基础代数

（第二卷）

席南华　编著

科学出版社

基 础 代 数

（第二卷）

席南华　编著

科 学 出 版 社

北 京

内 容 简 介

本书是作者为中国科学院大学一年级本科生讲授线性代数课程时，根据作者本人授课的课堂录音和学生的课堂笔记整理修订完善而成的. 作者吸收借鉴了柯斯特利金《代数学引论》的优点和框架，在内容的选取和组织，贯穿内容的观点等方面都有特色. 本书分为三卷，本册为第二卷，主要内容包括：向量空间，线性算子，内积空间，仿射空间与欧几里得仿射空间，二次曲面，张量等，每章节附有适当的习题，可供读者巩固练习使用.

本书可供数学类各专业及相关专业的本科生、研究生和教师使用，也可作为数学爱好者的参考读物.

图书在版编目（CIP）数据

基础代数. 第二卷/席南华编著. —北京：科学出版社，2018.1

ISBN 978-7-03-056033-9

Ⅰ. ①基… Ⅱ. ①席… Ⅲ. ①代数 Ⅳ. ①O15

中国版本图书馆 CIP 数据核字（2017）第 324241 号

责任编辑：张中兴 梁 清／责任校对：李 彤
责任印制：师艳茹／封面设计：迷底书装

科学出版社 出版
北京东黄城根北街 16 号
邮政编码 100717
http://www.sciencep.com

保定市中画美凯印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2018 年 1 月第 一 版 开本：720 × 1000 1/16
2018 年 1 月第一次印刷 印张：21
字数：423 000

定价：49.00 元

（如有印装质量问题，我社负责调换）

前　　言

本书的写作动因在第一卷的前言中已经说明. 和第一卷一样, 本书基本上沿用了柯斯特利金所著的《代数学引论》的框架和内容, 只是在表述和细节上 (希望) 更符合汉语读者的习惯, 有些地方的处理也和原教材的不一样, 同时, 贯穿内容的观点也时有不同的地方, 习题的安排上有较大的差别. 本书在写作过程中, 还参考了若干其他的中外教科书, 详见书后的参考文献. 本书的习题主要选自这些参考书, 还有个别自己加上的习题.

本书根据录音稿和学生的笔记 (主要是杨昊天、陈冰露、梁晨旭、金彦浩、张晓敏等同学的笔记) 整理而成. 本书的初稿于 2017 年春季学期在国科大试用. 曹桂兰老师, 助教陈攀、高剑伟, 2016 级线性代数 II-A1 班的同学们指出了初稿中很多的错误. 在此对以上各位的帮助一并致谢.

席南华

2017 年 6 月于玉泉路

目　　录

第 1 章　向 量 空 间

天地间的向量空间数不清, 我们仅讨论了实数域上的行和列的向量空间 $\mathbb{R}^n$, 它从线性方程组的系数中自然产生. 对其他的域, 也有线性方程组, 所以可以考虑这个域上的行和列的向量空间. 但这仍然有局限, 而且这个局限是不必要的. 注意到以前对 $\mathbb{R}^n$ 的讨论是基于两个运算: 行 (或列) 向量之间的加法, 纯量乘行 (或列) 向量及这两个运算的一些性质, 可以知道这两个运算和那些性质才是本质的. 把它们抽象出来, 就得到一般的向量空间的定义, 于是前面发展的若干概念和方法的适用范围大大扩展. 本章和随后的五章将对一般的向量空间做一些讨论.

1.1　定义与例子

一　**定义 1.1**　称集合 V 是域 K 上的**向量空间**(**或线性空间**), 如果 V 是加法 (交换) 群且 K 中的元素与 V 中的元素可以相乘得到 V 的元素, 相乘具有以下性质:

(1) $(ab)v = a(bv)$;

(2) $a(u+v) = au + av$;

(3) $(a+b)v = av + bv$;

(4) $1v = v$.

其中 a, b 是 K 中的任意元素, 1 是 K 中的乘法单位元, u, v 是 V 中任意元素. 向量空间中的元素称为**向量**, 其中的零元常称为**零向量**. 域 K 中的元素称为**纯量**, 纯量与向量的相乘称为**纯量乘**. 域 K 也常称为 V 的**基域**.

说明　在定义中加号 + 既表示 V 中的加法也表示 K 中的加法, 这样做是方便的, 且不会引起歧义. 讨论抽象的 (或抽象地讨论) 向量空间时, 向量空间的交换群运算一般用 + 表示, 基域的元素与向量相乘的运算符号一般省略, 或用小圆点 $\cdot$ 表示. 在讨论一些具体的向量空间时, 可能需要用其他的记号表示向量空间的加法和基域的元素与向量相乘的运算, 以避免混乱. 例如, 正实数全体 $\mathbb{R}_+$ 在数的乘法下是交换群, 对 $u, v \in \mathbb{R}_+$, $a \in \mathbb{R}$, 定义 $u \oplus v = uv$, $a \odot v = v^a$. 容易验证, 在运算 $\oplus, \odot$ 下 $\mathbb{R}_+$ 是实数域上的向量空间.

另一个例子更有意义. 设 V 是复数域上的向量空间. 我们定义复数域上一个新的向量空间 $\bar{V}$, 称为 V 的**复共轭向量空间**. 作为加法群, 它和 V 是一样的, 但纯量乘不一样, 定义为 $\lambda \odot x = \bar{\lambda}x$, 其中 $\bar{\lambda}$ 是复数 λ 的共轭. 由于共轭运算是复数域

的自同构, 容易看出 $\bar{V}$ 是复数域上的向量空间. 要同时研究 V 和 $\bar{V}$, 纯量乘的符号就得有差别.

二　设 V 是域 K 上的向量空间, 用 0 记 K 和 V 中的零元 (这样做是方便的, 虽然刚开始不习惯), 用 $-v$ 记向量 v 的负元. 对任意向量 $u, v \in V$, 定义 $u - v = u + (-v)$. 下面的等式是定义的简单推论, 其中 $a, b \in K$, $v, u \in V$.

(1) $0v = 0$, (等式左边的 0 是域中的零元, 右边的 0 是零向量.)

(2) $a0 = 0$, (等式两边的 0 都是零向量.)

(3) $(-1)v = -v$,

(4) $a(-v) = -av$,

(5) $a(u - v) = au - av$,

(6) $(a - b)v = av - bv$.

现给出最后一个等式的证明, 其余的作为练习. 从向量空间的定义知

$$(a - b)v + bv = (a - b + b)v = av.$$

于是

$$av + (-bv) = (a - b)v + bv + (-bv) = (a - b)v,$$

即 $av - bv = (a - b)v$.

练习 1.2　证明上面的等式 (1) 至 (5).

三　向量空间的例子是丰富的, 如欧几里得平面几何或立体几何中的向量全体形成的集合在向量的加法和实数与向量的乘法下成为向量空间, 分别记作 E^2 和 E^3. 这是向量空间名称的来源. 又如 $\mathbb{R}^n$ 是实数域上的向量空间. 更一般地, 域 K 上长为 n 的行 $(a_1, \cdots, a_n)$ 全体形成的集合 K^n 是 K 上的向量空间, 加法和纯量乘是

$$(a_1, \cdots, a_n) + (b_1, \cdots, b_n) = (a_1 + b_1, \cdots, a_n + b_n),$$
$$\lambda(a_1, \cdots, a_n) = (\lambda a_1, \cdots, \lambda a_n).$$

它称为 K 上的一个**坐标空间**. 类似地, 域 K 上高为 n 的列 $[a_1, \cdots, a_n]$ 全体形成的集合是 K 上的向量空间, 也记作 K^n, 同样称为 K 上的一个**坐标空间**.

下面的几个例子也是常用的.

例 1.3　平凡的交换群就是只含一个元素的群, 如果用 0 记这个元素, 定义纯量乘为 $a0 = 0$, 这个交换群就可以作为任何域的向量空间, 称为**零向量空间或零维空间**, 常记作**0**或 0.

例 1.4　如果 F 是域 K 的子域, 通过域中的加法和乘法, K 自然成为 F 上的向量空间. 如复数域 $\mathbb{C}$ 是实数域 $\mathbb{R}$ 上的向量空间, 复数域和实数域都是有理数

域上的向量空间. 特别, K 上的向量空间自然成为 F 上的向量空间, 域 K 是它自身上的向量空间.

实数域 (或复数域) 上的向量空间常称为**实向量空间**(或**复向量空间**).

例 1.5 设 U 是域 K 上的向量空间. 从集合 X 到 U 的映射全体记作 U^X. 定义映射之间的加法和 K 中元素与映射的纯量乘如下:

$$(f+g)(x)=f(x)+g(x),\quad \text{对任意的 } x\in X;$$
$$(af)(x)=a(f(x)),\quad \text{对任意的 } a\in K,\ x\in X.$$

易见在这些运算下 U^X 成为域 K 上的向量空间. 特别, K^X 是域 K 上的向量空间.

例 1.6 设 X 是实数集中的一个 (开, 闭, 或半开半闭) 区间, 区间 X 上的所有连续函数全体 $C(X)$ 在函数的加法和函数与实数的乘法下成为实向量空间.

例 1.7 设 K 是域, 则 m 元多项式环 $K[x_1,\cdots,x_m]$ 是 K 上的向量空间, 加法和纯量乘就是环中的加法和 K 中元素与多项式的乘法. 同样, 环 $K[x_1,\cdots,x_m]$ 中次数为 n 的齐次多项式添上零多项式是 K 上的向量空间; 环 $K[x_1,\cdots,x_m]$ 中次数不超过 n 的多项式全体也是 K 上的向量空间.

例 1.8 微分方程 $\dfrac{d^2y}{dt^2}+a\dfrac{dy}{dt}+by=0$ 的解集是实向量空间, 其中 a,b 是给定常数.

例 1.9 域 K 上的 $m\times n$ 矩阵全体在矩阵的加法和 K 中元素与矩阵的乘法下成为 K 上的向量空间, 记作 $M_{m,n}(K)$. 特别, 方阵环 $M_n(K)=M_{n,n}(K)$ 是 K 上的向量空间.

四 线性子空间 设 U 是向量空间 V 的子集, 称 U 为 V 的**线性子空间**(常简称为**子空间**), 如果 U 是 V 的加法子群且对任意的 $a\in K$ 和任意的 $v\in U$ 有 $av\in U$.

每个向量空间都有两个平凡的子空间: 它自身和只含零向量的子空间, 后者也记作**0**或 0, 称为向量空间的**零子空间**. 显然, V 的子空间自身是一个向量空间.

例 1.10 在向量空间 E^3 中和一个给定向量平行的向量全体形成一个子空间.

例 1.11 设 X 是实数集中的一个 (开, 闭或半开半闭) 区间, 区间 V 上的所有可微函数全体 $C^1(X)$ 是这个区间上连续函数空间 $C(X)$ 的子空间.

例 1.12 设 V 是定义在实数轴上的实函数全体, 在函数的加法和实数与函数的乘法下成为实向量空间. V 中的偶函数全体是子空间, 以 π 为周期的周期函数全体是子空间, 形如 ae^x+be^{-x}, $a,b\in\mathbb{R}$ 的函数全体是子空间.

例 1.13 设 A 是域 K 上的 $m\times n$ 矩阵, 那么齐次线性方程组 $AX=0$ 的解集是 K^n 的子空间.

例 1.14 $\mathbb{R}^2$ 上的直线 $y=ax$ 是子空间, 见图 1.1.

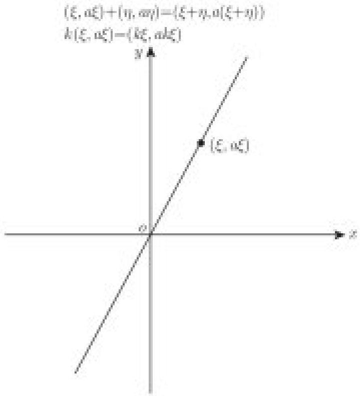


图 1.1

习　题　1.1

判断下列习题 1–6 中的集合是否为实向量空间, 其中的函数都是以实数域为定义域的实值函数.

1. (1) 满足条件 $f(1)=0$ 的函数全体; (2) 满足条件 $f(0)=1$ 的函数全体.

2. (1) 偶函数全体: $f(x)=f(-x)$; (2) 奇函数全体: $f(-x)=-f(x)$.

3. 定义在实数域上的所有的递增函数.

4. 满足条件 $f(x)=f(2-x)$ 所有函数.

5. 以 2π 为周期的函数全体: $f(x+2\pi)=f(x)$.

6. (1) 满足条件 $\int_a^b f(x)dx=0$ 的所有实值连续函数 f;

(2) 闭区间 $[a,b]$ 上满足条件 $\int_a^b f(x)dx\leqslant 0$ 的所有实值连续函数 f;

(3) 满足条件 $\lim\limits_{x\to\infty} f(x)=0$ 的所有函数 f.

设 K 是域. 判断习题 7–9 中的集合是否为原空间的线性子空间.

7. (1) K^n 中坐标满足方程 $x_1+\cdots+x_n=1$ 的所有向量形成的集合; (2) 坐标满足方程 $x_1+\cdots+x_n=0$ 的所有向量形成的集合.

8. $M_{m,n}(K)$ 中秩为 1 的矩阵全体.

9. (1) $M_n(K)$ 中的对称矩阵全体 (${}^tA=A$);

(2)$M_n(K)$ 中的斜对称矩阵全体 (${}^tA=-A$);

(3) $M_n(K)$ 中行列式为 0 的矩阵全体;

(4)$M_n(K)$ 中的迹为 0 的矩阵全体 (矩阵 $A=(a_{ij})$ 的迹定义为对角线中的元素之和:

$\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$).

10. 设 V 是向量空间. 证明: V 的任意一组子空间的交仍是 V 的子空间.

11. 对复数域上的坐标空间 $\mathbb{C}^n$, 定义新的纯量乘为 $b \circ (a_1, \cdots, a_n) = (b\bar{a}_1, \cdots, b\bar{a}_n)$. 在运算 + 和 $\circ$ 下 $\mathbb{C}^n$ 是否为向量空间?

12. 设 M 是有限集, 它的所有的子集形成的集合记作 2^M. 命 $K = \mathbb{Z}_2$. 定义 2^M 中的加法和 K 与 2^M 中元素的乘法如下:

$$A + B = (A \cup B) \backslash (A \cap B), \quad 1A = A, \quad 0A = \varnothing.$$

证明: 在这两个运算下, 2^M 成为域 $\mathbb{Z}_2$ 上的向量空间.

13. 交换群 A 能成为域 $\mathbb{Z}_p$ 上的向量空间当且仅当对任意的 $x \in A$ 有 $px = 0$.

14. 交换群 A 能成为域 $\mathbb{Q}$ 上的向量空间当且仅当 A 中的非零元都是无限阶的, 且对任意的正整数 n 和 $a \in A$, 方程 $nx = a$ 在 A 中有解.

15. 设 K 是有限域, $|K| = q$. 问: 坐标空间 K^n 有多少个元素. 在 K^n 中方程

$$a_1x_1 + \cdots + a_nx_n = b, \quad b \in K, \text{ 系数 } a_i \in K \text{ 不全为 } 0$$

有多少个解?

1.2 向量间的线性关系

一 向量间的线性关系对于向量空间的讨论是基本的. 第一个重要的概念是线性组合. 设 $v_1, v_2, \cdots, v_n \in V$ 是向量, 它们的**线性组合**是指形如

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

的向量, 其中 $a_1, a_2, \cdots, a_n$ 是纯量, 称为这个线性组合的**系数**.

向量 $v_1, v_2, \cdots, v_n$ 的所有线性组合形成的集合称为这些向量的**线性包络**, 记作 $\langle v_1, v_2, \cdots, v_n \rangle$. 易见这个线性包络是 V 的子空间, 且是包含 $v_1, v_2, \cdots, v_n$ 的最小的子空间, 它也称为这些向量**张成的**(线性) 子空间.

一般地, 向量空间 V 的任意子集 M 的线性包络定义为

$$\langle M \rangle = \{a_1x_1 + \cdots + a_kx_k \mid a_1, \cdots, a_k \in K, \quad x_1, \cdots, x_k \in M, \quad k = 1, 2, 3, \cdots\}.$$

即, $\langle M \rangle$ 是 M 中的所有有限子集的线性包络的并集. 易见

$$\langle M \rangle = \left\{ \sum_{x \in M} a_x x \mid \text{诸 } a_x \in K, \text{ 且这些系数 } a_x \text{ 仅有限个不为零} \right\},$$

所以它是 V 的子空间, 且是包含 M 的最小的子空间, 也称它为由 M**张成的**(线性) 子空间.

二　线性相关　向量空间 V 中的向量 (组)$v_1,\cdots,v_n$ 称为**线性相关**如果在基域 K 中有不全为零的元素 $a_1,\cdots,a_n$ 使得线性组合 $a_1v_1+a_2v_2+\cdots+a_nv_n$ 为零向量; 称为**线性无关**如果只要纯量 $a_1,\cdots,a_n\in K$ 不全为零, 线性组合 $a_1v_1+a_2v_2+\cdots+a_nv_n$ 就不是零向量 (即 $a_1v_1+a_2v_2+\cdots+a_nv_n=0$ 蕴含 $a_1=\cdots=a_n=0$).

例 1.15　在函数空间中 $\sin 3t$, $\sin t$, $\sin^3 t$ 线性相关, 因为 $\sin 3t-3\sin t+4\sin^3 t=0$.

例 1.16　在函数空间中 $\sin t$, $\sin 2t$, $\sin 3t$ 线性无关. 假设 $a\sin t+b\sin 2t+c\sin 3t=0$. 分别取 $t=\dfrac{\pi}{6}$, $\dfrac{\pi}{3}$, $\dfrac{\pi}{2}$, 得到线性方程组

$$\begin{aligned}\frac{1}{2}a+\frac{\sqrt{3}}{2}b+c=0,\\ \frac{\sqrt{3}}{2}a+\frac{\sqrt{3}}{2}b=0,\\ a-c=0.\end{aligned}$$

该方程组只有零解 $a=b=c=0$. 所以这些函数线性无关. 更简单的方法是利用公式

$$\int_0^{2\pi}\sin jt\sin kt\,dt=\begin{cases}\pi, & \text{如果}\ j=k,\\ 0, & \text{否则}.\end{cases}$$

从等式 $a\sin t+b\sin 2t+c\sin 3t=0$ 得 $a\sin t\sin jt+b\sin 2t\sin jt+c\sin 3t\sin jt=0$. 取 $j=1,2,3$, 然后对等式在区间 $[0,2\pi]$ 做定积分, 即得 $a\pi=b\pi=c\pi=0$.

三　下面是一些关于线性相关和线性无关的简单性质, 它们绝大部分在第一卷 3.1 节关于 $\mathbb{R}^n$ 的讨论中已经出现过.

定理 1.17　含有零向量的向量组是线性相关的. 至少有两个元素的向量组线性相关当且仅当其中有一个是其余的线性组合. 如果一组向量的一部分线性相关, 则这组向量线性相关; 换句话说, 如果一组向量线性无关, 则这组向量的任何部分向量是线性无关的.

我们省略证明, 因为直接从定义就可以看出, 除第一个显然的结论外, 其余结论的证明也是第一卷 3.1 节的相关证明的重复. 下一个结论是经常用到的.

定理 1.18　设向量 $u_1,\cdots,u_m$ 都是向量 $v_1,\cdots,v_n$ 的线性组合.

(1) 如果 $m>n$, 那么 $u_1,\cdots,u_m$ 线性相关;

(2) 如果 $u_1,\cdots,u_m$ 线性无关, 那么 $m\leqslant n$.

证明　(1) 和 (2) 等价. 现证 (2). 用反证法, 假设 $m>n$.

首先有 $u_1=a_1v_1+\cdots+a_nv_n$. 因为 $u_1\neq 0$, 所以线性组合中的系数不全为零, 不妨设 $a_{i_1}\neq 0$, 那么 v_{i_1} 是 $u_1,v_1,\cdots,\hat{v}_{i_1},\cdots,v_n$ 的线性组合 (本证明中向量上带 ˆ 的含义是在序列或集合中去掉这个向量). 于是有

$$\langle v_1,\cdots,v_n\rangle=\langle u_1,v_1,\cdots,\hat{v}_{i_1},\cdots,v_n\rangle.$$

从而 u_2 是 $u_1, v_1, \cdots, \hat{v}_i, \cdots, v_n$ 的线性组合. 由于 u_1, u_2 线性无关, 所以这个线性组合中某个 v_{i_2} 的系数不为零. 这样 v_{i_2} 就是 $u_1, u_2, v_1, \cdots, \hat{v}_{i_1}, \cdots, \hat{v}_{i_2}, \cdots, v_n$ 的线性组合. 于是

$$\langle v_1, \cdots, v_n\rangle = \langle u_1, v_1, \cdots, \hat{v}_i, \cdots, v_n\rangle = \langle u_1, u_2, v_1, \cdots, \hat{v}_{i_1}, \cdots, \hat{v}_{i_2}, \cdots, v_n\rangle.$$

如此下去, 可得

$$\langle v_1, \cdots, v_n\rangle = \langle u_1, \cdots, u_n\rangle.$$

于是 u_{n+1} 是 $u_1, \cdots, u_n$ 的线性组合. 这与条件矛盾, 所以 $m \leqslant n$. □

注 也可以像第一卷 3.1 节引理 3.7 的证明那样通过某个齐次线性方程组的非零解以证明 (1) 成立.

练习 1.19 证明:

(1) 如果向量 $v_1, \cdots, v_n$ 线性无关, $v_1, \cdots, v_n, v$ 线性相关, 则 v 是 $v_1, \cdots, v_n$ 的线性组合;

(2) 如果向量 $v_1, \cdots, v_n$ 线性无关, 向量 v 不能表示成 $v_1, \cdots, v_n$ 的线性组合, 则 $v, v_1, \cdots, v_n$ 线性无关.

(3) 如果 u 是 $v_1, \cdots, v_n$ 的线性组合, 而每个 v_i 又是 $w_1, \cdots, w_m$ 的线性组合, 那么 u 是 $w_1, \cdots, w_m$ 的线性组合, 即

$$u \in \langle v_1, \cdots, v_n\rangle \text{ 且每个 } v_i \in \langle w_1, \cdots, w_m\rangle \Longrightarrow u \in \langle w_1, \cdots, w_m\rangle.$$

习 题 1.2

1. 设 u, v 是向量, a, b 是纯量. 证明:

(1) $au = 0$ 当且仅当 $a = 0$ 或 $u = 0$;

(2) $au + bv = bu + av$ 当且仅当 $a = b$ 或 $u = v$.

2. (1) 向量 u, v 线性无关是否蕴含 $au + v$, $u + av$ 线性无关;

(2) 向量组 $v_1, \cdots, v_n$ 线性无关是否蕴含向量组 $v_1 + v_2, v_2 + v_3, \cdots, v_{n-1} + v_n, v_n + av_1$ 线性无关.

3. 证明下列函数组线性无关:

(1) $\sin x, \cos x$;

(2) $1, \sin x, \cos x$;

(3) $\sin x, \sin 2x, \cdots, \sin nx$;

(4) $\sin x, \sin^2 x, \cdots, \sin^n x$;

(5) $1, \sin x, \cos x, \sin 2x, \cos 2x, \cdots, \sin nx, \cos nx$.

4. 证明下列函数组线性无关:

(1) $e^{a_1 x}, \cdots, e^{a_n x}$;

(2) $x^{a_1}, \cdots, x^{a_n}$;

(3) $(1-a_1x)^{-1}, \cdots, (1-a_nx)^{-1}$.

其中 $a_1, \cdots, a_n$ 是互不相同的实数.

5. 在单变量实函数空间中, 向量 $f_1, \cdots, f_n$ 线性无关当且仅当存在实数 $a_1, \cdots, a_n$ 使得 $\det(f_i(a_j)) \neq 0$.

1.3　基与维数

一　有了线性无关的概念就可以定义向量空间的基和维数.

定义 1.20　称向量组 $v_1, \cdots, v_n$ 是 (或说形成、构成) 向量空间 V 的**基**如果这些向量线性无关且张成 V. 诸 v_i 称为 (这个)**基的向量**, 简称为**基向量**. 在不产生歧义的情况下, 有时我们会简单地用记号 (v_i) 表示基 $v_1, \cdots, v_n$.

例 1.21　在欧几里得平面几何或立体几何中的向量全体形成的空间 E^2 和 E^3 中, 平行于同一条直线的向量称为**共线的**, 平行于同一个平面的向量称为**共面的**. 在 E^2 中不共线的两个向量形成一个基, 在 E^3 中不共面的三个向量形成一个基.

例 1.22　向量 $\varepsilon_1 = (1, 0, 0, \cdots, 0)$, $\varepsilon_2 = (0, 1, 0, \cdots, 0)$, $\cdots$, $\varepsilon_n = (0, 0, 0, \cdots, 1)$ 形成空间 K^n 的一个基, 称为 K^n 的**标准基**.

例 1.23　在例 1.5 中, 如果 X 是有限集, 对每个 $x \in X$, 定义

$$\delta_x(y) = \begin{cases} 1, & \text{如果 } y = x, \\ 0, & \text{否则}. \end{cases}$$

那么向量 δ_x, $x \in X$ 线性无关. 对任意的 $f \in K^X$, 有 $f = \sum\limits_{x \in X} f(x)\delta_x$. 所以, 向量 δ_x, $x \in X$ 是 K^X 的基.

向量空间的基对于该空间的向量间的关系的讨论是很有用的, 利用基可以建立抽象的向量空间与空间 K^n 的联系.

定理 1.24　假设 $v_1, \cdots, v_n$ 是向量空间 V 的基. 那么 V 中每一个向量 v 都以唯一的方式表示成 $v_1, \cdots, v_n$ 的线性组合; 线性组合的系数称为 v 关于这个基 (或在这个基下) 的**坐标**.

证明　根据定义, v 可以表示成 $v_1, \cdots, v_n$ 的线性组合. 假设 v 有两种方式表示成 $v_1, \cdots, v_n$ 的线性组合:

$$a_1v_1 + \cdots + a_nv_n = v = b_1v_1 + \cdots + b_nv_n.$$

两个线性组合相减, 得

$$(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0.$$

由于 $v_1, \cdots, v_n$ 线性无关, 得

$$a_1 - b_1 = 0, \ \cdots, \ a_n - b_n = 0.$$

所以 $a_1 = b_1, \ \cdots, \ a_n = b_n$. 即 v 表示成 $v_1, \cdots, v_n$ 的线性组合的方式是唯一的. □

定理 1.25 假设 $v_1, \cdots, v_n$ 是向量空间 V 的基, V 的基域为 K, 那么映射

$$\varphi : V \to K^n, \quad a_1v_1 + \cdots + a_nv_n \to (a_1, \cdots, a_n)$$

是双射, 且对任意的 $a, b \in K$, $u, v \in V$ 有 $\varphi(au + bv) = a\varphi(u) + b\varphi(v)$.

证明 根据定理 1.24, V 中向量都是以唯一的方式表示成 $v_1, \cdots, v_n$ 的线性组合, 所以映射 φ 是确切定义的且是单射. 显然这个映射是满射, 所以它是双射.

命 $u = a_1v_1 + \cdots + a_nv_n$, $v = b_1v_1 + \cdots + b_nv_n$, 那么

$$au + bv = (aa_1 + bb_1)v_1 + \cdots + (aa_n + bb_n)v_n.$$

于是

$$\begin{aligned}\varphi(au + bv) &= (aa_1 + bb_1, \cdots, aa_n + bb_n) \\ &= a(a_1, \cdots, a_n) + b(b_1, \cdots, b_n) \\ &= a\varphi(u) + b\varphi(v).\end{aligned}$$

定理得证. □

定义 1.26 称域 K 上的两个向量空间 V, U **同构**如果存在双射 $\varphi : V \to U$ 使得对任意的 $a, b \in K$, $u, v \in V$ 有

$$\varphi(au + bv) = a\varphi(u) + b\varphi(v).$$

映射 φ 也称为**(线性) 同构**. 注意条件 $\varphi(au+bv) = a\varphi(u)+b\varphi(v)$ 等价于 $\varphi(u+v) = \varphi(u) + \varphi(v)$ 且 $\varphi(au) = a\varphi(v)$, 即映射保持加法和纯量乘两个运算.

定理 1.25 表明如果域 K 上的向量空间 V 的一个基有 n 个向量, 那么 V 与 K^n 同构.

练习 1.27 证明: 如果 $\varphi : V \to U$, $\psi : U \to W$ 是同构, 那么 $\varphi^{-1} : U \to V$ 和 $\psi\varphi : V \to W$ 都是同构.

练习 1.28 设 $\varphi : V \to U$ 是同构, $v_1, \cdots, v_m$ 是 V 中的向量. 证明这些向量线性无关当且仅当 $\varphi(v_1), \cdots, \varphi(v_m)$ 线性无关.

二 接下来对基的性质做进一步的讨论.

命题 1.29 假设 $u_1, \cdots, u_m$ 和 $v_1, \cdots, v_n$ 是向量空间 V 的两个基, 那么 $m = n$. 从而 V 的所有基含有同样数量的向量. 一个基中所含向量的个数称为 V 的**维数**, 记作

$$\dim V.$$

零向量空间的基约定为空集, 维数为 0.

证明 这是定理 1.18 的直接推论. □

定理 1.26 和练习 1.27 及练习 1.28 表明, 基域相同的向量空间是同构的当且仅当它们的维数相同, 而且都同构于某个 K^n. 虽然抽象地看, 有限维向量空间与向量空间 K^n 没有差别, 但具体的向量空间常有更丰富的结构和直观的性质, 如空间 E^2, E^3, 多项式空间, 函数空间 (参考例 1.16) 等.

定理 1.30 假设 V 的维数是 n, 那么

(1) V 中任意 $n+1$ 个向量都线性相关;

(2) V 中任意 n 个线性无关的向量都构成 V 的一个基.

证明 (1) 由定理 1.18(1) 推出.

(2) 设 $u_1,\cdots,u_n$ 线性无关, 从定理 1.18 的证明知这 n 个向量张成 V, 所以构成 V 的基. □

定理 1.31 假设 V 的维数是 n.

(1) 如果 V 中的 n 个向量张成 V, 那么这 n 个向量线性无关, 从而构成 V 的基;

(2) V 中任意一组线性无关的向量都可以扩充为 V 的一个基, 特别, 任何非零向量都在某个基中.

证明 (1) 如果这 n 个向量 $u_1,\cdots,u_n$ 线性相关, 那么其中的一个是其余向量的线性组合, 从而 V 可由 $n-1$ 个向量张成. 于是一个基中的 n 个向量都是这 $n-1$ 个向量的线性组合. 根据定理 1.18(1), 基中的 n 个向量线性相关. 这是一个矛盾, 所以向量 $u_1,\cdots,u_n$ 必须线性无关. 由定理 1.30(2) 知, $u_1,\cdots,u_n$ 是 V 的基.

(2) 假设 $u_1,\cdots,u_m\in V$ 线性无关. 根据定理 1.18(2), $m\leqslant n$. 从定理 1.18 的证明知, 对 V 的任意基 $v_1,\cdots,v_n$, 存在 $i_1,\cdots,i_m$ 使得

$$\langle v_1,\cdots,v_n\rangle=\langle u_1,\cdots,u_m,v_1,\cdots,\hat{v}_{i_1},\cdots,\hat{v}_{i_m},\cdots,v_n\rangle.$$

由定理 1.30(2) 知, $u_1,\cdots,u_m,v_1,\cdots,\hat{v}_{i_1},\cdots,\hat{v}_{i_m},\cdots,v_n$ 是 V 的基. □

如果向量空间 V 的任何有限个向量都不能张成 V, 那么称 V 是**无限维空间**, 维数 $\dim V=\infty$. 域 K 上的多项式环 $K[t]$ 和实数轴上的连续函数全体分别是 K 和 $\mathbb{R}$ 上的无限维空间. 本书主要讨论有限维向量空间.

有限维向量空间的一个向量组 (可能含有无限多的向量) 的一部分向量称为这个向量组的**极大线性无关组**如果这部分向量线性无关, 而向量组任何其他向量都是这部分向量的线性组合. 例如, 在 K^4 中, 向量组 (1,1,0,0), (0,1,0,1), (1,0,1,0), (1,2,0,1) 的前三个向量和后三个向量都是极大线性无关组.

定理 1.32 设 S 是有限维向量空间 V 的子集, 含有非零向量, 那么

(1) 向量组 S 有极大线性无关组;

(2) 向量组 S 中任意两个极大线性无关组所含的向量的个数相同;

(3) 如果 $U = \langle S\rangle$, 那么 S 的极大线性无关组是 U 的基.

证明 (1) 取 S 中的非零向量 v_1. 如果 $\langle S\rangle = \langle v_1\rangle$, 那么 v_1 就是 S 的极大线性无关组. 否则, 取 $v_2 \in \langle S\rangle \setminus V_1$, 其中 $V_1 = \langle v_1\rangle$. 由练习 1.19(2) 知, v_1, v_2 线性无关. 如果 $V_2 = \langle v_1, v_2\rangle$ 等于 $\langle S\rangle$, 那么 v_1, v_2 就是 S 的极大线性无关组. 否则取 $v_3 \in \langle S\rangle \setminus V_2$, 根据练习 1.19(2), v_1, v_2, v_3 线性无关. 如此重复下去, 假设在 S 中取出了线性无关的向量 $v_1, v_2, \cdots, v_k$. 由于 $S \subset V$, 由定理 1.30 得 $k \leqslant \dim V = n$. 这样一来, S 中有线性无关的向量组, 每个线性无关的向量组中的向量的个数不超过 n.

设 $v_1, v_2, \cdots, v_r$ 是 S 中的线性无关向量组, 所含的向量个数最多. 根据练习 1.19(2), S 中的每一个向量都是 $v_1, v_2, \cdots, v_r$ 的线性组合, 从而 $v_1, v_2, \cdots, v_r$ 是 S 的极大线性无关组.

(2) 设 $u_1, u_2, \cdots, u_s$ 是 S 的另一个极大线性无关组. 由于 $u_1, u_2, \cdots, u_s$ 都是 $v_1, v_2, \cdots, v_r$ 的线性组合, 由 1.18(2), $s \leqslant r$. 反过来, $v_1, v_2, \cdots, v_r$ 都是 $u_1, u_2, \cdots, u_s$ 的线性组合, 所以 $r \leqslant s$. 于是 $r = s$.

(3) 从基和极大线性无关组的定义推出. □

向量组的秩定义为向量组张成的空间的维数, 它等于向量组的极大线性无关组所含的向量个数. 我们用记号 $\operatorname{rank} S$ 或 $\operatorname{rk} S$ 表示向量组 S 的秩.

三 转换矩阵 设 V 是域 K 上的 n 维向量空间, $u_1, \cdots, u_n$ 和 $v_1, \cdots, v_n$ 是它的两个基. 一个基的向量可以表示成另一个基的向量的线性组合, 而且表示方式唯一.

$$\begin{aligned} u_1 &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n, \\ u_2 &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n, \\ &\cdots\cdots \\ u_n &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n. \end{aligned} \tag{1.3.1}$$

线性组合中的系数给出矩阵

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \tag{1.3.2}$$

称为从基 $v_1, \cdots, v_n$ 到基 $u_1, \cdots, u_n$ 的**转换矩阵**. *注意 u_j 在基 $v_1, \cdots, v_n$ 下的坐标是转换矩阵 A 的第 j 列.*

根据练习 1.28 和定理 1.25, 矩阵 A 的列向量是线性无关的, 从而矩阵 A 的秩是 n, 于是它可逆 (参看第一卷 5.4 节第四部分). 反过来, 对于域 K 上的任意一个

n 阶可逆矩阵 $A=(a_{ij})$, 通过等式 (1.3.1) 给出了 V 的一个新基 $u_1,\cdots,u_n$.

练习 1.33 证明矩阵 A 的逆矩阵就是从基 $u_1,\cdots,u_n$ 到基 $v_1,\cdots,v_n$ 的转换矩阵.

向量在不同基下的坐标通过转换矩阵建立联系. 设 $\lambda_1,\cdots,\lambda_n$ 和 $\mu_1,\cdots,\mu_n$ 分别是向量 x 在基 $u_1,\cdots,u_n$ 和基 $v_1,\cdots,v_n$ 下的坐标, 即

$$\lambda_1u_1+\cdots+\lambda_nu_n=x=\mu_1v_1+\cdots+\mu_nv_n.$$

利用等式 (1.3.1) 得

$$\begin{aligned}x&=\mu_1v_1+\cdots+\mu_nv_n\\&=\lambda_1(a_{11}v_1+a_{21}v_2+\cdots+a_{n1}v_n)+\cdots+\lambda_n(a_{1n}v_1+a_{2n}v_2+\cdots+a_{nn}v_n).\end{aligned}$$

比较系数得

$$\begin{aligned}\mu_1&=a_{11}\lambda_1+a_{12}\lambda_2+\cdots+a_{1n}\lambda_n,\\\mu_2&=a_{21}\lambda_1+a_{22}\lambda_2+\cdots+a_{2n}\lambda_n,\\&\cdots\cdots\\\mu_n&=a_{n1}\lambda_1+a_{n2}\lambda_2+\cdots+a_{nn}\lambda_n.\end{aligned}\tag{1.3.3}$$

可以把上面这些等式写得更紧凑

$$Y=AX,\tag{1.3.4}$$

其中 $X=[\lambda_1,\cdots,\lambda_n]$ 和 $Y=[\mu_1,\cdots,\mu_n]$ 分别是 x 在两个基下的坐标列. 上面的讨论可以总结成如下的结论.

定理 1.34 设 A 是从向量空间 V 的基 $v_1,\cdots,v_n$ 到基 $u_1,\cdots,u_n$ 的转换矩阵. 那么

(1) A 是可逆矩阵;

(2) 向量 $x\in V$ 在基 $u_1,\cdots,u_n$ 下的坐标列向量 X 可通过 A 的逆矩阵和 x 在基 $v_1,\cdots,v_n$ 下的坐标列向量 Y 计算:

$$X=A^{-1}Y.$$

四 线性组合可以写成矩阵的乘积

$$\mu_1v_1+\mu_2v_2+\cdots+\mu_nv_n=(v_1,v_2,\cdots,v_n)\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix},$$

等式 (1.3.1) 也能写成矩阵的形式

$$(u_1,u_2,\cdots,u_n)=(v_1,v_2,\cdots,v_n)A.$$

采用矩阵的记号, 等式 (1.3.4) 的推导就变得简明了:

$$x=(u_1,u_2,\cdots,u_n)X=(v_1,v_2,\cdots,v_n)AX=(v_1,v_2,\cdots,v_n)Y.$$

所以 $Y=AX$.

练习 1.35 设 A 是从向量空间 V 的基 $v_1,\cdots,v_n$ 到基 $u_1,\cdots,u_n$ 的转换矩阵; B 是从 V 的基 $u_1,\cdots,u_n$ 到基 $w_1,\cdots,w_n$ 的转换矩阵. 证明: 从基 $v_1,\cdots,v_n$ 到基 $w_1,\cdots,w_n$ 的转换矩阵是 AB.

五 关于基向量的序 说到向量组 $v_1,\cdots,v_n$ 是空间 V 的基时, *一般默认它是有序的向量组*. 很多的问题, 如基之间的转换矩阵, 向量在基下的坐标, 基向量之间的顺序都是重要的. 当然, 也有很多问题, 基向量的序无关紧要, 如一个向量可以写成基向量的线性组合这个事实就与基向量的序没有关系. 以后 (包括下面的习题), 如果不加说明, 基向量的序是固定的.

习 题 1.3

1. (1) 找出二维向量空间的所有非平凡线性子空间; (2) 找出三维向量空间的所有线性子空间.

2. 设 V 是 q 元域 K 上的 n 维向量空间 $(n\geqslant 2)$. 求

(1) V 的基的个数;

(2) K 上 n 阶可逆方阵的个数;

(3) K 上 n 阶退化矩阵的个数;

(4) 对 $1\leqslant k\leqslant n-1$, V 中 k 维子空间的个数.

3. 求出满足下列条件的复 n 阶方阵形成的向量空间的维数:

(1) 对称矩阵; (2) 斜对称矩阵; (3) 迹为 0 的矩阵.

4. 设 P_n 为域 K 上次数不超过 $n-1$ 的多项式形成的向量空间. 证明 P_n 中满足下列条件的多项式全体形成 P_n 的子空间并找出这些子空间的一个基和维数:

(1) $c\in K$ 是多项式的一个根;

(2) $c_1,\cdots,c_k\in K$ 是多项式的根.

5. 设 $K[x_1,\cdots,x_m]$ 是域 K 上的 m 元多项式环. 求下列空间的维数:

(1) 所有齐次 k 次多项式张成的空间;

(2) 所有对称的齐次 k 次多项式张成的空间.

6. 设 P_n 是次数不超过 $n-1$ 的实多项式形成的空间. 找出从基 $1,t-c,\cdots,(t-c)^{n-1}$ 到基 $1,t,\cdots,t^{n-1}$ 的转换矩阵.

7. 设 θ 是有理数域上一个既约多项式 f 的复根. 把复数域看作有理数域上的向量空间, 求出其子空间 $\mathbb{Q}[\theta]=\langle 1,\theta,\cdots,\theta^k,\cdots\rangle$ 的维数.

8. 证明实数域上的连续函数全体是无限维实向量空间.

9. 证明: 向量空间 V 是无限维的当且仅当 V 中有一个无限的向量序列 $v_1, v_2, v_3, \cdots$ 使得对每个正整数 k, 向量 $v_1, v_2, \cdots, v_k$ 线性无关.

10. 设 S 和 T 是一个向量空间的子集. 证明: $\langle S\rangle \cap \langle T\rangle \supseteq \langle S\cap T\rangle$.

11. 求下列向量张成的线性子空间的一个基和维数: (1,1,1,1,0), (1,1,−1,−1,−1), (2,2,0,0,−1), (1,1,5,5,2), (1,−1,−1,0,0).

1.4 子 空 间

一　显然, 向量空间 V 的两个子空间 U, W 的交集 $U\cap W$ 是 V 的子空间, 也是 U 和 W 的子空间. 但是两个子空间的并集 $U\cup W$ 一般不是子空间, 例如实平面的横轴与纵轴都是子空间, 但并集只是垂直的两条直线, 不是子空间.

定义 1.36　两个子空间 U 和 W 的**和**是所有形如 $u+w$, $u\in U$, $w\in W$ 的向量形成的集合, 记做 $U+W$. 类似地, V 的有限个子空间 $U_1, \cdots, U_m$ 的和定义为所有形如 $u_1+\cdots+u_m$, $u_1\in U_1, \cdots, u_m\in U_m$, 的向量形成的集合, 记作 $U_1+\cdots+U_m$.

显然, $U_1+\cdots+U_m$ 是 V 中含诸 U_i 的最小子空间. 以下假设 V 是维数为 n 的向量空间, U 是 V 的子空间. 根据定理 1.32, 如果 U 不是零子空间, 那么 U 的极大线性无关组是 U 的基. 根据定理 1.31(2), U 的任何基可以扩充为 V 的基, 于是我们有

命题 1.37　设 U 是 n 维向量空间 V 的子空间, 那么 U 有基, $\dim U\leqslant \dim V$. 等号成立当且仅当 $U=V$.

下面考虑 V 的基与子空间的关系. 向量空间 V 的一个基称为与子空间 U **相合**如果 U 由这个基的部分向量张成. 例如在 $\mathbb{R}^3$ 中, 基 $(1,-1,0)$, $(0,0,1)$, $(0,1,-1)$ 与平面

$$U=\{(x_1,x_2,x_3)\,|\,x_1+x_2+x_3=0\}$$

相合.

寻找与若干子空间相合的基是很有意思的问题. 前些年发现的量子群的典范基就与量子群的很多有意思的子空间相合, 这个结果引起一系列的后续研究. 对一个子空间而言, 这个问题是简单的, 因为子空间的基都可以扩充成全空间 V 的基, 所以与一个子空间相合的 V 的基是很多的. 令人惊讶的是, 这个结果可以推广到两个子空间的情形.

定理 1.38　对 V 的任意两个子空间 U, W, 存在 V 的基, 它与 U 和 W 都相合.

证明　设 $u_1, \cdots, u_r$ 是 $U\cap W$ 的基. 根据定理 1.31(2), 它可以扩充为 U 的基 $u_1, \cdots, u_r, v_1, \cdots, v_s$, 还可以扩充为 W 的基 $u_1, \cdots, u_r, w_1, \cdots, w_t$. 下面证明

向量 $u_1, \cdots, u_r, v_1, \cdots, v_s, w_1, \cdots, w_t$ 线性无关. 于是, 这组向量就可以扩充为 V 的基, 所得到的基与 U 和 W 都相合.

假设

$$\sum_{i=1}^{r}\alpha_i u_i + \sum_{j=1}^{s}\beta_j v_j + \sum_{k=1}^{t}\gamma_k w_k = 0.$$

考虑向量

$$x = \sum_{i=1}^{r}\alpha_i u_i + \sum_{j=1}^{s}\beta_j v_j = -\sum_{k=1}^{t}\gamma_k w_k.$$

向量 x 的第一个表达式说明它在 U 中, 第二个表达式意指它在 W 中. 于是 $x \in U\cap W$, 从而

$$x = \sum_{i=1}^{r}\lambda_i u_i = -\sum_{k=1}^{t}\gamma_k w_k.$$

但 $u_1, \cdots, u_r, w_1, \cdots, w_t$ 线性无关, 所以 $\gamma_1, \cdots, \gamma_t$ 均为 0. 因为 $u_1, \cdots, u_r, v_1, \cdots, v_s$ 线性无关, 从等式

$$\sum_{i=1}^{r}\alpha_i u_i + \sum_{j=1}^{s}\beta_j v_j = 0$$

知诸 α_i, β_j 均为 0. 定理得证. □

推论 1.39 $\dim(U+W) = \dim U + \dim W - \dim U\cap W$.

证明 沿用定理 1.38 证明中的记号, 则 $u_1, \cdots, u_r, v_1, \cdots, v_s, w_1, \cdots, w_t$ 是 $U+W$ 的基. 于是

$$\dim(U+W) = r+s+t = \dim U + \dim W - \dim U\cap W.$$ □

练习 1.40 举例说明定理 1.38 不能推广到三个子空间的情形.

1 维的向量空间称为**直线**, 2 维的向量空间称为**平面**. 设 V 是有限维空间, U 是 V 的子空间. 如果 U 的维数是 k, 则称 U 为 V 的k **维平面**. 当 $k=1,2$ 时, U 也分别称为 V 的直线和平面. 维数差

$$\operatorname{codim} U = \dim V - \dim U$$

称为子空间 U 的**余维数**. 余维数为 1 的子空间称为**超平面**. 所有的子空间链

$$0\subset V_1\subset V_2\subset\cdots\subset V_{n-1}\subset V_n = V,\quad \dim V_i = i,\quad i=1,2,\cdots,n$$

形成的集合称为**旗簇**, 是几何中非常重要的研究对象.

二 直和 练习 1.40 从一个侧面说明多个子空间的相互关系是复杂的, 但最让人感兴趣的一个极端情况是容易处理的.

定义 1.41　称子空间 $U_1, \cdots, U_m$ **线性无关**如果等式 $u_1 + \cdots + u_m = 0$, $u_i \in U_i$, 蕴含 $u_1 = \cdots = u_k = 0$. 换句话说, 称这些空间线性无关如果 $u_1 \in U_1, \cdots, u_m \in U_m$ 不全为零时, 有 $u_1 + \cdots + u_m \neq 0$. 否则称这些空间**线性相关**.

两个子空间 U 和 W 线性无关等价于 $U \cap W = 0$. 但对多个子空间的情形, 直接的推广并不成立.

练习 1.42　举例说明存在三个线性相关的子空间, 它们中的任何两个的交集都是零向量空间.

定理 1.43　子空间 $U_1, \cdots, U_m$ 线性无关当且仅当如下条件满足

$$U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m) = 0, \quad i = 1, 2, \cdots, m. \tag{1.4.5}$$

(子空间 U_i 上带上 ˆ 表示该子空间在和中不出现, 即

$$U_1 + \cdots + \hat{U}_i + \cdots + U_m = U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m.)$$

证明　假设 $U_1, \cdots, U_m$ 线性无关. 对任意的 i, 取 $u \in U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m)$. 于是 $u = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_m, u_j \in U_j, j \neq i$. 从而

$$0 = u_1 + \cdots + u_{i-1} + (-u) + u_{i+1} + \cdots + u_m.$$

根据定义, $u = 0$, 所以 $U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m) = 0$.

反之, 假设条件 (1.4.5) 满足. 如果有 $u_i \in U_i$, $i = 1, \cdots, m$ 使得

$$u_1 + \cdots + u_i + \cdots + u_m = 0,$$

则有

$$u_i = -(u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_m) \in U_i \cap (U_1 + \cdots + \hat{U}_i + \cdots + U_m) = 0.$$

即 $u_i = 0$ 对所有的 i 成立. □

线性无关的子空间的和的维数与这些子空间的维数的关系特别简单.

定理 1.44　子空间 $U_1, \cdots, U_m$ $(m \geqslant 2)$ 线性无关当且仅当

$$\dim(U_1 + \cdots + U_m) = \dim U_1 + \cdots + \dim U_m. \tag{1.4.6}$$

证明　当 $m = 2$ 时, 由推论 1.39 和定理 1.43 知定理成立. 现用归纳法证明一般情形. 如果 $U_1, \cdots, U_m$ 线性无关, 由推论 1.39, 定理 1.43 和归纳假设得

$$\begin{aligned}\dim(U_1 + U_2 + \cdots + U_m) &= \dim U_1 + \dim(U_2 + \cdots + U_m) \\ &= \dim U_1 + \dim U_2 + \cdots + \dim U_m.\end{aligned}$$

反过来, 假设等式 (1.4.6) 成立, 那么所有子空间 U_i 的基合起来是 $U_1+\cdots+U_m$ 的基. 如果 $u_i \in U_i$ 不为零, 则 u_i 可以扩充为 U_i 的基. 于是, 如果 $u_1 \in U_1, \cdots, u_m \in U_m$ 不全为零, 则那些非零的向量可以扩充为 $U_1+\cdots+U_m$ 的基, 从而它们的和不为零. □

线性无关的子空间 $U_1, \cdots, U_m$ 的和称为它们的**直和**, 记作 $U_1 \oplus \cdots \oplus U_m$. 诸 U_i 称为直和的**直和项**. 直和中的向量分解成直和项中的向量的和的方式是唯一的, 即对直和中的向量 u, 存在唯一的 $u_i \in U_i$ 使得 $u = u_1+\cdots+u_m$.

利用直和的概念可以定义子空间的补空间.

定理 1.45 设 V 是 n 维向量空间, U 是 V 的 m 维子空间, 那么存在 V 的 $n-m$ 维子空间 W 使得 $V = U \oplus W$. 子空间 W 称为 U 的**补空间**, U 和 W 称为**互补**的子空间.

证明 取 U 的基 $u_1, \cdots, u_m$ 并把它扩充为 V 的基 $u_1, \cdots, u_m, w_1, \cdots, w_{n-m}$, 那么 $W = \langle w_1, \cdots, w_{n-m}\rangle$ 满足要求. □

刚才所讨论的直和的直和项都在一个向量空间内. 这样的直和常称为**内直和**. 与内直和相对的还有外直和的概念. 设 U, V 是域 K 上的任意两个向量空间 (可以相同). 在它们的笛卡尔积 $W = U \times V$ 上定义加法和纯量乘如下:

$$(u, v) + (u', v') = (u+u', v+v'),$$
$$\alpha(u, v) = (\alpha u, \alpha v).$$

易见, W 成为域 K 上的向量空间, 称为 U 和 V 的**外直和**, 也记作 $U \oplus V$. 例如 $\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$. 外直和的构造类似于用横轴与纵轴构造平面.

在 U 和 V 的外直和 W 中, 形如 $(u, 0)$ 的向量全体 $\tilde{U}$ 是 W 的一个子空间, 与 U 同构; 而形如 $(0, v)$ 的向量全体 $\tilde{V}$ 也是 W 的子空间, 与 V 同构. 显然 $\tilde{U}$, $\tilde{V}$ 线性无关, 且有

$$\underbrace{U \oplus V}_{\text{外直和}} = W = \underbrace{\tilde{U} \oplus \tilde{V}}_{\text{内直和}}.$$

因此, 外直和与内直和没有本质的差别, 两者常简单称为**直和**.

习 题 1.4

1. 如果 U_1, U_2, U_3 是有限维向量空间 V 的子空间. 证明

$$\dim(U_1+U_2+U_3) \leqslant \dim U_1 + \dim U_2 + \dim U_3$$
$$- \dim U_1 \cap U_2 - \dim U_1 \cap U_3 - \dim U_2 \cap U_3 + \dim U_1 \cap U_2 \cap U_3.$$

举例说明上式左端比右端小的情况可以发生.

2. 设 U 和 W 是 V 的有限维子空间, $\dim(U+W)=1+\dim U\cap W$. 证明: $U+W$ 是这两个子空间中的一个, $U\cap W$ 是这两个子空间中的另一个.

3. 证明: 子空间 $U_1,U_2,\cdots,U_m$ 线性无关当且仅当

$$(U_1+U_2+\cdots+U_{i-1})\cap U_i=0,\quad 1<i\leqslant m.$$

4. 设 U,V,W 是一个向量空间的子空间. 下列等式是否成立?

$$U\cap(V+W)=(U\cap V)+(U\cap W).$$

5. 证明直和不满足消去律, 即等式 $U\oplus W_1=U\oplus W_2$ 并不意味着 $W_1=W_2$.

6. 设 $V=M_n(K)$, U 和 W 分别是 V 中对称矩阵和斜对称矩阵形成的子空间. 证明: 如果 $\operatorname{char}K\neq 2$, 那么 $V=U\oplus W$.

7. 找出子空间 $\langle u_1,u_2,u_3\rangle$ 和子空间 $\langle v_1,v_2,v_3\rangle$ 的和与交的一个基:

(1) $u_1=(1,2,1)$, $u_2=(1,1,-1)$, $u_3=(1,3,3)$,
$v_1=(1,2,2)$, $v_2=(2,3,-1)$, $v_3=(1,1,-3)$;

(2) $u_1=(-1,6,4,7,-2)$, $u_2=(-2,3,0,5,-2)$, $u_3=(-3,6,5,6,-5)$,
$v_1=(1,1,2,1,-1)$, $v_2=(0,-2,0,-1,-5)$, $v_3=(2,0,2,1,-3)$.

1.5　商　空　间

一般说来, 子空间有很多的补空间. 例如, 在 $\mathbb{R}^2$ 中, 过原点的直线只要斜率不为零, 就都是横轴的补空间. 由于补空间的维数都相同, 所以它们是同构的. 但这些同构映射的构造都依赖基选取. 不依赖基的选取的同构一般更能反映空间的内在性质. 对于向量空间的子空间, 可以构造一个与补空间维数相同且自然同构于每一个补空间的向量空间, 称为向量空间模子空间的**商空间**.

设 U 是向量空间 V 的子空间. 定义 V 中的二元关系 $\sim$ 如下:

$$x\sim y\Longleftrightarrow x-y\in U.$$

容易验证, 这个二元关系是等价关系, 即有

(1) 对称性: 如果 $x\sim y$, 则 $y\sim x$;

(2) 自反性: $x\sim x$ 对所有的 $x\in V$;

(3) 传递性: 如果 $x\sim y$, $y\sim z$, 那么 $x\sim z$.

这个等价关系的等价类全体记作 V/U. 向量 x 所在的等价类记作 $\bar{x}$. 容易看出

$$\bar{x}=x+U=\{x+u\,|\,u\in U\}.$$

定义等价类之间的加法和 V 的基域 K 与等价类的纯量乘如下:

$$\bar{x}+\bar{y}=\overline{x+y},$$

$$\alpha\bar{x} = \overline{\alpha x}.$$

容易验证, 在这两个运算下, V/U 是域 K 上的向量空间, 称为 V 模 U 的**商空间**. 商空间的零向量就是 V 中的零向量所在的等价类 $\bar{0} = U$.

在 $\mathbb{R}^2$ 中, 横轴的一个补空间就是一条过原点的斜率不为零的直线 M. 直线 M 所定义的等价关系的等价类就是平行于横轴的直线, 平行于横轴的直线全体是 M 所定义的等价关系的等价类全体. 每条平行于横轴的直线与 M 交于一点, 交点是这个等价类的自然的代表元.

定理 1.46 设 W 是 V 的子空间 U 的补空间, 那么映射 $f: W \to V/U$, $x \to \bar{x} = x + U$ 是线性同构.

证明 首先 f 保持加法和纯量乘

$$f(\alpha x + \beta y) = \alpha x + \beta y + U = \alpha(x + U) + \beta(y + U) = \alpha f(x) + \beta f(y).$$

因为 $V = W \oplus U$, 对任意的 $v \in V$, 存在 $x \in W$, $y \in U$ 使得 $v = x + y$. 于是

$$v + U = \bar{v} = \bar{x} = x + U = f(x),$$

所以 f 是满射. 对 $x, x' \in W$, 如果 $f(x) = \bar{x} = \bar{x}' = f(x')$, 那么 $x - x' \in U \cap W = 0$, 从而 $x = x'$. 所以 f 也是单射. □

推论 1.47 如果 U 是有限维向量空间 V 的子空间, 那么

$$\dim V/U = \dim V - \dim U.$$

证明 取 U 的一个补空间 W, 有 $\dim V - \dim U = \dim W = \dim V/U$. □

习 题 1.5

1. 设 U 是 V 的子空间. 证明: V 与 $U \oplus V/U$ 同构.

2. 对下列子空间 L 说明商空间 $K[t]/L$ 是否为有限维的:

(1) L 是所有次数不超过 $n-1$ 的多项式形成的子空间;

(2) L 是所有可以被 t^n 整除的多项式形成的子空间;

(3) L 是所有 t^2 的多项式形成的子空间.

3. 称向量空间 V 的子空间 U 是**余有限维的**如果商空间 V/U 是有限维的, 此时 U 的**余维数**定义为 $\operatorname{codim} U = \dim V/U$. 如果 U 和 W 都是 V 的余有限维子空间, 证明:

(1) $U + W$ 和 $U \cap W$ 也是余有限维子空间;

(2) $\operatorname{codim}(U + W) + \operatorname{codim}(U \cap W) = \operatorname{codim} U + \operatorname{codim} W$.

1.6 线性函数

一 齐次线性多项式 $a_1x_1 + \cdots + a_nx_n$ 可以看作是从向量空间 K^n 到基域的映射, 它把 K^n 中的向量 $(\xi_1, \cdots, \xi_n)$ 映到 $a_1\xi_1 + \cdots + a_n\xi_n$. 实数轴的闭区间 $[a, b]$

上的连续函数 $f(x)$ 可以作定积分 $\int_a^b f(x)dx$. 这个定积分可以看作是闭区间上的连续函数空间到实数域的映射. 域 K 上的多项式在 K 中一个元素的取值可以看作是多项式形成的空间到域 K 的映射. 方阵取迹可以看作是从域 K 上的 n 阶方阵构成的向量空间到 K 的映射.

这些映射都有一个共性: 某个域 K 上向量空间 V 到 K 的映射 α, 具有性质 ① $\alpha(u+v)=\alpha(u)+\alpha(v)$, ② $\alpha(au)=a\alpha(u)$, 其中 $u,v\in V$, $a\in K$. 上面那些例子足以说明具有这些性质的映射是重要的, 所以有专门的术语.

定义 1.48 设 V 是域 K 上的向量空间. 映射 $\alpha: V\to K$ 称为**线性函数**如果它具有如下性质:

$$\alpha(u+v)=\alpha(u)+\alpha(v),\quad \alpha(au)=a\alpha(u),$$

其中 $u,v\in V$, $a\in K$.

随意取定 V 的基 $e_1,\cdots,e_n$. 如果向量 $x\in V$ 在这个基下的坐标是 $x_1,\cdots,x_n$, 那么线性函数 α 在 x 处的值为

$$\alpha(x)=a_1x_1+\cdots+a_nx_n, \tag{1.6.7}$$

其中 $a_i=\alpha(e_i)$. 这表明线性函数完全由它在一个基的向量上的值确定. 反过来, 对于取定的基 $e_1,\cdots,e_n$ 和任意取定的纯量 $a_1,\cdots,a_n\in K$, 公式 (1.6.7) 定义了 V 上的一个线性函数, 它在基向量 e_i 处的值为 a_i.

向量空间 V 的线性函数全体是 V 的所有 K 值函数空间 K^V(参见例 1.5) 的子空间, 因为线性函数相加, 与 K 中元素相乘仍是线性函数. 为方便起见, 重述加与纯量乘的定义:

$$(\alpha+\beta)(v)=\alpha(v)+\beta(v),$$
$$(a\alpha)(v)=a(\alpha(v)).$$

定义 1.49 向量空间 V 的所有的线性函数形成的空间称为 V 的**对偶空间**, 记作 V^*.

设 $e_1,\cdots,e_n$ 是 V 的基, 定义 V 的线性函数 $e^1,\cdots,e^n$ 如下:

$$e^i(x_1e_1+\cdots+x_ne_n)=x_i.$$

这些线性函数就是关于基 $e_1,\cdots,e_n$ 的坐标函数. 它们的线性组合 $a_1e^1+\cdots+a_ne^n$ 在 e_i 处的值为 a_i, 由此立见 $e^1,\cdots,e^n$ 线性无关. 对于 $\alpha\in V^*$, 有

$$\alpha=\alpha(e_1)e^1+\cdots+\alpha(e_n)e^n.$$

所以这些坐标函数形成 V^* 的一个基. 上面的讨论证明了下面的定理.

定理 1.50 设 V 是域 K 上的 n 维向量空间, 那么它的对偶空间 V^* 也是 n 维的. 如果 $e_1, \cdots, e_n$ 是 V 的基, 那么满足下面条件的线性函数 $e^1, \cdots, e^n$,

$$e^i(e_j) = \delta_{ij} = \begin{cases} 1, & \text{如果} i = j, \\ 0, & \text{否则}, \end{cases}$$

构成 V^* 的基, 称为基 $e_1, \cdots, e_n$ 的**对偶基**.

对一些具体的向量空间, 有很多自然的线性函数.

例 1.51 设 $V = P_n = \langle 1, t, \cdots, t^{n-1} \rangle$ 是次数不超过 $n-1$ 的实多项式构成的向量空间. 多项式在实数 a 处取值给出了一个映射:

$$\theta_a : V \to \mathbb{R}, \quad f \to f(a).$$

易见它是线性函数. 只要 $a_1, \cdots, a_n \in \mathbb{R}$ 是互不相同的实数, $\theta_{a_1}, \cdots, \theta_{a_n}$ 就线性无关, 从而构成 V^* 的基. 实际上, 如果 $\theta_{a_1}, \cdots, \theta_{a_n}$ 的一个线性组合为 0, 那么它在 $1, t, \cdots, t^{n-1}$ 处的取值也为 0. 利用范德蒙德行列式知这个线性组合的系数都是 0.

利用导数, 可以方便地表达 $1, t, \cdots, t^{n-1}$ 的对偶基

$$e^i : f \to \frac{f^{(i)}(0)}{i!} = f_i, \quad i = 0, 1, \cdots, n-1,$$

其中 $f^{(i)}$ 是 f 的 i 阶导数. 基 $1, t-\lambda, \cdots, (t-\lambda)^{n-1}$ 的对偶基则是

$$e^i_\lambda : f \to \frac{f^{(i)}(\lambda)}{i!}, \quad i = 0, 1, \cdots, n-1.$$

更一般地, 对 $\lambda, \mu \in \mathbb{R}$, 映射 $f \to \mu f^{(i)}(\lambda)$ 是 V 上的线性函数.

例 1.52 设 $V = \langle \sin t,\ \sin 2t,\ \cdots,\ \sin nt \rangle$. 它是实数轴上的函数空间的子空间. 例 1.16 的方法表明 V 的维数是 n, $\sin t$, $\sin 2t$, $\cdots$, $\sin nt$ 是一个基, 对偶基 $e^1, \cdots, e^n$ 由如下公式确定:

$$e^i : \varphi \to \frac{1}{\pi} \int_0^{2\pi} \varphi \sin it \, dt.$$

二 有限维向量空间和它的对偶空间维数相同, 所以同构. 但同构映射的构造一般依赖基的选取, 没有自然的同构. 不过, 把 V^* 再做一次对偶, 得到的向量空间 $(V^*)^* = V^{**}$ 却是自然同构于 V.

对 $x \in V$, 定义 V^* 上的函数 f_x 如下:

$$f_x(\alpha) = \alpha(x).$$

从 V^* 上的加法和纯量乘可以看出这个函数是线性的.

定理 1.53 映射 $f: x \to f_x$ 是从空间 V 到空间 V^{**} 的同构.

证明 首先这个映射是线性的.

$$f_{x+y}(\alpha) = \alpha(x+y) = \alpha(x) + \alpha(y) = f_x(\alpha) + f_y(\alpha) = (f_x + f_y)(\alpha),$$
$$f_{\lambda x}(\alpha) = \alpha(\lambda x) = \lambda\alpha(x) = \lambda f_x(\alpha) = (\lambda f_x)(\alpha),$$

即 $f(x+y) = f(x) + f(y)$, $f(\lambda x) = \lambda f(x)$.

下面证明它是双射. 取 V 的基 $e_1, \cdots, e_n$, 它在 V^* 中的对偶基是 $e^1, \cdots, e^n$. 有

$$f_{e_i}(e^j) = e^j(e_i) = \delta_{ij}.$$

所以, $f_{e_1}, \cdots, f_{e_n}$ 是 V^{**} 的基, 对偶于 $e^1, \cdots, e^n$. 于是映射 f 把 V 的基 $e_1, \cdots, e_n$ 映到 V^{**} 的基 $f_{e_1}, \cdots, f_{e_n}$. 由此可知 f 是双射. □

定理 1.53 中的同构是自然的, 即其构造不依赖基的选取. 因此, 通过这个同构可以把 V 与 V^{**} 等同起来, 即 V 中的元素可以看作 V^* 的线性函数. 这样一来, V 和 V^* 的地位就对称了. 一个有用的推论是

推论 1.54 空间 V^* 的每一个基都是 V 的某个基的对偶基.

证明 设 $f_1, \cdots, f_n$ 是 V^* 的基, 那么在 V^{**} 中有它的对偶基. 等同 V^{**} 与 V 后, 这个对偶基成为 V 的某个基 $e_1, \cdots, e_n$, 且有等式: $f_i(e_j) = \delta_{ij}$. □

注 对无限维空间, 定理 1.50 和定理 1.53 不成立, 因为此时 V^* 和 V^{**} 比 V 大得多. 例如, 考虑实数域上的多项式环 $V = \mathbb{R}[t]$. 这个空间的维数是可数的. 对每一个实数无限序列 $a = (a_0, a_1, \cdots, a_n, \cdots)$, 可唯一定义 V 的线性函数 ε_a 使得 $\varepsilon_a(t^i) = a_i$. 反之, 每一个线性函数也唯一确定了这样一个无限序列. 于是 V^* 与所有这样的无限序列形成的空间 M 同构. 可以验证 M 的维数是不可数的. 所以 V^* 与 V 不会同构, V^{**} 就更不会与 V 同构了.

三 空间上的线性函数可以用于讨论向量的线性无关和线性相关性.

引理 1.55 如果 V 中的向量 $v_1, \cdots, v_m$ 线性相关, 那么对 V 上的任意的线性函数 $\alpha_1, \cdots, \alpha_m$, 都有

$$\det(\alpha_i(v_j)) = 0, \quad 1 \leqslant i, j \leqslant m.$$

证明 因为 $v_1, \cdots, v_m$ 线性相关, 其中的某个向量 v_k 是其余向量的线性组合. 于是矩阵 $(\alpha_i(v_j))$ 的第 k 列就是其他列的线性组合, 从而矩阵是退化的. □

引理 1.56 如果 $\alpha_1, \cdots, \alpha_n$ 是 V^* 的基, 那么 V 中的向量 $v_1, \cdots, v_n$ 线性无关当且仅当

$$\det(\alpha_i(v_j)) \neq 0, \quad 1 \leqslant i, j \leqslant n.$$

证明 根据引理 1.55, 引理中的行列式如果不为 0, 那么 $v_1,\cdots,v_n$ 线性无关. 现证向量组线性无关蕴含行列式不为 0. 根据推论 1.54, $\alpha_1,\cdots,\alpha_n$ 对偶于 V 的某个基 $e_1,\cdots,e_n$. 命 $a_{1j},\cdots,a_{nj}$ 为 v_j 在这个基下的坐标, 那么

$$(\alpha_i(v_j))=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

就是从基 $e_1,\cdots,e_n$ 到基 $v_1,\cdots,v_n$ 的转换矩阵. 根据定理 1.34, 这个转换矩阵是可逆的, 从而行列式不为 0. □

定理 1.57 设 $\alpha_1,\cdots,\alpha_n$ 是 V^* 的基. 那么 V 中的向量组 $v_1,\cdots,v_m$ 的秩等于 $n\times m$ 矩阵 $(\alpha_i(v_j))$ 的秩, 也就是所有形如

$$\det(\alpha_i(v_j))\neq 0,\qquad 1\leqslant i=i_1,\cdots,i_k\leqslant n,\ \ 1\leqslant j=j_1,\cdots,j_k\leqslant m \tag{1.6.8}$$

的非零行列式的最大阶数.

证明 设向量组的秩为 r. 那么向量组中任何 $r+1$ 个向量线性相关, 由引理 1.55 知如果 $k>r$, 则定理中的行列式等于 0.

接下来证明存在形如 (1.6.8) 的非零行列式, 其阶为 r. 设 $v_{j_1},\cdots,v_{j_r}$ 线性无关. 用记号 $u_1,\cdots,u_r$ 表示这 r 个向量, 并扩充成 V 的基 $u_1,\cdots,u_n$. 由引理 1.56 知 n 阶方阵 $(\alpha_i(u_j))$ 非退化, 所以它前面的 r 列向量线性无关, 它们组成的矩阵 $(\alpha_i(u_j))$, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant r$, 的秩为 r. 于是这个 $n\times r$ 矩阵有一个 r 阶子式不为 0, 且具有 (1.6.8) 的形式. □

四 空间 V 的子空间和它的对偶 V^* 的子空间之间有一个自然的一一对应: V 的 k 维子空间对应到 V^* 的 $n-k$ 维子空间 (这里 $n=\dim V$).

定义 1.58 设 U 是 V 的子空间, 其**零化子**定义为 V^* 的子空间

$$U^\circ=\{\alpha\in V^*\,|\,\alpha(u)=0,\ \forall u\in U\}.$$

定理 1.59 $\dim U^\circ=\dim V-\dim U$.

证明 取 V 的基 $e_1,\cdots,e_n$ 使得 $U=\langle e_1,\cdots,e_k\rangle$, 所取基的对偶基记为 $e^1,\cdots,e^n$, 那么 U° 由 $e^{k+1},\cdots,e^n$ 张成. □

对 V^* 的子空间 W, 其零化子 W° 是 V^{**} 的子空间. 等同 V^{**} 与 V 后, W° 就是 V 的子空间

$$W^\circ=\{v\in V\,|\,\alpha(v)=0,\ \forall\alpha\in W\}.$$

定理 1.60 对 V 的每一个子空间 U 有 $(U^\circ)^\circ=U$.

证明　沿用定理 1.59 证明的记号, 显然有 $(U^\circ)^\circ = \langle e_1, \cdots, e_k \rangle$. □

推论 1.61　V 的每个子空间都是 V^* 的某个子空间的零化子. 反之也对, 即 V^* 的每个子空间都是 V 的某个子空间的零化子.

五　齐次线性方程组解的几何解释　带有 n 个未知元的齐次线性方程组的解可以看作基域 K 上的列 (或行) 空间 K^n 的向量, 方程组的解集是 K^n 的子空间. 齐次线性多项式 $a_1x_1 + \cdots + a_nx_n$ 可以看作 K^n 上的线性函数. 于是齐次线性方程组可以写成

$$\alpha_1(x) = 0, \ \cdots, \ \alpha_m(x) = 0,$$

其中 $x \in K^n$, $\alpha_1, \cdots, \alpha_m \in (K^n)^*$.

把 K^n 换成一般的向量空间 V. 对 V 上的线性函数 $\beta_1, \cdots, \beta_m$, 也可以考虑方程组

$$\beta_1(x) = 0, \ \cdots, \ \beta_m(x) = 0. \tag{1.6.9}$$

方程组的解是 V 中的向量, 解集是 V 的子空间, 称为方程组的**解空间**. 取定 V 的一个基, 那么该方程组就成为向量坐标的线性方程组, 其形态是我们熟悉的. 下面的结论是定理 1.59 和推论 1.61 的另一种表述.

定理 1.62　(1) 如果向量组 $\beta_1, \cdots, \beta_m$ 的秩为 r, 那么方程组 (1.6.9) 的解空间 $U \subset V$ 的维数是 $n - r$(这里 $n = \dim V$).

(2) V 的每一个子空间都是某个形如 (1.6.9) 的方程组的解空间.

习　题　1.6

1. 证明: 对 n 维向量空间的非零线性函数 f, 存在基 $e_1, \cdots, e_n$ 使得对基域中任意的纯量 $x_1, \cdots, x_n$, 有

$$f(x_1e_1 + \cdots + x_ne_n) = x_1.$$

2. 证明: n 维向量空间的任何 k 维子空间都是 $n-k$ 个线性函数的核的交.

3. 证明: n 维向量空间的 n 个线性函数线性无关当且仅当它们的核的交是零子空间.

4. 设 f 和 g 是向量空间 V 的两个线性函数. 证明: 如果对所有的向量 $x \in V$ 都有 $f(x)g(x) = 0$, 那么这两个线性函数中有一个是零函数.

5. 设 $\alpha(t)$ 是实数域上的无限次可微函数. 说明下面的函数是否为实数域上所有无限次可微函数形成的空间 $C^\infty(\mathbb{R})$ 上的线性函数:

(1) $f(u) = \int_a^b \alpha(t)u(t)dt$;

(2) $f(u) = \int_a^b \alpha(t)[u(t)]^2 dt$;

(3) $f(u) = \int_a^b \alpha(t)u(t^2)dt$;

(4) $f(u)=\dfrac{d^3}{dt^3}[\alpha(t)u(t)]\Big|_{t=a}$.

6. 命 $V=M_n(K)$ 是域 K 上所有 n 阶方阵形成的空间. 证明: 对 V 上的每个线性函数 f 都能找到唯一的方阵 A 使得 $f(X)=\operatorname{tr}(AX)$.

7. 证明: 如果一个向量空间的两个线性函数的核相同, 则这两个线性函数成比例.

1.7 双线性型和二次型

一 多重线性映射 对 $\alpha\in V^*$ 和 $v\in V$, 用记号 (α,v) 代替 $\alpha(v)$. 这个记号意指可以把 v 和 α 都看作变量, 从而得到 V^* 和 V 的笛卡尔积 $V^*\times V$ 上的一个自然的函数

$$(\ ,\ ):V^*\times V\to K,\quad (\alpha,v)=\alpha(v).$$

这个函数对每个变量都是线性的.

$$(a\alpha+b\beta,v)=a(\alpha,v)+b(\beta,v),\quad (\alpha,au+bv)=a(\alpha,u)+b(\alpha,v).$$

这样的函数称为双线性函数. $\mathbb{R}^3\times\mathbb{R}^3$ 上的函数

$$(x,y)=x_1y_1+x_2y_2+x_3y_3$$

就是 $\mathbb{R}^3$ 上的**点积**, 也是双线性函数. 一般地, 对域 K 上的向量空间 $V_1,\cdots,V_p,U$, 映射

$$f:V_1\times V_2\times\cdots\times V_p\to U$$

称为**多重线性的**(在此处也说 p**重线性的**) 如果映射对每个变量都是线性的, 即对任意指标 i 和固定的向量 $v_j\in V_j$, $1\leqslant j\leqslant p$, $j\neq i$, 映射

$$f_i:V_i\to U,\quad x\to f(v_1,\cdots,v_{i-1},x,v_{i+1},\cdots,v_n)$$

是线性的:

$$f_i(ax+by)=af_i(x)+bf_i(y),\quad \forall x,y\in V_i,\quad a,b\in K.$$

如果 $p=1$, 则简单称作**线性映射**. 如果 $U=K$, 多重线性映射一般称为**多重线性函数**, 或 p**重线性函数**如果笛卡尔积中的因子数是 p 个.

例 1.63 当 $V_1=\cdots=V_p=U=K$ 时, $f(x_1,\cdots,x_p)=x_1\cdots x_p$ 是 p 重线性函数.

例 1.64 如果 α_i 是 V_i 上的线性函数, 那么

$$f(v_1,\cdots,v_p)=\alpha_1(v_1)\cdots\alpha_p(v_p)$$

是 $V_1\times V_2\times\cdots\times V_p$ 上的多重线性函数.

当 $V_1=\cdots=V_p=V$ 时, $V^p=V\times\cdots\times V$ (p 个因子) 上的多重 (或 p 重) 线性映射 (函数) 一般简称为 V 上的多重 (或 p 重) 线性映射 (函数).

例 1.65　行列式函数 $\det: M_n(K)\to K$ 可以看作矩阵的行 (或列) 的函数. 此时 $\det$ 就是 K^n 上的 n 重线性函数, 而且具有斜对称性.

称 V 上 p 重线性映射 (函数) 为**对称的**如果对任意的向量 $v_1,v_2,\cdots,v_p\in V$ 和任意的置换 $\pi\in S_p$, 有

$$f(v_{\pi(1)},v_{\pi(2)},\cdots,v_{\pi(p)})=f(v_1,v_2,\cdots,v_p).$$

如果

$$f(v_{\pi(1)},v_{\pi(2)},\cdots,v_{\pi(p)})=\varepsilon_\pi f(v_1,v_2,\cdots,v_p),$$

则称 f 为**斜对称的**(**或交错的**), 这里 ε_π 是置换的符号.

所有的多重线性映射 $f:V_1\times V_2\times\cdots\times V_p\to U$ 形成的集合记作 $\mathcal{L}(V_1,V_2,\cdots,V_p;U)$ 或 $\mathrm{Hom}\,(V_1,V_2,\cdots,V_p;U)$, 它是 U^X 的子空间 (参见例 1.6), 其中 $X=V_1\times V_2\times\cdots\times V_p$. 当 $V_1=V_2=\cdots=V_p=V$ 时, 多重线性映射空间也记为 $\mathcal{L}_p(V,U)$.

二　双线性型　多重线性函数中最简单的情况是特别有意思的, 这时 $p=2$, $V_1=V_2=V$. 空间 V 上的二重线性函数 $f:V\times V\to K$ 称为 V 上的**双线性型**(或**双线性函数**). 双线性型 f 的定义性质是

$$\begin{aligned} f(au+bv,w)&=af(u,w)+bf(v,w),\\ f(w,au+bv)&=af(w,u)+bf(w,v), \end{aligned}\tag{1.7.10}$$

其中 $u,v,w\in V$, $a,b\in K$.

例 1.66　空间 $\mathbb{R}^3$ 上的点积是双线性函数.

例 1.67　函数

$$f(\varphi,\psi)=\int_a^b\varphi(x)\psi(x)dx$$

是闭区间 $[a,b]$ 上的连续函数空间 $C(a,b)$ 上的双线性型.

例 1.68　函数

$$f(A,B)=\mathrm{tr}\,AB$$

是空间 $M_n(K)$ 上的双线性型.

设 $e_1,\cdots,e_n$ 是 V 的基. 对 V 中的向量

$$x=x_1e_1+\cdots+x_ne_n,\quad y=y_1e_1+\cdots+y_ne_n,$$

利用性质 (1.7.10), 得

$$\begin{aligned} f(x,y) &= f\left(\sum_i x_i e_i, \sum_j y_j e_j\right) = \sum_i x_i f\left(e_i, \sum_j y_j e_j\right) \\ &= \sum_i x_i \sum_j y_j f(e_i, e_j) = \sum_{i,j} f_{ij} x_i y_j, \end{aligned} \tag{1.7.11}$$

其中 $f_{ij} = f(e_i, e_j)$.

矩阵 $F = (f_{ij})$ 称为双线性型f **在基** $e_1, \cdots, e_n$ **下的矩阵**. 命 $X = [x_1, \cdots, x_n]$ 和 $Y = [y_1, \cdots, y_n]$ 分别是 x 和 y 在基 $e_1, \cdots, e_n$ 下的坐标列向量, 那么公式 (1.7.11) 可以简洁地写成

$$f(x,y) = {}^t XFY. \tag{1.7.12}$$

公式 (1.7.11) 表明双线性型由它的矩阵完全确定. 并且, 映射

$$f \rightarrow (f_{ij})$$

是 V 上的双线性型空间 $\mathcal{L}_2(V,K)$ 到 n 阶方阵空间 $M_n(K)$ 的线性同构. 实际上, 对 $f, g \in \mathcal{L}_2(V,K)$, 如果

$$f(x,y) = {}^t XFY, \quad g(x,y) = {}^t XGY,$$

那么

$$(af + bg)(x,y) = {}^t X(aF + bG)Y.$$

三　双线性型理论的主要目标是寻找适当的基使得双线性型的矩阵尽可能简单, 从而双线性型可以通过公式 (1.7.11) 方便地计算. 为此需要考虑双线性型在不同基下的矩阵的联系, 得到这些矩阵的共有性质, 这些性质与基的选取无关.

如果从基 $e_1, \cdots, e_n$ 到基 $e'_1, \cdots, e'_n$ 的转换矩阵是 A, 向量 x 和 y 在新基下的坐标列向量分别是 X' 和 Y', 则有

$$X = AX', \quad Y = AY'.$$

代入公式 (1.7.11), 得

$$f(x,y) = {}^t X' \cdot {}^t AFA \cdot Y'.$$

命 F' 为双线性型 f 在基 $e'_1, \cdots, e'_n$ 下的矩阵, 则有 $f(x,y) = {}^t X'F'Y'$. 比较 $f(x,y)$ 的两个公式, 我们得到

定理 1.69　双线性型 f 在基 (e_i) 和基 (e'_i) 下的矩阵 F 和 F' 之间有如下的联系:

$$F' = {}^t AFA, \tag{1.7.13}$$

其中 A 是从 (e_i) 到 (e'_i) 的转换矩阵.

双线性型在不同基下的联系给出了方阵之间的一个等价关系, 称为合同关系.

定义 1.70 方阵 $F, F' \in M_n(K)$ 称为**合同的**如果存在可逆矩阵 $A \in M_n(K)$ 使得 $F' = {}^tAFA$.

根据第一卷推论 3.28, 合同的矩阵有相同的秩. 定理 1.69 表明双线性型 f 在不同基下的矩阵是合同的, 它们的秩相等, 从而这些矩阵的秩是 f 的一个内在的量, 称为 f 的**秩**, 记作 $\operatorname{rank} f$. 双线性型 f 称为**非退化的**如果它在某个基 (或说任意基) 下的矩阵是非退化的, 即 $\operatorname{rank} f = \dim V$. 双线性型的秩还有另一个解释.

定义 1.71 双线性型 f 的**左核** (或说**左根**) 定义为

$$L_f = \{x \in V \mid f(x, y) = 0, \forall y \in V\}.$$

显然 L_f 是 V 的子空间, 其定义仅依赖 f, 不依赖 V 的任何基. 如果 $e_1, \cdots, e_n$ 是 V 的基, 那么 $x \in L_f$ 等价于

$$f(x, e_1) = 0, \ \cdots, \ f(x, e_n) = 0.$$

这是由线性函数 $f_j : x \to f(x, e_j)$, $j = 1, \cdots, n$ 决定的方程组. 根据定理 1.62(1), 方程组的解空间 L_f 的维数等于

$$\dim V - \operatorname{rank}\{f_1, \cdots, f_n\}.$$

运用定理 1.57 知 V^* 中的向量组 $f_1, \cdots, f_n$ 的秩就是矩阵 $F = (f_{ij})$ 的秩. (在定理 1.57 中用 V^* 代替 V, 注意 V^{**} 等同于 V, 从而 $v \in V$ 看作 V^* 的线性函数: $f \to f(v)$, 即 $v(f) = f(v)$. 于是 $e_i(f_j) = f_j(e_i, e_j) = f_{ij}$.) 由此可见 $\operatorname{rank} F = \dim V - \dim L_f$ 不依赖基的选取.

练习 1.72 (1) 定义双线性型的右核并证明右核的维数等于左核的维数; (2) 双线性型非退化当且仅当其左核 (或右核) 为零空间.

四 对称双线性型与斜对称双线性型 最重要的双线性型是对称双线性型和斜对称双线性型, 它们是第一部分定义的对称多重线性函数和斜对称多重线性函数的特例. 称双线性型 $f : V \times V \to K$ 是**对称的**如果对 V 的任意向量 x, y 有 $f(x, y) = f(y, x)$; 是**斜对称的**如果 $f(x, y) = -f(y, x)$.

例 1.66 和例 1.67 中的双线性型是对称的, 例 1.68 中的双线性型也是对称的因为 $\operatorname{tr} AB = \operatorname{tr} BA$. 二阶方阵的行列式函数可以看作 K^2 上的双线性型, 它是斜对称的.

回忆方阵 $A = (a_{ij})$ 称为对称的如果对所有的 i, j 有 $a_{ij} = a_{ji}$(即 ${}^tA = A$), 称为斜对称的如果 $a_{ij} = -a_{ji}$(即 ${}^tA = -A$). 从定义立见 (斜) 对称双线性型在任何基下的矩阵都是 (斜) 对称的.

反之, 假设 f 在某个基下的矩阵 A 是 (斜) 对称的, 那么 ${}^tA = \varepsilon A$, 其中 $\varepsilon = 1$ 或 -1. 注意作为 1×1 矩阵有 $f(x,y) = {}^tf(x,y)$. 于是

$$\begin{aligned} f(x,y) &= {}^tXAY = {}^tf(x,y) = {}^t({}^tXAY) \\ &= {}^tY \cdot {}^tA \cdot {}^t({}^tX) = \varepsilon {}^tYAX \\ &= \varepsilon f(y,x). \end{aligned}$$

从而 f 是 (斜) 对称的. 所以双线性型是 (斜) 对称的当且仅当它在任意基下的矩阵是 (斜) 对称的.

如果域 K 的特征不等于 2, 任意的方阵都可以写成对称方阵与斜对称方阵的和

$$A = \frac{1}{2}(A + {}^tA) + \frac{1}{2}(A - {}^tA).$$

由此可以想到此时每一个双线性型都是对称双线性型与斜对称双线性型的和

$$f(x,y) = {}^tXAY = {}^tX \cdot \frac{1}{2}(A + {}^tA) \cdot Y + {}^tX \cdot \frac{1}{2}(A - {}^tA) \cdot Y.$$

当然这个事实也可以直接验证:

$$f(x,y) = \frac{1}{2}[f(x,y) + f(y,x)] + \frac{1}{2}[f(x,y) - f(y,x)].$$

如果 f 既是对称的也是斜对称的, 则有 $f(x,y) = f(y,x) = -f(x,y)$, 于是 $2f(x,y) = 0$. 因为 $\operatorname{char} K \neq 2$, 所以对任意的 $x, y \in V$ 有 $f(x,y) = 0$, 即 $f = 0$. 我们把上面的讨论总结为

定理 1.73 如果域 K 的特征不等于 2, 那么 K 上的向量空间的双线性型空间 $\mathcal{L}_2(V,K)$ 是对称双线性型子空间 $\mathcal{L}_2^+(V,K)$ 和斜对称双线性型子空间 $\mathcal{L}_2^-(V,K)$ 的直和:

$$\mathcal{L}_2(V,K) = \mathcal{L}_2^+(V,K) \oplus \mathcal{L}_2^-(V,K).$$

注 当 $\operatorname{char} K = 2$ 时, $-1 = 1$, 从而对双线性型, 斜对称与对称是一回事. 矩阵 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 不是对称的, 从而它对应的双线性型不是对称的. 这说明定理 1.73 在特征为 2 的域上是不成立的. 特征为 2 的域上的双线性型 f 称为**交错的**如果 $f(x,x) = 0, \forall x \in V$.

以下总是假设 K 的特征不等于 2.

五 二次型 双线性型是讨论二次型的一个合适工具, 二次型广泛出现在数学不同的分支中.

定义 1.74 设 V 是域 K 上的有限维向量空间, f 是 V 上的双线性型. 定义函数 $q: V \to K$ 如下

$$q(x) = f(x, x).$$

称它为相伴于 f 的**二次型**(或**二次函数**), 也说它是 f 给出的二次型.

取定 V 的一个基, f 在基下的矩阵是 $F = (f_{ij})$, 向量 x 的坐标列向量是 $X = [x_1, \cdots, x_n]$. 那么

$$q(x, x) = \sum_{i,j} f_{ij} x_i x_j. \tag{1.7.14}$$

这是一个次数为 2 的齐次多项式. 由于 $x_i x_j = x_j x_i$, 所以 $x_i x_j$ 的系数是 $f_{ij} + f_{ji}$ 如果 $i \neq j$, 而 x_i^2 的系数是 f_{ii}. 由此可见, 不同的双线性型可以给出相同的二次型, 比如以 ${}^tF$ 为矩阵的双线性型和 f 给出的二次型相等.

在给出同一个二次型 q 的双线性型中, 只有一个是对称的, 称为 q 的**极化** (polarization), 由下式定义:

$$f_q(x, y) = \frac{1}{2}[q(x + y) - q(x) - q(y)], \tag{1.7.15}$$

如果 $q(x) = f(x, x)$, 则有 $f_q(x, y) = \dfrac{1}{2}[f(x, y) + f(y, x)]$, 且 $f_q(x, x) = f(x, x) = q(x)$.

于是, 在二次型和对称双线性型之间有一一对应: $q \to f_q$. 这样, 双线性型的有关概念就可以运用到二次型上.

定义 1.75 设 q 是对称双线性型 f 给出的二次型, 那么, 二次型 q 的**秩**定义为 f 的秩: $\operatorname{rank} q = \operatorname{rank} f$. 二次型 q 在空间 V 的基 $e_1, \cdots, e_n$ 的**矩阵**定义为 f 在这个基下的矩阵 F.

六 二次型的典范式 如果对称双线性型 f 在某个基下的矩阵是对角的, 那么 f 和它给出的二次型就有如下整齐简洁的形式:

$$f(x, y) = f_1 x_1 y_1 + \cdots + f_n x_n y_n, \tag{1.7.16}$$

$$q(x) = f_1 x_1^2 + \cdots + f_n x^n. \tag{1.7.17}$$

表达式 (1.7.16) 和 (1.7.17) 分别称为双线性型 f 和二次型 q 的一个**典范式**(canonical form, 也译成典型式、规范式或规范型), 相应的基称为 f 和 q 的**典范基**.

在典范式中, 二次型 q 的秩就是非零系数的个数. 从第二部分的结尾处知 $\operatorname{rank} q = \dim V - \dim L_q$, 其中 $L_q = L_f$ 是 f 的核 (因 f 对称, 其左核与右核相同). 子空间 $L_q \subset V$ 也称为二次型 q 的**核**, 可以用 q 刻画如下:

$$L_q = \{u \in V \mid q(u + v) = q(u) + q(v), \forall v \in V\}.$$

定理 1.76 空间 V 上的每个对称双线性型 f 都有典范基.

证明 对 V 的维数 n 做归纳法. 当 n 为 1 时, 结论是显然的. 假设 $n-1$ 维的空间上的对称双线性型都有典范基. 如果 $f=0$, 那么 V 的任意基都是 f 的典范基. 现设 $f\neq 0$. 从公式 (1.7.15) 知它给出的二次型 q 不为 0. 于是存在 $e_1\in V$ 使得 $q(e_1)=f(e_1,e_1)\neq 0$. 从而线性函数 $\alpha: x\to f(x,e_1)$ 不是零函数. 根据定理 1.62(1), 子空间

$$L=\ker\alpha=\{x\in V\mid \alpha(x)=0\}$$

的维数是 $n-1$. 根据归纳假设, f 在 L 上的限制有典范基 $e_2,\cdots,e_n$. 于是, 对向量组 $e_1,\cdots,e_n$ 有

$$f(e_i,e_j)=0,\quad \text{如果} 1\leqslant i\neq j\leqslant n.$$

如果这组向量线性无关, 它就是 f 的一个典范基.

根据练习 1.19(2), 要说明这组向量线性无关, 只要说明 e_1 不是 $e_2,\cdots,e_n$ 的线性组合. 假设不然, 则有 $e_1=a_2e_2+\cdots a_ne_n$. 于是 $0\neq\alpha(e_1)=a_2\alpha(e_2)+\cdots+a_n\alpha(e_n)=0$. 这个矛盾说明向量组 $e_1,\cdots,e_n$ 线性无关. □

推论 1.77 如果域 K 上的向量空间 V 的二次型 q 的秩为 r, 那么存在 V 的基使得 q 有如下的典范式:

$$q(x)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2.\tag{1.7.18}$$

推论 1.78 每个对称矩阵都合同于某个对角矩阵.

在某些情况下, 可以通过对称双线性型 f 在一个基下的矩阵直接写出相应的二次型 q 的典范式. 设 $e_1,\cdots,e_n$ 是 V 的基, f 在该基下的矩阵为 $F=(f_{ij})$. 矩阵 F 的 k 阶**主子式**定义为

$$\Delta_k=\begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1k}\\ f_{21} & f_{22} & \cdots & f_{2k}\\ \vdots & \vdots & & \vdots\\ f_{k1} & f_{k2} & \cdots & f_{kk}\end{vmatrix},\quad 1\leqslant k\leqslant n.$$

命 $\Delta_0=1$.

定理 1.79(雅可比方法) 设 f 是空间 V 上的对称双线性型, F 是它在某个基下的矩阵. 如果 F 的主子式 $\Delta_1,\cdots,\Delta_n$ 均不为 0, 那么存在 V 的基使得 f 和它给出的二次型 q 在这个基下有典范式

$$f(x,y)=\frac{\Delta_1}{\Delta_0}x_1y_1+\frac{\Delta_2}{\Delta_1}x_2y_2+\cdots+\frac{\Delta_n}{\Delta_{n-1}}x_ny_n.\tag{1.7.19}$$

$$q(x)=\frac{\Delta_1}{\Delta_0}x_1^2+\frac{\Delta_2}{\Delta_1}x_2^2+\cdots+\frac{\Delta_n}{\Delta_{n-1}}x_n^2.\tag{1.7.20}$$

证明 两个公式是等价的, 我们证明第二个公式. 当 V 的维数为 1 时, 结论显然成立. 对维数 n 作归纳法. 设 F 是 f 在基 $e_1,\cdots,e_n$ 下的矩阵. 命 $L=\langle e_1,\cdots,e_{n-1}\rangle$, $q'=q|_L$ 是 q 在 L 上的限制. 那么 q' 在基 $e_1,\cdots,e_{n-1}$ 下的矩阵 F' 就是 F 去掉最后一行与最后一列得到的矩阵. 于是 F' 的主子式是 $\Delta_1,\cdots,\Delta_{n-1}$. 由归纳假设, 有 L 的基 $e_1',\cdots,e_{n-1}'$ 使得在这个基下对 $x=x_1e_1'+\cdots+x_{n-1}e_{n-1}'\in L$, $q'(x)$ 有典范式

$$q'(x)=q(x)=\frac{\Delta_1}{\Delta_0}x_1^2+\frac{\Delta_2}{\Delta_1}x_2^2+\cdots+\frac{\Delta_{n-1}}{\Delta_{n-2}}x_{n-1}^2.$$

需要找出 e_n' 使得 $e_1',\cdots,e_{n-1}',e_n'$ 是 V 的基, 且在这个基下 q 有要求的典范式. 考虑线性方程组

$$f(x,e_1')=0,\cdots,\ f(x,e_{n-1}')=0,\quad x\in V.$$

方程组涉及的线性函数只有 $n-1$ 个, 根据定理 1.62(1), 解空间的维数大于或等于 $n-(n-1)=1$. 取方程组的非零解 e_n'. 由于

$$f(e_i',e_i')=\frac{\Delta_i}{\Delta_{i-1}},\quad f(e_i',e_j')=0,\quad 1\leqslant i\neq j\leqslant n-1,$$

所以 e_n' 不是 $e_1',\cdots,e_{n-1}'$ 的线性组合, 从而 $e_1',\cdots,e_{n-1}',e_n'$ 线性无关, 即这些向量构成 V 的基. 可以选取 e_n'(否则用 e_n' 的适当倍数代替 e_n') 使得从基 (e_i) 到基 (e_i') 的转换矩阵 A 的行列式的值为 1.

设 $F'=(f_{ij}')$ 是 f 在基 (e_i') 下的矩阵. 当 $i\neq j$ 时, $f_{ij}'=f(e_i',e_j')=0$, 所以

$$\begin{aligned}\det F'&=f(e_1',e_1')\cdots f(e_n',e_n')=\Delta_{n-1}f(e_n',e_n')\\&=\det({}^tAFA)=(\det A)^2\det F=\Delta_n.\end{aligned}$$

于是

$$f(e_n',e_n')=\frac{\Delta_n}{\Delta_{n-1}}.$$

从而对 $x=x_1e_1'+\cdots+x_ne_n'\in V$, 二次型的值 $q(x)$ 由公式 (1.7.20) 给出. □

容易看出矩阵 A 是上三角形的:

$$\begin{aligned}e_1'&=a_{11}e_1,\\e_2'&=a_{12}e_1+a_{22}e_2,\\&\cdots\cdots\\e_n'&=a_{1n}e_1+a_{2n}e_2+\cdots+a_{nn}e_n.\end{aligned}$$

但我们并不需要这个事实.

七　配方法　二次型是向量坐标的二次函数. 可以直接通过配方法把二次函数化成仅含平方项的二次函数, 这就是二次型的一个典范式. 设

$$q(x_1,\cdots,x_n):=q(x)=\sum_{i,j=1}^{n}f_{ij}x_ix_j.$$

把含 x_1 的项合在一起, 得

$$q(x_1,\cdots,x_n)=f_{11}x_1^2+2f_{12}x_1x_2+2f_{13}x_1x_3+\cdots+2f_{1n}x_1x_n+\sum_{i,j\neq 1}^{n}f_{ij}x_ix_j.$$

(注意 $f_{ij}=f_{ji}$, 所以当 $j\neq 1$ 时, x_1x_j 的系数为 $f_{1j}+f_{j1}=2f_{1j}$.)

如果 $f_{11}\neq 0$, 可以把含 x_1 的项配成平方项

$$q(x_1,\cdots,x_n)=\frac{1}{f_{11}}(f_{11}x_1+f_{12}x_2+f_{13}x_3+\cdots+f_{1n}x_n)^2+\sum_{i,j\neq 1}^{n}f'_{ij}x_ix_j.$$

命

$$x'_1=f_{11}x_1+f_{12}x_2+f_{13}x_3+\cdots+f_{1n}x_n,\quad x'_2=x_2,\ \cdots,\ x'_n=x_n.$$

二次型 q 就变成了

$$q(x_1,\cdots,x_n)=\frac{1}{f_{11}}{x'}_1^2+q_1(x'_2,\cdots,x'_n),$$

其中 $q_1(x'_2,\cdots,x'_n)=\sum\limits_{i,j=2}^{n}f'_{ij}x'_ix'_j$ 是 $n-1$ 维空间上的二次型 (或说 $n-1$ 个变元的二次函数). 如果 $f'_{22}\neq 0$, 通过变量替换

$$x''_1=x'_1,\quad x''_2=f'_{22}x'_2+f'_{23}x'_3+\cdots+f'_{2n}x'_n,\quad x''_3=x'_3,\cdots,x''_n=x'_n,$$

就得到如下的表达式:

$$q(x_1,\cdots,x_n)=\frac{1}{f_{11}}{x''}_1^2+\frac{1}{f'_{22}}{x''}_2^2+q_2(x''_3,\cdots,x''_n),$$

其中 $q_2(x''_2,\cdots,x''_n)=\sum\limits_{i,j=3}^{n}f''_{ij}x''_ix''_j$ 是 $n-2$ 维空间上的二次型 (或说 $n-2$ 个变元的二次函数).

如此下去, 就可以把 $q(x)$ 化成只含变元平方项的二次函数. 变量替换实际上就是从向量空间一个基转到另一个基的向量坐标变换, 坐标变换和基变换之间的联系由定理 1.34 给出. 需要说明的是如果 $f_{11}=0$, 该怎么做变换. 如果某个 $f_{kk}\neq 0$,

互换 x_1 和 x_k 的指标 (即变动相应基向量的顺序), 就可以像上面那样做变换了. 如果 $q \neq 0$ 但所有的 f_{kk} 都为 0, 则有某些 $2f_{ij}$ 不为 0, 不妨设 $i=1$, $j=2$. 做变换

$$x_1 = x_1' + x_2',\quad x_2 = x_1' - x_2',\quad x_i = x_i',\quad i \geqslant 3.$$

二次型 q 的表达式中 ${x'}_1^2$ 的系数是 $2f_{12}$, 前面的方法就可以用了.

八 在二次型的典范式

$$q(x) = f_1x_1^2 + \cdots + f_nx_n^2$$

中, 诸系数 f_i 显然不是唯一的. 例如, 取非零元素 $a_i \in K$, 把 x_i 换成 a_ix_i'(相当于典范基向量 e_i 换成它的 a_i 倍), 得到 q 的另一个典范式

$$q(x) = f_1a_1^2{x'}_1^2 + \cdots + f_na_n^2{x'}_n^2.$$

当变动典范基向量的顺序后, 典范式的变量也变动顺序, 从而变量的系数 (顺序) 又发生了变化. 典范式的系数的变化范围取决于基域.

如果 $K=\mathbb{C}$. 适当选取典范基向量的倍数, 就可以让诸 f_i 取值 0 或 1, 再适当变动基向量的顺序, 二次型的典范式就可以取成如下的**标准形**(normal form):

$$q(x) = x_1^2 + \cdots + x_r^2.$$

如果 $K=\mathbb{R}$. 适当选取典范基向量的倍数, 就可以让诸 f_i 取值 0 或 ±1, 再适当变动基向量的顺序, 二次型的典范式就可以取成如下的**标准形**:

$$q(x) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2. \tag{1.7.21}$$

在上面这个标准形中, r 是二次型的秩, 不依赖基的选取. 那 s 呢, 是否也不依赖基的选取. 答案: 是.

定理 1.80(惯性定律) 设 V 是 n 维实向量空间, q 是 V 上的二次型, 那么在其标准形 (1.7.21) 中, 整数 s 与 r 都仅依赖 q, 不依赖典范基的选取.

证明 只要证明 s 不依赖典范基的选取. 假设 q 在基 $e_1, \cdots, e_n$ 下的表达式是标准形 (1.7.21), 另有基 $e_1', \cdots, e_n'$ 使得对 $x = \sum x_i'e_i' \in V$ 有

$$q(x) = {x'}_1^2 + \cdots + {x'}_t^2 - {x'}_{t+1}^2 - \cdots - {x'}_r^2.$$

如果 $t \neq s$, 则不妨设 $s < t$. 考虑 V 的两个子空间:

$$L = \langle e_{s+1}, \cdots, e_n\rangle,\quad L' = \langle e_1', \cdots, e_t'\rangle.$$

因为 $\dim(L+L')\leqslant\dim V\leqslant n$, 由推论 1.39 得

$$\begin{aligned}\dim(L\cap L')&=\dim L+\dim L'-\dim(L+L')\\&\geqslant n-s+t-n=t-s>0.\end{aligned}$$

所以 $L\cap L'$ 含有非零向量 v:

$$v=a_{s+1}e_{s+1}+\cdots+a_ne_n=a_1'e_1'+\cdots+a_t'e_t'.$$

利用 q 的两个标准形分别计算 $q(v)$, 得

$$q(v)=-a_{s+1}^2-\cdots-a_r^2\leqslant 0,\quad q(v)={a_1'}^2+\cdots+{a_t'}^2>0,$$

(有可能发生 $r<n$, $a_{s+1}=\cdots=a_r=0$). 这个矛盾意味着 $s=t$. □

实向量空间上的二次型称为**实二次型**. 惯性定律中的整数 r,s 是实二次型 q 的重要不变量, 分别称为 q 的**惯性指数**和**正惯性指数**. 它们的差 $r-s$ 称为 q 的**负惯性指数**. **符号差**(signature) 是指整数对 $(s,r-s)$, 也指正惯性指数与负惯性指数的差 $2s-r$. 惯性定律有时也称西尔维斯特 (Sylvester) 惯性定律, 他在 1852 年发表了定律的证明.

例 1.81 考虑二次型 $q(x)=x_1x_2+x_2x_3$. 做 (可逆) 变量替换:

$$x_1=x_1'+x_2'-x_3',\quad x_2=x_1'-x_2',\quad x_3=x_3',$$

得 $q(x)={x_1'}^2-{x_2'}^2$. 于是二次型的符号是 $(1,1)$.

实数域和复数域上的向量空间的二次型是线性代数讨论的对象, 有理数域和有限域上的向量空间的二次型是数论研究的对象.

九 总是取非负值或非正值的实二次型特别有意思.

定义 1.82 实二次型 $q:V\to\mathbb{R}$ 称为**正定的**如果对 V 中任意非零向量 x 有 $q(x)>0$; **半正定的**如果对任意的 $x\in V$ 有 $q(x)\geqslant 0$; **负定的**如果 $-q$ 是正定的, 即 $q(x)<0$ 如果 $x\neq 0$; **半负定的**如果对任意的 $x\in V$ 有 $q(x)\leqslant 0$. 最后, 二次型 q 称为**不定的**如果它在有些向量处取正值, 在另外有些向量处取负值.

需要注意的是这些概念和基的选取没有关系. 如果 $n=\dim V$, 那么相应的标准形是

正定情形: $x_1^2+x_2^2+\cdots+x_n^2$;

半正定情形: $x_1^2+x_2^2+\cdots+x_r^2, r\leqslant n$;

负定情形: $-x_1^2-x_2^2-\cdots-x_n^2$;

半负定情形: $-x_1^2-x_2^2-\cdots-x_r^2, r\leqslant n$;

不定情形: $x_1^2+x_2^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2, 1\leqslant s<r\leqslant n$.

如果实二次型 q 在典范基 (e_i) 下的表达式是 $\lambda_1 x_1^2 + \cdots + \lambda_n x^n$, 那显然 q 半正定当且仅当所有的系数都是非负的, 正定当且仅当所有的系数为正. 注意 λ_i 就是 q 在 e_i 处的值.

对称双线性型和对称矩阵也有正定 (半正定) 的概念. 称对称双线性型 f 是**正定 (半正定) 的**如果它给出的二次型是正定 (半正定) 的, 这时 f 在任意基下的矩阵 F 也称为**正定 (半正定) 的**.

由于正定的双线性型在某个基下的矩阵是单位矩阵, 由定理 1.69 得

定理 1.83 任意的正定矩阵具有如下形式:

$$F = {}^tA \cdot A,$$

其中 A 是非退化实矩阵. 反过来也成立: 如果 A 是非退化实矩阵, 那么 ${}^tA \cdot A$ 是正定矩阵.

如果二次型在某个基下的表达式是

$$q(x) = \sum_{i,j} f_{ij} x_i x_j,$$

其中 $f_{ij} = f_{ji}$, 那么通过这些系数直接判断 q 的正定性或其他性质的想法是诱人的. 如果矩阵 $F = (f_{ij})$ 的主子式都是非零的, 雅可比方法 (定理 1.79) 给出了二次型的负惯性指数.

定理 1.84 假设实二次型在某个基下的矩阵的主子式 Δ_k, $1 \leqslant k \leqslant n$ 都是非零的, 那么二次型的负惯性指数等于下面数列的变号数:

$$1,\ \Delta_1,\ \Delta_2,\ \cdots,\ \Delta_n.$$

证明 应用定理 1.79. □

注意 $\Delta_n \neq 0$ 意味着二次型是非退化的. 这是正负惯性指数的和就是 n. 上述定理表明如果主子式都是正的, 则二次型是正定的, 反之也对.

定理 1.85(西尔维斯特准则) 实二次型是正定的当且仅当它 (在任意基下) 的矩阵的主子式都是正的.

证明 我们仅需证明如果实二次型是正定的, 则它的矩阵的主子式都是正的. 设 q 是实向量空间 V 上的正定二次型, $e_1, \cdots, e_n$ 是 V 的基, q 在这个基下的矩阵为 F. 对向量空间 V 的维数 n 作归纳法证明 F 的主子式都是正的. 当 $n = 1$ 时, 结论显然. 命 $U = \langle e_1, \cdots, e_{n-1} \rangle$, 那么 q 在 U 上的限制 $q' = q|_U$ 仍是正定的. 显然 q' 在基 $e_1, \cdots, e_{n-1}$ 下的矩阵 F' 就是 F 去掉最后一行和最后一列得到的矩阵, 所以其主子式是 F 的主子式 $\Delta_1, \cdots, \Delta_{n-1}$. 由归纳假设, 这 $n-1$ 个主子式都是正的. 只剩下 Δ_n 的正性要证明. 根据定理 1.82, 有非退化矩阵 A 使得 $F = {}^tA \cdot A$. 于是

$$\Delta_n = \det F = \det {}^tA \cdot \det A = (\det A)^2 > 0.$$ □

例 1.86 设 $\varphi(x)=\varphi(x_1,\cdots,x_n)$ 是 n 元实变量函数, 在点 $c=(c_1,\cdots,c_n)$ 附近有收敛的泰勒展开. 如果 c 是**临界点**, 即对所有的 i 有 $\dfrac{\partial\varphi}{\partial x_i}(c)=0$, 那么 φ 在 a 处的泰勒展开是

$$\varphi(x)=\varphi(c)+\frac{1}{2}\sum_{i,j=1}^{n}a_{ij}(x_i-c_i)(x_j-c_j)+\text{高次项},$$

其中$a_{ij}=\dfrac{\partial^2\varphi}{\partial x_i\partial x_j}(c)$. 在点 c 足够小的邻域, 高次项的贡献是微小的, 可以忽略不计. 从而 φ 在 c 附近的形态与 $\varphi(a)+q(y)$ 差不多, 此处

$$q(y)=\sum_{i,j=1}^{n}a_{ij}y_iy_j,\quad y_i=x_i-c_i$$

是二次型. 如果 $\operatorname{rank}q=n$ 则称 c 是非退化的临界点. 如果 q 正定, 那么 $\varphi(c)$ 是**极小值**. $\varphi(c)$ 是**极大值**意味着 q 是负定的. 如果 q 是不定的非退化二次型, 则称临界点 c 为**鞍点**.

十 斜对称双线性型的典范式 讨论了二次型和对称双线性型后, 定理 1.73 建议我们转向斜对称双线性型. 换句话说, 要考虑 n 维向量空间 V 上的双线性型 f, 满足如下条件:

$$f(x,y)=-f(y,x),\quad \forall x,y\in V.$$

目标是寻找 V 的基使得 f 的矩阵尽可能简单. 令人惊讶的是斜对称双线性型的典范式和 V 的基域 K 没有关系 (注意我们已经假设了 $\operatorname{char}K\neq 2$).

如果 $f=0$, 无事可做. 现设 $f\neq 0$, 那么存在 $e_1,e_2\in V$ 使得 $f(e_1,e_2)=f_{12}\neq 0$. 可以要求 $f_{12}=1$, 否则用 $f_{12}^{-1}e_2$ 代替 e_2. 于是有

$$f(e_1,e_2)=1,\quad f(e_2,e_1)=-1,\quad f(e_1,e_1)=f(e_2,e_2)=0. \tag{1.7.22}$$

由此可见 e_1,e_2 是线性无关的. 考虑 V 的线性函数 $\alpha_i:x\to f(x,e_i)$, $i=1,2$, 给出的线性方程组

$$f(x,e_1)=0,\quad f(x,e_2)=0.$$

易见 α_1 和 α_2 线性无关. 根据定理 1.62(1), 该方程组的解空间 $\langle e_1,e_2\rangle^{\perp}$ 的维数是 $n-2$. 由于 $f(ae_1+a_2e_2,e_1)=a_1$, $f(ae_1+a_2e_2,e_2)=a_2$, 所以 $\langle e_1,e_2\rangle$ 中的非零元不是方程组的解. 于是

$$V=\langle e_1,e_2\rangle\oplus\langle e_1,e_2\rangle^{\perp}.$$

双线性型 f 在 $\langle e_1,e_2\rangle^{\perp}$ 上的限制仍是斜对称的, 可以重复上面的过程. 对 n 做归纳法秩, 我们得到如下结论.

定理 1.87　设 V 是域 K 上的 n 维向量空间, f 是 V 上的斜对称双线性型. 那么存在 V 的基 $e_1, e_2, \cdots, e_{2m-1}, e_{2m}, e_{2m+1}, \cdots, e_n$ 使得

$$\begin{aligned} &f(e_{2k-1}, e_{2k}) = -f(e_{2k}, e_{2k-1}) = -1, \quad 对 \ \ k = 1, \cdots, m, \\ &f(e_i, e_j) = 0, \quad 所有其他情况. \end{aligned}$$

从而 f 在这个基下的矩阵是

$$\begin{pmatrix} 0 & 1 & & & & & & & & \\ -1 & 0 & & & & & & & & \\ & & 0 & 1 & & & & 0 & & \\ & & -1 & 0 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 0 & 1 & & & \\ & & & & & -1 & 0 & & & \\ & & 0 & & & & & 0 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 0 \end{pmatrix}.$$

特别地, 如果 f 是非退化的, 则 V 的维数是偶数.

推论 1.88　任意非退化的斜对称的 $2m$ 阶方阵 F 合同于如下矩阵:

$$J = \begin{pmatrix} 0 & 1 & & & & & & \\ -1 & 0 & & & & & 0 & \\ & & \ddots & & & & & \\ & & & \ddots & & & & \\ & & & & \ddots & & & \\ & 0 & & & & & 0 & 1 \\ & & & & & & -1 & 0 \end{pmatrix},$$

也就是说, 存在非退化的矩阵 A 使得 ${}^tAFA = J$.

定理 1.87 中的基称为斜对称双线性型 f 的**辛基**(symplectic basis). 空间 V 的二维子空间 U 称为 f 的**辛平面**如果它有基 u, v 使得

$$f(u, v) = -f(v, u) = 1, \quad f(u, u) = f(v, v) = 0.$$

空间 V 的两个线性无关的子空间 U 和 W 称为**斜正交的**如果 $f(u, w) = 0$ 对所有的 $u \in U$, $w \in W$ 成立. 双线性型 f 的核定义为

$$\ker f = \{v \in V \mid f(v, x) = 0, \ \forall x \in V\}.$$

定理 1.87 表明斜对称双线性型 f 是其核及若干斜正交的辛平面的直和.

常选择

$$J_0=\begin{pmatrix}0 & I_m\\ -I_m & 0\end{pmatrix}$$

作为标准的非退化斜对称矩阵, 其中 I_m 是 m 阶单位矩阵. 只要适当变动基向量的顺序, 非退化斜对称双线性型的矩阵就从 J 变成了 J_0.

十一 在辛基下, 斜对称双线性型 f 的表达式是**典范的**, 十分简单

$$f(x,y)={}^tXJY=(x_1y_2-x_2y_1)+\cdots+(x_{2m-1}y_{2m}-x_{2m}y_{2m-1}).$$

在一般的基下, f 的表达式是

$$f(x,y)={}^tXFY=\sum_{1\leqslant i<j\leqslant n}f_{ij}(x_iy_j-x_jy_i). \tag{1.7.23}$$

可以通过标准的方法把上式化成典范式. 零双线性型无需考虑. 现设 $f\neq 0$, 适当变动基向量的顺序, 可设 $f_{12}\neq 0$. 在 (1.7.23) 中含 x_1 和 y_1 的项是

$$x_1(f_{12}y_2+\cdots+f_{1n}y_n)-(f_{12}x_2+\cdots+f_{1n}x_n)y_1.$$

命

$$x_2'=f_{12}x_2+\cdots+f_{1n}x_n,\quad y_2'=f_{12}y_2+\cdots+f_{1n}y_n,$$

其他变量 $x_1,y_1,x_3,y_3,\cdots,x_n,y_n$ 记号保持. 考虑 $f(x,y)$ 中含有 x_2',y_2' 的项, 有

$$\begin{aligned}f(x,y)&=(x_1-f_{23}'x_3-\cdots-f_{2n}'x_n)y_2'-x_2'(y_1-f_{23}'y_3-\cdots-f_{2n}'y_n)+\cdots\\&=x_1'y_2'-x_2'y_1'+\cdots,\end{aligned}$$

其中

$$x_1'=x_1-f_{23}'x_3-\cdots-f_{2n}'x_n,\quad y_1'=y_1-f_{23}'y_3-\cdots-f_{2n}'y_n,$$

省略的加项只含有变量 $x_3,y_3,\cdots,x_n,y_n$, 仍是一个斜对称双线性型 f'. 如果 $f'\neq 0$, 可用同样的方法处理. 最后, f 就化成了典范式

$$f(x,y)=(x_1'y_2'-x_2'y_1')+\cdots+(x_{2m-1}'y_{2m}'-x_{2m}'y_{2m-1}').$$

十二 普法夫多项式 (Pfaffian) 根据推论 1.88, 对域 K 上的每个非退化的斜对称方阵 F 存在矩阵 A 使得

$${}^tAFA=\begin{pmatrix}0 & I_m\\ -I_m & 0\end{pmatrix}.$$

从而 $(\det A)^2 \det F = 1$. 由此可知 $\det F$ 是 K 中的平方元. 虽然对复数域或实数域, 成为域中的平方元不是稀奇的事情, 但对很多的域, 平方元并不寻常, 比如, 有理数域, 分式域等. 斜对称矩阵可以看作一个由未知元构成的斜对称矩阵 $T = (t_{ij})$ 的取值, 其中, $t_{ji} = -t_{ij}$, 且 t_{ij}, $1 \leqslant i < j \leqslant n = 2m$, 是未知元. 而 T 是整环

$$K[t] = K[t_{12}, t_{23}, \cdots, t_{n-1,n}]$$

的分式域

$$K(t) = K(t_{12}, t_{23}, \cdots, t_{n-1,n})$$

上的斜对称矩阵, 具有非退化性. 于是存在 $P_n(t) \in K(t)$ 使得

$$\det T = P_n(t)^2.$$

取 $K = \mathbb{Q}$. 由于 $\mathbb{Q}(t)$ 是 $\mathbb{Z}[t]$ 的分式域, $\det T$ 的系数都是整数, 且 $\mathbb{Z}[t]$ 是唯一因子分解整环, 所以 $P_n(t) = P_n(t_{12}, t_{23}, \cdots, t_{n-1,n}) \in \mathbb{Z}[t]$ 是整系数多项式. 由于 $\mathbb{Z}[t]$ 是整环, $\det T$ 在其中只有两个平方根. 我们可以选定一个根使其在 $T = J_0$ 处的取值为 $(-1)^{m(m-1)/2}$ $(n = 2m)$, 记这个根为 $\mathrm{Pf}(T)$ 或 $\mathrm{Pf}_n(t)$, 称为 T 的**普法夫**(Pfaffian) 或 **n 维普法夫多项式**. 例如

$$\mathrm{Pf}_2(t) = t_{12}, \quad \mathrm{Pf}_4(t) = t_{12}t_{34} - t_{13}t_{24} + t_{14}t_{23}.$$

它们的平方分别是下列矩阵的行列式:

$$\begin{pmatrix} 0 & t_{12} \\ -t_{12} & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & t_{12} & t_{13} & t_{14} \\ -t_{12} & 0 & t_{23} & t_{24} \\ -t_{13} & -t_{23} & 0 & t_{34} \\ -t_{14} & -t_{24} & -t_{34} & 0 \end{pmatrix}.$$

奇数维的普法夫多项式 $\mathrm{Pf}_{2m+1}(t)$ 定义为 0. 对于域 K 上的 n 阶斜对称矩阵 $F = (f_{ij})$, 有唯一的环同态

$$\varphi_F : \mathbb{Z}[t] = \mathbb{Z}[t_{12}, t_{23}, \cdots, t_{n-1,n}] \to K$$

把 t_{ij} 映到 f_{ij}. 在这个映射下, $\mathrm{Pf}_n(t)$ 的像记作 $\mathrm{Pf}(F)$, 称为 F 的**普法夫**.

定理 1.89 设 F 是斜对称方阵, 那么

$$\det F = \mathrm{Pf}(F)^2.$$

而且

$$\mathrm{Pf}({}^{t}AFA) = \det A \cdot \mathrm{Pf}(F).$$

证明　映射 φ_F 把 $\det T$ 映到 $\det F$. 由此可知第一个等式成立. 为证第二个等式, 考虑与 F 同阶的方阵 $U=(u_{ij})$, 其中 u_{ij} 都是未知元, 相互独立. 那么

$$\mathrm{Pf}({}^tUTU)^2=\det({}^tUTU)=(\det U)^2\det T=(\det U)^2\mathrm{Pf}(T)^2.$$

由于整环上的多项式环是整环, 所以

$$(*)\ \ \mathrm{Pf}({}^tUTU)=\det U\cdot\mathrm{Pf}(T)\quad 或\quad (\star)\ \ \mathrm{Pf}({}^tUTU)=-\det U\cdot\mathrm{Pf}(T).$$

取 U 为单位矩阵即知 $(*)$ 成立. 于是对特殊的矩阵 $U=A$ 和 $T=F$, 等式 $(*)$ 也成立. □

注意 $\det T$ 是 $2m$ 次齐次多项式, 故 T 的普法夫是 m 次齐次多项式. 普法夫多项式在微分几何中有用, 出现在高维的 Gauss-Bonnet 公式中.

习　题　1.7

1. 说明下面的函数是否是某个空间的双线性型, 其中 K 是域:

(1) $f(x,y)={}^tx\cdot y\,(x,y\in K^n$ 是列向量$)$;

(2) $f(A,B)=\mathrm{tr}\,({}^tAB)\,(A,B\in M_{m,n}(K))$;

(3) $f(A,B)=\det(AB)\,(A,B\in M_n(K))$;

(4) $f(A,B)$ 是矩阵 AB 在 (i,j) 处的值;

(5) $f(u,v)=\displaystyle\int_a^b uvt^2dt\quad(u,v$ 是闭区间 $[a,b]$ 上的连续函数$)$;

(6) $f(u,v)=\displaystyle\int_a^b (u+v)^2dt$;

(7) $f(u,v)=(uv)(a)\quad(u,v\in K[x],\ a\in K)$;

(8) $f(u,v)=\dfrac{d}{dt}(uv)(a)$.

2. 对列向量空间 $\mathbb{R}^2$ 上的双线性型 $f(x,y)={}^tx\begin{pmatrix}2&1\\1&3\end{pmatrix}y$ 求典范基.

3. 运用雅可比方法把双线性型化成典范式:

$$f(x,y)=x_1y_1-x_1y_2+x_1y_3-x_2y_1+x_2y_2+2x_2y_3+x_3y_1+2x_3y_2+x_3y_3.$$

4. 运用雅可比方法判断下面两个矩阵是否合同:

$$\begin{pmatrix}1&2&3\\2&0&-1\\3&-1&3\end{pmatrix},\quad\begin{pmatrix}1&3&0\\3&1&1\\0&1&5\end{pmatrix}.$$

5. 举例说明:

(1) 正定矩阵 (a_{ij}) 可以在某些 (i,j) 处的值 a_{ij} 是负的;

(2) 实对称矩阵 $A=(a_{ij})$ 所有的值都是正的, 但 A 可以不是正定的.

6. 证明: 实二次型

$$q(x)=\sum_{i,j=1}^{n}a_{ij}x_ix_j,\quad a_{ij}=a_{ji},$$

是半正定的当且仅当对任意的指标 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$, k 阶行列式 $\det(a_{i_ri_s})$ 是非负的.

7. 已知实二次型 $f(x)=ax_1^2+x_2^2+3x_3^2-2x_1x_2$ 的秩为 2,

(1) 写出二次型所对应的矩阵 A, 并求参数 a;

(2) 把二次型化成标准形.

8. 实二次型

$$\lambda x_1^2-2x_2^2-3x_3^2+2x_1x_2-2x_1x_3+2x_2x_3$$

在 λ 取什么值时是负定的.

9. 设 $A=(a_{ij})\in M_n(\mathbb{R})$ 是对称矩阵,

$$L(f)=\sum_{i,j=1}^{n}a_{ij}\frac{\partial^2 f}{\partial x_i\partial x_j}$$

是多项式环 $\mathbb{R}[x_1,\cdots,x_n]$ 上的微分算子. 证明

(1) 如果 $C\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$, $C\in GL_n(\mathbb{R})$, 是变量替换, 那么

$$L(f)=\sum_{i,j=1}^{n}b_{ij}\frac{\partial^2 f}{\partial y_i\partial y_j},$$

其中 $(b_{ij})=CA{}^tC$;

(2) 可以找到某个 $C\in GL_n(\mathbb{R})$ 使得

$$L(f)=\frac{\partial^2 f}{\partial y_1^2}+\cdots+\frac{\partial^2 f}{\partial y_s^2}-\frac{\partial^2 f}{\partial y_{s+1}^2}-\cdots-\frac{\partial^2 f}{\partial y_t^2},$$

其中 $0\leqslant s\leqslant t\leqslant n$.

10. 设 $x=[x_1,x_2,x_3]\in\mathbb{C}^3$, $q(x)=x_1^3+x_2^3+x_3^3-3x_1x_2x_3$, ε 是 1 的三次本原根. 利用等式

$$q(x)=(x_1+x_2+x_3)(x_1+\varepsilon x_2+\varepsilon^2x_3)(x_1+\varepsilon^2x_2+\varepsilon x_3)$$

证明: (1) 存在对称双线性型 $z_i=z_i(x,y)=\sum_{j,k}a_{jk}^{(i)}x_jy_k$ 使得 $q(x)q(y)=q(z)$, 其中 $z=[z_1,z_2,z_3]$; (2) 求出这些双线性型.

11. 设 A 是域 K 上的 n 阶方阵, $\operatorname{char}K\neq 2$. 证明 A 是斜对称的当且仅当对所有的 $X=[x_1,\cdots,x_n]$ 有 ${}^tXAX=0$.

12. 把双线性型

$$x_1y_2 - y_2x_1 + x_1y_3 - y_3x_1 + x_2y_3 - y_3x_2 + x_2y_4 - y_4x_2$$

化成典范式.

13. 设 A 是 n 阶斜对称实方阵, E 是 n 阶单位矩阵, λ 是复数. 如果 $\det(\lambda E - A) = 0$, 那么 λ 的实部为 0.

14. 设 A 是实对称矩阵, E 是同阶单位矩阵, ε 是充分小的实数. 证明方阵 $E + \varepsilon A$ 是正定的.

第2章　线 性 算 子

我们生活的空间中, 取定原点后, 绕原点的旋转, 以过原点的平面为镜面的反射, 伸缩变换, 微积分中的求导数等都是线性映射的例子. 第 1 章关于线性函数的内容和第一卷 3.3 节已经显示向量空间之间的线性映射是研究向量空间不可缺少的工具, 而且极大地丰富和加深了线性代数的内容. 一类特殊的线性映射, 称为线性算子, 是本章重点讨论的内容, 在线性代数中占有核心的位置, 应用广泛.

2.1　向量空间的线性映射

一　本节的向量空间都是域 K 上的向量空间. 向量空间之间的线性映射就是保持加法和纯量乘的映射.

定义 2.1　设 V 和 W 是域 K 上的向量空间. 称映射 $f: V \to W$ 为**线性映射**, 如果

(1) $f(u+v) = f(u) + f(v)$,

(2) $f(\lambda u) = \lambda f(u)$,

其中 u, v 是 V 中任意的元素, λ 是 K 中的任意元素.

定义中的条件 (1) 和 (2) 可以合并成一个条件: $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$. 线性函数是一类特殊的线性映射.

例 2.2　在实平面 $\mathbb{R}^2$ 上绕原点的旋转是从 $\mathbb{R}^2$ 到自身的线性映射 (其实还是同构).

例 2.3　三维空间 $\mathbb{R}^3$ 到二维子空间 $V = \{(a,b,0)\,|\,a,b \in \mathbb{R}\}$ 的投影是线性映射.

例 2.4　命 $C^k(\mathbb{R})$ 为实轴上的 k 次可微函数全体, $\mathbb{R}_{k+1}[x]$ 是次数不超过 k 的实多项式全体. 那么对任意取定的实数 a, 映射

$$\varphi \to \varphi(a) + \varphi'(a)(x-a) + \frac{\varphi''(a)}{2}(x-a)^2 + \cdots + \frac{\varphi^{(k)}(a)}{k!}(x-a)^k,$$

是从 $C^k(\mathbb{R})$ 到 $\mathbb{R}_{k+1}[x]$ 的线性映射.

例 2.5　记号同例 2.4. 求导数是 $C^k(\mathbb{R})$ 到自身的线性映射, 也是 $\mathbb{R}_{k+1}[x]$ 到自身的线性映射.

例 2.6 命 $C(\mathbb{R})$ 为实轴上的连续函数全体. 那么映射

$$\mathscr{A}: C(\mathbb{R}) \to C(\mathbb{R}), \quad (\mathscr{A}\varphi)(x) = \int_0^1 \varphi(y)(x-y)^3 dy,$$

是线性映射.

每一个线性映射 $f: V \to W$ 自然产生两个子空间: f 的**核**与 f 的**像**, 定义如下:

$$\ker f = \{v \in V \mid f(v) = 0\}, \quad \operatorname{Im} f = \{f(v) \mid v \in V\}.$$

显然核是 V 的子空间, 像是 W 的子空间.

除非特别说明, 以下讨论的向量空间 V, W, U 等都是有限维的. 关于线性映射的一个基本结果是

定理 2.7 设 $f: V \to W$ 是线性映射, 那么

$$\dim \ker f + \dim \operatorname{Im} f = \dim V.$$

证明 假设 V 的维数为 n. 取 $\ker f$ 的一个基 $v_1, \cdots, v_k$, 把它扩充为 V 的基 $v_1, \cdots, v_n$. 我们断言 $f(v_{k+1}), \cdots, f(v_n)$ 是 $\operatorname{Im} f$ 的基, 即它们张成 $\operatorname{Im} f$ 且线性无关.

对 V 中的任意向量 $v = a_1 v_1 + \cdots + a_n v_n$, 有 $f(v) = a_{k+1} f(v_{k+1}) + \cdots + a_n f(v_n)$. 所以 $f(v_{k+1}), \cdots, f(v_n)$ 张成 $\operatorname{Im} f$. 如果

$$\lambda_{k+1} f(v_{k+1}) + \cdots + \lambda_n f(v_n) = 0, \quad \lambda_i \in K,$$

那么 $f(\lambda_{k+1} v_{k+1} + \cdots + \lambda_n v_n) = 0$. 即 $\sum\limits_{i=k+1}^{n} \lambda_i v_i \in \ker f$. 于是 $\sum\limits_{i=k+1}^{n} \lambda_i v_i = \sum\limits_{j=1}^{k} \lambda_j v_j$. 但 $v_1, \cdots, v_n$ 线性无关, 所以等式中的系数 λ_i 全为 0, 从而 $f(v_{k+1}), \cdots, f(v_n)$ 线性无关. 于是, 要证的等式成立. □

注 上述定理对任意维数的向量空间之间的线性映射都是成立的.

推论 2.8 设 $f: V \to W$ 是线性映射, 那么 $\dim \operatorname{Im} f \leqslant \dim V$, 等号成立当且仅当 f 是单射.

二 线性映射的矩阵和线性映射的秩 我们曾经对两类特殊的线性映射有过较详细的讨论, 它们分别是线性函数 (见 1.6 节) 和坐标空间之间的线性映射 (见第一卷 3.3 节). 线性函数完全由它在一个基上的值确定, 而且在一个基上的值可以取任意给定的值. 这一点对一般的线性映射也是成立的. 也就是说, 如果 $v_1, \cdots, v_n$ 是向量空间 V 的基, 那么对向量空间 W 中任意给定的向量组 $\xi_1, \cdots, \xi_n$, 存在唯一的线性映射 $f: V \to W$ 使得 $f(v_i) = \xi_i$, $i = 1, \cdots, n$.

事实上，如果线性映射 f 满足条件，那么对任意的向量 $v = \sum_{i=1}^{n} a_i v_i$，有

$$f(v) = \sum_{i=1}^{n} a_i f(v_i) = \sum_{i=1}^{n} a_i \xi_i. \tag{2.1.1}$$

所以 f 是唯一的．反之，给定向量组 $\xi_1, \cdots, \xi_n$，通过上式定义的映射 $f: V \to W$ 是线性映射，因为对任意的 $v = \sum_{i=1}^{n} a_i v_i$ 和 $v' = \sum_{i=1}^{n} a_i' v_i$ 有

$$\begin{aligned} f(cv + c'v') &= f\left(\sum_{i=1}^{n}(ca_i + c'a_i')v_i\right) = \sum_{i=1}^{n}(ca_i + c'a_i')\xi_i \\ &= \sum_{i=1}^{n} ca_i\xi_i + \sum_{j} c'a_i'\xi_i = cf(v) + c'f(v'). \end{aligned}$$

从而 f 是满足要求的映射．

坐标空间之间的线性映射的一个特点是与矩阵紧密相连．一般的线性映射也是如此．取定 V 的基 $v_1, \cdots, v_n$ 和 W 的基 $w_1, \cdots, w_m$．设 $f: V \to W$ 是线性映射．我们有

$$\begin{gathered} f(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m, \\ \cdots\cdots \\ f(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m. \end{gathered} \tag{2.1.2}$$

系数矩阵

$$M_f = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{2.1.3}$$

称为**线性映射 $f: V \to W$ 关于基 (v_i) 和 (w_j) 的矩阵**，也说 M_f 是 f**在基 (v_i) 和 (w_j) 下的矩阵**．

例 2.9　对实平面 $\mathbb{R}^2$ 取标准的基 $e_1 = (1, 0)$，$e_2 = (0, 1)$．命 $f: \mathbb{R}^2 \to \mathbb{R}^2$ 是绕原点旋转角度 α 的线性映射．那么

$$\begin{aligned} f(e_1) &= \quad e_1 \cos\alpha + e_2 \sin\alpha, \\ f(e_2) &= -e_1 \sin\alpha + e_2 \cos\alpha. \end{aligned}$$

所以，f 关于标准基的矩阵是

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{2.1.4}$$

例 2.10 三维空间 $\mathbb{R}^3$ 到二维子空间 $V=\{(a,b,0)\,|\,a,b\in\mathbb{R}\}$ 的投影 P 关于 $\mathbb{R}^3$ 的标准基 e_1,e_2,e_3 和 V 的基 e_1,e_2 的矩阵是

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

因为 $P(e_1)=e_1$，$P(e_2)=e_2$，$P(e_3)=0$.

注意矩阵 M_f 的第 j 列是向量 $f(v_j)=\xi_j$ 的坐标形成的列向量. 由定理 1.25 知向量组 $\xi_1,\cdots,\xi_n$ 的秩是矩阵 M_f 的秩. 由于 $\xi_1,\cdots,\xi_n$ 张成 $\operatorname{Im} f$, 所以 $\dim\operatorname{Im} f=\operatorname{rank} M_f$. 因此, 像 $\operatorname{Im} f$ 的维数也称为线性映射 f 的**秩**, 记作 $\operatorname{rank} f$.

以上所讨论的结果可以总结如下.

定理 2.11 设 $v_1,\cdots,v_n$ 和 $w_1,\cdots,w_m$ 分别是向量空间 V 和 W 的基. 对线性映射 $f:V\to W$ 用 M_f 记其关于 (v_i) 和 (w_j) 的矩阵, 那么

(1) 从 V 到 W 的线性映射全体和基域 K 上的 $m\times n$ 矩阵全体之间有一个双射

$$\varphi: f\to M_f.$$

(2) $\dim\operatorname{Im} f=\operatorname{rank} M_f$, 即线性映射 $f:V\to W$ 的秩等于 f 在 V 和 W 的任意的基下的矩阵的秩.

(3) 对任意的向量组 $\xi_1,\cdots,\xi_n\in W$, 存在唯一的线性映射 $f:V\to W$ 使得 $f(v_i)=\xi_i$，$i=1,\cdots,n$.

三　从 V 到 W 的线性映射全体记作 $\mathcal{L}(V,W)$, 也记作 $\operatorname{Hom}(V,W)$. 例 1.5 说明可以在这些线性映射中定义加法和纯量乘法:

$$(\lambda f+\mu g)(v)=\lambda f(v)+\mu g(v), \tag{2.1.5}$$

其中 $f,g\in\mathcal{L}(V,W)$，$\lambda,\mu\in K$，$v\in V$. 直接验证知 $\lambda f+\mu g$ 是线性映射. 于是 $\mathcal{L}(V,W)$ 是 W^V 的子空间. 我们已经知道域 K 上的 $m\times n$ 全体 $M_{m,n}(K)$ 在矩阵的加法和纯量与矩阵的乘法下成为向量空间.

不会让人惊讶的事实是: 定理 2.11(1) 中的双射是线性同构. 验证是直截了当的: 从 M_f 的定义和 (2.1.5) 容易看出, 对于任意的 $f,g\in\mathcal{L}(V,W)$ 和 $\lambda,\mu\in K$, 有

$$M_{\lambda f+\mu g}=\lambda M_f+\mu M_g.$$

由此可见 $\dim\mathcal{L}(V,W)=\dim V\cdot\dim W$. 线性映射与矩阵的紧密联系不仅体现在加法和纯量乘上, 还体现在乘法上.

利用矩阵 M_f 可以把 (2.1.2) 中的等式写成矩阵的形式:

$$f(v_1,\cdots,v_n)=(f(v_1),\cdots,f(v_n))=(w_1,\cdots,w_m)M_f. \tag{2.1.6}$$

设 $g: U \to V$ 是线性映射. 又设 (u_k) 是 U 的基, g 关于基 (u_k) 和 (v_i) 的矩阵是 M_g. 置 $\dim U = p$. 于是

$$\begin{aligned} g(u_1, \cdots, u_p) &= (v_1, \cdots, v_n)M_g \\ (fg)(u_1, \cdots, u_p) &= f((v_1, \cdots, v_n)M_g) \\ &= f(v_1, \cdots, v_n) \cdot M_g \\ &= (w_1, \cdots, w_m)M_f \cdot M_g \\ &= (w_1, \cdots, w_m) \cdot M_f M_g. \end{aligned}$$

所以 fg 关于基 (u_k) 和 (w_j) 的矩阵是 $M_f M_g$, 即

$$M_{fg} = M_f M_g. \tag{2.1.7}$$

利用定理 2.11(2) 和公式 (2.1.7) 知, 下面的结论如果用矩阵的秩表述则是我们熟悉的结论: 矩阵乘积的秩小于因子的秩.

定理 2.12 设 $g: U \to V$ 和 $f: V \to W$ 是线性映射, 那么

(1) $\dim \mathrm{Im}\,(fg) \leqslant \dim \mathrm{Im}\, f$;

(2) $\dim \mathrm{Im}\,(fg) \leqslant \dim \mathrm{Im}\, g$.

证明 因为 $\mathrm{Im}\,(fg) \subset \mathrm{Im}\, f$, 所以第一个不等式成立. 由于 $\mathrm{Im}\,(fg) = f(\mathrm{Im}\, g)$, 推论 2.8 给出第二个不等式. □

四 线性映射与矩阵的紧密联系也体现在坐标变换上. 如果 $v \in V$ 在基 (v_i) 下的坐标列向量为 X, 则有 $v = (v_1, \cdots, v_n)X$, 从而

$$f(v) = f((v_1, \cdots, v_n)X) = (f(v_1), \cdots, f(v_n))X = (w_1, \cdots, w_m)M_f X.$$

所以 $f(v)$ 在基 (w_j) 下的坐标列向量为 $M_f X$. 于是线性映射 f 产生了线性映射

$$\theta_f : K^n \to K^m,\ X \to M_f X.$$

根据定理 1.25, V 的基 (v_i) 和 W 的基 (w_j) 分别确定了线性同构 $\varphi: V \to K^n$ 和 $\psi: W \to K^m$. 易见

$$\psi f = \theta_f \varphi.$$

这个等式常用如下交换图表示

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\varphi"'] & W \arrow[d, "\psi"] \\ K^n \arrow[r, "\theta_f"'] & K^m \end{tikzcd}$$

(映射图交换的含义: 起点和终点之间不同的路线给出同样的结果.)

习 题 2.1

1. 设 $f: V \to W$ 是线性映射. 证明:

(1)V 的任意子空间 U 的像 $f(U) = \{f(u) \mid u \in U\}$ 是 W 的子空间;

(2)W 的任意子空间 Z 的逆像 $f^{-1}(Z) = \{v \in V \mid f(v) \in Z\}$ 是 V 的子空间.

2. 设 $f: V \to W$ 是线性映射. 证明存在 V 的基和 W 的基, 使得 f 关于这两个基的矩阵具有如下形式

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

其中 I_r 是 r 阶单位矩阵, $r = \operatorname{rank} f$.

3. 设 $f: V \to W$ 是线性映射, (v_i) 和 (v_i') 是 V 的基, (w_j) 和 (w_j') 是 W 的基, f 关于 (v_i) 和 (w_j) 的矩阵是 M_f, 关于 (v_i') 和 (w_j') 的矩阵是 M_f'. 又设从 (v_i) 到 (v_i') 的转换矩阵是 P, 从 (w_j) 到 (w_j') 的转换矩阵是 Q. 证明:

$$M_f' = Q^{-1} M_f P.$$

4. 给出一个非线性函数 $f: \mathbb{R}^2 \to \mathbb{R}$ 使得对任意的 $a \in \mathbb{R}$ 和 $v \in \mathbb{R}^2$ 有

$$f(av) = af(v).$$

5. 设 U 是有限维向量空间 V 的子空间. 证明:

(1) 任何线性映射 $g: U \to W$ 都可以延拓成 V 到 W 的线性映射, 即存在线性映射 $f: V \to W$ 使得对任何 $u \in U$ 有 $f(u) = g(u)$.

(2) 自然映射 $V \to V/U$, $x \to x + U$ 是线性映射.

(3) 存在线性映射 $h: V \to U$ 使得 h 限制在 U 上是恒等映射.

6. 设 F 是 q 元域. 找出

(1) 从 F^n 到 F^k 的线性映射的个数;

(2) 从 F^n 到 F^k 的单线性映射的个数;

(3) 从 F^n 到 F^k 的满线性映射的个数.

7. 设 $f: V \to W$ 是线性映射. 证明: 如果 V 是无限维向量空间, 则 $\ker f$ 和 $\operatorname{Im} f$ 中至少有一个是无限维的.

8. 设 P_n 是域 K 上次数不超过 $n-1$ 的多项式全体形成的空间. 定义 $\varphi: P_n \to P_n$, $u(t) \to tu'(t) - u(t)$. 证明 φ 是线性映射, 并求出其核与秩.

9. 设 $A \in M_n(K)$ 是非退化矩阵. 证明: 映射 $f_C: M_n(K) \to M_n(K)$, $X \to C^{-1}XC$ 是线性映射且 $f_C(XY) = f_C(X)f_C(Y)$.

10. 设 V 是区间 $[-\pi, \pi]$ 上所有的连续实函数形成的向量空间. 命 U 为 V 中满足如下三个条件的所有函数 f 形成的集合:

$$\int_{-\pi}^{\pi} f(t)dt = 0, \quad \int_{-\pi}^{\pi} f(t)\cos t dt = 0, \quad \int_{-\pi}^{\pi} f(t)\sin t dt = 0.$$

证明:

(1) U 是 V 的子空间;

(2) U 包含所有的函数 $\cos nx$ 和 $\sin nx$, $n = 2, 3, \cdots$;

(3) U 是无限维空间.

定义线性映射 $T : V \to V$ 如下: 对 $f \in V$, $g = T(f)$, 有

$$g(x) = \int_{-\pi}^{\pi} [1 + \cos(x - t)] f(t) dt.$$

(4) 证明 $\operatorname{Im} T$ 是有限维向量空间并求出它的一个基;

(5) 求出 T 的核;

(6) 求所有的实数 $c \neq 0$ 和 V 中所有的非零元素 f 使得 $T(f) = cf$. (注意这样的 f 在 T 的像中.)

2.2 线性算子代数

一 从向量空间到自身的线性映射称为**线性算子**(**或线性变换**). 我们将用花体字母 $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \cdots$ 表示线性算子. 本节的向量空间都是有限维的. 设 $\mathcal{A} : V \to V$ 是线性算子, 向量 $v \in V$ 的像 $\mathcal{A}(v)$ 将简单记作 $\mathcal{A}v$. 设 $e_1, \cdots, e_n$ 是 V 的基, 那么

$$\mathcal{A}e_j = \sum_i a_{ij} e_i. \tag{2.2.8}$$

矩阵 $A = (a_{ij})$ 称为 $\mathcal{A}$ 在基 (e_i) 下的**矩阵**或关于基 (e_i) 的矩阵. *很多时候将简单称 A 是 $\mathcal{A}$ 的矩阵*. 注意 A 的第 j 列由 $\mathcal{A}e_j$ 在基 (e_i) 下的坐标构成. 把 (2.2.8) 中的等式全体用矩阵形式表达是方便的.

$$(\mathcal{A}e_1, \cdots, \mathcal{A}e_n) = (e_1, \cdots, e_n)A. \tag{2.2.9}$$

如果 $v \in V$ 在基 (e_i) 下的坐标列向量为 X, 由上一节第四部分的讨论知 $\mathcal{A}v$ 在基 (e_i) 下的坐标列向量为 AX.

用 $\mathcal{E}$ 记恒等映射 $\mathrm{Id}_V : V \to V$, $v \to v$. 称算子 $\mathcal{B}$ 是算子 $\mathcal{A}$ 的逆算子如果 $\mathcal{BA} = \mathcal{AB} = \mathcal{E}$. 显然 $\mathcal{A}$ 有逆算子当且仅当它是同构, 这时也称 $\mathcal{A}$ 是**可逆的或非退化的或非奇异的**. 称 $\mathcal{A}$ 是**退化的或奇异的**如果它不是可逆的. 上一节中关于对线性映射的定理 2.7 自然在这里是适用的, 而且更精细一些.

定理 2.13 设 $\mathcal{A} : V \to V$ 是线性算子, 那么

(1) $\dim \ker \mathcal{A} + \dim \operatorname{Im} \mathcal{A} = \dim V$;

(2) $\mathcal{A}$ 是同构当且仅当 $\mathcal{A}$ 是单射;

(3) $\mathcal{A}$ 是同构当且仅当 $\mathcal{A}$ 是满射.

证明 (1) 由定理 2.7 推出. 同构当然是单射和满射. 如果 $\mathcal{A}$ 是单射, 则 $\ker\mathcal{A}=0$, 于是 $\dim\operatorname{Im}\mathcal{A}=\dim V$, 从而 $\operatorname{Im}\mathcal{A}=V$, 即 $\mathcal{A}$ 也是满射, 所以 (2) 成立. 类似可证 (3) 成立. □

注意定理 2.13(1) 中的等式仅是关于维数的等式, 不意味着 V 是 $\ker\mathcal{A}$ 与 $\operatorname{Im}\mathcal{A}$ 的直和. 例如, 设 e_1,e_2 是一个二维空间 V 的基, 定义 $\mathcal{A}(\lambda e_1+\mu e_2)=\mu e_1$. 那么 $\mathcal{A}$ 是 V 上的线性算子, 有 $\ker\mathcal{A}=\operatorname{Im}\mathcal{A}=\langle e_1\rangle$.

二 设 $e_1',\cdots,e_n'$ 是 V 的另一个基, 从基 (e_i) 到 (e_i') 的转换矩阵是 B. 那么

$$(e_1',\cdots,e_n')=(e_1,\cdots,e_n)B.$$

由于 $\mathcal{A}$ 是线性的, 有

$$\begin{aligned}(\mathcal{A}e_1',\cdots,\mathcal{A}e_n')&=(\mathcal{A}e_1,\cdots,\mathcal{A}e_n)B\\&=(e_1,\cdots,e_n)AB=(e_1',\cdots,e_n')B^{-1}AB.\end{aligned}$$

所以 $\mathcal{A}$ 在基 (e_i') 下的矩阵是

$$A'=B^{-1}AB. \tag{2.2.10}$$

定义 2.14 称方阵 A' 和方阵 A 是**相似的**, 记作 $A'\sim A$, 如果存在可逆矩阵 B 使得 $A'=B^{-1}AB$. (这里的方阵可以相乘意味着它们都是同一个域上的同阶方阵.)

定理 2.15 向量空间 V 上的线性算子 $\mathcal{A}$ 在不同基下的矩阵是相似的.

这一结论使得我们能够定义线性算子的行列式和迹. 相似的矩阵有相同的行列式: $\det(B^{-1}AB)=\det A$, 所以线性算子 $\mathcal{A}$ 的所有矩阵的行列式值相等, 这个值称为 $\mathcal{A}$ 的**行列式**, 记作 $\det\mathcal{A}$.

方阵 $A=(a_{ij})$ 的**迹**定义为对角线的全体元素的和. 如果 A 是线性算子 $\mathcal{A}$ 在某个基下的矩阵, 则定义算子 $\mathcal{A}$ 的**迹**为 A 的迹:

$$\operatorname{tr}\mathcal{A}=\operatorname{tr}A=\sum_{i=1}^{n}a_{ii}.$$

对于同阶的方阵 A,B, 容易证明如下等式

$$\operatorname{tr}AB=\operatorname{tr}BA. \tag{2.2.11}$$

把这个公式应用到 $B^{-1}A$ 和 B, 其中 B 非退化, 得

$$\operatorname{tr}(B^{-1}AB)=\operatorname{tr}(BB^{-1}A)=\operatorname{tr}A.$$

即相似的矩阵有相同的迹, 所以线性算子 $\mathcal{A}$ 的所有的矩阵有相同的迹, 从而线性算子的迹不依赖定义中所选取的基和相应的矩阵.

三　方阵的相似关系是等价关系. 首先方阵 A 与自身相似 (在定义中取 B 为单位矩阵 I 即可). 其次 $A' = B^{-1}AB$ 意味着 $A = D^{-1}A'D$, 其中 $D = B^{-1}$, 所以相似关系是对称的: $A' \sim A \Rightarrow A \sim A'$. 再者, 如果 $A' = B^{-1}AB$, $A'' = C^{-1}A'C$, 则有 $A'' = (BC)^{-1}A(BC)$, 所以相似关系有传递性: $A'' \sim A'$, $A' \sim A \Rightarrow A'' \sim A$.

于是方阵全体在相似关系下被分成了不相交的等价类. 在相似等价类中寻找简单的矩阵是一个重要的问题, 对实际应用很有价值. 比如要计算方阵 A 的高次幂不是简单的事情, 甚至对于二阶方阵. 如果知道 A 与一个对角矩阵相似: $A = B^{-1}A'B$, 其中 $A' = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$, 那么 A 的幂的计算就变得简单了

$$A^k = B^{-1}{A'}^kB = B^{-1}\mathrm{diag}\,(\lambda_1^k, \cdots, \lambda_n^k)B.$$

第一卷例 3.31 就是利用这个想法计算矩阵 $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ 的幂, 从而求出斐波那契数列的通项. 如果能计算方阵 A 的幂, 那方阵 A 的多项式

$$a_mA^m + a_1A^{m-1} + \cdots + a_1A + a_0$$

的计算变得有可能.

定理 2.15 表明, 线性算子 $\mathcal{A}$ 在不同基下的矩阵形成一个相似 (等价) 类. 每个相似类都有一个线性算子与之对应. 于是寻找矩阵相似类中简单的矩阵就可以通过寻找合适的基使得相似类对应的线性算子的矩阵具有简单的形式. 我们后面将回到这个问题. 先看一些简单算子的例子.

例 2.16　把 V 中任何向量都映到 0 的算子称为**零算子**, 记作 $\mathcal{O}$. 它在任何基下的矩阵都是零矩阵.

例 2.17　实平面 $\mathbb{R}^2$ 的算子 $(a, b) \to (a, -b)$ 是关于直线 $\langle e_1 \rangle = \mathbb{R}e_1$ 的反射, 其中 $e_1 = (1, 0)$. 这个算子关于标准基的矩阵是 $\mathrm{diag}\,(1, -1)$. 更一般地, 设 $V = U \oplus W$, 那么线性算子 $\mathcal{R} : u + w \to u - w$ 是关于子空间 W 的一个反射. 取 V 的基使得该基与 U 和 W 都相合, 则在该基下算子 $\mathcal{R}$ 的矩阵是 $\mathrm{diag}\,(1, \cdots, 1, -1, \cdots, -1)$, 其中 1 出现的次数为 $\dim U$, -1 出现的次数为 $\dim W$.

例 2.18　取定纯量 $\lambda \in K$. 伸缩变换 $\mathcal{A} : V \to V$, $x \to \lambda x$, 关于任何基的矩阵都是纯量矩阵 λE.

例 2.19　设 $P_n[t]$ 是域 K 上的次数不超过 $n-1$ 的多项式全体形成的空间. 微分算子 $\mathcal{D} = \dfrac{d}{dt}$ 是 $P_n[t]$ 上的线性算子. 它在基 $1, t, \cdots, t^{n-1}$ 的矩阵是

$$\begin{pmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 2 & & 0 & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & 0 & 0 & n-1 \\ & & & 0 & 0 & 0 \end{pmatrix}.$$

如果 K 的特征为 0, 那么算子 $\mathcal{D}$ 在基 $1,\ t,\ \dfrac{t^2}{2},\ \cdots,\ \dfrac{t^{n-1}}{(n-1)!}$ 下的矩阵是

$$\begin{pmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & 0 & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & 0 & 0 & 1 \\ & & & 0 & 0 & 0 \end{pmatrix}.$$

四　现考虑不同线性算子之间的联系. 向量空间 V 上的线性算子全体 $\mathcal{L}(V,V)$ 可以简单记作 $\mathcal{L}(V)$. 由于 V 上的两个线性算子在映射合成下仍是 V 的线性算子, 所以 $\mathcal{L}(V)$ 不仅是向量空间, 还有乘法. 为了强调, 把加法、纯量乘、乘法的定义列出

$$(\mathcal{A}+\mathcal{B})(x)=\mathcal{A}x+\mathcal{B}x,\quad (\lambda\mathcal{A})x=\lambda(\mathcal{A}x),\quad (\mathcal{A}\mathcal{B})x=\mathcal{A}(\mathcal{B}x).$$

容易验证, 这些运算具有如下性质:

$$\begin{aligned}
&\lambda(\mathcal{A}+\mathcal{B})=\lambda\mathcal{A}+\lambda\mathcal{B},\quad (\lambda+\mu)\mathcal{A}=\lambda\mathcal{A}+\lambda\mathcal{B},\\
&(\lambda\mu)\mathcal{A}=\lambda(\mu\mathcal{A}),\quad 1\cdot\mathcal{A}=\mathcal{A},\\
&\mathcal{A}(\mathcal{B}\mathcal{C})=(\mathcal{A}\mathcal{B})\mathcal{C},\\
&\mathcal{A}(\mathcal{B}+\mathcal{C})=\mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C},\quad (\mathcal{A}+\mathcal{B})\mathcal{C}=\mathcal{A}\mathcal{C}+\mathcal{B}\mathcal{C},\\
&\lambda(\mathcal{A}\mathcal{B})=(\lambda\mathcal{A})\mathcal{B}=\mathcal{A}(\lambda\mathcal{B}),
\end{aligned}$$

其中, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ 是 V 上的线性算子, $\lambda,\mu\in K$. 最后一个性质是关于纯量乘与乘法的关系. 上面的性质比环的公理性质更强, 所以需要新的术语.

定义 2.20　环 R 称为域 K 上的**代数**如果

(1) K 中的元素可以对 R 中的元素作纯量乘法,

(2) 在这个纯量乘法和环的加法下 R 成为 K 上的向量空间,

(3) 对任意的 $\lambda\in K$ 和 $a,b\in R$ 有 $\lambda(ab)=(\lambda a)b=a(\lambda b)$.

称 R 的子集 L 为**子代数**如果 L 是 R 的子空间且对任意的 $a,b\in L$ 有 $ab\in L$.

域 K 上的两个代数 R 和 R' 称为**同构的**如果存在线性空间的同构 $\varphi: R\to R'$ 使得对任意的 $a,b\in R$ 有 $\varphi(ab)=\varphi(a)\varphi(b)$. 此时映射 φ 也称为**同构映射**.

于是 $\mathcal{L}(V)$ 是 K 上的代数. 代数的其他例子包括 K 上的 n 阶方阵全体 $M_n(K)$, 多项式代数 $K[t]$ 等. 域 K 自身也是 K 代数.

取定 V 的基 $e_1,\cdots,e_n$, 那么 V 上的算子 $\mathcal{A}$ 就有一个矩阵 A. 根据上一节第三部分的讨论, 映射

$$\varphi:\mathcal{L}(V)\to M_n(K),\quad \mathcal{A}\to A \tag{2.2.12}$$

是向量空间的同构. 根据公式 (2.1.7), 这个映射保持乘法, 即如果 B 是算子 $\mathcal{B}$ 在基 (e_i) 下的矩阵, 那么 $\mathcal{AB}$ 在这个基下的矩阵是 AB, 所以 φ 是代数同构.

算子的行列式和迹给出了两个函数, 分别称为**行列式函数**和**迹函数**:

$$\det:\mathcal{L}(V)\to K,\quad \mathcal{A}\to\det\mathcal{A};$$
$$\mathrm{tr}:\mathcal{L}(V)\to K,\quad \mathcal{A}\to\mathrm{tr}\,\mathcal{A}.$$

行列式函数保持乘法: $\det(\mathcal{AB})=(\det\mathcal{A})\cdot(\det\mathcal{B})$. 算子 $\mathcal{A}$ 的行列式不为零意味着算子在任何一个基下的矩阵都是非退化的, 从而 $\mathrm{rank}\,\mathcal{A}=\dim\mathrm{Im}\,\mathcal{A}=\dim V$, 即 $\mathcal{A}$ 是满射, 于是 $\mathcal{A}$ 为同构. 反之, 如果 $\mathcal{A}$ 是同构, 那么存在算子 $\mathcal{B}$ 使得 $\mathcal{AB}=\mathcal{BA}$ 为恒等变换 (也称为**单位算子**)$\mathcal{E}$, 从而 $\det\mathcal{A}\cdot\det\mathcal{B}=\det\mathcal{E}=1$, 所以 $\det\mathcal{A}\neq 0$. 于是, 算子的行列式不为零等价于算子可逆. 空间 V 上的可逆算子全体记作 $\mathrm{Aut}V$, 在乘法下是群, 称为 V 的**自同构群**. 式 (2.2.12) 中的映射限制在 $\mathrm{Aut}V$ 上给出了群同构

$$\varphi:\mathrm{Aut}V\to GL_n(K).$$

迹函数 tr 是线性的:

$$\mathrm{tr}\,(\lambda\mathcal{A}+\mu\mathcal{B})=\lambda\mathrm{tr}\,\mathcal{A}+\mu\mathrm{tr}\,\mathcal{B}.$$

(容易验证.) 而且, 式 (2.2.11) 表明

$$\mathrm{tr}\,(\mathcal{AB})=\mathrm{tr}\,(\mathcal{BA}).$$

迹函数的这些性质使用方便, 反映了算子很多重要的特征, 所以有着广泛的应用. 群表示论中的重要组成部分 —— 特征标理论就是建立在算子的迹的概念上.

五 在代数 $\mathcal{L}(V)$ 的子代数中, 最简单的就是由一个算子 $\mathcal{A}$**生成的子代数**$K[\mathcal{A}]$, 它是包含 $\mathcal{A}$ 的子代数中最小的那个. 显然, $\mathcal{A}$ 的所有的正次数幂都在 $K[\mathcal{A}]$ 中, 需要说明的是 $\mathcal{A}$ 的零次幂约定为恒等变换 $\mathcal{E}$, 也是 $K[\mathcal{A}]$ 的元素. 于是 $\mathcal{A}$ 的各种幂的线性组合

$$a_m\mathcal{A}^m+a_{m-1}\mathcal{A}^{m-1}+\cdots+a_1\mathcal{A}+a_0\mathcal{E}$$

也是 $K[\mathcal{A}]$ 的元素. 可以把上式看作多项式

$$f(t)=a_mt^m+a_{m-1}t^{m-1}+\cdots+a_1t+a_0\in K[t]$$

在 $t=\mathcal{A}$ 处的值 $f(\mathcal{A})$. 对 $f,g\in K[t],\lambda\in K$, 有

$$f(\mathcal{A})+g(\mathcal{A})=(f+g)(\mathcal{A}),\quad f(\mathcal{A})g(\mathcal{A})=(fg)(\mathcal{A}),\quad \lambda(f(\mathcal{A}))=(\lambda f)(\mathcal{A}).$$

所以, $\mathcal{A}$ 的各种幂的线性组合全体既是 $\mathcal{L}(V)$ 的子空间, 也在乘法下封闭, 从而是子代数. 由于这些线性组合都在 $K[\mathcal{A}]$ 中, 而 $K[\mathcal{A}]$ 是包含 $\mathcal{A}$ 的的子代数中最小的那个, 所以 $K[\mathcal{A}]$ 由这些线性组合构成:

$$K[\mathcal{A}]=\{f(\mathcal{A})\,|\,f(t)\in K[t]\}.$$

由于多项式的乘法是交换的, 故 $K[\mathcal{A}]$ 是交换代数. 它作为 $\mathcal{L}(V)$ 的子空间, 维数不超过 $n^2=\dim\mathcal{L}(V)=(\dim V)^2$. 于是算子 $\mathcal{A}$ 的幂之间有线性关系 (即有些幂是线性相关的). 也就是说, 有些多项式在 $\mathcal{A}$ 处的值为 0. 称 $K[t]$ 中的多项式 $f(t)$**零化**线性算子 $\mathcal{A}$ 如果 $f(\mathcal{A})=\mathcal{O}$.

定义 2.21 多项式 $f(t)$ 称为算子 $\mathcal{A}$ 的**极小多项式**如果

(1) $f(\mathcal{A})=\mathcal{O}$;

(2) $f(t)$ 的首项系数为 1;

(3) 在零化 $\mathcal{A}$ 的非零多项式中, $f(t)$ 的次数最低.

定理 2.22 每个线性算子都有极小多项式 $\mu_{\mathcal{A}}(t)$. 它的次数等于代数 $K[\mathcal{A}]$ 的维数. 算子 $\mathcal{A}$ 可逆当且仅当它的极小多项式 $\mu_{\mathcal{A}}(t)$ 的常数项不为 0.

证明 考虑序列 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^m$, $\cdots$. 由于这些算子都是 n^2 维向量空间 $\mathcal{L}(V)$ 中向量, $\mathcal{E}\neq 0$, 所以必有正整数 m 使得 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^m$ 线性相关, 但 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^{m-1}$ 线性无关. 于是 $\mathcal{A}^m$ 是 $\mathcal{A}$ 的低次幂的线性组合:

$$\mathcal{A}^m+\mu_1\mathcal{A}^{m-1}+\cdots+\mu_{m-1}\mathcal{A}+\mu_m\mathcal{E}=\mathcal{O}.$$

命 $\mu_{\mathcal{A}}(t)=t^m+\mu_1t^{m-1}+\cdots+\mu_1t+\mu_m$. 那么, 这个多项式零化 $\mathcal{A}$, 首项系数为 1. 由于 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^{m-1}$ 线性无关, 任何次数低于 m 的非零多项式不会零化 $\mathcal{A}$. 所以 $\mu_{\mathcal{A}}(t)$ 是 $\mathcal{A}$ 的极小多项式.

由于 $\mathcal{A}^m$ 是 $\mathcal{A}$ 的低次幂的线性组合, 所以只要 $k>m$, $\mathcal{A}^k$ 就是 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^{m-1}$ 的线性组合. 从而 $\mathcal{E}$, $\mathcal{A}$, $\mathcal{A}^2$, $\cdots$, $\mathcal{A}^{m-1}$ 是 $K[\mathcal{A}]$ 的基, $\dim K[\mathcal{A}]=m$.

如果常数项 $\mu_m=0$, 则有

$$\mathcal{A}(\mathcal{A}^{m-1}+\mu_1\mathcal{A}^{m-2}+\cdots+\mu_{m-1}\mathcal{E})=\mathcal{O}.$$

因为 $\mathcal{A}^{m-1}+\mu_1\mathcal{A}^{m-2}+\cdots+\mu_{m-1}\mathcal{E}\neq\mathcal{O}$, 所以 $\mathcal{A}$ 有零因子, 从而不是可逆的. 反之, 如果 $\mu_m\neq 0$, 则有

$$\mathcal{A}(-\mu_m^{-1}\mathcal{A}^{m-1}-\mu_m^{-1}\mu_1\mathcal{A}^{m-2}-\cdots-\mu_m^{-1}\mu_{m-1}\mathcal{E})=\mathcal{E}.$$

这个等式给出了 $\mathcal{A}$ 的逆元. □

定理 2.23　线性算子 $\mathcal{A}$ 的极小多项式整除任何零化 $\mathcal{A}$ 的多项式.

证明　根据带余除法, 对任何多项式 $f(t)$, 存在多项式 $q(t)$, $r(t)$ 使得

$$f(t)=q(t)\mu_{\mathcal{A}}(t)+r(t),\quad \deg r(t)<\deg\mu_{\mathcal{A}}(t).$$

如果 $f(\mathcal{A})=\mathcal{O}$, 则有 $r(\mathcal{A})=f(\mathcal{A})-q(\mathcal{A})\mu_{\mathcal{A}}(\mathcal{A})=\mathcal{O}$. 由极小多项式的定义知 $r(t)=0$. □

定义 2.24　称线性算子 $\mathcal{A}$ 是**幂零的**, 如果 $\mathcal{A}$ 的某个幂成为零算子. 称 m 是幂零算子 $\mathcal{A}$ 的**幂零指数**如果 $\mathcal{A}^m=\mathcal{O}$ 但 $\mathcal{A}^{m-1}\neq\mathcal{O}$.

根据定理 2.23, 算子 $\mathcal{A}$ 是幂零的当且仅当其极小多项式是单项式. 此时, 极小多项式的次数就是算子的幂零指数. 例 2.19 中的微分算子是幂零的, 如果 K 的特征为 0, 这个算子的幂零指数是 n.

习　题　2.2

下面习题中的向量空间都是有限维的, 除非另有说明.

1. 设 $A, B\in M_2(K)$. 验证映射 $\mathcal{T}: X\to AXB$ 是 $M_2(K)$ 上的线性算子并确定这个算子在基 E_{11}, E_{12}, E_{21}, E_{22} 下的矩阵.

2. 对次数不超过 n 的多项式形成的空间, 验证映射 $f(x)\to f(x+1)-f(x)$ 是线性算子.

3. 证明微分算子

(1) 在次数不超过 n 的多项式形成的空间上是退化的;

(2) 在函数 $\cos t$, $\sin t$, $\cdots$, $\cos nt$, $\sin nt$ 张成的空间上是非退化的.

4. 证明: 向量空间 V 的任何子空间 U 都可以是:

(1) 某个线性算子的核;

(2) 某个线性算子的像.

5. 证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}$, $\mathcal{B}$, 有等式

$$\operatorname{rank}\mathcal{A}=\operatorname{rank}\mathcal{BA}+\dim(\operatorname{Im}\mathcal{A}\cap\ker\mathcal{B}).$$

6. 利用上一题的等式证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, 弗罗宾纽斯不等式

$$\operatorname{rank}\mathcal{BA}+\operatorname{rank}\mathcal{AC}\leqslant\operatorname{rank}\mathcal{A}+\operatorname{rank}\mathcal{BAC}$$

成立.

7. 证明: 对有限维向量空间 V 上的任意线性算子 $\mathcal{A}$ 和任意正整数 i 都有等式

$$\dim(\operatorname{Im}\mathcal{A}^{i-1}\cap\ker\mathcal{A})=\dim\ker\mathcal{A}^{i}-\dim\ker\mathcal{A}^{i-1}.$$

(由于总是约定 $\mathcal{A}^0=\mathcal{E}$, 所以对 $i=1$, 公式是显然的.)

8. 证明: 域 K 上的两个同阶方阵 A, B 中如果有一个是非退化的, 那么 AB 和 BA 相似.

举例说明如果 A, B 都是退化的, 则 AB 和 BA 可以不相似.

9. 命 $gl_n(K)$ 为代数闭域 K 上的 n 阶方阵全体形成的向量空间. 这个空间中的元素 A 称为半单的如果它与对角矩阵相似. 设 $A\in gl_n(K)$. 证明:

(1) 映射 $f_A: X\to AX-XA$ 是 $gl_n(K)$ 上的线性算子;

(2) f_A 是退化的;

(3) 如果 A 是半单的, 那么 f_A 在某个基下的矩阵是对角矩阵.

10. 证明: 两个 n 阶实矩阵如果在复数域上相似, 则一定在实数域上相似.

11. 证明: 如果 A, B, C 分别是 $n\times p$, $p\times q$, $q\times n$ 矩阵, 那么

$$\operatorname{tr}(ABC)=\operatorname{tr}(BCA)=\operatorname{tr}(CAB).$$

12. 设 K 是 q 元域, V 是 K 上的 n 维向量空间. 求出 $\operatorname{Aut}V$ 的阶数 $|\operatorname{Aut}V|$. 由此知道 $|GL_n(K)|$. 进而求出 $GL_n(K)$ 中行列式为 1 的矩阵全体形成的群 $SL_n(K)$ 的阶数.

13. 证明:

(1) 微分算子在次数不超过 n 的多项式形成的空间上是幂零算子. 如果域的特征为 0, 这个算子的幂零指数为 $n+1$.

(2) 第二题中的算子是幂零的. 如果域的特征为 0, 那个算子的幂零指数为 $n+1$.

14. 假设域的特征是 0, 对第二题中的算子找一个基使得算子的矩阵是

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

15. 设 $\mathcal{A}$ 是二维向量空间上的线性算子. 证明多项式 $t^2-(\operatorname{tr}\mathcal{A})t+\det\mathcal{A}$ 零化 $\mathcal{A}$, 即 $\mathcal{A}^2-(\operatorname{tr}\mathcal{A})\mathcal{A}+(\det\mathcal{A})\mathcal{E}=0$.

16. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子. 称多项式 $f(t)$ 是算子 $\mathcal{A}$ 对 $v\in V$ 的**零化多项式**如果 $f(\mathcal{A})v=0$. 在算子 $\mathcal{A}$ 对 v 的零化多项式中次数最小且首项系数为 1 者称为 $\mathcal{A}$ 对向量 v 的**极小多项式**, 记为 $\mu_{\mathcal{A},v}(t)$. 假设基域 K 是无限域. 证明:

(1) $\mu_{\mathcal{A},v}(t)$ 整除 $\mu_{\mathcal{A}}(t)$;

(2) 存在向量 $u\in V$ 使得 $\mu_{\mathcal{A},u}(t)=\mu_{\mathcal{A}}(t)$.

17. 设 A 是特征为 0 的域 K 上的 n 阶方阵. 证明: 如果 A 的迹为零, 那么 A 相似于某个主对角线均取零值的矩阵 $B(B=(b_{ij}),\ b_{11}=b_{22}=\cdots=b_{nn}=0.)$

18. 上一题中域的特征为 0 是必要条件么?

2.3 不变子空间和特征向量

一 不变子空间 讨论线性算子的主要工具是不变子空间和特征向量. 我们的目标是对 V 上的算子 $\mathcal{A}$, 找出 V 的基, 使得 $\mathcal{A}$ 的矩阵尽可能简单.

定义 2.25 设 $\mathcal{A}$ 是向量空间 V 上的线性算子. 空间 V 的子空间 U 称为 $\mathcal{A}$ 的**不变子空间**如果 $\mathcal{A}U \subset U$. (根据一般的定义, $\mathcal{A}U = \{\mathcal{A}x \mid x \in U\}$.)

空间 V 本身和它的零子空间总是不变子空间, 称为**平凡的不变子空间**.

例 2.26 算子 $\mathcal{A}$ 的像 $\mathrm{Im}\,\mathcal{A}$ 与核 $\ker\mathcal{A}$ 是 $\mathcal{A}$ 的不变子空间, 它们可能是平凡的.

例 2.27 设 V 是子空间 U 与 W 的直和. 定义**沿着 W 到 U 的投影算子** $\mathcal{P}: V \to V$, $u + w \to u$, 其中 $u \in U$, $w \in W$. 那么 U 和 W 的子空间都是 $\mathcal{P}$ 的不变子空间.

例 2.28 在例 2.19 中, 对 $i = 1, 2, \cdots, n-1$, 空间 $V_i = \langle 1, t, \cdots, t^{i-1}\rangle$ 是算子 $\dfrac{d}{dt}$ 的不变子空间.

设 U 是 $\mathcal{A}$ 的不变子空间, 那么 $\mathcal{A}$ 在 U 上的限制 $\mathcal{A}_U$ 是 U 上的线性算子. 取 U 的基 $e_1, \cdots, e_k$, 然后扩充为 V 的基 $e_1, \cdots, e_n$. 在这个基下 $\mathcal{A}$ 的矩阵具有如下形式:

$$A = \begin{pmatrix} A_1 & A_0 \\ 0 & A_2 \end{pmatrix}, \tag{2.3.13}$$

其中 $A_1 = A_U$ 是 $\mathcal{A}_U$ 在基 $e_1, \cdots, e_k$ 下的矩阵, A_2 是 $n-k$ 阶方阵, A_0 是 $k \times (n-k)$ 矩阵. 反过来, 如果 $\mathcal{A}$ 在基 $e_1, \cdots, e_n$ 的矩阵具有形式 (2.3.13), 那么 $U = \langle e_1, \cdots, e_k\rangle$ 是 $\mathcal{A}$ 的不变子空间.

如果 $W = \langle e_{k+1}, \cdots, e_n\rangle$ 也是 $\mathcal{A}$ 的不变子空间, 那么 $\mathcal{A}$ 在基 $e_1, \cdots, e_n$ 下的矩阵就有分块对角的形式:

$$A = \begin{pmatrix} A_U & 0 \\ 0 & A_W \end{pmatrix} = A_U \dotplus A_W, \tag{2.3.14}$$

其中 A_U 是 $\mathcal{A}_U$ 在基 $e_1, \cdots, e_k$ 下的矩阵, A_W 是 $\mathcal{A}_W$ 在基 $e_{k+1}, \cdots, e_n$ 下的矩阵. 此时, 也说矩阵 A 是矩阵 A_U 和 A_W 的**对角和**; 算子 $\mathcal{A}$ 是 $\mathcal{A}_U$ 和 $\mathcal{A}_W$ 的**直和**, 记作

$$\mathcal{A} = \mathcal{A}_U \dotplus \mathcal{A}_W. \tag{2.3.15}$$

上面的讨论已经证明了下面的定理.

定理 2.29 线性算子 $\mathcal{A}: V \to V$ 在某个基下的矩阵具有分块对角的形式 (2.3.15) 当且仅当 V 是两个 $\mathcal{A}$ 不变子空间 U 和 W 的直和.

更一般地, 设 V 是 m 个不变子空间 $V_1, \cdots, V_m$ 的直和. 取 V 的基, 要求这个基与每个子空间 V_i 相合. 那么在这个基下, 算子 $\mathcal{A}$ 的矩阵就有如下的形式:

$$\begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{pmatrix}, \tag{2.3.16}$$

其中 A_i 是 $\mathcal{A}_{V_i}$ 的矩阵. 当每个子空间 V_i 的维数都是 1 时, 这个矩阵就是对角矩阵.

不过, 算子 $\mathcal{A}$ 未必有非平凡的直和分解, 例如对于例 2.19 中的微分算子 $\mathcal{D}$, 任何两个非零的不变子空间的交都是非零的. 又如实平面上绕原点的旋转一般只有平凡的不变子空间.

如果 U 是 $\mathcal{A}$ 的不变子空间, 那么 U 是 $\mathcal{A}^k$ 的不变子空间, 从而对任何多项式 $f(t)$, 都有 $f(\mathcal{A})U \subset U$. 如果 U 是 $\mathcal{A}$ 和 $\mathcal{B}$ 共同的不变子空间, 那么 U 也是 $\mathcal{AB}$ 和 $\mathcal{BA}$ 的不变子空间.

二　特征向量、特征多项式　从式 (2.3.16) 中的矩阵我们已经看到, 一维的不变子空间有特别的价值.

定义 2.30　空间 V 中的非零向量 v 称为线性算子 $\mathcal{A}: V \to V$ 的**特征向量**(eigenvector) 如果存在 $\lambda \in K$ 使得

$$\mathcal{A}v = \lambda v.$$

纯量 λ 称为 $\mathcal{A}$ 的 (对应于特征向量 v 的)**特征值**(eigenvalue). 特征向量和特征值有时也分别称为**本征向量**和**本征值**.

典型的例子是 $\dfrac{d}{dt}e^{\lambda t} = \lambda e^{\lambda t}$. 在二元可微函数空间中求算子 $\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$ 的特征值和特征向量是重要的问题.

显然, $\mathcal{A}v = \lambda v \Rightarrow \mathcal{A}^k v = \lambda^k v$, 所以对任意的多项式 $f \in K[t]$, 有

$$f(\mathcal{A})v = f(\lambda)v. \tag{2.3.17}$$

特别,

$$f(\mathcal{A}) = \mathcal{O} \ \Rightarrow \ f(\lambda) = 0. \tag{2.3.18}$$

一维不变子空间的任何非零向量都是特征向量. 反过来, 每个特征向量张成的子空间是一维不变子空间. 如果 V 的一个基由特征向量构成, 那么在这个基下算子的矩阵是对角的. 这当然是件美好的事情, 不过, 并不经常发生.

如果 u, v 都是 $\mathcal{A}$ 的以 λ 为特征值的特征向量, 那么对任意的 $a, b \in K$ 有

$$\mathcal{A}(au + bv) = a\mathcal{A}u + b\mathcal{A}v = \lambda(au + bv).$$

所以只要 $au + bv \neq 0$, 它就仍然是以 λ 为特征值的特征向量. 由于 $\mathcal{A}0 = \lambda 0 = 0$, 所以

$$V^{\lambda} = \{v \in V \mid \mathcal{A}v = \lambda v\}$$

是 V 的子空间, 称为对应于特征值 λ 的**特征 (子) 空间**(eigenspace). 这个特征空间的维数称为 λ 的**几何重数**. 显然有

$$V^{\lambda} = \ker(\mathcal{A} - \lambda\mathcal{E}) = \{x \in V \mid (\mathcal{A} - \lambda\mathcal{E})x = 0\}.$$

于是, λ 的几何重数等于 $n - \mathrm{rk}\,(\mathcal{A} - \lambda\mathcal{E})$. 以 λ 为特征值的特征向量存在等价于

$$\ker(\mathcal{A} - \lambda\mathcal{E}) \neq 0,$$

即算子 $\mathcal{A} - \lambda\mathcal{E}$ 是退化的:

$$\det(\mathcal{A} - \lambda\mathcal{E}) = 0. \tag{2.3.19}$$

从而 λ 是多项式 $\chi_{\mathcal{A}}(t) = \det(t\mathcal{E} - \mathcal{A}) = (-1)^n \det(\mathcal{A} - t\mathcal{E})$ 的根. 反过来, 如果 $\lambda \in K$ 是多项式 $\chi_{\mathcal{A}}(t)$ 的根, 那么 $\det(\mathcal{A} - \lambda\mathcal{E}) = 0$, 所以算子 $\mathcal{A} - \lambda\mathcal{E}$ 是退化的. 于是有 V 中的非零向量 v 使得 $(\mathcal{A} - \lambda\mathcal{E})v = 0$, 即 $\mathcal{A}v = \lambda v$. 因此 λ 是 $\mathcal{A}$ 的特征值. 我们已经证明了

定理 2.31 纯量 $\lambda \in K$ 是线性算子 $\mathcal{A} : V \to V$ 的特征值当且仅当算子 $\mathcal{A} - \lambda\mathcal{E}$ 是退化的, 即 λ 是多项式

$$\chi_{\mathcal{A}}(t) = \det(t\mathcal{E} - \mathcal{A})$$

的根. 这个多项式称为 $\mathcal{A}$ 的**特征多项式**(characteristic polynomial).

要求出算子的特征多项式必须把它化成矩阵的行列式形式. 设 $\mathcal{A}$ 在某个基 (e_i) 下的矩阵为 $A = (a_{ij})$. 根据定义, 有

$$\det(t\mathcal{E} - \mathcal{A}) = \det(tE - A).$$

从行列式的定义得

$$\det(tE - A) = \sum_{\sigma \in S_n} \varepsilon_\sigma (\delta_{1\sigma(1)}t - a_{1\sigma(1)})(\delta_{2\sigma(2)}t - a_{2\sigma(2)}) \cdots (\delta_{n\sigma(n)}t - a_{n\sigma(n)}).$$

所以

$$\chi_{\mathcal{A}}(t) = \det(tE - A) = t^n + \chi_1 t^{n-1} + \cdots + \chi_{n-1}t + \chi_n. \tag{2.3.20}$$

多项式 $\det(tE-A)$ 也称为 A 的**特征多项式**, 记作 $\chi_A(t)$. 于是算子 $\mathcal{A}$ 的特征多项式就是算子在某个基下的矩阵的特征多项式. 虽然我们知道 $\det(t\mathcal{E}-\mathcal{A})$ 不依赖基的选取的本质原因是相似的矩阵有相同的行列式, 但还是值得把这个结果用矩阵的语言再介绍一遍.

定理 2.32 相似的矩阵有相同的特征多项式.

证明 设 $A'=B^{-1}AB$. 那么

$$\begin{aligned}\det(tE-A')&=\det(tB^{-1}EB-b^{-1}AB)\\&=\det[B^{-1}(tE-A)B]=\det B^{-1}\cdot\det(tE-A)\cdot\det B\\&=\det(tE-A).\end{aligned}$$

□

算子的特征多项式可能没有根. 这时算子就没有特征向量. 正次数的复多项式一定有根, 所以复向量空间上的线性算子一定有特征向量. 不变子空间对于理解算子的特征多项式是很有益处的.

定理 2.33 设 V 的子空间 U 是线性算子 $\mathcal{A}:V\to V$ 的不变子空间, 那么 $\mathcal{A}_U$ 的特征多项式 $\chi_{\mathcal{A}_U}(t)$ 整除 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)$.

证明 取 U 的基, 然后扩充为 V 的基, 那么在 V 的这个基下 $\mathcal{A}$ 的矩阵具有形式 (2.3.13). 于是

$$\begin{aligned}\chi_{\mathcal{A}}(t)&=\det(tE_k-A_1)\cdot\det(tE_{n-k}-A_2)\\&=\chi_{\mathcal{A}_U}(t)\chi_{A_2}(t),\end{aligned}$$

其中 E_k 和 E_{n-k} 分别是 k 阶和 $n-k$ 阶单位矩阵. □

定义 2.34 线性算子 $\mathcal{A}$ 的特征值 λ 的**代数重数**就是 λ 作为 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)$ 的根的重数.

推论 2.35 一个特征值的几何重数不超过它的代数重数.

证明 设线性算子 $\mathcal{A}$ 的特征值 λ 的几何重数为 k. 根据定义, $k=\dim V^\lambda$. 由于 $U=V^\lambda$ 是 $\mathcal{A}$ 的不变子空间, 由定理 2.33 知 $\chi_{\mathcal{A}_U}(t)$ 整除 $\chi_{\mathcal{A}}(t)$ 但 $A_1=\lambda E_k$, 所以 $\chi_{\mathcal{A}_U}(t)=(t-\lambda)^k$. 从而 λ 的代数重数至少是 k. □

定理 2.36 设 V 是线性算子 $\mathcal{A}$ 的不变子空间 $V_1,\cdots,V_m$ 的直和, $\mathcal{A}_i$ 为 $\mathcal{A}$ 在 V_i 上的限制. 那么

$$\chi_{\mathcal{A}}(t)=\chi_{\mathcal{A}_1}(t)\chi_{\mathcal{A}_2}(t)\cdots\chi_{\mathcal{A}_m}(t).$$

证明 对每个 V_i, 取一个基. 把诸 V_i 的基合在一起得到 V 的一个基. 此时 $\mathcal{A}$ 的矩阵具有形式 (2.3.16). 于是

$$\begin{aligned}\chi_{\mathcal{A}}(t)&=\chi_{A_1}(t)\chi_{A_2}(t)\cdots\chi_{A_m}(t)\\&=\chi_{\mathcal{A}_1}(t)\chi_{\mathcal{A}_2}(t)\cdots\chi_{\mathcal{A}_m}(t).\end{aligned}$$

□

三　线性算子可对角化的判别准则　称线性算子 $\mathcal{A}$ 是**可对角化的**如果它在某个基下的矩阵具有如下形式:

$$A=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

显然, 算子可对角化等价于空间有特征向量构成的基.

引理 2.37　具有不同特征值的特征向量线性无关. 从而, 如果 $\lambda_1,\cdots,\lambda_m$ 是线性算子不同的特征值, 则特征空间 $V^{\lambda_1},\cdots,V^{\lambda_m}$ 线性无关, 即和 $V^{\lambda_1}+\cdots+V^{\lambda_m}$ 为直和.

证明　设 $\lambda_1,\cdots,\lambda_m$ 是线性算子 $\mathcal{A}$ 的特征值, 互不相同. 在每个 V^{λ_i} 中取非零元 v_i. 要证明向量 $v_1,\cdots,v_m$ 线性无关. 对 m 作归纳法. 当 $m=1$ 时, 结论显然. 假设

$$a_1v_1+a_2v_2+\cdots+a_mv_m=0,\quad a_i\in K.$$

用算子 $\mathcal{A}$ 作用到这个等式, 得

$$a_1\lambda_1v_1+a_2\lambda_2v_2+\cdots+a_m\lambda_mv_m=0.$$

用 λ_1 乘第一个等式, 然后减去第二个等式, 得

$$a_2(\lambda_1-\lambda_2)v_2+\cdots+a_m(\lambda_1-\lambda_m)v_m=0.$$

根据归纳假设, $a_i(\lambda_1-\lambda_i)=0$, $i=2,\cdots,m$. 但 $\lambda_1-\lambda_i\neq 0$ 如果 $i\neq 1$, 所以 $a_2=\cdots=a_m=0$. 进而, 从第一个等式得 $a_1v_1=0$. 于是 a_1 也为 0.

取 $u_i\in V^{\lambda_i}$. 如果 $u_1+\cdots+u_m=0$, 则上面的讨论表明所有的 u_i 都是零向量. 所以特征空间 $V^{\lambda_1},\cdots,V^{\lambda_m}$ 线性无关. □

显然, 算子的特征值及其几何重数都是算子的重要信息.

定义 2.38　算子 $\mathcal{A}$ 的**谱**, 记作 $\operatorname{Spec}\mathcal{A}$, 就是 $\mathcal{A}$ 的特征值全体, 但每个特征值按几何重数计算其个数 (即如果一个特征值的几何重数是 k, 那么这个特征值在谱中算为 k 个特征值). 方阵 A 的谱的定义类似. 谱中的元素称为**谱点**. 一个谱点称为**单的**如果它出现的次数是 1(即这个特征值的几何重数为 1). 如果一个算子的谱点都是单的, 则称算子的谱为**单的**.

定理 2.39　设 $\mathcal{A}$ 是 n 维向量空间 V 上的线性算子. 如果 $\mathcal{A}$ 有 n 个不同的特征值 (即 $\operatorname{Spec}\mathcal{A}$ 有 n 个单谱点), 那么 $\mathcal{A}$ 是可对角化的.

证明 设 $\lambda_1,\cdots,\lambda_n$ 是 $\mathcal{A}$ 的特征值. 对每个特征值 λ_i 取一个特征向量 v_i. 根据引理 2.35, 这些向量线性无关, 所以构成 V 的基. 在这个基下, $\mathcal{A}$ 的矩阵是 $\mathrm{diag}\,(\lambda_1,\cdots,\lambda_n)$. □

这个定理的条件只是一个充分条件, 重要的不是谱点的单性, 例 2.19 中的微分算子只有一个谱点 0, 它是单的, 但微分算子不是可对角化的, 又如恒等算子显然是可对角化的, 但它的谱点 1 不是单的. 其实重要的是按重数计算后有 $\dim V$ 个谱点.

定理 2.40 设 $\mathcal{A}$ 是 n 维向量空间 V 上的线性算子. 算子 $\mathcal{A}$ 可对角化的充要条件是它的谱 $\mathrm{Spec}\,\mathcal{A}$ 有 n 个谱点 (按重数计算), 等价的说法: 特征多项式 $\chi_{\mathcal{A}}(t)$ 的根都在 K 中, 且每个特征值的几何重数与其代数重数相等.

证明 先证必要性. 假设 $\mathcal{A}$ 可对角化, 那么 V 有一个基由特征向量构成. 设 $\lambda_1,\cdots,\lambda_m$ 是 $\mathcal{A}$ 的全部特征值, 则有

$$V=V^{\lambda_1}+\cdots+V^{\lambda_m}.$$

根据引理 2.36, 这个和是直和. 对每个 V^{λ_i} 取一个基. 把诸 V^{λ_i} 的基合起来得到 V 的一个基, 在这个基下, $\mathcal{A}$ 的矩阵是

$$A=\mathrm{diag}\,(\underbrace{\lambda_1,\cdots,\lambda_1}_{l_1\text{个}},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{l_m\text{个}}),$$

其中 $l_i=\dim V^{\lambda_i}$, $1\leqslant i\leqslant m$. 于是

$$\chi_{\mathcal{A}}(t)=(t-\lambda_1)^{l_1}\cdots(t-\lambda_m)^{l_m}.$$

由此可见, $\mathcal{A}$ 的特征多项式的根就是特征值, 全在 K 中, 且每个特征值 λ_i 的代数重数等于它的几何重数 l_i.

再看充分性. 如果 $\lambda_1,\cdots,\lambda_m$ 是 $\mathcal{A}$ 的所有的根, 重数分别是 $k_1,\cdots,k_m$, 那么它们都在 K 中, 所以都是算子 $\mathcal{A}$ 的特征值. 根据条件, 有

$$\dim V^{\lambda_i}=k_i,\quad k_1+\cdots+k_m=n.$$

根据引理 2.36, 这些特征空间线性无关, 所以它们的和是直和, 维数等于 n. 从而

$$V=V^{\lambda_1}\oplus\cdots\oplus V^{\lambda_m}.$$

对每个 V^{λ_i} 取一个基. 把诸 V^{λ_i} 的基合起来得到 V 的一个基. 空间 V 的这个基由特征向量构成, 从而 $\mathcal{A}$ 在这个基下的矩阵是对角的. □

四 不变子空间的存在性 算子的非平凡不变子空间的存在性与其特征多项式在 $K[t]$ 中的可分解性密不可分. 定理 2.33 表明, 如果算子 $\mathcal{A}$ 有非平凡的不变子

空间, 那么算子的特征多项式在 $K[t]$ 中可以分解成两个次数小于 $n = \dim V$ 的多项式的乘积. 这个结果的逆命题也是对的, 我们在下一节证明它 (定理 2.46). 多项式的可分解性与基域密切相关. 代数闭域 (特别, 复数域) 上的一元多项式可以分解成线性多项式的乘积, 实数域上的多项式可以分解成次数不超过 2 的多项式的乘积. 这些事实指向代数闭域和实数域上的向量空间的线性算子一般会有非平凡的不变子空间. 更确切地说, 有如下的结论.

定理 2.41 代数闭域上的向量空间的线性算子有一维的不变子空间.

证明 当域 K 是代数闭域时, 系数在 K 中的正次数多项式的根都在 K 中. 由定理 2.31 知这个算子有特征值, 所以有特征向量. □

定理 2.42 实向量空间上的线性算子一定有一维或二维的不变子空间.

证明 设 $\mathcal{A}$ 是实向量空间 V 上的线性算子, 它的极小多项式 $\mu_{\mathcal{A}}(t)$ 有一次或二次因子. 如果 $\mu_{\mathcal{A}}(t)$ 有一次因子 $t-a$, 那么 $\mu_{\mathcal{A}}(t) = (t-a)\nu(t)$. 而且多项式 $\nu(t)$ 不能零化 $\mathcal{A}$ 因为其次数小于 $\mu_{\mathcal{A}}(t)$ 的次数. 取 $u \in V$ 使得 $v = \nu(\mathcal{A})u \neq 0$. 那么

$$(\mathcal{A} - a\mathcal{E})v = (\mathcal{A} - a\mathcal{E})\nu(\mathcal{A})u = \mu_{\mathcal{A}}(\mathcal{A})u = 0.$$

于是 v 是 $\mathcal{A}$ 的特征向量, 特征值为 a.

如果 $\mu_{\mathcal{A}}(t)$ 有二次因子 $t^2 - at - b$, 那么 $\mu_{\mathcal{A}}(t) = (t^2 - at - b)\nu(t)$. 而且多项式 $\nu(t)$ 不能零化 $\mathcal{A}$ 因为其次数小于 $\mu_{\mathcal{A}}(t)$ 的次数. 取 $u \in V$ 使得 $v = \nu(\mathcal{A})u \neq 0$. 假设 $\mu_{\mathcal{A}}(t)$ 没有一次因子, 那么 $\mathcal{A}v$ 和 v 线性无关. 由于

$$\begin{aligned}
\mathcal{A}^2 v - a\mathcal{A}v - bv &= (\mathcal{A}^2 - a\mathcal{A} - b\mathcal{E})v \\
&= (\mathcal{A}^2 - a\mathcal{A} - b\mathcal{E})\nu(\mathcal{A})u \\
&= \mu_{\mathcal{A}}(\mathcal{A})u = 0,
\end{aligned}$$

所以 $\mathcal{A}(\mathcal{A}v) = a\mathcal{A}v + bv$. 从而 $\mathcal{A}v$ 和 v 张成的空间是二维不变子空间. □

习 题 2.3

1. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子. 证明: 如果 $V_1, \cdots, V_m$ 是 $\mathcal{A}$ 的不变子空间, 那么这些子空间的和与交都是 $\mathcal{A}$ 的不变子空间.

2. 证明: 如果 V 的子空间 U 在 V 的每个线性算子下都是不变的, 则 U 是 V 的平凡子空间, 即 $U = \{0\}$ 或 $U = V$.

3. 设 $\mathcal{A}, \mathcal{B}$ 是向量空间 V 上的线性算子. 证明: 如果 $\mathcal{AB} = \mathcal{BA}$, 那么对任意的 $\lambda \in K$, $\ker(\lambda\mathcal{E} - \mathcal{A})$ 是 $\mathcal{B}$ 的不变子空间.

4. 证明: 如果 $\mathcal{A}$ 是可逆线性算子, 那么 $\mathcal{A}$ 的不变子空间也是 $\mathcal{A}^{-1}$ 的不变子空间. 特别, $\mathcal{A}$ 的特征向量也是 $\mathcal{A}^{-1}$ 的特征向量.

5. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 其平方是恒等算子. 证明: 对任意的向量 $v \in V$, 向量 $v - \mathcal{A}v$ 是零向量或以 -1 为特征值的特征向量. 如果 $\operatorname{char} K \neq 2$, 证明 V 是特征空间 V^1 和 V^{-1} 的直和.

6. 设 $\mathcal{A}$ 是向量空间 V 上的**幂等线性算子**(即 $\mathcal{A}^2 = \mathcal{A}$). 证明:

(1) $\mathcal{B} = \mathcal{E} - \mathcal{A}$ 也是幂等线性算子, 而且 $\mathcal{AB} = \mathcal{BA}$, $\mathcal{E} = \mathcal{A} + \mathcal{B}$;

(2) $\operatorname{Im} \mathcal{A}$ 是 $\mathcal{A}$ 的以 1 为特征值的特征空间 V^1, $\ker \mathcal{A}$ 是 $\mathcal{A}$ 的以 0 为特征值的特征空间;

(3) $\ker \mathcal{A} = \operatorname{Im} \mathcal{B}$;

(4) $V = \operatorname{Im} \mathcal{A} \oplus \ker \mathcal{A} = \operatorname{Im} \mathcal{A} \oplus \operatorname{Im} (\mathcal{E} - \mathcal{A})$.

7. 更一般地, 设 $\mathcal{A}_1, \cdots, \mathcal{A}_{m-1}$ 是向量空间 V 上的幂等线性算子, 且相互正交 (即 $\mathcal{A}_i\mathcal{A}_j = \mathcal{A}_j\mathcal{A}_i = \mathcal{O}$ 如果 $i \neq j$). 证明:

(1) $\mathcal{A} = \mathcal{A}_1 + \cdots + \mathcal{A}_{m-1}$ 是幂等线性算子, 且 $\mathcal{A}\mathcal{A}_i = \mathcal{A}_i\mathcal{A} = \mathcal{A}_i$, $1 \leqslant i \leqslant m-1$;

(2) 置 $\mathcal{A}_m = \mathcal{E} - \mathcal{A}$, 则有 $\mathcal{A}_m\mathcal{A}_i = \mathcal{A}_i\mathcal{A}_m = \mathcal{O}$, $1 \leqslant i \leqslant m-1$;

(3) $V = \operatorname{Im} \mathcal{A}_1 \oplus \operatorname{Im} \mathcal{A}_2 \oplus \cdots \oplus \operatorname{Im} \mathcal{A}_m$.

8. 设 $\mathcal{D}: M_n(K) \to M_n(K)$ 是域 K 上的 n 阶方阵空间上的非零线性算子. 证明: 如果它保持乘法, 即对所有 $A, B \in M_n(K)$ 有

$$\mathcal{D}(AB) = \mathcal{D}(A)\mathcal{D}(B),$$

那么, 存在非退化矩阵 $C \in M_n(K)$ 使得 $\mathcal{D} = f_C$(见习题 2.1 第 9 题) 找出 f_C 的以 1 为特征值的特征空间.

9. 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 对某个正整数 k 有 $\operatorname{Im} \mathcal{A}^k = \operatorname{Im} \mathcal{A}^{k+1}$. 证明: 在这种情形有 $V = \ker \mathcal{A}^p \oplus \operatorname{Im} \mathcal{A}^p$ 是 $\mathcal{A}$ 的两个不变子空间的直和.

10. 找出下列线性算子的特征值和特征空间:

(1) 次数不超过 n 的实多项式空间上的微分算子;

(2) 次数不超过 n 的实多项式空间上的算子 $t\dfrac{d}{dt}$;

(3) 次数不超过 n 的实多项式空间上的算子 $\dfrac{1}{t}\displaystyle\int_0^t f(x)dx$;

(4) 方阵空间 $M_n(K)$ 上的算子 $X \to {}^tX$;

(5) 函数空间 $\langle \cos t, \sin t, \cdots, \cos mt, \sin mt \rangle$ 上的算子 $f \to \dfrac{d^n f}{dt^n}$;

(6) 函数空间 $\langle \cos t, \sin t, \cdots, \cos mt, \sin mt \rangle$ 上的算子 $f \to \displaystyle\int_0^t f(x)dx$.

11. 证明: 次数不超过 n 的实多项式空间上的算子 $f(t) \to f(at+b)$ 的特征值是 $1, a, \cdots, a^n$.

12. 证明: 如果 λ^2 是线性算子 $\mathcal{A}^2$ 的特征值, 那么 λ 和 $-\lambda$ 中有一个是 $\mathcal{A}$ 的特征值.

13. 设 A, B 是域 K 上的 n 阶方阵, 且 $AB = BA$, B 幂零. 证明: $\det(tE - (A+B)) = \det(tE - A)$.

14. 证明: 幂等方阵 A(即 $A^2 = A$) 相似于分块对角矩阵 $\operatorname{diag}(I_r, 0)$, 其中 $r = \operatorname{rank} A$, I_r 是 r 阶单位矩阵. 利用这一结果证明: 对特征为 0 的域上的幂等矩阵 A 有 $\operatorname{rank} A = \operatorname{tr} A$.

15. 证明: 如果 n 维向量空间上的线性算子 $\mathcal{E}$, $\mathcal{A}$, $\cdots$, $\mathcal{A}^{n-1}$ 线性无关, 那么, 存在向量 $v \in V$ 使得

$$V = \langle v, \mathcal{A}v, \cdots, \mathcal{A}^{n-1}v \rangle.$$

(此时说 V 是 $\mathcal{A}$ 的**循环空间**.) 而且, $\mathcal{A}$ 的特征多项式就是 $\mathcal{A}$ 的极小多项式.

16. 设 A 是 n 阶实方阵, 没有实的特征根, 特别, n 是偶数且 A 可逆. 证明: 存在实矩阵 B 使得 $AB = BA$ 且 $B^2 = -E$, 其中 E 是单位矩阵.

17. 证明: 对任意的方阵 $A, B \in M_n(K)$, 方阵 AB 和 BA 的特征多项式相同. 举例说明 AB 和 BA 的极小多项式可以不同.

18. 利用简单的关系式

$$A = a_0E + a_1B + a_2B^2, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

求出复数域上的循环矩阵

$$A = \begin{pmatrix} a_0 & a_1 & a_2 \\ a_2 & a_0 & a_1 \\ a_1 & a_2 & a_0 \end{pmatrix}$$

的特征根.

19. 设 $A, B \in M_n(K)$, $\operatorname{char} K \neq 2$. 若有 $D = \operatorname{diag}(\theta_1, \cdots, \theta_n)$, $\theta_i = \pm 1$, 使得 $B = DA$, 则记为 $B \sim A$. 显然 $\sim$ 是等价关系. 如果 $S(A)$ 表含 A 的等价类, 那么 $\operatorname{Card} S(A) \leqslant 2^n$.

证明: 每个等价类中都含有矩阵, 1 不是其特征根.

20. 对 $A \in M_n(K)$, 称非零列向量 $C = [c_1, \cdots, c_n] \in K^n$ 是 A 的特征向量如果存在 $\lambda \in K$ 使得 $AC = \lambda C$. 纯量 λ 称为 A 的特征值. 证明: A 的特征值一定是 A 的特征多项式的根.

21. 计算特征多项式并求出复特征值和特征向量:

(1) $\begin{pmatrix} -2 & 2 \\ -2 & 3 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$; (3) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$;

(4) $\begin{pmatrix} -2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$; (5) $\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$.

22. 方阵 A 和 ${}^tA$ 是否有相同的特征向量? 是否有相同的特征值?

23. 对三阶方阵 $A = (a_{ij})$, 证明其特征多项式中一次项的系数为 2 阶对称子式之和:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$

24. 设 $\mathcal{A}, \mathcal{B}$ 是有限维复向量空间上的线性算子. 证明: 如果 $\mathcal{AB} - \mathcal{BA}$ 的秩不超过 1, 那么这两个算子有共同的特征向量.

2.4 商算子和对偶算子

一 商算子 设 U 是 V 的子空间. 如果 U 是 V 上的线性算子 $\mathcal{A}$ 的不变子空间, 那么在商空间 $\bar{V}=V/U$ 上可以定义 $\mathcal{A}$ 的**商算子**$\bar{\mathcal{A}}$ 如下:

$$\bar{\mathcal{A}}\bar{x}=\overline{\mathcal{A}x},\quad \text{即}\quad \bar{\mathcal{A}}(x+U)=\mathcal{A}x+U.$$

这个定义是明确的: 如果 $x+U=y+U$, 那么 $x-y=z\in U$, 从而 $\mathcal{A}x+U=\mathcal{A}y+\mathcal{A}z+U=\mathcal{A}y+U$ 因为 $\mathcal{A}z\in U$.

取 V 的基 $e_1,\cdots,e_n$ 使得 $e_1,\cdots,e_k$ 是 U 的基, 那么 $\bar{e}_{k+1},\cdots,\bar{e}_n$ 是 $\bar{V}$ 的基. 于是算子 $\mathcal{A}$ 在基 (e_i) 下的矩阵具有式 (2.3.13) 的形式

$$A=\begin{pmatrix} A_1 & A_0 \\ 0 & A_2 \end{pmatrix},$$

其中 A_1 是算子 $\mathcal{A}_U$ 在基 $e_1,\cdots,e_k$ 的矩阵. 显然 A_2 是 $\bar{\mathcal{A}}$ 在基 $(\bar{e}_j)$ 下的矩阵. 因此定理 2.33 有如下更精细的形式.

定理 2.43 设 V 的子空间 U 是线性算子 $\mathcal{A}:V\to V$ 的不变子空间, 那么

$$\chi_{\mathcal{A}}(t)=\chi_{\mathcal{A}_U}(t)\cdot\chi_{\bar{\mathcal{A}}}(t),$$

即 $\mathcal{A}$ 的特征多项式是 $\mathcal{A}_U$ 的特征多项式与 $\bar{\mathcal{A}}$ 的特征多项式的乘积.

如果 $e_{k+1},\cdots,e_n$ 张成的子空间 W 也是 $\mathcal{A}$ 的不变子空间, 那么 $\mathcal{A}_W$ 在基 $e_{k+1},\cdots,e_n$ 下的矩阵就是 $\bar{\mathcal{A}}$ 在基 $\bar{e}_{k+1},\cdots,\bar{e}_n$ 下的矩阵. 所以这两个算子本质上是一样的. 这一点通过定理 1.46 建立的同构

$$f:W\to\bar{V},\quad x\to x+U$$

可以说得更清楚, 我们有如下的映射交换图:

$$\begin{tikzcd} W \arrow[r, "\mathcal{A}_W"] \arrow[d, "f"'] & W \arrow[d, "f"] \\ V/U \arrow[r, "\bar{\mathcal{A}}"'] & V/U \end{tikzcd}$$

即 $f\cdot\mathcal{A}_W=\bar{\mathcal{A}}\cdot f$, 或等价地 $\bar{\mathcal{A}}=f\cdot\mathcal{A}_W\cdot f^{-1}$.

二 作为商算子的一个应用, 我们证明下面的重要结论.

定理 2.44(凯莱–哈密顿定理) 设 $\mathcal{A}$ 是向量空间 V 上的线性算子, 那么 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)$ 零化 $\mathcal{A}$. 等价的说法是, 方阵 A 的特征多项式 $\chi_A(t)$ 零化 A, 即 $\chi_A(A)=0$.

证明 对 $n=\dim V$ 的维数作归纳法. 设 K 是 V 的基域. 当 $n=1$ 时, V 由一个非零向量 v 张成, 从而存在 $\lambda\in K$ 使得 $\mathcal{A}v=\lambda v$, 于是 $\chi_{\mathcal{A}}(t)=t-\lambda$ 且 $\chi_{\mathcal{A}}(\mathcal{A})=\mathcal{A}-\lambda\mathcal{E}=0$. 结论这时成立.

假设 $n>1$ 且结论对维数小于 n 的向量空间成立.

如果 V 有非平凡的不变子空间 U, 那么 U 的维数和商空间 $\bar{V}=Y/U$ 的维数都小于 n. 由归纳假设, $\chi_{\mathcal{A}_U}(t)$ 零化 $\mathcal{A}_U$, $\chi_{\bar{\mathcal{A}}}(t)$ 零化 $\bar{\mathcal{A}}$. 于是, 对任意的 $u\in U$ 和 $x\in V$, 有

$$\chi_{\mathcal{A}_U}(\mathcal{A})u=0,\quad \chi_{\bar{\mathcal{A}}}(\bar{\mathcal{A}})\bar{x}=\bar{0}.$$

第二个等式的含义是 $u=\chi_{\bar{\mathcal{A}}}(\mathcal{A})x\in U$. 于是, 对任意的 $x\in V$, 利用定理 2.43 和上面的两个等式得

$$\begin{aligned}\chi_{\mathcal{A}}(\mathcal{A})x&=\chi_{\mathcal{A}_U}(\mathcal{A})\cdot\chi_{\bar{\mathcal{A}}}(\mathcal{A})x\\&=\chi_{\mathcal{A}_U}(\mathcal{A})u\\&=0.\end{aligned}$$

所以此时 $\chi_{\mathcal{A}}(\mathcal{A})=\mathcal{O}$.

如果 $\mathcal{A}$ 没有非平凡的不变子空间, 那么对任意的非零向量 $v\in V$, 向量 v, $\mathcal{A}v,\cdots,\mathcal{A}^{n-1}v$ 线性无关. 否则某个 $\mathcal{A}^iv$, $1\leqslant i\leqslant n-1$, 是 $v,\mathcal{A}v,\cdots,\mathcal{A}^{i-1}v$ 的线性组合, 从而 $\langle v,\mathcal{A}v,\cdots,\mathcal{A}^{i-1}v\rangle$ 是 $\mathcal{A}$ 的非平凡不变子空间, 这与假设矛盾. 于是 $v,\mathcal{A}v,\cdots,\mathcal{A}^{n-1}v$ 形成 V 的基, $\mathcal{A}^nv$ 是这些向量的线性组合:

$$\mathcal{A}^nv=a_0v+a_1\mathcal{A}v+\cdots+a_{n-1}\mathcal{A}^{n-1}v. \tag{2.4.21}$$

这样一来, $\mathcal{A}$ 在这个基下的矩阵是

$$A=\begin{pmatrix}0&0&0&\cdots&0&a_0\\1&0&0&\cdots&0&a_1\\0&1&0&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&a_{n-2}\\0&0&0&\cdots&1&a_{n-1}\end{pmatrix}.$$

于是

$$\chi_{\mathcal{A}}(t)=\chi_A(t)=\det(tE-A)=t^n-a_{n-1}t^{n-1}-\cdots-a_1t-a_0.$$

等式 (2.4.21) 表明 $\chi_{\mathcal{A}}(\mathcal{A})v=0$. 所以对任意的非负整数 k, 有

$$\chi_{\mathcal{A}}(\mathcal{A})\mathcal{A}^kv=\mathcal{A}^k\chi_{\mathcal{A}}(\mathcal{A})v=0.$$

由于 $v, \mathcal{A}v, \cdots, \mathcal{A}^{n-1}v$ 是 V 的基, 从而对任意的 $x \in V$, 有 $\chi_{\mathcal{A}}(\mathcal{A})x = 0$. 即 $\chi_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. □

推论 2.45 算子 $\mathcal{A}$ 的极小多项式 $\mu_{\mathcal{A}}(t)$ 是其特征多项式 $\chi_{\mathcal{A}}(t)$ 的因子. 如果 λ 是 $\mathcal{A}$ 的特征值, 那么 $t-\lambda$ 整除 $\mu_{\mathcal{A}}(t)$.

证明 因为 $\mu_{\mathcal{A}}(t)$ 整除任何零化 $\mathcal{A}$ 的多项式 (定理 2.23), 由定理 2.44 知第一个结论成立. 如果 λ 是 $\mathcal{A}$ 的特征值, 根据式 (2.3.18), 有

$$\mathcal{A}v = \lambda v \ \Rightarrow \ 0 = \mu_{\mathcal{A}}(\mathcal{A})v = \mu_{\mathcal{A}}(\lambda)v \ \Rightarrow \ \mu_{\mathcal{A}}(\lambda) = 0 \ \Rightarrow \ (t-\lambda) \mid \mu_{\mathcal{A}}(t).$$

□

下面的结论是定理 2.33 的逆命题.

定理 2.46 假设算子 $\mathcal{A}$ 的特征多项式在 $K[t]$ 中能分解成两个正次数的多项式的乘积, 那么 $\mathcal{A}$ 有非平凡的不变子空间.

证明 设 $\chi_{\mathcal{A}}(t) = \xi(t)\eta(t)$, 其中 ξ, η 的次数分别是正整数 p, q. 取 V 中的非零向量 v. 如果 $\eta(\mathcal{A})v = 0$, 那么 $\mathcal{A}^q v$ 是 $\mathcal{A}^{q-1}v, \cdots, \mathcal{A}v, v$ 的线性组合. 所以 $\mathcal{A}^{q-1}, \cdots, \mathcal{A}v, v$ 张成的子空间 U 是 $\mathcal{A}$ 的不变子空间, 且是非平凡的, 因为 $n > q \geqslant \dim U \geqslant 1$.

假设 $u = \eta(\mathcal{A})v \neq 0$. 由凯莱–哈密顿定理知 $\xi(\mathcal{A})u = \xi(\mathcal{A})\eta(\mathcal{A})v = \chi_{\mathcal{A}}(\mathcal{A})v = 0$. 如同讨论 $\eta(\mathcal{A})v = 0$ 情形知 $\mathcal{A}^{p-1}u, \cdots, \mathcal{A}u, u$ 张成的子空间 W 是 $\mathcal{A}$ 的非平凡的不变子空间. □

三 商算子的另一个应用是证明下面的结论.

定理 2.47 设 $\mathcal{A}$ 是代数闭域上的向量空间 V 上的线性算子, 那么 $\mathcal{A}$ 是可以**上三角化的**, 即 $\mathcal{A}$ 在某个基下的矩阵是上三角矩阵.

证明 对 $n = \dim V$ 的维数作归纳法. 当 $n = 1$ 时, 结论是显然的. 假设 $n > 1$ 且结论对 $n-1$ 维的向量空间成立. 根据定理 2.41, $\mathcal{A}$ 有一维的不变子空间. 设 U 是 $\mathcal{A}$ 的一维的不变子空间, 那么 $\bar{V} = V/U$ 是 $n-1$ 维的. 由归纳假设, 存在 $\bar{V}$ 的基 $\bar{e}_2, \cdots, e_n$ 使得商算子 $\bar{\mathcal{A}}$ 在该基下的矩阵 A_2 是上三角的. 取 U 中的非零向量 e_1, 那么 $e_1, \cdots, e_n$ 是 V 的基, 在这个基下 $\mathcal{A}$ 的矩阵

$$A = \begin{pmatrix} \lambda & A_0 \\ 0 & A_2 \end{pmatrix}$$

是上三角的. □

四 对偶算子 定理 2.47 还可以通过对偶空间的算子的性质证明. 设 V 是域 K 上的有限维向量空间, $\mathcal{A}$ 是 V 上的线性算子. 线性映射的合成仍是线性映射, 所以 V 上的线性函数与 $\mathcal{A}$ 合成给出了映射 $\mathcal{A}^*: V^* \to V^*$, $f \to f\mathcal{A}$. 由于

$$(\alpha f)\mathcal{A} = \alpha(f\mathcal{A}), \quad (f+g)\mathcal{A} = f\mathcal{A} + g\mathcal{A},$$

其中 $f,g\in V^*$, $\alpha\in K$, 所以 $\mathcal{A}^*$ 是 V^* 上的线性算子, 称为 $\mathcal{A}$ 的**对偶算子**.

于是我们得到了一个从 $\mathcal{L}(V)$ 到 $\mathcal{L}(V^*)$ 的映射 $*$: $\mathcal{A}\to\mathcal{A}^*$. 采用 1.7 节中第一部分的记号, 则有

$$(\mathcal{A}^*f,x)=(f,\mathcal{A}x),$$

其中 $f\in V^*$, $x\in V$. 由此可见

$$\begin{aligned}&\mathcal{O}_V^*=\mathcal{O}_{V^*},\quad \mathcal{E}_V^*=\mathcal{E}_{V^*},\quad (\alpha\mathcal{A})^*=\alpha\mathcal{A}^*,\\&(\mathcal{A}+\mathcal{B})^*=\mathcal{A}^*+\mathcal{B}^*,\quad (\mathcal{A}\mathcal{B})^*=\mathcal{B}^*\mathcal{A}^*.\end{aligned}\tag{2.4.22}$$

利用算子的对偶可以给定理 2.47 另一个证明. 取算子 $\mathcal{A}^*$ 的一个特征向量 f, 相应的特征值为 λ. 根据定理 2.41, 这是可以办到的. 那么 $\ker f$ 是 V 的 $n-1$ 维子空间. 对 $x\in\ker f$, 有

$$(f,\mathcal{A}x)=(\mathcal{A}^*f,x)=(\lambda f,x)=\lambda(f,x)=\lambda f(x)=0.$$

所以 $U=\ker f$ 是 $\mathcal{A}$ 的不变子空间. 归纳假设说 $\mathcal{A}_U$ 在 U 的某个基 $e_1,\cdots,e_{n-1}$ 下的矩阵 A_1 是上三角形的. 取 $e_n\in V\setminus U$, 那么 $e_1,\cdots,e_n$ 是 V 的基, 且在这个基下 $\mathcal{A}$ 的矩阵

$$A=\begin{pmatrix}A_1&A_0\\0&\lambda_n\end{pmatrix}$$

是上三角形的. □

利用对偶基, 可以建立算子 $\mathcal{A}$ 和 $\mathcal{A}^*$ 的矩阵之间的联系. 设 $\mathcal{A}$ 在 V 的基 (e_i) 下的矩阵是 $A=(a_{ij})$, $\mathcal{A}^*$ 在对偶基 (e^j) 下的矩阵是 $A^*=(a^*_{ij})$. 那么

$$\mathcal{A}e_j=\sum_{k=1}^n a_{kj}e_k,\quad \mathcal{A}^*e^j=\sum_{k=1}^n a^*_{kj}e^k.$$

于是

$$\begin{aligned}(e^i,\mathcal{A}e_j)&=\sum_{k=1}^n a_{kj}(e^i,e_k)=\sum_{k=1}^n a_{kj}\delta_{ik}=a_{ij},\\(e^i,\mathcal{A}e_j)&=(\mathcal{A}^*e^i,e_j)=\sum_{k=1}^n a^*_{ki}(e^k,e_j)=\sum_{k=1}^n a^*_{ki}\delta_{kj}=a^*_{ji}.\end{aligned}$$

所以 $a_{ij}=a^*_{ji}$. 即 A^* 是 A 的转置 ${}^tA$. 刚才的讨论可以总结如下.

定理 2.48　设 V 上的线性算子 $\mathcal{A}$ 在基 (e_i) 下的矩阵是 A, 那么对偶算子 $\mathcal{A}^*$ 在对偶基 (e^i) 下的矩阵 A^* 是 A 的转置矩阵: $A^*={}^tA$.

矩阵转置后再转置就成为它自身了, 所以可以想象算子的对偶的对偶就是算子自身. 如果把 V 与 V^{**} 等同起来, 算子的对偶的对偶的确是它自身:

$$(f, \mathcal{A}x) = (\mathcal{A}^* f, x) = (f, \mathcal{A}^{**} x),$$

即 $\mathcal{A}^{**} = \mathcal{A}$. 于是映射 $* : \mathcal{L}(V) \to \mathcal{L}(V^*)$, $\mathcal{A} \to \mathcal{A}^*$ 是双射, 它保持了加法和纯量乘, 对乘法则反了一个顺序: $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$. 这样的双射称为代数的**反同构**.

习　题　2.4

1. 利用推论 2.45 求出下列矩阵的极小多项式 (其定义和算子的极小多项式的定义类似):

$$(1)\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}; \qquad (2)\begin{pmatrix} 4 & -2 & 2 \\ -5 & 7 & -5 \\ -6 & 7 & -4 \end{pmatrix}.$$

2. 凯莱–哈密顿定理 (定理 2.44) 可以用于计算多项式在线性算子和方阵处的值. 设 $\mathcal{A}$ 是向量空间上 V 上的线性算子, $f(t)$ 是多项式. 带余除法给出

$$f(t) = q(t)\chi_{\mathcal{A}}(t) + r(t).$$

于是

$$f(\mathcal{A}) = r(\mathcal{A}).$$

如果 λ 是 $\chi_{\mathcal{A}}(t)$ 的 k 重根, 那么, 对 $i = 0, 1, \cdots, k-1$, 多项式 f 的 i 阶导数在 λ 处的值 $f^{(i)}(\lambda)$ 等于 r 的 i 阶导数在 λ 处的值 $r^{(i)}(\lambda)$. (约定 $f^{(0)} = f$.) 这个性质对计算 r 是有用的.

计算:

$$(1)\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{100}; \qquad (2)\begin{pmatrix} 7 & -4 \\ 14 & -8 \end{pmatrix}^{96}; \qquad (3)\begin{pmatrix} -2 & 8 & 6 \\ -4 & 10 & 6 \\ 4 & -8 & -4 \end{pmatrix}^{50}.$$

3. 假设基域是代数闭域. 证明: 线性算子 $\mathcal{A}$ 是幂零的当且仅当其特征多项式没有非零根.

2.5　约当标准形

我们曾经提到, 对 V 上的算子 $\mathcal{A}$, 要找 V 的基, 使得 $\mathcal{A}$ 的矩阵尽可能简单. 这对理论探讨和实际的计算及应用都是特别有意义的, 也是算子理论的目标. 虽然不同的地方对矩阵的简单性有不同的含义, 但一般而言, 简单的矩阵有很多的零, 而且这些零处在恰当位置从而可以很方便地做计算. 在定理 2.44 的证明中出现的矩阵虽然有很多的 0, 但那不是一个简单的矩阵, 因为计算那个矩阵的幂是很不容易的.

对角矩阵是简单的矩阵, 例 2.19 和定理 2.39 表明一般的线性算子不是可对角化的. 2.3 节中的讨论, 如定理 2.29 和式 (2.3.13) 等, 显示不变子空间对寻找符合要求的基是非常有价值的. 引理 2.37 说不同的特征空间是线性无关的. 这给出一个解决问题的思路, 把一般的情形归结到只有一个特征值的情形, 再考虑只有一个特征值的算子.

一　根子空间　如果基域是代数闭的, 那么只有一个特征值的算子的特征多项式具有形式 $(t-\lambda)^n$. 这时特征空间 V^λ 一般不是整个空间, 如例 2.19 中的微分算子只有一个特征值 0, 在基域 K 的特征为 0 时, 相应的特征空间是一维的. 不过, 定理 2.44 说对 V 中任意的向量 v, 有 $(\mathcal{A}-\lambda\mathcal{E})^n v=0$. 这启发了下面的定义.

定义 2.49　设 λ 是线性算子 $\mathcal{A}: V\to V$ 的特征值. 它的**根子空间**定义为

$$V(\lambda)=\{v\in V \mid \text{对某个正整数 } k \text{ 有 } (\mathcal{A}-\lambda\mathcal{E})^k v=0\}.$$

容易验证 $V(\lambda)$ 的确是子空间. 对 $u,v\in V(\lambda)$, $\alpha,\beta\in K$, 如果 $(\mathcal{A}-\lambda\mathcal{E})^s u=0$, $(\mathcal{A}-\lambda\mathcal{E})^t v=0$, 那么

$$(\mathcal{A}-\lambda\mathcal{E})^{s+t}(\alpha u+\beta v)=\alpha(\mathcal{A}-\lambda\mathcal{E})^{s+t}u+\beta(\mathcal{A}-\lambda\mathcal{E})^{s+t}v=0.$$

所以 $\alpha u+\beta v\in V(\lambda)$. 显然特征空间 $V^\lambda=\{v\in V \mid (\mathcal{A}-\lambda\mathcal{E})v=0\}$ 是 $V(\lambda)$ 的子空间. 如果 $(\mathcal{A}-\lambda\mathcal{E})^k v=0$, 那么 $(\mathcal{A}-\lambda\mathcal{E})^k\mathcal{A}v=\mathcal{A}(\mathcal{A}-\lambda\mathcal{E})^k v=0$. 所以 $V(\lambda)$ 是 $\mathcal{A}$ 的不变子空间.

练习 2.50　证明算子 $\mathcal{A}-\lambda\mathcal{E}$ 在 $V(\lambda)$ 上的限制是幂零的, 且幂零指数不超过 n.

我们将看到根子空间是 V 的不变子空间直和分解的直和项. 先证明如下的一般结论.

定理 2.51　假设多项式 $\theta(t)\in K[t]$ 零化算子 $\mathcal{A}$. 如果 $\theta(t)=\xi(t)\eta(t)$, 其中 $\xi(t)$ 和 $\eta(t)$ 是互素的多项式, 那么

$$U=\{v\in V \mid \xi(\mathcal{A})v=0\},\quad W=\{v\in V \mid \eta(\mathcal{A})v=0\}$$

是 $\mathcal{A}$ 不变子空间且 V 是它们的直和:

$$V=U\oplus W.$$

另外, $\eta(\mathcal{A})$ 在 U 上的限制是可逆算子, $\xi(\mathcal{A})$ 在 W 上的限制是可逆算子.

证明　(1) 因为对任意的多项式 $\sigma(t)$, 有 $\sigma(\mathcal{A})\mathcal{A}=\mathcal{A}\sigma(\mathcal{A})$, 所以 U 和 W 都是 $\mathcal{A}$ 的不变子空间. 由于 $\xi(t)$ 和 $\eta(t)$ 是互素的多项式, 所以存在多项式 $\pi(t)$ 和 $\tau(t)$ 使得

$$\pi(t)\xi(t)+\tau(t)\eta(t)=1. \tag{2.5.23}$$

于是

$$\pi(\mathcal{A})\xi(\mathcal{A})+\tau(\mathcal{A})\eta(\mathcal{A})=\mathcal{E}. \tag{2.5.24}$$

对任意的 $x\in V$, 有

$$x=\mathcal{E}x=\pi(\mathcal{A})\xi(\mathcal{A})x+\tau(\mathcal{A})\eta(\mathcal{A})x. \tag{2.5.25}$$

由于 $\theta(\mathcal{A})=\xi(\mathcal{A})\eta(\mathcal{A})=\mathcal{O}$, 所以 $\pi(\mathcal{A})\xi(\mathcal{A})x\in W$, $\tau(\mathcal{A})\eta(\mathcal{A})x\in U$. 于是 $V=U+W$. 如果 $x\in U\cap W$, 那么 $x=\mathcal{E}x=\pi(\mathcal{A})\xi(\mathcal{A})x+\tau(\mathcal{A})\eta(\mathcal{A})x=0$, 所以 U 和 V 线性无关, 从而 $V=U\oplus W$.

从式 (2.5.24) 可以看出 $V=\xi(\mathcal{A})V+\eta(\mathcal{A})V$. 由于 $\xi(\mathcal{A})V\subset W$, $\eta(\mathcal{A})V\subset U$. 所以有 $\xi(\mathcal{A})V=W$, $\eta(\mathcal{A})V=U$.

(2) 如果 $x\in U$, 从等式 (2.5.25) 得 $\tau(\mathcal{A})\eta(\mathcal{A})x=x$, 所以 $\eta(\mathcal{A})$ 在 U 上的限制是可逆算子, 逆为 $\tau(\mathcal{A})$ 在 U 上的限制. 类似地, $\xi(\mathcal{A})$ 在 W 上的限制是可逆算子, 逆为 $\pi(\mathcal{A})$ 在 W 上的限制. □

定理 2.52 设线性算子 $\mathcal{A}:V\to V$ 的特征多项式在 $K[t]$ 中能分解成线性多项式的乘积

$$\chi_{\mathcal{A}}(t)=\prod_{i=1}^{p}(t-\lambda_i)^{n_i},\quad \lambda_k\neq\lambda_j \text{ 如果 } k\neq j.$$

对 $i=1,2,\cdots,p$, 命

$$U_i=\{v\in V\mid(\mathcal{A}-\lambda_i\mathcal{E})^{n_i}v=0\},\quad U_i'=\left\{v\in V\;\middle|\;\prod_{j\neq i}(\mathcal{A}-\lambda_j\mathcal{E})^{n_j}v=0\right\}.$$

那么诸 U_i 和 U_i' 是 $\mathcal{A}$ 的不变子空间, V 是这些不变子空间的直和: $V=U_1\oplus\cdots\oplus U_p$. 而且 $\dim U_i=n_i$, 算子 $\mathcal{A}-\lambda_i\mathcal{E}$ 在 U_i 上的限制是幂零的, 在 U_i' 上的限制是非退化的 (即可逆的). 另外, λ_i 是算子 $\mathcal{A}|_{U_i}$ 的唯一的特征值.

证明 根据定理 2.51, 诸 U_i 和 U_i' 是 $\mathcal{A}$ 的不变子空间. 对 p 做归纳法证明 V 是诸 U_i 的直和分解. 当 $p=1$ 时, 结论是显然的. 当 $p=2$ 时, 根据上一个定理, 有

$$V=U_i\oplus U_i'. \tag{2.5.26}$$

对 $\mathcal{A}|_{U_i'}$ 和多项式 $\prod\limits_{j\neq i}(t-\lambda_j)^{n_j}$ 应用归纳假设, 知 U_i' 是 $U_1,\cdots,U_{i-1},U_{i+1},\cdots,U_p$ 的直和. 所以 U 是诸 U_i 的直和.

从 U_i 的定义知算子 $\mathcal{A}-\lambda_i\mathcal{E}$ 在其上的限制是幂零. 由上一个定理知 $(\mathcal{A}-\lambda_i\mathcal{E})^{n_i}$ 在 U_i' 上的限制是可逆算子, 所以 $\mathcal{A}-\lambda_i\mathcal{E}$ 在 U_i' 上是非退化的.

根据定理 2.33, $\chi_{\mathcal{A}_{U_i}}(t)=\prod\limits_{k=1}^{p}(t-\lambda_k)^{m_{i,k}}$, $0\leqslant m_{i,k}\leqslant n_k$. 如果对某个 $k\neq i$ 有 $m_{i,k}>0$, 定理 2.31 知, λ_k 是 $\mathcal{A}|_{U_i}$ 的特征值. 从而有非零向量 $v\in U_i$ 使得

$\mathcal{A}v = \lambda_k v$. 但 $(\mathcal{A} - \lambda_i\mathcal{E})^{n_i} v = (\lambda_k - \lambda_i)^{n_i} v \neq 0$, 这与 $v \in U_i$ 矛盾. 所以 $m_{i,k} = 0$ 如果 $k \neq i$. 由定理 2.36 知 $\chi_{\mathcal{A}}(t) = \chi_{\mathcal{A}_{U_1}}(t)\chi_{\mathcal{A}_{U_2}}(t)\cdots\chi_{\mathcal{A}_{U_p}}(t)$. 这迫使

$$\chi_{\mathcal{A}_{U_i}}(t) = (t - \lambda_i)^{n_i}.$$

于是 $\dim U_i = n_i$, 且 λ_i 是算子 $\mathcal{A}|_{U_i}$ 的唯一的特征值. □

推论 2.53 保持定理 2.52 的假设和记号, 那么 $V(\lambda_i) = U_i$. 特别, 此时 V 是所有根子空间的直和, $\dim V(\lambda_i)$ 就是 λ_i 的代数重数 n_i, 而 $(\mathcal{A} - \lambda_i\mathcal{E})^{n_i}$ 在 $V(\lambda_i)$ 上的限制为 0, 在其他根子空间上的限制是非退化的.

证明 显然 $U_i \subset V(\lambda_i)$. 对任意的 $x \in V(\lambda_i)$, 我们证明 $x \in U_i$. 根据分解 (2.5.26), 有

$$x = y + z, \quad y \in U_i, \ z \in U_i'.$$

因为算子 $\mathcal{A} - \lambda_i\mathcal{E}$ 在 U_i' 上的限制是非退化的, 如果 $z \neq 0$, 那么对任意的正整数 s 有 $(\mathcal{A} - \lambda_i\mathcal{E})^s z \neq 0$, 从而 $(\mathcal{A} - \lambda_i\mathcal{E})^s x \neq 0$. 这与 $x \in V(\lambda_i)$ 矛盾. 所以 $z = 0$, $x = y \in U_i$. □

二 约当矩阵 在定理 2.52 的条件下, 寻找算子的最简矩阵的问题就归结到如下情形: 算子 $\mathcal{A}$ 只有一个特征值 λ 且 $\mathcal{A} - \lambda\mathcal{E}$ 是幂零的. 下面是此类算子的一个自然的例子.

例 2.54 设 $W_n(\lambda)$ 是所有形如 $e^{\lambda t}f(t)$ 的复函数形成的复向量空间, 其中 λ 是复数, $f(t)$ 是次数不超过 $n-1$ 的多项式. 对 $e^{\lambda t}f(t) \in W_n(\lambda)$, 有

$$\frac{d}{dt}(e^{\lambda t}f(t)) = \lambda e^{\lambda t}f(t) + e^{\lambda t}f'(t) \in W_n(\lambda).$$

所以, 求导数 $\mathcal{D} = \dfrac{d}{dt}$ 是 $W_n(\lambda)$ 上的一个线性算子. 对 $i = 0, 1, \cdots, n-1$, 命 $e_{i+1} = \dfrac{t^i}{i!}e^{\lambda t}$. 那么 $e_1, \cdots, e_n$ 是 $W_n(\lambda)$ 的基. 显然

$$\mathcal{D}e_{i+1} = \frac{t^{i-1}}{(i-1)!}e^{\lambda t} + \lambda\frac{t^i}{i!}e^{\lambda t} = e_i + \lambda e_{i+1}.$$

(约定 $e_0 = 0$.) 于是, 算子 $\mathcal{D}$ 在基 $e_1, \cdots, e_n$ 下的矩阵是

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}. \tag{2.5.27}$$

不用说, 对任意的域, 这样形状的矩阵都是可以定义的.

定义 2.55 设 λ 是域 K 中的元素, 则式 (2.5.27) 中的矩阵 $J_n(\lambda)$ 称为 (域 K 上的) 以 λ 为特征值的 n 阶**(上) 约当块**.

约当矩阵就是一个分块对角矩阵

$$\begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{m_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J_{m_s}(\lambda_s) \end{pmatrix}. \tag{2.5.28}$$

约当块和约当矩阵的价值在于它的幂和它的多项式都很容易计算. 比如, 当 K 的特征为 0 时, 对任意的多项式 $f(t) \in K[t]$, 有

$$f(J_n(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda)/1! & f''(\lambda)/2! & \cdots & f^{(n-2)}(\lambda)/(n-2)! & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & f'(\lambda)/1! & \cdots & f^{(n-3)}(\lambda)/(n-3)! & f^{(n-2)}(\lambda)/(n-2)! \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & f(\lambda) & f'(\lambda)/1! \\ 0 & 0 & 0 & \cdots & 0 & f(\lambda) \end{pmatrix}.$$

这个等式对特征 p 的域也是成立的, 只是要明确 $f^{(i)}/i!$ 的含义. 而这是很自然可以做到的, 因为 $k(k-1)\cdots(k-i+1)/i!$ 是整数, 在任意域中都是有意义的.

于是, 对算子 $\mathcal{A}$ 能找到基使得其矩阵为约当矩阵对这个算子的计算和性质讨论都是很有价值的.

定理 2.56 设 V 是域 K 上的向量空间, $\mathcal{A}$ 是其上的线性算子. 如果 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)$ 在 $K[t]$ 中能分解成线性多项式的乘积 (对代数闭域, 如复数域, 这个条件总是成立的), 那么 $\mathcal{A}$ 在某个基下的矩阵是形如式 (2.5.28) 的约当矩阵. 这个基称为 $\mathcal{A}$ 的**约当基**, 在约当基下的矩阵称为 $\mathcal{A}$ 的**约当标准形**, 常记作 $J(\mathcal{A})$. 如果不计约当块之间的置换, $\mathcal{A}$ 的约当标准形是唯一的.

约当矩阵 J 如果与方阵 A 相似, 则称 J 是 A 的一个**约当标准形**. 定理 2.56 用矩阵的语言表述就是

定理 2.57 设 A 是域 K 上的 n 阶方阵, 其特征多项式在 $K[t]$ 中能分解成线性多项式的乘积 (对代数闭域, 如复数域, 这个条件总是成立的), 那么 A 有约当标准形, 即存在 K 上的可逆方阵 C 使得 $C^{-1}AC = J(A)$ 是约当矩阵. 如果不计约当块之间的置换, A 的约当标准形是唯一的.

定理中的存在性和唯一性分别在下面三、四部分证明.

三　定理 2.56 中约当基的存在性证明　设 $\lambda_1, \cdots, \lambda_p$ 是 $\chi_{\mathcal{A}}(t)$ 的根, 重数分别是 $n_1, \cdots, n_p$. 根据定理 2.52 和推论 2.53, V 是 $\mathcal{A}$ 的不变子空间 $V(\lambda_1), \cdots, V(\lambda_p)$

的直和. 根据式 (2.3.16), 只要证明 $\mathcal{A}$ 在 $V(\lambda_i)$ 上的限制有约当基. 这等价于寻找幂零算子 $(\mathcal{A}-\lambda_i\mathcal{E})|_{V(\lambda_i)}$ 的约当基因为 $\lambda_i\mathcal{E}$ 在任何基下的矩阵都是纯量矩阵. 于是问题就归结对幂零算子 $\mathcal{B}: V \to V$ 寻找约当基.

对 V 的维数 $n = \dim V$ 用归纳法. 当 $n=1$ 时, 结论显然成立. 假设 $n>1$. 算子 $\mathcal{B}$ 只有一个特征值 0. 取特征向量 $v \in V$. 命 $Y = V/\langle v\rangle$, $\mathcal{C}$ 是 $\mathcal{B}$ 在 Y 上的商算子. 根据定理 2.43, $\mathcal{C}$ 仍是幂零的. 由归纳假设, $\mathcal{C}$ 有约当基. 相应地, Y 能分解成 $\mathcal{C}$ 的不变子空间的直和:

$$Y = Y_1 \oplus \cdots \oplus Y_s,$$

使得每个 Y_i 都有一个基, $\mathcal{C}|_{Y_i}$ 在这个基下的矩阵是一个约当块:

$$\begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & 0 & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & 0 & & 0 & 1 \\ & & & & & 0 \end{pmatrix}.$$

于是存在 $e_i \in V$ 使得 $\mathcal{C}^{m_i-1}\bar{e}_i, \cdots, \mathcal{C}\bar{e}_i, \bar{e}_i$ 是 Y_i 的基, 且 $\mathcal{C}^{m_i}\bar{e}_i = 0$. 其中 $m_i = \dim Y_i$.

向量组 v, $\mathcal{B}^{m_i-1}e_i, \cdots, \mathcal{B}e_i, e_i$, 其中 i 取遍 $1, 2, \cdots, s$, 是 V 的基, 且

$$\mathcal{B}^{m_i}e_i = a_i v, \quad a_i \in K.$$

如果所有的 a_i 都是 0, 那么 V 的这个基就是 $\mathcal{B}$ 的约当基, 向量 v 贡献了一个约当块, 其他的约当块和 $\mathcal{C}$ 的约当块相同. 假设某些 a_i 不为零, 取 r 使得 $a_r \neq 0$ 且 $m_r \geqslant m_i$ 如果 $a_i \neq 0$. 定义

$$e_i' = \begin{cases} e_i, & i = r \text{ 或 } a_i = 0, \\ e_i - a_i a_r^{-1}\mathcal{B}^{m_r-m_i}e_r, & i \neq r \text{ 且 } a_i \neq 0. \end{cases}$$

令 $m_r' = m_r + 1$, $m_i' = m_i$ 如果 $i \neq r$. 那么向量组 $\mathcal{B}^{m_i'-1}e_i', \cdots, \mathcal{B}e_i', e_i'$, 其中 i 取遍 $1, 2, \cdots, s$, 是 V 的基, 而且对所有的 i 有 $\mathcal{B}^{m_i'}e_i' = 0$. 从而在这个基下 $\mathcal{B}$ 的矩阵是约当矩阵. 存在性得证.

四　算子的约当标准形的唯一性　定理 2.56 中算子 $\mathcal{A}$ 的约当标准形的唯一性是指 $\mathcal{A}$ 任意两个约当标准形的约当块除了顺序可能不同外, 其他的如阶、特征根等都是一样的. 由于 $\mathcal{A}$ 在根子空间 $V(\lambda)$ 上有唯一的特征根, 所以 $\mathcal{A}$ 的约当基中, 能贡献以 λ 为特征值的约当块的基向量一定在 $V(\lambda)$ 中. 于是, 唯一性问题归

结到 $\mathcal{A}$ 在根子空间上的限制的约当标准形的唯一性. 如同存在性的证明那样, 只要对幂零算子 $\mathcal{B}: V \to V$ 证明约当标准形的唯一性.

假设在某个基下 $\mathcal{B}$ 的矩阵是约当标准形:

$$J(\mathcal{B}) = \operatorname{diag}(J_{m_1}(0), \cdots, J_{m_s}(0)).$$

相应地, V 分解成 $\mathcal{B}$ 的不变子空间的直和:

$$V = V_1 \oplus \cdots \oplus V_s,$$

其中 V_i 贡献了约当块 $J_{m_i}(0)$.

要证明对任意正整数 k, 约当块 $J_k(0)$ 在 $J(\mathcal{B})$ 中出现的个数 $N(k)$ 不依赖约当基的选取. 显然

$$N(k) = \operatorname{Card}\{V_i \mid m_i = k\}.$$

容易看出 $\mathcal{B}_i = \mathcal{B}|_{V_i}$ 的幂零指数是 m_i, 对小于 m_i 的非负整数 k, $\mathcal{B}_i^k V_i$ 是 $m_i - k$ 维向量空间.

我们有

$$\begin{aligned}
\dim \mathcal{B}^{k-1}V/\mathcal{B}^k V &= \dim \mathcal{B}^{k-1}V - \dim \mathcal{B}^k V \\
&= \sum_{m_i \geqslant k} (m_i - k + 1) - \sum_{m_i \geqslant k+1} (m_i - k) \\
&= \sum_{m_i \geqslant k} 1 \\
&= N(k) + N(k+1) + \cdots + N(m),
\end{aligned}$$

其中 m 是 $J(\mathcal{B})$ 中最大的约当块的阶数. 所以

$$\begin{aligned}
N(k) &= \dim \mathcal{B}^{k-1}V - \dim \mathcal{B}^k V - (\dim \mathcal{B}^k V - \dim \mathcal{B}^{k+1}V) \\
&= \dim \mathcal{B}^{k-1}V - 2\dim \mathcal{B}^k V + \dim \mathcal{B}^{k+1}V.
\end{aligned} \tag{2.5.29}$$

这个数显然与基的选取无关. 唯一性得证, 这也完成了定理 2.56 的证明.

五　求算子和矩阵的约当标准形　如果知道算子或矩阵的特征值, 唯一性的证明实际上也给出了求约当标准形的方法. 设算子 $\mathcal{A}$ 在某个基下的矩阵是 A, λ 是 A 的特征值. 只有根子空间 $V(\lambda)$ 对以 λ 为特征值的约当块有贡献. 设 V' 是其他的根子空间的直和. 根据推论 2.53, 有

$$V = V(\lambda) \oplus V',$$

而且 $\mathcal{B} = \mathcal{A} - \lambda\mathcal{E}$ 在 $V(\lambda)$ 上的限制是幂零的, 在 V' 上的限制是可逆的. 于是

$$\begin{aligned}
\dim \mathcal{B}^{k-1}V - \dim \mathcal{B}^k V &= \dim \mathcal{B}^{k-1}V(\lambda) + \dim \mathcal{B}^{k-1}V' - \dim \mathcal{B}^k V(\lambda) - \dim \mathcal{B}^k V' \\
&= \dim \mathcal{B}^{k-1}V(\lambda) + \dim V' - \dim \mathcal{B}^k V(\lambda) - \dim V' \\
&= \dim \mathcal{B}^{k-1}V(\lambda) - \dim \mathcal{B}^k V(\lambda).
\end{aligned}$$

根据式 (2.5.29), 在 $\mathcal{A}$ 的约当标准形中, $J_k(\lambda)$ 出现的次数是

$$N(\lambda,k)=\dim(\mathcal{A}-\lambda\mathcal{E})^{k-1}V-2\dim(\mathcal{A}-\lambda\mathcal{E})^{k}V+\dim(\mathcal{A}-\lambda\mathcal{E})^{k+1}V. \tag{2.5.30}$$

注意对 V 上的任意的算子 $\mathcal{C}$, 有 $\operatorname{rk}\mathcal{C}=\dim\mathcal{C}V=\operatorname{rk}C$, 其中 C 是 $\mathcal{C}$ 在某个基下的矩阵. 于是式 (2.5.30) 的另一个形式是

$$N(\lambda,k)=\operatorname{rk}(\mathcal{A}-\lambda\mathcal{E})^{k-1}-2\operatorname{rk}(\mathcal{A}-\lambda\mathcal{E})^{k}+\operatorname{rk}(\mathcal{A}-\lambda\mathcal{E})^{k+1}. \tag{2.5.31}$$

即

$$N(\lambda,k)=\operatorname{rk}(A-\lambda E)^{k-1}-2\operatorname{rk}(A-\lambda E)^{k}+\operatorname{rk}(A-\lambda E)^{k+1}. \tag{2.5.32}$$

例 2.58　假设算子 $\mathcal{A}$ 在某个基下的矩阵是

$$A=\begin{pmatrix}6&-9&5&4\\7&-13&8&7\\8&-17&11&8\\1&-2&1&3\end{pmatrix}.$$

求 $\mathcal{A}$ 的约当标准形.

通过计算知 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)=\chi_A(t)=(t-1)(t-2)^3$. 于是根子空间 $V(1)$ 是一维的, 它也是特征子空间, 贡献一个约当块 $J_1(1)$. 要看出 $\mathcal{A}$ 的以 2 为特征值的约当块, 我们计算矩阵 $A-2E$ 的幂的秩. 由于 $(\mathcal{A}-2\mathcal{E})^3$ 在 $V(2)$ 上的限制为 0, 在 $V(1)$ 上的限制是可逆的, 所以 $\operatorname{rk}(A-2E)^k=1$ 如果 $k\geqslant 3$. 直接计算 $A-2E$ 的秩和 $(A-2E)^2$ 的秩, 知它们的秩分别是 2 和 1. 根据公式 (2.5.32) 知 $J_1(2)$ 和 $J_2(2)$ 在 $\mathcal{A}$ 的约当标准形中出现的次数都是 1. 于是 $\mathcal{A}$ 和 A 的约当标准形是

$$J(\mathcal{A})=J(A)=\begin{pmatrix}1&&&\\&2&&\\&&2&1\\&&&2\end{pmatrix}.$$

矩阵中未写出的值均为 0.(求矩阵的秩的有效方法是通过初等变换化简矩阵, 对乘积 AB 的秩, 则可以对 A 和 B 分别做行初等变换和列初等变换化简后再相乘.)

六　求约当基　如果要计算方阵的幂, 仅知道矩阵 A 的约当标准形还不够, 还需要找出矩阵 C 使得 $C^{-1}AC$ 是约当标准形. 用算子的语言, 这个问题就是寻找约当基. 我们对约当标准形的唯一性的证明其实给出了如何寻找约当基的线索.

假设算子 $\mathcal{A}$ 有约当标准形, λ 是它的特征值, 代数重数是 m. 在 $\mathcal{A}$ 的约当标准形中以 λ 为特征值的约当块都由 $V(\lambda)$ 的子空间贡献. 设 V' 是其他的根子空间,

则有 $V = V(\lambda) \oplus V'$. 命

$$V(\lambda) = W_1 \oplus \cdots \oplus W_s,$$

其中每个 W_i 都是 $\mathcal{A}$ 的不变子空间, 贡献 $J(\mathcal{A})$ 的一个约当块. 设 W_i 的维数是 m_i. 于是有向量 e_i 使得 e_i, $(\mathcal{A}-\lambda\mathcal{E})e_i$, $\cdots$, $(\mathcal{A}-\lambda\mathcal{E})^{m_i-1}e_i$ 是 W_i 的基, 且 $(\mathcal{A}-\lambda\mathcal{E})^{m_i}e_i = 0$. 显然 $\ker(\mathcal{A}-\lambda\mathcal{E}) \cap W_i$ 是一维的, 由 $(\mathcal{A}-\lambda\mathcal{E})^{m_i-1}e_i$ 张成, 而 $\mathcal{A}-\lambda\mathcal{E}$ 在 V' 上的限制是可逆的, 所以我们有

命题 2.59 假设算子 $\mathcal{A}$ 有约当标准形, λ 是它的特征值. 那么 $J(\mathcal{A})$ 中以 λ 为特征值的约当块的个数是

$$N(\lambda) = \dim\ker(\mathcal{A}-\lambda\mathcal{E}).$$

如果 A 是 $\mathcal{A}$ 的一个矩阵, n 是 A 的阶数, 那么

$$N(\lambda) = n - \mathrm{rk}\,(A-\lambda E).$$

证明 第一个等式已经证明了. 利用定理 2.11(2) 和定理 2.13(1), 知第二个等式与第一个等式是等价的. □

观察 W_i 的基, 便知道可以通过 $\ker(\mathcal{A}-\lambda\mathcal{E})$ 求得 $\mathcal{A}|_{V(\lambda)}$ 的约当基. 设 $\mathcal{B} = (\mathcal{A}-\lambda\mathcal{E})|_{V(\lambda)}$ 的幂零指数是 p. 对正整数 i, 命

$$V(\lambda)_i = \mathrm{Im}\,\mathcal{B}^{i-1} \cap \ker\mathcal{B} = (\mathcal{A}-\lambda\mathcal{E})^{i-1}V(\lambda) \cap \ker(\mathcal{A}-\lambda\mathcal{E}).$$

那么

$$\ker(\mathcal{A}-\lambda\mathcal{E}) = V(\lambda)_1 \supset V(\lambda)_2 \supset \cdots \supset V(\lambda)_p \neq 0, \quad V(\lambda)_{p+1} = 0.$$

在 $V(\lambda)$ 中选取线性无关的向量 e_1, e_2, $\cdots$, e_{s_p} 使得 $\mathcal{B}^{p-1}e_1$, $\cdots$, $\mathcal{B}^{p-1}e_{s_p}$ 是 $V(\lambda)_p$ 的基. 接下来在 $V(\lambda)$ 中取线性无关的向量 e_{s_p+1}, e_{s_p+2}, $\cdots$, $e_{s_{p-1}}$ 使得向量组

$$\mathcal{B}^{p-1}e_1, \cdots, \mathcal{B}^{p-1}e_{s_p}, \mathcal{B}^{p-2}e_{s_p+1}, \cdots, \mathcal{B}^{p-2}e_{s_{p-1}}$$

是 $V(\lambda)_{p-1}$ 的基. 如此下去, 最后得到向量组

$$e_1, \cdots, e_{s_p}, e_{s_p+1}, \cdots, e_{s_{p-1}}, \cdots, e_{s_2+1}, \cdots, e_{s_1},$$

使得对于不超过 p 的正整数 k, 向量组

$$\mathcal{B}^{p-1}e_1, \cdots, \mathcal{B}^{p-1}e_{s_p}, \cdots, \mathcal{B}^{k-1}e_{s_{k+1}+1}, \cdots, \mathcal{B}^{k-1}e_{s_k}$$

是 $V(\lambda)_k$ 的基. (注意: 任何算子的零次幂都约定为恒等算子 $\mathcal{E}$.)

把算子 $\mathcal{B}=\mathcal{A}-\lambda\mathcal{E}$ 的幂作用在向量组 (e_i) 上, 得到如下的向量组:

$$\begin{array}{lllll}
e_1 & \mathcal{B}e_1 & \cdots & \mathcal{B}^{p-2}e_1 & \mathcal{B}^{p-1}e_1 \\
\vdots & \vdots & & \vdots & \vdots \\
e_{s_p} & \mathcal{B}e_{s_p} & \cdots & \mathcal{B}^{p-2}e_{s_p} & \mathcal{B}^{p-1}e_{s_p} \\
e_{s_p+1} & \mathcal{B}e_{s_p+1} & \cdots & \mathcal{B}^{p-2}e_{s_p+1} & \\
\vdots & \vdots & & \vdots & \\
e_{s_{p-1}} & \mathcal{B}e_{s_{p-1}} & \cdots & \mathcal{B}^{p-2}e_{s_{p-1}} & \\
\vdots & \vdots & & & \\
e_{s_3+1} & \mathcal{B}e_{s_3+1} & & & \\
\vdots & \vdots & & & \\
e_{s_2} & \mathcal{B}e_{s_2} & & & \\
e_{s_2+1} & & & & \\
\vdots & & & & \\
e_{s_1} & & & &
\end{array} \tag{2.5.33}$$

这些向量是线性无关的. 事实上, 设

$$\sum_{i=1}^{s_1}\alpha_{i,1}e_i+\sum_{i=1}^{s_2}\alpha_{i,2}\mathcal{B}e_i+\cdots+\sum_{i=1}^{s_p}\alpha_{i,p}\mathcal{B}^{p-1}e_i=0. \tag{2.5.34}$$

把算子 $\mathcal{B}^{p-1}$ 作用到上式, 得

$$\sum_{i=1}^{s_p}\alpha_{i,1}\mathcal{B}^{p-1}e_i=0,$$

从而

$$\alpha_{i,1}=0,\quad 1\leqslant i\leqslant s_p.$$

把算子 $\mathcal{B}^{p-2}$ 作用到式 (2.5.34), 得

$$\sum_{i=s_{p+1}}^{s_{p-1}}\alpha_{i,1}\mathcal{B}^{p-2}e_i+\sum_{i=1}^{s_p}\alpha_{i,2}\mathcal{B}^{p-1}e_i=0,$$

根据向量 e_i 的选择, 得

$$\alpha_{i,1}=\alpha_{j,2}=0,\qquad s_p+1\leqslant i\leqslant s_{p-1},\ \ 1\leqslant j\leqslant s_p.$$

如此下去, 可知式 (2.5.34) 中的系数全为 0. 这个向量组所含的向量的个数是

$$
\begin{aligned}
& ps_p+(p-1)(s_{p-1}-s_p)+\cdots+2(s_2-s_3)+(s_1-s_2)\\
=& s_1+s_2+\cdots+s_p\\
=& \dim V(\lambda)_1+\dim V(\lambda)_2+\cdots+\dim V(\lambda)_p\\
=& \sum_{i=1}^{p}\dim(\operatorname{Im}\mathcal{B}^{i-1}\cap\ker\mathcal{B})=\sum_{i=1}^{p}(\dim\ker\mathcal{B}^i-\dim\ker\mathcal{B}^{i-1})\\
=& \dim\ker\mathcal{B}^p=\dim V(\lambda).
\end{aligned}
$$

上面倒数第二行的等式来自满射 $\ker\mathcal{B}^i\to\operatorname{Im}\mathcal{B}^{i-1}\cap\ker\mathcal{B}$, $x\to\mathcal{B}^{i-1}x$ 的核是 $\ker\mathcal{B}^{i-1}$. 于是向量组 (2.5.33) 是 $V(\lambda)$ 的基, 其构造方式表明这是 $\mathcal{B}$ 的约当基, (2.5.33) 中每一行向量给出 $\mathcal{B}$(等价地, $\mathcal{A}$) 的一个约当块.

例 2.60 假设 $\mathcal{A}$ 在基 v_1,v_2,v_3,v_4 下的矩阵是例 2.58 中的矩阵 A. 我们用刚才的方法求出 $\mathcal{A}$ 的一个约当基. 利用同构 $V\to K^4$, 不妨假设 $V=K^4$ 是列向量空间, 所给的基是标准基. 于是 $\mathcal{A}$ 就是 K^4 上的算子 $X\to AX$. 这个算子有两个特征值: 1 和 2, 代数重数分别是 1 和 3. 所以 $V(1)=\ker(\mathcal{A}-\mathcal{E})$ 是一维的.

解方程组 $(A-E)X=0$, 得到特征值 1 的特征向量 $e_1=[3,6,7,1]$. 由于 $(\mathcal{A}-2\mathcal{E})^3$ 在 $V(2)$ 上的限制是零算子 (推论 2.53), 所以对大于 2 的整数 k, $(\mathcal{A}-2\mathcal{E})^kV$ 由 e_1 张成. 矩阵 $(A-2E)$ 的秩是 2, 它的第一列和第四列相同, 后三列之和是 0, 所以方程组 $(A-2E)X=0$ 的解空间 $\ker(\mathcal{A}-2\mathcal{E})$ 的一个基是 $[1,0,0,-1]$, $[0,1,1,1]$. 于是 $\mathcal{A}$ 以 2 为特征值的约当块有两个, 其中有一个必是 2 阶的. 算子 $(\mathcal{A}-2\mathcal{E})$ 的像空间 $\operatorname{Im}(\mathcal{A}-2\mathcal{E})$ 由

$$
A-2E=\begin{pmatrix}4&-9&5&4\\7&-15&8&7\\8&-17&9&8\\1&-2&1&1\end{pmatrix}
$$

的列向量张成, 第 3,4 列形成一个基. 所以 $[1,1,1,0]=[5,8,9,1]-[4,7,8,1]=(A-2E)[0,0,1,-1]$ 在 $\operatorname{Im}(\mathcal{A}-2\mathcal{E})$ 中. 注意这个向量也在 $\ker(\mathcal{A}-2\mathcal{E})$ 中. 根据前面所讨论的, 向量组

$$
e_1=[3,6,7,1],\quad e_2=[1,0,0,-1],\quad (A-2E)e_3=[1,1,1,0],\quad e_3=[0,0,1,-1]
$$

是 $\mathcal{A}$ 的约当基. 约当基中的向量就是从标准基 v_1,v_2,v_3,v_4 到这个约当基的转换矩

阵 C 的列向量, 所以

$$C^{-1}AC=\begin{pmatrix}3&1&1&0\\6&0&1&0\\7&0&1&1\\1&-1&0&-1\end{pmatrix}^{-1}\cdot\begin{pmatrix}6&-9&5&4\\7&-13&8&7\\8&-17&11&8\\1&-2&1&3\end{pmatrix}\cdot\begin{pmatrix}3&1&1&0\\6&0&1&0\\7&0&1&1\\1&-1&0&-1\end{pmatrix}$$

$$=\begin{pmatrix}1&&&\\&2&&\\&&2&1\\&&&2\end{pmatrix}=J(A)=J(\mathcal{A}).$$

理论上寻找约当基可以通过解矩阵方程 $XJ(A)=AX$ 实现, 此处 $X=(x_{ij})$ 是由未知元 x_{ij} 构成的 n 阶方阵. 但这是一个有 n^2 个未知元的线性方程组, 求解这个方程组并在解空间中找到非退化的矩阵是不太可行的事情.

例 2.61 矩阵

$$S=\sum_{i,j}E_{ij}=\begin{pmatrix}1&\cdots&1\\\vdots&&\vdots\\1&\cdots&1\end{pmatrix}$$

满足等式 $S^2=nS$, 所以它的极小多项式是 t^2-nt 的因子, 但 t 和 $t-n$ 都不零化 S, 故 S 的极小多项式为 $\mu_S(t)=t^2-nt$. 另外, S 的特征多项式是 $\chi_S(t)=t^{n-1}(t-n)$. 所以 S 的约当标准形中有一个约当块 $J_1(n)$, 其他的约当块的特征值都是 0. 由于 $\mathrm{rk}\,S=1$ 是 S 的约当标准形中所有约当块的秩的和, 所以 $J(S)=\mathrm{diag}\,(n,0,\cdots,0)$. 显然 $S[1,\cdots,1]=n[1,\cdots,1]$, $SX=0$ 的解空间是 $\left\{[a_1,\cdots,a_n]\,\middle|\,\sum_i a_i=0\right\}$, 它的一个基是 $[-1,1,0,\cdots,0]$, $\cdots$, $[0,\cdots,0,-1,1]$, 所以有

$$\begin{pmatrix}1&-1&0&\cdots&0&0\\1&1&-1&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\1&0&0&\cdots&1&-1\\1&0&0&\cdots&0&1\end{pmatrix}^{-1}\cdot S\cdot\begin{pmatrix}1&-1&0&\cdots&0&0\\1&1&-1&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\1&0&0&\cdots&1&-1\\1&0&0&\cdots&0&1\end{pmatrix}$$

$$=\begin{pmatrix}n&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}=J(S).$$

习 题 2.5

1. 记号同定理 2.51. 如果 $\theta=\chi_{\mathcal{A}}(t)$ 是特征多项式, $\xi(t)$ 和 $\eta(t)$ 的首项系数都是 1, 那么

$$\chi_{\mathcal{A}_U}(t)=\xi(t),\quad \chi_{\mathcal{A}_W}(t)=\eta(t);$$
$$\dim U=\deg\xi,\quad \dim W=\deg\eta(t).$$

2. 证明如果 $\mathcal{A}$ 有 k 维不变子空间, 那么它有 $n-k$ 维不变子空间.

3. 非零的 4 阶幂零方阵的约当标准形只有下面四个:

$$A_1=J_2(0)\dotplus J_1(0)\dotplus J_1(0),\qquad A_2=J_2(0)\dotplus J_2(0),$$
$$A_3=J_3(0)\dotplus J_1(0),\qquad A_4=J_4(0).$$

下列矩阵都是幂零的, 它们各相似于哪个 A_i:

$$\begin{pmatrix}0&0&0&1\\0&0&0&1\\0&0&0&1\\0&0&0&0\end{pmatrix},\quad \begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&1\\0&2&0&0\end{pmatrix},\quad \begin{pmatrix}0&0&0&0\\1&0&0&0\\2&1&0&0\\3&2&1&0\end{pmatrix},\quad \begin{pmatrix}0&0&0&1\\0&0&0&0\\0&0&0&-1\\0&7&0&0\end{pmatrix}.$$

4. (1) 如果矩阵的特征多项式为 $(t+3)^2(t-2)^3$, 它的约当标准形有哪些可能?

(2) 如果矩阵的特征多项式为 $(t+3)^2(t-2)^3$, 以 -3 为特征值的特征空间的维数是 1, 以 2 为特征值的特征空间的维数是 2, 这个矩阵的约当标准形有哪些可能?

5. 求出下列矩阵的约当标准形:

$$(1)\begin{pmatrix}1&-3&4\\4&-7&8\\6&-7&7\end{pmatrix},\quad (2)\begin{pmatrix}1&-3&3\\-2&-6&13\\-1&-4&8\end{pmatrix},\quad (3)\begin{pmatrix}-5&-3&-2&4\\2&0&1&-1\\10&7&-4&9\\2&0&1&0\end{pmatrix}.$$

6. (1) 验证矩阵

$$A=\begin{pmatrix}6&2&-2\\-2&2&2\\2&2&2\end{pmatrix},\qquad B=\begin{pmatrix}6&2&2\\-2&2&0\\0&0&2\end{pmatrix}$$

有相同的特征多项式;

(2) 求出它们的极小多项式;

(3) 求出它们的约当标准形.

7. 设域 K 的特征为 0, A 是 K 上的 n 阶方阵. 证明: A 是幂零的当且仅当 $\operatorname{tr}(A^k)=0,\ 1\leqslant k\leqslant n$.

8. 如果线性算子 $\mathcal{A}$ 的约当标准形只有一个约当块, 确定 $\mathcal{A}$ 的所有不变子空间.

9. 证明: 方阵 $A\in M_n(\mathbb{C})$ 和它的转置 ${}^tA$ 相似.

10. 设 $A\in M_n(\mathbb{C})$ 的特征根都等于 1. 证明: 对任意非零整数 k, A 和 A^k 相似.

11. 证明: 对于矩阵 $A \in M_n(\mathbb{C})$, 等式 $A^m = E$ 成立当且仅当 A 可对角化且它的特征值都是 m 次单位根. 举例说明这个结论对 $\mathbb{Z}_p$ 上的方阵不成立, 即存在 $\mathbb{Z}_p$ 上的方阵, 其某个幂为单位矩阵, 但它不能对角化 (即不相似于对角矩阵).

12. 把矩阵 $A = J_1(\lambda) + J_2(\mu)$, $\lambda \neq \mu$, 写成 $A = S + N$ 的形式, 其中 $S = \mathrm{diag}\,(\lambda, \mu, \mu)$, 也就是说

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{pmatrix}, \quad S = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

把 S 和 N 分别写成 A 的多项式 $s(A)$ 和 $m(A)$. (根据凯莱–哈密顿定理, 可以要求 $s(t)$ 和 $m(t)$ 的次数都不超过 2, 按惯例, $A^0 = E$.)

13. 设 $\mathcal{A}$ 是 n 维复向量空间 V 上的线性算子. 证明: 存在唯一的可对角化算子 $\mathcal{S}$ 和幂零算子 $\mathcal{N}$ 使得

$$\mathcal{A} = \mathcal{S} + \mathcal{N}, \quad \mathcal{S}\mathcal{N} = \mathcal{N}\mathcal{S}.$$

而且, $\mathcal{S}$ 和 $\mathcal{N}$ 都可以表达成 $\mathcal{A}$ 的多项式. (可对角化算子也称为**半单算子**.)

14. 对任意正整数 k, 计算 $(J_n(\lambda))^k$.

15. 如果线性算子 $\mathcal{A}$ 在基 $e_1, \cdots, e_n$ 下的矩阵是

$$(1) \begin{pmatrix} 3 & 2 & 3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}, \qquad (2) \begin{pmatrix} 0 & 1 & -1 & 1 \\ -1 & 2 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}.$$

求出 $\mathcal{A}$ 的一个约当基和 $\mathcal{A}$ 在约当基下的矩阵.

16. 证明: 如果复向量空间 V 上的线性算子 $\mathcal{A}$ 和 $\mathcal{B}$ 满足关系 $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \mathcal{B}$, 那么 $\mathcal{B}$ 是幂零的.

17. 算子 $\mathcal{A} = \dfrac{\partial}{\partial x} + \dfrac{\partial}{\partial y}$ 作用在次数不超过 n 的二元复多项式形成的空间上. 找出 $\mathcal{A}$ 的约当标准形.

18. 解矩阵方程:

$$(1)\ X^2 = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}, \qquad (2)\ X^2 = \begin{pmatrix} 6 & 2 \\ 3 & 7 \end{pmatrix}.$$

19. 命 $\mathscr{N}$ 为 $M_n(\mathbb{C})$ 中的幂零矩阵全体, $\mathscr{U}$ 为 $M_n(\mathbb{C})$ 中特征值都等于 1 的矩阵全体. 证明映射

$$\exp : \mathscr{N} \to \mathscr{U}, \quad A \to \exp A = E + A + \frac{A^2}{2!} + \cdots + \frac{A^{n-1}}{(n-1)!}$$

是双射并求出这个双射的逆映射.

第3章 内积空间

向量空间有加法和纯量乘, 已经可以在其上做很多的事情了, 但基本上都是代数的事情, 能涉及的几何概念仅是线、面等. 要想在向量空间中有更丰富的几何结构, 需要长度和角度的概念, 它们是我们现实生活的空间中描绘各种形状的基础. 在实向量空间上, 可以通过正定的双线性型定义长度和角度, 在复向量空间上可以通过正定的半双线性型定义长度和角度. 相应的空间称为内积空间, 也称为纯量乘积空间.

3.1 欧几里得向量空间

一 在平面 $\mathbb{R}^2$ 上, 两点之间的距离通过勾股定理计算:

$$d(A,B)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}.$$

类似地, 在三维空间 $\mathbb{R}^3$ 中, 两点间的距离公式是

$$d(A,B)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}.$$

平移不改变两点间的距离. 如果把 $\mathbb{R}^2$ 中的向量看作是始于原点的有向线段, 那么向量 $x=(x_1,x_2)\in\mathbb{R}^2$ 的长度为 $\|x\|=\sqrt{x_1^2+x_2^2}$.

向量 $x=(x_1,x_2,x_3)\in\mathbb{R}^3$ 的长度为 $\|x\|=\sqrt{x_1^2+x_2^2+x_3^2}$. 长度的平方 $\|x\|^2=x_1^2+x_2^2+x_3^2$ 是一个正定二次型. 相应的极化 $(x|y)=x_1y_1+x_2y_2+x_3y_3$ 就是解析几何中的点积, 它可以用于确定两个向量间的夹角. 这些是我们熟悉的. 设 $\varphi=\widehat{xy}$ 是向量 $x,y\in\mathbb{R}^3$ 的夹角, $z=x-y$ 是两个向量的差. 把平面几何的余弦公式应用到以 x,y,z 为边的三角形上, 得

$$\|z\|^2=\|x\|^2+\|y\|^2-2\|x\|\|y\|\cos\varphi.$$

另一方面, 可用双线性型 $(\cdot|\cdot)$ 计算 z 的长度的平方, 得 $\|z\|^2=(z|z)=(x-y|x-y)=\|x\|^2+\|y\|^2-2(x|y)$. 比较 $\|z\|^2$ 的两个表达式, 得

$$(x|y)=\|x\|\|y\|\cos\varphi.$$

由此可见, 某些对称双线性型不仅可以定义向量的长度, 还能定义向量间的夹角. 这启示在一般的坐标空间 $\mathbb{R}^n$ 中, 应该考虑双线性型

$$(x|y) = \sum_{i=1}^{n} x_i y_i. \tag{3.1.1}$$

但是, 我们的讨论不能仅限于坐标空间, 因为其他的实向量空间也需要长度和夹角的概念.

定义 3.1　带有给定的正定双线性型的实向量空间称为**欧几里得向量空间**, 简称为**欧氏空间**. 这个双线性型称为**内积**(inner product) 或**纯量乘积**(scalar product), 常记作 $(\cdot|\cdot)$.

例 3.2　带着内积 (3.1.1), 坐标空间 $\mathbb{R}^n$ 成为欧氏空间, 称为**标准欧氏空间**, 其内积称为**标准内积**.

例 3.3　设 $V = P_n$ 是次数不超过 $n-1$ 的实多项式全体. 对任意两个向量 $f, g \in V$, 定义

$$(f|g) = \int_a^b f(t)g(t)dt \tag{3.1.2}$$

($b > a$ 是实数). 这是一个正定的双线性型, 从而 V 成为欧氏空间. 值得注意的是这个双线性型的定义没有用到空间的基.

同样的公式可对闭区间 $[a,b]$ 上任意的连续函数 f, g 定义, 从而 $[a,b]$ 上的连续函数空间 $C(a,b)$ 是无限维欧氏空间.

二　基本的度量概念　欧氏空间 V 中的向量 v 的**长度**(或**范数**(norm)) 定义为

$$\|v\| = \sqrt{(v|v)}. \tag{3.1.3}$$

由于内积是正定双线性型, 所以非零向量的长度是正数, 且对任意实数 a 有 $\|av\| = \sqrt{(av|av)} = |a| \cdot \|v\|$.

一维的坐标空间就是实数域本身, 这时向量的长度就是实数的绝对值. 由于正定双线性型在子空间上的限制仍是正定双线性型, 所以欧氏空间的子空间也是欧氏空间.

长度为 1 的向量称为**单位向量**(unit vector). 任意非零向量 v 都和一个单位向量成比例 (即线性相关), 因为 $v' = v/\|v\|$ 是单位向量:

$$\|v'\| = \left\| \frac{1}{\|v\|} v \right\| = \frac{1}{\|v\|}\|v\| = 1.$$

在 $\mathbb{R}^2$ 中, 单位向量全体是一个圆周, 半径是 1, 圆心是原点. 在 $\mathbb{R}^3$ 中, 单位向量全体是一个球面, 半径是 1, 球心是原点. 一般地, 在欧氏空间 V 中, 所有单位向量构成的集合称为**单位球面**, 所有长度不超过 1 的向量构成的集合称为**单位球**.

为定义向量间的夹角, 我们需要证明如下结论.

定理 3.4(柯西–施瓦茨不等式)① 对欧氏空间中任意两个向量 u 和 v, 有

$$|(u|v)| \leqslant \|u\|\,\|v\|. \tag{3.1.4}$$

取等式当且仅当 u 和 v 成比例.

证明 如果 $u = av$, 那么 $|(u|v)| = |(av|v)| = |a||(v|v)| = |a|\|v\|^2 = \|u\|\|v\|$.

如果 u 和 v 不成比例, 那么它们是空间 $U = \langle u, v\rangle$ 的基. 内积在 U 上的限制仍是正定的, 从而它在这个基下的矩阵

$$\begin{pmatrix} (u|u) & (u|v) \\ (u|v) & (v|v) \end{pmatrix}$$

是正定的. 特别, 矩阵的行列式是正的, 从而有 $(u|v)^2 < (u|u)\cdot(v|v) = \|u\|^2\,\|v\|^2$. 取平方根, 即得 $|(u|v)| < \|u\|\,\|v\|$. □

例 3.5 在欧氏空间 $\mathbb{R}^n$ 中, 柯西–施瓦茨不等式成为

$$\left|\sum_{i=1}^n x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2}. \tag{3.1.5}$$

在连续函数空间 $C(a,b)$ 中, 柯西–施瓦茨不等式变成

$$\left|\int_a^b f(t)g(t)dt\right| \leqslant \sqrt{\int_a^b f(t)^2 dt}\sqrt{\int_a^b g(t)^2 dt}. \tag{3.1.6}$$

如果 u, v 是欧氏空间中的非零向量, 柯西–施瓦茨不等式意味着

$$-1 \leqslant \frac{(u|v)}{\|u\|\,\|v\|} \leqslant 1.$$

由此可见, 比值 $(u|v)/(\|u\|\,\|v\|)$ 是余弦函数的一个值.

定义 3.6 如果 u, v 是欧氏空间中的非零向量, 那么它们的**夹角** φ 由下式定义

$$\cos\varphi = \frac{(u|v)}{\|u\|\,\|v\|}, \qquad 0 \leqslant \varphi \leqslant \pi. \tag{3.1.7}$$

如果 u, v 中有一个是零向量, 那么它们没有夹角. 称向量 u 和 v 是**正交的**(orthogonal), 记作 $u \perp v$, 如果它们的内积 $(u|v)$ 为 0. 如果 u 和 v 都是非零向量, 正交等价于它们的夹角是 $\pi/2$.

① 也称为施瓦茨不等式, 或柯西–布尼亚科夫斯基不等式.

注意零向量正交于任何向量. 两个非零向量的夹角是 0 或 π 当且仅当它们成比例.

内积对表达和证明初等几何的结论是很方便的. 例如勾股定理说的是

$$u \perp v \implies \|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

推广也是直接的, 对互相正交的向量 $u, v, w, z, \cdots$, 有

$$\|u+v+w+z+\cdots\|^2 = \|u\|^2 + \|v\|^2 + \|w\|^2 + \|z\|^2 + \cdots.$$

而菱形的对角线互相垂直可以表达成

$$\|u\| = \|v\| \implies (u+v) \perp (u-v).$$

平行四边形等式就是

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

三角形三条边的长度之间有一个不等式:

命题 3.7(三角不等式) 对欧氏空间中的向量 u 和 v, 有

$$\|u \pm v\| \leqslant \|u\| + \|v\|. \tag{3.1.8}$$

证明 用内积的双线性和柯西–施瓦茨不等式, 得

$$\begin{aligned}\|u \pm v\|^2 &= \|u\|^2 + \|v\|^2 \pm 2(u|v) \leqslant \|u\|^2 + \|v\|^2 + 2|(u|v)| \\ &\leqslant \|u\|^2 + \|v\|^2 + 2\|u\|\,\|v\| = (\|u\| + \|v\|)^2.\end{aligned}$$

□

例 3.8 在欧氏空间 $\mathbb{R}^n$ 中, 三角不等式成为

$$\sqrt{\sum_{i=1}^n (x_i \pm y_i)^2} \leqslant \sqrt{\sum_{i=1}^n x_i^2} + \sqrt{\sum_{i=1}^n y_i^2}. \tag{3.1.9}$$

在连续函数空间 $C(a, b)$ 中, 三角不等式变成

$$\sqrt{\int_a^b (f(t) \pm g(t))^2 dt} \leqslant \sqrt{\int_a^b f(t)^2 dt} + \sqrt{\int_a^b g(t)^2 dt} \tag{3.1.10}$$

(闵可夫斯基不等式).

三 标准正交基 在欧氏空间 $\mathbb{R}^n$ 中, 标准基 $(1, 0, \cdots, 0)$, $(0, 1, \cdots, 0)$, $\cdots$, $(0, 0, \cdots, 1)$ 中的向量互相正交, 且都是单位向量. 这给各种计算带来很大的方便. 自然, 也希望在其他的欧氏空间中能有这样的基.

定义 3.9 欧氏空间 V 的基 $e_1,\cdots,e_n$ 称为**正交的**如果基中的向量互相正交. 进一步, 如果基中的向量都是单位向量, 这个基则称为**标准正交基**.

从正交基 (e_i) 很容易得到标准正交基: 只要把每个基向量 e_i 换成与它成比例的单位向量 $e_i/\|e_i\|$. 向量的正交性有很多好的性质.

定理 3.10 如果欧氏空间 V 中的非零向量 $v_1,\cdots,v_m$ 互相正交, 那么这些向量线性无关. 如果 $m=\dim V$, 那么这些向量构成 V 的一个正交基.

证明 设 $a_1v_1+a_2v_2+\cdots+a_mv_m=0$. 要证明所有的系数 a_i 是 0. 用 v_i 在等式两边做内积, 得

$$\begin{aligned}0=(0|v_i)&=(a_1v_1+\cdots+a_mv_m|v_i)\\&=a_1(v_1|v_i)+\cdots+a_i(v_i|v_i)+\cdots+a_m(v_m|v_i)=a_i(v_i|v_i).\end{aligned}$$

由于 v_i 是非零向量, 所以 $(v_i|v_i)>0$. 于是 $a_i=0$. 第一个结论得证. 再利用定理 1.30(2) 知第二个结论成立. □

命题 3.11 如果 $u_1,\cdots,u_k$ 正交于 v, 那么任何线性组合 $a_1u_1+\cdots+a_ku_k$ 和 v 正交.

证明 仅需注意到 $(a_1u_1+\cdots+a_ku_k|v)=a_1(u_1|v)+\cdots+a_k(u_k|v)=0$. □

这个命题表明如果 v 与 $u_1,\cdots,u_k$ 正交, 那么 v 与子空间 $U=\langle u_1,\cdots,u_k\rangle$ 中任何向量正交. 此时, 称向量 v 与子空间 U**正交**. 一般说来, 对欧氏空间 V 的子集 Y, 称向量 v**正交于**Y, 记作 $v\perp Y$, 如果 v 正交于 Y 中每一个向量. 根据命题 3.11, 所有正交于 Y 的向量形成 V 的一个子空间. Y 是子空间的情形值得特别关注. 我们引入一个概念.

定义 3.12 设 U 是欧氏空间 V 的线性子空间. 所有正交于 U 的向量形成的子空间称为 U 的**正交补**, 记作 $U^\perp$.

如果 $e_1,\cdots,e_n$ 是欧氏空间的标准正交基, 对向量 $x=x_1e_1+\cdots+x_ne_n$, 有 $\|x\|^2=x_1^2+\cdots+x_n^2$. 这其实就是二次型的标准形式. 于是, 可以考虑利用二次型的标准形建立标准正交基的存在性.

定理 3.13 有限维欧氏空间中有标准正交基.

证明 设 V 是欧氏空间, 维数是 n. 由于 V 上的二次型

$$q(x)=(x|x)=\|x\|^2$$

是正定的, 故存在 V 的基 $e_1,\cdots,e_n$ 使得 q 在这个基下有标准的形式

$$q(x)=x_1^2+x_2^2+\cdots+x_n^2,$$

$(x=x_1e_1+\cdots+x_ne_n)$. 于是 e_i 和 e_j 的内积是

$$(e_i|e_j)=\frac{1}{2}[q(e_i+e_j)-q(e_i)-q(e_j)]=\delta_{ij}.$$

所以 $e_1,\cdots,e_n$ 是标准正交基. □

在标准正交基 $e_1,\cdots,e_n$ 下, 两个向量的内积就是坐标分量分别相乘, 然后加起来: $(x|y)=x_1y_1+\cdots+x_ny_n$. 这与公式 (3.1.1) 没有差别. 向量 x 在这个基下的坐标就是基向量与 x 的内积:

$$(x|e_i)=x_i. \tag{3.1.11}$$

一般地, 对单位向量 e, 称向量 $(x|e)e$ 为 x 在直线 $\langle e\rangle=\mathbb{R}e$ 上的**投影**.

我们看到了标准正交基的价值, 建立了其存在性, 那怎么找到它们? 下面的结论给出这个问题的答案, 证明中的算法称为**格拉姆–施密特正交化方法**.

定理 3.14(正交化方法)　设 $v_1,\cdots,v_m$ 是欧氏空间 V 中线性无关的向量, 那么存在相互正交的单位向量 $e_1,\cdots,e_m$, 使得对每个 k, 向量 e_k 是 $v_1,\cdots,v_k$ 的线性组合.

证明　取 $e_1=v_1/\|v_1\|$. 可见结论在 $m=1$ 的情况成立. 假设 $m>1$, 且已经构造了 $e_1,\cdots,e_i$, 我们看如何找出 e_{i+1}. 注意 v_{i+1} 不是 $e_1,\cdots,e_i$ 的线性组合. 考虑线性组合

$$u=v_{i+1}-\sum_{k=1}^{i}a_ke_k,$$

其中 $a_1,\cdots,a_i$ 是待定的系数. 那么 $u\neq 0$. 要恰当选择 a_i 使得 u 与向量 $e_1,\cdots,e_i$ 都正交. 这等价于如下条件:

$$\begin{aligned}0=(u|e_j)&=(v_{i+1}|e_j)-\left(\sum_{k=1}^{i}a_ke_k|e_j\right)\\&=(v_{i+1}|e_j)-\sum_{k=1}^{i}a_k(e_k|e_j)=(v_{i+1}|e_j)-a_j,\quad j=1,2,\cdots,i.\end{aligned}$$

由此可见, 取 $a_j=(v_{k+1}|e_j)$, 则 u 与向量 $e_1,\cdots,e_i$ 都正交. 命 $e_{i+1}=u/\|u\|$, 就得到了相互正交的单位向量 $e_1,\cdots,e_i,e_{i+1}$. 如此下去, 最后, 就得到了所要求的向量组 $e_1,\cdots,e_m$. □

一个向量组称为**标准正交向量组**如果向量组中的向量都是单位向量且互相正交.

推论 3.15　欧氏空间 V 中任何标准正交向量组都可以扩充为 V 的标准正交基.

证明　设向量组 $v_1,\cdots,v_m$ 是标准正交的. 根据定理 3.10, 这些向量线性无关, 所以可以扩充为 V 的一个基向量组 $v_1,\cdots,v_m,v_{m+1},\cdots,v_n$. 运用正交化方法, 可以得到 V 的标准正交基 $e_1,\cdots,e_n$. 因为 $v_1,\cdots,v_m$ 是标准正交的, 从定理 3.12 的证明知, 正交化方法得到的标准正交基含有这 m 个向量. □

正交化方法的证明中的一个技巧可以用于证明下面的结论.

定理 3.16 设 U 是有限维欧氏空间 V 的子空间, $U^\perp$ 是它的正交补, 那么

$$V = U \oplus U^\perp, \quad U^{\perp\perp} = U. \tag{3.1.12}$$

证明 取 U 的标准正交基 $e_1, \cdots, e_m$. 对 $w \in V$, 命

$$v = w - \sum_{i=1}^{m} (w|e_i)e_i.$$

对 $j = 1, \cdots, m$, 有

$$(v|e_j) = (w|e_j) - \sum_{i=1}^{m} (w|e_i)(e_i|e_j) = (w|e_j) - (w|e_j) = 0.$$

所以 v 正交于 U, 即 $v \in U^\perp$. 于是 $w = \sum\limits_{i=1}^{m} (w|e_i)e_i + v \in U + U^\perp$. 这说明 $V = U + U^\perp$.

设 $x \in U \cap U^\perp$. 因为 $x \in U^\perp$, 所以对任意的向量 $y \in U$, 有 $(x|y) = 0$. 但 x 也是 U 中的向量, 所以 $(x|x) = 0$. 于是 $x = 0$. 从而 $V = U \oplus U^\perp$.

现证 $U^{\perp\perp} = U$. 因为对任意的 $x \in U$, $y \in U^\perp$, 有 $(x|y) = 0$, 所以 $U \subset U^{\perp\perp}$. 由第一个结论知 $V = U^\perp \oplus U^{\perp\perp} = U \oplus U^\perp$. 所以 U 和 $U^{\perp\perp}$ 的维数相同. 这迫使 U 和 $U^{\perp\perp}$ 相等. □

四 正交矩阵 欧氏空间的重要特征是有度量, 这包括长度与角度. 当考虑基变换时, 自然希望能保持向量的长度和两个非零向量的角度. 这意味着标准正交基要变到标准正交基. 正交基之间的转换矩阵因此也就让人十分感兴趣.

设 $e_1, \cdots, e_n$ 和 $v_1, \cdots, v_n$ 是欧氏空间 V 的两个标准正交基. 命

$$v_j = a_{1j}e_1 + a_{2j}e_2 + \cdots + a_{nj}e_n, \quad 1 \leqslant j \leqslant n. \tag{3.1.13}$$

我们就得到了从基 (e_j) 到基 (v_j) 的转换矩阵 $A = (a_{ij})$, 它的第 k 列是 v_k 在基 (e_j) 下的坐标列向量. 利用所给基的标准正交性质得

$$\delta_{kj} = (v_k|v_j) = \left(\sum_{i=1}^{n} a_{ik}e_i \,\middle|\, \sum_{i=1}^{n} a_{ij}e_i \right) = \sum_{i=1}^{n} a_{ik}a_{ij}.$$

也就是说, 矩阵 A 的两个列向量的分量乘积之和有如下的关系:

$$\sum_{i=1}^{n} a_{ik}a_{ij} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases} \tag{3.1.14}$$

用矩阵表述就是

$${}^tA \cdot A = E. \tag{3.1.15}$$

由此可见 ${}^{t}A = A^{-1}$. 因为 $A \cdot A^{-1} = E$, 于是又有

$$A \cdot {}^{t}A = E. \tag{3.1.16}$$

这意味着矩阵 A 的两个行向量的分量乘积之和有如下的关系:

$$\sum_{j=1}^{n} a_{ij} a_{kj} = \begin{cases} 1, & k = i, \\ 0, & k \neq i. \end{cases} \tag{3.1.17}$$

定义 3.17 称方阵 A 为正交矩阵, 如果它的转置矩阵 ${}^{t}A$ 是其逆矩阵, 即 ${}^{t}AA = E$, 或等价地 $A\,{}^{t}A = E$.

从等式 (3.1.15) 和等式 (3.1.17) 立即得到下面的结论.

定理 3.18 (1) n 阶方阵 A 是正交矩阵当且仅当它的列向量形成欧氏空间 $\mathbb{R}^n$ 的标准正交基.

(2) n 阶方阵 A 是正交矩阵当且仅当它的行向量形成欧氏空间 $\mathbb{R}^n$ 的标准正交基.

如果 A 和 B 是 n 阶正交矩阵, 那么

$$\begin{aligned} &{}^{t}(AB)AB = {}^{t}B\,{}^{t}AAB = {}^{t}B\,B = E, \\ &{}^{t}(A^{-1})A^{-1} = ({}^{t}A)^{-1}A^{-1} = (A\,{}^{t}A)^{-1} = E^{-1} = E. \end{aligned}$$

所以 n 阶正交矩阵全体成为群, 称为 (n 阶)**正交群**, 记作 $O(n)$. 从等式 ${}^{t}AA = E$ 可见 $(\det A)^2 = 1$, 因此正交矩阵的行列式为 1 或 -1. 正交群 $O(n)$ 中行列式为 1 的矩阵全体形成一个子群, 称为 (n 阶)**特殊正交群**, 记作 $SO(n)$.

如果 A 是 n 阶正交矩阵, $e_1, \cdots, e_n$ 是标准正交基, 那么, 由公式 (3.1.13) 确定的向量组 $v_1, \cdots, v_n$ 也是标准正交基. 我们已经证明了下面的结论.

定理 3.19 欧氏空间中两个标准正交基之间的转换矩阵是正交矩阵. 且每一个正交矩阵都是某两个标准正交基之间的转换矩阵.

正交矩阵 A 有着丰富的几何意义. 比如, 从公式 (3.1.13) 知它的元素 a_{ij} 是两个单位向量的内积:

$$a_{ij} = (e_i|v_j) = \cos\varphi_{ij}, \tag{3.1.18}$$

其中 φ_{ij} 是原来的基向量 e_i 和新的基向量 v_j 之间的夹角. 向量在两个正交基下的坐标之间的联系仅用一个转换矩阵就可以表达: 如果 $x = \sum\limits_i x_i e_i = \sum\limits_i y_i v_i$, 那么

$$x_i = \sum_j a_{ij} y_j, \quad i = 1, 2, \cdots, n;$$

$$y_i = \sum_j a_{ji} x_j, \quad i = 1, 2, \cdots, n.$$

第二个等式来自 $A^{-1} = {}^tA$.

五 欧氏空间的同构 在抽象的欧氏空间中, 有勾股定理, 平行四边形等式, 三角不等式等. 这说明抽象欧氏空间的度量性质和我们熟知的三维欧氏空间 $\mathbb{R}^n$ 的度量性质有很多的共性. 对于 n 维的欧氏空间, 在标准正交基下, 内积的公式和标准欧氏空间的内积公式没有差别. 由于欧氏空间的度量由内积决定, 所以同维数的欧氏空间就度量性质而言应是没有本质差别的. 恰当的语言是欧氏空间的同构.

定理 3.20 维数相同的欧氏空间是同构的, 即如果 V, W 是两个 n 维欧氏空间, 那么存在向量空间的同构映射 $\varphi: V \to W$, 它还保持内积:

$$(u|v) = (\varphi(u)\|\varphi(v)), \quad \forall u, v \in V, \tag{3.1.19}$$

其中 $(\cdot\|\cdot)$ 是 W 上的内积.

证明 取 V 的标准正交基 $e_1, \cdots, e_n$ 和 W 的标准正交基 $w_1, \cdots, w_n$. 那么映射

$$\varphi: V \to W, \quad a_1e_1 + \cdots + a_ne_n \to a_1w_1 + \cdots + a_nw_n$$

是线性同构. 两个向量的内积就是在任一个标准正交基下的坐标分量分别相乘然后加起来, 所以等式 (3.1.19) 成立. □

这个定理表明一个欧氏空间上的度量性质在另一个同维数的欧氏空间上也是成立的. 比如公式 (3.1.5) 和 (3.1.6) 其实是等价的, 公式 (3.1.9) 和 (3.1.10) 也是等价的, 虽然形式上看上去差别挺大的.

借助内积, 可以建立欧氏空间 V 和它的对偶空间 V^* 的一个自然同构. 对任意的 $v \in V$, 从内积的双线性性质知映射

$$\Phi_v = (v|*) : V \to \mathbb{R}, \quad x \to (v|x)$$

是 V 上的线性函数.

定理 3.21 映射 $\Phi: V \to V^*$, $v \to \Phi_v$ 是线性同构. 这个同构把 V 的标准正交基 $e_1, \cdots, e_n$ 映到它的对偶基 $e^1, \cdots, e^n$.

证明 由于内积 $(v|x)$ 对 v 是线性的, 所以映射 Φ 是线性的:

$$\Phi(au + bv) = \Phi_{au+bv} = (au + bv|*) = a(u|*) + b(v|*) = a\Phi_u + b\Phi_v = a\Phi(u) + b\Phi(v).$$

而且 Φ 是单射, 因为 $v \neq 0$ 意味着 $\Phi_v(v) = (v|v) \neq 0$, 即 $\ker\Phi = 0$.

设 $e_1, \cdots, e_n$ 是 V 的标准正交基. 由于

$$\Phi_{e_i}(e_j) = (e_i|e_j) = \delta_{ij},$$

所以 $\Phi_{e_1},\cdots,\Phi_{e_n}$ 是 $e_1,\cdots,e_n$ 的对偶基. 这既证明了映射 Φ 的满射性, 也证明了定理的第二个断言. □

利用这个定理中的同构, 可以自然定义 V^* 上的内积: $(\Phi_u|\Phi_v)^*=(u|v)$. 从而, 定理 3.1.21 中的同构就成为欧氏空间同构.

六 辛空间 欧氏空间的内积是正定双线性型. 其他的非退化对称双线性型同样可以用于定义向量空间的度量, 比如, 狭义相对论中的闵可夫斯基空间. 非退化的斜对称双线性型也可以定义空间的度量, 称为**辛结构**(symplectic structure). 相应的几何称为辛几何. 带有辛结构的向量空间称为**辛空间**. 考虑辛空间时, 相应的斜对称双线性型常记做 $[*|*]$, 也称为**辛内积**.

如同欧氏空间, 辛空间中可以定义正交、辛基等概念. 设 $(V,[*|*])$ 是辛空间. 空间中的两个向量 u,v 称为**斜正交的**如果 $[u|v]=0$. 子空间 U 的**斜正交补**定义为

$$U^{\perp}=\{x\in V\,|\,[x|y]=0,\forall y\in U\}.$$

比较欧氏空间的正交 (补) 和辛空间的正交 (补) 的性质是有意思的. 在辛空间中, 由于辛内积的斜对称性, 任何向量都与自身斜正交, 从而向量的长度无法通过辛内积定义. 这也意味着 $U\cap U^{\perp}$ 一般不是零子空间. 不过仍然有

$$\dim U+\dim U^{\perp}=\dim V. \tag{3.1.20}$$

定义 3.22 设 $(V,[*|*])$ 是辛空间, U 是 V 的子空间.

(1) 称 U 是**辛子空间**如果 $U\cap U^{\perp}=0$;

(2) 称 U 是**迷向子空间**① 如果 $U\subseteq U^{\perp}$, 即 $[U|U]=0$;

(3) 称 U 是**余迷向子空间**②如果 $U^{\perp}\subseteq U$, 即 $[U^{\perp}|U^{\perp}]=0$;

(4) 称 U 是**拉格朗日子空间**③如果 $U=U^{\perp}$.

从定义可见, 拉格朗日子空间既是迷向的也是余迷向的, 和辛子空间是截然不同的子空间: 辛内积在辛子空间上的限制是非退化的, 在拉格朗日子空间上的限制是 0 内积. 从维数公式 (3.1.20) 知拉格朗日子空间是极大的迷向子空间, 也是极大的余迷向子空间, 其维数是 V 的维数的一半. 假设 V 的维数是 $2m$. 根据推论 1.88, 更确切地说, 一个等价的形式, 存在 V 的基 $e_1,\cdots,e_m,e_{m+1},\cdots,e_{2m}$, 使得辛内积 $[*|*]$ 在该基下的矩阵具有标准的形式:

$$J_0=\begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix},$$

① isotropic subspace.

② coisotropic subspace.

③ Lagrangian subspace.

其中 I_m 是 m 阶单位矩阵.

例 3.23 记号同上. 容易验证, 对 $1 \leqslant i \leqslant m$, 有:

向量组 $e_1, \cdots, e_i, e_{m+1}, \cdots, e_{m+i}$ 张成的子空间是辛子空间; 向量组 $e_1, \cdots, e_i$ 张成的子空间和向量组 $e_{m+1}, \cdots, e_{m+i}$ 张成的子空间是迷向子空间; 向量组 $e_1, \cdots, e_m, e_{m+1}, \cdots, e_{m+i}$ 张成的子空间是余迷向子空间; 向量组 $e_1, \cdots, e_m$ 张成的子空间和向量组 $e_{m+1}, \cdots, e_{2m}$ 张成的子空间是拉格朗日子空间.

如同欧氏空间的情形, 自然要讨论保持辛结构的线性算子.

定义 3.24 设 $(V, [*|*])$ 是辛空间. 称线性算子 $\mathcal{A}: V \to V$ 是**辛算子**如果它保持辛结构, 即

$$[\mathcal{A}x|\mathcal{A}y] = [x|y], \quad \forall x, y \in V.$$

容易验证, 如果 $\mathcal{A}$ 是 V 上的辛算子, 那么 $\mathcal{A}$ 是可逆的, 且其逆也是辛算子. 如果 $\mathcal{B}$ 也是 V 上的辛算子, 那么 $\mathcal{AB}$ 是辛算子. 于是 V 上的辛算子全体是 $GL(V)$ 的子群, 记作 $Sp(V)$, 称为**辛群**. 取定 V 的一个辛基使得辛内积在这个基下的矩阵是 J_0. 那么 $\mathcal{A}$ 保持辛内积等价于如下条件:

$${}^tX\,{}^tAJ_0AY = {}^tXJ_0Y,$$

即

$${}^tA\begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}A = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}, \tag{3.1.21}$$

其中 X 和 Y 分别是向量 x, y 在辛基下的坐标列向量, A 是 $\mathcal{A}$ 在辛基下的矩阵. 满足等式 (3.1.21) 的方阵 A 称为**辛矩阵**. 从而辛群 $Sp(V)$ 的矩阵形式就是阶数为 $\dim V$ 的辛矩阵全体形成的群:

$$Sp_{2m}(\mathbb{R}) = \{A \in GL_{2m}(\mathbb{R}) \mid {}^tAJ_0A = J_0\}.$$

从等式 (3.1.21) 知辛矩阵的行列式为 ± 1. 事实上, 辛矩阵的行列式为 1. 这可以利用斜对称矩阵的普法夫 (见定理 1.89) 计算出来:

$$1 = \mathrm{Pf}(J_0) = \mathrm{Pf}({}^tAJ_0A) = (\det A)(\mathrm{Pf}(J_0)) = \det A.$$

于是 $Sp_{2m}(\mathbb{R})$ 是 $SL_{2m}(\mathbb{R})$ 的子群. 当 $m = 1$ 时, 这两个群是一样的. 因为, 对实数 a, b, c, d, 辛矩阵条件

$${}^t\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & bc-ad \\ ad-bc & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

等价于 $ad - bc = 1$.

显然, 辛算子把辛基变成辛基, 反之, 把辛基变成辛基的线性算子是辛算子. 另外, 两个辛基之间的转换矩阵是辛矩阵, 而且任意的辛矩阵是某两个辛基的转换矩阵. 于是, 辛空间 V 上的辛算子和 V 的辛基之间可以建立自然的一一对应: $\mathcal{A} \to (e_i)A$, 其中 (e_i) 是一个取定的辛基, A 是 $\mathcal{A}$ 在这个基下的矩阵. 这个一一对应和正交群与相应的欧氏空间中的正交基之间的一一对应是类似的, 也和 $GL(V)$ 与 V 的基之间的一一对应类似.

辛算子 (辛矩阵) 的特征多项式满足一个有趣的等式.

定理 3.25　设 $\mathcal{A}$ 是 $2m$ 维辛空间上的辛算子, 那么其特征多项式有如下性质:

$$\chi_{\mathcal{A}}(t) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right).$$

也就是说, $\chi_{\mathcal{A}}(t) = \sum_{i=0}^{2m} a_i t^i$ 的系数具有对称性: $a_i = a_{2m-i}$, 其中 $i = 0, 1, \cdots, m$.

证明　设 $\mathcal{A}$ 在某个辛基下的矩阵是 A. 那么 ${}^tAJ_0A = J_0$. 从而 $A = J_0^{-1} \cdot {}^tA^{-1} \cdot J_0$. 于是

$$\begin{aligned}
\chi_{\mathcal{A}}(t) &= \det(tE - A) = \det(tE - J_0^{-1} \cdot {}^tA^{-1} \cdot J_0) = \det(J_0^{-1}(tE - {}^tA^{-1})J_0) \\
&= \det(tE - {}^tA^{-1}) = \det({}^t(tE - A^{-1})) = \det(tE - A^{-1}) \\
&= \det(tE - A^{-1})\det A = \det(tA - E) = \det(E - tA) \\
&= t^{2m}\det\left(\frac{1}{t}E - A\right) = t^{2m}\chi_{\mathcal{A}}\left(\frac{1}{t}\right).
\end{aligned}$$

□

从这个定理可知, 如果 λ 是辛算子 $\mathcal{A}$ 的特征多项式 $\chi_{\mathcal{A}}(t)$ 的根 (可能是虚数), 那么 λ^{-1} 也是 $\chi_{\mathcal{A}}(t)$ 的根. 当 λ 是虚数时, 因为这个特征多项式是实系数的, 所以其共轭 $\bar{\lambda}$ 和共轭的逆 $\bar{\lambda}^{-1}$ 也是 $\chi_{\mathcal{A}}(t)$ 的根. 从而, 取逆与共轭就可以把 $\chi_{\mathcal{A}}(t)$ 的根分成不同的组, 一个组 λ, λ^{-1}, $\bar{\lambda}$, $\bar{\lambda}^{-1}$ 所含根的个数在 4, 2, 1 之中, 根据 λ 的虚实和模是否等于 1 而定.

相同维数的辛空间之间的线性映射如果把辛基映到辛基, 那么这个线性映射保持辛内积. 在这个意义下, 相同维数的辛空间都是同构的. 所以, 要讨论辛空间, 很多时候可以仅讨论坐标空间 $\mathbb{R}^{2m}$ 的辛结构.

设 V 是 $2m$ 维的辛空间. 在取定的辛基下, 辛内积有如下的形式:

$$[x|y] = \sum_{i=1}^{m}(x_iy_{i+m} - x_{i+m}y_i).$$

作为对比, 引入正定对称双线性型

$$(x|y) = \sum_{i=1}^{2m} x_iy_i.$$

于是, V 又有欧氏空间的结构. 用 X 和 Y 分别记 x 和 y 在给定辛基 (e_i) 下的坐标列向量. 定义 V 的算子 $\mathcal{J}: x=(e_1,\cdots,e_{2m})X \to (e_1,\cdots,e_{2m})J_0X$. 那么

$$[x|y] = {}^tXJ_0Y = (x|\mathcal{J}y).$$

因为 $J_0 \in Sp_{2m}(\mathbb{R})$ 且 $J_0^2=-E$, 所以 $\mathcal{J}\in Sp(V)$ 且 $\mathcal{J}^2=-\mathcal{E}$. 注意子空间 U 是迷向的等价于 U 与 $\mathcal{J}U$ 在 V 的欧氏结构中是正交的.

在 3.6 节中我们将看到满足条件 $\mathcal{J}^2=-\mathcal{E}$ 的算子 $\mathcal{J}$ 可以用来定义 V 上的复结构. 于是辛空间可以自然地成为复向量空间.

习 题 3.1

1. 在欧几里得平面几何和立体几何中解释格拉姆–施密特正交化方法的几何意义.

2. 在次数不超过 n 的实多项式形成的向量空间 P_{n+1} 中定义

$$(f|g)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)g\left(\frac{k}{n}\right).$$

(1) 证明: $(\cdot|\cdot)$ 是 P_{n+1} 的内积;

(2) 对 $f=t$, $g=at+b$, 求 $(f|g)$;

(3) 求出所有正交于 t 的多项式.

3. 在次数不超过 2 的实多项式形成的向量空间 P_3 中, 对于内积 $(f|g)=\int_{-1}^{1} f(t)g(t)dt$, 向量 1 和向量 t 是正交的. 找出:

(1) 子空间 $\langle 1,\ t\rangle^{\perp}$;

(2) 向量 1, $t+1$ 之间的夹角; 向量 t, $t+1$ 之间的夹角;

(3) P_3 的一个标准正交基.

4. 在所有实多项式形成的向量空间 $\mathbb{R}[t]$ 中, 定义 $(f|g)=\int_0^{\infty} e^{-t}f(t)g(t)dt$.

(1) 证明: 这是 $\mathbb{R}[t]$ 的一个内积;

(2) 证明: $(t^n\,|\,t^m)=(m+n)!$;

(3) 求出与 $t+1$ 正交的所有线性多项式 $at+b$.

5. 在向量空间 $\mathbb{R}[t]$ 中, 定义 $(f|g)=\int_1^{e}(\ln t)f(t)g(t)dt$.

(1) 证明: 这是 $\mathbb{R}[t]$ 的一个内积;

(2) 求出与 1 正交的一个线性多项式 $at+b$.

6. 验证: 在 3 维欧几里得向量空间 V 中, 二次型

$$\|x\|^2=(x|x)=3x_1^2+2x_2^2+x_3^2-4x_1x_2-2x_1x_3+2x_2x_3,\quad x\in V$$

是正定的, 从而相应的双线性型定义了 V 的一个内积. 对这个内积, 找出向量 $x=[1,1,1]$, $y=[2,2,1]$ 之间的夹角, 并找出所有与 x 正交的向量.

7. 如果实方阵的列向量都是非零的且相互正交的, 求其逆矩阵.

8. 运用正交化方法, 求出 $\mathbb{R}^4$ 中由向量 (1,2,1,3), (4,1,1,1), (3,1,1,0) 张成的线性子空间的一个标准正交基.

9. 把向量 $X_1 = \dfrac{1}{2}(1, -1, 1, 1)$ 扩充为 $\mathbb{R}^4$ 的标准正交基.

10. 运用正交化方法证明: 任意的非退化矩阵 $A = (a_{ij}) \in M_n(\mathbb{R})$ 都可以分解成 $A = BC$ 的形式, 其中 B 是正交矩阵, C 是上三角矩阵, 且 $\det A = \pm \det C$.

11. 命 $\mathfrak{o}(n)$ 为 n 阶斜对称实方阵全体形成的空间. 凯莱变换

$$K = (E - A)^{-1}(E + A), \quad A = (E - K)^{-1}(E + K), \tag{3.1.22}$$

建立了正交群 $O(n)$ 和向量空间 $\mathfrak{o}(n)$ 之间一个很好的联系.

证明: 如果 $A \in O(n)$, $\det(E - A) \neq 0$, 那么 $K \in \mathfrak{o}(n)$. 反之, 如果 $K \in \mathfrak{o}(n)$, 那么上式定义的矩阵 A 是正交矩阵, 且 $-1 \notin \operatorname{Spec} A$.

正交群 $O(n)$ 是 n^2 维向量空间 $M_n(\mathbb{R})$ 中的子集, 由 (3.1.14) 或 (3.1.17) 中 $n(n+1)/2$ 个方程定义. 所以, $O(n)$ 应该是具有维数 $n^2 - n(n+1)/2 = n(n-1)/2 = \dim \mathfrak{o}(n)$ 的一个 "代数流形" (代数簇). 在 $O(n)$ 中满足方程 $\det(E - A) = 0$ 的矩阵 A 形成 $O(n)$ 的一个 "超曲面" S. 凯莱变换建立了 $O(n)\backslash S$ 和 $\mathfrak{o}(n)$ 之间的联系.

可以提出另一个凯莱变换:

$$K = (E + A)^{-1}(E - A), \quad A = (E + K)^{-1}(E - K). \tag{3.1.23}$$

这时, 需要从 $O(n)$ 中挖去方程 $\det(E + A) = 0$ 确定的 "超曲面" T, 即, 挖去谱中含 -1 的那些正交矩阵.

证明: 变换 (3.1.23) 建立了 $O(n)\backslash T$ 和 $\mathfrak{o}(n)$ 之间的联系.

12. 证明: 在凯莱变换 (3.1.22) 或 (3.1.23) 中得到的正交矩阵的行列式均为 1(用 $SO(n)$ 表示所有这样的矩阵的集合).

13. 设 K 是 n 阶斜对称实方阵, λ 是其非零的纯虚特征根, X 是 λ 的特征向量: $KX = \lambda X$. 记 $X = Y + iZ$, 其中 Y, Z 都是列向量欧氏空间 $\mathbb{R}^n$ 中的元素. 证明: Y 和 Z 正交且长度相等.

14. 证明: n 阶正交矩阵 A 的特征多项式 $\chi_A(t)$ 具有性质:

$$t^n \chi_A\left(\frac{1}{t}\right) = \pm \chi_A(t).$$

15. 设方阵 $A = [A_{(1)}, \cdots, A_{(n)}] \in M_n(\mathbb{R})$ 的行向量相互正交. 证明:

$$|\det A| = \|A_{(1)}\| \cdot \|A_{(2)}\| \cdots \|A_{(n)}\|.$$

(标准欧氏空间 $\mathbb{R}^n$ 中的范数.)

16. 对任意的实方阵 $X = [X_{(1)}, \cdots, X_{(n)}] \in M_n(\mathbb{R})$, 证明:

$$|\det X| \leqslant \|X_{(1)}\| \cdot \|X_{(2)}\| \cdots \|X_{(n)}\|.$$

(阿达马不等式.)

17. 证明公式 (3.1.20).

18. 设 V 是辛空间. 证明

(1) 每一个迷向子空间都落在某个拉格朗日子空间中;

(2) 迷向子空间的维数不超过 V 的维数之半;

(3) 迷向子空间是拉格朗日子空间当且仅当其维数是 V 的维数之半;

(4) 子空间 U 是余迷向的当且仅当辛内积自然给出了商空间 $U/U^{\perp}$ 上的辛内积;

(5) 子空间 U 是辛子空间当且仅辛内积在 U 上的限制是非退化的;

(6) 子空间的斜正交补的斜正交补是子空间自身: $(U^{\perp})^{\perp}=U$.

3.2 埃尔米特向量空间

复数域既是代数闭域, 又有度量的完备性 (即柯西序列都是收敛的), 复数域的上数学理论丰富又深刻, 复向量空间是其中的一个基础部分.

尝试用对称双线性型定义复向量空间的度量是不能令人满意的. 复向量空间上的非退化二次型的标准形只有一个: $x_1^2+\cdots+x_n^2$. 它不是正定的, 因为 $i^2=-1$, 从而用它定义向量的长度会带来混乱.

复向量空间也是实向量空间, 可以通过它的实向量空间结构定义正定对称双线性型, 引入欧氏内积, 进而定义向量的长度和角度. 这样做不是自然的, 但富有启发性. 以坐标空间 $\mathbb{C}^n$ 作为例子. 当 $n=1$ 时, 复数 $z=x+iy$ 的模 $|z|=\sqrt{x^2+y^2}$ 就是以原点为起点, z 为终点的向量的长度. 这个长度的平方如果仅用复数表达就是 $|z|^2=z\bar{z}$. 它并不是二次型. 一般地, $\mathbb{C}^n$ 中的向量 $z=(z_1,\cdots,z_n)$(还是以原点为起点) 的长度的平方由勾股定理给出:

$$z_1\bar{z}_1+\cdots+z_n\bar{z}_n=x_1^2+y_1^2+\cdots+x_n^2+y_n^2, \tag{3.2.24}$$

其中 $z_k=x_k+iy_k$. 这不是二次型. 表达式 (3.2.22) 对应于 $\mathbb{C}^n\times\mathbb{C}^n$ 上的函数

$$f(u,v)=u_1\bar{v}_1+\cdots+u_n\bar{v}_n. \tag{3.2.25}$$

这个函数不是双线性型, 但与双线性型差得不多: 第一个变量是线性的, 第二个变量是半线性的, 所以是一个半线性型. 这个例子说明对复向量空间, 有必要考虑一个半线性型. 这类线性型的英文是 sesquilinear①, 被奇怪地翻译成半双线性型, 似乎也没法改了.

一 半双线性型

定义 3.26 设 V 是复向量空间. 函数 $f:V\times V\to\mathbb{C}$ 称为 (V 上的)**半双线性型**如果它对第一个变量是线性的, 对第二个变量是半线性的. 即对任意复数 a,b 和 V 中的向量 x,y,z, 有:

① sesqui 源自拉丁文, 意思是“一个半”.

(1) $f(ax+by,z)=af(x,z)+bf(y,z)$,

(2) $f(x,ay+bz)=\bar{a}f(x,y)+\bar{b}f(x,z)$,

其中 $\bar{a}$ 和 $\bar{b}$ 分别表示 a 和 b 的共轭复数.

注　有时半双线性型定义为对第一个变量半线性, 第二个变量线性.

半双线性型的理论和双线性型的理论类似.

设 $e_1,\cdots,e_n$ 是 V 的基, 那么半双线性型 f 在诸 (e_i,e_j) 处的值完全确定了 f:

$$f(x,y)=\sum_{i,j}f_{ij}x_i\bar{y}_j,\quad f_{ij}=f(e_i,e_j),$$

其中 $x=\sum\limits_i x_ie_i,\ y=\sum\limits_j y_je_j$. 矩阵 $F=(f_{ij})$ 称为 f 在基 (e_i) 下的**矩阵**. 如果从基 $e_1,\cdots,e_n$ 到另一个基 $e_1',\cdots,e_n'$ 的转换矩阵是 A,

$$(e_1',\cdots,e_n')=(e_1,\cdots,e_n)A.$$

那么 f 在基 (e_i') 下的矩阵为

$$F'={}^tA\cdot F\cdot\bar{A},\tag{3.2.26}$$

其中 $\bar{A}=(\bar{a}_{ij})$. 由此可见, f 在任意两个基下的矩阵都有相同的秩, 这个值称为 f 的**秩**, 记作 $\operatorname{rank}f$. 半双线性型 f 称为**非退化的**如果其核是零:

$$\ker f=\{x\in V\mid f(x,y)=0,\ \ \forall y\in V\}=0.$$

这个条件等价于 F 是非退化的.

称半双线性型 f 是**埃尔米特型**如果 $f(x,y)=\overline{f(y,x)}$, 是**斜埃尔米特型**如果 $f(x,y)=-\overline{f(y,x)}$. 这两种半双线性型没有本质的差别: 乘以 i 即把埃尔米特型变成斜埃尔米特型, 反之亦然.

半双线性型 f 的埃尔米特性质等价于它的矩阵 $F=(f_{ij})$ 满足条件 $f_{ij}=\bar{f}_{ji}$, 即

$$F^*=F,\tag{3.2.27}$$

其中 $F^*={}^t\bar{F}$ 是 F 的共轭转置 (或转置共轭). 类似地, 半双线性型 f 的斜埃尔米特性质等价于它的矩阵 F 有如下性质:

$$F^*=-F.\tag{3.2.28}$$

满足条件 (3.2.27) 的矩阵 F 称为**埃尔米特矩阵**, 满足条件 (3.2.28) 的矩阵 F 称为**斜埃尔米特矩阵**.

每个埃尔米特型 f 确定了一个**埃尔米特二次型**

$$q(x) = f(x, x).$$

易见这是 V 上的实值函数. 由于

$$\begin{aligned} q(x+y) &= q(x) + q(y) + f(x,y) + f(y,x), \\ q(x+iy) &= q(x) + q(y) - if(x,y) + if(y,x), \end{aligned} \tag{3.2.29}$$

所以从 q 可以求出 f. 特别, 如果 q 恒为零, 那么 f 恒为零.

二　埃尔米特空间　埃尔米特型 f 称为**正定的**如果它定义的二次型 $q(x)$ 是正定的, 即 $q(x) > 0$ 如果 $x \neq 0$.

定义 3.27　称复向量空间 V 为**埃尔米特空间**(或**酉空间**), 简称为**埃氏空间**, 如果它带有给定的正定埃尔米特型. 这个埃尔米特型称为 V 的**内积**(或**纯量积**), 记作 $(\,|\,)$. 复数 $(x|y)$ 称为向量 x, y 的内积.

有必要把内积定义的性质列出来:

$$\begin{aligned} &(x|y) = \overline{(y|x)}; \\ &(ax + by|z) = a(x|y) + b(y|z); \\ &(x|x) > 0, \quad 如果 x \neq 0. \end{aligned}$$

例 3.28　带着内积

$$(x|y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n,$$

坐标空间 $\mathbb{C}^n$ 成为埃尔米特空间. 这个内积在标准基下的矩阵是单位矩阵 E. 根据公式 (3.2.24), 这个内积在一般的基下的矩阵具有形式 $F' = {}^tA \cdot \bar{A}$, 其中 A 是可逆 n 阶复方阵.

例 3.29　闭区间 $[a, b]$ 上的复值连续函数空间 $C_2(a, b)$ 有内积

$$(f|g) = \int_a^b f(t)\overline{g(t)}dt,$$

从而是埃氏空间.

如同欧氏空间, 在埃尔米特空间中, 向量 v 的长度定义为

$$\|v\| = \sqrt{(v|v)}.$$

长度为 1 的向量称为**单位向量**. 欧氏空间中关于向量长度的性质在埃氏空间中也是成立的. 首先, 对复数 λ 和向量 x, 有

$$\|\lambda x\| = \sqrt{(\lambda x|\lambda x)} = \sqrt{\lambda\bar{\lambda}(x|x)} = \sqrt{|\lambda|^2(x|x)} = |\lambda|\sqrt{(x|x)},$$

所以

$$\|\lambda x\| = |\lambda| \cdot \|x\|. \tag{3.2.30}$$

其次, 柯西–施瓦茨不等式成立:

$$|(x|y)| \leqslant \|x\| \cdot \|y\|, \tag{3.2.31}$$

等式成立当且仅当 x, y 成比例 (即线性相关).

证明 如果 x, y 成比例, 那么 $y = \lambda x$, 从而有

$$|(x|y)| = |\bar{\lambda}(x|x)| = |\lambda| \, \|x\|^2 = \|x\| \cdot |\lambda| \|x\| = \|x\| \cdot \|y\|.$$

假设 x, y 不成比例. 把内积 $(x|y)$ 写成三角形式 $|(x|y)|e^{i\varphi}$, 其中 $e^{i\varphi} = \cos\varphi + i\sin\varphi$. 对任意实数 $t \in \mathbb{R}$, 有

$$0 < (xt + e^{i\varphi}y|xt + e^{i\varphi}y) = \|x\|^2t^2 + [(x|y)e^{-i\varphi} + \overline{(x|y)}e^{i\varphi}]t + \|y\|^2.$$

由于 $(x|y)e^{-i\varphi} = |(x|y)| = \overline{(x|y)}e^{i\varphi}$, 所以上面的不等式可以写成:

$$0 < \|x\|^2t^2 + 2|(x|y)|t + \|y\|^2.$$

上式右边的二次多项式的判别式为负, 故此时有 $|(x|y)| < \|x\| \cdot \|y\|$. □

由柯西–施瓦茨不等式可以得到三角不等式:

$$\|x \pm y\| \leqslant \|x\| + \|y\|. \tag{3.2.32}$$

它的一个常用的等价形式是

$$\|x - z\| \leqslant \|x - y\| + \|y - z\|.$$

例 3.30 把柯西–施瓦茨不等式和三角不等式用到埃氏空间 $\mathbb{C}^n$, 得

$$\left|\sum_{i=1}^n x_i\bar{y}_i\right| \leqslant \sqrt{\sum_{i=1}^n |x_i|^2} \cdot \sqrt{\sum_{i=1}^n |y_i|^2},$$

$$\sqrt{\sum_{i=1}^n |x_i \pm y_i|^2} \leqslant \sqrt{\sum_{i=1}^n |x_i|^2} + \sqrt{\sum_{i=1}^n |y_i|^2};$$

应用到 $C_2(a, b)$, 得

$$\left|\int_a^b f(t)\overline{g(t)}dt\right| \leqslant \sqrt{\int_a^b |f(t)|^2dt} \cdot \sqrt{\int_a^b |g(t)|^2dt},$$

$$\sqrt{\int_a^b |f(t)\pm g(t)|^2dt}\leqslant\sqrt{\int_a^b |f(t)|^2dt}+\sqrt{\int_a^b |g(t)|^2dt}\,.$$

利用柯西–施瓦茨不等式可以知道, 对于非零向量 x 和 y, 有唯一的角 φ, $0\leqslant\varphi\leqslant\pi/2$, 使得

$$\cos\varphi=\frac{|(x|y)|}{\|x\|\cdot\|y\|}.$$

三 标准正交基 对有限维埃氏空间, 如果内积在某个基下的矩阵是单位矩阵, 那么这个基就称为**标准正交基**. 也就是说, 称基 $e_1,\cdots,e_n$ 为标准正交基如果 $(e_i|e_j)=\delta_{ij}$.

运用格拉姆–施密特正交化方法可以知道有限维埃氏空间有标准正交基. 首先任取一个基 $v_1,\cdots,v_n$. 命 $e_1=v_1/\|v_1\|$. 如果已经确定了 $e_1,\cdots,e_{k-1}$ 使得 $(e_i|e_j)=\delta_{ij}$ 对所有的 $i,j=1,2,\cdots,k-1$, 则定义 $e_k=e_k'/\|e_k'\|$, 其中

$$e_k'=v_k-(v_k|e_1)e_1-\cdots-(v_k|e_{k-1})e_{k-1}.$$

那么 e_k 是单位向量且 $(e_k|e_i)=0$ 对 $i=1,\cdots,k$. 如此下去, 就可以得到标准正交基 $e_1,\cdots,e_n$.

如同欧氏空间, 埃氏空间 V 中的向量 u,v 称为**正交的**如果 $(u|v)=0$, 子空间 U 的正交补是

$$U^{\perp}=\{v\in V\mid (v|u)=0\ \forall u\in U\}.$$

下面的性质是意料之中的, 也是容易验证的.

(1) 如果埃氏空间 V 中的非零向量 $v_1,\ \cdots,v_m$ 互相正交, 那么这些向量线性无关. 如果 $m=\dim V$, 那么这些向量构成 V 的一个正交基.

(2) $V=U\oplus U^{\perp}$, $U^{\perp\perp}=U$.

标准正交基对计算是很有用的.

定理 3.31 设 $e_1,\cdots,e_n$ 是埃氏空间 (或欧氏空间)V 的标准正交基. 那么, 对 $x,y\in V$ 有:

(1) $x=\sum\limits_i(x|e_i)e_i$;

(2) $(x|y)=\sum\limits_i(x|e_i)(e_i|y)$ (帕塞瓦尔等式);

(3) $\|x\|^2=\sum\limits_i|(x|e_i)|^2$.

证明 留作练习.

通过标准正交基可以建立同维数的埃氏空间的同构. 设 $e_1,\cdots,e_n$ 和 $w_1,\cdots,w_n$ 分别是埃氏空间 V 和 W 的标准正交基, 那么映射

$$\varphi:V\to W,\quad \sum_i x_ie_i\to\sum_i x_iw_i$$

不仅是线性空间的同构, 还保持内积: 对 $x, y \in V$, 有 $(x|y) = (\varphi(x)|\varphi(y))$. 特别, 任意 n 维埃氏空间都与坐标空间 $\mathbb{C}^n$ 同构.

埃氏空间的内积也可以用于讨论埃氏空间 V 和它的对偶的联系, 不过无法得到它们之间的自然同构. 设 $v \in V$, 那么

$$f_v : V \to \mathbb{C}, \quad x \to (x|v)$$

是 V 上的线性函数. 如果 $e_1, \cdots, e_n$ 是 V 的基, 则 $f_{e_1}, \cdots, f_{e_n}$ 线性无关, 且对 $v = \sum_i x_i e_i$ 有

$$f_v = \sum_i \bar{x}_i f_{e_i}.$$

这表明所有的 f_v $(v \in V)$ 构成 V 的对偶空间 V^*, 且自然的映射 $v \to f_v$ 是**半线性的**: $f_{au+bv} = \bar{a} f_u + \bar{b} f_v$.

对 $v \in V$, 考虑

$$g_v : V \to \mathbb{C}, \quad x \to (v|x).$$

易见 g_v 是 V 的**半线性函数**: $g_v(ax + by) = \bar{a} g_v(x) + \bar{g}_v(y)$. V 上的半线性函数 g 完全由它在任意一个基上的值确定. 对复数 α, β 和 V 上的半线性函数 g, h, 定义 $(\alpha g + \beta h)(x) = \alpha g(x) + \beta h(x)$. 结果仍是 V 上的半线性函数. 直接验证可知 V 上的半线性函数全体因此成为复向量空间, 记作 $V^\star$, 维数等于 $\dim V$. 简单的计算表明, 映射

$$g : V \to V^\star, \quad v \to g_v$$

是线性同构.

上面的讨论给出了埃氏空间 V 上的线性函数和半线性函数一个自然的联系:

$$V^* \to V^\star, \quad f_v \to g_v.$$

注意我们有等式:

$$f_v(x) = (x|v) = \overline{(v|x)} = \overline{g_v(x)}.$$

所以, 对每个线性函数 f, 存在唯一的半线性函数 g, 使得 $g(x) = \overline{f(x)}$ 对任意的 $x \in V$. 自然, 半线性函数 g 可以称作线性函数 f 的**共轭**.

四　酉矩阵　欧氏空间中标准正交基之间的转换矩阵有一些独特的性质, 有专门的名称–正交矩阵. 而且同阶的正交矩阵全体形成一个群. 对埃氏空间, 情况是类似的.

设 V 是埃氏空间, 从标准正交基 $e_1, \cdots, e_n$ 到标准正交基 $v_1, \cdots, v_n$ 的转换矩阵是 $A = (a_{ij})$:

$$(v_1, \cdots, v_n) = (e_1, \cdots, e_n)A. \tag{3.2.33}$$

那么 $v_j = \sum_i a_{ij} e_i$. 从而

$$\delta_{jk} = (v_j|v_k) = \sum_{i,s} a_{ij}\bar{a}_{sk}(e_i|e_s) = \sum_i a_{ij}\bar{a}_{ik}. \tag{3.2.34}$$

写成矩阵形式就是

$$A^* \cdot A = E, \tag{3.2.35}$$

其中 $A^* = {}^t\bar{A}$ 是 A 的共轭转置. 于是 $A^* = A^{-1}$. 因为 $A \cdot A^{-1} = E$, 得

$$A \cdot A^* = E. \tag{3.2.36}$$

定义 3.32 满足条件 (3.2.35) 或 (3.2.36) 的矩阵 A 称为**酉矩阵**.

容易看出, n 阶复方阵是酉矩阵当且仅当它的列向量形成埃氏空间 $\mathbb{C}^n$ 的标准正交基, 也当且仅当它的行向量形成埃氏空间 $\mathbb{C}^n$ 的标准正交基. 显然实的酉矩阵就是正交矩阵. 由于 $\det \bar{A} = \overline{\det A}$, $\det {}^tA = A$, 所以 $\det A^* = \overline{\det A}$. 由式 (3.2.35) 知 $|\det A| = 1$.

如果 A 是 n 阶酉矩阵, $e_1, \cdots, e_n$ 是 V 的标准正交基, 那么通过公式 (3.2.33) 定义的向量组 $v_1, \cdots, v_n$ 是 V 的基, 且满足关系式 (3.2.34), 从而是标准正交基. 所以, n 阶复方阵是酉矩阵当且仅当它是标准正交基之间的转换矩阵.

设 (e_i), (v_i), (u_i) 是 n 维埃氏空间 V 的标准正交基, 从 (e_i) 到 (v_i) 的转换矩阵是 A, 从 (v_i) 到 (u_i) 的转换矩阵是 B. 那么, 从 (e_i) 到 (u_i) 的转换矩阵是 AB, 从 (v_i) 到 (e_i) 的转换矩阵是 A^{-1}. 所以 n 阶酉矩阵全体在矩阵乘法下形成一个群, 称为 (n 阶)**酉群**, 记作 $U(n)$. 我们已经知道, 它包含正交群 $O(n)$.

当然, 可以直接从酉矩阵的定义看出 n 阶酉矩阵全体形成 $GL_n(\mathbb{C})$ 的子群. 设 A, B 是 n 阶酉矩阵. 我们有

$$\begin{aligned} &A^* = A^{-1}, \quad (A \cdot B)^* = B^* \cdot A^*, \\ &A^{-1}(A^{-1})^* = A^*(A^{-1})^* = (A^{-1}A)^* = E^* = E, \\ &(AB)(AB)^* = A(BB^*)A^* = AEA^* = AA^* = E. \end{aligned}$$

所有行列式为 1 的 n 阶酉矩阵构成 $U(n)$ 的一个子群, 称为**特殊酉群**, 记作 $SU(n)$. 类似地, 所有行列式为 1 的 n 阶正交矩阵构成 $O(n)$ 的一个子群, 称为**特殊正交群**, 记作 $SO(n)$. 我们有如下的关系:

$$SO(n) = O(n) \cap SL_n(\mathbb{R}) \subset SU(n) = U(n) \cap SL_n(\mathbb{C}).$$

利用其他的非退化的实二次型也可以定义一些很有意思的群, 它们都是 $GL_n(\mathbb{R})$ 的子群. 同样, 对一般的域 K, 利用二次型 $x_1^2 + \cdots + x_n^2$ 可以定义正交群

$O(n,K)$. 如果域 K 有自同构 $\sigma: K \to K$, 其平方是恒等映射, 那么也可以定义酉群 $U(n,K)$. 例如, 有理数域的二次扩张 $\mathbb{Q}(\sqrt{d})$ 有这样的自同构: $a+b\sqrt{d} \to a-b\sqrt{d}$, 其中 d 是一个非平方整数.

五 度量空间和赋范向量空间 欧氏空间和埃氏空间中的三角不等式 (命题 3.7 和式 (3.2.30)) 是度量的基本性质, 抽象出来就得到一般的度量空间的概念.

定义 3.33 设 E 是集合. 称函数 $d: E\times E \to \mathbb{R}$ 为 E 的一个**度量**如果以下条件成立:

(1) $d(u,v) \geqslant 0$ (非负性);

(2) $d(u,v)=0 \Longleftrightarrow u=v$ (同一性原则);

(3) $d(u,v)=d(v,u)$ (对称性);

(4) $d(u,w) \leqslant d(u,v)+d(v,w)$ (三角不等式).

对 (E,d) 称为一个**度量空间**. 抽象地讨论时, 集合 E 中的元素常称**点**, 数 $d(u,v)$ 称为 u,v 间的**距离**.

例 3.34 欧氏空间和埃氏空间都是度量空间, 向量 u,v 间的距离是 $d(u,v)=\|u-v\|$.

例 3.35 对 $E=C_2(a,b)$, 除了内积 (见例 3.29) 给出的度量外, 下面的两个函数

$$d_1(f,g)=\max_{a\leqslant t\leqslant b}|f(t)-g(t)|, \quad d_2(f,g)=\int_a^b |f(t)-g(t)|dt,$$

也是 E 的度量.

有了度量, 就可以建立一些基本的几何概念和分析概念 (极限、收敛等). 度量空间 (E,d) 的子空间

$$B(a,r)=\{x\in E \mid d(a,x)<r\},$$
$$\bar{B}(a,r)=\{x\in E \mid d(a,x)\leqslant r\},$$
$$S(a,r)=\{x\in E \mid d(a,x)=r\},$$

分别称为以点 a 为中心, r 为半径的**开球**、**闭球**、**球面**. 子集 F 称为**有界的**如果它落在一个半径有限的球中.

称 (E,d) 中的点序列 $a_1,a_2,\cdots,a_n,\cdots$ **收敛到**E 中的点 a 如果 $\lim\limits_{n\to\infty} d(a_n,a)=0$. 称这个序列是**柯西序列**如果对任意的 $\varepsilon>0$, 存在 $N=N(\varepsilon)$ 使得 $m,n>N$ 时有 $d(a_m,a_n)<\varepsilon$. 称度量空间是**完备的**如果其中的柯西序列都是收敛的. 数学分析中已经证明了实数域和复数域都是完备度量空间. 由此可以知道, 向量空间 $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 对下面三个度量

$$d(x,y)=\left(\sum_{i=1}^n |x_i-y_i|^2\right)^{\frac{1}{2}},$$

$$d_1(x,y)=\max_i(|x_i-y_i|),\quad d_2(x,y)=\sum_{i=1}^n|x_i-y_i|,$$

都是完备的.

在欧氏空间和埃氏空间中, 前面定义的度量 d 是通过向量的长度函数 $\|\cdot\|$ 定义的. 这说明长度函数 $\|\cdot\|$ 是更基本的对象. 它具有以下的性质:

$$\|x\|>0 \text{ 如果} x\neq 0,\quad \|0\|=0;$$
$$\|\lambda x\|=|\lambda|\cdot\|x\|,\quad \text{对任意的数}\lambda,\ \text{向量 } x;$$
$$\|x+y\|\leqslant\|x\|+\|y\|,\quad \text{对任意的向量} x,y.$$

一般地, 我们有如下定义.

定义 3.36 对复 (或实) 向量空间 V, 称函数 $\|\cdot\|:V\to\mathbb{R}$ 为 V 的一个**范数**(norm) 如果它满足上面三个条件. 配对 $(V,\|\cdot\|)$ 则称为**赋范向量空间**. 范数 $\|\cdot\|$ 提供了 V 的度量 $d(x,y)=\|x-y\|$. 如果 (V,d) 是完备度量空间, 则称 $(V,\|\cdot\|)$ 为**巴拿赫空间**.

设 d 是复 (或实) 向量空间 V 上的一个度量. 如果它具有平移不变性和伸缩性:

$$d(x,y)=d(x+z,y+z),\quad d(\lambda x,\lambda y)=|\lambda|d(x,y),$$

那么 $\|x\|=d(x,0)$ 是 V 上的范数. 反之, 范数提供的度量具有平移不变性和伸缩性. 所以 V 带有范数等价于带着具有平移不变性和伸缩性的度量.

对欧氏空间或埃氏空间, 由于度量来自长度函数, 长度函数来自内积, 所以序列的收敛性可以通过内积刻画.

定理 3.37 设 V 是有限维欧氏空间或埃氏空间, 那么 V 中的序列 $x_1,x_2,\cdots,x_k,\cdots$ 收敛到 $x\in V$ 当且仅当

$$\lim_{k\to\infty}(x-x_k|y)=0,\quad \forall y\in V.$$

证明 必要性: 当 $k\to\infty$ 时, 由假设和柯西–施瓦茨不等式得

$$|(x-x_k|y)|\leqslant\|x-x_k\|\cdot\|y\|\to 0.$$

充分性: 取 V 的标准正交基 $e_1,\cdots,e_n$. 当 $k\to\infty$ 时, 由假设和定理 3.29(3) 得

$$\|x-x_k\|^2=\sum_{i=1}^n|(x-x_k|e_i)|^2\to 0,$$

从而 $\|x-x_k\|\to 0$. □

赋范向量空间由于有加法, 故可以定义级数的收敛性, 利用范数则可以进一步定义绝对收敛性. 称向量级数 $\sum_{i=1}^{\infty} x_i$ 是**收敛的**如果序列 $x_1, x_1+x_2, \cdots, \sum_{i=1}^{k} x_i, \cdots$, 是收敛的. 称这个级数**绝对收敛**如果级数 $\sum_{i=1}^{\infty} \|x_i\|$ 是收敛的.

习 题 3.2

1. 设 f 是复向量空间 V 上的埃尔米特型. 证明:

(1) 存在 V 的基使得 f 在这个基下的矩阵是

$$\begin{pmatrix} I_s & & \\ & -I_{r-s} & \\ & & 0 \end{pmatrix},$$

其中 I_s 是 s 阶单位矩阵, $r = \operatorname{rank} f$. 这样的基称为 f 的典范基.

(2) 整数 s 仅依赖 f(或等价地, 仅依赖 q), 不依赖典范基的选取.

整数 s 和 $r-s$ 分别称为 f 及其二次型 $q(x) = f(x,x)$ 的正惯性指数和负惯性指数.

2. 设 f 是 n 维复向量空间 V 上的埃尔米特型, 它在某个基下的矩阵是 $F = f_{ij})$. 证明: f 是正定的当且仅当 F 的主子式 (定义见 1.7 节第六部分)$\Delta_1, \cdots, \Delta_n$ 都是正的; 而且雅可比方法适用.

3. 在向量空间 $M_n(\mathrm{C})$ 上定义半双线性型

$$f(X,Y) = \operatorname{tr}(X \cdot {}^t\bar{Y}).$$

(1) 证明: 它是正定的埃尔米特型, 从而是 $M_n(\mathrm{C})$ 的一个内积. 记这个内积为 $(\cdot|\cdot)$.

找出如下子空间的正交补:

(2) 迹为零的矩阵全体;

(3) 所有的上三角矩阵.

4. 把下面的向量组扩充为埃尔米特空间的正交基:

(1) $(1, 1-i, 2)$, $(2, -1+3i, 1-i)$;

(2) $(-i, 2, -2-i)$, $(4-i, -i, i)$.

5. 运用正交化方法找出埃氏空间 C^4 的子空间的一个标准正交基:

(1) $\langle(2,1,-i,1), (1,-i,2,0), (-i,0,1,-i)\rangle$;

(2) $\langle(0,1-i,2,0), (1,0,2,i), (1-i,-1,0,-i)\rangle$.

6. 在 C^3 中找出子空间 $\langle(0,1+2i,-i), (1,-1,2-i)\rangle$ 的正交补.

7. 设 V 是欧氏空间或埃氏空间, 其中的两个基 $e_1, \cdots, e_n$ 和 $v_1, \cdots, v_n$ 互称为**对偶基**, 如果

$$(e_i|v_j) = \begin{cases} 1, & 当 i = j, \\ 0, & 当 i \neq j. \end{cases}$$

证明: 对任何基 $e_1, \cdots, e_n$, 对偶基存在且唯一.

8. 对欧氏空间或埃氏空间 V, 命 A 为从基 $e_1, \cdots, e_n$ 到基 $e_1', \cdots, e_n'$ 的转移矩阵. 求从 (e_i) 的对偶基 (v_i) 到 (e_i') 的对偶基 (v_i') 的转移矩阵, 分欧氏空间和埃氏空间两种情况讨论.

9. 复正交群定义为 $CO(n) = \{A \in GL_n(\mathbb{C}) \mid {}^tA \cdot A = E\}$. 显然, $O(n) \subset CO(n)$ 且 $SO(n) \subset SCO(n)$(后二者分别是正交群和复正交群中行列式为 1 的矩阵形成的子群). 问题: 是否有 $CO(n) \subset U(n)$ 和 $SCO(n) \subset SU(n)$.

10. 对任意实数 $p \geqslant 1$, 在空间 $\mathbb{R}^n$ 上定义

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

证明:

(1) $\|\cdot\|_p$ 是 $\mathbb{R}^n$ 的一个范数;

(2) $\lim\limits_{p\to\infty} \left(\sum\limits_{i=1}^n |x_i|^p\right)^{1/p} = \max\limits_{1\leqslant i\leqslant n} |x_i|$.

于是, 范数 $\|x\|_p$ 包括了前面提到的三个范数: $\|x\|_\infty = d_1(x,0)$, $\|x\|_2 = d(x,0)$, $\|x\|_1 = d_2(x,0)$.

在闭区间 $[a,b]$ 上的连续函数全体 $C(a,b)$ 上可以类似定义范数:

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt\right)^{1/p}.$$

同样有

$$\|f\|_\infty = \max_{a\leqslant t\leqslant b} |f(t)|, \quad \|f\|_1 = \int_a^b |f(t)| dt.$$

11. 设 p 和 q 是正数, 满足等式 $\dfrac{1}{p} + \dfrac{1}{q} = 1$. 证明: 对标准欧氏空间 $\mathbb{R}^n$ 中的向量 x, y, 有赫尔德不等式

$$|(x|y)| \leqslant \|x\|_p \cdot \|y\|_q.$$

(当 $p = q = 2$ 时, 赫尔德不等式就是柯西–施瓦茨不等式.)

12. 有限维实或复向量空间 V 上的两个范数 $\|\cdot\|_*$ 和 $\|\cdot\|_\circ$ 称为**等价的**如果存在常数 c 使得对所有的 $x \in V$ 有

$$\|x\|_* \leqslant c\|x\|_\circ, \quad \|x\|_\circ \leqslant c\|x\|_*.$$

证明: 有限维实或复向量空间 V 上的任意两个范数都是等价的.

3.3 内积空间上的线性算子, I——自伴随算子

欧氏空间和埃氏空间将统一称作**内积空间**. 前面已经看到内积空间和它的对偶通过内积建立了自然的映射 (定理 3.21 和 3.2 节中第三部分的讨论). 本节将借

助内积, 建立内积空间的线性算子和内积空间上的双线性型或半双线性型的自然联系. 这个联系对讨论内积空间上的线性算子的性质是很有益处的.

本节和接下来的两节, 内积空间都是有限维的.

一 线性算子与 θ 线性型的联系 设 V 是内积空间, 其内积记作 $(\cdot|\cdot)$, $\mathcal{A}$ 是 V 上的线性算子. 定义映射

$$f_{\mathcal{A}}: V\times V\to \mathbb{K}\ \ (\mathbb{K}=\mathbb{R}\text{ 或 }\mathbb{C}),\qquad (x,y)\to(\mathcal{A}x|y). \tag{3.3.37}$$

如果 V 是欧氏空间, 那么 $f_{\mathcal{A}}$ 是双线性型, 如果 V 是埃氏空间, 那么 $f_{\mathcal{A}}$ 是半双线性型.

为方便起见, 欧氏空间上的双线性型和埃氏空间上的半双线性型将统一称为**θ线性型**. 对欧氏空间, θ 可以理解为 2; 对埃氏空间, θ 则可以理解为 3/2. 内积空间 V 上的 θ 线性型全体是 $\mathbb{K}$ 上的向量空间 (参见 1.7 节对多重线性函数的讨论), 记做 $\mathcal{L}_\theta(V,\mathbb{K})$.

取 V 的标准正交基 $e_1,\cdots,e_n$. 设 $\mathcal{A}$ 在这个基下的矩阵是 $A=(a_{ij})$. 那么

$$f_{\mathcal{A}}(e_i,e_j)=(\mathcal{A}e_i|e_j)=(a_{1i}e_1+\cdots+a_{ni}e_n|e_j)=a_{ji}.$$

所以 $f_{\mathcal{A}}$ 在基 (e_i) 下的矩阵是 ${}^tA$. 于是

$$f_{\mathcal{A}}(x,y)={}^tX\cdot{}^tA\cdot\overline{Y}={}^t(AX)\cdot\overline{Y}={}^tX\cdot\overline{{}^t\bar{A}Y}, \tag{3.3.38}$$

其中, X 和 Y 分别是向量 x 和 y 在基 (e_i) 下的坐标列向量, $\bar{A}=(\bar{a}_{ij})$ 是复共轭 (对欧氏空间不需要). 在基 (e_i) 下的矩阵为 ${}^t\bar{A}$ 的算子记作 $\mathcal{A}^*$, 等式 (3.3.36) 就可以写成

$$f_{\mathcal{A}}(x,y)=(\mathcal{A}x|y)=(x|\mathcal{A}^*y). \tag{3.3.39}$$

对这个等式所涉及的算子, θ 线性型, 有如下的结论.

定理 3.38 (1) 对 V 上的任意线性算子 $\mathcal{A}$, 存在唯一的线性算子 $\mathcal{A}^*: V\to V$, 称为 $\mathcal{A}$ 的**伴随算子**①, 使得

$$f_{\mathcal{A}}(x,y)=(\mathcal{A}x|y)=(x|\mathcal{A}^*y),\quad \forall x,y\in V.$$

在同一个标准正交基下, $\mathcal{A}^*$ 的矩阵 A^* 是 $\mathcal{A}$ 的矩阵 A 的转置共轭 (或共轭转置): $A^*={}^t\bar{A}$.

(2) 映射 $\varphi:\mathcal{L}(V)\to\mathcal{L}_\theta(V,\mathbb{K})$, $\mathcal{A}\to f_{\mathcal{A}}$, 是线性空间的同构, 其逆映射记作 $\psi:\mathcal{L}_\theta(V,\mathbb{K})\to\mathcal{L}(V)$, $f\to\mathcal{A}_f$. 有如下等式:

$$f_{\mathcal{A}}(x,y)=(\mathcal{A}x|y),\quad f(x,y)=(\mathcal{A}_fx|y),\quad \forall x,y\in V.$$

① 也称为共轭算子.

在同一个标准正交基下, $\mathcal{A}_f$ 的矩阵 A 是 f 的矩阵 F 的转置: $A = {}^tF$.

证明 (1) 前面的讨论已经证明了存在性: 设 A 在 V 的标准正交基 $e_1, \cdots, e_n$ 下的矩阵是 A, 那么在这个基下矩阵为 $A^* = \overline{{}^tA}$ 的算子 $\mathcal{A}^*$ 就满足要求. 如果对任意 $x, y \in V$ 有 $(x|\mathcal{B}y) = (x|\mathcal{A}^*y)$, 那么 $(x|(\mathcal{B} - \mathcal{A}^*)y) = 0$. 取 $x = (\mathcal{B} - \mathcal{A}^*)y$, 则有 $(x|x) = 0$, 于是 $x = 0$. 即对任意 $y \in V$, 有 $(\mathcal{B} - \mathcal{A}^*)y = 0$. 从而 $\mathcal{B} = \mathcal{A}^*$. 唯一性得证.

假设 $v_1, \cdots, v_n$ 是 V 的另一个标准正交基, 从基 (e_i) 到基 (v_i) 的转换矩阵是 C. 那么,

$$\mathcal{A}(v_i) = \mathcal{A}(e_i)C = (e_i)AC = (v_i)C^{-1}AC,$$
$$\mathcal{A}^*(v_i) = \mathcal{A}^*(e_i)C = (e_i)A^*C = (v_i)C^{-1}A^*C.$$

由于 C 是酉矩阵, 所以 $C^* = C^{-1}$, 从而 $\mathcal{A}^*$ 在基 (v_i) 下的矩阵

$$C^{-1}A^*C = C^*A^*C = (C^*AC)^* = (C^{-1}AC)^*$$

是 $\mathcal{A}$ 在基 (v_i) 下的矩阵的共轭转置.

(2) 先证 φ 是单射. 假设 $f_{\mathcal{A}} = f_{\mathcal{B}}$. 那么对任意 $x, y \in V$, 有 $(\mathcal{A}x|y) = (\mathcal{B}x|y)$, 从而 $(\mathcal{A}x - \mathcal{B}x|y) = 0$. 取 $y = \mathcal{A}x - \mathcal{B}x$, 则 $(y|y) = 0$. 于是 $y = 0$, 即 $\mathcal{A}x - \mathcal{B}x = 0$. 所以这两个算子相等.

可以直接验证 φ 是线性的.

现证映射的满性. 取定 V 的一个基, 那么 V 上的线性算子和 θ 线性型都由它在这个基下的矩阵完全确定. 由此可知 $\mathcal{L}(V)$ 和 $\mathcal{L}_\theta(V, \mathfrak{K})$ 都与 $\mathfrak{K}$ 上 n 阶方阵全体形成的空间 $M_n(\mathfrak{K})$ 同构, 其中 $n = \dim V$. 特别, 这些空间的维数都是 n^2. 根据定理 2.7, φ 的单性蕴含其是满射, 所以是线性同构.

根据定义, 对任意 $x, y \in V$, 有 $f_{\mathcal{A}}(x, y) = (\mathcal{A}x|y)$, $f(x, y) = (\mathcal{A}_f x|y)$. 设 f 在标准正交基 $e_1, \cdots, e_n$ 下矩阵为 F. 从等式

$$f(x, y) = {}^tXF\bar{Y} = {}^t({}^tFX)\bar{Y} = (\mathcal{A}_f x|y)$$

知 $\mathcal{A}_f x$ 在基 (e_i) 的坐标为 ${}^tFX$. 由此可见, 在这个基下 $\mathcal{A}_f$ 的矩阵 A 是 ${}^tF$. □

借助 θ 线性型, 容易验证映射 $\mathcal{A} \to \mathcal{A}^*$ 有如下性质:

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \quad (\lambda\mathcal{A})^* = \bar{\lambda}\mathcal{A}^*, \quad (\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*, \quad (\mathcal{A}^*)^* = \mathcal{A}. \qquad (3.3.40)$$

以第三个等式为例, $(x|(\mathcal{A}\mathcal{B})^*y) = (\mathcal{A}\mathcal{B}x|y) = (\mathcal{B}x|\mathcal{A}^*y) = (x|\mathcal{B}^*\mathcal{A}^*y)$, 所以 $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$.

这些性质和 (2.4.22) 是几乎一样的.

当 V 是欧氏空间时, V 可以自然同构于其对偶 V^*: $x \to (x|*) = \Phi_x$(见定理 3.21). 在这个等同下, 2.4 节中第四部分定义的对偶算子和这里定义的伴随算子是一致的:

$$(\mathcal{A}x|y) = (y|\mathcal{A}x) = (\Phi_y, \mathcal{A}x) = (\mathcal{A}^*\Phi_y, x) = (x|\mathcal{A}^*y) = (\Phi_{\mathcal{A}^*y}, x),$$

所以 $\mathcal{A}^*\Phi_y = \Phi_{\mathcal{A}^*y}$.

二 自伴随算子 伴随算子的价值在于它可以刻画很多重要的算子. 更确切地说, 几类重要的算子都可以通过算子本身和它的伴随算子之间的简单联系而刻画. 本节先讨论其中一类, 称为自伴随算子.

定义 3.39 称内积空间 V 上的线性算子 $\mathcal{A}$ 是**自伴随的**(或**埃尔米特的**) 如果 $\mathcal{A}^* = \mathcal{A}$. 当 $\mathfrak{K} = \mathbb{R}$(即 V 为欧氏空间) 时, 埃尔米特线性算子也称为**对称线性算子**.

称线性算子 $\mathcal{A}$ 是**反自伴随的**(**斜埃尔米特的**) 如果 $\mathcal{A}^* = -\mathcal{A}$. 当 $\mathfrak{K} = \mathbb{R}$(即 V 为欧氏空间) 时, 反自伴随算子也称为**斜对称线性算子**.

如果 $\mathcal{A}$ 是 V 上的埃尔米特线性算子, 那么

$$(\mathcal{A}x|y) = (x|\mathcal{A}y) = \overline{(\mathcal{A}y|x)}.$$

从而, 相应的 θ 线性型 $f_{\mathcal{A}}$ 满足条件

$$f_{\mathcal{A}}(x, y) = (\mathcal{A}x|y) = \overline{(\mathcal{A}y|x)} = \overline{f(y, x)},$$

即 $f_{\mathcal{A}}$ 是埃尔米特型. 反过来, $f_{\mathcal{A}}$ 是埃尔米特型意味着 $\mathcal{A}^* = \mathcal{A}$. 所以, 线性算子的埃尔米特性质等价于相应的 θ 线性型的埃尔米特性质.

我们已经知道在标准正交基下, $\mathcal{A}^*$ 的矩阵 A^* 是 $\mathcal{A}$ 的矩阵 A 的转置共轭 (即埃尔米特共轭). 所以 $\mathcal{A}$ 的埃尔米特性质等价于 ${}^t\bar{A} = A$, 即 A 是埃尔米特矩阵. 这样一来, 矩阵、线性算子及 θ 线性型的埃尔米特性质本质上是一回事. 类似地, 矩阵、线性算子及 θ 线性型的斜埃尔米特性质本质上是一回事. 我们已经证明了如下结论.

命题 3.40 (1) 每一个埃尔米特矩阵都是某个埃尔米特空间 V 上的埃尔米特算子在某个标准正交基下的矩阵. 特别, 埃尔米特矩阵的特征多项式是某个埃尔米特算子的特征多项式.

(2) 每一个斜埃尔米特矩阵都是某个埃尔米特空间 V 上的斜埃尔米特算子在某个标准正交基下的矩阵. 特别, 斜埃尔米特矩阵的特征多项式是某个斜埃尔米特算子的特征多项式.

由于 $\mathcal{A}^{**} = \mathcal{A}$, 所以 $\mathcal{A} + \mathcal{A}^*$ 是埃尔米特的, $\mathcal{A} - \mathcal{A}^*$ 是斜埃尔米特的. 而且, $\mathcal{A}$ 是埃尔米特的等价于 $i\mathcal{A}$ 是斜埃尔米特的. 下面的结论是自然的.

命题 3.41 埃氏空间上的线性算子 $\mathcal{A}$ 可以写成埃尔米特算子与斜埃尔米特算子之和.

证明 $\mathcal{A}=\frac{1}{2}(\mathcal{A}+\mathcal{A}^*)+\frac{1}{2}(\mathcal{A}-\mathcal{A}^*)=\frac{1}{2}(\mathcal{A}+\mathcal{A}^*)+\frac{i}{2}(i\mathcal{A}^*-i\mathcal{A})$. 注意 $\frac{1}{2}(i\mathcal{A}^*-i\mathcal{A})$ 是埃尔米特的. □

三 自伴随算子的典范形式 自伴随算子 (即埃尔米特算子) 的一个重要性质是如下有点让人惊讶的结论.

定理 3.42 设 $\mathcal{A}$ 是内积空间 V 的自伴随算子, 那么存在 V 的标准正交基使得 $\mathcal{A}$ 在这个基下的矩阵是对角矩阵, 而且对角线的元素全是实数 (即 $\mathcal{A}$ 的特征多项式的根全是实数).

根据命题 3.40, 这个定理可用矩阵的语言表达.

定理 3.42′ 设 A 是埃尔米特 (或实对称) 矩阵, 那么存在酉矩阵 (相应地, 正交矩阵)B 使得 $B^{-1}AB$ 是实对角矩阵.

下面证明这个定理. 先证如下结论.

引理 3.43 埃尔米特矩阵的特征多项式的根都是实数. 特别, 实对称矩阵的特征多项式的根都是实数.

证明 设 A 是一个埃尔米特矩阵, 那么 A 是某个埃尔米特空间上的埃尔米特算子 $\mathcal{A}$ 在某个标准正交基下的矩阵. 设 λ 是 A 的特征多项式的根, 则 λ 是 $\mathcal{A}$ 的特征值. 设 e 是属于 λ 的特征向量. 按定义, 有

$$\lambda(e|e)=(\lambda e|e)=(\mathcal{A}e|e)=(e|\mathcal{A}^*e)=(e|\mathcal{A}e)=(e|\lambda e)=\bar{\lambda}(e|e).$$

由于 $(e|e)\neq 0$, 所以 $\lambda=\bar{\lambda}$, 即 λ 是实数. □

推论 3.44 内积空间上的自伴随线性算子 (即埃尔米特算子) 有特征向量.

证明 自伴随算子的矩阵是埃尔米特矩阵, 由上面的引理知, 算子的特征多项式的根全是实数. 所以, 不论这个空间是埃氏空间还是欧氏空间, 算子都有特征值, 从而有特征向量. □

引理 3.45 设 $\mathcal{A}$ 是内积空间 V 的自伴随线性算子, L 是 $\mathcal{A}$ 的不变子空间, 那么 L 的正交补空间 $L^{\perp}$ 也是 $\mathcal{A}$ 的不变子空间.

证明 设 $x\in L^{\perp}$. 对一切的 $y\in L$, 因为 $\mathcal{A}y\in L$, 所以 $(\mathcal{A}y|x)=0$. 但是 $\mathcal{A}$ 是自伴随的, 所以 $(y|\mathcal{A}x)=(\mathcal{A}y|x)=0$. 即 $\mathcal{A}x\in L^{\perp}$. □

现在我们可以完成定理的证明了. 设 e_1 是 $\mathcal{A}$ 的特征向量, 其相应的特征值必为实数. 不妨要求 $\|e_1\|=1$. 一维子空间 $L=\langle e_1\rangle$ 的正交补空间 $L^{\perp}$ 是 $\mathcal{A}$ 不变子空间, 维数为 $\dim V-1$. 而且 $\mathcal{A}$ 在 $L^{\perp}$ 上的限制, 记做 $\mathcal{A}_1$, 仍是自伴随的. 利用归纳假设知存在 $L^{\perp}$ 的标准正交基 $e_2,\cdots,e_n$ 使得 $\mathcal{A}_1$ 在这个基下的矩阵是对角矩阵, 而且对角线的元素全是实数. 结果得到 V 的标准正交基 $e_1,e_2,\cdots,e_n$, 在这个基下 $\mathcal{A}$ 的矩阵是对角矩阵, 且对角线的元素全是实数. □

四 埃尔米特二次型的典范形式 前面关于埃尔米特算子的典范形式可以导出相应的埃尔米特二次型的典范形式. 根据定理 3.38, 内积空间 V 上的埃尔米特型 $f(x,y)$ 有相应的埃尔米特算子 $\mathcal{A}=\mathcal{A}_f$, 满足条件

$$f(x,y)=(\mathcal{A}x|y).$$

定理 3.42 说, 存在由 $\mathcal{A}$ 的特征向量组成的标准正交基 $e_1,\cdots,e_n$. 设 $\mathcal{A}e_i=\lambda_i e_i$, 那么 $f(e_i,e_j)=(\mathcal{A}e_i|e_j)=(\lambda_i e_i|e_j)=\delta_{ij}\lambda_i$. 如果

$$x=x_1e_1+\cdots+x_ne_n,\quad y=y_1e_1+\cdots+y_ne_n,$$

则有

$$f(x,y)=\sum_{i,j}f(e_i,e_j)x_i\overline{y_j}=\sum_{i=1}^n\lambda_i x_i\overline{y_i}.$$

取 $x=y$, 就得到如下的结论.

定理 3.46 设 V 是埃尔米特空间, $q(x)$ 是其上的埃尔米特二次型, 那么存在 V 的标准正交基, 在这个基下, $q(x)$ 有如下形式

$$q(x)=\sum_{i=1}^n\lambda_i|x_i|^2,$$

其中 λ_i 都是实数, $x_1,\cdots,x_n$ 是 x 在这个基下的坐标.

可以把定理 3.46 中的标准正交基的基向量确定的 1 维子空间看作埃尔米特二次型 $q(x)$ 的主轴, 从而这个定理有时也称作**化二次型到主轴上** (reducing quadratic forms (functions) to principal axes, reduction of quadratic forms (functions) to principal axes).

对埃尔米特二次型

$$q(x)=\sum_{i,j=1}^n a_{ij}x_i\bar{x}_j, \tag{3.3.41}$$

其矩阵 $A=(a_{ij})$ 是埃尔米特矩阵. 称坐标变换

$$x_i=\sum_{j=1}^n c_{ij}x'_j,\quad i=1,\ \cdots,\ n,\quad c_{ij}\in\mathbb{R}, \tag{3.3.42}$$

为**保距的**如果矩阵 $C=(c_{ij})$ 是酉矩阵. 定理 3.46 表明对二次型 (3.3.39), 存在保距变换使得 $q(x)$ 具有典范式

$$q(x)=\sum_{i=1}^n\lambda_i|x'_i|^2, \tag{3.3.43}$$

其中诸 λ_i 是矩阵 A 的特征多项式 $\chi_A(t)$ 的根 (按重数计算).

要求出埃尔米特二次型在保距变换下的典范式, 只需要求出系数矩阵的特征值, 一般说来, 这仅在理论上是可行的. 如果要求出把二次型化成典范式的保距变换, 那么二次型与埃尔米特空间的联系是很有用的, 也就是说, 要利用定理 3.44. 首先定义埃尔米特空间 $V=\mathbb{C}^n$ 上的埃尔米特型和线性算子

$$f(x,y)=\sum_{i,j=1}^{n}a_{ij}x_i\bar{y}_j;\quad \mathcal{A}:x\to\left(\sum_{i=1}^{n}a_{1i}x_i,\ \cdots,\ \sum_{i=1}^{n}a_{ni}x_i\right),$$

其中 $x=(x_1,\cdots,x_n)$, $y=(y_1,\cdots,y_n)\in\mathbb{C}^n$. 接下来, 计算矩阵 $A=(a_{ij})$ 的特征多项式 $\chi_A(t)=\det(tE-A)$, 并找出它的根. 然后, 对每个根 λ, 求解齐次线性方程组:

$$\begin{aligned}
(a_{11}-\lambda)\xi_1+a_{12}\xi_2+\cdots+a_{1n}\xi_n&=0,\\
a_{21}\xi_1+(a_{22}-\lambda)\xi_2+\cdots+a_{1n}\xi_n&=0,\\
\cdots\cdots&\\
a_{n1}\xi_1+a_{n2}\xi_2+\cdots+(a_{nn}-\lambda)\xi_n&=0.
\end{aligned}$$

这个方程组的解空间就是算子 $\mathcal{A}$ 的以 λ_i 为特征值的特征空间 V^{λ_i}, 其维数等于 λ_i 的代数重数 (因为 $\mathcal{A}$ 是可对角化的). 把格拉姆–施密特正交化方法应用到 V^{λ_i} 的任意一个基 (比如一个基础解系) 上, 得到 V^{λ_i} 的一个标准正交基. 由于 $\mathcal{A}$ 是自伴随算子 (即埃尔米特算子), 特征根都是实数, 于是分别属于不同特征值 λ,μ 的特征向量 u,v 是正交的:

$$(\mathcal{A}u|v)=(u|\mathcal{A}v)\Longrightarrow(\lambda u|v)=(u|\mu v)\Longrightarrow(\lambda-\mu)(u|v)=0\Rightarrow(u|v)=0.$$

所以, 把不同的特征子空间的标准正交基合起来, 就得到 V 的一个标准正交基 $e_1',\ \cdots,\ e_n'$. 设

$$e_j'=(c_{1j},\ \cdots,\ c_{nj})=c_{1j}e_1+\cdots+c_{nj}e_n,\quad j=1,\ \cdots,\ n.$$

那么, 矩阵 $C=(c_{ij})$ 就是从标准正交基 $e_1=(1,0,\cdots,0,0)$, $\cdots$, $e_n=(0,0,\cdots,0,1)$ 到标准正交基 e_1', $\cdots$, e_n' 的转换矩阵. 它是一个酉矩阵. 二次型 $q(x)$ 用向量 $x=(x_1,\cdots,x_n)=x_1e_1+\cdots+x_ne_n$ 在标准正交基 e_1', $\cdots$, e_n' 的坐标表达就是典范式 (3.3.43), 从坐标 $x_1,\cdots,x_n$ 到坐标 $x_1',\cdots,x_n'$ 的变换是保距变换, 由公式 (3.3.42) 给出.

对实数域上的二次型和一般内积空间上的 (埃尔米特) 二次型, 上面的方法也是适用的.

五 把两个二次型同时化成典范式 有时, 需要把一个空间 V 上的两个二次型同时化成典范式, 即在某个基下两个二次型都有典范式. 实平面 $\mathbb{R}^2$ 上的两个二次型 $q(x) = x_1^2 - x_2^2$ 和 $r(x) = x_1x_2$ 不能同时化成典范式, 所以这个问题并非总有解答. 但如果这两个二次型中有一个是正定的, 答案是肯定的.

定理 3.47 设 $q(x)$ 和 $r(x)$ 是 n 维实向量空间或复向量空间 V 上的二次型或埃尔米特二次型. 如果它们中有一个是正定的, 那么存在 V 的基, 在这个基下, 两个二次型都具有典范形式.

证明 不妨设 $q(x)$ 是正定的. 命 $f(x, y)$ 是 $q(x)$ 对应的 θ 线性型. 定义

$$(x|y) = f(x, y),$$

那么 V 成为内积空间. 根据定理 3.46, 二次型 $r(x)$ 在 V 的某个标准正交基 $e_1, \cdots, e_n$ 具有典范形式

$$r(x) = \sum_{i=1}^{n} \mu_i |x_i|^2.$$

在这个基下, $q(x)$ 具有如下的典范形式:

$$q(x) = f(x, x) = (x|x) = \sum_{i=1}^{n} |x_i|^2. \qquad \square$$

六 正定算子 我们已经看到内积空间 V 上的自伴随算子 $\mathcal{A}$ 在某个标准正交基 $e_1, \cdots, e_n$ 下的矩阵是对角矩阵. 如果 $\mathcal{A}$ 的特征值都是正数, 那么相应的二次型 $q(x) = (\mathcal{A}x|x)$ 就是正定的. 下面的定义是自然的.

定义 3.48 内积空间 V 上的自伴随算子 $\mathcal{A}$ 称为**正定的**(或**半正定的**) 如果二次型 $q(x) = (\mathcal{A}x|x)$ 是正定的 (或半正定的), 即: 对任意非零向量 $x \in V$, 有 $(\mathcal{A}x|x) > 0$ (或 $(\mathcal{A}x|x) \geqslant 0$).

由于自伴随算子 $\mathcal{A}$ 在某个标准正交基的矩阵是实对角矩阵 $A = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$, 所以在这个基下二次型 $q(x)$ 有典范式

$$q(x) = (\mathcal{A}x|x) = \sum_{i=1}^{n} \lambda_i |x_i|^2.$$

于是 $\mathcal{A}$ 正定 (半正定) 当且仅当它的特征值都是正实数 (非负实数). 我们将用记号 $\mathcal{A} \succ 0$ 表示 $\mathcal{A}$ 是正定的, 记号 $\mathcal{A} \succcurlyeq 0$ 表示 $\mathcal{A}$ 是半正定的.

由于正定算子的特征值都是正实数, 所以正定算子是非退化的. 正定算子的非退化性也可以直接从柯西–施瓦茨不等式得出

$$|(\mathcal{A}x|x)| \leqslant \|\mathcal{A}x\| \cdot \|x\|.$$

另外, 非退化的半正定算子一定是正定的. 正定算子可以开方, 这和正数类似.

命题 3.49 内积空间 V 上的任何 (半) 正定算子 $\mathcal{A}$ 都是另一个 (半) 正定算子 $\mathcal{B}$ 的平方: $\mathcal{A}=\mathcal{B}^2$. (半) 正定算子 $\mathcal{B}$ 是唯一的, 称为 $\mathcal{A}$ 的平方根, 记作 $\mathcal{B}=\sqrt{\mathcal{A}}$.

证明 取 V 的标准正交基, 使得 $\mathcal{A}$ 在这个基下的矩阵是 $A=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. 命 $\mathcal{B}$ 是 V 上的算子, 其在这个基下的矩阵是 $B=\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})$. 那么 $\mathcal{B}$ 是 (半) 正定算子且 $\mathcal{A}=\mathcal{B}^2$. 存在性得证.

假设另有 (半) 正定算子 $\mathcal{C}$ 使得 $\mathcal{A}=\mathcal{C}^2$. 那么 $\mathcal{C}$ 的特征值 (按重数计算) 是 $\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n}$. 于是, 对 $\mathcal{A}$ 的任意特征值 λ, 有

$$\{v\in V\,|\,\mathcal{A}v=\lambda v\}=\{v\in V\,|\,\mathcal{B}v=\sqrt{\lambda}v\}=\{v\in V\,|\,\mathcal{C}v=\sqrt{\lambda}v\}.$$

这迫使 $\mathcal{B}=\mathcal{C}$, 因为 $\mathcal{B}$ 和 $\mathcal{C}$ 都是特征子空间的直和. □

命题 3.50 设 $\mathcal{C}$ 是内积空间 V 上的 (不) 可逆算子. 那么 $\mathcal{C}\mathcal{C}^*$ 和 $\mathcal{C}^*\mathcal{C}$ 都是 (半) 正定算子.

证明 首先 $\mathcal{C}\mathcal{C}^*$ 是自伴随的: $(\mathcal{C}\mathcal{C}^*)^*=(\mathcal{C}^*)^*\mathcal{C}^*=\mathcal{C}\mathcal{C}^*$. 于是

$$(\mathcal{C}\mathcal{C}^*x|x)=(\mathcal{C}^*x|\mathcal{C}^*x)\geqslant 0,\quad \forall x\in V.$$

这意味着 $\mathcal{C}\mathcal{C}^*$ 是半正定的. 当 $\mathcal{C}$ 可逆时, 对 V 中的非零向量 x, 有 $\mathcal{C}^*x\neq 0$, 于是 $(\mathcal{C}\mathcal{C}^*x|x)=(\mathcal{C}^*x|\mathcal{C}^*x)>0$. 从而 $\mathcal{C}\mathcal{C}^*$ 是正定的. 类似可知 $\mathcal{C}^*\mathcal{C}$ 是 (半) 正定算子. □

对半正定算子, 上面两个结论合在一起就是如下结论.

定理 3.51 对内积空间 V 上的线性算子 $\mathcal{A}$, 下列条件等价:

(1) $\mathcal{A}=\mathcal{B}^2$ 且 $\mathcal{B}^*=\mathcal{B}$;

(2) $\mathcal{A}=\mathcal{C}\mathcal{C}^*$;

(3) $\mathcal{A}$ 自伴随且 $(\mathcal{A}x|x)\geqslant 0\quad \forall x\in V$.

可以看出定理中的性质 (1), (2) 与非负实数在复数域中的性质是类似的.

习 题 3.3

1. 设 V 是内积空间, $\mathcal{A}$ 是 V 上的线性算子. 证明: 如果 $U\subset V$ 是 $\mathcal{A}$ 的不变子空间, 那么 U 的正交补是 $\mathcal{A}^*$ 的不变子空间.

2. 证明: 线性算子 $\mathcal{A}$ 的核与像分别是 $\mathcal{A}^*$ 的像与核的正交补.

3. 设 $\mu(t)=a_0+a_1t+\cdots+a_mt^m$ 是线性算子 $\mathcal{A}$ 的极小多项式. 证明: $\bar{\mu}(t)=\bar{a}_0+\bar{a}_1t+\cdots+\bar{a}_mt^m$(小横杆表示复共轭) 是线性算子 $\mathcal{A}^*$ 的极小多项式.

4. 证明: 如果 $\mathcal{A}$ 和 $\mathcal{B}$ 都是向量空间 V 上的自伴随线性算子, 则线性变换

$$\mathcal{A}+\mathcal{B},\quad i(\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A})$$

也是自伴随的.

5. 设 $\mathcal{A}$ 是内积空间 V 的沿着子空间 W 到子空间 U 的投影. 证明: $\mathcal{A}$ 是自伴随的当且仅当 U 和 W 是正交的.

6. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是域 K 上的向量空间的算子. 证明: 如果它们交换且可对角化, 那么, 它们可同时对角化, 也就是说, 存在 V 的一个基, 基中的向量既是 $\mathcal{A}$ 的特征向量, 也是 $\mathcal{B}$ 的特征向量.

7. 证明: 如果 $\mathcal{A}$ 和 $\mathcal{B}$ 是正定算子且 $\mathcal{AB}=\mathcal{BA}$, 那么 $\mathcal{AB}$ 也是正定的.

8. 假设 A 是可逆斜对称实矩阵. 证明: $-A^2$ 是对称的正定矩阵. 特别, 斜对称实矩阵的非零特征值必为纯虚数.

9. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是两个半正定的自伴随算子, 其中一个是可逆的. 证明: $\mathcal{AB}$ 的特征值都是非负实数.

10. 假设内积空间上线性算子在某个标准正交基下的矩阵如下:

$$(1)\begin{pmatrix}2&1\\1&2\end{pmatrix},\quad (2)\begin{pmatrix}11&2&-8\\2&2&10\\-8&10&5\end{pmatrix},\quad (3)\begin{pmatrix}5&-1&-1\\-1&5&-1\\-1&-1&5\end{pmatrix},$$

$$(4)\begin{pmatrix}3&2+2i\\2-2i&1\end{pmatrix},\quad (5)\begin{pmatrix}3&-i\\i&3\end{pmatrix}.$$

求一个标准正交基使得线性算子在这个基下的矩阵是对角的.

11. 三维欧氏空间 V 上的 $\mathcal{A}$ 在某个标准正交基下的矩阵是

$$\begin{pmatrix}13&14&4\\14&24&18\\4&18&29\end{pmatrix}.$$

求 $\mathcal{A}$ 的平方根 $\mathcal{B}$(即 $\mathcal{B}$ 正定且 $\mathcal{B}^2=\mathcal{A}$) 在这个基下的矩阵.

12. 找出把二次型化成典范式的一个保距变换:

(1) $6x_1^2+5x_2^2+7x_3^2-4x_1x_2+4x_1x_3$;

(2) $9x_1^2+5x_2^2+5x_3^2+8x_4^2+8x_2x_3-4x_2x_4+4x_3x_4$;

(3) $5|x_1|^2+i\sqrt{3}x_1\bar{x}_2-i\sqrt{3}x_2\bar{x}_1+6|x_2|^2$;

(4) $2|x_1|^2+2|x_2|^2+2|x_3|^2-2ix_1\bar{x}_2+2i3x_2\bar{x}_1+2ix_2\bar{x}_3+2ix_3\bar{x}_2$.

13. 给定欧氏空间上的二次型 $f(x)=\sum\limits_{i=1}^{n}\lambda_i x_i^2$. 证明

$$\max\{|\lambda_1|,\cdots,|\lambda_r|\}=\max_{|x|=1}|f(x)|.$$

设 $q(x)$ 是欧氏空间 V 上的二次型. 问: 在单位球面 $S=\{x\in V\mid \|x\|=1\}$ 上哪些点处二次型 q 能达到极大和极小. 更一般地, 单位球面上哪些点是 q 的静止点 (stationary point), 即 q 在该点的各个方向的导数都是 0? 证明下述论断正确:

二次型 $q(x)$ 在单位球面上的静止点恰是二次型 $q(x)=(\mathcal{F}x|x)$ 确定的对称算子 $\mathcal{F}$ 的单位特征向量.

特别, 二次型 $q(x)$ 在单位球面上的极大值就是 $\mathcal{F}$ 的特征值中的最大者, 极小值就是 $\mathcal{F}$ 的特征值中的最小者. 也就是说, 通过保距变换把 $q(x)$ 化成典范式, 那么典范式的系数中的最大者 (最小者) 就是二次型 $q(x)$ 在单位球面上的极大值 (极小值).

14. 如往常, E 是 n 阶单位矩阵, E_{ij} 是在 (i,j) 处的值为 1, 其他地方的值为 0 的 n 阶方阵. 验证: 矩阵族

$$\mathfrak{S} = \{E_{ij} \mid 1 \leqslant i \leqslant [n/2], [n/2]+2 \leqslant j \leqslant n\} \cup \{E\}$$

含有 $[n^2/4]+1$ 个元素, 它们线性无关. 而且, 对 $E_{ij}, E_{kl} \in \mathfrak{S}$, 有 $E_{ij}E_{kl} = E_{kl}E_{ij} = 0$. (对实数 a, 记号 $[a]$ 表示比 a 小的整数中的最大者, 即 $[a]$ 是整数, 且 $0 \leqslant a-[a] < 1$. 如 $[-1.2]=-2$, $[3.1]=1$, $[\pm 5]=\pm 5$.)

15. [I. Schur, 1905] 在复代数 $M_n(\mathbb{C})$ 中交换子代数的最大维数是 $[n^2/4]+1$. 实际上, 可以证明更强的结论: 在 $M_n(\mathbb{C})$ 中如果一个向量组中的向量两两交换, 那么这个向量组的秩不超过 $[n^2/4]+1$.

容易看出, $\mathbb{C}$ 可以用任何域代替.

16. 设 V 是偶数 $2m$ 维欧氏空间, $f(x,y)$ 是 V 上的非退化斜对称双线性型. 证明: 可将 V 分解成两个 m 维子空间的直和 $V=V_1 \oplus V_2$, 并找到一个对称的非退化线性算子 $\mathcal{A}: V \to V$ 使得

$$f(x,y) = (x_1|\mathcal{A}y_2) - (x_2|\mathcal{A}y_1),$$

此处 $x=x_1+x_2$, $y=y_1+y_2$, $x_i, y_i \in V_i$, $i=1,2$.

3.4 内积空间上的线性算子, II——保距算子

一 内积空间上的算子中最重要的部分很可能是保距算子.

定义 3.52 称内积空间 V 上的线性算子 $\mathcal{A}$ 是**保距的**如果 $\|\mathcal{A}x\| = \|x\|$ 对所有的 $x \in V$.

假设 $\mathcal{A}$ 保距. 对 $x,y \in V$, 我们有 $\|\mathcal{A}x-\mathcal{A}y\| = \|\mathcal{A}(x-y)\| = \|x-y\|$, 即 $\mathcal{A}x$, $\mathcal{A}y$ 之间的距离和 x, y 之间的距离相等. 这解释了保距的含义.

命题 3.53 内积空间 V 上的线性算子 $\mathcal{A}$ 是保距的当且仅当它保持内积: $(\mathcal{A}x|\mathcal{A}y) = (x|y)$, 从而当且仅当 $\mathcal{A}^*\mathcal{A} = \mathcal{E}$.

证明 如果 $\mathcal{A}$ 保持内积, 那么 $(\mathcal{A}x|\mathcal{A}x) = (x|x)$, 即 $\|\mathcal{A}x\| = \|x\|$. 反过来, 假设 $\mathcal{A}$ 是保距的, 即 $\|\mathcal{A}z\| = \|z\|$, 对所有的 $z \in V$. 对埃氏空间, 有

$$2(x|y) = \|x+y\|^2 + i\|x+iy\|^2 - (1+i)(\|x\|^2+\|y\|^2),$$

所以

$$\begin{aligned}2(\mathcal{A}x|\mathcal{A}y) &= \|\mathcal{A}x+\mathcal{A}y\|^2 + i\|\mathcal{A}x+i\mathcal{A}y)\|^2 - (1+i)(\|\mathcal{A}x\|^2+\|\mathcal{A}y\|^2)\\ &= \|\mathcal{A}(x+y)\|^2 + i\|\mathcal{A}(x+iy)\|^2 - (1+i)(\|\mathcal{A}x\|^2+\|\mathcal{A}y\|^2)\\ &= \|x+y\|^2 + i\|x+iy\|^2 - (1+i)(\|x\|^2+\|y\|^2)\\ &= 2(x|y).\end{aligned}$$

即 $\mathcal{A}$ 保持内积.

对欧氏空间, 有

$$2(x|y) = \|x+y\|^2 - (\|x\|^2+\|y\|^2).$$

所以

$$\begin{aligned}2(\mathcal{A}x|\mathcal{A}y) &= \|\mathcal{A}x+\mathcal{A}y\|^2 - (\|\mathcal{A}x\|^2+\|\mathcal{A}y\|^2)\\ &= \|\mathcal{A}(x+y)\|^2 - (\|\mathcal{A}x\|^2+\|\mathcal{A}y\|^2)\\ &= \|x+y\|^2 - (\|x\|^2+\|y\|^2) = 2(x|y).\end{aligned}$$

同样 $\mathcal{A}$ 保持内积.

由于 $(\mathcal{A}x|\mathcal{A}y) = (x|\mathcal{A}^*\mathcal{A}y)$, 所以 $\mathcal{A}$ 保持内积当且仅当 $(x|\mathcal{A}^*\mathcal{A}y) = (x|y)$, 即 $(x|(\mathcal{A}^*\mathcal{A}-\mathcal{E})y) = 0$. 取 $x = (\mathcal{A}^*\mathcal{A}-\mathcal{E})y$, 则有 $\|x\|^2 = 0$, 所以 $(\mathcal{A}^*\mathcal{A}-\mathcal{E})y = 0$. 从而 $\mathcal{A}^*\mathcal{A} = \mathcal{E}$. □

保距算子对内积空间的几何研究是很重要的, 对埃尔米特空间上的保距算子还有另一个术语.

定义 3.54 埃氏空间上的线性算子 $\mathcal{A}$ 称为**酉算子**如果 $\mathcal{A}^*\mathcal{A} = \mathcal{E}$.

定义 3.55 欧氏空间上的线性算子 $\mathcal{A}$ 称为**正交算子**如果 $\mathcal{A}^*\mathcal{A} = \mathcal{E}$.

定理 3.36 和命题 3.51 表明, 酉算子在标准正交基下的矩阵 A 有性质 ${}^t\bar{A}A = E$, 所以是酉矩阵. 反过来, 取定一个标准正交基 (e_i), 那么酉矩阵 A 对应的算子: $(e_i) \to (e_i)A$ 是酉算子. 于是 n 维埃尔米特空间上的酉算子全体形成的群与酉群 $U(n)$ 是同构的. 对欧氏空间而言, 正交算子在标准正交基下的矩阵是正交矩阵, 于是 n 维欧氏空间 V 上的正交算子全体形成的群 $O(V)$ 与正交群 $O(n)$ 是同构的. 我们已经知道 $O(n)$ 是 $U(n)$ 的子群.

二 保距算子的典范形式 内积空间的保距算子, 即酉算子或正交算子, 同样具有典范形式. 首先, 有如下结论.

引理 3.56 保距算子的特征值的模等于 1(实的情形为 ± 1).

证明 设 $\mathcal{A}: V \to V$ 是保距算子, $e \in V$ 是属于特征值 λ 的特征向量. 那么

$$(\mathcal{A}e|\mathcal{A}e) = (\lambda e|\lambda e) = \lambda\bar{\lambda}(e|e).$$

根据命题 3.53, 有

$$(\mathcal{A}e|\mathcal{A}e) = (e|e).$$

所以, $\lambda\bar{\lambda} = 1$. 即 $\bar{\lambda} = 1$. 在实的情形 (欧氏空间), λ 是实数, 故只能是 1 或 -1. □

引理 3.57 保距算子的不变子空间的正交补也是不变子空间.

证明 设 $\mathcal{A}: V \to V$ 是保距算子, $U \subset V$ 是 $\mathcal{A}$ 的不变子空间. 根据定义, U 的正交补是

$$U^{\perp} = \{v \in V \mid (u|v) = 0, \forall u \in U\}.$$

由于 $\mathcal{A}$ 在 U 上的限制 $\mathcal{A}_U$ 仍是保距的, 所以 $\mathcal{A}_U$ 是可逆的. 于是对任意的 $u \in U$, 存在 $u' \in U$ 使得 $\mathcal{A}u' = u$. 如果 $v \in U^{\perp}$, 我们有

$$(u|\mathcal{A}v) = (\mathcal{A}u'|\mathcal{A}v) = (u'|v) = 0.$$

也就是说, $\mathcal{A}v \in U^{\perp}$. □

1. 酉算子情形 结论和定理 3.42 类似

定理 3.58 设 $\mathcal{A}$ 是埃尔米特空间 V 上的酉算子 (即保距算子), 那么存在 V 的标准正交基使得 $\mathcal{A}$ 在这个基下的矩阵是对角矩阵, 而且对角线的元素模都是 1.

酉算子在标准正交基下的矩阵是酉矩阵 (参见定理 3.38 和命题 3.53). 由于标准正交基之间的转换矩阵是酉矩阵, 所以定理 3.56 用矩阵的语言表述就是下面的结论.

定理 3.58′ 设 A 是酉矩阵, 那么存在酉矩阵 B 使得

$$B^{-1}AB = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$$

是对角矩阵, 且 $|\lambda_i| = 1, \quad i = 1, \cdots, n$.

证明 取 $\mathcal{A}$ 的单位特征向量 e_1. 它存在, 因为基域 $\mathbb{C}$ 是代数闭域. 由引理 3.55, $\langle e_1\rangle^{\perp} = U$ 是 $\mathcal{A}$ 的不变子空间. 由于 V 是 $\langle e_1\rangle$ 和 $\langle e_1\rangle^{\perp}$ 的直和 (参见 3.2 节中第三部分), 对 V 的维数作归纳法即知: 存在 V 的标准正交基使得 $\mathcal{A}$ 在这个基下的矩阵是对角矩阵. 根据引理 3.56, 对角线的元素模都是 1. □

利用欧拉公式 $e^{i\varphi} = \cos\varphi + i\sin\varphi$, 定理 3.58′ 中的对角矩阵 $\mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$ 就可以记作

$$\begin{pmatrix} e^{i\varphi_1} & & & \\ & e^{i\varphi_2} & & \\ & & \ddots & \\ & & & e^{i\varphi_n} \end{pmatrix}.$$

2. 正交算子情形

这种情形比酉算子稍微复杂一些, 因为基域 $\mathbb{R}$ 不是代数闭的. 不过, 还不算糟糕, 因为引理 3.56 和引理 3.57 是可用的, 且实向量空间上的算子有 1 维或 2 维的不变子空间. 于是, 对 n 维欧氏空间 V 上的正交算子 $\mathcal{A}$, 借助引理 3.57, 可以把 V 分解成互相正交的 1 维或 2 维的不变子空间的直和:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m. \tag{3.4.44}$$

算子 $\mathcal{A}$ 在每个不变子空间上的限制都是正交算子. 对每个 V_i, 取一个标准正交基, 然后都合在一起, 就得到 V 的一个标准正交基. 如果在分解式 (3.4.42) 中要求直和项的个数 m 是最大的, 那么每个 V_i 就不能分解成更小的不变子空间的直和. 这时, 所得到 V 的标准正交基就是正交算子 $\mathcal{A}$ 的**典范基**.

现在讨论算子 $\mathcal{A}$ 在典范基下的矩阵. 如果 A_i 是算子 $\mathcal{A}|_{V_i}$ 在 V_i 上的所取的标准正交基下的矩阵, 那么 $\mathcal{A}$ 在所得的标准正交基下的矩阵是

$$A = A_1 \dotplus \cdots \dotplus A_m = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}.$$

于是, 问题归结为讨论 1 维和 2 维欧氏空间 V 上的没有非平凡不变子空间的正交算子 $\mathcal{A}$ 的矩阵. 如果 V 是 1 维的, 那么 V 中的非零向量是 $\mathcal{A}$ 的特征向量, 相应的特征值只能是 1 或 -1(引理 3.56). 如果 V 是 2 维的, e_1, e_2 是标准正交基, 那么 $\mathcal{A}$ 在这个基下的矩阵

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R},$$

是正交矩阵.

如果 $\det A = ad - bc = -1$, 那么, 特征多项式 $\chi_{\mathcal{A}}(t) = t^2 - (a+d)t - 1$ 有两个实根. 从而, 算子 $\mathcal{A}$ 有特征向量, 这与 $\mathcal{A}$ 没有非平凡的不变子空间矛盾. 所以有 $\det A = 1$. 此时 A 的伴随矩阵就是 A 的逆:

$$A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

另一方面, A 是正交矩阵意味着

$$A^{-1} = {}^tA = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

比较 $\mathcal{A}^{-1}$ 的两个表达式, 得

$$A=\begin{pmatrix} a & -c \\ c & a \end{pmatrix},\quad a^2+c^2=1.$$

于是, 存在角 φ 使得 $a=\cos\varphi$, $c=\sin\varphi$. 这样一来, 有

$$A=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

也就是说, 正交算子 $\mathcal{A}$ 是平面 V 上的一个旋转.

把上面的讨论应用到分解式 (3.4.42). 适当安排直和中子空间的顺序, 可以假设直和项 $V_1,\cdots,V_r$ 都是不可分解的 2 维不变子空间, 剩下的直和项是 1 维的不变子空间, 并把特征值为 -1 的 1 维不变子空间排在前面. 又设 $\mathcal{A}$ 在 2 维不变子空间 V_i 上的旋转角为 φ_i. 这样处理后, 矩阵 A 就是下面定理中的形式.

定理 3.59　设 $\mathcal{A}$ 是 n 维欧氏空间上的正交算子, 那么, 存在 V 的标准正交基使得算子 $\mathcal{A}$ 在这个基下的矩阵是

$$\begin{pmatrix}
\cos\varphi_1 & -\sin\varphi_1 & & & & & \\
\sin\varphi_1 & \cos\varphi_1 & & & & & \\
 & & \ddots & & & & \\
 & & & \cos\varphi_r & -\sin\varphi_r & & \\
 & & & \sin\varphi_r & \cos\varphi_r & & \\
 & & & & & -I_k & \\
 & & & & & & I_l
\end{pmatrix},\quad 2r+k+l=n,$$

其中 I_k 和 I_l 分别是 k 阶和 l 阶单位矩阵.

三　极化分解　内积空间上的半正定算子有点类似于非负实数, 酉算子和正交算子则类似于模为 1 的复数. 有意思的是, 内积空间的算子中也有一个公式类似于复数的三角形式: $z=|z|e^{i\varphi}=\sqrt{z\bar{z}}e^{i\varphi}$.

定理 3.60　内积空间 V 上的非退化算子 $\mathcal{A}$ 可以分解成

$$\mathcal{A}=\mathcal{P}\mathcal{Q},\tag{3.4.45}$$

其中 $\mathcal{P}$ 是正定算子, $\mathcal{Q}$ 是保距算子. 这个分解式是唯一的, 称为 $\mathcal{A}$ 的**极化分解**.

证明　正定算子 $\mathcal{P}$ 的选择范围很小, 因为 $\mathcal{A}$ 能产生的正定算子是 $\mathcal{A}\mathcal{A}^*$ 和 $\mathcal{A}^*\mathcal{A}$. 命 $\mathcal{P}$ 为 $\mathcal{A}\mathcal{A}^*$ 的平方根, $\mathcal{Q}=\mathcal{P}^{-1}\mathcal{A}$, 我们就得到分解 (3.4.43). 需要证明 $\mathcal{Q}$ 是保距算子. 这可以直接验证:

$$\mathcal{Q}\mathcal{Q}^*=\mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*(\mathcal{P}^{-1})^*=\mathcal{P}^{-1}\mathcal{A}\mathcal{A}^*\mathcal{P}^{-1}=\mathcal{E}.$$

这里我们用到了等式 $\mathcal{P}^{-1}=(\mathcal{P}^{-1})^*$, 它从等式 $\mathcal{P}=\mathcal{P}^*$ 推出:

$$\mathcal{P}\mathcal{P}^{-1}=\mathcal{E}=\mathcal{E}^*=(\mathcal{P}^{-1})^*\mathcal{P}^*\Rightarrow(\mathcal{P}^{-1})^*=(\mathcal{P}^*)^{-1}=\mathcal{P}^{-1}.$$

如果 $\mathcal{P}_1\mathcal{Q}_1$ 是 $\mathcal{A}$ 的另一个极化分解, 那么 $\mathcal{P}\mathcal{Q}=\mathcal{P}_1\mathcal{Q}_1$. 于是 $\mathcal{Q}^*\mathcal{P}=\mathcal{Q}_1^*\mathcal{P}_1$. 从而 $\mathcal{P}\mathcal{Q}\cdot\mathcal{Q}^*\mathcal{P}=\mathcal{P}_1\mathcal{Q}_1\cdot\mathcal{Q}_1^*\mathcal{P}_1$. 所以 $\mathcal{P}^2=\mathcal{P}_1^2$. 正定算子的平方根是唯一的, 因此得 $\mathcal{P}=\mathcal{P}_1$. 这也意味着 $\mathcal{Q}=\mathcal{Q}_1$. 唯一性得证. □

注 显然

$$\mathcal{A}=\mathcal{P}\mathcal{Q}=\mathcal{Q}\mathcal{Q}^{-1}\mathcal{P}\mathcal{Q}.$$

由此得到 $\mathcal{A}$ 的分解

$$\mathcal{A}=\mathcal{Q}\mathcal{P}_1,$$

其中 $\mathcal{Q}$ 是保距的, $\mathcal{P}_1=\mathcal{Q}^*\mathcal{P}\mathcal{Q}$ 是正定算子. 其实, $\mathcal{P}_1$ 是 $\mathcal{A}^*\mathcal{A}$ 的平方根: $\mathcal{P}_1^2=\mathcal{Q}^*\mathcal{P}^2\mathcal{Q}=\mathcal{Q}^*\mathcal{P}\cdot\mathcal{P}\mathcal{Q}=\mathcal{A}^*\mathcal{A}$.

对退化的算子, 极化分解 (3.4.43)($\mathcal{P}$ 是半正定的, 且不要求唯一性) 也是成立的. 与复数的三角形式不同, 算子的极化分解中的两个因子一般是不交换的. 例如取

$$P=\begin{pmatrix}1&0\\0&2\end{pmatrix},\quad Q=\begin{pmatrix}\dfrac{1}{2}&-\dfrac{\sqrt{3}}{2}\\\dfrac{\sqrt{3}}{2}&\dfrac{1}{2}\end{pmatrix},$$

它们分别是正定算子 $\mathcal{P}$ 和保距算子 $\mathcal{Q}$ 在某个标准正交基下的矩阵. 那么 $\mathcal{P}\mathcal{Q}\neq\mathcal{Q}\mathcal{P}$.

如果算子 $\mathcal{A}$ 的极化分解中的因子是交换的, 则有 $\mathcal{A}=\mathcal{P}\mathcal{Q}=\mathcal{Q}\mathcal{P}$. 从而

$$\mathcal{A}\mathcal{A}^*=\mathcal{P}\mathcal{Q}\mathcal{Q}^*\mathcal{P}^*=\mathcal{P}^2=\mathcal{P}\mathcal{Q}^*\mathcal{Q}\mathcal{P}=(\mathcal{Q}\mathcal{P})^*\mathcal{Q}\mathcal{P}=\mathcal{A}^*\mathcal{A},$$

即, $\mathcal{A}$ 与它的伴随算子交换.

反之,

$$\mathcal{A}\mathcal{A}^*=\mathcal{A}^*\mathcal{A}\Rightarrow\mathcal{P}^2=\mathcal{P}\mathcal{Q}\mathcal{Q}^*\mathcal{P}^*=\mathcal{Q}^*\mathcal{P}^*\mathcal{P}\mathcal{Q}=\mathcal{Q}^{-1}\mathcal{P}^2\mathcal{Q}=(\mathcal{Q}^{-1}\mathcal{P}\mathcal{Q})^2.$$

由命题 3.49 知 $\mathcal{P}=\mathcal{Q}^{-1}\mathcal{P}\mathcal{Q}$, 即 $\mathcal{P}$ 与 $\mathcal{Q}$ 是交换的. 于是我们有结论: 算子的极化分解中的因子是交换的当且仅当算子与它的伴随算子是交换的.

习 题 3.4

1. 证明: 如果内积空间 V 中的两个向量 x 和 y 的长度相同, 那么存在保距变换把 x 映到 y.

2. 设 $x_1, \cdots, x_k$ 和 $y_1, \cdots, y_k$ 是内积空间 V 中的两组向量. 证明: 存在保距算子把 x_i 映到 y_i $(i=1,2,\cdots,k)$ 的充要条件是 $(x_i|x_j)=(y_i|y_j)$, $1\leqslant i,j\leqslant k$.

3. 设 w 是内积空间 V 中的非零向量. 对任何向量 $x\in V$, 命

$$\mathcal{R}_w(x)=x-2\frac{(x|w)}{(w|w)}w.$$

证明:

(1) $\mathcal{R}_w(w)=-w$;

(2) $\mathcal{R}_w(y)=y$ 如果 $y\in\langle w\rangle^{\perp}$;

(3) $\mathcal{R}_w$ 是保距算子.

(4) 对 V 中线性无关的向量 x,y, 存在向量 v 使得 $\mathcal{R}_v(x)=\dfrac{\|x\|}{\|y\|}y$ 当且仅当 $(x|y)$ 是实数.

4. 证明: 如果内积空间 V 上的线性算子 $\mathcal{A}$ 具有下列三个性质中的任何两个:

(a) 自伴随; (b) 保距; (c) 对合变换 (即 $\mathcal{A}^2=\mathcal{E}$),

则它也具有第三个性质. 求出具有这三个性质的变换全体.

5. 酉矩阵 A 何时能表成换位子 (commutator)$A=XYX^{-1}Y^{-1}$ 的形式, 其中 X, Y 也是酉矩阵.

6. 假设保距算子在某个标准正交基下的矩阵如下, 找出这个算子的一个典范基和算子在这个基下的矩阵:

$$(1)\ \frac{1}{3}\begin{pmatrix}2&2&-1\\2&-1&2\\-1&2&2\end{pmatrix},\qquad (2)\ \frac{1}{2}\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix},$$

$$(3)\ \frac{1}{\sqrt{3}}\begin{pmatrix}1+i&1\\-1&1-i\end{pmatrix},\qquad (4)\ \frac{1}{9}\begin{pmatrix}4+3i&4i&-6-2i\\-4i&4-3i&-2-6i\\6+2i&-2-6i&1\end{pmatrix}.$$

7. 设 $\mathcal{A}$ 是酉算子. 证明: 如果 $\mathcal{A}-\mathcal{E}$ 是可逆的, 那么 $i(\mathcal{A}-\mathcal{E})^{-1}(\mathcal{A}+\mathcal{E})$ 是自伴随的.

8. 证明: 如果酉空间上的线性算子 $\mathcal{A}$ 是自伴随的, 那么 $(\mathcal{A}-i\mathcal{E})^{-1}(\mathcal{A}+i\mathcal{E})$ 是酉算子.

9. 已经知道 (见习题 3.2 中第 3 题), 带着内积 $(X|Y)=\mathrm{tr}\,(X\,{}^t\bar{Y})$ 下, $M_n(\mathbb{C})$ 成为埃尔米特空间. 证明:

(1) $M_n(\mathbb{C})$ 中的酉矩阵的长度是 $\sqrt{n}$;

(2) $M_n(\mathbb{C})$ 上的算子 $X\to AX$ 的伴随算子是 $X\to {}^t\bar{A}X$;

(3) 算子 $X\to AX$ 是酉算子当且仅当 A 是酉矩阵.

10. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是正定自伴随算子, $\mathcal{C}$ 是保距算子. 证明: 如果 $\mathcal{A}=\mathcal{BC}$, 那么 $\mathcal{C}=\mathcal{E}$.

11. 假设算子 $\mathcal{A}$ 在某个标准正交基下的矩阵如下, 找出 $\mathcal{A}$ 的极化分解:

$$(1)\ \begin{pmatrix}2&-1\\2&1\end{pmatrix},\qquad (2)\ \begin{pmatrix}4&-2&2\\4&4&-1\\-2&4&2\end{pmatrix}.$$

3.5　内积空间上的线性算子, III——正规算子

一　算子 $\mathcal{A}$ 的极化分解中因子的交换性等价于 $\mathcal{A}$ 与它的伴随算子 $\mathcal{A}^*$ 是交换的. 埃尔米特算子和保距算子都满足这个条件. 可见, 满足这个条件的算子是有趣的.

定义 3.61　内积空间 V 上的线性算子 $\mathcal{A}$ 称为**正规的**(normal), 如果它与其伴随算子交换:

$$\mathcal{A}\cdot\mathcal{A}^* = \mathcal{A}^*\cdot\mathcal{A}. \tag{3.5.46}$$

正规算子在标准正交基下的矩阵称为**正规矩阵**.

很容易举出埃尔米特空间上正规算子的例子, 它既不是埃尔米特算子也不是保距算子, 比方说, 在某个标准正交基的矩阵为 $\operatorname{diag}(2, i, 0, \cdots, 0)$ 的算子. 正规算子的重要性质是它们具有不同特征值的特征向量是正交的 (习题 1). 这一点对无限维内积空间 (如希尔伯特空间) 的正规算子仍然成立, 所有它们在无限维内积空间中起重要的作用.

定理 3.62　设 V 是埃尔米特空间, $\mathcal{A}: V\to V$ 是线性算子. 那么, 下述条件等价:

(1) $\mathcal{A}$ 在 V 的某个标准正交基下的矩阵是对角矩阵;

(2) $\mathcal{A}$ 是正规算子.

在证明这个定理之前, 我们先看正规算子的一些简单性质. 算子 $\mathcal{A}$ 的正规性意味着对任意的 $x\in V$ 有

$$\|\mathcal{A}x\|^2 = (\mathcal{A}x|\mathcal{A}x) = (\mathcal{A}^*\mathcal{A}x|x) = (\mathcal{A}\mathcal{A}^*x|x) = (\mathcal{A}^*x|\mathcal{A}^*x) = \|\mathcal{A}^*x\|^2. \tag{3.5.47}$$

反之, 如果对任意 $x\in V$ 有 $\|\mathcal{A}x\| = \|\mathcal{A}^*x\|$, 由上式知 $(\mathcal{B}x|x)=0$, 其中 $\mathcal{B} = \mathcal{A}^*\mathcal{A} - \mathcal{A}\mathcal{A}^*$. 于是, 对任意的 $x, y\in V$, 有

$$0 = (\mathcal{B}(x+y)|x+y) = (\mathcal{B}x|x) + (\mathcal{B}x|\mathcal{B}y) + (\mathcal{B}y|x) + (\mathcal{B}y|y) = (\mathcal{B}x|y) + (\mathcal{B}y|x),$$

$$0 = (\mathcal{B}(ix+y)|ix+y) = (\mathcal{B}x|x) + i(\mathcal{B}x|y) - i(\mathcal{B}y|x) + (\mathcal{B}y|y) = i(\mathcal{B}x|y) - i(\mathcal{B}y|x).$$

由此可见 $(\mathcal{B}x|y)=0, \forall x, y\in V$. 所以 $\mathcal{B}=\mathcal{O}$, 即 $\mathcal{A}^*\mathcal{A} = \mathcal{A}\mathcal{A}^*$. 从而条件 (3.5.46) 和条件 (3.5.47) 是等价的.

根据公式 (3.3.38), 有

$$(\lambda\mathcal{E})^* = \bar{\lambda}\mathcal{E}, \quad (\mathcal{A}-\lambda\mathcal{E})^* = \mathcal{A}^* - \bar{\lambda}\mathcal{E}.$$

由此可见, 如果 $\mathcal{A}$ 是正规的, 那么 $\mathcal{A}-\lambda\mathcal{E}$ 也是正规的, 于是

$$\|\mathcal{A}x - \lambda x\| = \|\mathcal{A}^*x - \bar{\lambda}x\|.$$

由此可得

$$\mathcal{A}x = \lambda x \iff \mathcal{A}^* x = \bar{\lambda} x. \tag{3.5.48}$$

定理 3.62 的证明 (1) $\Rightarrow$ (2): 如果 $e_1, \cdots, e_n$ 是 V 的标准正交基, 且 $\mathcal{A}e_i = \lambda_i e_i$, 那么, 由 (3.5.46) 知 $\mathcal{A}^* e_i = \bar{\lambda}_i e_i$. 因此, $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$.

(2) $\Rightarrow$ (1): 取 $\mathcal{A}$ 的特征值 λ 并考虑特征子空间

$$V^\lambda = \{x \in V | \mathcal{A}x = \lambda x\}.$$

由条件 (3.5.46) 知

$$\mathcal{A}^*(V^\lambda) \subseteq V^\lambda.$$

于是, 对任意的 $x \in V^\lambda$ 和 $y \in (V^\lambda)^\perp$, 有

$$(x|\mathcal{A}y) = (\mathcal{A}^* x|y) = 0, \quad (x|\mathcal{A}^* y) = (\mathcal{A}x|y) = 0.$$

这说明, $(V^\lambda)^\perp$ 也是 $\mathcal{A}$ 和 $\mathcal{A}^*$ 不变的. 显然, 算子 $\mathcal{A}$ 和 $\mathcal{A}^*$ 在 $(V^\lambda)^\perp$ 上的限制是交换的. 从而 $\mathcal{A}$ 在 $(V^\lambda)^\perp$ 上的限制是正规的. 因为 $V = V^\lambda \oplus (V^\lambda)^\perp$, 对 $\dim V$ 做归纳法即知 $\mathcal{A}$ 在 V 的某个标准正交基下的矩阵是对角矩阵. □

可以比较定理 3.62 的证明和定理 3.42 及定理 3.58, 其本质是类似的: 把空间分解成不变子空间的直和, 然后对维数用归纳法. 由于在特征子空间上算子的作用就是纯量作用, 十分简单, 所以正规算子, 或更一般地, 可对角化算子, 能分解成一些很简单的算子的和. 下面我们更详细地讨论这一点.

二 可对角化算子的谱分解 这一部分的结果对任意域上向量空间都是成立的. 考虑域 K 上的有限维向量空间 V. 如果 V 是子空间 $V_1, \cdots, V_r$ 的直和, 那么可以定义沿着 $\sum\limits_{j\neq i} V_j$ 到 V_i 的投影算子 $\mathcal{P}_i$(参见例 2.27):

$$\mathcal{P}_i : V \to V, \ \sum_{j=1}^{r} v_j \to v_i, \quad \text{其中} v_j \in V_j, \ j = 1, \cdots, r.$$

这些投影算子满足如下条件:

$$\mathcal{P}_i\mathcal{P}_j = \delta_{ij}\mathcal{P}_i, \quad \mathcal{P}_1 + \cdots + \mathcal{P}_m = \mathcal{E}. \tag{3.5.49}$$

算子 $\mathcal{P} : V \to V$ 是投影算子当且仅当 $\mathcal{P}^2 = \mathcal{P}$. 此时有 $V = \ker\mathcal{P} + \operatorname{Im}\mathcal{P}$. 两个投影算子 $\mathcal{P}$ 和 $\mathcal{Q}$ 称为**正交**的如果 $\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P} = \mathcal{O}$. 下面的结论称为可对角化算子的谱定理.

定理 3.63 设 V 是域 K 上的 n 维空间, $\mathcal{A} : V \to V$ 是可对角化算子, 互不相同的特征值为 $\lambda_1, \cdots, \lambda_r$. 那么, 存在相互正交的投影算子 $\mathcal{P}_1, \cdots, \mathcal{P}_r$, 使得

(1) $\sum_i \mathcal{P}_i = \mathcal{E}$;

(2) $\sum_i \lambda_i \mathcal{P}_i = \mathcal{A}$ 是算子 $\mathcal{A}$ 的谱分解式;

(3) 上面的谱分解式是唯一的. 即, 如果 $\mathcal{Q}_1, \cdots, \mathcal{Q}_m$ 是相互正交的投影算子, 满足 $\sum_j \mathcal{Q}_j = \mathcal{E}$, $\sum_j \mu_j \mathcal{Q}_j = \mathcal{A}$, 诸 $\mu_j \in K$ 互不相同, 则 $\{\lambda_1, \cdots, \lambda_r\} = \{\mu_1, \cdots, \mu_m\}$, 当 $\lambda_i = \mu_j$ 时有 $\mathcal{P}_i = \mathcal{Q}_j$;

(4) 存在 K 上的多项式 $f_1, \cdots, f_m$, 使得

$$f_i(\lambda_j) = \delta_{ij}, \quad f_i(\mathcal{A}) = \mathcal{P}_i.$$

证明 由于 $\mathcal{A}$ 是可对角化的, 互不相同的特征值为 $\lambda_1, \cdots, \lambda_r$, 所以有

$$V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_r}.$$

命 $\mathcal{P}_i$ 为沿着 $\sum_{j \neq i} V_j$ 到 V_i 的投影算子. 那么 $\mathcal{P}_1, \cdots, \mathcal{P}_r$ 相互正交, 且和为恒等算子. 结论 (1) 得证.

对任意向量 $v \in V$, 根据 (1), 有 $v = v_1 + \cdots + v_r$, 其中 $v_i = \mathcal{P}_i v \in V^{\lambda_i}$. 于是,

$$\mathcal{A}v = \mathcal{A}\mathcal{E}v = \mathcal{A}\left(\sum_i \mathcal{P}_i v\right) = \sum_i \mathcal{A}v_i = \sum_i \lambda_i v_i = \sum_i \lambda_i(\mathcal{P}_i v) = \left(\sum_i \lambda_i \mathcal{P}_i\right) v.$$

所以, 结论 (2) 成立.

现证谱分解式的唯一性. 假设$\mathcal{Q}_1, \cdots, \mathcal{Q}_m$是相互正交的投影算子, 满足 $\sum_j \mathcal{Q}_j = \mathcal{E}$, $\sum_j \mu_j \mathcal{Q}_j = \mathcal{A}$. 命 $U_i = \mathcal{Q}_i V$. 由于 $\mathcal{Q}_j \mathcal{Q}_i = \mathcal{O}$ 如果 $j \neq i$, 所以对 $x = \mathcal{Q}_i y \in U_i = \mathcal{Q}_i U$, 有

$$\mathcal{A}x = \sum_k \mu_k \mathcal{Q}_k x = \mu_i \mathcal{Q}_i x = \mu_i x.$$

即, μ_i 是 $\mathcal{A}$ 的特征值, U_i 是特征子空间 V^{μ_i} 的子空间. 但是,

$$V = \mathcal{E}V = \sum_j \mathcal{Q}_j V = \sum_j U_j.$$

不同的特征子空间的和是直和. 这迫使 $U_i = V^{\mu_i}$, $m = r$, 且 $\{\lambda_1, \cdots, \lambda_r\} = \{\mu_1, \cdots, \mu_m\}$, 当 $\lambda_i = \mu_j$ 时有 $\mathcal{P}_i = \mathcal{Q}_j$. 结论 (3) 得证.

结论 (4) 中的多项式 $f_1(t), \cdots, f_r(t) \in K[t]$ 可以清晰地写出来:

$$f_i(t) = \prod_{j \neq i} \frac{t - \lambda_j}{\lambda_i - \lambda_j}.$$

由投影算子 $\mathcal{P}_i$ 之间的正交性知

$$\mathcal{A}^2 = \left(\sum_i \lambda_i \mathcal{P}_i\right)\left(\sum_j \lambda_j \mathcal{P}_j\right) = \sum_j \lambda_j^2 \mathcal{P}_j^2,$$

$$\mathcal{A}^3 = \left(\sum_i \lambda_i \mathcal{P}_i\right)\left(\sum_j \lambda_j^2 \mathcal{P}_j^2\right) = \sum_j \lambda_j^3 \mathcal{P}_j^3,$$

$$\cdots\cdots$$

$$\mathcal{A}^k = \left(\sum_i \lambda_i \mathcal{P}_i\right)\left(\sum_j \lambda_j^{k-1} \mathcal{P}_j^{k-1}\right) = \sum_j \lambda_j^k \mathcal{P}_j^k.$$

(当 $k=0$ 时, 有 $\mathcal{A}^0 = \sum\limits_j \lambda_j^0 \mathcal{P}_j = \sum\limits_j \mathcal{P}_j = \mathcal{E}$.) 这样一来, 对任意的多项式 $f(t) \in K[t]$, 有

$$f(\mathcal{A}) = \sum_j f(\lambda_j)\mathcal{P}_j.$$

特别, 有

$$f_i(\mathcal{A}) = \sum_j f_i(\lambda_j)\mathcal{P}_j = f_i(\lambda_i)\mathcal{P}_i = \mathcal{P}_i. \qquad \square$$

三 我们回到正规算子定义中的交换性. 结论 (3.5.46) 说正规算子 $\mathcal{A}$ 的特征向量也是算子 $\mathcal{A}^*$ 的特征向量. 这个结论的证明用到的条件是 $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$. 一般地, 我们有如下的结论.

命题 3.64 复向量空间 V 上的两个算子 $\mathcal{A}$ 和 $\mathcal{B}$ 如果是交换的, 那么它们有共同的特征向量.

证明 设 λ 是 $\mathcal{A}$ 的特征值. 那么 $\mathcal{A}$ 的特征子空间 V^λ 是 $\mathcal{B}$ 的不变子空间, 因为, 对任意的 $x \in V^\lambda$, 有

$$\mathcal{A}(\mathcal{B}x) = (\mathcal{A}\mathcal{B})x = (\mathcal{B}\mathcal{A})x = \mathcal{B}(\mathcal{A}x) = \lambda\mathcal{B}x,$$

即, $\mathcal{B}x \in V^\lambda$.

线性算子 $\mathcal{B}$ 限制在 V^λ 上仍是线性算子, 所以有特征向量 $y \in V^\lambda$: $\mathcal{B}y = \mu y$. 向量 y 就是 $\mathcal{A}$ 和 $\mathcal{B}$ 共同的特征向量: $\mathcal{A}y = \lambda y$, $\mathcal{B}y = \mu y$. $\square$

定理 3.65 埃尔米特空间 V 上的正规算子 $\mathcal{A}$ 和 $\mathcal{B}$ 能在某个标准正交基下同时对角化当且仅当它们是交换的.

证明 如果 $\mathcal{A}$ 和 $\mathcal{B}$ 在某个标准正交基下同时对角化, 那么它们在这个基下的矩阵是交换的, 所以两个算子是交换的.

反过来, 如果这两个算子交换: $\mathcal{AB}=\mathcal{BA}$, 我们证明它们在某个标准正交基下同时对角化. 设 λ 是 $\mathcal{A}$ 的特征值. 从命题 2.62 的证明知, $\mathcal{A}$ 的以 λ 为特征值的特征子空间 $V_{\mathcal{A}}^{\lambda}$ 是 $\mathcal{B}$ 的不变子空间. 由式 (3.5.46) 知 $V_{\mathcal{A}}^{\lambda}$ 也是 $\mathcal{A}^*$ 的以 $\bar{\lambda}$ 为特征值的特征子空间 $V_{\mathcal{A}^*}^{\bar{\lambda}}$. 由于 $\mathcal{A}^*\mathcal{B}^*=\mathcal{B}^*\mathcal{A}^*$, 仍由命题 2.62 的证明知 $V_{\mathcal{A}}^{\lambda}=V_{\mathcal{A}^*}^{\bar{\lambda}}$ 是 $\mathcal{B}^*$ 的不变子空间. 所以, $\mathcal{B}$ 在 V^{λ} 上的限制是正规算子, 在 $V_{\mathcal{A}}^{\lambda}$ 的某个标准正交基下的矩阵是对角矩阵. 对 $\mathcal{A}$ 的每个特征子空间, 取这样一个标准正交基, 然后合起来, 就得到 V 的一个基 (e_i), 在这个基下, $\mathcal{A}$ 和 $\mathcal{B}$ 的矩阵都是对角的. 由于正规算子的特征子空间相互正交 (参见定理 3.62 的证明), 刚才得到的基 (e_i) 是 V 的标准正交基. □

对两个埃尔米特矩阵 $\mathcal{A}$ 和 $\mathcal{B}$ 而言, 可交换等价于它们的乘积是埃尔米特算子: $\mathcal{AB}=\mathcal{BA}=\mathcal{B}^*\mathcal{A}^*=(\mathcal{AB})^*$.

对任意的算子 $\mathcal{A}$, 算子

$$\mathcal{B}=\frac{1}{2}(\mathcal{A}+\mathcal{A}^*),\quad \mathcal{C}=\frac{1}{2i}(\mathcal{A}-\mathcal{A}^*)$$

都是埃尔米特算子. 易见 $\mathcal{A}=\mathcal{B}+i\mathcal{C}$, $\mathcal{A}^*=\mathcal{B}-i\mathcal{C}$. 如果 $\mathcal{A}$ 是正规的, 那么 $\mathcal{BC}=\mathcal{CB}$, 所以 $\mathcal{B}$ 和 $\mathcal{C}$ 可以在某个标准正交基下同时对角化. 反之, 如果 $\mathcal{B}$ 和 $\mathcal{C}$ 可交换, 则有

$$\mathcal{AA}^*=\mathcal{B}^2+\mathcal{C}^2=\mathcal{A}^*\mathcal{A},$$

即 $\mathcal{A}$ 是正规的.

习　题　3.5

1. 证明: 内积空间上的正规算子具有不同特征值的特征向量是正交的.

2. 证明: 埃尔米特空间上的算子 $\mathcal{A}$ 是正规的当且仅当 $\mathcal{A}^*=p(\mathcal{A})$ 对某个多项式 $p(t)$.

3. 设 $\mathcal{A}$ 是欧氏空间 V 上的正规算子. 证明: 如果 $\mathcal{A}^2=-\mathcal{E}$, 那么 $\mathcal{A}^*=-\mathcal{A}$.

4. 设 A 和 B 是交换的 n 阶复矩阵. 如果它们都与对角矩阵相似, 那么存在可逆矩阵 C 使 CAC^{-1} 和 CBC^{-1} 都是对角矩阵.

5. 设 $p(t)=t^2+at+b$ 是实不可约多项式, $\mathcal{A}$ 是欧氏空间上的正规算子. 证明: 如果 $p(\mathcal{A})=0$, 那么 $\mathcal{A}^*=-\mathcal{A}-a\mathcal{E}$.

6. 证明: 埃氏空间上的正规算子 $\mathcal{A}$ 如果与算子 $\mathcal{B}$ 交换, 那么 $\mathcal{A}$ 与 $\mathcal{B}^*$ 交换.

7. 设埃氏空间 V 上的正规算子 $\mathcal{A}$ 和 $\mathcal{B}$ 有相同的特征多项式. 证明: 这两个算子在任何基下的矩阵都是相似的.

8. 设 $\mathcal{A}$ 是欧氏空间 V 上的正规算子. 证明: 存在 V 的标准正交基, 使得 $\mathcal{A}$ 在这个基下

的矩阵是对角分块的:

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

其中每个块 A_i 是方阵, 阶至多为 2, 且 2 阶的块 A_i 具有如下形式

$$\begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}.$$

9. 证明: 有限维复向量空间上任意一族交换的线性算子都有共同的特征向量.

10. 设 A 是一个 n 阶实雅可比矩阵, 即具有如下形式的矩阵

$$\begin{pmatrix} a_1 & b_1 & 0 & \cdots & 0 & 0 \\ c_1 & a_2 & b_2 & \cdots & 0 & 0 \\ 0 & c_2 & a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \cdots & c_{n-1} & a_n \end{pmatrix}, \quad b_ic_i > 0, \quad i = 1, \cdots, n-1.$$

证明: A 的特征根都是实数且重数都是 1, 即 $\operatorname{Spec} A$ 有 n 个单谱点.

$\Bigg($在分析数学中, m 个 n 元可微函数 $f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n)$ 的雅可比矩阵是 $m \times n$ 矩阵 $\left(\dfrac{\partial f_i}{\partial x_j}\right)$.$\Bigg)$

3.6 复化与实化

复数域是代数闭域, 很多问题变得较简单, 如向量空间上算子的标准形. 而且, 很多实数域上的问题借助复数域能得到更简洁的处理. 如实多项式可以分解成 1 次或 2 次实多项式的乘积. 又如把实对称方阵看成埃尔米特空间上的算子的矩阵, 很容易就证明了这些方阵的特征多项式的根都是实数 (引理 3.43) 且与对角矩阵相似, 进而也就知道实向量空间上的算子的矩阵如果是对称矩阵, 则可对角化. 实对称矩阵这个例子背后涉及一个观点: 借助矩阵, 实向量空间上的算子自然对应到复向量空间的算子. 这个对应有很多好的性质, 是**函子**的一个例子. 本节仔细考察这个观点背后实向量空间与复向量空间的联系.

一 复结构 复向量空间自然是实向量空间. 实向量空间能成为复向量空间的条件是什么?

定义 3.66 实向量空间 V 上的一个**复结构**就是 V 上的一个线性算子 $\mathcal{J}$, 其平方 $\mathcal{J}^2 = -\mathcal{E}$.

例 3.67　算子 $\mathcal{J}:\mathbb{R}^2\to\mathbb{R}^2$, $(a,b)\to(-b,a)$ 是 $\mathbb{R}^2$ 上的一个复结构. 更一般地, 如果 $2n$ 维实向量空间 V 上的算子 $\mathcal{J}$ 在某个基下的矩阵是

$$\begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}, \tag{3.6.50}$$

那么 $\mathcal{J}$ 是 V 的一个复结构.

把定义中的算子称为复结构的原因是借助它, 向量空间 V 自然成为复向量空间:

$$(a+bi)v=av+b\mathcal{J}v,\quad a,b\in\mathbb{R},\quad v\in V.$$

分配律

$$\alpha(u+v)=\alpha u+\alpha v,\quad (\alpha+\beta)v=\alpha v+\beta v,\quad \alpha,\beta\in\mathbb{C},$$

显然成立. 由于 $\mathcal{J}$ 是线性算子, 且 $\mathcal{J}^2=-\mathcal{E}$, 所以纯量乘的结合律也成立:

$$\begin{aligned}(a+bi)[(c+di)v]&=(a+bi)(cv+d\mathcal{J}v)=a(cv+d\mathcal{J}v)+b\mathcal{J}(cv+d\mathcal{J}v)\\&=acv+ad\mathcal{J}v+bc\mathcal{J}v-bdv=(ac-bd)v+(ad+bc)\mathcal{J}v\\&=[ac-bd+(ad+bc)i]v=[(a+bi)(c+di)]v.\end{aligned}$$

向量空间其他的公理是自然满足的, 因为 V 是实向量空间. 这个复向量空间将记作 $V_{\mathcal{J}}$.

辛空间带有算子 $\mathcal{J}$, 其平方 $\mathcal{J}^2=-\mathcal{E}$(见 3.1 节第六部分), 所以辛空间自然成为复向量空间. 下面的结论表明实向量空间的复结构和辛结构有密切的联系.

命题 3.68　如果向量空间 V 有复结构 $\mathcal{J}$, 那么 V 的实维数 $(\dim_{\mathbb{R}}V)$ 是偶数, 在某个基下, $\mathcal{J}$ 的矩阵形如式 (3.6.50). 而且有

$$\dim_{\mathbb{C}}V_{\mathcal{J}}=\frac{1}{2}\dim_{\mathbb{R}}V. \tag{3.6.51}$$

证明　取 V 的非零向量 e_1, 那么 $e_1,\mathcal{J}e_1$ 线性无关. 事实上, 如果 $ae_1+b\mathcal{J}e_1=0$, 用算子 $\mathcal{J}$ 作用在这个等式上, 得 $-be_1+a\mathcal{J}e_1=0$. 于是 $(a^2+b^2)e_1=0$. 从而 $a=b=0$.

假设已经找到了向量 $e_1,\cdots,e_k$ 使得向量组 $e_1,\ \cdots,\ e_k,\ \mathcal{J}e_1,\ \cdots,\ \mathcal{J}e_k$ 线性无关. 命

$$V_k=\langle e_1,\ \cdots,\ e_k,\ \mathcal{J}e_1,\ \cdots,\ \mathcal{J}e_k\rangle.$$

如果 $V_k\neq V$, 取 $e_{k+1}\in V\backslash V_k$. 那么, 向量组 $e_1,\ \cdots,\ e_k,\ e_{k+1},\ \mathcal{J}e_1,\ \cdots,\ \mathcal{J}e_k$ 线性无关. 假设

$$\mathcal{J}e_{k+1}=ae_{k+1}+v,\quad a\in\mathbb{R},\quad v\in V_k.$$

用算子 $\mathcal{J}$ 作用到上式两端, 得

$$-e_{k+1} = a\mathcal{J}e_{k+1} + \mathcal{J}v, \quad a \in \mathbb{R}, \quad v \in V_k.$$

注意 V_k 是 $\mathcal{J}$ 不变的, 所以 $\mathcal{J}v \in V_k$. 把前一个等式乘以 a 后加到后一个等式, 可得

$$(a^2+1)e_{k+1} = -av - \mathcal{J}v \in V_k.$$

由于 $a^2+1 \neq 0$, 导致 $e_{k+1} \in V_k$, 这与 e_{k+1} 的取法矛盾. 于是, 向量组 $e_1, \cdots, e_k$, e_{k+1}, $\mathcal{J}e_1, \cdots, \mathcal{J}e_k, \mathcal{J}e_{k+1}$ 线性无关.

继续这一过程, 最后, 会有某个 n 使得

$$V = V_n = \langle e_1, \cdots, e_n, \mathcal{J}e_1, \cdots, \mathcal{J}e_n \rangle.$$

于是, $\dim_{\mathbb{R}} V = 2n$, 且 $\mathcal{J}$ 在基 $e_1, \cdots, e_n, \mathcal{J}e_1, \cdots, \mathcal{J}e_n$ 下的矩阵形如式 (3.6.50). 在复向量空间 $V_{\mathcal{J}}$ 中, $\mathcal{J}e_k = ie_k$, 所以等式 (3.6.51) 成立. □

可以看出, 这个证明实质上就是斜对称双线性型寻找典范基的过程.

二 复化 实向量空间上的算子自然对应到复向量空间的算子的过程中, 两个向量空间的维数是一致的. 所以, 对每一个实向量空间, 应该有一个同维数的复向量空间与之对应.

设 V 是 n 维实向量空间. 考虑外直和 $V \oplus V$ (参见 1.4 节最后的讨论). 作为集合, $V \oplus V$ 与笛卡尔积 $V \times V$ 是一样的, 数乘与加法的运算分别是

$$a(u,v) = (au, av), \quad (u,v) + (u',v') = (u+u') + (v,v'), \quad a \in \mathbb{R}.$$

例 3.67 的映射在这里有如下形式:

$$\mathcal{J} : (u,v) \to (-v, u).$$

这是 $V \oplus V$ 的一个复结构, 称为 $V \oplus V$ 的**典范复结构**.

定义 3.69 复空间 $(V \oplus V)_{\mathcal{J}}$ 称为实向量空间 V 的**复化空间**, 记作 $V^{\mathbb{C}}$.

用张量积的语言表达复化更简单和方便:

$$V^{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} V = (V \oplus V)_{\mathcal{J}}.$$

在第 6 章会简单介绍张量积. 因为 $\dim_{\mathbb{R}}(V \oplus V) = 2n$, 所以等式 (3.6.49) 给出

$$\dim_{\mathbb{C}} V^{\mathbb{C}} = n = \dim_{\mathbb{R}} V.$$

在 1.4 节讨论外直和时, 我们注意到形如 $(u, 0)$ 的向量全体 V' 是 $V \oplus V$ 的子空间, 与 V 同构. 利用这个同构, 我们可以等同 V' 与 V. 于是, $(u, 0)$ 可以简单记

作 u. 根据定义, $i(u,v)=\mathcal{J}(u,v)=(-v,u)$. 所以 $(u,v)=(u,0)+i(v,0)=u+iv$. 在这样的记号下, 向量加法成为

$$(u+iv)+(u'+iv')=u+u'+i(v+v').$$

数乘运算则是

$$(a+bi)(u+iv)=(au-bv)+i(av+bu),$$

这是因为

$$(a\mathcal{E}+b\mathcal{J})(u,v)=a(u,v)+b\mathcal{J}(u,v)=(au-bv,av+bu).$$

这些运算和复数的运算形式上是一致的.

定义 3.70 实向量空间 V 上的线性算子 $\mathcal{A}$ 的**复化算子**$\mathcal{A}^{\mathbb{C}}:V^{\mathbb{C}}\to V^{\mathbb{C}}$ 由下式定义:

$$\mathcal{A}^{\mathbb{C}}(u+iv)=\mathcal{A}u+i\mathcal{A}v.$$

如果 $e_1,\cdots,e_n$ 是 V 的基, 那么这个向量组也是 $V^{\mathbb{C}}$ 的基. 由此可见, 算子 $\mathcal{A}$ 和算子 $\mathcal{A}^{\mathbb{C}}$ 在基 $e_1,\cdots,e_n$ 下的矩阵是一样的. 特别, 有 $\det\mathcal{A}^{\mathbb{C}}=\det\mathcal{A}$, $\operatorname{tr}\mathcal{A}^{\mathbb{C}}=\operatorname{tr}\mathcal{A}$.

作为复化的一个应用, 我们证明 $\mathcal{A}$ 如果没有 1 维不变子空间, 则一定有 2 维不变子空间 (定理 2.42). 算子 $\mathcal{A}^{\mathbb{C}}$ 是复向量空间上的线性算子, 所以有特征向量. 设 $x+iy$ 是属于特征值 $a+bi$ 的特征向量, $x,y\in V$, $a,b\in\mathbb{R}$. 那么

$$\mathcal{A}^{\mathbb{C}}(x+iy)=\mathcal{A}x+i\mathcal{A}y=(a+bi)(x+iy)=(ax-by)+i(bx+ay).$$

由此可知

$$\mathcal{A}x=ax-by,\quad \mathcal{A}y=bx+ay.$$

也就是说, 在 V 中, 向量 x 和 y 张成的子空间是 $\mathcal{A}$ 的不变子空间, 它必然是 2 维的.

现在关注映射 $\mathcal{A}\to\mathcal{A}^{\mathbb{C}}$ 的性质. 意料之中的, 我们有

$$(\mathcal{A}+\mathcal{B})^{\mathbb{C}}=\mathcal{A}^{\mathbb{C}}+\mathcal{B}^{\mathbb{C}},\quad (\mathcal{A}\mathcal{B})^{\mathbb{C}}=\mathcal{A}^{\mathbb{C}}\mathcal{B}^{\mathbb{C}}.$$

验证是直截了当的:

$$\begin{aligned}(\mathcal{A}+\mathcal{B})^{\mathbb{C}}(u+iv)&=(\mathcal{A}+\mathcal{B})u+i(\mathcal{A}+\mathcal{B})v\\&=(\mathcal{A}u+\mathcal{B}u)+i(\mathcal{A}v+\mathcal{B}v)=(\mathcal{A}u+i\mathcal{A}v)+(\mathcal{B}u+i\mathcal{B}v)\\&=\mathcal{A}^{\mathbb{C}}(u+iv)+\mathcal{B}^{\mathbb{C}}(u+iv)=(\mathcal{A}^{\mathbb{C}}+\mathcal{B}^{\mathbb{C}})(u+iv),\end{aligned}$$

类似验证第二个等式.

实向量空间 V 上的双线性型 f 可以直接延拓成 $V^{\mathbb{C}}$ 上双线性型. 不过, 我们更感兴趣把 f 延拓成 $V^{\mathbb{C}}$ 的半双线性型:

$$f^{\mathbb{C}}(x+iy,u+iv)=f(x,u)+f(y,v)-if(x,v)+if(y,u).$$

容易验证, 如果 f 是 (斜) 对称的, 那么 $f^{\mathbb{C}}$ 是 (斜) 埃尔米特的.

假设 f 是正定的, 那么 $f^{\mathbb{C}}$ 也是正定的. 所以, 如果 V 是带内积 $(\cdot|\cdot)$ 的欧氏空间, 那么, $V^{\mathbb{C}}$ 带着内积

$$(x+iy|u+iv)^{\mathbb{C}}=(x|u)+(y|v)-i(x|v)+i(u|y)$$

就成为埃尔米特空间. 特别, $V^{\mathbb{C}}$ 上的模 $\|\cdot\|^{\mathbb{C}}$ 由下式给出:

$$\|x+iy\|^{\mathbb{C}}=\|x\|^2+\|y\|^2.$$

每一个 n 维复向量空间 U 都是某个 n 维实向量空间的复化. 实际上, 任取 U 的基 $e_1,\cdots,e_n$, 所有的实线性组合 $\sum\limits_i a_ie_i$, $a_i\in\mathbb{R}$, 形成的集合 V 是 n 维实向量空间. 而且, 有

$$U=V^{\mathbb{C}},\quad 即\quad (e_1,\cdots,e_n)_{\mathbb{C}}=((e_1,\cdots,e_n)_{\mathbb{R}})^{\mathbb{C}}.$$

三　实化　对复向量空间 U, 如果在数乘运算中仅考虑实数乘向量, 那么, U 就成为一个实向量空间, 称为 U 的**实化**, 记作 $U_{\mathbb{R}}$. 同样, U 的线性算子 $\mathcal{A}$ 如果在数乘运算中仅考虑实数乘向量, 就成为 $U_{\mathbb{R}}$ 的线性算子, 称为 $\mathcal{A}$ 的**实化**, 记作 $\mathcal{A}_{\mathbb{R}}$. 这些都是平淡无奇的.

如果 $e_1,\cdots,e_n$ 是 U 的基, 那么 $e_1,\cdots,e_n,ie_1,\cdots,ie_n$ 就是 $U_{\mathbb{R}}$ 的基. 算子 $\mathcal{A}$ 和 $\mathcal{A}_{\mathbb{R}}$ 在相应的基下的矩阵分别记作 A 和 $A_{\mathbb{R}}$. 命 $A=A_1+iA_2$, 其中 A_1,A_2 都是实矩阵. 又设 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 为 U 上的线性算子, 它们在基 (e_i) 下的矩阵分别是 A_1 和 A_2. 则有 $\mathcal{A}=\mathcal{A}_1+i\mathcal{A}_2$,

$$\mathcal{A}(ie_k)=i\mathcal{A}(e_k)=i(\mathcal{A}_1+i\mathcal{A}_2)e_k=-\mathcal{A}_2e_k+i\mathcal{A}_1e_k.$$

这意味着, 算子 $\mathcal{A}_{\mathbb{R}}$ 在基 $e_1,\cdots,e_n,ie_1,\cdots,ie_n$ 下的矩阵是

$$A_{\mathbb{R}}=\begin{pmatrix}A_1 & -A_2\\ A_2 & A_1\end{pmatrix}. \tag{3.6.52}$$

我们看到, $U_{\mathbb{R}}$ 上的线性算子能成为某个算子的实化是不太容易的.

直观上看, $\mathcal{A}_{\mathbb{R}}$ 的作用就是 $\mathcal{A}$ 作用在 U 的基 $e_1, \cdots, e_n$ 和基 $ie_1, \cdots, ie_n$ 的作用的合并, 所以行列式应该有下面的结论.

命题 3.71　$\det \mathcal{A}_{\mathbb{R}} = |\det \mathcal{A}|^2$.

证明　我们计算这些算子的矩阵的行列式. 利用等式 (3.6.52) 得

$$
\begin{aligned}
\det \mathcal{A}_{\mathbb{R}} = \det A_{\mathbb{R}} = \det \begin{pmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{pmatrix} &= \det \begin{pmatrix} A_1 + iA_2 & -A_2 + iA_1 \\ A_2 & A_1 \end{pmatrix} \\
&= \det \begin{pmatrix} A_1 + iA_2 & 0 \\ A_2 & A_1 - iA_2 \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ A_2 & \bar{A} \end{pmatrix} \\
&= \det A \det \bar{A} = \det A \cdot \overline{\det A} = |\det A|^2,
\end{aligned}
$$

其中 $\bar{A}$ 记 A 的共轭.　□

算子的实化并不改变算子在向量上的作用, 所以, 对于 U 上的算子 $\mathcal{A}$ 和 $\mathcal{B}$ 有

$$
(\mathcal{A} + \mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} + \mathcal{B}_{\mathbb{R}}, \quad (\mathcal{A}\mathcal{B})_{\mathbb{R}} = \mathcal{A}_{\mathbb{R}} \cdot \mathcal{B}_{\mathbb{R}}, \quad (a\mathcal{A})_{\mathbb{R}} = a\mathcal{A}_{\mathbb{R}}, \quad a \in \mathbb{R}.
$$

也就是说, 通过实化所有 U 的算子得到的集合 $\mathcal{L}(U)_{\mathbb{R}}$ 是 $U_{\mathbb{R}}$ 上的线性算子全体构成的代数 $\mathcal{L}(U_{\mathbb{R}})$ 的子代数. 作为向量空间, $\mathcal{L}(U)_{\mathbb{R}}$ 就是 $\mathcal{L}(U)$(复向量空间 U 的线性算子全体) 的实化, 所以

$$
\dim \mathcal{L}(U)_{\mathbb{R}} = 2 \dim_{\mathbb{C}} \mathcal{L}(U) = 2n^2 = \frac{1}{2} \dim \mathcal{L}(U_{\mathbb{R}}).
$$

用虚数单位 $i = \sqrt{-1}$ 乘 U 中的向量是 U 的线性算子, 记其实化为 $\mathcal{J}$, 则有

$$
\mathcal{J} e_k = ie_k, \quad \mathcal{J}(ie_k) = -e_k, \quad 1 \leqslant k \leqslant n.
$$

这是 $U_{\mathbb{R}}$ 的一个复结构. 显然, 我们有

$$
U_{\mathbb{R}, \mathcal{J}} = (U_{\mathbb{R}})_{\mathcal{J}} = U.
$$

显然, $U_{\mathbb{R}}$ 的线性算子 $\mathcal{C}$ 是 U 的某个算子的实化当且仅当它与 $\mathcal{J}$ 是交换的: $\mathcal{C}\mathcal{J} = \mathcal{J}\mathcal{C}$. 这个条件在基 $e_1, \cdots, e_n, ie_1, \cdots, ie_n$ 下的矩阵形式是

$$
\begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix},
$$

即

$$
\begin{pmatrix} C_2 & -C_1 \\ C_4 & -C_3 \end{pmatrix} = \begin{pmatrix} -C_3 & -C_4 \\ C_1 & C_2 \end{pmatrix},
$$

换句话说, $C_1 = C_4$, $C_2 + C_3 = 0$. 这正是 (3.6.52) 中矩阵的条件. 上面所讨论的可总结如下:

命题 3.72 $\mathcal{L}(U_{\mathbb{R}})$ 的子代数 $\mathcal{L}(U)_{\mathbb{R}}$ 正好由 $\mathcal{L}(U_{\mathbb{R}})$ 中所有与 $\mathcal{J}$ 交换的线性算子组成. 如果 $e_1, \cdots, e_n$ 是 U 的基, 那么 $\mathcal{L}(U)_{\mathbb{R}}$ 中的算子在基 $e_1, \cdots, e_n, ie_1, \cdots, ie_n$ 下的矩阵具有式 (3.6.52) 的形式.

这个命题实际上还回答了如下的问题: 偶数维的实向量空间 V 上的算子 $\mathcal{A}$ 在什么时候与 V 的一个复结构**相容**, 即存在复结构 $\mathcal{J}$ 和线性算子 $\mathcal{B}: V_{\mathcal{J}} \to V_{\mathcal{J}}$ 使得 $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$. 答案是: 在 V 的某个 $e_1, \cdots, e_n, e_{n+1}, \cdots, e_{2n}$ 基下, $\mathcal{A}$ 的矩阵具有式 (3.6.52) 的形式. 这时, $\mathcal{J}e_i = e_{n+i}$, $\mathcal{J}e_{n+i} = -e_i$, $1 \leqslant i \leqslant n$. 不过, 要具体找出这个基是不容易的. 要找出复结构 $\mathcal{J}$, 还可以从所有与 $\mathcal{A}$ 交换的算子中找出平方为 $-\mathcal{E}$ 的算子. 二维的情况是简单的.

定理 3.73 设 V 是二维实向量空间, $\mathcal{A}: V \to V$ 是没有特征向量的线性算子. 那么, $\mathcal{A}$ 与 V 的某个复结构相容.

证明 我们需要找到 V 的一个算子 $\mathcal{J}$ 使得 $\mathcal{J}^2 = -\mathcal{E}$ 且 $\mathcal{A}\mathcal{J} = \mathcal{J}\mathcal{A}$. 设 $\mathcal{A}$ 的特征多项式的两个根是 $a \pm bi$, $a, b \in \mathbb{R}$, 那么, $b \neq 0$, 且 $\mathcal{A}$ 的特征多项式是 $\chi_{\mathcal{A}}(t) = (t-a)^2 + b^2$. 根据凯莱–哈密顿定理, 有

$$(\mathcal{A} - a\mathcal{E})^2 + b^2\mathcal{E} = 0.$$

命

$$\mathcal{J} = b^{-1}(\mathcal{A} - a\mathcal{E}).$$

则有 $\mathcal{J}^2 = -\mathcal{E}$, $\mathcal{A} = a\mathcal{E} + b\mathcal{J}$. 显然有 $\mathcal{A}\mathcal{J} = \mathcal{J}\mathcal{A}$. 根据命题 3.72, 存在 $\mathcal{B}: V_{\mathcal{J}} \to V_{\mathcal{J}}$ 使得 $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$. 由于 $V_{\mathcal{J}}$ 是一维的, 所以 $\mathcal{B}$ 是用某个复数相乘的算子, 这个复数显然是 $a + bi$. □

四 复化–实化 把实向量空间复化, 再实化, 就能得到实平面和虚平面的概念. 设 V 是 n 维实向量空间, 复化后得到 n 维复向量空间 $V^{\mathbb{C}}$, 再实化, 得到 $2n$ 维实向量空间

$$W = (V^{\mathbb{C}})_{\mathbb{R}}.$$

易见

$$W = V \oplus iV. \tag{3.6.53}$$

自然, 称 V 是 W 的**实平面**, iV 是 W 的**虚平面**. 命 $\mathcal{J} = (i\mathcal{E})_{\mathbb{R}}$, 这是 $V^{\mathbb{C}}$ 上用 i 相乘所得到的算子 $i\mathcal{E}$ 的实化, 它交换实平面和虚平面.

当 $V = \mathbb{R}$ 时, $V^{\mathbb{C}} = \mathbb{C}^1$, 从而 $W = (\mathbb{C}^1)_{\mathbb{R}} = \mathbb{R} + i\mathbb{R}$. 这时, 复数取共轭就是 W 的一个算子:

$$a + bi \to a - bi = \overline{a + bi}.$$

对一般的实向量空间 V, 在 $W=(V^{\mathbb{C}})_{\mathbb{R}}=V\oplus iV$ 上, 有类似的算子:

$$\mathcal{S}: u+iv \to u-iv=\overline{u+iv}.$$

它在自然选取的基下的矩阵都是

$$S=\begin{pmatrix} I_n & 0 \\ 0 & -I_n \end{pmatrix}.$$

进一步, 可以定义线性算子 $\mathcal{A}: V^{\mathbb{C}} \to V^{\mathbb{C}}$ 的**复共轭算子**$\bar{\mathcal{A}}$ 如下:

$$\bar{\mathcal{A}}(u+iv)=\overline{\mathcal{A}(u+iv)}.$$

显然, 算子 $\bar{\mathcal{A}}$ 是半线性的, 而且有

$$(\bar{\mathcal{A}})_{\mathbb{R}}=\mathcal{S}\cdot\mathcal{A}_{\mathbb{R}}.$$

在 $V^{\mathbb{C}}$ 的实平面 V 中取一个基. 如果 $\mathcal{A}$ 在这个基下的矩阵是 $A=A_1+iA_2$, 其中 A_1 和 A_2 都是实矩阵, 那么 $\bar{\mathcal{A}}$ 在这个基下的矩阵就是 $\bar{A}=A_1-iA_2$. 由此可见, $\mathcal{A}=\bar{\mathcal{A}}$ 的充要条件是 $\mathcal{A}$ 把实平面映入实平面, 即 $\mathcal{A}=\mathcal{B}^{\mathbb{C}}$ 是某个线性算子 $\mathcal{B}: V \to V$ 的复化. 算子 $\mathcal{B}$ 没有任何神秘的, 它就是 $\mathcal{A}$ 在实平面 V 上的限制.

利用复共轭算子的概念, 借助矩阵 (3.6.52), 可以方便地写出实化算子的迹:

$$\operatorname{tr}\mathcal{A}_{\mathbb{R}}=2\operatorname{tr}A_1=\operatorname{tr}(\mathcal{A}+\bar{\mathcal{A}})=\operatorname{tr}\mathcal{A}+\operatorname{tr}\bar{\mathcal{A}}.$$

五　实化–复化　对复向量空间, 实化, 再复化, 也是很有意思的. 这时需要复共轭向量空间的概念, 我们在 1.1 节中的第一部分已经遇到了. 设 V 复向量空间, 其复共轭向量空间 $\bar{V}$ 的加法群和 V 一样, 但纯量乘有差别: $\lambda\odot x=\bar{\lambda}x$.

复共轭向量空间可以很自然地出现. 设 $\mathcal{J}$ 是实向量空间 V 的复结构, 那么, $-\mathcal{J}$ 也是 V 的复结构, 结果有 $V_{-\mathcal{J}}=\overline{V_{\mathcal{J}}}$.

现在, 对复向量空间 V, 实化, 再复化, 就得到一个 $2\dim V$ 维的复向量空间 $(V_{\mathbb{R}})^{\mathbb{C}}$. 下面我们证明作为复向量空间, $(V_{\mathbb{R}})^{\mathbb{C}}$ 和 $V\oplus\bar{V}$ 是同构的:

$$(V_{\mathbb{R}})^{\mathbb{C}}\simeq V\oplus\bar{V}.$$

为此, 需要在 $(V_{\mathbb{R}})^{\mathbb{C}}$ 中找出与 V 和 $\bar{V}$ 分别同构的子空间.

在实向量空间 $V_{\mathbb{R}}\oplus V_{\mathbb{R}}$ 上有两个明显的复结构, 一个是典范的复结构算子 $\mathcal{J}(x,y)=(-y,x)$, 它定义了复向量空间 $(V_{\mathbb{R}})^{\mathbb{C}}$. 用 $\cdot$ 表示 $(V_{\mathbb{R}})^{\mathbb{C}}$ 中的复数乘, 则在 $(V_{\mathbb{R}})^{\mathbb{C}}=(V_{\mathbb{R}}\oplus V_{\mathbb{R}})_{\mathcal{J}}$ 中有 $i\cdot(x,y)=\mathcal{J}(x,y)=(-y,x)$. 另一个来自 V 的复向量空间结构: $i(x,y)=(ix,iy)$, 它给出的复向量空间就是 $V\oplus V$. 因为 $\mathcal{J}$ 与 i 是交换

的, 所以它是 $V\oplus V$ 上的线性算子. 由于 $\mathcal{J}^2=-\mathcal{E}$, 所以 $\mathcal{J}$ 的特征值是 $\pm i$. 我们写出相应的特征子空间, 并把它们看作 $(V_{\mathbb{R}})^{\mathbb{C}}$ 的子集:

$$\begin{aligned}V^{1,0}&=\{(x,y)\in(V_{\mathbb{R}})^{\mathbb{C}}\mid\mathcal{J}(x,y)=i(x,y)\},\\V^{0,1}&=\{(x,y)\in(V_{\mathbb{R}})^{\mathbb{C}}\mid\mathcal{J}(x,y)=-i(x,y)\}.\end{aligned}$$

先证明这些子集是 $(V_{\mathbb{R}})^{\mathbb{C}}$ 的子空间, 分别同构于 V 和 $\bar{V}$.

首先, $V^{1,0}$ 和 $V^{0,1}$ 对于加法是封闭的. 在 $(V_{\mathbb{R}})^{\mathbb{C}}$ 中, 用 i 乘就是用 $\mathcal{J}$ 作用. 由于 $\mathcal{J}$ 和 $V\oplus V$ 中的 i 乘运算交换, 所以, 如果 $\mathcal{J}(x,y)=\pm i(x,y)$, 则有

$$\begin{aligned}\mathcal{J}[(a+bi)\cdot(x,y)]&=\mathcal{J}[(a\mathcal{E}+b\mathcal{J})(x,y)]=(a\mathcal{E}+b\mathcal{J})\mathcal{J}(x,y)\\&=(a\mathcal{E}+b\mathcal{J})[\pm i(x,y)]=\pm i[(a\mathcal{E}+b\mathcal{J})(x,y)]\\&=\pm i[(a+bi)\cdot(x,y)].\end{aligned}$$

即, $V^{1,0}$ 和 $V^{0,1}$ 对复数乘 $\cdot$ 是封闭的. 于是, 它们都是 $(V_{\mathbb{R}})^{\mathbb{C}}$ 的子空间.

从定义可以看出, $V^{1,0}$ 由形如 $(x,-ix)$, $x\in V_{\mathbb{R}}$, 的向量全体组成, 而 $V^{1,0}$ 则由形如 (y,iy), $y\in V_{\mathbb{R}}$, 的向量全体组成. 容易验证, 映射 $x\to(x,-ix)$ 和 $y\to(y,iy)$ 分别是 V 到 $V^{1,0}$ 和 $\bar{V}$ 到 $V^{0,1}$ 的同构.

由于方程 $(u,v)=(x,-ix)+(y,iy)$ 有唯一的解: $x=(u+iv)/2$, $y=(u-iv)/2$, 所以有 $V=V^{1,0}\oplus V^{0,1}$. 至此, 我们要的同构已经清楚了.

习 题 3.6

1. 设 $\mathcal{A}$ 是欧氏空间 V 上的正交算子, 它没有特征向量 (此时必有 $\dim V=2m$ 为偶数). 证明: 存在 V 的复结构 $\mathcal{J}$ 和复向量空间 $V_{\mathcal{J}}$ 上的内积, 以及关于这个内积的一个酉算子 $\mathcal{B}:V_{\mathcal{J}}\to V_{\mathcal{J}}$, 使得 $\mathcal{A}$ 是 $\mathcal{B}$ 的实化算子. (注意酉空间实化后自然是一个欧氏空间.)

2. 利用定理 2.47(代数闭域上的向量空间的线性算子可以上三角化) 证明命题 3.71.

3. 设 V 是欧氏空间, $V^{\mathbb{C}}$ 是其复化空间, $\mathcal{A}$ 是 $V^{\mathbb{C}}$ 的半线性算子, 定义法则是: 对任意 $u,v\in V$, $\mathcal{A}(u+iv)=u-iv$. 它实化后就是 $(V^{\mathbb{C}})_{\mathbb{R}}$ 的线性算子. 问: $\mathcal{A}_{\mathbb{R}}$ 是否是对称算子, 是否为正交算子?

4. 设 V 是有限维实向量空间. 证明:$(V^*)^{\mathbb{C}}$ 自然同构于 $(V^{\mathbb{C}})^*$. (所谓自然同构指其构造不依赖基的选取.)

3.7 正 交 展 开

一 在有限维内积空间中, 把向量通过标准正交基分解对很多问题的讨论都十分方便. 对无限维内积空间, 类似的分解同样是非常有用的. 无限维内积空间一般

很难写出基, 例如闭区间 $[a,b]$ 上的连续函数空间 $C(a,b)$. (对一般的向量空间, 也可以定义基的概念. 首先, 向量空间 V 中的一组向量称为线性无关的向量组如果其中任何有限个向量都是线性无关的. 如果 S 是 V 的一个线性无关的向量组, 且 V 中任何向量都可表示成 S 中有限个向量的线性组合, 那么 S 称为 V 的基. 例如 $1,t,t^2,t^3,\cdots$ 是多项式环 $K[t]$ 的一个基. 可以证明, 任何向量空间都有基.)

设 V 是内积空间. 空间中的一组向量 $e_1,e_2,\cdots$(可能有无穷多个) 称为**正交系**如果这些向量互相正交且都非零. 一个正交系称为**标准正交系**如果其中的向量都是单位向量 (即长度都是 1). 内积空间 V 中的一个正交系 $e_1,e_2,\cdots$ 称为**完备的**如果 V 中没有非零向量与该正交系中所有的向量正交, 即如果 $(v|e_i)=0$ 对所有的 e_i, 则可以推出 $v=0$. 完备的 (标准) 正交系有时简称为**(标准) 完备系**.

设 V 是无限维内积空间, $e_1,e_2,\cdots$ 是 V 的一个完备的标准正交系, v 是空间 V 的向量. 假定有展开式

$$v=\sum_{i=1}^{\infty}a_ie_i, \tag{3.7.54}$$

并且这个级数按照空间 V 的度量是收敛的 (见例 3.33 和 3.2 节中第五部分关于赋范向量空间的讨论). 我们找出这个展开式的系数. 为此, 先建立下面的引理.

引理 3.74 在内积空间 V 中, 内积 $(x|y)$ 是变量 x 和 y 的连续函数, 即: 当 $x_n\to x$, $y_n\to y$ 时, $(x_n|y_n)\to(x|y)$.

证明 设 $x-x_n=h_n$, $y-y_n=k_n$. 则有 $\|h_n\|\to 0$, $\|k_n\|\to 0$. 应用三角不等式和柯西–施瓦茨不等式, 得

$$\begin{aligned}|(x|y)-(x_n|y_n)|&=|(x-x_n|y)+(x_n|y-y_n)|=|(h_n|y)+(x|k_n)-(h_n|k_n)|\\&\leqslant\|h_n\|\,\|y\|+\|k_n\|\,\|x\|+\|h_n\|\,\|k_n\|\to 0.\end{aligned}$$

所以 $(x_n|y_n)\to(x|y)$. □

把展开式 (3.7.54) 中前 m 项的和记作 v_m. 对固定的向量 e_i, 由引理 3.74 知 $(v_m|e_i)\to(v|e_i)$. 当 $m>i$ 时, 由公式

$$(e_j|e_i)=\begin{cases}0, & 当\ j\neq i\ 时,\\ 1, & 当\ j=i\ 时,\end{cases}$$

得 $(v_m|e_i)=(a_1e_1+\cdots+a_me_m|e_i)=a_i$. 因此 $(v|e_i)=\lim\limits_{m\to\infty}(v_m|e_i)=a_i$. 于是有等式

$$a_i=(v|e_i). \tag{3.7.55}$$

由此可见, 如果展开式 (3.7.54) 存在, 则是唯一的.

数 $a_i=(v|e_i)$ 称为向量 v 对于标准正交系 (e_i) 的**傅里叶系数**. 它的几何意义是明显的:

$$a_i=(v|e_i)=\|v\|\,\|e_i\|\cos\widehat{ve_i}=\|v\|\,\cos\widehat{ve_i},$$

其中 $\widehat{ve_i}$ 是向量 v 和 e_i 的夹角. 从而, a_i 就是 v 在向量 e_i 方向上的投影.

因为

$$\begin{aligned}
0\leqslant\|v-v_m\|^2&=\left(v-\sum_{j=1}^m a_je_j\,\Big|\,v-\sum_{s=1}^m a_se_s\right)\\
&=\|v\|^2-\sum_j a_j(e_j|v)-\sum_s \bar{a}_s(v|e_s)+\sum_{j,s}c_j\bar{c}_s(e_j|e_s)\\
&=\|v\|-\sum_j a_j\overline{(v|e_j)}-\sum_s\bar{a}_sa_s+\sum_{j,s}a_j\bar{a}_s\delta_{js}\\
&=\|v\|^2-\sum_j a_j\bar{a}_j-\sum_s\bar{a}_sa_s+\sum_j a_j\bar{a}_j=\|v\|^2-\sum_j|a_j|^2,
\end{aligned}$$

所以, 永远有

$$\sum_{j=1}^m|a_j|^2\leqslant\|v\|^2.$$

由于这个不等式对任意的 m 成立, 让 m 趋于无穷大, 可以得出结论: 以 $|a_j|^2=|(v|e_j)|^2$ 为通项的无穷级数是收敛的, 而且满足不等式

$$\sum_{j=1}^\infty|(v|e_j)|^2\leqslant\|v\|^2. \tag{3.7.56}$$

这个不等式称为**贝塞尔不等式**.

二 现在看正交展开的存在性. 假设内积空间 V 对于其 (由内积确定的) 度量是完备的. 对 V 的完备的标准正交系 (e_i) 和向量 v, 考虑下面的级数

$$\sum_{i=1}^\infty a_ie_i, \tag{3.7.57}$$

其中 $a_i=(v|e_i)$. 我们将看到, 这个级数按照 V 的度量是收敛的, 而且收敛到 v.

设 v_m 是级数 (3.7.57) 前 m 项的和. 对两个正整数 p,q, 如果 $q>p$, 则有

$$\|v_q-v_p\|=\left\|\sum_{i=p+1}^q a_ie_i\right\|^2=\sum_{i=p+1}^q|a_i|^2.$$

由于以 $|a_i|^2$ 为通项的级数是收敛的, 所以当 p,q 充分大时, 上面的数充分小. 因此, $v_1, v_2, \cdots, v_m, \cdots$ 是柯西序列. 按照假定, V 是完备的, 所以当 $m\to\infty$ 时, v_m 有一个极限 $u \in V$. 我们证明 $u = v$. 根据引理 3.74, 对固定的 m 和 $p > m$, 有

$$\begin{aligned}(u|e_m) &= \lim_{p\to\infty}(v_p|e_m) = \lim_{p\to\infty}\left(\sum_{i=1}^{p} a_i e_i | e_m\right)\\ &= \lim_{p\to\infty} a_m = a_m = (v|e_m).\end{aligned}$$

因此, 对任意的 m, 有

$$(u - v|e_m) = (u|e_m) - (v|e_m) = 0. \tag{3.7.58}$$

由于标准正交系 (e_i) 是完备的, 由等式 (3.7.58) 得 $u = v$. 于是

$$v = \lim_{p\to\infty} v_p = \sum_{i=1}^{\infty} a_i e_i.$$

我们已经证明了下面的结论.

定理 3.75 在完备的内积空间 V 中, 对完备的标准正交系 (e_i), 向量 v 的傅里叶级数 $\sum_{i=1}^{\infty}(v|e_i)e_i$ 在 V 的度量下收敛到 v.

在这个定理的条件下, 对 $v = \sum_{i=1}^{\infty} a_i e_i$ 和 $u = \sum_{i=1}^{\infty} b_i e_i$, 运用引理 3.74 得

$$\begin{aligned}(v|u) &= \left(\lim_{p\to\infty}\sum_{i=1}^{p} a_i e_i \,\Big|\, \lim_{p\to\infty}\sum_{i=1}^{p} b_i e_i\right)\\ &= \lim_{p\to\infty}\left(\sum_{i=1}^{p} a_i e_i | \sum_{i=1}^{p} b_i e_i\right) = \lim_{p\to\infty}\sum_{i=1}^{p} a_i \bar{b}_i = \sum_{i=1}^{\infty} a_i \bar{b}_i.\end{aligned}$$

特别, 当 $u = v$ 时, 有

$$(v|v) = \|v\|^2 = \sum_{i=1}^{\infty} |a_i|^2.$$

于是, 正交展开式 (3.7.57) 把贝塞尔不等式 (3.7.56) 变成等式, 即, 向量 v 的长度的平方等于它在一个完备的标准正交系的所有方向 $e_1, e_2, \cdots, e_m, \cdots$ 上的投影的长度的平方和. 这是勾股定理在无穷维空间的推广.

三 在定理 3.75 中, 要求内积空间是完备的, 虽然很多的内积空间不是完备的, 例如 $C(a,b)$ 就不是, 但这其实并不要紧, 原因是下面的结论.

定理 3.76 对每个内积空间 V, 存在完备的内积空间 $\hat{V}$ 使得

(1) $\hat{V}$ 中有一个子空间 V_1 同构于 V;

(2) $\hat{V}$ 中每个向量都是 V_1 的某个柯西序列的极限. (这个性质称为 V_1 在 $\hat{V}$ 中**稠密**.)

空间 $\hat{V}$ 称为 V 的一个**完备化**. 我们将等同 V_1 和 V, 从而 V 是 $\hat{V}$ 的子空间.

证明 命 $\mathscr{C}$ 为 V 中的柯西序列全体. 对两个柯西序列 $\boldsymbol{x}=(x_1,x_2,\cdots)$ 和 $\boldsymbol{y}=(y_1,y_2,\cdots)$, 定义它们的加法及纯量 a 与 $\boldsymbol{x}$ 的乘法如下:

$$\boldsymbol{x}+\boldsymbol{y}=(x_1+y_1,x_2+y_2,\cdots),\qquad a\boldsymbol{x}=(ax_1,ax_2,\cdots).$$

容易验证, $\mathscr{C}$ 是向量空间. 命

$$\mathscr{C}_0=\{\boldsymbol{x}=(x_i)\in\mathscr{C}\mid \lim_{i\to\infty}\|x_i\|=0\}.$$

显然, $\mathscr{C}_0$ 是 $\mathscr{C}$ 的子空间. 命 $\hat{V}=\mathscr{C}/\mathscr{C}_0$. 这是 $\mathscr{C}$ 的一个商空间. 按惯例, 用记号 $\bar{\boldsymbol{x}}$ 表示含 $\boldsymbol{x}$ 的等价类. 对 $\boldsymbol{x}=(x_i)$ 和 $\boldsymbol{y}=(y_i)$, 我们有

$$\begin{aligned}|(x_n|y_n)-(x_m|y_m)|&=|(x_n|y_n-y_m)+(x_n-x_m|y_m)|\\&\leqslant\|x_n\|\,\|y_n-y_m\|+\|x_n-x_m\|\,\|y_m\|.\end{aligned}$$

因为 $\|x_i\|$ 和 $\|y_i\|$ 是有界的, 所以数列 $(x_1|y_1)$, $(x_2|y_2)$, $\cdots$, $(x_i|y_i)$, $\cdots$, 是柯西序列, 从而有极限. 我们定义

$$(\bar{\boldsymbol{x}}|\bar{\boldsymbol{y}})=\lim_{i\to\infty}(x_i|y_i).\tag{3.7.59}$$

这个定义是合理的, 因为如果 $\boldsymbol{x}-\boldsymbol{x}'\in\mathscr{C}_0$, $\boldsymbol{y}-\boldsymbol{y}'\in\mathscr{C}_0$, 由于 $\|x_i\|$ 和 $\|y_i'\|$ 是有界的, 我们有

$$\begin{aligned}|(x_i'|y_i')-(x_i|y_i)|&=|(x_i'-x_i|y_i')+(x_i|y_i'-y_i)|\\&\leqslant\|(x_i'-x_i\|\,\|y_i'\|+\|x_i\|\,\|y_i'-y_i\|\to 0,\end{aligned}$$

所以 $\lim\limits_{i\to\infty}(x_i'|y_i')=\lim\limits_{i\to\infty}(x_i|y_i)$.

直接验证可知 (3.7.59) 是 $\hat{V}$ 上的内积. 我们证明内积空间 $\hat{V}$ 具有要求的性质.

(a) $\hat{V}$ 中有一个子空间 V_1 同构于 V. 对任意向量 $v\in V$, $\boldsymbol{v}=(v,v,\cdots)$ 显然是柯西序列. 命

$$V_1=\{\bar{\boldsymbol{v}}\in\hat{V}\mid v\in V\}.$$

显然, V_1 与 V 作为内积空间是同构的.

(b) V_1 在 $\hat{V}$ 中是稠密的. 设 $\bar{\boldsymbol{x}}\in\hat{V}$, $\boldsymbol{x}=(x_i)$. 命 $\boldsymbol{x}_i=(x_i,x_i,\cdots)\in\mathscr{C}$. 那么 $\bar{\boldsymbol{x}}_i\in V_1$. 而且容易看出 $\bar{\boldsymbol{x}}_1,\bar{\boldsymbol{x}}_2,\cdots,\bar{\boldsymbol{x}}_i,\cdots$ 是 $\hat{V}$ 中的柯西序列. 按柯西序列的定义,

对每一个自然数 n, 都能找到自然数 k_n 使得 $\|x_m - x_i\| \leqslant 1/n$ 如果 $m, i \geqslant k_n$. 于是, 对 $m \geqslant k_n$, 有

$$\begin{aligned}\|\bar{\boldsymbol{x}}_m - \bar{\boldsymbol{x}}\|^2 &= (\bar{\boldsymbol{x}}_m - \bar{\boldsymbol{x}}|\bar{\boldsymbol{x}}_m - \bar{\boldsymbol{x}}) = \lim_{i\to\infty}(x_m - x_i|x_m - x_i)\\ &= \lim_{i\to\infty}\|x_m - x_i\|^2 \leqslant 1/n^2.\end{aligned}$$

所以 $\lim\limits_{m\to\infty}\|\bar{\boldsymbol{x}}_m - \bar{\boldsymbol{x}}\| = 0$, 即

$$\lim_{m\to\infty}\bar{\boldsymbol{x}}_m = \bar{\boldsymbol{x}}.$$

(c) $\hat{V}$ 是完备的. 设 $\xi_1, \xi_2, \cdots, \xi_i, \cdots$ 是 $\hat{V}$ 中的柯西序列. 根据 (b) 中的讨论, 对任意的项 ξ_i, 存在 $z_i \in V$ 使得 $\boldsymbol{z}_i = (z_i, z_i, \cdots)$ 所在的等价类 $\bar{\boldsymbol{z}}_i$ 与 ξ_i 的差的长度小于 $1/i$. 由于

$$\|z_i - z_j\| = \|\bar{\boldsymbol{z}}_i - \bar{\boldsymbol{z}}_j\| = \|\xi_i - \xi_j + (\bar{\boldsymbol{z}}_i - \xi_i) - (\bar{\boldsymbol{z}}_j - \xi_j)\| \leqslant \|\xi_i - \xi_j\| + \frac{1}{i} + \frac{1}{j},$$

所以序列 $\boldsymbol{z} = (z_1, z_2, \cdots)$ 是柯西序列. 命 $\eta = \bar{\boldsymbol{z}} \in \hat{V}$. 对任意小的正数 ε, 存在 $i_0 > 1/\varepsilon$, 使得当 $p, i > i_0$ 时有 $\|z_p - z_i\| < \varepsilon$. 于是, 我们有

$$\begin{aligned}\|\eta - \xi_i\| &\leqslant \|\eta - \bar{\boldsymbol{z}}_i\| + \|\bar{\boldsymbol{z}}_i - \xi_i\|\\ &\leqslant \lim_{p\to\infty}\|z_p - z_i\| + \frac{1}{i} < \varepsilon + \varepsilon = 2\varepsilon.\end{aligned}$$

从而 $\lim\limits_{i\to\infty}\xi_i = \eta$. 这样, $\hat{V}$ 中的每个柯西序列都在 $\hat{V}$ 中有极限. □

内积空间的完备化在同构的意义下是唯一的, 也就是说, 如果 $\hat{V}$ 和 W 是内积空间 V 的完备化, 那么 $\hat{V}$ 和 W 是同构的内积空间. 事实上, 设 $V_1 \subset \hat{V}$ 和 $W_1 \subset W$ 是同构于 V 的子空间, 那么 V_1 和 W_1 同构, 设 $\varphi_1 : V_1 \to W_1$ 是一个保持内积的线性同构. 这个同构可以导出内积空间 $\hat{V}$ 和 W 的同构. 原因如下. 由 V_1 在 $\hat{V}$ 中的稠密性知, 对任意的 $\xi \in \hat{V}$, 存在 V_1 中的柯西序列 $v_1, v_2, \cdots, v_i, \cdots$ 使得 $\lim\limits_{i\to\infty} v_i = \xi$. 由于 φ_1 是保持内积的, 所以 $\varphi_1(v_1), \varphi_1(v_2), \cdots, \varphi_1(v_i), \cdots$ 是 W_1 的柯西序列. 由于 W 是完备的, 所以这个柯西序列有极限, 记为 η. 定义 $\varphi(\xi) = \eta$. 显然, 如果 V_1 中的两个柯西序列的极限相等, 那么通过同构 φ_1 得到的 W_1 的柯西序列的极限也相等. 这样, 我们得到了一个映射 $\varphi : \hat{V} \to W$, 它是 φ_1 的延拓. 容易看出, φ 是一个双射, 而且保持内积.

四 为了方便地应用定理 3.75, 除了定理 3.76, 我们还需要下面两个引理.

引理 3.77 如果内积空间 V 的子空间 U 在 V 中稠密, 那么 U 在 $\hat{V}$ 中也稠密.

证明 沿用定理 3.76 的证明中的记号. 从定理 3.76 的证明中的 (b) 可知只要证明对 V 的任一个柯西序列 $\boldsymbol{x}=(x_1,x_2,\cdots,x_i,\cdots)$, 存在 U 中的柯西序列 $\boldsymbol{y}$ 使得 $\boldsymbol{x}-\boldsymbol{y}\in\mathscr{C}_0$, 从而在 $\hat{V}$ 中有 $\bar{\boldsymbol{x}}=\bar{\boldsymbol{y}}$.

由于 U 在 V 中稠密, 所以, 对每任意的 i, x_i 是 U 中的某个柯西序列的极限, 从而存在 $y_i\in U$ 使得 $\|x_i-y_i\|\leqslant 1/i$. 命 $\boldsymbol{y}=(y_1,y_2,\cdots,y_i,\cdots)$. 对任意的 $\varepsilon>0$, 存在整数 n_0 使得当 $i,j\geqslant n_0$ 时, 有 $\|x_i-x_j\|<\varepsilon/3$, $1/i<\varepsilon/3$, $1/j<\varepsilon/3$. 于是

$$\|y_i-y_j\|=\|y_i-x_i+x_i-x_j+x_j-y_j\|\leqslant\|y_i-x_i\|+\|x_i-x_j\|+\|x_j-y_j\|<\varepsilon.$$

所以 $\boldsymbol{y}$ 是 U 中的柯西序列. 显然 $\lim\limits_{i\to\infty}\|x_i-y_i\|=0$, 所以 $\boldsymbol{x}-\boldsymbol{y}\in\mathscr{C}_0$. □

回忆一下, 向量空间中含有无限个向量的向量组称为**线性无关的**如果向量组中任何有限个向量都是线性无关的. 设 $v_1,v_2,\cdots v_i,\cdots$ 是线性无关的向量组, 运用正交化方法, 可以得到一个标准正交系 $e_1,e_2,\cdots,e_i,\cdots$.

引理 3.78 设 V 是内积空间, $v_1,v_2,\cdots,v_i,\cdots$ 是其中的线性无关的向量组. 通过正交化方法得到标准正交系 $e_1,e_2,\cdots,e_i,\cdots$.

(1) 如果 $v_1,v_2,\cdots v_i,\cdots$ 张成的空间 $\langle v_1,v_2,\cdots v_i,\cdots\rangle$ 在 V 中是稠密的, 那么 $e_1,e_2,\cdots,e_i,\cdots$, 是 V 和 $\hat{V}$ 的完备标准正交系;

(2) 假设 V 是完备的, $e_1,e_2,\cdots,e_i,\cdots$ 是 V 的完备标准正交系, 那么 $\langle e_1,e_2,\cdots e_i,\cdots\rangle$ 在 V 中是稠密的 (等价的说法是 $\langle v_1,v_2,\cdots v_i,\cdots\rangle$ 在 V 中稠密).

证明 (1) 由正交化方法的过程知 $U=\langle e_1,e_2,\cdots e_i,\cdots\rangle$ 等于 $\langle v_1,v_2,\cdots v_i,\cdots\rangle$. 由引理 3.75, 如果这个子空间在 V 中稠密, 则在 $\hat{V}$ 中也稠密. 于是, 对任意的 $v\in\hat{V}$, 存在 U 中的柯西序列 $\xi_1,\xi_2,\cdots,\xi_i,\cdots$, 它以 v 为极限. 如果 v 和所有的 e_i 都正交, 那么, v 和所有的 ξ_i 都正交. 根据引理 3.72, 有

$$(v|v)=\lim_{i\to\infty}(v|\xi_i)=0.$$

这迫使 $v=0$. 所以 $e_1,e_2,\cdots,e_i,\cdots$ 是 $\hat{V}$ 的完备标准正交系, 当然更是 V 的完备标准正交系.

(2) 根据定理 3.75, $e_1,e_2,\cdots,e_i,\cdots$ 是 V 的完备标准正交系意味着 V 中的任意向量的傅里叶级数 $\sum\limits_{i=1}^{\infty}a_ie_i$ 收敛到这个向量, 因此是 $\langle e_1,e_2,\cdots e_i,\cdots\rangle$ 中的柯西序列的极限. 所以线性包络 $\langle e_1,e_2,\cdots e_i,\cdots\rangle$ 在 V 中是稠密的. □

推论 3.79 设 $e_1,e_2,\cdots,e_i,\cdots$ 是内积空间 V 的标准正交系. 如果这个标准正交系中的向量张成的空间在 V 中稠密, 那么任何向量 $v\in V$ 的傅里叶级数 $\sum\limits_{i=1}^{\infty}(v|e_i)e_i$ 是收敛的且收敛到 v.

证明 由引理 3.78(1) 和定理 3.75 推出. □

五 例 我们给出两个常用的例子.

例 3.80 根据魏尔斯特拉斯定理, 在闭区间 $[a,b]$ 上的任意连续函数 $f(t)$ 都是某个多项式序列 $P_i(t)$ 的极限, 而且, 当 $i\to\infty$ 时, $P_i(t)$ 是一致收敛到 $f(t)$ 的. 所以在 $[a,b]$ 上的连续函数空间 $C(a,b)$ 中, 多项式函数全体形成一个稠密子空间 $\mathbf{P}$. 显然, $\mathbf{P}$ 由 $1,t,\cdots,t^i,\cdots$ 张成. 对这组多项式实施正交化方法, 就得到 $C(a,b)$ 的一个完备标准正交系. 而且, 任何 $f\in C(a,b)$ 的傅里叶级数收敛到 f.

例 3.81 根据傅里叶级数理论, 三角函数系 $1,\cos t,\sin t,\cos 2t,\sin 2t,\cdots$ 在 $C(-\pi,\pi)$ 中是完备的. 而且, 这些函数张成的空间是稠密的. 标准化即得到 $C(-\pi,\pi)$ 的一个由三角函数构成的完备标准正交系

$$\frac{1}{\sqrt{2\pi}},\ \frac{\cos t}{\sqrt{\pi}},\ \frac{\sin t}{\sqrt{\pi}},\ \frac{\cos 2t}{\sqrt{\pi}},\ \frac{\sin 2t}{\sqrt{\pi}},\ \cdots.$$

习 题 3.7

1. 三角函数级数 $\displaystyle\sum_{k=1}^{\infty}\frac{\sin kt}{\sqrt{k}}$ 在 $\mathbb{R}$ 上收敛. 利用贝塞尔不等式证明它不是任何一个函数 $f\in C_2(-\pi,\pi)$(闭区间 $[-\pi,\pi]$ 上的二次可微函数全体) 的傅里叶级数.

2. 设 V 是完备的内积空间 (即希尔伯特空间), $\mathcal{A}$ 是其上的线性算子. 证明

$$\ker\mathcal{A}^*=(\operatorname{Im}\mathcal{A})^{\perp}.$$

3.8 正交投影和最小二乘法

一 正交投影 对向量空间 V 的子空间 U, 有很多的补空间, 每一个补空间 W 都确定了一个投影映射 $P:V\to V,\ u+w\to u$, 其中 $u\in U,\ w\in W$. 如果 V 是内积空间, U 是 V 的子空间, 那么 U 的正交补

$$U^{\perp}=\{x\in V\mid (x|u)=0\ \forall u\in U\}$$

是 U 的补空间. 由正交补确定的投影称为**正交投影**. 如果 U 是有限维的, $e_1,\cdots,e_n$ 是它的标准正交基, 那么正交投影 $\mathcal{P}_U$ 可以描述如下:

$$\mathcal{P}_U:V\to V,\quad v\to\sum_{i=1}^{m}(v|e_i)e_i. \tag{3.8.60}$$

二 正交投影与函数逼近及线性方程组的最小二乘解都密切相关. 在数学和其他科学如物理学、力学等中常会遇到用给定的函数的线性组合逼近某个函数. 如泰勒展开就是用多项式逼近无限可微函数 (即光滑函数), 傅里叶展开则是用正、余弦

函数的和逼近周期函数. 在逼近问题中, 显然, 判断哪个逼近最优是重要的问题. 最优的含义是什么? 对内积空间而言, 最优的含义很多时候指函数与被逼近的函数的差的模 (长度) 最小.

以 $V = C(a,b)$ 为例. 内积一般由下式给出:

$$(f|g) = \int_a^b f(t)g(t)dt.$$

对 V 中的函数 $f(t)$, 考虑用给定的函数 $\varphi_1(t), \cdots, \varphi_m(t)$ 的线性组合逼近, 要求选取系数 a_i 使得模平方 $\left\| f(t) - \sum\limits_{i=1}^m a_i\varphi_i(t) \right\|^2$ 最小. 问题的几何意义是明显的: 要在子空间 $U = \langle \varphi_1, \cdots, \varphi_m \rangle$ 中找一个向量 φ 使得 f 到 φ 的距离 $\|f - \varphi\|$ 最小.

这个问题可以称为**点到子空间的最短距离问题**, 或等价地, **垂直线问题**. 在直和分解 $V = U \oplus U^\perp$ 中, 有 $f = f_0 + f_1$, 其中 f_0 是 f 在 U 中的正交投影, $f_1 \in U^\perp$ 是以 "f 为起点到 U 的垂线" (不再有明显的几何直观), 或说是 f 在 $U^\perp$ 中的正交投影. 如果 $\varphi \in U$, $\varphi \neq f_0$, 那么

$$\|f - \varphi\| > \|f - f_0\|. \tag{3.8.61}$$

事实上, $f_0 - \varphi \in U$, 从而, 它与 $f_1 = f - f_0$ 正交. 由勾股定理知

$$\|f - \varphi\|^2 = \|f - f_0 + f_0 - \varphi\|^2 = \|f - f_0\|^2 + \|f_0 - \varphi\|^2.$$

由此得到不等式 (3.8.61). 所以, 垂直线问题就是寻找向量 f 的正交投影 f_0. 把 f_0 写成如下形式:

$$f_0 = x_1\varphi_1 + \cdots + x_m\varphi_m.$$

由条件

$$(f - f_0|\varphi_i) = 0, \quad i = 1, 2, \cdots, m,$$

得到线性方程组

$$x_1(\varphi_1|\varphi_i) + \cdots + x_m(\varphi_m|\varphi_i) = (f|\varphi_i), \quad 1 \leqslant i \leqslant m. \tag{3.8.62}$$

解这个线性方程组即得要求的系数 x_j.

如果 $\varphi_1, \cdots, \varphi_m$ 是标准正交系, 那么方程组 (3.8.62) 的解就是傅里叶系数:

$$x_j = (f|\varphi_j) = c_j, \quad 1 \leqslant j \leqslant m.$$

还容易看出, 此时有

$$\left\| f - \sum_{j=1}^m a_j\varphi_j \right\|^2 = \|f\|^2 - \sum_{j=1}^m |c_j|^2 + \sum_{j=1}^m |c_j - a_j|^2.$$

如果 $\varphi_1, \varphi_2 \cdots, \varphi_n, \cdots$ 是完备标准正交系且这些向量张成的子空间是稠密的, 根据推论 3.77 和定理 3.73, 得

$$f = \sum_{j=1}^{\infty}(f|\varphi_j)\varphi_j, \quad \|f\|^2 = \sum_{j=1}^{\infty}|(f|\varphi_j)|^2.$$

例 3.82①　找一个次数不超过 5 的实多项式, 使其在区间 $[-\pi, \pi]$ 上尽可能好地逼近正弦函数 $\sin t$. 为此, 考虑欧氏空间 $C(-\pi, \pi)$, 其内积定义为

$$(f|g) = \int_{-\pi}^{\pi} f(t)g(t)dt.$$

命 $U = \langle 1, t, t^2, t^3, t^4, t^5 \rangle$. 考虑 $\sin t$ 的正交投影. 可以先对 $1, t, t^2, t^3, t^4, t^5$ 实施正交化方法, 得到标准正交向量组 $e_1, e_2, e_3, e_4, e_5, e_6$. 于是, $v = \sin t$ 在 U 中的投影是 $P_U(v) = (v|e_1)e_1 + \cdots + (v|e_6)e_6$. 经计算, 并取 π 的近似值, 得

$$P_U(\sin t) = 0.987862t - 0.155271t^3 + 0.00564312t^5. \tag{3.8.63}$$

为了看出正交投影给出的逼近有多好, 我们把它与 $\sin t$ 的泰勒多项式给出的 5 次逼近比较:

$$t - \frac{1}{3!}t^3 + \frac{1}{5!}t^5.$$

在 0 附近, 两者都是 $\sin t$ 很好的逼近, 但当 $|t| > 2$ 时, 泰勒多项式就不是很好的逼近. 例如, 对 $t = 3$, 正交投影 (3.8.63) 给出的逼近与 $\sin 3$ 的误差大概是 0.001, 但 5 次泰勒多项式给出的逼近与 $\sin 3$ 的误差大概是 0.4, 比式 (3.8.63) 的误差大几百倍. 线性代数帮助我们发现了 $\sin t$ 的一个逼近, 它优于微积分中学到的泰勒多项式逼近!

三　最小二乘解　正交投影与最优逼近的联系可以带来审视线性方程组的新角度. 来源于实际生活的线性方程组由于测量、观察、数据处理等原因, 得到的线性方程组的系数和常数项都是近似值, 从而, 得到的方程组很可能不相容, 或者即便相容, 在系数和常数项都是近似值的情况下, 得到的解未必能符合实际. 于是, 需要寻找误差最小的解. 常用的方法就是最小二乘法, 得到的解称为最小二乘解.

定义 3.83　设 $A = (a_{ij})$ 是 $m \times n$ 实矩阵, $B = [b_1, \cdots, b_m] \in \mathbb{R}^m$, $X = [x_1, \cdots, x_n]$. 线性方程组

$$AX = B$$

的**最小二乘解**为 $X_0 \in \mathbb{R}^n$, 使得

$$\min_{X \in \mathbb{R}^n} \|AX - B\|^2 = \|AX_0 - B\|^2.$$

① 本例取自《线性代数应该这样学》(S. Axler, 人民邮电出版社, 第二版), 114-116.

符合条件的解可能不唯一. 其中模最小的解称为**模最小解**. 此处 $\mathbb{R}^m$ 的内积是标准的.

先看最小二乘解的几何意义. 矩阵 A 定义了一个线性映射:

$$\theta : \mathbb{R}^n \to \mathbb{R}^m, \quad X \to AX.$$

命 $U = \theta(\mathbb{R}^n) = \mathrm{Im}\,\theta$, 则有直和分解

$$\mathbb{R}^m = U \oplus U^{\perp}.$$

设 $B = B_1 + B_2$, 其中 B_1 是 B 在 U 中的正交投影, $B_2 \in U^{\perp}$. 对 $X \in \mathbb{R}^n$, 由勾股定理得

$$\|AX - B\|^2 = \|AX - B_1 - B_2\|^2 = \|AX - B_1\|^2 + \|B_2\|^2. \tag{3.8.64}$$

由此可见, 最小二乘解就是线性方程组

$$AX = B_1$$

的解. 根据构造, 这个方程组一定有解, 但解显然未必唯一.

由于 $AX = B_1$ 等价于 $AX - B \in U^{\perp}$, 子空间 $U = \theta(\mathbb{R}^n)$ 是 A 的列向量张成的空间, 所以最小二乘解就是如下线性方程组的解.

$$(x_1e_1 + \cdots + x_ne_n - B|e_i) = 0, \quad i = 1, \cdots, n,$$

其中 $e_1, \cdots, e_n$ 是 A 的列向量, 此处用到等式 $AX = x_1e_1 + \cdots + x_ne_n$. 写成下面的形式更习惯一些:

$$\begin{aligned}
(e_1|e_1)x_1 + (e_2|e_1)x_2 + \cdots + (e_n|e_1)x_n &= (B|e_1),\\
(e_1|e_2)x_1 + (e_2|e_2)x_2 + \cdots + (e_n|e_2)x_n &= (B|e_2),\\
&\cdots\cdots\\
(e_1|e_n)x_1 + (e_2|e_n)x_2 + \cdots + (e_n|e_n)x_n &= (B|e_n).
\end{aligned} \tag{3.8.65}$$

系数矩阵的行列式记作 $\det\|(e_i|e_j)\|$, 称为向量组 $e_1, \cdots, e_n$ 的**格拉姆行列式**. 这些向量线性无关当且仅当方程组 $AX = B_1$ 有唯一解, 即方程组 (3.8.65) 有唯一解, 这等价于系数矩阵非退化. 我们已经证明了如下结论.

定理 3.84 欧氏空间中的向量组 $e_1, \cdots, e_n$ 线性无关当且仅当它的格拉姆行列式不等于 0.

显然这个定理可以通过格拉姆–施密特正交化方法证明. 其实, 这个定理的主要部分很早就知道了: 正定对称双线性型在任何基下的矩阵是正定的, 特别, 其行

列式不等于 0. 线性相关时, 某个向量 e_i 是其余向量的线性组合, 相应地, 格拉姆行列式中的第 i 行是其余行的线性组合, 所以行列式等于零.

显然, 上面的讨论对复系数的线性方程组也是适用的, 当然需要把欧几里得空间 $\mathbb{R}^m$ 换成埃尔米特空间 $\mathbb{C}^m$.

习　题　3.8

对下面给定的 A 和 B, 求线性方程组 $AX = B$ 的最小二乘解.

1. $A = \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$.

2. $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$, $B = \begin{pmatrix} 2 \\ 5 \\ 6 \\ 6 \end{pmatrix}$.

3.9　正交多项式

例 3.82 和例 3.80 都说明正交多项式是很有意思的. 对多项式 $1, t, \cdots, t^m, \cdots$ 采用正交化方法就可以得到由多项式构成的完备标准正交系, 其中的多项式和它们的数值倍数都称为**正交多项式**. 在正交化方法的实际计算中, 刚开始不把得到的向量化成单位向量是方便的. 也就是说, 对线性无关的向量组 $v_1, v_2, \cdots, v_m, \cdots$, 按如下方法构造正交系:

$$e_1 = v_1,\ e_2 = v_2 - \frac{(v_2|e_1)}{\|e_1\|^2}e_1,\ \cdots,\ e_n = v_n - \sum_{i=1}^{n-1}\frac{(v_n|e_i)}{\|e_i\|^2}e_i,\ \cdots. \tag{3.9.66}$$

一　勒让德多项式　很容易构造 $C(a,b)$ 和 $C(c,d)$ 之间的同构. 下面考虑空间 $C(-1,1)$ 中的正交多项式. 对 $1, t, \cdots, t^m, \cdots$ 应用正交化方法 (3.9.66), 直接计算得

$$u_0(t) = 1,\ u_1(t) = t,\ u_2(t) = t^2 - \frac{1}{3},\ u_3(t) = t^3 - \frac{3}{5}t,\ \cdots. \tag{3.9.67}$$

这些多项式是法国数学家勒让德在 1785 年因考虑位势理论的问题而首先研究的. 但是, 这些多项式的一般公式迟至 1814 年才被罗德里格斯找到. 结果是, 多项式 $u_n(t)$ 乘上一个数值因子后, 等于多项式

$$p_n(t) = \frac{d^n}{dt^n}(t^2-1)^n, \quad n = 0, 1, 2, \cdots. \tag{3.9.68}$$

现在证明这个结论. 首先, $(t^2-1)^n$ 是 $2n$ 次多项式, 经过 n 次求导数后, 得到的多项式 $p_n(t)$ 是 n 次多项式. 所以有

$$V_n := \langle 1,\ t,\ \cdots,\ t^n\rangle = \langle p_0(t),\ p_1(t),\ \cdots,\ p_n(t)\rangle.$$

命 U_n 为 V_{n-1} 在 V_n 中的正交补, 那么 $u_n(t)\in U_n$, 且 $\dim U_n = \dim V_n - \dim V_{n-1} = 1$. 于是, 只要证明 $p_n(t)\in U_n$ 就知道它和 $u_n(t)$ 相差一个数值因子.

由于 ± 1 都是 $(t^2-1)^n$ 的 n 重根, 第一卷定理 7.11 说, 当 $k=0,1,\cdots,n-1$ 时, ± 1 都是 $\dfrac{d^k}{dt^k}(t^2-1)^n$ 的根. 假设 $0\leqslant k\leqslant n-1$, 下面计算 $p_n(t)$ 与 t^k 的内积. 用分部积分法, 得

$$\begin{aligned}
(t^k|p_n(t)) &= \int_{-1}^{1} t^k \frac{d^n}{dt^n}(t^2-1)^n dt\\
&= t^k \frac{d^{n-1}}{dt^{n-1}}(t^2-1)^n\bigg|_{-1}^{1} - k\int_{-1}^{1} t^{k-1}\frac{d^{n-1}}{dt^{n-1}}(t^2-1)^n dt\\
&= -kt^{k-1}\frac{d^{n-2}}{dt^{n-2}}(t^2-1)^n\bigg|_{-1}^{1} + k(k-1)\int_{-1}^{1} t^{k-2}\frac{d^{n-2}}{dt^{n-2}}(t^2-1)^n dt.
\end{aligned}$$

继续下去, 最后得到

$$(t^k|p_n(t)) = (-1)^k k!\frac{d^{n-k-1}}{dt^{n-k-1}}(t^2-1)^n\bigg|_{-1}^{1} = 0,\quad k=0,1,\cdots,n-1.$$

由此可见, 确有 $p_n(t)\in U_n$. 比较 $u_n(t)$ 和 $p_n(t)$ 的首项系数可得

$$u_n(t) = \frac{n!}{(2n)!}p_n(t),\quad n=0,1,2,\cdots.$$

看一下 $p_n(1)$ 的值. 为此, 把求乘积的 n 阶导数的莱布尼茨公式用到多项式 $(t^2-1)^n = (t-1)^n(t+1)^n$ 上, 得

$$\frac{d^n}{dt^n}(t^2-1)^n = \sum_{k=0}^{n}\binom{n}{k}\frac{d^k}{dt^k}(t-1)^n\cdot\frac{d^{n-k}}{dt^{n-k}}(t+1)^n.$$

当 $k<n$ 时, 多项式 $\dfrac{d^k}{dt^k}(t-1)^n$ 可以被 $t-1$ 整除, 从而, 它在 1 处的值为 0. 所以

$$p_n(1) = \binom{n}{n}\left[\frac{d^n}{dt^n}(t-1)^n\right](t+1)^n\Big|_{t=1} = 2^n\,n!.$$

为方便计算, 可以要求正交多项式在 1 处的值为 1. 这需要引入因子 $\dfrac{1}{2^n\,n!}$. 得到的多项式称为**勒让德多项式**. 用记号 $P_n(t)$ 表示 n 次勒让德多项式, 则

$$P_n(t) = \frac{1}{2^n\,n!}\frac{d^n}{dt^n}(t^2-1)^n.$$

可以验证，

$$\|P_n(t)\|^2 = \int_{-1}^{1} [P_n(t)]^2 dt = \frac{2}{2n+1}. \tag{3.9.69}$$

函数系 (3.9.67) 中的多项式

$$u_n(t) = \alpha_n P_n(t), \quad \alpha_n = \frac{2^n (n!)^2}{(2n)!},$$

具有如下有趣的极小性. 在所有的首项系数为 1 的实多项式中, 平方在区间 $-1 \leqslant t \leqslant 1$ 上的积分最小者是 $u_n(t)$.

事实上, 设 $f(t) = t^n + \cdots \in \mathbb{R}[t]$, 问题是要确定积分

$$I(f) = \int_{-1}^{1} f(t)^2 dt$$

的极小值. 把勒让德多项式之间的正交系和关于 $\|P_n(t)\|$ 的公式 (3.9.69) 应用到分解式

$$f(t) = \alpha_n P_n(t) + \beta_{n-1} P_{n-1}(t) + \cdots + \beta_1 P_1(t) + \beta_0, \quad \beta_i \in \mathbb{R}$$

上, 就可得到表达式

$$I(f) = \frac{2\alpha_n^2}{2n+1} + 2\sum_{i=0}^{n-1} \frac{\beta_i^2}{2i+1}.$$

显然, 在 $\beta_i = 0$, $0 \leqslant i \leqslant n-1$ 时, 上面的积分取极小值.

二 加权正交 很多时候需要考虑加权正交的多项式. 也就是说, 在区间 (a, b) 上给定一个非负 (连续) 函数 $p(t)$, 称为**权函数**. 考虑 $C(a, b)$ 的子空间

$$V(\sqrt{p(t)}) = \langle \sqrt{p(t)}t^k \mid k = 0, 1, 2, \cdots \rangle$$

或它的 n 维子空间

$$V_n(\sqrt{p(t)}) = \langle \sqrt{p(t)}t^k \mid 0 \leqslant k \leqslant n-1 \rangle$$

一个显然的问题就是求这些空间的标准正交基. 对向量组 $\sqrt{p(t)}$, $\sqrt{p(t)}t$, $\cdots$, $\sqrt{p(t)}t^m, \cdots$, 实施正交化方法, 得到满足条件

$$(\sqrt{p(t)}Q_k(t) \mid \sqrt{p(t)}Q_m(t)) = \int_a^b p(t)Q_k(t)Q_m(t)dt = \delta_{km}$$

的函数类

$$\{\sqrt{p(t)}Q_k(t) \mid \ Q_k(t) \in \mathbb{R}[t], \deg Q_k = k, k = 0, 1, 2, \cdots\}$$

可以称多项式 $Q_k(t),\ k=0,1,\cdots$ **是相应于权** $p(t)$ **的 (标准) 正交多项式**. 在这个意义下, 勒让德多项式就是相应于权 1 的正交多项式.

如果一开始就针对 $p(t)$ 引入新的内积

$$(f|g)_{p(t)}=\int_a^b p(t)f(t)g(t)dt,$$

那么, 前面意义下的加权正交 (多项式) 就是这个内积的正交 (多项式).

三　切比雪夫多项式　切比雪夫多项式在逼近论中有重要意义的. 第一类切比雪夫多项式 $T_n(t)$ 递归定义如下:

$$T_0(t)=1,\quad T_1(t)=t,\quad T_{n+1}(t)=2tT_n(t)-T_{n-1}(t).$$

由此可得

$$T_2(t)=2t^2-1,\quad T_3(t)=4t^3-3t,\quad T_4(t)=8t^4-8t^2+1.$$

一般项的公式是

$$T_n(t)=\frac{(-2)^n n!}{(2n)!}\sqrt{1-t^2}\frac{d^n}{dt^n}(1-t^2)^{\frac{n-1}{2}}=\cos(n\arccos t),\quad -1\leqslant t\leqslant 1.\qquad(3.9.70)$$

于是有

$$T_n(\cos t)=\cos nt.$$

第一类切比雪夫多项式是在区间 $[-1,1]$ 上相应于权 $p(t)=1/\sqrt{1-t^2}$ 的正交多项式, 因为有如下等式:

$$\int_{-1}^{1}\frac{T_k(t)T_m(t)dt}{\sqrt{1-t^2}}=\begin{cases}0, & 当\ k\neq m,\\ \pi/2, & 当\ k=m\neq 0,\\ \pi, & 当\ k=m=0.\end{cases}$$

等式的证明很简单: 命 $t=\cos x$, 再利用公式 $T_n(\cos t)=\cos nt$.

第二类切比雪夫多项式 $U_n(t)$ 递归定义如下:

$$U_0(t)=1,\quad U_1(t)=2t,\quad U_{n+1}(t)=2tU_n(t)-U_{n-1}(t).$$

由此可得

$$U_2(t)=4t^2-1,\quad U_3(t)=8t^3-4t,\quad U_4(t)=16t^4-12t^2+1.$$

一般项的公式是

$$U_n(t)=\frac{(t+\sqrt{t^2-1})^{n+1}(t-\sqrt{t^2-1})^{n+1}}{2\sqrt{t^2-1}}.\qquad(3.9.71)$$

于是有

$$U_n(\cos t) = \frac{\sin(n+1)t}{\sin t}.$$

第二类切比雪夫多项式则是在区间 $[-1,1]$ 上相应于权 $q(t) = \sqrt{1-t^2}$ 的正交多项式. 因为有如下等式:

$$\int_{-1}^{1} U_k(t)U_m(t)\sqrt{1-t^2}dt = \begin{cases} 0, & \text{当 } k \neq m, \\ \pi/2, & \text{当 } k = m. \end{cases}$$

两类切比雪夫多项式有很多的联系. 其中一个是

$$T_n(t) = U_n(t) - tU_{n-1}(t).$$

四　埃尔米特多项式　我们简单讨论一下埃尔米特多项式 $H_n(t)$(更恰当的名称应是拉普拉斯–切比雪夫–埃尔米特多项式). 它们也是一类加权正交多项式. 此时的区间是整个实数轴, 即 $a=-\infty$, $b=\infty$, 权函数是 $p(t)=e^{-t^2}$. 通过对单项式序列

$$1,\ 2t,\ 4t^2,\ \cdots,\ 2^n t^n,\ \cdots$$

实施加权正交化方法即可得到埃尔米特多项式, 即对序列

$$e^{-t^2/2},\ 2te^{-t^2/2},\ 4t^2e^{-t^2/2},\ \cdots,\ 2^n t^n e^{-t^2/2},\ \cdots$$

实施正交化方法 (3.9.66), 得到的函数再除以 $e^{-t^2/2}$. 前 5 个埃尔米特多项式是

$$H_0(t) = 1, \quad H_1(t) = 2t, \quad H_2(t) = 4t^2 - 2,$$
$$H_3(t) = 8t^3 - 12t, \quad H_4(t) = 16t^4 - 48t^2 + 12.$$

一般项的公式为

$$H_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n} e^{-t^2}.$$

加权内积是

$$\int_{\infty}^{\infty} e^{-t^2} H_k(t)H_m(t)dt = \begin{cases} 0, & \text{当 } k \neq m, \\ 2^n n!\sqrt{\pi}, & \text{当 } k = m. \end{cases}$$

这个积分的计算除了用分部积分外, 还用到下面的两个公式:

$$\int_{-\infty}^{\infty} e^{-t^2}dt = \sqrt{\pi}; \quad \lim_{t\to\infty} f(t)e^{-t^2} = 0, \quad \text{对任意多项式} f(t).$$

埃尔米特多项式也可以递归定义如下:

$$H_0(t) = 1, \quad H_1(t) = 2t, \quad H_{n+1}(t) = 2tH_n(t) - H_n'(t).$$

习 题 3.9

1. 证明勒让德多项式满足递归公式:

$$P_{n+1}(t)=\frac{2n+1}{n+1}tP_n(t)-\frac{n}{n+1}P_{n-1}(t).$$

2. 证明如下对勒让德多项式的积分很有用的公式:

$$(2n+1)P_n(t)=\frac{d}{dt}[P_{n+1}(t)-P_{n-1}(t)].$$

3. 利用归纳法和递归公式证明

$$P_n(t)=\sum_{k=0}^{n}(-1)^n\binom{n}{k}^2\left(\frac{1+t}{2}\right)^{n-k}\left(\frac{1-t}{2}\right)^k.$$

4. 证明: $f(x,t)=\dfrac{1}{\sqrt{1-2xt+x^2}}$ 是勒让德多项式的生成函数:

$$f(x,t)=\sum_{k=0}^{\infty}P_k(t)x^k.$$

5. 证明切比雪夫多项式的生成函数是

(1) $\displaystyle\sum_{k=0}^{\infty}T_k(t)x^k=\frac{1-tx}{1-2tx+x^2}$;

(2) $\displaystyle\sum_{k=0}^{\infty}U_k(t)x^k=\frac{1}{1-2tx+x^2}$.

6. 证明:

$$T_n(t)=\sum_{k=0}^{[n/2]}\binom{n}{2k}(t^2-1)^k t^{n2-k};$$

$$U_n(t)=\sum_{k=0}^{[n/2]}\binom{n+1}{2k+1}(t^2-1)^k t^{n-2k}.$$

7. 证明: $\max\limits_{-1\leqslant t\leqslant 1}|T_n(t)|=1$.

8. 证明: 两类切比雪夫多项式的所有根都是实的, 在区间 $[-1,1]$ 内, 且互不相同. 写出这些根的表达式.

9. 证明埃尔米特多项式满足递归公式

$$H_{n+1}(t)=2tH_n(t)-2nH_{n-1}(t).$$

10. 证明:

$$H_n(t)=n!\sum_{k=0}^{[n/2]}\frac{(-1)^k}{k!(n-2k)!}(2t)^{n-2k}.$$

11. 证明:

$$\exp(2tx-x^2)=\sum_{k=0}^{\infty}H_k(t)\frac{x^k}{k!}.$$

3.10　几个自伴随算子

对有限维内积空间, 自伴随可以对角化, 而且, 属于不同特征值的特征向量是正交的, 因而有由特征向量组成的正交基. 对无限维空间, 自伴随算子属于不同特征值的特征向量是正交的, 不过, 自伴随算子可以没有特征值和特征向量. 例如, 在空间 $C(a,b)$ 上, 用 t 相乘是一个算子, 记作 $\mathcal{F}_t$. 这个算子是自伴随的:

$$(\mathcal{F}_t f(t)|g(t)) = \int_a^b t f(t)g(t)dt = \int_a^b f(t)\cdot tg(t)dt = (f(t)|\mathcal{F}_t g(t)).$$

由于 $\mathcal{F}_t e(t) = \lambda e(t) \Rightarrow e(t) = 0$, 所以自伴随算子 $\mathcal{F}_t$ 没有特征向量. 这提示了无限维内积空间 (包括希尔伯特空间) 上线性算子的复杂性.

可是, 例 3.81 中的三角函数系中的三角函数都是算子 d^2/dt^2 的特征向量. 由此可见, 对无限维内积空间, 还是有些线性算子有充分多的特征向量, 也有由特征向量组成的完备正交基. 的确, 许多在数学上和物理上很重要的算子是自伴随的, 且有足够多的特征向量, 从而有特征向量组成的完备正交系.

一　考虑区间 $[-\pi,\pi]$ 上的二次连续可微函数空间 $C_2(-\pi,\pi)$, 带有通常的内积

$$(f|g) = \int_{\pi}^{\pi} f(t)g(t)dt.$$

算子

$$\mathcal{A} = \frac{d^2}{dt^2}$$

在其上不是自伴随的, 原因很简单, 它的特征向量 $\cos at$ 和特征向量 $\sin bt$ 不是正交的, 只要 a 和 b 中有一个不是整数.

考虑子空间

$$\Omega = \{f \in C''(a,b) \mid f(-\pi) = f(\pi),\ f'(-\pi) = f'(\pi)\}.$$

用分部积分法, 得

$$\begin{aligned}(\mathcal{A}f|g) &= \int_{\pi}^{\pi} f''(t)g(t)dt \\ &= f'(t)g(t)\big|_{-\pi}^{\pi} - \int_{\pi}^{\pi} f'(t)g'(t)dt = -\int_{\pi}^{\pi} f'(t)g'(t)dt \\ &= -f(t)g'(t)\big|_{-\pi}^{\pi} + \int_{\pi}^{\pi} f(t)g''(t)dt = \int_{\pi}^{\pi} f(t)g''(t)dt = (f|\mathcal{A}g).\end{aligned}$$

所以, 在空间 Ω 上算子 $\mathcal{A}$ 是自伴随的. 对它的特征值和特征向量能说什么呢? 设

$$\frac{d^2 f(t)}{dt^2} = \lambda f(t), \quad f \in \Omega. \tag{3.10.72}$$

对 $\lambda=-k^2$ $(k=0,1,2,\cdots)$, 函数族

$$a\cos kt+b\sin kt,\quad a,b\in\mathbb{R} \tag{3.10.73}$$

是方程 (3.10.72) 的解. 即, 它们是算子 $\mathcal{A}=d^2/dt^2$ 的属于特征值 $-k^2$ 的特征向量. 如果还有另外的特征值, 那么其特征向量与所有的三角函数 $\cos kt$, $\sin kt$ $(k=0,1,2,\cdots)$ 正交. 但这不会发生, 因为那些三角函数全体构成空间 Ω 的完备正交系. 同样的原因, 特征值 $-k^2$ 的特征子空间由函数族 (3.10.73) 构成, 即这些函数穷尽了方程 (3.10.72) 所有的解.

这样一来, 下面的定理成立.

定理 3.85 微分方程

$$\frac{d^2f(t)}{dt^2}=\lambda f(t)$$

如果要求其解是区间 $[-\pi,\pi]$ 上的二次连续可微函数且满足边界条件 $f(-\pi)=f(\pi)$, $f'(-\pi)=f'(\pi)$, 那么这个方程仅在 $\lambda=-k^2$ $(k=0,1,2,\cdots)$ 时有解, 解空间为 $\langle\cos kt,\ \sin kt\rangle$. 三角函数解 1, $\cos t$, $\sin t$, $\cdots$, 构成 $C(-\pi,\pi)$ 的一个完备正交系.

刚才的讨论说明微分算子的自伴随性和其作用的空间密切相关, 而且对很多的空间, 其伴随算子也不容易写出. 比较一般的微分算子具有如下形式:

$$\mathcal{D}=a_m(t)\frac{d^m}{dt^m}+a_{m-1}(t)\frac{d^{m-1}}{dt^{m-1}}+\cdots+a_1(t)\frac{d}{dt}. \tag{3.10.74}$$

它作用在区间 $[a,b]$ 上充分光滑的函数组成的空间. 在这个空间上并不容易写出 $\mathcal{D}$ 的伴随算子. 考虑满足条件 $f^{(k)}(a)=f^{(k)}(b)$, $k=0,1,\cdots,m-1$ 的函数 f 形成的子空间 Ω_m. 利用分部积分若干次可知, 作为空间 Ω_m 上的算子, $\mathcal{D}$ 的伴随算子是

$$\mathcal{D}^*=\sum_{i=1}^{m}(-1)^i\frac{d^i}{dt^i}\circ a_i(t), \tag{3.10.75}$$

其中算子 $\dfrac{d^i}{dt^i}\circ a_i(t)$ 作用在函数 $f(t)$ 处的值为 $\dfrac{d^i}{dt^i}[a_i(t)f(t)]$, 即先用 $a_i(t)$ 乘, 再求 i 次导数. 由此很容易判断一些微分算子 $\mathcal{D}$ 在空间 Ω_m 上是自伴随的.

二 我们感兴趣勒让德多项式是否为某个自伴随算子的特征向量. 设 $w_n=(t^2-1)^n$. 对恒等式

$$(t^2-1)\frac{d}{dt}w_n=2ntw_n$$

的两端求 $n+1$ 次导数, 利用莱布尼茨公式得

$$\begin{aligned}&(t^2-1)\frac{d^{n+2}}{dt^{n+2}}w_n+2(n+1)t\frac{d^{n+1}}{dt^{n+1}}w_n+n(n+1)\frac{d^n}{dt^n}w_n\\=&2nt\frac{d^{n+1}}{dt^{n+1}}w_n+2n(n+1)\frac{d^n}{dt^n}w_n.\end{aligned}$$

这个等式两端均乘以 $1/(2^n n!)$, 并使用记号

$$\frac{1}{2^n n!}\frac{d^n}{dt^n}w_n = P_n(t),$$

可得到微分关系式

$$(t^2-1)\frac{d^2}{dt^2}P_n(t) + 2t\frac{d}{dt}P_n(t) - n(n+1)P_n(t) = 0. \tag{3.10.76}$$

和过去一样, 区间 $[-1,1]$ 上二次连续可微函数形成的空间记作 $C_2''(-1,1)$, 它的内积 (参见 3.8 节中第二部分第二段) 我们已经多次讨论. 在这个空间上研究线性微分算子

$$\mathcal{S} = (t^2-1)\frac{d^2}{dt^2} + 2t\frac{d}{dt} = \frac{d}{dt}\left[(t^2-1)\frac{d}{dt}\right].$$

由于 t^2-1 在 ± 1 处的值为 0, 利用分部积分公式知 $\mathcal{S}$ 是自伴随算子. 公式 (3.10.76) 可以改写成

$$\mathcal{S}P_n(t) = n(n+1)P_n(t). \tag{3.10.77}$$

所以, 多项式 $P_n(t)$ 是自伴随算子 $\mathcal{S}$ 的属于特征值 $\lambda = n(n+1)$ 的特征函数. 换句话说, 微分方程 $\mathcal{S}y(t) = n(n+1)y(t)$ 有非零解 $y(t) = P_n(t)$. 解空间 (或说 $\mathcal{S}$ 的特征子空间)V^λ 的维数是 1, 否则 V^λ 中就会有非零向量 $x(t)$ 与 $P_n(t)$ 正交. 因为自伴随算子的特征子空间互相正交:

$$(V^\lambda \,|\, V^\mu) = 0, \quad \lambda \neq \mu,$$

从而, 向量 $x(t)$ 与所有的勒让德多项式正交. 但这样的函数不存在, 因为根据例 3.80, 勒让德多项式全体形成 $C(-1,1)$ 中的完备正交系. 同样的原因, 算子 $\mathcal{S}$ 没有异于 $n(n+1)$, $n = 0,1,2,\cdots$ 的特征值.

上面的讨论已经证明了如下断言.

定理 3.86　定义于区间 $-1 \leqslant t \leqslant 1$ 上的二次连续可微函数全体上的微分方程

$$\frac{d}{dt}\left[(t^2-1)\frac{d}{dt}y(t)\right] = \lambda y(t),$$

仅在 $\lambda = n(n+1)$, $n = 0,1,2,\cdots$ 时有解. 在差一个常数因子的意义下, 每个 n 都对应唯一的解 $y(t) = P_n(t)$. 这些解全体在 $C(-1,1)$ 中构成一个完备正交系.

我们简短讨论过的算子 $\dfrac{d^2}{dt^2}$ 和 $\dfrac{d}{dt}(t^2-1)\dfrac{d}{dt}$ 都是**斯图姆–刘维尔算子**的特殊情况. 后者定义在空间 $C''(a,b)$ 上, 由公式

$$\mathcal{L}y(t) = \frac{d}{dt}\left[p(t)\frac{d}{dt}y(t)\right] + q(t)y(t)$$

确定. 这类算子在数学和物理中都十分重要.

三　在数学物理中, 所谓的**埃尔米特函数**

$$\psi_n(t) = e^{-t^2/2}H_n(t) = (-1)^n e^{t^2/2}\frac{d^n}{dt^n}e^{-t^2}$$

是很有用的. 我们证明它是算子

$$\mathcal{H} = \frac{d^2}{dt^2} - t^2$$

的属于特征值 $-(2n+1)$ 的特征向量. 为此, 考虑辅助算子

$$\mathcal{M} = \frac{d}{dt} - t.$$

容易验证

$$[\mathcal{H}, \mathcal{M}] = \mathcal{HM} - \mathcal{MH} = -2\frac{d}{dt} - t = -2\mathcal{M}.$$

由此可见, 如果 f 是算子 $\mathcal{H}$ 的属于特征值 λ 的特征向量, 那么 $\mathcal{M}f$ 是算子 $\mathcal{H}$ 的属于特征值 $\lambda - 2$ 的特征向量:

$$\mathcal{HM}f = [\mathcal{H}, \mathcal{M}] + \mathcal{MH}f = -2\mathcal{M}f + \lambda\mathcal{M}f = (\lambda-2)\mathcal{M}f.$$

对 n 做归纳法, 可以推出等式

$$\mathcal{HM}^n f = (\lambda - 2n)\mathcal{M}^n f.$$

因为 $\mathcal{H}e^{-t^2/2} = -e^{-t^2/2}$, 所以, 对任意的整数 $n \geqslant 0$, $\mathcal{M}^n e^{-t^2/2}$ 是算子 $\mathcal{H}$ 的属于特征值 $-(2n+1)$ 的特征向量.

另一方面, 对 n 归纳, 可以直接算出

$$e^{t^2/2}\frac{d^n}{dt^n}e^{-t^2} = \mathcal{M}^n e^{-t^2/2},$$

也就是

$$\psi_n(t) = (-1)^n\mathcal{M}^n e^{-t^2/2}.$$

我们要证明的结论已经清楚了.

习　题　3.10

1. 证明: 第一类切比雪夫多项式 $T_n(t)$ 是微分算子

$$(t^2-1)\frac{d^2}{dt^2} + t\frac{d}{dt}$$

的一个属于特征值 n^2 的特征向量.

2. 证明: 第二类切比雪夫多项式 $T_n(t)$ 是微分算子

$$(t^2-1)\frac{d^2}{dt^2}+3t\frac{d}{dt}$$

的一个属于特征值 $n(n+2)$ 的特征向量.

第 4 章　仿射空间与欧几里得仿射空间

向量的本意是有方向的线段，方向相同、长度一致的向量被认为是相等的. 向量空间讨论的是有向线段的等价类，每个类由方向相同、长度相等的有向线段构成，而且，每个类用其中以原点为起点的有向线段的终点表示. 同时，在向量空间中，讨论的子空间都是过原点的，包括线与面等. 可是，在现实生活中，空间由点构成，不同位置的线段是需要区分的. 现实的空间中没有原点，即便选了原点，绝大部分线、面等不过原点. 这一切说明需要一个更方便讨论现实世界的空间，那就是仿射空间与欧几里得仿射空间.

4.1　仿 射 空 间

仿射空间是点的集合，空间中有线、面等. 仿射几何关注点、线、面之间的关系 (如点共线、线共点、平行等). 仿射几何可以像中学的几何那样：建立一套公理后讨论点、线、面及其关系. 这里用线性代数的方法处理，本质的原由是两点确定一个向量.

一　设 V 是域 K 上的向量空间.

定义 4.1　称非空集合 $\mathbb{A}$ 是与向量空间 V **关联**(或说**相伴**)的**仿射空间**如果有一个加法映射 $\mathbb{A} \times V \to \mathbb{A}$, $(p, x) \to p + x$, 满足如下条件：

(1) $(p+x)+y = p+(x+y)$, $p \in \mathbb{A}$, $x, y \in V$;

(2) $p + 0 = p, p \in S$, 0 是零向量;

(3) 对任意的 $p, q \in \mathbb{A}$, 存在唯一的向量 $x \in V$ 使得 $p + x = q$.

图 4.1

集合 $\mathbb{A}$ 中的元素称为点. 条件 (3) 中的向量 x 称为**连接点** p **和** q **的向量**，记作 $\overrightarrow{pq}$, 也记作 $q-p$. 仿射空间 $\mathbb{A}$ 的**维数**定义为它关联的向量空间 V 的维数 n. 有时，用记号 $\mathbb{A}^n$ 表明仿射空间 $\mathbb{A}$ 是 n 维的. 无论是理论还是应用，最令人感兴趣是 $K = \mathbb{R}$ 和 $\mathbb{C}$ 的情形，相应的仿射空间分别称为**实仿射空间**和**复仿射空间**.

例 4.2　每个向量空间 V 天然是仿射空间, 其中的元素既看作向量也看作点, 点与向量的加法就是向量间的加法. 此时, 向量 $\overline{pq}$ 是向量 p 和 q 的差 $q-p$.

例 4.3　设 V 是向量空间, $v_0 \in V$ 是取定的向量, U 是 V 的子空间. 那么 U 的陪集 $\mathbb{A} = v_0 + U = \{v_0 + u \mid u \in U\}$ 是与 U 关联的仿射空间, 加法映射是显然的:

$$(v_0 + U) \times U \to v_0 + U, \quad (v_0 + u) + x \to v_0 + (u + x).$$

常称 $\mathbb{A}$ 为 V 的一个**仿射线性流形或线性流形**, 而子空间 U 称为线性流形 $\mathbb{A}$ 的**方向**.

对仿射空间 $\mathbb{A}$ 中的点 p, q, 用记号 $\overline{pq}$ 或 $q-p$ 表示连接 p 和 q 的向量是很方便和直观的, 且便于记忆和计算. 由定义直接得出如下运算规则:

$$p + \overline{pq} = q, \quad \overline{pq} + \overline{qr} = \overline{pr}, \quad \overline{pq} = -\overline{qp}, \quad \overline{pp} = 0,$$

其中 p, q, r 是 $\mathbb{A}$ 中的任意点. 这些规则的另一个写法是

$$p + (q-p) = q, \quad (q-p) + (r-q) = r - p, \quad q - p = -(p-q), \quad p - p = 0.$$

二　仿射空间的定义中的加法允许点 p 和向量 v 相加, 结果仍是一个点. 固定一个点 $p \in \mathbb{A}$, 让向量 v 变化, 就得到一个映射 $V \to \mathbb{A}$, $v \to p + v$. 定义中的条件 (3) 表明这个映射是双射. 如果固定向量 v, 让点 p 变化, 则得到双射

$$t_v : \mathbb{A} \to \mathbb{A}, \quad p \to p + v = t_v(p),$$

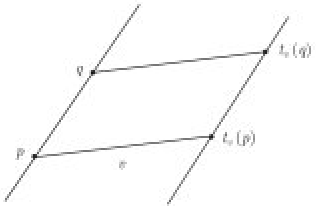


图 4.2　平移

称为**沿着向量 v 的平移 (算子)**, 也称为**由向量 v 确定的平移 (算子)**, 或简单称作 $\mathbb{A}$ 的一个**平移(算子)**. t_v 对 p 的作用称为**用向量 v 平移 p**. 由定义 4.1(1) 和 (2) 知

$$t_u \cdot t_v = t_{u+v}, \quad t_{-v} \cdot t_v = e$$

($e := t_0$ 是恒等映射). 也就是说, 两个平移的合成还是平移, 平移的逆映射也是平移. 所以, 平移全体形成一个群, 它同构于向量空间 V 的加法群. 如果令

$$\alpha t_u+\beta t_v=t_{\alpha u+\beta v},$$

那么, $\mathbb{A}$ 的平移全体就形成一个向量空间. 这个向量空间由 $\mathbb{A}$ 唯一确定且与 V 同构, 记作 $\mathbb{A}^\sharp$.

三 仿射空间之间的映射我们主要关注那些能保持运算的映射.

定义 4.4 设 $\mathbb{A}$ 和 $\mathbb{A}'$ 是域 K 上的仿射空间, 分别关联到向量空间 V 和 V'. 称映射 $f:\mathbb{A}\to\mathbb{A}'$ 是**仿射映射**(或**仿射线性映射**) 如果它满足条件

$$f(p+v)=f(p)+Df\cdot v,\quad p\in\mathbb{A},\ v\in V, \tag{4.1.1}$$

其中 $Df:V\to V'$ 是线性映射. 有时, 也称映射 Df 是映射 f 的**线性部分**(或**微分**). 如果 f 是双射, 则称 f 是 (**仿射**)**同构**, 并说 $\mathbb{A}$ 和 $\mathbb{A}'$ 是**同构的**. 从 $\mathbb{A}$ 到自身的仿射同构称为$\mathbb{A}$ **的自同构**, 也称为**仿射变换**.

例 4.5 (1) 把 $\mathbb{A}$ 映到一个点的映射是仿射映射, 称为**常值映射**, 其微分是零映射.

(2) 设 $\mathbb{A}=\mathbb{A}'=K$, 那么映射 $x\to ax+b$ 是仿射映射, 其微分是映射 $x\to ax$. 当 $a\neq 0$ 时, 它是仿射同构.

例 4.6 设 o 是 $\mathbb{A}$ 中的点, $\lambda\in K$. 定义 $f:\mathbb{A}\to\mathbb{A}$, $o+v\to o+\lambda v$, 即 $\overline{of(p)}=\lambda\overline{op}$, $\forall p\in\mathbb{A}$. 那么 f 是仿射映射, 称为**以 o 为中心的伸缩变换**(dilatation; 或**相似变换**, homothety), 伸缩率为 λ, 有时简单称为**伸缩变换**. 见图 4.3. 当 $\lambda\neq 0$, 伸缩变换是仿射自同构.

如果用 q 表示 $p+v$, 那么 $v=\overline{pq}$, 从而 (4.1.1) 可以写成

$$Df\cdot\overline{pq}=\overline{f(p)f(q)}. \tag{4.1.2}$$

所以, Df 由 f 唯一确定.

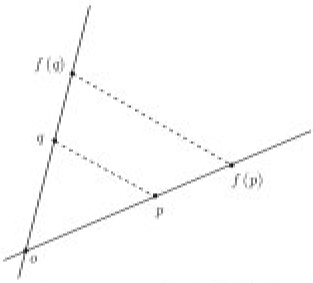


图 4.3 以 o 为中心的伸缩变换

命题 4.7　仿射映射 f 是双射当且仅当其微分 Df 是双射.

证明　当 Df 是双射时, 显然 f 是双射. 反之, 如果 f 是双射, 取定 $\mathbb{A}$ 中的一个点 o. 根据 (4.1.2), 对任意 $p \in \mathbb{A}$ 有 $Df \cdot \overline{op} = \overline{f(o)f(p)}$. 而且当 $p \neq p'$ 时, 有 $\overline{op} \neq \overline{op'}$, $\overline{f(o)f(p)} \neq \overline{f(o)f(p')}$. 由于 $V = \{\overline{op} | p \in \mathbb{A}\}$, 且 f 是双射意味着 $\mathbb{A}' = \{f(p) | p \in \mathbb{A}\}$, 从而 $V' = \{\overline{f(o)f(p)} | p \in \mathbb{A}\}$. 所以 $D(f)$ 是双射.　□

定理 4.8　域 K 上的两个仿射空间 $\mathbb{A}$ 和 $\mathbb{A}'$ 同构当且仅当它们的维数相等.

证明　如果仿射映射 $f : \mathbb{A} \to \mathbb{A}'$ 是同构, 那么 Df 是双射, 所以这两个仿射空间关联的线性空间是同构的, 从而它们的维数相等.

现假设 $\mathbb{A}$ 和 $\mathbb{A}'$ 的维数相等, 它们关联的向量空间分别是 V 和 V'. 那么, V 和 V' 的维数相同, 所以有线性同构 $\mathcal{F} : V \to V'$. 取定 $o \in \mathbb{A}$ 和 $o' \in \mathbb{A}'$. 由于任意一点 $p \in \mathbb{A}$ 能以唯一的方式写成 $p = o + v$, 我们可以定义映射

$$f : \mathbb{A} \to \mathbb{A}', \quad o + v \to o' + \mathcal{F}v.$$

如果 $o+v$ 取遍 $\mathbb{A}$ 中的点, v 则走遍 V 中的向量, 由于 $\mathcal{F}$ 是双射, 所以 $o' + \mathcal{F}v$ 取遍 $\mathbb{A}$ 中的点. 同理, f 把 $\mathbb{A}$ 中不同的点映到 $\mathbb{A}'$ 中不同的点. 从而, f 是双射. 剩下的事情是验证映射的仿射性. 从映射 f 的定义得

$$\begin{aligned} f(p+u) &= f((o+v)+u) = f(o+(v+u)) \\ &= o' + \mathcal{F}(v+u) = o' + (\mathcal{F}v + \mathcal{F}u) \\ &= (o' + \mathcal{F}v) + \mathcal{F}u = f(p) + \mathcal{F}u. \end{aligned}$$

所以, f 是以 $\mathcal{F}$ 为其微分的仿射映射.　□

四　坐标　一个点 $o \in \mathbb{A}$ 与 V 的一个基 $e_1, \cdots, e_n$ 合在一起称为 $\mathbb{A}$ 的一个**坐标系**(或**坐标架**). 向量 $\overline{op}$ 在基 $e_1, \cdots, e_n$ 下的坐标称为点 p 在坐标系 $\{o; e_1, \cdots, e_n\}$ 下的**坐标**. 点 o 称为这个坐标系的**原点**.

说明　坐标系也可以用 $n+1$ 个点 $p_0, p_1, \cdots, p_n$ 给出, 只要向量组 $\overline{p_0p_1}, \cdots, \overline{p_0p_n}$ 是 V 的一个基.

在仿射空间 $\mathbb{A}$ 中取定一个点 o(原点), 那就可以把点 p 与向量 $\overline{op}$ 等同起来. 然后, 点与向量的加法就是向量间的加法. 这种把点与向量等同的做法称为仿射空间的**向量化**. 当然, 这依赖原点的选取. 取定 n 维仿射空间的一个坐标系, 点与坐标的对应就是仿射空间到空间 K^n 的同构.

由等式 $\overline{pq} = \overline{oq} - \overline{op}$ 知下面的结论成立.

命题 4.9　给定仿射空间 $\mathbb{A}$ 的一个坐标系 $\{o; e_1, \cdots, e_n\}$.

(1) 如果 $x_1, \cdots, x_n$ 是点 p 的坐标, $y_1, \cdots, y_n$ 是点 q 的坐标, 那么 $\overline{pq}$ 的坐标就是 $y_1 - x_1, \cdots, y_n - x_n$; 而且

(2) 对任意向量 $v = a_1e_1 + \cdots + a_ne_n$, 点 $p+v$ 的坐标是 $x_1+a_1, \cdots, x_n+a_n$.

经常需要考虑点在不同坐标系之间的联系. 设 $\{o; e_1, \cdots, e_n\}$ 和 $\{o'; e_1, \cdots, e_n\}$ 是仿射空间 $\mathbb{A}$ 的两个坐标系 (图 4.4).

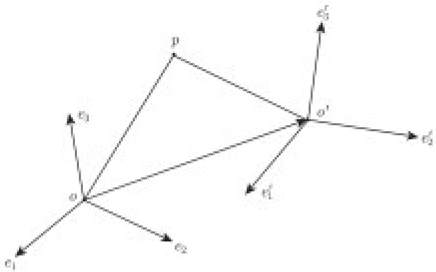


图 4.4 坐标变换

假设 $\mathbb{A}$ 中的点 p 在两个坐标系下的坐标分别是 $x_1, \cdots, x_n$ 和 $x_1', \cdots, x_n'$. 又假定点 o' 在前一个坐标系下的坐标是 $b_1, \cdots, b_n$, 空间 V 的基 $e_1, \cdots, e_n$ 到基 $e_1', \cdots, e_n'$ 的转换矩阵是 A. 由等式

$$\begin{aligned}\sum_i x_ie_i = \overline{op} = \overline{oo'} + \overline{o'p} &= \sum_i b_ie_i + \sum_j x_j'e_j' \\ &= \sum_i b_ie_i + \sum_j x_j' \sum_i a_{ij}e_i \\ &= \sum_i \left(\sum_j a_{ij}x_j'\right) e_i + \sum_i b_ie_i\end{aligned}$$

知

$$x_i = \sum_{j=1}^{n} a_{ij}x_j' + b_i, \quad i = 1, \cdots, n. \tag{4.1.3}$$

这些等式写成如下矩阵的形式更便于计算:

$$X = AX' + B,$$

其中 $X = [x_1, \cdots, x_n]$, $X' = [x_1', \cdots, x_n']$, $B = [b_1, \cdots, b_n]$. 由于 A 是可逆矩阵, 所以

$$X' = A^{-1}X + A^{-1}B.$$

五 仿射子空间 例 4.3 实际上是一个不过原点的仿射子空间. 现在对这类空间做进一步的讨论.

定义 4.10　设 p 是仿射空间 $(\mathbb{A}, V)$ 的点, U 是 V 的线性子空间. 那么集合

$$\Pi = p + U = \{p + u \mid u \in U\} \tag{4.1.4}$$

称为 $\mathbb{A}$ 的一个 $\dim U$ 维**平面**(或**仿射子空间**). 子空间 U 称为平面 Π 的**方向子空间**, Π 也称为过点 p 的以 U 为方向的平面或仿射子空间. 自然, 0 维的平面就是点, 1 维的平面是直线, $\dim V - 1$ 维的平面称为超平面 (与向量空间中的术语是相对应的).

下面的结论是意料之中的.

定理 4.11　一个仿射空间 $\mathbb{A}$ 中的平面 $\Pi = p + U$ 本身也是仿射空间, 其关联的向量空间为 U.

证明　映射 $\Pi \times U \to \Pi$, $(q, u) \to q + u$, 显然满足仿射空间定义中的条件 (1) 和 (2). 对 Π 中的两个点 $q = p + u$ 和 $r = p + w$, 有

$$r = p + (u + w - u) = p + u + (w - u) = q + (w - u).$$

由于 $w - u = \overline{qr} \in U$ 在 V 中是唯一的, 所以在 U 中更是唯一的, 于是仿射空间定义中的条件 (3) 也成立. □

因为对任意 $u_0 \in U$, 有 $(p + u_0) + U = p + U$, 所以可取 Π 中的任何点作为式 (4.1.4) 中的 p. 从这个简单的事实可以得到下面的结论.

命题 4.12　仿射空间 $\mathbb{A}$ 的两个平面的交集或是空集, 或是 $\mathbb{A}$ 的平面.

证明　设 $\Pi = p + U$ 和 $\Pi' = p' + U'$ 是 $\mathbb{A}$ 的两个平面. 如果它们的交集非空, 那么存在 $p_0 \in \Pi \cap \Pi'$. 于是 $\Pi = p_0 + U$, $\Pi' = p_0 + U'$. 由此可见 $\Pi \cap \Pi' = p_0 + U \cap U'$ 是 $\mathbb{A}$ 的平面. □

从定理 4.11 的证明可以看出, 对平面 $\Pi = p + U$ 中的任意两点 q, r 有 $\overline{qr} \in U$. 对任意的 $u \in U$, 命 $q = p + u$, 则有 $u = \overline{pq}$. 所以有

$$U = \{\overline{qr} \mid q, r \in \Pi\} = \{\overline{pq} \mid q \in \Pi\}. \tag{4.1.5}$$

对 Π 中不同的点 q, r, 由于 $\Pi = q + U$, 所以过 q, r 的直线

$$L_{q,r} = q + \{\lambda \overline{qr} \mid \lambda \in K\} \tag{4.1.6}$$

在 Π 中. 如果基域 K 的特征不等于 2, 这个结论的逆也是对的. 也就是说, 有如下结论.

定理 4.13　假设 $\operatorname{char} K \neq 2$. 那么仿射空间 $\mathbb{A}$ 的子集 Π 是平面, 当且仅当, 子集中任意不同的两点 q, r 确定的直线 $L_{q,r}$ 在子集中.

证明 仅需证明充分性. 取 $p \in \Pi$, 由于 $p + \overline{pq} = q$, 所以

$$\Pi = p + \{\overline{pq} \mid q \in \Pi\}.$$

需要证明 $U = \{\overline{pq} \mid q \in \Pi\}$ 是 V 的线性子空间. 由假设, 过 p, q 的直线

$$L_{p,q} = p + \{\lambda\overline{pq} \mid \lambda \in K\}$$

在 Π 中. 所以, 对任意的 $\lambda \in K$, $t = p + \lambda\overline{pq}$ 是 Π 中的点. 于是 $\lambda\overline{pq} = \overline{pt} \in U$. 即 U 对纯量乘是封闭的. 下面证明 U 对于加法是封闭的. 设 $q, r \in \Pi$. 由假设, 过 q, r 的直线

$$L_{q,r} = q + \{\lambda\overline{qr} \mid \lambda \in K\}$$

在 Π 中. 于是, 点 $s = q + \dfrac{1}{2}\overline{qr}$ 在 Π 中. 这意味着 $\overline{ps}$ 在 U 中. 因为 U 对纯量乘封闭, 所以 $2\overline{ps}$ 也在 U 中. 我们有

$$2\overline{ps} = 2\overline{pq} + 2\overline{qs} = 2\overline{pq} + \overline{qr} = \overline{pq} + \overline{pr} \in U.$$

所以, U 对加法封闭. □

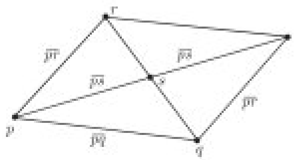


图 4.5

利用这个定理可以给命题 4.12 另一个证明.

六 对仿射空间 $\mathbb{A}$ 的子集 M 和任意的点 $p_0 \in M$, 平面

$$A(M) = p_0 + \langle \overline{p_0p} \mid p \in M \rangle$$

是包含 M 的最小的平面, 从而与点 p_0 的选取无关. 这个平面称为集合 M 的**仿射包络**.

定理 4.14 对仿射空间 $\mathbb{A}$ 中任意 $k+1$ 个点, 存在一个维数 $\leqslant k$ 的平面过这些点 (即包含这些点). 如果任何维数小于 k 的平面都不能全包含这些点, 那么存在唯一的 k 维平面过这些点.

证明 设 $p_0, p_1, \cdots, p_k$ 是 $\mathbb{A}$ 中的点, 那么

$$\Pi = p_0 + \langle \overline{p_0p_1}, \cdots, \overline{p_0p_k} \rangle$$

是过点 $p_0, p_1, \cdots, p_k$ 的平面, 且维数 $\leqslant k$. 如果 Π 的维数是 k, 那么 $\overrightarrow{p_0p_1}, \cdots, \overrightarrow{p_0p_k}$ 线性无关. 根据式 (4.1.5), 过这些点的平面的方向子空间都含有 $\overrightarrow{p_0p_1}, \cdots, \overrightarrow{p_0p_k}$, 所以 Π 是唯一的过点 $p_0, p_1, \cdots, p_k$ 的 k 维平面. □

例 4.15　设 p, q 是二维仿射空间 $\mathbb{A}$ 中两个不同的点, 那么过它们的最小平面就是直线 $L_{p,q}$. 又设 r 是另外一个点. 如果 r 在直线 $L_{p,q}$ 中, 则过 p, q, r 的最小平面还是 $L_{p,q}$; 否则, 就是整个空间 $\mathbb{A}$.

称仿射空间 $\mathbb{A}$ 中的点 $p_0, p_1, \cdots, p_k$ **仿射相关**如果它们落在某个 $k-1$ 维平面中, 否则称这些点**仿射无关或处于一般位置**. 定理 4.14 的证明显示 $p_0, p_1, \cdots, p_k$ 仿射无关当且仅当向量 $\overrightarrow{p_0p_1}, \cdots, \overrightarrow{p_0p_k}$ 线性无关. 从定义知仿射相关和仿射无关不依赖如何给点编号, 特别, 不依赖 p_0 的选取.

七　重心坐标　在仿射空间中, 一般而言, 点的线性组合没有含义, 只有 $\mathbb{A} = V$ 的情况可以例外. 不过, 点的某些线性组合却可以有明确的含义. 在第四部分关于坐标系的说明中, 已经看出坐标系 $\{o; e_1, \cdots, e_n\}$ 可以用仿射无关的点 $p_0, p_1, \cdots, p_n$ 代替. 任意点 $p \in \mathbb{A}$ 的坐标由表达式 $p = p_0 + \sum_{i=1}^{n} x_i(p_i - p_0)$ 确定. 该表达式形式上可以写成:

$$p = \left(1 - \sum_{i=1}^{n} x_i\right) p_0 + \sum_{i=1}^{n} x_i p_i.$$

定义 4.16　设 $p_0, p_1, \cdots, p_k$ 是仿射空间 $\mathbb{A}$ 中的点. 这些点的一个**重心组合**是指一个表达式 $\sum_{i=0}^{k} \alpha_i p_i$, 其中 $\alpha_0, \cdots, \alpha_k \in K$ 是纯量, 满足条件 $\sum_{i=0}^{k} \alpha_i = 1$. 任取 $\mathbb{A}$ 中的点 p, 令

$$\sum_{i=0}^{k} \alpha_i p_i = p + \sum_{i=0}^{k} \alpha_i (p_i - p),$$

从而, 这个重心组合是 $\mathbb{A}$ 中的点.

这个定义是合理的, 因为如下结论成立.

命题 4.17　如果 $\sum_{i=0}^{k} \alpha_i = 1$, 那么表达式

$$p + \sum_{i=0}^{k} \alpha_i (p_i - p)$$

与 p 的选取无关.

证明　的确, 如果 $q \in \mathbb{A}$ 是另一个点, 那么 $q = p + v$, $v \in V$. 于是

$$q + \sum_{i=0}^{k} \alpha_i (p_i - q) = p + v + \sum_{i=0}^{k} \alpha_i \overrightarrow{qp_i} = p + v + \sum_{i=0}^{k} \alpha_i (\overrightarrow{qp} + \overrightarrow{pp_i})$$

$$
\begin{aligned}
&= p + v + \sum_{i=0}^{k} \alpha_i (p_i - p - v) \\
&= p + v + \sum_{i=0}^{k} \alpha_i (p_i - p) - \left(\sum_{i=0}^{k} \alpha_i \right) v \\
&= p + \sum_{i=0}^{k} \alpha_i (p_i - p).
\end{aligned}
$$

□

点系 (system) $p_1, \cdots, p_k$ 的**质量中心**定义为

$$\mathrm{cent}(p_1, \cdots, p_k) = \frac{1}{k}(p_1 + \cdots + p_k).$$

$\mathrm{cent}(p, q) = \frac{1}{2}p + \frac{1}{2}q$ 可以认为是连接 p, q 的线段的**中点**. 又设 $\alpha_1, \cdots, \alpha_k$ 是纯量. 集合 $\{(p_1, \alpha_1), \cdots, (p_k, \alpha_k)\}$ 称为一个**质量点系**(或**加权点系**).

练习 4.18 设 $\{(p_1, \alpha_1), \cdots, (p_k, \alpha_k)\}$ 是一个质量点系, $\sum \alpha_i \neq 0$. 那么 $\mathbb{A}$ 中存在唯一的点 q 使得

$$\sum_{i=1}^{k} \alpha_i \overline{qp_i} = 0.$$

而且, 对 $\mathbb{A}$ 中任意的点 o, 有

$$\left(\sum_{i=1}^{k} \alpha_i \right) \overline{oq} = \sum_{i=1}^{k} \alpha_i \overline{op_i}.$$

点 q 称为这个质量点系的**重心**.

定义 4.19 称仿射空间 $\mathbb{A}$ 中的点组 $p_0, p_1, \cdots, p_n$ 是 $\mathbb{A}$ 的一个**重心坐标系**如果空间中任意的点 p 都能以唯一的方式表达成这些点的重心组合:

$$p = \sum_{i=0}^{n} x_i p_i, \quad x_i \in K, \quad \sum_{i=0}^{n} x_i = 1.$$

重心组合的系数 $x_0, \cdots, x_n$ 称为点 p 的**重心坐标**.

把点 p 的表达式写成 $p = p_0 + \sum_{i=1}^{n} x_i (p_i - p_0)$, 可以看出重心组合的唯一性等价于向量组 $p_1 - p_0, \cdots, p_n - p_0$ 是 V 的基, 即点组 $p_0, p_1, \cdots, p_n$ 是仿射无关的, 或说处于一般位置. 点 $p_0 + v$ 的坐标 $x_1, \cdots, x_n$ 和它的重心坐标 $x_0 = 1 - \sum_{i=1}^{n} x_i, x_1, \cdots, x_n$ 显然是互相确定的. 重心组合和重心坐标系对讨论仿射空间和仿射映射都是方便的. 例如, 直线 $L_{q,r}$ 就可以写成如下的形式:

$$L_{q,r} = \{\lambda q + (1-\lambda) r \mid \lambda \in K\}. \tag{4.1.7}$$

更一般地, 点组 $p_1, \cdots, p_k$ 的仿射包络是

$$A(p_1, \cdots, p_k) = \left\{ \sum_{i=1}^{k} \alpha_i p_i \,\middle|\, \alpha_i \in K,\ \sum_{i=1}^{k} \alpha_i = 1 \right\}. \tag{4.1.8}$$

定理 4.20 (1) 设 $f : \mathbb{A} \to \mathbb{A}'$ 是仿射映射, $p_0, \cdots, p_k$ 是 $\mathbb{A}$ 中的点; $x_i \in K$, $\sum\limits_{i=0}^{k} x_i = 1$. 那么

$$f\left(\sum_{i=0}^{k} x_i p_i\right) = \sum_{i=0}^{k} x_i f(p_i).$$

(2) 如果 $p_0, \cdots, p_n$ 是 $\mathbb{A}$ 的重心坐标系, 那么对任意的点 $q_0, q_1, \cdots, q_n \in \mathbb{A}'$, 存在唯一的仿射映射 f 使得

$$f(p_i) = q_i, \quad i = 0, 1, \cdots, n.$$

证明 (1) 对任意的 $p \in \mathbb{A}$, 按定义和式 (4.1.2), 有

$$\begin{aligned}
f\left(\sum_{i=0}^{k} x_i p_i\right) &= f\left(p + \sum_{i=0}^{k} x_i (p_i - p)\right) \\
&= f(p) + Df\left(\sum_{i=0}^{k} x_i (p_i - p)\right) = f(p) + \sum_{i=0}^{k} x_i Df(p_i - p) \\
&= f(p) + \sum_{i=0}^{k} x_i (f(p_i) - f(p)) = \sum_{i=0}^{k} x_i f(p_i).
\end{aligned}$$

(2) 由于 $\mathbb{A}$ 中每个点都以唯一的方式表成 $p_0, \cdots, p_n$ 的重心组合, 所以, 公式

$$f\left(\sum_{i=0}^{n} x_i p_i\right) = \sum_{i=0}^{n} x_i q_i$$

定义了一个映射 $f : \mathbb{A} \to \mathbb{A}'$. 要证明 f 是仿射映射需要找到它的微分. 因为 $p_1 - p_0, \cdots, p_n - p_0$ 是 V 的基, 所以存在唯一的线性映射 $\varphi : V \to V'$ 使得

$$\varphi(p_i - p_0) = q_i - q_0, \quad i = 1, \cdots, n.$$

我们证明它就是 f 的微分.

对 $\mathbb{A}$ 中的点 $p = \sum\limits_{i=0}^{n} x_i p_i$ 和 V 中的向量 $v = \sum\limits_{i=1}^{n} y_i (p_i - p_0)$, 有

$$\begin{aligned}p+v&=\sum_{i=0}^{n}x_ip_i+\sum_{i=1}^{n}y_i(p_i-p_0)\\&=p_0+\sum_{i=1}^{n}x_i(p_i-p_0)+\sum_{i=1}^{n}y_i(p_i-p_0)\\&=p_0+\sum_{i=1}^{n}(x_i+y_i)(p_i-p_0)\\&=(x_0-y_1-\cdots-y_n)p_0+\sum_{i=1}^{n}(x_i+y_i)p_i.\end{aligned}$$

于是

$$\begin{aligned}f(p)+\varphi(v)&=\sum_{i=0}^{n}x_iq_i+\sum_{i=1}^{n}y_i(q_i-q_0)\\&=(x_0-y_1-\cdots-y_n)q_0+\sum_{i=1}^{n}(x_i+y_i)q_i\\&=f(p+v).\end{aligned}$$

所以 f 是仿射映射, φ 就是 f 的微分. 显然 f 满足要求. 根据 (1), 它是唯一的. □

八　仿射线性函数　仿射映射的一个特殊情形是仿射线性函数. 设 $\mathbb{A}$ 是域 K 上的仿射空间, 其关联的向量空间是 V. 称映射 $f:\mathbb{A}\to K$ 是**仿射线性函数**如果

$$f(p+v)=f(p)+Df\cdot v,\quad \forall p\in\mathbb{A},\ v\in V,$$

其中 $Df\in V^*$ 是 V 上的线性函数. 按前面的说法, 可称它是 f 的线性部分或微分. 仿射空间上的常值函数是线性部分为零的仿射线性函数.

向量空间的坐标函数是线性函数. 对仿射空间, 有类似的结论.

命题 4.21　重心坐标函数是仿射线性函数.

证明　设 $p_0,\cdots,p_n$ 是仿射空间 $\mathbb{A}$ 的重心坐标系. 要证明函数

$$f_k:\mathbb{A}\to K,\quad \sum_{i=0}^{n}x_ip_i\to x_k$$

是仿射线性函数. 在定理 4.20(2) 中取 $\mathbb{A}'=K$, $q_k=1$, $q_i=0$ 如果 $i\neq k$, 即得函数 f_k. 所以重心坐标是仿射线性函数. □

坐标系对讨论仿射线性函数是有用的. 选取一个坐标系 $\{o;\ e_1,\cdots,e_n\}$, 点 $p\in\mathbb{A}$ 的坐标记为 $x_1,\cdots,x_n$. 那么仿射线性函数 f 的值就是

$$f(p)=f(o+\overline{op})=\sum_{i=1}^{n}\alpha_ix_i+\alpha_0,\tag{4.1.9}$$

其中 $\alpha_0=f(o)$, $\alpha_i=Df\cdot e_i$, $\overline{op}=x_1e_1+\cdots+x_ne_n$.

反过来, 如果函数 $f:\mathbb{A}\to K$ 的值由公式 (4.1.9) 给出, 那么, 对任意的向量 $v=y_1e_1+\cdots+y_ne_n$, 有

$$\begin{aligned}f(p+v)&=\sum_{i=1}^{n}\alpha_i(x_i+y_i)+\alpha_0\\&=\left(\sum_i\alpha_ix_i+\alpha_0\right)+\sum_i\alpha_iy_i=f(p)+Df\cdot v.\end{aligned}$$

也就是说, f 是仿射线性函数.

如同例 1.5, 对仿射线性函数 $f,g:\mathbb{A}\to K$ 和纯量 λ, μ, 可以定义 $(\lambda f+\mu g)(p)=\lambda f(p)+\mu g(p)$. 对任意的 $v\in V$, 有

$$\begin{aligned}(\lambda f+\mu g)(p+v)&=\lambda f(p+v)+\mu g(p+v)\\&=\lambda\,[\,f(p)+Df\cdot v]+\mu\,[g(p)+Dg\cdot v]\\&=(\lambda f+\mu g)(p)+(\lambda Df+\mu Dg)\cdot v.\end{aligned}$$

由此可见, $\lambda f+\mu g$ 是仿射线性函数, 其微分是 $\lambda Df+\mu Dg$. 从而, $\mathbb{A}$ 上的仿射线性函数全体 $\mathcal{S}$ 是向量空间. 现在, 微分 D 可以解释为线性映射 $\mathcal{S}\to V^*$, $f\to Df$. 核 $\ker D$ 就是 $\mathcal{S}$ 中所有的常值函数形成的直线 $\mathcal{S}^0$.

九 线性方程组 向量空间的线性函数和齐次线性方程组密切相关. 公式 (4.1.9) 提示仿射线性函数应该与一般的线性方程组密切相关. 的确如此.

线性方程组

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=b_1,\\&\cdots\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n&=b_m\end{aligned}\tag{4.1.10}$$

可以写成

$$f_1(p)=0,\quad\cdots,\quad f_m(p)=0,\tag{4.1.11}$$

其中 f_i 是仿射线性函数

$$f_i(p)=\sum_{j=1}^{n}a_{ij}x_j-b_i.$$

假设方程组 (4.1.10) 是相容的, ξ_1, $\cdots$, ξ_n 是它的一个解. 取定 $\mathbb{A}$ 的坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$. 那么这个解是某个点 p_0 的坐标 (因此 $f(o)=-b_i$). 也就是说, 点 p_0 是方程组 (4.1.11) 的解. 为方便起见, 也说 p_0 是方程组 (4.1.10) 的解. 方程组 (4.1.10) 的任意两个解的差是其相伴的齐次线性方程组的解, 而且, 相伴的齐次线性方程组的解与方程组 (4.1.10) 的解的和还是 (4.1.10) 的解. 所以方程组 (4.1.10)

或 (4.1.11) 的其他解具有形式 $p = p_0 + x$, 其中 $x \in V$ 且满足线性方程组

$$Df_1 \cdot x = 0, \quad \cdots, \quad Df_m \cdot x = 0. \tag{4.1.12}$$

方程组 (4.1.12) 中的 $Df_i \in V^*$ 是 f_i 的微分, 对 $x = x_1e_1 + \cdots + x_ne_n \in V$, 有 $Df_i \cdot x = \sum_j a_{ij}x_j$. 根据定理 1.62(1), 方程组 (4.1.12) 的解集 U 是 V 的子空间, 维数等于 $n-r$, 其中 r 是 (V^* 中的) 向量组 $Df_1, \cdots, Df_m$ 的秩. 于是, 方程组 (4.1.10) 或 (4.1.11) 的解集就是 $n-r$ 维的平面 $\Pi = p_0 + U$.

不出意料, 仿射空间 $\mathbb{A}$ 中的任意平面 $\Pi = p_0 + U$ 都是某个线性方程组的解. 事实上, 根据定理 1.62(2), U 是某个形如式 (4.1.12) 的方程组的解. 显然, p 属于 Π 当且仅当 $\overrightarrow{p_0p} \in U$(参见式 (4.1.5)). 如果 p 和 p_0 在选定的坐标系下的坐标分别是 $x_1, \cdots, x_n$ 和 $\alpha_1, \cdots, \alpha_n$, 那么 $\overrightarrow{p_0p} = \sum_j (x_j - \alpha_j)e_j$. 这样一来, 方程组 (4. 1.12) 就可以写成如下形式.

$$\sum_{j=1}^n a_{ij}(x_j - \alpha_j) = 0, \quad i = 1, \cdots, m.$$

命 $b_i = \sum_j \alpha_j$. 移项, 得

$$\sum_{j=1}^n a_{ij}x_j = b_j, \quad i = 1, \cdots, m,$$

这个方程组的系数矩阵 (和增广矩阵) 的秩等于 U 的余维数 $\dim V - \dim U$.

上面的讨论已经证明了下述结论.

定理 4.22 设 $\mathbb{A}$ 是 n 为仿射空间, 任意给定一个坐标系. 那么 $\mathbb{A}$ 中坐标满足一个系数矩阵秩为 r 的相容的线性方程组的点全体构成 $\mathbb{A}$ 的一个 $n-r$ 维的平面. $\mathbb{A}$ 中任意的平面都可以这样得到.

特别, 一个线性方程 $a_1x_1 + \cdots + a_nx_n = b$ 的解空间是超平面. 于是, 线性方程组的解空间就是若干个超平面的交集, 而且 $n-r$ 维的平面是 r 个超平面的交集. 对不相容的线性方程组, 其中的线性方程确定的超平面的交集是空集.

十 平面位置关系 在欧氏几何中, 点、线、面之间的关系是重要的研究内容. 在一般的仿射空间中, 由于没有度量, 所以没有垂直的概念, 但平行的概念是可以定义的.

定义 4.23 设 $\Pi = p + U$ 和 $\Pi' = p' + U'$ 是仿射空间 $\mathbb{A}$ 中的平面. 称这两个平面是**平行的**如果 $U \subseteq U'$ 或 $U' \subseteq U$.

在 $U = U'$ 的情形, 这个概念和以前的平行概念是一致的. 如果 $\Pi \subseteq \Pi'$ 或 $\Pi' \subseteq \Pi$, 那么, 平行性条件是满足的.

例 4.24 设 $f: V \to W$ 是线性映射, 那么, 所有的平面 $f^{-1}(w), w \in \operatorname{Im} f$, 是平行的, 因为它们的方向子空间都是 $\ker f$.

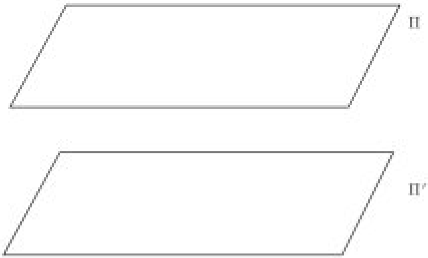


图 4.6　平行平面

命题 4.25　如果 $\Pi = p + U$ 和 $\Pi' = p' + U'$ 是平行的平面, 那么, 或者它们不相交, 或者它们之间有包含关系.

证明　如果它们相交, 取交集中的点 q, 则有

$$\Pi = q + U, \quad \Pi' = q + U'.$$

根据定义, 这两个平面有包含关系. □

平行关系的一个重要结论是平行公理.

定理 4.26　对一个仿射空间中任意的平面 Π 和任意的点 q, 存在唯一的平面 Π', 它过点 q, 维数等于 $\dim \Pi$, 且平行于 Π. 如果 $q \in \Pi$, 则 $\Pi = \Pi'$; 如果 $q \notin \Pi$, 则 Π 和 Π' 不相交.

证明　取定仿射空间的一个坐标系. 根据定理 4.22, 存在线性方程组

$$AX = B$$

使得 $p \in \Pi$ 当且仅当 p 的坐标是这个方程组的解. 设 q 的坐标列向量是 Y_0, $AY_0 = B'$, 那么 q 是方程组 $AX = B'$ 的解. 而且, 方程组 $AX = B'$ 的解空间 Π' 是一个平面. 由于 Π 和 Π' 的方向子空间都是方程组 $AX = 0$ 确定的子空间, 所以这两个平面平行且维数相等. 由于 $B = B'$ 当且仅当 $q \in \Pi$, 所以, 当 $q \in \Pi$ 时, 有 $\Pi = \Pi'$; 当 $q \notin \Pi$ 时, Π 和 Π' 不相交. □

在同一个坐标系下, 方程

$$a_1x_1 + \cdots + a_nx_n = b, \quad a_1'x_1 + \cdots + a_n'x_n = b',$$

给出的两个超平面是平行的等价于系数行向量成比例: $a_i' = \lambda a_i$, $i = 1, \cdots, n$. 两个平面重合则意味着常数项也是同样的比例: $b' = \lambda b$.

定义 4.27　一个仿射空间中的两个平面称为**偏斜的**(skew) 如果它们既不平行也不相交.

平面间的关系如平行、相交、偏斜等可以通过一些数值刻画. 设有两个平面 $\Pi' = p + U'$ 和 $\Pi'' = q + U''$, 维数分别是 k 和 l. 为方便计, 要求 $k \leqslant l$.

首先, 考虑这两个平面相交的情况. 设 o 是它们的一个公共点. 正如我们所知, 两个平面的并的仿射包络为

$$\Pi := A(\Pi', \Pi'') = o + W, \quad W = U' + U''.$$

根据推论 1.39, 此时有

$$m := \dim \Pi = \dim W = k + l - i, \quad i = \dim(U' \cap U'').$$

如果交集 $\Pi' \cap \Pi''$ 是空的, 那就需要考虑直线 $V_1 = \{\lambda \overline{pq} \mid \lambda \in K\}$ 和子空间

$$W^\circ = U' + U'' + V_1.$$

显然, 平面 $\Pi^\circ = p + W^\circ$ 包含 $\Pi' = p + U'$ 和 $\Pi'' = q + U'' = p + \overline{pq} + U'' \subset p + U'' + V_1$. 另一方面, 根据 (4.1.5), 包含 Π' 和 Π'' 的平面必然含有 $\overline{pq}$ 和直线 V_1, 从而包含 Π°. 这说明, 当 Π' 和 Π'' 的交集为空时, $\Pi^\circ = A(\Pi', \Pi'')$. 由于此时有 $\overline{pq} \notin U' + U''$ (见习题 1), 所以 Π' 和 Π'' 的并的仿射包络的维数是

$$m = \dim A(\Pi', \Pi'') = \dim(U' + U'') + \dim V_1 = k + l - i + 1.$$

四个整数形成的组

$$(i, k, l, m) = (\dim U' \cap U'', \dim U', \dim U'', \dim A(\Pi', \Pi'')), \quad 0 \leqslant i \leqslant k \leqslant l \leqslant m \leqslant n,$$

完全刻画了平面 Π' 和 Π'' 的相互位置:

(1) 平行, 如果 $i = k$;

(2) 相交, 如果 $m = k + l - i$;

(3) 不相交, 如果 $m = k + l - i + 1$;

(4) 偏斜, 如果 $i < k$ 且 $m = k + l - i + 1$;

(5) 交的维数是 i 如果 $m = k + l - i$.

例 4.28 两条直线的相互位置有四种情况: 重合、平行、相交但不重合、偏斜, 如图 4.7 所示.

(a) 重合 (b) 平行 (c) 相交但不重合 (d) 偏斜

图 4.7

习　题　4.1

1. 证明: 平面 $\Pi'=p+U'$ 和 $\Pi''=q+U''$ 的交集非空等价于 $\overline{pq}\in U'+U''$.

2. 设 $A(\Pi_1,\cdots,\Pi_m)$ 是 n 维仿射空间 $\mathbb{A}$ 中直线 $\Pi_1,\cdots,\Pi_m$ 的仿射包络. 如果 $A(\Pi_1,\cdots,\Pi_m)=\mathbb{A}$, 那么 m 的极小值是多少?

3. 设 $f:\mathbb{A}\to\mathbb{A}'$ 是仿射映射, Π 和 Π' 分别是 $\mathbb{A}$ 和 $\mathbb{A}'$ 中的平面. 证明:

(1) $f(\Pi)$ 是 $\mathbb{A}'$ 中的平面;

(2) 如果 $f^{-1}(\Pi')$ 非空, 那么 $f^{-1}(\Pi')$ 是 $\mathbb{A}$ 中的平面.

4. 设 $\mathbb{A}$ 和 $\mathbb{A}'$ 是仿射空间, 关联的向量空间分别是 V 和 V'. 证明: 对任意的线性映射 $\varphi:V\to V'$ 和任意的点 $p\in\mathbb{A}$, $p'\in\mathbb{A}'$, 存在唯一的仿射映射 $f:\mathbb{A}\to\mathbb{A}'$ 使得 $f(p)=p'$, $Df=\varphi$.

5. 设 $f:\mathbb{A}\to\mathbb{A}'$ 和 $g:\mathbb{A}'\to\mathbb{A}''$ 是仿射映射. 证明: $D(gf)=D(g)\circ D(f)$.

6. 设 $f:\mathbb{A}\to\mathbb{A}$ 是仿射映射, $o\in\mathbb{A}$. 证明: 存在唯一的向量 $u\in V$ 和保持 o 不动的仿射映射 $g:\mathbb{A}\to\mathbb{A}$ 使得

$$f=t_u\circ g.$$

7. 设 $p_0,\cdots,p_m$ 是仿射空间中的点. 证明:

(1) 这些点仿射无关当且仅当

$$p_i\notin A(p_0,\cdots,p_{i-1},p_{i+1},\cdots,p_m),\quad i=0,1,\cdots,m;$$

(2) 这些点仿射无关当且仅当 $p_i\notin A(p_0,\cdots,p_{i-1})$, $i=0,1,\cdots,m$.

8. 证明: 点组 $p_0,\cdots,p_m$ 的有限多个重心组合的重心组合仍然是这些点的一个重心组合.

9. 利用重心组合证明: 三角形的三条中线交于一点.

10. 如果仿射空间中有两个二维平面是偏斜的, 这个仿射空间的最小维数是多少?

11. 证明: 平行四边形的对角线交于它们的中点.

12. 设 $p_0,\cdots,p_n$ 是仿射空间 $\mathbb{A}$ 的仿射坐标系. 证明: $\mathbb{A}$ 中任何一点 q 都是某个质量点系 $\{(p_0,\alpha_0),\cdots,(p_n,\alpha_n)\}$ 的重心.

4.2　欧几里得仿射空间

一　结合欧几里得向量空间的公理和仿射空间的公理, 就能得到更合适讨论现实世界的空间——欧几里得仿射空间, 其中的几何涵盖了全部的初等几何.

定义 4.29　称仿射空间 $\mathbb{E}$ 为**欧几里得仿射空间**如果其关联的空间 V 是欧几里得向量空间. 如果不产生歧义, 欧几里得仿射空间常简称为**欧几里得空间**.

欧几里得空间 $\mathbb{E}$ 中两点的距离定义为

$$\rho(p,q)=\|\overline{pq}\|=\sqrt{(\overline{pq}\,|\,\overline{pq})},\tag{4.2.13}$$

这里 $(\cdot|\cdot)$ 是 V 的内积. 由欧几里得向量空间中长度的性质 (见 3.1 节) 知, 距离函数 $\rho(\cdot,\cdot)$ 是 $\mathbb{E}$ 的一个度量, 即有如下性质:

(1) $\rho(p,q)\geqslant 0$;

(2) $\rho(p,q)=0 \Leftrightarrow p=q$;

(3) $\rho(p,q)=\rho(q,p)$;

(4) $\rho(p,r)\leqslant \rho(p,q)+\rho(q,r)$ (三角不等式).

定义 4.30 欧几里得空间中向量 $\overline{pq}$ 和 $\overline{rs}$ 之间的夹角 φ 也称为直线 $L_{p,q}$ 和 $L_{r,s}$ 的夹角, 它由下式确定

$$\cos\varphi=\frac{(\overline{pq}\,|\,\overline{rs})}{\|\overline{pq}\|\cdot\|\overline{rs}\|}.$$

定义 4.31 欧几里得空间 $(\mathbb{E},V)$ 中的坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 称为**直角坐标系**, 如果 $e_1,\ \cdots,\ e_n$ 是欧几里得向量空间 V 的标准正交基: $(e_i|e_j)=\delta_{ij},\ i,j=1,2,\cdots,n$.

在直角坐标系下, 两个点的距离可以通过坐标计算. 设 $\mathbb{E}$ 中的点 p 和 q 在一个直角坐标系下的坐标分别为 $x_1,\ \cdots,\ x_n$ 和 $y_1,\ \cdots,\ y_n$. 那么向量 $\overline{pq}$ 的坐标就是 $y_1-x_1,\ \cdots,\ y_n-x_n$. 由定义等式 (4.2.13) 得

$$\rho(p,q)=\sqrt{(y_1-x_1)^2+\cdots+(y_n-x_n)^2} \tag{4.2.14}$$

称两个欧几里得空间 $\mathbb{E}$ 和 $\mathbb{E}'$ 是**同构的**如果存在仿射空间的同构映射 $f:\mathbb{E}\to\mathbb{E}'$, 它保持点之间的距离:

$$\rho(p,q)=\rho'(f(p),f(q)). \tag{4.2.15}$$

其中, ρ' 是 $\mathbb{E}'$ 上的距离函数. 借助直角坐标系可以建立同维数的欧几里得空间之间的同构.

定理 4.32 两个有限维的欧几里得空间 $\mathbb{E}$ 和 $\mathbb{E}'$ 是同构的当且仅当它们的维数相等.

证明 当它们同构时维数显然相等. 现假设它们的维数相等. 在 $\mathbb{E}$ 中取直角坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$, 在 $\mathbb{E}'$ 中取直角坐标系 $\{o';\ e_1',\ \cdots,\ e_n'\}$. 由定理 4.8 的证明知, 有仿射同构 $f:\mathbb{E}\to\mathbb{E}'$ 使得

$$f(o)=o',\quad Df(x_1e_1+\cdots+x_ne_n)=x_1e_1'+\cdots+x_ne_n'. \tag{4.2.16}$$

显然, 这个映射保持点的坐标, 即点 $p'=f(p)$ 在坐标系 $\{o';\ e_1',\ \cdots,\ e_n'\}$ 下的坐标就是点 p 在坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 的坐标. 于是, $\mathbb{E}$ 中点 p 和 q 之间的距离和 $\mathbb{E}'$ 中点 p' 和 q' 之间的距离都可以用相同的公式 (4.2.14) 计算. 这样, 映射 f 满足欧几里得空间同构的条件 (4.2.15), 所以它建立了欧几里得空间 $\mathbb{E}$ 和 $\mathbb{E}'$ 之间的同构. $\square$

我们引进一些在中学就知道的概念.

定义 4.33　实数域上的仿射空间中的集合

$$pq = \{p + \lambda\overline{pq} \mid 0 \leqslant \lambda \leqslant 1\} = \{(1-\lambda)p + \lambda q \mid 0 \leqslant \lambda \leqslant 1\}$$

称为连接点 p 和 q 的**线段**.

按定义, 有 $pq = qp$. 满足条件 $\overline{pr} = \overline{rq}$ 的点 $r \in pq$ 称为线段 pq 的**中点**. 易见

$$r = p + \frac{1}{2}\overline{pq} = \frac{1}{2}p + \frac{1}{2}q.$$

在欧几里得空间的情形, 线段 pq 的长度就是点 p 和 q 之间的距离,

$$|pq| := \|\overline{pq}\| = \rho(p,q).$$

二　点到平面的距离　设 $\Pi = o + U$ 和 $\Pi' = o' + U'$ 是欧几里得空间 $\mathbb{E}$ 的平面, $W = U \cap U'$. 又设 $U_1 = W^{\perp} \cap U$ 和 $U_1' = W^{\perp} \cap U'$ 分别是 W 在 U 中和在 U' 中的正交补.

定义 4.34　记号同上. 称欧几里得空间中两个平面 $\Pi = o+U$ 和 $\Pi' = o'+U'$**垂直**, 记做 $\Pi \perp \Pi'$, 如果 U_1 和 U_1' 均不是零子空间且正交.

设 p 是平面 Π 外的点, q 是 Π 中的点. 根据定义, 直线 $L_{p,q}$**垂直于**平面 Π 当且仅当对任意的 $r, s \in \Pi$ 都有 $(\overline{pq}|\overline{rs}) = 0$. 此时, 称值 $\rho(p,q)$ 为**点 p 到平面 Π 的距离**, 线段 pq 称为点 p 到平面 Π 的**垂 (直) 线**, 记作 $pq \perp \Pi$, q 称为垂线的**垂足**. 平面 Π 中的点到 Π 的距离定义为 0.

可以预料, 垂线的长度是点 p 到平面的最短距离, 即对任意不同于 q 的点 $r \in \Pi$, 有 $\rho(p,r) > \rho(p,q)$. 事实上, 如图 4.8 所示, $\overline{pr} = \overline{pq} + \overline{qr}$ 是两个正交的向量之和. 从而, 只要 $r \neq q$, 就有

$$\rho(p,r) = (\overline{pr} \mid \overline{pr}) = (\overline{pq} \mid \overline{pq}) + (\overline{qr} \mid \overline{qr}) = \rho(p,q)^2 + \rho(q,r)^2 > \rho(p,q)^2$$

(函数 ρ 的性质 (2)). 这说明垂线如果存在则是唯一的.

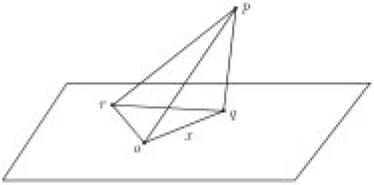


图 4.8

现在看一下过点 p 到平面 $\Pi = o + U$ 的垂线的存在性. 由于 $V = U \oplus U^{\perp}$, 所以 $\overline{po} = -x + y$, $x \in U$, $y \in U^{\perp}$. 命 $q = o + x \in \Pi$, 那么 $\overline{pq} = \overline{po} + \overline{oq} = y \in U^{\perp}$. 根据式 (4.1.5), 对任意的 $r, s \in \Pi$ 都有 $(\overline{pq} \,|\, \overline{rs}) = 0$. 即 pq 是垂线.

找出垂线 pq 的本质是找出垂足 q. 这等价于确定向量 $\overline{oq}$. 它由条件

$$(\overline{pq} \,|\, u) = (\overline{po} + \overline{oq} \,|\, u) = 0, \quad \forall u \in U \tag{4.2.17}$$

唯一确定. 由于点 o, p 是已知的, 所以向量 $\overline{po} = -\overline{op}$ 是已知的. 通过一个坐标系, 条件 (4.2.17) 就转化成向量 $\overline{oq}$ 的坐标的一个线性方程组, 从而找垂足 q 就是解线性方程组.

在 $\mathbb{E}$ 中选取直角坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 使得 $e_1,\ \cdots,\ e_m$ 构成 U 的基. 命

$$\overline{oq} = x = x_1 e_1 + \cdots + x_m e_m, \quad \overline{po} = -\overline{op} = -v,$$

那么条件 (4.2.17) 就是

$$(x - v \,|\, e_i) = 0, \quad i = 1,\ \cdots,\ m. \tag{4.2.18}$$

进而得到 $x_i = (v|e_i)$, $i = 1,\ \cdots,\ m$.

如果所选的坐标系不是直角坐标系, 那么, 方程组 (4.2.18) 的坐标形式就复杂一些:

$$(e_1|e_i)x_1 + (e_2|e_i)x_2 + \cdots + (e_m|e_i)x_m = (v|e_i), \quad i = 1,\ \cdots,\ m. \tag{4.2.19}$$

方程组 (4.2.18) 的解的存在性和唯一性不依赖基的选取. 已经知道 $\overline{oq}$ 是存在且唯一, 所以方程组 (4.2.19) 有唯一解. 这等价于方程组的系数矩阵的行列式不等于零. 这个行列式是向量 $e_1,\ \cdots,\ e_m$ 的格拉姆行列式

$$G(e_1, \cdots, e_m) = \begin{vmatrix} (e_1|e_1) & \cdots & (e_1|e_m) \\ \vdots & & \vdots \\ (e_m|e_1) & \cdots & (e_m|e_m) \end{vmatrix}. \tag{4.2.20}$$

它不等于零其实早就知道, 参见定理 3.84. 这里的推理过程和那里的没有差别.

刚才所得出关于垂线的结论可以总结成下述定理.

定理 4.35 给定欧几里得空间中的一个平面和平面外的一点, 那么存在过这一点到平面的垂线. 垂线的垂足在给定的平面上, 垂线的长度就是给定的点到给定的平面的最短距离. 垂足的坐标可以通过解线性方程组得到. 下面用符号表达这些内容.

设 p 是欧几里得空间 $\mathbb{E}$ 的平面 $\Pi = o + U$ 外的一点, 那么存在唯一的点 $q \in \Pi$ 使得 $(\overline{pq} \,|\, u) = 0$, $\forall u \in U$. 线段 pq 就是 p 到 Π 的垂线, 其长度 $|pq|$ 是 p 到 Π 的最

短距离. 点 o 可以在 Π 中任意选定. 这时, 存在向量 $x \in U$ 使得 $\overline{pq} = x - \overline{op}$. 任取 $\mathbb{E}$ 的坐标系 $\{o; e_1, \cdots, e_n\}$ 使得 $e_1, \cdots, e_m$ 构成 U 的基. 那么 $x = x_1e_1 + \cdots + x_me_m$ 在这个基下的坐标可以通过克拉默法则计算:

$$x_i = \frac{1}{G(e_1, \cdots, e_m)} \begin{array}{c} \begin{vmatrix} (e_1|e_1) & \cdots & (e_1|v) & \cdots & (e_1|e_m) \\ (e_2|e_1) & \cdots & (e_2|v) & \cdots & (e_2|e_m) \\ \vdots & & \vdots & & \vdots \\ (e_m|e_1) & \cdots & (e_m|v) & \cdots & (e_m|e_m) \end{vmatrix} \\ i \end{array}, \quad v = \overline{op}.$$

如果所取的坐标系是直角坐标系, 则有 $x_i = (e_i \,|\, \overline{op})$.

三　平面间的距离　有了垂线, 就可以讨论两个平面间的距离. 设 $\Pi = p + U$ 和 $\Pi' = p' + U'$ 是欧几里得空间 $\mathbb{E}$ 的两个平面. 由于 p' 可以用 Π' 中任意的点代替, 所以, 不妨要求 pp' 是点 p 到 Π' 的垂线, 即 $pp' \perp \Pi'$. 如果还有 $pp' \perp \Pi$, 则称 pp' 是 Π 和 Π' 间的**公垂线**.

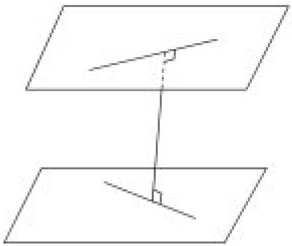


图 4.9

两个平面间的公垂线的长度称为这两个平面的**距离**, 原因如下.

引理 4.36　如果线段 pp' 是平面 Π 和 Π' 间的公垂线, 那么对任何点 $q \in \Pi$ 和 $q' \in \Pi'$ 都有

$$\rho(p, p') \leqslant \rho(q, q'). \tag{4.2.21}$$

证明　设 $q = p + u$, $q' = p' + u'$. 因为 $p' = p + \overline{pp'}$, 所以 $q' = p + \overline{pp'} + u'$. 从而有

$$\overline{qq'} = \overline{qp} + \overline{pq'} = \overline{pp'} + u' - u.$$

由条件 $(\overline{pp'} \,|\, u) = 0$ 和 $(\overline{pp'} \,|\, u') = 0$ 知 $(\overline{pp'} \,|\, u' - u) = 0$. 根据勾股定理, 得

$$\rho(q, q')^2 = \|\overline{qq'}\|^2 = \|u' - u\|^2 + \|\overline{pp'}\|^2 \geqslant \|\overline{pp'}\|^2 = \rho(p, p')^2.$$ □

引理 4.37 欧几里得空间中两个平面间的公垂线存在.

证明 设 $\Pi = o + U$ 和 $\Pi' = o' + U'$ 是欧几里得空间 $\mathbb{E}$ 中的两个平面. 根据定义, 需要找到 $p \in \Pi$ 和 $p' \in \Pi'$ 使得 $\overline{pp'} \in U^\perp$ 和 $\overline{pp'} \in U'^\perp$, 即

$$\overline{pp'} \in (U + U')^\perp.$$

由于 $V = (U+U')+(U+U')^\perp$, 所以 $\overline{oo'} = v+w$, 其中 $v \in U+U'$, $w \in (U+U')^\perp$. 可取 $u \in U$ 和 $u' \in U'$ 使得 $v = u + u'$. 命 $p = o + u$, $p' = o' - u'$, 那么

$$\overline{pp'} = \overline{po} + \overline{oo'} + \overline{o'p'} = -u + u + u' + w - u' = w \in (U + U')^\perp. \qquad \square$$

有这两个引理, 下面的结论几乎就是显然的了.

定理 4.38 给定欧几里得空间中两个平面, 可在两个平面上各找一点, 使得连接这两点的线段垂直这两个平面. 该线段是所给平面的公垂线, 公垂线的长度是这两个平面间的最短距离. 两个平面间的公垂线唯一当且仅当两个平面的方向子空间线性无关.

证明 前面两个断言已经证明了. 设 pp' 和 qq' 是平面 Π 和 Π' 间的公垂线, 其中 $p, q \in \Pi$ 和 $p', q' \in \Pi'$. 那么 $q = p + u$, $q' = p' + u'$, 其中 u 落在 Π 的方向子空间 U 中, u' 落在 Π' 的方向子空间 U' 中. 从引理 4.36 的证明知必有 $u = u'$. 所以 Π 和 Π' 间的公垂线与 $U \cap U'$ 中的向量自然地一一对应. 公垂线是唯一的当且仅当 $U \cap U' = 0$, 即方向子空间线性无关. 当平面 Π' 是一个点时, $U' = 0$, 事情回到第二部分的情形. $\square$

四 格拉姆行列式和平行六面体的体积 在 3.7 节中讨论最小二乘法和前面讨论点到平面的距离时都出现了格拉姆行列式 $G(e_1, \cdots, e_m)$. 这个行列式有很好的几何意义: 它是某个平行六面体的体积的平方. 在揭示这一点之前先看它的一个代数性质.

命题 4.39 设 $e_1, \cdots, e_m$ 是欧几里得向量空间 V 中的向量, 那么

$$G(e_1, \cdots, e_m) \geqslant 0, \tag{4.2.22}$$

等式成立当且仅当 $e_1, \cdots, e_m$ 线性相关. 当 $m = 2$ 时, 这个结论就是柯西–施瓦茨不等式.

证明 已经知道这个行列式等于零当且仅当 $e_1, \cdots, e_m$ 线性相关 (定理 3.84). 当这些向量线性无关时, 它们的格拉姆行列式是空间 $U = \langle e_1, \cdots, e_m \rangle$ 上的正定二次型 $q(v) = (v|v)$ 的矩阵的主子式 Δ_m. 西尔维斯特准则 (定理 1.85) 说: 此时有 $G(e_1, \cdots, e_m) = \Delta_m > 0$. $\square$

设 $\mathbb{E}$ 是欧几里得空间, V 是其关联的向量空间, $e_1, \cdots, e_m$ ($m \leqslant \dim V$) 是 V 中的向量. 任取 $\mathbb{E}$ 的一个点 o, 命 $p_i = o + e_i$, $i = 1, \cdots, m$. 格拉姆行列式

$G(e_1, \cdots, e_m)$ 可以解释为以向量 $\overline{op_1}, \cdots, \overline{op_m}$ 为边的平行六面体

$$P = P(\overline{op_1}, \cdots, \overline{op_m}) = \{o + t_1\overline{op_1} + \cdots + t_m\overline{op_m} \,|\, 0 \leqslant t_1, \cdots, t_m \leqslant 1\}$$

的体积的平方 S_P^2. 下面说明这一点.

向量组 $e_1, \cdots, e_m$ 落在 V 的一个 m 维子空间 U 中. 取 U 的标准正交基 $v_1, \cdots, v_m$. 令

$$e_j = \sum_{k=1}^{m} a_{kj} v_k, \quad j = 1, \cdots, m.$$

由第一卷 4.1 节第一部分的讨论知平行六面体 $P = P(\overline{op_1}, \cdots, \overline{op_m})$ 的体积 S_P 就是行列式 $\det(a_{ij})$ 的绝对值:

$$S_P = |\det(a_{ij})|. \tag{4.2.23}$$

由于 $v_1, \cdots, v_m$ 是标准正交向量组, 所以

$$(e_i|e_j) = \sum_{k=1}^{m} a_{ki} a_{kj}.$$

于是

$$\begin{aligned} S_P^2 &= \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix}^2 = \begin{vmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{mm} \end{vmatrix} \cdot \begin{vmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{vmatrix} \\ &= \begin{vmatrix} (e_1|e_1) & \cdots & (e_1|e_m) \\ \vdots & & \vdots \\ (e_m|e_1) & \cdots & (e_m|e_m) \end{vmatrix} \\ &= G(e_1, \cdots, e_m). \end{aligned}$$

由此可见, 格拉姆行列式是某个平行六面体的体积的平方.

习　题　4.2

1. 证明下面关于平面对 Π, Π' 的两个性质是等价的:

(1) 一个平面上的任何直线与另一个平面上的任何直线垂直;

(2) 平面 Π, Π' 互相垂直, 且至多有一个交点.

2. 设 $\Pi \subset \Gamma$ 是欧几里得空间 $\mathbb{E}$ 的两个平面. 证明: 如果平面 $\Gamma' \subset \mathbb{E}$ 与 Γ 垂直, 且 $\Gamma' \cap \Gamma = \Pi$, 那么 $\dim \Gamma' \leqslant \dim \mathbb{E} - \dim \Gamma + \dim \Pi$. 存在唯一的平面, 它具有这个性质, 且维数等于 $\dim \mathbb{E} - \dim \Gamma + \dim \Pi$.

3. 求出点 $p=(2,1,-3,4)$ 到平面

$$\Pi:\ 2x-4y-8z+13w+19=0,\quad x+y-z+2w-1=0$$

的距离.

4. 求出平面

$$\Pi_1:\ x_1+x_2+x_4-2x_5=2,\quad x_2+x_3-x_4-x_5=3,\quad x_1-x_2+2x_3-x_5=3$$

和平面

$$\Pi_2:\ (1,-2,5,8,2)+\langle(0,1,2,1,2),\ (2,1,2,-1,1)\rangle$$

之间的距离.

5. 证明: 公式 (4.2.23) 中的平行六面体的体积 S_P 可利用公式

$$S_P=\|\overrightarrow{op_1}\|\cdot l_1\cdot l_2\cdots l_{m-1}$$

计算, 其中 l_k 是点 p_{k+1} 到仿射包络 $A(o,p_1,\cdots,p_k)$ 的垂直线的长度.

6. 欧几里得向量空间 $\mathbb{R}^3$ 看作仿射空间. 考虑以原点为中心的单位球面 S. 曲面上两个大圆分别在两个平面 Π 和 Π' 上, 这两个平面的夹角 (在 0 和 π 之间) 定义为这两个圆周的夹角. (两个平面的交是直线 D, 其正交补 $D^{\perp}$ 与 Π 和 Π' 的交都是直线, 这两条直线的夹角就是这两个平面的夹角. 两个大圆的夹角也可定义为在它们在交点处的切线的夹角.)

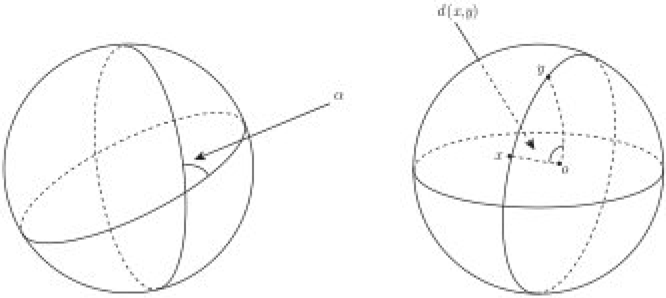


图 4.10

对 $x,y\in S$, 定义

$$d(x,y)=\arccos(x|y)$$

(它是向量 x 和 y 之间的夹角).

(1) 证明: 如果 x,y,z 是球面 S 上的三个点, α 是球面三角形 xyz 在顶点 x 的角 (即 x,y 所在的大圆与 x,z 所在的大圆的夹角), $a=d(y,z)$, $b=d(z,x)$, $c=d(x,y)$ 是三边的长度, 那么

$$\cos a=\cos b\cos c+\sin b\sin c\cos\alpha.$$

(2) 由此证明 d 是 S 上的长度函数.

(3) 证明: 如果 $y \neq -x$, 从 x 到 y 的折线中有一条最短的, 那就是含 x 和 y 的大圆上连接 x 和 y 的劣弧线.

7. 记号同上一题. 证明: 如果映射 $\varphi: S \to S$ 保持距离 d, 那么 φ 是欧几里得空间某个保距变换在 S 上的限制. 由此推出 (S, d) 的保距群和 $O(V)$ 同构.

4.3 群与几何

几何是最为古老的数学分支之一, 群在探索解一元高次方程的过程中于 19 世纪 30 年代诞生. 很快, 群论在几何的研究中就起突出的作用. 历史上, 阐述群与几何联系最有名文章应该是 F. 克莱因 1872 年发表的文章"Vergleichende Betrachtungen über neuere geometrische Forschungen"(A Comparative Review of Recent Researches in Geometry"(俗称埃尔朗根纲领). 这个纲领认为群论是组织几何知识最有益的方式: 不同的几何 (欧几里得几何、仿射几何、射影几何、非欧几何等) 其实研究的是图形在相应的空间变换群下不变的性质. 克莱因的这个观点在纯数学中影响非常广, 在物理中变换群也已是一个常用的工具.

一　仿射群　仿射几何研究图形在仿射自同构映射下不变的性质. 一个仿射空间的仿射自同构 (即仿射变换) 全体构成一个群. 仿射几何主要研究的就是图像在这个群下不变的性质. 先看一个简单的仿射群的例子.

例 4.40　实数集 $\mathbb{R}$ 是一维欧几里得向量空间. 根据例 4.2, 可以把它看作仿射直线. 实数域上的任何仿射直线与 $\mathbb{R}$ 都是同构的. 利用这个同构, 可以把一维的实仿射直线 $\mathbb{A}$ 与 $\mathbb{R}$ 等同. 从而, $\mathbb{A}$ 中的点 p 就用实数 x 表示. 仿射直线的几何, 需要通过仿射直线 $\mathbb{A}$ 的自同构研究. 定义映射 $\Phi_{\alpha,\beta}: \mathbb{A} \to \mathbb{A}$ 如下:

$$\Phi_{\alpha,\beta}: x \to \alpha x + \beta, \quad \alpha \in \mathbb{R}^*, \ \beta \in \mathbb{R}. \tag{4.3.24}$$

容易看出它是仿射自同构. 如果 $f: \mathbb{A} \to \mathbb{A}$ 是仿射自同构, 那么 $Df: \mathbb{R} \to \mathbb{R}$ 是线性同构, 所以有非零实数 α 使得 $Df(x) = \alpha x$. 于是 $f(x) = f(0 + x) = f(0) + Df(x) = \alpha x + f(0) = \Phi_{\alpha,\beta}(x)$, 其中 $\beta = f(0)$. 所以, $\mathbb{A}$ 的仿射自同构就是那些 $\Phi_{\alpha,\beta}$.

如果需要, 可以认为 x 就是点 p 在某个坐标系 $\{o; e\}$ 下的坐标, 即 $p = o + x$. 于是 $\Phi_{\alpha,\beta}(p) = \Phi_{\alpha,\beta}(o + x) = \Phi_{\alpha,\beta}(o) + D\Phi_{\alpha,\beta}(x) = (o + \beta) + \alpha x = o + (\alpha x + \beta)$. 实仿射直线 $\mathbb{A}$ 的所有仿射自同构组成的集合记作 A_1 或 $\mathrm{Aff}(\mathbb{A})$, 即

$$A_1 = \mathrm{Aff}(\mathbb{A}) = \{\Phi_{\alpha,\beta} \mid \alpha \in \mathbb{R}^*, \ \beta \in \mathbb{R}\}.$$

两个仿射变换 $\Phi_{\alpha,\beta}$, $\Phi_{\sigma,\tau} \in A_1$ 的合成仍是仿射变换:

$$\Phi_{\alpha,\beta} \cdot \Phi_{\sigma,\tau} = \Phi_{\alpha\sigma, \alpha\tau+\beta}. \tag{4.3.25}$$

恒等变换 $e = \Phi_{1,0} \in A_1$, 而 $\Phi_{\alpha,\beta}$ 的逆变换是 $\Phi_{\alpha^{-1},-\alpha^{-1}\beta}$. 所以集合 A_1 在映射的合成运算下成为群, 称为**1 阶仿射群**. 从等式 (4.3.25) 知 A_1 是非交换群, 而且映射

$$D : \Phi_{\alpha,\beta} \to \alpha$$

是从 A_1 到 $\mathbb{R}^*$ 的满同态, 其中 $\mathbb{R}^*$ 是非零实数全体形成的乘法群. 明显地, $\ker D = \{\Phi_{1,\beta} \mid \beta \in \mathbb{R}\}$ 是 $\mathbb{A}$ 的平移全体, 它同构于实数全体构成的加法群 $\mathbb{R}^+ = (\mathbb{R}; +)$. 于是, 我们有一个**群正合列**:

$$0 \to \mathbb{R}^+ \to A_1 \to \mathbb{R}^* \to 1.$$

(正合列的含义是其中任一个同态的像都等于下一个同态的核.)

现在转向一般情形. 设 $(\mathbb{A}, V)$ 是域 K 上的 n 维仿射空间, $f : \mathbb{A} \to \mathbb{A}$ 是仿射变换 (即仿射自同构). 按定义 4.4, 有

$$f(p+v) = f(p) + \mathcal{F}v,$$

其中 $\mathcal{F} = Df$ 是 f 的线性部分 (微分). 由于 f 是双射, 所以 $\mathcal{F}$ 是可逆的 (命题 4.7), 其逆算子 $\mathcal{F}^{-1}$ 是仿射变换 f^{-1} 的线性部分:

$$f^{-1}(p+v) = f^{-1}(p) + \mathcal{F}^{-1}v.$$

设 f 和 g 是 $\mathbb{A}$ 的两个仿射变换. 它们的合成

$$h = f \cdot g : p \to f(g(p))$$

仍是仿射变换, 而且, 其线性部分是 $\mathcal{F}\mathcal{G}$($\mathcal{F}$ 和 $\mathcal{G}$ 分别是 f 和 g 的线性部分). 实际上, 有

$$\begin{aligned}(f \cdot g)(p+v) &= f(g(p+v)) = f(g(p)) + \mathcal{G}v \\ &= f(g(p)) + \mathcal{F}(\mathcal{G}v) = (f \cdot g)(p) + \mathcal{F}\mathcal{G}v.\end{aligned} \tag{4.3.26}$$

显然, $\mathbb{A}$ 上的恒等变换 e 是仿射变换. 用 $\mathrm{Aff}\,\mathbb{A}$ 或 $A_n(K)$ 记 $\mathbb{A}$ 上的仿射变换全体. 由于映射的合成具有结合性, 所以前面的讨论表明 $A_n(K) = \mathrm{Aff}\,\mathbb{A}$ 在映射合成运算下成为群, 称为**n 阶仿射群**. 下面的定理显示这个群的结构和 1 阶仿射群的结构有类似之处.

定理 4.41 设 $(\mathbb{A}, V)$ 是域 K 上的 n 维仿射空间.

(1) 任取定点 $o \in \mathbb{A}$. 在 $A_n(K)$ 中保持点 o 不动的仿射变换全体记作 $A_n(K)_o$. 那么 $A_n(K)_o$ 是 $A_n(K)$ 的子群, 且与一般线性群 $GL(V) = GL_n(K)$ 同构;

(2) 空间 $\mathbb{A}$ 的平移全体 $T=\{t_v\,|\,v\in V\}$ 是 $A_n(K)$ 的正规子群, 即 T 是 $A_n(K)$ 的子群, 且对任意的 $f\in A_n(K)$ 和平移 t_v 有 $ft_vf^{-1}\in T$;

(3) 子群 T 是微分映射 $D: A_n(K)\to GL(V)$, $f\to Df$ 的核, 于是有正合列

$$e\to T\xrightarrow{\theta}A_n(K)\xrightarrow{D}GL(V)\to\bar{e},$$

其中 θ 是嵌入映射, 即 $\theta(t_v)=t_v$.

证明　(1) 如果 $f,g\in A_n(K)_o$, 那么

$$(f\cdot g)(o)=f(g(o))=f(o)=o,\quad f^{-1}(o)=o.$$

显然, $e(o)=o$. 所以 $A_n(K)_o$ 是 $A_n(K)$ 的子群.

对 $f\in A_n(K)_o$, 由于 $f(o+v)=o+Df\cdot v$, 所以 f 完全由其线性部分确定. 于是, 映射 $D_o: A_n(K)_o\to GL(V)$, $f\to Df$ 是单射. 容易验证 (参见定理 4.8 的证明), 对任意的线性同构 $\mathcal{F}: V\to V$, 映射 $o+v\to o+\mathcal{F}v$ 是 $\mathbb{A}$ 的仿射变换, 且保持 o 不动. 所以 D_o 也是满射, 给出了 $A_n(K)_o$ 到 $GL(V)$ 的群同构.

(2) 在 4.1 节第二部分就知道 $\mathbb{A}$ 的平移全体 T 是群, 与向量空间 V 的加法群同构. 设 t_v 是平移, f 是仿射变换, 其线性部分记作 $\mathcal{F}$. 那么

$$\begin{aligned}(ft_vf^{-1})(p)&=(ft_v)(f^{-1}(p))=f(f^{-1}(p)+v)\\&=f(f^{-1}(p))+\mathcal{F}v=p+\mathcal{F}v=t_{\mathcal{F}v}(p).\end{aligned}$$

由于 p 是 $\mathbb{A}$ 中任意的点, 从而可以推出

$$ft_vf^{-1}=t_{\mathcal{F}v}.$$

这个等式表明 T 是 $A_n(K)$ 的正规子群.

(3) 公式 (4.3.26) 说明微分映射 $D: A_n(K)\to GL(V)$ 是群同态, 而且是满射, 因为 $A_n(K)_o$ 的像就是 $GL(V)$. 要说明 T 等于这个同态的核 $\ker D=\{f\in A_n(K)\,|\,Df=\mathcal{E}\}$, 其中 $\mathcal{E}$ 是 V 的恒等映射. 对 $\mathbb{A}$ 中任意两点 p,q, 由仿射空间的定义知存在唯一的向量 $v\in V$ 使得 $q=p+v$. 如果 $f\in\ker D$, 那么 $f(q)=f(p+v)=f(p)+v$. 从而, 向量 $u=\overline{qf(q)}=\overline{(p+v)f(p+v)}=\overline{(p+v)(f(p)+v)}=\overline{pf(p)}$ 对任何 p,q 都是一样的. 而且 $f(p+v)=f(p)+v=(p+u)+v=(p+v)+u$. 所以 $f=t_u$ 是向量 u 给出的平移. 这说明 $\ker D\subset T$. 显然, $\mathbb{A}$ 的平移在 $\ker D$ 中. 于是 $T=\ker D$. □

从定理 4.41 可以推出下面的结论.

定理 4.42　取定点 $o\in\mathbb{A}$. 那么任意的仿射自同构 $f:\mathbb{A}\to\mathbb{A}$ 能以唯一的方式分解成平移和保持 o 不动的仿射自同构的乘积: $f=t_vg$, 其中 $g\in A_n(K)_o$. 易见 $v=\overline{of(o)}$.

证明 设 $Df=\mathcal{F}\in GL(V)$. 根据定理 4.41(1) 的证明, 存在唯一的 $g\in A_n(K)_o$ 使得 $Dg=\mathcal{F}$. 于是 $D(f\cdot g^{-1})=\mathcal{E}$, 即 $f\cdot g^{-1}\in\ker D$. 根据定理 4.41(3), 存在唯一的 $v\in V$ 使得 $f\cdot g^{-1}=t_v$. 于是 $f=t_vg$. 因为 $f(o)=t_v(g(o))=t_v(o)=o+v$, 所以 $v=\overline{of(o)}$. □

由这个定理可知, 对任意两点 $o,o'\in\mathbb{A}$ 和线性同构 $\mathcal{F}\in GL(V)$, 存在唯一的仿射自同构 $f:\mathbb{A}\to\mathbb{A}$ 使得 $f(o)=o'$, $Df=\mathcal{F}$. 实际上 $f=t_vg$, 其中 $v=\overline{oo'}$, $g\in A_n(K)_o$ 以 $\mathcal{F}$ 为其微分.

对 $f\in A_n(K)$ 和 $o,o'\in\mathbb{A}$, 定理 4.42 给出两个分解 $f=t_vg=t_{v'}g'$. 很容易看出向量 v,v' 的联系:

$$v'=\overline{o'f(o')}=\overline{o'o}+\overline{of(o)}+\overline{f(o)f(o')}=v-\overline{oo'}+Df(\overline{oo'}).$$

定理 4.42 表明, 取定点 $o\in\mathbb{A}$, 有自然的一一对应

$$\varphi:A_n(K)\to V\times GL(V),\quad f\to(\overline{of(o)},Df).$$

于是, 可以让 $V\times GL(V)$ 中的元素 $(v,\mathcal{F})$ 作用在 $\mathbb{A}$ 上:

$$(v,\mathcal{F})(o+x)=o+v+\mathcal{F}x.$$

这个作用的复合规则是

$$(v,\mathcal{F})(u,\mathcal{G})=(v+\mathcal{F}u,\mathcal{F}\mathcal{G}).$$

它和 $A_n(K)$ 中元素的乘积的作用规则是一致的. 实际上, 如果 $\varphi(f)=(v,\mathcal{F})$, $\varphi(g)=(u,\mathcal{G})$, 那么

$$\begin{aligned}(f\cdot g)(o+x)&=f(o+u+\mathcal{G}x)\\&=o+v+\mathcal{F}(u+\mathcal{G}x)=o+(v+\mathcal{F}u)+\mathcal{F}\mathcal{G}x.\end{aligned}$$

很多时候, 仿射映射的计算要借助坐标. 设 $\{o;\ e_1,\ \cdots,\ e_n\}$ 是仿射空间 $\mathbb{A}$ 的坐标系. 现在看一下 $f(p)$ 的坐标与 p 的坐标的联系. 按定义, 点 $p\in\mathbb{A}$ 的坐标就是向量 $\overline{op}=x=\sum\limits_i x_ie_i$ 的坐标 $x_1,\ \cdots,\ x_n$. 用 $\mathcal{F}$ 记 f 的线性部分, 则有

$$f(p)=f(o+x)=f(o)+\mathcal{F}x=o+\overline{of(o)}+\mathcal{F}x.$$

用 $y_1,\ \cdots,\ y_n$ 记点 $f(p)$ 的坐标, 并设 $\overline{of(o)}=\sum\limits_i b_ie_i$, 且 $F=(f_{ij})$ 是算子 $\mathcal{F}$ 在基 (e_i) 下的矩阵. 那么向量 $\mathcal{F}x$ 的第 i 个坐标是

$$(\mathcal{F}x)_i=\sum_{j=1}^n f_{ij}x_j.$$

综上所述, 得

$$y_i = \sum_{j=1}^{n} f_{ij}x_j + b_i, \quad i = 1, \cdots, n. \tag{4.3.27}$$

简记为

$$Y = FX + B,$$

其中 Y, X, B 分别是向量 $f(p)$, p, $\overline{of(o)}$ 的坐标列向量. 可以把这些等式与 (4.1.3) 中的等式和 2.1 节中第四部分中 $f(v)$ 的坐标的表达式比较.

二　三个平面几何定理　现在稍微放松休息一会儿, 利用仿射映射, 主要是投影、伸缩变换、平移, 证明三个平面几何的定理.

设 $L = o + U$ 是平面 $(\mathbb{A}, V)$(二维仿射空间) 中的直线, U' 是 V 中的直线, 不等于 U. 那么 $V = U \oplus U'$. 定义映射 $\pi : \mathbb{A} \to \mathbb{A}, o + x + y \to o + x$, 其中 $x \in U, y \in U'$. 这是一个仿射映射 (参见习题 4.1 第 4 题), 称为**沿着方向 $o + U'$(或方向 U') 到 L 的投影**.

对仿射空间 $\mathbb{A}$ 中共线的四点 p, q, r, s, 如果 $r \neq s$, 则有纯量 λ 使得 $\overline{pq} = \lambda\overline{rs}$, 于是可以定义向量 $\overline{pq}$ 与向量 $\overline{rs}$ 的比值为 $\overline{pq}/\overline{rs} = \lambda$.

定理 4.43 (Thales)　设 L, L', L'' 是平面 $(\mathbb{A}, V)$ 中三条平行的直线, D_1 和 D_2 是平面中与 L 不平行的直线. 令 $p_i = D_i \cap L$, $p_i' = D_i \cap L'$, $p_i'' = D_i \cap L''$. 那么下面的等式成立:

$$\frac{\overline{p_1p_1''}}{\overline{p_1p_1'}} = \frac{\overline{p_2p_2''}}{\overline{p_2p_2'}}.$$

反之, 如果 $q \in D_1$ 满足下述等式

$$\frac{\overline{p_1q}}{\overline{p_1p_1'}} = \frac{\overline{p_2p_2''}}{\overline{p_2p_2'}},$$

则 q 落在 L'' 中 (且 $q = p_1''$).

证明　见图 4.11.

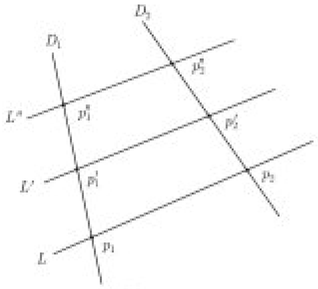


图 4.11

设 $\pi:\mathbb{A}\to\mathbb{A}$ 是沿着方向 L 到 D_2 的投影, $\mathcal{P}:V\to V$ 是其线性部分. 那么 π 把 p_1 映到 p_2, 等. 由于 $\mathcal{P}$ 是线性的, 所以 $\overline{p_1p_1''}=\lambda\overline{p_1p_1'}$ 意味着 $\mathcal{P}(\overline{p_1p_1''})=\lambda\mathcal{P}(\overline{p_1p_1'})$. 这就是说, $\overline{p_2p_2''}=\lambda\overline{p_2p_2'}$. 于是要证的等式成立.

反过来的结论从第一个结论推出: 因为有

$$\overline{p_1q}=\frac{\overline{p_2p_2''}}{\overline{p_2p_2'}}\overline{p_1p_1'}=\overline{p_1p_1''},$$

所以 $q=p_1''$. □

推论 4.44 设两条直线 D_1 和 D_2 交于点 p, L 和 L' 是两条平行的直线, 分别交 D_i 于 p_i 和 p_i', 点 p_i 和 p_i' 异于 p, 那么

$$\frac{\overline{pp_1}}{\overline{pp_1'}}=\frac{\overline{pp_2}}{\overline{pp_2'}}=\frac{\overline{p_1p_2}}{\overline{p_1'p_2'}}.$$

证明 见图 4.12.

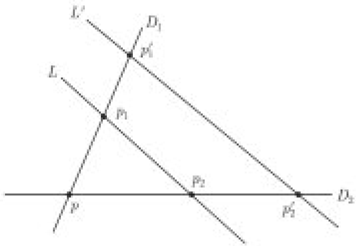


图 4.12

过点 p 作平行于 L 的直线. 运用 Thales 定理得到第一个等式, 而且以 p 为中心把 p_1 映到 p_1' 的伸缩变换把 p_2 映到 p_2', 于是第二个等式也成立. □

定理 4.45 (Pappus) 设 p,q,r 是直线 D 上的三点, p',q',r' 是另一条直线 D' 上的三点. 如果 $L_{p,q'}$ 和 $L_{q,p'}$ 平行, $L_{q,r'}$ 和 $L_{r,q'}$ 平行, 那么 $L_{p,r'}$ 和 $L_{r,p'}$ 平行.

证明 见图 4.13.

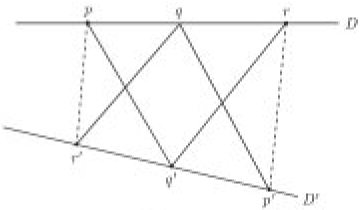


图 4.13

先考虑 D 和 D' 相交的情形. 设 o 是它们的交点. 设 φ 是以 o 为中心把 p 映到 q 的伸缩变换, ψ 是以 o 为中心把 q 映到 r 的伸缩变换. 运用 Thales 定理 (或由于伸缩变化的微分把向量映到共线而得向量) 知 φ 把 q' 映到 p', ψ 把 r' 映到 q'. 于是 $\psi\circ\varphi$ 把 p 映到 r, $\varphi\circ\psi$ 把 r' 映到 p'. 由于 φ 和 ψ 是同一中心的伸缩变换, 所以它们交换. 从而. 它们的合成就是以 o 为中心把 p 映到 r, r' 映到 p' 的伸缩变换. 由 Thales 定理 (的第二部分) 知 $L_{p,r'}$ 和 $L_{r,p'}$ 平行.

如果 D 和 D' 平行, 用平移代替伸缩变换即可证明结论. □

定理 4.46 (Desargues)　设 $\triangle pqr$ 和 $\triangle p'q'r'$ 是两个三角形, 没有公共顶点, 相应的边是平行的. 那么三直线 $L_{p,p'}$, $L_{q,q'}$, $L_{r,r'}$ 共点或平行.

证明　见图 4.14.

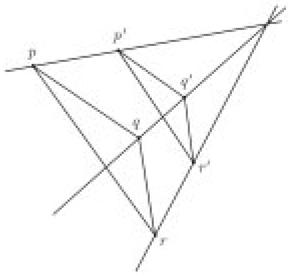


图 4.14

如果 $L_{p,p'}$ 和 $L_{q,q'}$ 交于点 o, 则以点 o 为中心把 p 映到 p' 的伸缩变换 φ 把 q 映到 q'(Thales 定理). 设 λ 是 φ 的伸缩率, $r''=\varphi(r)$, 则有 $\overline{or''}=\lambda\overline{or}$. 由于 $\overline{op'}=\lambda\overline{op}$, 所以 $L_{p',r''}$ 和 $L_{p,r}$ 平行. 于是 r'' 在过点 p' 的平行于 $L_{p,r}$ 的直线 $L_{p',r'}$ 上. 可是, r'' 也在过点 q' 的平行于 $L_{q,r}$ 的直线 $L_{q',r'}$ 上. 所以 $r''=r'$. 而且, o,r,r' 共线, 所以 $L_{r,r'}$ 过点 o.

如果 $L_{p,p'}$ 与 $L_{q,q'}$ 平行, 在上面的论证中用平移代替伸缩变换即可知道 $L_{r,r'}$ 与 $L_{p,p'}$ 平行. □

三　欧几里得空间的运动　设 $\mathbb{E}$ 是欧几里得空间, 其关联的欧几里得向量空间是 V, 度量记为 ρ. 无疑, 此时基域是 $K=\mathbb{R}$. 在欧几里得空间中度量是其最重要的特性.

定义 4.47　双射 $f:\mathbb{E}\to\mathbb{E}$ 称为 $\mathbb{E}$ 的**运动**(或**保距变换**、**保距映射**) 如果它保持距离, 即对任意 $p,q\in\mathbb{E}$, 有

$$\rho(f(p),f(q))=\rho(p,q). \tag{4.3.28}$$

下面的结论表明, 保距是比仿射更强的条件.

定理 4.48 映射 $f:\mathbb{E}\to\mathbb{E}$ 是运动, 当且仅当, f 是仿射变换且其线性部分是正交算子.

证明 断言的充分性部分几乎是显然的. 事实上, 以正交算子 $\mathcal{F}$ 为其线性部分的仿射变换 f 具有性质 (4.3.28):

$$\begin{aligned}\rho(f(p),f(q))&=\rho(f(p),f(p+v))=\|\overline{f(p)f(p+v)}\|\\&=\|\mathcal{F}v\|=\|v\|=\|\overline{p(p+v)}\|=\|\overline{pq}\|=\rho(p,q).\end{aligned}$$

特别, 要注意到平移是运动.

必要性的证明要费点周折, 分成四步.

1) 简单的验证即可看出, 任意两个运动的合成仍是运动. 设 f 是运动, 取定 $o\in\mathbb{E}$. 命 $v=\overline{f(o)o}$, 那么 $g=t_v\cdot f$ 仍是运动, 且保持点 o 不动:

$$g(o)=t_v(f(o))=f(o)+v=o.$$

于是问题归结为证明保持点 o 不变的运动 g 是仿射变换, 其微分是正交算子.

2) 定义映射 $\mathcal{G}:V\to V$, $x\to\mathcal{G}x=\overline{og(o+x)}$. 于是有

$$g(o+x)=o+\mathcal{G}x.\tag{4.3.29}$$

而且, 映射 $\mathcal{G}$ 有性质

$$\mathcal{G}0=0,\quad \|\mathcal{G}x-\mathcal{G}y\|=\|x-y\|.\tag{4.3.30}$$

实际上, $g(o)=o\Rightarrow\mathcal{G}0=0$. 对 $p=o+x$, $q=o+y$, 有

$$q=p+\overline{pq}=p+\overline{po}+\overline{oq}=p+y-x,$$

所以 $\overline{pq}=y-x$. 从而 $\rho(p,q)=\|y-x\|$. 另一方面, 由 (4.3.27) 得 $g(p)=o+\mathcal{G}x$, $g(q)=o+\mathcal{G}y$. 于是 $\rho(g(p),g(q))=\|\mathcal{G}y-\mathcal{G}x\|$. 由于 g 是运动, 故有 $\rho(g(p),g(q))=\rho(p,q)$. 我们说明了公式 (4.3.30) 成立.

设 $y=0$, 则得

$$\|\mathcal{G}x\|=\|x\|.\tag{4.3.31}$$

3) 映射 $\mathcal{G}$ 保持内积, 即

$$(\mathcal{G}x\,|\,\mathcal{G}y)=(x\,|\,y).\tag{4.3.32}$$

事实上, 根据式 (4.3.30), 有

$$\begin{aligned}\|x\|^2-2(x|y)+\|y\|^2&=(x-y\,|\,x-y)=\|x-y\|^2=\|\mathcal{G}x-\mathcal{G}y\|^2\\&=(\mathcal{G}x-\mathcal{G}y\,|\,\mathcal{G}x-\mathcal{G}y)=\|\mathcal{G}x\|^2-2(\mathcal{G}x|\mathcal{G}y)+\|\mathcal{G}y\|^2.\end{aligned}$$

由此和公式 (4.3.31) 即知式 (4.3.32) 成立.

4) 映射 $\mathcal{G}$ 是线性的. 首先它保持加法. 设 $z = x + y$. 则有 $\|z - x - y\|^2 = 0$. 展开, 得

$$\|z\|^2 + \|x\|^2 + \|y\|^2 - 2(z|x) - 2(z|y) + 2(x|y) = 0.$$

由这个等式和公式 (4.3.31) 及 (4.3.32) 知

$$\|\mathcal{G}z\|^2 + \|\mathcal{G}x\|^2 + \|\mathcal{G}y\|^2 - 2(\mathcal{G}z|\mathcal{G}x) - 2(\mathcal{G}z|\mathcal{G}y) + 2(\mathcal{G}x|\mathcal{G}y) = 0.$$

它等价于 $\|\mathcal{G}z - \mathcal{G}x - \mathcal{G}y\| = 0$. 也就是 $\mathcal{G}z = \mathcal{G}x + \mathcal{G}y$.

要看出 $\mathcal{G}$ 保持纯量乘, 还是利用公式 (4.3.29) 及 (4.3.30), 得

$$\begin{aligned}\|\mathcal{G}(\lambda x) - \lambda\mathcal{G}x\|^2 &= \|\mathcal{G}(\lambda x)\|^2 - 2(\mathcal{G}(\lambda x)|\lambda\mathcal{G}x) + \|\lambda\mathcal{G}x\|^2 \\ &= \|\lambda x\|^2 - 2\lambda(\lambda x|x) + |\lambda|^2\|x\|^2 = 0.\end{aligned}$$

所以 $\mathcal{G}(\lambda x) = \lambda x$.

于是, $\mathcal{G}$ 是正交算子. 由定理 4.8 的证明知 g 是以 $\mathcal{G}$ 为线性部分的仿射同构. 从而, $f = t_v \cdot g$ 是以 $\mathcal{G}$ 为线性部分的仿射变换. 定理证完. □

根据定理 4.42, 取定原点 $o \in \mathbb{E}$ 后, 每个运动可以写成平移与保持原点不动的运动的乘积. 这个分解中的平移依赖于原点的选取. 下面的定理给出运动的一个典范展示.

定理 4.49 对任何运动 f, 存在唯一确定的平面 $\Pi = p_0 + U$ 使得

(1) $f(\Pi) = \Pi$ 且 $f|_\Pi$ 是平移 (可能是平凡的, 即等于恒等变换, 或 Π 是一个点);

(2) Df 不固定任何 $U^\perp$ 中的非零元.

证明 如果平面 Π 存在, 它的方向子空间就是算子 $\mathcal{F} = Df$ 的以 1 为特征值的特征子空间 (也称为 $\mathcal{F}$ 不变的向量形成的子空间). 记这个子空间为 U. 取一个原点 o, 并把 f 表示成 $f = t_v \cdot g$, $g(o) = o$. 于是

$$f(o + x) = o + v + \mathcal{F}x.$$

令 $v = u + w$, $u \in U$, $w \in U^\perp$. 由于 $\mathcal{F} - \mathcal{E}$ 在 $U^\perp$ 上的限制是非退化的, 所以存在唯一的向量 $x_0 \in U^\perp$ 使得

$$(\mathcal{F} - \mathcal{E})x_0 = -w.$$

命 $p_0 = o + x_0$. 那么

$$f(p_0) = o + v + \mathcal{F}x_0 = o + u + w + x_0 - w = p_0 + u.$$

于是, 对任意的 $p = p_0 + x \in \Pi = p_0 + U$, 有 $f(p) = f(p_0) + \mathcal{F}x = p_0 + u + x \in \Pi$, 即 $f(\Pi) = \Pi$ 且 $f|_\Pi = t_u|_\Pi$. 显然 Π 是唯一满足条件 (1) 和 (2) 的平面. □

定理 4.49 中的平面 Π 称为运动 f 的**轴**. 运动 f 由它的轴 $\Pi = p_0 + U$, 向量 $u \in U$ 和正交变换 $\mathcal{B} = \mathcal{D}f|_{U^\perp}$ 确定.

运动的合成仍是运动, 运动的逆映射也是运动, 恒等变换是运动, 所以欧几里得空间 $\mathbb{E}$ 的运动全体形成群, 称为空间 $\mathbb{E}$ 的**保距群**, 记作 $\mathrm{Iso}(\mathbb{E})$. 由于维数相同的欧几里得空间是同构的 (定理 4.32), 所以, 在同构的意义下, 每一个维数的空间只有一个保距群. 显然, $\mathrm{Iso}(\mathbb{E})$ 是 $\mathrm{Aff}(\mathbb{E})$ 的子群. 空间 $\mathbb{E}$ 的平移全体形成的群 T 是 $\mathrm{Iso}(\mathbb{E})$ 的子群. 保持一个固定点 $o \in \mathbb{E}$ 不变的运动全体形成 $\mathrm{Iso}(\mathbb{E})$ 的一个子群 $\mathrm{Iso}(\mathbb{E})_o$, 它与 V 上正交算子群体形成的群 $O(V)$ 同构. 定理 4.41 中的正合列限制在保距群上则是下面的正合列:

$$e \to T \xrightarrow{\theta} \mathrm{Iso}(\mathbb{E}) \xrightarrow{D} O(V) \to \varepsilon.$$

称保距变换 f 为**刚体运动**(rigid motion) 或**规矩运动**((proper motion) 如果其微分的行列式是 1. 空间 $\mathbb{E}$ 的所有刚体运动形成一个群, 记作 $\mathrm{Iso}_+(\mathbb{E})$. 微分的行列式为 -1 的运动称为**非规矩运动**(improper motion) 或**非刚体运动**. 非刚体运动的重要例子是反射.

例 4.50 设 $\Pi = p_0 + U$ 是欧几里得空间 $\mathbb{E}$ 的超平面, $e \in U^\perp$ 是单位向量. 于是每个点 $p \in \mathbb{E}$ 可以写成

$$p = q + \lambda e, \quad q \in \Pi.$$

通过 Π 的 (正交) 反射 $R_\Pi : \mathbb{E} \to \mathbb{E}$ 定义为

$$R_\Pi p = q - \lambda e.$$

见图 4.15.

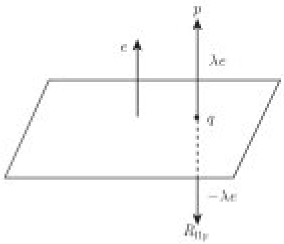


图 4.15

反射 R_H 的微分是空间 V 的通过 Π 的方向子空间 U 的 (正交) 反射. 如果 Π 和 Π' 是平行的超平面, 那么 $DR_\Pi = DR_{\Pi'}$, 从而

$$D(R_\Pi R_{\Pi'}) = DR_\Pi \cdot DR_{\Pi'} = \mathcal{E}.$$

于是, 在这种情况 $R_\Pi \cdot R_{\Pi'}$ 是平移, 平移的向量是 Π 和 Π' 间公垂线对应的向量的 2 倍.

如果超平面 Π 和 Π' 的交是 $n-2$ 维的平面 P, 那么 $R_\Pi R_{\Pi'}$ 是旋转, 旋转角是 Π 和 Π' 间夹角的 2 倍, 即映射保持 P 的每个点不动, 在每个正交于 P 的二维平面上是以夹角 2 倍为旋转角的旋转.

如果 $\{o; e_1, \cdots, e_n\}$ 是 $\mathbb{E}$ 的直角坐标系, 那么运动 $f \in \mathrm{Iso}(\mathbb{E})$ 的坐标形式是

$$Y = FX + B, \tag{4.3.33}$$

其中 $X = [x_1, \cdots, x_n]$, $Y = [y_1, \cdots, y_n]$, $B = [b_1, \cdots, b_n]$ 分别是点 p, $f(p)$, 向量 $\overline{of(o)}$ 的坐标列向量, F 是 Df 在标准正交基 $e_1, \cdots, e_n$ 下的矩阵, 它是正交矩阵.

四 低维欧几里得空间的运动 欧几里得空间的运动在几何学和物理学中经常遇到, 所以, 有必要对低维的情形给予较详细的讨论. 定理 4.49 对描述低维欧几里得空间的运动是方便的. 以下用 Π 记运动 f 的轴, $\mathcal{B}$ 记 $Df = \mathcal{F}$ 在 $U^\perp$ 上的限制, 其中 U 是 Π 的方向子空间.

一维情形 由于 $\mathbb{E}$ 是直线, 运动 f 的轴 Π 的维数有两个可能.

(1) $\dim \Pi = 1$. 这时 $\Pi = \mathbb{E}$, 从而 f 是平移.

(2) $\dim \Pi = 0$. 这时 Π 是点, $Df = -\mathcal{E}$, 所以 f 是通过点 Π 的反射.

用 f 的坐标形式讨论也是简单的, 可以和刚才的讨论做一比较. 根据 (4.3.33), 有

$$y = \epsilon x + a, \tag{4.3.34}$$

其中 $\epsilon = \pm 1$ (1 阶的正交矩阵), a 是某个常数. 如果 $\epsilon = 1$, 那么 f 就是向量 a 确定的平移. 如果 $\epsilon = -1$, (4.3.34) 的变形

$$y - a/2 = -(x - a/2)$$

有清晰的几何意义, 它指出 f 的轴是 $a/2$(更确切地说, 是以 $a/2$ 为坐标的点), 且 f 是过轴的反射: $f(a/2 + x) = a/2 - x$.

以 $o' = a/2$ 为新的原点, 那坐标变换就是 $x = x' + a/2$, $y = y' + a/2$. 新坐标系下 f 的坐标形式是 $y' = -x'$, 这是过新原点的反射.

二维情形 运动 f 的轴 Π 的维数有三个可能.

(1) $\dim \Pi = 2$. 这时 $\Pi = \mathbb{E}$, 从而 f 是平移.

(2) $\dim \Pi = 1$. 这时 Π 是直线, $\mathcal{B} = -\mathcal{E}$ ($\mathcal{E}$ 总是记恒等变换, 涉及的空间一般是自明的). 所以 f 是通过直线 Π 的反射, 或滑动反射 (glide reflection), 即过直线 Π 的反射和一个沿着 Π 的平移的合成.

(3) $\dim \Pi = 0$. 这时 Π 是点, f 是绕点 Π 的 (非平凡的, 即旋转角不等于 2π 的倍数的) 旋转.

同直线的情形一样, 用 f 的坐标形式讨论也是简单的. 选取直角坐标系 $\{o;\ e_1,\ e_2\}$ 使得 Df 具有典范形式 (定理 3.57). 在这个坐标系下, f 的坐标形式有如下三种可能:

$$\text{(i)}\ \begin{aligned} x' &= x+a, \\ y' &= y+b; \end{aligned} \qquad \text{(ii)}\ \begin{aligned} x' &= x+a, \\ y' &= -y+b; \end{aligned} \qquad \text{(iii)}\ \begin{aligned} x' &= x\cos\varphi - y\sin\varphi + a, \\ y' &= x\sin\varphi + y\cos\varphi + b. \end{aligned}$$

需要注意到 (iii) 包含 $Df = -\mathcal{E}$ 的情形, 此时 $\varphi = \pi$.

在情形 (i), f 是向量 $ae_1 + be_2$ 确定的平移, 其轴是二维的. 在情形 (ii), f 的轴是 $\Pi = o + \dfrac{b}{2}e_2 + \mathbb{R}e_1$. 以 $p_0 = o + \dfrac{b}{2}e_2$ 为新原点, 在坐标系 $\{p_0;\ e_1,\ e_2\}$ 下的坐标 ξ, η 与原来坐标的联系如下:

$$x = \xi \ \ (x' = \xi'), \quad y = \eta + b/2 \ \ (y' = \eta' + b/2).$$

在新坐标系下, f 的坐标形式是

$$\xi' = \xi + a, \quad \eta' = -\eta.$$

其几何含义是: 当 $a = 0$ 时, f 就是通过轴 Π 的反射; 当 $a \neq 0$, f 的作用等于先通过轴 Π 做反射, 然后再沿着向量 ae_1 做平移.

在情形 (iii), φ 不是 2π 的倍数, 否则就是情形 (i). 由于 1 不是 Df 的特征值, 所以 Df 不变的向量只有零向量. 根据定理 4.49, f 的轴是一个点, 这个点 p_0 是 f 唯一的不动点, 其坐标 x_0, y_0 由方程组

$$\begin{aligned} x_0\cos\varphi - y_0\sin\varphi + a &= x_0, \\ x_0\sin\varphi + y_0\cos\varphi + b &= y_0 \end{aligned}$$

确定. 以 $p_0 = o + x_0e_1 + y_0e_2$ 为新原点, 在坐标系 $\{p_0;\ e_1,\ e_2\}$ 下的坐标 ξ, η 与原来坐标的联系是:

$$x = \xi + x_0 \ \ (x' = \xi' + x_0), \quad y = \eta + y_0 \ \ (y' = \eta' + y_0).$$

在新坐标系下, f 的坐标形式是

$$\begin{aligned} \xi' &= \xi\cos\varphi - \eta\sin\varphi, \\ \eta' &= \xi\sin\varphi + \eta\cos\varphi. \end{aligned}$$

它就是绕新原点 p_0 的旋转, 旋转角度是 φ.

在情形 (i) 和 (iii), f 是刚体运动; 在情形 (ii), f 是非刚体运动. 特别, 欧几里得平面上的刚体运动只有平移和绕某个不动点的旋转.

三维情形　运动 f 的轴 Π 的维数有四个可能.

(1) $\dim\Pi=3$. 这时 $\Pi=\mathbb{E}$, 从而 f 是平移.

(2) $\dim\Pi=2$. 这时 Π 是平面, $\mathcal{B}=-\mathcal{E}$. 所以 f 是通过直线 Π 的反射, 或滑动反射 (glide reflection), 即过平面 Π 的反射和一个平移的合成, 这个平移由 Π 的方向子空间中的一个向量确定.

(3) $\dim\Pi=1$. 这时 Π 是直线, f 是绕直线 Π 的 (非平凡的) 旋转, 或是螺旋运动 (spiral motion), 即, 一个绕 Π 的旋转与一个平移的合成, 这个平移由 Π 的方向子空间中的一个向量确定.

(4) $\dim\Pi=0$. 这是 Π 是一个点, f 是镜像旋转 (mirror rotation), 即绕一条直线的 (非平凡的) 旋转与过一个平面的反射的合成, 这个平面和这条直线垂直, 且交于点 Π.

同直线、平面的情形一样, 可以用 f 的坐标形式讨论. 选取直角坐标系 $\{o;\ e_1,\ e_2,\ e_3\}$ 使得 Df 具有典范形式 (定理 3.59). 在这个坐标系下, f 的坐标形式有如下四种可能:

$$\begin{aligned}
&\text{(i)}\ x'=x+a, && \text{(iii)}\ x'=x\cos\varphi-y\sin\varphi+a,\\
&\quad\ \ y'=y+b, && \qquad\ \ y'=x\sin\varphi+y\cos\varphi+b,\\
&\quad\ \ z'=z+c; && \qquad\ \ z'=z+c;\\
&\text{(ii)}\ x'=x+a, && \text{(iv)}\ x'=x\cos\varphi-y\sin\varphi+a,\\
&\quad\ \ y'=y+b, && \qquad\ \ y'=x\sin\varphi+y\cos\varphi+b,\\
&\quad\ \ z'=-z+c; && \qquad\ \ z'=-z+c.
\end{aligned}$$

需要注意到取 $\varphi=\pi$, 则 (iv) 就是 $Df=-\mathcal{E}$ 的情形, (iii) 就是 $Df+\mathcal{E}$ 的核为二维的情形.

在情形 (i), f 是向量 $ae_1+be_2+ce_3$ 确定的平移, 其轴是三维的. 在情形 (ii), f 的轴是 $\Pi=o+\dfrac{c}{2}e_3+\mathbb{R}e_1+\mathbb{R}e_2$. 以 $p_0=o+\dfrac{c}{2}e_3$ 新原点, 在坐标系 $\{p_0;\ e_1,\ e_2,\ e_3\}$ 下, f 的坐标形式是

$$\xi'=\xi+a,\quad \eta'=\eta+b,\quad \mu'=-\mu.$$

其几何含义是: 当 $a=b=0$ 时, f 就是通过轴 Π 的反射; 当 $a\neq 0$ 或 $b\neq 0$ 时, f 的作用等于先通过轴 Π 做反射, 然后再沿着向量 ae_1+be_2 做平移.

在情形 (iii), 如同二维的情形 (iii), φ 不是 2π 的倍数, 在平面 $o+\mathbb{R}e_1+\mathbb{R}e_2$ 上 f 有唯一的不动点 $p_0=o+x_0e_1+y_0e_2$. 运动 f 的轴是 $\Pi=p_0+\mathbb{R}e_3$. 以 p_0 为新原点, 在坐标系 $\{p_0;\ e_1,\ e_2,\ e_3\}$ 下, f 的坐标形式是

$$\xi'=\xi\cos\varphi-\eta\sin\varphi,$$

$$\eta' = \xi \sin\varphi + \eta\cos\varphi,$$
$$\mu' = \mu + c.$$

当 $c=0$, 它就是绕轴 Π 的旋转, 旋转角度是 φ. 当 $c \neq 0$ 时, f 是绕轴 Π 旋转角度 φ, 然后再用向量 ce_3 做平移. 这种运动就是力学中的螺旋运动.

情形 (iv) 是情形 (ii) 和 (iii) 的合成, 这时 f 有唯一的不动点 $p_0 = o + x_0e_1 + y_0e_2 + \dfrac{c}{2}e_3$, 其中 x_0, y_0 与二维的情形中 (iii) 中的一样. 在坐标系 $\{p_0;\ e_1,\ e_2,\ e_3\}$ 下, f 的坐标形式是

$$\xi' = \xi\cos\varphi - \eta\sin\varphi,$$
$$\eta' = \xi \sin\varphi + \eta\cos\varphi,$$
$$\mu' = -\mu.$$

由此可见, f 是绕直线 $p_0 + \mathbb{R}e_3$ 旋转角度 φ, 然后再过平面 $p_0 + \mathbb{R}e_1 + \mathbb{R}e_2$ 做反射. 这个反射和旋转是交换的, 所以旋转后反射还是反射后旋转都是一样的.

在情形 (i) 和 (iii), f 是刚体运动; 在情形 (ii) 和 (iv), f 是非刚体运动. 特别, 三维欧几里得空间的刚体运动只有平移、旋转和螺旋运动三种. 于是, 三维欧几里得空间的刚体运动如果有不动点, 那一定是旋转, 旋转轴是由运动的不动点形成的直线 (欧拉定理, 1776 年). 而且, 任何物体的移动都可以通过先沿着一个方向平移, 后绕某条直线旋转来实现, 也可以先旋转, 再平移 (沙勒定理 (Chasles' theorem), 1830 年).

五　群对应的几何　按照克莱因在 "埃尔朗根纲领" 中的观点, 几何学研究在变换群下图形的不变的性质. 不同的变换群对应不同的几何. 换句话说, 给了集合 Γ, 和它的一些变换 (即双射 $\Gamma \to \Gamma$) 组成的群 G, 与 G 相应的几何就是研究 Γ 中的图形 (或说点组成的空间构形) 的在 G 的作用下不变的性质.

初等几何中全等这一概念是重要的. 仔细想一想就会发现, 中学几何中两个图形全等就是可以通过平移、旋转、反射这样的等距变换把一个图形变成另一个图形. 这个想法适用于其他群的几何. 集合 Γ 中的两个图形 Φ_1 和 Φ_2 称为 **G-全等的**或 **G-叠合的**(congruence) 如果存在 $g \in G$ 使得 $\Phi_2 = g(\Phi_1)$. 我们用记号 $\Phi_1 \overset{G}{\simeq} \Phi_2$ 表示图形 Φ_1 和 Φ_2 是 G-全等的. 在不会产生歧义的情况下 (即知道所涉及的群), G- 叠合就简单说成叠合. 由群的公理可以直接推出叠合关系是等价关系, 即有下列性质:

i) 自反性; 这是因为 $\Phi = e(\Phi) \Rightarrow \Phi \overset{G}{\simeq} \Phi$, 其中 e 是 G 的单位元.

ii) 对称性; 这是因为 $\Phi_2 = g(\Phi_1) \Leftrightarrow \Phi = g^{-1}(\Phi_2)$, 也就是 $\Phi_1 \overset{G}{\simeq} \Phi_2 \Leftrightarrow \Phi_2 \overset{G}{\simeq} \Phi_1$.

iii) 传递性. 这是因为 $\Phi_2 = g(\Phi_1)$, $\Phi_3 = h(\Phi_2) \Rightarrow \Phi_3 = (hg)(\Phi_1)$.

仿射群和保距群对应的几何分别是仿射几何和欧几里得几何. 仿射几何研究图

形在仿射变换下不变的性质. 仿射变换把平面映到维数相同的平面, 点组的重心组合映到点组的像的重心组合, 系数不变. 于是, 平面与重心组合是仿射几何的概念, 由它们衍生出来的概念如平行、平行六面体、区间、区间的中点、质量点系的重心等都是仿射几何的概念. 但是, 正方形、圆、角度等不是仿射几何的概念, 因为仿射变换可以把正方形变成长方形、圆变成椭圆、改变角的大小等. 下面将看到凸集、单形等也是仿射几何的概念. 类似地, 欧几里得几何研究图形在保距变换下不变的性质, 由于保距群是仿射群的子群, 所以仿射几何中的概念也是欧几里得几何的概念, 而且正方形、圆、椭圆、抛物线、角度等也是欧几里得几何的概念, 因为它们在保距变换下是不变的. 现在研究一下这两种几何最简单的一些性质. 对仿射几何, 空间是 $\Gamma = \mathbb{A}$, 群是 $\mathrm{Aff}(\mathbb{A})$; 对欧几里得几何, 空间是 $\mathbb{E}$, 群是 $\mathrm{Iso}(\mathbb{E})$.

定理 4.51　设 $(\mathbb{E}, V)$ 是欧几里得空间. 两个平面 $\Pi, \Pi' \subset \mathbb{E}$ 是叠合的 ($G = \mathrm{Iso}(\mathbb{E})$) 当且仅当它们的维数相等. 特别, 任意两点是叠合的. 对仿射空间 $(\mathbb{A}, V)$ 和群 $G = \mathrm{Aff}(\mathbb{A})$, 同样的结论成立.

证明　如果 $\mathbb{E}$ 中的平面 $\Pi = p + U$ 和 $\Pi' = p' + U'$ 是叠合的, 那么存在 $f \in G$ 使得 $\Pi' = f(\Pi)$. 于是 $Df(U) = U'$. 因为 Df 是非退化的, 所以 U 和 U' 的维数相等. 从而, 按定义 Π 和 Π' 的维数相等.

反过来, 假设 Π 和 Π' 的维数相同. 在它们的方向子空间 U 和 U' 中各选取标准正交基 $e_1, \cdots, e_m$ 和 $e'_1, \cdots, e'_m$, 然后分别扩充为空间 V 的标准正交基 (e_i) 和 (e'_i). 有唯一的正交算子 $\mathcal{F} : V \to V$ 使得 $\mathcal{F}e_i = e'_i$, $i = 1, \cdots, n = \dim V$. 进而, 满足 $f(p) = p'$ 且 $Df = \mathcal{F}$ 的运动就把 Π 变到 Π'.

对仿射空间, 论证是类似的, 而且, 不用为度量操心. □

定理 4.51 关于点的叠合性常用另一个术语表达: 空间 $\mathbb{E}$(相应地 $\mathbb{A}$) 中的点在群 G 的作用下是**可迁的**(transitive). 点的可迁性是与群 G 相应的几何的最重要的性质. 没有这个性质, 就无法比较不同位置的图形. 在仿射几何, 不仅点具有可迁性, 很多的点组也具有可迁性.

定理 4.52　设 $p_0, \cdots, p_m$ 和 $q_0, \cdots, q_m$ 是仿射空间 $\mathbb{A}$ 中任意两个仿射无关的点组, 那么它们是叠合的.

证明　把这两个点组分别扩充为 $\mathbb{A}$ 的重心坐标系 $p_0, \cdots, p_n$ 和 $q_0, \cdots, q_n$, 此处 $n = \dim \mathbb{A}$. 根据定理 4.20(2), 存在唯一的仿射映射 $f : \mathbb{A} \to \mathbb{A}$ 使得 $f(p_i) = q_i$, $i = 1, \cdots, n$; 还存在唯一的仿射映射 $g : \mathbb{A} \to \mathbb{A}$ 使得 $g(q_i) = p_i$, $i = 1, \cdots, n$. 利用定理 4.20(2) 知 $fg = gf = e$ 是 $\mathbb{A}$ 的恒等映射. 所以 f 和 g 都是仿射变换. □

明显地, 这个定理在欧几里得几何中不对, 甚至在 $m = 1$ 的情形, 因为点对 p, q 和点对 p', q' 是 $\mathrm{Iso}(\mathbb{E})$-叠合的意味着 $\rho(p, q) = \rho(p'q')$. 在仿射几何中, 没有距离的概念, 因为不同两点组成的点对可以通过仿射变换映到任意不同两点组成的点对. 不过, 仿射变换 $f : \mathbb{A} \to \mathbb{A}$ 保持点的共线性质, 而且从 (4.1.2) 知它有下述更强的

性质:

$$\overrightarrow{rs} = \lambda\overrightarrow{pq} \Rightarrow \overrightarrow{f(r)f(s)} = \lambda\overrightarrow{f(p)f(q)}, \quad \forall p,q,r,s \in \mathbb{A},\ \lambda \in K. \tag{4.3.35}$$

这个性质其实刻画了仿射变换.

定理 4.53 双射 $f:\mathbb{A}\to\mathbb{A}$ 是仿射变换当且仅当它满足性质 (4.3.35).

证明 性质的必要性已经清楚了. 下面证明充分性. 就是说, 如果双射 $f:\mathbb{A}\to\mathbb{A}$ 有性质 (4. 3.35), 那么它是仿射变换.

首先, 对这样的双射 f, 定义映射

$$\mathcal{F}: V\to V, \quad \overrightarrow{pq} = \overrightarrow{f(p)f(q)}.$$

性质 (4.3.35) 说, 如果 $\overrightarrow{rs}=\overrightarrow{pq}$, 那么 $\overrightarrow{f(r)f(s)}=\overrightarrow{f(p)f(q)}$. 所以 $\mathcal{F}$ 是确切定义的: 在 $\overrightarrow{pq}$ 上的作用由向量 $\overrightarrow{pq}$ 决定, 与选取表达这个向量的方式无关. 性质 (4.3.35) 还说映射 $\mathcal{F}$ 保持纯量乘: $\mathcal{F}(\lambda v)=\lambda\mathcal{F}v$, $\forall v\in V$. 其实, 它还能推出 $\mathcal{F}$ 保持加法. 对 $u,v\in V$, 可以找到 $p,q,r\in\mathbb{A}$ 使得 $u=\overrightarrow{pq}$, $v=\overrightarrow{qr}$. 于是 $u+v=\overrightarrow{pq}+\overrightarrow{qr}=\overrightarrow{pr}$. 从而

$$\mathcal{F}(u+v)=\overrightarrow{f(p)f(r)}=\overrightarrow{f(p)f(q)}+\overrightarrow{f(q)f(r)}=\mathcal{F}u+\mathcal{F}v.$$

可见, $\mathcal{F}$ 是 V 上的线性算子. 由于 f 是双射, 所以 $\mathcal{F}$ 是双射, 从而 $\mathcal{F}$ 是可逆的线性算子.

对任意的点 $p\in\mathbb{A}$ 和向量 $x\in V$, 命 $q=p+x$. 有

$$f(q)=f(p)+\overrightarrow{f(p)f(q)}, \quad \overrightarrow{f(p)f(q)}=\mathcal{F}(\overrightarrow{pq})=\mathcal{F}x.$$

即 $f(p+x)=f(p)+\mathcal{F}x$. 也就是说, f 是仿射变换. □

仿射变换一个常用的性质是它保持共线三点的简单比率.

定义 4.54 设仿射空间 $\mathbb{A}$ 中的三点 p,q,r 落在同一条直线上, 且 $p\neq q$, 那么有纯量 $\lambda\in K$ 使得

$$\overrightarrow{pr}=\lambda\overrightarrow{pq}. \tag{4.3.36}$$

纯量 λ 称为 p,q,r 的**简单比率**(simple ratio), 记作 $[p,q,r]$.

由定理 4.53, 更确切说是性质 (4.3.35), 立即得到

定理 4.55 仿射变换保持点的共线性质和共线三点组的简单比率.

把等式 (4.3.36) 改写成 $r-p=\lambda(q-p)$, 就可以用 p,q 的重心组合表达 $r=(1-\lambda)p+\lambda q$. 如果基域 K 是实数域, 那么仿射变换把线段 pq 映到线段 $f(p)f(q)$, 而且把中点映到中点, 虽然这两条线段的长度一般不相等. 由此可见, 线段的中点是仿射几何的概念, 长度却不是. 当然, 长度是欧几里得几何的概念.

六 欧几里得空间的仿射变换 我们生活的空间可以看作欧几里得空间. 这个空间并非只有刚体运动, 还有很多其他不保持长度、角度的变化, 如投影、橡皮筋的伸缩等. 所以, 考察欧几里得空间的仿射变换是有必要的.

我们感兴趣的是仿射变换的几何性质. 仿射变换总是把平面映到平面: $f(p_0+U)=f(p_0)+Df(U)$. 已经知道的是欧几里得空间的仿射变换 f 把线段 pq 映到线段 $f(p)f(q)$, 把线段的中点映到线段的中点. 更一般地, 欧几里得空间的仿射变换 f 把平行六面体

$$P=P(\overline{op_1},\cdots,\overline{op_m})=\{o+t_1\overline{op_1}+\cdots+t_m\overline{op_m}\,|\,0\leqslant t_1,\cdots,t_m\leqslant 1\}$$

映到平行六面体

$$\begin{aligned}f(P)&=P(\overline{f(o)f(p_1)},\cdots,\overline{f(o)f(p_m)})\\&=\{f(o)+t_1\overline{f(o)f(p_1)}+\cdots+t_m\overline{f(o)f(p_m)}\,|\,0\leqslant t_1,\cdots,t_m\leqslant 1\},\end{aligned}$$

其中 $o,\ p_1,\ \cdots,\ p_m\in\mathbb{E}$. 这两个平行六面体的体积之间的联系是有意思的问题. 为简单起见, 仅考虑 $m=n=\dim\mathbb{E}$ 的情况.

取 V 的标准正交基 $v_1,\ \cdots,\ v_n$. 令

$$Df(v_i)=\sum_{k=1}^n f_{ki}v_k,\quad i=1,\cdots,n;$$

$$e_j=\overline{op_j}=\sum_{i=1}^n a_{ij}v_i,\quad j=1,\cdots,n.$$

那么, 对 $j=1,\cdots,n$, 有

$$\begin{aligned}Df(e_j)=\overline{f(o)f(p_j)}&=\sum_{i=1}^n a_{ij}Df(v_i)\\&=\sum_{i=1}^n\sum_{k=1}^n f_{ki}v_k=\sum_{k=1}^n\left(\sum_{i=1}^n f_{ki}a_{ij}\right)v_k.\end{aligned}$$

命 $b_{kj}=\sum\limits_{i=1}^n f_{ki}a_{ij}$, 则有

$$B=FA,$$

其中 $B=(b_{kj})$, $F=(f_{ki})$, $A=(a_{ij})$. 根据式 (4.2.23), P 的体积是 $S_p=|\det A|$, $f(P)$ 的体积是

$$S_{f(P)}=|\det B|=|\det(FA)|=|\det F|\cdot|\det A|=|\det F|\cdot S_p.$$

这个等式的几何含义是清晰的.

定理 4.56　欧几里得空间 $\mathbb{E}$ 的仿射变换 f 把平行六面体 P 变到平行六面体. 如果平行六面体 P 由 $n=\dim\mathbb{E}$ 个向量构建, 那么它的像 $f(P)$ 的体积是 P 的体

积的倍数. 这个倍数对所有由 n 个向量构建的平行六面体都是一样的, 它等于 Df 的矩阵的行列式的绝对值. 由此可知, 仿射变换保持两个平行六面体的体积比.

对其他形状的几何体, 有类似的结果. 利用内积空间上的算子的结果可以赋予欧几里得空间上的仿射变换直观的几何意义.

定理 4.57 欧几里得空间 $(\mathbb{E}, V)$ 的仿射变换如果其微分的行列式为正, 则它是三个变换的合成:

(1) 平移;

(2) 绕某个点 o 的旋转;

(3) 以 o 为原点的一个直角坐标系的各个轴方向同时做伸缩.

如果这个仿射变换的微分的行列式为负, 则在还需要加一个过含 o 的平面的反射, 从而是四个有直观几何意义的变换的合成.

证明 根据定理 4.42, $\mathbb{E}$ 上的仿射变换 f 有分解 $f = t_v g$, 其中 g 保持某个点 o 不动. 按照定理 3.58, 有 $Dg = \mathcal{D}\mathcal{H}$, 其中 $\mathcal{D}$ 是 V 上的正交算子, $\mathcal{H}$ 是正定算子. 于是, 在某个直角坐标系 $\{o;\ e_1, \cdots, e_n\}$ 下, 算子 $\mathcal{H}$ 有典范形式

$$\mathcal{H}e_i = \lambda_i e_i, \quad \lambda_i > 0,\ i = 1, \cdots, n.$$

这样一来, 就有

$$f = t_v \cdot d \cdot h, \quad d(o+x) = o + \mathcal{D}x, \quad h(o+x) = o + \mathcal{H}x.$$

那么 h 在直角坐标系 $\{o;\ e_1, \cdots, e_n\}$ 的各个轴方向同时做伸缩. 如果 Dg 的行列式为正, 根据定理 3.59, d 是绕点 o 的旋转 (可以分解成一些绕含 o 的平面的旋转, 过余维数为偶数的平面的反射也是旋转的特殊情况), 我们得到了要的分解. 如果 Dg 的行列式为负, 根据定理 3.59, d 可以分解成绕点 o 的旋转和过某个含点 o 的平面的反射. 从而要求的分解存在. □

注 由于 $t_v g = g t_{Dg^{-1}(v)}$, $Dg = \mathcal{H}_1\mathcal{D}$, 其中 $\mathcal{H}_1$ 是正定算子, 所以合成 f 的方式可以是平移 · 旋转 (· 反射)· 伸缩, 可以是其他顺序如伸缩 ·(反射 ·) 旋转 · 平移, 等.

习 题 4.3

1. 证明: 中心不同且伸缩率 λ 和 μ 不相同的两个伸缩变换 (定义见例 4.6) 的乘积是: (1) 伸缩变换, 如果 $\lambda\mu \neq 1$; (2) 非平凡的平移, 如果 $\lambda\mu = 1$ 且 $\lambda, \mu \neq 1$.

2. 在一个 n 维实仿射空间中, 所有的 n 维平行六面体全等 (即对任意两个 n 维平行六面体都有仿射变换把一个变到另一个).

3. 设 $\Pi_1, \Pi_2, \Pi_1', \Pi_2' \subset \mathbb{A}$ 是平面, 方向子空间分别是 U_1, U_2, U_1', U_2'. 证明: 如果 $\dim \Pi_1 = \dim \Pi_1'$, $\dim \Pi_2 = \dim \Pi_2'$, $\dim U_1 \cap U_2 = \dim U_1' \cap U_2'$, 且交集 $\Pi_1 \cap \Pi_2$ 和 $\Pi_1' \cap \Pi_2'$ 均为空集或均为非空集, 那么存在仿射变换 $f \in \mathrm{Aff}(\mathbb{A})$, 它把 Π_1 映到 Π_1', Π_2 映到 Π_2'.

4. 仿射直线上三个点的排列顺序变动后它们的简单比率是怎么变化的. 在 $K = \mathbb{R}$ 的情形, 它的最大值和最小值是什么?

5. 设 $\lambda, \mu, \nu \in K$ 且它们及其乘积 $\lambda\mu\nu$ 都不等于 -1. 在平面上给定不在一条直线上的三个点 x, y, z, 构造一个三角形 abc 使得 x, y, z 分别在边 bc, ca, ab(或它们的延长线) 上且把相应的边分成两段, 比率分别是 $\lambda : 1$, $\mu : 1$, $\nu : 1$. (提示: 考虑分别以点 x, y, z 为中心, 把 c 映到 b, a 映到 c, b 映到 a 的伸缩变换的乘积.)

6. 描述欧几里得平面上绕不同的点的旋转 f 和 g 的合成是怎样的运动. (提示: 计算 $D(fg)$.)

7. 证明: 如果欧几里得空间的保距变换保持两个偏斜的平面不变, 且这两个平面的方向子空间的交集是零子空间, 那么它有不动点.

8. 如果欧几里得平面上的规矩运动 f 满足

$$Df = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad f(o) = (1, 1),$$

给出 f 的几何描述.

9. 如果欧几里得平面上的非规矩运动 f 满足

$$Df = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \quad f(o) = (1, -\sqrt{3}),$$

给出 f 的几何描述.

10. 对 4 维欧几里得空间的规矩运动分类.

4.4　凸　　集

一　线段、三角形、平行四边形、梯形、平行六面体等都是凸集的例子. 一般的定义如下.

定义 4.58　实数域上的仿射空间 $\mathbb{A}$ 中的集合 M 称为**凸集**如果连接 M 中任意两点的线段仍在 M 中.

在实平面上, 三角形有丰富的性质. 对实平面的多边形, 很容易分开成若干个三角形: 把一个顶点和其余的顶点都用线段连起来. 这说明, 三角形在多边形中是基本的. 高维的情形类似. 如果三角形 $\triangle$ 的三个顶点是 p_0, p_1, p_2, 那么它的三条边是 p_0p_1, p_1p_2, p_0p_2. 三角形内部的点 p 就是线段 p_0q 的一个内点, 其中 q 是线段

p_1p_2 的一个内点, 见图 4.16. 即有

$$p = \lambda_0 p_0 + \lambda q, \quad \lambda_0 + \lambda = 1, \quad \lambda_0 > 0,\ \lambda > 0,$$
$$q = \alpha_1 p_1 + \alpha_2 p_2, \quad \alpha_1 + \alpha_2 = 1, \quad \alpha_1 > 0,\ \alpha_2 > 0.$$

可见

$$p = \lambda_0 p_0 + \lambda(\alpha_1 p_1 + \alpha_2 p_2) = \lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2,$$

其中 $\lambda_1 = \lambda\alpha_1 > 0$, $\lambda_2 = \lambda\alpha_2 > 0$, 且 $\lambda_0 + \lambda_1 + \lambda_2 = \lambda_0 + \lambda = 1$.

反过来, 假设 $p = \lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2$ 及 $\lambda_0 + \lambda_1 + \lambda_2 = 1$, $\lambda_i > 0$, $i = 0,1,2$. 命 $q = \alpha_1 p_1 + \alpha_2 p_2$, 其中 $\alpha_1 = \lambda_1/(\lambda_1 + \lambda_2)$, $\alpha_2 = \lambda_2/(\lambda_1 + \lambda_2)$. 那么 q 是线段 p_1p_2 的内点, 且 $p = \lambda_0 p_0 + (\lambda_1 + \lambda_2)q$. 于是 p 是线段 p_0q 的内点, 从而是三角形 $\triangle$ 的一个内点. 于是, 以 p_0, p_1, p_2 为顶点的三角形 $\triangle$ 的内点集合 $\triangle^\circ$ 可以表述如下:

$$\triangle^\circ = \{\lambda_0 p_0 + \lambda_1 p_1 + \lambda_2 p_2 \mid \lambda_0 + \lambda_1 + \lambda_2 = 1, \quad \lambda_i > 0,\ i = 0,1,2\}.$$

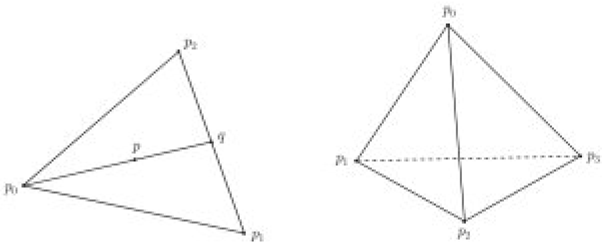


图 4.16

用类似的方式可以得到四面体. 一般情形, 则是单纯形 (simplex) 的概念.

定义 4.59 设 $p_0, p_1, \cdots, p_m$ 是实仿射空间 $\mathbb{A}$ 中 $m+1$ 个仿射无关的点. 集合

$$\{\lambda_0 p_0 + \lambda_1 p_1 + \cdots + \lambda_m p_m \mid \lambda_0 + \lambda_1 + \cdots + \lambda_m = 1, \quad \lambda_i > 0,\ i = 0,1,\cdots,m\}$$

称为以 $p_0, p_1, \cdots, p_m$ 为顶点的 m**维开单纯形**. 集合

$$\{\lambda_0 p_0 + \lambda_1 p_1 + \cdots + \lambda_m p_m \mid \lambda_0 + \lambda_1 + \cdots + \lambda_m = 1, \quad \lambda_i \geqslant 0,\ i = 0,1,\cdots,m\}$$

称为以 $p_0, p_1, \cdots, p_m$ 为顶点的 m**维闭单纯形**.

单纯形是凸集的基本例子. 在实仿射几何中, 本质上每个维数只有一个闭 (开) 单纯形.

定理 4.60　单纯形在仿射变换下的像仍是单纯形. 在同一个实仿射空间中, 任意两个 m 维闭 (开) 单纯形是叠合的.

证明　由于仿射映射保持重心组合 (定理 4.20), 仿射变换把任何仿射无关的点组都映成仿射无关的点组, 所以第一个断言成立. 利用定理 4.20 可以看出, 本质上, 第二个断言就是定理 4.52 所说的结论. □

显然, 任意多个凸集的交集仍是凸集.

定义 4.61　包含集合 M 的所有凸集的交集称为集合 M 的**凸闭包**, 记作 $C(M)$. 集合 M 的凸闭包也可以刻画为包含 M 的最小凸集.

显然, $C(M)=M$ 的充要条件是 M 本身为凸集. 以 $p_0,p_1,\cdots,p_m$ 为顶点的 m 维闭单纯形就是这组点的凸闭包.

命题 4.62　设 $M\in\mathbb{A}$ 是凸集, $p\in\mathbb{A}$. 那么

$$C(M\cup\{p\})=\bigcup_{q\in M}pq.$$

证明　对任意的 $q\in M$, 按定义, 线段 pq 落在每一个包含 M 和 p 的凸集中, 所以有 $M\cup\{p\}\subset\bigcup\limits_{q\in M}pq\subset C(M\cup p)$. 于是只要证明 $\bigcup\limits_{q\in M}pq$ 是凸集.

任取 $q_1,q_2\in M$. 对任意的点 $r_1\in pq_1$, $r_2\in pq_2$, 需要说明线段 r_1r_2 中的任何一个内点属于某个线段 pq, $q\in M$.

首先, 设点 p,q_1,q_2 不在一条直线上, 那么可以这三个点为顶点作三角形. 直观上看 (见图 4.17), 线段 pr 的延长线与线段 q_1q_2 相交于某一点 q. 由 M 的凸性可知 $q\in M$. 从而 $r\in pq$. 把这个直观用公式写下来也是简单的. 我们有

$$r_1=\lambda p+\lambda_1q_1,\quad r_2=\mu p+\mu_2q_2,\quad r=\nu_1r_1+\nu_2r_2.$$

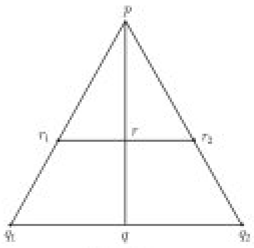


图 4.17

如果系数 λ_1, μ_2, ν_1, ν_2 中有等于 0 的, 取 $q=q_1$ 或 q_2, 可知 $r\in pq$. 如果那四个系数均不为 0, 令

$$q=\frac{\lambda_1\nu_1}{\lambda_1\nu_1+\mu_2\nu_2}q_1+\frac{\mu_2\nu_2}{\lambda_1\nu_1+\mu_2\nu_2}q_2\in q_1q_2.$$

则有

$$r = \nu_1 p(\lambda + \mu) + (\lambda_1\nu_1 + \mu_2\nu_2)q \in pq.$$

如果 p, q_1, q_2 在一条直线上, 那么线段 pq_1 和 pq_2 至少有一个含 r, 于是取 q_1 或 q_2 作为 q 即可. □

二 多面体 超平面和半空间也是凸集中基本的例子. 对实仿射空间 $\mathbb{A}$ 上的非常数仿射线性函数 $f : \mathbb{A} \to \mathbb{R}$, 命

$$\begin{aligned} H_f &= \{p \in \mathbb{A} \mid f(p) = 0\}, \\ H_f^+ &= \{p \in \mathbb{A} \mid f(p) \geqslant 0\}, \\ H_f^- &= \{p \in \mathbb{A} \mid f(p) \leqslant 0\} = H_{-f}^+. \end{aligned}$$

那么 H_f 是超平面, 它显然是凸集. 集合 H_f^+ 和 H_f^- 称为以超平面 H_f 为边界的**(闭的) 半空间**(half-space). 把定理 4.20 (1) 应用到仿射线性函数可知半空间是凸集. 一个线段如果端点分别在 H_f^+ 和 H_f^- 中, 那么这个线段穿过超平面 H_f, 即线段与超平面的交是一个点.

定义 4.63 有限多个半空间的交集称为**凸多面体**(polyhedron).

仿射空间中的集合 M 称为**有界的**如果它落在某个单纯形中. 闵可夫斯基–外尔的一个定理说: 有界的凸多面体一定是有限多个点的凸闭包, 反之亦然.[①]

凸多面体不仅在数学的理论上很重要, 也有很多的应用, 尤其是在线性规划中. 线性规划的基石之一是下面的结论.

定理 4.64 有限个点的凸闭包上的仿射线性函数的极大值一定在某个顶点上达到.

证明 设 $M = C(p_0, p_1, \cdots, p_m)$ 是 m 个点的凸闭包. 当 $m = 0$ 时, 结论是平凡的. 现对 m 作归纳法, 于是 M 上的仿射线性函数 f 在 $M_1 = C(p_0, p_1, \cdots, p_{m-1})$ 上的极大值是 $\max\limits_{0 \leqslant i < m} f(p_i)$. 由命题 4.62, M 的任何点 s 必然属于某个线段 $p_m q$, $q \in M_1$. 这意味着

$$s = \lambda p_m + (1 - \lambda)q, \quad 0 \leqslant \lambda \leqslant 1.$$

从而

$$f(s) = \lambda f(p_m) + (1 - \lambda) f(q) \leqslant \max\{f(p_m), f(q)\} \leqslant \max_{0 \leqslant i \leqslant m} f(p_i).$$ □

下面两个例子 [②] 显示线性规划的问题是自然产生的.

例 4.65 (最大收益问题) 某公司有原材料 $R_1, \cdots, R_m$, 数量分别是 $b_1, \cdots, b_m$, 并计划生产物品 $P_1, \cdots, P_n$, 数量分别是 $x_1, \cdots, x_n$. 命 a_{ij} 是产出一个单位 P_j 所

① 证明可见 E.B. Vinberg, A Course in Algebra, p.254. American Mathematical Society, 2003.
② 取自 E.B. Vinberg, A Course in Algebra, p.257-258.

需原材料 R_i 的数量, c_j 是产品 P_j 的销售单价. 很清楚, 下面的不等式成立:

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i=1,\cdots,m, \quad x_j \geqslant 0, \quad j=1,\cdots,n.$$

这些不等式确定了一个以 $x_1,\cdots,x_n$ 为坐标的 n 维空间中的凸多面体 M. 要取得最大收益, 就需要找到点 $(x_1,\cdots,x_n)\in M$ 使得线性函数 $\sum\limits_{j=1}^{n} c_jx_j$(总的销售价) 取最大值.

例 4.66 (运输问题)　某产品在供货点 $A_1,\cdots,A_m$ 分别有数量 $a_1,\cdots,a_m$. 顾客 $B_1,\cdots,B_n$ 需要这种产品的数量分别为 $b_1,\cdots,b_n$. 同时, 供求平衡: $\sum\limits_{i=1}^{m} a_i=\sum\limits_{j=1}^{n} b_j$. 假设从供货点 A_i 送到顾客 B_j 的产品数量是 x_{ij}, 每件产品的运费是 c_{ij}. 有下面的等式:

$$\sum_{j=1}^{n} x_{ij}=a_i, \quad \sum_{i=1}^{m} x_{ij}=b_j, \quad x_{ij}\geqslant 0.$$

它们定义了以 x_{ij}, $i=1,\cdots,m$, $j=1,\cdots,m$ 为坐标的 mn 维空间中的一个凸多面体. 问题是要确定线性函数 $\sum\limits_{i,j} c_{ij}x_{ij}$ 在这个多面体上的最小值.

习　题　4.4

1. 证明: 凸集在仿射映射下的像是凸集.

2. 证明: 凸集在仿射映射下的逆像是凸集.

3. 设 $p_0,p_1,\cdots,p_m$ 是实仿射空间 $\mathbb{A}$ 中的点. 证明: 集合

$$\left\{\sum_{i=0}^{m}\lambda_ip_i \,\middle|\, \sum_{i=0}^{m}\lambda_i=1,\ \lambda_i\geqslant 0\ \forall i\right\}$$

是凸集, 是有限个单纯形的并, 是点组 $p_0,p_1,\cdots,p_m$ 的凸闭包.

4. 欧几里得向量空间 V 同时看作仿射空间. 设 $C\subset V$ 是凸集, 它的对偶 $C^\vee$ 定义为

$$C^\vee=\{u\in V\mid \langle u|v\rangle\leqslant 1,\forall v\in C\}.$$

证明:

(1) $C^\vee$ 是凸集;

(2) 如果 C 的内部含有 0, 那么 $C^\vee$ 是有界的; (称集合 S 中的点 p 在 S 中的内部如果 S 包含某个含 p 的开球.)

(3) 如果 C 有界, 那么 $C^\vee$ 的内部含有 0;

(4) 如果 C 是有界闭集且 0 在 C 的内部, 那么 C 是 $C^\vee$ 的对偶; (称集合 $S\subset V$ 是闭集如果 S 中的任何柯西序列的极限都在 S 中.)

(5) 如果 P 是凸多面体且 0 在其内部, 那么 $P^\vee$ 也是凸多面体.

5. 设 $\mathbb{A}$ 是维数大于 1 的实仿射空间, $f:\mathbb{A}\to\mathbb{A}$ 是双射. 证明: 如果 f 把任何凸集映成凸集, 那么 f 是仿射变换.

6. (Kakutani(角谷) 引理) 设 C_1 和 C_2 是两个不相交的凸集, 点 $x\notin C_1\bigcup C_2$. 用 Γ_i 记 $\{x\}\bigcup C_i$ $(i=1,2)$ 的凸闭包. 证明: 集合 $\Gamma_1\cap C_2$ 与 $\Gamma_2\cap C_1$ 中至少有一个是空集.

4.5 伪欧几里得空间和闵可夫斯基空间

一　伪欧几里得向量空间和伪欧几里得仿射空间　欧几里得向量空间的本质是带着一个固定的正定二次型的实向量空间. 对欧几里得向量空间和欧几里得仿射空间的讨论就是把向量空间与仿射空间和这个正定的二次型放在一起研究. 这个想法可以沿用到带有其他二次型的任意域上的向量空间. 不过, 眼下我们的兴趣还是在实向量空间.

设 V 是 n 维实向量空间, q 是其上的非退化二次型. 那么, 存在 V 的基 (e_i) 使得 q 在这个基下有标准形 (参见定理 1.80):

$$q(x)=x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2. \tag{4.5.37}$$

相应的纯量积是

$$(x|y)=x_1y_1+\cdots+x_sy_s-x_{s+1}y_{s+1}-\cdots-x_ny_n. \tag{4.5.38}$$

空间 V 与二次型 q 合在一起 (V,q) 称为一个 (符号差为 $(s,n-s)$ 的**伪欧几里得向量空间**. 和过去一样, V 中向量 x 的模 (长度) 的平方定义为 $\|x\|^2=(x|x)$. 在负惯性指数 $n-s$ 大于 0 时, 它可以是正的, 也可以是 0, 还可以取负值. 如果 $\|x\|^2=0$, 则说 x 是**迷向的**.

与 V 关联的仿射空间称为**伪欧几里得仿射空间**, 简称为**伪欧几里得空间**. 在这个仿射空间中, 两点 $p(x_1,\cdots,x_n)$ 和 $q(x_1,\cdots,x_n)$ 的 "距离" 的平方是

$$\rho(p,q)^2=\sum_{i=1}^{s}(y_i-x_i)^2-\sum_{i=s+1}^{n}(y_i-x_i)^2.$$

二　伪欧几里得空间中的运动　伪欧几里得空间 $\mathbb{E}$ 的几何应该通过伪欧几里得空间的运动群 $\mathrm{Iso}(\mathbb{E})$ 进行研究. 和欧几里得空间中的运动一样, 伪欧几里得空间中的运动是一个仿射变换 $f:\mathbb{E}\to\mathbb{E}$, 满足条件

$$\rho(f(p),f(q))=\rho(p,q).$$

这等价于 f 的微分 $\mathcal{F}$ 保持纯量积 (4.5.38):

$$(\mathcal{F}x|\mathcal{F}y) = (x|y).$$

也等价于 $\mathcal{F}$ 保持二次型 (4.5.37):

$$q(\mathcal{F}x) = q(x).$$

向量空间 V 上的保持纯量积 (4.5.38) 的可逆线性算子全体形成的群称为**伪正交群**, 记作 $O(V)$ 或 $O(s, n-s)$. 它也称为二次型 q 的**自同构群**. 当 $s = n$ 或 $s = 0$ 时, 它就是正交群 $O(n)$.

在运动群 $\mathrm{Iso}(\mathbb{E})$ 中保持某个固定点 o 不动的运动全体形成 $\mathrm{Iso}(\mathbb{E})$ 的一个子群. 这个子群中的运动完全由其微分决定, 所以这个子群与 $O(s, n-s)$ 同构. 平移显然是伪欧几里得空间中的运动. 于是, 定理 4.41(3) 中的正合列限制在 $\mathrm{Iso}(\mathbb{E})$ 上给出了正合列:

$$e \to T \xrightarrow{\theta} \mathrm{Iso}\,(\mathbb{E}) \xrightarrow{D} O(s, n-s)) \to \bar{e}.$$

其中 θ 是嵌入映射, 即 $\theta(t_v) = t_v$; 映射 D 把 f 映到其微分.

选取 V 的基 $e_1, \cdots, e_n$ 使得纯量积有典范形式 (4.5.38). 那么, 纯量积 (作为双线性型) 在这个基下的矩阵是

$$D_s = \begin{pmatrix} I_s & 0 \\ 0 & -I_{n-s} \end{pmatrix}.$$

从而算子 $\mathcal{F} \in O(s, n-s)$ 在这个基下的矩阵 F 满足 (参见定理 1.69)

$${}^tF \cdot D_s \cdot F = D_s,$$

等式左边是纯量积在基 $\mathcal{F}e_1, \cdots, \mathcal{F}e_n$ 下的矩阵. 由此可知, $\det \mathcal{F} = \det F = \pm 1$. 如果 $\det \mathcal{F} = 1$, 则说 $\mathcal{F}$ 是二次型 q 的**规矩自同构**, 以 $\mathcal{F}$ 为微分的仿射变换 $f: \mathbb{E} \to \mathbb{E}$ 则称为**规矩的伪欧几里得运动**. 值得注意的是, 二次型 q 的自同构把迷向的向量变到迷向的向量, 因为 $q(\mathcal{F}x) = \|\mathcal{F}x\|^2 = (\mathcal{F}x|\mathcal{F}x) = (x|x) = \|x\|^2 = q(x) = 0$.

三　闵可夫斯基空间和洛伦兹变换　正惯性指数为 1 和负惯性指数为 1 的情形需要特别关注. 相应的伪欧几里得空间称为**闵可夫斯基空间**. 由于 q 的负惯性指数为 1 意味着 $-q$ 的正惯性指数为 1, 所以只需关注正惯性指数为 1 的情形. 四维的闵可夫斯基空间在狭义相对论中非常重要, 是时间和空间的统一体, 称为**闵可夫斯基时空**(spacetime) 或**四维时空**. 在这种情况下, 下面的表达方式是标准的:

$$\begin{aligned} &V = \langle e_0, e_1, e_2, e_3 \rangle, \\ &x = te_0 + x_1e_1 + x_2e_2 + x_3e_3, \\ &\|x\|^2 = q(x) = t^2 - x_1^2 - x_2^2 - x_3^2. \end{aligned}$$

空间 V 自然是仿射空间 (参见例 4.2), 看作仿射空间时记作 $\mathbb{E}$. 四维闵可夫斯基空间的运动群 $\mathrm{Iso}(\mathbb{E})$ 称为**庞加莱群**, 保持原点不动的运动全体形成的群称为**洛伦兹群**, 记为 $\mathbf{L}$. 洛伦兹群自然同构于伪正交群 $O(1,3)$, 利用这个同构, 可以把它们等同起来. 庞加莱群和洛伦兹群在相对论中都是十分重要的.

狭义相对论中为人熟知的**洛伦兹变换**(事实上出现在相对论之前) 是

$$t' = \gamma\left(t - \frac{vx_1}{c^2}\right), \quad x_1' = \gamma(x_1 - vt), \quad x_2' = x_2, \quad x_3' = x_3, \tag{4.5.39}$$

其中 $\gamma = c/\sqrt{c^2 - v^2}$, v 是速度, c 是真空中的光速. 它描述了一个事件在一个参照系中的时空坐标和在另一个参照系中的时空坐标之间的关系. 洛伦兹变换保持二次型 $c^2t^2 - x_1^2 - x_2^2 - x_3^2$ 不变, 其实只是一个二维时空的变换. 如果以光速 c 为单位, 那么洛伦兹变换保持不变的二次型具有标准形式.

二维时空可以看成四维时空的"玩具模型". 洛伦兹变换表明这个玩具模型是有趣的. 二维时空中保持度量

$$(te_0 + x_1e_1 \,|\, te_0 + x_1e_1) = t^2 - x_1^2 \tag{4.5.40}$$

的可逆线性算子全体形成伪正交群 $O(1,1)$, 也称为**1 阶洛伦兹群**, 记作 $\mathbf{L}_1$. 它刻画了直线上的物理运动. 显然, 所有的迷向向量都与 $e_0 + e_1$ 或 $e_0 - e_1$ 成比例. 于是, 对于保持度量 (4.5.39) 的可逆线性算子 $\mathcal{F}$, 只有两种可能:

$$\mathcal{F}(e_0 + e_1) = \alpha(e_0 + e_1), \quad \mathcal{F}(e_0 - e_1) = \beta(e_0 - e_1);$$
$$\mathcal{F}(e_0 + e_1) = \alpha(e_0 - e_1), \quad \mathcal{F}(e_0 - e_1) = \beta(e_0 + e_1).$$

如果是第一种情况, 则有

$$\mathcal{F}e_0 = \frac{\alpha+\beta}{2}e_0 + \frac{\alpha-\beta}{2}e_1,$$
$$\mathcal{F}e_1 = \frac{\alpha-\beta}{2}e_0 + \frac{\alpha+\beta}{2}e_1.$$

算子的矩阵是对称的:

$$F = \begin{pmatrix} \dfrac{\alpha+\beta}{2} & \dfrac{\alpha-\beta}{2} \\ \dfrac{\alpha-\beta}{2} & \dfrac{\alpha+\beta}{2} \end{pmatrix},$$

其行列式为 $\det \mathcal{F} = \alpha\beta$. 我们限于讨论规矩洛伦兹变换的情形, 也就是说 $\alpha\beta = 1$. 算子 $\mathcal{F}$ 给出的坐标变换是

$$\begin{pmatrix} t' \\ x_1' \end{pmatrix} = F\begin{pmatrix} t \\ x_1 \end{pmatrix} = \begin{pmatrix} \dfrac{\alpha+\alpha^{-1}}{2} & \dfrac{\alpha-\alpha^{-1}}{2} \\ \dfrac{\alpha-\alpha^{-1}}{2} & \dfrac{\alpha+\alpha^{-1}}{2} \end{pmatrix}\begin{pmatrix} t \\ x_1 \end{pmatrix}.$$

从而

$$t' = \frac{\alpha+\alpha^{-1}}{2}\left(t + \frac{\alpha-\alpha^{-1}}{\alpha+\alpha^{-1}}x_1\right);$$
$$x_1' = \frac{\alpha+\alpha^{-1}}{2}\left(\frac{\alpha-\alpha^{-1}}{\alpha+\alpha^{-1}}t + x_1\right).$$

引进表达式

$$v = \frac{\alpha^{-1}-\alpha}{\alpha+\alpha^{-1}} = \frac{1-\alpha^2}{1+\alpha^2}. \tag{4.5.41}$$

注意此处 v 是纯量, 而非像以前那样表示向量. 这个纯量对应物理上的速度, 遵循物理的规矩, 速度用记号 v 表示. 由于 $\alpha \neq 0$, 有 $|v| < 1$. 从而, 关系式 (4.5.42) 导出的表达式

$$\alpha^2 = \frac{1-v}{1+v}, \quad \alpha = \sqrt{\frac{1-v}{1+v}},$$
$$\frac{\alpha+\alpha^{-1}}{2} = \frac{1}{\sqrt{1-v^2}},$$

是有意义的.

最后得到

$$t' = \frac{t - vx_1}{\sqrt{1-v^2}}, \quad x_1' = \frac{x_1 - vt}{\sqrt{1-v^2}}. \tag{4.5.42}$$

它就是洛伦兹变换 (4.5.39) 在光速 c 取成单位 1 的形式. 在通常的尺度下, 二次型 $t^2 - x_1^2$ 要换成 $c^2t^2 - x_1^2$, 从而要用 v/c 代替 v, ct 代替 t, ct' 代替 t'. 这时, 式 (4.5.42) 就有式 (4.5.39) 的形式. 麦克斯韦电磁学方程在洛伦兹变换下是不变的. 更一般地, 狭义相对性原理说: 一切物理定律在洛伦兹变换下都是不变的.

看似简单的洛伦兹变换的物理意义是丰富的. 如果速度 v 接近 0(即与光速比较很小), 伽利略变换

$$t' = t, \quad x_1' = x - ct$$

就成为洛伦兹变换 (4.5.42) 的近似形式.

如果 t 取同一值 t_0, x_1 取不同的值 a 和 b, 那么 t' 的取值分别是

$$t_a' = \frac{t_0 - va}{\sqrt{1-v^2}}, \quad t_b' = \frac{t_0 - vb}{\sqrt{1-v^2}};$$

x_1' 的取值分别是

$$a' = \frac{a - vt_0}{\sqrt{1-v^2}}, \quad b' = \frac{b - vt_0}{\sqrt{1-v^2}}.$$

由此可知, 在同一个参照系下两个不同地方同时发生的事件在另一个参照系下看可

能不是同时发生的 (同时性的相对性, relativity of simultaneity). 同样, 物体的长度在不同的参照系下也可以是不一样的 (length contraction):

$$a' - b' = \frac{a - vt_0}{\sqrt{1-v^2}} - \frac{b - vt_0}{\sqrt{1-v^2}} = \frac{a-b}{\sqrt{1-v^2}}.$$

当 x_1 取相同的值, t 取不同的值 t_0 和 t_1 时, 就得到时间伸缩律 (time dilation)

$$\Delta t' = t_0' - t_1' = \frac{t_0 - vx_1}{\sqrt{1-v^2}} - \frac{t_1 - vx_1}{\sqrt{1-v^2}} = \frac{t_0 - t_1}{\sqrt{1-v^2}} = \frac{\Delta t}{\sqrt{1-v^2}}.$$

以 f_v 记变换 (4.5.42). 两个洛伦兹变换的复合仍是洛伦兹变换 (都保持二次型 $t^2 - x_1^2$), 所以有

$$f_{v_2} \cdot f_{v_1} = f_v.$$

设

$$\begin{aligned} t' &= \frac{t - v_1 x_1}{\sqrt{1-v_1^2}}, \quad x_1' = \frac{x_1 - v_1 t}{\sqrt{1-v_1^2}}; \\ t'' &= \frac{t' - v_2 x_1'}{\sqrt{1-v_2^2}}, \quad x_1' = \frac{x_1' - v_2 t'}{\sqrt{1-v_2^2}}, \end{aligned}$$

得

$$t'' = \frac{t - v_1 x_1 - v_2(x_1 - v_1 t)}{\sqrt{1-v_1^2} \cdot \sqrt{1-v_2^2}} = \frac{t - (v_1+v_2)x_1/(1+v_1v_2)}{\sqrt{1-(v_1+v_2)^2/(1+v_1v_2)^2}}.$$

这意味着速度的复合律是

$$v = \frac{v_1 + v_2}{1 + v_1 v_2}.$$

四　洛伦兹群　现在转向洛伦兹群 $\mathbf{L}$, 即二次型

$$q(x) = t^2 - x_1^2 - x_2^2 - x_3^2 \tag{4.5.43}$$

的自同构群, 群中的元素将称为**洛伦兹变换**. 保持这个二次型的变换的公式并不简单, 这也意味着直接描述洛伦兹群中所有的元素并不容易. 不过, 对保持空间定向和时间方向的洛伦兹变换, 有漂亮的方式描述. 称一个洛伦兹变换是**规矩的**(proper) 如果它的行列式等于 1, 这等价于变换保持空间的定向 (参见第一卷 4.1 节关于有向体积的讨论). 所谓保持时间方向就是保持正向光锥 (稍后定义) 不变.

对四维实空间 $\mathbb{R}^4$ 中的向量 $x = (t, x_1, x_2, x_3)$, 命

$$P_x = \begin{pmatrix} t - x_1 & x_2 - ix_3 \\ x_2 + ix_3 & t + x_1 \end{pmatrix}.$$

这是一个 2 阶埃尔米特矩阵. 所有这样的矩阵形成一个 4 维的实向量空间 U. 由于

$$P_{\alpha x+\beta y}=\alpha P_x+\beta P_y,$$

所以, 映射 $x\to P_x$ 是从 $\mathbb{R}^4$ 到 U 的线性同构.

通过矩阵构造 U 上的线性算子是容易的. 对每个非退化的二阶复方阵 A, 定义

$$\Gamma_A(P_x)=A\cdot P_x\cdot A^*.$$

由于

$$[\Gamma_A(P_x)]^*=A^{**}\cdot P_x^*\cdot A^*=A\cdot P_x\cdot A^*=\Gamma_A(P_x),$$
$$\Gamma_A(\alpha P_x+\beta P_y)=\alpha\Gamma_A(P_x)+\beta\Gamma_A(P_y),$$

所以 Γ_A 是 U 上的可逆线性算子.

如果 A 的行列式为 1, 即 A 是群 $SL_2(\mathbb{C})$ 中的元素, 则有

$$\det\Gamma_A(P_x)=\det A\cdot\det P_x\cdot\det A^*=\det P_x.$$

注意

$$\det P_x=t^2-x_1^2-x_2^2-x_3^2.$$

所以, 此时 Γ_A 保持二次型 (4.5.43), 从而是洛伦兹变换. 特别, $\det\Gamma_A=\pm1$. 不难验证, $\det\Gamma_A=1$. 从拓扑学的角度看这事显然: 因为函数 det 是连续的, 且群 $SL_2(\mathbb{C})$ 是连通的. 于是 Γ_A 保持空间的方向.

方程

$$t^2-x_1^2-x_2^2-x_3^2=0 \tag{4.5.44}$$

的解集构成 $\mathbb{R}^4$ 中的一个锥面–一个特殊的二次曲面 (第 5 章会对二次曲面做更多的探讨). 它包含过坐标原点和曲面上任意其他点的直线. 锥面中时间为正的部分: $q(x)=0$, $t>0$, 所围住的空间称为**正向光锥**(forward light cone). 正向光锥由两个不等式确定:

$$t>0,\quad t^2-x_1^2-x_2^2-x_3^2>0. \tag{4.5.45}$$

矩阵 P_x 的特征值是 $t\pm\sqrt{x_1^2+x_2^2+x_3^2}$. 这意味着, 矩阵 P_x 是正定的, 更确切地说, 以它为矩阵的埃尔米特型是正定的 (参见 3.2 节, 尤其是其第二部分), 当且仅当式 (4.5.45) 中的不等式成立. 显然, 变换 Γ_A 保持这个正定性条件, 即 P_x 正定意味着

$$\Gamma_A(P_x)=A\cdot P_x\cdot A^*$$

正定. 所以, 变换 Γ_A 保持正向光锥不变, 即保持时间方向不变.

当 $A \in SL_2(\mathbb{C})$ 时, 变换 Γ_A 具有如下性质:

(1) 为洛伦兹变换, 即保持二次型 (4.5.43) 不变;

(2) 行列式等于 1;

(3) 保持正向光锥不变.

定义 4.67 满足条件 (1)—(3) 的线性算子 $\mathcal{F}: \mathbb{R}^4 \to \mathbb{R}^4$ 称为**真洛伦兹变换**. 所有这种变换构成的群称为**受限洛伦兹群**(restricted Lorentz group), 记作 $\mathbf{L}^+$.

真洛伦兹变换都有形式 Γ_A, $A \in SL_2(\mathbb{C})$. 也就是说, 同态映射

$$\Gamma : SL_2(\mathbb{C}) \to \mathbf{L}^+, \quad A \to \Gamma_A$$

是满射. 现在确定它的核. 条件 $\Gamma_A = \mathcal{E}$ 意味着对任意的 $x \in \mathbb{R}^4$ 有

$$A \cdot P_x \cdot A^* = P_x.$$

对 $e = (1, 0, 0, 0)$, 有 $P_e = E$. 从而 $AA^* = E$, 进而 $A^* = A^{-1}$. 由此可知

$$AP_x = P_x A.$$

任何二阶复方阵都是埃尔米特矩阵的线性组合 (参见定理 3.39 的证明), 因此上式蕴含着 A 和所有的二阶复方阵交换. 这只有纯量矩阵能做到, 所以 $A = \alpha E$. 又因为 A 的行列式为 1, 所以 $\alpha = \pm 1$. 换句话说 $\ker \Gamma = \{\pm E\}$.

我们把上面的讨论总结如下.

定理 4.68 行列式值为 1 的二阶复矩阵与规矩洛伦兹变换之间的对应 $\Gamma : A \to \Gamma_A$ 是群 $SL_2(\mathbb{C})$ 到所有真洛伦兹变换构成的群 $\mathbf{L}^+$ 的一个同态映射. 每个真洛伦兹变换恰由两个矩阵 A 和 $-A$ 与之对应, 它们相差一个负号.

由于这个定理, 常常把 $SL_2(\mathbb{C})$ 称为洛伦兹群, 尽管更准确的说法应该是它的商群 $SL_2(\mathbb{C})/\{\pm E\}$ 才对.

受限洛伦兹群 $\mathbf{L}^+$ 和洛伦兹群 $\mathbf{L}$ 的差别不大, 前者是后者的含单位元的连通分支. 变换 $\mathcal{P} : (t, x_1, x_2, x_3) \to (t, -x_1, -x_2, -x_3)$ 和 $\mathcal{T} : (t, x_1, x_2, x_3) \to (-t, x_1, x_2, x_3)$ 都是洛伦兹变换, 但不是规矩洛伦兹变换. 不难验证

$$\mathbf{L} = \mathbf{L}^+ \cup \mathcal{P}\mathbf{L}^+ \cup \mathcal{T}\mathbf{L}^+ \cup \mathcal{P}\mathcal{T}\mathbf{L}^+. \tag{4.5.46}$$

二次型 $q(x)$ 除了给出前面的锥面外, 还可以通过方程

$$t^2 - x_1^2 - x_2^2 - x_3^2 = c, \quad c \in \mathbb{R}$$

定义曲面 S_c. 按定义, 洛伦兹变换保持 $q(x)$ 不变, 所以把 S_c 变成自身. 如果 $c > 0$, S_c 是双叶双曲面, $c < 0$ 时, S_c 是单叶双曲面, 我们已经知道 S_0 是锥面.

(这些术语源自三维空间的解析几何.) 在这些曲面上同样可以考虑运动 (即保持距离的变换). 曲面的一个典型例子是欧几里得向量空间 $\mathbb{R}^n$ 中的单位球面 $S^{n-1} = \left\{(x_1, \cdots, x_n) \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^{n} x_i^2 = 1\right\}$, 其上的运动都是 $\mathbb{R}^n$ 上的正交算子在球面上的限制.

双叶双曲面 S_1 的上叶

$$t^2 - x_1^2 - x_2^2 - x_3^2 = 1, \quad t \geqslant 1$$

连同它确定的运动群 $\mathbf{L}^+$(或者 $SL_2(\mathbb{C})$) 可以作为三维罗巴切夫斯基空间 Λ^3 的一个模型. 这里不宜开展罗巴切夫斯基几何的讨论, 但适宜讨论群在曲面上的作用的性质, 更确切地说, 我们关心群作用的可迁性, 即, 对曲面上任意两点 p, q 是否有群中的元素把一个点变到另一个点: $g(p) = q$. 在定理 4.51 那儿已经知道仿射群 $\mathrm{Aff}(\mathbb{A})$ 在仿射空间 $\mathbb{A}$ 上是可迁的, 保距群 $\mathrm{Iso}(\mathbb{E})$ 在欧几里得空间 $\mathbb{E}$ 上是可迁的. 正交群 $O(n)$ 在球面 S^{n-1} 上是可迁的.

我们说真洛伦兹群 $\mathbf{L}^+$ 在 Λ^3 上是可迁的. 和过去一样, 把 $\mathbb{R}^4$ 中的点 $x = (t, x_1, x_2, x_3)$ 和埃尔米特矩阵 P_x 等同起来. 当 $x \in \Lambda^3$ 时, P_x 是正定的, 所以存在二阶可逆复方阵 A 使得

$$P_x = AA^* = A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^*.$$

此时有 $\det P_x = t^2 - x_1^2 - x_2^2 - x_3^2 = 1$, 所以 $\det A \cdot \det A^* = \det A \cdot \overline{\det A} = 1$. 命 $A' = A \cdot \mathrm{diag}\,(\det A^*, 1)$, 那么 $A' \in SL_2(\mathbb{C})$ 且 $P_x = A'A'^*$. 所以, 运动 $\Gamma_{A'}$ 把矩阵 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 变到 P_x, 等价的说法是 $\Gamma_{A'}$ 把点 $(1, 0, 0, 0) \in \Lambda^3$ 变到点 $x = (t, x_1, x_2, x_3) \in \Lambda^3$. 于是 $\mathbf{L}^+$ 在 Λ^3 上是可迁的.

点 $e = (1, 0, 0, 0)$ 的稳定子群 $\mathbf{L}_e^+$ 由那些 Γ_A 组成, 其中的矩阵 $A \in SL_2(\mathbb{C})$ 满足条件

$$A \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

也就是 $AA^* = E$. 这样的矩阵全体形成特殊酉群 $SU(2)$. 所以

$$\mathbf{L}_e^+ \simeq SU(2)/\{\pm E\} \simeq SO(3).$$

(后面那个同构是有趣的结论, 不过, 我们目前并不需要, 证明可在柯斯特利金著的《代数学引论》第三卷找到.) 罗巴切夫斯基几何也称为双曲几何, 空间 Λ^3 的运动还被称为**双曲旋转**.

习 题 4.5

1. 讨论一下算子 $\mathcal{F} \in \mathbf{L}_1$ 作用的第二种情况.
2. 证明: 第四部分中定义的算子 Γ_A 的行列式等于 1.
3. 证明: 同态映射

$$\Gamma : SL_2(\mathbb{C}) \to \mathbf{L}^+$$

是满射.

4. 证明等式 (4.5.46).

第 5 章 二 次 曲 面

说起来, 二次曲线, 如圆、椭圆、抛物线等, 二次曲面, 如球面、椭球面、锥面、双曲面、抛物面等不仅很常见, 也很早很早就是数学研究的对象, 古人, 尤其是古希腊人对它们的研究就已经达到很高的理论水平. 中学里接触到主要是平面上的二次曲线和三维空间中的球面. 有了一般的向量空间、仿射空间和欧几里得空间理论, 现在可以讨论高维空间的二次曲面. 上一章关于闵可夫斯基空间的研究表明高维空间的二次曲面的研究是必要的.

从另一个角度说, 线性方程 (组) 的解空间是仿射空间和欧几里得空间的平面, 它们是仿射几何和欧几里得几何最简单的对象. 自然, 接下来要考虑二次多项式方程组的解空间. 不过, 这个问题要难得多. 作为第一步, 先考虑一个二次多项式方程的解空间, 这是本章的主题. 至于二元多项式方程组的解空间, 或一般的多项式方程 (组) 的解空间, 称为**(代数) 簇**, 则是代数几何研究的主题. 代数几何过去一百多年发展迅速, 对很多的主流分支如数论、群表示论、微分几何等有深刻的影响.

假设 $\mathrm{char}K \neq 2$.

5.1 二 次 函 数

一　仿射空间上的二次函数　设 $\mathbb{A}$ 是域 K 上的仿射空间, 其关联的向量空间 V 的维数是 n. 基域 K 看作仿射空间时记作 K_a. 于是, 一阶的仿射群作用在仿射直线 K_a 上.

定义 5.1　称函数 $Q:\mathbb{A}\to K$ 是**二次函数**如果

$$Q(o+x)=q(x)+l(x)+c, \tag{5.1.1}$$

其中 o 是取定的一个点, q 是 V 上的非零二次型, l 是 V 的线性函数, c 是常量.

命 f 为二次型 q 的极化, 即给出 q 的对称双线性型.

引理 5.2　如果取另一个原点 $o'\in\mathbb{A}$, 那么等式 (5.1.1) 有如下形式:

$$Q(o'+x)=q(x)+[2f(v_0,x)+l(x)]+[q(v_0)+l(v_0)+c], \tag{5.1.2}$$

其中 $v_0=\overline{oo'}$, 即 $o'=o+v_0$.

证明　我们有

$$Q(o'+x)=Q(o+v_0+x)=q(v_0+x)+l(v_0+x)+c$$

$$= q(v_0) + 2f(v_0, x) + q(x) + l(v_0) + l(x) + c$$
$$= q(x) + [2f(v_0, x) + l(x)] + [q(v_0) + l(v_0) + c]. \qquad \square$$

由此可见二次型 q 与原点的选取无关, 但线性函数 l 和常量 c 与原点的选取有关. 可以把 q 称为 Q**确定的二次型**, 它的秩定义为 Q 的秩:

$$\operatorname{rank} Q := \operatorname{rank} q = r.$$

在 $\mathbb{A}$ 中取以 o 为原点的坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$, 得到 Q 的坐标形式:

$$Q(o+x) = \sum_{i,j=1}^{n} a_{ij}x_ix_j + \sum_{i=1}^{n} b_ix_i + c, \quad a_{ij} = a_{ji}. \tag{5.1.3}$$

常量 c 和诸系数 b_i 有如下含义:

$$c = Q(o), \quad b_i = \frac{\partial Q}{\partial x_i}(o). \tag{5.1.4}$$

线性函数

$$l(x) = \sum_{i=1}^{n} b_ix_i$$

称为 Q 在点 o 处的**微分**, 记作 d_oQ. 当 K 是实数域时, 这个定义与微分的标准定义一致.

二　双仿射函数　如同 V 上的二次型可以通过双线性函数定义, 仿射空间上的二次函数也可以通过双仿射函数实现.

定义 5.3　称函数 $\Phi: \mathbb{A}\times\mathbb{A} \to K_a = K$ 是**双仿射的**, 如果 $\Phi(p,q)$ 在点 p 固定或点 q 固定时就是 $\mathbb{A}$ 上的仿射函数. 也就是说, 对任意取定的 $p_0, q_0 \in \mathbb{A}$, 函数

$$\mathbb{A} \to K_a,\ q \to \Phi(p_0, q);$$
$$\mathbb{A} \to K_a,\ p \to \Phi(p, q_0)$$

都是仿射的. 称双仿射函数是**对称的**如果

$$\Phi(p,q) = \Phi(q,p), \quad \forall p, q \in \mathbb{A}.$$

取 $\mathbb{A}$ 的坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$. 令 $p_0 = o$, $p_1 = o + e_1, \cdots, p_n = o + e_n$, 那么 $p_0,\ p_1,\ \cdots,\ p_n$ 是 $\mathbb{A}$ 的仿射坐标系. 假设 $p = o + x$, $q = o + y$ 的坐标分别是 $x_1, \cdots, x_n$ 和 $y_1, \cdots, y_n$. 则有

$$p = \left(1 - \sum_{i=1}^{n} x_i\right)p_0 + x_1p_1 + \cdots + x_np_n, \quad q = \left(1 - \sum_{j=1}^{n} y_j\right)p_0 + y_1p_1 + \cdots + y_np_n.$$

根据定义, 有

$$\begin{aligned}
\Phi(p,q) &= \sum_{j=1}^{n}\Phi(p,p_j)y_j+\Phi(p,p_0)\left(1-\sum_{j=1}^{n}y_j\right)\\
&=\sum_{i,j=1}^{n}\Phi(p_i,p_j)x_iy_j+\sum_{j=1}^{n}\Phi(p_0,p_j)\left(1-\sum_{i=1}^{n}x_i\right)y_j\\
&\quad+\sum_{i=1}^{n}\Phi(p_i,p_0)x_i\left(1-\sum_{j=1}^{n}y_j\right)+\Phi(p_0,p_0)\left(1-\sum_{i=1}^{n}x_i\right)\left(1-\sum_{j=1}^{n}y_j\right)\\
&=\sum_{i,j=1}^{n}a_{ij}x_iy_j+\sum_{i=1}^{n}a_ix_i+\sum_{j=1}^{n}b_jy_j+\Phi(p_0,p_0).
\end{aligned}$$

由此可见, 双仿射函数 Φ 可以表达成

$$\Phi(o+x,o+y)=f(x,y)+l_1(x)+l_2(y)+c, \tag{5.1.5}$$

其中 f 是 V 上的双线性函数, l_1 和 l_2 是 V 的线性函数, $c=\Phi(o,o)$ 是常量. 于是

$$Q(o+x)=\Phi(o+x,o+x)=f(x,x)+[l_1(x)+l_2(x)]+c$$

是二次函数. 容易看出, Φ 是对称的当且仅当 f 是对称的且 $l_1=l_2$.

三　二次函数的中心 (点)　把圆、椭圆和双曲线的中心概念延伸到一般的二次函数, 得到下面的定义.

定义 5.4　点 $o\in\mathbb{A}$ 称为二次函数 Q 的**中心 (点)**如果

$$Q(o+x)=Q(o-x),\qquad \forall x\in V. \tag{5.1.6}$$

用 $C(Q)$ 记 Q 的全体中心 (点) 形成的集合. 如果它不是空集, 则称 Q 是**中心的**.

非中心的二次函数的最简单的例子就是抛物线对应的函数, 例如 $Q(o+x)=x_1^2-x_2$.

从式 (5.1.3) 和式 (5.1.4) 立见 o 是 Q 的中心当且仅当 $d_oQ=0$. 于是全体中心形成的集合 $C(Q)$ 是线性方程组

$$\frac{\partial Q}{\partial x_1}=\cdots=\frac{\partial Q}{\partial x_n}=0 \tag{5.1.7}$$

的解空间. 所以, $C(Q)$ 是平面或空集. 容易看出, 这个方程组的系数矩阵是矩阵 (a_{ij}) 的 2 倍. 从而, Q 有唯一的中心当且仅当 q 是非退化的.

与方程组 (5.1.7) 相伴的齐次线性方程组是

$$\frac{\partial q}{\partial x_1}=\cdots=\frac{\partial q}{\partial x_n}=0. \tag{5.1.8}$$

它的解空间正是 q 的核 $\ker q = L_q$, 或等价地说, q 的极化 f 的核 $\ker f = L_f$. (回忆一下, $L_q = \ker f = \{x \in V \mid f(x,y) = 0\ \forall y \in V\}$.) 这说明, 如果 $C(Q)$ 不是空集, 那么它的方向平面 U 是二次型 q 的核, 即 q 的极化的核, 从而维数是 $n - \operatorname{rank} q$.

假设 $o \in C(Q)$. 那么 $C(Q) = o + \ker f$. 此时, 式 (5.1.1) 中的线性函数 l 是零函数. 于是, 对任意的 $p = o + x \in C(Q)$ 和 $y \in V$, 有

$$\begin{aligned} Q(p+y) &= Q(o+x+y) = q(x+y) + l(x+y) + c \\ &= q(x) + q(y) + c = Q(p) + q(y). \end{aligned} \tag{5.1.9}$$

反过来, 如果 p 满足条件 (5.1.9), 那么 $Q(p-y) = Q(p) + q(-y) = Q(p) + q(y) = Q(p+y)$, 所以 p 是 Q 的中心. 这表明, 条件 (5.1.9) 可以作为中心点的另一个定义.

上面的讨论可以总结如下.

定理 5.5 二次函数 Q 的中心集 $C(Q)$ 是线性方程组 (5.1.7) 的解空间. 如果 $C(Q)$ 不是空集, 那么它是 $\mathbb{A}$ 的平面, 方向平面是 q 的核, 维数等于 $n - r$, 其中 n 是 $\mathbb{A}$ 的维数, r 是 Q 的秩.

另外, Q 只有一个中心当且仅当 $r = n$, 即 q 是非退化的; Q 没有中心意味着 $r < n$.

四 二次函数的典范式 仿射空间 $\mathbb{A}$ 上的两个二次函数 Q 和 Q' 称为**仿射等价的**如果存在仿射变换 $g \in \operatorname{Aff}(\mathbb{A})$ 使得 $Q' = Q \cdot g$. 我们希望二次函数 Q 有尽可能简单的表达式. 一般而言, 这需要借助坐标.

定理 5.6 设 Q 是 $\mathbb{A}$ 上的二次函数, 秩为 r.

(1) 如果 Q 是中心的, 那么, 可适当选取以中心点 o 为原点的坐标系 $\{o; e_1, \cdots, e_n\}$ 使得 Q 具有典范式

$$Q(o+x) = a_1x_1^2 + \cdots + a_rx_r^2 + c, \tag{5.1.10}$$

其中 $a_1, \cdots, a_r$ 是非零纯量, $x_1, \cdots, x_n$ 是 x 的坐标. 此时, q 的核是方程组 $x_1 = \cdots = x_r = 0$ 的解空间 (q 是 Q 确定的二次型). 而且, 对任意的中心点 $o' \in C(Q)$, 有 $Q(o') = c$.

(2) 如果 $C(Q)$ 是空集, 那么存在 $\mathbb{A}$ 的坐标系 $\{o; v_1, \cdots, v_n\}$ 使得 Q 具有典范式

$$Q(o+x) = a_1x_1^2 + \cdots + a_rx_r^2 + x_{r+1}, \tag{5.1.11}$$

其中 $x_1, \cdots, x_n$ 是 x 的坐标.

(3) 有中心的二次函数与无中心的二次函数不是仿射等价的.

证明 在空间 V 中取二次型 q 的典范基 $e_1, \cdots, e_n$.

(1) 任取 $o \in C(Q)$ 作为原点, 那么在坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 下, Q 有下述形式

$$Q(o+x) = a_1x_1^2 + \cdots + a_rx_r^2 + l(x) + c.$$

其中 $a_1,\ \cdots,\ a_r$ 是非零纯量, $x_1,\ \cdots,\ x_n$ 是 x 的坐标. 由于 o 是 Q 的中心, 所以线性函数 l 为 0, 从而 Q 具有形式 (5.1.10).

显然 $Q(o) = c$. 如果 o' 是中心, 那么 $v = \overline{oo'}$ 在 q 的核中, 所以 $q(v) = 0$. 这意味着 $Q(o') = Q(o+v) = q(v) + c = c$.

因为 $q(x) = a_1x_1^2 + \cdots + a_rx_r^2$, 所以, 它的核是方程组 $x_1 = \cdots = x_r = 0$ 确定的 $n-r$ 维子空间.

(2) 任取 $p_0 \in \mathbb{A}$ 为原点. 那么在坐标系 $\{p_0;\ e_1,\ \cdots,\ e_n\}$ 下, Q 有下述形式

$$Q(p_0+x) = a_1\xi_1^2 + \cdots + a_r\xi_r^2 + b_1\xi_1 + \cdots + b_n\xi_n + Q(p_0),$$

其中诸纯量 a_i 均不等于零, $\xi_1,\ \cdots,\ \xi_n$ 是 x 的坐标, 诸 b_i 是纯量.

命 $u = -\sum_{i=1}^{r} \dfrac{b_i}{2a_i}e_i$, $p_1 = p_0 + u$. 那么, 在坐标系 $\{p_1;\ e_1,\ \cdots,\ e_n\}$ 下, Q 的表达式为

$$\begin{aligned}Q(p_1+x) = Q(p_0+u+x) &= a_1\left(\xi_1 - \frac{b_1}{2a_1}\right)^2 + \cdots + a_r\left(\xi_r - \frac{b_r}{2a_r}\right)^2 \\ &\quad + b_1\left(\xi_1 - \frac{b_1}{2a_1}\right) + \cdots + \left(\xi_r - \frac{b_r}{2a_r}\right) + b_{r+1}\xi_{r+1} + \cdots + b_n\xi_n + b_0 \\ &= a_1\xi_1^2 + \cdots + a_r\xi_r^2 + b_{r+1}\xi_1 + \cdots + b_n\xi_n + Q(p_1).\end{aligned}$$

由于 Q 没有中心, 所以 $b_{r+1}, \cdots, b_n$ 不全为零, 否则 p_0 是中心.

假设 $b_{r+k} \neq 0$. 命

$$\begin{aligned}&o = p_1 - \frac{Q(p_1)}{b_{r+k}}e_{r+k}; \quad v_i = e_i, \quad i = 1, \cdots, r; \quad v_{r+1} = \frac{1}{b_{r+k}}e_{r+k}, \\ &v_{r+2} = e_{r+1} - \frac{b_{r+1}}{b_{r+k}}e_{r+k}, \quad \cdots, \quad v_{r+k} = e_{r+k-1} - \frac{b_{r+k-1}}{b_{r+k}}e_{r+k}, \\ &v_{r+k+i} = e_{r+k+i} - \frac{b_{r+k+i}}{b_{r+k}}e_{r+k}, \quad i \geqslant 1.\end{aligned}$$

那么, 在坐标系 $\{o;\ v_1,\ \cdots,\ v_n\}$ 下, Q 的表达式为

$$\begin{aligned}Q(o+x) &= Q\left(p_1 - \frac{Q(p_1)}{b_{r+k}}e_{r+k} + x\right) \\ &= a_1x_1^2 + \cdots + a_rx_r^2 + (-Q(p_1) + x_{r+1}) + Q(p_1) \\ &= a_1x_1^2 + \cdots + a_rx_r^2 + x_{r+1}.\end{aligned}$$

其中 $x_1, \cdots, x_n$ 是 x 在坐标系 $\{o;\ v_1,\ \cdots,\ v_n\}$ 下的坐标.

(3) 如果 o 是 Q 的中心, g 是 $\mathbb{A}$ 上的仿射变换, $Q' = Q \cdot g$, 那么

$$\begin{aligned}Q'(g^{-1}(o)+x) &= Q(g(g^{-1}(o)+x)) = Q(o+Dg(x)) \\ &= Q(o-Dg(x)) = Q(g(g^{-1}(o)-x)) = Q'(g^{-1}(o)-x).\end{aligned}$$

这说明, $g^{-1}(o)$ 是 Q' 的中心. 所以, 有中心的二次函数与无中心的二次函数不是仿射等价的. □

注 在上面定理的证明中, 清楚地写出了有关的坐标系. 如果仅是为了说明存在, 不必清楚写出坐标系, 只要实施坐标变换即可, 因为根据 4.1 节第四部分的讨论, 从坐标变换可以求出基变换, 反之亦然. 在讨论欧几里得空间上的二次函数的典范式时将只写出坐标变换, 不再具体写出基的变化.

推论 5.7 在实数域 $\mathbb{R}$ 上, 每个二次函数 Q 可以通过选取空间 $\mathbb{A}$ 的适当的基 $\{o;\ e_1,\ \cdots,\ e_n\}$ 使得它有标准形:

$$Q(o+x) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 + c; \tag{5.1.12}$$

$$Q(o+x) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2 + x_{r+1}. \tag{5.1.13}$$

而且, Q 的标准形是唯一的, 即 s, r, c 仅依赖 Q, 不依赖基的选取.

由于实数域上二次函数的标准形的唯一性, 所以实仿射空间 $\mathbb{A}$ 上的两个二次函数等价, 当且仅当, 它们有相同的秩和符号差, 并且都没有有中心, 或都有中心且在中心点上的值是一样的.

五 欧几里得空间上的二次函数 对欧几里得空间 $(\mathbb{E}, V)$ 上的二次函数, 需要把度量考虑进去, 也就是说, 要考虑它们在直角坐标系下的典范式, 在保距作用下的等价性.

定义 5.8 称欧几里得空间 $\mathbb{E}$ 上的 Q_1 和 Q_2 是**保距等价的**如果存在 $g \in \operatorname{Iso}(\mathbb{E})$ 使得 $Q_2 = Q_1 \cdot g$, 即 $Q_2(p) = Q_1(g(p))$.

定理 5.9 设 Q 是 n 维欧几里得空间 $\mathbb{E}$ 上的二次函数. 那么存在直角坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 使得 Q 的表达式是下列两种形式之一:

$$Q(o+x) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + c, \quad o \in C(Q); \tag{5.1.14}$$

$$Q(o+x) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + \mu x_{r+1}, \quad \mu > 0. \tag{5.1.15}$$

其中诸 λ_i 是非零实数. 上述形式, 在不计变量 x_i 的编号顺序的意义下, 是唯一的.

证明 先考虑 Q 有中心的情况. 在欧几里得向量空间 V 中有标准正交基 $e_1, \cdots, e_n$ 使得 Q 确定的二次型 q 有典范式 (参见定理 3.46)

$$q(x) = \sum_{i=1}^{r} \lambda_i x_i^2.$$

取 $o \in C(Q)$. 由于 Q 在点 o 处的微分 d_oQ 为零, 所以 Q 在直角坐标系 $\{o; e_1, \cdots, e_n\}$ 下的表达式是典范的:

$$Q(o+x) = \lambda_1 x_1^2 + \cdots + \lambda_r x_r^2 + c.$$

这些 λ_i 是 q 在标准正交基下的矩阵的非零特征值, 所以由 q 唯一确定. 而 $c = Q(o)$ 与中心点 o 的选取无关. 由此可见, 对有中心的二次函数, 上述典范式是唯一的.

下面假设 Q 是非中心的. 取 $p_0 \in \mathbb{E}$, 那么有直角坐标系 $\{p_0;\ v_1,\ \cdots,\ v_n\}$ 使得 Q 的表达式为

$$Q(p_0+y) = \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2 + b_1 y_1 + \cdots + b_n y_n + Q(p_0),$$

其中诸 b_i 是常数.

做平移 (它显然只改变原点, 不改变坐标系的直角性和坐标轴的方向):

$$\begin{aligned} z_i &= y_i + b_i/2\lambda_i, \quad i = 1, 2, \cdots, r, \\ z_i &= y_i, \quad i = r+1, \cdots, n, \end{aligned}$$

得

$$Q(p_1+z) = \lambda_1 z_1^2 + \cdots + \lambda_r z_r^2 + b_{r+1} z_{r+1} + \cdots + b_n z_n + Q(p_1).$$

由于 Q 是非中心的, $b_{r+1}, \cdots, b_n$ 不能都等于零. 引入线性函数 $b_{r+1}z_{r+1} + \cdots + b_n z_n$ 的 "模":

$$\mu = \sqrt{b_{r+1}^2 + \cdots + b_n^2} > 0.$$

考虑坐标变换

$$\begin{aligned} x_i &= z_i, \quad i = 1, 2, \cdots, r, \\ x_{r+1} &= \frac{1}{\mu}(b_{r+1} z_{r+1} + \cdots + b_n z_n + Q(p_1)), \\ x_i &= \alpha_{r+1,i} z_{r+1} + \cdots + \alpha_{n,i} z_n, \quad i = r+2, \cdots, n. \end{aligned}$$

它可以由 $(n-r) \times (n-r)$ 矩阵

$$A = \begin{pmatrix} b_{r+1}/\mu & \alpha_{r+1,r+2} & \cdots & \alpha_{r+1,n} \\ b_{r+2}/\mu & \alpha_{r+2,r+2} & \cdots & \alpha_{r+2,n} \\ \vdots & \vdots & & \vdots \\ b_n/\mu & \alpha_{n,r+2} & \cdots & \alpha_{n,n} \end{pmatrix}$$

来实现.

由于变换要把直角坐标系变到直角坐标系, 所以矩阵 A 需要选择成正交矩阵. 矩阵 A 的第一列是欧几里得向量空间 $\mathbb{R}^{n-r}$ 中的单位向量, 它可以扩成为 $\mathbb{R}^{n-r}$ 的

标准正交基. 把这个基中其他的向量作为 A 的其他列, 那么 A 是正交矩阵. 做了这个变换后, 得到一个直角坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$, 且有

$$Q(o+x)=\lambda_1 x_1^2+\cdots+\lambda_r x_r^2+\mu x_{r+1},\quad x=\sum_{i=1}^{n}x_i e_i.$$

这正是要求的典范式.

和前面一样, 诸 λ_i 是 q 在标准正交基下的矩阵的非零特征值, 所以由 q 唯一确定. 剩下要说明 μ 的唯一性. 假设另有直角坐标系 $\{o';\ e_1',\ \cdots,\ e_n'\}$ 使得 Q 有典范式

$$Q(o'+x)=\lambda_1 {x'}_1^2+\cdots+\lambda_r {x'}_r^2+\mu' x'_{r+1},\quad \mu'>0.$$

二次型 q 的极化 f 对应了一个线性算子 $\mathcal{F}:V\to V$ 使得 (参见定理 3.38)

$$f(x,y)=(\mathcal{F}x|y).$$

双线性型 f 在标准正交基 $e_i)$ 和 (e_i') 下的矩阵是对角矩阵

$$F=\operatorname{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0).$$

它也是 $\mathcal{F}$ 在那两个基下的矩阵, 因为 ${}^tF=F$. 这意味着

$$\begin{aligned}\operatorname{Im}\mathcal{F}&=\langle e_1,\cdots,e_r\rangle=\langle e_1',\cdots,e_r'\rangle,\\ \ker\mathcal{F}&=\langle e_{r+1},\cdots,e_n\rangle=\langle e_{r+1}',\cdots,e_n'\rangle.\end{aligned}$$

由此可见, 从 (e_i) 到 (e_i') 的转换矩阵形如

$$C=\begin{pmatrix}C_1&0\\0&C_2\end{pmatrix},$$

其中 C_1 是 r 阶正交矩阵, C_{n-r} 是 $n-r$ 阶正交矩阵.

设点 $p\in\mathbb{E}$ 在直角坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 和 $\{o;\ e_1,\ \cdots,\ e_n\}$ 下的坐标分别是 $X=[x_1,\cdots,x_n]$ 和 $X'=[x_1',\cdots,x_n']$, 点 o' 在坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 下的坐标是 $B=[b_1,\cdots,b_n]$. 根据式 (4.1.3), 有

$$X=CX'+B=\begin{pmatrix}C_1&0\\0&C_2\end{pmatrix}X'+B.$$

由此可知,

$$\begin{aligned}\sum_{i}^{r}\lambda_i x_i^2&=\sum_{i}^{r}\lambda_i {x'}_i^2+\nu,\\ \mu x_{r+1}&=\mu' x'_{r+1}-\nu,\quad \nu\in\mathbb{R}.\end{aligned}$$

从而

$$x_{r+1} = \frac{\mu'}{\mu}x'_{r+1} - \frac{\nu}{\mu}.$$

这说明, C_2 的第一列是 $[\mu'/\mu, 0, \cdots, 0]$. 由于 C_2 是正交矩阵, 它的每一列的所有值的平方和等于 1, 所以 $(\mu'/\mu)^2 = 1$. 这给出所需要的 $\mu' = \mu$, 因为 μ 和 μ' 都是正的. □

习　题　5.1

1. 设 Q 是仿射空间 $(\mathbb{A}, V)$ 上的二次函数. 证明: 对 $x, y \in V$ 和 $p_0 \in \mathbb{A}$ 有

$$Q(o + x + y) = q(y) + 2f(x, y) + l(y) + Q(o + x).$$

(本题和下面三题中记号的含义参见定义 5.1.)

2. 证明:

$$f(x, y) = \frac{1}{2}\sum_{i=1}^{n} y_i \frac{\partial q}{\partial x_i}(x_1, \cdots, x_n).$$

3. 设 f 和 g 是 V 上的线性函数. 证明: 如果 $\dim V \geqslant 3$, 那么二次型

$$q(x) = f(x)g(x)$$

是退化的.

4. 设 $p = o + v\ (v \in V)$ 是二次函数 Q 的中心点, 那么 V 的线性函数 $y \to 2f(v, y) + l(y)$ 是零函数.

5. 实仿射空间 $\mathbb{A}$ 上的二次函数 Q 在某个坐标系下由下面的二次多项式给出, 求出其中心:

(1) $2\sum\limits_{1 \leqslant i < j \leqslant n} x_i x_j + 2\sum\limits_{i=1}^{n} x_i + 1$;

(2) $x_1^2 + 2\sum\limits_{1 \leqslant i < j \leqslant n} x_i x_j + 2\sum\limits_{i=1}^{n} x_i + 1$;

(3) $\sum\limits_{i=1}^{n-1} x_i x_{i+1} + x_1 + x_n + 1$;

(4) $\sum\limits_{i=1}^{n} x_i^2 + 2\sum\limits_{1 \leqslant i < j \leqslant n} x_i x_j + x_1$.

6. 找出 n 维仿射空间上二次函数的等价类的个数:

(1) 基域是 $\mathbb{C}$;

(2) 基域是 $\mathbb{R}$.

5.2　仿射空间和欧几里得空间中的二次曲面

一　仿射空间 $\mathbb{A}$ 上的二次函数 Q 的零点集

$$S_Q = \{p \in \mathbb{A} \mid Q(p) = 0\}$$

称为**二次曲面**(或**二次超曲面**). 如果 $\mathbb{A}$ 的维数是 2, 二次曲面也称为**二次曲线**或**圆锥曲线**. 进一步讨论前需要排除几个无趣的极端情形:

(1) $\dim \mathbb{A} = 1$, 此时 S_Q 至多是两个点;

(2) S_Q 是空集, 例如基域 K 是实数域时, 二次函数 $x_1^2 + 2x_2^2 + 1$ 没有零点;

以下总假定 S_Q 非空, $\dim \mathbb{A} = n \geqslant 2$, 且 $\operatorname{rank} S_Q := r = \operatorname{rank} Q = \operatorname{rank} q > 0$. 先证明几个简单的几何结论.

命题 5.10 如果一条直线与一个二次曲面相交至少三个点, 那么整个这条直线在曲面上.

证明 取直线上的一点 o 作为原点, 这条直线可以写成 $L = o + \langle u \rangle = \{o + tu \mid t \in K\}$, 其中 u 是 V 的向量; 二次函数 Q 可以写成 $Q(o + x) = q(x) + l(x) + c$. 直线 L 与二次曲面 S_Q 的交点由下面的方程确定:

$$Q(o + tu) = t^2 q(u) + tl(u) + c = 0. \tag{5.2.16}$$

这是关于 t 的二次方程. 如果它的系数都是 0, 那么 $L \subset S_Q$. 否则, 它至多有两个根, 这意味着交集 $L \cap S_Q$ 至多含有两个点. □

如果方程 (5.2.16) 有重根, 相应的交点称为 L 与 S_Q 的**二重交点**. 如果 L 与 S_Q 有唯一的二重交点, 可以说 L 是 S_Q 的切线. 由于目前我们只在微积分中定义了一般曲面的切线, 所以只能在 $K = \mathbb{R}$ 的情形讨论重根和切线的联系.

定义 5.11 点 o 称为二次曲面 S_Q 的一个**中心**如果这个二次曲面关于点 o 对称, 即如果 $o + x$ 是 S_Q 上的点, 那么 $o - x$ 也是 S_Q 上的点. 如果二次曲面的中心也是二次曲面上的点, 那么它称为 (二次曲面的) 一个**顶点**.

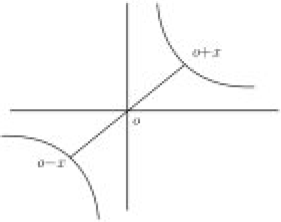


图 5.1

命题 5.12 假设 $K = \mathbb{R}$. 直线 L 与二次曲面 S_Q 有唯一的二重交点 p_0. 如果 p_0 不是 S_Q 的顶点, 那么 L 是 S_Q 在点 p_0 处的切线.

证明 设 $L = p_0 + \langle u \rangle$, $Q(p_0 + x) = q(x) + l(x)$. 直线 L 与二次曲面 S_Q 的交点由下面的方程确定:

$$Q(p_0+tu)=t^2q(u)+tl(u)=0.$$

所以二次函数 Q 在 p_0 处的微分是 $d_{p_0}Q=l$. 交点 p_0 是唯一的二重交点意味着 $l(u)=0$ 且 $q(u)\neq 0$. 从而 $d_{p_0}Q(u)=0$. 因为 p_0 不是 S_Q 的顶点, 所以 $l\neq 0$. 而 $d_{p_0}Q(u)=0$ 正是直线 L 为 S_Q 在点 p_0 处的切线的条件. □

命题 5.13 如果 o 是二次曲面 S 的顶点, $p\neq o$ 是 S 上的另一点, 那么直线 $L_{o,p}$ 全在曲面 S 上.

证明 设 $p=o+x$, $x\in V$. 那么 S 含有直线 $L_{o,p}$ 上的三点: o, $o+x$, 以及 $o-x$. 所以, S 包含整条直线 $L_{o,p}$. □

推论 5.14 如果二次曲面 S 上的每一个点都是顶点, 那么这个二次曲面是仿射空间中的平面.

证明 由命题 5.13 知曲面上任意两点连成的直线都在曲面上. 根据定理 4.13, S 是平面. □

仿射空间中的子集 X 称为以 o 为顶点的**锥**如果 X 含有 o 且对 X 中任意其他的点 p 有 $L_{o,p}\subset X$. 二次曲面称为**二次锥面**如果它有顶点且不是平面. 根据推论 5.14, 每个二次曲面只要不是平面就有异于顶点的点.

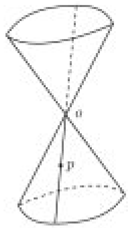


图 5.2

二 很清楚, 两个成比例的二次函数给出同样的二次曲面. 反过来未必对, 比如, 对实数域上的仿射空间, 方程 $x_1^2+x_2^2=0$ 和 $2x_1^2+5x_2^2=0$ 定义的二次曲面都是方程组 $x_1=0$, $x_2=0$ 确定的 $n-2$ 维平面. 但排除这一情形后, 反过来的结论也是对的. 为证明这一结论先把第一卷定理 7.17 推广至多元多项式.

定理 5.15 如果 K 是无限域, 那么, 不同的多项式 $f,g\in K[x_1,\cdots,x_n]$ 给出不同的函数 $K^n\to K$.

证明 只要证明非零多项式确定的函数是非零的. 对 n 作归纳法.

一元多项式的情形结论成立 (第一卷定理 7.17). 现设 $f\in K[x_1,\cdots,x_n]$ 是非零多项式, 给出零函数 $K^n\to K$. 多项式 f 可以写成

$$f=f_0+f_1x_n+\cdots+f_mx_n^m,$$

其中诸 f_i 是 $x_1,\cdots,x_{n-1}$ 的多项式. 对 $x_1,\cdots,x_{n-1}$ 任意取定 K 中的值, f 都成为 x_n 的一个多项式 ξ. 根据条件, x_n 用 K 中任何元素代替, ξ 的值都是 0. 由第一卷定理 7.17 知, ξ 的系数都是 0. 于是, 不论 $x_1,\cdots,x_{n-1}$ 怎样取 K 中的值, 诸 f_i 的值都是 0. 由归纳假设, 这些 f_i 都是 0 多项式. 从而 $f=0$. □

注 如果 K 是有限的, 多项式 f,g 在每个未知元的次数都小于 K 中元素的个数, 那么定理的结论依然成立, 证明是一样的 (参见第一卷定理 7.16).

定理 5.16 假设基域 K 是无限域. 如果 $\mathbb{A}$ 上二次函数 Q_1 和 Q_2 给出相同的二次曲面

$$S=S_{Q_1}=S_{Q_2},$$

那么: Q_1 和 Q_2 成比例, 或 S 是平面.

证明 假设 S 不是平面, 要证明 Q_1 和 Q_2 成比例. 根据推论 5.14, S 中含有异于顶点的点. 取一个这样的点 o 作为原点. 那么

$$Q_1(o+x)=q_1(x)+l_1(x),\quad Q_2(o+x)=q_2(x)+l_2(x),$$

其中 l_1,l_2 均是 V 的非零的线性函数 (否则 o 是中心, 且在 S 上, 从而是顶点, 与假设矛盾). 直线 $o+\langle x\rangle$ 与曲面 S 的交点由下面任一个方程确定:

$$t^2q_1(x)+tl_1(x)=0,\quad t^2q_2(x)+tl_2(x)=0.$$

由于这两个关于 t 的方程有相同的解, 当 $l_1(x)\neq 0$ 且 $l_2(x)\neq 0$ 时, 必有

$$\frac{q_1(x)}{l_1(x)}=\frac{q_2(x)}{l_2(x)}.$$

于是, 此时有

$$q_1(x)l_2(x)=q_2(x)l_1(x). \tag{5.2.17}$$

上式两边同乘以 $l_1(x)l_2(x)$, 得

$$q_1(x)l_2(x)l_1(x)l_2(x)=q_2(x)l_1(x)l_1(x)l_2(x).$$

这个等式对所有的 x 成立. 由于 K 是无限域, 由定理 5.15, K 上的 n 元多项式函数环与 n 元多项式环同构, 所以没有零因子. 因此可以消去上式两边的公因子. 从而式 (5.2.17) 对所有的 x 成立.

如果 l_1 和 l_2 不成比例, 那么 $l_1,l_2\in V^*$ 线性无关, 可以扩充为 V^* 的基. 取这个基在 V 中的对偶基 $e_1,\cdots,e_n$, 则有

$$l_1\left(\sum_i x_ie_i\right)=x_1,\quad l_2\left(\sum_i x_ie_i\right)=x_2.$$

这时式 (5.2.17) 成为如下等式:

$$q_1(x)x_2 = q_2(x)x_1.$$

由此可见, 有线性函数 $l \in V^*$ 使得

$$q_1(x) = l(x)x_1, \quad q_2(x) = l(x)x_2.$$

于是,

$$Q_1(x) = (l(x)+1)x_1, \quad Q_2(x) = (l(x)+1)x_2.$$

由于 $S = S_{Q_1}$, 所以 S 含有超平面 $x_1 = 0$. 由于 $S = S_{Q_2}$, 所以 Q_2 在这个超平面上恒等于 0. 可是很容易找到超平面 $x_1 = 0$ 上的点使得 $l(x)+1 \neq 0$ 和 $x_2 \neq 0$, 例如取 $a \in K$ 使得 $a \neq 0$, $al(e_2)+1 \neq 0$, 那么点 $p = o + ae_2$ 在超平面 $x_1 = 0$ 上, 且有 $Q_2(p) = a(al(e_2)+1) \neq 0$. 这个矛盾说明 l_1 和 l_2 线性相关.

于是, $l_2 = \lambda l_1$, $\lambda \in K^\times = K - \{0\}$. 由式 (5.2.17) 知 $q_2 = \lambda q_1$. 由此得到 $Q_2 = \lambda Q_1$. □

为简单起见, 除非另有说明, 以下假定二次曲面 S_Q 不是平面.

推论 5.17 二次曲面 S_Q 的中心也是二次函数 Q 的中心, 反之亦然.

证明 假设 o 是 S_Q 的中心, 那么 $S_Q = S_{\bar{Q}}$, 其中

$$\bar{Q}(o+x) = Q(o-x).$$

于是 $\bar{Q} = \lambda Q$, $\lambda \in K^\times$. 比较 Q 和 $\bar{Q}$ 的二次项即知 $\lambda = 1$. 从而 $Q = \bar{Q}$, 所以 o 是函数 Q 的中心. 显然 Q 的中心是 S_Q 的中心. □

推论 5.18 如果二次曲面 S_Q 在一个平移下不变, 那么 Q 在这个平移下也不变.

证明 设 S_Q 在平移 t_v 下不变, 即对任意的 $p \in S_Q$, $p+v \in S_Q$. 命 $\bar{Q} = Q \cdot t_v$. 那么 $S_Q = S_{\bar{Q}}$. 如同前面推论的证明中的原由, 有 $Q = \bar{Q}$. □

注 如果 K 是三元域 $\mathbb{Z}_3$, 那么定理 5.16 中的结论不成立, 反例如下: 方程 $x_1^2 + x_1x_2 + x_2 = 0$ 和 $x_2^2 + x_1x_2 + x_1 = 0$ 给出 $\mathbb{Z}_3^2$ 中相同的二次曲面 (由点 (0,0) 和 (1,1) 组成). 仔细检查定理 5.16 的证明, 并考虑定理 5.15 证明后的注, 就会发现除了 $\mathbb{Z}_3$, 定理的结论对其他有限域也是成立的 (当然, 要满足本章开始的假设 $\operatorname{char} K \neq 2$). 不过, 推论 5.17 和推论 5.18 对 $\mathbb{Z}_3$ 也成立.

我们看一下二次曲面 S_Q 能在什么平移下不变. 对 $Q(o+x) = q(x) + l(x) + c$, 命

$$\ker Q = \ker q \cap \ker l$$

(回顾: $\ker q$ 就是 q 的极化的核).

命题 5.19 函数 Q 在平移 t_v 下不变当且仅当 $v \in \ker Q$. 特别, $\ker Q$ 不依赖原点的选取.

证明 函数 Q 在平移 t_v 下不变等价于 $Q(o+v+x)=Q(o+x)$. 由引理 5.2 知, 这正是 $v\in\ker Q$ 的条件. □

由此可见, 如果 $U=\ker Q\neq 0$, 那么对任意的 $p\in S_Q$, 平面 $p+U$ 都在曲面 S_Q 上. 这样的二次曲面称为由 U 生成的一个**柱形二次曲面**(cylindrical quardric) 或**柱面**. 取 V 的基使得后 d 个基向量构成子空间 U 的一个基. 那么 Q 的坐标形式 (5.1.3) 不含后 d 个坐标. 方程 $Q=0$ 因此可以看作 $(n-d)$ 维空间中一个二次曲面 S_0 的方程. 于是, S_Q 的点的后 d 个坐标是任意的, 前 $n-d$ 个坐标是 S_0 的点的坐标. 曲面 S_0 称为 S_Q 的**底面**.

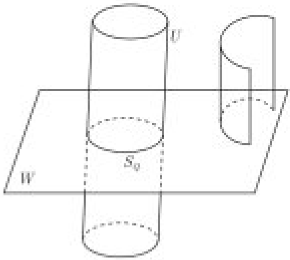


图 5.3

所以, 二次曲面的描述归结为非柱形二次曲面的描述, 即描述核 $\ker Q=0$ 的二次曲面.

命题 5.20 非柱形二次曲面至多含有一个中心.

证明 假设二次曲面 S 有两个中心 o 和 o'. 那么 $Q(o+x)=Q(o-x)$, $Q(o'+x)=Q(o'-x)$. 由此可得

$$\begin{aligned}Q(o'+x)&=Q(o+\overline{oo'}+x)=Q(o-\overline{oo'}-x)\\&=Q(o'+\overline{o'o}-\overline{oo'}-x)=Q(o'-2\overline{oo'}-x)\\&=Q(o'+2\overline{oo'}+x).\end{aligned}$$

于是 Q 在平移 $t_{2\overline{oo'}}$ 下是不变的. 这与 $\ker Q=0$ 矛盾.

另一个证明是考虑反射 $\tau: o+x\to o-x$ 和 $\tau': o'+x\to o'-x$. 因为 $\tau(S)=\tau'(S)=S$, 所以 $\tau\tau'(S)=S$. 由于

$$D(\tau\tau')=D(\tau)D(\tau')=(-\mathcal{E})^2=\mathcal{E},$$

知 $\tau\tau'$ 是平移. 于是, S 在一个非平凡的平移下不变. 这与 S 为非柱形的假设矛盾. □

三　仿射空间中二次曲面的分类　根据前面的讨论, 只要讨论非柱形二次曲面, 即 $\ker Q=\ker q\cap\ker l=0$ 的情形.

(1) 假设二次曲面有中心. 以中心 o 为原点, 则二次函数的形式是

$$Q(o+x)=q(x)+c,$$

其中 q 是非退化的二次型 (即 q 的秩等于空间 V 的维数).

如果 $c=0$, 那么二次曲面 S_Q 的方程是

$$q(x_1,\cdots,x_n)=0. \tag{5.2.18}$$

这是一个锥面, 原点是顶点.

如果 $c\neq 0$, 则二次曲面 S_Q 的方程是

$$\bar{q}(x_1,\cdots,x_n)=1, \tag{5.2.19}$$

其中 $\bar{q}=-c^{-1}q$. 这是一个非锥形的中心二次曲面, 没有顶点.

(2) 假设二次曲面没有中心. 根据定理 5.6(2) 知在某个坐标系下, 二次函数的形式是 $Q(o+x)=a_1x_1^2+\cdots+a_{n-1}x_{n-1}^2+x_n$, 其中诸 a_i 均不等于 0. 于是 S_Q 的方程可以写成

$$\bar{q}(x_1,\cdots,x_{n-1})=x_n, \tag{5.2.20}$$

其中 $\bar{q}=-q|_{\ker l}$ 是有 $n-1$ 个变量的非退化二次函数.

对实数域上的非柱形二次曲面, 其定义方程可以化成很简单的形式.

命题 5.21　设 $\mathbb{A}$ 是 n 维实仿射空间, S 是其中的非柱形二次曲面. 那么, 在适当的坐标系下, 等价的说法是, 通过适当仿射变换, S 的方程可以化成下列标准形之一:

$\mathrm{I}_{s,n}:x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=1, 0<s\leqslant n;$

$\mathrm{I}'_{s,n}:x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_n^2=0, \dfrac{n}{2}\leqslant s<n;$

$\mathrm{II}_{s,n-1}:x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_{n-1}^2=x_n, \dfrac{n-1}{2}\leqslant s\leqslant n-1.$

前面两型二次曲面有中心, 后面一型二次曲面没有中心. 这三型二次曲面互不仿射等价, 即任何一型不能通过仿射变换转成另一型.

证明　这几乎是显然的: 把方程 (5.2.19), (5.2.18) 和 (5.2.20) 中的二次型化成标准形, 并注意 $S_Q=S_{\lambda Q}$, $\lambda\in K^\times$, 且 S 不是平面. 对非中心的情况, 需要时, 可以把坐标系中最后一个基向量 e_n 换成 $-e_n$, 以保证 $(n-1)/2\leqslant s\leqslant n-1$.

定理 5.6 和推论 5.17 说有中心的二次曲面与无中心的二次曲面不是仿射等价的. $\mathrm{I}'_{s,n}$ 型二次曲面包含中心, $\mathrm{I}_{s,n}$ 型二次曲面的中心不在曲面上. 由于仿射变换把

曲面的中心变到中心 (参见定理 5.6(3) 的证明), 所以 $\mathrm{I}_{s,n}$ 型二次曲面和 $\mathrm{I}'_{s,n}$ 型二次曲面不是仿射等价的. □

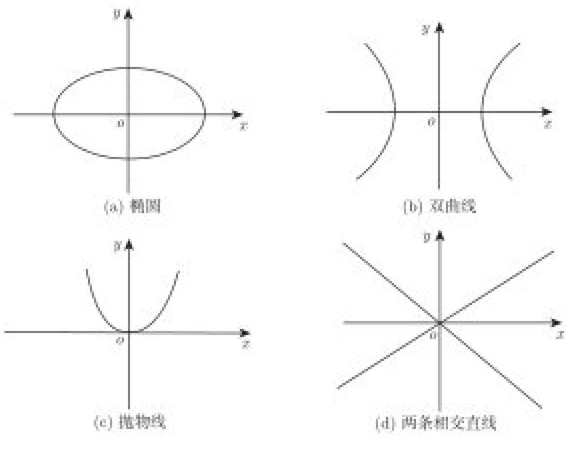
定义 5.22 称 $\mathrm{I}_{n,n}$ 型二次曲面为**椭球面**, 称 $\mathrm{I}_{s,n}$ $(0<s<n)$ 型二次曲面为**双曲面**, $\mathrm{I}'_{s,n}$ 型二次曲面是**二次锥面**, 称 $\mathrm{II}_{n-1,n-1}$ 型二次曲面为**椭球抛物面**, 称 $\mathrm{II}_{s,n}$ $((n-1)/2 \leqslant s < n-1)$ 型二次曲面为**双曲抛物面**.

椭球面、双曲面、椭球抛物面和双曲抛物面都称为**非退化的二次曲面**, 锥面和柱形的二次曲面则称为**退化的二次曲面**.

特别, 对 $n=2$ 或 3, 我们得到了熟知的非柱形的实二次曲线与曲面的分类, 列表和示图如下.

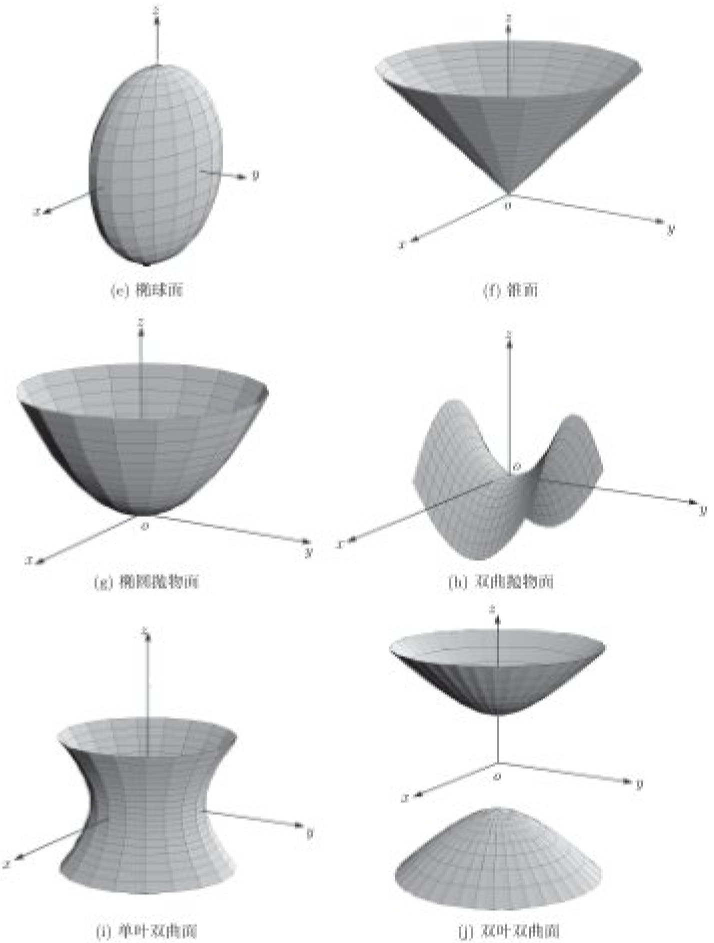
表 5.1

空间维数 n	曲面类型	正惯性指数 s	曲线或曲面名称
2	I	2	椭圆
2	I	1	双曲线
2	I′	1	两条相交直线
2	II	1	抛物线
3	I	3	椭球面
3	I	2	单叶双曲面
3	I	1	双叶双曲面
3	I′	2	二次锥面
3	II	2	椭圆抛物面
3	II	1	双曲抛物面



(a) 椭圆 (b) 双曲线

(c) 抛物线 (d) 两条相交直线



(e) 椭球面　(f) 锥面

(g) 椭圆抛物面　(h) 双曲抛物面

(i) 单叶双曲面　(j) 双叶双曲面

图 5.4

在 $n=2$ 时柱形的二次曲线没有 (就是一条直线或两条直线). 在 $n=3$ 的情形, 柱形的二次曲面的底面是平面的二次曲线和直线的两点 (可以重合), 所以有下

面五种情形:

i) 椭圆柱面, ii) 双曲柱面, iii) 抛物柱面, iv) 两个相交的平面, v) 平面.

四 二次曲面的一般方程的直接研究 前面通过二次函数的典范式对二次曲面做了分类. 然而, 遇到的曲面常常由一般的方程定义, 而非典范式方程. 怎样直接从一般的二次方程得到二次曲面的信息是有必要探讨的.

按照定义, 二次曲面在一个坐标系 $\{o;\ e_1,\ \cdots,\ e_n\}$ 的方程的一般形式是

$$Q(o+x)=\sum_{i,j=1}^{n}a_{ij}x_ix_j+2\sum_{i=1}^{n}b_ix_i+c=0. \tag{5.2.21}$$

从这个方程得到两个矩阵: 二次型 q 的矩阵 $A=(a_{ij})$ 和加边矩阵

$$\bar{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \\ b_1 & \cdots & b_n & c \end{pmatrix}.$$

如果 A 是非退化的, 那么 q 是非退化的, 从而 S 是非柱形的中心二次曲面. 这时曲面方程的典范式中变量的平方项的系数是 A 的特征值, 而且 S 是锥面当且仅当 $\operatorname{rank}\bar{A}=\operatorname{rank}A=n$.

仅最后一个断言是新结论, 下面对它给予证明.

方程 (5.2.21) 通过坐标变换可以化成典范形式. 一般的坐标变换是

$$x_1=c_{11}x_1'+\cdots+c_{1n}x_n'+d_1,$$
$$\cdots\cdots$$
$$x_n=c_{n1}x_1'+\cdots+c_{nn}x_n'+d_n.$$

假设变换后坐标系成为 $\{o';\ e_1',\ \cdots,\ e_n'\}$, $o+x=o'+x'$. 在这个坐标系下方程是典范形式:

$$Q(o'+x')=\sum_{i,j=1}^{n}a_i'x_i'^2+c'=0. \tag{5.2.22}$$

它的加边矩阵是 $\bar{A}'=\operatorname{diag}(a_1',\cdots,a_n',c')$. 命

$$\bar{C}=\begin{pmatrix} c_{11} & \cdots & c_{1n} & d_1 \\ \vdots & & \vdots & \vdots \\ c_{n1} & \cdots & c_{nn} & d_n \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

那么有

$$ {}^t\bar{C}\cdot\bar{A}\cdot\bar{C}=\bar{A}'. \tag{5.2.23} $$

由此可见如果 A 是非退化的, 则 S 是锥面当且仅当 $\operatorname{rank}\bar{A}=n$.

等式 (5.2.23) 可以从另一个角度阐述. 为记号的统一, 置

$$ a_{i,n+1}=a_{n+1,i}:=b_i,\quad i=1,2,\cdots,n;\quad a_{n+1,n+1}:=c. $$

那么 $\bar{A}=(a_{ij})_{1\leqslant i,j\leqslant n+1}$. 它可以看作二次型

$$ \bar{q}=\bar{Q}(o+x)=\sum_{i,j=1}^{n+1}a_{ij}x_ix_j $$

的矩阵. 从坐标系 $\{o;\ e_1,\ \cdots,\ e_n,\ e_{n+1}\}$ 向坐标系 $\{o;\ e_1',\ \cdots,\ e_n',\ e_{n+1}'=e_{n+1}\}$ 的转换带来的坐标变换是

$$ \bar{X}=\bar{C}\bar{X}', $$

其中 $\bar{X}=[X,x_{n+1}]$, $\bar{X}'=[X',x_{n+1}']$. 变换后二次型 $\bar{q}$ 的矩阵成为 $\bar{A}'$, 所以等式 (5.2.23) 成立.

上面的论证方式适用于 A 退化的情形. 再比较二次曲面方程的典范形式, 得到如下结论.

定理 5.23 方程式 (5.2.21) 给出的矩阵 A 和 $\bar{A}$ 的秩 r 和 $\bar{r}$ 在仿射变换下是不变的. 于是曲面 S 有中心当且仅当 $r=\bar{r}$ 或 $\bar{r}=r+1$. 曲面 S 无中心当且仅当 $\bar{r}=r+2$.

当 $\bar{r}=r$ 时, 曲面的中心在曲面上; 当 $\bar{r}=r+1$, 曲面的中心不在曲面上. 有中心的二次曲面是锥面当且仅当 $\bar{r}=r=n$, 是柱面当且仅当 $r<n$.

无中心的二次曲面在 $r=n-1$ 时是抛物面, 在 $r<n-1$ 时是柱面.

当 $K=\mathbb{R}$ 时, 曲面的类型由矩阵 A 的特征值确定.

五 假设 $K=\mathbb{R}$. 命题 5.11 的结论是关于非顶点的切线的一个充要条件. 顶点其实是奇点, 它没有切线, 当然也没有切平面. 回忆我们在定理 5.16 的证明后约定 S_Q 不是平面, 除非特别说明. 一般地, 称 $p\in S_Q$ 是二次曲面的**奇点**如果 $d_pQ=0$. 由此可见, 二次曲面的奇点就是在曲面上的中心点. 例如, 原点是锥面 $x_1^2+\cdots+x_s^2-x_{s+1}^2-\cdots x_n^2=0\ (0<s<n)$ 的唯一奇点. 如果 $p\in S_Q$ 不是奇点, S_Q 在 p 处的切平面由下述方程式确定:

$$ D_p[Q(p+x)]=0. $$

没有奇点的二次曲面是光滑的.

在平面解析几何中双曲线 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ 的渐近线是 $\dfrac{x}{a}\pm\dfrac{y}{b}=0$.

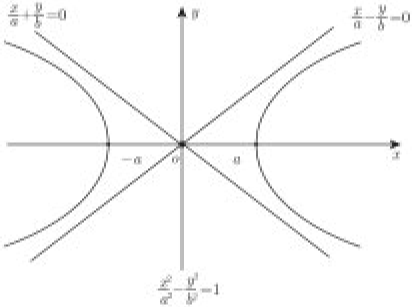


图 5.5

有了这个例子, 就不难理解下面的定义.

定义 5.24 称非零向量 $v \in V$ 是二次曲面 S_Q 的**渐近向量**或**渐近方向**如果 $q(v)=0$. 方程

$$q(x)=0$$

的解集称为二次曲面 S_Q 的**渐近锥面**.

如果 u 不是渐近方向, 那么 $q(u) \neq 0$, 从方程 (5.2.16) 知, 直线 $L=p_0+\langle u\rangle$ 与 S_Q 的交可能是两个点, 一个二重点, 或没有交点 (一对虚交点). 如果 v 是渐近方向, 那么 L 与 S_Q 可能交于一点或没有交点, 还可能整条直线在曲面上.

实 (或复) 二次曲面有一个令人惊讶的性质: 仿射对称很多. 一般的高次超曲面没有这个性质.

设 S 是实数域上的二次曲面. 用 $G(S)$ 记保持 S 不变的仿射变换全体形成的群.

定理 5.25 如果 S 没有顶点, 那么 $G(S)$ 可迁地作用在 S 上. 如果 S 有顶点, 那么 $G(S)$ 可迁地作用在 S 的顶点集在 S 中的补集.

证明 如果 $S=S_Q$ 是柱形的, $\ker Q=U$. 那么 $G(S)$ 包含 U 中所有向量确定的平移群 $T_U=\{t_u \mid u \in U\}$. 这个平移群可迁地作用在任何以 U 为方向子空间的平面 $p+U$ 上. 曲面 S 的底面 S_0 是较低维数空间中非柱形二次曲面, 而且 S 是所有平面 $p+U$, $p \in S_0$ 的并集. 于是, 问题归结为对非柱形二次曲面证明定理的结论.

如果 S 是椭球面, 那么它的方程可以写成 $q(x)=1$, 其中 q 是正定二次型. 用 q 的极化作为 V 的内积, V 成为欧几里得向量空间. 于是 $\mathbb{A}$ 成为欧几里得空间, S 是 $\mathbb{A}$ 中以 o 为中心的单位球面. 显然, $G(S)$ 包含保持 o 不动的保距变换全体形成的群 $\mathrm{Iso}(\mathbb{A})_o$ (实际上两者相等, 但这里不需要这么强的结论), 后者与正交群 $O(V)$ 同构

(参见 4.3 节中第三部分). 对 S 上的两点 $o+e_1$ 和 $o+e_1'$, 把 e_1 和 e_1' 扩充为 V 的标准正交基 (e_i) 和 (e_i'). 有唯一的正交变换 $\mathcal{F}: V \to V$ 使得 $\mathcal{F}e_i = e_i'$, $i = 1, \cdots, n$. 于是 $\mathrm{Iso}(\mathbb{A})_o$ 中以 $\mathcal{F}$ 为微分的保距变换 f 把 $o+e_1$ 映到 $o+e_1'$. (把 $\mathbb{A}$ 与 V 等同起来, S 就成为 V 中的单位球面, 这样做可以使叙述简单一些.)

双曲面的情形是类似的. 这时二次型 q 的极化确定的内积让 V 成为伪欧几里得向量空间, 符号差为 $(s, n-s)$. 映射 $\mathbb{A} \to V$, $o+x \to x$ 是仿射空间的同构. 把 S 看作 V 中的二次曲面. 对 S 中的任何点 e 和 e', 有 $q(e) = q(e') = 1$. 所以 $\langle e \rangle$ 和 $\langle e' \rangle$ 的正交补 $\langle e \rangle^{\perp}$ 和 $\langle e' \rangle^{\perp}$ 都是符号差为 $(s-1, n-s)$ 的伪欧几里得空间, 从而同构. 考虑 V 的线性同构 φ, 它把 e 映到 e', 把 $\langle e \rangle^{\perp}$ 保距映到 $\langle e' \rangle^{\perp}$. 那么 $\varphi \in O(v, q) \subset G(S)$ 且 $\varphi(e) = e'$.

假设 S 是二次锥面, 方程为 $q(x) = 0$, q 的符号差是 $(s, n-s)$. 如同双曲面的情形, 通过 q 的极化让 V 成为伪欧几里得空间, 并把 $\mathbb{A}$ 与 V 等同. 对任何非零向量 $u \in S$, 存在 $v' \in V$ 使得 $(u|v') \neq 0$. 适当调整比例可要求 $(u|v') = 1$. 命$v = v' - \dfrac{1}{2}(v'|v')u$, 则有 $(v|v) = 0$ 且 $(u|v) = 1$. 于是, 在子空间 $U = \langle u, v \rangle$ 上, 内积的矩阵是 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. 从而, 内积在 U 上是非退化的, 符号差为 (1,1). 由此可得

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

其中 $U^{\perp}$ 是符号差为 $(s-1, n-s-1)$ 的伪欧几里得空间 (或欧几里得空间). 同样, 对另一个非零向量 $u' \in S$, 可以得到分解

$$V = U' \oplus U'^{\perp}.$$

考虑线性变换 $\varphi \in GL(v)$, 它把 u 映到 u', v 映到 v', 子空间 U 保距映到 U'. 于是 $\varphi \in Q(V, q) \subset G(S)$, 且 $\varphi(u) = u'$.

剩下的情形是抛物面, 曲面的方程是式 (5.2.20). 把 $\mathbb{A}$ 与 V 等同. 任何向量 $x \in V$ 都可以表成 $x = y + te$, 其中 $y \in \ker l$, $t \in \mathbb{R}$, e 是 $\ker q$ 的一个基向量使得 $l(e) = 1$. 于是, $x \in S$ 当且仅当 $\bar{q}(y) = t$. 对任意向量 $v \in \ker l$, 考虑仿射映射

$$f_v : y + te \to y + v + (t + 2\pi(v, y) + \bar{q}(v))e,$$

其中 π 是 $\bar{q}$ 的极化. 易见, 对 $u \in \ker l$, 有 $f_u \cdot f_v = f_{u+v}$, 且 $f_0 = \mathcal{E}$. 所以 f_v 是仿射变换. 如果 $\bar{q}(y) = t$, 那么

$$\bar{q}(y + v) = t + 2\pi(v, y) + \bar{q}(v),$$

反之亦然. 这意味着 $f_v \in G(S)$. 这些变换 f_v, $v \in \ker l$, 形成一个群, 可迁地作用在 S 上. □

对抛物面 $S=S_Q$, 可以典范地关联到 V 的 1 维子空间 $\ker q$, 称为 S 的**特殊方向**(special direction). 不论怎么选择原点, 都有 $\ker q\not\subset\ker l$. 因此, 对任意非零向量 $u\in\ker q$, 方程 (5.2.16) 只有一个解. 所以, 在特殊方向上的直线与抛物面恰好交于一点.

六　欧几里得空间中的二次曲面　设 $\mathbb{E}$ 是 n 维欧几里得空间, V 是其关联的实向量空间. 如同仿射空间, $\mathbb{E}$ 上的二次函数 Q 的零点集 S_Q 称为 $\mathbb{E}$ 中的二次曲面. 需要考虑二次曲面的方程 $Q(p)$ 在直角坐标系的典范形式. 下面的结论是定理 5.9 的直接应用.

定理 5.26　设 S 是 n 维欧几里得空间 $\mathbb{E}$ 的二次曲面但不是平面. 那么, 适当选取 $\mathbb{E}$ 的直角坐标系, S 的方程能且只能化成如下标准形中的一个:

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}=1,\quad 0<s\leqslant r\leqslant n;\tag{5.2.24}$$

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}=0,\quad \frac{r}{2}\leqslant s<r\leqslant n;\tag{5.2.25}$$

$$\frac{x_1^2}{a_1^2}+\cdots+\frac{x_s^2}{a_s^2}-\frac{x_{s+1}^2}{a_{s+1}^2}-\cdots-\frac{x_r^2}{a_r^2}=x_{r+1},\quad \frac{r}{2}\leqslant s\leqslant r<n.\tag{5.2.26}$$

其中诸 a_i 均为正数. 前两种情况原点是曲面的中心, 后一种情况曲面没有中心.

前两个标准形来自定理 5.9 中的典范式

$$Q(o+x)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+c,\quad o\in C(Q).$$

这些 λ_i 是 Q 确定的二次型 q 在标准正交基下的矩阵的非零特征值.

当 $c\neq 0$ 时, 可以适当安排变元 x_i 的下标以使下面的不等式成立:

$$\lambda_1c<0,\ \cdots,\ \lambda_sc<0,\ \lambda_{s+1}c>0,\ \cdots,\ \lambda_rc>0.$$

从而标准形 (5.2.24) 中的 a_i 由下式确定

$$a_i=\sqrt{|c\lambda_i^{-1}|},\quad i=1,\cdots,r.$$

当 $c=0$ 时, 可以适当安排变元 x_i 的下标使得 $\lambda_1,\ \cdots,\ \lambda_s$ 同号, $s\geqslant r/2$; 另外那些 λ_i 与 λ_1 异号. 然后令

$$a_i=\sqrt{|\lambda_i^{-1}|},\quad i=1,\cdots,r.$$

当然, 也可以取 $a_i=a\sqrt{|\lambda_i^{-1}|}$, 其中 a 是一个固定的正数.

第三个标准形来自定理 5.9 中的典范式

$$Q(o+x)=\lambda_1x_1^2+\cdots+\lambda_rx_r^2+\mu x_{r+1},\quad \mu>0.$$

可以适当安排变元 x_i 的下标使得 $\lambda_1, \cdots, \lambda_s$ 同号, $s \geqslant r/2$; 另外, 那些 λ_i 与 λ_1 异号. 然后有

$$a_i = \sqrt{|\mu\lambda_i^{-1}|}, \quad i = 1, \cdots, r.$$

必要时可以用 $-x_{r+1}$ 代替 x_{r+1}(即做变量替换 $x_{r+1} = -x'_{r+1}$), 就得到第三个标准形.

欧几里得空间中二次曲面的名称是直观自然的: **椭球面**(式 (5.2.24), $s = n$), **双曲面**(式 (5.2.24), $0 < s < r = n$), **二次锥面**(式 (5.2.25), $0 < s < r = n$), **椭球抛物面**(式 (5.2.26), $s = r = n-1$), **双曲抛物面**(式 (5.2.26), $0 < s < r = n-1$), **柱形二次曲面或柱面**(其余情况). 这些名称和直观仿射空间中二次曲面的名称是一致的. 历史上, 先有欧几里得几何中二次曲面的名称, 然后仿射几何贴切地使用了欧几里得几何的那些名称.

欧几里得空间的二次曲面方程的标准形中的诸 a_i 称为曲面的**半轴**. 在非锥面的情形, 这些半轴都是 $\mathrm{Iso}(\mathbb{E})$ 不变的, 因为它们由 λ_i, c, μ 等确定, 而这些量都是 $\mathrm{Iso}(\mathbb{E})$ 不变的. 在锥面的情形, 半轴的比值是 $\mathrm{Iso}(\mathbb{E})$ 不变的. 椭球面的一个极端情况是诸半轴都相等, 这时椭球面就是球面, 半轴是球面的半径. 在欧几里得几何中, 由于有度量的约束, 几何体具有刚性. 这与仿射几何是有差别的. 例如, 在欧几里得几何中, 半轴不同的椭球面是不一样的; 但在仿射几何中, 任何椭球面都等价于单位球面, 从而椭球面和球面没有差别.

在欧几里得空间中, 椭球面和球面的关系是密切的: 每个椭球面都有内切球面和外切球面. 例如, 对以 $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n > 0$ 为半轴的椭球面, 它的内切球面的半径是 a_n, 外切球面的半径是 a_1. 的确, 假设椭球面 S 的中心是原点, 那么它的方程是

$$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_n^2}{a_n^2} = 1.$$

由于

$$\frac{1}{a_1^2}(x_1^2 + \cdots + x_n^2) \leqslant \frac{x_1^2}{a_1^2} + \cdots + \frac{x_n^2}{a_n^2} \leqslant \frac{1}{a_n^2}(x_1^2 + \cdots + x_n^2),$$

所以椭球面上的点 $(\xi_1, \cdots, \xi_n)$ 满足不等式

$$a_n^2 \leqslant \xi_1^2 + \cdots + \xi_n^2 \leqslant a_1^2.$$

即: 以原点为中心, 以 a_n 为半径的球面在椭球面 S 的里面, 且有两个共同点 $(0, \cdots, 0, \pm a_n)$; 以原点为中心, 以 a_1 为半径的球面在椭球面 S 的外面, 且有两个共同点 $(a_1, 0, \cdots, 0)$. 椭球面与这些球面在交点处相切 (相切的含义是在交点处切平面相同).

在双曲面情形, a_{s+1}, $\cdots$, a_n 称为**副半轴**(semi-minor axis). 它指出一个事实: 平面 $x_1 = \cdots = x_s = 0$ 与双曲面没有实交点. 自然, a_1, $\cdots$, a_s 称为**主半轴**(semi-major axis).

下面证明使抛物面的方程具有标准形 (5.2.26) 的原点是唯一的. 假设在直角坐标系 $\{o; e_1, \cdots, e_n\}$ 下抛物面 S 的方程是 (5.2.26). 那么, 抛物面 S 的特殊方向是 $\langle e_n\rangle$, 在原点 o 处的切平面 Π_n 由方程 $x_n = 0$ 确定. 于是, 在点 o 处的切平面与特殊方向正交. 这样的点 o 称为抛物面的顶点 (尽管这里的意思与定义 5.12 中顶点的含义不一致). 特殊方向上过这点的直线称为抛物面的**轴**. 注意这些概念仅针对欧几里得空间中的抛物面.

命题 5.27 对欧几里得空间中的抛物面, 顶点是唯一的.

证明 设 p 是抛物面上的点, 坐标为 ξ_1, $\cdots$, ξ_n. 对方程 (5.2.26) 微分, 可知抛物面在点 p 处的法向量 (与切平面正交的向量) 的坐标是

$$2a_1^{-2}\xi_1, \ \cdots, \ 2a_s^{-2}\xi_s, \ \cdots, \ -2a_{s+1}^{-2}\xi_{s+1}, \ \cdots, \ -2a_{n-1}^{-2}\xi_{n-1}, \ -1.$$

如果 p 是顶点, 那么法向量要与 e_n 成比例. 这意味着 $\xi_1 = \cdots = \xi_{n-1} = 0$, 即 $p = o$ 是原点. □

七 高于三维的空间中的二次曲面难有直观的形状. 二维空间和三维空间的几何直观对高维几何的想象是很有帮助的. 二维的情况简单. 现在看一看三维空间中, 对 $s = 1, 2, 3$, 各种中心二次曲面之间的差别. 在双叶双曲面 ($s = 1$) 上, 存在两个点, 它们不能通过点在曲面上的连续移动叠合 (即连接这两点任何连续曲线都不全在曲面上), 要得到这样一对点, 只需在曲面的两页各取一点. 在单叶双曲面 ($s = 2$) 上, 任意两点可以通过点在曲面上的连续移动而叠合, 但曲面上有闭曲线 (例如双曲面的腰线) 不能在曲面上通过连续变形退化成一点. 在椭球面上, 每一条闭曲线都可以在曲面上连续变形退化成一点. 这些事实, 可以用于描述高维空间的各种中心二次曲面的几何差别的出发点.

借助截面研究高维几何对象是行之有效的常用方法, 可以获得一定程度的几何直观. 用超平面 $x_i =$ 常数截取椭球面, 在常数的绝对值小于 a_i 时, 截面是 $n-1$ 维空间中的椭球面, 常数 $=\pm a_i$ 时, 截面退化成一个点 (此时超平面 $x_i = \pm a_i$ 是椭球面的两个切平面). 由此可见, n 维椭球面可以看作 $n-1$ 空间中一族连续的椭球面 (包括退化情形的顶点) 的并.

双曲面的截面千差万别, 但有一个共同点: 如果截面非空, 则都有中心. 三维空间中以原点为中心的双叶双曲面和单叶双曲面的方程分别是

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1, \quad \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1.$$

用平面 $x_1=a$ 截取双叶双曲面. 当 $|a|>a_1$ 时, 截面是椭圆; $|a|=a_1$, 截面退化成一个点; 当 $|a|<a_1$ 时, 截面为空集. 用平面 $x_2=a$ 或 $x_3=a$ 截取双叶双曲面, 截面总是双曲线.

同样的方法可以讨论其他的二次曲面和高维的情形.

习　题　5.2

1. 证明: 在实 2 维仿射空间中, 任意的二次曲线都可以在适当的坐标系下, 有下列的方程式之一给出:

(1) $x_1^2+x_2^2=1$;　(2) $x_1^2-x_2^2=1$;　(3) $x_1^2=x_2$;

(4) $x_1^2+x_2^2=-1$;　(5) $x_1^2-x_2^2=0$;　(6) $x_1^2-1=0$;

(7) $x_1^2+x_2^2=0$;　(8) $x_1^2+1=0$;　(9) $x_1^2=0$.

在实 2 维欧几里得空间, 非退化的二次曲线在直角坐标系下的标准形有哪些?

2. 在实 3 维仿射空间中, 任意的二次曲面都可以在适当的坐标系下, 有下列的方程式之一给出:

(1) $x_1^2+x_2^2+x_3^2=1$;　(2) $x_1^2+x_2^2-x_3^2=1$;　(3) $x_1^2-x_2^2-x_3^2=1$;

(4) $x_1^2+x_2^2=x_3$;　(5) $x_1^2-x_2^2=x_3$;　(6) $x_1^2+x_2^2+x_3^2=-1$;

(7) $x_1^2+x_2^2-x_3^2=0$;　(8) $x_1^2+x_2^2+x_3^2=0$;　(9) $x_1^2+x_2^2=1$;

(10) $x_1^2-x_2^2=1$;　(11) $x_1^2=x_2$;　(12) $x_1^2+x_2^2=-1$;

(13) $x_1^2-x_2^2=0$;　(14) $x_1^2-1=0$;　(15) $x_1^2+x_2^2=0$;

(16) $x_1^2+1=0$;　(17) $x_1^2=0$.

在实 3 维欧几里得空间, 非退化的二次曲面在直角坐标系下的标准形有哪些?

3. 在 3 维欧几里得空间中把二次曲面化成标准形并判定曲面的类型:

(1) $2xy+2xz+2yz+2x+2y+2z+1=0$;

(2) $2x^2+y^2-3z^2+12xy+4xz+8yz+18=0$;

(3) $5x^2+2y^2+2z^2-2xy-4yz+2xz-4y-4z+4=0$;

(4) $4x^2+4y^2-8z^2-10xy+4xz+4yz-16x-16y+10z-2=0$;

(5) $6x^2+5y^2+7z^2+4xy-4xz-8x-10y+14z-6=0$.

4. 变量 t 取什么值时, 二次曲面

$$x^2+y^2+z^2+2txy+2txz+2tyz-4t=0$$

是椭球面.

5. 两个双曲面有相同的渐近锥面的充要条件是什么?

6. 找出二次曲面与平面的相交曲线的仿射类型:

(1) $3y^2+4z^2+24x+12y-72z+360=0,\quad x-y+z=1$.

(2) $x^2+5y^2+z^2+2xy+6xz+2yz-2x+6y+2z=0,\quad 2x-y+z=0$.

7. 设 p_0 是二次曲面 S 上的点, $v\in V$ 是 S 的一个渐近方向. 证明: 直线 $x=p_0+tv$ 或者整个在曲面 S 上, 或者与曲面 S 只有一个交点.

8. 证明: 二次曲面的中心全体是平面或空集.

9. 证明: 如果 S 是方程 (5.2.20) 确定的抛物面, 那么群 $G(S)$ 可迁地作用在区域 $q(x_1,\cdots,x_{n-1})<x_n$ 上.

10. 证明: 对抛物面 S 的非特殊方向, 存在这个方向上的直线, 它与 S 不相交.

5.3 射影空间

在日常生活中一条铁路线中的两根铁轨是平行线, 放眼望去, 平行的铁轨会在很远很远的地方汇在一起. 这是一种透视的效果. 在光线下, 物体的投影, 在绘画中, 把立体的东西符合视觉效果地呈现在平面上等都出现透视. 文艺复兴时期画家和建筑师对透视的研究结果导致了一种新的几何产生: 射影几何. 刚开始是平面的射影几何. 在 19 世纪, 射影几何成为一个严格的几何体系, 它的发展对几何学的影响是巨大的. 射影簇是代数几何研究的研究对象, 射影微分几何则研究射影变换的微分不变量. 我们将考虑域 K 上的射影空间并假设 $\operatorname{char} K\neq 2$.

一 射影平面 在仿射平面上, 任意两个不同的点确定唯一的直线, 任意两条不平行的直线交于一点. 射影平面可以看作是普通平面添加上在"无穷远处的点形成的直线", 普通平面上平行的直线在无穷远处交于一点. 于是, 射影平面上的几何是一种非欧几何, 在它上面没有平行的直线. 射影平面中的基本对象是点和线, 满足如下公理:

i) 任给两个不同的点, 有且只有一条直线附随 (incident with) 这两个点;

ii) 任给两条不同的直线, 有且只有一个点附随这两条直线;

iii) 存在四个点使得任何直线至多附随其中两点.

这里用附随而非"过"和"交"等常用的术语是为了强调点与线的关系的对称性质并方便抽象的讨论. 在下面具体的讨论中, 还是用更习惯的"过"与"交". 最后一条公理是为了排除一些没意思的退化情形如只有一个点, 一堆直线, 这个点附随所有的直线; 又如只有一条直线, 一堆点, 这些点都附随这条直线, 等.

直观上看, 域 K 上的射影平面 $\mathbb{P}^2=K\mathbb{P}^2$ 的就是在仿射平面 K^2 加上无穷远直线. 最常用的模型之一是通过 3 维向量空间构造. 任取域 K 上的 3 维向量空间 V. 命 $\mathbb{P}^2=\mathbb{P}(V)$ 是 V 的所有一维子空间形成的集合, $\mathbb{P}(V)$ 的直线就是 V 中的 2 维子空间. 射影平面 $\mathbb{P}(V)$ 中的点 p(V 中的 1 维子空间) 与直线 L(V 中 2 维子空间) 如果在 V 中有包含关系, 则称 L 过 (附随) 点 p 或 p 位于 L 上 (附随 L). 如果 p 和 q 是不同的点, 那么它们在 V 中是不同的直线, 从而, 它们的和是 2 维子空间 L, 也就是 $\mathbb{P}(V)$ 的直线, 它是唯一的过 p,q 的直线. 于是公理 i) 满足. 其次, $\mathbb{P}(V)$ 中两条不同的直线 L 和 M 是 V 中两个不同的 2 维子空间, 所以和 $L+M$ 是整个

空间 V. 由推论 1.39 得

$$\dim(L\cap M)=\dim L+\dim M-\dim(L+M)=2+2-3=1.$$

这意味着 $L\cap M$ 是 V 的 1 维子空间, 它是 L 与 M 唯一的交点. 于是公理 ii) 满足. 设 u,v,w 是 V 的基, 那么 $\mathbb{P}(V)$ 中的四点 $\langle u\rangle$, $\langle v\rangle$, $\langle u+w\rangle$, $\langle v+w\rangle$ 满足公理 iii).

取 V 的一个基, 然后通过向量的坐标把 V 与 K^3 等同. 那么 $\mathbb{P}^2$ 可以看做 $K^3-\{0\}$ 的等价类集合, 等价关系定义为 $x\sim y$ 如果存在 $\lambda\in K^\times$ 使得 $x=\lambda y$. 那么 K^2 可以与 $\mathbb{P}^2$ 中的子集 $X=\{\bar{x}\,|\,x\in K^3,\ \ x_3=1\}$ 等同起来. 子集 X 在 $\mathbb{P}^2$ 中的余集 $\{\bar{x}\,|\,x\in K^3,\ x_3=0\}$ 就是无穷远直线.

对实射影平面, 还有一个常用的模型, 利用 3 维实空间的球面实现. 本质上, 它和上面的模型是一致的. 不过, 球面模型在有些方面有更好的直观性.

设 $K=\mathbb{R}$ 是实数域. 在欧几里得空间 $\mathbb{E}=\mathbb{R}^3$ 中考虑单位球面

$$S^2:x^2+y^2+z^2=1.$$

经过原点 o 的每条直线和单位球面交于两个对径点 (antipodal points, 直径的两个端点), 包含原点的平面与单位球面的交是大圆.

可以选择单位球面的对径点对 $\{t,t'\}$ 作为射影平面 $\mathbb{RP}^2$ 的点, 单位球面上的大圆为射影平面的直线. 如果大圆 L 包含对径点 t,t', 那么射影平面的点 $\{t,t'\}$ 就属于 (即附随)L. 很清楚, 两个不同的大圆的交点是两个对径点. 任何两个不同的对径点对 $\{t,t'\}$ 和 $\{s,s'\}$ 确定两条不同的过原点的直线, 从而确定唯一的过原点的平面, 所以只有唯一的大圆包含这两个对径点对. 公理 i) 和 ii) 得到满足. 至于公理 iii), 取 $t=(1,0,0)$, $s=(0,1,0)$, $u=(0,0,1)$, $v=(s+u)/\sqrt{2}$, $w=(t+u)/\sqrt{2}$, 那么射影平面中的点 $\{t,t'\}$, $\{s,s'\}$, $\{v,v'\}$ 和 $\{w,w'\}$ 满足公理 iii).

可以把单位球面换成它的下半球面来实现射影平面. 这个下半球面 S^2_- 由所有满足条件 $S^2:x^2+y^2+z^2=1,\ z\leqslant 0$ 的点 $x,y,z)$ 组成. 它的边界是赤道, 即满足方程 $x^2+y^2=1,\ z=0$ 的点集 S^1. 单位球面上的对径点对中至少一个点属于 S^2_-, 而且只有它们在赤道 S^1 上时对径点对的两个点都属于 S^2_-. 如果把 S^2_- 非赤道上的点和赤道上的对径点对看作射影平面的点, 大圆与 S^2_- 的交集看作射影平面的直线, 特别, 赤道 S^1 本身是射影平面的直线, 那么下半球面 S^2_- 就实现了射影平面 $\mathbb{RP}^2$.

半球面 S^2_- 在南极点 $(0,0,-1)$ 的切平面 Π 是 $z=-1$. 这个半球面立在 Π 上面. 对任何点 $t\in S^2_-\backslash S^1$, 原点 o 和 t 的连线与平面 Π 有唯一的交点 t^*. 显然, 映射

$$\sigma:S^2_-\backslash S^1\to\Pi,\ \ t\to t^*$$

是双射. 除了赤道, 映射 σ 把 $\mathbb{RP}^2$ 上每一条直线, 也就是 S^2_- 上不在赤道上的大圆周的弧, 映到 Π 上的直线, 并且, σ 保持点与直线的附随性. 可以想象, 原点与赤道上的对径点对的连线交平面 Π 于无穷远处, 从而 σ 如果延拓到整个下半球面, 赤道应该映到无穷远处, 那里有一条直线 L_0. 这样一来, 射影平面可以由仿射平面 Π 添加一条无穷远直线 L_0 得到, L_0 的点和赤道 S^1 上的对径点对一一对应. 在 $S^2_-\backslash S^1$ 上不相交的两条直线一定交于赤道, 经过 σ 映成 Π 上的两条平行直线, 它们一定交于无穷远直线. 可见, 对仿射平面添加无穷远直线实际上就是添加仿射平面上平行直线的交点.

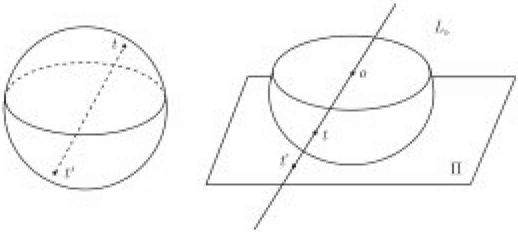


图 5.6

射影平面可以看作三维空间中过原点的直线全体形成的线束 (line pencil). 射影直线 $\mathbb{RP}^1$ 可以类似构造: 平面上过原点 o 的直线全体形成的线束. 如果点 o 在圆周 S^1 上, 那么圆周上任何其他点 p 都可以与 o 连成一条直线, 点 o 是圆周在 o 处的切线与圆周的二重交点, 所以, 这条切线 l 可以看作点 o 与自己的连线. 这样就可以得到所有过点 o 的直线. 圆周也是射影直线 $\mathbb{RP}^1$ 的一个模型. 如果以点 o 的对径点 p_0 为基点, 那么 p_0 与直线 l 的点的连线与圆周有唯一的交点. 这样可以得到圆周上除 p_0 外的所有点, 这也得到一个双射 $\tau: S^1-\{p_0\}\to l$. 同样可以想象, τ 延拓到整个圆周应该把 p_0 映到无穷远处. 所以, 直线 l 添加无穷远点就得到了射影直线, 无穷远点把直线两头连起来得到一个圆.

二　任意维的射影空间　可以用公理化方法 (也称综合法, synthetic method) 建立射影几何, 优点是直观性好, 局限是难以处理复杂和高维的问题. 现在一般是通过线性空间定义射影空间, 并使用坐标.

定义 5.28　域 K 上的 $n+1$ 维向量空间 V 的 1 维子空间全体形成的集合 $\mathbb{P}(V)$ 称为域 K 上的一个 n 维**射影空间**, 其中的元素将称为射影空间的点. 如果 $U\subset V$ 是 V 的 $m+1$ 维子空间, 那么 U 中的 1 维子空间全体形成的集合 $\mathbb{P}(U)\subset\mathbb{P}(V)$ 称为 $\mathbb{P}(V)$ 的 m 维**(射影) 子空间**(或**平面**). 射影空间 $\mathbb{P}(V)$ 的 1 维子空间也称为 (射影) 直线, $n-1$ 维子空间也称为**(射影) 超平面**. 约定 $\mathbb{P}(\{0\})=\varnothing$.

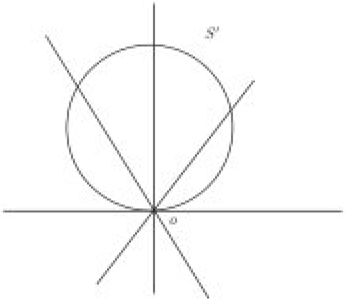


图 5.7

若干个子空间的交集只要非空就仍是子空间, 因为

$$\bigcap_{i\in I}\mathbb{P}(U_i)=\mathbb{P}\left(\bigcap_{i\in I}U_i\right).$$

如果 $V=K^{n+1}$, 那么 $\mathbb{P}(V)$ 常记作 $K\mathbb{P}^n$.

射影空间还可以通过非零向量之间的一个等价关系定义. 设 $V^\times=V-\{0\}$. 定义其上的等价关系 $\sim$ 如下:

$$x\sim y \Leftrightarrow \exists\lambda\in K^\times,\ \ x=\lambda y.$$

这个等价关系的商集 $V/\sim$ 就是射影空间 $\mathbb{P}(V)$.

非零向量 $x\in V$ 所在的等价类 $\bar{x}$ 成为射影空间 $\mathbb{P}(V)$ 的一个点. 根据定义, 有

$$\overline{\lambda x}=\bar{x},\quad \forall\lambda\in K^\times. \tag{5.3.27}$$

映射 $\pi:x\to\bar{x}$ 称为 $V^\times$ 到 $\mathbb{P}(V)$ 的**自然映射**. 应当强调, 在 $\mathbb{P}(V)$ 上没有定义线性运算, 例如, 不能规定 $\bar{x}+\bar{y}=\widetilde{x+y}$.

对任何集合 $S\in\mathbb{P}(V)$, 在包含 S 的子空间中存在一个最小的子空间 $\mathbb{P}(U)$. 如果 $S=\{\bar{x}_1,\bar{x}_2,\cdots\}$, 那么 $U=\langle x_1,x_2,\cdots\rangle$.

三　齐次坐标　射影空间中的点看成非零向量的等价类的益处之一是引入齐次坐标. 设 $e_0,e_1,\cdots,e_n$ 是 V 的基. 如果

$$x=\xi_0e_1+\xi_1e_1+\cdots+\xi_ne_n\in V^\times,$$

那么, 我们直接把 $\xi_0,\xi_1,\cdots,\xi_n$ 称为点 $\bar{x}$ 关于基 (e_i) 的**齐次坐标**, 并记作

$$\bar{x}=(\xi_0:\xi_1:\cdots:\xi_n).$$

由于 $\widetilde{\lambda x} = \tilde{x}, \forall \lambda \in K^\times$, 所以 $\lambda\xi_0, \lambda\xi_1, \cdots, \lambda\xi_n$ 也是 $\tilde{x}$ 的齐次坐标, 即

$$(\xi_0 : \xi_1 : \cdots : \xi_n) = (\lambda\xi_0 : \lambda\xi_1 : \cdots : \lambda\xi_n), \qquad \forall \lambda \in K^\times. \tag{5.3.28}$$

易见, (ξ_i) 和 (η_i) 都是 $\tilde{x}$ 的齐次坐标当且仅当存在 $\lambda \in K^\times$ 使得 $\eta_i = \lambda\xi_i$, $i = 0, 1, \cdots, n$. 可见, 点 $\tilde{x}$ 的齐次坐标不是唯一的, 但不同的齐次坐标成比例. 显然, 域 K 中任意 $n+1$ 个不全为零的元素组 (ξ_i) 必是 $\mathbb{P}(V)$ 中某个点关于基 (e_i) 的齐次坐标. 这样, 我们建立了 $\mathbb{P}(V)$ 与 $K\mathbb{P}^n$ 之间的一一对应. 这个对应把子空间映到子空间. 由于 K^{n+1} 中的子空间都是齐次线性方程组的解空间, 所以 $\mathbb{P}(V)$ 的子空间 $\mathbb{P}(U)$(的齐次坐标集合) 由一个线性方程组

$$\begin{gathered} a_{10}\xi_0 + a_{11}\xi_1 + \cdots + a_{1n}\xi_n = 0, \\ \cdots\cdots \\ a_{r0}\xi_0 + a_{r1}\xi_1 + \cdots + a_{rn}\xi_n = 0 \end{gathered} \tag{5.3.29}$$

给出.

四　仿射图　射影空间与仿射空间是分不开的, 齐次坐标也可以转化为非齐次坐标. 设 $e_0, e_1, \cdots, e_n$ 是 V 的基. 考虑向量子空间

$$V_0 = \langle e_1, e_2, \cdots, e_n \rangle.$$

在仿射空间 $(\mathbb{A} = V, V)$ 中选出超平面

$$\mathbb{A}_0 = e_0 + V_0 = \{e_0 + x \mid x \in V_0\}.$$

正如我们已经知道 (参见例 4.3), $\mathbb{A}_0$ 是与 V_0 关联的仿射空间. 如果 $x \notin V_0$, 那么直线 $\langle x \rangle$ 与 $\mathbb{A}_0$ 有唯一的交点. 实际上,

$$x \notin V_0 \;\Leftrightarrow\; x = \xi_0 e_0 + \xi_1 e_1 + \cdots + \xi_n e_n, \quad \xi_0 \neq 0.$$

由此可知, $\lambda x = \lambda\xi_0 e_0 + \lambda\xi_1 e_1 + \cdots + \lambda\xi_n e_n \in \mathbb{A}_0$ 的充要条件是 $\lambda\xi_0 = 1$.

这意味着, 取交点建立了 V_0 外的直线 $\langle x \rangle$ 和仿射空间 $\mathbb{A}_0$ 的点之间的一一对应:

$$\Phi_0 : \langle x \rangle \to \langle x \rangle \cap \mathbb{A}_0.$$

换句话说, 我们得到了一个双射

$$\Phi_0 : \mathbb{P}(V) \backslash \mathbb{P}(V_0) \to \mathbb{A}_0, \quad \langle x \rangle \to \langle x \rangle \cap \mathbb{A}_0. \tag{5.3.30}$$

定义 5.29　仿射空间 $\mathbb{A}_0$ 连同映射 Φ_0 一起称为射影空间 $\mathbb{P}(V)$ 的一个**仿射图**(affine chart). 超平面 $\mathbb{P}(V_0)$ 中的点和平面分别称为**相对于图** $\mathbb{A}_0$ **的无穷远点和平面**, 特别, $\mathbb{P}(V_0)$ 是相对于 $\mathbb{A}_0$ 的**无穷远超平面**.

由于基 (e_i) 可以任意取, 所以 $\mathbb{A}$ 中任意的不过原点的超平面 (连同合适的映射) 都可以成为仿射图. 仿射图可以用于定义 $\mathbb{P}(V)$ 中的点的非齐次坐标. 与齐次坐标不同, 一个点的非齐次坐标是唯一确定的; 但并非每个点都有非齐次坐标, 就是说, 那些相对于给定的仿射图的无穷远点没有非齐次坐标. 我们看一下齐次坐标与非齐次坐标的联系. 取定 V 的基 $e_0, e_1, \cdots, e_n$. 如果点 $x \in \mathbb{P}(V)$ 的齐次坐标是 $\bar{x} = (\xi_0 : \xi_1 : \cdots : \xi_n)$, 那么直线 (x) 与 $\mathbb{A}_0$ 的交点是

$$e_0 + \frac{\xi_1}{\xi_0} e_1 + \cdots + \frac{\xi_n}{\xi_0} e_n.$$

这意味着, 点 $\bar{x}$ 在 $\mathbb{A}_0$ 的坐标系 $\{e_0;\ e_1,\ \cdots,\ e_n\}$ 下的仿射坐标 (非齐次坐标) 是 $\frac{\xi_1}{\xi_0}, \cdots, \frac{\xi_n}{\xi_0}$. 于是, 对于仿射图 $\mathbb{A}_0$ 和取定的坐标系, 点 $(\xi_0 : \xi_1 : \cdots : \xi_n)$ 的非齐次坐标是如下比值: $\frac{\xi_1}{\xi_0}, \cdots, \frac{\xi_n}{\xi_0}$. 那些 $\xi_0 = 0$ 的点就是相对于 $\mathbb{A}_0$ 的无穷远点.

用向量 e_i 替换 e_0, 用超平面

$$V_i = \langle e_0, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n \rangle$$

替换 V_0, $\mathbb{A}_i = e_i + V_i$ 替换 $\mathbb{A}_0$, 得到仿射图 $(\mathbb{A}_i, \Phi_i)$. 这时, 点 $(\xi_0 : \xi_1 : \cdots : \xi_n)$ 在 $\mathbb{A}_i$ 的坐标系 $\{e_i;\ e_0,\ \cdots,\ e_{i-1},\ e_{i+1}, \cdots,\ e_n\}$ 下的仿射坐标是

$$\left(\frac{\xi_0}{\xi_i},\ \cdots,\ \frac{\xi_{i-1}}{\xi_i},\ \frac{\xi_{i+1}}{\xi_1},\ \cdots,\ \frac{\xi_n}{\xi_0}\right).$$

容易看出, 仿射图

$$(\mathbb{A}_i, \Phi_i), \quad i = 0, 1, \cdots, n,$$

全体组成 $\mathbb{P}(V)$ 的一张地图 (atlas), 即它们一起完全覆盖了 $\mathbb{P}(V)$. 事实上, 对射影空间 $\mathbb{P}(V)$ 中任意的点 $\bar{x} = (\xi_0 : \xi_1, \cdots, \xi_n)$, 至少有一个坐标 ξ_i 不等于零. 这意味着 $\Phi_i(\bar{x}) \in \mathbb{A}_i$. 完全可以把 $\mathbb{P}(V) \backslash \mathbb{P}(V_i)$ 与 $\mathbb{A}_i$ 等同起来, 从而有

$$\mathbb{P}(V) = \bigcup_{i=0}^{n} \mathbb{A}_i.$$

练习 5.30　在任何一张 $\mathbb{P}(V)$ 的地图中, 仿射图的数量不能少于 $n+1$.

定理 5.31　对射影空间 $\mathbb{P}(V)$ 中任意 $k+1$ 个点, 存在维数 $\leqslant k$ 的子空间, 它包含所有这些点. 进一步, 如果这些点不在任何维数 $< k$ 的子空间中, 那么有唯一的 k 维子空间包含这些点.

证明　用向量空间的语言表述, 结论是显然的: 任何 $k+1$ 个向量落在一个维数 $\leqslant k+1$ 的子空间; 如果它们不全属于任何维数 $< k+1$ 的子空间中, 那么它们张成的子空间就是唯一包含它们全体的 $k+1$ 维子空间. □

定理 5.32 设 Π_1 和 Π_2 是 n 维射影空间 $\mathbb{P}(V)$ 的子空间. 如果 $\dim\Pi_1+\dim\Pi_2\geqslant n$, 那么 $\Pi_1\cap\Pi_2\neq\varnothing$, 且有

$$\dim(\Pi_1\cap\Pi_2)\geqslant\dim\Pi_1+\dim\Pi_2-n. \tag{5.3.31}$$

例如, 射影平面上任意两条直线相交.

证明 如果 $\Pi_1=\mathbb{P}(U_1)$, $\Pi_2=\mathbb{P}(U_2)$, 那么

$$\dim U_1+\dim U_2=\dim\Pi_1+\dim\Pi_2+2\geqslant n+2>\dim V.$$

于是, $\Pi_1\cap\Pi_2=\mathbb{P}(U_1\cap U_2)\neq\varnothing$. 实际上, 有

$$\dim(U_1\cap U_2)\geqslant\dim U_1+\dim U_2-\dim V.$$

这个不等式蕴含式 (5.3.31). □

五 射影簇 (代数流形) 的概念 一个 $n+1$ 元多项式 $f(t_0,t_1\cdots,t_n)\in K[t_0,t_1,\cdots,t_n]$ 可以自然成为 K^{n+1} 上的函数. 在任意的 $n+1$ 维向量空间 V 取定一个基后, 就有同构 $V\simeq K^{n+1}$, 从而, f 也可以成为 V 上的函数. 但一般没法让 f 成为 $\mathbb{P}(V)$ 上的函数. 例如, 当 K 是无限域时, 因为对 $\bar{x}=(\xi_0:\xi_1:\cdots:\xi_n)\in\mathbb{P}(V)$, 有 $\bar{x}=(\lambda\xi_0:\lambda\xi_1:\cdots:\lambda\xi_n)$, $\lambda\in K^\times$, 从而满足条件

$$f(\xi_0,\xi_1,\cdots,\xi_n)=f(\lambda\xi_0,\lambda\xi_1,\cdots,\lambda\xi_n),\quad\forall\ \bar{x}\in\mathbb{P}(V),\lambda\in K^\times$$

的多项式 f 只有零多项式. 如果 f 是 d 次齐次多项式, 那么有

$$f(\lambda\xi_0,\lambda\xi_1,\cdots,\lambda\xi_n)=\lambda^d f(\xi_0,\xi_1,\cdots,\xi_n),\quad\forall\ \lambda\in K^\times.$$

虽然仍不能得到 $\mathbb{P}(V)$ 上的函数, 但可以谈论 f 的零点, 因为此时有

$$f(\lambda\xi_0,\lambda\xi_1,\cdots,\lambda\xi_n)=0\ \Leftrightarrow\ f(\xi_0,\xi_1,\cdots,\xi_n)=0.$$

我们对多项式的零点总是很好奇的, 下面的定义是自然的事情.

定义 5.33 设 g_r, $r\in R$, 是 $K[t_0,t_1,\cdots,t_n]$ 中的一族齐次多项式. 满足方程组

$$g_r(\alpha_0,\alpha_1,\cdots,\alpha_n)=0,\quad r\in R,$$

的点 $(\alpha_0:\alpha_1:\cdots:\alpha_n)$ 全体形成的集合 $S\subset\mathbb{P}(V)$ 称为一个**射影簇**(或**代数流形**).

射影簇是一类代数簇. 可以在 $\mathbb{P}(V)$ 上引入查理斯基 (Zariski) 拓扑, 射影簇是这个拓扑中的闭集, 所以是几何体. 它们是代数几何研究的对象, 这里无法细说. 不过, 对一个方程的情形, 可以做一些简单的探讨.

设 $S \subset \mathbb{P}(V)$ 是由齐次多项式方程

$$g(t_0, t_1, \cdots, t_n) = 0$$

定义的射影簇. 我们对交集 $S_i = S \cap \mathbb{A}_i$ 感兴趣. 简单起见, 考虑 S_0 的情况. 如果 $\bar{x} = (\alpha_0 : \alpha_1 : \cdots : \alpha_n) \in S_0$, 那么 $\alpha_0 \neq 0$. 因此, 条件 $g(\alpha_0, \alpha_1, \cdots, \alpha_n) = 0$ 等价于条件

$$g\left(1, \frac{\alpha_1}{\alpha_0}, \cdots, \frac{\alpha_n}{\alpha_0}\right) = 0.$$

这说明 S_0 中的点 $\bar{x}$ 的仿射坐标 $\frac{\alpha_1}{\alpha_0}, \cdots, \frac{\alpha_n}{\alpha_0}$ 是多项式

$$g(1, x_1, \cdots, x_n) = 0$$

的解. 反过来, 这个多项式的任何解是 S_0 中某个点的仿射坐标.

另一方面, 如果知道 $S_0 \subset \mathbb{A}_0$ 的点的仿射坐标集合是多项式方程

$$f(x_1, \cdots, x_n) = 0$$

的解集合, 其中 f 的次数是 m, 但 f 不一定是齐次的, 那么

$$g(t_0, t_1, \cdots, t_n) = t_0^m f\left(\frac{t_1}{t_0}, \cdots, \frac{t_n}{t_0}\right)$$

是 m 次齐次多项式. 实际上, f 中任何单项式

$$x_1^{k_1} \cdots x_n^{k_n}, \quad k_1 + \cdots + k_n \leqslant m,$$

给出 g 中的单项式 $t_0^{m-k_1-\cdots-k_n} t_1^{k_1} \cdots t_n^{k_n}$ 都是 m 次的. 而且

$$g(1, x_1, \cdots, x_n) = f(x_1, \cdots, x_n).$$

所以, 如果射影簇 $S \subset \mathbb{P}(V)$ 由方程 $g = 0$ 定义, 那么 $S \cap \mathbb{A}_0 = S_0$.

对其他的 S_i, 可类似讨论并得到相应的结果. 注意某些 S_i 可以是空集.

例 5.34 考虑实射影平面 $\mathbb{RP}^2$ 中由方程 $g(t_0, t_1, t_2) = t_0^2 - t_1^2 - t_2^2 = 0$ 定义的射影簇 S.

(1) 在仿射图 $\mathbb{A}_0$ 中, S_0 的方程是 $g(1, x_1, x_2) = 1 - x_1^2 - x_2^2 = 0$, 所以 S_0 是单位圆周.

(2) 在仿射图 $\mathbb{A}_1$ 中, S_1 的方程是 $g(x_0, 1, x_2) = x_0^2 - 1 - x_2^2 = 0$, 所以 S_1 是双曲线.

(3) 在仿射图 $\mathbb{A}_2$ 中, S_2 的方程是 $g(x_0, x_1, 1) = x_0^2 - x_1^2 - 1 = 0$, 所以 S_2 也是双曲线.

如果取坐标变换 $\alpha_0 = t_0 - t_1$, $\alpha_1 = t_0 + t_1$, $\alpha_2 = t_2$, 那么 S 的定义方程变为 $\alpha_0\alpha_1 = \alpha_2^2$. 新坐标相应的仿射图分别记作 $\mathbb{A}_0'$, $\mathbb{A}_1'$, $\mathbb{A}_0'$.

(4) 在仿射图 $\mathbb{A}_0'$ 中, S_0' 的方程是 $\alpha_1 = \alpha_2^2$, 所以 S_0' 是抛物线.

(5) 在仿射图 $\mathbb{A}_1'$ 中, S_1' 的方程是 $\alpha_0 = \alpha_2^2$, 所以 S_0' 也是抛物线.

(6) 在仿射图 $\mathbb{A}_2$ 中, S_2' 的方程是 $\alpha_0\alpha_1 = 1$, 所以 S_2' 是双曲线.

以上的分析表明, 仿射空间中的二次曲线 (椭) 圆、双曲线和抛物线的方程齐次化后得到的方程本质上都是一样的. 这些二次曲线只是一个射影簇在不同的仿射图中的投影.

六　射影群　每一个非退化的线性算子 $\mathcal{A}: V \to V$ 把 V 的 1 维子空间映到 1 维子空间, 从而定义了射影空间 $\mathbb{P}(V)$ 到自身的一个双射 $\tilde{\mathcal{A}}$.

定义 5.35　非退化的线性算子 $\mathcal{A} \in GL(V)$ 给出的变换

$$\tilde{\mathcal{A}}: \mathbb{P}(V) \to \mathbb{P}(V), \quad \tilde{\mathcal{A}} \cdot \tilde{x} = \widetilde{\mathcal{A}x} \tag{5.3.32}$$

称为一个**射影变换**.

不同的非退化线性算子可以给出同样的射影变换. 对 $\lambda \in K^\times$, 有

$$\widetilde{\lambda\mathcal{A}} \cdot \tilde{x} = \widetilde{\lambda \cdot \mathcal{A}x} = \widetilde{\mathcal{A}x} = \tilde{\mathcal{A}} \cdot \tilde{x}, \tag{5.3.33}$$

所以 $\widetilde{\lambda\mathcal{A}} = \tilde{\mathcal{A}}$. 这也是一个必要条件.

定理 5.36　等式 $\tilde{\mathcal{B}} = \tilde{\mathcal{A}}$ 当且仅当 $\mathcal{B} = \lambda\mathcal{A}$, 其中 $\lambda \in K^\times$.

证明　只需说明 $\tilde{\mathcal{B}} = \tilde{\mathcal{A}} \Rightarrow \mathcal{B} = \lambda\mathcal{A}$. 对 V 中的非零向量 x, 有 $\widetilde{\mathcal{B}x} = \tilde{\mathcal{B}} \cdot \tilde{x} = \tilde{\mathcal{A}} \cdot \tilde{x} = \widetilde{\mathcal{A}x}$. 所以存在某个由 x 确定的纯量 $\lambda_x \neq 0$ 使得 $\mathcal{B}x = \lambda_x\mathcal{A}x$. 如果 $y = \alpha x$, 那么

$$\lambda_y\mathcal{A}y = \mathcal{B}y = \alpha\mathcal{B}x = \alpha\lambda_x\mathcal{A}x = \lambda_x\mathcal{A}y,$$

从而有 $\lambda_y = \lambda_x$. 如果向量 x 和 y 线性无关, 那么 $\mathcal{A}x$ 和 $\mathcal{A}y$ 也线性无关. 由关系式

$$\begin{aligned}\lambda_x\mathcal{A}x + \lambda_y\mathcal{A}y &= \mathcal{B}x + \mathcal{B}y = \mathcal{B}(x+y)\\ &= \lambda_{x+y}\mathcal{A}(x+y) = \lambda_{x+y}\mathcal{A}x + \lambda_{x+y}\mathcal{A}y\end{aligned}$$

知 $\lambda_x = \lambda_{x+y} = \lambda_y$. 这意味着 $\lambda_x = \lambda$ 对所有的非零向量 $x \in V$ 都一样, 从而有 $\mathcal{B} = \lambda\mathcal{A}$. □

从定义 5.35 可以看出, $\widetilde{\mathcal{A}\mathcal{B}} = \tilde{\mathcal{A}}\tilde{\mathcal{B}}$, 且 $\tilde{\mathcal{E}}$ 是射影变换. 这表明 $\mathbb{P}(V)$ 的射影变换全体形成一个群, 称为 $\mathbb{P}(V)$ 的**一般射影群**(general projective group), 记作 $PGL(V)$. 从定理 5.36 立即可得

定理 5.37　映射 $\pi: GL(V) \to PGL(V)$, $\mathcal{A} \to \tilde{\mathcal{A}}$ 是群同态且是满射. 同态 π 的核与 $K^\times$ 同构, 从而有短正合列

$$1 \to K^\times \xrightarrow{\psi} GL(V) \xrightarrow{\pi} PGL(V) \to 1, \tag{5.3.34}$$

其中 $\psi(\lambda)=\lambda\mathcal{E}$.

此处用短正合列的语言仅是为了温习这一概念的含义: ψ 和 π 分别是单同态和满同态, 且 $\operatorname{Im}\psi=\ker\pi$. 短正合列在进一步的数学中广泛使用.

借助于向量空间 V 的基, 射影变换可以通过坐标表达. 设 $e_0,e_1,\cdots,e_n$ 是 V 的基, $A=(a_{ij})$ 是线性算子 $\mathcal{A}\in GL(V)$ 在这个基下的矩阵:

$$\mathcal{A}e_j=\sum_{i=1}^{n}a_{ij}e_i.$$

如果 $\bar{x}=(\alpha_0:\alpha_1:\cdots:\alpha_n)$, 而 $\bar{\mathcal{A}}\bar{x}=(\beta_0:\beta_1:\cdots:\beta_n)$, 那么

$$\beta_i=\lambda\sum_{j=0}^{n}a_{ij}\alpha_j,\quad i=0,1,\cdots,n, \tag{5.3.35}$$

其中 $\lambda\neq 0$ 是纯量. 原因是 $\sum\limits_{i=0}^{n}\beta_ie_i$ 与 $\mathcal{A}x=\sum\limits_{j=0}^{n}\alpha_j\mathcal{A}e_j=\sum\limits_{j=0}^{n}\sum\limits_{i=0}^{n}a_{ij}\alpha_je_i$ 成比例. 可见, 齐次坐标的变换规则与线性算子作用下向量空间的坐标变换规则是一致的.

考察射影变换 $\bar{\mathcal{A}}$ 限制在仿射图 $\mathbb{A}_0$ 上是有趣的. 首先, $\bar{\mathcal{A}}\mathbb{A}_0$ 也是 $\mathbb{P}(V)$ 的仿射图. 对于点 $\bar{x}=(\alpha_0:\alpha_1:\cdots:\alpha_n)\in\mathbb{A}_0$, 如果 $\bar{\mathcal{A}}\bar{x}\in\mathbb{A}_0$, 那么式 (5.3.35) 中的 β_0 不等于零. 此时, 点 $\bar{x}$ 的仿射坐标 $x_i=\dfrac{\alpha_i}{\alpha_0}$, $i=1,2,\cdots,n$ 和点 $\bar{\mathcal{A}}\bar{x}$ 的仿射坐标 $y_i=\dfrac{\beta_i}{\beta_0}$, $i=1,2,\cdots,n$ 之间有如下的关系:

$$y_i=\frac{a_{i1}x_1+\cdots+a_{in}x_n+a_{i0}}{a_{01}x_1+\cdots+a_{0n}x_n+a_{00}},\quad i=1,2,\cdots,n. \tag{5.3.36}$$

如果 $\beta_0=0$, 即 $\bar{\mathcal{A}}\bar{x}\in\mathbb{P}(V_0)$ 是图 $\mathbb{A}_0$ 的无穷远点, 那么公式 (5.3.36) 就没有意义了, 不过可以理解 y_i 为 ∞ 如果 $\beta_i\neq 0$, 这倒不失为理解 $\bar{\mathcal{A}}\bar{x}$ 为图 $\mathbb{A}_0$ 的无穷远点含义的一个方式.

在 1 维情形, 公式 (5.3.36) 具有如下形式:

$$y=\frac{ax+b}{cx+d},\quad ad-bc\neq 0. \tag{5.3.37}$$

($x=x_1$, $y=y_1$.) 这时无穷远点是 $(\alpha_0:\alpha_1)=(0:1)$. 如果 $c\neq 0$, 那么射影变换 $\bar{\mathcal{A}}$ 把点 $(c,-d)$ 变到无穷点, 无穷远点变到点 (c,a). 如果 $c=0$, 公式 (5.3.37) 总是有意义的. 所以 1 维射影空间的射影变换和仿射空间上的线性分式变换没有差别: 每个射影变换给出一个线性分式变换, 每个线性分式变换确定唯一的射影变换. 此时, 射影变换可以用公式 (5.3.37) 定义.

一般情形, 如果 $\bar{\mathcal{A}}\mathbb{A}_0=\mathbb{A}_0$, 即 $a_{00}\neq 0$; $a_{0j}=0$, $1\leqslant j\leqslant n$, 那么式 (5.3.36) 就成为 $\mathbb{A}_0$ 的仿射变换. 下面的结论说明 $\mathbb{A}_0$ 的仿射变换都可以这样得到, 从而仿射变换可以看作是射影变换的特殊情形.

命题 5.38 设 $\mathbb{A}_0\subset V$ 是不含零向量的超平面, 那么 $\mathbb{A}_0$ 上的任何仿射变换都可以唯一地延拓成 V 的线性变换.

证明 设 $e_0\in\mathbb{A}_0$, U 是 $\mathbb{A}_0$ 的方向平面, 那么 $\mathbb{A}_0=e_0+U$. 如果 $e_1,\cdots,e_n$ 是 U 的基, 则 $\{e_0;e_1,\cdots,e_n\}$ 是 $\mathbb{A}_0$ 的坐标系. 如果 f 是 $\mathbb{A}_0$ 的仿射变换, 那么 $f(e_0),Df(e_1),\cdots,Df(e_n)$ 构成 $\mathbb{A}_0$ 的坐标系, 从而是 V 的基. 所以 f 可以唯一地延拓成 V 的线性变换

$$\mathcal{A}:V\to V,\quad e_0\to f(e_0),\quad e_i\to Df(e_i),\quad i=1,2,\cdots,n.\qquad\square$$

通过映射 (5.3.30) 把 $\mathbb{A}_0$ 看作 $\mathbb{P}(V)$ 的一部分, 就可以说仿射群 $\mathrm{Aff}(\mathbb{A}_0)$ 是射影群 $PGL(V)$ 的子群.

七 射影几何 由射影变换群确定的几何称为**射影几何**. 它研究射影空间 $\mathbb{P}(V)$ 中的图形在射影变换下不变的性质. 这样的性质称为**射影性质**. 直线的平行性和平面的平行性, 显然, 与射影性质没有关系. 由于射影空间中没有长度与角度的概念, 所以在射影几何中没有勾股定理. 由此可见, 射影几何与欧几里得几何、仿射几何都有很大的差别. 不过, 射影几何是一种内容极其丰富又是极其必要的几何学. 例如, 与定理 4.52 相比较, 下面的定理显示射影变换群较仿射群更多彩. 稍后将讨论射影几何的一个重要概念——交比, 并建立共线四点的一个重要射影性质.

由于 V 的线性变换在 $V\backslash\{0\}$ 上的作用是可迁的, 所以 $PGL(V)$ 在 $\mathbb{P}(V)$ 上的作用是可迁的, 即可以把任意一点变到另一点. 对一些特殊的点组, $PGL(V)$ 的作用也是可迁的. 称 n 维射影空间 $\mathbb{P}(V)$ 中的点 $\bar{u}_0,\bar{u}_1,\cdots,\bar{u}_{n+1}$ 处于**一般位置**, 如果它们中的任意 $n+1$ 个点都不在一个超平面上. 换句话说, 任意 $n+1$ 个向量

$$u_0,\ u_1,\ \cdots,\ u_{i-1},\ u_{i+1},\ \cdots,\ u_{n+1}$$

线性无关.

定理 5.39 假设 $\bar{u}_0,\bar{u}_1,\cdots,\bar{u}_{n+1}$ 和 $\bar{v}_0,\bar{v}_1,\cdots,\bar{v}_{n+1}$ 是两个处于一般位置的点组, 那么存在唯一的射影变换 $\bar{\mathcal{A}}\in PGL(V)$ 使得 $\bar{\mathcal{A}}\bar{u}_i=\bar{v}_i$, $i=0,1,\cdots,n+1$.

证明 按照定义, 向量组 $u_0,\ u_1,\ \cdots,\ u_n$ 和向量组 $v_0,\ v_1,\ \cdots,\ v_n$ 都是 V 的基. 于是有,

$$u_{n+1}=\sum_{i=0}^{n}a_iu_i,\quad v_{n+1}=\sum_{i=0}^{n}b_iv_i.\tag{5.3.38}$$

由于所给的两个点组都处于一般位置, 所以这些系数 a_i, b_i 都不等于 0. 这意味着向量组 $a_0u_0,\ a_1u_1,\ \cdots,\ a_nu_n$ 和向量组 $b_0v_0,\ b_1v_1,\ \cdots,\ b_nv_n$ 都是 V 的基. 从而

有非退化的线性算子 $\mathcal{A}$ 使得

$$\mathcal{A}a_iu_i = b_iv_i, \quad i = 0, 1, \cdots, n. \tag{5.3.39}$$

由于 $\mathcal{A}u_{n+1} = v_{n+1}$, $\widetilde{a_iu_i} = \tilde{u}_i$, $\widetilde{b_iv_i} = \tilde{v}_i$, $i = 0, 1, \cdots, n$, 所以 $\tilde{\mathcal{A}}$ 满足要求.

如果另有射影变换 $\tilde{\mathcal{B}}$ 满足要求, 那么 $\tilde{\mathcal{B}}^{-1}\tilde{\mathcal{A}}\tilde{u}_i = \tilde{u}_i$, $i = 0, 1, \cdots, n+1$. 所以

$$\mathcal{B}^{-1}\mathcal{A}u_i = \alpha_iu_i, \quad i = 0, 1, \cdots, n+1.$$

由式 (5.3.38) 知 $\alpha_i = \alpha_{n+1}$, $i = 0, 1, \cdots, n$, 所以 $\mathcal{B}^{-1}\mathcal{A} = \alpha_{n+1}\mathcal{E}$, 从而 $\tilde{\mathcal{B}} = \tilde{\mathcal{A}}$. 唯一性得证. □

推论 5.40　对射影直线 $\mathbb{P}^1$ 上任意两个三点组 $\tilde{u}_0, \tilde{u}_1, \tilde{u}_2$ 和 $\tilde{v}_0, \tilde{v}_1, \tilde{v}_2$(每个组中的点互不相同), 都存在唯一的射影变换 $\mathbb{P}^1 \to \mathbb{P}^1$, 它把 $\tilde{u}_i$ 分别变到 $\tilde{v}_i$, $i = 0, 1, 2$.

这个结论表明射影变换不保持直线上三个点的简单比率; 它可以把中间的点变成边上的点. 另外, 这个结论也说明 (由变换的唯一性), 射影变换一般不能把射影直线上的任意四个点变到给定的四个点.

命题 5.41　(1) 设 $\mathbb{P}(U)$ 和 $\mathbb{P}(W)$ 是 $\mathbb{P}(V)$ 的两个 m 维平面, 那么它们是 $PGL(V)$**全等的**(congruent), 也就是说, 存在 $\mathbb{P}(V)$ 的射影变换, 它把一个平面变到另一个平面.

(2) 平面 $\mathbb{P}(U) \subset \mathbb{P}(V)$ 上的所有射影变换都可以延拓成整个空间 $\mathbb{P}(V)$ 的射影变换.

证明　(1) 设 $u_0, \cdots, u_m$ 和 $w_0, \cdots, w_m$ 分别是 U 和 W 的基. 把它们分别扩充为 V 的基 $u_0, \cdots, u_m, \cdots, u_n$ 和 $w_0, \cdots, w_m, \cdots, w_n$. 有唯一的非退化线性算子 $\mathcal{A}: V \to V$ 使得 $\mathcal{A}u_i = w_i$, $i = 0, 1, \cdots, n$. 显然有 $\mathcal{A}U = W$, 所以 $\tilde{\mathcal{A}}(\mathbb{P}(U)) = \mathbb{P}(W)$.

(2) 设 $\tilde{\mathcal{D}}$ 是 $\mathbb{P}(U)$ 上的射影变换, $u_0, \cdots, u_m$ 是 U 的基. 那么 $\mathcal{D}u_0, \cdots, \mathcal{D}u_m$ 也是 U 的基. 把 U 的这两个基都扩充为 V 的基

$$u_0, \cdots, u_m, u_{m+1}, \cdots, u_n; \quad \mathcal{D}u_0, \cdots, \mathcal{D}u_m, w_{m+1}, \cdots, w_n.$$

有 V 上的非退化线性算子 $\mathcal{A}$ 使得

$$\mathcal{A}u_i = \mathcal{D}u_i, \quad i = 0, 1, \cdots, m,$$
$$\mathcal{A}u_i = w_i, \quad i = m+1, \cdots, n.$$

那么 $\mathbb{P}(V)$ 的射影变换 $\tilde{\mathcal{A}}$ 就是 $\tilde{\mathcal{D}}$ 的一个到 $\mathbb{P}(V)$ 的延拓: 它在 $\mathbb{P}(U)$ 上的作用与 $\tilde{\mathcal{D}}$ 的作用相同. □

八　交比　虽然射影变换不保持一条直线上三个点的简单比率, 但射影变换保持一条直线上四个点的交比. 这是射影几何中一个重要的不变量.

设 $\mathbb{P}^n = \mathbb{P}(V)$ 是射影空间, $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4$ 是射影直线 $\mathbb{P}^1 = \mathbb{P}(U) \subset \mathbb{P}^n$ 上的四个点, 而且

$$\tilde{u}_1 \neq \tilde{u}_3, \quad \tilde{u}_1 \neq \tilde{u}_4, \quad \tilde{u}_2 \neq \tilde{u}_3, \quad \tilde{u}_2 \neq \tilde{u}_4.$$

这意味着四个向量组 $\{u_1, u_3\}$, $\{u_1, u_4\}$, $\{u_2, u_3\}$ 和 $\{u_2, u_4\}$ 都是 U 的基:

$$\langle u_1, u_3 \rangle = \langle u_1, u_4 \rangle = U = \langle u_2, u_3 \rangle = \langle u_2, u_4 \rangle.$$

把二维向量空间的基 u, v 到另一个基 x, y 的转换矩阵

$$x = au + bv, \quad y = cu + dv$$

的行列式记作

$$\left(\frac{x, y}{u, v}\right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}. \tag{5.3.40}$$

由加在点 $\tilde{u}_i$ 上的条件, 可以建立一个表达式

$$[\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4] = \left(\frac{u_1, u_3}{u_1, u_4}\right) : \left(\frac{u_2, u_3}{u_2, u_4}\right) = \left(\frac{u_1, u_3}{u_1, u_4}\right) \cdot \left(\frac{u_2, u_3}{u_2, u_4}\right)^{-1}. \tag{5.3.41}$$

定义 5.42 表达式 (5.3.41) 称为四点 $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4$ 的**交比**(cross ratio), 也称为**二重比**(double ratio) 或**不和谐比**(anharmonic ratio).

自然, 需要说明定义的合理性, 即 $[\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4]$ 仅与诸 $\tilde{u}_i$ 有关, 而与诸向量 u_i 的选取无关 (回忆一下, $\widetilde{\lambda u_i} = \tilde{u}_i$), 也就是说, 分别用诸 λu_i 代替 u_i, $[\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4]$ 是不变的.

实际上, 用 $v_1 = \lambda u_1$ 代替 u_1, 如果 $u_1 = u_1$, $u_3 = cu_1 + du_4$, 那么 $v_1 = v_1$, $u_3 = c\lambda^{-1}v_1 + du_4$, 从而

$$\left(\frac{u_1, u_3}{u_1, u_4}\right) = \begin{vmatrix} 1 & 0 \\ c & d \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ c\lambda^{-1} & d \end{vmatrix} = \left(\frac{v_1, u_3}{v_1, u_4}\right).$$

此时, 显然因子

$$\left(\frac{u_2, u_3}{u_2, u_4}\right)$$

也是不变的. 用 λu_2 代替 u_2 时, 可以建立同样的事实.

现在用 $v_3 = \lambda u_3$ 代替 u_3. 如果

$$\begin{aligned} &u_1 = u_1, \quad u_2 = u_2, \\ &u_3 = cu_1 + du_4, \quad u_3 = c'u_2 + d'u_4, \end{aligned}$$

那么

$$
\begin{aligned}
&u_1 = u_1, \quad u_2 = u_2, \\
&v_3 = \lambda c u_1 + \lambda d u_4, \quad v_3 = \lambda c' u_2 + \lambda d' u_4.
\end{aligned}
$$

由此可得

$$\left(\frac{u_1, u_3}{u_1, u_4}\right) = \begin{vmatrix} 1 & 0 \\ c & d \end{vmatrix} = \lambda^{-1} \begin{vmatrix} 1 & 0 \\ \lambda c & \lambda d \end{vmatrix} = \lambda^{-1} \left(\frac{u_1, v_3}{u_1, u_4}\right),$$

$$\left(\frac{u_2, u_3}{u_2, u_4}\right) = \begin{vmatrix} 1 & 0 \\ c' & d' \end{vmatrix} = \lambda^{-1} \begin{vmatrix} 1 & 0 \\ \lambda c' & \lambda d' \end{vmatrix} = \lambda^{-1} \left(\frac{u_2, v_3}{u_2, u_4}\right).$$

此时, 公式 (5.3.41) 右端的值没有变化. 用 λu_4 代替 u_4, 会得到同样的结果. 同时用诸 λu_i 分别代替 u_i 等同于依次用 λu_i 代替 u_i, $i = 1, 2, 3, 4$, 也就是说, 可以等同于连续做四次单个向量的替换, 所以交比是不变的. 这说明了射影直线上四个点的交比的定义是合理的.

定理 5.43　在射影变换下交比不变, 即对任意的 $\bar{\mathcal{A}} \in PGL(V)$ 都有

$$[\bar{\mathcal{A}}\tilde{u}_1, \bar{\mathcal{A}}\tilde{u}_2, \bar{\mathcal{A}}\tilde{u}_3, \bar{\mathcal{A}}\tilde{u}_4] = [\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4]. \tag{5.3.42}$$

证明　设 $\mathcal{A} \in GL(V)$, U 是 V 的 2 维子空间, $U' = \mathcal{A}(U)$. 由于 $\mathcal{A}$ 是非退化的, 线性无关的向量在 $\mathcal{A}$ 的作用下, 得到的向量仍是线性无关的. 如果

$$\langle u, v \rangle = U = \langle x, y \rangle,$$

那么

$$\langle \mathcal{A}u, \mathcal{A}v \rangle = U' = \langle \mathcal{A}x, \mathcal{A}y \rangle.$$

联系向量组 u, v 与向量组 x, y 的关系式

$$x = au + bv, \quad y = cu + dv,$$

在平面 U' 上对应着相同的关系式

$$\mathcal{A}x = a\mathcal{A}u + b\mathcal{A}v, \quad \mathcal{A}y = c\mathcal{A}u + d\mathcal{A}v.$$

这意味着

$$\left(\frac{x, y}{u, v}\right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \left(\frac{\mathcal{A}x, \mathcal{A}y}{\mathcal{A}u, \mathcal{A}v}\right).$$

应用这个等式到交比, 得到

$$\begin{aligned}[\tilde{\mathcal{A}}\tilde{u}_1, \tilde{\mathcal{A}}\tilde{u}_2, \tilde{\mathcal{A}}\tilde{u}_3, \tilde{\mathcal{A}}\tilde{u}_4] &= [\widetilde{\mathcal{A}u_1}, \widetilde{\mathcal{A}u_2}, \widetilde{\mathcal{A}u_3}, \widetilde{\mathcal{A}u_4}] \\ &= \left(\frac{\mathcal{A}u_1, \mathcal{A}u_3}{\mathcal{A}u_1, \mathcal{A}u_4}\right) \cdot \left(\frac{\mathcal{A}u_2, \mathcal{A}u_3}{\mathcal{A}u_2, \mathcal{A}u_4}\right)^{-1} = \left(\frac{u_1, u_3}{u_1, u_4}\right) \cdot \left(\frac{u_2, u_3}{u_4, u_4}\right)^{-1} \\ &= [\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4].\end{aligned}$$

□

九　交比的坐标表达　交比可以通过齐次坐标计算. 设 e_0, e_1 是平面 U 的基, $u_i = a_i e_0 + b_i e_1$. 那么在射影直线 $\mathbb{P}(U)$ 中, 诸 $\tilde{u}_i$ 的齐次坐标是

$$\tilde{u}_i = (a_i : b_i), \quad i = 1, 2, 3, 4. \tag{5.3.43}$$

如果 $u_1 = u_1$, $u_3 = cu_1 + du_4$, 则有

$$\left(\frac{u_1, u_3}{u_1, u_4}\right) = d. \tag{5.3.44}$$

此外, 由等式

$$a_3 e_0 + b_3 e_1 = u_3 = c(a_1 e_0 + b_1 e_1) + d(a_4 e_0 + b_4 e_1)$$

得 $a_3 = ca_1 + da_4$, $b_3 = cb_1 + db_4$. 所以

$$\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ ca_1 + da_4 & cb_1 + db_4 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ da_4 & db_4 \end{vmatrix} = d \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix},$$

再结合等式 (5.3.44), 就得到

$$\left(\frac{u_1, u_3}{u_1, u_4}\right) = \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix}^{-1}.$$

类似地, 有

$$\left(\frac{u_2, u_3}{u_2, u_4}\right) = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix}^{-1}.$$

这样一来, 按照交比的定义有

$$[\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{u}_4] = \frac{\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot \begin{vmatrix} a_2 & b_2 \\ a_4 & b_4 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_4 & b_4 \end{vmatrix} \cdot \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}}. \tag{5.3.45}$$

如果 $a_i \neq 0,\ 1 \leqslant i \leqslant 4$, 且 $x_i = b_i/a_i$, 那么, 由上式得出

$$[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = \frac{\begin{vmatrix} 1 & x_1 \\ 1 & x_3 \end{vmatrix} \cdot \begin{vmatrix} 1 & x_2 \\ 1 & x_4 \end{vmatrix}}{\begin{vmatrix} 1 & x_1 \\ 1 & x_4 \end{vmatrix} \cdot \begin{vmatrix} 1 & x_2 \\ 1 & x_3 \end{vmatrix}},$$

也就是

$$[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = \frac{(x_3 - x_1)(x_4 - x_2)}{(x_4 - x_1)(x_3 - x_2)}. \tag{5.3.46}$$

表达式 (5.3.46) 或它的等价形式

$$[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = \frac{x_1 - x_3}{x_2 - x_3} : \frac{x_1 - x_4}{x_2 - x_4}$$

可以作为交比的定义. 根据交比的定义, 表达式 (5.3.46) 的右端与仿射图的选取 (等价于无穷远点的选取) 和坐标系的选取都没关系. 注意当所有的 a_i 都不为 0 时, 不妨认为 $u_i = e_0 + x_i e_1$, 于是诸 u_i 都在直线 $L = e_0 + \langle e_1 \rangle$ 上, 见图 5.8.

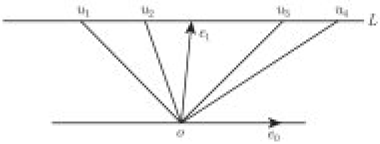


图 5.8

如果取 $e_1 = \lambda u_4$, 那么 $\bar{u}_4 = (0 : \lambda^{-1}) = (0 : 1)$ 是无穷远点. 由 (5.3.45) 得

$$[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = \frac{x_3 - x_1}{x_3 - x_2}.$$

除此之外, 如果还取 $e_0 = \mu u_3$, 那么 $\bar{u}_3 = (\mu^{-1} : 0) = (1 : 0)$ 是原点. 再选择 $\bar{u}_2$ 为单位点 (1:1). 那么 $x_3 = 0,\ x_2 = 1$, 且

$$[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = x_1. \tag{5.3.47}$$

它是点 $\bar{u}_1$ 在仿射图 $e_0 + \langle e_1 \rangle$ 的坐标系 $\{e_0;\ e_1\}$ 下的坐标. 在这个坐标系下, $\bar{u}_4$ 是无穷远点, $\bar{u}_3$ 是原点, $\bar{u}_2$ 是单位点.

上面的讨论是很有意思的. 我们不在 $\mathbb{P}^1 = \mathbb{P}(U)$ 中选取点, 而是选取一个坐标系. 设 $\bar{u}_1, \bar{u}_2, \bar{u}_3$ 是直线 $\mathbb{P}^1$ 上三个互不相同的点. 令 $e_0 = \alpha u_1,\ e_1 = \beta u_2$, 其中 α 和 β 是纯量使得 $u_3 = e_0 + e_1$. 于是 $\bar{u}_1 = (1 : 0),\ \bar{u}_2 = (0 : 1),\ \bar{u}_3 = (1 : 1)$. 现在, 如

果 $\bar{u}_4=(a:b)$ 是 $\mathbb{P}^1$ 上的任意点, 那么由公式 (5.3.45), 并在其中取

$$a_1=1,\ b_1=0;\qquad a_2=0,\ b_2=1;$$
$$a_3=1,\ b_3=1;\qquad a_4=a,\ b_4=b,$$

就得到

$$[\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4]=a/b.$$

由此可见, 点 $\bar{u}_4$ 的齐次坐标 a,b 的关系是由交比 $[\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4]$ 唯一确定的. 这样一来, 就有

定理 5.44 取定射影直线 $\mathbb{P}^1$ 上三个互不相同的点 $\bar{u}_1,\bar{u}_2,\bar{u}_3$, 那么直线上的任意第四个点 $\bar{u}_4$ 都由交比 $[\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4]$ 唯一确定.

现在, 我们已经准备好了, 可以证明一个结论, 就其本质而言, 细化了推论 5.40.

定理 5.45 在射影空间 $\mathbb{P}(V)$ 中, 两个分别共线的四点组 $\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4$ 和 $\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{v}_4$ 是 $PGL(V)$ 全等的, 当且仅当

$$[\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4]=[\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{v}_4]. \tag{5.3.48}$$

证明 条件 (5.3.48) 的必要性由定理 5.43 推出.

现证条件 (5.3.48) 的充分性. 根据命题 5.41(1), 有射影变换 $\bar{B}:\mathbb{P}(V)\to\mathbb{P}(V)$, 它把诸点 $\bar{u}_i$ 所在的直线 $\mathbb{P}(U)$ 映到诸点 $\bar{v}_i$ 所在的直线 $\mathbb{P}(W)$. 由推论 5.40, 存在直线 $\mathbb{P}(W)$ 的射影变换 $\bar{D}$, 它把 $\bar{B}\bar{u}_i$ 映到 $\bar{v}_i$, $i=1,2,3$. 利用命题 5.41(2), 把 $\bar{D}$ 延拓成整个空间 $\mathbb{P}(V)$ 的射影变换 $\bar{A}_1$, 那么变换 $\bar{A}=\bar{A}_1\bar{B}$ 把点 $\bar{u}_i$ 映到 $\bar{v}_i$, $i=1,2,3$. 同时, 点 $\bar{u}_4$ 被映到 $\mathbb{P}(W)$ 上的某个点 $\bar{x}$. 根据定理 5.43, 有

$$[\bar{u}_1,\bar{u}_2,\bar{u}_3,\bar{u}_4]=[\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{x}].$$

利用条件 (5.3.48), 得到

$$[\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{v}_4]=[\bar{v}_1,\bar{v}_2,\bar{v}_3,\bar{x}].$$

由定理 5.44 知 $\bar{x}=\bar{v}_4$. □

习 题 5.3

1. 设 $x_1,\cdots,x_n$ 是点 $\bar{x}\in\mathbb{P}(V)$ 在仿射图 $\mathbb{A}_0$ 上的仿射坐标, 求 $\bar{x}$ 在仿射图 $\mathbb{A}_1$ 上的仿射坐标.

2. 求出 q 元域上 n 维射影空间的点的个数.

3. 设 $K=\mathbb{F}_q$ 是 q 元域, 求出 $PGL(K^{n+1})$.

4. 射影平面到自身的双射如果把直线映到直线并且保持每条直线上的交比不变, 那么它是射影变换.

5. 证明: 复射影空间上的射影变换有不动点.

6. 证明: 对复射影空间上的每个射影变换, 都存在一个仿射图使得该射影变换在仿射图上的作用是仿射变换的作用.

7. 根据推论 5.40, 对射影直线

$$\mathbb{P}^1 = \{(\alpha : \beta) \mid \alpha, \beta \in K \text{ 且不全为零}\} = K \cup \{\infty\}, \quad \text{其中}\infty = (0:1)\}$$

上三个给定的互不相同的点 $\bar{u}_1, \bar{u}_2, \bar{u}_3$, 存在唯一的射影变换把 $\bar{u}_1, \bar{u}_2, \bar{u}_3$ 分别映到 ∞, 0, 1 (齐次坐标分别是 (0:1), (1:0), (1:1)). 证明:

(1) 如果 $\bar{u}_4$ 是 $\mathbb{P}^1$ 上的点, 那么它在这个射影变换下的像 $\lambda \in K \cup \{\infty\}$ 就是四点 $\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4$ 的交比;

(2) 对任意的 $\lambda \in K \cup \{\infty\}$, 存在唯一的点 $\bar{u}_4$ 使得 $[\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4] = \lambda$;

(3) 四个点的交比值在 $K - \{0, 1\}$ 中当且仅当四个点是互不相同的.

第一个结论可以作为交比的另一个定义, 并以此为基础建立交比的理论.

8. 假设 $\bar{x}, \bar{y}, \bar{z}, \bar{w}$ 是射影直线上四个不同的点. 证明

$$[\bar{x}, \bar{y}, \bar{z}, \bar{w}] = [\bar{y}, \bar{x}, \bar{z}, \bar{w}]^{-1} = [\bar{x}, \bar{y}, \bar{w}, \bar{z}]^{-1};$$
$$[\bar{x}, \bar{y}, \bar{z}, \bar{w}] + [\bar{x}, \bar{z}, \bar{y}, \bar{w}] = 1.$$

5.4　射影空间的二次曲面

射影空间中的二次曲面理论比起仿射空间中的二次曲面理论要简单得多. 在 19 世纪的数学家看来, 这是射影几何完美性的一个体现. 他们对射影几何那么着迷, 甚至相信所有的几何必须由它演绎出来.

一　分类　向量空间 V 中的子集 Z 称为**锥**如果它在纯量乘下是不变的 (即如果这个子集 Z 含有一个向量, 那么它含有这个向量所有的倍数, 等价的说法是这个向量张成的子空间含在 Z 中). 设 V 是域 K 上维数 $\geqslant 2$ 的向量空间. 显然, V 上的二次型 q 的零点集

$$S_q = \{x \in V \mid q(x) = 0\}$$

是一个锥, 零向量是它的顶点.

对于 V 中的锥 Z, 所有在 Z 中的 1 维子空间形成 $\mathbb{P}(V)$ 的一个子集, 称为 Z 的**射影化**(projectivization), 记作 $\mathbb{P}(Z)$ 或 $\bar{Z}$. 显然, $\bar{Z}$ 是 $Z\backslash\{0\}$ 在自然映射 $\pi : V^{\times} \to \mathbb{P}(V)$, $x \to \bar{x}$ 下的像.

定义 5.46　向量空间 V 上的二次型 q 的零点集 S_q 的射影化 $\bar{S}_q$ 称为**射影二次曲面**(当 $\dim V = 3$ 时, 也称为**射影二次曲线**).

这个定义和前面射影簇的一般定义是吻合的. 这说的是, q 作为 V 上的二次齐次多项式, 它在 $\mathbb{P}(V)$ 中的零点集是确切定义的, 它等于 $\bar{S}_q$. 射影二次曲面 $\bar{S}_q$ 称为**非退化的**如果 q 是非退化的且 S_q 含有非零向量. 在 K 是实数域时, 平面是退化的二次曲面, 例如, 取定 V 的基后, 二次型 $q(x)=x_0^2+\cdots+x_r^2$ 的零点集是平面 $x_0=\cdots=x_r=0$.

在 V 中随意取一个基 $e_0,e_1,\cdots,e_n$, 那么锥 S_q 的方程有如下形式:

$$q(x)=q(x_0,x_1,\cdots,x_n)=\sum_{i,j=0}^{n}a_{ij}x_ix_j=0. \tag{5.4.49}$$

这个方程也是射影二次曲面 $\bar{S}_q$ 的齐次坐标方程式. 如果所取的基是 q 的典范基 (更确切地说是 q 的极化 f 的典范基, 见定理 1.76), 那么 S_q 和 $\bar{S}_q$ 的方程有如下典范形式

$$q(x)=\lambda_0x_0^2+\lambda_1x_1^2+\cdots+\lambda_rx_r^2. \tag{5.4.50}$$

在 $K=\mathbb{C}$ 的情形, 射影二次曲面的方程的典范式更简单, 它完全由二次型的秩决定:

$$q(x)=x_0^2+x_1^2+\cdots+x_r^2=0. \tag{5.4.51}$$

在 $K=\mathbb{R}$ 的情形, 射影二次曲面的方程的典范式略微复杂一些:

$$x_0^2+\cdots+x_s^2-x_{s+1}^2-\cdots-x_r^2=0,\quad \frac{r-1}{2}\leqslant s\leqslant r. \tag{5.4.52}$$

(可以要求 $(r+1)/2\leqslant s$ 是因为必要时可以对方程两端乘以 -1.) 在 $s=r$ 时, $\bar{S}_q$ 是平面或是空集 ($s=r=n$ 的情形.) 于是, 实射影二次曲面由二次型的秩和符号差完全确定 (已假定正惯性指数不小于负惯性指数).

可以借助线性变换 $\mathcal{A}\in GL(V)$ 实现从一个基到另一个基的转换, 而 $\mathcal{A}$ 对应着射影变换 $\bar{\mathcal{A}}\in PGL(V)$. 所以, 在射影等价的意义下 (即两个二次曲面等价如果有射影变换把一个变成另一个), 复射影空间中的非退化二次曲面只有一个.

由定理 5.16 和实二次型的惯性定律知, 在射影等价的意义下, 实射影空间中非退化的二次曲面有 $\left[\dfrac{\dim V}{2}\right]$ 个. 正惯性指数对实射影二次曲面的几何结构也是有影响的. 下面是有关的一个结论.

定理 5.47 假设非退化实射影二次曲面 S 的方程是 (5.4.52), 那么在 S 中的平面的最大维数是 $n-s-1$.

证明 由于 S 是非退化的, 所以 $r=n=\dim V-1$. 显然, 方程组

$$x_{s+1}=\cdots=x_n=0$$

定义的 s 维平面 $\Pi_0 \subset \mathbb{P}(V)$ 与 S 不相交. 由定理 5.32 知, $\mathbb{P}(V)$ 中每个维数 $\geqslant n-s$ 的平面与 Π_0 都相交, 所以它们都不在 S 中. 另一方面, 方程组

$$x_0 = x_{s+1},\ x_1 = x_{s+2},\ \cdots,\ x_{n-s-1} = x_n,\ x_{n-s} = 0,\ \cdots,\ x_s = 0,$$

定义了 $\mathbb{P}(V)$ 中一个 $n-s-1$ 维平面, 它在 S 中. □

这个定理告诉我们, 非退化的实射影二次曲面不含直线当且仅当它的正惯性指数 $s+1=\dim V-1$, 这时二次曲面称为**椭球面**(ellipsoid). 其余的非退化实射影二次曲面称为**直纹面**(ruled surface).

下面的结论和定理 5.25 密切相关, 也是它的射影类似物.

定理 5.48 对每个非退化的实射影二次曲面 $S=\bar{S}_q$, 保持 S 不变的射影变换全体形成的群 $G(S)$ 可迁地作用在 S 上.

证明 非退化意味着零向量是 S_q 的唯一顶点. 根据定理 5.23, 保持 S_q 不变的线性变换全体形成的群可迁地作用在 $S_q\backslash\{0\}$ 上. 射影化后即得所要的结论. □

二 在仿射图上的截图 由方程 (5.4.49) 定义的射影二次曲面 $\bar{S}_q$ 在仿射图 $\mathbb{A}_0$ 上的截图 (即与仿射图的交集) 由仿射坐标的方程

$$q(1, x_1, \cdots, x_n) = 0 \tag{5.4.53}$$

确定, 它与 (相对于 $\mathbb{A}_0$ 的) 无穷远平面的交集的齐次坐标方程是

$$q(0, x_1, \cdots, x_n) = 0. \tag{5.4.54}$$

值得注意的是, 仿射图 $\mathbb{A}_0$(因而, 任何仿射图) 上的二次曲面 S 如果不是平面, 那么它确定了一个射影二次曲面 $\bar{S}$. 实际上, 在 S 的坐标方程中, 每个线性项乘入 x_0, 常数项乘入 x_0^2, 就得到曲面 $\bar{S}$ 的齐次坐标方程. 由于 S 不是平面, 定理 5.16 蕴含 $\bar{S}$ 由 S 唯一确定: 如果 $\bar{T}$ 是另一个射影二次曲面使得 $\bar{T}\cap\mathbb{A}_0=S$, 那么 $\bar{T}=\bar{S}$. 由于这些事实, 射影二次曲面在仿射图上的截图也称为该曲面的**仿射表达**. 当 $\bar{S}$ 是平面时, S 是平面或是空集. 如果 S 是非空的平面, 那么其定义方程也是 $\bar{S}$ 的定义方程, 至少在实数域的情况是如此.

射影二次曲面的分类一方面比仿射几何和欧几里得几何中的二次曲面的分类简单, 另一方面, 例 5.34 又说明一个射影二次曲线可以把若干个仿射二次曲线 (椭圆、双曲、抛物) 统一起来. 高维的情形有类似的现象. 例如, 在 3 维实射影空间 $\mathbb{RP}^3$ 中, 有两个不同的非退化二次曲面, 分别由方程式

$$x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0, \quad x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0 \tag{5.4.55}$$

定义.

二次曲面 $\bar{S}: x_0^2+x_1^2+x_2^2-x_3^2=0$ 在仿射图上的截图可以是椭球面、双叶双曲面和椭圆抛物面. 第一种情形对应的仿射图是 $\mathbb{A}_3=e_3+V_3$, 第二种情形对应的仿射图是 $\mathbb{A}_0$, $\mathbb{A}_1$ 或 $\mathbb{A}_2$. 第三种情形对应的仿射图在新的坐标系

$$y_0=x_0,\quad y_1=x_1,\quad y_2=x_2-x_3,\quad y_3=x_2+x_3$$

下更容易描述. 此时, 二次曲面 $\bar{S}$ 的方程如下:

$$y_0^2+y_1^2+y_2y_3=0.$$

曲面与仿射图 $\mathbb{A}_3'=e_3'+V_3'$ 的交集是椭圆抛物面. 在这三种情形, 射影二次曲面 $\bar{S}$ 与无穷远平面 $\mathbb{P}(V_i)$ 的交集是有区别的:

(1) $\bar{S}\cap\mathbb{P}(V_3)=\varnothing$, 这时 $\bar{S}\cap\mathbb{A}_3$ 是椭球面;

(2) $\bar{S}\cap\mathbb{P}(V_2)$ 是圆周, 这时 $\bar{S}\cap\mathbb{A}_2$ 是双叶双曲面;

(3) $\bar{S}\cap\mathbb{P}(V_3')$ 是点 (0:0:1:0), 这时 $\bar{S}\cap\mathbb{A}_3'$ 是椭圆抛物面.

同样的方式可以说明, 曲面 $\bar{T}: x_0^2+x_1^2-x_2^2-x_3^2=0$ 在仿射图上的截图可以是单叶双曲面、双叶双曲面, 但不会有椭球面和椭圆抛物面.

还有一个事实也是值得注意的, 在射影空间中, 柱面和锥面的差别都被磨光了. 原因在于, 柱面的直母线从仿射的观点看是平行的, 在射影的观点下, 它们都交于无穷远点. 比如, 三维仿射空间中的锥面 $x_1^2+x_2^2-x_3^2=0$, 椭圆柱面 $x_1^2+x_2^2-1=0$, 抛物柱面 $x_1^2+x_2=0$ 都是 3 维射影空间中柱面 $\alpha^2+\beta^2-\gamma^2=0$ 在不同的仿射图中的截图, 这些仿射二次曲面确定的射影二次曲面就是这个柱面.

三 直线与射影二次曲面的交 在射影二次曲面的分类中我们已经看到复射影空间的独特优越性: 二次型的秩完全确定了曲面的典范式方程. 不仅如此, 我们即将看到, 复射影空间中的二次曲面和任意直线都有交点.

假设复射影空间 $\mathbb{CP}^n$ 中的二次曲面 $\bar{S}$ 的由方程 (5.4.49) 定义, 直线 L 过点 $\bar{u}$ 和 $\bar{v}$, 从而它的齐次坐标有如下形式

$$x_i=\alpha a_i+\beta b_i,\quad i=0,1,\cdots,n, \tag{5.4.56}$$

其中 $\bar{u}=(a_0:a_1:\cdots:a_n)$, $\bar{v}=(b_0:b_1:\cdots:b_n)$. 把上式代入式 (5.4.49), 得

$$q(u)\alpha^2+2f(u,v)\alpha\beta+q(v)\beta^2=0. \tag{5.4.57}$$

如果 $q(u)=0$ 或 $q(v)=0$, 那么 u 或 v 已经在曲面 $\bar{S}$ 上了. 当 $q(u)=q(v)=f(u,v)=0$ 时, 方程式 (5.4.57) 是恒等式, 于是直线 L 全在二次曲面 $\bar{S}$ 上, 成为它的一条直母线.

假设 $q(u)$ 和 $q(v)$ 都不等于零. 当判别式

$$D = f(u,v)^2 - q(u)q(v)$$

不等于零时, 式 (5.4.57)(在 $\mathbb{CP}^1$ 中) 有两个非零解, 对应直线 L 与二次曲面 $\bar{S}$ 的两个交点. 当 $D=0$ 时, 直线与二次曲面有一个二重交点. 如果这个二重交点 $\bar{p}$ 不是曲面的奇点, 即 $\dfrac{\partial q}{\partial x_i}|_{\bar{p}} \neq 0$ 至少对某个 i 成立, 那么直线 L 在交点处与二次曲面相切.

自然, 渐近方向的概念对射影二次曲面是没有意义的.

习 题 5.4

1. 证明: 实射影空间 $\mathbb{RP}^3$ 中二次曲面的等价类有 8 个 (包括零二次曲面、退化二次曲面和二次型定义的平面), 在适当的齐次坐标系下, 这些等价类可由方程式

(1) $\xi_0^2+\xi_1^2+\xi_2^2+\xi_3^2=0$;

(2) $\xi_0^2+\xi_1^2+\xi_2^2-\xi_3^2=0$;

(3) $\xi_0^2+\xi_1^2-\xi_2^2-\xi_3^2=0$;

(4) $\xi_0^2+\xi_1^2+\xi_2^2=0$;

(5) $\xi_0^2+\xi_1^2-\xi_2^2=0$;

(6) $\xi_0^2+\xi_1^2=0$;

(7) $\xi_0^2-\xi_1^2=0$;

(8) $\xi_0^2=0$

分别给出.

2. 利用上题和习题 5.2 中的第 2 题, 指出 (2)、(3) 和 (5) 型二次曲面在仿射图 $\mathbb{A}_i$ $(i=0,1,2,3)$ 中的截图 (或说仿射表达).

3. 设 $\bar{v}_0 \in \mathbb{P}(V)$ 是射影二次曲面 $\bar{S}$ 上的光滑点. 证明: 过 $\bar{v}_0$ 的切线全体形成 $\mathbb{P}(V)$ 的一个超平面.

4. 证明: 对实射影空间 $\mathbb{RP}^n$, 仿射图 $\mathbb{A}_0$ 中的抛物面与无穷远平面 $\mathbb{P}(V_0)$ 相切.

5. 实射影空间 $\mathbb{RP}^n$ 中方程

$$x_1^2+\cdots+x_n^2-x_{n+1}^2=0$$

定义了一个椭球面 S. 证明: 保持 S 不变的射影变换形成的群 $G(S)$ 可迁地作用在 S 的内部

$$x_1^2+\cdots+x_n^2<x_{n+1}^2.$$

(参见 5.2 节中的习题 9.)

第6章　张　　量

到目前为止, 我们基本上都是在和向量空间打交道, 出现的角色有纯量、向量、向量空间、线性函数、线性映射、线性算子、双线性型, 甚至还有半线性函数和映射. 从一个向量空间衍生出它的对偶空间、双线性型空间、线性算子空间等. 在1.7节中还定义了多重线性映射和多重线性函数, 不过没有开展一般的讨论. 现在, 我们准备用新的观点 —— 张量, 统一描述这些对象, 而且, 把它们放入一个代数结构中.

张量 (代数) 与其说是内容丰富的理论, 不如说是一套非常有用且不可缺少的语言. 它源于 (微分) 几何, 在物理学和几何中都十分重要.

6.1　张量计算初步

有几种不同的方式定义张量, 它们说的事情都是一样的, 只是用不同的语言, 抽象程度不同而已. 这里先用多重线性函数的语言定义张量, 给一些简单的讨论, 然后用张量坐标刻画张量, 从而给出张量的另一个定义. 在下一节引入线性空间的张量积, 得到张量的第三个定义.

一　张量的概念 为了与过去所学的内容衔接, 也为了更容易说明张量的本质和若干性质, 先讨论一个最重要的特殊情况.

定义 6.1　设 K 是域, V 是 K 上的向量空间, 对偶是 V^*, p 和 q 均为非负整数,

$$V^p \times (V^*)^q = \underbrace{V \times \cdots \times V}_{p} \times \underbrace{V^* \times \cdots \times V^*}_{q}$$

是 p 个空间 V 和 q 个空间 V^* 的笛卡尔积. 所有的 $p+q$ 重线性函数

$$f : V^p \times (V^*)^q \to K$$

都称为 V 上的 (p,q) 型**张量**(tensor), 或 $p+q$ 阶 (order) 张量. 这些线性函数也说成 p 次共变 (covariant) 和 q 次反变 (contravariant) 的**混合张量**(mixed tensor). 当 $p=0$ 时, 简单称 f 是**反变的**, 当 $q=0$ 时, 则说 f 是**共变的**.

下面我们看到, 在 $p+q \leqslant 2$ 的情况, 张量就已经把线性函数、向量、双线性型、线性算子、纯量等概念统一起来.

按定义, (1,0) 型张量就是通常的 V 上的线性函数, 也就是 V^* 的元素. 而 (0,1) 型张量是 V^* 上的线性函数, 也就是 V^{**} 的元素. 由于有限维空间的自反性 (定理 1.53), V 和 V^{**} 之间存在自然同构, 允许把 $\varphi \in V^{**}$ 与向量 $x_\varphi \in V$ 等同起来 (或者把向量 x 与线性函数 $f_x \in V^{**}$ 等同起来). 这个等同可以在记号

$$\alpha(x) = (\alpha, x) \tag{6.1.1}$$

下实现, 它在 1.7 节中用过. 当 $\alpha \in V^*$ 固定时, 它是 V 上的线性函数 α, 当向量 $x \in V$ 固定时, 它是 V^* 上的线性函数 f_x. 这样一来, 可以认为 (0,1) 型张量是向量, 也就是 V 的元素.

显然, (2,0) 型张量是 V 上的双线性型, (0,2) 型张量是 V^* 上的双线性型. 这两类张量中前者是共变的, 后者是反变的.

最简单的混合型张量是 (1,1) 型张量, 它的解读是挺有趣的. 根据定义, 函数 $f(x,\alpha)$ 对 $x \in V$ 和 $\alpha \in V^*$ 均是线性的. 固定 x, 那么 $f(x,\alpha)$ 就成了 V^* 的线性函数, 从而有 V 中的向量与之对应 (定理 1.53), 该向量由 x 确定, 可记作 $\mathcal{F}x$. 于是, $f(x,\alpha) = \alpha(\mathcal{F}x)$. 由 (6.1.1), 得

$$f(x,\alpha) = (\alpha, \mathcal{F}x). \tag{6.1.2}$$

因为 $f(ax+by,\alpha) = af(x,\alpha) + bf(y,\alpha)$, 所以

$$(\alpha, \mathcal{F}(ax+by)) = a(\alpha,\mathcal{F}x) + b(\alpha,\mathcal{F}y) = (\alpha, a\mathcal{F}x + b\mathcal{F}y),$$

从而有

$$\mathcal{F}(ax+by) = a\mathcal{F}x + b\mathcal{F}y,$$

也就是说, $\mathcal{F}$ 是 V 上的线性算子. 反过来, 对每个线性算子 $\mathcal{F} \in \mathcal{L}(V)$, 按照公式 (6.1.2) 定义函数 $f: V \times V^* \to K$, 那么 f 对 $x \in V$ 和 $\alpha \in V^*$ 都是线性的. 容易看出, $f \to \mathcal{F}$ 是从 V 上的 (1,1) 型张量到 V 上的线性算子的双射. 这样一来, (1,1) 型张量和线性算子本质上是一回事.

自然, 应该约定 (0,0) 型张量就是纯量. 现在可以看出, 张量确实能把以前遇到的若干重要概念统一起来.

空间 V 上的所有 (p,q) 型张量形成的集合 $\mathbb{T}^{p,q} = \mathbb{T}_p^q(V)$ 是一个向量空间. 事实上, 对 $f,g \in \mathbb{T}_p^q(V)$ 和 $a,b \in K$, 定义

$$\begin{aligned}&(af+bg)(v_1,\cdots,v_p;\alpha_1,\cdots,\alpha_q)\\ =&af(v_1,\cdots,v_p;\alpha_1,\cdots,\alpha_q) + bg(v_1,\cdots,v_p;\alpha_1,\cdots,\alpha_q),\end{aligned} \tag{6.1.3}$$

则有 $af+bg \in \mathbb{T}_p^q(V)$.

二 多重线性函数的张量积 任意给定两个多重线性函数

$$f : V_1 \times \cdots \times V_r \to K, \qquad g : W_1 \times \cdots \times W_s \to K.$$

定义 6.2 多重线性函数 f 和 g 的张量积是一个多重线性函数

$$f \otimes g : V_1 \times \cdots \times V_r \times W_1 \times \cdots \times W_s \to K,$$

它由公式

$$(f \otimes g)(v_1, \cdots, v_r; w_1, \cdots, w_s) = f(v_1, \cdots, v_r)g(w_1, \cdots, w_s)$$

定义. 需要强调的是, 此处诸空间 V_i 和 W_j 之间没有任何关系, 诸变量 v_i 和 w_j 互相独立.

例 6.3 如果 f 和 g 是 V 上的线性函数, 那么

$$(f \otimes g)(x, y) = f(x)g(y)$$

是 V 上的双线性型. 这个简单的例子可以打消张量积交换的幻想. 事实上,

$$(g \otimes f)(x, y) = g(x)f(y) \neq f(x)g(y) = (f \otimes g)(x, y).$$

不过, 张量积满足结合律. 如果 $h : U_1 \times \cdots \times U_t \to K$ 是多重线性函数, 那么

$$(f \otimes g) \otimes h = f \otimes (g \otimes h). \tag{6.1.4}$$

这是因为, 作为函数, 上式左右边的值都是

$$f(v_1, \cdots, v_r)g(w_1, \cdots, w_s)h(u_1, \cdots, u_t).$$

现在设 f 和 g 分别是 V 上的 (p, q) 型张量和 (r, s) 型张量, 那么 $f \otimes g$ 是笛卡尔积

$$V^p \times (V^*)^q \times V^r \times (V^*)^s$$

上的多重线性函数. 把这个笛卡尔积与

$$V^{p+r} \times (V^*)^{q+s}$$

等同起来. 对任意的 $v_i \in V$ 和 $\alpha_j \in V^*$, 可要求

$$\begin{aligned} &(f \otimes g)(v_1, \cdots, v_{p+r}; \alpha_1, \cdots, \alpha_{q+s}) \\ &= f(v_1, \cdots, v_p; \alpha_1, \cdots, \alpha_q)g(v_{p+1}, \cdots, v_{p+r}; \alpha_{q+1}, \cdots, \alpha_{q+r}), \end{aligned} \tag{6.1.5}$$

从而 $f\otimes g$ 是 V 上的 $(p+r,q+s)$ 型张量, 称为 f 和 g 的**(张量) 乘积**.

在下面, 用来区别不同类型变量的分号, 依照惯例, 将用逗号代替.

例 6.4 设 f,g 是 V 上的线性函数, u,v,w 是 V 中的向量. 正如前面所说明的, 它们分别是 (1,0) 型张量和 (0,1) 型张量. 于是, 乘积

$$\theta=f\otimes g\otimes u\otimes v\otimes w$$

是 V 上的 (2,3) 型张量. 如果 $x,y\in V$, $\alpha,\beta,\gamma\in V^*$, 那么

$$\theta(x,y,\alpha,\beta,\gamma)=f(x)g(y)\alpha(u)\beta(v)\gamma(w).$$

(注意 $\alpha(u)$ 还可以写成 $u(\alpha)=(\alpha,u)=(u,\alpha)$.)

由公式 (6.1.5) 和张量的线性组合 $af+bg$ 的定义公式 (6.1.3) 可以看出, 张量的乘积具有分配律

$$\begin{aligned}(af+bg)\otimes h&=af\otimes h+bg\otimes h,\\ h\otimes(af+bg)&=ah\otimes f+bh\otimes g.\end{aligned}\tag{6.1.6}$$

以上内容可小结如下:

1) 定义了张量的乘积运算 $\otimes$;

2) 乘积的阶等于各因子的阶的和;

3) 张量的乘积运算是结合的, 满足分配律, 但不是交换的.

三 张量的坐标 在有限维情形, 张量可以通过坐标刻画, 这同向量可以通过坐标刻画一样. 从而, 张量具有坐标形式, 这是张量在物理学和几何中常出现的形式, 也是历史上张量出现的最初形式.

通常, 在 V 和 V^* 中选择互相对偶的基

$$V=\langle e_1,\cdots,e_n\rangle,\quad V^*=\langle e^1,\cdots,e^n\rangle.$$

注意这里 V 中基向量的指标在右下角, V^* 中基向量的指标在右上角. 为了直观和计算的便利, 空间中向量的坐标的指标与基向量的指标是上下对立的, 即对 $x\in V$ 和 $f\in V^*$, 有

$$x=\sum_{i=1}^n\xi^ie_i,\quad f=\sum_{j=1}^n\eta_je^j.$$

我们有

$$(e_i,e^j)=(e^j,e_i)=\delta_i^j=\begin{cases}0, & 如果\ i\neq j,\\ 1, & 如果\ i=j.\end{cases}\tag{6.1.7}$$

从而

$$f(x)=(f,x)=(x,f)=\sum_{i=1}^n\xi^i\eta_i.$$

上指标与可以判断的方幂不能混淆. 不过, 我们这里不会有这种事情发生.

在张量分析中, 爱因斯坦 1916 年引入的**和式约定**(Einstein summation convention, 或 Einstein notation) 在物理学中经常使用, 在数学中, 尤其是微分几何中也很常用. 在这个约定下, 求和 $\sum\limits_{i=1}^{n}\xi^i e_i$, $\sum\limits_{i=1}^{n}\eta_j e^j$, $\sum\limits_{i=1}^{n}\xi^i\eta_i$ 等会被简单地分别记作 $\xi^i e_i$, $\eta_j e^j$, $\xi^i\eta_i$. 这里的要点是: 如果一个指标出现两次, 就对这个指标求和, 指标的范围是默认清楚的; 如果一个指标仅出现一次, 不对这个指标求和. 在和式约定中, 求和的指标也称为**哑指标**(dummy index), 它可以用任何记号替代而不改变表达式的含义; 不求和的指标也称为**自由指标**(free index). 例如, 在这个约定下, 表达式 $a^i_j b^j_k$ 对指标 j 求和, 但对指标 i 和 k 都不求和, j 是哑指标, i 和 k 都是自由指标.

爱因斯坦的和式约定有它的优越性: 简洁, 计算高效. 我们这儿不用爱因斯坦和式约定, 不过, 约定按不同的指标求和可以简单地用一个求和号表示, 即

$$\sum_i\sum_j\cdots\sum_k=\sum_{i,j,\cdots,k},$$

并且不指明求和指标的取值范围, 因为, 根据上下文, 指标的取值范围是自明的 (通常从 1 到 $n=\dim V$).

由于张量是多重线性的, 所以 V 的一个 (p,q) 型张量由它的如下值完全确定

$$T^{j_1\cdots j_q}_{i_1\cdots i_p}:=T(e_{i_1},\cdots,e_{i_p},e^{j_1},\cdots,e^{j_q}),\quad 1\leqslant i_1,\cdots,i_p,j_1,\cdots,j_q\leqslant n. \tag{6.1.8}$$

(这与线性函数、双线性型的情形是类似的.)

定义 6.5 诸纯量 $T^{j_1\cdots j_q}_{i_1\cdots i_p}$ 称为张量 T 在基 $e_1,\cdots,e_n$ 下的**坐标**(**系数**, 或**分量**).

在我们的印象中, 向量的坐标是和基联系在一起的. 借助对偶基的思想, 我们应当考虑 (p,q) 型张量

$$e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}, \tag{6.1.9}$$

这里, 如同以前做过的一样, 把 $e_{j_1},\cdots,e_{j_q}$ 与 V^* 的线性函数 $e_{j_k}(f)=f(e_{j_k})$ 等同起来. 因为 $(e^i,e_{i'})=\delta^i_{i'}$, $(e_k,e^{k'})=\delta^{k'}_k$, 所以

$$(e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q})(e_{i'_1},\cdots,e_{i'_p},e^{j'_1},\cdots,e^{j'_q})=\delta^{i_1}_{i'_1}\cdots\delta^{i_p}_{i'_p}\delta^{j'_1}_{j_1}\cdots\delta^{j'_q}_{j_q}. \tag{6.1.10}$$

构造张量

$$T_1=\sum_{i,j}T^{j_1\cdots j_q}_{i_1\cdots i_p}e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q},$$

它是张量 (6.1.9) 的线性组合, 以 T 的坐标 (6.1.8) 为系数. 由公式 (6.1.3) 和 (6.1.10) 知

$$T_1(e_{i_1},\cdots,e_{i_p},e^{j_1},\cdots,e^{j_q})=T^{j_1\cdots j_q}_{i_1\cdots i_p}.$$

也就是说, 张量 T_1 的坐标和张量 T 的坐标一致. 由于张量完全由它的坐标确定, 所以 $T_1=T$, 从而有

$$T=\sum_{i,j}T^{j_1\cdots j_q}_{i_1\cdots i_p}e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}. \tag{6.1.11}$$

把上式应用到 V 的双线性型 f, 得

$$f=\sum_{i,j}f_{ij}e^i\otimes e^j.$$

还需要说明, 形如式 (6.1.9) 的张量全体是线性无关的. 这可由公式 (6.1.9) 推出. 的确, 线性关系式

$$\sum_{i,j}\lambda^{j_1\cdots j_q}_{i_1\cdots i_p}e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}=0,\quad \lambda^{j_1,\cdots,j_q}_{i_1,\cdots,i_p}\in K$$

左右两边在 $(e_{i'_1},\cdots,e_{i'_p},e^{j'_1},\cdots,e^{j'_q})$ 处的值分别是 $\lambda^{j'_1\cdots j'_q}_{i'_1\cdots i'_p}$ 和 0, 所以左边的系数全为零.

于是, 形如式 (6.1.9) 的张量全体构成向量空间 $\mathbb{T}^q_p(V)$ 的一个基. 基向量中的指标 $i_1,\cdots,i_p,j_1,\cdots,j_q$ 可以在 1 到 n 间的整数可以相互独立地任意取值, 所以, 这个基含有 n^{p+q} 个向量.

直观上看, 可以把张量 T 的坐标排成一个空间立方体表, 立方体的维数就是向量的阶数, 我们熟知的行向量, 列向量, 方阵都是这种 $p+q$ 维表的特例.

上面的讨论可总结如下.

定理 6.6 向量空间 V 上的 (p,q) 型张量全体形成一个 n^{p+q} 维的向量空间 $\mathbb{T}^q_p(V)$. 向量组

$$e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q},\quad 1\leqslant i_1,\cdots,i_p,j_1,\cdots,j_q\leqslant n$$

是它的一个基, 其中 $e_1,\cdots,e_n$ 是 V 的基, $e^1,\cdots,e^n$ 是 V^* 中相应的对偶基.

存在唯一的张量, 它具有预先任意给定的坐标 $T^{j_1\cdots j_q}_{i_1\cdots i_p}$.

四 张量的坐标变换 我们需要知道对 V 的不同的基, 张量的坐标是怎样变换的. 设 $e'_1,\cdots,e'_n$ 是 V 的另一个基, $e'^1,\cdots,e'^n$ 是 V^* 中相应的对偶基. 从 (e_i) 到 (e'_j) 的转换矩阵记作 $A=(a^i_j)$. 矩阵 $A=(a^i_j)$ 中的元素的上指标表示行, 下指标表示列. 也就是说,

$$e'_k=\sum_i a^i_k e_i,\quad k=1,\cdots,n. \tag{6.1.12}$$

从 (e^i) 到 (e'^i) 的转换矩阵的转置矩阵记作 $B=(b^i_j)$. 于是有

$$e'^k=\sum_i b^k_i e^i, \quad k=1,\cdots,n. \tag{6.1.13}$$

这里只有做了转置, 才能保证坐标变换时指标的上下正确对应. 再引入辅助矩阵 $C=B^{-1}=(c^i_j)$:

$$e^k=\sum_i c^k_i e'^i.$$

由对偶基的性质, 得

$$c^k_j=\left(\sum_i c^k_i e'^i, e'_j\right)=(e^k,e'_j)=\left(e^k,\sum_i a^i_j e_i\right)=a^k_j.$$

由此可见, $C=A$, 从而

$$e^k=\sum_i a^k_i e'^i,$$

且有 $B=A^{-1}$. 于是, (e^i) 到 $e'^i)$ 的转换矩阵是 ${}^tB={}^t(A^{-1})={}^tA^{-1}$, 它称为 A 的**转置逆矩阵**.

现在, 可以求出张量 T 在基

$$e'^{i_1}\otimes\cdots\otimes e'^{i_p}\otimes e'_{j_1}\otimes\cdots\otimes e'_{j_q}, \quad 1\leqslant i_1,\cdots,i_p,j_1,\cdots,j_q\leqslant n$$

之下的坐标 $T'^{j_1\cdots j_q}_{i_1\cdots i_p}$. 按公式 (6.1.8), 有

$$\begin{aligned}
T&=\sum_{i,j} T^{j_1\cdots j_q}_{i_1\cdots i_p}\, e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}\\
&=\sum_{i',j'} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p}\, e'^{i'_1}\otimes\cdots\otimes e'^{i'_p}\otimes e'_{j'_1}\otimes\cdots\otimes e'_{j'_q}\\
&=\sum_{i,j}\left(\sum_{i',j'} b^{i'_1\cdots i'_p}_{i_1\cdots i_p} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p}\, a^{j_1\cdots j_q}_{j'_1\cdots j'_q}\right) e^{i_1}\otimes\cdots\otimes e^{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}.
\end{aligned}$$

这说明, 下面的定理成立.

定理 6.7 设 (e_i) 和 (e^i) 分别是空间 V 和 V^* 的基, 它们相互对偶的. 如果按公式 (6.1.12) 和 (6.1.13) 把它们转换成另一对相互对偶的基 (e'_i) 和 (e'^i), 那么 V 的 (p,q) 型张量 T 的坐标按公式

$$T^{j_1\cdots j_q}_{i_1\cdots i_p}=\sum_{i',j'} b^{i'_1\cdots i'_p}_{i_1\cdots i_p} T'^{j'_1\cdots j'_q}_{i'_1\cdots i'_p}\, a^{j_1\cdots j_q}_{j'_1\cdots j'_q} \tag{6.1.14}$$

转换, 其中 $b_{i_1\cdots i_p}^{i'_1\cdots i'_p}=b_{i_1}^{i'_1}\cdots b_{i_p}^{i'_p}$, $a_{j'_1\cdots j'_q}^{j_1\cdots j_q}=a_{j'_1}^{j_1}\cdots a_{j'_q}^{j_q}$, 而且矩阵 $B=(b_j^i)$ 是矩阵 $A=(a_j^i)$ 的逆矩阵.

从公式 (6.1.14) 可以看出, 矩阵 $A=(a_j^i)$ 的上指标与 $B=(b_j^i)=A^{-1}$ 的下指标不参与求和, 所以它们出现在求和后的张量坐标.

定理 6.7 提供了张量的另一个定义. 所谓 V 上的一个 (p,q) 型张量 T, 是指从 V 的基全体形成的集合 $\mathbf{B}$ 到 n^{p+q} 元纯量组空间

$$\{(\lambda_{i_1\cdots i_p}^{j_1\cdots j_q})\in K^{n^{p+q}}\,|\,1\leqslant i_1,\cdots,i_p,j_1,\cdots,j_q\leqslant n=\dim V\}$$

的一个映射, 不同的基对应的纯量组之间由公式 (6.1.14) 联系. 这个定义有它方便之处. 比如说, 对同型张量 S 和 T, 它们的线性组合 $aS+bT$ 就可以说是以

$$aS_{i_1\cdots i_p}^{j_1\cdots j_q}+bT_{i_1\cdots i_p}^{j_1\cdots j_q}$$

为坐标的张量. 任意两个张量 $(Q_{i_1\cdots i_p}^{j_1\cdots j_q})$ 和 $(R_{k_1\cdots k_s}^{l_1\cdots l_t})$ 的乘积的坐标是

$$T_{i_1\cdots i_pk_1\cdots k_s}^{j_1\cdots j_ql_1\cdots l_t}=Q_{i_1\cdots i_p}^{j_1\cdots j_q}\cdot R_{k_1\cdots k_s}^{l_1\cdots l_t}. \tag{6.1.15}$$

容易看出, 这个公式与定理 6.7 中的公式是相容的.

例 6.8 从第一部分的讨论知道, 空间 V 上的线性算子 $\mathcal{F}$ 可以等同于一个 (1,1) 型张量. 如果 $e_1,\cdots,e_n$ 是 V 的基且 $\mathcal{F}e_k=\sum\limits_i f_k^ie_i$, 那么算子的张量形式是

$$T=\sum_{i,j}f_j^ie^j\otimes e_i.$$

按公式

$$e'_k=\sum_i a_k^ie_i,\quad e_k=\sum_i b_k^ie'_i,\quad \sum_i b_k^ia_i^l=\delta_k^l$$

转化基时, 有

$$\sum_i {f'}_j^ie'_i=\mathcal{F}e'_j=\sum_l a_j^l\mathcal{F}e_l=\sum_{k,l}a_j^lf_l^ke_k=\sum_{k,l,i}a_j^lf_l^kb_k^ie'_i.$$

因此得

$${f'}_j^i=\sum_{k,l}a_j^lf_l^kb_k^i. \tag{6.1.16}$$

这正是张量 T 的坐标形式 (f_j^i) 的转换公式 (6.1.14). 当然, 上式的矩阵形式就是我们很早就证明了的公式 (2.2.10). 用爱因斯坦的和式约定, 上式可以简单写成 ${f'}_j^i=a_j^lf_l^kb_k^i$.

在数学和物理学中, 张量大部分情况是服从变换规则 (6.1.14) 的自然的几何对象. 在实际的应用中, 尤其是在微分几何与物理学中, 需要考虑的不是单个张量, 而是张量场, 如曲率张量, 引力场张量, 应力能张量场等. 也就是说, 在一个空间或流形上, 每一点都有一个张量与之对应, 从而张量的坐标都是函数, 而非纯量. 这时, 坐标变换规则会涉及坐标函数的偏导数.

习 题 6.1

1. 克罗内克符号: 验证: δ_i^j 是一个张量, 更确切地说, 它是 $\mathbb{T}_1^1(V)$ 的一个元素, 对应于 V 到自身的恒等映射.

2. 计算张量 $T=e^1\otimes e_2+e^2\otimes(e_1+3e_3)\in\mathbb{T}_1^1(V)$ 的取值 $T(v,\alpha)$, 其中 $v=e_1+2e_2+5e_3$, $\alpha=e^1-e^2+2e^3$.

3. 求张量 $T\otimes S-S\otimes T\in\mathbb{T}_5^0(V)$ 在 $(v_1,\cdots,v_5)$ 处的值:

(1) $T=e^1\otimes e^2+e^2\otimes e^3-e^2\otimes e^2\in\mathbb{T}_2^0(V)$,

$S=e^1\otimes e^2\otimes(e^2-e^3)\in\mathbb{T}_3^0(V)$,

$v_1=e_1,\quad v_2=e_1+e_2+e_3,\quad v_3=e_1+e_2,\quad v_4=e_2+e_3,\quad v_5=e_2-e_3.$

(2) $T=e^1\otimes e^2-e^2\otimes e^3-e^3\otimes e^1\in\mathbb{T}_2^0(V)$,

$S\in\mathbb{T}_3^0(V)$, 它的坐标都是 1,

$v_1=e_1+e_2,\quad v_2=e_2+e_3,\quad v_3=e_1-e_2,\quad v_4=e_2,\quad v_5=e_3.$

4. 假设 $T\in\mathbb{T}_3^2(V)$ 的坐标都是 2, 求 $T(v,v,v,\alpha,\alpha)$, 其中 $v=e_1+2e_2-3e_3+e_4$, $\alpha=e^1-e^4$.

5. 假设张量 $T\in\mathbb{T}_3^2(V)$ 在基 e_1,e_2,e_3 下的坐标都是 3, 求它在基

$$(\tilde{e}_1,\tilde{e}_2,\tilde{e}_3)=(e_1,e_2,e_3)\begin{pmatrix}1&2&3\\0&1&2\\0&0&1\end{pmatrix}$$

下的坐标 $\tilde{T}_{123}^{12}$.

6. 求张量的坐标:

(1) $(e_1+e_2)\otimes(e_1-2e_2)$;

(2) $(e_1+2e_2)\otimes(e_3-2e_4)-(e_1-e_2)\otimes(e_3+e_4)$.

7. 假设 $\dim V=4$, $T=e^1\otimes e_2+e^2\otimes e_3+e^3\otimes e_4\in\mathbb{T}_1^1(V)$. 求出所有的

(1) $\alpha\in V^*$ 使得 $T(v,\alpha)=0,\ \forall v\in V$;

(2) $v\in V$ 使得 $T(v,\alpha)=0,\ \forall\alpha\in V^*$.

8. 求双线性型 $(e^1+e^2)\otimes(e^1+e^3)-e^1\otimes e^1-e^2\otimes e^2$ 的秩.

9. (度量张量) 设 V 是 n 维欧几里得向量空间, $e_1,\cdots,e_n$ 是它的基. 内积由正定双线性型

$$(x\,|\,y)=g(x,y)=\sum_{i,j}g_{ij}x^iy^j,\quad g_{ij}=g(e_i,e_j)$$

给出. (这里遵循几何学的习惯采用记号 g 和 g_{ij}.) 通常, 称 $G_0=(g_{ij})$ 为空间 V 的 (2,0) 型**度量张量**. 可见, 度量张量的分量是内积在某个基下的格拉姆矩阵的元素.

空间 V 的对偶 V^* 自然是欧几里得向量空间 (参见 3.1 节第五部分), 内积仍记为 g. 设 $e^1,\cdots,e^n$ 是 $e_1,\cdots,e_n$ 的对偶基. 命 $g^{ij}=g(e^i,e^j)$. 矩阵 $G^0=(g^{ij})$ 也被称为空间 V 的度量张量, 它是 (0,2) 型的. 从而, g_{ij} 和 g^{ij} 分别是度量张量 G 的共变坐标和反变坐标.

验证: $G^0G_0=E$, 也就是

$$\sum_j g^{ij}g_{jk}=\delta^i_k.$$

10. 三维欧几里得空间以矩阵 $(g_{ij})=\begin{pmatrix}1&1&1\\1&6&3\\1&3&2\end{pmatrix}$ 为度量张量. 求顶点为 $A(1,0,1)$, $B(2,1,1)$, $C(3,1,2)$ 的三角形的面积以及从顶点 C 引向 AB 的高. 这里坐标和张量都是在同一个基下的.

6.2 向量空间的张量积

本节中的向量空间都是域 K 上的向量空间, 除非另有说明.

一 形如 (6.1.9) 的张量可以从另一个角度看, 它们直接从空间 V 和 V^* 构造, 不经过多重线性函数这个环节. 现在对这个观点做更细致的分析. 首先, 对任意的元素 $\alpha_1,\cdots,\alpha_p\in V^*$, $v_1,\cdots,v_q\in V$, 张量乘积

$$\alpha_1\otimes\cdots\otimes\alpha_p\otimes v_1\otimes\cdots\otimes v_q$$

是 V 上的 (p,q) 型张量. 于是, 我们得到一个自然的映射

$$(V^*)^p\times V^q\to \mathrm{T}^q_p(V),\quad (\alpha_1,\cdots,\alpha_p,v_1,\cdots,v_q)\to\alpha_1\otimes\cdots\otimes\alpha_p\otimes v_1\otimes\cdots\otimes v_q.$$

因为张量乘积有分配律与结合律, 这个映射是多重线性映射的. 它不是单射, 但有一些独特的性质. 为简单起见, 考虑 $p+q=2$ 的情况, 这足以看出事情的本质. 我们可以把讨论放到更一般的框架.

命题 6.9 设 V 和 W 是向量空间, 向量组 v_i, $i\in I$ 和向量组 w_j, $j\in J$ 分别是 V 和 W 的基. 那么, 双线性映射 $\varphi:V\times W\to U$ 的如下性质是等价的:

(1) 向量组 $\varphi(v_i,w_j)$, $i\in I$, $j\in J$, 是 U 的基;

(2) 每个向量 $z\in U$ 以唯一的方式分解成 $z=\sum\limits_i\varphi(v_i,y_i)$, $y_i\in W$;

(3) 每个向量 $z\in U$ 以唯一的方式分解成 $z=\sum\limits_j\varphi(x_j,w_j)$, $x_j\in V$.

证明 如果 $z=\sum\limits_{i,j}z_{ij}\varphi(v_i,w_j)$, 那么 $z=\sum\limits_i\varphi(v_i,y_i)$, 其中 $y_i=\sum\limits_j z_{ij}w_j\in W$. 反之亦然. 由此可知, 性质 (1) 和 (2) 等价. 类似地, 性质 (1) 和 (3) 等价. □

推论 6.10 如果性质 (1) 对 V 和 W 的某些基成立, 那么它对任何基成立.

定义 6.11 向量空间 V 和 W 的**张量积**是一个向量空间 T, 带着一个双线性映射

$$\otimes : V \times W \to T, \quad (x, y) \to x \otimes y,$$

满足如下条件: 存在 V 的基 v_i, $i \in I$, 和 W 的基 w_j, $j \in J$, 使得向量组 $v_i \otimes w_j$, $i \in I$, $j \in J$, 是 T 的基.

前面的推论表明定义中对双线性映射要求的条件不依赖基的选取. 显然, 两个空间 V 和 W 的张量积存在. 的确, 命 T 为向量空间, 有一个基 t_{ij}, $i \in I$, $j \in J$. 可定义双线性映射 $\otimes : V \times W \to T$, $(x, y) \to x \otimes y$ 使得 $v_i \otimes w_j = t_{ij}$.

张量积在如下意义是唯一的: 如果 $(T, \otimes)$ 和 $(T', \otimes')$ 都是 V 和 W 的张量积, 那么, 存在 (唯一的) 同构

$$\psi : T \to T'$$

使得对任意的 $x \in V$ 和 $y \in W$ 有

$$\psi(x \otimes y) = x \otimes' y, \qquad \begin{tikzcd} & V \times W \arrow[dl, "\otimes"'] \arrow[dr, "\otimes'"] & \\ T \arrow[rr, "\psi"] & & T' \end{tikzcd} \tag{6.2.17}$$

的确, 这个同构可直接构造

$$\psi\left(\sum_{i,j} a_{ij} v_i \otimes w_j\right) = \sum_{i,j} a_{ij} v_i \otimes' w_j, \qquad a_{ij} \in K.$$

它显然满足式 (6.2.17).

向量空间的张量积是个灵巧的语言, 在代数、群表示论、几何、物理学中都有广泛的应用. 下面几个例子用张量积重新解读我们熟悉的一些空间.

例 6.12 考虑双线性映射

$$\otimes : K[x] \times K[y] \to K[x, y], \qquad (f, g) \to f \otimes g,$$

其中 $f \otimes g$ 定义如下

$$(f \otimes g)(x, y) = f(x)g(y).$$

乘积 $x^i \otimes y^j = x^i y^j$, $i, j = 0, 1, 2, \cdots$, 构成 $K[x, y]$ 的基, 所以 $K[x, y] = K[x] \otimes K[y]$. 类似地, 有

$$K[x_1, \cdots, x_m, y_1, \cdots, y_n] = K[x_1, \cdots, x_m] \otimes K[y_1, \cdots, y_n]. \tag{6.2.18}$$

从空间 V 到 W 的线性映射全体将记作 $\mathrm{Hom}\,(V;W)$, 在文献中它比 $\mathcal{L}(V,W)$ 用得更普遍. 在下面两个例子中 V 和 W 都是有限维的, 分别有基 $v_1,\cdots,v_n$ 和 $w_1,\cdots,w_m$. 这些基的对偶基分别记作 $v^1,\cdots,v^n$ 和 $w^1,\cdots,w^m$.

例 6.13 对任意的 $\alpha\in V^*$ 和 $y\in W$, 定义从 V 到 W 的线性映射 $\alpha\otimes y$ 如下:

$$(\alpha\otimes y)(x)=\alpha(x)y. \tag{6.2.19}$$

这构造了一个双线性映射

$$\otimes: V^*\times W\to \mathrm{Hom}\,(V;W).$$

由于 $v^i\otimes w_j$ 把 v_k 映到 $\delta_{ki}w_j$, 所以这些线性映射全体形成 $\mathrm{Hom}\,(V;W)$ 的一个基. 这意味着我们有

$$\mathrm{Hom}\,(V;W)=V^*\otimes W. \tag{6.2.20}$$

当 $V=W$ 时, 这给出张量与线性算子的一个直接对应, 它就是 (6.1.2) 给出的对应.

例 6.14 对任意的 $\alpha\in V^*$ 和 $\beta\in W^*$, 定义 $V\times W$ 上的双线性函数 $\alpha\otimes\beta$ 如下

$$(\alpha\otimes\beta)(x,y)=\alpha(x)\beta(y). \tag{6.2.21}$$

用 $\mathrm{Hom}\,(V,W;K)$ 记 $V\times W$ 上的双线性函数全体, 就得到一个双线性映射

$$\otimes: V^*\times W^*\to \mathrm{Hom}\,(V,W;K).$$

如果向量 $x\in V$ 和 $y\in W$ 的坐标分别是 $x_1,\cdots,x_n$ 和 $y_1,\cdots,y_m$, 那么 $(v^i\otimes w^j)(x,y)=x_iy_j$. 由于每个 $V\times W$ 上的双线性函数 γ 在 (x,y) 上的取值是 $\gamma(x,y)=\sum\limits_{i,j}c_{ij}x_iy_j$, 所以, 诸函数 $v^i\otimes w^j$ 全体形成 $\mathrm{Hom}\,(V,W;K)$ 的一个基. 因此有

$$\mathrm{Hom}\,(V,W;K)=V^*\otimes W^*. \tag{6.2.22}$$

向量空间的张量积的一个重要例子是**基域扩张**. 在 3.6 节中我们讨论了最简单的情况: 实向量空间的复化.

设 V 是域 K 上的向量空间, L 是 K 的扩域. 把 L 看作 K 的向量空间, 可以构造张量积

$$V(L)=L\otimes V.$$

根据定义, 它是域 K 上的向量空间. 有意思的是, L 中的元素可以乘向量空间 $V(L)$ 中的元素:

$$\lambda(\mu\otimes v)=\lambda\mu\otimes v,\quad \lambda,\mu\in L,\ v\in V,$$

从而 $V(L)$ 成为域 L 上的向量空间.

二 一些性质 空间的张量积运算有很多方便使用的性质. 第一个是它的普遍性 (universality), 也说成泛性.

定理 6.15 (向量空间张量积的普遍性) 对任意的双线性映射 $\varphi: V\times W\to U$, 存在唯一的线性映射 $\psi: V\otimes W\to U$ 使得

$$\varphi(x,y)=\psi(x\otimes y),\quad \forall x\in V,\ y\in W. \tag{6.2.23}$$

证明 线性映射 ψ 的存在性和唯一性都由它在基向量 $v_i\otimes w_j$ 上的值确定了:

$$\psi(v_i\otimes w_j)=\varphi(v_i,w_j).$$

□

推论 6.16 向量空间的张量积是交换的, 即存在唯一的同构

$$V\otimes W\xrightarrow{\sim}W\otimes V,$$

它把 $x\otimes y$ $(x\in V,\ y\in W)$ 映到 $y\otimes x$.

证明 映射 $V\times W\to W\otimes V$, $(x,y)\to y\otimes x$ 是双线性的. 根据定理 6.15, 存在唯一的线性映射 $V\otimes W\to W\otimes V$, 它把 $x\otimes y$ $(x\in V,\ y\in W)$ 映到 $y\otimes x$. 这个线性映射把 $V\otimes W$ 的基向量 $v_i\otimes w_j$ 映到基向量 $w_j\otimes v_i$, 所以是同构. □

命题 6.17 向量空间的张量积满足结合律, 即存在唯一的同构

$$(U\otimes V)\otimes W\xrightarrow{\sim}U\otimes(V\otimes W),$$

它把 $(x\otimes y)\otimes z$ $(x\in U,\ y\in V,\ z\in W)$ 映到 $x\otimes(y\otimes z)$.

证明 由于形如 $(x\otimes y)\otimes z$ 的向量张成 $(U\otimes V)\otimes W$, 所以定理中的线性同构的唯一性是显然的. 下证存在性. 对 $z\in W$, 映射

$$\theta_z: U\times V\to U\otimes(V\otimes W),\quad (x,y)\to x\otimes(y\otimes z)$$

是双线性的. 根据定理 6.15, 有线性映射

$$\bar\theta_z: U\otimes V\to U\otimes(V\otimes W),\quad x\otimes y\to x\otimes(y\otimes z).$$

容易看出, 映射

$$(U\otimes V)\times W\to U\otimes(V\otimes W),\quad (\alpha,z)\to\bar\theta_z(\alpha)$$

是双线性的. 再次利用定理 6.15, 得到线性映射

$$(U\otimes V)\otimes W\to U\otimes(V\otimes W)$$

它有所需的性质. □

命题 6.18　向量空间的张量积满足分配律, 即存在唯一的同构

(1) $(U\oplus V)\otimes W\xrightarrow{\sim}(U\otimes W)\oplus(V\otimes W)$, 它把 $(x,y)\otimes z$ $(x\in U,\ y\in V,\ z\in W)$ 映到 $(x\otimes z,y\otimes z)$.

(2) $U\otimes(V\oplus W)\xrightarrow{\sim}(U\otimes V)\oplus(U\otimes W)$, 它把 $x\otimes(y,z)$ $(x\in U,\ y\in V,\ z\in W)$ 映到 $(x\otimes y,x\otimes z)$.

证明　(1) 显然, 映射

$$(U\oplus V)\times W\to(U\otimes W)\oplus(V\otimes W),\quad ((x,y),z)\to(x\otimes z,y\otimes z)$$

是双线性映射, 通过定理 6.15, 就得到要求的同构.

(2) 的证明是类似的. □

三　算子的张量 (乘) 积　设 $\mathcal{A}:V\to V$ 和 $\mathcal{B}:W\to W$ 是线性算子. 映射

$$V\times W\to V\otimes W,\qquad (x,y)\to\mathcal{A}x\otimes\mathcal{B}y$$

是双线性的, 根据定理 6.15, 存在唯一的线性算子

$$\mathcal{A}\otimes\mathcal{B}:V\otimes W\to V\otimes W$$

使得

$$(\mathcal{A}\otimes\mathcal{B})(x\otimes y)=\mathcal{A}x\otimes\mathcal{B}y.\tag{6.2.24}$$

这个线性算子称为 $\mathcal{A}$ 和 $\mathcal{B}$ 的**张量积**.

由于 $V\otimes W$ 由所有的 $x\otimes y$ $(x\in V,y\in W)$ 张成, 由公式 (6.2.24) 直接计算得

$$\begin{aligned}&(\mathcal{A}\otimes\mathcal{B})(\mathcal{C}\otimes\mathcal{D})=\mathcal{AC}\otimes\mathcal{BD},\\&(\mathcal{A}+\mathcal{C})\otimes\mathcal{B}=\mathcal{A}\otimes\mathcal{B}+\mathcal{C}\otimes\mathcal{B},\\&\mathcal{A}\otimes(\mathcal{B}+\mathcal{D})=\mathcal{A}\otimes\mathcal{B}+\mathcal{A}\otimes\mathcal{D},\\&\mathcal{A}\otimes\lambda\mathcal{B}=\lambda\mathcal{A}\otimes\mathcal{B}=\lambda(\mathcal{A}\otimes\mathcal{B}).\end{aligned}$$

细节留给读者验证.

设 $v_1,\cdots,v_n$ 是 V 的基, $w_1,\cdots,w_m$ 是 W 的基. 在 $V\otimes W$ 的基

$$v_1\otimes w_1,\ \cdots,\ v_1\otimes w_m,\ \cdots,\ v_n\otimes w_1,\ \cdots,\ v_n\otimes w_m$$

之下, 算子 $\mathcal{A}\otimes\mathcal{B}$ 的矩阵 $A\otimes B$ 是 nm 阶方阵. 命

$$\mathcal{A}v_j=\sum_i a_{ij}v_i,\quad \mathcal{B}w_l=\sum_k b_{kl}w_k,$$

则有

$$(\mathcal{A}\otimes\mathcal{B})(v_j\otimes w_l)=\sum_{i,j}a_{ij}b_{kl}v_j\otimes w_l.$$

这是说, 由 $A=(a_{ij})$, $B=(b_{kl})$, 可以得到

$$A\otimes B=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{pmatrix}. \tag{6.2.25}$$

由此可知, 对于迹, 我们有公式

$$\mathrm{tr}A\otimes B=a_{11}\mathrm{tr}B+\cdots+a_{nn}\mathrm{tr}B=\mathrm{tr}A\cdot\mathrm{tr}B. \tag{6.2.26}$$

对行列式, 则有

$$\begin{aligned}\det A\otimes B&=\det[(A\otimes E)(E\otimes B)]=\det(A\otimes E)\cdot\det(E\otimes B)\\&=(\det A)^m\cdot(\det B)^n.\end{aligned} \tag{6.2.27}$$

所以, 线性算子 $\mathcal{A}$ 和 $\mathcal{B}$ 的非退化性蕴含它们的张量积 $\mathcal{A}\otimes\mathcal{B}$ 的非退化性.

公式 (6.2.26) 在群的特征标理论中经常使用.

四 可分解张量 设向量组 v_i, $i\in I$ 和向量组 w_j, $j\in J$ 分别是 V 和 W 的基. 那么 $V\otimes W$ 中的任意元素 z 都可以唯一地分解成

$$z=\sum_{i,j}z_{ij}v_i\otimes w_j,\qquad z_{ij}\in K. \tag{6.2.28}$$

纯量 z_{ij} 称为 z 在所给基下的坐标. 在有限维的情形, z 可以通过 $n\times m$ 矩阵 (z_{ij}) 刻画, 此处 $n=\dim V$, $m=\dim W$.

元素 $z\in V\otimes W$ 称为**可分解的**(decomposable) 如果存在 $x\in V$ 和 $y\in W$ 使得

$$z=x\otimes y. \tag{6.2.29}$$

当 $x=\sum\limits_i x_iv_i$, $y=\sum\limits_j y_jw_j$ 时, 有 $z_{ij}=x_iy_j$. 在有限维情形, 这意味着 $\mathrm{rk}\,(z_{ij})\leqslant 1$. 所以可分解元素在 $V\otimes W$ 中是很少的 (除非 V 和 W 中有一个是 1 维的); 不过, 它们张成 $V\otimes W$.

命题 6.9 其实给出了两个有用的分解. 任何元素 $z\in V\otimes W$ 都可以唯一地分解成

$$z=\sum_i v_i\otimes\eta_i,\qquad \eta_i\in W, \tag{6.2.30}$$

也可以唯一地分解成

$$z=\sum_j \xi_j\otimes w_j,\qquad \xi_j\in V. \tag{6.2.31}$$

无疑, 形如式 (6.1.9) 的张量应该称作**可分解张量**. 一般地, 称 V 上的 (p,q) 型张量 T 是**可分解的**如果它能表成如下形式

$$T=\alpha_1\otimes\cdots\otimes\alpha_p\otimes v_1\otimes\cdots\otimes v_q,\quad \alpha_i\in V^*,\ v_j\in V.$$

空间 V 的 p 重张量积 $\underbrace{V\otimes\cdots\otimes V}_{p}$ 常记作 $V^{\otimes p}$. 如果 $p=0$, **约定**$V^{\otimes p}=K$. 下面的结论现在应该是意料之中的了.

定理 6.19 设 V 是有限维向量空间, 那么

$$\mathbb{T}_p^q(V)\simeq\underbrace{V^*\otimes\cdots\otimes V^*}_{p}\otimes\underbrace{V\otimes\cdots\otimes V}_{q}. \tag{6.2.32}$$

证明 对 $p+q$ 用归纳法. 当 $p+q\leqslant 1$ 时结论是显然的. 假设 $p+q=2$, 这儿的结论是例 6.14 的特殊情况. 假设 $p+q\geqslant 3$. 如果 $p\geqslant 2$, 由归纳假设得

$$\mathbb{T}_{p-1}^q(V)\simeq(V^*)^{\otimes(p-1)}\otimes V^{\otimes q}.$$

映射

$$V^*\times\mathbb{T}_{p-1}^q(V)\to\mathbb{T}_p^q(V),\quad (\alpha,f)\to\alpha\otimes f$$

是双线性的, 所以 (定理 6.15) 存在唯一的线性映射

$$\psi: V^*\otimes\mathbb{T}_{p-1}^q(V)\to\mathbb{T}_p^q(V),\quad \alpha\otimes f\to\alpha\otimes f.$$

(这里前一个 $\alpha\otimes f$ 是张量积空间中的元素, 后一个是 (p,q) 型张量.) 显然 ψ 的像包含所有形如式 (6.1.9) 的元素, 它们张成 $\mathbb{T}_p^q(V)$, 所有 ψ 是满射. 由于 $V^*\otimes\mathbb{T}_{p-1}^q(V)$ 和 $\mathbb{T}_p^q(V)$ 的维数都是 $(\dim V)^{p+q}$, 所以 ψ 是同构. 由归纳假设和张量积的结合性知 $\mathbb{T}_p^q(V)\simeq(V^*)^{\otimes p}\otimes V^{\otimes q}$. 对 $q\geqslant 2$ 的情况类似处理. □

注 定理 6.19 对无限维空间也是成立的.

习 题 6.2

1. 设非零向量 $z=x\otimes y\in V\otimes W$ 是可分解的. 证明: 如果 $z=x'\otimes y'$ 是另一个分解, 那么存在 $\lambda\in K^*$ 使得 $x=\lambda x'$, $y=\lambda^{-1}y'$.

2. 证明: 如果空间 V 和 W 的维数都大于 1, 那么 $V\otimes W$ 存在不可分解的向量.

3. 设 V 是域 K 上的向量空间. 证明: $K\otimes V=V=V\otimes K$.

4. 设 $V_1, \cdots, V_k, U$ 是域 K 上的向量空间, k 重线性映射 $V_1 \times \cdots \times V_k \to U$ 全体形成一个向量空间, 记作 $\mathrm{Hom}\,(V_1, \cdots, V_k; U)$ 或 $\mathcal{L}(V_1, \cdots, V_k; U)$. 证明:

(1) $\mathrm{Hom}\,(V_1, \cdots, V_k; U) \simeq \mathrm{Hom}\,(V_1 \otimes \cdots \otimes V_k; U)$;

(2) $\mathrm{Hom}\,(V_1, \cdots, V_k; U) \simeq V_1^* \otimes \cdots \otimes V_k^* \otimes U$;

(3) $V_1^* \otimes \cdots \otimes V_k^* \simeq (V_1 \otimes \cdots \otimes V_k)^*$.

5. 设 V 是域 K 上的 n 维向量空间, f 是 V 上的非退化对称双线性型.

(1) 证明: 对 V 的每一个基 $v_1, \cdots, v_n$, 存在一个基 $v_1', \cdots, v_n'$ 使得 $f(v_i, v_j') = \delta_{ij}$. 这个基称为 (v_i) 的 f- 对偶基.

(2) 记号同 (1). 证明 $V \otimes V$ 中的元素 $\sum\limits_{i=1}^{n} v_i \otimes v_i'$ 与基的选取无关, 即对任何基, 这个元素都是一样的. 它称为**卡西米尔元素**(Casimir element).

6. 把分块矩阵

$$\begin{pmatrix} b_{11}A & b_{12}A & \cdots & b_{1m}A \\ b_{21}A & b_{22}A & \cdots & b_{2m}A \\ \vdots & \vdots & & \vdots \\ b_{m1}A & b_{m2}A & \cdots & b_{mm}A \end{pmatrix}$$

与分块矩阵 (6.2.25) 加以比较.

7. 假设 $K = \mathbb{C}$ 是复数域. 利用复矩阵 A, B 可三角化的性质证明等式 (6.2.27).

8. 证明: 如果 $\mathcal{A}$ 是可对角化的, 那么 $\mathcal{A} \otimes \mathcal{A}$ 也是可对角化的.

9. 假设空间 V 是有限维的. 证明 $(\mathbb{T}_p^q(V))^*$ 与 $\mathbb{T}_q^p(V)$ 同构.

6.3 张量的收缩、对称化与交错化、张量代数

本节 V 始终是 n 维向量空间. 通过定理 6.19 我们等同 $\mathbb{T}_p^q(V)$ 和 $(V^*)^{\otimes p} \otimes V^{\otimes q}$. 特别, 有 (记号 Hom 含义参见 1.7 节第一部分最后一段)

$$\mathbb{T}_p^0(V) = \mathrm{Hom}\,(\underbrace{V, \cdots, V}_{p}; K) = \underbrace{V^* \otimes \cdots \otimes V^*}_{p}; \tag{6.3.33}$$

$$\mathbb{T}_0^q(V) = \mathrm{Hom}\,(\underbrace{V^*, \cdots, V^*}_{q}; K) = \underbrace{V \otimes \cdots \otimes V}_{q}. \tag{6.3.34}$$

一 张量乘积给出一个双线性映射

$$\otimes : \mathbb{T}_p^q(V) \times \mathbb{T}_r^s(V) \to \mathbb{T}_{p+r}^{q+s}(V),$$

使得

$$\begin{aligned}&(\alpha_1 \otimes \cdots \otimes \alpha_p \otimes v_1 \otimes \cdots \otimes v_q) \otimes (\alpha_{p+1} \otimes \cdots \otimes \alpha_{p+r} \otimes v_{q+1} \otimes \cdots \otimes v_{q+s})\\ =&\alpha_1 \otimes \cdots \otimes \alpha_{p+r} \otimes v_1 \otimes \cdots \otimes v_{q+s}.\end{aligned}$$

例 6.20　空间

$$\mathbb{T}_2^2(V) = V^* \otimes V^* \otimes V \otimes V = (V \otimes V)^* \otimes (V \otimes V)$$

可与 Hom $(V \otimes V, V \otimes V)$ 等同起来 (参见 (6.2.20)). 在这个角度下, 张量乘积

$$\mathbb{T}_1^1(V) \times \mathbb{T}_1^1(V) \to \mathbb{T}_2^2(V)$$

和线性算子的张量乘积 (见 6.2 节第三部分) 是一致的. 的确, 双线性意味着仅需对可分解的线性算子验证这个一致性. 命 $\mathcal{A} = \alpha \otimes u$, $\mathcal{B} = \beta \otimes v$, $\alpha, \beta \in V^*$, $u, v \in V$. 根据线性算子的张量积的定义 (6.2.24) 和线性算子与 (1,1) 型张量的联系 (6.1.2), 对 $x, y \in V$, 有

$$\begin{aligned}(\mathcal{A} \otimes \mathcal{B})(x \otimes y) &= \mathcal{A}x \otimes \mathcal{B}y = \alpha(x)\beta(y)u \otimes v \\ &= [(\alpha \otimes \beta)(x \otimes y)]u \otimes v = [(\alpha \otimes \beta) \otimes (u \otimes v)](x \otimes y).\end{aligned}$$

所以

$$\mathcal{A} \otimes \mathcal{B} = \alpha \otimes \beta \otimes u \otimes v.$$

二　收缩　张量中另一个重要的运算是**收缩 (映射)**(contraction), 也称为**卷积**(convolution). 它是一个线性映射

$$\mathbb{T}_p^q(V) \to \mathbb{T}_{p-1}^{q-1}(V), \quad p, q > 0,$$

定义如下. 考虑映射

$$\underbrace{V^* \times \cdots \times V^*}_{p} \times \underbrace{V \times \cdots \times V}_{q} \longrightarrow \mathbb{T}_{p-1}^{q-1}(V),$$
$$(\alpha_1, \cdots, \alpha_p, v_1, \cdots, v_q) \to (\alpha_1, v_1)\alpha_2 \otimes \cdots \otimes \alpha_p \otimes v_2 \otimes \cdots \otimes v_q.$$

它是多重线性的, 所以有线性映射

$$\mathrm{tr}_1^1 : \mathbb{T}_p^q(V) \to \mathbb{T}_{p-1}^{q-1}(V) \tag{6.3.35}$$

使得

$$\alpha_1 \otimes \cdots \otimes \alpha_p \otimes v_1 \otimes \cdots \otimes v_q \to (\alpha_1, v_1)\alpha_2 \otimes \cdots \otimes \alpha_p \otimes v_2 \otimes \cdots \otimes v_q.$$

线性映射 tr_1^1 称为一个**收缩 (映射)**.

类似可定义收缩 (映射)

$$\mathrm{tr}_r^s : \mathbb{T}_p^q(V) \to \mathbb{T}_{p-1}^{q-1}(V), \tag{6.3.36}$$

使得

$$\alpha_1 \otimes \cdots \otimes \alpha_p \otimes v_1 \otimes \cdots \otimes v_q \to (\alpha_r, v_s)\alpha_1 \otimes \cdots \otimes \hat{\alpha}_r \otimes \cdots \otimes \hat{v}_s \otimes \cdots \otimes v_q,$$

记号上有 ˆ 的表示这个记号在张量积中不出现. 它称为按共变指标 r 和反变指标 s 的收缩.

三　张量收缩的坐标形式　对张量 $T \in \mathrm{T}_p^q(V)$, 考察它的收缩 $\mathrm{tr}_r^s(T)$ 的坐标形式是很有意思的. 设 (参见式 (6.1.11))

$$T = \sum_{i,j} T_{i_1\cdots i_p}^{j_1\cdots j_q} e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q}.$$

那么

$$\begin{aligned}\tilde{T} = \mathrm{tr}_r^s(T) &= \sum_{i,j} T_{i_1\cdots i_p}^{j_1\cdots j_q}\, \mathrm{tr}_r^s(e^{i_1} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes e_{j_q}) \\ &= \sum_{i,j} T_{i_1\cdots i_p}^{j_1\cdots j_q}\, \delta_{j_s}^{i_r} e^{i_1} \otimes \cdots \otimes \widehat{e^{i_r}} \otimes \cdots \otimes \widehat{e_{j_s}} \otimes \cdots \otimes e_{j_q} \\ &= \sum_{i,j} \tilde{T}_{i_1\cdots \hat{i}_r \cdots i_p}^{j_1\cdots \hat{j}_s \cdots j_q} e^{i_1} \otimes \cdots \otimes \widehat{e^{i_r}} \otimes \cdots \otimes \widehat{e_{j_s}} \otimes \cdots \otimes e_{j_q},\end{aligned}$$

其中

$$\tilde{T}_{i_1\cdots \hat{i}_r \cdots i_p}^{j_1\cdots \hat{j}_s \cdots j_q} = \tilde{T}_{i_1\cdots i_{r-1} i_{r+1}\cdots i_p}^{j_1\cdots j_{s-1} j_{s+1}\cdots j_q} = \sum_k T_{i_1\cdots i_{r-1} k i_{r+1}\cdots i_p}^{j_1\cdots j_{s-1} k j_{s+1}\cdots j_q}. \tag{6.3.37}$$

我们知道, 可分解张量

$$e^{i_1} \otimes \cdots \otimes \widehat{e^{i_r}} \otimes \cdots \otimes e^{i_p} \otimes e_{j_1} \otimes \cdots \otimes \widehat{e_{j_s}} \otimes \cdots \otimes e_{j_q}$$

全体构成 $\mathrm{T}_{p-1}^{q-1}(V)$ 的一个基. 因此 $\tilde{T}_{i_1\cdots i_{r-1} i_{r+1}\cdots i_p}^{j_1\cdots j_{s-1} j_{s+1}\cdots j_q}$ 是张量 $\tilde{T}$ 的坐标. 我们已经证明了如下结论.

定理 6.21　(p,q) 型混合张量 T 共变指标 r 和反变指标 s 的收缩是 $(p-1,q-1)$ 型张量 $\tilde{T}$, 它的坐标由公式 (6.3.37) 确定.

对一个张量 T, 连续做 $m = \min\{p,q\}$ 次收缩, 最后会得到一个纯量, 或者是共变张量, 或者是反变张量. 最终得到的张量称为原来张量的**完全收缩**. 张量的收缩把很多概念如算子的迹, 矩阵的乘法等统一起来.

例 6.22　沿用例 6.8 的记号. 设 $\mathcal{F}$ 是 V 上的线性算子, $e_1, \cdots, e_n$ 是 V 的基, $\mathcal{F}e_k = \sum_i f_k^i e_i$. 它的张量形式是

$$T = \sum_{i,j} f_j^i e^j \otimes e_i.$$

张量 T 是 (1,1) 型的, 它的收缩 (0,0) 型张量, 即是纯量. 张量 T 的坐标是 f^i_j. 按公式 (6.3.37), T 的收缩是纯量

$$\bar{T} = \sum_k f^k_k.$$

它正是算子 $\mathcal{F}$ 的迹 $\mathrm{tr}(\mathcal{F})$. 这说明算子的迹只是张量收缩的一个特殊情况.

如果 $\alpha \in V^*$, $x \in V$, 根据例 6.13, (1,1) 型张量 $\alpha \otimes x$ 对应的线性算子是 $\mathcal{A}: v \to \alpha(v)x$. 刚才的讨论表明算子 $\mathcal{A}$ 的迹就是张量 $\alpha \otimes x$ 的收缩 $\alpha(x)$.

例 6.23　对 $\alpha \in V^*$, $u, v \in V$, 考虑 (1,2) 型张量 $\alpha \otimes u \otimes v$ 按共变指标 1 和反变指标 2 收缩, 得到向量 $\alpha(v)u$. 它正是算子 $\mathcal{A} = \alpha \otimes u$ 在 v 上的作用 $\mathcal{A}v$, 参见例 6.13.

例 6.24　两个线性算子的乘积也是张量收缩的一个特殊情况. 设 $\mathcal{A}$ 和 $\mathcal{B}$ 是 V 上的两个线性算子, 它们在某个基下的矩阵分别是 $A = (a^i_j)$ 和 $B = (b^i_j)$. 于是, 作为 (1,1) 型张量, 它们的坐标分别是

$$T_A = a^i_j, \qquad T_B = b^i_j.$$

张量积的坐标是

$$T = T_A \otimes T_B = T^{ik}_{jl} = a^i_j b^k_l.$$

按张量 T_A 的共变指标和张量 T_B 的反变指标收缩, 得到张量 $T_C = c^i_l$, 其中

$$c^i_l = \sum_j T^{ij}_{jl} = \sum_j a^i_j b^j_l.$$

容易看出, c^i_l 是矩阵 AB 在 (i, l) 处的值. 所以 $\mathcal{AB}$ 是 $\mathcal{A} \otimes \mathcal{B}$ 的一个收缩.

四　代数的结构张量　设 R 是域 K 上的一个有限维代数 R(见定义 2.20). 这里的乘法运算 $(a, b) \to a * b$ 可以不满足结合律. 如果 $e_1, \cdots, e_n$ 是 R 的一个基, 那么

$$e_i * e_j = \sum_k \gamma^k_{ij} e_k.$$

定义 6.25　纯量 $\gamma^k_{ij} \in K$ 称为代数 R 在给定基下的**结构常数**.

结构常数的意义在于它们完全确定了代数 R 的乘法, 这由乘法的双线性性质推出. 而且, 结构常数还构成一个 (2,1) 型张量 (的坐标). 要看出这一点, 需要考察在不同基下结构常数的联系. 设 $e'_1, \cdots, e'_n$ 是 R 的另一个基, 且

$$e'_i * e'_j = \sum_k \gamma'^k_{ij} e'_k.$$

又设

$$e'_i = \sum_s a^s_i e_s, \quad e_j = \sum_t b^t_j e'_t,$$

则 $A=(a_i^s)$, $B=(b_j^t)=A^{-1}$. 于是

$$\sum_k \gamma_{ij}^k e_k = e_i * e_j = \left(\sum_s b_i^s e'_s\right) * \left(\sum_t b_j^t e'_t\right)$$
$$= \sum_{s,t} b_i^s b_j^t e'_s * e'_t = \sum_{s,t,r} a_i^s a_j^t \gamma'^{r}_{st} e'_r = \sum_{s,t,r,k} b_i^s b_j^t \gamma'^{r}_{st} a_r^k e_k,$$

由此可得

$$\gamma_{ij}^k = \sum_{s,t,r} b_i^s b_j^t \gamma'^{r}_{st} a_r^k.$$

比较公式 (6.1.14) 知结构常数 γ_{ij}^k 确实构成一个 (2,1) 型张量, 称为 R 的**结构张量**. 代数 R 的结构张量完全确定了 R 的乘法, 反之亦然.

对固定的元素 $u\in R$, 映射 $L_u: x\to u*x$ 是 R 上的线性算子. 对称双线性型

$$f(u,v)=\mathrm{tr}(L_uL_v)$$

是研究 R 的代数结构的重要工具. 它的双线性性来自运算 $*$ 的双线性性, 对称性则来自 $\mathrm{tr}L_uL_v=\mathrm{tr}L_vL_u$(参见公式 (2.2.11)).

这个对称双线性型可以通过结构张量的收缩表达. 记

$$u=\sum_i \alpha^i e_i,\quad v=\sum_i \beta^i e_i,$$
$$L_uL_ve_k = u*(v*e_k) = \sum_{i,j}\alpha^i\beta^j e_i*(e_j*e_k) = \sum_{i,j,s,t}\alpha^i\beta^j\gamma_{jk}^s\gamma_{is}^t e_t.$$

要计算 $f(u,v)$ 需要取矩阵

$$(H_k^t),\quad H_k^t=\sum_{i,j,s}\alpha^i\beta^j\gamma_{jk}^s\gamma_{is}^t$$

对角线上的元素并把它们加起来:

$$f(u,v)=\sum_{i,j,s,k}\alpha^i\beta^j\gamma_{jk}^s\gamma_{is}^k. \tag{6.3.38}$$

这是 (4,4) 型张量 $\alpha^i\beta^j\gamma_{pq}^k\gamma_{rs}^l$ 的一个完全收缩.

五　对称张量　在双线性型的理论中, 我们主要讨论了对称的和斜对称的. 对一般的张量

$$T: V^p\times (V^*)^q\to K$$

也考虑它的对称性. 显然, 对称性只能在同类的变量中考虑, 不同类型的变量之间考虑对称性一般是没有意义的. 任何 (p,q) 型张量可以分解成 $(p,0)$ 型张量和 $(0,q)$

型张量乘积的线性组合. 这样一来, 不妨仅对共变张量或反变张量讨论对称性. 为方便起见, 以下假设域 K 的特征是 0.

设 $T \in \mathbb{T}_p^0(V)$ 是 $(p,0)$ 型共变张量, 即它是一个 p 重线性函数

$$T : V^p \to K.$$

对任意置换 $\pi \in S_p$, 命

$$f_\pi(T)(v_1, \cdots, v_p) = T(v_{\pi 1}, \cdots, v_{\pi p}) \tag{6.3.39}$$

因为 T 是多重线性的, 所以 $f_\pi(T)$ 也是多重线性的, 从而是 $(p,0)$ 型张量. 事实上, 对某个指标, 比如说 p, 有 $\pi k = p$, 那么

$$\begin{aligned} f_\pi(T)(v_1, \cdots, v_{p-1}, av_p + bu_p) &= T(v_{\pi 1}, \cdots, av_p + bu_p, \cdots, v_{\pi p}) \\ &= aT(v_{\pi 1}, \cdots, v_p, \cdots, v_{\pi p}) + bT(v_{\pi 1}, \cdots, u_p, \cdots, v_{\pi p}) \\ &= af_\pi(T)(v_1, \cdots, v_{p-1}, v_p) + bf_\pi(T)(v_1, \cdots, v_{p-1}, u_p). \end{aligned}$$

比较对称双线性型的定义, 对称函数和对称多项式的定义, 就会感到下面的定义是一脉相承的.

定义 6.26 称张量 $T \in \mathbb{T}_p^0(V)$ 是**对称的**如果对任意的置换 $\pi \in S_p$ 有 $f_\pi(T) = T$. 类似地, 可以定义对称的反变张量.

容易看出,

$$f_\pi(aT' + bT'') = af_\pi(T') + bf_\pi(T''), \tag{6.3.40}$$

所以 $\mathbb{T}_p^0(V)$ 中的对称张量全体形成一个子空间, 记作 $\mathbb{T}_p^+(V)$. 同样, $\mathbb{T}_0^q(V)$ 中的对称张量全体形成一个子空间, 记作 $\mathbb{T}_+^q(V)$.

另一方面, 公式 (6.3.40) 表明 $\pi \in S_p$ 引出一个线性算子 $f_\pi : \mathbb{T}_p^0(V) \to \mathbb{T}_p^0(V)$. 当 $\pi = e$ 是恒等置换时, f_π 是恒等算子. 直接验证可知 (参见第一卷引理 2.33)$f_\sigma \circ f_\tau = f_{\sigma\tau}$. 于是, 群 S_p 通过诸算子 f_π 作用在 $\mathbb{T}_p^0(V)$ 上. 从而, 有标准的方法从一般的共变张量 (或反变张量) 得到对称张量.

定义 6.27 称映射

$$\mathscr{S} = \frac{1}{p!} \sum_{\pi \in S_p} f_\pi : \mathbb{T}_p^0(V) \to \mathbb{T}_p^0(V) \tag{6.3.41}$$

为 $\mathbb{T}_p^0(V)$ 的**对称化算子**. 对张量 $T \in \mathbb{T}_p^0(V)$, 称 $\mathscr{S}(T)$ 为 T 的对称化.

每个 $(p,0)$ 型张量 $T \in \mathbb{T}_p^0(V)$ 经过对称化后是对称的:

$$f_\sigma(\mathscr{S}(T)) = \frac{1}{p!} \sum_{\pi \in S_p} f_\sigma(f_\pi(T)) = \frac{1}{p!} \sum_{\pi \in S_p} f_{\sigma\pi}(T) = \frac{1}{p!} \sum_{\tau \in S_p} f_\tau(T) = \mathscr{S}(T).$$

也就是说 $\operatorname{Im}\mathscr{S}\subset\mathbb{T}_p^+(V)$. 此处的计算用到如下事实, 对固定的置换 $\sigma\in S_p$, 当 π 取遍群 S_p 中所有元素时, $\sigma\pi$ 也取遍 S_p 中所有元素, 而且每个元素只取一次.

反过来, 从定义可以看出, 在对称张量实施对称化算子得到的就是原来的对称张量, 即 $T\in\mathbb{T}_p^+(V)\Rightarrow\mathscr{S}(T)=T$. 可见有

定理 6.28　对称化算子 $\mathscr{S}$ 在 $\mathbb{T}_p^0(V)$ 上的作用具有性质 $\mathscr{S}^2=\mathscr{S}$ 且 $\operatorname{Im}\mathscr{S}=\mathbb{T}_p^+(V)$.

六　通过张量的坐标计算线性算子 f_π 的作用是很方便的. 设 $e_1,\cdots,e_n$ 是 V 的基,

$$T=\sum_{i_1,\cdots,i_p}T_{i_1\cdots i_p}e^{i_1}\otimes\cdots\otimes e^{i_p}\tag{6.3.42}$$

是 V 上的 $(p,0)$ 型张量. 它的坐标为 $T_{i_1\cdots i_p}=T(e_{i_1},\cdots,e_{i_p})$. 于是, 张量 $f_\pi(T)$ 在同样的基下的坐标是

$$f_\pi(T)_{i_1\cdots i_p}=f_\pi(T)(e_{i_1},\cdots,e_{i_p})=T(e_{i_{\pi 1}},\cdots,e_{i_{\pi p}})=T_{i_{\pi 1}\cdots i_{\pi p}}.$$

这意味着

$$f_\pi(T)=\sum_{i_1,\cdots,i_p}T_{i_{\pi 1}\cdots i_{\pi p}}e^{i_1}\otimes\cdots\otimes e^{i_p}.$$

它的一个等价形式是

$$f_\pi(T)=\sum_{i_1,\cdots,i_p}T_{i_1\cdots i_p}e^{i_{\pi^{-1}1}}\otimes\cdots\otimes e^{i_{\pi^{-1}p}}.\tag{6.3.43}$$

不用说, 对反变张量 $S\in T_0^p(V)$, 有类似的公式:

$$f_\pi(S)=\sum_{i_1,\cdots,i_p}S^{i_{\pi 1}\cdots i_{\pi p}}e_{i_1}\otimes\cdots\otimes e_{i_p}.$$

它等价于

$$f_\pi(S)=\sum_{i_1,\cdots,i_p}S^{i_1\cdots i_p}e_{i_{\pi^{-1}1}}\otimes\cdots\otimes e_{i_{\pi^{-1}p}}.\tag{6.3.44}$$

例 6.29　张量 $e^1\otimes e^2\otimes e^1$ 的对称化是

$$\mathscr{S}(e^1\otimes e^2\otimes e^1)=\frac{1}{3}(e^1\otimes e^2\otimes e^1+e^1\otimes e^1\otimes e^2+e^2\otimes e^1\otimes e^1).$$

在物理学中很多张量是对称的, 如惯性张量、应力张量等.

七　齐次函数　双线性型给出二次函数, 多重线性函数则给出一般的齐次函数.

定义 6.30 向量空间 V 上的函数 $Q: V \to K$ 称为 p**次齐次函数**, 如果

$$Q(x) = F(x, \cdots, x),$$

其中 $F: V^p \to K$ 是 V 上的 p 重线性函数.

把对称化算子作用到 p 重线性函数 F 上, 得到一个对称的 p 重线性函数 $\mathscr{S}(F)$:

$$\mathscr{S}(F)(v_1, \cdots, v_n) = \frac{1}{p!} \sum_{\pi \in S_p} F(v_{\pi 1}, \cdots, v_{\pi p}).$$

明显地,

$$Q(x) = (\mathscr{S}(F))(x, \cdots, x). \tag{6.3.45}$$

称对称的 p 重线性函数 G 为 p 次齐次函数 Q 的一个**极化**(polarization) 如果 $Q(x) = G(x, \cdots, x)$.

定理 6.31 向量空间 V 上的齐次函数的极化存在且唯一.

证明 公式 (6.3.45) 表明齐次函数的极化存在. 下证唯一性. 设 Q 是 p 次齐次函数, F 是对称的 p 重线性函数 (即对称的 $(p,0)$ 型张量) 使得

$$Q(x) = F(x, \cdots, x).$$

设 $e_1, \cdots, e_n$ 是 V 的基, $x = x_1 e_1 + \cdots + x_n e_n$,

$$F = \sum_{i_1, \cdots, i_p} F_{i_1 \cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p}.$$

则有

$$Q(x) = F(x, \cdots, x) = \sum_{i_1, \cdots, i_p} F_{i_1 \cdots i_p} x_{i_1} \cdots x_{i_p}.$$

由于 $x_{i_1} \cdots x_{i_p} = x_{j_1} \cdots x_{j_p}$ 当且仅当 $j_1 \cdots j_p$ 是 $i_1 \cdots i_p$ 的一个置换, 即存在 $\pi \in S_p$ 使得 $j_1 \cdots j_p = i_{\pi 1} \cdots i_{\pi p}$, 而 F 是对称的, 即 $F_{i_1 \cdots i_p} = F_{i_{\pi 1} \cdots i_{\pi p}}$, 所以 Q 的系数和 F 的系数互相确定. 更确切地说, 如果

$$Q(x) = \sum_{i_1 \leqslant \cdots \leqslant i_p} f_{i_1 \cdots i_p} x_{i_1} \cdots x_{i_p},$$

那么

$$f_{i_1, \cdots, i_p} = c \cdot F_{i_1 \cdots i_p},$$

其中 $c = c(i_1, \cdots, i_p) \in \mathbb{N}$ 是指标列 $i_1 \cdots i_p$ 所有不同的排列的个数. 例如, 对 $p = 4$, $c(i,j,k,l) = 24$, $c(i,i,i,j) = 4$, 等等. 唯一性得证. □

这个定理实际上建立了对称张量空间 $\mathbb{T}_p^+(V)$ 和 V 上的 p 次齐次函数全体形成的空间 $S^p(V)$ 之间的线性同构. 取定 V 的一个基后, 齐次函数就是坐标的多项式, 所以 $S^p(V)$ 与 K 上的 p 次齐次多项式全体张成的空间 $K[x_1, \cdots, x_n]_p$ 是线性同构的. 于是, $\mathbb{T}_p^+(V)$ 与 $K[x_1, \cdots, x_n]_p$ 是线性同构的. 对空间 $\mathbb{T}_+^p(V)$, 有同样的结论. 在这些同构中应该注意到

$$\dim K[x_1, \cdots, x_n]_p = \binom{n+p-1}{p}.$$

例 6.32　多项式

$$Q(x) = x_1^2 x_2 + x_3^3$$

的极化是对称的 3 重线性函数

$$F(x, y, z) = \frac{1}{3}(x_1 y_2 z_2 + x_2 y_1 z_2 + x_2 y_2 z_1) + x_3 y_3 z_3.$$

(此处 x, y, z 是三维向量空间中以 x_i, y_i, z_i $(i = 1, 2, 3)$ 为坐标的向量.)

八　斜对称张量　和双线性型一样, 可以讨论斜对称的共变张量或反变张量. 置换 $\pi \in S_p$ 通过公式 (6.3.39) 或 (6.3.43) 作用在张量 $T \in \mathbb{T}_p^0(V)$ 上; 通过公式 (6.3.44) 作用在张量 $S \in \mathbb{T}_0^p(V)$ 上 (也可以类似于式 (6.3.39) 那样定义作用).

定义 6.33　称张量 T 是**斜对称的**或**反对称的**如果

$$f_\pi(T) = \varepsilon_\pi T, \quad \forall \pi \in S_p, \tag{6.3.46}$$

其中 ε_π 是置换 π 的 (奇偶性) 符号.

诸君应该还记得: $\varepsilon : \pi \to \varepsilon_\pi$ 是群 S_p 到 $\{\pm 1\}$ 的同态且任何对换 τ 的符号 ε_τ 等于 -1. 由于置换可以分解成对换的乘积, $f_{\sigma\pi}(T) = f_\sigma(f_\pi(T))$, 所以条件 (6.3.46) 等价于

$$f_\tau(T) = -T, \qquad \forall \text{ 对换 } \tau = (i\, j) \in S_p. \tag{6.3.47}$$

上式的含义是

$$T(\cdots, v_i, \cdots, v_j, \cdots) = -T(\cdots, v_j, \cdots, v_i, \cdots), \qquad \forall v_i, v_j \in V, \tag{6.3.48}$$

其中省略号中的向量可以是任意的向量, 但等式两侧相同位置的被省略的向量是相等的. 由于 $\operatorname{char} K = 0$, 当 $v_i = v_j = v$ 时, 由式 (6.3.48) 得

$$T(\cdots, v, \cdots, v, \cdots) = 0. \tag{6.3.49}$$

在式 (6.3.49) 中令 $v = v_i + v_j$, 利用 T 的多重线性性质, 得

$$\begin{aligned}0 &= T(\cdots, v_i + v_j, \cdots, v_i + v_j, \cdots)\\&= T(\cdots, v_i, \cdots, v_i, \cdots) + T(\cdots, v_j, \cdots, v_j, \cdots)\\&\quad + T(\cdots, v_i, \cdots, v_j, \cdots) + T(\cdots, v_j, \cdots, v_i, \cdots)\\&= T(\cdots, v_i, \cdots, v_j, \cdots) + T(\cdots, v_j, \cdots, v_i, \cdots).\end{aligned}$$

可见, 式 (6.3.49) 蕴含式 (6.3.48). 所以, 这两个条件等价.

张量 T 的斜对称性用坐标表述更容易看出内在的特性. 当 $p = 2$ 时, 斜对称意味着 $T_{ij} = -T_{ji}$ (即坐标矩阵的反对称性). 同时, 这个性质与基的选取无关, 这是我们看重的. 在一般情形,

$$T_{i_{\pi 1} \cdots i_{\pi p}} = \varepsilon_\pi T_{i_1 \cdots i_p}.$$

这意味着在坐标的下指标 $i_{\pi 1}, \cdots, i_{\pi p}$ 中如果有相同的, 那么, 这个坐标值为 0. 如果这些下指标互不相同, 那么它们不同的排列对应的坐标值至多差一个负号, 而且互相确定. 从而, 张量 T 完全由如下的坐标

$$T_{i_1 i_2 \cdots i_p}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n \tag{6.3.50}$$

确定. 这样的坐标共有 $\binom{n}{p}$ 个, 它们是相互独立的. 上面的讨论可以总结成如下的结论.

命题 6.34　张量 $T \in \mathrm{T}_p^0(V)$ 是斜对称的当且仅当它可以写成如下形式:

$$T = \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} T_{i_1 \cdots i_p} \left(\sum_{\pi \in S_p} \varepsilon_\pi e^{i_{\pi 1}} \otimes \cdots \otimes e^{i_{\pi p}} \right). \tag{6.3.51}$$

如果 $f_\pi T = \varepsilon_\pi T$, $f_\pi S = \varepsilon_\pi S$, 那么 $f_\pi(aT + bS) = a f_\pi T + b f_\pi S = a\varepsilon_\pi T + b\varepsilon_\pi S = \varepsilon_\pi(aT + bS)$. 由此可见, 空间 $\mathrm{T}_p^0(V)$ 中的斜对称张量全体形成一个子空间, 记作 $\Lambda^p(V^*)$. 类似地, $\mathrm{T}_0^q(V)$ 中的斜对称张量全体形成一个子空间, 记作 $\Lambda^q(V)$.

为从一般的 $(p, 0)$ 型张量得到斜对称张量, 我们引入交错化算子.

定义 6.35　映射

$$\mathscr{A} = \frac{1}{p!} \sum_{\pi \in S_p} \varepsilon_\pi f_\pi : \mathrm{T}_p^0(V) \to \mathrm{T}_p^0(V) \tag{6.3.52}$$

称为**交错化算子**, 也称为**斜对称化算子**.

定理 6.36　交错化算子 $\mathscr{A}$ 具有如下性质:

(1) $\mathscr{A}^2 = \mathscr{A}$;

(2) $\operatorname{Im}\mathscr{A} = \Lambda^p(V^*)$;

(3) $\mathscr{A}(f_\sigma(T)) = \varepsilon_\sigma \mathscr{A}(T)$.

证明 (1) 根据公式 (6.3.52), 有

$$\mathscr{A}^2 = \frac{1}{p!^2}\sum_{\sigma,\pi\in S_p}\varepsilon_\sigma\varepsilon_\pi f_\sigma\circ f_\pi = \frac{1}{p!^2}\sum_{\sigma,\pi\in S_p}\varepsilon_{\sigma\pi}f_{\sigma\pi} = \frac{1}{p!}\sum_{\rho\in S_p}\varepsilon_\rho f_\rho = \mathscr{A}.$$

此处用到一个事实, 任意元素 $\rho\in S_p$ 都有 $p!$ 种方式表成乘积 $\sigma\pi$: 先任意选取 σ(有 $p!$ 种选法), 然后必须取 $\pi=\sigma^{-1}\rho$. 当然, 还用到 ε 和 f 与 S_p 的乘法的相容性.

(2) 对任意的 $T\in \mathrm{T}^0_p(V)$, 有

$$f_\sigma(\mathscr{A}(T)) = \frac{1}{p!}\sum_{\pi\in S_p}\varepsilon_\pi f_\sigma(f_\pi(T)) = \varepsilon_\sigma\sum_{\pi\in S_p}\varepsilon_{\sigma\pi}f_{\sigma\pi}(T) = \varepsilon_\sigma\mathscr{A}(T).$$

所以, $\operatorname{Im}\mathscr{A}\subset\Lambda(V*)$. 另一方面,

$$T\in\Lambda(V^*)\Rightarrow\mathscr{A}(T) = \frac{1}{p!}\sum_{\pi\in S_p}\varepsilon_\pi f_\pi(T) = \frac{1}{p!}\sum_{\pi\in S_p}\varepsilon_\pi^2 T = \frac{1}{p!}\sum_{\pi\in S_p}T = T.$$

从而, 断言成立.

(3) 在证明 (ii) 的过程中已经验证 $f_\sigma\mathscr{A} = \varepsilon_\sigma\mathscr{A}$. 同样的方式可以说明 $\mathscr{A}f_\sigma = \varepsilon_\sigma\mathscr{A}$. □

九 张量代数 把向量空间 V 上的共变张量 (或反变张量) 合在一起有更丰富的代数结构. 考虑无限的外直和

$$\mathrm{T}(V^*) = \bigoplus_{p=0}^{\infty}\mathrm{T}^0_p(V) = K\oplus\mathrm{T}^0_1(V)\oplus\mathrm{T}^0_2(V)\oplus\cdots, \tag{6.3.53}$$

这个直和的元素可以认为是序列

$$(\xi_0,\xi_1,\xi_2,\cdots) = \sum_{i=0}^{\infty}\xi_i,\quad \xi_i\in\mathrm{T}^0_i(V),$$

它的项几乎全为零 (即不为零的项仅有限个), 加法和纯量乘的规则是

$$\sum_i\xi_i+\sum_i\eta_i = \sum_i(\xi_i+\eta_i),\quad a\sum_i\xi_i = \sum_i a\xi_i,\quad a\in K.$$

重要的是张量积 (见定义 6.2) 给出了 $\mathrm{T}(V^*)$ 上的乘法:

$$\left(\sum_{i=0}^{\infty}\xi_i\right)\otimes\left(\sum_{j=0}^{\infty}\eta_j\right) = \sum_{k=0}^{\infty}\left(\sum_{i+j=0}^{k}\xi_i\otimes\eta_j\right). \tag{6.3.54}$$

(注意上式中的无限和其实都是有限和, 因为不等于零的项仅有有限个.) 自然, 这个乘法满足结合律 (参见式 (6.1.4)), 对加法有分配律 (参见式 (6.1.6)). 容易验证, 乘法还满足条件

$$a(\xi \otimes \eta) = (a\xi) \otimes \eta = \xi \otimes (a\eta), \quad a \in K.$$

可见, $\mathbb{T}(V^*)$ 是 K 上的一个无限维结合代数, 称为 V^* 的**张量代数**, 也称为 V 的**共变张量代数**. 乘法规则 (6.3.54) 令人想起多项式的乘法规则, 差别仅在于这里的非交换性.

在张量代数 $\mathbb{T}(V^*)$ 中由对称张量张成的子空间是

$$S(V^*) = \mathbb{T}^+(V^*) = \bigoplus_{p=0}^{\infty} \mathrm{T}_p^+(V) = K \oplus \mathrm{T}_1^+(V) \oplus \mathrm{T}_2^+(V) \oplus \cdots. \tag{6.3.55}$$

对称张量的乘积不一定是对称的, 简单的例子如 $\dim V = 2$ 时, 张量 e^1 和 e^2 是对称的, 但 $e^1 \otimes e^2$ 不是对称的. 所以 $S(V^*)$ 不是 $\mathbb{T}(V^*)$ 的子代数. 不过, 可以在 $S(V^*)$ 中按公式

$$T_1T_2 = \mathscr{S}(T_1 \otimes T_2), \quad T_1 \in T_p^+(V),\ T_2 \in T_q^+(V),$$

对齐次元素定义乘法, 再线性拓展到一般元素的乘法. 因此, $S(V^*)$ 成为一个交换的结合代数, 称为 V^* 的**对称代数**, 它与域 K 上的 $\dim V$ 个变元的多项式代数同构, $T_p^+(V)$ 对应到 p 次齐次多项式张成的空间.

反变张量代数$\mathbb{T}(V)$ 的定义完全类似:

$$\mathbb{T}(V) = \bigoplus_{p=0}^{\infty} \mathrm{T}_0^p(V) = K \oplus \mathrm{T}_0^1(V) \oplus \mathrm{T}_0^2(V) \oplus \cdots \tag{6.3.56}$$

它也称为 V 的**张量代数**. 对称的反变张量张成的子空间是

$$S(V) = \mathbb{T}_+(V) = \bigoplus_{p=0}^{\infty} \mathrm{T}_0^p = K \oplus \mathrm{T}_+^1(V) \oplus \mathrm{T}_+^2(V) \oplus \cdots. \tag{6.3.57}$$

按公式

$$T_1T_2 = \mathscr{S}(T_1 \otimes T_2), \quad T_1 \in T_+^p(V),\ T_2 \in T_+^q(V),$$

对齐次元素定义乘法, 再线性拓展到一般元素的乘法. 那么, $S(V)$ 成为一个交换的结合代数, 称为 V 的**对称代数**, 它也与域 K 上的 $\dim V$ 个变元的多项式代数同构, $T_+^p(V)$ 对应到 p 次齐次多项式张成的空间.

说明　本节定义的对称化算子和斜对称化算子要求基域 K 的特征为 0. 虽然对称化算子并不能直接在正特征的域上的张量空间上定义对称算子, 但有直接的方式定义向量空间的对称代数.

习 题 6.3

1. **指标的提升与下放**(raising of subscript, lowering of superscript). 对欧几里得向量空间，度量张量可以用于定义张量指标的提升与下放. 如同习题 6.1 第 9 题，设 (V, g) 是欧几里得向量空间，度量张量是 $G_0 = (g_{ij})$ 和 $G^0 = (g^{ij})$. 张量 G^0 与张量 $T \in \mathbb{T}^q_p(V)$ 的乘积 $G^0 \otimes T$ 的如下收缩

$$\tilde{T}^{i_1 j_1 \cdots j_q}_{i_2 \cdots i_p} = \sum_k g^{i_1 k} T^{j_1 \cdots j_q}_{k i_2 \cdots i_p}$$

得到的张量 $\tilde{T} \in \mathbb{T}^{q+1}_{p-1}$ 称为**第一个下指标的提升**(raising of the first subscript). 类似地定义其他下指标的提升.

张量 G_0 与张量 $T \in \mathbb{T}^q_p(V)$ 的乘积 $G_0 \otimes T$ 的如下收缩

$$\tilde{T}^{j_1 j_3 \cdots j_q}_{i_1 i_2 \cdots i_p j_2} = \sum_k g_{j_2 k} T^{j_1 k j_3 \cdots j_q}_{i_1 \cdots i_p}$$

得到的张量 $\tilde{T} \in \mathbb{T}^{q-1}_{p+1}$ 称为**第二个上指标的下放** (lowering of the second subscript). 类似地定义其他上指标的下放.

一般地，提升第 s 个下指标和下放第 t 个上指标分别给出线性映射

$$\mathbb{T}^q_p(V) \to \mathbb{T}^{q+1}_{p-1} \quad 和 \quad \mathbb{T}^q_p(V) \to \mathbb{T}^{q-1}_{p+1}.$$

把这些线性映射应用到 $T \in \mathbb{T}^q_p(V)$ 上得到的张量的坐标可以分别记作

$$\tilde{T}_{i_1 \cdots i_{s-1}}{}^{i_s}{}_{i_{s+1} \cdots i_p}{}^{j_1 \cdots j_q}, \quad \tilde{T}_{i_1 \cdots i_p}{}^{j_1 \cdots j_{t-1}}{}_{j_t}{}^{j_{t+1} \cdots j_q},$$

这里的指标使用了分块排列.

对标准正交基，有 $g_{i_1 k} = \delta_{i_1 k}$, $\quad g^{j_2 k} = \delta_{j_2 k}$, 因此

$$\tilde{T}^{i_1 j_1 \cdots j_q}_{i_2 \cdots i_p} = T^{j_1 \cdots j_q}_{i_1 i_2 \cdots i_p}, \qquad \tilde{T}^{j_1 j_3 \cdots j_q}_{i_1 i_2 \cdots i_p j_2} = T^{j_1 j_2 j_3 \cdots j_q}_{i_1 \cdots i_p}.$$

这意味着，在标准正交基下，张量的上指标和下指标没有差别.

可对一个张量多次应用指标提升和下放. 对度量张量自身也可以用提升和下放算子. 对张量 G^0 和 G_0 分别施行提升和下放算子. 证明，在和式约定下，有

$$g^{ik} g^{jl} g_{kl} = g^{ij}, \qquad g_{ik} g_{jl} g^{kl} = g_{ij}.$$

通过内积 g, 可以等同 V 与 V^*(参见定理 3.19). 那么，对 $x \in V = V^*$, 有

$$x = x^1 e_1 + \cdots + x^n e_n = x_1 e^1 + \cdots + x_n e^n$$

(如往常，$e^1, \cdots, e^n$ 是 $e_1, \cdots, e_n$ 的对偶基). 证明：在和式约定下有 $g^{ik} x_k = x^i$. 这是说向量 x 的反变指标 x^i 可以在同一个向量的共变坐标 x_k 上使用指标提升算子得到.

2. 记号同第 1 题. 对张量 $x \otimes u \in V \otimes V$ 下放 u 的指标，得到 V 的一个线性函数 α_u. 证明：$\alpha_u(x) = (x|u)$, 从而映射 $u \to \alpha_u$ 是定理 3.19 给出的同构 $V \xrightarrow{\sim} V^*$.

3. 设 $\mathcal{A}$ 是欧几里得向量空间 V 的线性算子. 证明: 对 $\mathcal{A}$ 下放指标得到双线性函数

$$\alpha(x,y) = g_{ij}x^i\mathcal{A}^j_k y^k = (x|\mathcal{A}y).$$

4. 假设 $\dim V > 1$. 证明向量空间 $\mathbb{T}^q_0(V)$ 上的对称化算子和斜对称化算子有如下性质:

(1) 它们的核的交等于零如果 $q = 2$, 不等于零如果 $q > 2$;

(2) $\mathscr{S} \cdot \mathscr{A} = \mathscr{A} \cdot \mathscr{S} = 0$;

(3) 算子 $\mathcal{P} = (\mathcal{E} - \mathscr{S})(\mathcal{E} - \mathscr{A})$ 是投影算子;

(4) 如果 $q = 3$ 且 V 是有限维空间, 求 $\mathcal{P}$ 的秩.

5. 设 $\mathcal{A}$ 是 n 维向量空间 V 上的线性算子. 证明:

(1) $\mathbb{T}^2_+(V) \subset V \otimes V$ 是 $\mathcal{A}\otimes\mathcal{A}$ 的不变子空间. 算子 $\mathcal{A}\otimes\mathcal{A}$ 在 $\mathbb{T}^2_+(V)$ 上的限制记作 $\mathscr{S}^2\mathcal{A}$.

(2) 如果 $\mathcal{A}$ 的特征多项式的根是 $\lambda_1, \cdots, \lambda_n$(按重数计算), 那么 $\mathscr{S}^2\mathcal{A}$ 的特征多项式的根是 $\lambda_i\lambda_j$, $1 \leqslant i \leqslant j \leqslant n$.

(3) $\operatorname{tr}\mathscr{S}^2\mathcal{A} = \frac{1}{2}[(\operatorname{tr}\mathcal{A})^2 + \operatorname{tr}\mathcal{A}^2]$.

6. 设 V 和 W 是向量空间. 证明 $\mathbb{T}^q_+(V \oplus W)$ 与 $\displaystyle\bigoplus_{i=0}^{q}\left(\mathbb{T}^i_+(V) \otimes \mathbb{T}^{q-i}_+(W)\right)$ 同构.

7. 假设 $\dim V > 1$. 证明: $\mathbb{T}(V)$ 是无零因子的非交换环, 其中的可逆元只有非零纯量.

6.4 外 代 数

现在, 我们把注意力转向斜对称张量. *本节假设基域 K 的特征为 0, V 是 K 上的有限维向量空间, 除非另有说明.* 首先引进两个术语并回顾两个记号. 空间 V 上的斜对称的 $(p,0)$ 张量称为**p-形式**(p-form, 也称为 p 次外形式, exterior p-form), 它们全体形成的空间记作 $\Lambda^p(V^*)$. 空间 V 上的斜对称的 $(0,p)$ 张量称为**p-向量**(p-vector), 它们全体形成的空间记作 $\Lambda^p(V)$.

一　外积　在张量代数 $\mathbb{T}(V^*)$ 和 $\mathbb{T}(V)$ 中所有斜对称张量张成的空间分别是

$$\Lambda(V^*) = K \oplus \Lambda^1(V^*) \oplus \Lambda^2(V^*) \oplus \cdots \subset \mathbb{T}(V^*) \tag{6.4.58}$$

和

$$\Lambda(V) = K \oplus \Lambda^1(V) \oplus \Lambda^2(V) \oplus \cdots \subset \mathbb{T}(V^*). \tag{6.4.59}$$

注意 $\Lambda^1(V^*) = V^* = \mathbb{T}^0_1(V)$, $\Lambda^1(V) = V = \mathbb{T}^1_0(V)$.

不失一般性, 我们仅讨论斜对称反变张量和它们张成的空间 $\Lambda(V)$. 与对称张量的情形类似, 斜对称张量的乘积一般不是斜对称的, 例如 $\dim V = 2$ 时, 张量 e_1 和 e_2 是斜对称的, 但 $e_1 \otimes e_2$ 不是斜对称的. 同样与对称张量的情形类似, 借助交错算子, 可以对两个斜对称张量定义它们的一个乘积, 称为外积.

定义 6.37　任意的 q-向量 Q 和 r-向量 R 的**外积**(wedge product) 是

$$Q \wedge R := \mathscr{A}(Q \otimes R). \tag{6.4.60}$$

对 $\lambda \in K$, 由于 $\lambda \otimes R = R \otimes \lambda = \lambda R$, 所以有 $\lambda \wedge R = R \wedge \lambda = \lambda R$.

在外积的定义中, $Q \otimes R$ 是 $(0, q+r)$ 张量, 交错化算子 $\mathscr{A}$(定义 6.35 中的 p 在这里是 $q+r$) 作用在 $Q \otimes R$ 上得到的张量 $\mathscr{A}(Q \otimes R)$ 是斜对称的 (参见定理 6.36), 所以公式 (6.4.60) 给出了映射

$$\wedge : \Lambda^q(V) \times \Lambda^r(V) \to \Lambda^{q+r}(V).$$

张量乘积是双线性的, 交错化算子是线性的, 这蕴含外积是双线性的. 也就是说, 对 $T \in \Lambda^r(V)$, $S \in \Lambda^q(V)$, 以及 $a, b \in K$, 有

$$\begin{aligned} Q \wedge (aR + bT) &= \mathscr{A}(Q \otimes (aR + bT)) = \mathscr{A}(aQ \otimes R + bQ \otimes T) \\ &= a\mathscr{A}(Q \otimes R) + b\mathscr{A}(Q \otimes T) = a(Q \wedge R) + b(Q \wedge T); \end{aligned}$$

以及

$$(aQ + bS) \wedge R = a(Q \wedge R) + b(S \wedge R).$$

特别, 外积对张量的加法有分配律. 外积可以很自然地延拓成整个空间 $\Lambda(V)$ 的外积. 对 $\Lambda(V)$ 中任意两个元素 Q' 和 R', 有唯一的分解

$$Q' = \sum_{i \geqslant 0} Q_i, \quad R' = \sum_{j \geqslant 0} R_j; \quad Q_i \in \Lambda^i(V), \quad R_j \in \Lambda^j(V).$$

令

$$Q' \wedge R' = \sum_{i,j \geqslant 0} Q_i \wedge R_j.$$

从而, 我们得到了 $\Lambda(V)$ 的外积运算

$$\wedge : \Lambda(V) \times \Lambda(V) \to \Lambda(V), \quad (Q', R') \to Q' \wedge R'.$$

容易验证, 这个运算是双线性的, 特别对加法有分配律.

二　外代数　外积的双线性赋予 $\Lambda(V)$ 一个 K 上的代数结构.

定义 6.38　域 K 上的代数 $\Lambda(V)$ 称为向量空间 V 的**外代数**(exterior algebra), 也称为**格拉斯曼代数**(Grassmann algebra), 因为格拉斯曼在 1844 年引入这个代数.

外代数 $\Lambda(V)$ 有单位元, 就是 K 中的乘法单位元 1. 外代数特别重要的性质是乘法的结合性, 将由下面的定理 6.40 中给出.

定理 6.39　对任意张量 $Q \in \mathrm{T}_0^q(V)$ 和 $R \in \mathrm{T}_0^r(V)$, 有

$$\mathscr{A}(\mathscr{A}(Q) \otimes R) = \mathscr{A}(Q \otimes \mathscr{A}(R)) = \mathscr{A}(Q \otimes R).$$

证明　按定义, 有

$$\mathscr{A}(R) = \frac{1}{r!}\sum_{\pi\in S_r}\varepsilon_\pi f_\pi(R),$$

以及

$$\mathscr{A}(Q\otimes\mathscr{A}(R)) = \frac{1}{(q+r)!}\sum_{\sigma\in S_{q+r}}\varepsilon_\sigma f_\sigma(Q\otimes\mathscr{A}(R)).$$

由于张量乘积是双线性的, 并且交错化算子是线性的, 得

$$\mathscr{A}(Q\otimes\mathscr{A}(R)) = \frac{1}{r!}\sum_{\pi\in S_r}\varepsilon_\pi\mathscr{A}(Q\otimes f_\pi(R)). \tag{6.4.61}$$

考虑嵌入映射 $\varphi: S_r\to S_{q+r}$. 置换 $\pi\in S_r$ 的像 $\bar\pi=\varphi(\pi)$ 是 S_{q+r} 中的置换, 作用规则为

$$\bar\pi = \begin{cases} i, & \text{如果 } 1\leqslant i\leqslant q,\\ \pi i, & \text{如果 } q+1\leqslant i\leqslant q+r.\end{cases}$$

这样一来, 就有 $Q\otimes f_\pi(R) = f_{\bar\pi}(Q\otimes R)$(参见公式 (6.3.44)). 由定理 6.36 得

$$\mathscr{A}(Q\otimes f_\pi(R)) = \mathscr{A}f_{\bar\pi}(Q\otimes R) = \varepsilon_\pi\mathscr{A}(Q\otimes R).$$

再注意到 $\varepsilon_{\bar\pi}=\varepsilon_\pi$, 就可以把等式 (6.4.61) 化成我们需要的形式

$$\mathscr{A}(Q\otimes\mathscr{A}(R)) = \frac{1}{r!}\sum_{\pi\in S_r}\varepsilon_\pi^2\mathscr{A}(Q\otimes R) = \mathscr{A}(Q\otimes R),$$

因为 $\varepsilon_\pi^2=1$. 类似地, 可以证明 $\mathscr{A}(\mathscr{A}(Q)\otimes R) = \mathscr{A}(Q\otimes R)$. □

定理 6.40　外代数 $\Lambda(V)$ 是结合代数, 即它的乘法运算 $\wedge$ 满足结合律.

证明　需要对 $\Lambda(V)$ 中任意的元素 P,Q,R 验证如下等式

$$(P\wedge Q)\wedge Q = P\wedge(Q\wedge R). \tag{6.4.62}$$

由于外积 $\wedge$ 是双线性的, 所以只需讨论

$$P\in\Lambda^p(V),\quad Q\in\Lambda^q(V),\quad R\in\Lambda^r(V)$$

的情形. 按定义中的公式 (6.4.60), 有

$$(P\wedge Q)\wedge Q = \mathscr{A}(\mathscr{A}(P\otimes Q)\otimes R.$$

运用定理 6.39, 得

$$\mathscr{A}(\mathscr{A}(P\otimes Q)\otimes R) = \mathscr{A}((P\otimes Q)\otimes R).$$

利用张量乘积的结合性, 并再次利用定理 6.39, 得

$$\begin{aligned}\mathscr{A}((P\otimes Q)\otimes R)&=\mathscr{A}(P\otimes(Q\otimes R))\\&=\mathscr{A}(P\otimes\mathscr{A}(Q\otimes R))=P\wedge(Q\wedge R).\end{aligned}$$

把刚才建立的那些等式合起来看, 可知等式 (6.4.62) 成立. □

三 由于外积满足结合律, 所以若干个斜对称张量的外积无论怎样安排括号的位置, 最终得到的张量都是一样的, 这和群中多个元素相乘是一样的, 参见第一卷命题 5.7. 于是, 乘积 $P_1\wedge P_2\wedge\cdots\wedge P_k$ 有明确的意义. 值得指出 1- 向量间的外积运算公式. 对 $x,y\in V=\mathbb{T}_0^1(V)$, 由公式 (6.4.60), 得

$$x\wedge y=\mathscr{A}(x\otimes y)=\frac{1}{2}(x\otimes y-y\otimes x). \tag{6.4.63}$$

可见

$$x\wedge y=-y\wedge x,\quad x\wedge x=0. \tag{6.4.64}$$

更一般地, 有下面的结论.

命题 6.41 (1) 设 $v_1,\cdots,v_k$ 是 V 中的向量, 那么

$$\begin{aligned}v_1\wedge v_2\wedge\cdots\wedge v_k&=\mathscr{A}(v_1\otimes v_2\otimes\cdots\otimes v_k)\\&=\frac{1}{k!}\sum_{\pi\in S_k}\varepsilon_\pi(v_{\pi1}\otimes v_{\pi2}\otimes\cdots\otimes v_{\pi k}).\end{aligned} \tag{6.4.65}$$

(2) 对任意置换 $\pi\in S_k$, 有

$$v_{\pi1}\wedge\cdots\wedge v_{\pi k}=\varepsilon_\pi(v_1\wedge\cdots\wedge v_k). \tag{6.4.66}$$

(3) 如果存在不同的指标 i 和 j 使得 $v_i=v_j$, 那么

$$v_1\wedge v_2\wedge\cdots\wedge v_k=0.$$

证明 (1) 当 $k=1$ 时, 等式是显然的. 当 $k=2$ 时, 要证的等式和式 (6.4.63) 是一致的. 当 $k>2$ 时, 对 k 做归纳法. 由定理 6.39 和定理 6.40 得

$$\begin{aligned}v_1\wedge v_2\wedge\cdots\wedge v_k&=(v_1\wedge v_2\wedge\cdots v_{k-1})\wedge v_k=\mathscr{A}((v_1\wedge v_2\wedge\cdots v_{k-1})\otimes v_k)\\&=\mathscr{A}(\mathscr{A}(v_1\otimes v_2\otimes\cdots\otimes v_{k-1})\otimes v_k)=\mathscr{A}(v_1\otimes v_2\otimes\cdots\otimes v_k).\end{aligned}$$

第一个等式得证. 第二个等式来自交错化算子的定义, 并注意 $\varepsilon_\pi=\varepsilon_{\pi^{-1}}$.

(2) 由 (1) 得

$$\begin{aligned}
v_{\pi 1}\wedge\cdots\wedge v_{\pi k} &= \mathscr{A}(v_{\pi 1}\otimes\cdots\otimes v_{\pi k})\\
&= \mathscr{A} f_{\pi^{-1}}(v_1\otimes\cdots\otimes v_k) \qquad \text{由定义, 参见公式 (6.3.44)}\\
&= \varepsilon_\pi \mathscr{A}((v_1\otimes\cdots\otimes v_k) \qquad \text{利用定理 6.36(3) 和} \varepsilon_\pi=\varepsilon_{\pi^{-1}}\\
&= \varepsilon_\pi (v_1\wedge\cdots\wedge v_k).
\end{aligned}$$

(3) 命 $\pi=(i\,j)\in S_k$. 那么 π 互换 i 和 j, 保持其他指标不变, 且 $\varepsilon_\pi=-1$. 由于 $v_i=v_j$, 所以有

$$v_1\wedge v_2\wedge\cdots\wedge v_k=v_{\pi 1}\wedge v_{\pi 2}\wedge\cdots\wedge v_{\pi k}.$$

由 (2) 得

$$v_{\pi 1}\wedge v_{\pi 2}\wedge\cdots\wedge v_{\pi k}=-(v_1\wedge v_2\wedge\cdots\wedge v_k).$$

因为 K 的特征是 0, 所以上式蕴含 $v_1\wedge v_2\wedge\cdots\wedge v_k=0$. □

现在, 下面的结论是容易想到的.

定理 6.42　设 $e_1,\cdots,e_n$ 是向量空间 V 的基, 那么 p- 向量

$$e_{i_1}\wedge e_{i_2}\wedge\cdots\wedge e_{i_p},\quad 1\leqslant i_1<i_2\cdots<i_p\leqslant n \tag{6.4.67}$$

形成空间 $\Lambda^p(V)$ 的一个基.

证明　根据定理 6.36(2), $\Lambda^p(V)=\mathscr{A}(\mathrm{T}_0^p(V))$. 由于张量

$$e_{j_1}\otimes e_{j_2}\otimes\cdots\otimes e_{j_p},\quad 1\leqslant j_1,j_2,\cdots,j_p\leqslant n,$$

形成 $\mathrm{T}_0^p(V)$ 的基, 交错化算子 $\mathscr{A}$ 是线性的, 所以 $\Lambda^p(V)$ 中的元素都是如下张量

$$e_{j_1}\wedge e_{j_2}\wedge\cdots\wedge e_{j_p}=\mathscr{A}(e_{j_1}\otimes e_{j_2}\otimes\cdots\otimes e_{j_p}),\quad 1\leqslant j_1,j_2,\cdots,j_p\leqslant n,$$

的线性组合.

如果指标 $j_1,j_2,\cdots,j_p$ 中有相同的, 根据命题 6.41(3) 知 $e_{j_1}\wedge e_{j_2}\wedge\cdots\wedge e_{j_p}=0$. 如果这些指标都不相同, 把它们从小到大排列起来: $r_1,r_2,\cdots,r_p$. 根据命题 6.41(2), 得

$$e_{j_1}\wedge e_{j_2}\wedge\cdots\wedge e_{j_p}=\pm e_{r_1}\wedge e_{r_2}\wedge\cdots\wedge e_{r_p}.$$

由此可知, $\Lambda^p(V)$ 中的元素都是形如式 (6.4.67) 的张量的线性组合. 剩下的事情是证明那些张量线性无关.

设

$$\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda^{i_1\cdots i_p}e_{i_1}\wedge\cdots\wedge e_{i_p}=0.$$

由命题 6.41(1) 得

$$\frac{1}{p!}\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda^{i_1\cdots i_p}\sum_{\pi\in S_p}\varepsilon_\pi(e_{i_{\pi 1}}\otimes e_{i_{\pi 2}}\otimes\cdots\otimes e_{i_{\pi p}})=0.\tag{6.4.68}$$

如果 $\pi\neq e$, 那么指标序列 $i_{\pi 1},\cdots,i_{\pi p}$ 就不是从小到大的排列. 于是, 等式 (6.4.68) 可以改写成

$$\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\lambda^{i_1\cdots i_p}e_{i_1}\otimes\cdots\otimes e_{i_p}+T'=0,\tag{6.4.69}$$

其中 T' 是指标序列 $j_1,\cdots,j_p$ 有反序 (即有数对 r,s 使得 $r<s$ 但 $j_r>j_s$) 的张量 $e_{j_1}\otimes\cdots\otimes e_{j_p}$ 的一个线性组合. 由于这些张量 $e_{i_1}\otimes\cdots\otimes e_{i_p}$ 和 $e_{j_1}\otimes\cdots\otimes e_{j_p}$ 全体形成 $\mathbb{T}_0^p(V)$ 的基 (参见定理 6.6), 所以它们线性无关. 由此可知, 等式 (6.4.69) 左边的系数 $\lambda^{i_1\cdots i_p}$ 都等于零. 定理得证. □

本节开始部分的等式 (6.4.59) 告诉我们, $\Lambda(V)$ 是诸子空间 $\Lambda^p(V)$ 的直和, 所以有

$$\dim\Lambda(V)=\sum_{p\geqslant 0}\dim\Lambda^p(V).$$

进而有

推论 6.43 设 V 是域 K 上的 n 维向量空间.

(1) 如果 $p>n$, 那么 $\Lambda^p(V)=0$;

(2) 如果 $0\leqslant p\leqslant n$, 那么

$$\dim\Lambda^p(V)=\binom{n}{p}.$$

特别, 空间 $\Lambda^n(V)$ 是一维的, 由 n- 向量 $e_1\wedge e_2\wedge\cdots\wedge e_n$ 张成;

(3) 空间 V 的外代数 $\Lambda(V)$ 的维数是 2^n.

证明 (1) 从定理 6.42 的证明知 $\Lambda^p(V)$ 的元素都是张量

$$e_{j_1}\wedge e_{j_2}\wedge\cdots\wedge e_{j_p}=\mathscr{A}(e_{j_1}\otimes e_{j_2}\otimes\cdots\otimes e_{j_p}),\quad 1\leqslant j_1,j_2,\cdots,j_p\leqslant n$$

的线性组合. 当 $p>n$ 时, 任何序列 $j_1,\cdots,j_p\in\{1,2,\cdots,n\}$ 中都有某个数至少出现两次. 根据命题 6.41(3), 有 $e_{j_1}\wedge\cdots\wedge e_{j_p}=0$. 所以 $p>n$ 时有 $\Lambda^p(V)=0$. (1) 得证.

(2) 如果 $p=0$, 那么 $\Lambda^p(V)=K$, 所以它是一维的. 当 $1\leqslant p\leqslant n$, 定理 6.42 中的基向量的个数就是集合 $\{1,2,\cdots,n\}$ 中取 p 个元素的组合数, 这个数是 $\binom{n}{p}$.

(3) 由 (1) 和 (2) 知

$$\dim\Lambda(V)=\sum_{p=0}^{n}\binom{n}{p}=2^n.$$

□

公式 (6.4.64) 表明外代数 $\Lambda(V)$ 不是交换代数, 但离交换也不远. 那个公式更一般的形式是

$$Q \in \Lambda^q(V), \quad R \in \Lambda^r(V) \implies Q \wedge R = (-1)^{qr} R \wedge Q. \tag{6.4.70}$$

这个性质称为外积的**阶化反交换性**(graded anticommutative) 或**超交换性**(supercommutative) 是很有用的.

由于运算 $\wedge$ 是双线性的, 要证明式 (6.4.70), 考虑

$$Q = e_{i_1} \wedge \cdots \wedge e_{i_q}, \quad R = e_{j_1} \wedge \cdots \wedge e_{j_r}$$

的情形足矣. 利用外积的结合性, 并把公式 (6.4.64) 用 q 次, 可得

$$(e_{i_1} \wedge \cdots \wedge e_{i_q}) \wedge e_{j_k} = (-1)^q e_{j_k} \wedge (e_{i_1} \wedge \cdots \wedge e_{i_q}).$$

也就是说, $Q \wedge e_{j_k} = (-1)^q e_{j_k} \wedge Q$. 由此可见

$$\begin{aligned} Q \wedge R &= Q \wedge (e_{j_1} \wedge \cdots \wedge e_{j_r}) = (-1)^q e_{j_1} \wedge (Q \wedge e_{j_2} \wedge \cdots \wedge e_{j_r}) \\ &= (-1)^{2q} e_{j_1} \wedge e_{j_2} \wedge (Q \wedge e_{j_3} \wedge \cdots \wedge e_{j_r}) = \cdots = (-1)^{qr} R \wedge Q. \end{aligned}$$

由公式 (6.4.70), 当 p 是奇数时, 对任何的 p-向量 P, 有

$$P \wedge P = 0. \tag{6.4.71}$$

这是式 (6.4.64) 中第二个等式的推广. 当 p 为偶数时, 等式 (6.4.71) 一般不成立.

例 6.44 设 $V = \langle e_1, \cdots, e_n \rangle$, $n \geqslant 4$, 那么

$$\begin{aligned} &(e_1 \wedge e_3 - e_2 \wedge e_4) \wedge (e_1 \wedge e_3 - e_2 \wedge e_4) \\ &= -e_1 \wedge e_3 \wedge e_2 \wedge e_4 - e_2 \wedge e_4 \wedge e_1 \wedge e_3 = 2e_1 \wedge e_2 \wedge e_3 \wedge e_4 \neq 0. \end{aligned}$$

四 与行列式的联系 外积在代数学、在微分几何、流形上的分析、代数几何等分支都是重要的, 其中一个原因是它与行列式的联系. 我们已经知道, 行列式可以认为是平行六面体的体积. 外积运算能简洁地刻画向量组的线性无关性, 并导出行列式的所有重要性质.

定理 6.45 设 V 是域 K 上的 n 维向量空间. 空间 V 中的向量组 $v_1, v_2, \cdots, v_p$ 线性无关当且仅当它们的外积不等于零, 即

$$v_1 \wedge v_2 \wedge \cdots \wedge v_p \neq 0.$$

证明 如果向量组 $v_1, v_2, \cdots, v_p$ 线性相关, 那么至少有一个, 比如说 v_1, 是其余向量的线性组合. 这时候外积 $v_1 \wedge v_2 \wedge \cdots \wedge v_p$ 可以分解成外积

$$a_i v_i \wedge v_2 \wedge \cdots \wedge v_p, \quad 2 \leqslant i \leqslant p$$

的和, 其中每一项都含有相同的因子, 从而都是零.

反之, 如果向量组 $v_1, v_2, \cdots, v_p$ 线性无关, 那么这些向量可以作为 V 的某个基的前 p 个向量. 于是, p- 向量 $v_1 \wedge v_2 \wedge \cdots \wedge v_p$ 就是 $\Lambda^p(V)$ 某个基中的向量, 从而不等于零. □

设 $e_1, \cdots, e_n$ 是 V 的基. 根据推论 6.43(2), $\Lambda^n(V)$ 是 1 维的, $e_1 \wedge \cdots \wedge e_n$ 是其中的一个非零向量. 于是, 对任意 n 个向量 $v_1, \cdots, v_n \in V$, 有

$$v_1 \wedge \cdots \wedge v_n = \Delta e_1 \wedge \cdots \wedge e_n,$$

其中 $\Delta = \Delta(v_1, \cdots, v_n) \in K$ 是纯量, 它可以看成向量 $v_1, \cdots, v_n$ 的函数. 根据外积的性质, 这个函数是多重线性的, 斜对称的, 而且, $v_i = e_i$, $i = 1, \cdots, n$ 时, 有 $\Delta = 1$. 如同第一卷定理 4.1 的证明, 可以看出 Δ 是诸 v_i 关于基 (e_j) 的坐标形成的矩阵的行列式. 也就是说, 命

$$v_j = \sum_{i=1}^{n} a_{ij} e_i, \quad j = 1, \cdots, n,$$

那么 $\Delta(v_1, \cdots, v_n) = \det(a_{ij})$. 从而有

$$v_1 \wedge \cdots \wedge v_n = \det(a_{ij}) e_1 \wedge \cdots \wedge e_n. \tag{6.4.72}$$

从这个公式很容易推出行列式的所有性质, 而且可以得到若干定理的最自然的证明. 把这个公式应用到 V 的 p 维子空间 U, 则知如下等式成立:

$$w_1 \wedge \cdots \wedge w_p = \lambda u_1 \wedge \cdots \wedge u_n, \quad \lambda \in K, \tag{6.4.73}$$

其中 $w_1, \cdots, w_p$ 和 $u_1, \cdots, u_p$ 都是 U 的基, λ 是从基 (u_i) 到基 (w_i) 的转换矩阵的行列式.

如果只考虑 p 个向量 $v_1, \cdots, v_p$ 的外积, 则有

$$\begin{aligned}
v_1 \wedge \cdots \wedge v_p &= \sum_{i_1, \cdots, i_p} a_{i_1 1} \cdots a_{i_p p} e_{i_1} \wedge \cdots \wedge e_{i_p} \\
&= \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \left(\sum_{\pi \in S_p} a_{i_{\pi 1} 1} \cdots a_{i_{\pi p} p} e_{i_{\pi 1}} \wedge \cdots \wedge e_{i_{\pi p}} \right) \\
&= \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \left(\sum_{\pi \in S_p} a_{i_{\pi 1} 1} \cdots a_{i_{\pi p} p} (\varepsilon_\pi e_{i_1} \wedge \cdots \wedge e_{i_p}) \right) \\
&= \sum_{1 \leqslant i_1 < \cdots < i_p \leqslant n} \left(\sum_{\pi \in S_p} \varepsilon_\pi a_{i_{\pi 1} 1} \cdots a_{i_{\pi p} p} \right) e_{i_1} \wedge \cdots \wedge e_{i_p}.
\end{aligned}$$

根据第一卷推论 4.4, 上式最后一行的括号中的和恰好是一个行列式的值:

$$\begin{aligned}\Delta_{i_1\cdots i_p}(v_1,\cdots,v_p):&=\sum_{\pi\in S_p}\varepsilon_\pi a_{i_{\pi 1}1}\cdots a_{i_{\pi p}p}\\&=\begin{vmatrix}a_{i_11}&\cdots&a_{i_1p}\\\vdots&&\vdots\\a_{i_p1}&\cdots&a_{i_pp}\end{vmatrix}=\det(a_{i_kj})_{1\leqslant j,k\leqslant p}.\end{aligned}\tag{6.4.74}$$

由此可见, 下面的结论成立.

定理 6.46　设 $e_1,\cdots,e_n$ 是空间 V 的基, 且

$$v_j=\sum_{i=1}^n a_{ij}e_i,\quad j=1,2,\cdots,p,$$

是 V 中任意 p 个向量. 那么

$$v_1\wedge\cdots\wedge v_p=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}\Delta_{i_1\cdots i_p}(v_1,\cdots,v_p)e_{i_1}\wedge\cdots\wedge e_{i_p},\tag{6.4.75}$$

其中 $\Delta_{i_1\cdots i_p}(v_1,\cdots,v_p)$ 由公式 (6.4.74) 定义的行列式.

这个定理中的公式显然是公式 (6.4.72) 的推广. 行列式 $\Delta_{i_1\cdots i_p}(v_1,\cdots,v_p)$ 可以看作向量 $v_1,\cdots,v_p$ 的函数, 它是多重线性的, 斜对称的, 所以 $\Delta_{i_1\cdots i_p}$ 是一个 $(p,0)$ 型斜对称张量, 即是一个 p- 形式. 特别, $\det=\Delta_{1\cdots n}$ 是一个 n- 形式. 当然, 这些形式都依赖选定的基 $e_1,\cdots,e_n$.

五　向量子空间与 p- 向量　公式 (6.4.73) 提示 V 的 p 维子空间可能可以通过 p- 向量刻画. 设 $P\in\Lambda^p(V)$ 是非零 p- 向量. 我们对集合

$$\operatorname{Ann}P=\{\,x\in V\mid x\wedge P=0\,\}$$

感兴趣, 称它为 P 的**零化子**(annihilator). 首先, 它是 V 的子空间. 的确, 对 $x,y\in\operatorname{Ann}P$, $a,b\in K$, 有

$$(ax+by)\wedge P=ax\wedge P+by\wedge P=0,$$

所以 $ax+by\in\operatorname{Ann}P$.

定理 6.47　设非零 p- 向量 $P\in\Lambda^p(V)$ 的零化子 $\operatorname{Ann}P=\langle e_1,\cdots,e_r\rangle$ 是 r 维子空间. 那么 $r\leqslant p$, 且存在一个 $(p-r)$- 向量 Q 使得

$$P=Q\wedge e_1\wedge\cdots\wedge e_r.$$

等式 $r=p$ 成立当且仅当 P 是可分解 p- 向量 (即存在向量 $u_1,\cdots,u_p$ 使得 $P=u_1\wedge\cdots\wedge u_p$, 定义见 6.2 节第四部分).

证明 把向量组 $e_1,\cdots,e_r$ 扩充为 V 的基 $e_1,\cdots,e_n$. 根据定理 6.42, 有

$$P=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}P^{i_1\cdots i_p}e_{i_1}\wedge\cdots\wedge e_{i_p}. \tag{6.4.76}$$

于是, 零化子条件可以写成

$$P=\sum_{1\leqslant i_1<\cdots<i_p\leqslant n}P^{i_1\cdots i_p}e_j\wedge e_{i_1}\wedge\cdots\wedge e_{i_p},\quad j=1,2,\cdots,r. \tag{6.4.77}$$

如果 j 与某个 i_k 相同, 那么 $e_j\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}=0$, 从而 (6.4.77) 的左边出现的非零基向量是 $e_j\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}$, $j\neq i_1,\ \cdots,\ j\neq i_p$. 这些非零基向量线性无关, 所以, 只要 $1,2,\cdots,r$ 中有某个指标 j 不出现在 $i_1,\cdots,i_p$ 中, 就有 $e_j\wedge e_{i_1}\wedge\cdots\wedge e_{i_p}\neq 0$, 从而 $P^{i_1\cdots i_p}=0$. 由于 $P\neq 0$, 所以 $r\leqslant p$, 而且

$$\begin{aligned}P&=\sum P^{12\cdots ri_{r+1}\cdots i_p}e_1\wedge\cdots\wedge e_r\wedge e_{i_{r+1}}\wedge\cdots\wedge e_{i_p}\\&=\sum(-1)^{r(p-r)}P^{12\cdots ri_{r+1}\cdots i_p}e_{i_{r+1}}\wedge\cdots\wedge e_{i_p}\wedge e_1\wedge\cdots\wedge e_r\\&=Q\wedge e_1\wedge\cdots\wedge e_r,\end{aligned} \tag{6.4.78}$$

其中 $Q=\sum(-1)^{r(p-r)}P^{12\cdots ri_{r+1}\cdots i_p}e_{i_{r+1}}\wedge\cdots\wedge e_{i_p}$ 是一个 $(p-r)$- 向量.

如果 $r=p$, 那么 Q 是纯量, 所以 $P=\lambda e_1\wedge\cdots\wedge e_p$ 是可分解向量.

反过来, 假设 $P\neq 0$ 可分解, 即

$$P=u_1\wedge\cdots\wedge u_p.$$

那么 $u_1,\cdots,u_p\in\operatorname{Ann}P$. 根据定理 6.45, 这些向量线性无关, 所以 $\dim\operatorname{Ann}P\geqslant p$. 前面已经证明了 $\dim\operatorname{Ann}P\leqslant p$, 因而此时必须有 $\dim\operatorname{Ann}P=p$. □

注 从证明中可以看出, 也存在 $(p-r)$- 向量 Q' 使得 $P=e_1\wedge\cdots\wedge e_r\wedge Q'$, 且 Q' 与 Q 至多差一个负号.

定理 6.48 设 $U=\langle u_1,\cdots,u_p\rangle$ 和 $W=\langle u_1,\cdots,u_p\rangle$ 是 V 的具有相同维数的子空间. 那么 U 和 W 相等的充要条件是 $\Lambda^p(U)$ 和 $\Lambda^p(W)$ 相等, 即 p- 向量 $P=u_1\wedge\cdots\wedge u_p$ 和 p- 向量 $Q=w_1\wedge\cdots\wedge w_p$ 成比例.

证明 如果 $U=W$, 由公式 (6.4.73) 知 P 和 Q 成比例. 反之, 假设 $P=\lambda Q$, 根据定理 6.47 得

$$U=\operatorname{Ann}P=\operatorname{Ann}Q=W.$$ □

下面的定理是子空间与 p- 向量的关系的一些更精细的结论.

定理 6.49　设 $U = \mathrm{Ann}\, P$, $W = \mathrm{Ann}\, Q$, 其中 P 是可分解的 p- 向量, Q 是可分解的 q- 向量. 那么

(1) $U \supseteq W \Leftrightarrow P = R \wedge Q$ (R 是某个 $(p-q)$- 向量);

(2) $U \cap W = 0 \Leftrightarrow P \wedge Q \neq 0$;

(3) 当 $U \cap W = 0$ 时, 有 $U \oplus W = \mathrm{Ann}\,(P \wedge Q)$.

证明　结论 (1) 是定理 6.47 的一部分.

假设 $P = u_1 \wedge \cdots \wedge u_p$, $Q = w_1 \wedge \cdots \wedge w_p$, 且 $P \wedge Q \neq 0$. 根据定理 6.45, 向量 $u_1, \cdots, u_p, w_1, \cdots, w_q$ 线性无关, 从而它们构成 $U \oplus W$ 的一个基. 此时有 $U \cap W = 0$, 且定理 6.47 告诉我们 $U \oplus W = \mathrm{Ann}\,(P \wedge Q)$.

如果 $P \wedge Q = 0$, 根据定理 6.45, 存在非平凡的线性关系式

$$a_1 u_1 + \cdots + a_p u_p + b_1 w_1 + \cdots + b_q w_q = 0.$$

由此可知 $0 \neq a_1 u_1 + \cdots + a_p u_p \in U \cap W$. □

六　可分解向量　上一部分的讨论表明可分解向量有丰富的性质. 定理 6.47 通过 p-向量的零化子的维数刻画向量的可分解性. 我们先利用这一刻画讨论一些简单且有意思的情形.

定理 6.50　设 V 是 n 维向量空间. 那么 V 上的任意非零 $(n-1)$-向量 P 是可分解的.

证明　根据推论 6.43(2) 或公式 (6.4.72), 有

$$P \wedge x = f(x) e_1 \wedge \cdots \wedge e_n, \quad f(x) \in K,$$

其中 $f: V \to K$, $x \to f(x)$ 是纯量函数, (e_i) 是 V 的基. 外积运算对因子的线性性质蕴含 f 是线性函数. 此外, $\mathrm{Ann}\, P = \ker f$. 从而有 $\dim \mathrm{Ann}\, P = \dim \ker f \geqslant n-1$. 根据定理 6.47, $\dim \mathrm{Ann}\, P \leqslant n-1$. 所有必须有 $\dim \mathrm{Ann}\, P = n-1$. 这保证了 P 是可分解的 (参见定理 6.47). □

例 6.51　设 V 是 3 维欧几里得向量空间, e_1, e_2, e_3 是它的标准正交基. 如果 $P = u \wedge v \neq 0$, 那么

$$u \wedge v \wedge x = (w|x) \cdot e_1 \wedge e_2 \wedge e_3,$$

其中 $w = u \times v$ 是 $u = a_1 e_1 + a_2 e_2 + a_3 e_3$ 和 $v = b_1 e_1 + b_2 e_2 + b_3 e_3$ 的叉积(cross product), 定义为

$$u \times v = \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2) e_1 - (a_1 b_3 - a_3 b_1) e_2 + (a_1 b_2 - a_2 b_1) e_3.$$

容易看出, 对 $x = x_1e_1 + x_2e_2 + x_3e_3$ 有

$$(u \times v \,|\, x) = \begin{vmatrix} x_1 & x_2 & x_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

而且, $u \times v$ 与 u 和 v 都正交.

更一般地, 设 V 是 n 维欧几里得向量空间, $e_1, \cdots, e_n$ 是它的标准正交基. 如果 $P = v_1 \wedge \cdots \wedge v_{n-1} \neq 0$, 那么

$$v_1 \wedge \cdots \wedge v_{n-1} \wedge x = (w|x) \cdot e_1 \wedge \cdots \wedge e_n.$$

设 $v_i = a_{i1}e_1 + \cdots + a_{in}e_n$, $i = 1, \cdots, n-1$, 那么上式中的 w 由下面的公式给出:

$$w = \begin{vmatrix} e_1 & \cdots & e_n \\ a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n} \end{vmatrix}.$$

容易看出, w 与 $v_1, \cdots, v_{n-1}$ 都正交.

显然, 对任何可分解向量 P 有 $P \wedge P = 0$. 对**二重向量**(bivector, 即 2-向量), 这个条件也是充分的.

定理 6.52 二重向量

$$P = \sum_{1 \leqslant i < j \leqslant n} P^{ij} e_i \wedge e_j \tag{6.4.79}$$

可分解当且仅当 $P \wedge P = 0$.

证明 需要证明, 如果 $P \wedge P = 0$, 那么 P 是可分解的. 对 $n = \dim V$ 做归纳法. 当 $n = 3$ 时, 由定理 6.50 知 P 是可分解的.

把式 (6.4.79) 中所有含 e_1 的项放在一起, 它们可以写成 $e_1 \wedge u$ 的形式. 命

$$Q = \sum_{2 \leqslant i < j \leqslant n} P^{ij} e_i \wedge e_j. \tag{6.4.80}$$

则有

$$P = e_1 \wedge u + Q,$$

其中 $u = P^{12}e_2 + \cdots + P^{1n}e_n$. 如果 $u = 0$ 或 $Q = 0$, 则 P 已经是可分解的. 下设 u 和 $Q0$ 均不等于零. 由 $P \wedge P = 0$ 得

$$(e_1 \wedge u) \wedge Q + Q \wedge (e_1 \wedge u) + Q \wedge Q = 0.$$

利用外积的阶化反交换性式 (6.4.70) 得

$$2e_1 \wedge u \wedge Q + Q \wedge Q = 0. \tag{6.4.81}$$

上式左端第一项是若干基向量 $e_1 \wedge e_j \wedge e_k \wedge e_l$ 的线性组合, 第二项是若干基向量 $e_i \wedge e_j \wedge e_k \wedge e_l$ $(l > k > j > i \geqslant 2)$ 的线性组合. 这两组基向量没有共同的元素, 因此, 上式可以分成两个等式

$$e_1 \wedge u \wedge Q = 0, \quad Q \wedge Q = 0. \tag{6.4.82}$$

但是 $Q \in \Lambda^2(U)$, 其中 $U = \langle e_2, \cdots, e_n \rangle$. 由归纳假设, Q 可分解. 设 $U = v \wedge w$, 并把它代入式 (6.4.82), 得

$$e_1 \wedge u \wedge v \wedge w = 0. \tag{6.4.83}$$

于是 e_1, u, v, w 线性相关. 由于 $u, v, w \in U$, $e_1 \notin U$, 所以 u, v, w 线性相关. 注意我们已经假设 $u \neq 0$, $Q = v \wedge w \neq 0$. 如果 $v = \lambda u$, 那么 $P = e_1 \wedge u + \lambda u \wedge w = (e_1 - \lambda w) \wedge u$ 是可分解的. 如果 u, v 线性无关, 则 w 是 u, v 的线性组合. 设 $w = au + bv$, 则有

$$Q = v \wedge w = v \wedge (au + bv) = av \wedge u$$

且

$$P = e_1 \wedge u + Q = e_1 \wedge u + av \wedge u = (e_1 + av) \wedge u. \qquad \square$$

七　格拉斯曼簇　定理 6.52 有很好的几何意义. 空间 $\Lambda^2(V)$ 中的可分解向量全体称为一个**格拉斯曼锥**. 这个锥的射影化称为一个**格拉斯曼簇**(Grassmann variety), 记作 $Gr_2(V)$. 从而 $Gr_2(V)$ 是 $\mathbb{P}(\Lambda^2(V))$ 的子集. 定理 6.52 说明 $Gr_2(V)$ 中的元素的齐次坐标 $(P^{11}, P^{12}, \cdots, P^{n-1,n})$ 形成的集合正好是一个二次齐次方程组的解集, 从而 $Gr_2(V)$ 是一个射影簇. 根据定理 6.48, $Gr_2(V)$ 的点与 V 的 2 维子空间是一一对应的. 这说明 V 的 2 维子空间全体可以自然地成为一个射影簇. 这是一件有意思的事情.

当 $n = \dim V = 4$ 时, 二重向量具有如下的形式:

$$P = P^{12}e_1 \wedge e_2 + P^{12}e_1 \wedge e_2 + P^{13}e_1 \wedge e_3 + P^{14}e_1 \wedge e_4 + P^{23}e_2 \wedge e_3 + P^{24}e_2 \wedge e_4 + P^{34}e_3 \wedge e_4.$$

计算它与自身的外积, 得

$$P \wedge P = (P^{12}P^{34} - P^{13}P^{24} + P^{14}P^{23})e_1 \wedge e_2 \wedge e_3 \wedge e_4.$$

根据定理 6.52, 知 P 可分解当且仅当

$$P^{12}P^{34} - P^{13}P^{24} + P^{14}P^{23} = 0. \tag{6.4.84}$$

可见，此时格拉斯曼簇 $Gr_2(V)$ 是 5 维射影空间中的二次曲面.

对 p-可分解向量，有同样的概念. 空间 $\Lambda^p(V)$ 中的可分解向量全体称为一个**格拉斯曼锥**. 这个锥的射影化称为**格拉斯曼簇**(Grassmann variety)，记作 $Gr_p(V)$. 定理 6.48 表明 $Gr_p(V)$ 的点与 V 的 p 维子空间是一一对应的.

取定 V 的基 $e_1, \cdots, e_n$. 那么 $\Lambda^p(V)$ 的一个基是 $e_{i_1} \wedge \cdots \wedge e_{i_p}$, $1 \leqslant i_1 < \cdots < i_p \leqslant n$. 定理 6.46 给出了 p- 向量 $v_1 \wedge \cdots \wedge v_p$ 在这个基下的坐标，称为 $U = \langle v_1, \cdots, v_p \rangle$ 的**普吕克坐标**(Plücker coordinates). 根据定理 6.48 和定理 6.47，这些坐标唯一确定了子空间 U. 当然，它们是这个 p- 向量在 $\Lambda^p(V)$ 在 $\mathbb{P}(\Lambda^p(V))$ 中对应的点的齐次坐标. 除了 $p = 0, 1, n-1, n$ 的情形，可分解 p- 向量形成的集合是 $\Lambda^p(V)$ 的真子集，所以子空间的普吕克坐标之间有某些关系，不能是任意的.

为了说清楚这些关系，我们约定一些记号. 假设给了一组纯量 $\xi_{i_1 \cdots i_p}$, $1 \leqslant i_1 < \cdots < i_p \leqslant n = \dim V$，对任意的置换 $\pi \in S_p$ 和任意的指标 $j_1, \cdots, j_p$，约定

$$\xi_{j_{\pi 1} \cdots j_{\pi p}} = \varepsilon_\pi \xi_{j_1 \cdots j_p}. \tag{6.4.85}$$

特别，如果指标 $j_1, \cdots, j_p$ 中有相同的，那么 $\xi_{j_1 \cdots j_p} = 0$. 对 $1 \leqslant i_1 < \cdots < i_p \leqslant n$，命 $\xi_{i_1 \cdots i_p} = \Delta_{i_1 \cdots i_p} = \Delta_{i_1 \cdots i_p}(v_1, \cdots, v_p)$(记号见定理 6.46)，它是一个 $n \times p$ 矩阵 A 中第 $i_1, \cdots, i_p$ 行组成的方阵的行列式. 那么，对任意的指标 $1 \leqslant j_1, \cdots, j_p \leqslant n$，$\xi_{j_{\pi 1} \cdots j_{\pi p}} = \Delta_{j_{\pi 1} \cdots j_{\pi p}}$ 就是这个矩阵 A 中第 $j_1, \cdots, j_p$ 行组成的方阵的行列式.

定理 6.53 一组纯量 $\xi_{i_1 \cdots i_p}$ 是某个 p 维子空间 $U \subset V$ 的普吕克坐标，当且仅当，它们不全为 0 且对任意的指标 $1 \leqslant i_1, \cdots, i_{p+1}, j_1, \cdots, j_{p-1} \leqslant n = \dim V$ 如下关系式成立:

$$\sum_{k=1}^{p+1} (-1)^k \xi_{i_1 \cdots \hat{i}_k \cdots i_{p+1}} \xi_{i_k j_1 \cdots j_{p-1}} = 0, \tag{6.4.86}$$

公式中符号 ^ 表示它下面的指标不出现. 这些关系 (式) 称为**普吕克关系 (式)**.

证明 由于普吕克关系式的左边对指标组 $i_1, \cdots, i_{p+1}$ 和 $j_1, \cdots, j_{p-1}$ 是斜对称的，所以不妨要求 $i_1 < \cdots < i_p < i_{p+1}$ 且 $j_1 < \cdots < j_{p-1}$.

首先证明对 p 维子空间 $U \subset V$ 的普吕克坐标 $\Delta_{i_1 \cdots i_p} = \Delta_{i_1 \cdots i_p}(v_1, \cdots, v_p)$，普吕克关系式成立.

行列式 $\Delta_{i_k j_1 \cdots j_{p-1}}$ 按第一行展开，得

$$\Delta_{i_k j_1 \cdots j_{p-1}} = \sum_{s=1}^{p} a_{i_k s} M_s,$$

其中 M_s 与 k 无关. 由此可知，只要证明

$$\sum_{k=1}^{p+1} (-1)^k \Delta_{i_1 \cdots \hat{i}_k \cdots i_{p+1}} a_{i_k s} = 0, \quad s = 1, \cdots, p. \tag{6.4.87}$$

把矩阵 $A=(a_{ij})_{1\leqslant i\leqslant n,\ 1\leqslant j\leqslant p}$ 的第 s 列 A^s 添加到 A 中去, 得到一个 $n\times(p+1)$ 矩阵 $A_s=(A,A^s)$, 即 A_s 得前 p 列与 A 的列是一样的, 第 $p+1$ 列是 A 的第 s 列. 矩阵 A_s 的第 $i_1,\cdots,i_{p+1}$ 行形成一个方阵, 其行列式按最后一列展开, 至多差一个负号, 它等于表达式 (6.4.87) 的左边. 由于这个行列式有两列相等, 所以值等于零. 于是等式 (6.4.87) 成立.

反过来, 假设纯量 $\xi_{i_1\cdots i_p}$ 不全为零, 且满足关系式 (6.4.86). 需要证明存在 $n\times p$ 矩阵 A 使得

$$\xi_{i_1\cdots i_p}=\Delta_{i_1\cdots i_p},\quad 1\leqslant i_1,\cdots,i_p\leqslant n \tag{6.4.88}$$

是 A 的第 $i_1,\cdots,i_p$ 行形成的方阵的行列式值.

对任意的 $\sigma\in S_n$, 命 $\eta_{i_1\cdots i_p}=\xi_{\sigma(i_1)\cdots\sigma(i_p)}$. 那么 $\eta_{i_1\cdots i_p}$ 对指标 $i_1,\cdots,i_p$ 有斜对称性, 且这些纯量之间也满足关系式 (6.4.86). 于是, 不妨设 $\xi_{1\cdots p}\neq 0$. 关系式 (6.4.86) 是齐次的, 命 $\xi'_{i_1\cdots i_p}=\xi^{-1}_{1\cdots p}\xi_{i_1\cdots i_p}$, 那么 $\xi'_{i_1\cdots i_p}$ 也满足关系式 (6.4.86). 从而, 可以进一步要求 $\xi_{i_1\cdots i_p}=1$. 我们寻找一个形如

$$A=\begin{pmatrix}1&0&\cdots&1\\0&1&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&1\\a_{p+1,1}&a_{p+1,2}&\cdots&a_{p+1,p}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{np}\end{pmatrix}$$

使得 (6.4.88) 成立. 对 $j>p$ 的情况, 此时有

$$\Delta_{1\cdots\hat{i}\cdots pj}=(-1)^i a_{ji}.$$

所以, 必须设

$$a_{ji}=(-1)^{p-i}\xi_{1\cdots\hat{i}\cdots pj}.$$

那么等式 (6.4.88) 成立如果集合 $\{i_1,\cdots,i_p\}$ 与 $\{1,\cdots,p\}$ 相差至多一个元素.

接下来要证明如果集合 $\{i_1,\cdots,i_p\}$ 与 $\{1,\cdots,p\}$ 相差至 m 个元素, 等式 (6.4.88) 成立. 对 m 用归纳法. 设 $i_1\notin\{1,\cdots,p\}$. 由 $\xi_{1\cdots p}=1$ 和 (6.4.86) 知下面等式成立:

$$\xi_{i_1\cdots i_p}=\xi_{1\cdots p}\xi_{i_1\cdots i_p}=\sum_{k=1}^{p}(-1)^{k+1}\xi_{i_1 1\cdots\hat{k}\cdots p}\xi_{k i_2\cdots i_p}. \tag{6.4.89}$$

另一方面, 把行列式 $\Delta_{i_1\cdots i_p}$ 按第一行展开, 或利用 (6.4.87), 可得

$$\Delta_{i_1\cdots i_p}=\sum_{k=1}^{p}(-1)^{k+1}\Delta_{i_1 1\cdots\hat{k}\cdots p}\Delta_{ki_2\cdots i_p}. \tag{6.4.90}$$

由归纳假设, 等式 (6.4.89) 和 (6.4.90) 的右边相等. 因此, $\xi_{i_1\cdots i_p}=\Delta_{i_1\cdots i_p}$. □

八 回首再看普法夫 外积不仅可以用于定义和研究行列式, 还可以构造和研究偶数阶斜对称矩阵的普法夫.

假设 $n=2m$, $A=(a_{ij})$ 是 n 阶斜对称方阵. 考虑二重向量

$$P=\sum_{i<j}a_{ij}(e_i\wedge e_j)=\frac{1}{2}\sum_{i,j}a_{ij}(e_i\wedge e_j),$$

其中 $e_1,\cdots,e_n$ 是空间 V 的一个给定的基. 计算 P 在 $\Lambda(V)$ 中的 m 次幂, 得

$$\begin{aligned}P^m=\underbrace{P\wedge\cdots\wedge P}_{m}&=\frac{1}{2^m}\sum_{i_1,\cdots,i_n}a_{i_1i_2}\cdots a_{i_{n-1}i_n}e_{i_1}\wedge\cdots\wedge e_{i_n}\\&=\frac{1}{2^m}\left(\sum_{(i_1,\cdots,i_n)}\operatorname{sgn}(i_1,\cdots,i_n)a_{i_1i_2}\cdots a_{i_{n-1}i_n}\right)e_1\wedge\cdots\wedge e_n,\end{aligned}$$

最后的和式是对指标 $1,\cdots,n$ 所有的置换 $(i_1,\cdots,i_n)$ 求和. 由于 $a_{ij}=-a_{ji}$, 置换 $(i_1,\cdots,i_n)$ 中两个指标交换位置后, 它的符号会改变, 所以, 如果置换 $(j_1,\cdots,j_n)$ 与置换 $(i_1,\cdots,i_n)$ 有如下的关系:

存在 $\sigma\in S_n$ 使得

$$\{\sigma(i_{2k-1}),\ \sigma(i_{2k})\}=\{j_{2k-1},\ j_{2k}\},\quad k=1,\cdots,m,$$

那么

$$\operatorname{sgn}(i_1,\cdots,i_n)a_{i_1i_2}\cdots a_{i_{n-1}i_n}=\operatorname{sgn}(j_1,\cdots,j_n)a_{j_1j_2}\cdots a_{j_{n-1}j_n}.$$

于是有

$$P^m=m!\left(\sum_{(i_1i_2|\cdots|i_{n-1}i_n)}\operatorname{sgn}(i_1,\cdots,i_n)a_{i_1i_2}\cdots a_{i_{n-1}i_n}\right)e_1\wedge\cdots\wedge e_n, \tag{6.4.91}$$

其中的和式对集合 $\{1,\cdots,n\}$ 的所有划分 $\{i_1,i_2\},\cdots,\{i_{n-1},i_n\}$ 求和 (这些二元集合之间的序随意取定, 二元集合里面的元素的序同样是随意取定).

定理 6.54 等式 (6.4.91) 中和式就是在 1.7 节中第九部分定义的矩阵 A 的普法夫, 即有

$$\operatorname{Pf}(A)=\sum_{(i_1i_2|\cdots|i_{n-1}i_n)}\operatorname{sgn}(i_1,\cdots,i_n)a_{i_1i_2}\cdots a_{i_{n-1}i_n}. \tag{6.4.92}$$

证明 等式 (6.4.92) 右端的表达式记为 $p(A)$. 我们先证明定理 1.89 中的两个公式:

(1) 对任何方阵 $C=(c_{ij})$ 有 $p(C\cdot A\cdot {}^tC)=\det C\cdot p(A)$;

(2) $p(A)^2=\det A$.

假设基域 K 的特征为 0. 设 $v_1,\cdots,v_n$ 是 V 的基, 且

$$(e_1,\cdots,e_n)=(v_1,\cdots,v_n)C.$$

把二重向量 $P=\dfrac{1}{2}\sum\limits_{i,j}a_{ij}e_i\wedge e_j$ 通过 $v_1,\cdots,v_n$ 表达, 得

$$P=\frac{1}{2}\sum_{i,j,k,l}a_{ij}c_{ki}c_{lj}v_k\wedge v_l=\frac{1}{2}\sum_{k,l}b_{kl}v_k\wedge v_l,$$

其中

$$b_{kl}=\sum_{i,j}a_{ij}c_{ki}c_{lj}.$$

命 $B=(b_{kl})$, 则有

$$B=C\cdot A\cdot {}^tC.$$

于是

$$P^m=m!\cdot p(A)e_1\wedge\cdots\wedge e_n=m!\cdot p(B)v_1\wedge\cdots\wedge v_n.$$

另一方面, 根据公式 (6.4.72) 有

$$e_1\wedge\cdots\wedge e_n=(\det C)v_1\wedge\cdots\wedge v_n.$$

于是得

$$p(A)\cdot\det C=p(B)=p(C\cdot A^t\cdot C).$$

如果把 a_{ij} $(i<j)$ 和 b_{kl} 都看作不定元, 那么上式可以看作是整数环 $\mathbb{Z}$ 上的多项式等式. 通过自然环的同态 $\mathbb{Z}\to\mathbb{Z}_p$, 可知上式也可以看成 $\mathbb{Z}_p$ 上的多项式等式, 从而对任意特征 p 的域也成立.

在定理 1.87 中把基的顺序适当改动, 可知存在非退化矩阵 C 使得

$$A=C\cdot F\cdot {}^tC,$$

其中 F 是如下形式的矩阵

$$F=\begin{pmatrix}0 & I_k & 0\\ -I_k & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

I_k 是 k 阶单位矩阵. 容易看出,

$$p(F) = \det F = \begin{cases} 1, & \operatorname{rank} F = n, \\ 0, & \operatorname{rank} F < n. \end{cases} \tag{6.4.93}$$

于是总有 $\det F = p(F)^2$. 根据已证明的断言 (1), 有

$$p(A) = \det C \cdot p(F).$$

同时,

$$\det A = (\det C)^2 \det F.$$

所以有 $\det A = p(A)^2$.

由 1.7 节中第十二部分的讨论知斜对称矩阵的普法夫完全由性质 (1) 和式 (6.4.93) 确定, 所以 $\operatorname{Pf}(A) = p(A)$. □

这个定理给了一个直接计算普法夫多项式的公式.

例 6.55 假设 $n = 4$. 集合 $\{1,2,3,4\}$ 划分成二元集合的方式有三种 $\{1,2\}$, $\{3,4\}$; $\{1,3\}$, $\{2,4\}$; 以及 $\{1,4\}$, $\{2,3\}$, 所以此时有

$$\operatorname{Pf}(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}.$$

把这个公式与式 (6.4.84) 比较即知二重向量 $P = \sum\limits_{i<j} a_{ij} e_i \wedge e_j$ 可分解当且仅当 $\operatorname{Pf}(A) = 0$. 根据定理 1.89(或定理 6.54 的证明) 知, 这等价于 $\det A = 0$. 由于斜对称矩阵的秩总是偶数, 所以 $\det A = 0$ 当且仅当 $\operatorname{rank} A \leqslant 2$. 当然, 很容易直接证明 P 可分解当且仅当 $\operatorname{rank} A \leqslant 2$.

九 关于基域特征的说明 本节关于外代数的讨论要利用斜对称化算子, 所以要求基域的特征为 0, 但很容易摆脱这一限制. 可以直接定义任意域 K 上的向量空间的外代数 $\Lambda(V)$ 如下. 它是域 K 上的一个向量空间, 有外积 $\wedge$, 其元素都是如下元素

$$\lambda,\ v_1 \wedge \cdots \wedge v_p, \quad p = 1, 2, \cdots,\ v_1, \cdots, v_p \in V$$

的线性组合; 外积运算满足结合律和反交换性: $v \wedge v = 0$ 对任何 $v \in V$. 而且, 如果有线性映射 $\varphi: V \to R$, 其中 R 是一个结合 K 代数, 满足 $\varphi(v)^2 = 0$ 对一切 $v \in V$, 那么存在唯一的 K 代数同态

$$\psi: \Lambda(V) \to R$$

使得 $\psi(u \wedge v) = \varphi(u)\varphi(v)$.

除了命题 6.41(1), 公式 (6.4.64) 及后面的公式和后面的关于外代数的所有结论对正特征域上的外代数 $\Lambda(V)$ 都是成立的. 在习题中我们将不再要求域的特征为 0.

习 题 6.4

1. 设 $V = K^n$ 是列坐标空间, $A_1, \cdots, A_n$ 是 V 的基, B 是 V 中任意的元素. 证明: 向量方程式

$$\sum_k \lambda_k A_k = B$$

的解参数 λ_k 可由关系式

$$\lambda_k(A_1 \wedge \cdots \wedge A_n) = A_1 \wedge \cdots \wedge A_{k-1} \wedge B \wedge A_{k+1} \wedge \cdots \wedge A_n$$

给出. 试由此导出克拉默法则 (第一卷定理 4.19).

2. 假设基域 K 的特征不等于 2. 证明:

$$\mathrm{T}_2^0(V) = \mathrm{T}_2^+(V) \oplus \Lambda^2(V^*),$$

但如果 $\dim V > 1$, 对 $p > 2$ 有 $\mathrm{T}_p^0(V) \neq \mathrm{T}_p^+(V) + \Lambda^p(V^*)$. (注意 $\Lambda^p(V^*)$ 是 $\mathrm{T}_p^0(V)$ 中的斜对称张量全体形成的向量空间, 参见 6.3 节第八部分的讨论.)

3. 外代数 $\Lambda(V)$ 的中心定义为

$$\mathcal{Z}(\Lambda(V)) = \{\, \xi \in \Lambda(V) \mid \xi \wedge \eta = \eta \wedge \xi, \forall \eta \in \Lambda(V) \,\}.$$

假设 V 是无限维向量空间且基域 K 的特征不等于 2. 证明:

$$\mathcal{Z}(\Lambda(V)) = \Lambda^0(V) + \Lambda^2(V) + \Lambda^4(V) + \cdots = \bigoplus_{k=0}^{\infty} \Lambda^{2k}(V).$$

4. 设 $\xi = \xi_0 + \xi_1 + \cdots + \xi_n \in \Lambda(V)$, 其中 $\xi_p \in \Lambda^p(V)$, $p = 0, 1, \cdots, n$. 证明 ξ 在 $\Lambda(V)$ 中可逆当且仅当 $\xi_0 \neq 0$.

5. 设 $\mathcal{A} : V \to V$ 是线性算子, $\dim V = n$. 线性算子

$$\wedge^p \mathcal{A} : \Lambda^p(V) \to \Lambda^p(V)$$

在可分解 p- 向量上的像定义为

$$(\wedge^p \mathcal{A})(v_1 \wedge \cdots \wedge v_p) = \mathcal{A}v_1 \wedge \cdots \wedge \mathcal{A}v_p.$$

称 $\wedge^p \mathcal{A}$ 为 $\mathcal{A}$ 的p **次外幂**(pth exterior power). 证明:

$$\det(\wedge^p \mathcal{A}) \cdot \det(\wedge^{n-p} \mathcal{A}) = (\det \mathcal{A})^{\binom{n}{p}}.$$

6. 记号同上题. 证明: V 的 p 维子空间 U 是 $\mathcal{A}$ 的不变子空间当且仅当 $\Lambda^p(U)$ 是 $\wedge^p \mathcal{A}$ 的不变子空间.

7. 记号同第 5 题. 证明:

(1) $\mathrm{tr} \wedge^2 \mathcal{A} = \frac{1}{2}[(\mathrm{tr}\mathcal{A})^2 - \mathrm{tr}\mathcal{A}^2]$;

(2) $\mathcal{A}$ 的特征多项式

$$\chi_{\mathcal{A}}(t)=\sum_{k=0}^{n}\operatorname{tr}(\wedge^k\mathcal{A})(-t)^{n-k}.$$

8. 设 V 和 W 是向量空间. 证明 $\Lambda^p(V\oplus W)$ 与 $\bigoplus_{i=0}^{p}\left(\Lambda^i(V)\otimes\Lambda^{p-i}(W)\right)$ 同构.

9. 设 V 是 n 维向量空间, q 是其上的非退化二次型, f 是 q 的极化. 在空间 $\Lambda^p(V)$ 上可以用公式

$$q^{\wedge 0}=1,\quad q^{\wedge p}(v_1\wedge\cdots\wedge v_p)=\begin{vmatrix} f(v_1,v_1) & \cdots & f(v_1,v_p)\\ \vdots & & \vdots\\ f(v_p,v_1) & \cdots & f(v_p,v_p)\end{vmatrix}$$

定义二次型 $q^{\wedge p}$. 证明: $q^{\wedge p}$ 是非退化的.

10. 设 V 是 n 维欧几里得向量空间, 内积由正定的二次型 q 给出. 称元素 $d\in\Lambda^n(V)$ 是 V 的一个**定向**(orientation) 如果 $q^{\wedge n}(d)=1$.

另一方面, 设 $(e_1,\cdots,e_n)$ 和 $(e_1',\cdots,e_n')$ 是空间 V 的两个有序基. 如果从一个基到另一个基的转换矩阵的行列式是正的, 则说这两个基是**同向的**或**定向相同**. 于是, V 的有序基分成两类, 每一类由同向的有序基构成. 每一类都称为 V 的一个**定向**.

定向的这两个定义有联系么?

11. 假设 V 是 n 维向量空间. 证明: 对 $p\leqslant q\leqslant n$, 存在双线性映射

$$\varphi:\Lambda^p(V)\times\Lambda^q(V^*)\to\Lambda^{q-p}(V^*),$$

它在可分解元素上的定义是

$$\begin{aligned}&\varphi(v_1\wedge\cdots\wedge v_p,\alpha_1\wedge\cdots\wedge\alpha_q)\\ =&\sum_{(i_1,\cdots,i_p)}\operatorname{sign}(i_1,\cdots,i_p,j_1,\cdots,j_{q-p})\alpha_{i_1}(v_1)\cdots\alpha_{i_p}(v_p)\alpha_{j_1}\wedge\cdots\wedge\alpha_{j_{q-p}},\end{aligned}$$

其中和式对所有不同的有序指标组 $(i_1,\cdots,i_p)$ 求和; $\{j_1,\cdots,j_{q-p}\}$ 是集合 $\{i_1,\cdots,i_p\}$ 在 $\{1,\cdots,q\}$ 的补集, 随意排一个序.

12. 记号同上一题. 证明对每一个非零元素 $\xi\in\Lambda^n(V^*)$, 映射

$$\Lambda^p(V)\to\Lambda^{n-p}(V^*),\quad u\to\varphi(u,\xi)$$

是线性同构, 且这个同构把可分解元素映到可分解元素. 由此可知任何 $(n-1)$- 向量是可分解的, 参见定理 6.50.

参考文献

普洛斯库列柯夫 N B, 1982. 线性代数习题集. 北京: 人民教育出版社.

希洛夫 G E, 2013. 线性空间引论. 2 版. 王梓坤, 等译. 北京: 高等教育出版社.

许以超, 2008. 线性代数与矩阵论. 2 版. 北京: 高等教育出版社.

APOSTOL T M, 2010. 线性代数及其应用导论, 沈灏, 沈佳辰, 译. 北京: 人民邮电出版社.

ARTIN M, 2011. Algebra. 2nd ed. 北京: 机械工业出版社.

AUDIN M, 2003. Geometry. New York: Springer-Verlag.

AXLER S, 2016. 线性代数应该这样学. 2 版, 杜现昆, 等译. 北京: 人民邮电出版社.

BERGER M, 1987. Geometry I, II. New York: Springer-Verlag.

KOSTRIKIN A I. 1996. Exercises in Algebra: A Collection of Exercises in Algebra, Linear Algebra and Geometry. Algebra, Logic and Applications series, Vol. 6. Gordon and Breach Publishers.

LANG S, 2002. Algebra, revised 3rd. Graduate Texts in Mathematics. Springer-Verlag, 世界图书出版公司.

LAX P D, 2007. Linear Algebra and Its Application, New York: John Wiley & Sons, Inc.

LAY D C, 2005. 线性代数及其应用. 3 版, 刘深泉, 等译. 北京: 机械工业出版社.

SHILOV G E, 1971. Linear Algebra. New York: Dover Publications, Inc.

VINBERG E B, 2003. A Course in Algebra. Graduate Studies in Mathematics, Vol. 56. American Mathematical Society.

附　　录

基础数学的一些过去和现状

席南华

摘要：本文试图通过人们对一些基本的数学研究对象如素数、圆、球、方程、函数等的探索历程展示基础数学的特点、部分思想和发展及现在活跃的一些研究方向.

谈论整个数学或者基础数学的发展趋势已经超出一个人的能力，庞加莱和希尔伯特被认为是数学领域最后两个全才. 后来还有一些杰出的数学家如外尔、冯·诺依曼、柯尔莫格罗夫和 I.M. 格尔方德等对纯数学和应用数学都做出巨大的贡献，但现在这样的数学家也很难寻到了.

基础数学大致分为代数 (含数论)、几何、分析 (基于微积分的数学) 三部分，但看一看前几届国际数学家大会的报告目录及其分组就知道现代数学的分支繁多，各个部分之间的融合与交叉也是日趋深入. 有些方向是非常活跃的，如代数几何、数论、表示理论、动力系统、偏微分方程、几何分析、调和分析、微分几何、微分拓扑、复几何、拓扑、组合、数学物理等等.

数学当然是研究数与形的科学，也研究结构. 逻辑支撑着数学的大厦，而逻辑本身也是数学研究的对象，与计算机科学密切相关.

1. 数学理论的起始

形是容易感知的，我们一睁开眼睛就会看到各种各样形状的物体. 数却是一个抽象的概念，但其形成也有很长历史了，据考证和研究，人类在洞穴时代就已经有数的概念了，若干动物也有数的概念. 刚开始时，实际的需要产生了加法、减法、乘法、除法等运算，长度、面积等概念. 到公元前三千年，数学的应用范围就很广了，如税收、建筑、天文等. 数学从理论上系统研究始于古希腊人，在公元前六百年至公元前三百年期间，代表人物有毕达哥拉斯、欧几里得等. 欧几里得的《几何原理》采用公理化体系系统整理了古希腊人的数学成就，两千多年来一直是数学领域的教科书，其体系、数学理论的表述方式和书中体现的思维方式对数学乃至科学的发展影响深远.

2. 数和多项式方程及相关的数学分支

我们认识数学基本上都是从数开始的，然后是简单的几何与多项式方程. 数中

间有无穷的魅力、奥秘和神奇, 始终吸引着最富智慧的数学家和业余爱好者. 多项式方程是从实际问题和数的研究中自然产生的. 在对数和多项式方程的认识和探究过程中, 代数、数论、组合、代数几何等数学分支逐步产生.

2.1 素数

素数有无穷多个, 在《几何原理》中有一个优美的证明. 素数是数学永恒的研究对象, 而且是最难以琢磨的数学研究对象, 很多最为深刻的数学都与素数 (或其复杂的其他形式如素理想等) 有关. 我们熟知的孪生素数猜想和哥德巴赫猜想, 到现在仍未解决, 孪生素数猜想的巨大突破由张益唐做出 (2014 年), 哥德巴赫猜想目前最好的结果是陈景润的 (1973 年). 但奇数哥德巴赫猜想由维诺格拉多夫于 1937 年基本解决. 哈代–利特伍德猜想是比孪生素数猜想更强的猜想.

对于素数在自然数中的比例, 有著名的素数定理, 曾是勒让德的猜想 (1808 年), 阿达马和德拉瓦勒–普森最先分别证明该定理 (1896 年). 1949 年赛尔伯格和厄尔迪斯分别给出素数定理的初等证明. 这是赛尔伯格获 1950 年菲尔兹奖的重要工作之一.

2004 年陶哲轩和本合作证明了存在任意长的等差素数数列. 这项工作极大地激发了人们对解析数论的新热情, 也是陶获 2006 年菲尔兹奖的重要工作之一.

18 世纪欧拉对素数有无穷多个给出一个深刻的证明, 他用到无穷级数 $1+2^{-1}+3^{-1}+\cdots$ 的发散性. 他还对实数 s 考虑了级数 $1+2^{-s}+3^{-s}+\cdots$. 1859 年, 为研究素数的分布, 黎曼对复数 s 考虑这个级数, 证明了它可以延拓成复平面上的亚纯函数, 现称为黎曼 ζ 函数, 给出了函数方程, 建立了这个函数的零点和素数分布的联系, 提出了著名的黎曼猜想. 这个猜想断言黎曼 ζ 函数的零点除平凡的外实部均为二分之一. 黎曼对素数和 ζ 函数的研究工作影响深远. 一般认为黎曼猜想是数学中最有名的猜想, 也是克雷数学研究所的悬赏百万美元的千禧年问题之一, 自它提出之时起就在数学研究中占有突出的位置, 很多问题与它有关, 还与算子代数、非交换几何、统计物理等有深刻的联系, 在阿达马和德拉瓦勒–普森对素数定理的证明中起关键的作用.

黎曼的工作对 L 函数和代数几何也有巨大的影响. L 函数已是数论的一个中心研究对象, 与分析、几何及表示论的联系极深, 其在一些特殊点的值含有很多深刻的算术信息. 我们先从狄利赫列的 L 函数说起.

2.2 L 函数和朗兰之纲领

对有限循环群的特征, 狄利赫列构造了与黎曼 ζ 函数类似的函数, 现称为狄利赫列 L 函数. 利用这些函数, 他证明了一个有趣的结论 —— 很多算术数列含有无限多个素数. 具体说来就是：如果两个正整数 a 和 m 互素, 那么算术数列 $a+m$,

$a+2m, a+3m\cdots, a+km, \cdots$ 里有无穷多个素数.

后来阿丁对数域的有限扩张域的伽罗华群的表示, 类似地也定义了一类 L 级数并解析延拓得到一个 L 函数, 现称为阿丁 L 函数. 利用这些 L 函数, 他证明了交换类域论里面很有名的阿丁互反律. 20 世纪六七十年代朗兰之想把阿丁的工作延伸到非交换的类域论去. 雅各和朗兰之对 p 进域上的简约代数群的不可约表示和整体域上的简约代数群的自守表示也定义了 L 函数. 朗兰之给出了一系列的猜想, 这就是现在非常热闹的朗兰之纲领.

这个纲领的中心是函子性 (functoriality) 猜想, 该猜想描述了不同代数群的自守表示之间深刻的联系. 函子性猜想蕴含了很多著名的猜想, 如阿丁猜想, 拉玛努金猜想, 佐藤–塔特猜想等. 函子性猜想的一个重要特殊情况是朗兰之互反律, 或说朗兰之对应. 通过整体域上简约代数群的自守表示定义的 L 函数称为自守 L 函数. 还有一种 L 函数称为模体 (motivic) L 函数, 是哈塞–韦伊 L 函数的推广, 例子包括阿丁 L 函数和哈塞–韦伊 L 函数. 本质上朗兰之纲领的中心问题就是证明所有的模体 L 函数均是自守 L 函数.

在最简单的情形, 函子性猜想就是阿丁互反律, 类域论的实质. 函子性猜想仅在一些很特别的情形得到证明, 离完全解决遥远得很. 但对函数域上的一般线性群, 拉佛格在 2002 年证明了朗兰之的互反律猜想 (即建立了朗兰之对应), 并因此获得当年的菲尔兹奖. 2010 年发表的基本引理的证明也是这个纲领中的一个巨大进展, 有意思的是来自代数群表示论的仿射斯普林格纤维和因研究可积系统而产生的希钦纤维化之间的联系在吴宝珠的证明中起一个关键的作用. 吴宝珠因其对基本引理的证明获得 2010 年的菲尔兹奖.

研究函子性猜想的重要工具是赛尔伯格–亚瑟迹公式. 赛尔伯格迹公式 1956 年得出, 与黎曼 ζ 函数的联系导致他引进了赛尔伯格 ζ 函数. 赛尔伯格迹公式后由亚瑟在 1974 年至 2003 年间做出各种推广, 它在数学物理中也有很好的应用.

2.3　一元高次方程和群论

人们很早就会解一元一次方程和一元二次方程, 一元三次方程和四次方程的公式解在 16 世纪被找到. 在尝试得到更高次方程的根式解时, 数学家的探索失败了, 其中包括 18 世纪一流的数学家拉格朗日. 答案原来是否定: 1824 年挪威数学家阿贝尔证明了五次及更高次的方程一般没有根式解. 稍后几年法国数学家伽罗瓦给出的证明影响深远, 一个重要的数学分支 —— 群论因此而诞生. 我们可以简单说一下伽罗瓦的证明. 5 个人排队的排法有 120 种, 一种排法按另一种方法重排就会产生第三种排法, 于是这 120 种排法成为一个群, 而且是不可解的, 所以五次及更高次的方程一般没有根式解.

群论的影响几乎遍及整个数学, 在物理和化学及材料科学中有很多的应用, 是

研究对称的基本工具. 1872 年克莱因提出著名的埃尔朗根纲领, 用群来分类和刻画几何, 对几何的发展影响巨大. 拓扑学中同调群和同伦群是极其重要的研究工具和研究对象. 代数几何中的阿贝尔簇是一类特别重要的几何对象. 很多空间具有一些自然的群作用, 从而可以作相应的商空间. 这些商空间在几何、数论和表示论中极其重要. 齐性空间和志村簇是其中两类例子. 几何不变量则是一个有关的重要数学分支.

群论自身的研究同样是非常深刻的. 20 世纪一项伟大的数学成就是对有限单群的分类. 这是一项庞大的工作, 第一个证明主要的工作发表于 1960 年至 1983 年期间, 前后有 100 多位数学家参与, 数百篇发表的论文, 总长度超过 10000 页. 到 2004 年, 群论专家完成第二个证明, 总长度也有 5000 页. 现在, 他们正试图进一步简化. 汤普森因其在单群分类中的杰出工作于 1974 年获菲尔兹奖, 他最出名的工作是与费特合作证明了伯恩赛德猜想: 非交换的有限单群的阶是偶数, 论文发表于 1963 年, 占了太平洋数学杂志整个一期. 阿西巴赫因其在有限单群分类的杰出工作获 2010 年沃尔夫奖. 在有限单群中有一个非常大的单群, 称为魔群, 其中元素的个数大约是 8×10^{53}, 与数学中的月光猜想密切相关. 1992 年波谢兹证明了这个猜想, 为此他引进了广义卡茨–穆迪代数, 与他人一起引进了顶点算子代数. 现在, 这些代数都是重要的研究对象. 主要因为这项工作, 波谢兹于 1998 年获菲尔兹奖.

如果把所有整系数的一元多项式方程的根放在一起, 我们得到一个数的集合, 比有理数全体大, 称为有理数域的代数闭包. 有理数域的代数闭包的绝对伽罗华群及其表示的研究是现代数学尤其是数论中极其重要的研究课题.

如果一个数不是任何整系数一元多项式的根, 则称这个数是超越数, 圆周率 π 就是一个超越数. 超越数的研究也是数论的重要组成部分, 贝克尔曾因对超越数的研究获得 1970 年的菲尔兹奖. 一些自然产生的数如某些无穷级数的和与某些函数的值等是否为超越数是人们特别感兴趣的.

在群论中, 李群和代数群的理论与其他数学分支的联系十分广泛和深刻. 群表示论, 尤其是李群和代数群的表示论是现在非常活跃的分支. 李群和代数群的离散子群特别有意思, 与数论和遍历论等分支的联系极密切, 马古利斯因其在半单李群的离散子群上的深刻工作获得 1978 年的菲尔兹奖.

2.4　*不定方程和数论*

不定方程是数论研究的中心对象之一. 直角三角形三边的关系 $X^2+Y^2=Z^2$ 就是一个不定方程, 它与圆方程类似. 它有很多的整数解, 勾三股四弦五就给出一组. 一般的解很容易给出: $X=a^2-b^2$, $Y=2ab$, $Z=a^2+b^2$, 其中 a, b 是任意整数. 高次的情形就是方程 $X^n+Y^n=Z^n$, 其中 n 是大于 2 的整数. 1637 年, 费马在一本书内的边页写道他有一个此方程无非平凡整数解的证明, 但太长, 边页空

白处写不下. 人们怎么也没找出费马说的那个证明, 一般认为费马在书中注记说的证明可能有问题, 于是此方程无非平凡整数解成为一个猜想, 称为费马大定理问题. 这个猜想一直吸引着数学家的强烈兴趣, 费马本人对 4 次的情形的证明流传下来, 3 次的情形是欧拉在 1770 年证明的, 5 次的情形于 1825 年由勒让德和狄利赫列独立证明, 等等. 19 世纪库莫对这个问题的研究导致了代数数论的诞生. 1920 年, 莫德尔提出一个猜想：有理数域上亏格大于 1 的代数曲线的有理点只有有限多个. 这个猜想被法尔廷斯于 1983 年证明, 它蕴含了费马的方程在 n 比 2 大时至多存在有限多个本原整数解. 法尔廷斯主要因此获得 1986 年的菲尔兹奖. 费马大定理最后在 1995 年被外尔斯证明, 这是 20 世纪一项伟大的数学成就. 代数数论现在是非常有活力的数学分支.

在外尔斯对费马大定理的证明中, 椭圆曲线起了关键的作用. 椭圆曲线的方程其实很简单：$Y^2 = X^3 + aX + b$, 其中 a, b 是常数, 如 1, 2 等等. 它们有群结构, 在射影空间中的几何图形就是环面, 与汽车轮胎一个形状. 对椭圆曲线也能定义 L 函数. BSD 猜想断言这个 L 函数在 1 处的值与椭圆曲线的群结构密切相关. 这个猜想是克雷数学研究所悬赏百万美元的千禧年问题之一, 自然是数学的研究热点之一.

BSD 猜想还和一个古老的问题有关. 如果考虑方程 $X^2 + Y^2 = Z^2$ 的正数解, 那么解是一个直角三角形的三个边长. 有一个古老的问题：什么时候这个三角形的面积 $XY/2$ 是整数, 而且 X, Y, Z 都是有理数. 这样的整数称为和谐数 (congruent number). 数组 (3,4,5) 和 (3/2, 20/3, 41/6) 是方程的解, 所以 6 和 5 都是和谐数. 塔奈尔 1983 年的一个结果告诉我们如果 BSD 猜想成立, 有可行的计算办法判定一个整数是否为和谐数.

椭圆曲线还与数学的几何密切相关. 巴嘎瓦在数的几何中发展了一些强有力的方法, 并把这些方法用于小秩环的计数和估计椭圆曲线的平均秩. 他因此于 2014 年获菲尔兹奖.

2.5　*多项式方程和代数几何*

我们已经看到解方程, 哪怕是一个一元的或简单的二元方程, 都不是容易的事情, 其研究给数学已经而且还要带来巨大的发展. 多项式方程组的求解显然是更为困难, 甚至一般说来是毫无希望的. 我们需要换一个角度, 把一组多项式方程的零点集看作一个整体, 就会得到一个几何空间, 称为簇. 研究簇的数学分支就是代数几何, 一个庞大深刻又极富活力的分支. 我们读中学时就知道一个二元一次方程和直线是一回事, $X^2 + Y^2 = 1$ 则是单位圆周的方程. 代数几何的踪迹可以追溯到公元前, 17 世纪笛卡尔建立的解析几何可以看作代数几何的先声.

代数几何的中心问题是对代数簇分类. 但这个问题太大太难, 现阶段没希望完

全解决, 人们只能从不同的角度考虑更弱的问题. 一维的情形是代数曲线, 其分类很容易, 在 19 世纪就知道光滑的射影曲线可以用他们的亏格来分类, 这时还有著名的黎曼–洛赫定理. 在 1885 年至 1935 年期间, 代数几何史上著名的意大利学派对二维的情形研究了分类, 也得到了二维情形的黎曼–洛赫定理. 意大利学派的特点是几何直观思想丰富深刻, 后期的工作严格性不足. 后来, 20 世纪四五十年代韦伊和查里斯基用新的语言严格表述代数几何的基础. 小平邦彦和沙法列维奇及其学生在 20 世纪 60 年代重新整理了代数曲面的分类. 小平在代数几何和复流形上的工作十分有影响, 早在 1954 年, 他就获得菲尔兹奖, 沙法列维奇在代数数论和代数几何上都做出了重要的贡献, 有著名的沙法列维奇猜想, 至今未解决.

曼福德和庞比利在 20 世纪六七十年代把意大利学派对曲面的分类工作做到了特征 p 域上. 曼福德在代数几何方面的贡献是多方面的, 构造了给定亏格的曲线的模空间、几何不变量的研究等, 因为这些贡献, 他于 1974 年获菲尔兹奖. 庞比利则因其在解析数论、代数几何和分析数学上的杰出工作于 1974 年获菲尔兹奖.

三维情形的分类直到 20 世纪 80 年代才由日本数学家森重文完成, 他因此于 1990 年获菲尔兹奖. 如何把这些分类的工作推广到高维的情形是非常活跃的研究方向.

前面提到的黎曼–洛赫定理是极其重要的定理, 它计算了某些函数空间的维数. 1954 年希茨布茹赫把它推广到高维, 现称为希茨布茹赫–黎曼–洛赫定理. 这是他最为人知的工作, 其实他对拓扑、复分析和代数几何都做出重要的贡献, 1988 年获沃尔夫奖. 希茨布茹赫–黎曼–洛赫定理很快被格罗登迪克进一步推广成格罗登迪克–希茨布茹赫–黎曼–洛赫定理. 为此, 格罗登迪克定义了 K 群, 这是 K 理论的开始. 后来阿梯亚和希茨布茹赫发展了拓扑 K 理论, 它被阿梯亚和辛格用于证明阿梯亚 - 辛格指标定理. 希茨布茹赫–黎曼–洛赫定理也是 1963 年出现的阿梯亚–辛格指标定理的先声. 阿梯亚于 1966 年获菲尔兹奖, 这个指标定理是他最为有名的结果. K 理论已成为代数、数论、几何、拓扑等分支的重要工具, 奎楼因为在 20 世纪 70 年代建立了高阶 K 理论而于 1978 年获菲尔兹奖, 沃尔沃兹基因其对米尔诺关于 K 群的一个猜想的证明和相关的工作获得 2002 年菲尔兹奖.

对有限域上的代数簇, 韦伊 1949 年提出了一个猜想, 其中一部分可以看作黎曼猜想在有限域上的形式, 对以后代数几何的发展影响巨大, 包括塞尔和格罗登迪克在代数几何上的工作. 20 世纪五六十年代格罗登迪克用概型的语言改写了代数几何, 在此基础上极大地发展了代数几何, 包括为证明韦伊猜想而建立的 l 进制上同调理论. 他于 1966 年获菲尔兹奖. 他的思想和工作对代数几何与数学的发展产生了深远的影响. 1974 年格罗登迪克的学生德林用 l 进制上同调证明了韦伊猜想中的黎曼假设部分并主要因此于 1978 年获菲尔兹奖.

如果一个代数簇有奇点, 那么很多对研究无奇点的代数簇有效的工具就失效

了. 1964 年広中平祐找到一个办法解消奇点, 为此他于 1970 年获得菲尔兹奖. 几何中的奇点非常有意思, 常常蕴含了丰富的信息, 与其他的分支有出人意料的联系, 如舒伯特簇的奇点和李代数的表示的联系就是一个例子.

2.6　群和李代数的表示理论

前面我们看到因为一元高次方程的研究产生了群论, 它的应用很广泛. 很多时候, 群是通过它的表示应用到其他分支和领域. 表示在数学中间是随处可见的, 比如说我们熟悉的多项式环, 分析里面的平方可积函数空间, 拓扑里面的上同调群和 K 群等等, 就有丰富的表示结构. 在物理和化学中也很常见, 例如, 在单粒子模型中, 单电子的轨道波函数生成三阶正交群的表示, 自旋波函数生成二阶酉群的表示. 20 世纪 60 年代吉尔–曼用三阶酉群的十维表示预言了 Ω 粒子的存在, 后来很快被实验证实.

群表示理论是一个庞大而且非常活跃的研究领域, 在数学和物理中应用广泛. 李群和代数群在单位元处的切空间是李代数, 可以看作李群和代数群的线性化. 李代数和相关的代数如顶点算子代数等及其表示同样在数学和物理中应用广泛. 有限群的表示可以通过其群代数的模来研究. 过去几十年, 代数的表示论有很大的发展, 尤其是林格尔发现代数表示论与量子群的联系之后. I.M. 格尔方德似乎对这个领域有独特的感受, 曾经说 "所有的数学就是某类表示论" (All of mathematics is some kind of representation theory). 他是伟大的数学家, 从研究的广度和深度来说, 20 世纪后半叶能和他相提并论的数学家是非常少的, 对表示论做出的贡献广泛深刻.

表示论的基本的思想有两点: 一个是对称, 一个是线性化. 这个领域关心的主要问题有: 最基本的表示的性质, 如分类、维数、特征标等; 一般的表示如何从最基本的表示构建; 如何构造最基本的表示; 一些自然得到的表示的性质; 等等. 大致说来表示论就是要弄清楚这些事情.

表示论一直吸引着最优秀的数学家, 早期如 S. 李、E. 嘉当 (陈省身先生的老师)、外尔, 后来有 I.M. 格尔方德、哈里西–钱德拉、赛尔伯格等, 现在有朗兰之、卡兹但、俊菲尔德、拉佛格、路兹梯格、吴宝珠等. 奥昆寇夫的工作揭示了概率论、表示论和代数几何之间的一些深刻联系, 并因此获 2006 年菲尔兹奖.

表示论过去几十年的发展可能给人印象最深的是几何方法在代数群和量子群表示理论中的运用并由此产生的几何表示论、用表示论研究数论的朗兰之纲领和一个平行的几何朗兰之纲领、李 (超) 代数及其表示的发展与在理论物理和数学物理中的应用 (包括标准模型), 还有近二十年的一股范畴化潮流. 另外, 传统的李群表示理论、代数表示论和有限群的模表示理论也是很活跃的. 这些依然是表示论的主要研究方向. 几何中的相交上同调、反常层理论和 K 理论在表示论中的运用给表示论带来巨大的进展, 很多困难的问题得到解决, 也带来了很多新的研究课题.

这个方向的一个代表性人物是路兹梯格. 正是用几何的方法, 他建立了有限李型群的特征标理论, 或许这是目前有限群表示理论中最为深入的部分.

2.7　计数、集合论和数理逻辑

计算一些物品的数量当然是我们日常生活经常要做的事情. 对有限集合, 确定其中元素的个数理论上不是问题, 一个一个数就行了. 组合论的一部分就是研究计数, 和数论密切相关. 但对无限集合, 事情显然并不简单. 例如某人有个面积无穷的王国, 国土增加一两平方公里的面积对他显然没什么意义. 无限集合的计数理论是德国人康托在 19 世纪后半叶建立的, 称为集合论. 其中一个核心的概念是等势: 两个集合称为等势的如果它们之间能建立一一对应. 有意思的一件事情是自然数集合和有理数集合等势, 但与实数集合不等势. 1874 年, 康托尔提出有名的连续统假设: 实数集合的任何无穷子集要么与实数集合等势, 要么与自然数集合等势. 1940 年哥德尔证明了这个假设与现有的公理体系不矛盾. 20 世纪 60 年代, 科恩建立了强有力的力迫法, 证明了连续统假设之否与现有的公理体系不矛盾, 他因为这项工作获得 1966 年的菲尔兹奖.

现代数学是建立在集合论上的, 集合论也是数理逻辑的重要组成部分. 连续统假设表明我们的逻辑体系并不能对每一个陈述断定真伪. 事实上更早以前就有各种各样的悖论和哥德尔的不完全定理表明数学逻辑体系的危机. 数学家为补救这些缺陷做了巨大的努力, 这包括罗素和怀特海德的三大卷《数学原理》等. 罗素获得 1950 年的诺贝尔文学奖. 与数理逻辑密切相关的一个问题是 P 和 NP 问题, 这是克雷数学研究所的千禧年问题之一, 也是理论计算机科学领域最有名的问题. 简单说, P 和 NP 本质上问的是如下事情: 给了一些整数, 能否有很快捷的方法 (即多项式时间算法) 判断这些整数的某一部分的和为零.

模型论是数理逻辑的一个分支, 在代数和代数几何有深刻的应用, 有些代数几何的结果还是最先用模型论发现并证明的. 赫鲁晓夫斯基 1996 年用模型论证明了函数域上的莫德尔–朗猜想, 名噪一时.

3. 形与几何、拓扑

最简单的形无疑是线段、直线、多边形、多面体、圆、球、椭圆、抛物线、双曲线等, 它们也是几何与拓扑的起点, 人类很早就研究它们了. 我们做一点简单的游戏: 多边形的顶点的个数等于边的个数, 凸多面体的面的个数加上顶点的个数等于棱的个数加二. 后一个等式称为欧拉公式, 虽然并不是欧拉最早发现的. 这些公式被认为是拓扑学的起源. 拓扑学研究几何空间的整体性质, 就是说那些在连续变形下不变的性质, 是数学的主流分支, 在数学的其他分支和物理中的应用极其广泛, 有时是研究一些问题必不可少的工具, 如广义相对论中的一般性的时空奇点定理就

是彭罗斯把拓扑学引入广义相对论而证明的.

如果把多面体的棱角磨平, 再整理一下, 我们就得到球了. 欧拉公式本质上是说球面的欧拉示性数等于 2. 一个几何空间的欧拉示性数是通过空间的同调群定义的. 球面当然是一个光滑的曲面. 对于一般的光滑曲面, 有高斯—伯内特公式, 它把曲面的曲率和欧拉示性数联系起来, 从而把微分几何与拓扑联系起来, 非常深刻, 对以后数学的发展影响很大. 20 世纪 40 年代, 阿冷多尔费尔和韦伊把它推广到高维的情形. 陈省身对高维情形的高斯–伯内特公式的证明则是整体微分几何一个开端, 影响深远.

上面提到同调群, 它们是研究拓扑的主要手段之一, 也是代数拓扑研究的主要对象之一. 为了不同的目的, 人们定义了各种各样的同调群和上同调群. 在好的空间如流形上, 这些 (上) 同调群都是一样, 而且有著名的庞加莱对偶. 但对有奇点的空间, 如何定义好的 (上) 同调群, 花了人们很长的时间. 直到 20 世纪 80 年代, 高热斯基和曼可菲森才找到对空间奇点研究很有意义的一种上同调, 称为相交上同调. 后来伯恩斯坦、贝林森和德林三人用层的语言处理相交上同调, 形成了反常层理论. 很快相交上同调和反常层理论成为研究代数几何、拓扑和表示论的强有力工具. 夫洛尔同调在低维拓扑和辛几何中是有力的研究工具, 它是夫洛尔为研究辛几何中的阿诺德猜想而引进的.

同调群中有一些特别的元素对研究认识空间的几何结构非常重要, 这些元素就是示性类. 最著名的示性类有陈类、史提芬–惠特尼类、庞特列亚金类等. 对光滑的复代数簇的德拉姆上同调, 其中一些元素称为霍奇类. 代数几何中一个未解决的主要问题是霍奇猜想, 它断言霍奇类都是一些代数圈类的有理线性组合. 这也是克雷数学研究所的千禧年问题之一.

圆和球是我们熟悉的基本形状, 在数学上的意义是非凡的. 圆周在三维空间的嵌入称为纽结. 通俗说来纽结就是一根首尾相连的柔软绳子, 在不弄断绳子, 也不打结的情况下, 它在三维空间中的各种样子. 纽结理论是拓扑学中非常活跃的分支, 一个重要的问题是寻找纽结不变量. 20 世纪 20 年代发现的亚历山大多项式是纽结不变量, 纽结补的基本群是纽结不变量, 称为纽结群. 20 世纪 70 年代, 瑟斯顿把双曲几何引入纽结的研究中, 从而定义了新的有力的不变量. 20 世纪 80 年代琼斯发现了新的多项式不变量 —— 琼斯多项式. 威腾和孔策维奇等人一系列的后续工作则揭示了纽结和统计力学、量子场论之间的深刻联系. 琼斯多项式是琼斯 1990 年获菲尔兹奖的重要工作之一. 图拉耶夫等人用量子群研究纽结, 得到新的不变量, 很有影响. 以上是圆周给我们带来的深刻数学的一部分. 下面我们看一下高维的情形 —— 球面.

关于球面, 最有名的应该是庞加莱 1904 年提出的猜想, 它断言一个单连通的闭三维流形与球面同胚. 在 2003 年被解决前, 这个猜想是拓扑学中的一个中心问

题. 在此之前, 数学家做过很多的努力. 既然三维的情形证明不了, 人们就对高维的情形考虑类似的问题. 1961 年, 斯梅尔证明了当维数大于四时, 高维的庞加莱猜想成立, 因此他获得 1966 年的菲尔兹奖. 1982 年弗里德曼对四维的情形证明了庞加莱猜想, 于是他获得 1986 年的菲尔兹奖. 庞加莱猜想最后在 2003 年被佩雷曼证明, 这是轰动一时的结果, 标志了数学中一个大问题的终结, 也是克雷数学研究所七个千禧年问题中到目前为止唯一被证明的. 佩雷曼证明这个猜想所用的工具是非常有意思的, 那就是几何分析. 几何分析是微分几何与微分方程的交叉学科, 丘成桐、后来还有哈密顿等人在其中的建立和发展起了突出的作用, 是一个有力的工具, 也是非常活跃的研究方向. 2007 年布仁德尔和舍恩用几何分析的方法证明了微分球定理, 是流形理论中一个重要结论.

球面带来的深刻数学还很多. 1956 年, 米尔诺发现七维球面上有非标准的微分结构. 这一发现对拓扑学的发展影响很大, 是米尔诺最有名的工作, 也是他 1962 年获菲尔兹奖的主要工作之一. 六维球面是否有复结构则是困扰数学家很多年的一个问题, 至今未解决. 球面的同伦群也是拓扑学研究的重要问题, 至今未完全解决. 20 世纪 50 年代初, 塞尔成功计算了球面的很多同伦群, 这是他获 1954 年菲尔兹奖的重要工作之一. 同伦群现在仍是拓扑学研究的一个主要方向.

在几何与拓扑中, 一个基本问题是对流形分类. 流形有各种各样的, 如拓扑流形、微分流形、复流形、黎曼流形、辛流形、无穷维流形, 等等, 这里面的问题和结果都是非常丰富的. 闭二维拓扑流形是曲面, 其分类很早就知道, 结果很漂亮：可定向闭曲面的同构类由曲面的亏格完全确定, 不可定向的闭曲面则同胚于一些实射影线的联通和. 曲面的亏格就是曲面所围的空洞的个数, 如汽车轮胎是亏格为 1 的曲面, 它只围了一个空洞.

黎曼面是一维的复流形, 一直是非常重要的研究对象. 米扎哈尼因其在黎曼面及其模空间的动力系统和几何上的杰出工作获得 2014 年菲尔兹奖. 她是第一位获此奖的女性.

三维流形的研究中, 瑟斯顿的工作非常重要, 他发现双曲几何在三维流形的研究中起突出的作用. 瑟斯顿提出的几何化猜想是比庞加莱三维球面猜想更广泛的猜想, 后与庞加莱猜想一起得到证明. 瑟斯顿因其在三维流形上的开创性工作获得 1982 年的菲尔兹奖.

4. 切线、面积、速度、加速度等和微积分、分析数学

我们会求一些简单图形如多边形、圆等的面积, 也会求圆的切线, 但对更复杂的图形, 这就不是一件容易的事情了. 在物理中, 对于非匀速运动, 求加速度和路程同样不是一件容易的事情. 对这些问题探索最后导致牛顿和莱布尼茨在 17 世纪分别独立建立了微积分. 用微积分我们能轻易求出一些复杂图形的面积、体积, 确定

物体的加速度、路程, π 的精确值等等. 微积分及在其上发展起来的分析数学成为认识和探索世界奥秘最有力的数学工具之一, 为数学带来全面的大发展, 促进了很多新分支的产生如解析数论、实分析、复分析、调和分析、微分几何、微分拓扑、微分方程等等.

微积分的基本概念有极限、微分和积分, 分析数学的基本研究对象是函数. 1927 年物理学家狄拉克在研究量子力学时引进了 δ 函数, 它不是经典意义下的函数, 给当时的数学家带来很大的困惑. 施瓦兹建立的分布理论使得 δ 函数变得容易理解并能严格处理, 他因此获 1950 年的菲尔兹奖. 分布理论在现代偏微分方程理论中极其重要.

正弦函数和余弦函数都是周期函数. 傅里叶认为它们是描述周期运动的基本函数并在 19 世纪初建立了相应的理论, 现称为傅里叶分析. 傅里叶分析及其更一般的理论调和分析是内容非常丰富且应用很广泛的数学分支. 如果注意到正弦和余弦函数可以看作圆周上的函数并把单位圆周与模长为一的复数等同起来, 就知道傅里叶分析与李群表示论是密切相关的. 卡尔松因其在调和分析上的重要工作于 1992 年获沃尔夫奖, 特别他理清了函数与其傅里叶级数表示的关系. 陶哲轩在调和分析上的工作也是他获菲尔兹奖的工作的一部分. 李群和拓扑群上的调和分析是一个重要的分支, 与泛函分析密切相关, 在数论中的深刻应用使人惊叹.

大自然很多的奥秘是通过微分方程表述的, 描写电磁运动的麦克斯韦方程, 描写微观世界的薛定谔方程, 描写流体运动的纳维尔–斯托克斯方程, 描写宏观世界的爱因斯坦方程等等. 这些方程都是非线性微分方程, 有很多人研究. 纳维尔–斯托克斯方程是否有整体光滑解则是克雷数学研究所的千禧年问题之一.

在线性偏微分方程上, 赫曼德的工作可能是最深刻和突出的, 他因此获得 1962 年的菲尔兹奖. P.-L. 里翁斯在非线性方程上的杰出工作使他获得了 1994 年的菲尔兹奖. 丘成桐发展一些强有力的偏微分方程技巧用以解决微分几何的一些重要问题如卡拉比猜想等, 在这些工作的基础上, 几何分析逐步发展起来. 因为这些工作, 丘成桐获得 1982 年的菲尔兹奖, 另外, 他的工作在理论物理和数学物理中有极大的影响. 偏微分方程领域引人入胜的深刻问题比比皆是, 一流的数学家很多, 如拉克斯、卡发热利等等. 过去这些年, 随机偏微分方程发展迅速. 海热尔因其在随机偏微分方程方面的工作尤其是建立了这类方程的正则性理论获得 2014 年菲尔兹奖.

只有一个独立变量的微分方程称为常微分方程, 很多这类方程来自经典力学, 如牛顿第二定律, 独立变量很多时候就是时间. 混沌理论来自常微分方程的研究. 事情起源于 19 世纪末, 自 17 世纪以来人们一直试图弄清太阳系行星运行轨道的稳定性. 如果只有两个星球, 那么牛顿的万有引力定律很容易导出星球的轨道行为, 但太阳系是多体的, 极其复杂. 庞加莱想先把三体问题解决, 但发现问题太困难, 清楚写出微分方程的解是没希望的, 只能考虑解的定性研究, 发现解的混沌性. 对一

些微分方程的解混沌性, 有一个通俗的说法 —— 蝴蝶效应, 意指在一定的约束下, 刚开始时很小的差别可以导致后来巨大的差异. 混沌理论的应用十分广泛, 气象预报是其中之一. 三体问题的一个幂级数解在 1912 年由逊德曼给出, 但对初始值有很强的要求, 而且收敛得很慢. 逊德曼的结果被王秋东 (音译) 在 1991 年推广到多体的情形, 但没考虑奇点问题.

常微分方程解的定性研究与动力系统密切相关. 太阳系的运动是一个动力系统 (运动和力之间关系的系统), 由万有引力决定, 所以是一个常微分方程的动力系统. 庞加莱对太阳系和三体问题的研究是动力系统史上非常重要的工作. 动力系统是很活跃的研究领域, 其中一个研究方向是复动力系统, 研究函数的迭代. 约科兹因其在动力系统的杰出工作获 1994 年菲尔兹奖. 曼克木棱在复动力系统方面的重要工作是他获 1998 年菲尔兹奖的原因之一. 部分因其在动力系统方面的重要工作, 斯米尔诺夫获得 2010 年菲尔兹奖. 阿维拉因其在动力系统上的深刻工作于 2014 年获菲尔兹奖. 研究有不变测度的动力系统的分支称为遍历论, 与调和分析、李群及其表示、代数群、数论有密切的联系. 林德施特劳斯因其在遍历论中的出色工作获得 2010 年的菲尔兹奖, 另外马古利斯获 1978 年菲尔兹奖的工作中遍历论起了重要的作用.

在 19 世纪对常微分方程的研究导致了李群和李代数的诞生, 后者在数学和物理中的应用广泛深刻.

无限维空间上的分析是泛函分析, 巴拿赫空间和希尔伯特空间及其上面的算子是基本的研究对象, 其中的希尔伯特空间对量子力学有着基本的重要性. 泛函分析的重要一支是算子代数, 与表示论、微分几何等有深入的联系. 孔内斯因对一些算子代数的分类获得 1982 年的菲尔兹奖. 他还把泛函分析引入非交换微分几何的研究中. 高韦尔斯主要因其在巴拿赫空间上的重要工作获 1998 年的菲尔兹奖.

5. 数学物理

物理一直是给数学发展带来最为强大推动力量的学科, 在这里有着无穷无尽的问题, 提供非常鲜活、生动的思想, 它永远给数学带来很多特别深刻的东西. 弦理论、量子场论和规范场论是非常活跃的领域. 弦理论能统一四种基本的作用力, 把量子力学和相对论统一起来. 卡拉比–丘流形在超弦理论中非常重要, 因为额外的时空被认为是六维卡拉比–丘流形. 杨–米尔斯理论是一种规范场论, 共形场论则是一种量子场论.

20 世纪 80 年代初期, 唐纳森利用杨–米尔斯理论中的方程的一类特别的解, 称为瞬子, 研究四维流形的微分结构, 证明了一大类四维流形没有光滑结构, 而有些则有无穷多的微分结构. 唐纳森因其在四维流形上的开创性工作获得 1986 年的菲尔兹奖. 结合他的结果和弗里德曼关于四维流形分类的结果, 1987 年陶贝斯证明

了四维欧氏空间有不可数多的微分结构. 注意我们生存的三维空间加上一维的时间就是四维欧氏空间, 而其他维数的欧氏空间则仅有一种微分结构. 瞬子在数学和物理中都有很多的用处. 杨–米尔斯理论在数学上则可能是最受重视的规范场理论, 是否对任意的紧单的规范群在四维欧氏空间存在质量间隙非负的量子杨–米尔斯理论是克雷数学研究所千禧年问题之一.

在共形场论的研究中, 群论、李代数、顶点算子代数、维那索拉代数等代数结构是描述对称的工具, 十分重要.

也是在 20 世纪 80 年代, 数学物理中对量子可积系统和杨–巴克斯特方程的研究导致了俊菲尔德和神保 (相互独立) 在 20 世纪 80 年代中期定义了量子群, 随后引发了世界范围的研究热潮, 产生了很多深刻的结果如典范基和晶体基, 新的纽结不变量等, 引出很多新的研究问题. 俊菲尔德因其在量子群和表示论上的工作获 1990 年菲尔兹奖.

在过去几十年的数学物理进展中必须提到威腾的工作, 他带来很多新的深刻思想, 在数学和物理中架起桥梁, 为相关研究方向带来全新的面貌和很多问题, 给数学和物理两者都带来巨大的影响, 因为其深刻的工作他于 1990 年获得菲尔兹奖. 在对两个假设的量子场论作比较时, 威腾对代数曲线的模空间提出一个猜想, 后被孔策维奇证明. 同样基于量子场论的考虑, 威腾认为存在一些可通过某些积分计算的纽结和三维流形不变量, 此事后被孔策维奇. 这些工作影响很大, 是孔策维奇获得 1998 年菲尔兹奖的部分主要工作.

最近这些年, 统计力学及相关的研究方向包括随机过程等非常活跃, 有很多突出的进展. 2006 年沃纳因其在随机洛马纳演化和二维布朗运动的几何等方面的工作获菲尔兹奖, 2010 年维那尼因其关于波尔兹曼方程和兰道阻尼的工作获得菲尔兹奖, 斯米尔诺夫获菲尔兹奖的部分工作也与统计力学有关.

结束语

以上对基础数学进展的介绍是很不全面的. 不过, 从以上的介绍可以看出, 数学的发展始终贯穿在对基本问题和基本对象的探索认识中. 好的问题对数学的发展起了巨大的推动作用. 在数学研究中, 我们需要考虑好的问题, 基本的问题, 同时要有好的数学思想. 写完这篇文章后, 一个强烈的感受是在数学的发展中, 我们做出的贡献太少. 缺乏好的传统和数学思想乃至背后的哲学思想和思考可能是一个重要的原因, 在这些方面我们还有很大的差距. 可能我国已有很多数学家感受到我们还未形成中文数学的思考体系和语言体系. 我们对数学的认识仍然很不足, 在努力成为数学强国的路途上我们有很多的东西需要弥补, 需要时间, 需要国家的支持, 更需要数学家的努力.

本文主要参考资料

关于数学史和数学思想的通俗读物可以看以下两套书:

[1] 克莱因 M,《古今数学思想》. 1~4 卷, 北京大学数学系, 译, 上海: 上海科学技术出版社, 2009.

[2] 亚历山大洛夫 A D, 等,《数学: 它的内容方法》. 1~3 卷, 孙小礼、赵孟养、裘光明, 等译. 北京: 科学出版社, 2010.

下文观点深邃, 思维流畅.

[3] Weil H, History of mathematics: Why and how, proceedings of the international congress of mathematicians, Helsinki, 1978, 1: 227-236.

如果希望对现代数学了解得更深入, 可以看下面的丛书.

[4] 《国外数学名著系列》(影印版), 北京: 科学出版社, 2006 年及以后.

国际数学家大会的论文集是了解现代数学进展的一个窗口, 可喜的是国际数学联盟已经在其网站汇集了 2006 年及以前大会论文集的电子版.

[5] 国际数学家大会论文集. http://mathunion.org/ICM/.

克雷数学研究所的七个千禧问题的描述见其网站.

[6] http://www.claymath.org/millennium/.

[7] 维基百科. http://en.wikipedia.org/. 维基百科的文章经常也是很有参考价值的.

数学家的传记对了解数学家和数学都是有益的, 下面列三本.

[8] Reid C. "Hilbert". New York: Springer-Verlag LLC, 1996. 有中译本.

[9] 王元.《华罗庚》, 南昌: 江西教育出版社, 1999.

[10] 张奠宙, 王善平.《陈省身传》(修订版). 天津: 南开大学出版社, 2011.

数学大家谈数学的文章或书籍多有对数学的深刻感悟和认识. 很有名的两本书是:

[11] Poincaré H, Mathematics and Science Last Essays, BiblioBazaar, 2009. 有中译本:《最后的沉思》.

[12] Hardy G H. A Mathematician's Apology. Cambridge: Cambridge University Press, 1940 年第一版, 1992 年 Canto 版. 有中译本:《一个数学家的辩白》.

与本文有关的较专门的参考文献和论文太多, 因篇幅有限, 就不罗列了.

说明: 本文首次发表于《中国科学院院刊》2012 年第 27 卷第 2 期, 134-144. 后略加修改, 转载于 "数学与人文" 系列从第十四卷《数学与科学》.

谨以此书献给我的妻子刘桂菊，她尽了一切努力保证我本职工作之余写作此书的时间.

(O-7168.01)

基础代数（第二卷）

科学出版社互联网入口
高教数理分社：(010)64015178　销售：(010)64031535
E-mail：mph@mail.sciencep.com
销售分类建议：数学

www.sciencep.com

ISBN 978-7-03-056033-9

定价：49.00 元