Banach-Steinhaus)

设 $\mathscr{X}$ 是 B 空间, $\mathscr{Y}$ 是 B^* 空间, M 是 $\mathscr{X}$ 的某个稠密子集. 若 $A_n(n = 1, 2, \cdots), A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 则 $\forall x \in \mathscr{X}$ 都有

$$\lim_{n \to \infty} A_n x = Ax \tag{3.12}$$

的充要条件是:

(i) $\sup\limits_{n \in \mathbb{N}} \|A_n\| < \infty$;

(ii) (3.12)式对 $\forall x \in M$ 成立.

证明

当 $\forall x \in \mathscr{X}(\lim\limits_{n \to \infty} A_n x = Ax)$, 此时 (ii) 显然成立, 而该极限式本身表明 $\forall x \in \mathscr{X}(\sup\limits_{n \in \mathbb{N}} \|A_n x\| < \infty)$, 故由共鸣定理得 $\sup\limits_{n \in \mathbb{N}} \|A_n\| < \infty$.

当条件 (i) 和 (ii) 成立, 由 (i) 设 $\sup\limits_{n \in \mathbb{N}} \|A_n\| \leq C$, 并由 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 不妨设 $\|A\| \leq C$. 既然 M 在 $\mathscr{X}$ 中稠密, 知

$$\forall x \in \mathscr{X} \forall \varepsilon > 0 \exists y \in M(\|x - y\|_{\mathscr{X}} < \varepsilon)$$

得到

$$\|A_n x - Ax\|_{\mathscr{Y}} \leq \|A_n x - A_n y\|_{\mathscr{Y}} + \|A_n y - Ay\|_{\mathscr{Y}} + \|Ax - Ay\|_{\mathscr{Y}} < C\varepsilon + \|A_n y - Ay\| + C\varepsilon, \quad \forall n \in \mathbb{N}$$

又因为 $y \in M$, 根据条件 (ii), 由 $\lim\limits_{n \to \infty} A_n y = Ay$ 的定义知可取 N 足够大使得

$$\|A_n y - Ay\| < \varepsilon, \quad \forall n \geq N$$

得到

$$\|A_n x - Ax\| < (2C + 1)\varepsilon, \quad \forall n \geq N$$

也即

$$\lim_{n \to \infty} A_n x = Ax \quad (\forall x \in \mathscr{X})$$

□

注 Banach-Steinhaus 定理中(3.12)式推条件 (i) 这一步自己在独立推导时的过程如下:

当 $\forall x \in \mathscr{X}(\lim\limits_{n \to \infty} A_n x = Ax)$, 此时 (ii) 显然成立, 要说明 (i) 成立, 将(3.12)式按定义写开即

$$\forall x \in \mathscr{X} \forall \varepsilon > 0 \exists N = N(\varepsilon, x) > 0 \forall n > N(\|A_n x - Ax\|_{\mathscr{Y}} < \varepsilon)$$

特别取 $\|x\|_{\mathscr{X}} = 1$ 即得 $\|A_n x - Ax\|_{\mathscr{Y}} < \varepsilon\|x\|_{\mathscr{X}} \Rightarrow \|A_n - A\| \leq \varepsilon$. 若设

$$U = \max\{\|A_1\|, \cdots, \|A_N\|, \|A\| + \varepsilon\} < \infty$$

则知 $\sup\limits_{n\in\mathbb{N}}\|A_n\| \le U < \infty$, 因而 (i) 成立.

但这个过程是有问题的! 因为 U 依赖于 N, 而 N 依赖于 x, 这说明 $U<\infty$ 这一件事未必成立.

注 [**XDX**] 中介绍了 Banach-Steinhaus 定理(3.4.9)的一个更强的结果, 为此先介绍一个引理.

引理 3.4.1

设 $\mathscr{X}$ 是赋范线性空间, $\mathscr{Y}$ 是 Banach 空间, $T_n \in \mathscr{L}(\mathscr{X},\mathscr{Y})(n=1,2,\cdots)$. 如果有常数 M 使得 $\|T_n\| < M(n = 1,2,\cdots)$, 且有 $\mathscr{X}$ 的稠密子集 D, 使得当 $x\in D$ 时序列 $\{T_nx\}_{n\in\mathbb{N}}$ 收敛, 则存在算子 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 使得 $\{T_n\}_{n\in\mathbb{N}}$ 强收敛于 T, 即

$$\forall x\in\mathscr{X}(\lim_{n\to\infty}\|T_nx-Tx\|=0)$$

且

$$\|T\| \le \varliminf_{n\to\infty}\|T_n\|.$$

证明

任取 $x\in\mathscr{X}$, 根据 D 在 $\mathscr{X}$ 中稠密的定义知:

$$\forall\varepsilon>0\exists x'\in D(\|x-x'\|<\varepsilon)$$

因为序列 $\{T_nx'\}_{n\in\mathbb{N}}$ 收敛, 故根据 Cauchy 准则知存在 $N>0$ 使得 $n,m\ge N$ 时有

$$\|T_mx'-T_nx'\|<\varepsilon$$

故

$$\begin{aligned}\|T_mx-T_nx\| &\le \|T_mx-T_mx'\|+\|T_mx'-T_nx'\|+\|T_nx'-T_nx\|\\ &\le (\|T_m\|+\|T_n\|)\|x-x'\|+\varepsilon<(2M+1)\varepsilon\end{aligned}$$

由 ε 的任意性知 $\{T_nx\}_{n\in\mathbb{N}}$ 是 Banach 空间 $\mathscr{Y}$ 中的基本列, 因而其为收敛列. 作

$$T:\mathscr{X}\to\mathscr{Y},\quad x\mapsto Tx=\lim_{n\to\infty}T_nx$$

则易证 T 为线性算子. 因为对任意 $n\in\mathscr{N}$ 均有 $\|T_nx\|\le\|T_n\|\cdot\|x\|$, 故

$$\|Tx\|=\lim_{n\to\infty}\|T_nx\|\le\varliminf_{n\to\infty}\|T_n\|\|x\|$$

从而 T 有界, 且

$$\|T\|\le\varliminf_{n\to\infty}\|T_n\|.$$

□

容易发现上述引理与 Banach-Steinhaus 定理对 $\mathscr{X},\mathscr{Y}$ 的完备性要求是相反的. 这其中的区别在于: Banach-Steinhaus 定理中从极限式 $\lim\limits_{n\to\infty}A_nx=Ax$ 推出 $\sup\limits_{n\in\mathbb{N}}\|A_n\|<\infty$ 的过程本身需要用到共鸣定理, 而共鸣定理基于等价范数定理就要求 $\mathscr{X}$ 完备, 对 $\mathscr{Y}$ 就只需为赋范线性空间; 上述引理中, $\mathscr{Y}$ 的完备性主要用于断言算子 T 的存在性, 而 Banach-Steinhaus 定理是直接把 T 存在作为条件 (亦即 $\lim\limits_{n\to\infty}A_nx=Ax$) 来用的. 总而言之, 上述引理的贡献在于给出了 Banach-Steinhaus 定理中算子 A 范数的控制, 得到下述定理:

补充定理 3.4.7 ([XDX]Banach-Steinhaus)

设 $\mathscr{X}$ 是 B 空间, $\mathscr{Y}$ 是赋范线性空间, $\{T_n\}\subset\mathscr{L}(\mathscr{X},\mathscr{Y})$, 且对每个 $x\in\mathscr{X}$, 序列 $\{T_nx\}_{n\in\mathbb{N}}$ 均收敛, 则存在 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 使得 $\{T_n\}$ 强收敛于 T(即 $\forall x\in\mathscr{X}(\lim\limits_{n\to\infty}T_nx=Tx)$), 且 $\|T\|\le\varliminf\limits_{n\to\infty}\|T_n\|$.

这里 $\mathscr{Y}$ 是赋范线性空间与前面引理的完备性并不矛盾, 因为 T 的存在性是通过" 对每个 $x\in\mathscr{X}$, 序列 $\{T_nx\}_{n\in\mathbb{N}}$ 均收敛" 这句话得到的.

3.4.1.6 应用: Lax-Milgram 定理与 Lax 等价定理

定理 3.4.10 (Lax-Milgram)

设 $a(x,y)$ 是 Hilbert 空间 $\mathscr{X}$ 上的一个共轭双线性函数, 满足:

1. $\exists M>0$, 使得 $|a(x,y)|\le M\|x\|\cdot\|y\|(\forall x,y\in\mathscr{X})$,
2. $\exists\delta>0$, 使得 $|a(x,x)|\ge\delta\|x\|^2(\forall x\in\mathscr{X})$,

则必存在唯一的有连续逆的连续线性算子 $A\in\mathscr{L}(\mathscr{X})$, 满足

$$a(x,y)=(x,Ay)\quad 且\quad \|A^{-1}\|\le\frac{1}{\delta}$$

证明

根据定理(3.3.2), 满足 $a(x,y)=(x,Ay)$ 的算子 $A\in\mathscr{L}(\mathscr{X})$ 存在唯一, 为了证明 A 有连续逆, 考虑应用 Banach 定理(3.4.3), 这需要说明 A 既是单射又是满射.

首先, A 是单射. 这是因为若 $y_1,y_2\in\mathscr{X}$ 满足 $Ay_1=Ay_2$, 则

$$a(x,y_1)=a(x,y_2)\quad(\forall x\in\mathscr{X})$$

故

$$a(x,y_1-y_2)=0\quad(\forall x\in\mathscr{X})$$

取 $x=y_1-y_2$, 根据条件 2. 知只能有 $\|y_1-y_2\|=0$, 也即 $y_1=y_2$.

其次, A 是满射, 这便需要说明 $R(A)=\mathscr{X}$. 因为根据练习(2.49), 对于闭线性子空间 M 而言 $(M^\perp)^\perp=M$, 而 $\mathscr{X}^\perp=\{\theta\}$, 故可以通过说明 $R(A)$ 是闭的, 以及 $R(A)^\perp=\{\theta\}$ 来达到欲证.

先证明 $R(A)$ 是闭的. 任取 $w\in\overline{R(A)}$, 根据闭包的定义知存在 $\{v_n\}\subset\mathscr{X}$ 使得

$$w=\lim_{n\to\infty}Av_n \tag{3.13}$$

根据条件 2. 有

$$\delta\|v_{n+p}-v_n\|^2\le|a(v_{n+p}-v_n,v_{n+p}-v_n)|=|(v_{n+p}-v_n,A(v_{n+p}-v_n))|\le\|v_{n+p}-v_n\|\cdot\|Av_{n+p}-Av_n\|\quad(\forall n,p\in\mathbb{N})$$

得到

$$\|v_{n+p}-v_n\|\le\frac{1}{\delta}\|Av_{n+p}-Av_n\|\to 0\quad(\forall p\in\mathbb{N},n\to\infty)$$

这说明 $\{v_n\}$ 是基本列, 又因为 $\mathscr{X}$ 是 Hilbert 空间, 故存在 $v^*\in\mathscr{X}$ 使得 $v_n\to v^*$, 再根据 A 的连续性与(3.13)式得到 $w=Av^*$, 这说明 $w\in R(A)$, 也即 $R(A)$ 是闭的.

再证明 $R(A)^\perp=\{\theta\}$. 若 $w\in R(A)^\perp$, 根据定义知

$$(w,Av)=0\quad(\forall v\in\mathscr{X})$$

也即 $a(w,v)=0(\forall v\in\mathscr{X})$, 取 $v=w$, 根据条件 2. 知

$$\delta\|w\|^2\le|a(w,w)|=0$$

这说明 $w=\theta$, 也即 $R(A)^\perp=\{\theta\}$.

综上, A 是满射. 最后根据 Banach 定理(3.13)得到 $A^{-1}\in\mathscr{L}(\mathscr{X})$. 因为

$$\delta\|x\|^2\le|a(x,x)|=|(x,Ax)|\le\|x\|\cdot\|Ax\|$$

故

$$\forall x\in\mathscr{X}(\delta\|x\|\le\|Ax\|)\Rightarrow\|A^{-1}\|\le\frac{1}{\delta}$$

□

下面讨论 Lax 等价定理. 数值分析中, 要求一个方程的解, 往往都是用近似方程的解来代替的, 比如用差分方程或有限元方程近似代替微分方程. 现在问近似方程的解是否收敛到原方程的解? 如果是, 则称该近似格式具有收敛性.

把上述背景用泛函分析的语言描述即, 设 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 其中 $\mathscr{X}, \mathscr{Y}$ 是 B 空间. 给定 $y \in \mathscr{Y}$, 希望求解 $x \in \mathscr{X}$ 使得

$$Tx = y \tag{3.14}$$

首先设方程的解存在唯一, 也即 $\forall y \in \mathscr{Y}$, 存在唯一的 $x \in \mathscr{X}$ 满足(3.14)式. 这说明 T 既是单射又是满射, 从而根据 Banach 定理(3.4.3)得到 $T^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$. 现在考虑(3.14)的近似方程: $\forall n \in \mathbb{N}$, 设 $T_n \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 希望求解 $x_n \in \mathscr{X}$ 使得

$$T_n x_n = y \tag{3.15}$$

依旧设方程的解存在唯一, 也即 $\forall y \in \mathscr{Y}$, 存在唯一的 $x_n \in \mathscr{X}$ 满足(3.15)式, 仿照前述步骤得到 $T_n^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$.

T_n 是 T 的近似是指:

$$\forall x \in \mathscr{X}(\|Tx - T_n x\| \to 0, T \to \infty)$$

这在数值分析中称为近似格式具有相容性.

同时数值分析中还有稳定性的概念: 称近似格式具有稳定性, 是指

$$\exists C > 0 \forall n \in \mathbb{N}(\|T_n^{-1}\| \leq C)$$

在相容性的前提下, Lax 等价定理说的是近似格式的收敛性与稳定性等价.

定理 3.4.11 (Lax 等价定理)

在方程

$$Tx = y \tag{3.16}$$

与

$$T_n x_n = y \tag{3.17}$$

中, 若

$$\forall x \in \mathscr{X}(\|Tx - T_n x\| \to 0, n \to \infty) \tag{3.18}$$

那么为了 $x_n \to x(n \to \infty)$, 其中 x_n, x 分别是(3.17)式与(3.16)式的解, 必须且仅须 $\exists C > 0$, 使得

$$\|T_n^{-1}\| \leq C \quad (\forall n \in \mathbb{N}) \tag{3.19}$$

成立.

证明

当 $\exists C > 0 \forall n \in \mathbb{N}(\|T_n^{-1}\| \leq C)$, 由 $\|Tx - T_n x\| \to 0 (n \to \infty)$ 知

$$\|x_n - x\| = \|T_n^{-1}y - T_n^{-1}T_n x\| = \|T_n^{-1}Tx - T_n^{-1}T_n x\| \leq \|T_n^{-1}\| \cdot \|Tx - T_n x\| \leq C\|Tx - T_n x\| \to 0 \quad (n \to \infty)$$

当 $x_n \to x(n \to \infty)$, 对任意的 $y \in \mathscr{Y}$, 根据(3.16)式与(3.17)式知 $x_n = T_n^{-1}y, x = T^{-1}y$, 进而

$$\forall y \in \mathscr{Y}(x_n \to x, n \to \infty) \Rightarrow \forall y \in \mathscr{Y}(T_n^{-1}y \to T^{-1}y, n \to \infty)$$

这说明 $\forall y \in \mathscr{Y}(\sup\limits_{n\in\mathbb{N}}\|T_n^{-1}y\| < \infty)$, 进而根据共鸣定理(3.4.8)知 $\|T_n{-}1\|$ 有界. □

3.4.1.7 补充: 共鸣定理的更多应用

本节选自 [**PGC**], 旨在介绍共鸣定理 (尤其是 Banach-Steinhaus 定理) 的更多应用.

Polya 定理: 数值求积公式的收敛性

给定一个权函数 $w \in L^1(0, 1)$, 现在希望对任意函数 $f \in C[0, 1]$, 用" 容易计算" 的有限和" 尽可能精确地" 逼

近积分

$$l(f) := \int_0^1 f(x)w(x)dx$$

其中一个自然的方式是考虑定积分的定义, 亦即固定 $n \in \mathbb{N}$, 适当选取 $n+1$ 个不同的节点 $0 \le x_0^n < x_1^n < \cdots < x_n^n \le 1$ 与 $n+1$ 个权 $\{w_j^n\}_{j=0}^n \in \mathbb{R}$, 用数值求积公式

$$l_n(f) := \sum_{j=0}^n w_k^n f(x_j^n)$$

来逼近积分 $l(f)$. 注意在 $C[0,1]$ 装备一致范数时, 上面定义的映射 $l: C[0,1] \to \mathbb{R}, l_n: C[0,1] \to \mathbb{R}$ 显然是连续线性泛函.

数值积分中的一个基本问题就是在选定权函数 w 与权 $\{w_j\}$ 后, 讨论数值求积公式的收敛性. 下述定理是这方面的一个很强的结果:

定理 3.4.12 (Polya)

给定权函数 $w \in L^1(0,1)$ 与连续线性泛函序列

$$l_n : C[0,1] \to \mathbb{R}, f \mapsto l_n(f) := \sum_{j=0}^n w_j^n f(x_j^n) \in \mathbb{R}, \quad 0 \le x_0^n < x_1^n < \cdots < x_n^n \le 1$$

如果序列 $\{l_n\}_{n\in\mathbb{N}}$ 满足:

$$\lim_{n\to\infty} |\int_0^1 p(x)w(x)dx - l_n(p)| = 0, \quad \forall p \in \mathcal{P}[0,1]$$

其中 $\mathcal{P}[0,1]$ 表示 [0,1] 上的全体多项式, 则

$$\lim_{n\to\infty} |\int_0^1 f(x)w(x)dx - l_n(f)| = 0, \quad \forall f \in C[0,1] \tag{3.20}$$

的充要条件是

$$\sup_{n\ge 0}(\sum_{j=0}^n |w_j^n|) < \infty. \tag{3.21}$$

♡

证明

当(3.20)式成立, 显见对任意 $f \in C[0,1]$ 有 $|l_n(f)| \le (\sum_{j=0}^n |w_j^n|)\|f\|$, 故 $\|l_n\| \le \sum_{j=0}^n |w_j^n|$. 另一方面, 设 $f_0 \in C[0,1]$ 是下述分段仿射连续函数

$$\begin{cases} f_0(0) = \operatorname{sgn} w_0^n \\ f_0(x_j^n) = \operatorname{sgn} w_j^n, \quad 0 \le j \le n \\ f_0(1) = \operatorname{sgn} w_n^n \end{cases}$$

则 $\|f_0\| = 1$, 且 $|l_n(f_0)| = \sum_{j=0}^n |w_j^n|$, 于是

$$\|l_n\| \ge \frac{|l_n(f_0)|}{\|f_0\|} = \sum_{j=0}^n |w_j^n|, \quad \forall n \in \mathbb{N}$$

综上有 $\|l_n\| = \sum_{j=0}^n |w_j^n|$. 根据 Banach-Steinhaus 定理, 此时有 $\sup_{n\ge 0}\|l_n\| < \infty$, 此即(3.21)式.

当(3.21)式成立, 任取 $f \in C[0,1], p \in \mathcal{P}[0,1]$, 有:

$$\begin{aligned}|l_n(f) - \int_0^1 f(x)w(x)dx| &\leq |l_n(f-p)| + |l_n(p) - \int_0^1 p(x)w(x)dx| + |\int_0^1 (f(x)-p(x))w(x)dx| \\ &\leq \sup_{n\geq 0}\|l_n\| \cdot \|f-p\| + |l_n(p) - \int_0^1 p(x)w(x)dx| + \|w\|_{L^1(0,1)}\|f-p\| \\ &= (\sup_{n\geq 0}\|l_n\| + \|w\|_{L^1(0,1)})\|f-p\| + |l_n(p) - \int_0^1 p(x)w(x)dx|\end{aligned}$$

现固定 $f \in C[0,1]$, 根据 Weierstrass 逼近定理知对任意 $\varepsilon > 0$ 存在多项式 $p(f;\varepsilon) \in \mathcal{P}[0,1]$ 使得

$$\|f-p\| < \varepsilon$$

根据序列 $\{l_n\}_{n\in\mathbb{N}}$ 的性质知存在 $N > 0$ 使得

$$\forall n > N(|l_n(p) - \int_0^1 p(x)w(x)dx| < \varepsilon)$$

于是

$$|l_n(f) - \int_0^1 f(x)w(x)dx| < (\sup_{n\geq 0}\|l_n\| + \|w\|_{L^1(0,1)})\varepsilon + \varepsilon, \quad \forall n > N$$

根据 ε 的任意性即得欲证. □

Fourier 级数的发散

约定 $C_{per}[0,2\pi]$ 表示所有连续的 2π 周期函数 $g \in \mathbb{R} \to \mathbb{R}$ 组成的 Banach 空间, 其装备一致收敛范数 $\|g\| = \sup_{\theta\in[0,2\pi]}|g(\theta)|$; $Q_n[0,2\pi]$ 表示所有阶数不大于 n 的实 2π 周期三角多项式构成的空间. 定义 g 的第 n 个 Fourier 部分和为:

$$S_n(g)(\theta) := \frac{a_0}{2} + \sum_{k=1}^n a_k \cos k\theta + \sum_{k=1}^n b_k \sin k\theta, \quad \theta \in [0,2\pi]$$

其中

$$\begin{aligned}a_k &:= \frac{1}{\pi}\int_0^{2\pi} g(\theta)\cos k\theta d\theta, \quad k \geq 0 \\ b_k &:= \frac{1}{\pi}\int_0^{2\pi} g(\theta)\sin k\theta d\theta, \quad k \geq 1\end{aligned}$$

Fourier 级数的发散指的是下述定理:

定理 3.4.13

存在函数 $g \in C_{per}[0,2\pi]$, 其第 n 个 Fourier 部分和 $S_n(g)$ 满足:

$$\sup_{n\geq 0}\|S_n(g)\| = \infty$$

从而 $S_n(g)$ 不可能一致收敛于 g.

证明

任取 $f \in C_{per}[0,2\pi]$, 知

$$
\begin{aligned}
S_n(f)(t) &= \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos kt + b_k \sin kt) \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(u)du + \sum_{k=1}^{\infty}(\frac{1}{\pi}(\int_0^{2\pi} f(u)\cos kudu)\cos kt + \frac{1}{\pi}(\int_0^{2\pi} f(u)\sin kudu)\sin kt) \\
&= \frac{1}{\pi}\int_0^{2\pi} f(u)(\frac{1}{2} + \sum_{k=1}^{n}(\cos ku\cos kt + \sin ku\sin kt))du \\
&= \frac{1}{\pi}\int_0^{2\pi} f(u)(\frac{1}{2} + \sum_{k=1}^{n}\cos k(u-t))du \\
&= \frac{1}{\pi}\int_0^{2\pi} f(s+t)\frac{\sin\left(n+\frac{1}{2}\right)s}{2\sin\frac{s}{2}}ds
\end{aligned}
$$

其中 $D_n(s) := \frac{\sin\left(n+\frac{1}{2}\right)s}{2\sin\frac{s}{2}}$ 称为 Dirhchlet 核. 现取 $t=0$, 得到

$$
S_n(f)(0) = \frac{1}{\pi}\int_0^{2\pi} f(s)D_n(s)ds
$$

容易验证对每个固定的 n, $S_n(f)(0)$ 都是 $C_{per}[0,2\pi]$ 到 $\mathbb{R}$ 的有界线性算子. 先前验证过 $\|S_n\| = \frac{1}{\pi}\int_0^{2\pi}|D_n(s)|ds$, 有:

$$
\begin{aligned}
\|S_n\| &= \frac{1}{\pi}\int_0^{2\pi}|D_n(s)|ds = \frac{2}{\pi}\int_0^{\pi}|D_n(s)|ds \\
&\geq \frac{2}{\pi}\int_0^{\pi}\frac{|\sin\left(n+\frac{1}{2}\right)s|}{2\cdot\frac{s}{2}}ds = \frac{2}{\pi}\int_0^{(n+\frac{1}{2})\pi}\frac{|\sin v|}{v}dv \\
&= \frac{2}{\pi}\sum_{k=0}^{n-1}\int_{k\pi}^{(k+1)\pi}\frac{|\sin v|}{v}dv + \frac{2}{\pi}\int_{n\pi}^{(n+\frac{1}{2})\pi}\frac{|\sin v|}{v}dv \\
&\geq \frac{2}{\pi}\sum_{k=0}^{n-1}\int_{k\pi}^{(k+1)\pi}\frac{|\sin v|}{(k+1)\pi}dv = \frac{4}{\pi^2}\sum_{k=0}^{n-1}\frac{1}{k+1} \to \infty, \quad n\to\infty
\end{aligned}
$$

故 $\sup\limits_{n\geq 0}\|S_n\| = \infty$, 因而由共鸣定理知存在 $C_{per}[0,2\pi]$, 使得 $\sup\limits_{n\geq 0}|S_n(f)(0)| = \infty$, 亦即 $\{S_n(f)(0)\}_{n\in\mathbb{N}}$ 不收敛. □

注 在 Fourier 分析中, 证明 $\|S_n\| \to \infty (n\to\infty)$ 实际上就是证明 Dirichlet 核不是好核. 具体想法可参见 **[ST1]**.

3.4.2 习题

练习 3.15 设 $\mathscr{X}$ 是 B 空间, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的闭子空间. 映射 $\varphi:\mathscr{X}\to\mathscr{X}/\mathscr{X}_0$ 定义为

$$
\varphi: x\mapsto [x] \quad (\forall x\in\mathscr{X})
$$

其中 $[x]$ 表示含 x 的商类, 求证 φ 是开映射.

证明

已经知道 $\mathscr{X}/\mathscr{X}_0$ 是 B 空间, 根据开映射定理知只需证明 φ 是满射. 任取 $[x]\in\mathscr{X}/\mathscr{X}_0$, 根据 φ 的构造即知 $[x]$ 存在原像 $x\in\mathscr{X}$, 故 φ 是满射, 进而是开映射.

练习 3.16★ 设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, 又设方程 $Ux=y$ 对 $\forall y\in\mathscr{Y}$ 有解 $x\in\mathscr{X}$, 其中 $U\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 并且 $\exists m>0$, 使得

$$
\|Ux\| \geq m\|x\| \quad (\forall x\in\mathscr{X})
$$

求证: U 有连续逆 U^{-1}, 并且 $\|U^{-1}\| \leq \frac{1}{m}$.

证明

要说明 U 有连续逆, 根据 Banach 定理只需说明 U 既是满射又是单射. U 是满射可由方程 $Ux=y$ 对任意

$y \in \mathscr{Y}$ 推知, 而 U 是单射是因为若 $x_1 \neq x_2$, 则

$$\|Ux_1 - Ux_2\| \geq m\|x_1 - x_2\| > 0 \Rightarrow Ux_1 \neq Ux_2$$

故 U 有连续逆 $U^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$. 既然对每个 $y \in \mathscr{Y}$ 总有对应的解 $x \in \mathscr{X}$, 知每个 $x \in \mathscr{X}$ 总能写成 $U^{-1}y, y \in \mathscr{Y}$ 的形式, 代入条件有

$$\|y\| \geq m\|U^{-1}y\| \Rightarrow \|U^{-1}\| \leq \frac{1}{m}$$

练习 3.17★ 设 H 是 Hilbert 空间, $A \in \mathscr{L}(H)$, 并且 $\exists m > 0$, 使得

$$|(Ax, x)| \geq m\|x\|^2 \quad (\forall x \in H)$$

求证: $\exists A^{-1} \in \mathscr{L}(H)$.

证明

首先说明 A 是满射, 进一步只需说明 $R(A)$ 是闭集, 且 $R(A)^{\perp} = \{\theta\}$ 即可. 因为 $A \in \mathscr{L}(H)$, 故 $D(A)$ 总可以看成闭集, 下面说明 $R(A)$ 是闭集. 任取 $y \in \overline{R(A)}$, 根据闭包的定义知存在 $\{y_n\} = \{Ax_n\} \subset R(A)$ 使得

$$y = \lim_{n\to\infty} Ax_n$$

根据 $\{Ax_n\}_{n\in\mathbb{A}} \subset H$ 收敛知其为基本列, 因而根据 Cauchy 准则有

$$\forall \varepsilon > 0 \exists N > 0 \forall m > n > N(\|Ax_m - Ax_n\| < \varepsilon)$$

在题目条件中代入 $x = x_m - x_n$ 并利用 Cauchy-Schwarz 不等式得到

$$m\|x_m - x_n\|^2 \leq |(A(x_m - x_n), x_m - x_n)| \leq \|A(x_m - x_n)\|\|x_m - x_n\|$$

因而

$$m\|x_m - x_n\| \leq \|Ax_m - Ax_n\| < \varepsilon \Rightarrow \|x_m - x_n\| < \frac{\varepsilon}{m}$$

故 $\{x_n\}_{n\in\mathbb{N}}$ 是 H 中的基本列, 由 $D(A)$ 的闭性知其存在极限 x, 现根据 A 的连续性有

$$y = \lim_{n\to\infty} Ax_n = A(\lim_{n\to\infty} x_n) = Ax$$

从而 $y \in R(A)$, 亦即 $R(A)$ 是闭集.

再设 $y \in H$ 使得 $\forall x \in H((Ax, y) = 0)$, 取 $x = y$ 知 $0 = |(Ay, y)| \geq m\|y\|^2$, 因而只能有 $y = \theta$, 故 $R(A)^{\perp} = \{\theta\}$, 亦即 $R(A) = H$.

最后证明 A 是单射, 由 Cauchy-Schwarz 不等式知

$$m\|x\|^2 \leq |(Ax, x)| \leq \|Ax\| \cdot \|x\| \Rightarrow \|Ax\| \geq m\|x\|$$

故若 $x_1 \neq x_2$, 则

$$\|Ax_1 - Ax_2\| \geq m\|x_1 - x_2\| > 0 \Rightarrow Ax_1 \neq Ax_2$$

这说明 A 是单射, 因而由 Banach 定理知 A 有连续逆 $A^{-1} \in \mathscr{L}(H)$.

练习 3.18★ 设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, D 是 $\mathscr{X}$ 的线性子空间, 并且 $A: D \to \mathscr{Y}$ 是线性映射. 求证:

(1) 如果 A 连续且 D 是闭的, 那么 A 是闭算子.

证明

由条件知此时 D 是闭线性子空间, 也即 D 中的任意基本列都在 D 中收敛. 由 A 连续知存在 $M > 0$ 使得

$$\|Ax\| \leq M\|x\|, \quad \forall x \in D$$

故若在 D 中 $x_n \to x(n \to \infty)$, 则

$$\|Ax_n - Ax\| \leq M\|x_n - x\| \to 0, \quad n \to \infty$$

这说明 $Ax_n \to Ax$, 若另设 $Ax_n \to y \in \mathscr{Y}$, 显见 $y = Ax$, 故 A 是闭算子.

(2) 如果 A 连续且是闭算子, 那么 $\mathscr{Y}$ 完备蕴含 D 闭.

证明

由 A 连续知存在 $M > 0$ 使得

$$\|Ax\| \le M\|x\|, \quad \forall x \in D$$

任取 D 中的基本列 $\{x_n\}$, 根据基本列的定义知

$$\|x_{n+p} - x_n\| \to 0, \quad n \to \infty, \forall p \in \mathbb{N}$$

得到

$$\|Ax_{n+p} - Ax_n\| \le M\|x_{n+p} - x_n\| \to 0, \quad n \to \infty, \forall p \in \mathbb{N}$$

这说明 $\{Ax_n\}$ 是 $\mathscr{Y}$ 中的基本列, 由 $\mathscr{Y}$ 的完备性可设 $Ax_n \to y \in \mathscr{Y}(n \to \infty)$. 当 $x_n \to x(n \to \infty)$, 根据 A 的闭性知 $x \in D, y = Ax$, 这说明 D 是闭的.

(3) 如果 A 是单射的闭算子, 那么 A^{-1} 也是闭算子;

证明

由 A 是单射知 A^{-1} 良定义. 根据 A 是闭算子的定义知:

$$D \ni x_n \to x, Ax_n \to y \Rightarrow x \in D, y = Ax$$

取 $x_n = A^{-1}y_n, y_n \in \mathscr{Y}$ 知

$$D \ni A^{-1}y_n \to x, y_n \to y \Rightarrow x \in D, y = Ax$$

既然 $y = Ax$, 知至少有 $y \in \mathscr{Y}$, 且 $x = A^{-1}y$, 这说明 A^{-1} 也是闭算子.

(4) 如果 $\mathscr{X}$ 完备, A 是单射的闭算子, $R(A)$ 在 $\mathscr{Y}$ 中稠密, 并且 A^{-1} 连续, 那么 $R(A) = \mathscr{Y}$.

证明

既然 A 是单射的闭算子, 由 (3) 知 A^{-1} 也是闭算子, 又因为 A^{-1} 连续且 $\mathscr{X}$ 完备, 故由 (2) 知 $R(A)$ 是闭的, 也即 $R(A) = \overline{R(A)} = \mathscr{Y}$.

✍ **练习 3.19★** 用等价范数定理(3.4.1)证明: $(C[0,1], \|\cdot\|_1)$ 不是 B 空间, 其中 $\|f\|_1 = \int_0^1 |f(t)|dt, \forall f \in C[0,1]$.

证明

用反证法. 若 $(C[0,1], \|\cdot\|_1)$ 是 B 空间, 已知 $(C[0,1], \|\cdot\|)$ 也是 B 空间, 其中

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

且

$$\|f\|_1 = \int_0^1 |f(t)|dt \le \int_0^1 \max_{x \in [0,1]} |f(t)|dt \max_{x \in [0,1]} |f(x)| = \|f\|$$

故 $\|\cdot\|$ 比 $\|\cdot\|_1$ 强, 由等价范数定理知 $\|\cdot\|$ 与 $\|\cdot\|_1$ 应等价, 往证不存在常数 $M > 0$ 使得

$$\|f\|_1 \ge M\|f\|, \quad \forall f \in C[0,1]$$

亦即证明对任意 M 总存在 $f \in C[0,1]$ 使得

$$\|f\|_1 < M\|f\|.$$

事实上, 若 $M \ge 1$, 则构造

$$C[0,1] \ni f(x) = -M^2(x-1)$$

知 $\|f\|_1 = \frac{M^2}{2} < M\|f\|$. 而若 $M < 1$, 构造

$$C[0,1] \ni f(x) = \begin{cases} -M(x-M), & x \in [0, M] \\ 0, & x \in (M, 1] \end{cases}$$

同样有 $\|f\|_1 < M\|f\|$, 故 $\|\cdot\|_1$ 不可能比 $\|\cdot\|$ 强, 矛盾! 故 $(C[0,1], \|\cdot\|_1)$ 不是 B 空间.

练习 3.20(Gelfand 引理) 设 $\mathscr{X}$ 是 B 空间, $p: \mathscr{X} \to \mathbb{R}$ 满足

1. $p(x) \geq 0 (\forall x \in \mathscr{X})$;
2. $p(\lambda x) = \lambda p(x) (\forall \lambda > 0, \forall x \in \mathscr{X})$;
3. $p(x_1 + x_2) \leq p(x_1) + p(x_2) (\forall x_1, x_2 \in \mathscr{X})$;
4. 当 $x_n \to x$ 时, $\varliminf\limits_{n\to\infty} p(x_n) \geq p(x)$.

求证: $\exists M > 0$, 使得 $p(x) \leq M\|x\|, \forall x \in \mathscr{X}$.

证明

取

$$\|x\|_p := \|x\| + \sup_{|\lambda|=1} |p(\lambda x)|$$

首先说明 $\sup\limits_{|\lambda|=1} |p(\lambda x)|$ 是 $\mathscr{X}$ 上的范数, 正定性和三角不等式显然, 对齐次性知

$$\sup_{|\lambda|=1} |p(\lambda \cdot \mu x)| = \sup_{|\operatorname{sgn}(\mu)\lambda|=1} |p(\operatorname{sgn}(\mu)\lambda \cdot \mu x)| = \sup_{|\lambda|=1} |\mu p(\lambda x)| = |\mu| \sup_{|\lambda|=1} |p(\lambda x)|$$

从而易证 $\|\cdot\|_p$ 也是 $\mathscr{X}$ 上的范数, 且 $\|\cdot\|_p$ 强于 $\|\cdot\|$, 下面说明 $(\mathscr{X}, \|\cdot\|_p)$ 完备. 取 $\{x_n\}$ 是 $(\mathscr{X}, \|\cdot\|_p)$ 中的基本列, 根据定义知

$$\|x_{n+m} - x_n\| + \sup_{|\lambda|=1} |p(\lambda(x_{n+m} - x_n))| \to 0, \quad n \to \infty, \forall m \in \mathbb{N}$$

因而至少有

$$\|x_{n+m} - x_n\| \to 0, \quad n \to \infty, \forall m \in \mathbb{N}$$

因为 $(\mathscr{X}, \|\cdot\|)$ 完备, 故可设 $x_n \to x \in \mathscr{X}$, 从而只需说明

$$\sup_{|\lambda|=1} |p(\lambda(x_n - x))| \to 0, \quad n \to \infty$$

即可. 因为已经知道

$$\sup_{|\lambda|=1} |p(\lambda(x_{n+m} - x_n))| \to 0, \quad n \to \infty, \forall m \in \mathbb{N}$$

且

$$\lim_{m\to\infty} (x_n - x_{n+m}) = x_n - x$$

故

$$p(\lambda(x_n - x)) \leq \varliminf_{m\to\infty} p(\lambda(x_n - x_{n+m})) \leq \sup_{|\lambda|=1} |p(\lambda(x_{n+m} - x_n))| \to 0, \quad n \to \infty, \forall |\lambda| = 1$$

得到

$$\sup_{|\lambda|=1} |p(\lambda(x_n - x))| \to 0, \quad n \to \infty$$

故 $(\mathscr{X}, \|\cdot\|_p)$ 完备, 因而由等价范数定理知存在 $M > 1$ 使得

$$\|x\| + \sup_{|\lambda|=1} |p(\lambda x)| \leq M\|x\| \Rightarrow p(x) \leq \sup_{|\lambda|=1} |p(\lambda x)| \leq (M-1)\|x\|.$$

练习 3.21★ 设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是 B 空间, $A_n \in \mathscr{L}(\mathscr{X}, \mathscr{Y}) (n = 1, 2, \cdots)$, 又对 $\forall x \in \mathscr{X}$, $\{A_n x\}$ 在 $\mathscr{Y}$ 中收敛. 求证:

$\exists A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 使得

$$A_n x \to Ax(\forall x \in \mathscr{X}), \quad 且\|A\| \leq \varliminf_{n\to\infty} \|A_n\|.$$

证明

取

$$Ax := \lim_{n\to\infty} A_n x, \quad \forall x \in \mathscr{X}$$

现在证明 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. 既然 $A_n \in \mathscr{L}(\mathscr{X}, \mathscr{Y})(n = 1, 2, \cdots)$, 知 A 首先是线性的. 现在说明 A 是闭算子, 也即

$$\mathscr{X} \ni x_n \to x, Ax_n \to y \Rightarrow x \in \mathscr{X}, y = Ax$$

因为 $\mathscr{X}$ 是 B 空间, 显见 $x \in D$. 任取趋向 x 的基本列 $\{x_p\}$, 下面证明 $\{Ax_p\}$ 是基本列. 这是因为根据基本列的定义本身有

$$\|x_{p+q} - x_p\| \to 0, \quad p \to \infty, \forall q \in \mathbb{N}$$

既然对每个 $x \in \mathscr{X}$, $\{A_n x\}$ 均收敛, 根据一致有界定理知存在 $M > 0$ 使得 $\forall n \in \mathbb{N}(\|A_n\| \leq M)$, 得到

$$\|A_n x_{p+q} - A_n x_p\| \leq \|A_n\| \cdot \|x_{p+q} - x_p\| \leq M\|x_{p+q} - x_p\| \to 0, \quad p \to \infty, \forall q, n \in \mathbb{N}$$

故令 $n \to \infty$ 得到

$$\|Ax_{p+q} - Ax_p\| \leq M\|x_{p+q} - x_p\| \to 0, \quad p \to \infty, \forall q \in \mathbb{N}$$

这说明 $\{Ax_p\}$ 是基本列, 又因为 $\mathscr{Y}$ 是 B 空间, 故只能有 $y = Ax$. 这说明 A 是闭线性算子, 又根据 $\mathscr{X}$ 是 B 空间与闭图像定理知 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$.

最后说明 $\|A\| \leq \varliminf_{n\to\infty} \|A_n\|$, 这是因为任取 $x \in \mathscr{X}$ 有[6]:

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| = \varliminf_{n\to\infty} \|A_n x\| \leq \varliminf_{n\to\infty} \|A_n\| \cdot \|x\|$$

故只要 $x \neq \theta$, 则

$$\frac{\|Ax\|}{\|x\|} = \|A\| \leq \varliminf_{n\to\infty} \|A_n\|$$

练习 3.22★ 设 $1 < p < \infty$, 并且 $\frac{1}{p} + \frac{1}{q} = 1$. 如果序列 $\{\alpha_k\}$ 使得对 $\forall x = \{\xi_k\} \in l^p$ 保证 $\sum\limits_{k=1}^{\infty} \alpha_k \xi_k$ 收敛, 求证: $\{\alpha_k\} \in l^q$. 又若 $f : x \mapsto \sum\limits_{k=1}^{\infty} \alpha_k \xi_k$, 求证: f 作为 l^p 上的线性泛函, 有

$$\|f\| = (\sum_{k=1}^{\infty} |\alpha_k|^q)^{\frac{1}{q}}.$$

证明

记 $\alpha^n = (\alpha_1, \cdots, \alpha_n, 0, \cdots), \alpha = \{\alpha_k\}_{k\in\mathbb{N}}$, 显见 $\alpha^n \in l^q$. 现定义

$$f_n(x) = \sum_{k=1}^{n} \alpha_k \xi_k, \quad \forall x = \{\xi_k\}_{k\in\mathbb{N}} \in l^p$$

由 Hölder 不等式知

$$|f_n(x)| \leq \sum_{k=1}^{n} |\alpha_k \xi_k| \leq (\sum_{k=1}^{n} |\alpha_k|^q)^{\frac{1}{q}} (\sum_{k=1}^{n} |\xi_k|^p)^{\frac{1}{p}} \leq \|\alpha^n\|_{l^q} \|x\|_{l^p} \Rightarrow \|f_n\| \leq \|\alpha^n\|_{l^q}$$

于是 f_n 是 l^p 上的有界线性泛函. 再取

$$x^n = \left(\frac{|\alpha_1|^{q-1} \operatorname{sgn}(\alpha_1)}{(\sum\limits_{k=1}^{n} |\alpha_k|^q)^{\frac{1}{p}}}, \cdots, \frac{|\alpha_n|^{q-1} \operatorname{sgn}(\alpha_n)}{(\sum\limits_{k=1}^{n} |\alpha_k|^q)^{\frac{1}{p}}}, 0, \cdots \right)$$

[6]这里第一步就相当于对点列 $\{a_n\}$ 而言, 若 $a_n \to a(n \to \infty)$, 则 $\|\lim\limits_{n\to\infty} a_n\| = \lim\limits_{n\to\infty} \|a_n\|$.

知

$$\|x^n\|_{l^p} = \sum_{k=1}^{n} \frac{|\alpha_k|^{pq-p}}{\sum\limits_{k=1}^{n}|\alpha_k|^q} = \sum_{k=1}^{n} \frac{|\alpha_k|^q}{\sum\limits_{k=1}^{n}|\alpha_k|^q} = 1$$

从而 $x^n \in l^p$, 有

$$|f_n(x_n)| = \sum_{k=1}^{n} \frac{|\alpha_k|^q}{(\sum\limits_{k=1}^{n}|\alpha_k|^q)^{\frac{1}{p}}} = (\sum_{k=1}^{n}|\alpha_k|^q)^{1-\frac{1}{p}} = (\sum_{k=1}^{n}|\alpha_k|^q)^{\frac{1}{q}} = \|\alpha^n\|_{l^q}$$

故 $\|f_n\| = \|\alpha^n\|_{l^q}$. 现因对任意 $x \in l^q$, 均有

$$\lim_{n\to\infty} f_n(x) = \sum_{k=1}^{\infty} \alpha_k \xi_k \text{收敛}$$

故根据一致有界定理知 $\sup\limits_{n\geq 1}\{\|f_n\|\} = \sup\limits_{n\geq 1}\|\alpha^n\|_{l^q} < \infty$, 亦即 $\|\alpha\|_{l^q} < \infty$, 故 $\alpha \in l^q$. 根据练习3.21的结果知

$$\|f\| \leq \varliminf_{n\to\infty} \|f_n\| = \|\alpha\|_{l^q}$$

再取

$$x = \left\{ \frac{|\alpha_k|^{q-1}\operatorname{sgn}(\alpha_k)}{(\sum\limits_{k=1}^{n}|\alpha_k|^q)^{\frac{1}{p}}} \right\}_{k\in\mathbb{N}}$$

仿照前述过程可以说明 $x \in l^p, \|x\|_{l^p} = 1$, 代入 f 有

$$|f(x)| = \sum_{k=1}^{\infty} \frac{|\alpha_k|^q}{(\sum\limits_{k=1}^{n}|\alpha_k|^q)^{\frac{1}{p}}} = (\sum_{k=1}^{\infty}|\alpha_k|^q)^{1-\frac{1}{p}} = (\sum_{k=1}^{\infty}|\alpha_k|^q)^{\frac{1}{q}} = \|\alpha\|_{l^q}$$

故 $\|f\| = \|\alpha\|_{l^q}$, 命题得证.

练习 3.23 如果序列 $\{\alpha_k\}$ 使得对 $\forall x = \{\xi_k\} \in l^1$, 保证 $\sum\limits_{k=1}^{\infty}\alpha_k\xi_k$ 收敛, 求证: $\{\alpha_k\} \in l^\infty$. 又若 $f : x \mapsto \sum\limits_{k=1}^{\infty}\alpha_k\xi_k$ 作为 l^1 上的线性泛函, 求证:

$$\|f\| = \sup_{k\geq 1}|\alpha_k|.$$

证明

记 $\alpha^n = (\alpha_1, \cdots, \alpha_n, 0, \cdots), \alpha = \{\alpha_k\}_{k\in\mathbb{N}}$, 显见 $\alpha^n \in l^\infty$. 现定义

$$f_n(x) = \sum_{k=1}^{n}\alpha_k\xi_k, \quad \forall x = \{\xi_k\}_{k\in\mathbb{N}} \in l^1$$

知

$$|f_n(x)| \leq \sum_{k=1}^{n}|\alpha_k\xi_k| \leq \sup_{1\leq k\leq n}|\alpha_k| \cdot \sum_{k=1}^{n}|\xi_k| \leq \|\alpha^n\|_{l^\infty}\|x\|_{l^1} \Rightarrow \|f_n\| \leq \|\alpha^n\|_{l^\infty}$$

于是 f_n 是 l^1 上的有界线性泛函. 因为 $\sup\limits_{1\leq k\leq n}|\alpha_k|$ 是有限个数中取上确界, 故总可以找到指标 i 使得 $|\alpha_i| = \sup\limits_{1\leq k\leq n}|\alpha_k|$, 取 $x_n = \operatorname{sgn}\alpha_i e_i$ 是第 i 项为 $\operatorname{sgn}\alpha_i$, 其余项均为 0 的元素, 显见 $x_n \in l^1, \|x_n\|_{l^1} = 1$, 且

$$|f_n(x_n)| = |\alpha_i| = \sup_{1\leq k\leq n}|\alpha_k| = \|\alpha^n\|_{l^\infty}$$

故 $\|f_n\| = \|\alpha^n\|_{l^\infty}$. 现因对任意 $x \in l^1$, 均有

$$\lim_{n\to\infty} f_n(x) = \sum_{k=1}^{\infty} \alpha_k \xi_k \text{收敛}$$

故根据一致有界定理知 $\sup\limits_{n\geq 1}\{\|f_n\|\} = \sup\limits_{n\geq 1}\|\alpha^n\|_{l^\infty} < \infty$, 亦即 $\|\alpha\|_{l^\infty} < \infty$, 故 $\alpha \in l^\infty$. 根据练习3.21的结果知

$$\|f\| \leq \varliminf_{n\to\infty} \|f_n\| = \|\alpha\|_{l^\infty}$$

若存在指标 i 使得 $|\alpha_i| = \|\alpha\|_{l^\infty}$, 则取 $x = \operatorname{sgn}\alpha_i e_i$, 知 $x \in l^1, \|x\|_{l^1} = 1$, 代入 f 有

$$|f(x)| = |\alpha_i| = \|\alpha\|_{l^\infty}$$

而若不存在指标 i 使得 $|\alpha_i| = \|\alpha\|_{l^\infty}$, 则必存在子列 $\{\alpha_{n_r}\}$ 使得 $\lim\limits_{r\to\infty} |\alpha_{n_r}| = \|\alpha\|_{l^\infty}$, 简洁起见不妨就设

$$\lim_{k\to\infty} |\alpha_k| = \|\alpha\|_{l^\infty}$$

根据极限的定义知

$$\forall\varepsilon > 0 \exists N > 0 \forall k > N(||\alpha_k| - \|\alpha\|_{l^\infty}| < \varepsilon \Rightarrow |\alpha_k| > \|\alpha\|_{l^\infty} - \varepsilon)$$

此时令

$$x_k = \operatorname{sgn}\alpha_k e_k$$

则 $x_k \in l^1, \|x_k\|_{l^1} = 1$, 且

$$|f(x_k)| = |\alpha_k| \ge \|\alpha\|_{l^\infty} - \varepsilon$$

由 ε 的任意性知 $\|f\| \ge \|\alpha\|_{l^\infty}$, 从而 $\|f\| = \|\alpha\|_{l^\infty}$, 命题即证.

练习 3.24 用 Gelfand 引理 (练习(3.20)) 证明共鸣定理 (定理(3.4.8)).

- 证明

取 $p(x) = \sup\limits_{A\in W} \|Ax\| < \infty, \forall x \in \mathscr{X}$. 现在证明 $p(x)$ 满足 Gelfand 引理(3.20)的条件.

显见 $p(x) \ge 0, p(\lambda x) = \lambda p(x), \forall x \in \mathscr{X}, \forall\lambda > 0$. 由 $\|\cdot\|$ 的三角不等式易得 $p(x_1+x_2) \le p(x_1)+p(x_2), \forall x_1, x_2 \in \mathscr{X}$. 现需证明当 $x_n \to x$, 有 $\varliminf\limits_{n\to\infty} p(x_n) \ge p(x)$[7].

练习 3.25 设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间, $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 是满射的. 求证: 如果在 $\mathscr{Y}$ 中 $y_n \to y_0$, 则 $\exists C > 0$ 与 $x_n \to x_0$, 使得 $Ax_n = y_n$, 且 $\|x_n\| \le C\|y_n\|$.

- 证明

考虑 $N(A) = \{x \in \mathscr{X} : Ax = \theta\}$, 由 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 知 $N(A)$ 是 $\mathscr{X}$ 的闭线性子空间, 从而可以定义商空间 $\mathscr{X}/N(A)$, 再在 $\mathscr{X}/N(A)$ 中定义

$$A^* : \mathscr{X}/N(A) \ni [x] \mapsto Ax \in \mathscr{Y}$$

可以证明 $A^* \in \mathscr{L}(\mathscr{X}/N(A), \mathscr{Y})$. 由 A 是满射知 A^* 是满射. 同时 A^* 是单射, 这是因为当 $[x_1] \ne [x_2]$, 根据商空间的构造知 $Ax_1 \ne Ax_2$, 也即 $A^*[x_1] \ne A^*[x_2]$. 已经知道 $\mathscr{X}/N(A)$ 依旧是 B 空间, 故由 Banach 定理知 A^* 有连续逆, 设 $C > 0$ 使得 $\|(A^*)^{-1}\| \le C$. 当 $y_n \to y_0$, 根据连续性知存在 $\{[x_n]\} \subset \mathscr{X}/N(A)$, 使得 $[x_n] \to [x_0] = (A^*)^{-1}y_0, n \to \infty$. 现在说明从 $\{[x_n]\}$ 中可以找出 $\mathscr{X}$ 中的收敛列, 为此考虑 $[x_n]$ 已经是 $\mathscr{X}/N(A)$ 中的基本列, 根据定义有:

$$\|[x_{n+p}] - [x_n]\|_0 = \inf_{\alpha_n\in N(A)} \|x_{n+p} - \alpha_{n+p} - (x_n - \alpha_n)\| \to 0, \quad n \to \infty, \forall p \in \mathbb{N}$$

根据下确界的定义知存在 α_{n_k}, 使得

$$\|x_{n_k+p} - \alpha_{n_k+p} - (x_{n_k} - \alpha_{n_k})\| \to 0, \quad k \to \infty, \forall p \in \mathbb{N}$$

现取 $x_k^* = x_{n_k} - \alpha_{n_k}$, 则 $\{x_k^*\}$ 是 $\mathscr{X}$ 中的基本列, 由 $\mathscr{X}$ 的完备性与 $[x_n] \to [x_0]$ 知存在 $x_0 \in [x_0]$ 使得 $x_n \to x_0, n \to \infty$. 最后, 由 $y_n = Ax_n$ 知[8]

练习 3.26★ 设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间, T 是闭线性算子, $D(T) \subset \mathscr{X}, R(T) \subset \mathscr{Y}, N(T) := \{x \in \mathscr{X} : Tx = \theta\}$.

(1) 求证: $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间.

证明

显见 $N(T)$ 是 $\mathscr{X}$ 的线性子空间. 要证明 $N(T)$ 的闭性, 则需证明若 $N(T) \ni x_n \to x, n \to \infty$, 则 $x \in N(T)$. 为此

[7] 到这一步不会证了

[8] 最后的不等式不会证了, 感觉这里定义的 x_n 还是有问题.

考虑 T 作为闭线性算子的定义:

$$D(T) \ni x_n \to x, Tx_n \to y \Rightarrow x \in D(T), y = Tx$$

现在已知 $Tx_n = \theta, \forall n \in \mathbb{N}$, 故 $y = Tx = \theta$, 也即 $x \in N(T)$, 故 $N(T)$ 是闭的.

(2) 求证: $N(T) = \{\theta\}, R(T)$ 在 $\mathscr{Y}$ 中闭的充要条件是, $\exists \alpha > 0$, 使得

$$\|x\| \le \alpha \|Tx\| \quad (\forall x \in D(T)).$$

证明

既然 T 是闭线性算子, 故 $D(T)$ 是 $\mathscr{X}$ 的闭线性子空间, 因而可将 T 视作 $D(T)$ 到 $\mathscr{Y}$ 的算子, 且根据闭图像定理知 $T \in \mathscr{L}(D(T), \mathscr{Y})$.

当 $N(T) = \{\theta\}, R(T)$ 在 $\mathscr{Y}$ 中闭, 知此时 $R(T)$ 是 $\mathscr{Y}$ 的闭线性子空间, 故考虑 $T \in \mathscr{L}(D(T), R(T))$. 显见在这个意义下 T 是满射, 且由 $N(T) = \{\theta\}$ 知 T 是单射, 故由 Banach 定理知 T 有连续逆, 设 $\exists \alpha > 0$ 使得 $\|T^{-1}\| \le \alpha$, 得到

$$\|x\| = \|T^{-1}Tx\| \le \|T^{-1}\| \cdot \|Tx\| \le \alpha \|Tx\|, \quad \forall x \in D(T)$$

当 $\exists \alpha > 0 \forall x \in D(T)(\|x\| \le \alpha \|Tx\|)$, 任取 $x_1 \ne x_2$, 得到

$$\|Tx_1 - Tx_2\| \ge \frac{1}{\alpha}\|x_1 - x_2\| > 0$$

这说明 T 是单射, 因而 $N(T) = \{\theta\}$. 下面证明 $R(T)$ 在 $\mathscr{Y}$ 中闭, 任取 $R(T)$ 中的基本列 $\{y_n\}$, 既然 $\mathscr{Y}$ 是 B 空间, 可设 $y_n \to y \in \mathscr{Y}, n \to \infty$, 现需证明 $y \in R(T)$. 设 $x_n = T^{-1}y_n$, 得到

$$\|x_{n+p} - x_n\| \le \alpha \|y_{n+p} - y_n\| \to 0, \quad n \to \infty, \forall p \in \mathbb{N}$$

故 $\{x_n\}$ 是 $D(T)$ 中的基本列, 根据 $D(T)$ 的闭性可设 $x_n \to x \in D(T), n \to \infty$, 再根据 T 的闭性知 $y = Tx \in R(T)$, 故 $R(T)$ 在 $\mathscr{Y}$ 中闭.

(3) 如果用 $d(x, N(T))$ 表示点 $x \in \mathscr{X}$ 到集合 $N(T)$ 的距离 ($\inf\limits_{z \in N(T)} \|z - x\|$). 求证: $R(T)$ 在 $\mathscr{Y}$ 中闭的充要条件是, $\exists \alpha > 0$, 使得

$$d(x, N(T)) \le \alpha \|Tx\| \quad (\forall x \in D(T))$$

证明

定义

$$T^* : D(T)/N(T) \ni [x] \mapsto Tx \in R(T)$$

易证 $T^* \in \mathscr{L}(D(T)/N(T), R(T))$. 当 $R(T)$ 在 $\mathscr{Y}$ 中闭, 则此时 $R(T)$ 是 B 空间, 又已知 $D(T)/N(T)$ 是 B 空间, 故由 Banach 定理知 T^* 有连续逆, 设 $\exists \alpha > 0(\|(T^*)^{-1}\| \le \alpha)$, 得到

$$d(x, N(T)) = \|[x]\| = \|(T^*)^{-1}T^*[x]\| \|(T^*)^{-1}\| \cdot \|T^*[x]\| \le \alpha \|Tx\|, \quad (\forall x \in D(T))$$

当 $\exists \alpha > 0 \forall x \in D(T)(\|[x]\| = d(x, N(T)) \le \alpha \|Tx\| = \alpha \|T^*[x]\|)$, 根据 (2) 的结论知此时 $N(T^*) = \{\theta\}, R(T^*)$ 在 $\mathscr{Y}$ 中闭. 根据 T^* 的构造知 $R(T^*) = R(T)$, 故 $R(T)$ 在 $\mathscr{Y}$ 中闭.

✍ **练习 3.27★** 设 $a(x, y)$ 是 Hilbert 空间 H 上的一个共轭双线性泛函, 满足:

1. $\exists M > 0$, 使得 $|a(x, y)| \le M\|x\| \cdot \|y\| (\forall x, y \in H)$;
2. $\exists \delta > 0$, 使得 $|a(x, x)| \ge \delta \|x\|^2 (\forall x \in H)$.

求证: $\forall f \in H^*, \exists | y_f \in H$, 使得

$$a(x, y_f) = f(x) \quad (\forall x \in H)$$

而且 y_f 连续地依赖于 f.

证明

由 Lax-Milgram 定理, 存在 $A \in \mathscr{L}(H)$ 使得

$$a(x,y) = (x, Ay), \quad \|A^{-1}\| \le \frac{1}{\delta}$$

现对任意的 $f \in H^*$, 由 Riesz 表示定理知存在 $z_f \in H$ 使得 $f(x) = (x, z_f)$. 现取 $y_f = A^{-1}z_f$, 知 $f(x) = a(x, y_f), \forall x \in H$. 再证明 y_f 连续地依赖于 f. 知当 $\|f\| \to 0$, 有 $\|z_f\| \to 0$, 得到

$$\|y_f\| \le \|A^{-1}\| \cdot \|z_f\| \le \frac{1}{\delta}\|z_f\| \to 0$$

故 y_f 连续地依赖于 f.

练习 3.28 设 Ω 是 $\mathbb{R}^2$ 中边界光滑的有界开区域, $\alpha : \Omega \to \mathbb{R}$ 有界可测并满足 $0 < \alpha_0 \le \alpha, f \in L^2(\Omega)$. 规定:

$$a(u,v) := \int_\Omega (\nabla u \cdot \nabla v + auv)dxdy \quad (\forall u, v \in H^1(\Omega))$$

$$F(v) := \int_\Omega f \cdot vdxdy \quad (\forall v \in L^2(\Omega))$$

求证: $\exists | u \in H^1(\Omega)$ 满足

$$a(u,v) = F(v) \quad (\forall v \in H^1(\Omega))$$

●[9]证明

3.5 Hahn-Banach 定理

3.5.1 知识梳理

给定无穷维赋范线性空间 $\mathscr{X}$, Hahn-Banach 定理的动机在于说明是否有当 $x_1 \ne x_2, x_1, x_2 \in \mathscr{X}$ 时, 必定存在 $\mathscr{X}$ 上的一个连续线性泛函 $f(\cdot)$, 使得 $f(x_1) \ne f(x_2)$. 下面从线性泛函的延拓入手来解决这个问题.

课堂笔记 (♡)

高志强老师与 [**JD1**] 均介绍了 Hahn-Banach 定理的最初动机. 最初, E.Schmidt 于 1908 年研究了 l^2 中无穷维线性方程组的一个问题: 对于方程组

$$(\alpha_n, x) = c_n (n = 1, 2, \cdots) \tag{3.22}$$

其中 $\{\alpha_n\}_{n\in\mathbb{N}}$ 是 l^2 中线性无关的向量, $\{c_n\}_{n\in\mathbb{N}}$ 是给定的常数列, 希望证明如果存在常数 $M > 0$ 使得对任意 $n \in \mathbb{N}$ 与序列 $\{\lambda_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$, 均有:

$$|\sum_{k=1}^n \lambda_k c_k| \le M\|\sum_{k=1}^n \lambda_k \alpha_k\| \tag{3.23}$$

那么方程组(3.22)就有一个解. F.Riesz 提出了下述思路:
设由 $\{\alpha_k\}_{k\in\mathbb{N}}$ 张成的子空间为 A, 定义泛函:

$$f : A \to \mathbb{R}, \quad \alpha_k \mapsto f(\alpha_k) = (\alpha_k, x) = c_k$$

则条件(3.23)源于 f 是有界线性泛函, 这是因为 f 的线性性由构造即得, 因而 $f(\sum_{k=1}^n \lambda_k \alpha_k) = \sum_{k=1}^n \lambda_k f(\alpha_k) = \sum_{k=1}^n \lambda_k c_k$, 从而

$$|\sum_{k=1} \lambda_k c_k| \le M\|\sum_{k=1}^n \lambda_k \alpha_k\| \Leftrightarrow |f(\sum_{k=1}^n \lambda_k \alpha_k)| \le M\|\sum_{k=1}^n \lambda_k \alpha_k\|, \quad \forall n \in \mathbb{N}, \forall \{\lambda_k\}_{k\in\mathbb{N}} \subset \mathbb{R}$$

[9]方程的感觉很浓厚, 初步感觉还是得回到变分不等式和弱解那一章, 目前没有实力写这道.

注意此时 f 是定义在 A 上的, 而 A 最多也只能是 l^2 的有限维线性子空间. 如果能够说明给定一个 A 之后, f 可延拓到 l^2 上, 成为 l^2 上的有界线性泛函 (记为 F), 那么根据 Riesz 表示定理与 l^2 作为 Hilbert 空间的性质, 就存在 $x_0 \in l^2$ 使得 $\forall \alpha \in l^2(F(\alpha) = (\alpha, x_0))$. 此时

$$(\alpha_n, x_0) = F(\alpha_n) = f(\alpha_n) = c_n, \quad n = 1, 2, \cdots$$

从而方程组(3.22)就有解 x_0.

Riesz 本人也研究了上述问题的连续形式与其推广: 对于方程组

$$\int_I f(x) g_j(x) dx = c_j, \quad j = 1, 2, \cdots$$

其中 $\{g_j\}_{j\in\mathbb{N}} \subset L^q(I)$ 线性无关, $c_j(j = 1, 2, \cdots)$ 为常数, 问可否找到满足方程组的 $f \in L^p(I)$. 这个问题同样可以转化成上面介绍的泛函问题, 其关键步骤在于能不能把 $\mathrm{span}\{g_i\}_{i=1}^n$ 上的泛函 f 延拓到 L^p 上. Hahn-Banach 定理的一个动机就在于解决这个延拓问题.

3.5.1.1 线性泛函的延拓定理

回忆命题2.5.5, 复线性空间 $\mathscr{X}$ 上的均衡吸收凸集决定该空间上的一个半范数 $p(x)$, 现在研究从 $p(x)$ 中能否产生 $\mathscr{X}$ 上的一个连续线性泛函. 设 $x_0 \in \mathscr{X}$ 使得 $p(x_0) \neq 0$, 规定

$$\mathscr{X}_0 := \{\lambda x_0 : \lambda \in \mathbb{K}\}, f_0(\lambda x_0) := \lambda p(x_0), \quad \forall \lambda \in \mathbb{K}$$

则 f_0 就是 $\mathscr{X}_0$ 上的一个非零线性泛函, 满足有界性条件:

$$|f_0(\lambda x_0)| \leq |\lambda p(x_0)| = p(\lambda x_0), \quad \forall \lambda \in \mathbb{K}$$

这是因为若设 C 是 $\mathscr{X}$ 上的均衡吸收凸集, 知 $\theta \in C$, 因而根据命题2.5.5知 C 的 Minkowski 泛函决定了 $\mathscr{X}$ 上的一个半范数. 现在断言 $(\forall x \in \mathscr{X}(p(x) = 0)) \Leftrightarrow C = \mathscr{X}$.

当 $\forall x \in \mathscr{X}(p(x) = 0)$, 因为 $p(x) = 0 < 1$, 而根据 Minkowski 泛函本身的定义知若 $\lambda > P(x)$, 则 $\frac{x}{\lambda} \in C$, 故 $x \in C$, 这说明 $\mathscr{X} \subset C \subset \mathscr{X}$, 故 $C = \mathscr{X}$.

当 $C = \mathscr{X}$, 任取 $x \in \mathscr{X}, n \in \mathbb{N}$, 因为 $\mathscr{X}$ 是线性空间, 故 $nx \in \mathscr{X}$. 又因为 $C = \mathscr{X}$, 故 $nx \in C$, 故

$$0 \leq p(x) \leq \inf\{\frac{1}{n} : nx \in C, n \in \mathbb{N}\} = 0$$

从而 $p(x) = 0$, 上述断言因而成立. 特别若 $C \subsetneq \mathscr{X}$, 取 $x_0 \in \mathscr{X} \backslash C$, 则 $p(x_0) \geq 1$, 否则 $x_0 \in C$ 矛盾!

再来说明 f_0 是 $\mathscr{X}_0$ 上的非零线性泛函. 因为 $f_0(x_0) = p(x_0) \neq 0$, 故 $f_0 \not\equiv 0$. 又因为对任意 $x, y \in \mathscr{X}_0$, 出于 $x_0 \neq \theta$ 知存在唯一 $\lambda_x, \lambda_y \in \mathbb{K}$ 使得 $x = \lambda_x x_0, y = \lambda_y x_0$, 于是对任意 $\alpha, \beta \in \mathbb{K}$ 有:

$$\begin{aligned} f_0(\alpha x + \beta y) &= f_0(\alpha \lambda_x x_0 + \beta \lambda_y x_0) = f_0((\alpha \lambda_x + \beta \lambda_y) x_0) \\ &= (\alpha \lambda_x + \beta \lambda_y) f(x_0) = (\alpha \lambda_x + \beta \lambda_y) p(x_0) \\ &= \alpha \lambda_x p(x_0) + \beta \lambda_y p(x_0) = \alpha f_0(x) + \beta f_0(y) \end{aligned}$$

从而 f_0 是 $\mathscr{X}_0$ 上的线性泛函, 因而其为 $\mathscr{X}_0$ 上的非零线性泛函, 有界性条件易得.

现在希望把这个定义在 $\mathscr{X}_0$ 上的连续线性泛函延拓成整个空间 $\mathscr{X}$ 上的连续线性泛函.

定理 3.5.1 (实 Hahn-Banach 定理)

设 $\mathscr{X}$ 是实线性空间, p 是定义在 $\mathscr{X}$ 上的次线性泛函[a], $\mathscr{X}_0$ 是 $\mathscr{X}$ 的实线性子空间, f_0 是 $\mathscr{X}_0$ 上的实线性泛函并满足 $f_0(x) \leq p(x)(\forall x \in \mathscr{X}_0)$. 那么 $\mathscr{X}$ 上必有一个实线性泛函 f, 满足

1. $f(x) \leq p(x)(\forall x \in \mathscr{X})$(受 p 控制条件);
2. $f(x) = f_0(x)(\forall x \in \mathscr{X}_0)$(延拓条件),

[a]满足次可加性与正齐次性的实值函数

证明

任取 $y_0 \in \mathscr{X} \backslash \mathscr{X}_0$, 记 $\mathscr{X}_1 := \{x + \alpha y_0 : x \in \mathscr{X}_0, \alpha \in \mathbb{R}\}$, 现在考虑将 f_0 延拓到 $\mathscr{X}_1$. 设延拓后的线性泛函记为 f_1, 根据线性性的要求有

$$f_1(x + \alpha y_0) = f_0(x) + \alpha f_1(y_0), \quad \forall x \in \mathscr{X}_0, \forall \alpha \in \mathbb{R} \tag{3.24}$$

故只需确定 $f_1(y_0)$ 的值. 既然要求 f_1 满足受 p 控制条件, 有:

$$f_1(x + \alpha y_0) \le p(x + \alpha y_0), \quad \forall x \in \mathscr{X}_0, \forall \alpha \in \mathbb{R} \tag{3.25}$$

在(3.25)式两边同时除以 $|\alpha|$, 得到

$$\begin{cases} f_1(\frac{x}{\alpha} + y_0) \le p(\frac{x}{\alpha} + y_0), & \forall x \in \mathscr{X}_0, \forall \alpha > 0 \\ f_1(\frac{x}{\alpha} - y_0) \le p(\frac{x}{\alpha} - y_0), & \forall x \in \mathscr{X}_0, \forall \alpha < 0 \end{cases}$$

整理有

$$\begin{cases} f_1(y_0 - z) \le p(y_0 - z), & \forall z \in \mathscr{X}_0 \\ f_1(-y_0 + y) \le p(-y_0 + y), & \forall y \in \mathscr{X}_0 \end{cases}$$

再根据 f_1 的线性性得到

$$f_0(y) - p(-y_0 + y) \le f_1(y_0) \le f_0(z) + p(y_0 - z), \quad \forall y, z \in \mathscr{X}_0$$

进而要使得(3.25)式成立, 只需令上述不等式有解即可, 也即

$$\sup_{y \in \mathscr{X}_0} \{f_0(y) - p(-y_0 + y)\} \le \inf_{z \in \mathscr{X}_0} \{f_0(z) + p(y_0 - z)\} \tag{3.26}$$

注意到

$$f_0(y) - f_0(z) = f_0(y - z) \le p(y - z) \le p(y - y_0) + p(y_0 - z)$$

得到

$$f_0(y) - p(-y_0 + y) \le f_0(z) + p(y_0 - z), \quad \forall y, z \in \mathscr{X}_0$$

这说明(3.26)式恒成立. 现在任意取定 $f_1(y_0)$ 是满足(3.26)的某值, 即可根据(3.24)式得到 f_0 在 $\mathscr{X}_1$ 上的延拓 f_1. 因为(3.26)两端不一定相等, 故 $f_1(y_0)$ 的取法不唯一, 因而这个延拓不一定唯一.

现在考虑把 f_0 延拓到整个 $\mathscr{X}$ 上去, 下面借助 Zorn 引理(2.6.2). 令

$$\mathscr{F} := \{(\mathscr{X}_\Delta, f_\Delta) : \mathscr{X}_0 \subset \mathscr{X}_\Delta \subset \mathscr{X}; \forall x \in \mathscr{X}_0(f_\Delta(x) = f_0(x)); \forall x \in \mathscr{X}_\Delta(f_\Delta(x) \le p(x))\}$$

在 $\mathscr{F}$ 中引入序关系: $(\mathscr{X}_{\Delta_1}, f_{\Delta_1}) \prec (\mathscr{X}_{\Delta_2}, f_{\Delta_2})$ 是指

$$\mathscr{X}_{\Delta_1} \subset \mathscr{X}_{\Delta_2} \quad 且 \quad f_{\Delta_1}(x) = f_{\Delta_2}(x)(\forall x \in \mathscr{X}_{\Delta_1})$$

于是 $\mathscr{F}$ 是半序集, 又设 M 是 中的任意一个全序子集, 令

$$\mathscr{X}_M := \bigcup_{(\mathscr{X}_\Delta, f_\Delta) \in M\{\mathscr{X}_\Delta\}} \{\mathscr{X}_\Delta\}, \quad f_M(x) = f_\Delta(x)(\forall x \in \mathscr{X}_\Delta, (\mathscr{X}_\Delta, f_\Delta) \in M)$$

因为 M 是全序子集, 知 $\mathscr{X}_M$ 是 $\mathscr{X}$ 的包含 $\mathscr{X}_0$ 的子空间, 且 f_M 在 $\mathscr{X}_M$ 上唯一确定, 满足 $f_M(x) \le p(x)$. 故 $(\mathscr{X}_M, f_M) \in \mathscr{F}$ 且是 M 的一个上界. 根据 Zorn 引理(2.6.2), $\mathscr{F}$ 存在极大元, 记之为 $(\mathscr{X}_\Lambda, f_\Lambda)$.

最后证明 $\mathscr{X}_\Lambda = \mathscr{X}$, 用反证法. 如若不然, 根据先前对 $\mathscr{X}_1$ 的讨论, 类似可作 $(\widetilde{\mathscr{X}_\Lambda}, \widetilde{f_\Lambda}) \in \mathscr{F}$ 使得

$$\mathscr{X}_\Lambda \subset \widetilde{\mathscr{X}_\Lambda}, \quad \mathscr{X}_\Lambda \ne \widetilde{\mathscr{X}_\Lambda}$$

这说明 $(\widetilde{\mathscr{X}_\Lambda}, \widetilde{f_\Lambda}) \succ (\mathscr{X}_\Lambda, f_\Lambda)$, 但 $(\widetilde{\mathscr{X}_\Lambda}, \widetilde{f_\Lambda}) \ne (\mathscr{X}_\Lambda, f_\Lambda)$, 这与 $(\mathscr{X}_\Lambda, f_\Lambda)$ 的极大性矛盾! 故 $\mathscr{X}_\Lambda = \mathscr{X}$, 因而 f_Λ 就是欲求的延拓. □

注 **[ZL]** 上的叙述有些意识流与混乱, 下面介绍高志强老师上课采用的叙述再证一次.

不妨设 $\mathscr{X}_0 \subsetneq \mathscr{X}$, 否则定理已经证完了. 证明的第一部分在于说明可以做出一个" 较小的" 延拓. 因为 $\mathscr{X} \backslash \mathscr{X}_0$ 此时非空, 故可取 $x_0 \in \mathscr{X} \backslash \mathscr{X}_0$, 作 $\mathscr{X}_1 = \{x + \lambda x_0 : x \in \mathscr{X}_0, \lambda \in \mathbb{R}\}$, 则容易证明 $\mathscr{X}_1$ 为包含 $\mathscr{X}_0$ 与 x_0 的最小线性子

空间. 同时注意任取 $y \in \mathscr{X}_1$, 都存在唯一的 $x \in \mathscr{X}_0, \lambda \in \mathbb{R}$ 使得 $y = x + \lambda x_0$. 这是因为

$$x_1 + \lambda_1 x_0 = x_2 + \lambda_2 x_0 \Rightarrow x_1 - x_2 = (\lambda_2 - \lambda_1) x_0$$

但左式中 $x_1, x_2 \in \mathscr{X}_0$, 根据 $\mathscr{X}_0$ 是线性子空间知 $x_1 - x_2 \in \mathscr{X}_0$; 右式中 $x_0 \notin \mathscr{X}_0$, 因而 $\operatorname{span}\{x_0\} \cap \mathscr{X}_0 = \{\theta\}$. 于是左右两式要想相等, 只可能是它们均为 θ, 亦即 $x_1 = x_2, \lambda_1 = \lambda_2$, 唯一性得证. 下面证明在 $\mathscr{X}_1$ 上存在实线性泛函 f_1 满足

$$\begin{cases} f_1(x) = f(x), & \forall x \in \mathscr{X}_0 \\ f_1(x) \le p(x), & \forall x \in \mathscr{X}_1 \end{cases}$$

根据 $\mathscr{X}_1$ 的构造, 只需研究 $f_1(x + \lambda x_0) = f(x) + \lambda f_1(x)$ 即可, 其中 $x_0 \in \mathscr{X} \setminus \mathscr{X}_0, x \in \mathscr{X}_0$. 设 $f_1(x_0) = c$ 是一个常数, 则

$$f_1(x + \lambda x_0) = f_1(x) + \lambda f_1(x_0) = f(x) + \lambda c$$

而 $f(x)$ 是已知的, 所以 f_1 的不确定性全在于 c 的不确定性. 现在要想说明存在这样的 f_1 使得对任意的 $y = \lambda x_0 + x \in \mathscr{X}_1$ 都有 $f_1(y) \le p(y)$, 就是要说明存在常数 c 使得

$$\forall \lambda \in \mathbb{R}(f(x) + \lambda c \le p(x + \lambda x_0)) \tag{3.27}$$

现在开始讨论, 当 $\lambda = 0$ 时, c 可以任意取. 当 $\lambda > 0$ 时, 不等式(3.27)等价于

$$c \le \frac{1}{\lambda}(p(x + \lambda x_0) - f(x))$$

根据 f 的线性性与 p 的正齐次性 (即 $p(\lambda x) = \lambda p(x)(\lambda > 0)$) 有

$$c \le p(\frac{x}{\lambda} + x_0) - f(\frac{x}{\lambda})$$

记 $x' = \frac{x}{\lambda} \in \mathscr{X}_0$, 则上式即要求

$$c \le p(x' + x_0) - f(x'), \quad \forall x' \in \mathscr{X}_0$$

当 $\lambda < 0$ 时, 不等式(3.27)等价于

$$c \ge \frac{1}{\lambda}(p(x + \lambda x_0) - f(x)) = -p(-\frac{x}{\lambda} - x_0) + f(-\frac{x}{\lambda})$$

同样记 $x'' = -\frac{x}{\lambda} \in \mathscr{X}_0$, 上式即要求

$$c \ge -p(x'' - x_0) + f(x''), \quad \forall x'' \in \mathscr{X}_0$$

于是要说明存在常数 c 满足(3.27)式, 就是要说明存在常数 c 满足:

$$-p(x'' - x_0) + f(x'') \le c \le p(x' + x_0) - f(x'), \quad \forall x', x'' \in \mathscr{X}_0$$

进而只需说明上右式恒不大于上左式即可, 亦即要说明

$$-p(x'' - x_0) + f(x'') \le p(x' + x_0) - f(x'), \quad \forall x', x'' \in \mathscr{X}_0$$

这等价于说明

$$f(x'') + f(x') \le p(x' + x_0) + p(x'' - x_0), \quad \forall x', x'' \in \mathscr{X}_0 \tag{3.28}$$

已知

$$f(x'') + f(x') \overset{\text{(i)}}{=} f(x'' + x') \overset{\text{(ii)}}{\le} p(x'' + x') \overset{\text{(iii)}}{\le} p(x'' - x_0) + p(x' + x_0)$$

其中 (i) 是 f 的线性性, (ii) 是 f 被 p 控制的条件, (iii) 是 p 作为次线性泛函满足的次线性性. 于是(3.28)式成立, 因而 f_1 存在.

证明的第二部分在于说明从上面介绍的"较小的"延拓出发, 一步步延拓下去, 最终总可以得到在 $\mathscr{X}$ 上的延拓. 这个最终延拓的存在性就依赖于 Zorn 引理, 而 Zorn 引理要求的半序关系实际上也是很自然的.

在图3.3中, 把 $\{x\} = \{\theta\}$ 视作 $\mathscr{X}_0$, 则总可以在 $\mathscr{X}_0$ 外找到一点 x_1, 证明的第一部分相当于说明从 $\{x\}$ 延拓到 $\mathscr{X}_1 = \{y : y = \lambda x_1\}$(即橙色直线) 的这一步成立. 再把 $\mathscr{X}_1$ 进一步按证明的第一部分延拓成橙色平面 $\mathscr{X}_2$, 最后由 $\mathscr{X}_2$ 即可延拓成全空间 $\mathscr{X}$. 约定如果对满足 Hahn-Banach 定理3.5.1条件的实线性泛函 f 而言, 存在实线性泛函 F

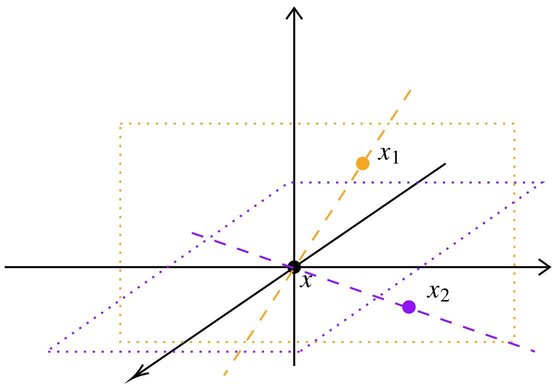


图 3.3: 半序关系的启发示意图

使得 $D(f) \subset D(F)$, 且 $\forall x \in D(f)(F(x) = f(x))$, 就称 F 是 f 的线性延拓. 如果记

$$\mathscr{F} := \{F : F\text{为}f\text{的线性延拓, 且}\forall x \in D(F)(F(x) \le p(x))\}$$

那么证明的第一部分实际上是在 $\mathscr{F}$ 上定义了一个半序 $\prec$: 称 $F_1 \prec F_2$, 是指

$$D(F_1) \subset D(F_2) \quad \text{且} \quad \forall x \in D(F_1)(F_2(x) = F_1(x))$$

注意 $\prec$ 之所以是半序, 是因为每一步延拓都不一定唯一. 比如在图3.3中另选 x_2, 同样可以有这么一条" 从 $\{x\}$ 到紫色直线, 再从紫色直线到紫色平面, 最后从紫色平面到全空间" 的延拓路线, 而这条路线和前面橙色的路线之间是没法比较的. 同时, $\mathscr{F}$ 中的全序子集其实就是上面介绍的一条条延拓路线. 现在要利用 Zorn 引理, 就是要说明 $\mathscr{F}$ 中的每个全序子集都有上界. 设 $M \subset \mathscr{F}$ 是一个全序子集, 则构造线性泛函 φ 满足:

$$D(\varphi) = \bigcup_{F \in M} D(F), \quad \forall F \in M \forall x \in D(F)(\varphi(x) = F(x))$$

注意 φ 的良定义性是由 M 全序推出来的. 现在显见 $\forall x \in D(\varphi)(\varphi(x) \le p(x))$, 于是 $\varphi \in \mathscr{F}$, 同时可以验证 $\forall F \in M(F \prec \varphi)$, 故 φ 正是 M 的上界. 从而根据 Zorn 引理, $\mathscr{F}$ 存在极大元 $\tilde{f}$.

最后, 证明的第三部分就是说明 $\tilde{f}$ 恰为欲求, 这需要说明 $D(\tilde{f}) = \mathscr{X}$ 的同时 $\forall x \in D(\tilde{f})(\tilde{f}(x) \le p(x))$. 根据 $\mathscr{F}$ 的构造, 第二条已经成立了, 所以只需说明第一条. 用反证法, 如果 $D(\tilde{f}) \subsetneq \mathscr{X}$, 就说明存在 $\tilde{x} \in \mathscr{X} \backslash D(\tilde{f})$, 于是根据证明的第一部分, $\tilde{f}$ 可以进一步延拓到 $\widetilde{\mathscr{X}} = \{x + \lambda\tilde{x} : x \in D(\tilde{f}), \lambda \in \mathbb{R}\}$ 上, 对应 $\tilde{f}$ 本身延拓为 $\hat{f}$. 显见 $D(\tilde{f}) \subsetneq \widetilde{\mathscr{X}} \subset \mathscr{X}$, 且 $\forall x \in D(\tilde{f})(\hat{f}(x) = \tilde{f}(x))$, 这说明 $\hat{f} \in \mathscr{F}$, $\tilde{f} \prec \hat{f}$, 且 $\tilde{f} \ne \hat{f}$, 这与 $\tilde{f}$ 的极大性矛盾! 故 $D(\tilde{f}) = \mathscr{X}$, 命题即证. □

在复线性空间的情形, 因为复数不能比较大小, 故定理需要做出修改.

定理 3.5.2 (复 Hahn-Banach 定理)

设 $\mathscr{X}$ 是复线性空间, p 是 $\mathscr{X}$ 上的半范数. $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间, f_0 是 $\mathscr{X}_0$ 上的线性泛函, 并满足 $|f_0(x)| \le p(x), \forall x \in \mathscr{X}_0$, 那么 $\mathscr{X}$ 上必有一个线性泛函 f 满足:

1. $|f(x)| \le p(x)(\forall x \in \mathscr{X})$;
2. $f(x) = f_0(x)(\forall x \in \mathscr{X}_0)$.

♡

证明

把 $\mathscr{X}$ 看成实线性空间, 相应地把 $\mathscr{X}_0$ 也看成实线性子空间, 令

$$g_0(x) := \operatorname{Re} f_0(x), \quad \forall x \in \mathscr{X}_0$$

则由 $|f(x)| \le p(x)$ 得到 $g_0(x) \le p(x)(\forall x \in \mathscr{X}_0)$, 进而根据实 Hahn-Banach 定理(3.5.1), 必有 $\mathscr{X}$ 上的实线性泛函 g, 使得

$$g(x) = g_0(x) \quad (\forall x \in \mathscr{X}_0) \tag{3.29}$$

且

$$g(x) \le p(x) \quad (\forall x \in \mathscr{X}) \tag{3.30}$$

现令

$$f(x) := g(x) - ig(ix) \quad (\forall x \in \mathscr{X}) \tag{3.31}$$

则由(3.29)知

$$f(x) = g_0(x) - ig_0(ix) = \operatorname{Re} f_0(x) + i \operatorname{Im} f_0(x) = f_0(x) \quad (\forall x \in \mathscr{X}_0)$$

又因为

$$f(ix) = g(ix) - ig(-x) = i(g(x) - ig(ix)) = if(x) \quad (\forall x \in \mathscr{X})$$

故 f 是复齐性的. 现在说明在 $\mathscr{X}$ 上 $|f(x)|$ 受 $p(x)$ 控制. 当 $f(x) = 0$, 该命题显然. 当 $f(x) \neq 0$, 令

$$\theta := \arg f(x)$$

由(3.30)式知

$$|f(x)| = e^{i\theta} f(x) = f(e^{i\theta}x) = g(e^{i\theta}x) \le p(e^{i\theta}x) = p(x) \quad (\forall x \in \mathscr{X})$$

其中 $f(e^{i\theta}x) = g(e^{i\theta}x)$ 是因为正数 $f(e^{i\theta}x) = |f(x)|$ 的虚部为 0. □

注 与前注同样的理由, 下面按照课堂思路再整理一遍复 Hahn-Banach 定理的证明. 首先回忆半范数的定义:

定义 3.5.1 (半范数)

若 $\mathscr{X}$ 是 $\mathbb{K}$ 上的线性空间, $p : \mathscr{X} \to \mathbb{R}$ 满足条件:

(i) $\forall x, y \in \mathscr{X}(p(x+y) \le p(x) + p(y))$;

(ii) $\forall x \in \mathscr{X} \forall \alpha \in \mathbb{K}(p(\alpha x) = |\alpha| p(x))$,

则称 p 是 $\mathscr{X}$ 上的半范数. ♣

在复 Hahn-Banach 定理中, 因为 f 现在成为复线性泛函, 故不能直接写 $f(x) \le p(x)$, 而只能是取模后写 $|f(x)| \le p(x)$. 现在开始证明.

证明

回忆在实 Hahn-Banach 定理中, 只要 $\mathscr{X}$ 是实线性空间, $\forall x \in \mathscr{X}_0(|f(x)| \le p(x)) \Rightarrow \forall x \in \mathscr{X}_0(f(x) \le p(x))$, 从而存在 $\tilde{f}$ 使得

(a) $\forall x \in \mathscr{X}_0(\tilde{f}(x) = f(x))$;

(b) $\forall x \in \mathscr{X}(\tilde{f}(x) \le p(x))$.

而 $-\tilde{f}(x) = \tilde{f}(-x) \le p(-x) = p(x)$, 于是 $\forall x \in \mathscr{X}(|\tilde{f}(x)| \le p(x))$. 现在当 $\mathscr{X}$ 是复线性空间, 保持 $\mathscr{X}$ 中的元素和计算不变, 而仅把数乘限制在实数范围内, 得到实线性空间 $\mathscr{X}_r$, 则在 $\mathscr{X}_r$ 上即可应用实 Hahn-Banach 定理.

现在说明复线性空间 $\mathscr{X}$ 上的复线性泛函 f 由其实部唯一决定. 考察 $f(ix)$, 一方面根据实部和虚部的定义有

$$f(ix) = \operatorname{Re} f(ix) + i \operatorname{Im} f(ix)$$

另一方面根据 f 的线性性有

$$f(ix) = if(x) = i \operatorname{Re} f(x) + i^2 \operatorname{Im} f(x) = i \operatorname{Re} f(x) - \operatorname{Im} f(x)$$

比对两式实部即得 $\operatorname{Im} f(x) = -\operatorname{Re} f(ix)$, 从而

$$f(x) = \operatorname{Re} f(x) - i \operatorname{Re} f(ix)$$

现在便只需讨论 $\operatorname{Re} f$, 设 $\mathscr{X}_0 \subset \mathscr{X}$ 是线性子空间, f 是 $\mathscr{X}_0$ 上的线性泛函, 且满足

$$|f(x)| \le p(x), \quad \forall x \in \mathscr{X}_0$$

记 $f_1(x) = \operatorname{Re} f(x)$, 并记 $\mathscr{X}_{0,r}$ 是 $\mathscr{X}_0$ 对应的实线性空间, 则 f_1 是 $\mathscr{X}_{0,r}$ 上的实线性泛函, 其满足条件

$$|f_1(x)| \le p(x), \quad \forall x \in \mathscr{X}_{0,r}$$

于是由实 Hahn-Banach 定理知存在 f_1 由 $\mathscr{X}_{0,r}$ 到 $\mathscr{X}_r$ 的实线性延拓 $\widetilde{f}_1$, 其满足

(a) $\forall x \in \mathscr{X}_{0,r}(\widetilde{f}_1(x) = f(x))$;

(b) $\forall x \in \mathscr{X}_r(|\widetilde{f}_1(x)| \le p(x))$

从而令 $\widetilde{f}(x) = \widetilde{f}_1(x) - i\widetilde{f}_1(ix)$, 则知 $\widetilde{f}$ 是 f 由 $\mathscr{X}_0$ 到 $\mathscr{X}$ 的线性延拓.

最后只需证明 $\forall x \in \mathscr{X}(|\widetilde{f}(x)| \le p(x))$, 在 $\widetilde{f}(x) \ne 0$ 时, 根据复数的指数记法知可设 $\widetilde{f}(x) = |\widetilde{f}(x)|e^{i\theta}$, 从而

$$|\widetilde{f}(x)| = \widetilde{f}(x)e^{-i\theta} = \widetilde{f}(-e^{i\theta}x)$$

注意 $\widetilde{f}(-e^{i\theta}x) \in \mathbb{R}$, 故

$$\widetilde{f}(-e^{i\theta}x) = \widetilde{f}_1(-e^{i\theta}x) \le p(-e^{i\theta}x) = p(x)$$

故 $\forall x \in \mathscr{X}(|\widetilde{f}(x)| \le p(x))$, 命题得证. □

结合命题2.5.5可得下述定理.

定理 3.5.3

为了复线性空间 $\mathscr{X}$ 上至少有一个非零线性泛函, 只要 $\mathscr{X}$ 中含有某一个均衡吸收真凸子集. ♡

证明 ★

若 $\mathscr{X}$ 含有一个均衡吸收真凸子集, 则根据命题2.5.5知存在 $\mathscr{X}$ 上的一个半范数 $p(x)$, 且 p 必不恒等于零. 现任取 $x_0 \in \mathscr{X}$ 满足 $x_0 \ne \theta, p(x_0) \ne 0$, 并取 $f_0(x_0) \in \mathbb{R}, f_0(x_0) \ne 0, |f_0(x_0)| \le p(x_0)$, 知对任意的 $\alpha \in \mathbb{K}$, 设 $f_0(\alpha x_0) := \alpha f_0(x_0)$, 并取 $\mathscr{X}_0 = \operatorname{span}\{x_0\}$, 则根据本节最初的讨论知 $\forall y \in \mathscr{X}_0(|f_0(y)| \le p(y))$, 因而由复 Hahn-Banach 定理3.5.2知存在 $\mathscr{X}$ 上的线性泛函 f, 它是 f_0 在 $\mathscr{X}$ 上的延拓, 取用 $f \not\equiv 0$ 即为欲求. □

在 B^* 空间上, 因为不再需要讨论半范数或次线性泛函 $p(x)$ 的存在性, 而是有现成的泛函 $\|\cdot\|$, 故此时 Hahn-Banach 延拓定理有更特殊的形式.

定理 3.5.4 (Hahn-Banach)

设 $\mathscr{X}$ 是 B^* 空间, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间, f_0 是定义在 $\mathscr{X}_0$ 上的有界线性泛函, 则在 $\mathscr{X}$ 上必有有界线性泛函 f 满足:

(i) $f(x) = f_0(x)(\forall x \in \mathscr{X}_0)$(延拓条件),

(ii) $\|f\| = \|f_0\|_0$(保范条件),

其中 $\|f_0\|_0$ 表示 f_0 在 $\mathscr{X}_0$ 上的范数. ♡

注 因为 f 满足 (i),(ii) 两个条件, 一般称 f 为 f_0 的保范延拓.

证明

在 $\mathscr{X}$ 上定义 $p(x) := \|f_0\|_0 \cdot \|x\|$, 则 $p(x)$ 是 $\mathscr{X}$ 上的半范数, 且在 $\mathscr{X}_0$ 上有

$$|f_0(x)| \le \|f_0\|_0 \cdot \|x\| = p(x), \quad \forall x \in \mathscr{X}_0$$

因而根据复 Hahn-Banach 定理(3.5.2), 必存在 $\mathscr{X}$ 上的线性泛函 $f(x)$ 满足

$$f(x) = f_0(x), \quad \forall x \in \mathscr{X}_0 \tag{3.32}$$

及

$$|f(x)| \le p(x) = \|f_0\|_0 \cdot \|x\| \Rightarrow \|f\| \le \|f_0\|_0, \quad \forall x \in \mathscr{X} \tag{3.33}$$

又因为(3.32)式表明

$$|f(x)| = |f_0(x)| \le \|f\| \cdot \|x\| \Rightarrow \|f_0\|_0 \le \|f\|, \quad \forall x \in \mathscr{X}_0 \tag{3.34}$$

故由(3.33),(3.34)式知 $\|f\| = \|f_0\|_0$. □

推论 3.5.1 (连续线性泛函足以区分元素)

每个 B^* 空间中必有足够多的连续线性泛函. 也即当 $x_1 \neq x_2(x_1, x_2 \in \mathscr{X})$ 时, 必存在该空间中的一个连续线性泛函 $f(\cdot)$, 使得 $f(x_1) \neq f(x_2)$.

证明

任取 $x_1, x_2 \in \mathscr{X}$, 若 $x_1 \neq x_2$, 则定义 $x_0 := x_1 - x_2 \neq \theta$. 令 $\mathscr{X}_0 := \{\lambda x_0 : \lambda \in \mathbb{C}\}$, 并在 $\mathscr{X}_0$ 上定义

$$f_0(\lambda x_0) = \lambda \|x_0\| \quad (\forall \lambda \in \mathbb{C})$$

可知 $f_0(x_0) = \|x_0\|, \|f_0\|_0 = 1$. 根据 B^* 空间上的 Hahn-Banach 定理(3.5.4), 存在 $\mathscr{X}$ 上的连续线性泛函 f, 使得

$$f(x_0) = f_0(x_0) = \|x_0\|, \quad \|f\| = \|f_0\|_0 = 1$$

此时

$$f(x_1) - f(x_2) - f(x_1 - x_2) = f(x_0) \neq 0$$

□

推论 3.5.2 (总存在保单点范数且范数为 1 的连续线性泛函)

设 $\mathscr{X}$ 是 B^* 空间, $\forall x_0 \in \mathscr{X} \backslash \{\theta\}$, 必 $\exists f \in \mathscr{X}^*$, 使得

$$f(x_0) = \|x_0\| \quad 且 \quad \|f\| = 1.$$

注 该推论给出了一个判别 B^* 空间零元的一种方法: 要使得 $x_0 = \theta$, 必须且仅须对任意的 $f \in \mathscr{X}^*$ 有 $f(x_0) = 0$.

在 Hilbert 空间中, 根据 Riesz 表示定理, 对任意的连续线性泛函 f, 都存在 $y \in H$ 使得

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

若记 $M := \{x : f(x) = 0\}$, 则对任意的 $x_0 \in H$ 有

$$f(x_0) = (x_0, y) = (x_0 - P_M x_0, y)$$

其中 $P_M x_0$ 是 x_0 在 M 上的投影, 故

$$|f(x_0)| \leq \|x_0 - P_M x_0\| \cdot \|y\| = \rho(x_0, M) \cdot \|f\| \tag{3.35}$$

在一般的 B^* 空间 $\mathscr{X}$ 中, $\rho(x_0, M) := \inf\limits_{y \in M} \|x_0 - y\|$, 下面说明(3.35)式此时仍然成立. 这是因为对任意的 $n \in \mathbb{N}, x_0 \in \mathscr{X}$, 由下确界定义知存在 $x_n \in M$ 使得

$$\rho(x_0, M) \leq \rho(x_0, x_n) < \rho(x_0, M) + \frac{1}{n}$$

故

$$|f(x_0)| = |f(x_n - x_0)| \leq \|f\| \cdot \|x_n - x_0\| \leq \|f\| \cdot (\rho(x_0, M) + \frac{1}{n})$$

令 $n \to \infty$ 即得(3.35)式.

现在希望了解在 B^* 空间 $\mathscr{X}$ 上, 给定子空间 M 与 $x_0 \in \mathscr{X} \backslash M$, 是否存在 $f \in \mathscr{X}^*$, 使得 f 在 M 上为 0, 并且(3.35)式中的等号成立? 有下述结论成立.

定理 3.5.5

设 $\mathscr{X}$ 是 B^* 空间, M 是 $\mathscr{X}$ 的线性子空间. 若 $x_0 \in \mathscr{X}$, 且

$$d := \rho(x_0, M) > 0$$

则必 $\exists f \in \mathscr{X}^*$ 满足条件:

(i) $f(x) = 0 (\forall x \in M)$;

(ii) $f(x_0) = d$;

(iii) $\|f\| = 1$.

证明

考虑 $\mathscr{X}_0 := \{x = x' + \alpha x_0 : x' \in M, \alpha \in \mathbb{K}\}$. 对任意的 $x \in \mathscr{X}_0$, 定义

$$f_0(x) = \alpha d$$

对条件 (i), 知此时 $\alpha = 0$, 因而 $f(x) = 0$.

对条件 (ii), 知此时 $x' = \theta, \alpha = 1$, 得到 $f(x_0) = d$.

现在验证条件 (iii). 当 $x = x' + \alpha x_0 (x' \in M, \alpha \neq 0)$, 有

$$|f_0(x)| = |\alpha| d = |\alpha| \rho(x_0, M) = |\alpha| \cdot \inf_{z \in M} \|x_0 - z\| \leq |\alpha| \|x_0 + \frac{x'}{\alpha}\| = \|x' + \alpha x_0\| = \|x\|$$

这说明 $\|f_0\| \leq 1$, 进而 f_0 是 $\mathscr{X}_0$ 上的一个有界线性泛函. 由 B^* 空间上的 Hahn-Banach 定理(3.5.4), 将 f_0 保范延拓成 $f \in \mathscr{X}$, 知 f 满足条件 (i) 和 (ii), 且 $\|f\| = \|f_0\| \leq 1$. 又因为 $f \in \mathscr{X}^*$ 且满足条件 (ii), 由(3.35)式知

$$|f(x_0)| = d \leq \|f\| \cdot \rho(x_0, M) = \|f\| \cdot d \Rightarrow \|f\| \geq 1$$

故只能有 $\|f\| = 1$, 条件 (iii) 得证. □

推论 3.5.3

设 M 是 B^* 空间 $\mathscr{X}$ 的一个子集, 又设 x_0 是 $\mathscr{X}$ 中的任一个非零元素, 那么

$$x_0 \in \overline{\operatorname{span} M}$$

的充要条件是: 对 $\forall f \in \mathscr{X}^*$ 有

$$\forall x \in M (f(x) = 0) \Rightarrow f(x_0) = 0.$$

证明

当 $x_0 \in \overline{\operatorname{span} M}$, 命题无所可证.

当 $\forall f \in \mathscr{X}^* (\forall x \in M (f(x) = 0) \Rightarrow f(x_0) = 0)$, 考虑反证法. 设 $x_0 \notin \overline{\operatorname{span} M}$, 则

$$d := \rho(x_0, \overline{\operatorname{span} M}) > 0$$

故由定理(3.5.5)知, 存在 $f \in \mathscr{X}^*$ 使得 $f(x) = 0 (\forall x \in M)$, 且 $f(x_0) = d > 0$. 但根据前述证明此时应有 $f(x_0) = 0$, 矛盾! □

例 3.24 若 $M = \{x_1, x_2, \cdots, x_n, \cdots\}$, 问可否用形如 $\sum_{i=1}^{n} c_i x_i$ 的线性组合的序列极限去逼近给定的元素 x_0? 现在根据推论(3.5.3)知这个逼近存在当且仅当对所有的在 $x_1, x_2, \cdots, x_n, \cdots$ 上为 0 的连续线性泛函 f, 都有 $f(x_0) = 0$.

3.5.1.2 Hahn-Banach 定理的几何形式: 凸集分离定理

平面上两个互不相交的凸集 A 与 B, 若 $A \cap B = \emptyset$, 则必存在一条直线 l 分离 A 与 B, 即存在直线 l 使得 A 与 B 各在 l 的一侧.

现在希望在一般的线性空间 $\mathscr{X}$ 中推广这条性质. 为简单起见, 下面总假定 $\mathscr{X}$ 是实的, 其上的线性泛函取实值. $\mathscr{X}$ 上对应平面情况下过原点直线的概念是极大线性子空间.

定义 3.5.2 (极大)

在线性空间 $\mathscr{X}$ 中, $\mathscr{X}$ 的线性子空间 M 称为是极大的, 如果对于任何一个以 M 为真子集的线性子空间 M_1 必有 $M_1 = \mathscr{X}$.

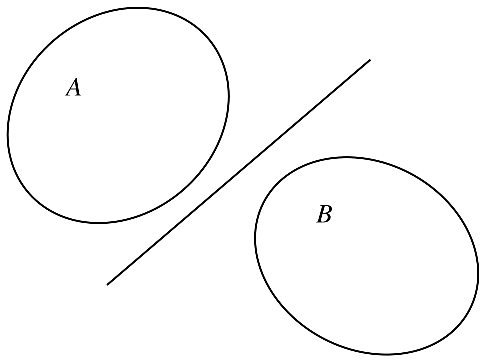


图 3.4: 互不相交的凸集总能被直线分离.

命题 3.5.1

M 是极大线性子空间的充要条件是, $M \subsetneq \mathscr{X}$ 是线性子空间, 并且 $\exists x_0 \in \mathscr{X} \backslash M$ 使得

$$\mathscr{X} = \{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M.$$

证明

当 M 是极大线性子空间, 结论显然成立.

当 $M \subsetneq \mathscr{X}$ 是线性子空间, 且 $\exists x_0 \in \mathscr{X} \backslash M (\mathscr{X} = \{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M)$, 任取 $\widetilde{x}_0 \in \mathscr{X} \backslash M$, 因为

$$\widetilde{x}_0 \in \mathscr{X} = \{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M$$

故存在 $\lambda_M \in \mathbb{R}$ 与 $x_M \in M$ 满足

$$\widetilde{x}_0 = \lambda_M x_0 + x_M$$

显见 $\lambda_M \neq 0$, 否则 $\widetilde{x}_0 = x_M \in M$ 矛盾! 故

$$x_0 = \frac{\widetilde{x}_0}{\lambda_M} - \frac{x_M}{\lambda_M}$$

从而

$$\mathscr{X} = (\{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M) \subset (\{\mu \widetilde{x}_0 : \mu \in \mathbb{R}\} \oplus M) \subset \mathscr{X}$$

于是

$$\mathscr{X} = \{\mu \widetilde{x}_0 : \mu \in \mathbb{R}\} \oplus M$$

根据定义即得 M 是 $\mathscr{X}$ 的极大线性子空间. □

定义 3.5.3 (超平面)

$\mathscr{X}$ 的极大线性子空间 M 对向量 $x_0 \in \mathscr{X}$ 的平移

$$L := x_0 + M$$

称为极大线性流形, 或简称超平面.

注 超平面是平面情况下一般直线的推广. 平面上的直线 l 可以通过线性函数表示:

$$l = \{x = (\xi, \eta) : a\xi + b\eta = c\}$$

同样, 超平面也可以通过线性泛函来刻画.

定理 3.5.6 (超平面的刻画)

要使得 L 是线性 (B^*) 空间 $\mathscr{X}$ 上的一个 (闭) 超平面, 必须且仅须存在非零 (连续) 线性泛函 f 及 $r \in \mathbb{R}$, 使

得

$$L=H_f^r:=\{x\in\mathscr{X}:f(x)=r\}.$$

♡

证明

当存在非零线性泛函 f 满足条件, 知此时 H_f^0 必是线性子空间, 现在说明 H_f^0 是极大的. 这是因为若取 $x_1\in\mathscr{X}\backslash H_f^0$, 对任意的 $x\in\mathscr{X}$, 必定有分解

$$x=a+b,\quad a\in\mathscr{X}\backslash H_f^0, b\in H_f^0$$

现在要求 $f(x)=f(a+b)$, 根据 f 的线性性知 $f(a+b)=f(a)+f(b)=f(a)$, 故可取 $a=\frac{f(x)}{f(x_1)}x_1$. 这说明 $\mathscr{X}$ 中的任意元素总能表成 $\lambda x_1+b, x_1\in\mathscr{X}\backslash H_f^0$ 的形式, 因而 $\mathscr{X}=\{\lambda x_1:\lambda\in\mathbb{R}\}\oplus H_f^0$. 又因为 x_1 是任取的, 故 H_f^0 是极大的.

再说明 H_f^0 是超平面. 因为 f 非零, 知存在 $x_0\in\mathscr{X}$ 使得 $f(x_0)\neq 0$. 不妨设 $f(x_0)=r$, 则根据 f 的线性性, 对任意的 $x\in H_f^r$ 有

$$f(x-x_0)=f(x)-f(x_0)=0$$

故 $x-x_0\in H_f^0$, 这说明 $H_f^r=x_0+H_f^0$ 是超平面. 当 f 连续, 显见 H_f^r 是闭的.

当 L 是 (闭) 超平面, 根据定义设 $L=x_0+M$, 其中 M 是 (闭) 极大线性子空间, $x_0\in\mathscr{X}\backslash M$. 根据 M 极大的定义知 $\mathscr{X}$ 中的任意元素 x 总能表成

$$x=\lambda x_0+y\quad(\lambda\in\mathbb{R}, y\in M)$$

再定义线性泛函 $f:\mathscr{X}\to\mathbb{R}$ 为

$$f(x)=f(\lambda x_0+y)=\lambda\quad(\lambda\in\mathbb{R}, y\in M)$$

显见 f 为 $\mathscr{X}$ 上的线性泛函, 满足 $M=H_f^0, f(x_0)=1$. 故 $L=H_f^1$. 可以证明当 L 是闭的, 则 H_f^0 是闭的, 得到 f 连续. □

现在要说明超平面 $L=H_f^r$ 使一个集合 E 在它的一侧, 用线性泛函描述就是

$$\forall x\in E\Rightarrow f(x)\le r(\text{或}f(x)\ge r)$$

定义 3.5.4 (分离, 严格分离)

所谓超平面 $L=H_f^r$ 分离集合 E 和 F 是指:

$$\forall x\in E\Rightarrow f(x)\le r(\text{或}f(x)\ge r)$$

$$\forall x\in F\Rightarrow f(x)\ge r(\text{或}f(x)\le r)$$

如果在上述两个式子中用严格不等号代替不等号, 就说 H_f^r 严格分离 E 与 F.

♣

Hahn-Banach 定理可以用来讨论如何用超平面分离两个互不相交的凸集, 这呈现于下述定理.

定理 3.5.7 (Hahn-Banach 定理的几何形式)

设 E 是实 B^* 空间 $\mathscr{X}$ 上以 θ 为内点的真凸子集, 又设 $x_0\notin E$, 则必存在一个超平面 H_f^r 分离 x_0 与 E.

♡

证明

根据命题2.5.6, 由 E 作为真凸子集以 θ 为内点知, E 的 Minkowski 泛函 $p(x)$ 是一个非零的连续次线性泛函, 满足

$$\forall x\in E\Rightarrow p(x)\le 1$$

若还存在一点 $x_0\in\mathscr{X}\backslash E$, 则由 $p(x)$ 的定义与 E 是以 θ 为内点的凸集知: $p(x_0)\ge 1$. 下面证明存在超平面 H_f^r 分离

E 与 x_0, 现在根据 Minkowski 泛函寻找欲求的线性泛函. 先在一维线性空间

$$\mathscr{X}_0 := \{\lambda x_0 : \lambda \in \mathbb{R}\}$$

上定义

$$f_0(\lambda x_0) := \lambda p(x_0) \quad (\forall \lambda \in \mathbb{R})$$

显见 f_0 是 $\mathscr{X}_0$ 上的线性泛函. 注意 $p(x)$ 本身是次线性泛函, 且有

$$f_0(x) = f_0(\lambda x_0) = \lambda p(x_0) \le p(\lambda x_0) = p(x) \quad (\forall x \in \mathscr{X}_0)$$

根据实 Hahn-Banach 定理(3.5.1)知, 必存在 $\mathscr{X}$ 上的线性泛函 $f(x)$ 满足

$$f(x_0) = f_0(x_0) = p(x_0) \ge 1$$

$$f(x) \le p(x) \quad (\forall x \in \mathscr{X})$$

故 $f(x) \le p(x) \le 1 (\forall x \in E)$, 因而 H_f^1 就是分离 E 与 x_0 的超平面. □

注

1. 因为只要通过平移即可把任意一点变为 θ 点, 故定理3.5.7对于含有任意内点的真凸子集都成立. 但对于无穷维空间 $\mathscr{X}$, 不能忽略 E 有内点. 譬如对 l^2 中的真凸子集 E, 若 E 无内点, 则存在 $x_0 \notin E$, 使得 x_0 与 E 不可通过闭超平面分离. 构造
$$E = \{x = (x_1, \cdots, x_n, \cdots) \in l^2 : \exists i_0 \in \mathbb{N} \forall i \ge i_0 (x_i = \frac{1}{i})\}$$
显见 $\theta = (0, \cdots, 0, \cdots) \notin E$, 且 E 是凸集, 现在说明 E 无内点. 任取 $x_0 \in E$, 不妨设存在 $n_0 \in \mathbb{N}$ 使得
$$x_0 = (x_1, \cdots, x_{n_0}, \frac{1}{n_0+1}, \frac{1}{n_0+2}, \cdots)$$
对任意 $\varepsilon > 0$, 取 $n_1 \in [n_0, \infty) \cap \mathbb{N}$ 充分大, 使得
$$(\sum_{i=n_1+1}^{\infty} \frac{1}{i^2})^{\frac{1}{2}} < \varepsilon$$
现在令
$$y := (x_1. \cdots, x_{n_0}, \frac{1}{n_0+1}, \cdots, \frac{1}{n_1}, 0, \cdots)$$
则 $y \notin E$, 且 $\|y - x_0\|_{l^2} = (\sum\limits_{i=n_1+1}^{\infty} \frac{1}{i^2})^{\frac{1}{2}} < \varepsilon$, 故 $B_{l^2}(x_0, \varepsilon) \not\subset$, 从而 x_0 不是 E 的内点. 由 x_0 的任意性即知 E 无内点.
因为 l^2 本身是 Hilbert 空间, 故由 Hilbert 空间上的 Riesz 表示定理3.3.1知对任意 $y = (y_1, \cdots, y_n, \cdots) \in l^2$ 总存在 $f_y \in (l^2)^*$ 使得 $\forall x \in l^2 (\langle f, x\rangle = \langle x, y_f\rangle)$.
现若 $f_y \neq 0$ 分离 θ 与 E, 则 $f_y(\theta) = \sum\limits_{i=1}^{\infty} 0 y_i = 0$, 又因为 $f_y \neq 0$, 故存在 $i_0 \in \mathbb{N}$ 使得 $y_{i_0} \neq 0$. 考虑
$$x = (0, \cdots, 0, x_{i_0}, 0, \cdots, 0, \frac{1}{j_0}, \frac{1}{j_0+1}, \cdots) \in E \tag{3.36}$$
则
$$f_y(x) = x_{i_0} y_{i_0} + \sum_{k=j_0}^{\infty} \frac{y_k}{k}$$
取 $j_0 > i_0$ 充分大, 使得
$$|\sum_{k=j_0}^{\infty} \frac{y_k}{k}| \le (\sum_{k=j_0}^{\infty} \frac{1}{k^2})^{\frac{1}{2}} (\sum_{k=j_0}^{\infty} y_k^2)^{\frac{1}{2}} < \frac{1}{2}|y_{i_0}|$$
现在对这样的 j_0, 总可以找到 E 中形如(3.36)的点 $x^1 = x, x^2 = x - 2x_{i_0} e_{i_0} (x_{i_0} > 0)$ 使得 $f_y(x^1) > 0, f_y(x^2) < 0$, 故 f_y 不能分离 θ 与 E.
2. 可以证明定理3.5.7中存在的超平面 $L := H_f^r$ 是闭的, 这只需证明 f 是连续的. 事实上, 由 $f(x) \le p(x) (\forall x \in$

$\mathscr{X}$) 知

$$|f(x)| \le \max(p(x), p(-x)) \quad (\forall x \in \mathscr{X})$$

故由 $p(x)$ 的连续性可得 f 在 θ 点连续. 又因为 f 是线性的, 故 f 在 $\mathscr{X}$ 上连续.

前面考虑的是点和凸集的分离问题, 现在考虑两个凸集之间的分离问题.

定理 3.5.8 (凸集分离定理)

设 E_1 和 E_2 是 B^* 空间中两个互不相交的非空凸集, E_1 有内点, 那么 $\exists s \in \mathbb{R}$ 及非零连续线性泛函 f, 使得超平面 H_f^s 分离 E_1 和 E_2. 也就是说, 存在一个非零线性泛函 f, 使得

$$f(x) \le s(\forall x \in E_1), \quad f(x) \ge s(\forall x \in E_2).$$

证明

在 B^* 空间 $\mathscr{X}$ 中, 若 E_1, E_2 是两个互不相交的凸集, 且 E_1 有内点, 则知集合

$$E = E_1 - E_2 := \{x - y : x \in E_1, y \in E_2\}$$

是一个有内点的非空凸集. 同时 $\theta \notin E$, 这是因为如若 $\theta \in E$, 则存在 $x \in E_1, y \in E_2$ 使得 $x - y = \theta$, 也即 $x = y$, 这与 $E_1 \cap E_2 = \varnothing$ 矛盾.

现在根据 Hahn-Banach 定理的几何形式(3.5.7), 存在非零连续泛函 f 与闭超平面 H_f^r 分离 E 和 θ, 不妨设

$$f(x) \le r(\forall x \in E), \quad f(\theta) \ge r$$

因为 $f(\theta) = 0$, 故 $f(x) \le r \le 0(\forall x \in E)$, 也即 $\forall x \in E_1 \forall y \in E_2(f(x - y) \le 0)$, 根据 f 的线性性知

$$\forall x \in E_1 \forall y \in E_2(f(x) \le f(y))$$

从不等式有解的角度看, 这说明存在 $s \in \mathbb{R}$ 使得

$$\sup_{x \in E_1} f(x) \le s \le \inf_{y \in E_2} f(y)$$

于是 H_f^s 分离 E_1 和 E_2. 根据 H_f^r 的闭性知 H_f^s 也是闭的. □

注 条件 $E_1 \cap E_2 = \varnothing$ 可以减弱成 $E_1^\circ \cap E_2 = \varnothing$. 这是因为从 E_1 有内点知 E_1° 有内点, 故 E_1° 是有内点的凸集. 对 E_1° 和 E_2 应用凸集分离定理(3.5.8)得到超平面 H_f^s 使得

$$f(x) \le s(\forall x \in E_1^\circ), \quad f(x) \ge s(\forall x \in E_2)$$

由 f 的连续性知

$$\forall x \in E_1^\circ(f(x) \le s) \Rightarrow \forall x \in \overline{E_1^\circ}(f(x) \le s)$$

根据练习(2.39)知 $\overline{E_1^\circ} = \overline{E_1}$, 得到

$$f(x) \le s \quad (\forall x \in E_1)$$

这说明 H_f^s 分离 E_1, E_2.

注 [**HB**] 中给出了 Hahn-Banach 几何形式的另外一种表述: 第一几何形式与第二几何形式.

补充定理 3.5.1 (Hahn-Banach, 第一几何形式[HB])

设 E 是 B^* 空间, $A \subset E, B \subset E$ 是两个不相交的非空凸集. 设 A 是开集, 则存在闭的超平面分离 A 和 B.

这其实就是 [**ZL**] 中讲的凸集分离定理, 只是上述定理的条件还要强一些: [**ZL**] 中只要求 $A^\circ \neq \varnothing$ 即可. 不过它们的证明思路都是考虑集合的差集.

补充定理 3.5.2 (Hahn-Banach, 第二几何形式[HB])

设 E 是 B^* 空间, $A \subset E, B \subset E$ 是两个不相交的非空凸集. 设 A 是闭集, B 是紧集, 则存在闭超平面严格分离 A 与 B.

证明

任取 $\varepsilon > 0$, 设 $A_\varepsilon := A + B(0,\varepsilon) = \{x+y : x \in A, y \in B(0,\varepsilon)\}, B_\varepsilon := B + B(0,\varepsilon) = \{x+y : x \in B, y \in B(0,\varepsilon)\}$, 则显见 $A_\varepsilon \neq \varnothing, B_\varepsilon \neq \varnothing$, 且由 $A, B, B(0,\varepsilon)$ 的凸性知 $A_\varepsilon, B_\varepsilon$ 都是凸集. 另外, 任取 $x+y \in A_\varepsilon (x \in A, y \in B(0,\varepsilon))$, 因为 $B(0,\varepsilon)$ 是开集, 故总存在 $\delta > 0$ 使得 $B(y,\delta) \subset B(0,\varepsilon)$, 因而 $x + B(y,\delta) = B(x,\delta) \subset A_\varepsilon$, 故 A_ε 是开集, 类似可证 B_ε 是开集, 于是 $A_\varepsilon, B_\varepsilon$ 均为非空凸开集.

再证明 $\varepsilon > 0$ 足够小时, $A_\varepsilon, B_\varepsilon$ 依旧不相交. 用反证法, 如果 $\forall \varepsilon > 0 \exists z_\varepsilon \in A_\varepsilon \cap B_\varepsilon$, 则存在 $x_A^\varepsilon \in A, y_A^\varepsilon \in B(0,\varepsilon), x_B^\varepsilon \in B, y_B^\varepsilon \in B(0,\varepsilon)$ 使得:

$$z_\varepsilon = x_A^\varepsilon + y_A^\varepsilon = x_B^\varepsilon + y_B^\varepsilon$$

于是

$$\|x_A^\varepsilon - x_B^\varepsilon\| = \|y_A^\varepsilon - y_B^\varepsilon\| \leq \|y_A^\varepsilon\| + \|y_B^\varepsilon\| < 2\varepsilon$$

从而根据闭性与紧集必为自列紧集的性质知存在 $z \in A \cap B$, 矛盾! 故 $A_\varepsilon, B_\varepsilon$ 在 ε 足够小时是不相交的非空开凸集, 因而由凸集分离定理3.5.8知存在闭超平面 H_f^α 分离 $A_\varepsilon, B_\varepsilon$, 即

$$f(x+\varepsilon z) \leq \alpha \leq f(y+\varepsilon z), \quad \forall x \in A, y \in B, z \in B(0,1)$$

又注意到

$$f(x+\varepsilon z) = f(x) + \varepsilon f(z) \leq \alpha \Rightarrow f(x) + \varepsilon \sup_{z \in B(0,1)} f(z) \leq \alpha \Rightarrow f(x) + \varepsilon \|f\| \leq \alpha, \quad \forall x \in A$$

且

$$\alpha \leq f(y+\varepsilon z) = f(y) + \varepsilon f(z) \leq f(y) + \varepsilon \|f\|, \quad \forall y \in B$$

于是

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

又因为 $\|f\| \neq 0$, 这便说明 H_f^α 严格分离 A 和 B. □

推论 3.5.4 (Ascoli 定理)

设 E 是实 B^* 空间 $\mathscr{X}$ 中的闭凸集, 则 $\forall x_0 \in \mathscr{X} \backslash E, \exists f \in \mathscr{X}^*$ 及 $\alpha \in \mathbb{R}$, 满足

$$f(x) < \alpha < f(x_0) \quad (\forall x \in E)$$

证明

因为 $x_0 \in \mathscr{X} \backslash E$, 而 E 是闭集, 故存在 $\delta > 0$ 使得

$$B(x_0,\delta) \subset \mathscr{X} \backslash E$$

又因为 $B(x_0,\delta)$ 是有内点的凸集, 故对 E 和 $B(x_0,\delta)$ 应用凸集分离定理(3.5.8)知, 存在非零连续线性泛函 f 满足

$$\sup_{x \in E} f(x) \leq \inf_{y \in B(x_0,\delta)} f(y)$$

现在说明

$$\inf_{y \in B(x_0,\delta)} f(y) < f(x_0) \tag{3.37}$$

如若不然, 则

$$\forall y \in B(x_0,\delta)(f(y) \geq f(x_0))$$

这说明 $f(x_0)$ 是 $f(y)$ 在 $B(x_0,\delta)$ 中的极小值, 但这与练习(3.9)矛盾! 故(3.37)式成立. 取 $\alpha \in (\inf_{y \in B(x_0,\delta)} f(y), f(x_0))$,

得到

$$\sup_{x\in E} f(x) \le \inf_{y\in B(x_0,\delta)} f(y) < \alpha < f(x_0)$$

这说明

$$f(x) < \alpha < f(x_0), \quad \forall x \in E$$

命题得证. □

推论 3.5.5 (Mazur 定理)

设 E 是 B^* 空间 $\mathscr{X}$ 上的一个有内点的闭凸集, F 是 $\mathscr{X}$ 上的一个线性流形, 又设 $E^\circ \cap F = \varnothing$, 则存在一个包含 F 的闭超平面 L, 使得 E 在 L 的一侧.

证明

设 $F = x_0 + \mathscr{X}_0$, 其中 $x_0 \in \mathscr{X}$, $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间. 由凸集分离定理(3.5.8)知, 存在 H_f^r 分离 E 与 F, 即

$$f(E) \le r, \quad f(x_0, \mathscr{X}_0) \ge r$$

记 $r_0 := r - f(x_0)$, 得到 $\forall x \in \mathscr{X}_0 (f(x) \ge r_0)$. 注意因为 $\mathscr{X}_0$ 是线性子空间, 故必有 $\theta \in \mathscr{X}_0$, 从而至少有 $r_0 \le 0$. 现在若 f 可取负值, 设 $f(\zeta) < 0, \zeta \in \mathscr{X}_0$, 根据线性性总能取 $\lambda = \frac{2r_0}{f(\zeta)}$, 使得 $\lambda\zeta \in \mathscr{X}_0, f(\lambda\zeta) = \lambda f(\zeta) = 2r_0 \le r_0$, 等号当且仅当 $r_0 = 0$ 成立. 当 $r_0 = 0$, 自然与 ζ 的取定矛盾, 而当 $r_0 < 0$, 则与 f 满足的条件矛盾. 这说明 f 不能取负值, 又因为 θ 此时成为 f 在开球 $B(\theta, 1)$ 内的极小值点, 故只能有

$$f(x) \equiv 0, \quad \forall x \in \mathscr{X}_0$$

故 $\mathscr{X}_0 \subset H_f^0$, 进而 $F \subset x_0 + H_f^0$, $s := f(x_0)$, 从而

$$f(E) \le r \le f(x_0 + \mathscr{X}_0) = f(x_0) = s$$

这说明 $K := H_f^s$ 即为欲求. □

注 Mazur 定理(3.5.5)可以理解成若在 B^* 空间 $\mathscr{X}$ 上给定有内点的闭凸集 E 与线性流形 F, 且 $E^\circ \cap F = \varnothing$, 则存在超平面 $L := H_f^s$ 使得 E 在 L 的一侧, 而 F 恰好在 L 中.

下面推广平面情况里直线和圆相切的概念.

定义 3.5.5 (承托超平面)

超平面 $L := H_f^r$ 称为凸集 E 在点 x_0 的承托超平面, 是指 E 在 L 的一侧, 且 $\overline{E}$ 与 L 有公共点 x_0. 也即

$$f(x) \le r = f(x_0) \quad (\forall x \in E)$$

或

$$f(x) \ge r = f(x_0) \quad (\forall x \in E)$$

例 3.25(B^* 空间中的球与超平面" 相切") 设 $\mathscr{X}$ 是 B^* 空间, $E = \{x \in \mathscr{X} : \|x\| \le r\}$, $\|x_0\| = r$, 则 E 在 x_0 有一个承托超平面.

证明

由推论(3.5.2)知存在 $f \in \mathscr{X}^*$ 使得 $f(x_0) = \|x_0\|, \|f\| = 1$, 则 H_f^r 就是欲求的超平面, 这是因为

$$f(x) \le \|f\| \cdot \|x\| \le r = f(x_0) \quad (\forall x \in E)$$

□

一般情形由下述定理给出.

定理 3.5.9

设 E 是实 B^* 空间中含有内点的闭凸集, 那么通过 E 的每个边界点都可以作出 E 的一个承托超平面.

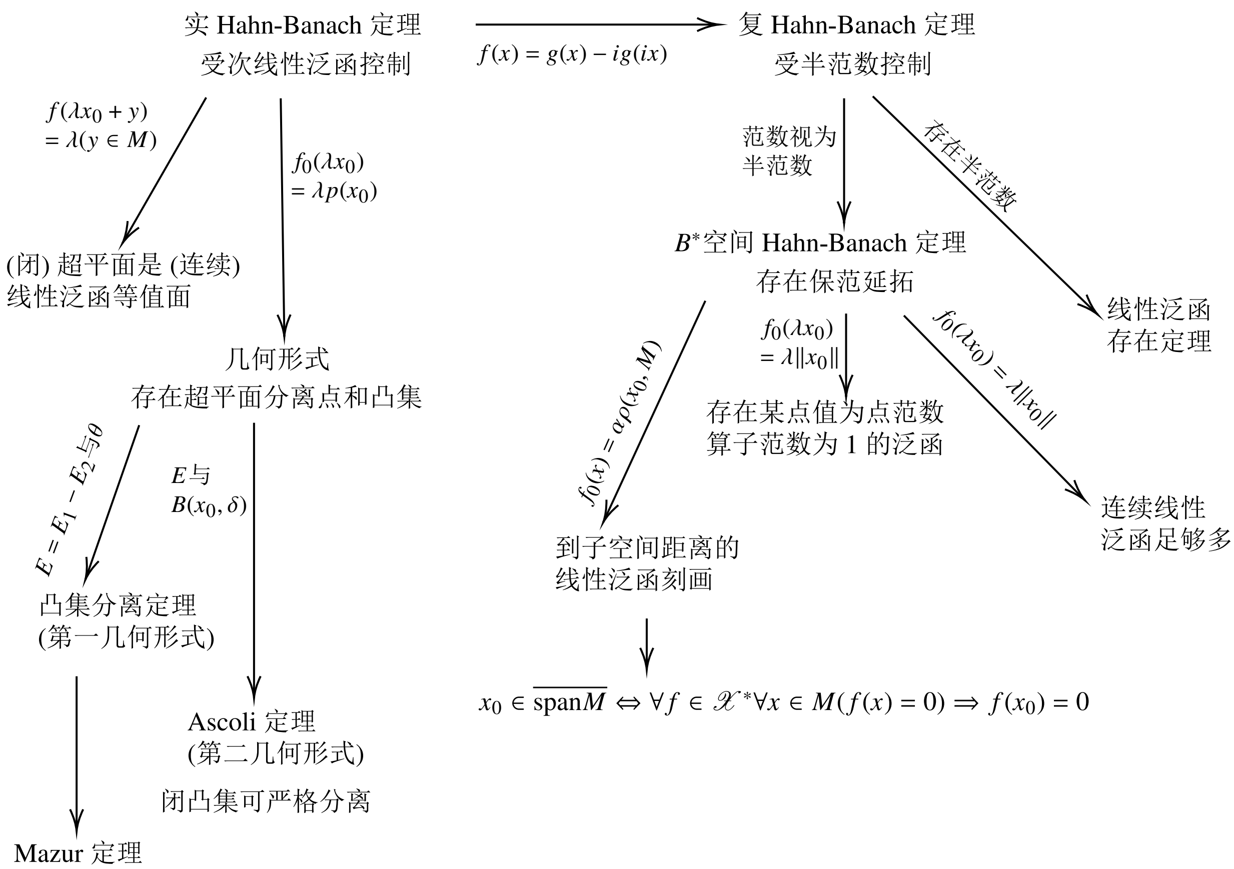
证明

任取 $x_0 \in E \backslash E^\circ$, 取 $F := \{x_0\}$, 知 F 可以看做是线性子空间 $\{\theta\}$ 的平移, 因而是线性流形. 由 Mazur 定理(3.5.5)知存在 $f \in \mathscr{X}^* \backslash \{\theta\}$ 与 $s \in \mathbb{R}$, 使得

$$f(x) \le s = f(x_0), \quad \forall x \in E$$

故 H_f^s 是 E 在 x_0 的承托超平面. □

注 Hahn-Banach 定理一节的定理与推论庞杂繁多, 下面在节末另作一个梳理图.



3.5.1.3 应用: 抽象可微函数的中值定理

定义 3.5.6 (抽象函数与其微商, 可微)

设 $\mathscr{Y}$ 是 B^* 空间, $t \in (a, b)$. 称 $f : (a, b) \to \mathscr{Y}$ 为数值变数 t 的抽象函数. 若在 $\mathscr{Y}$ 中存在极限

$$\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

则定义该极限为 f 在 t 点的微商, 记为 $f'(t)$. 又若 f 在 (a, b) 内处处有微商, 则称 f 在 (a, b) 内可微.

定理 3.5.10 (抽象可微函数中值定理)

设 $\mathscr{Y}$ 是 B^* 空间, 抽象函数 $f : (a, b) \to \mathscr{Y}$ 在 (a, b) 内可微, 则

$$\forall t_1, t_2 \in (a, b) \exists \theta \in (0, 1)(\|f(t_2) - f(t_1)\| \le \|f'(\theta t_2 + (1 - \theta)t_1)\| \cdot |t - 2 - t_1|)$$

证明

由推论(3.5.2)知存在 $y^* \in \mathscr{Y}^*$ 使得 $\|y^*\| = 1$, 且

$$y^*(f(t_2) - f(t_1)) = \|f(t_2) - f(t_1)\|$$

记

$$\varphi(\eta) = y^*(t_1 + \eta(t_2 - t_1))$$

则 $\varphi(\eta)$ 是在 $[0,1]$ 上连续, 在 $(0,1)$ 内可微的实函数. 为表述方便, 根据 Riesz 表示定理, 取 $y \in \mathscr{Y}$ 使得

$$y^*(x) = \langle y, x\rangle \Rightarrow \varphi(\eta) = \langle y, f(t_1 + \eta(t_2 - t_1))\rangle$$

得到

$$\varphi'(\eta) = \langle y, f'(t_1 + \eta(t_2 - t_1))(t_2 - t_1)\rangle$$

对 $\varphi(\eta)$ 应用微分中值定理, 得到 $\theta \in (0,1)$ 使得

$$\varphi(1) - \varphi(0) = \varphi'(\theta) = \langle y, f'(t_1 + \theta(t_2 - t_1))(t_2 - t_1)\rangle$$

注意 $\|y\| = \|y^*\| = 1$, 有

$$\|f(t_2) - f(t_1)\| = \varphi(1) - \varphi(0) \le \|y\| \cdot \|f'(t_1 + \theta(t_2 - t_1))(t_2 - t_1)\| = \|f'(t_1 + \theta(t_2 - t_1))\| \cdot |t_2 - t_1|$$

此即欲证. □

3.5.1.4 应用: 凸规划问题的 Lagrange 乘子

凸集分离定理(3.5.8)是数学规划理论的基础, 本节介绍 Kuhn-Tucker 定理.

定义 3.5.7 (凸泛函)

设 $\mathscr{X}$ 是一个线性空间, $C \subset \mathscr{X}$ 是一个凸集. 称 $f: C \to \mathbb{R}$ 是一个凸泛函, 是指 f 满足

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \quad (\forall x, y \in C, \forall \lambda \in (0,1))$$

♣

注 该定义可等价地表述为: 上方图

$$\operatorname{epi}(f) := \{(x,y) \in C \times \mathbb{R} : f(x) \le t\}$$

是 $C \times \mathbb{R}$ 中的凸集.

凸规划问题 (简便起见记作 P) 指: 给定凸集 C 上的凸函数 $f, g_1, g_2, \cdots, g_n$, 求 $x_0 \in C$, 满足

$$g_i(x_0) \le 0 \quad (i = 1, 2, \cdots, n)$$

且

$$f(x_0) = \min\{f(x) : x \in C, g_i(x) \le 0 (i = 1, 2, \cdots, n)\}$$

其中条件 $g_i(x) \le 0 (i = 1, 2, \cdots, n)$ 称为约束.

过去对于带约束的极值问题, 可以通过 Lagrange 乘子法将其转化为无约束的极值问题, 现在也用类似的思路来考虑问题 P, 也即寻找一组 $(\hat{\lambda}_1, \cdots, \hat{\lambda}_n) \in \mathbb{R}^n$, 使得若 x_0 是问题 P 的解, 则

$$f(x_0) + \sum_{i=1}^n \hat{\lambda}_i g_i(x_0) = \min\{f(x) + \sum_{i=1}^n \hat{\lambda}_i g_i(x) : x \in C\}$$

这等价于说

$$f(x_0) + \sum_{i=1}^n \hat{\lambda}_i g_i(x_0) \le f(x) + \sum_{i=1}^n \hat{\lambda}_i g_i(x) \quad (\forall x \in C) \tag{3.38}$$

为了转化成凸集分离问题, 针对 $f(x)$ 再引进一个参数 $\hat{\lambda}_0$, 考察不等式

$$\hat{\lambda}_0 f(x_0) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \le \hat{\lambda}_0 f(x) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x) \quad (\forall x \in C) \tag{3.39}$$

如果能证明 $\hat{\lambda}_0 > 0$, 则(3.39)式就等价于(3.38)式.

现在要寻找非零的 $(\hat{\lambda}_0, \hat{\lambda}_1, \cdots, \hat{\lambda}_n) \in \mathbb{R}^{n+1}$ 满足不等式(3.39), 在几何上就等价于在 $\mathbb{R}^{n+1}$ 上找一个超平面, 这个超平面分离集合

$$E := \{(t_0, t_1, \cdots, t_n) \in \mathbb{R}^{n+1} : t_0 \le f(x_0), t_i \le 0 (i = 1, 2, \cdots, n)\}$$

与

$$F := \{(t_0, t_1, \cdots, t_n) \in \mathbb{R}^{n+1} : \exists x \in C (t_0 \ge f(x), t_i \ge g_i(x)(i = 1, 2, \cdots, n))\}$$

因为 $f, g_1, \cdots, g_n$ 都是凸函数, 知 F 是 $\mathbb{R}^{n+1}$ 中的一个凸集, 又显见 E 有内点, 其内部为

$$E^\circ = \{(t_0, t_1, \cdots, t_n) \in \mathbb{R}^{n+1} : t_0 < f(x_0), t_i < 0 (i = 1, 2, \cdots, n)\}$$

因为 x_0 是问题 P 的解, 故 $E^\circ \cap F = \varnothing$, 现在对线性泛函

$$p(x) := t_0 f(x) + \sum_{i=1}^{n} t_i g_i(x) \quad ((t_0, \cdots, t_n) \in \mathbb{R}^{n+1})$$

应用凸集分离定理(3.5.8)得到 $\mathbb{R}^{n+1}$ 中的点 $(\hat{\lambda}_0, \hat{\lambda}_1, \cdots, \hat{\lambda}_n)$, 它们满足:

$$\hat{\lambda}_0 f(x_0) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \le \hat{\lambda}_0 f(x) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x) \quad (\forall x \in C) \tag{3.40}$$

为了方便后续讨论, 从(3.40)式容易推知下述不等式

$$\hat{\lambda}_0 f(x_0) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \le \hat{\lambda}_0 (f(x) + \xi_0) + \sum_{i=1}^{n} \hat{\lambda}_i (g_i(x) + \xi_i) \quad (\forall x \in C, \forall \xi_i \ge 0 (i = 0, 1, \cdots, n)) \tag{3.41}$$

现令 $x = x_0$, 得到

$$\sum_{i=0}^{n} \hat{\lambda}_i \xi_i \ge 0 \quad (\forall \xi_i \ge 0 (i = 0, 1, \cdots, n)) \tag{3.42}$$

如若有某个 $\hat{\lambda}_k < 0$, 则取 $\xi_i = 0 (i \ne k), \xi_k > 0$ 即得矛盾, 这说明 $\hat{\lambda}_i \ge 0 (i = 0, 1, \cdots, n)$, 且(3.39)式确实成立. 现在说明

$$\hat{\lambda}_i g_i(x_0) = 0 \quad (i = 1, 2, \cdots, n)$$

这是因为一方面, 根据凸集分离定理(3.5.8)知[10]

$$\hat{\lambda}_0 f(x_0) \le \hat{\lambda}_0 f(x_0) + \sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \Rightarrow \sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \ge 0$$

另一方面根据假设中的 $g_i(x_0) \le 0$ 与 $\hat{\lambda}_i \ge 0 (i = 1, 2, \cdots, n)$ 知

$$\sum_{i=1}^{n} \hat{\lambda}_i g_i(x_0) \le 0$$

故

$$\hat{\lambda}_i g_i(x_0) = 0 \quad (i = 1, 2, \cdots, n)$$

这说明指标 i 在约束上不起作用. 下面确定使 $\hat{\lambda}_0 > 0$ 的条件.

引理 3.5.1

若 $\exists \hat{x} \in C$ 满足

$$g_i(\hat{x}) < 0 \quad (i = 1, 2, \cdots, n)$$

[10] 这里完全没有想明白是怎么成立的

则 $\hat{\lambda}_0 > 0$.

3.5.2 习题

练习 3.29 设 p 是实线性空间 $\mathscr{X}$ 上的次线性泛函, 求证:

(1) $p(\theta) = 0$.

证明

根据次线性泛函的定义知

$$\forall \lambda > 0 \forall x \in \mathscr{X}(p(\lambda x) = \lambda p(x))$$

代入 $x = \theta, \lambda = 2$ 即得 $p(\theta) = 0$.

(2) $p(-x) \geq -p(x)$

证明

根据次线性泛函的定义知

$$\forall x, y \in \mathscr{X}(p(x+y) \leq p(x) + p(y))$$

令 $y = -x$, 结合 (1) 有

$$p(\theta) = 0 \leq p(x) + p(-x)$$

也即

$$p(-x) \geq -p(x)$$

(3) 任意给定 $x_0 \in \mathscr{X}$, 在 $\mathscr{X}$ 上必有实线性泛函 f, 满足 $f(x_0) = p(x_0)$, 以及 $f(x) \leq p(x)(\forall x \in \mathscr{X})$.

证明

记 $\mathscr{X}_0 := \{\lambda x_0 : \lambda \in \mathbb{R}\}$, 并令 $f_0(\lambda x_0) := \lambda p(x_0), \forall \lambda \in \mathbb{R}$. 从定义中显见 $\mathscr{X}_0$ 是 $\mathscr{X}$ 的线性子空间, 且 $f_0 \in \mathscr{X}_0^*$, 下面证明 $\forall x \in \mathscr{X}_0(f_0(x) \leq p(x))$. 这是因为任取 $x \in \mathscr{X}_0$, 根据 $\mathscr{X}_0$ 的构造可设 $x = \alpha_x x_0$. 当 $\alpha_x \geq 0$, 命题无所可证, 而当 $\alpha_x < 0$, 知

$$f_0(x) = \alpha_x p(x_0) = -(-\alpha_x p(x_0)) = -p(-\alpha_x x_0) \leq p(\alpha_x x_0) = p(x)$$

故 $\forall x \in \mathscr{X}_0(f_0(x) \leq p(x))$, 从而根据实 Hahn-Banach 定理知存在 $f \in \mathscr{X}^*$ 满足

$$\forall x \in \mathscr{X}(f(x) \leq p(x)), \quad \forall x \in \mathscr{X}_0(f(x) = f_0(x))$$

特别代入 $x = x_0$ 即得 $f(x_0) = p(x_0)$.

练习 3.30 设 $\mathscr{X}$ 是由实数列 $x = \{a_n\}$ 全体组成的实线性空间, 其元素间相等和线性运算都按坐标定义, 并定义

$$p(x) = \overline{\lim_{n\to\infty}} \alpha_n \quad (\forall x = \{\alpha_n\} \in \mathscr{X})$$

求证: $p(x)$ 是 $\mathscr{X}$ 上的次线性泛函.

证明

任取 $x = \{\alpha_n\} \in \mathscr{X}, y = \{\beta_n\} \in \mathscr{X}$, 知

$$p(x+y) = \overline{\lim_{n\to\infty}}(\alpha_n + \beta_n) \leq \overline{\lim_{n\to\infty}} \alpha_n + \overline{\lim_{n\to\infty}} \beta_n = p(x) + p(y)$$

任取 $\lambda > 0$ 知

$$p(\lambda x) = \overline{\lim_{n\to\infty}} \lambda\alpha_n = \lambda \overline{\lim_{n\to\infty}} \alpha_n = \lambda p(x)$$

命题即证.

练习 3.31★♡ 设 $\mathscr{X}$ 是复线性空间, p 是 $\mathscr{X}$ 上的半范数. $\forall x_0 \in \mathscr{X}, p(x_0) \neq 0$[11]. 求证: 存在 $\mathscr{X}$ 上的线性泛函 f 满足

1. $f(x_0) = 1$;
2. $|f(x)| \leq \frac{p(x)}{p(x_0)} (\forall x \in \mathscr{X})$.

证明

固定 $x_0 \in \mathscr{X}$, 知 $l(x) := \frac{p(x)}{p(x_0)}$ 依旧是 $\mathscr{X}$ 上的半范数, 定义 $\mathscr{X}_0 := \{\lambda x_0 : \lambda \in \mathbb{C}\}$, 同时取 $f_0(\lambda x_0) := \lambda l(x_0) = \lambda$, 显见 $f_0 \in \mathscr{X}_0^*$. 现在说明 $\forall x \in \mathscr{X}_0(|f_0(x)| \leq l(x))$, 这是因为任取 $x \in \mathscr{X}_0$, 根据 $\mathscr{X}_0$ 的构造知可设 $x = \alpha_x x_0$, 进而

$$|f_0(x)| = |\alpha_x| = |\alpha_x| \cdot l(x_0) \leq l(\alpha_x x_0) = l(x)$$

故由复 Hahn-Banach 定理知存在 $\mathscr{X}$ 上的线性泛函满足

$$\forall x \in \mathscr{X}\left(|f(x)| \leq l(x) = \frac{p(x)}{p(x_0)}\right), \quad \forall x \in \mathscr{X}_0(f(x) = f_0(x))$$

特别取 $x = x_0$ 即得 $f(x_0) = 1$.

练习 3.32★♡ 设 $\mathscr{X}$ 是 B^* 空间, $\{x_n\}(n = 1, 2, \cdots)$ 是 $\mathscr{X}$ 中的点列. 如果 $\forall f \in \mathscr{X}^*$, 数列 $\{f(x_n)\}$ 有界, 求证: $\{x_n\}$ 在 $\mathscr{X}$ 内有界.

证明

首先证明

$$\forall x_0 \in \mathscr{X}\left(\|x_0\| = \sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} |f(x_0)|\right) \tag{3.43}$$

一方面, $\forall f \in \mathscr{X}^*(\|f\| = 1)$ 说明 $\forall f \in \mathscr{X}^*(\sup_{x \in \mathscr{X}} |f(x)| = \|x\|)$, 从而 $\forall f \in \mathscr{X}^*(|f(x_0)| \leq \|x_0\|)$, 亦即 $\sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} |f(x_0)| \leq \|x_0\|$.

另一方面, 任取 $\theta \neq x_0 \in \mathscr{X}$, 知 $\mathscr{X}_0 = \{\lambda x_0 : \lambda \in \mathbb{R}\}$ 是 $\mathscr{X}$ 的闭线性子空间, 定义 $\forall x \in \mathscr{X}_0(f(x) = f(\lambda x_0) = \lambda\|x_0\|)$, 则 f 首先是线性泛函, 下证 f 有界. 注意 $\|\cdot\|$ 是 $\mathscr{X}$ 上的次线性泛函, 且

$$\frac{|f(x)|}{\|x\|} = \frac{|f(\lambda x_0)|}{\|\lambda x_0\|} = \frac{|\lambda\|x_0\||}{|\lambda|\|x_0\|} = 1 \Rightarrow \|f\| = 1$$

故一方面 f 有界, 另一方面 $|f|$ 被 $\mathscr{X}$ 上的一个次线性泛函控制, 因而由 Hahn-Banach 定理知存在 $\widetilde{f} \in \mathscr{X}^*$ 使得

(1) $\forall x \in \mathscr{X}_0(\widetilde{f}(x) = f(x))$;

(2) $\forall x \in \mathscr{X}(|\widetilde{f}(x)| \leq \|x\|)$.

特别有 $|\widetilde{f}(x_0)| = |f(x_0)| = \|x_0\|$.

综上即证(3.97)式. 现注意任取 $x_0 \in \mathscr{X}$, 总存在 $\mathscr{X}^{**}$ 中的元素 x_0^{**}, 使得

$$\langle x_0^{**}, f\rangle = f(x_0), \quad \forall f \in \mathscr{X}^*$$

且由(3.97)式知 $\|x_0^{**}\|_{\mathscr{X}^{**}} = \|x_0\|_{\mathscr{X}}$. 因为 $\mathscr{X}$ 是 B^* 空间, 故 $\mathscr{X}^* = \mathscr{L}(\mathscr{X}, \mathbb{R})$ 是 B 空间, 进一步 $\mathscr{X}^{**} = \mathscr{L}(\mathscr{X}^*, \mathbb{R})$ 是 B 空间, 从而条件即

$$\forall f \in \mathscr{X}^*\left(\sup_{n \geq 1} |\langle x_n^{**}, f\rangle| < \infty\right)$$

根据共鸣定理知存在常数 M 使得 $\sup_{n \geq 1} \|x_n^{**}\| < \infty$, 亦即 $\sup_{n \geq 1} \|x_n\| < \infty$, 命题得证.

练习 3.33★♡ 设 $\mathscr{X}_0$ 是 B^* 空间 $\mathscr{X}$ 的闭子空间, 求证:

$$\rho(x, \mathscr{X}_0) = \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \quad (\forall x \in \mathscr{X})$$

其中 $\rho(x, \mathscr{X}_0) = \inf_{y \in \mathscr{X}_0} \|x - y\|$.

证明

[11] 这句话的意思是对任意固定的 x_0.

固定 $x_0 \in \mathscr{X}$, 根据定理(3.5.5)知存在 $f \in \mathscr{X}^*$ 满足

$$\forall x \in M(f(x) = 0), \quad f(x_0) = \rho(x_0, \mathscr{X}_0), \quad \|f\| = 1$$

这说明

$$\rho(x_0, \mathscr{X}_0) \leq \sup\{|f(x_0)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\}$$

注意 $x_0 \in \mathscr{X}$ 是任取的, 因而

$$\rho(x, \mathscr{X}_0) \leq \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \quad (\forall x \in \mathscr{X})$$

再说明

$$\rho(x, \mathscr{X}_0) \geq \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \quad (\forall x \in \mathscr{X})$$

这是因为根据练习(3.8)(2), 注意到 $M \subset H_f^0$, 故

$$|f(x)| = \rho(x, H_f^0) \leq \rho(x, \mathscr{X}_0) \quad (\forall x \in \mathscr{X})$$

这说明

$$\rho(x, \mathscr{X}_0) \geq \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \quad (\forall x \in \mathscr{X})$$

综上即得

$$\rho(x, \mathscr{X}_0) = \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \quad (\forall x \in \mathscr{X})$$

✍ **练习 3.34**♡ 设 $\mathscr{X}$ 是 B^* 空间. 给定 $\mathscr{X}$ 中 n 个线性无关的元素 $x_1, x_2, \cdots, x_n$ 与数域 $\mathbb{K}$ 中的 n 个数 $C_1, C_2, \cdots, C_n$, 及 $M > 0$. 求证: 为了 $\exists f \in \mathscr{X}^*$ 适合 $f(x_k) = C_k (k = 1, 2, \cdots, n)$, 以及 $\|f\| \leq M$, 必须且仅须对任意的 $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$, 有

$$|\sum_{k=1}^n \alpha_k C_k| \leq M \|\sum_{k=1}^n \alpha_k x_k\|.$$

证明

当

$$\exists f \in \mathscr{X}^*(f(x_k) = C_k (k = 1, 2, \cdots, n), \|f\| \leq M)$$

知

$$|f(\sum_{k=1}^n \alpha_k x_k)| = |\sum_{k=1}^n \alpha_k C_k| \leq \|f\| \cdot \|\sum_{k=1}^n \alpha_k x_k\| \leq M \|\sum_{k=1}^n \alpha_k x_k\|$$

当

$$\forall \alpha_1, \cdots, \alpha_n \in \mathbb{K}(|\sum_{k=1}^n \alpha_k C_k| \leq M \|\sum_{k=1}^n \alpha_k x_k\|)$$

记 $E = \operatorname{span}\{x_k : k = 1, 2, \cdots, n\}$, 取 $f_0(\sum_{k=1}^n \alpha_k x_k) := \sum_{k=1}^n \alpha_k C_k (\forall \alpha_i \in \mathbb{K}(i = 1, 2, \cdots, n))$. 显见 E 是 $\mathscr{X}$ 的线性子空间, 且 $|f_0(x)| \leq p(x) := M\|x\|, \forall x \in E$. 故由 Hahn-Banach 定理知存在 $f \in \mathscr{X}^*$ 满足

$$\forall x \in E(f(x) = f_0(x)), \quad \forall x \in \mathscr{X}(|f(x)| \leq p(x) = M\|x\|) \Rightarrow \|f\| \leq M$$

特别代入 $x = x_k (k = 1, 2, \cdots, n)$ 得

$$f(x_k) = C_k (k = 1, 2, \cdots, n)$$

✍ **练习 3.35**★♡ 给定 B^* 空间 $\mathscr{X}$ 中 n 个线性无关的元素 $x_1, x_2, \cdots, x_n$, 求证: $\exists f_1, f_2, \cdots, f_n \in \mathscr{X}^*$, 使得

$$\langle f_i, x_j \rangle = \delta_{ij} \quad (i, j = 1, 2, \cdots, n).$$

证明

只需证明存在 $f_1 \in \mathscr{X}^*$ 使得 $\langle f_1, x_j\rangle = \delta_{1j}(j = 1, 2, \cdots, n)$ 即可. 设 $(x_1, x_2, \cdots, x_n)$ 经过线性变换 A 后变为了正交集 $(e_1, e_2, \cdots, e_n)$, 其中 $e_1 = x_1$. 设 $E = \text{span}\{e_1, e_2, \cdots, e_n\}$, 并取 $f_0(\sum\limits_{k=1}^{n} \alpha_k e_k) = \alpha_1 e_1$. 显见 E 是 $\mathscr{X}$ 的线性子空间, 且 $f_0 \in E^*$. 根据 Hahn-Banach 定理知存在 $f \in \mathscr{X}^*$ 使得

$$\forall x \in E(f(x) = f_0(x)), \quad \|f\| = \|f_0\|$$

现在取 $f_1(x) = f(A^{-1}x)$ 即得满足条件的 f_1. $f_k(k = 2, \cdots, n)$ 同理.

✍ **练习 3.36★** 设 $\mathscr{X}$ 是线性空间, 求证: 为了 M 是 $\mathscr{X}$ 的极大线性子空间, 必须且仅须 $\dim(\mathscr{X}/M) = 1$.

证明

由命题(3.5.1)知 M 是极大线性子空间当且仅当 M 是线性真子空间, 且对任意的 $x_0 \in \mathscr{X}\backslash M$ 有

$$\mathscr{X} = \{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M$$

故当 M 是极大线性子空间, 知取定 x_0 有

$$\mathscr{X}/M = \{\lambda x_0 : \lambda \in \mathbb{R}\} \Rightarrow \dim(\mathscr{X}/M) = 1$$

当 $\dim(\mathscr{X}/M) = 1$, 设 $\mathscr{X}/M$ 的基为 x_0, 知 $\mathscr{X}$ 中的任意元素 x 都可表成

$$x = x_0 + b, \quad b \in M$$

进一步对任意的 $y \in \mathscr{X}\backslash M$, 总可将 x 换成 $x + x_0 - y$, 从而 $\mathscr{X}$ 中的元素可表成

$$x = y + b, \quad b \in M$$

这里 $y \in \mathscr{X}_0\backslash M$ 可以任意取, 这便符合了

$$\mathscr{X} = \{\lambda x_0 : \lambda \in \mathbb{R}\} \oplus M, \quad \forall x_0 \in \mathscr{X}\backslash M$$

同时易知 M 是线性真子空间, 故 M 是 $\mathscr{X}$ 的极大线性子空间.

✍ **练习 3.37★** 设 $\mathscr{X}$ 是复线性空间, E 是 $\mathscr{X}$ 中的非空均衡集, f 是 $\mathscr{X}$ 上的线性泛函. 求证:

$$|f(x)| \leq \sup_{y\in E} \text{Re}\, f(y) \quad (\forall x \in E).$$

证明

知

$$f(x) = |f(x)|e^{i\theta_x}, \quad \forall x \in \mathscr{X}$$

故

$$|f(x)| = f(x)e^{-i\theta_x} = f(e^{-i\theta_x}x) \leq \sup_{\substack{x\in\mathscr{X}\\ \theta_x\in\mathbb{R}}} f(e^{-i\theta_x}x) = \sup_{\substack{x\in\mathscr{X}\\ \theta_x\in\mathbb{R}}} \text{Re}\, f(e^{-i\theta_x}x)$$

其中最后一步是因为 $f(e^{-i\theta_x}x)$ 是实数. 根据 E 的均衡性知

$$x \in E \Rightarrow \forall\theta \in \mathbb{R}(e^{i\theta}x \in E)$$

因而

$$\sup_{\substack{x\in\mathscr{X}\\ \theta_x\in\mathbb{R}}} \text{Re}\, f(e^{-i\theta_x}x) = \sup_{y\in E} \text{Re}\, f(y) \quad (\forall x \in E)$$

综上即得

$$|f(x)| \leq \sup_{y\in E} \text{Re}\, f(y) \quad (\forall x \in E).$$

✍ **练习 3.38★** 设 $\mathscr{X}$ 是 B^* 空间, $E \subset \mathscr{X}$ 是非空的均衡闭凸集, $\forall x_0 \in \mathscr{X} \backslash E$, 求证: $\exists f \in \mathscr{X}^*$ 及 $\alpha > 0$, 使得

$$|f(x)| < \alpha < |f(x_0)| \quad (\forall x \in E).$$

证明

若将 $\mathscr{X}$ 视作实 B^* 空间, 则因为 E 是闭凸集, 由 Ascoli 定理知存在 $\widetilde{f} \in \mathscr{X}^*$ 与 $\beta \in \mathbb{R}$ 使得

$$\widetilde{f}(x) < \beta < \widetilde{f}(x_0), \quad \forall x \in E$$

于是

$$\sup_{x \in E} \widetilde{f}(x) \le \beta < \widetilde{f}(x_0).$$

现取 $f(x) = \widetilde{f}(x) + i\widetilde{f}(-ix)(x \in \mathscr{X})$, 则 f 是复 B^* 空间 $\mathscr{X}$ 上的复线性泛函, 进一步有 $f \in \mathscr{X}^*$. 设 $f(x) = e^{i\theta_x}|f(x)|$, 则

$$|f(x)| = e^{-i\theta_x} f(x) = f(e^{-i\theta_x} x) \overset{\text{(i)}}{=} \widetilde{f}(e^{-i\theta_x} x) \overset{\text{(ii)}}{\le} \beta, \quad \forall x \in E$$

其中 (i) 是因为 $f(e^{-i\theta_x}x)$ 是实数, 故其与其实部 $\widetilde{f}(e^{-i\theta_x}x)$ 相等; (ii) 是因为 E 是均衡的, 故 $x \in E \Rightarrow e^{-i\theta_x}x \in E$, 从而 $\widetilde{f}(e^{-i\theta_x}x) \le \sup\limits_{x \in E} \widetilde{f}(x) \le \beta$.

另一边, $|f(x_0)| \ge \widetilde{f}(x_0) > \beta$, 故只需取 $\alpha \in (\beta, \widetilde{f}(x_0))$, 则有:

$$|f(x)| \le \beta < \alpha < \widetilde{f}(x_0) \le |f(x_0)|, \quad \forall x \in E$$

命题即证.

✍ **练习 3.39★** 设 E, F 是实的 B^* 空间 $\mathscr{X}$ 中的两个互不相交的非空凸集, 并且 E 是开的和对称的. 求证: $\exists f \in \mathscr{X}^*$, 使得

$$|f(x)| < \inf_{y \in F} |f(y)|, \quad \forall x \in E.$$

证明

因为 E 是开对称凸集, 故 θ 是 E 的内点, 从而根据凸集分离定理知存在不为常值的 $f \in \mathscr{X}^*$ 使得

$$\sup_{x \in E} f(x) \le \inf_{x \in F} f(x)$$

因为 E 是对称凸集, 故 $\sup\limits_{x \in E} f(x) = \sup\limits_{x \in E} |f(x)|$. 而由 $\forall x \in F(f(x) \le |f(x)|)$ 可知 $\inf\limits_{x \in F} f(x) \le \inf\limits_{x \in F} |f(x)|$, 故有

$$\sup_{x \in E} |f(x)| \le \inf_{x \in F} |f(x)|$$

再来说明 E 中不存在 x_0 使得 $|f(x_0)| = \sup\limits_{x \in E} |f(x)|$. 用反证法, 若存在这样的 x_0, 根据前述讨论显见 $f(x_0) \ge 0$, 故 $|f(x_0)| = f(x_0) = \sup\limits_{x \in E} |f(x)|$. 因为 E 是开集, 故由 $x_0 \in E$ 知存在 $\delta > 0$ 使得 $B(x_0, \delta) \subset E$, 进而 $f(x_0)$ 至少是开球 $B(x_0, \delta)$ 中的极大值, 从而 f 在 $B(x_0, \delta)$ 上是常值泛函, 因而根据 f 的线性性知其在 $B(\theta, \delta)$ 内也是常值泛函. 但显见 $f(\theta) = 0$, 故 $f(\mathscr{X}) = 0$, 矛盾! 从而

$$\forall x \in E(f(x) < \sup_{x \in E} |f(x)|)$$

最后根据 E 的对称性即得

$$|f(x)| < \sup_{x \in E} |f(x)| \le \inf_{x \in F} |f(x)|$$

命题即证.

✍ **练习 3.40★** 设 C 是实 B^* 空间 $\mathscr{X}$ 中的一个凸集, 并设 $x_0 \in C^\circ, x_1 \in \partial C, x_2 = m(x_1 - x_0) + x_0 (m > 1)$, 求证 $x_2 \notin C$.

证明

用反证法, 若 $x_2 \in C$, 设 $\lambda = \frac{1}{m}$, 则 $x_1 = \lambda x_2 + (1 - \lambda)x_0$, 往证 $x_1 \notin \partial C$. 因为 $m > 1$, 故 $\lambda \in (0, 1)$, 从而根据凸集的定义知必有 $x_1 \in C$, 进而只需证明存在 $d > 0$, 使得 $B(x_1, d) \subset C$.

因为 $x_0 \in C^\circ$, 故存在 $\delta > 0$ 使得 $B(x_0, \delta) \subset C$. 现在任取 $y \in B(x_0, \delta)$, 因为 $x_2 \in C$, 根据凸性总有

$$z = \lambda x_2 + (1-\lambda)y \in C$$

因而

$$\begin{cases} x_1 = \lambda x_2 + (1-\lambda)x_0 \\ z = \lambda x_2 + (1-\lambda)y \end{cases} \Rightarrow z - x_1 = (1-\lambda)(y - x_0) \Rightarrow \|z - x_1\| = (1-\lambda)\|y - x_0\| < (1-\lambda)\delta$$

现取 $\delta = (1-\lambda)\delta$, 知当 $\|z - x_1\| < d$ 时, 取辅助点

$$y = x_0 + \frac{z - x_1}{1-\lambda} \in B(x_0, \delta) \subset C$$

于是

$$z = x_1 + (1-\lambda)(y - x_0) = \lambda x_2 + (1-\lambda)x_0 + (1-\lambda)(y - x_0) = \lambda x_2 + (1-\lambda)y \in C$$

这说明 $B(x_1, d) \subset C$, 这与 $x_1 \in \partial C$ 矛盾! 命题即证.

练习 3.41★ 设 M 是 B^* 空间 $\mathscr{X}$ 中的闭凸集, 求证: $\forall x \in \mathscr{X} \backslash M$, 必 $\exists f_1 \in \mathscr{X}^*$, 满足 $\|f_1\| = 1$, 并且

$$\sup_{y \in M} f_1(y) \leq f_1(x) - d(x)$$

其中 $d(x) = \inf\limits_{x \in M} \|x - z\|$.

证明

因为 M 是闭集, 故 $x \in \mathscr{X} \backslash M$ 说明存在 $d > 0$ 使得 $B(x, d) \cap M = \varnothing$. 根据凸集分离定理, 存在 $f \in \mathscr{X}^*, \alpha \in \mathbb{R}$, 使得

$$\forall y \in M \forall z \in B(x, d)(f(y) \leq \alpha \leq f(z))$$

进而

$$\begin{aligned} \sup_{y \in M} f(y) \leq \inf_{z \in B(x,d)} f(z) &= \inf_{w \in B(\theta,1)} f(x - d(x)w) \\ &= \inf_{w \in B(\theta,1)} (f(x) - d(x)f(w)) = f(x) - d(x) \sup_{w \in B(\theta,1)} f(w) \\ &= f(x) - d(x)\|f\| \end{aligned}$$

取用 $f_1(x) = \frac{f(x)}{\|f\|}$ 即得欲求.

练习 3.42 设 M 是实 B^* 空间 $\mathscr{X}$ 内的闭凸集, 求证:

$$\inf_{z \in M} \|x - z\| = \sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} \{f(x) - \sup_{z \in M} f(z)\}, \quad \forall x \in \mathscr{X}.$$

证明

当 $x \in M$, 欲证显然成立. 当 $M = \{z\}$ 时, 欲证即

$$\|x - z\| = \sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} \{f(x) - f(z)\}, \quad \forall x \in \mathscr{X}$$

将 $x - z$ 看作 $\mathscr{X}^{**}$ 上的元素即得欲证. 现设 M 中至少有两个不同的元素, 因为 M 是闭集, 故必存在 $z_0 \in M$ 使得 $\inf\limits_{z \in M} \|x - z\| = \|x - z_0\|$. 容易证明 $z_0 \in \partial M$, 否则存在 $\lambda > 0$ 使得 $z_0' = \lambda z_0 + (1-\lambda)x$, 且 $\|x - z_0'\| \leq \|x - z_0\|$. 现在在 z_0 处知存在 $f \in \mathscr{X}^*$ 使得

$$\forall z \in M(f(z_0) \geq f(z)) \Rightarrow f(z_0) = \sup_{z \in M} f(z)$$

于是欲证即

$$\|x - z_0\| = \sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} \{f(x) - f(z_0)\}, \quad \forall x \in \mathscr{X}$$

将 $x - z_0$ 视作 $\mathscr{X}^{**}$ 中的元素即得结论.

练习 3.43 设 $\mathscr{X}$ 是一个 B 空间, $f : \mathscr{X} \to \overline{R}(:= \mathbb{R} \cup \{\infty\})$ 是连续的凸泛函, 并且 $f(x) \not\equiv \infty$. 若定义 $f^* : \mathscr{X}^* \to \overline{\mathbb{R}}$ 为

$$f^*(x^*) = \sup_{x \in \mathscr{X}} \{\langle x^*, x\rangle - f(x)\}, \quad \forall x^* \in \mathscr{X}^*,$$

求证: $f^*(x^*) \not\equiv \infty$.

练习 3.44 设 $\mathscr{X}$ 是 B 空间, $x(t) : [a,b] \to \mathscr{X}$ 是连续的抽象函数, 又设 Δ 表示 $[a,b]$ 的分割:

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b, \quad \|\Delta\| := \max_{0 \le i \le n-1} \{|t_{i+1} - t_i|\}$$

求证: 在 $\mathscr{X}$ 中存在极限

$$\lim_{\|\Delta\| \to 0} \sum_{i=0}^{n-1} x(t_i)(t_{i+1} - t_i)$$

此极限称为抽象函数 $x(t)$ 在 $[a,b]$ 上的 Riemann 积分.

练习 3.45(推广的 Cauchy 定理) 设 $\mathscr{X}$ 是 Banach 空间, G 是由 $\mathbb{C}$ 中的简单闭曲线 L 围成的开区域. 如果 $x(z) : \overline{G} \to \mathscr{X}$ 在 G 内解析, 且在 $\overline{G}$ 上连续, 求证:

$$\int_L x(z)dz = 0.$$

练习 3.46 求证:

(1) $|x|$ 在 $\mathbb{R}$ 中是凸的;

(2) $|x|$ 在 $x = 0$ 点的次微分 $\partial|x|(0) = [-1, 1]$.

3.5.3 补充: 线性空间微分学概要

本节选自 [**AHCB**], 旨在完善前面提到的抽象函数微分中值定理, 也希望展示非线性泛函分析的一些概念与应用.

3.5.3.1 线性空间中的微分法

定义 3.5.8 (强微分 (Frechet 微分))

设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, $F : \mathscr{X} \supset O \to \mathscr{Y}$ 是映射, 其中 O 是开集. 称 F 在 $x \in O$ 处可微, 如果存在有界线性算子 $L_x \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 满足:

$$\forall \varepsilon > 0 \exists \delta > 0 (\|h\| \Rightarrow \|F(x+h) - F(x) - L_x h\| \le \varepsilon \|h\|)$$

这可以缩写成

$$F(x+h) - F(x) - L_x h = o(h)$$

表达式 $L_x h$ 称为映射 F 在 x 处的强微分 (或 Frechet 微分), 线性算子 L_x 本身称为映射 F 在 x 点处的导数 (强导数). 一般用 $F'(x)$ 表示这个导数.

从定义中显见在点 x 处有强微分的映射在 x 处连续, 且如果映射 F 知其点 x 处强可微, 那么它在 x 处的强导数唯一确定, 这是因为 $\|L_1 h - L_2 h\| = o(h)$ 在 $L_1, L_2 \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 时可导出 $L_1 = L_2$. 同时根据定义显见下述性质.

命题 3.5.2

1. 若 $F(x) = y_0$ 是常值, 那么 $F'(x) \equiv 0$. 也即常值映射的强导数是零算子.
2. 若 L 是连续线性映射, 那么它的强导数就是它自己:

$$L'(x) = L(x).$$

同时链式法则与加法数乘法则也成立.

定理 3.5.11 (复合函数的导数)

设 $\mathscr{X},\mathscr{Y},\mathscr{Z}$ 是 B^* 空间, $U(x_0)$ 是点 $x_0\in\mathscr{X}$ 的一个邻域, $F:U(x_0)\to\mathscr{Y}$ 是映射; $V(y_0)$ 是点 $y_0=F(x_0)\in\mathscr{Y}$ 的一个邻域, $G:V(y_0)\to\mathscr{Z}$ 是映射. 若 F 在点 x_0 处强可微, G 在点 y_0 处强可微, 那么映射 $H=GF:U(x_0)\to\mathscr{Z}$ 在点 x_0 处强可微, 且

$$H'(x_0)=G'(y_0)F'(x_0).$$

证明

根据强可微的定义有

$$F(x_0+\xi)=F(x_0)+F'(x_0)\xi+o_1(\xi)$$

$$G(y_0+\eta)=G(y_0)+G'(y_0)\eta+o_2(\eta)$$

根据 F 的连续性[12]与 F',G' 的有界性有

$$\begin{aligned}H(x_0+\xi)=G(F(x_0+\xi))&=G(y_0+F'(x_0)+o_1(\xi))\\&=G(y_0)+G'(y_0)(F'(x_0)+o_1(\xi))+o_2(F'(x_0)+o_1(\xi))\\&=G(y_0)+G'(y_0)F'(x_0)+o_3(\xi)\end{aligned}$$

这正是 $H'(x_0)=G'(y_0)F'(x_0)$ 的定义. □

定理 3.5.12 (加法与数乘法则)

设 $F,G:\mathscr{X}\to\mathscr{Y}$ 都是连续映射. 如果 F,G 在 x_0 处强可微, 那么映射 $F+G$ 与 $aF(a\in\mathbb{K})$ 也在 x_0 处强可微, 且

$$(F+G)'(x_0)=F'(x_0)+G'(x_0),\quad (aF)'(x_0)=aF'(x_0)$$

这根据定义与加法和数乘的性质立得. □

定义 3.5.9 (弱微分 (Gateaux 微分))

设 $F:\mathscr{X}\to\mathscr{Y}$ 是映射, 所谓 F 在点 x 处 (增量为 h) 的弱微分 (Gateaux 微分) 是指极限

$$DF(x,h)=\frac{d}{dt}F(x+th)|_{t=0}=\lim_{t\to0}\frac{F(x+th)-F(x)}{t}$$

其中极限收敛是在 $\mathscr{Y}$ 的范数下的. 表达式 $DF(x,h)$ 叫做 F 在点 x 的初级变分 (第一变分). 注意弱微分 $DF(x,h)$ 关于 h 不一定是线性的, 而如果线性性成立, 即

$$DF(x,h)=F_c'(x)h$$

其中 $F_c'(x)$ 是一个有界线性算子, 那么就称这个算子为弱导数 (或 Gateaux 导数).

注 对弱导数而言, 链式法则一般不成立.

现在回过头来在弱导数的背景下讨论微分中值定理. 设 O 是 $\mathscr{X}$ 中的一个开集, 又设闭区间 $[x_0,x]\subset O$, 并设 $F:\mathscr{X}\supset O\to\mathscr{Y}$ 在闭区间 $[x_0,x]$ 的每一点有弱导数 F_c'. 令 $\Delta x=x-x_0$, 任取泛函 $\varphi\in\mathscr{Y}^*$, 考虑定义在 $[0,1]$ 上的数值函数:

$$f(t)=\varphi(F(x_0+t\Delta x))$$

知

$$\frac{f(t+\Delta t)-f(t)}{\Delta t}=\varphi(\frac{1}{\Delta t}(F(x_0+t\Delta x+\Delta t\Delta x)-F(x_0+t\Delta x)))$$

令 $\Delta t\to0$, 根据 φ 的连续性有

$$f'(t)=\lim_{\Delta t\to0}\frac{f(t+\Delta t)-f(t)}{\Delta t}=\varphi(\lim_{\Delta t\to0}\frac{1}{\Delta t}(F(x_0+t\Delta x+\Delta t\Delta x)-F(x_0+t\Delta x)))=\varphi(F_c'(x_0+t\Delta x)\Delta x)$$

[12]这是为了保证 $\|F'(x_0)+o_1(\xi)\|$ 足够小.

从而对 $[0,1]$ 上的函数 f 应用微分中值定理得到

$$f(1)-f(0)=f'(\theta),\quad 0\leq\theta\leq 1$$

代入即

$$\varphi(F(x)-F(x_0))=\varphi(F_c'(x_0+\theta\Delta x)\Delta x) \tag{3.44}$$

注意 $\varphi\in\mathscr{Y}^*$ 是任意取的, 从(3.44)式中得到

$$|\varphi(F(x)-F(x_0))|\leq\|\varphi\|\cdot\sup_{0\leq\theta\leq 1}\|F_c'(x_0+\theta\Delta x)\|\cdot\|\Delta x\| \tag{3.45}$$

现在根据推论(3.5.2)知存在非零泛函 φ 使得

$$\varphi(F(x)-F(x_0))=\|\varphi\|\cdot\|F(x)-F(x_0)\|$$

代入(3.45)式有

$$\|F(x)-F(x_0)\|\leq\sup_{0\leq\theta\leq 1}\|F_c'(x_0+\theta\Delta x)\|\cdot\|\Delta x\|,\quad \Delta x=x-x_0 \tag{3.46}$$

(3.46)式正是微分中值定理的类似公式, 将(3.46)式应用到映射

$$x\mapsto F(x)-F_c'(x_0)\Delta x$$

得到

$$\|F(x)-F(x_0)-F_c'(x_0)\Delta x\|\leq\sup_{0\leq\theta\leq 1}\|F_c'(x_0+\theta\Delta x)-F_c'(x_0)\|\cdot\|\Delta x\| \tag{3.47}$$

注意强可微和弱可微是两个不同的概念, 这在有限维空间中也是成立的, 而这是因为对数值函数 $f(x)=f(x_1,\cdots,x_n)$ 而言, 当 $n\geq 2$ 时从导数 $\frac{d}{dt}f(x+th)$ 对任意固定的 $h=(h_1,\cdots,h_n)$ 的存在性, 还不能推出 f 的可微性.

例 3.26(弱可微但不强可微的例子) 取函数

$$f(x_1,x_2)=\begin{cases}\frac{x_1^3x_2}{x_1^4+x_2^2}, & (x_1,x_2)\neq(0,0)\\ 0, & (x_1,x_2)=(0,0)\end{cases}$$

$f(x)$ 在 $(0,0)$ 处存在弱微分, 这是因为

$$\lim_{t\to 0}\frac{f(0+th)-f(0)}{t}=\lim_{t\to 0}\frac{t^4h_1^3h_2}{(t^4h_1^4+t^2h_2^2)t}=0$$

但它并非 $f(x)$ 在 $(0,0)$ 处增量的线性主部, 这是因为若令 $h_2=h_1^2$, 可得

$$\lim_{\|h\|\to 0}\frac{f(h_1,h_2)-f(0,0)}{\|h\|}=\frac{1}{2}\neq 0$$

这说明式子

$$f(h_1,h_2)-f(0,0)=f'(0,0)h+o(h)\quad (f'(0,0)=0)$$

并不成立.

但强可微可推知弱可微, 这是因为对于强可微映射 F 有:

$$F(x+th)-F(x)=F'(x)(th)+o(th)=tF'(x)h+o(th)$$

得到

$$\frac{F(x+th)-F(x)}{t}=F'(x)h+\frac{o(th)}{t}\to F'(x)h$$

下面讨论弱可微推强可微的条件.

定理 3.5.13

如果映射 F 的弱导数 $F_c'(x)$ 在点 x_0 的某一邻域 U 内存在, 且在该邻域内是 x 的一个 (算子) 函数, 这个函数在 x_0 连续, 那么在点 x_0 处强导数 $F'(x_0)$ 存在, 并且与弱导数相同.

证明

根据 $F_c'(x)$ 在 x_0 处连续知:

$$\forall\varepsilon > 0\exists\delta > 0(\|h\| < \delta \Rightarrow \|F_c'(x_0+h) - F_c'(x_0)\| < \varepsilon)$$

对 F 应用有限增量公式(3.47)知

$$\|F(x_0+h) - F(x_0) - F_c'(x_0)h\| \leq \sup_{0\leq\theta\leq 1} \|F_c'(x_0+\theta h) - F_c'(x_0)\| \cdot \|h\| < \varepsilon\|h\|$$

这恰是强导数的定义, 故强导数 $F'(x_0)$ 存在, 且它就等于弱导数 $F_c'(x_0)$. □

注 之后若无特殊说明, 默认可微指的是强可微.

例 3.27 在实 Hilbert 空间 H 中讨论泛函 $F(x) = \|x\|^2$, 知

$$\|x+h\|^2 - \|x\|^2 = 2(x,h) + \|h\|^2$$

这说明量 $2(x,h)$ 是上述表达式关于 h 的线性主部, 得到

$$F'(x) = F_c'(x) = 2x.$$

定义 3.5.10 (抽象函数与其导数)

假定变元空间 $\mathscr{X}$ 为数轴, $\mathscr{Y}$ 是 B 空间, 映射 $F: \mathscr{X} \to \mathscr{Y}$ 就称为抽象函数. 抽象函数的导数 $F'(x)$ 存在时是 $\mathscr{Y}$ 中的元素, 它代表曲线 $F(x)$ 的切向量.

定义 3.5.11 (抽象函数的积分)

设 $F: [a,b] \to \mathscr{Y}$ 是抽象函数, $[a,b]$ 有分割:

$$a = t_0 < t_1 < \cdots < t_n = b$$

则定义 F 在闭区间 $[a,b]$ 上的积分为

$$\lim_{\max\{t_{k+1}-t_k\}_{k=0}^{n-1}\to 0} \sum_{k=0}^{n-1} F(\xi_k)(t_{k+1}-t_k), \quad \xi_k \in [t_k, t_{k+1}], k = 0, 1, \cdots, n-1$$

记为

$$\int_a^b F(t)dt.$$

可以看出上述定义与 Riemann 积分极为相似, 事实上它的性质也类似于 Riemann 积分.

命题 3.5.3

1. (积分的线性性) 如果 U 是一个由空间 $\mathscr{Y}$ 到某一空间 $\mathscr{Z}$ 内的固定的线性连续映射, 那么
$$\int_a^b UF(t)dt = U\int_a^b F(t)dt$$
2. (与 *Riemann* 积分的关系) 若 $F(t)$ 形如 $f(t)y_0$, 其中 $f(t)$ 是一个数值函数, 而 $y_0 \in \mathscr{Y}$ 是一个固定的元素, 则
$$\int_a^b F(t)dt = y_0\int_a^b f(t)dt$$
3. (一个积分不等式) $\|\int_a^b F(t)dt\| \leq \int_a^b \|F(t)\|dt$.

设 $\mathscr{X},\mathscr{Y}$ 是 B^* 空间, $BC(\mathscr{X},\mathscr{Y})$ 是由 $\mathscr{X}$ 到 $\mathscr{Y}$ 的全体连续有界映射[13]所构成的线性空间. 在 $BC(\mathscr{X},\mathscr{Y})$ 中可以引入由算子范数诱导的拓扑, 进而可以取集合

$$U_{n,\varepsilon} = \{F: \sup_{\|x\|\leq n} \|F(x)\| < \varepsilon\}$$

[13]映射 $F: \mathscr{X} \to \mathscr{Y}$ 称为有界映射, 如果对于每一有界集 $Q \subset \mathscr{X}$, 集合 $F(Q)$ 在 $\mathscr{Y}$ 中均有界. 注意非线性连续映射不一定有界.

作为零映射的邻域. 设 $J = [x_0, x_0 + \Delta x]$ 是 $\mathscr{X}$ 中的一条直线段, 现在假设给定了一个由线段 J 到空间 $BC(\mathscr{X}, \mathscr{Y})$ 的连续映射, 也即设某个连续依赖于向量参量 $x \in J$ 的映射 $F(x) \in BC(\mathscr{X}, \mathscr{Y})$ 与每一点 $x \in J$ 相对应, 此时可以定义 $F(x)$ 沿线段 J 的积分, 令

$$\int_{x_0}^{x_0+\Delta x} F(x)dx = \int_0^1 F(x_0 + t\Delta x)(\Delta x)dt$$

这里对每个固定的 $t \in [0,1]$ 而言, $F(x_0 + t\Delta x)(\Delta x)$ 是 $\mathscr{Y}$ 中的元素, 它表示 $\Delta x \in \mathscr{X}$ 在映射 $F(x_0 + t\Delta x)$ 下的象. 显见 $\int_0^1 F(x_0 + t\Delta x)(\Delta x)dt \in \mathscr{Y}$.

下面来说明如何从映射的导数来还原映射.

定理 3.5.14 (广义 Newton-Leibniz 公式)

设 $F : \mathscr{X} \to \mathscr{Y}$ 在线段 $[x_0, x_0 + \Delta x]$ 上有连续依赖于 x 的强导数 $F'(x)$, 则

$$\int_{x_0}^{x_0+\Delta x} F'(x)dx = F(x_0 + \Delta x) - F(x_0)$$

♡

证明

既然 $F'(x)$ 关于 x 连续, 显见 $\int_{x_0}^{x_0+\Delta x} F'(x)dx$ 存在, 根据定义有:

$$\int_{x_0}^{x_0+\Delta x} F'(x)dx = \lim_{\Delta\to 0}\sum_{k=0}^{n-1} F'(x_0 + t_k\Delta x)(\Delta x)(t_{k+1} - t_k) = \lim_{\delta\to 0}\sum_{k=0}^{n-1} F'(x_k)(\Delta x_k)$$

其中

$$x_k = x_0 + t_k\Delta x, \quad \Delta x_k = (t_{k+1} - t_k)\Delta x, \quad \delta = \max_k(t_{k+1} - t_k)$$

但同时对于线段 $0 \le t \le 1$ 的任何分割有

$$F(x_0 + \Delta x) - F(x_0) = \sum_{k=0}^{n-1}[F(x_0 + t_{k+1}\Delta x) - F(x_0 + t_k\Delta x)] = \sum_{k=0}^{n-1}[F(x_{k+1}) - F(x_k)]$$

应用有限增量公式(3.47)得到

$$\|\sum_{k=0}^{n-1}[F(x_{k+1}) - F(x_k) - F'(x_k)\Delta x_k]\| \le \|\Delta x\| \cdot \sum_{k=0}^{n-1}(t_{k+1} - t_k)\sup_\theta\|F'(x_k + \theta_k\Delta x_k) - F'(x_k)\| \tag{3.48}$$

因为 $F'(x)$ 关于 x 连续, 进而其在闭区间 $[x_0, x_0 + \Delta x]$ 上一致连续, 因而可以说明 $\delta \to 0$ 时, (3.48)式的右端趋零, 因而

$$F(x_0 + \Delta x) - F(x_0) = \sum_{k=0}^{n-1}[F(x_{k+1}) - F(x_k)] \sim \sum_{k=0}^{n-1} F'(x_k)\Delta x_k \to \int_{x_0}^{x_0+\Delta x} F'(x)dx, \quad \delta \to 0$$

命题进而得证. □

定义 3.5.12 (二阶导数)

设 $F : \mathscr{X} \to \mathscr{Y}$ 可微, 它的导数 $F'(x)$ 对每个固定的 $x \in \mathscr{X}$ 都是 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 中的一个元素. 如果 $F'(x)$ 也可微, 就称它的导数为映射 F 的二阶导数, 记为 F''. 可见 $F''(x)$ 本质上是由 $\mathscr{X}$ 到 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 的线性算子, 也即 $F'' \in \mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$.

♣

二阶导数用双线性映射的表述更方便一些.

定义 3.5.13 (双线性映射)

称给定了一个由 $\mathscr{X}$ 到 $\mathscr{Y}$ 的双线性映射, 如果 $\mathscr{X}$ 中的每一对有序对 (x, x') 与元素 $y = B(x, x') \in \mathscr{Y}$ 建立了对应关系, 并且该关系满足:

1. (双线性性) 对任意的 $x_1, x_2, x_1', x_2' \in \mathscr{X}, \alpha, \beta \in \mathbb{R}$ 有

$$B(\alpha x_1 + \beta x_2, x_1') = \alpha B(x_1, x_1') + \beta B(x_2, x_1')$$
$$B(x_1, \alpha x_1' + \beta x_2') = \alpha B(x_1, x_1') + \beta B(x_1, x_2')$$

2. (连续性) 对所有的 $x, x' \in \mathscr{X}$, 存在正数 M 使得

$$\|B(x, x')\| \leq M\|x\| \cdot \|x'\|$$

满足条件 2. 的最小数 M 称为双线性映射 B 的范数, 记为 $\|B\|$.

空间 $\mathscr{X}$ 到空间 $\mathscr{Y}$ 的双线性映射全体也构成一个 B^* 空间, 记为 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$. 当 $\mathscr{Y}$ 完备时, $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 也完备. 现在任取 $A \in \mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$, 知其与 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 中的元素建立关系, 为此令

$$B(x, x') = (Ax)x' \tag{3.49}$$

这个对应首先显见是线性的, 现在证明它还是等距的, 且把空间 $\mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$ 映到 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 上. 这是因为若 $y = B(x, x') = (Ax)x'$, 则

$$\|y\| \leq \|Ax\| \cdot \|x'\| \leq \|A\| \cdot \|x\| \cdot \|x'\|$$

这说明

$$\|B\| \leq \|A\|$$

另一方面, 当给定双线性映射 B, 则固定 $x \in \mathscr{X}$, 映射 $x' \mapsto (Ax)x' = B(x, x')$ 就是 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性映射. 这样一来, 每个 $x \in \mathscr{X}$ 与空间 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 中的元素 Ax 对应, 显见 Ax 线性依赖于 x, 这说明双线性映射 B 确定了空间 $\mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$ 的某一元素 A. 此时根据(3.49)式有

$$\|Ax\| = \sup_{\|x'\| \leq 1} \|(Ax)x'\| = \sup_{\|x'\| \leq 1} \|B(x, x')\| \leq \|B\| \cdot \|x\|$$

这说明

$$\|A\| \leq \|B\|$$

综上得到 $\|A\| = \|B\|$. 这说明由(3.49)式确定的 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 与 $\mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$ 之间的对应是线性且等距的, 因而是一一对应. 此时 $\mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \mathscr{Y}))$ 的像正是 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$.

例 3.28 设 $\mathscr{X}, \mathscr{Y}$ 分别是维数为 m, n 的有限维欧氏空间, 则 $\mathscr{X}$ 到 $\mathscr{Y}$ 内的每一个线性映射都可以用某个 $n \times m$ 矩阵给出. 这样一来, $F : \mathscr{X} \to \mathscr{Y}$ 的导数 $F'(x)$ 就是依赖于 $x \in \mathscr{X}$ 的矩阵. 如果选定 $\mathscr{X}, \mathscr{Y}$ 的基

$$e_1, \cdots, e_m \in \mathscr{X}, \quad f_1, \cdots, f_n \in \mathscr{Y}$$

则记

$$x = x_1 e_1 + \cdots + x_m e_m, \quad y = y_1 f_1 + \cdots + y_n f_n$$

于是映射 $y = F(x)$ 可以写成

$$\begin{gathered} y_1 = F_1(x_1, \cdots, x_m) \\ \vdots \\ y_n = F_n(x_1, \cdots, x_m) \end{gathered}$$

进而

$$F'(x) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_m} \end{pmatrix}$$

二阶导数 $F''(x)$ 由 $n \times m \times m$ 个量 $a_{k,i,j} = \frac{\partial^2 y_k}{\partial x_i \partial x_j}$ 来确定. 一方面, 如果要把 $F''(x)$ 看成 $\mathscr{X}$ 到 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 的线性映

射, 则可取

$$b_{k,j} = \sum_{i=1}^{m} a_{k,i,j} x_i$$

这样的 $b_{k,j}$ 构成一个矩阵. 同时, 要把 $F''(x)$ 看成 $\mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 中的元素, 则可取

$$y_k = \sum_{i,j=1}^{m} a_{k,i,j} x_i x_j'$$

显见 $(y_1, \cdots, y_m)$ 是 $\mathscr{Y}$ 中的元素.

二阶导数的概念可以进一步 (归纳地) 推广到 n 阶导数: n 阶导数定义为 $(n-1)$ 阶导数的导数. 此时, n 阶导数是空间 $\mathscr{L}(\mathscr{X}, \mathscr{L}(\mathscr{X}, \cdots, \mathscr{L}(\mathscr{X}, \mathscr{Y})))$ 中的元素. 类似于二阶导数时的讨论, 可以把这个空间中的每个元素与 n 线性映射空间 $\mathscr{N}(\mathscr{X}^n, \mathscr{Y})$ 的元素建立关系, 此时 n 线性映射理解成 $\mathscr{X}$ 中的有序数组 $(x', x'', \cdots, x^{(n)})$ 与 $\mathscr{Y}$ 的元素之间的对应关系 $y = N(x', x'', \cdots, x^{(n)})$, 其中在其余元素固定时, N 对每个 x^i 都是线性的, 且存在 $M > 0$ 满足

$$\|N(x', x'', \cdots, x^{(n)})\| \le M\|x'\| \cdot \|x''\| \cdots \|x^{(n)}\|$$

这说明 F 的 n 阶导数也可以看成 $\mathscr{N}(\mathscr{X}^n, \mathscr{Y})$ 的元素.

定义 3.5.14 (高阶微分)

映射 F 的 (强) 微分定义为线性算子 $F'(x)$ 作用在元素 $h \in \mathscr{X}$ 得到的结果. 二阶微分定义为 $d^2F = F''(x)(h,h)$, 这是映射 $F''(x) \in \mathscr{B}(\mathscr{X}^2, \mathscr{Y})$ 作用在 $(h,h) \in \mathscr{X}^2$ 上的结果. 类似地, n 阶微分指 $d^nF = F^{(n)}(x)(h,h,\cdots,h)$, 这是 $F^{(n)}(x) \in \mathscr{N}(\mathscr{X}^n, \mathscr{Y})$ 作用在 $(h,h,\cdots,h) \in \mathscr{X}^n$ 上的结果.

♣

最终, 可以得到映射的 Taylor 公式.

定理 3.5.15 (Taylor 公式)

设 $F: \mathscr{X} \supset O \to \mathscr{Y}$, $F^{(n)}(x)$ 在 O 内存在, 且是 x 的一直连续函数, 那么有等式:

$$F(x+h) - F(x) = F'(x)h + \frac{1}{2!}F''(x)(h,h) + \cdots + \frac{1}{n!}F^{(n)}(x)(h,h,\cdots,h) + \omega(x,h) \tag{3.50}$$

其中 $\|\omega(x,h)\| = o(\|h\|^n)$.

♡

证明

用归纳法. $n = 1$ 时(3.50)式恰是强微分的定义. 现在任意固定 n, 设 $n-1$ 时(3.50)式已经成立, 即此时

$$F(x+h) - F(x) = F'(x)h + F''(x)(h,h) + \cdots + \frac{1}{(n-1)!}F^{(n-1)}(x)(h,h,\cdots,h) + \omega(x,h)$$

两边求导有[14]

$$F'(x+h) - F'(x) = F''(x)h + \frac{1}{2!}F'''(x)(h,h) + \cdots + \frac{1}{(n-1)!}F^{(n)}(x)(h,h,\cdots,h) + \omega_1(x,h)$$

3.5.3.2 隐函数定理及其应用

在笔记最开始曾经讨论过隐函数的存在定理, 本节补充的就是隐函数微分定理.

定理 3.5.16 (隐函数定理)

设 $\mathscr{X}, \mathscr{Y}, \mathscr{Z}$ 是 B 空间, U 是点 $(x_0, y_0) \in \mathscr{X} \times \mathscr{Y}$ 的一个邻域, F 是 U 到 $\mathscr{Z}$ 内的映射, 其满足:

1. F 在点 (x_0, y_0) 连续;
2. $F(x_0, y_0) = 0$;
3. $F_y'(x,y)$ 在 U 内存在, 在 (x_0, y_0) 处连续, 算子 $F_y'(x_0, y_0)$ 有有界逆算子.

[14] 在这里有巨大疑问: $F''(x)h$ 中的 h 是在哪个空间中的?

那么存在 $y=f(x)$ 使得在 (x_0,y_0) 的某一邻域内有 $F(x,f(x))=0$. 当 F'_x 在 U 内存在, 在点 (x_0,y_0) 处连续时, 可得 f 在 x_0 处可微, 且

$$f'(x_0)=-[F'_y(x_0,y_0)]^{-1}F'_x(x_0,y_0) \tag{3.51}$$

证明

只需验证(3.51)式. 注意到(3.51)式右端 (记为 Λ) 是由 $\mathscr{X}$ 到 $\mathscr{Y}$ 内的线性算子, 现在希望说明 Λ 正是 f 在 x_0 的导数, 根据定义这需要说明:

$$\forall\varepsilon>0\exists\delta>0(\|x-x_0\|<\delta\Rightarrow\|f(x)-f(x_0)-\Lambda(x-x_0)\|<\varepsilon\|x-x_0\|)$$

记 $f(x)=y$, 用 y_0 表示 $f(x_0)$, 代入 Λ 得到

$$\begin{aligned}f(x)-f(x_0)-\Lambda(x-x_0)&=y-y_0+[F'_y(x_0,y_0)]^{-1}F'_x(x_0,y_0)(x-x_0)\\&=[F'_y(x_0,y_0)]^{-1}[F'_x(x_0,y_0)(x-x_0)+F'_y(x_0,y_0)(y-y_0)]\end{aligned}$$

根据条件有 $F(x,y)=F(x_0,y_0)=0$, 故根据有限增量公式(3.47)得

$$\begin{aligned}&\|f(x)-f(x_0)-\Lambda(x-x_0)\|\\\le&\|[F'_y(x_0,y_0)]^{-1}\|\cdot\|F(x,y)-F(x_0,y_0)-F'_x(x_0,y_0)(x-x_0)-F'_y(x_0,y_0)(y-y_0)\|\\\le&\|[F'_y(x_0,y_0)]^{-1}\|\cdot\sup_{\substack{x<\xi_x<x_0\\y<\xi_y<y_0}}\|F'_x(\xi_x,\xi_y)(x-x_0)+F'_y(\xi_x,\xi_y)(y-y_0)-F'_x(x_0,y_0)(x-x_0)-F'_y(x_0,t_0)(y-y_0)\|\\\le&\|[F'_y(x_0,y_0)]^{-1}\|\cdot\sup_{\substack{x<\xi_x<x_0\\y<\xi_y<y_0}}\|F'_x(\xi_x,\xi_y)-F'_x(x_0,y_0)\|\cdot\|x-x_0\|\\&+\|[F'_y(x_0,y_0)]^{-1}\|\cdot\sup_{\substack{x<\xi_x<x_0\\y<\xi_y<y_0}}\|F'_y(\xi_x,\xi_y)-F'_y(x_0,y_0)\|\cdot\|y-y_0\|\\\le&\eta(\|x-x_0\|+\|y-y_0\|)\end{aligned}$$

代入 $y=f(x)$ 得

$$\|f(x)-f(x_0)-\Lambda(x-x_0)\|\le\eta(\|x-x_0\|+\|f(x)-f(x_0)\|)\le\eta(\|x-x_0\|+\|\Lambda(x-x_0)\|+\|f(x)-f(x_0)-\Lambda(x-x_0)\|) \tag{3.52}$$

根据 F'_x,F'_y 的连续性, 只要 δ 足够小, η 就是足够小的. 当 η 充分小时, 从(3.52)式可解得

$$\|f(x)-f(x_0)-\Lambda(x-x_0)\|\le\frac{\eta}{1-\eta}(1+\|\Lambda\|)\|x-x_0\|$$

最后选取足够小的 η 使得 $\frac{\eta}{1-\eta}(1+\|\Lambda\|)\le\varepsilon$ 即得定理. □

隐函数定理的一个应用在笔记开头已经介绍过, 即微分方程解对初值的依赖性. 现在讨论隐函数定理的另一个应用.

设 $F(x),x=(x_1,x_2)$ 是平面上的一个可微函数. 方程 $F(x)=0$ 确定了平面上的某一曲线 C. 设 x_0 是该曲线的一个点. 曲线 C 在给定点的切线有两种表述: 其一是形如 x_0+th 的向量的全体, 其中 h 是垂直于向量 $F'(x_0)$(也即函数 F 在 x_0 处的梯度) 的向量; 其二是点 x_0+th 的全体, 其中这些点到 C 的距离是 t 高于一阶的无穷小. Liusternik 定理指出切线的这两个定义对于任意的 Banach 空间中的流形也成立.

定义 3.5.15 (正则映射, 切线性流形)

设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, $F:\mathscr{X}\to\mathscr{Y}$. 设 M_0 是 $\mathscr{X}$ 中满足 $F(x)=0$ 的全体 x 构成的集合, $x_0\in M_0$. 假定 F 在点 x_0 的某一邻域 U 内连续可微, 称 F 在点 x_0 处是正则的, 如果线性算子 $F'(x_0)$ 将空间 $\mathscr{X}$ 映到全空间 $\mathscr{Y}$. 用 T_0 表示满足条件 $F'(x_0)h=0$ 的元素 $h\in\mathscr{X}$ 的全体, 也即 $T_0=\ker F'(x_0)$. 显见 T_0 是 $\mathscr{X}$ 的线性子空间. 用 T_{x_0} 表示将 T_0 平移一个向量 x_0, 也即流形 x_0+T_0. 称 T_{x_0} 为在点 x_0 处与集合 M_0 相切的线性流形.

定理 3.5.17 (Liusternik)

设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, 若 $F:\mathscr{X}\to\mathscr{Y}$ 在 x_0 处正则, $F(x_0)=0$, 则元素 x_0+h 属于 F 在 x_0 处的切流形 T_{x_0} 的充要条件是: 元素 x_0+th 到集合 $M_0:=\{x\in\mathscr{X}:F(x)=0\}$ 的距离是 t 的高于一阶的无穷小量.

简洁起见, 只对" 可展" 情形的 Liusternik 定理给出证明, 此时 $\mathscr{X}$ 可表为子空间 $T_0=\ker F'(x_0)$ 与某一空间 T_ξ 的直和[15]:

$$\mathscr{X}=T_0\oplus T_\xi$$

此时 Liusternik 定理可被重新表述成下述形式.

定理 3.5.18

如果

$$\mathscr{X}=T_0\oplus T_\xi$$

而映射 $F:\mathscr{X}\to\mathscr{Y}$ 在 x_0 处正则, $F(x_0)=0$, 那么存在一个从 $M_0:=\{x\in\mathscr{X}:F(x)=0\}$ 中点 x_0 的邻域到 T_{x_0} 中同一点的邻域的同胚映射, 使得彼此对应的点之间的距离是比它们到点 x_0 的距离更高阶的无穷小量.

证明

用 A 表示算子 $F'(x_0)$ 在子空间 T_ξ 上的限制, 即

$$\forall\xi\in T_\xi(A\xi=F'(x_0)\xi)$$

下面证明 A 将子空间 T_ξ 映射到全空间 $\mathscr{Y}$ 上. 既然根据条件 $\mathscr{X}=T_0\oplus T_\xi$, 知任取 $x\in\mathscr{X}$, 都可以将其写成

$$x=h+\xi,\quad h\in T_0=\ker F'(x_0),\xi\in T_\xi$$

这说明

$$F'(x_0)x=F'(x_0)(h+\xi)=F'(x_0)\xi=A\xi$$

但根据条件, $F'(x_0)$ 应该将 $\mathscr{X}$ 映到 $\mathscr{Y}$ 上, 这说明当 ξ 跑遍 T_ξ 时, $A\xi$ 跑遍 $\mathscr{Y}$. 其次, 映射 $A:T_\xi\to\mathscr{Y}$ 是一一对应, 这是因为若 $A\xi_1=A\xi_2$, 则 $F'(x_0)(\xi_2-\xi_1)=0$, 从而 $\xi_2-\xi_1\in T_0$, 得到 $\xi_2-\xi_1\in T_0\cap T_\xi=\{0\}$, 也即 $\xi_1=\xi_2$. 从而 A 可逆, 由 Banach 定理(3.4.3)得到 A^{-1} 是有界线性算子.

现在固定 x_0, 若将 $\mathscr{X}$ 的元素 x 表成

$$x=x_0+h+\xi,\quad h\in T_0,\xi\in T_\xi$$

则可以将定义流形 M 的方程 $F(x)=0$ 改写成

$$\Phi(h,\xi)=F(x_0+h+\xi)=0$$

这个函数在 $(0,0)$ 处对应 ξ 的增量 $\Delta\xi$ 的偏微分形如

$$\Phi'_\xi(0,0)\Delta\xi=F'(x_0)\Delta\xi=A\Delta\xi$$

根据前述讨论, $A=\Phi'_\xi(0,0)$ 有逆算子, 进而根据隐函数定理, 方程 $\Phi(h,\xi)=0$ 在点 $(0,0)$ 的某邻域内与形如

$$\xi=\psi(h)$$

的方程等价, 其中 $\psi(h)$ 是一个满足 $\psi(0)=0$ 的可微映射.

上面证明的是每个充分靠近 x_0 的点 $x\in M_0$ 都具有形式

$$x=x_0+h+\psi(h),\quad h\in T_0,\psi(h)\in T_\xi$$

这便得到映射

$$x_0+h\mapsto x_0+h+\psi(h)$$

[15] 与 Hilbert 空间不同, 在 Banach 空间中, 并非每个子空间都有直补子空间, 也即正交分解定理并不一定成立. 可以证明如果空间 $\mathscr{X}$ 中每一线性子空间都有直补子空间, 那么 $\mathscr{X}$ 是 Hilbert 空间.

它把 T_{x_0} 中点 x_0 的某一邻域映射到 M_0 中同一点的邻域, 显见这个映射是一一对应且连续的. 下面证明彼此对应的点之间的距离, 也即量 $\|\psi(h)\|$ 是 $\|h\|$ 的高阶无穷小量.

对等式 $\Phi(h,\psi(h))=0$ 微分有

$$\Phi'_h(0,0)h+\Phi'_\xi(0,0)\psi'(0)h=\Phi'_h(0,0)h+A\psi'(0)h=0$$

这说明

$$\psi'(0)h=-A^{-1}\Phi'_h(0,0)h=-A^{-1}F'(x_0)h=-A^{-1}0=0$$

注意先前所提到的 $\psi(0)=0$, 可得

$$\psi(h)=\psi(0)+\psi'(0)h+o(\|h\|)=o(\|h\|)$$

此即欲证. □

现在转过头来, 在已经介绍了微分学的背景下再讨论一下极值问题.

定义 3.5.16 (极值)

设 F 是定义在 Banach 空间 $\mathscr{X}$ 上的某个实泛函, 称 F 在点 x_0 达到极小值, 如果对于所有充分接近 x_0 的 x, 都有不等式

$$F(x)-F(x_0)\geq 0$$

类似可以定义泛函的极大值. 如果在给定点 x_0 泛函 F 达到了极小值或极大值, 就称在这一点泛函 F 有极值.

♣

对 n 元函数而言, 已经学过它在某点取极值的必要条件了: 如果函数 f 在点 $x_0=(x_1^0,x_2^0,\cdots,x_n^0)$ 处可微且在该点有极值, 则在该点 $df=0$, 即

$$\frac{\partial f}{\partial x_1}=\frac{\partial f}{\partial x_2}=\cdots=\frac{\partial f}{\partial x_n}=0$$

这个条件可以推广到任意赋范空间的泛函上去.

定理 3.5.19

可微泛函 F 在点 x_0 达到极值的必要条件是: 对于所有的 h, 它的微分在该点等于 0:

$$F'(x_0)h\equiv 0$$

换言之, 泛函 F 在点 x_0 有极值的必要条件是

$$F'(x_0)=0$$

这里的 0 表示零映射.

♡

证明

根据可微的定义知

$$F(x_0+h)-F(x_0)=F'(x_0)h+o(h)$$

如果对某个 h, 有 $F'(x_0)h\neq 0$, 则取 λ 充分小, 使得 $F'(x_0)(\lambda h)+o(\lambda h)$ 与 $F'(x_0)(\lambda h)$ 的符号相同. 但既然 $F'(x_0)$ 是线性泛函, 知 $F'(x_0)(\lambda h)=\lambda F'(x_0)h$. 这说明若 $F'(x_0)\neq 0$, 则 $F'(x_0)(\lambda h)+o(\lambda h)$ 在 h 充分小时既可取正又可取负, 也即 $F(x_0+h)-F(x_0)$ 没有固定符号, 与极值相悖. □

例 3.29 设

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

其中 f 连续可微. 知 $F: C[a,b] \to \mathbb{R}$ 是可微的, 这是因为

$$F(x+h) - F(x) = \int_a^b [f(t, x(t) + h(t)) - f(t, x(t))]dt = \int_a^b f'_x(t, x(t))h(t)dt + o(h)$$

进而

$$dF = \int_a^b f'_x(t, x(t))h(t)dt$$

要想讨论 F 的极值, 则需令 dF 对所有的 $h \in C[a,b]$ 全为零, 这说明 $f'_x(t, x(t)) = 0$. 从而泛函 F 的极值需要在曲线 $f'_x(t, x) = 0$ 上寻找.

例 3.30 在 $C[a,b]$ 上考虑泛函

$$F(x) = \int_a^b d\xi_1 \int_a^b K(\xi_1, \xi_2)x(\xi_1)x(\xi_2)d\xi_2$$

其中 $K(\xi_1, \xi_2)$ 满足 $K(\xi_1, \xi_2) = K(\xi_2, \xi_1)$ 且连续. 不难计算这个泛函的微分为

$$dF = 2\int_a^b d\xi_1 \int_a^b K(\xi_1, \xi_2)x(\xi_1)h(\xi_2)d\xi_2$$

在讨论 F 的极值时, 希望对于每个 $h \in C[a,b]$ 而言, $dF = 0$. 这说明对所有的 $\xi_2 \in [a,b]$ 有

$$\int_a^b K(\xi_1, \xi_2)x(\xi_1)d\xi_1 = 0$$

这个方程的其中一个解是 $x \equiv 0$. 至于其他的问题需要具体研究 $K(\xi_1, \xi_2)$.

例 3.31 考虑定义在 $C^1[a,b]$ 上的泛函

$$F(x) = \int_a^b f(t, x(t), x'(t))dt$$

其中 $x'(t) = \frac{dx(t)}{dt}$, 而 $f(t, x, x')$ 是对其变元二次可微的函数. 下面求解 F 的微分.

$$F(x+h) - F(x) = \int_a^b [f(t, x+h, x'+h') - f(t, x, x')]dt = \int_a^b (f'_x h + f'_{x'} h)dt + o(\|h\|)$$

其中 $\|h\|$ 是 $h \in C^1[a,b]$ 的范数. 这说明泛函 F 有极值的必要条件为

$$dF = \int_a^b (f'_x h + f'_{x'} h')dt = 0$$

因为积分形式下想要求解 x 是困难的, 下面将其变换一下形式. 分部积分有:

$$\int_a^b f'_{x'} h' dt = \int_a^b f'_{x'} dh = f'_{x'} h|_a^b - \int_a^b h \frac{d}{dt} f'_{x'} dt$$

得到

$$dF = \int_a^b (f'_x - \frac{d}{dt} f'_{x'})hdt + f'_{x'} h|_a^b = 0 \tag{3.53}$$

这个等式对全体 h 都应满足, 从而对所有使得 $h(a) = h(b) = 0$ 的 h 都有

$$\int_a^b (f'_x - \frac{d}{dt} f'_{x'})hdt = 0$$

进而得到

$$f'_x - \frac{d}{dt} f'_{x'} = 0 \tag{3.54}$$

从而(3.53)式简化为

$$f'_{x'} h|_a^b = 0$$

回忆 F 是定义在 $C^1[a,b]$ 上的, 可以选取 h 使得 $h(a) = 0, h(b) \neq 0$, 代入得

$$f'_{x'}|_{t=b} = 0 \tag{3.55}$$

若令 $h(b)=0, h(a)\neq 0$ 即得

$$f'_{x'}|_{t=a}=0 \tag{3.56}$$

这说明使得 F 达到极值的函数 x, 应该满足方程(3.54), 边值条件满足(3.55),(3.56)两式.

现在考虑求 n 元函数的极值, 设 $f(x_1,\cdots,x_n)$ 在点 $(x_1^0,\cdots,x_n^0)$ 满足条件 $df=0$, 则极值的判别需要下述条件.

1. 若 $f(x_1,\cdots,x_n)$ 在 $(x_1^0,\cdots,x_n^0)$ 有极小值, 那么在该点 $d^2f\geq 0$, 类似地, 若 f 在点 $(x_1^0,\cdots,x_n^0)$ 有极大值, 那么在该点 $d^2f\leq 0$.
2. 如果在点 $(x_1^0,\cdots,x_n^0)$ 处有
$$df=0,\quad d^2f=\sum_{i,k=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_k}dx_i dx_k>0$$
且 dx_i 不全为零, 则在该点处 $f(x)$ 有极小值. 类似地, 当 $d^2f<0$ 时 $f(x)$ 有极大值.

简而言之, 二阶微分的非负性是有极小值的必要条件, 而它的正定性是有极小值的充分条件. 现在把这些命题推广到 Banach 空间上.

定理 3.5.20

设 F 是一个给定在 B 空间 $\mathscr{X}$ 中的实泛函, 且在点 x_0 的某一邻域内有连续的二阶导数. 如果这个泛函在 x_0 处达到极小值, 那么 $d^2F(x_0)\geq 0$, 也即对全体 h 都有 $F''(x_0)(h,h)\geq 0$. ♡

证明

由 Taylor 公式知

$$F(x_0+h)-F(x_0)=F'(x_0)h+\frac{1}{2}F''(x_0)(h,h)+p(\|h\|^2)$$

如果在 x_0 处泛函 F 有极小值, 那么 $F'(x_0)=0$, 得到

$$F(x_0+h)-F(x_0)=\frac{1}{2}F''(x_0)(h,h)+o(\|h\|^2)$$

当 h 的范数足够小时, $F(x_0+h)-F(x_0)$ 的符号仅由 $F''(x_0)(h,h)$ 确定. 进而根据极小值的定义, 要想对任意足够小的 h 有 $F(x_0+h)\geq F(x_0)$, 需要 $F''(x_0)(h,h)\geq 0$, 也即 $d^2F(x_0)\geq 0$. □

对极大值而言有类似的定理.

上述定理是关于有限变量函数相应定理的推广. 注意上面提到的条件 $F''(x_0)(h,h)>0$ 是 n 元函数有极小值的充分条件, 但并非定义在无穷维 Banach 空间上的泛函达到极小值的充分条件.

例 3.32 在 Hilbert 空间 l^2 中给定泛函

$$F(x)=\sum_{n=1}^{\infty}\frac{x_n^2}{n^3}-\sum_{n=1}^{\infty}x_n^4$$

这个泛函在 $x=0$ 的一阶微分为 0, 二阶微分为 $2\sum\limits_{n=1}^{\infty}\frac{h_n^2}{n^3}$, 进而其为正定泛函. 但它在 $x=0$ 处没有极小值, 这是因为 $F(0)=0, F(0,\cdots,0,\frac{1}{n},0,\cdots)=\frac{1}{n^5}-\frac{1}{n^4}<0$, 这说明在 0 的任意邻域内都有满足 $F(x)<0$ 的点.

定义 3.5.17 (强正)

二次泛函 B 称为是强正的, 如果存在常数 c, 使得对于所有的 x 都有 $B(x,x)\geq c\|x\|^2$. ♣

定理 3.5.21

如果定义在 B 空间 $\mathscr{X}$ 的泛函 F 满足条件:

1. $dF(x_0)=0$,
2. $d^2F(x_0)$ 是强正二次泛函,

那么 F 在点 x_0 有极小值.

证明

根据强正性的定义, 设 $F''(x_0)(h,h) \geq c\|h\|^2$, 选取 ε 足够小, 使得对于 $\|h\| < \varepsilon$ 的 $o(\|h\|^2)$ 满足 $|o(\|h\|^2)| < \frac{c}{4}\|h\|^2$. 于是当 $\|h\| < \varepsilon$ 时有

$$F(x_0+h) - F(x_0) = \frac{1}{2}F''(x_0)(h,h) + o(\|h\|^2) > \frac{c}{4}\|h\|^2 > 0$$

□

在有限维空间中, 二次型的强正性与正定性是等价的, 而在无穷维的情形, 强正性是比正定性更强的条件.

3.5.4 补充: Hahn-Banach 定理的更多应用

本节选自 [**ST4**],[**PGC**], [**HB**] 等, 旨在进一步介绍 Hahn-Banach 定理的应用.

3.5.4.1 有界线性算子共轭的保范性

在闭图像定理一节最开始的讨论中, 我们曾经讨论过 B^* 空间到 B 空间的连续线性算子定义域总可以延拓成闭集, 于是在讨论连续线性算子时可以默认它是 B 空间之间的算子 (即把 B 空间映成 B 空间). 现定义 B 空间上算子的共轭.

定义 3.5.18 (共轭算子)

设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间, $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 则其诱导了一个共轭算子 $T^* : \mathscr{Y}^* \to \mathscr{X}^*$, 其满足对任意的 $Y \in \mathscr{Y}^*$ 而言有

$$T^*(Y)(x) = Y(Tx), \quad \forall x \in \mathscr{X} \tag{3.57}$$

定理 3.5.22 (ST4)

在定义3.5.18的条件下, 由(3.57)式定义的算子 $T^* : \mathscr{Y}^* \to \mathscr{X}^*$ 是有界线性算子, 且 $\|T\| = \|T^*\|$.

证明

任取 $x \in \mathscr{X}$, 知

$$|T^*(Y)(x)| = |Y(Tx)| \leq \|Y\|\|Tx\|_{\mathscr{Y}} \leq \|Y\|\|T\|\|x\|_{\mathscr{X}}$$

于是

$$\|T^*(Y)\| \leq \|Y\|\|T\| \Rightarrow \|T^*\| \leq \|T\|$$

再证明 $\|T^*\| \geq \|T\|$. 根据算子范数的上确界定义知对任意 $\varepsilon > 0$, 均存在 $x \in \mathscr{X}$ 使得 $\|x\|_{\mathscr{X}} = 1$, 且

$$\|Tx\|_{\mathscr{Y}} \geq \|T\| - \varepsilon$$

现在因为 $Tx \in \mathscr{Y}$, 根据 Hahn-Banach 定理的推论3.5.2知存在 $Y \in \mathscr{Y}^*$ 使得 $Y(Tx) = \|Tx\|_{\mathscr{Y}}$ 且 $\|Y\| = 1$. 但前面已经有 $\|Tx\|_{\mathscr{Y}} \geq \|T\| - \varepsilon$, 这便说明

$$Y(Tx) = T^*(Y)(x) \geq \|T\| - \varepsilon$$

因为 $\|x\|_{\mathscr{X}} = 1$, 故

$$\|T^*(Y)\| \geq \|T\| - \varepsilon$$

最后因为 $\|Y\| = 1$, 故 $\|T^*\| \geq \|T\| - \varepsilon$, 由 ε 的任意性即得 $\|T^*\| \geq \|T\|$, 命题即证. □

3.5.4.2 证明子空间稠密的一种方法

[HB] 中给出了 Hahn-Banach 第二几何形式的一个推论:

推论 3.5.6 (HB)

设 E 是 B^* 空间, $F \subset E$ 是线性子空间, 且满足 $\overline{F} \neq E$, 则存在 $f \in E^*, f \neq 0$ 使得

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

证明

设 $x_0 \in E \backslash \overline{F}$, 知 $A = \overline{F}, B = \{x_0\}$ 是不相交的非空凸集, 且 A 是闭集, B 是紧集, 故存在 $f \in E^*, f \neq 0$ 使得闭超平面 H_f^α 严格分离 $\overline{F}$ 和 $\{x_0\}$, 也就是说

$$f(x) < \alpha < f(x_0), \quad \forall x \in F$$

又因为任取 $\theta \neq x \in F$ 有 $f(\lambda x) = \lambda f(x)(\lambda \in \mathbb{R})$, 故为使上式对任意 $x \in F$ 均成立, 只能要求

$$f(x) = 0, \quad \forall x \in F$$

$\square$

根据这个推论, 如果想要证明线性子空间 $F \subset E$ 稠密, 就可以去寻找 $f \in E^*$ 满足 $\forall x \in F(f(x) = 0)$, 再证明 f 在 E 上恒为零.

3.5.4.3 线性算子的延拓

[ST4] 指出, 只要线性算子是定义在稠密子空间上的, 那么这个线性算子就可以延拓.

命题 3.5.4 (线性算子延拓定理 ST4)

设 $\mathscr{X}, \mathscr{Y}$ 是 Banach 空间, $S \subset \mathscr{X}$ 是 $\mathscr{X}$ 的一个稠密线性子空间. 若 $T_0 : S \to \mathscr{Y}$ 满足 $\forall f \in S(\|T_0 f\|_{\mathscr{Y}} \leq M\|f\|_{\mathscr{X}})$, 其中 M 是常数, 则 T_0 在 $\mathscr{X}$ 上有唯一的延拓 T, 且 $\forall f \in \mathscr{X}(\|Tf\|_{\mathscr{Y}} \leq M\|f\|_{\mathscr{X}})$.

证明

任取 $f \in \mathscr{X}$, 根据 S 在 $\mathscr{X}$ 中稠密的定义知存在 $\{f_n\}_{n\in\mathbb{N}} \subset S$ 使得 $\|f_n - f\|_{\mathscr{X}} \to 0(n \to \infty)$. 根据 T_0 的条件知

$$\|T_0(f_n - f_m)\|_{\mathscr{Y}} \leq M\|f_n - f_m\|_{\mathscr{X}}, \quad \forall m, n \in \mathbb{N}$$

故 $\{T_0 f_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{Y}$ 中的收敛列, 由 $\mathscr{Y}$ 的完备性知其具有极限, 定义该极限为 y, 并构造算子

$$T : \mathscr{X} \to \mathscr{Y}, \quad \lim_{n\to\infty} f_n = f \mapsto Tf := \lim_{n\to\infty} T_0 f_n$$

显见 Tf 并不依赖于 $\{f_n\}_{n\in\mathbb{N}}$ 的选取, 故 $T : \mathscr{X} \to \mathscr{Y}$ 即为欲求. $\square$

[HB] 中提到了上述定理的更弱形式:

定理 3.5.23

设 $\mathscr{X}, \mathscr{Y}$ 是 Banach 空间, $G \subset \mathscr{X}$ 是闭子向量空间, $T \in \mathscr{L}(G, \mathscr{Y})$, 则

(a) 若 $\dim \mathscr{Y} < \infty$;

(b) 若 G 有拓扑余子空间, 如 $\dim G < \infty$ 或 $\operatorname{codim} G < \infty$;

(c) 若 $\mathscr{X}$ 是 Hilbert 空间,

三者有其一成立, 那么存在 $T' \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 是 T 在 $\mathscr{X}$ 上的延拓.

3.5.5 补充: 拓扑余子空间

本节选自 [**HB**], 旨在介绍拓扑余子空间这一之后会在算子谱论中碰到的内容. 事实上, 只要是讨论算子定义域和值域的方向, 都是绕不过拓扑余子空间的. 根据开映射定理, Banach 空间中的闭子空间有下述分解性质:

定理 3.5.24

设 E 是 Banach 空间, 若 G 和 L 是 E 的两个闭子空间, 满足 $G+L=\{g+l : g\in G, l\in L\}$ 是 E 中的闭集, 则存在常数 $C\geq 0$ 使得对任意 $z=x+y\in G+L(x\in G, y\in L)$, 都有 $\|x\|\leq C\|z\|, \|y\|\leq C\|z\|$.

证明

考虑乘积空间 $G\times L$ 赋范

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

而空间 $G+L$ 赋 E 的函数, 定义映射

$$T: G\times L\to G+L, \quad (x,y)\mapsto x+y$$

则 T 是连续线性的满射. 根据 Banach 逆算子定理知 $T^{-1}: z=x+y\mapsto (x,y)$ 也是连续线性满射, 因而根据 T^{-1} 的有界性知存在 $C>0$ 使得

$$\forall z=x+y\in G+L(\|(x,y)\|=\|x\|+\|y\|=\|T^{-1}z\|\leq C\|z\|)$$

命题即证. □

推论 3.5.7

在定理3.5.24的条件下, 存在常数 C 使得

$$\rho(x, G\cap L)\leq C(\rho(x,G)+\rho(x,L)), \quad \forall x\in E \tag{3.58}$$

其中 $\rho(x,G)$ 表示 x 到空间 G 的距离.

证明

设 $x\in E, \varepsilon>0$, 根据 $\rho(x,G), \rho(x,L)$ 的下确界定义知存在 $a\in G, b\in L$ 使得

$$\|x-a\|\leq \rho(x,G)+\varepsilon, \quad \|x-b\|\leq\rho(x,L)+\varepsilon$$

对 $z=a-b$ 应用定理3.5.24知存在 $a'\in G, b'\in L$ 使得

$$a-b=z=a'-b', \quad \|a'\|\leq C\|a-b\|, \quad \|b'\|\leq C\|a-b\|$$

故 $a-a'\in G\cap L$(因为显见 $a-a'\in G$, 而 $a-a'=b-b'\in L$), 且

$$\begin{aligned}\rho(x,G\cap L) &\leq \|x-(a-a')\|\leq\|x-a\|+\|a'\| \\ &\leq \|x-a\|+C\|a-b\|\leq\|x-a\|+C(\|x-a\|+\|x-b\|) \\ &\leq \rho(x,G)+\varepsilon+C(\rho(x,G)+\varepsilon+\rho(x,L)+\varepsilon) \\ &= (1+C)(\rho(x,G)+\rho(x,L))+(1+2C)\varepsilon\end{aligned}$$

令 $\varepsilon\to 0$ 即得欲证. □

注 推论3.5.7的逆命题也是正确的: 若 G, L 是 E 中满足(3.58)式的两个闭子空间, 那么 $G+L$ 是闭集.

定义 3.5.19 (拓扑余子空间)

设 $G\subset E$ 是 Banach 空间 E 的闭子空间, 称 E 中的子空间 L 为 G 的拓扑余子空间, 如果:

(i) L 是闭集;

(ii) $G\cap L=\{\theta\}$, 且 $G+L=E$.

在定义3.5.19下, 任意 $z \in E$ 可以被唯一地写成 $z = x + y$, 其中 $x \in G, y \in L$. 根据定理3.5.24知投影算子 $z \mapsto x$ 与 $z \mapsto y$ 都是连续线性算子.

例 3.33(拓扑余子空间的例子)

(1) E 的任意有限维子空间 G 都具有拓扑余子空间. 事实上, 取 G 的基底为 $\{e_1, \cdots, e_n\}$, 则任意 $x \in G$ 总能写成 $x = \sum_{i=1}^{n} x_i e_i$. 现在定义 $\varphi_i(x) = x_i$, 把每个 φ_i 按照 Hahn-Banach 定理延拓成 E 上的连续线性泛函 $\widetilde{\varphi}_i$, 则可以证明 $L = \bigcap_{i=1}^{n} (\widetilde{\varphi}_i)^{-1}(0)$(也就是 $\bigcap_{i=1}^{n} \{x \in E : \widetilde{\varphi}_i(x) = 0\}$) 是 G 的拓扑余子空间.

(2) 任意余维有限 (即 $\operatorname{codim} G = \dim(E/G) < \infty$) 的闭子空间 G 具有拓扑余子空间, 这种情况下取 G 在代数意义下的补空间即可. 这种情况还有一个典例: 设空间 $N \subset E^*$ 满足 $\dim N = p$, 则

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

是 E 的闭子空间, 且其具有余维数 p. 事实上, 取 N 的基底 $\{f_1, \cdots, f_p\}$, 根据练习3.35的结果知存在 $e_1, \cdots, e_p \in E$ 使得

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

可以验证 $\{e_i\}_{1 \le i \le p}$ 是线性无关的, 且 $\operatorname{span}\{e_i\}_{1 \le i \le p}$ 恰是 G 的拓扑余子空间.

(3) 在 Hilbert 空间中, 任意闭子空间均有拓扑余子空间.

注 即使在自反空间中, 也可以构造某些闭子空间, 使得它们不具有任何拓扑余子空间. **[HB]** 中指明 Lindenstrauss-Tzafriri 有结果: 任何不同构于 Hilbert 空间的 Banach 空间中, 存在不具有拓扑余子空间的闭子空间.

设 $T \in \mathscr{L}(E, F)$ 是满射, 由开映射定理知存在 $C > 0$ 使得

$$\forall f \in F \exists x \in E (Tx = f \wedge \|x\| \le C\|f\|)$$

现在问是否存在 $S \in \mathscr{L}(F, E)$ 使得 $T \circ S = id_F$, 这样的 S 称为 T 的右逆算子.

定理 3.5.25

设 $T \in \mathscr{L}(E, F)$ 是满射, 则以下性质等价:

(i) T 有右逆算子;

(ii) $N(T) = T^{-1}(0)$ 在 E 中有拓扑余子空间.

证明

(i)⇒(ii) 设 $S \in \mathscr{L}(F, E)$ 为 T 的右逆算子, 则易证 $R(S) = S(F)$ 是 $N(T)$ 在 E 中的拓扑余子空间.

(ii)⇒(i) 设 L 是 $N(T)$ 的拓扑余子空间, 用 P 表示从 E 到 L 的投影, 显见 $P \in \mathscr{L}(E, L)$. 给定 $f \in F$, 设 x 是满足 $Tx = f$ 的任意一个解, 令 $Sf = Px$, 注意 S 的定义实际上与 x 的选取无关. 可以证明 $S \in \mathscr{L}(F, E)$, 且 $T \circ S = id_F$.□

类似地, 称 S 是 $T \in \mathscr{L}(E, F)$ 的左逆算子, 如果 $S \in \mathscr{L}(F, E)$ 满足 $S \circ T = id_E$.

定理 3.5.26

设 $T \in \mathscr{L}(E, F)$ 是单射, 则以下性质等价:

(i) T 有左逆算子;

(ii) $R(T) = T(E)$ 是闭的, 且在 F 中有拓扑余子空间.

证明

(i)⇒(ii) 设 S 是 T 的左逆算子, 易证 $R(T)$ 是闭的, 且 $N(S)$ 是 $R(T)$ 的拓扑余子空间.

(ii)⇒(i) 设 P 是从 F 到 $R(T)$ 的连续投影, 设 $f \in F$. 因为 $Pf \in R(T)$, 故存在唯一的 $x \in E$ 使得 $Pf = Tx$. 定义 $Sf = x$, 则 $S \circ T = id_E$, 且由 Banach 逆算子定理知 S 是连续算子. □

3.6 共轭空间, 弱收敛, 自反空间

3.6.1 知识梳理

3.6.1.1 共轭空间的表示及应用

定义 3.6.1 (共轭空间)

设 $\mathscr{X}$ 是一个 B^* 空间, $\mathscr{X}$ 上的所有连续线性泛函全体 $\mathscr{X}^*$, 按范数

$$\|f\| = \sup_{\|x\|=1} |f(x)|$$

构成一个 B 空间, 称为 $\mathscr{X}$ 的共轭空间. ♣

例 3.34(L^p 的共轭空间 $(1 \le p < \infty)$) 设 q 是 p 的共轭指数, 下面证明

$$L^p[0,1]^* = L^q[0,1] \tag{3.59}$$

任取 $g \in L^q[0,1]$, 根据 Hölder 不等式知

$$|\int_0^1 f(x)g(x)d\mu| \le (\int_0^1 |f(x)|^p d\mu)^{\frac{1}{p}} (\int_0^1 |g(x)|^q d\mu)^{\frac{1}{q}}$$

其中 μ 是 $[0,1]$ 上的 Lebesgue 测度. 知

$$F_g(f) := \int_0^1 f(x)g(x)d\mu \quad (\forall f \in L^p[0,1])$$

定义了 $L^p[0,1]$ 上的一个连续线性泛函, 且

$$\|F_g\|_{L^p[0,1]^*} = \sup_{\|f\|_{L^p[0,1]} \le 1} |\int_0^1 f(x)g(x)d\mu| \le \sup_{\|f\|_{L^p[0,1]} \le 1} \|f\|_{L^p[0,1]} \cdot \|g\|_{L^q[0,1]} \le \|g\|_{L^q[0,1]}$$

这说明映射 $g \mapsto F_g$ 将 $L^q[0,1]$ 连续地嵌入 $L^p[0,1]^*$. 下面说明映射 $g \mapsto F_g$ 是等距且一一的, 也即对给定的 $F \in L^p[0,1]^*$, 希望找一个 $g \in L^q[0,1]$, 使得

$$F(f) = \int_0^1 f(x)g(x)d\mu \quad (\forall f \in L^p[0,1]) \tag{3.60}$$

且

$$\|g\|_{L^q[0,1]} = \|F\|$$

对任意的可测集 $E \subset [0,1]$, 令

$$\nu(E) := F(\chi_E)$$

其中 χ_E 是 E 的特征函数:

$$\chi_E(x) = \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases}$$

下面验证 ν 是一个可数可加测度. 根据 $F \in L^p[0,1]^*$ 知 ν 是有限可加测度, 现在设 $\{E_n\} \subset [0,1]$ 满足

$$E_1 \supset E_2 \supset \cdots \supset E_n \supset \cdots$$

及

$$\bigcap_{n=1}^{\infty} E_n = \varnothing$$

则

$$\nu(E_n) = F(\chi_{E_n}) \le \|F\| \cdot \|\chi_{E_n}\|_{L^p[0,1]} = \|F\|(\int_0^1 |\chi_{E_n}|^p d\mu)^{\frac{1}{p}} = \|F\|\mu(E_n)^{\frac{1}{p}} \to 0, \quad n \to \infty$$

此外可以证明 $\mu(E) = 0 \Rightarrow \nu(E) = 0$, 也即 ν 关于 μ 是绝对连续的.

现在应用 Radon-Nikodym 定理(3.3.5)知, 存在可测函数 g, 使得对任意的可测集 E 有

$$\nu(E) = \int_E g d\mu$$

得到

$$F(\chi_E) = \nu(E) = \int_E g d\mu = \int_0^1 \chi_E(x) g(x) d\mu$$

进而对一切简单函数 f 都有

$$F(f) = \int_0^1 f(x) g(x) d\mu$$

下面证明

$$\|g\|_{L^q[0,1]} \le \|F\| \tag{3.61}$$

并从(3.61)式推出(3.60)式.

当 $1 < p < \infty$, 任取 $t > 0$, 记

$$E_t := \{x \in [0,1] : |g(x)| \le t\}$$

令 $f = \chi_{E_t} |g|^{q-2} g$, 有

$$\int_{E_t} |g|^q d\mu = \int_0^1 f \cdot g d\mu = F(f) \le \|F\| \cdot \|f\|_{L^p[0,1]} = \|F\| (\int_{E_t} |g|^q d\mu)^{\frac{1}{p}}$$

即

$$(\int_{E_t} |g|^q d\mu)^{\frac{1}{q}} \le \|F\|$$

令 $t \to \infty$ 即得(3.61)式.

当 $p = 1$, 此时 $q = \infty$, 对任意的 $\varepsilon > 0$, 令

$$A := \{x \in [0,1] : |g(x)| > \|F\| + \varepsilon\}$$

对任意的 $t > 0$, 令

$$E_t := \{x \in [0,1] : |g(x)| \le t\}$$

记

$$f = \chi_{E_t \cap A} \operatorname{sgn} g$$

得到

$$\|f\|_{L^1[0,1]} = \mu(E_t \cap A)$$

且

$$\mu(E_t \cap A)(\|F\| + \varepsilon) \le \int_{A \cap E_t} |g| d\mu = \int_0^1 f \cdot g d\mu \le \|F\| \mu(E_t \cap A)$$

令 $t \to \infty$ 得到

$$\mu(A)(\|F\| + \varepsilon) \le \|F\| \mu(A)$$

这说明只能有 $\mu(A) = 0$, 得到

$$\|g\|_{L^\infty[0,1]} \le \|F\|$$

这便是 $q = \infty$ 情况的(3.61)式.

最后从(3.61)式推导(3.60)式. 因为简单函数集在 $L^p[0,1]$ 中稠密, 故对任意的 $f \in L^p[0,1]$, 存在简单函数列 $f_n \to f$, 进而

$$F(f) = \int_0^1 \lim_{n \to \infty} f_n(x) g(x) d\mu$$

由 Hölder 不等式与(3.61)知

$$|\int_0^1 \lim_{n\to\infty} f_n(x)g(x)d\mu| \le \|f_n\|_{L^p[0,1]}\cdot\|g\|_{L^q[0,1]} \le \|F\|\cdot\|f_n\|_{L^p[0,1]} < \infty$$

从而由控制收敛定理得

$$F(f) = \lim_{n\to\infty}\int_0^1 f_n(x)g(x)d\mu = \lim_{n\to\infty} F(f_n)$$

又因为

$$|\int_0^1 [f(x)-f_n(x)]g(x)d\mu| \le (\int_0^1 |f(x)-f_n(x)|^p d\mu)^{\frac{1}{p}}(\int_0^1 |g(x)|^q d\mu)^{\frac{1}{q}} \le \|F\|\cdot\|f-f_n\|_{L^p[0,1]} \to 0, \quad n\to\infty$$

这说明

$$F(f) = \lim_{n\to\infty}\int_0^1 f_n(x)g(x)d\mu = \int_0^1 f(x)g(x)d\mu.$$

(3.60)式得证. □

注

1. (3.59)式可以推广到一般的可数可加, $\sigma-$ 有限的测度空间. 设 $(\Omega,\mathscr{B},\mu)$ 是这样的测度空间, 则

$$L^p(\Omega,\mathscr{B},\mu)^* = L^q(\Omega,\mathscr{B},\mu) \quad (1\le p<\infty).$$

2. ⋆**[FL]** 指出, 即使 $L^p(\Omega)$ 是复线性空间, 也可以像实线性空间的情形定义泛函, 亦即 $g\mapsto Fg=\int_\Omega f(x)g(x)dx(f\in L^q(\Omega))$, 而不必特别利用复共轭:$g\mapsto Fg=\int_\Omega f(x)\overline{g(x)}dx$. 在复线性空间的情况下按后者定义, 只是因为在把 $L^2(\Omega)$ 当成 Hilbert 空间时, 这里得到的结论可以和 Riesz 表示定理3.3.1保持一致.

例 3.35(补充: L^∞ 的共轭空间YOS) 设 $f\in L^\infty(\Omega,\mathscr{B},m)$ 定义为 $\forall B\in\mathscr{B}(f(\chi_B)=\psi(B))$, 其中 χ_B 是 B 的示性函数. 有:

(a) $B_1\cap B_2=\varnothing \Rightarrow \psi(B_1\cup B_2)=\psi(B_1)+\psi(B_2)$, 亦即 ψ 是有限可加的集函数. 这是因为

$$f(\chi_{B_1\cup B_2}) = f(\chi_{B_1}+\chi_{B_2}) = f(\chi_{[}B_1]) + f(\chi_{B_2}) = \psi(B_1)+\psi(B_2)$$

(b) $\psi(B)$ 的实部 $\psi_1(B)$ 与虚部 $\psi_2(B)$ 均有界, 亦即 $\sup\limits_B|\psi_i(B)|<\infty(i=1,2)$. 这是因为

$$\max\{|\psi_1(B)|,|\psi_2(B)|\} \le |\psi(B)| = |f(\chi_B)| \le \|f\|\cdot\|\chi_B\| < \infty$$

(c) ψ 是 $m-$ 绝对连续的[16], 亦即 $m(B)=0\Rightarrow\psi(B)=0$, 这同样是因为 $|\psi(B)|\le\|f\|\cdot\|\chi_B\|$.

现在用上述语言来刻画 $L^\infty(\Omega,\mathscr{B},m)$ 中的任意元素 x. 根据复平面中的闭圆盘 $\{z:|z|\le\|x\|\}$ 的紧性, 任取 $\varepsilon>0$, 总存在有限个不交的 Borel 集 $A_1,\cdots,A_n$ 使得 $\{z:|z|\le\|x\|\}\subset\bigcup\limits_{k=1}^n A_k$, 且 $\sup\limits_{1\le i\le n}\operatorname{diam}(A_i)<\varepsilon$. 若设 $B_i=\{s\in\Omega:x(s)\in A_i\}$, 则在 $A_i(i=1,\cdots,n)$ 中任取点 $\alpha_i(i=1,\cdots,n)$, 知对任意的 $s\in\Omega$, 根据 $\{A_i\}_{i=1}^n$ 的不交性与 $\{z:|z|\le\|x\|\}\subset\bigcup\limits_{i=1}^n A_i$ 知存在唯一 A_{i_0} 使得 $s\in B_{i_0}$, 于是

$$|x(s)-\sum_{i=1}^n\alpha_i\chi_{B_i}(s)| = |x(s)-\alpha_{i_0}| \le \operatorname{diam}(A_{i_0}) < \varepsilon$$

于是

$$|f(x)-\sum_{i=1}^n\alpha_i\psi(B_i)| \le \|f\|\cdot\varepsilon$$

从而若令 $\varepsilon\to0$, 亦即 $n\to\infty$, 可得

$$f(x) = \lim_{n\to\infty}\sum_{i=1}^n\alpha_i\psi(B_i) \tag{3.62}$$

[16] **ZMQ** 设 $f(x)$ 是 $[a,b]$ 上的实值函数. 若对任给的 $\varepsilon>0$, 存在 $\delta>0$, 使得当 $[a,b]$ 中任意有限个互不相交的开区间 $(x_i,y_i)(i=1,2,\cdots,n)$ 满足 $\sum\limits_{i=1}^n(y_i-x_i)<\delta$ 时, 有 $\sum\limits_{k=1}^n|f(y_i)-f(x_i)|<\varepsilon$, 则称 $f(x)$ 是 $[a,b]$ 上的绝对连续函数.

可以说明 f 的值与 $\{z : |z| \le \|x\|\}$ 的分解 $\bigcup_{i=1}^{n} A_i$ 和 α 的选取均无关. (3.62)右式称为 $x(s)$ 关于有限可加测度 ψ 的 Radon 积分, 记为

$$f(x) = \int_S x(s)\psi(ds), \quad x \in L^\infty(\Omega, \mathscr{B}, m) \tag{3.63}$$

可得

$$\|f\| = \sup_{\operatorname{ess\,sup}|x(s)| \le 1} \left| \int_S x(s)\psi(ds) \right| \tag{3.64}$$

反之, 也可以说明任意满足 (a)-(c) 的集函数 ψ 都可以通过 Radon 积分(3.63)定义 $f \in L^\infty(\Omega, \mathscr{B}, m)^*$, 且(3.64)式成立. 这说明 $L^\infty(\Omega, \mathscr{B}, m)^*$ 是全体满足 (a)-(c) 的集函数构成的空间, 其范数由(3.64)式定义.

例 3.36(连续函数空间的共轭空间) 现在讨论 $C[0,1]$ 的共轭空间. 设

$$BV[0,1] := \{g : [0,1] \to \mathbb{C} : g(0) = 0, g(t) = g(t+0)(\forall t \in (0,1)), \bigvee_0^1(g) < \infty\}$$

其中

$$\bigvee_0^1(g) = \sup \sum_{j=1}^{n-1} |g(t_{j+1}) - g(t_j)|$$

这里上确界对所有 $[0,1]$ 的分割

$$\Delta : 0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$$

来取. 在 $BV[0,1]$ 赋范数

$$\|g\|_v = \bigvee_0^1(g), \quad \forall g \in BV[0,1]$$

可以证明 $BV[0,1]$ 是 B 空间.

回顾对任意的 $\varphi \in C[0,1], g \in BV[0,1]$, Stieltjes 积分的定义:

$$\int_0^1 \varphi(t)dg(t) = \lim_{\|\Delta\| \to 0} \sum_{j=1}^{n-1} \varphi(t_j^*)[g(t_{j+1}) - g(t_j)]$$

其中 Δ 是 [0,1] 的分割, 而

$$\|\Delta\| := \max_{1 \le j \le n} |t_j - t_{j-1}|$$

从定义可知任意固定 $g \in BV[0,1]$, Stieltjes 积分都对应 $C[0,1]$ 上的一个连续线性泛函:

$$\varphi \mapsto \langle f, \varphi \rangle = \int_0^1 \varphi(t)dg(t)$$

其满足

$$\|f\| \le \int_0^1 |dg(t)| = \bigvee_0^1(g) = \|g\|_v$$

现在希望证明对任意的 $f \in C[0,1]^*$, 都存在唯一的 $g \in BV[0,1]$, 使得

$$\langle f, \varphi \rangle = \int_0^1 \varphi(t)dg(t), \quad \forall \varphi \in C[0,1] \tag{3.65}$$

且

$$\|g\|_v \le \|f\| \tag{3.66}$$

现在的关键在于如何根据 f 确定 g. 如果允许 φ 取成特征函数 $\chi_E(t)(E \subset [0,1])$, 则有

$$g(s) = \int_0^1 \chi_{(0,s]}(t)dg(t)$$

故现将 $C[0,1]$ 看成 $L^\infty[0,1]$ 的一个闭子空间, 应用 Hahn-Banach 定理(3.5.4)知对给定的 $f \in C[0,1]^*$, 存在 $\tilde{f} \in$

$L^\infty[0,1]^*$, 使得

$$\langle \tilde{f}, \chi_{\{0\}}\rangle = 0, \quad \langle \tilde{f}, \varphi\rangle = \langle f, \varphi\rangle (\forall \varphi \in C[0,1])$$

且 $\|\tilde{f}\| = \|f\|$. 因为 $\chi_{(0,s]} \in L^\infty[0,1]$, 故可令

$$g(s) = \langle \tilde{f}, \chi_{(0,s]}\rangle (0 < s \leq 1), \quad g(0) = 0 \tag{3.67}$$

下面证明由(3.67)式定义的 $g(x)$ 在 $BV[0,1]$ 中, 且满足(3.65)式与(3.66)式.

对 $[0,1]$ 的任一分割

$$\Delta : 0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$$

记

$$\omega_k := g(t_{k+1}) - g(t_k), \lambda_k := \begin{cases} \frac{\overline{\omega_k}}{|\omega_k|}, & \omega_k \neq 0 \\ 0, & \text{其他} \end{cases} \quad (k = 0, 1, 2, \cdots, n-1)$$

$$h_\Delta(t) := \sum_{k=0}^{n-1} \lambda_k \chi_{(t_k, t_{k+1}]}(t)$$

代入 $g(s) = \langle \tilde{f}, \chi_{(0,s]}\rangle$ 有

$$\sum_{k=0}^{n-1} |\omega_k| = \sum_{k=0}^{n-1} \lambda_k \omega_k = \langle \tilde{f}, h_\Delta\rangle \leq \|\tilde{f}\| \cdot \|h_\Delta\|_{L^\infty[0,1]} \leq \|\tilde{f}\| = \|f\| < \infty$$

故 $g \in BV[0,1]$, 且满足(3.66)式.

要证明(3.65)式成立, 对任意的 $\varphi \in C[0,1], \varepsilon > 0$, 取分割 Δ 使得

$$|\varphi(t) - \varphi(t')| < \frac{\varepsilon}{2\|\tilde{f}\|} \quad (\forall t, t' \in [t_j, t_{j+1}], j = 0, 1, \cdots, n-1)$$

以及

$$\left|\int_0^1 \varphi(t) dg(t) - \sum_{j=0}^{n-1} \varphi(t_j)(g(t_{j+1}) - g(t_j))\right| < \frac{\varepsilon}{2}$$

令

$$\varphi_\Delta := \sum_{j=0}^{n-1} \varphi(t_j) \chi_{(t_j, t_{j+1}]} + \varphi(0) \chi_{\{0\}}$$

代入 $g(s) = \langle \tilde{f}, \chi_{(0,s]}\rangle$ 得到

$$\begin{aligned} \left|\langle f, \varphi\rangle - \int_0^1 \varphi(t) dg(t)\right| &\leq |\langle f, \varphi\rangle - \langle \tilde{f}, \varphi_\Delta\rangle| + \left|\langle \tilde{f}, \varphi_\Delta\rangle - \int_0^1 \varphi(t) dg(t)\right| \\ &\leq \|\tilde{f}\| \cdot \|\varphi - \varphi_\Delta\|_{L^\infty[0,1]} + \left|\sum_{j=0}^{n-1} \varphi(t_j)(g(t_{j+1}) - g(t_j)) - \int_0^1 \varphi(t) dg(t)\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

由 ε 的任意性即得(3.65)式.

根据 Jordan 分解定理, 有界变差函数是单增函数的差, 也即 $g = g_1 - g_2$. 因为 g_1, g_2 的右极限总存在, 可以在不改变 $\int_0^1 \varphi(t) dg_i(t), i = 1, 2$ 值的情况下改变函数值使其右连续, 从而可设 g 是右连续的. 总而言之, $g \mapsto \int_0^1 \varphi(t) dg(t)$ 是 $BV[0,1] \to C[0,1]^*$ 的一个等距同构, 也即

$$C[0,1]^* = BV[0,1].$$

□

注 可用测度或更一般的可数可加集函数代替有界变差函数得到下述定理.

定理 3.6.1 (Riesz 表示定理 (连续函数空间))

若 M 是一个 Hausdorff 紧空间, 则 $\forall f \in C(M)^*$, 有唯一的复值 Baire 测度, 即可数可加的集函数 μ, 满足 $|\mu|(M) < \infty$, 且

$$\langle f, \varphi \rangle = \int_M \varphi(m) d\mu \quad (\forall \varphi \in C(M)).$$

从 Hahn-Banach 定理(3.5.4)和 Riesz 表示定理(3.6.1)中可推知下述复变逼近论中的定理.

定理 3.6.2 (Runge)

设 K 是复平面 $\mathbb{C}$ 上的一个紧子集, 记 $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$. 又设 E 是 $\mathbb{C}_\infty \backslash K$ 中的一个子集, 它与 $\mathbb{C}_\infty \backslash K$ 的每一个连通部分都相交. 若 f 是 K 的一个邻域内的任意解析函数, 则必有有理函数列 f_n, 其极点都在 E 内, 使得 f_n 在 K 上一致收敛到 f.

证明

引入 K 上的连续函数空间 $C(K)$, 并记 $R(K,E)$ 为极点在 E 内的有理函数集在 $C(K)$ 中的闭包. 显见 $R(K,E)$ 是 $C(K)$ 的一个闭线性子空间. 由推论(3.5.3)知, 要证明 $f \in R(K,E)$, 只需证明对任意的 $F \in C(K)^*$ 有:

$$\forall g \in R(K,E)(F(g) = 0) \Rightarrow F(f) = 0$$

但根据 Riesz 表示定理(3.6.1), $C(K)^*$ 是由 K 上的可数可加的复值测度组成的, 故只需证: 对任意的 $\mu \in M(K)$

$$\forall g \in R(K,E)(\int_K g d\mu = 0) \Rightarrow \int_K f d\mu = 0 \tag{3.68}$$

现在暂时引入下述引理.

引理 3.6.1

对 $\forall \mu \in M(K)$, 若设

$$\hat{\mu}(w) = \int_K \frac{d\mu(z)}{w - z}$$

那么对 $\forall R > 0, \hat{\mu} \in L^1(B_R)$(其中 B_R 是中心在原点的半径为 R 的圆), 且 $\hat{\mu}$ 在 $\mathbb{C}_\infty \backslash K$ 上解析, 还满足 $\hat{\mu}(\infty) = 0$.

在引理成立的条件下, 可得

$$(\frac{d}{dw})^n \hat{\mu}(w_0) = n! \int_K (z - w_0)^{-n+1} d\mu(z) \quad (\forall w_0 \in \mathbb{C} \backslash K) \tag{3.69}$$

在 ∞ 点附近它有展开

$$\hat{\mu}(w) = -\frac{1}{w} \sum_{n=0}^{\infty} \int_K (\frac{z}{w})^n d\mu(z) = -\sum_{n=0}^{\infty} \frac{a_n}{w^{n+1}}, \quad a_n = \int_K z^n d\mu(z) \tag{3.70}$$

进一步, 如果 $\mu \in M(K)$ 满足

$$\int_K g(z) d\mu(z) = 0 \tag{3.71}$$

其中 g 是极点在 E 内的有理函数, 则有[17]

$$\hat{\mu}(w) = 0, \quad \forall w \in \mathbb{C}_\infty \backslash K \tag{3.72}$$

这是因为对任意的 $w_0 \in E$, 设 $\Omega(w_0)$ 是 $\mathbb{C}_\infty \backslash K$ 中含 w_0 的部分, 根据 E 最初的构造知[18]

$$\mathbb{C}_\infty \backslash K = \bigcup_{w_0 \in E} \Omega(w_0)$$

[17]这个推断可以形象理解成: 当对任意极点在 E 中有理函数 g 而言, $\int_K g(z) d\mu(z) = 0$ 时, $\frac{1}{w-z}$ 在刨去可去奇点的情况下在 K 内是解析的, 因而积分为 0.

[18]这个式子没有看懂

如果 $w_0 \neq \infty$, 根据(3.69)式与(3.70)式可知 $\hat{\mu}$ 在 w_0 的各阶导数均为 0, 故 $\hat{\mu}$ 在 $\Omega(w_0)$ 内恒为 0; 如果 $w_0 = \infty$, 则由(3.70)式与(3.71)式知 $\hat{\mu}$ 在 $\Omega(w_0)$ 内恒为 0, 因而(3.72)式成立.

现在考察在 K 的某邻域 G 内解析的任意函数 f, 根据 Cauchy 积分公式知存在含于 $G\backslash K$ 内的折线 $\gamma_1, \cdots, \gamma_n$ 满足

$$f(z) = \sum_{k=1}^{n} \frac{1}{2\pi i} \int_{\gamma_k} \frac{f(w)}{w-z} dw$$

进而由 Fubini 定理知

$$\int_K f(z) d\mu(z) = \sum_{k=1}^{n} \frac{1}{2\pi i} \int_K \int_{\gamma_k} \frac{f(w)}{w-z} dw d\mu(z) = \sum_{k=1}^{n} \frac{1}{2\pi i} \int_{\gamma_k} f(w)\hat{\mu}(w) dw = 0$$

从而(3.68)式成立, 命题得证.

最后补充引理(3.6.1)的证明, 首先证明 $\hat{\mu} \in L^1(B_R)$. 这是因为根据定义有

$$|\hat{\mu}(w)| \leq \int_K \frac{d|\mu|(z)}{|w-z|}$$

故

$$\int_{B_R} |\hat{\mu}(w)| dx dy \leq \int_{B_R} \int_K \frac{d|\mu|(z)}{|w-z|} dx dy = \int_K \int_{B_R} \frac{dx dy}{|w-z|} d|\mu|(z) \leq \int_K \int_{B(x,\rho)} \frac{dx dy}{|w-z|} d|\mu|(z) \leq 2\pi\rho|\mu|(K) < \infty$$

其中 $\rho > R + \max\limits_{z\in K}|z|$.

再证明 $\hat{\mu}$ 在 $\mathbb{C}\backslash K$ 内解析, 这是因为 K 是紧的, 故可在积分号下求导.

最后说明 $\hat{\mu}$ 在 $\{\infty\}$ 解析, 且 $\hat{\mu}(\infty) = 0$. 这是因为 K 是紧集, 故 $|w| \to \infty$ 时, 在积分号下取极限即得 $\hat{\mu}(w) = 0$, 进而 ∞ 是可去奇点. □

课堂笔记 (♡ 更多共轭空间的例子)

高志强老师在课堂上给出了更多共轭空间的例子, 整理如下:

(i) 设 C_0 是以 0 为极限的数列全体, 其范数为

$$\|\cdot\| : C_0 \ni \{\xi_k\}_{k\in\mathbb{N}} \mapsto \sup_{k\geq 1} |\xi_k|$$

则 $(C_0)^* = l^1$, 亦即对任意 $f \in (C_0)^*$, 存在唯一 $\eta = \{\eta_k\}_{k\in\mathbb{N}} \in l^1$ 使得

$$f(\{x_k\}_{k\in\mathbb{N}}) = \sum_{k=1}^{\infty} x_k \eta_k, \quad \forall \{x_k\}_{k\in\mathbb{N}} \in C_0$$

这是因为首先对任意 $x = \{x_k\}_{k\in\mathbb{N}} \in C_0$ 总存在表示 $x = \sum\limits_{k=1}^{\infty} x_k e_k$, 这是因为对余项有

$$\|\sum_{k=N}^{\infty} x_k e_k\| = \sup_{k\geq N} \{x_k\} \to 0, \quad N \to \infty$$

于是根据线性性有 $f(x) = \sum\limits_{k=1}^{\infty} x_k f(e_k)$, 取 $\eta = \{f(e_k)\}_{k\in\mathbb{N}}$, 若 $\eta \notin l^1$, 则可取 $x^n = \sum\limits_{k=1}^{n} \operatorname{sgn}(\eta_1) e_k$, 知 $\|x^n\| = 1 (n \in \mathbb{N})$, 而 $f(x^n) \to \infty (n \to \infty)$, 这与 $f \in (C_0)^*$ 矛盾! 于是 $\eta \in l^1$. 反之, 任取 l^1 中的序列 $\eta = \{\eta_k\}_{k\in\mathbb{N}}$, 总可以构造

$$f(x) = \sum_{k=1}^{\infty} x_k \eta_k, \quad x = \{x_k\}_{k\in\mathbb{N}} \in C_0$$

可以验证 $f \in (C_0)^*$, 这便有 $(C_0)^* = l^1$.

(ii) 设 C 为收敛数列的全体, 其范数为

$$\|\cdot\| : C \ni \{\xi_k\}_{k\in\mathbb{N}} \mapsto \sup_{k\geq 1} |\xi_k|$$

则 $C^*=l^1$, 亦即对任意 $f\in C^*$, 存在唯一 $\eta=\{\eta_k\}_{k\in\mathbb{N}}\in l^1$ 使得

$$f(\{x_k\}_{k\in\mathbb{N}})=\sum_{k=1}^{\infty}(x_k-x_0)\eta_k+x_0\eta_0,\quad \forall\{x_k\}_{k\in\mathbb{N}}\in C,\lim_{k\to\infty}x_k=x_0,e_0=(1,\cdots,1,\cdots)$$

这是因为对任意 $x=\{x_k\}_{k\in\mathbb{N}}\in C$, 设 $\lim\limits_{k\to\infty}x_k=x_0$, 则存在表示 $x=\sum\limits_{k=1}^{\infty}(x_k-x_0)e_k+x_0e_0$, 其中 $e_0=(1,\cdots,1,\cdots)$. 进而令 $\eta_i=f(e_i)(i=0,1,2,\cdots)$, 知

$$f(\{x_k\}_{k\in\mathbb{N}})=\sum_{k=1}^{\infty}(x_k-x_0)f(e_k)+x_0f(e_0)=\sum_{k=1}^{\infty}x_k\eta_k+x_0(-\sum_{k=1}^{\infty}\eta_k+\eta_0)$$

若 $\eta\notin l^1$, 则取 $x^n=\sum\limits_{k=1}^{n}(\operatorname{sgn}(\eta_k)+x_0)e_k$, 知 $\|x^n\|=x_0+1(n\in\mathbb{N})$, 而 $f(x^n)\to\infty(n\to\infty)$ 矛盾! 故 $\eta\in l^1$.
反之, 任取 l^1 中的序列 $\eta=\{\eta_k\}_{k\in\mathbb{N}}$, 总可以构造

$$f(x)=\sum_{k=1}^{\infty}(x_k-x_0)\eta_k+x_0(-\sum_{k=1}^{\infty}\eta_k+\eta_0),\quad x=\{x_k\}_{k\in\mathbb{N}}\in C$$

可以验证 $f\in C^*$, 这便有 $C^*=l^1$.

(iii) $(\mathbb{K}^n,\|\cdot\|_{l^p})^*=(\mathbb{K}^n,\|\cdot\|_{l^q})$, 这由 Hölder 不等式即证.

共轭空间提供了一个有效的判定空间是否可分的方法:

补充定理 3.6.1 (♡Banach)

对于 B^* 空间 $\mathscr{X}$, 若 $\mathscr{X}^*$ 可分, 则 $\mathscr{X}$ 可分.

证明

取 $\mathscr{X}^*$ 的单位球面 $\mathbb{S}_1^*:=\{f\in\mathscr{X}^*:\|f\|=1\}$, 下面说明 S_1^* 可分. 因为 $\mathscr{X}^*$ 可分, 故存在 $\{f_n\}_{n\in\mathbb{N}}\subset\mathscr{X}^*$ 使得

$$\forall f\in\mathbb{S}_1^*\exists\{n_k\}_{k\in\mathbb{N}}\subset\mathbb{N}(\lim_{k\to\infty}f_{n_k}=f)$$

令 $g_n:=\frac{f_n}{\|f_n\|}$ 可得

$$\begin{aligned}\|f-g_{n_k}\|&\le\|f-f_{n_k}\|+\|f_{n_k}-g_{n_k}\|=\|f-f_{n_k}\|+\|f_{n_k}-\frac{f_{n_k}}{\|f_{n_k}\|}\|\\&=\|f-f_{n_k}\|+\|f_{n_k}\|\cdot|1-\frac{1}{\|f-n_k\|}|=\|f-f_{n_k}\|+|\|f_{n_k}\|-1|\\&=\|f-f_{n_k}\|+|\|f_{n_k}\|-\|f\||\le\|f-f_{n_k}\|+\|f-f_{n_k}\|\to0,\quad k\to\infty\end{aligned}$$

故

$$\forall f\in\mathbb{S}_1^*\exists\{n_k\}_{k\in\mathbb{N}}\subset\mathbb{N}(\lim_{k\to\infty}g_{n_k}=f)$$

这说明 $\{g_n\}_{n\in\mathbb{N}}$ 是 $\mathbb{S}_1^*$ 的可数稠密子集, 因而 $\mathbb{S}_1^*$ 可分.

因为 $\forall n\in\mathbb{N}(\|g_n\|=\sup\limits_{\|x\|=1}|g_n(x)|=1)$, 故根据上确界的定义知, 对任意 g_n 均对应存在 $x_n\in\mathscr{X}$ 使得 $\|x_n\|=1$, 且 $|g_n(x_n)|\ge1-\varepsilon$, 其中 $\varepsilon>0$ 是一个选定的常数. 现记 $\mathscr{X}_0=\overline{\operatorname{span}\{x_n\}_{n\in\mathbb{N}}}$, 下面说明 $\mathscr{X}_0=\mathscr{X}$. 用反证法, 若 $\mathscr{X}_0\neq\mathscr{X}$, 则因为 $\mathscr{X}_0$ 是 $\mathscr{X}$ 的一个闭线性子空间, 由 Hahn-Banach 定理的推论知存在 $f_0\in\mathscr{X}^*$, 使得 $\|f_0\|=1$, 且 $\forall x\in\mathscr{X}_0(f_0(x)=0)$. 知 $f_0\in\mathbb{S}_1^*$, 有:

$$\|g_n-f_0\|=\sup_{\|x\|=1}|g_n(x)-f(x)|\ge|g_n(x_n)-f_0(x_n)|=|g_n(x_n)|\ge1-\varepsilon,\quad\forall n\in\mathbb{N}$$

这与 $\{g_n\}_{n\in\mathbb{N}}$ 在 $\mathbb{S}_1^*$ 中稠密矛盾! 故 $\mathscr{X}_0=\mathscr{X}$, 命题得证. □

因为 B^* 空间 $\mathscr{X}$ 的共轭空间 $\mathscr{X}^*$ 是 B 空间, 当然也可以考虑 $\mathscr{X}^*$ 的共轭空间:

定义 3.6.2 (第二共轭空间, 自然映射)

B^* 空间 $\mathscr{X}$ 有共轭空间 $\mathscr{X}^*$, $\mathscr{X}^*$ 的共轭空间记作 $\mathscr{X}^{**}$, 称为 $\mathscr{X}$ 的第二共轭空间. 任意固定 $x \in \mathscr{X}$, 可以定义

$$X(f) = \langle f, x \rangle$$

也即 $\mathscr{X}^*$ 中的元素 f 在 x 上的取值. 称映射 $T : x \mapsto X$ 为自然映射[a].

[a] 这里的逻辑在于: 现给定 $\mathscr{X}$ 中的一个元素 x, 那么 $\mathscr{X}^*$ 中的所有元素 f 在 x 上都有取值, 即 $\langle f, x \rangle$. 这就给出了一个从 $\mathscr{X}^*$ 到 $\mathbb{R}$ 的泛函 $X(f)$, 它吃进一个 $f \in \mathscr{X}^*$, 给出的是 f 在前面固定的 x 处的取值. 可以知道 x 不同, X 的取值大多情况下也是不同的, 而 $T : x \mapsto X$ 刻画的正是这个关系.

♣

现在来研究自然映射的性质. 首先知

$$|X(f)| = |\langle f, x \rangle| \le \|f\| \cdot \|x\| \Rightarrow \|X\| \le \|x\|$$

这说明

$$\|Tx\| = \|X\| \le \|x\| \Rightarrow \|T\| \le 1$$

也即自然映射是连续映射, T 是 $\mathscr{X}$ 到 $\mathscr{X}^{**}$ 的连续嵌入.

自然映射还是线性映射, 这是因为若 $\alpha, \beta \in \mathbb{C}, x, y \in \mathscr{X}$, 记 $X = Tx, Y = Ty$, 有

$$\begin{aligned} T(\alpha x + \beta y)(f) &= f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha X(f) + \beta Y(f) \\ &= (\alpha X + \beta Y)(f) = (\alpha Tx + \beta Ty)(f) \end{aligned}$$

这说明 T 是线性同构.

最后说明自然映射还是等距的, 这只需说明

$$\|X\| \ge \|x\| \tag{3.73}$$

根据 Hahn-Banach 定理的推论(3.5.2)知, 对每个固定的 $x_0 \in \mathscr{X}$, 都存在 $f_{x_0} \in \mathscr{X}^*$ 使得

$$f_{x_0}(x_0) = \|x_0\|, \quad \|f_{x_0}\| = 1$$

故

$$\|x_0\| = f_{x_0}(x_0) = X(f_{x_0}) \le \|X\| \cdot \|f_{x_0}\| = \|X\|$$

进而(3.73)式成立[19], 从而

$$\|X\| = \|x\|$$

又因为 $Tx = X$, 这说明 $\|Tx\| = \|x\|$, 也即 T 是等距映射. 这导出下述定理.

定理 3.6.3

B^* 空间 $\mathscr{X}$ 与它的第二共轭空间 $\mathscr{X}^{**}$ 的一个子空间等距同构.

♡

推论 3.6.1 (⋆)

B 空间 $\mathscr{X}$ 与 $\mathscr{X}^{**}$ 的一个闭子空间等距同构.

♡

证明

记 $\mathscr{A} := \{Tx : x \in \mathscr{X}\}$, 其中 T 是定义3.6.2中的自然映射, 下面证明若 $\mathscr{X}$ 是 Banach 空间, 则 $\mathscr{A}$ 闭包含于 $\mathscr{X}^{**}$. 设 $\{Tx_n\}_{n\in\mathbb{N}} \subset \mathscr{A}$, 且 $\{Tx_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}^{**}$ 中的基本列. 因为 $\mathscr{X}^{**}$ 首先是完备的, 故存在 $y \in \mathscr{X}^{**}$, 使得 $\|Tx_n - y\|_{\mathscr{X}^{**}} \to 0 (n \to \infty)$. 又因为定理3.6.3说明 T 是等距同构, 故

$$\|x_n - x_m\|_{\mathscr{X}} = \|Tx_n - Tx_m\|_{\mathscr{X}^{**}}$$

[19] 特别标注 f_{x_0} 正是为了说明: 任意取 x_0, 都会存在这么一个 f_{x_0} 使得式子成立, 而最后式子与 f 无关了, 就说明对任意的 $x \in \mathscr{X}$ 式子都是成立的.

从而 $\{x_n\}_{n\in\mathbb{N}}$ 也是 $\mathscr{X}$ 中的基本列, 因而由 $\mathscr{X}$ 的完备性知存在 $x\in\mathscr{X}$ 使得 $\|x_n-x\|_{\mathscr{X}}\to 0(n\to\infty)$. 由此知:

$$\|y-Tx\|_{\mathscr{X}^{**}}\leq\|y-Tx_n\|_{\mathscr{X}^{**}}+\|Tx_n-Tx\|_{\mathscr{X}^{**}}\overset{(i)}{\leq}\|y-Tx_n\|_{\mathscr{X}^{**}}+\|x_n-x\|_{\mathscr{X}}\to 0,\quad n\to\infty$$

其中 (i) 是出于(3.73)式. 故 $y=Tx$, 因而 $y\in\mathscr{A}$, 从而 $\mathscr{A}$ 是 $\mathscr{X}^{**}$ 中的闭集. □

注 有时不区别 x 与 X, 就记成 $\mathscr{X}\subset\mathscr{X}^{**}$.

定义 3.6.3 (自反)

如果 $\mathscr{X}$ 到 $\mathscr{X}^{**}$ 的自然映射 T 是满射的, 则称 $\mathscr{X}$ 是自反的, 记作 $\mathscr{X}=\mathscr{X}^{**}$.

注 特别注意自反的定义指明必须是自然映射为满射才行, 事实上存在这样的空间 $\mathscr{X}$, 使得 $\mathscr{X}$ 与 $\mathscr{X}^{**}$ 等距同构, 但 $\mathscr{X}$ 并不是自反空间. 下面给出 **[WL]** 介绍的 James 空间这一例子.

设 J 为满足下述条件的实数列 $x=\{\xi_n\}_{n\in\mathbb{N}}$ 所组成的实线性空间: $x=\{\xi_n\}_{n\in\mathbb{N}}\in J$ 当且仅当 $\lim\limits_{n\to\infty}\xi_n=0$, 且

$$\|x\|=\sup\Big(\sum_{i=1}^{n}(\xi_{k_{2i-1}}-\xi_{k_{2i}})^2+(\xi_{k_{2n+1}})^2\Big)^{\frac{1}{2}}<\infty \tag{3.74}$$

其中 sup 是对一切 $n\in\mathbb{N}$ 及所有可能的有限多个正整数 $k_1,\cdots,k_{2n+1}(k_1<k_2<\cdots<k_{2n+1})$ 取的, 由(3.74)式规定的 $\|\cdot\|$ 为 J 上的范数, 下面说明 $(J,\|\cdot\|)$ 是 Banach 空间.

首先说明 $\|\cdot\|$ 确为范数, 进一步只需验证三角不等式成立. 设 $x=\{\xi_n\},y=\{\eta_n\}\in J$, 由离散的 Minkowski 不等式知

$$\begin{aligned}&\Big(\sum_{i=1}^{n}(\xi_{k_{2i-1}}+\eta_{k_{2i-1}}-\xi_{k_{2i}}-\eta_{k_{2i}})^2+(\xi_{k_{2n+1}}+\eta_{k_{2n+1}})^2\Big)^{\frac{1}{2}}\\ \leq&\Big(\sum_{i=1}^{n}(\xi_{k_{2i-1}}-\xi_{k_{2i}})^2+\xi_{k_{2n+1}}^2\Big)^{\frac{1}{2}}+\Big(\sum_{i=1}^{n}(\eta_{k_{2i-1}}-\eta_{k_{2i}})^2+(\eta_{k_{2n+1}})^2\Big)^{\frac{1}{2}}\leq\|x\|+\|y\|\end{aligned}$$

故 $\|x+y\|\leq\|x\|+\|y\|$.

再说明 J 在 $\|\cdot\|$ 下的完备性. 设 $x_n=\{\xi_i^{(n)}\}_{i\in\mathbb{N}}(n=1,2,\cdots)$ 是 J 中的基本列, 则根据基本列的定义知任给 $\varepsilon>0$, 存在正整数 n_0, 当 $m,s\geq n_0$ 时有

$$\|x_s-x_m\|=\sup\Big(\sum_{i=1}^{n}(\xi_{k_{2i-1}}^{(s)}-\xi_{k_{2i-1}}^{(m)}-\xi_{k_{2i}}^{(s)}+\xi_{k_{2i}}^{(m)})^2+(\xi_{k_{2n+1}}^{(s)}-\xi_{k_{2n+1}}^{(m)})^2\Big)^{\frac{1}{2}}<\varepsilon \tag{3.75}$$

于是 $|\xi_{k_{2n+1}}-\xi_{k_{2n+1}}^{(m)}|<\varepsilon$, 而根据 J 的构造本身有 $\xi_{k_{2n+1}}^{(m)}\to 0(n\to\infty)$, 故 $\xi_{k_{2n+1}}\to 0(n\to\infty)$. 同样, 由 $|\xi_{k_{2i-1}}-\xi_{k_{2i-1}}^{(m)}-\xi_{k_{2i}}+\xi_{k_{2i}}^{(m)}|<\varepsilon$ 与 $\xi_i^{(m)}\to 0(i\to\infty)$ 可得 $\xi_{k_{2i}}\to 0(i\to\infty)$. 这样一来, $(\xi_1,\cdots,\xi_n,\cdots)$ 的偶数项与奇数项组成的子列均收敛于 0, 因此 $\xi_i\to 0(i\to\infty)$.

现在在(3.75)式中令 $s\to\infty$, 得到 $\|x-x_m\|\leq\varepsilon(m\geq n_0)$, 因此 $x-x_m\in J$, 但 $x_m\in J$, 故 $x\in J$, 且 $x_m\to x$, 于是 J 为 Banach 空间.

J 称为 James 空间, 它有下述基本性质:

(a) J 是可分的 Banach 空间, $\{e_n\}_{n\in\mathbb{N}}$ 是 J 的收缩的单调基其中 $e_n=(0,\cdots,0,1,0,\cdots)$.

(b) $\{e_n\}_{n\in\mathbb{N}}$ 不是有界完全的, 因而 J 不是自反空间.

(c) J 没有无条件基.

(d) J^* 没有无条件基, 也就是说 $\{e_n\}_{n\in\mathbb{N}}$ 的坐标泛函列是 J^* 的有界完全基, 但不是收缩基.

(e) J 与 J^{**} 等距同构, J 的所有共轭空间均可分.

补充命题 3.6.1

Hilbert 空间总是自反的.

证明

对 Hilbert 空间 $\mathscr{X}$ 而言, 根据 Riesz 表示定理知任意 $f \in \mathscr{X}^*$ 均存在唯一 $y_f \in \mathscr{X}$ 使得

$$f(x) = (x, y_f)$$

另外, 对任意 $y \in \mathscr{X}$ 总可以通过定义 $f_y(x) = (x, y)$ 唯一地得到 $f \in \mathscr{X}^*$, 这说明 $\mathscr{X}$ 与 $\mathscr{X}^*$ 是等距同构的. 现取 $\mathscr{X}$ 在 x 处经过自然映射得到的元素 $X : X(f) = f(x)$, 只需说明 X 是满射即可. 这是因为任取 $r \in \mathbb{R}$, 总存在元素 $x_r = (r, 0, \cdots)$ 与 $e_1 = (1, 0, \cdots)$, 使得 $(e_1, x_r) = r$. 现取 x_r 对应的 $f_{x_r} \in \mathscr{X}^*$, 则知 f_{x_r} 是 r 在 e_1 对应的自然映射下的原像, 亦即自然映射是满射, 命题即证. □

从前面对 L^p 空间的讨论可知: 当 $1 < p < \infty$ 时, $L^p(\Omega, \mathscr{B}, \mu)$ 是自反的, 而当 $p = 1, \infty$ 时, $L^p(\Omega, \mathscr{B}, \mu)$ 不是自反的.

例 3.37★(有限维线性赋范空间都是自反空间) 设 $\mathscr{X}$ 是有限维 B^* 空间, $\dim \mathscr{X} = n$, $\{x_i\}_{i=1}^n$ 是 $\mathscr{X}$ 的一组基, 则由习题3.35知存在 $\{f_i\}_{i=1}^n \subset \mathscr{X}^*$ 使得 $f_i(x_j) = \delta_{ij}(i, j \in \{1, \cdots, n\})$. 从而任取 $x \in \mathscr{X}$, 只要 $x = \sum\limits_{i=1}^n a_i x_i$, 其中 $\{a_i\}_{i=1}^n \subset \mathbb{K}$, 则 $f_j(x) = \sum\limits_{i=1}^n a_i f_j(x_i) = a_j (\forall j \in \{1, \cdots, n\})$, 故 $f(x) = \sum\limits_{i=1}^n a_i f(x_i) = \sum\limits_{i=1}^n f(x_i) f_i(x)$, 于是 $f = \sum\limits_{i=1}^n f(x_i) f_i$, 因此 $\dim \mathscr{X}^* \leq n$. 又若存在 $\{\widetilde{a}_j\}_{j=1}^n \subset \mathbb{K}$ 使得 $\forall x \in \mathscr{X} (\sum\limits_{i=1}^n \widetilde{a}_i f_i(x) = 0)$, 则

$$\forall j \in \{1, \cdots, n\} (\sum_{i=1}^n \widetilde{a}_i f_i(x_j) = 0) \Rightarrow \sum_{i=1}^n \widetilde{a}_i \delta_{ij} = \widehat{a}_j = 0$$

这说明 $\{f_i\}_{i=1}^n \subset \mathscr{X}^*$ 线性无关, 故 $\dim \mathscr{X}^* = n$. 由此进一步有 $\dim \mathscr{X}^{**} = n$, 又因为 $\mathscr{X} \subset \mathscr{X}^{**}$, 故 $\mathscr{X} = \mathscr{X}^{**}$, 从而有限维线性空间均自反.

注 ★ 前面说明了 Hilbert 空间总是自反的, 但有限维线性赋范空间未必是 Hilbert 空间. 例如在 $\mathbb{R}^2$ 上赋范数 $\|(x_1, x_2)\|_1 := |x_1| + |x_2| (\forall (x_1, x_2) \in \mathbb{R}^2)$, 取 $x = (1, 2), y = (2, 1)$ 可知该范数并不满足平行四边形法则, 故 $(\mathbb{R}^2, \|\cdot\|_1)$ 不是内积空间, 更不可能是 Hilbert 空间.

注 ♡ 判定空间自反与否也可以从可分性与 Banach 定理3.6.1入手, 下面给出两例.

(i) $L^1[a, b]$ 不自反. 这是因为 $L^1[a, b]^* = L^\infty[a, b]$, 若 $L^1[a, b]$ 自反, 就应有 $L^\infty[a, b]^* = L^1[a, b]$. 现在根据 Banach 定理3.6.1, $L^1[a, b]$ 是可分空间, 从而 $L^\infty[a, b]$ 也应为可分空间. 但任取 $\xi \in [a, b]$, 构造

$$h_\xi(x) = \begin{cases} 1, & x \in (\xi, b] \\ 0, & x \in [a, \xi] \end{cases}$$

知 $\{h_\xi\}_{\xi \in [a, b]}$ 与 $[a, b]$ 等势, 且 $\xi_1 \neq \xi_2 \Rightarrow \|h_{\xi_1} - h_{\xi_2}\|_{L^\infty} = 1$. 现若存在 $L^\infty[a, b]$ 的可数稠密子集 $\{g_n\}_{n \in \mathbb{N}}$, 则应有 $L^\infty[a, b] \subset \bigcup\limits_{n \in \mathbb{N}} B(g_n, \frac{1}{2})$, 进一步

$$\{h_\xi\}_{\xi \in [a, b]} \subset \bigcup_{n \in \mathbb{N}} B(g_n, \frac{1}{2})$$

因为前者的势比后者大, 故根据鸽笼原理, 必存在 $\xi_1 \neq \xi_2$ 使得 h_{ξ_1}, h_{ξ_2} 在同一个 $B(g_n, \frac{1}{2})$ 内, 但这说明

$$\|h_{\xi_1} - h_{\xi_2}\|_{L^\infty} \leq \|h_{\xi_1} - g_n\|_{L^\infty} + \|h_{\xi_2} - g_n\|_{L^\infty} < 1$$

矛盾! 于是 $L^\infty[a, b]$ 不可分, 导出矛盾.

(ii) $C[0, 1]$ 不自反. 这是因为 $C[0, 1]^* = BV[0, 1]$, 若 $C[0, 1]$ 自反, 就应有 $BV[0, 1]^* = C[0, 1]$. 现在根据 Banach 定理3.6.1, $C[0, 1]$ 是可分空间, 从而 $BV[0, 1]$ 也应为可分空间. 但任取 $\tau \in (0, 1)$. 构造

$$h_\tau(t) = \begin{cases} 0, & 0 \leq t < \tau \\ 1, & \tau \leq t \leq 1 \end{cases}$$

则 $\|h_\tau\| = 1$, 且 $\forall \tau_1 \neq \tau_2 (\|h_{\tau_1} - h_{\tau_2}\| = 2)$. 注意 $\{h_\tau\}_{\tau \in (0, 1)}$ 与 $(0, 1)$ 等势, 类似于前例理由即可说明 $BV[0, 1]$

不可分, 导出矛盾.

特别需要注意的是, 上两例中否定空间可分性时构造的函数是依赖于区域的. 对不同的区域而言上述结论不一定永远成立.

3.6.1.2 共轭算子

共轭算子是有穷维空间中转置矩阵的推广. $n \times m$ 矩阵 $A = (a_{ij})$ 可以看成 $\mathbb{K}^m \to \mathbb{K}^n$ 的线性算子:

$$(Ax)_i = \sum_{j=1}^{m} a_{ij}x_j \quad (\forall x = (x_1, x_2, \cdots, x_m) \in \mathbb{K}^m, i = 1, 2, \cdots, n)$$

其转置矩阵定义的 $m \times n$ 矩阵 $A^* = (a_{ji})$, 它是 $\mathbb{K}^n \to \mathbb{K}^m$ 的线性算子:

$$(A^*y)_j = \sum_{i=1}^{n} a_{ij}y_i \quad (\forall y = (y_1, y_2, \cdots, y_n) \in \mathbb{K}^n, j = 1, 2, \cdots, m)$$

现在希望把这个关系推广到 B 空间, 注意到

$$\langle y, Ax\rangle_n = \sum_{i=1}^{n}(\sum_{j=1}^{m} a_{ij}x_j)y_i = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}x_jy_i = \sum_{j=1}^{m}(\sum_{i=1}^{n} a_{ij}y_i)x_j = \langle A^*x, y\rangle_m$$

其中

$$\langle y, z\rangle_n = \sum_{i=1}^{n} y_iz_i, \quad \forall y = (y_1, y_2, \cdots, y_n), z = (z_1, z_2, \cdots, z_n) \in \mathbb{K}^n$$

$$\langle w, x\rangle_m = \sum_{j=1}^{m} w_jx_j, \quad \forall w = (w_1, w_2, \cdots, w_m), x = (x_1, x_2, \cdots, x_m) \in \mathbb{K}^m$$

从中诱导出共轭算子的定义.

定义 3.6.4 (共轭算子)

设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, 算子 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. 算子 $T^* : \mathscr{Y}^* \to \mathscr{X}^*$ 称为是 T 的共轭算子是指:

$$f(Tx) = (T^*f)(x), \quad \forall f \in \mathscr{Y}^*, \forall x \in \mathscr{X}$$

♣

注 对任意的 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, T^* 都是唯一存在的, 且属于 $\mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$. 这是因为对任意的 $f \in \mathscr{Y}^*$, 令

$$g(x) = f(Tx), \quad \forall x \in \mathscr{X}$$

显见 g 是线性的, 且 g 有界:

$$|g(x)| \leq \|f\| \cdot \|T\| \cdot \|x\|, \quad \forall x \in \mathscr{X}$$

这说明 $g \in \mathscr{X}^*$, 对应 $f \mapsto g$ 又是线性的, 该对应正是 T^*. 根据定义知

$$\|T^*f\| = \|g\| \leq \|T\| \cdot \|f\|, \quad \forall f \in \mathscr{Y}$$

故 $T^* \in \mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$. 同时有

$$\|T^*\| \leq \|T\|$$

这说明 T^* 是唯一的, 因为此时 $* : T \mapsto T^*$ 是单射.

定理 3.6.4

映射 $* : T \mapsto T^*$ 是 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 到 $\mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$ 内的等距同构.

♡

证明

首先证明 $* : T \mapsto T^*$ 是线性的:

$$[(\alpha_1T_1+\alpha_2T_2)^*f](x) = f[(\alpha_1T_1+\alpha_2T_2)x] = \alpha_1f(T_1x)+\alpha_2f(T_2x) = [(\alpha_1T_1^*+\alpha_2T_2^*)f](x), \quad \forall x \in \mathscr{X}, \forall f \in \mathscr{Y}^*, \forall \alpha_1, \alpha_2 \in \mathbb{K}$$

再证明 * 是等距的, 因为已知 $\|T^*\| \le \|T\|$, 故只需证 $\|T\| \le \|T^*\|$. 对任意固定的 $x \in \mathscr{X}$, 只要 $Tx \neq \theta$, 根据推论(3.5.2)知必存在 $f \in \mathscr{Y}^*$ 使得

$$f(Tx) = \|Tx\|, \quad \|f\| = 1$$

得到

$$\|Tx\| = f(Tx) = (T^*f)(x) \le \|T^*f\| \cdot \|x\| \le \|T^*\| \cdot \|x\|$$

此即

$$\|T\| \le \|T^*\|$$

□

对 T^* 而言, 当然也可以考察它的共轭算子 $T^{**} = (T^*)^* \in \mathscr{L}(\mathscr{X}^{**}, \mathscr{Y}^{**})$. 注意 $\mathscr{X} \subset \mathscr{X}^{**}, \mathscr{Y} \subset \mathscr{Y}^{**}$, 并设它们的自然嵌入映射分别为 U, V[20], 则

$$\langle T^{**}Ux, f\rangle = \langle Ux, T^*f\rangle = \langle T^*f, x\rangle = \langle f, Tx\rangle = \langle VTx, f\rangle, \quad \forall f \in \mathscr{Y}^*, \forall x \in \mathscr{X}$$

得到

$$T^{**}Ux = VTx$$

这意味着下述图是交换图

$$\begin{tikzcd} \mathscr{X} \arrow[r, "T"] \arrow[d, "U"'] & \mathscr{Y} \arrow[d, "V"] \\ \mathscr{X}^{**} \arrow[r, "T^{**}"'] & \mathscr{Y}^{**} \end{tikzcd}$$

从而 T^{**} 是 T 在 $\mathscr{X}^{**}$ 上的扩张, 进而得到下述定理.

定理 3.6.5

设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 则 $T^{**} \in \mathscr{L}(\mathscr{X}^{**}, \mathscr{Y}^{**})$ 是 T 在 $\mathscr{X}^{**}$ 上的延拓, 且满足 $\|T^{**}\| = \|T\|$. ♡

例 3.38 设 $(\Omega, \mathscr{B}, \mu)$ 是一个测度空间, 又设 $K(x, y)$ 是 $\Omega \times \Omega$ 上的二元平方可积函数:

$$\iint_{\Omega\times\Omega} |K(x, y)|^2 d\mu(x)d\mu(y) < \infty$$

定义算子

$$T : u \mapsto (Tu)(x) = \int_\Omega K(x, y)u(y)d\mu(y) \quad (\forall u \in L^2(\Omega, \mu))$$

则有 $T \in \mathscr{L}(L^2(\Omega, \mu))$, 且

$$(T^*v)(x) = \int_\Omega K(y, x)v(y)d\mu(y) \quad (\forall v \in L^2(\Omega, \mu))$$

这是因为

$$\begin{aligned} \|Tu\|^2 &= \int_\Omega |\int_\Omega K(x, y)u(y)d\mu(y)|^2 d\mu(x) \\ &\le \int_\Omega (\int_\Omega |K(x, y)|^2 d\mu(y) \int_\Omega |u(y)|^2 d\mu(y))d\mu(x) \\ &\le (\iint_{\Omega\times\Omega} |K(x, y)|^2 d\mu(x)d\mu(y))\|u\|^2, \quad (\forall u \in L^2(\Omega, \mu)) \end{aligned}$$

[20]注意 $\langle f, x\rangle = \langle Ux, f\rangle$, 对 V 同理.

其中 $\|\cdot\|$ 是 $L^2(\Omega,\mu)$ 上的范数, 这说明 $T\in\mathscr{L}(L^2(\Omega,\mu))$. 同时

$$\begin{aligned}\langle T^*v,u\rangle=\langle v,Tu\rangle&=\int_\Omega(\int_\Omega K(x,y)u(y)d\mu(y))v(x)d\mu(x)\\&=\iint_{\Omega\times\Omega}K(x,y)u(y)v(x)d\mu(x)d\mu(y)\\&=\int_\Omega(\int_\Omega K(x,y)v(x)d\mu(x))u(y)d\mu(y),\quad\forall u,v\in L^2(\Omega,\mu)\end{aligned}$$

其中 Fubini 定理成立是因为

$$\iint_{\Omega\times\Omega}|K(x,y)||u(y)||v(x)|d\mu(x)d\mu(y)\le\|u\|\cdot\|v\|\cdot(\iint_{\Omega\times\Omega}|K(x,y)|^2d\mu(x)d\mu(y))^{\frac{1}{2}}<\infty$$

综上有

$$(T^*v)(y)=\int_\Omega K(x,y)v(x)d\mu(x),\quad\forall v\in L^2(\Omega,\mu)$$

□

下面考察卷积算子与其共轭算子.

例 3.39 设 $K(x)$ 是 $\mathbb{R}$ 上的 L^1 函数, 考察空间 $L^p(\mathbb{R})(1\le p\le\infty)$ 上的卷积算子:

$$(K*f)(x):=\int_\mathbb{R}K(x-y)f(y)dy$$

并求其共轭. 首先说明 $K*$ 是 $L^p(\mathbb{R})$ 到自身的有界线性算子, 为此引入 Young 不等式:

引理 3.6.2 (Young 不等式)

设 $f\in L^p(\mathbb{R})(1\le p\le\infty), K\in L^1(\mathbb{R})$, 则

$$\|K*f\|_p\le\|K\|_1\cdot\|f\|_p$$

其中 $\|\cdot\|_p$ 表示 $L^p(\mathbb{R})(1\le p\le\infty)$ 的范数.

证明

当 $1<p<\infty$ 时, 由 Hölder 不等式知

$$|\int_\mathbb{R}K(x-y)f(y)dy|\le\int_\mathbb{R}|K(x-y)|^{\frac{1}{q}}\cdot|K(x-y)|^{\frac{1}{p}}|f(y)|dy\le(\int_\mathbb{R}|K(x-y)|dy)^{\frac{1}{q}}\cdot(\int_\mathbb{R}|K(x-y)|\cdot|f(y)|^pdy)^{\frac{1}{p}}\quad(3.76)$$

注意

$$\|K*f\|_p=(\int_\mathbb{R}|\int_\mathbb{R}K(x-y)f(y)dy|^pdx)^{\frac{1}{p}}$$

而代入(3.76)式有

$$\begin{aligned}\int_\mathbb{R}|\int_\mathbb{R}K(x-y)f(y)dy|^pdx&\le\int_\mathbb{R}((\int_\mathbb{R}|K(x-y)|dy)^{\frac{p}{q}}\cdot\int_\mathbb{R}|K(x-y)|\cdot|f(y)|^pdy)dx\\&=\|K\|_1^{\frac{p}{q}}\cdot\iint_{\mathbb{R}^2}|K(x-y)|\cdot|f(y)|^pdydx=\|K\|_1^{\frac{p}{q}}\cdot\|f\|_p^p\cdot\|K\|_1<\infty\end{aligned}$$

根据 Fubini 定理知 $(K*f)(x)$ 几乎处处存在有限, 且

$$\|K*f\|_p^p\le\|K\|_1^{1+\frac{p}{q}}\cdot\|f\|_p^p\Rightarrow\|K*f\|_p\le\|K\|_1\cdot\|f\|_p.$$

当 $p=1$ 或 ∞ 时, 只需套用 $\|\cdot\|_\infty$ 的定义即得结论. □

现在已经说明了卷积算子 $K*$ 是有界线性算子, 下面来求它的共轭算子. 记 $\check{K}(x):=K(-x)$, 根据 Fubini 定理知

$$\int_\mathbb{R}(\int_\mathbb{R}K(x-y)f(y)dy)g(x)dx=\int_\mathbb{R}f(y)(\int_\mathbb{R}K(x-y)g(x)dx)dy=\int_\mathbb{R}(\check{K}*g)(y)f(y)dy$$

这说明 $T:=K*$ 的共轭算子正是 $T^*=\check{K}*$. □

注 上面所考虑的共轭空间和共轭算子都是在 $\mathbb{R}$ 上的. 如果在 $\mathbb{C}$ 上, 则每个 L^q 函数 g 对应 L^p 空间上的一个反连

续线性泛函:

$$F_g(f) = \int_\Omega f \cdot \overline{g} d\mu$$

此时复共轭算子为 $T^* = \breve{\overline{K}}*$.

3.6.1.3 弱收敛及 * 弱收敛

有穷维空间中, 任意收敛点列必有收敛子列, 但无穷维空间没有这条性质. 故为了推广有穷维空间具备的性质, 引入弱收敛和 * 弱收敛的概念.

定义 3.6.5 (弱收敛, 弱极限)

设 $\mathscr{X}$ 是一个 B^* 空间, $\{x_n\} \subset \mathscr{X}, x \in \mathscr{X}$. 称 $\{x_n\}$ 弱收敛到 x, 记作 $x \rightharpoonup x$, 如果 $\forall f \in \mathscr{X}^*$ 都有

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

此时 x 称作点列 $\{x_n\}$ 的弱极限.

注 区别起见, 此后称 $x_n \to x$(按范数收敛) 为 $\{x_n\}$ 强收敛到 x, 即 x 是 $\{x_n\}$ 的强极限.

命题 3.6.1

若 $\dim \mathscr{X} < \infty$, 则弱收敛与强收敛等价.

证明

设 $e_1, e_2, \cdots, e_m$ 是 $\mathscr{X}$ 的一组基, 并设

$$x_n = \xi_1^{(n)} e_1 + \xi_2^{(n)} e_2 + \cdots + \xi_m^{(n)} e_m \quad (n = 1, 2, \cdots)$$
$$x = \xi_1^{(0)} e_1 + \xi_2^{(0)} e_2 + \cdots + \xi_m^{(0)} e_m$$

根据练习(3.35), 可取 $f_i \in \mathscr{X}^*(i = 1, 2, \cdots, m)$ 使得 $f_i(e_j) = \delta_{ij}(i, j = 1, 2, \cdots, m)$, 得到

$$f_i(x_n) = \xi_i^{(n)}, f_i(x) = \xi_i^{(0)}, \quad i = 1, 2, \cdots, m$$

现在若 $x_n \rightharpoonup x$, 则 $\lim_{n\to\infty} f(x_n) = f(x)(\forall f \in \mathscr{X}^*)$, 代入上述 f_i 即

$$\lim_{n\to\infty} f_i(x_n) = \lim_{n\to\infty} \xi_i^{(n)} = \xi_i^{(0)} = f_i(x), \quad i = 1, 2, \cdots, m$$

这说明 $\{x_n\}$ 强收敛到 x. 而当 $x_n \to x$, 根据 f 的有界线性性立得 $x_n \rightharpoonup x$. □

命题 3.6.2

(i) 弱极限若存在必唯一.

(ii) 强极限若存在必是弱极限.

证明

(i) 当 $x_n \rightharpoonup x, x_n \rightharpoonup y(n \to \infty)$, 根据定义知

$$f(x) = \lim_{n\to\infty} f(x_n) = f(y), \quad \forall f \in \mathscr{X}^*$$

于是 $\forall f \in \mathscr{X}^*(f(x-y) = 0)$. 由推论(3.5.2)知存在 $f_0 \in \mathscr{X}^*$ 使得 $\|f_0\| = 1, f_0(x-y) = \|x-y\|$, 于是 $\|x-y\| = 0$, 亦即 $x = y$.

(ii) 若 $x_n \to x(n \to \infty)$, 则对任意的 $f \in \mathscr{X}^*$ 有

$$|f(x_n) - f(x)| \le \|f\| \cdot \|x_n - x\| \to 0, \quad n \to \infty$$

即

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

故 $x_n \rightharpoonup x(n\to\infty)$. □

例 3.40(弱极限存在但强极限不存在) 在 $L^2[0,1]$ 中, 设 $x_n = x_n(t) = \sin n\pi t$, 由 Riemann-Lebesgue 定理知

$$\langle f, x_n\rangle = \int_0^1 f(t)\sin n\pi t dt \to 0, \quad \forall f \in L^2[0,1]$$

也即 $x_n \rightharpoonup \theta(n\to\infty)$. 但 $\|x_n\|_{L^2} = \frac{1}{\sqrt{2}}$, 不可能有 $x_n \to \theta(m\to\infty)$.

注意如果 $x_n \rightharpoonup x(n\to\infty)$, 则可以找到 $\{x_n\}$ 的凸组合序列, 使其强收敛到 x.

定理 3.6.6 (Mazur)

设 $\mathscr{X}$ 是实 B^* 空间, $x_n \rightharpoonup x_0(n\to\infty)$, 则 $\forall\varepsilon>0, \exists\lambda_i\geq 0(i=1,2,\cdots,n), \sum_{i=1}^n\lambda_i = 1$, 使得

$$\|x_0 - \sum_{i=1}^n \lambda_i x_i\| \leq \varepsilon.$$

证明

设 $M := \overline{\mathrm{co}(\{x_n\}_{n=1}^\infty)}$, 知 M 是 $\mathscr{X}$ 中的一个闭凸集. 若 $x_0 \notin M$, 则由 Ascoli 定理3.5.4知, 存在 $f\in\mathscr{X}^*, \alpha\in\mathbb{R}$, 使得

$$f(x) < \alpha < f(x_0), \quad \forall x\in M$$

得到

$$f(x_n) < \alpha < f(x_0), \quad \forall n\in\mathbb{N}$$

于是由 $x_n \rightharpoonup x(n\to\infty)$ 知 $\lim_{n\to\infty} f(x_n) = f(x_0) \leq \alpha < f(x_0)$, 矛盾! □

注 ⋆ 事实上, Mazur 定理3.6.6对复 B^* 空间依旧成立, 为此先阐明下述命题:

补充命题 3.6.2 (⋆)

设 $\mathscr{X}$ 是赋范线性空间, M 是 $\mathscr{X}$ 的一个子集, $x_0\in\mathscr{X}$, 则 $x_0\in\overline{\mathrm{co}(M)}$ 的那个且仅当对任意 $f\in\mathscr{X}^*$ 与任意 $\alpha\in\mathbb{R}$, 只要 $f(x)$ 满足 $\forall x\in M(\mathrm{Re}\, f(x)\leq\alpha)$, 就有 $\mathrm{Re}\, f(x_0)\leq\alpha$.

证明

当 $x_0\in\overline{\mathrm{co}(M)}$, 根据闭包的定义知存在 $\{y_n\}_{n\in\mathbb{N}}\subset\mathrm{co}(M)$ 使得 $y_n\to x_0$. 根据凸包的定义, 记 $y_n = \sum_{l=1}^{k_n}\lambda_l^{(n)}\xi_l^{(n)}$, 其中 $\xi_l^{(n)}\in M, \sum_{l=1}^{k_n}\lambda_l^{(n)} = 1, 0\leq\lambda_l^{(n)}\leq 1$, 得到

$$f(y_n) = f(\sum_{l=1}^{k_n}\lambda_l^{(n)}\xi_l^{(n)}) = \sum_{l=1}^{k_n}\lambda_l^{(n)} f(\xi_l^{(n)}) = \sum_{l=1}^{k_n}\lambda_l^{(n)}\,\mathrm{Re}\, f(\xi_l^{(n)}) + i\sum_{l=1}^{k_n}\lambda_l^{(n)}\,\mathrm{Im}\, f(\xi_l^{(n)})$$

现在对任意 $f\in\mathscr{X}^*$ 与任意 $\alpha\in\mathbb{R}$, 若 $\forall x\in M(\mathrm{Re}\, f(x)\leq\alpha)$, 则

$$\mathrm{Re}\, f(x_0) = \lim_{n\to\infty}\mathrm{Re}\, f(y_n) = \lim_{n\to\infty}\sum_{l=1}^{k_n}\lambda_l^{(n)}\,\mathrm{Re}\, f(\xi_l^{(n)}) \leq \alpha$$

当 $\forall f\in\mathscr{X}^*\forall\alpha\in\mathbb{R}(\forall x\in M(\mathrm{Re}\, f(x)\leq\alpha)\Rightarrow\mathrm{Re}\, f(x_0)\leq\alpha)$, 往证 $x_0\in\overline{\mathrm{co}(M)}$. 若 $x_0\notin\overline{\mathrm{co}(M)}$, 首先考虑 $\mathscr{X}$ 是实线性赋范空间的情况. 根据 Ascoli 定理3.5.4知存在 $g\in\mathscr{X}^*, \alpha\in\mathbb{R}$ 使得

$$g(x) < \alpha < g(x_0), \quad \forall x\in\overline{\mathrm{co}(M)}$$

现令 $f(x) = g(x) - ig(ix)$, 则 $f(ix) = g(ix) - ig(-x) = i(g(x) - ig(ix)) = if(x)$, 故 f 是复线性泛函, 且对任意 $x\in\mathscr{X}$ 有

$$|f(x)| = \sqrt{g(x)^2 + g(ix)^2} \leq \sqrt{\|g\|^2\|x\|^2 + \|g\|^2\|ix\|^2} = \sqrt{2}\|g\|\|x\|$$

故 $f \in \mathscr{X}^*$, 亦即 f 是复线性空间 $\mathscr{X}$ 上的有界线性泛函, 且满足

$$\operatorname{Re} f(x) = g(x) < \alpha < g(x_0) = \operatorname{Re} f(x_0), \quad \forall x \in M$$

但这与 $\operatorname{Re} f(x_0) \leq \alpha$ 矛盾! 故 $x_0 \in \overline{\operatorname{co}(M)}$, 命题得证. □

现在给出复 B^* 空间上 Mazur 定理3.6.6的证明.

证明

设 $M = \overline{\operatorname{co}(\{x_n\}_{n\in\mathbb{N}})}$, 则 M 是 $\mathscr{X}$ 中的复凸集. 要说明 $x_0 \in M$, 根据前述命题3.6.2, 只需说明 $\forall f \in \mathscr{X}^* \forall \alpha \in \mathbb{R}(\forall n \in \mathbb{N}(\operatorname{Re} f(x_n) \leq \alpha) \Rightarrow \operatorname{Re} f(x_0) \leq \alpha)$. 事实上, 因为 $x_n \rightharpoonup x_0$, 故 $f(x_n) \to f(x_0)(n \to \infty)$, 于是 $\operatorname{Re} f(x_n) \to \operatorname{Re} f(x_0)$, 又由 $\operatorname{Re} f(x_n) \leq \alpha (n \in \mathbb{N})$ 知 $\operatorname{Re} f(x_0) \leq \alpha$, 故 $x_0 \in M$, 命题证毕. □

补充命题 3.6.3 (♡)

设 $x_n \rightharpoonup x_0(n \to \infty)$, 则 $\|x_0\| \leq \varliminf_{n\to\infty} \|x_n\|$. ♠

证明

任取 $f \in \mathscr{X}^*$, 设 $\|f\| \leq 1$, 则

$$|f(x_0)| = \lim_{n\to\infty} |f(x_n)| = \varliminf_{n\to\infty} |f(x_n)| \leq \varliminf_{n\to\infty} \|f\| \cdot \|x_n\| = \varliminf_{n\to\infty} \|x_n\|$$

于是

$$\|x_0\| = \sup_{\|f\|\leq 1} |f(x_0)| \leq \varliminf_{n\to\infty} \|x_n\|$$

□

既然 $\mathscr{X}^*$ 也是 B 空间, 在 $\mathscr{X}^*$ 上也可以讨论弱收敛. 称弱收敛 $f_n \rightharpoonup f$, 如果 $\forall x^{**} \in \mathscr{X}^{**}(x^{**}(f_n) \to x^{**}(f))$. 有时为了不涉及 $\mathscr{X}^{**}$, 只选择考察 $\mathscr{X}$.

定义 3.6.6 (* 弱收敛, * 弱极限)

设 $\mathscr{X}$ 是 B^* 空间, $\{f_n\} \subset \mathscr{X}^*, f \in \mathscr{X}^*$. 称 $\{f_n\}$ * 弱收敛到 f, 记作 $w^* - \lim_{n\to\infty} f_n = f$, 如果对于 $\forall x \in \mathscr{X}$, 都有 $\lim_{n\to\infty} f_n(x) = f(x)$. 这时 f 称作泛函序列 $\{f_n\}$ 的 * 弱极限. ♣

补充命题 3.6.4 (⋆)

若 $\mathscr{X}$ 是 B^* 空间, 则 $(\mathscr{X}^* \ni f_n \rightharpoonup f \in \mathscr{X}^*) \Rightarrow w^* - \lim_{n\to\infty} f_n = f$, 亦即 $\mathscr{X}^*$ 上的弱收敛强于 * 弱收敛. ♠

证明

根据 $\mathscr{X}^*$ 上弱收敛的定义知

$$\forall y \in \mathscr{X}^{**}(\langle y, f_n\rangle \to \langle y, f\rangle(n \to \infty))$$

现记 U 是 $\mathscr{X} \to \mathscr{X}^{**}$ 的自然映射, 亦即

$$\forall x \in \mathscr{X} \forall g \in \mathscr{X}^*(\langle Ux, g\rangle = \langle g, x\rangle).$$

现取 $y = Ux$ 知:

$$\langle Ux, f_n\rangle \to \langle Ux, f\rangle, \quad n \to \infty$$

亦即

$$\langle f_n, x\rangle \to \langle f, x\rangle, \quad \forall x \in \mathscr{X}, n \to \infty$$

从而 $w^* - \lim_{n\to\infty} f_n = f$. □

补充命题 3.6.5 (⋆)

若 $\mathscr{X}$ 是自反 B^* 空间, 则 $(\mathscr{X}^* \ni f_n \rightharpoonup f \in \mathscr{X}^*) \Leftrightarrow w^* - \lim\limits_{n\to\infty} f_n = f$, 亦即 $\mathscr{X}^*$ 上的弱收敛与 * 弱收敛等价.

证明

由 $\mathscr{X}$ 自反知 $\mathscr{X} = \mathscr{X}^{**}$. 设 U 是 $\mathscr{X} \to \mathscr{X}^{**}$ 的自然映射, 则 $x \in \mathscr{X} \Leftrightarrow Ux \in \mathscr{X}^{**}$. 于是任取 $y \in \mathscr{X}^{**}$, 均存在 $x \in \mathscr{X}$ 使得 $y = Ux$, 有:

$$\langle y, f_n\rangle \to \langle y, f\rangle(n\to\infty) \Leftrightarrow \langle Ux, f_n\rangle \to \langle Ux, f\rangle(n\to\infty) \Leftrightarrow \langle f_n, x\rangle \to \langle f, x\rangle(n\to\infty)$$

命题即证. □

已经知道 $\mathscr{X} \subset \mathscr{X}^{**}$, 故 $\mathscr{X}^*$ 上的弱收敛可以推出 $\mathscr{X}^*$ 上的 * 弱收敛. 另外, 当 $\mathscr{X}$ 是一个自反空间时, * 弱收敛与弱收敛等价. 这是因为 $\mathscr{X}$ 自反意味着 $\mathscr{X} = \mathscr{X}^{**}$, 于是 $x \in \mathscr{X} \Leftrightarrow Ux \in \mathscr{X}^{**}$, 其中 $U : \mathscr{X} \to \mathscr{X}^{**}$ 是自然映射. 于是

$$\begin{aligned}
&\forall y \in \mathscr{X}^{**}(y(f_n) \to y(f)(n\to\infty))\\
\Leftrightarrow&\forall x \in \mathscr{X}((Ux)(f_n) \to (Ux)(f)(n\to\infty))\\
\Leftrightarrow&\forall x \in \mathscr{X}(f_n(x) \to f(x)(n\to\infty))
\end{aligned}$$

故 $\mathscr{X}^*$ 上的 * 弱收敛与弱收敛等价.

现在把 Banach-Steinhaus 定理3.4.9应用到下述情形.

定理 3.6.7

设 $\mathscr{X}$ 是一个 B^* 空间, 又设 $\{x_n\} \subset \mathscr{X}, x \in \mathscr{X}$, 则要使得 $x_n \rightharpoonup x(n\to\infty)$, 必须且仅须:

1. $\|x_n\|$ 有界;
2. 对 $\mathscr{X}^*$ 中的一个稠密子集 M^* 上的一切 f 都有
$$\lim_{n\to\infty} f(x_n) = f(x).$$

证明

把 x_n 看成 $\mathscr{X}^*$ 上的有界线性泛函:

$$A_n(f) := \langle x_n, f\rangle := f(x_n), \quad \forall f \in \mathscr{X}^*$$

现在证明 $\|A_n\|_{\mathscr{X}^{**}} = \|x_n\|_{\mathscr{X}}$. 事实上, 若 $x_n = \theta$, 则 $A_n = \theta$, 故 $\|A_n\|_{X^{**}} = 0 = \|x_n\|_{\mathscr{X}}$. 而若 $x_n \neq \theta$, 则由

$$|A_n(f)| = |\langle x_n, f\rangle| \leq \|x_n\|_{\mathscr{X}}\|f\|_{\mathscr{X}^*}$$

可知 $\|A_n\|_{\mathscr{X}^{**}} \leq \|x_n\|_{\mathscr{X}}$. 另一方面, 根据 Hahn-Banach 定理的推论知存在 $f_0 \in \mathscr{X}^*, \|f_0\|_{\mathscr{X}^*} = 1$ 使得 $f_0(x_n) = \|x_n\|_{\mathscr{X}} = A_n(f_0)$, 故 $\|x_n\|_{\mathscr{X}} \leq \|A_n\|_{\mathscr{X}^{**}}$, 从而 $\|A_n\|_{\mathscr{X}^{**}} = \|x_n\|_{\mathscr{X}}$. 现对 $\{A_n\}_{n\in\mathbb{N}}$ 套用 Banach-Steinhaus 定理3.4.9即可. □

定理 3.6.8

设 $\mathscr{X}$ 是一个 B 空间, 又设 $\{f_n\} \subset \mathscr{X}^*, f \in \mathscr{X}^*$, 则要使得 $w^* - \lim\limits_{n\to\infty} f_n = f$, 必须且仅须:

1. $\|f_n\|$ 有界;
2. 对 $\mathscr{X}$ 中的一个稠密子集 M 上的一切 x 都有
$$\lim_{n\to\infty} f_n(x) = f(x).$$

证明

这正是 Banach-Steinhaus 定理的一个特殊情况. □

上面讨论的是连续线性泛函的收敛性, 对于连续线性算子列 $\{T_n\} \subset \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 其中 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, 同样可以考察各种收敛性.

定义 3.6.7 (一致收敛, 一致极限, 强收敛, 强极限, 弱收敛, 弱极限)

设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, 又设 $T_n(n = 1, 2, \cdots), T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$.

(a) 若 $\|T_n - T\| \to 0$, 则称 T_n 一致收敛于 T, 记作 $T_n \rightrightarrows T$, 此时 T 称作 $\{T_n\}$ 的一致极限.

(b) 若 $\|(T_n - T)x\| \to 0 (\forall x \in \mathscr{X})$, 则称 T_n 强收敛于 T, 记作 $T_n \to T$, 此时 T 称作 $\{T_n\}$ 的强极限.

(c) 如果对于 $\forall x \in \mathscr{X}, \forall f \in \mathscr{Y}^*$ 都有

$$\lim_{n\to\infty} f(T_n x) = f(Tx)$$

则称 T_n 弱收敛于 T, 记作 $T_n \rightharpoonup T$, 此时 T 称作 $\{T_n\}$ 的弱极限.

♣

显见一致收敛 $\Rightarrow$ 强收敛 $\Rightarrow$ 弱收敛, 且每种极限存在必唯一. 特别地, 若 $\mathscr{Y} = \mathbb{K}$, 则此时 $\{T_n\}_{n\in\mathbb{N}}$ 与 T 均为 $\mathscr{X}$ 上的有界线性泛函. 因为 $\mathbb{K}$ 本身是 Hilbert 空间, 根据 Hilbert 空间上的 Riesz 表示定理3.3.1知 $f \in \mathbb{K}^*$ 当且仅当存在唯一 $\alpha \in \mathbb{K}$ 使得对任意 $x \in \mathbb{K}$ 均有 $f(x) = \alpha x (\mathbb{K} = \mathbb{R})$ 或 $f(x) = \overline{\alpha} x (\mathbb{K} = \mathbb{C})$, 此时 $T_n \to T$ 和 $T_n \rightharpoonup T$ 均等价于 $w^* - \lim\limits_{n\to\infty} T_n = T$.

另外, 因为 $\mathscr{Y}$ 是有限维 B^* 空间时, 其上的强收敛和弱收敛等价, 故此时上述定义中的 (ii), (iii) 是等价的. 但若 $\mathscr{Y}$ 是无穷维 B^* 空间, 该等价性不再成立. 下面给出一些反例.

例 3.41(强收敛但不一致收敛) 在空间 l^2 上考察左推移算子:

$$T : x = (x_1, x_2, \cdots, x_n, \cdots) \mapsto Tx = (x_2, x_3, \cdots, x_n, \cdots)$$

令 $T_n := T^n$, 得到

$$T_n x = (x_{n+1}, x_{n+2}, \cdots), \quad \forall x = (x_1, x_2, \cdots, x_n, \cdots) \in l^2$$

下面证明 $T_n \to 0$, 但 $T_n \not\rightrightarrows 0 (n \to \infty)$. 若取

$$e_n = (\underbrace{0, 0, \cdots, 0, 1}_{n}, 0, \cdots)$$

则 $T_n e_{n+1} = e_1$, 且 $\|e_n\| = 1 (\forall n \in \mathbb{N})$. 这说明

$$\|T_n\| = \sup_{\|x\|=1} \|T_n x\| \geq \|T_n(e_{n+1})\| = 1$$

故 $T_n \not\rightrightarrows 0$. 但对任意的 $x = (x_1, x_2, \cdots, x_n, \cdots) \in l^2$ 有

$$\|T_n x\| = (\sum_{i=1}^{\infty} |x_{n+i}|^2)^{\frac{1}{2}} \to 0, \quad n \to \infty$$

得到 $T_n \to 0$. □

例 3.42(弱收敛但不强收敛) 在空间 l^2 上考察右推移算子:

$$S : x = (x_1, x_2, \cdots, x_n, \cdots) \mapsto Sx = (0, x_1, x_2, \cdots, x_n, \cdots)$$

令 $S_n := S^n$, 得到

$$S_n x = (\underbrace{0, 0, \cdots, 0}_{n}, x_1, x_2, \cdots), \quad \forall x \in l^2$$

显见 $\|S_n x\| = \|x\| (\forall x \in l^2)$, 这说明 $S_n \nrightarrow 0$. 但对任意的 $f = (y_1, y_2, \cdots, y_n, \cdots) \in (l^2)^* = l^2$, 有

$$|\langle f, S_n x\rangle| = |\sum_{i=1}^{\infty} y_{i+n} x_i| \leq (\sum_{i=1}^{\infty} |y_{i+n}|^2)^{\frac{1}{2}} \cdot \|x\| \to 0, \quad n \to \infty$$

此即 $S_n \rightharpoonup 0 (n \to \infty)$. □

3.6.1.4 弱列紧性与 * 弱列紧性

引进弱收敛与 * 弱收敛的目的之一就是从有界性导出某种紧性.

定义 3.6.8 (弱列紧, * 弱列紧)

称集合 A 是弱列紧的, 如果 A 中任意点列有一个弱收敛子列. 称 A 是 * 弱列紧的, 如果 A 中任意点列有一个 * 弱收敛的子列. ♣

定理 3.6.9

设 $\mathscr{X}$ 是可分的 B^* 空间, 那么 $\mathscr{X}^*$ 上的任意有界列 $\{f_n\}$ 必有 * 弱收敛的子列. ♡

证明

由 $\mathscr{X}$ 可分知 $\mathscr{X}$ 有可数稠密子集 $\{x_m\}$. 因为 $\{f_n\}$ 有界, 故对每个固定的 m, 数集

$$\{\langle f_n, x_m\rangle : n \in \mathbb{N}\}$$

是有界的. 根据对角线法则可抽出子列 $\{f_{n_k}\}_{k=1}^{\infty}$, 使得对任意的 $m \in \mathbb{N}$ 而言, $\{\langle f_{n_k}, x_m\rangle\}_{k=1}^{\infty}$ 是收敛数列. 再由 $\{x_m\}$ 在 $\mathscr{X}$ 中稠密及 $\{f_n\}$ 有界知, 对任意的 $x \in \mathscr{X}$, $\{\langle f_{n_k}, x\rangle\}_{k=1}^{\infty}$ 都是收敛数列. 记 $F(x) := \lim\limits_{k\to\infty}\langle f_{n_k}, x\rangle$. 显见 F 是线性的, 且

$$|F(x)| \le \sup_n \|f_n\| \cdot \|x\|, \quad \forall x \in \mathscr{X}$$

故存在 $f \in \mathscr{X}^*$ 使得

$$\langle f, x\rangle = F(x) = \lim_{k\to\infty}\langle f_{n_k}, x\rangle, \quad \forall x \in \mathscr{X}$$

也即 $w^* - \lim\limits_{k\to\infty} f_{n_k} = f$. □

注 ♡ 定理3.6.9可以加强为下述结论.

补充定理 3.6.2

若 $(\mathscr{X}, \|\cdot\|)$ 是可分 B^* 空间, 那么 $\mathscr{X}^*$ 中任意有界集都是 * 弱列紧的. ♡

证明

设 Φ 是 $\mathscr{X}^*$ 中的有界集, 任取 Φ 中的序列 $\{\phi_n\}_{n\in\mathbb{N}}$, 根据 Φ 的有界性知存在 $M > 0$, 使得

$$\|\phi_n\|_{\mathscr{X}^*} \le M, \quad n = 1, 2, \cdots \tag{3.77}$$

因为 $\mathscr{X}$ 可分, 故设 $\{x_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}$ 的稠密子集, 由(3.77)式知 $\{\phi_n(x_1)\}_{n\in\mathbb{N}}$ 是 $\mathbb{R}$ 上的有界序列, 故其存在收敛子列 $\{\phi_{1\nu}(x_1)\}_{\nu\in\mathbb{N}}$. 再考虑 $\{\phi_{2\nu}(x_2)\}_{\nu\in\mathbb{N}}$, 知其同样是 $\mathbb{R}$ 上的有界序列, 故其存在收敛子列 $\{\phi_{2\mu}(x_2)\}_{n\in\mathbb{N}}$. 一直这样下去, 根据对角线法则选取子列 $\{\phi_{\nu\nu}\}_{\nu\in\mathbb{N}}$, 可知对任意 $k \in \mathbb{N}$, $\{\phi_{\nu\nu}(x_k)\}_{\nu\in\mathbb{N}}$ 均收敛, 而 $\{x_n\}_{n\in\mathbb{N}}$ 是 $\mathscr{X}$ 的稠密子集, $\{\phi_{\nu\nu}\}_{\nu\in\mathbb{N}}$ 满足 $\sup\limits_{\nu\ge1}\|\phi_{\nu\nu}\|_{\mathscr{X}^*} \le M$, 故由 Banach-Steinhaus 定理知存在 $\varphi \in \mathscr{X}^*$ 使得

$$w^* - \lim_{n\to\infty}\phi_{\nu\nu} = \varphi$$

命题即证. □

事实上空间不需要可分也可以导出 * 列紧性. 下面从自反性导出 * 列紧.

定理 3.6.10 (Banach)

设 $\mathscr{X}$ 是 B^* 空间. 若 $\mathscr{X}$ 的共轭空间 $\mathscr{X}^*$ 是可分的, 则 $\mathscr{X}$ 本身必可分. ♡

证明

首先考察 $\mathscr{X}^*$ 的单位球面 $S_1^* := \{f \in \mathscr{X}^* : \|f\| = 1\}$, 下面证明 S_1^* 是可分的. 这是因为由 $\mathscr{X}^*$ 可分知, 存在

$\{f_n\} \subset \mathscr{X}^*$ 使得

$$\forall f \in S_1^* \exists \{n_k\}_{k=1}^{\infty} (\lim_{k\to\infty} f_{n_k} = f)$$

现在令 $g_n := \frac{f_n}{\|f_n\|}$(这里设 $f_n \neq \theta$), 得到

$$\|f - g_{n_k}\| \leq \|f - f_{n_k}\| + \|f_{n_k} - g_{n_k}\| = \|f - f_{n_k}\| + |1 - \|f_{n_k}\|| \to 0, \quad k \to \infty$$

这说明 S_1^* 有可数的稠密子集 $\{g_n\}$.

对每个 g_n, 因为 $\|g_n\| = 1$, 根据算子范数的定义知可选取 $x_n \in \mathscr{X}$ 使得

$$\|x_n\| = 1, \quad g_n(x_n) \geq \frac{1}{2}$$

记 $\mathscr{X}_0 := \overline{\operatorname{span}\{x_n\}}$, 知 $\mathscr{X}_0$ 可分, 下面证明 $\mathscr{X}_0 = \mathscr{X}$. 如若不然, 则存在 $x_0 \in \mathscr{X} \backslash \mathscr{X}_0$, 不妨设 $\|x_0\| = 1$. 由 Hahn-Banach 定理的推论3.5.5知存在 $f_0 \in \mathscr{X}^*$, 使得 $\|f_0\| = 1$, 且 $\forall x \in \mathscr{X}_0 (f_0(x) = 0)$. 注意此时 $f_0 \in S_1^*$, 有

$$\|g_n - f_0\| = \sup_{\|x\|=1} |g_n(x) - f_0(x)| \geq |g_n(x_n) - f_0(x_n)| = |g_n(x_n)| \geq \frac{1}{2}$$

这与 $\{g_n\}$ 在 S_1^* 中稠密矛盾! 故 $\mathscr{X} = \mathscr{X}_0$, 因而 $\mathscr{X}$ 是可分的. □

定理 3.6.11 (Pettis)

自反空间 $\mathscr{X}$ 的闭子空间 $\mathscr{X}_0$ 必是自反空间.

证明

欲证 $z_0 \in \mathscr{X}^{**} \Rightarrow z_0 \in \mathscr{X}_0$, 此即证明存在 $x \in \mathscr{X}_0$ 使得

$$\langle z_0, f_0 \rangle = \langle f_0, x \rangle, \quad \forall f_0 \in \mathscr{X}_0^*$$

现在对任意的 $f \in \mathscr{X}^*$, 考察其在 $\mathscr{X}_0$ 上的限制 $Tf = f_0 \in \mathscr{X}_0^*$. 注意

$$\|f_0\| \leq \|f\|$$

故 $T \in \mathscr{L}(\mathscr{X}^*, \mathscr{X}_0^*)$, 于是 $z := T^* z_0 \in \mathscr{X}_0^{**}$, 又因为 $\mathscr{X}$ 自反, 故存在 $x \in \mathscr{X}$ 使得

$$\langle z, f \rangle = \langle f, x \rangle, \quad \forall f \in \mathscr{X}^*$$

现在证明 $x \in \mathscr{X}_0$. 如若不然, 由 Hahn-Banach 定理的推论3.5.5知存在 $f \in \mathscr{X}^*$, 使得

$$f(\mathscr{X}_0) = 0, \quad \langle f, x \rangle = 1$$

由 $f(\mathscr{X}_0) = 0$ 知 $Tf = \theta$, 但这说明

$$0 = \langle z_0, Tf \rangle = \langle T^* z_0, f \rangle = \langle z, f \rangle = \langle f, x \rangle = 1$$

矛盾! 故存在 $x \in \mathscr{X}_0$ 使得

$$\langle z, f \rangle = \langle f, x \rangle, \quad \forall f \in \mathscr{X}^*$$

现在还需要证明这个 x 满足

$$\langle z_0, f_0 \rangle = \langle f_0, x \rangle, \quad \forall f_0 \in \mathscr{X}_0^*$$

这是因为对任意的 $f_0 \in \mathscr{X}_0^*$, 由 Hahn-Banach 定理(3.5.4)知, 存在 $f \in \mathscr{X}^*$ 使得 $f_0 = Tf$, 从而

$$\langle z_0, f_0 \rangle = \langle z_0, Tf \rangle = \langle z, f \rangle$$

以及

$$\langle f_0, x \rangle = \langle f, x \rangle, \quad \forall x \in \mathscr{X}_0$$

联立即得欲证. □

定理 3.6.12 (Eberlein-Smulian)

自反空间的单位 (闭) 球是弱 (自) 列紧的.

证明

首先证明自反空间 $\mathscr{X}$ 中的任意有界点列 $\{x_n\}$ 必有一个在 $\mathscr{X}$ 中弱收敛的子列. 令

$$\mathscr{X}_0 := \overline{\operatorname{span}\{x_n\}}$$

根据 Pettis 定理3.6.11, 由 $\mathscr{X}$ 自反知 $\mathscr{X}_0$ 自反. 又显见 $\mathscr{X}_0$ 可分, 故 $\mathscr{X}_0^{**} = \mathscr{X}_0 = (\mathscr{X}_0^*)^*$ 可分, 进而由 Banach 定理3.6.10知 $\mathscr{X}_0^*$ 可分. 记 $\{g_n\}$ 为 $\mathscr{X}_0^{**}$ 中的元素, 其满足

$$\langle g_n, f\rangle = \langle f, x_n\rangle, \quad \forall f \in \mathscr{X}_0^*$$

则

$$|\langle g_n, f\rangle| \leq \|f\| \cdot \|x_n\| \Rightarrow \|g_n\| \leq \|x_n\|$$

这说明 $\{\|g_n\|\}$ 有界. 设 M_0^* 是 $\mathscr{X}_0^*$ 的可数稠密子集, 根据对角线法则可在 $\{g_n\}$ 中抽出子列 $\{g_{n_k}\}$, 并在 $\mathscr{X}_0^{**}$ 中找到某个 g 使得

$$\lim_{k\to\infty}\langle g_{n_k}, f\rangle = \langle g, f\rangle, \quad \forall f \in M_0^*$$

再根据定理3.6.8知

$$\lim_{k\to\infty}\langle g_{n_k}, f\rangle = \langle g, f\rangle, \quad \forall f \in \mathscr{X}_0^*$$

根据 $\mathscr{X}_0$ 的自反性知, 存在 $x_0 \in \mathscr{X}_0$ 使得

$$\langle g, f\rangle = \langle f, x_0\rangle, \quad \forall f \in \mathscr{X}_0^*$$

联立得到

$$\lim_{k\to\infty}\langle f, x_{n_k}\rangle = \lim_{k\to\infty}\langle g_{n_k}, f\rangle = \langle f, x_0\rangle, \quad \forall f \in \mathscr{X}_0^* \tag{3.78}$$

现在把结果延拓到 $\mathscr{X}^*$ 上, 对任意的 $\tilde{f} \in \mathscr{X}^*$, 记 $f := T\tilde{f}$ 为 $\tilde{f}$ 在 $\mathscr{X}_0$ 上的限制. 因为 $\{x_{n_k}\} \subset \mathscr{X}_0, x_0 \in \mathscr{X}_0$, 代入(3.78)式得

$$\lim_{k\to\infty}\langle \tilde{f}, x_{n_k}\rangle = \lim_{k\to\infty}\langle f, x_{n_k}\rangle = \langle f, x_0\rangle = \langle \tilde{f}, x_0\rangle, \quad \forall \tilde{f} \in \mathscr{X}^*$$

这便说明 $x_{n_k} \rightharpoonup x_0$, 故 $\mathscr{X}$ 中的任意有界集都是弱列紧集, 特别单位球也是弱列紧集, 因而单位闭球也是弱列紧集.

下面说明单位闭球是弱自列紧的. 设 $x_{n_k} \rightharpoonup x_0$, 且 $\|x_{n_k}\| \leq 1$, 由 Hahn-Banach 定理的推论3.5.2知存在 $f \in \mathscr{X}^*$ 满足

$$f(x_0) = \|x_0\|, \quad \|f\| = 1$$

故

$$\|x_0\| = f(x_0) = \lim_{k\to\infty} f(x_{n_k}) \leq \|f\| \cdot \sup_{k\geq 1}\|x_{n_k}\| \leq 1$$

这说明 x_0 也在单位闭球内, 故单位闭球弱自列紧. □

注 从直观上理解, 当 $\mathscr{X}$ 自反时, $\mathscr{X}^*$ 上的 * 弱收敛性与弱收敛性等价, 因而 $\mathscr{X}^*$ 上的 * 弱列紧性与弱列紧性应该等价. 根据 Eberlein-Smulian 定理3.6.12, 已经知道自反空间的单位球弱列紧, 故必有自反空间的单位球 * 弱列紧.

定理 3.6.13 (Alaoglu)

设 $\mathscr{X}$ 是 B^* 空间, 则 $\mathscr{X}^*$ 中的单位闭球是 * 弱紧的.

下面刻画 $L^p[0,2\pi](1<p<\infty)$ 上函数的 Fourier 级数.

设 $f \in L^2[0,2\pi]$, 称

$$c_n := \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx, \quad n = 0, \pm 1, \pm 2, \cdots$$

为 f 的 Fourier 系数, 并称级数

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

为 f 的 Fourier 级数. 当 $f \in L^2[0,2\pi]$ 时, 其 Fourier 级数在 $L^2[0,2\pi]$ 上收敛. 一般来说, 对 $L^2[0,2\pi]$ 中的函数 f, 其 Fourier 级数未必 L^1 收敛, 这在一定程度上是因为 Fourier 级数对应的算子, 即 Dirichlet 核并非好核[ST1]. 于是转而 Fourier 部分和取算术平均, 考察下述 Cesaro 和:

$$\sigma_n(f)(x) = \frac{1}{n+1}\sum_{k=0}^{n} S_k(f)(x) = \sum_{k=-n}^{n} c_k(1-\frac{|k|}{n+1})e^{ikx}, \quad S_n(f)(x) := \sum_{k=-n}^{n} c_k e^{ikx}$$

计算可得

$$\sigma_n(f)(x) = \int_0^{2\pi} f(y)K_n(x-y)dy$$

其中

$$K_n(x) := \frac{1}{2\pi}\sum_{k=-n}^{n}(1-\frac{|k|}{n+1})e^{ikx} = \frac{1}{2\pi(n+1)}(\frac{\sin(n+1)\frac{x}{2}}{\sin\frac{x}{2}})^2$$

为 Fejer 核. 注意 $K_n(x) \geq 0$, 且其 (作为好核) 满足

$$\int_0^{2\pi} K_n(x)dx = \frac{1}{2\pi}\sum_{k=-n}^{n}(1-\frac{|k|}{n+1})\int_0^{2\pi} e^{ikx}dx = 1$$

根据 Young 不等式(3.6.2)知, 对任意的 $f \in L^p[0,2\pi](1 \leq p < \infty)$ 有

$$\|\sigma_n(f)\|_{L^p} \leq \|K_n\|_{L^1} \cdot \|f\|_{L^p} = \|f\|_{L^p}$$

故若 $f \in L^p$, 则其 Cesaro 和的 L^p 范数是一致有界的. □

现在希望证明上述断言的逆命题.

定理 3.6.14

若 $1 < p \leq \infty$, 且级数

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

的 Cesaro 部分和

$$\sigma_n(x) = \sum_{k=-n}^{n} c_k(1-\frac{|k|}{n+1})e^{ikx}$$

的 L^p 范数是一致有界的, 也即

$$\sup_{n\geq 1}\|\sigma_n\|_{L^p} < \infty$$

则必存在 $f \in L^p[0,2\pi]$, 使得 σ_n 是 f 的 Fourier 级数的 Cesaro 部分和.

证明

根据定理(3.6.9), 既然 $L^q[0,2\pi]$ 可分, 故 $L^p[0,2\pi] = L^q[0,2\pi]^*(\frac{1}{p}+\frac{1}{q}=1)$ 上的任意有界列必有 * 弱收敛的子列. 现在知道

$$\sup_{n\geq 1}\|\sigma_n\|_{L^p} < \infty$$

故存在 $f \in L^p[0,2\pi]$ 与子列 $\{\sigma_{n_k}\}$ 使得在 $L^p[0,2\pi]$ 中有

$$w^* - \lim_{k\to\infty}\sigma_{n_k}(x) = f(x)$$

注意 $e^{imx} \in L^q[0,2\pi](m = 0, \pm1, \pm2, \cdots)$, 故根据 * 弱收敛的定义有

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-imx}dx = \lim_{k\to\infty}\frac{1}{2\pi}\int_0^{2\pi}\sigma_{n_k}(x)e^{-imx}dx = \lim_{k\to\infty}(1-\frac{|m|}{n_k+1})c_m = c_m$$

故 $\sigma_n(x)$ 是 f 的 Fourier 级数的 Cesaro 部分和 $\sigma_n(f)(x)$. □

3.6.1.5 弱收敛的例子

常见的弱收敛而不强收敛的例子有三种类型: 振荡, 平移, 集中.

例 3.43(振荡)

1. 取 $u_n(x) = \sin n\pi x$, 由 Riemann-Lebesgue 引理知, 当 $n \to \infty$ 时, u_n 在 $L^2[0,1]$ 中弱收敛但不强收敛于 0.
2. 锯齿型函数序列:
$$u_n(x) = \begin{cases} x - \frac{k}{n}, & x \in [\frac{k}{n}, \frac{2k+1}{n}], \\ -x + \frac{k+1}{n}, & x \in [\frac{2k+1}{n}, \frac{k+1}{n}] \end{cases}$$
可以证明 $\{u_n\}$ 在 $H^{1,2}(0,1)$ 中弱收敛但不强收敛于 0, 这主要是因为当 $n \to \infty$, u_n 在 $(0,1)$ 中的振荡越来越剧烈.

例 3.44(平移) 设 $f \in L^p(\mathbb{R})\backslash\{0\}, 1 < p < \infty$, 令
$$f_n(x) = f(x+n), \quad x \in \mathbb{R}, n = 1, 2, \cdots$$
则 $f_n \rightharpoonup 0$, 但 $\|f_n\| = \|f\|$, 因而 f_n 弱收敛但不强收敛于 0.

证明

要证明 $f_n \rightharpoonup 0$, 即证明对任意的 $g \in (L^p(\mathbb{R}))^* = L^q(\mathbb{R}), \frac{1}{p} + \frac{1}{q} = 1$, 有
$$\int_{\mathbb{R}} f_n(x)g(x)dx = \int_{\mathbb{R}} f(x+n)g(x)dx \to 0, \quad n \to \infty$$
又因为 $C_0^\infty(\mathbb{R})$ 在 $L^q(\mathbb{R})$ 中稠密, 故根据 Banach-Steinhaus 定理(3.4.9), 只需证明
$$\forall g \in C_0^\infty(\mathbb{R})(\int_{\mathbb{R}} f(x+n)g(x)dx \to 0, n \to \infty) \tag{3.79}$$
即可. 根据变量代换知
$$\int_{\mathbb{R}} f(x+n)g(x)dx = \int_{\mathbb{R}} f(x)g(x-n)dx$$
要证明(3.79)式, 先设 $f \in C_0^\infty(\mathbb{R})$, 令 $\varphi_n(x) = f(x)g(x-n)$. 既然 g 紧支且连续, 知 g 必有界, 从而存在依赖于 g 的常数 $C_g > 0$, 使得
$$|\varphi_n(x)| \leq C_g|f(x)|, \quad x \in \mathbb{R}$$
同时根据紧支性知 $\varphi_n(x) \to 0$, a.e. $n \to \infty$. 进而由 Lebesgue 控制收敛定理得
$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)g(x)dx = \lim_{n\to\infty} \int_{\mathbb{R}} f(x+n)g(x)dx = \lim_{n\to\infty} \int_{\mathbb{R}} f(x)g(x-n)dx = \int_{\mathbb{R}} \lim_{n\to\infty} \varphi_n(x)dx = 0$$
对一般的 $f \in L^p(\mathbb{R})\backslash\{0\}, 1 < p < \infty$, 既然 $C_0^\infty(\mathbb{R})$ 在 $L^p(\mathbb{R})\backslash\{0\}$ 中稠密, 知对任意的 $\varepsilon > 0$, 均可取 $f_\varepsilon \in C_0^\infty(\mathbb{R})$ 使得
$$\|f - f_\varepsilon\|_{L^p} < \varepsilon$$
进而
$$\begin{aligned} |\int_{\mathbb{R}} f_n(x)g(x)dx| &= |\int_{\mathbb{R}} f(x)g(x-n)dx| \leq |\int_{\mathbb{R}} (f(x) - f_n(x))g(x-n)dx| + |\int_{\mathbb{R}} f_n(x)g(x-n)dx| \\ &\leq \|f - f_\varepsilon\|_{L^p} \cdot \|g\|_{L^q} + |\int_{\mathbb{R}} f_\varepsilon(x)g(x-n)dx| \leq \|g\|_{L^q} \cdot \varepsilon + |\int_{\mathbb{R}} f_\varepsilon(x)g(x-n)dx| \end{aligned}$$
结合前述结论即知存在 $N > 0$, 使得当 $n \geq N$ 时有
$$|\int_{\mathbb{R}} f_n(x)g(x)dx| < \varepsilon$$
命题即证. □

例 3.45(集中)[21] 考虑由平面上定义的函数空间 $\mathscr{H}$, 它由旋转不变函数 $f(x,y)=u(r)$ 组成, 其中 (x,y) 为平面的直角坐标, $r=\sqrt{x^2+y^2}$, $u\in L^2_{\frac{4rdr}{(1+r^2)^2}}(\mathbb{R}^1_+)$[22], 其导函数 $u'\in L^2_{rdr}(\mathbb{R}^1_+)$, 其范数平方为

$$\|u\|^2=\int_{\mathbb{R}^1_+}|u(r)|^2\frac{4rdr}{(1+r^2)^2}+\int_0^\infty|u'(r)|^2rdr$$

现在给定参数 $\lambda>0$, 函数 $\phi_\lambda(r)=\frac{-1+\lambda^2r^2}{1+\lambda^2r^2},\psi_\lambda(r)=\frac{2\lambda r}{1+\lambda^2r^2}$ 都属于 $\mathscr{H}$. 进一步有

$$|\phi_\lambda(r)|^2+|\psi_\lambda(r)|^2=1$$

进而

$$\int_{\mathbb{R}^1_+}(|\phi_\lambda(r)|^2+|\psi_\lambda(r)|^2)\frac{4rdr}{(1+r^2)^2}=2$$

又因为

$$\begin{cases}(\phi'_\lambda)_r=\frac{4\lambda^2r}{(1+\lambda^2r^2)^2}\\(\psi'_\lambda)_r=\frac{2\lambda(1-\lambda^2r^2)}{(1+\lambda^2r^2)^2}\end{cases}$$

故可计算知

$$\int_{\mathbb{R}^1_+}(|(\phi'_\lambda)_r|^2+|(\psi'_\lambda)_r|^2)rdr=2$$

这说明对任意一个趋于无穷的序列 $\lambda_j\to\infty$, $(\phi_{\lambda_j},\psi_{\lambda_j})$[23]必有弱收敛子列[24], 但 ϕ_{λ_j} 并不强收敛. 这是以为除了 $r=0$ 外, 当 $\lambda\to\infty$ 时, $(\phi_\lambda,\psi_\lambda)$ 逐点收敛到常值函数 $(1,0)$.

注 上例在几何上表明: 一族能量有界的调和映射, 能量可以集中到一点.

3.6.2 习题

练习 3.47★ 求证: $(l^p)^*=l^q(1\le p<\infty,\frac{1}{p}+\frac{1}{q}=1)$.

证明

先证明 $1<p<\infty$ 的情况.

首先找一个合适的 l^q 到 $(l^p)^*$ 的映射. 任取 $y=(y_1,y_2,\cdots,y_n,\cdots)\in l^q$, 取

$$F_y(x)=\sum_{k=1}^\infty x_ky_k,\quad \forall x=(x_1,x_2,\cdots,x_n,\cdots)\in l^p$$

显见 $F_y:l^p\to\mathbb{R}$ 是线性的, 且由 Hölder 不等式知

$$|F_y(x)|\le(\sum_{k=1}^\infty x_k^p)^{\frac{1}{p}}\cdot(\sum_{k=1}^\infty y_k^q)^{\frac{1}{q}}=\|x\|_{l^p}\cdot\|y\|_{l^q}$$

这说明

$$\|F_y\|\le\|y\|_{l^q}<\infty\tag{3.80}$$

因而 $F_y\in(l^p)^*$. 现取映射 $T:y\mapsto F_y$, 需要证明 T 是一一的等距映射, 此即说明对任意一个给定的 $F\in(l^p)^*$, 存在 $y=(y_1,\cdots,y_n,\cdots)\in l^q$ 使得

$$F(x)=\sum_{k=1}^\infty x_ky_k,\quad \forall x=(x_1,\cdots,x_n,\cdots)\in l^p$$

同时

$$\|y\|_{l^q}=\|F\|_{(l^p)^*}$$

[21] 这个例子完全没有看懂.

[22] 不知道这是怎么定义的?

[23] 这是怎么定义的?

[24] 为什么?

记 $x=\sum\limits_{k=1}^{\infty}x_ke_k$, 其中 $e_k=\{\underbrace{0,\cdots,0,1}_{k},0,\cdots\}$, $x_k\in\mathbb{R}$, 知 $F(x)=\sum\limits_{k=1}^{\infty}x_kF(e_k)$. 现在对任意的 $F\in(l^p)^*$, 取

$$y_k=F(e_k),\quad y=\sum_{k=1}^{\infty}y_ke_k$$

取 $x_k=|y_k|^{q-2}y_k$ 有

$$\begin{aligned}\sum_{k=1}^{\infty}|y_k|^q&=\sum_{k=1}^{\infty}x_ky_k=F(x)\leq\|F\|_{(l^p)^*}\cdot\|x\|_{l^p}\\&=\|F\|_{(l^p)^*}\cdot(\sum_{k=1}^{\infty}|y_k|^{pq-p})^{\frac{1}{p}}=\|F\|_{(l^p)^*}\cdot(\sum_{k=1}^{\infty}|y_k|^q)^{\frac{1}{p}}\end{aligned}$$

此即

$$\|y\|_{l^q}\leq\|F\|_{(l^p)^*}$$

结合(3.80)式即得

$$\|y\|_{l^q}=\|F\|_{(l^p)^*}$$

再证明 y 是唯一的, 设另有 $y'=(y_1',\cdots,y_n',\cdots)\in l^q$ 满足 $F(x)=\sum\limits_{k=1}^{\infty}x_ky_k'(\forall x\in l^p)$, 则

$$\sum_{k=1}^{\infty}x_k(y_k-y_k')=0,\quad\forall x\in l^p$$

令 $x=e_k,k=1,2,\cdots$ 即得 $y_k=y_k'$, 故映射 T 是一一且等距的, 命题得证.

再说明 $p=1$ 的情况, F_y 与前述定义相同, 此时由 l^∞ 范数的定义知

$$|F_y(x)|\leq\|x\|_{l^1}\cdot\|y\|_{l^\infty}$$

得到 $\|F_y\|\leq\|y\|_{l^\infty}$, 这说明 $F_y\in(l^1)^*$. 取 T 与前述定义相同, 并按相同方式构造 y_k. 注意此时 $\|y\|_{l^\infty}=\sup\limits_{k\geq1}|y_k|$, 显然有不等式

$$\|y\|_{l^\infty}\leq\sup_{\|x\|_{l^1}=1}\sum_{k=1}^{\infty}|x_ky_k|=\|F\|_{(l^1)^*}$$

这便得到 T 的等距性. T 作为一一映射的证明与前述相同.

综上, 命题得证.

练习 3.48★ 设 C 是收敛数列的全体, 赋以范数

$$\|\cdot\|:\{\xi_k\}\in C\mapsto\sup_{k\geq1}|\xi_k|$$

求证: $C^*=l^1$.

证明

任取 $y=(y_1,\cdots,y_n,\cdots)\in l^1$, 取

$$F_y(x)=\sum_{k=1}^{\infty}x_ky_k,\quad\forall x=(x_1,\cdots,x_n,\cdots)\in C$$

首先说明 $F_y(x)$ 有定义, 这是因为

$$|F_y(x)|\leq\sum_{k=1}^{\infty}|x_ky_k|\leq\|x\|\cdot\sum_{k=1}^{\infty}|y_k|=\|x\|\cdot\|y\|_{l^1}<\infty\tag{3.81}$$

这说明 $F_y(x)$ 收敛. 显见 $F_y:C\to\mathbb{C}$ 是线性的, 且由(3.81)式可知

$$\|F_y\|\leq\|y\|_{l^1}\tag{3.82}$$

得到 F_y 连续, 因而 $F_y\in C^*$. 取映射 $T:y\mapsto F_y$, 需要证明 T 是一一的等距映射, 此即说明对任意一个给定的

$F \in C^*$, 存在 $y = (y_1, \cdots, y_n, \cdots) \in l^1$ 满足

$$F(x) = \sum_{k=1}^{\infty} x_k y_k, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C$$

同时

$$\|y\|_{l^1} = \|F\|$$

记 $x = \sum\limits_{k=1}^{\infty} x_k e_k$, 其中 $e_k = \{\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots\}$, $x_k \in \mathbb{R}$, 知 $F(x) = \sum\limits_{k=1}^{\infty} x_k F(e_k)$. 现在对任意的 $F \in C^*$, 取

$$y_k = F(e_k), \quad y = \sum_{k=1}^{\infty} y_k e_k$$

取

$$x_k^n = \begin{cases} e^{-i\arg(F(e_k))}, & k \leq n \\ 0, & k > n \end{cases}$$

知 $\{x_k^n\}_{k=1}^{\infty} \in C$, 且 $\forall n \in \mathbb{N}(\|x_k^n\| = 1)$. 得到

$$F(x^n) = \sum_{k=1}^{\infty} x_k^n y_k = \sum_{k=1}^{n} |F(e_k)| e^{i\arg(F(e_k))} \cdot e^{-i\arg(F(e_k))} = \sum_{k=1}^{n} |F(e_k)| = \sum_{k=1}^{n} |y_k|$$

根据算子范数的定义有

$$\|F\|_{C^*} = \sup_{\|x\|=1} |F(x)| \geq |F(x^n)| = \sum_{k=1}^{n} |y_k|, \quad \forall n \in \mathbb{N}$$

令 $n \to \infty$ 得

$$\|y\|_{l^1} = \sum_{k=1}^{\infty} |y_k| \leq \|F\|_{C^*}$$

结合(3.82)式得

$$\|y\|_{l^1} = \|F\|_{C^*}$$

这说明 T 等距, 下面说明 T 是一一映射, 这只需证明 y 的唯一性即可. 设另有 $y' \in l^1$ 满足

$$F_y(x) = \sum_{k=1}^{\infty} x_k y_k', \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C$$

得到

$$\sum_{k=1}^{\infty} x_k (y_k - y_k') = 0, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C$$

令 $x = e_k (k = 1, 2, \cdots)$ 即得 $y = y'$, 命题得证.

练习 3.49★ 设 C_0 是以 0 为极限的数列全体, 赋以范数

$$\|\cdot\| : \{\xi_k\} \in C \mapsto \sup_{k \geq 1} |\xi_k|$$

求证: $C_0^* = l^1$.

证明

任取 $y = (y_1, \cdots, y_n, \cdots) \in l^1$, 取

$$F_y(x) = \sum_{k=1}^{\infty} x_k y_k, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C_0$$

首先说明 $F_y(x)$ 有定义, 这是因为

$$|F_y(x)| \leq \sum_{k=1}^{\infty} |x_k y_k| \leq \|x\| \cdot \sum_{k=1}^{\infty} |y_k| = \|x\| \cdot \|y\|_{l^1} < \infty \tag{3.83}$$

这说明 $F_y(x)$ 收敛. 显见 $F_y : C_0 \to \mathbb{C}$ 是线性的, 且由(3.83)式可知

$$\|F_y\| \le \|y\|_{l^1}$$

得到 F_y 连续, 因而 $F_y \in C_0^*$. 取映射 $T : y \mapsto F_y$, 需要证明 T 是一一的等距映射, 此即说明对任意一个给定的 $F \in C_0^*$, 存在 $y = (y_1, \cdots, y_n, \cdots) \in l^1$ 满足

$$F(x) = \sum_{k=1}^{\infty} x_k y_k, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C_0$$

同时

$$\|y\|_{l^1} = \|F\|_{C_0^*}$$

记 $x = \sum\limits_{k=1}^{\infty} x_k e_k$, 其中 $e_k = \{\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots\}$, $x_k \in \mathbb{R}$, 知 $F(x) = \sum\limits_{k=1}^{\infty} x_k F(e_k)$. 现在对任意的 $F \in C_0^*$, 取

$$y_k = F(e_k), \quad y = \sum_{k=1}^{\infty} y_k e_k$$

取

$$x_k^n = \begin{cases} e^{-i\arg(F(e_k))}, & k \le n \\ 0, & k > n \end{cases}$$

知 $\{x_k^n\}_{k=1}^{\infty} \in C_0$, 且 $\forall n \in \mathbb{N}(\|x_k^n\| = 1)$. 得到

$$F(x^n) = \sum_{k=1}^{\infty} x_k^n y_k = \sum_{k=1}^{n} |F(e_k)| e^{i\arg(F(e_k))} \cdot e^{-i\arg(F(e_k))} = \sum_{k=1}^{n} |F(e_k)| = \sum_{k=1}^{n} |y_k|$$

根据算子范数的定义有

$$\|F\|_{C^*} = \sup_{\|x\|=1} |F(x)| \ge |F(x^n)| = \sum_{k=1}^{n} |y_k|, \quad \forall n \in \mathbb{N}$$

令 $n \to \infty$ 得

$$\|y\|_{l^1} = \sum_{k=1}^{\infty} |y_k| \le \|F\|_{C^*}$$

结合(3.6.2)式得

$$\|y\|_{l^1} = \|F\|_{C_0^*}$$

这说明 T 等距, 下面说明 T 一一, 这只需证明 y 的唯一性即可. 设另有 $y' \in l^1$ 满足

$$F_y(x) = \sum_{k=1}^{\infty} x_k y'_k, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C_0$$

得到

$$\sum_{k=1}^{\infty} x_k (y_k - y'_k) = 0, \quad \forall x = (x_1, \cdots, x_n, \cdots) \in C_0$$

令 $x = e_k (k = 1, 2, \cdots)$ 即得 $y = y'$, 命题得证.

练习 3.50★ 求证: 有限维 B^* 空间必是自反的.

证明

设 $\dim \mathscr{X} < \infty$, 进一步设 $\mathscr{X}$ 的一组基为 $\{e_1, \cdots, e_n\}$. 已经知道 $\mathscr{X} \subset \mathscr{X}^{**}$, 现在希望证明 $z \in \mathscr{X}^{**} \Rightarrow z \in \mathscr{X}$, 也即存在 $x_z \in \mathscr{X}$ 使得

$$\langle z, f \rangle = \langle f, x_z \rangle, \quad \forall f \in \mathscr{X}^*$$

根据练习(3.35)知存在 $f_1,\cdots,f_n\in\mathscr{X}^*$ 使得

$$\langle f_i,e_j\rangle=\delta_{ij},\quad 1\le i,j\le n$$

故任意固定 $x\in\mathscr{X}$, 设 $x=\sum\limits_{k=1}^{n}x_ke_k$, 有

$$\langle f_i,x\rangle=\sum_{k=1}^{n}x_kf_i(e_k)=x_i,\quad i=1,2,\cdots,n$$

从而对任意的 $f\in\mathscr{X}^*$ 有

$$\langle f,x\rangle=\sum_{k=1}^{n}x_if(e_i)=\sum_{k=1}^{n}\langle f_i,x\rangle f(e_i)=\sum_{k=1}^{n}\langle f(e_i)f_i,x\rangle$$

得到

$$f=\sum_{k=1}^{n}f(e_i)f_i$$

现取 $z\in\mathscr{X}^{**}$, 有

$$\langle z,f\rangle=\langle z,\sum_{k=1}^{n}f(e_i)f_i\rangle=\sum_{k=1}^{n}f(e_i)\langle z,f_i\rangle=f(\sum_{k=1}^{n}\langle z,f_i\rangle e_i)$$

故若取 $x_z=\sum\limits_{k=1}^{n}\langle z,f_i\rangle e_i\in\mathscr{X}$, 则

$$\langle z,f\rangle=\langle f,x_z\rangle$$

从而 $z\in\mathscr{X}$, 得到 $\mathscr{X}=\mathscr{X}^{**}$, 命题得证.

练习 3.51 求证: B 空间是自反的, 当且仅当它的共轭空间是自反的.

证明

当 B 空间 $\mathscr{X}$ 自反, 根据定义知对任意固定的 $z\in\mathscr{X}^{**}$, 都存在 $x_z\in\mathscr{X}$ 使得

$$\langle z,f\rangle=\langle f,x_z\rangle,\quad \forall f\in\mathscr{X}^*$$

希望证明对任意固定的 $g\in\mathscr{X}^{***}$, 都存在 $f_g\in\mathscr{X}^*$ 使得

$$\langle g,z\rangle=\langle z,f_g\rangle,\quad \forall z\in\mathscr{X}^{**}$$

设 $z=Tx_z$, 此时 $T\in\mathscr{L}(\mathscr{X},\mathscr{X}^{**})$, 故 $T^*\in\mathscr{L}(\mathscr{X}^{***},\mathscr{X}^*)$, 故考察 $f_g=T^*g$, 下面证明

$$\langle g,z\rangle=\langle z,T^*g\rangle,\quad \forall z\in\mathscr{X}^{**}\tag{3.84}$$

知

$$\langle g,z\rangle=\langle g,Tx_z\rangle=\langle T^*g,x_z\rangle=\langle z,T^*g\rangle$$

(3.84)式得证, 从而 $\mathscr{X}^*=\mathscr{X}^{***}$.

当 $\mathscr{X}^*$ 自反, 由前述结论知 $\mathscr{X}^{**}$ 自反, 故 $\mathscr{X}\subset\mathscr{X}^{**}$ 作为闭子空间也自反, 命题得证.

练习 3.52 设 $\mathscr{X}$ 是 B^* 空间, T 是从 $\mathscr{X}$ 到 $\mathscr{X}^{**}$ 的自然映射, 求证: $R(T)$ 是闭的充要条件是 $\mathscr{X}$ 是完备的.

证明

当 $\mathscr{X}$ 完备, 根据定义知任取 $\mathscr{X}$ 中的 Cauchy 列 $\{x_n\}$, 其都有极限 $x\in\mathscr{X}$. 现在任取 $R(T)$ 中的收敛列 $\{Tx_n\}$, 设 $Tx_n\to y$, 希望证明存在 $x\in\mathscr{X}$ 使得 $y=Tx$. 根据 T 的连续性知

$$\lim_{n\to\infty}Tx_n=T(\lim_{n\to\infty}x_n)=Tx$$

故 $y=Tx$, 因而 $R(T)$ 是闭的.

当 $R(T)$ 是闭的, 知若 $R(T)\ni Tx_n\to y$, 则存在 $x\in\mathscr{X}$ 使得 $y=Tx$. 现在对 $\mathscr{X}$ 中的 Cauchy 列 $\{x_n\}$, 知 $\{Tx_n\}$

必是 $R(T)$ 中的 Cauchy 列, 因而有极限 $y = Tx \in R(T)$, 现在证明 $x_n \to x$, 这是因为

$$T(x_n - x) \to y - Tx = \theta, \quad n \to \infty$$

这说明 $x_n - x \to \theta, n \to \infty$, 也即 $x_n \to x$, 因而 $\mathscr{X}$ 是完备的.

✍ **练习 3.53★** 在 l^1 中定义算子

$$T : (x_1, x_2, \cdots, x_n, \cdots) \mapsto (0, x_1, x_2, \cdots, x_n, \cdots)$$

求证: $T \in \mathscr{L}(l^1)$ 并求 T^*.

证明

T 显然是线性的, 下面说明 T 连续. 知

$$\|Tx\| = \sum_{k=1}^{\infty} |x_k| = \|x\|, \quad \forall x \in l^1$$

故

$$\|T\| = 1 < \infty$$

从而 T 连续, 因而 $T \in \mathscr{L}(l^1)$.

根据 T^* 的定义知

$$\langle f, Tx \rangle = \langle T^* f, x \rangle, \quad \forall f \in (l^1)^*, \forall x \in l^1$$

设 $l^1 \ni x = \sum\limits_{k=1}^{\infty} x_k e_k, e_k = (\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots)$, 则若取 $x_0 = 0$, 有

$$\langle f, Tx \rangle = \langle f, \sum_{k=1}^{\infty} x_{k-1} e_k \rangle = \sum_{k=1}^{\infty} x_{k-1} \langle f, e_k \rangle$$

现在希望

$$\langle T^* f, x \rangle = \sum_{k=1}^{\infty} x_k \langle T^* f, e_k \rangle = \sum_{k=1}^{\infty} x_{k-1} \langle f, e_k \rangle$$

比对得

$$\langle T^* f, e_k \rangle = \langle f, e_{k+1} \rangle, \quad k \in \mathbb{N}$$

已经知道 $(l^1)^* = l^\infty$, 故 $T^* \in \mathscr{L}(l^\infty)$. 由上式可知

$$T^* : l^\infty \to l^\infty, (y_1, \cdots, y_n, \cdots) \mapsto (y_2, \cdots, y_n, \cdots)$$

✍ **练习 3.54★** 在 l^2 中定义算子

$$T : (x_1, x_2, \cdots, x_n) \mapsto (x_1, \frac{x_2}{2}, \cdots, \frac{x_n}{n}, \cdots)$$

求证: $T \in \mathscr{L}(l^2)$ 并求 T^*.

证明

T 显然是线性的, 下面证明 T 连续, 知

$$\|Tx\|_{l^2} = (\sum_{k=1}^{\infty} \frac{x_k^2}{k^2})^{\frac{1}{2}} \le (\sum_{k=1}^{\infty} x_k^2)^{\frac{1}{2}} \cdot (\sum_{k=1}^{\infty} \frac{1}{k^2})^{\frac{1}{2}} = \frac{\pi}{\sqrt{6}} \|x\|_{l^2}, \quad \forall x \in l^2$$

得到

$$\|T\| \le \frac{\pi}{\sqrt{6}} < \infty$$

故 $T \in \mathscr{L}(l^2)$. 注意到 $(l^2)^* = l^2$, 现在欲求 $T^* \in \mathscr{L}(l^2)$ 满足

$$\langle f, Tx \rangle = \langle T^* f, x \rangle, \quad \forall f \in (l^2)^*$$

设 $l^2 \ni x = \sum\limits_{k=1}^{\infty} x_k e_k, e_k = (\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots)$, 则

$$\langle f, Tx\rangle = \langle f, \sum_{k=1}^{\infty} \frac{x_k}{k} e_k\rangle = \sum_{k=1}^{\infty} \frac{x_k}{k} \langle f, e_k\rangle$$

现在希望

$$\langle T^* f, x\rangle = \sum_{k=1}^{\infty} x_k \langle T^* f, e_k\rangle = \sum_{k=1}^{\infty} \frac{x_k}{k} \langle f, e_k\rangle$$

比对知

$$\langle T^* f, e_k\rangle = \frac{1}{k}\langle f, e_k\rangle, \quad k \in \mathbb{N}$$

注意 $f \in (l^2)^* = l^2$, 故有

$$T^* : l^2 \to l^2, (y_1, y_2, \cdots, y_n, \cdots) \mapsto (y_1, \frac{y_2}{2}, \cdots, \frac{y_n}{n}, \cdots)$$

练习 3.55★ 设 H 是 Hilbert 空间, $A \in \mathscr{L}(H)$ 并满足

$$(Ax, y) = (x, Ay) \quad (\forall x, y \in H)$$

求证:

(1) $A^* = A$.

证明

根据共轭算子的定义知

$$(Ax, y) = (x, A^* y), \quad \forall x, y \in H$$

又因为

$$(Ax, y) = (x, Ay), \quad \forall x, y \in H$$

故

$$\forall x, y \in H((x, A^* y - Ay) = 0) \Rightarrow \forall y \in H(A^* y = Ay) \Rightarrow A^* = A$$

(2) 若 $R(A)$ 在 H 中稠密, 则方程 $Ax = y$ 对 $\forall y \in R(A)$ 存在唯一解.

证明

既然 $y \in R(A)$, 故 $y = Ax$ 必有解, 现在证明解唯一. 设 $x' \in H$ 也满足 $y = Ax'$, 则知 $A(x - x') = \theta$, 现在证明 $N(A) = \{\theta\}$. 因为 $R(A)$ 在 H 中稠密, 故 $\overline{R(A)}^{\perp} = \{\theta\}$. 而任取 $x \in N(A)$, 根据定义有:

$$Ax = 0 \Rightarrow \forall y \in H((Ax, y) = 0 = (x, Ay)) \Rightarrow x \in \overline{R(A)}^{\perp}$$

故 $N(A) \subset \overline{R(A)}^{\perp} = \{\theta\}$, 也即 $N(A) = \{\theta\}$, 故 $x = x'$, 方程解唯一.

练习 3.56 设 $\mathscr{X}, \mathscr{Y}$ 是 B^* 空间, $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, 又设 A^{-1} 存在且 $A^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$, 求证:

(1) $(A^*)^{-1}$ 存在, 且 $(A^*)^{-1} \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*)$.

证明

既然 $T : A \mapsto A^*$ 是 $\mathscr{L}(\mathscr{X}, \mathscr{Y})$ 到 $\mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$ 的等距同构, 故 A 可逆可推出 A^* 可逆, 且 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 可推出 $A^* \in \mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$. 因为同构保持双射, 且由 $A^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$ 知 A 是双射, 故 A^* 是双射. 又因为 $\mathscr{X}^*, \mathscr{Y}^*$ 都是 B 空间, 故根据 Banach 定理知 $(A^*)^{-1} \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*)$.

(2) $(A^*)^{-1} = (A^{-1})^*$.

证明

首先知 $(A^*)^{-1} \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*), (A^{-1})^* \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*)$. 现在任取 $f \in \mathscr{X}^*, y \in \mathscr{Y}$, 设 $y = Ax$, 根据共轭算子的定

义有:

$$\langle (A^{-1})^* f, y\rangle = \langle f, A^{-1}y\rangle = \langle f, x\rangle$$

与此同时:

$$\langle (A^*)^{-1} f, y\rangle = \langle (A^*)^{-1} f, Ax\rangle = \langle (A^*)^{-1}A^* f, x\rangle = \langle f, x\rangle$$

这说明

$$\forall y \in \mathscr{Y} \forall f \in \mathscr{X}^*(\langle (A^{-1})^* f - (A^*)^{-1} f, y\rangle = 0) \Rightarrow \forall f \in \mathscr{X}^*((A^{-1})^* f = (A^*)^{-1} f) \Rightarrow (A^{-1})^* = (A^*)^{-1}.$$

练习 3.57 设 $\mathscr{X},\mathscr{Y},\mathscr{Z}$ 是 B^* 空间, 而 $B \in \mathscr{L}(\mathscr{X},\mathscr{Y})$ 以及 $A \in \mathscr{L}(\mathscr{Y},\mathscr{Z})$, 求证: $(AB)^* = B^*A^*$.

证明

知 $AB \in \mathscr{L}(\mathscr{X},\mathscr{Z})$, 因而 $(AB)^* \in \mathscr{L}(\mathscr{Z}^*,\mathscr{X}^*)$. 任取 $h \in \mathscr{Z}^*, x \in \mathscr{X}^*$, 设 $y = Bx, z = Ay$, 根据共轭算子的定义知

$$\langle (AB)^* h, x\rangle = \langle h, ABx\rangle = \langle h, z\rangle$$

另一方面知

$$\langle B^*A^* h, x\rangle = \langle A^* h, Bx\rangle = \langle h, ABx\rangle = \langle h, z\rangle$$

这说明

$$\forall x \in \mathscr{X} \forall h \in \mathscr{X}^*(\langle (AB)^* h - B^*A^* h, x\rangle = 0) \Rightarrow \forall h \in \mathscr{X}^*((AB)^* h = B^*A^* h) \Rightarrow (AB)^* = B^*A^*.$$

练习 3.58★♡ 设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, T 是 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性算子, 又设对 $\forall g \in \mathscr{Y}^*$, $g(Tx)$ 是 $\mathscr{X}$ 上的有界线性泛函, 求证: T 是连续的.

† 错误的证明其一

既然 $* : T \mapsto T^*$ 是等距同构, 故 $*$ 保连续, 也即 T 连续当且仅当 T^* 连续. 根据共轭算子的定义知 $g(Tx) = (T^*g)(x)$, 因而 $\forall g \in \mathscr{Y}^*(T^*g \in \mathscr{X}^*)$, 根据有界性知

$$\|T^*g\|_{\mathscr{X}^*} \leq C, \quad \forall g \in \mathscr{Y}^*$$

特别有

$$\|T^*g\|_{\mathscr{X}^*} \leq C, \quad \|g\|_{\mathscr{Y}^*} = 1$$

故

$$\sup_{\|g\|=1} \|T^*g\|_{\mathscr{X}^*} \leq C$$

此即 $\|T^*\| \leq C$, 故 T^* 连续, 因而 T 连续.

注 上面这个证明的错误原因在于共轭算子是基于连续线性算子来定义的, 只要谈一个算子的共轭算子, 就是暗示这个算子是连续线性算子.

† 错误的证明其二

设 $\|x_n - x\|_{\mathscr{X}} \to 0(n \to \infty)$, 往证 $\|Tx_n - Tx\|_{\mathscr{X}} \to 0(n \to \infty)$. 因为对任意 $g \in \mathscr{Y}^*$, $g(T\circledast)$ 均为 $\mathscr{X}$ 上的有界线性泛函, 故

$$\|x_n - x\|_{\mathscr{X}} \to 0(n \to \infty) \Rightarrow |g(Tx_n) - g(Tx)| \to 0(n \to \infty)$$

根据极限定义即

$$\forall g \in \mathscr{Y}^* \forall \varepsilon > 0 \exists N > 0 \forall n > N(|g(Tx_n - Tx)| < \varepsilon)$$

对每个固定的 n, 根据 Hahn-Banach 定理知存在 $g_n \in \mathscr{Y}^*$ 使得 $|g_n(Tx_n - Tx)| = \|Tx_n - Tx\|_{\mathscr{Y}}$, 于是

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N(\|Tx_n - Tx\|_{\mathscr{Y}} < \varepsilon) \Leftrightarrow \|Tx_n - Tx\|_{\mathscr{Y}} \to 0(n \to \infty)$$

此即欲证.

注 上面这个证明的错误原因在于默认 ε 是不随 g 的变动而变动的. 事实上从条件出发并不能得到任何这类" 一致性" 的结果, 故当在变动 g 为 g_n 时, ε 事实上要做相应改动, 而这个改动也没有办法去刻画, 从而这个方法从源头上就是行不通的.

该题的困难点在于书上并没有介绍太多证明算子连续的方法. 一般来讲, 要证明算子连续只能通过闭图像定理或定义本身, 前者需要验证算子是闭算子, 后者可以通过证明算子有界线性. **[Lin]** 上给出的两个答案分别对应了这两种方法.

正确的证明其一 (利用闭图像定理)

既然 $\mathscr{X}$ 完备, 只需证明 T 是闭算子即可. 设 $\mathscr{X}$ 中的基本列 $\{x_n\}_{n\in\mathbb{N}}$ 满足 $x_n \to x_0, Tx_n \to y_0(n \to \infty)$, 只需说明 $y_0 = Tx_0$ 即可. 因为 $\forall g \in \mathscr{Y}^*(g(T\circledast) \in \mathscr{X}^*)$, 故

$$\lim_{n\to\infty} g(Tx_n) = g(Tx), \quad \forall g \in \mathscr{Y}^*$$

这说明 $Tx_n \rightharpoonup Tx$, 又因为 $Tx_n \to y(n \to \infty)$, 故必有 $Tx_n \rightharpoonup y(n \to \infty)$, 根据弱极限的唯一性即知 $y = Tx$, 从而 T 是闭算子, 进而根据闭图像定理知 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. $\square$

正确的证明其二 (利用有界线性性)

设 $U : \mathscr{Y} \to \mathscr{Y}^{**}$ 是自然映射. 对任意 $x \in \mathscr{X}, Tx \in \mathscr{Y}$, 令 $F_T(x) = U(Tx)$, 于是对任意 $g \in \mathscr{Y}^*$ 与 $x \in \mathscr{X}$ 有:

$$\begin{cases} \langle F_T(x), g\rangle = \langle g, Tx\rangle \\ \|F_T(x)\|_{\mathscr{Y}^{**}} = \|Tx\|_{\mathscr{Y}} \end{cases}$$

现在要说明 T 的有界性, 就是说明 $\sup\limits_{\|x\|_{\mathscr{X}}=1} \|Tx\|_{\mathscr{Y}}$ 的有界性, 亦即说明 $\sup\limits_{\|x\|_{\mathscr{X}}=1} \|F_T(x)\|_{\mathscr{Y}^{**}}$ 的有界性. 任取 $x \in \mathscr{X}$ 满足 $\|x\|_{\mathscr{X}} = 1$, 因为 $g(T\circledast) \in \mathscr{X}^*$, 故

$$\sup_{\|x\|_{\mathscr{X}}=1} |\langle F_T(x), g\rangle| = \sup_{\|x\|_{\mathscr{X}}=1} |\langle g, Tx\rangle| = \sup_{\|x\|_{\mathscr{X}}=1} |\langle g(T\circledast), x\rangle| < \infty, \quad \forall g \in \mathscr{Y}^*$$

而 $\mathscr{Y}^*$ 是 B 空间, $F_T(x)$ 是 $\mathscr{Y}^*$ 上的有界线性泛函, 故根据共鸣定理3.4.8知存在 $M > 0$ 使得 $\sup\limits_{\|x\|_{\mathscr{X}}=1} \|F_T(x)\|_{\mathscr{Y}^{**}} \le M$, 亦即

$$\|T\| = \sup_{\|x\|_{\mathscr{X}}=1} \|Tx\|_{\mathscr{Y}} = \sup_{\|x\|_{\mathscr{X}}=1} \|F_T(x)\|_{\mathscr{Y}^{**}} \le M$$

故 T 是有界线性算子, 因而其为连续线性算子. $\square$

练习 3.59 设 $\{x_n\} \subset C[a,b], x \in C[a,b]$ 且 $x_n \rightharpoonup x(n \to \infty)$, 求证:

$$\lim_{n\to\infty} x_n(t) = x(t) \quad (\forall t \in [a,b])$$

也即 x_n 点态收敛到 x.

证明

设 $f_t(x) = x(t), x \in C[a,b], t \in [a,b]$, 下面证明 $f_t \in C[a,b]^*$. 线性性显然, 在有界性上知

$$|f_t(x)| = |x(t)| \le \max_{t\in[a,b]} |x(t)| = \|x\|_{C[a,b]} \Rightarrow \|f_t\|_{C[a,b]^*} \le 1$$

故 $f_t \in C[a,b]^*$. 根据弱收敛的定义有

$$\lim_{n\to\infty} f(x_n) = f(x), \quad \forall f \in C[a,b]^*$$

代入 f_t 即得

$$\lim_{n\to\infty} x_n(t) = x(t), \quad \forall t \in [a,b].$$

练习 3.60 已知在 B^* 空间中 $x_n \rightharpoonup x_0(n \to \infty)$, 求证:

$$\varliminf_{n\to\infty} \|x_n\| \geq \|x_0\|.$$

证明

设 $x_n, x_0 \in \mathscr{X}, n = 1, 2, \cdots$, 根据弱收敛的定义知

$$\lim_{n\to\infty} f(x_n) = f(x_0), \quad \forall f \in \mathscr{X}^*$$

现设 $y_n = Tx_n \in \mathscr{X}^{**}, y_0 = Tx_0 \in \mathscr{X}^{**}$, 知 $\|y_n\|_{\mathscr{X}^{**}} = \|x_n\|_{\mathscr{X}}, \|y_0\|_{\mathscr{X}^{**}} = \|x_0\|_{\mathscr{X}}$, 且有

$$\langle y_n, f\rangle = \langle f, x_n\rangle, \langle y_0, f\rangle = \langle f, x_0\rangle, \quad \forall f \in \mathscr{X}^*$$

得到

$$|\langle y_0, f\rangle| = |\lim_{n\to\infty} \langle y_n, f\rangle| \leq \lim_{n\to\infty} \|y_n\|_{\mathscr{X}^{**}} \cdot \|f\|_{\mathscr{X}^*}, \quad \forall f \in \mathscr{X}^*$$

这说明

$$|\langle y_0, f\rangle| \leq \varliminf_{n\to\infty} \|y_n\|_{\mathscr{X}^{**}} \cdot \|f\|_{\mathscr{X}^*}, \quad \forall f \in \mathscr{X}^*$$

此即

$$\|y_0\|_{\mathscr{X}^{**}} \leq \varliminf_{n\to\infty} \|y_n\|_{\mathscr{X}^{**}}$$

也即

$$\varliminf_{n\to\infty} \|x_n\| \geq \|x_0\|$$

练习 3.61★ 设 H 是 Hilbert 空间, $\{e_n\}$ 是 H 的正交规范基, 求证: 在 H 中 $x_n \rightharpoonup x_0(n \to \infty)$ 的充要条件是

1. $\|x_n\|$ 有界;
2. $(x_n, e_k) \to (x_0, e_k)(n \to \infty)(k = 1, 2, \cdots)$.

证明

当 $x_n \rightharpoonup x_0(n \to \infty)$, 注意到 H 是 Hilbert 空间, 因而根据弱收敛的定义与 Riesz 表示定理知

$$\lim_{n\to\infty} (x_n, y) = (x_0, y), \quad \forall y \in H$$

由 Banach-Steinhaus 定理定理即得 $\|x_n\|$ 有界. 另取 $y = e_k, k = 1, 2, \cdots$ 即得 $(x_n, e_k) \to (x_0, e_k)(n \to \infty)(k = 1, 2, \cdots)$.

当题设的两个条件满足, 注意 $\{e_n\}$ 是 H 的正交规范基, 进而 H 可分, 且 $H_0 := \operatorname{span}\{e_n\}_{n=1}^{\infty}$ 是 H 的一个可数稠密子集. 知任取 $y \in H_0$, 由条件 2. 可得

$$(x_n, y) \to (x_0, y), \quad n \to \infty$$

进而由 Banach-Steinhaus 定理即得

$$(x_n, y) \to (x_0, y), \quad n \to \infty, \forall y \in H$$

这正是 $x_n \rightharpoonup x_0$ 的定义.

练习 3.62 设 S_n 是 $L^p(\mathbb{R})(1 \le p < \infty)$ 到自身的算子:

$$(S_n u)(x) = \begin{cases} u(x), & |x| \le n \\ 0, & |x| > n \end{cases}$$

其中 $u \in L^p(\mathbb{R})$ 是任意的, 求证: $\{S_n\}$ 强收敛于恒同算子 I, 但不一致收敛到 I.

证明

知任取 $u \in L^p(\mathbb{R})$, 有

$$(S_n u - u)(x) = \begin{cases} 0, & |x| \le n \\ -u(x), & |x| > n \end{cases}$$

进而当 $n \to \infty$, 有 $S_n u - u \overset{a.e.}{\to} 0$, 也即 $\|(S_n - I)u\| \to 0 (\forall u \in L^p(\mathbb{R}), n \to \infty)$. 这正是强收敛的定义.

对于一致收敛, 注意

$$\|S_n - I\| = \sup_{\|u\|_{L^p(\mathbb{R})}=1} \|S_n u - u\|_{L^p(\mathbb{R})} = \sup_{\|u\|_{L^p(\mathbb{R})}=1} (\int_{(-\infty,-n]\cup[n,\infty)} |u(x)|^p dx)^{\frac{1}{p}}$$

取

$$u_n(x) = \begin{cases} 0, & x \in [-n, n] \\ (\frac{1}{\pi((|x|-n)^2+1)})^{\frac{1}{p}}, & x \in (-\infty, -n) \cup (n, \infty) \end{cases}$$

知

$$\|u\|_{L^p(\mathbb{R})} = 1 = (\int_{(-\infty,-n]\cup[n,\infty)} |u(x)|^p dx)^{\frac{1}{p}}$$

故 $\|S_n - I\| \nrightarrow 0$, 也即 S_n 并不一致收敛到 I.

练习 3.63★ 设 H 是 Hilbert 空间, 在 H 中 $x_n \rightharpoonup x_0 (n \to \infty)$, 而且 $y_n \to y_0 (n \to \infty)$, 求证: $(x_n, y_n) \to (x_0, y_0)(n \to \infty)$.

证明

根据弱收敛的定义与 Riesz 表示定理有

$$(x_n, y) \to (x_0, y), \quad \forall y \in H, n \to \infty$$

从 $y_n \to y_0 (n \to \infty)$ 知

$$\|y_n - y\| \to 0, \quad n \to \infty$$

进而

$$\begin{aligned} |(x_n, y_n) - (x_0, y_0)| &= |(x_n, y_n) - (x_0, y_n) + (x_0, y_n) - (x_0, y_0)| \\ &\le |(x_n, y_n) - (x_0, y_n)| + |(x_0, y_n) - (x_0, y_0)| \\ &\le |(x_n, y_n) - (x_0, y_n)| + \|x_0\| \cdot \|y_n - y_0\| \to 0, \quad n \to \infty \end{aligned}$$

此即 $(x_n, y_n) \to (x_0, y_0)(n \to \infty)$.

练习 3.64 设 $\{e_n\}$ 是 Hilbert 空间 H 中的正交规范基[25], 求证: 在 H 中 $e_n \rightharpoonup \theta(n \to \infty)$, 但 $e_n \nrightarrow \theta(n \to \infty)$.

证明

既然 $\{e_n\}$ 是正交规范基, 设 $H \ni x = \sum_{k=1}^{\infty} x_k e_k$, 进而由级数收敛与规范性知 $\|x_n e_n\| = |x_n| \to 0 (n \to \infty)$. 由 Riesz 表示定理与正交性知

$$\langle e_n, x \rangle = x_n \to 0, \quad n \to \infty$$

这说明 $e_n \rightharpoonup \theta(n \to \infty)$. 但 $\|e_n\| \equiv 1 \ne 0$, 故 $e_n \nrightarrow \theta(n \to \infty)$.

[25] **[ZL]** 此处有误: " 正交规范集".

练习 3.65★♡ 设 H 是 Hilbert 空间, 求证: 在 H 中 $x_n \to x(n \to \infty)$ 的充要条件是

1. $\|x_n\| \to \|x\|(n \to \infty)$;
2. $x_n \rightharpoonup x(n \to \infty)$.

† 错误的证明

当 $x_n \to x(n \to \infty)$, 因为强极限存在必为弱极限, 故 $x_n \rightharpoonup x(n \to \infty)$, 而范数收敛恰是强极限的定义.

当 $\|x_n\| \to \|x\|$ 且 $x_n \rightharpoonup x$, 根据弱收敛的定义与 Riesz 表示定理有

$$\lim_{n\to\infty} (x_n, y) = (x, y), \quad \forall y \in H$$

代入 $y = x_n$ 有

$$\lim_{n\to\infty} (x_n, x_n) = \lim_{n\to\infty} (x, x_n)$$

代入 $y = x$ 有

$$\lim_{n\to\infty} (x_n, x) = (x, x)$$

故

$$\lim_{n\to\infty} (x_n, x_n) = \lim_{n\to\infty} \|x_n\|^2 = \|x\|^2 = (x, x)$$

此即

$$\|x_n\| \to \|x\|, \quad n \to \infty$$

也即 $x_n \to x(n \to \infty)$.

注 上述证明的问题在于代入 $y = x_n$ 这一步, 这是不符合极限过程的.

正确的证明

$x_n \to x(n \to \infty)$ 时的情况是显然的, 下设 (1), (2) 成立. 根据弱收敛的定义有

$$\lim_{n\to\infty} f(x_n) = f(x), \quad \forall f \in H^*$$

因为 H 是 Hilbert 空间, 根据 Riesz 表示定理, 对固定的 x 可取 $f_x(y) = (y, x) \in H^*$, 故

$$\lim_{n\to\infty} (x_n, x) = (x, x) = \|x\|^2$$

根据内积的对称性有

$$\lim_{n\to\infty} (x, x_n) = \|x\|^2$$

于是

$$\lim_{n\to\infty} \|x_n - x\|^2 = (x_n - x, x_n - x) = \|x_n\|^2 - (x_n, x) - (x, x_n) + \|x\|^2 = 0$$

故 $x_n \to x(n \to \infty)$, 命题即证. □

注 本题的结论实际上对一致凸的 B^* 空间都适用. 参见 **[PGC]** 定理 5.12-3.

练习 3.66★ 求证: 在自反的 B 空间中, 集合的弱列紧性与有界性是等价的.

证明

从 Eberlein-Smulian 定理的证明中知自反空间的任意有界集必弱列紧, 因而证明弱列紧集必有界即可. 用反证法, 设 $A \subset \mathscr{X}$ 无界, 则存在 $\{x_n\} \subset A$ 满足 $\|x_n\| > n(n \to \infty)$. 既然 A 弱列紧, 知 $\{x_n\}$ 存在弱收敛的子列 $\{x_{n_k}\}$, 设 $x_{n_k} \rightharpoonup x(k \to \infty)$. 现设 T 是 $\mathscr{X}$ 到 $\mathscr{X}^{**}$ 的自然映射, 由 $\mathscr{X}$ 的自反性知

$$\langle f, x_n \rangle = \langle Tx_n, f \rangle, \langle f, x \rangle = \langle Tx, f \rangle, \quad \forall f \in \mathscr{X}^*$$

进而由弱收敛知

$$\lim_{k\to\infty} \langle Tx_{n_k}, f \rangle = \langle Tx, f \rangle, \quad \forall f \in \mathscr{X}^*$$

既然 x 有意义, 知

$$\sup_{k\geq 1}|\langle Tx_{n_k}, f\rangle| < \infty, \quad \forall f \in \mathscr{X}^*$$

故由共鸣定理知存在 $M > 0$ 使得 $\|Tx_{n_k}\| \leq M(k = 1, 2, \cdots)$. 但 $\|Tx_{n_k}\| = \|x_{n_k}\| \to \infty$, 矛盾! 命题即证.

✍ **练习 3.67** 求证: B^* 空间中的闭凸集是弱闭的, 即若 M 是闭凸集, $\{x_n\} \subset M$, 且 $x_n \rightharpoonup x_0(n \to \infty)$, 则 $x_0 \in M$.

证明

由 Mazur 定理, 既然 M 是凸集, 且 $x_n \rightharpoonup x_0$, 则在 M 中存在序列 $\{y_n\}$ 满足 $y_n \to x_0(n \to \infty)$. 又因为 M 是闭的, 故 $x_0 \in M$.

✍ **练习 3.68★** 设 $\mathscr{X}$ 是自反的 B 空间, M 是 $\mathscr{X}$ 中的有界闭凸集, $\forall f \in \mathscr{X}^*$, 求证: f 在 M 上达到最大值和最小值.

证明

设 $L = \sup\limits_{x\in M}\langle f, x\rangle$, 根据上确界的定义知

$$\forall n \in \mathbb{N}\exists x_n \in M(\langle f, x_n\rangle > L - \frac{1}{n})$$

既然 M 有界, 知 M 是弱列紧的, 又由练习3.67知 M 是弱闭的. 现在 $\{x_n\}$ 中取弱收敛子列 $\{x_{n_k}\}$, 设 $x_{n_k} \rightharpoonup x_0 \in M$, 有

$$L \leq \lim_{k\to\infty}\langle f, x_{n_k}\rangle = \langle f, x_0\rangle \leq L$$

故 $\langle f, x_0\rangle = L$, 也即 f 在 M 上达到最大值. 最小值同理.

✍ **练习 3.69** 设 $\mathscr{X}$ 是自反的 B 空间, M 是 $\mathscr{X}$ 中的非空闭凸集, 求证: $\exists x_0 \in M$, 使得 $\|x_0\| = \inf\{\|x\| : x \in M\}$.

证明

设 $l = \inf\{\|x\| : x \in M\}$, 根据下确界的定义知

$$\forall n \in \mathbb{N}\exists x_n \in M(\|x_n\| < l + \frac{1}{n})$$

根据构造知 $\{x_n\}$ 有界, 因而 $M_0 := \overline{\mathrm{co}\{x_n\}} \subset M$ 是有界闭凸集, 故其是弱列紧且弱闭的. 取弱收敛子列 $\{x_{n_k}\}$, 设 $x_{n_k} \rightharpoonup x_0 \in M_0$, 现在希望说明 $\|x_0\| = l$. 由 Hahn-Banach 定理的推论知存在 $f \in \mathscr{X}^*$ 使得

$$\|f\| = 1, \quad f(x_0) = \|x_0\|$$

进而由弱收敛知

$$l \leq \lim_{k\to\infty}\langle f, x_{n_k}\rangle = \langle f, x_0\rangle \leq l$$

此即 $\|x_0\| = l$, 命题得证.

3.6.3 补充: 弱拓扑与弱 * 拓扑

本节选自 [**HB**], 旨在从拓扑的角度出发去理解弱收敛, 弱 * 收敛及相关的紧性.

3.6.3.1 原始的想法: 使一族映射连续的最粗糙的拓扑

从一般的拓扑概念出发, 设 X 是一个集合, 而 $\{Y_i\}_{i\in I}$ 是一族拓扑空间, 对每个 $i \in I$, 给定一个映射 $\varphi_i : X \to Y_i$. 现在希望知道在 X 上是否可以定义一个拓扑, 使得所有的映射 $\{\varphi_i\}_{i\in I}$ 都是连续映射. 如果这样的拓扑存在, 又能否构造出最粗糙 (也就是包含最少开集) 的拓扑 τ 使得 $\{\varphi_i\}_{i\in I}$ 连续? 当 X 具有离散拓扑 (即 X 的所有子集都是开集) 时, 显然每个 φ_i 都是连续映射, 因为每个开集的原像必定是开集. 但这个拓扑并不粗糙, 实际上它是最不粗糙的. 从连续映射的等价形式出发, 设 $\Omega_i \subset Y_i$ 是开集, 则 $\varphi_i^{-1}(\Omega_i)$ 在拓扑 τ 下必须是开集. 在 Ω_i 取遍 Y_i 的开集族, i 跑遍 I 时, $\{\varphi_i^{-1}(\Omega_i)\}$ 就组成拓扑 τ 中的开集形成的 X 中的一个子集族, 记该子集族为 $\{U_\lambda\}_{\lambda\in\Lambda}$, 则要谈 τ 是最粗糙的, 就是要谈 τ 是使得 $\{U_\lambda\}_{\lambda\in\Lambda}$ 全为开集的最弱的拓扑. 这个问题可以归结为:

问题构造 X 上的子集族 $\mathcal{F}$, 使得 $\mathcal{F}$ 尽可能包含少的集合, 它在这些集合的有限交合任意并下封闭, 且 $\forall\lambda\in\Lambda(U_\lambda\in\mathcal{F})$.

这个问题的答案如下: 首先考虑 $\mathcal{F}$ 包含有限交, 也就是 $\bigcap\limits_{\lambda\in\Gamma}U_\lambda,\Gamma\subset\Lambda$ 是有限集. 取遍全体可能的有限交可以得到 X 的一个子集族 Φ, 显见它在 $\{U_\lambda\}_{\lambda\in\Lambda}$ 的有限交运算下是封闭的. 再对 Φ 中的元素取任意并, 得到新的子集族 $\mathcal{F}$, 显见 $\mathcal{F}$ 对任意并运算是封闭的, 但此时 $\mathcal{F}$ 在有限交运算下的封闭性就不再显而易见了. 这由下面的引理给出:

引理 3.6.3

子集族 $\mathcal{F}$ 在有限交运算下封闭.

证明从略. 从这个思路出发可以刻画拓扑 τ: 首先取形如 $\varphi_i^{-1}(\Omega_i)$ 的集合的有限交, 其中 Ω_i 是 Y_i 的开集, 对得到的子集族再取任意并, 这样得到的子集族就是 τ 中的开集族. 给定点 $x\in X$, 就可以谈 x 在拓扑 τ 中的邻域基, 它们由形如 $\varphi_i^{-1}(V_i)$ 的有限交组成, 其中 V_i 是 $\varphi_i(x_i)$ 在 Y_i 中的邻域.

下面对 X 赋拓扑 τ, 先讨论它的一些基本性质.

命题 3.6.3

设 $\{x_n\}_{n\in\mathbb{N}}$ 是 X 中的点列, 则 $x_n\to x$ 当且仅当对任意 $i\in I$ 均有 $\varphi_i(x_n)\to\varphi_i(x)$.

证明

若 $x_n\to x$, 因为 $\varphi_i(i\in I)$ 都是连续映射, 故 $\varphi_i(x_n)\to\varphi_i(x)$ 对任意 $i\in I$ 成立. 反之, 若 $\forall i\in I(\varphi_i(x_n)\to\varphi_i(x))$, 则设 U 是 x 的邻域, 根据前面对邻域基的刻画知总有:

$$U=\bigcap_{i\in J}\varphi_i^{-1}(V_i)$$

其中 $J\subset I$ 是有限集, V_i 是 $\varphi_i(x)$ 在 Y_i 中的邻域. 根据 $\forall i\in I(\varphi_i(x_n)\to\varphi_i(x))$ 的定义, 对每个 $i\in J$, 存在整数 N_i 使得使得 $\varphi_i(x_n)\in V_i$ 对任意 $n\ge N_i$ 都成立. 现取 $N=\max\limits_{i\in J}N_i$, 知 $\forall n\ge N(x_n\in U)$, 此即 $x_n\to x$. □

命题 3.6.4

设 Z 是拓扑空间, 设 ψ 是从 Z 到 X 的映射, 则 ψ 是连续映射当且仅当对任意 $i\in I,\varphi_i\circ\psi$ 是从 Z 到 Y_i 的连续映射.

证明

若 ψ 是连续的, 则根据 $\{\varphi_i\}_{i\in I}$ 的连续性显见对任意 $i\in I$ 而言, $\varphi_i\circ\psi$ 也是连续的. 相反, 设 U 是 X 的开集, 要证明 $\psi^{-1}(U)$ 是 Z 的开集. 已经知道 U 可以写成:

$$U=\bigcup_{j\in J}\bigcap_{i\in I_j}\varphi_i^{-1}(\Omega_i)$$

其中对每个 j 而言 I_j 均为有限集, Ω_i 是 Y_i 中的开集, 因此

$$\psi^{-1}(U)=\bigcup_{j\in J}\bigcap_{i\in I_j}\psi^{-1}(\varphi_i^{-1}(\Omega_i))=\bigcup_{j\in J}\bigcap_{i\in I_j}(\varphi_i\circ\psi)^{-1}(\Omega_i)$$

因为对任意 $i\in I,\varphi_i\circ\psi$ 都是连续映射, 故 $(\varphi_i\circ\psi)^{-1}(\Omega_i)$ 是开集, 从而 $\psi^{-1}(U)$ 是开集, 此即欲证. □

3.6.3.2 弱拓扑

设 E 是 Banach 空间且 $f\in E^*$, 记 $\varphi_f:E\to\mathbb{R}$ 是由 $\varphi_f(x)=\langle f,x\rangle$ 定义的映射. 当 f 跑遍 E^* 时, 可以得到一族从 E 到 $\mathbb{R}$ 的映射 $\{\varphi_f\}_{f\in E^*}$.

定义 3.6.9 (弱拓扑)

E 上的弱拓扑 $\sigma(E,E^*)$ 是 E 上使得所有映射 $\{\varphi_f\}_{f\in E^*}$ 都连续的最粗糙的拓扑.

注 回忆弱收敛的定义:

$$x_n \rightharpoonup x \Leftrightarrow \forall f \in E^*(f(x_n) \to f(x))$$

于是弱拓扑中" 所有映射" 就对应这里的 $\forall f \in E^*$," 都连续" 对应这里的 $f(x_n) \to f(x)$.

命题 3.6.5

弱拓扑 $\sigma(E, E^*)$ 是可分离的, 亦即任取 $x_1, x_2 \in E$, 若 $x_1 \neq x_2$, 就必存在弱拓扑 $\sigma(E, E^*)$ 中的开集 O_1, O_2, 使得 $x_1 \in O_1, x_2 \in O_2, O_1 \cap O_2 = \varnothing$.

证明

设 $x_1, x_2 \in E$, 根据 Hahn-Banach 定理 (第二几何形式) 知, 存在闭超平面严格分离 $\{x_1\}$ 与 $\{x_2\}$, 从而存在 $f \in E^*$ 与 $\alpha \in \mathbb{R}$ 使得

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

令

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

则 O_1, O_2 是 $\sigma(E, E^*)$ 中的开集, 且满足 $x_1 \in O_1, x_2 \in O_2, O_1 \cap O_2 = \varnothing$. □

命题 3.6.6

设 $x_0 \in E$, 则 x_0 在拓扑 $\sigma(E, E^*)$ 中的邻域基可由所有形如

$$V = \{x \in E : |\langle f_i, x - x_0 \rangle| < \varepsilon, \forall i \in I\}$$

的集合组成, 其中 I 是有限集, $f_i \in E^*, \varepsilon > 0$.

证明

显见 $V = \bigcap\limits_{i \in I} \varphi_{f_i}^{-1}((a_i - \varepsilon, a_i + \varepsilon))$ 是拓扑 $\sigma(E, E^*)$ 中包含 x_0 的开集, 其中 $a_i = \langle f_i, x_0 \rangle$. 现设 U 是拓扑 $\sigma(E, E^*)$ 中包含 x_0 的邻域, 知存在 x_0 的邻域 W 满足 $W \subset U$, 且 $W = \bigcap\limits_{i \in I} \varphi_{f_i}^{-1}(\Omega_i)$, 其中 I 有限, Ω_i 是 $\langle f_i, x_0 \rangle = a_i$ 在 $\mathbb{R}$ 中的邻域, 故存在 $\varepsilon > 0$ 使得对任意 $i \in I$ 都有 $(a_i - \varepsilon, a_i + \varepsilon) \subset \Omega_i$. 因此 $x_0 \in V \subset W \subset U$. □

从弱拓扑出发, 弱收敛可以定义为序列 $\{x_n\}_{n \in \mathbb{N}}$ 在弱拓扑 $\sigma(E, E^*)$ 下收敛于 x, 记作 $x_n \rightharpoonup x$. 弱收敛有下述 (已经知道的) 性质:

命题 3.6.7

设 $\{x_n\}_{n \in \mathbb{N}}$ 是 Banach 空间 E 中的序列, 则

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

(ii) $x_n \to x \Rightarrow x_n \rightharpoonup x$;

(iii) 若 $x_n \rightharpoonup x$, 则 $\sup\limits_{n \geq 1} \|x_n\| < \infty$, 且 $\|x\| \leq \varliminf\limits_{n \to \infty} \|x_n\|$;

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

3.6.3.3 弱 * 拓扑

设 E 是 Banach 空间, 对每个 $x \in E$, 考虑映射 $\varphi_x : E^* \to \mathbb{R}$ 定义为 $f \mapsto \varphi_x(f) = \langle f, x \rangle$. 当 x 取遍 E 时, 可以得到从 E^* 到 $\mathbb{R}$ 的映射族 $\{\varphi_x\}_{x \in E}$.

定义 3.6.10 (弱 * 拓扑)

弱 * 拓扑记为 $\sigma(E^*, E)$, 它是使得所有映射 $\{\varphi_x\}_{x \in E}$ 都连续的最粗糙的拓扑.

因为 $E \subset E^{**}$, 故拓扑 $\sigma(E^*, E)$ 比拓扑 $\sigma(E^*, E^{**})$ 更粗糙, 也就是说拓扑 $\sigma(E^*, E)$ 比拓扑 $\sigma(E^*, E^{**})$ 拥有更少的开集和闭集. 这样不懈追求更粗糙的拓扑的意义在于, 如果一个拓扑具有更少的开集, 它就会拥有更多的紧集, 而紧集在建立存在性定理方面起到一些基本的作用.

命题 3.6.8

弱 * 拓扑 $\sigma(E^*, E)$ 是可分离的.

证明

设 $f_1, f_2 \in E, f_1 \neq f_2$, 则存在 $x \in E$ 使得 $\langle f_1, x\rangle \neq \langle f_2, x\rangle$. 不妨设 $\langle f_1, x\rangle < \langle f_2, x\rangle$, 则可取 $\alpha \in \mathbb{R}$ 满足

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

令

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

则 O_1, O_2 是 $\sigma(E^*, E)$ 中的开集, 且满足 $f_1 \in O_1, f_2 \in O_2, O_1 \cap O_2 = \varnothing$. □

类似于命题3.6.6, 可以证明:

命题 3.6.9

对 $f_0 \in E^*$, f_0 在拓扑 $\sigma(E^*, E)$ 中的邻域基可由所有形如

$$V = \{f \in E^* : |\langle f - f_0, x_i\rangle| < \varepsilon, \forall i \in I\}$$

的集合组成, 其中 I 是有限集, $x_i \in E, \varepsilon > 0$.

从弱 * 拓扑出发, 弱 * 收敛可以定义为序列 $\{f_n\}_{n\in\mathbb{N}}$ 在弱 * 拓扑 $\sigma(E^*, E)$ 下收敛到 f, 记作 $f_n \overset{*}{\rightharpoonup} f$. 类似于命题3.6.7, 弱 * 拓扑有下述性质:

命题 3.6.10

设 $\{f_n\}_{n\in\mathbb{N}}$ 是 Banach 空间 E^* 中的序列, 则

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

(ii) $f_n \rightharpoonup f \Rightarrow f_n \overset{*}{\rightharpoonup} f$;

(iii) 若 $f_n \overset{*}{\rightharpoonup} f$, 则 $\sup\limits_{n\geq 1} \|f_n\| < \infty$, 且 $\|f\| \leq \varliminf\limits_{n\to\infty} \|f_n\|$;

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

3.7 线性算子的谱

3.7.1 知识梳理

线性算子的谱可以类比成矩阵的特征值, 它在求振动频率, 判定系统稳定性与刻画方程解的构造上都有广泛应用.

3.7.1.1 定义与例子

在维数不低于 1 的复 Banach 空间 $\mathscr{X}$ 上考察闭线性算子 $A : D(A) \subset \mathscr{X} \to \mathscr{X}$.

定义 3.7.1 (特征值, 特征元)

$\lambda \in \mathbb{C}$ 称为是 A 的特征值, 如果 $\exists x_0 \in D(A)\backslash\{\theta\}$ 满足:

$$Ax_0 = \lambda x_0$$

称相应的 x_0 为对应于 λ 的特征元.

♣

从线性代数中知, 当 $\dim \mathscr{X} < \infty$ 时, 对任意的 $\lambda \in \mathbb{C}$, 下面两句话有且仅有一个成立:

(i) λ 是 A 的特征值;

(ii) $\lambda I - A$ 可逆, 也即 $(\lambda I - A)^{-1} \in \mathscr{L}(\mathscr{X})$.

定义 3.7.2 (预解集, 正则值)

设 $\mathscr{X}$ 是 B 空间, $A : D(A) \subset \mathscr{X} \to \mathscr{X}$ 是闭线性算子, 称集合

$$\rho(A) := \{\lambda \in \mathbb{C} : (\lambda I - A)^{-1} \in \mathscr{L}(\mathscr{X})\}$$

为 A 的预解集, $\rho(A)$ 中的 λ 称为 A 的正则值.

♣

根据上述定义, 当 $\dim \mathscr{X} < \infty$, 对任意的 $\lambda \in \mathbb{C}$, 知它要么是 A 的特征值, 要么是 A 的正则值. 但当 $\dim \mathscr{X} = \infty$ 时情况就复杂了:

(i) $(\lambda I - A)^{-1}$ 不存在, 此时 λ 是特征值.

(ii) $(\lambda I - A)^{-1}$ 存在, 且 $\lambda I - A$ 的值域 $R(\lambda I - A) := (\lambda I - A)D(A) = \mathscr{X}$, 此时 $\lambda I - A$ 既是满射又是单射, 因而根据 Banach 逆算子定理(3.4.3)知 $(\lambda I - A)^{-1} \in \mathscr{L}(\mathscr{X})$, 也即 λ 是正则值.

(iii) $(\lambda I - A)^{-1}$ 存在, $R(\lambda I - A) \neq \mathscr{X}$, 但 $\overline{R(\lambda I - A)} = \mathscr{X}$. 这部分 λ 称为 A 的连续谱.

(iv) $(\lambda I - A)^{-1}$ 存在, 且 $\overline{R(\lambda I - A)} \neq \mathscr{X}$. 这部分 λ 称为 A 的剩余谱.

定义 3.7.3 (谱集, 谱点, 点谱)

记 $\sigma(A) := \mathbb{C}\backslash\rho(A)$, 称 $\sigma(A)$ 为 A 的谱集. $\sigma(A)$ 中的点称为 A 的谱点. 对应于上述情形 (i) 的 λ 的集合 (也即 $(\lambda I - A)^{-1}$ 不存在), 记作 $\sigma_p(A)$, 称为 A 的点谱. A 的连续谱记作 $\sigma_c(A)$, A 的剩余谱记作 $\sigma_r(A)$.

♣

从定义显见

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A).$$

下面说明当 $\dim \mathscr{X} = \infty$ 时这些谱都可能出现.

例 3.46 设 $\mathscr{X} = L^2[0,1]$, 考虑算子 $A : u(t) \mapsto -\frac{d^2}{dt^2}u(t)$. 要得到闭算子, 首先需要明确定义域. 对 $u \in \mathscr{X}$, 有 Fourier 展开:

$$u(t) = \sum_{n=-\infty}^{\infty} u_n e^{2\pi int}$$

其中

$$u_n = \int_0^1 u(t)e^{-2\pi int}dt \quad (n \in \mathbb{Z})$$

现在定义

$$(Au)(t) = \sum_{n=-\infty}^{\infty} (2\pi n)^2 u_n e^{2\pi int}$$

显见 $u \in C^2[0,1]$ 时有 $(Au)(t) = -\frac{d^2}{dt^2}u(t)$. 令

$$D(A) = \{u \in \mathscr{X} : Au \in \mathscr{X}\}$$

则 $A: D(A) \to \mathscr{X}$ 是闭线性算子. 下面说明

$$\sigma(A) = \sigma_p(A) = \{(2n\pi)^2 : n = 0, 1, 2, \cdots\}$$

这是因为一方面有

$$-\frac{d^2}{dt^2}\sin 2n\pi t = (2n\pi)^2 \sin 2n\pi t, -\frac{d^2}{dt^2}\cos 2n\pi t = (2n\pi)^2 \cos 2n\pi t, \quad n = 0, 1, 2, \cdots$$

故

$$((2n\pi)^2 I - A)(\sin 2n\pi t) = (2n\pi)^2 \sin 2n\pi t - (2n\pi)^2 \sin 2n\pi t = 0 = ((2n\pi)^2 I - A)(\cos 2n\pi t), \quad n = 0, 1, 2, \cdots$$

得到

$$((2n\pi)^2 I - A)u(t) \equiv 0, \quad n = 0, 1, 2, \cdots$$

这说明 $(2n\pi)^2 I - A = \theta(n = 0, 1, \cdots)$, 也即 $((2n\pi)^2 I - A)^{-1}$ 不存在, 从而 $(2n\pi)^2(n = 0, 1, \cdots)$ 依照定义是点谱.

另一方面, 当 $\lambda \neq (2n\pi)^2$ 时, 根据常微分方程的理论知对任意的 $f \in L^2[0,1]$, 对方程

$$(-\frac{d^2}{dt^2} - \lambda)u(t) = f(t)$$

两边同取第 n 个 Fourier 系数有

$$((2n\pi)^2 - \lambda)u_n = f_n = \int_0^1 f(t)e^{-2\pi int}dt$$

得到

$$u_n = \frac{\int_0^1 f(t)e^{-2\pi int}dt}{(2n\pi)^2 - \lambda}$$

因而

$$u(t) = \sum_{n=-\infty}^{\infty} \frac{\int_0^1 f(t)e^{-2\pi int}dt}{(2n\pi)^2 - \lambda}$$

可以验证 $u \in D(A)$, 且根据 Parseval 恒等式有

$$\|u\|^2 = \sum_{n=-\infty}^{\infty} \frac{|\int_0^1 f(t)e^{-2\pi int}dt|^2}{|(2n\pi)^2 - \lambda|^2} \leq M_\lambda^2 \sum_{n=-\infty}^{\infty} |f_n|^2 = M_\lambda^2 \|f\|^2$$

其中

$$M_\lambda = \sup_{n\in\mathbb{Z}} \frac{1}{|(2n\pi)^2 - \lambda^2|} < \infty$$

这说明 $\lambda \in \rho(A)$.

综上即得

$$\sigma(A) = \sigma_p(A) = \{(2n\pi)^2 : n = 0, 1, 2, \cdots\}$$

例 3.47 设 $\mathscr{X} = C[0,1]$, $A: u(t) \mapsto t \cdot u(t)$. 则 A 是一个有界算子, 且

$$\sigma(A) = \sigma_r(A) = [0,1].$$

证明

对任意的 $\lambda \notin [0,1]$, 对乘法算子 $(\lambda - t)^{-1}$ 而言, 知

$$\|\frac{1}{\lambda - t}x(t)\| \leq \sup_{t\in[0,1]} \frac{1}{|\lambda - t|}\|x\|$$

因而 $(\lambda - t)^{-1} \in \mathscr{L}(\mathscr{X})$, 这说明 $\sigma(A) \subset [0,1]$. 而对任意的 $\lambda \in [0,1]$, 方程

$$(\lambda - t)u(t) = 0$$

只有 θ 解, 即 $u(t) \equiv 0(t \in [0,1])$. 同时对 $v \in R(\lambda I - A)$, 根据定义设 $v(t) = (\lambda - t)u(t)$, 代入 $t = \lambda$ 知必有 $v(\lambda) = 0$, 这说明至少 $v(t) \equiv 1 \notin \overline{R(\lambda I - A)}$, 故

$$[0,1] \subset \sigma_r(A) \subset \sigma(A) \subset [0,1]$$

命题即证. □

例 3.48 设 $\mathscr{X} = L^2[0,1]$, $A : u(t) \mapsto tu(t)$, 则 A 是有界线性算子, 且

$$\sigma(A) = \sigma_c(A) = [0,1].$$

证明

该例与前例类似, 不同在于需要重新讨论 $\overline{R(\lambda I - A)}$. 知此时 $v(t) \equiv 1 \notin R(\lambda I - A)$, 如若不然, 则存在 $u \in L^2[0,1]$ 满足 $1 = (\lambda - t)u(t) \Rightarrow u(t) = \frac{1}{\lambda - t}, t \in [0,1]$, 但 $\frac{1}{\lambda - t} \notin L^2[0,1]$, 矛盾! 另一方面, 注意到 $R(\lambda I - A)$ 中的函数在 $t = \lambda$ 的任意一个小邻域外依旧可以成为 $L^2[0,1]$ 中的函数, 这说明 $\overline{R(\lambda I - A)} = L^2[0,1]$, 也即欲证. □

3.7.1.2 Gelfand 定理

现在研究谱集 $\sigma(A)$. 当 $\dim \mathscr{X} < \infty$ 时, 根据线性代数的知识知必定有 $\sigma(A) \neq \varnothing$, 这是因为特征多项式 $\det(\lambda I - A) = 0$ 根据代数基本定理必有根. 但代数基本定理不能推广到无穷维空间, 故这个方法在无穷维空间行不通. 注意代数基本定理可以由 Liouville 定理给出证明, 故现在可以考虑函数的解析性.

定义 3.7.4 (预解式)

算子值函数 $R_\lambda(A) : \rho(A) \to \mathscr{L}(\mathscr{X})$ 定义为

$$\lambda \mapsto (\lambda I - A)^{-1} \quad (\forall \lambda \in \rho(A))$$

称为 A 的预解式.

定义 3.7.5 (解析)

设 $\mathscr{X}$ 是 Banach 空间, $G \subset \mathbb{C}$. 称函数 $x(z) : G \to \mathscr{X}$ 解析, 如果 $\forall z_0 \in G$, 在 $\mathscr{X}$ 中存在极限

$$\lim_{z \to z_0} \frac{x(z) - x(z_0)}{z - z_0}.$$

现在希望证明:

(i) $\rho(A)$ 是开集;

(ii) $R_\lambda(A)$ 是 $\rho(A)$ 内的算子值解析函数.

要证明 (i), 引入下述引理.

引理 3.7.1

设 $T \in \mathscr{L}(\mathscr{X})$, $\|T\| < 1$, 则 $(I - T)^{-1} \in \mathscr{L}(\mathscr{X})$, 且

$$\|(I - T)^{-1}\| \le \frac{1}{1 - \|T\|}.$$

证明

首先说明 $(I - T)^{-1}$ 一定存在, 这是因为若 $(I - T)x = 0$, 则 $Tx = x$, 但

$$\| Tx \| \le \| T \| \| x \| < \| x \|, \quad \forall x \neq \theta$$

这说明只能有 $x = \theta$, 故 $I - T$ 是单射, 进而其逆有定义.

其次说明

$$\forall y \in \mathscr{X} \exists! x \in \mathscr{X} ((x \text{是 } Sx := y + Tx \text{ 的不动点}) \wedge (\| x \| \le M \| y \|)) \Rightarrow \| (I - T)^{-1} \| \le M$$

这是因为

$$y = Sx - Tx = (I - T)x \Rightarrow (I - T)^{-1} y = x \Rightarrow \sup_{y \in \mathscr{X}} \frac{\| (I - T)^{-1} y \|}{\| y \|} \le \sup_{y \in \mathscr{X}} \frac{\| x \|}{\| y \|} \le M \Rightarrow \| (I - T)^{-1} \| \le M$$

现在

$$\| Sx - Sx' \| = \| Tx - Tx' \| \le \| T \| \| x - x' \|, \quad \forall x, x' \in \mathscr{X}$$

根据 $\| T \| < 1$ 知 S 是压缩映射, 从而存在唯一 $x = Sx = y + Tx$, 得到

$$\| x \| = \| y + Tx \| \le \| y \| + \| Tx \| \le \| y \| + \| T \| \| x \| \Rightarrow \| x \| \le \frac{\| y \|}{1 - \| T \|}$$

取 $M = \frac{1}{1-\|T\|}$ 即得

$$\| (I - T)^{-1} \| \le \frac{1}{1 - \| T \|}$$

□

注 事实上, 若定义 $Sx := y + Tx, y \in \mathscr{X}$, 则容易证明 Sx 也是压缩映射, 且

$$x = \lim_{n\to\infty} S^n y = \sum_{k=0}^{\infty} T^k y (= y + Ty + T^2 y + \cdots)$$

进而当 $\|T\| < 1$ 时有

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k \tag{3.85}$$

称 $\sum\limits_{k=0}^{\infty} T^k$ 为 Neuman 级数.

推论 3.7.1

设 A 是闭线性算子, 则 $\rho(A)$ 是开集.

证明

从开集的定义着手. 设 $\lambda_0 \in \rho(A)$, 知

$$\lambda I - A = (\lambda - \lambda_0)I + (\lambda_0 I - A) = (\lambda_0 I - A)[I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]$$

套用引理(3.7.1), 当 $\| -(\lambda - \lambda_0)(\lambda_0 I - A)^{-1} \| < 1$, 即 $|\lambda - \lambda_0| < \frac{1}{\|(\lambda_0 I - A)^{-1}\|}$ 时, 有

$$B := [I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]^{-1} \in \mathscr{L}(\mathscr{X})$$

注意因为已经有 $\lambda_0 \in \rho(A)$, 故 $R_{\lambda_0}(A) \in \mathscr{L}(\mathscr{X})$, 知

$$(\lambda I - A)^{-1} = [I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]^{-1}(\lambda_0 I - A)^{-1} = BR_{\lambda_0}(A) \in \mathscr{L}(\mathscr{X}) \in \mathscr{L}(\mathscr{X})$$

这说明 $\lambda \in \rho(A)$, 这便找到了 λ_0 在 $\rho(A)$ 中的邻域, 命题得证. □

下面研究 $R_\lambda(A)$ 的解析性, 考虑对 $R_\lambda(A)$ 求导.

引理 3.7.2 (第一预解公式)

设 $\lambda, \mu \in \rho(A)$, 则有

$$R_\lambda(A) - R_\mu(A) = (\mu - \lambda) R_\lambda(A) R_\mu(A) \tag{3.86}$$

证明

直接计算有

$$\begin{aligned}(\lambda I - A)^{-1} &= (\lambda I - A)^{-1}(\mu I - A)(\mu I - A)^{-1} \\ &= (\lambda I - A)^{-1}[(\mu - \lambda)I + \lambda I - A](\mu I - A)^{-1} \\ &= (\mu - \lambda)(\lambda I - A)^{-1}(\mu I - A)^{-1} + (\mu I - A)^{-1}\end{aligned}$$

套用 $R_\lambda(A), R_\mu(A)$ 的定义即得结论. □

定理 3.7.1

预解式 $R_\lambda(A)$ 在 $\rho(A)$ 内是算子值解析函数.

证明

首先证明 $R_\lambda(A)$ 关于 λ 连续. 设 $\lambda_0 \in \rho(A)$, 从推论(3.7.1)的证明过程知, 当 $|\lambda - \lambda_0| < \frac{1}{2\|R_{\lambda_0}(A)\|}$ 时有

$$\|R_\lambda(A)\| \leq \|R_{\lambda_0}(A)\| \cdot \|[I + (\lambda - \lambda_0)R_{\lambda_0}(A)]^{-1}\|$$

由引理(3.7.1)知

$$\|[I + (\lambda - \lambda_0)R_{\lambda_0}(A)]^{-1}\| \leq \frac{1}{1 - \|(\lambda_0 - \lambda)R_{\lambda_0}(A)\|} < \frac{1}{1 - \frac{1}{2\|R_{\lambda_0}(A)\|} \cdot \|R_{\lambda_0}(A)\|} = 2$$

得到

$$\|R_\lambda(A)\| \leq 2\|R_{\lambda_0}(A)\|$$

再由第一预解公式(3.86)得

$$\|R_\lambda(A) - R_{\lambda_0}(A)\| \leq |\lambda - \lambda_0| \cdot \|R_\lambda(A)\| \cdot \|R_{\lambda_0}(A)\| \leq 2\|R_{\lambda_0}(A)\|^2 \cdot |\lambda - \lambda_0| \to 0, \quad \lambda \to \lambda_0$$

至此连续性得证, 再证 $R_\lambda(A)$ 关于 λ 的可微性. 由第一预解公式(3.86)知

$$\lim_{\lambda \to \lambda_0} \frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} = -\lim_{\lambda \to \lambda_0} R_\lambda(A) \cdot R_{\lambda_0}(A) = -(R_{\lambda_0}(A))^2.$$

注意 $\lambda_0 \in \rho(A)$ 是任取的, 故 $R_\lambda(A)$ 在 $\rho(A)$ 内解析. □

下面讨论谱点的存在性.

定理 3.7.2

设 A 是有界线性算子, 则 $\sigma(A) \neq \varnothing$.

证明

用反证法. 若 $\rho(A) = \mathbb{C}$, 则由定理(3.7.1)知 $R_\lambda(A)$ 在 $\mathbb{C}$ 上解析, 且根据 $\rho(A)$ 的定义知 $R_\lambda(A) \in \mathscr{L}(\mathscr{X})(\forall \lambda \in \mathbb{C})$. 取 Neuman 级数(3.85)知当 $|\lambda| > \|A\|$ 时有

$$R_\lambda(A) = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} A^n \tag{3.87}$$

且此时由引理(3.7.1)知

$$\|R_\lambda(A)\| \leq \frac{1}{|\lambda| - \|A\|} \tag{3.88}$$

当 $|\lambda| \leq \|A\|$, 注意此时 λ 取在有界闭区域中, 故 $\|R_\lambda(A)\|$ 必能取得最大值, 这说明 $\|R_\lambda(A)\|$ 在复平面上有界.

任取 $f \in \mathscr{L}(\mathscr{X})^*$, 考察 (数值) 解析函数

$$u_f(\lambda) := f(R_\lambda(A))$$

注意 $u_f(\lambda)$ 是在全平面有界的解析函数, 因而由 Liouville 定理知 $u_f(\lambda)$ 是常数, 其与 λ 无关. 但若 $R_\lambda(A)$ 不是常值, 根据推论(3.5.1)知, 必能取到某个 f 使得 $f(R_\lambda(A))$ 至少不是常数, 矛盾! 故 $R_\lambda(A)$ 是常值算子, 其与 λ 无关. 但这与第一预解公式(3.86)矛盾! 命题即证. □

下面估计谱集的范围. 从(3.87)式与(3.88)式知, 当 A 是有界线性算子时, 只要 $|\lambda| > \|A\|$, 则 $R_\lambda(A) \in \mathscr{L}(\mathscr{X})$[26], 故 $\sigma(A) \subset \overline{B}(0, \|A\|)$. 又由推论(3.7.1)与定理(3.7.2)知 $\sigma(A)$ 是非空紧集.

[26]这里没有想明白为什么.

定义 3.7.6 (谱半径)

设 $A \in \mathscr{L}(\mathscr{X})$, 称数

$$r_\sigma(A) := \sup\{|\lambda| : \lambda \in \sigma(A)\}$$

为 A 的谱半径.

从定义知 $r_\sigma(A) \le \|A\|$, 现在希望得到更精确的估计. 由(3.87)与 Cauchy-Hadamard 公式可得: 当

$$|\lambda| > \overline{\lim_{n\to\infty}} \|A^n\|^{\frac{1}{n}}$$

时, $R_\lambda(A) \in \mathscr{L}(\mathscr{X})$. 故

$$r_\sigma(A) \le \overline{\lim_{n\to\infty}} \|A^n\|^{\frac{1}{n}}. \tag{3.89}$$

可以说明(3.89)式是 $r_\sigma(A)$ 的最佳估计. 事实上有 Gelfand 定理.

定理 3.7.3 (Gelfand)

设 $\mathscr{X}$ 是 B 空间, $A \in \mathscr{L}(\mathscr{X})$, 那么

$$r_\sigma(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

证明

记 $a := r_\sigma(A)$, 首先证明

$$\overline{\lim_{n\to\infty}} \|A^n\|^{\frac{1}{n}} \le a \tag{3.90}$$

任取 $f \in \mathscr{L}(\mathscr{X})^*$, 考虑复函数

$$u_f(\lambda) := f(R_\lambda(A))$$

因为根据定理(3.7.1), $R_\lambda(A)$ 在 $\rho(A)$ 内解析, 故 $u_f(\lambda)$ 在 $|\lambda| > a$ 解析. 知此时 $u_f(\lambda)$ 有 Laurent 展开式:

$$u_f(\lambda) = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} f(A^n)$$

根据 Abel 第一定理知对任意的 $\varepsilon > 0$ 都有

$$\sum_{n=0}^{\infty} \frac{1}{(a+\varepsilon)^{n+1}} |f(A^n)| < \infty$$

这说明对任意的 $f \in \mathscr{L}(\mathscr{X})^*$ 与任意的 $n \in \mathbb{N}$, $|f(\frac{A^n}{(a+\varepsilon)^{n+1}})|$ 都有界. 特别对每个 $n \in \mathscr{N}$, 由 Hahn-Banach 定理的推论(3.5.2), 总能取 $f_n \in \mathscr{L}(\mathscr{X})^*$ 使得 $f_n(\frac{A^n}{(a+\varepsilon)^{n+1}}) = \|\frac{A^n}{(a+\varepsilon)^{n+1}}\|, \|f\| = 1$. 这说明

$$\sup_{n\ge1} \|\frac{A^n x}{(a+\varepsilon)^{n+1}}\| < \infty, \quad \forall x \in \mathscr{X}$$

进而由共鸣定理(3.4.8)知存在常数 $M > 0$ 使得

$$\forall n \in \mathbb{N}(\frac{\|A^n\|}{(a+\varepsilon)^{n+1}} \le M)$$

进而

$$\overline{\lim_{n\to\infty}} \|A^n\|^{\frac{1}{n}} \le a+\varepsilon$$

注意 $\varepsilon > 0$ 是任取的, 故(3.90)式成立. 结合(3.89)式可得

$$r_\sigma(A) = \overline{\lim_{n\to\infty}} \|A^n\|^{\frac{1}{n}} \tag{3.91}$$

现在记

$$P_\lambda(A) = \sum_{j=1}^{n} \lambda^{j-1} A^{n-j}$$

则对任意的 $\lambda \in \mathbb{C}$ 有

$$\lambda^n I - A^n = (\lambda I - A)(\lambda^{n-1} I + \lambda^{n-2} A + \cdots + A^{n-1}) = (\lambda I - A) P_\lambda(A) = P_\lambda(A)(\lambda I - A)$$

于是当 $\lambda^n \in \rho(A^n)$, 知 $\lambda^n I - A^n = P_\lambda(A)(\lambda I - A) \in \mathscr{L}(\mathscr{X})$, 又因为 $P_\lambda(A) \in \mathscr{L}(\mathscr{X})$, 故 $\lambda I - A \in \mathscr{L}(\mathscr{X})$, 即 $\lambda \in \rho(A)$. 这说明

$$\lambda \in \sigma(A) \Rightarrow \lambda^n \in \sigma(A^n)$$

此时 $|\lambda^n| \le \|A^n\|$, 也即 $|\lambda| \le \varliminf_{n\to\infty} \|A^n\|^{\frac{1}{n}}$, 根据上确界定义得到

$$r_\sigma(A) \le \varliminf_{n\to\infty} \|A^n\|^{\frac{1}{n}} \tag{3.92}$$

联合(3.91)式与(3.92)式即得 $\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}$ 存在, 且

$$r_\sigma(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}.$$

□

注 定理(3.7.2)与 Gelfand 定理(3.7.3)可以推广到一般的 Banach 代数上.

3.7.1.3 例子

现在特别考察几个算子的谱集.

例 3.49(推移算子) 在空间 l^2 上, 考察右推移算子

$$A : x = (x_1, x_2, \cdots, x_n, \cdots) \mapsto (0, x_1, x_2, \cdots, x_n, \cdots)$$

则

$$\sigma_p(A) = \varnothing, \sigma_r(A) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}, \sigma_c(A) = \{\lambda \in \mathbb{C} : |\lambda = 1\}.$$

证明

因为 $\|A\| = 1$, 由 Gelfand 定理(3.7.3)知 $r_\sigma(A) = 1$, 进而

$$\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \le 1\}$$

可以验证 A 的共轭算子是:

$$A^* : l^2 \to l^2, (x_1, x_2, \cdots, x_n, \cdots) \mapsto (x_2, x_3, \cdots, x_{n+1}, \cdots)$$

根据共轭算子的定义有

$$((\lambda I - A)x, y) = (x, (\bar{\lambda} I - A^*)y) \quad (\forall x, y \in l^2)$$

进而

$$(\bar{\lambda} I - A^*)y = 0 \Leftrightarrow y \in \overline{R(\lambda I - A)}^{\perp} = R(\lambda I - A)^{\perp}$$

这说明

$$\overline{R(\lambda I - A)}^{\perp} = N(\bar{\lambda} I - A^*).$$

首先证明 $\sigma_p(A) = \varnothing$. 设 $(\lambda I - A)x = 0$, 代入 A 知 $\lambda x_n = x_{n-1} (n \ge 2), \lambda x_1 = 0$. 故

$$\lambda \ne 0 \Rightarrow x = 0, \quad \lambda = 0 \Rightarrow x = 0$$

这说明 $(\lambda I - A)^{-1}$ 总存在, 故 $\sigma_p(A) = \varnothing$.

再证明 $\sigma_r(A) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}, \sigma_c(A) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$. 由前述结论此即证明

$$|\lambda| < 1 \Rightarrow \overline{R(\lambda I - A)}^{\perp} = N(\bar{\lambda} I - A^*) \ne \{0\}$$

$$|\lambda| = 1 \Rightarrow R(\lambda I - A)^{\perp} = N(\bar{\lambda} I - A^*) = \{0\}$$

要求 $N(\bar{\lambda} I - A^*)$, 即求解 $(\bar{\lambda} I - A^*)x = 0$, 代入 A^* 有

$$\bar{\lambda} x_n = x_{n+1}, \quad n = 1, 2, \cdots$$

亦即

$$x_{n+1} = \overline{\lambda}^n x_1, \quad n = 1, 2, \cdots$$

这说明当 $|\lambda| < 1$ 时:

$$N(\overline{\lambda} I - A^*) = \{c(1, \overline{\lambda}, \overline{\lambda}^2, \cdots, \overline{\lambda}^n, \cdots) : c \in \mathbb{C}\}$$

即 $N(\overline{\lambda} I - A^*) \neq \{0\}$, $R(\lambda I - A)$ 在 l^2 中不稠密, $\lambda \in \sigma_r(A)$. 故

$$\{\lambda \in \mathbb{C} : |\lambda| < 1\} \subset \sigma_r(A), \quad \sigma(A) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$$

当 $|\lambda| = 1$, 从级数收敛的必要条件知 $x = x_1(1, \overline{\lambda}, \cdots, \overline{\lambda}^n, \cdots) \neq 0 \Rightarrow x \notin l^2$, 故 $N(\overline{\lambda} I - A^*) = \{0\}$, 即 $R(\lambda I - A)$ 在 l^2 中稠密. 因为 $\lambda \in \sigma(A)$, 故只能有 $\lambda \in \sigma_c(A)$.

综上, 命题即证. □

例 3.50(对称算子) 设 $(\mathscr{X}, (\cdot, \cdot))$ 是 Hilbert 空间, $A \in \mathscr{L}(\mathscr{X})$ 称为对称算子, 如果 $A^* = A$, 即

$$(Ax, y) = (x, Ay) \quad (\forall x, y \in \mathscr{X}) \tag{3.93}$$

若 A 是对称算子, 则 $\sigma(A) \subset \mathbb{R}$, 且 $\sigma_r(A) = \varnothing$.

证明

由(3.93)式知 $\forall x \in \mathscr{X}((Ax, x) \in \mathbb{R})$. 设 $\lambda = a + ib, b \neq 0$, 下面证明 $\lambda I - A$ 有有界逆, 考虑应用 Banach 逆算子定理, 这需要证明 $\lambda I - A$ 是双射.

先证明 $\lambda I - A$ 是单射. 因为

$$\|(\lambda I - A)x\|^2 = \|(aI - A)x\|^2 + b^2\|x\|^2 \tag{3.94}$$

故 $N(\lambda I - A) = \{0\}$, 因而 $\lambda I - A$ 是单射.

再证明 $\lambda I - A$ 是满射, 为此先证明 $R(\lambda I - A)$ 是闭的. 设 $x_n \in \mathscr{X}$, 且

$$(\lambda I - A)x_n = y_n \to y \quad (n \to \infty)$$

由(3.94)式得

$$\|(\lambda I - A)(x_n - x_m)\|^2 \geq b^2\|x_n - x_m\|^2$$

这说明 $\{x_n\}$ 是 $\mathscr{X}$ 中的 Cauchy 列, 设 $x_n \to x (n \to \infty)$, 此时

$$y_n = (\lambda I - A)x_n \to (\lambda I - A)x = y \in R(\lambda I - A), \quad n \to \infty$$

这说明

$$R(\lambda I - A) = \overline{R(\lambda I - A)}$$

也即 $R(\lambda I - A)$ 是闭的. 再说明 $R(\lambda I - A)^\perp = \{0\}$. 设 $y \in R(\lambda I - A)^\perp$, 根据定义有

$$((\lambda I - A)x, y) = 0 \quad (\forall x \in \mathscr{X})$$

根据 A 的对称性有

$$(x, (\overline{\lambda} I - A)y) = 0 \quad (\forall x \in \mathscr{X})$$

故 $(\overline{\lambda} I - A)y = 0$, 即 $y \in N(\overline{\lambda} I - A)$. 将 $\overline{\lambda}$ 代入(3.94)式得 $N(\overline{\lambda} I - A) = \{0\}$, 这说明只能有 $y = 0$. 故

$$R(\lambda I - A)^\perp = \{0\}$$

综上有 $R(\lambda I - A) = \mathscr{X}$, 从而 $\lambda I - A$ 是满射, 因而由 Banach 逆算子定理(3.4.3)知当 $\lambda = a + ib, b \neq 0$ 时必有 $(\lambda I - A)^{-1} \in \mathscr{L}(\mathscr{X})$. 故 $\sigma(A) \subset \mathbb{R}$.

再说明 $\sigma_r(A) = \varnothing$. 设 $\lambda \in \sigma(A)$, 但 $\lambda \notin \sigma_p(A)$, 也即 $(\lambda I - A)^{-1}$ 存在, 此时 $N(\lambda I - A) = \{0\}$, 有

$$R(\lambda I - A)^\perp = N(\overline{\lambda} I - A^*) = N(\lambda I - A) = \{0\}$$

故 $R(\lambda I - A)$ 在 $\mathscr{X}$ 中稠密, 得到 $\lambda \in \sigma_c(A)$, 即 $\sigma_r(A) = \varnothing$. □

例 3.51(酉算子) 设 $(\mathscr{X},(\cdot,\cdot))$ 是 Hilbert 空间, $U\in\mathscr{L}(\mathscr{X})$ 称为酉算子, 若

$$\forall x,y\in\mathscr{X}((Ux,Uy)=(x,y))\quad 且\quad R(U)=\mathscr{X} \tag{3.95}$$

若 U 是酉算子, 则 $\sigma(U)\subset S^1=\{e^{i\theta}:\theta\in[0,2\pi]\}$, 且 $\sigma_r(U)=\varnothing$.

证明

从酉算子的定义(3.95)知 $\|Ux\|^2=\|x\|^2$, 故 U 首先是单射, 且

$$\|U\|=1\Rightarrow\sigma(A)\subset\{\lambda\in\mathbb{C}:|\lambda|\leq 1\}$$

同时根据定义知 U 是满射, 因而由 Banach 逆算子定理(3.4.3)知 U 存在有界逆 U^{-1}, 且 $\|U^{-1}\|=1$. 设 $|\lambda|<1$, 知

$$\lambda I-U=-U(I-\lambda U^{-1}) \tag{3.96}$$

注意 $\|\lambda U^{-1}\|=|\lambda|\cdot\|U^{-1}\|=|\lambda|<1$, 故

$$\|\lambda U^{-1}(x-y)\|\leq\|\lambda U^{-1}\|\cdot\|x-y\|$$

进而由压缩映射原理(2.1.1)知存在唯一的 $x\in\mathscr{X}$ 使得 $\lambda U^{-1}x=x$, 也即 $(I-\lambda U^{-1})x=0$ 有唯一解. 代入 $x=\theta$ 知方程成立, 故 $N(I-\lambda U^{-1})=\{\theta\}$, 也即 $I-\lambda U^{-1}$ 是单射. 由 U 的满射性知 $I-\lambda U^{-1}$ 也是满射, 故由 Banach 逆算子定理知 $I-\lambda U^{-1}$ 存在有界逆, 进而由(3.96)式知 $\lambda I-U$ 有有界逆, 且

$$(\lambda I-U)^{-1}=(I-\lambda U^{-1})^{-1}U^{-1}$$

这说明 $\lambda\notin\sigma(A)$. 注意此时 $|\lambda|<1$, 故 $\sigma(A)\subset\{\lambda\in\mathbb{C}:|\lambda|=1\}$.

设 $\lambda\in\sigma(U)$, 但 $\lambda\notin\sigma_p(A)$, 此时 $N(\lambda I-U)=\{0\}$. 可以证明

$$R(\lambda I-U)^{\perp}=N(\bar{\lambda}I-U^{-1})=N(\lambda I-U)=\{0\}$$

这说明 $R(\lambda I-U)$ 在 $\mathscr{X}$ 中稠密, 从而 $\lambda\in\sigma_c(U)$, 即 $\sigma_r(U)=\varnothing$. □

3.7.2 习题

练习 3.70 设 $\mathscr{X}$ 是 B 空间, 求证: $\mathscr{L}(\mathscr{X})$ 中可逆 (存在有界逆) 算子集是开的.

证明

任意取定 $A_0\in\mathscr{L}(\mathscr{X})$, 希望证明存在 $\delta>0$, 使得当 $\|A-A_0\|<\delta$ 时 A^{-1} 存在. 知条件 $\|A-A_0\|<\delta$ 可强化为 $\|A_0\|\cdot\|\lambda I-A_0^{-1}A\|<\delta$, 也即 $\|\lambda I-A_0^{-1}A\|<\frac{\delta}{\|A_0\|}$. 既然 $A_0^{-1},A\in\mathscr{L}(\mathscr{X})$, 知 $I-A_0^{-1}A\in\mathscr{L}(\mathscr{X})$. 现套用引理(3.7.1), 取 $\delta<\|A_0\|$, 则知 $(I-(I-A_0^{-1}A))^{-1}=A^{-1}A_0\in\mathscr{L}(\mathscr{X})$, 得到 $A^{-1}\in\mathscr{L}(\mathscr{X})$, 即 A 可逆, 命题即证.

练习 3.71 设 A 是闭线性算子, $\lambda_1,\lambda_2,\cdots,\lambda_n\in\sigma_p(A)$ 两两互异, 又设 x_i 是对应于 λ_i 的特征元 $(i=1,2,\cdots,n)$. 求证: $\{x_1,x_2,\cdots,x_n\}$ 是线性无关的.

证明

用反证法, 若 $\{x_1,\cdots,x_n\}$ 线性相关, 根据特征元的定义知

$$Ax_i=\lambda_i x_i,\quad i=1,2,\cdots,n$$

如果 $x_1,\cdots,x_n$ 两两相关, 不妨设 $x_1=kx_2$, 则

$$Ax_1=\lambda_1x_1,Ax_2=\lambda_2x_2=k\lambda_2x_1\Rightarrow k(\lambda_1-\lambda_2)x_1=0$$

因为 $x_1,\cdots,x_n$ 都是特征元, 故它们均非零, 因而只能有 $\lambda_1=\lambda_2$, 矛盾! 故 $\{x_1,\cdots,x_n\}$ 的极大无关组必非空, 不妨就设其为 $\{x_1,\cdots,x_m\}(m<n)$. 知 $x_{m+1},\cdots,x_n$ 均能由 $x_1,\cdots,x_m$ 线性表示, 设

$$x_{m+i}=\sum_{j=1}^{m}k_{i,j}x_j,\quad i=1,\cdots,n-m$$

对每个 i 有

$$A(\sum_{j=1}^{m} k_{i,j}x_j) = Ax_{m+i} = \sum_{j=1}^{m} k_{i,j}\lambda_j x_j = \lambda_{m+i}\sum_{j=1}^{m} k_{i,j}x_j$$

因为 $x_1,\cdots,x_m$ 无关, 故只能有 $k_{i,j}\lambda_j = k_{i,j}\lambda_{m+i}(i=1,\cdots,n-m, j=1,\cdots,m)$, 注意必能找到非零的 $k_{i,j}$, 得到矛盾, 命题即证.

练习 3.72 在双边 l^2 空间上, 考察右推移算子

$$A: x = (\cdots,\xi_{-n},\xi_{-n+1},\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},\xi_n,\cdots)\in l^2 \mapsto Ax = (\cdots,\eta_{-n},\eta_{-n+1},\cdots,\eta_{-1},\eta_0,\eta_1,\cdots,\eta_{n-1},\eta_n,\cdots)$$

其中 $\eta_m = \xi_{m-1}(m\in\mathbb{Z})$. 求证: $\sigma_c(A) = \sigma(A)$ 是单位圆周.

证明

首先知

$$\forall x\in\mathscr{X}(\|Ax\|_{l^2} = \|x\|_{l^2}) \Rightarrow \|A\| = 1$$

故由 Gelfand 定理知 $\sigma(A)\subset\{\lambda\in\mathbb{C}:|\lambda|\le 1\}$. 当 $|\lambda|\le 1$, 将 A 代入方程 $(\lambda I - A)x = 0$ 知

$$\lambda x_i = x_{i-1} \Rightarrow x_n = \lambda^n x_0 \quad (\lambda\neq 0)$$

这说明 $N(\lambda I - A)\subset\{x_0(\cdots,\lambda^{-n},\cdots,\lambda^{-1},1,\lambda,\cdots): x_0\in\mathbb{C}\}\cap\{\theta\}$. 注意 $\{x_0(\cdots,\lambda^{-n},\cdots,\lambda^{-1},1,\lambda,\cdots): x_0\in\mathbb{C}\}\cap l^2 = \theta$, 故 $N(\lambda I - A) = \theta$, 这说明 $\lambda I - A$ 首先是单射, 进而 $(\lambda I - A)^{-1}$ 总是存在的, 也即 $\sigma_p(A) = \varnothing$. 下面说明 $\overline{R(\lambda I - A)} = l^2$, 可以证明 A^* 正是左推移算子, 根据共轭算子的定义有

$$((\lambda I - A)x, y) = (x, (\bar{\lambda}I - A^*)y) \quad (\forall x, y\in l^2)$$

故 $(\bar{\lambda}I - A^*)y = 0 \Leftrightarrow y\in\overline{R(\lambda I - A)}^{\perp} = R(\lambda I - A)^{\perp}$. 得到

$$\overline{R(\lambda I - A)}^{\perp} = N(\bar{\lambda}I - A^*)$$

可以证明 $N(\bar{\lambda}I - A^*) = \{\theta\}$, 进而 $\overline{R(\lambda I - A)}^{\perp} = \{\theta\}$, 此即 $\overline{R(\lambda I - A)} = l^2$. 这说明 $\sigma_r(A) = \varnothing$, 从而 $\sigma(A) = \sigma_c(A)$.

最后说明 $\sigma_c(A)$ 是单位圆周[27].

练习 3.73 在 l^2 空间上, 考察左推移算子

$$A: (\xi_1,\xi_2,\cdots)\mapsto(\xi_2,\xi_3,\cdots)$$

求证: $\sigma_p(A) = \{\lambda\in\mathbb{C}:|\lambda|<1\}$, $\sigma_c(A) = \{\lambda\in\mathbb{C}:|\lambda| = 1\}$, 且

$$\sigma(A) = \sigma_p(A)\cup\sigma_c(A).$$

证明

首先知

$$\forall x\in\mathscr{X}(\|Ax\|\le\|x\|) \Rightarrow \|A\|\le 1$$

而显见 $\|A\|\neq 0$, 故 $r_\sigma(A) = \lim\limits_{n\to\infty}\|A^n\|^{\frac{1}{n}}\le 1$, 由 Gelfand 定理知

$$\sigma(A)\subset\{\lambda\in\mathbb{C}:|\lambda|\le 1\}.$$

现将 A 代入方程 $(\lambda I - A)x = 0$, 解得

$$\lambda x_n = x_{n+1}, \quad n = 1,2,\cdots$$

得到

$$N(\lambda I - A) = \{x_1(1,\lambda,\cdots,\lambda^n,\cdots): x_1\in\mathbb{C}\}\cap l^2$$

当 $|\lambda|<1$, 知 $\{x_1(1,\lambda,\cdots,\lambda^n,\cdots): x_1\in\mathbb{C}\}\cap l^2\neq\{\theta\}$, 也即 $N(\lambda I - A)\neq\{\theta\}$, 这说明 $(\lambda I - A)^{-1}$ 不存在, 也即 $\lambda\in\sigma_p(A)$. 故 $\sigma_p(A) = \{\lambda\in\mathbb{C}:|\lambda|<1\}$.

[27] 到这里卡住了: 单位圆周如果定义为 $\{\lambda:|\lambda| = 1\}$, 要怎么否定 $|\lambda|<1$ 的情况?

当 $|\lambda| = 1$, 知 $\{x_1(1, \lambda, \cdots, \lambda^n, \cdots) : x_1 \in \mathbb{C}\} \cap l^2 = \{\theta\}$, 也即 $N(\lambda I - A) = \{\theta\}$, 这说明 $\lambda I - A$ 是单射, 因而 $(\lambda I - A)^{-1}$ 存在. 现需说明

$$R(\lambda I - A) \neq l^2, \quad \overline{R(\lambda I - A)} = l^2$$

注意 A^* 恰是右推移算子, 根据共轭算子的定义有

$$((\lambda I - A)x, y) = (x, (\overline{\lambda} I - A^*)y)$$

这说明 $(\overline{\lambda} I - A^*)y = 0 \Leftrightarrow y \in \overline{R(\lambda I - A)}^{\perp} = R(\lambda I - A)^{\perp}$. 故

$$\overline{R(\lambda I - A)}^{\perp} = N(\overline{\lambda} I - A^*)$$

从而欲证转化为

$$N(\overline{\lambda} I - A^*) \neq \{\theta\}, \quad \overline{N(\overline{\lambda} I - A^*)} = \{\theta\}$$

根据右推移算子的性质, $\sigma_c(A^*) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$, 根据连续谱的性质即得

$$N(\overline{\lambda} I - A^*) \neq \{\theta\}, \quad \overline{N(\overline{\lambda} I - A^*)} = \{\theta\}$$

故 $\sigma_c(A) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

最后, 因为 $\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq 1\} = \sigma_p(A) \cup \sigma_c(A)$, 故 $\sigma(A) = \sigma_p(A) \cup \sigma_c(A)$.

3.8 泛函分析 (高志强老师班) 第二章作业与解答合集

3.8.1 第九次作业

1. 在 $C[0,1]$ 上赋予范数 $\|\cdot\|_1$, 令

$$f(x) = \int_0^1 sx(s)ds, \quad x \in C[0,1]$$

证明 f 是 $C[0,1]$ 的连续线性泛函, 并求 $\|f\|$.

证明

根据定义, 任取 $x, x_1 \in C[0,1]$, 知

$$|f(x) - f(x_1)| = |\int_0^1 (sx(s) - sx_1(s))ds| \leq \int_0^1 s|x(s) - x_1(s)|ds \leq \int_0^1 |x(s) - x_1(s)|ds = \|x - x_1\|_1$$

故 $\|x - x_1\|_1 \to 0 \Rightarrow |f(x) - f(x_1)| \to 0$, 这便说明 f 在 $C[0,1]$ 上连续.

对于 $\|f\|$, 首先由 Hölder 不等式知

$$|f(x)| \leq \int_0^1 s|x(s)|ds \leq \max_{s\in[0,1]} s \cdot \int_0^1 |x(s)|ds = \|x\|_{L^1[0,1]}$$

于是

$$\frac{|f(x)|}{\|x\|_{L^1[0,1]}} \leq 1 \Rightarrow \|f\| = \sup_{x \in \mathscr{X} \setminus \{\theta\}} \frac{|f(x)|}{\|x\|_{L^1[0,1]}} \leq 1$$

现在任取 $1 < p < \infty$, 设 $\frac{1}{p} + \frac{1}{q} = 1$, 由 Hölder 不等式知

$$|f(x)| \leq \int_0^1 s|x(s)|ds \overset{\text{(i)}}{\leq} (\int_0^1 s^p ds)^{\frac{1}{p}} (\int_0^1 |x(s)|^q ds)^{\frac{1}{q}}$$

(i) 式取等当且仅当 $s^p = Cx(s)^q$, 其中 C 是常数, 亦即 $x(s) = Cs^{\frac{p}{q}}$, 其中 C 是常数. 令 $\|x\|_{L^1[0,1]} = 1$ 知

$$\int_0^1 Cs^{\frac{p}{q}}ds = C\frac{q}{p+q} = 1 \Rightarrow C = \frac{p+q}{q}$$

代回 $x(s) = \frac{p+q}{q}s^{\frac{p}{q}}$ 知

$$|f(x)| = \frac{p+q}{q}(\int_0^1 s^p ds)^{\frac{1}{p}}(\int_0^1 s^p ds)^{\frac{1}{q}} = \frac{p}{p+1}$$

于是 $\forall p \in (1, \infty)(\|f\| \geq \frac{p}{p+1})$, 因而 $\|f\| \geq 1$, 故 $\|f\| = 1$. □

2. T 是 $(\mathbb{R}^n, l^1)$ 到 $(\mathbb{R}^n, l^1)$ 的线性算子, $\mathbb{R}^n$ 赋以 l^1 范数, 即

$$\forall x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n, \quad \|x\|_{l^1} = \sum_{i=1}^{n} |x_i|$$

试求 $\|T\|$.

证明

届于 T 是有穷维空间之间的线性算子, 可设 $T = (t_{ij})_{n\times n}$. 现有:

$$Tx = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} t_{1j}x_j \\ \vdots \\ \sum\limits_{j=1}^{n} t_{nj}x_j \end{pmatrix}$$

因而

$$\begin{aligned} \|Tx\|_{l^1} &= \sum_{i=1}^{n} |\sum_{j=1}^{n} t_{ij}x_j| \le \sum_{i=1}^{n}\sum_{j=1}^{n} |t_{ij}||x_j| = \sum_{j=1}^{n}(\sum_{i=1}^{n} |t_{ij}|)|x_j| \\ &\le \max_j(\sum_{i=1}^{n} |t_{ij}|) \sum_{j=1}^{n} |x_j| = \max_j(\sum_{i=1}^{n} |t_{ij}|)\|x\|_{l^1} \end{aligned}$$

亦即

$$\|T\| \le \max_j(\sum_{i=1}^{n} |t_{ij}|)$$

现在要说明等号成立, 就是说明存在 $x \in (\mathbb{R}^n, l^1)$ 使得 $\|x\|_{l^1} = 1$, 且 $\|Tx\|_{l^1} = \max\limits_j(\sum\limits_{i=1}^{n} |t_{ij}|)$. 因为 $1 \le j \le n$, 故必存在 j_0 使得

$$\sum_{i=1}^{n} |t_{ij_0}| = \max_j(\sum_{i=1}^{n} |t_{ij}|)$$

现取 $e^{j_0} = (\underbrace{0, \cdots, 0, 1}_{j_0}, 0, \cdots, 0)$, 显见 e^{j_0} 满足要求, 且 $\|Te^{j_0}\| = \max\limits_j(\sum\limits_{i=1}^{n} |t_{ij}|)$, 故

$$\|T\| = \max_j(\sum_{j=1}^{n} |t_{ij}|)$$

□

3. 设 $x(t) \in C[a,b]$, $f(x) = x(a) - x(b)$, 证明 f 是 $C[a,b]$ 上的有界线性泛函, 并求 $\|f\|$.

证明

f 的线性性显见. 任取 $x_1, x_2 \in C[a,b]$, 知

$$\begin{aligned} |f(x_1) - f(x_2)| &= |x_1(a) - x_1(b) - x_2(a) + x_2(b)| = |(x_1(a) - x_2(a)) - (x_1(b) - x_2(b))| \\ &\le |x_1(a) - x_2(a)| + |x_1(b) - x_2(b)| \le 2\max_{[a,b]}(x_1 - x_2) = 2\|x_1 - x_2\| \end{aligned}$$

故 $\|f\| \le 2$, 因而 f 是 $C[a,b]$ 上的有界线性泛函. 另一方面, 取 $x(t) = \frac{2}{a-b}(t - \frac{a+b}{2})$, 知 $x(a) = 1, x(b) = -1, \|x\| = 1$, 且 $|f(x)| = 2$. 故

$$\|f\| = 2$$

□

4. 求泛函 $f(x) = \int_0^1 \sqrt{t}x(t^2)dt$ 在以下两种情形下的范数 $\|f\|$.

(1) $x(t) \in C[0,1]$;

解

知

$$|f(x)| \le \int_0^1 \sqrt{t}|x(t^2)|dt \le \max_{[0,1]}|x| \cdot \int_0^1 \sqrt{t}dt = \frac{2}{3}\|x\|$$

另取 $x \equiv 1$, 知 $x \in C[0,1], \|x\| = 1$, 且 $|f(x)| = \frac{2}{3}$. 故 $\|f\| = \frac{2}{3}$. □

(2) $x(t) \in L^2[0,1]$.

解

令 $s = t^2$, 换元知

$$f(x) = \frac{1}{2}\int_0^1 s^{-\frac{1}{4}}x(s)ds \Rightarrow |f(x)| \le \frac{1}{2}\int_0^1 s^{-\frac{1}{4}}|x(s)|ds$$

由 Cauchy-Schwarz 不等式知

$$\frac{1}{2}\int_0^1 s^{-\frac{1}{4}}|x(s)|ds \le \frac{1}{2}(\int_0^1 s^{-\frac{1}{2}}ds)^{\frac{1}{2}}\|x\|_{L^2} = \frac{\sqrt{2}}{2}\|x\|_{L^2}$$

等号成立当且仅当 $x(s) = cs^{-\frac{1}{4}}$. 显见存在 $c > 0$ 使得 $\|x\|_{L^2} = 1$, 因而 $\|f\| = \frac{\sqrt{2}}{2}$. □

5. 设 $\phi(t) \in C[0,1]$, 在 $C[0,1]$ 上定义泛函

$$\Phi(f) = \int_0^1 \phi(t)f(t)dt, \quad \forall f \in C[0,1],$$

求 $\|\Phi\|$.

解

知

$$|\Phi(f)| \le \int_0^1 |\phi(t)||f(t)|dt \le \|\phi\|_{L^1}\|f\|$$

这说明 $\|\Phi\| \le \|\phi\|_{L^1}$. 当 ϕ 在 (0,1) 上有可列个零点时, 根据连续性知 $\phi(t) \equiv 0$, 从而显见 $\|\Phi\| = 0$. 下面讨论 ϕ 在 (0,1) 上只有有限个零点的情况. 当 $\phi(t)$ 在 $(0,1)$ 上无零点, 知取用 $f(t) = \operatorname{sgn}\phi(t)$ 即得 $|\Phi(f)| = \|\phi\|_{L^1}, \|f\| = 1$. 当 $\phi(t)$ 在 $(0,1)$ 上有唯一零点 t_0, 任取 $\varepsilon > 0$ 足够小, 使得 $(t_0-\varepsilon, t_0+\varepsilon) \subset (0,1)$, 记

$$f_\varepsilon(t) = \begin{cases} \operatorname{sgn}\phi(t), & t \in [0,1]\backslash(t_0-\varepsilon, t_0+\varepsilon) \\ \frac{\operatorname{sgn}\phi(t_0-\varepsilon)-\operatorname{sgn}\phi(t_0+\varepsilon)}{2\varepsilon}(t-t_0), & t \in [t_0-\varepsilon, t_0+\varepsilon] \end{cases}$$

则

$$\begin{aligned} |\Phi(f_\varepsilon)| &\ge (\int_0^{t_0-\varepsilon} + \int_{t_0+\varepsilon}^1)|\phi(t)|dt - \int_{t_0-\varepsilon}^{t_0+\varepsilon}|\phi(t)||f_\varepsilon(t)| \\ &\ge \|\phi\|_{L^1} - 4\varepsilon \max_{[t_0-\varepsilon, t_0+\varepsilon]}|\phi| \end{aligned}$$

亦即存在序列 $\{f_\varepsilon\} \subset C[0,1]$ 使得 $\|f_\varepsilon\| = 1$, 且 $|\Phi(f_\varepsilon)| \to \|\phi\|_{L^1}(\varepsilon \to 0^+)$, 这便说明 $\|\Phi\| = \|\phi\|_{L^1}$. 最后, 若 $\phi(t)$ 在 (0,1) 上有多个零点 $t_1, \cdots, t_n$, 类似构造 f_ε 亦得 $\|\Phi\| = \|\phi\|_{L^1}$, 综上得到:

$$\|\Phi\| = \|\phi\|_{L^1}$$

□

† 订正

上述解答中" 可列个零点" 推出 $\phi(t) \equiv 0$ 这一步显然是有问题的. 事实上, 该题不需要具体构造出连续函数列 f_ε, 而是借助 Luzin 定理即可完成. 取 $f(t) = \operatorname{sgn}\phi(t)$, 因为 $[0,1]$ 是 $\mathbb{R}$ 上的紧集, 故根据 Luzin 定理知存在连续函数列 $f_\varepsilon(t)$ 使得

$$f_\varepsilon(t) \to f(t)(\forall t \in [0,1]), m(\{t : f_\varepsilon(t) \ne f(t)\}) \to 0, \quad \varepsilon \to 0^+$$

因此

$$\begin{aligned}|\Phi(f_\varepsilon)| &= |\int_0^1 \phi(t)f_\varepsilon(t)dt| \\ &\geq \int_0^1 |\phi(t)|\cdot|f(t)|dt - \int_0^1 |\phi(t)|\cdot|f_\varepsilon(t)-f(t)|dt \\ &\geq \int_0^1 |\phi(t)|dt - m(\{t: f_\varepsilon(t)\neq f(t)\})\max_{t\in[0,1]}|\phi(t)| \to \|\phi\|_{L^1}, \quad \varepsilon\to 0^+\end{aligned}$$

另一方面容易说明 $|\Phi(f)| \leq \|\phi\|_{L^1}\|f\|$, 因而 $\|\Phi\| = \|\phi\|_{L^1}$. □

3.8.2 第十次作业

1. 设 f 是 $\mathscr{X}$ 上的非零有界线性泛函, 令

$$d = \inf\{\|x\| : f(x) = 1, x\in\mathscr{X}\}$$

求证: $\|f\| = \frac{1}{d}$.

证明

任取 $x\in\mathscr{X}$ 使得 $f(x)\neq 0$, 根据 f 的线性性知 $f(\frac{x}{f(x)}) = 1$, 于是由 d 的定义知:

$$d \leq \|\frac{x}{f(x)}\| = \frac{\|x\|}{|f(x)|} \Rightarrow |f(x)| \leq \frac{\|x\|}{d} \Rightarrow \|f\| \leq \frac{1}{d}$$

另一方面, 由 d 作为下确界的定义知对任意 $\varepsilon > 0$, 存在 $n_\varepsilon\in\mathbb{N}$ 使得 $f(x_{n_\varepsilon}) = 1$, 且

$$\|x_{n_\varepsilon}\| < d+\varepsilon = (d+\varepsilon)|f(x_{n_\varepsilon})| \Rightarrow |f(x_{n_\varepsilon})| > \frac{\|x_{n_\varepsilon}\|}{d+\varepsilon} \Rightarrow \|f\| > \frac{1}{d+\varepsilon}$$

根据 ε 的任意性知 $\|f\| \geq \frac{1}{d}$, 故 $\|f\| = \frac{1}{d}$. □

2. 设 $f\in\mathscr{X}^*$, 求证: $\forall\varepsilon > 0, \exists x_0\in\mathscr{X}$, 使得 $f(x_0) = \|f\|$, 且 $\|x_0\| < 1+\varepsilon$.

证明

根据算子范数的定义知

$$\|f\| = \sup_{\|x\|=1}|f(x)|$$

根据上确界定义知

$$\forall\varepsilon > 0\exists x_\varepsilon\in\mathscr{X}(\|x_\varepsilon\| = 1 \wedge |f(x_\varepsilon)| \geq \|f\| - \frac{1}{2}\varepsilon|f(x_\varepsilon)|)$$

不妨设 $f(x_\varepsilon) > 0$, 取 $x_0 = \frac{x_\varepsilon}{f(x_\varepsilon)}\cdot\|f\|$, 知 $f(x_0) = \frac{f(x_\varepsilon)}{f(x_\varepsilon)}\cdot\|f\| = \|f\|$, 且

$$\|x_0\| = \frac{\|x_\varepsilon\|}{|f(x_\varepsilon)|}\cdot\|f\| \leq \frac{|f(x_\varepsilon)| + \frac{1}{2}\varepsilon|f(x_\varepsilon)|}{|f(x_\varepsilon)|} = 1+\frac{1}{2}\varepsilon < 1+\varepsilon$$

命题即证. □

3. 设 $T:\mathscr{X}\to\mathscr{Y}$ 是线性的, 令

$$N(T) := \{x\in\mathscr{X} : Tx = \theta\}.$$

(1) 若 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 求证: $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间.

证明

显见 $T\theta = \theta$, 故 $\theta\in N(T)$, 亦即 $N(T)\neq\varnothing$. 若 $x_1, x_2\in N(T)$, 则任取 $a, b\in\mathbb{K}$ 知

$$T(ax_1+bx_2) = aTx_1 + bTx_2 = \theta \Rightarrow ax_1+bx_2\in N(T)$$

故 $N(T)$ 是 $\mathscr{X}$ 的线性子空间. 最后, 只要 $x\in\overline{N(T)}$, 根据极限点的定义知存在 $\{x_n\}\subset N(T)$ 使得 $\|x_n - x\|_{\mathscr{X}}\to 0(n\to\infty)$, 根据 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$ 知 $\|Tx_n - Tx\|_{\mathscr{Y}}\to 0(n\to\infty)$, 又因为 $Tx_n = \theta(\forall n\in\mathbb{N})$, 故 $Tx = \theta$, 亦即 $x\in N(T)$, 故 $\overline{N}(T)\subset N(T)$, 从而 $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间. □

(2) 问 $N(T)$ 是 $\mathscr{X}$ 的闭线性子空间能否推出 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$?

解

未必如此. 设 $\mathscr{X} = (l^1, \|\cdot\|_{l^\infty})$, 另取 $a = (1, -1, 0, \cdots) \in \mathscr{X}$ 与线性泛函:

$$f(x) = \sum_{n=1}^{\infty} \xi_n, \quad \forall x = (\xi_1, \cdots, \xi_n, \cdots) \in \mathscr{X}$$

知 $\|a\|_{l^\infty} = 1, f(a) = 0$. 取用:

$$Tx := x - af(x), \quad \forall x = (\xi_1, \cdots, \xi_n, \cdots) \in \mathscr{X}$$

现讨论 $N(T)$, 令 $Tx = x - af(x) = \theta$, 得到 $x = af(x)$, 两边作用 f 得 $f(x) = f(a)f(x) = 0$, 因而 $f(x) = 0$, 从而 $Tx = x = \theta$, 故 $N(T) \subset \{\theta\}$. 这便得到 $N(T) = \{\theta\}$.

下面说明 T 的无界性, 首先说明线性泛函 f 的无界性. 取 $e_k = (\underbrace{0, \cdots, 0, 1}_{k}, 0, \cdots)$, 并设 $x_n = \sum\limits_{k=1}^{n} e_k \in \mathscr{X}$, 知

$$\|x_n\| = 1, \quad f(x_n) = n, \quad \forall n \in \mathbb{N}$$

故 $\|f\| \geq n (\forall n \in \mathbb{N})$, 因而 f 无界.

再证明 T 无界, 用反证法. 如若 T 有界, 知

$$\exists M > 0 \forall x \in \mathscr{X} (\|Tx\| \leq M\|x\|)$$

则

$$\begin{aligned}
&Tx = x - af(x) \\
\Rightarrow & af(x) = x - Tx \\
\Rightarrow & \|a\||f(x)| \leq \|x\| + \|Tx\| \leq (1 + M)\|x\| \\
\Rightarrow & |f(x)| \leq (1 + M)\|x\|
\end{aligned}$$

其中最后一步是因为 $\|a\| = 1$. 故 f 有界, 但这是矛盾的! 从而 $T \notin \mathscr{L}(\mathscr{X}, \mathscr{Y})$. □

(3) 若 f 是线性泛函, 求证:

$$f \in \mathscr{X}^* \Leftrightarrow N(f) \text{是闭线性子空间.}$$

证明

当 $f \in \mathscr{X}^*$, 这便是 (1) 所证, 下设 $N(f)$ 是闭线性子空间, 往证 $f \in \mathscr{X}^*$, 根据 f 的线性性知只需说明 f 有界. 用反证法, 若 f 无界, 亦即

$$\forall n \in \mathbb{N} \exists x_n \in \mathscr{X} (\|x_n\| = 1 \wedge |f(x_n)| > n)$$

现取 $y_n = \frac{x_n}{f(x_n)} - \frac{x_1}{f(x_1)}$, 知 $f(y_n) = 1 - 1 = 0 \Rightarrow y_n \in N(f)$. 同时 $\|\frac{x_n}{|f(x_n)|}\| < \frac{\|x_n\|}{n} \to 0 (n \to \infty)$, 故 $\|y_n + \frac{x_1}{f(x_1)}\| \to 0 (n \to \infty)$, 亦即 $y_n \to -\frac{x_1}{f(x_1)} (n \to \infty)$. 根据 $N(f)$ 的闭性知 $-\frac{x_1}{f(x_1)} \in N(f)$, 但 $f(-\frac{x_1}{f(x_1)}) = -1 \neq 0$, 矛盾! 故 f 有界, 因而 $f \in \mathscr{X}^*$. □

3.8.3 第十一次作业

1. 设 T 为定义在复 Hilbert 空间 X 上的有界线性算子, 若存在常数 $\alpha_0 > 0$, 使 $(Tx, x) \geq \alpha_0(x, x)$, 则称 T 为正定的. 证明: 正定算子 T 必有有界逆算子 T^{-1}, 并且

$$\|T^{-1}\| \leq \frac{1}{\alpha_0}.$$

证明

当 T 是正定算子, 知其必为单射, 因为

$$Tx = \theta \Rightarrow (Tx, x) = \theta \geq \alpha_0(x, x) \Rightarrow \|x\| = 0 \Rightarrow x = \theta$$

下面说明 T 为满射, 先说明 $R(T)$ 是闭的. 任取 $w \in \overline{R(T)}$, 根据闭包的定义知存在 $\{v_n\}_{n\in\mathbb{N}} \in X$ 使得

$$w = \lim_{n\to\infty} Tv_n$$

进而

$$\alpha_0\|v_{n+p} - v_n\|^2 \leq (T(v_{n+p} - v_n), v_{n+p} - v_n) \leq \|T(v_{n+p} - v_n)\| \cdot \|v_{n+p} - v_n\|$$

故

$$\|v_{n+p} - v_n\| \leq \frac{1}{\alpha_0}\|T(v_{n+p} - v_p)\|$$

由 $\{Tv_n\}_{n\in\mathbb{N}} \subset X$ 收敛知其为基本列, 因而 $\{v_n\}_{n\in\mathbb{N}} \subset X$ 是基本列, 由 X 的完备性知可设 $v_n \to v^*(n \to \infty)$, 进而由 T 的有界性 (进而是连续性) 知

$$w = \lim_{n\to\infty} Tv_n = T(\lim_{n\to\infty} v_n) = Tv^*$$

从而 $w \in R(T)$, 因而 $R(T)$ 是闭的.

再证明 $R(T) = X$, 这只需证明 $R(T)^\perp = \{\theta\}$ 即可. 若 $w \in R(A)^\perp$, 则

$$\forall v \in X((Tv, w) = 0)$$

取 $v = w$ 知

$$(Tw, w) = 0 \geq \alpha_0(w, w) \Rightarrow w = \theta$$

故 $R(T) = X$, 因而 T 是满射, 从而由 Banach 逆算子定理知 T 有有界线性逆 T^{-1}, 且任取 $x \in X$ 有

$$\alpha_0\|x\|^2 \leq (Tx, x) \leq \|Tx\| \cdot \|x\| \Rightarrow \|x\| \leq \frac{1}{\alpha_0}\|Tx\| \Rightarrow \|T^{-1}x\| \leq \frac{1}{\alpha_0}\|x\| \Rightarrow \|T^{-1}\| \leq \frac{1}{\alpha_0}$$

□

2. 设数列 $\{a_k\}$ 满足 $\sup\limits_{k\geq 1}|a_k| < \infty$. 作 l^1 到自身的算子 T 如下:

$$\forall x = \{\xi_k\} \in l^1, \quad y = Tx = \{a_k\xi_k\}$$

证明: T 存在有界逆算子的充分必要条件是 $\inf\limits_{k\geq 1}|a_k| > 0$.

证明

当 $\inf\limits_{k\geq 1}|a_k| > 0$, 取

$$W : l^1 \to l^1, \quad l^1 \ni \{\xi_k\} \mapsto \{\frac{\xi_k}{a_k}\} \in l^1$$

显见 $TWx = WTx = x$, 且

$$\|Wx\| = \sum_{k=1}^{\infty}|\frac{\xi_k}{a_k}| \leq \sup_{k\geq 1}|\frac{1}{a_k}|\|x\|_{l^1} = \frac{1}{\inf\limits_{k\geq 1}|a_k|}\|x\|_{l^1}, \quad \forall x = \{\xi_k\}$$

故 W 是 T 的有界逆算子.

当 $\inf\limits_{k\geq 1}|a_k| = 0$, 若存在 $n > 0$ 使得 $|a_n| = 0 = a_n$, 则取 $x_1 = (\underbrace{0, \cdots, 0, 1}_{n}, 0, \cdots), x_2 = 0$, 知 $Tx_1 = Tx_2 = 0$, 因而 T 并非单射, 其不存在逆算子. 若 $\forall k \in \mathbb{N}(|a_k| \neq 0)$, 则必存在子列 $\{a_{k_n}\}$ 满足 $\lim\limits_{n\to\infty}|a_{k_n}| = 0$, 为简便不妨就设 $\lim\limits_{k\to\infty}|a_k| = 0$. 现取 $x_n = (\underbrace{0, \cdots, 0, \frac{1}{a_n}}_{n}, 0, \cdots) \in l^1(n = 1, 2, \cdots)$, 设 T 存在有界逆算子 T^{-1}, 知至少应有 $\forall x \in l^1(T^{-1}Tx = x)$. 注意 $Tx_n = 1(n = 1, 2, \cdots)$, 从而

$$\|T^{-1}Tx\| = \|x\| = \frac{1}{|a_n|} = \frac{1}{|a_n|}\|Tx\|, \quad \forall n = 1, 2, \cdots$$

这说明

$$\|T^{-1}\| \geq \frac{1}{|a_n|}, \quad \forall n = 1, 2, \cdots$$

因而 T^{-1} 无界, 矛盾! 命题即证. □

3.8.4 第十二次作业

1. 设 $\mathscr{X}$ 是复线性空间, p 是 $\mathscr{X}$ 上的半范数. $\forall x_0 \in \mathscr{X}, p(x_0) \neq 0$. 求证: 存在 $\mathscr{X}$ 上的线性泛函 f 满足

(1) $f(x_0) = 1$;

(2) $|f(x)| \leq \frac{p(x)}{p(x_0)}(\forall x \in \mathscr{X})$.

证明

由 p 是 $\mathscr{X}$ 上的半范数显见 $\frac{p(x)}{p(x_0)}$ 也是 $\mathscr{X}$ 上的半范数. 取 $\mathscr{X}_0 = \{\lambda x_0 : \lambda \in \mathbb{C}\}$, 并定义 $f(x) = f(\lambda x_0) = \lambda \in \mathbb{C}(x \in \mathscr{X}_0)$. 显见 f 是线性泛函, $f(x_0) = 1$, 且

$$\frac{p(x)}{p(x_0)} = \frac{p(\lambda x_0)}{p(x_0)} = \frac{|\lambda| p(x_0)}{p(x_0)} = |\lambda| = |f(x)|$$

亦即 $|f(x)| \leq \frac{p(x)}{p(x_0)}(\forall x \in \mathscr{X}_0)$. 于是由复 Hahn-Banach 定理知存在 $\mathscr{X}$ 上的线性泛函 $\widetilde{f}$ 满足

(1) $\forall x \in \mathscr{X}_0(\widetilde{f}(x) = f(x))$;

(2) $\forall x \in \mathscr{X}(|\widetilde{f}(x)| \leq \frac{p(x)}{p(x_0)})$

特别有 $\widetilde{f}(x_0) = f(x_0) = 1$, 命题即证. □

2. 设 $\mathscr{X}$ 是 B^* 空间, $\{x_n\}(n = 1, 2, \cdots)$ 是 $\mathscr{X}$ 中的点列. 如果 $\forall f \in \mathscr{X}^*$, 数列 $\{f(x_n)\}$ 有界, 求证: $\{x_n\}$ 在 $\mathscr{X}$ 内有界.

证明

首先证明

$$\forall x_0 \in \mathscr{X}(\|x_0\| = \sup_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} |f(x_0)|) \tag{3.97}$$

一方面, $\forall f \in \mathscr{X}^*(\|f\| = 1)$ 说明 $\forall f \in \mathscr{X}^*(\sup\limits_{x \in \mathscr{X}} |f(x)| = \|x\|)$, 从而 $\forall f \in \mathscr{X}^*(|f(x_0)| \leq \|x_0\|)$, 亦即 $\sup\limits_{\substack{f \in \mathscr{X}^* \\ \|f\|=1}} |f(x_0)| \leq \|x_0\|$.

另一方面, 任取 $\theta \neq x_0 \in \mathscr{X}$, 知 $\mathscr{X}_0 = \{\lambda x_0 : \lambda \in \mathbb{R}\}$ 是 $\mathscr{X}$ 的闭线性子空间, 定义 $\forall x \in \mathscr{X}_0(f(x) = f(\lambda x_0) = \lambda \|x_0\|)$, 则 f 首先是线性泛函, 下证 f 有界. 注意 $\|\cdot\|$ 是 $\mathscr{X}$ 上的次线性泛函, 且

$$\frac{|f(x)|}{\|x\|} = \frac{|f(\lambda x_0)|}{\|\lambda x_0\|} = \frac{|\lambda|\|x_0\||}{|\lambda|\|x_0\|} = 1 \Rightarrow \|f\| = 1$$

故一方面 f 有界, 另一方面 $|f|$ 被 $\mathscr{X}$ 上的一个次线性泛函控制, 因而由 Hahn-Banach 定理知存在 $\widetilde{f} \in \mathscr{X}^*$ 使得

(1) $\forall x \in \mathscr{X}_0(\widetilde{f}(x) = f(x))$;

(2) $\forall x \in \mathscr{X}(|\widetilde{f}(x)| \leq \|x\|)$.

特别有 $|\widetilde{f}(x_0)| = |f(x_0)| = \|x_0\|$.

综上即证(3.97)式. 现注意任取 $x_0 \in \mathscr{X}$, 总存在 $\mathscr{X}^{**}$ 中的元素 x_0^{**}, 使得

$$\langle x_0^{**}, f \rangle = f(x_0), \quad \forall f \in \mathscr{X}^*$$

且由(3.97)式知 $\|x_0^{**}\|_{\mathscr{X}^{**}} = \|x_0\|_{\mathscr{X}}$. 因为 $\mathscr{X}$ 是 B^* 空间, 故 $\mathscr{X}^* = \mathscr{L}(\mathscr{X}, \mathbb{R})$ 是 B 空间, 进一步 $\mathscr{X}^{**} = \mathscr{L}(\mathscr{X}^*, \mathbb{R})$ 是 B 空间, 从而条件即

$$\forall f \in \mathscr{X}^*(\sup_{n \geq 1} |\langle x_n^{**}, f \rangle| < \infty)$$

根据共鸣定理知存在常数 M 使得 $\sup\limits_{n \geq 1} \|x_n^{**}\| < \infty$, 亦即 $\sup\limits_{n \geq 1} \|x_n\| < \infty$, 命题得证. □

3. 设 $\mathscr{X}_0$ 是 B^* 空间 $\mathscr{X}$ 的闭子空间, 求证:

$$\rho(x, \mathscr{X}_0) = \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\}, \quad \forall x \in \mathscr{X}$$

其中 $\rho(x, \mathscr{X}_0) = \inf\limits_{y \in \mathscr{X}_0} \|x - y\|$.

证明

当 $x \in \mathscr{X}_0$ 或 $\mathscr{X}_0 = \mathscr{X}$ 时欲证式显然成立, 下设 $x \in \mathscr{X} \backslash \mathscr{X}_0, \mathscr{X}_0 \subsetneq \mathscr{X}$. 一方面, 由 $\|f\| = 1$ 知 $\sup\limits_{x \in \mathscr{X}} |f(x)| = \|x\|$. 现任取 $x \in \mathscr{X}, y \in \mathscr{X}_0$, 由 $f(\mathscr{X}_0) = 0$ 知 $|f(x)| = |f(x-y)| \leq \|x-y\|$, 亦即

$$\sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \leq \inf_{y \in \mathscr{X}_0} \|x-y\| = \rho(x, \mathscr{X}_0), \quad \forall x \in \mathscr{X}$$

另一方面, 取 $x_0 \in \mathscr{X} \backslash \mathscr{X}_0$, 设 $\mathscr{X}_1 = \{y + \lambda x_0 : y \in \mathscr{X}_0, \lambda \in \mathbb{R}\}$, 令 $f(x) = f(y + \lambda x_0) = \lambda \rho(x_0, \mathscr{X}_0)$, 显见 f 是线性的. 注意 $\rho(\cdot, \mathscr{X}_0)$ 是 $\mathscr{X}$ 上的次线性泛函, 且

$$\frac{|f(x)|}{\rho(x, \mathscr{X}_0)} = \frac{|f(y+\lambda x_0)|}{\rho(y+\lambda x_0, \mathscr{X}_0)} = \frac{|\lambda \rho(x_0, \mathscr{X}_0)|}{\rho(\lambda x_0, \mathscr{X}_0)} = 1, \quad \forall x \in \mathscr{X}_1$$

这说明 $|f|$ 被 $\mathscr{X}$ 上的一个次线性泛函控制, 因而由 Hahn-Banach 定理知存在 $\widetilde{f} \in \mathscr{X}^*$, 使得

(1) $\forall x \in \mathscr{X}_1 (\widetilde{f}(x) = f(x))$;

(2) $\forall x \in \mathscr{X} (|\widetilde{f}(x)| \leq \rho(x, \mathscr{X}_0))$.

下面证明 $\|\widetilde{f}\| = 1$. 已知

$$\forall x \in \mathscr{X} (|\widetilde{f}(x)| \leq \rho(x, \mathscr{X}_0) = \inf_{y \in \mathscr{X}_0} \|x-y\| \leq \|x\|) \Rightarrow \|f\| \leq 1$$

现在先证明

$$\forall x \in \mathscr{X} (|\widetilde{f}(x)| \leq \rho(x, \mathscr{X}_0) \|\widetilde{f}\|)$$

这是因为根据 ρ 作为下确界的定义知对任意的 $\varepsilon > 0$, 存在 $x_\varepsilon \in \mathscr{X}_0$ 使得

$$\rho(x, \mathscr{X}_0) \leq \rho(x, x_\varepsilon) \leq \rho(x, \mathscr{X}_0) + \varepsilon$$

于是

$$|\widetilde{f}(x)| = |\widetilde{f}(x - x_\varepsilon)| \leq \|\widetilde{f}\| \cdot \|x - x_\varepsilon\| \leq \|\widetilde{f}\| (\rho(x, \mathscr{X}_0) + \varepsilon)$$

由 ε 的任意性即得 $|\widetilde{f}(x)| \leq \|\widetilde{f}\| \rho(x, \mathscr{X}_0)$. 特别取 $x = x_0$ 有

$$|f(x_0)| = \rho(x_0, \mathscr{X}_0) \leq \rho(x_0, \mathscr{X}_0) \|\widetilde{f}\| \Rightarrow \|\widetilde{f}\| \geq 1$$

综上, $\|\widetilde{f}\| = 1$. 最后对这个固定的 $x_0 \in \mathscr{X} \backslash \mathscr{X}_0$ 有

$$\sup\{|f(x_0)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \geq |\widetilde{f}(x_0)| = |f(x_0)| = \rho(x_0, \mathscr{X}_0)$$

这说明任取 $x \in \mathscr{X} \backslash \mathscr{X}_0$, 总可以依照上述过程构造 $\widetilde{f}$, 使得 $|\widetilde{f}(x)| = \rho(x, \mathscr{X}_0)$, 这便有

$$\sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\} \geq \rho(x, \mathscr{X}_0), \quad \forall x \in \mathscr{X}$$

综上即得

$$\rho(x, \mathscr{X}_0) = \sup\{|f(x)| : f \in \mathscr{X}^*, \|f\| = 1, f(\mathscr{X}_0) = 0\}, \quad \forall x \in \mathscr{X}$$

命题得证. □

4. 设 $\mathscr{X}$ 是 B^* 空间. 给定 $\mathscr{X}$ 中 n 个线性无关的元素 $x_1, x_2, \cdots, x_n$ 与数域 $\mathbb{K}$ 中的 n 个数 $C_1, C_2, \cdots, C_n$, 及 $M > 0$. 求证: 为了 $\exists f \in \mathscr{X}^*$ 适合 $f(x_k) = C_k (k = 1, 2, \cdots, n)$, 以及 $\|f\| \leq M$, 必须且仅须对任意的 $\alpha_1, \alpha_2, \cdots, \alpha_n \in \mathbb{K}$, 有

$$|\sum_{k=1}^n \alpha_k C_k| \leq M \|\sum_{k=1}^n \alpha_k x_k\|.$$

证明

当 $\exists f \in \mathscr{X}^* ((f(x_k) = C_k (k = 1, 2, \cdots, n)) \wedge (\|f\| \leq M))$, 根据 f 的线性性知

$$|\sum_{k=1}^n \alpha_k C_k| = |f(\sum_{k=1}^n \alpha_k x_k)| \leq M \|\sum_{k=1}^n \alpha_k x_k\|, \quad \forall \alpha_1, \cdots, \alpha_n \in \mathbb{K}$$

当 $\forall\alpha_1,\cdots,\alpha_n\in\mathbb{K}(|\sum\limits_{k=1}^{n}\alpha_kC_k|\le M\|\sum\limits_{k=1}^{n}\alpha_kx_k\|)$, 记 $A=\mathrm{span}\{x_k\}_{k=1}^n$, 对任意 $y=\sum\limits_{k=1}^{n}y^kx_k\in A$, 定义

$$f(y)=\sum_{k=1}^{n}y^kC_k$$

f 的良定义性由 $x_1,\cdots,x_n$ 的线性无关性保证. 显见 f 是线性的, 又注意到

$$|f(y)|=|\sum_{k=1}^{n}y^kC_k|\le M\|\sum_{k=1}^{n}y^kx_k\|=M\|y\|\Rightarrow\|f\|\le M$$

现在把 f 延拓到 $\mathscr{X}$ 上, 知 $M\|\cdot\|$ 是 $\mathscr{X}$ 上的次线性泛函, 故根据 Hahn-Banach 定理知存在 $\widetilde{f}\in\mathscr{X}^*$ 使得

(1) $\forall x\in A(\widetilde{f}(x)=f(x))$;

(2) $\forall x\in\mathscr{X}(|\widetilde{f}(x)|\le M\|x\|)$,

特别有 $\widetilde{f}(x_k)=f(x_k)=C_k(k=1,2,\cdots,n)$, 命题即证. □

5. 给定 B^* 空间 $\mathscr{X}$ 中 n 个线性无关的元素 $x_1,x_2,\cdots,x_n$, 求证: $\exists f_1,f_2,\cdots,f_n\in\mathscr{X}^*$ 使得

$$\langle f_i,x_j\rangle=\delta_{ij},\quad i,j=1,2,\cdots,n$$

证明

根据 $x_1,\cdots,x_n$ 的线性无关性, 对 x_j 构造 $A_j=\mathrm{span}\{x_k\}_{k\in\{1,\cdots,n\}\setminus\{j\}}$, 根据习题 **3** 知存在 $\widetilde{f_j}\in\mathscr{X}^*$ 使得 $\widetilde{f_j}(x_j)=\rho(x_j,A_j),\widetilde{f_j}(x)=0(\forall x\in A_j)$. 再令 $f_j(x)=\frac{\widetilde{f_j}(x)}{\rho(x_j,A_j)}$ 即得 f_j 满足 $\langle f_j,x_i\rangle=\delta_{ij}(i,j=1,2,\cdots,n)$. 令 j 取遍 $1,\cdots,n$ 即得欲求. □

3.8.5 第十三次作业

1. 设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, T 是 $\mathscr{X}$ 到 $\mathscr{Y}$ 的线性算子, 又设对 $\forall g\in\mathscr{Y}^*$, $g(Tx)$ 是 $\mathscr{X}$ 上的有界线性泛函, 求证: T 是连续的.

证明

既然 $\mathscr{X}$ 完备, 只需证明 T 是闭算子即可. 设 $\mathscr{X}$ 中的基本列 $\{x_n\}_{n\in\mathbb{N}}$ 满足 $x_n\to x_0,Tx_n\to y_0(n\to\infty)$, 只需说明 $y_0=Tx_0$ 即可. 因为 $\forall g\in\mathscr{Y}^*(g(T\circledast)\in\mathscr{X}^*)$, 故

$$\lim_{n\to\infty}g(Tx_n)=g(Tx),\quad\forall g\in\mathscr{Y}^*$$

这说明 $Tx_n\rightharpoonup Tx$, 又因为 $Tx_n\to y(n\to\infty)$, 故必有 $Tx_n\rightharpoonup y(n\to\infty)$, 根据弱极限的唯一性即知 $y=Tx$, 从而 T 是闭算子, 进而根据闭图像定理知 $T\in\mathscr{L}(\mathscr{X},\mathscr{Y})$. □

2. 设 H 是 Hilbert 空间, 求证: 在 H 中 $x_n\to x(n\to\infty)$ 的充要条件是

(1) $\|x_n\|\to\|x\|(n\to\infty)$;

(2) $x_n\rightharpoonup x(n\to\infty)$.

证明

$x_n\to x(n\to\infty)$ 时的情况是显然的, 下设 (1), (2) 成立. 根据弱收敛的定义有

$$\lim_{n\to\infty}f(x_n)=f(x),\quad\forall f\in H^*$$

因为 H 是 Hilbert 空间, 根据 Riesz 表示定理, 对固定的 x 可取 $f_x(y)=(y,x)\in H^*$, 故

$$\lim_{n\to\infty}(x_n,x)=(x,x)=\|x\|^2$$

根据内积的对称性有

$$\lim_{n\to\infty}(x,x_n)=\|x\|^2$$

于是

$$\lim_{n\to\infty}\|x_n-x\|^2=(x_n-x,x_n-x)=\|x_n\|^2-(x_n,x)-(x,x_n)+\|x\|^2=0$$

故 $x_n \to x(n \to \infty)$, 命题即证. $\square$

第四章 紧算子与 Fredholm 算子

4.1 紧算子的定义和基本性质

4.1.1 知识梳理

定义 4.1.1 (紧算子)

设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, $A:\mathscr{X}\to\mathscr{Y}$ 是线性算子. 称 A 是紧算子, 如果 $\overline{A(B_1)}$ 在 $\mathscr{Y}$ 中是紧集, 其中 B_1 是 $\mathscr{X}$ 中的单位球. 一切紧算子的集合记作 $\mathfrak{C}(\mathscr{X},\mathscr{Y})$, 当 $\mathscr{X}=\mathscr{Y}$ 时记作 $\mathfrak{C}(\mathscr{X})$.

注 紧算子有下述两个等价定义:

(a) $A\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$ 当且仅当对 $\mathscr{X}$ 中的任意有界集 B, $\overline{A(B)}$ 在 $\mathscr{Y}$ 中都是紧集. 这是因为由 B 有界知存在 $\rho_B>0$ 使得 $B\subset B(\theta;\rho_B)$, 而由 $\overline{A(B_1)}$ 紧知 $\overline{A(B(\theta;\rho_B))}$ 紧, 且 $\overline{A(B)}\subset\overline{A(B(\theta;\rho_B))}$. 根据紧集与自列紧集等价定理2.3.3知 $\overline{A(B(\theta;\rho_B))}$ 是自列紧集, 因而 $\overline{A(B)}$ 首先是列紧集. 又因为 $\overline{A(B)}$ 是闭的, 故其为自列紧集, 从而 $\overline{A(B)}$ 是紧集.

(b) $A\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$ 当且仅当对任意有界点列 $\{x_n\}\subset\mathscr{X}$, $\{Ax_n\}$ 中有收敛子列. 这恰是紧集与自列紧集等价定理2.3.3的结论.

紧算子有下述简单性质.

命题 4.1.1

(i) $\mathfrak{C}(\mathscr{X},\mathscr{Y})\subset\mathscr{L}(\mathscr{X},\mathscr{Y})$.

(ii) 若 $A,B\in\mathfrak{C}(\mathscr{X},\mathscr{Y}),\alpha,\beta\in\mathbb{C}$, 则
$$\alpha A+\beta B\in\mathfrak{C}(\mathscr{X},\mathscr{Y}).$$

(iii) $\mathfrak{C}(\mathscr{X},\mathscr{Y})$ 在 $\mathscr{L}(\mathscr{X},\mathscr{Y})$ 中闭.

(iv) 设 $A\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$, 又设 $\mathscr{X}_0\subset\mathscr{X}$ 是一个闭线性子空间, 则 $A_0:=A|_{\mathscr{X}_0}\in\mathfrak{C}(\mathscr{X}_0,\mathscr{Y})$.

(v) 若 $A\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$, 则 $R(A)$ 可分.

(vi) 若 $A\in\mathscr{L}(\mathscr{X},\mathscr{Y})$, 而 $B\in\mathscr{L}(\mathscr{Y},\mathscr{Z})$, 且这两个算子中有一个是紧的, 则 $BA\in\mathfrak{C}(\mathscr{X},\mathscr{Z})$.

证明

(i) 因为 $y\mapsto\|y\|$ 是连续算子, 故 $\|y\|$ 在紧集 $\overline{A(B_1)}$ 上可取得最大值. 从而若 $A\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$, 根据紧算子的定义知
$$M:=\sup_{x\in B_1}\|Ax\|=\sup_{y\in\overline{A(B_1)}}\|y\|<\infty$$
从而
$$\|Ax\|\leq M\|x\|\quad(\forall x\in\mathscr{X})$$

(ii) 这根据紧算子的等价定义 (b) 即得.

(iii) 从闭空间的定义出发, 设 $T_n\in\mathfrak{C}(\mathscr{X},\mathscr{Y})(n=1,2,\cdots)$, 且 $\|T_n-T\|\to0(n\to\infty)$, 现在需要证明 $T\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$. 根据极限的定义知
$$\forall\varepsilon>0\exists N>0\forall n>N(\|T_n-T\|<\varepsilon)$$
取定 n, 既然 $\overline{T_n(B_1)}$ 是紧的, 知其首先是列紧的, 因而由 Hausdorff 定理(2.3.1)知其是完全有界集, 从而对任意的

$\varepsilon > 0$, $\overline{T_n(B_1)}$ 都有有穷的 $\varepsilon-$ 网, 设之为 $\{y_1, y_2, \cdots, y_m\}$, 根据 $\varepsilon-$ 网的定义与上述极限定义有

$$\overline{T(B_1)} \subset \bigcup_{i=1}^{m} B(y_i, 2\varepsilon)$$

由 ε 的任意性知 $\overline{T(B_1)}$ 对任意的 $\varepsilon > 0$ 有有穷的 $\varepsilon-$ 网, 进而由 Hausdorff 定理与其闭性即得 $\overline{T(B_1)}$ 是紧的.

(iv) 任取 $\{x_n\} \subset \mathscr{X}_0$ 是有界点列, 根据 $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$ 知 $\{Ax_n\}$ 有收敛子列. 又因为 $Ax_n = A_0 x_n$, 根据紧算子的等价定义 (b) 便有 $A_0 \in \mathfrak{C}(\mathscr{X}_0, \mathscr{Y})$.

(v) 知

$$R(A) = \bigcup_{n=1}^{\infty} A(\{x : |x| \leq n\}) = \bigcup_{n=1}^{\infty} nA(B_1)$$

现在根据 $A(B_1)$ 列紧, 由 Hausdorff 定理知 $A(B_1)$ 是完全有界的, 因而其可分, 从而 $R(A)$ 也可分.

(vi) 注意连续线性算子保有界性和紧性即可. □

定义 4.1.2 (全连续)

称 $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$ 是全连续的, 如果

$$x_n \rightharpoonup x \Rightarrow Ax_n \to Ax, \quad n \to \infty$$

命题 4.1.2

若 $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$, 则 A 是全连续的. 反之, 若 $\mathscr{X}$ 自反, 且 A 是全连续的, 则 $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$.

证明

当 $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$, 根据全连续的定义, 设 $x_n \rightharpoonup x(n \to \infty)$, 现在需要证明 $Ax_n \to Ax =: y(n \to \infty)$. 用反证法, 若

$$\exists \varepsilon_0 > 0 \forall N > 0 \exists n_i > N(\|Ax_{n_i} - y\| \geq \varepsilon_0)$$

由 $x_n \rightharpoonup x(n \to \infty)$ 与共鸣定理(3.4.8)知 $\{x_n\}$ 一致有界, 进而 $\{x_{n_i}\}$ 一致有界. 又因为 A 是紧算子, 根据紧算子的等价定义 (b) 知 $\{x_{n_i}\}$ 中存在子列 (为简便起见, 不妨就设该子列为 $\{x_{n_i}\}$ 本身) 满足 $Ax_{n_i} \to z$. 但任取 $y^* \in \mathscr{Y}^*$ 有

$$\langle y^*, Ax_{n_i} - y\rangle = \langle A^* y^*, x_{n_i} - x\rangle \to 0$$

这说明 $Ax_{n_i} \rightharpoonup y$, 进而 $y = z$, 矛盾! 故 A 是全连续的.

当 $\mathscr{X}$ 自反且 A 全连续, 根据 Eberlein-Smulian 定理(3.6.12)知 $\mathscr{X}$ 的单位 (闭) 球是弱 (自) 列紧的, 因而其任意有界集都是弱列紧的. 现设 $\{x_n\}$ 有界, 则其有弱收敛子列 $x_{n_i} \rightharpoonup x$. 由 A 全连续知 $Ax_{n_i} \to Ax$, 这恰是 A 作为紧算子的等价定义 (b). □

定理 4.1.1

$T \in \mathfrak{C}(\mathscr{X}, \mathscr{Y}) \Leftrightarrow T^* \in \mathfrak{C}(\mathscr{Y}^*, \mathscr{X}^*)$.

证明

当 $T \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$, 套用紧算子的等价定义 (b), 希望证明若 $\{y_n^*\}$ 是有界数列 (设 $\|y_n^*\| \leq M(n = 1, 2, \cdots)$), 则 $\{T^* y_n^*\}$ 中有收敛子列. 任取 $n \in \mathbb{N}$, 令

$$\varphi_n(y) := \langle y_n^*, y\rangle \quad (\forall y \in \overline{T(B_1)})$$

显见 $\varphi_n \in C(\overline{T(B_1)})$, 现在证明 $\{\varphi_n\}$ 作为 $C(\overline{T(B_1)})$ 上的函数列有一致收敛子列. 既然 $y \in \overline{T(B_1)}$, 可设 $y = Tx, x \in \overline{B_1}$, 有

$$|\varphi_n(y)| = |\langle y_n^*, y\rangle| \leq \|y_n^*\| \cdot \|y\| \leq M \cdot \|Tx\| \leq M \cdot \|T\| \cdot \|x\| = M\|T\|, \quad \forall n \in \mathbb{N}, \forall y \in \overline{T(B_1)}$$

这说明 $\{\varphi_n\}$ 是一致有界的. 而

$$|\varphi_n(y)-\varphi_n(z)|\le\|y_n^*\|\cdot\|y-z\|\le M\|y-z\|$$

这说明 $\{\varphi_n\}$ 是等度连续的. 故由 Arzela-Ascoli 定理(2.3.4)知 $\{\varphi_n\}$ 是列紧集, 从而可设其有收敛子列 $\{\varphi_{n_k}\}$, 根据 Cauchy 准则知此即

$$\forall\varepsilon>0\exists K>0\forall p,q>K(\|\varphi_{n_p}-\varphi_{n_q}\|_{C(\overline{T(B_1)})}=\max_{y\in\overline{T(B_1)}}|\varphi_{n_p}(y)-\varphi_{n_q}(y)|<\varepsilon)$$

根据 $\varphi_n(y)$ 的定义即得

$$\max_{y\in\overline{T(B_1)}}|\langle y_{n_p}^*-y_{n_q}^*,y\rangle|<\varepsilon$$

而

$$\langle(y_{n_p}^*-y_{n_q}^*),y\rangle=\langle T^*(y_{n_p}^*-y_{n_q}^*),x\rangle$$

这说明

$$\max_{x\in\overline{B_1}}|\langle T^*(y_{n_p}^*-y_{n_q}^*),x\rangle|<\varepsilon$$

而

$$\|T^*y_{n_p}^*-T^*y_{n_q}^*\|=\max_{x\in\overline{B_1}}|\langle T^*(y_{n_p}^*-y_{n_q}^*),x\rangle|$$

故 $\|T^*y_{n_p}^*-T^*y_{n_q}^*\|<\varepsilon$, 也即 $\{T^*y_{n_k}^*\}$ 是 Cauchy 列, 根据 $\mathscr{X}^*$ 是 B 空间知 $\{T^*y_{n_k}^*\}$ 收敛, 此即欲求.

当 $T^*\in\mathfrak{C}(\mathscr{Y}^*,\mathscr{X}^*)$, 根据前述结论知 $T^{**}\in\mathfrak{C}(\mathscr{X}^{**},\mathscr{Y}^{**})$. 知 $T=T^{**}|_{\mathscr{X}}$, 套用命题(4.1.1)(iv) 可得 $T\in\mathfrak{C}(\mathscr{X},\mathscr{Y})$. □

下面给出一些紧算子的例子.

例 4.1 设 $\Omega\subset\mathbb{R}^n$ 是一个有界闭集, $K\in C(\Omega\times\Omega)$, 取 $\mathscr{X}=\mathscr{Y}=C(\Omega)$. 若令

$$T:u\mapsto\int_\Omega K(x,y)u(y)dy\quad(\forall u\in C(\Omega))$$

则 $T\in\mathfrak{C}(\mathscr{X})$.

证明

只需证明 $\overline{T(B_1)}$ 是紧的, 又根据其闭性只需证明 $\overline{T(B_1)}$ 是列紧的. 注意 $\overline{T(B_1)}\subset C(\Omega)$, 故考虑用 Arzela-Ascoli 定理(2.3.4). 既然 Ω 是有界闭集, 而 $K\in C(\Omega\times\Omega)$, 知 $|K|$ 在 $\Omega\times\Omega$ 上可取最大值, 设

$$M:=\max_{x,y\in\Omega}|K(x,y)|$$

则

$$\|Tu\|\le M|\int_\Omega u(y)dy|\le M\cdot\|u\|\cdot\mu(\Omega)$$

其中 $\mu(\Omega)$ 是 Ω 的 Lebesgue 测度. 因为现在讨论的是 $\overline{T(B_1)}$, 故 $\|u\|\le 1$, 进而 $\|Tu\|\le M\mu(\Omega)$, 这说明 $\overline{T(B_1)}$ 一致有界.

注意 $K(x,y)$ 在 $\Omega\times\Omega$ 中一致连续, 根据定义有

$$\forall\varepsilon>0\exists\delta>0\forall y\in\Omega(|x-x'|<\delta\Rightarrow|K(x,y)-K(x',y)|<\varepsilon)$$

故

$$|(Tu)(x)-(Tu)(x')|\le\int_\Omega|K(x,y)-K(x',y)|\cdot|u(y)|dy\le\varepsilon\cdot\|u\|\cdot\mu(\Omega),\quad|x-x'|<\delta$$

这说明 $\overline{T(B_1)}$ 等度连续, 从而 $\overline{T(B_1)}$ 是列紧集, 因而由闭性知其为紧集, 从而根据定义有 $T\in\mathfrak{C}(\mathscr{X})$. □

下面讨论紧算子的构造.

定义 4.1.3 (有穷秩算子)

设 $T \in \mathscr{L}(\mathscr{X},\mathscr{Y})$, 若 $\dim R(T) < \infty$, 则称 T 是有穷秩算子, 一切有穷秩算子的集合记作 $F(\mathscr{X},\mathscr{Y})$. ♣

知

$$F(\mathscr{X},\mathscr{Y}) \subset \mathfrak{C}(\mathscr{X},\mathscr{Y}).$$

这是因为既然 $\dim R(T) < \infty$, 故可以将其视作 $\mathbb{R}^n$, 而 $\overline{T(B_1)}$ 作为其上的有界闭集必是紧集.

定义 4.1.4 (秩 1 算子)

设 $f \in \mathscr{X}^*, y \in \mathscr{Y}$, 用 $y \otimes f$ 表示下述算子:

$$x \mapsto \langle f, x\rangle y \quad (\forall x \in \mathscr{X})$$

称它为秩 1 算子. ♣

可以用秩 1 算子来表示 $F(\mathscr{X},\mathscr{Y})$.

定理 4.1.2

要使得 $T \in F(\mathscr{X},\mathscr{Y})$, 必须且仅须 $\exists y_i \in \mathscr{Y}, f_i \in \mathscr{X}^*(i = 1, 2, \cdots, n)$, 使得

$$T = \sum_{i=1}^{n} y_i \otimes f_i$$

♡

当 $\exists y_i \in \mathscr{Y}, f_i \in \mathscr{X}^*(i = 1, 2, \cdots, n)$, 使得

$$T = \sum_{i=1}^{n} y_i \otimes f_i$$

知此时 $R(T) = \mathrm{span}\{y_1, \cdots, y_n\}$, 故 $\dim R(T) < \infty$.

当 $T \in F(\mathscr{X},\mathscr{Y})$, 在 $R(T)$ 上取基 $\{y_1, y_2, \cdots, y_n\}$, 根据基的定义知

$$\forall x \in \mathscr{X} \exists | \{l_i(x)\}_{i=1}^n (Tx = \sum_{i=1}^{n} l_i(x) y_i)$$

下面证明 $l_i(i = 1, 2, \cdots, n)$ 是 $\mathscr{X}$ 上的连续线性泛函.

首先说明线性性, 这是因为 T 本身是线性的, 而 l_i 作为基表示元素是唯一的.

其次说明有界性. 注意到 $\|Tx\|$ 与 $\sum_{i=1}^{n} |l_i(x)|$ 都是 $R(T)$ 上的范数, 又因为 $\dim R(T) < \infty$, 故这两个范数等价, 于是存在 $M > 0$ 使得

$$\sum_{i=1}^{n} |l_i(x)| \leq M\|Tx\| \leq M\|T\| \cdot \|x\|, \quad \forall x \in \mathscr{X}$$

故 $l_i \in \mathscr{X}^*(i = 1, 2, \cdots, n)$, 也即存在 $f_i \in \mathscr{X}^*(i = 1, 2, \cdots, n)$ 满足

$$\langle f_i, x\rangle = l_i(x), \quad \forall x \in \mathscr{X}, i = 1, 2, \cdots, n$$

得到

$$Tx = \sum_{i=1}^{n} y_i \langle f_i, x\rangle = (\sum_{i=1}^{n} y_i \otimes f_i)(x), \quad \forall x \in \mathscr{X}$$

也即

$$T = \sum_{i=1}^{n} y_i \otimes f_i.$$

□

现在希望利用 $F(\mathscr{X},\mathscr{Y})$ 研究 $\mathfrak{C}(\mathscr{X},\mathscr{Y})$. 显见 $\overline{F(\mathscr{X},\mathscr{Y})} \subset \mathfrak{C}(\mathscr{X},\mathscr{Y})$, 下面探讨是否有 $\overline{F(\mathscr{X},\mathscr{Y})} = \mathfrak{C}(\mathscr{X},\mathscr{Y})$. 不妨设 $\mathscr{Y} = \mathscr{X}$.

(i) 当 $\mathscr{X}$ 是一个 Hilbert 空间, 则 $\overline{F(\mathscr{X})} = \mathfrak{C}(\mathscr{X})$ 成立. 这是因为当 $T \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$, 根据定义 $\overline{T(B_1)}$ 是紧的, 因而由 Hausdorff 定理知其完全有界. 任取 $\varepsilon > 0$, 知其有有穷的 $\varepsilon-$ 网 $\{y_1, y_2, \cdots, y_n\}$, 即

$$\overline{T(B_1)} \subset \bigcup_{i=1}^{n} B(y_i, \varepsilon) \tag{4.1}$$

令 $E_\varepsilon = \operatorname{span}\{y_1, y_2, \cdots, y_n\}$, 并令 P_ε 为到 E_ε 上的正交投影, 则显见 $P_\varepsilon T \in F(\mathscr{X}, \mathscr{Y})$. 从(4.1)式中可知

$$\forall x \in B_1 \exists y_i (1 \leq i \leq n)(\|Tx - y_i\| < \varepsilon)$$

从而

$$\|P_\varepsilon Tx - y_i\| = \|P_\varepsilon(Tx - y_i)\| < \varepsilon$$

得到

$$\|Tx - P_\varepsilon Tx\| \leq \|Tx - y_i\| + \|y_i - P_\varepsilon Tx\| < 2\varepsilon, \quad \forall x \in B_1$$

也即 $\|T - P_\varepsilon T\| \leq \varepsilon$. 由 ε 的任意性即得 $\mathfrak{C}(\mathscr{X}) = \overline{F(\mathscr{X})}$.

(ii) 当 $\mathscr{X}$ 是一个 Banach 空间, 根据命题(4.1.1)(v) 知, 只需讨论 $\mathscr{X} = R(T)$ 是可分空间的情形.

定义 4.1.5 (Schauder 基)

设 $\mathscr{X}$ 是可分的 Banach 空间, 称 $\{e_n\}_{n=1}^{\infty} \subset \mathscr{X}$ 为 $\mathscr{X}$ 的一组 Schauder 基, 如果 $\forall x \in \mathscr{X}$, 存在唯一的一个序列 $\{C_n(x)\}$, 使得在 $\mathscr{X}$ 中有

$$x = \lim_{N \to \infty} \sum_{n=1}^{N} C_n(x) e_n$$

♣

由于 $\forall n \in \mathscr{N}$, 对应 $x \mapsto C_n(x)$ 都是唯一的, 故 $C_n(x)$ 是 $\mathscr{X}$ 上的线性函数. 进一步可以先承认下述引理.

引理 4.1.1

$C_n(x)(n \in \mathbb{N})$ 是 $\mathscr{X}$ 上的连续泛函.

♡

由该引理可得下述定理.

定理 4.1.3

若可分 B 空间 $\mathscr{X}$ 上有一组 Schauder 基, 则

$$\overline{F(\mathscr{X})} = \mathfrak{C}(\mathscr{X}).$$

♡

证明

对任意的 $N \in \mathbb{N}$, 令

$$S_n x = \sum_{n=1}^{N} C_n(x) e_n, \quad \forall x \in \mathscr{X}$$

令 $R_N = I - S_n$[1], 根据 Schauder 基的定义与共鸣定理(3.4.8)知, 存在 $M > 0$ 使得

$$\|S_n\| \leq M$$

进而有

$$\|R_N\| \leq 1 + M$$

现在设 $T \in \mathfrak{C}(\mathscr{X})$, 任取 ε, 希望找到 $T_\varepsilon \in F(\mathscr{X})$ 使得 $\|T - T_\varepsilon\| < \varepsilon$. 因为 $\overline{T(B_1)}$ 是紧的, 故由 Hausdorff 定理知存在有穷的 $\varepsilon-$ 网 $\{y_1, y_2, \cdots, y_m\}$, 也即

$$\forall x \in B_1 \exists y_i (1 \leq i \leq m)(\|Tx - y_i\| < \varepsilon) \tag{4.2}$$

[1] 没看懂这里的 R_N 起什么作用.

又由 Schauder 基的定义知

$$\exists N \in \mathbb{N}(\|y_j - S_N y_j\| < \varepsilon, j = 1, 2, \cdots, m) \tag{4.3}$$

因为 $\|S_N\| \leq M$, 故

$$\|S_N(Tx) - S_N y_i\| \leq \|S_N\| \cdot \|Tx - y_i\| < M\varepsilon \tag{4.4}$$

联立(4.2),(4.3),(4.4)三式得

$$\|Tx - (S_N T)x\| < (M+2)\varepsilon$$

取 $T_\varepsilon = S_N T$, 由 ε 的任意性即得欲证.

最后补充引理(4.1.1)的证明. 在 $\mathscr{X}$ 上引入另一个范数:

$$\Box x \Box = \sup_{N \in \mathbb{N}} \|S_N x\|$$

其中 $S_N(x) = \sum\limits_{n=1}^{N} C_n(x) e_n (\forall x \in \mathscr{X})$. 可以证明 $\mathscr{X}$ 按 $\Box \cdot \Box$ 是完备的, 且

$$\|x\| = \lim_{N \to \infty} \|S_N x\| \leq \Box x \Box, \quad \forall x \in \mathscr{X}$$

由等价范数定理(3.4.1)知, 存在 $M_1 > 0$ 使得

$$\Box x \Box \leq M_1 \|x\|, \quad \forall x \in \mathscr{X}$$

故对任意的 $n \in \mathbb{N}$ 都有

$$\|C_n(x) e_n\| = \|S_n x - S_{n-1} x\| \leq 2 \Box x \Box \leq 2M_1 \|x\|, \quad \forall x \in \mathscr{X}$$

这说明

$$|C_n(x)| \leq 2M_1 \|e_n\|^{-1} \cdot \|x\|, \quad \forall x \in \mathscr{X}, n \in \mathbb{N}$$

故 $C_n(x)(n \in \mathbb{N})$ 是连续的. □

注 可分的 Banach 空间不一定都有 Schauder 基.

4.1.2 习题

练习 4.1 设 $\mathscr{X}$ 是一个无穷维 B 空间, 求证: 若 $A \in \mathfrak{C}(\mathscr{X})$, 则 A 没有有界逆.

证明

既然 $\mathscr{X}$ 是无穷维的, 知其单位球 B_1 不是列紧集. 如若 A 有有界逆 A^{-1}, 则 $A^{-1} \in \mathscr{L}(\mathscr{X})$, 根据连续算子保紧性知 $A^{-1}(\overline{A(B_1)})$ 是紧集, 因而 $B_1 \subset A^{-1}(\overline{A(B_1)})$ 至少是列紧集, 矛盾! 命题即证.

练习 4.2 设 $\mathscr{X}$ 是一个 B 空间, $A \in \mathscr{L}(\mathscr{X})$ 满足

$$\|Ax\| \geq \alpha \|x\|, \quad \forall x \in \mathscr{X}$$

其中 α 是正常数. 求证: $A \in \mathfrak{C}(\mathscr{X})$ 的充要条件是 $\mathscr{X}$ 是有穷维的.

证明

既然 $\|Ax\| \geq \alpha \|x\| (\alpha > 0)$, 知 A 是单射, 进而其有有界逆, 根据练习(4.1)知 $\mathscr{X}$ 是有穷维的.

当 $\mathscr{X}$ 有穷维, 知其单位球 B_1 是列紧的, 根据 $A \in \mathscr{L}(\mathscr{X})$ 知 $A(B_1)$ 是列紧的, 因而 $\overline{A(B_1)}$ 是列紧的. 又因为 $\overline{A(B_1)}$ 是闭的, 故其是紧集, 得到 $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$.

练习 4.3 设 $\mathscr{X}, \mathscr{Y}$ 是 B 空间, $A \in \mathscr{L}(\mathscr{X}, \mathscr{Y}), K \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$, 如果 $R(A) \subset R(K)$, 求证: $A \in \mathfrak{C}(\mathscr{X}, \mathscr{Y})$.

证明[2]

[2] 证明困在了 K 是否有有界逆上, 只要可逆即可说明命题.

既然 $K \in \mathfrak{C}(\mathscr{X},\mathscr{Y})$, 根据紧算子的第二个等价定义知对任意的有界序列 $\{x_n\} \in \mathscr{X}$, $\{Kx_n\} \subset R(K)$ 均有收敛子列.

练习 4.4 设 H 是 Hilbert 空间, $A : H \to H$ 是紧算子, 又设 $x_n \rightharpoonup x_0, y_n \rightharpoonup y_0$, 求证:

$$(x_n, Ay_n) \to (x_0, Ay_0) \quad (n \to \infty)$$

证明

因为 $A \in \mathfrak{C}(H)$, 知 A 首先是全连续的, 因而由 $y_n \rightharpoonup y_0$ 知 $Ay_n \to Ay_0$. 根据 $x_n \rightharpoonup x_0$ 与共鸣定理知 $\|x_n\|$ 一致有界, 设 $\|x_n\| < M$, 可得

$$\forall \varepsilon > 0 \exists N > 0 \forall n > N (|(x_n, Ay_n) - (x_n, Ay_0)| \leq \|x_n\| \cdot \|Ay_n - Ay_0\| \leq M\varepsilon)$$

再证明 $(x_n, Ay_0) \to (x_0, Ay_0)(n \to \infty)$, 而这根据 Riesz 表示定理与弱收敛的定义即得. 故

$$|(x_n, Ay_n) - (x_0, Ay_0)| \leq |(x_n, Ay_n) - (x_n, Ay_0)| + |(x_n, Ay_0) - (x_0, Ay_0)| \to 0, \quad n \to \infty$$

命题即证.

练习 4.5 设 $\mathscr{X},\mathscr{Y}$ 是 B 空间, $A \in \mathscr{L}(\mathscr{X},\mathscr{Y})$. 如果 $R(A)$ 闭且 $\dim R(A) = \infty$, 求证: $A \notin \mathfrak{C}(\mathscr{X},\mathscr{Y})$.

证明

用反证法, 若 $A \in \mathfrak{C}(\mathscr{X},\mathscr{Y})$, 则取 B_1 是 $\mathscr{X}$ 中的单位球, 知 $\overline{A(B_1)}$ 作为有界集是紧集, 因而其至少是列紧的. 将 $R(A)$ 视作 $\mathscr{Y}$ 的闭线性子空间, 由推论(2.4.9)知 $\overline{A(B_1)}$ 列紧与 $\dim R(A) = \infty$ 相悖! 命题即证.

练习 4.6 设 $\omega_n \in \mathbb{K}, \omega_n \to 0(n \to \infty)$, 求证: 映射

$$T : \{\xi_n\} \mapsto \{\omega_n \xi_n\} \quad (\forall \{\xi_n\} \in l^p)$$

是 $l^p(p \geq 1)$ 上的紧算子.

证明

取 B_1 是 l^p 中的单位球, 只需证明 $\overline{T(B_1)} \subset l^p$ 是紧的即可, 进一步只需证明其列紧.

4.2 Riesz-Fredholm 理论

本节着重研究与紧算子有关的算子方程的可解性问题. 设 $\mathscr{X}$ 是 B 空间, $A \in \mathfrak{C}(\mathscr{X})$, 又设 $T = I - A$, 其中 I 是恒同算子. 问方程

$$Tx = y \tag{4.5}$$

对哪些 $y \in \mathscr{X}$ 有解? 解的结构如何?

† 线性代数方程的一些回顾

当 $\mathscr{X} = \mathbb{R}^n$, 知此时 T 作为连续线性算子是矩阵, 这一套理论就是线性代数所对应的解空间理论. 设 $T = (t_{ij})_{n\times n}, x = \{x_j\}_{j=1}^n, y = \{y_i\}_{i=1}^n$. 根据线性代数的知识, 要使得方程(4.5)有界, 亦即

$$\sum_{j=1}^n t_{ij}x_j = y_i, \quad i = 1, 2, \cdots, n$$

必须且仅须

$$y = \sum_{j=1}^n x_j T_j$$

其中 $T_j = \{t_{ij}\}_{i=1}^n \in \mathbb{R}^n (j = 1, 2, \cdots, n)$, 也就是说 y 可以通过 $T_j(j = 1, \cdots, n)$ 线性表出, 这又等价于若 $z \in \mathbb{R}^n$, 则

$$z \perp y \Leftrightarrow z \perp T_j, \quad j = 1, 2, \cdots, n$$

即

$$\langle z, y\rangle = 0 \Leftrightarrow \sum_{i=1}^{n} t_{ij} z_i = 0, \quad j = 1, 2, \cdots, n \tag{4.6}$$

这导出下述两个结论:

(1) 要使得 $y \in \mathbb{R}^n$ 是方程(4.5)的解, 必须且仅须对任意满足 $T^* z = \theta$ 的 $z \in \mathbb{R}^n$ 有

$$\langle z, y\rangle = 0$$

(2) 关于方程(4.5)只可能有下述两种情形:

  (a) 要么 $\forall y \in \mathbb{R}^n$, 方程(4.5)总有唯一解;

  (b) 要么 $Tx = \theta$ 有非零解, 此时 $Tx = \theta$ 的非零解的极大线性无关组个数与 $T^* x = \theta$ 的非零解的极大线性无关组个数相等.

† Fredholm 的工作

Fredholm 研究过下述积分方程: 设 $K \in C([0,1] \times [0,1])$, 考察:

$$x(t) = \int_0^1 K(t,s)x(s)ds + y(t), \quad K \in C([0,1] \times [0,1]) \tag{4.7}$$

及其共轭方程

$$f(t) = \int_0^1 K(s,t)f(s)ds + g(t) \tag{4.8}$$

其中 $x, y, f, g \in L^2[0,1]$. 他得到下述结论:

(1) 关于方程(4.7)只可能有下述两种情形:

  (a) 要么 $\forall y \in L^2[0,1]$, 方程(4.7)存在唯一解 $x \in L^2[0,1]$;

  (b) 要么当 $y = \theta$ 时, 方程(4.7)有非零解.

(2) 方程(4.8)与方程(4.7)的情形一样, 也就是说它们同时满足结论 (1) 的 (a) 或 (b), 且它们在同时满足结论 (1)(b) 时, 它们所对应的齐次方程的线性无关解的个数是相同的有穷数.

(3) 在方程(4.7),(4.8)同时满足结论 (1)(b) 时, 要使得方程(4.7)有解, 必须且仅须

$$\int_0^1 f(t)y(t)dt = 0$$

其中 f 是方程(4.8)的齐次方程的解. 要使得方程(4.8)有解, 必须且仅须

$$\int_0^1 g(t)x(t)dt = 0$$

其中 x 是方程(4.7)的齐次方程的解.

比较上面代数方程和积分方程的结论, 可以发现它们是极为相似的. 事实上, 这种相似性对一般的算子方程也成立, 下面先引进记号.

† 一些记号

对任意 $T \in \mathscr{L}(\mathscr{X})$, 记

$$R(T) := T(\mathscr{X})$$

$$N(T) := \{x \in \mathscr{X} : Tx = \theta\}$$

对任意的 $M \subset \mathscr{X}, N \subset \mathscr{X}^*$, 记

$${}^{\perp}M := \{f \in \mathscr{X} : \langle f, x\rangle = 0, \forall x \in M\}$$

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

若 $f \in \mathscr{X}^*, x \in \mathscr{X}$ 满足 $\langle f, x\rangle = 0$, 就简记为

$$f \perp x$$

根据这些几号, 当 $T = I - A$, 其中

$$A : x(t) \mapsto \int_0^1 K(t,s)x(s)ds$$

则三个 Fredholm 结论可以简洁表为:

(1) $N(T) = \{\theta\} \Rightarrow R(T) = \mathscr{X}$.

(2) $\sigma(A) = \sigma(A^*)$, 且 $\dim N(T) = \dim N(T^*) < \infty$.

(3) $R(T) = N(T^*)^{\perp}, R(T^*) = {}^{\perp}N(T)$.

下面对一般的 $T = I - A (A \in \mathfrak{C}(\mathscr{X}))$ 证明上面三个 Fredholm 结论.

定理 4.2.1 (Riesz-Fredholm)

设 $\mathscr{X}$ 是 B 空间, $A \in \mathfrak{C}(\mathscr{X}), T = I - A$, 则

(1) $\sigma(T) = \sigma(T^*)$;

(2) $\dim N(T) = \dim N(T^*) < \infty$;

(3)

$$R(T) = N(T^*)^{\perp} = \{x \in \mathscr{X} : f(x) = 0, \forall f \in N(T^*)\},$$

$$R(T^*) = {}^{\perp}N(T) = \{f \in \mathscr{X}^* : f(x) = 0, \forall x \in N(T)\}.$$

♡

† 结论 (1) 的证明

首先来证明结论 (1) 对任意有界算子都成立, 也就是下面的定理:

定理 4.2.2

若 $T \in \mathscr{L}(\mathscr{X})$, 则 $\sigma(T) = \sigma(T^*)$.

♡

证明

回忆

$$\sigma(T) = \mathbb{C} \backslash \rho(T), \quad \rho(T) = \{\lambda \in \mathbb{C} : (\lambda I - T)^{-1} \in \mathscr{L}(\mathscr{X})\}$$

$$\sigma(T^*) = \mathbb{C} \backslash \rho(T^*), \quad \rho(T^*) = \{\lambda \in \mathbb{C} : (\lambda I^* - T^*)^{-1} \in \mathscr{L}(\mathscr{X}^*)\}$$

于是 $\sigma(T) = \sigma(T^*) \Leftrightarrow \rho(T) = \rho(T^*)$. 注意如果记 $S_\lambda = \lambda I - T \in \mathscr{L}(\mathscr{X})$, 则 $\rho(T) = \rho(T^*)$ 实际上说明的是 $S_\lambda^{-1} \in \mathscr{L}(\mathscr{X}) \Leftrightarrow (S_\lambda^*)^{-1} \in \mathscr{L}(\mathscr{X})$, 于是可以考虑先证明下述引理:

引理 4.2.1

若 $S \in \mathscr{L}(\mathscr{X})$, 则 $S^{-1} \in \mathscr{L}(\mathscr{X}) \Leftrightarrow (S^*)^{-1} \in \mathscr{L}(\mathscr{X}^*)$.

♡

首先说明若 $S \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, $\mathscr{X}, \mathscr{Y}$ 均为 B^* 空间, $S^{-1} \in \mathscr{L}(\mathscr{Y}, \mathscr{X})$ 存在, 则 $(S^*)^{-1} = (S^{-1})^*$. 梳理有 $S^* \in \mathscr{L}(\mathscr{Y}^*, \mathscr{X}^*)$, 从而 $(S^*)^{-1} \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*)$, 同理可得 $(S^{-1})^* \in \mathscr{L}(\mathscr{X}^*, \mathscr{Y}^*)$. 现在任取 $f \in \mathscr{X}^*$, 根据 S^* 的定义, 任取 $y \in \mathscr{Y} . x \in \mathscr{X}$ 满足 $y = Sx$, $g \in \mathscr{Y}^*$ 满足 $f = S^* g$, 有

$$\langle f, x \rangle = \langle f, S^{-1} y \rangle = \langle (S^{-1})^* f, y \rangle$$

另一方面

$$\langle f, x \rangle = \langle S^* g, x \rangle = \langle g, Sx \rangle = \langle g, y \rangle$$

这说明

$$\langle (S^{-1})^* f, y \rangle = \langle g, y \rangle = \langle (S^*)^{-1} f, y \rangle$$

从而 $(S^{-1})^* = (S^*)^{-1}$.

回到引理, 若 $S^{-1} \in \mathscr{L}(\mathscr{X})$, 知 $(S^{-1})^* \in \mathscr{L}(\mathscr{X}^*)$, 根据前述说明知 $(S^{-1})^* = (S^*)^{-1}$, 故 $(S^*)^{-1} \in \mathscr{L}(\mathscr{X}^*)$.

若 $(S^*)^{-1} \in \mathscr{L}(\mathscr{X}^*)$, 根据刚刚证明的结论知 $(S^{**})^{-1} \in \mathscr{L}(\mathscr{X}^{**})$, 又因为 $S = S^{**}|_{\mathscr{X}}$, 故 S 在 $\mathscr{X}$ 上是一一映射. 且根据 $S \in \mathscr{L}(\mathscr{X}^{**})$ 知 $R(S) \subset \mathscr{X} \subset \mathscr{X}^{**}$ 至少是闭线性子空间.

再证明 $R(S)=\mathscr{X}$. 若不然, 则存在 $\theta\neq x_0\in\mathscr{X}\setminus R(S)$. 根据 Hahn-Banach 定理知存在 $\theta\neq f\in\mathscr{X}^*$ 使得

$$f(x_0)=\|x_0\|,\quad \forall x\in R(S)(f(x)=0)$$

于是

$$\forall y\in\mathscr{X}(\langle f,Sy\rangle=0=\langle S^*f,y\rangle)$$

这说明 $S^*f=\theta$, 亦即 $f\in N(S^*)$. 又因为 $(S^*)^{-1}$ 存在, 故 $N(S^*)=\{\theta\}$, 从而 $f=\theta$, 矛盾! 故 $S^{-1}\in\mathscr{L}(\mathscr{X})$.

现在只要 $\lambda\in\rho(T)$, 则因为 $T\in\mathscr{L}(\mathscr{X})\Rightarrow\forall\lambda\in\mathbb{C}((\lambda I-T)\in\mathscr{L}(\mathscr{X}))$, 且根据 $\rho(T)$ 的定义有 $(\lambda I-T)^{-1}\in\mathscr{L}(\mathscr{X})$, 从而根据引理有 $(\lambda I^*-T^*)^{-1}\in\mathscr{L}(\mathscr{X}^*)$, 于是 $\lambda\in\rho(T^*)$, 亦即 $\rho(T)\subset\rho(T^*)$.

而若 $\lambda\in\rho(T^*)$, 知 $T^*\in\mathscr{L}(\mathscr{X}^*)\Rightarrow\forall\lambda\in\mathbb{C}((\lambda I^*-T^*)\in\mathscr{L}(\mathscr{X}^*))$, 且根据 $\rho(T^*)$ 的定义有 $(\lambda I^*-T^*)^{-1}\in\mathscr{L}(\mathscr{X}^*)$, 从而根据引理有 $(\lambda I-T)^{-1}\in\mathscr{L}(\mathscr{X})$, 故 $\lambda\in\rho(T)$, 亦即 $\rho(T^*)\subset\rho(T)$.

综上, $\rho(T)=\rho(T^*)$, 故 $\sigma(T)=\sigma(T^*)$, 定理4.2.2至此得证, 结论 (1) 顺势得证. □

† 中场的准备工作

先引进一个定义:

定义 4.2.1 (闭值域算子)

称 $T\in\mathscr{L}(\mathscr{X})$ 是闭值域算子, 如果

$$R(T)=\overline{R(T)}.$$

♣

定理 4.2.3

若 $A\in\mathfrak{C}(\mathscr{X})$, 则 $T=I-A$ 是闭值域算子.

♡

证明

因为 $N(T)$ 是 $\mathscr{X}$ 的闭子空间, 故商空间 $\mathscr{X}/N(T)$ 有意义, 考察

$$\widetilde{T}:\mathscr{X}/N(T)\to\mathscr{X},[x]\mapsto\widetilde{T}[x]=Tx.$$

显见 $R(\widetilde{T})=R(T)$, 且 $\widetilde{T}$ 是有界线性的, 其满足 $N(\widetilde{T})=\{[\theta]\}$, 这说明 $\widetilde{T}$ 存在逆算子. 回忆闭图像定理一节曾经介绍过: 如果 T 是 B^* 空间到 B 空间的连续线性算子, 那么 $D(T)$ 总可以视作闭子空间. 现在要说明 $R(T)$ 是 $\mathscr{X}$ 的闭子空间, 注意 $R(T)=R(\widetilde{T})$ 恰是 $\widetilde{T}^{-1}$ 的定义域, 从而只需说明 $\widetilde{T}^{-1}$ 是连续线性算子. 现在其线性性是显然的, 对于连续性用反证法, 如若 $\widetilde{T}^{-1}$ 不连续, 根据线性性知其在 θ 处也不连续, 根据连续的定义知此时

$$\exists\varepsilon_0>0\forall\delta>0\exists y_\delta\in R(\widetilde{T})(\|y_\delta\|<\delta\wedge\|\widetilde{T}^{-1}y_\delta\|\geq\varepsilon_0)$$

设 $y_\delta=\widetilde{T}[w_\delta]$, 则此即

$$\|[w_\delta]\|\geq\varepsilon_0\wedge\|\widetilde{T}[w_\delta]\|<\delta$$

取 $\delta=\frac{1}{n}(n=1,2,\cdots)$, 对应 $[w_\delta]$ 记为 $[w_n]$, 则知此时存在序列 $\{[w_n]\}_{n\in\mathbb{N}}$ 使得

$$\|[w_n]\|\geq\varepsilon_0\wedge\|\widetilde{T}[w_n]\|\to 0(n\to\infty)$$

设 $[x_n]=\frac{[w_n]}{\|w_n\|}(n\in\mathbb{N})$, 则

$$\|[x_n]\|=1(n\in\mathbb{N})\wedge\widetilde{T}[x_n]\to\theta(n\to\infty)$$

根据商空间范数的定义知对任意 $n\in\mathbb{N}$, 均存在 $x_n\in[x_n]$ 使得

$$\|x_n\|<2(n=1,2,\cdots)\wedge Tx_n=(I-A)x_n\to\theta(n\to\infty)$$

这说明 $\{Ax_n\}_{n\in\mathbb{N}}$ 最终是有界序列. 因为 A 是紧算子, 故存在子列 $\{x_{n_k}\}_{k\in\mathbb{N}}$ 使得 $\{Ax_{n_k}\}_{k\in\mathbb{N}}$ 收敛, 不妨设 $Ax_{n_k}\to z(k\to\infty)$, 从而

$$x_{n_k}=Ax_{n_k}+(I-A)x_{n_k}\to z,\quad k\to\infty$$

根据 T 的连续性知 $Tz = \theta = \widetilde{T}[z]$, 于是 $[z] = [\theta]$, 故

$$\|[x_{n_k}]\| = \|[x_{n-k} - z]\| \le \|x_{n_k} - z\| \to 0, \quad k \to \infty$$

但这与 $\|[x_{n_k}]\| = 1(\forall k \in \mathbb{N})$ 矛盾! 定理4.2.3至此得证. □

定理 4.2.4

若 $A \in \mathfrak{C}(\mathscr{X}), T = I - A$, 且 $N(T) = \{\theta\}$, 则 $R(T) = \mathscr{X}$.

证明

用反证法. 如若不然, 设 $\mathscr{X}_0 = \mathscr{X}, \mathscr{X}_k = T(\mathscr{X}_{k-1})(k = 1, 2, \cdots)$, 则因为反证的结论是 $R(T) \subsetneq \mathscr{X}$, 知 $\mathscr{X}_1 \neq \mathscr{X}_0$. 又因为 $N(T) = \{\theta\}$ 说明 T 是一一映射, 故

$$\mathscr{X}_0 \supsetneq \mathscr{X}_1 \supsetneq \mathscr{X}_2 \supsetneq \cdots$$

根据真闭子空间的 Riesz 引理知, 对任意 $k \in \mathbb{N}$, 均存在 $y_k \in \mathscr{X}_k$, 使得 $\|y_k\| = 1$, 但

$$\rho(y_k, \mathscr{X}_{k+1}) \ge \frac{1}{2}$$

于是对任意的 $p, n \in \mathbb{N}$, 因为

$$Ty_n - Ty_{n+p} + y_{n+p} \in \mathscr{X}_{n+1}$$

故

$$\|Ay_n - Ay_{n+p}\| = \|y_n - Ty_n + Ty_{n+p} - y_{n+p}\| \ge \frac{1}{2}$$

这说明序列 $\{Ay_n\}_{n\in\mathbb{N}}$ 没有收敛子列, 这与 $A \in \mathfrak{C}(\mathscr{X})$ 矛盾! 定理4.2.4至此得证. □

† 一个特殊情况的简单说明

现在说明至少 $\dim N(T) = 0$ 时, Riesz-Fredholm 定理4.2.1已经成立了. 回忆要证明的剩下两件事:

(1) $\dim N(T) = \dim N(T^*) < \infty$;

(2)

$$R(T) = N(T^*)^{\perp} = \{x \in \mathscr{X} : f(x) = 0, \forall f \in N(T^*)\}$$

$$R(T^*) = {}^{\perp}N(T) = \{f \in \mathscr{X}^* : f(x) = 0, \forall x \in N(T)\}$$

现在 $\dim N(T) = 0 \Rightarrow N(T) = \{\theta\}$, 这说明 T 是一一映射, 又根据定理4.2.4知 $R(T) = \mathscr{X}$, 故 T 是满射. 根据 Banach 逆算子定理有 $T \in \mathscr{L}^{-1}(\mathscr{X})$, 这说明 $0 \notin \sigma(T)$. 又因为已经说明了 $\sigma(T) = \sigma(T^*)$, 故 $0 \notin \sigma(T^*)$, 亦即 $(T^*)^{-1} \in \mathscr{L}(\mathscr{X}^*)$, T^* 也是一一的满射, 故 $\dim N(T^*) = 0 = \dim N(T)$ 且 $R(T^*) = \mathscr{X}^*$. 注意

$${}^{\perp}N(T) = \{f \in \mathscr{X}^* : f(x) = 0, \forall x \in N(T)\}$$

而 $x \in N(T) \Rightarrow x = \theta$, 故 $f(x) = 0$ 对任意 $f \in \mathscr{X}^*$ 均成立, 也就是 ${}^{\perp}N(T) = \mathscr{X}^*$, 于是

$$R(T^*) = \mathscr{X}^* = {}^{\perp}N(T)$$

同时

$$N(T^*)^{\perp} = \{x \in \mathscr{X} : f(x) = 0, \forall f \in N(T^*)\}$$

现在 $\dim N(T^*) = 0 \Rightarrow N(T^*) = \{\theta\}$, 于是只要 $x \in \mathscr{X}$, 就有 $f(x) = 0(\forall f \in N(T^*))$, 这说明 $N(T^*)^{\perp} = \mathscr{X}$, 因而

$$R(T) = \mathscr{X} = N(T^*)^{\perp}$$

至此 $\dim N(T) = 0$ 时的 Riesz-Fredholm 定理证毕. □

† 对结论 (2),(3) 的证明

首先说明结论 (2) 中, 至少 $\dim N(T), \dim N(T^*)$ 均有限.

引理 4.2.2

若 $A \in \mathfrak{C}(\mathscr{X}), T = I - A$, 则

$$\dim N(T) < \infty, \quad \dim N(T^*) < \infty$$

证明

设 $\mathscr{X}_0 = N(T), B_1 = \{x \in \mathscr{X}_0 : \|x\| \leq 1\}$, 此时因为任取 $x \in B_1 \subset \mathscr{X}_0$ 有

$$Tx = (I - A)x = 0 \Rightarrow Ax = x \Rightarrow \|Ax\| = \|x\| \leq 1 \Rightarrow Ax \in B_1$$

于是一方面由 $x = Ax$ 知每个在 B_1 中的元素都可表成 $Ax(x \in B_1)$ 的形式, 进而 $B_1 \subset A(B_1) \subset \overline{A(B_1)}$, 另一方面由 $\|Ax\| \leq 1$ 知 $A(B_1) \subset B_1$, 进一步由 A 的连续性知 $\overline{A(B_1)} \subset B_1$, 于是 $B_1 = \overline{A(B_1)}$. 现在因为 A 是紧算子, 故 $\overline{A(B_1)}$ 是紧集, 因而 $B_1 = \overline{A(B_1)}$ 是紧集, 根据有穷维 B^* 空间的刻画, 现在 $\mathscr{X}_0$ 的单位球面 B_1 是紧集, 故 $\mathscr{X}_0$ 是有穷维 B 空间, 亦即

$$\dim \mathscr{X}_0 = \dim N(T) < \infty$$

因为 $T^* = I^* - A^*$, 且 $A \in \mathfrak{C}(\mathscr{X}) \Rightarrow A^* \in \mathfrak{C}(\mathscr{X}^*)$, 利用同样的方法可以证明 $\dim(N^*) < \infty$. □

再来说明 $\dim N(T) = \dim N(T^*)$. 设 $x_1, \cdots, x_n \in N(T)$ 是 $N(T)$ 的一组基, $f_1, \cdots, f_m \in N(T^*)$ 是 $N(T^*)$ 的一组基, 此即说明 $n = m$. 这需要下面两个引理.

引理 4.2.3

存在闭线性子空间 $\mathscr{X}_1 \subset \mathscr{X}$, 使得

$$\mathscr{X} = \operatorname{span}\{x_1, x_2, \cdots, x_n\} \oplus \mathscr{X}_1.$$

证明

由 Hahn-Banach 定理, 因为 $x_1, \cdots, x_n$ 线性无关, 故存在 $g_1, \cdots, g_n \in \mathscr{X}^*$ 使得

$$g_i(x_j) = \delta_{ij}, \quad 1 \leq i, j \leq n$$

令 $\mathscr{X}_1 = \bigcap_{i=1}^{n} N(g_i)$, 其中 $N(g) = \{x \in \mathscr{X} : g(x) = 0\}$, 则 $\mathscr{X}_1$ 是闭线性子空间, 且

(a) $\operatorname{span}\{x_1, \cdots, x_n\} \cap \mathscr{X}_1 = \{\theta\}$;

(b) 对任意 $x \in \mathscr{X}$ 有 $x - \sum_{i=1}^{n} g_i(x)x_i \in \mathscr{X}_1$. 这是因为设 $x = \sum_{i=1}^{n} c_i x_i$, 有

$$g_j(x - \sum_{i=1}^{n} g_i(x)x_i) = c_j - \sum_{i=1}^{n} g_i(x)\delta_{ij} = c_j - g_j(x) = c_j - c_j = 0$$

综合 (a),(b) 即知 $\mathscr{X} = \operatorname{span}\{x_1, \cdots, x_n\} \oplus \mathscr{X}_1$. 引理4.2.3至此证毕. □

引理 4.2.4

存在 $y_1, y_2, \cdots, y_m \in \mathscr{X}$ 使得

$$f_i(y_j) = \delta_{ij}, \quad 1 \leq i, j \leq m.$$

证明

设线性连续映射

$$V : \mathscr{X} \to \mathbb{K}^m, \quad x \mapsto (\langle f_1, x\rangle, \langle f_2, x\rangle, \cdots, \langle f_m, x\rangle)$$

则引理结论等价于 V 是满射, 下面用反证法证明 V 是满射. 如若不然, 则 $V(\mathscr{X})$ 是 $\mathbb{K}^m$ 的一个真子空间, 从而根据 Hahn-Banach 定理知存在 $g \in (\mathbb{K}^m)^*$ 使得 $g(V(\mathscr{X})) = \{0\}$. 又因为 $\mathbb{K}^m$ 是 Hilbert 空间, 根据 Riesz 表示定理知存在 $0 \neq \alpha = (\alpha_1, \cdots, \alpha_m) \in \mathbb{K}^m$ 使得

$$g(x) = (\alpha, x), \quad \forall x \in \mathbb{K}^m$$

这便有

$$(\alpha, Vx) = 0, \quad \forall x \in \mathscr{X}$$

亦即

$$\langle \sum_{j=1}^{m} \alpha_j f_j, x \rangle = 0, \quad \forall x \in \mathscr{X}$$

于是 $\sum\limits_{j=1}^{m} \alpha_j f_j = \theta$, 但这与 $\{f_j\}_{j=1}^{m}$ 的线性无关性矛盾! 引理4.2.4至此证毕. $\square$

†† 结论 (2) 的收尾

现在来说明 $\dim N(T) = \dim N(T^*)$, 这只需证明 $n = m$. 不妨设 $n < m$, 考虑

$$\widetilde{T} : \mathscr{X} = \operatorname{span}\{x_1, \cdots, x_n\} \oplus \mathscr{X}_1 \to \operatorname{span}\{y_1, \cdots, y_n\} \oplus R(T) \hookrightarrow \mathscr{X}$$

$$\sum_{i=1}^{n} c_i x_i + y \mapsto \widehat{T}(\sum_{i=1}^{n} c_i x_i + y) = \sum_{i=1}^{n} c_i y_i + Ty$$

因为前面说明过 T 是一一的满射, 故 $Ty(y \in \mathscr{X}_1)$ 在 $y \neq \theta$ 时跑遍 $\mathscr{X}$ 的全体非零元, 这说明 $\widehat{T}$ 是满射. 但同时因为 $n < m$, 故

第五章 广义函数与 Sobolev 空间

在讨论广义函数前, 首先说明其引入动机. 这里选用 **[LS]** 的引言介绍.

五十多年以前, 工程师 Heaviside 在一篇大胆的学术论文中引入了他的符号计算法则, 利用一些没有得到证实的数学计算来求解物理问题. 此后, 这种符号计算或者说是运算演算, 得到了不断发展, 并成为电学家们理论研究的基础. 工程师们系统地使用此方法, 虽然每个人对它都有各自的理解, 但在使用时均或多或少地感到心安理得; 于是上述符号计算方法就成了一种" 虽不严格但却很成功的" 技巧. 自 Dirac 引入在 $x = 0$ 以外处处为零而在 $x = 0$ 处为无穷且使得

$$\int_{-\infty}^{+\infty} \delta(x)dx = +1$$

的著名函数 $\delta(x)$ 后, 符号计算公式就更不为强调严格性的数学家们所接受. 不但声称在 $x < 0$ 时等于 0, 而在 $x \geq 0$ 时等于 1 的 Heaviside 函数 $Y(x)$ 的导函数就是其定义本身在数学上就矛盾的 Dirac 函数 $\delta(x)$, 还要谈论这个缺乏实际存在意义的函数的导数 $\delta'(x), \delta''(x), \cdots$, 所有这些都超出了我们的容忍极限. 但怎么解释这些方法所取得的成功呢? 每当这种矛盾现象出现时, 很少不因此产生一种新的, 经过修改后就可以用来解释物理学家们的语言的数学理论; 这甚至是数学和物理取得进步的重要源泉.

首先需要引进一些记号. 记多重指标 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 其中 $\alpha_1, \alpha_2, \cdots, \alpha_n \geq 0$ 是整数, 记:

$$|\alpha| := \sum_{i=1}^{n} \alpha_i, \quad \alpha! := \alpha_1!\alpha_2! \cdots \alpha_n!,$$

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \quad \forall x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n,$$

$$\partial^\alpha = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_n}^{\alpha_n}$$

且若 $\beta \leq \alpha$, 也即 $\beta_j \leq \alpha_j, j = 1, 2, \cdots, n$, 则

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!} = \binom{\alpha_1}{\beta_1} \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n}$$

5.1 广义函数的概念

5.1.1 知识梳理

广义函数的定义需要借助一类性质优良的函数空间, 它是定义在这个空间上的连续线性泛函. 所以首先需要引进该函数空间.

5.1.1.1 基本空间 $\mathscr{D}(\Omega)$

定义 5.1.1 (支集)

设 $\Omega \subset \mathbb{R}^n$ 是一个开集, $u \in C(\overline{\Omega})$, 称集合

$$F = \{x \in \Omega : u(x) \neq 0\}$$

的闭包 (关于 Ω) 为 u 的关于 Ω 的支集, 记作 $\mathrm{supp}(u)$. 也即, 连续函数 u 的支集是在这个集合外 u 恒为 0 的相对于 Ω 的最小闭集.

♣

对于整数 $k \geq 0$(可以是 ∞), $C_0^k(\Omega)$ 表示在 Ω 内有紧支集的全体 $C^k(\overline{\Omega})$ 函数所组成的集合, 显见

$$C_0^\infty(\Omega) \subset \cdots \subset C_0^{k+1}(\Omega) \subset C_0^k(\Omega) \subset \cdots \subset C_0^0(\Omega).$$

例 5.1(磨光算子) 设

$$j(x) := \begin{cases} C_n e^{-\frac{1}{1-|x|^2}}, & |x| < 1 \\ 0, & |x| \ge 1 \end{cases}$$

其中

$$C_n := (\int_{|x|\le 1} e^{-\frac{1}{1-|x|^2}} dx)^{-1}$$

是一个只依赖于维数的常数. 则 $j(x) \in C_0^\infty(\mathbb{R}^n)$, 且

$$\int_{\mathbb{R}^n} j(x)dx = 1.$$

从 $j(x)$ 出发, 可以得到很多 $C_0^\infty(\mathbb{R}^n)$ 的函数. 譬如对任意的 $\delta > 0$, 令

$$j_\delta(x) := \frac{1}{\delta^n} j(\frac{x}{\delta})$$

则有下述命题.

命题 5.1.1

设 $u(x)$ 是一个可积函数, 并在 Ω 的一个紧子集 K 外恒为 0, 则当 $\delta > 0$ 足够小时, 函数

$$u_\delta(x) := \int_\Omega u(y) j_\delta(x-y)dy$$

是 $C_0^\infty(\Omega)$ 的函数.

证明

当 δ 足够小时, 根据 u 的条件知 $u_\delta(x) = \int_K u(y)j_\delta(x-y)dy$. 根据 $j_\delta(x)$ 的定义知 $j_\delta(x) \neq 0 \Rightarrow \rho(x,K) \le \delta$. 从而可设 $K_\delta := \{x \in \mathbb{R}^n : \rho(x,K) \le \delta\}$, 另设 δ 足够小使得 $K_\delta \subset \Omega$. 知 $\forall x \notin K_\delta (u_\delta(x) = 0)$.

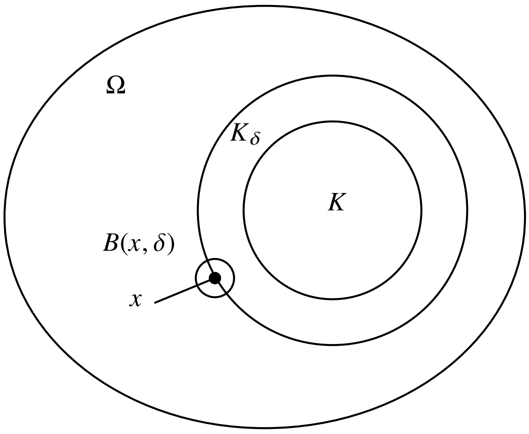


图 5.1: K_δ 构造示意

现在求解 $\partial^\alpha u_\delta(x)$. 对指标 $\alpha_0 = (1,0,\cdots,0)$, 根据偏导数的定义知

$$\begin{aligned}\partial^{\alpha_0} u_\delta(x) &= \lim_{h\to 0}\int_\Omega \frac{1}{h}[j_\delta(x+he_1-y) - j_\delta(x-y)]u(y)dy \\ &= \lim_{h\to 0}\int_\Omega \partial^{\alpha_0} j_\delta(x+\theta h e_1 - y)u(y)dy, \quad \theta = \theta(x,y) \in (0,1), e_1 = (1,0,\cdots,0) \in \mathbb{R}^n\end{aligned}$$

鉴于 j_δ 紧支连续可微, 知存在 $M_{\alpha_0} > 0$ 使得

$$|\partial^{\alpha_0} j_\delta(z)| \le M_{\alpha_0}, \quad \forall z \in \mathbb{R}^n$$

进而由 Lebesgue 控制收敛定理知极限与积分号可交换, 有

$$\partial^{\alpha_0} u_\delta(x) = \int_\Omega \lim_{h\to 0} \partial^{\alpha_0} j_\delta(x+\theta h e_1 - y)u(y)dy = \int_\Omega \partial^{\alpha_0} j_\delta(x-y)u(y)dy.$$

归纳可得对任意的指标 α 有

$$\partial^\alpha u_\delta(x) = \int_\Omega u(y)\partial^\alpha j_\delta(x-y)dy, \quad \forall x \in K_\delta$$

由 $j_\delta \in C_0^\infty(\Omega)$ 知 $u_\delta \in C_0^\infty(\Omega)$. □

定理 5.1.1

若 $u \in C_0^k(\Omega)$, 则

$$\|u_\delta(x) - u(x)\|_{C^k(\overline{\Omega})} \to 0, \quad \delta \to 0$$

证明

回忆

$$\|u\|_{C^k(\overline{\Omega})} = \sup_{|\alpha|\le k}\sup_{x\in\overline{\Omega}}|\partial^\alpha u(x)|$$

把 $u(x)$ 延拓到 $\mathbb{R}^n$ 上. 既然 u 是紧支的, 可设 $\forall x \notin \Omega(u(x)=0)$. 任取 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 当 $|\alpha| \le k$ 时有

$$\begin{aligned}\partial^\alpha u_\delta(x) &= \int_{\mathbb{R}^n} u(y)\partial_x^\alpha j_\delta(x-y)dy \\ &= (-1)^{|\alpha|}\int_{\mathbb{R}^n} u(y)\partial_y^\alpha j_\delta(x-y)dy \\ &= \int_{\mathbb{R}^n} \partial_y^\alpha u(y) j_\delta(x-y)dy \\ &= \int_{\mathbb{R}^n} \partial_y^\alpha u(x-y) j_\delta(y)dy \\ &= \int_{\mathbb{R}^n} \partial_y^\alpha u(x-\delta y) j(y)dy\end{aligned}$$

得到

$$|\partial^\alpha u_\delta(x) - \partial^\alpha u(x)| \le \int_{\mathbb{R}^n} |\partial^\alpha u(x-\delta y) - \partial^\alpha u(x)| j(y)dy$$

根据 $j(y)$ 的定义知 $j(y) = 0(|y| \ge 1)$, 而 $\partial^\alpha u(z)$ 在

$$(\mathrm{supp}(u))_1 := \{x \in \mathbb{R}^n : \rho(x, \mathrm{supp}(u)) \le 1\}$$

上是连续的, 因而由 $(\mathrm{supp}(u))_1$ 的紧性推知 $\partial^\alpha u(z)$ 在其上一致连续, 从而根据一致连续的定义知

$$\forall \varepsilon > 0 \exists 0 < \delta_0 < \frac{1}{2} \forall 0 < \delta < \delta_0 (|\partial^\alpha u(x-\delta y) - \partial^\alpha u(x)| < \varepsilon, \forall x \in \mathbb{R}^n, |y| \le 1)$$

得到

$$\sup_{x\in\mathbb{R}^n} |\partial^\alpha u_\delta(x) - \partial^\alpha u(x)| < \varepsilon \cdot \int_{\mathbb{R}^n} j(y)dy = \varepsilon$$

此即

$$\|u_\delta(x) - u(x)\|_{C^k(\overline{\Omega})} \to 0, \quad \delta \to 0$$

□

推论 5.1.1

若 μ 是 Ω 上的一个完全可加测度, 则由

$$\int_\Omega \varphi d\mu = 0, \quad \forall \varphi \in C_0^\infty(\Omega)$$

可得

$$\int_\Omega \varphi d\mu = 0, \quad \forall \varphi \in C_0^0(\Omega).$$

证明

当 μ 满足

$$\int_\Omega \varphi d\mu = 0, \quad \forall \varphi \in C_0^\infty(\Omega)$$

对于 $\varphi \in C_0^0(\Omega)$, 考虑其磨光函数 $\varphi_\delta := \varphi * j_\delta$, 根据命题(5.1.1)知 $\varphi_\delta \in C_0^\infty(\Omega)$, 因而

$$\int_\Omega \varphi_\delta d\mu = 0$$

再由定理(5.1.1), 只需说明

$$\lim_{\delta\to 0}\int_\Omega \varphi_\delta d\mu = \int_\Omega \lim_{\delta\to 0}\varphi_\delta d\mu$$

即可, 而这是由 μ 的完全可加性保证的[1]. □

定义 5.1.2 ($C_0^\infty(\Omega)$ 上的收敛, 基本空间 $\mathscr{D}(\Omega)$)

称序列 $\{\varphi_j\} \subset C_0^\infty(\Omega)$ 收敛于 φ_0, 如果

1. 存在一个相对于 Ω 的紧集 $K \subset \Omega$, 使得

$$\mathrm{supp}(\varphi_j) \subset K, \quad j = 0, 1, 2, \cdots$$

2. 对于任意指标 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 都有

$$\max_{x\in K}|\partial^\alpha \varphi_j(x) - \partial^\alpha \varphi_0(x)| \to 0, \quad j \to \infty$$

带上述收敛性的线性空间 $C_0^\infty(\Omega)$ 称为基本空间 $\mathscr{D}(\Omega)$. ♣

注 $\mathscr{D}(\Omega)$ 上只有收敛性, 而没有给定拓扑. 事实上 $\mathscr{D}(\Omega)$ 上的收敛不能由任何范数, 甚至是任何准范数导出. 这说明 $\mathscr{D}(\Omega)$ 并非 B^* 空间.

命题 5.1.2 ($\mathscr{D}(\Omega)$ 的序列完备性)

若 $\{\varphi_j\}_{j=1}^\infty$ 是一个基本列, 其满足

1. 存在公共的紧支集 K, 使得 $\mathrm{supp}(\varphi_j) \subset K$.
2. $\forall \varepsilon > 0 \exists N = N(\varepsilon, \alpha) \in \mathbb{N} \forall m, n > N(\max_{x\in K}|\partial^\alpha \varphi_n(x) - \partial^\alpha \varphi_m(x)| < \varepsilon)$

则必有 $\varphi_0 \in \mathscr{D}(\Omega)$ 使得 $\varphi_j \to \varphi_0 (j \to \infty)$. ♠

证明[2]

知 $\{\varphi_j\}_{j=1}^\infty \subset \mathscr{D}(\Omega) \subset C_0^0(\Omega)$, 令 $\alpha = (0, \cdots, 0)$ 可知 $\{\varphi_j\}$ 在 C_0^0 中是基本列, 因而可设其依照 C_0^0 的范数收敛到 $\varphi_0 \in C_0^0$.

5.1.1.2 广义函数的定义和基本性质

定义 5.1.3 (广义函数)

$\mathscr{D}(\Omega)$ 上的一切连续线性泛函都称为广义函数, 也即广义函数作为泛函 $f : \mathscr{D}(\Omega) \to \mathbb{R}$ 满足:

1. 线性性:

$$\langle f, \lambda_1\varphi_1 + \lambda_2\varphi_2\rangle = \lambda_1\langle f, \varphi_1\rangle + \lambda_2\langle f, \varphi_2\rangle, \quad \forall \varphi_1, \varphi_2 \in \mathscr{D}(\Omega), \forall \lambda_1, \lambda_2 \in \mathbb{R}$$

2. 连续性:

$$\forall\{\varphi_j\} \subset \mathscr{D}(\Omega)(\varphi_j \to \varphi_0(\mathscr{D}(\Omega)) \Rightarrow \langle f, \varphi_j\rangle \to \langle f, \varphi_0\rangle (j \to \infty))$$

[1] 自己并不特别清楚完全可加性怎么保证极限与积分号交换, 但是感觉可能会用到这里.

[2] 在确定 φ_0 之后的步骤就想不到了: 自己的思路是通过前面的磨光算子把 φ_0 磨光化, 但问题在于磨光化得到的对应于 φ_0 的光滑函数要怎么定义? 不确定 $\lim\limits_{\delta\to\infty}\varphi_0^\delta$ 是否存在.

其中 $\varphi_j \to \varphi_0(\mathscr{D}(\Omega))$ 表示收敛是在 $\mathscr{D}(\Omega)$ 的意义下的. 一切广义函数所组成的集合记作 $\mathscr{D}'(\Omega)$.

下面给出几个广义函数的例子.

例 5.2(δ 函数) 设 $\theta \in \Omega$, 定义

$$\langle \delta, \varphi \rangle = \varphi(\theta), \quad \forall \varphi \in \mathscr{D}(\Omega)$$

显见 δ 是线性的, 且当 $\varphi_j \to \varphi_0(\mathscr{D})$ 时, 有

$$|\varphi_j(\theta) - \varphi_0(\theta)| \to 0, \quad j \to \infty$$

故

$$\langle \delta, \varphi_j \rangle = \varphi_j(\theta) \to \varphi_0(\theta) = \langle \delta, \varphi_0 \rangle, \quad j \to \infty$$

这说明 δ 是 $\mathscr{D}(\Omega)$ 上的连续线性泛函, 因而是广义函数.

例 5.3($\delta^{(\alpha)}$) 对任意多重指标 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 定义

$$\langle \delta^{(\alpha)}, \varphi \rangle = (-1)^{|\alpha|}(\partial^\alpha \varphi)(\theta), \quad \varphi \in \mathscr{D}(\Omega)$$

容易证明 $\delta^{(\alpha)}$ 也是广义函数.

在介绍下一个例子前先引进局部可积的概念.

定义 5.1.4 (局部可积)

称 $f(x)$ 是 Ω 上的局部可积函数, 如果对任意相对于 Ω 的紧集 K, 都有

$$\int_K |f(x)| dx < \infty$$

此时记 $f \in L^1_{loc}(\Omega)$.

例 5.4(局部可积函数对应的泛函) 若 $f \in L^1_{loc}(\Omega)$, 则 $f(x)$ 对应一个广义函数

$$\langle f, \varphi \rangle = \int_\Omega f(x)\varphi(x) dx, \quad \forall \varphi \in \mathscr{D}(\Omega). \tag{5.1}$$

证明

显见 f 是线性的, 下面验证连续性. 若 $\varphi_j \to \varphi_0(\mathscr{D}(\Omega))$, 根据 $\mathscr{D}(\Omega)$ 意义收敛的定义知存在紧集 $K \subset \Omega$, 使得

$$\operatorname{supp}(\varphi_j) \subset K, \quad \max_{x \in K} |\varphi_j(x) - \varphi_0(x)| \to 0 (j \to \infty)$$

进而由 Lebesgue 控制收敛定理知

$$|\langle f, \varphi_j \rangle - \langle f, \varphi_0 \rangle| \le \int_K |f(x)| \cdot |\varphi_j(x) - \varphi_0(x)| dx \to 0, \quad j \to \infty$$

□

注

1. 如果把几乎处处相等的函数视作同一函数, 则对应

$$T : L^1_{loc}(\Omega) \to \mathscr{D}'(\Omega), f(x) \mapsto f$$

是单射. 要证明这件事, 只需证明若 $f \in L^1_{loc}(\Omega)$, 且

$$\int_\Omega f \cdot \varphi dx = 0, \quad \forall \varphi \in \mathscr{D}(\Omega)$$

则 $f(x) = 0$(在 Ω 上 a.e.). 进一步只需证明对任意的闭球 $B(x_0, \delta) \subset \Omega$ 都有 $f(x) = 0$(在 $B(x_0, \delta)$ 上 a.e.). 为此

考察函数

$$\tilde{f}(x) := \begin{cases} \operatorname{sgn} f(x), & x \in B(x_0, \delta) \\ 0, & x \notin B(x_0, \delta) \end{cases}$$

显见 $\tilde{f} \in L^1(\Omega)$, 且在 Ω 的一个紧集 $B(x_0, \delta)$ 外为 0. 可以证明 $C_0^\infty(\Omega)$ 在 $L^1(\Omega)$ 中稠密, 因而可以用 $C_0^\infty(\Omega)$ 中的函数:

$$\tilde{f}_\delta(x) := \int_\Omega \tilde{f}(y) j_\delta(x-y) dy$$

来逼近 $\tilde{f}$. 知

$$\|\tilde{f}_\delta - \tilde{f}\|_{L^1(\Omega)} \to 0, \quad \delta \to 0$$

且 $\tilde{f}_\delta \in \mathscr{D}(\Omega)$. 代入条件有

$$\int_\Omega f(x) \cdot \tilde{f}_{\delta_i}(x) dx = 0, \quad \forall i \in \mathbb{N}$$

其中 $\{\delta_i\}$ 满足 $\tilde{f}_{\delta_i} \to \tilde{f} (i \to \infty)$(在 Ω 上 a.e.) 且 $\delta_i > 0$. 由 Lebesgue 控制收敛定理知

$$\int_{B(x_0,\delta)} |f(x)| dx = \int_\Omega f(x) \cdot \tilde{f}(x) dx = \lim_{i\to\infty} \int_\Omega f(x) \tilde{f}_{\delta_i}(x) dx = 0$$

得到 $f(x) = 0$(在 $B(x_0, \delta)$ 上 a.e.).

2. 并非所有的函数都是广义函数. 事实上不可测函数就不能看成广义函数, 更准确的说, 广义函数只是局部可积函数的推广. 每个局部可积函数按照(5.1)式对应一个广义函数, 进而可以把所有局部可积函数都看成广义函数. 特别注意: 广义函数并不一定是局部可积函数.

例 5.5 若 μ 是 Ω 上的一个完全可加测度, 则

$$\langle f, \varphi \rangle = \int_\Omega \varphi(x) d\mu(x), \quad \forall \varphi \in \mathscr{D}(\Omega)$$

定义了一个广义函数. 特别由推论(5.1.1)可知这个对应是单射.

例 5.6 若 $\Omega = (0, 1)$, 则

$$\langle f, \varphi \rangle = \sum_{j=1}^{\infty} \varphi^{(j)}(\frac{1}{j}), \quad \forall \varphi \in \mathscr{D}(\Omega)$$

是一个广义函数.

定理 5.1.2

要使得 $f \in \mathscr{D}'(\Omega)$, 必须且仅须对任意相对于 Ω 的紧集 K, 存在常数 C 及非负整数 m, 使得

$$|\langle f, \varphi \rangle| \le C \sum_{|\alpha| \le m} \sup_{x\in K} |\partial^\alpha \varphi(x)| \quad (\forall \varphi \in \mathscr{D}(\Omega), \operatorname{supp}(\varphi) \subset K)$$

证明

当对任意相对于 Ω 的紧集 K, 存在常数 C 及非负整数 m, 使得

$$|\langle f, \varphi \rangle| \le C \sum_{|\alpha| \le m} \sup_{x\in K} |\partial^\alpha \varphi(x)| \quad (\forall \varphi \in \mathscr{D}(\Omega), \operatorname{supp}(\varphi) \subset K)$$

显见 $f \in \mathscr{D}'(\Omega)$.

当 $f \in \mathscr{D}'(\Omega)$, 用反证法, 若存在紧集 K 使得

$$\forall C > 0 \forall m \in \mathbb{N} \exists \varphi \in \mathscr{D}(\Omega)(|\langle f, \varphi \rangle| > C \sum_{|\alpha| \le m} \sup_{x\in K} |\partial^\alpha \varphi(x)| \wedge \operatorname{supp}(\varphi) \subset K)$$

注意上式对 φ 是齐次的, 故任取 $j \in \mathbb{N}$, 令 $C = m = j$, 知存在 $\varphi_j \in \mathscr{D}(\Omega)$ 使得 $\operatorname{supp}(\varphi_j) \subset K$, 满足

$$\sup_{x\in K}|\partial^\alpha \varphi_j| \le \frac{1}{j}, \quad |\alpha| \le j$$

且 $\langle f, \varphi_j\rangle = 1$. 故 $\varphi_j \to 0(\mathscr{D}(\Omega))$, 因而

$$\langle f, \varphi_j\rangle \to 0, \quad j \to \infty$$

这与 f 的连续性相悖! 命题即证. □

5.1.1.3 广义函数的收敛性

在 $\mathscr{D}'(\Omega)$ 上可以规定加法与数乘:

$$\langle(\lambda_1 f_1 + \lambda_2 f_2), \varphi\rangle = \lambda_1\langle f_1, \varphi\rangle + \lambda_2\langle f_2, \varphi\rangle, \quad \forall \varphi \in \mathscr{D}(\Omega), \forall \lambda_1, \lambda_2 \in \mathbb{R}$$

故 $\mathscr{D}'(\Omega)$ 构成一个线性空间, 进一步在 $D'(\Omega)$ 上引入 * 弱收敛.

定义 5.1.5 (广义函数的 * 弱收敛)

称 $\{f_j\} \subset \mathscr{D}'(\Omega)$* 弱收敛到 $f_0 \in \mathscr{D}'(\Omega)$, 是指:

$$\langle f_j, \varphi\rangle \to \langle f_0, \varphi\rangle (j \to \infty), \quad \forall \varphi \in \mathscr{D}(\Omega)$$

♣

下面几个例子说明广义函数的收敛是很弱的.

例 5.7 在 $\mathbb{R}$ 上,

$$f_j(x) = \frac{1}{\pi}\cdot\frac{\sin jx}{x}, \quad j = 1, 2, \cdots$$

是一列 $L^1_{loc}(\mathbb{R})$ 函数, 因而可以看成广义函数列. 有:

$$f_j \to \delta(\mathscr{D}'(\Omega)), \quad j \to \infty$$

证明

因为

$$\lim_{T\to\infty}\frac{1}{\pi}\int_{-T}^{T}\frac{\sin jx}{x}dx = 1 \tag{5.2}$$

现在任取 $\varphi \in \mathscr{D}(\mathbb{R})$, 根据基本函数的定义知 φ 是紧支的, 也即存在 $T_0 > 0$ 使得 $\operatorname{supp}(\varphi) \subset [-T_0, T_0]$. 当 $T > T_0$ 时有

$$\int_{-\infty}^{\infty} f_j(x)\cdot\varphi(x)dx = \int_{-T}^{T} f_j(x)\cdot\varphi(x)dx$$

另一方面, 根据(5.2)式中极限的定义有

$$\forall \varepsilon > 0 \exists T_1 > 0 \forall T > T_1 (|\frac{1}{\pi}\int_{-T}^{T}\frac{\sin jx}{x}dx - 1| < \varepsilon)$$

故当 $T > \max\{T_0, T_1\}$ 时有

$$|\langle f_j, \varphi\rangle - \varphi(0)| \le |\frac{1}{\pi}\int_{-T}^{T}\frac{\sin jx}{x}[\varphi(x) - \varphi(0)]dx| + \varepsilon|\varphi(0)| = \frac{1}{\pi}|\int_{0}^{T}\sin jx\cdot\frac{\varphi(x)+\varphi(-x)-2\varphi(0)}{x}dx| + \varepsilon|\varphi(0)|$$

固定 T, 由 Riemann-Lebesgue 引理知存在 $N > 0$, 当 $j > N$ 时有

$$\frac{1}{\pi}|\int_{0}^{T}\sin jx\cdot\frac{\varphi(x)+\varphi(-x)-2\varphi(0)}{x}dx| < \varepsilon$$

进而可得 $\langle f_j, \varphi\rangle \to \varphi(0) = \langle\delta, \varphi\rangle (j \to \infty)(\forall \varphi \in \mathscr{D}(\mathbb{R}))$. □

例 5.8 设 $j_\delta(x)$ 是例(5.1)中定义的函数, $\delta > 0$, 则当 $\delta \to 0$ 时, $j_\delta \to \delta(\mathscr{D}'(\mathbb{R}^n))$. 若用 δ_{x_0} 表示广义函数:

$$\langle\delta_{x_0}, \varphi\rangle = \varphi(x_0) \quad (\forall \varphi \in \mathscr{D}(\mathbb{R}^n))$$

则 $j_\delta(x - x_0) \to \delta_{x_0}(\delta \to 0)$.

证明

根据 $\mathscr{D}'(\mathbb{R}^n)$ 上收敛的定义知, 此即证明

$$\forall\varphi\in\mathscr{D}(\mathbb{R}^n)(\langle j_\delta,\varphi\rangle=\int_{\mathbb{R}^n}j_\delta(x)\varphi(x)dx\to\varphi(0)=\langle\delta,\varphi\rangle,\delta\to 0)$$

注意

$$\int_{\mathbb{R}^n}j_\delta(x)\varphi(x)dx=\int_{\mathbb{R}^n}\varphi(x)j_\delta(0-x)dx$$

根据定理(5.1.1), 记 $\langle j_\delta,\varphi\rangle:=\varphi_\delta$, 知:

$$\|\varphi_\delta(0)-\varphi(0)\|_{C^k(\mathbb{R}^n)}\to 0,\quad \delta\to 0$$

故

$$\int_{\mathbb{R}^n}j_\delta(x)\varphi(x)dx\to\varphi(0)=\langle\delta,\varphi\rangle,\quad \delta\to 0$$

此即欲证. 对于 $j_\delta(x-x_0)$ 的讨论是类似的, 只需将上述过程中的 0 换成 x_0. □

注 对于在 $x_0=\theta$ 点的 δ 函数, 在不混淆的情况下一般就记作 δ.

例 5.9 设 $f_j(x)$ 是 Ω 上的一列局部可积函数, 且对任意相对于 Ω 的紧集 K, 存在常数 M_k, 使得

$$|f_j(x)|\le M_k,\quad \forall x\in K,j=0,1,2,\cdots$$

且 $f_j(x)\to f_0(x)(k\to\infty)$(a.e. $x\in\Omega$), 则 f_j 作为广义函数列满足

$$\langle f_j,\varphi\rangle=\int_\Omega f_j(x)\varphi(x)dx,\quad j=0,1,2\cdots$$

在广义函数的意义下收敛到 f_0.

证明

根据广义函数收敛的定义知需证明

$$\forall\varphi\in\mathscr{D}(\Omega)(\langle f_j,\varphi\rangle=\int_\Omega f_j(x)\varphi(x)dx\to\int_\Omega f_0(x)\varphi(x)dx=\langle f_0,\varphi\rangle,j\to\infty)$$

又只需说明任取 $\overline{B(x_0,r)}\subset\Omega$, 均有

$$\int_{\overline{B(x_0,r)}}f_j(x)\varphi(x)dx\to\int_{\overline{B(x_0,r)}}f(x)\varphi(x)dx,\quad j\to\infty$$

显见 $\overline{B(x_0,r)}$ 是相对于 Ω 的紧集, 进而依题目条件与 Lebesgue 控制收敛定理知

$$\lim_{j\to\infty}\int_{\overline{B(x_0,r)}}f_j(x)\varphi(x)dx=\int_{\overline{B(x_0,r)}}\lim_{j\to\infty}f_j(x)\varphi(x)dx=\int_{\overline{B(x_0,r)}}f(x)\varphi(x)dx$$

命题即证. □

例 5.10(Gauss 核) 设 $f_1(x)=e^{-\pi|x|^2}$, 令

$$f_j(x)=j^{\frac{n}{2}}f_1(\sqrt{j}x)=j^{\frac{n}{2}}e^{-j\pi|x|^2},\quad j=2,3,\cdots$$

则有

$$\langle f_j,\varphi\rangle\to\langle\delta,\varphi\rangle\quad(j\to\infty)(\forall\varphi\in\mathscr{D}(\mathbb{R}^n))$$

其中 f_j 是函数 $f_j(x)$ 所对应的广义函数 $(j=1,2,3,\cdots)$.

证明

此即证明

$$\forall\varphi\in\mathscr{D}(\mathbb{R}^n)(\int_{\mathbb{R}^n}f_j(x)\varphi(x)dx=\int_{\mathbb{R}^n}j^{\frac{n}{2}}e^{-j\pi|x|^2}\varphi(x)dx\to\varphi(0),j\to\infty)$$

任取 $\varepsilon > 0$ 有

$$\int_{\mathbb{R}^n} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(x)dx = \int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(x)dx + \int_{\mathbb{R}^n\backslash B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(x)dx := I_\varepsilon + J_\varepsilon$$

对 J_ε, 令 $j \to \infty$, 既然 $\varphi \in \mathscr{D}(\Omega)$ 是连续紧支函数, 由 Lebesgue 控制收敛定理知 $J_\varepsilon \to 0$.

对 I_ε, 由 $\varphi \in \mathscr{D}(\Omega)$ 知

$$\forall\varepsilon' > 0\exists\delta > 0\forall x \in B(0,\delta)(|\varphi(x) - \varphi(0)| < \varepsilon')$$

取 ε 足够小使得 $\varepsilon < \delta$, 此时

$$\begin{aligned}|\int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(x)dx - \int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(0)dx| &\le \int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}|\varphi(x) - \varphi(0)|dx \\ &< \varepsilon' \int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}dx \le \varepsilon'\end{aligned}$$

故

$$\int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(x)dx \to \int_{B(0,\varepsilon)} j^{\frac{n}{2}} e^{-j\pi|x|^2}\varphi(0)dx \to \varphi(0), \quad j \to \infty$$

命题即证. □

5.1.2 习题

练习 5.1 设 $1 \le p < \infty$, 求证: $C_0^\infty(\Omega)$ 在 $L^p(\Omega)$ 中稠密.

证明

若 $u \in L^p(\Omega)(1 \le p < \infty)$, 取 $u_\delta = u * j_\delta \in C_0^\infty(\Omega)$, 其中 j_δ 是磨光算子. 显见 $j_\delta \in L^1(\Omega)$, 由 Young 不等式(3.6.2)知

$$\|u_\delta\|_{L^p(\Omega)} = \|u * j_\delta\|_{L^p(\Omega)} \le \|u\|_{L^p(\Omega)} \cdot \|j_\delta\|_{L^1(\Omega)} = \|u\|_{L^p(\Omega)}$$

再证明 $C_0^0(\Omega)$ 在 $L^p(\Omega)$ 中稠密. 根据积分存在的必要条件, 取 $u \in L^p(\Omega)$, 知当 $|x|$ 足够大时, $u(x)$ 总是足够小的, 进而不妨就设 Ω 有界. 显见 u 可测, 由 Luzin 定理知任取 $\varepsilon > 0$, 总存在闭集 $K \subset \Omega$ 使得 $\mu(\Omega\backslash K) < \varepsilon$(其中 μ 是 Lebesgue 测度), 且 $u|_K$(即 u 在 K 上的限制) 是连续函数. 现取

$$u_\varepsilon(x) = \begin{cases} u(x), & x \in K \\ 0, & x \notin K \end{cases}$$

知 $u_\varepsilon \in C_0^0(\Omega)$, 且根据 K 的构造显见 $\|u_\varepsilon - u\|_{L^p(\Omega)} \to 0(\varepsilon \to 0)$.

现在着手命题. 设 $u \in L^p(\Omega)$, 因为 $C_0^0(\Omega)$ 在 $L^p(\Omega)$ 中稠密, 故任取 $\varepsilon > 0$, 存在 $\varphi \in C_0^0(\Omega)$ 使得

$$\|u - \varphi\|_{L^p(\Omega)} < \varepsilon \tag{5.3}$$

针对 φ, 取 $\varphi_\delta = \varphi * j_\delta$, 其中 j_δ 是磨光算子. 显见 $\varphi_\delta \in C_0^\infty$, 且既然由定理(5.1.1)知 $\|\varphi_\delta - \varphi\|_{C^0(\Omega)} \to 0(\delta \to 0)$, 自然可以取足够小的 δ 使得

$$\|\varphi_\delta - \varphi\|_{L^p(\Omega)} < \varepsilon \tag{5.4}$$

最后由 Young 不等式知

$$\|u_\delta - \varphi_\delta\|_{L^p(\Omega)} = \|(u - \varphi) * j_\delta\|_{L^p(\Omega)} \le \|u - \varphi\|_{^p(\Omega)} < \varepsilon \tag{5.5}$$

联立(5.3)-(5.5)式即得

$$\|u - u_\delta\| < 3\varepsilon$$

由 ε 的任意性即得欲证.

练习 5.2 求证: δ 函数不是局部可积函数.

证明

如若 δ 函数确为局部可积函数, 则存在 $f \in L^1_{loc}$ 满足

$$\langle \delta, \varphi \rangle = \int_\Omega f(x)\varphi(x)dx, \quad \forall \varphi \in \mathscr{D}(\Omega)$$

取 $\varphi_k(x) = j(kx)(k \in \mathbb{N})$, 其中 $j(x)$ 是磨光算子. 根据 $j(x)$ 的定义知此时 $f(x)\varphi_k(x)$ 的支集仅为 $B(0, \frac{1}{k})$. 有

$$|\int_\Omega f(x)\varphi_k(x)dx| = |\int_{B(0,\frac{1}{k})} f(x)\varphi_k(x)dx| \le \max_{x \in B(0,\frac{1}{k})} \varphi_k(x) \cdot |\int_{B(0,\frac{1}{k})} f(x)dx| \to 0, \quad k \to \infty$$

另一方面知 $\langle \delta, \varphi \rangle = \varphi(0) \neq 0$, 矛盾! 命题即证.

练习 5.3 设

$$f_j(x) = (1 + \frac{x}{j})^j \quad (j = 1, 2, \cdots)(x \in \mathbb{R})$$

求证:

$$f_j(x) \to e^x \quad (\mathscr{D}'(\mathbb{R})).$$

证明

任取 $\varphi \in \mathscr{D}(\mathbb{R})$, 设 K 是 φ 的支集, 显见 $f_j(x)(j = 1, 2, \cdots)$ 在 K 内一致有界. 根据 Lebesgue 控制收敛定理知:

$$\lim_{j\to\infty} \int_\mathbb{R} f_j(x)\varphi(x)dx = \lim_{j\to\infty} \int_K f_j(x)\varphi(x)dx = \int_K \lim_{j\to\infty} f_j(x)\varphi(x)dx = \int_K e^x\varphi(x)dx = \int_\mathbb{R} e^x\varphi(x)dx$$

命题即证.

练习 5.4 在 $\mathscr{D}'(\mathbb{R})$ 中, 求证:

(1) $\frac{1}{\pi} \cdot \frac{1}{x^2+\varepsilon^2} \to \delta(x)(\varepsilon \to 0^+)$.

证明

任取 $\varphi \in \mathscr{D}(\mathbb{R})$, 知

$$\begin{aligned}\int_\mathbb{R} \frac{1}{\pi} \cdot \frac{\varphi(x)}{x^2+\varepsilon^2}dx &= \int_\mathbb{R} \frac{1}{\pi\varepsilon} \cdot \varphi(x) \cdot \frac{1}{(\frac{x}{\varepsilon})^2+1}d(\frac{x}{\varepsilon}) = \int_\mathbb{R} \frac{1}{\pi\varepsilon} \cdot \varphi(\varepsilon x) \cdot \frac{1}{x^2+1}dx \\ &= \int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(\varepsilon x) \cdot \frac{1}{x^2+1}dx + \int_{|x|\ge\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(\varepsilon x) \cdot \frac{1}{x^2+1}dx := I_\varepsilon + J_\varepsilon\end{aligned}$$

其中

$$|J_\varepsilon| = |\int_{|x|\ge\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(\varepsilon x)d\arctan x| = |\int_{|x|\ge\frac{\pi}{2}\varepsilon} \frac{1}{\pi} \arctan x\varphi'(\varepsilon x)dx| = \frac{1}{2}\arctan\frac{\pi}{2}\varepsilon \cdot |\varphi(\varepsilon) + \varphi(-\varepsilon)| \tag{5.6}$$

根据 φ 的连续性知任取 $\varepsilon' > 0$, 存在 $1 > \delta > 0$, 使得 $|x| < \delta$ 时有

$$|\varphi(x) - \varphi(0)| < \varepsilon'$$

取 ε 满足 $\frac{\pi}{2}\varepsilon < \delta$, 进而

$$\begin{aligned}|I_\varepsilon - \int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(0) \cdot \frac{1}{x^2+1}dx| &\le \int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot |\varphi(\varepsilon x) - \varepsilon(0)| \cdot \frac{1}{x^2+1}dx \\ &< \varepsilon' \cdot \frac{1}{\pi\varepsilon} \cdot \pi\varepsilon \cdot 1 = \varepsilon'\end{aligned} \tag{5.7}$$

而

$$\int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(0) \cdot \frac{1}{x^2+1}dx \le \int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(0)dx = \varphi(0) \tag{5.8}$$

同时

$$\int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(0) \cdot \frac{1}{x^2+1}dx \ge \int_{|x|<\frac{\pi}{2}\varepsilon} \frac{1}{\pi\varepsilon} \cdot \varphi(0) \cdot \frac{1}{1+\frac{\pi^2}{4}\varepsilon^2}dx = \frac{\varphi(0)}{1+\frac{\pi^2}{4}\varepsilon^2} \tag{5.9}$$

故令 $\varepsilon', \varepsilon \to 0$, 联立(5.6)-(5.9)式即得

$$\int_\mathbb{R} \frac{1}{\pi} \cdot \frac{\varphi(x)}{x^2+\varepsilon^2}dx \to \varphi(0), \quad \varepsilon \to 0, \forall \varphi \in \mathscr{D}(\mathbb{R})$$

命题即证.

(2) $\frac{1}{2\sqrt{\pi t}}\exp\left(-\frac{x^2}{4t}\right) \to \delta(x)(t \to 0^+)$.

证明

在 Gauss 核(5.10)中取 $j = \frac{1}{4\pi t}$ 即可.

练习 5.5 设 $\Omega \subset \mathbb{R}^n$ 是一个开集, 又设 K 是 Ω 的一个紧子集, 求证: 存在一个函数 $\varphi \in C_0^\infty(\Omega)$, 使得 $0 \le \varphi(x) \le 1$, 且 $\varphi(x)$ 在 K 的一个邻域内恒为 1.

证明

根据实变函数论的结论知存在 $C_0^0(\Omega)$ 内的函数 φ 满足 $0 \le \varphi(x) \le 1$, 且在 K 的一个邻域内恒为 1. 现取 $\varphi_j = \varphi * j$, 其中 j 是磨光算子. 知 φ_j 即为欲求.

5.2 B_0 空间

下面对 $\mathscr{D}(\Omega)$ 的收敛性作更细致的研究. 设 K 是相对于 Ω 的紧集, 再设 $C^\infty(\Omega)$ 是 Ω 上的无穷次可微函数全体, 引入

$$\mathscr{D}_K = \{\varphi \in C^\infty(\Omega) : \operatorname{supp}(\varphi) \subset K\}$$

规定其收敛性 $\varphi_j \to \varphi_0$ 是指对任意的多重指标 α 都有

$$\max_{x\in K}|\partial^\alpha(\varphi_j - \varphi_0)(x)| \to 0, \quad j \to \infty$$

已经知道这种收敛性不能通过单一的范数来刻画, 但可以利用 $C^\infty(\Omega)$ 的性质, 引入可数个范数:

$$\|\varphi\|_m = \sum_{|\alpha|\le m}\max_{x\in K}|\partial^\alpha\varphi(x)|, \quad m = 1, 2, \cdots \tag{5.10}$$

$\mathscr{D}_K$ 上的收敛性可以由这可数多个范数 $\{\|\varphi\|_m\}$ 来刻画: $\varphi_j \to \theta(\mathscr{D}_K)$ 当且仅当对任意的 $m \in \mathbb{N}$, 有

$$\|\varphi_j\|_m \to 0, \quad j \to \infty$$

也即

$$\forall\varepsilon > 0 \forall m \in \mathbb{N} \exists N = N(\varepsilon, m) \forall j > N(\|\varphi_j\|_m < \varepsilon)$$

定义 5.2.1 (可数范数空间 (B_0^* 空间))

设 $\mathscr{X}$ 是一个线性空间, 称它是可数范数空间 (或 B_0^* 空间), 是指在它上面有可数个半范数 $\{\|\cdot\|_m\}_{m=1}^\infty$, 满足:

1. $\|x + y\|_m \le \|x\|_m + \|y\|_m (\forall x, y \in \mathscr{X})$;
2. $\|\lambda x\|_m = |\lambda| \cdot \|x\|_m (\lambda \in \mathbb{R}, x \in \mathscr{X})$;
3. $\|x\|_m \ge 0, \|\theta\|_m = 0 (\forall x \in \mathscr{X})$;
4. $\|x\|_m = 0 (m = 1, 2, \cdots) \Leftrightarrow x = \theta$.

注

1. 事实上每个 $\|\cdot\|_m$ 都是半范数.
2. 在上述定义中, 这些半范数之间可以有序关系:

$$\|x\|_1 \le \|x\|_2 \le \cdots \le \|x\|_m \le \cdots, \quad \forall x \in \mathscr{X}$$

定义 5.2.2 (可数范数的等价)

在线性空间 $\mathscr{X}$ 上给定两组可数个半范数

$$\{\|\cdot\|_m\}_{m=1}^{\infty} \quad 与 \quad \{\|\cdot\|'_m\}_{m=1}^{\infty}$$

如果它们导出相同的收敛性, 则称它们是等价的.

命题 5.2.1

在线性空间 $\mathscr{X}$ 上, 为了两组可数个半范数

$$\{\|\cdot\|_m\}_{m=1}^{\infty} \quad 与 \quad \{\|\cdot\|'_m\}_{m=1}^{\infty}$$

是等价的, 必须且仅须: $\forall m \in \mathbb{N}, \exists m' \in \mathbb{N}$ 及 $\exists C_{mm'} > 0$, 使得

$$\|x\|_m \leq C_{mm'}\|x\|'_{m'}, \quad \forall x \in \mathscr{X}$$

且 $\forall n' \in \mathbb{N}, \exists n \in \mathbb{N}$ 及 $\exists C'_{n'n} > 0$, 使得

$$\|x\|'_{n'} \leq C'_{n'n}\|x\|_n, \quad \forall x \in \mathscr{X}.$$

证明

当 $\{\|\cdot\|_m\}_{m=1}^{\infty}$ 与 $\{\|\cdot\|'_m\}_{m=1}^{\infty}$ 等价, 用反证法. 不妨设

$$\exists m \in \mathbb{N} \forall m' \in \mathbb{N} \forall C_{mm'} > 0 \exists x_{m'} \in \mathscr{X}(\|x_{m'}\|_m > C_{mm'}\|x_{m'}\|'_m)$$

此时知 $\|\cdot\|_m$ 与 $\|\cdot\|'_m$ 不等价, 自然 $\{\|\cdot\|_m\}_{m=1}^{\infty}$ 与 $\{\|\cdot\|'_m\}_{m=1}^{\infty}$ 也不等价, 矛盾.

当对任意的 m 而言, $\|\cdot\|_m$ 与 $\|\cdot\|'_m$ 均等价, 显见 $\{\|\cdot\|_m\}_{m=1}^{\infty}$ 与 $\{\|\cdot\|'_m\}_{m=1}^{\infty}$ 等价. 命题即证. □

命题 5.2.2

每个 B_0^* 空间 $\mathscr{X}$ 必是一个 F^* 空间. 即若 $\{\|\cdot\|_m\}_{m=1}^{\infty}$ 是可数个半范数, 则

$$\|x\| = \sum_{m=1}^{\infty} \frac{1}{2^m} \cdot \frac{\|x\|_m}{1+\|x\|_m}, \quad \forall x \in \mathscr{X}$$

是一个准范数, 且 $\|\cdot\|$ 导出的收敛性与 $\{\|\cdot\|_m\}_{m=1}^{\infty}$ 导出的收敛性一致.

下面举一些 B_0^* 空间的例子.

例 5.11 $\mathscr{D}_K$ 是 B_0^* 空间, 其可数范数 $\|\cdot\|_m$ 按(5.10)式定义.

例 5.12($\mathscr{E}(\Omega)$) 设 Ω 是 $\mathbb{R}^n$ 中的任意开集, 又设 K_m 是一列相对于 Ω 的紧集, 满足

$$K_1 \subset K_2^\circ \subset K_2 \subset K_3^\circ \subset \cdots \subset K_m \subset \cdots \subset \Omega, \quad \bigcup_{m=1}^{\infty} K_m = \Omega$$

并令

$$\|\varphi\|_m = \sum_{|\alpha|\leq m} \max_{x\in K_m} |\partial^\alpha \varphi(x)|, \quad m = 1, 2, \cdots$$

用 $\mathscr{E}(\Omega)$ 表示带有可数范数 $\{\|\varphi\|_m\}_{m=1}^{\infty}$ 的线性空间 $C^\infty(\Omega)$, 则显见 $\mathscr{E}(\Omega)$ 是一个 B_0^* 空间. 根据定义知:

$$\varphi_j \to 0(\mathscr{E}(\Omega)) \Leftrightarrow \forall \varepsilon > 0 \forall m \in \mathbb{N} \exists N = N(\varepsilon, m) \forall |\alpha| \leq m \forall j > N(\max_{x\in K_m} |\partial^\alpha \varphi_j(x)| < \varepsilon)$$

也即在每个相对于 Ω 的紧集 K_m 上, 一直到 m 次导数都是一致收敛到 0 的.

注 从 $\{K_m\}$ 的严格递增性知 $\mathscr{E}(\Omega)$ 上的收敛性与 $\{K_m\}$ 的具体选择是无关的.

例 5.13(Schwartz 速降函数空间 $\mathscr{S}(\mathbb{R}^n)$) 用 $\mathscr{S}(\mathbb{R}^n)$ 表示集合

$$\{\varphi \in C^\infty(\mathbb{R}^n) : \sup |(1+|x|^2)^{\frac{k}{2}} |\partial^\alpha \varphi(x) \leq M_{k,\alpha} < \infty, k, |\alpha| = 0, 1, 2, \cdots\}$$

定义半范数为

$$\|\varphi\|_m = \sup_{\substack{|\alpha|\leq m \\ x\in\mathbb{R}^n}} |(1+|x|^2)^{\frac{m}{2}} \partial^\alpha \varphi(x)|, \quad m = 0, 1, 2, \cdots$$

$\mathscr{S}(\mathbb{R}^n)$ 上的函数称为速降函数, 其满足对任意的非负整数 m 与多重指标 α 都有:

$$\lim_{|x|\to\infty} (1+|x|^2)^{\frac{m}{2}} |\partial^\alpha \varphi(x)| = 0.$$

定义 5.2.3 (完备性)

一个 B_0^* 空间 $\mathscr{X}$ 称为是完备的, 是指其中的任意基本列都是收敛的. 完备的 B_0^* 空间称为 B_0 空间. ♣

例 5.14 $\mathscr{S}(\mathbb{R}^n)$ 是 B_0 空间.

证明

设 $\{\varphi_\nu(x)\}$ 是 $\mathscr{S}(\mathbb{R}^n)$ 中的一个基本列. 由定义知

$$\forall m \in \mathbb{N} \forall \varepsilon > 0 \exists N = N(m, \varepsilon) \forall \mu, \nu > N(\sup_{\substack{|\alpha|\leq m \\ x\in\mathbb{R}^n}} (1+|x|^2)^{\frac{m}{2}} |\partial^\alpha (\varphi_\mu - \varphi_\nu)(x)| < \varepsilon)$$

固定 m, 知 $\{\|\varphi_\nu\|_m\}$ 有界, 进而存在常数 M_m 使得

$$\sup_{\substack{|\alpha|\leq m \\ x\in\mathbb{R}^n}} (1+|x|^2)^{\frac{m}{2}} |\partial^\alpha \varphi_\nu(x)| \leq M_m, \quad \nu = 1, 2, \cdots$$

同时在任意有界闭球 $|x| \leq R$ 上, $\{\partial^\alpha \varphi_\nu(x)\}$ 依一致收敛范数是基本列, 故在 $|x| \leq R$ 上 $\{\partial^\alpha \varphi_\nu(x)\}$ 有一致极限 $\psi_\alpha(x)$, 满足:

$$|\psi_\alpha(x)| \leq \frac{M_m}{(1+|x|^2)^{\frac{m}{2}}}, \quad |\alpha| \leq m$$

现在证明 $\psi_\alpha(x) = \partial^\alpha \psi_0(x)$, 这只需证明

$$\psi_{(1,0,\cdots,0)}(x) = \partial_{x_1} \psi_0(x)$$

即可, 其余部分用归纳法可推知. 根据前述讨论, 任取 $R > 0$, 当 $|x| \leq R$ 时, 有:

$$\partial_{x_1} \varphi_\nu(x) \rightrightarrows \psi_{(1,0,\cdots,0)}(x), \quad \nu \to \infty$$

及

$$\varphi_\nu(x) \rightrightarrows \psi_0(x), \quad \nu \to \infty$$

现需证明

$$\varphi_\nu(x) \to \int_{-\infty}^{x_1} \psi_{(1,0,\cdots,0)}(x', x_2, \cdots, x_n) dx', \quad \nu \to \infty$$

根据积分收敛的必要条件知

$$\forall \varepsilon > 0 \exists M > 0 (\int_{|x_1|>M} \frac{dx_1}{(1+|x_1|^2)^{\frac{1}{2}}} < \varepsilon)$$

由极限定义知

$$\exists N_0 > 0 \forall \nu > N_0 (|\psi_{(1,0,\cdots,0)}(x) - \partial_{x_1} \varphi_\nu(x)| < \varepsilon), \quad |x| \leq M$$

故

$$
\begin{aligned}
&|\int_{-\infty}^{x_1}\psi_{(1,0,\cdots,0)}(x',x_2,\cdots,x_n)dx'-\varphi_\nu(x_1,x_2,\cdots,x_n)|\\
=&|\int_{-\infty}^{x_1}(\psi_{(1,0,\cdots,0)}(x',x_2,\cdots,x_n)-\partial_{x_1}\varphi_\nu(x',x_2,\cdots,x_n))dx'|\\
\leq&\int_{|x'|>M}|\psi_{(1,0,\cdots,0)}(x',x_2,\cdots,x_n)|dx'+\int_{|x'|>M}|\partial_{x_1}\varphi_\nu(x',x_2,\cdots,x_n)|dx'\\
&+\int_{|x'|\leq M}|\psi_{(1,0,\cdots,0)}(x',x_2,\cdots,x_n)-\partial_{x_1}\varphi_\nu(x',x_2,\cdots,x_n)|dx'<2\varepsilon+2M\varepsilon
\end{aligned}
$$

由 ε 的任意性即知

$$\psi_0(x)=\int_{-\infty}^{x_1}\psi_{(1,0,\cdots,0)}(x',x_2,\cdots,x_n)dx'$$

也即

$$\psi_{(1,0,\cdots,0)}(x)=\partial_{x_1}\psi_0(x)$$

这同时说明 $\psi_0\in\mathscr{S}(\mathbb{R}^n)$.

最后说明 $\|\varphi_\nu-\psi_0\|_m\to 0(\nu\to\infty)$. 此即欲证

$$\forall\varepsilon>0\exists N>0\forall\nu>N(\sup_{|\alpha|\leq m}|\partial^\alpha(\varphi_\nu(x)-\psi_0(x))|<\varepsilon)$$

对任意的 $\varepsilon>0$, 知可取 R 足够大, 使得当 $|x|>R$ 时, 由 $\varphi_\nu,\psi_0\in\mathscr{S}(\mathbb{R}^n)$ 知:

$$\sup_{|\alpha|\leq m}|\partial^\alpha(\varphi_\nu(x)-\psi_0(x))|\leq\frac{2M_m}{(1+|x|^2)^{\frac{m}{2}}}<\varepsilon$$

又根据前述讨论, 当 $|x|\leq R$ 时存在 $N>0$ 使得

$$\sup_{|\alpha|\leq m}|\partial^\alpha(\varphi_\nu(x)-\psi_0(x))|<\varepsilon$$

故

$$\partial^\alpha\varphi_\nu(x)\rightrightarrows\partial^\alpha\psi_0(x),\quad \nu\to\infty,\forall x\in\mathbb{R}^n$$

从而基本列 $\{\varphi_\nu(x)\}$ 在 $\mathscr{S}(\mathbb{R}^n)$ 中有极限 $\psi_0(x)$. 命题得证. □

注意 $\mathscr{D}(\Omega)$ 不是 B_0^* 空间, 但其上的收敛性有下述命题.

命题 5.2.3

要使得 $\varphi_\nu\to\varphi_0(\mathscr{D}(\Omega))$, 必须且仅须存在紧集 $K\subset\Omega$, 使得 $\varphi_0,\varphi_\nu\subset\mathscr{D}_K$, 且

$$\|\varphi_\nu-\varphi_0\|_{m,K}\to 0,\quad \nu\to\infty,,=1,2,\cdots$$

其中 $\|\varphi\|_{m,K}$ 是 $\mathscr{D}_K$ 上的可数范数.

证明

当存在紧集 K 使得 $\varphi_0,\varphi_\nu\subset\mathscr{D}_K$, 且 $\varphi_\nu\to\varphi_0(\mathscr{D}_K(\Omega))$ 时, 因为 $\mathscr{D}_K(\Omega)$ 上的收敛性是强于 $\mathscr{D}(\Omega)$ 上的收敛性的, 故必有 $\varphi_\nu\to\varphi_0(\mathscr{D}(\Omega))$.

当 $\varphi_\nu\to\varphi_0(\mathscr{D}(\Omega))$, 可取紧集 K 满足 $\mathrm{supp}(\varphi_\nu)\subset K,\forall\nu\in\mathbb{N}$, 且 $\mathrm{supp}(\varphi_0)\subset K$. 套用 $\mathscr{D}_K(\Omega)$ 上收敛的定义即得命题. □

正如前面讨论 $\mathscr{D}(\Omega)$ 时顺带讨论了广义函数全体 $\mathscr{D}'(\Omega)$ 一样, 对 $\mathscr{D}_K(\Omega),\mathscr{E}(\Omega),\mathscr{S}(\mathbb{R}^n)$, 也可以考虑它们的共轭空间, 分别记作 $\mathscr{D}_K'(\Omega),\mathscr{E}'(\Omega),\mathscr{S}'(\mathbb{R}^n)$. 特别当 $\Omega=\mathbb{R}^n$ 时简单记作 $\mathscr{D}_K',\mathscr{E},\mathscr{S}'$. 显见

$$\mathscr{E}'\subset\mathscr{S}'\subset\mathscr{D}_K'.$$

首先需要论证这些空间中有足够多的元素, 再设法把它们表示出来. 下面把 B 空间上线性泛函连续性与有界

性的关系推广到 B_0 空间.

引理 5.2.1

设 $\mathscr{X}$ 是一个 B_0 空间. 为了 $\mathscr{X}$ 上的线性泛函 f 是连续的, 必须且仅须 $\exists m \in \mathbb{N}$ 及 $M_m > 0$, 使得

$$|\langle f, \varphi\rangle| \le M_m\|\varphi\|_m, \quad \forall \varphi \in \mathscr{X}$$

证明

从有界性条件显得连续性. 现设 f 是连续的, 用反证法. 若

$$\forall n \in \mathbb{N} \exists x_n \in \mathscr{X}(|\langle f, x_n\rangle| > n\|x_n\|_n)$$

当存在 $N \in \mathbb{N}$ 使得 $\|x_n\|_n \ne 0 (\forall n \ge N)$ 时, 令

$$y_n := \frac{x_n}{n\|x_n\|_n}, \quad n \ge N$$

则当 $n \ge \max(N, p)$ 时有

$$\|y_n\|_p = \frac{\|x_n\|_p}{n\|x_n\|_n} \le \frac{1}{n}$$

故 $y_n \to \theta (n \to \infty)$, 但同时 $|\langle f, y_n\rangle| > 1 (n \ge N)$, 这与 f 连续矛盾!

当存在 $\{n_i\}_{i=1}^{\infty}$ 使得 $\|x_{n_i}\|_{n_i} = 0 (i = 1, 2, \cdots)$, 则知 $|\langle f, x_{n_i}\rangle| > 0$. 令

$$y_{n_i} = \frac{x_{n_i}}{\langle f, x_{n_i}\rangle}, \quad i = 1, 2, \cdots$$

得 $\langle f, y_{n_i}\rangle = 1$, 但 $\|y_{n_i}\|_{n_i} = 0 (i = 1, 2, \cdots)$, 这与 f 连续矛盾! 命题即证. □

定理 5.2.1

任意 B_0 空间 $\mathscr{X}$ 都有足够多的连续线性泛函.

证明

若 f_0 在 $\mathscr{X}$ 的一个线性闭子空间 $\mathscr{X}_0$ 上有定义且是连续线性的, 则由引理(5.2.1)知存在半范数 $\|\cdot\|_m$ 及常数 $M_m > 0$, 使得

$$|\langle f_0, x\rangle| \le M_m\|x\|_m, \quad \forall x \in \mathscr{X}_0$$

对半范数 $\|\cdot\|_m$ 应用 Hahn-Banach 定理(3.5.4)知 f_0 可以扩张到全空间 $\mathscr{X}$ 上, 即存在 f 满足 $f|_{\mathscr{X}_0} = f_0$, 且

$$|\langle f, x\rangle| \le M_m\|x\|_m, \quad \forall x \in \mathscr{X}$$

故 $f \in \mathscr{X}^*$. 这说明

$$\forall x_0 \ne \theta \exists f \in \mathscr{X}^*(\langle f, x_0\rangle = 1).$$

□

由定义即得下述命题.

命题 5.2.4

设 $\mathscr{X}$ 是 B_0 空间, 若 $D(\Omega) \hookrightarrow \mathscr{X}$, 即

$$\mathscr{D}(\Omega) \subset \mathscr{X} \quad 且 \quad \varphi_j \to \varphi_0(\mathscr{D}(\Omega)) \Rightarrow \varphi_j \to \varphi_0(\mathscr{X}), j \to \infty$$

则 $\mathscr{X}$ 上的任意一个连续线性泛函 f 都满足 $f \in \mathscr{D}'(\Omega)$.

例 5.15 $\mathscr{X} = \mathscr{E}(\Omega), L^p(\Omega)(1 \le p < \infty)$ 或 $C^k(\Omega)(k \in \mathbb{N})$ 等都满足 $\mathscr{D}(\Omega) \hookrightarrow \mathscr{X}$, 故 $\mathscr{E}'(\Omega), L^q(\Omega)(\frac{1}{p}+\frac{1}{q} = 1), [C^k(\Omega)]^*$ 等都包含在 $\mathscr{D}'(\Omega)$ 中.

现在用积分表示 $\mathscr{S}'$ 中的广义函数.

定理 5.2.2

要使得 $f \in \mathscr{S}'$, 必须且仅须 $\exists m \in \mathbb{N}$ 及 $u_\alpha \in L^2(\mathbb{R}^n)(|\alpha| \le m)$, 使得

$$\langle f, \varphi \rangle = \sum_{|\alpha|\le m} \int_{\mathbb{R}^n} u_\alpha(x) \partial^\alpha \varphi(x)(1+|x|^2)^{\frac{m}{2}}\, dx, \quad \forall \varphi \in \mathscr{S}.$$

证明

只需证明对任意的 $f \in \mathscr{S}'$ 都存在 $m \in \mathbb{N}, u_\alpha \in L^2(\mathbb{R}^n)$ 满足题式即可. 首先证明

$$\|\varphi\|'_m := (\sum_{|\alpha|\le m} \int_{\mathbb{R}^n} (1+|x|^2)^n |\partial^\alpha \varphi(x)|^2 dx)^{\frac{1}{2}}, \quad m = 1, 2, \cdots$$

是 $\mathscr{S}$ 上的一组等价范数. 这是因为

$$\|\varphi\|'^2_m \le \sum_{|\alpha|\le m} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^{m+n} |\partial^\alpha \varphi(x)|^2 \cdot \int_{\mathbb{R}^n} \frac{dx}{(1+|x|^2)^n} \le c\|\varphi\|^2_{m+n}$$

同时设 $e = (1, \cdots, 1)$, 有

$$\begin{aligned}
\|\varphi_m\| &= \sup_{\substack{|\alpha|\le m\\ x\in\mathbb{R}^n}} |(1+|x|^2)^{\frac{m}{2}} \partial^\alpha \varphi(x)| = \sup_{\substack{|\alpha|\le m\\ x\in\mathbb{R}^n}} |\int_{\prod_{i=1}^n(-\infty, x_i]} \frac{\partial^n}{\partial x_1 \cdots \partial x_n} (1+|x|^2)^{\frac{m}{2}} \partial^\alpha \varphi(x) dx| \\
&\le \sup_{|\alpha|\le m} \int_{\mathbb{R}^n} |\frac{\partial^n}{\partial x_1 \cdots \partial x_n} (1+|x|^2)^{\frac{m}{2}} \partial^\alpha \varphi(x)| dx \le c \sup_{|\alpha|\le m} \int_{\mathbb{R}^n} (1+|x|^2)^{\frac{m}{2}} |\partial^{\alpha+e} \varphi(x)| dx \\
&= c \sup_{|\alpha|\le m} \int_{\mathbb{R}^n} (1+|x|^2)^{-\frac{n}{2}} \cdot (1+|x|^2)^{\frac{m+n}{2}} |\partial^{\alpha+e}\varphi(x)| dx \le c \sum_{|\alpha|\le m+n} \int_{\mathbb{R}^n} (1+|x|^2)^{-\frac{n}{2}} \cdot (1+|x|^2)^{\frac{m+n}{2}} |\partial^{\alpha+e}\varphi(x)| dx \\
&\le c \sum_{|\alpha|\le m+n} (\int_{\mathbb{R}^n} (1+|x|^2)^{m+n} |\partial^\alpha \varphi(x)|^2 dx)^{\frac{1}{2}} \cdot (\int_{\mathbb{R}^n} \frac{dx}{(1+|x|^2)^n})^{\frac{1}{2}} \le c_1 \|\varphi\|'_{m+n}
\end{aligned}$$

现在对 f 应用引理(5.2.1)知存在 $m \in \mathbb{N}$ 使得

$$|\langle f, \varphi \rangle| \le c_m \|\varphi\|'_m, \quad \forall \varphi \in \mathscr{S}$$

故 f 可以连续地扩张到以 $\|\cdot\|'_m$ 为范数的 Banach 空间 $\mathscr{X}_m$ 上. 注意 $\|\cdot\|'_m$ 满足平行四边形等式, 故可诱导内积

$$(\varphi, \psi)_m = \sum_{|\alpha|\le m} \int_{\mathbb{R}^n} \partial^\alpha \varphi(x) \cdot \partial^\alpha \psi(x) \cdot (1+|x|^2)^{\frac{m}{2}} dx$$

可以验证 $(\mathscr{X}_m, (\cdot, \cdot))$ 是 Hilbert 空间.

现在应用 Riesz 表示定理(3.3.1)知存在 $u \in \mathscr{X}_m$ 使得

$$\langle f, \varphi \rangle = \sum_{|\alpha|\le m} \int_{\mathbb{R}^n} \partial^\alpha u(x) \cdot \partial^\alpha \varphi(x) \cdot (1+|x|^2)^{\frac{m}{2}} dx$$

其中

$$\int_{\mathbb{R}^n} |\partial^\alpha u(x)|^2 (1+|x|^2)^m dx < \infty$$

令

$$u_\alpha(x) = (1+|x|)^{\frac{m}{2}} \partial^\alpha u(x), \quad |\alpha| \le m$$

则 $u_\alpha \in L^2(\mathbb{R}^n)$, 且对任意的 $\varphi \in \mathscr{S}$ 都有

$$\langle f, \varphi \rangle = \sum_{|\alpha|\le m} \int_{\mathbb{R}^n} u_\alpha(x) \partial^\alpha \varphi(x) (1+|x|^2)^{\frac{m}{2}} dx$$

□

5.3 广义函数的运算

5.3.1 知识梳理

定义 5.3.1 (广义函数的连续性)

设 $A:\mathscr{D}(\Omega)\to\mathscr{D}(\Omega)$ 是一个线性算子. 称 A 是连续的, 如果

$$\varphi_j\to\varphi(\mathscr{D}(\Omega))\Rightarrow A\varphi_j\to A\varphi(\mathscr{D}(\Omega)),\quad j\to\infty.$$

例 5.16(微分算子) 任意微分算子 ∂^α 是 $\mathscr{D}(\Omega)$ 上的连续线性算子.

证明

根据 $\mathscr{D}(\Omega)$ 收敛的定义, 存在相对于 Ω 的紧集 K 使得 $\varphi_k\to\varphi(\mathscr{D}_K(\Omega))$. 知 $\partial^\alpha:\mathscr{D}_K\to\mathscr{D}_K$, 且

$$\|\partial^\alpha\varphi\|_m=\sum_{|\beta|\le m}\max_{x\in K}|\partial^{\beta+\alpha}\varphi(x)|\le\|\varphi\|_{m+|\alpha|}$$

命题即证. □

例 5.17(乘法算子) 设 $\psi\in C^\infty(\Omega)$, 由 ψ 决定一个乘法算子:

$$A:\varphi\mapsto\psi\cdot\varphi,\quad\forall\varphi\in\mathscr{D}(\Omega)$$

那么 A 是连续线性的.

证明

知对任意的相对紧集 $K\subset\Omega$ 而言, $A:\mathscr{D}_K\to\mathscr{D}_K$, 且

$$\|\psi\cdot\varphi\|_m=\sum_{|\alpha|\le m}\max_{x\in K}|\partial^\alpha(\psi\cdot\varphi)(x)|\le\sum_{|\alpha|\le m}\max_{x\in K}\left|\sum_{\beta\le\alpha}\binom{\alpha}{\beta}\partial^\beta\psi(x)\cdot\partial^{\alpha-\beta}\varphi(x)\right|\le C(m,\psi)\cdot\|\varphi\|_m$$

在 $\mathscr{D}'(\Omega)$ 上定义算子 A^*:

$$\langle A^*f,\varphi\rangle=\langle f,A\varphi\rangle,\quad\forall\varphi\in\mathscr{D}(\Omega),\forall f\in\mathscr{D}'(\Omega)$$

显见 $A^*:\mathscr{D}'(\Omega)\to\mathscr{D}'(\Omega)$ 是连续的. 这是因为若 $f_j\to f(\mathscr{D}'(\Omega))(j\to\infty)$, 则对任意的 $\varphi\in\mathscr{D}(\Omega)$ 有

$$\langle A^*f_j,\varphi\rangle=\langle f_j,A\varphi\rangle\to\langle f,A\varphi\rangle=\langle A^*f,\varphi\rangle,\quad j\to\infty$$

此即 $A^*f_j\to A^*f$, $j\to\infty$. □

5.3.1.1 广义微商

定义 5.3.2 (广义微商)

称 $\tilde{\partial}^\alpha=(-1)^\alpha(\partial^\alpha)^*$ 为 α 阶广义微商运算, 即 $\forall f\in\mathscr{D}'(\Omega)$,

$$\langle\tilde{\partial}^\alpha f,\varphi\rangle=(-1)^{|\alpha|}\langle f,\partial^\alpha\varphi\rangle,\quad\forall\varphi\in\mathscr{D}(\Omega)$$

注

1. 若 $f(x)\in C^1(\Omega)$, 依照自然对应:

$$\langle f,\varphi\rangle=\int_\Omega f(x)\varphi(x)dx,\quad\forall\varphi\in\mathscr{D}(\Omega)$$

产生的广义函数仍然记作 f, 知 f 有广义微商 $\tilde{\partial}_{x_i}f$, 同时 $f(x)$ 也有通常意义下的微商 $\partial_{x_i}f$. 注意 $\tilde{\partial}_{x_i}f$ 正是 $\partial_{x_i}f$ 对应的广义函数, 这是因为

$$\langle\tilde{\partial}_{x_i}f,\varphi\rangle=-\langle f,\partial_{x_i}\varphi\rangle=-\int_\Omega f(x)\partial_{x_i}\varphi(x)dx=\int_\Omega\partial_{x_i}f(x)\varphi(x)dx,\quad\forall\varphi\in\mathscr{D}(\Omega)$$

2. 广义函数对微商运算是封闭的, 即任意广义函数 $f \in \mathscr{D}'(\Omega)$ 都可以做任意次广义微商. 注意对普通的局部可积函数而言, 就算没有通常意义下的微商, 也可以在广义函数的意义下求广义微商, 只是做广义微商后所得的函数未必对应通常意义下的函数.
3. 若 α, β 是任意两个多重指标, 则
$$\tilde{\partial}^\alpha \cdot \tilde{\partial}^\beta = \tilde{\partial}^{\alpha+\beta} = \tilde{\partial}^\beta \cdot \tilde{\partial}^\alpha$$
4. 由 $\tilde{\partial}^\alpha$ 的连续性:
$$f_j \to f_0 (j \to \infty) \Rightarrow \tilde{\partial}^\alpha f_j \to \tilde{\partial}^\alpha f_0 (f \to \infty)$$
知广义微商与极限总是可交换的.

例 5.18(Heaviside 函数) 若
$$Y(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0 \end{cases}$$
则
$$\tilde{\partial}_x Y(x) = \delta(x).$$

证明

知
$$\langle \tilde{\partial}_x Y(x), \varphi \rangle = -\langle Y, \partial_x \varphi \rangle = -\int_0^\infty \varphi'(x) dx = \varphi(0) = \langle \delta, \varphi \rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R})$$
□

例 5.19 若 $\delta^{(\alpha)}$ 是例(5.3)中引进的广义函数, 则
$$\tilde{\partial}^\alpha \delta = \delta^{(\alpha)}.$$

证明

知
$$\langle \tilde{\partial}^\alpha, \varphi \rangle = (-1)^{|\alpha|} \langle \delta, \partial^\alpha \varphi \rangle = (-1)^{|\alpha|} (\partial^\alpha \varphi)(\theta) = \langle \delta^{(\alpha)}, \varphi \rangle.$$
□

例 5.20 设 $\tilde{\Delta} = \tilde{\partial}_{x_1}^2 + \cdots + \tilde{\partial}_{x_n}^2$, 则
$$\begin{cases} \tilde{\Delta} |x|^{2-n} = (2-n)\Omega_n \delta(x), & n \ge 3 \\ \tilde{\Delta} \ln |x| = 2\pi \delta(x), & n = 2 \end{cases}$$
其中 Ω_n 是 $\mathbb{R}^n$ 中单位球面的面积.

证明

当 $n \ge 3$ 时, 对任意的 $\varphi \in \mathscr{D}(\mathbb{R}^n)$ 有
$$\begin{aligned} \langle \tilde{\Delta} |x|^{2-n}, \varphi \rangle = \langle |x|^{2-n}, \Delta \varphi \rangle &= \lim_{\varepsilon \to 0} \int_{|x| \ge \varepsilon} |x|^{2-n} \Delta \varphi(x) dx \\ &= \lim_{\varepsilon \to 0} [\int_{|x| \ge \varepsilon} \Delta |x|^{2-n} \varphi(x) dx + \int_{|x| = \varepsilon} (\varphi \frac{\partial |x|^{2-n}}{\partial |x|} - |x|^{2-n} \frac{\partial \varphi}{\partial |x|}) d\sigma] \end{aligned}$$
其中 $d\sigma$ 是球面 $\{x \in \mathbb{R}^n : |x| = \varepsilon\}$ 上的面积元. 注意
$$\Delta |x|^{2-n} = 0, \quad |x| \ne 0$$
且
$$\varepsilon^{2-n} \int_{|x|=\varepsilon} \frac{\partial \varphi}{\partial |x|} d\sigma = O(\varepsilon) \to 0, \quad \varepsilon \to 0$$
而
$$\varepsilon^{1-n} \int_{|x|=\varepsilon} \varphi(x) d\sigma \to \varphi(0) \Omega_n, \quad \varepsilon \to 0$$

即得

$$\langle \Delta |x|^{2-n}, \varphi \rangle = (2-n)\varphi(0)\Omega_n = (2-n)\Omega_n \langle \delta, \varphi \rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n)$$

类似可证

$$\langle \Delta \ln |x|, \varphi \rangle = 2\pi \langle \delta, \varphi \rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^2).$$

□

5.3.1.2 广义函数的乘法

定义 5.3.3 (广义函数乘法)

对于任意的 $\psi \in C^\infty(\Omega), f \in \mathscr{D}'(\Omega)$, 定义

$$\langle \psi f, \varphi \rangle = \langle f, \psi\varphi \rangle, \quad \forall \varphi \in \mathscr{D}(\Omega)$$

即定义广义函数 f 对 $C^\infty(\Omega)$ 函数 ψ 的乘法为 $\mathscr{D}(\Omega)$ 上惩罚算子的共轭算子. ♣

显见广义函数的乘法算子是连续算子.

例 5.21

$$x^n \tilde{\partial}^m \delta(x) = \begin{cases} (-1)^n \frac{m!}{(m-n)!} \delta^{(m-n)}(x), & m \geq n \\ 0, & m < n \end{cases}$$

证明

知

$$\begin{aligned}
\langle x^n \tilde{\partial}^m \delta(x), \varphi(x) \rangle &= (-1)^m \langle \delta(x), \partial^m (x^n \varphi(x)) \rangle = (-1)^m \sum_{r=0}^{m} \binom{m}{r} (\partial^r x^n)(\partial^{m-r} \varphi(x))|_{x=0} \\
&= \begin{cases} \frac{(-1)^m m!}{n!(m-n)!} n! (\partial^{m-n} \varphi)(0), & m \geq n \\ 0, & m < n \end{cases} \\
&= \begin{cases} \frac{(-1)^n m!}{(m-n)!} \langle \delta^{(m-n)}(x), \varphi \rangle, & m \geq n \\ 0, & m < n \end{cases}
\end{aligned}$$

□

可见当 $f \in L^1_{loc}(\Omega)$ 时, 对任意的 $\psi \in C^\infty(\Omega)$, ψf 的通常定义与作为广义函数的定义知一致的. 特别注意一般不能直接定义两个广义函数的乘积, 比如两个 δ 函数相乘的结果不再是广义函数.

5.3.1.3 平移算子与反射算子

定义 5.3.4 (平移算子)

$\forall x_0 \in \mathbb{R}^n$, 定义 $\tau_{x_0} : \mathscr{D}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)$ 为

$$(\tau_{x_0} \varphi)(x) = \varphi(x - x_0), \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n)$$

称 τ_{x_0} 为平移算子. 显见 $\tau_{x_0} \in \mathscr{L}(\mathscr{D}(\mathbb{R}^n))$. ♣

定义 5.3.5 (平移算子的推广)

$\forall x_0 \in \mathscr{R}^n$, $\tilde{\tau}_{x_0} := (\tau_{-x_0})^*$, 也即对 $\forall f \in \mathscr{D}'(\mathbb{R}^n)$ 有

$$\langle \tilde{\tau_{x_0}} f, \varphi \rangle = \langle f, \tau_{-x_0} \varphi \rangle = \langle f, \varphi(x + x_0) \rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n).$$

♣

注 之所以称 $\tilde{\tau}_{x_0}$ 是 τ_{x_0} 的推广，是因为取 $f(x) \in L^1_{loc}(\mathbb{R}^n)$，则对任意的 $x_0 \in \mathbb{R}^n$ 有

$$\int_{\mathbb{R}^n} f(x-x_0)\varphi(x)dx = \int_{\mathbb{R}^n} f(x)\varphi(x+x_0)dx = \langle f, \tau_{-x_0}\varphi\rangle = \langle \tilde{\tau}_{x_0} f, \varphi\rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n)$$

即 $\tilde{\tau}_{x_0} f = f(x-x_0) = \tau_{x_0} f$.

定义 5.3.6 (反射算子)

$\forall x \in \mathbb{R}^n$ 定义 $\sigma : \mathscr{D}(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)$ 为

$$(\sigma\varphi)(x) = \varphi(-x), \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n)$$

称 σ 为反射算子. 显见 $\sigma \in \mathscr{L}(\mathbb{D}(\mathbb{R}^n))$.

定义 5.3.7 (反射算子的推广)

$\tilde{\sigma} := \sigma^*$，即对 $\forall f \in \mathscr{D}'(\mathbb{R}^n)$，有

$$\langle \tilde{\sigma} f, \varphi\rangle = \langle \sigma^* f, \varphi\rangle = \langle f, \sigma\varphi\rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R}^n).$$

注 之所以称 $\tilde{\sigma}$ 是 σ 的推广，是因为取 $f(x) \in L^1_{loc}(\mathbb{R}^n)$，有

$$\int_{\mathbb{R}^n} f(-x)\varphi(x)dx = \int_{\mathbb{R}^n} f(x)\varphi(-x)dx = \langle f, \sigma\varphi\rangle = \langle \tilde{\sigma} f, \varphi\rangle, \quad \forall \varphi \in \mathscr{D}(\mathbb{R})^n$$

即 $\tilde{\sigma} f = f(-x) = \sigma f$.

5.4 $\mathscr{S}'$ 上的 Fourier 变换

5.4.1 知识梳理

定义 5.4.1 ($L^1(\mathbb{R}^n)$ 上的 Fourier 变换)

对于 $\varphi \in L^1(\mathbb{R}^n)$，定义 φ 的 Fourier 变换为:

$$(\mathscr{F}\varphi)(\xi) = \int_{\mathbb{R}^n} \varphi(x)\exp(-2\pi i x \cdot \xi)dx \tag{5.11}$$

其中 $x \cdot \xi = x_1\xi_1 + x_2\xi_2 + \cdots + x_n\xi_n$.

但注意对一般的 $L^p(\mathbb{R}^n)(p > 1)$ 函数而言，(5.11)式中的积分未必有意义. 同时就算令 $\varphi(x) \equiv 1$，(5.11)式也是没有意义的. 下面介绍广义函数的引进是可以扩大 Fourier 变换的定义的.

命题 5.4.1

$\mathscr{F} \in \mathscr{L}(\mathscr{S})$. 也即 Fourier 变换把 Schwartz 速降函数连续线性地映成 Schwartz 速降函数.

证明

容易验证:

$$\mathscr{F}(\partial^\alpha\varphi) = (2\pi i\xi)^\alpha(\mathscr{F}\varphi)(\xi), \quad \mathscr{F}((-2\pi i x)^\alpha\varphi)(\xi) = \partial^\alpha(\mathscr{F}\varphi)(\xi)$$

从而若取 m 为偶数, 则有[3]:

$$\begin{aligned}
\|\mathscr{F}\varphi\|_m &= \sup_{\substack{\xi\in\mathbb{R}^n \\ |\alpha|\le m}} (1+|\xi|^2)^{\frac{m}{2}}|\partial^\alpha(\mathscr{F}\varphi)(\xi)| = \sup_{\substack{\xi\in\mathbb{R}^n \\ |\alpha|\le m}} (1+|\xi|^2)^{\frac{m}{2}}|\mathscr{F}((-2\pi ix)^\alpha\varphi)(\xi)| \\
&= \sup_{\substack{\xi\in\mathbb{R}^n \\ |\alpha|\le m}} |(\mathscr{F}[(1-\frac{\Delta}{4\pi^2})^{\frac{m}{2}}(-2\pi ix)^\alpha\varphi])(\xi)| \le \sup_{|\alpha|\le m}\int_{\mathbb{R}^n}|(1-\frac{\Delta}{4\pi^2})^{\frac{m}{2}}(-2\pi ix)^\alpha\varphi|dx \\
&\le \sup_{|\alpha|\le m}\sum_{\substack{p\le\alpha \\ q\le m}} A_{\alpha,p,q}\int_{\mathbb{R}^n}|x^p\partial^q\varphi|dx \le M_m \sup_{\substack{x\in\mathbb{R}^n \\ |\alpha|\le m}} (1+|x|^2)^{\frac{m}{2}+n}|\partial^\alpha\varphi(x)| \le M_m\|\varphi\|_{m+2n}
\end{aligned}$$

其中 $A_{\alpha,p,q}$ 与 M_m 都是常数. □

定义 5.4.2 (Fourier 逆变换 (Fourier 积分))

称积分

$$(\overline{\mathscr{F}}(\varphi))(\xi) = \int_{\mathbb{R}^n}\varphi(x)\exp(2\pi ix\cdot\xi)dx$$

为函数 φ 的 Fourier 逆变换, 或 Fourier 积分.

♣

命题 5.4.2

$\overline{\mathscr{F}} = \sigma\mathscr{F}$, 故 $\overline{\mathscr{F}} \in \mathscr{L}(\mathscr{S})$.

♠

命题 5.4.3

若 $\varphi,\psi\in\mathscr{S}$, 则

$$\langle\mathscr{F}\varphi,\psi\rangle = \langle\varphi,\mathscr{F}\psi\rangle,$$
$$\langle\overline{\mathscr{F}}\varphi,\psi\rangle = \langle\varphi,\overline{\mathscr{F}}\psi\rangle.$$

♠

证明

知

$$\langle\mathscr{F}\varphi,\psi\rangle = \int_{\mathbb{R}^n}[\int_{\mathbb{R}^n}\varphi(x)\exp(-2\pi ix\cdot\xi)dx]\psi(\xi)d\xi = \int_{\mathbb{R}^n}\varphi(x)[\int_{\mathbb{R}^n}\exp(-2\pi ix\cdot\xi)\psi(\xi)d\xi]dx = \langle\varphi,\mathscr{F}\psi\rangle.$$

同理可证 $\langle\overline{\mathscr{F}}\varphi,\psi\rangle = \langle\varphi,\overline{\mathscr{F}}\psi\rangle$. □

定义 5.4.3 (广义 Fourier(逆) 变换)

在空间 $\mathscr{S}'$ 上定义 $\widetilde{\mathscr{F}} = \mathscr{F}^*(\widetilde{\overline{\mathscr{F}}} = (\overline{\mathscr{F}})^*)$, 称为广义 Fourier(逆) 变换. 在不引起混淆时, 记号 $\sim$ 可以略去. 为了方便简记 $\mathscr{F}\varphi = \hat{\varphi}$.

♣

根据定义可得下述命题.

命题 5.4.4

$\mathscr{F}\in\mathscr{L}(\mathscr{S}'), \overline{\mathscr{F}}\in\mathscr{L}(\mathscr{S}')$. 且当限制在 $\mathscr{S}$ 上时, 它们分别与通常意义下的 Fourier 变换, Fourier 逆变换一致.

♠

Fourier 变换与微商, 乘法, 平移, 相移之间都有联系.

命题 5.4.5

设 $f,g\in\mathscr{S}'$, 且 $g=\mathscr{F}f$, 则:

1. $\mathscr{F}(\tilde{\partial}^\alpha f) = (2\pi i\xi)^\alpha g$.
2. $\mathscr{F}((-2\pi ix)^\alpha f) = \tilde{\partial}^\alpha g$.

[3] 第三排从 $A_{\alpha,p,q}$ 到 M_m 的式子没有看懂.

3. $\mathscr{F}(\tilde{\tau}_a f) = \exp(-2\pi i a \xi) g$.
4. $\mathscr{F}(\exp(2\pi i a \cdot x) f) = \tilde{\tau}_a g$.

例 5.22(Gauss 核的 Fourier 变换)

$$\mathscr{F}(\exp(-\pi|x|^2)) = \exp(-\pi|\xi|^2).$$

证明

不妨设 $n = 1$, 令 $f(x) = \exp(-\pi x^2)$, 知

$$f'(x) + 2\pi x f(x) = 0, \quad f(0) = 1$$

方程两边做 Fourier 变换有

$$2\pi i \xi(\mathscr{F} f)(\xi) + i(\mathscr{F} f)'(\xi) = 0, \quad (\mathscr{F} f)(0) = \int_{\mathbb{R}} \exp(-\pi x^2) dx = 1$$

这说明 $\mathscr{F} f$ 与 f 满足同一方程且具有相同初值, 故由 ODE 初值问题解的唯一性即得

$$(\mathscr{F} f)(\xi) = f(\xi) = \exp(-\pi|\xi|^2), \quad \forall \xi \in \mathbb{R}$$

对任意的 $n \in \mathbb{N}$, 分离变量即得相同结论. □

例 5.23

$$\mathscr{F}\delta = 1, \quad \overline{\mathscr{F}}\delta = 1.$$

证明

注意对任意的 $\varphi \in \mathscr{S}$ 而言:

$$\langle \mathscr{F}\delta, \varphi \rangle = \langle \delta, \mathscr{F}\varphi \rangle = (\mathscr{F}\varphi)(0) = \int_{\mathbb{R}^n} \varphi(x) dx = \langle 1, \varphi \rangle.$$

类似可证 $\overline{\mathscr{F}}\delta = 1$. □

例 5.24

$$\mathscr{F}(1) = \delta, \overline{\mathscr{F}}(1) = \delta.$$

证明

由例(5.22)知

$$\mathscr{F}[\exp\left(-\pi \frac{|x|^2}{m}\right)] = m^{\frac{n}{2}} \exp(-m\pi|\xi|^2)$$

令 $m \to \infty$ 知

$$\exp\left(-\pi \frac{|x|^2}{m}\right) \to 1(\mathscr{S}'), \quad m^{\frac{n}{2}} \exp(-m\pi|\xi|^2) \to \delta(\mathscr{S}')$$

由 $\mathscr{F}$ 的连续性即得 $\mathscr{F}(1) = \delta$. $\overline{\mathscr{F}}(1) = \delta$ 同理可证. □

例 5.25

(i) $\mathscr{F}(p(x)) = p(\frac{i}{2\pi}\tilde{\partial})\delta(\xi)$, 其中 $p(\cdot)$ 表示多项式.

(ii) $\mathscr{F}(\tilde{\partial}^\alpha \xi) = (2\pi i \xi)^\alpha$.

证明

(i) 设 $p(x) = \sum\limits_{k=1}^{n} c_k x^k$, 知

$$\mathscr{F}(p(x)) = \sum_{k=1}^{n} c_k \mathscr{F}(x^k) = \sum_{k=1}^{n} \frac{c_k}{(-2\pi i)^k} \tilde{\partial}\delta(\xi) = p(\frac{i}{2\pi}\tilde{\partial})\delta(\xi)$$

(ii) 知

$$\mathscr{F}(\tilde{\partial}^\alpha \delta) = (2\pi i\xi)^\alpha \mathscr{F}\delta = (2\pi i\xi)^\alpha$$

□

定理 5.4.1

$\overline{\mathscr{F}} = \mathscr{F}^{-1}$, 也即 $\overline{\mathscr{F}}\mathscr{F} = \mathscr{F}\overline{\mathscr{F}} = I$.

证明

任取 $\varphi \in \mathscr{S}, y \in \mathbb{R}^n$, 有:

$$\begin{aligned}
\varphi(y) &= \langle \delta, \tau_{-y}\varphi \rangle = \langle \tilde{\tau}_y \delta, \varphi \rangle \\
&= \langle \mathscr{F}(\exp(2\pi i\xi \cdot y)), \varphi \rangle = \langle \exp(2\pi i\xi \cdot y), (\mathscr{F}\varphi)(\xi) \rangle \\
&= \int_{\mathbb{R}^n} \exp(2\pi i\xi \cdot y)(\mathscr{F}\varphi)(\xi) d\xi = (\overline{\mathscr{F}}\mathscr{F}\varphi)(y)
\end{aligned}$$

此即 $\varphi = \overline{\mathscr{F}}\mathscr{F}\varphi$, 同理可证 $\varphi = \mathscr{F}\overline{\mathscr{F}}\varphi$. 这说明对任意的 $f \in \mathscr{S}'$ 有

$$\begin{aligned}
\langle \mathscr{F}\overline{\mathscr{F}} f, \varphi \rangle &= \langle f, \overline{\mathscr{F}}\mathscr{F}\varphi \rangle = \langle f, \varphi \rangle, \quad \forall \varphi \in \mathscr{S}. \\
\langle \overline{\mathscr{F}}\mathscr{F} f, \varphi \rangle &= \langle f, \mathscr{F}\overline{\mathscr{F}}\varphi \rangle = \langle f, \varphi \rangle, \quad \forall \varphi \in \mathscr{S}.
\end{aligned}$$

此即 $\overline{\mathscr{F}} = \mathscr{F}^{-1}$. □

推论 5.4.1 (Plancherel 定理)

若 $f \in L^2$, 则 $\widetilde{\mathscr{F}} f \in L^2$, 且

$$\|f\|_{L^2} = \|\widetilde{\mathscr{F}} f\|_{L^2}$$

即 $\widetilde{\mathscr{F}}$ 保持范数不变.

证明

因为任取 $\varphi \in \mathscr{S}$, 有:

$$\|\varphi\|_{L^2}^2 = \langle \varphi, \overline{\varphi} \rangle = \langle \overline{\mathscr{F}}\mathscr{F}\varphi, \overline{\varphi} \rangle = \langle \mathscr{F}\varphi, \overline{\mathscr{F}\varphi} \rangle = \|\mathscr{F}\varphi\|_{L^2}^2$$

而已经知道 $\mathscr{S}$ 在 L^2 中稠密, 故对任意的 $f \in L^2$, 总存在 $\{\varphi_m\} \subset \mathscr{S}$ 使得

$$\|\varphi_m - f\|_{L^2} \to 0, \quad m \to \infty$$

故对任意的 $p \in \mathbb{N}$ 有

$$\|\mathscr{F}\varphi_{m+p} - \mathscr{F}\varphi_m\|_{L^2} \to 0, \quad m \to \infty$$

即 $\{\mathscr{F}\varphi_m\}_{m=1}^\infty$ 是 L^2 中的基本列. 因为 $\mathscr{F}$ 在 $\mathscr{S}'$ 中连续, 而 $L^2 \hookrightarrow \mathscr{S}'$, 知在 L^2 的意义下有

$$\widetilde{\mathscr{F}} f = \lim_{m\to\infty} \mathscr{F}\varphi_m$$

这说明 $\widetilde{\mathscr{F}} f \in L^2$, 且

$$\|\widetilde{\mathscr{F}} f\|_{L^2} = \lim_{m\to\infty} \|\mathscr{F}\varphi_m\|_{L^2} = \lim_{m\to\infty} \|\varphi_m\|_{L^2} = \|f\|_{L^2}.$$

□

注 由极化恒等式与 Plancherel 定理(9.3.1)可知若 $f, g \in L^2$, 则

$$(f, g)_{L^2} = (\widetilde{\mathscr{F}} f, \widetilde{\mathscr{F}} g)_{L^2}$$

也即 $\widetilde{\mathscr{F}}$ 保持内积不变.

5.5 Sobolev 空间与嵌入定理

定义 5.5.1 (Sobolev 空间)

设 $\Omega \subset \mathbb{R}^n$ 是一个开集, m 是非负整数, $1 \le p < \infty$, 称集合

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : \tilde{\partial}^\alpha u \in L^p(\Omega), |\alpha| \le m\}$$

按范数

$$\|u\|_{m,p} = (\sum_{|\alpha| \le m} \|\tilde{\partial}^\alpha u\|_{L^p(\Omega)}^p)^{\frac{1}{p}} = (\sum_{|\alpha| \le m} \int_\Omega |\tilde{\partial}^\alpha u(x)|^p dx)^{\frac{1}{p}}$$

构成的空间为 Sobolev 空间, 记作 $W^{m,p}(\Omega)$ 或 $W_p^m(\Omega)$.

♣

定理 5.5.1

空间 $W^{m,p}(\Omega)$ 是完备的.

♡

证明

设 $\{u_k\}_{k=1}^\infty$ 是 $W^{m,p}(\Omega)$ 中的基本列, 则从 Sobolev 空间范数的定义中可知 $\{\tilde{\partial}^\alpha u_k\}$ 是 $L^p(\Omega)$ 中的基本列. 任取 $\alpha(|\alpha| \le m)$, 由 $L^p(\Omega)$ 的完备性知存在 $g_\alpha \in L^p(\Omega)$, 使得

$$\|\tilde{\partial}^\alpha u_k - g_\alpha\|_{L^p(\Omega)} \to 0, \quad k \to \infty, |\alpha| \le m$$

故对任意的 $\varphi \in \mathscr{D}(\Omega)$, 当 $k \to \infty$ 时, 一方面

$$\langle \tilde{\partial}^\alpha u_k, \varphi \rangle \to \langle g_\alpha, \varphi \rangle, \quad (-1)^{|\alpha|} \langle u_k, \partial^\alpha \varphi \rangle \to (-1)^{|\alpha|} \langle g_0, \partial^\alpha \varphi \rangle$$

另一方面从广义微商的定义知

$$\langle \tilde{\partial}^\alpha u_k, \varphi \rangle = (-1)^{|\alpha|} \langle u_k, \partial^\alpha \varphi \rangle$$

故 $\langle g_\alpha, \varphi \rangle = (-1)^{|\alpha|} \langle g_0, \partial^\alpha \varphi \rangle$, 也即 $g_\alpha = \tilde{\partial}^\alpha g_0$. 注意 α 是任取的, 故 $g_0 \in W^{m,p}(\Omega)$, 且

$$\|u_k - g_0\|_{m,p} \to 0, \quad k \to \infty.$$

□

定理 5.5.2

$\mathscr{D}(\mathbb{R}^n)$ 在 $W^{m,p}(\mathbb{R}^n)$ 中是稠密的.

♡

证明

设 $u \in W^{m,p}(\Omega)$, 根据定义知 $\tilde{\partial}^\alpha u \in L^p(\mathbb{R}^n)(|\alpha| \le m)$. 对 u 磨光处理: 任取 $\delta > 0$, 令

$$u_\delta(x) = \int_{\mathbb{R}^n} u(y) \cdot j_\delta(x-y) dy$$

其中 j_δ 是磨光算子(5.1). 可以证明 $u_\delta \in C^\infty(\mathbb{R}^n)$, 且

$$\begin{aligned} \partial_x^\alpha u_\delta(x) &= \int_{\mathbb{R}^n} u(y) \cdot \partial_x^\alpha j_\delta(x-y) dy \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} u(y) \cdot \partial_y^\alpha j_\delta(x-y) dy \\ &= \int_{\mathbb{R}^n} \tilde{\partial}^\alpha u(y) \cdot j_\delta(x-y) dy \end{aligned}$$

也即 $\partial^\alpha u_\delta = (\tilde{\partial}^\alpha u)_\delta$. 由 Young 不等式(3.6.2)知, 对任意的 $v \in L^p(\mathbb{R}^n)$ 都有

$$\|v_\delta\|_{L^p} \le \|j_\delta\|_{L^1} \cdot \|v\|_{L^p} = \|v\|_{L^p} \tag{5.12}$$

既然已经证明了 $L^p(\mathbb{R}^n)$ 函数可以被 $C_0^0(\mathbb{R}^n)$ 函数任意逼近, 对 $\tilde{\partial}^\alpha u \in L^p(\mathbb{R}^n)$ 而言有

$$\forall \varepsilon > 0 \exists v_\alpha \in C_0^0(\mathbb{R}^n)(\|v_\alpha - \tilde{\partial}^\alpha u\|_{L^p} < \varepsilon) \tag{5.13}$$

将其中的函数磨光化并应用(5.12)式有

$$\|(v_\alpha)_\delta - (\tilde{\partial}^\alpha u)_\delta\|_{L^p} < \varepsilon \tag{5.14}$$

又由磨光化函数的逼近性质(5.1.1)知

$$\exists \delta_0 = \delta_0(\varepsilon, \alpha) \forall 0 < \delta < \delta_0 (\|(v_\alpha)_\delta - v_\alpha\|_{L^p} < \varepsilon) \tag{5.15}$$

联立(5.13)-(5.15)式知

$$\|(\tilde{\partial}^\alpha u)_\delta - \tilde{\partial}^\alpha u\|_{L^p} < 3\varepsilon$$

此即

$$\|\partial^\alpha u_\delta - \tilde{\partial}^\alpha u\|_{L^p} < \varepsilon, \quad |\alpha| \le m$$

而 $u_\delta \in \mathscr{D}(\mathbb{R}^n)$, 命题即证. □

注 如果用 $W_0^{m,p}(\mathbb{R}^n)$ 表示 $C_0^\infty(\mathbb{R}^n)$ 按范数 $\|\cdot\|_{m,p}$ 完备化后产生的空间, 则依据该定理有

$$W_0^{m,p}(\mathbb{R}^n) = W^{m,p}(\mathbb{R}^n)$$

但对一般的开集 Ω 而言, 未必有 $W_0^{m,p}(\Omega) = W^{m,p}(\Omega)$.

回忆第一章中曾定义过 $H^{m,p}(\Omega)$ 为集合

$$S := \{u \in C^\infty(\Omega) : \|u\|_{m,p} < \infty\} \tag{5.16}$$

在范数 $\|\cdot\|_{m,p}$ 下的完备化空间, 并称之为 Sobolev 空间. 根据 $W^{m,p}(\Omega)$ 的完备性(5.5.1)已经知道 $H^{m,p}(\Omega) \subset W^{m,p}(\Omega)$, 事实上它们是等价的. 要证明这件事需要用到微分流形中的 C^∞ 单位分解定理.

定理 5.5.3 (单位分解定理)

若 $A \subset \mathbb{R}^n$, 且 $\mathscr{O}$ 是 A 的一个开覆盖, 那么必有一族 C^∞ 函数 $\mathscr{F}$, 使得 $\forall \varphi \in \mathscr{F}$ 定义在包含 A 的一个开集上, 且满足:

1. $0 \le \varphi(x) \le 1 (\forall x \in A)$;
2. $\forall x \in A$, 存在含 x 的开集 V, 使得只有有穷多个 $\varphi \in \mathscr{F}$ 在其上非零;
3. $\sum\limits_{\varphi \in \mathscr{F}} \varphi(x) \equiv 1 (\forall x \in A)$;
4. $\forall \varphi \in \mathscr{F}, \exists U \in \mathscr{O}$, 使得 $\mathrm{supp}(\varphi) \subset U$.

称函数族 $\mathscr{F}$ 为 A 关于 $\mathscr{O}$ 的 C^∞ 单位分解.

证明[RAJF]

如若 A 是紧集, 根据定义知 $\mathscr{O}$ 中存在 A 的有限覆盖 $\bigcup\limits_{j=1}^N U_j$. 进一步可以构造紧集 $K_1 \subset U_1, \cdots, K_N \subset U_N$ 使得 $A \subset \bigcup\limits_{j=1}^N K_j$. 对每个 K_j, 取 $\delta_j > 0$ 足够小, 并令 $\phi_j = \chi_{K_j} * j_{\delta_j}$, 其中 j_δ 是磨光算子, 显见 $\phi_j \in C_0^\infty(U_j)$, 且当 $x \in K_j$ 时必有 $\phi_j(x) > 0$. 进一步, 可取到函数 $\phi \in C^\infty(\mathbb{R}^n)$ 满足: $\forall x \in \mathbb{R}^n (\phi(x) > 0)$, 且 $\forall x \in A (\phi(x) = \sum\limits_{j=1}^N \phi_j(x))$. 现在函数族

$$\mathscr{F} = \{f_n : f_j(x) = \frac{\phi_j(x)}{\phi(x)}, 1 \le j \le N\}$$

就是满足要求的函数族.

当 A 是开集时, 考虑分解:

$$A = \bigcup_{j=1}^\infty A_j, \quad A_j = \{x \in A : |x| \le j, \rho(x, \partial A) \ge \frac{1}{j}\}$$

知每个 A_j 都是紧集. 设 $A_0 = A_{-1} = \varnothing$, 则集族

$$\mathscr{O}_j = \{U \cap (A_{j+1} \cap A_{j-2}^c)^\circ : U \in \mathscr{O}\}$$

是 A_j 的开覆盖, 因而由前述结论知存在 A_j 关于 $\mathscr{O}_j$ 的有限个 C^∞ 单位分解 $\mathscr{F}_j$. 进一步考虑和式 $\sigma(x) = \sum\limits_{j=1}^{\infty} \sum\limits_{f \in \mathscr{F}_j} f(x)$, 知其在每个 $x \in A$ 处只有有限多项非零. 从而函数族

$$\mathscr{F} = \{f : \text{当 } x \in A \text{ 时} \phi \in \mathscr{F}_i, f(x) = \frac{\phi(x)}{\sigma(x)}; \text{当 } x \notin A \text{ 时} f(x) = 0\}$$

即为欲求.

最后, 若 A 是任意集合, 设 A 的某个开覆盖为 $\mathscr{O}$, 令 $B = \bigcup\limits_{U \in \mathscr{O}} U$, 显见 B 是开集, 套用前述结论即得命题. □

定理 5.5.4 (Meyers-Serrin)

若 $1 \le p < \infty$, 则

$$H^{m,p}(\Omega) = W^{m,p}(\Omega).$$

证明

只需说明(5.16)式定义的集合 S 在 $W^{m,p}(\Omega)$ 中稠密即可. 取

$$\Omega_k := \{x \in \Omega : \|x\| < k, \rho(x, \partial\Omega) > \frac{1}{k}\}, \quad k = 1, 2, \cdots$$

并记 $\Omega_0 = \Omega_{-1} = \varnothing$, 则

$$\mathscr{O} = \{U_k : U_k = (\Omega_{k+1} \cap \Omega_{k-1}^c)^\circ, k = 1, 2, \cdots\}$$

是 Ω 的一族开覆盖. 记 $\mathscr{F}$ 是单位分解定理(5.5.3)中关于 $\mathscr{O}$ 的一族 C^∞ 函数. 注意 $\overline{U}_k$ 是紧集, 因而根据 C^∞ 单位分解的性质知 $\mathscr{F}$ 中只有有限多个函数 φ 满足 $\operatorname{supp}(\varphi) \subset U_k$. 记 $\psi_k = \sum\limits_{\operatorname{supp}(\varphi) \subset U_k} \varphi$, 则

$$\psi_k \in C_0^\infty(U_k), \quad \sum_{k=1}^{\infty} \psi_k(x) \equiv 1 (\forall x \in \Omega)$$

现设 $u \in W^{m,p}(\Omega)$. 任取 $\varepsilon > 0$, 需找到 S 中的函数 φ, 使得

$$\|u - \varphi\|_{m,p} \le \varepsilon.$$

当 $0 < \delta < \frac{1}{(k+1)(k+2)}$ 时, 知

$$\operatorname{supp}(\psi_k u)_\delta \subset (\Omega_{k+2} \cap \Omega_{k-2})^\circ$$

从磨光化函数的逼近性质知存在 $\delta_k \in (0, \frac{1}{(k+1)(k+2)})$ 满足:

$$\|\psi_k u - (\psi_k u)_{\delta_k}\|_{m,p} < \frac{\varepsilon}{2^k}$$

令 $\varphi = \sum\limits_{k=1}^{\infty} (\psi_k u)_{\delta_k}$, 因为对任意的 $x \in \Omega$, 在 x 的某个邻域内这个和式只有有限项不为零, 故 $\varphi \in C^\infty(\Omega)$. 又对任意的 $k \in \mathbb{N}$, 在 Ω_k 上有

$$u(x) = \sum_{j=1}^{k+2} \psi_j(x) u(x), \quad \varphi(x) = \sum_{j=1}^{k+2} (\psi_j u)_{\delta_j}(x)$$

故

$$\|u - \varphi\|_{W^{m,p}(\Omega_k)} \le \sum_{j=1}^{k+2} \|(\psi_j u)_{\delta_j} - \psi_j u\|_{m,p} < \varepsilon$$

再令 $k \to \infty$ 即得 $\|u - \varphi\|_{m,p} \le \varepsilon$. 显见 $\varphi \in S$, 此即欲求. □

定义 5.5.2 (可扩张)

$\mathbb{R}^n$ 中的开区域 Ω 称为可扩张的, 如果 $\forall m \in \mathbb{N}, \forall p \in [1,\infty], \exists T : W^{m,p}(\Omega) \to W^{m,p}(\mathbb{R}^n)$ 是连续线性算子, 并满足

$$Tu|_\Omega = u, \quad \forall u \in W^{m,p}(\Omega).$$

例 5.26 $\mathbb{R}^n_+ = \{(x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : x_n > 0\}$ 是可扩张的. 扩张算子构造为: 任取 $m \in \mathbb{N}, u \in C^\infty(\mathbb{R}^n_+)$, 定义

$$E_m u(x) = \begin{cases} u(x), & x_n > 0 \\ \sum_{j=1}^{m+1} \lambda_j u(x_1, \cdots, x_{n-1}, -jx_n), & x_n \le 0 \end{cases}$$

其中系数 $\lambda_1, \lambda_2, \cdots, \lambda_{m+1}$ 是线性方程组

$$\sum_{j=1}^{m+1} (-j)^i \lambda_j = 1, \quad i = 0, 1, \cdots, m$$

的唯一解. 可以验证若 $u \in C^m(\overline{\mathbb{R}^n_+})$, 则 $E_m u \in C^m(\mathbb{R}^n)$, 且

$$\|E_m u\|_{W^{m,p}(\mathbb{R}^n)} \le M_{m,p} \|u\|_{W^{m,p}(\mathbb{R}^n_+)}$$

其中 $M_{m,p}$ 是一个常数. 知 E_m 可以连续地扩张到 $W^{m,p}(\mathbb{R}^n)$ 上去, 因而 $\mathbb{R}^n_+$ 是可扩张的. □

例 5.27 若 Ω 是有界开区域, 具有一致 C^m 光滑的边界, 则 Ω 是可扩张的.

定理 5.5.5 (Sobolev 嵌入定理)

若 $\Omega \subset \mathbb{R}^n$ 是一个可扩张的区域, $m > \frac{n}{2}$, 则 $W^{m,2}(\Omega)$ 可以连续地嵌入 $C(\overline{\Omega})$.

证明

首先证明 $\mathbb{R}^n$ 的情况. 已经知道

$$\|u\|'_m = \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^m |\hat{u}(\xi)|^2 d\xi\right)^{\frac{1}{2}}$$

是 $W^{m,2}(\mathbb{R}^n)$ 的一个等价范数. 又因为 $L^1(\mathbb{R}^n)$ 函数的 Fourier(逆) 变换是连续函数, 且根据 Fourier 反演公式知

$$\|u\|_{C(\mathbb{R}^n)} \le \int_{\mathbb{R}^n} |\hat{u}(\xi)| d\xi$$

当 $m > \frac{n}{2}$ 时知

$$\int_{\mathbb{R}^n} |\hat{u}(\xi)| d\xi \le \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^m |\hat{u}(\xi)|^2 d\xi\right)^{\frac{1}{2}} \cdot \left(\int_{\mathbb{R}^n} \frac{d\xi}{(1+|\xi|^2)^m}\right)^{\frac{1}{2}} \le c_{n,m} \|u\|'_m$$

故 $i : u \mapsto u$ 是 $W^{m,2}(\mathbb{R}^n) \to C(\mathbb{R}^n)$ 的一个连续嵌入.

现设 $u \in W^{m,2}(\Omega)$, 因为 Ω 是可扩张的, 根据定义知存在延拓算子 T, 有:

$$\|u\|_{C(\overline{\Omega})} \le \|Tu\|_{C(\mathbb{R}^n)} \le c_{n,m} \|Tu\|'_m \le c\|u\|_{W^{m,2}(\Omega)}$$

即 $i : u \mapsto u$ 是 $W^{m,2}(\Omega) \to C(\Omega)$ 的连续嵌入. □

注 更一般的嵌入定理不需要限制 $p = 2$:

$$W^{m,p}(\Omega) \hookrightarrow L^q(\Omega), \quad \frac{1}{q} = \frac{1}{p} - \frac{m}{n}, m \le \frac{n}{p}$$
$$W^{m,p}(\Omega) \hookrightarrow C(\overline{\Omega}), \quad m > \frac{n}{p}$$

其中 $\hookrightarrow$ 表示连续嵌入.

定理 5.5.6 (Rellich)

设 $\Omega \subset \mathbb{R}^n$ 是一个有界可扩张区域, 则 $W^{1,2}(\Omega)$ 中的单位球在 $L^2(\Omega)$ 中是列紧的.

证明

根据列紧的定义知需要证明: 当 $\{u_m\}_{m=1}^{\infty} \subset W^{1,2}(\Omega), \|u_m\|_{W^{1,2}} \le 1$ 时, 可从中抽出子列 $\{u_{m'}\}$ 使得其在 $L^2(\Omega)$ 中收敛. 为此记 $U_m = Tu_m$, 其中 T 是扩张算子. 不妨设 U_m 在一个足够大的 $\mathbb{R}^n$ 紧的集合 K 之外取 0, 现在要证

$$\|u_{p'} - u_{m'}\|_{L^2(\Omega)} \to 0, \quad m', p' \to \infty$$

注意由扩张的性质与 Plancherel 定理知

$$\|u_{p'} - u_{m'}\|_{L^2(\Omega)}^2 \le C\|U_{p'} - U_{m'}\|_{L^2(\mathbb{R}^n)}^2 = C\int_{\mathbb{R}^n} |\widehat{U}_{p'}(\xi) - \widehat{U}_{m'}(\xi)|^2 d\xi$$

而

$$\widehat{U}_m(\xi) = \int_K U_m(x)e^{-2\pi i x\cdot\xi}dx = \langle e^{-2\pi i x\cdot\xi}\chi_K, U_m\rangle$$

既然 $\|u_m\|_{L^2(\mathbb{R}^n)} \le \|u_m\|_{W^{1,2}} \le 1$, 而 T 是连续线性的, 故至少存在 $C > 0$ 使得 $\|U_m\|_{L^2(\mathbb{R}^n)} \le C$. 因为 $L^2(\mathbb{R}^n)$ 是自反空间, 而 $\{U_m\}$ 如今是 $L^2(\mathbb{R}^n)$ 的某个有界集中的序列, 故由 Eberlein-Smulian 定理(3.6.12)知存在子列 $\{U_{m'}\}$ 是弱收敛的. 因为 $e^{-2\pi i x\cdot\xi}\chi_K \in L^2(\mathbb{R}^n)$, 根据弱收敛的定义知对任意的 $\xi \in \mathbb{R}^n$, $\widehat{U}_{m'}(\xi)$ 是收敛的. 又注意到

$$|\widehat{U}_m(\xi)| \le (\int_K dx)^{\frac{1}{2}} \cdot (\int_K |U_m(x)|^2 dx)^{\frac{1}{2}} \le (m(K))^{\frac{1}{2}} \cdot \|U_m\|_{L^2} \le const$$

其中 $m(K)$ 是 K 的 Lebesgue 测度. 现在证明 $\{\widehat{U}_{m'}\}$ 是 L^2 收敛的. 任取 $r > 0$, 设

$$\int_{\mathbb{R}^n} |\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi = \int_{|\xi|\le r} |\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi + \int_{|\xi|\ge r} |\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi =: I + II$$

知

$$\begin{aligned} II &\le r^{-2}\int_{|\xi|\ge r}(1+r^2)|\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi \le r^{-2}\int_{|\xi|\ge r}(1+|\xi|^2)|\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi \\ &\le 2r^{-2}\int_{|\xi|\ge r}(1+|\xi|^2)(|\widehat{U}_{m'}(\xi)|^2 + |\widehat{U}_{p'}(\xi)|^2)d\xi \le 2r^{-2}(\|U_{m'}\|_{W^{1,2}}^2 + \|U_{p'}\|_{W^{1,2}}^2) \le Mr^{-2} \end{aligned}$$

其中 M 是一个仅依赖于扩张算子 T 的常数. 现在对任意的 $\varepsilon > 0$, 取 r 足够大满足 $Mr^{-2} < \varepsilon$, 固定 r, 既然 $\{\widehat{U}_{m'}(\xi)\}$ 收敛, 知当 m', p' 足够大时有

$$I = \int_{|\xi|\le r} |\widehat{U}_{m'}(\xi) - \widehat{U}_{p'}(\xi)|^2 d\xi < \varepsilon$$

得到 $p', m' \to \infty$ 时:

$$\|u_{p'} - u_{m'}\|_{L^2(\Omega)}^2 \le C\|\widehat{U}_{p'} - \widehat{U}_{m;}\|_{L^2(\mathbb{R}^n)} \to 0$$

根据 $L^2(\Omega)$ 的完备性即得命题. □

推论 5.5.1

若 $\Omega \subset \mathbb{R}^n$ 是任意的有界开集, 则 $W_0^{1,2}(\Omega)$ 的单位球在 $L^2(\Omega)$ 中是列紧的.

回忆第一章中曾定义过 Hilbert 空间 H_0^m, 其上的内积为

$$(u, v)_m = \sum_{|\alpha|=m}\int_\Omega \partial^\alpha u(x)\cdot\overline{\partial^\alpha v(x)}dx, \quad \forall u, v \in C_0^m(\Omega)$$

根据 Meyers-Serrin 定理(5.5.4), $W_0^{1,2}(\Omega)$ 就是 $H_0^1(\Omega)$. 注意

$$\mathscr{D}(\Omega) \hookrightarrow H_0^m(\Omega) \hookrightarrow L^2(\Omega)$$

可见它们的共轭空间有下述联系:

$$L^2(\Omega) \hookrightarrow H_0^m(\Omega)^* \hookrightarrow \mathscr{D}'(\Omega)$$

这说明 Sobolev 空间的共轭空间 $H_0^m(\Omega)^*$ 是由比 $L^2(\Omega)$ 函数更多地广义函数组成的. 下面记 $H_0^m(\Omega)^*$ 为 $H^{-m}(\Omega)$, 并将其中的元素用积分表示出来.

定理 5.5.7

要使得 $f \in H^{-m}(\Omega)$, 必须且仅须 $\exists g_\alpha \in L^2(\Omega)(|\alpha| \leq m)$, 使得

$$\langle f, \varphi \rangle = \sum_{|\alpha| \leq m} \int_\Omega g_\alpha(x) \cdot \partial^\alpha \varphi(x) dx, \quad \forall \varphi \in H_0^m(\Omega).$$

证明

当存在 g_α 满足条件时, 显见 $f \in H^{-m}(\Omega)$. 现设 $f \in H^{-m}(\Omega)$, 因为 $H_0^m(\Omega)$ 是 Hilbert 空间, 由 Riesz 表示定理(3.3.1)知存在 $h \in H_0^m(\Omega)$ 使得

$$\langle f, \varphi \rangle = (\varphi, h)_m = \sum_{|\alpha| \leq m} \int_\Omega \overline{\tilde{\partial}^\alpha h(x)} \cdot \tilde{\partial}^\alpha \varphi(x) ddx = \sum_{|\alpha| \leq m} \int_\Omega g_\alpha(x) \tilde{\partial}^\alpha \varphi(x) dx, \quad \forall \varphi \in H_0^m(\Omega)$$

其中 $g_\alpha(x) := \tilde{\partial}^\alpha h(x) \in L^2(\Omega)$. □

推论 5.5.2

每个 $f \in H^{-m}(\Omega)$ 都是 $L^2(\Omega)$ 函数的广义微商的有限和, 即

$$f = (-1)^{|\alpha|} \sum_{|\alpha| \leq m} \tilde{\partial}^\alpha g_\alpha, \quad g_\alpha \in L^2(\Omega) \tag{5.17}$$

证明

由广义微商的定义即得命题. □

注

1. 如果先给定 $g_\alpha \in L^2(\Omega)$, 则根据(5.17)式可以定义 $f \in \mathscr{D}'(\Omega)$, 注意这个 f 在 $H^m(\Omega)$ 上的延拓可能不是唯一的, 但它在 $H_0^m(\Omega)$ 上的连续延拓确实唯一, 这是因为 $C_0^\infty(\Omega)$ 在 $H_0^m(\Omega)$ 中稠密, 且对任意的 $|\alpha| \leq m$ 有
$$|\langle \tilde{\partial}^\alpha g_\alpha, \varphi \rangle| \leq \|g_\alpha\|_{L^2(\Omega)} \cdot \|\varphi\|_{H_0^m(\Omega)}, \quad \forall \varphi \in C_0^\infty(\Omega)$$
故(5.17)式给出了 $H^{-m}(\Omega)$ 中广义函数的特征.
2. 由 Sobolev 嵌入定理(5.5.5)知, 当 $m > \frac{n}{2}$ 时, $H_0^m(\Omega)$ 可以连续嵌入 $C(\overline{\Omega})$, 故当 $m > \frac{n}{2}$ 时, $C(\overline{\Omega})$ 上的连续线性泛函都属于 $H^{-m}(\Omega)$. 特别 δ 函数也属于 $H^{-m}(\Omega)$.

第二部分

偏微分方程

第六章 引言与预备知识

6.1 常微分方程的一些重要定理

本节选自 [**YR**], 旨在回顾 ODE 中一些经典的定理, 以便日后用于对比. 设初值问题

$$\begin{cases} \frac{d\boldsymbol{y}}{dx} = \boldsymbol{f}(x, \boldsymbol{y}) \\ \boldsymbol{y}(x_0) = \boldsymbol{y}_0 \end{cases}$$

为了简便, 将其写成标量形式:

$$\begin{cases} \frac{dy}{dx} = f(x, y) \\ y(x_0) = y_0 \end{cases} \tag{6.1}$$

其中 f 连续.

6.1.1 解的存在唯一性定理

定义 6.1.1 ((整体)Lipschitz 条件)

设函数 $f(x, y)$ 在区域 $D \subset \mathbb{R}^{n+1}$ 内满足

$$\|f(x, y_1) - f(x, y_2)\| \leq L\|y_1 - y_2\|, \quad \forall (x, y_1), (x, y_2) \in D$$

其中常数 $L > 0$, 则称函数 $f(x, y)$ 在区域 D 内对 y 满足整体 Lipschitz 条件, 简称为 Lipschitz 条件, 记为 $f \in Lip$.

♣

定义 6.1.2 (局部 Lipschitz 条件)

设 $f(x, y)$ 在区域 D 上连续. 若对 D 中的每一点 $P_0(x_0, y_0)$, 都有点 P_0 的小矩形邻域 $R_{P_0} \subset D$, 使得 f 在区域 R_{P_0} 上对 y 满足 Lipschitz 条件, 则称 f 在区域 D 上满足局部 Lipschitz 条件.

♣

定理 6.1.1 (Picard 定理)

设初值问题

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

其中 $f(x, y)$ 在矩形区域

$$R =: \{(x, y) : |x - x_0| \leq a, \|y - y_0\| \leq b\}$$

内连续, 且 f 对 y 满足 Lipschitz 条件, 则该初值问题在区间 $I = [x_0 - h, x_0 + h]$ 上有且只有一个解 $y = \phi(x)$, 其中 $h = \min\{a, \frac{b}{M}\}, M = \max\limits_{(x,y)\in R} \|f(x, y)\|$, 此时也称方程组(6.1)过点 (x_0, y_0) 有且只有一个解.

♡

Picard 定理证明的关键思路是把方程组(6.1)转化成积分方程组

$$y = y_0 + \int_{x_0}^{x} f(s, y)ds$$

并构造 Picard 序列

$$\begin{cases} \phi_k(x) = y_0 + \int_{x_0}^{x} f(s, \phi_{k-1}(s))ds, \quad x \in I, k = 1, 2, \cdots \\ \phi_0(x) = y_0 \end{cases}$$

利用 Lipchitz 条件得到结果. □

在没有 Lipschitz 条件时, 方程的解可以存在, 但就不一定唯一了.

定理 6.1.2 (Peano)

设初值问题

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

其中 $f(x, y)$ 在矩形区域

$$R =: \{(x, y) : |x - x_0| \leq a, \|y - y_0\| \leq b\}$$

内连续, 则该初值问题在区间 $|x - x_0| \leq h$ 上至少有一个解 $y = \phi(x)$, 其中 $h = \min\{a, \frac{b}{M}\}, M = \max\limits_{(x,y)\in R} \|f(x, y)\|$, 也即方程 $\frac{dy}{dx} = f(x, y)$ 过点 (x_0, y_0) 至少有一个解 $\phi(x)$.

Lipschitz 条件毕竟太强, 下面介绍弱于 Lipschitz 条件且能保证解的唯一性成立的 Osgood 条件.

定义 6.1.3 (Osgood 条件)

设函数 $f(x, y)$ 在区域 D 内连续, 满足不等式

$$\|f(x, y_1) - f(x, y_2)\| \leq F(\|y_1 - y_2\|)$$

其中 $F(r) > 0$ 是 $r > 0$ 的连续函数, 且[a]

$$\int_0^{r_1} \frac{dr}{F(r)} = \infty$$

其中 $r_1 > 0$ 为常数, 则称 $f(x, y)$ 在 D 内对 y 满足 Osgood 条件.

[a] 这个式子事实上在说 $F(r)$ 在 0 附近的无穷小阶数高于 1, 譬如 $F(r) = r^2$.

定理 6.1.3 (Osgood)

设 $f(x, y)$ 在区域 D 内对 y 满足 Osgood 条件, 则微分方程 $\frac{dy}{dx} = f(x, y)$ 在 D 内经过每一点的解都是唯一的.

6.1.2 解的延伸定理

定义 6.1.4 (延拓, 可延伸的, 饱和解, 最大存在区间)

若 $y = \phi(x), x \in (\alpha, \beta)$ 和 $y = \tilde{\phi}(x), x \in (\tilde{\alpha}, \tilde{\beta})$ 是方程(6.1)的解, 且

(i) $(\alpha, \beta) \subset (\tilde{\alpha}, \tilde{\beta})$, 但 $(\alpha, \beta) \neq (\tilde{\alpha}, \tilde{\beta})$;

(ii) $\tilde{\phi}(x) = \phi(x), x \in (\alpha, \beta)$

则称 $\tilde{\phi}(x)$ 是 $\phi(x)$ 在 $(\tilde{\alpha}, \tilde{\beta})$ 上的一个延拓, 此时称 $\phi(x)$ 是可延伸的. 若不存在满足上述条件的 $\tilde{\phi}(x)$, 则称 $\phi(x), x \in (\alpha, \beta)$ 是方程 $\frac{dy}{dx} = f(x, y)$ 的一个饱和解, 此时得到的存在区间称为解的最大存在区间.

定理 6.1.4

设 $f(x, y)$ 在 D 上连续且有界, 若 $\phi(x)$ 是方程 $\frac{dy}{dx} = f(x, y)$ 的一个解, 它在 (a, b) 上有定义, 则 $\phi(a+0), \phi(b-0)$ 必存在有限, 且当 $(b, \phi(b-0)) \in D((a, \phi(a+0)) \in D)$ 时, 解可以从 b 向右 (从 a 向左) 延伸.

推论 6.1.1

设 $f(x, y)$ 在 D 上连续, 若 $\phi(x)$ 是方程 $\frac{dy}{dx} = f(x, y)$ 的一个解, 它在 (a, b) 上有定义, 且当 $x \in (a, b)$ 时, $\|f(x, \phi(x))\| \leq M < \infty$, 则 $\phi(a+0), \phi(b-0)$ 必存在且有限, 并且当 $(b, \phi(b-0)) \in D((a, \phi(a+0)) \in D)$ 时, 解可以从 b 向右 (从 a 向左) 延伸.

针对方程过给定点的积分曲线有:

定理 6.1.5

设 P_0 为区域 D 内任一点, 并设 $\Gamma: y=\phi(x)$ 是方程 $\frac{dy}{dx}=f(x,y)$ 经过 P_0 点的任一条积分曲线, 则积分曲线 Γ 可以延伸成饱和解 $y=\psi(x), x\in I$.

定理 6.1.6 (解的延伸定理)

设 P_0 为区域 D 内任一点, 并设 $\Gamma: y=\phi(x)$ 为方程 $\frac{dy}{dx}=f(x,y)$ 经过 P_0 点的任一条积分曲线, 则积分曲线 Γ 可以延伸成饱和解 $\Gamma_1: y=\psi(x)$. 如果 I 是解 $\psi(x)$ 的最大存在区间, 则 I 必为开区间, 记为 $I=(a,b)$, 且当 $x\to a$ 或 $x\to b$ 时, $(x,\psi(x))$ 趋于 D 的边界. 也就是说, 对任意有界闭区域 D_1, 若 $P_0\in D_1\subset D$, 则积分曲线 Γ_1 总会延伸到 D_1 之外. 以 x 增大的一方延伸为例, 解的延伸只会出现下面两种情形之一:

1. 解 $y=\psi(x)$ 可以延伸到区间 $[x_0,+\infty)$, 也即 $b=+\infty$;
2. 解 $y=\psi(x)$ 只能延伸到区间 $[x_0,b)$, 其中 b 是有限数, 则当 $x\to b$ 时, 要么 $y=\psi(x)$ 无界, 要么点 $(x,\psi(x))$ 趋于区域 D 的边界.

将解的延伸定理与 Picard 定理结合即得:

推论 6.1.2

设函数 $f(x,y)$ 在区域 D 内连续, 且对 y 满足局部 Lipschitz 条件, 则方程 $\frac{dy}{dx}=f(x,y)$ 经过 D 内任一点 P_0 都存在唯一的积分曲线 Γ, 且 Γ 在 D 内延伸到边界.

引理 6.1.1 (Gronwall 不等式)

设 K 为非负常数, $f(t)$ 与 $g(t)$ 为区间 $\alpha\le t\le\beta$ 上的连续非负函数, 且

$$f(t)\le K+\int_\alpha^t f(s)g(s)ds,\quad \alpha\le t\le\beta$$

则

$$f(t)\le K\exp\Big(\int_\alpha^t g(s)ds\Big),\quad \alpha\le t\le\beta.$$

定理 6.1.7

设微分方程

$$\frac{dy}{dx}=f(x,y)$$

其中函数 $f(x,y)$ 在条形区域

$$S:=\{(x,y):\alpha<x<\beta, y\in\mathbb{R}^n\}$$

内连续, 且满足不等式

$$\|f(x,y)\|\le A(x)\|y\|+B(x)$$

其中 $A(x)\ge 0$ 和 $B(x)\ge 0$ 在区间 $\alpha<x<\beta$ 上是连续的, 则该微分方程的每一个解都以区间 $\alpha<x<\beta$ 为最大存在区间.

6.1.3 解对初值与参数的依赖性与可微性

现在在微分方程中加入参数 λ, 考虑初值问题:

$$\begin{cases}\frac{dy}{dx}=f(x,y,\lambda)\\ y(x_0)=y_0\end{cases}\tag{6.2}$$

Picard 存在唯一性定理说明该初值问题的解应存在唯一. 如果初值组 (x_0, y_0, λ) 不同, 一般来说 $\phi(x)$ 也会不同, 故 ϕ 应该是关于 (x, x_0, y_0, λ) 的函数.

如果 ϕ 关于 x_0, y_0, λ 是连续的, 那么当新的初值与 (x_0, y_0, λ) 差距很小时, 得到的解 ϕ 与原来的解的差距也应该很小, 这便暗含了解的某种稳定性.

为了简化问题, 做变换

$$\begin{cases} t = x - x_0 \\ u = y - y_0 \end{cases}$$

其中 t 是新的自变量, 而 $u = u(t)$ 是未知函数, 这便把初值问题(6.2)变为

$$\left\{ \frac{du}{dt} = f(t + x_0, u + y_0, \lambda) y(0) = 0 \right. \tag{6.3}$$

此时原先的初值 (x_0, y_0) 就也变成了参数形式, 而初值条件本质上是没有变的, 故为了简化问题, 只讨论初值问题

$$\begin{cases} \frac{dy}{dx} = f(x, y, \mu) \\ y(0) = 0 \end{cases} \tag{6.4}$$

的解 $y = \phi(x, \mu)$ 对参数 μ 的依赖性, 其中 μ 是 m 维的参数向量.

定理 6.1.8 (解对参数的连续性)

设 n 维向量值函数 $f(x, y, \mu)$ 在区域

$$G := \{(x, y, \mu) : |x| \le a, \|y\| \le b, \|\mu - \mu_0\| \le c\}$$

上连续, 且对 y 满足 Lipschitz 条件

$$\|f(x, y_1, \mu) - f(x, y_2, \mu)\| \le L\|y_1 - y_2\|$$

其中常数 $L \ge 0$. 令正数 M 为 $\|f(x, y, \mu)\|$ 在区域 G 的一个上界, 且令

$$h = \min\{a, \frac{b}{M}\}$$

则初值问题(6.4)的解 $y = \phi(x, \mu)$ 在区域

$$D := \{(x, \mu) : |x| \le h, \|\mu - \mu_0\| \le c\}$$

上是连续的.

定理 6.1.9 (解对初值的连续性)

设 $f : G \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ 连续, 且对 y 满足局部 Lipschitz 条件, 其中 G 为 $\mathbb{R} \times \mathbb{R}^n$ 中的一个开区域. 假设 $y = \xi(x)$ 是微分方程

$$\frac{dy}{dx} = f(x, y)$$

的一个解, 存在区间为 J. 在区间 J 内任取一个有界闭区间 $a \le x \le b$, 则存在常数 $\delta > 0$, 使得对任何初值 (x_0, y_0) 都有

$$a \le x_0 \le b, \quad \|y_0 - \xi(x_0)\| \le \delta$$

且初值问题

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

的解 $y = \phi(x, x_0, y_0)$ 也至少在区间 $a \le x \le b$ 上存在, 且它在闭区域

$$D_\delta := \{(x, x_0, y_0) : a \le x \le b, a \le x_0 \le b, \|y_0 - \xi(x_0)\| \le \delta\}$$

上是连续的.

下面研究初值问题

$$\begin{cases} \frac{dy}{dx} = f(x, y, \lambda) y(0) = 0 \end{cases} \tag{6.5}$$

的解 $y = \phi(x, \lambda)$ 对参数 λ 的可微性.

定理 6.1.10

设 n 维向量值函数 $f(x, y, \lambda)$ 在区域

$$G := \{(x, y, \lambda) : |x| \le a, \|y\| \le b, \|\lambda - \lambda_0\| \le c\}$$

上连续, 且对 y 和 λ 有连续的偏导, 则初值问题(6.5)的解 $y = \phi(x, \lambda)$ 在区域

$$D := \{(x, \lambda) : |x| \le h, \|\lambda - \lambda_0\| \le c\}$$

上是连续可微的, 其中 $h = \min\{a, \frac{b}{M}\}$, 正数 M 为 $\|f(x, y, \lambda)\|$ 在区域 G 上的一个上界.

推论 6.1.3

设 n 维向量值函数 $f(x, y)$ 在区域

$$R := \{(x, y) : |x - x_0| \le a, \|y - y_0\| \le b\}$$

上连续, 且对 y 有连续的偏导数 $f'_y(x, y)$, 则初值问题

$$\begin{cases} \frac{dy}{dx} = f(x, y) \\ y(x_0) = \eta \end{cases}$$

的解 $y = \phi(x, \eta)$ 在区域

$$D := \{(x, y) : |x - x_0| \le \frac{h}{2}, \|\eta - y_0\| \le \frac{b}{2}\}$$

上是连续可微的.

6.1.4 解的稳定性

考虑一般的微分方程

$$\frac{dx}{dt} = f(t, x) \tag{6.6}$$

其中 $f : \mathbb{R} \times G \to \mathbb{R}^n$ 连续, $G \subset \mathbb{R}^n$ 是区域, 且对 x 满足局部 Lipschitz 条件. 进而方程(6.6)满足解的存在唯一性定理与延伸定理, 从而方程(6.6)满足初值条件 $x(t_0) = x_0$ 的解存在唯一, 记之为 $x = x(t, t_0, x_0)$.

设 $x = \phi(t)$ 是方程(6.6)的一个解, 且在 $t_0 \le t < \infty$ 上有定义.

定义 6.1.5 ((Lyapunov 意义) 稳定)

如果对任意给定的 $\varepsilon > 0$, 都存在 $\delta = \delta(\varepsilon) > 0$, 使得只要 $\|x_0 - \phi(t_0)\| < \delta$, 就有

$$\|x(t, t_0, x_0) - \phi(t)\| < \varepsilon, \quad \forall t \ge t_0$$

就称解 $x = \phi(t)$ 是 (Lyapunov 意义下) 稳定的.

定义 6.1.6 ((Lyapunov 意义) 渐近稳定)

如果 $x = \phi(t)$ 是稳定的, 且是吸引的, 即存在 $\delta_1 (0 < \delta_1 \le \delta)$, 使得只要 $\|x_0 - \phi(t_0)\| < \delta_1$, 就有

$$\lim_{t \to +\infty} (x(t, t_0, x_0) - \phi(t)) = 0$$

就称解 $x = \phi(t)$ 是 (Lyapunov 意义下) 渐近稳定的.

定义 6.1.7 ((Lyapunov 意义) 全局渐近稳定)

如果 $x = \phi(t)$ 是稳定的, 且对任意的 $x_0 \in \mathbb{R}^n$ 都有

$$\lim_{t\to+\infty}(x(t,t_0,x_0) - \phi(t)) = 0$$

则称解 $x = \phi(t)$ 是 (Lyapunov 意义下) 全局渐近稳定的.

定义 6.1.8 (不稳定)

如果解 $x = \phi(t)$ 不是稳定的, 则称它是不稳定的, 也即存在 $\varepsilon > 0$, 使得对任意的 $\delta > 0$, 总存在 x_0 使得 $\|x_0 - \phi(t_0)\| < \delta$, 且存在 $t_1 > t_0$ 使得

$$\|x(t_1,t_0,x_0) - \phi(t_1)\| \geq \varepsilon.$$

针对常系数齐次线性方程组有下述结论:

定理 6.1.11

考虑常系数齐次线性微分方程组

$$\frac{dx}{dt} = Ax \tag{6.7}$$

其中 A 为常数矩阵, 则有下述结论成立:

(i) 方程组(6.7)的零解是渐近稳定的, 当且仅当 A 的全部特征根都具有负实部;

(ii) 方程组(6.7)的零解是稳定的, 当且仅当 A 的全部特征根的实部都是非正的, 且实部为 0 的特征根所对应的 Jordan 块都是一阶的;

(iii) 方程组(6.7)的零解是不稳定的, 当且仅当 A 的特征根中至少有一个实部为正, 或者至少有一个特征根实部为零, 且它所对应的 Jordan 块是高于一阶的.

实际应用中, 为了纳入对扰动的考虑, 下面研究方程组

$$\frac{dx}{dt} = Ax + R(t,x) \tag{6.8}$$

其中 A 是常数矩阵, 而 $R(t,x)$ 对 t 和 x 在区域

$$G := \{(t,x) : t \geq t_0, \|x\| \leq M\}$$

上连续, 对 x 满足局部 Lipschitz 条件, 且满足 $R(t,0) \equiv 0 (t \geq t_0)$ 与

$$\lim_{\|x\|\to 0}\frac{\|R(t,x)\|}{\|x\|} = 0 \quad (\text{对 } t \geq t_0 \text{ 一致成立}) \tag{6.9}$$

则方程组(6.8)称为(6.7)的扰动系统.

定理 6.1.12

设条件(6.9)成立, 则

(i) 当 A 的全部特征根的实部都是负数时, 方程组(6.8)的零解是渐近稳定的;

(ii) 当 A 的特征根中至少有一个实部为正时, 方程组(6.8)的零解是不稳定的.

6.2 偏微分方程

6.2.1 一些基本概念

总的来说, 一个偏微分方程 (PDE) 就是包含了两个或更多变元以及它们偏导数的方程. 下面给出一些符号上的定义, 约定 $k \geq 1$ 是整数, $U \subset \mathbb{R}^n$ 是开子集.

定义 6.2.1 (多重指标, 高阶导数)

形如 $\alpha=(\alpha_1,\cdots,\alpha_n),\alpha_i\in\mathbb{N}$ 的向量 α 称作具有阶

$$|\alpha|=\alpha_1+\cdots+\alpha_n$$

的多重指标. 当 $u:U\to\mathbb{R}$, 约定

$$D^\alpha u(x):=\frac{\partial^{|\alpha|}u(x)}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}=\partial_{x_1}^{\alpha_1}\cdots\partial_{x_n}^{\alpha_n}u$$

当 k 是非负整数, u 的全体 k 阶导数即

$$D^k u(x):=\{D^\alpha u(x):|\alpha|=k\}.$$

♣

如果假定不同的偏导之间是有顺序的, 就可以把 $D^k u(x)$ 视作 $\mathbb{R}^{n^k}$ 中的点, 其范数定义为:

$$|D^k u|=(\sum_{|\alpha|=k}|D^\alpha u|^2)^{\frac{1}{2}}$$

有时对 D 添加自变量的下标以表示对这个自变量求偏导, 如若 $u=u(x,y)(x\in\mathbb{R}^n,y\in\mathscr{R}^m)$, 则:

$$D_x u=(u_{x_1},\cdots,u_{x_n}),$$

$$D_y u=(u_{y_1},\cdots,u_{y_m}).$$

定义 6.2.2 (k 阶偏微分方程)

具有形式

$$F(D^k u(x),D^{k-1}u(x),\cdots,Du(x),u(x),x)=0\quad(x\in U)\tag{6.10}$$

的表达式称作 k 阶偏微分方程, 其中

$$F:\mathbb{R}^{n^k}\times\mathbb{R}^{n^{k-1}}\times\cdots\times\mathbb{R}^n\times\mathbb{R}\times U\to\mathbb{R}$$

已知, 而 $u:U\to\mathbb{R}$ 未知.

♣

今后用 PDE 简称偏微分方程.

定义 6.2.3 (线性, 齐次, 半线性, 拟线性, 完全非线性)

1. PDE(6.10)称为线性的, 如果其具有形式

$$\sum_{|\alpha|\le k}a_\alpha(x)D^\alpha u=f(x)$$

其中 $a_\alpha(|\alpha|\le k)$ 与 f 是给定的. 线性 PDE 称作是齐次的, 如果 $f\equiv 0$.

2. PDE(6.10)称为半线性的, 如果其具有形式

$$\sum_{|\alpha|=k}a_\alpha(x)D^\alpha u+a_0(D^{k-1}u,\cdots,Du,u,x)=0.$$

3. PDE(6.10)称为拟线性的, 如果其具有形式

$$\sum_{|\alpha|=k}a_\alpha(D^{k-1}u,\cdots,Du,u,x)D^\alpha u+a_0(D^{k-1}u,\cdots,Du,u,x)=0.$$

4. PDE(6.10)称为完全非线性的, 如果它并不线性依赖于最高阶导数.

♣

不严谨地来说, PDE 系统指的就是一堆含有几个未知函数的 PDE.

定义 6.2.4 (k 阶 PDE 系统)

具有形式

$$\boldsymbol{F}(D^k\boldsymbol{u}(x),D^{k-1}\boldsymbol{u}(x),\cdots,D\boldsymbol{u}(x),\boldsymbol{u}(x),x)=\boldsymbol{0}\quad(x\in U)$$

的表达式称作 k 阶 PDE 系统, 其中

$$\boldsymbol{F}:\mathbb{R}^{mn^k}\times\mathbb{R}^{mn^{k-1}}\times\cdots\times\mathbb{R}^{mn}\times\mathbb{R}^m\times U\to\mathbb{R}^m$$

给定, 而

$$\boldsymbol{u}:U\to\mathbb{R}^m,\quad \boldsymbol{u}=(u^1,\cdots,u^m)$$

未知.

上述定义中的方程数与自变量数相等, 有时会出现不相等的情况. PDE 系统的线性性与定义(6.2.3)是类似的.

课堂笔记 (2023.8.31)

- (关于边值问题的提法) 课堂上介绍了两类边值问题, 下面以 Laplace 方程为例来说明:
  (i) Dirichlet 问题: 给定了未知函数在边界上具体分布, 希望求解方程. 例如:
  $$\begin{cases}\Delta u=0, & \Omega\subset\mathbb{R}^n\\ u|_{\partial\Omega}=g(x)\end{cases}$$
  (ii) Neumann 问题: 给定了未知函数在边界上法向导数的分布, 希望求解方程. 例如:
  $$\begin{cases}\Delta u=0, & \Omega\subset\mathbb{R}^n\\ \frac{\partial u}{\partial \boldsymbol{n}}|_{\partial\Omega}=g(x)\end{cases}$$
  在上述两种边值问题中, 边值条件 g 可能会有各种各样的连续性, 诸如 $C^\infty, C^{1+\alpha}, C, \mathrm{Lip}$ 等等.
- (关于问题的适定性与解的稳定性[CZC]) 对于事先选定的某函数空间, 如果定解问题的解在该函数空间存在唯一且稳定, 就称该定解问题是适定的, 否则称为不适定的. 定解问题解的稳定性是指: 设线性赋范空间 H 的函数为 $\|\cdot\|_H$, u_1,u_2 是分别对应于定解数据 φ_1,φ_2 的同一定解问题的解, 则若对任意的 $\varepsilon>0$, 都存在 $\delta>0$, 使得只要 $\|\varphi_1-\varphi_2\|_H<\delta$, 就有 $\|u_1-u_2\|_H<\varepsilon$, 就称定解问题的解是稳定的.

6.2.2 预备知识

之后的 PDE 推导中会需要诸多重要的预备知识, 一并列在此处.

定理 6.2.1 (Gauss-Green-Ostrogradsky 公式 (高-奥公式))

1. 设 $u\in C^1(\overline{U})$, 则
$$\int_U u_{x_i}dx=\int_{\partial U}u\nu^i dS$$
其中 ν^i 是 ∂U 上某点处的单位外法向量 $\boldsymbol{\nu}$ 朝 x_i 代表方向的投影, $\nu^i dS$ 表示 dS 在 x_i 方向上的分解.
2. 对任意向量场 $\boldsymbol{u}\in C^1(\overline{U};\mathbb{R}^n)$ 有
$$\int_U \operatorname{div}\boldsymbol{u}\,dx=\int_{\partial U}\boldsymbol{u}\cdot\boldsymbol{\nu}dS$$
其中 $\boldsymbol{\nu}dS$ 表示 dS 朝 $x_1,\cdots,x_n$ 方向作投影后形成的向量.

为了更方便理解, 这里再给出经典的三重积分下的高-奥公式:
$$\iiint_D\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dxdydz=\iint_{\partial D}(Pdy\wedge dz+Qdz\wedge dx+Rdx\wedge dy)$$
其中 $\boldsymbol{u}=(P,Q,R)$, $\boldsymbol{\nu}dS=(dy\wedge dz,dz\wedge dx,dx\wedge dy)$.

将 u 换成 uv 可得下述定理.

定理 6.2.2 (分部积分法)

设 $u, v \in C^1(\overline{U})$, 则

$$\int_U u_{x_i} v dx = -\int_U u v_{x_i} dx + \int_{\partial U} u v v^i dS, \quad i = 1, 2, \cdots, n$$

定理 6.2.3 (Green 公式)

设 $u, v \in C^2(\overline{U})$, 则

(i) $\int_U \Delta u dx = \int_{\partial U} \frac{\partial u}{\partial v} dS$. 其中 $\frac{\partial u}{\partial v}$ 是 u 沿 $\boldsymbol{v}$ 方向的方向导数. 特别注意 $\frac{\partial u}{\partial v} = Du \cdot \boldsymbol{v} = \sum_{i=1}^{n} u_{x_i} v^i$.

(ii) $\int_U Dv \cdot Du dx = -\int_U u \Delta v dx + \int_{\partial U} \frac{\partial v}{\partial v} u dS$.

(iii) $\int_U u\Delta v - v\Delta u dS = \int_{\partial U} u\frac{\partial v}{\partial v} - v\frac{\partial u}{\partial v} dS$.

定理 6.2.4 (极坐标换元)

(i) 若 $f : \mathbb{R}^n \to \mathbb{R}^n$ 是连续可积的, 则对 $x_0 \in \mathbb{R}^n$ 有

$$\int_{\mathbb{R}^n} f dx = \int_0^{\infty} (\int_{\partial B(x_0, r)} f dS) dr$$

(ii) 特别对 $r > 0$ 有

$$\frac{d}{dr}(\int_{B(x_0,r)} f dx) = \int_{\partial B(x_0,r)} f dS.$$

6.2.3 广义函数, 弱解与一些统领性的想法

本节大部分叙述参照于 [**QMY**]. 相较于其余更专业与系统性的内容, 本节更倾向于是自己对于广义函数与物理思想在各个方程中的运用, 以及弱解的引入所牵扯到的想法的一些整理.

恰如前面谈及广义函数时所提到的 [**LS**] 中的引言, Heaviside 函数与 Dirac 函数的出现最初是为了满足物理学的需求. Dirac 函数的定义十分古怪:

$$\delta(x) = \begin{cases} 0, & x \neq 0 \\ \infty, & x = 0 \end{cases} \qquad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

同时其满足对任意连续函数 $\varphi(x)$ 有

$$\int_{-\infty}^{\infty} \delta(x)\varphi(x) dx = \varphi(0)$$

在物理上, $\delta(x)$ 有鲜明的意义: x 轴的原点上有一个质量为 1 的质点, 则 $\delta(x)$ 就是这种质量分布的密度函数.

现在再略微跑一下题谈谈变分[1]. 对于泛函 Φ 与空间中的曲线 $\gamma = \{(t, x) : x = x(t), t_0 \le t \le t_1\}$ 而言, 如果 $\Phi(\gamma + h) - \Phi(\gamma) = F + R$, 其中 F 线性依赖于 h, 而 $R(h, \gamma) = O(h^2)$, 就称泛函 Φ 是 (强[2]) 可微的. 这个增量的线性部分 $F(h)$ 称为 Φ 的微分. 可以证明如果一个泛函 Φ 可微, 则其微分必唯一. 泛函的微分又叫做泛函的变分, h 对应叫做曲线的变分.

在经典力学中, 如果把 x 视作一个质点的位置矢量, t 视作时间, 那么曲线 $\gamma = \{(t, x) : x = x(t), t_0 \le t \le t_1\}$ 就能表示这个质点的运动情况. 就像中学物理所讲的那样, 位移的导数等于速度, 亦即速度为 $\dot{x} = \frac{dx}{dt}$. 从一般的角度讲, 如果设 $L = L(a, b, c)$ 是一个三变量可微函数, 在不考虑任何动机的情况下, 可以 (随意地) 定义一个泛函 Φ:

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t) dt$$

如果为 L 赋予一些具体意义, 比如令 $L(a, b, c) = \sqrt{1 + b^2}$, 那么 $L(x(t), \dot{x}(t), t) = \sqrt{1 + \dot{x}^2(t)}$, 从而泛函 $\Phi(\gamma)$ 输出的就是 γ 的曲线长度.

[1] 这部分也可以在第三部分找到相关说明

[2] 强微分的意义在前面的补充: 线性空间微分学概要中有提及.

上面构造的泛函 $\Phi(\gamma)=\int_{t_0}^{t_1}L(x,\dot{x},t)dt$ 关于 γ 是可微的, 这是因为根据微分的定义:

$$\begin{aligned}\Phi(\gamma+h)-\Phi(\gamma)&=\int_{t_0}^{t_1}[L(x+h,\dot{x}+\dot{h},t)-L(x,\dot{x},t)]dt\\&=\int_{t_0}^{t_1}[L(x+h,\dot{x}+\dot{h},t)-L(x,\dot{x}+\dot{h},t)+L(x,\dot{x}+\dot{h},t)-L(x,\dot{x},t)]dt\\&=\int_{t_0}^{t_1}[\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}]dt+O(h^2)=F(h)+R\end{aligned}$$

从而泛函 $\Phi(\gamma)$ 的微分正是

$$F(h)=\int_{t_0}^{t_1}(\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h})dt$$

进一步处理, 对 $\frac{\partial L}{\partial \dot{x}}\dot{h}$ 分部积分有:

$$\int_{t_0}^{t_1}\frac{\partial L}{\partial \dot{x}}\dot{h}dt=-\int_{t_0}^{t_1}h\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})dt+h\frac{\partial L}{\partial \dot{x}}|_{t_0}^{t_1}$$

代回原式可得

$$F(h)=\int_{t_0}^{t_1}[\frac{\partial L}{\partial x}-\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}]hdt+(\frac{\partial L}{\partial \dot{x}}h)|_{t_0}^{t_1}$$

可以看出, 如果 $h(t_1)=h(t_0)=0$, 则 $F(h)=0$ 对一切 h 成立, 当且仅当沿着曲线 $x(t)$ 时有

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}-\frac{\partial L}{\partial x}=0$$

这个方程就称为泛函 $\Phi=\int_{t_0}^{t_1}L(x,\dot{x},t)dt$ 的 Euler-Lagrange 方程. 当对一切的 h 都有 $F(h,\gamma)=0$, 亦即泛函 Φ 的微分总为 0 时, 就称对应的曲线 γ 是泛函 Φ 的驻定曲线.

下面利用上面这些铺垫引出 Hamilton 最小作用量原理. 回忆 Newton 力学, 给定两个位置 M_0, M, 力 $\boldsymbol{F}$ 从 M_0 到 M 所产生的势能为

$$U(M)=-\int_{M_0}^{M}\boldsymbol{F}\cdot d\boldsymbol{S}$$

此即

$$\boldsymbol{F}=-\frac{\partial U}{\partial \boldsymbol{x}}$$

现在对于 n 质点 $m_i(i=1,\cdots,n)$ 的力学系统, 记第 i 个质点的位移为 $\boldsymbol{r}_i$, 则对每个质点应用 Newton 第二定律:

$$\frac{d}{dt}(m_i\dot{\boldsymbol{r}}_i)+\frac{\partial U}{\partial \boldsymbol{r}_i}=0$$

下面就说明这个力学系统的运动所对应的曲线就是泛函

$$\Phi(t)=\int_{t_0}^{t_1}Ldt$$

的驻定曲线, 其中 $L=T-U$ 是动能与位能之差. 回忆中学时学过的动能定义:

$$T_i=\frac{1}{2}m_i\dot{\boldsymbol{r}}_i^2$$

现在系统动能即各个部分动能之和:

$$T=\sum_i T_i=\sum_i m_i\dot{\boldsymbol{r}}_i^2$$

要验证一个曲线是驻定曲线, 就是要验证其满足对应的 Euler-Lagrange 方程, 知:

$$\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i}=\frac{\partial T}{\partial \dot{\boldsymbol{r}}_i}=m_i\dot{\boldsymbol{r}}_i,\quad \frac{\partial L}{\partial \boldsymbol{r}_i}=-\frac{\partial U}{\partial \boldsymbol{r}_i}$$

而 $-\frac{\partial U}{\partial \boldsymbol{r}_i}$ 依照前面定义的势能, 正是作用在第 i 个质点上的外力, 从而 Euler-Lagrange 方程此时就正是 Newton 第二定律, 断言证毕.

现在回顾这个断言: 力学系统的运动是某个泛函的驻定曲线, 这便是所谓 Hamilton 最小作用量原理. 其中作用量就是 $\int_{t_0}^{t_1}L(\boldsymbol{q},\dot{\boldsymbol{q}},t)dt$, 最小是因为这个驻定曲线在很多情况下能使得作用量达到最小值.

回到具体的问题, 下面来讨论两端固定的弦 $[0,l]$ 的运动, 其运动由函数 $u(x,t)$ 表示, 其中

$$u(0,t)=u(l,t)=0$$

其作用量定义为

$$I[u]=\frac{1}{2}\int_{t_0}^{t_1}(\int_0^l[\rho u_t^2-Tu_x^2]dx)dt$$

其中 ρ 是弦质量密度, T 是张力系数. 如果 $u(x,t)$ 确实是一个运动, 那它就应该满足 Hamilton 最小作用量原理, 亦即让 I 达到最小值. 不妨设 u 所在的函数类为 $\mathscr{A}$, 现在要讨论最小, 就是说

$$I[u]=\min_{w\in\mathscr{A}}I[w]$$

如果 $C_0^\infty\subset\mathscr{A}$, 就可以取用 C_0^∞ 中的函数 φ 作为某种" 方向", 研究 $I[u+\lambda\varphi](\lambda\in\mathbb{R})$ 与 $I[u]$ 的大小关系 (这就像在谈论 $\mathbb{R}^n\to\mathbb{R}$ 的函数的最小值时, 将 x 变为 $x+\lambda h$ 一样). 知:

$$I[u+\lambda\varphi]=\frac{1}{2}\int_{t_0}^{t_1}(\int_0^l[\rho(u_t+\lambda\varphi_t)^2-T(u_x+\lambda\varphi_x)^2]dx)dt$$

在 $\lambda=0$ 时 $I[u+\lambda\varphi]$ 作为 λ 的函数取得最小值, 这便要求 $\frac{d}{d\lambda}I[u+\lambda\varphi]=0$, 得到:

$$\frac{d}{d\lambda}I[u+\lambda\varphi]|_{\lambda=0}=\int_{t_0}^{t_1}(\int_0^l[\rho u_t\varphi_t-Tu_x\varphi_x]dx)dt=0$$

分部积分并令 $\varphi(x,t_0)=\varphi(x,t_1)=0,\varphi(0,t)=\varphi(l,t)=0$(因为 φ 只在 C_0^∞ 中取, 就算有了这个限制之后, 它能取的函数范围依旧非常广) 可得

$$\frac{d}{d\lambda}I[u+\lambda\varphi]|_{\lambda=0}=\int_{t_0}^{t_1}(\int_0^l[Tu_{xx}-\rho u_{tt}]\varphi dx)dt=0$$

根据 φ 的任意性可以得到 $Tu_{xx}-\rho u_{tt}=0$, 这正是弦振动方程!

注意上面的推导默认了 u 至少是 C^2 的. 如果 u 仅仅是 C^1 的, 那么就不对 φ 分部积分, 而改为对 u 分部积分, 得到:

$$\frac{d}{d\lambda}I[u+\lambda\varphi]=\int_{t_0}^{t_1}(\int_0^l u(T\varphi_{xx}-\rho\varphi_{tt})dx)dt=0$$

上式甚至都没有出现 u 的一阶微分, 所有对可微性的要求全部转移到光滑函数 φ 上了. 现在就可以畅想: u 只需要是可积函数, 且对任意 C_0^∞ 函数 φ 上式都成立. 此时称 u 是弦振动方程的弱解.

第七章 四种重要的线性 PDE

7.1 输运方程

常系数输运方程是最简单的 PDE 之一, 其形如

$$u_t + b \cdot Du = 0 \quad 在\mathbb{R}^n \times (0,\infty)中 \tag{7.1}$$

其中 $b = (b_1, \cdots, b_n) \in \mathbb{R}^n$ 是一个给定的向量, $u = u(x,t) : \mathbb{R}^n \times [0,\infty) \to \mathbb{R}$ 未知. 这里 $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ 表示空间中的某个点, 而 $t \geq 0$ 表示某个时间. 用 $Du = D_x u = (u_{x_1}, \cdots, u_{x_n})$ 表示 u 关于空间变量 x 的梯度.

现在要问怎样的 u 是方程(7.1)的解? 要求方程的解, 可以先设该解足够光滑. 注意 PDE(7.1)本身表明 u 的某个方向导数恰为 0. 这是因为任意固定 $(x,t) \in \mathbb{R}^n \times (0,\infty)$, 记

$$z(s) := u(x+sb, t+s) \quad (s \in \mathbb{R})$$

则

$$\dot{z}(s) = \frac{d}{ds}(z) = Du(x+sb, t+s) \cdot b + u_t(x+sb, t+s) = 0$$

这说明任意固定 (x,t), $z(s)$ 都是常值函数, 因而 u 在连接 (x,t) 与 $(b,1)$ 的直线上总是常值. 这说明只要求出 u 在这些直线中某一点的值, 就可得出它在全空间 $\mathbb{R}^n \times (0,\infty)$ 的取值.

7.1.1 初值问题

为了明确问题, 现在考虑初值问题:

$$\begin{cases} u_t + b \cdot Du = 0 & 在\mathbb{R}^n \times (0,\infty)中 \\ u = g & 在\mathbb{R}^n \times \{t=0\}上 \end{cases} \tag{7.2}$$

其中 $b \in \mathbb{R}^n, g : \mathbb{R}^n \to \mathbb{R}$ 都是已知的, 希望求解 u. 给定 (x,t), 知连接 (x,t) 与 $(b,1)$ 的直线有参数方程 $(x+sb, t+s)(s \in \mathbb{R})$. 这条直线与平面 $\Gamma := \mathbb{R}^n \times \{t=0\}$ 相交于 $(x-tb, 0)$, 此时 $s = -t$. 根据前述讨论, u 在这条直线上是常值函数, 而 $u(x-tb,0) = g(x-tb)$, 这说明

$$u(x,t) = g(x-tb) \quad (x \in \mathbb{R}^n, t \geq 0) \tag{7.3}$$

故若初值问题(7.2)有足够正则的解, 那这个解必定满足(7.3)式. 反之也可以验证: 如果 $g \in C^1$, 则(7.3)式确实是初值问题(7.2)的解.

但若 $g \notin C^1$, 则初值问题(7.2)显然没有 C^1 解. 但这种情况下(7.3)式也给出了一个强的, 且事实上唯一可能的解. 所以尽管 g 可能不是 C^1 的, 还是不严谨地称 $u(x,t) = g(x-tb)(x \in \mathbb{R}^n, t \geq 0)$ 是初值问题(7.2)的一个弱解. 就算 g, u 不连续, 这个断言也是成立的. 这个现象提醒我们: 就算是不连续的函数, 也可能是 PDE 的解.

7.1.2 非齐次问题

下面考虑输运方程的非齐次问题.

$$\begin{cases} u_t + b \cdot Du = f & 在\mathbb{R}^n \times (0,+\infty)中 \\ u = g & 在\mathbb{R}^n \times \{t=0\}上 \end{cases} \tag{7.4}$$

任意固定 $(x,t) \in \mathbb{R}^n \times (0,\infty)$, 记 $z(s) := u(x+sb, t+s)(s \in \mathbb{R})$, 有

$$\dot{z}(s) = Du(x+sb, t+s) \cdot b + u_t(x+sb, t+s) = f(x+sb, t+s)$$

从而

$$u(x,t) - g(x-tb) = z(0) - z(-t) = \int_{-t}^{0} \dot{z}(s)ds = \int_{-t}^{0} f(x+sb, t+s)ds = \int_{0}^{t} f(x+(s-t)b, s)ds$$

这说明

$$u(x,t) = g(x-tb) + \int_{0}^{t} f(x+(s-t)b, s)ds \quad (x \in \mathbb{R}^n, t \geq 0) \tag{7.5}$$

正是初值问题(7.4)的解.

之后还会把这个公式用到一维波方程中.

注 注意在推导(7.3)与(7.5)时, 事实上做的事是把 PDE 转化成了常微分方程, 即 ODE. 这种方法是特征值方法的一种特例, 后面也会加以讨论.

课堂笔记 (2023.8.31)

- (关于齐次输运方程的求解) 如果把 (t,x) 看成一个在 $(0,\infty)\times\mathbb{R}^n$ 上的整体变量, 则(7.1)式可写成

$$(1,b)\begin{pmatrix} u_t \\ Du \end{pmatrix} = 0 \tag{7.6}$$

(7.6)左式表示函数 $u(t,x)$ 在 $(1,b)$ 方向的方向导数, 因而(7.6)式说明 $u(t,x)$ 在 $(1,b)$ 方向上变化率为零, 此即前述 $z(s)$ 的构造动机.

再考虑初值问题(7.2), 如果把(7.2)写成:

$$\begin{cases} u_t = -b\cdot Du \\ u(0,x) = g(x) \end{cases}$$

并把式子中的 $-b\cdot D$ 形式地记为常数 a, 上述 PDE 就变成了一个关于变量 t 和参数 x 的 ODE 问题, 容易求得:

$$u(t,\cdot) = e^{at}g(\cdot)$$

再将 $a=-b\cdot D$ 代回有

$$u(t,x) = e^{-bt\cdot D}g(x)$$

从前面讨论的结果知本来应该是 $u(t,x)=g(x-tb)$, 这便在形式上得到

$$e^{-bt\cdot D}g(x) = g(x-tb)$$

也就是说, 算子 $e^{y\cdot D}$ 把函数 $f(\cdot)$ 变为 $f(\cdot+y)$. 这件事也可以从 Taylor 展开的角度上理解: 如果对算子 $e^{y\cdot D}$ 作形式的展开, 有

$$e^{y\cdot D} = \sum_{n=0}^{\infty}\frac{y^n D^n}{n!}$$

从而对函数 $f(x)$ 有

$$e^{y\cdot D}f(x) = \sum_{n=0}^{\infty}\frac{y^n D^n}{n!}f(x) = \sum_{n=0}^{\infty}\frac{y^n D^n f(x)}{n!}$$

如果将 y 视作增量, 这就恰是函数 $f(x+y)$ 在 $f(x)$ 处的 Taylor 展开.

- (关于叠加原理) 以输运方程为例, 对线性方程而言, 如果 v,w 都是齐次输运方程的解:

$$v_t + b\cdot Dv = 0$$

$$w_t + b\cdot Dw = 0$$

那么对任意的常数 λ,μ, 函数 $\lambda v+\mu w$ 也是同一个齐次输运方程的解:

$$(\lambda v+\mu w) + b\cdot D(\lambda v+\mu w) = 0$$

- (关于非齐次输运方程的求解) 对于非齐次问题(7.4), 仿照前述思路将其写成

$$(1, b)\begin{pmatrix} u_t \\ Du \end{pmatrix} = f(t, x) \tag{7.7}$$

(7.7)式说明函数 $u(t, x)$ 在 $(1, b)$ 方向上的变化率正是 $f(t, x)$, 根据这个思路才构造出了前文所提的

$$z(s) = u(t + s, x + bs)$$

而 z 的导数就是 u 在 $(1, b)$ 方向上的导数, 从而由(7.7)式得

$$\dot{z}(s) = f(t + s, x + bs)$$

两边对 s 从 0 到 s' 积分得

$$z(s') - z(0) = \int_0^{s'} f(t + s, x + bs)ds \Rightarrow u(t + s', x + bs') - u(x, t) = \int_0^{s'} f(t + s, x + bs)ds$$

最后希望利用 $t = 0$ 的初值条件, 令 $s' = -t$ 即得前文的(7.5)式.

与前面把 $-b \cdot D$ 形式看成 a 相同, 这里也介绍一次这个过程, 此时得到的方程即

$$\begin{cases} u_t - au = f \\ u(0, x) = g(x) \end{cases}$$

解上述关于变量 t 与参数 x 的 ODE, 利用常数变易法即得

$$u(t, x) = e^{at}g(x) + \int_0^t e^{a(t-s)}f(s, x)ds$$

同样代回 $a = -b \cdot D$ 即得解的另一种表达式:

$$u(t, x) = e^{-bt \cdot D}g(x) + \int_0^t e^{-b(t-s) \cdot D}f(s, x)ds$$

(7.5)式又称为 Duhamel 公式.

- (几何解释) 从图(7.1)可以看出, 输运方程的几何意义在于随着时间 t 的推移, 在 x 轴上以速度 b 平移函数 g. 特别对于图(7.1), 可以明显看出波峰连成了直线, 这些线对应的就是前面说的方向 $(1, b)$, 而在后面讨论特征线方法的时候, 这些直线就正是输运方程对应的特征线.

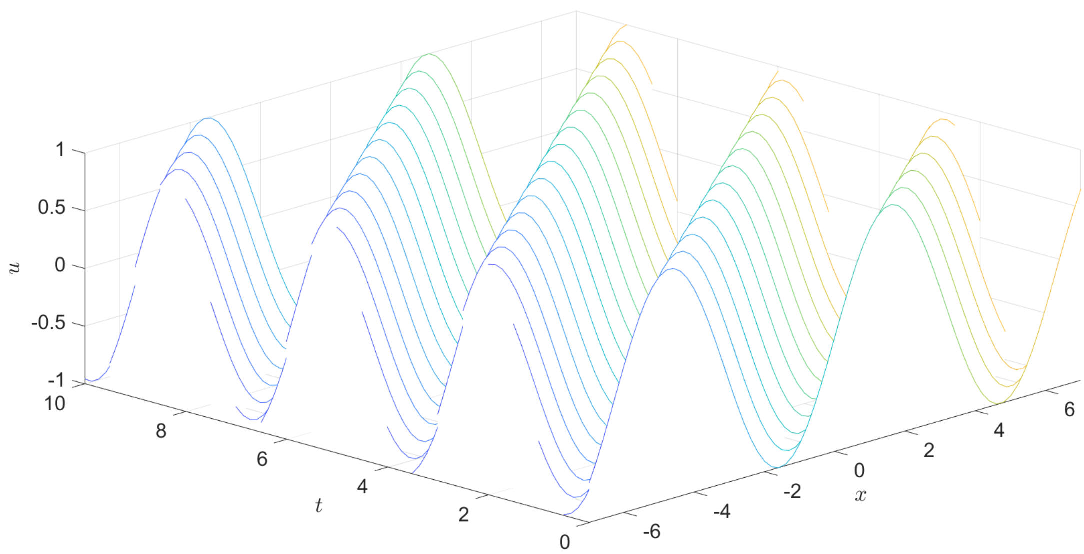


图 7.1: $u = g(x - 2t), g(x) = \sin x$ 的瀑布图

课堂笔记 (2023.9.4)

- (解与参函数的正则性) 对于齐次输运方程初值问题(7.2), 根据初值条件即知, 当 $g \in C^1$ 时, 方程就有 C^1 解. 而对于非齐次输运方程初值问题(7.4), 为了让方程依旧有 C^1 解, 针对 Duhamel 公式:
$$u(x,t) = g(x-tb) + \int_0^t f(x+(s-t)b, s)ds$$
两边对 t 求导可得
$$u_t(x,t) = -bg_t(x-tb) - bf(x,t) \cdot \int_0^t Df(x+(s-t)b, s)ds$$
因为 u_t 是连续的, 故 $f(x,t)$ 应是连续的. 而在 Duhamel 公式两边对 x 求导有:
$$Du(x,t) = Dg(x-tb) + \int_0^t Df(x+(s-t)b, s)ds$$
现在 $Du(x,t)$ 也是连续的, 故 $\int_0^t Df(x+(s-t)b, s)ds$ 至少应该对 x 连续 (对 t 的连续性已经由 f 本身的连续性保证了), 故 f 对 x 是 C^1 连续的.
- (关于算子 e^{-tbD} 在 Fourier 变换下的理解) 算子表达式
$$e^{-tbD}g(x) = g(x-tb)$$
也可以用 Fourier 变换的角度去理解. 为此注意 Fourier 变换[a]有性质:
$$\widehat{D^\alpha u} = (iy)^\alpha \widehat{u}$$
按 Fourier 变换的定义写开并约去系数即
$$\int_{\mathbb{R}^n} e^{-ix\cdot y} D^\alpha u(x)dx = \int_{\mathbb{R}^n} (iy)^\alpha e^{-ix\cdot y} u(x)dx$$
其中 α 是多重指标, $u \in L^2(\mathbb{R}^n)$. 上式说明至少在 Fourier 变换的意义下, 左乘算子 $(iy)^\alpha(\cdot)$ 与微分算子 $D^\alpha(\cdot)$ 是等价的. 现对 $e^{-tbD}g(x)$ 关于 x 作 Fourier 变换有:
$$\widehat{(e^{-tbD}g(x))} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} e^{-tbD} g(x)dx$$
将 D 换成左乘算子 iy 有
$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} e^{-tbD} g(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} e^{-tb\cdot(iy)} g(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i(x+tb)\cdot y} g(x)dx$$
换元有
$$\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i(x+tb)\cdot y} g(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\xi\cdot y} g(\xi - tb)d\xi = \widehat{g(\xi - tb)}$$
符号上将 ξ 改回 x, 即得
$$\widehat{(e^{-tbD}g(x))} = \widehat{g(x-tb)}$$
故在 Fourier 变换的意义下, $e^{-tbD}g(x) = g(x-tb)$.
- (非齐次输运方程的几何解释) 从图(7.2),(7.3)中可见, 非齐次的情况下输运方程的解沿特征线方向的值有所变化, 而某点相较于 $t=0$ 时变化的值正是 $f(x)$ 从 $t=0$ 到此时沿特征线的积分 (可参考图(7.3), 波峰连成的直线此时受到了 f 积分的影响.).

[a] 具体定义参见(9.3.1)

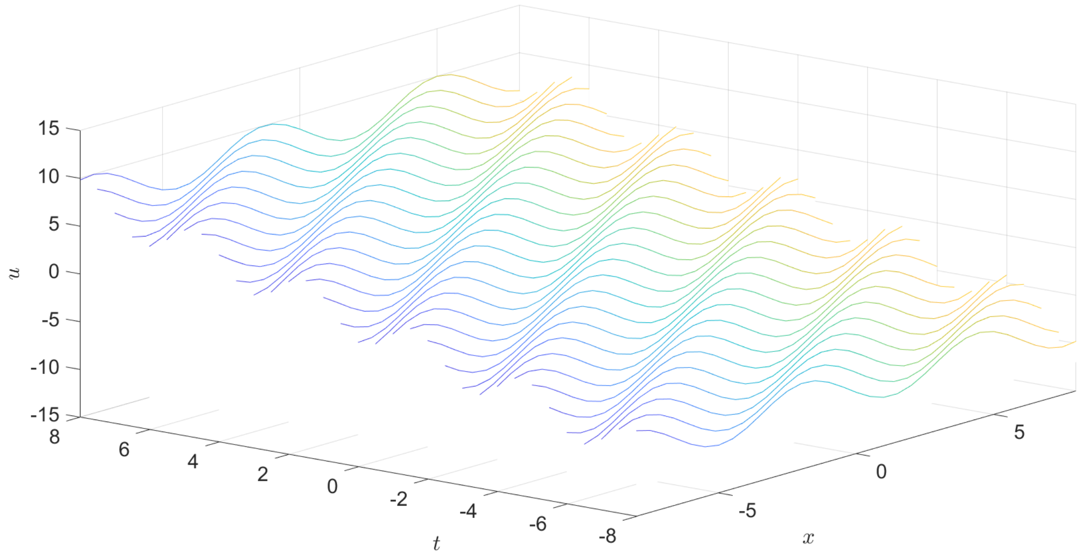


图 7.2: $u = g(x-2t) + \int_0^t f(x+2(s-t), s)ds, g = 2\sin x, f = 1$ 的瀑布图

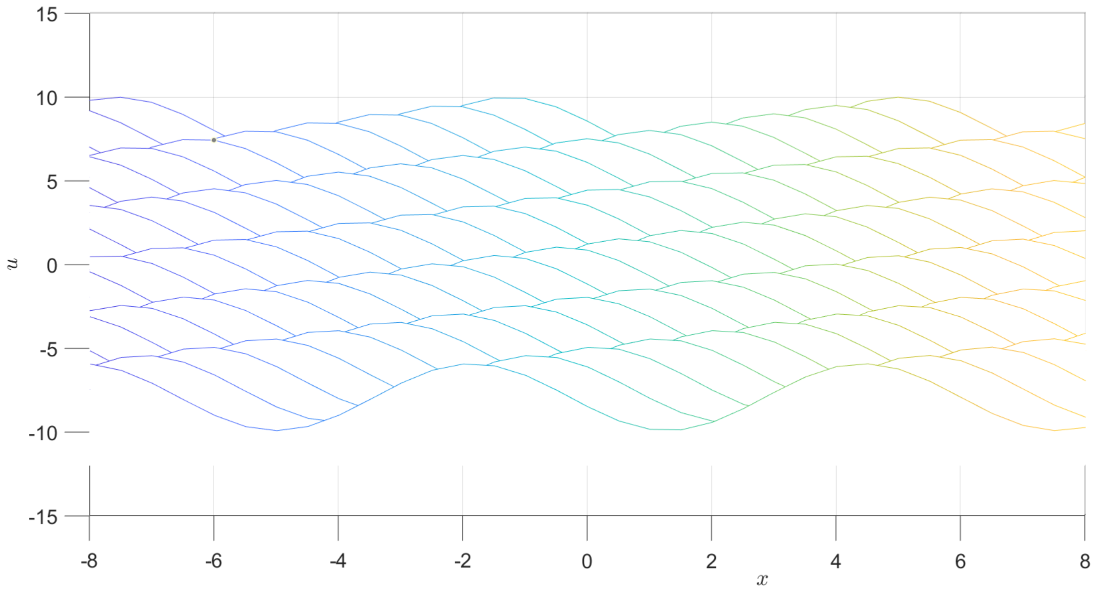


图 7.3: $u = g(x-2t) + \int_0^t f(x+2(s-t), s)ds, g = 2\sin x, f = 1$ 在 x 方向上的瀑布图投影

7.2 Laplace 方程

Laplace 方程是最重要的 PDE 之一, 其形如:

$$\Delta u = 0 \tag{7.8}$$

同样还有 Poisson 方程:

$$-\Delta u = f \tag{7.9}$$

在(7.8)式与(7.9)式中, $x \in U$ 是开集, 且 $u = u(x) : \overline{U} \to \mathbb{R}$ 未知. (7.9)式中 $f : U \to \mathbb{R}$ 是给定的. 回忆 $\Delta u = \sum_{i=1}^{n} u_{x_i x_i}$.

定义 7.2.1 (调和函数)

满足(7.8)式的 C^2 函数 u 称作调和函数. ♣

Laplace 方程出现在很多物理概念中. 一个经典解释是 u 表示平衡状态下某种量的密度 (比如化学物质的浓度). 从而如果 V 是 U 的任意光滑子区域, 则 u 穿过 ∂V 的净通量就是 0:

$$\int_{\partial V} \boldsymbol{F}\cdot\boldsymbol{\nu}dS = 0$$

其中 $\boldsymbol{F}$ 是通量密度, $\boldsymbol{\nu}$ 是单位外法向量. 根据高-奥公式(6.2.1)有

$$\int_V \operatorname{div}\boldsymbol{F}dx = \int_{\partial V} \boldsymbol{F}\cdot\boldsymbol{\nu}dS = 0$$

根据 V 的任意性知在 U 中有

$$\operatorname{div}\boldsymbol{F} = 0 \tag{7.10}$$

在很多例子中, 把通量 $\boldsymbol{F}$ 视作与梯度 Du 成比例但反向[1]是有物理意义的, 也即可设

$$\boldsymbol{F} = -aDu \quad (a>0) \tag{7.11}$$

代入(7.10)式有

$$\operatorname{div}(Du) = \Delta u = 0$$

这正是 Laplace 方程.

7.2.1 基本解

先从特殊解入手, 再根据 PDE 的线性性导出更复杂的解是研究 PDE 的一个常用方法. 在寻找这个特殊解时, 一般都对函数类作某种对称性的限制. 既然 Laplace 方程在旋转下是不变的, 就可以首先研究放射解, 也即关于 $r=|x|$ 的函数.

现在在 $U=\mathbb{R}^n$ 上寻找 Laplace 方程(7.8)的具有形式

$$u(x)=v(r),\quad r=|x|=(x_1^2+\cdots+x_n^2)^{\frac{1}{2}}$$

的解. 显见 v 的选取需要满足 $\Delta u=0$. 注意到对 $i=1,2,\cdots,n$ 有

$$\frac{\partial r}{\partial x_i}=\frac{x_i}{r},\quad x\neq 0$$

故

$$u_{x_i}=v'(r)\cdot\frac{x_i}{r}\Rightarrow u_{x_ix_i}=v''(r)\cdot\frac{x_i^2}{r^2}+v'(r)\cdot(\frac{1}{r}-\frac{x_i^2}{r^3})$$

进而

$$\Delta u=\sum_{i=1}^n u_{x_ix_i}=v''(r)+\frac{n-1}{r}\cdot v'(r)$$

故

$$\Delta u=0\Leftrightarrow v''+\frac{n-1}{r}v'=0$$

若 $v'\neq 0$, 根据 ODE 的理论可解得

$$v(r)=\begin{cases} b\ln r+c, & n=2\\ \frac{b}{r^{n-2}}+c, & n\geq 3\end{cases}$$

其中 b,c 都是常数.

---

[1]这是因为流总是从高密度到低密度, 而梯度是从低密度指向高密度的.

定义 7.2.2 (Laplace 方程的基本解)

函数

$$\Phi(x) := \begin{cases} -\frac{1}{2\pi}\ln|x|, & n=2 \\ \frac{1}{n(n-2)\alpha(n)}\frac{1}{|x|^{n-2}}, & n\geq 3 \end{cases} \quad x\in\mathbb{R}^n\backslash\{0\} \tag{7.12}$$

是 Laplace 方程的基本解.

有时为了强调这个基本解是放射的, 会记 $\Phi(x)=\Phi(|x|)$. 经过计算可得估计

$$|D\Phi(x)|\leq\frac{C}{|x|^{n-1}}, |D^2\Phi(x)|\leq\frac{C}{|x|^n}, \quad x\neq 0 \tag{7.13}$$

其中 $C>0$ 是某个固定的常数.

接下来讨论 Poisson 方程. 注意上面介绍的基本解 $\Phi(x)$ 关于 $x\neq 0$ 是调和函数. 如果把原点平移到一个新的点 y, 注意 PDE(7.8)是不变的, 故相应的基本解 $\Phi(x-y)$ 关于 $x\neq y$ 依旧是调和函数. 下面取 $f:\mathbb{R}^n\to\mathbb{R}$, 注意对每个固定的 $y\in\mathbb{R}^n$ 而言, 函数 $\Phi(x-y)f(y)$ 关于 $x\neq y$ 依旧是调和函数, 从而形如 $\sum\limits_y\Phi(x-y)f(y)$(即取不同的 y 求和) 的函数也是关于 $x\neq y$ 的调和函数.

出于前面讨论的式子形式, 可以猜想卷积

$$u(x)=\int_{\mathbb{R}^n}\Phi(x-y)f(y)dy=\begin{cases} -\frac{1}{2\pi}\int_{\mathbb{R}^2}\ln(|x-y|)f(y)dy, & n=2 \\ \frac{1}{n(n-2)\alpha(n)}\int_{\mathbb{R}^n}\frac{f(y)}{|x-y|^{n-2}}dy, & n\geq 3 \end{cases} \tag{7.14}$$

是 Laplace 方程(7.8)的解. 但这个推导并不正确, 因为估计式(7.13)已经暗示了 $D^2(x-y)$ 在奇点 $y=x$ 附近可能不可积, 从而积分号和求导号不能轻易交换 (在这个情况中交换事实上是不成立的).

课堂笔记 (2023.9.4)

- (关于 Laplace 方程的结构与一些拓展) 对 $x\in U\subset\mathbb{R}^n$ 来说, Laplace 方程可写成:

$$\Delta u=\sum_{i=1}^n\frac{\partial^2 u}{\partial x_i^2}=\sum_{i=j=1}^n\frac{\partial^2 u}{\partial x_i\partial x_j}$$

特别写成最右边的形式是为了强调 Laplace 方程中并未出现" 交叉项" $\frac{\partial^2 u}{\partial x_i\partial x_j}(i\neq j)$. 事实上, 这样的交叉项 $\frac{\partial^2 u}{\partial x_i\partial x_j}(i\neq j)$ 是可以被平方项 $\sum\limits_{i=j}\frac{\partial^2 u}{\partial x_i\partial x_j}$ 控制的, 这在调和分析中能通过 Riesz 变换来实现**LG1**. 要定义 Riesz 变换, 首先引入 $\mathbb{R}^n$ 上的缓增分布 W_j: 对 $1\leq j\leq n$, 设 $\mathscr{S}(\mathbb{R}^n)$ 是 $\mathbb{R}^n$ 上的速降函数空间, $\varphi\in\mathscr{S}(\mathbb{R}^n)$ 是测试函数, 则

$$\langle W_j,\varphi\rangle:=\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\lim_{\varepsilon\to 0}\int_{|y|\geq\varepsilon}\frac{y_j}{|y|^{n+1}}\varphi(y)dy$$

则函数 $f\in\mathscr{S}(\mathbb{R}^n)$ 的第 j 个 Riesz 变换定义为其与 W_j 的卷积:

$$R_j(f)(x)=(f*W_j)(x)=\frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\text{p.v.}\int_{\mathbb{R}^n}\frac{x_j-y_j}{|x-y|^{n+1}}f(y)dy$$

现在任取 $\varphi\in\mathscr{S}(\mathbb{R}^n)$, $1\leq j,k\leq n$, Riesz 变换就有性质:

$$\partial_j\partial_k\varphi(x)=-R_jR_k\Delta\varphi(x).$$

- (关于 Poisson 方程在 Newton 引力下的意义) 回忆万有引力公式:

$$F=-G\frac{Mm}{r^2}$$

将其写成矢量形式即:

$$\boldsymbol{F}=-G\frac{Mm}{r^2}\cdot\frac{\boldsymbol{r}}{r}=-G\frac{Mm}{r^3}\boldsymbol{r}$$

其中 $r=|\boldsymbol{r}|, \boldsymbol{r}$ 是位置向量, 负号是因为引力方向与坐标方向相反. 现在将 $\boldsymbol{g}=\frac{\boldsymbol{F}}{m}$ 代入有

$$\boldsymbol{g}=-\frac{GMm}{r^3m}\boldsymbol{r}=-\frac{GM}{r^3}\boldsymbol{r} \tag{7.15}$$

因为 $\boldsymbol{g}$ 与 $\boldsymbol{F}$ 同向, 故写成标量式有

$$g=\frac{GM}{r^2}$$

现在引力势即

$$\Phi=\int_r^{\infty} g\cdot dl=-\int_r^{\infty}\frac{GM}{r^2}dr=-\frac{GM}{r}$$

两边对 r 求导得

$$\frac{\partial}{\partial r}\Phi=\frac{GM}{r^2}$$

而引力场强方向为引力势的负梯度方向, 故上式的向量形式即

$$\boldsymbol{g}=-\nabla\Phi \tag{7.16}$$

将密度公式 $\rho=\frac{M}{V}$ 与球面体积公式 $V=\frac{4}{3}\pi r^3$ 代入(7.15)式得

$$\boldsymbol{g}=-\frac{GM}{r^3}\boldsymbol{r}=-\frac{G\rho V}{r^3}\boldsymbol{r}=-\frac{G\rho\frac{4}{3}\pi r^3}{r^3}\boldsymbol{r}=-\frac{4}{3}G\rho\pi\boldsymbol{r} \tag{7.17}$$

对(7.16)式求散度得

$$-\operatorname{div}\boldsymbol{g}=-\operatorname{div}(-\nabla\Phi)=\operatorname{div}(\nabla\Phi)=\nabla^2\Phi$$

另一方面, 对(7.17)式求散度得

$$\begin{aligned}\operatorname{div}(\frac{4}{3}G\rho\pi\boldsymbol{r})&=(\frac{\partial}{\partial r_x},\frac{\partial}{\partial r_y},\frac{\partial}{\partial r_z})\begin{pmatrix}\frac{4}{3}G\rho\pi r_x\\ \frac{4}{3}G\rho\pi r_y\\ \frac{4}{3}G\rho\pi r_z\end{pmatrix}=\frac{4}{3}G\rho\pi(\frac{\partial}{\partial r_x},\frac{\partial}{\partial r_y},\frac{\partial}{\partial r_z})\begin{pmatrix}r_x\\ r_y\\ r_z\end{pmatrix}\\&=\frac{4}{3}G\rho\pi\times 3=4G\rho\pi\end{aligned}$$

即得

$$\nabla^2\Phi=\Delta\Phi=4\pi G\rho$$

这便是 Newton 引力场方程, 而它正是一个 Poisson 方程.

- (关于 Laplace 方程在旋转下不变[a]) 设 $v(x)=u(Ox)$, 则

$$\begin{aligned}\Delta v(x)=(\operatorname{div}\nabla)v(x)&=(\operatorname{div}\nabla)(u(Ox))=\operatorname{div}(O(\nabla u)(Ox))\\&=O^T\operatorname{div}(O\nabla u)(Ox)=O^TO(\operatorname{div}\nabla)u(Ox)=\Delta u(Ox)=0\end{aligned}$$

需要注意的是上面的过程并不是简单地通过在 Laplace 算子下把 $v(x)$ 换成 $u(Ox)$ 得到的! 这是因为最开始的 Laplace 算子对应的 x, 与最后 Laplace 算子对应的 x 并不相同, 可以理解成若直接代换会出现 $\frac{\partial u(Ox)}{\partial(Ox)}$ 而非 $\frac{\partial u(Ox)}{\partial x}$.

- (关于 Poisson 方程的解是基本解与 $f(x)$ 卷积的想法) 后面在探讨 Green 函数的时候会提及 Laplace 方程的基本解 $\Phi(x)$ 所满足的方程:

$$-\Delta\Phi=\delta_0 \tag{7.18}$$

其中 δ_0 是 δ 函数. 因为 $(f*\delta_0)(x)=f(x)$, 故 Poisson 方程能写成:

$$-\Delta u=f(x)=(f*\delta_0)(x)=\int_{\mathbb{R}^n}f(x-y)\delta_0(y)dy$$

将(7.18)式代回上式得

$$-\Delta u=-\int_{\mathbb{R}^n}f(x-y)\Delta\Phi(y)dy$$

这便自然而然地希望

$$\int_{\mathbb{R}^n} f(x-y)\Delta\Phi(y)dy = \Delta\int_{\mathbb{R}^n} f(x-y)\Phi(y)dy$$

进而猜测

$$u = \int_{\mathbb{R}^n} f(x-y)\Phi(y)dy = (f*\Phi)(x)$$

[a]2023.9.21 记: 这一部分依旧存疑, 详情参见问题部分的练习(7.2).

下面假定 $f\in C_c^2(\mathbb{R}^n)$, 也即 f 是二阶连续可微的紧支函数. 进入讨论之前先给出积分平均的定义.

定义 7.2.3 (积分平均)

$$⨍_E f d\mu := \frac{1}{\mu(E)}\int_E f d\mu$$

称为 f 在 E 上的积分平均, 其中 $\mu(E)$ 是 E 的 μ 测度.

定理 7.2.1 (Poisson 方程的解)

按(7.14)式定义 u, 则

(i) $u\in C^2(\mathbb{R}^n)$,

(ii) 在 $\mathbb{R}^n$ 中有 $-\Delta u = f$.

证明

首先根据卷积的交换性有

$$u(x) = \int_{\mathbb{R}^n}\Phi(x-y)f(y)dy = \int_{\mathbb{R}^n}\Phi(y)f(x-y)dy \tag{7.19}$$

从而

$$\frac{1}{h}(u(x+he_i)-u(x)) = \int_{\mathbb{R}^n}\Phi(y)\cdot\frac{1}{h}(f(x+he_i-y)-f(x-y))dy$$

其中 $h\neq 0, e_i = (\underbrace{0,\cdots,0,1}_{i},0,\cdots)$. 根据 $f\in C_c^2(\mathbb{R}^n)$ 知

$$\frac{1}{h}(f(x+he_i-y)-f(x-y)) \rightrightarrows f_{x_i}(x-y),\quad h\to 0$$

其中一致性是因为趋向是从连续函数到连续函数的, 因而根据一致性知极限与积分号可换序, 得到

$$u_{x_i}(x) = \int_{\mathbb{R}^n}\Phi(y)f_{x_i}(x-y)dy,\quad i=1,2,\cdots,n$$

类似有

$$u_{x_ix_j}(x) = \int_{\mathbb{R}^n}\Phi(y)f_{x_ix_j}(x-y)dy,\quad i,j=1,2,\cdots,n \tag{7.20}$$

注意(7.20)式右侧是关于 x 的连续函数, 故 $u\in C^2(\mathbb{R})$.

因为 Φ 在 $x=0$ 处是爆破的, 故在 0 附近需要单独划分小邻域来讨论. 固定 $\varepsilon>0$, 有

$$\Delta u(x) = \int_{B(0,\varepsilon)}\Phi(y)\Delta_x f(x-y)dy + \int_{\mathbb{R}^n\backslash B(0,\varepsilon)}\Phi(y)\Delta_x f(x-y)dy =: I_\varepsilon + J_\varepsilon \tag{7.21}$$

根据 $\|\cdot\|_{L^\infty(\mathbb{R}^n)}$ 的定义有

$$|I_\varepsilon| \le C\|D^2f\|_{L^\infty(\mathbb{R}^n)}\int_{B(0,\varepsilon)}|\Phi(y)|dy \le \begin{cases} C\varepsilon^2|\ln\varepsilon|, & n=2\\ C\varepsilon^2, & n\ge 3\end{cases} \tag{7.22}$$

其中 C 是某常数. 针对 J_ε 分部积分有

$$J_\varepsilon = \int_{\mathbb{R}^n\backslash B(0,\varepsilon)}\Phi(y)\Delta_y f(x-y)dy = -\int_{\mathbb{R}^n\backslash B(0,\varepsilon)}D\Phi(y)\cdot D_y f(x-y)dy + \int_{\partial B(0,\varepsilon)}\Phi(y)\frac{\partial f}{\partial\nu}(x-y)dS(y) =: K_\varepsilon + L_\varepsilon$$

其中 $\boldsymbol{\nu}$ 表示 $\partial B(0,\varepsilon)$ 上的单位内法向量. 对 L_ε 有

$$|L_\varepsilon| \le \|Df\|_{L^\infty(\mathbb{R}^n)} \int_{\partial B(0,\varepsilon)} |\Phi(y)|dS(y) \le \begin{cases} C\varepsilon|\ln\varepsilon|, & n=2 \\ C\varepsilon, & n\ge 3 \end{cases} \tag{7.23}$$

对 K_ε, 继续分部积分有

$$K_\varepsilon = \int_{\mathbb{R}^n\backslash B(0,\varepsilon)} \Delta\Phi(y)f(x-y)dy - \int_{\partial B(0,\varepsilon)} \frac{\partial\Phi}{\partial\boldsymbol{\nu}}(y)f(x-y)dS(y) = -\int_{\partial B(0,\varepsilon)} \frac{\partial\Phi}{\partial\boldsymbol{\nu}}(y)f(x-y)dS(y)$$

其中第二个等号成立时因为在 $\mathbb{R}^n\backslash B(0,\varepsilon)$ 上 Φ 是调和的, 也即 $\Delta\Phi=0$. 现在在 $\partial B(0,\varepsilon)$ 上有:

$$D\Phi(y) = -\frac{1}{n\alpha(n)}\frac{y}{|y|^n}(y\neq 0), \quad \boldsymbol{\nu} = -\frac{y}{|y|} = -\frac{y}{\varepsilon}$$

故在 $\partial B(0,\varepsilon)$ 上有

$$\frac{\partial\Phi}{\partial\boldsymbol{\nu}}(y) = \boldsymbol{\nu}\cdot D\Phi(y) = \frac{1}{n\alpha(n)\varepsilon^{n-1}}$$

注意 $n\alpha(n)\varepsilon^{n-1}$ 恰是 $\partial B(x,\varepsilon)$ 的 Lebesgue 测度, 进而根据 f 的连续性有

$$K_\varepsilon = -\fint_{\partial B(x,\varepsilon)} f(y)dS(y) \to -f(x), \quad \varepsilon\to 0 \tag{7.24}$$

现联立(7.21)-(7.24)式, 令 $\varepsilon\to 0$ 即得 $-\Delta u(x) = f(x)$. □

对于 Poisson 方程的基本解, 有时在 $\mathbb{R}^n$ 中记

$$-\Delta\Phi_0 = \delta_0$$

其中 δ_0 表示 $\mathbb{R}^n$ 上的 Dirac 测度, 它在 0 点赋单位质量 (也就是广义函数中的 δ 函数). 根据这个记号有

$$-\Delta u(x) = \int_{\mathbb{R}^n} -\Delta_x\Phi(x-y)f(y)dy = \int_{\mathbb{R}^n} \delta_x f(y)dy = f(x), \quad x\in\mathbb{R}^n$$

注 在引入广义函数后, 基本解自身是有严格定义的, 下面的内容选自 **[OAO]**.

定义 7.2.4 (基本解)

如果函数 $V(x,x^0)$ 是空间 $\mathscr{D}'(\mathbb{R}^n_x)$ 中的元素 (也即广义函数), 且满足方程

$$\Delta V = \delta(x-x^0)$$

其中 δ 表示 δ 函数, 也即:

$$\langle\delta,\varphi\rangle = \varphi(0), \quad \langle\delta(x-x^0),\varphi(x)\rangle = \varphi(x^0)$$

则称 $V(x,x^0)$ 是 Laplace 方程的基本解. ♣

可以验证前面得到的 $\Phi(x-x^0)$(下文简记为 $\Phi^0(x)$) 是满足这个基本解的定义的. 显见 Φ^0 是 $\mathbb{R}^n_x$ 中的局部可积函数, 因而可以推广为广义函数 (例(5.4)), 故 $\Phi^0\in\mathscr{D}'(\mathbb{R}^n_x)$. 下面验证 $\Delta\Phi^0 = \delta(x-x^0)$. 任取 $\varphi\in\mathscr{D}(\mathbb{R}^n_x)$, 由广义微商的定义知:

$$\langle\Delta\Phi^0,\varphi\rangle = \langle\Phi^0,\Delta\varphi\rangle \tag{7.25}$$

又因为 $\Phi^0\in L^1_{loc}(\mathbb{R}^n_x)$, 可得

$$\langle\Phi^0,\Delta\varphi\rangle = \int_{\mathbb{R}^n_x}\Phi^0(x)\Delta\varphi(x)dx = \lim_{\varepsilon\to 0}\int_{\mathbb{R}^n_x\backslash B(x_0,\varepsilon)}\Phi^0(x)\Delta\varphi(x)dx \tag{7.26}$$

对(7.26)式右端用 Green 公式(6.2.3)(iii) 有:

$$\int_{\mathbb{R}^n_x\backslash B(x_0,\varepsilon)}\Phi^0(x)\Delta\varphi(x)dx = \int_{\mathbb{R}^n_x\backslash B(x_0,\varepsilon)}\Delta\Phi^0(x)\varphi(x)dx + \int_{\partial B(x_0,\varepsilon)}(\Phi\frac{\partial\varphi}{\partial\boldsymbol{\nu}} - \varphi\frac{\partial\Phi^0}{\partial\boldsymbol{\nu}})dS \tag{7.27}$$

其中 $\boldsymbol{\nu}$ 是 $\partial B(x_0,\varepsilon)$ 上的单位内法向量. 因为基本解本身满足 Laplace 方程, 故在 $\mathbb{R}^n_x\backslash B(x_0,\varepsilon)$ 内有 $\Delta\Phi^0(x)=0$, 因而

$$\int_{\mathbb{R}^n_x\backslash B(x_0,\varepsilon)}\Delta\Phi^0(x)\varphi(x)dx = 0 \tag{7.28}$$

另一方面知:

$$|\int_{\partial B(x_0,\varepsilon)} \Phi^0 \frac{\partial \varphi}{\partial \boldsymbol{\nu}} dS| \le \sup_{\partial B(x_0,\varepsilon)} |\Phi^0 \frac{\partial \varphi}{\partial \boldsymbol{\nu}}| \cdot C\varepsilon^{n-1} \to 0, \quad \varepsilon \to 0 \tag{7.29}$$

而根据基本解的式子(7.12)知当 $|x-x^0| \equiv \varepsilon$ 时有:

$$D\Phi^0(x) = -\frac{1}{n\alpha(n)} \frac{x-x^0}{|x-x^0|^n}, \quad \boldsymbol{\nu} = -\frac{x-x^0}{\varepsilon}$$

从而

$$\frac{\partial \Phi^0}{\partial \boldsymbol{\nu}} = D\Phi^0 \cdot \boldsymbol{\nu} = \frac{1}{n\alpha(n)\varepsilon^{n-1}}$$

得到

$$\lim_{\varepsilon\to 0} \int_{\partial B(x_0,\varepsilon)} \varphi \frac{\partial \Phi^0}{\partial \boldsymbol{\nu}} dS = \lim_{\varepsilon\to 0} \frac{1}{n\alpha(n)\varepsilon^{n-1}} \int_{\partial B(x_0,\varepsilon)} \varphi dS = \lim_{\varepsilon\to 0} \fint_{\partial B(x_0,\varepsilon)} \varphi dS = \varphi(x_0) \tag{7.30}$$

现联立(7.25)-(7.30)即得

$$\langle \Delta\Phi^0, \varphi \rangle = \varphi(x_0)$$

此即 $\Delta\Phi(x-x^0) = \delta(x-x^0)$.

7.2.2 均值定理

现在设 $U \subset \mathbb{R}^n$ 是开集, u 是 U 上的调和函数. 下面介绍调和函数的均值定理, 也即调和函数在某点处的值, 等于它在以该点为球心的球中的均值, 也等于它在该球球面上的均值. 这些公式进而可以用来导出调和函数的诸多性质.

定理 7.2.2 (Laplace 方程的均值定理)

若 $u \in C^2(U)$ 是调和函数, 则

$$\forall B(x,r) \subset U (u(x) = \fint_{\partial B(x,r)} udS = \fint_{B(x,r)} udy) \tag{7.31}$$

证明

记

$$\phi(r) := \fint_{\partial B(x,r)} u(y)dS(y) = \fint_{\partial B(0,1)} u(x+rz)dS(z)$$

则

$$\phi'(r) = \fint_{\partial B(0,1)} Du(x+rz) \cdot zdS(z)$$

进而根据高-奥公式(6.2.1), 令 $y = x + rz$ 有

$$\begin{aligned}\phi'(r) &= \fint_{\partial B(x,r)} Du(y) \cdot \frac{y-x}{r} dS(y) = \fint_{\partial B(x,r)} Du(y) \cdot \boldsymbol{\nu}(y) dS(y) = \fint_{\partial B(x,r)} \frac{\partial u}{\partial \boldsymbol{\nu}} dS(y) \\ &= \fint_{\partial B(x,r)} \sum_{i=1}^n \frac{\partial u}{\partial y^i} \cos\alpha_i dS(y) = \frac{r}{n} \fint_{B(x,r)} \Delta u(y) dy = \frac{r}{n} \fint_{B(x,r)} 0 dy = 0\end{aligned}$$

其中 $\cos\alpha_i$ 表示 y^i 方向的方向余弦. 故 $\phi(r)$ 是常数, 因而

$$\phi(r) = \lim_{t\to 0} \phi(t) = \lim_{t\to 0} \fint_{\partial B(x,t)} u(y)dS(y) = u(x)$$

再证明

$$u(x) = \fint_{B(x,r)} udy$$

将积分平均的定义代入前述结论知

$$n\alpha(n) r^{n-1} \cdot u(x) = \int_{\partial B(x,r)} udS \tag{7.32}$$

根据极坐标换元有

$$\int_{B(x,r)} udy = \int_0^r (\int_{\partial B(x,s)} udS)ds$$

将(7.32)式代入(7.2.2)式有

$$\int_{B(x,r)} udy = \int_0^r u(x)\cdot n\alpha(n)s^{n-1}ds = \alpha(n)r^n u(x)$$

此即

$$u(x) = ⨍_{B(x,r)} udy$$

□

定理 7.2.3 (均值定理的逆定理)

若 $u \in C^2(U)$ 满足对任意的 $B(x,r) \subset U$ 都有

$$u(x) = ⨍_{\partial B(x,r)} udS$$

则 u 是调和函数.

证明

用反证法, 若 $\Delta u \not\equiv 0$, 则存在 $B(x,r) \subset U$ 满足在其内 $\Delta u \neq 0$, 不妨设 $\Delta u > 0$, 取

$$\phi(r) = ⨍_{\partial B(x,r)} udS$$

则一方面由题式知 $\phi(r)$ 与 r 无关, 因而 $\phi'(r) = 0$, 另一方面根据前述结论有

$$\phi'(r) = \frac{r}{n}⨍_{B(x,r)} \Delta u(y)dy > 0$$

矛盾! 命题即证. □

7.2.3 调和函数的性质

设 $U \subset \mathbb{R}^n$ 是开的有界集, 下面通过均值定理导出调和函数的一系列性质.

7.2.3.1 强最大值原理与解的唯一性

首先说明调和函数的最大值原理, 即非常值的调和函数的最大值必在边界取得, 且在连通区域的内部无法取得最大值.

定理 7.2.4 (强最大值原理)

设 $u \in C^2(U) \cap C(\overline{U})$ 是 U 内的调和函数, 则

(i) $\max\limits_{\overline{U}} u = \max\limits_{\partial U} u$.

(ii) 若 U 是连通的, 且存在 $x_0 \in U$ 使得

$$u(x_0) = \max_{\overline{U}} u$$

则 u 是 U 内的常值函数.

结论 (i) 称为 Laplace 方程的最大值原理, 而结论 (ii) 称为 Laplace 方程的强最大值原理. 把 u 换成 $-u$, 也可以类似得到最小值原理与强最小值原理.

证明

设存在 $x_0 \in U$ 使得 $u(x_0) = M := \max_{\overline{U}} u$, 则取 $0 < r < \rho(x_0, \partial U)$, 由均值定理知

$$M = u(x_0) = \fint_{B(x_0,r)} u dy \le M \tag{7.33}$$

(7.33)式中的等号成立当且仅当在 $B(x_0, r)$ 中有 $u \equiv M$, 故 $\forall y \in B(x_0, r)(u(y) = M)$, 这说明集合 $\{x \in U : u(x) = M\}$ 是开集, 且在 U 内相对闭, 从而在 U 是连通集时它就是 U. 这便证明了结论 (ii), 结论 (i) 进而显见. □

强最大值原理说明如果 U 是连通的, 且 $u \in C^2(U) \cap C(\overline{U})$ 满足

$$\begin{cases} \Delta u = 0, & \text{在 } U \text{ 内} \\ u = g, & \text{在 } \partial U \text{ 上} \end{cases}$$

其中 $g \ge 0$, 则只要 g 在 ∂U 上可取正值, 那么 u 在 U 内就是处处正的.

课堂笔记 (2023.9.11)

- (关于弱极值原理与强极值原理) 结论 (i) 又称为弱极值原理, 其问题在于只确定边界上能取极值, 但不知道区域内是否也能取到极值, 而强极值原理 (ii) 就说明极值只能在边界上取到.
- (关于标蓝部分在拓扑意义下的解释) 当 U 是 (路) 连通集, 其有同时在 U 内为开集和闭集的子集 W. 现设 W 不是空集, 进而可设 $x \in W$. 如若 $W^c = U \backslash W$ 同样非空, 可设 $y \in W^c$, 根据 U 的连通性可设在 U 内存在连接 x, y 的曲线 l. 设 $l \cap \partial W = z$, 现在因为 W 是闭集, 知 $z \in W$, 又因为 ∂W 的定义表明 z 也可由 W^c 中的点来趋近, 故 $z \in W^c$, 矛盾! 故只能有 $W^c = \varnothing$, 也即 $W = U$ 是全集.
- (关于调和函数的变号问题) 当 $\Delta u = 0, u|_{\partial U} = g \ge 0$ 时, 根据最值原理:

$$0 \le \min_{\partial U} g = \min_{\partial U} u \le u \le \max_{\partial U} u$$

这说明 $u \ge 0$. 而当

$$g(x) \ge 0, \quad \exists x_0 \in \partial U (g(x_0) > 0)$$

则在 U 内必有 $u > 0$, 这是因为若存在 $y_0 \in U$ 使得 $u(y_0) = 0$, 则

$$0 = u(y_0) = \min_{\overline{U}} u = \min_{\partial U} u = \min_{\partial U} g$$

根据强极值原理, 当 y_0 在 U 内, 这便说明 u 在区域内取到极小值, 因而其只能为常数, 从而 $u \equiv 0 \Rightarrow g \equiv 0$, 矛盾!

注 [2]标蓝部分其实是拓扑连通的等价定义. 称集合 U 是 (拓扑) 连通的, 如果它不能被表成两个非空开集的并. 如果 U 有一个既开又闭的子集 W, 显见 W^c 也既开又闭, 于是 $U = W \cup W^c$ 是两个开集的并, 这便只能有 $W = \varnothing$ 或 $W^c = \varnothing$, 亦即 (拓扑) 连通集 U 中既开又闭的集合只可能是 $\varnothing$ 或 U.

注 注意调和函数还是允许边界变号的, 只是说边界不变号时调和函数也不会变号.

最大值原理的一个重要应用是说明 Poisson 问题的一类边值问题解的唯一性.

定理 7.2.5 (唯一性)

设 $g \in C(\partial U), f \in C(U)$, 则边值问题

$$\begin{cases} -\Delta u = f, & \text{在 } U \text{ 内} \\ u = g, & \text{在 } \partial U \text{ 上} \end{cases} \tag{7.34}$$

至多只有一个解 $u \in C^2(U) \cap C(\overline{U})$.

证明

[2]2023.12.3 注.

若边值问题(7.34)有满足条件的解 $u, \tilde{u}$, 则 $\pm(u-\tilde{u})$ 都是初值问题

$$\begin{cases} \Delta u = 0, & 在\ U\ 内 \\ u = 0, & 在\ \partial U\ 上 \end{cases}$$

的解, 但根据最大值原理(7.2.4)知在 ∂U 上有 $u-\tilde{u} \le 0, \tilde{u}-u \le 0$, 这说明 $u-\tilde{u} = \tilde{u}-u \equiv 0$, 也即 $u = \tilde{u}$. □

7.2.3.2 正则性

下面考虑调和函数的正则性, 我们来证明如果 $u \in C^2$ 是调和函数, 则必有 $u \in C^\infty$. 此即调和函数均无穷次可微, 与之相关的结论被称为正则性定理. 这件事的有趣之处在于从 Laplace 方程 $\Delta u = \sum_{i=1}^{n} u_{x_i x_i} = 0$ 的代数结构推出了 u 的所有偏导数全部存在这一分析结果, 而这是单靠 PDE 没法阐明的.

定理 7.2.6 (光滑性)

如果对每个球 $B(x,r) \subset U$, 函数 $u \in C(U)$ 都满足均值定理(7.31), 则

$$u \in C^\infty(U)$$

♡

注 条件中没有说明 u 在 ∂U 上是否光滑, 甚至是否连续.

证明

设 η 是磨光算子(5.1), 即

$$C^\infty(\mathbb{R}^n) \ni \eta(x) = \begin{cases} C\exp\left(\frac{1}{|x|^2-1}\right), & |x| < 1 \\ 0, & |x| \ge 1 \end{cases}, \quad C = (\int_{\mathbb{R}^n} \exp\left(\frac{1}{|x|^2-1}\right) dx)^{-1}$$

知 η 也是放射状的函数, 即只与 $r = |x|$ 有关的函数. 取

$$\eta_\varepsilon := \frac{1}{\varepsilon^n}\eta(\frac{x}{\varepsilon})$$

并在 $U_\varepsilon = \{x \in U : \rho(x, \partial U) > \varepsilon\}$ 中设

$$u^\varepsilon := \eta_\varepsilon * u$$

根据对磨光算子性质的讨论 (命题(5.1.1)) 知 $u^\varepsilon \in C^\infty(U_\varepsilon)$.

下面说明在 U_ε 上有 $u \equiv u^\varepsilon$. 设 $x \in U_\varepsilon$, 则:

$$\begin{aligned} u^\varepsilon(x) &= \int_U \eta_\varepsilon(x-y)u(y)dy = \frac{1}{\varepsilon^n}\int_{B(x,\varepsilon)} \eta(\frac{|x-y|}{\varepsilon})u(y)dy \\ &= \frac{1}{\varepsilon^n}\int_0^\varepsilon \eta(\frac{r}{\varepsilon}) \cdot (\int_{\partial B(x,r)} u dS)dr = \frac{1}{\varepsilon^n}u(x)\int_0^\varepsilon \eta(\frac{r}{\varepsilon})n\alpha(n)r^{n-1}dr \\ &= u(x)\int_{B(0,\varepsilon)} \eta_\varepsilon dy = u(x) \end{aligned}$$

这说明在 U_ε 上有 $u^\varepsilon \equiv u$, 故对任意的 $\varepsilon > 0$ 有 $u \in C^\infty(U_\varepsilon)$. □

7.2.3.3 调和函数的局部估计

下面利用均值定理来对调和函数的各阶偏导做更细致的估计.

定理 7.2.7 (调和函数的导数估计)

若 u 是 U 上的调和函数, 则对任意的球 $B(x_0,r) \subset U$, 任意的 k 阶多重指标 α 有:

$$|D^\alpha u(x_0)| \le \frac{C_k}{r^{n+k}}\|u\|_{L^1(B(x_0,r))}$$

其中

$$C_0 = \frac{1}{\alpha(n)}, \quad C_k = \frac{(2^{n+1}nk)^k}{\alpha(n)}, \quad k = 1, 2, \cdots$$

证明

用归纳法. 当 $k = 0$, 知此即

$$|u(x_0)| \leq \frac{1}{\alpha(n)r^n}\|u\|_{L^1(B(x_0,r))}$$

这由均值定理立得. 当 $k = 1$, 通过在 Laplace 方程两边求偏导可知 $u_{x_i}(i = 1, 2, \cdots, n)$ 也是调和函数, 故

$$\begin{aligned} |u_{x_i}(x_0)| &= |\fint_{B(x_0,\frac{r}{2})} u_{x_i}dx| = |\frac{1}{\alpha(n)(\frac{r}{2})^n}\int_{\partial B(x_0,\frac{r}{2})} u\nu_i dS| \\ &\leq \frac{2^n}{\alpha(n)r^n}\cdot n\alpha(n)(\frac{r}{2})^{n-1}\|u\|_{L^\infty(\partial B(x_0,\frac{r}{2}))} = \frac{2n}{r}\|u\|_{L^\infty(\partial B(x_0,\frac{r}{2}))} \end{aligned} \tag{7.35}$$

下面估计 $\|u\|_{L^\infty(\partial B(x_0,\frac{r}{2}))}$, 取 $x \in \partial B(x_0, \frac{r}{2})$, 知 $B(x, \frac{r}{2}) \subset B(x_0, r) \subset U$, 根据 $k = 0$ 时的结论有:

$$|u(x)| \leq \frac{1}{\alpha(n)}(\frac{2}{r})^n\|u\|_{L^1(B(x_0,r))} \tag{7.36}$$

联立(7.35), (7.36)式知 $|\alpha| = 1$ 时有

$$|D^\alpha u(x_0)| \leq \frac{2n}{r}\|u\|_{L^\infty(\partial B(x_0,\frac{r}{2}))} \leq \frac{2^{n+1}n}{\alpha(n)}\frac{1}{r^{n+1}}\|u\|_{L^1(B(x_0,r))}$$

$k = 1$ 的情况得证.

现设 $k \geq 2$, 命题对 $k-1$ 阶及以下的多重指标都成立. 固定 $B(x_0, r) \subset U$, 取 $|\alpha| = k$, 知存在 $n-1$ 阶多重指标 β 与下标 $i \in \{1, 2, \cdots, n\}$ 使得 $D^\alpha u = (D^\beta u)_{x_i}$. 类似于(7.35)式的过程可得

$$|D^\alpha u(x_0)| \leq \frac{nk}{r}\|D^\beta u\|_{L^\infty(\partial B(x_0,\frac{r}{k}))} \tag{7.37}$$

下面估计 $\|D^\beta u\|_{L^\infty(\partial B(x_0,\frac{r}{k}))}$, 取 $x \in \partial B(x_0, \frac{r}{k})$, 则 $B(x, \frac{k-1}{k}r) \subset B(x_0, r) \subset U$, 应用 $k-1$ 阶情况下的命题知

$$|D^\beta u(x)| \leq \frac{(2^{n+1}n(k-1))^{k-1}}{\alpha(n)(\frac{k-1}{k}r)^{n+k-1}}\|u\|_{L^1(B(x_0,r))} \tag{7.38}$$

联立(7.37), (7.38)式得到

$$|D^\alpha u(x_0)| \leq \frac{nk}{r}\|D^\beta u\|_{L^\infty(\partial B(x_0,\frac{r}{k}))} \leq \frac{(2^{n+1}nk)^k}{\alpha(n)r^{n+k}}\|u\|_{L^1(B(x_0,r))}$$

此即欲证. □

7.2.3.4 Liouville 定理

下面说明不存在 $\mathbb{R}^n$ 上的非平凡有界调和函数.

定理 7.2.8 (Liouville)

若 $u : \mathbb{R}^n \to \mathbb{R}^n$ 是有界的调和函数, 则 u 是常数.

证明

固定 $x_0 \in \mathbb{R}^n, r > 0$, 在 $B(x_0, r)$ 上应用估计(7.2.7)知

$$|Du(x_0)| \leq \frac{C_1}{r^{n+1}}\|u\|_{L^1(B(x_0,r))} \leq \frac{C_1\alpha(n)}{r}\|u\|_{L^\infty(B(x_0,r))} \to 0, \quad r \to \infty$$

故 $Du \equiv 0$, 也即 u 是常数. □

定理 7.2.9 (调和函数的表示公式)

若 $f\in C_c^2(\mathbb{R}^n), n\geq 3$, 则 $\mathbb{R}^n$ 上的方程

$$-\Delta u=f$$

的任意有界解具有下述形式:

$$u(x)=\int_{\mathbb{R}^n}\Phi(x-y)f(y)dy+C,\quad x\in\mathbb{R}^n$$

其中 C 是常数.

证明

从 $\Phi(x)$ 的定义知当 $n\geq 3$ 时, $\Phi(x)\to 0(|x|\to\infty)$. 这说明 $\tilde{u}(x):=\int_{\mathbb{R}^n}\Phi(x-y)f(y)dy$ 是命题方程的一个有界解. 若 u 是该方程的另一个有界解, 则 $u-\tilde{u}$ 就是一个有界的调和函数, 由 Liouville 定理知 $u-\tilde{u}$ 是常数, 命题即证.□

注 $n=2$ 时, 因为 $\Phi(x)=\frac{1}{2\pi}\ln|x|$ 在 $|x|\to\infty$ 时是无界的, 故 $u(x):=\int_{\mathbb{R}^2}\Phi(x-y)f(y)dy$ 也可能是无界的.

7.2.3.5 解析性

定理 7.2.10 (调和函数的解析性)

若 u 是 U 上的调和函数, 那么它在 U 上解析.

证明

任意固定 $x_0\in U$, 需要证明 u 在 x_0 的某一邻域内能被表成某个收敛的幂级数. 设 $r:=\frac{1}{4}\rho(x_0,\partial U)$, 显见

$$M:=\frac{1}{\alpha(n)r^n}\|u\|_{L^1(B(x_0,2r))}<\infty$$

任取 $x\in B(x,r)$, 知 $B(x,r)\subset B(x_0,2r)\subset U$, 根据调和函数的导数估计(7.2.7)知

$$\|D^\alpha u\|_{L^\infty(B(x_0,r))}\leq M(\frac{2^{n+1}n}{r})^{|\alpha|}|\alpha|^{|\alpha|}\tag{7.39}$$

注意对任意的正整数 k 都有 $\frac{k^k}{k!}<e^k$, 从而对 $|\alpha|^{|\alpha|}$ 有:

$$|\alpha|^{|\alpha|}\leq e^{|\alpha|}|\alpha|!\tag{7.40}$$

进一步由多项式定理知

$$n^k=(1+1+\cdots+1)^k=\sum_{|\alpha|=k}\frac{|\alpha|!}{\alpha!}$$

这表明

$$|\alpha|!\leq n^{|\alpha|}|\alpha|!$$

将(7.40)式代入(7.39)式有:

$$\|D^\alpha u\|_{L^\infty(B(x_0,r))}\leq M(\frac{2^{n+1}n^2e}{r})^{|\alpha|}|\alpha|\tag{7.41}$$

要得到收敛到 $u(x)$ 的幂级数, 最自然的便是 Taylor 级数. 现在考虑 u 在 x_0 处的 Taylor 级数

$$\sum_{\alpha}\frac{D^\alpha u(x_0)}{\alpha!}(x-x_0)^\alpha\tag{7.42}$$

其中和式取遍全体多重指标. 希望证明该级数在 x_0 的某邻域内收敛. 现在给定 N, 分析(7.42)式的余项:

$$R_N(x):=u(x)-\sum_{k=0}^{N-1}\sum_{|\alpha|=k}\frac{D^\alpha u(x_0)(x-x_0)^\alpha}{\alpha!}=\sum_{|\alpha|=N}\frac{D^\alpha u(x_0+t(x-x_0))(x-x_0)^\alpha}{\alpha!}$$

其中 $t \in [0,1]$ 是关于 x 的. 将(7.41)式代入知

$$|R_N(x)| \le \sum_{|\alpha|=N} M(\frac{2^{n+1}n^2e}{r})^{|\alpha|}|\alpha| \cdot \frac{|x-x_0|^\alpha}{\alpha!}$$

要使得 $N \to \infty$ 时有 $|R_N(x)| \to 0$, 注意当 α 遍历 N 阶多重指标时, 必有

$$\sum_{|\alpha|=N} \frac{|\alpha|}{\alpha!} < 1$$

得到:

$$\begin{aligned} |R_N(x)| &\le \sum_{|\alpha|=N} M(\frac{2^{n+1}n^2e}{r})^{|\alpha|} \cdot |x-x_0|^{|\alpha|} \\ &= \frac{(n+N-1)!}{(n-1)!N!} M(\frac{2^{n+1}n^2e}{r})^N \cdot |x-x_0|^N \\ &< n^N \cdot M(\frac{2^{n+1}n^2e}{r})^N \cdot |x-x_0|^N \end{aligned}$$

现令

$$|x-x_0| < \frac{r}{2^{n+2}n^3e}$$

可得

$$|R_N(x)| < M \cdot n^N \cdot (\frac{2^{n+1}n^2e}{r})^N \cdot (\frac{r}{2^{n+2}n^3e})^N = \frac{M}{2^N} \to 0, \quad N \to \infty$$

故级数(7.42)至少在 $B(x_0, \frac{r}{2^{n+2}n^3e})$ 中收敛, 根据 x_0 的任意性即得欲证. □

7.2.3.6 Harnack 不等式

定义 7.2.5 (紧包含)

若 $V \subset \overline{V} \subset U$, 且 $\overline{V}$ 是紧集, 则称 V 紧包含于 U, 记作 $V \subset\subset U$. ♣

定理 7.2.11 (Harnack 不等式)

设 $U \subset \mathbb{R}^n$ 是开集, 对每个连通开集 V, 若 $V \subset\subset U$, 则存在仅与 V 相关的正常数 C 满足

$$\sup_V u \le C \inf_V u$$

其中 u 是 U 上的任意非负调和函数. ♡

Harnack 不等式说明对任意的 $x, y \in V$ 都有

$$\frac{1}{C}u(y) \le u(x) \le Cu(y)$$

这表明非负调和函数在 V 上取值的波动较小: 如果 u 在 V 上处处取小值, 才可能在某点取值小. 一个直观的解释是既然 V 离 U 的边界还有一段距离 (这是 $V \subset\subset U$ 所保证的), 对每一点而言均值定理就都有用武之地.

证明

设 $r := \frac{1}{4}\rho(V, \partial U)$, 选定 $x, y \in V$ 满足 $|x-y| \le r$, 则根据 u 的非负性知

$$u(x) = ⨍_{B(x,2r)} udz = \frac{1}{\alpha(n)(2r)^n}\int_{B(x,2r)} udz \ge \frac{1}{\alpha(n)2^nr^n}\int_{B(y,r)} udz = \frac{1}{2^n}⨍_{B(y,r)} udz = \frac{1}{2^n}u(y)$$

同理可证 $u(y) \ge \frac{1}{2^n}u(x)$, 故

$$\forall x, y \in V(|x-y| \le r \Rightarrow 2^n u(y) \ge u(x) \ge \frac{1}{2^n}u(y))$$

因为 V 是连通的, 而 $\overline{V}$ 是紧集, 故根据紧集的定义知可用有限个开球 $\{B_i\}_{i=1}^N$ 覆盖 $\overline{V}$, 其中每个开球的半径为 $\frac{r}{2}$, 且 $B_i \cap B_{i-1} \ne \varnothing (i = 2, 3, \cdots, N)$.

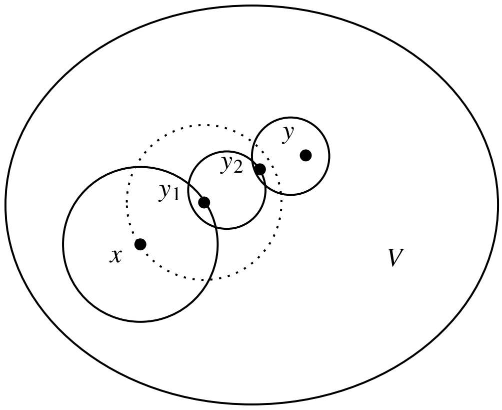


图 7.4

不断套用前述结论知[3]对任意的 $x, y \in V$ 有 (示意图如图(7.4)):

$$u(x) \geq \frac{1}{2^n(N+1)} u(y)$$

命题即证. □

课堂笔记 (2023.9.18)

- (关于 $n \geq 3$ 时 $(\Phi * f)(x)$ 是 Poisson 方程有界解的详细论述) 当 $n \geq 3$ 时有:
$$\Phi(x) = \frac{1}{n(n-2)\alpha(n)} |x|^{-(n-2)} \to 0, \quad |x| \to \infty$$
现在因为 $f \in C_c^2(\mathbb{R}^n)$, 故至少有
$$\overline{u} = \int_{\mathbb{R}^n} \Phi(x-y) f(y) dy \in C^2(\mathbb{R}^n)$$
设 f 的支集为 U, 则
$$\overline{u} = \int_U \Phi(x-y) f(y) dy$$
现若 x 在 y 附近, 则
$$|\overline{u}| \leq \int_U \Phi(x-y) \|f\|_{L^\infty(U)} dy = \|f\|_{L^\infty(U)} \int_{U'} \Phi(y) dy$$
其中 U' 是将 $x-y$ 换元后得到的新区域. 当 $0 \notin U'$, 自然有 $\int_{U'} \Phi(y) dy < \infty$; 当 $0 \in U'$, 在 $y \to 0$ 时, 记 $r = |y|$, 有:
$$\int_{U'} \Phi(y) dy \leq \int_{U'} \frac{C_1}{|y|^{n-2}} dy \leq \int_{U'} \frac{C_2}{r^{n-2}} r^{n-1} dr = \int_{U'} C_2 r dr$$
后者是有界的, 故 $\overline{u}$ 在 x 附近是良定义且有界的.
而当 x 离 y 足够远, 知对有界的 y 而言, $|x-y|$ 与 $|x|$ 关于 $|x|$ 同阶, 进而 $\Phi(x-y)$ 与 $\Phi(x)$ 分别作为 $|x-y|$ 与 $|x|$ 的函数, 也是同阶的, 故
$$|\overline{u}| \leq \int_U |\Phi(x-y)| \cdot |f(y)| dy \leq \int_U C_1 |\Phi(x)| \cdot |f(y)| dy \leq \frac{C_2}{|x|^{n-2}} \int_U |f(y)| dy$$
在 $|x|$ 足够大时, $\frac{C_2}{|x|}$ 与 $\int_U |f(y)| dy$ 都是有界的, 因而 $|\overline{u}| < \infty$.
综上即知 $\overline{u}$ 是 Poisson 方程的有界解.
- (关于 $|\alpha!| \leq n^{|\alpha|}\alpha!$ 的解释) 对于 $(x_1 + \cdots + x_m)^n$ 而言, 其展开后的单项式必形如 $x_1^{k_1} \cdots x_m^{k_m}$, 其中 $k_1 + \cdots + k_m = n$. 对于固定的幂指数指标 $(k_1, \cdots, k_m)$, 其对应的系数为 $C_n^{k_1} C_{n-k_1}^{k_2} \cdots C_{n-(k_1+\cdots+k_{m-1})}^{k_m}$,

[3]固定 x, 对 V 内的任意 y, 总是可以选出有限个 $y_1, y_2, \cdots$, 在 x, y 之间"搭一座桥", 其间每个点的距离也不超过 r(这是取覆盖球半径为 $\frac{r}{2}$ 的原因).

从而前述多项式展开式可表为:

$$(x_1+\cdots+x_m)^n=\sum_{\substack{(k_1,\cdots,k_m)\\ n=k_1+\cdots+k_m}} C_n^{k_1}C_{n-k_1}^{k_2}C_{n-(k_1+k_2)}^{k_3}\cdots C_{n-(k_1+\cdots+k_{m-1})}^{k_m}x_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$$

进一步计算组合数:

$$\begin{aligned} C_n^{k_1} &= \frac{n!}{k_1!(n-k_1)!} \\ C_{n-k_1}^{k_2} &= \frac{(n-k_1)!}{k_2!(n-(k_1+k_2))!} \\ &\vdots \\ C_{n-(k_1+\cdots+k_{m-1})}^{k_m} &= \frac{(n-(k_1+\cdots+k_{m-1}))!}{k_m!(n-(k_1+\cdots+k_m))!} \end{aligned}$$

故

$$C_{n-k_1}^{k_2}C_{n-(k_1+k_2)}^{k_3}\cdots C_{n-(k_1+\cdots+k_{m-1})}^{k_m}=\frac{n!}{k_1!k_2!\cdots k_m!}$$

这说明

$$(x_1+\cdots+x_m)^n=\sum_{\substack{(k_1,\cdots,k_m)\\ n=k_1+\cdots+k_m}}\frac{n!}{k_1!k_2!\cdots k_m!}x_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$$

代入 $x_1=\cdots=x_m=1$, 将 k_i 对应换成 $\alpha_i(1\le i\le m)$ 得:

$$(1+\cdots+1)^n=\sum_{\substack{(\alpha_1,\cdots,\alpha_m)\\ n=\alpha_1+\cdots+\alpha_m}}\frac{n!}{\alpha_1!\cdots\alpha_m!}$$

亦即

$$m^{|\alpha|}=\sum_{\alpha}\frac{|\alpha|!}{\alpha!}$$

最后将 m 换成 n, 对给定的某个 α 即得:

$$n^{|\alpha|}\ge\frac{|\alpha|!}{\alpha!}\Rightarrow|\alpha|!\le n^{|\alpha|}\alpha!$$

前述关于 Laplace 方程, Poisson 方程, 以及调和函数的诸多性质的阐述逻辑见图(7.5)

7.2.4 Green 函数

现在设 $U\subset\mathbb{R}^n$ 是有界开集, ∂U 是 C^1 的. 下面来推导 Poisson 方程

$$\begin{cases} -\Delta u=f, & 在U中 \\ u=g, & 在\partial U上 \end{cases}$$

的解的一般表示公式.

7.2.4.1 Green 函数的推导

任取 $u\in C^2(\overline{U})$, 固定 $x\in U$, 取 $\varepsilon>0$ 足够小使得 $B(x,\varepsilon)\subset U$. 在区域 $V_\varepsilon:=U\backslash B(x,\varepsilon)$ 中对 $u(y),\Phi(y-x)$ 应用 Green 公式 (iii) 有:

$$\int_{V_\varepsilon}u(y)\Delta\Phi(y-x)-\Phi(y-x)\Delta u(y)dy=\int_{\partial V_\varepsilon}u(y)\frac{\partial\Phi}{\partial\boldsymbol{\nu}}(y-x)-\Phi(y-x)\frac{\partial u}{\partial\boldsymbol{\nu}}(y)dS(y) \tag{7.43}$$

因为 Φ 本身是 Laplace 方程的解, 故 $x\ne y$ 时 $\Delta\Phi(x-y)=0$. 同时观察到

$$|\int_{\partial B(x,\varepsilon)}\Phi(y-x)\frac{\partial u}{\partial\boldsymbol{\nu}}(y)dS(y)|\le\max_{\partial B(0,\varepsilon)}|\Phi|\cdot\max_{\partial B(x,\varepsilon)}|\frac{\partial u}{\partial\boldsymbol{\nu}}|\cdot C\varepsilon^{n-1}=o(1),\quad\varepsilon\to0$$

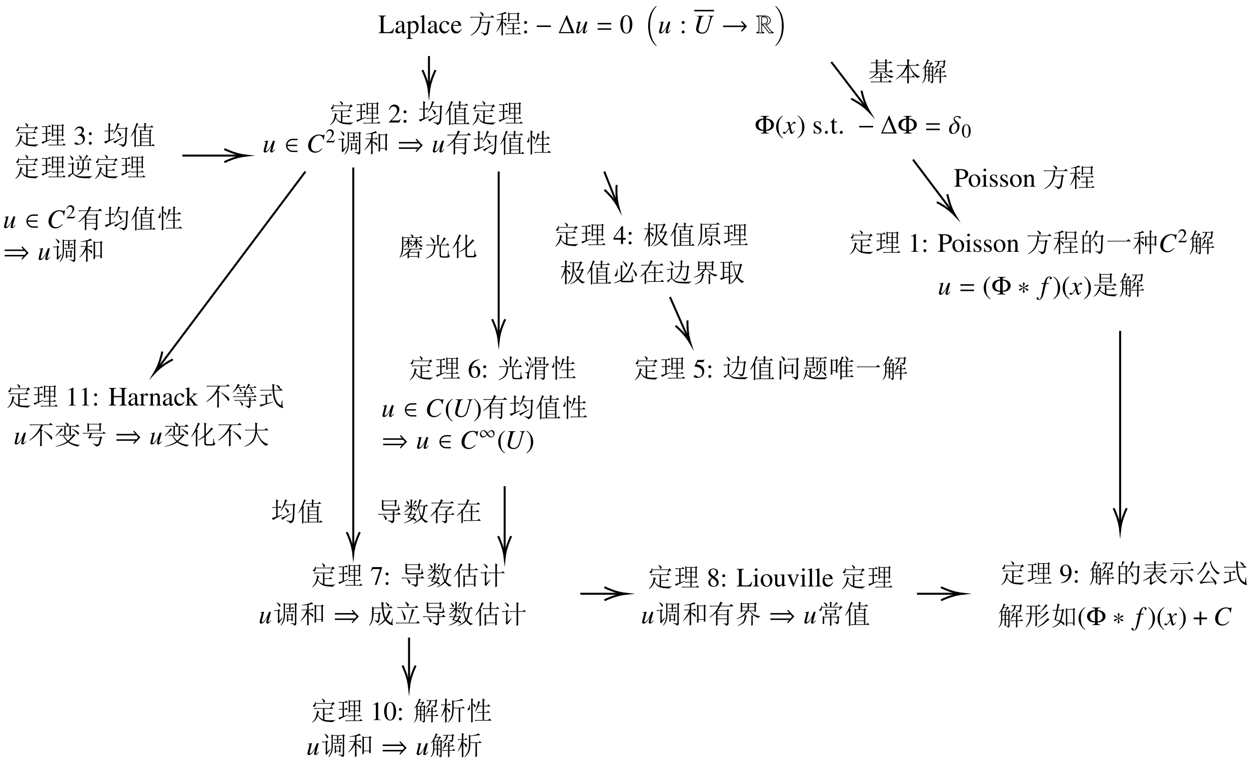


图 7.5

沿用(7.24)式有

$$\int_{\partial B(x,\varepsilon)} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x)dS(y) = ⨍_{\partial B(x,\varepsilon)} u(y)dS(y) \to u(x), \quad \varepsilon \to 0$$

现在在(7.43)式右端有:

$$\begin{aligned}
&\int_{\partial V_\varepsilon} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x) - \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) \\
=& \int_{\partial U} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x) - \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) + \int_{\partial B(x,\varepsilon)} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x) - \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) \\
\to& \int_{\partial U} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x) - \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) + u(x)
\end{aligned}$$

其中第二行 $\partial B(x,\varepsilon)$ 符号为正是因为在这个积分中, $\boldsymbol{\nu}$ 是 $\partial B(x,\varepsilon)$ 的单位外法向量, 这与 ∂V 的定向是一致的. 对左端有:

$$\begin{aligned}
\int_{V_\varepsilon} u(y)\Delta\Phi(y-x) - \Phi(y-x)\Delta u(y)dy &= -\int_U \Phi(y-x)\Delta u(y)dy + \int_{B(x,\varepsilon)} \Phi(y-x)\Delta u(y)dy \\
&\to -\int_U \Phi(y-x)\Delta u(y)dy, \quad \varepsilon \to 0
\end{aligned}$$

故

$$u(x) = \int_{\partial U} \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y) - u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x)dS(y) - \int_U \Phi(y-x)\Delta u(y)dy \tag{7.44}$$

这个等式对任意的 $u \in C^2(\overline{U}), x \in U$ 都是成立的.

现在如果知道 Δu 在 U 内的值, 以及 $u, \frac{\partial u}{\partial \nu}$ 在 ∂U 上的值, 就可以根据(7.44)式求出 u 了. 但在 Poisson 方程中, Δu 在 U 内的取值与 u 在 ∂U 上的取值确实知道, 而 $\frac{\partial u}{\partial \nu}$ 在 ∂U 上的取值仍然是未知的. 所以针对 Poisson 方程我们还要进一步修改(7.44)式.

修改的方法是对每个固定的 x 取一个修正函数 $\phi^x = \phi^x(y)$:

$$\begin{cases} \Delta\phi^x = 0, & 在U内 \\ \phi^x = \Phi(y-x), & 在\partial U上 \end{cases} \tag{7.45}$$

由 Green 公式(6.2.3)(iii) 知:

$$\begin{aligned} -\int_U \phi^x(y)\Delta u(y)dy &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \boldsymbol{\nu}}(y) - \phi^x(y)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) \\ &= \int_{\partial U} u(y)\frac{\partial \phi^x}{\partial \boldsymbol{\nu}}(y) - \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) \end{aligned} \tag{7.46}$$

注 ϕ^x 在这里的引入可能有点唐突, 下面是 [**OAO**] 的思路.

调和函数的表达式(7.44)已经给出:

$$u(x) = \int_{\partial U} \Phi(y-x)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) - \int_{\partial U} u(y)\frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y-x)dS(y) - \int_U \Phi(y-x)\Delta u(y)dy \tag{7.47}$$

现在的需求依旧是消除 $\frac{\partial u}{\partial \nu}(y)$ 的这一项. 先任取函数 $\phi(y) \in C^2(\overline{U})$, 由 Green 公式 (iii) 知:

$$\int_U (u\Delta\phi - \phi\Delta u)dy = \int_{\partial U} (u\frac{\partial \phi}{\partial \boldsymbol{\nu}} - \phi\frac{\partial u}{\partial \boldsymbol{\nu}})dS(y)$$

移项有

$$\int_U u(y)\Delta\phi(y)dy = -\int_{\partial U} \phi(y)\frac{\partial u}{\partial \boldsymbol{\nu}}(y)dS(y) + \int_{\partial U} u(y)\frac{\partial \phi}{\partial \boldsymbol{\nu}}(y)dS(y) + \int_U \phi(y)\Delta u(y)dy \tag{7.48}$$

比对(7.47),(7.48)两式, 为了消除等号右侧的第一项, 不妨就取 $\phi(y) = \Phi(y-x), y \in \partial U$. 另一方面, 为了将两式相加时保持等号左端依旧是 $u(x)$, 就取 $\Delta\phi(y) = 0, y \in U$. 这便得到了修正函数的方程(7.45). 如此相加的结果为:

$$u(x) = \int_{\partial U} u(y)(\frac{\partial \phi}{\partial \boldsymbol{\nu}}(y) - \frac{\partial \Phi}{\partial \boldsymbol{\nu}}(y))dS(y) + \int_U (\phi^x(y) - \Phi(y-x))\Delta u(y)dy \tag{7.49}$$

下面便可顺势引入 Green 函数.

定义 7.2.6 (Green 函数)

区域 U 的 Green 函数定义为

$$G(x,y) := \Phi(y-x) - \phi^x(y), \quad x, y \in U, x \neq y \tag{7.50}$$

将(7.50)式代入(7.49)式得

$$u(x) = -\int_{\partial U} u(y)\frac{\partial G}{\partial \boldsymbol{\nu}}(x,y)dS(y) - \int_U G(x,y)\Delta u(y)dy, \quad x \in U \tag{7.51}$$

其中 $\frac{\partial G}{\partial \nu}(x,y) = D_yG(x,y)\cdot\boldsymbol{\nu}(y)$ 是 G 关于变量 y 沿外法向量的方向导数. 现在(7.51)式中就不含 $\frac{\partial u}{\partial \nu}$ 了.

现设 $u \in C^2(\overline{U})$ 是边值问题

$$\begin{cases} -\Delta u = f, & 在U中 \\ u = g, & 在\partial U上 \end{cases} \tag{7.52}$$

的解, 其中 f, g 给定. 则将条件(7.52)代入(7.51)式即得:

定理 7.2.12 (使用了 Green 函数的 Possion 方程解的表示公式)

若 $u \in C^2(\overline{U})$ 是边值问题(7.52)的解, 则

$$u(x) = -\int_{\partial U} g(u)\frac{\partial G}{\partial \boldsymbol{\nu}}(x,y)dS(y) + \int_U f(y)G(x,y)dy, \quad x \in U \tag{7.53}$$

课堂笔记 (2023.9.21)

- (上课后新增的引出 Green 函数的想法) 前面调和函数的诸多性质与表示公式给出的仅仅是 Poisson 方程 $-\Delta u = f$ 的解, 现在要考虑的是如果圈定一个区域 U, 另外指定 u 在 ∂U 上的值 g, 又该如何表示解? 重申问题, 现在已知 Dirichlet 边值:

$$\begin{cases} -\Delta u = f, & U \\ u = g, & \partial U \end{cases}$$

希望求解 u. 因为 $-\Delta\Phi = \delta_0$, 故求解 u 正是求解 $-\int_U u(y)\Delta\Phi(x-y)dy$. 为了利用边界条件, 知:

$$-\int_U u(y)\Delta_y\Phi(x-y)dy = -\int_U \nabla_y\cdot(u(y)\nabla_y\Phi(x-y))dy + \int_U \nabla_y u(y)\cdot\nabla_y\Phi(x-y)dy$$

现在希望消掉 $\int_U \nabla_y u(y)\cdot\nabla_y\Phi(x-y)dy$, 注意到

$$-\int_U \Phi(x-y)\Delta_y u(y)dy = -\int_U \nabla_y\cdot(\Phi(x-y)\nabla_y u(y))dy + \int_U \nabla_y\Phi(x-y)\cdot\nabla_y u(y)dy$$

根据散度定理:

$$\int_\Omega \operatorname{div} f dx = \int_{\partial\Omega} f\cdot\boldsymbol{\nu} d\sigma$$

其中 $\boldsymbol{\nu} = (\nu^1,\cdots,\nu^n)$ 是 $\partial\Omega$ 上的单位外法向量. 代入 $f = u\nabla v$ 即得

$$\int_\Omega \operatorname{div}(u\cdot\nabla v)dx = \int_{\partial\Omega} u\nabla v\cdot\boldsymbol{\nu} d\sigma$$

而 $\nabla v\cdot\boldsymbol{\nu} = \frac{\partial v}{\partial\boldsymbol{\nu}}$, 故

$$\int_\Omega \operatorname{div}(u\cdot\nabla v)dx = \int_{\partial\Omega} u\frac{\partial v}{\partial\boldsymbol{\nu}} d\sigma$$

比对知

$$\int_U \nabla_y\cdot(u(y)\nabla_y\Phi(x-y))dy = \int_{\partial\Omega} u(y)\frac{\partial\Phi}{\partial\boldsymbol{\nu}}(x-y)dy$$
$$\int_U \nabla_y\cdot(\Phi(x-y)\nabla_y u(y))dy = \int_{\partial\Omega} \Phi(y)\frac{\partial u}{\partial\boldsymbol{\nu}}(x-y)dy$$

从而

$$\int_U (u(y)\Delta_y\Phi(x-y) - \Phi(x-y)\Delta_y u(y))dy = \int_{\partial U}(u(y)\frac{\partial\Phi}{\partial\boldsymbol{\nu}}(x-y) - \Phi(x-y)\frac{\partial u}{\partial\boldsymbol{\nu}}(y))dy$$

因为 $\Phi(x-y)$ 在 $y = x \in U$ 时有奇性, 故需将 U 分为 $B(x,\varepsilon)$ 与 $V_\varepsilon = U\backslash B(x,\varepsilon)$ 两部分讨论. 而在 Dirichlet 问题下, u 是待解函数, Φ 是已知函数, $\Delta_y u$ 与 $u|_{\partial U}$ 是给定条件, $\frac{\partial u}{\partial\boldsymbol{\nu}}$ 是未知量, 故剩下的任务就是估计已知部分并设法处理未知部分, 而这些内容正是前面所介绍的.

至此只要求解给定区域 U 的 Green 函数 G, 就可以解决边值问题(7.52)了. 但这在一般情形下很难做到, 确定 G 这件事也只能在 U 的几何结构很简单的时候才能进行. 下面就来讨论可以具体求出 G 的一些情况.

7.2.4.2 对 Green 函数的一些研究

固定 $x \in U$, 把 G 看成关于 y 的函数, 则由基本解的定义(7.2.4)知

$$\begin{cases} -\Delta G = \delta_x, & \text{在 } U \text{ 内} \\ G = 0, & \text{在 } \partial U \text{ 上} \end{cases} \tag{7.54}$$

定理 7.2.13 (Green 函数的对称性)

对任意的 $x, y \in U, x \neq y$, 有:

$$G(y,x) = G(x,y)$$

证明

固定 $x, y \in U, x \neq y$, 取

$$v(z) := G(x, z), \quad w(z) := G(y, z), \quad z \in U$$

则 $\Delta v(z) = 0(z \neq x), \Delta w(z) = 0(z \neq y)$, 且在 ∂U 上有 $w = v = 0$. 因而取足够小的 $\varepsilon > 0$, 在 $V := U \backslash (B(x, \varepsilon) \cup B(y, \varepsilon))$ 上应用 Green 公式(6.2.3)(iii) 有:

$$\begin{aligned} 0 = \int_U v\Delta w - w\Delta v dz &= (\int_{\partial U} - \int_{\partial B(x,\varepsilon)} - \int_{\partial B(y,\varepsilon)}) v\frac{\partial w}{\partial \boldsymbol{\nu}} - w\frac{\partial v}{\partial \boldsymbol{\nu}} dS(z) \\ &= -(\int_{\partial B(x,\varepsilon)} + \int_{\partial B(y,\varepsilon)}) v\frac{\partial w}{\partial \boldsymbol{\nu}} - w\frac{\partial v}{\partial \boldsymbol{\nu}} dS(z) \end{aligned}$$

得到

$$\int_{\partial B(x,\varepsilon)} \frac{\partial v}{\partial \boldsymbol{\nu}} w - \frac{\partial w}{\partial \boldsymbol{\nu}} v dS(z) = \int_{\partial B(y,\varepsilon)} \frac{\partial w}{\partial \boldsymbol{\nu}} v - \frac{\partial v}{\partial \boldsymbol{\nu}} w dS(z) \tag{7.55}$$

其中 $\boldsymbol{\nu}$ 是在 $\partial B(x, \varepsilon) \cup \partial B(y, \varepsilon)$ 上的内单位法向量. 因为 w 在 x 附近是光滑的, 故

$$|\int_{\partial B(x,\varepsilon)} \frac{\partial w}{\partial \boldsymbol{\nu}} v dS| \leq \sup_{\partial B(x,\varepsilon)} |v| \cdot \sup_{\partial B(x,\varepsilon)} |\frac{\partial w}{\partial \boldsymbol{\nu}}| \cdot C\varepsilon^{n-1} = o(1), \quad \varepsilon \to 0$$

另一方面, 知 $v(z) = \Phi(z - x) - \phi^x(z)$, 其中 ϕ^x 在 U 中光滑, 故根据(7.24)式可知

$$\lim_{\varepsilon \to 0} \int_{\partial B(x,\varepsilon)} \frac{\partial v}{\partial \boldsymbol{\nu}} w dS = \lim_{\varepsilon \to 0} \int_{\partial B(x,\varepsilon)} \frac{\partial \Phi}{\partial \boldsymbol{\nu}}(x - z) w(z) dS = w(x)$$

现在知(7.55)式左端在 $\varepsilon \to 0$ 时趋于 $w(x)$, 同理可知右端在 $\varepsilon \to 0$ 时趋于 $v(y)$, 故

$$G(y, x) = w(x) = v(y) = G(x, y).$$

□

Green 函数 $G(x, y)$ 也有一定的物理意义, 为此可以跑一下题, 先在热学上阐述一下 Laplace 方程具体的物理意义, 这对后面热方程的理解也是有帮助的. 下面的内容选自 [**Zo**].

设 U 是某个物体, ∂U 是这个物体的边界, 现在的目标是求 U 内每个点 (x, y, z) 在特定时刻 t 的温度: $T(x, y, z, t)$.

先考虑 U 内没有热源的情况, 此时 U 的内能只能通过边界的热交换来改变, 也即能量转移只发生在 ∂U 上. 中学物理已经学过, 要使得一个质量为 m 的均匀物体的温度升高 ΔT, 需要传递给它的热量为 $cm\Delta T$, 其中 c 是该物质的比热容. 这说明如果在时间间隔 Δt 内, 在 (x, y, z) 处 T 的增长量为 $\Delta T = T(x, y, z, t + \Delta t) - T(x, y, z, t)$, 那么 U 内的内能就增加

$$\iiint_U c\rho\Delta T dV \tag{7.56}$$

其中 ρ 是物体 U 的密度.

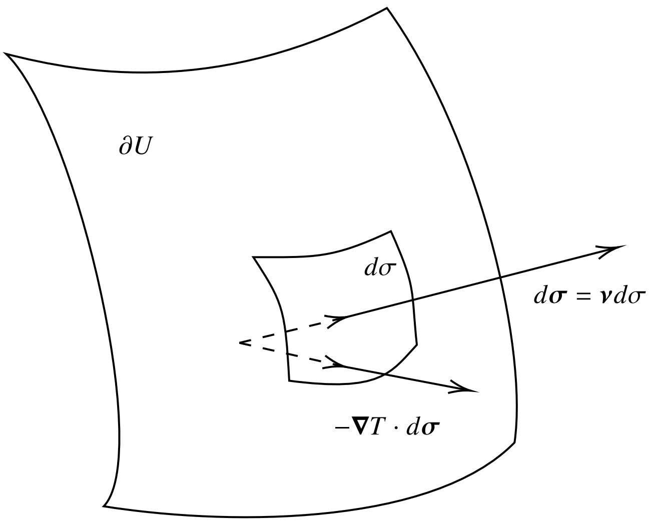


图 7.6

根据经验公式, 如果温度变化足够大, 则单位时间中经过 ∂U 上的一个小面元 $d\boldsymbol{\sigma} = \boldsymbol{\nu} d\sigma$ 的热量 (其中 $\boldsymbol{\nu}$ 是单位外法向量), 与向量场 $-\nabla T$ 通过这个小面元的流量 $-\nabla T \cdot d\boldsymbol{\sigma}$ 成正比 (其中梯度是关于 (x, y, z) 取的)(见图(7.6)). 比例系数 k 与 U 物体的性质有关, 称作热传导系数. ∇T 前面的负号是因为现在讨论的是流出的热量, 此时温度是内低外高, 因而正常的梯度方向是指向 U 内的. 现在, 在 Δt 时间内, 通过 ∂U 向外界沿外法向量 $\boldsymbol{\nu}$ 的方向逃出的热量可以近似 (具体来说是在 $o(\Delta t)$ 的精度上) 表为

$$\Delta t \iint_{\partial U} -k \nabla T \cdot d\boldsymbol{\sigma}$$

同理, 通过 ∂U 向 U 内沿内法向量 $\boldsymbol{\nu}$ 方向进入的热量就可以近似表为

$$\Delta t \iint_{\partial U} k \nabla T \cdot d\boldsymbol{\sigma} \tag{7.57}$$

根据" 在只有热交换的情况下, 内能的增加量就是传入物体的热量" 这一物理事实, 联立(7.56),(7.57)式有:

$$\iiint_U c\rho \Delta T dV = \Delta t \iint_{\partial U} k \nabla T \cdot d\boldsymbol{\sigma} \Rightarrow \iiint_U c\rho \frac{\Delta T}{\Delta t} dV = \iint_{\partial U} k \nabla T \cdot d\boldsymbol{\sigma}$$

令 $\Delta t \to 0$ 即得

$$\iiint_U c\rho \frac{\partial T}{\partial t} dV = \iint_{\partial U} k \nabla T \cdot d\boldsymbol{\sigma} \tag{7.58}$$

设 T 是足够光滑的函数, 并简便起见就设物体 U 的热传导是各向同性且均匀的, 此时 k 是常数, 进而对(7.58)右式应用高-奥公式(6.2.1)有

$$\iiint_U c\rho \frac{\partial T}{\partial t} dV = \iint_{\partial U} \operatorname{div}(k \nabla T) dV = \iint_{\partial U} k \Delta T dV \tag{7.59}$$

根据 U 的任意性得到

$$c\rho \frac{\partial T}{\partial t} = k \Delta T$$

但如果在 U 内还有热源 (或热汇), 其密度为 $F(x, y, z, t)$(即每单位时间都在 (x, y, z) 处产生热量 F), 则(7.59)式需改为

$$\iiint_U c\rho \frac{\partial T}{\partial t} dV = \iiint_U k \Delta T dV + \iiint_U F dV$$

相应地有

$$c\rho \frac{\partial T}{\partial t} = k \Delta T + F$$

注意到 $c, \rho, k > 0$ 后, 方程可以进一步简化为下述形式:

$$\frac{\partial T}{\partial t} = a^2 \Delta T + F \tag{7.60}$$

(7.60)式就称作热传导方程. 当热交换稳定时, T 与时间 t 是无关的, 因而 $\frac{\partial T}{\partial t} \equiv 0$, 得到 Poisson 方程

$$-\Delta T = f$$

在物体 U 内部无热源的情况下, $f \equiv 0$, 这便得到了 Laplace 方程

$$\Delta T = 0$$

在这样的背景下, 考察 Green 函数所满足的方程(7.54):

$$\begin{cases} -\Delta G = \delta_x, & \text{在 } U \text{ 内} \\ G = 0, & \text{在 } \partial U \text{ 上} \end{cases}$$

固定 x, $-G(x, y)$ 代表的就是在物体边界 ∂U 上温度等于零时, U 内部的温度分布, 其中固定点 x 处有一个产生热量为 1 的热源. 如果空间中仅仅存在这么一个热源, ∂U 上的温度当然不可能为 0, 所以要求解 Green 函数, 就是在问: 在物体 U 之外有没有一种热源分布, 使得在 ∂U 上这个热源分布与 δ_x 抗衡后, 达成的平衡是温度为 0?

前面这些关于热学的讨论也可以换成电学, 这时" 温度" 对应" 电势", " 热源" 对应" 正电荷", " 热量" 对应" 电量", δ_x 表示在 x 处有一个电量为 1 的点电荷. 现在问题就是: 在区域 U 之外有没有一种电荷分布, 使得在 ∂U 上电势为 0? 注意如果固定 x, Green 函数中的 $\Phi(y - x)$ 就已经起到了 δ_x 的作用了. 所以下面寻找 ϕ^x 的过程, 本质上

就是寻找这个在 U 之外未知的电荷分布的过程.

7.2.4.3 半空间的 Green 函数

考察半空间

$$\mathbb{R}^n_+ = \{x = (x_1, \cdots, x_n) \in \mathbb{R}^n : x_n > 0\}$$

尽管半空间的无界性导致前面对有界区域的讨论没法很好地在这里应用, 但它的 Green 函数的构造所需要的思想是类似的.

定义 7.2.7 (反射)

若 $x = (x_1, \cdots, x_{n-1}, x_n) \in \mathbb{R}^n_+$, 则它关于 $\partial\mathbb{R}^n_+$ 的反射是点

$$\tilde{x} = (x_1, \cdots, x_{n-1}, -x_n)$$

♣

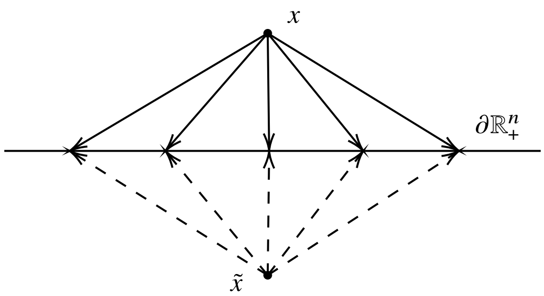


图 7.7

现在需要对半空间求解修正函数 ϕ^x. 从前面介绍的物理背景上看, 固定 $x \in \mathbb{R}^n_+$, 取 $\tilde{x}$ 是 x 关于 $\partial\mathbb{R}^n_+$ 的反射, 直觉上可以猜测: 只要在 $\tilde{x}$ 处安放一个电量为 1 的正电荷, $\partial\mathbb{R}^n_+$ 上的电势就恰为 0(如图(7.7)). 故取

$$\phi^x(y) := \Phi(y - \tilde{x}) = \Phi(y_1 - x_1, \cdots, y_{n-1} - x_{n-1}, y_n + x_n), \quad x, y \in \mathbb{R}^n_+$$

定义 7.2.8 (半空间 $\mathbb{R}^n_+$ 的 Green 函数)

半空间 $\mathbb{R}^n_+$ 的 Green 函数为:

$$G(x, y) := \Phi(y - x) - \Phi(y - \tilde{x}), \quad x, y \in \mathbb{R}^n_+, x \neq y$$

♣

注意 x 是固定的, 对 y_n 求偏导有:

$$G_{y_n}(x, y) = \Phi_{y_n}(y - x) - \Phi(y_n - \tilde{x}) = -\frac{1}{n\alpha(n)}[\frac{y_n - x_n}{|y - x|^n} - \frac{y_n + x_n}{|y - \tilde{x}|^n}]$$

故若 $y \in \partial\mathbb{R}^n_+$, 此时 $y_n = 0$, 有[4]:

$$\frac{\partial G}{\partial \boldsymbol{\nu}}(x, y) = -G_{y_n}(x, y) = -\frac{2x_n}{n\alpha(n)} \frac{1}{|x - y|^n}$$

现设 u 是边值问题

$$\begin{cases} \Delta u = 0, & \text{在 } \mathbb{R}^n_+ \text{ 中} \\ u = g, & \text{在 } \partial\mathbb{R}^n_+ \text{ 上} \end{cases} \tag{7.61}$$

的解, 则根据使用 Green 函数的解的表达式(7.53), 代入 $f \equiv 0, g = g$ 与 $G, \frac{\partial G}{\partial \nu}$, 有:

$$u(x) = \frac{2x_n}{n\alpha(n)} \int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x - y|^n} dy, \quad x \in \mathbb{R}^n_+ \tag{7.62}$$

[4] $\frac{\partial G}{\partial \nu} = -G_{y_n}$ 是因为如果把 $\frac{\partial G}{\partial \nu}$ 写成 $DG \cdot \boldsymbol{\nu}$, 会发现因为现在的曲面是 $\partial\mathbb{R}^n_+$, 所以外法向量就是 $(0, \cdots, 0, -1)$, 故 $DG \cdot \boldsymbol{\nu} = -G_{y_n}$.

观察(7.62)式的结构可知, 如果设

$$K(x,y):=\frac{2x_n}{n\alpha(n)}\frac{1}{|x-y|^n},\quad x\in\mathbb{R}^n_+, y\in\partial\mathbb{R}^n_+$$

则可将(7.62)式写成:

$$u(x)=\int_{\partial\mathbb{R}^n_+}g(y)K(x,y)dy,\quad x\in\mathbb{R}^n_+ \tag{7.63}$$

称 $K(x,y)$ 为 $\mathbb{R}^n_+$ 的 Poisson 核, 而(7.63)式为 Poisson 公式. 下面验证 Poisson 公式(7.63)确实是边值问题(7.61)的解.

定理 7.2.14 (半平面的 Poisson 公式)

若 $g\in C(\mathbb{R}^{n-1})\cap L^\infty(\mathbb{R}^{n-1})$, 且 u 由 Poisson 公式(7.63)给出, 则:

(i) $u\in C^\infty(\mathbb{R}^n_+)\cap L^\infty(\mathbb{R}^n_+)$,

(ii) 在 $\mathbb{R}^n_+$ 中有 $\Delta u=0$,

(iii) 对 $\partial\mathbb{R}^n_+$ 上的点 x^0 有 $\lim\limits_{\substack{x\to x^0\\ x\in\mathbb{R}^n_+}}u(x)=g(x^0)$.

证明

对任意固定的 x, 映射 $y\mapsto G(x,y)$ 在 $y=x$ 之外是调和映射 (也即 $\Delta_y G(x,y)=0$). 根据 Green 函数的对称性 $G(x,y)=G(y,x)$, 知映射 $x\mapsto G(x,y)$ 在 $x=y$ 之外也是调和映射. 故 $x\mapsto -\frac{\partial G}{\partial\nu}(x,y)=K(x,y)$ 对 $x\in\mathbb{R}^n_+, y\in\partial\mathbb{R}^n_+$ 是调和映射. 注意 Poisson 核满足:

$$\int_{\partial\mathbb{R}^n_+}K(x,y)dy=\int_{\partial\mathbb{R}^n_+}(-\frac{\partial G}{\partial\boldsymbol{\nu}}(x,y))dS(y)=\int_{\mathbb{R}^n_+}(-\Delta G(x,y))dy=\langle\delta_x,1\rangle=1,\quad\forall x\in\mathbb{R}^n_+ \tag{7.64}$$

因为 g 有界, 故

$$|u(x)|\le\sup_{\partial\mathbb{R}^n_+}|g(x)|\cdot|\int_{\partial\mathbb{R}^n_+}K(x,y)dy|=\sup_{\partial\mathbb{R}^n_+}|g(x)|$$

因而 u 也是有界的. 因为 $x\mapsto K(x,y)$ 是光滑的, 故 $u(x)$ 也是光滑的, 也即 $u\in C^\infty(\mathbb{R}^n_+)$, 且

$$\Delta u(x)=\int_{\partial\mathbb{R}^n_+}\Delta_x K(x,y)g(y)dy=0,\quad x\in\mathbb{R}^n_+$$

现只需证明 (iii), 固定 $x^0\in\partial\mathbb{R}^n_+,\varepsilon>0$, 根据 g 连续知可选取足够小的 $\delta>0$ 使得

$$\forall y\in\partial\mathbb{R}^n_+(|y-x^0|<\delta\Rightarrow|g(y)-g(x^0)|<\varepsilon) \tag{7.65}$$

现若 $|x-x^0|<\frac{\delta}{2}, x\in\mathbb{R}^n_+$, 则

$$\begin{aligned}|u(x)-g(x^0)|&=|\int_{\partial\mathbb{R}^n_+}K(x,y)[g(y)-g(x^0)]dy|\\&\le\int_{\partial\mathbb{R}^n_+\cap B(x^0,\delta)}K(x,y)|g(y)-g(x^0)|dy+\int_{\partial\mathbb{R}^n_+\setminus B(x^0,\delta)}K(x,y)|g(y)-g(x^0)|dy\\&=:I+J\end{aligned} \tag{7.66}$$

对 I, 由(7.64),(7.65)式知

$$|I|\le\varepsilon\int_{\partial\mathbb{R}^n_+}K(x,y)dy=\varepsilon \tag{7.67}$$

对 J, 注意当 $|x-x^0|\le\frac{\delta}{2}, |y-x^0|\ge\delta$ 时, 有

$$|y-x^0|\le|y-x|+|x-x^0|\le|y-x|+\frac{\delta}{2}\le|y-x|+\frac{1}{2}|y-x^0|$$

得到 $|y-x|\geq \frac{1}{2}|y-x^0|$, 因而

$$\begin{aligned}J &\leq 2\|g\|_{L^\infty}\int_{\partial\mathbb{R}^n_+\backslash B(x^0,\delta)} K(x,y)dy\\ &= 2\|g\|_{L^\infty}\int_{\partial\mathbb{R}^n_+\backslash B(x^0,\delta)} \frac{2x_n}{n\alpha(n)}\frac{1}{|x-y|^n}dy\\ &\leq 2\|g\|_{L^\infty}\int_{\partial\mathbb{R}^n_+\backslash B(x^0,\delta)} \frac{2x_n}{n\alpha(n)}\frac{2^n}{|y-x^0|^n}dy\\ &= x_n\cdot\frac{2^{n+2}\|g\|_{L^\infty}}{n\alpha(n)}\cdot\int_{\partial\mathbb{R}^n_+\backslash B(x^0,\delta)}|y-x^0|^{-n}dy\to 0,\quad x_n\to 0^+\end{aligned} \tag{7.68}$$

将(7.67),(7.68)式的结果代入(7.66)式即证命题. □

课堂笔记 (2023.9.25)

- (使用 Green 函数的解表达式的一些注记) 对表达式

$$u(x)=-\int_{\partial U} g(y)\frac{\partial G}{\partial \boldsymbol{\nu}}(x,y)dS(y)+\int_U f(y)G(x,y)dy$$

一般记 $-\frac{\partial G}{\partial \nu}(x,y)=:K(x,y)$, 称 $K(x,y)$ 为 Poisson 核. 上述式子可另写成:

$$u(x)=\int_{\partial U} g(y)K(x,y)dS(y)+\int_U f(y)G(x,y)dy$$

当 $f=0$ 时, u 所满足的方程即:

$$\begin{cases}-\Delta u=0, & U\\ u=g, & \partial U\end{cases}$$

从而 u 可视作 ∂U 上的函数 g 在 U 中的调和延拓. 更一般地, 若将 Laplace 算子改成分数阶 Laplace 算子:

$$\begin{cases}(-\Delta)^{\frac{\alpha}{2}}u=0, & U\\ u=g, & \partial U\end{cases}$$

就称 u 是 g 在 U 上的 α 阶调和延拓.

- (关于 Green 函数对称性研究的动机与一些算子表示) 由基本解 Φ 与修正子 ϕ^x 分别满足的方程:

$$\begin{cases}-\Delta_y\Phi(y-x)=\delta_0(y-x), & U\\ \Phi(y-x), & \partial U\end{cases}\qquad \begin{cases}-\Delta_y\phi^x(y)=0, & U\\ \phi^x(y)=\Phi(y-x), & \partial U\end{cases}$$

可得 Green 函数 $G(x,y)=\Phi(y-x)-\phi^x(y)$ 满足的方程:

$$\begin{cases}-\Delta_y G(x,y)=\delta_0(y-x), & U\\ G(x,y)=0, & \partial U\end{cases}$$

对于零边值 Dirichlet 问题

$$\begin{cases}-\Delta_x F(x)=f(x), & U\\ F(x)=0, & \partial U\end{cases}$$

可记 $F(x)=(-\Delta_D)^{-1}f(x)$, 从而有 $G(x,y)=(-\Delta_D)^{-1}\delta_0(y-x)$. 现在对于 Laplace 方程初边值问题:

$$\begin{cases}-\Delta u=f, & U\\ u=g, & \partial U\end{cases}$$

分别设 v,w 满足:

$$\begin{cases}-\Delta v=f, & U\\ v=0, & \partial U\end{cases}\qquad \begin{cases}-\Delta w=0, & U\\ w=g, & \partial U\end{cases}$$

由线性方程叠加原理知 $u=v+w$. 根据使用 Green 函数的解表达式可知:

$$v(x)=\int_U G(x,y)f(y)dy$$
$$w(x)=\int_{\partial U} K(x,y)g(y)dy$$

根据前面介绍的记号, 可记 $v(x)=(-\Delta_D)^{-1}f(x)$. 而根据 δ_0 的定义即知:

$$f(x)=\int_U \delta_0(x-y)d(y)dy$$

从而

$$\begin{aligned}v(x)&=(-\Delta_D)^{-1}\int_U \delta_0(x-y)f(y)dy\\&=\int_U(-\Delta_D)^{-1}\delta_0(x-y)f(y)dy\\&=\int_U G(y,x)f(y)dy\end{aligned}$$

这便要求

$$\int_U G(x,y)f(y)dy=\int_U G(y,x)f(y)dy$$

进而希望 $G(x,y)=G(y,x)$.

- (半平面 Poisson 公式的一些拓展) 半平面 Poisson 公式(7.2.14)意味着 u 正是 $\mathbb{R}^{n-1}$ 上的函数 g 在 $\mathbb{R}^n$ 中的调和延拓, 而其 (iii) 正是 Poisson 核 $K(x,y)$ 的恒等逼近性质, 进而要说明 (iii) 就是要说明 $K(x,y)$ 是好核 (恒等逼近元). 对于关于 $y\in\partial\mathbb{R}^n$ 的函数族 $K(x,y)$ 而言, 其作为好核需要满足下述三个条件:
  (1) $\int_{\partial\mathbb{R}^n_+}K(x,y)dy=1$;
  (2) $\int_{\partial\mathbb{R}^n_+}|K(x,y)|dy<\infty$;
  (3) $\forall\varepsilon>0\exists\delta>0(\int_{|y|>\delta}K(x,y)dy<\varepsilon)$.

  (1) 在已知 $u(x)=\int_{\partial\mathbb{R}^n_+}g(y)K(x,y)dy$ 时容易理解的, 设 $g(y)\equiv 1$, 因为 $g(y)$ 是 u 作为调和函数的边值, 故由极值原理知 u 必恒等于 1, 即得 $\int_{\partial\mathbb{R}^n_+}K(x,y)dy=1$. 而若从数学分析的角度来说明, 设 $\mathbb{R}^n_+\ni x=(x^1,\cdots,x^{n-1},x^n),\partial\mathbb{R}^n_+\ni y=(y^1,\cdots,y^{n-1},0),x'=(x^1,\cdots,x^{n-1}),y'=(y^1,\cdots,y^{n-1})$, 则:

$$K(x,y)=\frac{2x_n}{n\alpha_n}\frac{1}{|x-y|^n}=\frac{2x^n}{n\alpha_n}\frac{1}{(|x'-y'|^2+(x^n)^2)^{\frac{n}{2}}}=(x^n)^{-(n-1)}\widetilde{K}(\frac{|x'-y'|}{x^n})$$

  其中

$$\widetilde{K}(r)=\frac{2}{n\alpha(n)}\frac{1}{(r^2+1)^{\frac{n}{2}}}$$

  现在用极坐标换元可得:

$$\int_{\partial\mathbb{R}^n_+}K(x,y)dy=\frac{2}{n\alpha_n}(n-1)\alpha_{n-1}\int_0^{+\infty}\int_0^{+\infty}\frac{r^{n-2}}{(r^2+1)^{\frac{n}{2}}}dr$$

  最后由 B 积分的性质即得 $\int_{\partial\mathbb{R}^n_+}K(x,y)dy=1$.
  事实上, 如果绕过数学分析这些基础却繁冗的表达, 只从广义函数的角度上看, $K(x,y)$ 是活在 $\partial\mathbb{R}^n_+$ 上, 以 x_n 为参数, 定义为 $-\frac{\partial G}{\partial \nu}(x,y)$ 的函数族. 现在类似于(7.64)式的想法, 如果想研究 $K(x,y)$ 在 $\partial\mathbb{R}^n_+$ 的性质, 根据 Green 公式总能把该性质转化为 $-\Delta G(x,y)$ 在 $\mathbb{R}^n_+$ 上的相关性质, 但 $-\Delta G(x,y)$ 在 $\mathbb{R}^n_+$ 上就是 δ_x, 所以在谈论 $K(x,y)$ 作为核函数的性质时, 总能借用 δ_x 的相关性质. 至于 $-\Delta G(x,y)=\delta_x(U)$, 这又是源于基本解满足的方程. 在 **[OAO]** 中是直接将满足对应广义函数方程的解称为方程的基本解的, 所以"利用 δ_x 证明性质"这件事的合理性是落实到定义上, 随时可用的, 而非某种特殊的巧合.

注 半平面的 Poisson 公式还能写成

$$u(x,t)=\int_{\mathbb{R}^n_+}\frac{2t}{n\alpha_n((x-y)^2+t^2)^{\frac{n}{2}}}g(y)dy$$

两个公式由 Fubini 定理即可互换.

7.2.4.4 球的 Green 函数

定义 7.2.9 (反演)

若 $x \in \mathbb{R}^n \backslash \{0\}$, 则点

$$\tilde{x} = \frac{x}{|x|^2}$$

称为 x 关于 $\partial B(0,1)$ 的反演点, 映射 $x \mapsto \tilde{x}$ 称为关于 $\partial B(0,1)$ 的反演.

下面用球面的反演计算单位球 $U = B^0(0,1)$ 的 Green 函数. 固定 $x \in B^\circ(0,1)$, 注意我们要解的方程为

$$\begin{cases} \Delta \phi^x = 0, & \text{在 } B^\circ(0,1) \text{ 中} \\ \phi^x = \Phi(y-x), & \text{在 } \partial B(0,1) \text{ 上} \end{cases} \tag{7.69}$$

得到解后 Green 函数就能被表成

$$G(x,y) = \Phi(y-x) - \phi^x(y) \tag{7.70}$$

从前面介绍的物理意义的角度出发[5], 这相当于是给定单位球内的某个单位正电荷, 要求在球外找到一个负电荷 (也即寻找 $-\phi^x$), 使得单位球面上的电势为 0.

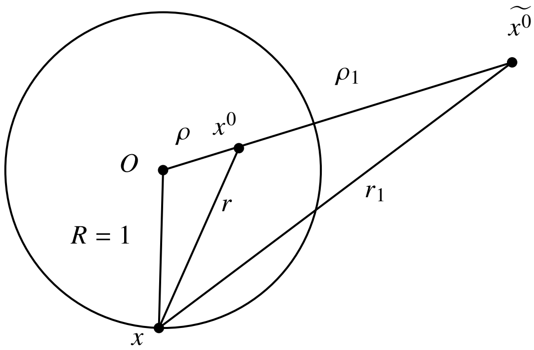


图 7.8

取固定点为 x^0, 其反演点为 $\widetilde{x^0}$, 任取 $\partial B(0,1)$ 上的一点 x. 设 $\rho = |x^0|, \rho_1 = |\widetilde{x^0}|, r = |x - x^0|, r_1 = |x - \widetilde{x^0}|$. 根据反演点的定义可知 $\rho\rho_1 = 1$, 也即若取 $R = 1$ 是球面半径, 则

$$\frac{|x^0|}{R} = \frac{R}{|\widetilde{x^0}|}$$

又因为 $\triangle Oxx^0$ 与 $\triangle O\widetilde{x^0}x$ 有公共角 $\angle O$, 故这两个三角形相似, 因而

$$\frac{\rho}{R} = \frac{R}{\rho_1} = \frac{r}{r_1}$$

得到 $r = \frac{\rho}{R} r_1$. 在 $n \geq 3$ 的情况下, 现在 x^0 在 x 处产生的电势为

$$\Phi(x - x^0) = \frac{1}{n(n-2)\alpha(n)} \frac{1}{r^{n-2}}$$

要想依靠 $\widetilde{x^0}$ 来抵消这个电势, 就是求解 $\widetilde{x^0}$ 的电量 q, 使得

$$q\Phi(x - \widetilde{x^0}) = \frac{1}{n(n-2)\alpha(n)} \frac{q}{r_1^{n-2}} = \frac{1}{n(n-2)\alpha(n)} \frac{1}{r^{n-2}}$$

得到

$$q = \frac{r_1^{n-2}}{r^{n-2}} = (\frac{R}{\rho})^{n-2} = \frac{1}{|x^0|^{n-2}}$$

[5] **[Ev]** 在这一块的介绍感觉有些含糊, 缺乏直观性, 所以还是选择了 **[OAO]** 的引入方式.

故

$$q\Phi(x-\widetilde{x^0})=\frac{1}{n(n-2)\alpha(n)}\frac{1}{|x^0|^{n-2}|x-\widetilde{x^0}|}=\Phi(|x^0|(x-\widetilde{x^0}))$$

现在回到原问题, 就可以设 $\phi^x(y):=\Phi(|x|(y-\tilde{x}))$. 可以验证 $\phi^x(y)$ 在 $U=B^\circ(0,1)$ 内是调和函数, 而当 $y\in\partial B(0,1), x\neq 0$ 时有:

$$|x^2||y-\tilde{x}|^2=|x|^2(|y|^2-\frac{2y\cdot x}{|x|^2}+\frac{1}{|x|^2})=|x|^2-2y\cdot x+1=|x-y|^2$$

故 $(|x||y-\tilde{x}|)^{-(n-2)}=|x-y|^{-(n-2)}$, 也即

$$\phi^x(y)=\Phi(y-x),\quad y\in\partial B(0,1)$$

这便验证了 ϕ^x 确实满足方程(7.69). $n=2$ 的情况是一样的.

定义 7.2.10 (球的 Green 函数)

单位球的 Green 函数为

$$G(x,y):=\Phi(y-x)-\Phi(|x|(y-\tilde{x})),\quad x,y\in B(0,1),x\neq y \tag{7.71}$$

♣

现在设 u 就是边值问题

$$\begin{cases}\Delta u=0, & \text{在 } B^\circ(0,1) \text{ 中}\\ u=g, & \text{在 } \partial B(0,1) \text{ 上}\end{cases} \tag{7.72}$$

的解, 将条件代入解的表示公式(7.53)有

$$u(x)=-\int_{\partial B(0,1)}g(y)\frac{\partial G}{\partial \boldsymbol{\nu}}(x,y)dS(y) \tag{7.73}$$

代入(7.71)式知

$$G_{y_i}(x,y)=\Phi_{y_i}(y-x)-\frac{\partial\Phi}{\partial y_i}(|x|(y-\tilde{x}))$$

计算有

$$\Phi_{y_i}(y-x)=\frac{1}{n\alpha(n)}\frac{x_i-y_i}{|x-y|^n}$$

且当 $y\in\partial B(0,1)$ 时有

$$\begin{aligned}\frac{\partial\Phi}{\partial y_i}(|x|(y-\tilde{x}))&=-\frac{1}{n\alpha(n)}\frac{y_i-\tilde{x}_i}{|x|^{n-2}|y-\tilde{x}|^n}=-\frac{1}{n\alpha(n)}\frac{y_i-\frac{x_i}{|x_i|^2}}{|x|^{n-2}|y-\tilde{x}|^n}\\&=-\frac{1}{n\alpha(n)}\frac{y_i|x_i|^2-x_i}{(|x||y-\tilde{x}|)^n}=-\frac{1}{n\alpha(n)}\frac{y_i|x|^2-x_i}{|x-y|^n}\end{aligned}$$

其中最后一步 $|x||y-\tilde{x}|=|x-y|$ 可以根据前面的物理意义分析中 $\frac{\rho}{R}=\frac{r}{r_1}$ 的关系式得到, 这里 $\rho=|x|, R=1, r=|x-y|, r_1=|y-\tilde{x}|$. 故

$$\begin{aligned}\frac{\partial G}{\partial \boldsymbol{\nu}}(x,y)&=\sum_{i=1}^n y_iG_{y_i}(x,y)=\sum_{i=1}^n y_i(\frac{1}{n\alpha(n)}\frac{x_i-y_i}{|x-y|^n}-\frac{1}{n\alpha(n)}\frac{y_i|x|^2-x_i}{|x-y|^n})\\&=-\frac{1}{n\alpha(n)}\frac{1}{|x-y|^n}\sum_{i=1}^n y_i((y_i-x_i)-y_i|x|^2+x_i)=-\frac{1}{n\alpha(n)}\frac{1-|x|^2}{|x-y|^n}\end{aligned} \tag{7.74}$$

将(7.74)式代入(7.73)式得

$$u(x)=\frac{1-|x|^2}{n\alpha(n)}\int_{\partial B(0,1)}\frac{g(y)}{|x-y|^n}dS(y)$$

如果现在球的半径不再是 1 而是 $r(r>0)$, u 是边值问题

$$\begin{cases}\Delta u=0, & \text{在 } B^\circ(0,r) \text{ 中}\\ u=g, & \text{在 } \partial B(0,r) \text{ 上}\end{cases} \tag{7.75}$$

的解, 则 $\tilde{u}(x) = u(rx)$ 就是单位球边值问题(7.72)的解, 其中 $\tilde{g}(x) = g(rx)$. 类似于前述过程可得

$$u(x) = \frac{r^2 - |x|^2}{n\alpha(n)r} \int_{\partial B(0,r)} \frac{g(y)}{|x-y|^n} dS(y), \quad x \in B^\circ(0,r) \tag{7.76}$$

这便是 Poisson 公式, 而函数

$$K(x,y) := \frac{r^2 - |x|^2}{n\alpha(n)r} \frac{1}{|x-y|^n}, \quad x \in B^\circ(0,r), y \in \partial B(0,r)$$

就称为球 $B(0,r)$ 的 Poisson 核.

注意(7.76)式是基于边值问题(7.75)有光滑解这一假设得到的, 下面声明这个式子就是边值问题的解.

定理 7.2.15 (球的 Poisson 公式)

若 $g \in C(\partial B(0,r))$, 且 u 由 Poisson 公式(7.76)给出, 则:

(i) $u \in C^\infty(B^\circ(0,r))$,

(ii) 在 $B^\circ(0,r)$ 中有 $\Delta u = 0$,

(iii) 对 $\partial B(0,r)$ 上的点 x^0 有 $\lim\limits_{\substack{x\to x^0\\ x\in B^0(0,r)}} u(x) = g(x^0)$.

♡

证明方法与半平面的情况是类似的. □

7.2.5 能量方法

到目前为止, 对调和函数的大多数分析都是依赖于特定的表达式, 比如基本解, Green 函数等. 本节介绍" 能量" 方法, 也即运用一些表达式的 L^2 范数的方法.

7.2.5.1 唯一性

首先考虑边值问题

$$\begin{cases} -\Delta u = f, & \text{在 } U \text{ 内} \\ u = g, & \text{在 } \partial U \text{ 上} \end{cases} \tag{7.77}$$

前面已经介绍过用最大值原理证明边值问题解的唯一性, 这里展示另一种简单的证明. 设 U 是有界开集, ∂U 是 C^1 的.

定理 7.2.16 (唯一性)

边值问题(7.77)至多只有一个解 $u \in C^2(\overline{U})$.

♡

证明

设 $\tilde{u}$ 是边值问题(7.77)的另一个解, 并设 $w := u - \tilde{u}$, 则在 U 内有 $\Delta w = 0$, 且在 ∂U 上有 $w = 0$, 进而根据分部积分法(6.2.3)(ii) 有

$$0 = -\int_U w\Delta w dx = \int_U Dw \cdot Dw dx - \int_{\partial U} \frac{\partial w}{\partial \boldsymbol{\nu}} w dS = \int_U |Dw|^2 dx$$

故在 U 内有 $Dw \equiv 0$, 又因为在 ∂U 上也有 $w = 0$, 故在 U 内有 $w = u - \tilde{u} \equiv 0$. □

7.2.5.2 Dirichlet 原理

下面, 我们来说明 Poisson 方程的边值问题(7.77)的解可以被当成是某个合适的泛函的最小值点. 要说明这一点, 可以定义能量泛函:

$$I[w] := \int_U \frac{1}{2}|Dw|^2 - wf dx$$

其中 w 从属于容许集

$$\mathcal{A} := \{w \in C^2(\overline{U}) : \text{在 } \partial U \text{ 上有} w = g\}$$

定理 7.2.17 (Dirichlet 原理 (变分原理))

若 $u \in C^2(\overline{U})$ 满足 Poisson 方程的边值问题(7.77), 则

$$I[u] = \min_{w \in \mathcal{A}} I[w] \tag{7.78}$$

相反地, 如果 $u \in \mathcal{A}$ 满足(7.78)式, 则它就是 Poisson 方程边值问题(7.77)的解.

也就是说, 如果 $u \in \mathcal{A}$, 则 PDE: $-\Delta u = f$ 就等价于 u 是能量 $I[\cdot]$ 的最小值点.

证明

如果 u 是边值问题(7.77)的解, 选定 $w \in \mathcal{A}$, 根据边值问题(7.77)知在 U 中有 $-\Delta u - f = 0$, 因而

$$\int_U (-\Delta u - f)(u - w)dx = 0$$

由 Green 公式(6.2.3)(ii) 与在 ∂U 上 $u - w = 0$ 知

$$\int_U (-\Delta u) \cdot (u - w)dx = \int_U Du \cdot D(u - w)dx - \int_{\partial U} \frac{\partial u}{\partial \boldsymbol{\nu}} \cdot (u - w)dS = \int_U Du \cdot D(u - w)dx$$

故

$$\int_U Du \cdot D(u - w) - f \cdot (u - w)dx = 0$$

现在式子中已经没有边值了. 整理可得:

$$\int_U |Du|^2 - uf dx = \int_U Du \cdot Dw - wf dx$$

根据 Cauchy-Schwarz 不等式有:

$$|Du \cdot Dw| \leq |Du| \cdot |Dw| \leq \frac{1}{2}|Du|^2 + \frac{1}{2}|Dw|^2$$

得到

$$\int_U |Du|^2 - uf dx \leq \int_U \frac{1}{2}|Du|^2 dx + \int_U \frac{1}{2}|Dw|^2 - wf dx$$

也即

$$\int_U \frac{1}{2}|Du|^2 - uf dx \leq \int_U \frac{1}{2}|Dw|^2 - wf dx$$

此即

$$I[u] \leq I[w], \quad w \in \mathcal{A} \tag{7.79}$$

因为 $u \in \mathcal{A}$, 故(7.78)式得证.

相反地, 如果 $u \in \mathcal{A}$ 满足(7.78)式成立, 任意固定 $v \in C_c^\infty(U)$, 记

$$i(\tau) := I[u + \tau v], \quad \tau \in \mathbb{R}$$

根据 v 的紧支性, 对任意 τ 都有 $u + \tau v \in \mathcal{A}$, 再根据(7.78)式知标量函数 $i(\cdot)$ 在 0 处取最小值, 这说明只要 i 可导, 就有

$$i'(0) = \frac{di}{d\tau}(0) = 0$$

但与此同时

$$\begin{aligned} i(\tau) &= \int_U \frac{1}{2}|Du + \tau Dv|^2 - (u + \tau v)f dx \\ &= \int_U \frac{1}{2}|Du|^2 + \tau Du \cdot Dv + \frac{\tau^2}{2}|Dv|^2 - (u + \tau v)f dx \end{aligned}$$

故根据 Green 公式(6.2.3)(ii) 有

$$0 = i'(0) = \int_U Du \cdot Dv - vf dx = \int_U (-\Delta u - f)v dx$$

因为 $v \in C_c^\infty(U)$ 是任意取的, 故在 U 中有 $-\Delta u = f$, 命题即证. □

7.3 热方程

下面我们在适当的初值与边值条件下研究热方程

$$u_t - \Delta u = 0 \tag{7.80}$$

与非齐次热方程

$$u_t - \Delta u = f \tag{7.81}$$

其中 $t > 0, x \in U$, $U \subset \mathbb{R}^n$ 是开的. $u = u(x,t) : \overline{U} \times [0,\infty) \to \mathbb{R}$ 是未知函数, Laplace 算子 Δ 是对空间变量 $x = (x_1, \cdots, x_n)$ 取的: $\Delta u = \Delta_x u = \sum_{i=1}^n u_{x_i x_i}$. 在非齐次方程(7.81)中函数 $f : U \times [0,\infty) \to \mathbb{R}$ 是已知的.

在前面讨论 Laplace 方程的物理意义时已经可以看出, Laplace 方程就是一类特殊的热方程, 这也表明调和函数的相关断言在热方程的解中有类似结论, 所以对热方程解性质的介绍与调和函数是类似的.

热方程又称作扩散方程, 它描述的是随着时间的推移, 某些物理量 (比如热量, 化学物质浓度等等) 的密度的变化情况. 如果 $V \subset U$ 是一个光滑子区域, 则 V 内物质总量的变化速率就等于经过 ∂V 的物质净通量的倒数:

$$\frac{d}{dt}\int_V u dx = -\int_{\partial V} \boldsymbol{F} \cdot \boldsymbol{\nu} dS$$

其中 $\boldsymbol{F}$ 是流密度, 进而根据高-奥公式与 V 的任意性有

$$u_t = -\operatorname{div} \boldsymbol{F} \tag{7.82}$$

在很多情况下, $\boldsymbol{F}$ 与 ∇u 是成正比但方向相反的 (前面提过, 这是因为流密度的方向是从高到低流):

$$\boldsymbol{F} = -aDu, \quad a > 0$$

把该式代回(7.82)式, 可得 PDE

$$u_t = a \operatorname{div}(Du) = a\Delta u$$

令 $a = 1$ 即得热方程. 热方程同样会出现在布朗运动的研究中.

7.3.1 基本解

7.3.1.1 基本解的推导

观察知热方程中包含了一个关于时间 t 的一阶偏导, 但对空间变量 $x_i (i = 1, \cdots, n)$ 却是二阶偏导. 这说明如果 u 是热方程(7.80)的解, 那么任取 $\lambda \in \mathbb{R}$, $u(\lambda x, \lambda^2 t)$ 也是热方程(7.80)的解. 从而可知比值 $\frac{r^2}{t} (r = |x|)$ 对于热方程而言是很重要的, 同时这暗示我们可以研究热方程(7.80)形如 $u(x,t) = v(\frac{r^2}{t}) = v(\frac{|x|^2}{t}) (t > 0, x \in \mathbb{R}^n)$ 的解.

就按 $v(\frac{|x|^2}{t})$ 的形式做下去的话, 确实是可以得到热方程(7.80)的解, 但这里还是介绍一个更快的方法. 考虑解具有下述结构:

$$u(x,t) = \frac{1}{t^\alpha} v(\frac{x}{t^\beta}), \quad x \in \mathbb{R}^n, t > 0 \tag{7.83}$$

其中常数 α, β 和函数 $v : \mathbb{R}^n \to \mathbb{R}$ 是要求的. 这个形式出于我们想让热方程的解 u 在膨胀的情况下:

$$u(x,t) \mapsto \lambda^\alpha u(\lambda^\beta x, \lambda t)$$

是不变的, 也即要求

$$u(x,t) = \lambda^\alpha u(\lambda^\beta x, \lambda t), \quad \forall \lambda > 0, x \in \mathbb{R}^n, t > 0$$

取 $\lambda = \frac{1}{t}$, 令 $v(y) := u(y, 1)$ 即得(7.83)式.

现在把(7.83)式代入方程(7.80)得

$$\alpha t^{-(\alpha+1)} v(y) + \beta t^{-(\alpha+1)} y \cdot Dv(y) + t^{-(\alpha+2\beta)} \Delta v(y) = 0 \tag{7.84}$$

其中 $y := \frac{x}{t^\beta}$. 希望(7.84)式能导出一个只关于 y 的式子, 故取 $\beta = \frac{1}{2}$, 可得

$$\alpha v + \frac{1}{2} y \cdot Dv + \Delta v = 0 \tag{7.85}$$

进一步猜测 v 是放射的, 也即存在函数 $w : \mathbb{R} \to \mathbb{R}$ 使得 $v(y) = w(|y|)$, 进而由(7.85)式得

$$\alpha w + \frac{1}{2} r w' + w'' + \frac{n-1}{r} w' = 0, \quad r = |y|, w' = \frac{dw}{dr}, w'' = \frac{d^2 w}{dr^2}$$

为了凑微分, 再设 $\alpha = \frac{n}{2}$, 得到

$$(r^{n-1} w')' + \frac{1}{2}(r^n w)' = 0$$

也即

$$r^{n-1} w' + \frac{1}{2} r^n w = a$$

其中 a 是常数. 设 $r \to \infty$ 时 $w', w \to 0$ 是足够快的, 则可得 $a = 0$, 得到

$$w' = -\frac{1}{2} r w$$

解 ODE 得

$$w = b e^{-\frac{r^2}{4}} \tag{7.86}$$

其中 b 是常数. 将(7.86)式代回(7.83)式可知 $\frac{b}{t^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}$ 是热方程(7.80)的解.

定义 7.3.1 (热方程的基本解)

函数

$$\Phi(x,t) := \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}}, & x \in \mathbb{R}^n, t > 0 \\ 0, & x \in \mathbb{R}^n, t < 0 \end{cases}$$

称为热方程的基本解[a].

[a] 热方程的基本解又称为 Gauss 核.

♣

注意 Φ 在 $(0,0)$ 处是有奇性的. 有时记 $\Phi(x,t) = \Phi(|x|,t)$ 来强调基本解关于 x 有放射性. 常数 $\frac{1}{(4\pi)^{\frac{n}{2}}}$ 的选择缘自下述引理.

引理 7.3.1 (基本解的积分)

对任意的 $t > 0$, 有

$$\int_{\mathbb{R}^n} \Phi(x,t) dx = 1$$

♡

证明

计算知

$$\int_{\mathbb{R}^n} \Phi(x,t) dx = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x|^2}{4t}} dx = \frac{1}{\pi^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-|z|^2} dz = \frac{1}{\pi^{\frac{n}{2}}} \prod_{i=1}^n \int_{\mathbb{R}} e^{-z_i^2} dz_i = 1$$

□

课堂笔记 (2023.10.9)

- (关于热方程中 Laplace 算子的符号问题) 热方程的形式之所以为 $u_t - \Delta u = 0$ 而非 $u_t + \Delta u = 0$, 也可以从前面输运方程介绍过的视角来看. 如果考虑方程 $u_t - au = 0, u(0) = c$, 可解得 $u = u(0) \cdot e^{at}$, 从而 $|u|$ 在 $a > 0$ 时随着 $t \to \infty$ 而趋无穷, 在 $a < 0$ 时随着 $t \to \infty$ 而趋常数. 因为 u 本身的物理意义是温度, 故随着时间推移, 温度应稳定在某一常数, 从而物理意义要求 $a < 0$. 又因为 Δ 本身是负算子 (形式上 $\langle \Delta u, u \rangle = -\langle Du, Du \rangle \leq 0$), 故形式上令 $a = \Delta$ 即得热方程.

- (关于热方程基本解形式的推导动机) 如果 $u_t(t,x)-\Delta u(t,x)=0$, 那么令 $u^\lambda(t,x):=u(\lambda^2 t,\lambda x)$, 知
$$u_t^\lambda(t,x)-\Delta u^\lambda(t,x)=\lambda^2(u_t(\lambda^2 t,\lambda x)-\Delta u(\lambda^2 t,\lambda x))=0$$
这对任意的 λ 都是成立的, 从而若令 $\lambda=\frac{1}{\sqrt{t}}$, 即知: 若 u 是热方程的解, 则 $\frac{1}{t}u(1,\frac{x}{\sqrt{t}})$ 也应该是热方程的解. 因而考虑 $u(x,t)=\frac{1}{t^\alpha}v(\frac{x}{t^\beta})$ 的形式来进一步确定指数.

7.3.1.2 初值问题

下面应用 Φ 来推导热方程的初值问题 (或 Cauchy 问题):
$$\begin{cases} u_t-\Delta u=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 中} \\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上} \end{cases} \tag{7.87}$$
注意 $(x,t)\mapsto\Phi(x,t)$ 在奇点 (0,0) 之外是热方程的解, 故只要 $t>0$, 则对任意取定的 $y\in\mathbb{R}^n$, $(x,t)\mapsto\Phi(x-y,t)$ 也是热方程的解, 从而卷积
$$u(x,t)=\int_{\mathbb{R}^n}\Phi(x-y,t)g(y)dy=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}}g(y)dy,\quad x\in\mathbb{R}^n,t>0 \tag{7.88}$$
也应该是热方程的解.

定理 7.3.1 (热方程初值问题的解)
若 $g\in C(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n)$, u 由(7.88)式给出, 则
(i) $u\in C^\infty(\mathbb{R}^n\times(0,\infty))$,
(ii) $u_t(x,t)-\Delta u(x,t)=0(x\in\mathbb{R}^n,t>0)$,
(iii) 对 $x^0\in\mathbb{R}^n$ 有 $\lim\limits_{\substack{(x,t)\to(0,0)\\ x\in\mathbb{R}^n,t>0}}u(x,t)=g(x^0)$.

证明

注意任取 $\delta>0$, 函数 $\frac{1}{t^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}$ 在 $\mathbb{R}^n\times[\delta,\infty)$ 上都是无穷可微, 且所有导函数均一致有界的, 因而从卷积算子的性质可知 $u\in C^\infty(\mathbb{R}^n\times(0,\infty))$. 进一步, 因为 Φ 本身是热方程的解, 故
$$u_t(x,t)-\Delta u(x,t)=\int_{\mathbb{R}^n}[(\Phi_t-\Delta_x\Phi)(x-y,t)]g(y)dy=0,\quad x\in\mathbb{R}^n,t>0 \tag{7.89}$$
现在证明 (iii), 固定 $x^0\in\mathbb{R}^n,\varepsilon>0$, 因为 $g\in C(\mathbb{R}^n)$, 知可取 $\delta>0$ 足够小使得
$$\forall y\in\mathbb{R}^n(|y-x^0|<\delta\Rightarrow|g(y)-g(x^0)|<\varepsilon) \tag{7.90}$$
故若 $|x-x^0|<\frac{\delta}{2}$, 则根据引理(7.3.1)知:
$$\begin{aligned}|u(x,t)-g(x^0)|&=|\int_{\mathbb{R}^n}\Phi(x-y,t)[g(y)-g(x^0)]dy|\\&\le\int_{B(x^0,\delta)}\Phi(x-y,t)|g(y)-g(x^0)|dy+\int_{\mathbb{R}^n\setminus B(x^0,\delta)}\Phi(x-y,t)|g(y)-g(x^0)|dy\\&=:I+J\end{aligned}$$
其中对 I, 代入引理(7.3.1)与(7.90)式有
$$I\le\varepsilon\int_{\mathbb{R}^n}\Phi(x-y,t)dy=\varepsilon \tag{7.91}$$
进一步, 若 $|x-x^0|\le\frac{\delta}{2}$ 且 $|y-x^0|\ge\delta$, 则
$$|y-x^0|\le|y-x|+|x-x^0|\le|y-x|+\frac{\delta}{2}\le|y-x|+\frac{1}{2}|y-x^0|$$

得到 $|y-x| \geq \frac{1}{2}|y-x^0|$, 故对 J 有

$$\begin{aligned} J &\leq 2\|g\|_{L^\infty} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} \Phi(x-y,t)dy \leq 2\|g\|_{L^\infty} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} \frac{e^{-\frac{|x-y|^2}{4t}}}{(4\pi t)^{\frac{n}{2}}} dy \\ &= \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} e^{-\frac{|x-y|^2}{4t}} dy \leq \frac{C}{t^{\frac{n}{2}}} \int_{\mathbb{R}^n \setminus B(x^0,\delta)} e^{-\frac{|y-x^0|^2}{16t}} dy \\ &= C \int_{\mathbb{R}^n \setminus B(0,\frac{\delta}{\sqrt{t}})} e^{-\frac{|z|^2}{16}} dz \to 0, \quad t \to 0^+ \end{aligned} \tag{7.92}$$

联立(7.91),(7.92)式即得命题. □

回顾泛函分析部分中的广义函数相关知识, 例(5.10)已经证明了:

$$\begin{cases} \Phi_t - \Delta\Phi = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 中} \\ \Phi = \delta_0, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

其中 δ_0 是 δ 函数, 也即 $\mathbb{R}^n$ 中在 0 点处赋单位质量的 Dirac 测度.

注意如果 g 有界连续, $g \geq 0, g \not\equiv 0$, 则

$$u(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} g(y)dy$$

对全体 $x \in \mathbb{R}^n, t > 0$ 都是正的. 一般称这个观察结果为: 热方程对初始扰动有无限传播速度. 如果初始温度非负, 且不恒为零, 则之后任意时刻的温度分布 (不管这个时间有多久) 都是处处正的.

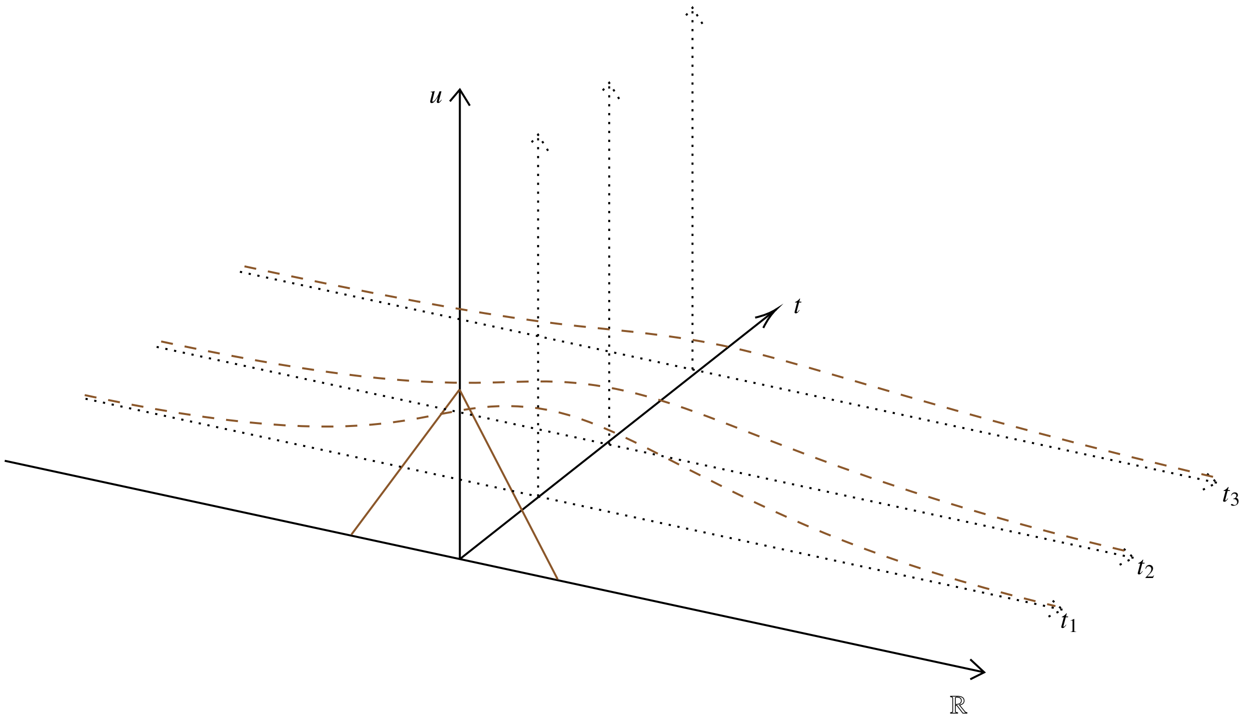


图 7.9: 热方程无限传播速度示意图

课堂笔记 (2023.10.12)

- (对无限传播速度的理解) 对热方程而言, 假设边值 g 是不恒为零的非负紧支函数, 支集为 D. 取定 $t > 0$, 任取 $x \in \mathbb{R}^n$, 根据定理(7.3.1)知:

$$u(x,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_D e^{-\frac{|x-y|^2}{4t}} g(y)dy > 0$$

这说明当计时开始时, 无论过了多短的时间, $u(x,t)$ 都已经成为了恒正函数, 亦即边值 g 的正部已经传遍了 $\mathbb{R}^n$, 故称之为无限传播速度.

- (基本解的一些简单性质与定理(7.3.1)的动机) 前面在半平面 Poisson 公式的笔记中已经介绍过用于调和延拓的 Poisson 核的本质是恒等逼近核, 事实上热核 (Gauss 核)Φ 也是恒等逼近核, 它是满足

$$\begin{cases} \Phi_t(x,t) - \Delta_x \Phi(x,t) = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ \Phi(x,t) = \delta_0(x), & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 中} \end{cases}$$

的一个函数. 定理(7.3.1)本质上依旧是恒等逼近定理的一个特例, 从而要证明定理(7.3.1), 只需证明 Φ 是好核 (恒等逼近核). 具体的阐释见下面的补充部分.

7.3.1.3 补充: 恒等逼近定理在 PDE 中的一个体现

Laplace 方程中曾经见过用 Poisson 核逼近解的定理(7.2.14), 而前面刚刚讲述的是热方程中用热核逼近解的定理(7.3.1). 事实上这些定理都是恒等逼近定理的体现, 前面在半空间 Poisson 公式的笔记中已经提到过相关思想, 这里系统整理出来并给予证明. 为了更清楚的表述 δ 函数的概念, 选用 **[Zo]** 中的定义:

定义 7.3.2 (关于基的 δ 型函数族)

由依赖于参变量 $\alpha \in A$ 的函数 $\Delta_\alpha : \mathbb{R}^n \to \mathbb{R}$ 构成的函数族 $\{\Delta_\alpha : \alpha \in A\}$ 称为 A 中的基 $\mathcal{B}$ 上的 δ 型函数族, 如果它满足下述三个条件:

(i) 函数族中所有函数均非负.

(ii) 函数族中任何函数 Δ_α 均满足 $\int_{\mathbb{R}^n} \Delta_\alpha(x)dx = 1$.

(iii) 对 $0 \in \mathbb{R}^n$ 的任意邻域 U 都有 $\lim\limits_{\mathcal{B}} \int_U \Delta_\alpha(x)dx = 1$.

上述定义也可以稍作修改, 一个等价的概念是 **[ST1]** 中提到的好核 (good kernal), 这里不再赘述. **[Zo]** 中基于 δ 型函数族的概念, 给出了点态意义下的恒等逼近定理:

引理 7.3.2 (恒等逼近定理)

设 $f : \mathbb{R}^n \to \mathbb{C}$ 是有界函数, $\{\Delta_\alpha : \alpha \in A\}$ 当 $\alpha \to \omega$ 时是 δ 型函数族. 如果对于任何 $\alpha \in A$, 卷积 $f * \Delta_\alpha$ 都存在, 并且函数 f 在 $x \in E$ 处连续, 则

$$(f * \Delta_\alpha)(x) \to f(x), \quad \alpha \to \omega.$$

若 f 在 $E \subset \mathbb{R}^n$ 内一致连续, 则上述极限在 E 内一致.

现在给出发展方程中一般的逼近命题:

命题 7.3.1

对于 m 阶线性发展方程 $(m \le k)$ 的边值问题:

$$\begin{cases} Tu(x,t) = 0, & (x,t) \in \mathbb{R}^n \times (0,+\infty) \\ u(x,t) = g(x), & (x,t) \in \mathbb{R}^n \times \{t=0\} \end{cases}$$

若存在 $\mathbb{R}^n$ 中以 $t \in (0,+\infty)$ 为参变量, 基 $t \to 0^+$ 上的 δ 型 $C^k(\mathbb{R}^n \times (0,\infty))$ 函数族 $\Phi(x,t)$ 使得:

$$T\Phi(x,t) = 0, \quad (x,t) \in \mathbb{R}^n \times (0,+\infty)$$

且 $\forall \alpha \in \mathbb{N}^{n+1} \forall t > 0 \exists M = M(\alpha) > 0((0 \le |\alpha| \le k \Rightarrow |D^\alpha \Phi(x,t)| \le M) \wedge \int_{\mathbb{R}^n} D^k \Phi(x-y,t) g(y)dy \exists)$, 则若 $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, 并取

$$w(x,t) = \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)dy, \quad x \in \mathbb{R}^n, t > 0$$

则有

(i) $w \in C^k(\mathbb{R}^n \times (0,+\infty))$,

(ii) $Tw(x,t) = 0$ 在 $\mathbb{R}^n \times (0,+\infty)$ 上成立,

(iii) 对任意 $x^0 \in \mathbb{R}^n$ 均有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n, t>0}} w(x,t) = g(x^0)$.

证明

对结论 (i), 只需证明 $\frac{\partial w}{\partial x_1}(x,t)$ 连续, 其余阶数的结论归纳即得. 记 $x' = (x_2,\cdots,x_n), y' = (y_2,\cdots,x_n)$, 知任取 $h>0$ 有:

$$\frac{1}{h}(w(x_1+h,x',t) - w(x_1,x',t)) = \frac{1}{h}\int_{\mathbb{R}^n}(\Phi(x_1+h-y_1,x'-y',t) - \Phi(x_1-y_1,x'-y',t))g(y)dy$$

注意

$$\begin{aligned}|\frac{1}{h}(w(x_1+h,x',t) - w(x_1,x',t))| &\le \int_{\mathbb{R}^n}|\frac{1}{h}(\Phi(x_1+h-y_1,x'-y',t) - \Phi(x_1-y_1,x'-y',t))|\cdot|g(y)|dy\\ &\le M\int_{\mathbb{R}^n}|g(y)|dy < \infty\end{aligned}$$

故由 Lebesgue 控制收敛定理知

$$\lim_{h\to0}\frac{1}{h}(w(x_1+h,x',t) - w(x_1,x',t)) = \int_{\mathbb{R}^n}\Phi_{x_1}(x-y,t)g(y)dy$$

存在, 从而

$$w_{x_1}(x,t) = \int_{\mathbb{R}^n}\Phi_{x_1}(x-y,t)g(y)dy$$

存在, 同时由 Φ_{x_1} 的连续性即得 w_{x_1} 连续.

对结论 (ii), 由结论 (i) 的证明可知

$$Tw(x,t) = \int_{\mathbb{R}^n}(T\Phi)(x-y,t)g(y)dy$$

又因为 $T\Phi \equiv 0$, 故 $Tw(x,t) = 0$.

对结论 (iii), 因为 $g \in C(\mathbb{R}^n)$, 故由恒等逼近定理即得结论. □

7.3.1.4 非齐次问题

下面研究非齐次初值问题

$$\begin{cases}u_t - \Delta u = f, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 中}\\ u = 0, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases} \tag{7.93}$$

回忆(7.88)式的形式, 不难发现映射 $(x,t)\mapsto\Phi(x-y,s-t)$ 也是热方程的解, 其中 $y\in\mathbb{R}^n, 0<s<t$ 都是给定的. 现在固定 s, 函数

$$u = u(x,t;s) = \int_{\mathbb{R}^n}\Phi(x-y,t-s)f(y,s)dy \tag{7.94}$$

就是初值问题

$$\begin{cases}u_t(\cdot;s) - \Delta u(\cdot;s) = 0, & \text{在 } \mathbb{R}^n\times(s,\infty) \text{ 中}\\ u(\cdot;s) = f(\cdot;s), & \text{在 } \mathbb{R}^n\times\{t=s\} \text{ 上}\end{cases} \tag{7.95}$$

的解. 这与前面讨论的初值问题(7.87)区别无非是把起始时间 $t=0$ 换成了 $t=s$, 并把 g 换成了 $f(\cdot;s)$. 当然, $u(\cdot;s)$ 这时还不是初值问题(7.93)的解.

但根据 Duhamel 原理, 可以从(7.95)的解中构造出(7.93)的解, 其方法在于对 s 积分. 考虑

$$u(x,t) = \int_0^t u(x,t;s)ds, \quad x\in\mathbb{R}^n, t\ge0$$

代入(7.94)式知

$$u(x,t)=\int_0^t\int_{\mathbb{R}^n}\Phi(x-y,t-s)f(y,s)dyds=\int_0^t\frac{1}{(4\pi(t-s))^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4(t-s)}}f(y,s)dyds,\quad x\in\mathbb{R}^n,t>0 \tag{7.96}$$

下面验证(7.96)式确实是初值问题(7.93)的解, 简便起见不妨设 $f\in C_1^2(\mathbb{R}^n\times[0,\infty))$[6]是紧支的.

定理 7.3.2 (非齐次初值问题的解)

若 u 由(7.96)式定义, 则

(i) $u\in C_1^2(\mathbb{R}^n\times(0,\infty))$,

(ii) $u_t(x,t)-\Delta u(x,t)=f(x,t)(x\in\mathbb{R}^n,t>0)$,

(iii) 对 $x^0\in\mathbb{R}^n$ 有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n,t>0}}u(x,t)=0$.

证明

既然 (0,0) 是 Φ 的奇点, 故不能直接交换积分号和微分号. 还是考虑取 $(0,0)$ 的一个小邻域研究.

首先验证 (i), 换元知

$$u(x,t)=\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)f(x-y,t-s)dyds$$

既然 $f\in C_1^2(\mathbb{R}^n\times[0,\infty))$ 紧支, 且 $\Phi=\Phi(y,s)$ 在 $s=t>0$ 附近光滑, 故

$$u_t(x,t)=\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy+\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)f_t(x-y,t-s)dyds$$

同时

$$u_{x_ix_j}(x,t)=\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)f_{x_ix_j}(x-y,t-s)dyds,\quad i,j=1,\cdots,n$$

这说明 $u_t,D_x^2u\in C(\mathbb{R}^n\times(0,\infty))$, 自然 $u,D_xu\in C(\mathbb{R}^n\times(0,\infty))$, 也即 $u\in C_1^2(\mathbb{R}^n\times(0,\infty))$.

下面验证 (ii), 知

$$\begin{aligned}u_t(x,t)-\Delta u(x,t)&=\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)f_t(x-y,t-s)dyds+\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy-\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)\Delta_xf(x-y,t-s)dyds\\&=\int_0^t\int_{\mathbb{R}^n}\Phi(y,s)[(\frac{\partial}{\partial t}-\Delta_x)f(x-y,t-s)]dyds+\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy\\&=\int_0^t\int_{\mathbb{R}^n}\Phi(t,s)[(-\frac{\partial}{\partial s}-\Delta_y)f(x-y,t-s)]dyds+\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy\\&=\int_\varepsilon^t\int_{\mathbb{R}^n}\Phi(y,s)[(-\frac{\partial}{\partial s}-\Delta_y)f(x-y,t-s)]dyds+\int_0^\varepsilon\int_{\mathbb{R}^n}\Phi(y,s)[(-\frac{\partial}{\partial s}-\Delta_y)f(x-y,t-s)]dyds\\&\quad+\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy\\&=:I_\varepsilon+J_\varepsilon+K\end{aligned} \tag{7.97}$$

现在对 J_ε, 由引理(7.3.1)知

$$|J_\varepsilon|\le\|f_t+D^2f\|_{L^\infty}\int_0^\varepsilon\int_{\mathbb{R}^n}\Phi(y,s)dyds\le C\int_0^\varepsilon 1ds=C\varepsilon \tag{7.98}$$

[6]若 $u=u(x,t)$, 则 $C_1^2(U_T)=\{u:U_T\to\mathbb{R}:u,D_xu,D_x^2u,u_t\in C(U_T)\}$, 也即上标代表对空间变量 x 的光滑程度, 下标代表对时间变量 t 的光滑程度.

分部积分知

$$\begin{aligned}&\int_{\varepsilon}^{t}\int_{\mathbb{R}^n}\Phi(y,s)(-\frac{\partial}{\partial s}f(x-y,t-s))dyds\\=&\int_{\mathbb{R}^n}dy\int_{\varepsilon}^{t}\Phi(y,s)(-\frac{\partial}{\partial s}f(x-y,t-s))ds\\=&\int_{\mathbb{R}^n}(-\Phi(y,t)f(x-y,0)+\Phi(y,\varepsilon)f(x-y,t-\varepsilon))dy+\int_{\mathbb{R}^n}dy\int_{\varepsilon}^{t}\frac{\partial}{\partial s}\Phi(y,s)\cdot f(x-y,t-s)ds\end{aligned}\tag{7.99}$$

因为 f 是紧支的, 故在 ∂I 上用分部积分法[7]同样可以验证

$$\int_{\varepsilon}^{t}\int_{\mathbb{R}^n}\Phi(y,s)(-\Delta_y f(x-y,t-s))dyds=\int_{\varepsilon}^{t}\int_{\mathbb{R}^n}(-\Delta_y\Phi(y,s))f(x-y,t-s)dyds\tag{7.100}$$

结合(7.99),(7.100)式并注意 Φ 本身是热方程(7.80)的解, 对 I_ε 有

$$\begin{aligned}I_\varepsilon=&\int_{\varepsilon}^{t}\int_{\mathbb{R}^n}[(\frac{\partial}{\partial s}-\Delta_y)\Phi(y,s)]f(x-y,t-s)dyds+\int_{\mathbb{R}^n}\Phi(y,\varepsilon)f(x-y,t-\varepsilon)dy-\int_{\mathbb{R}^n}\Phi(y,t)f(x-y,0)dy\\=&\int_{\mathbb{R}^n}\Phi(y,\varepsilon)f(x-y,t-\varepsilon)dy-K\end{aligned}\tag{7.101}$$

联立(7.97),(7.98),(7.101)式得

$$u_t(x,t)-\Delta u(x,t)=\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n}\Phi(y,\varepsilon)f(x-y,t-\varepsilon)dy=f(x,t),\quad x\in\mathbb{R}^n,t>0$$

最后的极限仿照例(5.10)的过程即得. 最后从 u 的构造看出 $\|u(\cdot,t)\|_{L^\infty}\le t\|f\|_{L^\infty}\to 0$, 这说明 u 是衰减的, 命题得证. □

结合定理(7.3.1)与定理(7.3.2)可知:

$$u(x,t)=\int_{\mathbb{R}^n}\Phi(x-y,t)g(y)dy+\int_0^t\int_{\mathbb{R}^n}\Phi(x-y,t-s)f(y,s)dyds\tag{7.102}$$

在相应定理中 g,f 的光滑性下, 正是一般非齐次初值问题

$$\begin{cases}u_t-\Delta u=f, & \text{在 } \mathbb{R}^n\times(0,\infty)\text{ 中}\\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\}\text{ 上}\end{cases}\tag{7.103}$$

的解.

课堂笔记 (2023.10.16)

- (关于热方程初值问题的形式求解) 对于热方程

$$\begin{cases}u_t(t,x)-\Delta_x u(t,x)=0, & (t,x)\in\mathbb{R}^+\times\mathbb{R}^n\\ u(t=0,x)=g(x), & x\in\mathbb{R}^n\end{cases}$$

其中 $\Delta_x=\frac{\partial^2}{\partial x_1^2}+\cdots+\frac{\partial^2}{\partial x_n^2}$. 如果依旧把 Δ_x 视作常数 a, 求解 $u_t(t,x)-au(t,x)=0$ 可得 $u(t,x)=C(x)e^{ta}$, 代入 $t=0$ 可知 $C(x)=g(x)$, 将 a 换回 Δ_x 可得形式上的表达式:

$$u(t,x)=e^{t\Delta_x}g(x)$$

回忆热方程的基本解满足方程:

$$\begin{cases}\Phi_t(t,x)-\Delta\Phi(t,x)=0 & (t,x)\in\mathbb{R}^+\times\mathbb{R}^n\\ \Phi(t=0,x)=\delta_x=\delta_0(x), & x\in\mathbb{R}^n\end{cases}$$

类似于前述思路可得

$$\Phi(t,x)=e^{t\Delta_x}\delta_x$$

[7]这里看成是对每个分量 y_i 分部积分也好, 从广义微商的角度来看也好, 都是可以验证的.

现在回忆

$$g(x) = \langle \delta_x, g \rangle = \int_{\mathbb{R}^n} \delta_0(x-y)g(y)dy$$

故如若 $e^{t\Delta_x}$ 与积分号可交换, 则有:

$$\begin{aligned} u(t,x) &= e^{t\Delta_x}\int_{\mathbb{R}^n} \delta_0(x-y)g(y)dy \\ &= \int_{\mathbb{R}^n} e^{t\Delta_{x-y}}\delta_0(x-y)g(y)dy \\ &= \int_{\mathbb{R}^n} \Phi(t,x-y)g(y)dy \end{aligned}$$

代回 $\Phi(t,x)$ 的表达式即得

$$u(t,x) = \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}}g(y)dy, \quad x \in \mathbb{R}^n, t > 0$$

- (关于 Duhamel 原理与叠加 (Superposition) 原理的区分) 对于非齐次热方程初值问题

$$\begin{cases} u_t - \Delta u = f, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

叠加原理所保证的是可以把上述问题的解拆分成下述两个问题的解之和:

$$\begin{cases} v_t - \Delta v = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ v = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \qquad \begin{cases} w_t - \Delta w = f, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ w = 0, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

其中前一个问题正是齐次热方程初值问题, 可以根据定理(7.3.1)求解. 而 Duhamel 原理主要解决的是后面这个非齐次热方程零初值问题, Duhamel 原理本身的断言是: 如果可以求解方程

$$\begin{cases} u_t(\cdot;s) - \Delta u(\cdot;s) = 0, & \text{在 } \mathbb{R}^n \times (s,\infty) \text{ 内} \\ u(\cdot;s) = f(\cdot;s), & \text{在 } \mathbb{R}^n \times \{t=s\} \text{ 上} \end{cases}$$

得到解族 $u_s(x,t;s)$, 那么函数

$$u(x,t) = \int_0^t u_s(x,t;s)ds, \quad x \in \mathbb{R}^n, t \geq 0$$

就正是原先的非齐次零边值问题的一个解.

7.3.2 均值定理

首先规定一些记号. 设 $U \subset \mathbb{R}^n$ 是有界开集, 固定时间 $T > 0$.

定义 7.3.3 (抛物柱 U_T, 抛物边界 Γ_T)

1. 定义抛物柱:

$$U_T := U \times (0,T].$$

2. U_T 的抛物边界为

$$\Gamma_T := \overline{U_T} \backslash U_T.$$

根据定义知 U_T 可以看做是 $\overline{U} \times [0,T]$ 的抛物内部, 特别注意 U_T 是包含顶端 $U \times \{t=T\}$ 的. 抛物边界 Γ_T 包含该柱面的底部与垂直侧边, 但不包含顶部 (参看图(7.10)).

下面希望导出热方程的解类似于调和函数均值定理(7.31)的性质. 当然, 此时的公式不会再像调和函数情况那么简单了, 不过注意如果固定 x, 则不同半径 r 的球面 $\partial B(x,r)$ 构成 Laplace 方程基本解 $\Phi(x-y)$ 的等值面. 这暗示我们对固定的 (x,t), 求解热方程基本解 $\Phi(x-y,t-s)$ 的等值面可能会有所帮助.

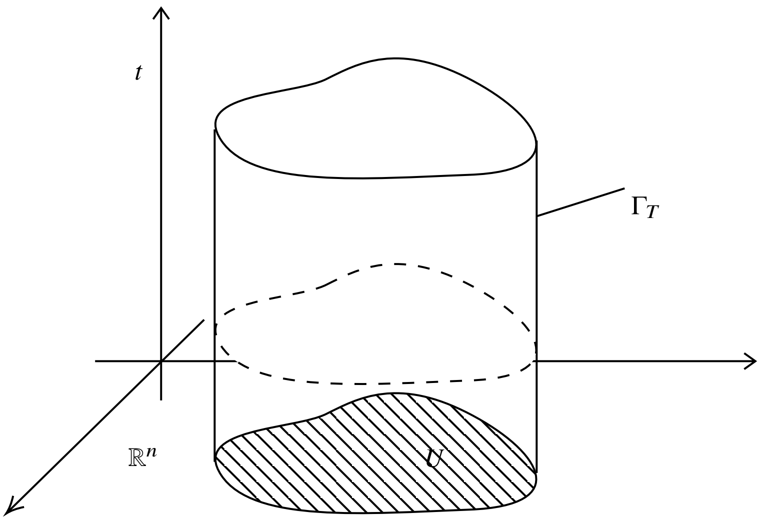


图 7.10

定义 7.3.4 (热球)

对固定的 $x \in \mathbb{R}^n, t \in \mathbb{R}, r > 0$, 定义

$$E(x,t;r) := \{(y,s) \in \mathbb{R}^{n+1} : s \leq t, \Phi(x-y,t-s) \geq \frac{1}{r^n}\}$$

♣

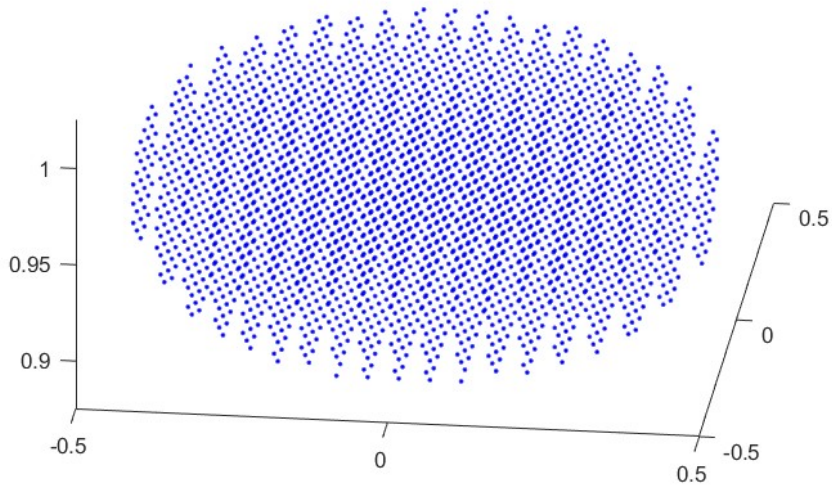


图 7.11: 按 $n=2, x=0, t=1, \frac{1}{r^n}=\frac{1}{2}$ 画出的热球

可以发现热球是一个时空区域, 其边界正是 $\Phi(x-y,t-s)$ 的一个等值面. 从图(7.11)中也可以观察到 (x,t) 是在热球的最顶端的.

定理 7.3.3 (热方程的均值性)

若 $u \in C_1^2(U_T)$ 是热方程的解, 则对任意的 $E(x,t;r) \subset U_T$ 有[a]

$$u(x,t) = \frac{1}{4r^n} \iint_{E(x,t;r)} u(y,s) \frac{|x-y|^2}{(t-s)^2} dyds. \tag{7.104}$$

^{a}一般把 $\frac{|x-y|^2}{(t-s)^2}dyds$ 称为抛物测度.

(7.104)式正是热方程类似于 Laplace 方程中均值定理的结论. 观察到该式右端只包含 $u(y,s)$ 在时间 $s \le t$ 的情况, 这是因为 $u(x,t)$ 不能由未来的情况来决定.

证明

平移时空坐标, 取 $x=0, t=0$. 因为可以对函数作磨光处理, 故不妨就设 u 是光滑的. 记 $E(r)=E(0,0;r)$, 另设 $y=r\widetilde{y}, s=r^2\widetilde{s}$, 取

$$\begin{aligned}
\phi(r) &:= \frac{1}{r^n}\iint_{E(r)} u(y,s)\frac{|y|^2}{s^2}dyds \\
&= \frac{1}{r^n}\iint_{\{(y,s)\in\mathbb{R}^{n+1}:s\le 0,(\frac{1}{-4\pi s})^{\frac{n}{2}}e^{\frac{|y|^2}{4\pi s}}\ge\frac{1}{r^n}\}} u(y,s)\frac{|y|^2}{s^2}dyds \\
&= \frac{1}{r^n}\iint_{\{(\widetilde{y},\widetilde{s})\in\mathbb{R}^{n+1}:\widetilde{s}\le 0,(\frac{1}{-4\pi(r^2\widetilde{s})})^{\frac{n}{2}}e^{\frac{|r\widetilde{y}|^2}{4\pi(r^2\widetilde{s})}}\ge\frac{1}{r^n}\}} u(\widetilde{y},\widetilde{s})\frac{|r\widetilde{y}|^2}{(r^2\widetilde{s})^2}d(r\widetilde{y})d(r^2\widetilde{s}) \\
&= \iint_{\{(\widetilde{y},\widetilde{s})\in\mathbb{R}^{n+1}:\widetilde{s}\le 0,(\frac{1}{-4\pi\widetilde{s}})^{\frac{n}{2}}e^{\frac{|\widetilde{y}|^2}{4\pi\widetilde{s}}}\ge 1\}} u(\widetilde{y},\widetilde{s})\frac{|\widetilde{y}|^2}{\widetilde{s}^2}d\widetilde{y}d\widetilde{s} \\
&= \iint_{E(1)} u(ry,r^2s)\frac{|y^2|}{s^2}dyds
\end{aligned} \tag{7.105}$$

求导有

$$\begin{aligned}
\phi'(r) &= \iint_{E(1)} (\sum_{i=1}^n u_{y_i}(ry,r^2s)\cdot y_i\frac{|y|^2}{s^2} + 2ru_s(ry,r^2s)\cdot\frac{|y|^2}{s})dyds \\
&= \frac{1}{r^{n+1}}\iint_{E(r)}((\sum_{i=1}^n u_{y_i}(y,s)\cdot y_i\frac{|y|^2}{s^2}) + 2u_s(y,s)\cdot\frac{|y|^2}{s})dyds =: A+B
\end{aligned}$$

下面先计算 $B=\frac{1}{r^{n+1}}\iint_{E(r)}2u_s(y,s)\cdot\frac{|y|^2}{s}dyds$. 把 $E(r)$ 定义中的 $\Phi(-y,-s)\ge\frac{1}{r^n}$ 写开:

$$\frac{1}{(-4\pi s)^{\frac{n}{2}}}e^{\frac{|y|^2}{4s}} \ge \frac{1}{r^n}$$

注意两边都至少是正的, 因而取对数有

$$-\frac{n}{2}\ln(-4\pi s)+\frac{|y|^2}{4s}\ge -n\ln r$$

从而可设函数

$$\psi(s,y,r) := -\frac{n}{2}\ln(-4\pi s)+\frac{|y|^2}{4s}+n\ln r \tag{7.106}$$

因为 $(y,s)\in\partial E(r)\backslash\{(0,0)\}$ 与 $\Phi(-y,-s)=\frac{1}{r^n}$ 等价, 从而 $\partial E(r)\backslash\{(0,0)\}=\{(y,s)\in\mathbb{R}^{n+1}:\psi(s,y,r)=0\}$. 计算易得 $\psi_{y_i}(s,y,r)=\frac{y_i}{2s}$, 从而整理有:

$$\begin{aligned}
B &= \frac{1}{r^{n+1}}\iint_{E(r)}2u_s(y,s)\cdot\frac{|y|^2}{s}dyds = \frac{1}{r^{n+1}}\iint_{E(r)}4u_s(y,s)(\sum_{i=1}^n y_i\cdot\frac{y_i}{2s})dyds \\
&= \frac{1}{r^{n+1}}\iint_{E(r)}4u_s(y,s)(\sum_{i=1}^n y_i\psi_{y_i}(s,y,r))dyds = \frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^n 4u_s(y,s)y_i\psi_{y_i}(s,y,r)dyds
\end{aligned}$$

对单个 i 而言, 对 y_i 分部积分有:

$$\begin{aligned}
&\frac{1}{r^{n+1}}\iint_{E(r)|_{y_i}} 4u_s(y,s)y_i\psi_{y_i}(s,y,r)dy_ids\\
=&\frac{1}{r^{n+1}}\iint_{E(r)|_{y_i}} 4u_s(y,s)y_id\psi(s,y,r)ds\\
=&\frac{1}{r^{n+1}}(4u_s(y,s)y_i\psi(s,y,r)|_{(y,s)\in\partial E(r)|_{y_i}}-\iint_{E(r)|_{y_i}} 4\psi(s,y,r)d(u_s(y,s)y_i)ds)\\
=&-\frac{1}{r^{n+1}}\iint_{E(r)|_{y_i}} 4\psi(s,y,r)(u_{sy_i}(y,s)y_i+u_s(y,s))dy_ids
\end{aligned}$$

其中 $E(r)|_{y_i}$ 表示 $E(r)$ 与 $y=y_i$ 相交所得截面, 加和得到:

$$\begin{aligned}
B&=-\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}(4\psi(s,y,r)u_{sy_i}(y,s)y_i+4\psi(s,y,r)u_s(y,s))dyds\\
&=-\frac{1}{r^{n+1}}\iint_{E(r)}((\sum_{i=1}^{n}4\psi(s,y,r)u_{sy_i}(y,s)y_i)+4n\psi(s,y,r)u_s(y,s))dyds
\end{aligned}\tag{7.107}$$

在前一项对 s 分部积分有:

$$\begin{aligned}
&-\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}4\psi(s,y,r)u_{sy_i}(y,s)y_idyds\\
=&-\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}4\psi(s,y,r)y_idydu_{y_i}(y,s)\\
=&-\frac{1}{r^{n+1}}((\sum_{i=1}^{n}4\psi(s,y,r)y_iu_{y_i}(y,s))|_{\partial E(r)|_s}-\iint_{E(r)}\sum_{i=1}^{n}4u_{y_i}(y,s)y_idyd\psi(s,y,r))\\
=&\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}4u_{y_i}(y,s)y_i\psi_s(s,y,r)dyds
\end{aligned}$$

容易算得 $\psi_s=-\frac{n}{2s}-\frac{|y|^2}{4s^2}$, 代入有:

$$\begin{aligned}
&\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}4u_{y_i}(y,s)y_i\psi_s(s,y,r)dyds\\
=&\frac{1}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}4u_{y_i}(y,s)y_i(-\frac{n}{2s}-\frac{|y|^2}{4s^2})dyds\\
=&-\frac{n}{r^{n+1}}\iint_{E(r)}\sum_{i=1}^{n}2u_{y_i}(y,s)\frac{y_i}{s}-A
\end{aligned}\tag{7.108}$$

将(7.108)式代入(7.107)式并整理知

$$B=\frac{1}{r^{n+1}}\iint_{E(r)}(-4nu_s(y,s)\psi(s,y,r)-\frac{2n}{s}\sum_{i=1}^{n}u_{y_i}(y,s)y_i)dyds-A$$

既然 u 是热方程的解, 知 $u_s=\Delta u$, 结合 Green 公式(6.2.3)(ii) 知

$$\begin{aligned}
\phi'(r)=A+B&=\frac{1}{r^{n+1}}\iint_{E(r)}(-4n\Delta u(y,s)\psi(s,y,r)-\frac{2n}{s}\sum_{i=1}^{n}u_{y_i}(y,s)y_i)dyds\\
&=\sum_{i=1}^{n}\frac{1}{r^{n+1}}\iint_{E(r)}(4nu_{y_i}(y,s)\psi_{y_i}(s,y,r)-\frac{2n}{s}u_{y_i}(y,s)y_i)dyds\\
&=\sum_{i=1}^{n}\frac{1}{r^{n+1}}\iint_{E(r)}(4nu_{y_i}(y,s)\cdot\frac{y_i}{2s}-\frac{2n}{s}u_{y_i}(y,s)y_i)dyds=0
\end{aligned}$$

这说明 ϕ 是常数, 因而

$$\phi(r)=\lim_{t\to0}\phi(t)=u(0,0)\cdot\lim_{t\to0}\frac{1}{t^n}\iint_{E(t)}\frac{|y|^2}{s^2}dyds$$

最后注意结论

$$\frac{1}{t^n}\iint_{E(t)}\frac{|y|^2}{s^2}dyds=\iint_{E(1)}\frac{|y|^2}{s^2}dyds=4 \tag{7.109}$$

即得

$$\phi(r)=4u(0,0)$$

命题即证. □

课堂笔记 (2023.10.19)

- (关于上述证明中蓝色部分的解释) 平移时空坐标在物理中对应的是时空的平移不变性, 具体可以追溯到 Galois 构造 (参见笔记的第三部分). 从数学上来讲, 如果需要验证 (x_0,t_0) 处的均值公式, 即验证:

$$u(x_0,t_0)=\frac{1}{4r^n}\iint_{E(x_0,t_0;r)}u(y,s)\frac{|x_0-y|^2}{(t_0-s)^2}dyds$$

回忆热球的定义:

$$E(x_0,t_0;r)=\{(y,s)\in\mathbb{R}^{n+1}:s\le t_0,\Phi(x_0-y,t_0-s)\ge\frac{1}{r^n}\}$$

将 y 换元为 $y+x_0$, s 换元为 $s+t_0$, 得到:

$$u(x_0,t_0)=\frac{1}{4r^n}\iint_{E(0,0;r)}u(y+x_0,s+t_0)\frac{|y|^2}{s^2}dyds$$

令 $v(x,t)=u(x+x_0,t+t_0)$, 根据热方程的线性性知 v 也是热方程的解. 现在如若验证了 v 在 (0,0) 处的均值公式:

$$v(0,0)=\frac{1}{4r^n}\iint_{E(0,0;r)}v(y,s)\frac{|y|^2}{s^2}dyds$$

将 v 换回 u, 自然得到欲证. 从而要验证均值公式对任意点成立, 只需要验证 (0,0) 处的均值公式成立就足够了.

- (关于结论(7.109)的推导) 欲证

$$\int_{E(1)}\frac{|y|^2}{s^2}dyds=4$$

回忆

$$E(1)=\{(y,s)\in\mathbb{R}^{n+1}:s\le 0,\Phi(-y,-s)=\frac{1}{(-4\pi s)^{\frac{n}{2}}}e^{\frac{|y|^2}{4s}}\ge 1\}$$

将 s 替换为 $-s$, y 替换为 $-y$, 设 $r=|y|$, 知欲证未变, $E(1)$ 变为:

$$E(1)=\{(y,s)\in\mathbb{R}^{n+1}:s\ge 0,r=|y|,\Phi(y,s)=\frac{1}{(4\pi s)^{\frac{n}{2}}}e^{-\frac{r^2}{4s}}\ge 1\}$$

知

$$\begin{aligned}\frac{1}{(4\pi s)^{\frac{n}{2}}}e^{-\frac{r^2}{4s}}\ge 1 &\Leftrightarrow e^{-\frac{r^2}{4s}}\ge(4\pi s)^{\frac{n}{2}}\\ &\Leftrightarrow -\frac{r^2}{4s}\ge\ln(4\pi s)^{\frac{n}{2}}\\ &\Leftrightarrow r\le(4ns\ln\left(\frac{1}{\sqrt{4\pi s}}\right))^{\frac{1}{2}}\end{aligned}$$

这说明对每个固定的 s, $E(1)$ 关于 y 的截面都可以表示成球

$$B_s:=\{y\in\mathbb{R}^n:|y|=r\le r_s:=(4ns\ln\left(\frac{1}{\sqrt{4\pi s}}\right))^{\frac{1}{2}}\}$$

另外为了保证 r 有解, 需 $\frac{1}{\sqrt{4\pi s}}\ge 1$, 亦即 $s\le\frac{1}{4\pi}$. 现在关于 r 应用 Fubini 定理知

$$I:=\int_{E(1)}\frac{|y|^2}{s^2}dyds=\int_0^{\frac{1}{4\pi}}\frac{1}{s^2}ds\int_0^{r_s}dt\int_{S_t}|y|^2d\sigma$$

其中 $r_s=(4ns\ln(\frac{1}{\sqrt{4\pi s}}))^{\frac{1}{2}}$, $S_t=\partial B_t$. 知在 S_t 上 $|y|^2=t^2$, 且已知 n 维单位球面面积为 $\omega_n=\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$, 有

$$\int_0^{r_s}dt\int_{S_t}|y|^2d\sigma=\int_0^{r_s}t^2\cdot\omega_nt^{n-1}dt=\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\int_0^{r_s}t^{n+1}dt=\frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\frac{r_s^{n+2}}{n+2}$$

从而代回 r_s 的定义有

$$I=\frac{2\pi^{\frac{n}{2}}}{(n+2)\Gamma(\frac{n}{2})}\int_0^{\frac{1}{4\pi}}s^{-2}(4ns\ln\left(\frac{1}{\sqrt{4\pi s}}\right))^{\frac{n+2}{2}}ds$$

换元 $\zeta=n\ln\left(\frac{1}{\sqrt{4\pi s}}\right)$, 知 $s=\frac{1}{4\pi}e^{-\frac{2\zeta}{n}}$, 同时 $d\zeta=-\frac{n}{2}\frac{ds}{s}$, 有

$$\begin{aligned}I&=\frac{2\pi^{\frac{n}{2}}}{(n+2)\Gamma(\frac{n}{2})}\int_0^{\frac{1}{4\pi}}4^{\frac{n+2}{2}}s^{\frac{n}{2}-1}(n\ln\left(\frac{1}{\sqrt{4\pi s}}\right)^{\frac{n+2}{2}})ds\\&=\frac{2\pi^{\frac{n}{2}}}{(n+2)\Gamma(\frac{n}{2})}\cdot4^{\frac{n+2}{2}}\int_\infty^0(\frac{1}{4\pi}e^{-\frac{2\zeta}{n}})^{\frac{n}{2}-1}\zeta^{\frac{n+2}{2}}(-\frac{2}{n}\cdot4\pi e^{\frac{2\zeta}{n}})d\zeta\\&=\frac{2^2\pi^{\frac{n}{2}}}{(n+2)\Gamma(\frac{n}{2})}\cdot\frac{4^{\frac{n}{2}+1}}{(4\pi)^{\frac{n}{2}}n}\int_0^\infty e^{-\zeta}\zeta^{\frac{n}{2}+1}d\zeta\\&=\frac{2^4}{n(n+2)\Gamma(\frac{n}{2})}\Gamma(\frac{n}{2}+2)=\frac{2^4}{n(n+2)\Gamma(\frac{n}{2})}(\frac{n}{2}+1)\frac{n}{2}\Gamma(\frac{n}{2})=4\end{aligned}$$

命题即证.

- (关于热球的进一步理解) 热球本质上称作" 抛物球", 其一大特点在于热球的中心在边界上. 一维热方程对应的热球实际上是可以画出的, 见图(7.12), 留意 $\Phi(x-y,t-s)=\frac{1}{r^n}$ 构成的等值面对图像的切割可以产生更好的理解.

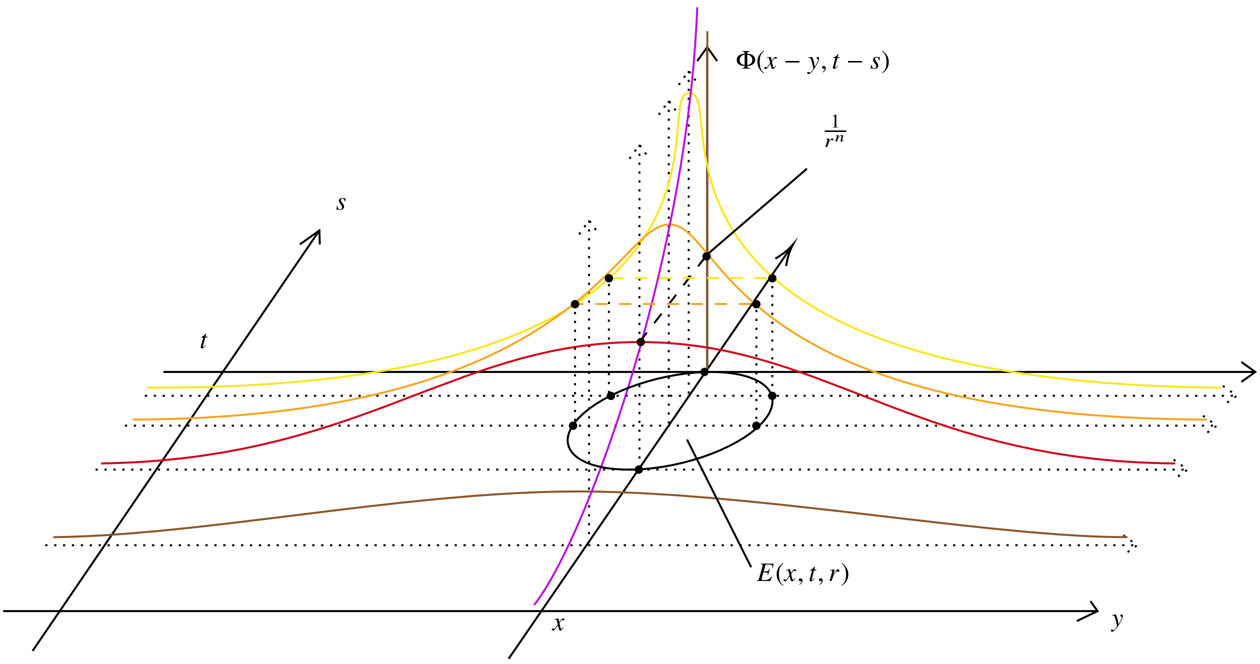


图 7.12: 一维热方程对应热球示意图

7.3.3 解的性质

7.3.3.1 强最大值原理与解的唯一性

定理 7.3.4 (热方程的强最大值原理)

若 $u \in C_1^2(U_T) \cap C(\overline{U_T})$ 是热方程在 U_T 上的解, 则

(i) $\max\limits_{\overline{U_T}} u = \max\limits_{\Gamma_T} u$.

(ii) 进一步, 若 U 是连通集, 且存在 $(x_0, t_0) \in U_T$ 满足

$$u(x_0, t_0) = \max_{\overline{U_T}} u$$

则 u 在 $\overline{U_{t_0}}$ 上是常数.

♡

断言 (i) 称作热方程的最大值原理, 而 (ii) 称作强最大值原理. 把 max 换成 min 也可以得到强最小值原理. 从物理意义上来讲, 强最大值原理强调的无非是如果 u 在某个内点取得最值, 那它在此刻及以前就全是常数. 这个解释缘自我们把 t 变量看成时间所产生的物理直觉: 只要在 $[0, t_0]$ 这段时间内初值和边值条件都是常数, 那么解自然也该是常数. 不过只要 t_0 之后边值条件发生变化, 这之后的解也同样会发生变化, 而只要边值条件保持不变, 解也就一直不变.

下面从 PDE 而不仅仅是物理直觉的角度来推导这件事情.

证明

设存在 $(x_0, t_0) \in U_T$ 满足 $u(x_0, t_0) = M := \max\limits_{\overline{U_T}} u$. 既然 (x_0, t_0) 此时是在 U_T 的内部, 总会存在充分小的 $r > 0$ 满足 $E(x_0, t_0; r) \subset U_T$. 应用均值定理(7.104)有:

$$M = u(x_0, t_0) = \frac{1}{4r^n} \iint_{E(x_0,t_0;r)} u(y, s) \frac{|x_0 - y|^2}{(t_0 - s)^2} dyds \le M$$

又因为由(7.109)式知

$$1 = \frac{1}{4r^n} \iint_{E(x_0,t_0;r)} \frac{|x_0 - y|^2}{(t_0 - s)^2} dyds$$

知等号成立当且仅当 u 在 $E(x_0, t_0; r)$ 内恒等于 M, 故

$$u(y, s) = M, \quad \forall (y, s) \in E(x_0, t_0; r)$$

现任取另一点 $(y_0, s_0) \in U_T$, 其中 $s_0 < t_0$, 用线段 L 连接 $(x_0, t_0), (y_0, s_0)$(如图(7.13)), 考虑

$$r_0 := \min\{s \ge s_0 : \forall (x, t) \in L(u(x, t) = M), s \le t \le t_0\}$$

因为 u 是连续的, 故最小值必能取到. 若 $r_0 > s_0$, 也即沿着 L 从 (x_0, t_0) 往 (y_0, s_0) 走的过程中, $u(x, t) = M$ 这件事会在中途某个点 (z_0, r_0) 后不再成立. 如图(7.13), 此时 $E(z_0, r_0; r)$ 对任意 r 还是会与 L 在 r_0 的下方有一段相交. 用标准的语言来说即存在 $\sigma > 0$ 使得 $E(z_0, r_0; r) \cap (L \cap \{r_0 - \sigma \le t \le r_0\}) \ne \emptyset$[8], 这说明存在 $r' < r_0$ 使得 $(x, t) \in L, r' \le t \le r_0$ 且 $u(x, t) = M$. 这与 r_0 的构造矛盾! 故只能有 $r_0 = s_0$, 因而在 L 上 $u \equiv M$.

现在任意固定 $x \in U$ 与时间 $t \in [0, t_0)$, 由 U 的连通性知总存在点 $x_0, x_1, \cdots, x_{m-1}, x_m = x$ 使得在 U 中存在一系列连接 x_{i-1} 与 $x_i (i = 1, \cdots, m)$ 的线段. 选取时间 $t_0 > t_1 > \cdots > t_m = t$, 按照在 U 中的方式依次连接 $(x_1, t_1), \cdots, (x_m, t_m) = (x, t)$, 显见这些线段必定在 U_T 内. 根据前述结论, u 在这些线段上一直为 M, 因而 $u(x, t) = M$. □

强最大值原理表明如果 U 连通, 且 $u \in C_1^2(U_T) \cap C(\overline{U_T})$ 满足

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } U_T \text{ 内} \\ u = 0, & \text{在 } \partial U \times [0, T] \text{ 上} \\ u = g, & \text{在 } U \times \{t = 0\} \text{ 上} \end{cases}$$

[8]这个性质并不平凡, 具体证明有待进一步思考.

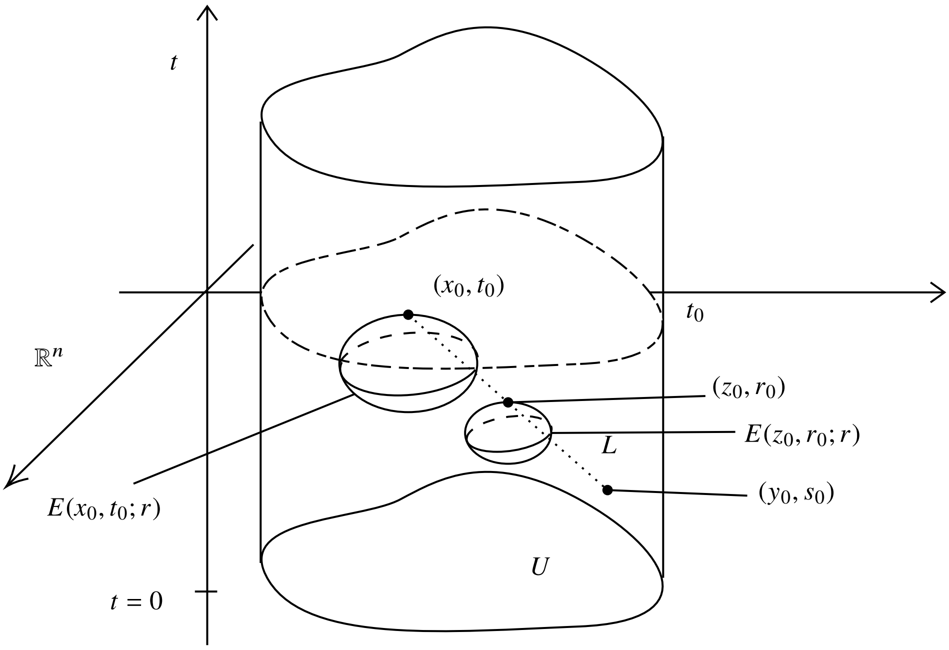


图 7.13

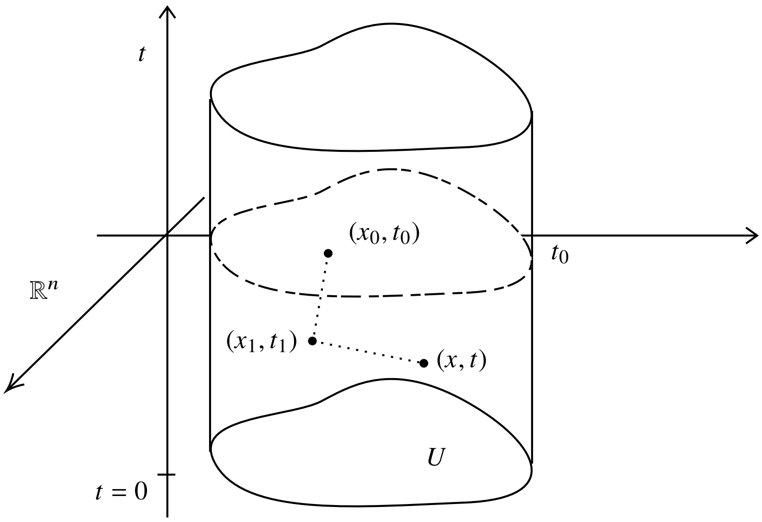


图 7.14: 某个 (x, t) 的连接示意图

其中 $g \geq 0$, 则只要 g 在某处为正, 那么 u 就在 U_T 中处处为正. 这也是对干扰的无限传播速度的另一种诠释.

最大值原理的一个重要应用就是唯一性的证明.

定理 7.3.5 (有界区域内的唯一性)

若 $g \in C(\Gamma_T), f \in C(U_T)$, 则初值/边值问题

$$\begin{cases} u_t - \Delta u = f, & \text{在 } U_T \text{ 内} \\ u = g, & \text{在 } \Gamma_T \text{ 上} \end{cases} \tag{7.110}$$

至多有一个解 $u \in C_1^2(U_T) \cap C(\overline{U_T})$.

证明

若 $u, \tilde{u}$ 是初边值问题(7.110)的两个解, 则由强最大值原理(7.3.4)知 $u - \tilde{u} \equiv 0$, 命题即证. □

下面把唯一性的断言扩充到 Cauchy 问题, 也即 $U = \mathbb{R}^n$ 时的初值问题. 因为现在区域已经无界了, 故对远处的 $|x|$ 需要做一些控制.

定理 7.3.6 (热方程 Cauchy 问题的最大值原理)

若 $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ 满足

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } \mathbb{R}^n \times (0,T) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \tag{7.111}$$

同时满足增长估计[a]:

$$u(x,t) \leq Ae^{a|x|^2}, \quad x \in \mathbb{R}^n, 0 \leq t \leq T \tag{7.112}$$

其中 A, a 是常数, $a > 0$. 则

$$\sup_{\mathbb{R}^n \times [0,T]} u = \sup_{\mathbb{R}^n} g.$$

[a]这又称为基本解的增长解

证明

总的来说, 只需证明 $u \leq \sup\limits_{\mathbb{R}^n} g$ 即可. 设

$$4aT < 1 \tag{7.113}$$

进而可以找到足够小的 $\varepsilon > 0$ 使得

$$4a(T+\varepsilon) < 1 \tag{7.114}$$

固定 $y \in \mathbb{R}^n, \mu > 0$, 定义

$$v(x,t) := u(x,t) - \frac{\mu}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{|x-y|^2}{4(T+\varepsilon-t)}}, \quad x \in \mathbb{R}^n, t > 0$$

因为 u 与具有基本解形式的 $\frac{\mu}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{|x-y|^2}{4(T+\varepsilon-t)}}$[9]均满足热方程, 故

$$v_t - \Delta v = 0, \quad \text{在 } \mathbb{R}^n \times (0,T] \text{ 内}$$

固定 $r > 0$, 取 $U := B^\circ(y,r), U_T = B^\circ(y,r) \times (0,T]$, 则根据强最大值原理(7.3.4)知

$$\max_{\overline{U_T}} v = \max_{\Gamma_T} v \tag{7.115}$$

[9]注意指数项应有的负号藏在分母处了.

现在目标是证明 $v \le \sup\limits_{\mathbb{R}^n} g$. 若 $x \in \mathbb{R}^n$, 则:

$$v(x,0) = u(x,0) - \frac{\mu}{(T+\varepsilon)^{\frac{n}{2}}} e^{\frac{|x-y|^2}{4(T+\varepsilon)}} \le u(x,0) = g(x) \tag{7.116}$$

而若 $|x-y| = r, 0 \le t \le T$, 则由(7.112)式知:

$$\begin{aligned} v(x,t) &= u(x,t) - \frac{\mu}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\varepsilon-t)}} \\ &\le Ae^{a|x|^2} - \frac{\mu}{(T+\varepsilon-t)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\varepsilon-t)}} \\ &\le Ae^{a(|y|+r)^2} - \frac{\mu}{(T+\varepsilon)^{\frac{n}{2}}} e^{\frac{r^2}{4(T+\varepsilon)}} \end{aligned} \tag{7.117}$$

现根据(7.114)式知存在 $\gamma > 0$ 使得 $\frac{1}{4(T+\varepsilon)} = a + \gamma$, 代入(7.117)式有

$$v(x,t) \le Ae^{a(|y|+r)^2} - \mu(4(a+\gamma))^{\frac{n}{2}} e^{(a+\gamma)r^2} \tag{7.118}$$

当 r 足够大, 计算可得[10]

$$Ae^{a(|y|+r)^2} - \mu(4(a+\gamma))^{\frac{n}{2}} e^{(a+\gamma)r^2} \le \sup_{\mathbb{R}^n} g \tag{7.119}$$

因而只要(7.113)式成立, 则由(7.115)-(7.119)式知:

$$v(y,t) \le \sup_{\mathbb{R}^n} g, \quad \forall y \in \mathbb{R}^n, 0 \le t \le T$$

令 $\mu \to 0$ 即得

$$u(y,t) \le \sup_{\mathbb{R}^n} g, \quad \forall y \in \mathbb{R}^n, 0 \le t \le T, 4aT < 1$$

最后, 取 $T_1 = \frac{1}{8a}$, 在 $[0,T_1], [T_1, 2T_1], \cdots$ 上重复运用上述结果即得命题. □

定理 7.3.7 (热方程 Cauchy 问题解的唯一性)

若 $g \in C(\mathbb{R}^n)$, $f \in C(\mathbb{R}^n \times [0,T])$, 则初值问题

$$\begin{cases} u_t - \Delta u = f, & 在\ \mathbb{R}^n \times (0,T)\ 内 \\ u = g, & 在\ \mathbb{R}^n \times \{t=0\}\ 上 \end{cases} \tag{7.120}$$

的满足增长估计

$$|u(x,t)| \le Ae^{a|x|^2}, \quad x \in \mathbb{R}^n, 0 \le t \le T \tag{7.121}$$

的解 $u \in C_1^2(\mathbb{R}^n \times (0,T]) \cap C(\mathbb{R}^n \times [0,T])$ 至多只有一个, 其中 A, a 是常数, $a > 0$.

证明

若 $u, \tilde{u}$ 均满足(7.120),(7.121), 则对 $w := \pm(u - \tilde{u})$ 应用最大值原理(7.3.6)即得命题. □

事实上, 初值问题

$$\begin{cases} u_t - \Delta u = 0, & 在\ \mathbb{R}^n \times (0,T)\ 中 \\ u = 0, & 在\ \mathbb{R}^n \times \{t=0\}\ 中 \end{cases} \tag{7.122}$$

是有无穷多个解的, 只是除了 $u \equiv 0$ 外, 其余的解在 $|x| \to \infty$ 时增长地都非常快.

有趣的是, 尽管 $u \equiv 0$ 肯定是初值问题(7.122)有物理意义的一个解, 这个初值问题本身还允许一些" 不具备物理意义" 的解的存在. 而定理(7.3.7)就通过加入了增长估计的条件, 限制了这些不具备物理意义的解的出现.

[10]这里没有推过去, $\sup\limits_{\mathbb{R}^n} g$ 是怎么引入的?

7.3.3.2 补充: 不依赖于均值定理的弱最大值原理证明与极值原理的更多拓展

回望之前 Laplace 方程的部分, 彼时证明调和函数的极值原理用的就是均值定理, 这也是 **[Ev]** 在热方程部分同样引入均值定理的动机: 为了凸显 Laplace 方程与热方程在某种意义上的统一性. 事实上, 调和函数的极值原理可以用不依赖于均值定理的方法证明, 这需要引入上调和函数和下调和函数[11]的概念.

定义 7.3.5 (上调和函数, 下调和函数)

称 $v \in C^2(\overline{U})$ 是上调和的, 如果在 U 内有

$$\Delta v \le 0,$$

称其为下调和的, 如果在 U 内有

$$\Delta v \ge 0.$$

♣

对于上调和函数而言, 有弱极小值原理成立:

定理 7.3.8 (上调和函数的弱极小值原理)

若 v 是有界区域 U 上的上调和函数, 则

$$\min_{\overline{U}} v = \min_{\partial U} v.$$

♡

证明

任取 $\varepsilon > 0$, 设

$$v^{\varepsilon}(x) = v(x) - \varepsilon |x|^2$$

现在说明 v^{ε} 在 U 内取不到最小值. 知:

$$Dv^{\varepsilon}(x) = (v_{x_1} - 2\varepsilon x_1, \cdots, v_{x_n} - 2\varepsilon x_n)$$

进一步有

$$D^2 v^{\varepsilon}(x) = \begin{pmatrix} v_{x_1x_1} - 2\varepsilon & v_{x_1x_2} & \cdots & v_{x_1x_n} \\ v_{x_2x_1} & v_{x_2x_2} - 2\varepsilon & \cdots & v_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{x_nx_1} & v_{x_nx_2} & \cdots & v_{x_nx_n} - 2\varepsilon \end{pmatrix}$$

注意 $\operatorname{tr} D^2 v^{\varepsilon} = \Delta v(x) - 2n\varepsilon = -2n\varepsilon < 0$, 故 $D^2 v^{\varepsilon}$ 必有负特征值, 因而其不为正定阵, 这说明 v^{ε} 在 U 内取不到最小值, 亦即

$$\min_{\overline{U}} v^{\varepsilon} = \min_{\partial U} v^{\varepsilon}$$

根据 U 的有界性知存在 $R > 0$ 使得 $U \subset B(0, R)$, 进而

$$\begin{aligned} \min_{\overline{U}} v \ge \min_{\overline{U}} v^{\varepsilon} = \min_{\partial U} v^{\varepsilon} \ge \min_{\partial U} v - \max_{\partial U} \varepsilon |x|^2 \\ \ge \min_{\partial U} v - \max_{B(0,R)} \varepsilon |x|^2 = \min_{\partial U} v - \varepsilon R^2 \end{aligned}$$

根据 ε 的任意性知

$$\min_{\overline{U}} v \ge \min_{\partial U} v$$

又由 $\partial U \subset \overline{U}$ 显见

$$\min_{\overline{U}} v \le \min_{\partial U} v$$

[11] 这个概念同样出现在了本节习题. 见习题(7.5).

故

$$\min_{\overline{U}} u = \min_{\partial U} u$$

命题即证. □

类似地, 对于下调和函数而言, 有弱极大值原理成立:

定理 7.3.9 (下调和函数的弱极大值原理)

若 v 是有界区域 U 上的下调和函数, 则

$$\max_{\overline{U}} v = \max_{\partial U} v.$$

♡

证明细节见习题(7.5). □

现在, 因为调和函数既是上调和函数又是下调和函数, 故调和函数同时满足弱极大值原理和弱极小值原理, 即得推论:

推论 7.3.1 (调和函数的弱极值原理)

若 u 是有界区域 U 上的调和函数, 则

$$\max_{\overline{U}} u = \max_{\partial U} u,$$

$$\min_{\overline{U}} u = \min_{\partial U} u.$$

♡

对热方程而言, 同样可以用这一套程序, 现在仿照上调和函数和下调和函数的概念引进热方程的上解和下解[12]:

定义 7.3.6 (热方程的上解, 下解)

设 U_T 是抛物柱, 称 $v \in C_1^2(U_T)$ 是热方程的上解, 如果

$$v_t - \Delta v \geq 0 \quad 在 \ U_T \ 内$$

称 $v \in C_1^2(U_T)$ 是热方程的下解, 如果

$$v_t - \Delta v \leq 0 \quad 在 \ U_T \ 内$$

♣

对于上解而言, 有弱极小值原理成立:

定理 7.3.10 (上解的弱极小值原理)

若 $v \in C_1^2(U_T) \cap C(\overline{U}_T)$ 是热方程的上解, 则

$$\min_{\overline{U}} v = \min_{\partial U} v.$$

♡

证明

先证明在 U_T 上 $\partial_t v - \Delta v > 0$ 的情况. 若存在 $(x_0, t_0) \in U_T$ 使得

$$v(x_0, t_0) = \min_{\overline{U}_T} v(x, t)$$

一方面, 从空间变量的角度说, 根据极值的必要条件有

$$D_x v(x, t) = 0$$

[12] 这个概念出现在本节习题(7.17)

进一步, 根据极小值的必要条件知

$$D_x^2 v(x,t) = \begin{pmatrix} v_{x_1x_1} & \cdots & v_{x_1x_n} \\ \vdots & & \vdots \\ v_{x_nx_1} & \cdots & v_{x_nx_n} \end{pmatrix}(x,t)$$

是正定阵, 这便要求 $\operatorname{tr} D_x^2 v(x,t) = \Delta v(x,t) \geq 0$. 另一方面, 从时间变量的角度说, 根据极值的必要条件有

$$\partial_t v(x_0,t_0) \begin{cases} = 0, & t_0 < T \\ \leq 0, & t_0 = T \end{cases} \Rightarrow \partial_t v(x_0,t_0) \leq 0$$

综上知 $\partial_t v(x_0,t_0) - \Delta v(x,t) \leq 0$, 这与条件矛盾! 故 U_T 内不存在使得 v 达到最小值的点, 从而 $\min\limits_{\Gamma_T} v \leq \min\limits_{\overline{U}_T} v$, 但另一方面显见 $\min\limits_{\Gamma_T} v \geq \min\limits_{\overline{U}_T} v$, 故 $\min\limits_{\Gamma_T} v = \min\limits_{\overline{U}_T} v$.

现在对于 $\partial_t v - \Delta v \geq 0$ 的情况, 任取 $\varepsilon > 0$, 构造 $v^\varepsilon(x,t) = v(x,t) + \varepsilon t$, 知 $\partial_t v^\varepsilon - \Delta v^\varepsilon > 0$, 因而根据前述结论有 $\min\limits_{\Gamma_T} v^\varepsilon = \min\limits_{\overline{U}_T} v^\varepsilon$, 从而:

$$\begin{aligned} \min_{\overline{U}_T} v &= \min_{\overline{U}_T}(v^\varepsilon(x,t) - \varepsilon t) \geq \min_{\overline{U}_T} v^\varepsilon - \max_{\overline{U}_T} \varepsilon t \\ &= \min_{\Gamma_T} v^\varepsilon - \varepsilon T \geq \min_{\Gamma_T} v - \varepsilon T \end{aligned}$$

根据 ε 的任意性知

$$\min_{\overline{U}_T} v \geq \min_{\Gamma_T} v$$

又显见 $\min\limits_{\overline{U}_T} v \leq \min\limits_{\Gamma_T} v$, 故 $\min\limits_{\overline{U}_T} v = \min\limits_{\Gamma_T} v$. 命题得证. □

类似地, 对于下解而言, 有弱极大值原理成立:

定理 7.3.11 (下解的弱极大值原理)

若 $v \in C_1^2(U_T) \cap C(\overline{U}_T)$ 是热方程的下解, 则

$$\max_{\overline{U}} v = \max_{\partial U} v.$$

♡

证明细节见习题(7.17). □

现在, 因为热方程的解既是上解又是下解, 故其同时满足弱极大值原理和弱极小值原理, 即得推论:

推论 7.3.2 (热方程解的弱极值原理)

若 $u \in C_1^2(U_T) \cap C(\overline{U}_T)$ 是热方程在 U_T 中的解, 则

$$\max_{\overline{U}_T} u = \max_{\Gamma_T} u,$$

$$\min_{\overline{U}_T} u = \min_{\Gamma_T} u.$$

♡

特别注意, 上述定理与推论之所以是弱极值原理而非强极值原理, 原因便在于它们均没有证明在连通区域内取得最大值, 则函数为常值函数, 这个缺陷是由构造本身产生的, 从而这个构造方法从本质上就无法证明强极值原理.

Laplace 方程是二阶椭圆方程的一种特殊情形. 事实上, 对于二阶椭圆方程, 弱极值原理与强极值原理依旧成立. 为此先介绍二阶椭圆方程的定义. 考虑边值问题

$$\begin{cases} Lu = f, & \text{在 } U \text{ 内} \\ u = 0, & \text{在 } \partial U \text{ 上} \end{cases} \tag{7.123}$$

其中 $U \subset \mathbb{R}^n$ 是开集, $u : \overline{U} \to \mathbb{R}$ 未知, $u = u(x)$, $f : \mathbb{U} \to \mathbb{R}$ 已知, L 表示形如下式的二阶微分算子

$$Lu = -\sum_{i,j=1}^{n}(a^{ij}(x)u_{x_i})_{x_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u \tag{7.124}$$

另一种记法为

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_ix_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} + c(x)u \tag{7.125}$$

其中系数函数 $a^{ij}, b^i, c(i,j = 1,\cdots,n)$ 给定. 一般称 L 形如(7.124)式的 PDE$Lu = f$ 为散度形式, 而称 L 形如(7.125)式的 PDE$Lu = f$ 为非散度形式. 边值问题(7.123)中 ∂U 上 $u = 0$ 的条件有时称为 Dirichlet 边值条件. 另外, 设 $(a^{ij}(x))$ 是对称的, 亦即 $a^{ij} = a^{ji}(i,j = 1,\cdots,n)$.

定义 7.3.7 ((一致) 椭圆算子)

称偏微分算子 L 是 (一致) 椭圆的, 如果存在常数 $\theta > 0$ 使得

$$\sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2 \tag{7.126}$$

对几乎处处的 $x \in U$ 与任意 $\xi \in \mathbb{R}^n$ 成立.

♣

换句话说, 算子的椭圆性意味着对任意 $x \in U$, 对称阵 $\boldsymbol{A}(x) = (a^{ij}(x))_{n\times n}$ 都是正定的, 其中最小的特征值大于等于 θ. 特别地, 当 $a^{ij} \equiv \delta_{ij}, b^i \equiv 0, c \equiv 0$ 时, $L = -\Delta$.

对于非散度形式的一些二阶椭圆 PDE, 其解同样具有弱极值原理:

定理 7.3.12 (二阶椭圆方程的弱极值原理)

设 $u \in C^2(U) \cap C(\overline{U})$, 且设在 U 内 $c \equiv 0$, 则:

(i) 若在 U 内 $Lu \leq 0$, 则 $\max\limits_{\overline{U}} u = \max\limits_{\partial U} u$.

(ii) 若在 U 内 $Lu \geq 0$, 则 $\min\limits_{\overline{U}} u = \min\limits_{\partial U} u$.

其中称满足 (i) 条件的解为下解, (ii) 条件的解为上解.

♡

证明

现设严格不等号成立, 即在 U 内有

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)u_{x_ix_j} + \sum_{i=1}^{n} b^i(x)u_{x_i} < 0$$

设存在 $x_0 \in U$ 使得

$$u(x_0) = \max_{\overline{U}} u$$

现根据内部极大值点的必要条件, 知

$$Du(x_0) = 0 \quad \text{且} \quad D^2u(x_0) \leq 0$$

根据椭圆偏微分算子的定义, 矩阵 $A = (a^{ij}(x_0))$ 是对称正定阵, 从而存在正交阵 $O = (o_{ij})$ 使得

$$OAO^T = \operatorname{diag}\{d_1,\cdots,d_n\}, \quad d_k > 0(k = 1,\cdots,n)$$

现记 $y = x_0 + O(x - x_0)$, 知 $x - x_0 = O^T(y - x_0)$, 因而

$$u_{x_i} = \sum_{k=1}^{n} u_{y_k}o_{ki}, u_{x_ix_j} = \sum_{k,l=1}^{n} u_{y_ky_l}o_{ki}o_{lj}, \quad i,j = 1,\cdots,n$$

故在 x_0 处有

$$\sum_{j=1}^{n} a^{ij}u_{x_ix_j} = \sum_{k,l=1}^{n}\sum_{i,j=1}^{n} a^{ij}u_{y_ky_l}o_{ki}o_{lj} = \sum_{n=1}^{n} d_ku_{y_ky_K} \leq 0$$

这说明在 x_0 处有

$$Lu = -\sum_{i,j=1}^{n} a^{ij} u_{x_i x_j} + \sum_{i=1}^{n} b^i u_{x_i} \geq 0$$

这与条件矛盾!

对于一般情况, 记

$$u^{\varepsilon}(x) := u(x) + \varepsilon e^{\lambda x_1}, \quad x \in U$$

其中 $\lambda > 0$ 是某常数, $\varepsilon > 0$. 回忆一致椭圆条件说明 $a^{ii}(x) \geq \theta(i = 1, \cdots, n, x \in U)$, 故在 U 内有

$$\begin{aligned} Lu^{\varepsilon} &= Lu + \varepsilon L(e^{\lambda x_1}) \leq \varepsilon e^{\lambda x_1}(-\lambda^2 a^{11} + \lambda b^1) \\ &\leq \varepsilon e^{\lambda x_1}(-\lambda^2 \theta + \|b\|_{L^\infty} \lambda) < 0 \end{aligned}$$

其中 λ 足够大. 从而 $\max\limits_{\overline{U}} u^{\varepsilon} = \max\limits_{\partial U} u$, 令 $\varepsilon \to 0$ 即知 $\max\limits_{\overline{U}} u = \max\limits_{\partial U} u$. 对于极小值的情况, 研究 $-u$ 即得欲证. □

定理 7.3.13 ($c \geq 0$ 时二阶椭圆方程的弱极值原理)

设 $u \in C^2(U) \cap C(\overline{U})$, 且在 U 内 $c \geq 0$, 则

(i) 若在 U 内 $Lu \leq 0$, 则 $\max\limits_{\overline{U}} u \leq \max\limits_{\partial U} u^+$.

(ii) 若在 U 内 $Lu \geq 0$, 则 $\min\limits_{\overline{U}} u \geq -\max\limits_{\partial U} u^-$.

证明

设 u 是下解, 取 $V := \{x \in U : u(x) > 0\}$, 则在 V 内

$$Ku := Lu - cu \leq -cu \leq 0$$

注意算子 K 没有零阶项, 因而对其运用定理(7.3.12)知 $\max\limits_{\overline{V}} u = \max\limits_{\partial V} u = \max\limits_{\partial U} u^+$, 而在 $V \neq \varnothing$ 时 $\max\limits_{\overline{U}} u = \max\limits_{\overline{V}} u$, (i) 即得证. 当 $V = \varnothing$, (i) 无需证明.

对于 (ii), 对 $-u$ 应用结论 (i) 即证. □

7.3.3.3 正则性

定理 7.3.14 (热方程解的光滑性)

若 $u \in C_1^2(U_T)$ 是热方程在 U_T 中的解, 则 $u \in C^\infty(U_T)$.

这个光滑性在边值不光滑时依旧成立.

证明

用

$$C(x, t; r) = \{(y, s) : |x - y| \leq r, t - r^2 \leq s \leq t\}$$

表示半径为 r, 高度为 r^2, 顶面圆心在 (x, t) 的闭圆柱体. 固定 $(x_0, t_0) \in U_T$, 选取 $r > 0$ 足够小使得 $C := C(x_0, t_0; r) \subset U_T$. 进一步选取 $C' := C(x_0, t_0; \frac{3}{4}r), C'' := C(x_0, t_0; \frac{1}{2}r)$, 它们有公共的顶部圆心 (x_0, t_0)(如图(7.15)).

根据实变函数的相关理论, 可以选取光滑的截断函数 $\zeta = \zeta(x, t)$ 满足

$$0 \leq \zeta \leq 1, \quad \text{在 } C' \text{ 上} \zeta \equiv 1, \quad \text{在 } C \text{ 的抛物边界附近} \zeta \equiv 0$$

再作延拓 $\zeta \equiv 0((x, t) \in \mathbb{R}^n \times [0, t_0] \backslash C)$.

暂时先设 $u \in C^\infty(U_T)$, 取

$$v(x, t) := \zeta(x, t) u(x, t), \quad x \in \mathbb{R}^n, 0 \leq t \leq t_0 \tag{7.127}$$

则计算得

$$v_t = \zeta u_t + \zeta_t u, \quad \Delta v = \zeta \Delta u + 2D\zeta \cdot Du + u \Delta \zeta$$

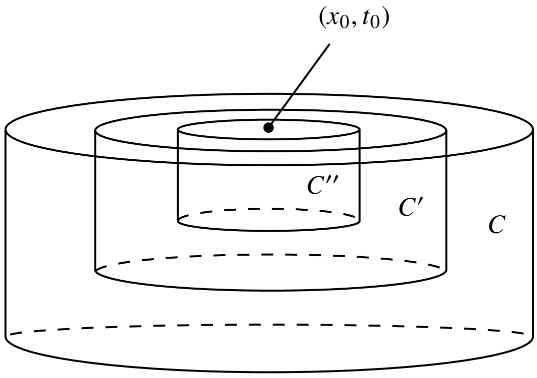


图 7.15

当 $t = 0$ 时, 根据 ζ 的构造知 $\zeta(t) \equiv 0$, 因而 $\zeta_t \equiv 0, D\zeta, \Delta\zeta \equiv 0$, 从而

$$v(x,t) = 0, \quad (x,t) \in \mathbb{R}^n \times \{t = 0\} \tag{7.128}$$

且由 u 是热方程的解知在 $\mathbb{R}^n \times (0, t_0)$ 中有

$$v_t - \Delta v = \zeta_t u - 2D\zeta \cdot Du - u\Delta\zeta =: \tilde{f} \tag{7.129}$$

只要现在可以解出 v, 根据 v 的构造即得 u 在 C' 内的表现. 注意只要 Du 可以变成 u, 则 $v_t - \Delta v$ 作为记录 v 关于时间一阶偏导与关于空间二阶偏导的量, 仅仅需要 u 连续即具有意义, 这正蕴藏着无穷可微性. 下面便尝试达成这件事. 首先求解 v, 设

$$\tilde{v}(x,t) := \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)dyds$$

根据热方程非齐次问题的求解定理(7.3.2)知

$$\begin{cases} \tilde{v}_t - \Delta\tilde{v} = \tilde{f}, & \text{在 } \mathbb{R}^n \times (0, t_0) \text{ 内} \\ \tilde{v} = 0, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 中} \end{cases} \tag{7.130}$$

因为由 u 的连续性与 ζ 的截断性知存在常数 A 使得 $|v|, |\tilde{v}| \leq A$, 故由 Cauchy 问题有界解的唯一性(7.3.7)知 $v \equiv \tilde{v}$, 也即

$$v(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y, t-s)\tilde{f}(y,s)dyds \tag{7.131}$$

下面设 $(x,t) \in C''$, 将(7.127),(7.129)式代入(7.131)式有

$$\zeta(x,t)u(x,t) = \iint_{\mathbb{R}^n \times (0,t)} \Phi(x-y, t-s)[(\zeta_s(y,s) - \Delta\zeta(y,s))u(y,s) - 2D\zeta(y,s) \cdot Du(y,s)]dyds$$

根据 ζ 的构造, ζ 只可能在 C 内不为 0, 因而上式被积函数的支集必在 C 内. 同时在 $(x,t) \in C'$ 时 $\zeta \equiv 1$, 故 $(x,t) \in C'$ 时有:

$$u(x,t) = \iint_C \Phi(x-y, t-s)[(\zeta_s(y,s) - \Delta\zeta(y,s))u(y,s) - 2D\zeta(y,s) \cdot Du(y,s)]dyds \tag{7.132}$$

现在只要设法将 $Du(y,s)$ 变成 $u(y,s)$, 根据积分对正则性的提升自然可得命题. 此时只研究 $(x,t) \in C''$ 的情况, 此时 $\zeta \equiv 1$, 故当 (y,s) 接近 $\Phi(x-y, t-s)$ 的奇点 (x,t) 的附近, 总有 $\zeta_s, \Delta\zeta, D\zeta$ 恒为零, 这说明(7.132)式方括号内的项在 Φ 的奇点 (x,t) 附近恒为零, Φ 的奇性便无需再考虑. 现在为了调整 Du, 对方括号内最后一项分部积分:

$$\begin{aligned} &\iint_C \Phi(x-y, t-s)D\zeta(y,s) \cdot Du(y,s)dyds \\ = &-\iint_C D(\Phi(x-y, t-s)D\zeta(y,s)) \cdot u(y,s)dyds + \int_{\partial C} \Phi(x-t, t-s)D\zeta(y,s) \cdot u(y,s)\sum_{i=1}^n \nu^i dyds \\ \overset{\text{(i)}}{=} &-\iint_C D_y\Phi(x-y, t-s) \cdot D\zeta(y,s) \cdot u(y,s)dyds - \iint_C \Phi(x-y, t-s) \cdot \Delta\zeta(y,s) \cdot u(y,s)dyds \end{aligned}$$

其中 (i) 前式后项边界积分中 $D\zeta \equiv 0$. 将上述结果代入(7.132)式得

$$u(x,t) = \iint_C [\Phi(x-y,t-s)(\zeta_s(y,s)+\Delta\zeta(y,s)) + 2D_y\Phi(x-y,t-s)\cdot D\zeta(y,s)]\cdot u(y,s)dyds \tag{7.133}$$

这便初步达到了优化 u 正则性的目标. 注意(7.133)还是基于 $u \in C^\infty$ 推出来的. 如果 u 仅仅满足命题的条件, 是 $C_1^2(U_T)$ 函数, 则用 $u^\varepsilon = \eta_\varepsilon * u$ 代替 u, 其中 η_ε 是关于 x,t 的磨光算子. 从 C 与诸函数的有界性, 根据 Lebesgue 控制收敛定理令 $\varepsilon \to 0$ 即得(7.133)式. 现在观察(7.133)的形式:

$$u(x,t) = \iint_C K(x,t,y,s)u(y,s)dyds, \quad (x,t) \in C'' \tag{7.134}$$

其中因为 $\zeta(y,s) \equiv 1((y,s) \in C')$, 故 $\zeta_s, \Delta\zeta, D\zeta$ 在 $(y,s) \in C'$ 时全为零, 从而

$$K(x,t,y,s) = \Phi(x-y,t-s)(\zeta_s(y,s)+\Delta\zeta(y,s)) + 2D_y\Phi(x-y,t-s)\cdot D\zeta(y,s) \equiv 0, \quad \forall (y,s) \in C'$$

得到(7.134)式被积函数的支集只可能在 $C\backslash C'$ 中, 有

$$u(x,t) = \iint_{C\backslash C'} K(x,t,y,s)u(y,s)dyds, \quad (x,t) \in C'' \tag{7.135}$$

最后, 因为 (y,s) 作为 Φ 的奇点只能在 C'' 中, 而现在积分的区域是 $C\backslash C'$, 故 Φ 在 $C\backslash C'$ 内无奇点, 因而是光滑函数. 又因为 ζ 在 $C\backslash C'$ 上也是光滑函数, 故知 K 在 $C\backslash C'$ 上光滑, 又因为(7.135)式表明 u 关于变量 (x,t) 的光滑性完全取决于 K 关于变量 (x,t) 的光滑性, 故可从 $K \in C^\infty(C'')$ 推知 $u \in C^\infty(C'')$, 命题得证. □

7.3.3.4 热方程解的局部估计

下面对热方程解的各阶导数做估计, 特别注意对 $x_i(i=1,\cdots,n)$ 的导数与对 t 的导数之间的差别, 这在物理上称作空间与时间的异性.

定理 7.3.15 (热方程解的导数估计)

对任意整数对 $k,l=0,1,\cdots$ 存在常数 C_{kl}, 使得对任意圆柱 $C(x,t;\frac{r}{2}) \subset C(x,t;r) \subset U_T$ 与热方程在 U_T 上的任一解 u 满足:

$$\max_{C(x,t;\frac{r}{2})} |D_x^k D_t^l u| \le \frac{C_{kl}}{r^{k+2l+n+2}} \|u\|_{L^1(C(x,t;r))}$$

♡

证明

为了确定圆柱 $C(x,t;r)$, 固定 $(x,t) \in U_T$, 进一步通过平移不妨设 $(x,t)=(0,0)$. 首先验证 $r=1$ 的情况, 设 $C(1) := C(0,0,1) \subset U_T$, 因而 $C(\frac{1}{2}) := C(0,0,\frac{1}{2}) \subset U_T$. 根据(7.134)式知存在光滑函数 K 满足

$$u(x,t) = \iint_{C(1)} K(x,t,y,s)u(y,s)dyds, \quad (x,t) \in C(\frac{1}{2})$$

因而

$$|D_x^k D_t^l u(x,t)| \le \iint_{C(1)} |D_t^l D_x^k K(x,t,y,s)|\cdot|u(y,s)|dyds \le \sup_{C(1)} |D_t^l D_x^k K|\cdot\|u\|_{L^1(C(1))} =: C_{kl}\|u\|_{L^1(C(1))} \tag{7.136}$$

下面讨论一般的 r, 设 $C(r) := C(0,0,r) \subset U_T$, 因而 $C(\frac{r}{2}) := C(0,0,\frac{r}{2}) \subset U_T$. 要化归到 $r=1$ 的情况, 只需令

$$v(x,t) := u(rx, r^2t)$$

现在在 $C(1)$ 内有 $v_t - \Delta v = 0$. 由(7.136)式知

$$|D_x^k D_t^l v(x,t)| \le C_{kl}\|v\|_{L^1(C(1))}, \quad (x,t) \in C(\frac{1}{2})$$

计算知

$$D_x^k D_t^l v(x,t) = r^{2l+k} D_x^k D_t^l u(rx, r^2t), \quad \|v\|_{L^1(C(1))} = \frac{1}{r^{n+2}}\|u\|_{L^1(C(r))}$$

故

$$\max_{C(\frac{r}{2})} |D_x^k D_t^l u| \le \frac{C_{kl}}{r^{2l+k+n+2}}\|u\|_{L^1(C(r))}$$

□

可以说明, 如果 u 是热方程在 U_T 内的解, 则对每个时间 $0<t\le T$ 而言, 映射 $x\mapsto u(x,t)$ 都是解析的. 不过映射 $t\mapsto u(x,t)$ 未必全局解析 (亦即在 $[0,+\infty)$ 上解析).

课堂笔记 (2023.10.30)

- (关于估计阶数的 scaling 分析[a]) 估计式中的 $\frac{1}{r^{k+2l+n+2}}$ 中的指数看似极其抽象, 下面探讨设置该指数的动机. 如果 $u(x,t)$ 是热方程的解, 从热方程的线性性知对任意 $\lambda>0$ 而言, $u_\lambda(x,t):=u(\lambda x,\lambda^2 t)$ 也应该是热方程的解. 所以如果估计式对 u 成立, 那它对 u_λ 也应该成立. 注意估计的导数是 $D_x^k D_t^l u$, 现在对 u_λ 取对应导数: D_x^k 表示对 x 分量求 k 阶导, 根据链式法则知 $D_x^k(u_\lambda(x,t))=\lambda^k(D_x^k u)(\lambda x,\lambda^2 t)$. 同样, 应用 D_t^l 可以进一步得到 $D_t^l(D_x^k u_\lambda(x,t))=\lambda^{k+2l}(D_t^l D_x^k u)(\lambda x,\lambda^2 t)$, 这便得到估计式左边形如 $\lambda^{k+2l}D_t^l D_x^k u$. 对估计式右边, 形式上计算 u_λ 的 L^1 范数知
$$\begin{aligned}\int|u_\lambda(x,t)|dxdt&=\int|u(\lambda x,\lambda^2 t)|dxdt=\int\frac{1}{\lambda^n}|u(y,\lambda^2 t)|dydt\\&=\int\frac{1}{\lambda^n}\frac{1}{\lambda^2}|u(y,s)|dyds=\frac{1}{\lambda^{n+2}}\int|u(y,s)|dyds\end{aligned}$$
从而估计式右边形如 $\frac{1}{\lambda^{n+2}}\|u\|_{L^1}$, 最后估计式整体便形如
$$\lambda^{k+2l}|D_t^l D_x^k u|\le\frac{1}{\lambda^{n+2}}\|u\|_{L^1}\Leftrightarrow|D_t^l D_x^k u|\le\frac{1}{\lambda^{k+2l+n+2}}\|u\|_{L^1}$$
得到估计式右式的指数.

[a]徐桂香老师一般直称 scaling 分析 (scaling analysis), 苗长兴译著中表述为" 尺度变换技术".

7.3.4 能量方法

7.3.4.1 唯一性

现在再来讨论初边值问题
$$\begin{cases}u_t-\Delta u=f, & \text{在 } U_T \text{ 内}\\ u=g, & \text{在 } \Gamma_T \text{ 上}\end{cases}\tag{7.137}$$
前面已经介绍了初边值问题解的唯一性可以由最大值原理推知, 而在 Laplace 方程的能量方法部分, 我们用分部积分法重新证明了一次边值问题解的唯一性. 现在对热方程作类似操作, 设 $U\subset\mathbb{R}^n$ 是有界开集, ∂U 是 C^1 的. 终止时间 $T>0$ 给定.

定理 7.3.16 (唯一性)

初边值问题(7.137)至多只有一个解 $u\in C_1^2(\overline{U_T})$.

证明

若 $\tilde{u}$ 是初边值问题(7.137)的另一个解, 则 $w:=u-\tilde{u}$ 是
$$\begin{cases}w_t-\Delta w=0, & \text{在 } U_T \text{ 内}\\ w=0, & \text{在 } \Gamma_T \text{ 上}\end{cases}\tag{7.138}$$
的解. 取
$$e(t):=\int_U w^2(x,t)dx,\quad 0\le t\le T$$
则
$$\frac{de}{dt}=\dot{e}(t)=2\int_U ww_t dx=2\int_U w\Delta w dx=-2\int_U|Dw|^2dx\le 0$$
故
$$e(t)\le e(0)=0,\quad 0\le t\le T$$

这说明在 U_T 上 $w = u - \tilde{u} \equiv 0$. □

可以观察到, 上述证明正是在 Laplace 方程部分中的证明加入了时间变量, 大致流程都是一样的.

7.3.4.2 反向唯一性

热方程的一个更微妙的特点在于其在时间上的反向唯一性. 设 $u, \tilde{u}$ 都是热方程在 U_T 上的光滑解, 边值条件相同:

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } U_T \text{ 内} \\ u = g, & \text{在 } \partial U \times [0,T] \text{ 上} \end{cases} \tag{7.139}$$

$$\begin{cases} \tilde{u}_t - \Delta \tilde{u} = 0, & \text{在 } U_T \text{ 内} \\ \tilde{u} = g, & \text{在 } \partial U \times [0,T] \text{ 上} \end{cases} \tag{7.140}$$

其中 g 是某个固定的函数. 特别注意此时并没有设在 $t = 0$ 处 $u = \tilde{u}$.

定理 7.3.17 (反向唯一性)

若 $u, \tilde{u} \in C^2(\overline{U_T})$ 是(7.139),(7.140)的解, 且

$$u(x,T) = \tilde{u}(x,T), \quad x \in U$$

则在 U_T 内 $u \equiv \tilde{u}$.

换句话说, 如果 U 上的两个温度分布在某一刻 $T > 0$ 相同了, 且在 $0 \le t \le T$ 时还有相同的边值条件, 那么这两个温度分布曾经在 U 上必定时时刻刻处处相等.

证明

记 $w = u - \tilde{u}$, 类似于在用能量方法证明唯一性时的操作, 取

$$e(t) := \int_U w^2(x,t)dx, \quad 0 \le t \le T$$

将条件(7.139),(7.140)代入知

$$\frac{de}{dt} = \dot{e}(t) = -2\int_U |Dw|^2 dx \tag{7.141}$$

进一步求导有

$$\ddot{e}(t) = -4\int_U Dw \cdot Dw_t dx = 4\int_U \Delta w \cdot w_t dx = 4\int_U (\Delta w)^2 dx \tag{7.142}$$

既然现在在 ∂U 上 $w = 0$, 有

$$\int_U |Dw|^2 dx = -\int_U w\Delta w dx \le (\int_U w^2 dx)^{\frac{1}{2}} \cdot (\int_U (\Delta w)^2 dx)^{\frac{1}{2}}$$

从而由(7.141),(7.142)式知:

$$(\dot{e}(t))^2 = 4(\int_U |Dw|^2 dx)^2 \le (\int_U w^2 dx)(4\int_U (\Delta w)^2 dx) = e(r)\ddot{e}(t)$$

故

$$\ddot{e}(t)e(t) \ge (\dot{e}(t))^2, \quad 0 \le t \le T \tag{7.143}$$

只要能说明在 $0 \le t \le T$ 上 $e(t) = 0$, 就能说明 $w \equiv 0$ 了. 用反证法, 假设存在区间 $[t_1, t_2] \subset [0,T]$ 使得

$$e(t) > 0(t_1 \le t < t_2), \quad e(t_2) = 0 \tag{7.144}$$

记

$$f(t) := \ln e(t), \quad t_1 < t < t_2 \tag{7.145}$$

则由(7.143)式知

$$\ddot{f}(t) = \frac{\ddot{e}(t)}{e(t)} - \frac{\dot{e}(t)^2}{e(t)^2} \geq 0$$

故 f 在 (t_1, t_2) 上是凸函数, 根据凸函数的定义知任取 $0 < \tau < 1, t_1 < t < t_2$ 有

$$f((1-\tau)t_1 + \tau t) \leq (1-\tau)f(t_1) + \tau f(t)$$

将(7.145)式代入知

$$e((1-\tau)t_1 + \tau t) \leq e(t_1)^{1-\tau} e(t)^{\tau}$$

故

$$0 \leq e((1-\tau)t_1 + \tau t_2) \leq e(t_1)^{1-\tau} e(t_2)^{\tau}, \quad 0 < \tau < 1$$

但根据假设, $e(t_2) = 0$, 这说明任取 $t_1 \leq t \leq t_2$ 都有 $e(t) = 0$, 矛盾! 命题即证. □

7.4 波方程

本节探讨合适初边值条件下的波方程

$$u_{tt} - \Delta u = 0 \tag{7.146}$$

与非齐次波方程

$$u_{tt} - \Delta u = f \tag{7.147}$$

其中 $t > 0, x \in U, U \subset \mathbb{R}^n$ 是开的. $u = u(x,t) : \overline{U} \times [0, \infty) \to \mathbb{R}$ 未知, Laplace 算子是对空间变量 $x = (x_1, \cdots, x_n)$ 取的. 在(7.147)中函数 $f : U \times [0, \infty) \to \mathbb{R}$ 给定. 一般把波方程简写为下述 d'Alembert 记号:

$$\Box u := u_{tt} - \Delta u$$

很快就会发现波方程的解与 Laplace 方程和热方程有很大的不同, 比如这些解一般都不是 C^∞ 的, 且有有限传播速度等等.

波方程本身是振动弦 $(n = 1)$, 薄膜 $(n = 2)$ 或弹性体 $(n = 3)$ 的简化模型. 在这些物理情景中, $u(x,t)$ 表示点 x 在时间 $t \geq 0$ 时朝某方向的位移.

设 V 表示 U 的任意光滑子区域, V 内的加速度为

$$\frac{d^2}{dt^2} \int_V u dx = \int_V u_{tt} dx$$

V 的净接触力为

$$-\int_{\partial V} \boldsymbol{F} \cdot \boldsymbol{\nu} dS$$

其中 $\boldsymbol{F}$ 是通过 ∂V 施加在 V 上的力, 取质量密度为单位密度. Newton 第二定律表明质量与加速度之积就是净力, 因而:

$$\int_V u_{tt} dx = -\int_{\partial V} \boldsymbol{F} \cdot \boldsymbol{\nu} dS$$

根据高-奥公式(6.2.1)知

$$\int_{\partial V} \boldsymbol{F} \cdot \boldsymbol{\nu} dS = \int_V \operatorname{div} \boldsymbol{F} dx$$

再根据 V 的任意性知

$$u_{tt} = -\operatorname{div} \boldsymbol{F}$$

对于弹性体而言, $\boldsymbol{F}$ 是位移梯度 Du 的函数, 也就是说

$$u_{tt} + \operatorname{div} \boldsymbol{F}(Du) = 0$$

对足够小的位移梯度 Du, 考虑近似 $\boldsymbol{F}(Du) \approx -aDu$, 得到

$$u_{tt} - a\Delta u = 0$$

当 $a = 1$ 时这就是波方程.

下面举一个来自 **[ST1]** 的一维波方程例子更直观地来介绍上面的过程. 设有一根材质均匀的弦放置在 (x, y)-平面上, 弦自身贴合 x 轴, 起点和终点分别为 $x = 0, x = L$. 当弦开始振动, 弦上位置 x 处的点在 t 时刻相对于 $t = 0$ 时刻的位移 y 就构成一个关于 x, t 的函数: $y = u(x, t)$. 现在的目的是找出 y 满足的 PDE.

高中物理已经学过微元法, 这里就是其用武之地. 固定时刻 t, 把这根弦等分成 N 份, 并把每一份都视作独立的质点 $x_n = \frac{nL}{N}$, 这就把连续的质量分布转变成了离散的质量分布, 进而可以考虑这条" 项链" 中每一个质点的运动. 注意每个质点都仅在竖直方向振荡, 而每个质点的运动理所当然是与它相邻的两个质点有关的.

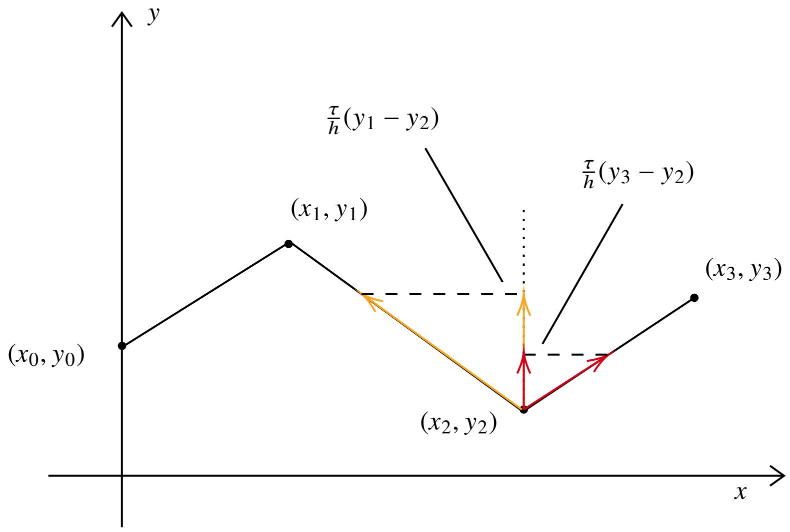


图 7.16

令 $y_n(t) = u(x_n, t)$, 记 $h = \frac{L}{N}$, 显见 $x_{n+1} - x_n = h$, 进而若设弦有恒定密度 ρ, 自然可分配每个质点的质量为 ρh. 根据 Newton 第二定律, 作用在 x_n 上的力就是 $\rho h y_n''(t)$, 其中 $y_n''(t)$ 自然是 x_n 垂直方向的加速度. 如图(7.16), 这个力由来自 x_{n-1} 与 x_{n+1} 两个质点生成的张力合成. 设右端的张力为 $\frac{\tau}{h}(y_{n+1} - t_n)$, $\tau > 0$ 是弦的张力系数. 同理可设左端张力为 $\frac{\tau}{h}(y_{n-1} - y_n)$, 从而在力的关系上有:

$$\rho h y_n''(t) = \frac{\tau}{h}(y_{n+1}(t) + y_{n-1}(t) - 2y_n(t)) \tag{7.148}$$

把 $y_n(t) = u(x_n, t)$ 代入有

$$y_{n+1}(t) + y_{n-1}(t) - 2y_n(t) = u(x_n + h, t) + u(x_n - h, t) - 2u(x_n, t) \tag{7.149}$$

如果 $u(x, t)$ 对 x 足够光滑, 知

$$\frac{1}{h^2}(u(x_n + h, t) + u(x_n - h, t) - 2u(x_n, t)) \to \frac{\partial^2 u}{\partial x^2}(x_n, t), \quad h \to 0 \tag{7.150}$$

把(7.149),(7.150)式代入(7.148)式即得:

$$\rho \frac{\partial^2 u}{\partial t^2} = \tau \frac{\partial^2 u}{\partial x^2}$$

最后对自变量作相应伸缩即得波方程(7.146).

波方程与连续介质动力学基本方程 (Euler 方程) 也有联系, 下面介绍 **[Zo]** 引入波方程的方法, 为此先介绍连续性方程与 Euler 方程.

设 $\mathbb{R}^3$ 中充满了某种介质, $\rho = \rho(x, y, z, t)$ 是介质的密度, $\boldsymbol{v} = \boldsymbol{v}(x, y, z, t)$ 是介质的速度场. D 是 $\mathbb{R}^3$ 中的区域,

边界为 ∂D. 根据密度的定义立即知道, 在一段时间 Δt 内, D 内的物质质量增加了

$$\iiint_D (\rho(x,y,z,t+\Delta t)-\rho(x,y,z,t))dV$$

在 Δt 内, 物质经过 ∂D 向 ∂D 外法线方向流出的流量 (在 $o(\Delta t)$ 的误差内) 为

$$\Delta t\cdot\iint_{\partial D}\rho\boldsymbol{v}\cdot d\boldsymbol{\sigma}$$

现在如果在 D 内没有源和汇, 则根据质量守恒得到

$$\iiint_D \Delta\rho dV=-\Delta t\iint_{\partial D}\rho\boldsymbol{v}\cdot d\boldsymbol{\sigma},\quad \Delta\rho=\rho(x,y,z,t+\Delta t)-\rho(x,y,z,t)$$

也即

$$\iiint_D\frac{\Delta\rho}{\Delta t}dV=-\iint_{\partial D}\rho\boldsymbol{v}\cdot d\boldsymbol{\sigma}\overset{\Delta t\to 0}{\Rightarrow}\iiint_D\frac{\partial\rho}{\partial t}dV=-\iint_{\partial D}\rho\boldsymbol{v}\cdot d\boldsymbol{\sigma}$$

对上右式应用高-奥公式得到

$$\iint_D\frac{\partial\rho}{\partial t}dV=-\iiint_D\operatorname{div}(\rho\boldsymbol{v})dV$$

由 D 的任意性得到

$$\frac{\partial\rho}{\partial t}=-\operatorname{div}(\rho\boldsymbol{v})\tag{7.151}$$

这就是连续介质的连续性方程.

再来说明连续介质在空间中运动的动力学方程. 除了连续性方程中的密度 ρ, 速度 $\boldsymbol{v}$ 外, 这里另外还需考虑压力 $p=p(x,y,z,t)$. 同样设 $D\subset\mathbb{R}^n$ 是区域, 现在考察在固定时刻作用在 D 中介质的体积上的力.

某些力场 (比如重力场或电场) 可以作用在介质的每个质量元素 ρdV 上, 这种场提供了所谓的质量力. 设 $\boldsymbol{F}=\boldsymbol{F}(x,y,z,t)$ 是由一些外场产生的质量力密度, 此时这些场作用在质量元 ρdV 上的力就是 $\boldsymbol{F}\rho dV$. 如果这个质量元在考察的这一时刻还有加速度 $\boldsymbol{a}$, 则根据 Newton 定律知这个质量元上还存在惯性质量力 $-\boldsymbol{a}\rho dV$. 最后在曲面 ∂D 的每个面元 $d\boldsymbol{\sigma}=\boldsymbol{n}d\sigma$($\boldsymbol{n}$ 为 ∂D 的外法线) 上, 还存在由压力引起的面力 $-pd\boldsymbol{\sigma}$, 它来自与 ∂D 贴合的那些介质粒子 (如图(7.17)).

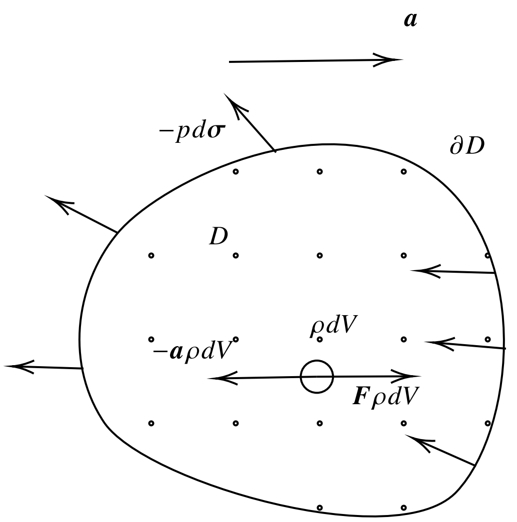


图 7.17

根据 d'Alembert 原理, 任何物质系统, 在其运动的每一瞬间, 作用在它上面的全部力 (包括惯性力) 是互相平衡的, 也就是说它们的合力为零. 此即

$$\iiint_D(\boldsymbol{F}-a)\rho dV-\iint_{\partial D}\rho d\boldsymbol{\sigma}=0$$

上式第一项是质量力与惯性力的合力, 第二项是作用在考察区域边界 ∂D 上的压力的合力. 简单起见, 我们认为现

在处理的流体是理想 (无粘性) 的, 这说明压力 p 与空间内面元的方向无关. 对第二项应用高-奥公式知

$$\iiint_D (\boldsymbol{F}-\boldsymbol{a})\rho dV - \iiint_D \boldsymbol{\nabla} p dV = 0$$

由 D 的任意性知

$$\rho \boldsymbol{a} = \rho \boldsymbol{F} - \boldsymbol{\nabla} p \tag{7.152}$$

这与物质质点运动的 Newton 方程是相当的, 其中 $\boldsymbol{a}$ 是质点速度 $\boldsymbol{v}$ 的导数 $\frac{d\boldsymbol{v}}{dt}$. 另设空间中粒子运动的规律为 $x=x(t), y=y(t), z=z(t)$, 则

$$\boldsymbol{a} = \frac{d\boldsymbol{v}}{dt} = \frac{\partial \boldsymbol{v}}{\partial t} + \frac{\partial \boldsymbol{v}}{\partial x}\frac{dx}{dt} + \frac{\partial \boldsymbol{v}}{\partial y}\frac{dy}{dt} + \frac{\partial \boldsymbol{v}}{\partial z}\frac{dz}{dt} = \frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$$

这样一来, 在(7.152)式两边同乘 $\frac{1}{\rho}$, 并将上式代入得到

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = \boldsymbol{F} - \frac{1}{\rho}\nabla p \tag{7.153}$$

这便是流体动力学的 Euler 方程.

最后回到波动方程, 现在考虑介质中有声波传播, 显见这个运动也符合 Euler 方程(7.153). 声音是介质状态的疏密交替, 且在音波中, 压力偏离其平均值的量非常小, 所以声音是介质的体元素偏离其平衡位置很小的小速度运动. 根据物理事实, 可以把气体中的声音运动看作是在其平衡位置附近的绝热的小震动. 根据这个事实, 因为每个质点的宏观速度 $\boldsymbol{v}$ 本身很小, 所以把(7.153)式中的 $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}$ 忽略即得

$$\rho\frac{\partial v}{\partial t} = \rho \boldsymbol{F} - \nabla p$$

注意

$$\rho\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\partial \rho}{\partial t}\boldsymbol{v} = \frac{\partial}{\partial t}(\rho \boldsymbol{v})$$

根据同样的理由忽略 $\frac{\partial\rho}{\partial t}\boldsymbol{v}$ 得到

$$\frac{\partial}{\partial t}(\rho \boldsymbol{v}) = \rho \boldsymbol{F} - \nabla p$$

把关于 x, y, z 的算子 ∇ 作用在式子两端得到

$$\frac{\partial}{\partial t}(\nabla\cdot\rho \boldsymbol{v}) = \nabla\cdot\rho \boldsymbol{F} - \Delta p$$

将连续性方程(7.151)代入, 并记 $\nabla\cdot\rho\boldsymbol{F} = -\Phi$, 得

$$\frac{\partial^2 \rho}{\partial t^2} = \Phi + \Delta p$$

注意 Φ 衡量的是外部的力场的影响, 如果外场影响可忽略, 那么上式即

$$\frac{\partial^2 \rho}{\partial t^2} = \Delta p \tag{7.154}$$

注意前面说过声波的传播过程是绝热的, 这说明状态方程 $f(p,\rho,T)=0$ 与时间 T 无关, 进而存在关系式 $\rho=\psi(p)$, 对 t 求二阶偏导有

$$\frac{\partial^2 \rho}{\partial t^2} = \psi'(p)\frac{\partial^2 p}{\partial t^2} + \psi''(p)(\frac{\partial p}{\partial t})^2$$

因为声波中压力的振动很小, 进而可以认为 $\psi'(p)\equiv\psi'(p_0)$, 其中 p_0 是平衡压力. 这说明 $\psi''=0$, 得到 $\frac{\partial^2\rho}{\partial t^2} \approx \psi'(p)\frac{\partial^2 p}{\partial t^2}$, 代入(7.154)式得

$$\frac{\partial^2 p}{\partial t^2} = a^2\Delta p$$

其中 $a=(\psi'(p_0))^{-\frac{1}{2}}$. 它刻画在声音运动状态下, 介质中的压力变化情况. 当然, 在受迫振动情形, 有某种力作用于介质的每一个体元素, 此时 Φ 不能忽略, 得到方程

$$\frac{\partial^2 p}{\partial t^2} = a^2\Delta p + f$$

当 $f\equiv 0$ 时这就是非齐次波动方程.

7.4.1 利用球均值求解

前面在研究 Laplace 方程与热方程时, 采用的方法都是寻找不变量构造基本解. 但对于波方程而言, 我们首先直接推导 $n = 1$ 的情况, 再用球均值的方法求解 $n \geq 2$ 的情况.

7.4.1.1 $n = 1$ 时的解, d'Alembert 公式

首先关注 $\mathbb{R}$ 上的一维波初值问题:

$$\begin{cases} u_{tt} - u_{xx} = 0, & \text{在 } \mathbb{R} \times (0, \infty) \text{ 内} \\ u = g, u_t = h, & \text{在 } \mathbb{R} \times \{t = 0\} \text{ 上} \end{cases} \tag{7.155}$$

其中 g, h 给定. 希望导出 u 关于 g, h 的表达式.

首先注意(7.155)中的 PDE 能被" 分解" 成:

$$(\frac{\partial}{\partial t} + \frac{\partial}{\partial x})(\frac{\partial}{\partial t} - \frac{\partial}{\partial x})u = u_{tt} - u_{xx} = 0 \tag{7.156}$$

记

$$v(x, t) := (\frac{\partial}{\partial t} - \frac{\partial}{\partial x})u(x, t) \tag{7.157}$$

则(7.156)式即

$$v_t(x, t) + v_x(x, t) = 0, \quad x \in \mathbb{R}, t > 0$$

这正是前面介绍的常系数输运方程, 不难解得

$$v(x, t) = a(x - t) \tag{7.158}$$

其中 $a(x) := v(x, 0)$. 联立(7.156)-(7.158)式知在 $\mathbb{R} \times (0, \infty)$ 中有

$$u_t(x, t) - u_x(x, t) = a(x - t)$$

这便回到了非齐次输运方程, 不难解得

$$u(x, t) = \int_0^t a(x + (t - s) - s)ds + b(x + t) = \frac{1}{2}\int_{x-t}^{x+t} a(y)dy + b(x + t) \tag{7.159}$$

其中 $b(x) := u(x, 0)$.

最后代入(7.155)的初值条件来求解 a, b. 在 $\mathbb{R} \times \{t = 0\}$ 上 $u = g$ 说明

$$b(x) = g(x), \quad x \in \mathbb{R}$$

同理, $u_t = h$ 说明

$$a(x) = v(x, 0) = u_t(x, 0) - u_x(x, 0) = h(x) - g'(x), \quad x \in \mathbb{R}$$

代入(7.159)式有

$$u(x, t) = \frac{1}{2}\int_{x-t}^{x+t} (h(y) - g'(y))dy + g(x = t)$$

整理即得

$$u(x, t) = \frac{1}{2}[g(x + t) + g(x - t)] + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy, \quad x \in \mathbb{R}, t \geq 0 \tag{7.160}$$

这便是 d'Alembert 公式.

上面的推导假设了 u 是初值问题(7.155)的 (足够光滑) 的解, 下面验证(7.160)式确实是它的解.

定理 7.4.1 (波方程的解, $n = 1$)

若 $g \in C^2(\mathbb{R}), h \in C^1(\mathbb{R}), u$ 按 d'Alembert 公式(7.160)定义, 则

1. $u \in C^2(\mathbb{R} \times [0, \infty))$,
2. 在 $\mathbb{R} \times (0, \infty)$ 中有 $u_{tt} - u_{xx} = 0$,

3. 对 $x^0 \in \mathbb{R}$ 有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ t>0}} u(x,t) = g(x^0)$, $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ t>0}} u_t(x,t) = h(x^0)$.

证明通过直接计算即可. □

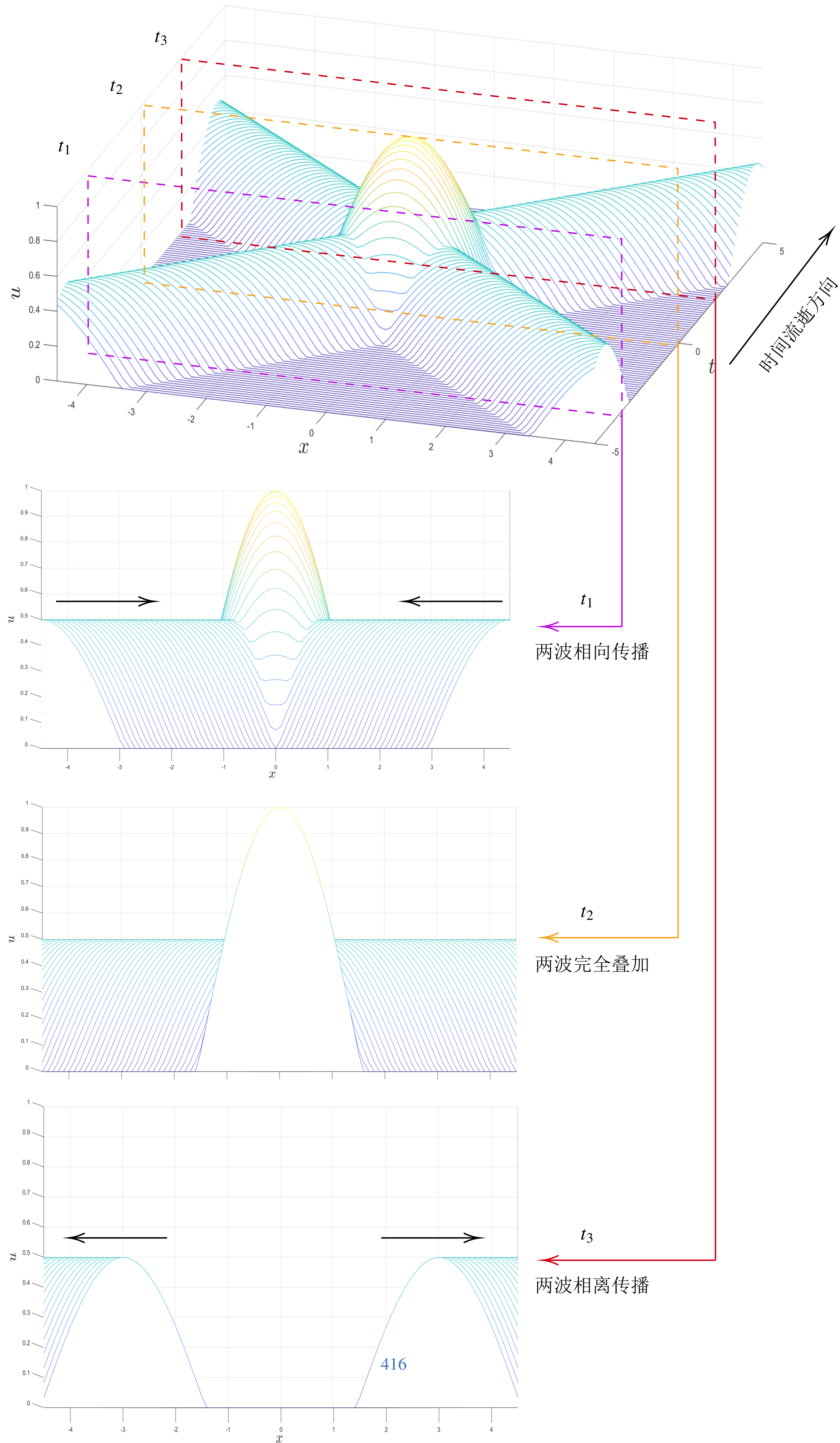
注

1. 在(7.160)式中, 解 u 具有形式
$$u(x,t) = F(x+t) + G(x-t)$$
其中 F, G 是适当的函数. 反之, 任何有这种形式的函数都是 $u_{tt} - u_{xx} = 0$ 的解. 这说明一维波方程的一般解就是 $u_t - u_x = 0$ 与 $u_t + u_x = 0$ 的一般解之和, 这也是分解(7.156)的结论.
2. 从(7.160)式中知若 $g \in C^k, h \in C^{k-1}$, 则 $u \in C^k$. 注意不同于前面的 Laplace 方程与热方程, u 在这里并没有太好的光滑性, 所以波方程对初值并不具有磨光化的作用.

课堂笔记 (2023.11.2)

- (关于一维波方程的一些其它处理思路)
  1. 如果记 $v(x,t) = \partial_t u(x,t)$, 则 $\partial_t v(x,t) = \partial_t^2 u(x,t) = \Delta u(x,t)$. 在形式上这便得到关系式:
  $$\partial_t \begin{pmatrix} u(x,t) \\ v(x,t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} u(x,t) \\ v(x,t) \end{pmatrix}$$
  这便将波方程的求解转化成了研究算子 $\begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}$ 的性质.
  2. [SA] 中提出研究求导算子矩阵的想法. 注意
  $$\begin{aligned} \partial_t \begin{pmatrix} \partial_x u \\ \partial_t u \end{pmatrix} &= \begin{pmatrix} \partial_x \partial_t u \\ \partial_t^2 u \end{pmatrix} = \begin{pmatrix} \partial_x \partial_t u \\ \partial_x^2 u \end{pmatrix} = \partial_x \begin{pmatrix} \partial_t u \\ \partial_x u \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_x \begin{pmatrix} \partial_x u \\ \partial_t u \end{pmatrix} = \begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix} \begin{pmatrix} \partial_x u \\ \partial_t u \end{pmatrix} \end{aligned}$$
  从而问题转化为研究算子 $\begin{pmatrix} 0 & \partial_x \\ \partial_x & 0 \end{pmatrix}$ 的性质.

注 因为模板的定理环境不便于画图, 这里单开一注说明一维情况的影响区域, 决定区域和依赖区域. 为了便于理解, 现设初速度 $h \equiv 0$. 首先介绍瀑布图的看法, 设
$$g(x) = \begin{cases} \cos x, & x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ 0, & x \notin [-\frac{\pi}{2}, \frac{\pi}{2}] \end{cases}$$
依照 d'Alembert 公式(7.160)画出的图如下页所见.


t_3
t_2
t_1
u
x
t
时间流逝方向
t_1
两波相向传播
t_2
两波完全叠加
t_3
两波相离传播

上图中第一幅图即 $u(x,t)$ 的瀑布图, t_1,t_2,t_3 表示不同的时刻. 在每一个时刻, 沿 x 方向对图像作切割 (如第一幅图虚线平面), 得到下三幅图所示的切面. 形象上, 可以把瀑布图视作一个过程在时间上的切片所形成的图像. 比如上图中下三幅图表示同两列波的传播过程, 这一整个随时间变化的过程被记录到了第一幅图中.

下面说明几个区域的意义与关系. 设

$$g(x)=\begin{cases}1, & x=x_0\\ 0, & x\neq x_0\end{cases}$$

现在希望研究 $u(x,t)$ 在 $t>0$ 时的表现. 根据 d'Alembert 公式, 此时有

$$u(x,t)=\frac{1}{2}(g(x+t)+g(x-t))$$

当 (x,t) 满足 $x+t=x_0$ 时, 注意 $u(x,t)=\frac{1}{2}(g(x_0)+g(x-t))=\frac{1}{2}(1+g(x-t))$, 在 $x-t\neq x_0$ 时即得 $u=\frac{1}{2}$. 同理, 当 $x-t=x_0, x+t\neq x_0$ 时也有 $u=\frac{1}{2}$, 这便确定了 u 在 $\mathbb{R}\times(0,+\infty)$ 中两条线 $x+t=x_0, x-t=x_0$ 上的取值. 而对于不在前述两条线中任意一条上的点, 因为 g 此时为 0, 故 u 只能在该点处为 0. 总而言之, 这样设置的 g 说明了 x 轴上一个点信号在何处对波产生影响: 只要后续研究的点 (x,t) 落在前述直线上, 根据前面说明的过程, 这个点信号就可以传到 (x,t) 处. 图见(7.18), 其中红色直线是点信号在往后的时间中传播的区域, 称为 $g(x)$ 在点 x_0 处的影响区域.

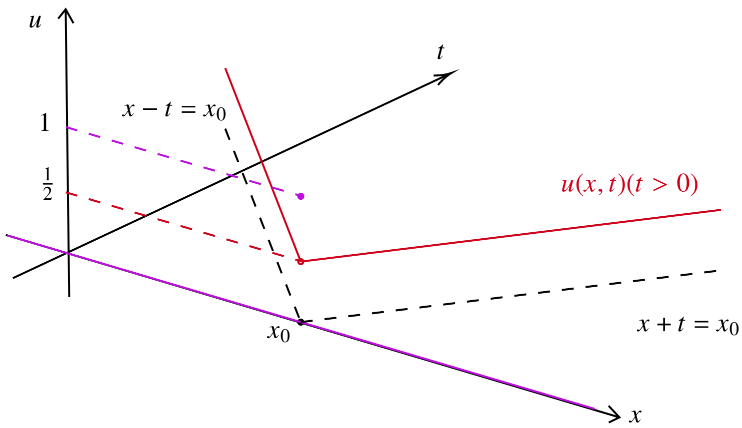


图 7.18: 影响区域示意图

再设

$$g(x)=\begin{cases}1, & x\in[x_1,x_2]\\ 0, & x\notin[x_1,x_2]\end{cases}$$

注意 $g(x)$ 是由一系列单点函数所组成的. 同时研究这样一群单点函数的影响区域, 如图(7.19):

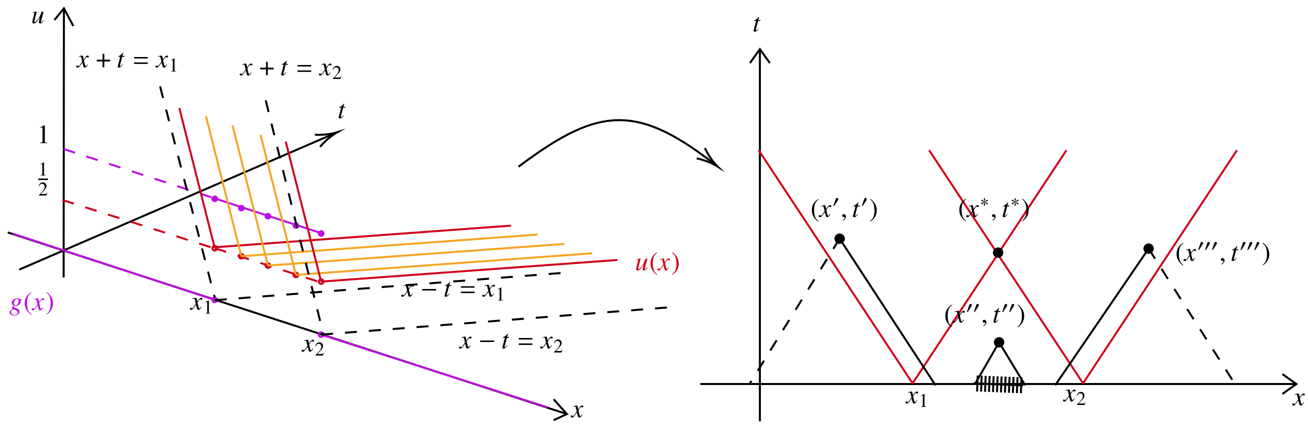


图 7.19: 影响区域与决定区域示意图

图(7.19)中左幅表示前设 $g(x)$ 对应的瀑布图. 在研究 $g(x)$ 在 $[x_1, x_2]$ 上的影响区域时, 便可将其视作一系列点信号 (黄色点) 的影响区域 (黄色线) 的总和, 得到右幅所表示的区域, 该区域即为 $g(x)$ 在区间 $[x_1, x_2]$ 上的影响区域. 在右幅中, (x', t') 点只能接收到来自 $g(x)$ 的向左传播的信号 (因为 $g(x)$ 在 $[x_1, x_2]$ 外都是 0, 这就说明 (x', t') 上虚线所对应的点没有信号, 但 (x', t') 作为单点, 能接收到的信号只有来自 $[x_1, x_2]$ 的, 实线对应的, 随着时间流逝而向左传播的信号, 与虚线对应的, (如果有的话) 随着时间流逝向右传播的信号). 同理, (x'', t'') 可以同时接收到向左传播与向右传播的信号, 而 (x''', t''') 只能接收到向右传播的信号. 现在对一个确定的点 $(\widetilde{x}, \widetilde{t})$ 来说, 它所可能接收到的信号, 只能来自于过它的两条直线 $x - t = \widetilde{x} - \widetilde{t}, x + t = \widetilde{x} + \widetilde{t}$ 与 x 轴相交的两点 (与两点之间的部分), 这一部分区间就称为该点的依赖区域. 譬如对于 (x'', t''), 它所接受的信号就只能来自于右幅图中 x 轴上标污的部分, 从而 x 轴上的这一部分区间就称为 (x'', t'') 的依赖区域. 特别地, 针对 $g(x)$ 而言, 只要时空图 xOt 中的点落在 $x_1, x_2, (x^*, t^*)$ 三点连成的三角形中, 该点所接收到的信号就只能来自于 $[x_1, x_2]$ 之间的部分, 亦即这一区域中的点接收到的信号是完全由 $g(x)$ 在 $[x_1, x_2]$ 上的取值决定的, 故将这一区域称为 $g(x)$ 在区间 $[x_1, x_2]$ 上的决定区域.

前面介绍的是 $\mathbb{R}$ 上的波方程初值问题, 下面来展示 d'Alembert 公式的一个应用, 考虑半直线 $\mathbb{R}_+ = \{x > 0\}$ 上的初边值问题:

$$\begin{cases} u_{tt} - u_{xx} = 0, & \text{在 } \mathbb{R}_+ \times (0, \infty) \text{ 中} \\ u = g, u_t = h, & \text{在 } \mathbb{R}_+ \times \{t = 0\} \text{ 上} \\ u = 0, & \text{在 } \{x = 0\} \times (0, \infty) \text{ 上} \end{cases} \tag{7.161}$$

其中 g, h 给定, $g(0) = h(0) = 0$.

为了把初边值问题(7.161)转变成(7.155)的形式, 对 u, g, h 在 $\mathbb{R}$ 上作奇延拓得到:

$$\tilde{u}(x,t) := \begin{cases} u(x,t), & x \geq 0, t \geq 0 \\ -u(-x,t), & x \leq 0, t \geq 0 \end{cases} \quad \tilde{g}(x) := \begin{cases} g(x), & x \geq 0 \\ -g(-x), & x \leq 0 \end{cases} \quad \tilde{h}(x) := \begin{cases} h(x), & x \geq 0 \\ -h(-x), & x \leq 0 \end{cases}$$

进而(7.161)变为

$$\begin{cases} \tilde{u}_{tt} = \tilde{u}_{xx}, & \text{在 } \mathbb{R} \times (0, \infty) \text{ 内} \\ \tilde{u} = \tilde{g}, \tilde{u}_t = \tilde{h}, & \text{在 } \mathbb{R} \times \{t = 0\} \text{ 中} \end{cases}$$

应用 d'Alembert 公式(7.160)得

$$\tilde{u}(x,t) = \frac{1}{2}[\tilde{g}(x+t) + \tilde{g}(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} \tilde{h}(y)dy$$

根据前面 $\tilde{u}, \tilde{g}, \tilde{h}$ 的构造可得

$$u(x,t) = \begin{cases} \frac{1}{2}[g(x+t) + g(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy, & 0 \leq t \leq x \\ \frac{1}{2}[g(x+t) - g(t-x)] + \frac{1}{2}\int_{t-x}^{t+x} h(y)dy, & 0 \leq x \leq t \end{cases} \tag{7.162}$$

若 $h \equiv 0$, 则可以把(7.162)式理解为初始位移 g 分成了两部分, 其中一部分以速度 1 向右移动, 另一部分以速度 1 向左移动, 而向左移动的部分在 $x = 0$ 处反射, 这是因为 $x = 0$ 处振动弦是固定的 (速度为 0).

注意如果 $g''(0) \neq 0$, 则解(7.162)甚至都不会是 C^2 的. □

课堂笔记 (2023.11.6)

- (关于奇延拓的动机) 对于半直线 $\mathbb{R}^+$ 上的问题

$$\begin{cases} u_{tt} - u_{xx} = 0, & \text{在 } \mathbb{R}_+ \times (0, \infty) \text{ 中} \\ u = g, u_t = h, & \text{在 } \mathbb{R}_+ \times \{t = 0\} \text{ 上} \end{cases}$$

只要边值 $x = 0$ 处恒有 $u = 0$, 则在将 u 延拓到为了保证延拓后的函数足够的正则性, 一般都采用奇

延拓，即

$$\widetilde{u}(x,t)=\begin{cases}u(x,t), & x\ge 0,t\ge 0\\ -u(-x,t), & x\le 0,t\ge 0\end{cases}\quad \widetilde{g}(x)=\begin{cases}g(x), & x\ge 0\\ -g(-x), & x\le 0\end{cases}\quad \widetilde{h}(x)=\begin{cases}h(x), & x\ge 0\\ -h(-x), & x\le 0\end{cases}$$

这是因为此时对初值有：

$$g'_-(0)=\lim_{\delta\to 0^+}\frac{g(0)-g(-\delta)}{\delta}=\lim_{\delta\to 0^+}\frac{g(\delta)-g(0)}{\delta}=g'_+(0)$$

而对初速度有：

$$h'_-(0)=\lim_{\delta\to 0^+}\frac{h(0)-h(-\delta)}{\delta}=\lim_{\delta\to 0^+}\frac{h(\delta)-h(0)}{\delta}=h'_+(0)$$

其中注意 $g(0)=h(0)=0$. 这说明在 $x=0$ 处奇延拓可保证 g,h 的连续性与可微性，这是 g,h 在 0 附近不恒为 0 时，偶延拓所无法做到的.

注 这里单开一注说明反射法的几何想法. 同样设 $h\equiv 0$, 针对 d'Alembert 公式

$$\widetilde{u}(x,t)=\frac{1}{2}[\widetilde{g}(x+t)+\widetilde{g}(x-t)]$$

作影响区域与决定区域如图(7.20).

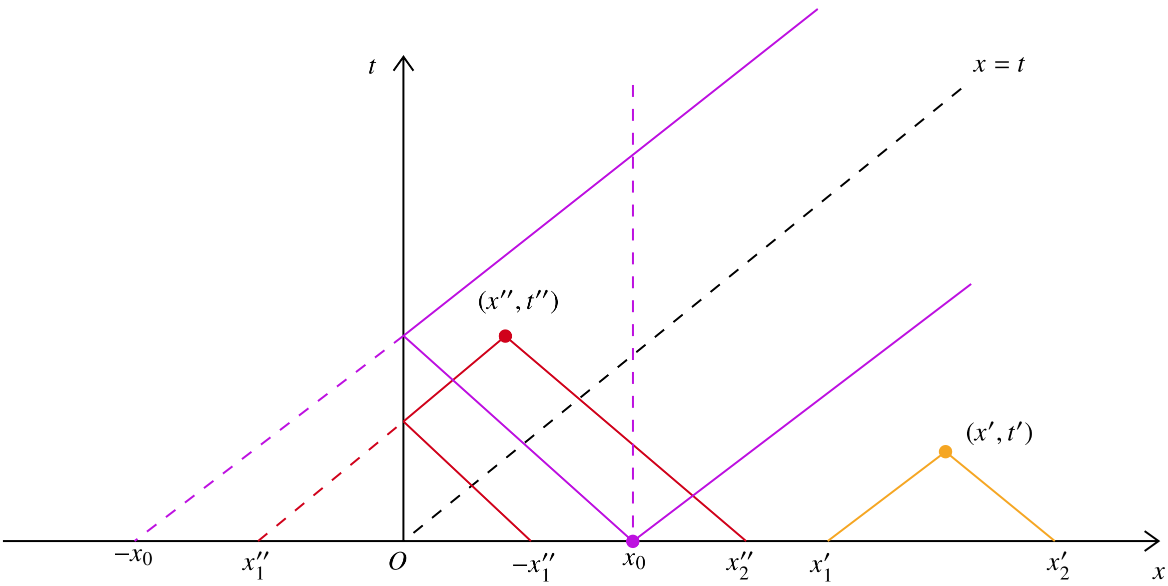


图 7.20: 反射法中的影响区域与决定区域示意图

针对 (x',t'), 因为它在直线 $x=t$ 下方，故它的依赖区域 $[x_1',x_2']$ 完全在 $\{(x,t):0\le t\le x\}$ 中，此时对应的公式恰为

$$u(x,t)=\frac{1}{2}[g(x+t)+g(x-t)]$$

这与前面讨论的 $\mathbb{R}$ 上的问题没有什么区别. 而针对 (x'',t''), 因为它在直线 $x=t$ 上方，故它接收到的信号中，向右传播的那一支的发出处 x_1'' 是小于 0 的，从而 $\widetilde{g}(x_1'')=-g(-x_1'')$, 这说明这一信号的值为 $-g(-x_1'')$, 而向左传输的信号的发出处 x_2'' 大于 0, 所以其信号值直接是 $g(x_2'')$. 综合来看，(x'',t'') 处的信号就应该是：

$$u(x'',t'')=\frac{1}{2}[\widetilde{g}(x_1'')+\widetilde{g}(x_2'')]=\frac{1}{2}[-g(-x_1'')+g(x_2'')]$$

而对于 x 轴上固定的点 x_0, 研究它发出的信号的传播时可以采取与上面类似的思路. 只要这个信号还在 $x>0$ 的区域，它就和全空间上波方程初值的信号传输没有什么区别. 而对于 xOt 系中还可能接收到来自 x_0 信号的点，它们正如前面讨论 (x'',t'') 的情况一样，接受到的信号是经过" 反射" 的，据此即可做出 x_0 点的影响区域 (紫色线条).

7.4.1.2 球均值

下面设 $n \geq 2, m \geq 2, u \in C^m(\mathbb{R}^n \times [0,\infty))$ 是初值问题

$$\begin{cases} u_{tt} - \Delta u = 0, & 在\ \mathbb{R}^n \times (0,\infty)\ 内 \\ u = g, u_t = h, & 在\ \mathbb{R}^n \times \{t = 0\}\ 上 \end{cases} \tag{7.163}$$

目的是导出 u 关于 g, h 的确切表达式. 首先引入球均值的概念, 这些均值作为时间 t 与半径 r 的函数是 Euler-Poisson-Darboux 方程的解, 而这个 PDE 在 n 为奇数的情况下就转换成了一维波方程. 应用 d'Alembert 公式或更精确的版本(7.162)就可以导出解的表达式.

定义 7.4.1 (球平均)

1. 若 $x \in \mathbb{R}^n, t > 0, r > 0$, 则定义

$$U(x, r; t) := \fint_{\partial B(x,r)} u(y,t) dS(y) \tag{7.164}$$

是 $u(\cdot, t)$ 在球面 $\partial B(x,r)$ 上的平均.

2. 类似记

$$G(x; r) := \fint_{\partial B(x,r)} g(y) dS(y), \quad H(x; r) := \fint_{\partial B(x,r)} h(y) dS(y) \tag{7.165}$$

♣

对固定的 x, 可以把 U 看作关于 r, t 的方程, 可以发现 U 是一个 PDE 的解:

引理 7.4.1 (Euler-Poisson-Darboux 方程)

固定 $x \in \mathbb{R}^n$, 设 u 是初值问题(7.163)的解, 则 $U \in C^m(\overline{\mathbb{R}_+} \times [0,\infty))$, 且

$$\begin{cases} U_{tt} - U_{rr} - \frac{n-1}{r} U_r = 0, & 在\ \mathbb{R}_+ \times (0,\infty)\ 内 \\ U = G, U_t = H, & 在\ \mathbb{R}_+ \times \{t = 0\}\ 上 \end{cases} \tag{7.166}$$

♡

(7.166)中所展示的 PDE 就是 Euler-Poisson-Darboux 方程. 注意 $U_{rr} + \frac{n-1}{r} U_r$ 是 Laplace 算子 Δ 的极坐标形式中关于半径的部分.

证明

回忆 Laplace 方程的均值定理的证明, 当 $r > 0$ 时, 知

$$U(x, r; t) = \frac{1}{n\alpha_n r^{n-1}} \int_{\partial B(x,r)} u(y,t) dS(y) = \frac{1}{n\alpha_n} \int_{\partial B(0,1)} u(x + rz, t) dS(z)$$

其中 α_n 是 n 维单位球体积. 故当 $r > 0$ 时有:

$$\begin{aligned} U_r(x, r; t) &= \frac{1}{n\alpha_n} \int_{\partial B(0,1)} z \cdot Du(x + rz, t) dS(z) \\ &= \frac{1}{n\alpha_n r^{n-1}} \int_{\partial B(x,r)} \frac{y - x}{r} \cdot Du(y, r) dS(y) \\ &= \frac{1}{n\alpha_n r^{n-1}} \int_{B(x,r)} \Delta u(y, r) dy = \frac{r}{n} \fint_{B(x,r)} \Delta u(y,t) dy \end{aligned}$$

综合得到

$$U_r(x; r, t) = \frac{r}{n} \fint_{B(x,r)} \Delta u(y,t) dy, \quad r > 0 \tag{7.167}$$

同时 $U_r(x; r, t) \to 0 (r \to 0^+)$, 这说明 U 在 $r = 0$ 处存在右导数. 下面对(7.167)进一步对 r 微分, 应用极坐标换

元(6.2.4)(ii) 有:

$$\begin{aligned} U_{rr}(x;r,t) &= \frac{1}{n}\fint_{B(x,r)} \Delta u(y,t)dy + \frac{r}{n}\cdot\frac{d}{dr}\Big(\frac{1}{\alpha(n)r^n}\int_{B(x,r)} \Delta u(y,t)dy\Big) \\ &= \frac{1}{n}\fint_{B(x,r)} \Delta u(y,t)dy - \frac{r}{n}\cdot\frac{n}{\alpha(n)r^{n+1}}\int_{B(x,r)} \Delta u(y,t)dy + \frac{r}{n}\cdot\frac{1}{\alpha(n)r^n}\cdot\int_{\partial B(x,r)} \Delta u(y,t)dS \\ &= \fint_{\partial B(x,r)} \Delta u(y,t)dS + \Big(\frac{1}{n}-1\Big)\fint_{B(x,r)} \Delta u dy, \quad r>0 \end{aligned} \tag{7.168}$$

同时 $U_{rr}(x;r,t)\to\frac{1}{n}\Delta u(x,t)(r\to 0^+)$, 这说明 U_r 在 $r=0$ 处存在右导数. 根据(7.168)式可以进一步计算 U_{rrr} 等等, 进而可以验证 $U\in C^m(\overline{\mathbb{R}_+}\times[0,\infty))$.

下面验证 U 是 Euler-Poisson-Darboux 方程的解. 因为 u 是初值问题(7.163)的解, 故由(7.167)式知

$$U_r = \frac{r}{n}\fint_{B(x,r)} u_{tt}(y,t)dy = \frac{1}{n\alpha(n)r^{n-1}}\int_{B(x,r)} u_{tt}(y,t)dy$$

因而

$$r^{n-1}U_r = \frac{1}{n\alpha(n)}\int_{B(x,r)} u_{tt}(y,t)dy$$

进而

$$(r^{n-1}U_r)_r = \frac{1}{n\alpha(n)}\int_{\partial B(x,r)} u_{tt}(y,t)dS = r^{n-1}\fint_{\partial B(x,r)} u_{tt}(y,t)dS = r^{n-1}U_{tt}$$

展开即得

$$U_{tt} - U_{rr} - \frac{n-1}{r}U_r = 0$$

□

7.4.1.3 $n=3,2$ 时的解, Kirchhoff 公式与 Poisson 公式

本节的目的是把 Euler-Poisson-Darboux 方程(7.166)转化成通常的一维波方程. 因为整个过程实在是太复杂了, 这里先讨论阶数 $n=3,2$ 这两个简单点的情况.

$n=3$ 时的解

设 $n=3$, 并设 $u\in C^2(\mathbb{R}^3\times[0,\infty))$ 是初值问题(7.163)的解. 回忆 U,G,H 的定义(7.164),(7.165), 设

$$\tilde{U} := rU \tag{7.169}$$

$$\tilde{G} := rG, \quad \tilde{H} := rH \tag{7.170}$$

下面说明 $\tilde{U}$ 满足:

$$\begin{cases} \tilde{U}_{tt} - \tilde{U}_{rr} = 0, & \text{在 } \mathbb{R}_+\times(0,\infty) \text{ 中} \\ \tilde{U} = \tilde{G}, \tilde{U}_t = \tilde{H}, & \text{在 } \mathbb{R}_+\times\{t=0\} \text{ 上} \\ \tilde{U} = 0, & \text{在 } \{r=0\}\times(0,\infty) \text{ 上} \end{cases} \tag{7.171}$$

这是因为将 $n=3$ 代入(7.166)式有

$$\tilde{U}_{tt} = rU_{tt} = r\Big(U_{rr}+\frac{2}{r}U_r\Big) = rU_{rr}+2U_r = (U+rU_r)r = \tilde{U}_{rr}$$

此即证毕. 这说明 $\tilde{U}$ 是满足一维波方程的, 同时注意 $\tilde{G}_{rr}(0)=0$, 现在将 d'Alembert 公式(7.160)代入, 知当 $0\le r\le t$ 时:

$$\tilde{U}(x;r,t) = \frac{1}{2}[\tilde{G}(r+t)-\tilde{G}(t-r)] + \frac{1}{2}\int_{-r+t}^{r+t}\tilde{H}(y)dy \tag{7.172}$$

从球平均的定义(7.164)可以看出 $u(x,t)=\lim_{r\to 0^+}U(x;r,t)$, 进而由(7.169),(7.170),(7.172)式可得

$$u(x,t) = \lim_{r\to 0^+}\frac{\tilde{U}(x;r,t)}{r} = \lim_{r\to 0^+}\Big[\frac{\tilde{G}(t+r)-\tilde{G}(t-r)}{2r} + \frac{1}{2r}\int_{t-r}^{t+r}\tilde{H}(y)dy\Big] = \tilde{G}'(t)+\tilde{H}(t)$$

再把 $\tilde{G},\tilde{H}$ 的定义(7.165)代入得

$$u(x,t)=\frac{\partial}{\partial t}\Big(t⨍_{\partial B(x,t)} g\,dS\Big)+t⨍_{\partial B(x,t)} h\,dS \tag{7.173}$$

但

$$⨍_{\partial B(x,t)} g(y)dS(y)=⨍_{\partial B(0,1)} g(x+tz)dS(z)$$

这便得到

$$\frac{\partial}{\partial t}\Big(⨍_{\partial B(x,t)} g\,dS\Big)=⨍_{\partial B(0,1)} Dg(x+tz)\cdot z\,dS(z)=⨍_{\partial B(x,t)} Dg(y)\cdot\Big(\frac{y-x}{t}\Big)dS(y)$$

回到(7.173)式, 现在便有

$$u(x,t)=⨍_{\partial B(x,t)}\big(th(y)+g(y)+Dg(y)\cdot(y-x)\big)dS(y),\quad x\in\mathbb{R}^3,t>0 \tag{7.174}$$

这便是三维情况下初值问题(7.163)的 Kirchhoff 公式.

$n=2$ 时的解

当 $n=2$ 时, 用类似于(7.169)式的变换把 Euler-Poisson-Darboux 方程转化成一维波方程的方法不再生效了. 转而代之, 我们就把 $n=2$ 时的初值问题(7.163)当成 $n=3$ 时的问题, 只是这个时候第三个空间变量 x_3 不出现.

设 $u\in C^2(\mathbb{R}^2\times[0,\infty))$ 是(7.163)在 $n=2$ 时的解, 记

$$\overline{u}(x_1,x_2,x_3,t):=u(x_1,x_2,t) \tag{7.175}$$

则(7.163)说明

$$\begin{cases}\overline{u}_{tt}-\Delta\overline{u}=0, & \text{在 } \mathbb{R}^3\times(0,\infty) \text{ 中}\\ \overline{u}=\overline{g},\overline{u}_t=\overline{h}, & \text{在 } \mathbb{R}^3\times\{t=0\} \text{ 上}\end{cases} \tag{7.176}$$

其中

$$\overline{g}(x_1,x_2,x_3):=g(x_1,x_2),\quad \overline{h}(x_1,x_2,x_3):=h(x_1,x_2)$$

如果记 $x=(x_1,x_2)\in\mathbb{R}^2,\overline{x}=(x_1,x_2,0)\in\mathbb{R}^3$, 则条件(7.176)与 Kirchhoff 公式(7.173)表明:

$$u(x,t)=\overline{u}(\overline{x},t)=\frac{\partial}{\partial t}\Big(t⨍_{\partial\overline{B}(\overline{x},t)}\overline{g}\,d\overline{S}\Big)+t⨍_{\partial\overline{B}(\overline{x},t)}\overline{h}\,d\overline{S} \tag{7.177}$$

其中 $\overline{B}(\overline{x},t)$ 表示 $\mathbb{R}^3$ 中以 $\overline{x}$ 为圆心, $t>0$ 为半径的球, 而 $d\overline{S}$ 表示 $\partial\overline{B}(\overline{x},t)$ 上的二维面测度.

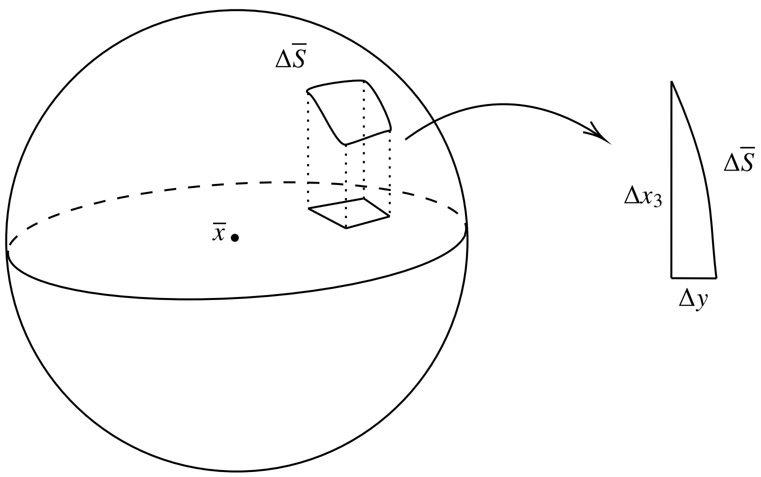


图 7.21

现在把讨论拉回到 $\mathbb{R}^2$ 上, 根据图(7.21)可知近似有

$$\Delta x_3^2+\Delta y^2=\Delta\overline{S}^2 \tag{7.178}$$

其中 $y=(x_1,x_2)$. 根据 $\overline{B}(\overline{x},t)$ 的方程可知

$$x_3=\pm(t^2-|y-x|^2)^{\frac{1}{2}} \tag{7.179}$$

记 $\gamma(y)=(t^2-|y-x|^2)^{\frac{1}{2}}$, 将(7.179)式代入(7.178)式有

$$\Delta\overline{S}^2=\Delta y^2+|\Delta\gamma(y)|^2\Rightarrow\Delta\overline{S}=(1+|\frac{\Delta\gamma(y)}{\Delta y}|^2)^{\frac{1}{2}}\cdot\Delta y$$

进而可得

$$d\overline{S}=(1+|D\gamma(y)|^2)^{\frac{1}{2}}dy$$

有

$$⨍_{\partial\overline{B}(\overline{x},t)}\overline{g}d\overline{S}=\frac{1}{4\pi t^2}\int_{\partial\overline{B}(\overline{x},t)}\overline{g}d\overline{S}=\frac{2}{4\pi t^2}\int_{B(x,t)}g(y)(1+|D\gamma(y)|^2)^{\frac{1}{2}}dy$$

其中 $\frac{1}{4\pi t^2}$ 变成 $\frac{2}{4\pi t^2}$ 是因为 $\partial\overline{B}(\overline{x},t)$ 有上下两个半球面. 计算知

$$(1+|D\gamma|^2)^{\frac{1}{2}}=t(t^2-|y-x|^2)^{-\frac{1}{2}}$$

故

$$⨍_{\partial\overline{B}(\overline{x},t)}\overline{g}d\overline{S}=\frac{1}{2\pi t}\int_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy=\frac{t}{2}⨍_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy \tag{7.180}$$

同理

$$⨍_{\partial\overline{B}(\overline{x},t)}\overline{h}d\overline{S}=\frac{t}{2}⨍_{B(x,t)}\frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy \tag{7.181}$$

将(7.180),(7.181)式代入(7.177)式得

$$u(x,t)=\frac{1}{2}\frac{\partial}{\partial t}(t^2⨍_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy)+\frac{t^2}{2}⨍_{B(x,t)}\frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy \tag{7.182}$$

而

$$t^2⨍_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy=t⨍_{B(0,1)}\frac{g(x+tz)}{(1-|z|^2)^{\frac{1}{2}}}dz$$

故

$$\begin{aligned}\frac{\partial}{\partial t}(t^2⨍_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy)&=⨍_{B(0,1)}\frac{g(x+tz)}{(1-|z|^2)^{\frac{1}{2}}}dz+t⨍_{B(0,1)}\frac{Dg(x+tz)\cdot z}{(1-|z|^2)^{\frac{1}{2}}}dz\\&=t⨍_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy+t⨍_{B(x,t)}\frac{Dg(y)\cdot(y-x)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy\end{aligned}$$

将结果代入(7.182)式得

$$u(x,t)=\frac{1}{2}⨍_{B(x,t)}\frac{tg(y)+t^2h(y)+tDg(y)\cdot(y-x)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy,\quad x\in\mathbb{R}^2,t>0 \tag{7.183}$$

这便是初值问题(7.163)在二维情况下解的 Poisson 公式.

这种先解决 $n=3$ 时的问题, 再应用到 $n=2$ 的情景的方法称为降维法.

课堂笔记 (2023.11.6)

- (关于三维 Kirchhoff 公式与降维法的动机) 要求解波方程初值问题

$$\begin{cases}u_{tt}-\Delta u=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 内}\\ u=g,u_t=h, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

根据引理(7.4.1), 就是要求解球平均的 Euler-Poisson-Darboux 方程

$$\begin{cases}U_{tt}-U_{rr}-\frac{n-1}{r}U_r=0, & \text{在 } \mathbb{R}_+\times(0,\infty) \text{ 内}\\ U=G,U_t=H, & \text{在 } \mathbb{R}_+\times\{t=0\} \text{ 上}\end{cases}$$

在求解上述方程得到 $U(x;r,t)$ 后, 令 $r \to 0^+$ 即得 u 满足的表达式. 但 Euler-Poisson-Darboux 方程中 $-\frac{n-1}{r}U_r$ 一项难以消去, 事实上在 n 为偶数时该项无法消除, 进而 Euler-Poisson-Darboux 方程难以求解. 然而 n 为奇数时可以消去 $-\frac{n-1}{r}U_r$ 这一项, 特别当 $n=3$ 时对应即取 $\widetilde{U}=rU$, 这便将 Euler-Poisson-Darboux 方程转化回了可应用反射法的半直线一维波方程:

$$\begin{cases} \widetilde{U}_{tt}-\widetilde{U}_{rr}=0, & \text{在 } \mathbb{R}_+\times(0,\infty) \text{ 中} \\ \widetilde{U}=\widetilde{G}, \widetilde{U}_t=\widetilde{H}, & \text{在 } \mathbb{R}_+\times\{t=0\} \text{ 上} \\ \widetilde{U}=0, & \text{在 } \{r=0\}\times(0,\infty) \text{ 上} \end{cases}$$

现在可应用反射法的半直线一维波方程有现成的 d'Alembert 公式, 于是 $\widetilde{U}$ 可求解, 因而 U 可求解, 再令 $r\to 0^+$ 便得到 u 的表达式, 这便是三维 Kirchhoff 公式 (与一般奇数维求解公式) 的推导动机.

因为前面提到 n 为偶数时 Euler-Poisson-Darboux 方程不好直接求解, 但 n 为奇数时可以直接求解, 故考虑在 n 为偶数时另增一维, 转化成奇数维问题后求解.

- (关于波方程对正则性的损失) 观察三维波方程的 Kirchhoff 公式:

$$u(x,t)=⨍_{\partial B(x,t)}(th(y)+g(y)+Dg(y)\cdot(y-x))dS(y)$$

可以知道要想 $u\in C^2(\mathbb{R}^3\times[0,+\infty))$, 就需要 $h\in C^2, g\in C^2, Dg\in C^2$, 进一步整理即得 $h\in C^2, g\in C^3$. 注意对初值正则性的要求是比对解正则性的要求高的, 故称波方程会损失正则性. 一般来说, 对于 n 维波方程, 要想保证 $u\in C^2$, 需要 $g\in C^{[\frac{n}{2}]+3}, h\in C^{[\frac{n}{2}]+2}$, 进而正则性会损失 $[\frac{n}{2}]$ 阶.

7.4.1.4 n 为奇数时的解

本节我们对奇数 $n\geq 3$ 求解 Euler-Poisson-Darboux 方程. 首先声明一些事实.

引理 7.4.2 (一些有用的性质)

若 $\phi:\mathbb{R}\to\mathbb{R}$ 是 C^{k+1} 的, 则对 $k=1,2,\cdots$ 有:

(i) $(\frac{d^2}{dr^2})(\frac{1}{r}\frac{d}{dr})^{k-1}(r^{2k-1}\phi(r))=(\frac{1}{r}\frac{d}{dr})^k(r^{2k}\frac{d\phi}{dr}(r))$,

(ii) $(\frac{1}{r}\frac{d}{dr})^{k-1}(r^{2k-1}\phi(r))=\sum_{j=0}^{k-1}\beta_j^k r^{j+1}\frac{d^j\phi}{dr^j}(r)$, 其中常数 $\beta_j^k(j=0,1,\cdots,k-1)$ 与 ϕ 独立.

(iii) $\beta_0^k=1\cdot 3\cdot 5\cdots(2k-1)$.

用归纳法即证命题. □

下面设 $n\geq 3$ 是奇整数, 记 $n=2k+1(k\geq 1)$. 现设 $u\in C^{k+1}(\mathbb{R}^n\times[0,\infty))$ 是初值问题(7.163)的解, 知由(7.164)式定义的 U 是 C^{k+1} 的.

记

$$\begin{cases} \tilde{U}(r,t):=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}U(x;r,t)) \\ \tilde{G}(r):=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}G(x;r)) \\ \tilde{H}(r):=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}H(x;r)) \end{cases} \quad (r>0,t\geq 0) \tag{7.184}$$

进而

$$\tilde{U}(r,0)=\tilde{G}(r),\quad \tilde{U}_t(r,0)=\tilde{H}(r) \tag{7.185}$$

下面结合引理(7.166)与引理(7.4.2)给出的性质来说明 U 到 $\tilde{U}$ 的变换(7.184)把 Euler-Poisson-Darboux 方程转化成了波方程.

引理 7.4.3 ($\tilde{U}$ 是一维波方程的解)

对 $\tilde{U}$ 有

$$\begin{cases}\tilde{U}_{tt}-\tilde{U}_{rr}=0, & 在\ \mathbb{R}_+\times(0,\infty)\ 中\\ \tilde{U}=\tilde{G},\tilde{U}=\tilde{H}, & 在\ \mathbb{R}_+\times\{t=0\}\ 上\\ \tilde{U}=0, & 在\ \{r=0\}\times(0,\infty)\ 上\end{cases}$$

证明

若 $r>0$, 则

$$\begin{aligned}\tilde{U}_{rr}&=(\frac{\partial^2}{\partial r^2})(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}U)\overset{(7.4.2)}{=}(\frac{1}{r}\frac{\partial}{\partial r})^{k}(r^{2k}U_r)\\&=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}U_{rr}+2kr^{2k-2}U_r)=(\frac{1}{r}\frac{\partial}{\partial r})[r^{2k-1}(U_{rr}+\frac{n-1}{r}U_r)]\\&=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}U_{tt})=\tilde{U}_{tt}\end{aligned}$$

其中倒数第二个等号是因为 U 满足 Euler-Poisson-Darboux 方程. 由引理(7.4.2)(ii) 知当 $r=0$ 时 $\tilde{U}=0$. □

现在从引理(7.4.3), (7.185)式与 d'Alembert 公式(7.162)知

$$\tilde{U}(r,t)=\frac{1}{2}[\tilde{G}(r+t)-\tilde{G}(t-r)]+\frac{1}{2}\int_{t-r}^{t+r}\tilde{H}(y)dy \tag{7.186}$$

根据 U 的定义知 $u(x,t)=\lim_{r\to 0}U(x;r,t)$, 进一步引理(7.4.2)(ii) 说明

$$\tilde{U}(r,t)=(\frac{1}{r}\frac{\partial}{\partial r})^{k-1}(r^{2k-1}U(x;r,t))=\sum_{j=0}^{k-1}\beta_j^k r^{j+1}\frac{\partial^j}{\partial r^j}U(x;r,t)$$

故

$$\lim_{r\to 0}\frac{\tilde{U}(r,t)}{\beta_0^k r}=\lim_{r\to 0}U(x;r,t)=u(x,t)$$

从而(7.185)式表明

$$u(x,t)=\frac{1}{\beta_0^k}\lim_{r\to 0}[\frac{\tilde{G}(t+r)-\tilde{G}(t-r)}{2r}+\frac{1}{2r}\int_{t-r}^{t+r}\tilde{H}(y)dy]=\frac{1}{\beta_0^k}[\tilde{G}'(t)+\tilde{H}(t)]$$

最后, 因为 $n=2k+1$, 考虑 $\tilde{G},\tilde{H}$ 的定义(7.184)与引理(7.4.2)(iii) 知对 $x\in\mathbb{R}^n,t>0$ 有

$$u(x,t)=\frac{1}{\gamma_n}[(\frac{\partial}{\partial t})(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(t^{n-2}\fint_{\partial B(x,t)}gdS)+(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(t^{n-2}\fint_{\partial B(x,t)}hdS)],\quad n\ 为奇数, \gamma_n=1\cdot 3\cdot 5\cdots(n-2) \tag{7.187}$$

可以计算 $\gamma_3=1$, 故(7.187)式与 $n=3$ 的情况下的 Kirchhoff 公式(7.174)是契合的.

还需要验证(7.187)式确实是初值问题(7.163)的解.

定理 7.4.2 (奇数维波方程的解)

若 n 是奇数, $n\geq 3$, $g\in C^{m+1}(\mathbb{R}^n)$, $h\in C^m(\mathbb{R}^n)$, $m=\frac{n+1}{2}$, u 由(7.187)式定义, 则:

1. $u\in C^2(\mathbb{R}^n\times[0,\infty))$,
2. 在 $\mathbb{R}^n\times(0,\infty)$ 中 $u_{tt}-\Delta u=0$,
3. 对 $x^0\in\mathbb{R}^n$ 有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\x\in\mathbb{R}^n,t>0}}u(x,t)=g(x^0)$, $\lim\limits_{\substack{(x,t)\to(x^0,0)\\x\in\mathbb{R}^n,t>0}}u_t(x,t)=h(x^0)$.

证明

先设 $g\equiv 0$, 则由(7.187)式知

$$u(x,t)=\frac{1}{\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(t^{n-2}H(x,t)) \tag{7.188}$$

由引理(7.4.2)(i) 知

$$u_{tt}=\frac{1}{\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-1}{2}}(t^{n-1}H_t) \tag{7.189}$$

知

$$\begin{aligned}H_t &= \frac{d}{dt}(\fint_{\partial B(x,t)} hdS) = \frac{d}{dt}\fint_{\partial B(0,1)} h(x+tz)dS(z) = \fint_{\partial B(0,1)} Dh(x+tz)\cdot zdS(z)\\ &= \fint_{\partial B(x,t)} Dh(y)\cdot \frac{y-x}{t}dS(y) = \frac{1}{n\alpha(n)t^{n-1}}\int_{B(x,t)}\Delta hdy = \frac{t}{n}\fint_{B(x,t)}\Delta hdy\end{aligned}$$

代入(7.189)式知

$$u_{tt} = \frac{1}{n\alpha(n)\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-1}{2}}(\int_{B(x,t)}\Delta hdy) = \frac{1}{n\alpha(n)\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(\frac{1}{t}\int_{\partial B(x,t)}\Delta hdS) \tag{7.190}$$

另一方面

$$\Delta H(x;t) = \Delta_x\fint_{\partial B(0,t)} h(x+y)dS(y) = \fint_{\partial B(0,t)}\Delta h(x+y)dS(y) = \fint_{\partial B(x,t)}\Delta hdS$$

这说明

$$\Delta u = \frac{1}{\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(t^{n-2}\Delta H(x;t)) = \frac{1}{\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(t^{n-2}\cdot\frac{1}{n\alpha(n)t^{n-1}}\int_{\partial B(x,t)}\Delta hdS) = \frac{1}{n\alpha(n)\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}}(\frac{1}{t}\int_{\partial B(x,t)}\Delta hdS) \tag{7.191}$$

从而由(7.190),(7.191)式即得在 $\mathbb{R}^n\times(0,\infty)$ 中有 $u_{tt}-\Delta u=0$.

$h\equiv 0$ 时结论是类似的, 随后由引理(7.4.2)(ii)-(iii) 可得一般的初值条件下命题也成立. □

注

1. 在计算 $u(x,t)$ 时, 只需要知道 g,h 与它们在球面 $\partial B(x,t)$ 上的导数这些信息, 而无需在整个球体 $B(x,t)$ 内讨论.
2. 把(7.187)式与 d'Alembert 公式(7.160)比较, 可以观察到 d'Alembert 公式并不包含 g 的导数. 这暗示当 $n>1$ 时, 波方程(7.163)的解在时间 $t>0$ 时不一定和初值 g 一样光滑: g 的非正则性可能会在某个 $t>0$ 时体现出来, 进而让 u 也变得不那么正则.
3. 和 $n=1$ 的情况一样, 可以发现此时初始扰动的传播速度是有限的.

7.4.1.5 n 为偶数时的解

现在设 n 是偶整数, 设 $u\in C^m$ 是(7.163)的解, $m=\frac{n+2}{2}$. 希望对 u 找到一个类似于(7.187)式的表达式, 方法其实跟 $n=2$ 时大同小异, 记

$$\bar{u}(x_1,\cdots,x_{n+1},t) := u(x_1,\cdots,x_n,t) \tag{7.192}$$

是 $\mathbb{R}^{n+1}\times(0,\infty)$ 上波方程的解, 初值条件为

$$\bar{u}=\bar{g},\bar{u}_t=\bar{h},\quad 在\ \mathbb{R}^{n+1}\times\{t=0\}$$

其中

$$\begin{cases}\bar{g}(x_1,\cdots,x_{n+1}) := g(x_1,\cdots,x_n)\\ \bar{h}(x_1,\cdots,x_{n+1}) := h(x_1,\cdots,x_n)\end{cases} \tag{7.193}$$

因为 $n+1$ 是奇数, 故可由(7.187)式给出 $\bar{u}$ 关于 $\bar{g},\bar{h}$ 的表达式, 进一步通过(7.192),(7.193)式可得 u 关于 g,h 的表达式. 这同样是所谓降维法.

下面详细介绍这个过程, 固定 $x\in\mathbb{R}^n, t>0$, 记 $\bar{x}=(x_1,\cdots,x_n,0)\in\mathbb{R}^{n+1}$, 则(7.187)式表明

$$u(x,t) = \frac{1}{\gamma_{n+1}}[\frac{\partial}{\partial t}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^{n-1}\fint_{\partial\bar{B}(\bar{x},t)\bar{g}d\bar{S}}) + (\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^{n-1}\fint_{\partial\bar{B}(\bar{x},t)}\bar{h}d\bar{S})] \tag{7.194}$$

其中 $\bar{B}(\bar{x},t)$ 表示 $\mathbb{R}^{n+1}$ 中以 $\bar{x}$ 为球心, 以 t 为半径的球, $d\bar{S}$ 表示 $\partial\bar{B}(\bar{x},t)$ 上的 n 维面积测度. 根据积分平均的定义知

$$\fint_{\partial\bar{B}(\bar{x},t)}\bar{g}d\bar{S} = \frac{1}{(n+1)\alpha(n+1)t^n}\int_{\partial B(\bar{x},t)}\bar{g}d\bar{S} \tag{7.195}$$

注意 $\partial\bar{B}(x,t)\cap\{y_{n+1}\geq 0\}$ 是函数 $\gamma(y):=(t^2-|y-x|^2)^{\frac{1}{2}}$ 当 $y\in B(x,t)\subset\mathbb{R}^n$ 时的图像. 同理 $\partial\bar{B}(x,t)\cap\{y_{n+1}\leq 0\}$

是 $-\gamma$ 的图像. 类似于图(7.21), 可推知

$$\fint_{\partial\overline{B}(\overline{x},t)} \overline{g}\, d\overline{S} = \frac{2}{(n+1)\alpha(n+1)t^n}\int_{B(x,t)} g(y)(1+|D\gamma(y)|^2)^{\frac{1}{2}}dy \tag{7.196}$$

其中分子 2 是因为 $\partial\overline{B}(\overline{x},t)$ 有上下两个半球面. 注意 $(1+|D\gamma(y)|^2)^{\frac{1}{2}} = t(t^2-|y-x|^2)^{-\frac{1}{2}}$, 代入(7.196)式有

$$\fint_{\partial\overline{B}(\overline{x},t)} \overline{g}\, d\overline{S} = \frac{2}{(n+1)\alpha(n+1)t^{n-1}}\int_{B(x,t)} \frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy = \frac{2t\alpha(n)}{(n+1)\alpha(n+1)}\fint_{B(x,t)} \frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy \tag{7.197}$$

对 h 同理可得

$$\fint_{\partial\overline{B}(\overline{x},t)} \overline{h}\, d\overline{S} = \frac{2}{(n+1)\alpha(n+1)t^{n-1}}\int_{B(x,t)} \frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy = \frac{2t\alpha(n)}{(n+1)\alpha(n+1)}\fint_{B(x,t)} \frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy \tag{7.198}$$

将(7.197),(7.198)式代入(7.194)式得

$$u(x,t) = \frac{1}{\gamma_{n+1}}\frac{2\alpha(n)}{(n+1)\alpha(n+1)}[\frac{\partial}{\partial t}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^n\fint_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy) + (\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^n\fint_{B(x,t)}\frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy)]$$

已知 $\gamma_{n+1} = 1\cdot 3\cdot 5\cdots(n-1)$, $\alpha(n) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n+2}{2})}$, 可得 $\gamma_n = 2\cdot 4\cdots(n-2)\cdot n$, 故当 n 为偶数时, 对 $x\in\mathbb{R}^n, t>0$ 有

$$u(x,t) = \frac{1}{\gamma_n}[\frac{\partial}{\partial t}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^n\fint_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy) + (\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-2}{2}}(t^n\fint_{B(x,t)}\frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy)] \tag{7.199}$$

可以计算 $\gamma_2 = 2$, 故 $n=2$ 时(7.199)式与 Poisson 公式(7.183)是契合的.

定理 7.4.3 (偶数维波方程的解)

若 n 是偶数, $n\geq 2$, $g\in C^{m+1}(\mathbb{R}^n)$, $h\in C^m(\mathbb{R}^n)$, $m=\frac{n+2}{2}$, u 由(7.199)式定义, 则:

1. $u\in C^2(\mathbb{R}^n\times[0,\infty))$,
2. 在 $\mathbb{R}^n\times(0,\infty)$ 中有 $u_{tt}-\Delta u = 0$,
3. 对 $x^0\in\mathbb{R}^n$ 有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n, t>0}} u(x,t) = g(x^0)$, $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n, t>0}} u_t(x,t) = h(x^0)$.

♡

证明同定理(7.4.2)类似, 但需要特别注意的是从(7.199)式中可知, 此时想要计算 u 需要 g,h 在 $B(x,t)$ 内的信息, 而不再需要 $\partial B(x,t)$ 上的信息. 这与奇数情况是截然相反的.

比较(7.187)式与(7.199)式, 观察知如果 n 是奇数且 $n\geq 3$, 则在给定点 $x\in\mathbb{R}^n$ 后, g,h 的数据仅在圆锥 $C=\{(y,t): t>0, |x-y|<t\}$ 的边界 $\{(y,t): t>0, |x-y|=t\}$ 处对解 u 造成影响. 相反, 如果 n 是偶数, g,h 的数据就会在圆锥 C 之内对 u 造成影响. 也就是说, 在 x 处的" 扰动" 在奇数维时是沿着一个尖波传播的, 而在偶数维时就算这个尖波的前缘已经过去了, 它还是会产生影响. 这就是所谓的 Huygens 原理.

注 [**Ev**] 上对 Huygens 原理的介绍有些笼统, 下面引入 [**CZC**] 的介绍.

首先研究三维波的传播, 回顾 Kirchhoff 公式(7.173)(其中积分平均已经展开了):

$$u(x,t) = \frac{1}{4\pi t}\int_{\partial B(x,t)} h dS + \frac{\partial}{\partial t}(\frac{1}{4\pi t}\int_{\partial B(x,t)} g dS)$$

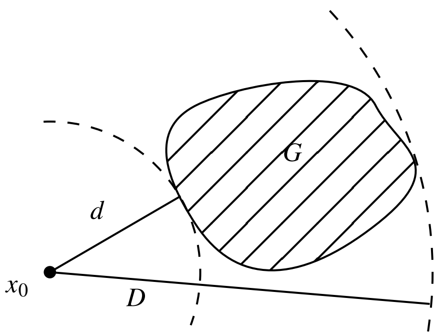


图 7.22

如图(7.22), 设 $t=0$ 时有界区域 $G\subset\mathbb{R}^3$ 内发生了扰动, 也即此时在 G 内 $g,h\neq 0$, 而在 G 外 $g=h=0$. 现在研究 G 外的一点 x_0 的振动情况. 取 $d=\rho(x_0,G)>0, D=\sup\limits_{x\in G}|x_0-x|$, 当 $0<t\leq d$ 时, 知 $\partial B(x,t)\cap G=\varnothing$, 故此时 $u(x_0,t)=0$, 也即 x_0 还保持静止状态. 当 $d<t<D$ 时, $\partial B(x_0,t)\cap G\neq\varnothing$, 一般来说此时 $u(x_0,t)\neq 0$, 也即 x_0 受到来自 G 的扰动而开始振动. 当 $t\geq D$, $\partial B(x_0,t)\cap G=\varnothing$, 此时 $u(x_0,t)=0$, 这说明来自 G 的扰动已经传过了 x_0, x_0 恢复静止.

再看 $t=t_0$ 时扰动的全局性态 $u(x,t_0)$. 此时受到扰动的区域是 (如图(7.23))

$$\Omega=\{x:\partial(x,t_0)\cap G\neq\varnothing\}=\bigcup_{x\in G}\partial B(x,t_0)$$

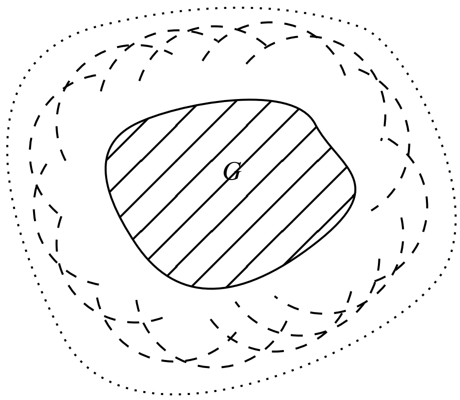


图 7.23

若记 d^* 为 G 的直径, 则当 $t>d^*$ 时, 空间中会出现明显的外包络面 (如图(7.23)所示) 与内包络面, 它们分别称为波的前阵面和后阵面. 如果 G 是球体, 则波的前阵面和后阵面都是球面, 这在一定意义上解释了为什么一般把三维波动方程的解称为球面波. 而这种" 初始扰动对空间中的每一点只影响有限的时间, 且波的传播有清晰的前阵面和后阵面" 的现象, 就称为 Huygens 原理.

再来看二维波的传播, 回顾 Poisson 公式(7.183)(其中积分平均已经展开):

$$u(x,t)=\frac{1}{2\pi}\frac{\partial}{\partial t}\int_{B(x,t)}\frac{g(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy+\frac{1}{2\pi}\int_{B(x,t)}\frac{h(y)}{(t^2-|y-x|^2)^{\frac{1}{2}}}dy$$

同样设 $t=0$ 时有界区域 $G\subset\mathbb{R}^2$ 内发生了扰动, 也即此时在 G 内 $g,h\neq 0$, 而在 G 外 $g=h=0$. 对平面上固定的一点 x_0, 当 $0<t<d=\rho(x_0,G)$ 时, $B(x_0,t)\cap G=\varnothing$, 故此时 $u(x_0,t)=0$, 但若 $t\geq d$, 知往后便一直有 $B(x_0,t)\cap G\neq\varnothing$, 从而一般都有 $u(x_0,t)\neq 0$. 这说明一旦 x_0 被扰动, 就会一直运动下去. 但又同时注意到, 若 $t\to\infty$, 可得 $u\to 0$, 这是因为 G 有界, 故当 t 足够大时积分是只在 G 内进行的. 这说明尽管 x_0 被扰动了, 但随着时间的推移它的振动还是会越来越微弱. 总的来说, 二维波相较于三维波, 保留了清晰的前阵面, 但没有后阵面了, 这种现象叫做波的弥散.

最后, 回顾奇数维波方程的解公式(7.187)与偶数维波方程的解公式(7.199), 不难借由上面的思想发现: 奇数维时 Huygens 原理都成立, 而偶数维时都会发生波的弥散.

7.4.1.6 补充: 强 (弱)Huygens 原理的解释

徐桂香老师关于 Huygens 原理的解释或许更为直观, 此处单开一节特别解释课堂笔记.

为了同时解释三维波与二维波, 将 $x=(x_1,x_2)$ 与 x_3 分别作两条轴. 现在当初值表现为三维, 特别当初值仅在点 (x_1',x_2',x_3') 处有值 (即单点信号) 时 (如图(7.24)), 若波速为 1, 则在时间为 t_1 时, 信号应传输到以 (x_1',x_2',x_3') 为圆心, 以 t_1 为半径的圆圈上 (亦即图(7.24)两幅图中的黄色圆圈). 同理, 在 t_2 时信号应传输到图(7.24)两幅图中的紫色圆圈处. 特别注意在特定时刻对应圆圈外的点是接收不到信号的, 譬如 t_2 时刻另有点 (x_1'',x_2'',x_3''), 它在空间中

位于信号圆圈内 (见图(7.24)右幅, 蓝色点在紫色圆圈围成的区域内), 但信号此时只存在于紫色圆圈上, 故蓝色点处在此时并无信号. 如此一来, 单点初值的影响区域在时空图中表现为一个延伸到无穷远处的圆锥面. 这种信号只存在于信号圆圈上, 源自某点 A 处初值的信号圆圈在经过空间中一点 B 后, B 点不再接收来自 A 的信号的性质, 就称为三维波的强 Huygens 原理.

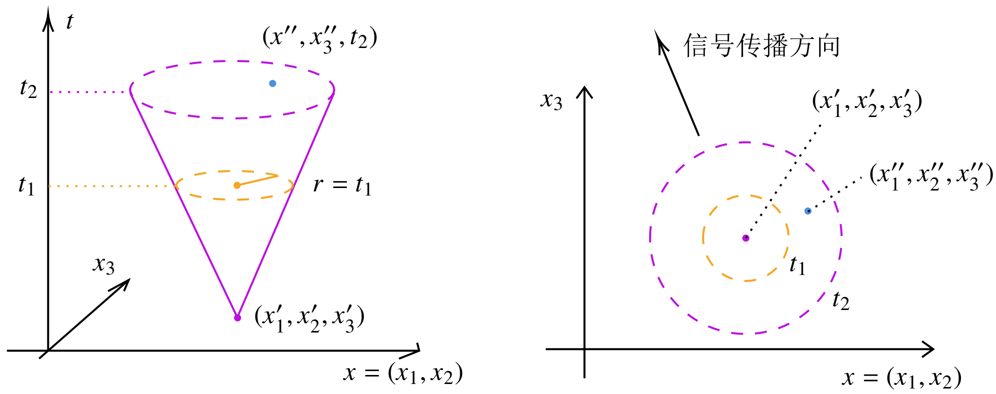


图 7.24: 单点初值的三维波传播示意图

如果讨论的是初值在某区域上均非零的三维波传播, 便可以采取与一维波讨论时类似的思路. 不妨设初值的支集是某圆盘 (如图(7.25)), 要研究空间中一点 (x', x_3') 接收信号的情况, 可以作经过 (x', x_3') 的截面, 截面内的情况便类似于一维波. 在图(7.25)右幅中, 强 Huygens 原理指的便是点 (x', x_3') 只会在图中所标示的时段接收到信号, 在经过该时段后该点再也不会接收到来自该区域的信号了.

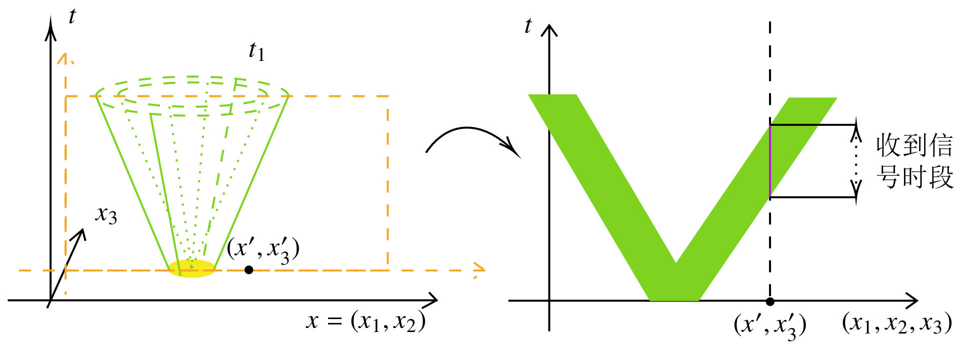


图 7.25: 区域初值的三维波传播示意图

区域初值的三维波传播在空间中的表现如图(7.26).

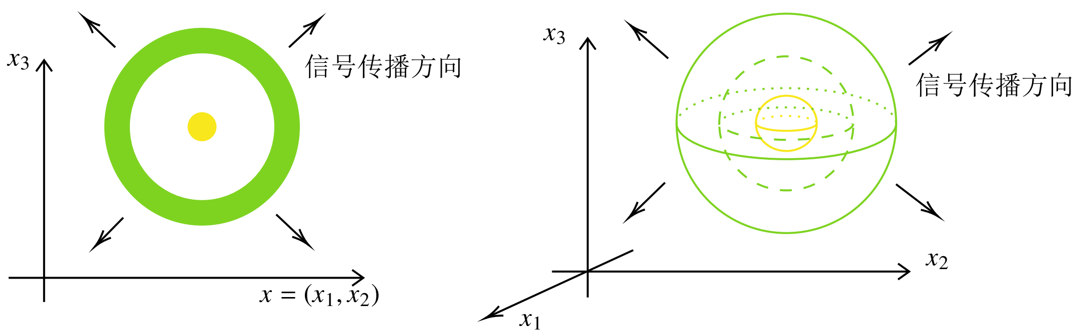


图 7.26: 区域初值的三维波在空间中的传播

现在来研究二维波的情况. 根据降维法的构造, 二维波与三维波在类似于图(7.24)的时空坐标系中, 最明显的区别在于不能再有 (三维意义下的) 单点初值了. 就算取二维意义下的单点初值, 因为降维法中初值在 x_3 方向上是常数, 故该单点初值在三维意义下表现为一条" 直线", " 直线" 上每一点的值都与 x 轴上对应的值相等. 如果

把这种情况也称作单点初值, 要研究单点初值的二维波传播, 就不可避免地出现顶点在直线上的一系列信号锥面 (如图(7.27)). 总而言之, 单点初值的二维波从本质上来说是直线初值的三维波.

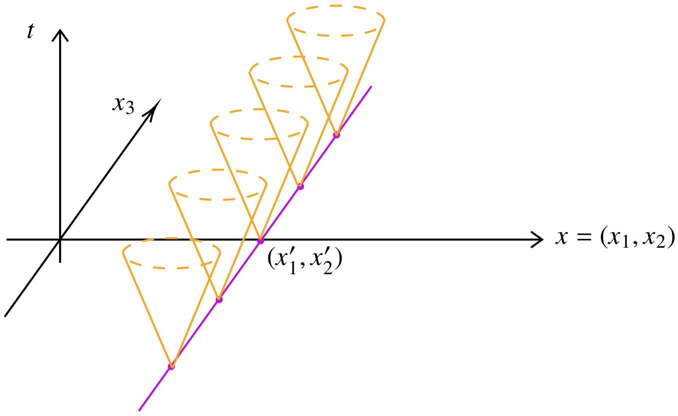


图 7.27: 单点初值的二维波在时空坐标中的传播示意图

现在对于二维空间 (即 x 轴) 上固定的一点 x'' 而言, 希望研究随着时间流逝它接收信号的情况. 因为从降维法的角度考虑, 单点初值二维波本质上是直线初值三维波, 故在 (虚拟的) 三维空间中, 现在研究的情况相当于是在二维点初值 (x_1', x_2') 处有一条平行于 x_3 的直线, 直线上的每一点都发出三维波. 于是开始计时后, 经过 t_1 时间, 来自 (x_1', x_2') 这一点的信号将会传到 x''; 经过 $t_2 = t_3$ 时间, 来自直线上另两点的信号将会传到 x''; 同样经过 $t_4 = t_5$ 时间, 来自直线上更远两点的信号将会传到 x''(如图(7.28))... 事实上, 在二维波的情况下, 只要一个点开始接收信号, 它在以后的时间里就会一直接受到来自这个信号发出者的信号, 这个现象的本质还是因为前面所说的二维波单点初值无法变成三维的单点, 而这个现象就称为二维波的弱 Huygens 原理.

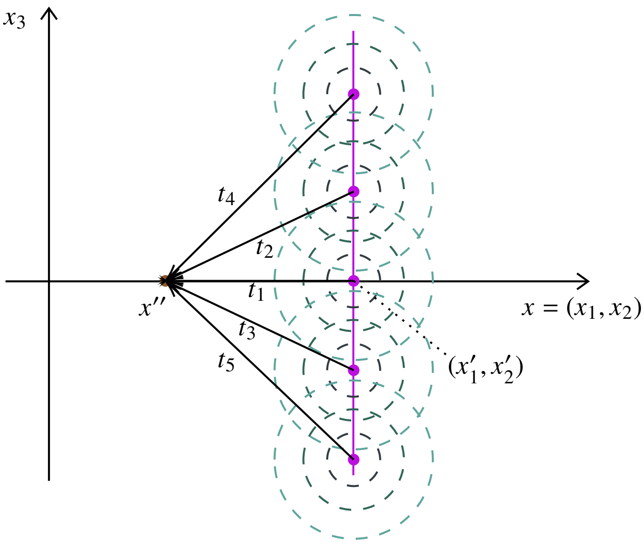


图 7.28: 单点初值的二维波在 (虚拟的) 三维空间中的传播示意图

7.4.2 非齐次问题

下面研究非齐次波方程的初值问题

$$\begin{cases} u_{tt} - \Delta u = f, & \text{在 } \mathbb{R}^n \times (0, \infty) \text{ 内} \\ u = 0, u_t = 0, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases} \tag{7.200}$$

还是采用先前提过但没细讲的 Duhamel 原理:

定理 7.4.4 (Duhamel 原理)

设 $U(t,\tau,x)$ 是初值问题

$$\begin{cases} U_{tt}(x,t;\tau) - \Delta U(x,t;\tau) = 0, & t > \tau, x \in \mathbb{R}^n \\ U|_{t=\tau} = U(x,\tau;\tau) = 0 \\ U_t|_{t=\tau} = U_t(x,\tau;\tau) = f(\tau,x) \end{cases} \tag{7.201}$$

的解, 那么函数

$$u(x,t) := \int_0^t U(x,t;\tau)d\tau \tag{7.202}$$

是非齐次问题

$$\begin{cases} u_{tt} - \Delta u = f, & t > 0, x \in \mathbb{R}^n \\ u|_{t=0} = u_t|_{t=0} = 0 \end{cases}$$

的解.

证明[OAO]

在(7.202)式两边同时对空间变量 x 取 Laplace 算子有

$$\Delta u(x,t) = \int_0^t \Delta U(x,t;\tau)d\tau$$

对 u 关于 t 求偏导有:

$$u_t(x,t) = \frac{\partial}{\partial t}\int_0^t U(x,t;\tau)d\tau = U(t,t,x) + \int_0^t U_t(x,t;\tau)d\tau$$

代入条件(7.201)知

$$u_t(x,t) = \int_0^t U_t(x,t;\tau)d\tau$$

进一步求导有

$$u_{tt}(x,t) = \frac{\partial}{\partial t}\int_0^t U_t(x,t;\tau)d\tau$$

同样代入条件(7.201)知

$$u_{tt}(t,x) = f(t,x) + \int_0^t U_{tt}(x,t;\tau)d\tau = f(t,x) + \int_0^t \Delta U(x,t;\tau)d\tau = f(t,x) + \Delta_x \int_0^t U(x,t;\tau)d\tau = f(t,x) + \Delta u$$

□

定理 7.4.5 (非齐次波方程的解)

若 $n \geq 2$, $f \in C^{[\frac{n}{2}]+1}(\mathbb{R}^n \times [0,\infty))$, u 由(7.202)式定义, 则

(i) $u \in C^2(\mathbb{R}^n \times [0,\infty))$,

(ii) 在 $\mathbb{R}^n \times (0,\infty)$ 上有 $u_{tt} - \Delta u = f$,

(iii) 对 $x^0 \in \mathbb{R}^n$ 有 $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n, t>0}} u(x,t) = 0$, $\lim\limits_{\substack{(x,t)\to(x^0,0)\\ x\in\mathbb{R}^n, t>0}} u_t(x,t) = 0$.

证明

只需证明 (i) 即可. 若 n 是奇数, 则 $[\frac{n}{2}]+1 = \frac{n+1}{2}$. 根据定理(7.4.2)知对每个 $s \geq 0$ 都有 $U(\cdot,\cdot;s) \in C^2(\mathbb{R}^n \times [s,\infty))$, 进而 $u \in C^2(\mathbb{R}^n \times [0,\infty))$. 若 n 是偶数, 则 $[\frac{n}{2}]+1 = \frac{n+2}{2}$, 由定理(7.4.3)同理可得 $u \in C^2(\mathbb{R}^n \times [0,\infty))$. □

一般的非齐次问题的解其实就是初值问题(7.163)的解与非齐次初值问题(7.200)的解之和.

例 7.1 下面演示 $n = 1$ 时对应的非齐次问题(7.200)具体怎么求解. 给定 s, 由 d'Alembert 公式(7.160)已经知道

$$u(x,t;s) = \frac{1}{2}\int_{x-t+s}^{x+t-s} f(y,s)dy \Rightarrow u(x,t) = \frac{1}{2}\int_0^t \int_{x-t+s}^{x+t-s} f(y,s)dyds$$

即

$$u(x,t) = \frac{1}{2}\int_0^t \int_{x-s}^{x+s} f(y,t-s)dyds, \quad x \in \mathbb{R}, t \geq 0.$$

例 7.2 $n = 3$ 时, Kirchhoff 公式(7.174)表明:

$$u(x,t;s) = (t-s)\fint_{\partial B(x,t-s)} f(y,s)dS$$

故

$$u(x,t) = \int_0^t (t-s)(\fint_{\partial B(x,t-s)} f(y,s)dS)ds = \frac{1}{4\pi}\int_0^t \int_{\partial B(x,t-s)} \frac{f(y,s)}{(t-s)}dSds = \frac{1}{4\pi}\int_0^t \int_{\partial B(x,r)} \frac{f(y,t-r)}{r}dSds$$

得

$$u(x,t) = \frac{1}{4\pi}\int_{B(x,t)} \frac{f(y,t-|y-x|)}{|y-x|}dy, \quad x \in \mathbb{R}^3, t \geq 0$$

这便是 $n = 3$ 时非齐次问题(7.200)的解. 上式右侧积分称为推迟势.

7.4.2.1 补充: Duhamel 原理的动机

远在热方程时我们就已经在用 Duhamel 原理解决非齐次问题, 本节旨在阐明 Duhamel 原理在热方程与波方程中的动机, 并给出一般常系数线性发展方程的 Duhamel 原理.

† **热方程**

首先回忆热方程 Duhamel 原理的内容:

定理 7.4.6 (Duhamel)

若 $w(x,t;s)$ 是方程

$$\begin{cases} w_t(x,t;s) - \Delta w(x,t;s) = 0, & \text{在 } \mathbb{R}^n \times (s,+\infty) \text{ 内} \\ w(x,t;s) = f(x,s), & \text{在 } \mathbb{R}^n \times \{t = s\} \text{ 上} \end{cases}$$

的解, 那么 $u(x,t) = \int_0^t w(x,t;s)ds$ 就是方程

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = f(x,t), & \text{在 } \mathbb{R}^n \times (0,+\infty) \text{ 内} \\ u(x,t) = 0, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases}$$

的解.

要理解 Duhamel 原理的动机, 本质上是理解非齐次方程 $u_t - \Delta u = f$ 中 f 的含义. 回忆在研究 Laplace 方程的 Green 函数法时, 曾经给出了热方程的物理意义式(7.60). 简单来说, $f(x,t)$ 表示研究区域内热源 (热汇) 的情况: 在点 x 处 t 时刻是否有额外的热源 (热汇) 提供 (吸收) 热量. 注意 u 表示一点的温度, 而温度的变化量与它内能的变化量成正比, 内能的变化量又体现在传输给它 (从它耗散) 的热量的累计. 现在确定需要研究的时刻 t_0, 对于初始时刻 $t = 0$ 到研究时刻 $t = t_0$ 之间的时间段, $f(x,t)$ 是在不断变化的. 对一个特定时刻 $s \in [0,t_0]$, 此时的热源[13] $f(x,s)$ 在瞬间产生了一个热量, 这个瞬时产生的热量依照热方程进行传播, 它在空间中产生的温度变化就用下述方程衡量:

$$\begin{cases} w_t(x,t;s) - \Delta w(x,t;s) = 0, & \text{在 } \mathbb{R}^n \times (s,+\infty) \text{ 内} \\ w(x,t;s) = f(x,s), & \text{在 } \mathbb{R}^n \times \{t = s\} \text{ 上} \end{cases}$$

其中 w 是 t 的函数这一点可以理解为: 在热源瞬间产生一个热量后, 这个热量随着时间流逝而逐渐扩散, 顺带着在扩散到的区域引起温度的变化, 特别在 t_0 时刻单个 w 对应的变化是固定的. 将这样一些变化沿时间 s 从 0 到 t_0

[13] 为了表述方便姑且就默认 $f(x,t) > 0$ 了, 在 $f(x,t) < 0$ 时它作为热汇有类似的效果.

累积起来, 得到的便是最终的温度 u, 因而

$$u(x,t_0)=\int_0^{t_0} w(x,t_0;s)ds.$$

† 波方程

这一部分内容在众多数学物理方程教材中都能找到, 这里参考了 **[GLC]**. 现在回忆波方程 Duhamel 原理的内容:

定理 7.4.7 (Duhamel)

若 $w(x,t;s)$ 是方程

$$\begin{cases} w_{tt}(x,t;s)-\Delta w(x,t;s)=0, & \text{在 } \mathbb{R}^n\times(s,+\infty) \text{ 内} \\ w(x,t;s)=0, w_t(x,t;s)=f(x,s), & \text{在 } \mathbb{R}^n\times\{t=s\} \text{ 上} \end{cases}$$

的解, 那么 $u(x,t)=\int_0^t w(x,t;s)ds$ 就是方程

$$\begin{cases} u_{tt}(x,t;s)-\Delta u(x,t;s)=f(x,t), & \text{在 } \mathbb{R}^n\times(0,+\infty) \text{ 内} \\ u(x,t;s)=0, u_t(x,t;s)=0, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上} \end{cases}$$

的解.

波方程的解 u 本质上是位移, 因而 $f(x,t)$ 作为位移关于时间的二阶导数本质上是力. 与热方程 Duhamel 原理的思路类似, 固定研究时刻 t_0, 考虑 $[0,t_0]$ 中的一个特定时刻 s, 现在希望探讨 s 时刻瞬间施加的力 $f(x,s)$ 对后续运动产生的影响, 其中" 瞬间" 理解为在微小时段内. 为了刻画" 瞬间" 的影响, 将研究时段 $[0,t_0]$ 分成微小时段 $[0,t_1],[t_1,t_2],\cdots,[t_n,t_0]$, 并记 $\delta t=t_0-t_n=t_i-t_{i-1}(i=2,\cdots,n)$. 对时段 $[t_i,t_{i+1}]$, 要考察力 $f(x,t)$ 产生的影响, 根据 f 的连续性可以近似视作考察 $f(x,t_i)$ 产生的影响, 从而回忆冲量定理:

$$f(x,t_i)\delta t=m\delta v$$

其中 δv 是速度的改变量. 如果把速度的改变量视作初速度, 则在该时段内施加的力 $f(x,t_i)$ 对后续的位移产生的位移变化量 $\widetilde{w}(x,t;t_i)$ 满足:

$$\begin{cases} \widetilde{w}_{tt}(x,t;t_i)-\Delta\widetilde{w}(x,t;t_i)=0, & \text{在 } \mathbb{R}^n\times(t_i,+\infty) \text{ 内} \\ \widetilde{w}(x,t;t_i)=0, \widetilde{w}_t(x,t;t_i)=f(x,t_i)\delta t, & \text{在 } \mathbb{R}^n\times\{t=t_i\} \text{ 上} \end{cases}$$

于是最后的位移应为

$$u(t,x)=\lim_{\delta t\to 0}\sum_{i=1}^n \widetilde{w}(t,x;t_i)$$

注意如果施加的力 $f(x,t_i)$ 是恒力, 这便意味着此时质点作恒加速度运动, 因而速度变化量 $\widetilde{w}_t(x,t;t_i)$ 与时间 δt 成正比, 故可设

$$\widetilde{w}(x,t;t_i)=w(x,t;t_i)\delta t$$

根据方程的线性性可得

$$\begin{cases} w_{tt}(x,t;t_i)-\Delta w(x,t;t_i)=0, & \text{在 } \mathbb{R}^n\times(t_i,+\infty) \text{ 内} \\ w(x,t;t_i)=0, w_t(x,t;t_i)=f(x,t_i), & \text{在 } \mathbb{R}^n\times\{t=t_i\} \text{ 上} \end{cases}$$

随着对 $[0,t_0]$ 分划的加细, 可知 $[0,t_0]$ 中的任意时刻 s 均可作为 w 的参数, 且有:

$$\begin{cases} w_{tt}(x,t;s)-\Delta w(x,t;s)=0, & \text{在 } \mathbb{R}^n\times(s,+\infty) \text{ 内} \\ w(x,t;s)=0, w_t(x,t;s)=f(x,s), & \text{在 } \mathbb{R}^n\times\{t=s\} \text{ 上} \end{cases}$$

同时

$$u(t,x)=\lim_{\delta t\to 0}\sum_{i=1}^n w(x,t;t_i)\delta t=\int_0^t w(x,t;s)ds.$$

† **一种一般的 Duhamel 原理**

诸多 PDE 的教科书中似乎没有介绍一般的 Duhamel 原理, 这里作出整理. 这部分内容也可以参考 **[CH]** 与 CSDN 博客: 线性叠加原理和齐次化原理 | 偏微分方程 (十二).

定理 7.4.8 (Duhamel)

设 L 是关于时间变量 $t \in [0, +\infty)$ 与空间变量 $x \in \mathbb{R}^n$ 的线性偏微分算子, 其中关于 t 的求导阶数不超过 $m-1$, 则若 w 满足初值问题

$$\begin{cases} \frac{\partial^m w}{\partial t^m}(x,t;s) - Lw(x,t;s) = 0, & \text{在 } \mathbb{R}^n \times (s, +\infty) \text{ 内} \\ w(x,t;s) = w_t(x,t;s) = \cdots = \frac{\partial^{m-2} w}{\partial t^{m-2}}(x,t;s) = 0, & \text{在 } \mathbb{R}^n \times \{t = s\} \text{ 上} \\ \frac{\partial^{m-1} w}{\partial t^{m-1}}(x,t;s) = f(x,s), & \text{在 } \mathbb{R}^n \times \{t = s\} \text{ 上} \end{cases}$$

则 $u(x,t) = \int_0^t w(x,t;s)ds$ 是非齐次初值问题

$$\begin{cases} \frac{\partial^m u}{\partial t^m}(x,t) - Lu(x,t) = f(x,t), & \text{在 } \mathbb{R}^n \times (0, +\infty) \text{ 内} \\ u(x,t) = u_t(x,t) = \cdots = \frac{\partial^{m-1} u}{\partial t^{m-1}}(x,t) = 0, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases}$$

的解.

♡

证明

知

$$\begin{aligned} \frac{\partial^m u}{\partial t^m} &= \sum_{i=1}^{m-1} \frac{\partial^i w}{\partial t^i}(x,t;t) + \int_0^t \frac{\partial^m w}{\partial t^m}(x,t;s)ds \\ &= f(x,t) + \int_0^t \frac{\partial^m w}{\partial t^m}(x,t;s)ds = f(x,t) + \int_0^t Lw(x,t;s)ds \\ &\overset{\text{(i)}}{=} f(x,t) + L\int_0^t w(x,t,s)ds = f(x,t) + Lu(x,t) \end{aligned}$$

其中 (i) 是因为若设 $L = \sum_{i=1}^{r} a_i(x) \frac{\partial^i}{\partial t^i} + L_x$, 其中 $r \le m-1$, L_x 是仅关于 x 的线性偏微分算子, 则

$$L \int_0^t w(x,t,s)ds = \sum_{i=1}^{r} a_i(x) \frac{\partial^i w}{\partial t^i} w(x,t;t) + \int_0^t Lw(x,t;s)ds = \int_0^t Lw(x,t;s).$$

命题即证. □

7.4.3 能量方法

奇数维波方程解公式(7.187)与偶数维波方程解公式(7.199)在维数 n 越来越大时, 都要求 g, h 的光滑性越来越好, 以此来保证波方程 C^2 解的存在性. 这暗示着或许还有其他更合适的方法用来衡量函数的大小和光滑性. 本节就来说明波方程 (对全体 n) 在特定" 能量模" 上都有良好性态.

7.4.3.1 唯一性

设 $U \subset \mathbb{R}^n$ 是有界开集, 且边界 ∂U 光滑. 同时设 $U_T = U \times (0, T], \Gamma_T = \overline{U_T} \backslash U_T$, 其中 $T > 0$. 我们感兴趣的是初边值问题

$$\begin{cases} u_{tt} - \Delta u = f, & \text{在 } U_T \text{ 内} \\ u = g, & \text{在 } \Gamma_T \text{ 上} \\ u_t = h, & \text{在 } U \times \{t = 0\} \text{ 上} \end{cases} \tag{7.203}$$

定理 7.4.9 (波方程解的唯一性)

初边值问题(7.203)至多存在一个解 $u \in C^2(\overline{U_T})$.

证明

设 $\tilde{u}$ 是另一个满足要求的解, 则 $w := u - \tilde{u}$ 是

$$\begin{cases} w_{tt} - \Delta w = 0, & \text{在 } U_T \text{ 内} \\ w = 0, & \text{在 } \Gamma_T \text{ 上} \\ w_t = 0, & \text{在 } U \times \{t = 0\} \text{ 上} \end{cases}$$

的解. 定义" 能量":

$$E(t) := \frac{1}{2} \int_U w_t^2(x,t) + |Dw(x,t)|^2 dx, \quad 0 \le t \le T$$

计算知

$$\frac{dE}{dt}(t) = \dot{E}(t) = \int_U (w_t w_{tt} + Dw \cdot Dw_t) dx = \int_U w_t (w_{tt} - \Delta w) dx = 0$$

其中因为 w 在 ∂U 上恒为零, 故上述结果中没有在 ∂U 上的积分. 同时因为在 $\partial U \times [0,T]$ 上 $w_t \equiv 0$, 故对任意的 $0 \le t \le T$ 都有 $E(t) = E(0) = 0$, 进而在 U_T 内只能有 $w_t, Dw \equiv 0$. 又因为在 $U \times \{t = 0\}$ 上也有 $w \equiv 0$, 故在 U_T 有 $w = u - \tilde{u} \equiv 0$. □

7.4.3.2 依赖区域

下面研究波方程在全空间中的依赖区域. 设 $u \in C^2$ 是

$$u_{tt} - \Delta u = 0, \quad \text{在 } \mathbb{R}^n \times (0, \infty) \text{ 中}$$

的解. 固定 $x_0 \in \mathbb{R}^n, t_0 > 0$, 考虑以 (x_0, t_0) 为顶点的反向波锥:

$$K(x_0, t_0) := \{(x,t) : 0 \le t \le t_0, |x - x_0| \le t_0 - t\}$$

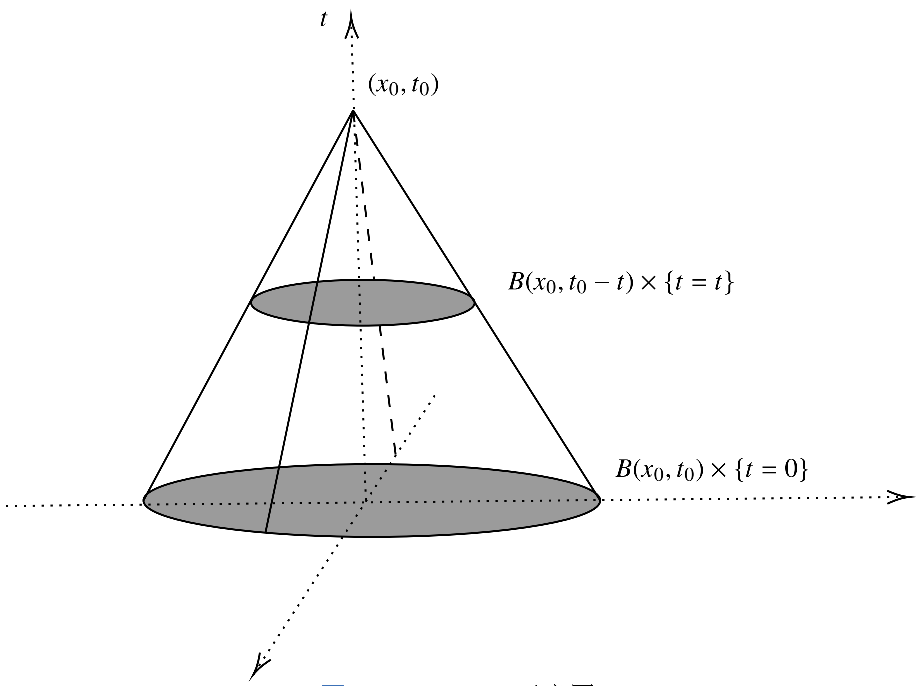


图 7.29: $K(x_0, t_0)$ 示意图

定理 7.4.10 (有限传播速度)

若在 $B(x_0,t_0)\times\{t=0\}$ 中有 $u\equiv u_t\equiv 0$, 则在锥体 $K(x_0,t_0)$ 中有 $u\equiv 0$.

这其实就是前面对波的 Huygens 原理的另一种阐述. 可以这么理解: 当 $t=0$ 时, 如果在 $B(x_0,t_0)$ 之外都发生了扰动, 显见这个扰动在 $t=0$ 时是传不到 $B(x_0,t_0)$ 内的. 但随着时间 t 的推移, 外部的扰动逐步向 $B(x_0,t_0)$ 内传播, 在 t 时刻就只剩 $B(x_0,t_0-t)$ 的区域还没有被扰动了. 当时间进一步推移, $t=t_0$ 时, 除了 (x_0,t_0) 这一点 (也即空间中的 x_0 点), 空间中的其余地方都被扰动了, 而在下一刻 x_0 处也会被扰动. 图(7.29)其实就是前面 Huygens 原理的示意图加上了时间轴 t. 这一点用能量方法其实很容易就能说明.

证明

定义局部能量

$$e(t):=\frac{1}{2}\int_{B(x_0,t_0-t)}(u_t^2(x,t)+|Du(x,t)|^2)dx,\quad 0\le t\le t_0$$

则

$$\begin{aligned}\dot{e}(t)&=\int_{B(x_0,t_0-t)}u_tu_{tt}+Du\cdot Du_tdx-\frac{1}{2}\int_{\partial B(x_0,t_0-t)}(u_t^2+|Du|^2)dS\\&=\int_{B(x_0,t_0-t)}u_t(u_{tt}-\Delta u)dx+\int_{\partial B(x_0,t_0-t)}\frac{\partial u}{\partial\nu}u_tdS-\frac{1}{2}\int_{\partial B(x_0,t_0-t)}(u_t^2+|Du|^2)dS\\&=\int_{\partial B(x_0,t_0-t)}(\frac{\partial u}{\partial\nu}u_t-\frac{1}{2}u_t^2-\frac{1}{2}|Du|^2)dS\end{aligned}\tag{7.204}$$

现在由 Cauchy-Schwarz 不等式得

$$|\frac{\partial u}{\partial\nu}u_t|\le|u_t|\cdot|Du|\le\frac{1}{2}u_t^2+\frac{1}{2}|Du|^2$$

这说明 $\dot{e}(t)\le 0$, 故对全体 $0\le t\le t_0$ 都有 $e(t)\le e(0)=0$, 进而只能有 $u_t,Du\equiv 0$, 因而在 $K(x_0,t_0)$ 内总有 $u\equiv 0$. □

注 有的教科书会把圆盘 $B(x_0,t_0)$ 称为点 (x_0,t_0) 的依赖区域, 这可以从解公式的积分区域来解释: $u(x_0,t_0)$ 对应的公式只在 $B(x_0,t_0)$ 内积分. 圆锥 $K(x_0,t_0)$ 称为圆盘 $B(x_0,t_0)$ 的决定区域, 这是因为根据定理(7.4.10), 圆锥 $K(x_0,t_0)$ 内的取值是由圆盘 $B(x_0,t_0)$ 上的情况决定的.

对于受扰动区域而言, 若在 $t=0$ 时点 x_0 受扰动, 根据前面对 Huygens 原理的讨论, 这个扰动随着时间的推移会向外扩散, 体现在时空图中即为图(7.30)中的圆锥. 称该圆锥体位初始平面上点 $(x_0,0)$ 的影响区域, 初始平面上给定区域 (比如讨论 Huygens 原理时的区域 G) 的影响区域, 就是该区域内每一点的影响区域的包络. 而 $(x_0,0)$ 对应影响区域的锥面就称为波动方程的特征锥.

7.5 问题

在下面的练习中, 给出的所有函数若无特殊说明, 都是光滑的.

练习 7.1♣ 求初值问题

$$\begin{cases}u_t+b\cdot Du+cu=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 中}\\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

的精确解 u, 其中 $v\in\mathbb{R}, b\in\mathbb{R}^n$ 是常数.

解

设 $z(s):=u(x+sb,t+s)$, 则知:

$$\dot{z}(s)=bDu(x+sb,t+s)+u_t(x+sb,t+s)=-cu(x+sb,t+s)=-cz(s)$$

解 ODE 得 $z(s)=Ce^{-cs}$, 其中 C 与 s 无关. 注意连接 (x,t) 与 $(b,1)$ 的直线交 $\mathbb{R}^n\times\{t=0\}$ 于 $(x-tb,0)$, 此时 $s=-t$. 故 $z(-t)=Ce^{ct}=g(x-tb)$, 得到 $C=g(x-tb)e^{-ct}$.

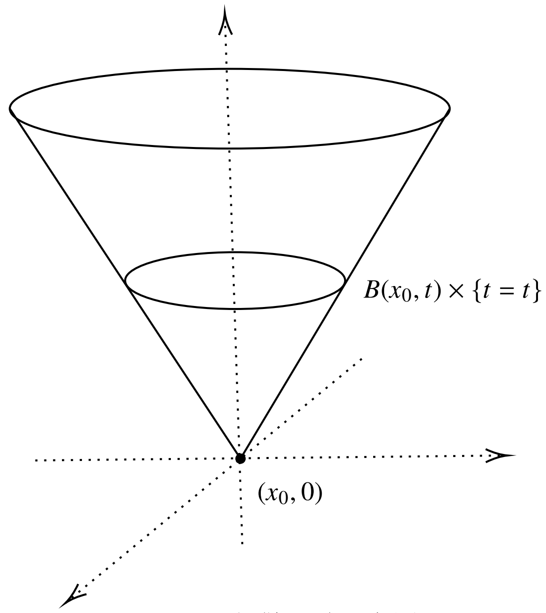


图 7.30: 扩散区域示意图

最后, 对任意的 $(x, t) \in \mathbb{R}^n \times (0, \infty)$, 知

$$u(x, t) = z(0) = g(x - tb)e^{-ct}$$

此即欲求. □

✍ **练习 7.2**♣ 证明 Laplace 方程 $\Delta u = 0$ 对旋转是不变的, 也就是说, 如果 O 是 $n \times n$ 正交矩阵, 定义

$$v(x) := u(Ox), \quad x \in \mathbb{R}^n$$

则 $\Delta v = 0$.

证明

设 $O = (a_{ij})_{n\times n}, x = (x_1, \cdots, x_n)^T$, 知

$$Ox = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k}x_k \\ \vdots \\ \sum_{k=1}^{n} a_{nk}x_k \end{pmatrix}$$

从而

$$\frac{\partial v}{\partial x_i}(x) = a_{1i}u_{x_1}(Ox) + \cdots + a_{ni}u_{x_n}(Ox) = \sum_{j=1}^{n} a_{ji}u_{x_j}(Ox)$$

进一步:

$$\frac{\partial}{\partial x_i}(\frac{\partial v}{\partial x_i}) = a_{1i}\sum_{j=1}^{n} a_{ji}u_{x_1x_j}(Ox) + \cdots + a_{ni}\sum_{j=1}^{n} a_{ji}u_{x_nx_j}(Ox) = \sum_{p=1}^{n}\sum_{j=1}^{n} a_{pi}a_{ji}u_{x_px_j}(Ox)$$

故

$$\Delta v(x) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(\frac{\partial v}{\partial x_i}) = \sum_{i=1}^{n}\sum_{p=1}^{n}\sum_{j=1}^{n} a_{pi}a_{ji}u_{x_px_j}(Ox)$$

每固定一组 $(p, j)(p \neq j)$, 知 $\Delta v(x)$ 中 $u_{x_px_j}$ 对应的系数为 $\sum_{i=1}^{n} a_{pi}a_{ji}$, 而这正是 O 中两不同行的内积, 故

$$\sum_{i=1}^{n} a_{pi}a_{ij} = 0$$

这说明

$$\Delta v(x) = \sum_{i=1}^{n} \sum_{1\le j=p\le n} a_{pi}a_{ji}u_{x_p x_j}(Ox) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ji}^2 u_{x_j x_j}(Ox)$$

根据正交阵的正交性, 知对每个固定的 j, 都有 $\sum\limits_{i=1}^{n} a_{ji}^2 = 1$, 故

$$\Delta v(x) = \sum_{j=1}^{n} u_{x_j x_j}(Ox)$$

而已知 $\Delta u = 0$, 故 $\Delta v(x) = 0$. □

注 [14]答案与老师上课的方法似乎仍存疑问: 这与把 $\Delta v(x) = 0$ 中的 $v(x)$ 按定义直接换成 $\Delta u(Ox) = 0$ 有何区别? 亦即如何由

$$\Delta u(x) = (\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2})u(x) = 0$$

直接推出

$$\Delta u(Ox) = (\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2})u(Ox) = 0$$

而后者似乎就是欲证? 如果考虑换元, 设 $O = (a_{ij})$ 是正交阵, $y = Ox = (y_1, \cdots, y_n)^T$, 则只有:

$$(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2})u(x_1, \cdots, x_n) = (\frac{\partial^2}{\partial y_1^2} + \cdots + \frac{\partial^2}{\partial y_n^2})u(y_1, \cdots, y_n) = 0$$

而欲证即

$$(\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2})u(y_1, \cdots, y_n) = 0$$

为了具体计算, 简记:

$$\nabla_y = \begin{pmatrix} \frac{\partial}{\partial y_1} \\ \vdots \\ \frac{\partial}{\partial y_n} \end{pmatrix}, \quad \nabla_x = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \quad \Delta_y = \nabla_y \cdot \nabla_y, \quad \Delta_x = \nabla_x \cdot \nabla_x$$

将 $y = Ox$ 写开有:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \Rightarrow \begin{cases} y_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ y_n = a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases}$$

不妨设 $O^{-1} = (b_{ij})(=: B)$, 知 B 也是正交阵, 解得:

$$\begin{cases} x_1 = b_{11}y_1 + \cdots + b_{1n}y_n \\ \vdots \\ x_n = b_{n1}y_1 + \cdots + b_{nn}y_n \end{cases}$$

从而:

$$\begin{aligned} \frac{\partial}{\partial y_1} &= \frac{\partial}{\partial x_1}\frac{dx_1}{dy_1} + \cdots + \frac{\partial}{\partial x_n}\frac{dx_n}{dy_1} = b_{11}\frac{\partial}{\partial x_1} + \cdots + b_{1n}\frac{\partial}{\partial x_1} \\ &\vdots \\ \frac{\partial}{\partial y_n} &= \frac{\partial}{\partial x_1}\frac{dx_1}{dy_n} + \cdots + \frac{\partial}{\partial x_n}\frac{dx_n}{dy_n} = b_{n1}\frac{\partial}{\partial x_1} + \cdots + b_{nn}\frac{\partial}{\partial x_n} \end{aligned}$$

[14]2023.9.21 记.

写成矩阵形式即:

$$\nabla_y = \begin{pmatrix} \frac{\partial}{\partial y_1} \\ \cdots \\ \frac{\partial}{\partial y_n} \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} = B \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} = B\nabla_x$$

从而:

$$\Delta_y = \nabla_y \cdot \nabla_y = (\nabla_y)^T \nabla_y = (B\nabla_x)^T (B\nabla_x) = (\nabla_x)^T B^T B \nabla_x = (\nabla_x)^T \nabla_x = \nabla_x \cdot \nabla_x = \Delta_x$$

这说明

$$\Delta_y u(y) = (\frac{\partial^2}{\partial y_1^2} + \cdots + \frac{\partial^2}{\partial y_n^2}) u(y_1, \cdots, y_n) = (\frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}) u(y_1, \cdots, y_n) = \Delta_x u(y) = \Delta_x u(Ox)$$

而由 $\Delta_y u(y) = 0$ 即得 $\Delta_x u(Ox) = 0$, 亦即 $\Delta_x v(x) = 0$. □

注 [15] 小组讨论后, 下面介绍另一位同学的想法: 记 $y = Ox$, 则有 $v(x) = u(y)$, 设 $O = (a_{ij})$, 写开有:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = Ox = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1i} x_i \\ \vdots \\ \sum_{i=1}^n a_{ni} x_i \end{pmatrix}$$

从而:

$$\frac{\partial}{\partial x_i} v(x) = \frac{\partial}{\partial x_i} u(y) = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial}{\partial y_k} u(y) = \sum_{k=1}^n a_{ki} \frac{\partial u(y)}{\partial y_k}$$

进一步有

$$\frac{\partial^2}{\partial x_i \partial x_j} v(x) = \frac{\partial}{\partial x_j} (\sum_{k=1}^n a_{ki} \frac{\partial u(y)}{\partial y_k}) = (\sum_{p=1}^n \frac{\partial y_p}{\partial x_j} \frac{\partial}{\partial y_p})(\sum_{k=1}^n a_{ki} \frac{\partial u(y)}{\partial y_k}) = \sum_{\substack{1 \le k \le n \\ 1 \le p \le n}} a_{pj} a_{ki} \frac{\partial^2 u(y)}{\partial y_k \partial y_p}$$

写成矩阵形式, 记 $v_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} v(x), (u_{ij}) = \frac{\partial^2}{\partial y_k \partial y_p} u(y)$, 有:

$$(v_{ij}) = O^T (u_{ij}) O$$

从而

$$\mathrm{tr}\left(v_{ij}\right) = \mathrm{tr}\left(O^T (u_{ij}) O\right) = \mathrm{tr}\left(O^T O (u_{ij})\right) = \mathrm{tr}\left(u_{ij}\right)$$

同时

$$\mathrm{tr}\left(v_{ij}\right) = \sum_{k=1}^n \frac{\partial^2}{\partial x_k^2} v(x) = \Delta_x v(x) \, \mathrm{tr}\left(u_{ij}\right) = \sum_{k=1}^n \frac{\partial^2}{\partial y_k^2} u(y) = \Delta_y u(y)$$

这便有

$$\Delta_x v(x) = \Delta_y u(y)$$

最后因为 $\Delta_y u(y) = 0$, 故 $\Delta_x v(x) = 0$, 命题即证. □

注 [16]课堂上的方法实际上并未出错. 前面 $\Delta v(x) = 0$ 直接替换成 $\Delta u(Ox) = 0$ 这句话的问题在于 Δ 算子是否真正应用到函数上了. 题目已知的条件应该写成 $(\Delta u)(x) = 0$, 而前面做的替换应该写成 $\Delta[u(Ox)] = 0$, 此时还没有具体做微分运算. 按照这种思路有:

$$\Delta[u(Ox)] = \nabla^T \nabla [u(Ox)] = \nabla^T (O(\nabla u)(Ox)) = OO^T (\nabla^T \nabla u)(Ox) = OO^T (\Delta u)(Ox)$$

虽说自变量 Ox 与求导分量 x 并不一致, 但 Δu 本身就是常值 0, 所以代入任何值得到的结果都是 0, 这才得到

[15]2023.9.28 记.

[16]2023.10.23 记.

$\Delta[u(Ox)] = 0$, 亦即 $\Delta v = 0$.

练习 7.3♣ 改进均值定理的证明, 说明 $n \geq 3$ 时只要

$$\begin{cases} -\Delta u = f, & \text{在 } B^\circ(0,r) \text{ 内} \\ u = g, & \text{在 } \partial B(0,r) \text{ 上} \end{cases}$$

则

$$u(0) = ⨍_{\partial B(0,r)} g dS + \frac{1}{n(n-2)\alpha(n)} \int_{B(0,r)} (\frac{1}{|x|^{n-2}} - \frac{1}{r^{n-2}}) f dx.$$

证明

同样研究 $x = 0$ 邻域内 u 的积分平均, 设

$$\phi(r) = ⨍_{\partial B(0,r)} u(x) dS(x) = ⨍_{\partial B(0,1)} u(ry) dS(y)$$

则

$$\begin{aligned} \phi'(r) &= ⨍_{\partial B(0,1)} Du(ry) \cdot y dS(y) = ⨍_{\partial B(0,r)} Du(x) \cdot \frac{x}{r} dS(x) \\ &= \frac{1}{n\alpha(n) r^{n-1}} \int_{\partial B(0,r)} Du(x) \cdot \frac{x}{r} dS(x) = \frac{1}{n\alpha(n) r^{n-1}} \int_{B(0,r)} \Delta u(x) dx \\ &= -\frac{1}{n\alpha(n) r^{n-1}} \int_{B(0,r)} f(x) dx \end{aligned}$$

因为 $\phi'(r)$ 在 $r = 0$ 处可能出现奇性, 故对任意的 $\varepsilon > 0$, 考虑

$$\begin{aligned} \phi(r) - \phi(\varepsilon) &= \int_\varepsilon^r \phi'(s) ds = -\int_\varepsilon^r \frac{1}{n\alpha(n) s^{n-1}} \int_{B(0,s)} f(x) dx ds = \frac{1}{n(n-2)\alpha(n)} \int_\varepsilon^r (\int_{B(0,s)} f(x) dx) ds^{2-n} \\ &= \frac{r^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,r)} f(x) dx - \frac{\varepsilon^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,\varepsilon)} f(x) dx - \frac{1}{n(n-2)\alpha(n)} \int_\varepsilon^r s^{2-n} d(\int_{B(0,s)} f(x) dx) \\ &= \frac{r^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,r)} f(x) dx - \frac{\varepsilon^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,\varepsilon)} f(x) dx - \frac{1}{n(n-2)\alpha(n)} \int_\varepsilon^r s^{2-n} (\int_{\partial B(0,s)} f(x) dx) ds \end{aligned}$$

现令 $\varepsilon \to 0$, 对 $\phi(\varepsilon)$, 根据 u 的连续性有:

$$\phi(\varepsilon) = ⨍_{\partial B(0,\varepsilon)} u(x) dS(x) \to u(0), \quad \varepsilon \to 0$$

而对 $-\frac{\varepsilon^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,\varepsilon)} f(x) dx$, 根据 f 的连续性有:

$$-\frac{\varepsilon^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,\varepsilon)} f(x) dx \sim -\frac{\varepsilon^{2-n}}{n(n-2)\alpha(n)} \cdot \alpha(n) \varepsilon^n f(\xi) = \frac{f(\xi)}{n(n-2)} \varepsilon^2 \to 0, \quad \varepsilon \to 0, \xi \in B(0,\varepsilon)$$

这说明:

$$\begin{aligned} \phi(r) - u(0) &= \frac{r^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,r)} f(x) dx - \frac{1}{n(n-2)\alpha(n)} \int_0^r s^{2-n} (\int_{\partial B(0,s)} f(x) dx) ds \\ &= \frac{r^{2-n}}{n(n-2)\alpha(n)} \int_{B(0,r)} f(x) dx - \frac{1}{n(n-2)\alpha(n)} \int_{B(0,r)} |x|^{2-n} f(x) dx \end{aligned}$$

整理得

$$u(0) = \phi(r) + \frac{1}{n(n-2)\alpha(n)} \int_{B(0,r)} (\frac{1}{|x|^{n-2}} - \frac{1}{r^{n-2}}) f(x) dx$$

最后, 根据 $\phi(r)$ 的定义与 $u(x) = g(x) (x \in \partial B(0,r))$, 得:

$$\phi(r) = ⨍_{\partial B(0,r)} g(x) dS(x)$$

综上得到

$$u(0) = ⨍_{\partial B(0,r)} g(x) dS(x) + \frac{1}{n(n-2)\alpha(n)} \int_{B(0,r)} (\frac{1}{|x|^{n-2}} - \frac{1}{r^{n-2}}) f(x) dx$$

此即欲证. □

✍ **练习 7.4**♣ 给出下述命题的直接证明: 若 $u \in C^2(U) \cap C(\overline{U})$ 是有界开集 U 中的调和函数, 则

$$\max_{\overline{U}} u = \max_{\partial U} u.$$

证明

任取 $\varepsilon > 0$, 设

$$u_\varepsilon(x) := u(x) + \varepsilon |x|^2$$

下面说明 u_ε 在 $\overline{U}$ 的内部取不到最大值. 知:

$$Du_\varepsilon(x) = Du(x) + \varepsilon \cdot (2x_1, 2x_2, \cdots, 2x_n) = (u_{x_1} + 2\varepsilon x_1, \cdots, u_{x_n} + 2\varepsilon x_n)$$

进一步有

$$D^2 u_\varepsilon(x) = \begin{pmatrix} u_{x_1x_1} + 2\varepsilon & u_{x_1x_2} & \cdots & u_{x_1x_n} \\ u_{x_2x_1} & u_{x_2x_2} + 2\varepsilon & \cdots & u_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{x_nx_1} & u_{x_nx_2} & \cdots & u_{x_nx_n} + 2\varepsilon \end{pmatrix}(x)$$

注意到 $\operatorname{tr} D^2 u_\varepsilon = \Delta u(x) + 2n\varepsilon = 2n\varepsilon > 0$, 故 $D^2 u_\varepsilon$ 必有正特征值, 进而其不为负定阵, 从而 $u_\varepsilon(x)$ 在 $\overline{U}$ 内部取不到最大值, 亦即

$$\max_{\overline{U}} u_\varepsilon = \max_{\partial U} u_\varepsilon$$

由 U 有界知存在 $R > 0$ 使得 $U \subset B(0, R)$, 进而

$$\begin{aligned} \max_{\overline{U}} u &\le \max_{\overline{U}} u_\varepsilon = \max_{\partial U} u_\varepsilon \le \max_{\partial U} u + \max_{\partial U} \varepsilon |x|^2 \\ &\le \max_{\partial U} u + \max_{B(0,R)} \varepsilon |x|^2 = \max_{\partial U} u + \varepsilon R^2 \end{aligned}$$

根据 ε 的任意性知

$$\max_{\overline{U}} u \le \max_{\partial U} u$$

又显见

$$\max_{\partial U} u \le \max_{\overline{U}} u$$

故

$$\max_{\overline{U}} u = \max_{\partial U} u$$

命题即证. □

✍ **练习 7.5(下调和函数)**♣ 称 $v \in C^2(\overline{U})$ 是下调和的, 如果在 U 内有

$$-\Delta v \le 0$$

(a) 证明对下调和函数 v, 任取 $B(x, r) \subset U$ 都有

$$v(x) \le \fint_{B(x,r)} v \, dy$$

证明

固定 x, 设

$$\phi(r) = \alpha(n) r^n v(x) - \int_{B(x,r)} v \, dy$$

则希望证明 $\phi'(r) \le 0$, 进而由 $\phi(r) \le \phi(0)$ 即得命题. 知

$$\phi'(r) = n\alpha(n) r^{n-1} v(x) - \int_{\partial B(x,r)} v \, dS$$

要证明 $\phi'(r) \le 0$, 只需证明

$$v(x) \le ⨍_{\partial B(x,r)} v dS$$

另设

$$\varphi(r) = ⨍_{\partial B(x,r)} v dS$$

知

$$\begin{aligned}\varphi'(r) &= \frac{d}{dr}(⨍_{\partial B(0,1)} v(x+rz)dS(z)) = ⨍_{\partial B(0,1)} Dv(x+rz)\cdot z dS(z)\\ &= ⨍_{\partial B(x,r)} Dv(y)\cdot \frac{y-x}{r} dS(y) = \frac{1}{n\alpha(n)r^{n-1}}\int_{B(x,r)} \Delta v dy \ge 0\end{aligned}$$

这说明

$$\varphi(r) \ge \lim_{r\to 0^+} \varphi(r) = v(x)$$

也即

$$v(x) \le ⨍_{\partial B(x,r)} v dS$$

命题即证.

(b) 证明 $\max\limits_{\overline{U}} v = \max\limits_{\partial U} v$.

证明

用反证法, 若 v 不是常值函数, 且在 $\overline{U}$ 的某个内点 x_0 处达到最大值, 根据 (a) 的结论知

$$v(x_0) \le ⨍_{B(x_0,r)} v dy$$

等号当且仅当在 $B(x_0,r)$ 内 $v \equiv v(x_0)$ 时取得. 与此同时, 因为 $v(x_0) = \max\limits_{\overline{U}} v$, 显见

$$⨍_{B(x_0,r)} v dy \le v(x_0)$$

故 $v(x_0) = ⨍_{B(x_0,r)} v dy$, 也即在 $B(x_0,r)$ 内 $v \equiv v(x_0)$. 现在在 U 内任取一点 x, 知总能选取球 $\{B(x_k,r_k)\}_{k=1}^n$ 使得 $B(x_k,r_k) \subset \overline{U}(k=1,\cdots,n)$, $B(x_k,r_k) \cap B(x_{k+1},r_{k+1})$ 不是零测集, 且 $x, x_0 \in \bigcup\limits_{k=1}^n B(x_k,r_k)$. 根据前述论证知 $v(x)$ 在 $\bigcup\limits_{k=1}^n B(x_k,r_k)$ 内为常值函数, 进而 $v(x) = v(x_0)$, 这说明 v 在 U 内是常值函数, 与前述假设矛盾! 从而 v 要么是常值函数, 要么不在 $\overline{U}$ 的内点取最大值. 但根据 v 的连续性与 $\overline{U}$ 的有界闭性 (进而在 $\mathbb{R}^n$ 中是紧性), 知 v 必在 $\overline{U}$ 上存在最大值, 故其只能在 ∂U 处取得最大值. 这便是

$$\max_{\overline{U}} v = \max_{\partial U} v$$

命题即证.

另证

构造 $v_\varepsilon(x) = v(x) + \varepsilon|x|^2$, 知

$$D^2 v_\varepsilon(x) = \begin{pmatrix} v_{x_1x_1}+2\varepsilon & \cdots & v_{x_1x_n} \\ \vdots & & \vdots \\ v_{x_nx_1} & \cdots & v_{x_nx_n}+2\varepsilon \end{pmatrix}$$

知 $\operatorname{tr} D^2 v_\varepsilon = \Delta v(x) + 2n\varepsilon > 0$, 这说明 $D^2 v_\varepsilon$ 必有正特征值, 因而其不为负定阵, 从而 v_ε 在 $\overline{U}$ 内部取不到最大值, 亦即

$$\max_{\overline{U}} v_\varepsilon = \max_{\partial U} v_\varepsilon$$

现若 U 有界, 则存在 $R>0$ 使得 $U\subset B(0,R)$, 因而

$$\begin{aligned}\max_{\overline{U}} v &\le \max_{\overline{U}} v_\varepsilon \le \max_{\partial U} v_\varepsilon = \max_{\partial U}(v(x)+\varepsilon|x|^2)\\ &\le \max_{\partial U} v + \max_{\partial U}\varepsilon|x|^2 \le \max_{\partial U} v + \varepsilon R^2\end{aligned}$$

根据 $R<+\infty$ 与 ε 的任意性即得

$$\max_{\overline{U}} v \le \max_{\partial U} v$$

但显见 $\max_{\overline{U}} v \ge \max_{\partial U} v$, 故 $\max_{\overline{U}} v = \max_{\partial U} v$, 命题得证.

(c) 设 $\phi:\mathbb{R}\to\mathbb{R}$ 是光滑的凸函数. 若 u 是调和函数, $v:=\phi(u)$, 证明 v 是下调和函数.

证明

既然 ϕ 是光滑凸函数, 知 $\phi''\ge 0$. 根据 u 是调和函数知 $u\in C^2(\overline{U})$, 进而知 $v\in C^2(\overline{U})$. 对 v 求导有:

$$Dv=\phi'(u)\cdot Du$$

进一步求导并结合 $\Delta u=0$ 知

$$\Delta v=\phi'(u)\cdot\Delta u+\phi''(u)\cdot|Du|^2=\phi''(u)\cdot|Du|^2\ge 0$$

此即 $-\Delta v\le 0$, 也即 v 是下调和函数, 命题得证.

(d) 当 u 是调和函数时, 证明 $v:=|Du|^2$ 是下调和函数.

证明

知

$$v=\sum_{i=1}^n (u_{x_i})^2 \Rightarrow Dv=(2\sum_{i=1}^n u_{x_i}u_{x_ix_1},2\sum_{i=1}^n u_{x_i}u_{x_ix_2},\cdots,2\sum_{i=1}^n u_{x_i}u_{x_ix_n})$$

进而

$$\Delta v=2\sum_{j=1}^n(\sum_{i=1}^n u_{x_i}u_{x_ix_jx_j}+\sum_{i=1}^n u_{x_ix_j}^2)$$

显见 $\sum_{j=1}^n\sum_{i=1}^n u_{x_ix_j}^2\ge 0$, 而

$$\sum_{j=1}^n\sum_{i=1}^n u_{x_i}u_{x_ix_jx_j}=\sum_{i=1}^n u_{x_i}\frac{\partial}{\partial x_i}(\Delta u)=0$$

故 $\Delta v\ge 0$, 也即 v 是下调和函数, 命题得证. □

练习 7.6♣ 设 U 是 $\mathbb{R}^n$ 上的有界开集, 证明存在仅依赖于 U 的常数 C, 使得

$$\max_{\overline{U}}|u|\le C(\max_{\partial U}|g|+\max_{\overline{U}}|f|)$$

其中 u 是光滑函数, 且满足

$$\begin{cases}-\Delta u=f, & \text{在 } U \text{ 内}\\ u=g, & \text{在 } \partial U \text{ 上}\end{cases}$$

证明

设 $v(x)=u(x)+\frac{|x|^2}{2n}\lambda\in C^2(\overline{U}),\lambda=\max_{\overline{U}}|f|$, 知当 $x\in U$ 时:

$$\Delta v(x)=\Delta u(x)+\lambda=-f(x)+\max_{\overline{U}}|f|\ge 0$$

这说明 v 在 U 内是下调和函数, 因而由上题 (b) 结论知:

$$\max_{\overline{U}} v(x)=\max_{\partial U} v(x)$$

亦即

$$\begin{aligned}
\max_{\overline{U}}(u(x)+\frac{|x|^2}{2n}\lambda) &= \max_{\partial U}(u(x)+\frac{|x|^2}{2n}\lambda) \\
&\le \max_{\partial U} u(x) + \max_{\partial U}\frac{|x|^2}{2n}\lambda \\
&= \max_{\partial U} g(x) + \frac{C_U}{2n}\max_{\overline{U}}|f(x)| \\
&\le \max\{1,\frac{C_U}{2n}\}(\max_{\partial U} g(x) + \max_{\overline{U}}|f(x)|)
\end{aligned}$$

其中 $C_U = \max_{\partial U}|x|^2$ 是只关于 U 的常数, 且由 U 有界知 $C_U < \infty$, 又因为显见

$$\max_{\overline{U}} u(x) \le \max_{\overline{U}}(u(x)+\frac{|x|^2}{2n}\lambda)$$

故存在只与 U 有关的常数 C 使得

$$\max_{\overline{U}} u(x) \le C(\max_{\partial U}|g(x)| + \max_{\overline{U}}|f(x)|)$$

在同样的记号下, 另记 $w(x) = -u(x) + \frac{|x|^2}{2n}\lambda$, 知依旧有 $-\Delta w(x) \le 0$, 这说明上述过程对 $-u(x)$ 也适用, 不妨设 C 足够大, 得到

$$\max_{\overline{U}}(-u(x)) \le C(\max_{\partial U}|g(x)| + \max_{\overline{U}}|f(x)|)$$

故

$$\max_{\overline{U}}|u(x)| \le C(\max_{\partial U}|g(x)| + \max_{\overline{U}}|f(x)|)$$

此即欲证. □

练习 7.7♣ 用球上的 Poisson 公式证明:

$$r^{n-2}\frac{r-|x|}{(r+|x|)^{n-1}}u(0) \le u(x) \le r^{n-2}\frac{r+|x|}{(r-|x|)^{n-1}}u(0)$$

其中 u 是 $B^\circ(0,r)$ 内的正调和函数. 这是 Harnack 不等式的一个精确形式.

证明

根据 Poisson 公式知

$$u(x) = \frac{r^2-|x|^2}{n\alpha(n)r}\int_{\partial B(0,r)}\frac{g(y)}{|x-y|^n}dS(y), \quad x\in B^\circ(0,r)$$

其中在 $\partial B(0,r)$ 上 $u=g$. 对左式此即证明

$$\begin{aligned}
&\frac{r^2-|x|^2}{n\alpha(n)r}\int_{\partial B(0,r)}\frac{g(y)}{|x-y|^n}dS(y) \ge r^{n-2}\frac{r-|x|}{(r+|x|)^{n-1}}u(0) \\
\Leftarrow\ & ⨍_{\partial B(0,r)}\frac{g(y)}{|x-y|^n}dS(y) \ge \frac{u(0)}{(r+|x|)^n}
\end{aligned}$$

既然 u 是正调和函数, 由最大值原理显见 $g(y)\ge u(0)>0$, 故只需证明

$$⨍_{\partial B(0,r)}\frac{dS(y)}{|x-y|^n} \ge \frac{1}{(r+|x|)^n}$$

进一步只需证明

$$|x-y|^n \le (r+|x|)^n, \quad x\in B^\circ(0,r)$$

而显见 $|x-y|\le|x|+|y| = r+|x|$ 成立, 左式即证.

对右式此即证明

$$\frac{r^2-|x|^2}{n\alpha(n)r}\int_{\partial B(0,r)}\frac{g(y)}{|x-y|^n}dS(y)\le r^{n-2}\frac{r+|x|}{(r-|x|)^{n-1}}u(0)$$
$$\Leftarrow\quad ⨍_{\partial B(0,r)}\frac{g(y)}{|x-y|^n}dS(y)\le\frac{u(0)}{(r-|x|)^n}$$

首先说明 $|x-y|^n\ge(r-|x|)^n$, 这是因为 $|x-y|\ge|y|-|x|=r-|x|$. 故只需证明

$$⨍_{\partial B(0,r)}g(y)dS(y)\le u(0)$$

因为 u 是调和函数, 故均值定理表明:

$$u(0)=⨍_{\partial B(0,r)}udS=⨍_{\partial B(0,r)}gdS$$

右式即证.

综上, 命题得证. □

练习 7.8♣ 证明球的 Poisson 公式(7.2.15).

证明

对初值问题

$$\begin{cases}\Delta u=0, & \text{在 } B^\circ(0,r) \text{ 内}\\ u=g, & \text{在 } \partial B(0,1) \text{ 上}\end{cases}$$

首先证明 $\int_{\partial B(0,r)}K(x,y)dS(y)=1$, 其中

$$K(x,y):=\frac{r^2-|x|^2}{n\alpha(n)r}\frac{1}{|x-y|^n},\quad x\in B^\circ(0,r),y\in\partial B(0,r)$$

注意从 $K(x,y)$ 的构造可知

$$K(x,y)=-\frac{\partial G}{\partial\boldsymbol{\nu}}(x,y)$$

其中 $\frac{\partial}{\partial\nu}$ 表示关于变量 y 沿 $\partial B(0,r)$ 外法方向的方向导数运算. 进而由 Green 公式:

$$\begin{aligned}\int_{\partial B(0,r)}K(x,y)dS(y)&=-\int_{\partial B(0,r)}\frac{\partial G}{\partial\boldsymbol{\nu}}(x,y)dS(y)=\int_{B^\circ(0,r)}(-\Delta_yG(x,y))dy\\&=\int_{B^\circ(0,r)}\delta_0(y-x)dy=1\end{aligned}$$

最后一步成立是因为 $x\in B^\circ(0,r)$.

显见对任意的 $x\in B^\circ(0,r)$, $K(x,y)$ 关于 x 都是 C^∞ 的, 故 $u\in C^\infty(B^\circ(0,r))$. 进一步, 因为

$$\Delta_x\int_{\partial B(0,r)}K(x,y)dS(y)=0=\int_{\partial B(0,r)}\Delta_xK(x,y)dS(y)$$

故

$$\Delta u=\int_{\partial B(0,r)}\Delta_xK(x,y)g(y)dS(y)=0$$

最后证明对 $x^0\in\partial B(0,r)$ 有 $\lim\limits_{\substack{x\to x^0\\x\in B^\circ(0,r)}}u(x)=g(x^0)$. 根据 g 的连续性知

$$\forall\varepsilon>0\exists\delta>0\forall y\in\partial B(0,r)(|y-x^0|<\delta\Rightarrow|g(y)-g(x^0)|<\varepsilon)$$

故

$$\begin{aligned}|u(x)-g(x^0)|&=|\int_{\partial B(0,r)}K(x,y)[g(y)-g(x^0)]dS(y)|\\&\le\int_{\partial B(0,r)\cap B(x^0,\delta)}K(x,y)|g(y)-g(x^0)|dS(y)+\int_{\partial B(0,r)\backslash B(x^0,\delta)}K(x,y)|g(y)-g(x^0)|dS(y)\\&=:I+J\end{aligned}$$

对 I 知

$$I \leq \varepsilon \int_{\partial B(0,r)} K(x,y)dS(y) = \varepsilon$$

对 J, 设 $|x-x^0| < \frac{\delta}{2}$, 则当 $|y-x^0| > \delta$, 有

$$|y-x^0| \leq |y-x| + |x-x^0| < |y-x| + \frac{\delta}{2} < |y-x| + \frac{1}{2}|y-x^0| \Rightarrow |y-x^0| < 2|y-x|$$

得到

$$\begin{aligned} J &\leq 2\|g\|_{L^\infty} \int_{\partial B(0,r)\setminus B(x^0,\delta)} \frac{r^2-|x|^2}{n\alpha(n)r} \frac{1}{|x-y|^n} dS(y) \\ &< 2\|g\|_{L^\infty} \cdot \frac{2^n}{n\alpha(n)r} \int_{\partial B(0,r)\setminus B(x^0,\delta)} \frac{r^2-|x|^2}{|y-x_0|^n} dS(y) \to 0, \quad |x| \to r \end{aligned}$$

故

$$\lim_{\substack{x\to x^0 \\ x\in B^\circ(0,r)}} u(x) = g(x^0)$$

命题得证. □

练习 7.9♣ 设 u 是初值问题

$$\begin{cases} \Delta u = 0, & \text{在 } \mathbb{R}^n_+ \text{ 内} \\ u = g, & \text{在 } \partial\mathbb{R}^n_+ \text{ 上} \end{cases}$$

由半平面 Poisson 公式给出的解. 设 g 有界, 且当 $x \in \partial\mathbb{R}^n_+, |x| \leq 1$ 时有 $g(x) = |x|$. 证明 Du 在 $x = 0$ 附近是无界的.

证明

由 Poisson 公式知

$$u(x) = \frac{2x_n}{n\alpha(n)} \int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|x-y|^n} dy, \quad x \in \mathbb{R}^n_+$$

得到

$$u(0) = 0, \quad u(\lambda e_n) = \frac{2\lambda}{n\alpha(n)} \int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|\lambda e_n - y|^n} dy$$

其中 $e_n = (0,\cdots,0,1)$. 现设 $\exists M > 0 \forall x \in \partial\mathbb{R}^n_+ (|g(x)| \leq M)$, 并记 $D_1 := \{x \in \partial\mathbb{R}^n_+ : |x| \leq 1\}, D_2 := \{x \in \partial\mathbb{R}^n_+ : |x| > 1\}$, 考察:

$$\begin{aligned} |\frac{u(\lambda e_n) - u(0)}{\lambda}| &= \frac{2}{n\alpha(n)} |\int_{\partial\mathbb{R}^n_+} \frac{g(y)}{|\lambda e_n - y|^n} dy| \\ &= \frac{2}{n\alpha(n)} |\int_{D_1} \frac{g(y)}{|\lambda e_n - y|^n} dy + \int_{D_2} \frac{g(y)}{|\lambda e_n - y|^n} dy| \\ &\geq \frac{2}{n\alpha(n)} (|\int_{D_1} \frac{g(y)}{|\lambda e_n - y|^n} dy| - |\int_{D_2} \frac{g(y)}{|\lambda e_n - y|^n} dy|) \\ &\geq \frac{2}{n\alpha(n)} (\int_{D_1} \frac{|y|}{|\lambda e_n - y|^n} dy - \int_{D_2} \frac{M}{|\lambda e_n - y|^n} dy) \\ &\geq \frac{2}{n\alpha(n)} (\int_{D_1} \frac{|y|}{(\lambda + |y|)^n} dy - \int_{D_2} \frac{M}{|y|^n} dy) \\ &= \frac{2}{n\alpha(n)} (\int_0^1 (\int_{\partial B(0,r)} \frac{|y|}{(\lambda+|y|)^n} d\sigma(y)) dr - \int_1^\infty (\int_{\partial B(0,r)} \frac{M}{|y|^n} d\sigma(y)) dr) \\ &= \frac{2(n-1)\alpha(n-1)}{n\alpha(n)} (\int_0^1 \frac{r^{n-1}dr}{(\lambda+r)^n} - \int_1^\infty \frac{M}{r^2} dr) \\ &\geq \frac{2(n-1)\alpha(n-1)}{n\alpha(n)} \int_0^1 \frac{r^{n-1}dr}{(\lambda+r)^n} + C \\ &= \frac{2(n-1)\alpha(n-1)}{n\alpha(n)} \int_0^1 \frac{dr}{r(\frac{\lambda}{r}+1)^n} + C \end{aligned}$$

其中 C 是常数. 现在估计 $\int_0^1 \frac{dr}{r(\frac{\lambda}{r}+1)^n}$:

$$\int_0^1 \frac{dr}{r(\frac{\lambda}{r}+1)^n} = \sum_{k=1}^{\infty}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{dr}{r(\frac{\lambda}{r}+1)^n} \geq \sum_{k=1}^{\infty}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{k\,dr}{(1+\lambda(k+1))^n}$$
$$= \sum_{k=1}^{\infty}\frac{1}{k(k+1)}\frac{k}{(1+\lambda(k+1))^n} = \sum_{k=2}^{\infty}\frac{1}{k(1+\lambda k)^n}$$

现任取 $K>0$, 取 $\lambda=\frac{1}{K}$, 则

$$\sum_{k=2}^{\infty}\frac{1}{k(1+\lambda k)^n} \geq \sum_{k=2}^{K}\frac{1}{k(1+\lambda k)^n} \geq \sum_{k=2}^{K}\frac{1}{k2^n}$$

令 $K\to\infty$ 即知 $\sum\limits_{k=2}^{\infty}\frac{1}{k(1+\lambda k)^n}=+\infty$, 因而 $\int_0^1 \frac{dr}{r(\frac{\lambda}{r}+1)^n}$ 随着 $\lambda\to 0^+$ 而发散, 又因为 $|Du|\geq|\lim\limits_{\lambda\to 0}\frac{u(\lambda e_n)-u(0)}{\lambda}|$, 故 Du 在 $x=0$ 附近无界. □

练习 7.10(反射原理)♣

(a) 设 U^+ 是开半球 $\{x\in\mathbb{R}^n : |x|<1, x_n>0\}$. 设 $u\in C^2(\overline{U^+})$ 在 U^+ 中调和, 且在 $\partial U^+\cap\{x_n=0\}$ 上 $u=0$. 记

$$v(x) := \begin{cases} u(x), & x_n\geq 0 \\ -u(x_1,\cdots,x_{n-1},-x_n), & x_n<0 \end{cases}$$

其中 $x\in U=B^\circ(0,1)$. 证明 $v\in C^2(U)$, 且 v 在 U 内调和.

证明

根据 v 本身的构造与 u 的正则性, 只需证明 v 在 $\{x_n=0\}$ 上是 C^2 的即可, 进一步只需说明 v 的全体与 x_n 有关的二阶偏导数连续即可. 显见至少有 $v\in C(\{x_n=0\})$. 取 $x^0=(x_1,\cdots,x_n,0)\in\{x_n=0\}$, 根据 v 的导数定义知:

$$\begin{aligned}
&\frac{1}{-\lambda}\cdot(v(x_1,\cdots,x_{n-1},-\lambda)-v(x_1,\cdots,x_{n-1},0))\\
=&\frac{1}{\lambda}\cdot u(x_1,\cdots,x_{n-1},\lambda)=\frac{1}{\lambda}\cdot(u(x_1,\cdots,x_{n-1},\lambda)-u(x_1,\cdots,x_{n-1},0))\\
=&\frac{1}{\lambda}\cdot(v(x_1,\cdots,x_{n-1},\lambda)-v(x_1,\cdots,x_{n-1},0))
\end{aligned}$$

现在已知 $\lim\limits_{\lambda\to 0^+}(u(x_1,\cdots,x_{n-1},\lambda)-u(x_1,\cdots,x_{n-1},0))$ 存在, 故

$$\begin{aligned}
&\lim_{\lambda\to 0^+}\frac{1}{\lambda}\cdot(v(x_1,\cdots,x_{n-1},0)-v(x_1,\cdots,x_{n-1},-\lambda))\\
=&\lim_{\lambda\to 0^+}\frac{1}{\lambda}\cdot(v(x_1,\cdots,x_{n-1},\lambda)-v(x_1,\cdots,x_{n-1},0))=\frac{\partial v}{\partial x_n}(x_1,\cdots,x_{n-1},0)
\end{aligned}$$

同时在 $\{x_n=0\}$ 上有 $v_{x_n}(x_1,\cdots,x_{n-1},0)=u_{x_n}(x_1,\cdots,x_{n-1},0)$, 进而根据 u_{x_n} 的连续性可知 $v\in C^1(U)$. 显见

$$v_{x_n}(x) = \begin{cases} u_{x_n}(x), & x_n\geq 0 \\ u_{x_n}(x_1,\cdots,x_{n-1},-x_n), & x_n<0 \end{cases}$$

而

$$\begin{aligned}
&\frac{1}{-\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},-\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))\\
=&\frac{1}{-\lambda}\cdot(u_{x_n}(x_1,\cdots,x_{n-1},\lambda)-u_{x_n}(x_1,\cdots,x_{n-1},0))\\
=&-\frac{1}{\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))
\end{aligned}$$

已知 $\lim\limits_{\lambda\to 0^+}\frac{1}{-\lambda}\cdot(u_{x_n}(x_1,\cdots,x_{n-1},\lambda)-u_{x_n}(x_1,\cdots,x_{n-1},0))$ 存在, 故 $\lim\limits_{\lambda\to 0^+}\frac{1}{-\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},-\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))$ 与 $\lim\limits_{\lambda\to 0^+}\frac{1}{-\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))$ 均存在. 但因为导函数本身不能出现第一类间断点, 故只能有

$$\lim_{\lambda\to 0^+}\frac{1}{-\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},-\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))=\lim_{\lambda\to 0^+}\frac{1}{-\lambda}\cdot(v_{x_n}(x_1,\cdots,x_{n-1},\lambda)-v_{x_n}(x_1,\cdots,x_{n-1},0))=0$$

故在 $\{x_n=0\}$ 上 $v_{x_nx_n}(x_1,\cdots,x_{n-1},0)=0$ 存在, 且显见 $v_{x_nx_n}$ 在 $|x_n|$ 足够小时有定义, 故 $v_{x_nx_n}$ 连续. $v_{x_ix_n}$ 的连续性可由 $u_{x_ix_n}$ 的连续性推知, 故 $v\in C^2(U)$.

要证明 v 是 U 内的调和函数, 根据 u 的调和性知只需证明 v 在 $\{x_n=0\}$ 上是调和函数. 考虑验证均值公式:

$$\begin{aligned}\fint_{B(x_1,\cdots,x_{n-1},0),r} v(y)dy &= \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_+} v(y)dy + \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_-} v(y)dy \\ &= \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_+} u(y)dy - \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_-} u(y_1,\cdots,y_{n-1},-y_n)dy \\ &= \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_+} u(y)dy - \fint_{B((x_1,\cdots,x_{n-1},0),r)\cap\mathbb{R}^n_+} u(y)dy = 0 = v(x_1,\cdots,x_{n-1},0)\end{aligned}$$

故均值公式成立, 再由 v 在 $\{x_n=0\}$ 上的连续性即知 v 是 U 内的调和函数.

(b) 现在设 $u\in C^2(U^+)\cap C(\overline{U^+})$, 证明 v 在 U 内调和.

证明

考虑 U 内的边值问题:

$$\begin{cases}\Delta w=0, & \text{在 } U \text{ 内}\\ w=v, & \text{在 } \partial U \text{ 上}\end{cases}$$

依据球上的 Poisson 公式可得

$$w(x)=\frac{1-|x|^2}{n\alpha(n)}\int_{\partial B(0,1)}\frac{v(y)}{|x-y|^n}dS(y),\quad x\in U$$

已知 w 是调和函数, 现在说明 $w(x)=v(x)(\forall x\in\{x_n=0\})$, 进一步只需说明 $w(x)=0(\forall x\in\{x_n=0\})$. 知:

$$\begin{aligned}&\frac{1-|x|^2}{n\alpha(n)}\int_{\partial B(0,1)}\frac{v(y)}{|x-y|^n}dS(y)\\ =&\frac{1-\sum\limits_{i=1}^{n-1}x_i^2}{n\alpha(n)}\int_{\partial B(0,1)}\frac{v(y)}{(\sum\limits_{i=1}^{n-1}(x_i-y_i)^2+y_n^2)^{\frac{n}{2}}}dS(y)\\ =&\frac{1-\sum\limits_{i=1}^{n-1}x_i^2}{n\alpha(n)}\int_{\partial B(0,1)\cap\mathbb{R}^n_+}\frac{u(y_1,\cdots,y_n)}{(\sum\limits_{i=1}^{n-1}(x_i-y_i)^2+y_n^2)^{\frac{n}{2}}}dS(y)-\frac{1-\sum\limits_{i=1}^{n-1}x_i^2}{n\alpha(n)}\int_{\partial B(0,1)\cap\mathbb{R}^n_-}\frac{u(y_1,\cdots,y_{n-1},-y_n)}{(\sum\limits_{i=1}^{n-1}(x_i-y_i)^2+y_n^2)^{\frac{n}{2}}}dS(y)\\ =&\frac{1-\sum\limits_{i=1}^{n-1}x_i^2}{n\alpha(n)}\int_{\partial B(0,1)\cap\mathbb{R}^n_+}\frac{u(y_1,\cdots,y_n)}{(\sum\limits_{i=1}^{n-1}(x_i-y_i)^2+y_n^2)^{\frac{n}{2}}}dS(y)-\frac{1-\sum\limits_{i=1}^{n-1}x_i^2}{n\alpha(n)}\int_{\partial B(0,1)\cap\mathbb{R}^n_+}\frac{u(y_1,\cdots,y_n)}{(\sum\limits_{i=1}^{n-1}(x_i-y_i)^2+y_n^2)^{\frac{n}{2}}}dS(y)=0\end{aligned}$$

因而 $w(x)=0(\forall x\in\{x_n=0\})$. 因为 $w=v=u$ 在 ∂U^+ 上也成立, 故根据调和函数的强极值原理知 $w=u$ 在 U^+ 内成立, 因而 $w=v$ 在 U^+ 内成立, 同理可证 $w=v$ 在 U^- 内也成立, 因而 $w=v$ 在 U 内均 (点态) 成立, 故 v 是调和函数, 命题证毕. □

练习 7.11(Laplace 方程的 Kelvin 变换)♣ 函数 $u:\mathbb{R}^n\to\mathbb{R}$ 的 Kelvin 变换 $\mathcal{K}u=\overline{u}$ 定义为:

$$\overline{u}(x):=u(\overline{x})|\overline{x}|^{n-2}=u(\frac{x}{|x|^2})|x|^{2-n},\quad x\neq 0$$

其中 $\overline{x}=\frac{x}{|x|^2}$. 证明如果 u 是调和函数, 那么 $\overline{u}$ 也是调和函数.

证明

首先约定符号, 设 $f(x) = (f_1, \cdots, f_n)^T(x)$:

$$\overline{x} = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix}, \nabla_x = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}, \nabla_x \boldsymbol{f} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} \\ \vdots \\ \frac{\partial f_n}{\partial x_n} \end{pmatrix}, D_x f = \begin{pmatrix} \frac{\partial}{\partial x_1}(f_1) & \cdots & \frac{\partial}{\partial x_n}(f_1) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1}(f_n) & \cdots & \frac{\partial}{\partial x_n}(f_n) \end{pmatrix}, \Delta_x = \nabla_x^T \nabla_x$$

为表示简洁, 对 $x \in \mathbb{R}^n$ 记 $h(x) = |x|^{2-n}$, 则 $\overline{u}(x) = u(\overline{x})h(x)$. 依照提示, 首先计算 $D_x\overline{x}$. 知 $i \neq j$ 时:

$$\frac{\partial}{\partial x_j}(\overline{x}_i) = \frac{-2x_j \cdot x_i}{|x|^4}$$

而 $i = j$ 时:

$$\frac{\partial}{\partial x_j}(\overline{x}_i) = \frac{|x|^2 - 2x_i \cdot x_i}{|x|^4}$$

综上可得:

$$\frac{\partial}{\partial x_j}(\overline{x}_i) = \frac{\delta_{ij}}{|x|^2} - \frac{2x_i x_j}{|x|^4}$$

其中 δ_{ij} 是 Kronecker 符号. 故

$$\begin{aligned} D_x\overline{x} = (\frac{\partial}{\partial x_j}(\overline{x}_i))_{n\times n} &= \begin{pmatrix} \frac{1}{|x|^2} - \frac{2x_1^2}{|x|^4} & \cdots & -\frac{2x_1x_n}{|x|^4} \\ \vdots & & \vdots \\ -\frac{2x_nx_1}{|x|^4} & \cdots & \frac{1}{|x|^2} - \frac{2x_n^2}{|x|^4} \end{pmatrix} \\ &= \frac{1}{|x|^2}I - \frac{2}{|x|^4}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}(x_1, \cdots, x_n) = |x|^{-2}I - 2|x|^{-4}xx^T \end{aligned}$$

对 $\Delta_x\overline{u}(x)$ 知:

$$\begin{aligned} \Delta_x\overline{u}(x) = 0 &\Leftrightarrow \nabla_x^T\nabla_x(u(\overline{x})h(x)) = 0 \\ &\Leftrightarrow \nabla_x^T(h(x)\nabla_x[u(\overline{x})] + u(\overline{x})(\nabla_x h(x))) = 0 \\ &\Leftrightarrow \nabla_x^T h(x)\nabla_x[u(\overline{x})] + h(x)\Delta_x[u(\overline{x})] + \nabla_x^T[u(\overline{x})](\nabla_x h(x)) + u(\overline{x})(\Delta_x h(x)) = 0 \\ &\Leftrightarrow h(x)\Delta_x[u(\overline{x})] + u(\overline{x})(\Delta_x h(x)) + 2\nabla_x^T h(x)\nabla_x[u(\overline{x})] = 0 \end{aligned}$$

其中

$$\nabla_x^T h(x)\nabla_x[u(\overline{x})] = (\frac{\partial h}{\partial x_1}(x), \cdots, \frac{\partial h}{\partial x_n}(x))\begin{pmatrix} \frac{\partial}{\partial x_1}[u(\overline{x})] \\ \vdots \\ \frac{\partial}{\partial x_n}[u(\overline{x})] \end{pmatrix}$$

下面分别计算这三项.

(i) 对于 $u(\overline{x})(\Delta_x h(x))$, 知:

$$\begin{aligned} \Delta_x h(x) = \Delta_x(|x|^{2-n}) &= \nabla_x^T((2-n)|x|^{-n}x) \\ &= (2-n)(\nabla_x^T(|x|^{-n})x + |x|^{-n}\nabla_x^T(x)) \\ &= (2-n)(-n|x|^{-n-2}x^Tx + n|x|^{-n}) = 0 \end{aligned}$$

这说明 $u(\overline{x})(\Delta_x h(x)) = 0$.

(ii) 对于 $2\nabla_x^T h(x)\nabla_x[u(\overline{x})]$, 知:

$$\nabla_x[u(\overline{x})] = (D_x\overline{x})(\nabla_x u(\overline{x})) = (|x|^{-2}I - 2|x|^{-4}xx^T)(\nabla_x u(\overline{x}))$$

$$\nabla_x^T h(x) = \nabla_x^T(|x|^{2-n}) = (2-n)|x|^{-n}x^T$$

因而

$$
\begin{aligned}
2\nabla_x^T h(x)\nabla_x[u(\overline{x})] &= 2(2-n)|x|^{-n}x^T(|x|^{-2}I - 2|x|^{-4}xx^T)(\nabla_x u(\overline{x})) \\
&= 2(2-n)|x|^{-n}(|x|^{-2}x^T - 2|x|^{-4}x^Txx^T)(\nabla_x u(\overline{x})) \\
&= 2(n-2)|x|^{-n-2}x^T\nabla_x u(\overline{x})
\end{aligned}
$$

(iii) 对于 $h(x)\Delta_x[u(\overline{x})]$, 展开计算 $\Delta_x[u(\overline{x})]$:

$$
\nabla_x u(\overline{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1}(u(\frac{x}{|x|^2})) \\ \vdots \\ \frac{\partial}{\partial x_n}(u(\frac{x}{|x|^2})) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n u_i(\frac{x}{|x|^2})\frac{\partial}{\partial x_1}(\frac{x_i}{|x|^2}) \\ \vdots \\ \sum_{i=1}^n u_i(\frac{x}{|x|^2})\frac{\partial}{\partial x_n}(\frac{x_i}{|x|^2}) \end{pmatrix}
$$

进一步

$$
\begin{aligned}
\nabla_x^T(\nabla_x u(\overline{x})) &= \sum_{j=1}^n \frac{\partial}{\partial x_j}(\sum_{i=1}^n u_i(\frac{x_i}{|x|^2})\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2})) \\
&= \sum_{j=1}^n\sum_{i=1}^n \frac{\partial}{\partial x_j}(u_i(\frac{x}{|x|^2})\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2})) \\
&= \sum_{j=1}^n\sum_{i=1}^n(\sum_{k=1}^n u_{ik}(\frac{x}{|x|^2})(\frac{\partial}{\partial x_j}(\frac{x_k}{|x|^2}))(\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2})) + u_i(\frac{x}{|x|^2})\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2}))
\end{aligned}
$$

其中:

$$
\begin{aligned}
\frac{\partial}{\partial x_j}(\frac{x_k}{|x|^2})(\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2})) &= |x|^{-4}(\delta_{kj} - 2x_kx_j|x|^{-2})(\delta_{ij} - 2x_ix_j|x|^{-2}) \\
&= |x|^{-4}(\delta_{kj}\delta_{ij} - 2x_kx_j\delta_{ij}|x|^{-2} - 2x_ix_j|x|^{-2}\delta_{kj} + 4x_kx_ix_j^2|x|^{-4})
\end{aligned}
$$

对 k 求和:

$$
\begin{aligned}
&\sum_{k=1}^n u_{ik}(\frac{x}{|x|^2})(\frac{\partial}{\partial x_j}(\frac{x_k}{|x|^2}))(\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2})) \\
=&|x|^{-4}\sum_{k=1}^n u_{ik}(\frac{x}{|x|^2})(\delta_{kj}\delta_{ij} - 2x_kx_j\delta_{ij}|x|^{-2} - 2x_ix_j|x|^{-2}\delta_{kj} + 4x_kx_ix_j^2|x|^{-4}) \\
=&|x|^{-4}(u_{ij}(\frac{x}{|x|^2})\delta_{ij} - 2\sum_{k=1}^n u_{ik}(\frac{x}{|x|^2})x_kx_j\delta_{ij}|x|^{-2} - 2u_{ij}(\frac{x}{|x|^2})x_ix_j|x|^{-2} + 4\sum_{k=1}^n u_{ik}(\frac{x}{|x|^2})x_kx_ix_j^2|x|^{-4})
\end{aligned}
$$

另外计算知 $i \neq j$ 时:

$$
\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2}) = -2x_i|x|^{-4} + 8x_ix_j^2|x|^{-6}
$$

$i = j$ 时:

$$
\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2}) = -6x_i|x|^{-4} + 8x_i^3|x|^{-6}
$$

综上有

$$
\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2}) = -4x_i|x|^{-4}\delta_{ij} - 2x_i|x|^{-4} + 8x_ix_j^2|x|^{-6}
$$

现在对 i 求和:

$$
\begin{aligned}
&\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})(\frac{\partial}{\partial x_j}(\frac{x_k}{|x|^2}))(\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2}))\\
=&|x|^{-4}(u_{jj}(\frac{x}{|x|^2})-2\sum_{k=1}^{n}u_{jk}(\frac{x}{|x|^2})x_kx_j|x|^{-2}-2\sum_{i=1}^{n}u_{ij}(\frac{x}{|x|^2})x_ix_j|x|^{-2}+4\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})x_kx_ix_j^2|x|^{-4})\\
=&|x|^{-4}(u_{jj}(\frac{x}{|x|^2})-4\sum_{i=1}^{n}u_{ij}(\frac{x}{|x|^2})x_ix_j|x|^{-2}+4\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})x_kx_ix_j^2|x|^{-4})
\end{aligned}
$$

而

$$
\begin{aligned}
&\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2})\\
=&\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})(-4x_i|x|^{-4}\delta_{ij}-2x_i|x|^{-4}+8x_ix_j^2|x|^{-6})\\
=&-4u_j(\frac{x}{|x|^2})x_j|x|^{-4}-2|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i+8|x|^{-6}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_ix_j^2
\end{aligned}
$$

最后对 j 求和:

$$
\begin{aligned}
&\sum_{j=1}^{n}\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})(\frac{\partial}{\partial x_j}(\frac{x_k}{|x|^2}))(\frac{\partial}{\partial x_j}(\frac{x_i}{|x|^2}))\\
=&|x|^{-4}(\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})-4\sum_{j=1}^{n}\sum_{i=1}^{n}u_{ij}(\frac{x}{|x|^2})x_ix_j|x|^{-2}+4\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})x_kx_i(\sum_{j=1}^{n}x_j^2)|x|^{-4})\\
=&|x|^{-4}(\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})-4\sum_{j=1}^{n}\sum_{i=1}^{n}u_{ij}(\frac{x}{|x|^2})x_ix_j|x|^{-2}+4\sum_{i=1}^{n}\sum_{k=1}^{n}u_{ik}(\frac{x}{|x|^2})x_kx_i|x|^{-2})\\
=&|x|^{-4}\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})
\end{aligned}
$$

而

$$
\begin{aligned}
&\sum_{j=1}^{n}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})\frac{\partial^2}{\partial x_j^2}(\frac{x_i}{|x|^2})\\
=&-4\sum_{j=1}^{n}u_j(\frac{x}{|x|^2})x_j|x|^{-4}-2n|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i+8|x|^{-6}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i(\sum_{j=1}^{n}x_j^2)\\
=&-4\sum_{j=1}^{n}u_j(\frac{x}{|x|^2})x_j|x|^{-4}-2n|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i+8|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i\\
=&-2(n+2)\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i|x|^{-4}+8|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i\\
=&2(2-n)|x|^{-4}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i
\end{aligned}
$$

综上得到:

$$
h(x)\Delta_x[u(\overline{x})]=|x|^{-2-n}\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})+2(2-n)|x|^{-2-n}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i
$$

最后, 综合 (i)-(iii), 可知

$$
\begin{aligned}
&h(x)\Delta_x[u(\overline{x})]+u(\overline{x})(\Delta_x h(x))+2\nabla_x^T h(x)\nabla_x[u(\overline{x})]\\
=&0+2(n-2)|x|^{-n-2}x^T\nabla_x u(\overline{x})+|x|^{-2-n}\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})+2(2-n)|x|^{-2-n}\sum_{i=1}^{n}u_i(\frac{x}{|x|^2})x_i\\
=&|x|^{-2-n}\sum_{j=1}^{n}u_{jj}(\frac{x}{|x|^2})
\end{aligned}
$$

因为 u 是调和函数, 故 $\sum_{j=1}^{n}u_{jj}(y)=0$ 对任意 $y\in\mathbb{R}^n$ 均成立, 因而

$$
h(x)\Delta_x[u(\overline{x})]+u(\overline{x})(\Delta_x h(x))+2\nabla_x^T h(x)\nabla_x[u(\overline{x})]=0
$$

故 $\Delta_x\overline{u}(x)=0$, $\overline{u}$ 是调和函数. □

练习 7.12♣ 设 u 是光滑的, 且是 $u_t-\Delta u=0$ 在 $\mathbb{R}^n\times(0,\infty)$ 上的解.

(a) 证明对任意的 $\lambda\in\mathbb{R}$, $u_\lambda(x,t):=u(\lambda x,\lambda^2 t)$ 也是热方程的解.

证明

知

$$
\begin{aligned}
(u_\lambda)_t(\lambda x,\lambda^2 t)&=\lambda^2 u_t(\lambda x,\lambda^2 t)\\
\Delta u_\lambda(\lambda x,\lambda^2 t)&=\lambda^2\Delta u(\lambda x,\lambda^2 t)
\end{aligned}
$$

得到

$$
(u_\lambda)_t-\Delta u_\lambda=\lambda^2(u_t-\Delta u)=0
$$

命题即证.

(b) 用 (a) 的结论证明 $v(x,t):=x\cdot Du(x,t)+2tu_t(x,t)$ 也是热方程的解.

证明

注意

$$
w(x,t,\lambda):=\frac{d}{d\lambda}u_\lambda(x,t)=x\cdot Du(\lambda x,\lambda^2 t)+2\lambda t u_t(\lambda x,\lambda^2 t)
$$

既然 u 是光滑的, 知

$$
\begin{aligned}
\frac{d}{d\lambda}((u_\lambda)_t)&=(\frac{d}{d\lambda}u_\lambda)_t=w_t\\
\frac{d}{d\lambda}(\Delta u_\lambda)&=\Delta(\frac{d}{d\lambda}u_\lambda)=\Delta w
\end{aligned}
$$

又因为 $(u_\lambda)_t-\Delta u_\lambda$ 已经获证为常值函数, 故 $\frac{d}{d\lambda}((u_\lambda)_t-\Delta u_\lambda)=0$, 亦即 $w_t-\Delta w=0$, 这说明对任意的 λ 而言, w 都是热方程的解. 特别取 $\lambda=1$ 即得 $v(x,t)$ 是热方程的解, 命题得证. □

练习 7.13♣ 设 $n=1$, $u(x,t)=v(\frac{x}{\sqrt{t}})$.

(a) 证明 $u_t=u_{xx}$ 当且仅当

$$
v''+\frac{z}{2}v'=0 \tag{*}
$$

证明 (*) 的通解为

$$
v(z)=c\int_0^z e^{-\frac{s^2}{4}}ds+d
$$

证明

当 $u_t = u_{xx}$ 时, 知

$$u_t = \frac{\partial}{\partial t}(v(\frac{x}{\sqrt{t}})) = -\frac{x}{2t^{\frac{3}{2}}} \cdot v'(\frac{x}{\sqrt{t}})$$
$$u_{xx} = \frac{\partial}{\partial x}(v(\frac{x}{\sqrt{t}})) = \frac{1}{t} \cdot v''(\frac{x}{\sqrt{t}})$$

故

$$\frac{1}{t} \cdot v''(\frac{x}{\sqrt{t}}) + \frac{x}{2t^{\frac{3}{2}}} \cdot v'(\frac{x}{\sqrt{t}}) = 0 \Rightarrow v''(\frac{x}{\sqrt{t}}) + \frac{x}{2\sqrt{t}}v'(\frac{x}{\sqrt{t}}) = 0 \Rightarrow v'' + \frac{z}{2}v' = 0$$

此即 (*) 式.

反之, 当 (*) 式成立, 代入 $z = \frac{x}{\sqrt{t}}$ 即知 $u_t = u_{xx}$.

最后求解 (*) 式, 知

$$\frac{dv'}{v'} = -\frac{z}{2}dz \Rightarrow \ln|v'| = -\frac{z^2}{4} \Rightarrow v' = ce^{-\frac{z^2}{4}} \Rightarrow v = c\int_0^z e^{-\frac{s^2}{4}}ds + d$$

此即欲证.

(b) 通过对 $u(x,t) = v(\frac{x}{\sqrt{t}})$ 作关于 x 的微分, 并适当选取常数 c, 导出 $n = 1$ 时的基本解 Φ. 解释为什么这个过程可以得到基本解.

解

知

$$u_x(x,t) = \frac{1}{\sqrt{t}} \cdot v'(\frac{x}{\sqrt{t}}) = \frac{1}{\sqrt{t}} \cdot v'(z) = \frac{z}{x} \cdot ce^{-\frac{z^2}{4}} = \frac{1}{\sqrt{t}} \cdot ce^{-\frac{x^2}{4t}}$$

当 $u_t - u_{xx} = 0$, 显见 $\frac{\partial}{\partial x}(u_t - u_{xx}) = 0$, 也即 $u_{xt} - u_{xxx} = 0$, 这说明 u_x 也应该是热方程的解. 归一化知 $c = \frac{1}{\sqrt{4\pi}}$, 这便得到了 $n = 1$ 时的基本解 Φ.

练习 7.14♣ 写出方程

$$\begin{cases} u_t - \Delta u + cu = f, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases}$$

的一个精确解, 其中 $c \in \mathbb{R}$.

解

定义 Fourier 变换 $\mathcal{F}[\cdot] = \hat{\cdot}$ 与 Fourier 逆变换 $\check{\mathcal{F}}[\cdot] = \check{\cdot}$:

$$\mathcal{F}[u](\xi,t) = \hat{u}(\xi,t) = \int_{\mathbb{R}^n} e^{-i(x|\xi)}u(x,t)dx$$
$$\check{\mathcal{F}}[u](x,t) = \check{u}(x,t) = \int_{\mathbb{R}^n} e^{i(x|\xi)}u(\xi,t)d\xi$$

其中 $(x|\xi)$ 表示 x 与 ξ 作内积. 若记 $i^\alpha := i^{|\alpha|}$, 则容易验证:

$$\widehat{D^\alpha u}(\xi,t) = (i\xi)^\alpha \hat{u}(\xi,t), \quad \check{\mathcal{F}}[\mathcal{F}[u]] = (2\pi)^n u, \quad \mathcal{F}(p * q) = \hat{p} \cdot \hat{q}$$

根据方程的线性性与叠加原理, 可分别设方程

$$\begin{cases} v_t - \Delta v + cv = f, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ v = 0, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases} \tag{7.205}$$

$$\begin{cases} w_t - \Delta w + cw = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ w = g, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases} \tag{7.206}$$

根据方程的线性性与 Duhamel 准则, 进一步只需求解方程(7.206), 在方程两边作 Fourier 变换得到:

$$\begin{cases} \widehat{w}_t(\xi,t)+|\xi|^2\widehat{w}(\xi,t)+c\widehat{w}(\xi,t)=0, & 在\ \mathbb{R}^n\times(0,\infty)\ 内 \\ \widehat{w}(\xi,t)=\widehat{g}(\xi), & 在\ \mathbb{R}^n\times\{t=0\}\ 上 \end{cases}$$

求解得

$$\widehat{w}(\xi,t)=\widehat{g}(\xi)e^{-(|\xi|^2+c)t},\quad t\in[0,+\infty) \tag{7.207}$$

对 $e^{-(|\xi|^2+c)t}$ 关于 ξ 作 Fourier 逆变换有:

$$\begin{aligned} \check{\mathcal{F}}[e^{-(|\xi|^2+c)t}](x,t) &= \int_{\mathbb{R}^d} e^{i(x|\xi)}e^{-(|\xi|^2+c)t}d\xi \\ &= \int_{\mathbb{R}^d} e^{-(|\xi|^2-\frac{i}{t}(x|\xi)-\frac{|x|^2}{4t^2})t}e^{-\frac{|x|^2}{4t}-ct}d\xi \\ &= e^{-\frac{|x|^2}{4t}-ct}\int_{\mathbb{R}^d} e^{-ts^2}ds=(\frac{\pi}{t})^{\frac{n}{2}}e^{-\frac{|x|^2}{4t}-ct} \end{aligned}$$

因而可得

$$\mathcal{F}[\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}-ct}](\xi,t)=e^{-(|\xi|^2+c)t}$$

在(7.207)式两端应用 Fourier 逆变换得到:

$$(2\pi)^n w(x,t)=(2\pi)^n(\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|\cdot|^2}{4t}-ct}*g)(x,t)\Rightarrow w(x,t)=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}g(y)dy$$

下面验证

$$w(x,t)=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}g(y)dy \tag{7.208}$$

确为初值问题(7.206)的解. 当 $t\to 0^+$ 时, 已知

$$\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}}g(y)dy\to g(x),\quad e^{-ct}\to 1$$

故 $w(x,t)\to g(x)(t\to 0)$. 而当 $t\in(0,\infty)$ 时, 求导有:

$$\begin{aligned} w_t(x,t) &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}(\frac{|x-y|^2}{4t^2}-c)g(y)dy-\frac{n}{2(4\pi)^{\frac{n}{2}}t^{\frac{n}{2}+1}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}g(y)dy \\ w_{x_i} &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}(-\frac{x_i-y_i}{2t})g(y)dy \\ w_{x_ix_i} &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}\frac{|x_i-y_i|^2}{4t^2}g(y)dy+\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}(-\frac{1}{2t})g(y)dy \\ \Delta w &= \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}\frac{|x-y|^2}{4t^2}g(y)dy-\frac{n}{2(4\pi)^{\frac{n}{2}}t^{\frac{n}{2}+1}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}g(y)dy \end{aligned}$$

显见 $w_t-\Delta w+cw=0$, 因而(7.208)式给出的 w 确实为初值问题(7.206)的解. 现在根据 Duhamel 原理, 对于非齐次初值问题(7.205), 可考虑初值问题

$$\begin{cases} v_t(\cdot,s)-\Delta v(\cdot,s)+cv(\cdot,s)=0, & 在\ \mathbb{R}^n\times(s,\infty)\ 内 \\ v(\cdot,s)=f(\cdot,s), & 在\ \mathbb{R}^n\times\{t=s\}\ 上 \end{cases} \tag{7.209}$$

在齐次初值问题(7.206)中将 t 替换为 $t-s$, g 替换为 f, 即得齐次初值问题(7.209)的一个解

$$v(\cdot,s)=\int_{\mathbb{R}^n}\frac{1}{(4\pi(t-s))^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4(t-s)}-c(t-s)}f(y,s)dy,\quad 0<s<t$$

由 Duhamel 原理知非齐次初值问题(7.205)的一个解为

$$v(x,t)=\int_0^t v(x,t,s)ds=\int_0^t\int_{\mathbb{R}^n}\frac{1}{(4\pi(t-s))^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4(t-s)}-c(t-s)}f(y,s)dyds$$

最后根据叠加原理知原问题的一个解为:

$$u(x,t)=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}-ct}g(y)dy+\int_0^t\int_{\mathbb{R}^n}\frac{1}{(4\pi(t-s))^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4(t-s)}-c(t-s)}f(y,s)dyds$$

□

另证

对于问题

$$\begin{cases} u_t - \Delta u + cu = f, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

构造 $v(x,t) = u(x,t)e^{ct}$, 则 v 满足

$$\begin{cases} v_t(x,t) - \Delta v(x,t) = e^{ct} f(x,t), & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ v(x,t) = e^{ct} g(x), & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \tag{7.210}$$

根据热方程非齐次初值问题的结论, 知问题(7.210)的一个解为

$$v(x,t) = \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)dy + \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s)f(y,s)dyds$$

代入 v 的定义知

$$u(x,t)e^{ct} = \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)dy + \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s)f(y,s)dyds$$

亦即

$$u(x,t) = e^{-ct} \int_{\mathbb{R}^n} \Phi(x-y,t)g(y)dy + \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s)f(y,s)e^{-ct}dyds$$

□

✍ **练习 7.15**♣ 给定 $g : [0,\infty) \to \mathbb{R}, g(0) = 0$, 对于初边值问题

$$\begin{cases} u_t - u_{xx} = 0, & \text{在 } \mathbb{R}^+ \times (0,\infty) \text{ 内} \\ u = 0, & \text{在 } \mathbb{R}^+ \times \{t=0\} \text{ 上} \\ u = g, & \text{在 } \{x=0\} \times [0,\infty) \text{ 上} \end{cases}$$

推出其解的表达式:

$$u(x,t) = \frac{x}{\sqrt{4\pi}} \int_0^t \frac{1}{(t-s)^{\frac{3}{2}}} e^{-\frac{x^2}{4(t-s)}} g(s)ds$$

证明[17]

将欲证理解为: 对任意紧集 $K \subset R^+$ 与任意 $T > 0$, 当 $(x,t) \in K \times [0,T]$, 则

$$u(x,t) = \frac{x}{\sqrt{4\pi}} \int_0^t \frac{1}{(t-s)^{\frac{3}{2}}} e^{-\frac{x^2}{4(t-s)}} g(s)ds$$

构造 $v(x,t) := u(x,t) - g(t)$, 知 $v(x,t)$ 满足:

$$\begin{cases} v_t - v_{xx} = -g'(t), & \text{在 } \mathbb{R}^+ \times (0,\infty) \text{ 内} \\ v = 0, & \text{在 } \mathbb{R}^+ \times \{t=0\} \text{ 上} \\ v = 0, & \text{在 } \{x=0\} \times [0,\infty) \text{ 上} \end{cases}$$

对 v 关于 x 作奇延拓, 取

$$\widetilde{v}(x,t) := \begin{cases} v(x,t), & x \geq 0 \\ -v(-x,t), & x < 0 \end{cases}$$

为简化符号, 依旧把 $\widetilde{v}$ 记作 v, 有:

$$\begin{cases} v_t(x,t) - v_{xx}(x,t) = -\operatorname{sgn}(x)g'(t) =: h(x,t), & \text{在 } \mathbb{R} \times (0,\infty) \text{ 内} \\ v = 0, & \text{在 } \mathbb{R} \times \{t=0\} \text{ 上} \end{cases}$$

[17] 依照书上的提示作出的奇延拓没法保证紧支连续性, 从而这里没有办法直接套用书上关于热方程非齐次问题解的结论. 为了绕开这一障碍, 只好弱化结论, 选用在紧集上证明这件事...

在将问题转化为 $K\times[0,T]$ 上的问题后, 记 $K_0=\{x\in\mathbb{R}:|x|\in K\}$, 知总存在 $\mathbb{R}$ 中的闭球 $\overline{B}(0,r)\supset K_0$, 记 $B_T=\overline{B}(0,r)\times[0,T]$. 现在在 B_T 内对每个固定的 t, 将 $h(x,t)$ 在 $\mathbb{R}$ 上磨光化, 设 $h^\varepsilon(x,t)$ 是支集在 B_T 中的磨光函数, 根据 $\overline{B}(0,r)$ 的紧性知 $h^\varepsilon(x,t)\rightrightarrows h(x,t)(\varepsilon\to 0)$. 现在对固定的指标 ε 考虑初值问题:

$$\begin{cases} v_t^\varepsilon(x,t)-v_{xx}^\varepsilon(x,t)=h^\varepsilon(x,t), & \text{在 } \mathbb{R}\times(0,T] \text{ 内} \\ v^\varepsilon=0, & \text{在 } \mathbb{R}\times\{t=0\} \text{ 上} \end{cases}$$

依照非齐次热方程解的结论与 h^ε 的紧支性知:

$$v^\varepsilon(x,t)=\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{\mathbb{R}}e^{-\frac{|x-y|^2}{4(t-s)}}h^\varepsilon(y,s)dyds$$

依照热方程在有界域上解的唯一性与恒等逼近的一致性, 令 $\varepsilon\to 0$ 可得:

$$v(x,t)=\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{\mathbb{R}}e^{-\frac{|x-y|^2}{4(t-s)}}h(y,s)dyds$$

代入 $h(y,s)$ 得:

$$\begin{aligned} v(x,t)&=\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{-\infty}^0e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds-\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_0^\infty e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds \\ &=-\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{\mathbb{R}}e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds+2\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{-\infty}^0e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds \\ &=-\int_0^t g'(s)ds+2\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{-\infty}^0e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds \end{aligned}$$

对于上后式, 知

$$\int_{-\infty}^0\frac{e^{-\frac{|x-y|^2}{4(t-s)}}}{(4\pi(t-s))^{\frac12}}dy=\frac{1}{\sqrt{\pi}}\int_{\frac{x}{\sqrt{4(t-s)}}}^\infty e^{-z^2}dz=:a(s)$$

$a(s)$ 对 s 求导可得:

$$a'(s)=-\frac{1}{\sqrt{\pi}}\cdot\frac{x}{4(t-s)^{\frac32}}e^{-\frac{x^2}{4(t-s)}}$$

进而

$$\begin{aligned} &2\int_0^t\frac{1}{(4\pi(t-s))^{\frac12}}\int_{-\infty}^0e^{-\frac{|x-y|^2}{4(t-s)}}g'(s)dyds \\ =&2\int_0^t a(s)g'(s)ds=2(\lim_{s\to t^-}a(s)g(s)-a(0)g(0)-\int_0^t g(s)a'(s)ds) \\ =&2\int_0^t g(s)\frac{1}{\sqrt{\pi}}\cdot\frac{x}{4(t-s)^{\frac32}}e^{-\frac{x^2}{4(t-s)}}ds=\frac{x}{\sqrt{4\pi}}\int_0^t\frac{1}{(t-s)^{\frac32}}e^{-\frac{x^2}{4(t-s)}}g(s)ds \end{aligned}$$

又显见 $-\int_0^t g'(s)ds=-g(t)$, 从而

$$u(x,t)=v(x,t)+g(t)=\frac{x}{\sqrt{4\pi}}\int_0^t\frac{1}{(t-s)^{\frac32}}e^{-\frac{x^2}{4(t-s)}}g(s)ds$$

取 $x>0$ 的部分即得欲证. □

✍ **练习 7.16**♣ 直接证明命题: 若 U 有界, $u\in C_1^2(U_T)\cap C(\overline{U}_T)$ 是热方程的解, 则

$$\max_{\overline{U}_T}u=\max_{\Gamma_T}u.$$

证明

首先证明: 若 U 有界, $u\in C_1^2(U_T)\cap C(\overline{U}_T)$ 满足

$$u_t(x,t)-\Delta u(x,t)=f(x,t)<0,\quad\forall(x,t)\in U_T$$

则 u 满足弱极值原理. 这是因为若存在 $(x_0,t_0)\in U_T$ 满足

$$u(x_0,t_0)=\max_{\overline{U}_T}u(x,t)$$

则根据极值的必要条件知:

$$D^2u(x_0,t_0)=\begin{pmatrix}u_{x_1x_1} & \cdots & u_{x_1x_n}\\ \vdots & & \vdots\\ u_{x_nx_1} & \cdots & u_{x_nx_n}\end{pmatrix}(x_0,t_0)$$

是负定阵, 这说明 $D^2u(x_0,t_0)$ 仅有非正特征值, 从而特征值之和 $\mathrm{tr}\big(D^2u(x_0,t_0)\big)=\Delta u(x_0,t_0)\le 0$. 同时

$$\frac{\partial u}{\partial t}(x_0,t_0)\begin{cases}=0, & t_0<T\\ \ge 0, & t_0=T\end{cases}$$

故 $u_t(x_0,t_0)-\Delta u(x_0,t_0)\ge 0$, 这与条件矛盾! 从而只能有 $(x_0,t_0)\in\Gamma_T$, 命题即证.

回到原题, 取 $u_\varepsilon(x,t):=u(x,t)-\varepsilon t(\varepsilon>0)$, 则 $(\partial_t-\Delta)u_\varepsilon(x,t)=-\varepsilon<0$, 这说明 u_ε 满足弱极值原理, 得到:

$$\max_{\overline{U}_T}u_\varepsilon(x,t)=\max_{\Gamma_T}u_\varepsilon(x,t)$$

代入 u_ε 的构造有:

$$\begin{aligned}\max_{\overline{U}_T}u(x,t)&=\max_{\overline{U}_T}(u_\varepsilon(x,t)+\varepsilon t)\le\max_{\overline{U}_T}u_\varepsilon(x,t)+\varepsilon T\\ &=\max_{\Gamma_T}u_\varepsilon(x,t)+\varepsilon T\le\max_{\Gamma_T}u(x,t)+\varepsilon T\end{aligned}$$

由 $T<\infty$ 与 ε 的任意性即得 $\max_{\overline{U}_T}u(x,t)\le\max_{\Gamma_T}u(x,t)$, 但显见应有 $\max_{\overline{U}_T}u(x,t)\ge\max_{\Gamma_T}u(x,t)$, 从而 $\max_{\overline{U}_T}u(x,t)=\max_{\Gamma_T}u(x,t)$, 命题得证. □

✍ **练习 7.17(下解)**♣ 称 $v\in C_1^2(U_T)$ 是热方程的下解, 如果

$$v_t-\Delta v\le 0\quad\text{在 } U_T \text{ 内}$$

(a) 证明下解 v 满足对任意的 $E(x,t;r)\subset U_T$ 都有

$$v(x,t)\le\frac{1}{4r^n}\iint_{E(x,t;r)}v(y,s)\frac{|x-y|^2}{(t-s)^2}dyds$$

证明

根据热传导过程的时空不变性, 记 $E(r):=E(0,0,r)$, 则只需验证

$$v(0,0)\le\frac{1}{4r^n}\iint_{E(r)}v(y,s)\frac{|y|^2}{s^2}dyds \tag{7.211}$$

现设

$$\phi(r):=\frac{1}{r^n}\iint_{E(r)}v(y,s)\frac{|y|^2}{s^2}dyds$$

已知经过伸缩变换有

$$\phi(r)=\iint_{E(1)}v(ry,r^2s)\frac{|y|^2}{s^2}dyds$$

对 r 求导得:

$$\phi'(r)=\frac{1}{r^{n+1}}\iint_{E(r)}((\sum_{i=1}^n v_{y_i}(y,s)y_i\frac{|y|^2}{s^2})+2v_s(y,s)\frac{|y|^2}{s})dyds \tag{7.212}$$

取

$$\psi(s,y,r):=-\frac{n}{2}\ln(-4\pi s)+\frac{|y|^2}{4s}+n\ln r$$

知 $\partial E(r)\backslash\{(0,0)\} = \{(y,s)\in\mathbb{R}^{n+1} : \psi(s,y,r)=0\}$, 计算有 $\psi_{y_i}(s,y,r) = \frac{y_i}{2s}$, 对(7.212)式应用分部积分可得:

$$\phi'(r) = -\frac{1}{r^{n+1}}\iint_{E(r)}(4nv_s(y,s)\psi(s,y,r)+\frac{2n}{s}\sum_{i=1}^n v_{y_i}(y,s)y_i)dyds$$

根据 Gauss 公式知

$$\begin{aligned}\iint_{E(r)}\frac{2n}{s}\sum_{i=1}^n v_{y_i}(y,s)y_i dyds &= 4n\iint_{E(r)}D_y v(y,s)\cdot D_y\psi(s,y,r)dyds\\ &= 4n(-\iint_{E(r)}\psi(s,y,r)\Delta v(y,s)dyds+\iint_{\partial E(r)}\psi(s,y,r)\frac{\partial v}{\partial \boldsymbol{\nu}}(y,s)d\sigma)\\ &= -\iint_{E(r)}4n\psi(s,y,r)\Delta v(y,s)dyds\end{aligned}$$

从而

$$\phi'(r) = -\frac{1}{r^{n+1}}(4nv_s(y,s)\psi(s,y,r)-\iint_{E(r)}4n\psi(s,y,r)\Delta v(y,s))dyds \ge 0$$

这说明 $\phi(r)\ge\lim\limits_{r\to0^+}\phi(r)$, 又已知 $\lim\limits_{r\to0^+}\phi(r)=4v(0,0)$, 故 $\phi(r)\ge 4v(0,0)$, (7.211)式得证, 因而命题得证. □

(b) 证明 $\max\limits_{\overline{U}_T} v = \max\limits_{\Gamma_T} v$[18].

证明

因为 U_T 是 $\mathbb{R}^{n+1}$ 中的有界区域, 故由恒等逼近定理可设 v 在 Γ_T 上足够光滑. 现设存在 $(x_0,t_0)\in U_T$ 使得 $v(x_0,t_0)=\max\limits_{\overline{U}_T}v$, 根据 U_T 的开性知存在 $r>0$ 使得 $E(x_0,t_0,r)\subset U_T$, 进而由 (a) 结论知

$$v(x_0,t_0)\le\frac{1}{4r^n}\iint_{E(x_0,t_0,r)}v(y,s)\frac{|x-y|^2}{(t-s)^2}dyds$$

等号取得当且仅当 $v(x,t)\equiv v(x_0,t_0)((x,t)\in E(x_0,t_0,r))$, 但另一方面

$$v(y,s)\le\max_{E(x_0,t_0,r)}v\le\max_{\overline{U}_T}v=v(x_0,t_0)$$

从而

$$\frac{1}{4r^n}\iint_{E(x_0,t_0,r)}v(y,s)\frac{|x-y|^2}{(t-s)^2}dyds\le\frac{1}{4r^n}\iint_{E(x_0,t_0,r)}v(x_0,t_0)\frac{|x-y|^2}{(t-s)^2}dyds=v(x_0,t_0)$$

因而只能有 $v(x,t)\equiv v(x_0,t_0)((x,t)\in E(x_0,t_0,r))$. 现在另任取 $(x_1,t_1)\in U_T$ 使得 $t_1<t_0$, 用线段 L 连接 $(x_0,t_0),(x_1,t_1)$, 设

$$r_0 := \min\{s\ge t_1 : \forall (x,t)\in L(v(x,t)=v(x_0,t_0)), s\le t\le t_0\}$$

根据 $v\in C_1^2(U_T)$ 知 r_0 是良定义的, 下面证明 $r_0=t_1$. 若 $r_0>t_1$, 则已知存在 $\sigma>0$ 使得 $E(x_0,t_0,r)\cap(L\cap\{r_0-\sigma\le t\le r_0\})\ne\varnothing$, 这说明依旧存在 $r'<r_0$ 使得 $u(x,t)=M((x,t)\in L, r'\le t\le t_0)$, 与 r_0 的构造矛盾! 故只能有 $r_0=t_1$, 因而在 L 上必有 $v(x,t)\equiv v(x_0,t_0)$. 最后, 在 $\overline{U}_T$ 的每个连通分支内, 对任意两点 $(x,t),(x',t')$, 总存在前述连线的程序保证 $v(x,t)=v(x',t')=v(x_0,t_0)$, 这便说明 v 在 $\overline{U}_T$ 的每个连通分支内都是常值函数, 故 $\max\limits_{\overline{U}_T}v=\max\limits_{\Gamma_T}v$. □

(c) 设 $\phi:\mathbb{R}\to\mathbb{R}$ 是光滑凸函数, 设 u 是热方程的解, $v:=\phi(u)$, 证明 v 是一个下解.

证明

知

$$\begin{aligned}v_t &= \phi'(u)u_t\\ v_{x_i} &= \phi'(u)u_{x_i}\\ v_{x_ix_i} &= \phi''(u)u_{x_i}^2+\phi'(u)u_{x_ix_i}\end{aligned}$$

[18] v 不一定在 Γ_T 上连续, 故本题疑需将 max 改为 sup, 不过结论是类似的.

从而

$$\begin{aligned} v_t - \Delta v &= \phi'(u)u_t - \phi''(u)\sum_{i=1}^n u_{x_i}^2 - \phi'(u)\Delta u \\ &= -\phi''(u)\sum_{i=1}^n u_{x_i}^2 \le 0 \end{aligned}$$

其中最后的不等号源自 ϕ 的光滑凸性, 故 v 是一个下解. □

(d) 证明 $v := |Du|^2 + u_t^2$ 是下解, 其中 u 是热方程的解.

证明

知

$$\begin{aligned} v_t &= 2\langle Du_t, Du\rangle + 2u_t u_{tt} \\ v_{x_i} &= 2\langle Du_{x_i}, Du\rangle + 2u_t u_{tx_i} \\ v_{x_i x_i} &= 2\langle Du_{x_i x_i}, Du\rangle + 2|Du_{x_i}|^2 + 2u_{tx_i}^2 + 2u_t u_{tx_i x_i} \end{aligned}$$

从而

$$\begin{aligned} v_t - \Delta v &= 2\langle D(u_t - \Delta u), Du\rangle + 2u_t u_{tt} - 2|Du_{x_i}|^2 - 2\sum_{i=1}^n u_{tx_i}^2 - 2u_t \Delta u_t \\ &= 2u_t(\frac{\partial}{\partial t} - \Delta)u_t - 2|Du_{x_i}|^2 - 2\sum_{i=1}^n u_{tx_i}^2 = -2\sum_{i=1}^n u_{tx_i}^2 \le 0 \end{aligned}$$

故 v 是一个下解. □

练习 7.18(Stokes 法则)♣ 设 u 满足初值问题

$$\begin{cases} u_{tt} - \Delta u = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = 0, u_t = h, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

证明 $v := u_t$ 满足

$$\begin{cases} v_{tt} - \Delta v = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ v = h, v_t = 0, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

这便是 Stokes 法则.

证明

当 $(x,t) \in \mathbb{R}^n \times (0,\infty)$ 时, 在 $u_{tt} - \Delta u = 0$ 两边对 t 求偏导即得:

$$(u_t)_{tt} - \Delta u_t = 0 \Rightarrow v_{tt} - \Delta v = 0$$

当 $(x,t) \in \mathbb{R}^n \times \{t=0\}$ 时, 知 $u_t(x,t) = h(x) = v(x,t)$, 而因为 h 与 t 无关, 故在 $\mathbb{R}^n \times \{t=0\}$ 上 $v_t = 0$, 得到

$$\begin{cases} v_{tt} - \Delta v = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ v = h, v_t = 0, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

□

练习 7.19

(a) 证明 PDE$u_{xy} = 0$ 的通解为

$$u(x,y) = F(x) + G(y)$$

其中 F, G 是任意函数.

(b) 利用换元 $\xi = x + t, \eta = x - t$, 证明 $u_{tt} - u_{xx} = 0$ 当且仅当 $u_{\xi\eta} = 0$.

(c) 利用 (a),(b) 重新推出 d'Alembert 公式.

(d) 初值 g, h 在什么条件下, u 成为右移波? 什么时候是左移波?

✍ **练习 7.20** 设对于衰减函数 $\alpha = \alpha(r)$ 与迟滞函数 $\beta = \beta(r) \geq 0$, 对全体波形 ϕ 都存在形如

$$u(x,t) = \alpha(r)\phi(t-\beta(r))$$

的波方程在 $(\mathbb{R}^n \backslash \{0\}) \times \mathbb{R}$ 上的解. 其中 $r = |x|$, 设 $\beta(0) = 0$.

证明这件事只在 $n = 1$ 或 3 时可能成立, 并计算函数 α, β 的形式.

✍ **练习 7.21**♣

(a) 设 $\boldsymbol{E} = (E^1, E^2, E^3), \boldsymbol{B} = (B^1, B^2, B^3)$ 满足 Maxwell 方程组:

$$\begin{cases} \boldsymbol{E}_t = \operatorname{curl} \boldsymbol{B}, \boldsymbol{B}_t = -\operatorname{curl} \boldsymbol{E} \\ \operatorname{div} \boldsymbol{B} = \operatorname{div} \boldsymbol{E} = 0 \end{cases}$$

证明

$$\boldsymbol{E}_{tt} - \Delta \boldsymbol{E} = 0, \quad \boldsymbol{B}_{tt} - \Delta \boldsymbol{B} = 0.$$

证明

对向量值函数 $\boldsymbol{f} = (f^1, f^2, f^3)^T \in \mathbb{R}^3$ 而言, 知

$$\operatorname{curl} \boldsymbol{f} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ f^1 & f^2 & f^3 \end{vmatrix} = \begin{pmatrix} f_2^3 - f_3^2 \\ f_3^1 - f_1^3 \\ f_1^2 - f_2^1 \end{pmatrix}$$

进一步

$$\operatorname{curl}(\operatorname{curl} \boldsymbol{f}) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ f_2^3 - f_3^2 & f_3^1 - f_1^3 & f_1^2 - f_2^1 \end{vmatrix} = \begin{pmatrix} (f_{12}^2 - f_{22}^1) - (f_{33}^1 - f_{13}^3) \\ (f_{23}^3 - f_{33}^2) - (f_{11}^2 - f_{21}^1) \\ (f_{31}^1 - f_{11}^3) - (f_{22}^3 - f_{32}^2) \end{pmatrix}$$

如若 $\operatorname{div} \boldsymbol{f} = f_1^1 + f_2^2 + f_3^3 = 0$, 则该等式两边分别对 x_1, x_2, x_3 求偏导有:

$$\begin{aligned} f_{11}^1 + f_{12}^2 + f_{13}^3 = 0 &\Rightarrow f_{12}^2 + f_{13}^3 = -f_{11}^1 \\ f_{21}^1 + f_{22}^2 + f_{23}^3 = 0 &\Rightarrow f_{21}^1 + f_{23}^3 = -f_{22}^2 \\ f_{31}^1 + f_{32}^2 + f_{33}^3 = 0 &\Rightarrow f_{31}^1 + f_{32}^2 = -f_{33}^3 \end{aligned}$$

从而

$$\begin{pmatrix} (f_{12}^2 - f_{22}^1) - (f_{33}^1 - f_{13}^3) \\ (f_{23}^3 - f_{33}^2) - (f_{11}^2 - f_{21}^1) \\ (f_{31}^1 - f_{11}^3) - (f_{22}^3 - f_{32}^2) \end{pmatrix} = \begin{pmatrix} -\Delta f^1 \\ -\Delta f^2 \\ -\Delta f^3 \end{pmatrix}$$

现由 $\operatorname{div} \boldsymbol{E} = 0$ 知

$$\boldsymbol{E}_{tt} = \operatorname{curl} \boldsymbol{B}_t = -\operatorname{curl}(\operatorname{curl} \boldsymbol{E}) = (\Delta E^1, \Delta E^2, \Delta E^3)^T = \Delta \boldsymbol{E}$$

亦即 $\boldsymbol{E}_{tt} - \Delta \boldsymbol{E} = 0$. 同理, 由 $\operatorname{div} \boldsymbol{B} = 0$ 知

$$\boldsymbol{B}_{tt} = -\operatorname{curl} \boldsymbol{E}_t = -\operatorname{curl}(\operatorname{curl} \boldsymbol{B}) = (\Delta B^1, \Delta B^2, \Delta B^3)^T = \Delta \boldsymbol{B}$$

亦即 $\boldsymbol{B}_{tt} - \Delta \boldsymbol{B} = 0$. □

(b) 设 $\boldsymbol{u} = (u^1, u^2, u^3)$ 满足线性弹力演化方程

$$\boldsymbol{u}_{tt} - \mu \Delta \boldsymbol{u} - (\lambda + \mu) D(\operatorname{div} \boldsymbol{u}) = \boldsymbol{0}, \quad \text{在 } \mathbb{R}^3 \times (0, \infty) \text{ 中}$$

证明 $w := \operatorname{div} \boldsymbol{u}, \boldsymbol{w} := \operatorname{curl} \boldsymbol{u}$ 都满足波方程, 但它们的传播速度不同.

证明

对于 w, 注意

$$\Delta \boldsymbol{u} = \begin{pmatrix} u_{11}^1 + u_{22}^1 + u_{33}^1 \\ u_{11}^2 + u_{22}^2 + u_{33}^2 \\ u_{11}^3 + u_{22}^3 + u_{33}^3 \end{pmatrix}, \quad D(\operatorname{div} \boldsymbol{u}) = \begin{pmatrix} u_{11}^1 + u_{12}^2 + u_{13}^3 \\ u_{21}^1 + u_{22}^2 + u_{23}^3 \\ u_{31}^1 + u_{32}^2 + u_{33}^3 \end{pmatrix}$$

知

$$\begin{aligned} w_{tt} &= \operatorname{div} \boldsymbol{u}_{tt} = \operatorname{div}(\mu\Delta\boldsymbol{u} + (\lambda+\mu)D(\operatorname{div}\boldsymbol{u})) \\ &= \mu(u_{111}^1 + u_{122}^1 + u_{133}^1 + u_{112}^2 + u_{222}^2 + u_{233}^2 + u_{311}^3 + u_{322}^3 + u_{333}^3) \\ &\quad + (\lambda+\mu)(u_{111}^1 + u_{112}^2 + u_{113}^3 + u_{221}^1 + u_{222}^2 + u_{223}^3 + u_{331}^1 + u_{332}^2 + u_{333}^3) \\ &= (\lambda+2\mu)(u_{111}^1 + u_{112}^2 + u_{113}^3 + u_{221}^1 + u_{222}^2 + u_{223}^3 + u_{331}^1 + u_{332}^2 + u_{333}^3) \end{aligned}$$

而

$$\begin{aligned} \Delta w &= \Delta(\operatorname{div} u) = \Delta(u_1^1 + u_2^2 + u_3^3) \\ &= u_{111}^1 + u_{211}^2 + u_{311}^3 + u_{122}^1 + u_{222}^2 + u_{322}^3 + u_{133}^1 + u_{233}^2 + u_{333}^3 \end{aligned}$$

故只要 u 三阶导数可换, 则有

$$w_{tt} - (\lambda+2\mu)\Delta w = 0$$

对于 $\boldsymbol{w}$, 注意

$$\begin{aligned} \operatorname{curl}(\Delta\boldsymbol{u}) &= \begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ u_{11}^1 + u_{22}^1 + u_{33}^1 & u_{11}^2 + u_{22}^2 + u_{33}^2 & u_{11}^3 + u_{22}^3 + u_{33}^3 \end{pmatrix} \\ &= \begin{pmatrix} (u_{112}^3 + u_{222}^3 + u_{332}^3) - (u_{113}^2 + u_{223}^2 + u_{333}^2) \\ (u_{113}^1 + u_{223}^1 + u_{333}^1) - (u_{111}^3 + u_{221}^3 + u_{331}^3) \\ (u_{111}^2 + u_{221}^2 + u_{331}^2) - (u_{112}^1 + u_{222}^1 + u_{332}^1) \end{pmatrix} \end{aligned}$$

同时

$$\begin{aligned} \operatorname{curl}(D \operatorname{div} \boldsymbol{u}) &= \begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ u_{11}^1 + u_{12}^2 + u_{13}^3 & u_{21}^1 + u_{22}^2 + u_{23}^3 & u_{31}^1 + u_{32}^2 + u_{33}^3 \end{pmatrix} \\ &= \begin{pmatrix} (u_{231}^1 + u_{232}^2 + u_{233}^3) - (u_{321}^1 + u_{322}^2 + u_{323}^3) \\ (u_{311}^1 + u_{312}^2 + u_{313}^3) - (u_{131}^1 + u_{132}^2 + u_{133}^3) \\ (u_{121}^1 + u_{122}^2 + u_{123}^3) - (u_{211}^1 + u_{212}^2 + u_{213}^3) \end{pmatrix} = 0 \end{aligned}$$

且

$$\Delta(\operatorname{curl}\boldsymbol{u}) = \Delta \begin{pmatrix} u_2^3 - u_3^2 \\ u_3^1 - u_1^3 \\ u_1^2 - u_2^1 \end{pmatrix} = \begin{pmatrix} (u_{112}^3 + u_{222}^3 + u_{332}^3) - (u_{113}^2 + u_{223}^2 + u_{333}^2) \\ (u_{113}^1 + u_{223}^1 + u_{333}^1) - (u_{111}^3 + u_{221}^3 + u_{331}^3) \\ (u_{111}^2 + u_{221}^2 + u_{331}^2) - (u_{112}^1 + u_{222}^1 + u_{332}^1) \end{pmatrix}$$

故

$$\begin{aligned} \boldsymbol{w}_{tt} &= \operatorname{curl} \boldsymbol{u}_{tt} = \operatorname{curl}(\mu\Delta\boldsymbol{u} + (\lambda+\mu)D(\operatorname{div}\boldsymbol{u})) \\ &= \mu \operatorname{curl}(\Delta\boldsymbol{u}) = \mu \begin{pmatrix} (u_{112}^3 + u_{222}^3 + u_{332}^3) - (u_{113}^2 + u_{223}^2 + u_{333}^2) \\ (u_{113}^1 + u_{223}^1 + u_{333}^1) - (u_{111}^3 + u_{221}^3 + u_{331}^3) \\ (u_{111}^2 + u_{221}^2 + u_{331}^2) - (u_{112}^1 + u_{222}^1 + u_{332}^1) \end{pmatrix} \end{aligned}$$

同时

$$\Delta\boldsymbol{w} = \Delta \operatorname{curl} \boldsymbol{u} = \begin{pmatrix} (u_{112}^3 + u_{222}^3 + u_{332}^3) - (u_{113}^2 + u_{223}^2 + u_{333}^2) \\ (u_{113}^1 + u_{223}^1 + u_{333}^1) - (u_{111}^3 + u_{221}^3 + u_{331}^3) \\ (u_{111}^2 + u_{221}^2 + u_{331}^2) - (u_{112}^1 + u_{222}^1 + u_{332}^1) \end{pmatrix}$$

这便有 $\boldsymbol{w}_{tt}-\mu\Delta\boldsymbol{w}=0$. 故 w 与 $\boldsymbol{w}$ 均满足波方程, 但前者的波速为 $\lambda+2\mu$, 后者的波速为 μ. □

✍ **练习 7.22** 设 u 表示实轴上以速度 1 向右移动的粒子的浓度, v 表示以速度 1 向左移动的粒子的浓度. 如果向右移动的粒子以比例 $d>0$ 随机变成向左移动, 向左移动的粒子以比例 d 随机变成向右移动, 则得到 PDE 系统

$$\begin{cases} u_t+u_x=d(v-u) \\ v_t-v_x=d(u-v) \end{cases}$$

证明 $w:=u$ 与 $w:=v$ 都满足电报方程

$$w_{tt}+2dw_t-w_{xx}=0.$$

✍ **练习 7.23♣** 设 S 表示在 $\mathbb{R}\times(0,\infty)$ 中顶点分别为 (0,1),(1,2),(0,3),(-1,2) 的正方形, 定义

$$f(x,t):=\begin{cases} -1, & (x,t)\in S\cap\{t>x+2\} \\ 1, & (x,t)\in S\cap\{t<x+2\} \\ 0, & \text{其它} \end{cases}$$

设 u 满足

$$\begin{cases} u_{tt}-u_{xx}=f, & \text{在 } \mathbb{R}\times(0,\infty) \text{ 内} \\ u=0, u_t=0, & \text{在 } \mathbb{R}\times\{t=0\} \text{ 上} \end{cases}$$

描述 $t>3$ 时 u 的形状.

解

显见 f 在 $\mathbb{R}\times(0,\infty)$ 中紧支, 因而存在光滑函数列 $f^\varepsilon\rightrightarrows f(\varepsilon\to0^+)$, 故只需求解

$$\begin{cases} u^\varepsilon_{tt}-u^\varepsilon_{xx}=f^\varepsilon, & \text{在 } \mathbb{R}\times(0,\infty) \text{ 内} \\ u^\varepsilon=0, u^\varepsilon_t=0, & \text{在 } \mathbb{R}\times\{t=0\} \text{ 上} \end{cases}$$

现考虑 Duhamel 原理, 先求解:

$$\begin{cases} w^\varepsilon_{tt}(x,t;s)-w^\varepsilon_{xx}(x,t;s)=0, & \text{在 } \mathbb{R}\times(s,\infty) \text{ 内} \\ w^\varepsilon(x,t;s)=0, w^\varepsilon_t(x,t;s)=f^\varepsilon(x,s), & \text{在 } \mathbb{R}\times\{t=s\} \text{ 上} \end{cases}$$

得

$$w^\varepsilon(x,t;s)=\frac{1}{2}\int_{x-(t-s)}^{x+(t-s)}f^\varepsilon(y,s)dy$$

因而

$$u^\varepsilon(x,t)=\int_0^t w^\varepsilon(x,t;s)ds=\frac{1}{2}\int_0^t\int_{x-(t-s)}^{x+(t-s)}f^\varepsilon(y,s)dyds$$

两边令 $\varepsilon\to0^+$ 并换元有

$$u(x,t)=\frac{1}{2}\int_0^t\int_{x-s}^{x+s}f(y,t-s)dyds$$

现在开始讨论.

当 $t>3$ 时, 任取 xOt 坐标系中的点 (x,t), 知 $u(x,t)$ 的值正是 $f(x,t)$ 在点 $(x,t),(x-t,0),(x+t,0)$ 所成的三角形内积分的一半. 如图7.31, 若设 $l_1: t=x+3, l_2: t=x+2, l_3: t=x+1$, 则当取点 (x,t) 于 $l_2,l_3,t=3$ 围成的区域 (如 A_1) 时, $u(x,t)$ 即对应三角形 $A_1B_1C_1$ 与正方形重叠部分面积的一半, 计算知 $u(x,t)=\frac{1}{2}(t-x-1)$; 当取点 (x,t) 于 $l_1,l_2,t=3$ 围成的区域 (如 A_2) 时, $u(x,t)$ 即对应正方形在 l_2,l_3 之间的部分与在 A_2B_2,l_2 之间的部分面积之差的一半, 计算知 $u(x,t)=\frac{1}{2}(3-t+x)$. 而若取前述未提及区域中的点 (如 A_3,A_4), 则 f 在依赖区域中的部分积分恰

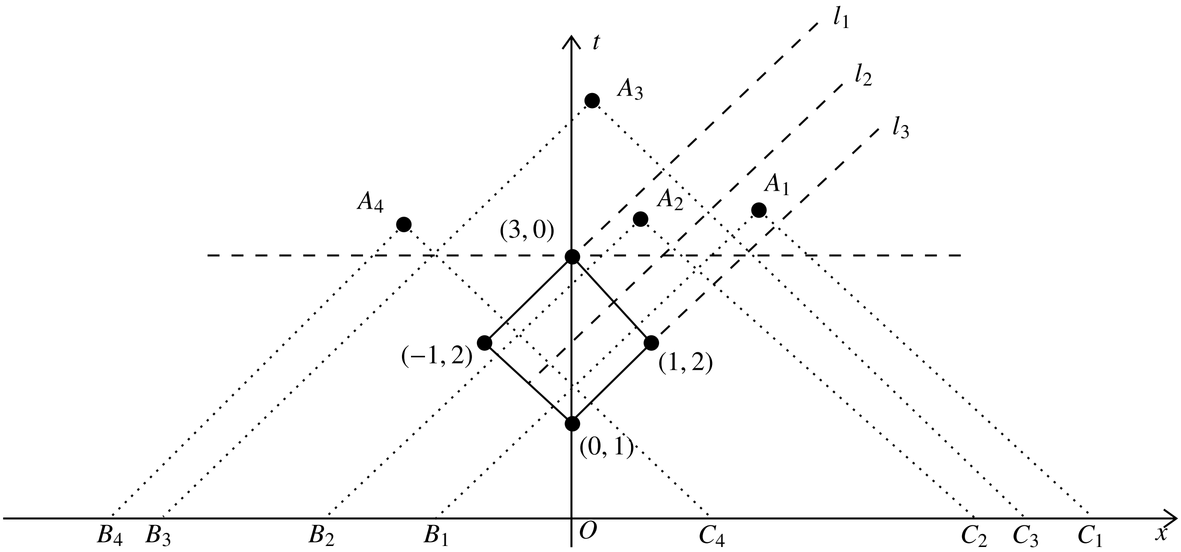


图 7.31: 依赖区域示意图

为 0, 从而 u 在这些点处取值均为 0. 综上, u 的表达式为:

$$u(x,t)=\begin{cases}\frac{1}{2}(t-x-1), & (x,t)\in\{(x,t):t\le x+2,t\ge x+1,t\ge 3\}\\ \frac{1}{2}(3-t+x), & (x,t)\in\{(x,t):t\le x+3,t\ge x+2,t\ge 3\}\\ 0, & \text{其它}\end{cases}$$

u 的形状参见图7.32.

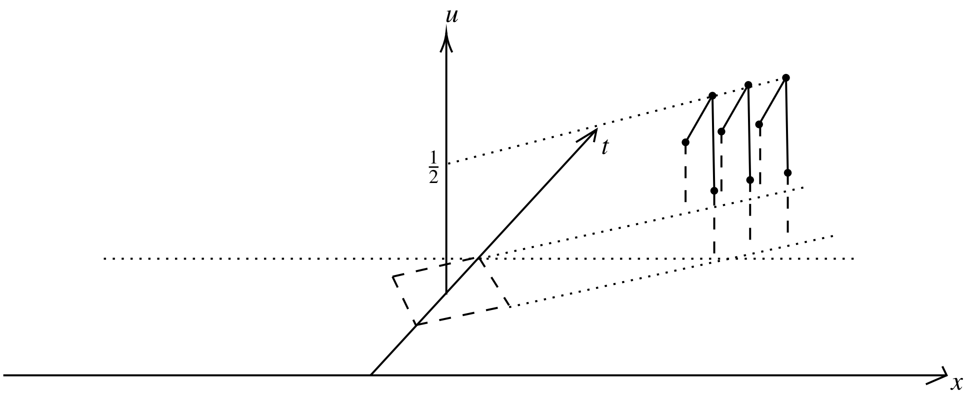


图 7.32: $u(x,t)$ 示意图

✍ **练习 7.24♣** 设 u 满足

$$\begin{cases}u_{tt}-u_{xx}=0, & \text{在 } \mathbb{R}\times(0,\infty) \text{ 内}\\ u=g,u_t=h, & \text{在 } \mathbb{R}\times\{t=0\} \text{ 上}\end{cases}$$

设 g,h 具有紧支集. 动能定义为 $k(t):=\frac{1}{2}\int_{-\infty}^{\infty}u_t^2(x,t)dx$, 势能定义为 $p(t):=\frac{1}{2}\int_{-\infty}^{\infty}u_x^2(x,t)dx$. 证明:

(a) $k(t)+p(t)$ 关于 t 是常数.

证明

记 $e(t)=k(t)+p(t)$, 有

$$\begin{aligned}\frac{d}{dt}e(t)&=\frac{1}{2}\int_{-\infty}^{\infty}2u_tu_{tt}dx+\frac{1}{2}\int_{-\infty}^{\infty}2u_xu_{tx}dx\\&=\int_{-\infty}^{\infty}u_tu_{tt}dx+\int_{-\infty}^{\infty}u_xu_{tx}dx\end{aligned}$$

因为 u_t 满足与 u 相同波速的波方程, 且 h 有紧支集, 故 u_t 对每个固定的 t 关于 x 是紧支的, 从而

$$\int_{-\infty}^{\infty} u_x u_{tx} dx = (u_x u_t)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} u_t u_{xx} dx = -\int_{-\infty}^{\infty} u_t u_{xx} dx$$

因而

$$\frac{d}{dt}e(t) = \int_{-\infty}^{\infty} u_t(u_{tt} - u_{xx})dx = 0$$

亦即 $e(t)$ 关于 t 是常数. □

(b) 对于足够大的 t 有 $k(t) = p(t)$.

证明

根据 d'Alembert 公式有

$$u(x,t) = \frac{1}{2}[g(x+t) + g(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy$$

故

$$\begin{aligned} u_t(x,t) &= \frac{1}{2}[g'(x+t) - g'(x-t)] + \frac{1}{2}h(x+t) + \frac{1}{2}h(x-t) \\ u_x(x,t) &= \frac{1}{2}[g'(x+t) + g'(x-t)] + \frac{1}{2}h(x+t) - \frac{1}{2}h(x-t) \end{aligned}$$

进一步知

$$\begin{aligned} u_t(x,t) - u_x(x,t) &= h(x-t) - g'(x-t) \\ u_t(x,t) + u_x(x,t) &= g'(x+t) + h(x+t) \end{aligned}$$

得到

$$\begin{aligned} u_t^2(x,t) - u_x^2(x,t) &= (u_t(x,t) - u_x(x,t))(u_t(x,t) + u_x(x,t)) \\ &= h(x-t)g'(x+t) + h(x-t)h(x+t) - g'(x-t)g'(x+t) - g'(x-t)h(x+t) \end{aligned}$$

两边对 x 在 $\mathbb{R}$ 上积分, 因为 g, h 均紧支, 故当 t 足够大时, $x-t$ 与 $x+t$ 不可能同时出现在 g, h 的支集中, 因而右式在 $\mathbb{R}$ 上的积分为 0, 得到

$$\int_{-\infty}^{\infty} (u_t^2(x,t) - u_x^2(x,t))dx = k(t) - p(t) = 0 \Rightarrow k(t) = p(t), \quad t \text{ 足够大时}$$

□

第八章 非线性一阶 PDE

本章介绍一般的非线性一阶 PDE, 其形如

$$F(Du,u,x)=0$$

其中 $x\in U, U\subset\mathbb{R}^n$ 是开集. $F:\mathbb{R}^n\times\mathbb{R}\times\overline{U}\to\mathbb{R}$ 是给定的, $u:\overline{U}\to\mathbb{R}$ 未知, $u=u(x)$.

特别引入记号:

$$F=F(p,z,x)=F(p_1,\cdots,p_n,z,x_1,\cdots,x_n)$$

其中 $p\in\mathbb{R}^n, z\in\mathbb{R}, x\in U$. 因此 p 替代了梯度 Du, 而 z 替代了 $u(x)$. 同时设 F 是光滑的, 记

$$\begin{cases} D_pF=(F_{p_1},\cdots,F_{p_n}) \\ D_zF=F_z \\ D_xF=(F_{x_1},\cdots,F_{x_n}) \end{cases}$$

希望解决的问题是找出 PDE$F(Du,u,x)=0$ 的解 u, 一般来说问题还会附上边值条件

$$u=g,\quad \text{在 } \Gamma \text{ 上}$$

其中 $\Gamma\subset\partial U$, 且 $g:\Gamma\to\mathbb{R}$ 给定.

非线性一阶 PDE 在很多物理定律中都有出现, 比如动力学 (用以产生典型的变换), 连续介质力学 (用以表示质量, 动量与能量守恒) 与光学 (用以描述波阵面). 尽管强非线性性一般都把得到通解公式的可能性排除在外了, 但我们依旧可以通过各种计算来收集解的信息. 前两节探讨的典型技巧一般都只能应用在局部上, 而后两节我们会用适当定义的弱解导出 Hamilton-Jacobi 方程组的解与守恒律的表示公式.

8.1 完全积分, 包络

8.1.1 完全积分

我们从非线性一阶 PDE

$$F(Du,u,x)=0 \tag{8.1}$$

开始分析, 讨论一些简单解类, 并学习怎么通过这些解构造更复杂的解.

设 $A\subset\mathbb{R}^n$ 是开集. 设对任意参数 $a=(a_1,\cdots,a_n)\in A$, PDE(8.1)都有一个 C^2 解 $u=u(x;a)$. 记

$$(D_au,D^2_{xa}u):=\begin{pmatrix} u_{a_1} & u_{x_1a_1} & \cdots & u_{x_na_1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{a_n} & u_{x_1a_n} & \cdots & u_{x_na_n} \end{pmatrix}_{n\times(n+1)} \tag{8.2}$$

定义 8.1.1 (完全积分)

C^2 函数 $u=u(x;a)$ 称为 $U\times A$ 中的完全积分, 如果

(i) 对每个 $a\in A$, $u(x;a)$ 都是 PDE(8.1)的解;

(ii) $\operatorname{rank}(D_au,D^2_{xa}u)=n(x\in U,a\in A)$.

♣

条件 (ii) 保证了 $u(x;a)$ 依赖于全体 n 个独立参数 $a_1,\cdots,a_n$. 要说明这一点, 设 $B\subset\mathbb{R}^{n-1}$ 是开的, 且对每个 $b\in B$ 设 $v=v(x;b)(x\in U)$ 是(8.1)的解. 再设存在 C^1 映射 $\boldsymbol{\psi}:A\to B, \boldsymbol{\psi}=(\psi^1,\cdots,\psi^{n-1})$, 满足

$$u(x;a)=v(x;\boldsymbol{\psi}(a)),\quad x\in U, a\in A \tag{8.3}$$

也就是说, 现在假设了一个函数 $u(x;a)$, 它仅仅依赖于 $b_1,\cdots,b_{n-1}$ 这 $n-1$ 个参数. 进一步有

$$u_{x_ia_j}(x;a)=\sum_{k=1}^{n-1}v_{x_ib_k}(x;\boldsymbol{\psi}(a))\psi_{a_j}^k(a),\quad i,j=1,\cdots,n$$

故

$$D_{xa}^2u=\begin{pmatrix}\sum\limits_{k=1}^{n-1}v_{x_1b_k}(x;\boldsymbol{\psi}(a))\psi_{a_1}^k(a) & \cdots & \sum\limits_{k=1}^{n-1}v_{x_1b_k}(x;\boldsymbol{\psi}(a))\psi_{a_n}^k(a)\\ \vdots & & \vdots\\ \sum\limits_{k=1}^{n-1}v_{x_nb_k}(x;\boldsymbol{\psi}(a))\psi_{a_1}^k(a) & \cdots & \sum\limits_{k=1}^{n-1}v_{x_nb_k}(x;\boldsymbol{\psi}(a))\psi_{a_n}^k(a)\end{pmatrix}$$

因而利用行列式的加法有

$$\det\left(D_{xa}^2u\right)=\sum_{k_1,\cdots,k_n=1}^{n-1}v_{x_1b_{k_1}}\cdots v_{x_nb_{k_n}}\det\begin{pmatrix}\psi_{a_1}^{k_1} & \cdots & \psi_{a_n}^{k_1}\\ & \ddots & \\ \psi_{a_1}^{k_n} & \cdots & \psi_{a_n}^{k_n}\end{pmatrix}$$

但注意到既然 $k_1,\cdots,k_n\in\{1,\cdots,n-1\}$, 知 $k_1,\cdots,k_n$ 中必有某两项是一样的, 因而上右式行列式中必有某两行是一样的, 故 $\det\left(D_{xa}^2u\right)=0$. 同时计算知

$$u_{a_j}(x;a)=\sum_{k=1}^{n-1}v_{b_k}(x;\psi(a))\psi_{a_j}^k(a),\quad j=1,\cdots,n$$

与 $u_{x_ia_j}(x;a)$ 的表达式比对即知 (D_au,D_{xa}^2u) 的任意 $n\times n$ 子矩阵的行列式都为 0, 因而 (D_au,D_{xa}^2u) 的秩严格小于 n.

注 完全积分的定义中参数 a 的引入动机, 包括秩条件的解释可能还是有些含糊. 下面借用视频3.1 Evans Complete integrals, envelops中的看法.

回忆 ODE 的相关理论, 在求解方程 $\frac{dy}{dx}=f(x)$ 后, 往往可以得到 $y(x)=g(x)+c$ 的形式, 这里 c 就可以看做一个参数, 因而通解 $u(x,c)=g(x)+c$ 是关于 x 和 c 的函数. 对于 n 阶 ODE, 其通解需要 n 个独立参数 $c_1,\cdots,c_n$, 故 $u(x,c)$ 是关于 $x\in\mathbb{R}$ 和 $c\in\mathbb{R}^n$ 的函数.

现在把目光放回到 PDE 上, 注意此时 $x\in\mathbb{R}^n$, 所以如果 PDE(8.1)中 u 与 F 的性质足够好 (比如 $Du=(f_1(x_1),f_2(x_2),\cdots,f_n(x_n))$, $F(Du,u,x)=Du+x$), 则可以把 PDE(8.1)看作 n 个 ODE 联立得到的方程组 (在前例中即 $f_i(x_i)+x_i=0(i=1,2,\cdots,n)$). 解这个方程组同样需要 n 个独立参数[1], 最后得到的一个特定函数 $u(x;a)$ 正是完全积分, 所以某种意义上可以把完全积分类比成 ODE 中的通解, 参数 a 的设置类比成通解中的”任意常数”.

再来探讨秩条件的动机. 在 ODE 理论中, 求解 n 阶 ODE

$$y^{(n)}+c_{n-1}y^{(n-1)}+\cdots+c_1y'+c_0x=0$$

可得

$$y=f(x,a_1,\cdots,a_n)$$

其中 $a_1,\cdots,a_n$ 是任意 (独立的) 常数. 而如果给定初值后, 是可以求解 $a_1,\cdots,a_n$ 的, 这说明至少需要 n 个方程, 因而考虑方程组:

$$\begin{cases}y=f_1(x,a_1,\cdots,a_n)\\ y'=f_2(x,a_1,\cdots,a_n)\\ \quad\vdots\\ y^{(n-1)}=f_n(x,a_1,\cdots,a_n)\end{cases}$$

现在想要说明独立参数的个数只能是 n, 也即不能存在某个参数可以被其余参数表示出来. 从函数独立性的角

[1]更准确的来说每个单独的 ODE 需要一个参数, 而方程组有 n 个 ODE, 故需 $1\times n$ 个参数. 后面会讨论二阶 PDE, 彼时就会需要 $2\times n$ 个参数了.

度[2]考虑, 这说明隐函数组 $f_1,\cdots,f_n$ 关于 $a_1,\cdots,a_n$ 的 Jacobi 矩阵必须满秩, 否则就存在某个函数 g 与参数 a_i, 使得 $a_i = g(a_1,\cdots,a_{i-1},a_{i+1},\cdots,a_n)$. 即

$$\left|\frac{\partial(y,y',\cdots,y^{(n-1)})}{\partial(a_1,a_2,\cdots,a_n)}\right| = \begin{vmatrix} \frac{\partial y}{\partial a_1} & \cdots & \frac{\partial y}{\partial a_n} \\ \vdots & & \vdots \\ \frac{\partial y^{(n-1)}}{\partial a_1} & \cdots & \frac{\partial y^{(n-1)}}{\partial a_n} \end{vmatrix} \neq 0$$

回到 PDE, 考虑 $u(x;a_1,\cdots,a_n)$ 是一阶 PDE$F(Du,u,x)=0$ 的完全积分, 考虑方程组

$$\begin{cases} u = f(x,a_1,\cdots,a_n) \\ u_{x_1} = f_1(x,a_1,\cdots,a_n) \\ u_{x_2} = f_2(x,a_1,\cdots,a_n) \\ \quad\vdots \\ u_{x_n} = f_n(x,a_1,\cdots,a_n) \end{cases}$$

同样考察参数的独立性, 知不能存在某个参数可以被其余参数表示. 与 ODE 的情况同理, 这要求隐函数组 $f,f_1,\cdots,f_n$ 对 $a_1,\cdots,a_n$ 的 Jacobi 矩阵满秩, 也即

$$\text{rank}\,\frac{\partial(u,u_{x_1},\cdots,u_{x_n})}{\partial(a_1,\cdots,a_n)} = \text{rank}\begin{pmatrix} u_{a_1} & u_{x_1a_1} & \cdots & u_{x_na_1} \\ \vdots & \vdots & & \vdots \\ u_{a_n} & u_{x_1a_n} & \cdots & u_{x_na_n} \end{pmatrix} = n$$

这便是秩条件 (ii).

例 8.1(Clairaut 方程) 微分几何中的 Clairaut 方程是 PDE

$$x\cdot Du + f(Du) = u \tag{8.4}$$

其中 $f:\mathbb{R}\to\mathbb{R}$ 给定. 它的一个完全积分为

$$u(x;a) = a\cdot x + f(a), \quad x\in U \tag{8.5}$$

其中 $a\in\mathbb{R}^n$.

例 8.2(eikonal 方程) 几何光学中的 eikonal 方程是 PDE

$$|Du| = 1 \tag{8.6}$$

它的一个完全积分为

$$u(x;a,b) = a\cdot x + b, \quad x\in U \tag{8.7}$$

其中 $a\in\partial B(0,1), b\in\mathbb{R}$.

例 8.3(Hamilton-Jacobi 方程组) 力学中的 Hamilton-Jacobi 方程组的最简形式为 PDE

$$u_t + H(Du) = 0 \tag{8.8}$$

其中 $H:\mathbb{R}^n\to\mathbb{R}$. 这里 u 依赖于 $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$ 与 $t\in\mathbb{R}$. 我们记 $t=x_{n+1}$, 并取 $Du = D_xu = (u_{x_1},\cdots,u_{x_n})$. 它的一个完全积分为

$$u(x,t;a,b) = a\cdot x - tH(a) + b, \quad x\in\mathbb{R}^n, t\geq 0 \tag{8.9}$$

其中 $a\in\mathbb{R}^n, b\in\mathbb{R}$.

[2]可以参考 **[Zo]** 的第一卷中的函数独立性相关部分.

8.1.2 包络中的新解

下面介绍怎么从非线性一阶 PDE(8.1)中构造更复杂的解, 这些解依赖于任意 $n-1$ 元函数, 且不只有 n 个参数. 我们将把这些新解构造成完全积分的包络, 或更一般地构造成关于另外 m 个参数的解族.

定义 8.1.2 (包络)

设 $u=u(x;a)$ 是关于 $x\in U, a\in A$ 的 C^1 函数, 其中 $U\subset\mathbb{R}^n, A\subset\mathbb{R}^m$ 是开集. 考虑向量方程

$$D_a u(x;a)=0,\quad x\in U, a\in A \tag{8.10}$$

设可从(8.10)式中解出参数 a 关于 x 的 C^1 函数

$$a=\phi(x) \tag{8.11}$$

则

$$D_a u(x;\phi(x))=0,\quad x\in U \tag{8.12}$$

我们就称

$$v(x):=u(x;\phi(x)),\quad x\in U \tag{8.13}$$

为函数族 $\{u(\cdot;a)\}_{a\in A}$ 的包络. ♣

通过构造包络, 就可以从非线性一阶 PDE 中构造出一些新解:

定理 8.1.1 (新解的构造)

设对每个满足定义(8.1.2)的 a, $u=u(\cdot;a)$ 都是 PDE(8.1)的解. 另设由(8.12)式与(8.13)式定义的包络 v 存在且是 C^1 函数, 则 v 也是方程(8.1)的解. ♡

上述定义的包络有时称为(8.1)的一个奇积分[3].

证明

根据包络定义知 $v(x)=u(x,\phi(x))$, 对 $i=1,\cdots,n$, 由(8.12)式知

$$v_{x_i}(x)=u_{x_i}(x;\phi(x))+\sum_{j=1}^{m}u_{a_j}(x,\phi(x))\phi^j_{x_i}(x)=u_{x_i}(x;\phi(x))$$

故任取 $x\in U$ 有

$$F(Dv(x),v(x),x)=F(Du(x,\phi(x)),u(x;\phi(x)),x)=0$$

□

包络的几何意义在于对任意的 $x\in U$, v 的图像总是与 $u(\cdot;a)$ 在 $a=\phi(x)$ 时的图像相切. 进而在 x 点处 $Dv=D_x u(\cdot;a)$, 其中 $a=\phi(x)$.

例 8.4 考虑 PDE

$$u^2(1+|Du|^2)=1 \tag{8.14}$$

它的一个完全积分为

$$u(x,a)=\pm(1-|x-a|^2)^{\frac{1}{2}},\quad |x-a|<1$$

计算知

$$D_a u=\frac{\mp(x-a)}{(1-|x-a|^2)^{\frac{1}{2}}}$$

当 $a=\phi(x)=x$ 时 $D_a u=0$, 代入 $u(x,a)$ 知 $v=\pm1$ 时(8.14)的奇异积分.

[3] 在 ODE 中对应奇解.

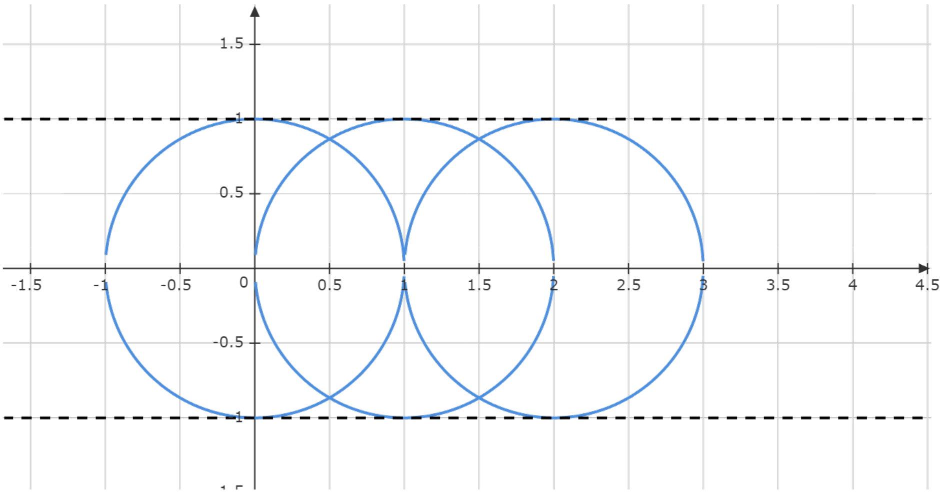


图 8.1: 例(8.4)示意图

要从一个完全积分中生成更多 PDE(8.1)的解, 可以考虑重复上述构造. 任取开集 $A' \subset \mathbb{R}^{n-1}$ 与 C^1 函数 $h: A' \to \mathbb{R}$, 则 h 的图像是在 A 中讨论的. 记

$$a = (a_1, \cdots, a_n) = (a', a_n), \quad a' = (a_1, \cdots, a_{n-1}) \in \mathbb{R}^{n-1}$$

定义 8.1.3 (通积分)

如果函数

$$u'(x; a') = u(x; a', h(a')), \quad x \in U, a' \in A'$$

的包络 $v' = v'(x)$ 存在且是 C^1 的, 那么就称 v' 是 (依赖于 h 的) 通积分[a].

[a]在 ODE 中对应通解.

换句话说, 在计算包络时我们还限定了只对形如 $a = (a', h(a'))$ 的参数 a 进行计算, 其中 h 是一个具体函数. 进而对一个依赖于 n 个任意常数 $a_1, \cdots, a_n$ 的完全积分, (在前面的构造行得通的情况下) 我们可以构造一个关于任意 $n-1$ 元函数 h 的解.

例 8.5 在例(8.3)中设 $H(p) = |p|^2, h \equiv 0$, 则

$$u'(x, t; a) = x \cdot a - t|a|^2$$

考虑

$$D_a u' = x - 2ta = 0 \Rightarrow a = \frac{x}{2t}$$

代入原方程得

$$v'(x, t) = x \cdot \frac{x}{2t} - t|\frac{x}{2t}|^2 = \frac{|x|^2}{4t}, \quad x \in \mathbb{R}^n, t > 0$$

这便是 Hamilton-Jacobi 方程 $v'_t + |Dv'|^2 = 0$ 的一个解.

大胆一点, 我们希望如果找到了 PDE(8.1)的类似上述依赖于任意函数 h 的解, 就得到了(8.1)的所有解. 但事实不一定如此, 比如 PDE 有下述结构

$$F(Du, u, x) = F_1(Du, u, x) F_2(Du, u, x) = 0$$

就算 $u_1(x, a)$ 是 PDE$F_1(Du, u, x) = 0$ 的完全积分, 且也找到了对应于任意函数 h 的通积分, 这个时候我们对 $F_2(Du, u, x) = 0$ 的解依旧是一无所知的.

8.2 特征线

8.2.1 特征 ODE 的推导

回到一开始的非线性一阶 PDE

$$F(Du,u,x)=0, \quad \text{在 } U \text{ 中} \tag{8.15}$$

其边值条件为

$$u=g, \quad \text{在 } \Gamma \text{ 上} \tag{8.16}$$

其中 $\Gamma \subset \partial U$, $g:\Gamma \to \mathbb{R}$ 是给定的. 设 F, g 都是光滑函数.

下面介绍特征线方法, 这种方法通过把 PDE 转化成一个合适的 ODE 系统来求解(8.15),(8.16). 设 u 是(8.15),(8.16)的解, 任意固定 $x \in U$, 我们希望通过找到一些在 U 内连接 $x^0 \in \Gamma$ 与 x 的曲线, 在曲线上求解 u. 既然在 Γ 上 $u=g$, 这条曲线的端点 x^0 处 u 的取值事实上是已知的. 下面希望求解 u 在整条曲线上的值, 进而得到在 x 处的值.

8.2.1.1 确定特征 ODE

怎么确定这样一条在 U 内的曲线呢? 设曲线的参数方程为 $\boldsymbol{x}(s)=(x^1(s),\cdots,x^n(s))$, 参数 $s \in I \subset \mathbb{R}$, I 是某个区间. 设 u 是(8.15)的一个 C^2 解, 定义

$$z(s) := u(\boldsymbol{x}(s)) \tag{8.17}$$

另设

$$\boldsymbol{p}(s) := Du(\boldsymbol{x}(s)) \tag{8.18}$$

也即 $\boldsymbol{p}(s)=(p^1(s),\cdots,p^n(s))$, 其中

$$p^i(s)=u_{x_i}(\boldsymbol{x}(s)), \quad i=1,\cdots,n \tag{8.19}$$

故 $z(\cdot)$ 给出 u 沿曲线的值, $\boldsymbol{p}(\cdot)$ 给出梯度 Du 的值. 现在希望选取 $\boldsymbol{x}(\cdot)$ 以便计算 $z(\cdot), \boldsymbol{p}(\cdot)$.

首先, 在(8.19)式两边微分:

$$\frac{d}{ds}(p^i(s))=\dot{p}^i(s)=\sum_{j=1}^n u_{x_i x_j}(\boldsymbol{x}(s))\dot{x}_j(s) \tag{8.20}$$

但这个表达式不怎么理想, 因为其中包含了 u 的二阶导数, 下面考虑怎么消掉这个二阶导. 在 PDE(8.15)两边对 x_i 微分有:

$$\sum_{j=1}^n F_{p_j}(Du,u,x)u_{x_j x_i}+F_z(Du,u,x)u_{x_i}+F_{x_i}(Du,u,x)=0 \tag{8.21}$$

可以通过(8.21)式消掉(8.20)式中的二阶导数. 首先设

$$\dot{x}^j(s)=F_{p_j}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s)), \quad j=1,\cdots,n \tag{8.22}$$

现在先假设(8.22)式成立, 计算(8.21)式在 $x=\boldsymbol{x}(s)$ 的情况. 把(8.17),(8.18)式代入知

$$\sum_{j=1}^n F_{p_j}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))u_{x_i x_j}(\boldsymbol{x}(s))+F_z(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))p^i(s)+F_{x_i}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))=0 \tag{8.23}$$

将(8.22),(8.23)式代入(8.20)式得

$$\dot{p}^i(s)=\sum_{j=1}^n u_{x_i x_j}(\boldsymbol{x}(s))F_{p_j}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))=-F_z(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))-F_{x_i}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s)), \quad i=1,\cdots,n \tag{8.24}$$

最后对(8.17)式两边微分并代入(8.18),(8.22)式:

$$\dot{z}(s)=\sum_{j=1}^n u_{x_j}(\boldsymbol{x}(s))\dot{x}^j(s)=\sum_{j=1}^n p^j(s)F_{p_j}(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s)). \tag{8.25}$$

8.2.1.2 特征方程

现在把(8.22), (8.23)与(8.25)式整理成向量形式:

$$\begin{cases}\text{(a) } \dot{\boldsymbol{p}}(s)=-D_xF(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))-D_zF(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))\\ \text{(b) } \dot{z}(s)=D_pF(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))\cdot\boldsymbol{p}(s)\\ \text{(c) } \dot{\boldsymbol{x}}(s)=D_pF(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))\end{cases}\tag{8.26}$$

进一步由 PDE 本身知

$$F(\boldsymbol{p}(s),z(s),\boldsymbol{x}(s))\equiv 0\tag{8.27}$$

这些等式对 $s\in I$ 均成立.

$2n+1$ 阶 ODE 系统(8.26)就包含了非线性一阶 PDE(8.15)的特征方程. 函数 $\boldsymbol{p}(\cdot)=(p^1(\cdot),\cdots,p^n(\cdot)),z(\cdot),\boldsymbol{x}(\cdot)=(x^1(\cdot),\cdots,x^n(\cdot))$ 称作特征. 有时把 $\boldsymbol{x}(\cdot)$ 称作投影特征, 它是全特征 $(\boldsymbol{p}(\cdot),z(\cdot),\boldsymbol{x}(\cdot))\subset\mathbb{R}^{2n+1}$ 在区域 $U\subset\mathbb{R}^n$ 上的投影.

前面已经证明了:

定理 8.2.1 (特征 ODE 的结构)

设 $u\in C^2(U)$ 是 U 上的非线性一阶 PDE(8.15)的解. 设 $\boldsymbol{x}(\cdot)$ 是 ODE(8.26)(c) 的解, $\boldsymbol{p}(\cdot)=Du(\boldsymbol{x}(\cdot)),z(\cdot)=u(\boldsymbol{x}(\cdot))$. 则对满足 $\boldsymbol{x}(s)\in U$ 的 s, $\boldsymbol{p}(\cdot)$ 是 ODE(8.26)(a) 的解, $z(\cdot)$ 是 ODE(8.26)(b) 的解. ♡

我们还需要系统(8.26)的初值来让上述定理真正发挥作用, 这就留到后面讨论.

注 特征 ODE 是很有用处的, 它对作为一般非线性 PDE(8.15)的光滑解 u 给出了关于 $\boldsymbol{x}(\cdot),z(\cdot)=u(\boldsymbol{x}(\cdot))$ 与 $p(\cdot)=DDu(\boldsymbol{x}(\cdot))$ 满足的一个确切的方程组. 推导中的关键步骤在于设 $\dot{\boldsymbol{x}}=D_pF$, 进而就如上面的过程所述, 表达式中的二阶导就被消掉了. 这便得到了某种封闭性, 不用引入关于 u 的二阶及更高阶导数的 ODE 了.

8.2.2 例子

在继续对特征方程(8.26)的讨论前, 先歇一歇看看一些简单结构方程的具体情况. 顺便我们也介绍一些情况, 其中给定边值条件, 可以通过解特征 ODE 得到特定一阶 PDE 的解.

8.2.2.1 F 是线性的

首先考虑 PDE(8.15)是线性齐次的, 其形如

$$F(Du,u,x)=\boldsymbol{b}(x)\cdot Du(x)+c(x)u(x)=0,\quad x\in U\tag{8.28}$$

进而 $F(p,z,x)=\boldsymbol{b}(x)\cdot p+c(x)z$, 因而

$$D_pF=\boldsymbol{b}(x)$$

在现在的情况下(8.26)(c) 变为

$$\dot{\boldsymbol{x}}(s)=\boldsymbol{b}(\boldsymbol{x}(s))\tag{8.29}$$

这是一个只包含 $\boldsymbol{x}(\cdot)$ 的 ODE. 进一步(8.26)(b) 变为

$$\dot{z}(s)=\boldsymbol{b}(\boldsymbol{x}(s))\cdot\boldsymbol{p}(s)\tag{8.30}$$

再根据(8.28)式, 把 $Du(x),u(x)$ 分别换成 $\boldsymbol{p},z$, 代入(8.30)式得到

$$\dot{z}(s)=-c(\boldsymbol{x}(s))z(s)\tag{8.31}$$

只要从(8.29)中解出 $\boldsymbol{x}(\cdot)$, 上述 ODE 对 $z(\cdot)$ 就是线性的了. 总之

$$\begin{cases} \text{(a) } \dot{\boldsymbol{x}}(s) = \boldsymbol{b}(\boldsymbol{x}(s)) \\ \text{(b) } \dot{z}(s) = -c(\boldsymbol{x}(s))z(s) \end{cases} \tag{8.32}$$

包含了线性一阶 PDE(8.28)的特征方程. (后面会看到 $\boldsymbol{p}(\cdot)$ 的方程实际上是多余的.) □

第九章 解的其它表示方法

本章收录诸多求得不同 PDE 精确解的技巧, 或给出解的表示公式.

9.1 分离变量法

分离变量法指对给定的 PDE 构造一个能表成更少变量的函数的组合的解. 也就是说, 其思想在于首先猜测 u 能表成诸如一些未知函数的和或积, 把这个猜测的表达式代回 PDE, 并选定一些简单函数来验证 u 确实是 PDE 的解.

9.1.1 例子

下面举一些例子来更好地理解分离变量法.

例 9.1 设 $U \subset \mathbb{R}^n$ 是有界开集, 其边界光滑. 考虑热方程的初边值问题:

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } U \times (0, \infty) \text{ 中} \\ u = 0, & \text{在 } \partial U \times [0, \infty) \text{ 上} \\ u = g, & \text{在 } U \times \{t = 0\} \text{ 上} \end{cases} \tag{9.1}$$

其中 $g : U \to \mathbb{R}$ 给定. 我们猜想存在一个形如

$$u(x,t) = v(t)w(x), \quad x \in U, t \geq 0 \tag{9.2}$$

的解. 也即, 我们希望找到一个形如(9.2)的解, 其中变量 $x = (x_1, \cdots, x_n) \in U$ 与 $t \in [0, T]$" 分开" 了.

这件事可行吗? 要得到回答, 下面计算

$$u_t(x,t) = v'(t)w(x), \quad \Delta u(x,t) = v(t)\Delta w(x)$$

故

$$0 = u_t(x,t) - \Delta u(x,t) = v'(t)w(x) - v(t)\Delta w(x)$$

上式成立当且仅当

$$\frac{v'(t)}{v(t)} = \frac{\Delta w(x)}{w(x)} \tag{9.3}$$

对任意满足 $w(x) \neq 0, v(t) \neq 0$ 的 $x \in U, t > 0$ 都成立. 观察知(9.3)式左式只与 t 有关, 右式只与 x 有关, 故(9.3)式的两个式子只能等于某个常数. 设

$$\frac{v'(t)}{v(t)} = \mu = \frac{\Delta w(x)}{w(x)}, \quad t \geq 0, x \in U$$

则

$$v' = \mu v \tag{9.4}$$

$$\Delta w = \mu w \tag{9.5}$$

现在求解关于 w, v, μ 的方程.

首先注意如果 μ 已知, 方程(9.4)的解就是 $v = de^{\mu t}$, 其中 d 是任意常数. 故只需讨论方程(9.5).

称 λ 是算子 $-\Delta$ 在 U 上 (在零边值条件下) 的特征值, 如果存在不恒等于零的函数 w, 满足

$$\begin{cases} -\Delta w = \lambda w, & \text{在 } U \text{ 内} \\ w = 0, & \text{在 } \partial U \text{ 上} \end{cases}$$

函数 w 称为特征函数.

若 λ 是特征值, w 是相应的特征函数, 则取 $\mu = -\lambda$ 即得

$$u = de^{-\lambda t} w \tag{9.6}$$

正是方程

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } U \times (0, \infty) \\ u = 0, & \text{在 } \partial U \times [0, \infty) \text{ 上} \end{cases} \tag{9.7}$$

在初值条件 $u(\cdot, 0) = dw$ 下的解. 故只要 $g = dw$, 由(9.6)式定义的函数 u 就是初边值问题(9.1)的解. 更一般地, 如果 $\lambda_1, \cdots, \lambda_m$ 是特征值, $w_1, \cdots, w_m$ 是对应的特征函数, $d_1, \cdots, d_m$ 是常数, 则

$$u = \sum_{k=1}^{n} d_k e^{-\lambda_k t} w_k \tag{9.8}$$

就是(9.7)在初值条件 $u(\cdot, 0) = \sum\limits_{k=1}^{m} d_k w_k$ 下的解. 现在只需找到 $m, w_1, \cdots$ 等等, 使之满足 $\sum\limits_{k=1}^{m} d_k w_k = g$ 即可.

进一步, 可设存在可数个特征值 $\lambda_1, \cdots$, 对应的特征函数为 $w_1, \cdots$, 进而对适当的常数 $d_1, \cdots$, 在 U 内有

$$\sum_{k=1}^{\infty} d_k w_k = g \tag{9.9}$$

进而

$$u = \sum_{k=1}^{\infty} d_k e^{-\lambda_k t} w_k \tag{9.10}$$

就很可能是初值问题(9.1)的解.

这已经是解的一个很好的表示式了, 但它取决于: (1) 是否能找到特征值, 特征函数与满足(9.9)式的常数; (2)(9.10)式中的级数在适当的意义下收敛. 之后会在 Galerkin 逼近的部分着重探讨这些内容. □

注意只有(9.6)式是由分离变量法得到的. (9.8),(9.10)式都是由热方程的线性性而来的.

例 9.2 下面把分离变量法应用到多孔介质方程

$$u_t - \Delta(u^\gamma) = 0, \quad \text{在 } \mathbb{R}^n \times (0, \infty) \text{ 中} \tag{9.11}$$

的求解中, 其中 $u \geq 0, \gamma > 1$ 是常数. 方程(9.11)是非线性扩散方程, 其中密度 u 的扩散速率与 u 自身相关. 这个 PDE 表示的是多孔介质中的流, 薄膜润滑与一些其它的现象.

正如前例, 考虑寻找形如

$$u(x, t) = v(t) w(x), \quad w \in \mathbb{R}^n, t \geq 0 \tag{9.12}$$

的解. 代入(9.11)有

$$\frac{v'(t)}{v(t)^\gamma} = \mu = \frac{\Delta w^\gamma(x)}{w(x)} \tag{9.13}$$

其中 μ 为常数, 式子对任意满足 $w(x), v(t) \neq 0$ 的 $x \in \mathbb{R}^n, t \geq 0$ 都成立.

求解关于 v 的 ODE 可得

$$v = ((1 - \gamma)\mu t + \lambda)^{\frac{1}{1-\gamma}}$$

其中 λ 为常数, 不妨设之为正数. 要求解 w, 即考虑求解 PDE

$$\Delta(w^\gamma) = \mu w \tag{9.14}$$

下面猜测

$$w = |x|^\alpha$$

其中 α 是需要确定的常数. 代入有

$$\mu w - \Delta(w^\gamma) = \mu |x|^\alpha - \alpha\gamma(\alpha\gamma + n - 2)|x|^{\alpha\gamma - 2} \tag{9.15}$$

因为(9.14)式在 $\mathbb{R}^n$ 中总成立, 首先便需要 $\alpha = \alpha\gamma - 2$, 解得

$$\alpha = \frac{2}{\gamma - 1} \tag{9.16}$$

代入(9.15)式得

$$\mu = \alpha\gamma(\alpha\gamma + n - 2) > 0 \tag{9.17}$$

总之, 对每个 $\lambda > 0$, 函数

$$u = ((1-\gamma)\mu t + \lambda)^{\frac{1}{1-\gamma}}|x|^{\alpha}$$

即为多孔介质方程(9.11)的解, 其中参数 α, μ 由(9.16),(9.17)式定义. □

注 观察知由于 $\gamma > 1$, 解在 $x \neq 0$ 时会在 $t \to t_* := \frac{\lambda}{(\gamma-1)\mu}$ 处爆破. 在物理意义上, 这里对应大量物质在有限时间内从无穷远处扩散进来.

课堂笔记 (2023.11.20)

- (关于热方程分离变量法思路的重述) 分离变量法中 $u(x,t) = v(t)w(x)$ 的形式比 $u(x,t) = v(t) + w(x)$ 的形式要更常用一些, 这是因为后者把 x 和 t 完全分开了. 先前对热方程的讨论大多是针对解的唯一性的, 而分离变量法更加着重于解的存在性. 对于 $u(x,t) = v(t)w(x)$ 的形式而言, 如果存在 $(x,t) \in U \times (0,\infty)$ 使得 $v(t) = 0$ 或 $w(x) = 0$, 那么讨论就是平凡的了, 所以在讨论过程中只需要研究对任意 $(x,t) \in U \times (0,\infty)$ 均有 $v(t)w(x) \neq 0$ 的情况. 根据前文, 在热方程中, 如果令 $u(x,t) = v(t)w(x)$, 设
$$\frac{v'(t)}{v(t)} = \frac{\Delta w(x)}{w(x)} = -\lambda, \quad \lambda \geq 0 \text{为常数}, \forall t \geq 0, \forall x \in U$$
则可解得 $v(t) = de^{-\lambda t}$, 其中 d 为常数. 如果能找到一组 w^k 满足
$$\begin{cases} -\Delta w^k(x) = \lambda w^k(x), & \text{在 } U \text{ 内} \\ w^k(x) = 0, & \text{在 } \partial U \text{ 上} \end{cases}$$
则 $u^k(x,t) = d_k e^{-\lambda t} w^k(x)$ 就满足初值问题
$$\begin{cases} u_t^k(x,t) - \Delta u^k(x,t) = 0, & \text{在 } U \text{ 内} \\ u^k(x,t) = d_k w^k(x), & \text{在 } \partial U \text{ 上} \end{cases}$$
现在希望求解初值问题
$$\begin{cases} \tilde{u}_t(x,t) - \Delta \tilde{u}(x,t) = 0, & \text{在 } U \text{ 内} \\ \tilde{u}(x,t) = g(x), & \text{在 } \partial U \text{ 上} \end{cases}$$
如果可以找到一组常数 $\{d_k\}_{k=1}^m$ 使得
$$g(x) = \sum_{k=1}^m d_k w^k(x)$$
则理应有
$$\tilde{u}(x,t) = \sum_{k=1}^m u^k(x,t) = \sum_{k=1}^m d_k e^{-\lambda t} w^k(x)$$
故问题转化为求解 $w^k(x)$, 这需要归结到算子 $-\Delta$ 的谱性质. 这样一套把 PDE 的初值问题转化成谱论问题的方法就称为 Galerkin 方法.

在前面的例子中, 分离变量法是通过非线性方程的齐次性起作用的, 这使得具有乘积形式(9.12)的函数能够被兼容. 在其它情况中, 可寻找被加法分离的解.

例 9.3 下面来研究 Hamilton-Jacobi 方程

$$u_t + H(Du) = 0, \quad 在\ \mathbb{R}^n \times (0,\infty)\ 中 \tag{9.18}$$

考虑寻找形如

$$u(x,t) = w(x) + v(t), \quad x \in \mathbb{R}^n, t \geq 0$$

的解 u. 代入知

$$0 = u_t(x,t) + H(Du(x,t)) = v'(t) + H(Dw(x))$$

上式成立当且仅当

$$H(Dw(x)) = \mu = -v'(t), \quad x \in \mathbb{R}^n, t > 0$$

其中 μ 是常数. 进而若

$$H(Dw) = \mu, \quad \mu \in \mathbb{R}$$

则

$$u(x,t) = w(x) - \mu t + b$$

就是 $u_t + H(Du) = 0$ 的解, 其中 b 是任意常数. 特别地, 若取 $w(x) = a \cdot x, a \in \mathbb{R}^n$, 并取 $\mu = H(a)$, 即得解

$$u = a \cdot x - H(a)t + b.$$

□

9.1.2 应用: Turing 不稳定性

在例(9.1)中讨论过的分离变量法与特征函数展开在纯数学与应用数学中都是非常强大的工具. 本节讨论其应用.

给定 $\mathbb{R}^2$ 上的光滑向量场 $\boldsymbol{f} = (f^1, f^2)$, 其中 0 为平衡点:

$$\boldsymbol{f}(0) = 0$$

现在希望比较 ODE

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \tag{9.19}$$

的解 $\boldsymbol{x}(x^1, x^2)$ 与其所对应的反应扩散 PDE

$$\begin{cases} \boldsymbol{u}_t - A\Delta \boldsymbol{u} = \boldsymbol{f}(\boldsymbol{u}), & 在\ U \times (0,\infty)\ 中 \\ \boldsymbol{u} = 0, & 在\ \partial U \times (0,\infty)\ 上 \end{cases} \tag{9.20}$$

的解 $\boldsymbol{u} = (u^1, u^2)$ 两者的稳定性, 其中 $U\mathbb{R}^2$ 是有界光滑区域. 矩阵

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$

记录了扩散系数 $a_1, a_2 \geq 0$.

9.1.2.1 线性化, 分离变量

(9.19)式在平衡点 $\boldsymbol{x} = 0$ 附近的线性化为 ODE 系统

$$\dot{\boldsymbol{y}} = D\boldsymbol{f}(0)\boldsymbol{y}, \quad t \geq 0 \tag{9.21}$$

其中 $\boldsymbol{y} = (y^1, y^2)$. 平衡点 $\boldsymbol{x} \equiv 0$ 在 $\boldsymbol{y} \to 0 (t \to \infty)$ 时是渐近稳定的. 这说明矩阵 $D\boldsymbol{f}(0)$ 的特征值有负实部.

类似地, (9.20)在 $\boldsymbol{u} \equiv 0$ 附近的线性化为线性 PDE 系统:

$$\boldsymbol{v}_t - A\Delta \boldsymbol{v} = D\boldsymbol{f}(0)\boldsymbol{v} \tag{9.22}$$

其中 $\boldsymbol{v} = (v^1, v^2)$. 下面用前面介绍过的分离变量法与特征函数展开法求解(9.22). 设 $\boldsymbol{v}(x,t) = \boldsymbol{s}(t)w(x)$[1], 代入(9.22)有

$$\boldsymbol{s}'(t)w(x) - A\boldsymbol{s}(t)\Delta w(x) = D\boldsymbol{f}(0)\boldsymbol{s}(t)w(x) \tag{9.23}$$

现若 $\{w_j\}_{j=1}^{\infty}$ 是 U 上的 Laplace 算子在零边界条件:

$$\begin{cases} -\Delta w_j = \lambda_j w_j, & \text{在 } U \text{ 中} \\ w_j = 0, & \text{在 } \partial U \text{ 上} \end{cases}$$

下的特征函数组, 代入(9.23)式得到

$$\boldsymbol{s}'(t)w_j(x) + A\lambda_j \boldsymbol{s}(t)w_j(x) = D\boldsymbol{f}(0)\boldsymbol{s}(t)w_j(x), \quad j = 1, 2, \cdots$$

当 $w_j(x) \neq 0$ 时, 两边约去 $w_j(x)$ 得 ODE:

$$\boldsymbol{s}'(t) + A\lambda_j \boldsymbol{s}(t) = D\boldsymbol{f}(0)\boldsymbol{s}(t) \tag{9.24}$$

对每一个 j, ODE(9.24)都可解出对应的 $\boldsymbol{s}_j(t)$, 故

$$\boldsymbol{v}(x,t) = \sum_{j=1}^{\infty} \boldsymbol{s}_j(t)w_j(x) \tag{9.25}$$

对这里特征值与特征函数的具体讨论会在之后进行, 此刻我们只需要知道有结论:

$$\lambda_j > 0, \quad j = 1, 2, \cdots$$

因而可令特征函数组 $\{w_j\}_{j=1}^{\infty}$ 在 $L^2(U)$ 中正交[2]:

$$\int_U w_i w_j dx = \delta_{ij}, \quad i, j = 1, 2, \cdots$$

将(9.25)式代回(9.22)式得

$$\sum_{j=1}^{\infty} \boldsymbol{s}_j'(t)w_j(x) - A\sum_{j=1}^{\infty} \boldsymbol{s}_j(t)\lambda_j w_j(x) = D\boldsymbol{f}(0)\sum_{j=1}^{\infty} \boldsymbol{s}_j(t)w_j(x)$$

根据 $\{w_j\}_{j=1}^{\infty}$ 的正交性可知

$$\boldsymbol{s}_j' = A_j \boldsymbol{s}_j \tag{9.26}$$

其中

$$A_j := D\boldsymbol{f}(0) - \lambda_j A \tag{9.27}$$

解 $\boldsymbol{v} \equiv 0$ 是稳定解当且仅当每个 $\boldsymbol{s}_j$ 在 $t \to \infty$ 时都衰减到 0, 这需要矩阵 A_j 的特征值均有负实部 ($j = 1, 2, \cdots$).

现在考虑下述问题: 如果 0 是 ODE 系统(9.19)的渐近稳定平衡点, 这是否能推出 0 也是 PDE 系统(9.20)的渐近稳定平衡点? 答案可能出乎意料: 并非如此. PDE(9.20)中的扩散项事实上把(9.19)中的稳定点变成了(9.20)中的不稳定点. 这个现象就称作 Turing 不稳定性.

9.1.2.2 $D\boldsymbol{f}(0)$ 的特征值

在讨论上面这种现象时, 首先需要 $D\boldsymbol{f}(0)$ 添加一些条件, 使得 0 确实是 ODE(9.19)的一个稳定点. 记

$$D\boldsymbol{f}(0) = \begin{pmatrix} f^1_{z_1}(0) & f^1_{z_2}(0) \\ f^2_{z_1}(0) & f^2_{z_2}(0) \end{pmatrix} =: \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

进而

$$\det(D\boldsymbol{f}(0) - \sigma I) = \sigma^2 - \sigma(\alpha + \delta) + \alpha\delta - \gamma\beta \tag{9.28}$$

[1]这样设的原因在于从前面的例子中可以看出, 要讨论 Laplace 算子的特征值, 最后回到的是一个形如 $\Delta w = \lambda w$ 的方程, 而这个方程的解是实值函数, 故设 $w(x)$ 为一维.

[2]为什么?

我们希望 0 是 $Df(0)$ 的一个稳定平衡点, 故需要方程(9.28)的根 σ_1, σ_2 有负实部, 即

$$\begin{cases} \alpha + \delta = \sigma_1 + \sigma_2 < 0 \\ \alpha\delta - \gamma\beta = \sigma_1\sigma_2 > 0 \end{cases} \tag{9.29}$$

注意上述条件同时包含了实特征值 $\sigma_1 \leq 0, \sigma_2 < 0$ 与复共轭特征值 $\sigma_1 = b + ic, \sigma_2 = b - ic(b < 0)$ 这两种情况.

9.1.2.3 A_j 的特征值

下面说明在加入扩散项 $a_1, a_2 \geq 0$ 后, 是否可以让 A_j 的特征值具有正实部.

由(9.27)式可知

$$\det\left(A_j - \sigma I\right) = \sigma^2 - \sigma(\alpha + \delta - \lambda_j(a_1 + a_2)) + p(\lambda_j) \tag{9.30}$$

其中

$$p(\lambda) := \lambda^2 a_1 a_2 - \lambda(a_1\delta + a_2\alpha) + \alpha\delta - \beta\gamma \tag{9.31}$$

由(9.29)中第一式与 $\lambda_j \geq 0, a_1, a_2 > 0$ 知, 多项式(9.30)的根 $\sigma_{1,j}, \sigma_{2,j}$ 应满足:

$$\sigma_{1,j} + \sigma_{2,j} = \alpha + \delta - \lambda_j(a_1 + a_2) < 0$$

故当 $\sigma_{1,j} = b_j + ic_j, \sigma_{2,j} = b_j - ic_j$ 是复共轭根时, 实部 b_j 必是负的. 此时 ODE(9.26)的解 s_j 在 $t \to \infty$ 时趋零, 因而得到渐近稳定性, 但这不是我们希望研究的 Turing 不稳定的情况.

9.1.2.4 稳定性的丢失

从上小节可知, 要想让 PDE 系统(9.22)丢失稳定性, 只能让 $\sigma_{2,j} \leq \sigma_{1,j}$ 是实根. 现在希望选取 $a_1, a_2 \geq 0$ 使得 $\sigma_{1,j} \geq 0$. 考虑先取 $a_1 = a_2 = 0$, 再逐步增长这两个扩散系数, 直到系统(9.22)首次在 $\sigma_{1,j} = 0$ 时丧失稳定性. 要使得 $\sigma_{1,j} = 0$, 需要

$$p(\lambda_j) = 0 \tag{9.32}$$

现在求解使(9.32)式成立的代数条件. 不失一般性设

$$\delta < 0 \tag{9.33}$$

进而若 $a_2 = 0$, 则由(9.29)式有

$$p(\lambda_j) = -\lambda_j a_1 \delta + \alpha\delta - \beta\gamma > 0$$

故必须要求 $a_2 > 0$, 也就是说, 要使得 PDE 系统(9.22)的 Turing 不稳定性发生, 我们必须在其第二式 (即涉及 a_2 的方程) 中引入扩散.

而若 $\alpha \leq 0$, 则由(9.29)式有

$$p(\lambda_j) = \lambda_j^2 a_1 a_2 - \lambda_j(a_1\delta + a_2\alpha) + (\alpha\delta - \beta\gamma) > 0$$

进而(9.32)式没法成立. 这说明只能有

$$\alpha > 0 \tag{9.34}$$

现在在(9.33),(9.34)式的条件下把 $p(\lambda_j) = 0$ 写成

$$a_2 = \frac{\alpha\delta - \gamma\beta - \delta\lambda_j a_1}{\lambda_j(\alpha - \lambda_j a_1)} \tag{9.35}$$

现在只要 $\lambda_j > 0$, 就很容易找到 $a_2 > 0, a_1 \geq 0$ 满足(9.35). 特别注意当 λ_j 远大于零时, 需要取 a_1 足够小以保证 $\alpha - \lambda_j a_1 > 0$.

9.1.2.5 现象的解释: 激发子与抑制子

符号条件

$$\begin{cases}\alpha = f^1_{z_1}(0) > 0, \beta = f^1_{z_2}(0) < 0\\ \gamma = f^2_{z_1}(0) > 0, \delta = f^2_{z_2}(0) < 0\end{cases}$$

是由(9.29),(9.33)与(9.34)式确定的. 进而可把 u^1 看成某种化学激发子的浓度, u^2 看成某种抑制子的浓度: 因为 $\alpha > 0$, 激发子本身是递增的, 但又因为 $\beta < 0$, 这个增量就可以被抑制子抵消. γ, δ 的正负性表示抑制子的增长只与激发子的量有关. (9.29)式表明对 ODE(9.19)而言, 激发子与抑制子的平衡至少在原点附近是可以维持的[3].

前面已经发现当 a_2 充分大, a_1 充分小时, 扩散系数是可以打破平衡的. 物理上的解释是: 抑制子 u^2 在任意给定点比激发子 u^1 扩散得都快, 进而在一点处抑制子的量不足以抵消激发子的量. 这样的反应扩散不平衡性有时能推广成生物模型.

9.2 自相似解

在研究 PDE 时, 通常都会找这样一些特殊解 u, 它们的结构能反映 PDE 结构的各种对称性. 这种思想在 Laplace 方程与热方程的基本解推导中已经出现过了, 下面介绍这种方法的更多应用.

9.2.1 平面波与行波, 孤波

首先考虑含有两个变量 $x \in \mathbb{R}, t \in \mathbb{R}$ 的 PDE, 形如

$$u(x,t) = v(x - \sigma t), \quad x \in \mathbb{R}, t \in \mathbb{R} \tag{9.36}$$

的解 u 称为 (速度为 σ, 波形为 v 的) 行波. 更一般地, 含有 $n+1$ 个变量 $x = (x_1, \cdots, x_n) \in \mathbb{R}^n, t \in \mathbb{R}$ 的 PDE 的解如果形如

$$u(x,t) = v(y \cdot x - \sigma t) \tag{9.37}$$

就称该解为 (波阵面与 $y \in \mathbb{R}^n$ 正交, 速度为 $\frac{\sigma}{|y|}$, 波形为 v 的) 平面波.

9.2.1.1 指数解

在 Fourier 变换的视角下, 研究线性 PDE 时考虑形如

$$u(x,t) = e^{i(y \cdot x - \sigma t)} \tag{9.38}$$

的复平面波解是很有启发意义的, 其中 $\sigma \in \mathbb{C}, y = (y_1, \cdots, y_n) \in \mathbb{R}^n$. σ 表示时频, $\{y_i\}_{i=1}^n$ 表示波数. 下面就用形如(9.38)的解去尝试解一些线性 PDE, 在此期间特别注意 y 与 $\sigma = \sigma(y)$ 是怎么受方程结构的影响的.

热方程

若 u 由(9.38)式给出, 代入热方程有

$$u_t - \Delta u = (-i\sigma + |y|^2)u = 0$$

解得 $\sigma = -i|y|^2$, 故

$$u = e^{iy \cdot x - |y|^2 t}$$

在任取 $y \in \mathbb{R}^n$ 时正是热方程的解. 将指数形式写成实复部形式后可进一步发现 $e^{-|y|^2 t}\cos(y \cdot x)$ 与 $e^{-|y|^2 t}\sin(y \cdot x)$ 同样是热方程的解. 注意在这个例子中, 因为 σ 是纯虚数, 故解得表达式中有一个实的负指数项 $e^{-|y|^2 t}$, 这个项对

[3]这一整段还是缺乏物理直觉...

应抑制项或耗散项.

波方程, Klein-Gordon 方程

将(9.38)式代入波方程得到

$$u_{tt} - \Delta u = (-\sigma^2 + |y|^2)u = 0$$

解得 $\sigma = \pm|y|$. 故

$$u = e^{i(y \cdot x \pm |y|t)}$$

就是波方程的解, 同上可得 $\cos(y \cdot x \pm |y|t)$ 与 $\sin(y \cdot x \pm |y|t)$ 也是波方程的解. 因为 σ 是实的, 这些解中就没有耗散效应, 且每个解的传播速度的绝对值都为 $\frac{|\sigma|}{|y|} = 1$.

下面研究 Klein-Gordon 方程:

$$u_{tt} - \Delta u + m^2 u = 0$$

代入(9.38)式得到

$$u_{tt} - \Delta u + m^2 u = (-\sigma^2 + |y|^2 + m^2)u = 0$$

解得 $\sigma = \pm(|y|^2 + m^2)^{\frac{1}{2}}$. 注意此时传播速度 $\frac{\sigma}{|y|}$ 非线性地依赖于初值 $e^{iy \cdot x}$ 的频率, 振荡越慢的解传播越快. 这种不同频率的波以不同的速度传播的现象说明 Klein-Gordon 方程将造成色散.

课堂笔记 (2023.11.20)

- (关于平面波的物理直觉) 注意平面波的定义中 v 是 $\mathbb{R}$ 到 $\mathbb{R}$ 的函数, 从而如果能将 $u(x,t)$ 写成 $v(y \cdot x - \sigma t)$ 的形式, 则只要 (x,t) 在超平面 $y \cdot x - \sigma t = c$ 上, u 就是一个常值, 从而 u 得名”平面波”.
- (关于热方程平面波法的一些补充) 如果仅求出 $u(x,t) = e^{i\xi \cdot x - |y|^2 t}$, 知 $u(x,t)$ 满足的初值问题为:

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u(x,t) = e^{i\xi \cdot x}, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

现在考虑初值问题

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u(x,t) = g(x), & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

根据叠加原理, 如果对 $g(x)$ 有表达式

$$g(x) = \sum_{k \in \mathbb{Z}} c_k e^{i\xi_k \cdot x}$$

设 $u^k(x,t)$ 满足的初值问题为

$$\begin{cases} u_t^k(x,t) - \Delta u^k(x,t) = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u^k(x,t) = e^{i\xi_k \cdot x}, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \tag{9.39}$$

则根据叠加原理应有

$$u(x,t) = \sum_{k \in \mathbb{Z}} c_k u^k(x,t)$$

事实上在连续的意义下, $g(x)$ 欲求的表达式正是其 Fourier 反演:

$$g(x) = \int_{\mathbb{R}^n} e^{i\xi \cdot x} \widehat{g}(\xi) d\xi$$

而根据正文的讨论, 初值问题(9.39)有解

$$u^k(x,t) = e^{i\xi_k \cdot x - |\xi|^2 t}$$

因而通过叠加原理得到的解应为

$$u(x,t)=\int_{\mathbb{R}^n}e^{i\xi\cdot x-|\xi|^2t}\widehat{g}(\xi)d\xi$$

在学习 Fourier 变换后, 可以验证

$$\int_{\mathbb{R}^n}e^{i\xi\cdot x-|\xi|^2t}\widehat{g}(\xi)d\xi=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}}g(y)dy$$

- (关于波方程平面波法的补充与一些符号上的约定) 根据正文讨论, 已经知道 $e^{i(\xi\cdot x+|\xi|t)}, e^{i(\xi\cdot x-|\xi|t)}$ 均满足波方程

$$u_{tt}(x,t)-\Delta u(x,t)=0$$

故 $u(x,t)=ae^{i(\xi\cdot x+|\xi|t)}+be^{i(\xi\cdot x-|\xi|t)}$ 也满足波方程, 其中 a,b 与 x,t 无关. 该形式满足的初值问题为

$$\begin{cases}u_{tt}(x,t)-\Delta u(x,t)=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 内}\\ u(x,t)=ae^{i\xi\cdot x}+be^{i\xi\cdot x}, u_t(x,t)=a|\xi|e^{i\xi\cdot x}-b|\xi|e^{i\xi\cdot x}, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

现在希望研究初值问题

$$\begin{cases}w_{tt}(x,t)-\Delta w(x,t)=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 内}\\ w(x,t)=g(x), w_t(x,t)=h(x), & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

于是为了应用叠加原理, 希望把 $g(x),h(x)$ 按 $\{e^{i\xi\cdot x}\}_{\xi\in\mathbb{R}^n}$ 的基底展开, 这便再一次回到 Fourier 变换, 根据 Fourier 反演公式有

$$g(x)=\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{g}(\xi)d\xi$$

$$h(x)=\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{h}(\xi)d\xi$$

于是要想利用平面波法, 就需要

$$\begin{cases}ae^{i\xi\cdot x}+be^{i\xi\cdot x}=e^{i\xi\cdot x}\widehat{g}(\xi)\\ a|\xi|e^{i\xi\cdot x}-b|\xi|e^{i\xi\cdot x}=e^{i\xi\cdot x}\widehat{h}(\xi)\end{cases}\Rightarrow\begin{cases}a=\frac{1}{2}(\widehat{g}(\xi)+\frac{\widehat{h}(\xi)}{i|\xi|})\\ b=\frac{1}{2}(\widehat{g}(\xi)-\frac{\widehat{h}(\xi)}{i|\xi|})\end{cases}$$

现在

$$u(x,t)=\frac{1}{2}(\widehat{g}(\xi)+\frac{\widehat{h}(\xi)}{i|\xi|})e^{i(\xi\cdot x+|\xi|t)}+\frac{1}{2}(\widehat{g}(\xi)-\frac{\widehat{h}(\xi)}{i|\xi|})e^{i(\xi\cdot x-|\xi|t)}$$

正是初值问题

$$\begin{cases}u_{tt}(x,t)-\Delta u(x,t)=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 内}\\ u(x,t)=e^{i\xi\cdot x}\widehat{g}(\xi), u_t(x,t)=e^{i\xi\cdot x}\widehat{h}(\xi), & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

的解, 自然应有

$$\begin{aligned}w(x,t)&=\int_{\mathbb{R}^n}u(x,t)d\xi\\&=\int_{\mathbb{R}^n}\frac{1}{2}(\widehat{g}(\xi)+\frac{\widehat{h}(\xi)}{i|\xi|})e^{i(\xi\cdot x+|\xi|t)}+\frac{1}{2}(\widehat{g}(\xi)-\frac{\widehat{h}(\xi)}{i|\xi|})e^{i(\xi\cdot x-|\xi|t)}d\xi\\&=\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{g}(\xi)\frac{e^{i|\xi|t}+e^{-i|\xi|t}}{2}d\xi+\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{h}(\xi)\frac{e^{i|\xi|t}-e^{-i|\xi|t}}{2i\xi}d\xi\\&=\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{g}(\xi)\cos(|\xi|t)d\xi+\int_{\mathbb{R}^n}e^{i\xi\cdot x}\widehat{h}(\xi)\frac{\sin(|\xi|t)}{|\xi|}d\xi\end{aligned}$$

[HJR] 中介绍了这样的记号[a]: 如果 f 是 $\mathbb{R}^d$ 上的可测函数, 在无穷远处以至多多项式速度增长, 则算子 $f(D)$ 定义为 $f(D)u=\mathcal{F}^{-1}(f\cdot\mathcal{F}u)$. 注意上式可以写成

$$w(x,t)=\mathcal{F}^{-1}(\cos(|\cdot|t)\mathcal{F}g)+\mathcal{F}^{-1}(\frac{\sin(|\cdot|t)}{|\cdot|}\mathcal{F}h)$$

于是波方程的解有记法:

$$w(x,t)=\cos(|\nabla|t)\cdot g(x)+\frac{\sin(|\nabla|t)}{|\nabla|}\cdot h(x) \tag{9.40}$$

[a]这个记号定义其实并不那么清楚, 不过在这一条注记中只需要停止在此即可. 在 Fourier 变换的部分会着重讲述这样的算子表示法.

其余色散方程

将 $u=e^{i(y\cdot x-\sigma t)}$ 代入 Schrödinger 方程

$$iu_t+\Delta u=0$$

得到

$$iu_t+\Delta u=(\sigma-|y|^2)u=0$$

其中 $\sigma=|y|^2$. 故

$$u=e^{i(y\cdot x-|y|^2t)}$$

同样, 这个解也表现出色散.

最后再举一个色散 PDE 的例子, 令 $n=1$, 将 $u=e^{i(yx-\sigma t)}$ 代入 Airy 方程

$$u_t+u_{xxx}=0$$

得

$$u_t+u_{xxx}=-i(\sigma+y^3)u=0$$

解得 $\sigma=-y^3$.

像与群速度

对一般的常系数色散线性 PDE, 都能类似于上述方法解出 $\sigma=\sigma(y)$. 有时把 $\frac{\sigma(y)}{|y|}$ 称作指数平面波解(9.38)的像速度: 这是单位向量 $\frac{y}{|y|}$ 方向上的传播速度.

但之后可以看到我们可以用 Fourier 变换来写出 PDE 的更多通解, 它们均形如指数平面波解:

$$u(x,t)=\int_{\mathbb{R}^n}e^{i(y\cdot x-\sigma(y)t)}a(y)dy$$

的线性叠加, 其中 a 是适当的函数. 要理解 u 的传播速度, 可以考虑极限 $t\to\infty$, 同时比值 $v:=\frac{x}{t}$ 固定. 后面在固定相的部分会介绍积分

$$\int_{\mathbb{R}^n}e^{i(y\cdot x-\sigma(y)t)}a(y)dy=\int_{\mathbb{R}^n}e^{it(y\cdot v-\sigma(y))}a(y)dy$$

的主要贡献在于波数 y, 其中 $D\sigma(y)=v$. 因此称 $D\sigma(y)$ 为群速度.

9.2.1.2 孤波

下面考虑 Korteweg-de Vries(KdV) 方程, 其形如

$$u_t+6uu_x+u_{xxx}=0,\quad 在\mathbb{R}\times(0,\infty)中 \tag{9.41}$$

这个非线性色散方程可以作为水的表面波模型. 下面希望找出一个具有结构

$$u(x,t)=v(x-\sigma t),\quad x\in\mathbb{R},t>0 \tag{9.42}$$

的行波解. 将(9.42)式代入(9.41)式得到 ODE

$$-\sigma v'+6vv'+v'''=0 \tag{9.43}$$

其中 $' = \frac{d}{ds}$. 对(9.43)式两边积分有

$$-\sigma v + 3v^2 + v'' = a \tag{9.44}$$

其中 a 表示某常数. (9.44)式两边同乘 v' 得

$$-\sigma v v' + 3v^2 v' + v'' v = a v'$$

积分得

$$\frac{(v')^2}{2} = -v^3 + \frac{\sigma}{2} v^2 + av + b \tag{9.45}$$

其中 b 是任意常数.

下面研究(9.45)式满足 $s \to \pm\infty$ 时 $v, v', v'' \to 0$ 的解 v(此时形如(9.42)式的函数 u 称为孤波). 令 $s \to \pm\infty$ 可由(9.44),(9.45)式知 $a = b = 0$. 进而(9.45)式简化为

$$\frac{(v')^2}{2} = v^2(-v + \frac{\sigma}{2})$$

也即 $v' = \pm v(\sigma - 2v)^{\frac{1}{2}}$.

为了计算简便, 考虑 $v' = -v(\sigma - 2v)^{\frac{1}{2}}$ 的情况, 得到:

$$-\frac{dv}{v(\sigma - 2v)^{\frac{1}{2}}} = ds$$

两边从 1 到 s 积分得到下述关于 v 的隐式表达式:

$$s = -\int_1^{v(s)} \frac{dz}{z(\sigma - 2z)^{\frac{1}{2}}} + c \tag{9.46}$$

其中 c 是某常数. 现考虑换元[4]$z = \frac{\sigma}{2}\operatorname{sech}^2\theta$, 进而 $\frac{dz}{d\theta} = -\sigma \operatorname{sech}^2\theta \tanh\theta$, 得到 $z(\sigma - 2z)^{\frac{1}{2}} = \frac{\sigma^{\frac{3}{2}}}{2}\operatorname{sech}^2\theta \tanh\theta$, 从而(9.46)式变为

$$s = \frac{2}{\sqrt{\sigma}}\theta + c \tag{9.47}$$

其中 θ 由隐式关系式

$$\frac{\sigma}{2}\operatorname{sech}^2\theta = v(s) \tag{9.48}$$

给出. 最后联立(9.47),(9.48)式得到

$$v(s) = \frac{\sigma}{2}\operatorname{sech}^2\left(\frac{\sqrt{\sigma}}{2}(s - c)\right), \quad s \in \mathbb{R}$$

反之, 也可以验证按上式定义的 v 确实是 ODE(9.43)的解. 最后结果表明

$$u(x,t) = \frac{\sigma}{2}\operatorname{sech}^2\left(\frac{\sqrt{\sigma}}{2}(x - \sigma t - c)\right), \quad x \in \mathbb{R}, t \geq 0$$

对任意的 $c \in \mathbb{R}, \sigma > 0$ 都是 KdV 方程的解. 这种形式的解就称为孤波, 注意孤波的速度依赖于其高度.

KdV 方程实际上极其特别, 这表现在它是可积的, 也即理论上来说一个给定的通解能接受任何一个初值[5].

9.2.1.3 半稳定方程的行波解

下面考虑标量反应扩散方程

$$u_t - u_{xx} = f(u), \quad \text{在 } \mathbb{R} \times (0, \infty) \text{ 中} \tag{9.49}$$

其中 $f : \mathbb{R} \to \mathbb{R}$” 形如三次曲线”.

---

[4] $\operatorname{sech} x = \frac{1}{\cosh x} = \frac{2}{e^x + e^{-x}}$, $(\operatorname{sech} x)' = -\operatorname{sech} x \tanh x$.

[5] 这句话不知道怎么翻译了...

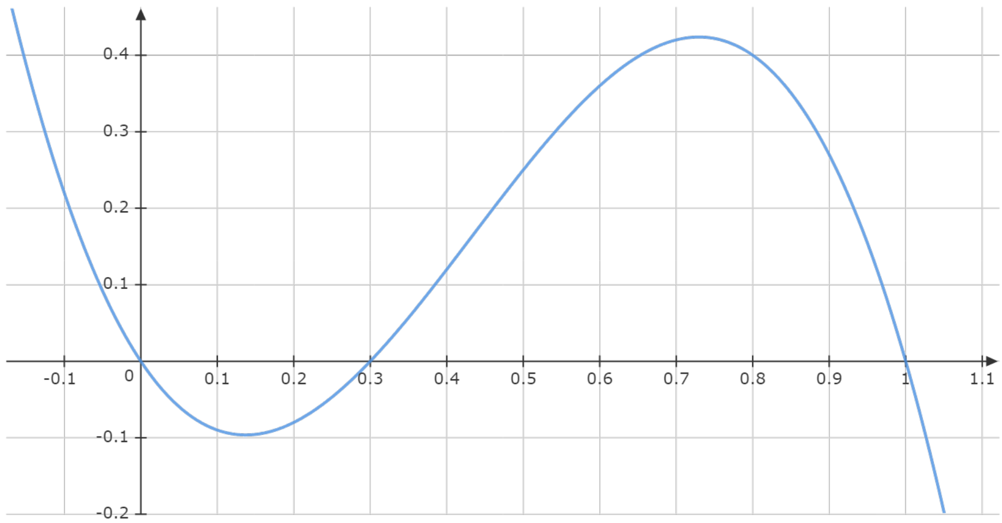


图 9.1: f 示意图

进一步设 f 光滑且满足:

$$\begin{cases} \text{(a) } f(0)=f(a)=f(1)=0 \\ \text{(b) 在}(0,a)\text{上}f<0, \text{在 }(a,1)\text{ 上}f>0 \\ \text{(c) } f'(0)<0, f'(1)<0 \\ \text{(d) } \int_0^1 f(z)dz>0 \end{cases} \tag{9.50}$$

其中 $0<a<1$.

现在希望找到形如

$$u(x,t)=v(x-\sigma t) \tag{9.51}$$

的行波解, 其中波形 v 与波速 σ 需满足:

$$u\to 0(x\to-\infty), \quad u\to 1(x\to+\infty)$$

回忆前面讨论 Turing 不稳定性时对反应扩散方程的研究, 套用(9.27)式, 知因为在 $z=0,1$ 处 $f'<0$, 故常值 0 和 1 就是 PDE 的稳定解. 同理, 因为在 $z=a$ 处 $f'\geq 0$, 故常值 a 就是 PDE 的不稳定解[6]. 根据这个性质, 我们希望行波解(9.51)在 $x=\mp\infty$ 处位于 $z=0, z=1$ 中间.

将(9.51)式代入(9.49)式, 知 v 需满足 ODE:

$$v''+\sigma v'+f(v)=0 \tag{9.52}$$

其中 $'=\frac{d}{ds}$, 同时需有

$$\lim_{s\to+\infty} v(s)=1, \lim_{s\to-\infty} v(s)=0, \lim_{s\to\pm\infty} v'(s)=0 \tag{9.53}$$

现在 (不完全带证明地) 对 ODE 问题(9.52),(9.53)进行相平面分析. 设

$$w:=v'$$

则(9.52),(9.53)转化成一阶系统:

$$\begin{cases} v'=w \\ w'=\sigma w-f(v) \end{cases} \tag{9.54}$$

[6]这里的稳定与不稳定是相对于时间 t 说的. 验证稳定性可以通过取近似 $f(u)=ku$, 利用 Fourier 方法解出 u, 发现当 $k>0$ 时表达式中会出现 e^{kt}, 这便说明 $t\to\infty$ 时 u 并没有得到很好的控制.

其中

$$\lim_{s\to\infty}(v,w)=(1,0),\ \lim_{s\to-\infty}(v,w)=(0,0) \tag{9.55}$$

现在 (0,0),(1,0) 就是系统(9.54)的驻点, 对应线性近似方程的特征值为

$$\lambda_0^{\pm}=\frac{-\sigma\pm(\sigma^2-4f'(0))^{\frac{1}{2}}}{2},\lambda_1^{\pm}=\frac{-\sigma\pm(\sigma^2-4f'(1))^{\frac{1}{2}}}{2} \tag{9.56}$$

根据前面的条件(9.50)(c) 知, $\lambda_0^{\pm},\lambda_1^{\pm}$ 都是实的, 且符号相异, 进而 (0,0),(1,0) 两点就是系统(9.54)相流的鞍点. 因此, ” 不稳定的曲线”W^u[7]

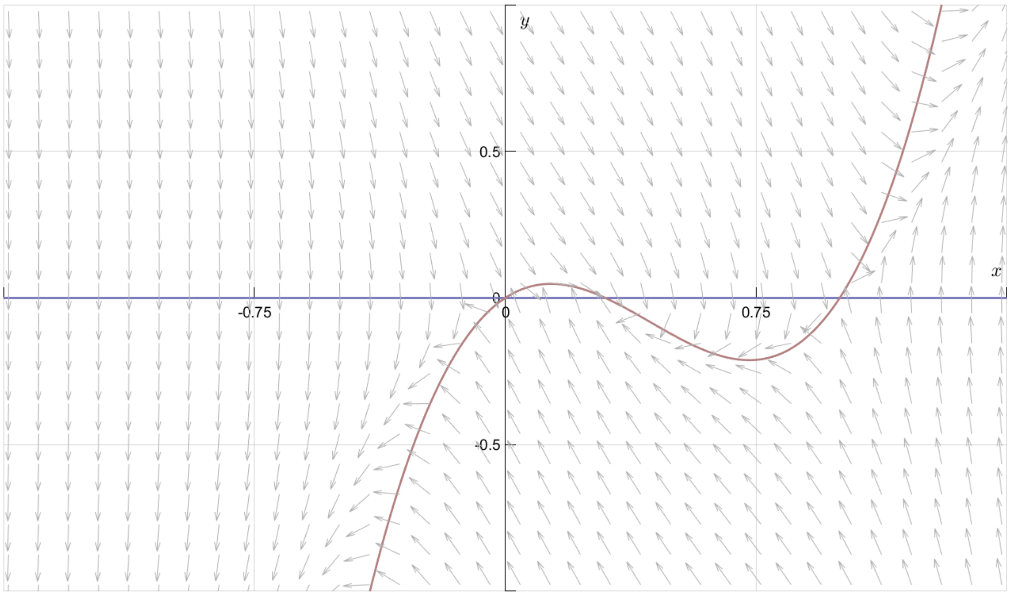


图 9.2: $\sigma=2$ 时系统(9.54)的相图

9.2.2 标量下的相似性

下面阐述 PDE 的一些其他形式的” 相似解” 的情况.

例 9.4(标量不变解) 再考虑一次多孔介质方程

$$u_t-\Delta(u^\gamma)=0,\quad 在\ \mathbb{R}^n\times(0,\infty)\ 中 \tag{9.57}$$

其中 $u\geq 0$, 而 $\gamma>1$ 是常数.

类似于前面推导热方程基本解的方法, 下面寻找形如

$$u(x,t)=\frac{1}{t^\alpha}v(\frac{x}{t^\beta}),\quad x\in\mathbb{R}^n,t>0 \tag{9.58}$$

的解 u, 其中常数 α,β 与函数 $v:\mathbb{R}^n\to\mathbb{R}$ 均需确定. 回忆之前推导热方程基本解的思路, 寻求形式(9.58)的思路也是在于我们想找到在膨胀情况:

$$u(x,t)\to\lambda^\alpha u(\lambda^\beta x,\lambda t)$$

中不变的解 u. 进而对全体 $\lambda>0,x\in\mathbb{R}^n,t>0$, 有

$$u(x,t)=\lambda^\alpha u(\lambda^\beta x,\lambda t).$$

取 $\lambda=\frac{1}{t}$ 即得 $v(y):=u(y,1)$.

[7] 到这里已经完全翻译不下去了, 所以直接跳到下一节, 不知道书上 W^u,W^s 到底是怎么设的?

将(9.58)式代回(9.57)式, 并取 $y=\frac{x}{t^{\beta}}$ 得

$$\alpha t^{-(\alpha+1)}v(y)+\beta t^{-(\alpha+1)}y\cdot Dv(y)+t^{-(\alpha\gamma+2\beta)}\Delta(v^{\gamma})(y)=0 \tag{9.59}$$

现在希望把(9.59)式化成只含有变量 y 的形式, 这便需要

$$\alpha+1=\alpha\gamma+2\beta \tag{9.60}$$

得到

$$\alpha v+\beta y\cdot Dv+\Delta(v^{\gamma})=0 \tag{9.61}$$

至此已经把问题从 $n+1$ 维化成 n 维 (约去 t) 了. 进一步设 v 是放射的, 也即 $v(y)=w(|y|)$, 其中 $w:\mathbb{R}\to\mathbb{R}$, 代入(9.61)式可得

$$\alpha w+\beta rw'+(w^{\gamma})''+\frac{n-1}{r}(w^{\gamma})'=0 \tag{9.62}$$

其中 $r=|y|, '=\frac{d}{dr}$. 现在若设

$$\alpha=n\beta \tag{9.63}$$

则(9.62)式简化为

$$(r^{n-1}(w^{\gamma})')'+\beta(r^{n}w)'=0$$

积分有

$$r^{n-1}(w^{\gamma})'+\beta r^{n}w=a$$

其中 a 为常数. 设 $r\to\infty$ 时 $w,w'\to 0$ 是足够快的 (至少快于 r^n), 则令 $r\to\infty$ 可得 $a=0$, 进而

$$(w^{\gamma})'=-\beta rw\Rightarrow \gamma w^{\gamma-1}w'=-\beta rw\Rightarrow (w^{\gamma-1})'=-\frac{\gamma-1}{\gamma}\beta r$$

积分得

$$w^{\gamma-1}=b-\frac{\gamma-1}{2\gamma}\beta r^{2}$$

其中 b 是常数, 故

$$w=((b-\frac{\gamma-1}{2\gamma}\beta r^{2})+)^{\frac{1}{\gamma-1}} \tag{9.64}$$

其中 $(f)^{+}$ 表示 f 的正部[8], 这样取是为了确保 $w\geq 0$. 回忆先前设过 $v(y)=w(r)$, 结合(9.58)式有

$$u(x,t)=\frac{1}{t^{\alpha}}((b-\frac{\gamma-1}{2\gamma}\beta\frac{|x|^{2}}{t^{2\beta}})^{+})^{\frac{1}{\gamma-1}},\quad x\in\mathbb{R}^{n},t>0 \tag{9.65}$$

且由(9.60),(9.63)式可解得

$$\alpha=\frac{n}{n(\gamma-1)+2},\quad \beta=\frac{1}{n(\gamma-1)+2} \tag{9.66}$$

(9.65),(9.66)式就是多孔介质方程的 Barenblatt-Kompaneetz-Zeldovich 解. □

观察到 Barenblatt-Kompaneetz-Zeldovich 解对每个 $t>0$ 都具有紧支集, 这是具紧支初值条件的多孔介质方程 (在合适的定义下的) 弱非负解的普遍特点. 非线性抛物型 PDE(9.57)在 $u=0$ 时就退化了, 集合 $\{u>0\}$ 以有限传播速度移动. 故多孔介质方程(9.57)常被当成是比线性热方程更优的耗散传播模型, 这是因为后者暗含了无限传播速度.

9.3 变换法

本节介绍 Fourier 变换 $\mathcal{F}$, Radon 变换 $\mathcal{R}$ 与 Laplace 变换 $\mathcal{L}$ 的一些理论, 它们是把一些线性 PDE 转化成代数方程或含更少变量的方程的强大工具.

[8] 回忆实变函数论中引出 Lebesgue 积分时的做法.

9.3.1 Fourier 变换

本节所有函数都是复值函数, $\overline{x}$ 表示 x 的复共轭.

9.3.1.1 定义与性质

定义 9.3.1 (Fourier 变换)

若 $u \in L^1(\mathbb{R}^n)$, 则定义其 Fourier 变换 $\mathcal{F}[u] = \widehat{u}$ 为:

$$\hat{u}(y) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} u(x)dx, \quad y \in \mathbb{R}^n \tag{9.67}$$

其 Fourier 逆变换 $\mathcal{F}^{-1}[u] = \check{u}$ 为

$$\check{u}(y) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ix\cdot y} u(x)dx, \quad y \in \mathbb{R}^n \tag{9.68}$$

因为 $|e^{\pm i\cdot y}| = 1, u \in L^1(\mathbb{R}^n)$, 故上述积分对每个 $y \in \mathbb{R}^n$ 均收敛.

注 在调和分析的教材中, Fourier 变换其实一开始是定义在 Schwartz 空间中的. Schwartz 空间的定义如下:

补充定义 9.3.1 (Schwartz 空间[LG1])

$\mathbb{R}^n$ 上的 C^∞ 函数 f 称为 Schwartz 函数, 如果对每一对多重指标 α, β 而言, 总存在正常数 $C_{\alpha,\beta}$ 使得

$$\rho_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^n} |x^\alpha D^\beta f(x)| = C_{\alpha,\beta} < \infty$$

$\rho_{\alpha,\beta}(f)$ 称为 f 的 Schwartz 半范, $\mathbb{R}^n$ 上全体 Schwartz 函数构成的空间便称为 Schwartz 空间, 记作 $\mathscr{S}(\mathbb{R}^n)$.

Schwartz 函数同时有下述刻画:

补充命题 9.3.1

C^∞ 函数 f 是 Schwartz 函数当且仅当对任意正整数 N 与任意多重指标 α, 存在正常数 $C_{\alpha,N}$ 使得

$$|(D^\alpha f)(x)| \leq C_{\alpha,N}(1+|x|)^{-N}$$

总的来说, Schwartz 函数 $f \in C^\infty$ 说明函数足够光滑, 而 $\forall\alpha\forall\beta(\rho_{\alpha,\beta}(f) < \infty)$ 说明 Schwartz 函数足够衰减 (在无穷远处比任意多项式衰减得都快). 在后面会说明 Fourier 变换实际上是 Schwartz 空间的自同构 (亦即把 Schwartz 函数保范地映成 Schwartz 函数). 特别注意: Schwartz 函数的 Fourier 变换总存在, 这便省去了讨论存在性的麻烦.

课堂笔记 (2023.11.27)

在进行接下来的讨论前, 有必要先说明仅根据 Fourier 变换与逆变换的定义本身即可得到的性质, 下面的性质选自 **[LG1]** 与上课的笔记. 设 $f, g \in \mathscr{S}(\mathbb{R}^n)$, 则:

(i) $\|\widehat{f}\|_{L^\infty} \leq \frac{1}{(2\pi)^{\frac{n}{2}}}\|f\|_{L^1}$.

证明: 根据 Fourier 变换的定义知

$$\begin{aligned}\|\widehat{f}\|_{L^\infty} &= \operatorname*{ess\ sup}_{\xi\in\mathbb{R}^n} \left|\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx\right| \leq \operatorname*{ess\ sup}_{\xi\in\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |e^{-ix\cdot\xi}||f(x)|dx \\ &= \operatorname*{ess\ sup}_{\xi\in\mathbb{R}^n} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |f(x)|dx = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |f(x)|dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\|f\|_{L^1}\end{aligned}$$

□

特别地, 这一条说明只要 $f \in L^1$, 那么 $\widehat{f} \in L^\infty$.

(ii) $\widehat{f+g} = \widehat{f} + \widehat{g}$,

证明: 为了清楚表示 Fourier 变换作用在怎样的函数上面, 一般对涉及运算符号的函数 u, 其 Fourier

变换均用 $\mathcal{F}[u]$ 这个形式表示. 现在根据 Fourier 变换的定义有

$$\begin{aligned}\mathcal{F}[f+g](\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}(f(x)+g(x))dx \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}f(x)dx + \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}g(x)dx = \widehat{f}(\xi)+\widehat{g}(\xi)\end{aligned}$$

□

(iii) $\widehat{bf} = b\widehat{f}(\forall b \in \mathbb{R})$,

证明: 根据 Fourier 变换的定义知

$$\mathcal{F}[bf](\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}(bf(x))dx = b\cdot\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}f(x)dx = b\widehat{f}(\xi)$$

□

(iv) $\mathcal{F}^{-1}[u](x) = \mathcal{F}[u](-x)$.

证明: 根据 Fourier 变换与逆变换的定义, 令 $\tau = -x$ 有

$$\mathcal{F}^{-1}[u](x) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}u(\xi)d\xi = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\tau\cdot\xi}u(\xi)d\xi = \mathcal{F}[u](\tau) = \mathcal{F}[u](-x).$$

□

(v) $\forall\lambda\in\mathbb{R}(\mathcal{F}[u(\lambda\circledast)](\xi) = \lambda^{-n}\mathcal{F}[u](\frac{\xi}{\lambda}))$,[a]

证明: 根据 Fourier 变换的定义, 令 $\tau = \lambda x$ 有

$$\begin{aligned}\mathcal{F}[u(\lambda\circledast)](\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}u(\lambda x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\frac{\tau}{\lambda}\cdot\xi}u(\tau)d\frac{\tau}{\lambda} \\ &= \frac{1}{\lambda^n}\cdot\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\tau\cdot\frac{\xi}{\lambda}}u(\tau)d\tau = \frac{1}{\lambda^n}\mathcal{F}[u](\frac{\xi}{\lambda})\end{aligned}$$

□

(vi) $\mathcal{F}[f(\circledast-h)](\xi) = e^{-i\xi\cdot h}\mathcal{F}[f](\xi)$,

证明: 根据 Fourier 变换的定义有

$$\mathcal{F}[f(\circledast-h)](\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}f(x-h)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i(x+h)\xi}f(x)dx = e^{-ih\cdot\xi}\mathcal{F}[f](\xi)$$

□

(vii) $\mathcal{F}[e^{ih\cdot\circledast}f(\circledast)](\xi) = \mathcal{F}[f](\xi-h)$,

证明: 根据 Fourier 变换的定义有

$$\mathcal{F}[e^{ih\cdot\circledast}f(\circledast)](\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}e^{ih\cdot x}f(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot(\xi-h)}f(x)dx = \mathcal{F}[f](\xi-h)$$

□

(viii) $\mathcal{F}[\overline{f}](\xi) = \overline{\mathcal{F}^{-1}[f](\xi)}$.

证明: 根据 Fourier 变换的定义有

$$\mathcal{F}[\overline{f}](\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\overline{f}(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} \overline{e^{ix\cdot\xi}f(x)}dx = \overline{\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}f(x)dx} = \overline{\mathcal{F}^{-1}[f](\xi)}$$

□

(ix) (卷积公式) $\widehat{f*g} = (2\pi)^{\frac{n}{2}}\widehat{f}\widehat{g}$.

证明: 根据 Fourier 变换的定义有

$$\begin{aligned}\mathcal{F}[f*g](\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}(f*g)(x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}(\int_{\mathbb{R}^n} f(x-y)g(y)dy)dx \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}f(x-y)g(y)dxdy = \int_{\mathbb{R}^n}(\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i(x-y)\cdot\xi}f(x-y)dx)e^{-iy\cdot\xi}g(y)dy \\ &= \int_{\mathbb{R}^n}\widehat{f}(\xi)e^{-iy\cdot\xi}g(y)dy = (2\pi)^{\frac{n}{2}}\widehat{f}(\xi)\widehat{g}(\xi)\end{aligned}$$

□

特别要注意的是, 上面提到的这些性质其实只要求 $f,g\in L^1(\mathbb{R}^n)$ 就足够了, 这里限制 $f,g\in\mathscr{S}(\mathbb{R}^n)$ 并在之后用 Schwartz 函数逼近 L^2 函数只是调和分析中的一种常见做法. [Ev] 中全书没有提起 Schwartz 函数, 而是用 $L^1\cap L^2$ 的逼近去定义 L^2 中的 Fourier 变换, 这两种做法本质上是一样的.

[a] $F[u(\lambda\circledast)]$ 中, $\circledast$ 表示在具体运算时这里应该填入自变量 x, 这种表示方式在前面泛函分析部分介绍范数时已经见过很多次了 (即 $\|\cdot\|$, 这里不用 $\cdot$ 而用 $\circledast$ 是因为怕和后面的内积搞混), 这个记法是因为函数的 Fourier 变换本身和单点没有任何关系.

在之后的证明中, 需要额外注意一个函数的 Fourier 变换:

命题 9.3.1

$\mathcal{F}(e^{-|\circledast|^2})(\xi)=\frac{1}{2^{\frac{n}{2}}}e^{-\frac{|\xi|^2}{4}}$.

证明

根据 Fourier 变换的定义知

$$\mathcal{F}(e^{-|\circledast|^2})(\xi)=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-ix\cdot\xi}e^{-|x|^2}dx=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-i\sum\limits_{k=1}^{n}x_k\xi_k}e^{-\sum\limits_{k=1}^{n}x_k^2}dx$$
$$\overset{(i)}{=}\frac{1}{(2\pi)^{\frac{n}{2}}}\prod_{k=1}^{n}\int_{\mathbb{R}}e^{-ix_k\xi_k-x_k^2}dx_k=\frac{1}{(2\pi)^{\frac{n}{2}}}\prod_{k=1}^{n}\int_{\mathbb{R}}e^{-(x_k+i\frac{\xi_k}{2})^2}e^{-\frac{\xi_k^2}{4}}dx_k$$

对于 $\int_{\mathbb{R}}e^{-(x_k+\frac{i\xi_k}{4})^2}dx_k$, 可以验证 e^{-z^2} 在复平面中四点 $(-p,0),(p,0),(p,\frac{\xi_k}{4}),(-p,\frac{\xi_k}{4})$ 连成的矩形中解析 $(p=1,2,\cdots)$, 因而根据留数定理并令 $p\to\infty$ 有

$$\int_{-\infty}^{\infty}e^{-x_k^2}dx_k+\int_{\infty}^{-\infty}e^{-(x_k+\frac{i\xi_k}{4})^2}dx_k=0\Rightarrow\int_{-\infty}^{\infty}e^{-(x_k+\frac{i\xi_k}{4})^2}dx_k=\int_{-\infty}^{\infty}e^{-x_k^2}dx_k=\pi^{\frac{1}{2}}$$

故

$$\frac{1}{(2\pi)^{\frac{n}{2}}}\prod_{k=1}^{n}\int_{\mathbb{R}}e^{-(x_k+i\frac{\xi_k}{2})^2}e^{-\frac{\xi_k^2}{4}}dx_k=\frac{1}{(2\pi)^{\frac{n}{2}}}\pi^{\frac{n}{2}}\prod_{k=1}^{n}e^{-\frac{\xi_k^2}{4}}=\frac{1}{2^{\frac{n}{2}}}e^{-\frac{|\xi|^2}{4}}$$

□

下面把 Fourier 变换与逆变换的定义(9.67),(9.68)延拓到 $L^2(\mathbb{R}^n)$ 上, 这依赖于下述极其重要的 Plancherel 定理.

定理 9.3.1 (Plancherel 定理)

设 $u\in L^1(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)$, 则 $\widehat{u},\check{u}\in L^2(\mathbb{R}^n)$, 且[a]

$$\|\widehat{u}\|_{L^2(\mathbb{R}^n)}=\|\check{u}\|_{L^2(\mathbb{R}^n)}=\|u\|_{L^2(\mathbb{R}^n)}. \tag{9.69}$$

[a] Plancherel 定理表明 Fourier 变换与 Fourier 逆变换都是保 L^2 范数的.

证明

因为 $\forall f\in L^1(\mathbb{R}^n)(\|\widehat{f}\|_{L^\infty}\le\|f\|_{L^1})$(补充性质 (i)), 故由 $v,w\in L^1(\mathbb{R}^n)$ 可知 $\widehat{v},\widehat{w}\in L^\infty(\mathbb{R}^n)$, 同时注意

$$\begin{aligned}\int_{\mathbb{R}^n}v(x)\widehat{w}(x)dx&=\int_{\mathbb{R}^n}v(x)\cdot(\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-iy\cdot x}w(y)dy)dx\\&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{1}{(2\pi)^{\frac{n}{2}}}v(x)e^{-iy\cdot x}w(y)dydx\\&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{1}{(2\pi)^{\frac{n}{2}}}v(x)e^{-iy\cdot x}w(y)dxdy\\&=\int_{\mathbb{R}^n}w(y)\cdot(\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-ix\cdot y}v(x)dx)dy\\&=\int_{\mathbb{R}^n}w(y)\widehat{v}(y)dy\end{aligned}$$

其中第二行到第三行的积分换序是 Fubini 定理, 由上述推导可得

$$\int_{\mathbb{R}^n} v(x)\widehat{w}(x)dx = \int_{\mathbb{R}^n} \widehat{v}(y)w(y)dy \tag{9.70}$$

根据命题9.3.1与补充性质 (v) 可得[9]:

$$\mathcal{F}[e^{-t|\circledast|^2}](\xi) = \frac{1}{t^{\frac{n}{2}}}\mathcal{F}[e^{-|\circledast|^2}](\frac{\xi}{\sqrt{t}}) = \frac{1}{(2t)^{\frac{n}{2}}}e^{-\frac{|\xi|^2}{4t}}, \quad \forall t > 0$$

故若取 $\varepsilon > 0, v_\varepsilon(x) := e^{-\varepsilon|x|^2}$, 可得 $\widehat{v}_\varepsilon(y) = \frac{1}{(2\varepsilon)^{\frac{n}{2}}}e^{-\frac{|y|^2}{4\varepsilon}}$, 再将 $v_\varepsilon(x), \widehat{v}_\varepsilon(y)$ 代入(9.70)式可得

$$\int_{\mathbb{R}^n} \widehat{w}(y)e^{-\varepsilon|y|^2}dy = \frac{1}{(2\varepsilon)^{\frac{n}{2}}}\int_{\mathbb{R}^n} w(x)e^{-\frac{|x|^2}{4\varepsilon}}dx, \quad \varepsilon > 0 \tag{9.71}$$

现取 $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, 取 $\widetilde{u}(x) := \overline{u}(-x)$, 设 $w := u * \widetilde{u} \in L^1(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, 根据补充性质 (ix) 有

$$\widehat{w}(y) = (2\pi)^{\frac{n}{2}}\widehat{u}(y)\widehat{\widetilde{u}}(y) \in L^\infty(\mathbb{R}^n)$$

但同时

$$\begin{aligned}\widehat{\widetilde{u}}(y) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot y}\overline{u}(-x)dx = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} \overline{e^{ix\cdot y}u(-x)}dx \\ &\overset{(i)}{=} \frac{1}{(2\pi)^{\frac{n}{2}}}\overline{\int_{\mathbb{R}^n} e^{-ix\cdot y}u(x)dx} = \overline{\widehat{u}}(y)\end{aligned}$$

其中 (i) 把 $-x$ 替换为 x, dx 不会出现负号是因为此时积分上下限也对应改变, 两个负号抵消了. 故

$$\widehat{w}(y) = (2\pi)^{\frac{n}{2}}\widehat{u}(y)\overline{\widehat{u}}(y) = (2\pi)^{\frac{n}{2}}|\widehat{u}(y)|^2 \tag{9.72}$$

现在已知 w 连续, 故

$$\lim_{\varepsilon\to 0}\frac{1}{(2\varepsilon)^{\frac{n}{2}}}\int_{\mathbb{R}^n} w(x)e^{-\frac{|x|^2}{4\varepsilon}}dx = (2\pi)^{\frac{n}{2}}w(0) \tag{9.73}$$

其中等号成立是基于热方程基本解的性质. 因为 $\widehat{w}(y) = (2\pi)^{\frac{n}{2}}|\widehat{u}|^2(y) \geq 0$, 故根据 Beppo Levi 定理, 在(9.71)式中令 $\varepsilon \to 0^+$, 得到 $\widehat{w}(y)$ 可积, 且结合(9.73)式有

$$\int_{\mathbb{R}^n} \widehat{w}(y)dy = (2\pi)^{\frac{n}{2}}w(0)$$

再将(9.72)式代入得:

$$\int_{\mathbb{R}^n} |\widehat{u}|^2(y)dy = w(0) = (u * \widetilde{u})(0) = \int_{\mathbb{R}^n} u(x)\widetilde{u}(-x)dx = \int_{\mathbb{R}^n} u(x)\overline{u}(x)dx = \int_{\mathbb{R}^n} |u|^2(x)dx$$

此即

$$\|\widehat{u}\|_{L^2} = \|u\|_{L^2}.$$

根据补充性质 (iv), 知

$$\int_{\mathbb{R}^n} |\widehat{u}|^2(y)dy = \int_{\mathbb{R}^n} |\check{u}|^2(-y)dy = \int_{\mathbb{R}^n} |\check{u}|^2(y)dy$$

亦即

$$\|\widehat{u}\|_{L^2} = \|\check{u}\|_{L^2} = \|u\|_{L^2}$$

□

前面的 Fourier 变换是定义在 L^1 上的, 下面讨论 L^2 上的 Fourier 变换. 根据 Plancherel 定理(9.3.1), 可以从逼近的方法来定义 L^2 中的 Fourier 变换. 现取序列 $\{u_k\}_{k=1}^\infty \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, 其中

$$u_k \to u(L^2(\mathbb{R}^n))$$

根据 Plancherel 定理(9.3.1):

$$\|\widehat{u}_k - \widehat{u}_j\|_{L^2(\mathbb{R}^n)} = \|\widehat{u_k - u_j}\|_{L^2(\mathbb{R}^n)} = \|u_k - u_j\|_{L^2(\mathbb{R}^n)}$$

[9]书上这一步有误: 应把 y 换成 $-y$.

故 $\{\widehat{u}_k\}_{k=1}^{\infty}$ 是 $L^2(\mathbb{R}^n)$ 中的基本列. 根据 $L^2(\mathbb{R}^n)$ 的完备性, 该序列收敛到某极限, 该极限定义为 $\mathcal{F}u = \widehat{u}$:

$$\widehat{u}_k \to \widehat{u}(L^2(\mathbb{R}^n))$$

显见 $\widehat{u}$ 的定义其实并不依赖于逼近序列 $\{\widehat{u}_k\}_{k=1}^{\infty}$ 的选取. 类似可以定义 $\check{u}$. 下面将其整理成正式定义:

定义 9.3.2 ($L^2(\mathbb{R}^n)$ 中的 Fourier 变换与 Fourier 逆变换)

若 $u \in L^2(\mathbb{R}^n)$, 且存在序列 $\{u_k\}_{k=1}^{\infty} \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$[a]满足 $u_k \to u(L^2(\mathbb{R}^n))$, 则定义 u 的 Fourier 变换 $\mathcal{F}u = \widehat{u}$ 为:

$$\widehat{u} := \lim_{k\to\infty} \widehat{u}_k(L^2(\mathbb{R}^n))$$

u 的 Fourier 逆变换 $\mathcal{F}^{-1}u = \check{u}$ 定义为:

$$\check{u} := \lim_{k\to\infty} \check{u}_k(L^2(\mathbb{R}^n)).$$

[a]也就是 u_k 在 $L^2(\mathbb{R}^n)$ 的意义下收敛到 u, 写成形式逻辑即 $\forall \varepsilon > 0 \exists N > 0 \forall k > N(\|u_k - u\|_{L^2(\mathbb{R}^n)} < \varepsilon)$.

♣

注 对于 $u \in L^2(\mathbb{R}^n)$ 而言, 因为

$$\widehat{u} := \lim_{k\to\infty} \widehat{u}_k$$

而根据 Plancherel 定理(9.3.1):

$$\|\widehat{u}_k\|_{L^2(\mathbb{R}^n)} = \|u_k\|_{L^2(\mathbb{R}^n)}, \quad k = 1, 2, \cdots, \{u_k\}_{k=1}^{\infty} \subset L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$$

故由 Lebesgue 控制收敛定理知:

$$\|\widehat{u}\|_{L^2(\mathbb{R}^n)} = \|\lim_{k\to\infty} \widehat{u}_k\|_{L^2(\mathbb{R}^n)} = \lim_{k\to\infty} \|\widehat{u}_k\|_{L^2(\mathbb{R}^n)} = \lim_{k\to\infty} \|u_k\|_{L^2(\mathbb{R}^n)} = \|\lim_{k\to\infty} u_k\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}$$

这说明 L^2 函数的 Fourier 变换也满足 Plancherel 定理.

下面介绍一些常用的公式:

定理 9.3.2 (Fourier 变换的一些性质)

设 $u, v \in L^2(\mathbb{R}^n)$, 则

(i) $\int_{\mathbb{R}^n} u\overline{v}dx = \int_{\mathbb{R}^n} \widehat{u}\overline{\widehat{v}}dy$.

(ii) 对每个多重指标 α, 若其满足 $D^\alpha u \in L^2(\mathbb{R}^n)$, 则[a]$(D^\alpha u)^\wedge = (iy)^\alpha \widehat{u}$.

(iii) 若 $u, v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, 则 $(u * v)^\wedge = (2\pi)^{\frac{n}{2}} \widehat{u}\widehat{v}$.

(iv) $u = (\widehat{u})^\vee$.

[a]注意 $f^\wedge$ 同样表示 f 的 Fourier 变换, 类似 $f^\vee$ 也表示 f 的 Fourier 逆变换.

♡

上述断言 (iv) 称为 Fourier 反演公式, 通过该公式, 只要 $\widehat{u} \in L^1(\mathbb{R}^n)$, 它就可以表成指数型平面波 $e^{ix\cdot y}$ 的形式

$$u(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ix\cdot y}\widehat{u}(y)dy \tag{9.74}$$

证明

(i): 设 $u, v \in L^2(\mathbb{R}^n), \alpha \in \mathbb{C}$, 则

$$\|u + \alpha v\|_{L^2(\mathbb{R}^n)}^2 = \|\widehat{u + \alpha v}\|_{L^2(\mathbb{R}^n)}^2 = \|\widehat{u} + \alpha\widehat{v}\|_{L^2(\mathbb{R}^n)}^2$$

把式子展开得

$$\int_{\mathbb{R}^n} (|u|^2 + |\alpha v|^2 + \overline{u} \cdot (\alpha v) + u \cdot (\overline{\alpha v}))dx = \int_{\mathbb{R}^n} (|\widehat{u}|^2 + |\widehat{\alpha v}|^2 + \overline{\widehat{u}} \cdot (\widehat{\alpha v}) + \widehat{u} \cdot (\overline{\widehat{\alpha v}}))dy$$

根据 Plancherel 定理(9.3.1)与前面的论述可得

$$\int_{\mathbb{R}^n} |u|^2 dx = \int_{\mathbb{R}^n} |\widehat{u}|^2 dy, \quad \int_{\mathbb{R}^n} |v|^2 dx = \int_{\mathbb{R}^n} |\widehat{v}|^2 dy$$

从而

$$\int_{\mathbb{R}^n} (\alpha \overline{u} v + \overline{\alpha} u \overline{v}) dx = \int_{\mathbb{R}^n} (\alpha \overline{\widehat{u}} \widehat{v} + \overline{\alpha} \widehat{u} \overline{\widehat{v}}) dy$$

分别取 $\alpha = 1, \alpha = i$ 得到:

$$\int_{\mathbb{R}^n} (\overline{u} v + u \overline{v}) dx = \int_{\mathbb{R}^n} (\overline{\widehat{u}} \widehat{v} + \widehat{u} \overline{\widehat{v}}) dy$$
$$\int_{\mathbb{R}^n} (\overline{u} v - u \overline{v}) dx = \int_{\mathbb{R}^n} (\overline{\widehat{u}} \widehat{v} - \widehat{u} \overline{\widehat{v}}) dy$$

进而

$$\int_{\mathbb{R}^n} u \overline{v} dx = \int_{\mathbb{R}^n} \widehat{u} \overline{\widehat{v}} dy$$

(i) 得证.

(ii): 如果 u 光滑且有紧支集, 则

$$\begin{aligned}
(D^\alpha u)^\wedge(y) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} D^\alpha u(x) dx \\
&\overset{(a)}{=} \frac{(-1)^{|\alpha|}}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} D_x^\alpha (e^{-ix\cdot y}) u(x) dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} (iy)^\alpha u(x) dx = (iy)^\alpha \widehat{u}(y)
\end{aligned}$$

其中 (a) 是分部积分. 当 $D^\alpha u \in L^2(\mathbb{R}^n)$, 根据光滑紧支函数在 L^2 函数下在 $L^2(\mathbb{R}^n)$ 稠密的结论与 Lebesgue 控制收敛定理即得命题 (ii).

(iii): 对于 $u, v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n), y \in \mathbb{R}^n$, 有

$$\begin{aligned}
(u * v)^\wedge(y) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot y} (\int_{\mathbb{R}^n} u(z) v(x-z) dx) dz \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-iz\cdot y} u(z) (\int_{\mathbb{R}^n} e^{-i(x-z)\cdot y} v(x-z) dx) dz \\
&= \int_{\mathbb{R}^n} e^{-iz\cdot y} u(z) dz \cdot \widehat{v}(y) = (2\pi)^{\frac{n}{2}} \widehat{u}(y) \widehat{v}(y)
\end{aligned}$$

(iii) 即证.

(iv): 由(9.70)式已知当 $u, v \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ 时有

$$\int_{\mathbb{R}^n} \check{u} v dx = \int_{\mathbb{R}^n} u \check{v} dx$$

现对 $u, v \in L^2(\mathbb{R}^n)$, 则从定义可知 $\widehat{u}, \widehat{v}, \check{u}, \check{v} \in L^2(\mathbb{R}^n)$. 分别设存在序列 $\{u_k\}_{k=1}^\infty, \{v_k\}_{k=1}^\infty$ 使得 $u_k \to u(L^2(\mathbb{R}^n)), v_k \to v(L^2(\mathbb{R}^n))(k \to \infty)$, 根据均值不等式:

$$|\check{u} v| \leq C_1(|\check{u}|^2 + |v|^2), |u \check{v}| \leq C_2(|u|^2 + |\check{v}|^2)$$

其中 C_1, C_2 是常数, 进而 $\check{u} v, u \check{v} \in L^1(\mathbb{R}^n)$, 从而根据 $L^2(\mathbb{R}^n)$ 上 Fourier 变换的定义与 Lebesgue 控制收敛定理, 在等式

$$\int_{\mathbb{R}^n} \check{u}_k v_m dx = \int_{\mathbb{R}^n} u_k \check{v}_m dx$$

两边令 $k, m \to \infty$ 即得:

$$\int_{\mathbb{R}^n} \check{u} v dx = \int_{\mathbb{R}^n} u \check{v} dx, \quad u, v \in L^2(\mathbb{R}^n) \tag{9.75}$$

同时, 注意

$$\begin{aligned}
\check{v}(x) &= \lim_{k\to\infty} \int_{\mathbb{R}^n} e^{ix\cdot y} v_k(y) dy = \overline{\lim_{k\to\infty} \overline{\int_{\mathbb{R}^n} e^{ix\cdot y} v(y) dy}} \\
&= \overline{\lim_{k\to\infty} \int_{\mathbb{R}^n} e^{-ix\cdot y} \overline{v}(y) dy} = \overline{(\overline{v})^\wedge}
\end{aligned}$$

即

$$\check{v}(x) = \overline{(\overline{v})^{\wedge}}, \quad v \in L^2(\mathbb{R}^n) \tag{9.76}$$

故应用结论 (i) 与(9.75),(9.76)式, 对任意的 $u, v \in L^2(\mathbb{R}^n)$ 有:

$$\int_{\mathbb{R}^n} (\widehat{u})^{\vee} v dx \overset{(9.75)}{=} \int_{\mathbb{R}^n} \widehat{u}\check{v} dx \overset{(9.76)}{=} \int_{\mathbb{R}^n} \widehat{u}\overline{(\overline{v})^{\wedge}} dx \overset{\text{(i)}}{=} \int_{\mathbb{R}^n} u\overline{\overline{v}} dx = \int_{\mathbb{R}^n} uv dx$$

注意上式对任意 $v \in L^2(\mathbb{R}^n)$ 都成立, 故命题 (iv) 得证. □

课堂笔记 (2023.11.30)

- (关于(9.73)式的理解) 远在讨论热方程时, 曾经给出过热方程初值问题解的逼近定理7.3.1. 如果对照来看, (9.73)式就是定理7.3.1中令 $x = 0, \varepsilon = t$ 的结论. 在现代分析中, 一般把这个结论叫做热核逼近, 这是恒等逼近定理的一个推论, 参见补充: 恒等逼近定理在 PDE 中的一个体现.
- (关于 Plancherel 定理与 Hahn-Banach 定理的联系) 远在笔记的泛函分析部分, 讨论 Hahn-Banach 定理的时候曾经补充过一个线性算子延拓定理3.5.4, 其内容为如果 $\mathscr{X}, \mathscr{Y}$ 是 Banach 空间, S 是 $\mathscr{X}$ 的稠密线性子空间, 则 $T_0 \in \mathscr{L}(S, \mathscr{Y})$ 可以唯一保范地延拓成 $T \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$. 现在 $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ 正是 $L^2(\mathbb{R}^n)$ 的稠密线性子空间, 而根据 Plancherel 定理, 如果在 $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ 上赋 L^2 范数, 就有 $\mathscr{F}, \mathscr{F}^{-1} \in \mathscr{L}(L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))$. 于是根据线性算子延拓定理, $\mathscr{F}, \mathscr{F}^{-1}$ 在 $L^2(\mathbb{R}^n)$ 上均唯一保范延拓, 这便是定义9.3.2.
- (关于 Fourier 变换性质9.3.2(i) 的注记) 在泛函分析课堂中曾证明过 $L^2(\mathbb{R}^n)$ 是 Hilbert 空间, 其内积为 $\langle f, g \rangle := \int_{\mathbb{R}^n} f(x)\overline{g(x)}dx$. 于是性质 (i) 可以写成:

  $$\langle u, v \rangle = \langle \widehat{u}, \widehat{v} \rangle, \quad \forall u, v \in L^2(\mathbb{R}^n)$$

  这个式子又称为 Parseval 等式. 因为内积与空间的几何结构 (也就是角度) 有关, 故变换前后的内积相等说明这个变换是保角的. 综合 Parseval 等式与 Plancherel 定理, 可以知道 Fourier 变换同时保范保角. 另一方面, 在泛函分析部分中曾给出 B^* 空间中内积与范数相容的充要条件: 平行四边形等式2.6.7, 而这个命题曾说明范数和内积是可以相互诱导的, 在现在的背景下这便说明 Parseval 等式和 Plancherel 定理其实是等价的.
- (关于 Fourier 变换性质9.3.2(ii) 的注记) 如果把 u 看成是物理空间中的函数, 那么 $D^\alpha u$ 同样是物理空间中的函数. $D^\alpha u \in L^2(\mathbb{R}^n)$ 实际上说的是物理空间函数 u 的光滑性, 亦即其具有满足条件的 α 阶导数. 而 Fourier 变换在物理上的意义是把物理空间转换成频率空间, 因为 $D^\alpha u \in L^2(\mathbb{R}^n)$, 故根据 Plancherel 定理知 $(D^\alpha u)^\wedge \in L^2(\mathbb{R}^n)$, 从而 $(D^\alpha u)^\wedge$ 至少在远处 (亦即 $|\xi|$ 足够大时) 是衰减 (亦即趋零) 的. 根据性质 (ii), $(D^\alpha u)^\wedge = (i\xi)^\alpha \widehat{u}$, 于是 $(i\xi)^\alpha \widehat{u}$ 在远处衰减, 其中 $\widehat{u}$ 在物理中称为物理空间函数 u 在频率空间中的对应. 这个衰减性说明当 $|\xi|$ 足够大时, $\widehat{u}$ 至少比 $(i\xi)^\alpha$ 更快趋零, 而后者是 α 阶多项式, 这便在一定程度上阐明了 $\widehat{u}$ 的衰减速度. 徐桂香老师上课所讲的" 物理空间的光滑 (物理空间函数具有 α 阶导数) 对应频率空间的衰减 (频率空间函数衰减速度至少快于 α 阶多项式)" 正是此意.

9.3.1.2 应用

Fourier 变换 $\mathcal{F}$ 在学习线性常系数 PDE 时是威力极其强大的工具.

例 9.5(Bessel 势函数) 下面首先研究 PDE

$$-\Delta u + u = f, \quad \text{在 } \mathbb{R}^n \text{ 中}$$

其中 $f \in L^2(\mathbb{R}^n)$. 现在希望找到 u 的精确表达式, 利用定理(9.3.2)(ii), 对方程两边作 Fourier 变换有

$$(1 + |y|^2)\widehat{u}(y) = \widehat{f}(y), \quad y \in \mathbb{R} \tag{9.77}$$

Fourier 变换的作用就是把 PDE 变成了代数方程(9.77), 易解得:

$$\widehat{u}=\frac{\widehat{f}}{1+|y|^2}$$

从而

$$u=\mathcal{F}^{-1}(\frac{\widehat{f}}{1+|y|^2}) \tag{9.78}$$

现在仅剩的问题就是把(9.78)右式写成更具体的形式. 考虑定理(9.3.2)(iii), 知可将(9.78)写成:

$$u=\frac{f*B}{(2\pi)^{\frac{n}{2}}} \tag{9.79}$$

其中

$$\widehat{B}=\frac{1}{1+|y|^2} \tag{9.80}$$

尽管 $\widehat{B}$ 并不在 L^1 或 L^2 中, 我们还是在形式上来求解 B. 因为对任意 $a>0$ 都有 $\frac{1}{a}=\int_0^\infty e^{-ta}dt$, 令 $a=1+|y|^2$ 得

$$\frac{1}{1+|y|^2}=\int_0^\infty e^{-t(1+|y|^2)}dt$$

进而

$$B=(\frac{1}{1+|y|^2})^\vee=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{ix\cdot y}(\int_0^\infty e^{-t(1+|y|^2)}dt)dy=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_0^\infty e^{-t}(\int_{\mathbb{R}^n}e^{ix\cdot y-t|y|^2}dy)dt \tag{9.81}$$

下面计算 $\int_{\mathbb{R}^n}e^{ix\cdot y-t|y|^2}dy$. 若取 $a,b\in\mathbb{R},b>0$, 设 $z=b^{\frac{1}{2}}x-\frac{a}{2b^{\frac{1}{2}}}i$, 则

$$\int_{-\infty}^\infty e^{iax-bx^2}dx=\frac{e^{-\frac{a^2}{4b}}}{b^{\frac{1}{2}}}\int_\Gamma e^{-z^2}dz$$

其中 Γ 表示复平面中的曲线 $\{\mathrm{Im}(z)=-\frac{a}{2b^{\frac{1}{2}}}\}$. 根据留数定理有

$$\int_\Gamma e^{-z^2}dz=\int_{-\infty}^\infty e^{-x^2}dx=\pi^{\frac{1}{2}}$$

从而

$$\int_{-\infty}^\infty e^{iax-bx^2}dx=e^{-\frac{a^2}{4b}}(\frac{\pi}{b})^{\frac{1}{2}} \tag{9.82}$$

作替换 $a\mapsto x_j,x\mapsto y_j(j=1,2,\cdots,n),b\mapsto t$ 并运用 Fubini 定理, 由(9.82)式得

$$\int_{\mathbb{R}^n}e^{ix\cdot y-t|y|^2}dy=\prod_{j=1}^n\int_{-\infty}^\infty e^{ix_jy_j-ty_j^2}dy_j=(\frac{\pi}{t})^{\frac{n}{2}}e^{-\frac{|x|^2}{4t}} \tag{9.83}$$

从而由(9.81),(9.83)式得

$$B(x)=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_0^\infty e^{-t}\cdot(\frac{\pi}{t})^{\frac{n}{2}}e^{-\frac{|x|^2}{4t}}dt=\frac{1}{2^{\frac{n}{2}}}\int_0^\infty\frac{e^{-t-\frac{|x|^2}{4t}}}{t^{\frac{n}{2}}}dt,\quad x\neq 0 \tag{9.84}$$

函数 B 就称为 Bessel 势函数. 将 B 代回(9.79)式可得

$$u(x)=\frac{1}{(4\pi)^{\frac{n}{2}}}\int_0^\infty\int_{\mathbb{R}^n}\frac{e^{-t-\frac{|x-y|^2}{4t}}}{t^{\frac{n}{2}}}f(y)dydt,\quad x\in\mathbb{R}^n \tag{9.85}$$

□

例 9.6(热方程的基本解) 再次考虑热方程的初值问题:

$$\begin{cases}u_t-\Delta u=0, & \text{在 } \mathbb{R}^n\times(0,\infty) \text{ 内}\\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases} \tag{9.86}$$

下面介绍利用 $\widehat{u}$ 求解问题(9.86)的方法, 其中 u 的 Fourier 变换仅对空间变量 x 作, 得到:

$$\begin{cases}\widehat{u}_t+|y|^2\widehat{u}=0, & t>0\\ \widehat{u}=\widehat{g}, & t=0\end{cases}$$

解关于 $\widehat{u}$ 与 t 的 ODE 得

$$\widehat{u}=e^{-t|y|^2}\widehat{g}$$

进而 $u=(e^{-t|y|^2}\widehat{g})^{\vee}$, 在定理(9.3.2)(iii) 两边作 Fourier 逆变换, 可记

$$u=\frac{g*F}{(2\pi)^{\frac{n}{2}}} \tag{9.87}$$

其中 $\widehat{F}=e^{-t|y|^2}$, 进一步根据(9.83)式有

$$F=\mathcal{F}^{-1}(e^{-t|y|^2})=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{ix\cdot y-t|y|^2}dy=\frac{1}{(2t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}$$

将上式代回(9.87)式得

$$u(x,t)=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}}g(y)dy,\quad x\in\mathbb{R}^n,t>0 \tag{9.88}$$

这与前面热方程一节得到的结果是一致的. 从而 Fourier 变换提供了一个新的导出热方程基本解的方法. □

课堂笔记 (2023.11.30)

- (关于 Bessel 势的注记) 徐桂香老师在课堂上给出了(9.85)式在 $n=3$ 时的进一步表达式, 这需要首先计算下述积分:

$$\begin{aligned}\int_0^{\infty}e^{-(x^2+\frac{a^2}{x^2})}dx&=e^{-2a}\int_0^{\infty}e^{-(x-\frac{a}{x})^2}dx=e^{-2a}\int_{-\infty}^{\infty}e^{-t^2}d(\frac{t+\sqrt{t^2+4a}}{2})\\&=e^{-2a}\int_{-\infty}^{\infty}e^{-t^2}(\frac{1}{2}+\frac{t}{2\sqrt{t^2+4a}})dt=\frac{1}{2}e^{-2a}(\int_{-\infty}^{\infty}e^{-t^2}dt+\int_{-\infty}^{\infty}e^{-t^2}\frac{t}{\sqrt{t^2+4a}}dt)\\&\overset{(i)}{=}\frac{\sqrt{\pi}}{2}e^{-2a}\end{aligned}$$

其中 $a>0$, (i) 是因为 $e^{-t^2}\frac{t}{\sqrt{t^2+4a}}$ 是奇函数. 于是当 $n=3$, (9.85)式即

$$\begin{aligned}u(x)&=\frac{1}{(4\pi)^{\frac{3}{2}}}\int_0^{\infty}\int_{\mathbb{R}^3}\frac{e^{-t-\frac{|x-y|^2}{4t}}}{t^{\frac{3}{2}}}f(y)dydt=\frac{1}{(4\pi)^{\frac{3}{2}}}\int_{\mathbb{R}^n}f(y)dy\int_0^{\infty}(-2)e^{-t-\frac{|x-y|^2}{4t}}dt^{-\frac{1}{2}}\\&=\frac{1}{4\pi^{\frac{3}{2}}}\int_{\mathbb{R}^n}f(y)dy\int_{\infty}^{0}(-e^{-(\frac{1}{p^2}+\frac{|x-y|^2}{4}p^2)})dp=\frac{1}{4\pi^{\frac{3}{2}}}\int_{\mathbb{R}^n}f(y)dy\int_0^{\infty}\frac{2}{|x-y|}e^{-(q^2+\frac{|x-y|^2}{4q^2})}dq\\&=\frac{1}{4\pi^{\frac{3}{2}}}\frac{\pi^{\frac{1}{2}}}{2}\int_{\mathbb{R}^n}f(y)\cdot\frac{2}{|x-y|}\cdot e^{-2\cdot\frac{|x-y|}{2}}dy=\frac{1}{4\pi}\int_{\mathbb{R}^n}\frac{e^{-|x-y|}}{|x-y|}f(y)dy\end{aligned}$$

从而在 $n=3$ 时 PDE$-\Delta u+u=f$ 有解

$$u(x)=\frac{1}{4\pi}\int_{\mathbb{R}^n}\frac{e^{-|x-y|}}{|x-y|}f(y)dy.$$

- (关于拟微分算子的注记) Bessel 势函数的研究背景是 PDE

$$-\Delta u+u=f$$

这个 PDE 可以另写成 $(1-\Delta)u=f$, 于是在形式上如果把 $-\Delta$ 看成常数, 它对应的解就是 $u=(1-\Delta)^{-1}f$. 前面在讨论平面波法的时候曾经介绍过 **[HJR]** 给出的算子表示法(9.40), 这里同样可以作此讨论. 因为在 PDE 两边作 Fourier 变换容易得到

$$\mathcal{F}[u](\xi)=\frac{1}{1+|\xi|^2}\cdot\mathcal{F}[f](\xi)$$

再作 Fourier 逆变换即得

$$u(x)=\mathcal{F}^{-1}[\frac{1}{1+|\circledast|^2}\cdot\mathcal{F}[f]](x)$$

于是仿照(9.40)式, 可以记

$$u(x)=(1+|D|^2)^{-1}f(x)$$

但形式上这便要求 $|D|^2 = -\Delta$, 虽说可以用 $-\Delta$ 是正算子来解释, 但在符号上还是会略显别扭, 仿佛左式应该是 $(i|D|)^2$ 才会更顺眼, 下面结合 Bessel 势对这个表示方法作一些正式介绍.

[**MCX1**] 的算子半群的乘子刻画一节给出了下述例子: 对于一般线性发展方程 Cauchy 问题

$$\begin{cases} u_t(x,t) - Q(\partial)u(x,t) = 0 \\ u(x,0) = \varphi(x) \end{cases}$$

其解可以写成 $u(x,t) = \mathcal{F}^{-1}[e^{Q(i\circledast)t} \cdot \mathcal{F}[\varphi]](x)$, 这暗示我们在还原算子表示的时候应该用 Fourier 变换后得到的最初形式去讨论. 在 Bessel 势的例子中, PDE 两边同时作 Fourier 变换得到的本质上是

$$\mathcal{F}[u](\xi) = \frac{1}{1-(i\xi)^2} \cdot \mathcal{F}[f](\xi)$$

Fourier 反演有

$$u(x) = \mathcal{F}^{-1}[\frac{1}{1-(i\circledast)^2} \cdot \mathcal{F}[f]](x) \tag{9.89}$$

一般来讲, 如果对给定的函数 f 与 g, 可以定义算子

$$T_g(f) = \mathcal{F}^{-1}[g(\circledast) \cdot \mathcal{F}[f]]$$

就称 T_g 是 g 所对应的 Fourier 乘子, 关于 Fourier 乘子的一些具体性质参见 [**LG1**]. 把 Fourier 乘子 T_g 与微分算子 ∇ 联系起来的课题称为拟微分算子 (Pseudodifferential Operators) 理论, 下面给出 [**TL2**] 的相关介绍. 回忆 Fourier 反演:

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

若在两边同时对 x 求导, 有

$$D^\alpha f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} (i\xi)^\alpha \widehat{f}(\xi) e^{ix\cdot\xi} d\xi$$

为了在右式消除这个" 不太自然的"i, 对上式稍作处理可得

$$(\frac{D}{i})^\alpha f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \xi^\alpha \widehat{f}(\xi) e^{ix\cdot\xi} d\xi$$

拟微分算子理论中的 D 其实表示的是这里的 $\frac{D}{i}$, 具体来说:

$$D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}, \quad D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$$

报告Fourier 分析及其在偏微分方程中的应用 | 张平院士中给出的记法是 $D = \frac{\nabla}{i}$. 于是如果要把前面定义的算子 T_m 写成拟微分算子的形式, 根据这个记法, 其实应该是把 $\circledast$ 替换成 $D = \frac{\nabla}{i}$, 亦即:

$$T_g(f) = g(D)f = g(\frac{\nabla}{i})f$$

根据上面的介绍, 特别对(9.89)式有

$$u(x) = (1-(i\cdot\frac{\nabla}{i})^2)^{-1}f = (1-\nabla^2)^{-1}f = (1-\Delta)^{-1}f$$

这便与形式上的记法契合了. 基于上述想法, 在 [**LG2**],[**HJR**] 等书中均直接约定 $(-\Delta)^{\frac{z}{2}}$ 正是 $|\xi|^z$ 对应的拟微分算子 (其中 $z \in \mathbb{C}$).

- (关于 Bessel 势的推广与 Riesz 势) 徐桂香老师的课堂与 [**LG2**] 均介绍了更广义的 Bessel 势. 书上的 Bessel 势旨在解决 $(1-\Delta)u = f$, 如果左式 $1-\Delta$ 这一项的幂次并非 1, 而是 $\frac{\alpha}{2}(\mathrm{Re}\,\alpha > 0)$, 对应的 PDE$(1-\Delta)^{\frac{\alpha}{2}}u = f$ 依旧可以形式求解得到 $u = (1-\Delta)^{-\frac{\alpha}{2}}f$. 现在用拟微分算子的表示把这个表达式还原得到

$$u(x) = \mathcal{F}^{-1}[(1+|\circledast|^2)^{-\frac{\alpha}{2}}\mathcal{F}[f]](x) = (2\pi)^{-\frac{n}{2}}(\mathcal{F}^{-1}[(1+|\circledast|^2)^{-\frac{\alpha}{2}}] * f)(x)$$

若设 $\widehat{B_\alpha}(\xi)=\frac{1}{(1+|\xi|^2)^{\frac{\alpha}{2}}}$, 对应得到的

$$B_\alpha(x)=\mathcal{F}^{-1}[\frac{1}{(1+|\circledast|^2)^{\frac{\alpha}{2}}}](x)=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\frac{e^{ix\cdot\xi}}{(1+|\xi|^2)^{\frac{\alpha}{2}}}d\xi \tag{9.90}$$

也称为 Bessel 势. 回忆 Γ 函数的定义:

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx$$

如果令 $x=tA(A>0)$, 则有

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}A^\alpha e^{-tA}dt\Rightarrow A^{-\alpha}=\frac{1}{\Gamma(\alpha)}\int_0^\infty e^{-tA}t^\alpha\frac{dt}{t}$$

现在把 A 换成 $1+|\xi|^2$, α 换成 $\frac{\alpha}{2}$, 则有

$$(1+|\xi|^2)^{\frac{\alpha}{2}}=\frac{1}{\Gamma(\frac{\alpha}{2})}\int_0^\infty e^{-t-t|\xi|^2}t^{\frac{\alpha}{2}}\frac{dt}{t}$$

代回(9.90)式得到

$$\begin{aligned}B_\alpha(x)&=\frac{1}{(2\pi)^{\frac{n}{2}}}\frac{1}{\Gamma(\frac{\alpha}{2})}\int_{\mathbb{R}^n}e^{ix\cdot\xi}(\int_0^\infty e^{-t-t|\xi|^2}t^{\frac{\alpha}{2}-1}dt)d\xi\\&=\frac{1}{(2\pi)^{\frac{n}{2}}\Gamma(\frac{\alpha}{2})}\int_0^\infty e^{-t}t^{\frac{\alpha}{2}-1}dt\int_{\mathbb{R}^n}e^{ix\cdot\xi}e^{-t|\xi|^2}d\xi\end{aligned}$$

对于 $\int_{\mathbb{R}^n}e^{ix\cdot\xi}e^{-t|\xi|^2}d\xi$, 可以用 $e^{-|\xi|^2}$ 的 Fourier 反演性质直接得到结果, 不过这里还是再演示一遍:

$$\begin{aligned}\int_{\mathbb{R}^n}e^{ix\cdot\xi-t|\xi|^2}d\xi&=t^{-\frac{n}{2}}\int_{\mathbb{R}^n}e^{i\frac{x}{\sqrt{t}}\cdot\tau-|\tau|^2}d\tau=t^{-\frac{n}{2}}e^{-\frac{|x|^2}{4t}}\prod_{i=1}^n\int_{\mathbb{R}}e^{-(\tau_i-\frac{ix_i}{\sqrt{t}})^2}d\tau_i\\&=t^{-\frac{n}{2}}e^{-\frac{|x|^2}{4t}}\prod_{i=1}^n\int_{\mathbb{R}}e^{-\mu_i^2}d\mu_i=(\frac{\pi}{t})^{\frac{n}{2}}e^{-\frac{|x|^2}{4t}}\end{aligned}$$

于是

$$B_\alpha(x)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{\alpha}{2})}\int_0^\infty e^{-t}t^{\frac{\alpha-n}{2}-1}e^{-\frac{|x|^2}{4t}}dt$$

代入 $\alpha=2$ 即得

$$B_2(x)=\frac{1}{2^{\frac{n}{2}}}\int_0^\infty e^{-t}t^{-\frac{n}{2}}e^{-\frac{|x|^2}{4t}}dt$$

这与书上是吻合的.

如果考虑的 PDE 不是 $(1-\Delta)^{\frac{\alpha}{2}}u=f$, 而是 $(-\Delta)^{\frac{\alpha}{2}}u=f(0<\operatorname{Re}\alpha<n)$, 事实上也可以仿照上述过程进行推导. 形式求解得到 $u=(-\Delta)^{-\frac{\alpha}{2}}f$, 还原有

$$u(x)=\mathcal{F}^{-1}[(|\circledast|^2)^{-\frac{\alpha}{2}}\mathcal{F}[f]](x)=(2\pi)^{-\frac{n}{2}}(\mathcal{F}^{-1}[|\circledast|^{-\alpha}]*f)(x)$$

若设 $\widehat{R_\alpha}(\xi)=\frac{1}{|\xi|^\alpha}$, 对应得到的

$$R_\alpha(x)=\mathcal{F}^{-1}[\frac{1}{|\circledast|^\alpha}](x)=\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\frac{e^{ix\cdot\xi}}{|\xi|^\alpha}d\xi$$

就称为 Riesz 势. Riesz 势的具体表达式稍显复杂, **[LG1]** 中记录了求解过程. 在现代 PDE 理论中, 算子

$$(-\Delta)^{\frac{s}{2}}:f\mapsto\mathcal{F}^{-1}[|\circledast|^{-s}]*f$$

就称为分数阶 Laplace 算子.

- (关于热方程基本解在拟微分算子上的注记) 对热方程初值问题

$$\begin{cases}u_t-\Delta u=0, & 在\ \mathbb{R}^n\times(0,\infty)\ 内\\ u=g, & 在\ \mathbb{R}^n\times\{t=0\}\ 上\end{cases}$$

形式上求解可得 $u(x,t)=e^{t\Delta}g(x)$, 依照前面介绍的拟微分算子表示还原有

$$u(x,t)=e^{-t(-\Delta)}g(x)=\mathcal{F}^{-1}[e^{-t|\circledast|^2}\mathcal{F}[g]](x,t)=(2\pi)^{-\frac{n}{2}}(\mathcal{F}^{-1}[e^{-t|\circledast|^2}]*g)(x,t)$$

前面在讨论 Bessel 势的推广时曾经演示了:

$$\int_{\mathbb{R}^n}e^{ix\cdot\xi-t|\xi|^2}d\xi=(\frac{\pi}{t})^{\frac{n}{2}}e^{-\frac{|x|^2}{4t}}\Rightarrow\mathcal{F}^{-1}[e^{-t|\circledast|^2}](x)=\frac{1}{(2t)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4t}}$$

于是

$$u(x,t)=(2\pi)^{-\frac{n}{2}}\int_{\mathbb{R}^n}\frac{1}{(2t)^{\frac{n}{2}}}e^{-\frac{|x-y|^2}{4t}}g(y)dy=\frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4t}}g(y)dy$$

这再一次回到了热方程基本解的内容. 在算子半群理论中, $\{e^{t\Delta}\}_{t\geq 0}$ 称为热半群**DY**, 其中的每一个算子 $e^{\varepsilon\Delta}$ 都定义为:

$$e^{\varepsilon\Delta}:g\mapsto\frac{1}{(4\pi\varepsilon)^{\frac{n}{2}}}e^{-\frac{|\circledast|}{4\varepsilon}}*g$$

- (关于二阶椭圆发展方程的补充) 考虑半平面上的二阶椭圆发展方程

$$\begin{cases}\partial_t^2u(x,t)+\Delta u(x,t)=0\\ u(x,t=0)=f(x)\end{cases}$$

在问题两边关于 x 作 Fourier 变换得到

$$\begin{cases}\partial_t^2\widehat{u}(\xi,t)-|\xi|^2\widehat{u}(\xi,t)=0\\ \widehat{u}(\xi,t=0)=\widehat{f}(\xi)\end{cases}$$

注意得到的 ODE 可以写成

$$(\partial_t-|\xi|)(\partial_t+|\xi|)\widehat{u}(\xi,t)=0$$

但形如 $e^{t|\xi|}$ 的解没有太大物理意义, 故不妨设 $(\partial_t-|\xi|)\widehat{u}(\xi,t)\neq 0$ 对任意 $t\in\mathbb{R}$ 均成立, 于是求解 $(\partial_t+|\xi|)\widehat{u}(\xi,t)=0$ 得到

$$\widehat{u}(\xi,t)=e^{-t|\xi|}\widehat{f}(\xi)\Rightarrow u(x,t)=(2\pi)^{-\frac{n}{2}}(\mathcal{F}^{-1}[e^{-t|\circledast|}]*f)(x,t)$$

前面介绍过在拟微分算子理论中, 约定 $|\xi|^z$ 对应的拟微分算子为 $(-\Delta)^{\frac{z}{2}}$, 于是上式可表为

$$u(x,t)=(2\pi)^{-\frac{n}{2}}e^{-t(-\Delta)^{\frac{1}{2}}}f(x)=(2\pi)^{-\frac{n}{2}}e^{-t\sqrt{-\Delta}}f(x)$$

在算子半群理论中, $\{e^{-t\sqrt{-\Delta}}\}_{t\geq 0}$ 称为 Poisson 半群, 其中每一个算子仿照热半群元素都有类似定义. ♠

例 9.7(Schrödinger 方程的基本解) 下面研究 Schrödinger 方程的初边值问题:

$$\begin{cases}iu_t+\Delta u=0, & \text{在 } \mathbb{R}^n\times(0,\infty)\text{ 内}\\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\}\text{ 上}\end{cases}\tag{9.91}$$

其中 u,g 均为复值函数.

如果形式上在(9.88)右式把 t 换成 it, 则得到表达式:

$$u(x,t)=\frac{1}{(4\pi it)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{\frac{i|x-y|^2}{4t}}g(y)dy,\quad x\in\mathbb{R}^n,t>0\tag{9.92}$$

其中 $i^{\frac{1}{2}}$ 记为 $e^{\frac{i\pi}{4}}$. 显见只要 $g\in L^1(\mathbb{R}^n)$, 这个表达式就对全体 $t>0$ 都成立. 进一步, 如果 $|y|^2g\in L^1(\mathbb{R}^n)$, 则可通过直接计算得到在 $\mathbb{R}^n\times(0,\infty)$ 内 u 满足 $iu_t+\Delta u=0$. (这里先不讨论 $u(\cdot,t)\to g(t\to 0^+)$ 的情况, 后面的章节与习题都会提到.)

下面把(9.92)式重写成

$$u(x,t)=\frac{e^{\frac{i|x|^2}{4t}}}{(4\pi it)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{\frac{-ix\cdot y}{2t}}e^{\frac{i|y|^2}{4t}}g(y)dy$$

因为 $|e^{\frac{i|x|^2}{4t}}, e^{\frac{i|y|^2}{4t}}| = 1$[10], 可类似于 Plancherel 定理(9.3.1)的证明过程说明: 若 $g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, 则

$$\|u(\cdot,t)\|_{L^2(\mathbb{R}^n)} = \|g\|_{L^2(\mathbb{R}^n)}, \quad t > 0 \tag{9.93}$$

从而映射 $g \mapsto u(\cdot,t)$ 保 L^2 范数. 故仿照 Fourier 变换延拓的方法, 可将(9.92)式延拓成对 $g \in L^2(\mathbb{R}^n)$ 也成立. □

称

$$\Psi(x,t) := \frac{1}{(4\pi i t)^{\frac{n}{2}}} e^{\frac{i|x|^2}{4t}}, \quad x \in \mathbb{R}^n, t \neq 0 \tag{9.94}$$

为 Schrödinger 方程的基本解. 注意(9.92)式, 亦即 $u = g * \Psi$ 对全体时间 $t \neq 0$ 都成立, 就算 $t < 0$ 也成立. 从而事实上我们解出了

$$\begin{cases} iu_t + \Delta u = 0, & \text{在 } \mathbb{R}^n \times (-\infty,\infty) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \tag{9.95}$$

更特别地, Schrödinger 方程关于时间可逆, 这是热方程所做不到的.

例 9.8(波方程) 下面研究波方程的初边值问题

$$\begin{cases} u_{tt} - \Delta u = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, u_t = h, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \tag{9.96}$$

其中为简便, 就设初始速度为零. 同样取 $\widehat{u}$ 为 u 关于变量 $x \in \mathbb{R}^n$ 的 Fourier 变换, 则

$$\begin{cases} \widehat{u}_{tt} + |y|^2 \widehat{u} = 0, & t > 0 \\ \widehat{u} = \widehat{g}, \widehat{u}_t = \widehat{h}, & t = 0 \end{cases} \tag{9.97}$$

对每个固定的 $y \in \mathbb{R}$ 而言, 上述方程是 ODE, 解得

$$\widehat{u} = \widehat{g}\cos(t|y|) + \frac{\widehat{h}}{|y|}\sin(t|y|) \tag{9.98}$$

两边求 Fourier 逆变换得

$$u(x,t) = [\widehat{g}\cos(t|y|) + \frac{\widehat{h}}{|y|}\sin(t|y|)]^{\vee}$$

特别当 $h \equiv 0$ 时, 有

$$u(x,t) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ix\cdot y}\widehat{g}(y)\cos(t|y|)dy = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \frac{\widehat{g}(y)}{2}(e^{i(x\cdot y + t|y|)} + e^{i(x\cdot y - t|y|)})dy, \quad x \in \mathbb{R}^n, t \geq 0 \tag{9.99}$$

后面将在某个渐近极限中回过头来讨论这个式子.

课堂笔记 (2023.12.4)

- (关于标蓝部分的解释) 对 Schrödinger 方程两边作 Fourier 变换可得

$$\begin{cases} i\widehat{u}_t - |\xi|^2 \widehat{u} = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ \widehat{u} = \widehat{g}, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

解得

$$\widehat{u}(\xi,t) = e^{\frac{|\xi|^2}{i}t}\widehat{g}(\xi) = e^{-it|\xi|^2}\widehat{g}(\xi) \tag{9.100}$$

而回忆热方程, 在两边作 Fourier 变换有:

$$\begin{cases} u_t - \Delta u = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases} \Rightarrow \begin{cases} \widehat{u}_t + |\xi|^2 \widehat{u} = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ \widehat{u} = \widehat{g}, & \text{在 } \mathbb{R}^n \times \{t=0\} \text{ 上} \end{cases}$$

[10] 两者的模均为 1, 亦即 $|e^{i\frac{|x|^2}{4t}}| = |e^{i\frac{|y|^2}{4t}}| = 1$.

解得

$$\widehat{u}(\xi,t)=e^{-t|\xi|^2}\widehat{g}(\xi) \tag{9.101}$$

形式上,从(9.101)式到(9.100)式只需要在右式把 t 换成 it,于是就有了把从(9.101)式解出的 $u(\xi,t)$ 表达式中的 t 形式上换成 it 的思路.

- (关于 Schrödinger 方程解的质量守恒) 对于 Schrödinger 方程的解

$$u(x,t)=\frac{1}{(4\pi it)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{-\frac{|x-y|^2}{4it}}g(y)dy$$

知

$$\widehat{u}(\xi,t)=e^{-it|\xi|^2}\widehat{g}(\xi)$$

于是

$$\|\widehat{u}(\circledast,t)\|_{L^2_\xi}=\|e^{-it|\circledast|^2}\widehat{g}\|_{L^2}$$

而一方面,根据 Plancherel 定理有

$$\|\widehat{u}(\circledast,t)\|_{L^2_\xi}=\|u(\circledast,t)\|_{L^2_x}$$

另一方面有

$$\|e^{-it|\circledast|^2}\widehat{g}\|_{L^2}=\|\widehat{g}\|_{L^2}=\|g\|_{L^2}$$

得到式子

$$\|u(\circledast,t)\|_{L^2_x}=\|g\|_{L^2}$$

注意式子右端与 t 无关,从而左端也应与 t 无关,这便有

$$\|u(\circledast,t=0)\|_{L^2_x}=\|u(\circledast,t)\|_{L^2_x}=\|g\|_{L^2}$$

如果把 $|u|^2$ 视作质量函数,上式表达的就正是质量守恒定律.

- (关于 Schrödinger 方程解的更精细估计) 根据习题9.1,对 Schrödinger 方程的解有初步估计:

$$\|u(\circledast,t)\|_{L^\infty_x}\le\frac{1}{(4\pi|t|)^{\frac{n}{2}}}\|g\|_{L^1}$$

在现代 PDE 理论中,若存在常数 C 使得 $f\le Cg$,就称 f 被 g 控制,记作 $f\lesssim g$,于是上式可写作 $\|u(\circledast,t)\|_{L^\infty_x}\lesssim|t|^{-\frac{n}{2}}\|g\|_{L^1}$. 此时同时有

$$\|u(\circledast,t)\|_{L^\infty_x}\lesssim|t|^{-\frac{n}{2}}\|g\|_{L^1}$$
$$\|u(\circledast,t)\|_{L^2_x}=\|g\|_{L^2}$$

如果将 $T:L^1+L^2\ni g\mapsto u=\Psi*g$ 视作线性算子,则上式可写作

$$\|Tg\|_{L^\infty}\lesssim|t|^{-\frac{n}{2}}\|g\|_{L^1}$$
$$\|Tg\|_{L^2}=\|g\|_{L^2}$$

于是由 Riesz-Thorin 插值定理**??**,对任意 $0<\lambda<1$,只要

$$\frac{1}{p}=\frac{1-\lambda}{\infty}+\frac{\lambda}{2}=\frac{\lambda}{2}$$
$$\frac{1}{q}=\frac{1-\lambda}{1}+\frac{\lambda}{2}=1-\frac{\lambda}{2}$$

就有

$$\|Tg\|_{L^p}\lesssim|t|^{-\frac{n}{2}(1-\lambda)}\|g\|_{L^q}$$

如果想要估计特定阶数,譬如 $\|Tg\|_{L^r}$,就令 $p=r$,解得

$$q=r',\quad \lambda=\frac{2}{r}$$

于是

$$\|Tg\|_{L^r} \lesssim |t|^{-\frac{n}{2}(1-\frac{2}{r})}\|g\|_{L^{r'}}$$

这便得到了 u 在 $2 \le r \le \infty$ 之间的任意阶估计.

- (关于 Fourier 变换的旋转不变性) Fourier 变换是保径向函数的, 亦即若 $F(x) = f(r), r = |x|(x \in \mathbb{R}^n)$, 则 $\widehat{F}(\xi)$ 也是径向函数. 这只需证明

$$\forall \xi_1, \xi_2 \in \mathbb{R}^n(|\xi_1| = |\xi_2| \Rightarrow \widehat{F}(\xi_1) = \widehat{F}(\xi_2))$$

把 $\widehat{F}(\xi_1), \widehat{F}(\xi_2)$ 按定义写开有

$$\widehat{F}(\xi_1) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi_1}F(x)dx$$
$$\widehat{F}(\xi_2) = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-ix\cdot\xi_2}F(x)dx$$

既然 $|\xi_1| = |\xi_2|$, 知存在正交变换 O 使得 $\xi_2 = O\xi_2$, 于是 $\langle x, \xi_2\rangle = \langle x, O\xi_1\rangle = \langle O^T x, \xi_1\rangle$, 得到

$$\begin{aligned}
\widehat{F}(\xi_2) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\langle x,\xi_2\rangle}F(x)dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\langle O^T x,\xi_1\rangle}F(x)dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\langle y,\xi_1\rangle}F((O^T)^{-1}y)d((O^T)^{-1}y) \\
&\overset{\text{(i)}}{=} \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-i\langle y,\xi_1\rangle}F(y)dy = \widehat{F}(\xi_1)
\end{aligned}$$

其中 (i) 是因为 $|(O^T)^{-1}y| = |y|$, 于是 F 作为径向函数, 理应有 $F((O^T)^{-1}y) = F(y)$.

- (关于波方程在拟微分算子上的注记) 已经知道对于波方程初值问题

$$\begin{cases} u_{tt} - \Delta u = 0, & \text{在 } \mathbb{R}^n \times (0,\infty) \text{ 内} \\ u = g, u_t = h, & \text{在 } \mathbb{R}^n \times \{t = 0\} \text{ 上} \end{cases}$$

可以利用 Fourier 变换方法求得

$$\begin{aligned}
u(x,t) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{g}(\xi)\cos(t|\xi|)d\xi + \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}\widehat{h}(\xi)\frac{\sin(t|\xi|)}{|\xi|}d\xi \\
&= \mathcal{F}^{-1}[\cos(t|\circledast|)\mathcal{F}[g]](x) + \mathcal{F}^{-1}[\frac{\sin(t|\circledast|)}{|\circledast|}\mathcal{F}[h]](x)
\end{aligned}$$

前面介绍过 $|\xi|^z$ 对应的拟微分算子为 $(-\Delta)^{\frac{z}{2}}$, 于是此处 $|\xi|$ 对应 $\sqrt{-\Delta}$, 形式上记 $\sqrt{-\Delta} = \nabla$, 得到波方程的拟微分算子表示:

$$u(x,t) = \cos(t|\nabla|)\cdot g(x) + \frac{\sin(t|\nabla|)}{|\nabla|}\cdot h(x).$$

下面还原得到 $g = 0$ 时的 d'Alembert 公式

$$u(x,t) = \frac{1}{2}\int_{x-t}^{x+t} h(y)dy = t\fint_{[x-t,x+t]} h(y)dy$$

与 Kirchhoff 公式

$$u(x,t) = \fint_{\partial B(x,t)} th(y)dy.$$

在此之前先承认一个公式: $\widehat{\delta}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \Rightarrow \widetilde{1}(\xi) = (2\pi)^{\frac{n}{2}}\delta(x)$, 其中 δ 是 Dirac 测度.

当 $n = 1$ 时, 因为

$$\frac{e^{-ix\xi}}{-i\xi}\Big|_{-1}^{1} = \frac{e^{-i\xi} - e^{i\xi}}{-i\xi} = \frac{2\sin\xi}{\xi}$$

故

$$\frac{\sin\xi}{\xi}=\frac{\sin|\xi|}{|\xi|}=\frac{1}{2}\int_{-1}^{1}e^{-ix\xi}dx\Rightarrow\frac{\sin t\xi}{\xi}=\frac{t}{2}\int_{-1}^{1}e^{-ixt\xi}dx$$

于是

$$\begin{aligned}u(x,t)=\frac{\sin(t|\nabla|)}{|\nabla|}\cdot h(x)&=\frac{1}{(2\pi)^{\frac{1}{2}}}\int_{\mathbb{R}}e^{ix\xi}\frac{\sin(t|\xi|)}{|\xi|}\widehat{h}(\xi)d\xi\\&=\frac{1}{(2\pi)^{\frac{1}{2}}}\int_{\mathbb{R}}e^{ix\xi}(\frac{t}{2}\int_{-1}^{1}e^{-izt\xi}dz)\widehat{h}(\xi)d\xi\\&=\frac{t}{2(2\pi)^{\frac{1}{2}}}\int_{\mathbb{R}}e^{ix\xi}(\int_{-1}^{1}e^{-izt\xi}dz)\cdot\frac{1}{(2\pi)^{\frac{1}{2}}}(\int_{\mathbb{R}}e^{-iy\xi}h(y)dy)d\xi\\&=\frac{t}{2(2\pi)^{\frac{1}{2}}}\int_{-1}^{1}dz\int_{\mathbb{R}}h(y)dy\int_{\mathbb{R}}\frac{1}{(2\pi)^{\frac{1}{2}}}e^{ix\xi-izt\xi-iy\xi}d\xi\\&=\frac{t}{2(2\pi)^{\frac{1}{2}}}\int_{-1}^{1}dz\int_{\mathbb{R}}h(y)\widetilde{1}(x-zt-y)dy=\frac{t}{2}\int_{-1}^{1}dz\int_{\mathbb{R}}h(y)\delta(x-zt-y)dy\\&=\frac{t}{2}\int_{-1}^{1}h(x-zt)dz=\frac{1}{2}\int_{x-t}^{x+t}h(y)dy.\end{aligned}$$

当 $n=3$ 时, 首先承认

$$\frac{\sin(t|\xi|)}{|\xi|}=\frac{1}{|\partial B(0,1)|t}\int_{\partial B(0,t)}e^{-iy\cdot\xi}d\sigma(y)$$

于是

$$\begin{aligned}u(x,t)=\frac{\sin(t|\nabla|)}{|\nabla|}\cdot h(x)&=\frac{1}{(2\pi)^{\frac{3}{2}}}\int_{\mathbb{R}^3_\xi}e^{ix\cdot\xi}\frac{\sin(t|\xi|)}{|\xi|}\widehat{h}(\xi)d\xi\\&=\frac{1}{(2\pi)^{\frac{3}{2}}}\int_{\mathbb{R}^3_\xi}e^{ix\cdot\xi}(\frac{1}{|\partial B(0,1)|t}\int_{\partial B(0,t)}e^{-iy\cdot\xi}d\sigma(y))(\frac{1}{(2\pi)^{\frac{3}{2}}}\int_{\mathbb{R}^3_z}e^{-iz\cdot\xi}h(z)dz)d\xi\\&=\frac{1}{|\partial B(0,1)|t}\frac{1}{(2\pi)^{\frac{3}{2}}}\int_{\partial B(0,t)}d\sigma(y)\int_{\mathbb{R}^3_z}h(z)dz\int_{\mathbb{R}^3_\xi}\frac{1}{(2\pi)^{\frac{3}{2}}}e^{ix\cdot\xi-iy\cdot\xi-iz\cdot\xi}d\xi\\&=\frac{1}{|\partial B(0,1)|t}\frac{1}{(2\pi)^{\frac{3}{2}}}\int_{\partial B(0,t)}d\sigma(y)\int_{\mathbb{R}^3_z}h(z)\widetilde{1}(x-y-z)dz\\&=\frac{1}{|\partial B(0,1)|t}\int_{\partial B(0,t)}d\sigma(y)\int_{\mathbb{R}^3_z}h(z)\delta(x-y-z)dz\\&=\frac{1}{|\partial B(0,1)|t}\int_{\partial B(0,t)}h(x-y)d\sigma(y)=\frac{t}{|\partial B(x,t)|}\int_{\partial B(x,t)}h(y)d\sigma(y)=t⨍_{\partial B(x,t)}h(y)d\sigma(y)\end{aligned}$$

再来说明

$$\frac{\sin(t|\xi|)}{|\xi|}=\frac{1}{|\partial B(0,1)|t}\int_{\partial B(0,t)}e^{-iy\cdot\xi}d\sigma(y)=\frac{1}{4\pi}\int_{\partial B(0,1)}e^{-ix\cdot\xi}d\sigma(x)$$

通过换元, 只需证明

$$\frac{\sin|\xi|}{|\xi|}=\frac{1}{4\pi}\int_{\partial B(0,1)}e^{-ix\cdot\xi}d\sigma(x)$$

注意左式是径向函数, 而对右式仿照证明 Fourier 变换旋转不变性的方法可得

$$|\xi_1|=|\xi_2|\Rightarrow\frac{1}{4\pi}\int_{\partial B(0,1)}e^{-ix\cdot\xi_1}d\sigma(x)=\frac{1}{4\pi}\int_{\partial B(0,1)}e^{-ix\cdot\xi_2}d\sigma(x)$$

于是只需验证满足 $|\xi|=\lambda$ 的一种情况, 特别取 $\xi=(\lambda,0,0)$ 知

$$\frac{1}{4\pi}\int_{\partial B(0,1)}e^{-ix\cdot\xi}d\sigma(x)=\frac{1}{4\pi t}\int_{\partial B(0,t)}e^{-ix_1\cdot\lambda}d\sigma(x)$$

作球坐标换元:

$$\begin{cases} x_1 = \cos\varphi_1 \\ x_2 = \sin\varphi_1\cos\varphi_2 \\ x_3 = \sin\varphi_1\sin\varphi_2 \\ \varphi_1 \in [0,\pi), \varphi_2 \in [0,2\pi) \end{cases}$$

知

$$\begin{aligned} \frac{1}{4\pi}\int_{\partial B(0,1)} e^{-ix_1\cdot\lambda}d\sigma(x) &= \frac{1}{4\pi}\int_0^\pi d\varphi_1\int_0^{2\pi} e^{-i\cos\varphi_1\cdot\lambda}\sin\varphi_1 d\varphi_2 \\ &= -\frac{1}{2}\int_0^\pi e^{-i\cos\varphi_1}\cdot\lambda d\cos\varphi_1 = \frac{\sin\lambda}{\lambda} \end{aligned}$$

命题即证, 至此便统一了波方程的拟微分算子表示, d'Alembert 公式与 Kirchhoff 公式.

- (关于输运方程在拟微分算子上的注记) 对输运方程的初值问题:

$$\begin{cases} u_t(x,t) + b\cdot\nabla u(x,t) = 0 \\ u(x,t=0) = g(x) \end{cases}$$

两边作 Fourier 变换得

$$\begin{cases} \widehat{u}_t(\xi,t) + b\cdot i\xi\widehat{u}(\xi,t) = 0 \\ \widehat{u}(\xi,t=0) = \widehat{g}(\xi) \end{cases}$$

解得

$$\widehat{u}(\xi,t) = e^{tb\cdot i\xi}\widehat{g}(\xi) \Rightarrow u(x,t) = e^{tb\cdot D}g(x)$$

另一方面, 把拟微分算子表示拆开有

$$\begin{aligned} u(x,t) &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{ix\cdot\xi}e^{-bt\cdot i\xi}\widehat{g}(\xi)d\xi \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{i(x-bt)}\widehat{g}(\xi)d\xi = \overline{\widehat{g}}(x-bt) = g(x-bt) \end{aligned}$$

这便统一了输运方程的拟微分算子表示与解的表示公式.

能量均分渐近线

设 $f, g, Dg \in L^2(\mathbb{R}^n)$, 回忆前面讨论四种基本 PDE 时提到的能量方法, 波方程(9.96)的解 u 的能量为:

$$E(t) := \frac{1}{2}\int_{\mathbb{R}^n}(u_t^2 + |Du|^2)dx, \quad t \geq 0$$

根据能量守恒定律, $E(t)$ 关于时间 t 是常数:

$$E(t) = E(0) = \frac{1}{2}\int_{\mathbb{R}^n}(h^2 + |Dg|^2)dx$$

作为表达式(9.98)的一个应用, 下面证明

$$\lim_{t\to\infty}\int_{\mathbb{R}^n}|Du|^2dx = \lim_{t\to\infty}u_t^2dx = E(0) \tag{9.102}$$

这说明在渐近意义下, 总能量平均地分配到势能和动能中.

要证明(9.102)式, 考虑利用(9.98)式与定理(9.3.2)(ii) 有:

$$\begin{aligned} \int_{\mathbb{R}^n}|Du|^2dx &= \int_{\mathbb{R}^n}|y|^2\cdot|\widehat{u}|^2dy \\ &= \int_{\mathbb{R}^n}(|y|^2|\widehat{g}|^2\cos^2(t|y|) + |\widehat{h}|^2\sin^2(t|y|))dy \\ &\quad + \int_{\mathbb{R}^n}\cos(t|y|)\sin(t|y|)|y|(\widehat{h}\overline{\widehat{g}} + \widehat{g}\overline{\widehat{h}})dy \end{aligned} \tag{9.103}$$

现若取 $f \in C_c^\infty(\mathbb{R}^n)$, 则[11]

$$\begin{aligned}\int_{\mathbb{R}^n} \cos(t|y|)\sin(t|y|) f dy &= \frac{1}{2}\int_{\mathbb{R}^n} \sin(2t|y|) f dy \\ &= \frac{1}{2}\int_0^\infty \sin(2tr)\left(\int_{\partial B(0,r)} f dS\right)dr \\ &= -\frac{1}{4t}\int_0^\infty \frac{d}{dr}(\cos(2tr))\left(\int_{\partial B(0,r)} f dS\right)dr \\ &= \frac{1}{4t}\int_0^\infty \cos(2tr)\frac{d}{dt}\left(\int_{\partial B(0,r)} f dS\right)dr = O(t^{-1})\end{aligned}$$

用光滑紧支函数逼近可积函数 $|y|(\widehat{h}\overline{\widehat{g}} + \widehat{g}\overline{\widehat{h}})$, 可得(9.103)式末尾的积分在 $t \to \infty$ 时趋零. 根据恒等式 $\cos^2(t|y|) = \frac{1}{2}(\cos(2t|y|)+1)$ 与 Riemann-Lebesgue 引理知:

$$\int_{\mathbb{R}^n} |y|^2|\widehat{g}|^2\cos^2(t|y|)dy \to \frac{1}{2}\int_{\mathbb{R}^n} |y|^2|\widehat{g}|^2 dy, \quad t \to \infty$$

类似有

$$\int_{\mathbb{R}^n} |\widehat{h}|^2\sin^2(t|y|)dy \to \frac{1}{2}\int_{\mathbb{R}^n} |\widehat{h}|^2 dy, \quad t \to \infty$$

从而由(9.103)式得

$$\lim_{t\to\infty}\int_{\mathbb{R}^n} |Du|^2 dx + \frac{1}{2}\int_{\mathbb{R}^n} (|y||\widehat{g}|^2 + |\widehat{h}|^2)dy = \frac{1}{2}\int_{\mathbb{R}^n} (|Dg|^2 + |h|^2)dx = E(0)$$

□

例 9.9(电报方程) 一维电报方程的初边值问题为:

$$\begin{cases} u_{tt} + 2du_t - u_{xx} = 0, & \text{在 } \mathbb{R}\times(0,\infty) \text{ 内} \\ u = g, u_t = h, & \text{在 } \mathbb{R}\times\{t=0\} \text{ 上}\end{cases}$$

其中 $d > 0$. $2du_t$ 这一项表示波传播的物理阻尼[12]. 对 u 关于 x 作 Fourier 变换得

$$\begin{cases} \widehat{u}_{tt} + 2d\widehat{u}_t + |y|^2\widehat{u} = 0, & t > 0 \\ \widehat{u} = \widehat{g}, \widehat{u}_t = \widehat{h}, & t = 0\end{cases}$$

再来寻找形如 $\widehat{u} = \beta e^{t\gamma}(\beta, \gamma \in \mathbb{C})$ 的解. 代入方程得

$$\gamma^2 + 2d\gamma + |y|^2 = 0 \Rightarrow \gamma = -d \pm (d^2 - |y|^2)^{\frac{1}{2}}$$

故

$$\widehat{u}(y,t) = \begin{cases} e^{-dt}(\beta_1(y)e^{\gamma(y)t} + \beta_2(y)e^{-\gamma(y)t}), & |y| \le d \\ e^{-dt}(\beta_1(y)e^{i\delta(y)t} + \beta_2(y)e^{-i\delta(y)t}), & |y| \ge d\end{cases}$$

其中 $\gamma(y) := (d^2 - |y|^2)^{\frac{1}{2}}(|y| \le d), \delta(y) := (|y|^2 - d^2)^{\frac{1}{2}}(|y| \ge d)$, 而 $\beta_1(y), \beta_2(y)$ 满足

$$\widehat{g}(y) = \beta_1(y) + \beta_2(y)$$

与

$$\widehat{h}(y) = \begin{cases} \beta_1(y)(\gamma(y) - d) + \beta_2(y)(-\gamma(y) - d), & |y| \le d \\ \beta_1(y)(i\delta(y) - d) + \beta_2(y)(-i\delta(y) - d), & |y| \ge d\end{cases}$$

从而得到表达式:

$$u(x,t) = \frac{e^{-dt}}{(2\pi)^{\frac{1}{2}}}\int_{\{|y|\le d\}} (\beta_1(y)e^{ixy+\gamma(y)t} + \beta_2(y)e^{ixy-\gamma(y)t})dy + \frac{e^{-dt}}{(2\pi)^{\frac{1}{2}}}\int_{\{|y|\ge d\}} (\beta_1(y)e^{i(xy+\delta(y)t)} + \beta_2(y)e^{i(xy-\delta(y)t)})dy$$

其中 e^{-dt} 这一项在 $t \to \infty$ 时就造成了阻尼.

[11] 最后一步是怎么估计的?

[12] 不确定是不是这么翻.

9.3.2 Radon 变换

Fourier 反演公式(9.74)对 PDE 理论的重要性是卓然的, 因为它把函数表成了指数波 $e^{ix\cdot y}$ 的形式. 本节介绍 Radon 变换 $\mathcal{R}$, 其在奇数维情况下给出了一种优雅的平面波代换分解. 这有时候是很管用的, 因为从一个函数的 Radon 变换中可以很清楚的看出其支集, 而这是 Fourier 变换所办不到的.

首先约定一些符号, 用 S^{n-1} 表示 $\mathbb{R}^n$ 中的单位球面 $\partial B(0,1)$, 特定的某点记作 $\omega = (\omega_1, \cdots, \omega_n)$. 与原点距离为 $s \in \mathbb{R}$, 单位法向量为 $\omega \in S^{n-1}$ 的平面为:

$$\Pi(s,\omega) := \{y \in \mathbb{R}^n : y\cdot\omega = s\}.$$

注意这个定义是允许 $s < 0$ 的.

定义 9.3.3 (Radon 变换)

函数 $u \in C_c^\infty(\mathbb{R}^n)$ 的 Radon 变换 $\mathcal{R}u = \widetilde{u}$ 为:

$$\widetilde{u}(s,\omega) := \int_{\Pi(s,\omega)} u dS, \quad s\in\mathbb{R}, \omega\in S^{n-1} \tag{9.104}$$

等式右边是在平面 $\Pi(s,\omega)$ 上的积分, 测度取 $n-1$ 维面测度. ♣

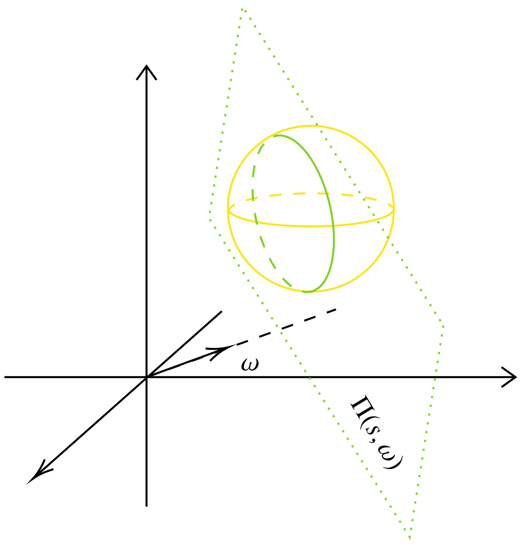


图 9.3: 一个特定方向的 Radon 变换示意图

定理 9.3.3 (Radon 变换的性质)

设 $u \in C_c^\infty(\mathbb{R}^n)$, 则

(i) $\widetilde{u}(-s,-\omega) = \widetilde{u}(s,\omega)$.

(ii) 对任意多重指标 α 有 $(D^\alpha u)^\sim = \omega^\alpha \frac{\partial^{|\alpha|}}{\partial s^{|\alpha|}}\widetilde{u}$.

(iii) $(\Delta u)^\sim = \frac{\partial^2}{\partial s^2}\widetilde{u}$.

(iv) 若在 $\mathbb{R}^n \backslash B(0,R)$ 内 $u \equiv 0$, 则对 $|s| \geq R$ 有 $\widetilde{u}(s,\omega) = 0$. ♡

证明

(i) 是显见的, 因为 $\Pi(s,\omega) = \Pi(-s,-\omega)$.

对于 (ii), 设 $\{b_1,\cdots,b_{n-1}\}$ 是超平面 $\Pi(s,\omega)$ 的一组正交基, 则 $\{b_1,\cdots,b_{n-1},\omega\}$ 是 $\mathbb{R}^n$ 上的一组正交基, 进而更换基底有

$$Du = \sum_{j=1}^{n-1}(Du\cdot b_j)b_j + (Du\cdot\omega)\omega$$

故

$$
\begin{aligned}
\widetilde{u_{x_i}} &= \int_{\Pi(s,\omega)} u_{x_i} dS \\
&= \int_{\Pi(s,\omega)} (\sum_{j=1}^{n-1}(Du \cdot b_j)(b_j \cdot e_i) + (Du \cdot \omega)(\omega \cdot e_i))dS \\
&= \sum_{j=1}^{n-1}(b_j \cdot e_i) \int_{\Pi(s,\omega)} Du \cdot b_j dS + \omega_i \int_{\Pi(s,\omega)} Du \cdot \omega dS \\
&= \omega_i \int_{\Pi(s,\omega)} Du \cdot \omega dS
\end{aligned}
$$

其中 $\sum_{j=1}^{n-1}(b_j \cdot e_i) \int_{\Pi(s,\omega)} Du \cdot b_j dS = 0$ 是因为 $b_j(1 \le j \le n-1)$ 与 $\Pi(s,\omega)$ 平行, 且 u 具有紧支集, 由 Newton-Leibniz 法则即得. 现在因为

$$
\widetilde{u}_s = \int_{\Pi(s,\omega)} Du \cdot \omega dS
$$

故这就证明了 $\alpha = e_i$ 的情况. 一般情况归纳即得, (iii) 进而根据 $|\omega| = 1$ 而成立.

最后, (iv) 是显然的, 因为在 $|s| > R$ 时 $\Pi(s,\omega) \cap B(0,R) = \varnothing$. □

定理 9.3.4 (Radon 变换与 Fourier 变换之间的关系)

设 $u \in C_c^\infty(\mathbb{R}^n)$, 则

$$
\overline{u}(r,\omega) := \int_{\mathbb{R}} \widetilde{u}(s,\omega)e^{-irs} ds = (2\pi)^{\frac{n}{2}} \widehat{u}(r\omega), \quad r \in \mathbb{R}, \omega \in S^{n-1} \tag{9.105}
$$

其中 $\widehat{u} = \mathcal{F}[u]$ 是 Fourier 变换.

证明

设 $\{b_1, \cdots, b_{n-1}\}$ 是 $\Pi(0,\omega)$ 上的正交基, 则更换基底有

$$
\widetilde{u}(s,\omega) = \int_{\mathbb{R}^{n-1}} u(\sum_{j=1}^{n-1} y_j b_j + s\omega)dy
$$

于是

$$
\int_{\mathbb{R}} \widetilde{u}(s,\omega)e^{-irs} ds = \int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} u(\sum_{j=1}^{n-1} y_j b_j + s\omega)e^{-irs} dyds
$$

现在换元并记 $x := \sum_{j=1}^{n-1} y_j b_j + s\omega$, 则 $x \cdot \omega = s|\omega|^2 = s$, 于是

$$
\int_{\mathbb{R}} \widetilde{u}(s,\omega)e^{-irs} ds = \int_{\mathbb{R}^n} u(x)e^{-ir(x\cdot\omega)} ds = (2\pi)^{\frac{n}{2}} \widehat{u}(r\omega).
$$

□

因为 Fourier 变换伴随着逆变换, 故我们自然地要问 Radon 变换是否有逆变换. 这由下面的定理给出

定理 9.3.5 (Radon 变换的逆变换)

(i) 设 $u \in C_c^\infty(\mathbb{R}^n)$, 则

$$
u(x) = \frac{1}{2(2\pi)^n} \int_{\mathbb{R}} \int_{S^{n-1}} \overline{u}(r,\omega)r^{n-1}e^{ir\omega\cdot x} dSdr \tag{9.106}
$$

其中 $\overline{u}$ 由(9.105)式定义.

(ii) 若 $n = 2k+1$ 是奇数, 则

$$
u(x) = \int_{S^{n-1}} r(x\cdot\omega,\omega)dS \tag{9.107}
$$

其中

$$r(s,\omega) := \frac{(-1)^k}{2(2\pi)^{2k}} \frac{\partial^{2k}}{\partial s^{2k}} \widetilde{u}(s,\omega) \tag{9.108}$$

证明

根据 $\overline{u}$ 的定义有

$$\begin{aligned}
\int_{\mathbb{R}} \int_{S^{n-1}} \overline{u}(s,\omega) r^{n-1} e^{ir\omega\cdot x} dS dr &= \int_{\mathbb{R}} \int_{S^{n-1}} ((2\pi)^{\frac{n}{2}} \widehat{u}(r\omega)) r^{n-1} e^{ir\omega\cdot x} dS dr \\
&= (2\pi)^{\frac{n}{2}} \int_{\mathbb{R}} \int_{S^{n-1}} \widehat{u}(r\omega) r^{n-1} e^{ir\omega\cdot x} dS dr \\
&\overset{(i)}{=} 2(2\pi)^{\frac{n}{2}} \int_0^\infty \int_{S^{n-1}} \widehat{u}(r\omega) r^{n-1} e^{ir\omega\cdot x} dS dr \\
&= 2(2\pi)^{\frac{n}{2}} \int_{\mathbb{R}^n} \widehat{u}(y) e^{iy\cdot x} dy = 2(2\pi)^n u(x)
\end{aligned}$$

其中 (i) 是单纯把 r 换成 $-r$, 把 ω 换成 $-\omega$.

如果把 $\overline{u}$ 的定义看成对 $\widetilde{u}$ 关于 s 的 Fourier 变换, 可知 $\overline{u} = \sqrt{2\pi}(\widetilde{u})^\wedge$, 此时符号 $\hat{\cdot}$ 表示固定 ω 时关于变量 r 的 Fourier 变换, 于是

$$(\frac{\partial^{2k}}{\partial s^{2k}} \widetilde{u})^\wedge = (ir)^{2k} (\widetilde{u})^\wedge = \frac{(-1)^k r^{2k}}{\sqrt{2\pi}} \overline{u}$$

两边应用 Fourier 逆变换有

$$\frac{\partial^{2k}}{\partial s^{2k}} \widetilde{u}(s,\omega) = \frac{(-1)^k}{2\pi} \int_{\mathbb{R}} \overline{u}(r,\omega) r^{2k} e^{irs} dr$$

现取 $s = x \cdot \omega$ 有

$$\frac{\partial^{2k}}{\partial s^{2k}} \widetilde{u}(\omega \cdot x, \omega) = \frac{(-1)^k}{2\pi} \int_{\mathbb{R}} \overline{u}(r,\omega) r^{n-1} e^{ir(\omega\cdot x)} dr$$

两边在单位球面上积分有

$$\int_{S^{n-1}} \frac{\partial^{2k}}{\partial s^{2k}} \widetilde{u}(\omega \cdot x, \omega) dS = \frac{(-1)^k}{2\pi} \int_{\mathbb{R}} \int_{S^{n-1}} \overline{u}(r,\omega) r^{n-1} e^{ir(\omega\cdot x)} dS dr = \frac{(-1)^k}{2\pi} 2(2\pi)^n u(x)$$

此即(9.107)式. □

9.3.3 Laplace 变换

记 $\mathbb{R}_+ = (0, \infty)$.

定义 9.3.4 (Laplace 变换)

若 $u \in L^1(\mathbb{R}_+)$, 则定义其 Laplace 变换 $\mathcal{L}u = u^\#$ 为

$$u^\#(s) := \int_0^\infty e^{-st} u(t) dt, \quad s \geq 0 \tag{9.109}$$

注意 Fourier 变换与 Radon 变换都是适用于定义在 $\mathbb{R}$(或 $\mathbb{R}^n$) 上的函数的, 而 Laplace 变换只适用于定义在 $\mathbb{R}_+$ 上的函数. 在实际应用中, Laplace 变换常常用于发展方程的研究, 将其应用在时间变量 t 上, 把 $n+1$ 元 PDE 变为 n 元 ODE. 这与 Fourier 变换的区别在于此时不必再考虑函数定义的区域了, 因而它适用于一般区域上 PDE 的讨论.

例 9.10(Laplace 变换与预解式) 考虑热方程

$$\begin{cases} v_t - \Delta v = 0, & \text{在 } U \times (0,\infty) \text{ 内} \\ v = f, & \text{在 } U \times \{t = 0\} \text{ 上} \end{cases} \tag{9.110}$$

对 v 关于 t 作 Laplace 变换有:

$$v^{\#}(x,s)=\int_0^{\infty}e^{-st}v(x,t)dt,\quad s>0$$

现在要问 $v^{\#}$ 满足什么方程? 计算知

$$\begin{aligned}\Delta v^{\#}(x,s)&=\int_0^{\infty}e^{-st}\Delta v(x,t)dt=\int_0^{\infty}e^{-st}v_t(x,t)dt\\&=s\int_0^{\infty}e^{-st}v(x,t)dt+e^{-st}v|_{t=0}^{t=\infty}=sv^{\#}(x,s)-f(x)\end{aligned}$$

固定 $s>0$, 记 $u(x):=v^{\#}(x,s)$, 则在 U 中有

$$-\Delta u+su=f \tag{9.111}$$

如果把 $-\Delta+s$ 看作一个算子, 自然地会考虑方程有形式解 $u=(-\Delta+s)^{-1}f$. 现在要讨论解的存在性与满足的估计, 就是要讨论算子 $-\Delta+s$ 的可逆性与满足的估计. 因为 $s\in\mathbb{R}_+$, 问题便转化为 $-\Delta+s$ 的预解估计. 随着 f 所处空间的不同, 可导出的 u 所在的空间也是不同的, 这便对应有不同的估计方法, 未来的分析学中会进一步发展这套理论.

9.4 问题

练习 9.1 证明若 u 是 Schrödinger 方程初值问题

$$\begin{cases}iu_t+\Delta u=0, & \text{在 } \mathbb{R}^n\times(-\infty,\infty) \text{ 内}\\ u=g, & \text{在 } \mathbb{R}^n\times\{t=0\} \text{ 上}\end{cases}$$

由式子

$$u(x,t)=\frac{1}{(4\pi it)^{\frac{n}{2}}}\int_{\mathbb{R}^n}e^{\frac{i|x-y|^2}{4t}}g(y)dy,\quad x\in\mathbb{R}^n,t>0$$

给出的解, 则

$$\|u(\cdot,t)\|_{L^{\infty}(\mathbb{R}^n)}\le\frac{1}{(4\pi|t|)^{\frac{n}{2}}}\|g\|_{L^1(\mathbb{R}^n)}$$

对任意 $t\neq0$ 均成立.

证明

知

$$\begin{aligned}\|u(\cdot,t)\|_{L^{\infty}(\mathbb{R}^n)}&\le\left|\frac{1}{(4\pi it)^{\frac{n}{2}}}\right|\cdot\left\|\int_{\mathbb{R}^n}e^{i\frac{|\cdot-y|^2}{4t}}g(y)dy\right\|_{L^{\infty}(\mathbb{R}^n)}\\&\le\frac{1}{(4\pi|t|)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\|e^{i\frac{|\cdot-y|^2}{4t}}g(y)\|_{L^{\infty}}dy\\&=\frac{1}{(4\pi|t|)^{\frac{n}{2}}}\|g\|_{L^1(\mathbb{R}^n)}\end{aligned}$$

□

9.5 补充: Schwartz 函数类与 Fourier 变换的更多补充

在泛函分析部分的第四章中, 曾经简单介绍过 Schwartz 函数类, 现在来系统介绍这一部分知识. 本节主要参考 [**LG1**].

9.5.1 Schwartz 函数类

首先回忆 Schwartz 函数的定义:

定义 9.5.1 (Schwartz 函数)

$\mathbb{R}^n$ 上的 $\mathscr{C}^\infty$ 复值函数 f 称为 Schwartz 函数, 如果对每一对多重指标 α, β 都存在正常数 $C_{\alpha,\beta}$ 使得

$$\rho_{\alpha,\beta}(f) = \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| = C_{\alpha,\beta} < \infty \tag{9.112}$$

$\rho_{\alpha,\beta}(f)$ 称为 f 的 Schwartz 半范数. $\mathbb{R}^n$ 上全体 Schwartz 函数的集合称为 Schwartz 空间, 记作 $\mathscr{S}(\mathbb{R}^n)$. ♣

定义 Schwartz 空间中的收敛如下:

定义 9.5.2 ($\mathscr{S}(\mathbb{R}^n)$ 中的收敛)

设 $f_k, f \in \mathscr{S}(\mathbb{R}^n)(k = 1, 2, \cdots)$. 称序列 $f_k \to f(\mathscr{S}(\mathbb{R}^n))$, 如果对任意多重指标 α, β 有

$$\rho_{\alpha,\beta}(f_k - f) = \sup_{x\in\mathbb{R}^n} |x^\alpha(\partial^\beta(f_k - f))(x)| \to 0, \quad k \to \infty$$

♣

例 9.11 对任意固定的 $x_0 \in \mathbb{R}^n$, 都有 $f(x + \frac{x_0}{k}) \to f(x)(\mathscr{S}(\mathbb{R}^n))(\forall f \in \mathscr{S}(\mathbb{R}^n), k \to \infty)$.

这里的收敛概念等价于 $\mathscr{S}(\mathbb{R}^n)$ 在代数运算 $(f, g) \mapsto f + g, (a, f) \mapsto af$ 下 (其中 a 是复标量, $f, g \in \mathscr{S}(\mathbb{R}^n)$) 保证 $f \mapsto \partial^\alpha f$ 连续的拓扑. 这个拓扑中那些包含 0 的开集的一个子基为

$$\{f \in \mathscr{S} : \rho_{\alpha,\beta}(f) < r\}, \quad \forall\alpha, \beta, \forall r \in \mathbb{Q}^+$$

观察知若 $\rho_{\alpha,\beta}(f) = 0$, 则 $f = 0$. 这说明 $\mathscr{S}(\mathbb{R}^n)$ 是局部凸拓扑空间, 其装备分离两点的半范数族 $\rho_{\alpha,\beta}$. 因为 $\mathscr{S}(\mathbb{R}^n)$ 的原点有可数拓扑基, 故由 Klee 定理知空间可度量化. 事实上下式就是 $\mathscr{S}(\mathbb{R}^n)$ 上的一个度量:

$$d(f, g) = \sum_{j=1}^\infty 2^{-j} \frac{\rho_j(f - g)}{1 + \rho_j(f - g)}$$

其中 ρ_j 跑遍半范数 $\rho_{\alpha,\beta}$, α, β 是多重指标. 显见 $\mathscr{S}$ 关于度量 d 完备. 事实上, $\mathscr{S}$ 中的基本列 $\{h_j\}_{j\in\mathbb{N}}$ 必为 L^∞ 中的基本列, 因而可以一致收敛到某函数 h. 另选 $\{\partial^\beta h_j\}_{j\in\mathbb{N}}, \{x^\alpha h_j\}_{j\in\mathbb{N}}$ 结论也成立, 可以证明它们的极限分别为 $\partial^\beta h, x^\alpha h(x)$. 从而 $\{h_j\}_{j\in\mathbb{N}}$ 收敛到 $h \in \mathscr{S}$. 故 $\mathscr{S}(\mathbb{R}^n)$ 是 Frechet 空间.

再说明接下来的命题前, 首先承认一个不等式[13]:

$$\forall k \in \mathbb{N}\forall x \in \mathbb{R}^n\backslash\{0\}(|x|^k \le C_{n,k} \sum_{|\beta|=k} |x^\beta|) \tag{9.113}$$

下面说明 $\mathscr{S}$ 中的收敛比任何 L^p 中的收敛都要强:

命题 9.5.1

设 $f, f_k, k = 1, 2, \cdots$ 都在 $\mathscr{S}(\mathbb{R}^n)$ 中. 若 $f_k \to f(\mathscr{S})$, 则 $f_k \to f(L^p)(\forall 0 < p \le \infty)$. 另外, 存在 $C_{p,n} > 0$ 满足

$$\|\partial^\beta f\|_{L^p} \le C_{p,n} \sum_{|\alpha|\le[\frac{n+1}{p}]+1} \rho_{\alpha,\beta}(f) \tag{9.114}$$

其中 f 保证右式收敛. ♠

证明

[13] 没有看懂 **[LG1]** 上是怎么证明这个式子的.

观察到当 $p<\infty$ 时有:

$$\begin{aligned}\|\partial^\beta f\|_{L^p} &= (\int_{\mathbb{R}^n} |\partial^\beta f(x)|^p dx)^{\frac{1}{p}} = (\int_{|x|\le 1} |\partial^\beta f(x)|^p dx + \int_{|x|\ge 1} |\partial^\beta f(x)|^p dx)^{\frac{1}{p}} \\ &= (\int_{|x|\le 1} |\partial^\beta f(x)|^p dx + \int_{|x|\ge 1} |x|^{n+1}|\partial^\beta f(x)|^p |x|^{-(n+1)} dx)^{\frac{1}{p}} \\ &\le (\|\partial^\beta f\|_{L^\infty}^p \int_{|x|\le 1} dx + (\sup_{|x|\ge 1} |x|^{n+1}|\partial^\beta f(x)|^p) \int_{|x|\ge 1} |x|^{-(n+1)} dx)^{\frac{1}{p}} \\ &\le C_{p,n}(\|\partial^\beta f\|_{L^\infty} + \sup_{|x|\ge 1}(|x|^{[\frac{n+1}{p}]+1}|\partial^\beta f(x)|))\end{aligned}$$

其中最后一步是因为 $\int_{|x|\le 1} dx, \int_{|x|\ge 1} |x|^{-(n+1)}dx$ 都是常数, 当 $0<p<1$ 时根据 $\|\cdot\|_{l^p}$ 的拟线性性即得, 而当 $1\le p<\infty$ 时根据 $|\cdot|_{l^p}$ 的三角不等式即得. 当 $p=\infty$ 时上式显然也成立 (取 $C_{p,n}\ge 1$ 即可). 现取 $m=[\frac{n+1}{p}]+1$, 由(9.113)式知

$$\sup_{|x|\ge 1} |x|^m |\partial^\beta f(x)| \le \sup_{|x|\ge 1} C_{n,m} \sum_{|\alpha|=m} |x^\alpha \partial^\beta f(x)| \le C_{n,m} \sum_{0<|\alpha|\le m} \rho_{\alpha,\beta}(f)$$

而 $\|\partial^\beta f\|_{L^\infty} \le \rho_{0,\beta}(f)$, (9.114)式立得, 因而 $\mathscr{S}$ 中的收敛强于 L^p 中的收敛. □

下面说明 Schwartz 函数类在乘法和卷积运算下封闭.

命题 9.5.2

设 $f,g\in\mathscr{S}(\mathbb{R}^n)$, 则 $fg, f*g\in\mathscr{S}(\mathbb{R}^n)$. 另有

$$\partial^\alpha(f*g) = (\partial^\alpha f)*g = f*(\partial^\alpha g) \tag{9.115}$$

对任意多重指标 α 成立.

证明

固定 $f,g\in\mathscr{S}(\mathbb{R}^n)$, 设 $e_j(0,\cdots,1,\cdots,0)$ 是第 j 个单位向量. 根据偏导定义:

$$\frac{f(y+he_j)-f(y)}{h} - (\partial_j f)(y) \to 0, \quad h\to 0 \tag{9.116}$$

同时(9.116)式可被关于 f 的常数控制, 故(9.116)式在 $g(x-y)dy$ 测度下的积分当 $h\to 0$ 时由 Lebesgue 控制收敛定理趋零. 因而 $\alpha=(0,\cdots,1,\cdots,0)$ 情况得证, 其余情况归纳即得, 这便证明了卷积与偏导的可换性.

现在证明两个 Schwartz 函数的卷积依旧在 $\mathscr{S}$ 内. 对任意 $N>0$ 存在常数 C_N 使得

$$|\int_{\mathbb{R}^n} f(x-y)g(y)dy| \le C_N \int_{\mathbb{R}^n} (1+|x-y|)^{-N}(1+|y|)^{-N-n-1}dy \tag{9.117}$$

显见

$$(1+|x-y|)^{-N} \le (1+|y|)^N(1+|x|)^{-N}$$

代入得

$$|(f*g)(x)| \le C_N(1+|x|)^{-N}\int_{\mathbb{R}^n}(1+|y|)^{-n-1}dy = C_N'(1+|x|)^{-N}$$

这说明 $f*g$ 与 $(1+|x|)^{-N}$ 在无穷处有类似衰减, 但因为 $N>0$ 是任意的, 故 $f*g$ 比任意多项式衰减都快.

因为 $\partial^\alpha(f*g)=(\partial^\alpha f)*g$, 在前面的证明中用 $\partial^\alpha f$ 换掉 f, 即得 $f*g$ 的任意阶导数在无穷处的衰减也快于任意多项式. 根据 $\mathscr{S}$ 类的等价定义即得 $f*g\in\mathscr{S}$. 最后, 根据 Leibniz 公式即得 $fg\in\mathscr{S}$.

9.5.2 Schwartz 函数的 Fourier 变换, Fourier 逆变换与 Fourier 反演

对 $x=(x_1,\cdots,x_n), y=(y_1,\cdots,y_n)$, 约定内积:

$$x\cdot y = \sum_{j=1}^n x_j y_j$$

定义 9.5.3 ((Schwartz 函数的)Fourier 变换)

给定 $f \in \mathscr{S}(\mathbb{R}^n)$, 记

$$\mathcal{F}[f](\xi) = \widehat{f}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx$$

称 $\widehat{f}$ 是 f 的 Fourier 变换.

♣

Schwartz 函数 Fourier 变换的诸多初等性质已经在前面的笔记中介绍过了, 这里补充一条:

命题 9.5.3

若 $f \in \mathscr{S}(\mathbb{R}^n), g \in \mathscr{S}(\mathbb{R}^m)$, 则

$$\mathcal{F}_{\mathbb{R}^{n+m}}[f(\circledast_1, \cdots, \circledast_n)g(\circledast_{n+1}, \cdots, \circledast_{n+m})] = \mathcal{F}_{\mathbb{R}^n}[f(\circledast_1, \cdots, \circledast_n)]\mathcal{F}_{\mathbb{R}^m}[g(\circledast_{n+1}, \cdots, \circledast_{n+m})]$$

其中 $\mathcal{F}_{\mathbb{R}^{n+m}}$ 表示 $\mathbb{R}^{n+m}$ 上的 Fourier 变换.

♠

定义 9.5.4 ((Schwartz 函数的)Fourier 逆变换)

给定 $f \in \mathscr{S}(\mathbb{R}^n)$, 定义

$$\mathcal{F}^{-1}[f](x) = f^{\vee}(x) = \widehat{f}(-x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(\xi)d\xi$$

算子 $f \mapsto f^{\vee}$ 就称为 Fourier 逆变换.

♣

从 Fourier 变换与逆变换的定义出发, 可以得到下述重要定理:

定理 9.5.1 (Fourier 反演)

给定 $f, g, h \in \mathscr{S}(\mathbb{R}^n)$, 有

(i) $\int_{\mathbb{R}^n} f(x)\widehat{g}(x)dx = \int_{\mathbb{R}^n} \widehat{f}(x)g(x)dx$,

(ii) (Fourier 反演) $(\widehat{f})^{\vee} = f = (f^{\vee})^{\wedge}$,

(iii) (Parseval 等式) $\int_{\mathbb{R}^n} f(x)\overline{h}(x)dx = \int_{\mathbb{R}^n} \widehat{f}(\xi)\overline{\widehat{h}(\xi)}d\xi$,

(iv) (Plancherel 等式) $\|f\|_{L^2} = \|\widehat{f}\|_{L^2} = \|f^{\vee}\|$,

(v) $\int_{\mathbb{R}^n} f(x)h(x)dx = \int_{\mathbb{R}^n} \widehat{f}(x)h^{\vee}(x)dx$.

♡

证明与前面介绍的 $L^1 \cap L^2$ 的情况是完全一样的, 这里就略过了. 根据 Parseval 等式与 Plancherel 等式, 可以得到下述推论:

推论 9.5.1

Fourier 变换是 $\mathscr{S}(\mathbb{R}^n)$ 上的自同构.

♡

9.5.3 $L^1 + L^2$ 上的 Fourier 变换

根据前面介绍的 Schwartz 函数 Fourier 变换的定义:

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

只要 $f \in L^1(\mathbb{R}^n)$, 上式都是有意义的, 这说明 Fourier 变换定义本身可以延拓到 L^1 上, 类似也可以延拓 Fourier 逆变换的定义.

在 **[Ev]** 的讲述中, 已经知道 L^2 上的 Fourier 变换与逆变换是依照 $L^1 \cap L^2$ 上的 Fourier 变换与逆变换定义的:

$$\widehat{f}(\xi) = \lim_{n\to\infty} \widehat{f_n}(\xi), \quad L^1 \cap L^2 \ni f_n \to f \in L^2(L^2)$$

在定义 L^1 与 L^2 上的 Fourier 变换后, 接下来就可以考虑 $L^p(1 < p < 2)$ 上的 Fourier 变换. 给定函数 $f \in L^p(\mathbb{R}^n)(1 < p < 2)$, 根据加和插值的结论**??**, 总有 $L^p(\mathbb{R}^n) \hookrightarrow L^1(\mathbb{R}^n) + L^2(\mathbb{R}^n)$, 亦即 f 可以唯一表成 $f_1 + f_2$, 其中

$f_1 \in L^1(\mathbb{R}^n), f_2 \in L^2(\mathbb{R}^n)$. 于是可以定义:

定义 9.5.5 (($L^p(1 < p < 2)$ 函数的)Fourier 变换)

若 $f = f_1 + f_2 \in L^p(\mathbb{R}^n)(1 < p < 2), f_1 \in L^1(\mathbb{R}^n), f_2 \in L^2(\mathbb{R}^n)$, 则 f 的 Fourier 变换定义为:

$$\mathcal{F}[f](\xi) = \widehat{f} = \widehat{f_1} + \widehat{f_2}.$$

♣

下面估计 L^p 函数的 Fourier 变换:

命题 9.5.4 (Hausdorff-Young)

对任意 $f \in L^p(\mathbb{R}^n)$, 都有估计

$$\| \widehat{f} \|_{L^{p'}} \leq \| f \|_{L^p}$$

其中 $1 \leq p \leq 2$.

♠

证明

根据 Riesz-Thorin 插值定理**??**, 将 Fourier 变换视作 $L^1 + L^2$ 上的算子, 根据 Fourier 变换的定义知

$$\|\widehat{f}\|_{L^\infty} \leq \|f\|_{L^1}$$

根据 Plancherel 定理有

$$\|\widehat{f}\|_{L^2} \leq \|f\|_{L^2}$$

于是只要

$$\frac{1}{p} = \frac{1-\lambda}{\infty} + \frac{\lambda}{2} = \frac{\lambda}{2}$$
$$\frac{1}{q} = \frac{1-\lambda}{1} + \frac{\lambda}{2} = 1 - \frac{\lambda}{2}$$

亦即 $\frac{1}{p} + \frac{1}{q} = 1 \Rightarrow p = q'$ 就有

$$\|\widehat{f}\|_{L^{q'}} \leq \|f\|_{L^q}, \quad 1 \leq q \leq 2$$

命题即证. □

第三部分

补充：经典力学的数学方法阅读笔记

第十章 Newton 力学

10.1 实验事实

本章给出一些作为力学基础的实验事实, 包含 Galois 相对性原理与 Newton 微分方程.

10.1.1 相对性原理和绝对性原理

下面是一些经典力学的实验事实, 可以当做建模课里的" 模型假设" 来看待.

- **空间与时间** 我们所在的空间是三维欧式空间, 时间则是一维的.
- **Galois 相对性原理** 存在一些参考系 (坐标系) 称为惯性系, 它们有下述两个性质:
  1. 一些自然规律在任何时刻, 在所有惯性系中都相同.
  2. 相对于一个惯性系作匀速直线运动的一切参考系也都是惯性系.
- **Newton 决定性原理** 一个力学系的初始状态 (即其中各点在某一时刻的位置与速度的总体) 唯一地决定其运动.

10.1.2 Galois 群与 Newton 方程

本节是前面实验事实的数学表述版本.

10.1.2.1 一些符号约定

约定 $\mathbb{R}$ 为全体实数的集合, $\mathbb{R}^n$ 为 n 维是向量空间.

定义 10.1.1 (n 维仿射空间)

n 维仿射空间可以看成" 没有固定原点的"$\mathbb{R}^n$. $\mathbb{R}^n$ 作为平移群作用在集合 A^n 上:

$$A^n \ni a \to a + \boldsymbol{b} \in \mathbb{R}^n, \quad \boldsymbol{b} \in \mathbb{R}^n$$

显见出于原点的缺失, A^n 中两点的和是没有定义的, 但两点之差有定义, 且它是 $\mathbb{R}^n$ 中的一个矢量. ♣

定义 10.1.2 (欧氏构造, 数量积, 距离, 欧氏空间)

矢量空间 $\mathbb{R}^n$ 上的一个欧氏构造指的是一个正定对称双线性型, 称为数量积, 一般记作 $\langle\cdot,\cdot\rangle$. 在给出数量积后, 就可以定义相应仿射空间 A^n 中的两点距离:

$$\rho(x,y) = \| x - y \| = \sqrt{\langle x-y, x-y\rangle}$$

仿射空间赋上这个距离函数就称为欧氏空间, 记为 E^n. ♣

10.1.2.2 Galois 构造

Galois 时空构造包含下述三个要素:

- **宇宙** 宇宙是一个四维仿射空间 A^4, A^4 中的点称为世界点或事件, 宇宙 A^4 中的平移构成一个矢量空间 $\mathbb{R}^4$.
- **时间** 时间是一个线性映射 $t : \mathbb{R}^4 \to \mathbb{R}$, 它把宇宙的平移矢量空间映到实的" 时间轴". 事件 $a \in A^4$ 到事件 $b \in A^4$ 的时间间隔指的就是 $t(b-a)$. 当 $t(b-a) = 0$ 时, 就称事件 a 和 b 是同时的.
  与一个固定事件同时的事件构成的集合是 A^4 的一个三维仿射子空间, 称为同时事件空间 A^3.
  $\ker t$(即使得 $t = 0$ 的那些点构成的集合) 包含 A^4 中的这么一些平移: 它们把某一事件变为与该事件同时的事件, 进而易知它们也把每个事件变成与它同时的事件. $\ker t$ 是矢量空间 $\mathbb{R}^4$ 的一个三维线性子空间 $\mathbb{R}^3$.

- 同时事件的距离

$$\rho(a,b) = \| a-b \| = \sqrt{\langle a-b, a-b\rangle}, \quad a,b \in A^3$$

由 $\mathbb{R}^3$ 上的数量积给出, 这个距离函数使得每个同时事件空间都成为三维欧式空间 E^3.

称具有 Galois 时空构造的空间 A^4 为 Galois 空间.

注 当然可以讨论"在不同位置同时发生的两个事件", 但是"在三维空间同一位置上发生的不同时事件 $a,b \in A^4$"就没有讨论意义了, 因为此时还没有选定一个坐标系.

定义 10.1.3 (Galois 群, Galois 变换)

Galois 群指 Galois 空间中保持其时空构造的一切变换所构成的群, 这个群的元素称为 Galois 变换. 显见 Galois 变换就是 A^4 中保持时间间隔和同时事件距离不变的仿射变换.

例 10.1(Galois 坐标空间中的 Galois 变换例子) 考虑时间轴和一个三维矢量空间 $\mathbb{R}^3$ 的直积 $\mathbb{R}\times\mathbb{R}^3$, 其中设 $\mathbb{R}^3$ 上有一个固定的欧氏构造, 这种空间具有自然的 Galois 构造, 称之为 Galois 坐标空间.

下面给出这个空间中的一些 Galois 变换例子, 其一为速度为 $\boldsymbol{v}$ 的匀速运动:

$$g_1(t,\boldsymbol{x}) = (t, \boldsymbol{x}+\boldsymbol{v}t), \quad \forall t\in\mathbb{R}, \boldsymbol{x}\in\mathbb{R}^3$$

其二为原点的平移:

$$g_2(t,\boldsymbol{x}) = (t+s, \boldsymbol{x}+\boldsymbol{s}), \quad \forall t\in\mathbb{R}, \boldsymbol{x}\in\mathbb{R}^3$$

其三为坐标轴的旋转:

$$g_3(t,\boldsymbol{x}) = (t, G\boldsymbol{x}), \quad \forall t\in\mathbb{R}, \boldsymbol{x}\in\mathbb{R}^3$$

其中 $G:\mathbb{R}^3\to\mathbb{R}^3$ 是正交变换.

命题 10.1.1

空间 $\mathbb{R}\times\mathbb{R}^3$ 的每个 Galois 变换都可以唯一地表成一个旋转变换, 一个平移变换和一个匀速运动的组合.

命题 10.1.2

所有 Galois 空间互相同构, 特别地它们都同构于坐标空间 $\mathbb{R}\times\mathbb{R}^3$.

定义 10.1.4 (Galois 坐标系, 均匀运动)

设 M 是一个集合, 一一对应 $\varphi_1: M\to\mathbb{R}\to\mathbb{R}^3$ 称为集 M 上的 Galois 坐标系. 如果对另一个坐标系 φ_2, 有 $\varphi_1\circ\varphi_2^{-1}:\mathbb{R}\times\mathbb{R}^3\to\mathbb{R}\times\mathbb{R}^3$ 是 Galois 变换, 就称 φ_2 对 φ_1 均匀地运动. Galois 坐标系 φ_1,φ_2 给 M 以相同的 Galois 构造.

10.1.2.3 运动, 速度, 加速度

定义 10.1.5 (运动, 速度, 加速度)

$\mathbb{R}^N$ 中的运动是一个可微映射 $\boldsymbol{x}: I\to\mathbb{R}^N$, 其中 I 是实轴的某个区间. 导数

$$\dot{\boldsymbol{x}}(t_0) = \frac{d\boldsymbol{x}}{dt}|_{t=t_0} = \lim_{h\to 0}\frac{\boldsymbol{x}(t_0+h)-\boldsymbol{x}(t_0)}{h} \in \mathbb{R}^N$$

称为 $t_0\in I$ 点的速度矢量. 二阶导数

$$\ddot{(\boldsymbol{x})}(t_0) = \frac{d^2\boldsymbol{x}}{dt^2}|_{t=t_0}$$

称为 t_0 点的加速度矢量. 一般都假设 $\boldsymbol{x}$ 有足够的光滑性, 映射 $\boldsymbol{x}: I\to\mathbb{R}^N$ 的像称为 $\mathbb{R}^N$ 中的一个轨迹或曲线.

下面定义三维欧式空间 E^3 中 n 个动点的力学系. 设 $\boldsymbol{x}:\mathbb{R}\to\mathbb{R}^3$ 是 $\mathbb{R}^3$ 中的运动, 其图像[1]是 $\mathbb{R}\times\mathbb{R}^3$ 中的曲线.

定义 10.1.6 (世界线)

对 Galois 空间中的曲线而言, 如果它是一个运动在某个 Galois 坐标系下的图像, 就称之为世界线.

因为一个点的运动可以给出一条世界线, 显见 n 点系统的运动在 Galois 空间中给出 n 条世界线. 在 Galois 坐标系中, 他们可以用 n 个映射 $\boldsymbol{x}_i:\mathbb{R}\to\mathbb{R}^3(i=1,2,\cdots,n)$ 来表示.

定义 10.1.7 (构形空间)

n 个 $\mathbb{R}^3$ 的直积称为 n 点系统的构形空间, 上述 n 个映射 $\boldsymbol{x}_i:\mathbb{R}\to\mathbb{R}^3$ 构成了由时间轴到构型空间的映射:

$$\boldsymbol{x}:\mathbb{R}\to\mathbb{R}^N,\quad N=3n$$

这个映射称为该 n 点系统在 Galois 坐标系 $\mathbb{R}\times\mathbb{R}^3$ 中的运动.

10.1.2.4 Newton 方程

根据 Newton 决定性原理, 一个系统的一切运动都由其初始位置 $\boldsymbol{x}(t_0)\in\mathbb{R}^N$ 与初速度 $\dot{\boldsymbol{x}}(t_0)\in\mathbb{R}^N$ 唯一决定. 特别地, 初始位置和初速度决定了加速度, 也即存在函数 $\boldsymbol{F}:\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}\to\mathbb{R}^N$ 使得

$$\ddot{\boldsymbol{x}}(s)=\boldsymbol{F}(\boldsymbol{x},\dot{\boldsymbol{x}},t)\tag{10.1}$$

Newton 力学以方程(10.1)为基础, 称之为 Newton 方程.

在合适的条件下, 根据 ODE 中解的存在唯一性定理, 函数 $\boldsymbol{F}$ 与初始条件 $\boldsymbol{x}(t_0),\dot{\boldsymbol{x}}(t_0)$ 唯一确定一个运动.

10.1.2.5 相对性原理对 Newton 方程的限制

Galois 相对性原理要求: 在物理时空中有一个特定的 Galois 构造, 称为惯性参考系类, 其满足若对任一力学系[2]各点的世界线作同一个 Galois 变换, 得到的依旧是这个力学系的世界线 (只是初始条件更新了). 这就要求 Newton 方程满足: 方程(10.1)对 Galois 变换群必须不变.

例 10.2 时间平移是 Galois 变换, Newton 方程对时间平移的不变性说明如果一个物理规律成立, 那么它时时刻刻都成立, 不会因为过了多久或在多久以前就是另一种规律. 数学上来说就是: 若 $\boldsymbol{x}=\boldsymbol{\varphi}(t)$ 是方程(10.1)的解, 则任取 $s\in\mathbb{R}$, $\boldsymbol{\varphi}(t+s)$ 也是(10.1)的一个解. 这说明(10.1)右式与 t 无关, 也即

$$\ddot{\boldsymbol{x}}=\boldsymbol{\Phi}(\boldsymbol{x},\dot{\boldsymbol{x}})$$

注 有时确实会出现 $\boldsymbol{F}$ 与 t 有关的情况, 比如要研究力学系 $I+II$ 的单个部分 I, 这时 II 对 I 的影响可能可以用描述 I 的方程组中参数对时间的变化来表示.

例 10.3 三维空间中的平移是 Galois 变换, 方程(10.1)对这种平移的不变性说明空间是均匀的, 也即" 空间各点有相同的性质". 从数学上来说, 如果 $\boldsymbol{x}_i=\boldsymbol{\varphi}_i(t)(i=1,\cdots,n)$ 是满足(10.1)的 n 个点的运动, 那么对任意的 $\boldsymbol{r}\in\mathbb{R}^3$, 运动 $\boldsymbol{\varphi}_i(t)+\boldsymbol{r}(i=1,\cdots,n)$ 同样满足方程(10.1). 这说明在惯性参考系中, 方程(10.1)右式只能依赖于" 相对坐标"$\boldsymbol{x}_j-\boldsymbol{x}_k$. 同样, 因为匀速运动也是 Galois 变换, 惯性参考系下的方程(10.1)右式就只能依赖于相对速度:

$$\ddot{\boldsymbol{x}}_i=\boldsymbol{f}_i(\{\boldsymbol{x}_j-\boldsymbol{x}_k,\dot{\boldsymbol{x}}_j-\dot{\boldsymbol{x}}_k\}),\quad i,j,k=1,\cdots,n$$

例 10.4 三维空间的旋转也是 Galois 变换, 方程(10.1)对这些旋转的不变性说明空间是各向同性的: 每个方向上的性质都一样. 从数学上来说, 若 $\boldsymbol{\varphi}_i:\mathbb{R}\to\mathbb{R}^3(i=1,\cdots,n)$ 是 n 点系统满足 Newton 方程(10.1)的运动, 而 $G:\mathbb{R}^3\to$

[1]映射 $f:A\to B$ 的图像是直积 $A\times B$ 中所有形如 $(a,f(a))(a\in A)$ 的元构成的子集.

[2]这里要求力学系是封闭的, 也就是说在研究一个给定的现象时, 所有与这个现象相互作用因而有所影响的物体, 都应该被包括在这个力学系内. 但一般来说很多物体都是能够近似忽略的. 力学系一词一般都表示封闭系.

$\mathbb{R}^3$ 是正交变换, 则运动 $G\boldsymbol{\varphi}_i : \mathbb{R} \to \mathbb{R}^3 (i = 1, \cdots, n)$ 也满足方程(10.1). 也就是说:

$$\boldsymbol{F}(G\boldsymbol{x}, G\dot{\boldsymbol{x}}) - G\boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}})$$

其中 $G\boldsymbol{x} = (G\boldsymbol{x}_1, \cdots, G\boldsymbol{x}_n), \boldsymbol{x}_i \in \mathbb{R}^3$.

命题 10.1.3 (Newton 第一定律)

若一个力学系仅由一点组成, 则其在惯性参考系中的加速度为 0.

证明

根据前述相对性原理对 Newton 方程的限制知, 对仅含一点的力学系 I 的 Newton 方程

$$\ddot{\boldsymbol{x}}(t) = \boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}})$$

而言, 有

$$\boldsymbol{F}(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \boldsymbol{F}(\boldsymbol{x} - \boldsymbol{x}, \dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}) = \boldsymbol{F}(0, 0)$$

这说明 $\ddot{\boldsymbol{x}}(t)$ 为常数, 设 $\ddot{\boldsymbol{x}}(t) \equiv \boldsymbol{a}$. 又根据 $\boldsymbol{F}$ 在正交变换下不变知 $\boldsymbol{a} = 0$, 即得 $\ddot{\boldsymbol{x}} = 0$. □

命题 10.1.4

设一个力学系由两点组成, 在初始时刻其速度 (在某惯性参考系中) 均为零, 证明它们将停留在初始时刻联结它们的直线上.

证明

对力学系 I 中的两点的运动 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 而言, 根据前述受限制的 Newton 方程可设

$$\begin{cases} \ddot{\boldsymbol{x}}_1 = \boldsymbol{F}(0, 0) \\ \ddot{\boldsymbol{x}}_2 = \boldsymbol{F}(\boldsymbol{x}_2 - \boldsymbol{x}_1, \dot{\boldsymbol{x}}_2 - \dot{\boldsymbol{x}}_1) \end{cases}$$

根据条件知 $\dot{\boldsymbol{x}}_1 = \dot{\boldsymbol{x}}_2 = 0$, 因而 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 均为常数, 自然停留在联结点 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 的直线上. □

10.1.2.6 例子: 保守系

在考察具体的力学系时, 一般不把宇宙里的所有物体都纳入这个系统. 比如在研究地球上的大多数现象时, 就可以忽略月球的影响, 同时也可以忽略研究的这个现象对地球运动本身产生的影响, 甚至可以把附着在地球上的一个坐标系看作固定坐标系. 显见此时相对性原理在这个坐标系里对运动方程就没有限制了.

设 $E^{3n} = E^3 \times \cdots \times E^3$ 是欧氏空间 E^3 中一个 n 点系统的构形空间, 令 $U : E^{3n} \to \mathbb{R}$ 为可微函数, $m_1, \cdots, m_n$ 为正数.

定义 10.1.8 (n 个质点在势场中的运动)

质量为 $m_1, \cdots, m_n$ 的 n 个点在位能为 U 的势场中的运动由下述微分方程组给出:

$$m_i \ddot{\boldsymbol{x}}_i = -\frac{\partial U}{\partial \boldsymbol{x}_i}, \quad i = 1, \cdots, n \tag{10.2}$$

10.2 运动方程的研究

10.2.1 具有一自由度的力学系

本节研究 Newton 方程(10.1)的相平面.

10.2.1.1 定义

定义 10.2.1 (具一自由度的力学系, 动能, 位能, 总能量)
具一自由度的力学系就是可以用单个微分方程描述的力学系:

$$\ddot{x} = f(x), \quad x \in \mathbb{R} \tag{10.3}$$

动能是二次型

$$T = \frac{1}{2}\dot{x}^2$$

位能是函数

$$U(x) = -\int_{x_0}^{x} f(\xi)d\xi$$

总能量即动能与位能之和:

$$E = T + U$$

故总能量是关于 x 和 $\dot{x}$ 的函数: $E(x, \dot{x})$. ♣

注 位能的符号是因为, 比如在研究石块的位能时, 当石块高于地面, 这样取符号可以让位能为正. 同时位能决定了 f, 故要刻画形如(10.3)的力学系, 给出位能就足够了. 从常数的导数为零可知在位能上加一个常数不会改变运动方程(10.3).

定理 10.2.1 (能量守恒定律)
按方程(10.3)运动的点的总能量守恒, 即 $E(x(t), \dot{x}(t))$ 与 t 无关. ♡

证明
知

$$\frac{dE}{dt} = \frac{d}{dt}(T + U) = \frac{d}{dt}(\frac{1}{2}\dot{x}^2 - \int_{x_0}^{x} f(\xi)d\xi) = \dot{x}\ddot{x} - \dot{x}f(x) = \dot{x}(\ddot{x} - f(x)) = 0$$

□

10.2.1.2 相平面

方程(10.3)可以等价为下述两个方程的方程组

$$\dot{x} = y, \quad \dot{y} = f(x) \tag{10.4}$$

定义 10.2.2 (相平面, 相点, 相速度矢量场)
以方程组(10.4)中的 x, y 为坐标的平面称为方程(10.3)的相平面, 相平面上的点称为相点. 方程组(10.4)的右式 $(y, f(x))$ 在相平面上决定了一个矢量场, 称为相速度矢量场. ♣

方程组(10.4)的解是相平面上一个相点的运动 $\varphi : \mathbb{R} \to \mathbb{R}^2$, 方程组(10.4)原先描述的动点在每一时刻的速度都等于其对应的相点在该时刻所在位置处的相速度矢量.

称 φ 的像为相曲线, 进而相曲线可以由下述参数方程给出:

$$x = \varphi(t), \quad y = \psi(t)$$

根据 ODE 中解的存在唯一性定理可得下述命题.

命题 10.2.1
经过每个相点的有且仅有一条相曲线, 也即不可能有两条不同的相曲线相交. ♠

注 有时相曲线可能缩为一点, 这种点称为平衡位置, 平衡位置上的相速度矢量为零.

因为每条相曲线都代表一个特定动点的运动, 而单点运动的能量是守恒的, 故每个相曲线都在一个等能集 $E(x, y) = h$ 上.

10.2.1.3 一些例子

例 10.5(振动) 振动理论的基本方程为

$$\ddot{x} = -x$$

其相平面如图(10.1)

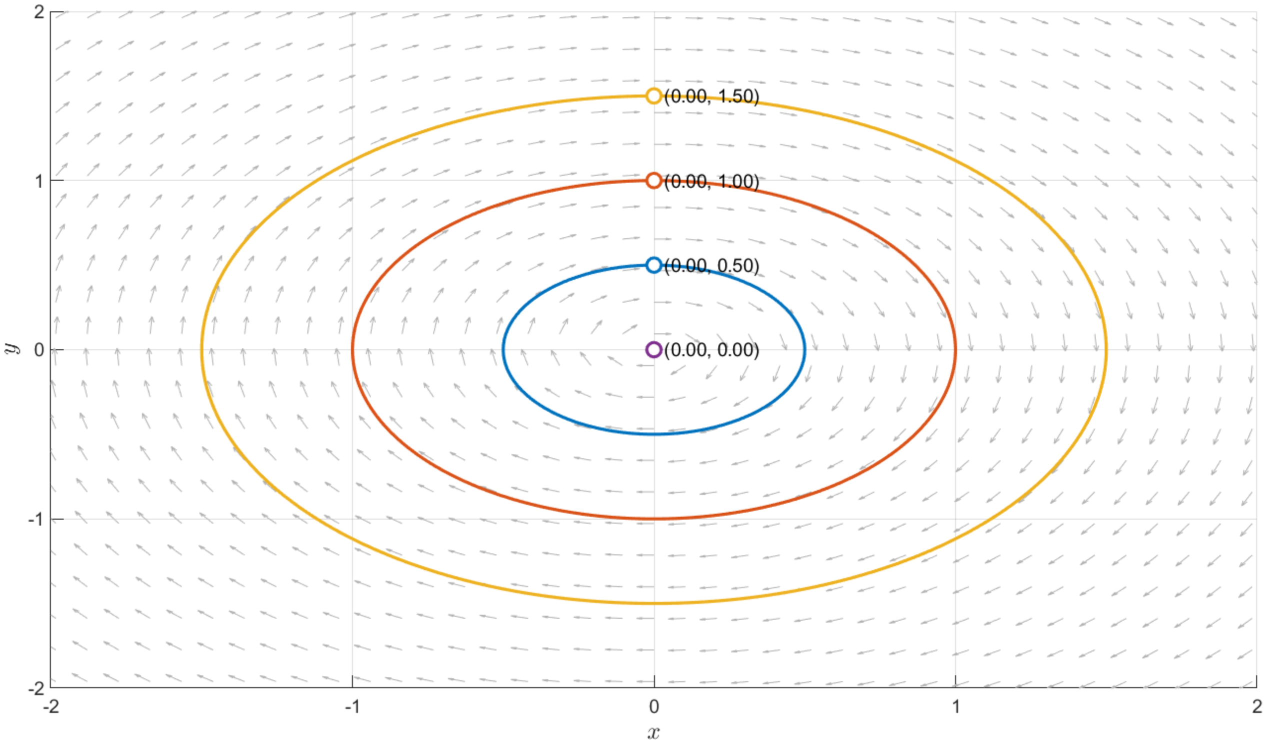


图 10.1: $\ddot{x} = -x$ 的相图与一些特定运动示意

若取 $x_0 = 0$, 则诸能量为:

$$T = \frac{\dot{x}^2}{2}, U = \frac{x^2}{2}, E = \frac{\dot{x}^2}{2} + \frac{x^2}{2}$$

如图(10.1)所示, 其等能集为同心圆和原点. 相点 (x, y) 处的相速度矢量为 $(y, -x)$, 它与动径 (x, y) 的方向垂直, 大小相等. 故相点在相平面上的运动是绕原点 O 的匀速旋转: $x = r_0 \cos(\varphi_0 - t), y = r_0 \sin(\varphi_0 - t)$, 每个等能集都是一条相曲线.

10.2.1.4 相流

设 M 是相平面上的一点, 现在研究方程组(10.4)的解, 其在 $t = 0$ 时的初始条件由点 M 表示. 设方程组的任一解都可以扩展到整个时间轴上, 知解在任意时刻 t 时所取的值依赖于初值 M. 记变动时刻 t 得到的相点 $M(t)$ 为

$$M(t) = g^t M$$

这便定义了相平面到自身的映射 $g^t : \mathbb{R}^2 \to \mathbb{R}^2$(如图(10.2)).

根据 ODE 中的定理[3]可知, 映射 g^t 是一个微分同胚 (即具有可微逆的一一可微映射). 这些微分同胚 $g^t (t \in \mathbb{R})$

[3]什么定理?

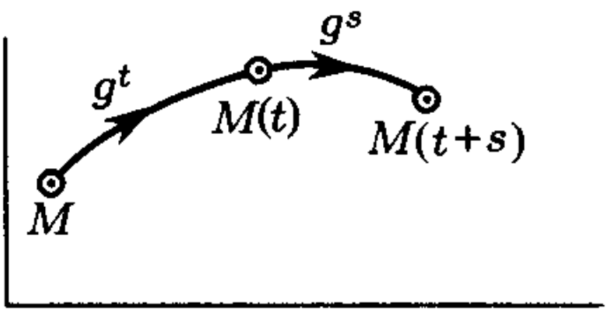


图 10.2: [AR1] 中对相流的示意图

构成一个群: $g^{t+s} = g^t \circ g^s$. 其中映射 g^0 是恒等映射, 即 $g^0 M = M$, g^{-t} 是 g^t 的逆. 由式子 $g(t, M) = g^t M$ 所定义的映射 $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ 是可微的. 综合这些性质, 可以说: 变换 g^t 构成相平面的微分同胚的一个单参数群. 这个群也称为由方程组(10.4)(或方程(10.3)) 给出的相流.

例 10.6 方程 $\ddot{x} = -x$ 给出的相流即绕原点旋转角 t 的旋转变换 g^t 构成的群.

10.2.2 具二自由度的力学系

10.2.2.1 定义

定义 10.2.3 (二自由度的力学系)

二自由度的力学系指的是由下述微分方程所定义的系统:

$$\ddot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}), \quad \boldsymbol{x} \in E^2 \tag{10.5}$$

其中 $\boldsymbol{f}$ 是一个平面矢量场. ♣

定义 10.2.4 (保守系)

如果一个力学系中存在单值函数 $U : E^2 \to \mathbb{R}$, 使得 $\boldsymbol{f} = -\frac{\partial U}{\partial \boldsymbol{x}}$, 就称该力学系为保守系. ♣

根据定义易知保守系的运动方程形如

$$\ddot{\boldsymbol{x}} = -\frac{\partial U}{\partial \boldsymbol{x}}$$

写成坐标形式即

$$\begin{cases} \ddot{x}_1 = -\frac{\partial U}{\partial x_1} \\ \ddot{x}_2 = -\frac{\partial U}{\partial x_2} \end{cases}$$

10.2.2.2 能量守恒定律

定理 10.2.2

保守系的总能量守恒, 即:

$$\frac{dE}{dt} = 0, \quad 其中 E = \frac{1}{2}\dot{\boldsymbol{x}}^2 + U(\boldsymbol{x}), \dot{\boldsymbol{x}}^2 = (\dot{\boldsymbol{x}}, \dot{\boldsymbol{x}})$$

♡

证明

根据保守系的运动方程知

$$\frac{dE}{dt}=\frac{1}{2}\frac{d}{dt}\langle\dot{\boldsymbol{x}},\dot{\boldsymbol{x}}\rangle+\frac{d}{dt}(U(\boldsymbol{x}))=\langle\dot{\boldsymbol{x}},\ddot{\boldsymbol{x}}\rangle+\langle\frac{\partial U}{\partial \boldsymbol{x}},\dot{\boldsymbol{x}}\rangle=\langle\ddot{\boldsymbol{x}}+\frac{\partial U}{\partial \boldsymbol{x}},\dot{\boldsymbol{x}}\rangle=0$$

□

注

1. 若初始时刻的总能量为 E, 则系统运动的一切轨道都在区域 $U(\boldsymbol{x})\leq E$ 中, 也即一点在任意时刻都位于势阱 $U(x_1,x_2)\leq E$ 内.
2. 对于具一自由度的力学系, 总是可以引入位能

$$U(x)=-\int_{x_0}^{x}f(\xi)d\xi$$

但对于具二自由度的力学系, 如果它不是保守系, 就不能这么做了.

10.2.2.3 相空间

二自由度力学系运动方程(10.5)可以化为方程组

$$\begin{cases}\dot{x}_1=y_1\\ \dot{y}_1=-\frac{\partial U}{\partial x_1}\\ \dot{x}_2=y_2\\ \dot{y}_2=-\frac{\partial U}{\partial x_2}\end{cases}\tag{10.6}$$

二自由度力学系的相空间即坐标为 x_1,x_2,y_1,y_2 的四维空间.

方程组(10.6)在四维空间中定义了相速度矢量场与相流 (它是四维空间中微分同胚的一个单参数群). (10.6)的相曲线是四维相空间的子集, 整个相空间被分解为相曲线的集合. 把相曲线由四维空间投影到 x_1,x_2 平面上即可给出动点在 x_1,x_2 平面上的轨迹, 这些轨迹称为轨道. 尽管相曲线彼此不相交, 但轨道是可以有交点的. 能量守恒定律方程

$$E=\frac{\dot{\boldsymbol{x}}^2}{2}+U(\boldsymbol{x})=\frac{y_1^2+y_2^2}{2}+U(x_1,x_2)$$

在四维空间中定义一个三维超曲面: $\pi_{E_0}:E(x_1,x_2,y_1,y_2)=E_0$, 曲面 π_{E_0} 在相流下是不变的: $g^t\pi_{E_0}=\pi_{E_0}$. 也就是说, 相流沿等能超曲面流动. 根据相速度矢量的定义, 相速度矢量在各点处都切于 π_{E_0}, 故 π_{E_0} 完全由相曲线组成.

10.2.3 保守力场

本节讨论功与位能的关系.

10.2.3.1 力场沿一条路径做的功

回忆力 $\boldsymbol{F}$ 沿路径 $\boldsymbol{S}$ 所做的功的定义: 恒定力在路径 $\boldsymbol{S}=\overrightarrow{M_1M_2}$ 上所做的功定义为数量积:

$$A=\langle\boldsymbol{F},\boldsymbol{S}\rangle=|\boldsymbol{F}|\cdot|\boldsymbol{S}|\cdot\cos\varphi$$

对于变力与曲线路径而言, 仿照上述定义可得:

定义 10.2.5 (功)

设有矢量场 $\boldsymbol{F}$ 与一条有限长曲线 l, 用分段为 $\Delta\boldsymbol{S}_i$ 的折线去逼近曲线 l, 并记 $\boldsymbol{F}_i$ 为力在 $\Delta\boldsymbol{S}_i$ 上某点的值, 则场 $\boldsymbol{F}$ 在 l 上做的功定义为:

$$A=\lim_{|\Delta\boldsymbol{S}_i|\to 0}\sum\langle\boldsymbol{F}_i,\Delta\boldsymbol{S}_i\rangle$$

若场是连续的且 l 可求长, 则上述极限存在, 记作 $\int_l\langle\boldsymbol{F},d\boldsymbol{S}\rangle$.

♣

10.2.3.2 场为保守场的条件

定理 10.2.3
向量场 $\boldsymbol{F}$ 为保守场当且仅当它沿路径 M_1M_2 作的功只依赖于路径的端点, 而与路径的形状无关. ♡

证明
当场 $\boldsymbol{F}$ 做的功与路径无关, 此时定义

$$U(M)=-\int_{M_0}^{M}\langle \boldsymbol{F}, d\boldsymbol{S}\rangle$$

根据假设, 该函数是良定义的, 且容易验证

$$\boldsymbol{F}=-\frac{\partial U}{\partial \boldsymbol{x}}$$

这说明场 $\boldsymbol{F}$ 是保守场, 且其位能为 U.

反之, 设场 $\boldsymbol{F}$ 是保守场, 其位能为 U, 则显见

$$\int_{M_0}^{M}\langle \boldsymbol{F}, d\boldsymbol{S}\rangle=-U(M)+U(M_0)$$

上右式与 $\boldsymbol{S}$ 无关, 故功不依赖于路径的形状. □

10.2.3.3 有心场

定义 10.2.6 (有心场)
平面 E^2 上的矢量场称为以 O 为心的有心场, 如果它对平面上保持 O 固定的运动的群是不变的. ♣

命题 10.2.2
有心场中的一切矢量都位于过 O 的射线上, 且矢量场在一点的大小只依赖于该点到场中心的距离. ♠

考虑在点 O 处无定义的有心场有时候也有意义.

例 10.7 Newton 场 $\boldsymbol{F}=-k(\frac{\boldsymbol{r}}{|\boldsymbol{r}|^3})$ 是有心场.

定理 10.2.4
每个有心场都是保守场, 且有心场的位能只依赖于到场中心的距离, 即 $U=U(r)$. ♡

证明
根据命题(10.2.2), 可设 $\boldsymbol{F}(\boldsymbol{r})=\Phi(r)\boldsymbol{e}_r$, 其中 $\boldsymbol{r}$ 是关于 O 的动径 (即向量 $O\boldsymbol{x}$), r 是 $\boldsymbol{r}$ 的长度, 单位矢量 $\boldsymbol{e}_r=\frac{\boldsymbol{r}}{|\boldsymbol{r}|}$ 是其方向, 进而

$$\int_{M_1}^{M_2}\langle \boldsymbol{F}, d\boldsymbol{S}\rangle=\int_{r(M_1)}^{r(M_2)}\Phi(r)dr$$

这个积分显然与 $\boldsymbol{S}$ 无关. □

10.2.4 角动量

之后会看到: 一个力学问题的过程对某个运动群的不变性总蕴含一个守恒律. 现在有心场对旋转群不变, 相应的首次积分称为角动量, 这个不变性就对应角动量守恒定律.

定义 10.2.7 (有心场运动方程)

(具单位质量的) 质点在平面有心场中的运动方程为

$$\ddot{\boldsymbol{r}} = \Phi(r)\boldsymbol{e}_r$$

其中 $\boldsymbol{r}$ 是以场中心 O 为起点的动径, r 为其长度, $\boldsymbol{e}_r$ 为其方向. 规定该平面在有向三维欧氏空间内.

定义 10.2.8 (角动量)

单位质量质点关于 O 点的角动量是矢量积

$$\boldsymbol{M} = \boldsymbol{r} \wedge \dot{\boldsymbol{r}}$$

矢量 $\boldsymbol{M}$ 垂直于上述平面, $\boldsymbol{M} = M\boldsymbol{n}$, 其中 $\boldsymbol{n} = \boldsymbol{e}_1 \wedge \boldsymbol{e}_2$ 是法向量, $\boldsymbol{e}_1, \boldsymbol{e}_2$ 给出了该平面的有向标架 (如图(10.3)).

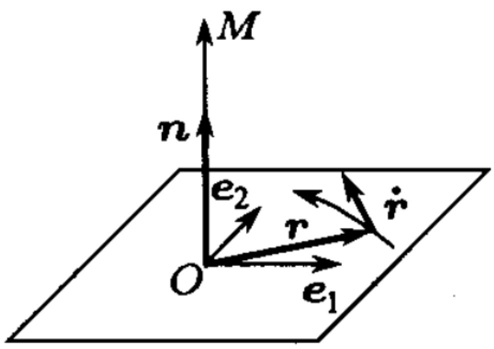


图 10.3: [AR1] 关于角动量的示意图

注 一般来说, " 作用在 $\boldsymbol{r}$ 点" 的矢量 $\boldsymbol{a}$ 关于 O 点的矩是 $\boldsymbol{r} \wedge \boldsymbol{a}$. 角动量就是动量矩.

10.2.4.1 角动量守恒定律

引理 10.2.1

令 $\boldsymbol{a}, \boldsymbol{b}$ 为有向欧氏空间 $\mathbb{R}^3$ 中随时间变化的矢量, 则

$$\frac{d}{dt}(\boldsymbol{a} \wedge \boldsymbol{b}) = \dot{\boldsymbol{a}} \wedge \boldsymbol{b} + \boldsymbol{a} \wedge \dot{\boldsymbol{b}}$$

定理 10.2.5 (角动量守恒定律)

在有心场中的运动, 其关于中心 O 的动量矩 $\boldsymbol{M}$ 不随时间而改变.

证明

此即证明 $\dot{\boldsymbol{M}} = 0$, 根据定义与前述引理知 $\boldsymbol{M} = \dot{\boldsymbol{r}} \wedge \dot{\boldsymbol{r}} + \boldsymbol{r} \wedge \ddot{\boldsymbol{r}}$. 因为现在是在有心场内, 故根据有心场运动方程知 $\ddot{\boldsymbol{r}}$ 与 $\boldsymbol{r}$ 共线, 因而 $\dot{\boldsymbol{M}} = 0$. □

10.2.4.2 Kepler 定律

角动量守恒定律最早是由 Kepler 在观察火星运动时提出的. 下面用 Kepler 的方式重新叙述角动量守恒定律.

在平面上引入极坐标 r, φ, 极点在中心 O 处. 在坐标为 $(|\boldsymbol{r}| = r, \varphi)$ 的点 $\boldsymbol{r}$ 处考虑两个单位矢量: 其中一个 $\boldsymbol{e}_r$ 沿着动径, 故

$$\boldsymbol{r} = r\boldsymbol{e}_r$$

另一个记作 $\boldsymbol{e}_\varphi$, 其与 $\boldsymbol{e}_r$ 垂直, 并指向 φ 增加的方向. 现在把速度矢量 $\dot{\boldsymbol{r}}$ 用基底 $\boldsymbol{e}_r, \boldsymbol{e}_\varphi$ 来表示:

引理 10.2.2

$$\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi$$

证明

显见 $\boldsymbol{e}_r, \boldsymbol{e}_\varphi$ 均以角速度 $\dot{\varphi}$ 旋转.

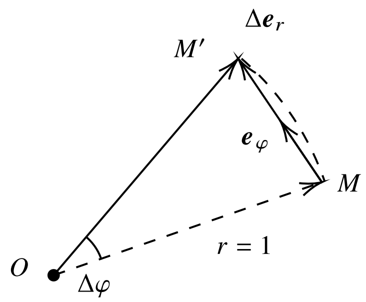


图 10.4

根据图(10.4), 当经过一个微小时间 Δt 后, 可近似将动径增量 $\Delta \boldsymbol{e}_r$ 视作与 OM 垂直, 因而其方向可近似看作 $\boldsymbol{e}_\varphi$. 而针对 $\boldsymbol{e}_r$ 的大小, 从图(10.4)中可见 $\overrightarrow{MM'}$ 与 $\overset{\frown}{MM'}$ 近似相等, 从而有

$$|\Delta \boldsymbol{e}_r| = |\overrightarrow{MM'}| \approx |\overset{\frown}{MM'}| = r\Delta\varphi = \Delta\varphi$$

得到

$$\Delta \boldsymbol{e}_r = \Delta\varphi \boldsymbol{e}_\varphi \Rightarrow \dot{\boldsymbol{e}}_r = \dot{\varphi}\boldsymbol{e}_\varphi$$

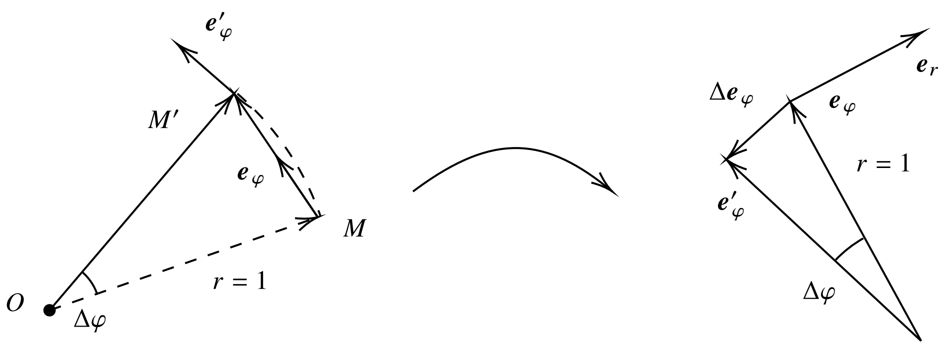


图 10.5

再根据图(10.5), 角增量 $\Delta \boldsymbol{e}_\varphi$ 的方向可近似看作 $-\boldsymbol{e}_r$, 而其大小与前文类似, 可以用弧长近似代替, 从而

$$|\Delta \boldsymbol{e}_\varphi| \approx r\Delta\varphi = \Delta\varphi$$

得到

$$\Delta \boldsymbol{e}_\varphi = -\Delta\varphi \boldsymbol{e}_r \Rightarrow \dot{\boldsymbol{e}}_\varphi = -\dot{\varphi}\boldsymbol{e}_r$$

综上有

$$\begin{cases} \dot{\boldsymbol{e}}_r = \dot{\varphi}\boldsymbol{e}_\varphi \\ \dot{\boldsymbol{e}}_\varphi = -\dot{\varphi}\boldsymbol{e}_r \end{cases}$$

现在对 $\boldsymbol{r} = r\boldsymbol{e}_r$ 两边求导有

$$\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\boldsymbol{e}}_r = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi$$

引理即证. □

根据引理, 对角动量有

$$\boldsymbol{M} = \boldsymbol{r} \wedge \dot{\boldsymbol{r}} = \boldsymbol{r} \wedge (\dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi) = \boldsymbol{r} \wedge \dot{r}\boldsymbol{e}_r + \boldsymbol{r} \wedge r\dot{\varphi}\boldsymbol{e}_\varphi = r\dot{\varphi}\boldsymbol{r} \wedge \boldsymbol{e}_\varphi = r^2\dot{\varphi}\boldsymbol{e}_r \wedge \boldsymbol{e}_\varphi$$

根据角动量守恒知 $r^2\dot{\varphi}$ 是守恒的, 下面说明这个量的几何意义.

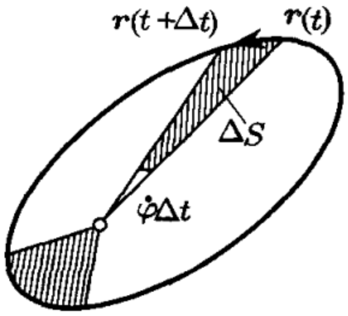


图 10.6: [**AR1**] 中对面积速度的示意图

Kepler 把动径扫过的面积 $S(t)$ 的变化率称为面积速度 C(见图(10.6)):

$$C = \frac{dS}{dt}$$

他通过观察行星运动发现, 在相等时间内, 动径扫过的面积相等, 故面积速度 $\frac{dS}{dt}$ 为常数. 又因为

$$\Delta S = S(t+\Delta t) - S(t) = \frac{1}{2}r^2\dot{\varphi}\Delta t + o(\Delta t)$$

故面积速度

$$C = \frac{dS}{dt} = \frac{1}{2}r^2\dot{\varphi} = \frac{1}{2}M$$

是单位质点角动量的一半, 进而是守恒的.

10.2.5 在有心力场中的运动的研究

角动量守恒定律使得我们能把有心力场中的运动问题化成一自由度的问题, 进而可以完全确定有心力场中的运动.

10.2.5.1 化为一维问题

考察质量为 1 的质点在平面有心力场中的运动:

$$\ddot{\boldsymbol{r}} = -\frac{\partial U}{\partial \boldsymbol{r}}, \quad U = U(r)$$

根据角动量守恒定律, 量 $M = \dot{\varphi}(t)r^2(t)$ 为常数, 也即 M 与 t 无关.

定理 10.2.6

对于单位质量质点在平面有心力场中的运动, 质点到中心的距离的变化, 与位能为

$$V(r) = U(r) + \frac{M^2}{2r^2}$$

的一维问题中质点到中心距离的变化方式相同.

证明

在式子

$$\dot{\boldsymbol{r}} = \dot{r}\boldsymbol{e}_r + r\dot{\varphi}\boldsymbol{e}_\varphi$$

两边对 t 微分有

$$\ddot{\boldsymbol{r}} = (\ddot{r} - r\dot{\varphi}^2)\boldsymbol{e}_r + (2\dot{r}\dot{\varphi} + r\ddot{\varphi})\boldsymbol{e}_\varphi$$

根据有心场的定理, 其位能只依赖于动径 r, 故

$$\frac{\partial U}{\partial \boldsymbol{r}} = \frac{\partial U}{\partial r}\boldsymbol{e}_r$$

故极坐标下的运动方程形如

$$\begin{cases} \ddot{r} - r\dot{\varphi}^2 = -\frac{\partial U}{\partial r} \\ 2\dot{r}\dot{\varphi} + r\ddot{\varphi} = 0 \end{cases}$$

又根据角动量守恒定律:

$$\dot{\varphi} = \frac{M}{r^2}$$

代入即得

$$\ddot{r} = -\frac{\partial U}{\partial r} + r\frac{M^2}{r^4} = -\frac{\partial V}{\partial r}$$

其中 $V = U + \frac{M^2}{2r^2}$. □

称 $V(r)$ 为有效位能.

注 上面化出的一维问题的总能量

$$E_1 = \frac{\dot{r}^2}{2} + V(r) \tag{10.7}$$

其实等于原问题的总能量

$$E = \frac{1}{2}\dot{\boldsymbol{r}}^2 + U(\boldsymbol{r})$$

这是因为

$$\frac{1}{2}\dot{\boldsymbol{r}}^2 = \frac{1}{2}\dot{r}^2 + \frac{1}{2}r^2\dot{\varphi}^2 = \frac{\dot{r}^2}{2} + \frac{M^2}{2r^2}.$$

10.2.5.2 运动方程的求积

由(10.7)式整理可得

$$\dot{r} = \sqrt{2(E - V(r))}$$

从而两边积分即得 r 对 t 的依赖性:

$$\frac{dr}{\sqrt{2(E - V(r))}} = dt \Rightarrow \int dt = \int \frac{dr}{\sqrt{2(E - V(r))}}$$

由角动量守恒知 $\dot{\varphi} = \frac{M}{r^2}$, 求导有

$$\frac{d\varphi}{dr} = \frac{M}{r^2\sqrt{2(E - V(r))}}$$

故极坐标下的轨道方程也可用求积法得出:

$$\varphi = \int \frac{M\,dr}{r^2\sqrt{2(E - V(r))}}$$

10.2.6 三维空间中质点的运动

本节定义关于一个轴的角动量, 并证明角动量在轴对称场中的运动是守恒的. 前面关于平面运动得到的结果可以很容易推广到空间运动.

10.2.6.1 保守场

考虑保守场中的运动

$$\ddot{\boldsymbol{r}} = -\frac{\partial U}{\partial \boldsymbol{r}}$$

其中 $U = U(\boldsymbol{r}), \boldsymbol{r} \in E^3$. 此时容易证明能量守恒定律仍然成立:

$$\frac{dE}{dt} = 0, \quad E = \frac{1}{2}\dot{\boldsymbol{r}}^2 + U(\boldsymbol{r})$$

10.2.6.2 有心场

三维空间有心场中的运动仍然满足角动量守恒定律, 也即 $\boldsymbol{M} = \boldsymbol{r} \wedge \dot{\boldsymbol{r}}$ 不变, 亦即 $\frac{d\boldsymbol{M}}{dt} = 0$.

和二维情况同理, 可以证明每个三维有心场都是保守的, 且

$$\frac{d\boldsymbol{M}}{dt} = \dot{\boldsymbol{r}} \wedge \dot{\boldsymbol{r}} + \boldsymbol{r} \wedge \ddot{\boldsymbol{r}} = 0$$

最后等于零是因为 $\ddot{\boldsymbol{r}} = -\frac{\partial U}{\partial \boldsymbol{r}}$, 而根据有心场知 $\frac{\partial U}{\partial \boldsymbol{r}}$ 与 $\boldsymbol{r}$ 共线.

推论 10.2.1

在有心场中的运动, 其轨道总是平面曲线.

证明

知

$$\langle \boldsymbol{M}, \boldsymbol{r} \rangle = \langle \boldsymbol{r} \wedge \dot{\boldsymbol{r}}, \boldsymbol{r} \rangle = 0$$

故 $\boldsymbol{r} \perp \boldsymbol{M}$. 又因为 $\boldsymbol{M}$ 本身是常向量, 故所有轨道都位于垂直于 $\boldsymbol{M}$ 的平面上. □

注 通过上述定理, 可以把三维空间有心场的轨道的研究划归回前面研究过的平面问题.

10.2.6.3 轴对称场

定义 10.2.9 (轴对称)

若 E^3 中的一个矢量场对一个旋转群不变, 而这些旋转使得某条轴上的各点不动, 就称该矢量场是轴对称的.

命题 10.2.3

如果一个场是轴对称且保守的, 那么其中的位能形如 $U = U(r, z)$, 其中 r, φ, z 是柱坐标. 特别由此可知场中的矢量均位于过 z 轴的平面中.

例 10.8 旋转体产生的引力场是轴对称场.

下面设旋转轴为 z 轴, 其方向与三维欧氏空间 E^3 中的矢量 $\boldsymbol{e}_z$ 同向, $\boldsymbol{F}$ 为欧氏线性空间 $\mathbb{R}^3$ 中的一个矢量, O 是 z 轴上一点, $\boldsymbol{r} = \boldsymbol{x} - O \in \mathbb{R}^3$ 为点 $\boldsymbol{x} \in E^3$ 相对于 O 的动径.

定义 10.2.10 (对轴的矩)

作用在点 $\boldsymbol{r}$ 的矢量 $\boldsymbol{F}$ 相对于 z 轴的矩 M_z, 指的是该矢量 $\boldsymbol{F}$ 对 O 的点矩 $\boldsymbol{r} \wedge \boldsymbol{F}$ 在 z 轴上的投影:

$$M_z = \langle \boldsymbol{e}_z, \boldsymbol{r} \wedge \boldsymbol{F} \rangle$$

命题 10.2.4

M_z 的大小不依赖于 z 轴上 O 点的选取.

证明

另设 z 轴上一点 O', 并设 $\boldsymbol{r}' = \boldsymbol{x} - O'$, 则

$$M_z' = \langle \boldsymbol{e}_z, \boldsymbol{r}' \wedge \boldsymbol{F} \rangle = \langle \boldsymbol{e}_z \wedge \boldsymbol{r}', \boldsymbol{F} \rangle = \langle \boldsymbol{e}_z \wedge \boldsymbol{r}, \boldsymbol{F} \rangle = M_z$$

□

定理 10.2.7

在对 z 轴为轴对称的保守场中的运动, 速度矢量关于 z 轴的矩守恒.

证明

知

$$M_z = \langle \boldsymbol{e}_z, \boldsymbol{r} \wedge \dot{\boldsymbol{r}} \rangle$$

因为 $\ddot{\boldsymbol{r}} = \boldsymbol{F}$ 与 $\boldsymbol{r}$ 共面, 进而与 $\boldsymbol{e}_z$ 共面, 从而 $\boldsymbol{r} \wedge \ddot{\boldsymbol{r}}$ 与 $\boldsymbol{e}_z$ 垂直, 故

$$\dot{M}_z = \langle \boldsymbol{e}_z, \dot{\boldsymbol{r}} \wedge \dot{\boldsymbol{r}} \rangle + \langle \boldsymbol{e}_z, \boldsymbol{r} \wedge \ddot{\boldsymbol{r}} \rangle = 0$$

□

10.2.7 n 质点力学系的运动

本节证明 E^3 中的质量组也有能量, 动量和角动量的守恒性.

10.2.7.1 内力和外力

质量为 m_i, 动径为 $\boldsymbol{r}_i \in E^3 (i = 1, \cdots, n)$ 的 n 质点力学系的 Newton 运动方程为:

$$m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i, \quad i = 1, 2, \cdots, n$$

其中矢量 $\boldsymbol{F}_i$ 称为作用在第 i 个质点上的力. 力 $\boldsymbol{F}_i$ 需要通过实验确定.

定义 10.2.11 (相互作用力)

在一个力学系中, 对于两个质点, 一对力大小相等, 作用在两点连线上而方向相反, 则这对力称为相互作用力.

定义 10.2.12 (封闭)

如果作用于力学系中的每个点上的力都是相互作用力, 就称这个力学系是封闭的.

根据定义, 作用在一个封闭系中的第 i 个质点上的力可以写成:

$$\boldsymbol{F}_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \boldsymbol{F}_{ij}$$

其中力 $\boldsymbol{F}_{ij}$ 是第 j 个质点作用在第 i 个质点上的力. 根据相互作用力的定义, 因为力 $\boldsymbol{F}_{ij}$ 与力 $\boldsymbol{F}_{ji}$ 的方向相反, 且 $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$, 故可把它们写成 $\boldsymbol{F}_{ij} = f_{ij}\boldsymbol{e}_{ij}$, 其中 $f_{ij} = f_{ji}$ 是力的大小, $\boldsymbol{e}_{ij}$ 是由第 i 个质点指向第 j 个质点的单位矢量.

若一个力学系不封闭, 则可以将作用在其中某个质点上的力写成:

$$\boldsymbol{F}_i = \sum \boldsymbol{F}_{ij} + \boldsymbol{F}'_i$$

其中 $\boldsymbol{F}_{ij}$ 是相互作用力, 而 $\boldsymbol{F}'_i(\boldsymbol{r}_i)$ 是所谓的外力.

10.2.7.2 动量守恒定律

定义 10.2.13 (动量)

一个力学系的动量定义为矢量:

$$\boldsymbol{P}=\sum_{i=1}^{m} m_i \dot{\boldsymbol{r}}_i$$

♣

定理 10.2.8

一个力学系的动量的变化率等于作用于其上各点的外力之和.

♡

证明

$$\frac{d\boldsymbol{P}}{dt}=\sum_{i=1}^{n} m_i \ddot{\boldsymbol{r}}_i=\sum_{i=1}^{n}\boldsymbol{F}_i=\sum_{i,j}\boldsymbol{F}_{ij}+\sum_{i}\boldsymbol{F}_i'=\sum_{i}\boldsymbol{F}_i'$$

其中最后一个等号是因为相互作用力相互抵消. □

推论 10.2.2

封闭系统的动量守恒.

♡

推论 10.2.3

如果作用于一个力学系的外力之和垂直于 x 轴, 则动量在 x 轴上的投影 P_x 守恒.

♡

定义 10.2.14 (质心)

一个力学系的质心定义为点

$$\boldsymbol{r}=\frac{\sum m_i \boldsymbol{r}_i}{\sum m_i}$$

♣

命题 10.2.5

力学系的动量等于位于其质心, 质量为 $\sum m_i$ 的质点的动量.

♠

定理 10.2.9

质心的运动等同于全部质量和外力均集中于它所得的运动.

♡

推论 10.2.4

如果一个力学系是封闭的, 则其质心作匀速直线运动.

♡

10.2.7.3 角动量守恒

定义 10.2.15 (质点对点的角动量, 力学系对点的角动量)

质量为 m 的质点相对于点 O 的角动量为其动量矢量关于 O 点的矩:

$$\boldsymbol{M}=\boldsymbol{r}\wedge m\dot{\boldsymbol{r}}$$

力学系关于 O 的角动量即其中各质点角动量之和:

$$\boldsymbol{M}=\sum_{i=1}^{n}\boldsymbol{r}_i\wedge m_i\dot{\boldsymbol{r}}_i$$

♣

定理 10.2.10

力学系角动量的变化率等于作用于其搁置电商的外力力矩之和.

证明

$$\frac{d\boldsymbol{M}}{dt}=\sum_{i=1}^{n}\dot{\boldsymbol{r}}_i\wedge m_i\dot{\boldsymbol{r}}_i+\sum_{i=1}^{n}\boldsymbol{r}_i\wedge m_i\ddot{\boldsymbol{r}}_i$$

其中第一项为 0, 而第二项代入 Newton 运动方程有

$$\sum_{i=1}^{n}\boldsymbol{r}_i\wedge\boldsymbol{F}_i=\sum_{i=1}^{n}\boldsymbol{r}_i\wedge(\sum_{i\neq j}\boldsymbol{F}_{ij}+\boldsymbol{F}'_i)=\sum_{i=1}^{n}\boldsymbol{r}_i\wedge\boldsymbol{F}'_i$$

其中

$$\boldsymbol{r}_i\wedge\boldsymbol{F}_{ij}+\boldsymbol{r}_j\wedge\boldsymbol{F}_{ji}=(\boldsymbol{r}_i-\boldsymbol{r}_j)\wedge\boldsymbol{F}_{ij}=0$$

故两相互作用力的力矩之和为零, 进而

$$\sum_{i=1}^{n}\boldsymbol{r}_i\wedge\sum_{i\neq j}\boldsymbol{F}_{ij}=0$$

也即一切相互作用力力矩之和均为零, 从而

$$\frac{d\boldsymbol{M}}{dt}=\sum_{i=1}^{n}\boldsymbol{r}_i\wedge\boldsymbol{F}'_i.$$

□

推论 10.2.5 (角动量守恒定律)

如果一个力学系是封闭的, 那么它的角动量 $\boldsymbol{M}$ 为常数.

推论 10.2.6

如果外力相对于 z 轴的矩为零, 那么 m_z 为常数.

10.2.7.4 能量守恒定律

定义 10.2.16 (质点与质点组的动能)

质量为 m 的质点的动能为

$$T=\frac{m\dot{\boldsymbol{r}}^2}{2}$$

质点组的动能为各点动能之和:

$$T=\sum_{i=1}^{n}\frac{m_i\dot{\boldsymbol{r}}_i^2}{2}$$

其中 m_i 为单个质点的质量, $\dot{\boldsymbol{r}}_i$ 为该质点的速度.

定理 10.2.11

力学系动能的增量等于作用于其各点上的力所做的功之和.

证明

$$\frac{dT}{dt}=\sum_{i=1}^{n}m_i\langle\dot{\boldsymbol{r}}_i,\ddot{\boldsymbol{r}}_i\rangle=\sum_{i=1}^{n}\langle\dot{\boldsymbol{r}}_i,m_i\ddot{\boldsymbol{r}}_i\rangle=\sum_{i=1}^{n}\langle\dot{\boldsymbol{r}}_i,\boldsymbol{F}_i\rangle$$

故两边积分有

$$T(t)-T(t_0)=\int_{t_0}^{t}\frac{dT}{dt}dt=\sum_{i=1}^{n}\int_{t_0}^{t}\langle\dot{\boldsymbol{r}}_i,\boldsymbol{F}_i\rangle dt=\sum_{i=1}^{n}A_i$$

□

E^3 中 n 个质点的力学系的构形空间是 n 个欧氏空间的直积: $E^{3n}=E^3\times\cdots\times E^3$, 构形空间也有欧式空间的构造. 令 $\boldsymbol{r}=(\boldsymbol{r}_1,\cdots,\boldsymbol{r}_n)$ 为构形空间中一点的动径, $\boldsymbol{F}=(\boldsymbol{F}_1,\cdots,\boldsymbol{F}_n)$ 是力矢量, 则可将前述定理写成:

$$T(t_1)-T(t_0)=\int_{\boldsymbol{r}(t_0)}^{\boldsymbol{r}(t_1)}\langle\boldsymbol{F},d\boldsymbol{r}\rangle=\int_{t_0}^{t_1}\langle\dot{\boldsymbol{r}},\boldsymbol{F}\rangle dt$$

也就是说, 动能的增量等于构形空间中的" 力"$\boldsymbol{F}$ 沿" 路径"$\boldsymbol{r}(t)$ 做的功.

定义 10.2.17 (保守系)

力学系称为保守的, 如果其中的力只依赖于该力学系中一点的位置, 即 $\boldsymbol{F}=\boldsymbol{F}(\boldsymbol{r})$, 且 $\boldsymbol{F}$ 沿任一路径的功只依赖于路径的起点与终点:

$$\int_{M_1}^{M_2}\langle\boldsymbol{F},d\boldsymbol{r}\rangle=\Phi(M_1,M_2)$$

♣

定理 10.2.12

力学系是保守系的那个且仅当其中存在位能, 也即存在函数 $U(\boldsymbol{r})$ 使得

$$\boldsymbol{F}=-\frac{\partial U(\boldsymbol{r})}{\partial\boldsymbol{r}}$$

♡

定理 10.2.13

保守系的总能量 $(E=T+U)$ 在运动中守恒, 也即 $E(t_1)=E(t_0)$.

♡

证明

知

$$T(t_1)-T(t_0)=\int_{\boldsymbol{r}(t_0)}^{\boldsymbol{r}(t_1)}\langle\boldsymbol{F},d\boldsymbol{r}\rangle=U(\boldsymbol{r}(t_0))-U(\boldsymbol{r}(t_1))$$

□

现在把作用在一个力学系上各点的力分为相互作用力和外力:

$$\boldsymbol{F}_i=\sum_{i\neq j}\boldsymbol{F}_{ij}+\boldsymbol{F}_i'$$

其中

$$\boldsymbol{F}_{ij}=-\boldsymbol{F}_{ji}=f_{ij}\boldsymbol{e}_{ij}$$

命题 10.2.6

如果相互作用力只依赖于距离, 也即 $f_{ij}=f_{ij}(|\boldsymbol{r}_i-\boldsymbol{r}_j|)$, 则这些力是保守的.

♠

证明[4]

如果一个力学系中只有两点 i,j, 则它们的相互位能为

$$U_{ij}(\boldsymbol{r})=-\int_{r_0}^{r_1}f_{ij}(p)dp$$

两边对 $\boldsymbol{r}$ 求导有

$$-\frac{\partial U_{ij}(|\boldsymbol{r}_i-\boldsymbol{r}_j|)}{\partial\boldsymbol{r}_i}=f_{ij}\frac{\partial|\boldsymbol{r}_i-\boldsymbol{r}_j|}{\partial\boldsymbol{r}_i}=f_{ij}\boldsymbol{e}_{ij}$$

[4] 这个证明没有看懂.

故所有点的相互作用位能为

$$U(\boldsymbol{r})=\sum_{i>j}U_{ij}(|\boldsymbol{r}_i-\boldsymbol{r}_j|)$$

□

对一个力学系而言, 如果它其中的相互作用力只依赖于距离, 其所受到的外力也是保守 (也即 $\boldsymbol{F}_i'=-\frac{\partial U_i'}{\partial \boldsymbol{r}_i}$), 则该力学系是保守系, 且其总位能为

$$U(\boldsymbol{r})=\sum_{i>j}U_{ij}+\sum_i U_i'$$

对这样的力学系而言, 总机械能

$$E=T+U=\sum_i\frac{\boldsymbol{r}_i^2}{2}+\sum_{i>j}U_{ij}+\sum_i U_i'$$

是守恒的.

但如果力学系不是保守的, 则总机械能一般来说不守恒.

定义 10.2.18 (非机械能增量)

机械能的减少量 $E(t_0)-E(t_1)$ 称为非机械能 E' 的增量:

$$E'(t_1)-E'(t_0)=E(t_0)-E(t_1).$$

♣

定理 10.2.14 (能量守恒定律)

总能量 $H=E+E'$ 守恒.

♡

第十一章 Lagrange 力学

11.1 变分原理

本节说明 Newton 有势力学系的运动是一个变分原理"Hamilton 最小作用原理" 中的驻定曲线, 在符号上, 默认使用 n 维坐标空间, 矢量 $\boldsymbol{x} = (x_1, \cdots, x_n)$, $\frac{\partial f}{\partial \boldsymbol{x}} = (\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n})$, $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = a_1 b_1 + \cdots + a_n b_n$.

11.1.1 变分法

定义 11.1.1 (泛函可微性, 变分)

若泛函 Φ 满足

$$\Phi(\gamma + h) - \Phi(\gamma) = F + R$$

其中 F 线性依赖于 h, 即当 γ 固定时有

$$F(c_1 h_1 + c_2 h_2) = c_1 F(h_1) + c_2 F(h_2)$$

而 $R(h, \gamma) = O(h^2)$, 即

$$|h| < \varepsilon, |\frac{dh}{dt}| < \varepsilon \Rightarrow |R| < C\varepsilon^2$$

则称 Φ 是可微的. 增量的线性部分 $F(h)$ 称为 Φ 的微分. 泛函的微分又称为其变分, h 称为曲线的变分. ♣

注 可以说明若 Φ 可微, 则其微分是唯一的.

例 11.1 设 $\gamma = \{(t, x) : x = x(t), t_0 \leq t \leq t_1\}$ 是 (t, x) 平面上的曲线, $\dot{x} = \frac{dx}{dt}$, $L = L(a, b, c)$ 是关于三个变元的可微函数. 定义泛函 Φ:

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t)dt$$

一个特例为 $L(a, b, c) = \sqrt{1 + b^2}$, 此时上述泛函 $\Phi(\gamma)$ 表示曲线 γ 的长度. 下面说明泛函 $\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt$ 可微, 且其微分为

$$F(h) = \int_{t_0}^{t_1} (\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}})hdt + (\frac{\partial L}{\partial \dot{x}}h)|_{t_0}^{t_1}$$

根据定义, 在 γ 上取微小增量 h 有

$$\begin{aligned}
\Phi(\gamma + h) - \Phi(\gamma) &= \int_{t_0}^{t_1} (L(x + h, \dot{x} + \dot{h}, t) - L(x, \dot{x}, t))dt = \int_{t_0}^{t_1} \\
&= \int_{t_0}^{t_1} (L(x + h, \dot{x} + \dot{h}, t) - L(x, \dot{x} + \dot{h}, t) + L(x, \dot{x} + \dot{h}, t) - L(x, \dot{x}, t))dt \\
&= \int_{t_0}^{t_1} (\frac{\partial L}{\partial x}h + \frac{\partial L}{\partial \dot{x}}\dot{h})dt + O(h^2) = F(h) + R
\end{aligned}$$

其中

$$F(h) = \int_{t_0}^{t_1} (\frac{\partial L}{\partial x}h + \frac{\partial L}{\partial \dot{x}}\dot{h})dt, \quad R = O(h^2)$$

对 $\int_{t_0}^{t_1} \frac{\partial L}{\partial \dot{x}}\dot{h}dt$ 分部积分得

$$\int_{t_0}^{t_1} \frac{\partial L}{\partial \dot{x}}\dot{h}dt = -\int_{t_0}^{t_1} h\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})dt + h\frac{\partial L}{\partial \dot{x}}|_{t_0}^{t_1}$$

代回 F 表达式即得欲证.

定义 11.1.2 (驻定曲线)

设可微泛函 Φ 的微分为 $F(h,\gamma)$, $\Phi(\gamma)$ 的驻定曲线 γ 指: 使得对一切 h 都有 $F(h,\gamma)=0$ 的曲线[a].

[a]类比到可微函数来看, 如果一个函数在某点微分为零, 则称该点为其驻点.

定理 11.1.1

曲线 $\gamma: x=x(t)$ 是泛函 $\Phi(\gamma)=\int_{t_0}^{t_1} L(x,\dot{x},t)dt$ 过 $x(t_0)=x_0, x(t_1)=x_1$ 两点的驻定曲线, 当且仅当沿曲线 $x(t)$ 有

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})-\frac{\partial L}{\partial x}=0$$

引理 11.1.1

若连续函数 $f(t)(t_0 \le t \le t_1)$ 满足: 对一切满足 $h(t_0)=h(t_1)=0$ 的连续函数 $h(t)$ 而言, 均有 $\int_{t_0}^{t_1} f(t)h(t)dt=0$, 则 $f(t)\equiv 0$.

引理是实变函数论中的结论, 下面利用引理证明前述定理.

证明

已经知道

$$F(h)=\int_{t_0}^{t_1}(\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})-\frac{\partial L}{\partial x})hdt+(\frac{\partial L}{\partial \dot{x}}h)|_{t_0}^{t_1}$$

现取 h 为满足引理条件的函数, 令 $h(t_0)=h(t_1)=0$, 得 $(\frac{\partial L}{\partial \dot{x}}h)|_{t_0}^{t_1}=0$. 当 γ 是驻定曲线, 根据驻定曲线的定义, 对满足 $h(t_0)=h(t_1)=0$ 的一切连续可微函数 h, 都有 $F(h)=0$, 记

$$f(t)=\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})-\frac{\partial L}{\partial x}$$

则 $F(h)=0$ 说明

$$\int_{t_0}^{t_1} f(t)h(t)dt=0$$

从而根据引理知 $f(t)\equiv 0$.

而当 $f(t)\equiv 0$, 显见 $F(h)\equiv 0$, 命题即证. □

从前面 $L(x,\dot{x},t)$ 微分的表达式与驻定曲线的定义可以抽象得到下述方程.

定义 11.1.3 (Euler-Lagrange 方程)

方程

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{x}})-\frac{\partial L}{\partial x}=0$$

称为泛函 $\Phi=\int_{t_0}^{t_1} L(x,\dot{x},t)dt$ 的 Euler-Lagrange 方程.

现设 $\boldsymbol{x}$ 是 n 维坐标空间 $\mathbb{R}^n$ 的一个矢量, $\gamma=\{(t,\boldsymbol{x}):\boldsymbol{x}=\boldsymbol{x}(t), t_0\le t\le t_1\}$ 是 $n+1$ 维空间 $\mathbb{R}\times\mathbb{R}^n$ 中的一条曲线, $L:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ 是一个 $2n+1$ 变量的函数. 于是前述定理的 n 维情况可以表述成:

定理 11.1.2

曲线 γ 是泛函 $\Phi(\gamma)=\int_{t_0}^{t_1} L(\boldsymbol{x},\dot{\boldsymbol{x}},t)$ 连接点 $(t_0,\boldsymbol{x}_0),(t_1,\boldsymbol{x}_1)$ 的驻定曲线, 当且仅当 γ 满足 Euler-Lagrange 方程.

特别注意曲线 γ 为某泛函驻定曲线的这条性质与坐标系的选取无关.

11.1.2 Lagrange 方程组

下面把 Newton 动力学方程

$$\frac{d}{dt}\langle m_i, \dot{\boldsymbol{r}}_i\rangle + \frac{\partial U}{\partial \boldsymbol{r}_i} = 0 \tag{11.1}$$

与 Euler-Lagrange 方程

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{x}}} - \frac{\partial L}{\partial \boldsymbol{x}} = 0$$

相比较.

11.1.2.1 Hamilton 最小作用原理

定理 11.1.3 (最小作用原理的 Hamilton 形式)

力学系(11.1)的运动正是下述泛函的驻定曲线

$$\Phi(\gamma) = \int_{t_0}^{t_1} L dt$$

其中 $L = T - U$ 是动能与位能之差.

证明

代入 $U = U(\boldsymbol{r}), T = \sum \frac{m_i \dot{\boldsymbol{r}}_i^2}{2}$ 有

$$\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} = \frac{\partial T}{\partial \dot{\boldsymbol{r}}_i} = m_i \dot{\boldsymbol{r}}_i, \quad \frac{\partial L}{\partial \boldsymbol{r}_i} = -\frac{\partial U}{\partial \boldsymbol{r}_i}$$

进一步有

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\boldsymbol{r}}_i} = m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i = -\frac{\partial U}{\partial \boldsymbol{r}}_i$$

这说明曲线 $\boldsymbol{r}$ 满足 Euler-Lagrange 方程, 命题即证. □

推论 11.1.1

令 $(q_1, \cdots, q_{3n})$ 是 n 质点力学系的构形空间中的任意坐标, 则 $\boldsymbol{q}$ 随时间的演化满足 Euler-Lagrange 方程:

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}) - \frac{\partial L}{\partial \boldsymbol{q}} = 0, \quad L = T - U$$

证明

根据前述定理, 运动曲线应为泛函 $\int L dt$ 的驻定曲线, 而驻定曲线与坐标系的选取无关, 故运动曲线在任意坐标系下都应满足该坐标系对应的 Euler-Lagrange 方程. □

定义 11.1.4 (Lagrange 函数, 广义坐标, 广义速度, 广义动量, 广义力, 作用量, Lagrange 方程组)

称 $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T - U$ 为 Lagrange 函数, q_i 为广义坐标, $\dot{q}_i$ 为广义速度, $\frac{\partial L}{\partial \dot{q}_i} = p_i$ 为广义动量, $\frac{\partial L}{\partial q_i}$ 为广义力, $\int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) dt$ 为作用量. 称

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = 0$$

为 Lagrange 方程组.

上述定理称为最小作用原理的 Hamilton 形式, 因为在诸多情况下, Euler-Lagrange 方程的解 $\boldsymbol{q}$ 不仅仅是作用量 $\int_{t_0}^{t_1} L dt$ 的驻定曲线, 还能使它达到最小值.

参考文献

[AHCB] A. H. 柯尔莫哥洛夫, C. B. 佛明. 函数论与泛函分析初步 (第七版)[M]. 北京: 高等教育出版社, 2006.

[AR1] V. I. 阿诺尔德. 经典力学的数学方法 (第 4 版)[M]. 北京: 高等教育出版社, 2006.

[CZC] 陈祖墀. 偏微分方程 (第三版)[M]. 北京: 高等教育出版社, 2008.

[CH] R. Courant, D. Hilbert. *Methods of Mathematical Physics(Volume II: Partial Differential Equations)*[M]. New York: John Wiley & Sons, Inc., 1966.

[DY] 丁勇. 现代分析基础 (第 2 版)[M]. 北京: 北京师范大学出版社, 2013.

[Ev] Lawrence C. Evans. *Partial Differential Equations*[M]. 北京: 高等教育出版社, 2017.

[FL] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*[M]. New York: John Wiley & Sons, Inc., 1999.

[GLC] 谷超豪, 李大潜, 陈恕行, 郑宋穆, 谭永基. 数学物理方程 (第三版)[M]. 北京: 高等教育出版社, 2012.

[HB] Haim Brezis. 泛函分析-理论和应用 [M]. 北京: 清华大学出版社, 2009.

[HJR] Hajer Bahouri, Jean-Yves Chemin, Raphael Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*[M]. Berlin: Springer, 2011.

[JD] Javier Duoandikoetxea. *Fourier Analysis*[M]. Rhode Island: American Mathematical Society, 1995.

[JDI] J. 迪厄多内. 泛函分析史 [M]. 北京: 高等教育出版社, 2016.

[LG1] Loukas Grafakos. *Classical Fourier Analysis*[M]. New York: Springer, 2014.

[LG2] Louksa Grafakos. *Modern Fourier Analysis*[M]. New York: Springer, 2014.

[Lin] 林源渠. 泛函分析学习指南 [M]. 北京: 北京大学出版社, 2009.

[LPD] 刘培德. 拓扑线性空间与算子谱理论 [M]. 北京: 高等教育出版社, 2013.

[LS] Laurent Schwartz. 广义函数论 [M]. 北京: 高等教育出版社, 2010.

[LWT] Loring W. Tu. *An Introduction to Manifolds*[M]. New York: Springer, 2011.

[MCX1] 苗长兴. 调和分析及其在偏微分方程中的应用 (第二版)[M]. 北京: 科学出版社, 2004.

[OAO] O. A. 奥列尼克. 偏微分方程讲义 (第 3 版)[M]. 北京: 高等教育出版社, 2008.

[PFHYS] Po-Fang Hsieh, Yasutaka Sibuya. *Basic Theory of Ordinary Differential Equations*[M]. New York: Springer, 1991.

[PGC] Philippe G. Ciarlet. 线性与非线性泛函分析及其应用 [M]. 北京: 高等教育出版社, 2020.

[QMY] 齐民友, 吴方同. 广义函数与数学物理方程 (第二版)[M]. 北京: 高等教育出版社, 1998.

[RAJF] Robert A. Adams, John J. F. Fournier. *Sobolev Spaces*[M]. Ocford: Elsevier Ltd., 2003.

[SA] Serge Alinhac. *Hyperbolic Partial Differential Equations*[M]. New York: Springer, 2009.

[SL] Serge Lang. *Real and Functional Analysis*[M]. New York: Springer, 1993.

[ST1] Elias M. Stein, Rami Shakarchi. *Fourier Analysis: An Introduction*[M]. New Jersey: Princeton University Press, 2003.

[ST2] Elias M. Stein, Rami Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*[M]. New Jersey: Princeton University Press, 2003.

[ST4] Elias M. Stein, Rami Shakarchi. *Functional Analysis: Introduction to Further Topics in Analysis*[M]. New Jersey: Princeton University Press, 2011.

[TL2] Michael E. Taylor. *Partial Differential Equations II: Qualitative Studies of Linear Equations*[M]. New York: Springer, 2011.

[WL] 汪林. 泛函分析中的反例 [M]. 北京: 高等教育出版社, 2014.

[WKY] 孙永生, 王昆扬. 泛函分析讲义 [M]. 北京: 北京师范大学出版社, 2007.

[WR] Walter Rudin. *Functional Analysis*[M]. New York: McGraw-Hill, Inc., 1991.

[XDX] 夏道行, 吴卓人, 严绍宗, 舒五昌. 实变函数论与泛函分析 (下册)[M]. 北京: 高等教育出版社, 2010.

[YOS] Kosaku Yosida. *Functional Analysis*[M]. Berlin: Springer, 1980.

[You] 尤承业. 基础拓扑学讲义 [M]. 北京: 北京大学出版社, 1997.

[YR] 袁荣. 常微分方程 [M]. 北京: 高等教育出版社, 2012.

[ZL] 张恭庆, 林源渠. 泛函分析讲义 (第二版)(上)[M]. 北京: 北京大学出版社, 2021.

[ZMQ] 周民强. 实变函数论 (第三版)[M]. 北京: 北京大学出版社, 2016.

[Zo] B. A. 卓里奇. 数学分析 (第二卷)(第四版)[M]. 北京: 高等教育出版社, 2019.