



面向 21 世纪 课 程 教 材
Textbook Series for 21st Century

高等数学

下 册

仇庆久 编

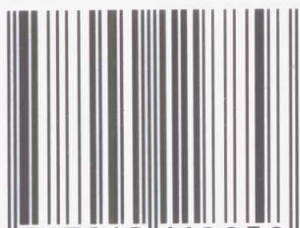


高等教育出版社
HIGHER EDUCATION PRESS

Advanced Mathematics



ISBN 7-04-011885-8



9 787040 118858 >

定价 15.10 元

面向 21 世纪课程教材
Textbook Series for 21st Century

高等数学

下册

仇庆久 编



高等教育出版社
HIGHER EDUCATION PRESS

内容提要

本书是教育部“面向 21 世纪教学内容和课程体系改革计划”的研究成果,是面向 21 世纪课程教材,同时也是普通高等教育“九五”国家级重点教材。

本教材思路新,观点高,基础扎实,知识面宽,并注意运用近代数学的思想解决问题。本书分上、下两册。下册内容包括:级数——函数的分析与研究Ⅲ,微分方程与傅里叶积分,空间的结构与现代数学等。

本书可作为高等院校理科各专业的教材,也可供各类专业人员学习参考。

图书在版编目(CIP)数据

高等数学.下册/仇庆久编. —北京:高等教育出版社,2003.5

ISBN 7-04-011885-8

I. 高... II. 仇... III. 高等数学-高等学校-教材 IV. 013

中国版本图书馆 CIP 数据核字(2003)第 009834 号

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010-82028899

购书热线 010-64054588
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>

经 销 新华书店北京发行所
排 版 高等教育出版社照排中心
印 刷 化学工业出版社印刷厂

开 本 787×960 1/16
印 张 14
字 数 250 000

版 次 2003 年 5 月第 1 版
印 次 2003 年 5 月第 1 次印刷
定 价 15.10 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

责任编辑	薛春玲
封面设计	张楠
责任绘图	吴文信
版式设计	史新薇
责任校对	尤静
责任印制	孔源

郑 重 声 明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 82028899 转 6897 (010)82086060

传真：(010) 82086060

E-mail: dd@hep.com.cn

通信地址：北京市西城区德外大街 4 号

高等教育出版社法律事务部

邮编：100011

购书请拨打读者服务部电话：(010)64054588



面向 21 世纪课程教材



普通高等教育“九五”
国家级重点教材

目 录

第四章 级数——函数的分析与研究Ⅲ	(1)
§ 4.1 数项级数	(1)
4.1.1 常数项级数	(1)
4.1.2 正项级数	(4)
4.1.3 任意项级数	(10)
习题	(14)
§ 4.2 函数项级数	(16)
4.2.1 函数项级数	(16)
4.2.2 幂级数、泰勒级数	(23)
4.2.3 多元函数的泰勒公式与泰勒级数	(38)
4.2.4 复级数、洛朗级数	(39)
习题	(48)
§ 4.3 傅里叶级数	(50)
4.3.1 直交(正交)函数系	(50)
4.3.2 傅里叶级数的定义	(54)
4.3.3 傅里叶级数的逐点收敛定理	(56)
4.3.4 任意有限区间上的傅里叶级数	(58)
4.3.5 傅里叶级数的逐项求积与逐项求导数	(60)
4.3.6 傅里叶级数的极值性质与贝塞尔不等式	(62)
习题	(65)
第五章 微分方程与傅里叶分析	(67)
§ 5.1 微分方程	(67)
5.1.1 建模实例,微分方程的通解、特解	(67)
5.1.2 一阶常微分方程的几种特殊类型	(73)
5.1.3 高阶常系数线性微分方程	(78)
5.1.4 隐函数方程	(92)
5.1.5 存在性与惟一性定理	(99)
5.1.6 常微分方程数值解法	(103)
习题	(108)
§ 5.2 偏微分方程	(110)
5.2.1 一阶方程	(110)
5.2.2 双曲、椭圆、抛物型方程	(118)
习题	(122)
§ 5.3 傅里叶变换	(122)

5.3.1 预备知识.....	(122)
5.3.2 $L^1(\mathbf{R})$ 上的傅里叶变换与卷积运算.....	(124)
5.3.3 $L^1(\mathbf{R})$ 上傅里叶变换的反演.....	(128)
5.3.4 $L^2(\mathbf{R})$ 上的傅里叶变换.....	(130)
习题.....	(131)
§ 5.4 小波变换简介.....	(131)
5.4.1 小波变换的引入.....	(131)
5.4.2 连续小波变换与时频分析.....	(134)
5.4.3 离散小波变换.....	(138)
5.4.4 小波反演公式.....	(139)
5.4.5 多分辨分析,小波应用概述.....	(141)
习题.....	(142)
第六章 空间结构与现代数学	(143)
§ 6.1 集合.....	(143)
§ 6.2 运算结构.....	(146)
6.2.1 群、环、域.....	(146)
6.2.2 线性空间.....	(149)
6.2.3 方阵空间.....	(150)
习题.....	(181)
§ 6.3 拓扑结构.....	(183)
6.3.1 距离空间.....	(183)
6.3.2 赋范线性空间.....	(185)
§ 6.4 分布理论.....	(185)
6.4.1 连续线性泛函.....	(186)
6.4.2 分布空间,分布的性质.....	(187)
6.4.3 $\mathcal{S}'$ 上的傅里叶分析.....	(189)
§ 6.5 Banach 空间微积分学.....	(191)
6.5.1 极限.....	(192)
6.5.2 导数.....	(192)
6.5.3 级数.....	(195)
6.5.4 偏导数和高阶导数.....	(197)
6.5.5 压缩映射定理和不动点定理.....	(198)
§ 6.6 非线性科学.....	(199)
附录	(206)
Fourier 变换表.....	(206)
习题答案.....	(209)
参考文献	(214)

第四章 级数

——函数的分析与研究Ⅲ

级数,与导数、微分、积分一样,是分析与研究函数的重要工具.通过把较为复杂的函数“分解”(或“表示”)为简单的函数,达到研究函数各种性质的目的.这里所指的“分解”或“表示”,是通过极限过程,引进收敛与发散概念进行的.本章首先介绍常数项级数的收敛与发散、判别法,包括正项级数与任意项级数;然后介绍函数项级数,包括幂级数、泰勒级数复级数与洛朗级数,以及它们的收敛与发散判别法、重要性质及应用.

§ 4.1 数项级数

4.1.1 常数项级数

1. 基本概念

定义 4.1.1 (级数、部分和、通项) 给定一个实数列 $a_1, a_2, \dots, a_n, \dots$, 我们把用加号依次联接它的各项而得到的式子

$$a_1 + a_2 + \dots + a_n + \dots$$

称为无穷级数,简称为级数,并记为

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \dots + a_n + \dots. \quad (4.1.1)$$

上式中的 a_n 是级数(4.1.1)的第 n 项,称为级数的一般项或通项.令

$$s_n = a_1 + a_2 + \dots + a_n;$$

s_n 为级数(4.1.1)的前 n 项的和,简称部分和.部分和 s_n 构成的数列

$$\{s_n\}: s_1, s_2, \dots, s_n, \dots \quad (4.1.2)$$

称为部分和数列.

级数通项 a_n 的下标也可从 0 或大于 1 的某正整数开始,例如级数

$$\sum_{n=k}^{\infty} a_n = a_k + a_{k+1} + \dots + a_{k+n} + \dots,$$

这里 k 是某个非负整数.

定义 4.1.2 (级数的敛散性(收敛性与发散性)、级数的和) 若级数(4.1.1)的部分和数列(4.1.2)当 $n \rightarrow \infty$ 时有有限极限 s , 即

$$\lim_{n \rightarrow \infty} s_n = s,$$

则称级数 $\sum_{n=1}^{\infty} a_n$ 收敛于 s . 记为

$$\sum_{n=1}^{\infty} a_n = s.$$

称 s 为级数(4.1.1)的和. 若数列 $\{s_n\}$ 的极限为 $\pm\infty$ 、或 ∞ 、或不存在, 则称级数(4.1.1)发散; 当此极限为 $+\infty$ (或 $-\infty$) 时, 也称级数(4.1.1)发散到 $+\infty$ (或 $-\infty$), 记为

$$\sum_{n=1}^{\infty} a_n = +\infty \text{ (或 } -\infty \text{)}.$$

例 1 讨论几何级数(等比级数)

$$\sum_{n=0}^{\infty} aq^n = a + aq + aq^2 + \cdots + aq^n + \cdots \quad (4.1.3)$$

的敛散性, 这里 $a \neq 0$.

解 当 $q=1$ 时, 部分和 $s_n = na \rightarrow \infty$. 当 $q=-1$ 时, 部分和

$$s_n = \begin{cases} a, & n \text{ 为奇数,} \\ 0, & n \text{ 为偶数,} \end{cases}$$

故 $\lim_{n \rightarrow \infty} s_n$ 不存在. 因此, 当 $|q|=1$ 时级数发散.

当 $|q| \neq 1$ 时, 因

$$s_n = a + aq + aq^2 + \cdots + aq^{n-1} = \frac{a(1-q^n)}{1-q},$$

故当 $|q| < 1$ 时, $s_n \rightarrow a$, 级数收敛; 当 $|q| > 1$ 时, $s_n \rightarrow \infty$, 级数发散.

综上所述, 几何级数(4.1.3)当且仅当 $|q| < 1$ 时收敛, 且其和为 $\frac{a}{1-q}$.

例 2 讨论级数

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} + \cdots$$

的敛散性.

解 因 $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, 于是, 部分和

$$\begin{aligned} s_n &= \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}. \end{aligned}$$

因此

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n+1}\right) = 1.$$

所以原级数收敛于 1.

为书写简便,在不引起混淆时,常把级数 $\sum_{n=1}^{\infty} a_n$ 简记为 $\sum_1^{\infty} a_n$,甚至为 $\sum a_n$.

2. 收敛级数的基本性质

性质 1(收敛级数的必要条件) 若级数 $\sum_{n=1}^{\infty} a_n$ 收敛,则

$$\lim_{n \rightarrow \infty} a_n = 0.$$

证 设 s_n 为级数的前 n 项和,则

$$a_n = s_n - s_{n-1}.$$

设 $s_n \rightarrow s$, s 为有限数. 同样有 $s_{n-1} \rightarrow s$, 于是

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = s - s = 0.$$

因此,级数通项的极限等于 0 是级数收敛的必要条件.

性质 2(收敛级数与常数之积) 设 $\sum_{n=1}^{\infty} a_n$ 收敛于 s , $c \neq 0$ 为常数,则 $\sum_{n=1}^{\infty} ca_n$ 收敛于 cs .

证 设级数 $\sum_{n=1}^{\infty} ca_n$ 的前 n 项和为 s'_n , 则

$$s'_n = ca_1 + ca_2 + \cdots + ca_n = c(a_1 + a_2 + \cdots + a_n) = cs_n.$$

于是

$$\sum_{n=1}^{\infty} ca_n = \lim_{n \rightarrow \infty} cs_n = c \lim_{n \rightarrow \infty} s_n = cs.$$

因此,当 $c \neq 0$ 时,级数 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} ca_n$ 具有相同的敛散性. 当其中有一个收敛时,则

$$\sum_{n=1}^{\infty} ca_n = c \sum_{n=1}^{\infty} a_n.$$

性质 3(收敛级数之和与差) 设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 均收敛,其和分别为 a 与 b , 则 $\sum_{n=1}^{\infty} (a_n \pm b_n)$ 收敛,且和为 $a \pm b$.

证 由收敛的定义便可证明.

性质 4(增、减有限项不影响敛散性) 在级数前加上或去掉有限多项,不改变级数的敛散性(但在收敛的情形,级数的和一般会改变).

证 只要考虑去掉前有限多项的情形就够了. 设级数(4.1.1)去掉前 k 项后得到级数

$$\sum_{n=k+1}^{\infty} a_n = a_{k+1} + a_{k+2} + \cdots + a_{k+m} + \cdots. \quad (4.1.4)$$

又设级数(4.1.1)与(4.1.4)的前 n 项分别为 s_n 与 s'_n , 于是 $s'_n = s_{k+n} - s_k$. 因 s_k 为常数, 由上式可知, 若级数(4.1.1)收敛于 s , 则

$$\sum_{n=k+1}^{\infty} a_n = \lim_{n \rightarrow \infty} s'_n = \lim_{n \rightarrow \infty} (s_{k+n} - s_k) = s - s_k.$$

反之, 若级数(4.1.4)收敛于 s' , 类似地可以证明级数(4.1.1)收敛于

$$s' + s_k.$$

由性质 4 可知, 改变某级数的有限多项不改变级数的敛散性.

性质 5 (收敛级数可加括号) 收敛级数任意“加上括号”后得到的级数仍收敛到原来的和.

例如级数(4.1.1)加上某种括号后得到级数

$$(a_1 + a_2) + a_3 + (a_4 + a_5 + a_6) + \cdots,$$

该级数第一项为 $(a_1 + a_2)$, 第二项为 a_3 , 第三项为 $(a_4 + a_5 + a_6)$, $\cdots$, 显然新级数的部分和数列是原级数部分和数列的子数列. 因收敛数列的任一子数列仍收敛, 并收敛到原数列的极限, 故加括号得到的级数仍收敛到原级数的和.

4.1.2 正项级数

如何判别级数的敛散性是我们要讨论的基本问题. 在上文中举了两个用级数敛散性的定义判别级数敛散性的例子, 但是能用定义方便地判别级数的敛散性并求得和是不多的, 因此需要建立敛散性的判别法. 当判定了某级数收敛以后, 再讨论如何计算其和才有意义. 我们先讨论正项级数的情形.

定义 4.1.3 (正项级数) 若级数 $\sum_{n=1}^{\infty} a_n$ 的各项均非负, 即 $a_n \geq 0, n \in \mathbf{N}$, 则称该级数为正项级数.

设 $\sum_{n=1}^{\infty} a_n$ 为正项级数, s_n 为它的部分和, 则

$$s_{n+1} = s_n + a_{n+1} \geq s_n.$$

故部分和数列 s_n 为递增数列, 因递增数列收敛的充要条件是该数列为有界数列 (第二章定理 2.2.2), 于是我们有

定理 4.1.1 (正项级数收敛的充要条件) 正项级数收敛的充要条件是其部分和数列有界.

由定理 4.1.1, 我们可以导出判别正项级数收敛的几种常用方法.

1. 比较判别法

设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 为两个正项级数. 若

$$a_n \leq b_n \quad (n=1, 2, \cdots), \quad (4.1.5)$$

则

(i) 当 $\sum_{n=1}^{\infty} b_n$ 收敛时, $\sum_{n=1}^{\infty} a_n$ 也收敛;

(ii) 当 $\sum_{n=1}^{\infty} a_n$ 发散时, $\sum_{n=1}^{\infty} b_n$ 也发散.

证 设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 前 n 项和分别为 s_n 与 s'_n . 若 $\sum_{n=1}^{\infty} b_n$ 收敛, 由定理 4.1.1, 数列 $\{s'_n\}$ 有界. 再由 (4.1.5) 式得, $s_n \leq s'_n$, 故数列 s_n 有界, 所以 $\sum_{n=1}^{\infty} a_n$ 收敛, 即 (i) 成立.

因 (ii) 是 (i) 的逆否命题, 它们是等价的, 由 (i) 即知 (ii) 成立. 定理得证.

注 因去掉级数若干项不改变级数的敛散性, 故 (4.1.5) 式只要对充分大的 n 成立, 就有同样的结论.

例 3 设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 均为收敛的正项级数. 求证 $\sum_{n=1}^{\infty} a_n^2$ 与 $\sum_{n=1}^{\infty} \sqrt{a_n b_n}$ 也收敛.

证 由收敛级数的必要条件知

$$\lim_{n \rightarrow \infty} a_n = 0.$$

故当 n 充分大时, $0 \leq a_n < 1$. 于是对充分大的 n 有

$$a_n^2 < a_n$$

成立. 由比较判别法 (i), 知 $\sum_{n=1}^{\infty} a_n^2$ 收敛. 又因

$$\sqrt{a_n b_n} \leq \frac{1}{2}(a_n + b_n) \leq a_n + b_n,$$

故由 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 收敛知 $\sum_{n=1}^{\infty} \sqrt{a_n b_n}$ 收敛.

例 4 设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 均为正项级数, 若存在正常数 c_1 与 c_2 , 使得

$$c_1 b_n < a_n < c_2 b_n$$

对充分大的 n 成立, 求证级数 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 有相同的敛散性.

证 由比较判别法, 若 $\sum_{n=1}^{\infty} a_n$ 收敛, 则 $\sum_{n=1}^{\infty} c_1 b_n$ 收敛, 故由收敛级数性质 2

知, $\sum_{n=1}^{\infty} b_n$ 收敛.

若 $\sum_{n=1}^{\infty} a_n$ 发散, 则 $\sum_{n=1}^{\infty} c_2 b_n$ 发散, 故 $\sum_{n=1}^{\infty} b_n$ 发散. 证毕.

2. 比较判别法的极限形式

设 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 均为正项级数, 且 $b_n \neq 0$. 若

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \lambda, \quad (4.1.6)$$

则当 $0 < \lambda < +\infty$ 时, $\sum_{n=1}^{\infty} b_n$ 与 $\sum_{n=1}^{\infty} a_n$ 有相同的敛散性; 当 $0 \leq \lambda < +\infty$ 时,

$\sum_{n=1}^{\infty} b_n$ 的收敛性蕴含 $\sum_{n=1}^{\infty} a_n$ 的收敛性; 当 $0 < \lambda \leq +\infty$ 时, $\sum_{n=1}^{\infty} b_n$ 的发散性蕴含

$\sum_{n=1}^{\infty} a_n$ 的发散性.

证 我们只给出假设(4.1.6)下定理的证明. 因 $\frac{\lambda}{2} < \lambda < 2\lambda$, 故当 n 充分大时, 有

$$\frac{\lambda}{2} < \frac{a_n}{b_n} < 2\lambda, \quad \text{或} \quad \frac{\lambda}{2} b_n < a_n < 2\lambda b_n.$$

由例 4 知 $\sum_{n=1}^{\infty} b_n$ 与 $\sum_{n=1}^{\infty} a_n$ 有相同敛散性, 其余情形请读者自行证明.

例 5 讨论级数 $\sum_{n=1}^{\infty} \sin \frac{1}{2^n}$ 的敛散性.

解 因

$$\lim_{n \rightarrow \infty} \frac{\sin \frac{1}{2^n}}{\frac{1}{2^n}} = 1,$$

故 $\sum_{n=1}^{\infty} \sin \frac{1}{2^n}$ 与 $\sum_{n=1}^{\infty} \frac{1}{2^n}$ 有相同的敛散性, 而等比级数 $\sum_{n=1}^{\infty} \frac{1}{2^n}$ 收敛, 故 $\sum_{n=1}^{\infty} \sin \frac{1}{2^n}$ 收敛.

3. 积分判别法

设 $f(x)$ 是区间 $[1, +\infty)$ 上正值连续的减函数, $F(x)$ 为 $f(x)$ 的一个原函数, 且 $\lim_{x \rightarrow +\infty} F(x) = M$. 令 $a_n = f(n)$, 则当 M 为有限数时, 级数 $\sum_{n=1}^{\infty} a_n$ 收敛, 而

当 $M = +\infty$ 时, 级数 $\sum_{n=1}^{\infty} a_n$ 发散.

证 由 $f(x)$ 的单调性知, 当 $n \leq x \leq n+1$ 时,

$$a_{n+1} = f(n+1) \leq f(x) \leq f(n) = a_n.$$

将上式各项从 n 到 $n+1$ 对 x 积分得

$$a_{n+1} = \int_n^{n+1} a_{n+1} dx \leq \int_n^{n+1} f(x) dx \leq \int_n^{n+1} a_n dx = a_n. \quad (4.1.7)$$

级数

$$\sum_{n=1}^{\infty} \int_n^{n+1} f(x) dx, \quad (4.1.8)$$

的前 n 项和为

$$\int_1^{n+1} f(x) dx = F(n+1) - F(1).$$

由条件知 $\lim_{n \rightarrow \infty} F(n+1) = M$, 于是正项级数 (4.1.8) 收敛或发散由 M 是有限或 $+\infty$ 而定. 由 (4.1.7) 式及比较判别法, 当 M 为有限时, 级数 (4.1.8) 收敛,

$\sum_{n=1}^{\infty} a_{n+1}$ 也收敛, 即 $\sum_{n=1}^{\infty} a_n$ 收敛; 当 $M = +\infty$ 时, 级数 (4.1.8) 发散, $\sum_{n=1}^{\infty} a_{n+1}$ 也

发散, 即 $\sum_{n=1}^{\infty} a_n$ 发散, 这就是要证明的.

例 6 称 $\sum_{n=1}^{\infty} \frac{1}{n^p}$ 为 p 级数, 这里 p 为实数. 试讨论 p 级数的敛散性.

解 当 $p \leq 0$ 时, 因 $\lim_{n \rightarrow \infty} \frac{1}{n^p} \neq 0$, 知 p 级数发散.

当 $p > 0$ 时, $\frac{1}{x^p}$ 为区间 $[1, +\infty)$ 上的正值连续减函数, 而 $\frac{1}{x^p}$ 的原函数可取

为

$$F(x) = \begin{cases} \ln x, & p = 1, \\ \frac{1}{1-p} x^{1-p}, & p \neq 1. \end{cases}$$

于是, 当 $0 < p \leq 1$ 时, $\lim_{x \rightarrow +\infty} F(x) = +\infty$, 级数发散; 当 $p > 1$ 时, $\lim_{x \rightarrow +\infty} F(x) = 0$, 级数收敛.

综上所述, p 级数当且仅当 $p > 1$ 时收敛.

p 级数的敛散性很重要. 我们常将某一个级数与 p 级数比较, 从而推知其敛散性. 特别地, 级数

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

称为调和级数, 由例 6 知, 调和级数发散于 $+\infty$, 即

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots = +\infty.$$

例7 判别级数 $\sum_{n=1}^{\infty} \frac{1}{2n-1}$ 的敛散性.

解 因 $\frac{1}{2n-1} > \frac{1}{2n} = \frac{1}{2} \cdot \frac{1}{n}$, 由调和级数发散, 知 $\sum_{n=1}^{\infty} \frac{1}{2n}$ 发散, 再由比较判别法, 知 $\sum_{n=1}^{\infty} \frac{1}{2n-1}$ 发散.

例8 对 α 讨论级数 $\sum_{n=1}^{\infty} \frac{\sqrt{n+1} - \sqrt{n}}{n^\alpha}$ 的敛散性.

解 因

$$a_n = \frac{\sqrt{n+1} - \sqrt{n}}{n^\alpha} = \frac{1}{(\sqrt{n+1} + \sqrt{n}) n^\alpha},$$

与

$$\lim_{n \rightarrow \infty} \frac{1}{(\sqrt{n+1} + \sqrt{n}) n^\alpha} : \frac{1}{n^{\alpha + \frac{1}{2}}} = \frac{1}{2},$$

故原级数与级数 $\sum_{n=1}^{\infty} \frac{1}{n^{\alpha + \frac{1}{2}}}$ 的敛散性相同. 由 p 级数的敛散知道, 当且仅当

$\alpha > \frac{1}{2}$ 时原级数收敛.

4. 柯西 (Cauchy) 根值判别法

设 $\sum_{n=1}^{\infty} a_n$ 为正项级数, 若

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = r \quad (0 \leq r \leq +\infty), \quad (4.1.9)$$

则当 $r < 1$ 时级数收敛, 当 $r > 1$ 时级数发散.

证 设 $r < 1$. 取 q , 使得 $r < q < 1$. 由条件 (4.1.9) 知, 当 n 充分大时, $\sqrt[n]{a_n} < q$, 即 $a_n < q^n$, 因几何级数 $\sum_{n=1}^{\infty} q^n$ 收敛, 所以 $\sum_{n=1}^{\infty} a_n$ 收敛.

设 $r > 1$. 当 n 充分大时必有 $\sqrt[n]{a_n} > 1$, 即 $a_n > 1$, 这与级数收敛的必要条件相违背, 所以级数 $\sum_{n=1}^{\infty} a_n$ 发散.

注 当 $r = 1$ 时, 柯西根值判别法失效. 例如, 对于级数 $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{1}{n^p}$, 得到 $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = 1$. 但是当 $p > 1$ 时级数收敛, 而当 $p \leq 1$ 时级数发散.

例9 判别级数 $\sum_{n=1}^{\infty} \left(\frac{n}{2n+1} \right)^n$ 的敛散性.

解 因 $\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = \lim_{n \rightarrow \infty} \frac{n}{2n+1} = \frac{1}{2}$, 由柯西判别法知此级数收敛.

5. 达朗贝尔(D'Alembert)比值判别法

设 $\sum_{n=1}^{\infty} a_n$ 为正项级数, 若

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = r \quad (0 \leq r \leq +\infty), \quad (4.1.10)$$

则当 $r < 1$ 时级数收敛, 当 $r > 1$ 时级数发散.

证 设 $r < 1$, 取 q 使得 $r < q < 1$, 由条件(4.1.10)知, 当 n 充分大时, 有

$$\frac{a_{n+1}}{a_n} < q.$$

因改变级数的有限项不影响其敛散性, 故可设上式对所有 $n \geq 1$ 成立. 于是,

$$\begin{aligned} a_2 &< a_1 q, \\ a_3 &< a_2 q < a_1 q^2, \\ a_4 &< a_3 q < a_1 q^3, \\ &\dots\dots \\ a_n &< a_{n-1} q < a_1 q^{n-1}, \\ &\dots\dots \end{aligned}$$

由于等比级数 $\sum_{n=1}^{\infty} a_1 q^{n-1}$ 收敛, 所以 $\sum_{n=1}^{\infty} a_n$ 收敛.

设 $r > 1$, 当 n 充分大时, $\frac{a_{n+1}}{a_n} > 1$, 即数列 a_n 当 n 充分大时单调增加, 故

$\lim_{n \rightarrow \infty} a_n \neq 0$, 所以 $\sum_{n=1}^{\infty} a_n$ 发散.

注 当 $r = 1$ 时, 达朗贝尔判别法失效. 例如, 设 $a_n = \frac{1}{n^p}$, 则 $\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = 1$, 而 p 级数可能收敛也可能发散.

例 10 判别级数 $\sum_{n=1}^{\infty} \frac{n!}{n^n}$ 的敛散性.

解 设 $a_n = \frac{n!}{n^n}$, 则当 $n \rightarrow \infty$ 时

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^{n+1}} \bigg/ \frac{n!}{n^n} = \frac{n^n}{(n+1)^n} \rightarrow \frac{1}{e}.$$

因 $\frac{1}{e} < 1$, 由达朗贝尔判别法知原级数收敛.

有时我们可通过证明级数 $\sum_{n=1}^{\infty} a_n$ 收敛来证明极限 $\lim_{n \rightarrow \infty} a_n = 0$. 例如, 由例 8

知 $\sum_{n=1}^{\infty} \frac{n!}{n^n}$ 收敛, 故

$$\lim_{n \rightarrow \infty} \frac{n!}{n^n} = 0.$$

例 11 设 $\{a_n\}$ 为单调减数列, 且 $\sum_{n=1}^{\infty} a_n$ 收敛. 求证 $\lim_{n \rightarrow \infty} na_n = 0$.

证 由假设条件可得出, $\sum_{n=1}^{\infty} a_n$ 为正项级数. 因若不然, 有某一个 $a_k < 0$, 则

当 $n > k$ 时, $a_n \leq a_k < 0$. 于是, $\lim_{n \rightarrow \infty} a_n \neq 0$, 这与 $\sum_{n=1}^{\infty} a_n$ 收敛相违.

设 s_n 为此正项级数 $\sum_{n=1}^{\infty} a_n$ 的前 n 项和. 由 $\{a_n\}$ 单调减少知,

$$s_{2n} - s_n = a_{n+1} + a_{n+2} + \cdots + a_{2n} \geq na_{2n} \geq 0.$$

当 $n \rightarrow \infty$ 时, 上式左边极限为零. 由第一章中极限的夹逼定理知

$$\lim_{n \rightarrow \infty} 2na_{2n} = 2 \lim_{n \rightarrow \infty} na_{2n} = 0. \quad (4.1.11)$$

另一方面

$$s_{2n+1} - s_n = a_{n+1} + a_{n+2} + \cdots + a_{2n+1} \geq (n+1)a_{2n+1} \geq \frac{1}{2}(2n+1)a_{2n+1} \geq 0.$$

同样可知

$$\lim_{n \rightarrow \infty} (2n+1)a_{2n+1} = 0. \quad (4.1.12)$$

由 (4.1.11) 与 (4.1.12) 两式, 知 $\lim_{n \rightarrow \infty} na_n = 0$.

柯西判别法与达朗贝尔判别法都是将所讨论的级数与几何级数进行比较而得到的判别法, 对那些收敛或发散的速度比几何级数慢的级数, 例如 p 级数, 这两个判别法就失效了. 读者可以从参考书中找到更为精细的判别法, 这里我们不再列举了.

4.1.3 任意项级数

现在我们讨论一般的常数项级数, 即它的各项可取任意正、负实数的级数. 首先讨论被称之为“交错级数”的特殊情形.

定义 4.1.4 (交错级数) 若某级数的正负项交替出现, 则称此种级数为交错级数.

例如, 当 $a_n > 0, n \in \mathbb{N}$,

$$\sum_{n=1}^{\infty} (-1)^{n-1} a_n = a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{n-1} a_n + \cdots \quad (4.1.13)$$

就是交错级数.

定理 4.1.2 (莱布尼茨 (G. W. Leibniz) 判别法) 若交错级数 (4.1.13) 满足下述条件:

(i) $a_n \geq a_{n+1} \geq 0, n \in \mathbf{N}$;

(ii) $\lim_{n \rightarrow \infty} a_n = 0$,

则交错级数(4.1.13)收敛.

证 设 s_n 为级数(4.1.13)的前 n 项和,因

$$s_{2n} = (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n}) \geq s_{2n-2}, \quad (4.1.14)$$

$$s_{2n} = a_1 - (a_2 - a_3) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \leq a_1, \quad (4.1.15)$$

故数列 $\{s_{2n}\}$ 为有界增数列,所以 $\{s_{2n}\}$ 收敛. 设

$$\lim_{n \rightarrow \infty} s_{2n} = s. \quad (4.1.16)$$

由于 $s_{2n+1} = s_{2n} + a_{2n+1}$, 依条件(ii)知

$$\lim_{n \rightarrow \infty} s_{2n+1} = \lim_{n \rightarrow \infty} s_{2n} + \lim_{n \rightarrow \infty} a_{2n+1} = s.$$

所以 $\lim_{n \rightarrow \infty} s_n = s$, 即交错级数(4.1.13)收敛.

易知交错级数

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + (-1)^{n-1} \frac{1}{n} + \cdots \quad (4.1.17)$$

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n^2} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \cdots + (-1)^{n-1} \frac{1}{n^2} + \cdots \quad (4.1.18)$$

均满足定理 4.1.2 中的条件,故必收敛.

若对上面两级数的各项取绝对值,分别得到级数 $\sum_{n=1}^{\infty} \frac{1}{n}$ 与 $\sum_{n=1}^{\infty} \frac{1}{n^2}$, 则前者是发散的,后者是收敛的. 为了区别具有这两种性质的级数,我们引进绝对收敛与非绝对收敛的概念.

定义 4.1.5 (绝对收敛与非绝对收敛) 设 $\sum_{n=1}^{\infty} a_n$ 为任意项级数. 若级数 $\sum_{n=1}^{\infty} |a_n|$ 收敛,则称 $\sum_{n=1}^{\infty} a_n$ 为绝对收敛级数. 若级数 $\sum_{n=1}^{\infty} |a_n|$ 发散,而 $\sum_{n=1}^{\infty} a_n$ 收敛,则称 $\sum_{n=1}^{\infty} a_n$ 为非绝对收敛级数,许多参考书上也称非绝对收敛级数为 $\sum_{n=1}^{\infty} a_n$ 条件收敛.

例如(4.1.17)与(4.1.18)分别为非绝对收敛与绝对收敛级数.

定理 4.1.3 (绝对收敛级数的收敛性) 设级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛,则级数

$$\sum_{n=1}^{\infty} a_n \text{ 收敛.}$$

证 令

$$x_n = \frac{1}{2}(|a_n| + a_n), \quad y_n = \frac{1}{2}(|a_n| - a_n).$$

显然 $\sum_{n=1}^{\infty} x_n$ 与 $\sum_{n=1}^{\infty} y_n$ 均为正项级数, 因为

$$x_n \leq |a_n|, \quad y_n \leq |a_n|,$$

由正项级数收敛的比较判别法知, $\sum_{n=1}^{\infty} x_n$ 与 $\sum_{n=1}^{\infty} y_n$ 均收敛. 而由

$$a_n = x_n - y_n,$$

以及收敛级数的差收敛, 得 $\sum_{n=1}^{\infty} a_n$ 收敛.

绝对收敛级数与非绝对收敛级数虽然都是收敛级数, 但它们有本质区别, 以下我们讨论绝对收敛级数所具有的一些性质.

若数列 $\{n_k\}$ 为正整数列的子数列, 则称级数 $\sum_{k=1}^{\infty} a_{n_k}$ 为级数 $\sum_{n=1}^{\infty} a_n$ 的子级数.

定理 4.1.4 (绝对收敛级数的子级数) 设级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛, 则它的任一子级数 $\sum_{k=1}^{\infty} a_{n_k}$ 也绝对收敛.

证 设 s_n 为 $\sum_{n=1}^{\infty} |a_n|$ 的前 n 项和, s'_n 为 $\sum_{k=1}^{\infty} |a_{n_k}|$ 的前 n 项和. 显然有

$$0 \leq s'_n = |a_{n_1}| + |a_{n_2}| + \cdots + |a_{n_k}|$$

$$\leq |a_1| + |a_2| + \cdots + |a_{n_k}| = s_{n_k}.$$

因正项级数 $\sum_{n=1}^{\infty} a_n$ 收敛, 故 $\{s_{n_k}\}$ 有界, 于是 $\{s'_n\}$ 也有界. 从而正项级数

$\sum_{k=1}^{\infty} |a_{n_k}|$ 收敛, 即 $\sum_{k=1}^{\infty} a_{n_k}$ 绝对收敛. 定理得证.

显然我们有下述推论.

推论 绝对收敛级数的正项与负项分别构成两个绝对收敛的子级数.

对于绝对收敛级数来说, 还有下述定理.

定理 4.1.5 (绝对收敛级数可交换次序) 任意改变绝对收敛级数各项的次序, 得到的级数仍然绝对收敛, 并收敛到原来的和.

该定理的证明较长, 这里从略.

定理 4.1.4 与定理 4.1.5 对于非绝对收敛级数是不成立的. 例如非绝对收敛级数的正项(或负项)构成的子级数为发散的级数. 其证明作为习题留给读者来完成. 关于交换非绝对收敛级数各项的次序, 我们有下述有趣的定理.

定理 4.1.6 (非绝对收敛级数的性质) 若级数 $\sum_{n=1}^{\infty} a_n$ 非绝对收敛, 任取实

数 B (或 $\pm\infty$), 都可以适当改变 $\sum_{n=1}^{\infty} a_n$ 各项的次序, 使得到的级数收敛于 B (或发散于 $\pm\infty$).

下面我们讨论两级数的乘法.

对有限多项的两个和式相乘, 我们有

$$\begin{aligned} & (a_1 + a_2 + \cdots + a_m) \cdot (b_1 + b_2 + \cdots + b_n) \\ &= a_1 b_1 + a_2 b_1 + \cdots + a_m b_n. \end{aligned}$$

上述乘法如何推广到两个无穷级数的乘积 $\left(\sum_{n=1}^{\infty} a_n\right) \left(\sum_{n=1}^{\infty} b_n\right)$ 呢? 我们有下述定理.

定理 4.1.7 (绝对收敛级数的柯西乘积) 若 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 是分别收敛于 A 与 B 的绝对收敛级数, 两级数的柯西乘积定义为下述级数

$$\begin{aligned} & a_1 b_1 + (a_1 b_2 + a_2 b_1) + (a_1 b_3 + a_2 b_2 + a_3 b_1) + \cdots \\ &= \sum_{n=1}^{\infty} (a_1 b_n + a_2 b_{n-1} + \cdots + a_n b_1), \end{aligned} \quad (4.1.19)$$

则级数(4.1.19)绝对收敛, 且其和为 AB .

特别指出, 若 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 均为非绝对收敛级数, 则不能保证级数(4.1.19)收敛. 若这两个级数收敛, 且至少一个绝对收敛, 则级数(4.1.19)仍收敛于 AB . 我们略去这一结论的证明.

上文建立的正项级数敛散性判别法对任意项级数一般不再适用了. 但这些判别法均可用来判别任意项级数的绝对收敛性. 例如, 我们有相应于达朗贝尔判别法的定理.

定理 4.1.8 (达朗贝尔判别法) 若极限

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = r,$$

则当 $r < 1$ 时, $\sum_{n=1}^{\infty} a_n$ 绝对收敛, 当 $r > 1$ 时, $\sum_{n=1}^{\infty} a_n$ 发散.

证 由正项级数收敛的达朗贝尔判别法知, 当 $r < 1$ 时, $\sum_{n=1}^{\infty} |a_n|$ 收敛, 故 $\sum_{n=1}^{\infty} a_n$ 绝对收敛. 而当 $r > 1$ 时, 必有 $\lim_{n \rightarrow \infty} a_n \neq 0$, 从而 $\sum_{n=1}^{\infty} a_n$ 发散.

第二章我们学过数列, 数列与级数这两个概念之间有着密切的联系, 任给常数项级数 $\sum_{n=1}^{\infty} a_n$, 它的前 n 项和构成了数列

$$\{s_n\}: s_1, s_2, \cdots, s_n, \cdots \quad (4.1.20)$$

反之,任给数列(4.1.20),若令

$$a_1 = s_1, a_n = s_n - s_{n-1} \quad (n=2, 3, \cdots),$$

便得到级数 $\sum_{n=1}^{\infty} a_n$, 该级数的前 n 项和恰为 s_n , 因此上述对应是级数与数列之间的一一对应关系, 互相对应的级数与数列有相同的敛散性, 并且收敛级数的和等于对应数列的极限. 正因为这样, 我们可以利用级数研究数列, 也可利用数列研究级数, 在第二章中证明的关于数列的许多性质, 都可以按照这种对应关系平移到级数中来, 而不需要重新证明.

在上述级数与数列的对应关系下, 正项级数(前有限项允许为负)对应于增数列. 因此“增数列收敛的充要条件为该数列有界”这一命题用级数语言叙述就成为“正项级数收敛的充要条件是其部分和数列有界”.

习题

1. 利用定义证明下列级数收敛, 并求其和:

$$(1) \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2^n}$$

$$(2) \sum_{n=1}^{\infty} \left(\frac{1}{2^n} + \frac{1}{3^n} \right)$$

$$(3) \sum_{n=1}^{\infty} \frac{1}{(3n-2)(3n+1)}$$

$$(4) \sum_{n=1}^{\infty} (\sqrt{n+2} - 2\sqrt{n+1} + \sqrt{n})$$

2. 判别下列正项级数的敛散性:

$$(1) \sum_{n=1}^{\infty} \frac{1}{2n+1}$$

$$(2) \sum_{n=1}^{\infty} \frac{2n-1}{2n}$$

$$(3) \sum_{n=1}^{\infty} \sqrt[n]{a}, \quad a > 0$$

$$(4) \sum_{n=1}^{\infty} \frac{1}{\sqrt[n]{n+n^2}}$$

$$(5) \sum_{n=1}^{\infty} \frac{n+1}{n(n+2)}$$

$$(6) \sum_{n=1}^{\infty} \tan \frac{1}{n}$$

$$(7) \sum_{n=1}^{\infty} \frac{n}{2^{n+1}}$$

$$(8) \sum_{n=1}^{\infty} \frac{3^n}{n!}$$

$$(9) \sum_{n=1}^{\infty} \frac{n^n}{2^n \cdot n!}$$

$$(10) \sum_{n=1}^{\infty} \frac{n^n}{3^n \cdot n!}$$

$$(11) \sum_{n=1}^{\infty} \frac{\sin \frac{1}{n}}{\ln(n+1)}$$

$$(12) \sum_{n=1}^{\infty} \left(\frac{n}{2n+1} \right)^n$$

$$(13) \sum_{n=1}^{\infty} \frac{n^3}{3^n}$$

$$(14) \sum_{n=2}^{\infty} \frac{1}{\ln(n!)}$$

3. 对 p, q , 讨论级数 $\sum_{n=1}^{\infty} (\ln n)^q \sin^2 \frac{1}{n^p}$ 的敛散性(其中 $p > 0$).

4. 若对任意正整数 n 有 $\frac{a_{n+1}}{a_n} < 1$ 成立, 能否断言正项级数 $\sum_{n=1}^{\infty} a_n$ 收敛? 为什么?

5. 试证明正项级数达朗贝尔判别法的一般形式:若存在常数 $r, 0 < r < 1$, 对充分大的 n 有 $\frac{a_{n+1}}{a_n} < r$ (或 $\frac{a_{n+1}}{a_n} \geq 1$) 成立, 则正项级数 $\sum_{n=1}^{\infty} a_n$ 收敛(或发散).

6. 举例说明级数性质 5 的逆命题不成立, 即若“加括号”后得到的级数收敛, 而原级数不一定收敛.

7. 证明下列极限:

$$(1) \lim_{n \rightarrow \infty} \frac{n^n}{(n!)^2} = 0$$

$$(2) \lim_{n \rightarrow \infty} \frac{e^n}{n!} = 0$$

8. 下列级数是否收敛? 若收敛, 是非绝对收敛还是绝对收敛?

$$(1) \sum_{n=2}^{\infty} (-1)^n \frac{1}{\ln n}$$

$$(2) \sum_{n=1}^{\infty} (-1)^{n-1} \frac{n}{2n+1}$$

$$(3) \sum_{n=1}^{\infty} \frac{1 + (-1)^n}{\sqrt{n}}$$

$$(4) \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n \sqrt{n+1}}$$

$$(5) \sum_{n=1}^{\infty} (-1)^n \arcsin \frac{1}{n}$$

$$(6) \sum_{n=1}^{\infty} (-1)^n \left(1 - \cos \frac{1}{n}\right)$$

$$(7) \sum_{n=2}^{\infty} \sin \left(n\pi + \frac{1}{\ln n}\right)$$

$$(8) \sum_{n=1}^{\infty} \frac{(-1)^n}{1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}}$$

9. 设当 $n \rightarrow \infty$ 时, 数列 $\{a_n\}$ 的通项 a_n 大于零趋于零, 试问交错级数 $\sum_{n=1}^{\infty} (-1)^n a_n$ 是否一定收敛? 考察 $a_n = \frac{1}{\sqrt{n}} - \frac{(-1)^n}{n}$ 的情形.

10. 设 a_n 与 b_n 恒正, 且当 $n \rightarrow \infty$ 时, a_n 与 b_n 为等价无穷小, 交错级数 $\sum_{n=1}^{\infty} (-1)^n a_n$ 与 $\sum_{n=1}^{\infty} (-1)^n b_n$ 是否有相同的敛散性? 试讨论当 $a_n = \frac{1}{\sqrt{n}}, b_n = \frac{1}{\sqrt{n}} - \frac{(-1)^n}{n}$ 的情形.

11. 求证非绝对收敛级数的正项(或负项)构成的子级数发散.

12. 若级数 $\sum_{n=1}^{\infty} |x_{n+1} - x_n|$ 收敛, 求证数列 $\{x_n\}$ 收敛.

13. 设 $\sum_1^{\infty} a_n$ 为非绝对收敛级数, 命 $p_n = \frac{1}{2}(|a_n| + a_n), q_n = \frac{1}{2}(|a_n| - a_n), P_n = \sum_{k=1}^n p_k, Q_n = \sum_{k=1}^n q_k$. 求证 $\lim_{n \rightarrow \infty} P_n/Q_n = 1$.

14. 将级数 $\sum_{n=0}^{\infty} \frac{1}{n!}$ 与 $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!}$ 按对角线法则相乘, 并求其和, 约定 $0! = 1$.

15. 设任意项级数 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} c_n$ 都收敛, 且对一切 n 均有 $a_n \leq b_n \leq c_n$ 成立. 试证 $\sum_{n=1}^{\infty} b_n$ 收敛.

16. 判别下列级数的敛散性, 若收敛, 试判别非绝对收敛还是绝对收敛.

$$(1) \sum_{n=1}^{\infty} \int_{n\pi}^{(n+1)\pi} \frac{\sin x}{x} dx$$

$$(2) \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n} - (-1)^n}$$

§ 4.2 函数项级数

4.2.1 函数项级数

1. 函数项级数的收敛性与一致收敛性

上一节我们研究了常数项级数,现在来研究函数项级数,即级数的各项都是某一个变量的函数.函数项级数一般可表示为下述形式

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \cdots + u_n(x) + \cdots, \quad (4.2.1)$$

式中 $u_n(x)$ 均为定义在某个集合 $X \subset \mathbf{R}$ 上的函数.

若 $x_0 \in X$, 以 $x = x_0$ 代入(4.2.1), 便得到常数项级数

$$\sum_{n=1}^{\infty} u_n(x_0) = u_1(x_0) + \cdots + u_n(x_0) + \cdots. \quad (4.2.2)$$

若级数(4.2.2)收敛, 则称 x_0 为函数项级数(4.2.1)的收敛点; 函数项级数(4.2.1)的收敛点的全体称为其收敛域, 记为 I . 若级数(4.2.2)发散, 则称 x_0 为其发散点. 对收敛域中任一点 $x \in I$, 级数(4.2.1)都有一个确定的和

$$s(x) = \sum_{n=1}^{\infty} u_n(x) \quad (x \in I).$$

这就是说, $s(x)$ 是定义在级数(4.2.1)的收敛域 I 上的函数, 称其为该函数项级数的和函数.

我们把(4.2.1)前 n 项之和记作 $s_n(x)$, 即

$$\lim_{n \rightarrow \infty} s_n(x) = s(x) \quad (x \in I).$$

把无穷序列

$$\{s_n(x)\}: s_1(x), s_2(x), \cdots, s_n(x), \cdots$$

称为函数项序列, 它是函数项级数(4.2.1)的部分和序列. 在(4.2.1)的收敛域 I 上, 该序列的极限为 $s(x)$.

例 1 几何级数

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \cdots + x^n + \cdots \quad (4.2.3)$$

是定义在实数集 $\mathbf{R}$ 上的函数项级数, 因为当且仅当 $|x| < 1$ 时, (4.2.3)收敛, 故其收敛域为 $(-1, 1)$, 和函数为

$$s(x) = \frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \cdots \quad (|x| < 1). \quad (4.2.4)$$

注意, 虽然当 $|x| > 1$ 与 $x = -1$ 时, 函数 $\frac{1}{1-x}$ 仍然有意义, 但此时级数

(4.2.3)为发散的级数,故此时(4.2.4)式不再成立.

例 2 设 $u_n(x) = x^n - x^{n-1}$, 函数项级数

$$\sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} (x^n - x^{n-1}) \quad (4.2.5)$$

在实数集 $\mathbf{R}$ 上有定义,它的前 n 项和为

$$s_n(x) = \sum_{k=1}^n (x^k - x^{k-1}) = (x - 1) + (x^2 - x) + \cdots + (x^n - x^{n-1}) = x^n - 1.$$

显然当 $x \in (-1, 1]$ 时,和函数

$$s(x) = \lim_{n \rightarrow \infty} s_n(x) = \begin{cases} 0, & x = 1, \\ -1, & |x| < 1; \end{cases}$$

而当 $x \leq -1$ 或 $x > 1$ 时,级数(4.2.5)发散,故函数项级数(4.2.5)的收敛域为 $(-1, 1]$.

从例 2 可见,虽然级数(4.2.5)中的每一项均为连续函数,可是它的和函数 $s(x)$ 不再连续了, $x = 1$ 是该和函数的间断点.这就是说,无穷多个连续函数的和不一定连续(读者记得,有限多个连续函数的和仍然连续).

还可以举出这样的例子,函数项级数(4.2.1)的每一项 $u_n(x)$ 均有导数,或在区间 $[a, b]$ 上可积,并且 $\sum_{n=1}^{\infty} u'_n(x)$ 与 $\sum_{n=1}^{\infty} \int_a^b u_n(x) dx$ 均收敛,但

$$\begin{aligned} \frac{d}{dx} \sum_{n=1}^{\infty} u_n(x) &\neq \sum_{n=1}^{\infty} u'_n(x), \\ \int_a^b \sum_{n=1}^{\infty} u_n(x) dx &\neq \sum_{n=1}^{\infty} \int_a^b u_n(x) dx. \end{aligned}$$

有限个连续函数的和仍然是连续函数,有限个函数的导数(或积分)的和等于它们的和的导数(或积分),但是,函数项级数是对无限多个函数求和,即所谓“无限和”,因此对有限多个函数求和所具有的性质,在无限和的情形下一般不再成立.那么在什么条件下,这些性质对无限和仍然成立呢?原来需要函数序列与函数项级数“一致收敛”.

定义 4.2.1 (函数序列的一致收敛性) 设函数项序列

$$\{f_n(x)\}: f_1(x), f_2(x), \cdots, f_n(x), \cdots,$$

定义在区间 I 上, $f(x)$ 为区间 I 上的某一函数.若对任给的 $\epsilon > 0$, 存在只依赖于 ϵ 而与 $x \in I$ 无关的 $N > 0$, 使得当 $n > N$ 时,不等式

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

对 I 中的一切 x 成立,则称函数序列 $\{f_n(x)\}$ 在 I 上一致收敛于极限函数 $f(x)$.

上述定义若用函数项级数的语言,可叙述成下述形式

定义 4.2.2 (函数项级数的一致收敛性) 设定义在区间 I 上的函数项级数

$\sum_{n=1}^{\infty} u_n(x)$ 的前 n 项和为 $s_n(x)$, 又设 $s(x)$ 为 I 上的某一函数. 若对任给的 $\epsilon > 0$, 存在只依赖于 ϵ 而与 $x \in I$ 无关的 $N > 0$, 使当 $n > N$ 时, 不等式

$$|r_n(x)| \equiv |s(x) - s_n(x)| < \epsilon$$

对 I 中的一切 x 成立, 则称函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 I 上一致收敛于和函数 $s(x)$.

2. 函数项级数的一致收敛性判别法

关于函数项级数的一致收敛性, 我们介绍一个常用而又简单的判别法——魏尔斯特拉斯(Weierstrass)判别法.

定理 4.2.1 (魏尔斯特拉斯判别法) 设 I 为某区间, 若存在收敛的常数项级数 $\sum_{n=1}^{\infty} M_n$, 使得

$$|u_n(x)| \leq M_n \quad (4.2.6)$$

对充分大的 n 及所有 $x \in I$ 成立, 则函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在 I 上绝对一致收敛, 亦即级数 $\sum_{n=1}^{\infty} |u_n(x)|$ 一致收敛.

证 因改变级数的有限多项不会改变级数的收敛性与一致收敛性, 故可设 (4.2.6) 式对所有的正整数 n 成立.

由 (4.2.6) 式及比较判别法知, $\sum_{n=1}^{\infty} |u_n(x)|$ 在 I 上收敛, 故 $\sum_{n=1}^{\infty} u_n(x)$ 在 I 上绝对收敛. 设 $\sum_{n=1}^{\infty} u_n(x)$ 的和函数为 $s(x)$, 前 n 项和为 $s_n(x)$.

因正项级数 $\sum_{n=1}^{\infty} M_n$ 收敛, 故对任给的 $\epsilon > 0$, 存在 $N > 0$, 使当 $n > N$ 时,

$$r_n = M_{n+1} + M_{n+2} + \cdots < \epsilon.$$

现设 $n > N$. 对任意正整数 j 及任意 $x \in I$, 有

$$\begin{aligned} |s_{n+j}(x) - s_n(x)| &= |u_{n+1}(x) + u_{n+2}(x) + \cdots + u_{n+j}(x)| \\ &\leq |u_{n+1}(x)| + |u_{n+2}(x)| + \cdots + |u_{n+j}(x)| \\ &\leq M_{n+1} + M_{n+2} + \cdots + M_{n+j} \leq r_n. \end{aligned}$$

令 $j \rightarrow \infty$, 则 $s_{n+j}(x) \rightarrow s(x)$. 由上式得

$$|s(x) - s_n(x)| \leq r_n < \epsilon \quad (x \in I).$$

所以 $\sum_{n=1}^{\infty} u_n(x)$ 在 I 上一致收敛于 $s(x)$.

上述判别法称为优级数判别法,级数 $\sum_{n=1}^{\infty} M_n$ 称为 $\sum_{n=1}^{\infty} u_n(x)$ 的控制级数,或优级数.所得的结果是受控制级数 $\sum_{n=1}^{\infty} u_n(x)$ 绝对一致收敛.请读者区分“绝对一致收敛”与“绝对收敛且一致收敛”.

例 3 求证函数项级数 $\sum_{n=1}^{\infty} \left(x + \frac{1}{n}\right)^n$ 在区间 $[-r, r]$ 上一致收敛,这里 $0 < r < 1$ 为常数.

证 取 q , 使得 $r < q < 1$, 于是当 n 充分大时, 对区间 $[-r, r]$ 上每一个 x 有 $\left|x + \frac{1}{n}\right| < q$, 即

$$\left|\left(x + \frac{1}{n}\right)^n\right| < q^n.$$

因几何级数 $\sum_{n=1}^{\infty} q^n$ 收敛, 由优级数判别法知, $\sum_{n=1}^{\infty} \left(x + \frac{1}{n}\right)^n$ 在 $[-r, r]$ 内一致收敛.

还有两个判别法, 对于判别级数的一致收敛性很有用处.

定理 4.2.2 (阿贝尔(N. H. Abel)判别法) 如果

- (i) 级数 $\sum_{n=1}^{\infty} a_n(x)$ 在区间 I 中收敛,
- (ii) $b_n(x)$ 在 I 上关于 n 单调减少且有界.

则级数 $\sum_{n=1}^{\infty} a_n(x)b_n(x)$ 在区间 I 中一致收敛.

定理 4.2.3 (狄利克雷(P. G. L. Dirichlet)判别法) 如果

- (i) 级数 $\sum_{n=1}^{\infty} a_n(x)$ 的部分和 $S_n(x)$ 对于 $x \in I$ 与 n 一致有界,
- (ii) $b_n(x)$ 在 I 上关于 n 单调减少, 且 $\lim_{n \rightarrow \infty} b_n(x) = 0$ 对 $x \in I$ 一致成立, 则

级数 $\sum_{n=1}^{\infty} a_n(x)b_n(x)$ 在区间 I 中一致收敛.

3. 函数项级数的性质

(1) 四则运算

给定两个函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 与 $\sum_{n=1}^{\infty} v_n(x)$, 并假设它们收敛域分别为 I_u 与 I_v .

线性运算

两函数项级数之和

$$\sum_{n=1}^{\infty} u_n(x) + \sum_{n=1}^{\infty} v_n(x) \quad (x \in I_u \cap I_v)$$

与数乘

$$\alpha \sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \alpha u_n(x) \quad (x \in I_u, \alpha \text{ 为常数})$$

在所指的集合内收敛.

柯西乘积

两函数项级数之柯西乘积定义为

$$\left(\sum_{n=1}^{\infty} u_n(x) \right) \left(\sum_{n=1}^{\infty} v_n(x) \right) = \sum_{n=1}^{\infty} \left(\sum_{j=1}^n u_j(x) v_{n-j}(x) \right) \quad (x \in I_u \cap I_v).$$

若两函数项级数绝对收敛, 则其柯西乘积也绝对收敛.

(2) 级数求和与极限、求导、积分的换序

现在讨论函数项级数和(即无穷多个函数求和)的连续性、求导、求积分的问题, 并解决“无限项求和”与极限、求导、求积分交换次序的问题.

定理 4.2.4 (和函数的连续性) 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 的每一项 $u_n(x)$ 均在某区间 I 上连续, 且该级数在 I 上一致收敛, 则它的和函数 $s(x)$ 在 I 连续.

证 设 $s_n(x)$ 为 $\sum_{n=1}^{\infty} u_n(x)$ 的前 n 项和 x_0 为 I 中的某一点. 由一致收敛定义知, 任给 $\varepsilon > 0$, 对某一个充分大的 n , 有

$$|s(x) - s_n(x)| < \frac{\varepsilon}{3} \quad (\forall x \in I)$$

成立, 显然也有

$$|s(x_0) - s_n(x_0)| < \frac{\varepsilon}{3}.$$

又因 $s_n(x)$ 在 x_0 点连续, 故存在 $\delta > 0$, 当 $|x - x_0| < \delta$ 时有

$$|s_n(x) - s_n(x_0)| < \frac{\varepsilon}{3}.$$

于是

$$\begin{aligned} |s(x) - s(x_0)| &\leq |s(x) - s_n(x)| + |s_n(x) - s_n(x_0)| + |s_n(x_0) - s(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

所以 $s(x)$ 在 x_0 点连续. 由 x_0 的任意性知, $s(x)$ 为 I 上的连续函数.

定理 4.2.4 用函数序列语言可以叙述为下述形式.

定理 4.2.5 (极限函数的连续性) 设函数项序列 $\{f_n(x)\}$ 在区间 I 上一致收敛于极限函数 $f(x)$, 且 $f_n(x)$ 均为 I 上的连续函数, 则 $f(x)$ 也在 I 上连续.

若 $x_0 \in I$, 上述论断也可表示成

$$\lim_{x \rightarrow x_0} \lim_{n \rightarrow \infty} s_n(x) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow x_0} s_n(x).$$

这就是说, 在定理 4.2.5 的条件下, 两种极限运算可以交换次序. 这里“一致收敛性”起了关键性的作用.

例 4 讨论函数 $s(x) = \sum_{n=1}^{\infty} \left(x + \frac{1}{n}\right)^n$ 的定义域与连续性.

解 当 $|x_0| > 1$ 或 $x_0 = 1$ 时, 对于任意正整数 n , 总有

$$\left|x_0 + \frac{1}{n}\right| > 1.$$

于是

$$\lim_{n \rightarrow \infty} \left(x_0 + \frac{1}{n}\right)^n \neq 0.$$

当 $x_0 = -1$ 时

$$\lim_{n \rightarrow \infty} \left(-1 + \frac{1}{n}\right)^n = \lim_{n \rightarrow \infty} (-1)^n \left(1 - \frac{1}{n}\right)^n$$

不存在. 故当 $|x| \geq 1$ 时, 级数 $\sum_{n=1}^{\infty} \left(x + \frac{1}{n}\right)^n$ 发散, $s(x)$ 没有定义.

当 $|x_0| < 1$ 时, 取 r , 使得 $|x_0| < r < 1$. 由例 3 知该级数在 $[-r, r]$ 上一致收敛, 于是和函数在 $x = x_0$ 处有定义且连续. 因为 x_0 是在区间 $(-1, 1)$ 内任取的, 故 $s(x)$ 的定义域为 $(-1, 1)$, 且为 $(-1, 1)$ 中的连续函数.

定理 4.2.6 (函数项级数逐项积分) 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 的每一项均在区间 $[a, b]$ 上连续, 且该级数在 $[a, b]$ 上一致收敛于函数 $s(x)$, 则

$$\int_a^b s(x) dx = \int_a^b \sum_{n=1}^{\infty} u_n(x) dx = \sum_{n=1}^{\infty} \int_a^b u_n(x) dx. \quad (4.2.7)$$

证 由定理 4.2.4 知, $s(x)$ 为 $[a, b]$ 上的连续函数, 故必在 $[a, b]$ 上可积. 由级数一致收敛, 任给 $\varepsilon > 0$, 存在 $N > 0$, 使当 $n > N$ 时

$$|s(x) - s_n(x)| < \frac{\varepsilon}{b-a} \quad (\forall x \in [a, b])$$

成立. 于是当 $n > N$ 时, 有

$$\begin{aligned} \left| \int_a^b s(x) dx - \sum_{k=1}^n \int_a^b u_k(x) dx \right| &= \left| \int_a^b s(x) dx - \int_a^b \sum_{k=1}^n u_k(x) dx \right| \\ &= \left| \int_a^b s(x) dx - \int_a^b s_n(x) dx \right| \\ &\leq \int_a^b |s(x) - s_n(x)| dx < \int_a^b \frac{\varepsilon}{b-a} dx = \varepsilon. \end{aligned}$$

所以,联合上面的结果,得到

$$\sum_{n=1}^{\infty} \int_a^b u_n(x) dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n \int_a^b u_k(x) dx = \int_a^b \lim_{n \rightarrow \infty} \sum_{k=1}^n u_k(x) dx = \int_a^b \sum_{n=1}^{\infty} u_n(x) dx.$$

定理得证.

若(4.2.7)式成立,我们称函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在区间 $[a, b]$ 上可逐项求积分.

定理 4.2.6 也可叙述为函数项序列的形式.

定理 4.2.7 (极限与积分交换次序) 设函数序列 $\{f_n(x)\}$ 在 $[a, b]$ 上一致收敛于函数 $f(x)$, 若 $f_n(x)$ 均为连续函数, 则

$$\int_a^b f(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx. \quad (4.2.8)$$

这就是说,在定理 4.2.7 的条件下,积分与求极限运算可交换次序.

定理 4.2.8 (函数项级数逐项求导) 设函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在某区间 I 上满足条件

- (i) 收敛于和函数 $s(x)$;
- (ii) $u_n(x)$ 均有连续的导数 $u'_n(x)$;
- (iii) $\sum_{n=1}^{\infty} u'_n(x)$ 一致收敛;

则函数 $s(x)$ 在 I 上有连续的导数 $s'(x)$, 且

$$s'(x) = \left(\sum_{n=1}^{\infty} u_n(x) \right)' = \sum_{n=1}^{\infty} u'_n(x). \quad (4.2.9)$$

证 设 $\sum_{n=1}^{\infty} u'_n(x)$ 一致收敛于 $s^*(x)$, 由定理 4.2.4, $s^*(x)$ 在 I 上连续.

任取 $a, x \in I$, 由定理 4.2.6 知,

$$\begin{aligned} \int_a^x s^*(x) dx &= \int_a^x \sum_{n=1}^{\infty} u'_n(x) dx = \sum_{n=1}^{\infty} \int_a^x u'_n(x) dx \\ &= \sum_{n=1}^{\infty} (u_n(x) - u_n(a)) = s(x) - s(a). \end{aligned}$$

上式两边对 x 求导数, 由 $s^*(x)$ 的连续性得

$$s'(x) = s^*(x) = \sum_{n=1}^{\infty} u'_n(x).$$

若(4.2.9)式成立,我们称函数项级数 $\sum_{n=1}^{\infty} u_n(x)$ 在区间 I 上可逐项求导数.

定理 4.2.8 也可叙述成函数项序列的形式.

定理 4.2.9 (极限与求导交换次序) 设函数项序列 $\{f_n(x)\}$ 在某区间 I 上满足条件

- (i) 收敛于极限函数 $f(x)$;
- (ii) $f_n(x)$ 均有连续的导数 $f'_n(x)$;
- (iii) $\{f'_n(x)\}$ 一致收敛;

则极限函数 $f(x)$ 在 I 上有连续导数 $f'(x)$, 且

$$f'(x) = \frac{d}{dx} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} f'_n(x). \quad (4.2.10)$$

这就是说, 在定理 4.2.9 的条件下, 求导数与求极限运算可交换次序.

4.2.2 幂级数、泰勒级数

幂级数是函数项级数中最简单而又最重要的一类级数, 幂级数的一般形式为

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + \cdots + a_n (x - x_0)^n + \cdots, \quad (4.2.11)$$

这里的系数 $a_0, a_1, \cdots, a_n, \cdots$ 以及 x_0 均为实常数. 在上式中若令 $t = x - x_0$, 则得

$$\sum_{n=0}^{\infty} a_n t^n = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n + \cdots.$$

因此, 我们只要讨论具有较简单形式的幂级数

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n + \cdots. \quad (4.2.12)$$

1. 幂级数的收敛区间与收敛半径

我们讨论幂级数 (4.2.12) 的收敛性问题. 显然, 当 $x = 0$ 时, (4.1.12) 收敛于 a_0 , 因此幂级数 (4.2.12) 的收敛域中至少有一点 $x = 0$.

阿贝尔定理对幂级数的收敛域给出有力的判别.

定理 4.2.10 (阿贝尔定理) 对于幂级数 $\sum_{n=0}^{\infty} a_n x^n$,

(1) 若当 $x = \alpha \neq 0$ 时, 幂级数 (4.2.12) 收敛, 则对满足 $|x| < |\alpha|$ 的任何 x , 级数 (4.2.12) 绝对收敛;

(2) 若当 $x = \beta$ 时, 幂级数 (4.2.12) 发散, 则对满足 $|x| > |\beta|$ 的任何 x , 级数 (4.2.12) 发散.

证 由假设, $\sum_{n=0}^{\infty} a_n \alpha^n$ 收敛, 故必有 $a_n \alpha^n \rightarrow 0$. 于是存在正数 $M > 0$, 使得

$$|a_n \alpha^n| \leq M \quad (n \in \mathbf{N}).$$

故有

$$|a_n \alpha^n| = |a_n \alpha^n| \cdot \left| \frac{x}{\alpha} \right|^n \leq M \left| \frac{x}{\alpha} \right|^n.$$

当 $|x| < |\alpha|$ 时, $\left| \frac{x}{\alpha} \right| < 1$, 几何级数 $\sum_{n=0}^{\infty} M \left| \frac{x}{\alpha} \right|^n$ 收敛, 所以级数 $\sum_{n=0}^{\infty} |a^n x^n|$ 收敛, 即幂级数 (4.2.12) 绝对收敛, (1) 得证.

其次, 假设有一点 $x_1, |x_1| > |\beta|$, 当 $x = x_1$ 时, (4.2.12) 收敛, 则由上述证明知当 $x = \beta$ 时, (4.2.12) 必收敛, 与假设矛盾, (2) 得证, 从而定理得证.

现在利用阿贝尔定理来观察幂级数 (4.2.12) 的收敛域 X .

显然 $0 \in X$. 若 $X \neq (-\infty, +\infty)$, 由定理 2.4.10 知, X 为有界集合. 设 $R \geq 0$ 为 X 的上确界.

当 $R = 0$ 时, 级数 (4.2.12) 仅当 $x = 0$ 时收敛; 当 $R > 0$ 时, 由定理 4.2.10 及上确界定义知, 在 $|x| > R$ 时级数 (4.2.12) 发散; $|x| < R$ 时, 级数 (4.2.12) 收敛.

我们称 R 为幂级数 (4.2.12) 的收敛半径. 当 $X = (-\infty, +\infty)$ 时, 约定收敛半径 $R = +\infty$. 因此对给定的幂级数 (4.2.12), 必有确定的收敛半径 R , R 或者为 $+\infty$, 或者为 0, 或者为正数.

下述定理是求幂级数收敛半径常用的方法.

定理 4.2.11 (幂级数的收敛半径) 若极限

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = r$$

存在, 为非负有限实数或 $+\infty$, 则幂级数 (4.2.12) 的收敛半径 $R = \frac{1}{r}$.

证 利用定理 4.1.8 (任意项级数达朗贝尔判别法), 把证明分三种情形讨论.

(1) 设 $r = 0$, 当 $x \neq 0$ 时,

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = |x| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = r = 0.$$

故幂级数 (4.2.12) 对任何 $x \in (-\infty, +\infty)$ 时收敛, 即收敛半径 $R = +\infty = \frac{1}{r}$.

(2) 设 $r = +\infty$, 当 $x \neq 0$ 时,

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = |x| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = r = +\infty.$$

故幂级数 (4.1.12) 对任何 $x \neq 0$ 发散, 即 $R = 0 = \frac{1}{r}$.

(3) 设 $0 < r < +\infty$, 当 $x \neq 0$ 时,

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = |x| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = |x| r.$$

于是, 当 $|x| < \frac{1}{r}$ 时, (4.2.12) 收敛, 当 $|x| > \frac{1}{r}$ 时, (4.2.12) 发散, 所以收敛半径

$R = \frac{1}{r}$. 定理得证.

上述定理并没有说明当极限 $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$ 不存在时收敛半径的计算的方法, 这超出我们的要求, 这里不再介绍了.

当收敛半径 R 为正实数时, 定理 4.2.11 没有给出 $x = \pm R$ 时幂级数的敛散性. 因此对于收敛区间端点 $x = \pm R$, 需要采用其他方法去判别.

除了收敛半径 $R = 0$ 的情形外, 幂级数的收敛域为一个区间. 因此常常把幂级数的收敛域称为其收敛区间.

例 5 确定幂级数 $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ 的收敛区间.

解 由定理 4.2.11, 立即求得 $r = 0$, 故收敛半径 $R = +\infty$, 该级数的收敛区间为 $(-\infty, +\infty)$.

例 6 确定幂级数 $\sum_{n=1}^{\infty} n^n x^n$ 的收敛区间.

解 可直接求比值 $\frac{a_n}{a_{n+1}}$ 的极限

$$\frac{1}{r} = \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n \rightarrow \infty} \frac{n^n}{(n+1)^{n+1}} = \lim_{n \rightarrow \infty} \frac{1}{\left(1 + \frac{1}{n}\right)^n} \cdot \frac{1}{n+1} = 0,$$

故收敛半径 $R = 0$, 该级数仅当 $x = 0$ 时收敛.

例 7 确定幂级数 $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}$ 的收敛区间.

解 因

$$\lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{\frac{1}{n+1}} = 1,$$

故收敛半径 $R = 1$.

当 $x = 1$ 时, 原级数成为

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1} \frac{1}{n} + \cdots,$$

该级数适合莱布尼茨判别法条件 (见定理 4.1.2), 故收敛.

当 $x = -1$ 时, 原级数成为

$$-1 - \frac{1}{2} - \frac{1}{3} - \cdots - \frac{1}{n} - \cdots,$$

该级数显然发散. 所以所求的收敛区间为 $(-1, 1]$.

例 8 求幂级数 $\sum_{n=1}^{\infty} 2^n (x-1)^{2n+1}$ 的收敛区间.

解 令 $t = (x-1)^2$, 提取 $(x-1)$ 至求和号外, 得

$$\sum_{n=1}^{\infty} 2^n (x-1)^{2n+1} = (x-1) \sum_{n=1}^{\infty} 2^n t^n.$$

因

$$\lim_{n \rightarrow \infty} \frac{2^n}{2^{n+1}} = \frac{1}{2},$$

故级数 $\sum_{n=1}^{\infty} 2^n t^n$ 的收敛半径为 $\frac{1}{2}$, 而当 $|t| = \frac{1}{2}$ 时, 该级数发散, 因此它的收敛区间为 $\left(-\frac{1}{2}, \frac{1}{2}\right)$. 于是原级数当且仅当 $|(x-1)^2| < \frac{1}{2}$ 时收敛, 故所求的收敛区间为

$$\left(1 - \frac{\sqrt{2}}{2}, 1 + \frac{\sqrt{2}}{2}\right).$$

2. 幂级数的运算

关于幂级数的运算, 我们先讨论四则运算. 设有两个幂级数

$$f(x) = \sum_{n=1}^{\infty} a_n x^n = a_0 + a_1 x + \cdots + a_n x^n + \cdots, \quad (4.2.13)$$

$$g(x) = \sum_{n=1}^{\infty} b_n x^n = b_0 + b_1 x + \cdots + b_n x^n + \cdots, \quad (4.2.14)$$

其收敛半径分别为 R_1 与 R_2 , $0 \leq R_1 \leq R_2$, 收敛区间分别为 I_1 与 I_2 .

(1) 幂级数的加法与减法

把幂级数

$$\sum_{n=1}^{\infty} (a_n + b_n) x^n = (a_0 + b_0) + (a_1 + b_1) x + \cdots + (a_n + b_n) x^n + \cdots, \quad (4.2.15)$$

$$\sum_{n=1}^{\infty} (a_n - b_n) x^n = (a_0 - b_0) + (a_1 - b_1) x + \cdots + (a_n - b_n) x^n + \cdots \quad (4.2.16)$$

分别称为幂级数(4.2.13)与(4.2.14)的和与差. 和与差的收敛半径可由所得级数(4.2.15)与(4.2.16)的表示式来确定.

例 9 求幂级数 $\sum_{n=0}^{\infty} x^{n+1}$ 与 $\sum_{n=1}^{\infty} x^n$ 的差, 并讨论此差的收敛域.

解 按幂级数的减法运算规则,将 x 同次幂的项归并在一起,先把第二个幂级数改写成 $1 + \sum_{n=0}^{\infty} x^{n+1}$, 则有

$$\sum_{n=0}^{\infty} x^{n+1} - \left(1 + \sum_{n=0}^{\infty} x^{n+1}\right) = -1.$$

上式右边 -1 也可看成 x 的幂级数,其中关于 x 正整数幂的各项系数均为 0 ,显然其收敛域为 $(-\infty, +\infty)$.

注 若把例 9 中的两个幂级数看成一般的函数项级数,按它们的项对应相减,它们的差可写成

$$\sum_{n=1}^{\infty} (x^{n+1} - x^n).$$

请读者自行确定此级数的收敛域,并与例 9 的答案作一比较.

(2) 幂级数的乘法

由定理 4.1.7 关于绝对收敛级数的乘法性质知,两个幂级数的乘积

$$\begin{aligned} f(x)g(x) &= \sum_{n=0}^{\infty} \sum_{k=0}^n a_k b_{n-k} x^n \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \cdots + (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0)x^n + \\ &\quad \cdots, \end{aligned} \quad (4.2.17)$$

也是幂级数,在某区间 I 内收敛.

(3) 幂级数的除法

除法是作为乘法的逆运算定义的.

设 $b_0 \neq 0$, 令

$$\frac{f(x)}{g(x)} = c_0 + c_1 x + \cdots + c_n x^n + \cdots. \quad (4.2.18)$$

为计算商的系数 $c_0, c_1, \cdots, c_n, \cdots$, 将上式两边乘以 $g(x)$ 得到

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n \cdot \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} \sum_{k=0}^n b_k c_{n-k} x^n.$$

令上式两边对应系数相等,得到方程组

$$\begin{aligned} a_0 &= b_0 c_0 \\ a_1 &= b_0 c_1 + b_1 c_0, \\ &\cdots \cdots \\ a_n &= b_0 c_n + b_1 c_{n-1} + \cdots + b_n c_0. \end{aligned}$$

由第一个方程得 $c_0 = \frac{a_0}{b_0}$, 代入第二个方程解得

$$c_1 = \frac{1}{b_0} \left(a_1 - \frac{a_0 b_1}{b_0} \right),$$

逐次代入,第 n 次后可得未知系数 c_n .

幂级数商(4.2.18)的收敛半径可能比 R_1 与 R_2 都小,难以给出一般的规律.我们只能对具体问题进行讨论.

3. 幂级数的性质

下面讨论幂级数的一致收敛性、和函数的连续性以及逐项积分与逐项求导数等问题.

定理 4.2.12 (幂级数的一致收敛性) 设幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径为 $R \neq 0$, 则对任意 $r, 0 < r < R$, 该级数在闭区间 $[-r, r]$ 上一致收敛.

证 由定理 4.2.9 知, $\sum_{n=0}^{\infty} a_n r^n$ 绝对收敛. 当 $|x| \leq r$ 时, 有不等式

$$|a_n x^n| \leq |a_n| r^n \quad (n = 1, 2, \dots).$$

根据优级数判别法知 $\sum_{n=0}^{\infty} a_n x^n$ 在 $[-r, r]$ 上一致收敛. 定理得证.

若幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径为 $R > 0$, 一般得不到该级数在 $(-R, R)$ 上的一致收敛的结论. 但我们有如下定理(在以下定理中, 用 I 表示幂级数的收敛域, 它可以是开区间、闭区间或半开半闭区间).

定理 4.2.13 (幂级数的内闭一致收敛性) 若幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛域为 I , $[a, b]$ 为包含在 I 内的任一闭区间, 则该幂级数在 $[a, b]$ 上一致收敛.

这个定理的证明是明显的.

由关于一致收敛级数和函数的连续性定理 4.2.4 与定理 4.2.13 立即可得:

定理 4.2.14 (幂级数的和的连续性) 若幂级数的收敛域为 I , 则该幂级数的和函数为 I 上的连续函数.

关于幂级数逐项积分问题是这样的:

设幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛域为 I , 任取 $x \in I$, 将积分

$$\int_0^x \sum_{n=0}^{\infty} a_n t^n dt \quad (4.2.19)$$

中求和运算与积分运算交换次序, 得到

$$\sum_{n=0}^{\infty} \int_0^x a_n t^n dt = a_0 x + \frac{a_1}{2} x^2 + \dots + \frac{a_{n-1}}{n} x^n + \dots \quad (4.2.20)$$

(4.2.19) 与 (4.2.20) 能否相等呢? 我们有下述定理:

定理 4.2.15 (幂级数的逐项可积性) 设幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛域为 I , 则

当 $x \in I$ 时,有

$$\int_0^x \sum_{n=0}^{\infty} a_n t^n dt = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n = a_0 x + \frac{a_1}{2} x^2 + \cdots + \frac{a_{n-1}}{n} x^n + \cdots, \quad (4.2.21)$$

且幂级数 $\sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n$ 与原级数 $\sum_{n=0}^{\infty} a_n x^n$ 有相同的收敛半径.

证 当 $x=0$ 时, (4.2.21) 式显然成立. 以下讨论 $x \neq 0$ 的情形.

设原级数 $\sum_{n=0}^{\infty} a_n x^n$ 与新级数 (4.2.20) 的收敛半径分别为 R 与 R' . 当 $x \in I$ 时, 由定理 4.2.13 知, 原级数在区间 $[0, x]$ (或 $[x, 0]$) 上一致收敛. 由定理 4.2.4 知, 原级数在该区间上可逐项积分, 于是 (4.2.21) 式成立, 且级数 (4.2.20) 在区间 I 上收敛, 因此 $R \leq R'$.

下面证明 $R > R'$ 对于 $0 < |x_0| < R'$, 我们来证明 $\sum_{n=0}^{\infty} a_n x_0^n$ 收敛. 取常数 $r > 0$, 使得 $|x_0| < r < R'$. 因

$$\lim_{n \rightarrow \infty} \left| \frac{x_0}{r} \right|^n = 0,$$

故当 n 充分大时, 有

$$n \left| \frac{x_0}{r} \right|^n < 1,$$

由此得到

$$|a_{n+1} x_0^{n+1}| \leq \left| \frac{a_{n+1}}{n+1} r^{n+1} \right|.$$

因为当 $x=r$ 时级数 $\sum_{n=0}^{\infty} \frac{a_{n+1}}{n+1} r^{n+1}$ 绝对收敛, 由上式知级数

$$\sum_{n=0}^{\infty} a_{n+1} x_0^{n+1} = x_0 \sum_{n=0}^{\infty} a_{n+1} x_0^n$$

绝对收敛, 因此 $\sum_{n=0}^{\infty} a_n x_0^n$ 收敛, 从而 $R \geq R'$, 于是必有 $R = R'$.

若在 (4.2.21) 式中将 a_{n-1} 换成 na_n ($n=1, 2, \cdots$), 然后两边对 x 求导数, 便得到下述定理:

定理 4.2.16 (幂级数的逐项可导性) 设幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛域为 I , 则在 I 上有

$$\frac{d}{dx} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1},$$

且 $\sum_{n=0}^{\infty} a_n x^n$ 与 $\sum_{n=1}^{\infty} n a_n x^{n-1}$ 有相同的收敛半径.

上述两个定理常叙述成“幂级数在其收敛域内可逐项积分与逐项求导数, 且

不改变其收敛半径”.细心的读者不难从定理 4.2.15 与 4.2.16 的证明中获知, 逐项积分不会使收敛域变小, 逐项求导数也不会使收敛域变大, 最多只会改变收敛域端点的敛散性.

例 10 求幂级数

$$f(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \quad (4.2.22)$$

的和, 并讨论其定义域.

解 将(4.2.22)式逐项求导数, 得

$$f'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1}.$$

上式右边是以 $-x$ 为公比的几何级数, 首项为 1, 故

$$f'(x) = \frac{1}{1+x}, \quad |x| < 1.$$

因 $f'(0) = 0$, 于是

$$f(x) = f(0) + \int_0^x \frac{1}{1+x} dx = \ln(1+x). \quad (4.2.23)$$

上式是在 $|x| < 1$ 条件下得到的, 由例 7 知级数(4.2.22)的收敛区间为 $(-1, 1)$. 由定理 4.2.14(幂级数和函数的连续性)知, (4.2.23)式当 $x = 1$ 成立, 于是(4.2.23)对 $x \in (-1, 1]$ 成立. 因此我们有关于 $\ln(1+x)$ 的展开式

$$\ln(1+x) = x - \frac{x^2}{2} + \cdots + (-1)^{n-1} \frac{x^n}{n} + \cdots, \quad -1 < x \leq 1. \quad (4.2.24)$$

在上式中若令 $x = 1$, 便得著名的公式

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1} \frac{1}{n} + \cdots. \quad (4.2.25)$$

例 11 求幂级数 $\sum_{n=0}^{\infty} (n+1)^2 x^n$ 的收敛区间, 并求和.

解 $R = \lim_{n \rightarrow \infty} \frac{(n+1)^2}{(n+2)^2} = 1$, 而当 $x = \pm 1$ 时, 该级数的通项不以零为极限, 故必发散. 于是收敛区间为 $(-1, 1)$.

令 $f(x) = \sum_{n=0}^{\infty} (n+1)^2 x^n$, 逐项求积分得

$$\int_0^x f(x) dx = \sum_{n=0}^{\infty} \int_0^x (n+1)^2 x^n dx = \sum_{n=0}^{\infty} (n+1) x^{n+1} = x \sum_{n=0}^{\infty} (n+1) x^n.$$

又因

$$\frac{d}{dx} \left(\frac{1}{1-x} \right) = \sum_{n=0}^{\infty} nx^{n-1},$$

于是

$$\int_0^x f(x) dx = x \frac{d}{dx} \left(\frac{1}{1-x} \right) = \frac{x}{(1-x)^2}.$$

将上式两边求导数, 即得

$$f(x) = \frac{1+x}{(1-x)^3} \quad (|x| < 1).$$

例 12 求证

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1} + \cdots. \quad (4.2.26)$$

证 上式左边为交错级数, 满足莱布尼茨判别法的条件, 故此级数收敛. 为求其和, 令

$$f(x) = x - \frac{x^3}{3} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots.$$

易判别上述幂级数的收敛区间为 $[-1, 1]$. 将其逐项求导数, 得

$$f'(x) = 1 - x^2 + x^4 - \cdots + (-1)^n x^{2n} + \cdots = \frac{1}{1+x^2}.$$

因 $f(0) = 0$, 故

$$f(x) = \int_0^x \frac{1}{1+x^2} dx = \arctan x,$$

即

$$\arctan x = x - \frac{x^3}{3} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots \quad (|x| \leq 1).$$

以 $x=1$ 代入上式, 得欲证之式 (4.2.26).

4. 函数的幂级数展式

我们已经讨论了如何确定幂级数的收敛区间与求和函数的方法, 现在讨论与之相反的问题, 即能否将已知函数表示成某一区间上的幂级数? 若可以表示, 幂级数应具有怎样的形式?

先证明下述定理.

定理 4.2.17 (可展开为幂级数的一个必要条件) 若函数 $f(x)$ 在 $x = x_0$ 点的某邻域 U 中可用幂级数 $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ 表示, 则 $f(x)$ 在 U 中存在任意阶导数, 且

$$a_0 = f(x_0), \quad a_n = \frac{f^{(n)}(x_0)}{n!} \quad (n = 1, 2, \cdots).$$

证 由定理 4.2.16 知, $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ 在区间 $I = (x_0 - R, x_0 + R)$ 内存在一阶至任意阶导数, 这里 R 为收敛半径, $U \subset I$. 因此 $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ 必在 U 中存在任意阶导数. 由级数逐项求导法则, 得

$$f'(x) = a_1 + 2a_2(x - x_0) + \cdots + na_n(x - x_0)^{n-1} + \cdots,$$

$$f''(x) = 2! a_2 + 3 \cdot 2a_3(x - x_0) + \cdots + n(n-1)a_n(x - x_0)^{n-2} + \cdots,$$

.....

$$f^{(n)}(x) = n! a_n + \frac{(n+1)!}{1!} a_{n+1}(x - x_0) + \frac{(n+2)!}{2!} a_{n+2}(x - x_0)^2 + \cdots,$$

.....

以 $x = x_0$ 代入上列各式, 便得

$$a_0 = f(x_0), \quad a_1 = \frac{f'(x_0)}{1!}, \cdots, a_n = \frac{f^{(n)}(x_0)}{n!}, \cdots.$$

定义 4.2.3 (泰勒级数) 若函数 $f(x)$ 在 x_0 的某邻域内可表示成幂级数, 则在该邻域中必有

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \cdots. \quad (4.2.27)$$

(4.2.27)式右边称为 $f(x)$ 在 $x = x_0$ 点的泰勒级数, 或称 $x - x_0$ 的幂级数展式.

由定理 4.2.17 可知, $f(x)$ 在 x_0 点的泰勒级数是惟一的, 但在不同点的泰勒级数有不同的表示式.

在(4.2.27)式中, 若 $x_0 = 0$, 则

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots. \quad (4.2.28)$$

(4.2.28)式右边称为 $f(x)$ 的麦克劳林级数. 它由函数 $f(x)$ 惟一确定.

若在(4.2.27)式中令 $t = x - x_0$, $g(t) = f(t + x_0)$, 则 $f(x)$ 的泰勒级数化为 $g(t)$ 的麦克劳林级数. 下文中为简便起见, 一般讨论函数的麦克劳林级数.

下式称为麦克劳林公式

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + r_n(x). \quad (4.2.29)$$

注意到, (4.2.28)与(4.2.29)的差别: (4.2.28)式右端是幂级数, 有无穷多项; (4.2.29)式只有有限多项, 我们称 $r_n(x)$ 为幂级数的余项. 若在 x_0 的某邻域 U 内, $f(x)$ 存在任意阶导数, 则(4.2.29)式对任意正整数 n 在 U 中成立, 但却不能保证(4.2.28)式在 U 中成立. 关于这一点, 可以研究下述函数

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

该函数在 $(-\infty, +\infty)$ 上存在任意阶导数, 且在原点各阶导数 $f^{(n)}(0) = 0$. 因此对任意正整数 n , 在区间 $(-\infty, +\infty)$ 上该函数可以按麦克劳林公式(4.2.29)展开, 它的展开式除了余项 $r_n(x)$ 外, 其余各项均为零. 但该函数的麦克劳林级数

(4.2.28) 仅仅当 $x=0$ 时成立, 此时 (4.2.28) 式右端各项全为零.

上述例子告诉我们, 若只假定 $f(x)$ 在原点的某邻域中存在任意阶导数还不能保证在该邻域中 (4.2.28) 式成立.

定理 4.2.18 (可展开为麦克劳林级数的充要条件) 设函数 $f(x)$ 在 $x=0$ 的某邻域 U 中存在任意阶导数, 则 $f(x)$ 在 U 中可展成麦克劳林级数 (4.2.28) 的充要条件是

$$\lim_{n \rightarrow \infty} r_n(x) = 0 \quad (x \in U),$$

式中 $r_n(x)$ 为 $f(x)$ 的麦克劳林公式 (4.2.29) 中的余项.

证 因 $f(x)$ 在 U 中存在任意阶导数, 对任意正整数 n 有

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + r_n(x) \quad (x \in U).$$

令 $n \rightarrow \infty$, 则

$$f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + \lim_{n \rightarrow \infty} r_n(x).$$

所以, 在 U 中, 等式

$$f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k$$

成立的充要条件是

$$\lim_{n \rightarrow \infty} r_n(x) = 0.$$

定理 4.2.18 中, 若 U 为闭区间或半开区间, 显然结论仍然成立. 因此该定理给出了使 (4.2.28) 式成立的 x 的取值范围. 由于判别余项 $r_n(x)$ 的极限是否为零常常很困难, 下述定理虽然仅仅是 $f(x)$ 可展成 x 的幂级数的充分条件, 但有时却十分有效.

定理 4.2.19 (可展开为麦克劳林级数的一个充分条件) 若 $f(x)$ 的各阶导数 $f^{(n)}(x)$ ($n=1, 2, \dots$), 在 $x=0$ 的某邻域 U 中一致有界, 则 $f(x)$ 在 U 中可展成 x 的麦克劳林级数.

证 在 (4.2.29) 式中取拉格朗日余项

$$r_n(x) = \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1} \quad (0 < \theta < 1).$$

由定理的假设条件, 得

$$|f^{(n)}(x)| \leq K \quad (n \in \mathbb{N}, x \in U),$$

则有

$$|r^n(x)| = \left| f^{(n+1)}(\theta x) \frac{x^{n+1}}{(n+1)!} \right| \leq K \frac{|x|^{n+1}}{(n+1)!}.$$

因为对任意 x , 级数 $\sum_{n=1}^{\infty} \frac{|x|^{n+1}}{(n+1)!}$ 收敛, 故

$$\lim_{n \rightarrow \infty} \frac{|x|^{n+1}}{(n+1)!} = 0.$$

于是当 $x \in U$ 时, $\lim_{n \rightarrow \infty} r_n(x) = 0$, 由定理 4.2.18 即得本定理.

下面我们讨论初等函数的幂级数展式.

例 13 求 $f(x) = e^x$ 在 $x=0$ 点的幂级数展式.

解 $f(x)$ 的任意阶导数均为 e^x , 以 $x=0$ 代入得 $f^{(n)}(0) = 1, n=0, 1, 2, \dots$.
于是

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \quad (4.2.30)$$

我们证明上式对任意实数 x 成立. 任取正数 M , 当 $|x| \leq M$ 时,

$$|f^{(n)}(x)| = e^x \leq e^M.$$

由定理 4.2.19 知, 当 $|x| \leq M$ 时, (4.2.30) 式成立. 再由 M 的任意性得到, 当 $x \in (-\infty, +\infty)$ 时 (4.2.30) 成立.

例 14 求 $f(x) = \sin x$ 在 $x=0$ 的幂级数展式.

解 因为 $f^{(n)}(x) = \sin\left(x + \frac{n\pi}{2}\right)$ ($n=0, 1, 2, \dots$), 以 $x=0$ 代入得 $f^{(n)}(0) = \sin \frac{n\pi}{2}$. 于是有

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots \quad (4.2.31)$$

因 $|f^{(n)}(x)| = \left| \sin\left(x + \frac{n\pi}{2}\right) \right| \leq 1$, 故上式对 $x \in (-\infty, +\infty)$ 成立.

上述两个例子都是利用定理 4.2.17 直接计算展开式的系数, 然后再讨论展开式成立的范围. 在很多情况下我们还可以利用已知函数的幂级数展式, 通过幂级数的四则运算、逐项微分与逐项积分等方法求出另一些函数的幂级数展式. 前一种方法称为直接展开法, 后一种方法称为间接展开法. 间接展开法有时十分方便, 尤其可以避免直接讨论余项的极限. 除了 e^x 与 $\sin x$ 外, 我们已经利用级数逐项求积分得到 $\ln(1+x)$ 与 $\arctan x$ 的幂级数展式, 见例 10 与例 12.

例 15 求 $f(x) = \cos x$ 关于 x 的幂级数展式.

解 将 (4.2.31) 式两边对 x 求导数即得

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots \quad (4.2.32)$$

由定理 4.2.16 知上式对 $x \in (-\infty, +\infty)$ 成立.

例 16 将 $f(x) = \frac{x-3}{2x^2+3x}$ 展成 $x+1$ 的幂级数.

解 因 $\frac{x-3}{2x^2+3x} = \frac{-1}{x} + \frac{3}{2x+3}$, 由几何级数求和公式得

$$\frac{-1}{x} = \frac{1}{1-(x+1)} = 1 + (x+1) + \cdots + (x+1)^n + \cdots \quad (|x+1| < 1),$$

$$\frac{1}{2x+3} = \frac{1}{1+2(x+1)} = 1 - 2(x+1) + \cdots + (-2)^n (x+1)^n + \cdots$$

$$\left(|x+1| < \frac{1}{2}\right).$$

所以

$$\frac{x-3}{2x^2+3x} = \sum_{n=0}^{\infty} (1+(-2)^n 3)(x+1)^n \quad \left(|x+1| < \frac{1}{2}\right).$$

例 17 将 $\frac{\ln x}{x}$ 展成 $x-1$ 的幂级数.

解 令 $y = x-1$, 得 $\frac{\ln x}{x} = \frac{1}{1+y} \ln(1+y)$. 已知

$$\frac{1}{1+y} = \sum_{n=0}^{\infty} (-1)^n y^n, \quad \ln(1+y) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} y^n}{n}.$$

上述两个幂级数当 $|y| < 1$ 时都是绝对收敛的. 在幂级数乘法公式

$$\sum_{n=0}^{\infty} a_n y^n \cdot \sum_{n=1}^{\infty} b_n y^n = \sum_{n=1}^{\infty} \sum_{k=1}^n a_{k-1} b_{n-k+1} y^n$$

中令 $a_n = (-1)^n$, $b_n = (-1)^{n-1} \frac{1}{n}$, 因

$$\sum_{k=1}^n a_{k-1} b_{n-k+1} = (-1)^{n-1} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right),$$

于是

$$\begin{aligned} \frac{1}{1+y} \ln(1+y) &= \sum_{n=0}^{\infty} (-1)^n y^n \cdot \sum_{n=1}^{\infty} \frac{(-1)^{n-1} y^n}{n} \\ &= \sum_{n=1}^{\infty} (-1)^{n-1} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) y^n. \end{aligned}$$

上式右边幂级数当 $y = \pm 1$ 时, 通项不趋于 0, 故仅当 $|y| < 1$ 时收敛. 以 $y = x-1$ 代入即得

$$\frac{\ln x}{x} = \sum_{n=1}^{\infty} (-1)^{n-1} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) (x-1)^n \quad (|x-1| < 1).$$

关于幂函数 $(1+x)^\alpha$ ($\alpha \neq 0$) 的麦克劳林级数, 我们有下面的公式

$$(1+x)^\alpha = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n \quad (|x| < 1). \quad (4.2.33)$$

由定理 4.2.17 容易验证, 若 $(1+x)^\alpha$ 在 $x=0$ 的某邻域中可展成 x 的幂级数, 该级数必有 (4.2.33) 的形式. 困难的是讨论它的收敛域. 可以证明 (从略), 对任意数 α , 当 $|x| < 1$ 时, (4.2.33) 式总是成立的. 当 $\alpha = n$ 为正整数时, (4.2.33) 式右边只有有限多项不等于零, 这就是我们熟悉的牛顿二项式公式

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + nx^{n-1} + x^n. \quad (4.2.34)$$

上式对任意实数 x 成立. 若 α 不是正整数, 进一步讨论可以证明, 当 $\alpha > 0$ 时, (4.2.33) 式的收敛域为 $[-1, 1]$, 当 $-1 < \alpha < 0$ 时, 收敛域为 $(-1, 1]$.

我们列出常用初等函数的麦克劳林级数如下.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (-\infty < x < +\infty);$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad (-\infty < x < +\infty);$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad (-\infty < x < +\infty);$$

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \quad (-1 < x \leq 1);$$

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad (-1 \leq x \leq 1);$$

$$(1+x)^\alpha = 1 + \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n \quad (-1 < x < 1).$$

关于 $x = \pm 1$ 的情形, 我们有如下表格:

$x = 1$	$\alpha > 0$	绝对收敛
	$0 > \alpha > -1$	非绝对收敛
	$\alpha \leq -1$	发散
$x = -1$	$\alpha > 0$	绝对收敛
	$\alpha < 0$	发散

以上初等函数的幂级数展式可以推广到复变量的情形(见 §4.2.4).

幂级数的应用非常广泛, 例如应用在微分方程、数值计算、数学物理方法, 以及物理学、天文学、地球科学等各个近代科学领域中. 下面是在近似计算中应用的例子.

例 18 计算 π 的近似值, 要求精确到 10^{-6} .

解 如果我们利用例 12 中推导的公式(4.2.26)

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1} + \cdots$$

计算 π 值, 至少要取前一百万项才能精确到 10^{-6} , 这样的计算量太大了. 这是因为该级数收敛速度太慢的缘故. 我们来推导一个收敛得较快的级数. 令 $\alpha =$

$\arctan \frac{1}{5}$, 于是 $\tan \alpha = \frac{1}{5}$, 则

$$\tan 2\alpha = \frac{2\tan \alpha}{1 - \tan^2 \alpha} = \frac{5}{12},$$

$$\tan 4\alpha = \frac{2\tan 2\alpha}{1 - \tan^2 2\alpha} = \frac{120}{119} \approx 1, \quad 4\alpha \approx \frac{\pi}{4}.$$

$$\text{令 } \beta = 4\alpha - \frac{\pi}{4}, \text{ 则 } \tan \beta = \frac{\left(\frac{120}{119} - 1\right)}{\left(1 + \frac{120}{119}\right)} = \frac{1}{239}. \text{ 于是}$$

$$\begin{aligned} \frac{\pi}{4} &= 4\alpha - \beta = 4\arctan \frac{1}{5} - \arctan \frac{1}{239} \\ &= 4\left(\frac{1}{5} - \frac{1}{3 \cdot 5^3} + \frac{1}{5 \cdot 5^5} - \cdots\right) - \left(\frac{1}{239} - \frac{1}{3 \cdot 239^3} + \cdots\right). \end{aligned}$$

因 $\frac{16}{11 \cdot 5^{11}} + \frac{4}{3 \cdot 239^3} < 1.3 \times 10^{-7}$, 故

$$\pi \approx 16\left(\frac{1}{5} - \frac{1}{3 \cdot 5^3} + \frac{1}{5 \cdot 5^5} - \frac{1}{7 \cdot 5^7} + \frac{1}{9 \cdot 5^9}\right) - \frac{4}{239} \approx 3.141\,593.$$

所列出的数字均是有效的.

例 19 计算积分 $\int_0^{0.1} e^{-x^2} dx$, 精确到 10^{-5} .

解 因 e^{-x^2} 的原函数不能用初等函数表示, 故不能直接用牛顿—莱布尼茨公式求解. 将 e^{-x^2} 展开成 x 的幂级数, 再逐项求积分

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \cdots,$$

从 0 到 0.1 积分得

$$\begin{aligned} \int_0^{0.1} e^{-x^2} dx &= \int_0^{0.1} \left(1 - x^2 + \frac{x^4}{2!} - \cdots\right) dx \\ &= \left[x - \frac{x^3}{3} + \frac{x^5}{10} - \cdots\right]_0^{0.1} \\ &= 0.1 - 0.000\,333\,3 + 0.000\,000\,1 - \cdots. \end{aligned}$$

不难看出, 若取前两项, 则其误差不超过 10^{-5} , 于是

$$\int_0^{0.1} e^{-x^2} dx \approx 0.1 - 0.000\,333 = 0.099\,67.$$

迄今为止, 我们都是在实数范围内展开讨论的. 在本节最后我们从幂级数公式推导复指数的欧拉公式.

例 20 利用指数函数、正弦函数、余弦函数的幂级数展式推导复指数的欧拉公式.

解 在 e^x 的幂级数展开式中, 以 ix 代换 x , 这里 $i = \sqrt{-1}$, 得

$$\begin{aligned}
e^{ix} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \cdots \\
&= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} + \cdots \\
&= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right) \\
&= \cos x + i \sin x,
\end{aligned}$$

即

$$e^{ix} = \cos x + i \sin x. \quad (4.2.35)$$

若以 $-x$ 换 x 可得

$$e^{-ix} = \cos x - i \sin x. \quad (4.2.36)$$

由上两式可解得

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}); \quad (4.2.37)$$

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix}). \quad (4.2.38)$$

上述(4.2.35)~(4.2.38)式均称为欧拉公式. 请读者将这些公式与复变量情形初等函数的定义相比较.

4.2.3 多元函数的泰勒公式与泰勒级数

我们以二元函数为例.

1. 二元函数的泰勒公式

定理 4.2.20 (二元函数的泰勒公式) 若函数 $f(x, y)$ 定义在二维平面区域 $D \subset \mathbb{R}^2$ 内. 在点 $(x_0, y_0) \in D$ 的某个邻域 $U_{(x_0, y_0)}$ 中有直到 $n+1$ 阶的所有连续偏导数, 则在此邻域 $U_{(x_0, y_0)}$ 中下述公式成立

$$\begin{aligned}
f(x, y) &= f(x_0, y_0) + \sum_{k=1}^n \frac{1}{k!} \left[(x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right]^k f(x_0, y_0) + \\
&\quad r_n(x, y),
\end{aligned} \quad (4.2.39)$$

其中余项为

$$\begin{aligned}
&r_n(x, y) \\
&= \frac{1}{(n+1)!} \left[(x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right]^{n+1} f(x_0 + \theta_n(x-x_0), y_0 + \theta_n(y-y_0)) \\
&\quad (0 < \theta_n < 1), \quad (4.2.40)
\end{aligned}$$

以上公式中的求偏导算子

$$\left[(x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right]^k \quad (k=1, 2, \cdots, n), \quad (4.2.41)$$

可按二项式公式展开. 例如, 若记 $x = x_0 + \Delta x$, $y = y_0 + \Delta y$, 则公式(4.2.39)可写为

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) &= \frac{1}{1!} [f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y] + \\ &\frac{1}{2!} [f''_{xx}(x_0, y_0)\Delta x^2 + 2f''_{xy}(x_0, y_0)\Delta x\Delta y + f''_{yy}(x_0, y_0)\Delta y^2] + \\ &\frac{1}{3!} [f'''_{xxx}(x_0, y_0)\Delta x^3 + 3f'''_{x^2y}(x_0, y_0)\Delta x^2\Delta y + 3f'''_{xy^2}(x_0, y_0)\Delta x\Delta y^2 + f'''_{yyy}\Delta y^3] + \\ &\cdots + r_n(x_0 + \Delta x, y_0 + \Delta y). \end{aligned}$$

多元函数的泰勒展开可类似得到.

2. 二元函数的泰勒级数

定理 4.2.21 (二元函数的泰勒级数) 若函数 $f(x, y)$ 定义在二维平面区域 $D \subset \mathbf{R}^2$ 内. 在点 $(x_0, y_0) \in D$ 的某个邻域 $U_{(x_0, y_0)}$ 中有任意阶的所有连续偏导数, 且余式满足

$$\lim_{n \rightarrow +\infty} r_n(x, y) = 0,$$

则函数 $f(x, y)$ 在邻域 $U_{(x_0, y_0)}$ 中可展开成泰勒级数

$$f(x, y) = f(x_0, y_0) + \sum_{k+j \geq 1}^{\infty} \frac{1}{k!j!} f^{(k+j)}_{x^k y^j}(x_0, y_0) (x - x_0)^k (y - y_0)^j. \quad (4.2.42)$$

公式(4.2.39)与(4.2.42)当 $x_0 = y_0 = 0$ 时, 就是熟知的麦克劳林展式与麦克劳林级数.

4.2.4 复级数、洛朗级数

前三小节所讨论的级数的项都是实值的, 本节转向复数项情形, 包括常数项的与函数项的复数项级数, 简称复级数.

1. 常数项复级数

当级数

$$\sum_{n=1}^{\infty} z_n = z_1 + z_2 + \cdots + z_n + \cdots \quad (4.2.43)$$

的通项 $z_n \in \mathbf{C}$ 是复数时, 称其为常数项复级数. 前 n 项的和为

$$\sigma_n = \sum_{k=1}^n z_k = z_1 + z_2 + \cdots + z_n.$$

部分和 σ_n 构成的复数列

$$\{\sigma_n\} : \sigma_1, \sigma_2, \cdots, \sigma_n, \cdots \quad (4.2.44)$$

为其部分和复数列.

函数项复级数的收敛性、绝对收敛性,其定义与函数项实级数类似.

进而,我们有以下定理:

定理 4.2.22 (复级数收敛、绝对收敛的充要条件) 复级数(4.2.43)收敛,当且仅当实级数 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 收敛;(4.2.43)绝对收敛,当且仅当实级数 $\sum_{n=1}^{\infty} a_n$ 与 $\sum_{n=1}^{\infty} b_n$ 绝对收敛,其中 $z_n = a_n + ib_n$.

此定理使我们把复级数的讨论及其性质的研究与实级数的情形联系起来.

2. 函数项复级数

函数项复级数一般可表示为下述形式

$$\sum_{n=1}^{\infty} f_n(z) = f_1(z) + f_2(z) + \cdots + f_n(z) + \cdots, \quad (4.2.45)$$

式中 $f_n(z)$ 均为定义在复平面上某区域 D 内的复函数.(4.2.45)可写为

$$\sum_{n=1}^{\infty} f_n(z) = \sum_{n=1}^{\infty} [u_n(x, y) + iv_n(x, y)], \quad (4.2.46)$$

这里 $u_n(x, y)$ 、 $v_n(x, y)$ 分别为 $f_n(z)$ 的实部与虚部.

函数项复级数(4.2.45)的收敛点、收敛域、绝对收敛性、一致收敛性、内闭一致收敛性等,其定义都与函数项实级数类似.类似于定理 4.2.22 的定理也成立,并且还有如下的

定理 4.2.23 (魏尔斯特拉斯判别法) 设 $f_n(z)$ ($n=1, 2, \cdots$) 定义在复平面 C 的区域 D 中.若存在收敛的正项常数项实级数 $\sum_{n=1}^{\infty} M_n$,使得

$$|f_n(z)| \leq M_n \quad (4.2.47)$$

对充分大的 n 及所有 $z \in D$ 成立,则复级数 $\sum_{n=1}^{\infty} f_n(z)$ 在 D 中绝对一致收敛.

下面的魏尔斯特拉斯定理在复级数的研究中非常重要.

定理 4.2.24 (魏尔斯特拉斯定理) 设 $f_n(z)$ ($n=1, 2, \cdots$) 在复平面 C 的区域 D 内解析,且复级数 $\sum_{n=1}^{\infty} f_n(z)$ 在 D 中内闭一致收敛于复函数 $f(z)$,则 $f(z)$ 在 D 内解析,并且成立

$$f^{(k)}(z) = \sum_{n=1}^{\infty} f_n^{(k)}(z) \quad (k=1, 2, 3, \cdots).$$

3. 复幂级数

形如

$$\sum_{n=0}^{\infty} \alpha_n (z - z_0)^n \quad (4.2.48)$$

的复级数称为复幂级数,其中 $z = x + iy$ 是复变数,系数 α_n 是复常数.

对于复幂级数,有类似于实系数幂级数的阿贝尔定理(定理 4.2.10),从而存在常数 R ($0 \leq R \leq \infty$),使得

(1) 若 $0 < R < \infty$,则当 $|z - z_0| < R$ 时,复幂级数(4.2.48)绝对收敛,而当 $|z - z_0| > R$ 时级数发散.此时,称 $R > 0$ 为复幂级数(4.2.48)的收敛半径,称 $|z - z_0| < R$ 为其收敛圆(或收敛圆盘);

(2) 若 $R = \infty$,则复幂级数(4.2.48)在复平面的每一点都绝对收敛.此时,称复幂级数的收敛半径为 $R = +\infty$,而其收敛圆为全复平面;

(3) 若 $R = 0$,则复幂级数(4.2.48)在复平面上除去 $z = z_0$ 外每一点都发散.此时,称复幂级数的收敛半径为 $R = 0$,而其收敛圆收缩为一点 z_0 .

定理 4.2.25 (复幂级数的和函数) 设复幂级数(4.2.48)的收敛半径 $R > 0$,收敛圆 $|z - z_0| < R$,则在此收敛圆中,级数内闭一致收敛于其和函数 $f(z)$,且 $f(z)$ 在收敛圆盘内解析,并成立

$$f^{(n)}(z) = n! \alpha_n + \frac{(n+1)!}{1!} \alpha_{n+1} (z - z_0) + \frac{(n+2)!}{2!} \alpha_{n+2} (z - z_0)^2 + \cdots \quad (n=0,1,2,\cdots) \quad (4.2.49)$$

4. 解析函数的泰勒展式

我们直接写出解析函数的泰勒展式:

设函数 $f(z)$ 在圆盘 $U: |z - z_0| < R$ 内解析,则在 U 内有

$$f(z) = f(z_0) + \frac{f'(z_0)}{1!} (z - z_0) + \frac{f''(z_0)}{2!} (z - z_0)^2 + \cdots + \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n + \cdots, \quad (4.2.50)$$

称(4.2.50)为解析函数 $f(z)$ 泰勒展式.由此,可证明

定理 4.2.26 (解析函数的充要条件) 函数 $f(z)$ 在一点 z_0 解析,当且仅当它在 z_0 的某一邻域内有惟一的(4.2.50)式的幂级数展式(泰勒级数).

常用的解析函数的幂级数展式有

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^n}{n!} + \cdots \quad (z \in \mathbb{C}),$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots \quad (z \in \mathbb{C}),$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \cdots \quad (z \in \mathbb{C}),$$

$$\arctan z = z - \frac{z^3}{3} + \frac{z^5}{5} - \cdots \quad (z \in \mathbb{C}),$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} + \cdots + (-1)^{n-1} \frac{z^n}{n} + \cdots \quad (|z| < 1),$$

$$(1+z)^\alpha = 1 + \alpha z + \frac{\alpha(\alpha-1)}{2!} z^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!} z^3 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} z^n + \cdots \quad (|z| < 1),$$

5. 复变函数的洛朗展式

在 § 3.7 与本节中, 我们研究了在点 $z = z_0$ 解析的函数. 然而, 许多函数并非解析, 例如 $\frac{1}{z - z_0}$, 在 $z = z_0$ 根本没有定义. 如何研究这类函数呢? 洛朗 (Laurent) 级数起了重要作用.

考虑形如

$$\sum_{n=-\infty}^{\infty} \alpha_n (z - z_0)^n \quad (4.2.51)$$

的“双向复级数”, 其中 z_0, α_n 都是复常数, 我们称 (4.2.51) 为洛朗级数. 当两个级数

$$\sum_{n=0}^{\infty} \alpha_n (z - z_0)^n \quad (4.2.52)$$

与

$$\sum_{n=-\infty}^{-1} \alpha_n (z - z_0)^n \quad (4.2.53)$$

都收敛时, 称洛朗级数 (4.2.51) 收敛.

设级数 (4.2.52) 的收敛半径为 R_1 , 亦即, 此级数在 $|z - z_0| < R_1$ 中绝对收敛且内闭一致收敛.

另一方面, 令 $w = \frac{1}{z - z_0}$, 则式 (4.2.53) 化为

$$\sum_{n=1}^{+\infty} \alpha_{-n} w^n. \quad (4.2.54)$$

假设复幂级数 (4.2.54) 的收敛半径为 $0 < R_3 < +\infty$, 令 $R_2 = \frac{1}{R_3}$, 则级数

(4.2.53) 在 $|z - z_0| > \frac{1}{R_3} = R_2$ 中绝对收敛并且内闭一致收敛, 在 $|z - z_0| < \frac{1}{R_3} = R_2$ 中发散.

于是, 作为两个级数 (4.2.52) 与 (4.2.53) 之和的复幂级数 (4.2.51), 分别在区域 $|z - z_0| < R_1$ 与 $|z - z_0| > R_2$ 中解析.

如果 $0 < R_2 < R_1$, 则洛朗级数 (4.2.51) 在环形区域

$$R_2 < |z - z_0| < R_1$$

中绝对收敛并且内闭一致收敛, 而且此级数的和函数在环形区域内解析.

我们有

定理 4.2.27 (在环中解析的函数的洛朗展式) 若函数 $f(z)$ 在圆环 $D: R_2 < |z - z_0| < R_1$ ($0 \leq R_2 < R_1 \leq +\infty$) 内解析, 则在 D 中有 $f(z)$ 的洛朗展式

$$f(z) = \sum_{n=-\infty}^{\infty} \alpha_n (z - z_0)^n, \quad (4.2.55)$$

其中

$$\alpha_n = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (n = 0, \pm 1, \pm 2, \dots),$$

Γ 是圆周 $|z - z_0| = \rho$, 而 ρ 是满足 $R_2 < \rho < R_1$ 的任一实数, 并且 $f(z)$ 的洛朗展式是惟一的.

定义 4.2.4 (洛朗级数的解析部分与主要部分) 在复变量的复值函数 $f(z)$ 的洛朗展式(4.2.54)中, 我们分别称

$$\sum_{n=0}^{\infty} \alpha_n (z - z_0)^n \quad \text{与} \quad \sum_{n=-\infty}^{-1} \alpha_n (z - z_0)^n$$

为 $f(z)$ 的解析部分与主要部分.

定义 4.2.5 (解析函数的孤立奇点) 设函数 $f(z)$ 在除去圆心的圆盘 $D: 0 < |z - z_0| < R$ ($0 < R \leq +\infty$) 内解析, 则我们称 z_0 为 $f(z)$ 的孤立奇点. 此时, 在圆盘 D 内 $f(z)$ 有洛朗展式

$$f(z) = \sum_{n=-\infty}^{\infty} \alpha_n (z - z_0)^n,$$

其中

$$\alpha_n = \frac{1}{2\pi i} \int_{C_\rho} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (n = 0, \pm 1, \pm 2, \dots),$$

C_ρ 是圆周 $|z - z_0| = \rho$ ($0 < \rho < R$).

(1) 当 $n = -1, -2, \dots$ 都有 $\alpha_n = 0$ 时, 则称 z_0 为 $f(z)$ 的可去奇点. 此时, 若补充令 $f(z_0) = \alpha_0$, 则 $f(z)$ 在整个圆盘 $|z - z_0| < R$ 内解析.

(2) 如果只有有限个 $n < 0$, 使得 $\alpha_n \neq 0$, 则称 z_0 为 $f(z)$ 的极点. 此时, 必存在正整数 $m > 0$, 使得

$$\alpha_n \begin{cases} \neq 0, & n \geq -m, \\ = 0, & n < -m, \end{cases}$$

我们称 z_0 为 $f(z)$ 的 m 阶极点, 并且当 $m = 1$ 时称 z_0 为 $f(z)$ 的单极点, $m > 1$ 时称 z_0 为 $f(z)$ 的 m 重极点.

(3) 如果有无穷多个 $n < 0$, 使得 $\alpha_n \neq 0$, 则称 z_0 为 $f(z)$ 的本性奇点.

由以上的定义知道, 如果函数 $f(z)$ 展开成 $z - z_0$ 的幂级数, $R > 0$ 为收敛半径, 则在收敛圆周 $|z - z_0| = R$ 上, 至少有函数 $f(z)$ 的一个奇点.

例 21 求函数 $f(z) = \frac{1}{(z-1)(z-2)}$ 分别在圆环 (i) $1 < |z| < 2$ 内;
(ii) $2 < |z| < +\infty$ 内的洛朗展式.

解 (i) 在 $1 < |z| < 2$ 内.

$$\begin{aligned}\frac{1}{(z-1)(z-2)} &= \frac{1}{z-2} - \frac{1}{z-1} = \frac{1}{2\left(1-\frac{z}{2}\right)} - \frac{1}{z\left(1-\frac{1}{z}\right)} \\ &= \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n - \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n \\ &= \sum_{n=0}^{\infty} \frac{z^n}{2^{n+1}} - \sum_{n=-\infty}^{-2} z^{n+1} = \sum_{n=-\infty}^{\infty} a_n z^n,\end{aligned}$$

其中 $a_n = \begin{cases} \frac{1}{2^{n+1}}, & n \geq 0, \\ -1, & n < 0 \end{cases}$.

(ii) 在 $2 < |z| < +\infty$ 内.

$$\begin{aligned}\frac{1}{(z-1)(z-2)} &= \frac{1}{z-2} - \frac{1}{z-1} = \frac{1}{z\left(1-\frac{2}{z}\right)} - \frac{1}{z\left(1-\frac{1}{z}\right)} \\ &= \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{2}{z}\right)^n - \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n \\ &= \sum_{n=2}^{\infty} \frac{2^{n-1} - 1}{z^n}.\end{aligned}$$

对于孤立奇点 z_0 属于那种类型, 我们有如下的判别法.

定理 4.2.28 (孤立奇点的类型) 设函数 $f(z)$ 在除去圆心的圆盘 $D: 0 < |z - z_0| < R$ ($0 < R \leq +\infty$) 内解析, 则

(1) z_0 是 $f(z)$ 的可去奇点, 当且仅当极限 $\lim_{z \rightarrow z_0} f(z) = \alpha_0$ 是一个复数;

(2) z_0 是 $f(z)$ 的极点, 当且仅当极限 $\lim_{z \rightarrow z_0} f(z) = +\infty$;

z_0 是 $f(z)$ 的 m 阶极点, 当且仅当极限 $\lim_{z \rightarrow z_0} (z - z_0)^m f(z) = \alpha_{-m} \neq 0$; 这里 m 为一正整数.

(3) z_0 是 $f(z)$ 的本性奇点, 当且仅当不存在有限的或无穷的极限 $\lim_{z \rightarrow z_0} f(z)$.

6. 复幂级数向收敛圆外的解析开拓

定理 4.2.25 告诉我们, 复幂级数 (4.2.48) 的和函数在其收敛圆 $|z - z_0| < R$ 内解析, 并成立

$$f^{(n)}(z) = n! \alpha_n + \frac{(n+1)!}{1!} \alpha_{n+1} (z - z_0) + \frac{(n+2)!}{2!} \alpha_{n+2} (z - z_0)^2 + \cdots \quad (n=0, 1, 2, \cdots).$$

在收敛圆的周界 $|z - z_0| = R$ 上情形如何呢? 我们有:

定理 4.2.29 (收敛圆周的性质) 设复幂级数(4.2.48)有收敛半径 $0 < R < +\infty$, 则在圆周 $|z - z_0| = R$ 上, 至少有 $f(z)$ 的一个奇点(或极点、或本性奇点、或支点).

现在来解决复幂级数(4.2.48)能否向收敛圆外“解析开拓”?

对于复幂级数(4.2.48), 在其收敛圆 $D: |z - z_0| < R$ 中, 和函数 $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ 解析. 在 D 内, 任取一点 $z_1 \neq z_0$, 见图 4.2.1. 将解析函数 $f(z)$ 展开成 $z - z_1$ 的复幂级数

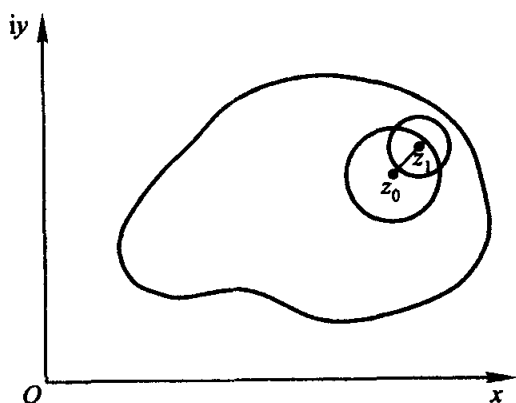


图 4.2.1

$$\sum_{n=0}^{\infty} \beta_n (z - z_1)^n, \quad (4.2.56)$$

其中

$$\beta_n = \frac{f^{(n)}(z_1)}{n!} \quad (n = 0, 1, 2, \dots).$$

设复幂级数(4.2.56)的收敛半径为 R_1 , 收敛圆为 $D_1: |z - z_1| < R_1$. 显然, z_1 与圆周 $|z - z_0| = R$ 上任一点的距离不小于 $R - |z_0 - z_1|$. 于是, 在圆盘 $|z - z_1| < R - |z_0 - z_1|$ 内, 函数 $f_1(z)$ 解析. 由此可见, 级数(4.2.56)的收敛半径 R_1 一定满足

$$R_1 \geq R - |z_0 - z_1|.$$

于是, 可能发生两种情况:

(1) 如果 $R_1 > R - |z_0 - z_1|$, 则(4.2.56)确定了一个在圆盘 $D_1: |z - z_1| < R_1$ 内解析的函数 $f_1(z)$, 并且 D_1 必定有一部分在 D 之外(图 4.2.1). 而在两个圆盘 D 与 D_1 的公共部分, 必有 $f(z) = f_1(z)$ (解析函数的惟一性). 这时,

我们说:解析函数 $f(z)$ 可以通过点 z_1 的半径方向, 解析开拓到 D 外.

(2) 如果 $R_1 = R - |z_0 - z_1|$, 则 D_1 包含在 D 内. 这时, 我们说: 解析函数 $f(z)$ 不能通过点 z_1 的半径方向, 开拓到 D 外. 圆盘 D_1 与 D 的边界必相切, 而切点正是函数 $f(z)$ 的一个奇点(或极点、或本性奇点、或支点).

例 22 函数

$$\frac{1}{1-z} = \sum_{n=0}^{+\infty} z^n,$$

在其收敛圆 $D: |z| < 1$ 内解析, 收敛半径 $R = 1$, 记其和函数为 $f(z) = \frac{1}{1-z}$. 试

求: (i) $f(z)$ 在点 $z = \frac{i}{2}$ 的幂级数展开式; (ii) $f(z)$ 在点 $z = a$ ($0 < a < 1$) 的幂级数展开式.

解 (i) 由

$$f\left(\frac{i}{2}\right) = \frac{2}{5}(2+i),$$

$$f^{(n)}\left(\frac{i}{2}\right) = n! \left(\frac{2}{5}\right)^{n+1} (2+i)^{n+1} \quad (n=1, 2, \dots).$$

于是, $f(z)$ 在 $z = \frac{i}{2}$ 的幂级数展开式为

$$f_1(z) = \sum_{n=0}^{+\infty} \left[\frac{2}{5}(2+i) \right]^{n+1} \left(z - \frac{i}{2} \right)^n,$$

其收敛半径为 $\frac{\sqrt{5}}{2}$, 而且 $f_1(z)$ 在圆盘 $D_1: \left| z - \frac{i}{2} \right| < \frac{\sqrt{5}}{2}$ 内解析. 圆盘 D_1 有一部分在圆盘 D 之外(图 4.2.2). 在 $D \cap D_1$ 中, $f(z) = f_1(z)$, 而 $f(z)$ 可以通过 $z = \frac{i}{2}$ 的半径方向解析开拓到 D 之外.

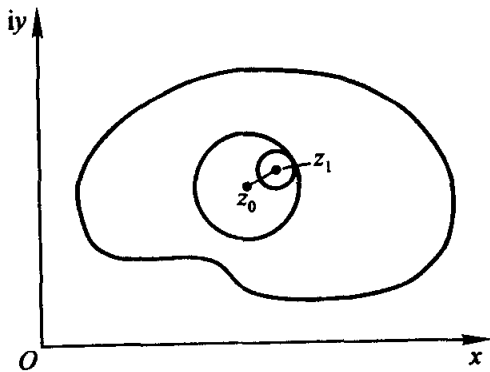


图 4.2.2

(ii) 容易求出函数 $f(z)$ 在 $z=a$ 的展开是为

$$f_1(z) = \frac{1}{1-a} + \frac{1}{(1-a)^2}(z-a) + \frac{1}{(1-a)^3}(z-a)^2 + \cdots,$$

其收敛半径为 $R_1 = 1-a$ 与 $R=1$ 相比较, 恰好有 $R_1 = 1-a = 1-|a-0| = R - |z_1 - z_0|$. 因此, $z=1$ 是 $f(z) = \frac{1}{1-z}$ 的奇点(显然是一个一阶极点).

例 23 实幂级数函数

$$\sum_{n=0}^{+\infty} (-1)^n x^{2n}$$

的收敛半径 $R=1$, 收敛区间为 $(-1, 1)$, 其和函数为 $\frac{1}{1+x^2}$. 我们有

$$\frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n x^{2n} \quad (x \in (-1, 1)).$$

显然在 $x = \pm 1$ 时, 右边级数发散; 但是, 左边的函数 $\frac{1}{1+x^2}$ 在 $(-\infty, +\infty)$ 中有意义. 试解释这个不一致发生的原因.

解 考虑复级数

$$\frac{1}{1+z^2} = \sum_{n=0}^{+\infty} (-1)^n z^{2n}, \quad (4.2.57)$$

令 $w = z^2$, 上式化为

$$\frac{1}{1+w} = \sum_{n=0}^{+\infty} (-1)^n w^n, \quad (4.2.58)$$

不难判断, 级数(4.2.58)的收敛半径 $r=1$, 收敛圆 $|w| < 1$. 因此, 级数(4.2.58)的收敛半径为 $R=1$, 收敛圆为 $|z| < 1$. 取收敛圆内的点 $z_1 = \frac{i}{2}$, 将复函数 $f(z) = \frac{1}{1+z^2}$ 在 $z_1 = \frac{i}{2}$ 处展开为:

$$f_1(z) = \sum_{n=0}^{+\infty} \left[\frac{f^{(n)}\left(\frac{i}{2}\right)}{n!} \left(z - \frac{i}{2}\right)^n \right],$$

其收敛半径的计算较繁, 但可以求出 $r_1 = \frac{1}{2}$, 并且有 $r_1 = \frac{1}{2} = 1 - \frac{1}{2} = 1 - \left| \frac{i}{2} - 0 \right|$, 所以, $z=i$ 为函数 $f(z) = \frac{1}{1+z^2}$ 的奇点(一级极点). 其实, $z=-i$ 为函数 $f(z) = \frac{1}{1+z^2}$ 的另一个奇点.

例 24 函数 $\ln(1+z)$ 在 $z=0$ 的幂级数展开式是

$$f_1(z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \cdots + (-1)^{n-1} \frac{z^n}{n} + \cdots,$$

收敛半径 $R=1$. 取 $z_0 = -\frac{1}{2}$ 的幂级数:

$$f_2(z) = -\ln 2 + \sum_{n=0}^{\infty} (-a)^{n-1} \frac{2^n}{n} \left(z + \frac{1}{2}\right)^n,$$

并获得其收敛半径为 $r_2 = \frac{1}{2} = 1 - \left| -\frac{1}{2} - 0 \right|$. 因此, $z = -1$ 是 $f_2(z)$ 的奇点 (其实是 $\ln(1+z)$ 的支点).

习题

1. 求下列级数的收敛域:

$$(1) \sum_{n=1}^{\infty} \frac{n^2}{x^n}$$

$$(2) \sum_{n=1}^{\infty} \frac{x^n}{1-x^n}$$

$$(3) \sum_{n=1}^{\infty} (\sin x)^n$$

$$(4) \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \left(\frac{x}{1+x}\right)^n$$

$$(5) \sum_{n=1}^{\infty} \frac{x^{n^2}}{2^n}$$

2. 讨论下列函数在给定区间上的一致收敛性

$$(1) \sum_{n=1}^{\infty} \frac{1}{x^2 + n^2}, x \in (-\infty, +\infty)$$

$$(2) \sum_{n=1}^{\infty} \frac{(-1)^n}{x + 2^n}, x \in [0, +\infty)$$

$$(3) \sum_{n=1}^{\infty} \frac{(-1)^n}{n + \sin^2 x}, x \in (-\infty, +\infty)$$

$$(4) \sum_{n=1}^{\infty} (x^n - x^{n-1}), x \in \left[0, \frac{1}{2}\right]$$

3. 证明函数 $f(x) = \sum_{n=1}^{\infty} e^{-n^2 x}$ 当 $x > 0$ 时有定义且连续.

4. 求下列幂级数的收敛域:

$$(1) \sum_{n=1}^{\infty} nx^n$$

$$(2) \sum_{n=1}^{\infty} \frac{2n}{n+1} x^n$$

$$(3) \sum_{n=1}^{\infty} n! x^n$$

$$(4) \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} x^n$$

$$(5) \sum_{n=1}^{\infty} \frac{(-1)^n}{n} (x-1)^n$$

$$(6) \sum_{n=1}^{\infty} \frac{2n-1}{2^n} x^{2n-1}$$

$$(7) \sum_{n=1}^{\infty} \left(\frac{3^n}{n^2} + \frac{2^n}{n}\right) x^n$$

$$(8) \sum_{n=1}^{\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) x^n$$

5. 若幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的收敛半径为 R , $a_n \neq 0$, 是否必有 $\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = R$? 试考察幂级

数 $\sum_{n=1}^{\infty} (2 + (-1)^n) x^n$, 并检验你的结论.

6. 已知两个幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 与 $\sum_{n=0}^{\infty} b_n x^n$ 的系数之比的绝对值的极限 $\lim_{n \rightarrow \infty} \left| \frac{a_n}{b_n} \right| = l$ 为

非零常数. 求证他们有相同的收敛半径.

7. 若 $a_n \neq 0$, 且

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = l \quad (0 \leq l \leq +\infty),$$

求证幂级数 $\sum_{n=0}^{\infty} \frac{x^n}{a_n}$ 的收敛半径为 $\frac{1}{l}$.

8. 利用第 7 题的结论求下列幂级数的收敛半径:

$$(1) \sum_{n=1}^{\infty} \left(\frac{x}{n}\right)^n \quad (2) \sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^{n^2} x^n$$

9. 利用逐项微分或逐项积分求下列幂级数的和函数:

$$\begin{aligned} (1) \sum_{n=0}^{\infty} (n+1)x^n & \quad (2) \sum_{n=1}^{\infty} n(n+1)x^n \\ (3) \sum_{n=1}^{\infty} \frac{x^n}{n(n+1)} & \quad (4) \sum_{n=0}^{\infty} (2n-1)x^n \\ (5) \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n-1)(2n+1)} & \quad (6) \sum_{n=1}^{\infty} \frac{2n-1}{2^n} x^{2n-1} \end{aligned}$$

10. 求证下列等式:

$$(1) 6 = \sum_{n=1}^{\infty} \frac{n^2}{2^n} \quad (2) 10 = \sum_{n=1}^{\infty} \frac{n(n+2)}{2^n}$$

11. 写出 $x^2 e^x$ 的麦克劳林级数, 并证明:

$$e = 2 + \sum_{n=0}^{\infty} \frac{1}{(n+3) \cdot n!}.$$

12. 求下列函数的麦克劳林级数, 并求收敛区间:

$$\begin{aligned} (1) \frac{x}{x^2 - 2x - 3} & \quad (2) \frac{1}{2x^2 - x - 1} \\ (3) \frac{1}{x^2 + x + 1} & \quad (4) \sin^2 x \\ (5) \operatorname{sh} x & \quad (6) \operatorname{ch} x \\ (7) (1+x)e^{-x} & \quad (8) \arcsin x \\ (9) \frac{1}{\sqrt{1+x^2}} & \quad (10) \int_0^x e^{-t^2} dt \\ (11) \int_0^x \frac{\sin t}{t} dt & \quad (12) \frac{e^x}{1-x} \end{aligned}$$

13. 验证下列展式(下式中 a 为某一常数):

$$\begin{aligned} (1) a^x &= \sum_{n=0}^{\infty} \frac{1}{n!} (\ln a)^n x^n, \quad a > 0, a \neq 1, -\infty < x < +\infty \\ (2) e^x &= e^a \sum_{n=0}^{\infty} \frac{1}{n!} (x-a)^n, \quad -\infty < x < +\infty \\ (3) \sin x &= \sum_{n=0}^{\infty} \frac{1}{n!} \sin\left(a + \frac{n\pi}{2}\right) (x-a)^n, \quad -\infty < x < +\infty \\ (4) \ln(a+x) &= \ln a - \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \left(\frac{x}{a}\right)^n, \quad a > 0, |x| < a. \end{aligned}$$

14. 展 $\frac{d}{dx} \left(\frac{e^x - 1}{x} \right)$ 为 x 的幂级数, 并求 $\sum_{n=1}^{\infty} \frac{n}{(n+1)!}$ 的和.

15. 将函数 $f(x) = \frac{1}{(1+x)^2}$ 在 $x=1$ 处展成幂级数, 并求收敛区间.

16. 利用函数幂级数展式, 计算下列积分之值, 精确到小数点后第三位:

$$(1) \int_0^1 \frac{\sin x}{x} dx$$

$$(2) \int_0^{\frac{1}{2}} \frac{dx}{\sqrt{1+x^4}}$$

17. 利用 $(1+x)^a$ 的展开式的前两项计算下列各值:

$$(1) \sqrt{1.002}$$

$$(2) \sqrt[4]{80}$$

18. 求 $\ln \frac{1+x}{1-x}$ 关于 x 的幂级数展式, 证明对自然数 n 有

$$\ln(1+n) = \ln n + \frac{2}{2n+1} + \frac{2}{3(2n+1)^3} + \frac{2}{5(2n+1)^5} + \cdots,$$

并求 $\ln 2$ 精确到 10^{-4} 的近似值.

§ 4.3 傅里叶级数

本节是继幂级数后要研究的另一类特殊的函数项级数——傅里叶(Fourier)级数. 傅里叶级数既是基础数学中的一个重要分支, 又是工程技术, 特别是无线电、通讯、数据处理中的一个不可缺少的重要数学工具.

4.3.1 直交(正交)函数系

从数学角度来看, 由简单的正整数幂函数 $(x-x_0)^n$ 与常数组成的幂级数在其收敛区间内可以表示一个较为复杂的函数. 反之, 如果一个复杂的函数在某一区间上能展开为一个幂级数, 则这种展开由该函数在某一点附近的值惟一确定. 事实上, 一个较为复杂的函数由 $f(x_0), f'(x_0), \cdots, f^{(n)}(x_0), \cdots$ 惟一确定, 就必须要求该函数具有很好的性质, 例如, 在该区间内至少无穷次可微. 于是对函数性质稍差的函数就不能用上面的分解, 必须寻找另外的分解. 为此可以从两个方面去考虑: 一是寻找代替 $\{(x-x_0)^n\}_{n=0}^{\infty}$ 的“函数系”, 其中的函数应当具有很好的性质; 二是对一个函数进行分解时, 不能只依赖于某一点附近的值, 而要依赖于几乎整个区间上的值. 人们发现, 三角函数系

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots$$

可以担当此任.

从物理或应用的角度讲, 一个简谐振动可以由正弦函数

$$y = A \sin(\omega t + \varphi)$$

来表示, 其中 A 是振幅, ω 是频率, φ 是初相, 振动的周期是 $T = \frac{2\pi}{\omega}$. 若令 $x = \omega t$

$= \frac{2\pi}{T}t$, 则 y 就变成了以 2π 为周期的正弦函数

$$y = A \sin(x + \varphi).$$

同样, 函数

$$y_k = A_k \sin(kx + \varphi_k) \quad (k=1, 2, \dots)$$

表示周期为 $\frac{2\pi}{k}$ 的简谐振动. 取 $y_0 = A_0 \sin \varphi_0$ 为一常数, 表示一个平衡在初相状态的振动. 由 $y_0, y_1, \dots, y_n$ 的叠加可以构成一个较复杂的振动

$$\sum_{k=0}^n y_k = \sum_{k=0}^n A_k \sin(kx + \varphi_k),$$

它仍具有周期 2π . 若有无穷多个 y_k 叠加, 便得一级数

$$\sum_{k=0}^{\infty} A_k \sin(kx + \varphi_k). \quad (4.3.1)$$

对 $\sin(kx + \varphi_k)$ 进行三角恒等变换

$$\sin(kx + \varphi_k) = \sin \varphi_k \cos kx + \cos \varphi_k \sin kx,$$

并令 $\frac{a_0}{2} = A_0 \sin \varphi_0, a_k = A_k \sin \varphi_k, b_k = A_k \cos \varphi_k$, 则 (4.3.1) 变为

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (4.3.2)$$

称为(实)三角级数. 如果此三角级数处处收敛, 它也代表一个周期为 2π 的振动. 在实际问题中, 人们经常遇到的是相反的问题: 一个复杂的、以 2π 为周期的振动能否表示成上述有限或无限个简谐振动 $y_0, y_1, \dots, y_n, \dots$ 的叠加? 如果能, 如何确定这些简谐振动? 用数学语言描述, 就是: 给定一个以 2π 为周期的函数 $f(x)$, 能否有等式

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \quad (4.3.3)$$

成立? 系数 a_k, b_k 如何确定? 傅里叶级数理论对这两个问题给出了有实际应用的回答.

如果有函数 $f(x)$ 能使等式 (4.3.3) 成立, 我们称 $f(x)$ 能展开成三角级数. 将函数 $f(x)$ 展开成三角级数, 可归结为两个问题:

(i) 给定 $f(x)$, 如何求系数 a_n, b_n ?

(ii) $f(x)$ 在什么条件下等式 (4.3.3) 成立?

问题(ii)就是通常函数项级数的收敛性和求和函数的问题. 而讨论 a_n, b_n 与 $f(x)$ 的关系要用到三角函数系的直交性.

首先给出一般函数系直交的概念.

定义 4.3.1 (直交函数系) 设 $S = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ 是一个函数集合. 若 S 中的每个函数 $\varphi_n(x)$ 是 $[a, b]$ 上定义的黎曼可积实(或复)值函数, 且满足

$$\int_a^b \varphi_n(x) \overline{\varphi_m(x)} dx = \begin{cases} 0, & n \neq m, \\ \lambda_n > 0, & n = m, \end{cases} \quad (4.3.4)$$

则称 S 是 $[a, b]$ 上的直交函数系 (或正交函数系), 简称直交系. 若 $\varphi_n(x)$ 是实值函数, 则 $\overline{\varphi_n(x)} = \varphi_n(x)$; 若 $\varphi_n(x)$ 是复值函数, $\varphi_n(x) = \alpha_n(x) + i\beta_n(x)$, 其中 $\alpha_n(x), \beta_n(x)$ 为实值函数, 其共轭函数为 $\overline{\varphi_n(x)} = \alpha_n(x) - i\beta_n(x)$, 且

$$\int_a^b \varphi_n(x) dx = \int_a^b \alpha_n(x) dx + i \int_a^b \beta_n(x) dx.$$

若 (4.3.4) 式中的 $\lambda_n = 1$ ($n = 0, 1, 2, \dots$), 则称 S 是标准直交 (正交) 系.

例 1 三角函数系 $S = \{\varphi_n(x)\}_{n=0}^{\infty}$, 其中

$$\varphi_0(x) = 1, \varphi_{2n-1}(x) = \cos nx, \varphi_{2n}(x) = \sin nx \quad (n = 1, 2, \dots)$$

是 $[-\pi, \pi]$ 上的直交系, 据周期性知, 也是 $[c, c + 2\pi]$ 上的直交系, c 是任何实数.

证 从等式

$$\begin{aligned} \int_{-\pi}^{\pi} 1 \cdot \sin nx dx &= \int_{-\pi}^{\pi} 1 \cdot \cos nx dx = 0, \\ \int_{-\pi}^{\pi} \cos nx \sin mx dx &= 0 \quad (n, m = 1, 2, \dots), \\ \int_{-\pi}^{\pi} \cos nx \cos mx dx &= \int_{-\pi}^{\pi} \sin nx \sin mx dx = 0 \quad (n \neq m), \\ \int_{-\pi}^{\pi} (\sin mx)^2 dx &= \int_{-\pi}^{\pi} (\cos mx)^2 dx = \pi \quad (m = 1, 2, \dots), \\ \int_{-\pi}^{\pi} 1^2 dx &= 2\pi, \end{aligned}$$

知函数系 S 满足 (4.3.4), 其中 $\lambda_0 = 2\pi, \lambda_n = \pi$ ($n = 1, 2, \dots$). 所以 S 是直交系.

由于对每个具有周期 T 的连续函数 $\varphi(x)$, 恒有

$$\int_0^T \varphi(x) dx = \int_c^{T+c} \varphi(x) dx \quad (c \text{ 为任何实数}),$$

故上述在 $[-\pi, \pi]$ 上的各个积分换成在 $[c, c + 2\pi]$ 上的积分后, 其结果不变, 从而 S 也是 $[c, c + 2\pi]$ 上的直交系.

由例 1 的证明易知, 三角函数系

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \quad \varphi_{2n-1}(x) = \frac{1}{\sqrt{\pi}} \cos nx, \quad \varphi_{2n}(x) = \frac{1}{\sqrt{\pi}} \sin nx, \quad (n = 1, 2, \dots)$$

是 $[c, c + 2\pi]$ 上的标准直交系.

例 2 复指数函数系 $S = \{\varphi_n(x)\}_{n \in \mathbb{Z}}$, 其中

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} = \frac{1}{\sqrt{2\pi}} (\cos nx + i \sin nx) \quad (n \in \mathbb{Z})$$

是 $[-\pi, \pi]$ 上、也是 $[c, c+2\pi]$ 上的标准直交系.

证 对任意的 $n, m \in \mathbb{Z}$, 由

$$\varphi_n(x) \overline{\varphi_m(x)} = \frac{1}{2\pi} e^{i(n-m)x} = \frac{1}{2\pi} [\cos(n-m)x + i \sin(n-m)x],$$

由此得: 若 $n \neq m$, 则 $\int_{-\pi}^{\pi} \varphi_n(x) \overline{\varphi_m(x)} dx = 0$; 若 $n = m$, 则

$$\int_{-\pi}^{\pi} \varphi_n(x) \overline{\varphi_n(x)} dx = \int_{-\pi}^{\pi} \frac{1}{2\pi} dx = 1.$$

故 S 是 $[-\pi, \pi]$ 上的标准直交系, 再由于 $\varphi_n(x)$ ($n \in \mathbb{Z}$), 都是 2π 周期的, 从而由例 1 的说明知, S 也是 $[c, c+2\pi]$ 上的标准直交系.

例 3 函数系

$$\left\{ 1, \cos \frac{n\pi x}{l}, \sin \frac{n\pi x}{l} \right\}_{n=1}^{\infty}$$

是 $[0, 2l]$ 或 $[-l, l]$ 上的直交系, 而

$$\left\{ \frac{1}{\sqrt{2l}}, \frac{1}{\sqrt{l}} \cos \frac{n\pi x}{l}, \frac{1}{\sqrt{l}} \sin \frac{n\pi x}{l} \right\}_{n=1}^{\infty}$$

是同一区间上的标准直交系. (证明留作习题)

注 1 直交函数系的概念很重要, 因为建立在三角函数系上的傅里叶级数理论与许多结果可移植到其他直交函数系上.

例 4 函数系 $S = \{\varphi_0(x), \varphi_1(x), \varphi_3, \dots\}$, 其中

$$\varphi_0(x) = \begin{cases} 1, & 0 \leq x \leq \frac{1}{2}, \\ -1, & \frac{1}{2} \leq x \leq 1; \end{cases} \quad \varphi_n(x) = \varphi_0(2^n x) \quad (n=1, 2, 3, \dots).$$

此函数系也是标准直交的, 在计算机与信号处理过程中经常用到.

注 2 在讨论直交函数系时, 另一个重要概念“完整性”, 也需要在此强调一下.

定义 4.3.2 (完整直交函数系) 设 $S = \{\varphi_0, \varphi_1, \varphi_2, \dots\}$ 是 $[a, b]$ 上一个直交函数系, 并且满足:

如果定义在 $[a, b]$ 的任何黎曼可积实(或复)值函数 $f(x)$, 对于 $n=0, 1, 2, 3, \dots$ 都成立

$$\int_a^b f(x) \varphi_n(x) dx = 0, \quad (4.3.5)$$

蕴含 $f(x) \equiv 0$, 则称 S 是 $[a, b]$ 上的完整直交函数系.

“完整性”的意义很直观: 直交函数系 S 中包含了所有的直交函数, 无一漏掉.

再回到三角级数上来,与(实)三角级数(4.3.2)相对应,我们称函数项级数

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx} \quad (4.3.6)$$

为复三角级数,其中 c_k 为复数. 实三角级数的收敛性定义与普通函数项级数的收敛性定义相同. 对(4.3.6)的收敛性定义,我们作如下规定:级数(4.3.6)在点 x 收敛的充要条件是对称部分和序列

$$S_n(x) = \sum_{k=-n}^n c_k e^{ikx}$$

在点 x 处收敛.

4.3.2 傅里叶级数的定义

我们先证明称为欧拉—傅里叶(Euler—Fourier)公式的定理.

定理 4.3.1 (欧拉—傅里叶公式) 设级数(4.3.2)在 $[-\pi, \pi]$ 上一致收敛于函数 $f(x)$, 即

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (4.3.7)$$

在 $[-\pi, \pi]$ 上成立, 且(4.3.7)的右边在 $[-\pi, \pi]$ 上一致收敛, 则

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx \quad (k=0, 1, 2, \dots), \quad (4.3.8)$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx \quad (k=1, 2, \dots). \quad (4.3.9)$$

(4.3.8)与(4.3.9)式称为欧拉—傅里叶公式.

证 将(4.3.7)式在 $[-\pi, \pi]$ 上积分, 由于右边级数在 $[-\pi, \pi]$ 上一致收敛, 从而可以逐项积分, 并得到

$$\int_{-\pi}^{\pi} f(x) dx = a_0 \pi + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos nx dx + b_n \int_{-\pi}^{\pi} \sin nx dx \right) = a_0 \pi,$$

$$\text{故 } a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx.$$

将(4.3.7)式两边乘以 $\cos nx$ ($n \geq 1$), 然后在 $[-\pi, \pi]$ 上积分, 由于级数(4.3.7)在 $[-\pi, \pi]$ 上一致收敛, 从而仍然可以逐项积分, 应用三角函数系的直交性得

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) dx &= \frac{a_0}{2} \int_{-\pi}^{\pi} \cos nx dx + \sum_{k=1}^{\infty} \left(a_k \int_{-\pi}^{\pi} \cos kx \cos nx dx \right) + \\ &\quad \sum_{k=1}^{\infty} \left(b_k \int_{-\pi}^{\pi} \sin kx \cos nx dx \right) = a_n \pi. \end{aligned}$$

$$\text{故 } a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx. \text{ 类似地, 将(4.3.7)的两边乘以 } \sin nx \quad (n \geq 1),$$

而后在 $[-\pi, \pi]$ 上积分, 得到 $b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx$.

定义 4.3.3 (傅里叶系数、傅里叶级数) 设函数 $f(x)$ 在区间 $[-\pi, \pi]$ 上可积, 由 (4.3.8)、(4.3.9) 确定的 a_n 、 b_n 称为函数 $f(x)$ 的 (实) 傅里叶系数, 或称其为 $f(x)$ 的有限傅里叶变换. 由傅里叶系数确定的三角级数 (4.3.2) 称为函数 $f(x)$ 的傅里叶级数, 记为

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (4.3.10)$$

在 (4.3.10) 的右边应用公式

$$\cos y = \frac{1}{2}(e^{iy} + e^{-iy}), \quad \sin y = \frac{-i}{2}(e^{iy} - e^{-iy}),$$

化简可得

$$\begin{aligned} f(x) &\sim \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k}{2}(e^{ikx} + e^{-ikx}) + \frac{-ib_k}{2}(e^{ikx} - e^{-ikx}) \right] \\ &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(\frac{a_k - ib_k}{2} e^{ikx} + \frac{a_k + ib_k}{2} e^{-ikx} \right). \end{aligned}$$

记 $c_0 = \frac{a_0}{2}$, $c_k = \frac{a_k - ib_k}{2}$, $c_{-k} = \frac{a_k + ib_k}{2}$ ($k=1, 2, \dots$), 则得

$$f(x) \sim c_0 + \sum_{k=1}^{\infty} (c_k e^{ikx} + c_{-k} e^{-ikx}) \equiv \sum_{k=-\infty}^{\infty} c_k e^{ikx}. \quad (4.3.11)$$

称 (4.3.11) 式为函数 $f(x)$ 的复数形式的傅里叶级数. 这时对 $k \geq 1$, 利用 (4.3.8)、(4.3.9) 可得

$$\begin{aligned} c_k &= \frac{a_k - ib_k}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx - \frac{i}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos kx - i \sin kx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx, \\ c_{-k} &= \frac{a_k + ib_k}{2} = \frac{1}{2} \left(\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx + \frac{i}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos kx + i \sin kx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{ikx} dx, \end{aligned}$$

而 $c_0 = \frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx$. 从而对任何 $k \in \mathbb{Z}$, 有

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx. \quad (4.3.12)$$

由 (4.3.12) 式确定的 c_k 称为 $f(x)$ 的复傅里叶系数.

现在, 我们指出两点: 第一, 实函数 $f(x)$ 具有实、复两种形式的傅里叶级数 (4.3.10) 与 (4.3.11), 但它们可以互换. 其实, 我们已经从 $f(x)$ 的 (实) 傅里叶

级数(4.3.10)出发得到了(4.3.11)与(4.3.12);反之,若从 $f(x)$ 的复数形式的傅里叶级数(4.3.11)与(4.3.12)出发,只要令

$$a_0 = 2c_0, \quad a_k = c_k + c_{-k}, \quad b_k = i(c_k - c_{-k}) \quad (k=1, 2, \dots),$$

并利用欧拉公式

$$e^{\pm i k x} = \cos kx \pm i \sin kx \quad (k=1, 2, \dots),$$

即可得(4.3.10)、(4.3.8)与(4.3.9).

第二,在一定条件下, $f(x)$ 的三角级数展式就是它的傅里叶级数. 例如,定理 4.3.1 中的 $f(x)$ 的三角级数展式就是 $f(x)$ 的(实)傅里叶级数.

关于傅里叶系数的性态,我们有黎曼—勒贝格引理.

定理 4.3.2 (黎曼—勒贝格引理) 设 $f(x)$ 在 $[a, b]$ 上黎曼可积(或反常积分绝对可积), $a > 0$, 则

$$\lim_{a \rightarrow +\infty} \int_a^b f(x) \cos ax dx = 0, \quad \lim_{a \rightarrow +\infty} \int_a^b f(x) \sin ax dx = 0.$$

证明从略. 当 $f(x)$ 连续可微、正常可积时,容易证明这一定理.

由黎曼—勒贝格引理知, $[-\pi, \pi]$ 上正常可积的函数 $f(x)$ 的傅里叶系数当 $k \rightarrow \infty$ 时, $a_k \rightarrow 0, b_k \rightarrow 0, c_k \rightarrow 0$.

此外,我们还注意到,函数 $f(x)$ 的傅里叶级数由其傅里叶系数完全确定,而 $f(x)$ 的傅里叶系数是由 $[-\pi, \pi]$ 上的积分表示的,它依赖于 $f(x)$ 在 $[-\pi, \pi]$ 上几乎所有点处的函数值,在有限多个点上改变 $f(x)$ 的函数值不会影响它的傅里叶系数,从而也不会影响它的傅里叶级数.

4.3.3 傅里叶级数的逐点收敛定理

给定函数 $f(x)$, 它的傅里叶级数是否收敛? 如果收敛,是否收敛于该函数 $f(x)$? 这是傅里叶级数理论中要讨论的基本问题. 这里仅介绍一个著名的收敛性定理——狄利克雷定理,证明从略.

定理 4.3.3 (狄利克雷定理) 设函数 $f(x)$ 在 $[-\pi, \pi]$ 上除有限多个第一类间断点外处处连续,且至多有有限个极值点(即分段单调),则 $f(x)$ 的傅里叶级数在 $(-\infty, +\infty)$ 上处处收敛,其和以 2π 为周期,在 $[-\pi, \pi]$ 上可表示为

$$S(x) = \begin{cases} f(x), & x \in (-\pi, \pi) \text{ 且为 } f(x) \text{ 的连续点;} \\ \frac{1}{2}[f(x-0) + f(x+0)], & x \text{ 为 } f(x) \text{ 的间断点;} \\ \frac{1}{2}[f(-\pi+0) + f(\pi-0)], & x = \pm\pi. \end{cases}$$

例 5 设 $f(x) = |x|, x \in (-\pi, \pi)$, 试求函数 $f(x)$ 的傅里叶级数及其和函数.

解 由于在有限多个点上改变 $f(x)$ 的函数值, 不会影响其傅里叶级数, 我们可设 $f(x) = |x|, x \in [-\pi, \pi]$.

由于 $f(x)$ 为偶函数, 所以 $f(x)\sin kx$ 为奇函数, 从而

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx = 0 \quad (k=1, 2, \dots).$$

又

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} |x| dx = \frac{2}{\pi}; \\ a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \cos kx dx = \frac{2}{\pi} \int_0^{\pi} x \cos kx dx \\ &= \frac{2}{k\pi} x \cos kx \Big|_0^{\pi} - \frac{2}{k\pi} \int_0^{\pi} \sin kx dx \\ &= \frac{2}{k^2\pi} (\cos k\pi - 1) = \begin{cases} -\frac{4}{\pi k^2}, & \text{当 } k \text{ 为奇数时;} \\ 0, & \text{当 } k \text{ 为偶数时.} \end{cases} \end{aligned}$$

于是 $f(x)$ 的傅里叶级数是

$$f(x) \sim \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2}.$$

再由狄利克雷定理知

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2} = |x| \quad (x \in [-\pi, \pi]).$$

例 6 设 $f(x) = x (x \in (-\pi, \pi))$, 试求 $f(x)$ 的傅里叶级数及其和函数.

解 同例 4, 可考虑 $f(x) = x (x \in [-\pi, \pi])$. 由于 $f(x)$ 为奇函数, 所以 $f(x)\cos kx$ 为奇函数, 故傅里叶系数中 $a_k = 0 (k=0, 1, 2, \dots)$. 又

$$\begin{aligned} b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin kx dx = \frac{2}{\pi} \int_0^{\pi} x \sin kx dx \\ &= \frac{2}{k\pi} \left[-x \cos kx \Big|_0^{\pi} + \frac{1}{k} \sin kx \Big|_0^{\pi} \right] = -\frac{2}{k} (-1)^k \quad (k=1, 2, \dots). \end{aligned}$$

于是所求傅里叶级数为

$$f(x) \sim \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin kx.$$

再由狄利克雷定理知

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin kx = \begin{cases} x, & -\pi < x < \pi, \\ 0, & x = \pm\pi. \end{cases}$$

在上面的例 4 与例 5 中, 函数 $f(x)$ 分别是偶函数与奇函数, 它们的傅里叶级数展式中只含余弦项, 或只含正弦项, 这样的傅里叶级数分别称为余弦级数与正弦级数.

有时候,函数 $f(x)$ 只在 $[0, \pi]$ 上定义,满足狄利克雷收敛定理的条件,我们可在区间 $[-\pi, 0]$ 上将函数补充定义,使 $f(x)$ 在 $[-\pi, \pi]$ 上有定义,且为偶函数(或奇函数),称为偶延拓(或奇延拓),则可将函数 $f(x)$ 在 $[0, \pi]$ 上展为余弦级数(或正弦级数).

下面再举一个求复数形式的傅里叶级数的例子.

例 7 设 $f(x)$ 在 $[-\pi, \pi]$ 上的表达式是

$$f(x) = \begin{cases} E, & 0 < x < \pi, \\ 0, & x = 0, \pm \pi, \\ -E, & -\pi < x < 0, \end{cases}$$

并设 $f(x)$ 具有周期 2π . 试求 $f(x)$ 的复数形式的傅里叶级数及其和函数.

解 由(4.3.12), 计算出 $f(x)$ 的傅里叶系数 $c_0 = 0$. 对 $k \neq 0$, 有

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx = \frac{E}{2\pi} \left(\int_{-\pi}^0 -e^{-ikx} dx + \int_0^{\pi} e^{-ikx} dx \right) \\ &= \frac{E}{2\pi} \left[\frac{1}{ik} e^{-ikx} \Big|_{-\pi}^0 + \left(-\frac{1}{ik} \right) e^{-ikx} \Big|_0^{\pi} \right] \\ &= \frac{E}{2k\pi i} [(1 - (-1)^k) - ((-1)^k - 1)] \\ &= \begin{cases} 0, & k \text{ 为偶数,} \\ \frac{2E}{k\pi i}, & k \text{ 为奇数,} \end{cases} \end{aligned}$$

从而得 $f(x)$ 的复数形式的傅里叶级数为

$$f(x) \sim \frac{2E}{\pi i} \sum_{k=-\infty}^{\infty} \frac{1}{2k-1} e^{i(2k-1)x}.$$

由于复数形式的傅里叶级数与实傅里叶级数收敛于相同的和, 且 $f(x)$ 与实、复傅里叶级数均具有周期 2π , 从而由狄利克雷定理得

$$f(x) = \frac{2E}{\pi i} \sum_{k=1}^{\infty} \frac{1}{2k-1} e^{i(2k-1)x}.$$

4.3.4 任意有限区间上的傅里叶级数

设函数 $f(x)$ 在 $[-l, l]$ 上可积, 则通过简单的变量代换可以得到在 $[-l, l]$ 上的实傅里叶级数

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{k\pi x}{l} + b_k \sin \frac{k\pi x}{l} \right), \quad (4.3.13)$$

其中

$$a_k = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{k\pi x}{l} dx \quad (k=0, 1, 2, \dots),$$

$$b_k = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{k\pi x}{l} dx \quad (k=1, 2, \dots). \quad (4.3.14)$$

与它的复数形式的傅里叶级数

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{i \frac{k\pi}{l} x}, \quad (4.3.15)$$

其中

$$c_k = \frac{1}{2l} \int_{-l}^l f(x) e^{-i \frac{k\pi}{l} x} dx, \quad k=0, \pm 1, \pm 2, \dots. \quad (4.3.16)$$

事实上, 令 $t = \frac{\pi x}{l}$, 则 $-l \leq x \leq l$ 化为 $-\pi \leq t \leq \pi$; 令 $F(t) = f\left(\frac{l}{\pi} t\right)$, 则函数 $F(t)$ 在 $[-\pi, \pi]$ 上有定义且可积, 其实傅里叶级数为

$$F(t) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt);$$

其复数形式的傅里叶级数为

$$F(t) \sim \sum_{k=-\infty}^{\infty} c_k e^{-i k t}.$$

将变量 t 换回到 x 即得(4.3.13)与(4.3.15). 其系数公式

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \cos kt dt, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} F(t) \sin kt dt,$$

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(t) e^{-i k t} dt,$$

在换元变换 $t = \frac{\pi x}{l}$ 下即为公式(4.1.14)与(4.1.16).

当函数 $f(x)$ 在 $[-l, l]$ 上满足狄利克雷收敛定理的条件时, $F(t)$ 在 $[-\pi, \pi]$ 上也满足狄利克雷收敛定理条件. 从而狄利克雷定理的收敛结果也适用于 $[-l, l]$ 上的 $f(x)$, 即若 $f(x)$ 在 $[-l, l]$ 上满足狄利克雷收敛定理的条件, 则 $f(x)$ 在 $[-l, l]$ 上的傅里叶级数在 $(-\infty, +\infty)$ 上处处收敛, 其和函数 $s(x)$ 以 $2l$ 为周期, 在 $[-l, l]$ 上可表示为

$$s(x) = \begin{cases} f(x), & x \text{ 为 } f(x) \text{ 的连续点,} \\ \frac{1}{2} [f(x-0) + f(x+0)], & x \text{ 为 } f(x) \text{ 的间断点,} \\ \frac{1}{2} [f(-l+0) + f(l-0)], & x = \pm l. \end{cases}$$

例 8 设 $f(x) = x^2, x \in [-1, 1]$, 试求 $f(x)$ 的傅里叶级数与其和函数.

解 由于 $f(x)$ 为偶函数, 故其傅里叶系数中 $b_k = 0 \quad (k=1, 2, \dots)$; 又 $l=1$, 所以

$$a_0 = \int_{-1}^1 x^2 dx = \frac{2}{3},$$

$$a_k = \int_{-1}^1 x^2 \cos k\pi x dx = (-1)^k \frac{4}{(k\pi)^2} \quad (k=1, 2, 3, \dots).$$

于是所求的傅里叶级数与和函数为

$$\frac{1}{3} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} (-1)^k \frac{1}{k^2} \cos k\pi x = x^2 \quad (x \in [-1, 1]).$$

例 9 把图 4.3.1 所示的宽为 τ , 高为 h , 周期为 T 的脉冲函数展为复数形式的傅里叶级数.

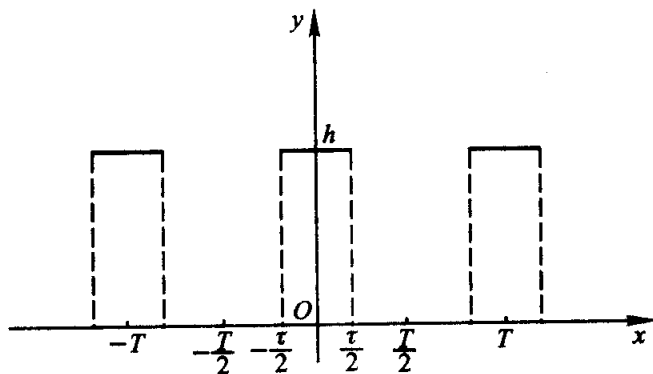


图 4.3.1

解 在区间 $[-\frac{T}{2}, \frac{T}{2}]$ 上函数的表达式为

$$f(t) = \begin{cases} 0, & \frac{T}{2} \leq t \leq -\frac{\tau}{2}, \text{ 或 } \frac{\tau}{2} < t \leq \frac{T}{2}, \\ h, & -\frac{\tau}{2} < t \leq \frac{\tau}{2}. \end{cases}$$

因 $l = \frac{T}{2}$, 由公式(4.3.16)计算复傅里叶系数为

$$c_0 = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) dt = \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} h dt = \frac{h\tau}{T},$$

$$\begin{aligned} c_k &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-i\frac{2\pi}{T}kt} dt = \frac{1}{T} \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} h e^{-i\frac{2\pi}{T}kt} dt \\ &= -\frac{h}{T} \frac{T}{i2\pi k} e^{-i\frac{2\pi}{T}kt} \Big|_{-\tau/2}^{\tau/2} = \frac{h}{k\pi} \sin \frac{k\pi\tau}{T} \quad (k = \pm 1, \pm 2, \dots). \end{aligned}$$

于是所求的复数形式的傅里叶级数为

$$f(t) \sim \frac{h\tau}{T} + \frac{h}{\pi} \sum_{k=-\infty, k \neq 0}^{\infty} \frac{1}{k} \sin \frac{k\pi\tau}{T} \exp \frac{i2\pi k}{T} t.$$

4.3.5 傅里叶级数的逐项求积与逐项求导数

对每个可积函数 $f(x)$ 都能形式上作出其傅里叶级数

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (4.3.17)$$

我们可以不管级数(4.3.17)是否收敛,当然更不谈一致收敛,只要 $f(x)$ 满足很少的条件,就可以进行逐项积分,且有下列的逐项求积分定理(证明从略)

定理 4.3.4 (傅里叶级数的逐项求积公式) 设 $f(x)$ 在 $[-\pi, \pi]$ 上黎曼可积, (4.3.17) 式是 $f(x)$ 的傅里叶级数, 则

$$\frac{a_0}{2}x + \sum_{k=1}^{\infty} \left(\frac{a_k}{k} \sin kx + \frac{b_k}{k} (1 - \cos kx) \right),$$

恒收敛, 且其和函数在 $[-\pi, \pi]$ 上恰为 $\int_0^x f(t)dt$, 亦即

$$\begin{aligned} \int_0^x f(t)dt &= \int_0^x \frac{a_0}{2}dt + \sum_{k=1}^{\infty} \int_0^x (a_k \cos kt + b_k \sin kt)dt \\ &= \frac{a_0}{2}x + \sum_{k=1}^{\infty} \left(\frac{a_k}{k} \sin kx + \frac{b_k}{k} (1 - \cos kx) \right), \quad x \in [-\pi, \pi]. \end{aligned}$$

例 10 把 $f(x) = x^2$ ($x \in [-\pi, \pi]$) 展开成傅里叶级数, 求其和; 并由此求级数 $\sum_{k=1}^{\infty} \frac{1}{k^2}$ 的和.

解 由例 5, 我们已求得

$$x = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} \sin kx \quad (x \in [-\pi, \pi]).$$

如果对上式从 0 到 x , $x \in [-\pi, \pi]$, 逐项积分, 有

$$\begin{aligned} \frac{x^2}{2} &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k^2} (-\cos kx + 1) \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k^2} + \sum_{k=1}^{\infty} (-1)^k \frac{2}{k^2} \cos kx \quad (x \in [-\pi, \pi]), \end{aligned}$$

于是

$$x^2 = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{4}{k^2} + \sum_{k=1}^{\infty} (-1)^k \frac{4}{k^2} \cos kx \quad (x \in [-\pi, \pi]).$$

这里 $\sum_{k=1}^{\infty} (-1)^{k+1} \frac{4}{k^2} = \frac{a_0}{2} = \frac{1}{2} \cdot \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{\pi^2}{3}$, 从而有

$$x^2 = \frac{\pi^2}{3} + \sum_{k=1}^{\infty} (-1)^k \frac{4}{k^2} \cos kx \quad (x \in [-\pi, \pi]).$$

在上式两端中令 $x = \pi$, 得

$$\pi^2 = \frac{\pi^2}{3} + \sum_{k=1}^{\infty} \frac{4}{k^2},$$

从而得

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

关于傅里叶级数的逐项求导数,我们给出下面的定理,证明从略.

定理 4.3.5 (傅里叶级数逐项求导公式) 设 $f(x)$ 在 $[-\pi, \pi]$ 上连续, $f'(x)$ 除有限个点外在 $[-\pi, \pi]$ 上处处存在且正常可积, 以及 $f(\pi) = f(-\pi)$, 则 $f'(x)$ 有傅里叶级数

$$f'(x) \sim \sum_{k=1}^{\infty} (kb_k \cos kx - ka_k \sin kx),$$

亦即它由 $f'(x)$ 的傅里叶级数逐项求导数而得.

4.3.6 傅里叶级数的极值性质与贝塞尔不等式

先考察较一般的情形.

定义 4.3.4 (广义傅里叶系数、广义傅里叶级数) 设

$$S = \{\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x), \dots\}$$

是定义 4.3.1 中给出的 $[a, b]$ 上的标准直交函数系, 即满足

$$\int_a^b \varphi_n(x) \overline{\varphi_m(x)} dx = \begin{cases} 0, & n \neq m, \\ 1, & n = m. \end{cases}$$

对实广义可积函数 $f(x)$, 称

$$c_k = \int_a^b f(x) \varphi_k(x) dx \quad (4.3.18)$$

为 $f(x)$ 的广义傅里叶系数, 且称级数 $\sum_{k=0}^{\infty} c_k \varphi_k(x)$ 为 $f(x)$ 的广义傅里叶级数, 记为

$$f(x) \sim \sum_{k=0}^{\infty} c_k \varphi_k(x), \quad (4.3.19)$$

这个级数的部分和记为

$$S_n(x) = \sum_{k=0}^n c_k \varphi_k(x). \quad (4.3.20)$$

此外, 我们还引入 n 次“ φ -多项式” $t_n(x)$

$$t_n(x) = \sum_{k=0}^n r_k \varphi_k(x), \quad (4.3.21)$$

其中 r_k 是一些已知的实数或复数.

定义 4.3.5 (均方误差) 设 $f(x)$ 是 $[a, b]$ ($b > a$) 上的可积且平方可积的实函数, 我们称

$$\delta_n^2 = \frac{1}{b-a} \int_a^b |f(x) - t_n(x)|^2 dx \quad (4.3.22)$$

为用 n 次 φ -多项式 $t_n(x)$ 逼近 $f(x)$ 的均方误差, 简称均方差.

现在讨论 $t_n(x)$ 中的系数 r_k 取何值时, 使 δ_n^2 为最小. 由于

$$\begin{aligned}
 (b-a)\delta_n^2 &= \int_a^b \left(f(x) - \sum_{k=0}^n r_k \varphi_k(x) \right) \left(f(x) - \sum_{k=0}^n \bar{r}_k \overline{\varphi_k(x)} \right) dx \\
 &= \int_a^b f^2(x) dx - \sum_{k=0}^n \bar{r}_k \int_a^b f(x) \overline{\varphi_k(x)} dx - \\
 &\quad \sum_{k=0}^n r_k \int_a^b f(x) \varphi_k(x) dx + \sum_{j,k=0}^n r_k \bar{r}_j \int_a^b \varphi_k(x) \overline{\varphi_j(x)} dx,
 \end{aligned}$$

据 $\varphi_k(x)$ 的标准直交性, 上式最后一项的值等于

$$\sum_{k=0}^n r_k \bar{r}_k = \sum_{k=0}^n |r_k|^2.$$

再由(4.3.18)得

$$\int_a^b f(x) \overline{\varphi_k(x)} dx = c_k,$$

与

$$\int_a^b f(x) \varphi_k(x) dx = \bar{c}_k.$$

于是有

$$(b-a)\delta_n^2 = \int_a^b f^2(x) dx - \sum_{k=0}^n (\bar{r}_k c_k + r_k \bar{c}_k) + \sum_{k=0}^n |r_k|^2.$$

又因

$$|r_k - c_k|^2 = |r_k|^2 + |c_k|^2 - (r_k \bar{c}_k + \bar{r}_k c_k),$$

最终有

$$(b-a)\delta_n^2 = \int_a^b f^2(x) dx - \sum_{k=0}^n |c_k|^2 + \sum_{k=0}^n |r_k - c_k|^2. \quad (4.3.23)$$

至此, 我们可得到

定理 4.3.6 (均方差取最小的充要条件) 使均方差 δ_n^2 取最小值的充要条件是(4.3.23)中最后的和式为 0

$$r_k = c_k \quad (k=0, 1, \dots, n).$$

且 δ_n^2 的最小值是

$$E_n = \frac{1}{b-a} \left(\int_a^b f^2(x) dx - \sum_{k=0}^n |c_k|^2 \right) = \frac{1}{b-a} \int_a^b |f(x) - s_n(x)|^2 dx.$$

由于 $E_n \geq 0$, 因此对任何 $n=0, 1, 2, \dots$, 有

$$\sum_{k=0}^n |c_k|^2 \leq \int_a^b f^2(x) dx.$$

从而

$$\sum_{k=0}^{\infty} |c_k|^2 \leq \int_a^b f^2(x) dx. \quad (4.3.24)$$

(4.3.24)称为(广义)贝塞尔(Bessel)不等式.

由(4.3.24)推得,级数 $\sum_{k=0}^{\infty} |c_k|^2$ 收敛,从而 $\lim_{k \rightarrow \infty} c_k = 0$,即 $f(x)$ 的广义傅里叶系数当 $k \rightarrow \infty$ 时趋于 0.

广义贝塞尔不等式是关于一般标准直交函数系的广义傅里叶系数的,在下面的例 10 中,我们将推导关于实傅里叶系数的贝塞尔不等式.

例 11 设 $f(x)$ 是 $[-\pi, \pi]$ 上的实黎曼可积函数(从而必平方可积).我们取一般标准直交系 S 为三角标准直交系

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}},$$

$$\varphi_{2k-1}(x) = \frac{1}{\sqrt{\pi}} \cos kx, \quad \varphi_{2k}(x) = \frac{1}{\sqrt{\pi}} \sin kx \quad (k=1, 2, \dots).$$

于是, $f(x)$ 的广义傅里叶级数 $f(x) \sim \sum_{k=0}^{\infty} c_k \varphi_k(x)$ 中的 c_k 成为

$$c_0 = \int_{-\pi}^{\pi} f(x) \cdot \frac{1}{\sqrt{2\pi}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) dx;$$

$$c_{2k-1} = \int_{-\pi}^{\pi} f(x) \cdot \frac{1}{\sqrt{\pi}} \cos kx dx = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos kx dx;$$

$$c_{2k} = \int_{-\pi}^{\pi} f(x) \cdot \frac{1}{\sqrt{\pi}} \sin kx dx = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin kx dx.$$

另一方面, $f(x)$ 的傅里叶级数为

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

并且 c_k 与 a_k, b_k 之间有如下关系

$$c_0 = \sqrt{\frac{\pi}{2}} a_0; \quad c_{2k-1} = \sqrt{\pi} a_k; \quad c_{2k} = \sqrt{\pi} b_k;$$

以及

$$|c_0|^2 = \frac{\pi a_0^2}{2}; \quad |c_{2k}|^2 = \pi b_k^2; \quad |c_{2k-1}|^2 = \pi a_k^2.$$

于是,由于 $\sum_{k=0}^{\infty} |c_k|^2$ 收敛,所以

$$\sum_{k=0}^{\infty} |c_k|^2 = \pi \left(\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \right).$$

从而,由(4.3.24)得实傅里叶系数的贝塞尔不等式

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx. \quad (4.3.25)$$

例 12 考察 $[-\pi, \pi]$ 上的复标准直交指数函数系 $S = \{\varphi_k(x)\}_{k=-\infty}^{\infty}$, 其中

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}, \quad \varphi_{2k+1}(x) = \frac{1}{\sqrt{2\pi}} e^{-ikx}.$$

设 $f(x)$ 是 $[-\pi, \pi]$ 上的实正常可积函数. 按照例 10 类似的方法, 便可以得到关于复傅里叶系数的贝塞尔不等式

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \frac{1}{2\pi} \int_a^b f^2(x) dx. \quad (4.3.26)$$

我们指出, 对于 $[-\pi, \pi]$ 上的实平方可积函数 $f(x)$, (4.3.24)、(4.3.25) 与 (4.3.26) 分别是关于一般标准直交函数系、实标准直交三角函数系、复标准直交指数函数系的傅里叶系数的贝塞尔不等式. 其中的等式成立的条件可叙述为下面的定理, 证明从略.

定理 4.3.7 (帕塞瓦尔 (Parseval) 等式) 设 $f(x)$ 是 $[-\pi, \pi]$ 上的平方可积的实函数, $f(x)$ 的实傅里叶级数与复数形式的傅里叶级数分别为

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

与

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx}.$$

则有下列帕塞瓦尔等式

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx, \quad (4.3.27)$$

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(x) dx, \quad (4.3.28)$$

亦即, 当 $n \rightarrow \infty$ 时, 有

$$\begin{aligned} E_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \frac{a_0}{2} - \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \right|^2 dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(x) - \sum_{k=-n}^n c_k e^{ikx} \right|^2 dx \rightarrow 0. \end{aligned}$$

习题

1. 试证函数系 $\left\{1, \cos \frac{n\pi}{l}x\right\}, \left\{\sin \frac{n\pi}{l}x\right\}_{n=1}^{\infty}$ 都是 $[0, l]$ 上的直交系.

2. 设拉德玛赫 (Rademarch) 函数系 $S = \{\varphi_0(x), \varphi_1(x), \dots\}$ 中每个函数具有周期 1, 在 $[0, 1]$ 上的定义

$$\varphi_0(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2}; \\ -1, & \frac{1}{2} \leq x < 1. \end{cases}$$

$$\varphi_n(x) = \varphi_0(2^n x) \quad (n=1, 2, \dots).$$

试作出 $\varphi_0(x), \varphi_1(x), \varphi_2(x)$ 的图形, 并证明 S 是 $[0, 1]$ 上的标准直交系.

3. 求极限 $\lim_{a \rightarrow +\infty} \int_{-\pi}^{\pi} f(x) \sin^2 ax dx$, 其中 $f(x)$ 在 $[-\pi, \pi]$ 上正常可积, 且 $\int_{-\pi}^{\pi} f(x) dx = 2$.

4. 将下列函数展为傅里叶级数, 并求和函数:

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

$$(2) f(x) = |\sin x|, \quad -\pi < x < \pi;$$

$$(3) f(x) = \begin{cases} 1, & 0 < x < \pi, \\ 0, & -\pi < x < 0. \end{cases}$$

5. 求函数 $f(x) = \begin{cases} 1, & 0 < x < \frac{\pi}{2}; \\ 0, & \frac{\pi}{2} < x < \pi \end{cases}$ 的正弦级数展式与和函数.

6. 求函数 $f(x) = \frac{1}{2}(\pi - x) \sin x, 0 \leq x \leq \pi$ 的余弦级数展式与和函数.

7. 求函数 $f(x) = \begin{cases} x, & 0 \leq x \leq 1; \\ 2-x, & 1 < x \leq 2. \end{cases}$ 的正弦级数与和函数.

8. 求函数 $f(x) = e^{ax} (a \neq 0), -\pi \leq x \leq \pi$ 的复数形式的傅里叶级数与和函数.

第五章 微分方程与傅里叶分析

本章讲述微分方程的基本知识,并由解微分方程的一种方法出发,介绍傅里叶分析初步及其最新发展——小波分析. § 5.1 引进数学建模的概念,并给出建立微分方程的例;简单类型的常微分方程的基本解法;常微分方程的存在惟一性与简单的数值方法. § 5.2 是偏微分方程初步,包括一阶偏微分方程的首次积分法,三类基本的二阶线性偏微分方程——双曲、抛物、椭圆型方程及它们的物理背景;从微分方程的傅里叶方法出发,获得热传导方程的傅里叶级数解. § 5.3 致力于傅里叶变换的基本理论、谱分析及其在解二阶偏微分方程定解问题中的应用. § 5.4 是小波变换,介绍其优越性,以及它在谱分析、图像显示中的应用概述.

§ 5.1 微分方程

微分方程是数学科学的一个重要组成部分,它不仅有严格的数学理论,而且有广泛的实际背景和应用.微分方程是永远活跃在数学科学最前沿、最富有创造性和生命力的学科分支之一.

在微积分出现不久,人们就开始了微分方程的研究.由于许多自然规律的数学表现形式都是微分方程,因此从它诞生起就成为人类认识自然的有力工具,成为数学科学联系实际的主要渠道之一.早在十七、十八世纪,微分方程就作为牛顿力学的重要助手,在天体力学、流体力学与其他机械力学领域中显示了巨大的功能.数学推动科学的发展,在科学发展史上是屡见不鲜的.例如,早在海王星被实际观测到之前,这颗行星的存在就被天文学家用微分方程方法推算出来,于是激励着天文学家去观测、去研究、去发现,终于在数年后真正找到了海王星,成为科学史上的美谈.时至高科技如此飞速发展的今天,微分方程在电子技术、自动控制、星际航行、计算机科学等各种尖端领域中起着更加重要的作用.

5.1.1 建模实例,微分方程的通解、特解

为解决实际问题,必须首先把自然界中的各种现象与各种规律用数学形式表示出来,也就是必须“建立数学模型”,然后利用各种数学方法,并加上计算机的高速运算,最终得到问题的答案.我们看几个建立数学模型的实例.

1. 弹簧悬物振动方程

如图 5.1.1, 假设质量为 m 的重物悬挂在弹簧上, 并设弹簧的质量与所悬重物的质量相比可以忽略不计. 当重物静止不动时, 它所受到的两个力, 重力 mg 和弹簧的恢复力 F , 互相平衡. 如果把重物向下拉(或向上推)一小段距离, 然后放手. 试求在不同时刻重物的位置.

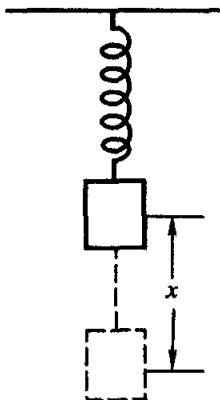


图 5.1.1

今取 x 轴为铅垂直线, 正方向向下, 取重物静止时的重心位置为 $x=0$, 在下拉(或上推)后开始振动. 此时重物受到三个力的作用:

① 重力 $F_1 = mg$, 方向向下;

② 弹簧恢复力 $F_2 = -mg - cx$, 其中 $c > 0$ 是弹簧的刚度, 即把它拉长一个单位长度所需的力, 此力的方向要看 $mg + cx$ 是正还是负而定. 若 $mg + cx > 0$, 弹簧的长度比没有悬挂重物时要长, 因此恢复力方向向上; $mg + cx < 0$ 时则恢复力方向向下;

③ 空气阻力 $F_3 = -\alpha \frac{dx}{dt}$, 这是由实验知道的, 空气阻力的大小与重物运动的速度成正比, 而方向与运动方向相反.

这样, 应用牛顿第二定律 $F = ma$ 就得到

$$m \frac{d^2 x}{dt^2} = F_1 + F_2 + F_3 = mg - (mg + cx) - \alpha \frac{dx}{dt} = -cx - \alpha \frac{dx}{dt},$$

亦即

$$m \frac{d^2 x}{dt^2} = -cx - \alpha \frac{dx}{dt}, \quad (5.1.1)$$

其中 $\alpha > 0$ 称为阻尼系数. (5.1.1) 是一个二阶线性常系数常微分方程, 它的解将在 § 5.2 中给出.

2. 弦振动方程

考察一根长为 l 的、两端固定的、拉紧的弦. 假设弦在未受到外力扰动时的平衡位置在 x 轴上, 并假设弦上各点在外力作用下所发生的位置变化是在 x, u 平面内垂直于 x 轴的方向上, 以 $u(x, t)$ 表示弦上点 x 在时刻 t 时的位置, 见图 5.1.2.

此外, 我们还假设

(1) 弦是柔软且有弹性的(即它不抗弯曲), 因此, 在任何时刻弦的张力总是沿着弦的切线方向.

(2) 弦只有微小的横向振动, 这并不是说

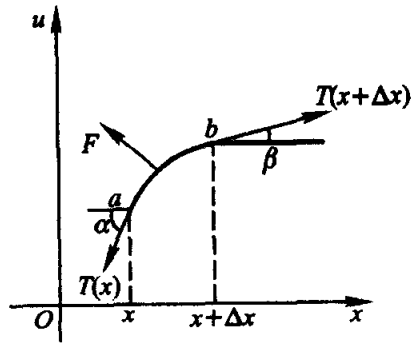


图 5.1.2

$u(x, t)$ 很小, 而是 $u_x(x, t)$ 很小, 从而 $u_{xx}^2(x, t)$ 可以忽略不计.

现在的问题是: 试求时刻 t 时的弦上各点的位移.

先列出作用在弦的一小段弧 ab 上的力:

- ① 作用在 a 点的张力 $T(x)$,
- ② 作用在 b 点的张力为 $T(x + \Delta x)$,
- ③ 作用在 ab 弧上的外力为 $F(\xi, t)\Delta x$,

这里 $F(\xi, t)$ 是在 $[x, x + \Delta x]$ 中的某点 ξ 处的力密度.

于是, 沿水平方向的分力是

$$\begin{aligned} T(x + \Delta x) \cos \beta - T(x) \cos \alpha &= T \left[\frac{1}{\sqrt{1 + \tan^2 \beta}} - \frac{1}{\sqrt{1 + \tan^2 \alpha}} \right] \\ &= T \left[\frac{1}{\sqrt{1 + u_x^2(b, t)}} - \frac{1}{\sqrt{1 + u_x^2(a, t)}} \right] = 0, \end{aligned}$$

沿垂直方向的分力是

$$T(x + \Delta x) \sin \beta - T(x) \sin \alpha + F(\xi, t) \Delta x \approx \rho \Delta x \frac{\partial^2 u(x, t)}{\partial t^2}.$$

由于振动微小, 故

$$\sin \alpha \approx \tan \alpha = \frac{\partial u(x, t)}{\partial x}, \quad \sin \beta \approx \tan \beta = \frac{\partial u(x + \Delta x, t)}{\partial x},$$

从而

$$\rho \Delta x \frac{\partial^2 u(x, t)}{\partial t^2} = T \frac{\partial^2 u(x + \tau \Delta x, t)}{\partial x^2} \Delta x + F(\xi, t) \Delta x \quad (0 < \tau < 1).$$

将上式两端除以 $\rho \Delta x$, 令 $\Delta x \rightarrow 0$, 此时有 $\xi \rightarrow x$, 取极限得

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad (5.1.2)$$

其中 $a^2 = \frac{T}{\rho}$, $f = \frac{F}{\rho}$. 方程(5.1.2)刻画了均匀弦的微小振动的一般规律, 通常称它为弦振动方程. 由于弦振动是波的传播, 因此方程(5.1.2)又称为(一维)波动方程. 它是二阶非齐次线性偏微分方程.

类似地, 可考虑平面的膜振动现象而得到二维波动方程

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + f(x, y, t).$$

当我们研究电磁波在空中传播时, 还可以得到三维波动方程

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) + f(x, y, z, t).$$

有关二阶偏微分方程的求解问题将在后行课《数学物理方程》中学习.

以上建立数学模型的过程中用到物理学、力学等方面的知识以及数学的常

规方法.下面我们再用热力学知识与积分知识来建立热的传导方程.

3. 热传导方程

现在研究质量均匀分布的物体的热传导问题.

设有一个温度不均匀分布的热物体(物体质量分布是均匀的),物体内部有密度为 $F(M, t)$ 的热源,这里 M 是物体上的一点.试建立物体内部热量的传导方程.

我们采用积分学中的“微元法”来建立该传导过程的数学模型.设物体 Ω 的质量分布是均匀的,假设其密度为 ρ 、比热为 c 、导热系数为 k ,均为常数, $u(M, t)$ 为 Ω 中的点 $M(x, y, z) \in \Omega$ 在时刻 t 的温度,见图 5.1.3.

任取一个体微元 dV ,其表面积记为 dS ,现在分析此微元 dV 内的热平衡情况,即热量的流入与温度升高.

① 流入 dV 的热量来自热的传导与热源

(i) 来自热传导的热量:由热力学中的傅里叶定律,热流强度与温度梯度成正比,即 $v = k \nabla u, k > 0$,于是,单位时间流入 dV 的总热量为

$$\theta_1 = \iint_S k \nabla u \cdot \mathbf{n} dS = \iiint_V k \operatorname{div} \nabla u dV,$$

$\mathbf{n}$ 为曲面 S 的外法向.

(ii) 来自热源的热量:

$$\theta_2 = \iiint_S F(M, t) dV.$$

② dV 温度升高所需热量 由热力学知,单位时间内温度升高所需热量为

$$\theta_3 = \iiint_S c\rho \frac{\partial u}{\partial t} dV.$$

于是,由热量守恒律 $\theta_1 + \theta_2 = \theta_3$ 得

$$\iiint_V k \operatorname{div} \nabla u dV + \iiint_S F(M, t) dV = \iiint_S c\rho \frac{\partial u}{\partial t} dV.$$

由微元的任意性,得

$$c\rho \frac{\partial u}{\partial t} = k \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) + F(M, t),$$

或

$$\frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) + f(x, y, z, t), \quad (5.1.3)$$

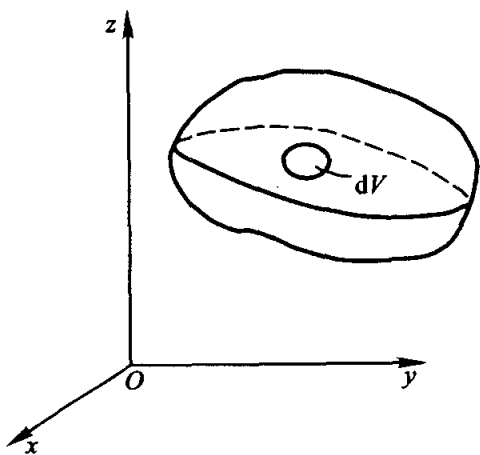


图 5.1.3

其中 $a^2 = \frac{k}{c\rho}$, $f(x, y, z, t) = \frac{F(M, t)}{c\rho}$. 方程(5.1.3)称为热传导方程,它是二阶非齐次线性偏微分方程.此方程不仅可以反映热的传导过程,而且还可以刻画扩散物质的浓度、粒子的扩散现象等物理过程.

还可以举许多数学建模的例,而且在今后的教学和科学研究过程中也是经常遇到的,当然要比现在举的例子复杂得多.

定义 5.1.1 (微分方程) 含有未知函数导数的方程称为微分方程.

微分方程的类型很多,并且可以从不同角度分类.

(1) 按照未知函数的自变量个数分类 未知函数的自变量个数为 1 的,称为常微分方程;多于一个自变量的,称为偏微分方程.

$$m \frac{d^2 x}{dt^2} = -cx - \alpha \frac{dx}{dt}, \quad \text{常微分方程.}$$

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad \text{偏微分方程.}$$

(2) 按照未知函数导数的阶数分类 未知函数导数的最高阶,称为微分方程的阶.

$$\frac{d^2 y}{dx^2} + \left(\frac{d^2 y}{dx^2} \right)^3 = -cx - \left(\alpha \frac{dy}{dx} \right)^4, \quad \text{二阶常微分方程.}$$

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^3 u}{\partial x^3} + b^2 \frac{\partial^2 u}{\partial y^2} + f(x, y, t), \quad \text{三阶偏微分方程.}$$

(3) 按照未知函数导数与未知函数的次数分类 未知函数导数与未知函数的次数都是 1 次的,称为线性微分方程,否则称为非线性微分方程.

$$\frac{\partial^5 u}{\partial t^5} = a^2 \frac{\partial^3 u}{\partial x^3} + \frac{\partial^2 u}{\partial z^2} + x + z + \sin t, \quad \text{五阶线性偏微分方程.}$$

$$\frac{d^3 y}{dx^3} + \left(\frac{d^2 y}{dx^2} \right)^3 = -ax^2 - \left(\alpha \frac{dy}{dx} \right)^4, \quad \text{三阶非线性常微分方程.}$$

(4) 按照未知函数的个数分类 未知函数个数为 1 的,称为方程,未知函数个数多于 1 的,称为方程组.

前面所列的都是微分方程,而

$$\begin{cases} s^2 \frac{d^2 x}{dt^2} + st \frac{dy}{dt} = sxy, \\ s^3 \left(\frac{d^3 x}{dt^3} \right)^2 + st \frac{dy}{dt} = sx^2 y^2 \end{cases}$$

是三阶非线性常微分方程组;类似地,也有偏微分方程组.

(5) 按照含自变量自身的函数与否分类 若有含自变量自身的函数的项,例如含有 $f(x, t)$ 的项,称为非齐次方程,否则称为齐次方程.

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad \text{二阶非齐次线性偏微分方程.}$$

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad \text{二阶齐次线性偏微分方程.}$$

定义 5.1.2 (常微分方程的通解、特解) 求常微分方程的解, 就是寻找那些函数, 使得函数连同其(各阶)导数满足所给的微分方程. 常微分方程的解可分为通解和特解. 通解是指微分方程的全部解的一个共同表达式, 其中包含一些任意常数, 一般情况下可以通过确定常数值而得到方程的每一个解; 特解是微分方程的满足给定条件的个别解.

例如, 方程

$$\frac{dy}{dx} + y = 0 \quad (5.1.4)$$

的通解为

$$y = y(x) = ce^{-x},$$

其中 c 为任意常数; 方程(5.1.4)的满足“初始条件” $y(0) = y_0$ 的特解为

$$y = y(x) = y_0 e^{-x}.$$

微分方程的通解有明显的几何意义. 我们以最简单的一阶常微分方程

$$\frac{dy}{dx} = f(x, y) \quad (5.1.5)$$

为例来说明.

把 x, y 看作平面上的直角坐标, 并设方程(5.1.5)中的函数 $f(x, y)$ 在平面内的某区域 D 中有定义. 在 D 中的每一点, 都可作一个以 $f(x, y)$ 在该点的值为斜率并一律指向 x 增加方向的有向线段(见图 5.1.4), 我们说在区域 D 上由方程(5.1.5)确定了一个“方向场”. 于是, 方程(5.1.5)的一个解 $y = y(x)$, 就是

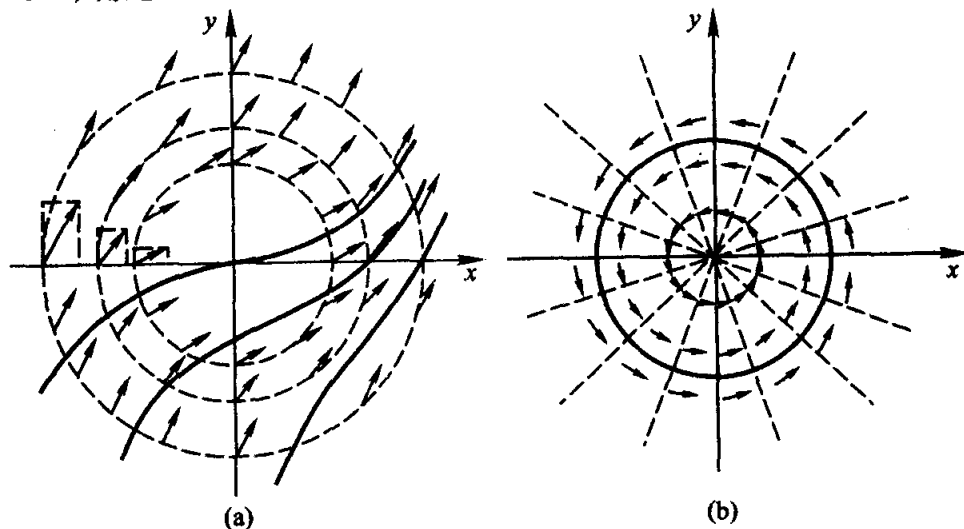


图 5.1.4

位于此方向场中的一条曲线,该曲线在所经过的每一点都与方向场在该点的方向相切.我们称这些曲线为微分方程的“积分曲线”.如果方程(5.1.5)的通解 $y = \varphi(x, C)$ 所对应的曲线记为 Γ_C , 当 C 变动时,得到一族“平行的曲线”,称为微分方程的“积分曲线族”.

例 1 求微分方程 $\frac{dy}{dx} = \frac{1}{x^2}$ 的解,并作其积分曲线.

解 将方程改写为

$$dy = \frac{dx}{x^2},$$

两边积分得 $y = -\frac{1}{x} + c$, 即为所求的通解. 积分曲线绘在图 5.1.4 中.

5.1.2 一阶常微分方程的几种特殊类型

对于形如

$$\frac{dy}{dx} = f(x, y)$$

的一阶常微分方程,以下省去“常”字,简称一阶微分方程,有几种特殊类型可以求得方程的通解.

1. 可按变量分离的情形

$f(x, y)$ 可按变量 x, y 分离成两个函数 $f(x, y) = \varphi(x)\psi(y)$. 若 $\varphi(x)\psi(y) \neq 0$, 则方程 $\frac{dy}{dx} = \varphi(x)\psi(y)$ 可改写为

$$\frac{dy}{\psi(y)} = \varphi(x)dx.$$

两边积分,得

$$\int \frac{dy}{\psi(y)} = \int \varphi(x)dx,$$

如果左、右分别各有原函数 $\Psi(y), \Phi(x)$, 则可得微分方程的隐函数解

$$\Psi(y) = \Phi(x) + C.$$

注意,当 $\varphi(x)\psi(y) = 0$ 时,方程可能有常数解.

例 2 求方程 $\frac{dy}{dx} = \frac{x}{y^2}$ 的通解.

解 把方程化为变量分离的形式

$$x dx = y^2 dy,$$

两边积分得

$$\frac{1}{2}x^2 + C = \frac{1}{3}y^3,$$

于是,得方程的含任意常数的通解

$$y(x) = \left(\frac{3}{2}x^2 + c \right)^{\frac{1}{3}}.$$

例 3 求方程 $\frac{dy}{dx} = \sqrt{1-y^2}$ 的通解.

解 首先设 $\sqrt{1-y^2} \neq 0$, 则可按以下步骤解方程

$$\frac{dy}{\sqrt{1-y^2}} = dx \Rightarrow \int \frac{dy}{\sqrt{1-y^2}} = x + C \Rightarrow \arcsin y = x + C.$$

最后得方程的解为

$$y(x) = \sin(x + C).$$

其次,当 $\sqrt{1-y^2} = 0$ 时, $y = \pm 1$, 验证后, $y = \pm 1$ 也是方程的解.

例 4 求方程 $\begin{cases} x(1-y^2)dx - y(1+x^2)dy = 0 \\ y(0) = 2 \end{cases}$ 的解.

解 对于 $1-y^2 \neq 0$, 我们有

$$\begin{aligned} x(1-y^2)dx - y(1+x^2)dy &= 0 \Rightarrow \frac{ydy}{1-y^2} = \frac{x dx}{1+x^2} \\ &\Rightarrow \ln(1+x^2) + \ln|1-y^2| = \ln|C|, \end{aligned}$$

得通解

$$y(x) = \pm \sqrt{1 - \frac{c}{\sqrt{1+x^2}}}.$$

验证后, 知 $y = \pm 1$ 也是方程的解.

为确定满足 $y(0) = 2$ 的特解, 也就是求通过点 $(0, 2)$ 的积分曲线, 我们可确定通解中的常数 $C = -3$, 故特解为

$$y(x) = \pm \sqrt{1 + \frac{3}{\sqrt{1+x^2}}}.$$

例 5 求方程 $\frac{dy}{dx} = (x+y)^2$ 的解.

解 令 $w = x + y$, 则由 $\frac{dy}{dx} = \frac{dw}{dx} - 1$, 我们有

$$\begin{aligned} \frac{dw}{dx} - 1 &= w^2 \Rightarrow \frac{dw}{w^2 + 1} = dx \Rightarrow \arctan w = x + C \\ &\Rightarrow w = \tan(x + C) \Rightarrow y(x) = \tan(x + C) - x. \end{aligned}$$

例 6 求方程 $\frac{dx}{dy} = \frac{xy^3}{2y^4 + x^4}$ 的解.

解 令 $x = yu$, 则由 $dx = u dy + y du$, 原方程可按下面路线求解

$$\frac{dx}{dy} = \frac{xy^3}{2y^4 + x^4} \Rightarrow u + y \frac{du}{dy} = \frac{u}{2 + u^4} \Rightarrow \frac{1}{y} dy + \frac{2 + u^4}{u + u^5} du = 0,$$

积分后得到

$$y^4 = c \frac{1 + u^4}{u^8},$$

于是方程的隐式解为

$$x^4 + y^4 = cx^8.$$

注 有些方程不是一下就能看出类型,但在经过变换后却可化为变量分离的情形.特别地,如果方程是 $\frac{dy}{dx} = \varphi\left(\frac{y}{x}\right)$ 形式,则可以化为变量分离的情形.

2. 一阶线性常微分方程的一般形式

$$\frac{dy}{dx} + p(x)y = q(x), \quad \text{一阶非齐次线性微分方程,} \quad (5.1.6)$$

$$\frac{dy}{dx} + p(x)y = 0, \quad \text{一阶齐次线性微分方程,} \quad (5.1.7)$$

这里假设 $p(x)$ 、 $q(x)$ 为 x 的连续函数.

容易看出,一阶齐次线性微分方程(5.1.7)是可以分离变量的,因此

$$\frac{dy}{dx} + p(x)y = 0 \Rightarrow \frac{dy}{y} + p(x)dx = 0 \Rightarrow \ln|y| + \int p(x)dx = 0,$$

于是得到一阶齐次线性微分方程(5.1.7)的通解

$$y(x) = C \exp \left\{ - \int p(x)dx \right\}, \quad (5.1.8)$$

其中 C 为常数.

为求非齐次线性微分方程(5.1.6)的解,我们介绍“常数变易法”.

常数变易法的思路是:把齐次方程(5.1.7)的通解(5.1.8)中的常数 C 看作 x 的函数 $C(x)$,确定 $C(x)$,使得

$$y(x) = C(x) \exp \left\{ - \int p(x)dx \right\} \quad (5.1.9)$$

是非齐次方程(5.1.6)的解.

为此,把(5.1.9)代入非齐次方程,得到

$$\frac{dy}{dx} = \frac{dC(x)}{dx} \exp \left\{ - \int p(x)dx \right\} - C(x)p(x) \exp \left\{ - \int p(x)dx \right\}$$

从而

$$C(x) = \int q(x) \exp \left\{ \int p(x)dx \right\} dx + C_1, \quad (5.1.10)$$

其中 C_1 是任意常数.将(5.1.10)代入(5.1.9),得非齐次方程的通解

$$y(x) = \exp \left\{ - \int p(x)dx \right\} \cdot \left[\int q(x) \exp \left\{ \int p(x)dx \right\} dx + C_1 \right]. \quad (5.1.11)$$

为了求非齐次方程(5.1.6)的满足初始条件 $y(x_0) = y_0$ 的特解, 只需改写(5.1.9)为

$$y(x) = C(x) \exp \left\{ - \int_{x_0}^x p(x) dx \right\}. \quad (5.1.12)$$

并代入(5.1.6)、化简、求解, 得

$$C(x) = \int_{x_0}^x q(t) \exp \left\{ \int_{x_0}^t p(s) ds \right\} dt + C_1, \quad (5.1.13)$$

再将(5.1.13)代入(5.1.9), 从而(5.1.6)满足 $y(x_0) = y_0$ 的特解为

$$y(x) = \exp \left\{ - \int_{x_0}^x p(s) ds \right\} \left[y_0 + \int_{x_0}^x q(t) \exp \left\{ \int_{x_0}^t p(s) ds \right\} dt \right]. \quad (5.1.14)$$

例7 求解方程

$$\frac{dy}{dx} - \frac{2}{x}y = \frac{1}{2}x.$$

解 将 $p(x) = -\frac{2}{x}$, $q(x) = \frac{x}{2}$ 代入(5.1.11)得

$$\begin{aligned} y(x) &= \exp \left\{ \int \frac{2}{x} dx \right\} \cdot \left[\int \frac{x}{2} \exp \left\{ \int -\frac{2}{x} dx \right\} dx + C_1 \right] \\ &= x^2 \left(\int \frac{1}{2x} dx + C_1 \right) = \frac{1}{2}x^2 \ln|x| + C_1 x^2, \end{aligned}$$

即为方程的通解.

例8 求解方程

$$(y^2 - 2x)dy - ydx = 0.$$

解 在这个方程中, 如果把 y 看作 x 的函数, 则是一个非线性方程, 但是, 如果把 x 看作 y 的函数, 它却是线性的:

$$\frac{dx}{dy} - \frac{2}{y}x + y = 0, \quad y \neq 0.$$

用常数变易法就可得到其解

$$\begin{aligned} x &= \exp \left\{ \int \frac{2}{y} dy \right\} \cdot \left[\int (-y) \exp \left(- \int \frac{2}{y} dy \right) dy + C_1 \right] \\ &= y^2 [c - \ln|y|]. \end{aligned}$$

$y=0$ 显然是原方程的解.

注 在求解微分方程时, 通常把 y 看作 x 的函数, 但为了方便起见, 也常把 x 、 y 平等看待, 这样所求得解往往是隐式形式的解.

3. 贝努里(Bernoulli)方程

有许多方程, 经过一定的变量代换可以化为线性方程, 贝努里方程就是其中

的一种.形如

$$y' + p(x)y = q(x)y^n \quad (n \text{ 为不等于 } 0, 1 \text{ 的实常数})$$

的方程称为贝努里方程.它的解法是:用 y^n 除方程的两边,得到

$$y^{-n}y' + p(x)y^{1-n} = q(x) \quad (n \neq 0, 1).$$

作变量代换 $z = y^{1-n}$,就可将原方程化为线性方程

$$z' + (1-n)p(x)z = (1-n)q(x) \quad (y \neq 0).$$

例 9 求解方程

$$\frac{dy}{dx} + \frac{y}{3} = \frac{(1-2x)y^4}{3}.$$

解 此方程显然不是线性的.然而,我们可以设法把它变成线性形式.在方程两边同乘 $3y^4$ ($y \neq 0$),得

$$3y^{-4}\frac{dy}{dx} + y^{-3} = 1 - 2x \quad (y \neq 0).$$

作变换 $z = y^{-3}$,方程化为以 z 为函数的线性方程

$$\frac{dz}{dx} - z = 2x - 1.$$

可求得解

$$z = e^x \left[c + \int (2x - 1)e^{-x} dx \right] = -1 - 2x + ce^x,$$

故原方程的解为

$$y^{-3} = -1 - 2x + ce^x,$$

$y = 0$ 显然也是原方程的解.

例 10 求解方程(黎卡堤(Riccati)方程)

$$\frac{dy}{dx} + p(x)y + q(x)y^2 = f(x).$$

解 初看起来,它不是贝努里方程,但是如果作函数的变换 $y = z + y_1$,这里的 y_1 是原方程的一个特解,则原方程化为

$$z' + (p(x) + 2q(x)y_1)z = -q(x)z^2,$$

上式是贝努里方程.

例如, $y' + y^2 = -\frac{2}{x^2}$,它是黎卡堤方程,易见它有一个特解 $y_1 = \frac{1}{x}$.于是,令

$y = z + \frac{1}{x}$,代入原方程并化简,得贝努里方程 $\frac{dz}{dx} - \frac{2}{x}z = z^2$.再令 $u = z^{-1}$,得到

u 的一阶线性方程 $u' + \frac{2}{x}u = -1$.求得其解 u 后,并代回原来的未知函数 y ,得

$$y = \frac{3x^2}{3C - x^3} + \frac{1}{x}.$$

5.1.3 高阶常系数线性微分方程

形如

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = f(x) \quad (5.1.15)$$

的方程(其中 a_i ($i=1, 2, \cdots, n$), 均为实常数), 称为 n 阶常系数线性微分方程. 如果 $f(x) \equiv 0$, 即

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0, \quad (5.1.16)$$

则称(5.1.16)式为(5.1.15)式对应的 n 阶常系数齐次线性方程. 相应地, (5.1.15)称为 n 阶常系数非齐次线性方程.

1. n 阶常系数齐次线性方程的解法

对一阶方程

$$y' + ay = 0, \quad (5.1.17)$$

其中 a 是常数, 容易求出它的特解

$$y = e^{-ax}.$$

比较(5.1.16)与(5.1.17), 我们猜想(5.1.16)也有形如

$$y = e^{\lambda x} \quad (5.1.18)$$

的解, 其中 λ 是待定常数. 为此, 将(5.1.18)代入方程(5.1.16)中, 便有

$$(\lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n) e^{\lambda x} = 0. \quad (5.1.19)$$

令

$$p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n.$$

定义 5.1.3 (特征多项式、特征方程、特征根) 称 $p(\lambda)$ 为方程(5.1.16)的特征多项式, $p(\lambda) = 0$ 为其特征方程, 而 $p(\lambda) = 0$ 的根为方程(5.1.16)的特征根.

定理 5.1.1 (常系数齐次线性方程解的充要条件) 方程(5.1.16)有形如(5.1.18)的解的充要条件是: λ 满足

$$p(\lambda) = 0. \quad (5.1.20)$$

亦即(5.1.18)中的 λ 应是特征方程(5.1.20)的根.

这样, 求方程(5.1.16)具有(5.1.18)形式的解的问题便归结为求(5.1.16)的特征根.

下面分两种情况加以讨论.

情形 1 特征根互异

首先假设特征方程(5.1.20)有 n 个互不相同的实根

$$\lambda_1, \lambda_2, \cdots, \lambda_n.$$

这时, 方程(5.1.16)有下列 n 个特解

$$y = e^{\lambda_1 x}, y = e^{\lambda_2 x}, \cdots, y = e^{\lambda_n x}. \quad (5.1.21)$$

可以证明,函数

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} + \cdots + c_n e^{\lambda_n x} \quad (5.1.22)$$

是方程(5.1.16)的通解,其中 $c_1, c_2, \cdots, c_n$ 是 n 个任意常数.

例 11 求方程

$$y''' + 2y'' - y' - 2y = 0$$

的通解及满足初值条件: $y(0) = 0, y'(0) = 1, y''(0) = 3$ 的特解.

解 特征方程为

$$\lambda^3 + 2\lambda^2 - \lambda - 2 = 0,$$

作因式分解,得

$$(\lambda + 2)(\lambda - 1)(\lambda + 1) = 0.$$

从而有特征根 $\lambda_1 = -2, \lambda_2 = 1, \lambda_3 = -1$. 因此通解为

$$y = c_1 e^{-2x} + c_2 e^x + c_3 e^{-x},$$

其中 c_1, c_2, c_3 为任意常数.

为求特解,我们把初值代入通解表达式,可得

$$\begin{cases} c_1 + c_2 + c_3 = 0, \\ -2c_1 + c_2 - c_3 = 1, \\ 4c_1 + c_2 + c_3 = 3. \end{cases}$$

由此解得 $c_1 = 1, c_2 = 1, c_3 = -2$. 故

$$y(x) = e^{-2x} + e^x - 2e^{-x}$$

为所求特解.

如果(5.1.20)的互异根中有复根,例如, $\lambda = a + ib$ 是一个复根,则由

$$(e^{\lambda x})' = \lambda e^{\lambda x}$$

知函数 $y = e^{(a+ib)x}$ 仍是方程(5.1.15)的解. 容易推知:无论 λ 是实数还是复数,只要方程(5.1.20)有 n 个互异的实的或复的特征根 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 时,方程(5.1.16)仍有形如(5.1.21)的 n 个特解. 同样可以证明,含有 n 个任意常数的线性组合

$$c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} + \cdots + c_n e^{\lambda_n x}$$

是方程(5.1.16)的通解.

例 12 求解方程

$$y''' - 5y'' + 17y' - 13y = 0.$$

解 特征方程 $\lambda^3 - 5\lambda^2 + 17\lambda - 13 = (\lambda - 1)(\lambda^2 - 4\lambda + 13) = 0$ 有特征根

$$\lambda_1 = 1, \lambda_2 = 2 + 3i, \lambda_3 = 2 - 3i,$$

故

$$y(x) = c_1 e^x + c_2 e^{(2+3i)x} + c_3 e^{(2-3i)x},$$

其中 c_1, c_2, c_3 是任意常数, 为所求通解.

对方程(5.1.15), 由于它是实系数的, 故当它的特征方程有 n 个互异的特征根(不论是实的还是复的)时, 我们还可以求出它的 n 个实特解, 这些特解的含 n 个任意常数的线性组合仍构成(5.1.16)的通解. 为此, 设 $\lambda = a + ib$ 是(5.1.16)的特征根, 注意到特征方程(5.1.20)是实系数的, 故 λ 的共轭复数 $\bar{\lambda} = a - ib$ 也是(5.1.16)的特征根, 从而

$$\left. \begin{aligned} y_1(x) &= e^{\lambda x} = e^{ax} (\cos bx + i \sin bx), \\ y_2(x) &= e^{\bar{\lambda} x} = e^{ax} (\cos bx - i \sin bx) \end{aligned} \right\} \quad (5.1.23)$$

都是方程(5.1.16)的解.

对于方程(5.1.16)的两个解 $y_1(x), y_2(x)$, 其线性组合

$$k_1 y_1(x) + k_2 y_2(x),$$

k_1, k_2 为任意常数(实值的或复值的), 仍是方程(5.1.16)的解. 这是因为

$$\begin{aligned} & [k_1 y_1 + k_2 y_2]^{(n)} + a_1 [k_1 y_1 + k_2 y_2]^{(n-1)} + \cdots + \\ & a_{n-1} [k_1 y_1 + k_2 y_2]' + a_n [k_1 y_1 + k_2 y_2] \\ &= k_1 [y_1^{(n)} + a_1 y_1^{(n-1)} + \cdots + a_{n-1} y_1' + a_n y_1] + \\ & k_2 [y_2^{(n)} + a_1 y_2^{(n-1)} + \cdots + a_{n-1} y_2' + a_n y_2] \\ &= k_1 \cdot 0 + k_2 \cdot 0 = 0. \end{aligned}$$

于是, 取适当的常数使得

$$\frac{1}{2} [y_1(x) + y_2(x)] = e^{ax} \cos bx, \quad \frac{1}{2i} [y_1(x) - y_2(x)] = e^{ax} \sin bx, \quad (5.1.24)$$

它们也是方程(5.1.16)的实函数解.

当方程(5.1.20)有复根时, 方程(5.1.16)就有形如(5.1.23)的解. 这时, 只要逐对地将(5.1.23)换成(5.1.24)的实解, 便可得到(5.1.16)的实通解. 例如, 在例12中, 已知道它有特征根 $\lambda_1 = 1, \lambda_2 = 2 + 3i, \lambda_3 = 2 - 3i$, 而 λ_2 与 λ_3 是共轭复数. 从而方程有特解

$$y_1(x) = e^x, \quad y_2(x) = e^{2x} \cos 3x, \quad y_3(x) = e^{2x} \sin 3x,$$

故有通解

$$y(x) = c_1 e^x + c_2 e^{2x} \cos 3x + c_3 e^{2x} \sin 3x.$$

例13 求方程 $y^{(4)} - 2y^{(3)} + 6y'' - 2y' + 5y = 0$ 的通解.

解 特征方程

$$\lambda^4 - 2\lambda^3 + 6\lambda^2 - 2\lambda + 5 = (\lambda^2 + 1)(\lambda^2 - 2\lambda + 5) = 0$$

有特征根 $\lambda_1 = i, \lambda_2 = -i, \lambda_3 = 1 + 2i, \lambda_4 = 1 - 2i$, 因而有特解

$$y_1(x) = \cos x, y_2(x) = \sin x, y_3(x) = e^x \cos 2x, y_4(x) = e^x \sin 2x.$$

从而

$$y(x) = c_1 \cos x + c_2 \sin x + c_3 e^x \cos 2x + c_4 e^x \sin 2x$$

为所求通解.

情形 2 特征根有重根

设 λ 是方程(5.1.20)的 k 重特征根(实的或复的),由前面讨论, $e^{\lambda x}$ 是方程(5.1.16)的一个解.下面证明,此时

$$e^{\lambda x}, x e^{\lambda x}, \dots, x^{k-1} e^{\lambda x}$$

是方程(5.1.16)的 k 个特解.

首先注意:如果 λ 是(5.1.20)的 k 重特征根,则有

$$p(\lambda) = p'(\lambda) = \dots = p^{(k-1)}(\lambda) = 0, p^{(k)}(\lambda) \neq 0.$$

其次,设 $u = e^{\lambda x}$, 则 $\frac{\partial^m u}{\partial \lambda^m} = x^m e^{\lambda x}$, 所以

$$\frac{\partial^j}{\partial x^j} (x^m e^{\lambda x}) = \frac{\partial^j}{\partial x^j} \left(\frac{\partial^m u}{\partial \lambda^m} \right) = \frac{\partial^m}{\partial \lambda^m} \left(\frac{\partial^j u}{\partial x^j} \right) = \frac{\partial^m}{\partial \lambda^m} (\lambda^j e^{\lambda x}).$$

由莱布尼茨公式

$$(uv)^{(m)} = \sum_{l=0}^m c_m^l u^{(l)} v^{(m-l)},$$

特别取 $u(\lambda) = p(\lambda)$, $v(\lambda) = e^{\lambda x}$, 则

$$\begin{aligned} & \frac{\partial^n}{\partial x^n} (x^m e^{\lambda x}) + a_1 \frac{\partial^{n-1}}{\partial x^{n-1}} (x^m e^{\lambda x}) + \dots + a_{n-1} \frac{\partial}{\partial x} (x^m e^{\lambda x}) + a_n (x^m e^{\lambda x}) \\ &= \frac{\partial^m}{\partial \lambda^m} \left(\frac{\partial^n}{\partial x^n} e^{\lambda x} \right) + a_1 \frac{\partial^m}{\partial \lambda^m} \left(\frac{\partial^{n-1}}{\partial x^{n-1}} e^{\lambda x} \right) + \dots + a_{n-1} \frac{\partial^m}{\partial \lambda^m} \left(\frac{\partial}{\partial x} e^{\lambda x} \right) + a_n \frac{\partial^m}{\partial \lambda^m} (e^{\lambda x}) \\ &= \frac{\partial^m}{\partial \lambda^m} [(\lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n) e^{\lambda x}] \\ &= \frac{\partial^m}{\partial \lambda^m} [p(\lambda) e^{\lambda x}] = \sum_{v=0}^m c_m^v p^{(v)}(\lambda) \frac{\partial^{m-v}}{\partial \lambda^{m-v}} (e^{\lambda x}). \end{aligned}$$

当 $m=0, 1, \dots, k-1, 0 \leq v \leq m$ 时, $p^{(v)}(\lambda) = 0$, 故当 $0 \leq m \leq k-1$ 时,

$$(x^m e^{\lambda x})^{(n)} + a_1 (x^m e^{\lambda x})^{(n-1)} + \dots + a_{n-1} (x^m e^{\lambda x})' + a_n (x^m e^{\lambda x}) = 0,$$

即 $x^m e^{\lambda x}$ 是方程(5.1.16)的特解.

一般地,我们有如下结论.

定理 5.1.2 (常系数齐次线性方程的通解) 如果方程(5.1.20)有两两互异的特征根 $\lambda_1, \dots, \lambda_\mu$, 并且它们的重数分别是 $k_1, \dots, k_\mu$ ($k_i \geq 1$), 其中 $k_1 + \dots + k_\mu = n$, 则与之对应的(5.1.16)的特解是

$$\begin{aligned} & e^{\lambda_1 x}, x e^{\lambda_1 x}, \dots, x^{k_1-1} e^{\lambda_1 x}; \\ & \dots \dots \dots \\ & e^{\lambda_\mu x}, x e^{\lambda_\mu x}, \dots, x^{k_\mu-1} e^{\lambda_\mu x}; \end{aligned}$$

且它们的含 n 个任意常数的线性组合构成(5.1.16)的通解.

若在定理 5.1.1 中出现复重根,那么如同单根时一样,也可求出实解.

例 14 求解方程 $y''' + 3y'' + 3y' + y = 0$.

解 特征方程 $\lambda^3 + 3\lambda^2 + 3\lambda + 1 = (\lambda + 1)^3 = 0$ 有三重根 $\lambda_1 = \lambda_2 = \lambda_3 = -1$, 故

$$y_1(x) = e^{-x}, y_2(x) = xe^{-x}, y_3(x) = x^2e^{-x}$$

是对应的特解.

$$y(x) = (c_1 + c_2x + c_3x^2)e^{-x}$$

为所求通解,其中 c_1, c_2, c_3 是任意常数.

例 15 求方程 $y^{(5)} - y^{(4)} + 4y^{(3)} - 4y'' + 4y' - 4y = 0$ 的通解.

解 特征方程

$$\lambda^5 - \lambda^4 + 4\lambda^3 - 4\lambda^2 + 4\lambda - 4 = (\lambda^2 + 2)^2(\lambda - 1) = 0$$

有特征根 $\lambda_1 = 1, \lambda_2 = \lambda_3 = \sqrt{2}i, \lambda_4 = \lambda_5 = -\sqrt{2}i$. 故

$$e^x, \cos\sqrt{x}, x\cos\sqrt{x}, \sin\sqrt{x}, x\sin\sqrt{x}$$

为对应的特解. 而

$$y(x) = c_1e^x + (c_2 + c_3x)\cos\sqrt{x} + (c_4 + c_5x)\sin\sqrt{x}$$

为所求通解,其中 c_1, c_2, c_3, c_4, c_5 是任意常数.

2. n 阶常系数非齐次线性方程的解法,算子解法

给定 n 阶常系数非齐次线性方程

$$y^{(n)} + a_1y^{(n-1)} + \cdots + a_{n-1}y' + a_ny = f(x), \quad (5.1.25)$$

我们研究它的解的结构及求解方法.

记

$$L(D) = D^n + a_1D^{n-1} + \cdots + a_{n-1}D + a_n,$$

其中 $D = \frac{d}{dx}$ 表示对 x 求微商的运算, $D^k = \frac{d^k}{dx^k}$ 表示对 x 求 k 次微商. $L(D)$ 称为

算子多项式, $L(D)$ 作用到函数 y 上是指

$$L(D)y = D^n y + a_1D^{n-1}y + \cdots + a_{n-1}Dy + a_ny. \quad (5.1.26)$$

因此,非齐次方程(5.1.25)可以简写成“微分算子方程”

$$L(D)y = f(x). \quad (5.1.27)$$

对应的齐次方程可以写成

$$L(D)y = 0. \quad (5.1.28)$$

定理 5.1.3 (算子 $L(D)$ 的性质) 算子 $L(D)$ 的性质:

(1) $L(D)[\alpha_1y_1 + \alpha_2y_2] = \alpha_1L(D)y_1 + \alpha_2L(D)y_2$, (线性性)

α_1, α_2 为任意常数;

(2) 若 $L(t) = L_1(t) + L_2(t) = L_2(t) + L_1(t)$, 则

$$L(D) = L_1(D) + L_2(D) = L_2(D) + L_1(D), \quad (\text{加法可交换性})$$

其中 $L_1(t), L_2(t)$ 是变量 t 的普通多项式, 于是

$$L(D)y = L_1(D)y + L_2(D)y = L_2(D)y + L_1(D)y. \quad (5.1.29)$$

(3) 若 $L(t) = L_1(t)L_2(t)$, 则

$$L(D) = L_1(D)L_2(D) = L_2(D)L_1(D), \quad (\text{乘法可交换性})$$

其中 $L(t), L_1(t), L_2(t)$ 是变量 t 的普通多项式, 于是

$$L(D)y = L_1(D)[L_2(D)y] = L_2(D)[L_1(D)y]. \quad (5.1.30)$$

利用定理 5.1.2, 我们可以得到下列两个关于 (5.1.25) 的解的结构定理.

定理 5.1.4 (解的结构) n 阶非齐次线性方程 (5.1.25)

$$L(D)y = y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = f(x)$$

的通解等于它所对应的齐次方程 (5.1.26)

$$L(D)y = y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0$$

的通解与 (5.1.25) 本身的一个特解之和.

证 首先证明, 若 $y_1(x)$ 是 (5.1.25) 的解, $y(x)$ 是 (5.1.26) 的解, 则 $y_1 + y$ 是 (5.1.25) 的解. 事实上, 由定理 5.1.2 的结论 1,

$$L(D)[y_1 + y] = L(D)y_1 + L(D)y.$$

因为

$$L(D)y_1 = f(x), \quad L(D)y = 0,$$

所以

$$L(D)[y_1 + y] = f(x),$$

即 $y_1 + y$ 是方程 (5.1.25) 的解.

其次, 证明对于 (5.1.25) 的任一个解 $\tilde{y}$, 必有 $\tilde{y} = y_1 + y^*$, 其中 y^* 是 (5.1.25) 的一个特解, y^* 是 (5.1.26) 的某个解. 这等价于证明 $\tilde{y} - y_1$ 是 (5.1.26) 的一个解. 实际上,

$$L(D)[\tilde{y} - y_1] = L(D)\tilde{y} - L(D)y_1 = f(x) - f(x) = 0.$$

故 $\tilde{y} - y_1$ 是方程 (5.1.26) 的解. 这就是要证明的.

上述结论表明, (5.1.25) 的任一解 $\tilde{y}$ 均可表为 (5.1.25) 的一个特解 y_1 与 (5.1.26) 的某个解 y^* 之和. 故定理得证.

根据定理 5.1.3, 为求 (5.1.25) 的通解, 只需求它的一个特解, 再加上相应的齐次线性方程 (5.1.26) 的通解便可.

下面我们对一些特殊的 $f(x)$, 给出求 (5.1.25) 的特解的某些方法.

定理 5.1.5 (非齐次方程解的充要条件) 给定复函数 $f(x) = f_1(x) +$

$if_2(x)$, 其中 $f_1(x), f_2(x)$ 为实函数, 则复值函数 $y^*(x) = y_1(x) + iy_2(x)$ 为方程 (5.1.25) 的解的充要条件是: $y_1(x)$ 与 $y_2(x)$ 是方程 $L(D)y = f_j(x)$ ($j = 1, 2$) 的解.

定理 5.1.4 的证明很容易, 只需注意方程 (5.1.25) 的系数 a_k ($k = 1, 2, \dots, n$), 均为实数即可. 我们把它留给读者自己完成.

为了利用算子多项式求方程 (5.1.25) 的特解, 先给出下面的定义.

定义 5.1.4 ($L(D)$ 的逆算子) 对微分算子方程 (5.1.27), 我们用

$$L^{-1}(D)f(x) \quad (5.1.31)$$

来记它的任一特解, 并称 $L^{-1}(D)$ 为 $L(D)$ 的逆算子.

证 这样定义的 $L^{-1}(D)f(x)$ 的值不惟一, 但这并不重要, 因为我们的目的是寻找 (5.1.25) 的一个特解. 通常总取 (5.1.31) 的表达式尽可能简单, 且不含常数.

例如, 我们总取

$$\begin{aligned} D^{-1}f(x) &= \int_{x_0}^x f(t)dt, \\ (D - \gamma)^{-1}f(x) &= e^{\gamma x} \int_{x_0}^x e^{-\gamma t} f(t)dt, \gamma \text{ 为常数.} \end{aligned} \quad (5.1.32)$$

根据上述定义与定理 5.1.3 可得

定理 5.1.6 (逆算子的性质) $L^{-1}(D)$ 具有下列性质:

$$(1) L(D)[L^{-1}(D)f(x)] = f(x), \quad (5.1.33)$$

$$(2) L^{-1}(D)[f_1(x) + f_2(x)] = L^{-1}(D)f_1(x) + L^{-1}(D)f_2(x), \quad (5.1.34)$$

$$\begin{aligned} (3) L_1^{-1}(D)L_2^{-1}(D)f(x) &= L_1^{-1}(D)[(L_2^{-1}(D)f(x))] \\ &= L_2^{-1}(D)[L_1^{-1}(D)f(x)]. \end{aligned} \quad (5.1.35)$$

证 (5.1.33) 式可直接由定义得到.

(5.1.34) 式的证明只要注意定义 5.1.1 及定理 5.1.3 中的性质便可得到.

要证 (5.1.35) 式, 注意到, 由 (5.1.30) 知, 若 $L(D) = L_1(D)L_2(D)$, 则三个方程

$$L(D)y = f(x), L_1(D)[L_2(D)y] = f(x), L_2(D)[L_1(D)y] = f(x)$$

是一样的, 而

$$L_1^{-1}(D)L_2^{-1}(D)f(x); L_1^{-1}(D)[L_2^{-1}(D)f(x)]; L_2^{-1}(D)[L_1^{-1}(D)f(x)]$$

分别是它们的一个特解, 按 (5.1.31) 式下面的注, 它们是相等的.

利用定理 5.1.6, 我们给出求 (5.1.25) 的一些特解的法则.

法则 I

$$L^{-1}(D)[e^{\alpha x}\varphi(x)] = e^{\alpha x}L^{-1}(D+\alpha)\varphi(x). \quad (5.1.36)$$

特别, 当 $L(\alpha) \neq 0$ 时, 即 α 不是特征方程 $L(\lambda) = 0$ 的根时, 我们有

$$L^{-1}(D)e^{\alpha x} = \frac{1}{L(\alpha)}e^{\alpha x}, \quad (5.1.37)$$

当 α 是特征方程 $L(\lambda) = 0$ 的 s 重特征根 ($s \geq 0$) 时, 即 $L(\lambda) = (\lambda - \alpha)^s L_1(\lambda)$, 而 $L_1(\alpha) \neq 0$, 或 $L^{(s)}(\alpha) = s! L_1(\alpha) \neq 0$, 则

$$L^{-1}(D)e^{\alpha x} = \frac{x^s e^{\alpha x}}{s! L_1(\alpha)} = \frac{x^s e^{\alpha x}}{L^{(s)}(\alpha)} \quad (s \geq 0). \quad (5.1.38)$$

证 先证当 $L(D) = D - \gamma$ 时, (5.1.36) 式成立. 因为

$$\begin{aligned} (D - \gamma)[e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x)] &= D[e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x)] - \\ &\quad \gamma[e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x)] \\ &= \alpha e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x) + e^{\alpha x}D[(D + \alpha - \gamma)^{-1}\varphi(x)] - \gamma e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x) \\ &= e^{\alpha x}(D + \alpha - \gamma)[(D + \alpha - \gamma)^{-1}\varphi(x)] = e^{\alpha x}\varphi(x). \end{aligned}$$

上式给出 $(D - \gamma)[e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x)] = e^{\alpha x}\varphi(x)$. 于是由定义

$$(D - \gamma)^{-1}[e^{\alpha x}\varphi(x)] = e^{\alpha x}(D + \alpha - \gamma)^{-1}\varphi(x).$$

其次应用定理 5.1.6 的 (5.1.35) 式及归纳法, 并用代数学基本定理将 $L(D)$ 因式分解

$$L(D) = (D - p_1)(D - p_2) \cdots (D - p_n),$$

这里 $p_1, \dots, p_n$ 为 n 个数, 我们有

$$\begin{aligned} L^{-1}(D)[e^{\alpha x}\varphi(x)] &= [(D - p_1) \cdots (D - p_n)]^{-1}[e^{\alpha x}\varphi(x)] \\ &= [(D - p_1) \cdots (D - p_{n-1})]^{-1}[e^{\alpha x}(D + \alpha - p_n)]^{-1}\varphi(x) \\ &\quad \dots \end{aligned}$$

$$= e^{\alpha x}[(D + \alpha - p_1) \cdots (D + \alpha - p_n)]^{-1}\varphi(x) = e^{\alpha x}L^{-1}(D + \alpha)\varphi(x),$$

这就是 (5.1.36) 要证 (5.1.37) 式, 只要在 (5.1.36) 式中取 $\varphi(x) \equiv 1$, 并注意 $L^{-1}(D + \alpha) \cdot 1$ 是方程 $L(D + \alpha)y = 1$ 的解, 故当 $L(\alpha) \neq 0$ 时, 得

$$L^{-1}(D)e^{\alpha x} = e^{\alpha x}L^{-1}(D + \alpha) \cdot 1 = \frac{e^{\alpha x}}{L(\alpha)}.$$

(5.1.37) 得证. 当 $L(\lambda) = (\lambda - \alpha)^s L_1(\lambda)$, 而 $L_1(\alpha) \neq 0$ 时, 在 (5.1.36) 式中取 $\varphi(x) \equiv 1$, 得

$$L^{-1}(D)e^{\alpha x} = e^{\alpha x}L^{-1}(D + \alpha) \cdot 1 = e^{\alpha x}(D^s)^{-1} \frac{1}{L_1(\alpha)}.$$

又因 $(D^s)^{-1} \cdot 1 = \frac{x^s}{s!}$, 故得

$$L^{-1}(D)e^{ax} = \frac{x^s e^{ax}}{s! L_1(\alpha)} \quad (s \geq 0),$$

即(5.1.38)式成立.

例 16 求 $y''' + y'' - 5y' + 3y = e^{2x}$ 的通解.

解 对应的齐次方程为 $y''' + y'' - 5y' + 3y = 0$, 其特征方程为

$$L(\lambda) = \lambda^3 + \lambda^2 - 5\lambda + 3 = (\lambda - 1)^2(\lambda + 3) = 0.$$

有特征根 $\lambda_1 = \lambda_2 = 1, \lambda_3 = -3$, 故齐次方程有通解

$$y(x) = c_1 e^{-3x} + (c_2 + c_3 x) e^x.$$

又因为 $k=2$ 不是特征方程的根, 故由(5.1.37)式得

$$y^*(x) = \frac{e^{2x}}{L(2)} = \frac{e^{2x}}{5}$$

为非齐次方程的一个特解, 从而

$$y(x) = c_1 e^{-3x} + (c_2 + c_3 x) e^x + \frac{e^{2x}}{5}$$

为所求通解.

例 17 求解方程 $(D^4 - 6D^2 - 8D - 3)y = e^{-x}$.

解 对应的齐次方程为 $(D^4 - 6D^2 - 8D - 3)y = 0$, 其特征方程

$$L(\lambda) = \lambda^4 - 6\lambda^2 - 8\lambda - 3 = (\lambda - 3)(\lambda + 1)^3 = 0$$

有特征根 $\lambda_1 = 3, \lambda_2 = \lambda_3 = \lambda_4 = -1$, 故齐次方程的通解为

$$c_1 e^{3x} + (c_2 + c_3 x + c_4 x^2) e^{-x}.$$

又 $k = -1$ 为特征方程的 3 重根, 故由(5.1.38)式得

$$y^*(x) = \frac{x^3 e^{-x}}{3! L_1(-1)} = \frac{x^3 e^{-x}}{6(-1-3)} = -\frac{1}{24} x^3 e^{-x}$$

为非齐次方程的一个特解. 从而

$$y(x) = c_1 e^{3x} + (c_2 + c_3 x + c_4 x^2) e^{-x} - \frac{1}{24} x^3 e^{-x}$$

为所求通解.

例 18 求 $(D^2 + 2D + 2)y = e^{-x} \cos x$ 的一个特解.

解 方程右边的非齐次项是 $e^{-x} \cos x$, 利用欧拉公式

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$$

得

$$e^{-x} \cos x = \frac{1}{2} e^{(-1+i)x} + \frac{1}{2} e^{(-1-i)x},$$

从而由定理 5.1.6 的(5.1.34)式知

$$y^*(x) = (D^2 + 2D + 2)^{-1} \left[\frac{1}{2} e^{(-1+i)x} \right] + (D^2 + 2D + 2)^{-1} \left[\frac{1}{2} e^{(-1-i)x} \right].$$

注意到, 特征方程 $\lambda^2 + 2\lambda + 2 = [\lambda - (-1+i)][\lambda - (-1-i)] = 0$ 有特征根 $\lambda_1 = -1+i, \lambda_2 = -1-i$, 从而

$$(D^2 + 2D + 2)^{-1} \left[\frac{1}{2} e^{(-1+i)x} \right] = \frac{1}{2} x e^{(-1+i)x} \frac{1}{(-1+i) - (-1-i)} = \frac{-ix}{4} e^{(-1+i)x},$$

$$(D^2 + 2D + 2)^{-1} \left[\frac{1}{2} e^{(-1-i)x} \right] = \frac{1}{2} x e^{(-1-i)x} \frac{1}{(-1-i) - (-1+i)} = \frac{ix}{4} e^{(-1-i)x}.$$

从而

$$y^*(x) = \frac{x \sin x}{2} e^{-x}$$

为所求的一个特解.

当方程(5.1.25)式右边的 $f(x)$ 是 x 的多项式时, 要计算 $L^{-1}(D)f(x)$, 需用到下面的法则.

法则 II 设 $f(x)$ 是 m 次多项式, $L(0) \neq 0$ (即 $L(D)$ 中的常数项 $a_n \neq 0$), 对于 $\frac{1}{L(t)}$ 按 t 的升幂展开为级数时, 前面的 m 次多项式记为 $Q(t)$, 于是 $Q(t)$ 中的常数项也不为零, 则

$$L^{-1}(D)f(x) = Q(D)f(x). \quad (5.1.39)$$

证 由假设有

$$\frac{1}{L(t)} = Q(t) + \frac{R(t)}{L(t)},$$

其中 $R(t)$ 是多项式, 首项的幂次大于 m , 以 $L(t)$ 乘上式两边, 得

$$1 = Q(t)L(t) + R(t).$$

于是, 由定理 5.1.3 也有

$$L(D)Q(D) + R(D) = 1.$$

把这个运算加到 $f(x)$ 上去, 并注意 $R(D)f(x) = 0$, 得到

$$L(D)Q(D)f(x) = f(x),$$

从而

$$L^{-1}(D)f(x) = Q(D)f(x),$$

当 $L(D)$ 中 D 的最低幂次为 s ($s \in \mathbb{N}$) 时 (此时 $L(0) = 0$), 可记

$$L(D) = D^s L_1(D) \quad (L_1(0) \neq 0),$$

于是由(5.1.39)式得

$$L^{-1}(D)f(x) = [D^s L_1(D)]^{-1} f(x) = (D^s)^{-1} Q_1(D)f(x) \quad (s \geq 0).$$

对于 $f(x) = e^{kx} p(x)$, 其中 $p(x)$ 为多项式, 要就方程(5.1.25)的特解, 可以先用法则 I, 再用法则 II. 注意 $f(x)$ 的这一形式也把 $e^{\alpha x} p(x) \cos \beta x$,

$e^{\alpha x} p(x) \sin \beta x$ 包含在其中了.

例 19 求方程 $(D^2 - 5D + 6)y = 6x^2 - 10x + 2$ 的通解.

解 对应的齐次方程为 $(D^2 - 5D + 6)y = 0$, 其特征方程

$$L(\lambda) = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) = 0$$

有特征根 $\lambda_1 = 2, \lambda_2 = 3$, 故齐次方程有通解 $c_1 e^{2x} + c_2 e^{3x}$.

由于 $f(x) = 6x^2 - 10x + 2$ 是二次多项式, 故取

$$\begin{aligned} \frac{1}{L(t)} &= \frac{1}{(t-2)(t-3)} = \frac{1}{2-t} - \frac{1}{3-t} \\ &= \frac{1}{2} \left(1 + \frac{1}{2}t + \frac{1}{4}t^2 \right) - \frac{1}{3} \left(1 + \frac{1}{3}t + \frac{1}{9}t^2 \right) + R(t), \end{aligned}$$

$R(t)$ 关于 t 的最低次幂大于 2, 从而

$$\begin{aligned} y^*(x) &= \frac{1}{2} \left(1 + \frac{1}{2}D + \frac{1}{4}D^2 \right) (6x^2 - 10x + 2) - \\ &\quad \frac{1}{3} \left(1 + \frac{1}{3}D + \frac{1}{9}D^2 \right) (6x^2 - 10x + 2) = x^2. \end{aligned}$$

故

$$y(x) = x^2 + c_1 e^{2x} + c_2 e^{3x}$$

为所求通解.

例 20 求方程 $(D^2 - 2D + 4)y = (x+2)e^{3x}$ 的通解.

解 对应的齐次方程为 $(D^2 - 2D + 4)y = 0$, 其特征方程为

$$\lambda^2 - 2\lambda + 4 = (\lambda - 1)^2 + 3 = 0,$$

有特征根 $\lambda_1 = 1 + i\sqrt{3}, \lambda_2 = 1 - i\sqrt{3}$. 从而齐次方程有通解

$$c_1 e^x \cos \sqrt{3}x + c_2 e^x \sin \sqrt{3}x.$$

由 $f(x) = (x+2)e^{3x}$, 所以根据法则 I 中的 (5.1.36) 式得

$$y^*(x) = e^{3x} (D^2 + 4D + 7)^{-1} (x+2).$$

又

$$\frac{1}{L(t)} = \frac{1}{t^2 + 4t + 7} = \left(\frac{1}{7} - \frac{4}{49}t \right) + \frac{R(t)}{t^2 + 4t + 7},$$

$R(t)$ 的最低幂次大于 1. 从而

$$y^*(x) = e^{3x} \left(\frac{1}{7} - \frac{4}{49}D \right) (x+2) = \left(\frac{1}{7}x + \frac{10}{49} \right) e^{3x}.$$

故

$$y(x) = \left(\frac{1}{7}x + \frac{10}{49} \right) e^{3x} + (c_1 \cos \sqrt{3}x + c_2 \sin \sqrt{3}x) e^x$$

为所求通解.

例 21 求 $(D^2 - 2D + 10)y = x^2 e^x \sin 3x$ 的特解.

解 先求 $(D^2 - 2D + 10)y = x^2 e^{(1+3i)x}$ 的特解.

$$\begin{aligned} & (D^2 - 2D + 10)^{-1} [x^2 e^{(1+3i)x}] \\ &= e^{(1+3i)x} ([D + (1+3i)]^2 - 2[D + (1+3i)] + 10)^{-1} x^2 \\ &= e^{(1+3i)x} [D(D+6i)]^{-1} x^2 = e^{(1+3i)x} D^{-1} \left[-\frac{i}{6} + \frac{D}{36} + \frac{iD^2}{216} \right] x^2 \\ &= e^{(1+3i)x} D^{-1} \left(-\frac{ix^2}{6} + \frac{x}{18} + \frac{i}{108} \right) = e^{(1+3i)x} \left(-\frac{ix^3}{18} + \frac{x^2}{36} + \frac{ix}{108} \right). \end{aligned}$$

由等式两边的虚部相等, 得到

$$\begin{aligned} y^*(x) &= (D^2 - 2D + 10)^{-1} [x^2 e^x \sin 3x] \\ &= x e^x \left[\left(-\frac{x^2}{18} + \frac{1}{108} \right) \cos 3x + \frac{x}{36} \sin 3x \right], \end{aligned}$$

它就是所求的特解.

3. n 阶变系数线性常微分方程

n 阶变系数线性方程的一般形式为

$$y^{(n)} + p_1(x)y^{(n-1)} + \cdots + p_{n-1}(x)y' + p_n(x)y = f(x), \quad (5.1.40)$$

其中 $f(x)$ 与 $p_1(x), \cdots, p_n(x)$ 在 $[a, b]$ 连续. 类似于常系数情形, 令

$$L(D) = D^n + p_1(x)D^{n-1} + \cdots + p_{n-1}(x)D + p_n(x),$$

其中 $D = \frac{d}{dx}$, $D^2 = \frac{d^2}{dx^2}$, $\cdots$, $D^n = \frac{d^n}{dx^n}$. 于是方程 (5.1.40) 可写为算子形式

$$L(D)y(x) = f(x). \quad (5.1.41)$$

易见, 算子 $L(D)$ 是线性算子

$$L(D)(\alpha y_1(x) + \beta y_2(x)) = \alpha L(D)y_1(x) + \beta L(D)y_2(x).$$

为讨论方程 (5.1.41) 的解, 我们需要函数组线性相关与线性无关的概念.

定义 5.1.5 (函数组的线性相关性) 设 n 个函数 $y_1(x), \cdots, y_n(x)$ 定义在区间 $[a, b]$ 上, 若存在 n 个不全为零的常数 $C_1, \cdots, C_n$, 使得在 $[a, b]$ 上成立恒等式

$$C_1 y_1(x) + \cdots + C_n y_n(x) = 0 \quad (x \in [a, b]), \quad (5.1.42)$$

则称函数 $y_1(x), \cdots, y_n(x)$ 线性相关;

若为使 (5.1.42) 关于 $x \in [a, b]$ 成为恒等式, 需当且仅当 $C_1 = \cdots = C_n = 0$,

则称 $y_1(x), \cdots, y_n(x)$ 线性无关.

线性相关与线性无关的例子很多, 如

$\sin x^2, \cos x^2, 1$ 是线性相关的;

$1, x, x^2, \cdots, x^n$ 是线性无关的;

$e^{\lambda_1 x}, e^{\lambda_2 x}, \cdots, e^{\lambda_n x}$, 当 λ_j ($j = 1, 2, \cdots, n$) 互异时是线性无关的.

定义 5.1.6 (函数组的朗斯基 (H. J. Wronsky) 行列式) 设定义在区间

$[a, b]$ 的 n 个函数 $y_1(x), \dots, y_n(x)$ 是 $n-1$ 次可微的, 称

$$W(y_1, \dots, y_n) = \begin{vmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{vmatrix}$$

为 $y_1(x), \dots, y_n(x)$ 的朗斯基行列式.

利用朗斯基行列式可以刻画 n 阶齐次方程 $L(D)y=0$ 解的线性相关性.

定理 5.1.7 (齐次方程解的线性相关性) 设 $y_1(x), \dots, y_n(x)$ 是 n 阶齐次线性方程 $L(D)y=0$ 的解, 则 $y_1(x), \dots, y_n(x)$ 在 $[a, b]$ 上线性相关, 当且仅当其朗斯基行列式为零

$$W(y_1, \dots, y_n) = 0 \quad (x \in [a, b]).$$

现在我们给出 n 阶变系数线性常微分方程的解的定理.

定理 5.1.8 (齐次方程的通解) n 阶变系数齐次线性常微分方程 $L(D)y=0$ 存在 n 个线性无关的解 $y_1(x), \dots, y_n(x)$, 且其通解为

$$y(x) = c_1 y_1(x) + \cdots + c_n y_n(x) = \sum_{j=1}^n c_j y_j(x),$$

其中 $c_1, \dots, c_n$ 为任意常数.

定理 5.1.9 (非齐次方程的通解) n 阶变系数非齐次线性常微分方程 $L(D)y=f(x)$ 的通解为

$$y(x) = y^*(x) + c_1 y_1(x) + \cdots + c_n y_n(x) = y^*(x) + \sum_{j=1}^n c_j y_j(x),$$

其中 $y^*(x)$ 是 $L(D)y(x)=f(x)$ 的一个特解, $\sum_{j=1}^n c_j y_j(x)$ 为 $L(D)y(x)=0$ 的通解, $c_1, \dots, c_n$ 为任意常数.

下面介绍求解 n 阶变系数齐次与非齐次线性微分方程的 3 种方法.

(1) 降阶法

对于 n 阶变系数齐次线性微分方程 $L(D)y(x)=0$, 如果已经知道它的一个特解 $\bar{y}(x)$, 则可以把 n 阶方程化为 $n-1$ 阶方程.

事实上, 令 $y(x) = \bar{y}(x)z(x)$, 便可将方程 $L(D)y(x)=0$ 化为关于未知函数 $z'(x)$ 的 $n-1$ 阶方程.

请读者通过习题来熟悉这个方法.

(2) 常数变易法

以二阶方程为例. 对于

$$L(D)y(x) = y'' + p_1(x)y' + p_2(x)y = f(x),$$

设相应的齐次方程 $L(D)y(x)=0$ 的通解已经求出

$$y(x) = c_1 y_1(x) + c_2 y_2(x).$$

考虑 $L(D)y(x)=f(x)$ 的形如 $y(x)=c_1(x)y_1(x)+c_2(x)y_2(x)$ 的解. 求导数, 得到

$$y'(x) = c_1'(x)y_1 + c_2'(x)y_2 + c_1 y_1' + c_2(x)y_2'.$$

我们要求 $c_1(x), c_2(x)$ 满足

$$c_1'(x)y_1 + c_2'(x)y_2 = 0, \quad (5.1.43)$$

于是, $y'(x) = c_1 y_1' + c_2(x)y_2'$. 再求导数.

$$y''(x) = c_1'(x)y_1' + c_2'(x)y_2' + c_1(x)y_1'' + c_2(x)y_2''.$$

由于 $y_1(x), y_2(x)$ 是 $L(D)y(x)=0$ 的解, 并且由 $L(D)y(x)=f(x)$, 将所求得的 y'', y', y 代入方程并化简, 配合 (5.1.43), 得到

$$\begin{cases} c_1'(x)y_1'(x) + c_2'(x)y_2'(x) = f(x), \\ c_1'(x)y_1 + c_2'(x)y_2 = 0, \end{cases} \quad (5.1.44)$$

解出 $c_1'(x), c_2'(x)$

$$c_1'(x) = -\frac{y_2(x)f(x)}{W(x)}, \quad c_2'(x) = \frac{y_1(x)f(x)}{W(x)},$$

其中 $W(x) = W(y_1(x), y_2(x))$ 为 y_1, y_2 的朗斯基行列式, 积分得到

$$c_1(x) = \bar{c}_1 - \int_{x_0}^x \frac{y_2(t)f(t)}{W(t)} dt, \quad c_2(x) = \bar{c}_2 + \int_{x_0}^x \frac{y_1(t)f(t)}{W(t)} dt,$$

其中 $x_0 \in [a, b]$, 而 $\bar{c}_1, \bar{c}_2$ 为常数. 最后得 $L(D)y(x)=f(x)$ 的通解为

$$y(x) = \bar{c}_1 y_1(x) + \bar{c}_2 y_2(x) + \int_{x_0}^x H(x, t) f(t) dt,$$

其中 $H(x, t) = \frac{\begin{vmatrix} y_1(t) & y_2(t) \\ y_1(x) & y_2(x) \end{vmatrix}}{W(t)}$ 称为格林函数, 而 $y(x) = \int_{x_0}^x H(x, t) dt$ 是

$L(D)y(x)=f(x)$ 的满足初始条件 $\begin{cases} y(x_0)=0 \\ y'(x_0)=0 \end{cases}$ 的特解.

例 22 求解方程 $y'' + y' - 2y = \frac{e^x}{1+e^x}$.

解 齐次方程 $y'' + y' - 2y = 0$ 的通解为 $y(x) = C_1 e^x + C_2 e^{-2x}$.

令 $y(x) = C_1(x)e^x + C_2(x)e^{-2x}$, 于是求满足 (5.1.44) 的 $C_1(x), C_2(x)$, 得到方程组

$$\begin{cases} C_1'(x)e^x + C_2'(x)e^{-2x} = 0, \\ C_1'(x)e^x - 2C_2'(x)e^{-2x} = \frac{e^x}{1+e^x}, \end{cases}$$

解出

$$C_1'(x) = \frac{1}{3(1+e^x)}, C_2'(x) = \frac{-e^{3x}}{3(1+e^x)},$$

积分后得到

$$\begin{cases} C_1(x) = c_1 + \frac{1}{3}x - \frac{1}{3}\ln(1+e^x), \\ C_2(x) = c_2 - \frac{1}{6}e^{2x} + \frac{1}{3}e^x - \frac{1}{3}\ln(1+e^x), \end{cases}$$

最后,得方程的解为

$$y(x) = c_1 e^x + c_2 e^{-2x} + \frac{1}{3}x e^x - \frac{1}{3}e^x \ln(1+e^x) - \frac{1}{6} + \frac{1}{3}e^{-x} - \frac{1}{3}e^{-2x} \ln(1+e^x).$$

(3) 欧拉方程

形如

$$a_n x^n y^{(n)} + a_{n-1} x^{n-1} y^{(n-1)} + \cdots + a_1 x y' + a_0 y = f(x) \quad (5.1.45)$$

的方程,其中 $a_n, a_{n-1}, \cdots, a_0$ 为常数,称为欧拉方程.欧拉方程可经过自变量变换

$x = e^t$ 化为常系数线性方程.事实上,若令 $D = \frac{d}{dt}, D^k = \frac{d^k}{dt^k}$,在变换 $x = e^t$

下,我们有

$$\begin{aligned} \frac{dy}{dx} &= \frac{1}{x} D y, \frac{d^2 y}{dx^2} = \frac{1}{x^2} D(D-1)y, \cdots, \\ \frac{d^k y}{dx^k} &= \frac{1}{x^k} D(D-1)\cdots(D-k)y, k=1,2,\cdots,n. \end{aligned}$$

从而,(5.1.45)化为

$$a_n D(D-1)\cdots(D-n)y + \cdots + a_1 D y + a_0 y = f(e^t).$$

例 23 求解方程 $x^3 y''' + 3x^2 y'' + x y' - y = 13x^{\frac{3}{2}}$

解 令 $x = e^t$, 方程化为

$$D(D-1)(D-2)y + 3D(D-1)y + (D-1)y = 13e^{3t},$$

化简得 $(D^3 - 1)y = 13e^{3t}$. 此为常系数非齐次线性方程,我们省略求解步骤,只写出其解为

$$y(x) = c_1 x + \frac{1}{\sqrt{x}} \left(c_2 \cos \frac{\sqrt{3}}{2} \ln x + c_3 \sin \frac{\sqrt{3}}{2} \ln x \right) + \frac{x^{\frac{3}{2}}}{2}.$$

5.1.4 隐函数方程

1. 一阶隐函数方程

现在考虑一阶隐函数方程

$$F(x, y, y') = 0. \quad (5.1.46)$$

对于隐式方程(5.1.46),有两种情况.

(1) 如果(5.1.46)可以就 $\frac{dy}{dx}$ 解出,成为一个或几个显式方程

$$y' = f_j(x, y) \quad (j=1, 2, \dots, n).$$

积分这些一阶显式微分方程,就得到方程(5.1.46)的解.

例 24 求解方程 $y'^2 - (x+y)y' + xy = 0$.

解 方程左边可以分解因式,得

$$(y' - x)(y' - y) = 0.$$

从而得到两个方程 $y' = x, y' = y$, 解这两个方程,得到

$$y = \frac{1}{2}x^2 + c, y = ce^{-x}.$$

(2) 如果无法就 $\frac{dy}{dx}$ 解出,或计算很繁,这时我们采用引入参数的方法直接来解(5.1.46).

其基本想法是通过变量变换(引入参数)把原来方程化为导数已解出的、且可求积分的方程. 下面我们就几种典型类型加以讨论.

① 方程(5.1.46)不显含 x 或不显含 y , 即可以写成

$$F(y, y') = 0, \quad (5.1.47)$$

或

$$F(x, y') = 0, \quad (5.1.48)$$

的形式. 我们记 $p = \frac{dy}{dx}$. 这时有 $dy - p dx = 0$.

如果能找到(5.1.47)的参数表达式

$$\begin{cases} p = \varphi(t), \\ y = \psi(t). \end{cases}$$

那么代入 $dy - p dx = 0$, 即得

$$\psi'(t)dt - \varphi(t)dx = 0, \quad \text{或} \quad \frac{dx}{dt} = \frac{\psi(t)'}{\varphi(t)}, \quad (5.1.49)$$

这是一个可求积的方程. 假设已求出(5.1.49)的通积分 $\Phi(x, t, c) = 0$, 则

$$y = \psi(t), \Phi(x, t, c) = 0$$

便是(5.1.47)的通积分表达式.

对方程(5.1.48), 可作类似的处理.

例 25 求 $4(1-y) = (3y-2)^2 y'^2$ 的通解.

解 令 $y = 1 - t$, 则由 $4t = (1-3t)^2 p^2$ 解得 $p = \pm \frac{2\sqrt{t}}{1-3t}$, 故原方程有下列

参数形式

$$\begin{cases} y = 1 - t, \\ p = \pm \frac{2\sqrt{t}}{1-3t}. \end{cases}$$

代入 $dy - p dx = 0$, 得

$$dx = \pm \frac{1-3t}{2\sqrt{t}} dt \quad (t = 1 - y \neq 0).$$

解得

$$x + c = \pm (t^{\frac{1}{2}} - t^{\frac{3}{2}}),$$

即

$$(x + c)^2 = t(1 - t)^2.$$

将 $t = 1 - y$ 代入上式, 得 $y^2(1 - y) = (x + c)^2$ 为原方程的通解.

此外, $y = 1$ 也是原方程的一个解.

例 26 求解 $x \sqrt{1 + y'^2} = y'$.

解 令 $y' = \tan t$, 则由 $x \sqrt{1 + y'^2} = y'$, 知 $x = \sin t$. 于是

$$\begin{cases} p = \tan t, \\ x = \sin t \end{cases}$$

为原方程的参数表达式, 代入 $dy - p dx = 0$, 得

$$dy - \tan t \cdot \cos t dt = 0,$$

即

$$dy - \sin t dt = 0,$$

解得 $y = -\cos t + c$. 从而通解为

$$\begin{cases} x = \sin t, \\ y = -\cos t + c, \end{cases} \quad \text{或} \quad x^2 + (y - c)^2 = 1.$$

从上面的例子我们看出, 对(5.1.47)与(5.1.48)的求解, 关键是找隐方程的参数表达式.

② 方程(5.1.46)可就 x 解出而写成

$$x = f(y, p), \quad (5.1.50)$$

或可就 y 解出而写成

$$y = g(x, p). \quad (5.1.51)$$

对方程(5.1.50), 可视 y, p 为参数, 将其代入 $dy - p dx = 0$ 中, 得到

$$dy - p \left(\frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial p} dp \right) = 0. \quad (5.1.52)$$

这是关于变量 y, p 的可解出 $\frac{dp}{dy}$ 的方程. 假如能求得通解

$$p = p(y, c),$$

代入(5.1.50)中,即得(5.1.50)的通解

$$x = f(y, p(y, c)), \text{ 或 } y = \tilde{f}(x, c).$$

对方程(5.1.51)方法是类似的.

例 27 求解方程 $y + xy' = x^4 y'^2$.

解 将 $y = x^4 p^2 - xp$ 代入 $dy - p dx = 0$ 中得到

$$4x^3 p^2 dx + 2x^4 p dp - p dx - x dp - p dx = 0,$$

即

$$(2x^3 p - 1)(2p dx + x dp) = 0.$$

由 $2x^3 p - 1 = 0$ 可得 $p = \frac{1}{2x^3}$, 代入 $y = x^4 p^2 - xp$, 得

$$y = -\frac{1}{4}x^{-2},$$

此为原方程的一个特解. 由 $2p dx + x dp = 0$ 解得 $p = \frac{c}{x^2}$, 代入 $y = x^4 p^2 - xp$, 得

$$y = c^2 - \frac{c}{x},$$

此为原方程的通解.

作为(5.1.51)的一个重要类型,我们来研究克莱罗(Clairet)方程

$$y = xy' + \varphi(y'). \quad (5.1.53)$$

由 $p = y'$, 先算出

$$dy = p dx + x dp + \varphi'(p) dp = (x + \varphi'(p)) dp + p dx.$$

代入 $dy - p dx = 0$ 得到

$$(p + \varphi'(p)) dp = 0. \quad (5.1.54)$$

对方程(5.1.54)由 $dp = 0$ 积分, 可得 $p = c$, 从而

$$y = cx + \varphi(c)$$

为(5.1.53)的通解, 它表示一族直线.

由 $x + \varphi'(p) = 0$ 得到(5.1.53)的一个特解

$$\begin{cases} x = -\varphi'(p), \\ y = \varphi(p) - p\varphi'(p). \end{cases} \quad (5.1.55)$$

容易看出: 曲线(5.1.55)与(5.1.53)中的每一条直线至少有一公共点

$$(-\varphi'(c), \varphi(c) - c\varphi'(c)),$$

且在该点相切. 我们称满足这种性质的曲线族(5.1.55)为直线族(5.1.54)的包络.

此外, 过曲线(5.1.55)上的每一点, 方程(5.1.53)至少有二个解. 换句话说, 方程(5.1.53)过曲线(5.1.55)上的点的解不惟一, 我们称这种在其上处处破坏

惟一性的解(5.1.55)为方程(5.1.53)的奇解.

例 28 求一曲线,使它的任一切线与两坐标所围成的三角形面积等于常数 a .

解 设 (x, y) 为所求曲线上的点,则过 (x, y) 的切线方程为

$$Y - y = y'(X - x).$$

该切线在 x, y 轴上的截距分别为 $a = \frac{y'x - y}{y'}$, $b = y - xy'$. 当 $a, b > 0$ 时,有

$$\frac{y'x - y}{y'} \times (y - xy') = 2a,$$

即

$$2ay' + (xy' - y)^2 = 0.$$

解出 y , 得到克莱罗方程

$$y = xy' \pm 2\sqrt{-ay'}.$$

由前面关于克莱罗方程的讨论,知

$$y = cx \pm 2\sqrt{-ax} \quad (c < 0)$$

为所求通解. 显然它是一族直线.

此外,由(5.1.55)式,它的特解由

$$\begin{cases} x = \pm \frac{a}{\sqrt{-ac}}, \\ y = -cx \pm \sqrt{-ac}, \end{cases}$$

或

$$\begin{cases} x = \pm \frac{a}{\sqrt{-ac}}, \\ y = \pm \sqrt{-ac} \end{cases}$$

所确定. 消去 c , 得 $xy = a$ 为所求曲线, 这是我们熟知的双曲线.

当 $a, b < 0$ 时, 可求得直线族

$$y = cx \pm \sqrt{ac} \quad (c > 0)$$

及奇解曲线

$$xy = -a.$$

2. 高阶隐函数方程

对高阶隐函数方程

$$F(x, y, y', \dots, y^{(n)}) = 0, \quad (5.1.56)$$

我们通常采用降阶法与引入参数法来求解. 下面介绍几种最主要的求解方法.

(1) F 中不显含 $y, y', \dots, y^{(n-k)}$ 情形

$$F(x, y^{(k)}, \dots, y^{(n)}) = 0. \quad (5.1.57)$$

这类方程的特点是出现的最低阶导数为 $y^{(k)}$. 作代换 $z = y^{(k)}$, 方程(5.1.57)被降为 $(n-k)$ 阶方程

$$F(x, z, \dots, z^{(n-k)}) = 0. \quad (5.1.58)$$

如果(5.1.58)能求出通解

$$z = \varphi(x, c_1, \dots, c_{n-k}),$$

那么, 对

$$y^{(k)} = \varphi(x, c_1, \dots, c_{n-k})$$

积分 k 次, 就可解出 y .

(2) F 中不显含 x 情形

$$F(y, y', \dots, y^{(n)}) = 0. \quad (5.1.59)$$

这类方程的特点是不显含自变量 x . 这时可以视 y 为自变量, 取 $y' = p$ 为未知函数而使方程(5.1.59)降阶一次. 以二阶方程 $F(y, y', y'') = 0$ 为例, 由 $y' = p$ 可得

$$y'' = \frac{dp}{dx} = \frac{dp}{dy} \cdot \frac{dy}{dx} = p \frac{dp}{dy}.$$

代入原方程, 就有

$$F(y, p, p \frac{dp}{dy}) = 0.$$

这是关于新未知函数 p 的一阶方程, 如它有通解

$$p = p(y, c),$$

则

$$y' = p = p(y, c),$$

或

$$\frac{dy}{p(y, c)} = dx.$$

两边求积分, 可求得原方程的通解.

例 29 求解方程 $y''' + 2y'' + 2x = 0$.

解 这个方程是常系数线性齐次方程, 可以用关于常系数线性方程方法求解, 但也可用降阶法求解. 令 $z = y''$, 则

$$\frac{dz}{dx} + 2z + 2x = 0.$$

解得 $z = c_1 e^{-2x} - x + \frac{1}{2}$. 由 $y'' = c_1 e^{-2x} - x + \frac{1}{2}$ 积分两次, 得

$$y(x) = c_1 e^{-2x} - \frac{1}{6}x^3 + \frac{1}{4}x^2 + c_2 x + c_3,$$

即为所求通解.

例 30 求解 $y''y - (y')^2 = 0$.

解 改取 y 为自变量, $p = y'$ 为未知函数, 则 $y'' = p \frac{dp}{dy}$. 原方程变为

$$py \frac{dp}{dy} - p^2 = p \left(y \frac{dp}{dy} - p \right) = 0,$$

由 $p = 0$ 可得含一个任意常数的解 $y = c_1$. 由 $y \frac{dp}{dy} - p = 0$, 得

$$\frac{dp}{p} = \frac{dy}{y},$$

积分得 $p = c_2 y$, 其中 $c_2 \neq 0$. 以 $p = \frac{dy}{dx}$ 代入, 得

$$\frac{dy}{dx} = c_2 y,$$

即

$$\frac{dy}{y} = c_2 dx.$$

对上式积分, 得

$$y = c_3 e^{c_2 x}.$$

若允许 $c_2 = 0$, 则 $y = c_3 e^{c_2 x}$ 便包含了特解 $y = c_1$, 故 $y = c_3 e^{c_2 x}$ 为所求通解, 其中 c_2, c_3 为任意常数.

3. 全微分方程与积分因子

如果方程(5.1.56)的左边是某一个 $n-1$ 阶微分表示式 $\Phi \left(x, y, y', \dots, \frac{d^{n-1}y}{dx^{n-1}} \right)$ 对 x 的全导数, 即

$$F \left(x, y, \frac{dy}{dx}, \dots, \frac{d^n y}{dx^n} \right) = \frac{d}{dx} \Phi \left(x, y, y', \dots, \frac{d^{n-1}y}{dx^{n-1}} \right),$$

我们就称(5.1.56)是全微分方程. 这时, 方程

$$\Phi \left(x, y, \frac{dy}{dx}, \dots, \frac{d^{n-1}y}{dx^{n-1}} \right) = 0 \quad (5.1.60)$$

的每一个解都是方程(5.1.56)的解. 方程(5.1.60)是 $n-1$ 阶的, 这样就降低了方程的(5.1.56)的阶数.

有时, 方程(5.1.56)本身不是全微分方程, 但是乘上某个适当的函数 $\mu(x, y)$ 后, 能化为全微分方程, 这时称函数 $\mu(x, y)$ 是(5.1.56)的积分因子.

例 31 求解方程 $xy'' + 3y' = 4x^3$.

解 先将方程两边同乘不为 0 的因子 x^2 , 则有

$$x''^3 + 3x^2 y' = 4x^5,$$

即

$$d(x^3 y' - \frac{2}{3} x^6) = 0.$$

故

$$x^3 y' - \frac{2}{3} x^6 = c_1,$$

从而

$$y = \frac{1}{6} x^4 - 2c_1 x^2 + c_2$$

为所求通解.

5.1.5 存在性与惟一性定理

本节我们研究微分方程解的存在性与惟一性. 主要对下列一阶显方程

$$\frac{dy}{dx} = f(x, y) \quad (5.1.61)$$

进行讨论.

定理 5.1.10 (存在与惟一性定理) 假设

(1) $f(x, y)$ 在闭矩形域 D

$$|x - x_0| \leq a, |y - y_0| \leq b \quad (5.1.62)$$

中连续, 记

$$M = \max_{(x, y) \in D} |f(x, y)|, h = \min\left(a, \frac{b}{M}\right).$$

(2) $f(x, y)$ 关于 y 满足利普希茨 (Lipshitz) 条件: 存在常数 $N > 0$, 使得对适合 $|x - x_0| \leq a$ 的任何 x , 适合 $|y - y_0| \leq b$ 的 y 的任意两个值 y_1, y_2 都有

$$|f(x, y_1) - f(x, y_2)| \leq N |y_1 - y_2|. \quad (5.1.63)$$

则方程 (5.1.61) 在区间 $|x - x_0| \leq h$ 上存在惟一的、满足初值条件

$$y(x_0) = y_0 \quad (5.1.64)$$

的解 $y^*(x)$.

证 我们首先证明, 求方程 (5.1.61) 的、满足初值条件 (5.1.64) 的解 $y^*(x)$, $x_0 - h \leq x \leq x_0 + h$, 等价于求积分方程

$$y(x) = y_0 + \int_{x_0}^x f(\eta, y(\eta)) d\eta \quad (5.1.65)$$

在区间 $x_0 - h \leq x \leq x_0 + h$ 上的连续解.

事实上, 如果 $y = y^*(x)$ 是方程 (5.1.61) 的解, 且 $y^*(x_0) = y_0$, 从 x_0 到 x 积分恒等式

$$\frac{dy^*}{dx} = f(x, y^*(x)),$$

就得到

$$y^* = y_0 + \int_{x_0}^x f(\eta, y^*(\eta)) d\eta,$$

即 $y^*(x)$ 适合积分方程(5.1.65). 另一方面, 假若连续函数 $y^*(x)$ 适合积分方程(5.1.65), 显然 $y^*(x_0) = y_0$. 又由于 $f(x, y^*(x))$ 是连续函数, 故由(5.1.65)所确定的函数 $y^*(x)$ 有连续的导数, 对(5.1.65)两边求微分, 就得到

$$\frac{dy^*}{dx} = f(x, y^*(x)),$$

即 $y^*(x)$ 满足方程(5.1.61)及初值条件(5.1.64).

因此, 下面只要证明积分方程(5.1.65)的连续解存在并且是惟一的就够了.

我们采用皮卡(Picard)逐次逼近法证明积分方程(5.1.65)的连续解的存在性. 证明分三步.

第一步, 在区间 $|x - x_0| \leq h$ 上构造一个连续函数序列 $\{y_n(x)\}$, 称为逐次逼近序列; 第二步, 证明 $\{y_n(x)\}$ 在区间 $|x - x_0| \leq h$ 上一致收敛; 第三步, 证明 $\{y_n(x)\}$ 的极限函数 $y^*(x)$ 满足方程(5.1.65).

步1 以 y_0 代替(5.1.65)右边 $f(x, y(x))$ 中的 $y(x)$. 记所得到的函数为 $y_1(x)$, 即

$$y_1(x) = y_0 + \int_{x_0}^x f(\eta, y_0) d\eta,$$

称它为第一次近似. 显然 $y_1(x)$ 在区间 $|x - x_0| \leq h$ 上定义且连续. 当 $|x - x_0| \leq h$ 时, 我们有

$$|y_1(x) - y_0| \leq \left| \int_{x_0}^x |f(\eta, y_0)| d\eta \right| \leq M |x - x_0| \leq M \cdot \frac{b}{M} = b, \quad (5.1.66)$$

所以在区间 $|x - x_0| \leq h$ 上, $f(x, y(x))$ 有意义. 现在再以 $y_1(x)$ 代替(5.1.65)式右边 $f(x, y(x))$ 中的 $y(x)$, 得到的函数记为 $y_2(x)$, 即

$$y_2(x) = y_0 + \int_{x_0}^x f(\eta, y_1(\eta)) d\eta,$$

称它为第二次近似. 显见 $y_2(x)$ 在区间 $|x - x_0| \leq h$ 上有定义且连续, 又

$$|y_2(x) - y_0| \leq M |x - x_0| \leq Mh \leq b,$$

故在区间 $|x - x_0| \leq h$ 上, $f(x, y_2(x))$ 有意义; 继续进行下去, 假设 $y_1(x), y_2(x), \dots, y_n(x)$ 已经取定, 且当 $|x - x_0| \leq h$ 时, $|y_n(x) - y_0| \leq b$. 以 $y_n(x)$ 代替方程(5.1.65)右边 $f(x, y(x))$ 中的 $y(x)$, 把所得到的函数记为 $y_{n+1}(x)$, 即

$$y_{n+1}(x) = y_0 + \int_{x_0}^x f(\eta, y_n(\eta)) d\eta. \quad (5.1.67)$$

显然 $y_{n+1}(x)$ 在区间 $|x - x_0| \leq h$ 上有定义、连续且

$$|y_{n+1}(x) - y_0| \leq M|x - x_0| \leq Mh \leq b, \quad (5.1.68)$$

从而当 $|x - x_0| \leq h$ 时, $f(x, y_{n+1}(x))$ 仍有意义. 由数学归纳法, 我们得到了一个确定的、在区间 $|x - x_0| \leq h$ 上的连续函数序列

$$y_1(x), y_2(x), \dots, y_n(x), \dots.$$

步 2 要证 $\{y_n(x)\}$ 在区间 $|x - x_0| \leq h$ 上一致收敛, 只需证明级数

$$y_0 + (y_1(x) - y_0) + (y_2(x) - y_1(x)) + \dots + (y_n(x) - y_{n-1}(x)) + \dots \quad (5.1.69)$$

在区间 $|x - x_0| \leq h$ 上一致收敛.

不等式(5.1.66)在区间 $|x - x_0| \leq h$ 上给出, 则

$$|y_1(x) - y_0| \leq M|x - x_0|. \quad (5.1.70)$$

由 $y_1(x)$ 与 $y_2(x)$ 的定义及(5.1.70), 有

$$\begin{aligned} |y_2(x) - y_1(x)| &= \left| \int_{x_0}^x (f(\eta, y_1(\eta)) - f(\eta, y_0)) d\eta \right| \\ &\leq \int_{x_0}^x |f(\eta, y_1(\eta)) - f(\eta, y_0)| d\eta \\ &\leq N \int_{x_0}^x |y_1(\eta) - y_0| d\eta \leq MN \left| \int_{x_0}^x |\eta - x_0| d\eta \right| \\ &\leq \frac{MN}{2!} |x - x_0|^2. \end{aligned}$$

一般地, 如果已经证明了

$$|y_n(x) - y_{n-1}| \leq \frac{MN^{n-1}}{n!} |x - x_0|^n, \quad (5.1.71)$$

则

$$\begin{aligned} |y_{n+1}(x) - y_n(x)| &\leq \left| \int_{x_0}^x |f(\eta, y_n(\eta)) - f(\eta, y_{n-1}(\eta))| d\eta \right| \\ &\leq N \left| \int_{x_0}^x |y_n(\eta) - y_{n-1}(\eta)| d\eta \right| \\ &\leq \frac{MN^n}{n!} \left| \int_{x_0}^x |\eta - x_0|^n d\eta \right| = \frac{MN^n}{(n+1)!} |x - x_0|^{n+1}. \end{aligned}$$

据数学归纳法知, 不等式(5.1.71)在区间 $|x - x_0| \leq h$ 上对一切正整数 n 成立.

由此可见, 级数(5.1.69)的每一项的绝对值不超过级数

$$|y_0| + \sum_{n=0}^{\infty} \frac{MN^n}{(n+1)!} |x - x_0|^{n+1} \quad (5.1.72)$$

的对应项. 但当 $|x - x_0| \leq h$ 时, 级数(5.1.72)是一致收敛的, 所以级数

(5.1.69)在区间 $|x - x_0| \leq h$ 上也是一致收敛的. 假定它的极限函数为 $y^*(x)$, 因为(5.1.69)的每一项在 $|x - x_0| \leq h$ 上连续, 所以 $y^*(x)$ 在 $|x - x_0| \leq h$ 上也连续.

步3 证明 $y^*(x)$ 满足方程(5.1.65). 由不等式(5.1.68), 取极限得

$$|y^*(x) - y_0| \leq b.$$

因为 $f(x, y^*(x))$ 有定义且连续, 又因为

$$|f(x, y_n(x)) - f(x, y^*(x))| \leq N|y_n(x) - y^*(x)|,$$

所以当 $n \rightarrow \infty$ 时, $f(x, y_n(x))$ 在区间 $|x - x_0| \leq h$ 上一致收敛到 $f(x, y^*(x))$. 从而在等式(5.1.67)中令 $n \rightarrow \infty$ 时, 可以通过积分号求极限, 而得到

$$y^*(x) = y_0 + \int_{x_0}^x f(\eta, y^*(\eta)) d\eta,$$

即 $y^*(x)$ 是积分方程(5.1.65)的连续解.

下面, 我们来证明(5.1.65)在区间 $|x - x_0| \leq h$ 上的连续解是惟一的.

假设除 $y^*(x)$ 外, 方程(5.1.65)还有另一个连续解 $\tilde{y}(x)$. 由

$$y_0(x) = y_0,$$

$$y_n(x) = y_0 + \int_{x_0}^x f(\eta, y_{n-1}(\eta)) d\eta \quad (n = 1, 2, \dots),$$

$$\tilde{y}(x) = y_0 + \int_{x_0}^x f(\eta, \tilde{y}(\eta)) d\eta,$$

可得

$$\begin{aligned} |\tilde{y}(x) - y_0(x)| &\leq \left| \int_{x_0}^x f(\eta, \tilde{y}(\eta)) d\eta \right| \leq M|x - x_0|, \\ |\tilde{y}(x) - y_1(x)| &\leq \left| \int_{x_0}^x |f(\eta, \tilde{y}(\eta)) - f(\eta, y_0(\eta))| d\eta \right| \\ &\leq N \left| \int_{x_0}^x |\tilde{y}(\eta) - y_0(\eta)| d\eta \right| \\ &\leq NM \left| \int_{x_0}^x |\eta - x_0| d\eta \right| = \frac{NM}{2!} |x - x_0|^2. \end{aligned}$$

与证明存在性时一样, 由数学归纳法可以证明不等式

$$|\tilde{y}(x) - y_n(x)| \leq \frac{MN^2}{(n+1)!} |x - x_0|^{n+1} \leq \frac{Mh}{(n+1)!} (Nh)^n$$

对一切正整数 n 成立. 上面不等式右端当 $n \rightarrow \infty$ 时趋于零, 因此, 当 $|x - x_0| \leq h$ 时, 就有

$$\lim_{n \rightarrow \infty} y_n(x) = \tilde{y}(x).$$

注意 $\lim_{n \rightarrow \infty} y_n(x) = y^*(x)$, 所以在区间 $|x - x_0| \leq h$ 上, $y^*(x) \equiv \tilde{y}(x)$. 定理证毕.

在结束这一节之前,我们再对定理的条件与结论作些说明.

注 1 在实际应用时,利普希茨条件的检验是比较费事的.我们能够用一个较强的,但易于验证的条件来替代它,即如果函数 $f(x, y)$ 在闭矩形域 D 上关于 y 的偏导数 $f'_y(x, y)$ 存在且有界,即 $|f'_y(x, y)| \leq N$, 则利普希茨条件成立. 如果 $f'_y(x, y)$ 在 D 上连续,它在 D 上必有界,从而也满足利普希茨条件.

注 2 定理中的数 h 有下列几何解释:由导数的几何意义,存在与惟一性定理就是要证明,存在惟一的一条经过点 (x_0, y_0) 的积分曲线

$$y = y^*(x),$$

使其上每一点 (x, y) 处的切线斜率为 $f(x, y)$. 现在如定理的假设, $f(x, y)$ 在 D 中最大值为 $M = \max_{(x, y) \in D} |f(x, y)|$, 于是,如果从 (x_0, y_0) 引两条斜率分别等于 M 与 $-M$ 的直线,则积分曲线 $y = y^*(x)$ (如果存在的话)一定位于图 5.1.5 中带阴影的两个区域内. 因此,我们只要取

$$h = \min\left(a, \frac{b}{M}\right),$$

则过点 (x_0, y_0) 的积分曲线 $y = y^*(x)$ 就能确保当 $|x - x_0| \leq h$ 时,位于 D 中.

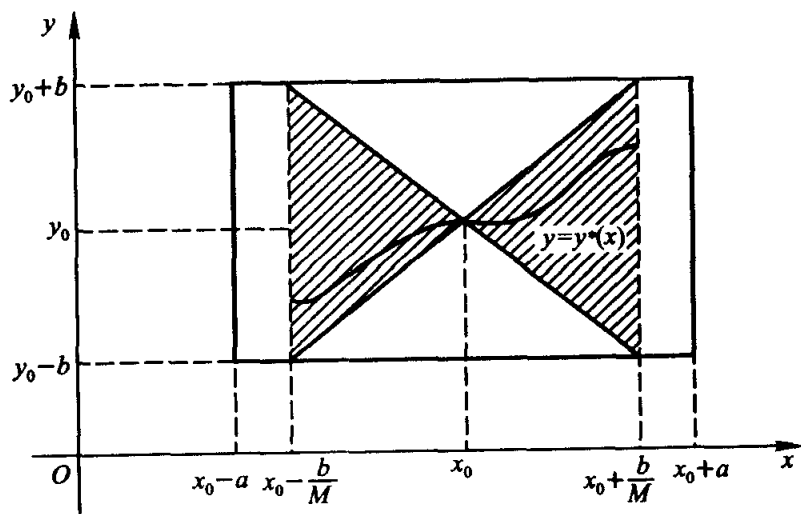


图 5.1.5

5.1.6 常微分方程数值解法

1. 建立数值方法的基本思想与途径

以一阶微分方程的初值问题

$$\begin{cases} y' = f(t, y) & (a \leq t \leq b) \\ y(a) = y_0 \end{cases} \quad (5.1.73)$$

为例来说明建立数值方法的基本思想.

求初值问题(5.1.73)的数值解,就是要求在区间 $[a, b]$ 上的若干个离散的点处,例如

$$a = t_0 < t_1 < \cdots < t_n = b,$$

计算出解方程(5.1.73)的解 $y(t)$ 的近似值

$$y_1, y_2, \cdots, y_n.$$

一般常取 $t_0, t_1, \cdots, t_n$ 为等距离的,即

$$t_j - t_{j-1} = h \quad (j = 1, 2, \cdots, n),$$

或

$$t_i = a + ih \quad (i = 0, 1, \cdots, n).$$

我们称 h 为步长.

建立数值方法的第一步就是离散化过程,即通过一定的方法把初值问题(5.1.73)化为在给定的 $n+1$ 个点上的近似的差分方程的初值问题.这里我们简单地介绍几种离散化过程,也称为差分方法.

(1) 化导数为差商,求解差分方程

在点 t_j 处的导数 $y'(t_j)$ 可以近似地表示成差商

$$y'(t_j) \approx \frac{y_{j+1} - y_j}{h},$$

从而就把微分方程初值问题(5.1.73)化为差分方程的初值问题

$$\begin{cases} \frac{y_{j+1} - y_j}{h} = f(t_j, y_j) & (j = 0, 1, 2, \cdots, n), \\ y_0 = y(a), \end{cases} \quad (5.1.74)$$

其中 y_j 表示解 $y(t)$ 在点 t_j 处的近似值,即 $y_j \approx y(t_j)$.

由(5.1.74),我们得到下列欧拉公式

$$\begin{cases} y_{j+1} = y_j + hf(t_j, y_j) \\ y_0 = y(a), \end{cases} \quad (5.1.75)$$

在公式(5.1.75)中,若已知 y_0 ,便可由递推公式(5.1.75)中的第一式求出 y_1 ,然后由 y_1 求出 y_2 ,依次类推,由 y_n 求出 y_{n+1} .为求后一点 t_{j+1} 上的数值 y_{j+1} ,只要知道前一点 t_j 上的数值 y_j 就够了,这个方法称为“单步法”.由欧拉公式确定的方法就是单步法.

一般地,如果给定一种方法,它在计算 y_{n+k} 时,需要知道以前的 y_{n+k-1} , y_{n+k-2} , $\cdots$, y_n 的 k 个函数值,则称这种方法为 k 步法.当 $k=1$ 时,就是单步法,而当 $k>1$ 时,通常叫做多步法.例如,由下列公式

$$\begin{aligned} & a_k y_{n+k} + a_{k-1} y_{n+k-1} + \cdots + a_0 y_n \\ & = h \{ \beta_k f(t_{n+k}, y_{n+k}) + \beta_{k-1} f(t_{n+k-1}, y_{n+k-1}) + \cdots + \beta_0 f(t_n, y_n) \} \end{aligned}$$

所确定的方法即为 k 步法.

(2) 利用泰勒展式, 求解差分方程

将 $y(t_j + h)$ 展成 t_j 点处的泰勒多项式, 并取其有限项作为 $y(t_j + h)$ 的近似值

$$y(t_j + h) \approx y(t_j) + hy'(t_j) + \cdots + \frac{h^p}{p!} y^{(p)}(t_j), \quad (5.1.76)$$

其中 p 为一正整数, 通过微分方程

$$y' = f(t, y)$$

可以逐次把各阶导数 $y', y'', \cdots$ 在 t_j 处的值表示出来

$$\begin{aligned} y'(t_j) &= f(t_j, y(t_j)), \\ y''(t_j) &= f'_t(t_j, y(t_j)) + f'_y(t_j, y(t_j))y'(t_j), \\ &\cdots, \end{aligned}$$

从而可知, 一阶与阶导数的近似值近似地满足

$$\begin{cases} y'_j = f(t_j, y_j), \\ y''_j = f'_t(t_j, y_j) + f'_y(t_j, y_j)y'_j, \end{cases} \quad j=0, 1,$$

其中 y_j 表示 $y(t_j)$ 的近似值, 而 $y', y'', \cdots$ 分别表示各阶导数 $y'(t_j), y''(t_j), \cdots$ 的近似值. 因此, 可以把微分方程初值问题(5.1.73)利用泰勒多项式(5.1.76)化为差分方程的初值问题

$$\begin{cases} y_{j+1} = y_j + hy'_j + \frac{h^2}{2!} y''_j + \cdots + \frac{h^p}{p!} y^{(p)}_j & (j=0, 1, \cdots, n), \\ y_0 = y(a). \end{cases} \quad (5.1.77)$$

若取 $p=1$, 则(5.1.77)成为

$$\begin{cases} y_{j+1} = y_j + hy'_j = y_j + hf(t_j, y_j), \\ y_0 = y(a), \end{cases}$$

它恰好是欧拉公式(5.1.75).

(3) 利用数值积分, 求解差分方程

在区间 $[t_j, t_{j+1}]$ 上, 将微分方程(5.1.68)化成下列积分方程

$$y(t_{j+1}) = y(t_j) + \int_{t_j}^{t_{j+1}} f(t, y(t)) dt. \quad (5.1.78)$$

于是初值问题(5.1.76)首先可以近似地化为

$$y^{j+1} = y_j + \int_{t_j}^{t_{j+1}} f(t, y(t)) dt. \quad (5.1.79)$$

进一步, 我们用数值积分方法近似计算(5.1.79)右边的积分. 例如, 我们在区间 $[t_j, t_{j+1}]$ 上利用矩形公式

$$\int_{t_j}^{t_{j+1}} f(t, y(t)) dt \approx f(t_j, y_j) h,$$

便可得到

$$y_{j+1} = y_j + hf(t_j, y_j),$$

这又得到欧拉公式(5.1.75). 在 $[t_j, t_{j+1}]$ 上利用梯形公式

$$\int_{t_j}^{t_{j+1}} f(t, y(t)) dt \approx \frac{h}{2} [f(t_j, y_j) + f(t_{j+1}, y_{j+1})],$$

则可得到

$$y_{j+1} = y_j + \frac{h}{2} [f(t_j, y_j) + f(t_{j+1}, y_{j+1})], \quad (5.1.80)$$

我们称公式(5.1.80)为后退的欧拉公式.

欧拉公式(5.1.75)是一个显式公式, 因为所要求的量 y_{n+1} 可以用 t_n 与 y_n 求得的 $f_n = f(t_n, y_n)$ 来表示, 后退欧拉公式(5.1.75)则是一个隐式公式, 因为所需求的 y_{n+1} 也包含在 f_{n+1} 中, 当 f 为 y 的非线性函数时, 一般无法将 y_{n+1} 表示成 y_n 的函数.

2. 数值方法的误差与阶数

给定一种数值解法, 我们需要了解它的精确度, 即估计数值解与精确解的误差.

定义 5.1.7 (离散误差) 我们把用数值方法求得的解在 t_n 点的近似值 y_n 与微分方程的解 $y(t_n)$ 之间的差叫做该方法在 t_n 点的离散误差, 记为 e_n , 即

$$e_n = y_n - y(t_n),$$

它是因微分方程的离散化而产生的, 所以又称为“方法误差”.

定义 5.1.8 (舍入误差) 我们把由近似公式计算得到的值 $\tilde{y}_n$ 与方法的理论解 y_n 之间的差叫做舍入误差, 记为

$$r_n = \tilde{y}_n - y_n.$$

当用数值解逼近微分方程的真正解时, 数值解对真正解所达到的精度的另一个重要标志是“方法的阶”.

为了给出“阶”的定义, 我们把精确解 $y(t)$ 在 t_n 点展成带余项泰勒公式

$$y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2} y''(\xi_n),$$

其中 $\xi \in (t_n, t_{n+1})$, 注意到 $y'(t_n) = f(t_n, y(t_n))$, 所以

$$y(t_{n+1}) = y(t_n) + hf(t_n, y(t_n)) + \frac{h^2}{2} y''(\xi_n). \quad (5.1.81)$$

将(5.1.81)式与欧拉公式(5.1.75)比较便知, 欧拉公式是由精确解的泰勒展开

中截去 $\frac{h^2}{2}y''(\xi_n)$ 这一项而得到的近似公式. 我们称 $\frac{h^2}{2}y''(\xi_n)$ 为欧拉公式的截断误差, 它是与 h^2 同阶的量, 记为 $O(h^2)$.

由于欧拉公式丢掉的项含有 $y(t)$ 的二阶以上的导数, 所以欧拉方法是一个一阶方法.

定义 5.1.9 (方法的阶数) 所有不超过 s 次的多项式所精确满足的公式称为 s 阶方法.

3. 二阶龙格-库塔(Runge-Kutta)方法

龙格-库塔方法属于单步法, 它有各种阶数的公式, 也有显式与隐式之分, 这里, 仅就 2 阶显式龙格-库塔公式加以介绍.

由方法阶数的定义, 我们提出下列问题: 如何形成

$$y_{n+1} = y_n + hy'_n + \frac{h^2}{2}y''_n \quad (5.1.82)$$

而得到的二阶方法?

首先, 由微分方程(5.1.73), 我们有

$$y'\left(t_n + \frac{h}{2}\right) = f\left(t_n + \frac{h}{2}, y\left(t_n + \frac{h}{2}\right)\right).$$

利用泰勒公式, 又有

$$\begin{aligned} y\left(t_n + \frac{h}{2}\right) &= y(t_n) + \frac{h}{2}y'(t_n) + O(h^2), \\ y''\left(t_n + \frac{h}{2}\right) &= y''(t_n) + O(h^2). \end{aligned}$$

因此

$$y(t_n) + hy(t_n) + \frac{h^2}{2}y''(t_n) = y(t_n) + hf\left(t_n + \frac{h}{2}, y(t_n) + \frac{h}{2}y'(t_n)\right) + O(h^3).$$

以 y_n 近似表示 $y(t_n)$, 由(5.1.76), 并舍去 $O(h^3)$ 项, 得

$$y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}y'_n\right).$$

若将 $y'_n = f(t_n, y_n)$ 代入, 即得

$$y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, y_n + \frac{h}{2}f(t_n, y_n)\right). \quad (5.1.83)$$

由公式(5.1.83)给出的方法是二阶龙格-库塔方法的一种, 称为中点方法.

一般地, 二阶龙格-库塔方法推导如下.

将(5.1.83)改写成

$$\begin{cases} q_1 = y_n + \frac{h}{2}f(t_n, y_n), \\ y_{n+1} = y_n + hf\left(t_n + \frac{h}{2}, q_1\right). \end{cases}$$

由此启发我们考虑

$$\begin{cases} q_1 = y_n + \alpha h f(t_n, y_n), \\ y_{n+1} = y_n + \beta h f(t_n, y_n) + \gamma h f(t_n + \eta h, q_1). \end{cases} \quad (5.1.84)$$

我们希望 y_{n+1} 的展式尽可能与泰勒级数一致. 由 (5.1.84), 得

$$\begin{aligned} y_{n+1} = & y_n + (\beta + \gamma) y'_n h + (\alpha \gamma f'_{y'} y'_n + \gamma \eta f'_t) h^2 \\ & + \left[\frac{\alpha^2}{2} \gamma f''_{yy} (y'_n)^2 + \alpha \gamma \eta f''_{yt} y'_n + \frac{\gamma}{2} \eta^2 f''_{tt} \right] h^3 + O(h^4). \end{aligned}$$

整理各项使之与 t_n 点处的泰勒级数对应项一致, 得到

$$\beta + \gamma = 1, \quad \alpha = \eta, \quad 2\alpha\gamma = 1. \quad (5.1.85)$$

由 (5.1.84) 与 (5.1.85) 所描述的方法, 称为二阶龙格-库塔方法. 可以选取自由参数 γ 的值, 使得方法的某个需要的性质达到最佳. 例如, 若取 $\gamma = 1$, 则得中点

法 (5.1.83). 若取 $\gamma = \frac{1}{2}$, 则给出二阶龙格-库塔方法的另一种形式, 即汉恩

(Heun) 方法.

4. 数值方法的收敛性与稳定性

当我们选取一个方法时, 由此得到的数值解对微分方程的精确解所达到的精度可能依赖于一个或多个参数, 例如步长 h . 是否可以选这些参数来达到需要的精确度, 是数值方法中的一个重要问题. 为此我们引入收敛性的概念.

定义 5.1.10 (数值方法的收敛性) 对任何适定的微分方程 (即解存在、惟一、稳定), 通过选取足够小的步长 h , 可使数值解达到任意的精确度, 便称该方法是收敛的.

定义中提到的稳定性, 是指“当定解条件作微小变化时, 相应的解也作微小变化”. 这种稳定性也常称为解对定解条件的连续依赖性.

对具体问题, 上述定义可以叙述得更加精确. 例如, 对问题 (5.1.77), 当 $f(t, y)$ 关于 y 满足利普希茨条件时, 由欧拉公式给出的方法是收敛的.

给定一个适当的初值问题, 我们还需要知道, 一个数值方法的稳定性, 也就是初值的微小变化是否对由给定的方法得到的数值解只产生有界的变化. 对差分方法, 稳定性可以定义为: 如果对于每个微分方程均存在 $h_0 > 0$, 使得初值中 小于一个固定量的变化在所有 $0 \leq h \leq h_0$ 的数值解中产生有界的变化, 则该方法 是稳定的.

习题

1. 解下列方程

$$(1) \quad xy' + (x+1)y = 4x^3 e^{-x}.$$

$$(2) \quad y' - \frac{1}{x-1}y = 3(x-1)^3$$

(3) $y' + y \tan x = \sec x$

(4) $x(y' - y) = e^y$

2. 解下列贝努里方程

(1) $\frac{dy}{dx} - \frac{1}{6}y = \frac{1}{6}xy^7$

(2) $y' + 2xy + \frac{1}{3}xy^4 = 0$

3. 求解 $\frac{dy}{dx} = \frac{1}{x^3 y^3 + xy}$ (提示: 将 x 看成 y 的函数).4. 求方程 $y' + x(y-x) + x^3(y-x)^2 = 1$ 满足 $y(0) = \frac{1}{2}$ 的特解 (提示: 令 $u = y - x$).5. 求方程 $y' = y^2 - 4x^2 + 2$ 的通解.

6. 试求下列各方程的通解

(1) $y'' + 7y' + 12y = 0$

(2) $y'' - 4y' + 4y = 0$

(3) $y^{(4)} - y'' = 0$

(4) $y^{(4)} + y = 0$

(5) $y''' - 2y'' + 2y' - y = 0$

(6) $y^{(6)} + 2y^{(4)} - y'' - 2y = 0$

7. 试求下列各方程满足给定的初始条件的解

(1) $y'' - 3y' + 2 = 0, y(0) = 2, y'(0) = -1$

(2) $y''' + y' = 0, y(0) = 1, y'(0) = -1, y''(0) = 0$

8. 求下列方程的通解

(1) $(D^2 - D - 2)y = 3e^x$

(2) $D^2 - 4D + 4 = e^{2x}$

(3) $(D^2 + 1)y = \cos x + \sin 2x$

(4) $(D^2 + 6D + 10)y = e^{-3x} \sin x$

(5) $(3D^2 + 2D + 1)y = 2x^2 + x + 1$

9. 求解下列方程

(1) $(xy''' - y'')^2 = y''^2 - 1$

(2) $xy'' + (x^2 - 1)(y' - 1) = 0$

(3) $a^2 y''^2 = (1 + y'^2)^3$

(4) $yy'' - y'^2 - y^2 y' = 0$

(5) $yy'' + y'^2 + 1 = 0$

(6) $3y''^2 - y' y''' = 0$

(7) $yy'' - y'^2 - 6xy^2 = 0$

(8) $x^2 yy'' - x^2 y'^2 - 4xyy' + 8y^2 = 0$

(9) $(y-x)y'' + y'^2 - y' = 1 - \sin x$

(10) $(x+y^2)y''' + 6yy'y'' + 3y'' + 2y'^3 = 0$

(11) $x^2 yy'' + (xy' - y)^2 = 0$ (提示: 令 $y = xz$)

(12) $\sin^2 xy'' - 2y = 0$ (提示: 令 $\cot x = z$)

(13) $x(yy'' + y'^2) + 3yy' = 2x^3$ (提示: 令 $y^2 = z$)

10. 一水槽内盛满酸性溶液, 水槽体积为 V , 注入清水以减弱酸性, 但却随时保持溶液均匀且体积 V 不变. 设在某一瞬间已经注入清水的总量为 x , S 表示这时刻槽内含有酸性溶液的浓度. 试问, 要使酸性减少一半, 应注入多少清水?

11. 依照牛顿冷却定律, 高温物体冷却的速度与它同周围温度之差成正比. 设周围温度保持为 20°C . 如果物体的初始温度为 100°C , 在 20 分钟内降低至 60°C , 问需要多少时间此物体温度降低至 30°C ?

12. 求一曲线, 使由其上任一点的切线与两个坐标轴所构成的三角形的面积等于常数 S .

13. 求一曲线, 使得其上两点 $(0, 1)$ 、 (x, y) 之间的弧长为 $\sqrt{y^2 - 1}$.

14. 一个装满燃料、质量为 M_0 的火箭从地面发射. 如果排气速率为常数 a , 排出气体的速度相对于火箭为 b . 试求: 在没有外力的作用下, 火箭在燃料用完前的飞行速度及飞行

距离.

15. 考虑某种物质 A 经化学反应全部生成另一种物质 B. 设 A 的初始质量为 m_0 克, 在 1 小时内生成 B 物质 g 克. 试求

(1) 经过 3 小时后, A 物质起反应的量是多少?

(2) 经过多少小时后, A 物质中 75 % 的量已经起了反应?

16. 一个质量均匀的圆柱形筒, 内壁与外壁半径分别为 a, b , 且内壁与外壁分别保持恒温 t_1, t_2 . 试求: 圆柱在半径为 $r (0 < a < r < b)$ 处的温度.

§ 5.2 偏微分方程

5.2.1 一阶方程

如果微分方程中的未知函数是多个 (至少是两个) 自变量的函数, 我们称这个方程为偏微分方程, 如

$$P(x, y) \frac{\partial u}{\partial x} + Q(x, y) \frac{\partial u}{\partial y} = 0, \quad (5.2.1)$$

$$P(x, y, z) \frac{\partial z}{\partial x} + Q(x, y, z) \frac{\partial z}{\partial y} = R(x, y, z), \quad (5.2.2)$$

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0, \quad (5.2.3)$$

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = f(x_1, x_2, x_3) \quad (5.2.4)$$

都是偏微分方程. 和常微分方程一样, 偏微分方程中出现的未知函数的偏导数的最高阶数称为该偏微分方程的阶. 如 (5.2.1)、(5.2.2) 为一阶偏微分方程, (5.2.3)、(5.2.4) 是二阶的. 进一步, 由线性方程的定义, 我们还知道, 方程 (5.2.1)、(5.2.3)、(5.2.4) 都是线性的. 方程 (5.2.2) 关于未知函数的偏导数是线性的, 对这类方程, 我们称为拟线性偏微分方程. 一阶线性偏微分方程与常微分方程 (组) 有密不可分的关系. 现在我们来阐明这种关系.

对下列最简单的一阶线性偏微分方程 (5.2.1) 进行讨论. 给定

$$P(x, y) \frac{\partial u}{\partial x} + Q(x, y) \frac{\partial u}{\partial y} = 0,$$

假定函数 $P(x, y), Q(x, y)$ 连续可微, 且不同时为零. 考虑下列常微分方程

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)} \quad (P(x, y) \neq 0),$$

或

$$\frac{dx}{dy} = \frac{P(x, y)}{Q(x, y)} \quad (Q(x, y) \neq 0).$$

将它们写成对称形式

$$\frac{dx}{P(x,y)} = \frac{dy}{Q(x,y)}. \quad (5.2.5)$$

设(5.2.5)的通解由隐函数 $F(x,y) = c$ (c 为任意常数) 所确定, 方程 $F(x,y) = c$ 亦称为(5.2.5)的通积分.

通积分 $F(x,y) = c$ 表示 xOy 平面上方程(5.2.5)的一族互不相交的解曲线, 过点 (x,y) 的解曲线 $F(x,y) = c^*$ 在 (x,y) 处的切线斜率 $k(x,y)$ 应为 $\frac{Q(x,y)}{P(x,y)}$. 另一方面, $k(x,y)$ 也可由曲线方程 $F(x,y) = c^*$ 本身求得

$$k(x,y) = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}},$$

因此, 对任何 (x,y) , 总有

$$\frac{Q(x,y)}{P(x,y)} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}},$$

或

$$P(x,y)\frac{\partial F}{\partial x} + Q(x,y)\frac{\partial F}{\partial y} = 0.$$

这表明函数 $u = F(x,y)$ 满足方程(5.2.1), 即 $u = F(x,y)$ 是方程(5.2.1)的解.

反之, 如果函数 $u = F(x,y)$ 是方程(5.2.1)的非零解, 用方程(5.2.5)的任意解 $y = y(x)$ (或 $x = x(y)$) 代入 $F(x,y)$ 中, 再求导数, 可得

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dx}$$

或

$$\frac{dF}{dy} = \frac{\partial F}{\partial x} \cdot \frac{dx}{dy} + \frac{\partial F}{\partial y}.$$

但因为(5.2.5)

$$\frac{dy}{dx} = \frac{Q(x,y)}{P(x,y)} \quad \left(\text{或} \frac{dx}{dy} = \frac{P(x,y)}{Q(x,y)} \right),$$

所以

$$\frac{dF}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{Q(x,y)}{P(x,y)} = \frac{1}{P(x,y)} \left[P(x,y)\frac{\partial F}{\partial x} + Q(x,y)\frac{\partial F}{\partial y} \right] = 0,$$

或

$$\frac{dF}{dy} = \frac{\partial F}{\partial x} \cdot \frac{P(x,y)}{Q(x,y)} + \frac{\partial F}{\partial y} = \frac{1}{Q(x,y)} \left[P(x,y)\frac{\partial F}{\partial x} + Q(x,y)\frac{\partial F}{\partial y} \right] = 0,$$

故 $F(x, y(x)) \equiv c$ (或 $F(x(y), y) \equiv c$). 注意到 $y = y(x)$ (或 $x = x(y)$) 是方程(5.2.5)的任一解, 这又表明 $F(x, y) = c$ 是(5.2.5)的通积分. 因此我们证明了如下定理.

定理 5.2.1 (偏微的解与常微的通积分的等价性) 函数 $u = F(x, y)$ 是一阶线性偏微分方程(5.2.1)的解, 当且仅当

$$F(x, y) = c \quad (5.2.6)$$

是常微分方程(5.2.5)的通积分.

对一阶线性偏微分方程, 仍可像常微分方程那样讨论通解问题.

定义 5.2.1 (偏微的通解) 若(5.2.6)是方程(5.2.5)的一个通积分, 则称

$$u = \Phi(F(x, y))$$

为方程(5.2.1)通解, 这里 Φ 是任意可微函数.

对于上述定义, 我们需要证明如下两个事实:

(1) $u = \Phi(F(x, y))$ 是方程(5.2.1)的解.

(2) 对方程(5.2.1)的任意解 $u_1 = F_1(x, y)$, 一定存在可微函数 Φ , 使得

$$F_1(x, y) = \Phi(F(x, y)). \quad (5.2.7)$$

(1) 的证明比较容易, 事实上, 由 Φ 是可微函数, 我们有

$$\begin{aligned} P(x, y) \frac{\partial u}{\partial x} + Q(x, y) \frac{\partial u}{\partial y} &= P(x, y) \Phi'(F) \frac{\partial F}{\partial x} + Q(x, y) \Phi'(F) \frac{\partial F}{\partial y} \\ &= \Phi'(F) \left[P(x, y) \frac{\partial F}{\partial x} + Q(x, y) \frac{\partial F}{\partial y} \right] = 0. \end{aligned}$$

故 $u = \Phi(F(x, y))$ 是方程(5.2.1)的解.

下面, 我们来证明(2).

首先, 由定理 5.2.1, $F_1(x, y) = c$, $F(x, y) = c$ 是方程(5.2.5)的两个通积分, 若令 $z = F(x, y)$. 不妨设 $\frac{\partial F}{\partial y} \neq 0$ (对于 $\frac{\partial F}{\partial x} \neq 0$ 可同样证明). 由隐函数存在定理, 可以从 $z = F(x, y)$ 解出 $y = G(x, z)$, 并将其代入 $F_1(x, y)$ 中, 得到 $F_1(x, G(x, z))$. 记

$$\Phi(x, z) = F_1(x, G(x, z)). \quad (5.2.8)$$

现在若能证明 $\Phi(x, z)$ 与 x 无关, 问题就解决了. 为此, 将(5.2.8)对 x 求偏微商, 得到

$$\frac{\partial \Phi}{\partial x} = \frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} \cdot \frac{\partial G}{\partial x}. \quad (5.2.9)$$

注意到 $y = G(x, z)$ 是由隐方程 $z - F(x, y) = 0$ 解出来的, 故

$$\frac{\partial G}{\partial x} = - \frac{\partial F / \partial x}{\partial F / \partial y}.$$

代入(5.2.9), 得到

$$\begin{aligned}\frac{\partial \Phi}{\partial x} &= \frac{\partial F_1}{\partial x} + \frac{\partial F_1}{\partial y} \cdot \left(-\frac{\partial F}{\partial x} / \frac{\partial F}{\partial y} \right) \\ &= \frac{1}{\frac{\partial F}{\partial y}} \left[\frac{\partial F_1}{\partial x} \frac{\partial F}{\partial y} - \frac{\partial F_1}{\partial y} \frac{\partial F}{\partial x} \right].\end{aligned}\quad (5.2.10)$$

另一方面, $F(x, y)$ 、 $F_1(x, y)$ 是方程(5.2.5)的两个解, 所以

$$\begin{cases} P(x, y) \frac{\partial F}{\partial x} + Q(x, y) \frac{\partial F}{\partial y} = 0, \\ P(x, y) \frac{\partial F_1}{\partial x} + Q(x, y) \frac{\partial F_1}{\partial y} = 0. \end{cases}$$

由上式推出

$$\frac{\partial F_1}{\partial x} \frac{\partial F}{\partial y} - \frac{\partial F_1}{\partial y} \frac{\partial F}{\partial x} = 0.$$

代入(5.2.10)即知

$$\frac{\partial \Phi}{\partial x} = 0.$$

所以

$$F_1(x, y) = F_1(x, G(x, z)) = \Phi(z) = \Phi(F(x, y)).$$

定理 5.2.1 给出了一阶线性偏微分方程(5.2.1)与常微分方程(5.2.5)之间的内在联系.

定义 5.2.2 (偏微分方程的特征方程与特征线) 我们称常微分方程 $\frac{dx}{P(x, y)} = \frac{dy}{Q(x, y)}$ (即(5.2.5)) 为偏微分方程 $P(x, y) \frac{\partial u}{\partial x} + Q(x, y) \frac{\partial u}{\partial y} = 0$ (即(5.2.1)) 的特征方程, 并把特征方程(5.2.5)的解称为(5.2.1)的特征曲线或简称为特征线.

这样, 偏微分方程(5.2.1)的求解问题化为解常微分方程(5.2.5)问题.

在讨论三个自变量的一阶线性偏微分方程之前, 需要对一阶常微分方程组有一个初步的认识.

给定一阶常微分方程组

$$\begin{cases} \frac{dy_1}{dx} = f_1(x, y_1, y_2), \\ \frac{dy_2}{dx} = f_2(x, y_1, y_2). \end{cases} \quad (5.2.11)$$

定义 5.2.3 (首次积分) 如果以(5.2.11)的任一组解 $y_1(x), y_2(x)$ 代入一个连续可微函数 $\Phi(x, y_1, y_2)$, 使函数 $\Phi(x, y_1(x), y_2(x))$ 恒等于某一常数 (此常数与所取的解有关), 则称 $\Phi(x, y_1, y_2) = c$ 为方程组(5.2.11)的一个首次积分, 其中 c 是任意常数.

例如,对下列常微分方程组

$$\begin{cases} \frac{dy_1}{dx} = y_2, \\ \frac{dy_2}{dx} = -y_1. \end{cases} \quad (5.2.12)$$

先将第一式两端同乘 y_1 , 第二式两端同乘 y_2 , 然后相加, 得到

$$y_1 \frac{dy_1}{dx} + y_2 \frac{dy_2}{dx} = 0,$$

或

$$\frac{d}{dx}(y_1^2 + y_2^2) = 0.$$

对后一式关于 x 积分, 得到 $y_1(x)$ 与 $y_2(x)$ 的一个关系式

$$y_1^2(x) + y_2^2(x) \equiv c. \quad (5.2.13)$$

从而, 若取 $\Phi(x, y_1, y_2) = y_1^2 + y_2^2$, 则 $\Phi(x, y_1, y_2) = c$ 便是(5.2.12)的一个首次积分.

如果对常微分方程组(5.2.12), 先将第一式两端同乘 y_2 , 第二式两端同乘 y_1 , 然后相减, 得到

$$y_2 \frac{dy_1}{dx} - y_1 \frac{dy_2}{dx} = y_2^2 + y_1^2,$$

或

$$\frac{y_2 \frac{dy_1}{dx} - y_1 \frac{dy_2}{dx}}{y_1^2 + y_2^2} = 1.$$

对后一式两边关于 x 积分, 便得到

$$\arctan \frac{y_1(x)}{y_2(x)} = x + c$$

或

$$\arctan \frac{y_1(x)}{y_2(x)} - x \equiv c \quad (c \text{ 是常数}).$$

因此 $\Psi(x, y_1, y_2) = \arctan \frac{y_1}{y_2} - x = c$ 也是(5.2.12)的一个首次积分. 进一步,

由

$$\begin{cases} y_1^2(x) + y_2^2(x) \equiv c_1, \\ \arctan \frac{y_1(x)}{y_2(x)} - x \equiv c_2 \end{cases}$$

可以解出

$$\begin{cases} y_1(x) = c_1 \sin(x + c_2), \\ y_2(x) = c_1 \cos(x + c_2) \end{cases}$$

为(5.2.11)的通解.

一般地,如果 $\Phi(x, y_1, y_2) = c_1, F(x, y_1, y_2) = c_2$ 是方程(5.2.11)的两个独立的首次积分,亦即矩阵

$$\begin{bmatrix} \frac{\partial \Phi}{\partial x} & \frac{\partial \Phi}{\partial y_1} & \frac{\partial \Phi}{\partial y_2} \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y_1} & \frac{\partial F}{\partial y_2} \end{bmatrix}$$

中含有行列式不为零的二阶子阵,则由

$$\begin{cases} \Phi(x, y_1, y_2) = c_1, \\ F(x, y_1, y_2) = c_2, \end{cases}$$

我们可以解出

$$\begin{cases} y_1 = Y_1(x, c_1, c_2), \\ y_2 = Y_2(x, c_1, c_2) \end{cases}$$

为(5.2.11)的通解.这时我们称

$$\Phi(x, y_1, y_2) = c_1, \quad F(x, y_1, y_2) = c_2$$

为方程组(5.2.11)的通积分.

有时,我们也把(5.2.11)写成下列形式

$$\frac{dy_1}{f_1(x, y_1, y_2)} = \frac{dy_2}{f_2(x, y_1, y_2)} = \frac{dx}{1}.$$

一般地,可写成

$$\frac{dx}{P(x, y, z)} = \frac{dy}{Q(x, y, z)} = \frac{dz}{R(x, y, z)}, \quad (5.2.14)$$

其中 P, Q, R 连续可微,且不同时为零.

与二个自变量一阶线性偏微分方程一样,我们称方程组(5.2.14)为下列一阶线性偏微分方程

$$P(x, y, z) \frac{\partial u}{\partial x} + Q(x, y, z) \frac{\partial u}{\partial y} + R(x, y, z) \frac{\partial u}{\partial z} = 0 \quad (5.2.15)$$

的特征方程组,称方程组(5.2.14)的解为(5.2.15)的特征曲线,并且有

定理 5.2.2 (偏微的解与特征方程的首次积分) 函数 $u = F(x, y, z)$ 是一阶线性偏微分方程(5.2.15)的解,当且仅当

$$F(x, y, z) = c$$

是特征方程(5.2.14)的一个首次积分.

与定义 5.2.1(偏微的通解)相比较,我们得到两个自变量的偏微分方程的

通解的定义:若 $F(x, y, z) = c_1, \Phi(x, y, z) = c_2$ 是常微分方程组(5.2.14)的两个独立的通积分,则称

$$u = \Psi(F(x, y, z), \Phi(x, y, z))$$

为偏微分方程(5.2.15)的通解,其中 $\Psi(\omega, v)$ 是 ω, v 的任意连续可微函数.

例1 求方程

$$x \frac{\partial u}{\partial x} - 2y \frac{\partial u}{\partial y} - z \frac{\partial u}{\partial z} = 0 \quad (5.2.16)$$

的通解.

解 (5.2.16)的特征方程组为

$$\frac{dx}{x} = \frac{dy}{-2y} = \frac{dz}{-z}.$$

由 $\frac{dx}{x} = \frac{dy}{-2y}$ 解得 $x = \frac{c_1}{\sqrt{y}}$, 可求得一个首次积分 $x\sqrt{y} = c_1$. 由 $\frac{dx}{x} = \frac{dz}{-z}$ 解得 $x =$

$\frac{c_2}{z}$, 可求得另一个首次积分 $xz = c_2$. 因为

$$\begin{bmatrix} \frac{\partial}{\partial x}(x\sqrt{y}) & \frac{\partial}{\partial y}(x\sqrt{y}) & \frac{\partial}{\partial z}(x\sqrt{y}) \\ \frac{\partial}{\partial x}(xz) & \frac{\partial}{\partial y}(xz) & \frac{\partial}{\partial z}(xz) \end{bmatrix} = \begin{bmatrix} \sqrt{y} & \frac{x}{2\sqrt{y}} & 0 \\ z & 0 & x \end{bmatrix}.$$

当 $xyz \neq 0$ 时, 显然存在行列式不为零的二阶方阵, 故在 $xyz \neq 0$ 的区域上

$$u = \Psi(x\sqrt{y}, xz)$$

为(5.2.16)的通解, 其中 Ψ 是任意二元连续可微函数.

例2 求解

$$(z-y)\frac{\partial u}{\partial x} + (x-z)\frac{\partial u}{\partial y} + (y-x)\frac{\partial u}{\partial z} = 0. \quad (5.2.17)$$

解 特征方程组为

$$\frac{dx}{z-y} = \frac{dy}{x-z} = \frac{dz}{y-x}.$$

由合比定理得到

$$\frac{dx}{z-y} = \frac{dy}{x-z} = \frac{dz}{y-x} = \frac{dx+dy+dz}{0}.$$

由此可得 $d(x+y+z)=0$, 于是得到一个首次积分

$$x+y+z=c_1.$$

另外, 我们有

$$\frac{2xdx}{2x(z-y)} = \frac{2ydy}{2y(x-z)} = \frac{2zdz}{2z(y-x)} = \frac{dx^2+dy^2+dz^2}{0}.$$

于是得到另一个首次积分

$$x^2 + y^2 + z^2 = c_2.$$

容易验证 $F = x + y + z = c_1$ 与 $\Phi = x^2 + y^2 + z^2 = c$ 是两个独立的首次积分. 所以(5.2.17)有通解

$$u = \Psi(x + y + z, x^2 + y^2 + z^2),$$

其中 Ψ 是任意的二元连续可微函数.

最后, 我们介绍一阶拟线性偏微分方程(5.2.2)

$$P(x, y, z) \frac{\partial z}{\partial x} + Q(x, y, z) \frac{\partial z}{\partial y} = R(x, y, z)$$

的求解方法.

设(5.2.2)有隐函数形式的解 $v(x, y, z) = 0$. 由隐函数求导法则, 有

$$\frac{\partial z}{\partial x} = -\frac{\partial v / \partial x}{\partial v / \partial z}, \quad \frac{\partial z}{\partial y} = -\frac{\partial v / \partial y}{\partial v / \partial z}, \quad (5.2.18)$$

将(5.2.18)代入(5.2.2), 就得到 $v(x, y, z)$ 应满足下列一阶线性偏微分方程

$$P(x, y, z) \frac{\partial v}{\partial x} + Q(x, y, z) \frac{\partial v}{\partial y} + R(x, y, z) \frac{\partial v}{\partial z} = 0. \quad (5.2.19)$$

假设

$$v(x, y, z) = \Psi(F(x, y, z), G(x, y, z))$$

是(5.2.19)的通解, 从而由

$$\Psi(F(x, y, z), G(x, y, z)) = 0$$

所确定的 x, y 的函数 $z(x, y)$ 便是(5.2.2)的解. 此外, 还可以证明: 当 P, Q, R 有一阶连续偏微商且它们不同时为零时, (5.2.2)的一切解都可由

$$\Psi(F(x, y, z), G(x, y, z)) = 0$$

给出. 因此, 要求一阶拟线性偏微分方程(5.2.2)的通解, 再令其为零便可.

例 3 求解方程

$$x^2 \frac{\partial z}{\partial x} - xy \frac{\partial z}{\partial y} + y^2 = 0. \quad (5.2.20)$$

解 首先求下列方程的通解

$$x^2 \frac{\partial v}{\partial x} - xy \frac{\partial v}{\partial y} + y^2 \frac{\partial v}{\partial z} = 0. \quad (5.2.21)$$

其特征方程为

$$\frac{dx}{x^2} = \frac{dy}{-xy} = \frac{dz}{-y^2}.$$

由 $\frac{dx}{x^2} = \frac{dy}{-xy}$ 解得 $xy = c_1$ 为一个首次积分;

由 $\frac{dx}{-xy} = \frac{dz}{-y^2}$ 与 $xy = c_1$ 得 $\frac{dx}{x^2} = -\frac{dz}{c_1^2}$, 从而 $z = \frac{y^2}{3x} + c_2$. 故 $z - \frac{y^2}{3x} = c_2$ 为

另一个首次积分.

可以验证 $F(x, y, z) \equiv xy = c_1, G(x, y, z) \equiv z - \frac{y^2}{3x} = c_2$ 为两个独立的首次积分, 于是

$$\Psi\left(xy, z - \frac{y^2}{3x}\right) = v(x, y, z)$$

为(5.2.21)的通解, 其中 Ψ 为任意连续可微函数, 而 $\Psi\left(xy, z - \frac{y^2}{3x}\right) = 0$ 则为(5.2.20)的通解.

5.2.2 双曲、椭圆、抛物型方程

在本章开始, 我们建立了弦振动、热传导等数学模型, 它们都是偏微分方程. 5.2.1 节简单地介绍了求解一阶偏微分方程的方法. 对于二阶以至高阶偏微分方程, 求解会更加困难. 当人们意识到求通解的难度、并且通常求特解更为有用的时候, 偏微分方程的各种定解问题就被提出来了. 本小节将介绍经典二阶线性偏微分方程及其主要定解问题.

1. 二阶线性偏微分方程的类型

两个自变量二阶线性偏微分方程的一般形式是

$$a_{11} \frac{\partial^2 u}{\partial x^2} + 2a_{12} \frac{\partial^2 u}{\partial y \partial x} + a_{22} \frac{\partial^2 u}{\partial y^2} + b_1 \frac{\partial u}{\partial x} + b_2 \frac{\partial u}{\partial y} + cu + f = 0. \quad (5.2.22)$$

经过一定的函数变换, 一般二阶线性偏微分方程总可化为三种类型: 双曲型、抛物型、椭圆型. 我们以两个自变量 t, x 的函数 $u(t, x)$ 为例, 写出二阶线性偏微分方程化简后的形式:

$$\text{双曲型} \quad \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(t, x), \quad (5.2.23)$$

$$\text{椭圆型} \quad \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = f(t, x), \quad (5.2.24)$$

$$\text{抛物型} \quad \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + f(t, x). \quad (5.2.25)$$

这些方程有明确的物理意义与定解问题.

(1) 弦振动方程

在推导弦振动方程时, 我们是从一根拉紧了的柔软均匀弦出发, 得到了

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(t, x).$$

然而均匀杆的纵振动也表现为

$$u_{tt} = a^2 u_{xx},$$

其中 $u(t, x)$ 为杆上坐标为 x 的点在时刻 t 的位移, $a = \frac{E}{\rho}$, E 为杨氏模量, ρ 为

杆的密度;

张紧了的柔软均匀薄膜振动的数学模型是

$$u_{tt} = a^2 (u_{xx} + u_{yy}),$$

其中 $u(t, x, y)$ 为薄膜上坐标为 (x, y) 的点在时刻 t 的位移, $a^2 = \frac{T}{\rho}$, T 为薄膜的张力, ρ 为薄膜的密度;

电报方程

$$\frac{\partial^2 j}{\partial t^2} = a^2 \frac{\partial^2 j}{\partial x^2},$$

其中 $a^2 = \frac{1}{LC}$, j 为电流强度, L 为自感, C 为电容.

其他, 如流体力学、声学等也都有类似的数学模型. 完全不同的物理过程, 有完全不同的物理规律, 然而所得的数学模型竟然是一样的! 它们都是双曲型方程.

(2) 调和方程

$$\Delta_n u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = f(x_1, x_2, \cdots, x_n).$$

这类方程所代表的物理现象是一种稳定的状态, 例如, 静电场中电位用 $\varphi(x, y, z)$ 表示, 电荷密度为 $\rho(x, y, z)$, 则按静电学的基本定理可得静电场方程

$$\Delta_3 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = -4\pi\rho(x, y, z),$$

特别, 对于自由电场(即 $\rho=0$), 我们有

$$\Delta_3 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0.$$

调和方程还可以描述许多物理现象, 如定常温度场、引力场、流体力学中的势、弹性力学中的调和势等等. 概括地说, 它所描述的自然现象是恒稳的、定常的, 亦即与时间无关的现象.

(3) 热传导方程

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(t, x)$$

在推导热传导方程时, 我们是从热量在导热物质中传导的物理过程出发的, 方程中的项 $f(t, x)$ 表示“热源”, 但是还有许多自然现象, 如三维无源扩散问题, 其数学模型是

$$u_t = (\alpha u_x)_x + (\beta u_y)_y + (\gamma u_z)_z,$$

其中 u 是扩散液体的浓度, α, β, γ 是扩散系数, 对于均匀物质, $a = \alpha = \beta = \gamma$, 方程化为

$$u_t = a^2 \Delta_3 u.$$

此外,放射性物质的蜕变过程也是用这类抛物型方程刻画的.至于链式反应,又相当于有正的源、并且源强度正比于物质的浓度,则化为

$$u_t = (\alpha u_x)_x + (\alpha u_y)_y + (\alpha u_z)_z + \beta u \quad (\beta > 0).$$

微分方程如此广泛地出现在各种自然科学领域中,求解微分方程的重要性不言而喻,各种定解条件也相应地提了出来.

我们称要解决的问题为一个系统,假设系统中函数所满足的方程已经按照科学规律建立起数学模型,是一个偏微分方程,要求该偏微分方程的、满足一定条件的解.所提条件应当恰好是系统的初始状态或边界上的物理状态,条件不能提得过多,也不能过少.就物理科学的观点来说,只要确定了系统的初始状态、边界上的物理状态,整个系统以后的发展情况也必是确定的.这些条件统称偏微分方程的定解条件.

① **初始条件**: 是系统的初始状态的数学表达,它说明整个系统的初始情况.

以 3 维情形为例.未知函数的初始条件为

$$u(t, M)|_{t=0} = u(t, x, y, z)|_{t=0} = \varphi(x, y, z),$$

它可以表示弦振动的初始位置,也可表示热传导过程的初始温度,等等.

未知函数的初始速度也可给定,记为

$$\frac{\partial u}{\partial t}(t, M)|_{t=0} = u_t(t, x, y, z)|_{t=0} = \psi(x, y, z),$$

它可以表示弦振动的初始速度,等等.

② **边界条件**: 用以表达系统的边界所处的物理状态.一般作为表征的物理量 u 在边界 Σ 上所满足的关系给出的.

通常有三种情况:

第一类边界条件 直接给出 u 在边界上的值

$$u(t, M)|_{M \in \Sigma} = \mu(t, M);$$

第二类边界条件 给出 u 沿边界法方向 n 的方向导数在边界上的值

$$\frac{\partial u}{\partial n}(t, M)|_{M \in \Sigma} = \theta(t, M);$$

第三类边界条件 给出 u 与其沿边界法方向导数在边界上的值的线性组合

$$\left[u + h \frac{\partial u}{\partial n}(t, M) \right] \Big|_{M \in \Sigma} = \theta(t, M).$$

偏微分方程加上它的定解条件(包括一定的初始条件和边界条件)统称一个定解问题.

2. 用傅里叶级数法解偏微分方程定解问题

(1) 单位圆盘上的狄利克雷(Dirichlet)问题

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 & (x^2 + y^2 < 1), \\ u(x, y)|_{\Gamma} = f(x, y)|_{\Gamma}, \end{cases}$$

其中 Γ 是圆周 $x^2 + y^2 = 1$ (图 5.2.1).

使用极坐标 $x = r \cos \theta, y = r \sin \theta, -\pi \leq \theta \leq \pi, 0 \leq r \leq 1$, 方程化为

$$\begin{cases} \frac{\partial^2 u(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} = 0 & (r < 1), \\ u(x, y)|_{r=1} \equiv f(r). \end{cases}$$

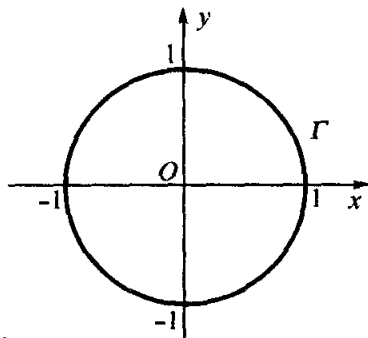


图 5.2.1

(5.2.26)

圆盘上的温度是稳定分布着的, 采用有限傅里叶变换法 (即傅里叶级数方法), 在 (5.2.26) 中对 θ 作有限傅里叶变换, 得

$$\left[\frac{\partial^2 u(r, \theta)}{\partial r^2} \right]^{\wedge}(k) + \left[\frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} \right]^{\wedge}(k) + \left[\frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} \right]^{\wedge}(k) = 0,$$

得

$$\frac{\partial^2 u^{\wedge}(r, k)}{\partial r^2} + \frac{1}{r} \frac{\partial u^{\wedge}(r, k)}{\partial r} - \frac{k^2}{r^2} u^{\wedge}(r, k) = 0 \quad (k = 0, \pm 1, \pm 2, \dots), \quad (5.2.27)$$

(5.2.27) 对每个 k 有一个二阶常微分方程, 只要对每个 $k = 0, \pm 1, \pm 2, \dots$, 按常微分方程的解法, 我们求得

$$u^{\wedge}(r, k) = \begin{cases} A(k)r^k + B(k)r^{-k}, & k \neq 0, \\ A(0) + B(0)\ln r, & k = 0, \end{cases} \quad (5.2.28)$$

其中 $A(k), B(k)$ 有定解条件 $u(x, y)|_{r=1} = f(r)$ 的有限傅里叶变换

$$u^{\wedge}(1, k) = f^{\wedge}(k) \quad (k = 0, \pm 1, \pm 2, \dots), \quad (5.2.29)$$

确定, 得

$$A(k) + B(k) = f^{\wedge}(k) \quad (k = \pm 1, \pm 2, \dots),$$

$$A(0) + B(0)\ln 1 = A(0) = f^{\wedge}(0).$$

据有限傅里叶变换的黎曼-勒贝格引理, 知 $u^{\wedge}(r, k)$ 有界, 故在 (5.2.28) 中确定

$$\begin{cases} A(k) = 0 & (k < 0), \\ B(k) = 0 & (k \geq 0), \end{cases}$$

从而 $u^{\wedge}(r, k) = f^{\wedge}(k)r^{|k|} \quad (k = 0, \pm 1, \pm 2, \dots)$. 最后得到单位圆盘上的狄利克雷问题的解

$$u(r, \theta) = \sum_{k=-\infty}^{+\infty} r^{|k|} f^{\wedge}(k) r^{ik\theta} \quad (r \in (0, 1), -\pi \leq \theta \leq \pi). \quad (5.2.30)$$

虽然已经求得定解问题的解, 但是还需要验证 (5.2.30) 确实是解, 并验证它

是问题的惟一的解. 进而, 还要验证当边界条件作微小变化时, 问题的解也作微小的变化, 这称为定解问题的稳定性. 由于在今后的后行课程中将要学习, 这里我们从略.

解的存在性、惟一性与稳定性, 合称偏微分方程定解问题的适定性.

以下两个定解问题也可用有限傅里叶变换解决.

(2) 有限长杆的热传导问题

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2}, & -\pi < x < \pi, t > 0, \\ u(x, t)|_{t=0} = f(x), & -\pi \leq x \leq \pi, \\ u(-\pi, t) = u(\pi, t), & t > 0, \\ \frac{\partial u(-\pi, t)}{\partial x} = \frac{\partial u(\pi, t)}{\partial x}, & t > 0. \end{cases} \quad (5.2.31)$$

(3) 有限长弦振动问题

$$\begin{cases} \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2}, & 0 < x < \pi, t > 0, \\ u(x, t)|_{t=0} = f(x), & 0 \leq x \leq \pi, \\ u_t(x, t)|_{t=0} = g(x), & 0 \leq x \leq \pi, \\ u(0, t) = u(\pi, t) = 0, & t > 0. \end{cases} \quad (5.2.32)$$

这两个求解问题留给读者.

习题

1. 求下列方程的通解

$$(1) y \frac{\partial z}{\partial x} - x \frac{\partial z}{\partial y} = 0$$

$$(2) y \frac{\partial z}{\partial x} = z$$

$$(3) \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = 0$$

$$(4) (x^2 + y) \frac{\partial u}{\partial x} + \frac{x(y+z)}{y} \frac{\partial u}{\partial z} = 0$$

$$(5) xz \frac{\partial z}{\partial x} + yz \frac{\partial z}{\partial y} = xy$$

$$(6) (y^2 + z^2 - x^2) \frac{\partial z}{\partial x} - 2xy \frac{\partial z}{\partial y} + 2xz = 0$$

2. 用有限傅里叶变换解问题(5.2.31)与(5.2.32).

§ 5.3 傅里叶变换

傅里叶级数对处理周期信号很有用, 但对非周期信号不大合适, 而要用傅里叶变换.

5.3.1 预备知识

傅里叶变换的理论与结果是建立在勒贝格测度与积分理论基础上的. 我们

总假设所考虑的函数是 L -可积的. 此为, 我们还假设定义在 $\mathbf{R}$ 上的函数 $f(x)$ 是逐段连续的, 在任一有限区间 $(a, b) \subset \mathbf{R}$ 中只有有限多个不连续点, 所遇到的积分为广义积分, 而累次积分均可交换次序.

定义 5.3.1 (函数集合 L^p) 对每个 $p, 1 \leq p < \infty$, 记

$$L^p = \left\{ f: \int_{\mathbf{R}} |f(x)|^p dx < \infty \right\},$$

$$L^\infty = \left\{ f: \operatorname{ess\,sup}_{-\infty < x < \infty} |f(x)| < \infty \right\},$$

及

$$\|f\|_p = \begin{cases} \left(\int_{\mathbf{R}} |f(x)|^p dx \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{-\infty < x < \infty} |f(x)|, & p = \infty, \end{cases}$$

其中 $\operatorname{ess\,sup}_{-\infty < x < \infty} |f(x)| = \inf_{me=0} \sup_{x \in \mathbf{R} \setminus e} |f(x)|$ 表示在 $f(x)$ 的全体连续点上取上确界, $\|f\|_p$ 称为 f 的 p -范数.

若 $p, q \geq 1$ 满足

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

则称 q 为 p 的共轭数, 当 $p=1$ 时规定 $q=\infty$. 下面证明著名的不等式: 赫尔德 (Holder) 不等式、施瓦兹 (Schwarz) 不等式与闵可夫斯基 (Minkowski) 不等式.

定理 5.3.1 (赫尔德不等式、施瓦兹不等式) 设 $f \in L^p(\mathbf{R}), g \in L^q(\mathbf{R})$,

$\frac{1}{p} + \frac{1}{q} = 1$, 则有赫尔不等式

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \quad (5.3.1)$$

当 $p=2$ 时, 有施瓦兹不等式

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2. \quad (5.3.2)$$

证 这里只证明施瓦兹不等式 (5.3.2). 不妨设 $f, g \geq 0$. 对任意 $\lambda \in \mathbf{R}$, 有

$$\begin{aligned} 0 &< \int_{\mathbf{R}} |f + \lambda g|^2 dx = \int_{\mathbf{R}} (|f|^2 + 2\lambda |fg| + \lambda^2 |g|^2) dx \\ &= \lambda^2 \int_{\mathbf{R}} |g|^2 dx + 2\lambda \int_{\mathbf{R}} |fg| dx + \int_{\mathbf{R}} |f|^2 dx, \end{aligned}$$

从而

$$\left(2 \int_{\mathbf{R}} |fg| dx \right)^2 - 4 \left(\int_{\mathbf{R}} |g|^2 dx \right) \left(\int_{\mathbf{R}} |f|^2 dx \right) \leq 0,$$

即有

$$\|fg\|_1^2 \leq \|f\|_2^2 \|g\|_2^2.$$

两边开方即得

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2.$$

定理得证.

定理 5.3.2 (闵可夫斯基不等式) 设 $p \geq 1, f, g \in L^p(\mathbf{R})$, 则有不等式

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \quad (5.3.3)$$

证 仍然只证 $p=2$ 的情况. 由施瓦兹不等式

$$\begin{aligned} \int_{\mathbf{R}} |f + g|^2 dx &\leq \int_{\mathbf{R}} |f|^2 dx + 2 \int_{\mathbf{R}} |fg| dx + \int_{\mathbf{R}} |g|^2 dx \\ &\leq \|f\|_2^2 + 2 \|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2. \end{aligned}$$

两边开方即得

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

不等式得证.

5.3.2 $L^1(\mathbf{R})$ 上的傅里叶变换与卷积运算

我们已经知道, 一个有限区间上的函数在一定条件下可展开成傅里叶级数. 例如, 设 $f(x)$ 是 $[-\pi, \pi]$ 只有有限个间断点的连续函数, 则展开为复数形式的傅里叶级数为

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \quad (x \in (-\pi, \pi)),$$

其中 $c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx$.

现在的问题是, 当 f 是 $\mathbf{R}$ 上的非周期函数时, 相应的“分解”应是怎样呢? 我们引入傅里叶变换与逆变换.

定义 5.3.2 (傅里叶变换) 设 $f \in L^1(\mathbf{R})$, 积分

$$f^\wedge(\xi) = Ff(\xi) \equiv \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{-i\xi x} f(x) dx \quad (\xi \in \mathbf{R}), \quad (5.3.4)$$

称为 $f \in L^1(\mathbf{R})$ 的傅里叶变换.

对每个 $f \in L^1(\mathbf{R})$, 显然 $f^\wedge(\xi)$ 在整个 $\mathbf{R}$ 上有定义, 它有如下的基本性质.

定理 5.3.3 (傅里叶变换的性质) 设 $f \in L^1(\mathbf{R})$, 则 $f^\wedge$ 有如下性质

- (i) $f^\wedge(\xi)$ 在 $\xi \in \mathbf{R}$ 上一致连续; (一致连续性)
- (ii) $f^\wedge \in L^\infty(\mathbf{R})$, 且 $\|f^\wedge\|_\infty \leq \|f\|_1$; (范数非增性)
- (iii) 当 $\xi \rightarrow +\infty$ 或 $\xi \rightarrow -\infty$ 时, $f^\wedge(\xi) \rightarrow 0$. ($R-L$ 引理)

证 为证(i), 我们设 $\delta > 0$, 考虑

$$\begin{aligned} \sup_{\xi} |f^\wedge(\xi + \delta) - f^\wedge(\xi)| &= \frac{1}{\sqrt{2\pi}} \sup_{\xi} \left| \int_{\mathbf{R}} e^{-i(\xi + \delta)x} (e^{-i\delta x} - 1) f(x) dx \right| \\ &\leq \int_{\mathbf{R}} |e^{-i\delta x} - 1| |f(x)| dx, \end{aligned}$$

由于 $|e^{-i\delta x} - 1| |f(x)| \leq 2|f(x)| \in L^1(\mathbf{R})$, 所以任给 $\varepsilon > 0$, 存在 $N > 0$, 使得

$$\int_{(-\infty, -N)} |e^{-i\delta x} - 1| |f(x)| dx + \int_{(N, +\infty)} |e^{-i\delta x} - 1| |f(x)| dx < \frac{\varepsilon}{2};$$

对此 N , 当 $\delta > 0$ 充分小时, 有 (不失一般性, 可设 $\|f\|_1 \neq 0$)

$$|e^{-i\delta x} - 1| \leq \frac{\epsilon}{2\|f\|_1}$$

对一切 $x \in [-N, N]$ 成立, 于是

$$\int_{-N}^N |e^{-i\delta x} - 1| |f(x)| dx < \frac{\epsilon}{2}.$$

从而当 δ 充分小时, 有

$$\sup_{\xi} |f^\wedge(\xi + \delta) - f^\wedge(\xi)| < \epsilon$$

对一切 $\xi \in \mathbf{R}$ 一致成立, 故 $f^\wedge$ 在 $\mathbf{R}$ 上一致连续.

对于(ii), 因为

$$|f^\wedge(\xi)| \leq \int_{\mathbf{R}} |e^{-i\xi x} f(x)| dx = \int_{\mathbf{R}} |f(x)| dx = \|f\|_1,$$

所以两边取上确界, 即得 $\|f^\wedge\|_\infty \leq \|f\|_1$, 从而也有 $f^\wedge \in L^\infty(\mathbf{R})$.

(iii) 中的 $R-L$ 引理, 就是著名的黎曼—勒贝格引理. 为证明此引理, 先设 f' 存在, 且 $f' \in L^1(\mathbf{R})$. 则由 $f \in L^1(\mathbf{R})$, 所以当 $x \rightarrow +\infty$ 或 $x \rightarrow -\infty$ 时有 $f(x) \rightarrow 0$, 于是利用分部积分

$$\begin{aligned} f^\wedge(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{-i\xi x} f(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{-i\xi} e^{-i\xi x} f(x) \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \frac{1}{i\xi} \int_{\mathbf{R}} f'(x) e^{-i\xi x} dx = \frac{1}{i\xi} f'^\wedge(\xi). \end{aligned}$$

从而得

$$f'^\wedge(\xi) = i\xi f^\wedge(\xi). \quad (5.3.5)$$

这样, 当 $\xi \rightarrow 0$ 时有

$$|f^\wedge(\xi)| = \frac{1}{|\xi|} |f'^\wedge(\xi)| \leq \frac{1}{|\xi|} \|f'\|_1 \rightarrow 0.$$

一般地, 对 $\epsilon > 0$, 可以找到函数 g 使得 $g, g^\wedge \in L^1(\mathbf{R})$ 且 $\|f - g\|_1 < \frac{\epsilon}{2}$. 于是由

(i), 得

$$\begin{aligned} |f^\wedge(\xi)| &\leq |f^\wedge(\xi) - g^\wedge(\xi)| + |g^\wedge(\xi)| \leq \|f - g\|_1 + |g^\wedge(\xi)| \\ &< \epsilon + |g^\wedge(\xi)|. \end{aligned}$$

而当 $|\xi|$ 充分大时有 $|g^\wedge(\xi)| < \frac{\epsilon}{2}$. 故当 ξ 充分大时有

$$|f^\wedge(\xi)| < \epsilon.$$

这就证明了(iii).

例 1 设 $a > 0$, 记 $e^x = \exp\{x\}$, 试证

$$\int_{\mathbf{R}} \exp\{-i\xi x\} \exp\{-ax^2\} dx = \sqrt{\frac{\pi}{a}} \exp\left\{-\frac{\xi^2}{4a}\right\}. \quad (5.3.6)$$

即:

$$(e^{-ax^2})^\wedge(\xi) = \sqrt{\frac{\pi}{a}} e^{-\frac{\xi^2}{4a}}.$$

证 对 $\xi \in \mathbf{R}$

$$\begin{aligned} & \int_{\mathbf{R}} \exp\{-i\xi x\} \exp\{-ax^2\} dx \\ &= \int_{\mathbf{R}} \exp\left\{-a\left[x^2 + \frac{i\xi}{a}x + \left(\frac{i\xi}{2a}\right)^2\right] + \frac{(i\xi)^2}{4a}\right\} dx \\ &= \int_{\mathbf{R}} \exp\left\{-a\left(x + \frac{i\xi}{2a}\right)^2\right\} \exp\left\{-\frac{\xi^2}{4a}\right\} dx \\ &= \exp\left\{-\frac{\xi^2}{4a}\right\} \int_{\mathbf{R}} \exp\left\{-a\left(x + \frac{i\xi}{2a}\right)^2\right\} dx. \end{aligned}$$

由复变函数论中的柯西定理可证明

$$\int_{\mathbf{R}} \exp\left\{-a\left(x + \frac{i\xi}{2a}\right)^2\right\} dx = \int_{\mathbf{R}} \exp\{-ax^2\} dx.$$

而

$$\int_{\mathbf{R}} \exp\{-ax^2\} dx = \frac{1}{\sqrt{a}} \int_{\mathbf{R}} \exp\{-x^2\} dx = \sqrt{\frac{\pi}{a}}.$$

从而可得(5.3.6).

注 在(5.3.6)中令 $a=1$, 即得高斯(Gauss)函数 e^{-x^2} 的傅里叶变换

$$(e^{-x^2})^\wedge(\xi) = \sqrt{\pi} e^{-\frac{\xi^2}{4}}.$$

取 $a = \frac{1}{2}$, 即得

$$(e^{-\frac{x^2}{2}})^\wedge(\xi) = \sqrt{2\pi} e^{-\frac{\xi^2}{2}}.$$

即, 若不计常数因子, 则 $e^{-\frac{x^2}{2}}$ 的傅里叶变换取同一形式.

在计算一些函数的傅里叶变换时, 需要用到傅里叶变换其他性质, 我们列出一些, 其余的重要性质可以从傅里叶变换表中查到.

定理 5.3.4 (傅里叶变换的运算性质) 设 $f \in L^1(\mathbf{R})$, 则

$$(i) [f(\cdot + h)]^\wedge(\xi) = e^{ih\xi} f^\wedge(\xi) \quad (h \in \mathbf{R}), \quad (\text{平移})$$

$$(ii) [e^{-ih\cdot} f(\cdot)]^\wedge(\xi) = f^\wedge(\xi + h) \quad (h \in \mathbf{R}), \quad (\text{乘 } e^{-ihx})$$

$$(iii) [\rho f(\rho \cdot)]^\wedge(\xi) = f^\wedge\left(\frac{\xi}{\rho}\right) \quad (h \in \mathbf{R}), \quad (\text{伸缩})$$

$$(iv) \overline{f(-\cdot)}^\wedge(\xi) = \overline{f^\wedge(\xi)}, \quad (\text{反射})$$

$$(v) f'^\wedge(\xi) = i\xi f^\wedge(\xi), \quad (\text{求导})$$

求导公式(v)需要对 f 加上一些条件, 由于涉及的知识很多, 这里不赘述, 请读

者在应用时,查看有关的参考书.

在傅里叶变换理论中还有一个必不可少的概念,就是“卷积”,在物理学与信号分析学起着极为重要的作用:一个信号,记为 $f(t)$,通过某个线性系统 S 时,如果线性系统的脉冲响应是 $h(t)$,则其输出信号就是 $f(t)$ 与 $h(t)$ 的卷积.

定义 5.3.3 (卷积) 设 $f, g \in L^1(\mathbf{R})$, 则 f 与 g 的卷积定义为

$$(f * g)(x) \equiv \int_{\mathbf{R}} f(x-y)g(y)dy, \quad x \in \mathbf{R}. \quad (5.3.7)$$

关于卷积,有如下定理

定理 5.3.5 (卷积的性质) 设 $f, g \in L^1(\mathbf{R})$, 则有

(i) $f * g \in L^1(\mathbf{R})$, 且 $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$; (封闭性)

(ii) $f * g = g * f$; (交换律)

(iii) 若还有 $u \in L^1(\mathbf{R})$, 则

$$(f * g) * u = f * (g * u). \quad (\text{结合律})$$

证 (i) 利用交换积分次序与积分换元法, 可得

$$\begin{aligned} \|f * g\|_1 &\leq \int_{\mathbf{R}} \left(\int_{\mathbf{R}} |f(x-y)| |g(y)| dy \right) dx \\ &= \int_{\mathbf{R}} |g(y)| \left(\int_{\mathbf{R}} |f(x-y)| dx \right) dy \\ &= \int_{\mathbf{R}} |g(y)| \|f\|_1 dy \\ &= \|f\|_1 \|g\|_1. \end{aligned}$$

(ii) 在(5.3.7)中令 $t = x - y$, 即可得

$$(f * g)(x) = (g * f)(x) \quad (x \in \mathbf{R}).$$

(iii) 留作习题.

定理 5.3.6 (卷积的傅立叶变换) 设 $f, g \in L^1(\mathbf{R})$, 则

$$(f * g)^{\wedge}(y) = f^{\wedge}(y)g^{\wedge}(y). \quad (5.3.8)$$

定理的证明只是交换积分次序的一个应用.

卷积作为 $L^1(\mathbf{R})$ 中元素之间的一种运算, 它有如定理 5.3.5 所述的封闭性, 交换律与结合律, 而且还容易验证它有关于 $L^1(\mathbf{R})$ 中函数的加法所具有线性性质与分配律, 但是卷积运算没有单位元, 即不存在 $L^1(\mathbf{R})$ 的元素 $e(x)$, 使得

$$f * e = f \quad (5.3.9)$$

对一切 $f \in L^1(\mathbf{R})$ 成立. 事实上, 如果(5.3.9)成立, 则由(5.3.8),

$$f^{\wedge}(y)e^{\wedge}(y) = f^{\wedge}(y)$$

对一切 $f \in L^1(\mathbf{R})$ 成立, 于是必须有 $e^{\wedge}(y) \equiv 1$, 这与黎曼-勒贝格引理(定理 5.3.3 的(iii))相矛盾.

然而, 高斯函数族

$$g_\alpha(x) = \frac{1}{2\sqrt{\pi\alpha}} e^{-\frac{x^2}{4\alpha}} \quad (\alpha > 0) \quad (5.3.10)$$

在极限意义下却起了单位元的作用. 事实上, 由例 1, 取 $\alpha = \frac{1}{4a}$, 则

$$g_\alpha^\wedge(y) = e^{-ay^2}, \quad (5.3.11)$$

我们有

定理 5.3.7 (高斯函数族的卷积定理) 设 $f \in L^1(\mathbf{R})$, 则

$$\lim_{\alpha \rightarrow 0} (f * g_\alpha)(x) = f(x)$$

在 f 的每个连续点 x 上成立.

证 假设 f 在 x 点连续, $\varepsilon > 0$ 为任意给定的正数, 选取 $\eta > 0$, 使得当 $|y| < \eta$ 时有

$$|f(x-y) - f(x)| < \varepsilon.$$

由 (5.3.11), $\int_{\mathbf{R}} g_\alpha(x) dx = g_\alpha^\wedge(0) = 1$, 因此

$$\begin{aligned} |(f * g_\alpha)(x) - f(x)| &= \left| \int_{\mathbf{R}} (f(x-y) - f(x)) g_\alpha(y) dy \right| \\ &\leq \int_{(-\eta, \eta)} |f(x-y) - f(x)| g_\alpha(y) dy + \\ &\quad \int_{|y| \geq \eta} |f(x-y) - f(x)| g_\alpha(y) dy \\ &\leq \varepsilon \int_{(-\eta, \eta)} g_\alpha(y) dy + \|f\|_1 \max_{|y| \geq \eta} g_\alpha(y) + \\ &\quad |f(x)| \int_{|y| \geq \eta} g_\alpha(y) dy \\ &\leq \varepsilon + \|f\|_1 g_\alpha(\eta) + |f(x)| \int_{|y| \geq \frac{\eta}{\sqrt{4\alpha}}} g_1(y) dy. \end{aligned}$$

因为 $g_\alpha(\eta)$ 与最后一项当 $\alpha \rightarrow 0^+$ 时, 都趋于 0, 所以, 要证的结论成立.

5.3.3 $L^1(\mathbf{R})$ 上傅里叶变换的反演

定理 5.3.3 告诉我们, 若函数 $f \in L^1(\mathbf{R})$, 则 $f^\wedge$ 在 $\mathbf{R}$ 上一致连续, 且 $y \rightarrow \pm\infty$ 时, $f^\wedge(y) \rightarrow 0$. 但 $f^\wedge$ 不一定属于 $L^1(\mathbf{R})$. 然而, 如果恰好有 $f^\wedge \in L^1(\mathbf{R})$, 则可以用 $f^\wedge$ 通过傅里叶逆变换来回复到 f .

定义 5.3.4 (傅里叶逆变换) 设 $f^\wedge$ 是 $f \in L^1(\mathbf{R})$ 的傅里叶变换, 则 $f^\wedge$ 的傅里叶逆变换定义为

$$(F^{-1}f^\wedge)(x) \equiv (f^\wedge)^\vee(x) \equiv \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{ix\xi} f^\wedge(\xi) d\xi. \quad (5.3.12)$$

一个重要的问题是:等式 $F^{-1}Ff \equiv (f^\wedge)^\vee = f$ 是否成立? 我们有

定理 5.3.8 (反演公式) 设 $f \in L^1(\mathbf{R})$, 且其傅里叶变换 $f^\wedge \in L^1(\mathbf{R})$, 则

$$f(x) = (F^{-1}f^\wedge)(x) \quad (5.3.13)$$

在 f 的每个连续点 x 上成立.

证 设 $x \in \mathbf{R}$ 为固定点, 令

$$g(t) = \frac{1}{\sqrt{2\pi}} e^{itx} e^{-at^2}. \quad (5.3.14)$$

利用例 1 中的 (5.3.6) 式, 得到

$$\begin{aligned} g^\wedge(y) &= \frac{1}{2\pi} \int_{\mathbf{R}} e^{-iyt} e^{itx} e^{-at^2} dt = \frac{1}{2\pi} \int_{\mathbf{R}} e^{-i(y-x)t} e^{-at^2} dt \\ &= \frac{1}{2\pi} \sqrt{\frac{\pi}{a}} e^{-\frac{(y-x)^2}{4a}} = g_a(x-y), \end{aligned}$$

其中 g_a 由 (5.3.10) 式定义, 演算下面的积分 (交换积分次序并由 (5.3.14))

$$\begin{aligned} (f * g_a)(x) &= \int_{\mathbf{R}} f(y) g_a(x-y) dy = \int_{\mathbf{R}} f(y) g^\wedge(y) dy \\ &= \int_{\mathbf{R}} f^\wedge(y) g(y) dy = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} e^{iyx} f^\wedge(y) e^{-ay^2} dy. \end{aligned}$$

令 $a \rightarrow 0^+$, 据定理 5.3.7 在 f 的连续点 x 上, 上式左边趋于 $f(x)$, 右边趋于 $(F^{-1}f^\wedge)(x)$. 故有

$$f(x) = (F^{-1}f^\wedge)(x).$$

这就是 (5.3.13). 注意到

$$\begin{aligned} (F^{-1}f^\wedge)(x) &= \frac{1}{2\pi} \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(t) e^{-iyt} dt \right) e^{ixy} dy \\ &= \frac{1}{2\pi} \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(t) e^{iy(x-t)} dt \right) dy \\ &= \frac{1}{2\pi} \left[\int_{(-\infty, 0)} + \int_{(0, \infty)} \right] \left(\int_{\mathbf{R}} f(t) e^{iy(x-t)} dt \right) dy, \end{aligned}$$

$$\text{而 } \int_{(-\infty, 0)} \left(\int_{\mathbf{R}} f(t) e^{iy(x-t)} dt \right) dy = \int_{(0, \infty)} \left(\int_{\mathbf{R}} f(t) e^{-iy(x-t)} dt \right) dy,$$

所以得到 $f(x)$ 的余弦表示

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{(0, \infty)} \left[\int_{\mathbf{R}} f(t) (e^{iy(x-t)} + e^{-iy(x-t)}) dt \right] dy \\ &= \frac{1}{\pi} \int_{(0, \infty)} \left(\int_{\mathbf{R}} f(t) \cos y(x-t) dt \right) dy. \end{aligned}$$

这个定理是当 $f, f^\wedge$ 都属于 $L^1(\mathbf{R})$ 时, 成立逆变换公式, 也叙述为: $f, f^\wedge \in L^1(\mathbf{R})$ 蕴含 $(f^\wedge)^\vee = f$. 因此有定理 5.3.8 的另一种形式, 称为傅里叶积分定理, 如下

定理 5.3.9 (傅里叶积分定理) 设 $f \in L^1(\mathbf{R})$, 如果其傅里叶变换 $f^\wedge \in$

$L^1(\mathbf{R})$, 则

$$f(x) = \frac{1}{\pi} \int_{(0, \infty)} \left(\int_{\mathbf{R}} f(t) \cos y(x-t) dt \right) dy \quad (5.3.15)$$

在 f 的每个连续点 x 处成立.

5.3.4 $L^2(\mathbf{R})$ 上的傅里叶变换

对 $f \in L^2(\mathbf{R})$, 积分 $\int_{\mathbf{R}} e^{-iyx} f(x) dx$ 可能发散, 所以不能直接用此积分来定义 $L^2(\mathbf{R})$ 上的傅里叶变换, 要作一些技术处理. 先给出平均极限的概念.

定义 5.3.5 (L^2 平均极限) 设 $\{f_N(x)\} \subset L^2(\mathbf{R})$ 为一函数项序列, 若存在 $f \in L^2(\mathbf{R})$, 使得

$$\lim_{N \rightarrow \infty} \|f_N - f\|_2 = 0,$$

则称 f 是 $f_N(x)$ 的平均极限, 或说 $f_N(x)$ 在 $L^2(\mathbf{R})$ 意义下收敛于 $f(x)$, 记为

$$\text{l.i.m.}_{N \rightarrow \infty} f_N(x) = f(x). \quad (5.3.16)$$

对 $f \in L^2(\mathbf{R})$, 令

$$f_N(x) = \begin{cases} f(x), & |x| \leq N, \\ 0, & |x| > N. \end{cases} \quad (5.3.17)$$

则 $f_N \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$, 因此 $f_N(x)$ 作为 $L^1(\mathbf{R})$ 上的傅里叶变换有定义, 且可以证明① $f_N \in L^2(\mathbf{R})$; ② 而当 $N, M \rightarrow \infty$ 时, $\|f_N - f_M\|_2 \rightarrow 0$, 并且③ 平均极限 $\text{l.i.m.}_{N \rightarrow \infty} f_N^\wedge$ 必存在, 记为 $f^\wedge$; ④ 也成立 $f^\wedge \in L^2(\mathbf{R})$.

定义 5.3.6 (L^2 的傅里叶变换) 设 $f \in L^2(\mathbf{R})$, $f_N(x)$ 由 (5.3.17) 式给出, 定义 $f(x)$ 的傅里叶变换为

$$f^\wedge(x) = \text{l.i.m.}_{N \rightarrow \infty} f_N^\wedge(x). \quad (5.3.18)$$

$L^2(\mathbf{R})$ 中的傅里叶变换, 有如下性质

定理 5.3.10 (L^2 傅里叶变换的性质) 设 $f \in L^2(\mathbf{R})$, 则有

- (i) $f^\wedge \in L^2(\mathbf{R})$; (封闭性)
- (ii) $\|f^\wedge\|_2^2 = \|f\|_2^2$; (保范性)
- (iii) 存在 $g \in L^2(\mathbf{R})$, 使得 $g^\wedge = f$; (逆元存在)
- (iv) 逆变换公式

$$f(x) = \text{l.i.m.}_{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{|y| \leq N} f^\wedge(y) e^{iyx} dy; \quad (5.3.19)$$

(v) 设 $g \in L^2(\mathbf{R})$, 则有一般的帕塞瓦尔等式

$$\int_{\mathbf{R}} f(x) \overline{g(x)} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} f^\wedge(\xi) \overline{g^\wedge(\xi)} d\xi \quad (5.3.20)$$

与普兰舍尔(Planchel)公式

$$\int_{\mathbf{R}} f(x) g^{\wedge}(x) dx = \int_{\mathbf{R}} f^{\wedge}(x) g(x) dx. \quad (5.3.21)$$

定理 5.3.10 的(ii)也称为帕塞瓦尔等式,它表示自然界中的能量守恒定律.

对于 $f \in L^p(\mathbf{R})$, $1 < p < 2$, 我们可以证明, f 可分解为

$$f = f_1(x) + f_2(x) \quad (f_1 \in L^1, f_2 \in L^2). \quad (5.3.22)$$

因此可定义

$$f^{\wedge} = f_1^{\wedge}(\xi) + f_2^{\wedge}(\xi), \quad (5.3.23)$$

上式右边 $f_1^{\wedge}$ 与 $f_2^{\wedge}$ 分别在 L^1 与 L^2 意义下.

关于更一般函数,以至“分布”的傅里叶变换,将在第六章中介绍.

习题

1. 设 $f(x) = \begin{cases} 1, & |x| \leq 1, \\ 0, & \text{其他.} \end{cases}$ 求 f 的傅里叶变换 $f^{\wedge}$, 并由此说明 $f \in L^1(\mathbf{R})$.

2. 设 $a \in \mathbf{R}$ 为常数, x 为变量. $f \in L^1(\mathbf{R})$, 试证

$$(1) (f(x-a))^{\wedge}(y) = e^{-ia y} f^{\wedge}(y)$$

$$(2) (f(a \cdot x))^{\wedge}(y) = \frac{1}{a} f^{\wedge}\left(\frac{y}{a}\right). (a > 0)$$

3. 设 $f, g, h \in L^1(\mathbf{R})$. λ_1, λ_2 为常数, 试证

$$h * (\lambda_1 f - \lambda_2 g) = \lambda_1 (h * f) - \lambda_2 (h * g).$$

4. 设 $f, g \in L^1(\mathbf{R})$. 并设所遇到的累次积分都能交换积分次序. 试证

$$(1) (f * g)^{\wedge}(y) = f^{\wedge}(y) g^{\wedge}(y)$$

$$(2) \int_{-\infty}^{\infty} f(x) g^{\wedge}(x) dx = \int_{-\infty}^{\infty} f^{\wedge}(x) g(x) dx$$

§ 5.4 小波变换简介

小波分析是 20 世纪 80 年代初期发展起来的一个数学分支,小波变换已被广泛应用于工程技术与科学领域中,作为具有丰富数学内涵的有效工具越来越受到人们的重视.

5.4.1 小波变换的引入

物理学中具有有限能量的模拟信号,在数学上表示为一个 L^2 可积函数 $f(x) \in L^2(\mathbf{R})$, 这里

$$L^2(\mathbf{R}) = \left\{ f: \|f\|_{L^2(\mathbf{R})} = \left\{ \int_{\mathbf{R}} |f(x)|^2 dx \right\}^{\frac{1}{2}} < \infty \right\} \quad (5.4.1)$$

表示 $\mathbf{R} = (-\infty, \infty)$ 上平方(勒贝格)可积的函数的全体. 若用 t 记时间变量, ω 记该模拟信号 f 的频率, 则 f 的“频谱”就是它的傅里叶变换(我们省去变换前的常数 $\frac{1}{\sqrt{2\pi}}$)

$$f^\wedge(\omega) = \int_{\mathbf{R}} e^{-i\omega t} f(t) dt \quad (\omega \in \mathbf{R}). \quad (5.4.2)$$

能量守恒定律表示为

$$\|f\|_2 = \|f^\wedge\|_2.$$

傅里叶变换是信号分析的重要工具, 在科学技术的各个领域起着极为重要的作用. 然而, 近年来科学技术的迅猛发展, 却使傅里叶变换感到力不从心, 也表现了其不足之处.

例如, 公式(5.4.2)告诉我们, 要从模拟信号 $f(t)$ 中获得其频谱信息, 甚至仅仅需要一个点 ω_0 的频谱 $f^\wedge$, 就得用上 $f(t)$ 在整个 $\mathbf{R}$ 的值.

但是, 实际问题所希望的是, 为求得频谱 $f^\wedge$ 在任一个频率范围 $[\omega, \omega + \Delta\omega]$ (频带)上的频谱信息, 是否存在一个时间间隔 $[t, t + \Delta t]$, 使得 $f^\wedge(t)$ 在 $[\omega, \omega + \Delta\omega]$ 的值由 $f(t)$ 在 $[t, t + \Delta t]$ 的值完全确定?

为了获得信号 $f(t)$ 的频谱 $f^\wedge(\omega)$ 的局部信息, 数学家 D. Gabor 在 1946 年引进了使得时间局部化的“窗函数”

$$g_a(t-b) = \frac{1}{2\sqrt{\pi a}} e^{-\frac{(t-b)^2}{4a}} \quad (a > 0, b \in \mathbf{R}), \quad (5.4.3)$$

其中 a, b 作为参数, 且

$$\int_{\mathbf{R}} g_a(t-b) db = \frac{1}{2\sqrt{\pi a}} \int_{\mathbf{R}} e^{-\frac{(t-b)^2}{4a}} db = \frac{1}{\sqrt{\pi}} \int_{\mathbf{R}} e^{-\xi^2} d\xi = 1.$$

若令

$$G_{b,\omega}^a(t) = e^{-i\omega t} g_a(t-b) = \overline{e^{i\omega t} g_a(t-b)}, \quad (5.4.4)$$

与

$$(G_b^a f)(\omega) = \int_{\mathbf{R}} e^{-i\omega t} f(t) g_a(t-b) dt = \int_{\mathbf{R}} f(t) G_{b,\omega}^a(t) dt, \quad (5.4.5)$$

则

$$f^\wedge(\omega) = \int_{\mathbf{R}} (G_b^a f)(\omega) db = \int_{\mathbf{R}} \int_{\mathbf{R}} f(t) G_{b,\omega}^a(t) dt db \quad (\omega \in \mathbf{R}),$$

这表明通过 $G_{b,\omega}^a(t)$ 可以从信号 $f(t)$ 中“提取”频谱 $f^\wedge$. 于是, 我们可以选取、调整参数 a, b , 构成一个“时域窗口”, 框出信号 $f(t)$ 的一部分, 同时构成一个“频域窗口”, 从中提取 $f^\wedge$.

由(5.4.3)式知道, $G_{b,\omega}^a(t)$ 源于 $g_a(t-b)$, 它是函数 $\psi(x) = e^{-x^2}$ 经过伸缩

与平移得到的,并且当 $x \rightarrow \pm \infty$ 时很快“衰减”为 0,后来人们把衰减很快的函数称为“小波”(wavelet).

另一方面,我们研究定义在有限区间,例如 $(-\pi, \pi)$ 上的平方可积函数类

$$L^2(-\pi, \pi) = \left\{ f: \|f\|_{L^2(-\pi, \pi)} = \left\{ \int_{\mathbf{R}} |f(x)|^2 dx \right\}^{\frac{1}{2}} < \infty \right\}, \quad (5.4.6)$$

任何函数 $f \in L^2(-\pi, \pi)$ 有复数形式的傅里叶级数展开,即

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad (5.4.7)$$

其中

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad (5.4.8)$$

(5.4.7) 式中的等号的意义是

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} \left| f(x) - \sum_{n=-N}^N c_n e^{inx} \right|^2 dx = 0, \quad (5.4.9)$$

也就是,级数 $\sum_{n=-\infty}^{\infty} c_n e^{inx}$ (平方)平均收敛于函数 $f(x)$. 由于(5.4.7)中的级数是 2π 周期的,因此我们总可以将 $L^2(-\pi, \pi)$ 中的函数 $f(x)$ 以周期 2π 扩张到整个实直线 $\mathbf{R}$ 上去,故可称 $L^2(-\pi, \pi)$ 为 2π 周期的平方可积函数空间.

若令

$$l^2 = \left\{ \{c_n\} \subset \mathbf{C}: \sum_{n=-\infty}^{\infty} |c_n|^2 < \infty \right\}. \quad (5.4.10)$$

其中 c_n 由(5.4.8)确定,则(5.4.7)式可以理解为:第一, $f(x)$ 在(平方)平均收敛的意义下可分解为无穷多相互直交的分量 $\varphi_n(x) \equiv c_n e^{inx}$ 的和,而 $w_n(x) \equiv e^{inx}$, $n = \dots, -1, 0, 1, \dots$ 是 $L^2(-\pi, \pi)$ 中的一个直交基(亦即,对任何函数 $f \in L^2(-\pi, \pi)$ 均有(5.4.8)的与(5.4.9)的表示). 第二,该直交基 $\{w_n\}$ 可以由同一个函数 $w(x) \equiv e^{ix}$ 经过整数伸缩,即 $w_n(x) = w(nx)$, $n \in \mathbf{Z}$ 而得到.

注意到,基本函数 $w(x) = e^{ix} = \cos x + i \sin x \in L^2(-\pi, \pi)$, 且表示一个周期为 2π 的正弦波, $w_n(x) = w(nx) \in L^2(-\pi, \pi)$ 当 $|n|$ 较大时其波的频率也较大,当 $|n|$ 较小时其波的频率较小,所以(5.4.7)也可说成“ $L^2(-\pi, \pi)$ 中的每个函数是由各种整数频率的波合成的”.

由于帕塞瓦尔等式

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |c_n|^2$$

与(5.4.9)式等价,于是,(5.4.7)可解释为“每个 2π 周期的平方可积函数 $f(x)$

可以用一个基本函数 $w(x) = e^{ix}$ 的整数伸缩 $w_n(x) \equiv e^{inx}$, $n \in \mathbb{Z}$, 经过 l^2 -线性组合而生成”.

经过以上分析, 我们看到, 无论在 $\mathbb{R}$ 情形, 还是在 $(-\pi, \pi)$ 情形, 都希望寻找一个“小波”, 将它经过伸缩与平移, 用来作为新的变换的基本函数, 并将要研究的函数进行“分解”. 这就是建立小波分析的思路.

于是, 问题归结为: 寻找一个满足一定条件的函数 $\psi(x)$, 对其自变量进行平移与伸缩

$$\psi\left(\frac{x-b}{a}\right) \quad (a, b \in \mathbb{R} \text{ 为参数, 且 } a \neq 0)$$

构成一种新的变换与相应的级数, 研究其性质并付诸应用.

5.4.2 连续小波变换与时频分析

连续小波变换, 或称积分小波变换, 简称小波变换, 是从定义一个“基小波”开始的.

定义 5.4.1 (基小波、积分小波变换) 设 $\psi \in L^2(\mathbb{R})$, 且满足条件

$$c_\psi = \int_{-\infty}^{\infty} \frac{|\psi^\wedge(\omega)|^2}{|\omega|} d\omega < \infty, \quad (5.4.11)$$

则称 ψ 是一个基小波, 对于定义积分

$$(W^\psi f)(b, a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad (f \in L^2(\mathbb{R})) \quad (5.4.12)$$

为 L^2 -函数 $f(x)$ 的关于基小波 ψ 的积分小波变换, 其中 $a, b \in \mathbb{R}$, 且 $a \neq 0$.

(5.4.11) 称为容许条件, c_ψ 称为容许常数. 函数族

$$\psi^{a,b}(t) = \frac{1}{\sqrt{|a|}} \overline{\psi\left(\frac{t-b}{a}\right)} \quad (5.4.13)$$

是一个双指标族, 称为相应于基小波 ψ 的小波族.

把(5.4.12)式记为

$$(W^\psi f)(b, a) = \langle f, \psi^{a,b} \rangle = |a|^{\frac{1}{2}} f * \overline{\psi(a^{-1} \cdot)}(b). \quad (5.4.14)$$

下面定义的窗函数在小波变换中与基小波起着同样重要的作用, 并且有重要的物理意义.

定义 5.4.2 (窗函数) 设 $g \in L^2(\mathbb{R})$, 且 $tg \in L^2(\mathbb{R})$, 则称 g 为一个窗函数.

现在假设 g 与 $g^\wedge$ 都是窗函数, 令

$$t_0 = \frac{1}{\|g\|_2^2} \int_{-\infty}^{\infty} t |g(t)|^2 dt,$$

$$\omega_0 = \frac{1}{\|g^\wedge\|_2^2} \int_{-\infty}^{\infty} \omega |g^\wedge(\omega)|^2 d\omega.$$

我们分别称 t_0 与 ω_0 为窗函数 g 与 $g^\wedge$ 的中心, 而称

$$(t_0, \omega_0) = \left(\frac{1}{\|g\|_2^2} \int_{\mathbf{R}} t |g(t)|^2 dt, \frac{1}{\|g^\wedge\|_2^2} \int_{\mathbf{R}} \omega |g^\wedge(\omega)|^2 d\omega \right) \quad (5.4.15)$$

为窗函数 g 在相空间 (t, ω) 中的中心 (也称相空间为时频空间); 并分别称

$$\Delta_g = \frac{1}{\|g\|_2} \left\{ \int_{\mathbf{R}} (t - t_0)^2 |g(t)|^2 dt \right\}^{\frac{1}{2}}, \quad (5.4.16)$$

$$\Delta_{g^\wedge} = \frac{1}{\|g^\wedge\|_2} \left\{ \int_{\mathbf{R}} (\omega - \omega_0)^2 |g^\wedge(\omega)|^2 d\omega \right\}^{\frac{1}{2}} \quad (5.4.17)$$

为 g 的时宽与频宽, 或称时半径与频半径.

在相空间 (t, ω) 中, 作一个以 (t_0, ω_0) 为中心、以 $2\Delta_g$ 为 t 轴方向的边长、以 $2\Delta_{g^\wedge}$ 为 ω 轴方向的边长的矩形, 此矩形称为窗函数 g 的时-频窗口; t 轴上的区间

$$T_g^0 = [t_0 - \Delta_g, t_0 + \Delta_g] \quad (5.4.18)$$

称为 g 的时间窗口, ω 轴上的区间

$$\Omega_{g^\wedge}^0 = [\omega_0 - \Delta_{g^\wedge}, \omega_0 + \Delta_{g^\wedge}] \quad (5.4.19)$$

称为 g 的频率窗口. 于是, g 的时-频窗口可写为

$$[T_g^0; \Omega_{g^\wedge}^0] = [t_0 - \Delta_g, t_0 + \Delta_g; \omega_0 - \Delta_{g^\wedge}, \omega_0 + \Delta_{g^\wedge}]. \quad (5.4.20)$$

窗函数有如下性质.

定理 5.4.1 (测不准原理) 设 g 与 $g^\wedge$ 都是窗函数, 则

$$\Delta_g \Delta_{g^\wedge} \geq \frac{1}{2};$$

等号成立当且仅当

$$g(t) = c e^{iat} \frac{1}{2\sqrt{\pi a}} e^{-\frac{(t-b)^2}{4a}} \quad (c \neq 0, a > 0, b \in \mathbf{R}). \quad (5.4.21)$$

这个定理称为测不准原理, 它告诉我们: ① 时频窗口的面积 $4\Delta_g \Delta_{g^\wedge} \geq 2$; ② 不可能有一个窗函数的时频窗口面积小于高斯函数 (5.4.21) 所决定的时频窗口的面积.

测不准原理的证明并不困难, 读者可参看有关文献, 其物理意义却非常重要, 体现了自然界的“测不准”的本质.

定理 5.4.2 (窗函数的性质) 设 g 与 $g^\wedge$ 都是窗函数, 则经平移与伸缩所得的函数

$$g^{a,b}(x) = \frac{1}{\sqrt{|a|}} g\left(\frac{x-b}{a}\right) \quad (a > 0, b \in \mathbf{R}) \quad (5.4.22)$$

也是窗函数, 其傅里叶变换

$$(g^{a,b})^\wedge(\omega) = \sqrt{|a|} e^{-ib\omega} g^\wedge(a\omega) \quad (5.4.23)$$

也是窗函数; $g^{a,b}$ 的中心与时宽分别为 $b + at_0$ 与 $a\Delta_g$; $(g^{a,b})^\wedge$ 的中心与频宽分别为 $\frac{\omega_0}{a}$ 与 $\frac{\Delta_g^\wedge}{a}$, 其时频窗口为

$$\begin{aligned} & [T_{g^{a,b}}^0; \Omega_{g^{a,b}}^0] \\ &= \left[b + at_0 - a\Delta_g, b + at_0 + a\Delta_g; \frac{\omega_0}{a} - \frac{\Delta_g^\wedge}{a}, \frac{\omega_0}{a} + \frac{\Delta_g^\wedge}{a} \right]. \end{aligned} \quad (5.4.24)$$

由(5.4.12)定义的 f 的小波变换 $(W^\psi f)(b, a)$, 从物理学的角度看, 有明显的意义. $(W^\psi f)(b, a)$ 可由在时间窗口(中心为 $b + at_0$ 、宽度为 $2a\Delta_\psi$)

$$T_{\psi^{a,b}}^0 = [b + at_0 - a\Delta_\psi, b + at_0 + a\Delta_\psi]$$

中的信号 $f(t)$ 的信息近似地反映出来, 在信号分析中称为“时间局部化”.

另一方面, 若令

$$\eta(\omega) = \psi^\wedge(\omega + \omega_0), \quad (5.4.25)$$

则由定理 5.4.2, η 也是一个窗函数, 且其中心为 0, 半径为 $\Delta_\psi^\wedge$. 利用(5.4.23)式与一般帕塞瓦尔等式

$$\int_{-\infty}^{\infty} f(x) \overline{g(x)} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} f^\wedge(x) \overline{g^\wedge(x)} dx,$$

可把小波变换 $(W^\psi f)(b, a)$ 表示为

$$(W^\psi f)(b, a) = c \int_{-\infty}^{\infty} f^\wedge(\omega) e^{ib\omega} \overline{\eta\left(a\left(\omega - \frac{\omega_0}{a}\right)\right)} d\omega. \quad (5.4.26)$$

因此, 在不计常数因子 c 与由时间窗口平移量 b 决定的线性相位移 $e^{ib\omega}$ 的情况下, $(W^\psi f)(b, a)$ 也可由在一个“频率窗口”(中心在 $\frac{\omega_0}{a}$ 、半径为 $\frac{\Delta_\psi^\wedge}{a}$)

$$\Omega_{\psi^{a,b}}^0 = \left[\frac{\omega_0}{a} - \frac{\Delta_\psi^\wedge}{a}, \frac{\omega_0}{a} + \frac{\Delta_\psi^\wedge}{a} \right]$$

中的谱 $f^\wedge(\omega)$ 的信息来近似反映, 我们称此为“频率局部化”.

通过(5.4.13)与(5.4.26), 我们得到一个时频窗口 $[T_{\psi^{a,b}}^0; \Omega_{\psi^{a,b}}^0]$ (见式(5.4.24)), 这个窗口的重要性在于: 对于“中心频率” $\frac{\omega_0}{a}$ 而言, 在较大的“中心频率”处较窄, 而在较小的“中心频率”处则较宽(见图 5.4.1), 因为此窗口的面积是一个常数 $4\Delta_\psi\Delta_\psi^\wedge$ (定理 5.4.1).

容许条件(5.4.12)也有重要的意义.

定理 5.4.3 (容许条件的意义) 设 ψ 是一个基小波, 则由 ψ 决定的连续小波变换 $(W^\psi f)(b, a)$ 有如下性质

$$J \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (W^\psi f)(b, a) \overline{(W^\psi g)(b, a)} db \frac{da}{a^2} = c_\psi \langle f, g \rangle \quad (5.4.27)$$

对所有 $f, g \in L^2(\mathbf{R})$ 成立.

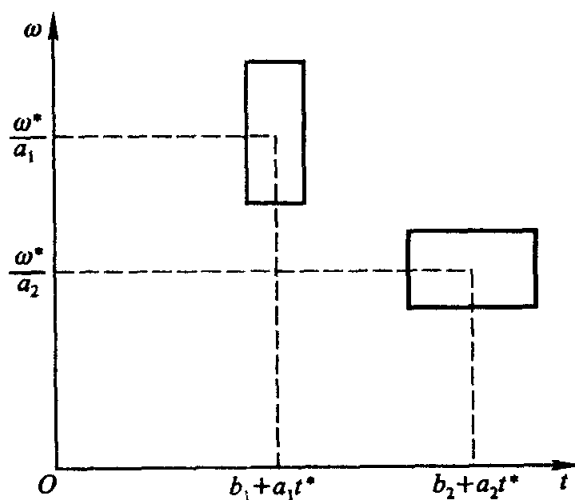


图 5.4.1

证 由富比尼定理与帕塞瓦尔公式, 计算

$$\begin{aligned}
 I &= \int_{\mathbf{R}} (W^\psi f)(b, a) \overline{(W^\psi g)(b, a)} da \\
 &= \frac{1}{|a|} \int_{\mathbf{R}} \left\{ \langle f, \psi^{b \cdot a} \rangle \int_{\mathbf{R}} \overline{g(s)} \psi\left(\frac{s-b}{a}\right) ds \right\} db \\
 &= |a| \left\{ \frac{1}{2\pi} \int_{\mathbf{R}} F(\omega) e^{ib\omega} d\omega \cdot \frac{1}{2\pi} \int_{\mathbf{R}} \overline{G(\tau)} e^{-ib\tau} d\tau \right\} db \\
 &= |a| \int_{\mathbf{R}} \left\{ \frac{1}{2\pi} \int_{\mathbf{R}} \overline{F(\omega)} e^{-ib\omega} d\omega \cdot \frac{1}{2\pi} \int_{\mathbf{R}} \overline{G(\omega)} e^{-ib\tau} d\tau \right\} db \\
 &= |a| \int_{\mathbf{R}} \overline{(F(\cdot)^\wedge)(b)} \cdot G(\cdot)^\wedge(b) db \\
 &= \frac{|a|}{2\pi} \int_{\mathbf{R}} F(x) \overline{G(x)} dx = \frac{a^2}{2\pi|a|} \int_{\mathbf{R}} \overline{G(x)} F(x) dx,
 \end{aligned}$$

其中

$$F(\omega) = f^\wedge(\omega) \overline{\psi^\wedge(a\omega)}, \quad G(\tau) = g^\wedge(\tau) \overline{\psi^\wedge(a\tau)}.$$

于是,

$$\begin{aligned}
 J &= \int_{\mathbf{R}} \int_{\mathbf{R}} (W^\psi f)(b, a) \overline{(W^\psi g)(b, a)} db \frac{da}{a^2} \\
 &= \int_{\mathbf{R}} I \frac{1}{a^2} da = \frac{1}{2\pi} \int_{\mathbf{R}} \frac{a^2}{|a|} \int_{\mathbf{R}} \overline{G(x)} F(x) \frac{1}{a^2} da \\
 &= \frac{1}{2\pi} \int_{\mathbf{R}} \frac{1}{|a|} \left\{ \int_{\mathbf{R}} \overline{g^\wedge(x)} \psi^\wedge(ax) \right\} \cdot f^\wedge(x) \overline{\psi^\wedge(ax)} dx da \\
 &= \frac{1}{2\pi} \int_{\mathbf{R}} f^\wedge(x) \overline{g^\wedge(x)} \left\{ \frac{|\psi^\wedge(ax)|^2}{|a|} da \right\} dx
 \end{aligned}$$

$$= c_\psi \langle f^\wedge, g^\wedge \rangle = c_\psi \langle f, g \rangle.$$

定理得证.

上述定理指出,对于“二元函数” $(W^\psi f)(b, a)$,在“权元”

$$db \frac{da}{a^2}$$

之下的“内积” $J = \langle W^\psi f, W^\psi g \rangle$ 等于 $c_\psi \langle f, g \rangle$,也就是,施行小波变换 $(W^\psi f)(b, a)$ 后的内积 $\langle W^\psi f, W^\psi g \rangle$ 与内积 $\langle f, g \rangle$ 相差一个常数 c_ψ 因子,这个因子就是容许条件(5.4.11)中的常数,也就是容许常数.

下面是基小波的几个例.

例1 哈尔(Haar)函数

$$\psi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2}, \\ 0, & \text{其他.} \end{cases}$$

例2 墨西哥(Mexico)帽子函数

$$\varphi(t) = \begin{cases} 1 - |x|, & 0 \leq |x| \leq 1, \\ 0, & \text{其他.} \end{cases}$$

5.4.3 离散小波变换

离散小波变换,又称二进小波变换,与连续小波变换有密切的联系.

给定一个基小波 ψ ,取 $(b, a) = \left(\frac{k}{2^j}, \frac{1}{2^j}\right)$, $j, k \in \mathbf{Z}$,并令

$$\psi_{jk}(x) = 2^{\frac{j}{2}} \psi(2^j x - k) \quad (j, k \in \mathbf{Z}) \quad (5.4.28)$$

定义 5.4.3 (二进小波变换、二进小波级数) 我们称

$$(W^\psi f)\left(\frac{k}{2^j}, \frac{1}{2^j}\right) \equiv \langle f, \psi_{jk} \rangle \equiv \alpha_{jk} \quad (j, k \in \mathbf{Z}) \quad (5.4.29)$$

为二进小波变换,与其对应,称

$$\sum_{j, k \in \mathbf{Z}} \alpha_{jk} \tilde{\psi}_{jk}(x), \quad (5.4.30)$$

为(二进)小波级数,其中

$$\tilde{\psi}_{jk}(x) = 2^{\frac{j}{2}} \tilde{\psi}(2^j x - k),$$

$\tilde{\psi}$ 是 ψ 的对偶函数,它们之间的关系由其傅里叶变换确定

$$(\tilde{\psi})^\wedge(\omega) = \frac{\psi^\wedge(\omega)}{\sum_{-\infty}^{\infty} |\psi^\wedge(\omega + 2\pi k)|^2}. \quad (5.4.31)$$

用克罗内克符号

$$\delta_{jk} = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases}$$

一般地,我们可以给出

定义 5.4.4 (L^2 的正交基、正交小波) 设 $\psi \in L^2(\mathbf{R})$, 用(5.4.28)中的记号, 取 $\{\psi_{jk}(x)\}_{j,k \in \mathbf{Z}}$, 如果它满足

$$(i) \langle \psi_{jk}, \psi_{lm} \rangle = \delta_{j,l} \cdot \delta_{k,m} \quad (j, k, l, m \in \mathbf{Z}), \quad (5.4.32)$$

(ii) 每个 $f \in L^2(\mathbf{R})$, 有

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{jk} \psi_{jk}(x), \quad (5.4.33)$$

上式中的级数收敛, 是指

$$\lim_{N, M \rightarrow \infty} \left\| f - \sum_{j=-M}^M \sum_{k=-N}^N c_{j,k} \psi_{j,k} \right\|_2 = 0, \quad (5.4.34)$$

则称 $\{\psi_{jk}\}$ 是 $L^2(\mathbf{R})$ 中的一个直交(正交)基, 并称 ψ 为一个正交小波. (5.4.33) 中的级数称为 f 的小波级数.

定义 5.4.5 (稳定条件) 若函数 $\psi \in L^2(\mathbf{R})$ 满足: 存在常数 $A, B, 0 < A \leq B < \infty$, 使得

$$A \leq \sum_{j=-\infty}^{\infty} |\psi^\wedge(2^j \omega)|^2 \leq B, \quad a.e., \quad (5.4.35)$$

则称 ψ 满足稳定条件(5.4.35); 当 $A = B$ 时, (5.4.35) 称为最稳定条件.

我们叙述两个定理.

定理 5.4.4 (稳定条件与基小波) 满足稳定条件的函数 $\psi \in L^2(\mathbf{R})$ 是一个满足容许条件的基小波. 且有不等式

$$A \ln 2 \leq \int_0^\infty \frac{|\psi^\wedge(\omega)|^2}{\omega} d\omega \leq B \ln 2;$$

$$A \ln 2 \leq \int_0^\infty \frac{|\psi^\wedge(-\omega)|^2}{\omega} d\omega \leq B \ln 2.$$

定理 5.4.5 (稳定条件的等价条件) 稳定条件等价于

$$A \|f\|_2^2 \leq \sum_{j=-\infty}^{\infty} \|W_j^\psi f\|_2^2 \leq B \|f\|_2^2 \quad (f \in L^2(\mathbf{R})). \quad (5.4.36)$$

5.4.4 小波反演公式

定义 5.4.6 (小波变换反演公式) 设 $\psi \in L^2(\mathbf{R})$, $(W^\psi f)(b, a)$ 是由(5.4.12)定义的积分小波变换. 若存在一个公式, 对任何 $f \in L^2(\mathbf{R})$ 都能通过 $(W^\psi f)(b, a)$ 来恢复 f (在(平方)平均收敛的意义下), 则称用于恢复 f 的公式为小波反演公式或逆变换公式.

下面对四种不同情况给出平均收敛意义下的反演公式,证明从略.

(1) 通过 $(W^\psi f)(b, a)$ ($a, b \in \mathbf{R}$) 变换来恢复 f

设 $\psi \in L^2(\mathbf{R})$ 是满足容许条件(5.4.11)的、具有单位长度的基小波,则成立反演公式:对于 $f \in L^2(\mathbf{R})$, 有

$$f(x) = \frac{1}{c_\psi} \iint_{\mathbf{R}^2} \{(W^\psi f)(a, b)\} \left\{ |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \right\} \frac{da db}{a^2}. \quad (5.4.37)$$

注意到,满足 $\psi \in L^2(\mathbf{R})$ 与容许条件(5.4.11)的 ψ 一定有 $\psi \in L(\mathbf{R})$, 故由傅里叶变换的性质知, $\psi^\wedge$ 是 $\mathbf{R}$ 上的连续函数,并且由(5.4.11)得到 $\psi^\wedge(0) = 0$, 因此也就有

$$\int_{-\infty}^{\infty} \psi(x) dx = \psi^\wedge(0) = 0.$$

于是, $\psi(x)$ 的图像具有上下波动、急速衰减且有小的波浪形态,如图 5.4.2 所示.

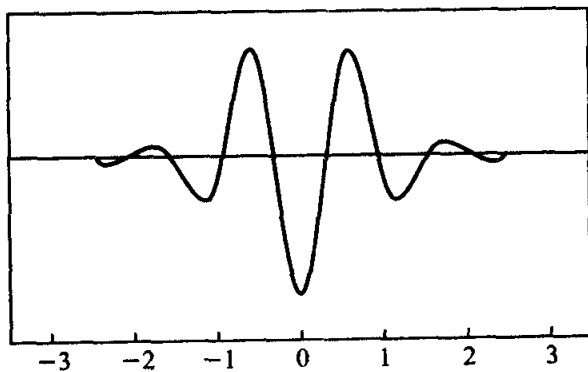


图 5.4.2

(2) 通过 $(W^\psi f)(b, a)$ ($b \in \mathbf{R}, a > 0$) 变换来恢复 f

设 $\psi \in L^2(\mathbf{R})$ 是满足容许条件(5.4.11)的、具有单位长度的基小波,且成立

$$\int_0^\infty \frac{|\psi^\wedge(\omega)|^2}{\omega} d\omega = \int_0^\infty \frac{|\psi^\wedge(-\omega)|^2}{\omega} d\omega = \frac{1}{2} c_\psi, \quad (5.4.38)$$

则成立反演公式:对于 $f \in L^2(\mathbf{R})$, 有

$$f(x) = \frac{2}{c_\psi} \int_0^\infty \int_{-\infty}^\infty \{(W^\psi f)(a, b)\} \left\{ |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \right\} db \frac{da}{a^2}. \quad (5.4.39)$$

(3) 通过 $(W^\psi f)(b, a)$ ($b \in \mathbf{R}, a = \frac{1}{2^j}, j \in \mathbf{Z}$) 变换来恢复 f

设 $\psi \in L^2(\mathbf{R})$ 是满足稳定性条件(5.4.35)的二进小波,则 ψ 是一个基小波,且对于 $f \in L^2(\mathbf{R})$, 成立

$$f(x) = \sum_{j=-\infty}^{\infty} \int_{-\infty}^{\infty} \{2^{\frac{j}{2}} (W^\psi f)(2^{-j}, b)\} \{2^j \psi^*(2^j(x-b))\} db, \quad (5.4.40)$$

其中 ψ^* 由它的傅里叶变换

$$(\psi^*)^\wedge(\omega) = \frac{\overline{\psi^\wedge(\omega)}}{\sum_{j=-\infty}^{\infty} |\psi^\wedge(2^j\omega)|^2} \quad (5.4.41)$$

所确定.

(4) 通过 $(W^\psi f)(b, a) \left(b = \frac{k}{2^j}, a = \frac{1}{2^j}, j \in \mathbb{Z} \right)$ 变换来恢复 f

设 $\psi \in L^2(\mathbb{R})$ 是满足稳定性条件(5.4.35)的二进小波, 则 ψ 是一个基小波, 且对于 $f \in L^2(\mathbb{R})$, 成立

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(x), \quad (5.4.42)$$

其中

$$\begin{aligned} \psi_{j,k}(x) &= 2^{\frac{j}{2}} \psi(2^j x - k) \quad (k \in \mathbb{Z}), \\ c_{j,k} &= \langle f, \tilde{\psi}_{j,k} \rangle = (W^{\tilde{\psi}} f) \left(\frac{k}{2^j}, \frac{1}{2^j} \right), \end{aligned} \quad (5.4.43)$$

$\tilde{\psi}$ 由它的傅里叶变换

$$(\tilde{\psi})^\wedge(\omega) = \frac{\psi^\wedge(\omega)}{\sum_{j=-\infty}^{\infty} |\psi^\wedge(2^j\omega)|^2} \quad (5.4.44)$$

所确定, 称为二进小波 ψ 的对偶.

现在给出公式(5.4.42)的对偶形式.

设 $\psi \in L^2(\mathbb{R})$ 是满足稳定性条件(5.4.35)的二进小波, 则 $\tilde{\psi}$ 是一个基小波, 且对于 $f \in L^2(\mathbb{R})$, 成立

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \tilde{\psi}_{j,k}(x), \quad (5.4.45)$$

其中

$$d_{j,k} = \langle f, \tilde{\psi}_{j,k} \rangle = (W^{\tilde{\psi}} f) \left(\frac{k}{2^j}, \frac{1}{2^j} \right). \quad (5.4.46)$$

5.4.5 多分辨分析, 小波应用概述

小波分析源于信号分析, 是数学科学中调和和分析领域近半个世纪以来的最重要的成果之一, 并且已经广泛应用于信号分析、图像处理、量子场论、地震勘探、语音识别、雷达噪声、CT 成像、彩色复印、流体湍流、天体探测、机器视觉、机械故障诊断、数字电视、音乐合成, 以及应用于分形分析的理论研究和计算, 等等.

由于小波分析具有时频局部化性质, 在时域和频域中同时具有良好的局部

化性质,这大大优于傅里叶分析,从原则上看,凡是传统上使用傅里叶分析的地方,都可以使用小波分析.

进而,多分辨分析与马拉特(Mallat)塔式算法也构成小波分析应用的基础,由于内容较深,我们不再详细介绍,只简述其大意.

把 $L^2(\mathbf{R})$ 分解为满足一定条件的一系列闭子空间 $\{V_j\}_{j \in \mathbf{Z}}$, 具有生成元 $g(x) \in V_0$. 由 g 通过公式

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \left[\left(\sum_{k \in \mathbf{Z}} |g^\wedge(\omega + 2k\pi)|^2 \right)^{-\frac{1}{2}} g^\wedge(\omega) \right]^\vee.$$

可以证明,如此生成的 φ 是一个尺度函数,而

$$\{\varphi_{jk}(x) = 2^{-j/2} \varphi(2^{-j}x - k) : j, k \in \mathbf{Z}\}$$

恰好是 V_j 的标准正交基. 再求出频率响应 $\{h_k : k \in \mathbf{Z}\}$ 与传递函数 $H(\omega)$ 以及 $\{V_j\}_{j \in \mathbf{Z}}$ 的正交补空间 $\{W_j\}_{j \in \mathbf{Z}}$ 与其标准正交基

$$\{\psi_{jk}(x) = 2^{-j/2} \psi(2^{-j}x - k) : j, k \in \mathbf{Z}\}.$$

在应用时,将输入信号 $f(x)$ 按照马拉特塔式算法展开成低频部分与高频部分(相当于通过低通滤波器与带通滤波器),然后,或进行编码数据压缩、或需要进行信噪分离与滤波,或其他应用等. 小波分析对高频成分采用逐渐精细的时域或空域取样步长,从而可以聚焦到信号的任意细节,以达到应用的要求.

习题

1. 试证:哈尔(Haar)函数

$$\varphi(t) = \begin{cases} 1, & 0 \leq t < \frac{1}{2}, \\ -1, & \frac{1}{2}, \\ 0, & \text{其他} \end{cases}$$

是一个基小波.

2. 试证:墨西哥(Mexico)帽子函数

$$\varphi(t) = \begin{cases} 1 - |x|, & 0 \leq |x| \leq 1, \\ 0, & \text{其他} \end{cases}$$

是一个基小波.

第六章 空间结构与现代数学

20 世纪数学的发展速度,是 19 世纪末任何有远见卓识的、有重要贡献的数学家始料未及的.然而,当今站在两个世纪之交的科学家们,确已看到 21 世纪整个科学发展对数学的需求程度的激增,这种需求,不仅包括对经典的,而且还包括对近代数学的依赖.为此,作为高等数学的最后一章,我们介绍近代数学的基础.

§ 6.1 集 合

一个重要的数学思维方法是:把研究对象按照一定的性质分成各种类型进行研究,在同一类型中的对象具有共同的性质,因此不需一个一个地考虑,而是按类进行研究,这样更能使我们了解一类事物的本质.例如,我们把全体有理数归为一类,称为有理数集,记为

$$Q = \{x: x \text{ 为有理数}\}.$$

又如,在区间 $[a, b]$ 上连续函数的全体,称为连续函数所成的集合,记为

$$C([a, b]) = \{f(x): f(x) \text{ 为 } [a, b] \text{ 上的连续函数}\}.$$

$[a, b]$ 上具有一阶连续导数的函数的全体,称为具有一阶连续导数的函数所成的集合,记为

$$C^1([a, b]) = \{f(x): f'(x) \text{ 为 } [a, b] \text{ 上的连续函数}\}$$

等等.一般地,我们把具有一定性质的对象的全体称为一个集合.集合用 $A, B, \dots$ 或 $X, Y, \dots$ 来记.集合中的每个对象称为“元素”,一般用小写字母 $x, y, \dots$ 或 $a, b, \dots$ 来记.若 x 是 A 的元素,记为 $x \in A$.若 x 不是 A 的元素,则记为 $x \notin A$.

可以看到,集合是一个非常有用的概念,更是很有用的数学工具,因为无论自然科学还是社会科学,所研究的都是具有一定性质的对象,常引进集合来描述这些对象,更便于用数学思维方法来处理.

1. 常用的数的集合

$$N = \{x: x \text{ 为自然数 } 1, 2, \dots\},$$

$$Z = \{x: x \text{ 为整数 } 0, \pm 1, \pm 2, \dots\},$$

$$Q = \{x: x \text{ 为有理数}\},$$

$$R = \{x: x \text{ 为实数}\},$$

$$C = \{x: x \text{ 为复数}\},$$

2. 集合之间的关系

包含: 设 A 与 B 为两个集合. 若 A 中的任一个元素 $a \in A$, 也属于 B , $a \in B$, 则称 A 含于 B , 或称 B 包含 A , 记为 $A \subseteq B$, 并称 A 为 B 的子集. 若 A 含于 B 但 $A \neq B$, 则称 A 为 B 的真子集, 记为 $A \subset B$.

例如, $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}$.

相等: 若 $A \subseteq B$, 且 $B \subseteq A$, 则称 $A = B$.

例如, $\{x \in \mathbf{N}: x^2 = 1\} = \{1\} \subset \{x \in \mathbf{Z}: x^2 = 1\} = \{1, -1\}$.

空集: 不含任何元素的集合称为空集, 记为 $\emptyset$.

例如, $\{x \in \mathbf{N}: x^2 = -1\} = \emptyset$.

3. 集合之间的运算

集合的交集: 设 A 与 B 为两个集合. A 与 B 的交集定义为一个集合, 其中的元素既属于 A 又属于 B . 记为 $A \cap B = \{x: x \in A \text{ 且 } x \in B\}$.

集合的并集: 设 A 与 B 为两个集合. A 与 B 的并集定义为一个集合, 其中的元素或属于 A 或属于 B . 记为 $A \cup B = \{x: x \in A \text{ 或 } x \in B\}$.

集合的差集: 设 A 与 B 为两个集合. A 与 B 的差集定义为一个集合, 其中的元素属于 A 但不属于 B . 记为 $A \setminus B = \{x: x \in A \text{ 但 } x \notin B\}$.

集合的补集: 如果 X 是一个基本集, $A \subset X$ 是 X 的子集, 称差集 $X \setminus A$ 为 A 关于 X 的补集, 记为 A' , 或 $\complement A$.

集合的以上运算可用下图示意(图 6.1.1)

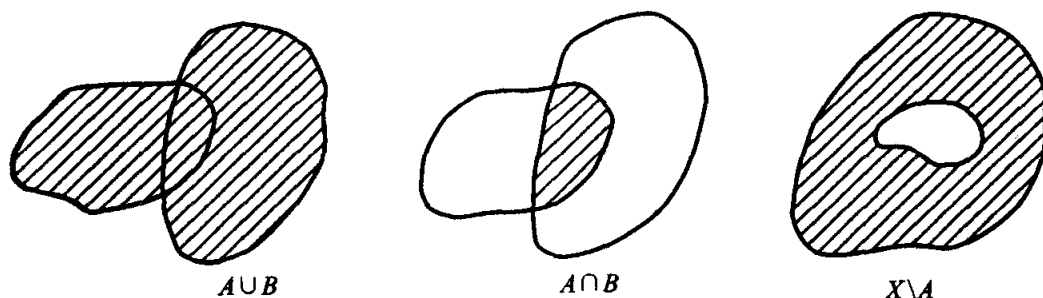


图 6.1.1

4. 集合的运算律

- | | |
|---|-------|
| (1) $A \cap B \subset A, A \cup B \supset A,$ | (等幂律) |
| (2) $A \cap B = B \cap A, A \cup B = B \cup A,$ | (交换律) |
| (3) $A \cap (B \cap C) = (A \cap B) \cap C, A \cup (B \cup C) = (A \cup B) \cup C,$ | (结合律) |
| (4) $A \cap (A \cup B) = A \cup (A \cap B) = A,$ | (吸收律) |
| (5) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C) = A,$ | (分配律) |
| $A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$ | (分配律) |

(6) $(A')' = A, (A \cap B)' = A' \cup B', (A \cup B)' = A' \cap B'.$ (对合律)

我们还可以考虑有限多个集合或无穷多个集合的并和交

$$\bigcup_{j \in I} A_j, \bigcap_{j \in I} A_j,$$

并且有性质

$$\begin{aligned} E \cap \left(\bigcup_{j \in I} A_j \right) &= \bigcup_{j \in I} (E \cap A_j), \\ \left(\bigcup_{j \in I} A_j \right)' &= \bigcap_{j \in I} (A_j)', \\ \left(\bigcap_{j \in I} A_j \right)' &= \bigcup_{j \in I} (A_j)', \end{aligned}$$

这些都留作练习.

5. 集合上的映射

在第二章中已经引进了函数的概念,那是经典情形下的,亦即,函数 f 的定义域 D 与值域 R 都仅在实数集合 $\mathbf{R}$ 中,或在 $\mathbf{R}^n$ 中.然而,近代科学技术中所遇到的量的变化范围、量与量的关系,远远超过了 $\mathbf{R}^n$.因此,在集合上引进“对应关系”就是必然的了.

定义 6.1.1 (映射、满射、单射) 设 A, B 是两个给定的集合,如果有一个“对应规则”,记为 f ,使对每个 $x \in A$,有惟一的 $y \in B$ 与之对应,记为 $f: A \rightarrow B$,或 $f(x) = y$,并称 A 为 f 的定义域, $f(A) = \{f(x): x \in A\} \subseteq B$ 为 f 的值域,或“象”.

对应关系 f 也称为“映射”,或“变换”,或“算子”等等.

易见, $f(A) \subseteq B$ 可能有 $f(A) = B$, 或 $f(A) \subset B$ 但 $f(A) \neq B$. 当 $f(A) = B$ 时,称 f 的象充满整个 B , 并称 f 为 A 到 B 满射. 在满射的情况下,对于每个 $y \in B$,至少存在一个 $x \in A$ 满足 $f(x) = y \in B$.

如果映射 f 是 A 到 B 的一一对应映射,亦即, $f(x) = f(y)$ 蕴含 $x = y$, 就称 f 是单射,或称一一映射. 对于一一映射 f , 它的逆映射 f^{-1} 定义为 $f^{-1}: f(A) \rightarrow A$, 满足

$$\text{当 } f(x) = y \text{ 时, } f^{-1}(y) = x.$$

例 1 设 $A = \{a, b, c\}, B = \{1, 2, 3, 4\}$. 则满足 $f: A \rightarrow B$ 的映射 f

$$f(a) = 1, f(b) = 2, f(c) = 3$$

是 A 到 B 的一一映射,但不是满射.

例 2 设 $A = B = \mathbf{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$, 则满足 $f: A \rightarrow B$ 的映射

$$f(n) = n + 1$$

是 A 到 B 的一一映射,故它是单射,并且是满射.

定义 6.1.2 (映射的复合) 设 A, B, C 是三个给定的集合,如果映射 $f: A \rightarrow B, g: B \rightarrow C$, 则称由 f 和 g 确定的、 $x \in A$ 到 $z \in C$ 的、满足 $z = h(x) =$

$g(f(x))$, 映射 $h: A \rightarrow C$, 为映射 f 与 g 的复合, 记为 $h = g \circ f$.

我们也常用记号

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

称为 f 的逆象集. 不要把逆象集 $f^{-1}(B)$ 与逆映射 f^{-1} 混淆了.

集合与集合上的映射的引进, 大大扩大了研究对象之间的关系的刻画方法, 例如, 不仅可以考虑数和数之间的函数关系, 而且还可刻画函数与数、函数与函数的关系, 等等, 从而使人们获得更为便利的数学工具.

例 3 设 $A = C[a, b]$, 令 $I: C[a, b] \rightarrow \mathbf{R}$ 为由下面的积分确定的映射:

$$I(f) = \int_a^b f(x) dx,$$

这里 $I(f)$ 为 f 的黎曼积分. 当 $f \in C[a, b]$ 时, $I(f)$ 视为 f 的函数, 称 $I(f)$ 为一个“(线性)泛函”. 下一节引入空间概念后, 对于线性泛函还要给出更一般的定义.

把具有一定性质的对象归在一起, 形成集合, 更便于研究该集合中的对象——亦即元素的性质以及元素之间的关系.

§ 6.2 运算结构

考虑集合中元素间的关系, 有两个重要方面, 一是元素之间的运算关系, 如元素的加法、乘法等; 另一方面是元素间的位置关系, 如两个元素间的距离, 即远近的程度等. 当然还有其他方面的关系, 读者会在以后的学习或研究中遇到.

本节介绍集合元素的一些代数运算, 被赋予代数运算的集合称为具有运算结构.

6.2.1 群、环、域

1. 群

从大家熟悉的实数集 $\mathbf{R}$ 开始. 在 $\mathbf{R}$ 中, 对两个实数 $x, y \in \mathbf{R}$, 它们的和 $x + y$ 可看作一种运算 $+$ 的结果. 这种运算有以下熟知的性质

- (i) $x, y \in \mathbf{R}$ 蕴含 $x + y \in \mathbf{R}$;
- (ii) $(x + y) + z = x + (y + z)$;
- (iii) 存在单位元 $0 \in \mathbf{R}$, 使对每个 $x \in \mathbf{R}$, 有 $x + 0 = 0 + x = x$;
- (iv) 对每个 $x \in \mathbf{R}$, 存在逆元 $-x \in \mathbf{R}$, 使 $x + (-x) = 0$.

这时, 我们称 $\mathbf{R}$ 在运算 $+$ 之下构成一个“群”. 一般地, 有

定义 6.2.1 (群) 如果在集合 G 上定义一个运算, 记为 $\cdot$, 满足

- (i) $x, y \in G$ 蕴含 $x \cdot y \in G$;

(封闭性)

(ii) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$; (结合律)

(iii) 存在单位元 $1 \in G$, 使对每个 $x \in G$, 有 $x \cdot 1 = x$; (单位元)

(iv) 对每个 $x \in G$, 存在逆元 $x^{-1} \in G$, 使 $x \cdot x^{-1} = 1$; ($\exists$ 逆元)

则称 G 在运算 $\cdot$ 之下构成一个群, 1 称为群的单位元. 如果运算 $\cdot$ 还满足

(v) $x \cdot y = y \cdot x$; (交换律)

则称 G 在运算 $\cdot$ 之下构成一个交换群.

定义 6.2.2 (子群) 设 G 在运算 $\cdot$ 之下构成一个群, $S \subset G$ 是 G 的子集, 若 S 在运算 $\cdot$ 之下也构成一个群, 则称 S 为 G 的子群.

在定义 6.2.1. 意义下, $\mathbf{R}$ 是运算 $\cdot$ 取作 $+$ 之下的一个“加群”, 单位元就是 0. 这里运算 $\cdot$ 的意义是很广泛的, 例如, 可取作加强乘法, 也可取作函数的复合, 等等.

例 1 令

$$\mathbf{R}^+ = \{x \in \mathbf{R}; x > 0\}.$$

把运算 $\cdot$ 取作数的乘法 $\times$. 不难验证, $\mathbf{R}^+$ 在运算 $\times$ 之下成为一个群 (请读者自行验证). 也请读者考虑, $(\mathbf{R}, \times)$ 是否能构成群?

例 2 取

$$C([a, b]) = \{f(x); f(x) \text{ 为 } [a, b] \text{ 上的连续函数}\}.$$

把运算 $\cdot$ 取作函数的加法 $f+g$, 即 $(f+g)(x) = f(x) + g(x)$, 则 $(C([a, b]), +)$ 构成一个加群, 但对乘法却不构成群 (为什么?).

例 3 设

$$\mathcal{F} = \{T; T(x) = ax + b, a, b \in \mathbf{R}, a \neq 0\},$$

$\mathcal{F}$ 是 $\mathbf{R}$ 到 $\mathbf{R}$ 的线性映射. $(\mathcal{F}, \circ)$ 构成一个群. 事实上, 对于 $S, T \in \mathcal{F}$, 令 $S(x) = ax + b, T(x) = cx + d$. 则

$$T(S(x)) = cS(x) + d = c(ax + b) + d = (ca)x + (cb + d),$$

因 $a, c \neq 0$, 故 $T \circ S \in \mathcal{F}$, 此即 (i). (ii) 显然成立. 对于 (iii), 单位元为 $I(x) = x$.

$T \in \mathcal{F}$ (若 $T(x) = ax + b$) 的逆元 $T^{-1} \in \mathcal{F}$ 为 $T^{-1}(x) = \frac{x}{a} - \frac{b}{a}$.

2. 同态、同构

定义 6.2.3 (群的同态、同构) 设 $(G, (\cdot_G))$ 与 $(H, (\cdot_H))$ 是两个群, $f: G \rightarrow H$ 是 G 到 H 的映射. 如果对于 $\alpha, \beta \in G$, 其象 $f(\alpha), f(\beta) \in H$ 满足

$$f(\alpha(\cdot_G)\beta) = f(\alpha)(\cdot_H)f(\beta),$$

则称映射 f 为 G 到 H 的同态. 如果 f 还是一一满射, 则称 f 为 G 到 H 的同构.

例 4 最熟悉的同构映射的例子是 $\varphi(x) = \log x$, 它是群 $(\mathbf{R}^+, \times)$ 到群 $(\mathbf{R}, +)$ 的同构: $\varphi(x \cdot y) = \varphi(x) + \varphi(y)$.

例 5 设 $G, \tilde{G}$ 为两个群, $\cdot$ 与 $\tilde{\cdot}$ 分别为其运算, $e, \tilde{e}$ 为其单位元, 令

$$f: G \rightarrow \tilde{G}, f(x) = \tilde{e}, \forall x \in G,$$

则 f 是 G 到 $\tilde{G}$ 的同态, 因为令 $z = x \cdot y \in G$,

$$f(x \cdot y) = f(z) = \tilde{e} = \tilde{e} \cdot \tilde{e} = f(x) \cdot f(y).$$

例 6 设 $(\mathbf{R}, +)$ 与 $(G, \times)$ 为两个群, 这里

$$G = \{z \in \mathbf{C} : |z| = 1\},$$

为复平面 $\mathbf{C}$ 上的单位圆周, 它在运算“ $\times$ ”之下成为一个群. 映射 $\varphi: \mathbf{R} \rightarrow \mathbf{C}$ 为 $\varphi(x) = e^{ix}$, 则 φ 是 $\mathbf{R}$ 到 $\mathbf{C}$ 上的同态.

如果集合是一个群, 就称它具有群结构. 群结构是运算结构的一种. 运算结构还有半群、环、域、线性空间等, 都是近代数学和各学科领域中常常遇到的. 例如“域”, 就是具有满足一定条件的两种运算的集合, 这种集合的最熟悉的例子是实数域 $\mathbf{R}$. 我们列举环与域的定义.

定义 6.2.4 (环、域) 如果在集合 R 上定义两个运算, 分别记为 $+$ 与 $\times$, 满足

- (i) R 对加法 $+$ 构成一个交换群; 其加法单位元记为 0 , 称为 R 的零元;
- (ii) R 对乘法 $\times$ 满足封闭性, 并满足结合律 $(x \times y) \times z = x \times (y \times z)$;
- (iii) R 对加法 $+$ 与乘法 $\times$ 满足加乘分配律

$$x \times (y + z) = x \times y + x \times z, \quad (y + z) \times x = y \times x + z \times x,$$

则称 R 在运算 $+$ 与 $\times$ 之下构成一个环. 若环 R 还满足:

- (iv) R 至少包含两个元素;
- (v) R 对乘法 $\times$ 满足交换律: $x \times y = y \times x$;
- (vi) R 对乘法 $\times$ 存在单位元, 记为 1 , 并且 R 的每个非零元 $x (x \neq 0)$ 都有逆元 x^{-1} ;

则称 R 在运算“ $+$ ”与“ $\times$ ”之下构成一个域. 环与域的实例很多很多, 我们仅列举以下的例.

例 7 $(\mathbf{R}, +, \times)$ 是最熟知的环, 同时它也是一个域. 这里的 $+$ 与 $\times$ 就是通常 $\mathbf{R}$ 上的加法与乘法.

其实, 以下集合

$$\mathbf{Z} = \{x : x \text{ 为整数 } 0, \pm 1, \pm 2, \dots\},$$

$$\mathbf{Q} = \{x : x \text{ 为有理数}\},$$

$$\mathbf{R} = \{x : x \text{ 为实数}\},$$

$$\mathbf{C} = \{x : x \text{ 为复数}\},$$

在通常加法与乘法运算之下构成环.

例 8 在计算机科学中有广泛应用的、仅含两个元素的域 $\mathbf{Z}_2 = \{0, 1\}$, 其加法与乘法可用表列出:

$+$	0	1	$\cdot$	0	1
0	0	1	0	0	0
1	1	0	1	0	1

群的理论在近代物理中也有重要的应用,为此,我们再给两个定理,让读者了解一下群与环的一些性质.

定理 6.2.1 (同态映群到群) 设 $f: G \rightarrow G'$ 是群 G 到 G' 的同态映射,则 f 把 G 的单位元映到 G' 的单位元 $e' = f(e)$; 把 $a \in G$ 的逆元 a^{-1} 映到 $f(a) \in G'$ 的逆元: $f(a^{-1}) = (f(a))^{-1}$; 把子群 $H \subset G$ 映到子群 $f(H) \subset G'$.

定理 6.2.2 (同态映环到环) 设 $f: R \rightarrow R'$ 是环 $(R, (+_R), (\times_R))$ 到 $(R', (+_{R'}), (\times_{R'}))$ 的同态映射,亦即,

$$f(x(+_R)y) = f(x)(+_R')f(y),$$

$$f(x(\times_R)y) = f(x)(\times_{R'})f(y),$$

则 f 把 R 的零元映到 R' 的零元 $0' = f(0)$; 把 $a \in R$ 的负元 $-a \in R$ 映到 $f(a) \in R'$ 的负元 $-f(a) \in R'$; 另一方面, f 把 R 的单位元 1 映到 R' 的单位元 $1' = f(1)$, 把 $a \in R$ 的逆元 $a^{-1} \in R$ 映到 $f(a) \in R'$ 的逆元 $(f(a))^{-1} \in R'$.

6.2.2 线性空间

1. 线性空间

另一种重要的运算结构是“线性空间”,它包含两种运算.我们可以从读者熟悉的向量空间出发.把解析几何中的三维空间 $\mathbf{R}^3$ 用以下集合的记号来表示:

$$\mathbf{R}^3 = \{x = (x_1, x_2, x_3): x_1, x_2, x_3 \text{ 为实数}\}.$$

在 $\mathbf{R}^3$ 中,元素之间的“加法”和“数乘”运算是大家熟悉的:对于 $x, y \in \mathbf{R}^3$,

$$x + y = (x_1, x_2, x_3) + (y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

$$\alpha \cdot x = \alpha \cdot (x_1, x_2, x_3) = (\alpha x_1, \alpha x_2, \alpha x_3) \quad (\alpha \in \mathbf{R}).$$

不难验证,上述“加法”满足定义 6.2.1,故 $\mathbf{R}^3$ 成为一个加群;对于上述“数乘”,满足分配律与结合律.由此可以在一般的集合上定义向量空间,或称线性空间.

定义 6.2.5 (线性空间) 设 X 为一集合,其元素为 $x, y, \dots \in X$,如果在 X 中定义其元素之间的“加法” $x + y$ 与“数乘” $\alpha \cdot x$ 运算,使其具有以下性质:

(I) “加法”运算 $x + y$,使得 X 成为一个交换群,亦即满足

$$(I_1) \quad x + y = y + x \quad (x, y \in X); \quad (\text{交换律})$$

$$(I_2) \quad (x + y) + z = x + (y + z) \quad (x, y, z \in X); \quad (\text{结合律})$$

(I₃) X 中存在惟一的元素 0 ,使对任何 $x \in X$,有 $x + 0 = x$,并称 0 为 X 的零元素;

(I₄) 对 X 中的任一元 $x \in X$,存在惟一的元素 $-x \in X$,满足 $x + (-x) = 0$,并称 $-x$ 为 X 的逆元素.

(II) “数乘”运算 $\alpha \cdot x$,也简记为 αx ,满足

$$(II_1) \quad \alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y \quad (x, y \in X, \alpha \text{ 为常数}); \quad (\text{分配律})$$

$$(\Pi_2) (\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x \quad (x \in X, \alpha, \beta \text{ 为常数}); \quad (\text{分配律})$$

$$(\Pi_3) (\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x) \quad (x \in X, \alpha, \beta \text{ 为常数}); \quad (\text{结合律})$$

$$(\Pi_4) 1 \cdot x = x.$$

则称 X 为向量空间, 或线性空间, 其元素也称为向量.

显然 $\mathbf{R}^3$ 为(三维)向量空间, 或称三维线性空间. 一般地, 可把

$$\mathbf{R}^n = \{x = (x_1, x_2, \dots, x_n) : x_1, x_2, \dots, x_n \text{ 为实数}\}$$

定义为(n 维)向量空间, 或称 n 维线性空间.

2. 矩阵空间

在第一章中已经学到矩阵, 它的定义, 性质以及它在解代数方程组中的应用. 现在我们要以向量空间的观点来看待 n 阶方阵所成的集合, 研究其性质及运算. 同时, 由于矩阵在自然科学各个领域的科学研究中、在工程技术和工业生产中有非常广泛的应用, 为此, 我们将对其进行较为深入的讨论.

记

$$M_{m \times n} = \left\{ A : A = (a_{ij}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \text{ 为 } m \times n \text{ 阶矩阵} \right\};$$

当 $m = n$ 时, 得到 n 阶方阵, 简记为 $M_n = M_{n \times n}$.

$M_{m \times n}$ 在矩阵加法与矩阵数乘运算之下成为一个线性空间.

对于 $A, B \in M_{m \times n}$ 与常数 α , 在第一章中已经定义:

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij});$$

$$\alpha A = \alpha (a_{ij}) = (\alpha a_{ij}).$$

易证, $M_{m \times n}$ 是一个线性空间, 称为矩阵空间.

在线性空间 $M_{m \times n}$ 与 $M_{n \times s}$ 之间, 可以定义两个矩阵 $A \in M_{m \times n}$ 与 $B \in M_{n \times s}$ 的乘法

$$\begin{aligned} AB &\equiv A \times B = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} (b_{kl})_{1 \leq k \leq n, 1 \leq l \leq s} \\ &= (c_{il})_{1 \leq i \leq m, 1 \leq l \leq s}, \end{aligned}$$

$$\text{其中 } c_{il} = \sum_{j=1}^n a_{ij} b_{jl} \quad (1 \leq i \leq m, 1 \leq l \leq s).$$

6.2.3 方阵空间

本小节的内容是第一章中“矩阵”部分的继续, 我们要以空间的观点来研究 n 阶方阵, $A = (a_{ij}) \in M_{n \times n}$, 并将 $M_{n \times n}$ 中的元素(方阵)简化, 使得既保持它的一些基本性质, 又具有简单的形式, 这对矩阵的定性讨论和实际计算都十分有益.

1. 方阵的标准型

(1) 相似方阵

定义 6.2.6 (相似方阵) 对于两方阵 $A, B \in M_{n \times n}$, 如果存在可逆方阵 $U \in M_{n \times n}$, 使得

$$B = U^{-1}AU,$$

则称 A 与 B 相似, 记为 $A \sim B$.

方阵相似关系的性质

性质 1 设 $A, B, C \in M_{n \times n}$, 则

- (i) $A \sim A$; (自反性)
- (ii) $A \sim B$ 蕴含 $B \sim A$; (对称性)
- (iii) $A \sim B$ 与 $B \sim C$ 蕴含 $A \sim C$. (传递性)

证 对于(i), 只要取 $U = I$, 即得 $I^{-1}AI = A$.

对于(ii), 由 $A \sim B$, 知存在 $U \in M_{n \times n}$, 使得 $B = U^{-1}AU$. 从而 $A = UBU^{-1} = (U^{-1})^{-1}BU^{-1}$, 此即 $B \sim A$.

至于(iii), 由 $A \sim B$, 知存在 $U \in M_{n \times n}$, 使得 $B = U^{-1}AU$; 由 $B \sim C$, 知存在 $V \in M_{n \times n}$, 使得 $C = V^{-1}BV$; 于是

$C = V^{-1}BV = V^{-1}(U^{-1}AU)V = (V^{-1}U^{-1})A(UV) = (UV)^{-1}AUV$. 所以 $A \sim C$.

性质 2 相似方阵的行列式的值相等.

证 由 $B = U^{-1}AU$, 得 $|B| = |U^{-1}||A||U| = |U|^{-1}|A||U| = |A|$.

性质 3 相似方阵或者都可逆, 或者都不可逆. 在都可逆的情形, 逆方阵也相似.

证 若 $A \sim B$, 则由性质 1 知, $|A| = |B|$, 于是这两个行列式同时为 0 或同时不为 0, 因此方阵 A 与 B 也同时可逆或同时不可逆.

在 A 与 B 同时可逆时, 由等式 $B = U^{-1}AU$ 两边取逆, 便得

$$B^{-1} = U^{-1}A^{-1}U,$$

故 $A^{-1} \sim B^{-1}$.

性质 4 若 $A \sim B$, 则对任意自然数 n , 有 $A^n \sim B^n$.

证 因为 $A \sim B$, 所以存在可逆矩阵 U , 使得 $U^{-1}AU = B$. 于是对任何正整数 n , 有

$$\begin{aligned} B^n &= (U^{-1}AU)(U^{-1}AU)\cdots(U^{-1}AU) \\ &= U^{-1}A(UU^{-1})A(UU^{-1})\cdots(UU^{-1})AU \\ &= U^{-1}A^nU, \end{aligned}$$

此即 $A^n \sim B^n$.

上述性质在求矩阵正整数幂时非常有用. 例如

例 9 若

$$A = \begin{bmatrix} 3 & -4 \\ -1 & -2 \end{bmatrix}; \quad U = \begin{bmatrix} 4 & 1 \\ -1 & -1 \end{bmatrix}.$$

且 $B = U^{-1}AU$. 试求 A^n , 其中 n 为正整数.

解 容易求得 $U^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & -4 \end{bmatrix}$, 于是

$$B = U^{-1}AU = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} 3 & -4 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

由 $B = U^{-1}AU$ 得 $A = UBU^{-1}$, 按性质 4 有

$$\begin{aligned} A^n &= UB^nU^{-1} = \frac{1}{3} \begin{bmatrix} 4 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}^n \begin{bmatrix} 1 & 1 \\ -1 & -4 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 4 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & (-1)^n \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -4 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 4(2^n) - (-1)^n & 4(2^n - (-1)^n) \\ -2^n - (-1)^n & -2^n + 4(-1)^n \end{bmatrix}. \end{aligned}$$

(2) 特征值与特征向量

例 9 对我们有一个重要的启发, 就是, 当 A 与一个对角阵 B 相似时, 计算 A 的正整数幂次就比较简单. 那么, 什么样的矩阵能相似于一个对角矩阵呢? 又如何求出对角矩阵 B 及可逆矩阵 U 呢? 关键在于本节中要介绍的方阵的特征值与特征向量.

让我们作一个简单的分析. 如果对于给定的 n 阶方阵 A , 存在可逆方阵 U , 使得 $U^{-1}AU = B$, 而 B 是对角阵

$$B = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

记 U 的列向量为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 即 $U = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 于是 $AU = UB$ 成为

$$AU = A(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)B = UB.$$

从而

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n).$$

这样得到

$$A\alpha_1 = \lambda_1\alpha_1, \quad A\alpha_2 = \lambda_2\alpha_2, \quad \dots, \quad A\alpha_n = \lambda_n\alpha_n. \quad (6.2.1)$$

这些关系是方阵 U 的列向量与 B 的主对角线上的元素所必须满足的. 于是就引

入下面的定义.

定义 6.2.7 (特征值与特征向量) 设 $A \in M_n$ 为 n 阶方阵, 如果常数 λ 与非零列向量 α 满足

$$A\alpha = \lambda\alpha,$$

则称数 λ 为方阵 A 的特征值, 称向量 α 为方阵 A 的属于 λ 的特征向量.

对于特征值与特征向量, 有以下说明.

从定义可见, 若向量 $\alpha \neq 0$, 使得 $A\alpha = \lambda\alpha$, 则对任意非零数 $k \neq 0$, 向量 $k\alpha$ 是非零向量, 且

$$A(k\alpha) = k(A\alpha) = k(\lambda\alpha) = \lambda(k\alpha).$$

可见, 当 α 是属于特征值 λ 的特征向量时, 则对于任何非零数 k , 向量 $k\alpha$ 也是属于 λ 的特征向量. 所以, 如果 λ 是 A 的特征值, 则属于 λ 的特征向量有无穷多个.

现在来求方阵 A 的全部特征值与特征向量. 由 $A\alpha = \lambda\alpha$ 得 $A\alpha - \lambda\alpha = 0$, 即

$$(\lambda I - A)\alpha = 0.$$

这表明 α 是齐次线性方程组

$$(\lambda I - A)x = 0 \quad (6.2.2)$$

的非零解, 从而必须有

$$|\lambda I - A| = 0. \quad (6.2.3)$$

(6.2.3) 中的行列式是 λ 的 n 次方程.

定义 6.2.8 (特征方程、特征多项式) 对于 n 阶方阵 $A \in M_n$, 我们称 λ 的 n 次方程 (6.2.3) 为 A 的特征方程. 此特征方程相应的行列式

$$|\lambda I - A| = |\lambda I - (a_{ij})| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

是 λ 的 n 次多项式, 称为 A 的特征多项式.

注 对于特征方程与特征多项式, 我们有以下说明:

(i) (6.2.3) 式表明, A 的特征值必定是其特征方程的根; 反之, 若 λ 是 (6.2.3) 的根, 则齐次线性方程组 (6.2.2) 有非零解 α , 亦即有非零向量 α , 使得 $A\alpha = \lambda\alpha$, 从而 λ 是方阵 A 的特征值.

(ii) A 的全部特征值就是其特征方程的全部的根 (实根与虚根); 而对每个特征值 λ , 齐次线性方程组 (6.2.2) 的一切非零解就是属于它的全部特征向量.

(iii) 求矩阵 A 的特征值、特征向量的步骤:

① 计算行列式 $|\lambda I - A|$, 求出 $|\lambda I - A| = 0$ 的全部根, 即方阵 A 的特征

值.

② 对于 A 的每一个特征值 λ , 求齐次线性方程组

$$(\lambda I - A)x = 0$$

的一个基础解系 $\alpha_1, \alpha_2, \dots, \alpha_s$, 并写成列向量的形式. 于是, 方阵 A 的全部特征向量可表示为“通解”的形式

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s, \quad (6.2.4)$$

其中 $k_1, k_2, \dots, k_s$ 为不全为 0 的任意常数.

例 10 求

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

的全部特征值与特征向量.

解 由

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2 (\lambda - 5) = 0$$

得 3 个特征值 $\lambda_1 = \lambda_2 = -1, \lambda_3 = 5$.

对于 $\lambda_1 = \lambda_2 = -1$, 齐次线性方程组为

$$\begin{cases} -2x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 - 2x_3 = 0, \end{cases}$$

求得它的基础解系为 $\alpha_1 = (-1, 1, 0)^T, \alpha_2 = (-1, 0, 1)^T$.

由此得知, 属于特征值 -1 的全部特征向量为 $k_1 \alpha_1 + k_2 \alpha_2$, 其中 k_1, k_2 是不全为 0 的任意常数.

类似地, 对于特征值 $\lambda_3 = 5$, 由

$$\begin{cases} 4x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 + 4x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 + 4x_3 = 0, \end{cases}$$

求得基础解系 $\alpha_3 = (1, 1, 1)^T$. 故属于特征值 5 的全部特征向量为 $k_3 \alpha_3$, 其中 $k_3 \neq 0$.

例 11 求方阵 $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ 的全部特征值与特征向量.

解 由

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -1 \\ 1 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2 + 1 = 0$$

得 $\lambda_1 = 1 + i, \lambda_2 = 1 - i$, 这是两个复特征值.

对于 $\lambda_1 = 1 + i$, 解齐次线性方程组 $[(1 + i)I - A]x = 0$, 即解

$$\begin{cases} ix_1 - x_2 = 0, \\ x_1 + ix_2 = 0. \end{cases}$$

由行变换将系数方阵化简

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix},$$

故一般解为 $x_1 = -ix_2$, 于是一个基础解系为 $\begin{bmatrix} -i \\ 1 \end{bmatrix}$, 因此属于 $1 + i$ 的全部特征向量为

$$k \begin{bmatrix} -i \\ 1 \end{bmatrix} \quad (k \neq 0 \text{ 为任意复数}).$$

类似地, 对应于 $\lambda_2 = 1 - i$, 得一个基础解系 $\begin{bmatrix} i \\ 1 \end{bmatrix}$, 故属于 $1 - i$ 的全部特征向量为

$$k \begin{bmatrix} i \\ 1 \end{bmatrix} \quad (k \neq 0 \text{ 为任意复数}).$$

若限制在实数范围内, 例 11 中的方阵 A 没有实特征值与实特征向量.

定理 6.2.3 (相似方阵的特征值与特征向量) 相似方阵有相同的特征多项式与特征值.

证 事实上, 设 $A \sim B$, 故存在可逆方阵 U , 使得 $U^{-1}AU = B$, 于是,

$$\begin{aligned} |\lambda I - B| &= |\lambda U^{-1}U - U^{-1}AU| = |U^{-1}(\lambda I - A)U| \\ &= |U^{-1}| |\lambda I - A| |U| = |U|^{-1} |\lambda I - A| |U| = |\lambda I - A|. \end{aligned}$$

这就是要证明的.

注意, 这个定理之逆并不成立. 特征多项式相同的方阵未必相似. 例如, 方阵

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

的特征多项式都是 $(\lambda - 1)^2$, 但是, 对于任何可逆方阵 U 都成立 $U^{-1}IU = I \neq B$, 故 I 与 B 并不相似.

(3) 方阵可对角化的条件

定义 6.2.9 (可对角化方阵) 若方阵 A 相似于对角方阵, 则称 A 可对角化.

什么方阵相似于对角方阵呢? 也就是什么方阵可以对角化呢?

定理 6.2.4 (可对角化的充要条件) n 阶方阵 A 可对角化的充要条件是: A 有 n 个线性无关的特征向量.

证 必要性: 设 A 可对角化, 即存在可逆方阵 U , 使得

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{bmatrix}.$$

此时, (6.2.1) 式 $A\alpha_1 = \lambda_1\alpha_1, A\alpha_2 = \lambda_2\alpha_2, \dots, A\alpha_n = \lambda_n\alpha_n$ 表明 U 的列向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 都是 A 的特征向量. 又由于 U 可逆, 故 U 的秩为 n . 从而 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关, 于是, A 有 n 个线性无关的特征向量.

充分性: 设 A 有 n 个线性无关的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 它们分别属于特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$, 即

$$A\alpha_1 = \lambda_1\alpha_1, \quad A\alpha_2 = \lambda_2\alpha_2, \quad \dots, \quad A\alpha_n = \lambda_n\alpha_n. \quad (6.2.5)$$

由于 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关, 得知以它们作为列向量的方阵 $U = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 的秩为 n , 因而是可逆的. 把 (6.2.5) 改写为

$$\begin{aligned} A(\alpha_1, \alpha_2, \dots, \alpha_n) &= (A\alpha_1, A\alpha_2, \dots, A\alpha_n) \\ &= (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)(\lambda_1, \lambda_2, \dots, \lambda_n)^T \\ &= (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{bmatrix}. \end{aligned}$$

也就是

$$AU = U \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{bmatrix}.$$

从而

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{bmatrix}.$$

定理得证.

例 12 方阵 $A = \begin{bmatrix} 2 & 5 & -1 \\ -1 & -3 & 0 \\ 2 & 3 & -2 \end{bmatrix}$ 能否对角化?

解 只需考虑 A 是否存在三个线性无关的特征向量. 由

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & -5 & 1 \\ 1 & \lambda + 3 & 0 \\ -2 & -3 & \lambda + 2 \end{vmatrix} = (\lambda + 1)^3 = 0$$

知 A 有三重特征值 $\lambda = -1$. 以 $\lambda = -1$ 代入齐次线性方程组 $(\lambda I - A)x = 0$, 对系数方阵 $\lambda I - A = -I - A$ 施行变换

$$\left| \begin{array}{ccc|ccc} -3 & -5 & 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & -3 & -5 & 1 \\ -2 & -3 & 1 & -2 & -3 & 1 \end{array} \right| \longrightarrow \left| \begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 2 & 0 \\ -3 & -5 & 1 & 0 & 1 & 1 \\ -2 & -3 & 1 & 0 & 0 & 0 \end{array} \right|.$$

系数方阵的秩为 2, 而未知数个数为 3, 故基础解系只含有 1 个线性无关的解, 因而 A 不能对角化.

定理 6.2.5 (特征向量的线性无关性) 方阵 A 的属于不同特征值的特征向量是线性无关的.

证 用数学归纳法证明. 当 $n = 1$ 时, 由定义, 特征向量是非零向量, 因此这个特征向量是线性无关的.

设分别属于不同的特征根 $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ 的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$ 线性无关. 现在假设与 $\lambda_1, \lambda_2, \dots, \lambda_{k-1}$ 都不相同的另一个特征值 λ_k , 及所对应的特征向量 α_k . 我们证明,

$$l_1 \alpha_1 + l_2 \alpha_2 + \dots + l_{k-1} \alpha_{k-1} + l_k \alpha_k = 0 \quad (6.2.6)$$

当且仅当 $l_1 = l_2 = \dots = l_{k-1} = l_k = 0$.

为此, 在式 (6.2.6) 两边左乘以矩阵 A , 得到

$$l_1 A \alpha_1 + l_2 A \alpha_2 + \dots + l_{k-1} A \alpha_{k-1} + l_k A \alpha_k = 0,$$

再以 $A \alpha_j = \lambda_j \alpha_j$ ($j = 1, 2, \dots, k$) 代入上式, 得到

$$l_1 \lambda_1 \alpha_1 + l_2 \lambda_2 \alpha_2 + \dots + l_{k-1} \lambda_{k-1} \alpha_{k-1} + l_k \lambda_k \alpha_k = 0. \quad (6.2.7)$$

从 (6.2.7) 式中减去 λ_k 乘以 (6.2.6), 得

$$\begin{aligned} & l_1 (\lambda_1 - \lambda_k) \alpha_1 + l_2 (\lambda_2 - \lambda_k) \alpha_2 + \dots + l_{k-1} (\lambda_{k-1} - \lambda_k) \alpha_{k-1} + l_k (\lambda_k - \lambda_k) \alpha_k \\ & = l_1 (\lambda_1 - \lambda_k) \alpha_1 + l_2 (\lambda_2 - \lambda_k) \alpha_2 + \dots + l_{k-1} (\lambda_{k-1} - \lambda_k) \alpha_{k-1} = 0. \end{aligned} \quad (6.2.8)$$

由归纳法假设, $\alpha_1, \alpha_2, \dots, \alpha_{k-1}$ 线性无关, 且 $\lambda_j \neq \lambda_k$ ($j = 1, 2, \dots, k-1$). 因此, 要使 (6.2.8) 成立, 必须

$$l_1 = l_2 = \dots = l_{k-1} = 0,$$

以此代入 (6.2.6) 得 $l_k \alpha_k = 0$. 但是 $\alpha_k \neq 0$, 故必有 $l_k = 0$. 这就证明了 $l_1 = l_2 = \dots =$

$l_{k-1} = l_k = 0$, 也就是 $\alpha_1, \alpha_2, \dots, \alpha_{k-1}, \alpha_k$ 线性无关. 从而由归纳法便得到定理的证明.

推论 若 n 阶方阵 A 有 n 个互异的特征值, 则 A 可对角化.

例 13 将方阵 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 3 & 6 \end{bmatrix}$ 对角化.

解 首先求特征值. 由

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 & -3 \\ -2 & \lambda - 1 & -3 \\ -3 & -3 & \lambda - 6 \end{vmatrix} = \lambda^3 - 8\lambda^2 - 9\lambda = 0,$$

得 3 个互异的特征值 $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = 9$, 由定理 6.2.5 的推论知, A 是可对角化的.

其次, 求特征向量. 以 $\lambda_1 = 0$ 代入齐次线性方程组 $\lambda I - A = 0$, 并对其系数方阵施行变换,

$$\begin{bmatrix} -1 & -2 & -3 \\ -2 & -1 & -3 \\ -3 & -3 & -6 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & -2 & -3 \\ 0 & 3 & 3 \\ 0 & 3 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & -2 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

得一般解 $x_2 = -x_3, x_1 = -x_3$. 令 $x_3 = 1$, 便得到属于 λ_1 的特征向量

$$\alpha_1 = (-2, -2, 1)^T.$$

以 $\lambda_2 = -1$ 代入齐次线性方程组 $\lambda I - A = 0$, 并对其系数方阵施行变换,

$$\begin{bmatrix} -2 & -2 & -3 \\ -2 & -2 & -3 \\ -3 & -3 & -7 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

得一般解 $x_1 = -x_2, x_3 = 0$. 令 $x_2 = 1$, 便得到属于 λ_2 的特征向量

$$\alpha_2 = (-1, 1, 0)^T.$$

类似地, 对于 $\lambda_3 = 9$, 齐次线性方程组 $\lambda I - A = 0$ 的系数方阵

$$\begin{bmatrix} 8 & -2 & -3 \\ -2 & 8 & -3 \\ -3 & -3 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

得一般解 $x_1 = x_2, 2x_2 = x_3$. 令 $x_3 = 2$, 得到属于 λ_3 的特征向量

$$\alpha_3 = (1, 1, 2)^T.$$

最后, 令

$$U = (\alpha_1, \alpha_2, \alpha_3) = \begin{bmatrix} -1 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix},$$

则

$$U^{-1}AU = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

有一个比定理 6.2.5 更为一般的结果.

定理 6.2.6 (特征向量的线性无关性) 若 $\lambda_1, \lambda_2, \dots, \lambda_m$ 是方阵 A 的不同的特征值, 而 $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ ($i = 1, 2, \dots, m$), 是 A 的属于 λ_i 的线性无关的特征向量, 则向量组

$$\alpha_{11}, \alpha_{12}, \dots, \alpha_{1s_1}, \alpha_{21}, \alpha_{22}, \dots, \alpha_{2s_2}, \dots, \alpha_{m1}, \alpha_{m2}, \dots, \alpha_{ms_m} \quad (6.2.9)$$

线性无关.

定理的证明从略.

这个定理告诉我们, 判别 A 可否对角化的最一般方法是: 求出 A 的全部不同的特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$, 对每一个特征值 λ_i , 求出齐次线性方程组 $(\lambda_i I - A)x = 0$ 的一个基础解系, 将所有这些基础解系依次排列起来, 得到向量组 (6.2.9), 如果这组向量的个数等于 A 的阶数 n , 则 A 可对角化, 否则 A 不能对角化.

在 A 可对角化时, 将向量组 (6.2.9) 作为列向量依次排列所得之方阵, 记为 U , 则 $U^{-1}AU = B$, 其中 B 是对角阵, 其主对角线上的第 i 个元素是 U 中第 i 个列向量所对应的特征值. 由于可能有重特征值, 因而, B 的主对角线上的元素可能有相同者.

由于每个齐次线性方程组的基础解系都不是惟一的, 因此 U 的取法也是不惟一的. 但是, $U^{-1}AU$ 的主对角线上的元素是 A 的全体特征值, 故除了主对角线上元素的次序不同外, $U^{-1}AU$ 是惟一确定的.

例 14 将方阵 $A = \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix}$ 对角化.

解 由

$$|\lambda I - A| = \begin{vmatrix} \lambda - 4 & -6 & 0 \\ 3 & \lambda + 5 & 0 \\ 3 & 6 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2(\lambda + 2) = 0,$$

得特征值 $\lambda = 1$ (二重)、 $\lambda = -2$ (单重). 对于 $\lambda = 1$, 齐次线性方程组 $(\lambda I - A)x = 0$ 成为

$$\begin{cases} -3x_1 - 6x_2 = 0, \\ 3x_1 + 6x_2 = 0, \\ 3x_1 + 6x_2 = 0. \end{cases}$$

此方程组的通解为 $x_1 = -2x_2$, 其中 x_2, x_3 可取任意值. 由此可得一个基础解系

$$\alpha_1 = (-2, 1, 0)^T, \quad \alpha_2 = (0, 0, 1)^T.$$

对于 $\lambda = -2$, 写出对应齐次线性方程组 $(\lambda I - A)x = 0$ 的系数方阵并作变换

$$\begin{bmatrix} -6 & -6 & 0 \\ 3 & 3 & 0 \\ 3 & 6 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

由此得 $x_1 = -x_2, x_3 = x_2$. 于是, 一个基础解系为 $\alpha_3 = (-1, 1, 1)^T$. 最后, 得到

$$U = \begin{bmatrix} -2 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad U^{-1}AU = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

(4) 若尔当标准型

由上一小节知道, 并不是每一个 n 阶方阵 A 都能对角化. 但是, 能否相似于某种比较简单的方阵呢? 有一种称为若尔当(Jordan)标准型的方阵可以起这样的作用.

定义 6.2.10 (若尔当块) 形如

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \dots & & \\ & & & \dots & \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{bmatrix}$$

的 n 阶方阵称为特征值 λ 的 n 阶若尔当块, 其中 λ 是实数或复数.

例如, 任意一个 1 阶方阵都是一阶若尔当块. 又如

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

分别为 2 阶与 3 阶若尔当块.

定义 6.2.11 (若尔当方阵) 形如

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \cdots & \\ & & & J_s \end{bmatrix}$$

的方阵称为若尔当方阵, 其中 $J_1, J_2, \dots, J_s$ 是若尔当块; 或说, 若方阵可表示成分块矩阵的形式, 使得主对角线上的子块均是若尔当块, 其余全是 0 子块.

例如,

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & J_3 & \\ & & & J_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & & & \\ & 2 & 0 & & \\ & & 2 & 1 & \\ & & & 2 & 0 \\ & & & & 2 & 1 \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{bmatrix},$$

其中

$$J_1 = [0], \quad J_2 = [2], \quad J_3 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad J_4 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix},$$

J 就是一个若尔当方阵.

由此知道, 对角阵是一个若尔当方阵.

定理 6.2.7 (若尔当标准型) 复数域上的任一个 n 阶方阵 A 都相似于一个若尔当方阵, 称之为若尔当标准型. 若不计若尔当块的排列次序, 则 A 的若尔当标准型是惟一确定的.

我们已经知道, 例 12 中的方阵 A 不能对角化. 但是, 若取

$$U = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & -2 \end{bmatrix},$$

则可以验证

$$U^{-1}AU = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

这就是 A 的若尔当标准型.

我们不再详述求若尔当标准型的方法, 只注意到, 由于相似方阵有相同的特征多项式, 容易看出, 方阵 A 的若尔当标准型主对角线上元素恰好是 A 的特征值, 并且任一特征值在主对角线上出现的次数, 就等于它的重数.

2. 实二次型

本教程的第一章解析几何中我们学习过,含有 2 个变量 x 与 y 的二次方程一般表示平面上的一条曲线,称为二次曲线;含有 3 个变量 x, y 与 z 的二次方程一般表示空间中的一个曲面,称为二次曲面.在那里,我们还讨论了二次曲面分类,但是需要把二次曲线或曲面的一般形式化为标准形式才能判定.化标准型,需要作变数变换,消去混合项(例如 xy, yz 等项),把二次式化为只含平方项的代数.现在我们利用矩阵工具来完成这个工作.

本节介绍四个内容:(1) 正交方阵;(2) 正交化方法;(3) 实二次型化简;(4) 正定实二次型.

(1) 正交方阵

首先,我们把三维向量的内积概念推广到 n 维情形.

定义 6.2.12 (n 维向量的内积) 设 $\alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n)$ 为 2 个 n 维向量,我们称

$$a_1 b_1 + a_2 b_2 + \dots + a_n b_n$$

为 α, β 的内积,记为 (α, β) .

由定义立即知道,三维向量内积的交换律、分配律、数乘律等都可推广到 n 维向量情形

$$(\alpha, \beta) = (\beta, \alpha), \quad (\text{交换律})$$

$$(\alpha_1 + \alpha_2, \beta) = (\alpha_1, \beta) + (\alpha_2, \beta), \quad (\text{分配律})$$

$$k(\alpha, \beta) = (k\alpha, \beta) \quad (k \text{ 为实数}). \quad (\text{数乘律})$$

也与三维向量内积类似可以定义正交概念.

定义 6.2.13 (正交性) 若 2 个 n 维向量 $\alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n)$ 的内积为 0,则称 2 个向量正交.

定义 6.2.14 (长度或模) 称 $\sqrt{(\alpha, \alpha)}$ 为向量 α 的长度或模,并记为 $|\alpha| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$.

长度为 1 的向量称为单位向量.对于任何非零向量 α ,向量 $\frac{\alpha}{|\alpha|}$ 是单位向量.

定义 6.2.15 (正交方阵) 若方阵 A 满足 $A^T A = I$,则称 A 为正交方阵.例如,

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}; \quad \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

都是正交方阵.

现在讨论正交方阵的性质.

定义 6.2.16 (正交向量组、法正交组) 设 $\{\alpha_1, \alpha_2, \dots, \alpha_s\}$ 为一个不含零向量的 n 维向量组, 若其中向量都两两彼此正交, 则称其为正交向量组. 由单位向量所组成的正交向量组称为单位正交向量组, 或法正交组.

正交向量组有一个重要性质.

定理 6.2.8 (线性无关性) 正交向量组是线性无关的.

证 设 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 是一正交向量组, 且

$$k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s = 0.$$

则

$$(\alpha_i, k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s) = 0 \quad (i = 1, 2, \dots, s),$$

即

$$k_1 (\alpha_i, \alpha_1) + k_2 (\alpha_i, \alpha_2) + \dots + k_s (\alpha_i, \alpha_s) = 0 \quad (i = 1, 2, \dots, s). \quad (6.2.10)$$

由于 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 两两正交, 且均非零向量, 故上式左边除了 $(\alpha_i, \alpha_i) \neq 0$ 外, 其余全部为 0. 因此 (6.2.10) 成为

$$k_i (\alpha_i, \alpha_i) = 0 \quad (i = 1, 2, \dots, s).$$

这样, 只有 $k_i = 0 \quad (i = 1, 2, \dots, s)$, 故 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 线性无关.

关于正交方阵, 我们有

定理 6.2.9 (正交方阵的行、列向量组) 正交方阵的行向量组与列向量组都是法正交组; 反之, 列向量组或行向量组为法正交组的方阵是正交方阵.

证 设 $A \in M_n$ 为正交方阵, 记为

$$A = (a_{ij}) = [\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n]; \quad A^T = (a_{ji}) = \begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T \end{bmatrix}.$$

由 $A^T A = I$ 得

$$\begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T \end{bmatrix} [\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \dots & \\ & & & 1 \end{bmatrix}.$$

可见

$$(\alpha_i, \alpha_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

此即正交方阵的列向量都是单位向量,且彼此两两正交.

对行向量组,只需利用“ $A^T A = I$ 等价于 $AA^T = I$ ”,便可得到.

反之,若列向量组或行向量组为法正交组,由定义便可得到方阵是正交方阵.

(2) 正交化方法

一组线性无关的向量组未必正交,一个方阵也未必是正交方阵.但是,从一组线性无关的向量组 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 出发,利用它们的线性组合,可以构造一组法正交组 $(\beta_1, \beta_2, \dots, \beta_s)$,这种方法称为施密特(O. Schmidt)正交化方法.

施密特正交化方法

第一步 由线性无关的向量组 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 产生正交向量组 $(\gamma_1, \gamma_2, \dots, \gamma_s)$,此即正交化过程.

(i) 取 $\gamma_1 = \alpha_1$;

(ii) 取 $\gamma_2 = \alpha_2 - \frac{(\alpha_2, \gamma_1)}{(\gamma_1, \gamma_1)} \gamma_1$.

则 γ_2 是 α_2 与 α_1 的线性组合,且

$$(\gamma_2, \gamma_1) = (\alpha_2, \gamma_1) - \frac{(\alpha_2, \gamma_1)}{(\gamma_1, \gamma_1)} (\gamma_1, \gamma_1) = 0,$$

即 γ_2 与 γ_1 正交.

(iii) 取 $\gamma_3 = \alpha_3 - \frac{(\alpha_3, \gamma_2)}{(\gamma_2, \gamma_2)} \gamma_2 - \frac{(\alpha_3, \gamma_1)}{(\gamma_1, \gamma_1)} \gamma_1$.

则 γ_3 是 $\alpha_3, \alpha_2, \alpha_1$ 的线性组合,且

$$(\gamma_3, \gamma_2) = (\alpha_3, \gamma_2) - \frac{(\alpha_3, \gamma_2)}{(\gamma_2, \gamma_2)} (\gamma_2, \gamma_2) - \frac{(\alpha_3, \gamma_1)}{(\gamma_1, \gamma_1)} (\gamma_2, \gamma_1) = 0,$$

$$(\gamma_3, \gamma_1) = (\alpha_3, \gamma_1) - \frac{(\alpha_3, \gamma_2)}{(\gamma_2, \gamma_2)} (\gamma_2, \gamma_1) - \frac{(\alpha_3, \gamma_1)}{(\gamma_1, \gamma_1)} (\gamma_1, \gamma_1) = 0,$$

即 γ_3 与 γ_2, γ_1 正交.

依次类推,直到

(iv) 取 $\gamma_s = \alpha_s - \frac{(\alpha_s, \gamma_{s-1})}{(\gamma_{s-1}, \gamma_{s-1})} \gamma_{s-1} - \dots - \frac{(\alpha_s, \gamma_1)}{(\gamma_1, \gamma_1)} \gamma_1$.

则 γ_s 是 $\alpha_s, \alpha_{s-1}, \dots, \alpha_1$ 的线性组合,且与 $\gamma_{s-1}, \gamma_{s-2}, \dots, \gamma_1$ 正交.

于是,得到了两两正交的向量组 $\gamma_1, \gamma_2, \dots, \gamma_s$,从而完成了正交化方法的第一步.

第二步 由正交向量组 $(\gamma_1, \gamma_2, \dots, \gamma_s)$,产生长度为 1 的单位正交向量组,即法正交组.

由第一步所得到的正交向量组 $(\gamma_1, \gamma_2, \dots, \gamma_s)$,作单位化如下:令

$$\beta_1 = \frac{\gamma_1}{|\gamma_1|}, \beta_2 = \frac{\gamma_2}{|\gamma_2|}, \dots, \beta_s = \frac{\gamma_s}{|\gamma_s|},$$

这样,就得到 $(\alpha_1, \alpha_2, \dots, \alpha_s)$ 的线性组合所构成的法正交组 $(\beta_1, \beta_2, \dots, \beta_s)$.

例 15 按施密特正交化方法,将向量组

$$\alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

法正交化.

首先正交化.取

$$\gamma_1 = \alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\gamma_2 = \alpha_2 - \frac{(\alpha_2, \gamma_1)}{(\gamma_1, \gamma_1)} \gamma_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix},$$

$$\begin{aligned} \gamma_3 &= \alpha_3 - \frac{(\alpha_3, \gamma_2)}{(\gamma_2, \gamma_2)} \gamma_2 - \frac{(\alpha_3, \gamma_1)}{(\gamma_1, \gamma_1)} \gamma_1 \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ -1 \end{bmatrix}. \end{aligned}$$

再单位化

$$\beta_1 = \frac{\gamma_1}{|\gamma_1|} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{bmatrix}, \quad \beta_2 = \frac{\gamma_2}{|\gamma_2|} = \begin{bmatrix} \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ 0 \end{bmatrix}, \quad \beta_3 = \frac{\gamma_3}{|\gamma_3|} = \begin{bmatrix} \frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}.$$

读者可以知道,方阵的正交化方法可以由施密特正交化方法得到.

(3) 实二次型化简

定义 6.2.17 (实二次型) 含有 n 个实变量 $x_1, x_2, \dots, x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \dots + 2a_{1n}x_1x_n \\ &\quad + a_{22}x_2^2 + 2a_{23}x_2x_3 + 2a_{24}x_2x_4 + \dots + 2a_{2n}x_2x_n \\ &\quad + \dots + a_{nn}x_n^2, \end{aligned}$$

其中诸系数 a_{ij} 都是实数, 称为实二次型, 或实二次齐式.

若令 $a_{ij} = a_{ji}$, 则二次型 f 可改写为

$$\begin{aligned} f &= a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \dots + a_{1n}x_1x_n \\ &\quad + a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + \dots + a_{2n}x_2x_n + \\ &\quad \dots \dots \dots \\ &\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + a_{n3}x_nx_3 + \dots + a_{nn}x_n^2. \end{aligned} \quad (6.2.11)$$

为简化上式的写法, 我们利用矩阵的乘法, 将 (6.2.11) 式改写成

$$f = \mathbf{x}^T \mathbf{A} \mathbf{x}, \quad (6.2.12)$$

(6.2.12) 式是二次型的矩阵表示, 其中

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \quad \mathbf{x}^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

$$\mathbf{A} = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad (6.2.13)$$

$\mathbf{A}$ 为实对称方阵.

易见, 给定一个实二次型 f (由式 (6.2.11)), 惟一确定一个实对称方阵 $\mathbf{A}$ (由式 (6.2.13)); 反之亦然. 因此, 实二次型与实对称方阵是一一对应的, 我们称 $\mathbf{A}$ 为实二次型 f 的方阵, 也称 f 为实对称方阵的二次型, 称方阵 $\mathbf{A}$ 的秩 $r = \text{rank}(\mathbf{A})$ 为二次型 f 的秩.

例 16 写出实二次型

$$\begin{aligned} f &= 2x_1^2 - 2x_1x_2 - 2x_1x_3 + 2x_1x_4 + 2x_2^2 + 2x_2x_3 - 2x_2x_4 + \\ &\quad 2x_3^2 - 2x_3x_4 + 2x_4^2 \end{aligned}$$

的方阵 $\mathbf{A}$, 并给出 f 的方阵表示.

解 只要按式 (6.2.13) 写出

$$A = \begin{bmatrix} 2 & -1 & -1 & 1 \\ -1 & 2 & 1 & -1 \\ -1 & 1 & 2 & -1 \\ 1 & -1 & -1 & 2 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

则由(6.2.12)即得 $f = x^T A x$.

定义 6.2.18 (线性变换) 两组变量 $x_1, x_2, \dots, x_n$ 与 $y_1, y_2, \dots, y_n$ 之间的关系

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \dots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \dots + c_{2n}y_n, \\ \dots\dots\dots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \dots + c_{nn}y_n, \end{cases} \quad (6.2.14)$$

称为从 $y_1, y_2, \dots, y_n$ 到 $x_1, x_2, \dots, x_n$ 的线性变换.

线性变换用矩阵表示为

$$x = Cy,$$

其中

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad C = (c_{ij}) = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}, \quad (6.2.15)$$

方阵 C 称为从 y 到 x 的线性变换方阵. 如果 C 是非异方阵, 则称相应的线性变换为非异线性变换, 或称为满秩线性变换. 特别地, 若 C 是正交方阵, 则称相应的变换为正交变换.

我们可以借助于非异线性变换来化简实二次型, 亦即化为各变量的平方项之代数和(称为实二次型的标准型)而消去其混合项(交叉项).

在非异线性变换 $x = Cy$ 之下, 二次型(6.2.11)化为

$$x^T A x = (Cy)^T A (Cy) = y^T (C^T A C) y.$$

若

$$C^T A C = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{bmatrix},$$

则

$$f = y \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{bmatrix} y^T = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

成为标准型.

由此可见,借助于非异线性变换化实二次型为标准型,可归结为寻找一个非异方阵 C ,使得 $C^T A C$ 成为对角方阵.这对于一般方阵 A 总可以办到.我们用配方法来实现这个要求.但是,如果 A 是实对称方阵,还可以用正交变换法来实现对角化.

① 配方法

例 17 把实二次型

$$f(x, y, z) = x^2 + 2y^2 + 4z^2 + 2xy + 6yz + 2zx$$

化为标准型,并写出相应的变换方阵.

解 在 f 中将含有 x 的项集中在一起进行配方

$$\begin{aligned} f &= x^2 + 2(y+z)x + (y+z)^2 - (y+z)^2 + 2y^2 + 4z^2 + 6yz \\ &= (x+y+z)^2 + y^2 + 4yz + 3z^2. \end{aligned}$$

再集中 y 的项进行配方,得

$$f = (x+y+z)^2 + (y+2z)^2 - z^2.$$

于是,只要作如下线性变换

$$\begin{cases} u = x + y + z, \\ v = y + 2z, \\ w = z, \end{cases} \quad \text{或} \quad \begin{cases} x = u - v + w, \\ y = v - 2w, \\ z = w, \end{cases}$$

就可把 f 化为标准型 $f = u^2 + v^2 - w^2$,变换方阵为

$$C = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix},$$

二次型 $f(x, y, z)$ 与 $g(u, v, w)$ 的方阵分别为

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

而且方阵 A, B, C 必适合

$$C^T A C = B.$$

例 18 化二次型 $f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$ 为标准型,并写出变换方阵.

解 因为 f 中不含平方项,需先作线性变换

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \end{cases}$$

相应的线性变换方阵为

$$C_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

将 x_1, x_2, x_3 代入 f , 得到 $f = 2y_1^2 - 4y_1y_3 - 2y_2^2 + 8y_2y_3$. 再对 f 逐次配方, 得到

$$\begin{aligned} f &= (2y_1^2 - 4y_1y_3 + y_3^2) - 2y_2^2 + 8y_2y_3 - 2y_3^2 \\ &= 2(y_1 - y_3)^2 - 2(y_2 - 2y_3)^2 + 6y_3^2. \end{aligned}$$

令

$$\begin{cases} z_1 = y_1 - y_3, \\ z_2 = y_2 - 2y_3, \\ z_3 = y_3, \end{cases} \quad \text{或} \quad \begin{cases} y_1 = z_1 + z_3, \\ y_2 = z_2 + 2z_3, \\ y_3 = z_3, \end{cases}$$

相应的变换方阵为

$$C_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

所求变换应为

$$C = C_1 C_2 = \begin{bmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

于是, 可作如下线性变换

$$\begin{cases} x_1 = z_1 + z_2 + 3z_3, \\ x_2 = z_1 - z_2 - z_3, \\ x_3 = z_3, \end{cases}$$

就把 f 化为标准型 $g = 2z_1^2 - 2z_2^2 + 6z_3^2$, 且有

$$C^T A C = C^T \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{bmatrix} C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

② 正交变换法

如果 A 是实对称方阵, 还可以要求 C 为正交方阵, 这不仅带来许多方便, 而且在应用中还常常这样要求.

为此, 我们先给出三个定理.

定理 6.2.10 (实对称方阵的特征值) 实对称方阵 A 的特征值都是实数.

证 设 λ 是实对称方阵 A 的任一个特征值, 以 $\bar{\lambda}$ 表示 λ 的共轭复数. 又设 A 的属于 λ 的任一特征向量为

$$\alpha = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \bar{\alpha} = \begin{bmatrix} \bar{a}_1 \\ \bar{a}_2 \\ \vdots \\ \bar{a}_n \end{bmatrix}.$$

于是, 一方面

$$\bar{\alpha}^T (A\alpha) = \bar{\alpha}^T (\lambda\alpha) = \lambda \bar{\alpha}^T \alpha;$$

另一方面, 记 $\bar{A} = (\bar{a}_{ij})_{n \times n}$, 则 $A^T = A = \bar{A}$, 又有

$$\bar{\alpha}^T (A\alpha) = (\bar{\alpha}^T \bar{A}^T) \alpha = (\bar{A}\bar{\alpha})^T \alpha = (\lambda\bar{\alpha})^T \alpha = \lambda \bar{\alpha}^T \alpha.$$

两式相减, 得

$$(\lambda - \bar{\lambda}) \bar{\alpha}^T \alpha = 0.$$

由 $\alpha \neq 0$, 得到

$$\bar{\alpha}^T \alpha = a_1 \bar{a}_1 + a_2 \bar{a}_2 + \cdots + a_n \bar{a}_n = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 \neq 0,$$

从而只有 $\lambda - \bar{\lambda} = 0$, 即 $\lambda = \bar{\lambda}$, 所以 λ 为实数.

定理 6.2.11 (实对称方阵的特征向量) 属于实对称方阵 A 的不同特征值的特征向量彼此正交.

证 设实对称方阵 A 的不同特征值为 $\lambda_1 \neq \lambda_2$. $\alpha_1 \neq 0$ 、 $\alpha_2 \neq 0$ 是分别属于 λ_1 、 λ_2 的特征向量, 即

$$A\alpha_1 = \lambda_1 \alpha_1, \quad A\alpha_2 = \lambda_2 \alpha_2,$$

于是,

$$\begin{aligned} \lambda_1 (\alpha_1, \alpha_2) &= (\lambda_1 \alpha_1, \alpha_2) = (A\alpha_1, \alpha_2) = (A\alpha_1)^T \alpha_2 \\ &= (\alpha_1)^T A^T \alpha_2 = (\alpha_1)^T A\alpha_2; \end{aligned}$$

$$\lambda_2 (\alpha_1, \alpha_2) = (\alpha_1, \lambda_2 \alpha_2) = (\alpha_1, A\alpha_2) = (\alpha_1)^T A\alpha_2;$$

两式相减得到

$$(\lambda_1 - \lambda_2) (\alpha_1, \alpha_2) = 0.$$

但是, $\lambda_1 \neq \lambda_2$, 故必有 $(\alpha_1, \alpha_2) = 0$, 所以 α_1 与 α_2 正交. 定理得证.

定理 6.2.12 (实对称方阵的正交化) 设 A 为 n 阶实对称方阵, 则必存在正交方阵 C , 使得

$$C^{-1}AC = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \vdots & \\ & & & \lambda_n \end{bmatrix}, \quad (6.2.16)$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的全体特征值 (k 重算作 k 个).

证 对 A 的阶数用数学归纳法.

$n=1$ 时, 1 阶方阵 $A=[a]$ 本身就是对角方阵, 取 C 为 1 阶方阵 $C=[1]$, 必有 $C^{-1}AC=[\lambda]$. 结论显然成立.

假设对于所有 $n-1$ 阶实对称方阵定理结论成立, 考虑 n 阶实对称方阵 A , 并且任取 A 的一个特征值 λ_1 , 以及属于 λ_1 的特征向量 $\alpha_1, |\alpha_1|=1$. 另取 $n-1$ 个向量, 使与 α_1 一起构成一个线性无关的向量组, 再按施密特正交化方法将它们正交化、单位化, 从而得到一组法正交组 $\alpha_1, \alpha_2, \dots, \alpha_n$, 并且获得一个正交方阵 $C_1=(\alpha_1, \alpha_2, \dots, \alpha_n)$. 于是,

$$AC_1=(A\alpha_1, A\alpha_2, \dots, A\alpha_n)=(\lambda_1\alpha_1, A\alpha_2, \dots, A\alpha_n).$$

由此,

$$C_1^{-1}AC=(C_1^{-1}\lambda_1\alpha_1, C_1^{-1}A\alpha_2, \dots, C_1^{-1}A\alpha_n).$$

因为 C_1 是正交方阵, 所以 $C_1^{-1}=C_1^T$. 又因 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是法正交组, 故上式右边方阵的第 1 列为

$$C_1^{-1}\lambda_1\alpha_1=\lambda_1C_1^T\alpha_1=\lambda_1\begin{bmatrix}\alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T\end{bmatrix}\alpha_1=\lambda_1\begin{bmatrix}\alpha_1^T\alpha_1 \\ \alpha_2^T\alpha_1 \\ \vdots \\ \alpha_n^T\alpha_1\end{bmatrix}=\lambda_1\begin{bmatrix}(\alpha_1, \alpha_1) \\ (\alpha_2, \alpha_1) \\ \vdots \\ (\alpha_n, \alpha_1)\end{bmatrix}=\lambda_1\begin{bmatrix}1 \\ 0 \\ \vdots \\ 0\end{bmatrix}.$$

这样, 我们可以假设

$$C_1^{-1}AC=\begin{bmatrix}\lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}.$$

利用 A 的对称性 $A=A^T$, 得

$$(C_1^{-1}AC)^T=(C_1^TAC)^T=C_1^TA^TC_1=C_1^{-1}AC_1,$$

这说明 $C_1^{-1}AC_1$ 也是对称方阵, 故

$$C_1^{-1}AC=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{bmatrix}.$$

其中 B 是 $n-1$ 阶实对称方阵, 按归纳法假设, 存在 $n-1$ 阶正交方阵 C_2 , 使得

$$C_2^{-1}BC_2 = \begin{bmatrix} \lambda_2 & & \\ & \lambda_3 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix}.$$

令

$$C = C_1 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix},$$

由分块方阵的逆方阵知

$$\begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix}^{-1} = \begin{bmatrix} 1^{-1} & 0 \\ 0 & C_2^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & C_2^T \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix}^T.$$

因此, $\begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix}$ 是正交方阵. 但是, 两个正交方阵的乘积也是正交方阵, 并且

$$\begin{aligned} C^{-1}AC &= \left(C_1 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} \right)^{-1} A \left(C_1 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} \right) \\ &= \left(\begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} \right)^{-1} C^{-1}AC_1 \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & C_2^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & C_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & C_2^{-1}BC_2 \end{bmatrix} = \begin{bmatrix} \lambda_2 & & \\ & \lambda_3 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix}. \end{aligned} \quad (6.2.17)$$

由数学归纳法, 定理的前一部分得证.

为证明定理的后一部分, 注意到 C 是正交方阵, 满足 $C^{-1} = C^T$, 故 (6.2.17) 式表示 A 相似于对角方阵, 因此对角方阵的主对角线上的元素都是 A 的特征值, 并且方阵 C 的第 i 个列向量是属于 λ_i ($i = 1, 2, \dots, n$) 的特征向量. 定理的后一部分得证.

上面的定理告诉我们, 任何实二次型 f 都可以通过正交变换化为标准型. 我们将这个过程归结为以下步骤:

- ① 写出实二次型 f 所对应的实对称方阵 A ;
- ② 求 A 的全部特征值, 亦即求 $|\lambda I - A| = 0$ 的所有不同的特征根 $\lambda_1, \lambda_2, \dots, \lambda_s$;
- ③ 对每个特征值 λ_j , 求齐次线性方程组 $(\lambda_j I - A)x = 0$ 的一个基础解系

$$\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jl_j},$$

并且按施密特正交化、单位化方法,求得一个法正交组

$$\beta_{j1}, \beta_{j2}, \dots, \beta_{jl_j} \quad (j=1, 2, \dots, s).$$

④ 以所得的 s 个法正交组的全部向量

$$\beta_{11}, \dots, \beta_{1l_1}, \beta_{21}, \dots, \beta_{2l_2}, \dots, \beta_{s1}, \dots, \beta_{sl_s}$$

为列向量,组成方阵 C ,则 C 为所求的正交方阵,而 $C^{-1}AC$ 为对角方阵,其主对角线上的元素 λ 是 C 的第 j 个列向量所属的特征值.

例 19 将例 16 中的实二次型 f 化为标准型,并写出其正交变换.

解 我们按上面的 4 个步骤进行.

① 实二次型 f 所对应的实对称方阵 A 为

$$A = \begin{bmatrix} 2 & -1 & -1 & 1 \\ -1 & 2 & 1 & -1 \\ -1 & 1 & 2 & -1 \\ 1 & -1 & -1 & 2 \end{bmatrix}.$$

② 求 A 的全部特征值,亦即求 $|\lambda I - A| = 0$ 的所有不同的特征根

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda - 2 & 1 & 1 & -1 \\ 1 & \lambda - 2 & -1 & 1 \\ 1 & -1 & \lambda - 2 & 1 \\ -1 & 1 & 1 & \lambda - 2 \end{vmatrix} = \begin{vmatrix} \lambda - 1 & \lambda - 1 & \lambda - 1 & \lambda - 1 \\ 1 & \lambda - 2 & -1 & 1 \\ 1 & -1 & \lambda - 2 & 1 \\ -1 & 1 & 1 & \lambda - 2 \end{vmatrix} \\ &= (\lambda - 1) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & \lambda - 3 & -2 & 0 \\ 1 & -2 & \lambda - 3 & 0 \\ -1 & 2 & 2 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^3 (\lambda - 5) \end{aligned}$$

故 A 有一个三重特征值 $\lambda = 1$ 与一个单重特征值 $\lambda = 5$.

③ 对每个特征值 λ_j ,求齐次线性方程组 $(\lambda_j I - A)x = 0$ 的一个基础解系

对于 $\lambda = 1$,由齐次线性方程组 $(1I - A)x = (I - A)x = 0$,求得一个基础解系

$$\alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

在例 15 中我们已经得到这 3 个向量的正交组:

$$\beta_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{bmatrix}, \quad \beta_2 = \begin{bmatrix} \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{3} \\ 0 \end{bmatrix}, \quad \beta_3 = \begin{bmatrix} \frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{6} \\ -\frac{\sqrt{3}}{2} \end{bmatrix}.$$

对于 $\lambda = 5$, 由齐次线性方程组 $(5I - A)x = 0$, 求得一个基础解系

$$\alpha_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

由于此基础解系中只含有一个向量, 故只要单位化就够了, 即得

$$\beta_4 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

④ 以所得的 s 个法正交组的全部向量为列向量, 组成正交方阵 C

$$C = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

而且

$$C^{-1}AC = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

最后, 我们得到正交变换 $x = Cy$ 为

$$\begin{cases} x_1 = \frac{\sqrt{2}}{2}y_1 + \frac{\sqrt{6}}{6}y_2 + \frac{\sqrt{3}}{6}y_3 + \frac{1}{2}y_4 \\ x_2 = \frac{\sqrt{2}}{2}y_1 - \frac{\sqrt{6}}{6}y_2 - \frac{\sqrt{3}}{6}y_3 - \frac{1}{2}y_4, \\ x_3 = \frac{\sqrt{6}}{3}y_2 - \frac{\sqrt{3}}{6}y_3 - \frac{1}{2}y_4, \\ x_4 = -\frac{\sqrt{3}}{2}y_3 + \frac{1}{2}y_4, \end{cases}$$

所得的二次型 f 的标准型为 $y_1^2 + y_2^2 + y_3^2 + 5y_4^2$.

(4) 正定实二次型

在第二章多元函数应用部分,学习关于极值判别时,利用二阶行列式恒大于0、小于0来判定极值的存在,其实这涉及二次型的“正定”与“负定”概念.

定义 6.2.19 (正定实二次型、正定方阵) 如果对于任何非零向量 x , 实二次型 $f = x^T A x$ 都大于0, 则称 f 为正定二次型, 相应的方阵 A 称为正定方阵;

如果对于任何非零向量 x , $f = x^T A x$ 都小于0, 则称 f 为负定二次型, 相应的方阵 A 也称为负定方阵;

若既有 $x \neq 0$ 使得 f 大于0, 也有 $x \neq 0$ 使得 f 小于0, 便称 f 为不定二次型, A 则称为不定方阵.

现在给出实二次型正定、负定判别定理. 显然, f 为负定二次型, 当且仅当 $-f$ 为正定的, 因此, 我们仅讨论正定二次型, 也只要讨论 n 阶实对称方阵 A 的正定性.

定理 6.2.13 (实对称方阵正定的充要条件) n 阶实对称方阵 A 为正定的, 当且仅当 A 的所有特征值都大于0.

证 必要性 设 λ 是实对称方阵的任一特征值, 取属于 λ 的非零特征向量 α , 则 $\alpha^T A \alpha > 0$ (因为 A 对应的实二次型 f 对任意非零向量 x 都成立 $f = x^T A x > 0$, 故对非零特征向量 α 也成立). 于是, 由

$$\lambda |\alpha|^2 = \lambda (\alpha, \alpha) = \lambda \alpha^T \alpha = \alpha^T \lambda \alpha = \alpha^T A \alpha > 0$$

得到 $\lambda > 0$.

充分性 设 A 的所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都是正数. 取 A 的一组法正交特征向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 则由定理 6.2.8 知, 它们是线性无关的. 于是任何非零向量 x 都是 $\alpha_1, \alpha_2, \dots, \alpha_n$ 的线性组合

$$x = k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n = 0,$$

其中 $k_1, k_2, \dots, k_n$ 不全为0, 于是

$$\begin{aligned} x^T A x &= (k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n)^T A (k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n) \\ &= (k_1 \alpha_1^T + k_2 \alpha_2^T + \dots + k_n \alpha_n^T) (k_1 A \alpha_1 + k_2 A \alpha_2 + \dots + k_n A \alpha_n) \end{aligned}$$

$$= k_1^2 \lambda_1 (\alpha_1, \alpha_1) + k_2^2 \lambda_2 (\alpha_2, \alpha_2) + \cdots + k_n^2 \lambda_n (\alpha_n, \alpha_n).$$

因为 $k_1, k_2, \dots, k_n$ 不全为 0, 而且 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都是正值, 故 $f = x^T A x > 0$. 最后, 由 $x \neq 0$ 是任意向量, 知 A 为正定方阵. 定理得证.

定理 6.2.14 (实对称方阵正定的充要条件) 实对称方阵

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

为正定的充要条件是: 对 A 的所有顺序主子式都大于 0, 亦即

$$a_{11} > 0; \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0; \cdots, \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} > 0; \quad (6.2.18)$$

A 为负定的充要条件是:

$$a_{11} < 0; \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0; \cdots, \quad (-1)^n \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} > 0. \quad (6.2.19)$$

证明从略.

例 20 判定下列二次型是否是正定的.

(i) $f(x, y, z) = 4x^2 - 2yz + 3y^2 + 3z^2$;

(ii) $g(x, y, z) = x^2 + 2y^2 - 3z^2 + 4xy + 2yz$.

解 (i) 相应于二次型 $f(x, y, z)$ 的实对称方阵为

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}.$$

各阶顺序主子式为

$$4 > 0; \quad \begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} = 12 > 0; \quad \begin{vmatrix} 4 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{vmatrix} > 0.$$

故 $f(x, y, z)$ 是正定二次型.

(ii) 二次型 $g(x, y, z)$ 的方阵为

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & -3 \end{bmatrix}.$$

各阶顺序主子式为

$$1 > 0; \quad \begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} = -2 < 0; \quad \begin{vmatrix} 1 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & -3 \end{vmatrix} = 5 > 0.$$

但是, g 既非正定, 也非负定的.

3. 方阵标准型的应用

(1) 坐标轴旋转变换

设三维空间 $\mathbf{R}^3$ 中有直角坐标系 $Oxyz$. 仍以 O 为原点, 构造一个新的直角坐标系 $Ox_1y_1z_1$, 原坐标系与新坐标系的基本向量分别为 i, j, k 与 i_1, j_1, k_1 . 它们之间的夹角由下表给出.

	i_1	j_1	k_1
i	α_1	α_2	α_3
j	β_1	β_2	β_3
k	γ_1	γ_2	γ_3

其中基本向量组 i, j, k 与 i_1, j_1, k_1 均为法正交组. 写成方程组形式

$$\begin{cases} i = i_1 \cos \alpha_1 + j_1 \cos \alpha_2 + k_1 \cos \alpha_3 \\ j = i_1 \cos \beta_1 + j_1 \cos \beta_2 + k_1 \cos \beta_3 \\ k = i_1 \cos \gamma_1 + j_1 \cos \gamma_2 + k_1 \cos \gamma_3 \end{cases} \quad (6.2.20)$$

若记

$$\delta = \begin{bmatrix} i \\ j \\ k \end{bmatrix}, \quad \delta_1 = \begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix}, \quad C = \begin{bmatrix} \cos \alpha_1 & \cos \alpha_2 & \cos \alpha_3 \\ \cos \beta_1 & \cos \beta_2 & \cos \beta_3 \\ \cos \gamma_1 & \cos \gamma_2 & \cos \gamma_3 \end{bmatrix}, \quad (6.2.21)$$

则(6.2.20)可简写为

$$\delta = C\delta_1, \quad (6.2.22)$$

基本向量组 i, j, k 与 i_1, j_1, k_1 均为法正交组, 易知 $CC^T = I$, 即 C 为正交方阵, 于是从 i_1, j_1, k_1 到 i, j, k 的线性变换(6.2.20)是正交变换.

反之, 设(6.2.21)中的 C 为 3 阶正交方阵, $\delta_1 = (i_1, j_1, k_1)$ 为某直角坐标系的三个基本向量, 则由(6.2.20)得到的三个分量 i, j, k 也是一个法正交组. 可以证明(从略), 当 $|C| = 1$ 时, i, j, k 仍然构成一个右手坐标系, 而当 $|C| =$

-1时, i, j, k 则构成左手坐标系.

以上是坐标轴之间的旋转变换. 在坐标轴的旋转变换之下, 空间任一点 P 在原坐标系与新坐标系中的坐标分别为 $(x, y, z), (x_1, y_1, z_1)$. 于是

$$\begin{aligned}\overrightarrow{OP} &= (x, y, z) = xi + yj + zk, \\ \overrightarrow{O_1P} &= (x_1, y_1, z_1) = x_1i_1 + y_1j_1 + z_1k_1,\end{aligned}$$

从而, 由内积公式得

$$(x_1, y_1, z_1)\delta_1 = (x, y, z)\delta = (x, y, z)C\delta_1.$$

因为 $\delta_1\delta_1^T = I$ 为 3 阶单位方阵, 在上式两边右乘 δ_1^T , 得到

$$(x_1, y_1, z_1) = (x, y, z)C.$$

或

$$(x_1, y_1, z_1) = (x, y, z) \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}. \quad (6.2.23)$$

记 $\xi = (x, y, z)^T, \xi_1 = (x_1, y_1, z_1)^T$, 则上式可写为 $\xi_1^T = \xi^T C$, 或 $\xi_1 = C^T \xi$.

又因 $C^{-1} = C^T$, 最后得到 $\xi = C\xi_1$, 亦即

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \quad (6.2.24)$$

(6.2.23)与(6.2.24)是坐标旋转后, 点 P 的原坐标与新坐标互相转换的关系式, 它通过一个 3 阶正交方阵 C 得以实现.

平面坐标的旋转变换是通过 2 阶正交方阵来完成的.

(2) 化二次曲面方程为标准型

三个变量 x, y, z 的二次方程一般形式为

$$\begin{aligned}a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz \\ + b_1x + b_2y + b_3z + c = 0.\end{aligned} \quad (6.2.25)$$

(一) 首先讨论, 如果经过坐标平移可以将(6.2.25)的一次项消去, 再将常数项移到方程右边, 使方程化为

$$f(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz = c. \quad (6.2.26)$$

这里 f 是含 3 个变量的二次型. 当 a_{ij} 与 c 为实数时, f 就是实二次型. 设 $A = (a_{ij})_{3 \times 3}$ 为实二次型 f 的实对称方阵, 其特征值为 $\lambda_1, \lambda_2, \lambda_3$, 由前面的讨论知, 存在正交方阵 C , 使得

$$C^T A C = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \quad (6.2.27)$$

不失一般性,可以设 $|C|=1$, 因若不然,则必有 $|C|=-1$, 我们用 $-C$ 取代 C , (6.2.27) 式仍然成立.

于是,在 $|C|=1$ 的假设下,令 $\xi = (x, y, z)^T$, $\xi_1 = (x_1, y_1, z_1)^T$, 且

$$\xi = C\xi_1,$$

则二次型 f 化为

$$f = \xi^T A \xi = (C\xi_1)^T A (C\xi_1) = \xi_1^T C^T A C \xi_1 = \lambda_1 x_1^2 + \lambda_2 y_1^2 + \lambda_3 z_1^2.$$

方程 $f=c$ 成为

$$\lambda_1 x_1^2 + \lambda_2 y_1^2 + \lambda_3 z_1^2 = c. \quad (6.2.28)$$

我们知道,正交变换对应于一个正交方阵,此方阵的行列式为 1, 这种变换在空间解析几何中表示坐标旋转,而使得一个不含一次项的二次方程总可化为其标准型的形式(6.2.28). 在第二章中已经看到,二次曲面可以化为柱面、锥面、椭圆面、椭圆抛物面、双曲抛物面、抛物柱面、单叶双曲面、双叶双曲面等标准型.

现在我们来讨论方程(6.2.28)的图形. 为简单起见,将其写为

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 = c \quad (c \text{ 为 } 0 \text{ 或 } 1). \quad (6.2.29)$$

① 当 $c=0$ 时

(i) 若 $\lambda_1, \lambda_2, \lambda_3$ 中有 1 个为 0, 例如,可设 $\lambda_3=0$, 而 λ_2, λ_3 同号时, (6.2.29) 成为一个点; 两者异号时, (6.2.29) 是退化锥面(即直线).

(ii) 若 $\lambda_1, \lambda_2, \lambda_3$ 中有 2 个为 0, 例如,可设 $\lambda_2=0, \lambda_3=0$, 而 $\lambda_1 \neq 0$, (6.2.29) 成为平面 $x=0$.

(iii) 若 $\lambda_1, \lambda_2, \lambda_3$ 全不为 0, 且同号时, 则(6.2.29)是一个点.

(iv) 若 $\lambda_1, \lambda_2, \lambda_3$ 全不为 0, 且异号时, 例如,可设 $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$, 则(6.2.29)成为

$$z^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

其图形是二次锥面.

② 当 $c=1$ 时

(i) 若 $\lambda_1, \lambda_2, \lambda_3$ 中有一个为 0, 例如,可设 $\lambda_3=0$:

(a) 若 $\lambda_1 > 0, \lambda_2 > 0$, 则(6.2.29)成为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

其图形是椭圆柱面.

(b) 若 $\lambda_1 > 0, \lambda_2 < 0$, 则(6.2.29)成为

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

其图形是双曲柱面.

(ii) 若 $\lambda_1, \lambda_2, \lambda_3$ 中有 2 个为 0, 例如, 可设 $\lambda_2 = 0, \lambda_3 = 0$, 而 $\lambda_1 \neq 0$, (6.2.29)成为平面 $x = 1$.

(iii) 若 $\lambda_1, \lambda_2, \lambda_3$ 全不为 0,

(a) 若 $\lambda_1, \lambda_2, \lambda_3 > 0$ 时, 即它们全正时, 则(6.2.29)成为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

其图形是椭圆面.

(b) 若 $\lambda_1, \lambda_2, \lambda_3 < 0$ 时, 即它们全负时, 则(6.2.29)成为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1,$$

其图形没有实轨迹.

(c) 若 $\lambda_1, \lambda_2, \lambda_3$ 异号, 且两正一负时, 例如

$\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$, 则(6.2.29)成为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

其图形是单叶双曲面.

(d) 若 $\lambda_1, \lambda_2, \lambda_3$ 异号, 且一正两负时, 例如

$\lambda_1 > 0, \lambda_2 < 0, \lambda_3 < 0$, 则(6.2.29)成为

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1,$$

其图形是双叶双曲面.

(二) 其次讨论, 如果二次方程(6.2.25)无论怎样作坐标平移, 都不能消去一次项, 则它们必可经过适当的坐标平移与旋转, 化为下述三种形式之一:

$$\textcircled{1} y^2 = 2px,$$

其图形是抛物柱面.

$$\textcircled{2} z = \frac{x^2}{a^2} + \frac{y^2}{b^2},$$

其图形是椭圆抛物面.

$$\textcircled{3} z = \frac{x^2}{a^2} - \frac{y^2}{b^2},$$

其图形是双曲抛物面.

以上图形已经在第一章中给出.

(3) 用二次型判别极值

这在第二章多元函数极值时介绍过,这里就不重复了.

(4) 惯性指数

我们再引进一个重要概念,就是惯性指数.因为实二次型化为标准型的非异线性变换并不是惟一的,因此所得到的标准型中平方项的系数并不是惟一的,这样,“惯性指数”概念就显得十分重要了.

定义 6.2.20 (正、负惯性指数) 二次型标准型中的平方项的正系数的个数,称为二次型的正惯性指数;负系数的个数,称为二次型的负惯性指数.

但是我们有如下定理.

定理 6.2.15 (惯性定理) 对于给定的实二次型,其任一标准型中的正、负惯性指数是确定不变的.

这个定理的证明较为复杂,我们从略.

习题

1. 若 $B_1 = U^{-1} A_1 U, B_2 = U^{-1} A_2 U$, 求证 $B_1 B_2 = U^{-1} (A_1 A_2) U, B_1 + B_2 = U^{-1} (A_1 + A_2) U$. 是否可以说:若 $A_1 \sim B_1, A_2 \sim B_2$, 则 $A_1 A_2 \sim B_1 B_2, A_1 + A_2 \sim B_1 + B_2$?

2. 求证:若 $A \sim B$, 求证 $A' \sim B'$, 若 A 还是可逆的, 则 $A^{-1} \sim B^{-1}$.

3. 若 $A_1 \sim B_1, A_2 \sim B_2$, 试证

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \sim \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}.$$

4. 求下列矩阵的特征值和全部特征向量:

$$(1) A = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{bmatrix}; \quad (2) A = \begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{bmatrix}.$$

5. 证明:若 $\alpha_1, \alpha_2, \dots, \alpha_s$ 是 A 的属于特征值 λ 的线性无关的特征向量, 则对任意一组不全为零的数 $k_1, k_2, \dots, k_s, k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_s \alpha_s$ 也是 A 属于 λ 的特征向量.

6. 如果存在正整数 k 使 $A^k = 0$, 则称 A 为幂零矩阵. 试证幂零矩阵的特征值是零.

7. 设 n 阶矩阵 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 又 A 的特征多项式 $|\lambda I - A| = \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$, 试证 $a_1 = -(\lambda_1 + \lambda_2 + \dots + \lambda_n), a_n = (-1)^n \lambda_1 \lambda_2 \dots \lambda_n$.

8. 设

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}.$$

其中 $a_{11}, a_{22}, a_{33}, a_{44}$ 各不相同, 试证 A 可对角化.

9. 求 A^n (n 为正整数), 设

$$(1) A = \begin{bmatrix} 1 & 4 & 2 \\ 0 & -3 & 4 \\ 0 & 4 & 3 \end{bmatrix}; \quad (2) A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

10. 将下列矩阵对角化:

$$(1) A = \begin{bmatrix} 3 & 6 & 6 \\ 0 & 2 & 0 \\ -3 & -12 & -6 \end{bmatrix}; \quad (2) A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}.$$

$$11. \text{ 设 } A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & a & 2 \\ 0 & 2 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & b \end{bmatrix},$$

(1) 试确定 a, b 使 $A \sim B$;

(2) 求一个可逆矩阵 U , 使 $U^{-1}AU = B$.

12. 若 α 与 $\beta_1, \beta_2, \dots, \beta_n$ 的每一个都正交, 求证 α 与 $\beta_1, \beta_2, \dots, \beta_n$ 的任一线性组合也正交.

13. 若 A, B 是同阶正交方阵, 试证 AB 也是正交方阵.

14. 若 A 是正交方阵, 求证 A^{-1} 也是正交方阵.

15. 求证: 若 A 是正交方阵, 则 $|A|^2 = 1$, 且当 $|A| = -1$ 时, -1 是 A 的一个特征值; 当 $|A| = 1$ 且 A 为奇数阶时 1 是 A 的一个特征值.

16. 若 A 为对称方阵且满足关系 $A^2 + 6A + 8I = 0$, 证明 $A + 3I$ 是正交方阵.

17. 试按施密特正交化方法从已知向量组出发构造法正交向量组(向量的个数与原组相同):

$$(1) \alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \alpha_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix};$$

$$(2) \alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \alpha_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \alpha_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

18. 求正交矩阵 C , 使 $C^{-1}AC$ 为对角矩阵:

$$(1) A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & -2 & 2 \\ 4 & 2 & 1 \end{bmatrix}; (2) A = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{bmatrix}.$$

19. 用配方法化下列二次型为标准型:

$$(1) f = x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3;$$

$$(2) f = x_1x_2 + x_2x_3 + x_3x_1.$$

20. 判别下列二次型是否为正定的或负定的:

$$(1) 3x^2 + 4y^2 + 5z^2 + 4xy - 4yz; \quad (2) -5x^2 - 6y^2 - 4z^2 + 4xy + 4yz.$$

21. 问 t 满足什么条件时, 二次型 $x^2 + 4y^2 + 2z^2 + 2txy + 2xz$ 为正定的?

22. 求一正交变换,化下列二次曲面为标准形:

$$(1) 4x^2 + 3y^2 + 3z^2 + 2yz = 1; \quad (2) 2x^2 + 2y^2 - 2xy - 2xz - 2yz = 1.$$

§ 6.3 拓扑结构

刻画集合中的元素之间的位置关系、远近程度,需要赋予集合以拓扑结构,使集合成为“拓扑空间”.拓扑空间在数学科学中占有重要地位,是近代数学中不可缺少的概念和工具.

本节通过几个常用的例子,介绍引进集合的拓扑结构的思路,几种拓扑空间的定义以及这些拓扑空间的结构.

三维欧氏空间 $\mathbf{R}^3$ 是人类生存的空间,当然也是人们最熟悉的空间.在第一章介绍数与形的结合时,已经将任一实数 x 与直线上的点建立了一一对应,称该直线为数轴,亦即一维欧氏空间 $\mathbf{R}$.平面是二维欧氏空间 $\mathbf{R}^2$.现在我们以欧氏空间作为模型来推广空间概念.

首先注意到, $\mathbf{R}$ 、 $\mathbf{R}^2$ 、 $\mathbf{R}^3$ 在向量加法“+”与数乘“ $\alpha \cdot$ ”之下是线性空间;其次,对于 $\mathbf{R}$ 中的任意两点 $x, y \in \mathbf{R}$,定义 x, y 的距离

$$\rho(x, y) = |x - y|.$$

类似地,在 $\mathbf{R}^n$ ($n=2,3$) 中任意两点 $x, y \in \mathbf{R}^n$ 间的距离 $\rho(x, y)$ 可表示为

$$\rho(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

可以举出许多的例子,它们不是欧氏空间,但却可以引进“距离”而刻画元素间的远近.利用 $\rho(x, y)$ 满足三个特征性质:非负性、对称性、三点不等式,把距离推广到一般集合,从而定义一般的距离空间.

6.3.1 距离空间

定义 6.3.1 (距离空间) 设 X 为一集合,其元素为 $x, y, \cdots \in X$,如果在 X 中给定其元素之间的“距离” $\rho(x, y)$,使其具有以下性质:

(i) $\rho(x, y) \geq 0, \rho(x, y) = 0$ 当且仅当 $x = y$, (非负性)

(ii) $\rho(x, y) = \rho(y, x)$, (对称性)

(iii) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$, (三点不等式)

则称 X 为一个距离空间, $\rho(x, y)$ 称为 x, y 之间的距离.此时我们说,集合 X 被赋予由距离 $\rho(x, y)$ 决定的拓扑结构.

显然, $\mathbf{R}^n$ 是以 $\rho(x, y) = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}$ 为距离的距离空间.

看几个重要的例子.

(1) l^2 空间

l^2 空间是欧氏空间的自然推广:

$$l^2 = \left\{ x = (x_1, x_2, \dots) : x_j \text{ 为实数, 且 } \sum_{j=1}^{\infty} x_j^2 < \infty \right\}.$$

l^2 在下述“加法”、“数乘”之下, 成为一个线性空间

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

$$\alpha \cdot x = (\alpha x_1, \alpha x_2, \dots).$$

进而, 如果赋予 l^2 如下“距离”, 则它也是一个距离空间

$$\rho(x, y) = \sqrt{\sum_{j=1}^{\infty} (x_j - y_j)^2},$$

其中 $x = (x_1, x_2, \dots), y = (y_1, y_2, \dots) \in l^2$.

(2) 连续函数空间 $C[a, b]$

$[a, b]$ 上实值连续函数的全体所成的集合:

$$C[a, b] = \{ f : f(x) \text{ 是定义在 } [a, b] \text{ 上的连续函数} \}.$$

对于 $f, g, h, \dots \in C[a, b]$, 我们定义

$$(f + g)(x) = f(x) + g(x),$$

$$(\alpha \cdot f)(x) = \alpha \cdot f(x),$$

不难验证, $C[a, b]$ 成为一个线性空间. 如果再定义如下距离

$$\rho(f, g) = \max \{ |f(x) - g(x)| : x \in [a, b] \},$$

则它也是一个距离空间.

(3) 平方可积函数空间 $L^2(E)$

如第三章 § 3.8 中, 称集合 E 上 L -可积函数的全体为 L -可积函数类, 记为

$$L^1(E) = \left\{ f : \int_E |f(x)| dm < +\infty \right\}.$$

而称 E 上函数 f 是 L -平方可积的, 如果该函数的平方是 L -可积的, 并称 L -平方可积函数的全体为 L -平方可积函数类, 记为

$$L^2(E) = \left\{ f : \int_E |f(x)|^2 dm < +\infty \right\}.$$

在函数类 $L(E)$ 和 $L^2(E)$ 中, 对于 $f, g, h, \dots \in L^2(E)$, 定义“加法”、“数乘”:

$$(f + g)(x) = f(x) + g(x),$$

$$(\alpha \cdot f)(x) = \alpha \cdot f(x),$$

则 $L(E)$ 和 $L^2(E)$ 成为线性空间; 若再分别定义“距离”:

$$\rho(f, g) = \int_E |f(x) - g(x)| dm,$$

$$\rho(f, g) = \left\{ \int_E |f(x) - g(x)|^2 dm \right\}^{\frac{1}{2}},$$

则它们也都成为距离空间.

6.3.2 赋范线性空间

在上节读者已经看到,距离空间中的元素之间不必有代数运算.但是如果一个距离空间也是一个线性空间,而且要求加法“+”与数乘“ $\alpha \cdot$ ”运算在距离拓扑下是连续的,则空间的结构就可表示为一种称为赋范线性空间的结构.赋范线性空间在现代科学技术各领域中都有重要的应用.

定义 6.3.2 (赋范线性空间) 若在线性空间 X 中,定义对应与每个元素 $x \in X$ 的非负实数 $\|x\|$,具有如下性质

(i) $\|x\| \geq 0$, $\|x\| = 0$ 当且仅当 $x = 0$, (非负性)

(ii) $\|\alpha x\| = |\alpha| \|x\|$, (绝对齐性)

(iii) $\|x + y\| \leq \|x\| + \|y\|$, (三角不等式)

则称 X 为一个赋范线性空间, $\|x\|$ 称为 x 的范数.此时我们说,线性空间 X 被赋予由范数 $\|x\|$ 决定的拓扑结构.

如果赋范线性空间是完备的(“完备性”概念较为复杂,这里不作介绍了),就称其为巴拿赫(Banach)空间.

$\mathbb{R}^n$ 、 l^2 、 $C[a, b]$ 、 $L^1(E)$ 、 $L^2(E)$ 都是巴拿赫空间,其范数分别为

$$x \in \mathbb{R}^n: \quad \|x\| = \sqrt{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2};$$

$$x \in l^2: \quad \|x\| = \sqrt{\sum_{j=1}^{\infty} (x_j)^2};$$

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

$$f \in L^1(E): \quad \|f\| = \int_E |f(x)| dm;$$

$$f \in L^2(E): \quad \|f\| = \left\{ \int_E |f(x)|^2 dm \right\}^{\frac{1}{2}}.$$

关于以上空间的性质、应用,在数学领域中有专门的研究,属于泛函分析、点集拓扑等的研究对象,有兴趣的读者可以参看有关的参考文献.

§ 6.4 分布理论

数学科学是一切科学的基础,而自然科学的发展又促进和推动了数学学科的发展,这种例子屡见不鲜.在本节中,我们要从一个著名而又有实用价值的例子开始,介绍在近代数学、物理学等自然科学领域中占有重要的地位的“分布理

论”.

物理学家们远在数学家对分布理论给出严格定义之前,就已使用了的“ δ -函数”(Dirac 函数),他们使用了如下定义

$$\delta(x) = \begin{cases} \infty, & x=0, \\ 0, & x \neq 0, \end{cases}$$

作为压力的强度,表示单位压力全部集中在原点 0 时,直线上任一点 x 处的压强就是 $\delta(x)$.但这个“ δ 函数”却久久不能被数学家接受,直到四十年代末法国数学家 Schwartz, L. 提出分布理论,才沟通了数学理论和物理应用的通道,使物理概念有了坚实的数学基础,而数学也有了强烈的物理背景.

要强调指出,分布理论所需的数学基础知识很多、面也很广,许多知识都已超出本教材的范围,我们只在较为特殊的情况下给出基本概念和重要性质,以满足非数学类学科的需要.有进一步需求或有兴趣的读者,可参考有关文献.

6.4.1 连续线性泛函

回忆在第一章中引进的函数概念 f ,是指由其定义域 $D_f \subset \mathbf{R}$ 到其值域 $R_f \subset \mathbf{R}$ 的对应关系.现在来考虑更一般的情况.

定义 6.4.1 (函数、映射、泛函) 设 $f: D_f \rightarrow R_f$ 为一个对应关系,

若 $D_f \subset \mathbf{R}$, $\mathbf{R}$ 为实数域, $R_f \subset \mathbf{R}$, $\mathbf{R}$ 为实数域,则称 f 为函数;

若 $D_f \subset X$, X 为一空间, $R_f \subset Y$, Y 为一空间,则称 f 为映射,或算子;

若 $D_f \subset X$, X 为一空间, $R_f \subset \mathbf{R}$, $\mathbf{R}$ 为实数域,则称 f 为泛函.

进而,我们有

定义 6.4.2 (线性泛函) 设 E 为一个赋范线性空间,称泛函 f 为定义在 E 上的线性泛函,如果对任意的 $x, y \in E$ 与 $\alpha \in \mathbf{R}$, 有

$$f(x+y) = f(x) + f(y), \quad (6.4.1)$$

$$f(\alpha x) = \alpha f(x). \quad (6.4.2)$$

这里 E 可以是 $\mathbf{R}, \mathbf{R}^n, l^2$ 等,也可以是函数空间 $C[a, b], L(E), L^2(E)$ 等.

于是,线性泛函就是取值于数集 $\mathbf{R}$ (也可能取值为复数集 $\mathbf{C}$) 的线性映射,或称线性算子. f 在 $x \in E$ 的取值 $f(x)$, 可视为 f 对 x 的作用,用记号 $\langle f, x \rangle$ 表示.

定义 6.4.3 (连续线性泛函) 赋范线性空间 E 上的线性泛函称为连续的,如果对任意的 $\epsilon > 0$, 存在 $\delta > 0$, 使当 $x, y \in E$ 且 $\|x - y\| < \delta$ 时, 有

$$|\langle f, x \rangle - \langle f, y \rangle| < \epsilon.$$

E 上连续线性泛函的全体记为 $\mathcal{L}(E) = \mathcal{L}(E, \mathbf{R})$.

例 1 设 $E = L^1(\mathbf{R})$, 则 $f \in \mathcal{L}(E)$ 对 $\varphi \in L^1(\mathbf{R})$ 的作用就是

$$\langle f, \varphi \rangle = \int_{\mathbf{R}} f(x) \varphi(x) dx.$$

6.4.2 分布空间, 分布的性质

取基本空间为 $X = \mathbf{R}^n$, 取 $\mathbf{R}^n$ 中的任一个闭的“立方体”(一般情形下应取为“紧集”)

$$K = \{x = (x_1, \dots, x_n) : a_1 \leq x_1 \leq b_1, \dots, a_n \leq x_n \leq b_n\},$$

记

$$C^\infty(\mathbf{R}^n) = \{\varphi : \varphi \text{ 是 } \mathbf{R}^n \text{ 上具有任意阶的连续偏导数的函数}\},$$

$$C^\infty(K) = \{\varphi \in C^\infty(\mathbf{R}^n) : \varphi \text{ 是在 } K \text{ 之外为 } 0 \text{ 的函数}\},$$

并记

$$C_0^\infty(\mathbf{R}^n) = \bigcup_K C^\infty(K),$$

这里并的运算 $\bigcup_K C^\infty(K)$ 是让 K 变化而取的, 表示存在一个 K , 使 f 在 K 外为 0.

在集合 $C_0^\infty(\mathbf{R}^n)$ 中赋予代数结构, 使其成为一个线性空间, 也赋予拓扑结构 (称为归纳极限拓扑, 这种结构比较复杂, 我们不再介绍), 使其成为拓扑空间, 记为 $D(\mathbf{R}^n) \equiv C_0^\infty(\mathbf{R}^n)$.

取空间 $D(\mathbf{R}^n)$ 的连续线性泛函空间 $\mathcal{L}(D(\mathbf{R}^n)) \equiv D^*(\mathbf{R}^n)$.

定义 6.4.4 (分布空间、分布) 在 $D(\mathbf{R}^n)$ 的连续线性泛函空间 $D^*(\mathbf{R}^n)$ 上, 赋予一定的拓扑 (称为弱* 拓扑), 所成的拓扑线性空间称为 $D(\mathbf{R}^n)$ 的分布空间, 其中的元 $f \in D^*(\mathbf{R}^n)$ 称为 D^* -分布, 或简称分布. 分布 f 对 $\varphi \in D(\mathbf{R}^n)$ 的作用记为 $\langle f, \varphi \rangle$.

按定义, 一个 D^* -分布 $f \in D^*(\mathbf{R}^n)$ 就是 $D(\mathbf{R}^n)$ 上的一个连续线性泛函, 从定义也可立即得到分布的如下重要性质:

1. 分布的线性性质 若 $f \in D^*(\mathbf{R}^n)$, 则

$$\langle f, \alpha_1 \varphi_1 + \alpha_2 \varphi_2 \rangle = \alpha_1 \langle f, \varphi_1 \rangle + \alpha_2 \langle f, \varphi_2 \rangle,$$

对 $\varphi_1, \varphi_2 \in D(\mathbf{R}^n)$ 都成立, 其中 $\alpha_1, \alpha_2 \in \mathbf{R}$.

2. 分布的线性运算 若 $f, g \in D^*(\mathbf{R}^n)$, 则

$$\langle \alpha_1 f + \alpha_2 g, \varphi \rangle = \alpha_1 \langle f, \varphi \rangle + \alpha_2 \langle g, \varphi \rangle,$$

对 $\varphi \in D(\mathbf{R}^n)$ 都成立, 其中 $\alpha_1, \alpha_2 \in \mathbf{R}$.

我们熟悉的函数中, 有许多都是分布, 有如下例子.

例 2 $L(\mathbf{R}^n) \subset D^*(\mathbf{R}^n)$.

亦即, L -可积的函数 f 都是 D^* 分布, 并且

$$\langle f, \varphi \rangle = \int_{\mathbf{R}^n} f(x) \varphi(x) dm,$$

更一般地,在任意有限区间上 L -可积的函数都是 D^* 分布.例如 Heaviside 函数 $H(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases}$ 就是一例.

例 3 δ -函数就是一个 D^* 分布

$$\langle \delta, \varphi \rangle = \varphi(0).$$

分布理论是近代科学中非常重要的数学工具,这个理论的建立不仅解决了科学家们对数学科学的期盼,而且也给出了严格的数学理论,对自然科学与人文科学都会产生深远的影响.

下面我们对分布来建立它的运算规则.

(1) 零分布 若 $f \in D^*(\mathbf{R}^n)$,

$$\langle f, \varphi \rangle = 0$$

对每个 $\varphi \in D(\mathbf{R}^n)$ 都成立,则称 f 为零分布.

(2) 分布的相等 若 $f, g \in D^*(\mathbf{R}^n)$, 等式

$$\langle f - g, \varphi \rangle = 0$$

对每个 $\varphi \in D(\mathbf{R}^n)$ 都成立,则称两个分布 f, g 相等,记为 $f = g$.

(3) 分布的平移 若 $f \in D^*(\mathbf{R}^n)$, 定义满足

$$\langle g, \varphi \rangle = \langle f, \varphi_{-a} \rangle, \quad \text{其中 } \varphi \in D(\mathbf{R}^n), \quad \varphi_{-a}(x) = \varphi(x + a)$$

的分布 $g \in D^*(\mathbf{R}^n)$ 称为分布 f 的平移,记为 $g = f_a$, 或 $\tau_a f$.

(4) 分布的反射 若 $f \in D^*(\mathbf{R}^n)$, 定义满足

$$\langle h, \varphi \rangle = \langle f, \tilde{\varphi} \rangle, \quad \text{其中 } \varphi \in D(\mathbf{R}^n), \quad \tilde{\varphi}(x) = \varphi(-x)$$

的分布 $h \in D^*(\mathbf{R}^n)$ 称为分布 f 的反射,记为 $h = \tilde{f}$.

(5) 分布与函数的乘积 若 $f \in D^*(\mathbf{R}^n)$, 定义满足

$$\langle k, \varphi \rangle = \langle f, a\varphi \rangle, \quad \text{其中 } \varphi \in D(\mathbf{R}^n), \quad a \in C^\infty(\mathbf{R}^n)$$

的分布 $k \in D^*(\mathbf{R}^n)$ 称为分布 f 与函数 a 的乘积,记为 $k = af$.

(6) 分布的导数 若 $f \in D^*(\mathbf{R}^n)$, 定义满足

$$\langle l, \varphi \rangle = -\langle f, \partial_{x_j} \varphi \rangle, \quad \text{其中 } \varphi \in D(\mathbf{R}^n), \quad \partial_{x_j} \varphi \text{ 是 } \varphi \text{ 的偏导数}$$

的分布 $l \in D^*(\mathbf{R}^n)$ 称为 f 关于 x_j 的一阶偏导数,记为 $l = \partial_{x_j} f$.

定理 6.4.1(分布的求导次序) 设 $f \in D^*(\mathbf{R}^n)$, 则 f 为无穷次可导分布, 并且其求导结果与求导次序无关,例如,

$$\partial_{x_j x_i}^2 f = \partial_{x_i x_j}^2 f.$$

例 4 对于在物理中颇有用途的 Heaviside 函数 $H(x)$, 它在经典意义下的导数没有多大的意义.然而,在分布意义下,视 $H' \in D^*(\mathbf{R}^n)$, 却有

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_0^\infty \varphi' dx = \varphi(0) = \langle \delta, \varphi \rangle,$$

故 $H' = \delta$.

以上逐条使我们能把函数类 $C^\infty(E)$ 中函数的许多性质推广到更一般的函数. 我们还要把傅里叶变换向分布空间推广.

6.4.3 $\mathcal{S}^*$ 上的傅里叶分析

在 §5.3 中讲述了傅里叶变换的理论与应用, 在那里, 对 $f \in L^1$ 与 $f \in L^2$ 定义了傅里叶变换, 现在我们要把傅里叶变换向 $\mathcal{S}^*$ 分布推广.

在分布空间 $\mathcal{S}^*(\mathbf{R}^n)$ 上建立傅里叶分析理论, 其理论与应用价值都是极为明显的.

经典傅里叶分析理论是从“急降函数类 $\mathcal{S}$ ”开始的.

$$\mathcal{S} = \{ \varphi \in C^\infty(\mathbf{R}^n) : \sup_{\mathbf{R}^n} |x^\alpha D^\beta \varphi(x)| < \infty, \forall \alpha, \beta \in \mathbf{R}^n \},$$

$\mathcal{S}$ 显然是一个线性空间, 在赋予一定的拓扑后, 它也是一个拓扑空间. 求导算子

$$\partial^\alpha = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n},$$

其中 $\alpha = (\alpha_1, \dots, \alpha_n), \beta = (\beta_1, \dots, \beta_n) \in \mathbf{R}^n$. 我们有

$$D(\mathbf{R}^n) \subset \mathcal{S}(\mathbf{R}^n) \subset D^*(\mathbf{R}^n). \quad (6.4.3)$$

定义 6.4.5 (分布的傅里叶变换) $\varphi \in \mathcal{S}(\mathbf{R}^n)$ 的傅里叶变式 $\varphi^\wedge$ 定义为

$$\varphi^\wedge(\xi) = \int_{\mathbf{R}^n} e^{-i(x, \xi)} \varphi(x) dm, \quad \xi \in \mathbf{R}^n, \quad (6.4.4)$$

其中 $(x, \xi) = \sum_{j=1}^n \xi_j x_j$. 记 $F: \varphi \rightarrow \varphi^\wedge$, 称映射 F 为 $\mathcal{S}$ 上的傅里叶变换, 而称 $\varphi^\wedge$ 为 φ 的傅里叶变换. 当 $n=1$ 时, 定义 6.4.4 与 (5.3.3) 相同.

定义 6.4.6 (分布的傅里叶逆变换) $\psi \in \mathcal{S}(\mathbf{R}^n)$ 的傅里叶逆变换 $\psi^\vee$ 定义为

$$F^{-1}(\psi)(\xi) = \psi^\vee(\xi) = (2\pi)^{-n} \int_{\mathbf{R}^n} e^{i(x, \xi)} \psi(x) dx,$$

我们列举 $\varphi \in \mathcal{S}$ 的傅里叶变换的主要性质:

(1) 若 $\varphi \in \mathcal{S}$, 则 $\varphi^\wedge \in \mathcal{S}$, 且满足

(i) $(D^\alpha \varphi)^\wedge(\xi) = \xi^\alpha \varphi^\wedge(\xi) \quad (\xi \in \mathbf{R}^n),$

(ii) $(x^\alpha \varphi)^\wedge(\xi) = (-D_\xi)^\alpha \varphi^\wedge(\xi) \quad (\xi \in \mathbf{R}^n),$

(2) 对于 $\varphi \in \mathcal{S}$, 我们有傅里叶变换的反演公式: $\varphi^\wedge \in \mathcal{S}$, 且

$$\varphi(x) = (2\pi)^{-n} \int_{\mathbf{R}^n} e^{i(x, \xi)} \varphi^\wedge(\xi) d\xi,$$

由此定义傅里叶逆变换 $F^{-1}: \psi \rightarrow \psi^\vee$.

(3) 对于 $\varphi \in \mathcal{S}(\mathbf{R}^n)$

- (i) $(\tilde{\varphi}(x))^\wedge(\xi) = (2\pi)^n \varphi^\vee(\xi),$
- (ii) $(\varphi^\wedge)^\vee(x) = \varphi(x),$
- (iii) $(\tau_h \varphi)^\wedge(\xi) = e^{-i(h, \xi)} \varphi^\wedge(\xi),$
- (iv) $(\varphi(cx))^\wedge(\xi) = |c|^{-n} \varphi^\wedge(c^{-1}\xi),$
- (v) $(\varphi * \psi)^\wedge(\xi) = \varphi^\wedge(\xi) \psi^\wedge(\xi).$

(4) 映射 $F: \mathcal{S}(\mathbf{R}^n) \rightarrow \mathcal{S}(\mathbf{R}^n)$ 是一个“一一对应”的映射: $\varphi \in \mathcal{S}(\mathbf{R}^n) \Rightarrow \varphi^\wedge \in \mathcal{S}(\mathbf{R}^n)$; 且 $F^{-1}: \varphi^\wedge \rightarrow \varphi$ 是“一一对应”的映射. 同时, $\mathcal{S}^*(\mathbf{R}^n)$ 也是一个分布空间, 满足 $\mathcal{S}^*(\mathbf{R}^n) \subset D^*(\mathbf{R}^n)$.

如何把傅里叶变换推广到 $\mathcal{S}^*$ 上去呢? 我们利用 $\mathcal{S}$ 的傅里叶变换所满足的一个性质: 若 $\varphi, \psi \in \mathcal{S}$, 则

$$\int_{\mathbf{R}^n} \varphi^\wedge(\xi) \overline{\psi^\wedge(\xi)} d\xi = 2\pi \int_{\mathbf{R}^n} \varphi(x) \overline{\psi(x)} dx, \quad (6.4.5)$$

这个关系式也称为帕塞瓦尔公式. 因为关系式(6.4.3)

$$D(\mathbf{R}^n) \subset \mathcal{S}(\mathbf{R}^n) \subset D^*(\mathbf{R}^n),$$

故(6.4.5)式可改写为

$$\langle \psi^\wedge, \varphi^\wedge \rangle = (2\pi)^n \langle \psi, \varphi \rangle.$$

于是, 对于分布 $f \in \mathcal{S}^*(\mathbf{R}^n)$, 定义分布 f 的傅里叶变式为满足下式的分布 $f^\wedge$,

$$\langle f^\wedge, \varphi^\wedge \rangle = (2\pi)^n \langle f, \varphi \rangle. \quad (6.4.6)$$

或, 等价地用下式代替(6.4.6),

$$\langle f^\wedge, \varphi \rangle = (2\pi)^n \langle f, \varphi^\wedge \rangle. \quad (6.4.7)$$

这个定义加上严格的数学推理, 使我们完成了傅里叶变换到 $\mathcal{S}^*$ 的推广及 $\mathcal{S}^*$ - 傅里叶变换性质的证明. 前述的性质(1)、(2)、(3)、(4)、(5)都可以推广到 $\mathcal{S}^*$ - 分布上去. 我们举一个例子, 说明这种推广是如何完成的.

例如, 我们证明性质

$$\begin{aligned} (x^a \varphi)^\wedge &= (-D)^a \varphi^\wedge, & \varphi \in \mathcal{S} \\ (D^a \varphi)^\wedge(\xi) &= \xi^a \varphi^\wedge(\xi), & \varphi \in \mathcal{S} \end{aligned}$$

对 $f \in \mathcal{S}^*$ 也成立.

先证明 $(x^a f)^\wedge = (-D)^a f^\wedge$. 任取 $\varphi \in \mathcal{S}^*$,

$$\begin{aligned} \langle (x^a f)^\wedge, \varphi \rangle &= \langle x^a f, \varphi^\wedge \rangle = \langle f, x^a \varphi^\wedge \rangle = \langle f, (D^a \varphi)^\wedge \rangle \\ &= \langle f^\wedge, D^a \varphi \rangle = \langle (-D)^a f^\wedge, \varphi^\wedge \rangle, \end{aligned}$$

从而有 $(x^a f)^\wedge = (-D)^a f^\wedge$.

再证 $(D^a f)^\wedge = \xi^a f^\wedge$, $f \in \mathcal{S}^*$. 任取 $\varphi \in \mathcal{S}^*$,

$$\begin{aligned} \langle (D_j f)^\wedge, \varphi \rangle &= \langle D_j f, \varphi^\wedge \rangle = -\langle f, D_j \varphi^\wedge \rangle \\ &= -\langle f, -(x_j \varphi)^\wedge \rangle = \langle f, (x_j \varphi)^\wedge \rangle = \langle f^\wedge, x_j \varphi \rangle \\ &= \langle \xi_j f^\wedge, \varphi \rangle, \end{aligned}$$

故得 $(D^\alpha f)^\wedge = \xi^\alpha f^\wedge$, $f \in \mathcal{S}^*$.

由于卷积至关重要,我们再介绍卷积的几个性质.

定理 6.4.2 (分布的卷积性质) $f, g, h \in D^*(\mathbf{R}^n)$, $\varphi, \psi, \chi \in D(\mathbf{R}^n)$, 定义卷积 $(f * \varphi)(x) = \langle f, \tilde{\varphi}_x \rangle$, 其中 $\tilde{\varphi}_x(y) = \tilde{\varphi}(y-x) = \varphi(x-y)$. 则

- (1) $(f * \varphi)(x) = (\varphi * f)(x)$, $f * \delta = \delta * f = f$, δ 为狄拉克函数,
- (2) $(\alpha f + \beta g) * \varphi = \alpha f * \varphi + \beta g * \varphi$,
- (3) $(f * \varphi) * \psi = f * (\varphi * \psi)$,
- (4) $\tau_h(f * \varphi) = \tau_h f * \varphi = f * \tau_h \varphi$,
- (5) $D^\alpha(f * \varphi) = D^\alpha f * \varphi = f * D^\alpha \varphi$,

为了熟悉和掌握分布的性质和运算,我们再证明如上的(1)、(3).

证 先证(3). 按卷积定义 $(f * \varphi)(x) = \langle f, \tilde{\varphi}_x \rangle$, 有

$$\begin{aligned} \langle f * (\varphi * \psi), \chi \rangle &= \langle f, \langle \varphi, \langle \psi, (\chi_x)_x \rangle \rangle \rangle \\ &= \langle f * \varphi, \langle \psi, \chi_x \rangle \rangle = \langle (f * \varphi) * \psi, \chi \rangle. \end{aligned}$$

次证(1) 我们断言, “ $f_1 * \varphi = f_2 * \varphi$ 蕴含 $f_1 = f_2$ ”. 事实上, 取 $\varphi = \tilde{\psi}$, 由定义

$$\langle f_1, \varphi \rangle = \langle f_1, \tilde{\psi} \rangle = (f_1 * \psi)(0) = (f_2 * \tilde{\psi})(0) = \langle f_2, \varphi \rangle$$

而得 $f_1 = f_2$.

于是, 由 $\psi * \chi = \chi * \psi \in D(\mathbf{R}^n)$ 便可推得 $f * \varphi = \varphi * f$.

对于 $f * \delta = f$, 当 $f = \varphi \in D(\mathbf{R}^n)$ 时, 我们有

$$\delta * \varphi(x) = \langle \delta, \tilde{\varphi}_x \rangle = \tilde{\varphi}_x(0) = \varphi(x).$$

一般地, 对 $f \in D^*(\mathbf{R}^n)$ 时, 有

$$(f * \delta) * \varphi = f * (\delta * \varphi) = f * \varphi,$$

从而, $f * \delta = f$ 得证. 这个式子理解为: δ 是卷积运算 $*$ 的单位元.

§ 6.5 Banach 空间微积分学

经典微积分仍然是高等数学的主要内容, 掌握其中的几个重要概念, 如极限、微分、积分、级数等, 以及相应的数学方法和技巧, 一方面是由于它们在各个学科领域的基础理论和实际应用中仍然大量地出现, 另一方面也是因为现代科学发展不断出现的新问题和新模型常常是以经典情形为雏形的. 在经典模型的启发下, 引进新的概念、理论、技巧、方法, 不仅解决了实际需求, 而且促进了数学科学自身的发展. 本节要以极限、级数、微分、积分概念和理论从 $\mathbf{R}$ 推广到 Banach 空间为例, 向读者展示数学的重要思维方法, 并使之在其自身的研究领域中能运用自如.

6.5.1 极限

回顾 § 2.2 中介绍的极限概念,我们可理解为“当自变量任意接近一个固定的量时,其函数也随之任意接近一个固定的量”.以函数的极限为例,

$$\lim_{x \rightarrow a} f(x) = A,$$

易见,要推广这个概念,可考虑推广以下方面:

(1) $x \rightarrow a$,只要把自变量 x 变化的范围一定义域(经典情形是 $x \in \mathbf{R}$ 或 $\mathbf{R}$ 的子集)推广到一个能测量元素之间“靠近”关系的集合.

(2) 类似地,对 $f(x)$ 任意接近 A ,也只要把因变量 $f(x)$ 变化的范围一值域(经典情形也是 $f(x) \in \mathbf{R}$ 或 $\mathbf{R}$ 的子集)推广到一个能测量元素之间“靠近”关系的集合.

(3) 还注意到,定义域和值域都只要能度量其元素之间的远近程度,并不一定是同一个集合.

注意到以上三点,我们以巴拿赫空间为例来推广已经熟悉的概念.

定义 6.5.1 (极限、连续性) 设 $(X, \|\cdot\|_X)$ 与 $(Y, \|\cdot\|_Y)$ 为两个巴拿赫空间, $f: X \rightarrow Y$ 为 X 到 Y 的映射.对于自变量 $x \in X$ 与因变量 $y = f(x) \in Y$,我们称 $b \in Y$ 为 $f(x)$ 当 $x \rightarrow a, x \neq a$ 时的极限,记为

$$\lim_{x \rightarrow a} f(x) = b,$$

如果对于任一给定的 $\epsilon > 0$,存在 $\delta > 0$,使得只要 $\|x - a\|_X < \delta$,就有

$$\|f(x) - b\|_Y < \epsilon$$

当 $f(x)$ 在 $a \in X$ 有定义并且 $f(a) = b$ 时,就称映射 f 在 a 处连续.

6.5.2 导数

现在来定义巴拿赫空间中的连续映射的导数.也回忆一下经典导数的定义

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0), \quad x \in (a, b) \subset \mathbf{R}.$$

这里极限值表示函数 $f: (a, b) \rightarrow \mathbf{R}$ 在 $x_0 \in (a, b)$ 的变化率.若将上式改写为

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - f'(x_0) \cdot (x - x_0)}{x - x_0} = 0,$$

而把 $y = f(x_0) + f'(x_0) \cdot (x - x_0)$ 理解为通过 x_0 的曲线的切线,就可以启示我们去定义巴拿赫空间中映射的相切概念,并且给出关于映射的可导概念.

定义 6.5.2 (映射的相切) 设 X 与 Y 为 Banach 空间, f, g 为 X 到 Y 的两个映射.若对于 $x_0 \in X$,

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{\|f(x) - g(x)\|_Y}{\|x - x_0\|_X} = 0,$$

则称 f 与 g 在 $x_0 \in A$ 相切.

显然,我们有如下简单性质.

定理 6.5.1 (映射相切的性质) (i) 若 f, g 在 x_0 相切, 并且 f, g 在 x_0 连续, 则

$$f(x_0) = g(x_0).$$

(ii) 相切有传递性, 亦即在 x_0 点, 若 f 与 g 相切, 又 g 与 h 相切, 则 f 与 h 在该点相切.

(iii) 与映射 f 在 x_0 相切的所有映射中, 最多只有一个形如

$$x \rightarrow f(x_0) + u(x - x_0)$$

的映射, 这里 u 为 X 到 Y 的线性映射.

于是, 我们有

定义 6.5.3 (F -可导) 设 X, Y 为巴拿赫空间, $f: X \rightarrow Y$ 为连续映射. 若对 $x_0 \in X$, 存在 X 到 Y 的线性映射 $u: X \rightarrow Y$, 使得映射 $g: x \rightarrow f(x_0) + u(x - x_0)$ 与 f 在 x_0 相切, 则称 f 在 $x = x_0$ 点 F -可导 (Fréchet 可导), 而线性映射 u 称为 f 在 x_0 点的 F -导数, 记为 $f'(x_0)$ 或 $Df(x_0)$. 若 f 在 A 的每一点都 F -可导, 就称 f 在 A 上 F -可导. 如不发生混淆, 我们简称 F -导数为导数.

例 1 常映射 $f: X \rightarrow Y$ 在 X 的任一点可导, 其导数为 $f'(x) = \theta \in \mathcal{L}(X; Y)$, θ 为 $\mathcal{L}(X; Y)$ 的零元.

例 2 X 到 Y 的连续线性映射 $u: X \rightarrow Y$ 在 X 的任一点可导, 且 $Du = u$.

例 3 若 f 为 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 上的可导映射. 设 f 在选定的一个基底下的分量为 $(f_1, \dots, f_m)$, 则 f 可导的充要条件是 f_j 可导, $j = 1, \dots, m$, 且其导数为

$$\left(\frac{\partial f_j}{\partial x_k} \right)_{j=1, \dots, m; k=1, \dots, n}.$$

如经典情形一样, 在 F -可导意义下, 也有复合映射及逆映射的求导法则.

以后使用如下方便的记法: 若 f 在 x_0 可导, 则任给 $\epsilon > 0$, 存在 $r > 0$, 使得

$$f(x_0 + \Delta x) - f(x_0) = f'(x_0)(\Delta x) + o(\Delta x)$$

对于 $\|\Delta x\| \leq r$ 成立; 或

$$f(x_0 + s) - f(x_0) = f'(x_0)(s) + o(s)$$

对于 $\|s\| \leq r$ 成立, 这里 $o(s) \in Y$, 且满足 $\|o(s)\| \leq \epsilon \|s\|$.

定理 6.5.2 (复合映射的导数) 设 X, Y, Z 为巴拿赫空间, $x_0 \in X, f: X \rightarrow Y$ 为连续映射, $y_0 = f(x_0)$. 又设 $B = \{y \in Y: \|y - f(x_0)\|_Y < r_0\} \subset Y$ 为包含 $f(x_0)$ 的球. $g: B \rightarrow Z$ 为连续映射. 若 f 在 x_0 可导, g 在 y_0 可导, 则 $h = g \circ f$ 在 x_0 可导, 并且成立

$$h'(x_0) = g'(y_0) \circ f'(x_0).$$

下面的证明不需要读者仔细阅读,仅供参考.

证 任给 $\varepsilon > 0$, 存在 $r > 0$, 使得当 $\|s\| \leq r, \|t\| \leq r$ 时有

$$f(x_0 + s) = f(x_0) + f'(x_0) \cdot s + o_1(s), \quad \|o_1(s)\| \leq \varepsilon \|s\|,$$

$$g(y_0 + t) = g(y_0) + g'(y_0) \cdot t + o_2(t), \quad \|o_2(t)\| \leq \varepsilon \|t\|,$$

由线性映射 $f'(x_0)$ 与 $g'(y_0)$ 的连续性, 存在 $a > 0, b > 0$, 使对任意 $s \in X, t \in Y$, 有

$$\|f'(x_0) \cdot s\| \leq a \|s\|, \quad \|g'(y_0) \cdot t\| \leq b \|t\|.$$

从而当 $\|s\| \leq r$ 时,

$$\|f'(x_0) \cdot s + o_1(s)\| \leq (a + 1) \|s\|.$$

而当 $\|s\| \leq \frac{r}{a+1}$ 时,

$$\|o_2(f'(x_0) \cdot s + o_1(s))\| \leq (a + 1) \varepsilon \|s\|, \quad \|g'(y_0) \cdot o_1(s)\| \leq b \varepsilon \|s\|.$$

上面两式复合的结果给出:

$$h(x_0 + s) = g(y_0) + g'(y_0)(f'(x_0) \cdot s) + o_3(s),$$

且 $\|o_3(s)\| \leq (a + b + 1) \varepsilon \|s\|$. 从而 $h'(x_0) = g'(y_0) \circ f'(x_0)$. 定理得证.

用类似的方法, 不难证明逆映射的求导法则:

定理 6.5.3 (逆映射的导数) 设 X, Y 为巴拿赫空间, 若 $f: X \rightarrow Y$ 为双方单值双方连续的映射, 则称其为同胚映射. 若 f 在 $x_0 \in X$ 可导, 且 $f'(x_0)$ 为 X 到 Y 的线性同胚, 则 f^{-1} 在 $y_0 = f(x_0)$ 可导, 且

$$(f^{-1})'(y_0) = (f'(x_0))^{-1}.$$

上述求导法则, 虽与经典情形的形式类似, 但却有新的理解. 例如复合函数求导法则, 理解为两个线性映射 $g'(y_0)$ 与 $f'(x_0)$ 的复合. 又如逆映射求导法则, 理解为线性映射 $f'(x_0)$ 的逆映射, 等.

当我们取 $X = Y = \mathbb{R}$ 时, 现在的结果也有相应的解释. 例如, 容易看出, 定理 6.5.3 中的条件“ $f'(x_0)$ 为 X 到 Y 的线性同胚”, 相当于经典情形的“ $f'(x_0) \neq 0$ ”. 此外, 请读者考虑为什么在经典情形下 $f'(x_0)$ 只是一个实数, 而现在的情形下它却是一个线性映射呢?

例 4 设 X, Y, Z 为巴拿赫空间, $f: X \times Y \rightarrow Z$ 为连续线性映射, 则 f 在 $X \times Y$ 的任一点 (x_0, y_0) 可导, 且

$$Df(x_0, y_0)(s, t) = f(x_0, t) + f(s, y_0).$$

例 5 微分算子 $A: C^1([0, 1]) \rightarrow C([0, 1])$, 其中

$$Af(x) = f'(x) + [f(x)]^2,$$

试求 $Af(x)$ 的导数.

计算后得

$$DA(g)(h)(x) = h'(x) + 2g(x)h(x).$$

例 6 乌里松(Урысон)算子

$$Af(x) = \int_0^1 k(x, y, f(y)) dy \quad (x, y \in [0, 1], f \in C([0, 1])).$$

$C([0, 1])$ 为 $[0, 1]$ 上的复值连续函数空间. 设函数 $k(x, y, z)$ 关于第三个变量 z 的偏导数 $\frac{\partial k}{\partial z}: [0, 1] \times [0, 1] \times \mathbb{C} \rightarrow \mathbb{C}$ 一致连续. 计算 A 在 $g \in C([0, 1])$ 处的 F -导数.

由于

$$\begin{aligned} & \left| k(x, y, g(y) + h(y)) - k(x, y, g(y)) - \frac{\partial k}{\partial z}(x, y, g(y))h(y) \right| \\ &= \left| \int_0^1 \left\{ \frac{\partial k}{\partial z}(x, y, g(y) + th(y)) - \frac{\partial k}{\partial z}(x, y, g(y)) \right\} h(y) dt \right| \\ &\leq \|h\| \int_0^1 \left| \frac{\partial k}{\partial z}(x, y, g(y) + th(y)) - \frac{\partial k}{\partial z}(x, y, g(y)) \right| dt, \end{aligned}$$

又由于 $\frac{\partial k}{\partial z}$ 的一致连续性, 故

$$\begin{aligned} & \left| A(g+h)(x) - Ag(x) - \int_0^1 \frac{\partial k}{\partial z}(x, y, g(y))h(y) dy \right| \\ &\leq \sup_{x \in [0, 1]} \left| \int_0^1 \left\{ k(x, y, g(y) + h(y)) - k(x, y, g(y)) - \frac{\partial k}{\partial z}(x, y, g(y))h(y) \right\} dy \right| \\ &= o(\|h\|). \end{aligned}$$

从而 $DA(g)h(x) = \int_0^1 \frac{\partial k}{\partial z}(x, y, g(y))h(y) dy \quad (h \in C([0, 1]))$.

还有很多应用的例, 我们不一一列举了. 读者可参考[4, 6, 9]等书目.

在现代数学、物理学以及许多应用科学中, 我们还会遇到其他类型的导数, 如伽多(Gâteaux)导数、吉布斯(Gibbs)导数、 p -型导数等概念, 这诸多导数概念各有各的背景、意义和应用. 读者多注意引进各种导数的思路是有益的.

6.5.3 级数

这里假设 X 为数域 K (实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$) 上的巴拿赫空间.

定义 6.5.4 (级数) 设 $x_n \in X$. 若对任意的 $n \geq 0$, 有 $x_0 + \cdots + x_n = \sum_{k=0}^n x_k = s_n$, 则称一对序列 $(x_n)_{n \geq 0}, (s_n)_{n \geq 0}$ 为一个级数, 记为 $\sum_{n=0}^{\infty} x_n$. 称 x_n 为此级数的第 n 项, 而称 s_n 为其第 n 部分和.

我们称级数 $\sum_{n=0}^{\infty} x_n$ 收敛到 s , 若 $\lim_{n \rightarrow \infty} s_n = s$. 精确地说, 对任给的 $\varepsilon > 0$, 存在 $n_0 \geq 0$, 使当 $n \geq n_0$ 时, 有 $\|s_n - s\| < \varepsilon$. 记为 $\sum_{n=0}^{\infty} x_n = s$, 并称 s 为级数的和.

类似于初等微积分中的一些定理, 我们有

定理 6.5.4 (柯西准则) 若 X 中的级数 $\sum_{n=0}^{\infty} x_n$ 收敛, 则对于任给的 $\varepsilon > 0$,

存在 $n_0 \geq 0$, 使当 $n \geq n_0, p \geq 0$ 时, 有 $\|s_{n+p} - s_n\| = \|x_{n+1} + \cdots + x_{n+p}\| \leq \varepsilon$.

反之, 若上述条件满足, 则级数 $\sum_{n=0}^{\infty} x_n$ 收敛于 $s \in X$.

定理 6.5.5 (收敛级数的和) 若 X 中的级数 $\sum_{n=0}^{\infty} x_n$ 与 $\sum_{n=0}^{\infty} y_n$ 分别收敛于 s 与 s' , 则级数 $\sum_{n=0}^{\infty} (x_n + y_n)$ 与 $\sum_{n=0}^{\infty} (\alpha x_n)$ 分别收敛于 $s + s'$ 与 αs , 这里 $\alpha \in K$.

定理 6.5.6 (收敛级数可加括号) 设 $0 = k_0 < k_1 < k_2 < \cdots$ 为非负整数序列. 若级数 $\sum_{n=0}^{\infty} x_n$ 收敛于 s , 且 $y_n = \sum_{j=k_n}^{k_{n+1}-1} x_j$, 则级数 $\sum_{n=0}^{\infty} y_n$ 也收敛于 s .

定义 6.5.5 (绝对收敛级数) 若 X 中的级数 $\sum_{n=0}^{\infty} x_n$ 的各项的范数所成的数项级数 $\sum_{n=0}^{\infty} \|x_n\|$ 收敛, 则称级数 $\sum_{n=0}^{\infty} x_n$ 绝对收敛.

显然成立如下定理.

定理 6.5.7 (绝对收敛级数必收敛) 若 X 中的级数 $\sum_{n=0}^{\infty} x_n$ 绝对收敛, 则 $\sum_{n=0}^{\infty} x_n$ 也收敛. 并且成立

$$\left\| \sum_{n=0}^{\infty} x_n \right\| \leq \sum_{n=0}^{\infty} \|x_n\|.$$

记 $N = \{0, 1, 2, 3, \cdots\}$ 为非负整数集, $N_+ = \{1, 2, \cdots\}$ 为正整数集. 我们有

定理 6.5.8 (绝对收敛级数的可交换性) 设 $\sum_{n=0}^{\infty} x_n$ 为 X 中的绝对收敛级数. 令 $\sigma: N \rightarrow N$ 为一对一的满映射, 若记 $y_n = x_{\sigma(n)}$, 则 $\sum_{n=0}^{\infty} y_n$ 也绝对收敛, 并且

$$\sum_{n=0}^{\infty} x_n = \sum_{n=0}^{\infty} y_n.$$

定理 6.5.9 (绝对收敛级数的可结合性) 设 $\sum_{n=0}^{\infty} x_n$ 为 X 中的绝对收敛级数. 若 $N = \bigcup_{n \in I} B_n$, I 为有限或无限集, B_n 为两两互不相交的集合, $n \in I$. 令 $z_n = \sum_{k \in B_j} x_k$, 则 $\sum_{n=0}^{\infty} z_n$ 也绝对收敛, 且

$$\sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} x_n.$$

6.5.4 偏导数和高阶导数

巴拿赫空间 X 到 Y 的连续映射 $f: X \rightarrow Y$ 的导数是 X 到 Y 的连续线性映射, 偏导数与高阶导数应当是什么呢? 现在讨论这个问题. 我们仍约定, 以下所讨论的导数均指 F -导数.

两个巴拿赫空间 X_1 与 X_2 的积集 $X_1 \times X_2$, 定义为

$$X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\}.$$

积集 $X \equiv X_1 \times X_2$ 在范数

$$\|x\|_X \equiv \|(x_1, x_2)\|_{X_1 \times X_2} = \max\{\|x_1\|, \|x_2\|\}$$

之下也成为巴拿赫空间. 在巴拿赫空间 $X_1 \times X_2$ 中, 定义

$$A = \{(x_1, x_2) \in X_1 \times X_2 : \|x_1 - a_1\| < r, \|x_2 - a_2\| < r\}$$

为以 (a_1, a_2) 为中心, 以 r 为半径的开球.

定义 6.5.6 (偏导数) 设 X_1, X_2, Y 为 Banach 空间, 集合 A 为积空间 $X = X_1 \times X_2$ 中的开球, $(a_1, a_2) \in A$. 我们称映射 $f: X \rightarrow Y$ 在 (a_1, a_2) 关于第一个 (或第二个) 变元 a_1 是可导的, 若偏映射 $x_1 \mapsto f(x_1, a_2)$ (或 $x_2 \mapsto f(a_1, x_2)$) 在 a_1 (或在 a_2) 可导. 亦即, 存在线性映射 $u: X_1 \rightarrow Y$ (或 $v: X_2 \rightarrow Y$) 使得

$$\lim_{x_1 \rightarrow a_1} \frac{\|f(x_1, a_2) - f(a_1, a_2) - u(x_1 - a_1)\|}{\|x_1 - a_1\|} = 0,$$

(或

$$\lim_{x_2 \rightarrow a_2} \frac{\|f(a_1, x_2) - f(a_1, a_2) - v(x_2 - a_2)\|}{\|x_2 - a_2\|} = 0, \quad)$$

并且称线性映射 u (或 v) 为 f 关于 x_1 (或 x_2) 的偏导数, 记为 $\partial_1 f(a_1, a_2)$, (或 $\partial_2 f(a_1, a_2)$). 读者清楚, 这里 u (或 v) 与 a_2 (或 a_1) 有关.

类似地, 可以证明, 若 f 在 (a_1, a_2) 关于第一个变元 x_1 可导, (或关于第二个变元 x_2 可导), 则

$$\partial_1 f(a_1, a_2) \in \mathcal{L}(X_1, Y), \quad (\text{或 } \partial_2 f(a_1, a_2) \in \mathcal{L}(X_2, Y).)$$

术语“连续可导”是指映射 $f: A \rightarrow Y$ 可导, 而且 f' 为 A 中的连续映射.

定理 6.5.10 (连续可导映射的条件) 设 X_1, X_2, Y 如上, A 是积空间 $X = X_1 \times X_2$ 中的开球, $f: A \rightarrow Y$ 为连续映射, 则 f 在 $X_1 \times X_2$ 中为连续可导映射, 当且仅当 f 关于第一个与第二个变元都可导, 且 $\partial_1 f$ 与 $\partial_2 f$ 在 A 中连续. 此时, 还成立公式

$$Df(x_1, x_2)(t_1, t_2) = \partial_1 f(x_1, x_2) \cdot t_1 + \partial_2 f(x_1, x_2) \cdot t_2.$$

现在我们讨论高阶导数. 此时它不仅是线性映射而是多重线性映射.

定义 6.5.7 (二阶导数) 设 X, Y 为巴拿赫空间, $A \subset X$ 为开球. $f: A \rightarrow Y$

为连续可导映射. 若 f' 作为 X 到 $\mathcal{L}(X, Y)$ 的映射在 $x_0 \in A$ 可导, 则称 f 在 x_0 为二次可导, 其二阶导数记为 $f''(x_0)$ 或 $D^2 f(x_0)$. 若 f 在 A 的每一点都是二次可导, 则称 f 在 A 上二次可导.

任一个连续映射 $f: A \rightarrow Y$, 若它在 $x_0 \in A$ 的 n 阶导数存在, 记为 $f^{(n)}(x_0)$, 它便是 $X \times X \times \cdots \times X$ 到 Y 的连续 n 重线性映射, 也就是

$$f^{(n)} \in \mathcal{L}(X \times \cdots \times X; Y).$$

本小节末, 我们介绍高阶导数如下的几个重要性质.

定理 6.5.11 (二阶导数的一个公式) 设 X, Y 为巴拿赫空间, $A \subset X$ 为开球. $f: A \rightarrow Y$ 为连续映射. 若 f 在 $x_0 \in A$ 为二次可导, 则对于任意固定元 $t \in X$, A 到 Y 的映射 $g: x \rightarrow Df(x) \cdot t$ 在 x_0 的导数为

$$g'(x_0) \cdot s = D^2 f(x_0) \cdot (s, t)$$

证 只需视映射 g 为如下两个映射的复合便得所需的结果

$$\varphi: Df(x) \rightarrow Df(x) \cdot t, \quad \psi: x \rightarrow Df(x),$$

其中 φ 为 $\mathcal{L}(X; Y)$ 到 Y 的映射, ψ 为 X 到 $\mathcal{L}(X; Y)$ 的映射, 而 $g = \varphi \circ \psi$.

定理 6.5.12 (二阶导数的对称性) 若 f 在 $x_0 \in A$ 为二次可导, 则连续二重线性映射 $D^2 f(x_0)$ 对称. 即

$$D^2 f(x_0) \cdot (s, t) = D^2 f(x_0) \cdot (t, s) \quad (s, t \in X).$$

需证明对任给的 $\varepsilon > 0$, 及任意的 $(s, t) \in X \times X$, 有

$$\|D^2 f(x_0) \cdot (s, t) - D^2 f(x_0) \cdot (t, s)\| \leq \varepsilon \|s\| \|t\|,$$

而由 ε 的任意性便得到所需的等式. 我们从略.

定理 6.5.13 (高阶可导) 设 $f = (f_1, \cdots, f_m)$ 为由巴拿赫空间 X 的开球 A 到巴拿赫空间 $Y_1, \cdots, Y_m$ 的积空间 $Y_1 \times \cdots \times Y_m$ 的连续映射, 则 f 在 A 中 p 次可导, 当且仅当每个 f_j 在 A 中 p 次可导, $1 \leq j \leq m$. 并且成立

$$D^p f = (D^p f_1, \cdots, D^p f_m).$$

6.5.5 压缩映射定理和不动点定理

在非线性分析理论中, 压缩映射定理和不动点原理对研究存在性定理等问题是十分重要的. 我们介绍其中应用非常广泛的不动点原理. 但是由于一般形式的压缩映射定理、布劳尔(Brouwer)不动点原理和绍德尔(Schauder)不动点原理都需要许多预备知识, 我们只给出一些特殊情况.

满足 $f(x) = x$ 的元 x , 称为映射 f 的不动点.

定理 6.5.14 (布劳尔) $\bar{B}(0, 1) = \{x \in \mathbf{R}^n : \|x\| \leq 1\}$ 为 $\mathbf{R}^n$ 中的闭单位球. 若此球上的映射 $f: \bar{B}(0, 1) \rightarrow \bar{B}(0, 1)$ 为连续映射, 则存在 $x_0 \in \bar{B}(0, 1)$ 使 $f(x_0) = x_0$.

定理 6.5.15 (绍德尔) 设 X, Y 为 Banach 空间, $U = B(0, \alpha)$ 与 $V = B(0, \beta)$ 分别为 X 与 Y 中的以零元为中心, α, β 为半径的开球, $v: U \times V \rightarrow Y$ 为连续映射. 若 v 满足

$$(1) \quad \|v(x, y_1) - v(x, y_2)\| \leq k \|y_1 - y_2\|,$$

$$0 < k < 1, \quad x \in U, \quad y_1, y_2 \in V;$$

$$(2) \quad \|v(x, 0)\| < \beta(1 - k), \quad x \in U,$$

则存在惟一的连续映射 $f: U \rightarrow V$, 满足

$$f(x) = v(x, f(x)).$$

用熟知的逐次逼近法可证得此定理. 为此, 取 $y_0 = 0, y_n = v(x, y_{n-1})$, 然后证明 $\{y_n\}$ 为柯西序列.

§ 6.6 非线性科学

众所周知, 线性系统满足叠加原理, 对象的整体等于部分之和. 过去近两个世纪, 数学科学的发展为线性系统的研究提供了强有力的分析方法和工具, 包括线性代数、线性微分方程、傅里叶分析、线性泛函和线性算子、随机过程的线性理论等等.

其实, 非线性问题并非近期才被科学家发现的, 从伽利略牛顿时代就提出并研究了非线性问题, 例如, 摆的运动、天体运动等等. 19 世纪科学家们继续了这类问题的研究, 并提出经典力学的两大难题: 刚体定点运动和三体问题. 然而, 由于非线性问题的解决并非易事, 当时的数学工具也远不能满足要求, 因此进展缓慢. 19 世纪末 20 世纪初, 法国数学力学家庞加莱 (Poincaré, H.) 在总结前期工作的基础上, 引进了新的数学概念和方法, 不仅推动了非线性科学的形成和发展, 而且对数学科学也起了巨大的推动作用. 20 世纪六十年代起, 混沌现象的发现引发了科学家们对复杂性问题研究的兴趣, 并且逐渐认识到非线性因素是事物复杂性的集中表现. 近二十余年, 计算机科学日新月异的飞速发展, 使对一个新的数学分支—分形几何的研究也快速发展起来. 此外, 19 世纪就已引起科学家们注意和研究的孤立波现象及相应的孤立子, 也都与非线性现象紧密联系着. 还可举出许多例子, 说明形成非线性科学领域的背景、研究的必要性.

并非所有非线性问题都属非线性科学的范围. 一个共同的理解: 非线性科学研究自然科学的各个不同领域内、各种非线性系统所具有的共性, 这种共性超越了不同学科领域的局限性, 而具有普适性、交叉性、综合性、前沿性. 科学家发现, 非线性现象的共性体现在: 混沌、分形、孤立子.

近期, 非线性科学研究的主要内容有: 保守系统与耗散系统的混沌、分形数学与分形物理、孤立子与拟序结构、动力系统的时间反演问题、自组织与耗散结

构、随机非线性微分方程、湍流、神经网络系统、复杂性探索。

展望 21 世纪,科学技术必将更加飞速发展,非线性科学将成为科学舞台上的主角.分形,是非线性现象的共性之一,是非线性科学的主题,我们对它作一些简单介绍.

分形分析的特点是:所研究的对象是传统数学视为不规则的、经典数学方法无力解决的、但在自然界和各科学领域中又大量存在着的现象、图形和函数关系(例如,图 6.6.1、6.6.2、6.6.3、6.6.4).

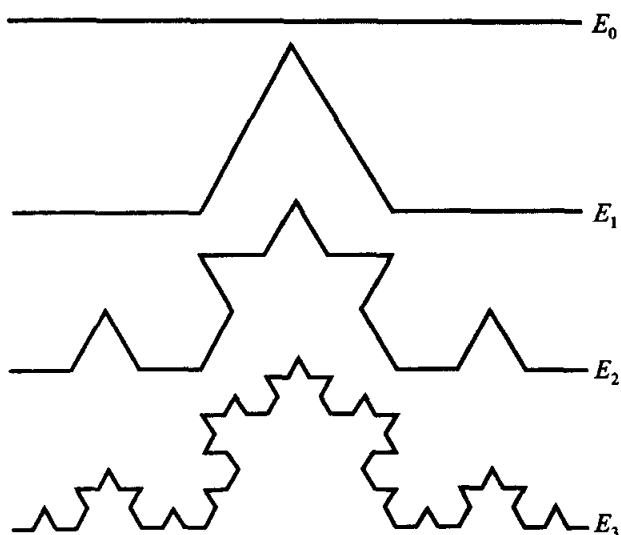


图 6.6.1

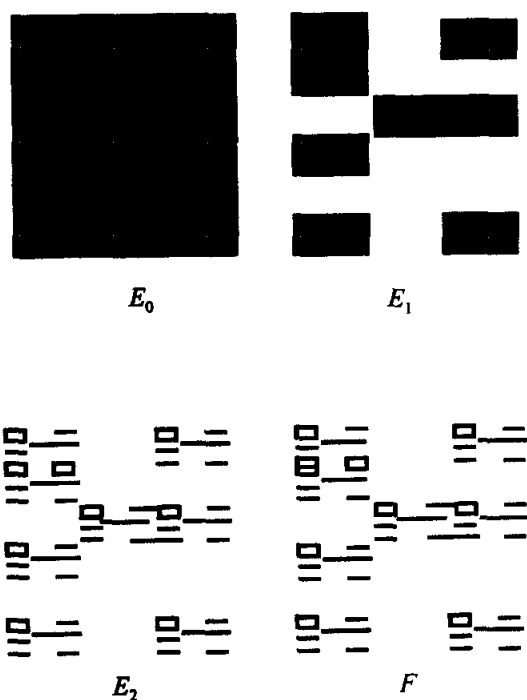


图 6.6.2

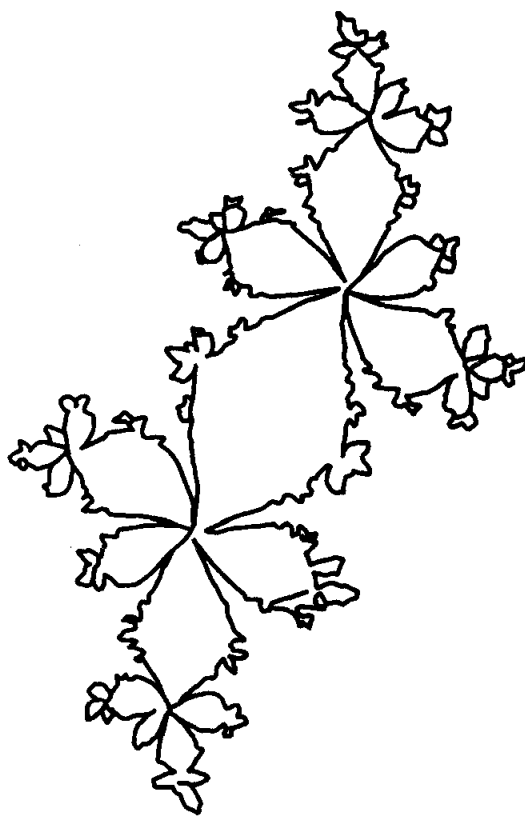


图 6.6.3

自然界中出现的诸如云层的边界,远山的轮廓,飘零的雪花,弯曲的海岸等“不规则”的几何形体,难以用经典几何中的直线、光滑曲线、光滑曲面来描述.同时,不同类型的不规则几何对象常常出现在自然科学的不同领域中:如非线性问

题所呈现的吸引子,滚滚大河中的湍流,物理学中临界现象与相变,化学中酶与蛋白质的结构,生物学中细胞的生长,工程技术中的信号处理、噪声分析……。长期以来,科学家用各种方法试图将它们纳入经典几何的框架中研究,但却发现,由此导出的模型即使在近似的情形,也很难处理所接触到的实际情形.另一方面,人们已注意到不规则集合往往能提供许多自然现象的更好的描述.20世纪八十年代初,由美籍数学家曼德布罗特(Mandelbrot, B. B.)所创立的分形几何(Fractal)提供了研究这类不规则几何对象的新的思想,方法和技巧.特别在近年来,这一新兴学科在数学、物理、化学、生物、医学、地质、材料、工程技术……诸学科中已获得成功的应用,并且促进了分形几何的深入发展.特别应当指出的是分形几何的诞生对整个科学的发展有极为重要的意义,它与非线性动力学的作用令人瞩目:后者涉及运动的非线性确定方程的一般普适行为,前者则是研究自相似或自仿射对象的几何以及该几何上的动力学.两者均已应用到一系列深刻的交叉学科问题中.

什么是分形几何(或更一般些,分形分析)?什么是分形维数?它们的意义是什么?我们作一个简要的描述.

分形几何的研究对象是一类不规则的几何形体,一方面,这类几何对象不能用经典几何来处理;另一方面,它们又具有某些能掌握的“规律性”,并适合于应用的目的.为能较清楚地阐述上述意思,我们考察一个典型的分形:柯克(Koch)曲线(图 6.6.1、6.6.4).

设 E_0 为单位区间,以 E_0 的中间三分之一线段为底,向上作等边三角形,然后去掉该底(保留端点).由此得到的四条线段组成的图形记为 E_1 ,对 E_1 的每一边重复上述过程,所得到的折线多边形记为 E_2 ;应用同样的方式,我们从 E_{k-1} 得到 E_k .当 k 趋于无穷时,折线多边形序列 E_k 趋于一极限曲线 E ,称为柯克曲线(图 6.6.1).

现在我们分析曲线 E 的性质.

(1) 曲线 E 具有“细结构”,亦即它包含任意小尺度下的细节:不管取多么小的尺度,60度的尖角仍然出现,只是边长相应减小.这个事实表明,曲线 E 的特征性质不随尺度的减小而消失.

(2) 曲线 E 难于用经典的方法刻画,从整体上看,它既不是满足某些简单几何条件的点的轨迹,亦不能作为任一简单方程的解的集合;从局部上看,它不能通过导数来描述曲线的变化率(事实上,曲线 E 上每点均无导数).

(3) 用通常的长度、面积、体积的测量方法去测量曲线 E ,得出曲线的长度为无穷大,而面积为零.从而我们不能用通常的测度来量度它的“大小”.

(4) 曲线 E 具有局部与整体的对称:它由四个与 E 相似的部分组成,其相

似因子为 $\frac{1}{4}$; 而每部分由四个更小的但仍与 E 相似的, 其相似因子为 $\frac{1}{4^{\frac{1}{2}}}$ 的部分组成……. 上述对称性亦称为自相似性. 尽管 E 具有复杂的细结构, 但它的定义非常直接. 特别, E 可以由简单的递归方式生成, 而且, 它的逐阶迭代 E_k 给出 E 的越来越好的近似.

我们看到, 性质(2)和(3)反映了柯克曲线的“不规则”性, 而性质(1)、(4)则给出了柯克曲线某些“规则”的性质. 一般说来, 我们所讨论的分形集合都具有前述的某些性质或是它们的变形. 因此, 我们的研究对象既从本质上有别于经典几何, 同时亦排除对那些毫无规则可遵循的几何形体的研究.

曼德布罗特刻画自然界中的分形是通过下述三个要素, 即: 形(form)、机遇(chance)与维数(dimension). 人们可以毫无困难地区别一片枫叶和一片柳叶, 因为它们“形”不一样; 人们亦认为两片大小不同的枫叶是“一样”的, 因为它们有相同的“形”. 前面提到的自相似性提供了将分形从形上进行分类的一个工具. 但迄今为止, 对分形的“形”的刻画的研究极为缓慢. 从图 6.6.1 与图 6.6.4 可以看到, 图 6.6.4 中的曲线与海岸线非常相似, 而图 6.6.1 与海岸线相比较则过于“规则”, 这种差异的产生是由于实际的海岸线受到自然界随机因素的影响而显示出更为“复杂”, 可是更为自然. 但这种复杂性究竟如何刻画与量度, 就是维数所应解决的问题了.

维数的研究在分形分析中起着极为重要的作用, 目前关于维数已经有较为丰富和深入的成果. 我们介绍几种常用的维数.

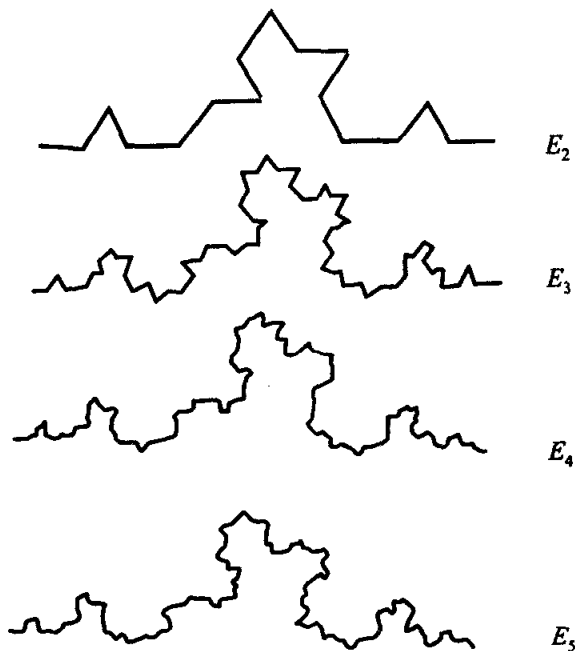


图 6.6.4

豪斯道夫(Hausdorff)维数

柯克曲线的性质(3)指出:柯克曲线的长度为无穷而面积为零.注意到我们测量几何图形的长度与面积时,分别用单位长线段与单位面积的正方形来测量,而线段与正方形的欧几里得维数分别为1与2.若用线段来测量正方形,其结果为无穷,说明所用尺度太“细”;反之,若用正方形为尺度来测量线段,其结果为零,亦即所用的尺度太“粗”.因此,在测量平面中的集合时,其测量结果与所采用尺度有关.特别,经典的几何对象的测量只容许用整数维的尺度.对于柯克曲线,一维尺度太“细”,二维尺度则太“粗”.因此,我们可以将柯克曲线想象为维数介于1维与2维间的几何对象,从而应当用非整维的尺度来对它进行测量.基于上述思想,豪斯道夫将传统的维数概念按下述方式推广到一般的非负实数.

设 A 为 n 维欧几里得空间 $\mathbf{R}^n$ 的一个子集. $s \geq 0$ 为非负实数.为测量 A , 一个方式是对于 $\delta > 0$, 我们用直径不大于 δ 的球来覆盖 A , 为使得覆盖中重叠产生的损失最小, 我们应取最经济的覆盖, 而相应采用 s 维尺度. 随着 δ 减小, 所得的测量愈精确, 这是 s 维豪斯道夫测度的基本思想. 它的定义如下:

设 $\{U_i\}_{i \geq 1}$ 为 $\mathbf{R}^n$ 中可数多个子集, 若 $\bigcup_{i \geq 1} U_i \supset A$, 且对任意 $i, |U_i| \leq \delta$, 则称 $\{U_i\}_{i \geq 1}$ 为 A 的一个 δ -覆盖. 令

$$\mathcal{G}_\delta^s(A) = \inf \sum_i |U_i|^s,$$

其中 $\inf$ 表示对 A 的所有 δ 覆盖取下确界, 令

$$\mathcal{G}(A) = \lim_{\delta \rightarrow 0} \mathcal{G}_\delta^s(A),$$

则 $\mathcal{G}^s(A)$ 称为 A 的 s 维 Hausdorff 测度. 它是一个度量外测度, 亦即, 仅满足“次可加性”: $\mu(\bigcup_i A_i) \leq \sum_i \mu(A_i)$.

进一步可证明, 对任一集 $A \subset \mathbf{R}^n$, 存在惟一的非负实数 $D_H(A)$, 满足下述性质:

若 $0 \leq s < D_H(A)$, 则 $\mathcal{G}^s(A) = \infty$;

若 $D_H(A) < s$, 则 $\mathcal{G}^s(A) = 0$,

值 $D_H(A)$ 称为 A 的豪斯道夫维数. 结合前面的分析, $D_H(A)$ 是测量 A 的一个合适的尺度, 若所用尺度大于它, 则过粗, 否则过于细. 因此, D_H 刻画了集合的度量性质与复杂程度, 从几何的角度看, 则表征了集合的填充空间的能力.

注 $\mathcal{G}^{D_H(A)}(A)$ 可能为一正有限数, 在这种情形, 称 A 为 s 集, 这类集合有很好的几何性质与分析性质. 它亦可取零或无穷, 对于这种情形, 我们还可考虑更细的尺度.

豪斯道夫维数具有下述基本性质:

(1) 单调性: 若 $A \subset B$, 则 $D_H(A) \leq D_H(B)$;

(2) 可列平稳性: $D_H(\bigcup_i A_i) = \sup_i D_H(A_i)$.

一个直接推论是可列集的豪斯道夫维数为零.

(3) 利普希茨不变性: 若 f 为双利普希茨变换, 即存在两个正数 c_1, c_2 , 使得

$$c_1 |x - y| \leq |f(x) - f(y)| \leq c_2 |x - y|.$$

则 $D_H(f(A)) = D_H(A)$. 从而上述等式对平移, 旋转及相似变换均成立.

(4) 若集 A 的 n 维勒贝格测度大于零, 则 $D_H(A) = n$, 特别, $\mathbf{R}^n$ 中的任一开集的维数 $D_H(A) = n$.

从上面的讨论, 我们可以看到豪斯道夫维数至少有下列意义:

① 测量几何对象适当的尺度;

② 充斥空间的能力;

③ 利普希茨映射的不变量 (因此, 两个集合的豪斯道夫维数不等, 它们不能通过一个利普希茨映射作相互变换).

帕克 (Packing) 维数

在维数定义中, 我们是通过覆盖的方式来测量一个集合. 作为一种“对偶”的方式, 也可以采用填充 (Packing) 的方式, 亦即利用互不相交的球来得到一个集合的近似测量. 具体定义如下: 设 $A \subset \mathbf{R}^n, s \geq 0, \delta > 0$. 令

$\mathcal{P}_\delta^s(A) = \sup \left\{ \sum_i |U_i|^s : \{U_i\} \text{ 为中心在 } A \text{ 中、互不相交、直径不大于 } \delta \text{ 的球} \right\},$

$$\mathcal{P}_0^s(A) = \lim_{\delta \rightarrow 0} \mathcal{P}_\delta^s(A).$$

此时, $\mathcal{P}_0^s(A)$ 不是一个外测度, 但可以通过下述方式成为一个外测度:

$$\mathcal{P}^s(A) = \inf \left\{ \sum_i \mathcal{P}_0^s(A_i), F \subset \bigcup_i A_i \right\}.$$

$\mathcal{P}^s(A)$ 称为 A 的 s 维帕克测度, 对应的帕克维数 $D_P(A)$ 定义为:

$$D_P(A) = \sup \{s : \mathcal{P}^s(A) = \infty\} = \inf \{s : \mathcal{P}^s(A) = 0\}.$$

它具有豪斯道夫维数所具有的性质 (1) — (4). 并且

$$D_H(A) \leq D_P(A).$$

盒维数

这一维数在理论及应用上均有重要意义, 它首先由布里冈 (Bouligand) 于 1929 年引入, 亦称计盒维数、熵维数、容量维数、度量维数、对数密度、信息维数、柯尔莫果洛夫 (Kolmogorov) 熵.

设 $A \subset \mathbf{R}^n$ 为有界非空子集, 令 $N_\delta(A)$ 为半径为 δ 的覆盖 A 的球的最小个数, 则 A 的上下 Bouligand 维数分别定义为

$$\Delta(A) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta},$$

$$\delta(A) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta},$$

若 $\Delta(A) = \delta(A)$, 则称这一公共值为 A 的布里冈维数, 并记为 $D_B(A)$. 布里冈维数除具有前述豪斯道夫维数的性质(1)、(3)、(4)外, 还具有下述性质:

(1) 有限平稳性 $D_B(A_1 \cup A_2) = \max\{D_B(A_1), D_B(A_2)\}$.

(2) $D_B(A) = D_B(\bar{A})$, $\bar{A}$ 表示 A 的闭包.

注 性质 2) 指出, 可列集的布里冈维数可不为零, 事实上 $D_B(Q) = 1$, Q 为有理数集.

(3) $D_B \geq D_P \geq D_H \geq D_T$. (D_T 为拓扑维数)

注 在许多情形下它们不相等. 若 $\Delta(A) \neq \delta(A)$, 在对布里冈维数作数值计算的时候, 要特别注意这一点.

这一性质亦说明布里冈维数较豪斯道夫维数更“细”.

布里冈维数尚有下列等价定义.

若我们用 $N_\delta^i(A)$ ($i = 1, 2, 3, 4$) 分别记: 半径至多为 δ 的可覆盖 A 的集的数目、边长为 δ 的可覆盖 A 的 n 维立方体的最小数目、与 A 相交的 δ -平行网中的立方体的数目、半径为 δ 互不相交的中心在 A 上的球的最大数目. 则

$$D_B(A) = \lim_{\delta \rightarrow 0} \frac{\log N_\delta^i(A)}{-\log \delta} \quad (i = 1, 2, 3, 4).$$

在应用中我们可根据需要, 来选择上述不同定义.

还有许多常用的维数, 如斯坦豪斯 (Steinhaus) 维数、贝欣科维奇 - 泰乐 (Besicovitch - Taylor) 维数、相似维数、测度维数、连通性维数等, 各有各的具体意义, 用来刻画研究对象的各种性质. 在此就不多介绍了.

附录

Fourier 变换表

表 1 Fourier 变换

$f(x)$	$f^\wedge(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-iuv} du$
运算规则	
$\alpha f(x) + \beta g(x) (\alpha, \beta \in \mathbf{C})$	$\alpha f^\wedge(v) + \beta g^\wedge(v)$
$f(x+h) \quad (h \in \mathbf{R})$	$e^{ihv} f^\wedge(v)$
$e^{-ihx} f(x) \quad (h \in \mathbf{R})$	$f^\wedge(v+h)$
$\frac{\rho f(\rho x)}{f(-x)} \quad (\rho > 0)$	$\frac{f^\wedge(v/\rho)}{f^\wedge(v)}$
$f(x) = f(-x)$	$f^\wedge(v) = f^\wedge(-v)$ $= \sqrt{\frac{2}{\pi}} \int_0^\infty f(u) \cos vud u$
$f(x) = -f(-x)$	$f^\wedge(v) = -f^\wedge(-v)$ $= -i \sqrt{\frac{2}{\pi}} \int_0^\infty f(u) \sin vud u$
$f(-x)$	$f^\wedge(-v)$
$f(hx) \quad (h \in \mathbf{R}, h \neq 0)$	$ h ^{-1} f^\wedge(v/h)$
$(f * g)(x)$	$f^\wedge(v) g^\wedge(v)$
$f^\wedge(x)$	$f(-v)$
$\left(\frac{d}{dx}\right)^r f(x)$	$(iv)^r f^\wedge(v)$
$x^r f(x)$	$\left(i \frac{d}{dv}\right)^r f^\wedge(v)$
$f^-(x) \quad (f \text{ 的 Hilbert 变换})$	$\{-i \operatorname{sgn} v\} f^\wedge(v)$
$\{-i \operatorname{sgn} x\} f(x)$	$f^-(-v)$

续表

$f(x)$	$f^\wedge(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-iuv} du$
基本函数	
$\sqrt{2\pi} \chi_{[-1,0]}(x)$	$\frac{e^{iv} - 1}{iv}$
$\sqrt{\frac{2}{\pi}} \frac{\sin ax}{x} \quad (a > 0)$	$\begin{cases} 1, & v < a \\ \frac{1}{2}, & v = a \\ 0, & v > a \end{cases}$
$\frac{1}{\sqrt{2\pi}} \left[\frac{\sin x/2}{x/2} \right]^2$	$\begin{cases} 1 - v , & v \leq 1 \\ 0, & v > 1 \end{cases}$
$\frac{3}{\sqrt{8\pi}} \left[\frac{\sin(x/2)}{x/2} \right]^4$	$\begin{cases} 1 - \frac{1}{2}v^2 + \frac{3}{4} v ^3, & v \leq 1 \\ (2 - v)^3/4, & 1 < v \leq 2 \\ 0, & v > 2 \end{cases}$
$\sqrt{\frac{2}{\pi}} (1+x^2)^{-1}$	$e^{- v }$
$(1/\sqrt{2}) e^{-x^2/4}$	e^{-v^2}
$H_n(x) \equiv (-1)^n e^{x^2/2} \frac{d^n}{dx^n} (e^{-x^2})$	$(-i)^n H_n(v)$
$\begin{cases} \sqrt{2\pi} e^{-x}, & 0 \leq x < \infty \\ 0, & -\infty < x < 0 \end{cases}$	$(1+iv)^{-1}$
$ x ^{-s} \quad (0 < \operatorname{Re} s < 1)$	$\sqrt{\frac{2}{\pi}} \frac{\Gamma(1-s)}{ v ^{1-s}} \sin \frac{\pi s}{2}$
$x^{-s} \chi_{(0,\infty)}(x) \quad (0 < \operatorname{Re} s < 1)$	$(\Gamma(s)/\sqrt{2\pi})(iv)^{-(1-s)}$
$\sqrt{\frac{2}{\pi}} \Gamma(1+\alpha) \operatorname{Re}(1-ix)^{-1-\alpha}$	$ v ^\alpha e^{- v }$
$\sqrt{\frac{2}{\pi}} \Gamma(1+\alpha) \operatorname{Im}(1-ix)^{-1-\alpha}$	$ -i \operatorname{sgn} v v ^\alpha e^{- v }$
$\frac{e^{-ihx} - 1}{ x } \quad (\alpha > 0)$	$\sqrt{\frac{2}{\pi}} \log \left \frac{v}{h+v} \right $
$\frac{2^{(2-x)/2}}{\Gamma(x/2)} x ^{(x-1)/2} \chi_{(1-x)/2}$	$(1+v^2)^{-x/2}$
$\times (x) \quad (x > 0)$	
$2^\lambda \Gamma(1+\lambda) \frac{J_{1/2+\lambda}(x)}{ x ^{1/2+\lambda}}$	$\begin{cases} (1-v^2)^\lambda, & v \leq 1 \\ 0, & v > 1 \end{cases}$
$(\lambda > 0)$	

续表

$f(x)$	$f^\wedge(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-iuv} du$
基本函数	
$\frac{1}{ x+h ^{1-a}} - \frac{1}{ x ^{1-a}}$ ($0 < a < 1$)	$\frac{A_\epsilon(a) e^{ihv} - 1}{\sqrt{2\pi} v ^a}$
$\frac{1}{x} \log \left(\frac{a+x}{a-x} \right)^2$ ($a > 0$)	$\sqrt{\frac{8}{\pi}} \int_{av}^{\infty} \frac{\sin u}{u} du$
$\sqrt{\frac{2}{\pi}} \int_{ x }^{\infty} \frac{\sin u}{u^2} du$	$\begin{cases} 1, & v \leq 1 \\ v ^{-1}, & v > 1 \end{cases}$
$\frac{1}{\sqrt{2\pi}} \log \frac{1+x^2}{x^2}$	$\frac{1 - e^{- v }}{ v }$
$\sqrt{2} \left(e^{-x^2/4} - x \int_{ x /2}^{\infty} e^{-u^2} du \right)$	$\frac{1 - e^{-v^2}}{v^2}$

表 2 有限 Fourier 变换

$f(x)$	$f^\wedge(k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(u) e^{-iku} du$
运算规则	
$\alpha f(x) + \beta g(x)$ ($\alpha, \beta \in \mathbb{C}$)	$\alpha f^\wedge(k) + \beta g^\wedge(k)$
$f(x+h)$ ($h \in \mathbb{R}$)	$e^{ikh} f^\wedge(k)$
$\frac{e^{-ijx} f(x)}{f(-x)}$ ($j \in \mathbb{Z}$)	$\frac{f^\wedge(k+j)}{f^\wedge(k)}$
$f(x) = f(-x)$	$f^\wedge(k) = f^\wedge(-k)$ $= \frac{1}{\pi} \int_0^\pi f(u) \cos ku du$
$f(x) = -f(-x)$	$f^\wedge(k) = -f^\wedge(-k)$ $= \frac{-i}{\pi} \int_0^\pi f(u) \sin ku du$
$(f * g)(x)$	$f^\wedge(k) g^\wedge(k)$
$\left(\frac{d}{dx} \right)^r f(x)$	$(ik)^r f^\wedge(k)$
$f^-(x)$	$\{-i \operatorname{sgn} k\} f^\wedge(k)$
$\sqrt{2\pi} \sum_{j=-\infty}^{\infty} \rho \chi(\rho(x+2j\pi))$	$\chi^\wedge(k/\rho)$

续表

$f(x)$	$f^{\wedge}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(u) e^{-iku} du$
基本函数	
$\sum_{j=-n}^n c_j e^{ijx}$	$\begin{cases} c_k, & k \leq n \\ 0, & k > n \end{cases}$
$2\pi \chi_{[-1,0]}(x) \quad x \in [-\pi, \pi]$	$\frac{e^{ik} - 1}{ik}$
$\frac{\sin((2n+1)x/2)}{\sin(x/2)}$	$\begin{cases} 1, & k \leq n \\ 0, & k > n \end{cases}$
$\frac{1}{n+1} \left[\frac{\sin((n+1)x/2)}{\sin(x/2)} \right]^2$	$\begin{cases} 1 - \frac{ k }{n+1}, & k \leq n \\ 0, & k > n \end{cases}$
$\frac{3}{n(2n^2+1)} \left[\frac{\sin(nx/2)}{\sin(x/2)} \right]^4$	$\frac{1}{2n(2n^2+1)} \begin{cases} 3 k ^3 - 6nk^2 - 3 k + 4n^3 + 2n, & k \leq n \\ - k ^3 + 6nk^2 - (12n^2 - 1) k + 8n^3 - 2n, & n < k \leq 2n-2 \\ 0, & k > 2n-2 \end{cases}$

习题答案

第四章 级数—函数的分析与研究 III

§ 4.1

- (1) 收敛, 3; (2) 收敛, $\frac{3}{2}$; (3) 收敛, $\frac{1}{3}$; (4) 收敛, $1 - \sqrt{2}$.
- (1) 发散; (2) 发散; (3) 发散; (4) 发散; (5) 发散; (6) 发散; (7) 收敛; (8) 收敛; (9) 发散; (10) 收敛; (11) 发散; (12) 收敛; (13) 收敛; (14) 发散.
- 能.
- 考虑级数 $1 - 1 + 1 - 1 + \cdots + (-1)^{n+1} + \cdots$.
- (1) 条件收敛; (2) 发散; (3) 发散; (4) 绝对收敛; (5) 条件收敛; (6) 绝对收敛; (7) 条件收敛; (8) 绝对收敛.
- 不一定.

10. 不一定

14. 1.

16. (1) 收敛; (2) 发散.

§ 4.2

1. (1) $(-\infty, -1) \cup (1, +\infty)$; (2) $(-1, 1)$; (3) $(-\infty, +\infty)$; (4) $(-\frac{1}{2}, +\infty)$; (5) $(-1, 1)$.

2. (1) 一致收敛; (2) 一致收敛; (3) 一致收敛; (4) 一致收敛 (本题中 $x=0$ 时, $n=1$ 时, 级数无意义).

4 (1) $(-1, 1)$; (2) $(-1, 1)$, (3) 收敛域为点 $x=0$; (4) $(-2, -2)$; (5) $(0, 2]$; (6) $(-\sqrt{2}, \sqrt{2})$; (7) $[-\frac{1}{3}, \frac{1}{3}]$; (8) $(-1, 1)$.

5. 是.

8. (1) 1; (2) e.

9. (1) $\frac{1}{(1-x)^2}, x \in (-1, 1)$; (2) $\frac{2x}{(1-x)^3}, x \in (-1, 1)$;

(3) $\begin{cases} \frac{1}{x} \ln(1-x) + 1, & x \in [-1, 1) \cup (0, 1); \\ 0, & x = 0 \\ 1, & x = 1; \end{cases}$ (4) $\frac{2x-3}{1-x}, x \in (-1, 1)$; (5) $-\frac{x}{1-x^2}, x \in$

$(-1, 1)$; (6) $\frac{1}{2\sqrt{2}} \ln \left| \frac{\sqrt{2}+x}{\sqrt{2}-1} \right|, x \in (-\sqrt{2}, \sqrt{2})$.

12. (1) $\frac{1}{4} \sum_{n=0}^{\infty} \left[(-1)^n - \left(\frac{1}{3} \right)^n \right] x^n, x \in (-1, 1)$; (2) $\frac{1}{3} \sum_{n=0}^{\infty} [(-1)^n 2^n - 1] x^{n-1}, x \in (-1, 1)$; (3) $\frac{2}{\sqrt{3}} \sum_{n=0}^{\infty} \left(\sin \frac{2(n+1)}{3} \pi \right) x^n, x \in (-1, 1)$; (4) $-\frac{1}{2} \sum_{n=1}^{\infty} (-1)^n \frac{(2x)^{2n}}{(2n)!}, x \in (-\infty, +\infty)$; (5) $\sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}, x \in (-\infty, +\infty)$, (6) $\sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, x \in (-\infty, +\infty)$,

(7) $1 + \sum_{n=2}^{\infty} \frac{n-1}{n!} x^n, x \in (-\infty, +\infty)$; (7) $x + \sum_{n=1}^{\infty} \frac{(2n-1)!!}{2^n \cdot n! (2n+1)} x^{2n+1}, x \in [-1, 1]$.

(9) $1 + \sum_{n=1}^{\infty} \frac{(-1)^n (2n-1)!!}{2^n \cdot n!} x^{2n}, x \in (-1, 1)$; (注: $2^n \cdot n! = (2n)!!$); (10) $\sum_{n=0}^{\infty} (-1)^n \cdot$

$\frac{x^{2n+1}}{n! (2n+1)}, x \in (-\infty, +\infty)$; (11) $\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)! (2n+1)}, x \in (-\infty, +\infty)$; (12)

$\sum_{n=0}^{\infty} \frac{1}{n!} x^n, x \in (-1, 1)$;

14. $\sum_{n=1}^{\infty} \frac{nx^{n-1}}{(n+1)!}, x \in (-\infty, +\infty)$; 1.

16. (1) 1.058; (2) 0.497.

17. (1) 1.000 10; (2) 2.990 7.

18. 0.693 2.

§ 4.3

3. 1.

$$4. (1) \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin(2n-1)x}{2n-1}; (2) \frac{2}{\pi} + \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{1-4n^2} \cos 2nx; (3) \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(2n-1)x}{2n-1}.$$

$$5. f(x) \sim \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1 - \cos \frac{n\pi}{2}}{n} \sin nx, S(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1 - \cos \frac{n\pi}{2}}{n} \sin nx, x \in \left(0, \frac{\pi}{2}\right) \cup \left(\frac{\pi}{2}, \pi\right).$$

$$6. f(x) \sim \frac{1}{2} + \frac{1}{4} \cos x + \sum_{k=2}^{\infty} \frac{1}{1-k^2} \cos kx, S(x) = \frac{1}{2} + \frac{1}{4} \cos x + \sum_{n=2}^{\infty} \frac{1}{1-n^2} \cos nx, x \in [-\pi, \pi].$$

$$7. f(x) \sim \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^2} \sin \frac{(2n-1)\pi}{2} x, S(x) = \frac{8}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^2} \sin \frac{(2n-1)\pi}{2} x.$$

$$8. f(x) \sim \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \frac{(-1)^n}{a-in} (e^{ax} - e^{-ax}) e^{inx}, S(x) = \begin{cases} e^{ax}, & x \in (-\pi, \pi), \\ \frac{1}{2}(e^{ax} + e^{-ax}), & x = \pm\pi. \end{cases}$$

第五章 微分方程与傅里叶分析

§ 5.1

1. (1) $y = x^3 e^{-x} + c e^{-x} \cdot x^{-1}$; (2) $y = (x-1)^4 + c(x-1)$; (3) $y = \sin x + c \cos x$; (4) $y = e^x (\ln |x| + c)$.

2. (1) $y = (1-x-ce^{-x})^{-\frac{1}{6}}$, $y=0$ 也是方程的根; (2) $y = \left(-\frac{1}{6} + ce^{3x^2}\right)^{-\frac{1}{3}}$, $y=0$ 也是方程的根.

$$3. x^{-2} = 1 - y^2 + ce^{-y^2}.$$

$$4. y(x) = (4e^{\frac{1}{2}x^2} - x^2 - 2)^{-1} + x.$$

$$5. y = e^{2x^2} \cdot \left(c - \int e^{2x^2} dx\right)^{-1} + 2x, y = 2x \text{ 也是方程的解.}$$

$$6. (1) y = c_1 e^{-3x} + c_2 e^{-4x}; (2) y = (c_1 + c_2 x) e^{2x}; (3) y = c_1 + c_2 x + c_3 e^{-x} + c_4 e^x; (4)$$

$$y = e^{\frac{x}{\sqrt{2}}} (c_1 \cos \frac{1}{\sqrt{2}} x + c_2 \sin \frac{1}{\sqrt{2}} x) + e^{-\frac{x}{\sqrt{2}}} (c_3 \cos \frac{1}{\sqrt{2}} x + c_4 \sin \frac{1}{\sqrt{2}} x); (5) y = c_1 e^x + c_2 e^{\frac{x}{2}} \cos \frac{\sqrt{3}}{2} x$$

$$+ c_3 e^{\frac{x}{2}} \sin \frac{\sqrt{3}}{2} x; (6) y = c_1 e^{\sqrt{2}x} + c_2 e^{-\sqrt{2}x} + c_3 e^x + c_4 e^{-x} + c_5 \cos x + c_6 \sin x.$$

$$7. (1) 5e^x - 3e^{2x}; (2) y = 1 - \sin x.$$

$$8. (1) c_1 e^{-x} + c_2 e^{2x} - \frac{3}{2} e^x; (2) c_1 e^{2x} + c_2 x e^{2x} + \frac{1}{2} x^2 e^{2x}; (3) c_1' \cos x + c_2' \sin x + \frac{1}{2} x \cdot \sin x - \frac{1}{3} \sin 2x; (4) c_1' e^{-3x} \cos x + c_2' e^{-3x} \sin x - \frac{x}{2} e^{-3x} \cos x; (5) c_1 e^{-\frac{1}{3}x} \cos \frac{\sqrt{2}}{3} x + c_2 e^{-\frac{1}{3}x} \cdot \sin \frac{\sqrt{2}}{3} x + 2x^2 - 7x + 3.$$

$$9. (1) y = -\frac{1}{6}(1-x^2)^{\frac{3}{2}} - \frac{1}{2}x \arccos x + \frac{1}{2}\sqrt{1-x^2} + c_3 x + c_3'; (2) y = -c_1 e^{-\frac{1}{2}x^2} + x + c_2; (3) (x-c')^2 + (y-c)^2 = a^2; (4) y = c \text{ 或 } y = \frac{c - c_2 e^{cx}}{1 - c_2 e^{cx}} \text{ 或 } y = \frac{-1}{x+c_2}; (5) 1 - c_1 y^2 = (\pm x + c_2)^2 c_1; (6) \frac{c_3}{2} y^2 + c_4 y = -x + c_5; y = c_2 e^{x^3 + cx} (c_2 \text{ 可以为 } 0); (8) y = c_2 x^{\frac{8}{3}} e^{\frac{5}{3}x^5} (c_2 \text{ 可以为 } 0); (9) y = -\frac{1}{c_1} [1 + \arctan(\cot x) \cdot \cot x] + c_2 \cot x; (10) xy + \frac{1}{3}y^3 = \frac{c_1}{2}x^2 + c_2 x + c_3; (11) y^2 = c_1 x^2 + c_2 x; (12) y = c_1 \cot x + c_2(1 - x \cot x); (13) y^2 = \frac{1}{6}x^4 - c_1 x^{-2} + 2c_2.$$

$$11. 60 \text{ 分钟}.$$

$$12. xy = \pm s.$$

§ 5.2

$$1. z = \Phi(x^2 + y^2), \Phi \text{ 为任意可微函数}.$$

$$2. F\left(y, \frac{x}{y} - \ln|z|\right) = 0, F \text{ 为任意可微函数}, z \equiv 0 \text{ 也是解或 } F(y, e^{-\frac{x}{y}}z) = 0, F \text{ 为任意可微函数}.$$

$$3. u = F(x-y, x-z) \text{ 或 } u = F(x-y, y-z), F \text{ 为任意可微函数}.$$

$$4. u = F\left(y, \frac{1}{2}\ln|x^2 + y| - y\ln|y + z|\right), F \text{ 为任意可微函数}.$$

$$6. x^2 + y^2 + z^2 = y\Phi\left(\frac{2}{y}\right).$$

§ 5.3

$$1. \tilde{f} = \frac{2}{\omega} \sin \omega.$$

第六章 空间的结构与现代数学

§ 6.2

$$4. (1) \lambda_1 = 1, \lambda_2 = -2, \alpha_1 = k_1(3, -6, 20)^T, \alpha_2 = k(0, 0, 1)^T, \text{ 其中 } k_1, k_2 \text{ 为任意非零常}$$

数; (2) $\lambda_1 = 2, \lambda_2 = -7, \alpha_1 = k_1(-2, 1, 0)^T + k_2(2, 0, 1)^T, \alpha_2 = k(1, 2, -2)^T$, 其中 k_1, k_2 为不全为零的任意常数, k 为非零常数.

$$9. (1) \begin{bmatrix} 1 & 2 \times 5^{n-1}(1 - (-1)^n) & 5^{n-1}(4 + (-1)^n) - 1 \\ 0 & 5^{n-1}(1 + 4(-1)^n) & 2 \times 5^{n-1}(1 - (-1)^n) \\ 0 & 2 \times 5^{n+1}(1 - (-1)^n) & 5^{n+1}(4 + (-1)^n) \end{bmatrix}; (2) \frac{1}{3}(5^n - (-1)^n)J_3 -$$

$(-1)^n I_3$, 这里 I_3 为三阶单位矩阵, J_3 为三阶所有元素均为 1 的“全 1 矩阵”.

$$10. (1) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}; (2) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix}.$$

$$11. (1) a = 3, b = 5; (2) U = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

$$17. (1) \frac{1}{\sqrt{3}}(1, 1, 1)^T, \frac{1}{\sqrt{2}}(1, 0, -1)^T, \frac{1}{\sqrt{6}}(-1, 2, -1)^T; (2) \frac{1}{2}(1, 1, 1, 1)^T,$$

$$\frac{1}{2}(1, -1, 1, -1)^T, \frac{1}{2}(1, 1, -1, -1)^T, \frac{1}{2}(1, -1, -1, 1)^T.$$

$$18. (1) \frac{1}{3\sqrt{5}} \begin{bmatrix} 3 & 4 & 2\sqrt{5} \\ -6 & 2 & \sqrt{5} \\ 0 & -5 & 2\sqrt{5} \end{bmatrix}, \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}; (2) \frac{1}{3} \begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ -2 & 1 & 2 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

$$19. (1) \begin{cases} x_1 = z_1 - \frac{1}{2}z_2 + \frac{5}{6}z_3, \\ x_2 = \frac{1}{2}z_2 - \frac{1}{6}z_3, \\ x_3 = \frac{1}{3}z_3, \end{cases} \quad f = z_1^2 + z_2^2 - z_3^2; (2) \begin{cases} x_1 = y_1 + y_2 - y_3, \\ x_2 = y_1 - y_2 - y_3, \\ x_3 = y_3, \end{cases} \quad f = y_1^2 - y_2^2 - y_3^2.$$

20. (1) 正定; (2) 负定.

21. $|t| < \sqrt{2}$.

$$22. (1) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}, 2u^2 + 4v^2 + 4w^2 = 1;$$

$$(2) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}, 2u^2 - v^2 + 3w^2 = 1.$$