高等代数实验班习题课

吴乘洋

2025 年 4 月 12 日

目录

第 0 章　前言

To be added.

第 1 章　群、域、线性空间

本章首先介绍三个代数学的基本概念: 群 (group), 域 (field), 线性空间 (linear space). 粗糙地说, 群是表达对称性的工具, 域是描述数和运算的范围, 线性空间是抽象作用对象的模型. 这三者的重要之处不仅在于它们各自本身, 更在于它们之间存在的紧密联系.

§1.1　群概念简介

1.1.1　代数方程发展史与群论起源

代数学最初的主要任务是解代数方程.

对于一元一次方程 $ax+b=0\,(a\neq 0)$, 我们熟知它有唯一解 $x=-\dfrac{b}{a}$.

对于一元二次方程 $ax^2+bx+c=0\,(a\neq 0)$, 早在古巴比伦的文字泥板中就给出了其解法, 即我们今天所谓的配方法. 记 $\Delta=b^2-4ac$, 则

$$\begin{aligned}&ax^2+bx+c=0\\ \Longleftrightarrow&x^2+\frac{b}{a}x+\frac{c}{a}=0\\ \Longleftrightarrow&\left(x+\frac{b}{2a}\right)^2=\frac{\Delta}{4a^2}\\ \Longleftrightarrow&x=\frac{-b\pm\sqrt{\Delta}}{2a}.\end{aligned}$$

以上过程提示我们, 通过对变量做适当的平移, 可消去一元方程中的次高项.

具体地说, 对于一元 n 次多项式 $P(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0\,(a_n\neq 0)$, 令 $x=y+t$ (其中参数 t 待定), 则由二项式定理知,

$$P(x)=a_n(y+t)^n+a_{n-1}(y+t)^{n-1}+\sum_{k=2}^{n}a_{n-k}(y+t)^{n-k}=a_ny^n+(na_nt+a_{n-1})y^{n-1}+Q(y),$$

其中 $Q(y)$ 是关于 y 的次数小于 $(n-1)$ 的多项式. 取 $t=-\dfrac{a_{n-1}}{na_n}$, 即 $y=x-\dfrac{a_{n-1}}{na_n}$, 则

$$P(x)=0\Longleftrightarrow y^n+Q(y)=0.$$

因此解一元 n 次方程 $P(x)=0$ 等价于解不含次高项 (即 y^{n-1} 项) 的一元 n 次方程 $y^n+Q(y)=0$.

在历史上, 一元高次方程的解法一度是个困难的问题. 意大利数学家 L. Pacioli(1447∼1517) 于 1494 年在威尼斯发表了文艺复兴时期最伟大的数学著作 “Summa de arithmetica, geometria, proportioni et proportionalita”, 他在书中认为以当时的数学求解一元三次方程是不可能的. Pacioli 曾于 1501∼1502 年在 Bologna 大学任教, 其间与 Bologna 大学的教授 S. del Ferro(1465∼1526) 讨论数学问题, 后者不久就会解了不含二次项的一元三次方程. 然而 Ferro 并不愿意公开自己的成果, 而只在去世前秘密传授给了他的少数几个朋友与学生, 意大利人 A. M. del Fiore 就是其中的一位. 戏剧性的是, Fiore 虽然吹嘘自己会解所有的一元三次方程, 但他实际上只会解某种特殊形式; 与此同时, 意大利人 N. Fontana(1499∼1557)(绰号: Tartaglia, 即口吃者) 独立发现了不含一次项的一元三次方程的解法. 于是 Fiore 与 Fontana 两人相约 1535 年 2 月 12 日在威尼斯 (或米兰) 的一个大教堂内公开竞赛, 各出 30 道题给对方. 恰在竞赛前几天, Fontana 又会解了不含二次项的一元三次方程, 从而他仅用两个小时就做出了 Fiore 的全部问题, 而 Fiore 却失败了. Fontana 由此扬名, 也吸引来了意大利人 G. Cardano(1501∼1576) 于 1539 年向他请教. 在 Cardano 发誓保密的前提下, Fontana 将一元三次方程的解法以诗行的形式告诉了他; 很快在 1540 年, Cardano 的学生 L. Ferrari(1522∼1565) 在此基础上找到了一元四次方程的解法, 同样因为保密而不能发表. 转折点在 1543 年, Cardano 和 Ferrari 在访问 Bologna 大学时, 得知 Ferro 才是第一个解出一元三次方程的人, 终于他们在 1545 年将一元三、四次方程的解法在著作 “Arsmagna” 中正式公布.

我们现在陈述不含二次项的一元三次方程 $x^3+px+q=0$ 的解法: 记 $x=a+b$, 则原方程化为

$$(p+3ab)(a+b)=-q-(a^3+b^3).$$

特别地, 以下希望选取合适的 a,b, 满足 $p+3ab=0=-q-(a^3+b^3)$.

由上式可得 $(a^3-b^3)^2=(a^3+b^3)^2-4(ab)^3=q^2+\dfrac{4}{27}p^3$. 记 $\Delta=\dfrac{q^2}{4}+\dfrac{p^3}{27}$, 则 $a^3-b^3=\pm2\sqrt{\Delta}$, 故解出 $\begin{cases}a^3=-\dfrac{q}{2}\pm\sqrt{\Delta}\\ b^3=-\dfrac{q}{2}\mp\sqrt{\Delta}\end{cases}$. 固定这样的一组解 a,b, 可验证此时它们满足 $\begin{cases}(ab)^3=-\dfrac{p^3}{27}\\ a^3+b^3=-q\end{cases}$.

记 $x^3-1=(x-1)(x-\omega)(x-\omega^2)$, 则 $ab=-\dfrac{p}{3}\cdot\omega^k\,(k\in\{0,1,2\})$. 任取 $l\in\{0,1,2\}$, 并记 $\begin{cases}a_l=\omega^l a\\ b_l=\omega^{-l-k}b\end{cases}$, 则此时 a_l,b_l 满足上述要求, 故 $x_l=a_l+b_l\,(l\in\{0,1,2\})$ 为原方程的三个解.

在此基础上, 对于不含三次项的一元四次方程 $x^4+px^2+qx+r=0$, 我们试图将它配方化为两个一元二次方程的形式. 引入待定参数 t, 则原方程化为

$$(x^2+t)^2=(2t-p)x^2-qx+(t^2-r).$$

希望选取 t 使得方程右端为关于 x 的完全平方式, 即要求判别式 $(-q)^2-4(2t-p)(t^2-r)=0$. 这是一个关于 t 的一元三次方程, 固定它的一个根为 t, 则上述方程化为

$$(x^2+t)^2=(2t-p)\left(x-\frac{q}{2(2t-p)}\right)^2,$$

于是只需解两个一元二次方程 $x^2+t=\pm\sqrt{2t-p}\left(x-\dfrac{q}{2(2t-p)}\right)$ 即可.

J. L. Lagrange(1736∼1813) 认为一元三、四次方程求根公式的发现多少带有偶然性, 他力图用自然而统一的观点来求解代数方程. 在他 1770 年的长文 “Reflexions sur la resolution algebrique des équation” 中, Lagrange 通过引入 Lagrange 预解方程给出了新的办法.

对于一元 n 次方程 $x^n+a_{n-1}x^{n-1}+\cdots+a_0=(x-x_1)\cdots(x-x_n)$, 记 $t(\zeta)=x_1+\zeta x_2+\cdots+\zeta^{n-1}x_n$, 其中 ζ 为任意 n 次单位根, 则 $\forall 1\le k\le n$,

$$\frac{1}{n}\sum_{\zeta^n=1}\zeta^{1-k}t(\zeta)=\frac{1}{n}\sum_{\zeta^n=1}\zeta^{1-k}\sum_{j=1}^{n}\zeta^{j-1}x_j=\frac{1}{n}\sum_{j=1}^{n}x_j\sum_{\zeta^n=1}\zeta^{j-k}=x_k,$$

故对所有 n 次单位根 ζ 求得 $t(\zeta)$ 后, 即知 $x_1,\cdots,x_n$. 因此 Lagrange 建议先解预解方程

$$\prod_{i_1,\cdots,i_n}(x-(x_{i_1}+\zeta x_{i_2}+\cdots+\zeta^{n-1}x_{i_n}))=0,$$

其中 ζ 为任意 n 次单位根, 乘积取遍 $\{1,\cdots,n\}$ 的所有全排列. 注意这个预解方程的系数是关于 $x_1,\cdots,x_n$ 的对称多项式, 从而可写成关于 $x_1,\cdots,x_n$ 的初等对称多项式 $\sigma_k=\sum_{1\le i_1<\cdots<i_k\le n}x_{i_1}\cdots x_{i_n}\,(1\le k\le n)$ 的多项式; 由 Vieta 定理知, 这也是关于原方程系数 $a_{n-1},\cdots,a_0$ 的多项式.

例如, 对于一元二次方程 $x^2+bx+c=0$, 记其根为 x_1,x_2, 则相应的两个 Lagrange 预解方程为

$$(x-(x_1+x_2))(x-(x_2+x_1))=(x-(x_1+x_2))^2=0,$$

$$(x-(x_1-x_2))(x-(x_2-x_1))=x^2-(x_1-x_2)^2=0.$$

由 $\begin{cases}x_1+x_2=-b\\ x_1x_2=c\end{cases}$ 可求得这两个预解方程的解, 从而可解出 x_1,x_2. 类似地, 不含二次项的一元三次方程的预解方程是 6 次的, 但为关于 x^3 的一元二次方程; 不含三次项的一元四次方程的预解方程是 24 次的, 但可整理为关于 x^2 的一元三次方程.

但 Lagrange 惊奇地发现这一办法对于一元五次以上的方程失效. 例如, 不含四次项的一元五次方程的预解方程是 120 次的, 只能整理为关于 x^5 的 24 次方程, 这比原本的一元五次方程更困难! 于是 Lagrange 认为一元五次以上方程的求解是上帝向人类智慧的挑战.

直到十九世纪二十年代, 挪威数学家 N. H. Abel (1802∼1829) 证明了以下否定性结果: $n\ge5$ 次复系数一般代数方程不是根式可解的, 即利用复数和方程系数进行加减乘除和开方运算不可能得到方程的所有根.

幸运的是, 法国天才数学家 E. Galois(1811∼1832) 创造性地引入 “群 (group)” 的概念, 证明了方程根式可解的判别准则: 特征 0 的域 F 上多项式 $f(X)\in F[X]$ 的根可由根式解, 当且仅当 $f(X)\in F[X]$ 的分裂域 E 在域 F 上的 Galois 群是可解群. 于是, $n\ge5$ 次复系数一般代数方程不是根式可解的, 当且仅当 $S_n\,(n\ge5)$ 不是

可解群. 而后者基于一个更强的事实: $A_n\,(n\geq 5)$ 为单群. 人们普遍认为, Galois 的工作不仅标志着经典代数方程论的结束, 也促使代数学转向研究 “群” 这样抽象的结构.

1.1.2 群的定义与例子

定义 1.1.1 (幺半群) 设 M 是一个非空集合, $p: M\times M\to M$ 为一个满足结合律的映射, 且存在 $1\in M$ 满足 $p(1,a)=a=p(a,1)$, $\forall a\in M$, 则称三元组 $(M,p,1)$ 是一个**幺半群** (monoid).

注 1.1.1

(1) 若在上述定义中去掉 p 满足结合律的条件, 则称三元组 $(M,p,1)$ 是一个 “monad”.

(2) 若在上述定义中去掉 $1\in M$ 的条件, 则称二元组 (M,p) 是一个 “semigroup”. (这里的术语遵从 N. Jacobson “Basic Algebra I” 的习惯, 在有些文献中 “semigroup” 指我们定义的 “monoid”.)

(3) 在幺半群 $(M,p,1)$ 中幺元 1 是唯一的, 这是因为 $1=p(1,1')=1'$.

例 1.1.1

(1) $(\mathbb{N}^*,\cdot,1)\leq(\mathbb{N},\cdot,1)\leq(\mathbb{Z},\cdot,1)$ 为幺半群; $(\mathbb{N},+,0)\leq(\mathbb{Z},+,0)$ 为幺半群. (这里的记号与安师讲义略有不同, $\mathbb{N}$ 为自然数集, 特别地它包含 0; $\mathbb{N}^*$ 为正整数集, 特别地它不包含 0.)

(2) 设 $n\in\mathbb{N}^*$, 则 $(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})\leq(\mathbb{Z}/n\mathbb{Z},\cdot,\overline{1})$ 为幺半群; $(\mathbb{Z}/n\mathbb{Z},+,\overline{0})$ 为幺半群.

例 1.1.2 任取一个 ≥ 2 元的集合 S, 令 $p(a,b)=b$, $\forall a,b\in S$, 则 (S,p) 是一个 “semigroup”. 但 S 中不存在幺元 1, 满足 $p(a,1)=a$, $\forall a\in S$, 否则 $1=p(a,1)=a$, $\forall a\in S$, 即 S 中只有一个元, 矛盾! 因此 (S,p) 不可能成为一个幺半群. 注意此例也说明, 若在幺元的定义中只考虑 “左幺性”, 即 $p(1,a)=a$, $\forall a\in S$, 则这样的 “幺元” 可能不唯一! 但这一现象在群的情形可以避免, 见群定义的注.

定义 1.1.2 (群) 设 $(G,p,1)$ 是一个幺半群, 满足 $\forall a\in G$, $\exists b\in G$, $s.t.\ p(a,b)=1=p(b,a)$, 则称三元组 $(G,p,1)$ 是一个**群** (group).

注:

(1) 在上述定义中, 设 $a\in G$, 则满足 $p(a,b)=1=p(b,a)$ 的 $b\in G$ 称为 a 的逆元, 它由 a 唯一确定:
$b=p(b,1)=p(b,p(a,b'))=p(p(b,a),b')=p(1,b')=b'$.

(2) 群的最简单定义可陈述为: 设 G 为一个非空集合, p 为 G 上的一个二元运算满足结合律, 以及
① $\exists 1\in G,\forall a\in G,\ p(1,a)=a$; ② 固定①中的某个 1, 则 $\forall a\in G$, $\exists b\in G$, $s.t.\ p(b,a)=1$,
则三元组 $(G,p,1)$ 是一个群. (显然正式定义蕴含此简单定义; 现设 $(G,p,1)$ 为简单定义的群, 任取 $a\in G$, 则 $\exists b,c\in G$, $s.t.\ p(b,a)=1$, $p(c,b)=1$, 故 $a=p(1,a)=p(p(c,b),a)=p(c,p(b,a))=p(c,1)$. 因此 $p(a,b)=p(p(c,1),b)=p(c,p(1,b))=p(c,b)=1$, 且 $p(a,1)=p(a,p(b,a))=p(p(a,b),a)=p(1,a)=a$.)

例 1.1.3

(1) $(\mathbb{Q}^*,\cdot,1)\leq(\mathbb{R}^*,\cdot,1)\leq(\mathbb{C}^*,\cdot,1)$ 为群; $(\mathbb{Z},+,0)\leq(\mathbb{Q},+,0)\leq(\mathbb{R},+,0)\leq(\mathbb{C},+,0)$ 为群.

(2) 设 $n\in\mathbb{N}^*$, 则 $(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})$ 为群; $(\mathbb{Z}/n\mathbb{Z},+,\overline{0})$ 为群.

群论新手常苦恼于如何举出更多群的例子. 回顾群的诞生故事, 它最初的想法是描述某种对称性的现象, 于是几何图形的对称群与抽象集合的自同构群应当是最自然的例子.

例 1.1.4 (置换群与自同构群) 设 $n\geq 1$, 记 n 元置换群

$$S_n\coloneqq\left\{\sigma=\begin{pmatrix}1&2&\cdots&n\\ \sigma(1)&\sigma(2)&\cdots&\sigma(n)\end{pmatrix},\ \text{其中}\{\sigma(1),\sigma(2),\cdots,\sigma(n)\}=\{1,2,\cdots,n\}\right\},$$

其中二元运算为置换的复合 $\circ$, 幺元为恒同置换 id. 显然 $|S_n|=n!$. 以映射的语言, S_n 中元 σ 为集合 $\{1,2,\cdots,n\}$ 到自身的一个双射, 则 $(S_n,\circ,\mathrm{id})$ 可视为任意 n 元集合到自身的双射全体关于映射的复合构成的群. 一般地, 任给集合 S, 记 $\mathrm{Aut}(S)\coloneqq\{S\text{ 到自身的双射全体}\}$, $\circ$ 为映射的复合, 1_S 为恒同映射, 则三元组 $(\mathrm{Aut}(S),\circ,1_S)$ 构成了一个群, 称为 S 的**自同构群** (automorphism group). (特别地, 若 $S=\varnothing$, 则 $\mathrm{Aut}(S)=\{\text{空映射}\}$.)

例 1.1.5 (几何图形的对称群) 平面上的 n 边形中对称性最强的是正 n 边形, 其中当 $n=1$ 时可以看成是一个点加一个圈, 当 $n=2$ 时是两个点加连接两点的两条边. 以其重心为原点, 可以发现保持正 n 边形不变的平面旋转恰构成 n 阶**循环群** (cyclic group), 即

$$\mathcal{C}_n \cong \langle b | b^n = 1 \rangle ,$$

其中 b 为绕原点旋转 $\frac{2\pi}{n}$. 如果在三维空间中考虑, 绕经过原点和每个顶点的直线旋转 π 也保持正 n 边形不变, 任取其中一个记为 a, 则正 n 边形的对称群为 **二面体群** (dihedral group)

$$\mathcal{D}_n \cong \langle a, b | a^2 = b^n = (ab)^2 = 1 \rangle .$$

古希腊人很早就知道凸正多面体只有五个, 分别是正四面体、正方体、正八面体、正十二面体和正二十面体. Plato(约公元前 427∼ 约公元前 347) 对其进行了研究, 后者称为 Plato 多面体. 以下我们考虑凸正多面体的对称群.

首先, 注意到任何一个凸正多面体的每个面的中心也张成一个凸正多面体, 我们称这两个凸正多面体是对偶的. 例如正方体和正八面体、正十二面体和正二十面体是对偶的, 而正四面体是自对偶的. 其次, 注意到保持正多面体不变的旋转都经过其重心, 旋转轴与正多面体的交点必须是正多面体的顶点、棱的中点或面的中心. 于是通过初等几何的讨论, 我们得到

(1) 正四面体的对称群 (称为**正四面体群** (tetrahedral group)) 为

$$\mathcal{T} \cong A_4 = \langle a, b | a^2 = b^3 = (ab)^3 = 1 \rangle ,$$

即 S_4 中的偶置换构成的群.

(2) 正八面体的对称群 (称为**正八面体群** (octahedral group)) 为

$$\mathcal{O} \cong S_4 = \langle a, b | a^2 = b^3 = (ab)^4 = 1 \rangle .$$

(3) 正二十面体的对称群 (称为**正二十面体群** (icosahedral group)) 为

$$\mathcal{I} \cong A_5 = \langle a, b | a^2 = b^3 = (ab)^5 = 1 \rangle ,$$

即 S_5 中的偶置换构成的群.

(以上书写群的方式称为**群的表现** (presentation), 即具有 $\langle$生成元$|$生成关系$\rangle$ 的形式.)

记 SO(3) 为实三维空间中所有绕过原点的直线旋转构成的群. 上述分析实际上给出了 SO(3) 的五类有限子群: $\mathcal{C}_n\,(n\geq 1), \mathcal{D}_n\,(n\geq 2), \mathcal{T}\cong A_4, \mathcal{O}\cong S_4, \mathcal{I}\cong A_5$. 事实上, 可以证明 SO(3) 的任何有限子群都同构于这五类群之一. 这个分类结果可以用简单连通图来表示, 恰为图论中的连通和谐图分类 (一个连通图称为和谐的, 如果它存在一种正整数赋值使得每个顶点的值等于与其相邻顶点值的和的一半), 也一一对应于复单 Lie 代数中 ADE 分类的 Dynkin 图. 此现象称为 **Mckay 对应**.

以下的 Cayley 定理告诉我们, 任意群都可视为某个集合的自同构群的子群, 从而研究集合的自同构群具有非常重要的意义. 为严格叙述并证明此 Cayley 定理, 我们需引入群同态的概念.

定义 1.1.3 (群同态) 设 $(M,p,1)$, $(M',p',1')$ 为两个幺半群, 若存在映射 $\varphi\colon M\to M'$, 满足 $\varphi(1)=1'$, 以及 $\forall a,b\in M,\ \varphi(p(a,b))=p'(\varphi(a),\varphi(b))$, 则称 $\varphi\colon (M,p,1)\to(M',p',1')$ 为一个**幺半群同态**. 两个群之间的幺半群同态称为**群同态** (group homomorphism). 双射的 (幺半) 群同态称为 (幺半) 群同构 (group isomorphism).

注: 设 $\varphi\colon (G,p,1)\to(H,p',1')$ 为一个群同态, 则

(1) $\varphi(a)^{-1}=\varphi(a^{-1})$, $\forall a\in G$, 即 φ 保持群中逆元;

(2) $\ker(\varphi)\coloneqq\{a\in H\colon \varphi(a)=1'\}$ 是 $(G,p,1)$ 的一个 (正规) 子群, 称为 φ 的**核** (kernel);
$\mathrm{Im}(\varphi)\coloneqq\{a'\in H\colon \exists a\in G,\ s.t.\ a'=\varphi(a)\}$ 是 $(H,p',1')$ 的一个子群, 称为 φ 的**像** (image).

定理 1.1.1 (Cayley) *任意群均可同构于某个集合的自同构群的子群.*

证明: 注意左乘 $L\colon\ (G,p,1)\longrightarrow(\mathrm{Aut}(G),\circ,\mathrm{id})$ 是一个单的群同态, 这里 $(\mathrm{Aut}(G),\circ,\mathrm{id})$ 是关于集合 G 的自

$$a\longmapsto (L_a\colon b\mapsto p(a,b))$$

同构群. 于是群 $(G,p,1)$ 与它在左乘映射 L 下的像为同构的群, 而后者为集合 G 的自同构群 $(\mathrm{Aut}(G),\circ,\mathrm{id})$ 的子群. □

注: Cayley 定理提供了一种重要的观点, 即将抽象的群元实现为在某个集合上的作用, 而抽象代数中群的作用恰在于群作用.

在任意范畴 $\mathscr{C}$ 中, 记对象 A 的自同构群

$$\mathrm{Aut}_{\mathscr{C}}(A) := \{\varphi \in \mathrm{Hom}_{\mathscr{C}}(A,A)\colon \exists \psi \in \mathrm{Hom}_{\mathscr{C}}(A,A),\ s.t.\ \psi\circ\varphi = \mathrm{id}_A = \varphi\circ\psi\},$$

其中二元运算为态射的复合 $\circ$, 幺元为恒同态射 id_A. 例如以上讨论的都是集合范畴 **Set** 中对象的自同构群. 接着我们将讨论群范畴 **Group** 中对象的自同构群, 即任意群到自身的群同构全体关于映射的复合构成的群.

例 1.1.6 *记 $\mathrm{Aut}(G,p,1)$ 为群 $(G,p,1)$ 的自同构全体构成的群.*

(1) $\mathrm{Aut}(\mathbb{Q},+,0)\cong(\mathbb{Q}^*,\cdot,1)$; $\mathrm{Aut}(\mathbb{R},+,0)\cong\mathrm{Aut}_{\mathbb{Q}-\mathrm{Mod}}(\mathbb{R})$; $\mathrm{Aut}(\mathbb{C},+,0)\cong\mathrm{Aut}_{\mathbb{Q}-\mathrm{Mod}}(\mathbb{C})$;

(2) $\mathrm{Aut}(\mathbb{Q}^*,\cdot,1)\cong\mathrm{Aut}(\mathbb{Q},+,0)$; $\mathrm{Aut}(\mathbb{R}^*,\cdot,1)\cong\mathrm{Aut}(\mathbb{R},+,0)$; $\mathrm{Aut}(\mathbb{C}^*,\cdot,1)\cong$?;

(3) $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z},+,\overline{0})\cong(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})$;

(4) $\mathrm{Aut}(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})\cong$?;

(5) $$\mathrm{Aut}(S_n,\circ,\mathrm{id})\cong\begin{cases}(S_n,\circ,\mathrm{id}), & n\neq 2,6\\ (\mathcal{C}_1,\cdot,1), & n=2\\ (S_6,\circ,\mathrm{id})\rtimes(\mathcal{C}_2,\cdot,1), & n=6\end{cases}.$$

证明: (1) 设 $F=\mathbb{Q},\mathbb{R},\mathbb{C}$, $\varphi\in\mathrm{Aut}(F,+,0)$, 则求 φ 可以先求解域 F 上的 Cauchy 函数方程

$$\varphi(x+y)=\varphi(x)+\varphi(y),\ \forall x,y\in F.$$

当 $F=\mathbb{Q}$ 时, 可以确定 $\varphi(x)=\varphi(1)x$, $\forall x\in\mathbb{Q}$, 则 φ 为双射 $\Longleftrightarrow\varphi(1)\neq 0$, 故 $\mathrm{Aut}(\mathbb{Q},+,0)\xrightarrow[\varphi\longmapsto\varphi(1)]{\cong}(\mathbb{Q}^*,\cdot,1)$ 为群同构. 当 $F=\mathbb{R}$ 或 $\mathbb{C}$ 时, F 上的 Cauchy 函数方程存在非 F-线性解: 事实上, 线性空间的理论表明存在群同构 $\mathrm{Aut}(F,+,0)\cong\mathrm{Aut}_{\mathbb{Q}-\mathrm{Mod}}(F)$, 而 $|\mathrm{Aut}_{\mathbb{Q}-\mathrm{Mod}}(\mathbb{R})|$, $|\mathrm{Aut}_{\mathbb{Q}-\mathrm{Mod}}(\mathbb{C})|$ 均为不可数集.

(2) 设 $F=\mathbb{Q}$ 或 $\mathbb{R}$, 考虑群同构 $(F^*,\cdot,1)\xrightarrow[r\longmapsto(\frac{r}{|r|},\log|r|)]{\cong}(\mathcal{C}_2,\cdot,1)\times(F,+,0)$, 则 $\mathrm{Aut}(F^*,\cdot,1)\cong\mathrm{Aut}((\mathcal{C}_2,\cdot,1)\times(F,+,0))$.

进一步地, 设 $\varphi\in\mathrm{Aut}((\mathcal{C}_2,\cdot,1)\times(F,+,0))$, 由于 $(-1,0)\in\mathcal{C}_2\times F$ 是唯一的二阶元, 则 $\varphi(-1,0)=(-1,0)$. 记 $\varphi=(\varphi_1,\varphi_2)$, $\varphi_1=\pi_1\circ\varphi$, $\varphi_2=\pi_2\circ\varphi$ 均为群同态. 由于 $\varphi_1(\epsilon,0)=\epsilon$, $\forall\epsilon\in\mathcal{C}_2$, 则

$$\varphi_1(\epsilon,x)=\varphi_1((\epsilon,0)(\cdot,+)(1,x))=\varphi_1(\epsilon,0)\varphi_1(1,x)=\epsilon\varphi_1(1,x);$$

而 $\varphi_1(1,x)=\varphi_1(1,x/2)^2\geq 0$, 则 $\varphi_1(1,x)=1$, 故 $\varphi_1(\epsilon,x)=\epsilon$, $\forall(\epsilon,x)\in\mathcal{C}_2\times F$. 另外, 注意 $\forall(\epsilon,x)\in\mathcal{C}_2\times F$,

$$\varphi_2(\epsilon,x)=\pi_2(\varphi(\epsilon,x))=\pi_2(\varphi(\epsilon,x)(\cdot,+)(-1,0))=\pi_2(\varphi(\epsilon,x)(\cdot,+)\varphi(-1,0))=\pi_2(\varphi(-\epsilon,x))=\varphi_2(-\epsilon,x),$$

即 φ_2 与第一分量无关, 可记为 $\varphi_2\colon(F,+,0)\to(F,+,0)$. 因此, $\varphi(\epsilon,x)=(\epsilon,\varphi_2(x))$, $\forall(\epsilon,x)\in\mathcal{C}_2\times F$, 即存在群同构 $\mathrm{Aut}((\mathcal{C}_2,\cdot,1)\times(F,+,0))\xrightarrow[\varphi\longmapsto\varphi_2]{\cong}\mathrm{Aut}(F,+,0)$.

设 $F=\mathbb{C}$, 考虑交换图 $\begin{array}{ccc}(\mathbb{C},+,0) & \xrightarrow{\widetilde{\varphi}} & (\mathbb{C},+,0)\\ \downarrow\exp & & \downarrow\exp\\ (\mathbb{C}^*,\cdot,1) & \xrightarrow{\varphi} & (\mathbb{C}^*,\cdot,1)\end{array}$, 则 $\mathrm{Aut}(\mathbb{C},+,0)$ 中元均可下降为 $\mathrm{Aut}(\mathbb{C}^*,\cdot,1)$ 中元. 问题是反过来这样的提升是否总能做到? 至少加连续性条件后由覆叠提升定理知可行.

(3) 设 $\varphi\in\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z},+,\overline{0})$, 则 $\varphi(\overline{k})=k\varphi(\overline{1})$, $\forall k\in\{0,1,\cdots,n-1\}$, 故 φ 为双射 $\Longleftrightarrow\varphi(\overline{1})$ 为循环群 $(\mathbb{Z}/n\mathbb{Z},+,\overline{0})$ 的生成元, 即 $\varphi(\overline{1})=\overline{m}$, 其中 $\gcd(m,n)=1$. 因此 $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z},+,\overline{0})\xrightarrow[\varphi\longmapsto\varphi(\overline{1})]{\cong}(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})$ 为群同构.

(4) 这没有直接的刻画方式. 注意 $(\{\mathbb{Z}/n\mathbb{Z}\text{中的乘法可逆元}\},\cdot,\overline{1})$ 为有限 Abel 群, 可以分解成有限个素数幂次阶循环群的直积, 而后者的自同构群已在 (3) 中确定. 一般地, 设 $(G,p,1)=\prod_{i=1}^{n}(G_i,p_i,1_i)$ 为有限个 Abel 群的直积, 则

$$\mathrm{Aut}(G,p,1)\cong\{A\in\mathrm{GL}(n)\colon A_{ij}\in\mathrm{Hom}_{\mathbf{Group}}((G_i,p_i,1_i),(G_j,p_j,1_j))\}.$$

(5) 这是一个纯群论的计算, 可以参考 Dummit, Foote "Abstract Algebra" 的相应章节习题. □

习题 1.1 (群的简单性质)

(1) 设 (G,p) 为一个 "semigroup", 满足 $\forall a,b\in G$, 方程 $p(a,x)=b$ 与 $p(y,a)=b$ 均在 G 中有解, 则 (G,p) 可成为一个群.

(2) 设 (G,p) 为一个 "semigroup", 满足 $|G|<+\infty$ 以及左右消去律, 则 (G,p) 可成为一个群.
(注意 $|G|<+\infty$ 是必要的, 例如 $(\mathbb{N}^*,\cdot,1)$ 与 $(\mathbb{N},+,0)$ 均为无穷幺半群, 且满足左右消去律, 但不为群.)

参考文献与补注 1.1

(1) 关于代数方程发展史的部分, 可以参考孙智伟 "近世代数" 或维基百科.

(2) 关于群、群同态等定义定理的部分, 可以参考 N. Jacobson"Basic Algebra I".

(3) 关于几何图形的对称群的部分, 可以参考朱富海 "有限群表示论".

§1.2 域与线性空间

1.2.1 从群到域

我们回忆群的概念, 并引入范畴化的描述: 群是集合范畴 **Set** 中的群对象 (group object). 具体地说, 设 $\mathscr{C}$ 是一个具有有限乘积 (product) 和终对象 (terminal object)$*$ 的范畴, 则 $\mathscr{C}$ 中的一个群对象是指 $G\in\mathrm{obj}(\mathscr{C})$, 以及 $\mu\in\mathrm{Hom}_{\mathscr{C}}(G\times G,G)$, $\eta\in\mathrm{Hom}_{\mathscr{C}}(G,G)$, $\epsilon\in\mathrm{Hom}_{\mathscr{C}}(*,G)$, 满足以下交换图:

① (结合律)
$$\begin{array}{ccc} G\times G\times G & \xrightarrow{1\times\mu} & G\times G \\ \downarrow{\scriptstyle \mu\times 1} & & \downarrow{\scriptstyle \mu} \\ G\times G & \xrightarrow{\mu} & G \end{array}$$

② (幺元)
$$\begin{array}{ccccc} G\times * & \xrightarrow{1\times\epsilon} & G\times G\times G & \xleftarrow{\epsilon\times 1} & *\times G \\ & {\scriptstyle \pi_1}\searrow & \downarrow{\scriptstyle \mu} & \swarrow{\scriptstyle \pi_2} & \\ & & G & & \end{array}$$

③ (逆元)
$$\begin{array}{ccccc} G & \xrightarrow{1\times\eta} & G\times G & \xleftarrow{\eta\times 1} & G \\ \downarrow & & \downarrow{\scriptstyle \mu} & & \downarrow \\ * & \xrightarrow{\epsilon} & G & \xleftarrow{\epsilon} & * \end{array}$$

这种范畴化语言的优点在于, 它摆脱了所有关于集合中元素的表述, 而只用一些态射来刻画关系, 并很容易推广.

现在我们借助群的定义重写教材中域的定义:

定义 1.2.1 (域) 设 F 为一个非空集合, $0\neq 1\in F$, $+,\cdot\colon F\times F\to F$ 为两个二元运算, 满足 $(F,+,0)$, $(F^*,\cdot,1)$ 均为交换群, 且 $+,\cdot$ 之间具有分配律, 则称五元组 $(F,+,0;\cdot,1)$ 为一个**域** (field).

定义 1.2.2 (域同态) 设 $(F,+,0;\cdot,1)$, $(F',+',0';\cdot',1')$ 为两个域, 若存在映射 $\varphi\colon F\to F'$, 满足

$$\varphi\colon (F,+,0)\to(F',+',0'),$$
$$\varphi\big|_{F^*}\colon (F^*,\cdot,1)\to(F'^*,\cdot',1')$$

均为群同态, 则称 $\varphi\colon (F,+,0;\cdot,1)\to(F',+',0';\cdot',1')$ 为一个**域同态** (field homomorphism). 双射的域同态称为域同构 (field isomorphism).

注: 在上述定义中, 可直接验证任意域同态必为单射. (这里的术语遵从 Nathan Jacobson"Basic Algebra I" 的习惯, 在有些文献中域同态并不要求 $\varphi(1)=1'$, 此时任意非零的域同态必为单射.) 因此域同态 $\varphi\colon (F,+,0;\cdot,1)\to(F',+',0';\cdot',1')$ 的作用是将前者视为后者的一个子域.

例 1.2.1 记 $\mathrm{Aut}(F,+,0;\cdot,1)$ 为域 $(F,+,0;\cdot,1)$ 的自同构全体构成的群.

(1) $\mathrm{Aut}(\mathbb{Q},+,0;\cdot,1)=\{\mathrm{id}\}$; $\mathrm{Aut}(\mathbb{R},+,0;\cdot,1)=\{\mathrm{id}\}$; $\mathrm{Aut}(\mathbb{C},+,0;\cdot,1)=?$;

(2) 设 p 为素数, 记 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$, 则 $\mathrm{Aut}(\mathbb{F}_p,+,0;\cdot,1)=\{\mathrm{id}\}$;

(3) 记 $\mathbb{F}_{p^n}$ 为多项式 $X^{p^n}-X$ 在 $\mathbb{F}_p$ 上的分裂域, 则 $\mathrm{Aut}(\mathbb{F}_{p^n},+,0;\cdot,1)\cong\mathbb{Z}/n\mathbb{Z}$.

证明: (1) 由定义知 $\mathrm{Aut}(\mathbb{Q},+,0;\cdot,1)=\mathrm{Aut}(\mathbb{Q},+,0)\cap\mathrm{Aut}(\mathbb{Q}^*,\cdot,1)=\{\mathrm{id}\}$.

设 $\varphi\in\mathrm{Aut}(\mathbb{R},+,0;\cdot,1)$, 则 $\varphi\big|_{\mathbb{Q}}\in\mathrm{Aut}(\mathbb{Q},+,0;\cdot,1)=\{\mathrm{id}\}$; 又由 $\forall x\geq 0$, $\varphi(x)=\varphi(\sqrt{x})^2\geq 0$ 知, φ 单调增, 故 $\varphi=\mathrm{id}$, 即 $\mathrm{Aut}(\mathbb{R},+,0;\cdot,1)=\{\mathrm{id}\}$. (事实上这体现了 $\mathbb{R}$ 从 $\mathbb{Q}$ 的 Dedekind 构造过程.)

但 $\mathrm{Aut}(\mathbb{C},+,0;\cdot,1)$ 的情形将会复杂得多, 这里有两部分的原因. 第一部分是由 $\mathbb{C}/\overline{\mathbb{Q}}^{\mathrm{alg}}$ 的超越性造成的, 即任取 $\varphi_0\in\mathrm{Aut}(\overline{\mathbb{Q}}^{\mathrm{alg}},+,0;\cdot,1)$, 以及 $\mathbb{C}$ 在 $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 上的两组超越基 B_1,B_2, 均存在 $\varphi\in\mathrm{Aut}(\mathbb{C},+,0;\cdot,1)$, 使得 $\varphi|_{\overline{\mathbb{Q}}^{\mathrm{alg}}}=\varphi_0$, 且 $\varphi(B_1)=B_2$. 第二部分是由 $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 本身的性质造成的, 根据域论中的同构延拓定理, $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 中任意两个子域的域同构均可延拓为 $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 的域自同构. 例如任取一个 ≥ 2 次的不可约多项式 $p(X)\in\mathbb{Q}[X]$ 及其在 $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 中的两根 $\alpha_1\neq\alpha_2$, 则存在子域的域同构 $\mathbb{Q}(\alpha_1)\xrightarrow{\cong}\mathbb{Q}(\alpha_2)$, 且将 α_1 映为 α_2, 这样延拓后可得 $\overline{\mathbb{Q}}^{\mathrm{alg}}$ 的非平凡域自同构. (可以进一步参考 P. B. Yale"Automorphisms of the Complex Numbers" (1966).)

(2) 由定义知 $\mathrm{Aut}(\mathbb{F}_p,+,0;\cdot,1)=\mathrm{Aut}(\mathbb{F}_p,+,0)\cap\mathrm{Aut}(\mathbb{F}_p^*,\cdot,1)=\{\mathrm{id}\}$.

(3) 注意 Frobenius 映射 $\mathrm{Frob}_p\colon x\mapsto x^p$ 为 n 阶域自同构, 则 $\mathrm{Aut}(\mathbb{F}_{p^n},+,0;\cdot,1)\supseteq\langle\mathrm{Frob}_p\rangle$. 又域扩张 $\mathbb{F}_{p^n}/\mathbb{F}_p$ 为 Galois 扩张, 则 $|\mathrm{Aut}(\mathbb{F}_{p^n},+,0;\cdot,1)|=|\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p)|=[\mathbb{F}_{p^n}:\mathbb{F}_p]=n$, 故 $\mathrm{Aut}(\mathbb{F}_{p^n},+,0;\cdot,1)=\langle\mathrm{Frob}_p\rangle\cong\mathbb{Z}/n\mathbb{Z}$. □

命题 1.2.1 *任意域 F 都包含一个素域 (即不含任何真子域的域), 且该素域* $\cong\begin{cases}\mathbb{Q}, & \text{若}\mathrm{char}(F)=0\\ \mathbb{F}_p, & \text{若}\mathrm{char}(F)=p\end{cases}$.

证明: 设 $(F,+,0;\cdot,1)$ 为一个域, 记 $R_0:=\{n\cdot 1\colon n\in\mathbb{Z}\}$, 则 $R_0\cong\begin{cases}\mathbb{Z}, & \text{若}\mathrm{char}(F)=0\\ \mathbb{Z}/p\mathbb{Z}, & \text{若}\mathrm{char}(F)=p\end{cases}$ 为 (整) 环同构 (即作为加法群同构, 且作为乘法幺半群同构). 记 F_0 为 R_0 在 F 中生成的子域, 则

$$F_0=\begin{cases}\left\{\dfrac{n\cdot 1}{m\cdot 1}\colon n\in\mathbb{Z},m\in\mathbb{Z}\setminus\{0\}\right\}, & \text{若}\mathrm{char}(F)=0\\ \left\{\dfrac{n\cdot 1}{m\cdot 1}\colon n\in\mathbb{Z},m\in\mathbb{Z}\setminus p\mathbb{Z}\right\}, & \text{若}\mathrm{char}(F)=p\end{cases}\cong\begin{cases}\mathbb{Q}, & \text{若}\mathrm{char}(F)=0\\ \mathbb{F}_p, & \text{若}\mathrm{char}(F)=p\end{cases}$$

为域同构. 注意域 F 的任意子域均包含 F_0, 故上述 F_0 即 F 中的素域. □

1.2.2 线性空间及其子空间

我们借助群、域的定义回忆域上线性空间的定义:

定义 1.2.3 (线性空间) 设 F 是一个域, V 为一个非空集合, 若存在一个二元运算 $+\colon V\times V\to V$, 以及 $0\in V$, 满足 $(V,+,0)$ 为一个交换群; 以及一个数乘作用 $F\times V\to V$, 满足加法与数乘之间的分配律, 则称 V 为域 F 上的一个**线性空间** (linear space), 或称 V 为一个 **F-模** (module).

注:

(1) 在域上线性空间的定义中, 域 F 与集合 V 是两个分立的对象, 注意域中的零元 0_F 与线性空间中的零元 0_V 不同, 域中的加法也与线性空间中的加法不同, 域中的乘法也与线性空间上的数乘不同.

(2) 在安师讲义中, 线性空间的八条公理陈述如下:

① $\forall\alpha,\beta\in V,\alpha+\beta=\beta+\alpha$;
② $\forall\alpha,\beta,\gamma\in V,(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;
③ $\exists 0_V\in V,\ \forall\alpha\in V,\ \alpha+0_V=\alpha$;
④ $\forall\alpha\in V,\ \exists\beta\in V,\ \alpha+\beta=0_V$;
⑤ $\forall\alpha\in V,\ 1_F\alpha=\alpha$;
⑥ $\forall c_1,c_2\in F,\ \forall\alpha\in V,\ (c_1c_2)\alpha=c_1(c_2\alpha)$;
⑦ $\forall c\in F,\ \forall\alpha,\beta\in V,\ c(\alpha+\beta)=c\alpha+c\beta$;
⑧ $\forall c_1,c_2\in F,\ \forall\alpha\in V,\ (c_1+c_2)\alpha=c_1\alpha+c_2\alpha$.

在这种写法中, 值得注意的是零元与负元的位置在同侧 (都在右侧), 这导致这八条公理并不是独立的, 例如① 可由②~⑧推出 (可见以下的引理). 如果将零元与负元的位置写在异侧, 则① 不可由②~⑧推出.

引理 1.2.2 *在安师讲义的线性空间八条公理中, ① 可由②~⑧推出.*

证明: 由于 $\forall\alpha\in V,\ 0_F\alpha+0_F\alpha=(0_F+0_F)\alpha=0_F\alpha$, 则

$$0_F\alpha=0_F\alpha+0_V=0_F\alpha+(0_F\alpha+(-0_F\alpha))=(0_F\alpha+0_F\alpha)+(-0_F\alpha)=0_F\alpha+(-0_F\alpha)=0_V,$$

故 $0_V+\alpha=0_F\alpha+1_F\alpha=(0_F+1_F)\alpha=1_F\alpha=\alpha$. 另外,

$$\begin{aligned}(-1_F)\alpha&=(-1_F)\alpha+0_V=(-1_F)\alpha+(\alpha+(-\alpha))\\&=((-1_F)\alpha+\alpha)+(-\alpha)=((-1_F)\alpha+1_F\alpha)+(-\alpha)\\&=(-1_F+1_F)\alpha+(-\alpha)=0_F\alpha+(-\alpha)=0_V+(-\alpha)=-\alpha,\end{aligned}$$

则 $(-\alpha)+\alpha=(-1_F)\alpha+1_F\alpha=(-1_F+1_F)\alpha=0_F\alpha=0_V$.

因此 $\forall\alpha,\beta\in V$,

$$\begin{aligned}\alpha+(\beta+\alpha)+\beta&=(\alpha+\beta)+(\alpha+\beta)=1_F(\alpha+\beta)+1_F(\alpha+\beta)=(1_F+1_F)(\alpha+\beta)\\&=(1_F+1_F)\alpha+(1_F+1_F)\beta=(1_F\alpha+1_F\alpha)+(1_F\beta+1_F\beta)\\&=(\alpha+\alpha)+(\beta+\beta)=\alpha+(\alpha+\beta)+\beta,\end{aligned}$$

则

$$\begin{aligned}\beta+\alpha&=0_V+(\beta+\alpha)+0_V=(-\alpha+\alpha)+(\beta+\alpha)+(\beta+(-\beta))\\&=-\alpha+(\alpha+(\beta+\alpha)+\beta)+(-\beta)=-\alpha+(\alpha+(\alpha+\beta)+\beta)+(-\beta)\\&=(-\alpha+\alpha)+(\alpha+\beta)+(\beta+(-\beta))=0_V+(\alpha+\beta)+0_V=\alpha+\beta.\end{aligned}$$

□

命题 1.2.3 *设 V 为域 F 上的线性空间, $W_1,W_2\subseteq V$ 为真子空间, 证明: $W_1\cup W_2\neq V$.*

证明: 不妨设 $W_1\not\subseteq W_2$ 且 $W_2\not\subseteq W_1$, 取 $\alpha_1\in W_1\setminus W_2,\alpha_2\in W_2\setminus W_1$, 则 $\alpha_1+\alpha_2\in V\setminus(W_1\cup W_2)$. □

注: 注意上述命题对于三个真子空间的情形未必成立, 反例如取 $V=\mathbb{F}_2^2$, $W_1=\{(x,y)\in V:x=0\}$, $W_2=\{(x,y)\in V:y=0\}$, $W_3=\{(x,y)\in V:x=y\}$, 则 $W_1\cup W_2\cup W_3=V$.

我们考虑上述命题的可能推广, 这是线性覆盖方向的经典问题:

定理 1.2.4 (线性覆盖)

(1) *无限域上的任意线性空间不可写成它的有限个真子空间之并.*

(2) *任意域 F 上的有限维线性空间若可写成它的 $|I|$ 个真子空间之并, 则 $|I|\geq|F|$.*

证明: (1) 设 F 为无限域, $V_1,\cdots,V_s$ 为 V 的有限个真子空间, 以下对 $s\geq1$ 归纳证明结论. 当 $s=1$ 时结论显然; 假设 $s\geq2$ 且当 $(s-1)$ 时结论成立, 则由归纳假设知 $\exists\alpha\in V\setminus\bigcup\limits_{i=1}^{s-1}V_i$, $\beta\in V\setminus V_s$. 考虑 $\alpha+c\beta\,(c\in F)$ 为 V 中无限个不同元, 若 $V=\bigcup\limits_{i=1}^{s}V_i$, 则由抽屉原理知, $\exists1\leq i\leq s$, $c_1\neq c_2\in F$, $s.t.\ \alpha+c_1\beta,\alpha+c_2\beta\in V_i$, 故 $\beta=\dfrac{1}{c_1-c_2}((\alpha+c_1\beta)-(\alpha+c_2\beta))\in V_i$, 且 $\alpha=(\alpha+c_1\beta)-c_1\beta\in V_i$, 这样的 i 不存在, 矛盾!

(2) 设 V 为任意域 F 上的有限维线性空间, 且 $V=\bigcup\limits_{i\in I}W_i$, 其中 W_i 为 V 的真子空间. 通过增大每个 W_i, 不妨设 $\forall i\in I$, $\dim_F(W_i)=\dim_F(V)-1$. 以下对 $n=\dim_F(V)$ 归纳证明结论. 当 $n=1$ 时条件不成立, 则命题虚空成立. 假设 $n\geq2$ 且当 $(n-1)$ 时结论成立, 固定 $i_0\in I$, 取 $\{\alpha_1,\cdots,\alpha_{n-1}\}$ 为 W_{i_0} 的基, 以及 $\beta\in V\setminus W_{i_0}$.

考虑 V 中一族两两不同的 $(n-1)$ 维子空间

$$\{U_c\coloneqq\mathrm{Span}_F\{\alpha_1+c\beta,\alpha_2,\cdots,\alpha_{n-1}\}:c\in F\}.$$

记 $I_c\coloneqq\{i\in I:U_c=W_i\}$, $\forall c\in F$, 以及 $S\coloneqq\{c\in F:I_c\neq\varnothing\}$. 由选择公理知, 存在选择映射 $\{I_c\}_{c\in S}\longrightarrow\bigcup\limits_{c\in S}I_c$, $I_c\longmapsto i(c)\in I_c$

则也存在复合映射 $S\longrightarrow\bigcup\limits_{c\in S}I_c\subseteq I$, $c\longmapsto i(c)\in I_c$. 注意此复合映射为单射, 故 $|S|\leq|I|$. 若 $S=F$, 则结论成立. 若 $S\subsetneqq F$, 则取 $c_0\in F\setminus S$, 此时 $\forall i\in I$, $U_{c_0}\neq W_i$, 即 $U_{c_0}\cap W_i\subsetneqq U_{c_0}$. 而 $U_{c_0}=U_{c_0}\cap V=U_{c_0}\cap\bigcup\limits_{i\in I}W_i=\bigcup\limits_{i\in I}(U_{c_0}\cap W_i)$, 故由归纳假设知 $|I|\geq|F|$. □

注: 关于线性覆盖更一般的结果, 可见 Pete L. Clark "Covering Numbers in Linear Algebra"(2012). 例如, 无限域上的无限维线性空间总可写成它的可数个真子空间之并.

在一般情形中, 一族子空间的并未必是子空间, 但它可线性生成一个子空间, 即这族子空间的和; 一族子空间的交一定还是子空间. 于是对于任意线性空间 V, 它的所有子空间组成的集合 L 上按照两种二元运算 "+" 与 "∩" 构成一个**格** (lattice). 注意格结构本质上是一种序关系, 例如这里 L 在包含偏序下的上确界运算即 "+", 下确界运算即 "∩".

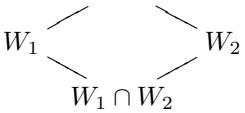
命题 1.2.5 (模律) *设 V 为线性空间, $N,M,W\subseteq V$ 为线性子空间, 若 $N\subseteq M$, 则 $M\cap(N+W)=N+(M\cap W)$.*
证明: 一方面, 由 $N\subseteq M$ 知 $N\subseteq M\cap(N+W)$, 且显然 $M\cap W\subseteq M\cap(N+W)$, 则 $N+M\cap W\subseteq M\cap(N+W)$. 另一方面, 设 $\alpha\in M\cap(N+W)$, 记 $\alpha=\alpha_1+\alpha_2$, 其中 $\alpha_1\in N,\alpha_2\in W$, 由 $N\subseteq M$ 知 $\alpha_2=\alpha-\alpha_1\in M\cap W$, 故 $\alpha=\alpha_1+\alpha_2\in N+(M\cap W)$, 这说明 $M\cap(N+W)\subseteq N+(M\cap W)$. □

注:

(1) 一般地, 如果去掉条件 $N\subseteq M$, 则只有一方面 $M\cap(N+W)\supseteq(M\cap N)+(M\cap W)$ 显然成立. 另一方面的反例如, 取 $V=\mathbb{R}^2$, $N=\{(x,y)\in V\colon x=0\}$, $M=\{(x,y)\in V\colon y=0\}$, $W=\{(x,y)\in V\colon x=y\}$, 则 $M\cap(N+W)=M\supsetneq\{(0,0)\}=(M\cap N)+(M\cap W)$.

(2) 注意上述模律的证明并未涉及域 F 在 V 上的数乘作用, 而只需 $(V,+,0)$ 的加法交换群结构, 于是它对于一般环上的模也成立. 对于任意环上的模, 它的所有子模组成的集合按照两种二元运算 "+" 与 "∩" 也构成一个格, 在格理论中满足上述模律的格称为**模格** (modular lattice).

一个经典的问题是, 有 n 个生成元的自由模格中有多少个元素? 以线性代数的语言, 在一个线性空间中由 n 个线性子空间按照 "+" 与 "∩" 运算至多可构造多少个线性子空间? 例如, 当 $n=1$ 时只有 1 个; 当 $n=2$ 时至多有 4 个:



(Hasse 图); 当 $n=3$ 时至多有 28 个; 当 $n\geq 4$ 时至多有 $+\infty$ 个. 这是 Dedekind 的经典结果!

参考文献与补注 1.2

(1) 关于范畴学基本定义的部分, 可以参考 N. Jacobson"Basic Algebra II".

(2) 关于域、域同态定义和更多域论的部分, 可以参考 Pete L. Clark"Field Theory".

(3) 关于线性空间公理的部分, 可以参考 V. Bruyant"Reducing Classical Axioms"(1971), 与 J. F. Rigby, J. Wiegold "Independent Axioms for Vector Spaces"(1973).

(4) 关于格论与自由模格的部分, 可以参考 N. Jacobson"Basic Algebra I", 与 R. Dedekind"Uber die von drei Moduln erzeugte Dualgruppe"(1900).

§1.3 线性空间的直和与商

本节主要讨论线性空间的一些构造方式, 包括线性空间的直积、(内外) 直和、对角线以及商.

1.3.1 线性空间的直积、直和、对角线

定义 1.3.1 (线性空间的直积) 设 $\{V_i\}_{i\in I}$ 为域 F 上的一族线性空间, 则诸 $V_i\,(i\in I)$ 的 Cartesian 积 $\prod\limits_{i\in I}V_i$ 按照逐分量的加法与数乘仍为域 F 上的线性空间, 称为 $\{V_i\}_{i\in I}$ 的**直积** (direct product).

注:

(1) 由集合 Cartesian 积的定义, $\prod\limits_{i\in I}V_i\coloneqq\left\{\alpha\colon I\to\bigcup\limits_{i\in I}V_i\middle|\,\forall i\in I,\ \alpha(i)\in V_i\right\}$; 为简化记号, 我们也将 $\prod\limits_{i\in I}V_i$ 中元写成 $(\alpha_i)_{i\in I}$, 其中 $\alpha_i\in V_i$.

(2) 以范畴论的语言, 线性空间的直积是线性空间范畴 **F-Mod** 中的一族对象的积 (product). 具体地说, 给定范畴 $\mathscr{C}$ 中的一族对象 $\{X_i\}_{i\in I}$, 它们的积是指 $X\in\mathrm{obj}(\mathscr{C})$, 以及 $\pi_i\in\mathrm{Hom}_{\mathscr{C}}(X,X_i)\,(i\in I)$, 满足以下的泛性质: 任给 $Y\in\mathrm{obj}(\mathscr{C})$, 以及 $f_i\in\mathrm{Hom}_{\mathscr{C}}(Y,X_i)\,(i\in I)$, 则存在唯一的 $f\in\mathrm{Hom}_{\mathscr{C}}(Y,X)$, 使得 $f_i=\pi_i\circ f$, $\forall i\in I$. 特别地, 在线性空间的直积中, π_i 可取为向第 i 个分量的投影映射 $:(\alpha_i)_{i\in I}\mapsto\alpha_i$.

例 1.3.1 直积空间多见于数列空间或函数空间. 例如, 设 F 为一个域, 则

(1) $F^\omega \coloneqq \{\text{数列}(a_n)_{n\geq 0}\colon a_n \in F\} = \{\text{级数}\sum\limits_{n=0}^{+\infty} a_n\colon a_n \in F\}$ 为可数个 F 的直积. 当 $F=\mathbb{R}$ 或 $\mathbb{C}$ 时, F^ω 有许多有趣的子空间, 如

$$\ell^p \coloneqq \{(a_n)_{n\geq 0} \in F^\omega\colon \left(\sum_{n=0}^{+\infty} |a_n|^p\right)^{1/p} < +\infty\}\,(1 \leq p < +\infty),$$
$$\ell^\infty \coloneqq \{(a_n)_{n\geq 0} \in F^\omega\colon \sup_{n\geq 0} |a_n| < +\infty\}.$$

(2) 任取一个非空集合 I, $F^I \coloneqq \{\text{映射} I \to F\}$ 为 $|I|$ 个 F 的直积. 当 $I \subseteq \mathbb{R}$ 为一个非退化区间, $F=\mathbb{R}$ 或 $\mathbb{C}$ 时, F^I 有许多有趣的子空间, 如

$$B(I;F) \supseteq R(I;F) \supseteq C^0(I;F) \supseteq C^k(I;F) \supseteq C^{k,\alpha}(I;F) \supseteq C^\infty(I;F) \supseteq C_c^\infty(I;F) \supseteq \cdots.$$

定义 1.3.2 (线性空间的直和) 设 $\{V_i\}_{i\in I}$ 为域 F 上的一族线性空间, 记 $\coprod\limits_{i\in I} V_i \coloneqq \left\{\alpha \in \prod\limits_{i\in I} V_i\colon |\mathrm{supp}(\alpha)| < +\infty\right\}$, 这里 $\mathrm{supp}(\alpha) \coloneqq \{i \in I\colon \alpha(i) \neq 0_{V_i}\}$, 则 $\coprod\limits_{i\in I} V_i$ 按照逐分量的加法与数乘为 $\prod\limits_{i\in I} V_i$ 的线性子空间, 称为 $\{V_i\}_{i\in I}$ 的**直和** (direct sum).

注:

(1) 若 $\forall i \in I$, $V_i \neq \{0\}$, 则 $\prod\limits_{i\in I} V_i = \coprod\limits_{i\in I} V_i \iff |I| < +\infty$.

(2) 以范畴论的语言, 线性空间的直和是线性空间范畴 **F-Mod** 中的一族对象的余积 (coproduct). 具体地说, 给定范畴 $\mathscr{C}$ 中的一族对象 $\{X_i\}_{i\in I}$, 它们的余积是指 $X \in \mathrm{obj}(\mathscr{C})$, 以及 $j_i \in \mathrm{Hom}_{\mathscr{C}}(X_i, X)\,(i \in I)$, 满足以下的泛性质: 任给 $Y \in \mathrm{obj}(\mathscr{C})$, 以及 $f_i \in \mathrm{Hom}_{\mathscr{C}}(X_i, Y)\,(i \in I)$, 则存在唯一的 $f \in \mathrm{Hom}_{\mathscr{C}}(X, Y)$, 使得 $f_i = f \circ j_i$, $\forall i \in I$. 特别地, 在线性空间的直和中, j_i 可取为向第 i 个分量的嵌入映射 : $\alpha_i \mapsto \left(\alpha\colon j \mapsto \begin{cases}\alpha_i, & \text{若} j = i\\ 0_{V_j}, & \text{若} j \neq i\end{cases}\right)$.

(3) 在通常的线性代数教材中, 上述定义的线性空间直和又叫做**外直和** (external direct sum), 是因为需要与以下的 **内直和** (internal direct sum) 相区别.

例 1.3.2 (域 F 上的一元多项式代数) 设 F 为一个域, 则 $\coprod\limits_{n=0}^{+\infty} F$ 为可数个 F 的直和. 我们试图在 $\coprod\limits_{n=0}^{+\infty} F$ 上以卷积的形式引入 "乘法": 令 $(a_0, a_1, \cdots, a_n, \cdots)\cdot(b_0, b_1, \cdots, b_m, \cdots) = (c_0, c_1, \cdots, c_{n+m}, \cdots)$, 其中

$$c_k \coloneqq \sum_{i+j=k} a_i b_j,$$

此时 $\coprod\limits_{n=0}^{+\infty} F$ 兼具线性空间与环的结构, 且这两者相容, 故为一个 F-**结合代数** (associative algebra), 称为域 F 上的一元多项式代数. 这是因为, 可将 $\coprod\limits_{n=0}^{+\infty} F$ 中元 $(a, 0, 0, \cdots)$ 视为 "常数"$a\,(a \in F)$, $(0, 1, 0, 0, \cdots)$ 视为 "未定元"X, 则上述 F-代数中的加法、数乘和乘法运算恰好对应通常多项式的加法、数乘和乘法运算, 此时 $\coprod\limits_{n=0}^{+\infty} F$ 也记为 $F[X]$.

定义 1.3.3 (线性子空间的内直和) 设 V 为域 F 上的线性空间, $\{V_i\}_{i\in I}$ 为 V 的一族线性子空间, 若对于子空间之和 $\sum\limits_{i\in I} V_i \coloneqq \left\{\sum\limits_{j=1}^{k} \alpha_{i_j} \in V\colon \alpha_{i_j} \in V_{i_j}\right\}$ 中的每个元 α, 都存在唯一的 $(\alpha_i)_{i\in I} \in \coprod\limits_{i\in I} V_i$, 使得 $\alpha = \sum\limits_{i\in I} \alpha_i$, 则称 $\sum\limits_{i\in I} V_i$ 为 $\{V_i\}_{i\in I}$ 的**内直和** (internal direct sum), 记为 $\bigoplus\limits_{i\in I} V_i$.

注:

(1) 在上述定义中, 语句 "$\alpha = \sum\limits_{i\in I} \alpha_i$" 的含义明确, 这是因为除了有限个 $i \in I$ 之外 α_i 均为 0, 于是 $\sum\limits_{i\in I} \alpha_i$ 实际上为有限和.

(2) 上述定义实际上直接给出了线性子空间的内、外直和同构: $\coprod\limits_{i\in I} V_i \longrightarrow \bigoplus\limits_{i\in I} V_i$, 于是在线性同构的意义下

$$(\alpha_i)_{i\in I} \longmapsto \sum_{i\in I} \alpha_i$$

可以不区分线性空间的内、外直和. 但严格地说, 线性空间的外直和是对于一族线性空间来谈的 (不涉及大空间), 而线性空间的内直和是对于给定大空间的一族线性子空间来谈的.

命题 1.3.1 设 V 为域 F 上的线性空间, $\{V_i\}_{i\in I}$ 为 V 的一族线性子空间, 则以下条件等价:

(1) $\sum\limits_{i\in I} V_i = \bigoplus\limits_{i\in I} V_i$;

(2) 映射 $\coprod\limits_{i\in I} V_i \longrightarrow \sum\limits_{i\in I} V_i$ 为线性同构;

$$(\alpha_i)_{i\in I} \longmapsto \sum_{i\in I} \alpha_i$$

(3) 对于 $0_V \in \sum\limits_{i\in I} V_i$, 存在唯一的 $(\alpha_i)_{i\in I} \in \coprod\limits_{i\in I} V_i$, 使得 $0_V = \sum\limits_{i\in I} \alpha_i$;

(4) $\forall i \in I$, $V_i \cap \sum\limits_{j\in I\setminus\{i\}} V_j = \{0_V\}$;

(5) $\forall \{i_1, \cdots, i_k\} \subseteq I$, $\sum\limits_{j=1}^{k} V_{i_j} = \bigoplus\limits_{j=1}^{k} V_{i_j}$;

(6) $\forall \{i_1, \cdots, i_k\} \subseteq I$, 映射 $\coprod\limits_{j=1}^{k} V_{i_j} \longrightarrow \sum\limits_{j=1}^{k} V_{i_j}$ 为线性同构;

$$(\alpha_i)_{i\in I} \longmapsto \sum_{j=1}^{k} \alpha_{i_j}$$

(7) $\forall \{i_1, \cdots, i_k\} \subseteq I$, 对于 $0_V \in \sum\limits_{j=1}^{k} V_{i_j}$, 存在唯一的 $(\alpha_{i_j})_{j=1}^{k} \in \coprod\limits_{j=1}^{k} V_{i_j}$, 使得 $0_V = \sum\limits_{j=1}^{k} \alpha_{i_j}$;

(8) $\forall \{i_1, \cdots, i_k\} \subseteq I$, $V_{i_1} \cap \sum\limits_{j=2}^{k} V_{i_j} = \{0_V\}$.

证明: (1)⇒(2)⇒(3) 显然;

现证 (3)⇒(4): $\forall i \in I$, 设 $\alpha_i = \sum\limits_{j=1}^{k} \alpha_{i_j} \in V_i \cap \sum\limits_{j\in I\setminus\{i\}} V_j$, 其中 $i_1, \cdots, i_k \in I \setminus \{i\}$ 两两不同, 则 $0_V = \sum\limits_{j=1}^{k} \alpha_{i_j} - \alpha_i$. 而 $0_V = \sum\limits_{i\in I} 0_V$, 故由 (3) 知, $\forall 1 \le j \le k$, $\alpha_{i_j} = 0_V = \alpha_i$, 则 $\forall i \in I$, $V_i \cap \sum\limits_{j\in I\setminus\{i\}} V_j = \{0_V\}$.

再证 (4)⇒(1): 任取 $\alpha \in \sum\limits_{i\in I} V_i$, 记 $\alpha = \sum\limits_{j=1}^{k} \alpha_{i_j}$, 令 $\alpha' \colon i \mapsto \begin{cases} \alpha_{i_j}, & 若 i = i_j, j = 1, \cdots, k \\ 0_V, & 若 i \ne i_1, \cdots, i_k \end{cases}$, 则 $\alpha' \in \coprod\limits_{i\in I} V_i$ 且 $\alpha = \sum\limits_{i\in I} \alpha'(i)$. 若 $\exists \beta \in \coprod\limits_{i\in I} V_i$, $s.t.\ \alpha = \sum\limits_{i\in I} \beta(i)$, 即 $\sum\limits_{j=1}^{k} \alpha_{i_j} = \sum\limits_{i\in I} \beta(i)$, 则 $\alpha_{i_1} - \beta(i_1) = \sum\limits_{i\in I\setminus\{i_1\}} \beta(i) - \sum\limits_{j=2}^{k} \alpha_{i_j} \in V_{i_1} \cap \sum\limits_{i\in I\setminus\{i_1\}} V_i$. 由 (4) 知 $\alpha_{i_1} - \beta(i_1) = 0_V = \sum\limits_{i\in I\setminus\{i_1\}} \beta(i) - \sum\limits_{j=2}^{k} \alpha_{i_j}$, 即 $\alpha_{i_1} = \beta(i_1)$ 且 $\sum\limits_{j=2}^{k} \alpha_{i_j} = \sum\limits_{i\in I\setminus\{i_1\}} \beta(i)$. 如此继续, 可知 $\forall 1 \le j \le k$, $\alpha_{i_j} = \beta(i_j)$, 且 $\forall i \in I \setminus \{i_1, \cdots, i_k\}$, $\beta(i) = 0$, 此即 $\alpha' = \beta$.

完全同理可证 (5)⇒(6)⇒(7)⇒(8)⇒(5). 最后 (3)⇔(7) 显然. □

注: 对于 V 的一族线性子空间 $\{V_i\}_{i\in I}$, 若仅满足 $\bigcap\limits_{i\in I} V_i = \{0_V\}$ 或者 $\forall i \ne j$, $V_i \cap V_j = \{0_V\}$, 并不能推出 $\sum\limits_{i\in I} V_i = \bigoplus\limits_{i\in I} V_i$. 反例如取 $V = \mathbb{R}^2$, $V_1 = \{(x, y) \in V \colon x = 0\}$, $V_2 = \{(x, y) \in V \colon y = 0\}$, $V_3 = \{(x, y) \in V \colon x = y\}$, 则 $(0, 0) = (0, 0) + (0, 0) + (0, 0) = (0, 1) + (1, 0) - (1, 1)$ 为两种 $(0, 0) \in \sum_{i=1}^{3} V_i$ 的表示方式.

在线性空间的构造中, 常被忽视的是对角线子空间, 它具有很强的几何直观.

定义 1.3.4 (对角线子空间) 设 V 为域 F 上的线性空间, 则 $\Delta(V) \coloneqq \{\alpha \in \prod\limits_{i\in I} V \big| \alpha$ 为常值映射$\}$ 按照逐分量的加法与数乘为 $\prod\limits_{i\in I} V$ 的线性子空间, 称为 $\prod\limits_{i\in I} V$ 的**对角线子空间** (diagonal subspace).

例 1.3.3 设 V 为域 F 上的线性空间, V_1, V_2 为 V 的线性子空间, 则 $\sum\limits_{i=1}^{2} V_i = V \iff \prod\limits_{i=1}^{2} V_i + \Delta(V) = \prod\limits_{i=1}^{2} V$.

证明: "⇒": 任取 $(\alpha_1, \alpha_2) \in \prod\limits_{i=1}^{2} V$, 由 $\alpha_1 - \alpha_2 \in V = \sum\limits_{i=1}^{2} V_i$ 知, $\exists (\beta_1, \beta_2) \in \prod\limits_{i=1}^{2} V_i$, $s.t.\ \alpha_1 - \alpha_2 = \beta_1 - \beta_2$.

再取 $\gamma = \alpha_1 - \beta_1 = \alpha_2 - \beta_2 \in V$, 则 $(\alpha_1, \alpha_2) = (\beta_1, \beta_2) + (\gamma, \gamma) \in \prod\limits_{i=1}^{2} V_i + \Delta(V)$, 故 $\prod\limits_{i=1}^{2} V_i + \Delta(V) = \prod\limits_{i=1}^{2} V$.

"⇔": 任取 $\alpha \in V$, 由 $(\alpha, 0) \in \prod\limits_{i=1}^{2} V = \prod\limits_{i=1}^{2} V_i + \Delta(V)$ 知, $\exists (\beta_1, \beta_2) \in \prod\limits_{i=1}^{2} V_i$, $\gamma \in V$, $s.t.\ (\alpha, 0) = (\beta_1, \beta_2) + (\gamma, \gamma)$,

则 $\alpha=\beta_1-\beta_2\in\sum\limits_{i=1}^{2}V_i$. □

注: 上例来源于微分拓扑中关于横截性的命题: "设 $f\colon M\to N$ 为光滑流形之间的光滑映射, $i\colon A\hookrightarrow N$ 为嵌入子流形, 则 $f\pitchfork A\iff(f\times i)\pitchfork\Delta(N)$." 这是对角线子空间几何意义的一次精彩呈现.

1.3.2 线性空间的商

取商的操作是许多数学范畴中常用的技巧, 即在原对象上引入一个等价关系, 则原对象关于此等价关系的所有等价类构成了一个商对象. 例如整数环 $\mathbb{Z}$ 上已有加法与乘法运算, 定义减法运算可视为在 $\mathbb{Z}\times\mathbb{Z}$ 上引入等价关系 $(a_1,b_1)\sim(a_2,b_2)\iff a_1+b_2=a_2+b_1$, 所得等价类 $[(a_1,b_1)]$ 称为 a_1 与 b_1(有序) 的差, 则 $\mathbb{Z}$ 关于差等价的商对象同构于 $\mathbb{Z}$; 定义除法运算可视为在 $\mathbb{Z}\times\mathbb{Z}^*$ 上引入等价关系 $(a_1,b_1)\sim(a_2,b_2)\iff a_1b_2=a_2b_1$, 所得等价类 $[(a_1,b_1)]$ 称为 a_1 与 b_1(有序) 的商, 则 $\mathbb{Z}$ 关于商等价的商对象同构于 $\mathbb{Q}$.

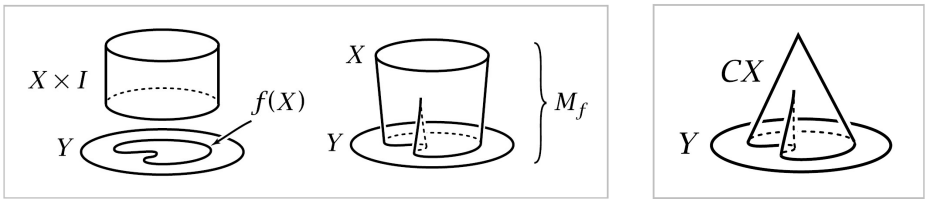
又如在几何中, 拓扑空间的商空间即将某个子空间粘合成一个点; 一般地, 设 $f\colon X\to Y$ 为拓扑空间之间的连续映射, 则 f 的映射柱

$$M_f\coloneqq(X\times I)\sqcup Y/[(x,0)\sim f(x)],$$

映射锥

$$C_f\coloneqq(X\times I)\sqcup Y/[(x,0)\sim f(x),(x,1)\sim(x',1)].$$

它们在同伦理论 (homotopy theory) 中是很重要的工具.



定义 1.3.5 (商空间) 设 V 为域 F 上的线性空间, W 为 V 的线性子空间, 在集合 V 上引入等价关系

$$\alpha_1\sim_W\alpha_2\iff\alpha_1-\alpha_2\in W,$$

则集合 V 关于此等价关系的商集为 $V/\sim_W\coloneqq\{[\alpha]\coloneqq\alpha+W\,\big|\,\alpha\in V\}$. 注意商集 $V/\sim_W$ 上仍可赋予域 F 上的线性空间结构: $\forall[\alpha],[\beta]\in V/\sim_W$, $\forall c\in F$, $[\alpha]+[\beta]\coloneqq[\alpha+\beta]$; $c[\alpha]\coloneqq[c\alpha]$, 称为 V 关于 W 的**商空间** (quotient space), 记为 V/W.

注:

(1) 利用 $W\subseteq V$ 为线性子空间, 可以验证上述等价关系 $\sim_W$ 以及商集 $V/\sim_W$ 上的加法与数乘都是定义良好的, 且满足线性空间的八条公理.

(2) 警告商空间一定不是子空间 (不是子集)! 例如, $V/\{0_V\}\xrightarrow{\cong}V$ 与 $V/V\xrightarrow{\cong}\{0_V\}$ 均为线性空间的同构, 而非线性空间的相同.

$$[\alpha]=\alpha+\{0_V\}\longmapsto\alpha\qquad[\alpha]=\alpha+V\longmapsto0_V$$

命题 1.3.2 *设 V 为域 F 上的线性空间, W 为 V 的线性子空间, 则存在 V 的另一线性子空间 U, 满足 $V=W\oplus U$, 且 $U\xrightarrow{\cong}V/W$ 为线性空间的同构.*

$$\alpha\longmapsto[\alpha]=\alpha+W$$

证明: 考虑集合族 $\mathscr{F}\coloneqq\{U\subseteq V\text{为线性子空间}\colon W\cap U=\{0_V\}\}$ 及其上的包含偏序, 由 $\{0_V\}\in\mathscr{F}$ 知 $\mathscr{F}\neq\varnothing$. 现设 $\{U_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链 (即一个全序子集), 则可直接验证 $\bigcup\limits_{i\in I}U_i=\sum\limits_{i\in I}U_i\subseteq V$ 也为线性子空间, 且 $W\cap\bigcup\limits_{i\in I}U_i=\{0_V\}$, 即 $\bigcup\limits_{i\in I}U_i$ 为 $\{U_i\}_{i\in I}$ 在 $\mathscr{F}$ 中的上界. 由 Zorn 引理知, $\mathscr{F}$ 中存在极大元, 记为 U. 断言: $V=W\oplus U$. 事实上, 假设 $\exists\alpha\in V\setminus(W\oplus U)$, 则可直接验证 $U+F\alpha\supsetneq U$ 且 $U+F\alpha\in\mathscr{F}$, 这与 U 在 $\mathscr{F}$ 中的极大性矛盾! 由此直和关系可直接验证 $U\xrightarrow{\cong}V/W$ 为线性空间的同构. □

$$\alpha\longmapsto[\alpha]=\alpha+W$$

注:

(1) 在上述命题中, 满足直和关系 $V=W\oplus U$ 的子空间 U 称为 W 在 V 中的**直和补空间** (direct complement). 它在线性同构的意义下由 V 关于 W 的商空间唯一确定, 于是可视为商空间 V/W 在原空间 V 中的一种实现方式.

(2) 但在线性空间相同的意义下, W 在 V 中的直和补空间不是唯一的. (上述证明过程无法提供唯一性, 这是因为 Zorn 引理考虑的偏序集中极大元一般不唯一.) 事实上, 反例如取 $V=\mathbb{R}^2$, $W=\{(x,y)\in V\colon x=0\}$, $U_1=\{(x,y)\in V\colon y=0\}$, $U_2=\{(x,y)\in V\colon x=y\}$, 则 $U_1\neq U_2$ 均为 W 在 V 中的直和补空间.

(3) 上述证明中 Zorn 引理偏序集的取法来源于半单模论中的技巧. 事实上, 域 F 上的线性空间是 F-**半单模** (semisimple module). 我们将会看到 (半单) 模论的观点和技术在线性代数的理论中几乎是致命的.

(4) 在上述证明中, 我们使用了等价于选择公理的 Zorn 引理来说明存在性, 这种方法将在下节讨论线性空间的基与维数时详细阐释. 值得注意的是, Zorn 引理通常还有另一种使用方法, 比如在上述命题中考虑集合族 $\mathscr{F}\coloneqq\{U\subseteq V\text{为线性子空间}\colon V=W+U\}$ 及其上的包含偏序, 由 $V\in\mathscr{F}$ 知 $\mathscr{F}\neq\varnothing$. 我们似应证明 $\mathscr{F}$ 中存在极小元, 而这无法直接使用 Zorn 引理得到. (事实上, 现设 $\{U_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链, 我们自然希望说明线性子空间 $\bigcap\limits_{i\in I}U_i\subseteq V$ 为 $\{U_i\}_{i\in I}$ 在 $\mathscr{F}$ 中的下界, 即证 $V=W+\bigcap\limits_{i\in I}U_i$, 而这由以下的构造知存在反例.)

例 1.3.4 *设 V 为域 F 上的线性空间, W 为 V 的线性子空间, $\{U_i\}_{i\in I}$ 为 V 的一条线性子空间链, 且满足 $V=W+U_i$, $\forall i\in I$, 但 $V=W+\bigcap\limits_{i\in I}U_i$ 未必成立.*

证明: 考虑 $V=\mathrm{Span}_F\{\alpha_i\}_{i=0}^{+\infty}$, 其中 $\{\alpha_i\}_{i=0}^{+\infty}$ 在域 F 上线性无关 (在下节线性空间的基与维数中很常见), $W\coloneqq\mathrm{Span}_F\{\alpha_0+\alpha_i\}_{i=1}^{+\infty}$, $U_i\coloneqq\mathrm{Span}_F\{\alpha_j\}_{j=i}^{+\infty}$ $(i\geq 1)$, 则 $V=W+U_i$, $\forall i\geq 1$, 但 $V\supsetneqq W$ 且 $\bigcap\limits_{i=1}^{+\infty}U_i=\{0_V\}$. □

注:

(1) 在此例中, 若线性子空间链 $\{U_i\}_{i\in I}$ 是有限长度的, 则 $V=W+\bigcap\limits_{i\in I}U_i$ 显然成立. 因此上述反例的构造使用了分析中 "将困难留在无穷远处" 的技巧.

(2) 此例实际上也说明了线性子空间运算 "+" 与 "∩" 的交换律一般不成立, 即 "$\bigcap\limits_{i\in I}(W+U_i)\supsetneqq W+\bigcap\limits_{i\in I}U_i$".

习题 1.3 (直积、直和的正合性) *设 $\{V_i\}_{i\in I}$ 为域 F 上的一族线性空间, W_i 为 V_i 的子空间 $(i\in I)$, 则*

(1) *存在线性同构* $\prod\limits_{i\in I}V_i/\prod\limits_{i\in I}W_i\xrightarrow{\cong}\prod\limits_{i\in I}(V_i/W_i)$;
$$(\alpha_i)_{i\in I}+\prod_{i\in I}W_i\longmapsto(\alpha_i+W_i)_{i\in I}$$

(2) *存在线性同构* $\coprod\limits_{i\in I}V_i/\coprod\limits_{i\in I}W_i\xrightarrow{\cong}\coprod\limits_{i\in I}(V_i/W_i)$.
$$(\alpha_i)_{i\in I}+\coprod_{i\in I}W_i\longmapsto(\alpha_i+W_i)_{i\in I}$$

参考文献与补注 1.3

(1) 关于范畴学中积与余积的部分, 可以参考 J. J. Rotman "An Introduction to Homological Algebra".

(2) 关于线性空间 (或者模) 内、外直和的部分, 可以参考 Dummit, Foote"Abstract Algebra".

(3) 关于流形横截性的部分, 可以参考 G. Pollack "Differential Topology".

(4) 关于拓扑空间映射柱、映射锥与同伦论的部分, 可以参考 A. Hatcher "Algebraic Topology".

§1.4 线性空间的基与维数

本节的主要目标是证明: "$\dim_F(V)\coloneqq|B|$, 其中 B 是域 F 上线性空间 V 的基" 是良定义的. 为此我们需要证明两部分: (1) 域上的线性空间总存在基; (2) 域上线性空间的任两基等势.

1.4.1 基的存在性: Zorn 引理

我们回忆域 F 上线性空间 V 的**基** (basis) 的概念, 即 V 中线性无关且可 F-线性生成 V 的子集. 直观上看, 这两个条件具有一定的 "互斥性", 即线性无关性要求基中元不能过多而产生线性冗余, 生成性要求基中元不能过少而遗漏部分方向. 由此先给出基的自然刻画:

命题 1.4.1 *设 V 为域 F 上的线性空间, $B \subseteq V$ 为子集, 则以下条件等价:*

(1) *B 为 V 的基;*

(2) *在包含偏序下 B 为极大的线性无关集;*

(3) *在包含偏序下 B 为极小的线性生成集.*

证明: "(1)⇒(2),(3)": 设 B 为 V 的基. 任取 $T \supsetneqq B$, 由 B 可线性生成 V 知, B 可线性表出 $T \setminus B$ 中元, 则 T 线性相关, 故 B 为极大线性无关的; 任取 $S \subsetneqq B$, 由 B 线性无关知, $B \setminus S$ 中元不可能由 S 线性表出, 则 S 不可线性生成 V, 故 B 为极小线性生成的.

"(2)⇒(1)": 设 $B \subseteq V$ 为极大的线性无关子集, 即 B 线性无关, 且 $\forall \alpha \in V \setminus B$, $B \cup \{\alpha\}$ 均线性相关, 则存在 $B \cup \{\alpha\}$ 的某个非空有限子集线性相关. 这一有限子集必包含 α, 否则它为 B 的子集, 这与 B 的线性无关性矛盾. 现取不全为零的 $c_1, \cdots, c_k, c \in F$, 以及 $\alpha_1, \cdots, \alpha_k \in B$, 满足 $\sum\limits_{i=1}^{k} c_i\alpha_i + c\alpha = 0_V$. 注意 $c \neq 0_F$, 否则由 $\sum\limits_{i=1}^{k} c_i\alpha_i = 0_V$ 以及 B 的线性无关性知 $c_1 = \cdots = c_k = 0_F = c$, 矛盾! 故 $\alpha = -c^{-1}\sum\limits_{i=1}^{k} c_i\alpha_i \in \mathrm{Span}_F(B)$. 这说明 $V \setminus B \subseteq \mathrm{Span}_F(B)$, 即 S 可线性生成 V, 故 B 为 V 的基.

"(3)⇒(1)": 设 $B \subseteq V$ 为极小的线性生成子集, 即 B 可线性生成 V, 且 $\forall \alpha \in B$, $B \setminus \{\alpha\}$ 不可线性生成 V. 假设 B 线性相关, 则 $B \neq \varnothing$ 且 B 的某个非空有限子集线性相关, 即 $\exists c_1, \cdots, c_k \in F$(不全为零), 以及 $\alpha_1, \cdots, \alpha_k \in B$, $s.t.$ $\sum\limits_{i=1}^{k} c_i\alpha_i = 0_V$. 取 $c_i \neq 0_F$, 则 $\alpha_i = -c_i^{-1}\sum\limits_{\substack{j=1 \\ j\neq i}}^{k} c_j\alpha_j$, 即 $\alpha_i \in B$ 可由 $B \setminus \{\alpha_i\}$ 线性表出, 故 $B \setminus \{\alpha_i\}$ 可线性表出 B 从而生成 V, 矛盾! 因此 B 线性无关, 故为 V 的基. □

于是以下只需说明 (1) V 中极大线性无关子集的存在性; 或 (2) V 中极小线性生成子集的存在性. 直观上看,

(1) 取 $\varnothing \subseteq V$ 为线性无关子集. 若 $V = \mathrm{Span}_F(\varnothing)(= \{0_V\})$, 则 $\varnothing$ 已是 V 中极大线性无关的; 若 $V \neq \mathrm{Span}_F(\varnothing)$, 则可取 $\alpha_1 \in V \setminus \mathrm{Span}_F(\varnothing)$, 此时 $\varnothing \cup \{\alpha_1\} = \{\alpha_1\} \subseteq V$ 为线性无关子集. 若 $V = \mathrm{Span}_F(\{\alpha_1\})$, 则 $\{\alpha_1\}$ 已是 V 中极大线性无关的; 若 $V \neq \mathrm{Span}_F(\{\alpha_1\})$, 则可取 $\alpha_2 \in V \setminus \mathrm{Span}_F(\{\alpha_1\})$, 此时 $\{\alpha_1\} \cup \{\alpha_2\} = \{\alpha_1, \alpha_2\} \subseteq V$ 为线性无关子集. 如此继续, 则此过程可能在某一时刻 (或无穷远时刻) 终止吗?

(2) 类似地, 取 $V \subseteq V$ 为线性生成子集, 显然 $V \setminus \{0_V\}$ 也为线性生成子集. 若 $V \setminus \{0_V\} = \varnothing$, 则 $V \setminus \{0_V\}$ 已是 V 中极小线性生成子集; 若 $V \setminus \{0_V\} \neq \varnothing$, 则可取 $\alpha_1 \in V \setminus \{0_V\}$, 此时 $(V \setminus \mathrm{Span}_F(\{\alpha_1\})) \cup \{\alpha_1\}$ 也为线性生成子集. 若 $V \setminus \mathrm{Span}_F(\{\alpha_1\}) = \varnothing$, 则 $\{\alpha_1\}$ 已是 V 中极小线性生成子集; 若 $V \setminus \mathrm{Span}_F(\{\alpha_1\}) \neq \varnothing$, 则可取 $\alpha_2 \in V \setminus \mathrm{Span}_F(\{\alpha_1\})$, 此时 $(V \setminus \mathrm{Span}_F(\{\alpha_1, \alpha_2\})) \cup \{\alpha_1, \alpha_2\}$ 也为线性生成子集. 如此继续, 则此过程可能在某一时刻 (或无穷远时刻) 终止吗?

为严格起见, 我们引入 Zorn 引理来描述这一过程.

引理 1.4.2 (Zorn) *设 X 为非空偏序集, 若 X 中的每条链 (即全序子集) 均在 X 中有上界, 则 X 中存在极大元.*

注: 安师讲义中已利用选择公理证明了 Zorn 引理, 思路与维基百科一致. 事实上, 在集合论的 ZF 公理系统下, Zorn 引理等价于选择公理, 等价于良序公理, 也等价于 Hausdorff 极大性原理. 此外, Zorn 引理还有许多应用, 例如可由 Zorn 引理证明任意线性空间均存在基; 任意非零交换环均存在极大理想; 以及拓扑学中的 Tychonoff 定理、泛函分析中的 Krein-Milman 定理. 令人惊奇的是这些应用也都能反过来推出 Zorn 引理.

最后我们利用 Zorn 引理证明以下稍显一般的命题, 从它的证明过程中可见 V 中极大线性无关子集的存在性.

命题 1.4.3 *设 V 为域 F 上的线性空间, $S \subseteq T \subseteq V$ 为子集, 若 S 线性无关, T 生成 V, 则存在 V 的基 B 满足 $S \subseteq B \subseteq T$.*

证明: 考虑集合族 $\mathscr{F} \coloneqq \{B \subseteq V\colon B\text{为线性无关集, 且}S \subseteq B \subseteq T\}$ 及其上的包含偏序, 由 $S \in \mathscr{F}$ 知 $\mathscr{F} \neq \varnothing$. 现设 $\{B_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链, 取 $B = \bigcup\limits_{i\in I} B_i \subseteq V$, 则 $S \subseteq B \subseteq T$. 断言 B 也为线性无关集. 这是因为, 若 $S = \varnothing$, 则 B 线性无关; 若 $B \neq \varnothing$, 任取 B 的非空有限子集 $\{\alpha_1, \cdots, \alpha_k\}$, 记 $\alpha_j \in B_{i_j}$, $j = 1, \cdots, k$, 则 $\{\alpha_1, \cdots, \alpha_k\} \subseteq \bigcup\limits_{j=1}^{k} B_{i_j} = B_i$, 其中 B_i 为 $\{B_{i_j}\}_{j=1}^{k}$ 中在包含全序下的最大元. 由 B_i 为线性无关集知, $\{\alpha_1, \cdots, \alpha_k\}$ 也为线性无关集, 因此 B 为线性无关集, 即 B 为 $\{B_i\}_{i\in I}$ 在 $\mathscr{F}$ 中的上界. 由 Zorn 引理知, $\mathscr{F}$ 有极大元, 记为 B. 断言 B 可线性生成 V. 这是因为, 若 $\mathrm{Span}_F(B) \supseteq T$, 则由 T 线性生成 V 知 B 也可线性生成 V; 若 $\mathrm{Span}_F(B) \not\supseteq T$, 则取 $\alpha \in T \setminus \mathrm{Span}_F(B)$, 此时 $B \cup \{\alpha\}$ 为线性无关集, 且 $S \subseteq B \cup \{\alpha\} \subseteq T$, 故 $B \cup \{\alpha\} \in \mathscr{F}$, 这与 B 的极大性矛盾! 因此 B 可线性生成 V, 故为 V 的基. □

推论 1.4.4 *设 V 为域 F 上的线性空间, 则:*

(1) *V 中任意线性无关集均可扩充为 V 的基.*

(2) *V 中任意线性生成集均可取子集为 V 的基.*

注: 类似地, 考虑集合族 $\mathscr{F} \coloneqq \{B \subseteq V\colon B\text{为线性生成集, 且}S \subseteq B \subseteq T\}$ 及其上的包含偏序, 由 $T \in \mathscr{F}$ 知 $\mathscr{F} \neq \varnothing$. 我们应证明 $\mathscr{F}$ 中存在极小元, 而这无法直接使用 Zorn 引理得到. (事实上, 现设 $\{B_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链, 我们自然希望说明子集 $\bigcap\limits_{i\in I} B_i \subseteq V$ 为 $\{B_i\}_{i\in I}$ 在 $\mathscr{F}$ 中的下界, 即证 $\bigcap\limits_{i\in I} B_i$ 为线性生成集, 而这由以下的构造知存在反例.)

例 1.4.1 *设 V 为域 F 上的线性空间, $\{B_i\}_{i\in I}$ 为 V 的一条子集链, 且满足 $V = \mathrm{Span}_F(B_i)$, $\forall i \in I$, 但 $V = \mathrm{Span}_F(\bigcap\limits_{i\in I} B_i)$ 未必成立.*

证明: 考虑 $V = \mathbb{R}$ 为 $\mathbb{R}$-线性空间, $B_i \coloneqq \{\alpha \in V : |\alpha| < 1/i\}\,(i \geq 1)$, 则 $V = \mathrm{Span}_{\mathbb{R}}(B_i)$, $\forall i \geq 1$, 但 $V \supsetneqq \{0_V\} = \mathrm{Span}_F(\{0_V\}) = \mathrm{Span}_F(\bigcap\limits_{i=1}^{+\infty} B_i)$. □

注:

(1) 在此例中, 若子集链 $\{B_i\}_{i\in I}$ 是有限长度的, 则 $V = \mathrm{Span}_F(\bigcap\limits_{i\in I} B_i)$ 显然成立. 因此上述反例的构造使用了分析中 "将困难留在无穷远处" 的技巧.

(2) 此例实际上也说明了子集运算 "Span_F" 与 "$\cap$" 的交换律一般不成立, 即 "$\bigcap\limits_{i\in I} \mathrm{Span}_F(B_i) \supsetneqq \mathrm{Span}_F(\bigcap\limits_{i\in I} B_i)$".

1.4.2 基的等势性: Steinitz 替换引理

为了证明线性空间中任两基等势, 我们并不直接构造任两基之间的双射, 而是证明任两基之间互相存在单射, 从而由以下的 Schröder-Bernstein 定理即知它们之间存在双射.

引理 1.4.5 (Schröder-Bernstein) *设 A, B 为两个集合, 若存在单射 $f\colon A \to B$, 以及单射 $g\colon B \to A$, 则存在双射 $h\colon A \to B$.*

证明: 令 $\mathscr{C} = \{C \subseteq A : C \cap g(B \setminus f(C)) = \varnothing\}$.

设 $C \in \mathscr{C}$, 则 $C \subseteq \bigcup\mathscr{C} \subseteq A$, 故 $f(C) \subseteq f(\bigcup\mathscr{C})$, 即 $B\setminus f(C) \supseteq B\setminus f(\bigcup\mathscr{C})$, 则 $g(B\setminus f(C)) \supseteq g(B\setminus f(\bigcup\mathscr{C}))$. 由 $C \cap g(B \setminus f(C)) = \varnothing$ 知, $C \cap g(B \setminus f(\bigcup\mathscr{C})) = \varnothing$, 即 $C \subseteq A \setminus g(B \setminus f(\bigcup\mathscr{C}))$.

记 $D = A \setminus g(B \setminus f(\bigcup\mathscr{C}))$, 则 $\bigcup\mathscr{C} = \bigcup\limits_{c\in\mathscr{C}} C \subseteq D$, 故 $f(\bigcup\mathscr{C}) \subseteq f(D)$, 即 $B \setminus f(\bigcup\mathscr{C}) \supseteq B \setminus f(D)$, 则 $g(B \setminus f(\bigcup\mathscr{C})) \supseteq g(B \setminus f(D))$, 即 $A \setminus D \supseteq g(B \setminus f(D))$, 故 $D \cap g(B \setminus f(D)) = \varnothing$. 因此 $D \in \mathscr{C}$, 则 $D \subseteq \bigcup\mathscr{C}$.

综上, $D = \bigcup\mathscr{C}$, 即 $A \setminus D = g(B \setminus f(D))$. 令 $h = f\big|_D \cup g^{-1}\big|_{g(B\setminus f(D))}\colon A \to B$, 则 h 为双射. □

注: 上述证明比维基百科上通常的证明简洁许多, 它来源于南京大学孙智伟教授的离散数学课程, 孙先生声称这是 A. Fraenkel 在 1954 年给出的方法, 但我们没有找到相应的参考文献, 于是只能将它暂时记录在这里.

现在证明线性空间 V 的任两基之间互相存在单射, 由基的定义知, 只需证明从 V 的任意线性无关集 S 到任意线性生成集 T 存在单射. 这依赖于著名的 Steinitz 替换引理, 以下先对 $|S|, |T|$ 均有限的情形完成证明.

定理 1.4.6 (Steinitz 替换引理的有限情形) *设 V 为域 F 上的线性空间, $S \subseteq V$ 为线性无关子集, $T \subseteq V$ 为线性生成子集, 且 $|S|, |T| < +\infty$, 则 $\forall 0 \leq k \leq \min\{|S|, |T|\}$, $\exists S_k \subseteq S$, $T_k \subseteq T$, $s.t.$ $|S_k| = |T_k| = k$, 且 $(T \setminus T_k) \cup S_k$ 也为线性生成子集.*

证明: 我们对 $0 \leq k \leq \min\{|S|, |T|\}$ 归纳证明结论. 当 $k = 0$ 时, 取 $S_0 = T_0 = \varnothing$ 即可, 结论显然.

现设 $1 \leq k \leq \min\{|S|, |T|\}$ 且当 $(k-1)$ 时结论成立, 即 $\exists S_{k-1} \subseteq S$, $T_{k-1} \subseteq T$, $s.t.$ $|S_{k-1}| = |T_{k-1}| = k-1$, 且 $(T \setminus T_{k-1}) \cup S_{k-1}$ 也为线性生成子集. 由于 $|S| \geq k > k-1 = |S_{k-1}|$, 可取 $\alpha \in S \setminus S_{k-1}$; 又 $(T \setminus T_{k-1}) \cup S_{k-1}$ 生成 V, 则 $\exists \beta_1, \cdots, \beta_t \in (T \setminus T_{k-1}) \cup S_{k-1}$ 两两不同, 以及 $c_1, \cdots, c_t \in F^*$, $s.t.$ $\alpha = \sum_{i=1}^{t} c_i \beta_i$. 注意 $\beta_1, \cdots, \beta_t$ 不可能全在 S_{k-1} 中, 否则 $\{\beta_1, \cdots, \beta_t, \alpha\} \subseteq S_{k-1} \cup \{\alpha\} \subseteq S$, 由 S 线性无关知 $\{\beta_1, \cdots, \beta_t, \alpha\}$ 也线性无关, 这与 α 可由 $\{\beta_1, \cdots, \beta_t\}$ 线性表出矛盾! 因此可取 $\beta = \beta_i \in (T \setminus T_{k-1}) \setminus S_{k-1}$.

现记 $S_k = S_{k-1} \cup \{\alpha\} \subseteq S$, $T_k = T_{k-1} \cup \{\beta\} \subseteq T$, 则 $|S_k| = |T_k| = k$, 且 $\beta = c_i^{-1}(\alpha - \sum_{\substack{j=1 \\ j \neq i}}^{k} c_j \beta_j) \in \mathrm{Span}_F((T \setminus T_k) \cup S_k)$. 又 $((T \setminus T_{k-1}) \setminus \{\beta\}) \cup S_{k-1} \subseteq (T \setminus T_k) \cup S_k$, 故 $(T \setminus T_{k-1}) \cup S_{k-1} \subseteq \mathrm{Span}_F((T \setminus T_k) \cup S_k)$. 因此, 由 $(T \setminus T_{k-1}) \cup S_{k-1}$ 生成 V 知 $(T \setminus T_k) \cup S_k$ 也生成 V. □

推论 1.4.7 *设 V 为域 F 上的线性空间, $S \subseteq V$ 为线性无关子集, $T \subseteq V$ 为线性生成子集, 且 $|S|, |T| < +\infty$, 则 $|S| \leq |T|$.*

证明: 假设 $|S| > |T|$, 则在 Steinitz 替换引理的有限情形中取 $k = |T|$ 知, $\exists S_k \subseteq S$, $s.t.$ $|S_k| = k$ 且 S_k 为线性生成子集. 由 $|S| > k = |S_k|$ 知, 可取 $\alpha \in S \setminus S_k$; 又 α 可由 S_k 线性表出, 这与 S 线性无关矛盾! □

在此基础上, 我们直接证明上述推论中 $|S|, |T|$ 均有限的条件可以去掉.

命题 1.4.8 *设 V 为域 F 上的线性空间, $S \subseteq V$ 为线性无关子集, $T \subseteq V$ 为线性生成子集, 则 $|S| \leq |T|$.*

证明: 考虑集合族

$$\mathscr{F} \coloneqq \{(C, f) \colon C \subseteq T,\ f \colon C \to S\text{为单射, 且}(S \setminus f(C)) \cup C\text{线性无关}\}$$

及其上的偏序

$$(C_1, f_1) \preccurlyeq (C_2, f_2) \Longleftrightarrow C_1 \subseteq C_2 \text{ 且} f_2\big|_{C_1} = f_1.$$

由 $(\varnothing, \text{空映射}) \in \mathscr{F}$ 知 $\mathscr{F} \neq \varnothing$. 现设 $\{(C_i, f_i)\}_{i \in I}$ 为 $\mathscr{F}$ 中的一条链, 记 $C = \bigcup_{i \in I} C_i$, $f = \bigcup_{i \in I} f_i$, 则 $C \subseteq T$, $f \colon C \to S$ 定义良好且为单射, 且 $\forall i \in I$, $(S \setminus f(C)) \cup C_i \subseteq (S \setminus f(C_i)) \cup C_i$ 为线性无关集, 故由命题 1.4.3的证明知, $(S \setminus f(C)) \cup C = \bigcup_{i \in I} (S \setminus f(C)) \cup C_i$ 也为线性无关集. 因此 $(C, f) \in \mathscr{F}$ 为 $\{(C_i, f_i)\}_{i \in I}$ 的上界. 由 Zorn 引理知, $\mathscr{F}$ 中存在极大元, 记为 (C, f). 断言: $S \setminus f(C) \subseteq C$.

这是因为, 假设 $S \setminus f(C) \not\subseteq C$, 则可取 $\alpha \in S \setminus (f(C) \cup C)$. 记 $S' = (S \setminus f(C)) \cup C$ 为线性无关集, 则 $S' \setminus \{\alpha\}$ 也为线性无关集. 若 $T \setminus C \subseteq \mathrm{Span}_F(S' \setminus \{\alpha\})$, 则

$$\alpha \in V = \mathrm{Span}_F(T) = \mathrm{Span}_F(T \setminus C) + \mathrm{Span}_F(C) \subseteq \mathrm{Span}_F(S' \setminus \{\alpha\}),$$

这与 S' 线性无关矛盾! 故 $T \setminus C \not\subseteq \mathrm{Span}_F(S' \setminus \{\alpha\})$, 可取 $\beta \in T \setminus (C \cup \mathrm{Span}_F(S' \setminus \{\alpha\}))$, 则 $(S' \setminus \{\alpha\}) \cup \{\beta\}$ 线性无关. 此时 $C \cup \{\beta\} \subseteq T$; 令 $g \colon C \cup \{\beta\} \subseteq T \to S$ 为 $g\big|_C = f$ 且 $g(\beta) = \alpha$, 则 g 为单射; 并且

$$(S \setminus g(C \cup \{\beta\})) \cup (C \cup \{\beta\}) = (S \setminus (f(C)) \cup \{\alpha\}) \cup C \cup \{\beta\} = ((S \setminus f(C) \cup C) \setminus \{\alpha\}) \cup \{\beta\} = (S' \setminus \{\alpha\}) \cup \{\beta\}$$

为线性无关集, 即 $(C \cup \{\beta\}, g) \in \mathscr{F}$, 这与 (C, f) 的极大性矛盾!

因此, $|S| \leq |f(C)| + |C| = 2|C| \leq 2|T|$. 若 $|T|$ 为无穷势, 则 $|S| \leq 2|T| = |T|$; 若 $|T| < +\infty$, 则 $|S| < +\infty$, 此时可由推论 1.4.7知 $|S| \leq |T|$. □

至此, 我们已经完成了线性空间中任两基等势的证明. 注意在这个过程中, 我们并未用到 Steinitz 替换引理的一般情形. 为完整起见, 以下将由 “线性空间中任两基等势” 证明 Steinitz 替换引理的一般情形.

定理 1.4.9 (Steinitz 替换引理的一般情形) *设 V 为域 F 上的线性空间, $S \subseteq V$ 为线性无关子集, $T \subseteq V$ 为线性生成子集, 则任取 S 的子集 S', 都存在单射 $f \colon S' \to T$, 使得 $(T \setminus f(S')) \cup S'$ 也为线性生成子集.*

证明: 由于线性无关集的子集仍为线性无关集, 故可不妨设 $S' = S$. 由于任意线性生成集总可取子集为基, 故可不妨设 T 为基. 于是以下只需证明: “设 V 为域 F 上的线性空间, $S \subseteq V$ 为线性无关子集, $B \subseteq V$ 为基, 则存在单射 $f \colon S \to B$, 使得 $(B \setminus f(S)) \cup S$ 为无交并, 且为 V 的基.”

考虑集合族 $\mathscr{F} \coloneqq \{C \subseteq B\colon C \cup S$为无交并, 且为线性无关集$\}$ 及其上的包含偏序, 由 $\varnothing \in \mathscr{F}$ 知 $\mathscr{F} \neq \varnothing$. 现设 $\{C_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链, 记 $C = \bigcup\limits_{i\in I} C_i \subseteq B$, 则 $C \cup S$ 也为无交并, 且 $C \cup S = \bigcup\limits_{i\in I}(C_i \cup S)$ 为线性无关集, 故 $C \in \mathscr{F}$ 为 $\{C_i\}_{i\in I}$ 的上界. 由 Zorn 引理知, $\mathscr{F}$ 中存在极大元, 记为 C. 断言: $C \cup S$ 可线性生成 V. 这是因为, 若 $\mathrm{Span}_F(C \cup S) \supseteq B$, 则由 B 线性生成 V 知 $C \cup S$ 可线性生成 V; 若 $\mathrm{Span}_F(C \cup S) \not\supseteq B$, 则可取 $\alpha \in B \setminus \mathrm{Span}_F(C \cup S)$, 故 $C \cup \{\alpha\} \subseteq B$, 且 $(C \cup \{\alpha\}) \cup S$ 为无交并且为线性无关集, 即 $C \cup \{\alpha\} \in \mathscr{F}$, 这与 C 的极大性矛盾! 因此, $C \cup S$ 可线性生成 V, 故为 V 的基.

注意 $C \cup (B \setminus C)$ 也为无交并, 且为 V 的基, 断言: $|S| = |B \setminus C|$. 这是因为, $V = \mathrm{Span}_F(C) \oplus \mathrm{Span}_F(S)$ 且 $V = \mathrm{Span}_F(C) \oplus \mathrm{Span}_F(B \setminus C)$, 这里 $\mathrm{Span}_F(S)$ 与 $\mathrm{Span}_F(B \setminus C)$ 均为 $\mathrm{Span}_F(C)$ 在 V 中的直和补空间, 故它们线性同构, 则它们的基等势, 即 $|S| = |B \setminus C|$. 于是可取单射 $f\colon S \to B$, 使得 f 的像集为 $B \setminus C$, 从而满足要求. □

1.4.3 维数的简单应用

设 V 为域 F 上的线性空间, 我们首先从基与维数的定义出发讨论 $|V|$, $|F|$, $\dim_F(V)$ 三者之间的关系.

引理 1.4.10 *设 V 为域 F 上的线性空间, 则* $|V| = \begin{cases} |F|^{\dim_F(V)}, & \text{若 } \dim_F(V) < +\infty \\ |F| \cdot \dim_F(V), & \text{若 } \dim_F(V) \text{为无穷势} \end{cases}$.

证明: 若 $\dim_F(V) = n < +\infty$, 取 $B = \{\alpha_i\}_{i=1}^n$ 为 V 的基, 则存在双射 $F^B \longrightarrow V$,

$$(c_1, \cdots, c_n) \longmapsto \sum_{i=1}^n c_i\alpha_i$$

故 $|V| = |F^B| = |F|^{|B|} = |F|^{\dim_F(V)}$.

若 $\dim_F(V)$ 为无穷势, 取 B 为 V 的基, 记 $S \coloneqq \{c\alpha \in V\colon c \in F, \alpha \in B\}$, 则 V 为由 S 生成的加法 Abel 群. 断言 $|V| = |S|$. 这是因为, 显然 $|V| \geq |S|$; 另一方面, 考虑满射 $\bigcup\limits_{n\in\mathbb{N}} S^n \longrightarrow V$, 则

$$(\alpha_1, \cdots, \alpha_n) \longmapsto \sum_{i=1}^n \alpha_i$$

$|V| \leq \left|\bigcup\limits_{n\in\mathbb{N}} S^n\right| = \sum\limits_{n\in\mathbb{N}} |S|^n$. 由 $\dim_F(V)$ 为无穷势知, $|S|$ 也为无穷势, 则上式 $= \sum\limits_{n\in\mathbb{N}} |S| = |\mathbb{N}| \cdot |S| = |S|$. 因此 $|V| = |S| = |F| \cdot |B| = |F| \cdot \dim_F(V)$. □

设 B 为线性空间 V 的基, 则由定义知 $\coprod\limits_{\alpha\in B} F \longrightarrow V$ 为线性同构, 即任意线性空间总可视为以基为

$$(c_\alpha)_{\alpha\in B} \longmapsto \sum_{\alpha\in B} c_\alpha\alpha$$

指标集的域上直和空间. 考虑其对偶即得线性同构 $V^* \longrightarrow \prod\limits_{\alpha\in B} F =: F^B$, 即任意线性空间的对偶空间总可视为以基为指标集的域上直积空间. 以下我们进一步讨论这些空间的维数.

引理 1.4.11 (Erdos-Kaplansky) *设 F 为一个域, 则 $\dim_F(F^I) = |F^I|$.*

证明: 可见 Nathan Jacobson "Lectures in Abstract Algebra: II. Linear Algebra". □

推论 1.4.12 *设 V 为域 F 上的线性空间, 则 $\dim_F(V^*) = |V^*|$.*

命题 1.4.13 (直积空间的维数) *设 $\{V_i\}_{i\in I}$ 为域 F 上的一族线性空间, 则*

$$\dim_F(\prod_{i\in I} V_i) = \begin{cases} \sum\limits_{i\in I} \dim_F(V_i), & \text{若} |\{i \in I\colon V_i \neq \{0\}\}| < +\infty \\ \left|\prod\limits_{i\in I} V_i\right|, & \text{若} |\{i \in I\colon V_i \neq \{0\}\}| \text{为无穷势} \end{cases}.$$

证明: 可不妨设 $V_i \neq \{0\}$, $\forall i \in I$. 若 $|I| < +\infty$, 则显然 $\dim_F(\prod\limits_{i\in I} V_i) = \sum\limits_{i\in I} \dim_F(V_i)$. 若 $|I|$ 为无穷势, 则一方面, 显然 $\dim_F(\prod\limits_{i\in I} V_i) \leq \left|\prod\limits_{i\in I} V_i\right|$; 另一方面, $\dim_F(\prod\limits_{i\in I} V_i) \geq \dim_F(F^I) = |F^I| \geq |F|$. 又 $\dim_F(\prod\limits_{i\in I} V_i)$ 为无穷势, 则 $\dim_F(\prod\limits_{i\in I} V_i) = \dim_F(\prod\limits_{i\in I} V_i)^2 \geq |F| \cdot \dim_F(\prod\limits_{i\in I} V_i) = \left|\prod\limits_{i\in I} V_i\right|$. □

注: 对于一般的直积空间, 虽然以上命题确定了其维数, 但我们很难显式写出它的基.

例 1.4.2 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则 $\ell^p\,(1\leq p<+\infty)$ 与 ℓ^∞ 都不是可数维线性空间.

证明: 在 $\ell^p\,(1\leq p<+\infty)$ 上引入度量 d:

$$d((a_n)_{n\geq 0},(b_n)_{n\geq 0}):=\left(\sum_{n=0}^{+\infty}|a_n-b_n|^p\right)^{1/p},\ \forall (a_n)_{n\geq 0},(b_n)_{n\geq 0}\in\ell^p;$$

在 ℓ^∞ 上引入度量 d:

$$d((a_n)_{n\geq 0},(b_n)_{n\geq 0}):=\sup_{n\geq 0}|a_n-b_n|,\ \forall (a_n)_{n\geq 0},(b_n)_{n\geq 0}\in\ell^\infty,$$

则 $\ell^p\,(1\leq p<+\infty)$ 与 ℓ^∞ 均为 Banach 空间, 记为 X. 显然 X 不是有限维线性空间. 假设 X 存在可数基 $\{\alpha_n\}_{n=1}^{+\infty}$, 记 $X_n=\mathrm{Span}_F(\{\alpha_i\}_{i=1}^n)\,(n\geq 1)$, 则 $X_n\subseteq X$ 为真闭子空间, 故 $\mathrm{int}(X_n)=\varnothing$. 由 Baire 纲定理知, $X=\bigcup\limits_{n=1}^{+\infty}X_n$ 也满足 $\mathrm{int}(X)=\varnothing$, 矛盾! 因此 X 不是可数维线性空间. □

命题 1.4.14 (直和空间的维数) 设 $\{V_i\}_{i\in I}$ 为域 F 上的一族线性空间, 取 V_i 的基为 $B_i\,(i\in I)$, 则 $\coprod\limits_{i\in I}V_i$ 的基为 $\bigsqcup\limits_{i\in I}B_i$, 这里 $\bigsqcup\limits_{i\in I}B_i:=\{\alpha\colon I\to\bigcup\limits_{i\in I}V_i\big|\,\exists i\in I,\ s.t.\ \alpha(i)\in B_i,\ \alpha\big|_{I\setminus\{i\}}=0\}$, 故 $\dim_F(\coprod\limits_{i\in I}V_i)=\left|\bigsqcup\limits_{i\in I}B_i\right|$.

例 1.4.3 域 F 上的一元多项式代数 $F[X]$ 是可数维 F-线性空间, 它的基为 $\{1,X,X^2,\cdots\}$.

命题 1.4.15 (对角线子空间的维数) 设 V 为域 F 上的线性空间, 则 $V\longrightarrow\Delta(V)$ 为线性同构, 故

$$\alpha\longmapsto(\alpha)_{i\in I}$$

$\dim_F(\Delta(V))=\dim_F(V)$.

命题 1.4.16 (商空间的维数) 设 V 为域 F 上的线性空间, $W\subseteq V$ 为线性子空间, 取 W 的基 B' 并扩充为 V 的基 B, 则 $\mathrm{Span}_F(B\setminus B')$ 为 W 在 V 中的一个直和补空间, 故 $\dim_F(V/W)=|B\setminus B'|$.

以下我们研究域论中一些涉及维数的例子.

例 1.4.4 (域扩张可视为基域上的线性空间) 设 E/F 为域扩张, 则 E 可视为域 F 上的线性空间, 记

$$[E\colon F]:=\dim_F(E)$$

为域扩张的次数. 例如 $\mathbb{C}/\mathbb{R}$ 为二次扩张; $\mathbb{R}/\mathbb{Q}$ 为不可数次扩张; $\mathbb{F}_{p^n}/\mathbb{F}_p$ 为 n 次扩张.

进一步地, 设 $K\supseteq E\supseteq F$ 为域扩张链, 取 K 的 E-基为 $\{\alpha_i\}_{i\in I}$, E 的 F-基为 $\{\beta_j\}_{j\in J}$, 则 K 的 F-基为 $\{\alpha_i\cdot\beta_j\}_{i\in I,j\in J}$, 故 $[K\colon F]=[K\colon E]\cdot[E\colon F]$. 因此 $\mathbb{C}/\mathbb{Q}$ 也为不可数次扩张; $\mathbb{F}_{p^m}\subseteq\mathbb{F}_{p^n}\Longleftrightarrow m\mid n$.

例 1.4.5 ($\dim_{\mathbb{Q}}(\mathbb{R})$ 非有限 (或不可数) 的证明方法)

(1) 取 $\mathbb{R}$ 中的一个 $\mathbb{Q}$-超越元 α (例如 π,e), 则 $\{\alpha^n\}_{n=0}^{+\infty}\subseteq\mathbb{R}$ 在 $\mathbb{Q}$ 上线性无关, 故 $\dim_{\mathbb{Q}}(\mathbb{R})$ 非有限.

(2) 设 $\{p_i\}_{i=1}^n$ 为不同的素数, 则 $\{\log p_i\}_{i=1}^n\subseteq\mathbb{R}$ 在 $\mathbb{Q}$ 上线性无关, 故由素数无穷多知 $\dim_{\mathbb{Q}}(\mathbb{R})$ 非有限.
(这是因为: 假设 $\exists\{c_i\}_{i=1}^n\subseteq\mathbb{Q},\ s.t.\ \sum\limits_{i=1}^n c_i\log p_i=0$, 通过消去分母可不妨设 $c_i\in\mathbb{Z}$, 则 $\prod\limits_{i=1}^n p_i^{c_i}=1$. 由算术基本定理知, $c_i=0,\ \forall 1\leq i\leq n$, 即 $\{\log p_i\}_{i=1}^n$ 在 $\mathbb{Q}$ 上线性无关.)

(3) 设 $\{p_i\}_{i=1}^n$ 为不同的素数, 则 $\{\sqrt{p_i}\}_{i=1}^n\subseteq\mathbb{R}$ 在 $\mathbb{Q}$ 上线性无关, 故由素数无穷多知 $\dim_{\mathbb{Q}}(\mathbb{R})$ 非有限.
(这是因为: 由命题 1.4.18知 $\mathbb{Q}\subseteq\mathbb{Q}(\sqrt{p_1})\subseteq\mathbb{Q}(\sqrt{p_1},\sqrt{p_2})\subseteq\cdots\subseteq\mathbb{Q}(\sqrt{p_1},\cdots,\sqrt{p_n})$ 为一个二次扩张链, 故 $\{\sqrt{p_i}\}_{i=1}^n$ 在 $\mathbb{Q}$ 上线性无关.)

(4) 设 V 为 $\mathbb{Q}$-线性空间, 则 $\dim_{\mathbb{Q}}(V)$ 为可数势 $\Longleftrightarrow|V|$ 为可数势 (由本节开头的引理即知), 故由 $|\mathbb{R}|$ 不可数知, $\dim_{\mathbb{Q}}(V)$ 不可数.

(5) 事实上, $\dim_{\mathbb{Q}}(\mathbb{R})=|\mathbb{R}|$. (这是因为: 显然 $\dim_{\mathbb{Q}}(\mathbb{R})\leq|\mathbb{R}|$; 另一方面, 固定 $\{q_n\}_{n=0}^{+\infty}$ 为 $\mathbb{Q}$ 的一个排列, 记 $S:=\{A_r:=\sum\limits_{\substack{n\geq 0\\ q_n<r}}\dfrac{1}{n!}\in\mathbb{R}\colon r\in\mathbb{R}\}$, 则 $\mathbb{R}\longrightarrow S$ 为单射, 故 $|S|=|\mathbb{R}|$. 又可直接验证 $S\subseteq\mathbb{R}$ 为 $\mathbb{Q}$-线性无关集, 故 $\dim_{\mathbb{Q}}(\mathbb{R})\geq|\mathbb{R}|$.)
$r\longmapsto A_r$

注: 在上例的 (5) 中, 我们显式给出了 $\mathbb{R}$ 的一个在 $\mathbb{Q}$ 上线性无关的不可数子集 S, 但 S 并不是 $\mathbb{R}$ 的 $\mathbb{Q}$-基.

引理 1.4.17 设 F 为一个域, 且 $\mathrm{char}(F)\neq 2$, $a,b\in F$ 满足 $\sqrt{a},\sqrt{b},\sqrt{ab}\notin F$, 则 $[F(\sqrt{a},\sqrt{b}):F]=4$.
证明: 由 $\sqrt{a}\notin F$ 知 $[F(\sqrt{a}):F]=2$. 下证 $[F(\sqrt{a},\sqrt{b}),F(\sqrt{a})]=2$. 假设 $F(\sqrt{a},\sqrt{b})=F(\sqrt{a})$, 即 $b\in F(\sqrt{a})$, 则 $\exists r,s\in F,\ s.t.\ \sqrt{b}=r+s\sqrt{a}$, 故 $F\ni b=r^2+s^2a+2rs\sqrt{a}$. 由 $\sqrt{a}\notin F$ 且 $\mathrm{char}(F)\neq 2$ 知, $rs=0$. 若 $r=0$, 则 $\sqrt{b}=s\sqrt{a}$, 故 $\sqrt{ab}=sa\in F$, 矛盾! 若 $s=0$, 则 $\sqrt{b}=r\in F$, 矛盾! 因此 $[F(\sqrt{a},\sqrt{b}),F(\sqrt{a})]=2$, 故 $[F(\sqrt{a},\sqrt{b}):F]=[F(\sqrt{a},\sqrt{b}),F(\sqrt{a})]\cdot[F(\sqrt{a}):F]=4$. □

命题 1.4.18 设 F 为一个域, 且 $\mathrm{char}(F)\neq 2$, $a_1,\cdots,a_n\in F$, 满足

$$\forall 1\leq k\leq n,\ \forall 1\leq i_1<\cdots<i_k\leq n,\ \sqrt{a_{i_1}\cdots a_{i_k}}\notin F,$$

则 $F\subseteq F(\sqrt{a_1})\subseteq F(\sqrt{a_1},\sqrt{a_2})\subseteq\cdots\subseteq F(\sqrt{a_1},\cdots,\sqrt{a_n})$ 为一个二次扩张链.
证明: 对 n 归纳证明结论. 当 $n=1$ 时结论显然. 现设 $n\geq 2$, 且结论当 $\leq(n-1)$ 时均成立. 记

$$E=F(\sqrt{a_1},\cdots,\sqrt{a_{n-2}}),\ K=F(\sqrt{a_1},\cdots,\sqrt{a_n}),$$

则 $K=E(\sqrt{a_{n-1}},\sqrt{a_n})$. 由归纳假设知 $[E:F]=2^{n-2}$, 下证 $[K:E]=4$. 事实上, 由归纳假设知 $[E(\sqrt{a_{n-1}}):F]=[E(\sqrt{a_n}):F]=[E(\sqrt{a_{n-1}a_n}):F]=2^{n-1}$, 则 $\sqrt{a_{n-1}},\sqrt{a_n},\sqrt{a_{n-1}a_n}\notin E$, 故由引理 1.4.17知 $[E(\sqrt{a_{n-1}},\sqrt{a_n}):E]=4$, 即 $[K:E]=4$, 因此 $[K:F]=[K:E]\cdot[E:F]=2^n$. □

最后, 我们用这个巧妙的例子结束本章 “群、域、线性空间” 的讨论.

例 1.4.6 设 $(G,\cdot,1)$ 为一个群, 且 $|G|>2$, 证明: $|\mathrm{Aut}(G,\cdot,1)|\geq 2$.

证明: 显然 $\mathrm{id}\in\mathrm{Aut}(G,\cdot,1)$. 若 G 非交换群, 则 $\forall a\in G\setminus Z(G)$, $c_a\colon G\longrightarrow G,\ g\longmapsto aga^{-1}$ 为一个非平凡的 (内) 自同构. 若 G 为交换群, 且 $\exists a\in G,\ s.t.\ a^2\neq 1$, 则 $\eta\colon G\longrightarrow G,\ g\longmapsto g^{-1}$ 为一个非平凡的自同构. 现设 $\forall a\in G,\ a^2=1$, 以下将 G 中二元运算记为 “+”, 幺元记为 “0”, 则 $(G,+,0)$ 为 $\mathbb{F}_2$ 上的线性空间. 由 $|G|>2$ 知 $\dim_{\mathbb{F}_2}(G)>1$, 此时交换 G 中两个基元的线性变换是一个非平凡的自同构. □

参考文献与补注 1.4

(1) 关于 Zorn 引理的部分, 可以参考维基百科或者 Keith Conrad “Zorn's Lemma and Some Applications” 及其参考文献.
(2) 关于集合势算术的部分, 可以参考 Gabriel Nagy“Real Analysis” 的附录.
(3) 关于 Banach 空间中 Baire 纲定理的部分, 可以参考 Walter Rudin“Functional Analysis”.
(4) 关于 $\dim_{\mathbb{Q}}(\mathbb{R})$ 非有限 (或不可数) 的证明方法的部分, 可以参考 https://math.stackexchange.com/上的相关问题.

第 2 章　线性方程组与矩阵

本章从线性方程组 (system of linear equations) 的求解出发, 介绍线性代数的基本工具: 矩阵 (matrix). 它往往为处理一般问题提供了最形象直观的思路.

§2.1　线性方程组的求解

固定一个域 F. 我们回忆域 F 上线性方程组的初等解法, 即通过对其中若干方程做线性组合, 得到尽可能简化的新方程组, 从而直接得到解. 一个基本的观察是, 如果变形前后的线性方程组是等价的 (即它们可以互相线性表出), 则这两个方程组同解, 于是这种变形过程并没有损失方程组解的信息. 以矩阵和线性空间的语言, 上述分析可抽象为:

引理 2.1.1 *设 $A \in F^{m_A \times n}$, $B \in F^{m_B \times n}$, 记 $\mathrm{row}(A) \subseteq F^{1\times n}$ 为 A 的行向量组生成的子空间, $\ker(A) \subseteq F^{n\times 1}$ 为齐次线性方程组 $AX = 0$ 的解空间, 则 $\mathrm{row}(A) \subseteq \mathrm{row}(B) \iff \ker(A) \supseteq \ker(B)$.*

证明: "⇒": 设 $\mathrm{row}(A) \subseteq \mathrm{row}(B)$, 则任取 A 的行向量 $\alpha \in F^{1\times n}$, 存在 B 的若干行向量 $\beta_1, \cdots, \beta_k$, 以及 $c_1, \cdots, c_k \in F$, 使得 $\alpha = \sum\limits_{j=1}^{k} c_j\beta_j$. 现任取 $X \in \ker(B)$, 由 $BX = 0$ 知 $\beta_j X = 0$, $j = 1, \cdots, k$, 则 $\alpha X = \sum\limits_{j=1}^{k} c_j\beta_j X = 0$, 故 $AX = 0$, 即 $X \in \ker(A)$. 因此 $\ker(B) \subseteq \ker(A)$.

"⇐": 假设 $\mathrm{row}(A) \not\subseteq \mathrm{row}(B)$, 则存在 A 的行向量 $\alpha \notin \mathrm{row}(B)$, 记 $B' = \begin{pmatrix} B \\ \alpha \end{pmatrix}$, 断言: $\ker(B') \subsetneqq \ker(B)$. 这是因为, 显然 $\ker(B') \subseteq \ker(B)$; 另一方面, 由 $\alpha \notin \mathrm{row}(B)$ 知, $\mathrm{row}(B') \supsetneqq \mathrm{row}(B)$, 则

$$\dim_F(\ker(B')) = n - \dim_F(\mathrm{row}(B')) \leq n - \dim_F(\mathrm{row}(B)) = \dim_F(\ker(B)),$$

故 $\ker(B') \subsetneqq \ker(B)$. 取 $X \in \ker(B) \setminus \ker(B')$, 则 $\alpha X \neq 0$, 故 $AX \neq 0$, 即 $X \notin \ker(A)$, 因此 $\ker(B) \not\subseteq \ker(A)$. □

注:

(1) 上述证明的非平凡部分 "⇐" 不可避免地用到了公式 $\dim_F(\ker(A)) = n - \dim_F(\mathrm{row}(A))$, 这是线性方程组理论的核心结果. 具体地说, 由已证的 "⇒" 部分知, 对矩阵 A 做初等行变换不改变 A 的行空间, 也不改变线性方程组 $AX = 0$ 的解空间, 故可不妨设 A 为行简化阶梯形的. 此时上述公式左端为线性方程组 $AX = 0$ 的自由未知量的个数, 也为矩阵 A 中零行的个数; 而右端为 $n-$ "矩阵 A 中非零行的个数", 故两者相等.

(2) 我们将一族 (无常数项的) 多元线性多项式的公共零点集称为一个 (过原点的) **线性簇** (linear variety). 上述引理表明, 对于无常数项的多元线性多项式, 若它在一个过原点的线性簇上取值为 0, 则它能写成定义这个线性簇的那些多项式的线性组合. 这一现象的非线性版本即著名的 (代数闭域上)Hilbert's Nullstellensatz.

推论 2.1.2 *设 $A \in F^{m\times n}$, 则对 A 做初等行变换不改变任意列之间的线性相关性; 对偶地, 对 A 做初等列变换不改变任意行之间的线性相关性.*

证明: 任取 A 的 s 列 $\alpha_{i_1}, \cdots, \alpha_{i_s} \in F^{m\times 1}$, 则考虑它们的线性相关性即求线性方程组 $(\alpha_{i_1}, \cdots, \alpha_{i_s})X = 0$ 的解. 当对矩阵 A 做初等行变换时, 也相当于对矩阵 $(\alpha_{i_1}, \cdots, \alpha_{i_s})$ 做初等行变换, 这并不改变矩阵 $(\alpha_{i_1}, \cdots, \alpha_{i_s})$ 的行空间, 故由引理知, 这也不改变线性方程组 $(\alpha_{i_1}, \cdots, \alpha_{i_s})X = 0$ 的解空间, 即不改变列向量组 $\{\alpha_{i_1}, \cdots, \alpha_{i_s}\}$ 的线性相关性. 再对 A^t 重复上述讨论即知对偶版本的结论. □

注: 此推论提供了求 $F^{1\times n}$ 中向量组的极大线性无关组的办法, 即将这些向量以列向量的形式排成一个 $m \times n$ 矩阵, 再对这个矩阵做初等行变换化为行简化阶梯形, 则主元 1 所在的那些列在原向量组中就是极大线性无关的.

命题 2.1.3 *对于解集非空的情形, 两个线性方程组等价当且仅当它们同解.*

证明: 一方面, 设线性方程组 $AX=Y$ 与 $BX=Z$ 等价, 则任取增广矩阵 (B,Z) 的行向量 (β,z), 存在增广矩阵 (A,Y) 的若干行向量 $(\alpha_1,y_1),\cdots,(\alpha_k,y_k)$, 以及 $c_1,\cdots,c_k\in F$, $s.t.$ $(\beta,z)=\sum\limits_{j=1}^{k}c_j(\alpha_j,y_j)$. 现任取 $AX=Y$ 的解 X, 则由 $\alpha_jX=y_j$, $j=1,\cdots,k$ 知, $\beta X=\sum\limits_{j=1}^{k}c_j\alpha_jX=\sum\limits_{j=1}^{k}y_j=z$, 故 X 也为 $BX=Z$ 的解, 即 $AX=Y$ 的解均为 $BX=Z$ 的解. 反之亦然.

另一方面, 设线性方程组 $AX=Y$ 与 $BX=Z$ 同解且解集非空, 此时解空间形如仿射子空间 $\gamma+W$, 其中 γ 为特解, W 为对应的齐次线性方程组的解空间. 断言: $AX=0$ 与 $BX=0$ 也同解. 这是因为, 取 $\gamma_{(A,Y)}$ 为 $AX=Y$ 的特解, 则也为 $BX=Z$ 的特解, 再取 $\gamma_{(B,Z)}$ 为 $BX=Z$ 的特解, 则 $\gamma_{(A,Y)}-\gamma_{(B,Z)}$ 为 $BX=0$ 的解, 即 $\gamma_{(A,Y)}-\gamma_{(B,Z)}\in W_B$, 故 $\gamma_{(A,Y)}+W_B=\gamma_{(B,Z)}+W_B$. 又 $\gamma_{(A,Y)}+W_A=\gamma_{(B,Z)}+W_B$, 则 $W_A=W_B$. 因此, 由上述引理知矩阵 A,B 行等价. 取 $AX=Y$ 与 $BX=Z$ 的公共特解 γ, 则可直接验证矩阵 $(A,A\gamma)$ 与 $(B,B\gamma)$ 也行等价, 即增广矩阵 (A,Y) 与 (B,Z) 行等价. □

注: 若两个 (非齐次) 线性方程组的解集均为空集, 则它们未必等价, 反例如 $\begin{cases}x_1+x_2=0\\x_1+x_2=1\end{cases}$ 与 $\begin{cases}x_1+2x_2=0\\x_1+2x_2=1\end{cases}$, 这里两个线性方程组均无解, 但它们的增广矩阵 $\begin{pmatrix}1&1&0\\1&1&1\end{pmatrix}$ 与 $\begin{pmatrix}1&2&0\\1&2&1\end{pmatrix}$ 显然不是行等价的.

以下命题将揭示: 线性方程组理论几乎是矩阵理论的应用.

命题 2.1.4 *设 $A\in F^{m\times n}$, 则:*

(1) *对于任意 $Y\in F^{m\times1}$, 非齐次线性方程组 $AX=Y$ 的解若存在必唯一 (可能不存在)*
$\iff$ *齐次线性方程组 $AX=0$ 只有零解*
$\iff$ *矩阵 A 的列向量组是 F-线性无关的*
$\iff$ *矩阵 A 的行向量组可 F-线性生成 $F^{1\times n}$.*

(2) *对于任意 $Y\in F^{m\times1}$, 非齐次线性方程组 $AX=Y$ 的解必存在*
$\iff$ *矩阵 A 的列向量组可 F-线性生成 $F^{m\times1}$*
$\iff$ *矩阵 A 的行向量组是 F-线性无关的.*

证明: 记 $A=(\alpha_1,\cdots,\alpha_n)$, 其中 $\alpha_1,\cdots,\alpha_n\in F^{m\times1}$ 为列向量, 则矩阵乘法 $AX=\sum\limits_{j=1}^{k}x_j\alpha_j$, 由此可知 (1)(2) 中除了各自与最后一行条件等价外均显然. 于是以下只需证明 "矩阵的行空间维数等于列空间维数", 我们将这部分推迟到矩阵的相抵标准形与秩一节中解释. □

最后我们讨论线性方程组的解与域扩张的关系.

命题 2.1.5 *设 E/F 为域扩张, $A\in F^{m\times n}$,*

(1) *记 $\mathrm{column}_F(A)\subseteq F^{m\times1}$, $\mathrm{column}_E(A)\subseteq E^{m\times1}$ 为 A 的列向量组生成的子空间, 则*
$$\mathrm{column}_F(A)=\mathrm{column}_E(A)\cap F^{m\times1}.$$

(2) *记 $\ker_F(A)\subseteq F^{n\times1}$, $\ker_E(A)\subseteq E^{n\times1}$ 为齐次线性方程组 $AX=0$ 的解空间, 则*
$$\mathrm{Span}_E(\ker_F(A))=\ker_E(A).$$

证明: (1) "$\subseteq$": 显然; "$\supseteq$": 设 $Y\in\mathrm{column}_E(A)\cap F^{n\times1}$, 即 $\exists X\in E^{n\times1}$, $s.t.$ $AX=Y$. 注意对于非齐次线性方程组 $AX=Y$, 可通过初等行变换将增广矩阵 (A,Y) 化为行简化阶梯形 (B,Z), 则原方程组有解当且仅当在 B 中零行的编号上 Z 的相应分量也为 0; 而 Z 的分量总是 Y 的诸分量的 F-线性组合, 故方程组的有解性与域扩张 E/F 无关. 因此也 $\exists X\in F^{n\times1}$, $s.t.$ $AX=Y$, 即 $Y\in\mathrm{column}_F(A)$.
(2) "$\subseteq$": 显然; "$=$": 先断言若 $\ker_E(A)\neq\{0\}$, 则 $\ker_F(A)\neq\{0\}$. 这是因为, 对于齐次线性方程组 $AX=0$, 可通过初等行变换将矩阵 A 化为行简化阶梯形 B, 则原方程组有非零解当且仅当 B 中非零行的个数小于 n; 而 B 的行向量总是 A 的行向量组的 F-线性组合, 故方程组非零解的存在性与域扩张 E/F 无关. 这也说明 $F^{m\times1}$ 中的向量组 F-线性无关当且仅当 E-线性无关. 因此
$$\dim_E(\mathrm{Span}_E(\ker_F(A)))=\dim_F(\ker_F(A))=n-\dim_F(\mathrm{row}_F(A))=n-\dim_E(\mathrm{row}_E(A))=\dim_E(\ker_E(A)).\square$$

推论 2.1.6 *设 E/F 为域扩张, $A_i \in F^{m_i\times n}\,(i\in I)$,*

(1) *设 $Y_i \in F^{m_i\times 1}\,(i\in I)$, 若 $\exists X\in E^{n\times 1}$, s.t. $A_iX=Y_i$, $\forall i\in I$, 则 $\exists X\in F^{n\times 1}$, s.t. $A_iX=Y_i$, $\forall i\in I$.*

(2) $\mathrm{Span}_E\left(\bigcap\limits_{i\in I}\ker_F(A_i)\right)=\bigcap\limits_{i\in I}\ker_E(A_i)$.

证明: 不妨设 $m_i=1$, $\forall i\in I$, 否则将每个 $A_i\in F^{m_i\times n}$ 与 $Y\in F^{m_i\times 1}$ 均拆成相应的行向量即可. 考虑 $F^{1\times(n+1)}$ 中子空间的和 $\sum\limits_{i\in I}\mathrm{Span}_F(A_i,Y_i)$, 由维数的有限性知, $\exists i_1,\cdots,i_k\in I$, s.t. $\sum\limits_{i\in I}\mathrm{Span}_F(A_i,Y_i)=\sum\limits_{j=1}^{k}\mathrm{Span}_F(A_{i_j},Y_{i_j})$, 则在 F 的任意扩域上, 一族线性方程组的联立 $A_iX=Y_i\,(i\in I)$ 等价于这有限个线性方程组的联立 $A_{i_j}X=Y_{i_j}$ $(1\le j\le k)$. 因此可不妨设 $|I|<+\infty$, 于是 (1),(2) 均由上述命题可知. □

习题 2.1 *证明以下两条陈述等价:*

(1) *设 $A\in F^{m\times n}\,(m<n)$, 则齐次线性方程组 $AX=0$ 有非零解;*

(2) *设 S,T 为线性空间 V 的有限子集, 若 S 线性无关, T 生成 V, 则 $|S|\le|T|$.*

参考文献与补注 2.1

(1) 关于代数簇与 Hilbert's Nullstellensatz 的部分, 可以参考 J. E. Humphreys "Linear Algebraic Groups".

(2) 关于执行 Gauss 消元法将线性方程组化为行简化阶梯形的过程, 可以参考 Hoffman, Kunze "Linear Algebra" 的例题, 我们在习题课上将具体演示.

§2.2 矩阵的乘法与逆

2.2.1 环与理想的引入

固定一个域 F. 事实上, 矩阵乘法的讨论只需矩阵的诸分量都在一个环中, 于是我们先补充环的定义:

定义 2.2.1 (环) 设 R 为一个非空集合, $0,1\in R$, $+,\cdot\colon R\times R\to R$ 为两个二元运算, 满足 $(R,+,0)$ 为交换群, $(R,\cdot,1)$ 为幺半群, 且 $+,\cdot$ 之间具有左、右分配律, 则称五元组 $(R,+,0;\cdot,1)$ 为一个**环** (ring).

注:

(1) 我们比较环与域的定义: 首先, 环 R 中 0, 1 未必为不同元, 且 $0=1\Leftrightarrow R=\{0\}$; 其次, 环中元未必有乘法逆元, 乘法也未必交换. (这里的术语遵从 Nathan Jacobson"Basic Algebra I" 的习惯, 在很多文献中环的定义与此不同, 为谨慎起见最好随时确认.)

(2) 若在上述定义中去掉 $1\in R$ 的条件, 则称三元组 $(R,0,+)$ 是一个 "rng", 形式上看即从环 (ring) 中去掉了幺元 (i=identity).

(3) 特别地, 满足 $0\neq 1\in R$ 且 $(R^*,\cdot,1)$ 是群的环称为一个 **除环** (division ring), 或**斜域** (skew field), 或**体** (corps).

例 2.2.1 (环 R 上的方阵代数) 设 R 为一个环, 则以 R 中元为分量的 m 行 n 列矩阵全体 $R^{m\times n}$ 构成了一个 R-模 (可类比域 F 上 m 行 n 列矩阵全体 $F^{m\times n}$ 构成了一个 F-线性空间). 特别地, 当 $m=n$ 时, 我们试图在 $R^{n\times n}$ 上以矩阵的形式引入 "乘法": 令 $(A_{pq})_{n\times n}\cdot(B_{kl})_{n\times n}=(C_{ij})_{n\times n}$, 其中 $C_{ij}=\sum\limits_{k=1}^{n}A_{ik}B_{kj}$, 此时 $R^{n\times n}$ 兼具 R-模与环的结构, 且这两者相容, 故为一个 R-结合代数, 称为环 R 上的方阵代数.

为了继续研究环 R 上方阵代数的结构, 我们引入环中理想的概念:

定义 2.2.2 (理想) 设 $(R,+,0;1,\cdot)$ 为一个环, $\mathfrak{a}\subseteq R$ 为子集, 满足 $(\mathfrak{a},+,0)$ 为 $(R,+,0)$ 的子群, 若 $\forall r\in R$, $r\mathfrak{a}\subseteq\mathfrak{a}$, 则称 $\mathfrak{a}$ 为环 R 的一个**左理想** (left ideal); 若 $\forall r\in R$, $\mathfrak{a}r\subseteq\mathfrak{a}$, 则称 $\mathfrak{a}$ 为环 R 的一个**右理想** (right ideal); 既是左理想又是右理想的子集称为环的 **(双边) 理想** ((two-sided) ideal).

例 2.2.2 (环 R 上方阵环的左理想) 设 R 为一个环, 则存在一一对应:

$$\{\text{秩为 } n \text{ 的自由左 } R\text{-模 } R^{1\times n} \text{ 的子模}\}\xleftrightarrow{1-1}\{\text{方阵环 } R^{n\times n} \text{ 的左理想}\}$$

$$S\longmapsto\mathfrak{a}(S):=\{M\in R^{n\times n}\colon \mathrm{row}(M)\subseteq S\}$$

$$S(\mathfrak{a}):=\{\alpha\in R^{1\times n}\colon \alpha\text{为某个 } M\in\mathfrak{a}\text{ 的第一行}\}\longmapsfrom\mathfrak{a}$$

特别地, 设 F 为一个域, 则上述一一对应为 $\{F$-线性空间 $F^{1\times n}$ 的子空间$\} \xleftrightarrow{1-1} \{$方阵环 $F^{n\times n}$ 的左理想$\}$.

例如, 取 $F^{1\times n}$ 的若干极小非零子空间 $\mathrm{Span}_F(\{(1,0\cdots,0)\}),\cdots,\mathrm{Span}_F(\{(0,\cdots,0,1)\})$, 则对应了 $F^{n\times n}$ 的若干极小非零理想 $\mathrm{Span}_F(\{E_{i1}\}_{i=1}^n),\cdots,\mathrm{Span}_F(\{E_{in}\}_{i=1}^n)$. 注意这些极小非零理想的直和恰为 $F^{n\times n}$, 故 $F^{n\times n}$ 是一个**半单环** (semisimple ring). 不平凡的是, 半单环论中的 Wedderburn-Artin 定理断言, 任意半单环都同构于有限个除环上的方阵环 $D_i^{n_i\times n_i}$ 的乘积, 且 D_i 与 n_i 都唯一确定. 因此方阵环在半单环论中起着重要的作用. 又如 $F^{1\times n}$ 的子空间维数恰为 $k=0,1,\cdots,n$, 则 $F^{n\times n}$ 的左理想维数恰为 $kn\,(k=0,1,\cdots,n)$.

进一步地, 上述一一对应具有更深的背景, 它构造了 R-模范畴与 $R^{n\times n}$-模范畴的范畴等价:

$$\begin{array}{ccc}\mathbf{R-Mod} & \xrightarrow{\cong} & \mathbf{R^{n\times n}-Mod} \\ X & \longmapsto & X^{n\times 1}\coloneqq\left\{\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}: x_i\in X\right\} \\ E_{11}.Y & \longmapsto & Y\end{array}$$

此时我们称环 R 与环 $R^{n\times n}$ 是 **Morita-等价**的.

例 2.2.3 (环 R 上方阵代数的双边理想) 设 R 为一个环, 考虑映射 $\{$环 R 的双边理想$\} \longrightarrow \{$环 $R^{n\times n}$ 的双边理想$\}$, $\mathfrak{a}\longmapsto \mathfrak{a}^{n\times n}$

环论的一个经典结果是上述为满射. 特别地, 若 R 为单环 (即只有平凡的双边理想的非零环), 则 $R^{n\times n}$ 也为单环. 例如, 设 F 为一个域, 则方阵代数 $F^{n\times n}$ 的双边理想均平凡, 这是域 F 上的单结合代数.

当然我们也可以直接证明 "域 F 上方阵代数 $F^{n\times n}$ 的双边理想均平凡". 这是因为, 设 $\{0\}\neq\mathfrak{a}\subseteq F^{n\times n}$ 为一个双边理想, 任取 $0\neq A\in\mathfrak{a}$, 记 $A_{ij}\neq 0$, 则 $E_{ij}=\dfrac{1}{A_{ij}}E_{ii}\cdot A\cdot E_{jj}\in\mathfrak{a}$. 再由 $E_{ij}\cdot E_{kl}=\delta_{jk}E_{il}$, $\forall 1\leq i,j,k,l\leq n$ 知, $\forall 1\leq k,l\leq n$, $E_{kl}\in\mathfrak{a}$, 故 $\mathfrak{a}=F^{n\times n}$.

半单环论中的 Wedderburn-Artin 定理的另一版本是, 域 F 上的任意有限维半单结合代数都同构于有限个域 F 上有限维可除代数上的方阵环 $D_i^{n_i\times n_i}$ 的乘积, 且 D_i 与 n_i 都唯一确定; 进一步地, 若 F 为代数闭域, 则 $D_i=F$. 特别地, 域 F 上的任意有限维单结合代数都同构于某个域 F 上有限维可除代数上的方阵环 $D^{n\times n}$, 且 D 与 n 都唯一确定.

2.2.2 矩阵的乘法

固定一个环 R. 矩阵乘法的最大特点是非交换性, 这可从以下基本矩阵的例子中略窥一二.

例 2.2.4 (基本矩阵的乘法) 设 R 为环, 记 $E_{ij}\in R^{m\times n}$ 是第 (i,j) 分量为 1, 其余分量均为 0 的矩阵 (称为基本矩阵), 则 $\{E_{ij}\colon i=1,\cdots,m,j=1,\cdots,n\}$ 是 $R^{m\times n}$ 的 R-基 (即 R-线性无关且 R-线性张成 $R^{m\times n}$), 故 $R^{m\times n}$ 是一个秩为 mn 的自由 R-模. 特别地, 当 $m=n$ 时, $R^{n\times n}$ 上的乘法完全由基本矩阵的乘法决定: 对于 $A=\sum\limits_{i,j=1}^{n}A_{ij}E_{ij}, B=\sum\limits_{k,l=1}^{n}B_{ij}E_{ij}$, 则 $A\cdot B=\sum\limits_{i,j,k,l=1}^{n}A_{ij}B_{kl}E_{ij}\cdot E_{kl}$; 而可直接验证 $E_{ij}\cdot E_{kl}=\delta_{jk}E_{il}$, 其中 $\delta_{jk}\coloneqq\begin{cases}1, & j=k\\ 0, & j\neq k\end{cases}$ 为 Kronecker 符号, 故 $A\cdot B=\sum\limits_{i,l=1}^{n}\left(\sum\limits_{k=1}^{n}A_{ik}B_{kl}\right)E_{il}$.

我们着重讨论一些特殊的矩阵运算, 包括转置、分块、对角、三角、括积与迹, 由此可深入体会矩阵乘法的性质.

定义 2.2.3 (转置) 设 R 为一个环, 则映射 $R^{m\times n}\longrightarrow R^{n\times m}$ (其中 $(A^t)_{ij}\coloneqq A_{ji}$) 称为矩阵的**转置** (transpose).
$$A\longmapsto A^t$$

注: 转置保持矩阵的加法与 R-数乘, 但并不保持矩阵的乘法; 特别地, 当 R 中乘法可交换时,
$$\forall A\in R^{m\times n},\ \forall B\in R^{n\times s},\ (A\cdot B)^t=B^t\cdot A^t.$$

定义 2.2.4 (分块) 设 R 为一个环, $m=\sum\limits_{k=1}^{p} m_k$, $n=\sum\limits_{l=1}^{q} n_l$ 为正整数的分划, 则映射 $\begin{array}{ccc} R^{m\times n} & \longrightarrow & \prod\limits_{k=1}^{p}\prod\limits_{l=1}^{q} R^{m_k\times n_l} \\ A & \longmapsto & \widetilde{A} \end{array}$ (其中 $\widetilde{A}_{kl}\coloneqq A_{I_k,J_l}$, $I_k=\left\{\sum\limits_{i=1}^{k-1} m_i+1,\cdots,\sum\limits_{i=1}^{k} m_i\right\}$, $J_l=\left\{\sum\limits_{j=1}^{l-1} n_j+1,\cdots,\sum\limits_{j=1}^{l} n_j\right\}$) 称为矩阵的**分块** (block).

注:

(1) 设 $A\in R^{m\times n}$, $B\in R^{n\times s}$, 且 A 的列与 B 的行分划方法相同, 则 $\widetilde{(A\cdot B)}=\widetilde{A}\cdot\widetilde{B}$.

(2) 一般地, 设 $A\in R^{m\times n}$, 则 $\widetilde{A^t}\neq(\widetilde{A})^t$, 这是因为前者是在 $R^{m\times n}$ 中转置, 矩阵分量为 R; 后者是在 $\prod\limits_{k=1}^{p}\prod\limits_{l=1}^{q} R^{m_k\times n_l}$ 中转置, 矩阵分量为 $R^{m_k\times n_l}$ $(1\leqslant k\leqslant p,\,1\leqslant l\leqslant q)$.

定义 2.2.5 (对角、三角) 设 R 为一个环, $A\in R^{n\times n}$, 若 $\forall i\neq(>;\geqslant;<;\leqslant)j$, $A_{ij}=0$, 则称 A 为**对角的** (diagonal) (**上三角的** (upper triangular); **严格上三角的** (strictly upper triangular); **下三角的** (lower triangular); **严格下三角的** (strictly lower triangular)).

注:

(1) 上述对角与三角的定义也可类比地应用于分块矩阵, 此时一般在矩阵类型前加 "准 (quasi-)" 字以示区分.

(2) 上述对角、三角的五种矩阵类型都各自对矩阵乘法封闭, 且对角分量即相应位置的对角分量直接相乘.

定义 2.2.6 (括积) 设 R 为一个环, 则映射 $\begin{array}{ccc} [\cdot,\cdot]\colon\ R^{n\times n}\times R^{n\times n} & \longrightarrow & R^{n\times n} \\ (A,B) & \longmapsto & [A,B]\coloneqq A\cdot B-B\cdot A \end{array}$ 称为矩阵的 **Lie 括积** (Lie bracket).

注: Lie 括积衡量了方阵乘法不可交换的程度, 具体地说来自两个原因: ①R 中乘法非交换; ②方阵乘法非交换. 后者是更本质的原因, 这也是历史上 A. Cayley(1821∼1895) 定义抽象矩阵运算时的突破性工作.

例 2.2.5 (一般线性 Lie 代数) 设 F 为一个域, 记 $\mathfrak{gl}(n,F)\coloneqq F^{n\times n}$, 其中运算为加法, F-数乘与 Lie 括积 $[\cdot,\cdot]$, 此时 $[\cdot,\cdot]$ 关于每个位置都是 F-线性的, 关于两个位置是交错的, 且满足 Jacobi 恒等式

$$[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0,\ \forall A,B,C\in\mathfrak{gl}(n,F),$$

则 $(\mathfrak{gl}(n,F),[\cdot,\cdot])$ 为域 F 上的一个 Lie 代数, 称为域 F 上的 **一般线性 Lie 代数** (general linear Lie algebra). 注意 $[\cdot,\cdot]$ 不满足结合律, 故 $(\mathfrak{gl}(n,F),[\cdot,\cdot])$ 为非结合代数.

例 2.2.6 (上三角线性 Lie 代数) 设 F 为一个域, $(\mathfrak{gl}(n,F),[\cdot,\cdot])$ 为域 F 上的一般线性 Lie 代数, 考虑 $\mathfrak{gl}(n,F)$ 的子集

$$\begin{aligned}
\mathfrak{t}(n,F)&\coloneqq\{A\in\mathfrak{gl}(n,F)\colon A_{ij}=0,\ \forall i>j\},\\
\mathfrak{n}(n,F)&\coloneqq\{A\in\mathfrak{gl}(n,F)\colon A_{ij}=0,\ \forall i\geqslant j\},\\
\mathfrak{d}(n,F)&\coloneqq\{A\in\mathfrak{gl}(n,F)\colon A_{ij}=0,\ \forall i\neq j\}.
\end{aligned}$$

由对角、三角阵运算知, 它们都是子代数. 另外可直接验证 $\mathfrak{t}(n,F)=\mathfrak{n}(n,F)\oplus\mathfrak{d}(n,F)$ 为线性空间的直和, 且 $[\mathfrak{n}(n,F),\mathfrak{d}(n,F)]=\mathfrak{n}(n,F)$, 故 $[\mathfrak{t}(n,F),\mathfrak{t}(n,F)]=\mathfrak{n}(n,F)$. 进一步地, 记 $L=\mathfrak{t}(n,F)$, 以及导列 $L^{(0)}=L,\cdots$, $L^{(k)}=[L^{(k-1)},L^{(k-1)}]$ $(k\geqslant 1)$, 则可直接验证 $L^{(k)}=\mathrm{Span}_F(\{E_{ij}\colon j-i\geqslant\lfloor 2^{k-1}\rfloor\})$ $(k\geqslant 0)$, 这给出了上三角线性 Lie 代数的一个自然分层结构.

注: 上述严格上三角线性 Lie 代数 $\mathfrak{n}(n,F)$ 是典型的**幂零** (nilpotent)Lie 代数, Lie 理论中的 Engel 定理表明, $\mathfrak{gl}(n,F)$ 中的幂零子代数都共轭于 $\mathfrak{n}(n,F)$ 的子代数. 而上述上三角线性 Lie 代数 $\mathfrak{t}(n,F)$ 是典型的**可解** (solvable)Lie 代数, Lie 理论中的 Lie 定理表明, 当 F 为代数闭域且 $\mathrm{char}(F)=0$ 时, $\mathfrak{gl}(n,F)$ 中的可解子代数都共轭于 $\mathfrak{t}(n,F)$ 的子代数.

定义 2.2.7 (迹) 设 R 为一个环, 则映射 $\begin{array}{ccc} \mathrm{tr}\colon\ R^{n\times n} & \longrightarrow & R \\ A & \longmapsto & \mathrm{tr}(A)\coloneqq\sum\limits_{i=1}^{n} A_{ii} \end{array}$ 称为矩阵的**迹** (trace).

注:

(1) 迹映射保持矩阵的加法与 R-数乘, 但并不保持矩阵的乘法.

(2) 迹映射关于矩阵的转置与分块都是不变的; 此外, 迹映射的一个重要特点是 $\mathrm{tr}\circ[\cdot,\cdot]=0$.

例 2.2.7 (特殊线性 Lie 代数) 设 F 为一个域, $(\mathfrak{gl}(n,F),[\cdot,\cdot])$ 为域 F 上的一般线性 Lie 代数, 考虑 $\mathfrak{gl}(n,F)$ 的子集

$$\mathfrak{sl}(n,F)\coloneqq\{A\in\mathfrak{gl}(n,F)\colon \mathrm{tr}(A)=0\}.$$

由迹映射保持加法与 F-数乘知 $\mathfrak{sl}(n,F)$ 为线性子空间; 再由 $\mathrm{tr}\circ[\cdot,\cdot]=0$ 知 $\mathfrak{sl}(n,F)$ 为子代数, 称为域 F 上的**特殊线性 Lie 代数** (special linear Lie algebra).

注意 $\mathfrak{sl}(n,F)$ 的 F-基可取为 $\{E_{ij}\colon i\neq j\}\cup\{E_{ii}-E_{i+1,i+1}\colon 1\leq i\leq n-1\}$, 则 $\dim_F(\mathfrak{sl}(n,F))=n^2-1$. 记 $\mathfrak{s}(n,F)\coloneqq\mathrm{Span}_F(\{I_n\})$. 显然 $\forall 1\leq i\leq n$, $\mathfrak{sl}(n,F)\oplus\mathrm{Span}_F(\{E_{ii}\})=\mathfrak{gl}(n,F)$; 但

$$\mathfrak{sl}(n,F)\oplus\mathfrak{s}(n,F)=\mathfrak{gl}(n,F)\Longleftrightarrow \mathrm{char}(F)\nmid n$$

(这是因为, 当 $\mathrm{char}(F)\nmid n$ 时, $\left(A-\dfrac{\mathrm{tr(A)}}{n}I_n,\dfrac{\mathrm{tr(A)}}{n}I_n\right)\mapsto A$ 给出了上述直和分解; 当 $\mathrm{char}(F)\mid n$ 时, $I_n\in\mathfrak{sl}(n,F)$.)

以下我们试图给出迹为 0 的矩阵的一个刻画, 即总能写成 Lie 括积的形式.

引理 2.2.1 *设 F 为一个域, 且 $\mathrm{char}(F)\nmid n$, 则域 F 上的迹为 0 的方阵总可相似于某个对角分量均为 0 的方阵.*

证明: 我们对 n 归纳证明结论. 当 $n=1$ 时, 迹为 0 的方阵为 (0), 结论显然; 假设 $n\geq 2$ 且当 $(n-1)$ 时结论成立, 设 $X\in F^{n\times n}$ 满足 $\mathrm{tr}(X)=0$. 若 $X=0$, 则结论显然; 若 $X\neq 0$, 断言: $\exists\alpha\in F^{n\times 1}$, *s.t.* $\{\alpha,X\alpha\}$ 为线性无关集.

(这是因为, 假设 $\forall\alpha\in F^{n\times 1}$, $\{\alpha,X\alpha\}$ 为线性相关集, 则 $\forall\alpha\in F^{n\times 1}\setminus\{0\}$, $\exists! c_\alpha\in F$, *s.t.* $X\alpha=c_\alpha\alpha$. 现任取 $\alpha,\beta\in F^{n\times 1}$ 为线性无关元, 则 $c_{\alpha+\beta}(\alpha+\beta)=X(\alpha+\beta)=X\alpha+X\beta=c_\alpha\alpha+c_\beta\beta$, 故由线性无关性知, $c_\alpha=c_{\alpha+\beta}=c_\beta$. 因此任取 $F^{n\times 1}$ 的基 $\{\alpha_i\}_{i=1}^n$, 则 $\exists c\in F$, *s.t.* $X\alpha_i=c\alpha_i$, $\forall 1\leq i\leq n$, 此即 $X=cI_n$, 故 $\mathrm{tr}(X)=nc$. 但由 $\mathrm{char}(F)\nmid n$ 知, $\mathrm{tr}(X)=0\Leftrightarrow c=0\Leftrightarrow X=0$, 矛盾.)

现将 $\{\alpha,X\alpha\}$ 扩充为 $F^{n\times 1}$ 的基 B, 记 $P_0=(\alpha,X\alpha,\cdots)$ 为 B 中元以列向量方式排成的方阵, 则 $P_0\in\mathrm{GL}(n,F)$, 且 $P_0^{-1}XP_0=\begin{pmatrix}0&*\\*&Y\end{pmatrix}$, 其中 $\mathrm{tr}(Y)=\mathrm{tr}(P_0^{-1}XP_0)=\mathrm{tr}(X)=0$. 由归纳假设知, $\exists Q_0\in\mathrm{GL}(n-1,F)$, *s.t.* $Q_0^{-1}YQ_0=Y'$, 这里 Y' 的对角分量均为 0, 则

$$\begin{pmatrix}1&0\\0&Q_0\end{pmatrix}^{-1}\cdot P_0^{-1}XP_0\cdot\begin{pmatrix}1&0\\0&Q_0\end{pmatrix}=\begin{pmatrix}0&*'\\*'&Y'\end{pmatrix}=:X',$$

这里 X' 的对角分量均为 0. □

注: 上述引理对于 $\mathrm{char}(F)\mid n$ 的情形不成立, 此时 I_n 的迹为 0 但只能相似于自身.

推论 2.2.2 *设 F 为一个域, 且 $\mathrm{char}(F)\nmid n$, 则域 F 上的方阵总可相似于某个对角分量均相等的方阵.*

命题 2.2.3 *设 F 为一个无限域, 且 $\mathrm{char}(F)\nmid n$, 则*

$$\{[A,B]\colon A,B\in\mathfrak{gl}(n,F)\}=\mathrm{Span}_F\{[A,B]\colon A,B\in\mathfrak{gl}(n,F)\}=\mathfrak{sl}(n,F).$$

证明: 显然 $\{[A,B]\colon A,B\in\mathfrak{gl}(n,F)\}\subseteq\mathrm{Span}_F\{[A,B]\colon A,B\in\mathfrak{gl}(n,F)\}\subseteq\mathfrak{sl}(n,F)$. 另一方面, 设 $X\in\mathfrak{sl}(n,F)$, 则由引理知, $\exists P\in\mathrm{GL}(n,F)$, 以及 $X'\in\mathfrak{gl}(n,F)$ 的对角分量均为 0, *s.t.*$P^{-1}XP=X'$. 由 F 为无限域知, 可取 $A'=\mathrm{diag}(a_1,\cdots,a_n)\in\mathfrak{gl}(n,F)$ 的对角分量两两不同, 以及 $B'\in\mathfrak{gl}(n,F)$ 满足

$$B'_{ii}=0,\ \forall 1\leq i\leq n;\ B'_{ij}=\frac{X'_{ij}}{a_i-a_j},\ \forall 1\leq i\neq j\leq n,$$

则 $X'=[A',B']$, 故 $X=[PA'P^{-1},PB'P^{-1}]$. □

注:

(1) 上述命题的证明依赖于域 F 的若干条件. 事实上利用矩阵的有理标准形可以去掉这些条件, 可见 Albert, Muckenhoupt "On Matrices of Trace Zero"(1956).

(2) 进一步地, 我们考虑以下等式的成立条件:

$$\{[A,B]\colon A,B\in\mathfrak{sl}(n,F)\}=\mathrm{Span}_F\{[A,B]\colon A,B\in\mathfrak{sl}(n,F)\}=\mathfrak{sl}(n,F).$$

事实上, 除了 $\mathrm{char}(F)=2$ 且 $n=2$ 的情形, 上述等式均成立, 可见 R. C. Thompson"Matrices with Zero Trace"(1965); 而当 $\mathrm{char}(F)=2$ 且 $n=2$ 时, $E_{12}\in\mathfrak{sl}(n,F)$ 不可能写成前两个集合中的形式.
(对于以上等式的后半部分, 我们在这里给出一种简单的证明: 当 $n=1$ 时, 结论显然; 当 $n=2$ 时, 由等式 $[E_{12},E_{21}]=E_{11}-E_{22}$, $[E_{11}-E_{22},E_{12}]=2E_{12}$, $[E_{11}-E_{22},E_{21}]=-2E_{21}$ 即知; 当 $n\geq 3$ 时, 由等式 $E_{ii}-E_{i+1,i+1}=[E_{i,i+1},E_{i+1,i}]$ 以及 $E_{ij}=[E_{ik},E_{kj}]\,(i,j,k$两两不同$)$ 即知.)

(3) 设 $L\subseteq\mathfrak{gl}(n,F)$ 为子代数, 记 $[L,L]\coloneqq\mathrm{Span}_F\{[A,B]\colon A,B\in L\}$, 若 $[L,L]=L$, 则称 L 为**完美** (perfect)Lie 代数. Lie 理论表明 **半单** (semisimple)Lie 代数都是完美的. 但半单 Lie 代数中元是否都能写成 Lie 括积的形式还是个开放的问题. 现已证明当 $|F|$ 充分大 (例如非有限) 时, 四种典型的**单** (simple)Lie 代数中元都能写成 Lie 括积的形式, 例如 $A_n\coloneqq\mathfrak{sl}(n+1,F)\,(n\geq 2)$ 要求 $|F|>2n-1$, 可见 G. Brown"On Commutators in a Simple Lie Algebra" (1962).

(4) 上述命题及注可类比有限群论中的 Ore 猜想 (1951):"有限非交换单群中每个元均为交换子", 它的证明依赖于有限单群的分类, 可见 Liebeck, O'Brien, Shalev, Tiep "The Ore Conjecture"(2010).

2.2.3 矩阵的逆

设 R 为一个环, 引入记号 $R^*\coloneqq R\setminus\{0\}$; $R^\times=U(R)\coloneqq\{R$中所有乘法可逆元$\}$. 特别地, $R^\times$ 是一个群.

定义 2.2.8 (可逆矩阵群) 设 R 为一个环, 则 $\mathrm{GL}(n,R)\coloneqq U(R^{n\times n})$ 称为环 R 上的 n 级**可逆矩阵群** (group of invertible matrices). 特别地, 若 F 为一个域, 则 $\mathrm{GL}(n,F)$ 称为域 F 上的 n 级 **一般线性群** (general linear group).

例 2.2.8 (可逆矩阵群关于转置的封闭性) 设 R 为一个环.
(1) 若 R 中乘法可交换, 则 $\forall A,B\in R^{n\times n}$, $(A\cdot B)^t=B^t\cdot A^t$, 故 $A\in\mathrm{GL}(n,R)\iff A^t\in\mathrm{GL}(n,R)$, 即 $\mathrm{GL}(n,R)$ 关于转置封闭.
(2) 但若 R 中乘法非交换, 则这一结论不对. 反例如设 F 为一个域, $R=F^{2\times 2}$, 考虑 $\begin{pmatrix}I_2 & A\\ -B & C\end{pmatrix}\in R^{2\times 2}$, 由初等列变换 $\begin{pmatrix}I_2 & A\\ -B & C\end{pmatrix}\cdot\begin{pmatrix}I_2 & -A\\ 0 & I_2\end{pmatrix}=\begin{pmatrix}I_2 & 0\\ -B & BA+C\end{pmatrix}$ 知, $\begin{pmatrix}I_2 & A\\ -B & C\end{pmatrix}\in\mathrm{GL}(2,R)\iff(BA+C)\in\mathrm{GL}(2,F)$.
另一方面, 注意 $\begin{pmatrix}I_2 & A\\ -B & C\end{pmatrix}^t=\begin{pmatrix}I_2 & -B\\ A & C\end{pmatrix}\in R^{2\times 2}$ (这是在 $R^{2\times 2}$ 中转置, 而不是在 $F^{4\times 4}$ 中转置), 由初等列变换 $\begin{pmatrix}I_2 & -B\\ A & C\end{pmatrix}\cdot\begin{pmatrix}I_2 & B\\ 0 & I_2\end{pmatrix}=\begin{pmatrix}I_2 & 0\\ A & AB+C\end{pmatrix}$ 知, $\begin{pmatrix}I_2 & -B\\ A & C\end{pmatrix}\in\mathrm{GL}(2,R)\iff(AB+C)\in\mathrm{GL}(2,F)$. 现取 $A=E_{12},B=E_{21},C=E_{11}\in R=F^{2\times 2}$, 则 $BA+C=I_2\in\mathrm{GL}(2,F)$, 但 $AB+C=2E_{11}\notin\mathrm{GL}(2,F)$, 故此时 $\mathrm{GL}(2,R)$ 关于转置不封闭.
(3) 一般地, $\forall n\in\mathbb{N}^*$, $\mathrm{GL}(n,R)$ 关于转置封闭 $\iff R/\mathrm{Rad}(R)$ 为交换环, 其中 $\mathrm{Rad}(R)$ 为 R 的 *Jacobson* 根, 可见 *R. N. Gupta, A. Khurana, D. Khurana, T. Y. Lam "Rings over which the transpose of every invertible matrix is invertible"(2009).*

例 2.2.9 (上三角 unipotent 矩阵的逆) 设 R 为一个环, 记

$$\mathrm{UT}(n,R)\coloneqq\{A\in R^{n\times n}\colon\forall i>j,\ A_{ij}=0;\ \forall 1\leq i\leq n,\ A_{ii}=1\},$$

则显然 $\mathrm{UT}(n,R)$ 关于矩阵乘法封闭; 进一步地, 注意 $\forall A\in\mathrm{UT}(n,R)$, $(A-I_n)^n=0$, 则由二项式定理知, $A\cdot\sum\limits_{k=1}^{n}(-1)^{k-1}\binom{n}{k}A^{k-1}=I_n$, 即 $A^{-1}=\sum\limits_{k=1}^{n}(-1)^{k-1}\binom{n}{k}A^{k-1}\in\mathrm{UT}(n,R)$, 故 $\mathrm{UT}(n,R)$ 关于矩阵的逆封闭, 为一般线性群 $\mathrm{GL}(n,R)$ 的子群, 称为**上三角 unipotent**(unitriangular) 矩阵群.

注:

(1) 由幂零 Lie 群的理论知, 设 F 为一个域, 则 $\mathrm{GL}(n,F)$ 中的 unipotent 子群都共轭于上三角 unipotent 群的子群. 这可视为有限群论中事实 "有限 p-群均为幂零群" 的推广, 因为当 $\mathrm{char}(F)=p$ 为素数时, $A\in\mathrm{GL}(n,F)$ 是 unipotent 矩阵当且仅当 $A\in\mathrm{GL}(n,F)$ 的阶为 p 的幂次.

(2) 上三角 unipotent 矩阵的逆中分量的具体表达式与量子物理中一维多体 Calogero-Sytherland 模型的解有关, 可见 Xiaoping Xu "Representations of Lie Algebras and Partial Differential Equations"(2016).

例 2.2.10 (Hilbert 矩阵的逆) 我们回忆求域 F 上矩阵的逆的具体算法: 设 $A\in F^{n\times n}$, 对矩阵 $(A,I_n)\in F^{n\times 2n}$ 做初等行变换, 将它化为行简化阶梯形矩阵 (X,B). 若 $X\in F^{n\times n}$ 的最后一行为 0, 则 $A\notin \mathrm{GL}(n,F)$; 若 $X=I_n$, 则 $A\in\mathrm{GL}(n,F)$ 且 $A^{-1}=B$. 教材 Hoffman, Kunze "Linear Algebra" 的例题求出了 3 阶 Hilbert 矩阵的逆, 我们在习题课上将具体演示. 注意一个神奇的现象是, 3 阶 Hilbert 矩阵的逆的每个分量均为 (大) 整数.

一般地, 记 $H_n\in\mathbb{Q}^{n\times n}$ 为 n 级 Hilbert 矩阵, 其中 $(H_n)_{ij}\coloneqq\displaystyle\int_0^1 t^{i+j-2}\mathrm{d}t=\frac{1}{i+j-1}$, 则 H_n 为可逆阵 (事实上为正定阵), 且 H_n^{-1} 的每个分量均为 (大) 整数: $(H_n^{-1})_{ij}=(-1)^{i+j}(i+j-1)\binom{n+i-1}{n-j}\binom{n+j-1}{n-i}\binom{i+j-1}{i-1}^2$.

Hilbert 矩阵是典型的病态矩阵 (即条件数很大), 它在数值计算中是臭名昭著的. 具体地说, 考虑线性方程组 $AX=Y$ 的机器求解, 当输入端 Y 产生微扰时, 输出端 $X(=A^{-1}Y)$ 是否变化很大? 这一现象可由相对误差 $\dfrac{\|A^{-1}\Delta Y\|/\|A^{-1}Y\|}{\|\Delta Y\|/\|Y\|}$ ($\|\cdot\|$ 为任意矩阵范数) 来描述, 则最大可能相对误差为 $\displaystyle\max_{\Delta Y,Y\neq 0}\frac{\|A^{-1}\Delta Y\|/\|A^{-1}Y\|}{\|\Delta Y\|/\|Y\|}$
$=\displaystyle\max_{\Delta Y\neq 0}\left\{\frac{\|A^{-1}\Delta Y\|}{\|\Delta Y\|}\right\}\cdot\max_{Y\neq 0}\left\{\frac{\|Y\|}{\|A^{-1}Y\|}\right\}=\|A^{-1}\|_{\mathrm{op}}\cdot\|A\|_{\mathrm{op}}=:\kappa(A)$ 称为矩阵 A 的 **条件数** (condition number). n 级 Hilbert 矩阵的条件数为 $O((1+\sqrt{2})^{4n}/\sqrt{n})$.

以下我们研究与矩阵的逆密切相关的矩阵打洞技术, 相传这是华罗庚先生 (及其学生) 的拿手好戏.

命题 2.2.4 (矩阵打洞技术) 设 $M\in F^{n\times n}$, 且可分块为 $\begin{pmatrix}A&B\\C&D\end{pmatrix}$, 其中 $A\in F^{r\times r}$.

(1) 若 $A\in\mathrm{GL}(r,F)$, 则

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}\cdot\begin{pmatrix}I_r&-A^{-1}B\\0&I_{n-r}\end{pmatrix}=\begin{pmatrix}A&0\\C&D-CA^{-1}B\end{pmatrix};$$

$$\begin{pmatrix}I_r&0\\-CA^{-1}&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A&B\\0&D-CA^{-1}B\end{pmatrix};$$

$$\begin{pmatrix}I_r&0\\-CA^{-1}&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}A&B\\C&D\end{pmatrix}\cdot\begin{pmatrix}I_r&-A^{-1}B\\0&I_{n-r}\end{pmatrix}=\begin{pmatrix}A&0\\0&D-CA^{-1}B\end{pmatrix};$$

特别地, 若 $A\in\mathrm{GL}(r,F)$ 且 $(D-CA^{-1}B)\in\mathrm{GL}(n-r,F)$, 则

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}^{-1}=\begin{pmatrix}I_r&-A^{-1}B\\0&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}A^{-1}&0\\0&(D-CA^{-1}B)^{-1}\end{pmatrix}\cdot\begin{pmatrix}I_r&0\\-CA^{-1}&I_{n-r}\end{pmatrix}.$$

(2) 若 $D\in\mathrm{GL}(n-r,F)$, 则

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}\cdot\begin{pmatrix}I_r&0\\-D^{-1}C&I_{n-r}\end{pmatrix}=\begin{pmatrix}A-BD^{-1}C&B\\0&D\end{pmatrix};$$

$$\begin{pmatrix}I_r&-BD^{-1}\\0&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}A-BD^{-1}C&0\\C&D\end{pmatrix};$$

$$\begin{pmatrix}I_r&-BD^{-1}\\0&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}A&B\\C&D\end{pmatrix}\cdot\begin{pmatrix}I_r&0\\-D^{-1}C&I_{n-r}\end{pmatrix}=\begin{pmatrix}A-BD^{-1}C&0\\0&D\end{pmatrix};$$

特别地, 若 $D\in\mathrm{GL}(n-r,F)$ 且 $(A-BD^{-1}C)\in\mathrm{GL}(r,F)$, 则

$$\begin{pmatrix}A&B\\C&D\end{pmatrix}^{-1}=\begin{pmatrix}I_r&0\\-D^{-1}C&I_{n-r}\end{pmatrix}\cdot\begin{pmatrix}(A-BD^{-1}C)^{-1}&0\\0&D^{-1}\end{pmatrix}\cdot\begin{pmatrix}I_r&-BD^{-1}\\0&I_{n-r}\end{pmatrix}.$$

注: 上述命题实际上给出了分块阵 $M=\begin{pmatrix}A&B\\C&D\end{pmatrix}$ 的 (粗糙)$L\cdot D\cdot U$ 分解, 从而提供了 (大型) 线性方程组 $\begin{pmatrix}A&B\\C&D\end{pmatrix}\cdot\begin{pmatrix}X_1\\X_2\end{pmatrix}=\begin{pmatrix}Y_1\\Y_2\end{pmatrix}$ 的一种约化解法. 例如, 若 $A\in\mathrm{GL}(r,F)$, 可以先消去 $X_1=A^{-1}(Y_1-BX_2)$, 得到约化方程 $(D-CA^{-1}B)X_2=-CA^{-1}Y_1+Y_2$. 这里 $M/A\coloneqq D-CA^{-1}B$ 称为子阵 A 在 M 中的 **Schur**

补

(Schur complement), 则 $r(A)+r(M/A)=r(M)$.

推论 2.2.5 (Woodbury) 设 $M\in F^{n\times n}$, 且可分块为 $\begin{pmatrix}A & B\\ C & D\end{pmatrix}$, 其中 $A\in F^{r\times r}$. 若 $A\in \mathrm{GL}(r,F)$, $D\in \mathrm{GL}(n-r,F)$, 且 $(D-CA^{-1}B)\in \mathrm{GL}(n-r,F)$, $(A-BD^{-1}C)\in \mathrm{GL}(r,F)$ 则

$$(A-BD^{-1}C)^{-1}=A^{-1}+A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1};$$
$$(D-CA^{-1}B)^{-1}=D^{-1}+D^{-1}C(A-BD^{-1}C)^{-1}BD^{-1}.$$

特别地, 取 $A=I_n, D=I_1$, $B=\alpha=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$, $C=\beta^t=\begin{pmatrix}b_1 & \cdots & b_n\end{pmatrix}$, 则 $(I_n-\alpha\cdot\beta^t)^{-1}=I_n+(1-\sum\limits_{i=1}^{n}a_ib_i)^{-1}\alpha\cdot\beta^t$.

例 2.2.11 (Cartan 矩阵的逆) 我们陈述几类典型的 Cartan 矩阵及其逆:

(1) $A_n=\begin{pmatrix}2&-1&&&\\-1&2&-1&&\\&-1&\ddots&\ddots&\\&&\ddots&2&-1\\&&&-1&2\end{pmatrix}_{n\times n}$ $(n\geq 1)$, 则 $(A_n^{-1})_{ij}=\min\{i,j\}-\dfrac{ij}{n+1}$;

(2) $B_n=\begin{pmatrix}2&-1&&&\\-1&2&-1&&\\&-1&\ddots&\ddots&\\&&\ddots&2&-2\\&&&-1&2\end{pmatrix}_{n\times n}$ $(n\geq 2)$, 则 $(B_n^{-1})_{ij}=\dfrac{\min\{i,j\}}{1-\min\{0,n-i-1\}}=\begin{cases}\min\{i,j\}, & 若i<n\\ \frac{j}{2}, & 若i=n\end{cases}$;

(3) $C_n=B_n^t$;

(4) $D_n=\begin{pmatrix}2&-1&&&&\\-1&\ddots&\ddots&&&\\&\ddots&2&-1&-1\\&&-1&2&0\\&&-1&0&2\end{pmatrix}_{n\times n}$ $(n\geq 4)$, 则 $(D_n^{-1})_{ij}=(D_n^{-1})_{ji}=\begin{cases}i, & 若1\leq i\leq j\leq n-2\\ \frac{i}{2}, & 若i<n-1, j=n-1或n\\ \frac{n-2}{4}, & 若i=n-1, j=n\\ \frac{n}{4}, & 若i=j=n-1或n\end{cases}$.

上述几类 Cartan 矩阵的逆的求法提示如下: 先用初等行变换求辅助矩阵 $S_n=\begin{pmatrix}2&-1&&&\\-1&2&-1&&\\&-1&\ddots&\ddots&\\&&\ddots&2&-1\\&&&-1&1\end{pmatrix}_{n\times n}$ $(n\geq 1)$ 的逆, 可知 $(S_n^{-1})_{ij}=\min\{i,j\}$, 再反复利用上述推论即可. 具体过程可见 Y. J. Wei, Y. M. Zou "Inverses of Cartan Matrices of Lie Algebras and Lie Superalgebras"(2017).

在 Lie 理论中, 抽象 Cartan 矩阵指满足以下条件的 n 级方阵 A: ① $\forall 1\leq i,j\leq n$, $A_{ij}\in\mathbb{Z}$; ② $\forall 1\leq i\leq n$, $A_{ii}=2$; ③ $\forall 1\leq i\neq j\leq n$, $A_{ij}\leq 0$; ④ $\forall 1\leq i,j\leq n$, $A_{ij}=0\Leftrightarrow A_{ji}=0$; ⑤ 存在对角分量均为正数的对角阵 D, 使得 DAD^{-1} 为对称正定阵. 事实上, 可以证明任何不可分解为两个准对角块的抽象 Cartan 矩阵都形如以上四种典型类以及五个例外类之一. 这个分类结果恰对应了复单 Lie 代数 (或者连通 Dynkin 图) 的分类.

习题 2.2 设 F 为一个域, $(\mathfrak{gl}(n,F),[\cdot,\cdot])$ 为域 F 上的一般线性 Lie 代数, $L\subseteq\mathfrak{gl}(n,F)$ 为下列子代数, 分别求 $Z(L):=\{A\in L\colon [A,B]=0,\forall B\in L\}$ 以及 $C_{\mathfrak{gl}(n,F)}(L):=\{A\in\mathfrak{gl}(n,F)\colon [A,B]=0,\forall B\in L\}$:

(1) $L=\mathfrak{gl}(n,F),\mathfrak{sl}(n,F),\mathfrak{s}(n,F)$.

(2) $L=\mathfrak{t}(n,F),\mathfrak{n}(n,F),\mathfrak{d}(n,F)$.

参考文献与补注 2.2

(1) 关于矩阵环论的部分, 可以参考 T. Y. Lam“Lectures on Modules and Rings”, 以及 N. Jacobson“Basic Algebra II”.

(2) 关于 Hilbert 矩阵及其条件数的部分, 可以参考维基百科以及 https://mathoverflow.net/上的相关问题.

(3) 关于 Lie 群 Lie 代数理论的部分, 可以参考 J. E. Humphreys“Introduction to Lie Algebras and Representation Theory”, 以及 A. W. Knapp “Lie Groups Beyond an Introduction”.

(4) 关于代数群理论的部分, 可以参考 J. E. Humphreys“Linear Algebraic Groups”.

§2.3 矩阵的相抵标准形与秩

定理 2.3.1 (域上矩阵的相抵标准形) *设 F 为一个域, 则 $\forall A\in F^{m\times n}$, $\exists P\in \mathrm{GL}(m,F)$, $Q\in \mathrm{GL}(n,F)$, 以及唯一的 $r=r(A)\in\{0,1,\cdots,\min\{m,n\}\}$, s.t. $P\cdot A\cdot Q=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$.*

证明: 我们熟知域 F 上矩阵可通过初等行变换化为行简化阶梯形的形式, 以矩阵的语言即 $\forall A\in F^{m\times n}$, $\exists P\in \mathrm{GL}(m,F)$, s.t. $P\cdot A=\begin{pmatrix} A_1 & *\\ 0 & 0\end{pmatrix}$, 其中 A_1 的每行除一个 1 外全为 0, 且这些 1 的列指标随着行指标的增大而增大; 再做初等列变换知, $\exists Q\in \mathrm{GL}(n,F)$, s.t. $P\cdot A\cdot Q=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$. 由于初等行变换不改变行空间, 初等列变换不改变任意行的线性相关性, 则 $\dim_F(\mathrm{row}(A))=\dim_F(\mathrm{row}(P\cdot A\cdot Q))=r$; 类似可知 $\dim_F(\mathrm{column}(A))=\dim_F(\mathrm{column}(P\cdot A\cdot Q))=r$, 故 $r=r(A)\in\{0,1,\cdots,\min\{m,n\}\}$ 唯一确定. □

注:

(1) 在上述定理中, 等式右端称为矩阵 A 的**相抵标准形** (canonical form of equivalent matrices), 它由矩阵的**秩** (rank)$r=r(A)$ 完全决定, 故 “秩是同形矩阵相抵变换的完全不变量”.

(2) 上述定理的证明实际上给出了 $\dim_F(\mathrm{row}(A))=r=\dim_F(\mathrm{column}(A))$, 即矩阵的行空间维数 (称为**行秩** (row rank)) 等于列空间维数 (称为**列秩** (column rank)), 均为矩阵的秩.

(3) 特别地, 当 $m=n$ 时, 取 $B=Q\cdot P\in \mathrm{GL}(n,F)$, 则 $ABA=\left(P^{-1}\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}Q^{-1}\right)\cdot(QP)\cdot\left(P^{-1}\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}Q^{-1}\right)$ $=P^{-1}\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}Q^{-1}=A$, 故 $F^{n\times n}$ 是一个 Von Neumann 正则环. (这是一个严格弱于半单环的概念, 事实上 $F^{n\times n}$ 也是半单环.)

(4) 上述矩阵的相抵标准形可推广至 **主理想整环** (principal ideal domain) 上: 设 R 为一个 P.I.D., 则 $\forall A\in R^{m\times n}$,
$\exists P\in \mathrm{GL}(m,R)$, $Q\in \mathrm{GL}(n,R)$, $d_1\mid\cdots\mid d_r\in R^*$, s.t. $P\cdot A\cdot Q=\mathrm{diag}(d_1,\cdots,d_r,0,\cdots,0)$, 其中 d_i 在相差乘法可逆元的意义下唯一. 这里等式右端称为矩阵 A 的 **Smith 标准形** (Smith normal form), 它由矩阵的**不变因子** (invariant factor) $d_i\,(1\le i\le r)$ 完全决定. (Smith 标准形的证明依赖于 P.I.D. 中的 Bezout 等式; 另外我们将会看到, 它可用于决定 P.I.D. 上有限生成模的结构定理中的 “不变因子”.)

推论 2.3.2 (满秩分解) *域 F 上的矩阵总可分解为某个列满秩阵与某个行满秩阵的乘积, 且这种分解在相差可逆阵的意义下唯一.*

证明: 由矩阵的相抵标准形, $\forall A\in F^{m\times n}$, $\exists P\in \mathrm{GL}(m,F)$, $Q\in \mathrm{GL}(n,F)$, s.t. $P\cdot A\cdot Q=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$. 现取 $C=P^{-1}\cdot\begin{pmatrix} I_r\\ 0\end{pmatrix}$ 为列满秩阵, $R=\begin{pmatrix} I_r & 0\end{pmatrix}\cdot Q^{-1}$ 为行满秩阵, 则 $A=C\cdot R$. 再设 $A=C'\cdot R'$ 为另一种满秩分解, 由列满秩阵有左逆, 行满秩阵有右逆知, $\exists P'\in F^{r\times m}$, $Q'\in F^{n\times r}$, s.t. $P'\cdot C'=I_r=R'\cdot Q'$, 则 $P'C\cdot RQ'=I_r$. 注意 $P'C$, $RQ'\in F^{r\times r}$, 故由定义知 $P'C$, $RQ'\in \mathrm{GL}(r,F)$. 记 $P'C=G\in \mathrm{GL}(r,F)$, 则 $RQ'=G^{-1}$, 因此 $C'=C'\cdot I_r=C'\cdot R'Q'=C\cdot RQ'=C\cdot G^{-1}$, $R'=I_r\cdot R'=P'C'\cdot R'=P'C\cdot R=G\cdot R$. □

注: 域 F 上矩阵的满秩分解有一种更直观的看法: 设 $A \in F^{m\times n}$, 取 $C \in F^{m\times r}$ 为在 A 中从左往右选取线性无关列排成的矩阵, $R \in F^{r\times n}$ 为在 A 的行简化阶梯形矩阵中去掉了零行, 则 $A = C \cdot R$.

推论 2.3.3 (秩一分解) 域 F 上矩阵的秩等于在它写成若干秩一矩阵的和式中矩阵的最小个数.

证明: 由矩阵的相抵标准形, $\forall A \in F^{m\times n},\ \exists P \in \mathrm{GL}(m,F),\ Q \in \mathrm{GL}(n,F),\ s.t.\ A = P^{-1} \cdot \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \cdot Q^{-1}$.
现取 $A_i = P^{-1} \cdot E_{ii} \cdot Q^{-1}\,(1 \le i \le r)$, 则诸 A_i 为秩一矩阵, 且 $A = \sum\limits_{i=1}^{r} A_i$. 若 $A = \sum\limits_{i=1}^{r} A'_i$, 其中诸 A'_i 为秩一矩阵, 则由 A 的每个列向量为诸 $A'_i\,(1 \le i \le r)$ 的列向量的线性组合知, $r(A) = \dim_F(\mathrm{column}(A)) \le \dim_F(\mathrm{column}(A'_1, \cdots, A'_r))$
$\le \sum\limits_{i=1}^{r} \dim_F(\mathrm{column}(A'_i)) = \sum\limits_{i=1}^{r} r(A_i) = r.$ □

注: 域 F 上矩阵的秩一分解有一种更直观的看法: 设 $A \in F^{m\times n}$, 任取 A 的一种满秩分解 $A = C \cdot R$, 取 $A_i \in F^{m\times n}$ 为 C 的第 i 列与 R 的第 i 行的乘积 $(i = 1, \cdots, r)$, 则 $A = \sum\limits_{i=1}^{r} A_i$.

接着我们讨论与矩阵的秩相关的一些不等式.

命题 2.3.4 (秩不等式 (I))

(1) 设 $A, B \in F^{m\times n}$, 则 $r(A+B) \le r(A) + r(B)$, 其中 "=" 成立 $\iff \begin{cases} \mathrm{row}(A) \cap \mathrm{row}(B) = \{0\} \\ \mathrm{column}(A) \cap \mathrm{column}(B) = \{0\} \end{cases}$
$\iff \exists P \in \mathrm{GL}(m,F),\ Q \in \mathrm{GL}(n,F),\ s.t.\ P \cdot A \cdot Q = \mathrm{diag}(I_{r(A)}, 0_{r(B)}, 0),\ P \cdot B \cdot Q = \mathrm{diag}(0_{r(A)}, I_{r(B)}, 0)$.

(2) 设 $A \in F^{m\times n},\ B \in F^{n\times s}$, 则 $r(A \cdot B) \le \min\{r(A), r(B)\}$, 其中
$r(A \cdot B) = r(A) \iff \exists Y \in F^{s\times n},\ s.t.\ ABY = A \iff F^{n\times 1} = \ker(A) + \mathrm{column}(B)$;
$r(A \cdot B) = r(B) \iff \exists X \in F^{n\times m},\ s.t.\ XAB = B \iff \ker(A) \cap \mathrm{column}(B) = \{0\}$.

证明: (1) 由于 $(A+B)$ 的每个列向量为 A, B 的列向量的线性组合, 则 $r(A+B) \le r(A,B) \le r(A) + r(B)$.
进一步地, 由秩一分解的存在性知, $\exists \alpha_i \in F^{m\times 1}, \beta_i \in F^{n\times 1}\,(1 \le i \le r(A)),\ s.t.\ A = \sum\limits_{i=1}^{r(A)} \alpha_i \cdot \beta_i^t$, 此时 $\mathrm{column}(A) = \mathrm{Span}_F(\{\alpha_i\}_{i=1}^{r(A)})$, $\mathrm{row}(A) = \mathrm{Span}_F(\{\beta_i^t\}_{i=1}^{r(A)})$; 以及 $\exists \gamma_j \in F^{m\times 1}, \delta_j \in F^{n\times 1}\,(1 \le j \le r(B))$, $s.t.\ B = \sum\limits_{j=1}^{r(B)} \gamma_j \cdot \delta_j^t$, 此时 $\mathrm{column}(B) = \mathrm{Span}_F(\{\gamma_j\}_{i=1}^{r(B)})$, $\mathrm{row}(B) = \mathrm{Span}_F(\{\delta_j^t\}_{j=1}^{r(B)})$.

现设 $r(A+B) = r(A)+r(B)$, 由 $A+B = \sum\limits_{i=1}^{r(A)} \alpha_i \cdot \beta_i^t + \sum\limits_{j=1}^{r(B)} \gamma_j \cdot \delta_j^t$ 以及秩一分解的最小性知, $\{\alpha_i\}_{i=1}^{r(A)} \cup \{\gamma_j\}_{j=1}^{r(B)}$ 为线性无关集, $\{\beta_i^t\}_{i=1}^{r(A)} \cup \{\delta_j^t\}_{j=1}^{r(B)}$ 为线性无关集, 故 $\begin{cases} \mathrm{row}(A) \cap \mathrm{row}(B) = \{0\} \\ \mathrm{column}(A) \cap \mathrm{column}(B) = \{0\} \end{cases}$.

现设 $\begin{cases} \mathrm{row}(A) \cap \mathrm{row}(B) = \{0\} \\ \mathrm{column}(A) \cap \mathrm{column}(B) = \{0\} \end{cases}$, 则 $\{\alpha_i\}_{i=1}^{r(A)} \cup \{\gamma_j\}_{j=1}^{r(B)}$ 为线性无关集, $\{\beta_i^t\}_{i=1}^{r(A)} \cup \{\delta_j^t\}_{j=1}^{r(B)}$ 为线性无关集, 故可将前者扩充为 $F^{m\times 1}$ 的基, 再以列向量形式排成 $m \times m$ 方阵 P, 则 $P \in \mathrm{GL}(m,F)$; 也可将后者扩充为 $F^{1\times n}$ 的基, 再以行向量形式排成 $n \times n$ 方阵 Q, 则 $Q \in \mathrm{GL}(n,F)$. 此时 $A = P \cdot \mathrm{diag}(I_{r(A)}, 0_{r(B)}, 0) \cdot Q$, $B = P \cdot \mathrm{diag}(0_{r(A)}, I_{r(B)}, 0) \cdot Q$.

现设 $\exists P \in \mathrm{GL}(m,F),\ Q \in \mathrm{GL}(n,F),\ s.t.\ P \cdot A \cdot Q = \mathrm{diag}(I_{r(A)}, 0_{r(B)}, 0),\ P \cdot B \cdot Q = \mathrm{diag}(0_{r(A)}, I_{r(B)}, 0)$, 则 $A + B = P^{-1} \cdot \mathrm{diag}(I_{r(A)}, I_{r(B)}, 0) \cdot Q^{-1}$, 故 $r(A+B) = r(A) + r(B)$.
(2) 由于 $A \cdot B$ 的每个列向量为 A 的列向量的线性组合, $A \cdot B$ 的每个行向量为 B 的行向量的线性组合, 则

$r(A\cdot B)\le \min\{r(A), r(B)\}$. 进一步地, 注意

$$\begin{aligned} r(A\cdot B)=r(A) &\iff \mathrm{column}(A\cdot B)\supseteq \mathrm{column}(A)\\ &\iff \exists Y\in F^{s\times n},\ s.t.\ A=ABY\\ &\iff \forall X\in F^{n\times 1},\ \exists Y\in F^{s\times 1},\ s.t.\ AX=ABY\\ &\iff F^{n\times 1}=\ker(A)+\mathrm{column}(B);\\ r(A\cdot B)=r(B) &\iff \mathrm{row}(A\cdot B)\supseteq \mathrm{row}(B)\\ &\iff \exists X\in F^{n\times m},\ s.t.\ B=XAB\\ &\iff \ker(A\cdot B)\subseteq \ker(B)\\ &\iff \ker(A)\cap \mathrm{column}(B)=\{0\}. \end{aligned}$$

□

引理 2.3.5 设 $A\in F^{m\times n}, B\in F^{p\times q}, C\in F^{p\times n}$, 则 $r\begin{pmatrix}A & 0\\ C & B\end{pmatrix}=r(A)+r(B)\iff \exists X\in F^{q\times n}, Y\in F^{p\times m}$, $s.t.\ BX+YA=C$.

证明: "⇐": 设 $\exists X\in F^{q\times n}, Y\in F^{p\times m},\ s.t.\ BX+YA=C$, 则由 $\begin{pmatrix}A & 0\\ C & B\end{pmatrix}=\begin{pmatrix}I & 0\\ Y & I\end{pmatrix}\cdot\begin{pmatrix}A & 0\\ 0 & B\end{pmatrix}\cdot\begin{pmatrix}I & 0\\ X & I\end{pmatrix}$ 知, $r\begin{pmatrix}A & 0\\ C & B\end{pmatrix}=r(A)+r(B)$.

"⇒": 由矩阵的相抵标准形, $\exists P_1\in \mathrm{GL}(m,F), P_2\in \mathrm{GL}(p,F), Q_1\in \mathrm{GL}(n,F), Q_2\in \mathrm{GL}(q,F),\ s.t.$

$P_1\cdot A\cdot Q_1=\begin{pmatrix}I_{r(A)} & 0\\ 0 & 0\end{pmatrix}$, $P_2\cdot B\cdot Q_2=\begin{pmatrix}I_{r(B)} & 0\\ 0 & 0\end{pmatrix}$, 则 $\begin{pmatrix}P_1 & 0\\ 0 & P_2\end{pmatrix}\cdot\begin{pmatrix}A & 0\\ C & B\end{pmatrix}\cdot\begin{pmatrix}Q_1 & 0\\ 0 & Q_2\end{pmatrix}=\begin{pmatrix}I_{r(A)} & 0 & 0\\ 0 & 0 & \\ & & I_{r(B)} \\ P_2CQ_1 & & 0\end{pmatrix}$.

记 $P_2CQ_1=\begin{pmatrix}C_1 & C_2\\ C_3 & C_4\end{pmatrix}$, 则由初等行、列变换知,

$$\begin{pmatrix}I_{r(A)} & 0 & 0 & 0\\ 0 & I & 0 & 0\\ -C_1 & 0 & I_{r(B)} & 0\\ -C_3 & 0 & 0 & I\end{pmatrix}\cdot\begin{pmatrix}I_{r(A)} & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ C_1 & C_2 & I_{r(B)} & 0\\ C_3 & C_4 & 0 & 0\end{pmatrix}\cdot\begin{pmatrix}I_{r(A)} & 0 & 0 & 0\\ 0 & I & 0 & 0\\ 0 & -C_2 & I_{r(B)} & 0\\ 0 & 0 & 0 & I\end{pmatrix}=\begin{pmatrix}I_{r(A)} & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & I_{r(B)} & 0\\ 0 & C_4 & 0 & 0\end{pmatrix},$$

故 $\begin{pmatrix}A & 0\\ C & B\end{pmatrix}$ 相抵于 $\begin{pmatrix}I_{r(A)} & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & I_{r(B)} & 0\\ 0 & C_4 & 0 & 0\end{pmatrix}$. 由 $r\begin{pmatrix}A & 0\\ C & B\end{pmatrix}=r(A)+r(B)$ 知, $C_4=0$, 因此

$$\begin{aligned} C&=P_2^{-1}\begin{pmatrix}C_1 & C_2\\ C_3 & 0\end{pmatrix}Q_2^{-1}=P_2^{-1}\begin{pmatrix}I_{r(B)} & 0\\ 0 & 0\end{pmatrix}\cdot\begin{pmatrix}C_1 & C_2\\ 0 & 0\end{pmatrix}Q_2^{-1}+P_2^{-1}\begin{pmatrix}0 & 0\\ C_3 & 0\end{pmatrix}\cdot\begin{pmatrix}I_{r(A)} & 0\\ 0 & 0\end{pmatrix}Q_1^{-1}\\ &=B\cdot Q_2\begin{pmatrix}C_1 & C_2\\ 0 & 0\end{pmatrix}Q_1^{-1}+P_2^{-1}\begin{pmatrix}0 & 0\\ C_3 & 0\end{pmatrix}P_1\cdot A. \end{aligned}$$

令 $X=Q_2\begin{pmatrix}C_1 & C_2\\ 0 & 0\end{pmatrix}Q_1^{-1}$, $Y=P_2^{-1}\begin{pmatrix}0 & 0\\ C_3 & 0\end{pmatrix}P_1$ 即可. □

命题 2.3.6 (秩不等式 (II))

(1) *(Sylvester)* 设 $A\in F^{m\times n}, B\in F^{n\times s}$, 则 $r(A)+r(B)\le r(AB)+n$, 其中 "=" 成立 $\iff \exists X\in F^{s\times n}, Y\in F^{n\times m}$, $s.t.\ BX+YA=I_n$.

(2) *(Frobenius)* 设 $A\in F^{m\times n}, B\in F^{p\times q}, C\in F^{n\times p}$, 则 $r(AC)+r(CB)\le r(ACB)+r(C)$, 其中 "=" 成立 $\iff \exists X\in F^{q\times p}, Y\in F^{n\times m},\ s.t.\ CBX+YAC=C$.

证明: 注意 (1) 可由 (2) 取 $C=I_n$ 得到. 以下由初等行、列变换证明 (2):

$$r(AC)+r(CB)=r\begin{pmatrix}AC&0\\0&CB\end{pmatrix}\le r\begin{pmatrix}AC&0\\C&CB\end{pmatrix}=r\left(\begin{pmatrix}I_m&-A\\0&I_n\end{pmatrix}\cdot\begin{pmatrix}AC&0\\C&CB\end{pmatrix}\cdot\begin{pmatrix}I_p&-B\\0&I_q\end{pmatrix}\right)$$
$$=r\begin{pmatrix}0&-ACB\\C&0\end{pmatrix}=r(ACB)+r(C).$$

再由引理 2.3.5即知等号成立的条件. □

注: 上述秩不等式可归纳推广到 $l\ge 2$ 个矩阵的情形:

$$\forall A_i\in F^{n_{i-1}\times n_i}\ (1\le i\le l),\ r\left(\prod_{i=1}^{l}A_i\right)\ge\sum_{i=1}^{l-1}r(A_iA_{i+1})-\sum_{i=2}^{l-1}r(A_i)\ge\sum_{i=1}^{l}r(A_i)-\sum_{i=1}^{l-1}n_i.$$

推论 2.3.7 *设 $m\ge n$, 则 $\max\{|r(AB)-r(BA)|\colon A\in F^{m\times n},B\in F^{n\times m}\}=\min\left\{\left\lfloor\frac{m}{2}\right\rfloor,n\right\}$.*

证明: 由 Sylvester 不等式知, $\forall A\in F^{m\times n},B\in F^{n\times m}$, $\begin{cases}\max\{0,r(A)+r(B)-n\}\le r(AB)\le\min\{r(A),r(B)\}\\ \max\{0,r(A)+r(B)-m\}\le r(BA)\le\min\{r(A),r(B)\}\end{cases}$,

则 $|r(AB)-r(BA)|\le\min\{r(A),r(B)\}-\max\{0,r(A)+r(B)-m\}\le\frac{r(A)+r(B)}{2}-\frac{r(A)+r(B)-m}{2}=\frac{m}{2}$. 又 $|r(AB)-r(BA)|\le\min\{r(A),r(B)\}\le n$, 故 $|r(AB)-r(BA)|\le\min\left\{\left\lfloor\frac{m}{2}\right\rfloor,n\right\}$. 特别地, 记 $k=\min\left\{\left\lfloor\frac{m}{2}\right\rfloor,n\right\}$, 取 $A=\begin{pmatrix}0&I_k\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&I_k\end{pmatrix}$, 则 $r(AB)=k$, $r(BA)=0$. □

以下我们研究方阵多项式的一些秩的性质.

例 2.3.1 (方阵多项式代数) 设 F 是一个域, $A\in F^{n\times n}$, 考虑代数同态

$$F[X]\longrightarrow F^{n\times n}\quad.$$
$$f(X)=a_mX^m+\cdots+a_1X+a_0\longmapsto f(A)=a_mA^m+\cdots+a_1A+a_0I_n$$

注意在一元多项式代数 $F[X]$ 中, 未定元 X 是域 F 上的超越元 (transcendental element) (即不满足任何域 F 上的代数关系), 但它在上述同态下的像 A 一定满足域 F 上的某个代数关系. 这是因为, 方阵代数 $F^{n\times n}$ 是域 F 上的 n^2 维线性空间, 而 $I_n,A,\cdots,A^{n^2}\in F^{n\times n}$ 为 (n^2+1) 个元, 故它们必线性相关, 即存在不全为零的 $a_0,a_1,\cdots,a_{n^2}\in F$, 满足 $a_0I_n+a_1A+\cdots+a_{n^2}A^{n^2}=0$, 因此 $M_A=\{f(X)\in F[X]\colon f(A)=0\}\ne\{0\}$. 记 $d=\min\limits_{f(X)\in M_A\setminus\{0\}}\deg(f)$, 以及 $p(X)\in M_A$ 满足 $\deg(p(X))=d$. 不妨设 $p(X)$ 首一, 则:

(1) $\{I_n,A,\cdots,A^{d-1}\}$ 为 $F[A]\coloneqq\{f(A)\colon f(X)\in F[X]\}$ 的 F-基; 特别地, $\dim_F(F[A])=d$;

(2) M_A 是 $F[X]$ 的理想, 以 $p(X)$ 为唯一的首一生成元, 故上述代数同态可分解为

$$F[X]\longrightarrow F[X]/M_A\xrightarrow{\simeq}F[A]\subseteq F^{n\times n}\ .$$
$$f(X)\longmapsto f(X)+M_A\longmapsto f(A)$$

引理 2.3.8 *设 F 为一个域, $A\in F^{n\times n}$, 则 $\dim_F(F[A])\le n$.*

证明: 固定 $\alpha\in F^{n\times 1}\setminus\{0\}$. 取 $m=\min\{i\in\mathbb{N}\colon\{\alpha,A\alpha,\cdots,A^i\alpha\}$线性相关$\}$, 则 $1\le m\le n$, 故 $\exists g(X)\in F[X]\setminus\{0\}$, $s.t.\ \deg(g(X))=m$ 且 $g(A)\alpha=0$. 注意 $\{\alpha,A\alpha,\cdots,A^{m-1}\alpha\}$ 线性无关, 且 $\forall 0\le i\le m-1$, $g(A)(A^i\alpha)=0$, 故 $\dim_F(\ker(g(A)))\ge m$. 因此 $\dim_F(\mathrm{column}(g(A)))=r(g(A))\le n-m$. 由对 n 归纳假设知, $\exists f(X)\in F[X]\setminus\{0\}$, $s.t.\ \deg(f(X))\le r(g(A))$ 且 $f(A)\big|_{\mathrm{column}(g(A))}=0$, 故 $\deg(f(X)g(X))\le n$, 且 $f(A)g(A)=0$. □

推论 2.3.9 *设 F 为一个域, $A\in F^{n\times n}$, 则 $\dim_F(F[A])\le\min\{n,\ r(A)+1\}$.*

证明: 记 $r=r(A)$. 由满秩分解知, $\exists C\in F^{n\times r}$, $R\in F^{r\times n}$, $s.t.\ A=C\cdot R$. 由 $R\cdot C\in F^{r\times r}$ 以及命题 2.3.8知, $\exists g(X)\in F[X]\setminus\{0\}$, $s.t.\ \deg(g(X))\le r$ 且 $g(R\cdot C)=0$. 记 $f(X)=X\cdot g(X)\in F[X]\setminus\{0\}$, 则 $\deg(f(X))\le r+1$, 且 $f(A)=f(C\cdot R)=CR\cdot g(CR)=C\cdot g(RC)\cdot R=0$. □

命题 2.3.10 *设 F 为一个域, $A\in F^{n\times n}$, 则*

(1) *$\{r(A^i)\}_{i\ge0}$ 是 $\mathbb{N}$ 中的递减序列; 特别地, 它必在某刻静止, 即 $\exists k\ge0$, $s.t.\ r(A^k)=r(A^{k+1})$.*

(2) *$\{r(A^i)\}_{i\ge0}$ 为 $\mathbb{N}$ 中的凸序列, 即 $\forall i\ge1$, $2r(A^i)\le r(A^{i+1})+r(A^{i-1})$.*

(3) 记 $s(A) \coloneqq \min\{k \geq 0 \colon r(A^k) = r(A^{k+1})\}$, 则 $s(A) \leq n$, 且 $\forall i \geq 0$, $r(A^{s(A)}) = r(A^{s(A)+i})$.

(4) $\dim_F(F[A]) = \sum\limits_{c \in \overline{F}^{\mathrm{alg}}} s(cI_n - A)$.

证明: (1) 由秩不等式 (I) 即知;

(2) 由 Frobenius 不等式即知.

(3) 由于 $n = r(I_n) > r(A) > \cdots > r(A^{s(A)-1}) > r(A^{s(A)}) \geq 0$, 则 $s(A) \leq n$. 由于 $r(A^{s(A)}) = r(A^{s(A)+1})$, 则由 (2) 知 $r(A^{s(A)+2}) \geq 2r(A^{s(A)+1}) - r(A^{s(A)}) = r(A^{s(A)})$; 又 $r(A^{s(A)+2}) \leq r(A^{s(A)})$, 则 $r(A^{s(A)+2}) = r(A^{s(A)})$. 类似知 $\forall i \geq 0$, $r(A^{s(A)+i}) = r(A^{s(A)})$.

(4) 不妨设 F 为代数闭域, 则由 Jordan 标准形即知. □

命题 2.3.11 设 F 为一个域, $A, B \in F^{n\times n}$, 则 $|s(AB) - s(BA)| \leq 1$, 且 $r((AB)^{s(AB)}) = r((BA)^{s(BA)})$.

证明: 由 $A(BA)^kB = (AB)^{k+1}$; $B(AB)^kA = (BA)^{k+1}$ 以及秩不等式 (I) 即知. □

命题 2.3.12 设 F 为一个域, $A \in F^{n\times n}$, $f(X), g(X) \in F[X]$ 互素, 则 $f(A)\cdot g(A) = 0 \iff r(f(A)) + r(g(A)) = n$.

证明: 由 $f(X), g(X) \in F[X]$ 互素以及 Bezout 定理知, $\exists u(X), v(X) \in F[X]$, $s.t.\ u(X)\cdot f(X) + v(X)\cdot g(X) = 1$, 则 $u(A)\cdot f(A) + v(A)\cdot g(A) = I_n$, 故由初等行、列变换知,

$$r(f(A)) + r(g(A)) = r\begin{pmatrix} f(A) & 0 \\ 0 & g(A) \end{pmatrix} = r\left(\begin{pmatrix} I_n & 0 \\ u(A) & I_n \end{pmatrix}\cdot\begin{pmatrix} f(A) & 0 \\ 0 & g(A) \end{pmatrix}\cdot\begin{pmatrix} I_n & 0 \\ v(A) & I_n \end{pmatrix}\right) = r\begin{pmatrix} f(A) & 0 \\ I_n & g(A) \end{pmatrix}$$

$$= r\left(\begin{pmatrix} I_n & -f(A) \\ 0 & I_n \end{pmatrix}\cdot\begin{pmatrix} f(A) & 0 \\ I_n & g(A) \end{pmatrix}\cdot\begin{pmatrix} I_n & -g(A) \\ 0 & I_n \end{pmatrix}\right) = r\begin{pmatrix} 0 & -f(A)\cdot g(A) \\ I_n & 0 \end{pmatrix} = r(f(A)\cdot g(A)) + n,$$

因此 $f(A)\cdot g(A) = 0 \iff r(f(A)) + r(g(A)) = n$. □

最后这个命题呼应首节: 将矩阵理论应用于线性方程组理论.

命题 2.3.13 设 $A \in F^{m\times n}$, 则:

(1) A 为列满秩阵, 即 $\dim_F(\mathrm{column}(A)) = n$

$\iff$ 齐次线性方程组 $AX = 0$ 只有零解

$\iff$ 若 $\{\alpha_1, \cdots, \alpha_s\} \subseteq F^{n\times 1}$ 线性无关, 则 $\{A\alpha_1, \cdots, A\alpha_s\} \subseteq F^{m\times 1}$ 线性无关

$\iff$ A 有左逆, 即 $\exists P \in F^{n\times m}$, $s.t.\ P\cdot A = I_n$

$\iff$ A 可扩充为 $(A, A') \in \mathrm{GL}(m, F)$

(2) A 为行满秩阵, 即 $\dim_F(\mathrm{row}(A)) = m$

$\iff$ 对于任意 $Y \in F^{m\times 1}$, 非齐次线性方程组 $AX = Y$ 的解必存在

$\iff$ 若 $\{\alpha_1, \cdots, \alpha_s\} \subseteq F^{n\times 1}$ 可线性生成 $F^{n\times 1}$, 则 $\{A\alpha_1, \cdots, A\alpha_s\} \subseteq F^{m\times 1}$ 可线性生成 $F^{m\times 1}$

$\iff$ A 有右逆, 即 $\exists Q \in F^{n\times m}$, $s.t.\ A\cdot Q = I_m$

$\iff$ A 可扩充为 $\begin{pmatrix} A \\ A'' \end{pmatrix} \in \mathrm{GL}(n, F)$

证明: (1) 第一个等价号是因为矩阵乘法 AX 即对 A 的列向量组做线性组合; 第二个等价号由定义显然; 第三个等价号可由基对应而构造线性映射的左逆; 第四个等价号由相抵标准形与秩的定义即知.

(2) 第一个等价号是因为矩阵乘法 AX 即对 A 的列向量组做线性组合, 而列空间的维数等于行空间的维数; 第二个等价号由定义显然; 第三个等价号可取定原像而构造线性映射的右逆; 第四个等价号由相抵标准形与秩的定义即知. □

注: 上述命题的证明暴露了目前的两个缺陷: 一是并未给出矩阵行空间的直观解释, 本质原因在于暂时无法理解行列空间的对偶关系; 二是发现构造矩阵运算往往并不容易, 本质原因在于矩阵技巧往往掩盖了线性映射的观点. 我们将在线性映射初步一章解决这两个问题.

习题 2.3 (对合阵与幂等阵的刻画) 设 $A \in F^{n\times n}$, 则:

(1) A 为幂等的 *(idempotent)*, 即 $A^2 = A \iff r(A) + r(A - I_n) = n \iff A$ 相似于 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

(2) 当 $\mathrm{char}(F) \neq 2$ 时, A 为对合的 *(involutary)*, 即 $A^2 = I_n \Longleftrightarrow r(A + I_n) + r(A - I_n) = n \Longleftrightarrow A$ 相似于 $\begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}$. 特别地, A 为幂等的 $\Longleftrightarrow (2A - I_n)$ 为对合的.

(提示: 利用秩不等式、矩阵打洞技术以及矩阵的相抵标准形.)

参考文献与补注 2.3

(1) 关于 Smith 标准形的部分, 可以参考维基百科或者 P. Morandi "The Smith Normal Form of a Matrix"(2005).

(2) 关于环论基本定义的部分, 可以参考 T. Y. Lam "Lectures on Modules and Rings".

第 3 章　线性映射初步

本章初步介绍线性代数的核心概念: 线性映射 (linear map), 与它相关的观点和技术不仅哺育了代数学领域的各个分支, 也在许多复杂的几何、物理模型中有基本而广泛的应用.

§3.1　线性映射基本定理

3.1.1　线性代数基本定理

设 V 为域 F 上的线性空间, B 为 V 的 F-基, 则坐标映射 Γ_B: $V \longrightarrow\!\!\!\!\!\rightarrow F^{(B)} \coloneqq \coprod\limits_{\alpha\in B} F$ 为

$$\text{(有限和)}\ \sum_{\alpha\in B} c_\alpha \alpha \longmapsto (c_\alpha)_{\alpha\in B}$$

F-线性同构. 于是, 对于任意 F-线性空间 W, 为决定 F-线性映射 $T\colon V\to W$, 只需决定 T 在集合 B 上的取值, 即存在自然双射 $\mathrm{Hom}_{\mathbf{F-Mod}}(F^{(B)},W)\cong \mathrm{Hom}_{\mathbf{Set}}(B,\mathscr{G}(W))$, 这里 $\mathscr{G}(W)$ 为 W 的集合结构. 这表明在范畴 **Set** 与 **F-Mod** 之间存在一对 **伴随函子** (adjoint functor) $F^{(\cdot)}\colon \mathbf{Set}\rightleftarrows \mathbf{F\text{-}Mod}\colon \mathscr{G}$, 其中 $F^{(\cdot)}$ 为在给定集合上自由生成 F-直和空间, $\mathscr{G}$ 为忘记给定线性空间的结构而只将其视为集合.

命题 3.1.1 *设 V,W 为域 F 上的两个线性空间, $T\in L(V,W)\coloneqq \mathrm{Hom}_{\mathbf{F-Mod}}(V,W)$, 则:*

(1) *T 为单射 $\Longleftrightarrow$ T 存在左逆;*

(2) *T 为满射 $\Longleftrightarrow$ T 存在右逆.*

证明: (1) "$\Leftarrow$": 显然; "$\Rightarrow$": 设 $T\in L(V,W)$ 为单射, 则限制到它的像上为 F-线性同构 $\widetilde{T}\colon V\to \mathrm{Im}(T)$, 记它的逆为 $\widetilde{S}\in L(\mathrm{Im}(T),V)$. 取 $\mathrm{Im}(T)$ 在 W 中的一个直和补空间 U, 则直和分解式给出了 F-线性同构 $L(W,V)\cong L(\mathrm{Im}(T),V)\oplus L(U,V)$. 令 $S\in L(W,V)$ 为 $S\big|_{\mathrm{Im(T)}}=\widetilde{S}$, $S\big|_U=0$, 则 $S\circ T=\mathrm{id}_V$.

(2) "$\Leftarrow$": 显然; "$\Rightarrow$": 设 $T\in L(V,W)$ 为满射, 则 $\forall \alpha'\in W$, $T^{-1}(\alpha')\neq\varnothing$. 取 W 的基 B', 由选择公理知, 存在选择映射 $\{T^{-1}(b')\}_{b'\in B'} \longrightarrow\!\!\!\!\!\rightarrow \bigcup\limits_{b'\in B'} T^{-1}(b')$, 则也存在复合映射 $B' \longrightarrow\!\!\!\!\!\rightarrow \bigcup\limits_{b'\in B'} T^{-1}(b')\subseteq V$. 再将它线性延拓到 W 上即得 $S\in L(W,V)$, 且 $T\circ S=\mathrm{id}_W$.

$$T^{-1}(b')\longmapsto \alpha_{b'}\in T^{-1}(b') \qquad\qquad b'\longmapsto \alpha_{b'}\in T^{-1}(b')$$

□

注:

(1) 上述命题实际上是集合论中相应命题的线性空间版本, 这种翻译的过程体现了上述范畴论中伴随函子的思想. 其中, (2) 的证明不可避免地用到了选择公理. 事实上, "任意两个非空集合之间的满射都存在右逆" 等价于选择公理.

(2) 特别地, 当 V,W 均为有限维线性空间时, 考虑线性映射的矩阵表示, 则上述命题回顾了上一章的最后命题.

接着我们换一种观点看待线性映射, 它启发了我们在一般 Lie 群中考虑格的刚性与算术性.

例 3.1.1 回忆域 F 上的线性空间 V 可视为在一个加法交换群 $(V,+,0)$ 上配备了与之相容的 F-数乘作用, 取 V 的 F-基 B 可生成 $(V,+,0)$ 的一个子群 $\left(\Gamma\coloneqq\bigoplus\limits_{b\in B}\mathbb{Z}b,+,0\right)$. 于是域 F 上两个线性空间 V,W 之间的 F-线性映射可视为保持 F-数乘的加法交换群同态, 且它完全由限制在子群 $(\Gamma,+,0)$ 上的取值决定.

特别地, 当 $F=\mathbb{R}$ 或 $\mathbb{C}$, $V=F^n$ 时, 在通常的拓扑下 Γ 为 V 中的一个离散子群, 且 V/Γ 具有有限的平移不变测度, 则 Γ 为 V 中的一个**格** (lattice) (注意需要与序理论中的格相区别). 此时对于任意线性空间 W, $T\in L(V,W)$ 完全由 $T\big|_\Gamma$ 决定, 这体现了格 Γ 在群 V 中的**刚性** (rigidity); 另外, 任给 V 中的一个格 Γ', 均存在 $(V,+,0)$ 的群自同构 T 满足 $T(\Gamma)=\Gamma'$, 这体现了格 Γ 在群 V 中的**算术性** (arithmeticity).

关于线性映射最基本的定理如下, 它包含了所谓的基本同构定理与秩-零度定理:

定理 3.1.2 (线性代数基本定理) 设 V,W 为域 F 上的两个线性空间, $T\in L(V,W)$, 则存在线性同构
$\widetilde{T}$: $V/\ker(T)\longrightarrow \mathrm{Im}(T)$. 特别地, 当 $\dim_F(V)<+\infty$ 时, $\dim_F(V)=\dim_F(\ker(T))+\dim_F(\mathrm{Im}(T))$.

$\alpha+\ker(T)\longmapsto T(\alpha)$

证明: 注意到上述商空间与像空间之间存在基对应即可. □

注:

(1) 上述定理的证明实际上先通过基对应构造了加法交换群的同构, 再延拓到了 F-线性空间上, 于是更一般的结论是: "设 $(G,p,1)$, $(H,p',1')$ 为两个群, φ: $(G,p,1)\to(H,p',1')$ 为群同态, 则存在群同构
$\widetilde{\varphi}$: $G/\ker(\varphi)\longrightarrow \mathrm{Im}(\varphi)$ ", 这是群的基本同构定理. 类似地也可陈述环、模、代数的基本同构定理.

$p(a,\ker(\varphi))\longmapsto\varphi(a)$

(2) 设 $T\in L(V,W)$, 则上述基本同构 $\widetilde{T}$ 实际上给出了 T 的**满单分解** (epi-mono factorization):
T: $V\longrightarrow V/\ker(T)\xrightarrow{\widetilde{T}}\mathrm{Im}(T)$, 且它在商空间同构的意义下是唯一的. 类似地, 群、环、模、代数的基本

$\alpha\longmapsto\alpha+\ker(T)\longmapsto T(\alpha)$

同构也都给出了相应的满单分解, 这是在**正则范畴** (regular category) 中的普遍现象.

(3) 设 $T\in L(V,W)$, 取 $\ker(T)$ 在 V 中的直和补空间 U, 则存在复合的线性同构 $U\xrightarrow{\cong}V/\ker(T)\xrightarrow{\widetilde{T}}\mathrm{Im}(T)$,

$\alpha\longmapsto\alpha+\ker(T)\longmapsto T(\alpha)$

故也存在线性同构 $V\cong\ker(T)\oplus\mathrm{Im}(T)$, 此式称为上述基本同构 $\widetilde{T}$ 的 **可裂性** (splitness). 注意上述群的基本同构未必具有可裂性, 这里有两部分的原因: 一是对于非 Abel 群来说, 子群未必具有补子群; 二是对于 Abel 群来说, 单的群同态未必具有左逆, 满的群同态未必具有右逆. 在一般的 Abel 范畴中, 一个短正合列的可裂性等价于 (左可裂性) 该单态射存在左逆, 也等价于 (右可裂性) 该满态射存在右逆.

(4) 我们将上述基本同构 $\widetilde{T}$ 写成**短正合列** (short exact sequence) 的形式: $0\to\ker(T)\hookrightarrow V\xrightarrow{T}\mathrm{Im}(T)\to 0$, 于是秩-零度定理即当 $\dim_F(V)<+\infty$ 时, 上述短正合列中各项维数的交错和为 0. 一般地, 设 $0\to V_1\to V_2\to\cdots\to V_s\to 0$ 为若干有限维线性空间的正合列, 则同理有 $\sum\limits_{i=1}^{s}(-1)^i\dim_F(V_i)=0$.

(5) 设 V,W 为域 F 上的两个有限维线性空间, $T\in L(V,W)$, 记 $\mathrm{index}(T)\coloneqq\dim_F(\ker(T))-\dim_F(\mathrm{coker}(T))$, 其中 $\mathrm{coker}(T)\coloneqq W/\mathrm{Im}(T)$. 直观上看, $\dim_F(\ker(T))$ 是线性方程 $T(\alpha)=0$ 的线性无关解的个数, $\dim_F(\mathrm{coker}(T))$ 是使得线性方程 $T(\alpha)=\beta$ 存在解需要对 $\beta\in W$ 加的线性无关限制条件的个数. 于是秩-零度定理即此时 $\mathrm{index}(T)=\dim_F(V)-\dim_F(W)$. 这表明即使并不仔细分析 $T\in L(V,W)$ 的性质, 我们也可以从它所涉及的线性空间直接读出 $\mathrm{index}(T)$ 的值. 这种现象在深刻的 Atiyah-Singer 指标定理中也有所体现: 定向闭流形上椭圆算子的解析指标 (与方程解空间的维数有关) 等于它的拓扑指标 (与流形、向量丛的拓扑量有关).

以下利用线性代数基本定理讨论一些与维数有关的问题:

例 3.1.2 (子空间复形) 设 V 为域 F 上的线性空间, $\{W_i\}_{i=1}^s$ 为 V 的若干子空间, 则存在子空间复形

$$0\to\bigcap_i W_i\to\coprod_i\bigcap_{j\neq i}W_j\to\cdots\to\coprod_{i,j}W_i\cap W_j\to\coprod_i W_i\to\sum_i W_i\to 0,$$

其中左端 0 后的第 $1\le k\le s$ 项是诸 $W_i\,(1\le i\le s)$ 的各种 $(s-k+1)$ 次交的直和, 第 $(s+1)$ 项是诸 $W_i\,(1\le i\le s)$ 的和; 第 $(s-k+1)$ 个非零映射将 $\alpha\in W_{j_1}\cap\cdots\cap W_{j_k}$ 映为 $((-1)^i\alpha)_i\in\coprod\limits_i(W_{j_1}\cap\cdots\cap\widehat{W_{j_i}}\cap\cdots\cap W_{j_k})$, 最后一个非零映射将 $(\alpha_i)_i\in\coprod\limits_i W_i$ 映为 $\sum\limits_i\alpha_i\in\sum\limits_i W_i$. 于是可直接验证上述任意相邻映射的复合均为 0.

例如当 $s=2$ 时, 上述子空间复形为

$$0\to W_1\cap W_2\to W_1\coprod W_2\to W_1+W_2\to 0,$$

其中第一个非零映射为 $\alpha\mapsto(\alpha,-\alpha)$, 第二个非零映射为 $(\alpha_1,\alpha_2)\mapsto\alpha_1+\alpha_2$. 这是正合列 (即任意前一映射的像等于后一映射的核), 故当 $\dim_F(W_i)<+\infty\,(1\le i\le 2)$ 时, 由秩-零度定理可得维数公式

$$\dim_F(W_1\cap W_2)-\dim_F(W_1)-\dim_F(W_2)+\dim_F(W_1+W_2)=0.$$

当 $s=3$ 时, 上述子空间复形为

$$0\to W_1\cap W_2\cap W_3\to(W_2\cap W_3)\coprod(W_1\cap W_3)\coprod(W_1\cap W_2)\to W_1\coprod W_2\coprod W_3\to W_1+W_2+W_3\to 0,$$

其中第一个非零映射为 $\alpha \mapsto (-\alpha, \alpha, -\alpha)$, 第二个非零映射为 $(\alpha_1, \alpha_2, \alpha_3) \mapsto (\alpha_2 + \alpha_3, \alpha_1 - \alpha_3, -\alpha_1 - \alpha_2)$, 第三个非零映射为 $(\beta_1, \beta_2, \beta_3) \mapsto \beta_1 + \beta_2 + \beta_3$. 这不是正合列 (注意第二个映射的像未必等于第三个映射的核, 即 "+" 与 "∩" 未必具有分配律), 故当 $\dim_F(W_i) < +\infty\,(1 \leq i \leq 3)$ 时, 由秩-零度定理可得维数不等式

$$\begin{aligned}&\dim_F(W_1 \cap W_2 \cap W_3) - \dim_F(W_2 \cap W_3) - \dim_F(W_1 \cap W_3) - \dim_F(W_1 \cap W_2)\\ &+ \dim_F(W_1) + \dim_F(W_2) + \dim_F(W_3) - \dim_F(W_1 + W_2 + W_3) \geq 0.\end{aligned}$$

当 $s \geq 4$ 时, 上述子空间复形中非正合的位置可能更多, 于是无法得到一般的维数不等式. 但总有最后一个非零映射为满射, 故 $\sum\limits_i \dim_F(W_i) \geq \dim_F(\sum\limits_i W_i)$.

例 3.1.3 (商空间复形) 设 V 为域 F 上的线性空间, $\{W_i\}_{i=1}^s$ 为 V 的若干子空间, 则存在商空间复形

$$0 \to V/\bigcap_i W_i \to \coprod_i V/\bigcap_{j\neq i} W_j \to \cdots \to \coprod_{i,j} V/W_i \cap W_j \to \coprod_i V/W_i \to V/\sum_i W_i \to 0,$$

其中左端 0 后的第 $1 \leq k \leq s$ 项是 V 关于诸 $W_i\,(1 \leq i \leq s)$ 的各种 $(s-k+1)$ 次交的直和的商, 第 $(s+1)$ 项是 V 关于诸 $W_i\,(1 \leq i \leq s)$ 的和的商; 第 $(s-k+1)$ 个非零映射将 $\alpha + W_{j_1} \cap \cdots \cap W_{j_k}$ 映为 $((-1)^i\alpha + (W_{j_1} \cap \cdots \cap \widehat{W_{j_i}} \cap \cdots \cap W_{j_k}))_i$, 最后一个非零映射将 $(\alpha_i + W_i)_i$ 映为 $\sum\limits_i \alpha_i + \sum\limits_i W_i$. 于是可直接验证上述任意相邻映射的复合均为 0.

例如当 $s = 2$ 时, 上述商空间复形为

$$0 \to V/W_1 \cap W_2 \to V/W_1 \coprod V/W_2 \to V/(W_1 + W_2) \to 0,$$

其中第一个非零映射为 $\alpha + W_1 \cap W_2 \mapsto (\alpha + W_1, -\alpha + W_2)$, 第二个非零映射为 $(\alpha_1 + W_1, \alpha_2 + W_2) \mapsto (\alpha_1 + \alpha_2) + (W_1 + W_2)$. 这是正合列 (即任意前一映射的像等于后一映射的核), 故当 $\operatorname{codim}_F(W_i) \coloneqq \dim_F(V/W_i) < +\infty\,(1 \leq i \leq 2)$ 时, 由秩-零度定理可得余维数公式

$$\operatorname{codim}_F(W_1 \cap W_2) - \operatorname{codim}_F(W_1) - \operatorname{codim}_F(W_2) + \operatorname{codim}_F(W_1 + W_2) = 0.$$

当 $s = 3$ 时, 上述商空间复形为

$$\begin{aligned}&0 \to V/W_1 \cap W_2 \cap W_3 \to V/(W_2 \cap W_3) \coprod V/(W_1 \cap W_3) \coprod V/(W_1 \cap W_2)\\ &\to V/W_1 \coprod V/W_2 \coprod V/W_3 \to V/(W_1 + W_2 + W_3) \to 0,\end{aligned}$$

其中第一个非零映射为 $\alpha + W_1 \cap W_2 \cap W_3 \mapsto (-\alpha + (W_2 \cap W_3), \alpha + (W_1 \cap W_3), -\alpha + (W_1 \cap W_2))$, 第二个非零映射为 $(\alpha_1 + (W_2 \cap W_3), \alpha_2 + (W_1 \cap W_3), \alpha_3 + (W_1 \cap W_2)) \mapsto ((\alpha_2 + \alpha_3) + W_1, (\alpha_1 - \alpha_3) + W_2, (-\alpha_1 - \alpha_2) + W_3)$, 第三个非零映射为 $(\beta_1 + W_1, \beta_2 + W_2, \beta_3 + W_3) \mapsto (\beta_1 + \beta_2 + \beta_3) + (W_1 + W_2 + W_3)$. 这不是正合列 (注意第一个映射的像未必等于第二个映射的核), 故当 $\operatorname{codim}_F(W_i) < +\infty\,(1 \leq i \leq 3)$ 时, 由秩-零度定理可得余维数不等式

$$\begin{aligned}&\operatorname{codim}_F(W_1 \cap W_2 \cap W_3) - \operatorname{codim}_F(W_2 \cap W_3) - \operatorname{codim}_F(W_1 \cap W_3) - \operatorname{codim}_F(W_1 \cap W_2)\\ &+ \operatorname{codim}_F(W_1) + \operatorname{codim}_F(W_2) + \operatorname{codim}_F(W_3) - \operatorname{codim}_F(W_1 + W_2 + W_3) \leq 0.\end{aligned}$$

当 $s \geq 4$ 时, 上述商空间复形中非正合的位置可能更多, 于是无法得到一般的余维数不等式. 但总有

$$\begin{array}{ccc} V/\bigcap\limits_i W_i & \longrightarrow & \coprod\limits_i V/W_i \\ \alpha + \bigcap\limits_i W_i & \longmapsto & (\alpha + W_i)_i \end{array} \quad \text{为单射, 故 } \operatorname{codim}_F(\bigcap_i W_i) \leq \sum_i \operatorname{codim}_F(W_i).$$

最后我们将线性代数基本定理应用于矩阵理论.

推论 3.1.3 *设 F 为一个域, $A \in F^{m\times n}$, 则存在线性同构 $\widetilde{L_A}$:*

$$\begin{array}{ccc} F^{n\times 1}/\ker(A) & \longrightarrow & \operatorname{column}(A) \\ \alpha + \ker(A) & \longmapsto & A\alpha \end{array}.$$

特别地, $n = \dim_F(\ker(A)) + \dim_F(\operatorname{column}(A))$.

注: 结合线性方程组理论 $\dim_F(\ker(A)) = n - \dim_F(\operatorname{row}(A))$, 我们再一次得到了 $\dim_F(\operatorname{row}(A)) = \dim_F(\operatorname{column}(A))$. 于是与矩阵的秩相关的一些不等式都可以用线性映射的观点重新证明.

例 3.1.4 (秩不等式 (II) 的重证)

(1) *(Sylvester) 设 $A \in F^{m\times n}, B \in F^{n\times s}$, 则 $r(A) + r(B) \leq r(AB) + n$, 其中 "=" 成立 $\iff \ker(A) \subseteq \operatorname{column}(B)$.*

(2) *(Frobenius)* 设 $A\in F^{m\times n},B\in F^{n\times p},C\in F^{p\times q}$, 则 $r(AB)+r(BC)\leq r(ABC)+r(B)$, 其中 "=" 成立 $\iff \ker(A)\cap \mathrm{column}(BC)=\ker(A)\cap\mathrm{column}(B)$.

证明: 注意 (1) 可由 (2) 取 $B=I_n$ 得到. 以下由线性代数基本定理证明 (2): 考虑线性映射 $T\colon \mathrm{column}(BC)\longrightarrow F^{m\times 1}$, $BC\alpha\longmapsto ABC\alpha$

则

$$\begin{aligned} r(BC)&=\dim_F(\mathrm{column}(BC))=\dim_F(\ker(A)\cap\mathrm{column}(BC))+\dim_F(\mathrm{column}(ABC))\\ &\leq \dim_F(\ker(A)\cap\mathrm{column}(B))+r(ABC)=\dim_F(\mathrm{column}(B))-\dim_F(\mathrm{column}(AB))+r(ABC)\\ &=r(B)-r(AB)+r(ABC). \end{aligned}$$

□

推论 3.1.4

(1) *(Sylvester)* 设 $A\in F^{m\times n},B\in F^{n\times s}$, 则 $\ker(A)\subseteq\mathrm{column}(B)\iff \exists X\in F^{s\times n},Y\in F^{n\times m},\ s.t.\ BX+YA=I_n$.

(2) *(Frobenius)* 设 $A\in F^{m\times n},B\in F^{n\times p},C\in F^{p\times q}$, 则 $\ker(A)\cap\mathrm{column}(BC)=\ker(A)\cap\mathrm{column}(B)$ $\iff \exists X\in F^{q\times p},Y\in F^{n\times m},s.t.\ BCX+YAB=B$.

例 3.1.5 (Kronecker-Weyl 定理的约化) 记 $\mathbb{T}^n:=\mathbb{R}^n/\mathbb{Z}^n$ 为实 n 维环面, $[x]:=x+\mathbb{Z}^n\,(x\in\mathbb{R}^n)$ 为环面中的点, 每个 $v\in\mathbb{R}^n$ 决定了环面上的一个光滑流, 即群同态 $\varphi\colon (\mathbb{R},+,0)\longrightarrow (C^\infty(\mathbb{T}^n;\mathbb{T}^n),\circ,\mathrm{id}_{\mathbb{T}^n})$, 则环面中每个点 $t\longmapsto(\varphi_t\colon [x]\mapsto[x+tv])$ 在此光滑流下的轨道闭包都是一个子环面. 具体地说, 存在线性子空间 $V\subseteq\mathbb{R}^n$, 满足:

(1) $v\in V$, 从而 $\forall x\in\mathbb{R}^n$, $\{\varphi_t([x])\in\mathbb{T}^n\colon t\in\mathbb{R}\}\subseteq[x+V]$;

(2) $\forall x\in\mathbb{R}^n$, $[x+V]\subseteq\mathbb{T}^n$ 为闭子集, 从而光滑同胚于 $\mathbb{T}^k$, 其中 $k=\dim_{\mathbb{R}}(V)=\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(\{v_1,\cdots,v_n\}))$;

(3) $\forall x\in\mathbb{R}^n$, $\{\varphi_t([x])\in\mathbb{T}^n\colon t\in\mathbb{R}\}\subseteq[x+V]$ 为稠子集.

证明: 我们利用线性代数将此定理约化为一个纯数论的定理.

记 $V:=\{w\in\mathbb{R}^n\colon \forall\lambda_1,\cdots,\lambda_n\in\mathbb{Q},\ \sum_{i=1}^n\lambda_iv_i=0\Rightarrow\sum_{i=1}^n\lambda_iw_i=0\}$, 则 $V\subseteq\mathbb{R}^n$ 为实线性子空间, 且 $v\in V$. 考虑线性映射 $T\colon \mathbb{Q}^n\longrightarrow\mathbb{R}$, 则 $\ker(T)\subseteq\mathbb{Q}^n$ 为 $\mathbb{Q}$-线性子空间, $\mathrm{Im}(T)=\mathrm{Span}_{\mathbb{Q}}(\{v_1,\cdots,v_n\})\subseteq\mathbb{R}$ 为

$$\lambda\longmapsto\sum_{i=1}^n\lambda_iv_i$$

$\mathbb{Q}$-线性子空间. 由秩-零度定理知, $\dim_{\mathbb{Q}}(\ker(T))=n-\dim_{\mathbb{Q}}(\mathrm{Im}(T))=n-\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(\{v_1,\cdots,v_n\}))$. 记线性函数 $f_\lambda\colon \mathbb{R}^n\longrightarrow\mathbb{R}$ $(\lambda\in\mathbb{R}^n)$, 则 $V=\bigcap_{\lambda\in\ker(T)}\ker(f_\lambda)=\mathrm{Span}_{\mathbb{R}}\left(\bigcap_{\lambda\in\ker(T)}\ker(f_\lambda\big|_{\mathbb{Q}^n})\right)$ (这是线性方程组 $w\longmapsto\sum_{i=1}^n\lambda_iw_i$ 理论), 故 $\dim_{\mathbb{R}}(V)=\dim_{\mathbb{Q}}\left(\bigcap_{\lambda\in\ker(T)}\ker(f_\lambda\big|_{\mathbb{Q}^n})\right)=n-\dim_{\mathbb{Q}}(\ker(T))=\dim_{\mathbb{Q}}(\mathrm{Span}_{\mathbb{Q}}(\{v_1,\cdots,v_n\}))$.

现取 $\bigcap_{\lambda\in\ker(T)}\ker(f_\lambda\big|_{\mathbb{Q}^n})\subseteq\mathbb{Q}^n$ 的一个 $\mathbb{Q}$-基 B , 则 B 也为 V 的 $\mathbb{R}$-基. 通过去分母可不妨设 $B\subseteq\mathbb{Z}^n$, 则 B 也为 $V\cap\mathbb{Z}^n$ 的 $\mathbb{Z}$-基. 记 $B=\{w^{(i)}\}_{i=1}^k\,(k=\dim_{\mathbb{R}}(V))$, $v=\sum_{i=1}^kv_i'w^{(i)}$ $(v_i'\in\mathbb{R})$, 断言: $\{v_i'\}_{i=1}^k$ 在 $\mathbb{Q}$ 上线性无关. (这是因为, 考虑 $\mathbb{Q}$-线性组合 $\sum_{i=1}^k\mu_iv_i'=0$ $(\mu_i\in\mathbb{Q})$, 由 $\ker(T)\oplus\mathrm{Span}_{\mathbb{Q}}(\{w^{(i)}\}_{i=1}^k)=\mathbb{Q}^n$ 知, 可取 $z\in V,\ s.t.\ f_{w^{(i)}}(z)=\mu_i$, $\forall 1\leq i\leq k$, 故 $f_v(z)=\sum_{i=1}^kf_{w^{(i)}}(z)v_i'=\sum_{i=1}^k\mu_iv_i'=0$. 因此由定义知 $\forall w\in V,\ f_w(z)=0$, 特别地, $f_z(z)=0$, 即 $z=0$, 故 $\mu_i=0,\ \forall 1\leq i\leq k$.)

于是, 为证明 $\{[tv]\in\mathbb{T}^n\colon t\in\mathbb{R}\}\subseteq[V]$ 为稠子集, 只需在 V 关于 $V\cap\mathbb{Z}^n$ 的一个基本区域内考虑, 即证明 $\left\{\sum_{i=1}^k(tv_i'\pmod 1)w^{(i)}\in\mathbb{R}^n\colon t\in\mathbb{R}\right\}\subseteq\left\{\sum_{i=1}^kc_iw^{(i)}\in V\colon c_i\in\mathbb{R}且0\leq c_i<1\right\}$ 为稠子集. 再通过坐标变换, 只需证明 $\{t(v_1',\cdots,v_k')\pmod 1\in\mathbb{R}^k\colon t\in\mathbb{R}\}\subseteq[0,1)^k$ 为稠子集. 这是数论中经典的 Kronecker 定理, 它的证明可见 Hardy, Wright"An introduction to the Theory of Numbers". □

注: 上述 Kronecker-Weyl 定理可视为著名的 Ratner 轨道闭包定理的特例, 后者可叙述为: "设 G 为 Lie 群,

$\Gamma \subseteq G$ 为格, 则 G/Γ 中每个点在任意幂幺流下的轨道闭包都是齐性的." Ratner 轨道闭包定理的证明可见 Morris "Ratner's Theorems on Unipotent Flows".

3.1.2 商、核、余核的泛性质

本节将补充范畴学中泛性质的观点, 即通过某些特征性质在同构的意义下唯一确定范畴中的对象与态射. 回忆在定义线性空间的直积与直和时, 我们已经提及了相应的泛性质; 以下主要讨论线性映射的商、核、余核的泛性质.

命题 3.1.5 (商的泛性质) *设 V,U 为域 F 上的线性空间, $W \subseteq V$ 为线性子空间, 记 $\pi\colon V \longrightarrow V/W,\ \alpha \longmapsto \alpha+W$ 为商映射,*
$T \in L(V,U)$, 则 $W \subseteq \ker(T) \Longleftrightarrow \exists!\, \widetilde{T} \in L(V/W,U)$, $s.t.\ T = \widetilde{T}\circ\pi$.
证明: "⇐": 显然; "⇒": 设 $W \subseteq \ker(T)$, 则 $\widetilde{T}\colon V/W \longrightarrow U,\ \alpha+W \longmapsto T(\alpha)$ 是定义良好的线性映射, 且 $T = \widetilde{T}\circ\pi$. □

注: 以范畴学的语言, 我们应把二元组 $(V/W,\pi)$ 称为线性空间 V 的一个商对象, 满足上述商的泛性质的二元组在同构意义下唯一.

推论 3.1.6 *设 V,W_1,W_2 为域 F 上的线性空间, $T_1 \in L(V,W_1)$, $T_2 \in L(V,W_2)$, 则 $\ker(T_1) \subseteq \ker(T_2) \Longleftrightarrow T_2$ 可经过 T_1 分解, 即 $\exists S \in L(W_1,W_2)$, $s.t.\ T_2 = S\circ T_1$.*
证明: "⇐": 显然; "⇒": 设 $\ker(T_1) \subseteq \ker(T_2)$, 则 $S_0\colon \mathrm{Im}(T_1) \longrightarrow W_2,\ T_1(\alpha) \longmapsto T_2(\alpha)$ 是定义良好的线性映射; 通过取 $\mathrm{Im}(T_1)$ 在 W_1 中的直和补空间, 可将 S_0 延拓为 $S \in L(W_1,W_2)$, 则 $T_2 = S\circ T_1$. □

注: 由上述证明可知, T_2 经过 T_1 的分解方式并不唯一.

命题 3.1.7 (核的泛性质) *设 V,W 为域 F 上的线性空间, $T \in L(V,W)$, 记 $i\colon \ker(T) \to V$ 为含入映射, 若存在域 F 上的线性空间 K, 以及 $j \in L(K,V)$, 满足 $T\circ j = 0$, 则 j 可唯一地经过 i 分解, 即 $\exists!\, S \in L(K,\ker(T))$, $s.t.\ j = i\circ S$.*

注: 以范畴学的语言, 我们应把二元组 $(\ker(T),i)$ 称为线性映射 T 的核, 满足上述核的泛性质的二元组在同构意义下唯一.

命题 3.1.8 (余核的泛性质) *设 V,W 为域 F 上的线性空间, $T \in L(V,W)$, 记 $p\colon V \to \mathrm{coker}(T) \coloneqq V/\mathrm{Im}(T)$ 为商映射, 若存在域 F 上的线性空间 C, 以及 $q \in L(W,C)$, 满足 $q\circ T = 0$, 则 q 可唯一地经过 p 分解, 即 $\exists!\, S \in L(\mathrm{coker}(T),C)$,*
$s.t.\ q = S\circ p$.

注: 以范畴学的语言, 我们应把二元组 $(\mathrm{coker}(T),p)$ 称为线性映射 T 的余核, 满足上述余核的泛性质的二元组在同构意义下唯一.

3.1.3 商映射与同构定理

本节主要补充几个与商映射相关的同构定理, 与线性代数基本定理一样, 它们也可推广为群、环、模、代数的相应同构定理. 这些同构定理的编号不是统一的 (可见维基百科), 这里的编号遵从 Nathan Jacobson"Basic Algebra I" 的习惯.

命题 3.1.9 (第一同构定理) *设 V 为域 F 上的线性空间, $W \subseteq V$ 为线性子空间, 则存在集合族之间的保序双射*

$$\{V\text{中包含}W\text{的线性子空间}\} \longrightarrow \{V/W\text{的线性子空间}\},$$
$$U \longmapsto U/W$$

且当 $U_2 \supseteq U_1 \supseteq W$ 时, 商映射诱导了线性空间同构 $\dfrac{U_2/W}{U_1/W} \cong U_2/U_1$.

命题 3.1.10 (第二同构定理) *设 V 为域 F 上的线性空间, $W,U\subseteq V$ 为线性子空间, 则存在线性空间同构*

$$\begin{aligned} U/(U\cap W) &\xrightarrow{\cong} (U+W)/W\ .\\ \alpha+U\cap W &\longmapsto \alpha+W \end{aligned}$$

推论 3.1.11 *设 V 为域 F 上的线性空间, $T\in L(V,V)$, $W,U\subseteq V$ 为线性子空间, 则存在线性空间同构*

$$U/(T^{-1}(W)\cap U)\cong (T(U)+W)/W.$$

证明: 注意这是线性空间同构的复合

$$\begin{aligned} U/(T^{-1}(W)\cap U) &\xrightarrow{\cong} (U+T^{-1}(W))/T^{-1}(W) \xrightarrow{\cong} (T(U)+W)/W\ .\\ \alpha+T^{-1}(W)\cap U &\longmapsto \alpha+T^{-1}(W) \longmapsto T(\alpha)+W \end{aligned}$$

□

命题 3.1.12 (Zassenhaus 引理) *设 V 为域 F 上的线性空间, $W_1\subseteq W\subseteq V, U_1\subseteq U\subseteq V$ 为线性子空间, 则存在线性空间同构*

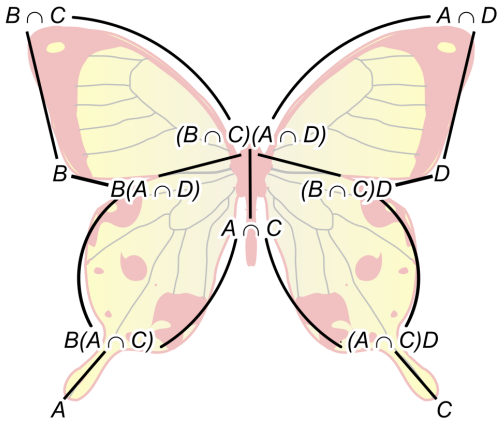
$$\frac{W_1+W\cap U}{W_1+W\cap U_1}\cong\frac{U_1+W\cap U}{U_1+W_1\cap U}.$$

证明: 记 $V_0\coloneqq (W_1\cap U)+(U_1\cap W)\subseteq W\cap U$, 则由第二同构定理与模律知,

$$\frac{W_1+W\cap U}{W_1+W\cap U_1}=\frac{(W_1+V_0)+W\cap U}{W_1+V_0}\cong\frac{W\cap U}{(W_1+V_0)\cap(W\cap U)}=\frac{W\cap U}{W_1\cap U+V_0}=\frac{W\cap U}{V_0};$$

同理 $\dfrac{U_1+W\cap U}{U_1+W_1\cap U}\cong\dfrac{W\cap U}{V_0}$, 故 $\dfrac{W_1+W\cap U}{W_1+W\cap U_1}\cong\dfrac{U_1+W\cap U}{U_1+W_1\cap U}$. □

注: Zassenhaus 引理又被称为蝴蝶引理, 这是因为当绘制涉及线性子空间的 Hasse 图时会出现一只蝴蝶 (如图记 $A=W$, $B=W_1, C=U, D=U_1$):



推论 3.1.13 (Schreier 加细定理) *设 V 为域 F 上的线性空间, 记*

$$\widetilde{\mathscr{F}}(V)\coloneqq\{F=(W_0,W_1,\cdots,W_n)\colon n\in\mathbb{N},\ \{0\}=W_0\subseteq W_1\subseteq\cdots\subseteq W_n=V\text{为线性子空间链}\}.$$

对于 $F=(W_0,\cdots,W_n)$, $F'=(U_0,\cdots,U_m)\in\widetilde{\mathscr{F}}(V)$, 记 $F\preccurlyeq F'\iff$ 存在单射 $f\colon\{0,1,\cdots,n\}\to\{0,1,\cdots,m\}$, $s.t.\ W_i=U_{f(i)}$, $\forall i\in\{0,1,\cdots,n\}$; 记 $F\sim F'\iff m=n$ 且 $\exists\sigma\in S_n$, $s.t.\ W_i/W_{i-1}\cong U_{\sigma(i)}/U_{\sigma(i-1)}$ $(i=1,\cdots,n)$. 于是 $\forall F,F'\in\widetilde{\mathscr{F}}(V)$, $\exists F_1,F_1'\in\widetilde{\mathscr{F}}(V)$, $s.t.\ F\preccurlyeq F_1\sim F_1'\succcurlyeq F'$.

证明: 记 $F=(W_0,W_1,\cdots,W_n)$, $F'=(U_0,U_1,\cdots,U_m)$. 令

$$W_{i,j}\coloneqq W_{i-1}+(W_i\cap U_j)\ (i=1,\cdots,n;j=0,\cdots,m),$$

$$U_{j,i}\coloneqq U_{j-1}+(U_j\cap W_i)\ (j=1,\cdots,m;i=0,\cdots,n),$$

则 $W_{i-1,m}=W_{i-1}=W_{i,0}$ $(i=1,\cdots,n)$, $U_{j-1,n}=U_{j-1}=U_{j,0}$ $(j=1,\cdots,m)$. 由 Zassenhaus 引理知, $\forall 1\le i\le n$, $\forall 1\le j\le m$, $W_{i,j}/W_{i,j-1}\cong U_{j,i}/U_{j,i-1}$. 现令

$$F_1=(W_{1,0},W_{1,1},\cdots,W_{1,m},W_{2,1},\cdots,W_{2,m},\cdots,W_{n,1},\cdots,W_{n,m}),$$

$$F_1'=(U_{1,0},U_{1,1},\cdots,U_{1,n},U_{2,1},\cdots,U_{2,n},\cdots,U_{m,1},\cdots,U_{m,n}),$$

则 $F_1,F_1'\in\widetilde{\mathscr{F}}(V)$, 且 $F\preccurlyeq F_1\sim F_1'\succcurlyeq F'$. □

参考文献与补注 3.1

(1) 关于范畴学中伴随函子、正合列可裂性与泛性质的部分, 可以参考 C. A. Weibel "An Introduction to Homological Algebra".

(2) 关于 Lie 群中格的刚性与算术性的部分, 可以参考 R. J. Zimmer"Ergodic Theory and Semisimple Groups".

(3) 关于 Atiyah-Singer 指标定理的部分, 可以参考梅加强 "流形与几何初步".

(4) 关于若干群同构定理及其应用的部分, 可以参考孙智伟 "近世代数".

§3.2 线性函数与双线性函数

3.2.1 线性函数与对偶空间

设 V 为域 F 上的线性空间, B 为 V 的 F-基, 则坐标映射 $\Gamma_B\colon V\xrightarrow{\cong}F^{(B)}\coloneqq\coprod\limits_{\alpha\in B}F$ 诱导了线性同构

$$V^*\coloneqq\mathrm{Hom}_{\mathbf{F-Mod}}(V,F)\xrightarrow{(\Gamma_B^{-1})^t}\mathrm{Hom}_{\mathbf{F-Mod}}(F^{(B)},F)=F^B\coloneqq\prod_{\alpha\in B}F,$$

于是为决定 $f\in V^*$ 只需决定 f 在集合 B 上的取值. 例如, 任取 $\alpha\in B$, 可令 $f_\alpha\colon\alpha'\mapsto\begin{cases}1, & \alpha'=\alpha\\ 0, & \alpha'\neq\alpha\end{cases}$ 并做线性延拓, 则可直接验证 $\{f_\alpha\}_{\alpha\in B}\subseteq V^*$ 为线性无关集, 称为 B 的**对偶集** (dual set). 注意

$$\mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})=V^*\Longleftrightarrow\dim_F(V)<+\infty.$$

(这是因为: 一方面, 若 $\dim_F(V)=|B|=n<+\infty$, 则 $\dim_F(V^*)=n=|\{f_\alpha\}_{\alpha\in B}|$, 故 $\mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})=V^*$. 另一方面, 若 $\dim_F(V)=|B|$ 非有限, 则考虑 $f=\sum\limits_{\alpha\in B}f_\alpha\colon$ $\begin{array}{ccc}V&\longrightarrow&F\\ (\text{有限和})\sum\limits_{\alpha\in B}c_\alpha\alpha&\longmapsto&(\text{有限和})\sum\limits_{\alpha\in B}c_\alpha\end{array}$, 则 f 定义良好且 $f\in V^*$, 但不可能写成有限个 $f_\alpha\,(\alpha\in B)$ 的线性组合.) 因此映射 $\alpha\in B\mapsto f_\alpha\in V^*$ 的线性延拓给出了单射 $\Phi_B\colon V\to V^*$, 且 Φ_B 为满射 $\Longleftrightarrow\dim_F(V)<+\infty$.

注意上述映射 Φ_B 的构造严重地依赖于基 B 的选取. 如果事先不给定基 B, 而是直接考虑将任意元 $\alpha\in V$ 映为它的对偶元 $f_\alpha\in V^*$, 则将面临以下困难: 例如对于 $\alpha\in V\setminus\{0\}$, 我们希望定义 $f_\alpha\in V^*$, 满足 $f_\alpha(\alpha)=1$, 且 f_α 在 $\mathrm{Span}_F(\{\alpha\})$ 的某个直和补空间上取值为 0, 但这样的对应 $\alpha\mapsto f_\alpha$ 很难保证是线性的. 于是我们不得不先借助基 B 来定义对偶集 $\{f_\alpha\}_{\alpha\in B}\subseteq V^*$, 再做线性延拓.

一个非平凡的观察是, 即使事先不给定基 B, 我们也可以直接对于任意元 $\alpha\in V$ 定义它的重对偶元 $L_\alpha\in V^{**}$, 即 $L_\alpha\colon$ $\begin{array}{ccc}V^*&\longrightarrow&F\\ f&\longmapsto&f(\alpha)\end{array}$, 且这个对应是线性映射 $\tau\colon$ $\begin{array}{ccc}V&\longrightarrow&V^{**}\\ \alpha&\longmapsto&L_\alpha\end{array}$. 断言: τ 为单射; τ 为满射 $\Longleftrightarrow\dim_F(V)<+\infty$.

(这是因为, 若 $\alpha\in V\setminus\{0\}$, 则可将 $\{\alpha\}$ 扩充为 V 的基 B, 取 B 的对偶集 $\{f_\alpha\}_{\alpha\in B}\subseteq V^*$, 则 $L_\alpha(f_\alpha)=f_\alpha(\alpha)=1\neq0$, 故 $L_\alpha\neq0$, 因此 τ 为单射. 现设 $\dim_F(V)=n<+\infty$, 则 $\dim_F(V^*)=n$, 故 $\dim_F(V^{**})=n$, 此时 τ 也为满射.

另一方面, 若 $\dim_F(V)=|B|$ 非有限, 则 $\dim_F(V^*)=\dim_F(F^B)=|F^B|>|B|=\dim_F(V)$ 也非有限, 故同理 $\dim_F(V^{**})>\dim_F(V^*)>\dim_F(V)$, 此时任意线性映射 $V\to V^{**}$ 都不可能是满射.)

现在我们给出当 $\dim_F(V)<+\infty$ 时, 线性同构 $\tau\colon V\xrightarrow{\cong}V^{**}$ 的一个应用: 任取 V^* 中的线性无关集 $\{f_i\}_{i=1}^s$, 均存在 $\{\alpha_i\}_{i=1}^s\subseteq V$, 满足 $f_j(\alpha_i)=\delta_{ij}$, $\forall1\leq i,j\leq s$. (这是因为, 可将 V^* 中的线性无关集 $\{f_i\}_{i=1}^s$ 扩充为 V^* 的基 $\{f_i\}_{i=1}^n$, 再取它的对偶集 $\{L_i\}_{i=1}^n\subseteq V^{**}$. 由 $\tau\colon V\xrightarrow{\cong}V^{**}$ 为线性同构知, $\exists\{\alpha_i\}_{i=1}^n\subseteq V$, $s.t.\ L_{\alpha_i}=L_i$, $\forall1\leq i\leq n$, 则 $\forall1\leq i,j\leq n$, $f_j(\alpha_i)=L_{\alpha_i}(f_j)=L_i(f_j)=\delta_{ij}$.) 此时可直接验证 $\{\alpha_i\}_{i=1}^s\subseteq V$ 为线性无关集, 也称为 $\{f_i\}_{i=1}^s\subseteq V^*$ 的**对偶集**.

我们将上述对偶集的讨论中关于线性无关性的验证整理为如下引理:

引理 3.2.1 *设 V 为域 F 上的线性空间, $\{\alpha_i\}_{i\in I}\subseteq V$, $\{f_i\}_{i\in I}\subseteq V^*$, 满足 $f_j(\alpha_i)=\delta_{ij}$, $\forall i,j\in I$, 则 $\{\alpha_i\}_{i\in I}\subseteq V$,*

$\{f_i\}_{i\in I}\subseteq V^*$ 均为线性无关集.

证明: 任取 $\{\alpha_i\}_{i\in I}$ 的非空有限子集 $\{\alpha_{i_l}\}_{l=1}^k$, 考虑其线性组合 $\sum\limits_{l=1}^k c_l\alpha_{i_l}=0\,(c_l\in F)$, 则 $\forall 1\le l\le k$, $c_l=f_{i_l}(c_1\alpha_{i_1}+\cdots+c_k\alpha_{i_k})=f_{i_l}(0)=0$, 故 $\{\alpha_{i_l}\}_{l=1}^k$ 线性无关, 因此 $\{\alpha_i\}_{i\in I}$ 线性无关. 同理知 $\{f_i\}_{i\in I}$ 线性无关.

□

以下我们主要以对偶的观点讨论线性函数与线性子空间的关系.

设 V 为域 F 上的线性空间, $f\in V^*$, 由线性代数基本定理知, 存在线性同构 $\widetilde{f}\colon\ V/\ker(f)\longrightarrow \mathrm{Im}(f)$. 注

$$\alpha+\ker(f)\longmapsto f(\alpha)$$

意 $\mathrm{Im}(f)\subseteq F$ 为线性子空间, 故当 $f\neq 0$ 时, $\mathrm{Im}(f)=F$, 则 $\ker(f)\subseteq V$ 为余一维子空间, 也称为超平面. 回忆线性覆盖的相关命题:

命题 3.2.2 *设 V 为域 F 上的线性空间,*

(1) *若 F 为无限域, 则 $\forall\{f_i\}_{i=1}^s\subseteq V^*\setminus\{0\}$, $\exists\alpha\in V$, s.t. $f_i(\alpha)\neq 0$, $\forall 1\le i\le s$.*

(2) *若 $|I|<|F|$, 且 $\dim_F(V)<+\infty$, 则 $\forall\{f_i\}_{i\in I}\subseteq V^*\setminus\{0\}$, $\exists\alpha\in V$, s.t. $f_i(\alpha)\neq 0$, $\forall i\in I$.*

另一方面, 任取 V 的余一维子空间 H, 通过取 H 在 V 中的直和补空间, 可构造 $f\in V^*$ 满足 $\ker(f)=H$. 显然这样的构造并不唯一. 一般地, 设 $W\subseteq V$ 为线性子空间, 则 $W^0\coloneqq\{f\in V^*\colon\ \ker(f)\supseteq W\}$ 为 V^* 的线性子空间, 称为 W 的**零化子空间** (annihilator); 对偶地, 设 $M\subseteq V^*$ 为线性子空间, 则 $M^\diamondsuit\coloneqq\bigcap\limits_{f\in M}\ker(f)$ 为 V 的线性子空间, 称为 M 的公共核空间.

命题 3.2.3 *设 V 为域 F 上的线性空间,*

(1) $\{0\}^0=V^*$, $V^0=\{0\}$; $\{0\}^\diamondsuit=V$, $(V^*)^\diamondsuit=\{0\}$.

(2) *设 $W\subseteq V$ 为线性子空间, 则 $(W^0)^\diamondsuit=W$; 设 $M\subseteq V^*$ 为线性子空间, 则 $(M^\diamondsuit)^0\supseteq M$.*

(3) *存在两个反序映射 $(\cdot)^0\colon\{V\text{的线性子空间}\}\rightleftarrows\{V^*\text{的线性子空间}\}\colon(\cdot)^\diamondsuit$, 其中 $(\cdot)^0$ 为单射, $(\cdot)^\diamondsuit$ 为满射.*

$$\begin{array}{ccc} W & \longmapsto & W^0 \\ M^\diamondsuit & \longleftarrow\!\shortmid & M \end{array}$$

(4) $\dim_F(V)<+\infty\iff$ (2) 中 "$=$" 总可取到 $\iff$ (3) 中映射为互逆的.

证明: (1) 显然;

(2) 设 $W\subseteq V$ 为线性子空间, 则显然 $(W^0)^\diamondsuit\supseteq W$; 另一方面, 假设 $\exists\alpha\in(W^0)^\diamondsuit\setminus W$, 则通过取 W 在 V 中包含 α 的直和补空间, 可构造 $f\in W^0$ 满足 $f(\alpha)\neq 0$, 这与 $\alpha\in(W^0)^\diamondsuit$ 矛盾! 设 $M\subseteq V^*$ 为线性子空间, 则显然 $(M^\diamondsuit)^0\supseteq M$.

(3) $(\cdot)^0$ 与 $(\cdot)^\diamondsuit$ 的反序性显然; 由 $((\cdot)^0)^\diamondsuit=\mathrm{id}$ 知, $(\cdot)^0$ 为单射, $(\cdot)^\diamondsuit$ 为满射.

(4) 设 $\dim_F(V)=n<+\infty$, 则 $\dim_F(V^*)=n<+\infty$. 设 $M\subseteq V^*$ 为线性子空间, 则可取 M 的基 $\{f_1,\cdots,f_m\}$, 并延拓为 V^* 的基 $\{f_1,\cdots,f_m,f_{m+1},\cdots,f_n\}$. 取它的对偶集 $\{\alpha_i\}_{i=1}^n\subseteq V$, 则 $\{\alpha_i\}_{i=1}^n$ 为 V 的基, 且其中 $\{\alpha_i\}_{i=m+1}^n$ 为 $M^\diamondsuit$ 的基, 故 $(M^\diamondsuit)^0$ 的基为 $\{f_i\}_{i=1}^m$, 因此 $(M^\diamondsuit)^0=M$. 特别地, 由 $((\cdot)^0)^\diamondsuit=\mathrm{id}$ 且 $((\cdot)^\diamondsuit)^0=\mathrm{id}$ 知, (3) 中映射互逆.

现设 (3) 中映射为互逆的, 则 (2) 中任取 $M\subseteq V^*$ 为线性子空间, $(M^\diamondsuit)^0=M$. 现设 B 为 V 的基, 取它的对偶集 $\{f_\alpha\}_{\alpha\in B}\subseteq V^*$, 令 $M=\mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})$, 则 $M^\diamondsuit=\{0\}$, 故 $(M^\diamondsuit)^0=V^*$. 因此由条件知 $V^*=\mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})$, 即 Φ_B 为满射, 故 $\dim_F(V)<+\infty$. □

推论 3.2.4 *设 V 为域 F 上的线性空间, $W\subseteq V$ 为线性子空间, $M\subseteq V^*$ 为线性子空间, 则:*

(1) *存在线性同构 $W^0\xrightarrow{\cong}(V/W)^*$. 特别地, 当 $\dim_F(V)<+\infty$ 时, 存在线性同构 $M\cong(V/M^\diamondsuit)^*$.*

$$f\longmapsto(\widetilde{f}\colon\alpha+W\mapsto f(\alpha))$$

(2) *存在线性同构 $V^*/W^0\xrightarrow{\cong}W^*$. 特别地, 当 $\dim_F(V)<+\infty$ 时, 存在线性同构 $V^*/M\cong(M^\diamondsuit)^*$.*

$$f+W^0\longmapsto f\big|_W$$

证明: (1) 是商的泛性质; (2) 是线性代数基本定理. □

运用上述对偶 (或零化) 的观点可轻松处理一些技术化的问题, 具体原理如下:

命题 3.2.5 *设 V 为域 F 上的线性空间, $\{f_i\}_{i=1}^s \subseteq V^*$, 则 $\mathrm{Span}_F(\{f_i\}_{i=1}^s) = \left(\bigcap\limits_{i=1}^s \ker(f_i)\right)^0$. 特别地,*
$\{f_i\}_{i=1}^s$ 线性无关 $\Longleftrightarrow \mathrm{codim}_V(\bigcap\limits_{i=1}^s \ker(f_i)) = s$; $\{f_i\}_{i=1}^s$ 线性生成 $V^* \Longleftrightarrow \bigcap\limits_{i=1}^s \ker(f_i) = \{0\}$.

证明: 显然 $\mathrm{Span}_F(\{f_i\}_{i=1}^s) \subseteq \left(\bigcap\limits_{i=1}^s \ker(f_i)\right)^0$. 另一方面, 考虑线性映射 $T\colon \begin{array}{ccc} V & \longrightarrow & F^{1\times s} \\ \alpha & \longmapsto & (f_1(\alpha),\cdots,f_s(\alpha)) \end{array}$, 则 $\ker(T) = \bigcap\limits_{i=1}^s \ker(f_i)$. 现任取 $f \in \left(\bigcap\limits_{i=1}^s \ker(f_i)\right)^0$, 则 f 可经过 T 分解, 即 $\exists g \in (F^{1\times s})^*$, $s.t.$ $f = g\circ T$. 记 $g\colon \begin{array}{ccc} F^{1\times s} & \longrightarrow & F \\ (\alpha_1,\cdots,\alpha_s) & \longmapsto & \sum\limits_{i=1}^s c_i\alpha_i \end{array}$, 则 $f = \sum\limits_{i=1}^s c_i f_i \in \mathrm{Span}_F(\{f_i\}_{i=1}^s)$. 最后由上述推论即知特别结论. □

注:

(1) 上例的结论对于无限个线性函数未必成立. 例如设 $\dim_F(V) = |B|$ 为无穷势, $\{f_\alpha\}_{\alpha\in B} \subseteq V^*$ 为 B 的对偶集, 则 $\sum\limits_{\alpha\in B} f_\alpha \in \left(\bigcap\limits_{\alpha\in B} \ker(f_\alpha)\right)^0 \setminus \mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})$.

(2) 上例提供了在一般的线性空间中对于有限个线性无关的线性函数取对偶集的方式: 事实上, 设 $\{f_i\}_{i=1}^s \subseteq V^*$ 线性无关, 则 $\forall 1\le i\le s$, $f_i \notin \mathrm{Span}_F(\{f_j\}_{j\neq i}) = \left(\bigcap\limits_{j\neq i} \ker(f_j)\right)^0$, 即 $\exists \alpha_i \in \bigcap\limits_{j\neq i} \ker(f_j) \setminus \ker(f_i)$. 通过标准化可不妨设 $f_j(\alpha_i) = \delta_{ij}$, $\forall 1\le i,j\le s$.

例 3.2.1 (矩阵行列空间的对偶关系) 设 F 为一个域, $A \in F^{m\times n}$, 考虑齐次线性方程组 $AX = 0$, 可将矩阵 A 的每个行向量 $\alpha_i = (A_{i1},\cdots,A_{in}) \in F^{1\times n}$ $(1\le i\le m)$ 视为 $(F^{n\times 1})^*$ 中元 $f_i\colon \begin{array}{ccc} F^{n\times 1} & \longrightarrow & F \\ X & \longmapsto & \alpha X \end{array}$, 这里将 $F^{1\times 1}$ 与 F 视为等同. 于是由定义知 $\mathrm{row}(A)^\diamondsuit = \ker(A)$, 此时应用上述命题可再次得到线性方程组理论的相应结果. 另外, 由上述命题、推论与线性代数基本定理知, $\mathrm{row}(A) = (\mathrm{row}(A)^\diamondsuit)^0 = \ker(A)^0 \cong (F^{n\times 1}/\ker(A))^* \cong \mathrm{column}(A)^*$, 此即矩阵行列空间的对偶关系.

最后我们再补充一些对偶 (零化) 与子空间交与和的关系.

命题 3.2.6 *设 V 为域 F 上的线性空间,*

(1) *设 $W_1, W_2 \subseteq V$ 为线性子空间, 则 $(W_1+W_2)^0 = W_1^0 \cap W_2^0$; $(W_1\cap W_2)^0 = W_1^0 + W_2^0$.*

(2) *设 $M_1, M_2 \subseteq V^*$ 为线性子空间, 则 $(M_1+M_2)^\diamondsuit = M_1^\diamondsuit \cap M_2^\diamondsuit$; $(M_1\cap M_2)^\diamondsuit \supseteq M_1^\diamondsuit + M_2^\diamondsuit$; 后者 "=" 总可取到 $\Longleftrightarrow \dim_F(V) < +\infty$.*

证明: (1) 只证明 $(W_1\cap W_2)^0 \subseteq W_1^0 + W_2^0$, 其余均显然. 任取 $f \in (W_1\cap W_2)^0$, 记 $M_1 \coloneqq \{h \in W_1^*\colon \ker(h) \supseteq W_1\cap W_2\}$, 则 $f_1 \coloneqq f\big|_{W_1} \in M_1$. 再记 $M_2 \coloneqq \{h \in (W_1+W_2)^*\colon \ker(h) \supseteq W_2\}$, 由上述推论 (1) 和第二同构定理知, 存在线性空间的同构

$$M_1 \cong (W_1/W_1\cap W_2)^* \cong ((W_1+W_2)/W_2)^* \cong M_2,$$

在此同构下记 $f_1 \in M_1$ 的像为 $g_2 \in M_2$, 则 $g_2(\alpha_1+\alpha_2) = f_1(\alpha_1)$, $\forall \alpha_1 \in W_1, \alpha_2 \in W_2$. 再将 $g_2 \in (W_1+W_2)^*$ 延拓为 $g \in V^*$, 则 $g \in W_2^0$, 且 $f - g \in W_1^0$. 因此 $(W_1\cap W_2)^0 \subseteq W_1^0 + W_2^0$.
(2) 只证明 $(M_1\cap M_2)^\diamondsuit = M_1^\diamondsuit + M_2^\diamondsuit$ 总成立 $\Longleftrightarrow \dim_F(V) < +\infty$, 其余均显然. 当 $\dim_F(V) < +\infty$ 时, 由上述命题知可记 $M_1 = W_1^0$, $M_2 = W_2^0$, 则 $(M_1\cap M_2)^\diamondsuit = (W_1^0\cap W_2^0)^\diamondsuit = ((W_1+W_2)^0)^\diamondsuit = W_1+W_2 = M_1^\diamondsuit + M_2^\diamondsuit$; 当 $\dim_F(V) = |B|$ 非有限时, 取 B 为 V 的基, 以及它的对偶集 $\{f_\alpha\}_{\alpha\in B} \subseteq V^*$, 令 $M_1 = \mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})$, $M_2 = \mathrm{Span}_F(\{\sum\limits_{\alpha\in B} f_\alpha\})$, 则 $(M_1\cap M_2)^\diamondsuit = \{0\}^\diamondsuit = V \supsetneqq \{0\} + \ker(\sum\limits_{\alpha\in B} f_\alpha) = M_1^\diamondsuit + M_2^\diamondsuit$. □

3.2.2 转置映射

记 **F-Mod** 为域 F 上的线性空间范畴. 上节取对偶的操作给出了 **F-Mod** 中对象的对应 $V \mapsto V^*$, 并进一步诱导了 **F-Mod** 中态射集的对应 $L(U,V) \longrightarrow L(V^*,U^*)$. 注意取转置的操作保持恒同态射, 但逆转了

$$T \longmapsto (T^t\colon g \mapsto g\circ T)$$

态射的复合, 故为一个**反变函子** (contravariant functor). 当然, 取重对偶的操作将会是一个**协变函子** (covariant functor). 取重对偶的特别之处在于, 存在从恒同函子到取重对偶函子的 **自然变换** (natural transformation), 它在每个线性空间对象上由典型映射 $\tau_V\colon V \longrightarrow V^{**}$ 给出, 即存在交换图 $\begin{array}{ccc} U & \xrightarrow{\tau_U} & U^{**} \\ \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle (T^t)^t} \\ V & \xrightarrow{\tau_V} & V^{**} \end{array}$.

$$\alpha \longmapsto L_\alpha$$

转置映射最大的特点在于它的核与像的关系, 这有别于由内积确定的伴随映射的相应性质.

命题 3.2.7 *设 U,V 为域 F 上的线性空间, $T \in L(U,V)$, 则:*

(1) $\ker(T^t) = \mathrm{Im}(T)^0$; $\ker(T^t)^\diamondsuit = \mathrm{Im}(T)$*; 特别地, T^t 为单射 $\Longleftrightarrow$ T 为满射.*

(2) $\mathrm{Im}(T^t) = \ker(T)^0$; $\mathrm{Im}(T^t)^\diamondsuit = \ker(T)$*; 特别地, T^t 为满射 $\Longleftrightarrow$ T 为单射.*

证明: 只证明 $\mathrm{Im}(T^t) \supseteq \ker(T)^0$, 其余均显然. 任取 $f \in \ker(T)^0$, 则 f 可经过 T 分解, 即 $\exists g \in W^*$, *s.t.* $f = g\circ T = T^t(g) \in \mathrm{Im}(T^t)$, 故 $\ker(T)^0 \subseteq \mathrm{Im}(T^t)$. □

转置映射的另一不平凡性质是, 它与映射的原像可通过零化的观点联系起来.

命题 3.2.8 *设 U,V 为域 F 上的线性空间, $T \in L(U,V)$,*

(1) *设 $W \subseteq V$ 为线性子空间, 则 $T^t(W^0) = (T^{-1}(W))^0$;*

(2) *设 $M \subseteq V^*$ 为线性子空间, 则 $T^t(M) \subseteq (T^{-1}(M^\diamondsuit))^0$, 且 "$=$" 总可取到 $\Longleftrightarrow \dim_F(\mathrm{Im}(T)) < +\infty$.*

证明: 只证明 (1)(2) 中的 "$\supseteq$" 方向, 因为 "$\subseteq$" 方向均显然.

(1) 任取 $f \in (T^{-1}(W))^0$. 特别地, $f \in \ker(T)^0 = \mathrm{Im}(T^t)$, 即 $\exists g_0 \in V^*$, *s.t.* $f = g_0\circ T$. 以下调整 $g_0 \in V^*$ 为 $g \in W^0$: 先考虑 V 的子空间 $\mathrm{Im}(T)+W$, 则存在子空间 $\mathrm{Im}_1 \subseteq \mathrm{Im}(T)$, 使得 $\mathrm{Im}_1 \oplus W = \mathrm{Im}(T)+W$. 令 $g \in V^*$ 为 $g\big|_{\mathrm{Im}_1} \coloneqq g_0\big|_{\mathrm{Im}_1}$, $g\big|_W \coloneqq 0$, 以及在 $\mathrm{Im}(T)+W$ 的直和补空间上任取线性函数, 则 $g \in W^0$. 最后验证 $f = g\circ T$: 事实上, 由 $\mathrm{Im}_1 \oplus W \supseteq \mathrm{Im}(T)$ 知, $U = T^{-1}(\mathrm{Im}_1)+T^{-1}(W)$. 任取 $\alpha \in U$, 记 $\alpha = \alpha_1+\alpha_2$, 其中 $\alpha_1 \in T^{-1}(\mathrm{Im}_1)$, $\alpha_2 \in T^{-1}(W)$, 则 $f(\alpha) = f(\alpha_1)+f(\alpha_2) = g_0(T(\alpha_1))+0 = g(T(\alpha_1))+g(T(\alpha_2)) = (g\circ T)(\alpha)$, 故 $f = g\circ T$.

(2) 当 $\dim_F(\mathrm{Im}(T)) < +\infty$ 时, 记 $T_1\colon U \longrightarrow \mathrm{Im}(T)$, $i\colon \mathrm{Im}(T) \hookrightarrow V$ 为含入映射, 则 $T = i\circ T_1$. 任取 $M \subseteq V^*$

$$\alpha \longmapsto T(\alpha)$$

为线性子空间, 则 $i^t(M) \subseteq \mathrm{Im}(T)^*$, 故 $i^t(M) = (i^t(M)^\diamondsuit)^0$. 由定义可直接验证 $i^t(M)^\diamondsuit = i^{-1}(M^\diamondsuit)$, 因此由 (1) 知

$$T^t(M) = T_1^t(i^t(M)) = T_1^t((i^t(M)^\diamondsuit)^0) = (T_1^{-1}((i^t(M)^\diamondsuit))^0 = (T_1^{-1}(i^{-1}(M^\diamondsuit))^0 = (T^{-1}(M^\diamondsuit))^0.$$

当 $\dim_F(\mathrm{Im}(T))$ 非有限时, 设 B_1 为 $\mathrm{Im}(T)$ 的基, 并延拓为 V 的基 B, 再取它的对偶集 $\{f_\alpha\}_{\alpha\in B} \subseteq V^*$. 令 $M = \mathrm{Span}_F(\{f_\alpha\}_{\alpha\in B})$, 则 $M^\diamondsuit = \{0\}$, 此时 $T^t(M) = \mathrm{Span}_F(\{f_\alpha\circ T\}_{\alpha\in B})$; $(T^{-1}(M^\diamondsuit))^0 = \ker(T)^0 = \mathrm{Im}(T^t)$. 断言: $\mathrm{Span}_F(\{f_\alpha\circ T\}_{\alpha\in B}) \subsetneqq \mathrm{Im}(T^t)$. 这是因为, 取 $T^t\left(\sum\limits_{\alpha\in B} f_\alpha\right) \in \mathrm{Im}(T^t)$. 假设 $\exists c_1,\cdots,c_k \in F$, $\alpha_1,\cdots,\alpha_k \in B$, *s.t.* $T^t\left(\sum\limits_{\alpha\in B} f_\alpha\right) = \sum\limits_{i=1}^k c_i f_{\alpha_i}\circ T$, 即 $\left(\sum\limits_{\alpha\in B} f_\alpha - \sum\limits_{i=1}^k c_i f_{\alpha_i}\right)\big|_{\mathrm{Im}(T)} = 0$, 则取 $\alpha \in B_1 \setminus \{\alpha_i\}_{i=1}^k$ 代入知 $1 = 0$, 矛盾! □

3.2.3 双线性函数

本节主要补充双线性函数的一些性质, 以此给出线性函数与对偶空间的另一种实现方式.

设 V,W 为域 F 上的两个线性空间, $\varphi\colon V\times W \to F$ 为映射, 满足 φ 关于每个分量都是线性函数, 则 φ 称为一个**双线性函数** (bilinear function). 它可视为一族由 V 作指标集的 W 上的线性函数 (或一族由 W 作指标集的 V 上的线性函数), 且该线性函数族关于指标集也是线性的. 换句话说, 一个双线性函数 $\varphi\colon V\times W \to F$ 等价于一个线性映射 $\varphi_L\colon V \longrightarrow W^*$ (或 $\varphi_R\colon W \longrightarrow V^*$). 若 φ_L 为单射, 则称 φ 是 **左非退化的** (left

$$\alpha \longmapsto \varphi(\alpha,\cdot) \qquad\qquad \beta \longmapsto \varphi(\cdot,\beta)$$

non-degenerate); 若 φ_R 为单射, 则称 φ 是 **右非退化的** (right non-degenerate). 既左非退化又右非退化的双线性函数称为非退化的.

一个最基本的非退化双线性函数为 $\varphi\colon \begin{array}{ccc} F^{(I)}\times F^{(I)} & \longrightarrow & F \\ ((\alpha_i)_{i\in I},(\beta_i)_{i\in I}) & \longmapsto & \sum\limits_{i\in I}\alpha_i\beta_i \end{array}$; 例如取 $F=\mathbb{R}$, $I=\{1,\cdots,n\}$, 则 φ 为 $\mathbb{R}^n$ 上通常的实内积. 另一个基本的非退化双线性函数为 $\mathrm{tr}\colon \begin{array}{ccc} F^{n\times n}\times F^{n\times n} & \longrightarrow & F \\ (A,B) & \longmapsto & \mathrm{tr}(A\cdot B) \end{array}$. 值得注意的是它们都是对称的.

一般来说, 双线性函数的左右非退化性并不等价. 一个平凡的例子是取 $\dim_F(V)>1$, $W=F$, $f\in V^*\setminus\{0\}$, 则 $\varphi\colon \begin{array}{ccc} V\times F & \longrightarrow & F \\ (\alpha,c) & \longmapsto & cf(\alpha) \end{array}$ 是一个双线性函数, 它左退化但右非退化. 另一个不太平凡的例子是取 $V=F^{(\mathbb{N})}$, 则 $\varphi\colon \begin{array}{ccc} V\times V & \longrightarrow & F \\ ((a_i)_{i\in\mathbb{N}},(b_i)_{i\in\mathbb{N}}) & \longmapsto & \sum\limits_{i\in\mathbb{N}}a_ib_{i+1} \end{array}$ 是一个双线性函数, 它左非退化但右退化. 以下说明: 当 $\dim_F(V)=\dim_F(W)<+\infty$ 时, 双线性函数 $\varphi\colon V\times W\to F$ 的左右非退化性等价.

引理 3.2.9 *设 V,W 为域 F 上的线性空间, 且 $\dim_F(V)=\dim_F(W)<+\infty$, $\varphi\colon V\times W\to F$ 为双线性函数, 则以下条件等价:*

(1) *φ 是左非退化的;*

(2) *$\varphi_L\colon \begin{array}{ccc} V & \longrightarrow & W^* \\ \alpha & \longmapsto & \varphi(\alpha,\cdot) \end{array}$ 为线性同构;*

(3) *任取 W 的基 $\{\beta_i\}_{i=1}^n$, 存在 V 的基 $\{\alpha_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(4) *φ 是右非退化的;*

(5) *$\varphi_R\colon \begin{array}{ccc} W & \longrightarrow & V^* \\ \beta & \longmapsto & \varphi(\cdot,\beta) \end{array}$ 为线性同构;*

(6) *任取 V 的基 $\{\alpha_i\}_{i=1}^n$, 存在 W 的基 $\{\beta_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(7) *存在 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 W 的基 $\{\beta_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(8) *任取 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 W 的基 $\{\beta_i\}_{i=1}^n$, 则 $(\varphi(\alpha_i,\beta_j))_{n\times n}\in\mathrm{GL}(n,F)$.*

证明: 我们先证 (1)⇒(2)⇒(3)⇒(7)⇒(8)⇒(1); 同理可证 (4)⇒(5)⇒(6)⇒(7)⇒(8)⇒(4).

(1)⇒(2): 由于 $\dim_F(V)=\dim_F(W)<+\infty$, 则 $\dim_F(V)=\dim_F(W^*)<+\infty$, 故线性映射 φ_L 为单射当且仅当它为满射, 也为线性同构.

(2)⇒(3): 任取 W 的基 $B=\{\beta_i\}_{i=1}^n$, 由 $\dim_F(W)<+\infty$ 知, 对偶集的构造给出了线性同构 $\Phi_B\colon W\to W^*$; 复合知 $\Phi_B^{-1}\circ\varphi_L\colon V\to W$ 也为线性同构. 取 $\{\beta_i\}_{i=1}^n\subseteq W$ 在此同构下的原像为 $\{\alpha_i\}_{i=1}^n\subseteq V$, 则 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$.

(3)⇒(7): 显然;

(7)⇒(8): 任取 V 的基 $\{\alpha_i'\}_{i=1}^n$ 与 W 的基 $\{\beta_i'\}_{i=1}^n$ 记 $(\alpha_1',\cdots,\alpha_n')=(\alpha_1,\cdots,\alpha_n)P$, $(\beta_1',\cdots,\beta_n')=(\beta_1,\cdots,\beta_n)Q$, 其中 $P,Q\in\mathrm{GL}(n,F)$, 则 $(\varphi(\alpha_i',\beta_j'))_{n\times n}=P^t\cdot(\varphi(\alpha_i,\beta_j))_{n\times n}\cdot Q=P^tQ\in\mathrm{GL}(n,F)$.

(8)⇒(1): 设 $\alpha\in V$ 满足 $\varphi(\alpha,\cdot)=0$, 即 $\varphi(\alpha,\beta)=0$, $\forall\beta\in V$. 记 $\alpha=\sum\limits_{i=1}^n c_i\alpha_i$, 则 $(c_1,\cdots,c_n)\in F^{1\times n}$ 是齐次线性方程组 $X\cdot(\varphi(\alpha_i,\beta_j))_{n\times n}=0$ 的解. 由 $(\varphi(\alpha_i,\beta_j))_{n\times n}\in\mathrm{GL}(n,F)$ 知 $(c_1,\cdots,c_n)=(0,\cdots,0)$, 即 $\alpha=0$, 故 φ_L 为单射. □

进一步地, 记 $\ker(\varphi_L)\subseteq V$ 为 φ 的**左根** (left radical), $\ker(\varphi_R)\subseteq W$ 为 φ 的**右根** (right radical). 注意一般左根和右根是不同线性空间的子空间. 但以下说明: 当 $\max\{\dim_F(V),\dim_F(W)\}<+\infty$ 时, $\mathrm{codim}_F(\ker(\varphi_L))=\mathrm{codim}_F(\ker(\varphi_R))$.

引理 3.2.10 *设 V,W 为域 F 上的有限维线性空间, $\varphi\colon V\times W\to F$ 为双线性函数, 则 $\dim_F(V)-\dim_F(\ker(\varphi_L))=\dim_F(W)-\dim_F(\ker(\varphi_R))$.*

证明: 记 $\overline{V}=V/\ker(\varphi_L)$, $\overline{W}=W/\ker(\varphi_R)$, 则 $\overline{\varphi}\colon \begin{array}{ccc}\overline{V}\times\overline{W}&\longrightarrow&F\\(\overline{\alpha},\overline{\beta})&\longmapsto&\varphi(\alpha,\beta)\end{array}$ 为定义良好的双线性函数, 且为左右非退化的. 由 $\overline{\varphi}_L$ 为单射知, $\dim_F(\overline{V})\le\dim_F(\overline{W}^*)$; 由 $\overline{\varphi}_R$ 为单射知, $\dim_F(\overline{W})\le\dim_F(\overline{V}^*)$. 因此由 $\overline{V},\overline{W}$ 有限维知 $\dim_F(\overline{V})=\dim_F(\overline{W})$, 即 $\dim_F(V)-\dim_F(\ker(\varphi_L))=\dim_F(W)-\dim_F(\ker(\varphi_R))$. □

注: 事实上, 当 $\max\{\dim_F(V),\dim_F(W)\}<+\infty$ 时, 任取 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 W 的基 $\{\beta_j\}_{j=1}^m$, 则 $\operatorname{codim}_F(\ker(\varphi_L))=\operatorname{codim}_F(\ker(\varphi_R))=r((\varphi(\alpha_i,\beta_j))_{n\times m})$.

通过分别做关于左根和右根的商空间, 一般的双线性函数可以化为非退化的. 以下考虑一种典型的非退化双线性函数, 它在某种程度上给出了对偶空间的另一种实现方式.

例 3.2.2 (典型的非退化双线性函数) 设 V 为域 F 上的线性空间, 映射 $\varphi\colon \begin{array}{ccc}V\times V^*&\longrightarrow&F\\(\alpha,f)&\longmapsto&f(\alpha)\end{array}$ 是一个双线性函数, 满足 φ_L 为单射 $\tau\colon \begin{array}{ccc}V&\longrightarrow&V^{**}\\\alpha&\longmapsto&L_\alpha\end{array}$, φ_R 为恒同 $\mathrm{id}\colon V^*\to V^*$, 故 φ 是非退化的. 现任取双线性函数 $\psi\colon V\times W\to F$, 则 ψ 必经过 φ 分解, 即 $\psi=\varphi\circ(\mathrm{id}\times\psi_R)$. 特别地, 当 $\dim_F(V)=\dim_F(W)<+\infty$ 且 ψ 非退化时, $\psi_R\colon W\to V^*$ 为线性同构, 故此时 $\psi\circ(\mathrm{id}\times\psi_R)^{-1}=\varphi$, 即上述典型的非退化双线性函数可由任意特殊的非退化双线性函数表示出来. 这启示我们对偶空间理论都可类比到双线性函数理论中并加以推广.

参考文献与补注 3.2 关于范畴学中自然变换的部分, 可以参考 J. J. Rotman "An Introduction to Homological Algebra".

§3.3 双线性映射与张量积

本节主要补充双线性映射的一些性质, 由此完成线性空间的张量积的构造.

设 V,W,U 为域 F 上的三个线性空间, $\Phi\colon V\times W\to U$ 为映射, 满足 Φ 关于每个分量都是线性映射, 则 Φ 称为一个**双线性映射** (bilinear map). 换句话说, 一个双线性映射 $\Phi\colon V\times W\to U$ 等价于一个线性映射 $\Phi_L\colon V\longrightarrow L(W,U)$ (或 $\Phi_R\colon W\longrightarrow L(V,U)$). 注意任取 $f\in U^*$, 则 $f\circ\Phi\colon V\times W\to F$ 是一个双线性函数, $\begin{array}{ccc}\alpha&\longmapsto&\Phi(\alpha,\cdot)\end{array}$ $\begin{array}{ccc}\beta&\longmapsto&\Phi(\cdot,\beta)\end{array}$ 故通过调整到达域, 双线性映射应该比双线性函数具有更灵活的性质. 以下考虑一种典型的双线性映射, 它的求迹给出了上节典型的非退化双线性函数.

命题 3.3.1 (典型的双线性映射) *设 V 为域 F 上的线性空间, 映射 $\Phi\colon \begin{array}{ccc}V\times V^*&\longrightarrow&L(V):=L(V,V)\\(\alpha,f)&\longmapsto&(T\colon\beta\mapsto f(\beta)\alpha)\end{array}$ 是双线性的, 证明: 任给双线性映射 $\Psi\colon V\times V^*\to W$, 存在线性映射 $S\colon L(V)\to W$, 满足 $\Psi=S\circ\Phi$.*

证明: 考虑 $L(V)$ 的线性子空间 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$ (注意这里 Φ 为双线性映射, 故 $\mathrm{Im}(\Phi)\subseteq L(V)$ 只是子集而非子空间).

取 V 的基 $\{\alpha_i\}_{i\in I}$, V^* 的基 $\{f_j\}_{j\in J}$, 断言: $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 为 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$ 的基. (这是因为, 由 Φ 的双线性性知, $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 线性生成 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$. 下证 $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 线性无关: 任取 $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 的有限非空子集, 考虑其线性组合 $c_1\Phi(\alpha_{i_1},f_{j_1})+\cdots+c_k\Phi(\alpha_{i_k},f_{j_k})=0$. 不妨设 $j_1=\cdots=j_l\notin\{j_{l+1},\cdots,j_k\}$, 则由 $\{f_j\}_{j\in J}$ 的线性无关性知, $\exists\beta\in V$, *s.t.* $f_{j_1}(\beta)=\cdots=f_{j_l}(\beta)\neq0=f_{j_{l+1}}(\beta)=\cdots=f_{j_k}(\beta)$, 故代入上式知 $c_1\alpha_{i_1}+\cdots+c_l\alpha_{i_l}=0$. 由于 $(i_1,j_1),\cdots,(i_k,j_k)$ 两两不同, 故 $i_1,\cdots,i_l$ 两两不同, 则由 $\{\alpha_i\}_{i\in I}$ 的线性无关性知, $c_1=\cdots=c_l=0$. 类似知 $c_{l+1}=\cdots=c_k=0$, 因此 $\{\Phi(\alpha_{i_1},f_{j_1}),\cdots,\Phi(\alpha_{i_k},f_{j_k})\}$ 为线性无关集, 从而 $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 为线性无关集.)

以下先构造线性映射 $S_0\colon\mathrm{Span}_F(\mathrm{Im}(\Phi))\to W$, 只需确定 S_0 在基 $\{\Phi(\alpha_i,f_j)\}_{i\in I,j\in J}$ 上的取值即可: 令 $S_0(\Phi(\alpha_i,f_j)):=\Psi(\alpha_i,f_j)$, $\forall i\in I$, $j\in J$. 注意这是定义良好的. (这是因为, 若 $\Phi(\alpha_i,f_j)=\Phi(\alpha_{i'},f_{j'})$, 则可取 $\beta\in V\setminus(\ker(f_j)\cap\ker(f_{j'}))$, 由 $f_j(\beta)\alpha_i=f_{j'}(\beta)\alpha_{i'}$ 知, $\{\alpha_i,\alpha_{i'}\}$ 线性相关, 则 $i=i'$, 故 $j=j'$.)

最后通过取 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$ 在 $L(V)$ 中的直和补空间, 可将 S_0 延拓为 $S\in L(L(V),W)$, 满足 $L\circ\Phi=\Psi$. □

注:

(1) 在上述命题的证明中, $L(V)$ 的线性子空间 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$ 起着非常关键的作用: 事实上, 任意双线性映射 $\Psi\colon V\times V^*\to W$ 都可唯一地经过 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$ 分解. 我们将看到这种泛性质在线性同构的意义下确定了 $\mathrm{Span}_F(\mathrm{Im}(\Phi))$, 它称为 V 与 V^* 的张量积.

(2) 特别地, 当 $\dim_F(V)=n<+\infty$ 时, 由 $\dim_F(\mathrm{Span}_F(\mathrm{Im}(\Phi)))=n^2=\dim_F(L(V))$ 知, $\mathrm{Span}_F(\mathrm{Im}(\Phi))=L(V)$. 而当 $\dim_F(V)=|B|$ 为无穷势时, $\mathrm{Span}_F(\mathrm{Im}(\Phi))\subseteq L(V)$ 未必能取等号. 以下简单给出它们的维数计算: $\dim_F(\mathrm{Span}_F(\mathrm{Im}(\Phi)))=\dim_F(V)\cdot\dim_F(V^*)=|B|\cdot|F^B|$; $\dim_F(L(V))=\dim_F(V^B)=|V^B|=(|F|\cdot|B|)^{|B|}$.

推论 3.3.2 *设 V 为域 F 上的有限维线性空间, $\Phi\colon V\times V^*\longrightarrow L(V)$ $(\alpha,f)\longmapsto(T\colon\beta\mapsto f(\beta)\alpha)$ 为双线性映射, $\varphi\colon V\times V^*\longrightarrow F$ $(\alpha,f)\longmapsto f(\alpha)$ 为双线性函数, 则存在唯一的 $\mathrm{tr}\in L(V)^*$, 满足 $\varphi=\mathrm{tr}\circ\Phi$.*

注:

(1) 在上述推论中, 由条件 $\varphi=\mathrm{tr}\circ\Phi$ 唯一确定的 $\mathrm{tr}\in L(V)^*$ 称为迹函数. 这与用矩阵定义的迹 tr 是一致的: 设 $T\in L(V)$ 在基 $\{\alpha_i\}_{i=1}^n$ 下的矩阵表示为 $A\in F^{n\times n}$, 即 $T(\alpha_j)=\sum\limits_{i=1}^n A_{ij}\alpha_i$, $\forall 1\le j\le n$; 再取 $\{\alpha_i\}_{i=1}^n$ 的对偶基 $\{f_i\}_{i=1}^n\subseteq V^*$, 则 $T=\sum\limits_{i,j=1}^n A_{ij}\Phi(\alpha_i,f_j)$, 故 $\mathrm{tr}(T)=\sum\limits_{i,j=1}^n A_{ij}\mathrm{tr}(\Phi(\alpha_i,f_j))=\sum\limits_{i,j=1}^n A_{ij}\varphi(\alpha_i,f_j)$ $=\sum\limits_{i,j=1}^n A_{ij}f_j(\alpha_i)=\sum\limits_{i=1}^n A_{ii}=\mathrm{tr}(A)$. 于是这也说明同一线性变换的不同矩阵表示的迹均相同.

(2) 当 $\dim_F(V)<+\infty$ 时, 记 V 与 V^* 的张量积 $V\underset{F}{\otimes}V^*\coloneqq L(V)$, 其中纯张量为 $\alpha\otimes f\coloneqq\Phi(\alpha,f)$, 则迹函数 $\mathrm{tr}\colon V\underset{F}{\otimes}V^*\longrightarrow F$ $\alpha\otimes f\longmapsto f(\alpha)$ 又称为张量指标的**缩并** (contraction), 它将一个协变指标和一个反变指标自然配对得到了一个标量. 这种张量指标的缩并与升降是流形上微积分的基本工具.

现在正式介绍线性空间的**张量积** (tensor product). 在最一般的情形中, 我们将给出线性空间的张量积的三种等价描述.

设 V,W 为域 F 上的两个线性空间. 最原始的想法是模仿上述命题的思路, 利用双线性映射的泛性质确定 V 与 W 的张量积. 具体地说, 线性空间 V 与 W 的张量积应当是一个线性空间 $V\underset{F}{\otimes}W$ 与一个双线性映射 $\Phi\colon V\times W\to V\underset{F}{\otimes}W$, 满足以下的泛性质: 任给线性空间 U 与双线性映射 $\Psi\colon V\times W\to U$, 存在唯一的线性映射 $S\colon V\underset{F}{\otimes}W\to U$, 使得 $\Psi=S\circ\Phi$.

上述泛性质可用更范畴化的语言叙述: 固定域 F 上的两个线性空间 V,W, 考虑范畴 $\mathscr{C}$ 如下:

$$\mathrm{obj}(\mathscr{C})\coloneqq\{(U,\Psi)\colon U\text{为域}F\text{上的线性空间},\ \Psi\colon V\times W\to U\text{为双线性映射}\},$$

以及 $\forall(U_1,\Psi_1),(U_2,\Psi_2)\in\mathrm{obj}(\mathscr{C})$,

$$\mathrm{Hom}_{\mathscr{C}}((U_1,\Psi_1),(U_2,\Psi_2))\coloneqq\{T\in L(U_1,U_2)\colon\Psi_2=T\circ\Psi_1\},$$

则线性空间 V 与 W 的张量积是上述范畴 $\mathscr{C}$ 中的始对象 (initial object).

上述泛性质也可用伴随函子的语言叙述: 任取域 F 上的三个线性空间 V,W,U, 均存在一个只与 V,W 有关的线性空间 $V\underset{F}{\otimes}W$, 以及一个双线性映射 $\Phi\colon V\times W\to V\underset{F}{\otimes}W$, 满足以下的线性空间同构:

$$\begin{aligned}\mathrm{Hom}_{\mathbf{F-Mod}}(V\underset{F}{\otimes}W,U)&\xrightarrow{\cong}\mathrm{Hom}_{\mathbf{F-Mod}}(V,\mathrm{Hom}_{\mathbf{F-Mod}}(W,U))\\ S&\longmapsto(S\circ\Phi)_L\end{aligned}$$

(或者

$$\begin{aligned}\mathrm{Hom}_{\mathbf{F-Mod}}(V\underset{F}{\otimes}W,U)&\xrightarrow{\cong}\mathrm{Hom}_{\mathbf{F-Mod}}(W,\mathrm{Hom}_{\mathbf{F-Mod}}(V,U))\\ S&\longmapsto(S\circ\Phi)_R\end{aligned}$$

). 这里显式给出了同构映射, 它们关于每个位置都是自然的. 这表明在范畴 **F-Mod** 上存在两对伴随函子

$$-\underset{F}{\otimes}W\colon\mathbf{F-Mod}\rightleftarrows\mathbf{F-Mod}\colon\mathrm{Hom}_{\mathbf{F-Mod}}(W,-)$$

与

$$V\underset{F}{\otimes}-\colon\mathbf{F-Mod}\rightleftarrows\mathbf{F-Mod}\colon\mathrm{Hom}_{\mathbf{F-Mod}}(V,-).$$

特别地, 由 Yoneda 引理知, 存在函子的自然同构 $-\underset{F}{\otimes} V \cong V \underset{F}{\otimes} -$, 即线性空间的张量积关于左右是平衡的.

最困难的部分在于显式构造出满足上述泛性质的线性空间 $V \underset{F}{\otimes} W$ 与双线性映射 $\Phi\colon V \times W \to V \underset{F}{\otimes} W$. 我们的方法是将线性空间 $V \underset{F}{\otimes} W$ 实现为某个自由线性空间关于某个线性子空间的商空间, 其中自由线性空间应当包含集合 $V \times W$ 中的所有元而忘记附带的线性结构, 线性子空间应当包含自由线性空间中的元之间关于双线性性的差距. 具体地说, 记 $\mathscr{G}\colon \mathbf{F\text{-}Mod} \to \mathbf{Set}$ 为忘记函子, $X \coloneqq F^{(\mathscr{G}(V\times W))}$ 为集合 $\mathscr{G}(V \times W)$ 上的自由线性空间, $R \subseteq X$ 为由以下元生成的线性子空间:

$$(\alpha_1 + \alpha_2, \beta) - (\alpha_1, \beta) - (\alpha_2, \beta),\ (\alpha, \beta_1 + \beta_2) - (\alpha, \beta_1) - (\alpha, \beta_2),$$
$$(c\alpha, \beta) - c(\alpha, \beta),\ (\alpha, c\beta) - c(\alpha, \beta),$$

其中 $\alpha, \alpha_1, \alpha_2 \in V,\ \beta, \beta_1, \beta_2 \in W,\ c \in F$, 则令线性空间 $V \underset{F}{\otimes} W \coloneqq X/R$, 以及双线性映射 $\Phi\colon V \times W \to V \underset{F}{\otimes} W$ 为含入映射 $V \times W \hookrightarrow X$ 与商映射 $X \twoheadrightarrow V \underset{F}{\otimes} W$ 的复合. 记 $\alpha \otimes \beta \coloneqq \Phi(\alpha, \beta)$ 为 $V \underset{F}{\otimes} W$ 中的纯张量. 于是可直接验证这种构造满足上述伴随函子描述的泛性质, 而由泛性质确定的对象在同构意义下是唯一的.

接着我们插入张量积空间的一些简单性质.

引理 3.3.3 (张量积的基) *设 V, W 为域 F 上的线性空间, $\{\alpha_i\}_{i\in I}$ 为 V 的基, $\{\beta_j\}_{j\in J}$ 为 W 的基, 则 $\{\alpha_i \otimes \beta_j\}_{i\in I, j\in J}$ 为 $V \underset{F}{\otimes} W$ 的基. 特别地, $\dim_F(V \underset{F}{\otimes} W) = \dim_F(V) \cdot \dim_F(W)$.*

证明: 由线性空间 $V \underset{F}{\otimes} W$ 的构造与映射 $\Phi\colon V \times W \to V \underset{F}{\otimes} W$ 的双线性性知, $\mathrm{Span}_F(\{\alpha_i \otimes \beta_j\}_{i\in I, j\in J}) = V \underset{F}{\otimes} W$. 下证 $\{\alpha_i \otimes \beta_j\}_{i\in I, j\in J}$ 为线性无关集. 事实上, 可取 $\{\alpha_i\}_{i\in I}$ 的对偶集 $\{f_{\alpha_i}\}_{i\in I} \subseteq V^*$, 以及 $\{\beta_j\}_{j\in J}$ 的对偶集 $\{g_{\beta_j}\}_{j\in J} \subseteq W^*$, 考虑双线性函数 $\begin{array}{ccc} V \times W & \longrightarrow & F \\ (\alpha, \beta) & \longmapsto & f_{\alpha_i}(\alpha) g_{\beta_j}(\beta) \end{array}$ $(i \in I, j \in J)$. 由线性空间的张量积的泛性质知, 它可经过线性函数 $f_{\alpha_i} \otimes g_{\beta_j}\colon \begin{array}{ccc} V \underset{F}{\otimes} W & \longrightarrow & F \\ \alpha \otimes \beta & \longmapsto & f_{\alpha_i}(\alpha) g_{\beta_j}(\beta) \end{array}$ 分解. 注意 $(f_{\alpha_i} \otimes g_{\beta_j})(\alpha_k \otimes \beta_l) = \delta_{ik}\delta_{jl}$, 则由对偶集的线性无关性引理知, $\{\alpha_i \otimes \beta_j\}_{i\in I, j\in J}$ 为线性无关集. □

引理 3.3.4 (张量积的结合性) *设 V, W, U 为域 F 上的线性空间, 则存在线性同构 $\begin{array}{ccc} (V \underset{F}{\otimes} W) \underset{F}{\otimes} U & \xrightarrow{\cong} & V \underset{F}{\otimes} (W \underset{F}{\otimes} U) \\ (\alpha \otimes \beta) \otimes \gamma & \longmapsto & \alpha \otimes (\beta \otimes \gamma) \end{array}$.*

证明: 通过基对应或者利用泛性质直接验证即可. □

引理 3.3.5 (张量积的交换性) *设 V, W 为域 F 上的线性空间, 则存在线性同构 $\begin{array}{ccc} V \underset{F}{\otimes} W & \xrightarrow{\cong} & W \underset{F}{\otimes} V \\ \alpha \otimes \beta & \longmapsto & \beta \otimes \alpha \end{array}$.*

证明: 通过基对应或者利用泛性质直接验证即可. □

以下我们给出上述纯张量的一种直观解释, 这可视为线性空间的张量积的另一种刻画. 任取 $\alpha \in V,\ \beta \in W$, 记 $\alpha \otimes' \beta\colon \begin{array}{ccc} V^* \times W^* & \longrightarrow & F \\ (f, g) & \longmapsto & f(\alpha) g(\beta) \end{array}$ 为双线性函数, 以及 $V \underset{F}{\otimes'} W \coloneqq \mathrm{Span}_F(\{\alpha \otimes' \beta\colon \alpha \in V, \beta \in W\})$ 为另一种张量积. 类似前述张量积的基的证明, 同样可说明: 设 $\{\alpha_i\}_{i\in I}$ 为 V 的基, $\{\beta_j\}_{j\in J}$ 为 W 的基, 则 $\{\alpha_i \otimes' \beta_j\}_{i\in I, j\in J}$ 为 $V \underset{F}{\otimes'} W$ 的基. 于是通过基对应 $\alpha_i \otimes \beta_j \mapsto \alpha_i \otimes' \beta_j$ $(i \in I, j \in J)$ 可给出线性空间同构 $V \underset{F}{\otimes} W \xrightarrow{\cong} V \underset{F}{\otimes'} W$. 当然, 利用线性空间的张量积的泛性质也可直接验证此同构.

最后我们补充线性映射的张量积, 以及它在有限维线性空间上的矩阵表示.

设 V_1, V_2, W_1, W_2 为域 F 上的线性空间, $T \in L(V_1, V_2),\ S \in L(W_1, W_2)$, 考虑双线性映射 $\begin{array}{ccc} V_1 \times W_1 & \longrightarrow & V_2 \underset{F}{\otimes} W_2 \\ (\alpha, \beta) & \longmapsto & T(\alpha) \otimes S(\beta) \end{array}$. 由线性空间的张量积的泛性质知, 它可经过线性映射 $T \otimes S\colon \begin{array}{ccc} V_1 \underset{F}{\otimes} W_1 & \longrightarrow & V_2 \underset{F}{\otimes} W_2 \\ \alpha \otimes \beta & \longmapsto & T(\alpha) \otimes S(\beta) \end{array}$ 分解, 后者称为线性映射 T 与 S 的张量积. 容易验证若 T, S 均为单射 (或满射), 则 $T \otimes S$ 也为单射 (或满射). 进一步地, 设 $V_1 = V_2 = V$,

$W_1=W_2=W$ 均为域 F 上的有限维线性空间, T 在 V 的基 $\{\alpha_i\}_{i=1}^n$ 下的矩阵表示为 $A\in F^{n\times n}$, S 在 W 的基 $\{\beta_j\}_{j=1}^m$ 下的矩阵表示为 $B\in F^{m\times m}$, 以下讨论 $T\otimes S$ 在 $V\underset{F}{\otimes}W$ 的基 $\{\alpha_i\otimes\beta_j\}_{1\leq i\leq n,\,1\leq j\leq m}$ 下的矩阵表示.

注意 $V\underset{F}{\otimes}W$ 的基 $\{\alpha_i\otimes\beta_j\}_{1\leq i\leq n,\,1\leq j\leq m}$ 有两种排序方式:

(1) $\alpha_1\otimes\beta_1,\cdots,\alpha_1\otimes\beta_m;\cdots;\alpha_n\otimes\beta_1,\cdots,\alpha_n\otimes\beta_m$;

(2) $\alpha_1\otimes\beta_1,\cdots,\alpha_n\otimes\beta_1;\cdots;\alpha_1\otimes\beta_m,\cdots,\alpha_n\otimes\beta_m$.

由于 $(T\otimes S)(\alpha_i\otimes\beta_j)=T(\alpha_i)\otimes S(\beta_j)=\left(\sum\limits_{k=1}^n A_{ki}\alpha_k\right)\otimes\left(\sum\limits_{l=1}^m B_{lj}\beta_l\right)=\sum\limits_{k=1}^n\sum\limits_{l=1}^m A_{ki}B_{lj}\alpha_k\otimes\beta_l$, 则:

$T\otimes S$ 在基 (1) 下的矩阵表示为

$$\begin{pmatrix} A_{11}B_{11} & \cdots & A_{11}B_{1m} & \cdots & A_{1n}B_{11} & \cdots & A_{1n}B_{1m}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{11}B_{m1} & \cdots & A_{11}B_{mm} & \cdots & A_{1n}B_{m1} & \cdots & A_{1n}B_{mm}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{n1}B_{11} & \cdots & A_{n1}B_{1m} & \cdots & A_{nn}B_{11} & \cdots & A_{nn}B_{1m}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{n1}B_{m1} & \cdots & A_{n1}B_{mm} & \cdots & A_{nn}B_{m1} & \cdots & A_{nn}B_{mm} \end{pmatrix}=\begin{pmatrix} A_{11}B & \cdots & A_{1n}B\\ \vdots & \vdots & \vdots\\ A_{n1}B & \cdots & A_{nn}B \end{pmatrix}=:A\otimes B;$$

$T\otimes S$ 在基 (2) 下的矩阵表示为

$$\begin{pmatrix} A_{11}B_{11} & \cdots & A_{1n}B_{11} & \cdots & A_{11}B_{1m} & \cdots & A_{1n}B_{1m}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{n1}B_{11} & \cdots & A_{nn}B_{11} & \cdots & A_{n1}B_{1m} & \cdots & A_{nn}B_{1m}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{11}B_{m1} & \cdots & A_{1n}B_{m1} & \cdots & A_{11}B_{mm} & \cdots & A_{1n}B_{mm}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ A_{n1}B_{m1} & \cdots & A_{nn}B_{m1} & \cdots & A_{n1}B_{mm} & \cdots & A_{nn}B_{mm} \end{pmatrix}=\begin{pmatrix} B_{11}A & \cdots & B_{1m}A\\ \vdots & \vdots & \vdots\\ B_{m1}A & \cdots & B_{mm}A \end{pmatrix}=:B\otimes A.$$

这里 $A\otimes B\in F^{(nm)\times(nm)}$ 称为矩阵 A 与 B 的 Kronecker 积. 由线性映射的张量积具有双线性性与结合性知, 矩阵的 Kronecker 积也具有双线性性与结合性. 由于同一线性映射在不同基下的矩阵表示是相似的, 则 $A\otimes B$ 相似于 $B\otimes A$. 再由线性映射的复合知, $\forall A_1,A_2\in F^{n\times n},B_1,B_2\in F^{m\times m}$, $(A_1\otimes B_1)\cdot(A_2\otimes B_2)=(A_1\cdot A_2)\otimes(B_1\cdot B_2)$; $\forall A\in\mathrm{GL}(n,F),B\in\mathrm{GL}(m,F)$, $(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$. 因此由矩阵的相抵标准形知, $\forall A\in F^{n\times n},B\in F^{m\times m}$, $r(A\otimes B)=r(A)r(B)$.

例 3.3.1 设 F 为一个域, $A\in F^{n\times n}$, $B\in F^{m\times m}$, 求线性映射 $T\colon\begin{array}{ccc}F^{m\times n}&\longrightarrow&F^{m\times n}\\ X&\longmapsto&B\cdot X\cdot A^t\end{array}$ 的一个矩阵表示.

证明: 考虑线性同构 $\mathrm{vec}\colon\begin{array}{ccc}F^{m\times n}&\xrightarrow{\cong}&F^{(mn)\times1}\\ X&\longmapsto&(X_{11},\cdots,X_{m1};\cdots;X_{1n},\cdots,X_{mn})^t\end{array}$, 则 $\mathrm{vec}(B\cdot X\cdot A^t)=(A\otimes B)\cdot\mathrm{vec}(X)$, 故存在交换图 $\begin{array}{ccc}F^{m\times n}&\xrightarrow{T}&F^{m\times n}\\ \downarrow{\scriptstyle\mathrm{vec}}&&\downarrow{\scriptstyle\mathrm{vec}}\\ F^{(mn)\times1}&\xrightarrow{L_{A\otimes B}}&F^{(mn)\times1}\end{array}$. 由 $L_{A\otimes B}$ 的一个矩阵表示为 $A\otimes B$ 知, T 也有一个矩阵表示为 $A\otimes B$.

□

例 3.3.2 设 F 为一个域, $A\in F^{n\times n}$, $B\in F^{m\times m}$, 求线性映射 $T\colon\begin{array}{ccc}F^{m\times n}&\longrightarrow&F^{m\times n}\\ X&\longmapsto&B\cdot X+X\cdot A^t\end{array}$ 的一个矩阵表示.

证明: 仍考虑线性同构 $\mathrm{vec}\colon F^{m\times n}\xrightarrow{\cong}F^{(mn)\times1}$, 则 $\mathrm{vec}(B\cdot X+X\cdot A^t)=(I_n\otimes B+A\otimes I_m)\cdot\mathrm{vec}(X)$. 记 A 与 B 的 Kronecker 和为 $A\oplus B:=I_n\otimes B+A\otimes I_m$, 则存在交换图 $\begin{array}{ccc}F^{m\times n}&\xrightarrow{T}&F^{m\times n}\\ \downarrow{\scriptstyle\mathrm{vec}}&&\downarrow{\scriptstyle\mathrm{vec}}\\ F^{(mn)\times1}&\xrightarrow{L_{A\oplus B}}&F^{(mn)\times1}\end{array}$. 由 $L_{A\oplus B}$ 的一个矩阵表示为 $A\oplus B$ 知, T 也有一个矩阵表示为 $A\oplus B$. □

参考文献与补注 3.3

(1) 关于流形上的张量运算的部分, 可以参考梅加强 “流形与几何初步”.
(2) 关于一般模的张量积的部分, 可以参考 Keith Conrad “Tensor Products”.

第 4 章　行列式

本章主要介绍线性代数的一大利器: 行列式 (determinant). 它当然可以视为线性方程组与矩阵理论的延续, 但这里的观点与技术更具线性映射的特色.

§4.1　置换群与群表示

4.1.1　置换群与辫群

本节继续补充群论中的简单定义, 并由此研究置换群的结构.

定义 4.1.1 (正规子群) 设 $(G,p,1)$ 是一个群, $(K,p,1)\leq(G,p,1)$ 为子群, 若 $\forall a\in G$, $p(a,K)=p(K,a)$, 则称 $(K,p,1)$ 为 $(G,p,1)$ 的一个**正规子群** (normal subgroup), 记为 $(K,p,1)\trianglelefteq(G,p,1)$.

注:

(1) 回忆群同态的概念, 它可给出正规子群的另一刻画: 子群 $(K,p,1)\leq(G,p,1)$ 为正规子群 $\Longleftrightarrow$ 存在另一群 $(H,p',1')$, 以及群同态 $\varphi\colon(G,p,1)\to(H,p',1')$, 使得 $\ker(\varphi)=K$. 这里群 $(H,p',1')$ 存在以下典范的取法: 取 $H\coloneqq\{p(a,K)\colon a\in G\}$, $p'(p(a,K),p(b,K))\coloneqq p(p(a,b),K)$, $1'\coloneqq p(1,K)$, 则 $(H,p',1')$ 是一个群, 称为群 $(G,p,1)$ 关于正规子群 $(K,p,1)$ 的 **商群** (quotient subgroup), 记为 $(G/K,\overline{p},\overline{1})$; 此时相应的群同态 $\varphi\colon\ (G,p,1)\longrightarrow(G/K,\overline{p},\overline{1})$ 称为**商同态** (quotient homomorphism).

$$a\longmapsto p(a,K)$$

(2) 通过关于群中的正规子群做商群 (即考虑群正合列 $1\to(K,p,1)\to(G,p,1)\to(G/K,\overline{p},\overline{1})\to 1$), 我们可不断约化群的结构. 在此意义下, 最简单的群为不含非平凡正规子群的群, 称为**单群** (simple group). 一般单群的分类是个过于复杂的问题, 甚至有限单群的分类也直至 2004 年才完全解决, 它们是素数阶循环群、5 个文字以上的交错群、Lie 型单群、26 个散在单群和 Tits 群.

定义 4.1.2 (特征子群) 设 $(G,p,1)$ 是一个群, $(K,p,1)\leq(G,p,1)$ 为子群, 若 $\forall\varphi\in\mathrm{Aut}(G,p,1)$, $\varphi(K)=K$, 则称 $(K,p,1)$ 为 $(G,p,1)$ 的一个**特征子群** (characteristic subgroup), 记为 $(K,p,1)\operatorname{char}(G,p,1)$.

注:

(1) 特征子群必为正规子群. 这是因为, 可取 $\varphi\in\mathrm{Aut}(G,p,1)$ 为群的内自同构 $\mathrm{Int}(a)\colon\ (G,p,1)\longrightarrow(G,p,1)$

$$b\longmapsto p(p(a,b),a^{-1})$$

$(a\in K)$, 则 $p(a,K)=p(K,a)\Longleftrightarrow\mathrm{Int}(a)(K)=K$.

(2) 不含非平凡特征子群的群称为**特征单群** (characteristically simple group). 例如单群必为特征单群. 可以证明: 一族彼此同构的单群的直积是特征单群; 具有极小非平凡正规子群的特征单群是一族彼此同构的单群的直积. 后者对于无极小非平凡正规子群的特征单群不成立, 反例如 $(\mathbb{Q},+,0)$.

引理 4.1.1 *设 $(K,p,1)\leq(H,p,1)\leq(G,p,1)$ 为子群,*

(1) *若 $(K,p,1)\trianglelefteq(H,p,1)\trianglelefteq(G,p,1)$, 则 $(K,p,1)\trianglelefteq(G,p,1)$ 未必成立;*

(2) *若 $(K,p,1)\operatorname{char}(H,p,1)\trianglelefteq(G,p,1)$, 则 $(K,p,1)\trianglelefteq(G,p,1)$;*

(3) *若 $(K,p,1)\operatorname{char}(H,p,1)\operatorname{char}(G,p,1)$, 则 $(K,p,1)\operatorname{char}(G,p,1)$.*

证明: (1) 取 $G=A_4$, $H=\{\mathrm{id},(1,2)\circ(3,4),(1,4)\circ(2,3),(1,3)\circ(2,4)\}$, $K=\{\mathrm{id},(1,2)\circ(3,4)\}$ 即为反例.
(2) 由 $(H,p,1)\trianglelefteq(G,p,1)$ 知, $\forall a\in G$, $\mathrm{Int}(a)\big|_H\in\mathrm{Aut}(H,p,1)$; 再由 $(K,p,1)\operatorname{char}(H,p,1)$ 知 $\mathrm{Int}(a)\big|_H(K)=K$, 即 $(K,p,1)\trianglelefteq(G,p,1)$.
(3) 由 $(H,p,1)\operatorname{char}(G,p,1)$ 知, $\forall\varphi\in\mathrm{Aut}(G,p,1)$, $\varphi\big|_H\in\mathrm{Aut}(H,p,1)$; 再由 $(K,p,1)\operatorname{char}(H,p,1)$ 知 $\varphi\big|_H(K)=K$, 即 $(K,p,1)\operatorname{char}(G,p,1)$. □

由于符号映射 $\operatorname{sgn}\colon (S_n,\circ,\mathrm{id}) \longrightarrow (\mathcal{C}_2,\cdot,1)$ 为群同态, 则 $A_n \coloneqq \ker(\operatorname{sgn})$ 为 $(S_n,\circ,\mathrm{id})$ 的正规子群. 以下通过 S_n 的一些性质说明 $(A_n,\circ,\mathrm{id})$ 也为 $(S_n,\circ,\mathrm{id})$ 的特征子群.

$$\sigma \longmapsto \prod_{1\le i<j\le n} \frac{\sigma(j)-\sigma(i)}{j-i}$$

引理 4.1.2 设 $n\ge 2$, 则 $(A_n,\circ,\mathrm{id})$ 是 $(S_n,\circ,\mathrm{id})$ 中唯一的指标为 2 的子群.

证明: 由陪集分解知, 群中指标为 2 的子群必为正规子群, 故只需考虑非平凡的群同态 $\varphi\colon (S_n,\circ,\mathrm{id})\to(\mathcal{C}_2,\cdot,1)$. 注意 S_n 中的对换都相互共轭, 而 $(\mathcal{C}_2,\cdot,1)$ 为交换群, 故 φ 在 S_n 中的对换上均取同一值. 又 S_n 可由对换生成, 且 φ 非平凡, 则 φ 在 S_n 中的对换上均取 -1, 故 $A_n=\{\sigma\in S_n\colon \sigma\text{为偶数个对换的乘积}\}=\ker(\varphi)$. □

推论 4.1.3 设 $n\ge 1$, 则 $(A_n,\circ,\mathrm{id})\operatorname{char}(S_n,\circ,\mathrm{id})$.

证明: 当 $n=1$ 时结论平凡. 当 $n\ge 2$ 时, 由于 $\forall\varphi\in\operatorname{Aut}(S_n,\circ,\mathrm{id})$, $(\varphi(A_n),\circ,\mathrm{id})$ 也为 $(S_n,\circ,\mathrm{id})$ 中指标为 2 的子群, 故由引理知 $\varphi(A_n)=A_n$. □

定义 4.1.3 (交换子群) 设 $(G,p,1)$ 为群, 记 $G'\coloneqq\left\langle g_1g_2g_1^{-1}g_2^{-1}\colon g_1,g_2\in G\right\rangle$, 则 $(G',p,1)$ 也是一个群, 称为 $(G,p,1)$ 的**交换子群** (commutator group) 或**导群** (derived group).

注: 显然 $(G',p,1)\operatorname{char}(G,p,1)$. 记 $G^{(0)}=G$, $G^{(k)}=(G^{(k-1)})'\,(k\ge 1)$, 则 $\{(G^{(k)},p,1)\}_{k\in\mathbb{N}}$ 为 $(G,p,1)$ 中的(特征) 子群降链. 若 $\exists k\in\mathbb{N}$, $s.t.$ $G^{(k)}=\{1\}$, 则称 $(G,p,1)$ 为**可解群** (solvable group). 一个非平凡的事实是, $(S_n,\circ,\mathrm{id})$ 为可解群 $\Longleftrightarrow$ $(A_n,\circ,\mathrm{id})$ 为可解群 $\Longleftrightarrow$ $n=1,2,3,4$.

引理 4.1.4 设 $n\ge 1$, 则 $A_n=S_n'\coloneqq\left\langle\sigma_1\circ\sigma_2\circ\sigma_1^{-1}\circ\sigma_2^{-1}\colon\sigma_1,\sigma_2\in S_n\right\rangle$.

证明: 当 $n=1,2$ 时结论平凡. 现设 $n\ge 3$. 一方面, 由于 $\forall\sigma_1,\sigma_2\in S_n$, $\sigma_1\circ\sigma_2\circ\sigma_1^{-1}\circ\sigma_2^{-1}\in\ker(\operatorname{sgn})$, 则 $S_n'\subseteq\ker(\operatorname{sgn})=A_n$. 另一方面, 注意 A_n 可由 S_n 中所有长度为 3 的轮换生成. (这是因为, 显然长度为 3 的轮换 $(i,j,k)=(i,k)\circ(i,j)\in A_n$; 而 A_n 中元 $(i,k)\circ(i,j)=(i,j,k)$, $(k,l)\circ(i,j)=(i,l,k)\circ(i,j,k)$ 均为长度为 3 的轮换的乘积.) 又长度为 3 的轮换 $(i,j,k)=(i,j)\circ(i,k)\circ(i,j)^{-1}\circ(i,k)^{-1}\in S_n'$, 故 $A_n\subseteq S_n'$. □

推论 4.1.5 设 $n\ge 1$, 则 $(A_n,\circ,\mathrm{id})\operatorname{char}(S_n,\circ,\mathrm{id})$.

证明: 由引理知, $\forall\varphi\in\operatorname{Aut}(S_n,\circ,\mathrm{id})$, $\varphi(A_n)=\varphi(S_n')=\varphi(S_n)'=S_n'=A_n$. □

最后, 我们研究置换群与辫群的关系. 注意 S_n 可由相邻对换生成, 记 $\sigma_i=(i,i+1)\,(1\le i\le n-1)$, 则它们恰满足关系 $\begin{cases}\sigma_i^2=\mathrm{id}\\ (\sigma_i\circ\sigma_{i+1})^3=\mathrm{id}\\ \sigma_i\circ\sigma_j=\sigma_j\circ\sigma_i,\ \forall|i-j|>1\end{cases}$, 这表明置换群 $(S_n,\circ,\mathrm{id})$ 具有有限生成元的有限表现形式:

$$(S_n,\circ,\mathrm{id})=\left\langle\sigma_1,\cdots,\sigma_{n-1}\colon\sigma_i^2=\mathrm{id};\ \sigma_i\circ\sigma_{i+1}\circ\sigma_i=\sigma_{i+1}\circ\sigma_i\circ\sigma_{i+1};\ \sigma_i\circ\sigma_j=\sigma_j\circ\sigma_i,\ \forall|i-j|>1\right\rangle.$$

抽象地说, 一个**辫群** (braid group) 是指以下有限生成元的有限表现群

$$(B_n,\cdot,1)\coloneqq\left\langle b_1,\cdots,b_{n-1}\colon b_ib_{i+1}b_i=b_{i+1}b_ib_{i+1};\ b_ib_j=b_jb_i,\ \forall|i-j|>1\right\rangle.$$

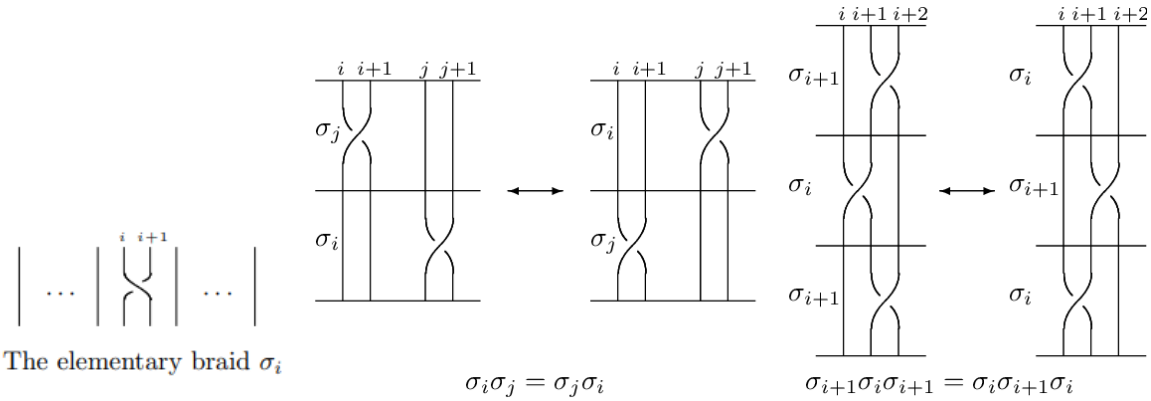
于是存在满的群同态 $(B_n,\cdot,1)\twoheadrightarrow(S_n,\circ,\mathrm{id})$, 它的核记为 PB_n. 注意当 $n\ge 2$ 时, B_n 与 PB_n 都是无限集, 且 $b_i\longmapsto\sigma_i$

$1\to(PB_n,\cdot,1)\to(B_n,\cdot,1)\to(S_n,\circ,\mathrm{id})\to 1$ 是个不可裂的群正合列.

形象地说, 辫群中元可由以下的**辫图** (braid diagram) 实现: 设 $\{p_1,\cdots,p_n\}\subseteq\mathbb{R}^2$ 为 n 个点, $(\phi_1,\cdots,\phi_n)$ 为 n 个连续映射, $\phi_i\colon[0,1]\to\mathbb{R}^2\,(1\le i\le n)$ 满足 $\phi_i(0)=p_i$, $\phi_i(1)=p_{\sigma(i)}\,(\sigma\in S_n)$, 且 $\operatorname{graph}(\phi_i)\,(1\le i\le n)$ 两两不交. 此时 $(\phi_1,\cdots,\phi_n)$ 称为一个辫子. 两个辫子 $(\phi_1,\cdots,\phi_n),(\psi_1,\cdots,\psi_n)$ 的乘积是一个新的辫子 $((\phi\cdot\psi)_1,\cdots,(\phi\cdot\psi)_n)$, 其中 $(\phi\cdot\psi)_i\colon[0,1]\longrightarrow\mathbb{R}^2$. 于是辫群中的关系如

$$t\longmapsto\begin{cases}\phi_i(2t), & 0\le t\le 1/2\\ \psi_j(2t-1), & 1/2\le t\le 1\end{cases}\quad(\phi_i(1)=p_j)$$

下图:



辫群的另一种实现方式如下: 设 X 为拓扑空间, 记 X 的 n **次有序构形空间** (the n^{th} ordered configuration space) $\mathrm{Conf}_n(X) \coloneqq \{\underline{x} = (x_1, \cdots, x_n) \in X^n \colon x_i \neq x_j,\ \forall i \neq j\}$, 此时群 $(S_n, \circ, \mathrm{id})$ 可作用于拓扑空间 $\mathrm{Conf}_n(X)$ 上: $S_n \times \mathrm{Conf}_n(X) \longrightarrow \mathrm{Conf}_n(X)$, 该作用的轨道空间记为 $\mathrm{Uconf}_n(X) \coloneqq S_n \backslash \mathrm{Conf}_n(X)$ (n **次无序构形空间** (the n^{th} unordered configuration space)). 一个经典的结果是, $B_n \cong \pi_1(\mathrm{Uconf}_n(\mathbb{C}))$; $PB_n \cong \pi_1(\mathrm{Conf}_n(\mathbb{C}))$.

$$(\sigma, \underline{x}) \longmapsto \sigma(\underline{x}) \coloneqq (x_{\sigma(1)}, \cdots, x_{\sigma(n)})$$

4.1.2 群的线性表示

群表示的思想是将抽象群作用于线性空间上, 从而更直观地获取群本身的一些性质.

例 4.1.1 ($(S_n, \circ, \mathrm{id})$ 在 n 维线性空间上的作用) 设 V 为域 F 上的 n 维线性空间, $B = \{\alpha_i\}_{i=1}^n$ 为 V 的基. 注意 $(S_n, \circ, \mathrm{id})$ 可视为 n 元集合 B 的置换群 $(\mathrm{Aut}(B), \circ, \mathrm{id})$, 即存在群作用 $S_n \times B \longrightarrow B$, 则可进一步诱导

$$(\sigma, \alpha_i) \longmapsto \alpha_{\sigma(i)}$$

群作用

$S_n \times V \longrightarrow V$. 记 $\mathrm{GL}(V) \coloneqq L(V)^\times$, 则存在群同态 $(S_n, \circ, \mathrm{id}) \longrightarrow \mathrm{GL}(V)$.

$$(\sigma, \sum_{i=1}^n c_i\alpha_i) \longmapsto \sum_{i=1}^n c_i\alpha_{\sigma(i)} \qquad \sigma \longmapsto \left(T_\sigma \colon \sum_{i=1}^n c_i\alpha_i \mapsto \sum_{i=1}^n c_i\alpha_{\sigma(i)}\right)$$

若考虑 T_σ 在基 B 下的矩阵表示 $R_\sigma = (\epsilon_{\sigma(1)}, \cdots, \epsilon_{\sigma(n)})$, 其中 $\epsilon_j \coloneqq (0, \cdots, 0, 1, 0, \cdots, 0)^t \in F^{n\times 1}$, 则存在群同态 $(S_n, \circ, \mathrm{id}) \longrightarrow \mathrm{GL}(n, F)$. 特别地, 上述两个群同态均为单射, 它们的作用在于将抽象群 $(S_n, \circ, \mathrm{id})$ 视为可逆线性变换群 $\mathrm{GL}(V)$ (或可逆矩阵群 $\mathrm{GL}(n, F)$) 的子群.

$$\sigma \longmapsto R_\sigma$$

定义 4.1.4 (群的线性表示) 设 $(G, p, 1)$ 为一个群, V 为域 F 上的线性空间, 若存在群同态 $\rho \colon (G, p, 1) \to \mathrm{GL}(V)$, 则称 (ρ, V) 为群 $(G, p, 1)$ 的一个线性表示 (linear representation), 其中 $\dim_F(V)$ 称为表示的维数.

注:

(1) 记 $F[G]$ 为以 G 为基生成的 F-线性空间, 即 $F[G]$ 中元形如 (有限和) $\sum\limits_{g\in G} c_g g$, 并将 G 上的二元运算 p 线性延拓为 $F[G]$ 上的乘法 $\cdot$, 即 $\left(\sum\limits_{g\in G} c_g g\right) \cdot \left(\sum\limits_{h\in G} c_h h\right) \coloneqq \sum\limits_{g\in G} c_g c_h p(g, h)$, 此时 $F[G]$ 为一个 F-结合代数, 称为群 G 在域 F 上的**群代数** (group algebra). 于是群 $(G, p, 1)$ 的线性表示 (ρ, V) 也可线性延拓为群代数 $F[G]$ 的线性表示 $(\widetilde{\rho}, V)$, 即存在结合代数同态 $F[G] \longrightarrow \mathrm{End}(V)$.

$$\sum_{g\in G} c_g g \longmapsto \sum_{g\in G} c_g \rho(g)$$

(2) 注意群同态 $\rho \colon (G, p, 1) \to \mathrm{GL}(V)$ 与结合代数同态 $\widetilde{\rho} \colon F[G] \to \mathrm{End}(V)$ 是一一对应的: 前者的线性延拓即为后者, 后者限制在群上即为前者. 但这两者的部分性质可能不同, 例如 $\widetilde{\rho}$ 为单射显然推出 ρ 为单射, 但 ρ 为单射无法推出 $\widetilde{\rho}$ 为单射. 反例如本节开头, $\rho \colon (S_n, \circ, \mathrm{id}) \to \mathrm{GL}(V)$ 总是单射, 但当 $n \geq 4$ 时, $\dim_F(F[S_n]) = |S_n| = n! > n^2 = \dim_F(\mathrm{End}_F(V))$, 故 $\widetilde{\rho} \colon F[S_n] \to \mathrm{End}_F(V)$ 不可能为单射.

定义 4.1.5 (不可约表示) 设 (ρ, V) 为群 $(G, p, 1)$ 的一个线性表示, 若存在线性子空间 $W \subseteq V$, 满足 $\forall g \in G$, $\rho(g)(W) \subseteq W$, 则称 (ρ, W) 为 (ρ, V) 的一个 **子表示** (subrepresentation). 不含非零真子表示的线性表示称为 **不可约的** (irreducible).

例 4.1.2 在本节开头的例子中, 记 $W_1 = \mathrm{Span}_F(\{\sum_{i=1}^{n} \alpha_i\})$, 则 $\forall \sigma \in S_n$, $T_\sigma\big|_{W_1} = \mathrm{id}_{W_1}$, 故 $W_1 \subseteq V$ 是 $(S_n, \circ, \mathrm{id})$ 的一维平凡表示. 记 $W_2 = \left\{ \sum_{i=1}^{n} c_i \alpha_i \in V \colon \sum_{i=1}^{n} c_i = 0 \right\}$, 则 $W_2 \subseteq V$ 是 $(S_n, \circ, \mathrm{id})$ 的 $(n-1)$ 维 (标准) 表示.

当 $\mathrm{char}(F) \nmid n$ 时, $V = W_1 \oplus W_2$ 为不可约子表示的直和. (这里 W_2 的不可约性是因为: 若 $n = 1$, 则 $W_2 = \{0\}$ 显然不可约; 现设 $n \geq 2$, 且 $\{0\} \neq U \subseteq W_2$ 为子表示. 由 $U \neq \{0\}$ 且 $W_1 \cap U = \{0\}$ 知, $\exists \alpha = \sum_{i=1}^{n} c_i \alpha_i \in U$ 满足 $\exists 1 \leq i < j \leq n$, $s.t.\ c_i \neq c_j$. 取 $\sigma \in S_n$ 满足 $\sigma(i) = 1$, $\sigma(j) = 2$, 记 $\alpha' = T_\sigma(\alpha) \in U$, 则 $\alpha' = \sum_{i=1}^{n} c_i' \alpha_i \in U$ 满足 $c_1' \neq c_2'$. 再取 $\tau = (1, 2) \in S_n$, 则 $U \ni \alpha' - T_\tau(\alpha') = (c_1' - c_2')(\alpha_1 - \alpha_2)$, 故 $\alpha_1 - \alpha_2 \in U$. 由轮换 $(1, 2, \cdots, n)$ 不断作用知, $\forall 1 \leq i \leq n-1$, $\alpha_i - \alpha_{i+1} \in U$, 故 $W_2 = \mathrm{Span}_F(\{\alpha_i - \alpha_{i+1}\}_{i=1}^{n-1}) \subseteq U$.)

当 $\mathrm{char}(F) \mid n$ 时, $W_1 \subseteq W_2$ 为子表示. 当 $n = 2$ 时, $W_1 = W_2$ 是不可约的. 当 $n \geq 3$ 时, W_2 是可约的; 但完全同理知, 不存在子表示 $U \subseteq W_2$ 满足 $W_1 \oplus U = W_2$.

注: 此例表明域特征与线性表示的可约性、可分解性密切相关. 对于有限群 $(G, p, 1)$, 这恰是常 (ordinary) 表示论 (即 $\mathrm{char}(F) \nmid |G|$) 与模 (modular) 表示论 (即 $\mathrm{char}(F) \mid |G|$) 的差异. 有限群的常表示论中最基本的结果是以下的 Maschke 定理.

定理 4.1.6 (Maschke) *设 $(G, p, 1)$ 为有限群, F 为一个域, 则以下条件等价:*

(1) $\mathrm{char}(F) \nmid |G|$;

(2) $F[G]$ *为域* F *上的半单结合代数;*

(3) *任取* $(G, p, 1)$ *的线性表示* (ρ, V), *则它的任意子表示* (ρ, W) *均存在直和补表示* (ρ, U);

(4) $(G, p, 1)$ *的任意线性表示都可分解为一族不可约子表示的直和.*

证明: "(1)⇒(2)": 设 $\mathrm{char}(F) \nmid |G|$, $M \subseteq F[G]$ 为一个 $F[G]$-子模, 则可取 $\pi \colon F[G] \twoheadrightarrow M$ 为 F-线性的到 M 上投影映射. 考虑映射 π 关于 G 作用的平均化 $\tilde{\pi} \colon F[G] \longrightarrow M$, 则 $\tilde{\pi}$ 仍为 F-线性的到 M 上投

$$x \longmapsto \frac{1}{|G|} \sum_{g \in G} g.\pi(g^{-1}.x)$$

影映射, 且与 G 的作用可交换, 故 $\tilde{\pi}$ 是 $F[G]$-模同态. 由可裂性引理知, $F[G] = M \oplus \ker(\tilde{\pi})$, 其中 $\ker(\tilde{\pi}) \subseteq F[G]$ 也是一个 $F[G]$-子模. 因此 $F[G]$ 为半单结合代数.

"(2)⇒(3)": 设 $F[G]$ 为半单结合代数, 即 $F[G]$ 自身为半单 $F[G]$-模, 则任意自由 $F[G]$-模都是半单 $F[G]$-模. 现任取一个 $F[G]$-模 V, 则 V 同构于某个自由 $F[G]$-模的商模; 又由于半单模的商模仍为半单模, 则 V 也为半单 $F[G]$-模, 即作为群 $(G, p, 1)$ 的表示空间, V 的任意子表示都存在直和补表示.

"(3)⇔(4)": 这是模论中的经典结果: 一个 R-模的任意子模均存在直和补子模 $\Longleftrightarrow$ 该 R-模可分解成不可约 R-子模的直和. 此证明是反复运用 Zorn 引理, 可见 J. J. Rotman"An Introduction to Homological Algebra".

"(3)⇒(2)": 考虑 $(G, p, 1)$ 的正则表示 $(\rho_{\mathrm{reg}}, F[G])$, 即 $\rho_{\mathrm{reg}} \colon (G, p, 1) \longrightarrow \mathrm{GL}(F[G])$. 由

$$g \longmapsto \left(\rho_{\mathrm{reg}} \colon \sum_{h \in G} c_h h \mapsto \sum_{h \in G} c_h p(g, h) \right)$$

条件 $(\rho_{\mathrm{reg}}, F[G])$ 的任意子表示都存在直和补表示, 即 $F[G]$ 的任意 $F[G]$-子模均存在直和补 $F[G]$-子模, 则 $F[G]$ 为半单结合代数.

"(2)⇒(1)": 考虑 F-线性函数 $f \colon F[G] \longrightarrow F$, 则 $K \coloneqq \ker(f) \subseteq F[G]$ 为 $F[G]$-子模. 假设 $\mathrm{char}(F) \mid |G|$.

$$\sum_{g \in G} c_g g \longmapsto \sum_{g \in G} c_g$$

断言: 任取 $\{0\} \neq M \subseteq F[G]$ 为 $F[G]$-子模, $M \cap K \neq \{0\}$. (这是因为, 由 $M \neq \{0\}$ 知, 可取 $0 \neq \alpha = \sum_{g \in G} c_g g \in M$. 若 $\alpha \in K$, 则结论显然. 若 $\alpha \notin K$, 则记 $\beta = \sum_{h \in G} h \in F[G]$, 由 $f(\beta) = \sum_{h \in G} 1 = |G| = 0$ 知, $\beta \in K$. 一方面, 注意

$\beta\cdot\alpha=\sum\limits_{h\in G}h.\alpha\in M$; 另一方面, $\beta\cdot\alpha=\left(\sum\limits_{h\in G}h\right)\cdot\left(\sum\limits_{g\in G}c_gg\right)=\sum\limits_{h\in G}\left(\sum\limits_{g\in G}c_gp(h,g)\right)=\sum\limits_{h\in G}\left(\sum\limits_{g\in G}c_{p(h^{-1},g)}g\right)$
$=\sum\limits_{g\in G}\left(\sum\limits_{h\in G}c_{p(h^{-1},g)}\right)g=\sum\limits_{g\in G}\left(\sum\limits_{h\in G}c_h\right)g=f(\alpha)\beta\in K\setminus\{0\}$. 因此 $M\cap K\neq\{0\}$.) 于是 $F[G]$ 不为半单结合代数. □

注:

(1) 上述有限群的线性表示的不可约分解一般不唯一. 例如设群的作用是平凡的, 则线性空间的不可约分解是它的一维子空间的直和, 这与基的选取有关.

(2) 当不加拓扑条件时, 上述 Maschke 定理一般无法推广. 这是因为, 设 $(G,p,1)$ 为无限群, R 为一个环, 则群环 $R[G]$ 一定不是半单环.

(3) 当加拓扑条件时, 上述 Maschke 定理可推广为抽象调和分析中的 Peter-Weyl 定理如下: "设 $(G,p,1)$ 为紧群, 则它在任意复 Hilbert 空间上的酉表示都可分解为一族有限维不可约酉表示的正交直和."

以下简单给出群的不可约表示与正则表示的关系. 当然, 利用特征标理论可以得到更精确的表述.

命题 4.1.7 *设 $(G,p,1)$ 为一个群, F 为一个域,*

(1) *$(G,p,1)$ 的任意不可约表示都同构于正则表示的某个商表示;*

(2) *若 $|G|<+\infty$ 且 $\mathrm{char}(F)\nmid|G|$, 则 $(G,p,1)$ 的任意不可约表示都同构于正则表示的某个子表示.*

证明: (1) 设 (ρ,V) 为群 $(G,p,1)$ 的一个不可约表示, 若 $V=\{0\}$, 则结论显然. 现设 $V\neq\{0\}$, 取 $\alpha\in V\setminus\{0\}$, 考虑 $F[G]$-模同态 φ: $F[G]\longrightarrow V$, 则 $\{0\}\neq\mathrm{Im}(\varphi)\subseteq V$ 是一个 $F[G]$-子模. 由 (ρ,V) 的不可约

$$\sum_{g\in G}c_gg\longmapsto\sum_{g\in G}c_g\rho(g)(\alpha)$$

性知, $\mathrm{Im}(\varphi)=V$, 故上述诱导了 $F[G]$-模同构 $F[G]/\ker(\varphi)\xrightarrow{\cong}V$.
(2) 由 Maschke 定理知, 若 $|G|<+\infty$ 且 $\mathrm{char}(F)\nmid|G|$, 则 $F[G]$ 为半单结合代数, 故 $F[G]$ 的任意商模都同构于 $F[G]$ 的某个子模. 因此由 (1) 即知结论. □

最后给出群的线性表示的构造方法, 包括直和、张量、商、对偶与 Hom.

定义 4.1.6 (直和表示) 设 $(G,p,1)$ 为一个群, (ρ_1,V_1), (ρ_2,V_2) 为群 $(G,p,1)$ 的两个线性表示, 记群同态
$\rho_1\oplus\rho_2$: $(G,p,1)\longrightarrow\mathrm{GL}(V_1\oplus V_2)$, 则称 $(\rho_1\oplus\rho_2,V_1\oplus V_2)$ 为线性表示
$g\longmapsto(\rho_1(g)\oplus\rho_2(g)\colon(\alpha_1,\alpha_2)\mapsto(\rho_1(g)(\alpha_1),\rho_2(g)(\alpha_2)))$
(ρ_1,V_1), (ρ_2,V_2) 的直和.

注: 当 $\dim_F(V_1),\dim_F(V_2)<+\infty$ 时, 取 V_1 的基 B_1, V_2 的基 B_2, 则 $\forall g\in G$, $(\rho_1\oplus\rho_2)(g)$ 在基 $B_1\sqcup B_2$ 下的矩阵表示为 $\begin{pmatrix}[\rho_1(g)]_{B_1}&0\\0&[\rho_2(g)]_{B_2}\end{pmatrix}$.

定义 4.1.7 (张量表示) 设 $(G,p,1)$ 为一个群, (ρ_1,V_1), (ρ_2,V_2) 为群 $(G,p,1)$ 的两个线性表示, 记群同态
$\rho_1\otimes\rho_2$: $(G,p,1)\longrightarrow\mathrm{GL}(V_1\underset{F}{\otimes}V_2)$, 则称 $(\rho_1\otimes\rho_2,V_1\underset{F}{\otimes}V_2)$ 为线性表示
$g\longmapsto(\rho_1(g)\otimes\rho_2(g)\colon\alpha_1\otimes\alpha_2\mapsto\rho_1(g)(\alpha_1)\otimes\rho_2(g)(\alpha_2))$
(ρ_1,V_1), (ρ_2,V_2) 的张量.

注: 当 $\dim_F(V_1),\dim_F(V_2)<+\infty$ 时, 取 V_1 的基 B_1, V_2 的基 B_2, 则 $\forall g\in G$, $(\rho_1\otimes\rho_2)(g)$ 在基 $B_1\otimes B_2$ 下的矩阵表示为 $[\rho_1(g)]_{B_1}\otimes[\rho_2(g)]_{B_2}$.

定义 4.1.8 (商表示) 设 $(G,p,1)$ 为一个群, (ρ,V) 为群 $(G,p,1)$ 的一个线性表示, (ρ,W) 为 (ρ,V) 的一个子表示, 记群同态 $\widetilde{\rho}$: $(G,p,1)\longrightarrow\mathrm{GL}(V/W)$, 则称 $(\widetilde{\rho},V/W)$ 为线性表示 (ρ,V) 关于子表示
$g\longmapsto(\widetilde{\rho}(g)\colon\alpha+W\mapsto\rho(\alpha)+W)$
(ρ,W) 的商.

注: 当 $\dim_F(V) < +\infty$ 时, 取 W 的基 B_0, 延拓为 V 的基 B, 则 $\forall g \in G$, $\widetilde{\rho}(g)$ 在基 $B \setminus B_0$ 下的矩阵表示由以下确定:

$$[\rho(g)]_B = \begin{pmatrix} [\rho(g)]_{B_0} & * \\ 0 & [\widetilde{\rho}(g)]_{B\setminus B_0} \end{pmatrix}.$$

定义 4.1.9 (对偶表示) 设 $(G,p,1)$ 为一个群, (ρ,V) 为群 $(G,p,1)$ 的一个线性表示, 记群同态

$$\begin{aligned} \rho^*\colon\ (G,p,1) &\longrightarrow \mathrm{GL}(V^*) \\ g &\longmapsto (\rho^*(g)\colon f \mapsto f\circ\rho(g^{-1})) \end{aligned}$$

, 则称 (ρ^*,V^*) 为线性表示 (ρ,V) 的对偶.

注: 当 $\dim_F(V) < +\infty$ 时, 取 V 的基 B, 以及它的对偶基 B^*, 则 $\forall g \in G$, $\rho^*(g)$ 在基 B^* 下的矩阵表示为 $([\rho(g)]_B^{-1})^t$.

定义 4.1.10 (Hom 表示) 设 $(G,p,1)$ 为一个群, (ρ_1,V_1), (ρ_2,V_2) 为群 $(G,p,1)$ 的两个线性表示, 记群同态

$$\begin{aligned} \mathrm{Hom}_F(\rho_1,\rho_2)\colon\ (G,p,1) &\longrightarrow \mathrm{GL}(\mathrm{Hom}_F(V_1,V_2)) \\ g &\longmapsto (\mathrm{Hom}_F(\rho_1,\rho_2)(g)\colon T \mapsto \rho_2(g)\circ T\circ\rho_1(g^{-1})) \end{aligned}$$

, 则称 $(\mathrm{Hom}_F(\rho_1,\rho_2),\mathrm{Hom}_F(V_1,V_2))$ 为线性表示 (ρ_1,V_1), (ρ_2,V_2) 的 Hom.

注: 当 $\dim_F(V_1),\dim_F(V_2) < +\infty$ 时, 取 V_1 的基 B_1, V_2 的基 B_2, 则 $\forall g \in G$, $\mathrm{Hom}_F(\rho_1,\rho_2)(g)$ 在基 $B_1^*\otimes B_2$ 下的矩阵表示为 $([\rho_1(g)]_{B_1}^{-1})^t \otimes [\rho_2(g)]_{B_2}$. 此时存在表示的同构 $(\mathrm{Hom}_F(\rho_1,\rho_2),\mathrm{Hom}_F(V_1,V_2)) \cong (\rho_1^*,V_1^*)\otimes(\rho_2,V_2)$.

例 4.1.3 在本节开头的例子中, 取 $B = \{\alpha_i\}_{i=1}^n$ 的对偶基 $B^* = \{f_{\alpha_i}\}_{i=1}^n$, 考虑其对偶表示

$$\begin{aligned} (S_n,\circ,\mathrm{id}) &\longrightarrow \mathrm{GL}(V^*) \\ \sigma &\longmapsto \left(T^t_{\sigma^{-1}}\colon \sum_{i=1}^n c_i f_{\alpha_i} \mapsto \sum_{i=1}^n c_i f_{\sigma(\alpha_i)}\right) \end{aligned}$$

, 这里 $T^t_{\sigma^{-1}}$ 在基 B^* 下的矩阵表示仍为 R_σ, 故 $(R_\sigma^{-1})^t = R_\sigma$.

习题 4.1

(1) 证明: $\mathrm{Aut}(A_n,\circ,\mathrm{id}) \cong \begin{cases} (S_n,\circ,\mathrm{id}), & n \neq 1,2,3,6 \\ (\mathcal{C}_1,\cdot,1), & n = 1,2 \\ (\mathcal{C}_2,\cdot,1), & n = 3 \\ (S_6,\circ,\mathrm{id}) \rtimes (\mathcal{C}_2,\cdot,1), & n = 6 \end{cases}$.

(2) 证明: $S_1 = \{\mathrm{id}\}'$; $S_2 \cong Q_8'$ (这里 Q_8 为四元数群(quaternion group)); 当 $n \geq 3$ 时, S_n 不可能同构于任意群的导群.

参考文献与补注 4.1

(1) 关于有限单群的分类部分, 可以参考维基百科或者 Aschbacher, Michae"The Status of the Classification of the Finite Simple Groups" (2004).
(2) 关于特征单群的部分, 可以参考 J. S. Robinson"A Course in the Theory of Groups".
(3) 关于辫群的部分, 可以参考 J. Wilson "The geometry and topology of braid groups".
(4) 关于有限群的线性表示的部分, 可以参考 J. P. Serre"Linear Representations of Finite Groups".
(5) 关于抽象调和分析的部分, 可以参考 G. B. Folland"A Course in Abstract Harmonic Analysis".

§4.2 外积与体积形式

4.2.1 从张量积到外积

我们回忆线性空间的张量积的构造, 并引入张量代数的记号: 设 V 为域 F 上的线性空间, 记 $T^0(V) = F$, 归纳定义 $T^r(V) = T^{r-1}(V) \underset{F}{\otimes} V$ $(r \in \mathbb{N}^*)$, 以及 $T(V) \coloneqq \coprod\limits_{r\in\mathbb{N}} T^r(V)$. 这里 $T(V)$ 为域 F 上的线性空间, 且可引入

乘法:

$$\begin{aligned} T^r(V)\times T^s(V)&\longrightarrow T^{r+s}(V)\quad (r,s\in\mathbb{N}) \qquad ,\\ (\alpha_1\otimes\cdots\otimes\alpha_r,\ \alpha_{r+1}\otimes\cdots\otimes\alpha_s)&\longmapsto \alpha_1\otimes\cdots\otimes\alpha_r\otimes\alpha_{r+1}\otimes\cdots\otimes\alpha_s \end{aligned}$$

此时 $T(V)$ 为一个分次的 F-结合代数, 称为线性空间 V 的**张量代数** (tensor algebra). 它有时也被称为线性空间 V 上的 **自由代数** (free algebra). 这是因为 F-结合代数 $T(V)$ 与含入映射 $i\colon V\hookrightarrow T(V)$ 满足以下的泛性质: 任给 F-结合代数 A 与线性映射 $T\colon V\to A$, 存在唯一的 F-结合代数同态 $\widetilde{T}\colon T(V)\to A$, 满足 $T=\widetilde{T}\circ i$. 换句话说,

记 $\mathscr{G}\colon$ **F-Alg** $\to$ **F-Mod** 为忘记函子, 则存在一对伴随函子 $T(\cdot)\colon$ **F-Mod** $\rightleftarrows$ **F-Alg**$\colon\mathscr{G}$. (请务必比较这里与第三章开头的伴随函子的区别!)

考虑张量代数 $T(V)$ 中由 "平方元"$\alpha\otimes\alpha\,(\alpha\in V)$ 生成的双边理想 $I(V)$. 注意 $I(V)$ 为分次理想: $I(V)=\coprod\limits_{r\in\mathbb{N}}I^r(V)$, 其中 $I^r(V)=I(V)\cap T^r(V)\,(r\in\mathbb{N})$. 于是 $T(V)$ 关于 $I(V)$ 的商代数 $\bigwedge(V):=T(V)/I(V)$ 也为分次结合代数: $\bigwedge(V)=\coprod\limits_{r\in\mathbb{N}}\bigwedge^r(V)$, 其中 $\bigwedge^r(V)=T^r(V)/I^r(V)\,(r\in\mathbb{N})$. 此时 $\bigwedge(V)$ 称为线性空间 V 的**外代数** (exterior algebra). 在商映射 $T(V)\twoheadrightarrow\bigwedge(V)$ 下, 纯张量 $\alpha_1\otimes\cdots\otimes\alpha_r$ 的像记为 $\alpha_1\wedge\cdots\wedge\alpha_r$, 张量积诱导了**外积** (exterior product)

$$\begin{aligned} \textstyle\bigwedge^r(V)\times\bigwedge^s(V)&\longrightarrow\textstyle\bigwedge^{r+s}(V)\quad (r,s\in\mathbb{N}) \qquad .\\ (\alpha_1\wedge\cdots\wedge\alpha_r,\ \alpha_{r+1}\wedge\cdots\wedge\alpha_{r+s})&\longmapsto \alpha_1\wedge\cdots\wedge\alpha_r\wedge\alpha_{r+1}\wedge\cdots\wedge\alpha_{r+s} \end{aligned}$$

由构造知外积满足多线性性、结合性与交错性; 外代数 $\bigwedge(V)$ 与含入映射 $i\colon V\hookrightarrow\bigwedge(V)$ 满足以下的泛性质: 任给 F-结合代数 A 与平方为零的线性映射 $T\colon V\to A$, 存在唯一的 F-结合代数同态 $\widetilde{T}\colon\bigwedge(V)\to A$, 满足 $T=\widetilde{T}\circ i$.

接着考虑张量代数 $T(V)$ 中的 "反对称元": 任取 $r\in\mathbb{N}$, 对于 $\sigma\in S_r$ 以及纯张量 $\alpha_1\otimes\cdots\otimes\alpha_r\in T^r(V)$, 记 $\sigma(\alpha_1\otimes\cdots\otimes\alpha_r):=\alpha_{\sigma(1)}\otimes\cdots\otimes\alpha_{\sigma(r)}\in T^r(V)$, 则可线性延拓为群表示 $(S_r,\circ,\mathrm{id})\to\mathrm{GL}(T^r(V))$. 于是 $A^r(V):=\{u\in T^r(V)\colon\sigma(u)=\mathrm{sgn}(\sigma)u,\ \forall\sigma\in S_r\}$ 为它的一个子表示. 记 $A(V)=\coprod\limits_{r\in\mathbb{N}}A^r(V)$, 这是域 F 上的线性空间, 且可引入乘法:

$$\begin{aligned} \widehat{\otimes}\colon\ A^r(V)\times A^s(V)&\longrightarrow A^{r+s}(V)\quad (r,s\in\mathbb{N}) \qquad ,\\ (u_1,\ u_2)&\longmapsto\sum_{\sigma\in\mathrm{Sh(r,s)}}\mathrm{sgn}(\sigma)\sigma(u_1\otimes u_2) \end{aligned}$$

此时 $A(V)$ 为一个分次的 F-结合代数, 称为线性空间 V 的 **反对称张量代数** (antisymmetric tensor algebra). (注意 $A(V)\subseteq T(V)$ 只是子空间, 而非子代数!)

记张量代数 $T(V)$ 上的**反称化** (antisymmetrization) 算子为 $\mathrm{Ant}^{(r)}=\sum\limits_{\sigma\in S_r}\mathrm{sgn}(\sigma)\sigma\in L(T^r(V))\,(r\in\mathbb{N})$, 以下试图用 $\mathrm{Ant}^{(r)}$ 将 $T^r(V),I^r(V),\Lambda^r(V),A^r(V)$ 四者联系起来.

命题 4.2.1 *设 V 为域 F 上的线性空间, $r\in\mathbb{N}$, 则:*

(1) $\ker(\mathrm{Ant}^{(r)})\supseteq I^r(V)$; $\mathrm{Im}(\mathrm{Ant}^{(r)})\subseteq A^r(V)$; $\mathrm{Ant}^{(\mathrm{r})}\circ\mathrm{Ant}^{(\mathrm{r})}=r!\mathrm{Ant}^{(\mathrm{r})}$;

(2) *当 $\mathrm{char}(F)\nmid r!$ 时, (1) 中取 "$=$", 则 $\dfrac{1}{r!}\mathrm{Ant}^{(\mathrm{r})}$ 可视为到 $A^r(V)$ 上的投影算子, 故存在线性空间的直和分解*

$T^r(V)=I^r(V)\oplus A^r(V)$;

(3) *当 $\mathrm{char}(F)=0$ 时, 存在结合代数同构 $A(V)\cong\bigwedge(V)$.*

证明: (1) 显然;

(2) 先断言: $\forall u\in T^r(V)$, $\forall\sigma\in S_r$, $\mathrm{sgn}(\sigma)\sigma(u)-u\in I^r(V)$. (这是因为, 由线性性可不妨设 $u=\alpha_1\otimes\cdots\otimes\alpha_r$. 以下对于 $\sigma\in S_r$ 可写成相邻对换的复合的最少个数归纳: 当 $\sigma=\mathrm{id}$ 时, 结论显然; 当 $\sigma=(i,i+1)\,(1\le i\le r-1)$ 时,

$$\begin{aligned} \mathrm{sgn}(\sigma)\sigma(u)-u&=-\alpha_1\otimes\cdots\otimes\alpha_{i+1}\otimes\alpha_i\otimes\cdots\otimes\alpha_r-\alpha_1\otimes\cdots\otimes\alpha_i\otimes\alpha_{i+1}\otimes\cdots\otimes\alpha_r\\ &=-\alpha_1\otimes\cdots\otimes(\alpha_i+\alpha_{i+1})\otimes(\alpha_i+\alpha_{i+1})\otimes\cdots\otimes\alpha_r\\ &\quad+\alpha_1\otimes\cdots\otimes\alpha_i\otimes\alpha_i\otimes\cdots\otimes\alpha_r+\alpha_1\otimes\cdots\otimes\alpha_{i+1}\otimes\alpha_{i+1}\otimes\cdots\otimes\alpha_r\\ &\quad\in I^r(V). \end{aligned}$$

现设对于任意 $k\ge1$ 个相邻对换的复合 $\sigma\in S_r$, $\mathrm{sgn}(\sigma)\sigma(u)-u\in I^r(V)$ 都成立, 则对于任意 $(k+1)$ 个相邻对

换的复合 $\sigma' \in S_r$, 记 $\sigma' = (i, i+1) \circ \sigma$, 其中 $1 \leq i \leq r-1$, $\sigma \in S_r$ 为 k 个相邻对换的复合. 由归纳假设知 $\mathrm{sgn}(\sigma)\sigma(u) - u \in I^r(V)$; 再由 $I^r(V)$ 在 S_r 作用下为不变子空间知, $\mathrm{sgn}(\sigma)\sigma'(u) - (i, i+1)(u) \in I^r(V)$. 因此 $\mathrm{sgn}(\sigma')(u) - u = -(\mathrm{sgn}(\sigma)\sigma'(u) - (i, i+1)(u)) + (-(i, i+1)(u) - u) \in I^r(V)$.)

再注意 $\forall u \in T^r(V)$, $\mathrm{Ant}^{(r)}(u) - r!u = \sum\limits_{\sigma \in S_r} (\mathrm{sgn}(\sigma)\sigma(u) - u)$ (这是因为: 由线性性可不妨设 u 为纯张量, 从而直接验证). 当 $\mathrm{char}(F) \nmid r!$ 时, $\forall u \in \ker(\mathrm{Ant}^{(r)})$, $u = -\dfrac{1}{r!} \sum\limits_{\sigma \in S_r} (\mathrm{sgn}(\sigma)\sigma(u) - u) \in I^r(V)$, 即 $\ker(\mathrm{Ant}^{(r)}) \subseteq I^r(V)$. 另外, 由于 $\forall u \in A^r(V)$, $\forall \sigma \in S_r$, $\sigma(u) = \mathrm{sgn}(\sigma)u$, 则 $u = \dfrac{1}{r!}\mathrm{Ant}^{(r)}(u) \in \mathrm{Im}(\mathrm{Ant}^{(r)})$, 即 $A^r(V) \subseteq \mathrm{Im}(\mathrm{Ant}^{(r)})$. 于是此时 $\dfrac{1}{r!}\mathrm{Ant}^{(\mathrm{r})}$ 为到 $A^r(V)$ 上的投影算子, 故给出了线性空间分解 $T^r(V) = I^r(V) \oplus A^r(V)$. 特别地, 存在线性空间同构 $A^r(V) \cong T^r(V)/I^r(V) = \bigwedge^r(V)$.

(3) 只需证明上述同构保持结合代数的乘法, 即证 $\forall u_1 \in T^r(V)$, $u_2 \in T^s(V)$, $\dfrac{1}{r!s!}\mathrm{Ant}^{(r+s)}(\mathrm{Ant}^{(r)}(u_1) \otimes \mathrm{Ant}^{(s)}(u_2))$ $= \mathrm{Ant}^{(r+s)}(u_1 \otimes u_2)$. 这是定义的直接验证. □

在安师讲义中, 记 $M^r(V) \coloneqq \{L \colon V^r \to F \text{为多重线性函数}\}$, $\Lambda^r(V) \coloneqq \{L \in M^r(V) \colon L \text{为交错的}\}$, 以及**交错化** (alternation) 算子 $\mathrm{Alt}^{(r)} \colon M^r(V) \longrightarrow M^r(V)$ $(r \in \mathbb{N})$. 我们

$$L \longmapsto (\mathrm{Alt}^{(r)}(L) \colon (\alpha_1, \cdots, \alpha_r) \mapsto \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) L(\alpha_{\sigma(1)}, \cdots, \alpha_{\sigma(r)}))$$

先从泛性质的角度说明这三者与前述概念的联系.

命题 4.2.2 *设 V 为域 F 上的线性空间, $r \in \mathbb{N}$, 则*

(1) *存在线性同构 $(T^r(V))^* \cong M^r(V)$; $(\bigwedge^r(V))^* \cong (I^r(V))^0 \cong \Lambda^r(V)$; 在此同构意义下, $(\mathrm{Ant}^{(\mathrm{r})})^t = \mathrm{Alt}^{(\mathrm{r})}$;*

(2) *当 $\mathrm{char}(F) \nmid r!$ 时, $\ker(\mathrm{Alt}^{(r)}) = (A^r(V))^0$; $\mathrm{Im}(\mathrm{Alt}^{(r)}) = \Lambda^r(V)$; 则 $\dfrac{1}{r!}\mathrm{Alt}^{(\mathrm{r})}$ 可视为到 $\Lambda^r(V)$ 上的投影算子, 故存在线性空间的直和分解 $M^r(V) = (A^r(V))^0 \oplus \Lambda^r(V)$.*

证明: (1) 由线性空间的张量积的泛性质即知线性同构 $(T^r(V))^* \cong M^r(V)$; 再由外积的定义知 $(I^r(V))^0 \cong \Lambda^r(V)$; 最后由商的泛性质知 $(\bigwedge^r(V))^* = (T^r(V)/I^r(V))^* \cong (I^r(V))^0$.

(2) 由上述命题与 (1) 即知. □

进一步地, 安师讲义对于多重线性函数定义了张量积

$$\begin{aligned} M^r(V) \times M^s(V) &\longrightarrow M^{r+s}(V) \quad (r, s \in \mathbb{N}) \quad , \\ (L_1, L_2) &\longmapsto (L_1 \otimes L_2 \colon (\alpha_1, \cdots, \alpha_{r+s}) \mapsto L_1(\alpha_1, \cdots, \alpha_r) L_2(\alpha_{r+1}, \cdots, \alpha_{r+s})) \end{aligned}$$

对于多重交错线性函数定义了外积

$$\begin{aligned} \Lambda^r(V) \times \Lambda^s(V) &\longrightarrow \Lambda^{r+s}(V) \quad (r, s \in \mathbb{N}) \quad . \\ (L_1, L_2) &\longmapsto (L_1 \wedge L_2 \colon (\alpha_1, \cdots, \alpha_{r+s}) \mapsto \sum_{\sigma \in \mathrm{Sh}(r,s)} \mathrm{sgn}(\sigma)(L_1 \otimes L_2)(\alpha_{\sigma(1)}, \cdots, \alpha_{\sigma(r+s)})) \end{aligned}$$

我们再从等变映射的角度说明这两者与前述概念的联系.

命题 4.2.3 *设 V 为域 F 上的线性空间, $r \in \mathbb{N}$, 则:*

(1) *存在线性空间正合列的交换图* $\begin{array}{ccccccccc} 0 & \longrightarrow & I^r(V^*) & \longrightarrow & T^r(V^*) & \longrightarrow & \bigwedge^r(V^*) & \longrightarrow & 0 \ ; \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \ker(\mathrm{Alt}^{(r)}) & \longrightarrow & M^r(V) & \xrightarrow{\mathrm{Alt}^{(r)}} & \Lambda^r(V) & \longrightarrow & 0 \end{array}$

(2) *当 $\mathrm{char}(F) \nmid r!$ 时, 上图右端的竖直映射也为单射;*

(3) *当 $\mathrm{char}(F) = 0$ 时, 存在分次 F-结合代数的交换图* $\begin{array}{ccc} T(V^*) & \twoheadrightarrow & \bigwedge(V^*) \\ \downarrow & & \downarrow \\ M(V) & \xrightarrow{\mathrm{Alt}} & \Lambda(V) \end{array}$.

证明: (1) 注意这里张量积符号 $\otimes$ 具有两种含义: 一是上一章抽象定义的 $(V^*)^r \longrightarrow T^r(V^*)$; 二是

$$(f_1, \cdots, f_r) \longmapsto f_1 \otimes \cdots \otimes f_r$$

安师讲义中具体定义的 $(V^*)^r \longrightarrow M^r(V)$. 为避免混淆, 我们暂时将后者记为 $\otimes''$. 由线性空间的张

$$(f_1, \cdots, f_r) \longmapsto f_1 \otimes \cdots \otimes f_r$$

量积的泛性质, 后者诱导了线性映射 S: $T^r(V^*) \longrightarrow M^r(V)$, 断言: S 为单射.

$$f_1 \otimes \cdots \otimes f_r \longmapsto f_1 \otimes'' \cdots \otimes'' f_r$$

这是因为, 注意 $\mathrm{Im}(S) = \mathrm{Span}_F(\{f_1 \otimes'' \cdots \otimes'' f_r \colon f_i \in V^*\})$, 取 V^* 的基 $\{f_i\}_{i\in I}$, 先说明: $B = \{f_{i_1} \otimes'' \cdots \otimes'' f_{i_r} \colon i_1, \cdots, i_r \in I\}$ 为 $\mathrm{Im}(S)$ 的基. 显然 B 可线性生成 $\mathrm{Im}(S)$; 现证 B 线性无关: 任取 B 的有限非空子集, 考虑其线性组合 $c_1 f_{i_{1,1}} \otimes'' \cdots \otimes'' f_{i_{1,r}} + \cdots + c_k f_{i_{k,1}} \otimes'' \cdots \otimes'' f_{i_{k,r}} = 0$, 其中 $(i_{1,1}, \cdots, i_{1,r}), \cdots, (i_{k,1}, \cdots, i_{k,r})$ 为两两不同的 r 元组. 将 $\{f_{i_{1,1}}, \cdots, f_{i_{1,r}}, \cdots, f_{i_{k,1}}, \cdots, f_{i_{k,r}}\}$ 去掉重复项后记为 $\{g_1, \cdots, g_l\} \subseteq V^*$, 则后者为线性无关集. 可取 $\{\beta_1, \cdots, \beta_l\} \subseteq V$, 满足 $g_i(\beta_j) = \delta_{ij}$, $\forall 1 \leq i, j \leq l$, 再按照之前的序号对应, 将 $\{\beta_1, \cdots, \beta_l\}$ 扩充重复项后记为 $\{\alpha_{i_{1,1}}, \cdots, \alpha_{i_{1,r}}, \cdots, \alpha_{i_{k,1}}, \cdots, \alpha_{i_{k,r}}\}$. 于是 $(f_{i_{j,1}} \otimes'' \cdots \otimes'' f_{i_{j,r}})(\alpha_{i_{j',1}}, \cdots, \alpha_{i_{j',r}}) = \delta_{j,j'}$, $\forall 1 \leq j, j' \leq k$, 由此知 $c_j = 0$, $\forall 1 \leq j \leq k$, 故 $\{f_{i_{1,1}} \otimes'' \cdots \otimes'' f_{i_{1,r}}, \cdots, f_{i_{k,1}} \otimes'' \cdots \otimes'' f_{i_{k,r}}\}$ 为线性无关集, 因此 B 为线性无关集. 此时对应 $f_{i_1} \otimes'' \cdots \otimes'' f_{i_r} \mapsto f_{i_1} \otimes \cdots \otimes f_{i_r}$ $(i_1, \cdots, i_r \in I)$ 可延拓为线性映射 $\mathrm{Im}(S) \to T^r(V^*)$, 再通过取 $\mathrm{Im}(S)$ 在 $M^r(V)$ 中的直和补空间, 可延拓为线性映射 $M^r(V) \to T^r(V^*)$, 此即 S 的左逆, 故 S 为单射.

于是我们将上图中间的竖直映射定义为 S. 断言: $S(I^r(V^*)) \subseteq \ker(\mathrm{Alt}^{(r)})$, 从而中间的竖直映射 S 可限制在左端为单射, 也可诱导右端的合理映射. 这是因为, 由于 $I^r(V^*) = \mathrm{Span}_F(\{f_1 \otimes \cdots \otimes f_r \colon f_i \in V^*;\ \exists 1 \leq i < j \leq r,\ s.t.\ f_i = f_j\})$, 只需验证 $\mathrm{Alt}^{(r)}(f_1 \otimes'' \cdots \otimes'' f_r) = 0$, 其中 $f_i \in V^*$ 且 $\exists 1 \leq i < j \leq r$, $s.t.\ f_i = f_j$. 考虑陪集分解 $S_r = A_r \sqcup A_r(i, j)$, 则 $\forall \alpha_1, \cdots, \alpha_r \in V$, $\mathrm{Alt}^{(r)}(f_1 \otimes'' \cdots \otimes'' f_r)(\alpha_1, \cdots, \alpha_r) = \sum\limits_{\sigma\in S_r} \mathrm{sgn}(\sigma) f_1(\alpha_{\sigma(1)}) \cdots f_r(\alpha_{\sigma(r)})$
$= \sum\limits_{\sigma\in A_r} f_1(\alpha_{\sigma(1)}) \cdots f_i(\alpha_{\sigma(i)}) \cdots f_j(\alpha_{\sigma(j)}) \cdots f_r(\alpha_{\sigma(r)}) - \sum\limits_{\sigma\in A_r} f_1(\alpha_{\sigma(1)}) \cdots f_i(\alpha_{\sigma(j)}) \cdots f_j(\alpha_{\sigma(i)}) \cdots f_r(\alpha_{\sigma(r)}) = 0.$

(2) 断言: $S^{-1}(\ker(\mathrm{Alt}^{(r)})) \subseteq I^r(V^*)$, 从而右端映射为单射. 这是因为, 任取 $u \in S^{-1}(\ker(\mathrm{Alt}^{(r)})) \subseteq T^r(V^*)$, 由第一个命题的证明知, $\forall \sigma \in S_r$, $\mathrm{sgn}(\sigma)\sigma(u) - u \in I^r(V^*)$. 对于 $\sigma \in S_r$ 以及 $L \in M^r(V)$, 记 $\sigma(L) \in M^r(V)$ 为
$(\alpha_1, \cdots, \alpha_r) \mapsto L(\alpha_{\sigma(1)}, \cdots, \alpha_{\sigma(r)})$, $\forall \alpha_1, \cdots, \alpha_r \in V$. 于是可直接验证: $\forall u \in T^r(V^*)$, $\sigma(S(u)) = S(\sigma(u))$, 以及
$\forall L \in M^r(V)$, $\mathrm{Alt}^{(r)}(L) - r!L = \sum\limits_{\sigma\in S_r} (\mathrm{sgn}(\sigma)\sigma(L) - L)$. 因此, 当 $u \in S^{-1}(\ker(\mathrm{Alt}^{(r)}))$ 且 $\mathrm{char}(F) \nmid r!$ 时,
$S(u) = -\dfrac{1}{r!} \sum\limits_{\sigma\in S_r} (\mathrm{sgn}(\sigma)\sigma(S(u)) - S(u)) = S\left(-\dfrac{1}{r!} \sum\limits_{\sigma\in S_r} (\mathrm{sgn}(\sigma)\sigma(u) - u)\right)$, 故由 S 为单射知,
$u = -\dfrac{1}{r!} \sum\limits_{\sigma\in S_r} (\mathrm{sgn}(\sigma)\sigma(u) - u) \in I^r(V^*)$.

(3) 显然映射 $S \colon T(V^*) \to M(V)$ 保持分次结合代数的乘法 (即两种张量积); 只需证明映射 $\widetilde{S} \colon \bigwedge(V^*) \to \Lambda(V^*)$ 保持分次结合代数的乘法 (即两种外积), 即证 $\forall u_1 \in T^r(V^*)$, $u_2 \in T^s(V^*)$, $\dfrac{1}{r!s!}\mathrm{Alt}^{(r+s)}(\mathrm{Alt}^{(r)}(S(u_1)) \otimes \mathrm{Alt}^{(s)}(S(u_2)))$
$= \mathrm{Alt}^{(r+s)}(S(u_1 \otimes u_2))$. 这是定义的直接验证. □

注:

(1) 在上述命题的条件下, 我们可不区分抽象定义的张量积 (或外积) 与安师讲义中具体定义的张量积 (或外积).

(2) 当 $\dim_F(V) < +\infty$ 时, 由维数关系知, 上述交换图中间的竖直映射 $S \colon T^r(V^*) \to M^r(V)$ 为线性同构, 从而两端的映射也为线性同构.

4.2.2 定向与体积形式

本节主要讨论有限维实线性空间的定向与体积形式, 并给出它们的几何解释.

引理 4.2.4 *设 V 为域 F 上的线性空间, $B = \{\alpha_i\}_{i\in I}$ 为 V 的基, 固定 I 上的一个严格全序 $<$, $r \in \mathbb{N}$, 则 $\{\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_r} \colon i_1 < \cdots < i_r\}$ 为 $\bigwedge^r(V)$ 的基. 特别地, $\dim_F(\bigwedge^r(V)) = \binom{\dim_F(V)}{r}$.*

证明: 由外积的构造、多线性性与交错性知, $\bigwedge^r(V) = \mathrm{Span}_F(\{\alpha_1 \wedge \cdots \wedge \alpha_r \colon \alpha_i \in V\}) = \mathrm{Span}_F(\{\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_r} \colon i_1 < \cdots < i_r\})$. 下证 $\{\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_r} \colon i_1 < \cdots < i_r\}$ 为线性无关集. 这是因为, 取 $B = \{\alpha_i\}_{i\in I}$ 的对偶集 $\{f_i\}_{i\in I}$, 则 $\forall i_1 < \cdots < i_r$, $f_{i_1} \wedge \cdots \wedge f_{i_r} \in \Lambda^r(V) \cong (\bigwedge^r(V))^*$. 在此同构的意义下, 可直接验证 $(f_{i_1} \wedge \cdots \wedge f_{i_r})(\alpha_{j_1} \wedge \cdots \wedge \alpha_{j_r})$
$= \delta_{i_1,j_1} \cdots \delta_{i_r,j_r}$, 故由对偶集的线性无关性引理知, $\{\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_r} \colon i_1 < \cdots < i_r\}$ 为线性无关集. □

注: 设 V 为域 F 上的 n 维线性空间, 当 $r>n$ 时, $\dim_F(\bigwedge^r(V))=\binom{n}{r}=0$, 即 $\bigwedge^r(V)=\{0\}$, 故 $\bigwedge(V)=\coprod_{r=0}^{n}\bigwedge^r(V)$. 此时一维线性空间 $\bigwedge^n(V)$ 称为**顶外积空间** (top exterior power), 其中的非零元之间只相差非零常数倍. 例如, 任取 V 的两个基 $B=\{\alpha_i\}_{i=1}^n$, $B'=\{\alpha'_j\}_{j=1}^n$, 记 $\alpha'_j=\sum\limits_{i=1}^{n}A_{ij}\alpha_i\,(1\leq j\leq n)$, 则 $\alpha'_1\wedge\cdots\wedge\alpha'_n=c(A)\alpha_1\wedge\cdots\wedge\alpha_n$, 其中 $c(A)=\sum\limits_{\sigma\in S_n}\operatorname{sgn}(\sigma)A_{\sigma(1),1}\cdots A_{\sigma(n),n}\in F\setminus\{0\}$.

定义 4.2.1 (定向) 设 V 为 n 维实线性空间, $B=(\alpha_1,\cdots,\alpha_n)$, $B'=(\alpha'_1,\cdots,\alpha'_n)$ 为 V 的两个有序基, 若 $\exists c>0$,
$s.t.\ \alpha'_1\wedge\cdots\wedge\alpha'_n=c\alpha_1\wedge\cdots\wedge\alpha_n$, 则它们称为**同向的** (same orientation); 若 $\exists c<0,\ s.t.\ \alpha'_1\wedge\cdots\wedge\alpha'_n=c\alpha_1\wedge\cdots\wedge\alpha_n$, 则它们称为**反向的** (opposite orientation).

记 $\mathscr{B}=\{V$中的全体有序基$\}$, 则同向性决定了 $\mathscr{B}$ 上的一个等价关系 $\sim$, 且当 $V\neq\{0\}$ 时恰有两个等价类, 当 $V=\{0\}$ 时只有一个等价类. 于是 V 上的一个**定向** (orientation) 是一个单射 $\mathscr{B}/\sim\to\{\pm1\}$, 即在一个等价类上取 $+1$, 在另一个等价类上取 -1, 于是恰有两种这样的定向. 固定一个有序基 B, 则它可决定 V 上的一个定向 $\mathscr{B}/\sim\longrightarrow\{\pm1\}$.

$$[B']\longmapsto\begin{cases}1, & 若B'\sim B\\ -1, & 若B'\not\sim B\end{cases}$$

定义 4.2.2 (体积形式) 设 V 为 n 维实线性空间, 则 $\Lambda^n(V)\cong(\bigwedge^n(V))^*$ 中的非零元称为 V 的一个**体积形式** (volume form). 固定一个体积形式 $L\in\Lambda^n(V)\backslash\{0\}$, 则它可决定 V 上的一个定向 $\mathscr{B}/\sim\longrightarrow\{\pm1\}$, $[(\alpha_1,\cdots,\alpha_n)]\longmapsto\operatorname{sgn}(L(\alpha_1,\cdots,\alpha_n))$
此时 $L\in\Lambda^n(V)\setminus\{0\}$ 称为一个**定向形式** (orientation form).

在安师讲义中, 设 V 为 n 维实线性空间, 任取 V 的有序基 $B=(\alpha_1,\cdots,\alpha_n)$, 记以 B 为顶点集的**广义平行多面体** (parallelotope) $P(\alpha_1,\cdots,\alpha_n)\coloneqq\left\{\sum\limits_{i=1}^{n}c_i\alpha_i\in V\colon 0\leq c_i\leq1\right\}$. 考虑在顶点集上的置换作用 $S_n\times B\longrightarrow B$, $(\sigma,\alpha_i)\longmapsto\alpha_{\sigma(i)}$, 它诱导了在顶外积空间上的置换作用 $S_n\times\bigwedge^n(V)\longrightarrow\bigwedge^n(V)$, $(\sigma,c\alpha_1\wedge\cdots\wedge\alpha_n)\longmapsto c\alpha_{\sigma(1)}\wedge\cdots\wedge\alpha_{\sigma(n)}$, 于是安师讲义中顶点序的等价 $(\alpha_1,\cdots,\alpha_n)\sim(\alpha_{\sigma(1)},\cdots,\alpha_{\sigma(n)})$ 即 $\sigma\in A_n$, 即顶形式的相同 $\alpha_1\wedge\cdots\wedge\alpha_n=\alpha_{\sigma(1)}\wedge\cdots\wedge\alpha_{\sigma(n)}$. 因此顶点序的等价类恰有 $[S_n\colon A_n]=\begin{cases}1, & n=1\\ 2, & n\geq2\end{cases}$ 个, 即 $P(\alpha_1,\cdots,\alpha_n)$ 上的 "定向" 恰有 $[S_n\colon A_n]$ 个. (注意这里广义平行多面体上的定向与线性空间上的定向的区别!)

另外, 任取体积形式 $L\in\Lambda^n(V)\setminus\{0\}$, 则 $L(\alpha_1,\cdots,\alpha_n)\in\mathbb{R}\setminus\{0\}$ 称为 $P(\alpha_1,\cdots,\alpha_n)$ 在 L 下的**有向体积** (oriented volume). 具体地说, $|L(\alpha_1,\cdots,\alpha_n)|>0$ 是 $P(\alpha_1,\cdots,\alpha_n)$ 在 L 下的 "绝对" 体积; $\operatorname{sgn}(L(\alpha_1,\cdots,\alpha_n))\in\{\pm1\}$ 反映了有序基 $B=(\alpha_1,\cdots,\alpha_n)$ 给出的定向与体积形式 L 给出的定向是否相同.

4.2.3 Grassmannian 簇与旗簇

本节将给出张量积与外积的一些应用, 包括 Grassmannian 簇、旗簇和它们的射影嵌入.

定义 4.2.3 (旗簇) 设 V 为域 F 上的线性空间, $0\leq n_1<\cdots<n_d\leq\dim_F(V)$, 记
$$\mathscr{F}_{n_1,\cdots,n_d}(V)\coloneqq\{F=(W_1,\cdots,W_d)\colon W_1\subseteq\cdots\subseteq W_d为V的线性子空间链,\text{且}\dim_F(W_i)=n_i\},$$
则 $\mathscr{F}_{n_1,\cdots,n_d}(V)$ 称为线性空间 V 的一个**部分旗簇** (partial flag variety). 特别地, 当 $\dim_F(V)=n$ 且 $n_i=i$, $\forall1\leq i\leq d=n$ 时, $\mathscr{F}_{1,\cdots,n}(V)$ 称为线性空间 V 的一个**全旗簇** (full flag variety).

定义 4.2.4 (Grassmannian 簇) 设 V 为域 F 上的线性空间, $0\leq r\leq\dim_F(V)$, 记
$$\mathscr{G}_r(V)\coloneqq\mathscr{F}_r(V)=\{W\subseteq V\colon W为V的线性子空间,\text{且}\dim_F(W)=r\},$$
则 $\mathscr{G}_r(V)$ 称为线性空间 V 的一个 **Grassmannian 簇** (Grassmannian variety). 特别地, 当 $r=1$ 时, $\mathbb{P}(V)\coloneqq\mathscr{G}_1(V)$ 称为线性空间 V 的**射影空间** (projective space).

命题 4.2.5 (Segre 嵌入) 设 $V_1,\cdots,V_d$ 为域 F 上的线性空间, 则映射 Φ: $\mathbb{P}(V_1)\times\cdots\times\mathbb{P}(V_d)\longrightarrow\mathbb{P}(V_1\otimes\cdots\otimes V_d)$

$$(\langle\alpha_1\rangle,\cdots,\langle\alpha_d\rangle)\longmapsto\langle\alpha_1\otimes\cdots\otimes\alpha_d\rangle$$

是单射, 其中 $\langle\alpha_i\rangle\coloneqq\mathrm{Span}_F(\{\alpha_i\})\subseteq V_i$.

证明: 由归纳法知可不妨设 $d=2$. 显然此映射定义良好; 现取 V_1 的基 $\{\alpha_i^{(1)}\}_{i\in I}$, V_2 的基 $\{\alpha_j^{(2)}\}_{j\in J}$, 记 $\alpha_1=\sum\limits_{i\in I}c_i^{(1)}\alpha_i^{(1)}$, $\alpha_2=\sum\limits_{j\in J}c_j^{(2)}\alpha_j^{(2)}$ (均为有限和), 则 $\alpha_1\otimes\alpha_2=\sum\limits_{i\in I,j\in J}c_i^{(1)}c_j^{(2)}\alpha_i^{(1)}\otimes\alpha_j^{(2)}$. 此公式说明, 在相差非零常数倍的意义下, α_1,α_2 的坐标可由 $\alpha_1\otimes\alpha_2$ 的坐标唯一确定, 故上述映射为单射. □

命题 4.2.6 (Plucker 嵌入) 设 V 为域 F 上的线性空间, $0\leq r\leq\dim_F(V)$ 且 $r\in\mathbb{N}$, 则映射 ι_r: $\mathscr{G}_r(V)\longrightarrow\mathbb{P}(\bigwedge^r(V))$

$$W\longmapsto\langle\alpha_1\wedge\cdots\wedge\alpha_r\rangle$$

是单射, 其中 $\{\alpha_i\}_{i=1}^r$ 为 W 的基, $\langle\alpha_1\wedge\cdots\wedge\alpha_r\rangle\coloneqq\mathrm{Span}_F(\{\alpha_1\wedge\cdots\wedge\alpha_r\})\subseteq\bigwedge^r(W)\subseteq\bigwedge^r(V)$.

证明: 由 $\bigwedge^r(W)$ 为顶外积空间知, 其中的非零元只相差非零常数倍, 故此映射定义良好. 现设 $\{\alpha_i\}_{i=1}^r$ 为 W 的基, 任取 $c\in F\setminus\{0\}$, 考虑线性映射 T: $V\longrightarrow\bigwedge^{r+1}(V)$, 则 $\ker(T)=\mathrm{Span}_F(\{\alpha_i\}_{i=1}^r)=W$, 即线性

$$\alpha\longmapsto\alpha\wedge(c\alpha_1\wedge\cdots\wedge\alpha_r)$$

子空间 W 可由它的顶外积空间 $\bigwedge^r(W)$ 唯一确定, 故上述映射为单射. □

推论 4.2.7 设 V 为域 F 上的线性空间, $0\leq n_1<\cdots<n_d\leq\dim_F(V)$, 则映射

$$\begin{aligned}\mathscr{F}_{n_1,\cdots,n_d}(V)&\longrightarrow\mathbb{P}(\textstyle\bigwedge^{n_1}(V)\otimes\cdots\otimes\bigwedge^{n_d}(V))\\(W_1,\cdots,W_d)&\longmapsto\Phi(\iota_{n_1}(W_1),\cdots,\iota_{n_d}(W_d))\end{aligned}$$

是单射.

注: 此推论说明线性空间的任意部分旗簇都可视为某个射影空间的子集; 甚至加 Zariski 拓扑条件时, 这还是一个闭子集, 于是确实是一个 (射影) 代数簇.

参考文献与补注 4.2

(1) 关于张量积与外积的部分, 可以参考 W. Greub“Multilinear Algebra”.

(2) 关于 Grassmannian 簇与旗簇的部分, 可以参考 Geck, Meinolf “An Introduction to Algebraic Geometry and Algebraic Groups”.

§4.3 行列式的意义与应用

4.3.1 行列式的定义

我们回忆有限维线性空间上线性映射的行列式的定义: 设 V 为域 F 上的 n 维线性空间, $T\in L(V)$, 令拉回映射

$$\begin{aligned}T^{(n)}\colon\ \Lambda^n(V)&\longrightarrow\Lambda^n(V)\\L&\longmapsto(T^{(n)}(L)\colon(\alpha_1,\cdots,\alpha_n)\mapsto L(T(\alpha_1),\cdots,T(\alpha_n)))\end{aligned}$$

, 则 $T^{(n)}\in L(\Lambda^n(V))$. 由 $\dim_F(\Lambda^n(V))=1$ 知, $\exists!c=c(T)\in F$, $s.t.\ T=c\mathrm{id}_{\Lambda^n(V)}$. 记 $\det(T)\coloneqq c(T)$ 为 $T\in L(V)$ 的**行列式** (determinant).

具体地说, 任取 $L\in\Lambda^n(V)$, 以及 $\alpha_1,\cdots,\alpha_n\in V$, 则 $L(T(\alpha_1),\cdots,T(\alpha_n))=\det(T)L(\alpha_1,\cdots,\alpha_n)$. 特别地, 若 $L\in\Lambda^n(V)\setminus\{0\}$ 为体积形式, $\{\alpha_i\}_{i=1}^n$ 为 V 的有序基, 则 $\det(T)=\dfrac{L(T(\alpha_1),\cdots,T(\alpha_n))}{L(\alpha_1,\cdots,\alpha_n)}$ 是在 L 下广义平行多面体 $P(T(\alpha_1),\cdots,T(\alpha_n))$ 与 $P(\alpha_1,\cdots,\alpha_n)$ 的体积的比值. 若考虑线性同构 $\Lambda^n(V)\cong(\bigwedge^n(V))^*$, 则前式也可写成 $T(\alpha_1)\wedge\cdots\wedge T(\alpha_n)=\det(T)\alpha_1\wedge\cdots\wedge\alpha_n$, 此即多元微积分中变量替换公式的线性版本.

特别地, 若取 $L=f_1\wedge\cdots\wedge f_n$, 其中 $\{f_i\}_{i=1}^n\subseteq V^*$ 为 $\{\alpha_i\}_{i=1}^n$ 的对偶基, 则前式也可写成

$$\det(T)=(f_1\wedge\cdots\wedge f_n)(T(\alpha_1),\cdots,T(\alpha_n))=(T^t(f_1)\wedge\cdots\wedge T^t(f_n))(\alpha_1,\cdots,\alpha_n).$$

回忆线性同构 $V\otimes V^*\xrightarrow{\cong}L(V)$, 其中将纯张量 $\alpha\otimes f$ 映为 T: $\beta\mapsto f(\beta)\alpha$, 则在此线性同构下, $\sum\limits_{i=1}^n\alpha_i\otimes T^t(f_i)$ 的像为 T. 于是行列式函数可视为关于线性函数组 $\{T^t(f_i)\}_{i=1}^n$ 的多线性、交错性、规范性函数. 以矩阵的语言, 记 T 在基 $\{\alpha_i\}_{i=1}^n$ 下的矩阵表示为 $A\in F^{n\times n}$, 则上式为 $\det(T)=\sum\limits_{\sigma\in S_n}\mathrm{sgn}(\sigma)A_{1,\sigma(1)}\cdots A_{n,\sigma(n)}$, 于是行列式函数又可视为关于矩阵表示的行向量组的多线性、交错性、规范性函数.

4.3.2 行列式与迹

由拉回映射的复合知, 行列式函数 $\det\colon (L(V),\circ,\mathrm{id}_V)\longrightarrow (F,\cdot,1)$ 是一个幺半群同态, 它的限制

$$T\longmapsto \det(T)$$

$\det\big|_{\mathrm{GL}(V)}\colon (\mathrm{GL}(V),\circ,\mathrm{id}_V)\longrightarrow (F^*,\cdot,1)$ 是一个群同态, 记 $\mathrm{SL}(V)\coloneqq\ker(\det\big|_{\mathrm{GL}(V)})$. 当取定 V 的基时, 行列

$$T\longmapsto \det(T)$$

式函数的限制也可视为 $\det\big|_{\mathrm{GL}(n,F)}\colon (\mathrm{GL}(n,F),\circ,\mathrm{id}_V)\longrightarrow (F^*,\cdot,1)$, 记 $\mathrm{SL}(n,F)\coloneqq\ker(\det\big|_{\mathrm{GL}(n,F)})$. 显然

$$A\longmapsto \det(L_A)$$

$\mathrm{SL}(V)\supseteq(\mathrm{GL}(V))'$; 同理 $\mathrm{SL}(n,F)\supseteq(\mathrm{GL}(n,F))'$. 类比 Lie 代数中的讨论, 我们先研究上述等号何时成立.

引理 4.3.1 *设 F 为一个域, $n\in\mathbb{N}^*$, 则 $\mathrm{SL}(n,F)=\langle I_n+cE_{ij}\colon c\in F,\ 1\le i\neq j\le n\rangle$.*

证明: 任取 $c\in F$, $1\le i\neq j\le n$, 由 $(I_n+cE_{ij})(I_n-cE_{ij})=I_n$ 知, $\det(I_n+cE_{ij})=1$, 即 $I_n+cE_{ij}\in\mathrm{SL}(n,F)$. 另一方面, 设 $A\in\mathrm{SL}(n,F)$. 由 $A\in\mathrm{GL}(n,F)$ 知, $\exists c\in F,\ 2\le i_0\le n,\ s.t.\ cA_{11}+A_{i_0,1}\neq 0$, 故

$$(I_n-(cA_{11}+A_{i_0,1})E_{i_0,1})\cdot\prod_{\substack{i=2\\ i\neq i_0}}^{n}(I_n-A_{i1}E_{i1})\cdot(I_n+\frac{1-A_{11}}{cA_{11}+A_{i_0,1}}E_{1,i_0})\cdot(I_n+cE_{i_0,1})\cdot A=\begin{pmatrix}1 & *\\ 0 & A_1\end{pmatrix},$$

其中 $A_1\in\mathrm{SL}(n-1,F)$. 由归纳法知, 存在矩阵 P_0 为若干初等矩阵 $I_n+cE_{ij}\,(c\in F,\ 1\le i\neq j\le n)$ 的乘积, 满足 P_0A 为对角分量均为 1 的上三角阵. 进一步地, 可再利用上述初等矩阵的乘积将 P_0A 的严格上三角部分消为 0, 即 $\exists P\in\langle I_n+cE_{ij}\colon c\in F,\ 1\le i\neq j\le n\rangle$, $s.t.\ PA=I_n$, 故 $A=P^{-1}\in\langle I_n+cE_{ij}\colon c\in F,\ 1\le i\neq j\le n\rangle$. □

命题 4.3.2 *设 V 为域 F 上的 n 维线性空间, 若 $n\neq 2$ 或 $F\neq\mathbb{F}_2$, 则 $\mathrm{SL}(V)=(\mathrm{SL}(V))'=(\mathrm{GL}(V))'$.*

证明: 只需证明 $\mathrm{SL}(V)\subseteq(\mathrm{SL}(V))'$; 当取定 V 的基时, 由上述引理知, 只需证明 $\forall c\in F$, $\forall 1\le i\neq j\le n$, $I_n+cE_{ij}\in(\mathrm{SL}(n,F))'$. 当 $n=1$ 时结论平凡. 当 $n=2$ 且 $F\neq\mathbb{F}_2$ 时, 可取 $c'\in F\setminus\{0,1\}$, 则

$$I_2+cE_{12}=\mathrm{diag}(c',c'^{-1})\cdot(I_2+\frac{c}{c'^2-1}E_{12})\cdot\mathrm{diag}(c',c'^{-1})^{-1}\cdot(I_2+\frac{c}{c'^2-1}E_{12})^{-1}\in(\mathrm{SL}(2,F))'.$$

同理知 $I_2+cE_{21}\in(\mathrm{SL}(2,F))'$. 当 $n\ge 3$ 时, 可取 $k\in\{1,\cdots,n\}\setminus\{i,j\}$, 则

$$I_n+cE_{ij}=(I_n+cE_{ik})(I_n+E_{kj})(I_n+cE_{ik})^{-1}(I_n+E_{kj})^{-1}\in(\mathrm{SL}(n,F))'.$$

□

注: 当 $n=2$ 且 $F=\mathbb{F}_2$ 时, 可直接验证 $\mathrm{GL}(2,\mathbb{F}_2)=\mathrm{SL}(2,\mathbb{F}_2)\cong S_3$, 故 $(\mathrm{GL}(2,\mathbb{F}_2))'=(\mathrm{SL}(2,\mathbb{F}_2))'\cong A_3$.

以下试图对行列式函数 "求导", 从而将上述讨论与它们的 Lie 代数版本联系起来. 为方便起见, 现设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 从而 $V\cong F^n$, 且赋予欧氏范数 $\|\cdot\|$.

固定 $S\in L(V)$, 考虑 $L(V)$ 中的一条曲线 $\gamma_S\colon (-\epsilon,\epsilon)\subseteq\mathbb{R}\longrightarrow L(V)$, 先断言: 当 $\epsilon=\epsilon(S)>0$ 充分

$$t\longmapsto \mathrm{id}_V+tS$$

小时, $\mathrm{Im}(\gamma_S)\subseteq\mathrm{GL}(V)$. (这是因为, $\det\circ\gamma_S$ 是连续函数, 且 $(\det\circ\gamma_S)(0)=1$, 故当 $\epsilon=\epsilon(S)>0$ 充分小时, $\forall t\in(-\epsilon,\epsilon)$,

$(\det\circ\gamma_S)(t)\neq 0$, 即 $\gamma_S(t)\in\mathrm{GL}(V)$.) 现考虑复合函数 $\det\big|_{\mathrm{GL}(V)}\circ\gamma_S\colon (-\epsilon,\epsilon)\subseteq\mathbb{R}\longrightarrow F\setminus\{0\}$, 再取 V

$$t\longmapsto \det(\mathrm{id}_V+tS)$$

的基 $\{\alpha_i\}_{i=1}^n$, 以及它的对偶基 $\{f_i\}_{i=1}^n \subseteq V^*$, 则

$$\begin{aligned}\det(\mathrm{id}_V + tS) =& (f_1 \wedge \cdots \wedge f_n)((\mathrm{id}_V + tS)(\alpha_1), \cdots, (\mathrm{id}_V + tS)(\alpha_n)) \\ =& \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) f_1((\mathrm{id}_V + tS)(\alpha_{\sigma(1)})) \cdots f_n((\mathrm{id}_V + tS)(\alpha_{\sigma(n)})) \\ =& \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)(f_1(\alpha_{\sigma(1)}) + tf_1(S(\alpha_{\sigma(1)}))) \cdots (f_n(\alpha_{\sigma(n)}) + tf_n(S(\alpha_{\sigma(n)}))) \\ =& \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) f_1(\alpha_{\sigma(1)}) \cdots f_n(\alpha_{\sigma(n)}) \\ & + t \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \sum_{i=1}^n f_1(\alpha_{\sigma(1)}) \cdots f_{i-1}(\alpha_{\sigma(i-1)}) f_i(S(\alpha_{\sigma(i)})) f_{i+1}(\alpha_{\sigma(i+1)}) \cdots f_n(\alpha_{\sigma(n)}) \\ & + O(t^2) \\ =& (f_1 \wedge \cdots \wedge f_n)(\alpha_1, \cdots, \alpha_n) + t \sum_{i=1}^n (f_1 \wedge \cdots \wedge f_n)(\alpha_1, \cdots, \alpha_{i-1}, S(\alpha_i), \alpha_{i+1}, \cdots, \alpha_n) + O(t^2) \\ =& 1 + t \sum_{i=1}^n f_i(S(\alpha_i)) + O(t^2),\end{aligned}$$

故 $\left.\dfrac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\det|_{\mathrm{GL}(V)} \circ \gamma_S) = \sum\limits_{i=1}^n f_i(S(\alpha_i)) = \mathrm{tr}(S)$. 于是存在以下交换图 $\begin{array}{ccc} \mathrm{Im}(\gamma_S) \subseteq \mathrm{GL}(V) & \xrightarrow{\det} & F^* \\ \downarrow D & & \downarrow \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \\ L(V) & \xrightarrow{\mathrm{tr}} & F \end{array}$.

反之, 给定 $S \in L(V)$, 考虑 $\exp(S)\colon V \longrightarrow V$, 先断言: 这是定义良好的可逆线性映射. (事实上, 记

$$\alpha \longmapsto \sum_{k=0}^{+\infty} \frac{S^k(\alpha)}{k!}$$

$\|S\|_{\mathrm{op}} \coloneqq \max\limits_{\alpha \in V \setminus \{0\}} \left\{ \dfrac{\|S(\alpha)\|}{\|\alpha\|} \right\}$, 由于 $\sum\limits_{k=0}^{+\infty} \dfrac{\left\|S^k\right\|_{\mathrm{op}}}{k!} \leq \sum\limits_{k=0}^{+\infty} \dfrac{\|S\|_{\mathrm{op}}^k}{k!} = e^{\|S\|_{\mathrm{op}}} < +\infty$, 则 $\left\{ \sum\limits_{k=0}^{m} \dfrac{\left\|S^k\right\|_{\mathrm{op}}}{k!} \right\}_{m=0}^{+\infty}$ 为 $\mathbb{R}$ 中的 Cauchy 列; 又 $\forall \alpha \in V$, $\forall l \geq m \geq 0$, $\left\| \sum\limits_{k=m}^{l} \dfrac{S^k(\alpha)}{k!} \right\| \leq \sum\limits_{k=m}^{l} \dfrac{\left\|S^k(\alpha)\right\|}{k!} \leq \sum\limits_{k=m}^{l} \dfrac{\left\|S^k\right\|_{\mathrm{op}} \|\alpha\|}{k!}$, 故 $\left\{ \sum\limits_{k=0}^{m} \dfrac{S^k(\alpha)}{k!} \right\}_{m=0}^{+\infty}$ 为 V 中的 Cauchy 列, 因此由 $(V, \|\cdot\|)$ 完备知 $\sum\limits_{k=0}^{+\infty} \dfrac{S^k(\alpha)}{k!} \in V$ 必存在, 且 $\sum\limits_{k=0}^{+\infty} \dfrac{S^k}{k!}$ 在 V 的任意紧子集上一致收敛. 特别地, $\exp(S) \coloneqq \sum\limits_{k=0}^{+\infty} \dfrac{S^k}{k!}$ 是线性映射. 进一步地, 可直接验证 $\forall S_1, S_2 \in L(V)$, $\exp(S_1) \circ \exp(S_2) = \exp(S_1 + S_2)$, 且 $\exp(0_V) = \mathrm{id}_V$, 故 $\forall S \in L(V)$, $\exp(S) \in \mathrm{GL}(V)$, 且指数映射 $\exp\colon (L(V), +, 0_V) \longrightarrow (\mathrm{GL}(V), \circ, \mathrm{id}_V)$ 为群同态.)

$$S \longmapsto \exp(S)$$

现考虑复合群同态 $\varphi\colon (\mathbb{R}, +, 0) \longrightarrow (F^*, \cdot, 1)$, 则同上知 $\det(\exp(tS)) = 1 + t\mathrm{tr}(S) + O(t^2)$, 故

$$t \longmapsto \det(\exp(tS))$$

$\left.\dfrac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \det(\exp(tS)) = \mathrm{tr}(S)$. (注意这里 $O(t^2)$ 中含无穷多高次项, 但由上述系数的一致收敛性知, 它仍可逐项求导, 故为 0.) 一般地,

$$\forall t_0 \in \mathbb{R},\ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} \det(\exp(tS)) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} \det(\exp((t-t_0)S)) \det(\exp(t_0 S)) = \mathrm{tr}(S) \det(\exp(t_0 S)),$$

即群同态 φ 满足一阶常系数微分方程 $\begin{cases} \dfrac{\mathrm{d}\varphi}{\mathrm{d}t} = \mathrm{tr}(S)\varphi \\ \varphi(0) = 1 \end{cases}$, 解得 $\varphi(t) = e^{t\mathrm{tr}(S)}$, $\forall t \in \mathbb{R}$. 特别地, $\det(\exp(S)) = e^{\mathrm{tr}(S)}$.

于是存在以下群正合列的交换图 $\begin{array}{ccccccccc} 1 & \longrightarrow & (\mathrm{SL}(V), \circ, \mathrm{id}_V) & \longrightarrow & (\mathrm{GL}(V), \circ, \mathrm{id}_V) & \xrightarrow{\det} & (F^*, \cdot, 1) & \longrightarrow & 1 \\ & & \uparrow \exp & & \uparrow \exp & & \uparrow \exp & & \\ 0 & \longrightarrow & (\mathfrak{sl}(V) \coloneqq \ker(\mathrm{tr}), +, 0_V) & \longrightarrow & (\mathfrak{gl}(V) \coloneqq L(V), +, 0_V) & \xrightarrow{\mathrm{tr}} & (F, +, 0) & \longrightarrow & 0 \end{array}$.

当取定 V 的基时, 此交换图也可视为 $\begin{array}{ccccccccc} 1 & \longrightarrow & (\mathrm{SL}(n, F), \circ, \mathrm{id}_V) & \longrightarrow & (\mathrm{GL}(n, F), \circ, \mathrm{id}_V) & \xrightarrow{\det} & (F^*, \cdot, 1) & \longrightarrow & 1 \\ & & \uparrow \exp & & \uparrow \exp & & \uparrow \exp & & \\ 0 & \longrightarrow & (\mathfrak{sl}(n, F), +, 0_V) & \longrightarrow & (\mathfrak{gl}(n, F), +, 0_V) & \xrightarrow{\mathrm{tr}} & (F, +, 0) & \longrightarrow & 0 \end{array}$.

最后, 我们再利用流形上向量场的相流和散度给出行列式与迹的一种几何解释. 现仍设 $F=\mathbb{R}$ 或 $\mathbb{C}$, V 为域 F 上的 n 维线性空间. 一方面, 通过赋予欧氏范数 $\|\cdot\|$, $V\cong F^n$ 可视为标准 n 维欧氏流形; 另一方面, 在流形 V 上的任一点 α 处, 切空间 $T_\alpha V$ 在坐标映射下线性同构于线性空间 V. 于是任意 $T\in L(V)$ 都决定了流形 V 上的一个向量场 $X_T\colon \begin{array}{ccc} V & \longrightarrow & \bigsqcup\limits_{\alpha\in V} T_\alpha V \\ \alpha & \longmapsto & T(\alpha)\in V\cong T_\alpha V \end{array}$, 即 X_T 在流形 V 上的任一点 α 处指定切方向 $T(\alpha)\in V\cong T_\alpha V$.

再取 V 的基
$B=\{\alpha_i\}_{i=1}^n$, 记 $T\in L(V)$ 在基 B 下的矩阵表示为 $A\in F^{n\times n}$, 则 X_T 的坐标表示为 L_A, 故向量场的**散度** (divergence) $\operatorname{div}(X_T)\coloneqq\sum\limits_{i=1}^n\dfrac{\partial(X_T)_i}{\partial x_i}=\sum\limits_{i=1}^n A_{ii}=\operatorname{tr}(A)$.

此外, 向量场 X_T 可生成流形 V 上的一条过点 α 的积分曲线 γ_α, 即光滑曲线 $\gamma_\alpha\colon\mathbb{R}\to V$ 满足 $\begin{cases}\dfrac{\mathrm{d}\gamma_\alpha}{\mathrm{d}t}=X_T\circ\gamma_\alpha,\\ \gamma_\alpha(0)=\alpha\end{cases}$
解得 $\gamma_\alpha(t)=\exp(tL_A)(\alpha)$, $\forall t\in\mathbb{R}$. 于是所有这样的积分曲线 $\{\gamma_\alpha\colon\alpha\in V\}$ 给出了一个群同态
$\varphi\colon \begin{array}{ccc} (\mathbb{R},+,0) & \longrightarrow & (\mathrm{GL}(V),\circ,\mathrm{id}_V) \\ t & \longmapsto & (\exp(tL_A)\colon\alpha\mapsto\gamma_\alpha(t)) \end{array}$, 称为该向量场生成的**相流** (phase flow). 注意 $\det(\exp(tL_A))=e^{t\operatorname{tr}(L_A)}$,
故综上知, 几何上该相流是保体积的 (即 $\forall t\in\mathbb{R}$, $\det(\varphi(t))=1$) $\Longleftrightarrow$ 该向量场是无源的 (即 $\operatorname{div}(X_T)=0$).

4.3.3 行列式与扰动

我们回忆域 F 上方阵的行列式的公式: 设 $A\in F^{n\times n}$, 则
$$\det(A)=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)A_{1,\sigma(1)}\cdots A_{n,\sigma(n)}=\sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)A_{\sigma(1),1}\cdots A_{\sigma(n),n}.$$
此式中 $\det(A)$ 关于每个 A_{ij} $(1\le i,j\le n)$ 都是一次多项式. 严格地说, 设 R 为一个交换环, $A\in R^{n\times n}$, 则上式定义了一个幺半群同态 $\det\colon(R^{n\times n},\cdot,I_n)\to(R,\cdot,1)$. 取 $R=F[X_{ij}]_{1\le i,j\le n}$ 为域 F 上的多元多项式环, 以及 $A=(X_{ij})_{1\le i,j\le n}$, 则 $\det((X_{ij})_{1\le i,j\le n})\in F[X_{ij}]_{1\le i,j\le n}$ 关于每个 X_{ij} $(1\le i,j\le n)$ 都是一次多项式. 一个非平凡的事实如下:

引理 4.3.3 *设 $\varphi\colon((F[X_{ij}]_{1\le i,j\le n})^{n\times n},\cdot,I_n)\to(F[X_{ij}]_{1\le i,j\le n},\cdot,1)$ 为幺半群同态, 且 $\varphi((X_{ij})_{1\le i,j\le n})\in F[X_{ij}]_{1\le i,j\le n}$ 非常值多项式, 则 $\deg\varphi((X_{ij})_{1\le i,j\le n})\ge n$, 且 "$=$" 取到 $\Longleftrightarrow\varphi=\det$.*

我们跳过这个事实的证明. 现考虑 $A\in F^{n\times n}$ 的扰动 $A+XI_n\in(F[X])^{n\times n}$, 则 $\det(A+XI_n)\in F[X]$ 是 n 次首一多项式. 记 $F(X)=\operatorname{Frac}(F[X])\coloneqq\left\{\dfrac{f(X)}{g(X)}\colon f(X)\in F[X],g(X)\in F[X]^*\right\}$ 为域 F 上一元多项式环的**分式域** (field of fractions). 由 $A+XI_n\in(F(X))^{n\times n}$ 且 $\det(A+XI_n)\in F(X)^*$ 知, $A+XI_n\in\mathrm{GL}(n,F(X))$. 于是通过添加未定元与域扩张, 我们将域 F 上的任意方阵 A 扰动成了域 $F(X)$ 上的可逆方阵 $A+XI_n$. 这一操作的好处在于, 如果需要证明某个关于域 F 上方阵的等式, 我们可以先将它扰动为一个关于环 $F[X]$ 上方阵的等式; 再在域 $F(X)$ 上看, 此时扰动所涉及的方阵是 $(F(X))^{n\times n}$ 中的可逆阵, 于是我们可以在 $(F(X))^{n\times n}$ 中证明此等式; 又由于扰动所涉及的方阵在 $(F[X])^{n\times n}$ 中, 故此等式在 $(F[X])^{n\times n}$ 中也成立; 最后通过取值同态 $\begin{array}{ccc} F[X] & \longrightarrow & F \\ f(X) & \longmapsto & f(0) \end{array}$ (或比较常数项), 此等式恢复成了 $F^{n\times n}$ 中的原等式.

引理 4.3.4 (Schur 公式) *设 $M\in F^{n\times n}$, 且可分块为 $\begin{pmatrix}A&B\\C&D\end{pmatrix}$, 其中 $A\in F^{r\times r}$, 则*

(1) *若 $A\in\mathrm{GL}(r,F)$, 则 $\det(M)=\det(A)\det(D-CA^{-1}B)$;*
(2) *若 $D\in\mathrm{GL}(n-r,F)$, 则 $\det(M)=\det(D)\det(A-BD^{-1}C)$.*

证明: 由矩阵分块零化技巧与行列式的定义即知. □

推论 4.3.5 *设 $M\in F^{2n\times 2n}$, 且可分块为 $\begin{pmatrix}A&B\\C&D\end{pmatrix}$, 其中 $A\in F^{n\times n}$, 则*

(1) *若 $AC=CA$, 则 $\det(M)=\det(AD-CB)$;*
(2) *若 $AB=BA$, 则 $\det(M)=\det(DA-CB)$;*

(3) 若 $BD=DB$, 则 $\det(M)=\det(DA-BC)$;

(4) 若 $CD=DC$, 则 $\det(M)=\det(AD-BC)$.

证明: 只证明 (1), 其余完全类似. 现设 $AC=CA$. 若 $A\in \mathrm{GL}(r,F)$, 则由上述 Schur 公式即知结论. 一般地, 考虑 $A\in F^{n\times n}$ 的扰动 $A+XI_n\in (F[X])^{n\times n}$, 则 $(A+XI_n)C=C(A+XI_n)$, 且在域 $F(X)$ 上的 Schur 公式

$$\det\begin{pmatrix} A+XI_n & B\\ C & D\end{pmatrix}=\det(A+XI_n)\det(D-C(A+XI_n)^{-1}B)=\det((A+XI_n)D-CB)$$

成立. 注意此等式的左端与右端都为 $F[X]$ 中元, 故 $\det\begin{pmatrix} A+XI_n & B\\ C & D\end{pmatrix}=\det((A+XI_n)D-CB)$ 在 $F[X]$ 中也成立. 最后通过取值同态 $\begin{array}{ccc} F[X] & \longrightarrow & F\\ f(X) & \longmapsto & f(0)\end{array}$ (或比较常数项) 知, $\det\begin{pmatrix} A & B\\ C & D\end{pmatrix}=\det(AD-CB)$. □

上述扰动法是更具代数特色的扰动. 当 $F\subseteq\mathbb{C}$ 为子域时, 利用范数 $|\cdot|$ 可以给出更具分析特色的扰动. 具体地说, 仍考虑 $A\in F^{n\times n}$ 的扰动 $A+XI_n\in (F[X])^{n\times n}$, 则 $\det(A+XI_n)\in F[X]$ 是 n 次首一多项式, 故它在域 F 内至多有 n 个根. 因此由根集的离散性知, $\exists\epsilon>0,\ \forall c\in F\,(0<|c|<\epsilon),\ \det(A+cI_n)\neq 0$, 即 $A+cI_n\in\mathrm{GL}(n,F)$. 这一操作的好处在于, 如果需要证明某个关于域 F 上方阵的等式, 我们可以先作如上扰动, 此时扰动所涉及的方阵是可逆的, 于是我们相对容易证明此等式; 又由于等式两端都是关于 c 的多项式函数, 且在无穷集 $\{c\in F\colon 0<|c|<\epsilon\}$ 上处处取值相等, 故等式两端作为多项式相等 (即多项式的系数对应相等). 最后通过取它们在 0 处的值 (或比较常数项), 此等式恢复成了原等式.

最后, 当 $F=\mathbb{R}$ 时, 利用多项式函数的介值性可以给出另一种扰动法. 具体地说, 设 $A\in\mathbb{R}^{n\times n}$, 则 $\begin{array}{ccc}\mathbb{R} & \longrightarrow & \mathbb{R}\\ t & \longmapsto & \det(A+tI_n)\end{array}$ 为 n 次首一多项式函数, 则 $\lim\limits_{t\to+\infty}\det(A+tI_n)=+\infty$. 由连续函数的介值定理知, 若 $\exists t_0\in\mathbb{R},\ s.t.\ \det(A+t_0I_n)<0$, 则 $\exists t_1>t_0,\ s.t.\ \det(A+t_1I_n)=0$. 这常常为我们判断行列式的正负性提供了帮助.

命题 4.3.6 (Levy-Desplanques) 设 $F\subseteq\mathbb{C}$ 为子域, $A\in F^{n\times n}$, 满足严格对角占优条件 $\forall 1\leq i\leq n,\ |A_{ii}|>\sum\limits_{j\neq i}|A_{ij}|$, 则:

(1) $\det(A)\neq 0$;

(2) 进一步设 $A=\overline{A^t}$, 且 $\forall 1\leq i\leq n,\ A_{ii}>0$, 则 $\det(A)>0$.

证明: (1) 假设 $\det(A)=0$, 即 $A\notin\mathrm{GL}(n,F)$, 则 $\exists 0\neq X\in F^{n\times 1},\ s.t.\ AX=0$. 记 $|x_i|=\max\limits_{1\leq j\leq n}|x_j|>0$, 则由 $\sum\limits_{j=1}^{n}A_{ij}x_j=0$ 知, $|A_{ii}|=\left|\dfrac{1}{x_i}\sum\limits_{j\neq i}A_{ij}x_j\right|\leq\dfrac{1}{|x_i|}\sum\limits_{j\neq i}|A_{ij}||x_j|\leq\sum\limits_{j\neq i}|A_{ij}|$, 这与 $|A_{ii}|>\sum\limits_{j\neq i}|A_{ij}|$ 矛盾!

(2) 由 $A=\overline{A^t}$ 知, $\forall t\geq 0,\ A+tI_n=\overline{(A+tI_n)^t}$, 则 $\det(A+tI_n)=\overline{\det(A+tI_n)}$, 故 $\det(A+tI_n)\in\mathbb{R}$. 考虑 n 次首一多项式函数 $\begin{array}{ccc}[0,+\infty) & \longrightarrow & \mathbb{R}\\ t & \longmapsto & \det(A+tI_n)\end{array}$, 由 $\forall 1\leq i\leq n,\ A_{ii}>0$ 知, $\forall t\geq 0,\ A+tI_n$ 仍为对角占优的, 故由 (1) 知 $\forall t\geq 0,\ \det(A+tI_n)\neq 0$. 又 $\lim\limits_{t\to+\infty}\det(A+tI_n)=+\infty$, 故由连续函数的介值定理知, $\forall t\geq 0,\ \det(A+tI_n)>0$. 特别地, $\det(A)>0$. □

推论 4.3.7 设 $F\subseteq\mathbb{C}$ 为子域, $A\in F^{n\times n}$, 满足 $A=\overline{A^t}$, 以及对角占优条件 $\forall 1\leq i\leq n,\ |A_{ii}|\geq\sum\limits_{j\neq i}|A_{ij}|$ 且 $A_{ii}\geq 0$, 则 $\det(A)\geq 0$.

证明: 任取 $t>0$, 则由条件知 $A+tI_n=\overline{(A+tI_n)^t}$, 且 $\forall 1\leq i\leq n,\ A_{ii}+t>\sum\limits_{j\neq i}|A_{ij}|$, 故由上述命题知, $\det(A+tI_n)>0$. 令 $t\to 0^+$ 即知 $\det(A)\geq 0$. □

注: 在上述命题与推论中, 若将 (严格) 对角占优条件减弱为 (严格)Brauer 条件 $|A_{ii}A_{jj}|\geq(>)\sum\limits_{k\neq i}|A_{ik}|\cdot\sum\limits_{l\neq j}|A_{jl}|$, $\forall 1\leq i\neq j\leq n$, 则上述结论仍成立, 证明完全类似.

4.3.4 附属方阵的性质

本节将总结附属方阵的若干性质, 它的作用在于当方阵不可逆时替代逆矩阵而完成一般的论证.

定义 4.3.1 (附属方阵) 设 R 为一个交换环, $A \in R^{n\times n}$, 记 C_{ij} 为 A 中 (i,j) 位置元的代数余子式, 令 $\mathrm{adj}(A)_{ij} \coloneqq C_{ji}$, 则称 $\mathrm{adj}(A) \in R^{n\times n}$ 为 A 的**附属方阵** (adjugate matrix).

注:

(1) 在不同文献中, "附属方阵" 具有不同的翻译方式: adjugate/adjunct/classical adjoint matrix. 为了避免与伴随映射的矩阵表示混淆, 我们这里采用 "附属方阵" 的表述方式.

(2) 由行列式的 Laplace 展开知, $\forall A \in R^{n\times n}$, $A\cdot\mathrm{adj}(A) = \det(A)I_n = \mathrm{adj}(A)\cdot A$, 则当 $A \in \mathrm{GL}(n,R)$ 时 (即当 $\det(A) \in U(R)$ 时), $A^{-1} = (\det(A))^{-1}\mathrm{adj}(A)$. 而一般地, 我们常可通过左 (或右) 乘附属方阵的方法将矩阵等式简化为纯量等式, 这一手段称为 "行列式技巧".

命题 4.3.8 *设 R 为一个交换环, $A \in R^{n\times n}$, 则*

(1) $\mathrm{adj}(I_n) = I_n$; $\mathrm{adj}(0_n) = \begin{cases} I_1, & n = 1 \\ 0_n, & n \geq 2 \end{cases}$;

(2) $\forall c \in R$, $\mathrm{adj}(cA) = c^{n-1}\mathrm{adj}(A)$;

(3) $\mathrm{adj}(A^t) = \mathrm{adj}(A)^t$;

(4) $\det(\mathrm{adj}(A)) = (\det(A))^{n-1}$;

(5) *当 $R = F$ 为一个域时*, $r(\mathrm{adj}(A)) = \begin{cases} n, & r(A) = n \\ 1, & r(A) = n-1 \\ 0, & r(A) < n-1 \end{cases}$;

(6) $\underbrace{\mathrm{adj}\cdots\mathrm{adj}}_{k}(A) = \det(A)^{\frac{(n-1)^k-(-1)^k}{n}}A^{(-1)^k}$;

(7) $\det(\underbrace{\mathrm{adj}\cdots\mathrm{adj}}_{k}(A)) = \det(A)^{(n-1)^k}$.

证明: (1)(2)(3) 均显然; (4)(5)(6)(7) 由 $A\cdot\mathrm{adj}(A) = \det(A)I_n$ 即知. □

引理 4.3.9 (Cauchy-Binet) *设 R 为一个交换环, $A \in R^{m\times n}$, $B \in R^{n\times s}$, 记 $C = A\cdot B \in R^{m\times s}$, 则*

$$\det(C_{\{i_1,\cdots,i_r\},\{j_1,\cdots,j_r\}}) = \begin{cases} 0, & r > n \\ \sum\limits_{1\leq k_1<\cdots<k_r\leq n} \det(A_{\{i_1,\cdots,i_r\},\{k_1,\cdots,k_r\}})\cdot\det(B_{\{k_1,\cdots,k_r\},\{j_1,\cdots,j_r\}}), & r \leq n \end{cases}.$$

证明: 由矩阵乘法 $C_{\{i_1,\cdots,i_r\},\{j_1,\cdots,j_r\}} = A_{\{i_1,\cdots,i_r\},\{1,\cdots,n\}}\cdot B_{\{1,\cdots,n\},\{j_1,\cdots,j_r\}}$ 知, 可不妨设 $m = r = s$. 以下只需由行列式函数关于行 (或列) 向量组的多线性、交错性与规范性即可完成证明, 具体可见维基百科. □

命题 4.3.10 *设 R 为一个交换环, $A, B \in R^{n\times n}$, 则*

(1) $\mathrm{adj}(A\cdot B) = \mathrm{adj}(B)\cdot\mathrm{adj}(A)$;

(2) *若 $AB = BA$, 则* $\mathrm{adj}(A)\cdot B = B\cdot\mathrm{adj}(A)$.

证明: (1) 一般地, 由 Cauchy-Binet 公式即知结论. 另外, 我们也可以仿照上节更具代数特色的扰动法证明如下. 记 $E = F(X_{ij}, Y_{ij})_{1\leq i,j\leq n}$ 为域 F 上多元多项式环的分式域, 以及 $A = (X_{ij})_{1\leq i,j\leq n}, B = (Y_{ij})_{1\leq i,j\leq n} \in E^{n\times n}$. 由于 $\det(A), \det(B) \in E^*$, 则 $A, B \in \mathrm{GL}(n,E)$, 故 $\mathrm{adj}(A\cdot B) = \det(AB)\cdot(AB)^{-1} = (\det(B)B^{-1})\cdot(\det(A)A^{-1}) = \mathrm{adj}(B)\cdot\mathrm{adj}(A)$. 注意等式两端的方阵分量都在环 $F[X_{ij}, Y_{ij}]_{1\leq i,j\leq n}$ 中, 故 $\mathrm{adj}(A\cdot B) = \mathrm{adj}(B)\cdot\mathrm{adj}(A)$ 对于 $A, B \in (F[X_{ij}, Y_{ij}]_{1\leq i,j\leq n})^{n\times n}$ 也成立. 特别地, 取 $X_{ij}, Y_{ij} \in R$ 则知原式成立.

(2) 与 (1) 同理. □

最后, 我们介绍域上以附属方阵为矩阵表示的线性映射. 设 V 为域 F 上的 n 维线性空间, 考虑双线性映射 $\Lambda^1(V)\times\Lambda^{n-1}(V) \longrightarrow \Lambda^n(V)$. 若将它复合线性同构

$$(f, L) \longmapsto (f\wedge L\colon (\alpha_1,\cdots,\alpha_{r+s}) \mapsto \sum_{i=1}^{n}(-1)^{i-1}f(\alpha_i)L(\alpha_1,\cdots,\widehat{\alpha_i},\cdots,\alpha_n))$$

$\Lambda^n(V)\cong F$, 则可验证上述双线性函数为非退化的, 故它诱导了线性空间的同构 $\phi\colon V^*=\Lambda^1(V)\longrightarrow(\Lambda^{n-1}(V))^*$.

$$f\longmapsto f\wedge(-)$$

再考虑它的转置映射, 利用重对偶可知 $\tau_V^{-1}\circ\phi^t\circ\tau_{\Lambda^{n-1}(V)}\colon\Lambda^{n-1}(V)\to V$ 也为线性空间的同构. 现设 $T\in L(V)$, 令 $\mathrm{adj}(T)\in L(V)$ 为 $\mathrm{adj}(T)=(\tau_V^{-1}\circ\phi^t\circ\tau_{\Lambda^{n-1}(V)})\circ T^{(n-1)}\circ(\tau_V^{-1}\circ\phi^t\circ\tau_{\Lambda^{n-1}(V)})^{-1}$, 断言: 若 T 在基 B 下的矩阵表示为 $A\in F^{n\times n}$, 则 $\mathrm{adj}(T)$ 在基 B 下的矩阵表示为 $\mathrm{adj}(A)\in F^{n\times n}$.

这是因为: 取 V 的基 $B=\{\alpha_i\}_{i=1}^n$, 以及它在 V^* 中的对偶基 $\{f_i\}_{i=1}^n$; 再任取线性同构 $\Lambda^n(V)\cong F$, 即 $\lambda\in F^*$, 记为 $cf_1\wedge\cdots\wedge f_n\mapsto\lambda c\,(c\in F)$. 任取 $1\le j\le n$, 考虑 $L_j\coloneqq(\tau_V^{-1}\circ\phi^t\circ\tau_{\Lambda^{n-1}(V)})^{-1}(\alpha_j)\in\Lambda^{n-1}(V)$, 即 $(\phi^t\circ\tau_{\Lambda^{n-1}(V)})(L_j)=\tau_V(\alpha_j)$, 则在上述同构下 $\forall 1\le i\le n,\ f_i\wedge L_j=\tau_{\Lambda^{n-1}(V)}(L_j)(f_i\wedge(-))=f_i(\alpha_j)=\delta_{ij}$, 故 $L_j=(-1)^{j-1}\lambda^{-1}f_1\wedge\cdots\wedge\widehat{f_j}\wedge\cdots\wedge f_n$. 此时

$$\begin{aligned}T^{(n-1)}(L_j)&=(-1)^{j-1}\lambda^{-1}T^t(f_1)\wedge\cdots\wedge\widehat{T^t(f_j)}\wedge\cdots\wedge T^t(f_n)\\&=(-1)^{j-1}\lambda^{-1}\left(\sum_{k_1=1}^n A_{1,k_1}f_{k_1}\right)\wedge\cdots\wedge\left(\sum_{k_j=1}^n\widehat{A_{j,k_j}f_{k_j}}\right)\wedge\cdots\wedge\left(\sum_{k_n=1}^n A_{n,k_n}f_{k_n}\right)\\&=(-1)^{j-1}\lambda^{-1}\sum_{k_1,\cdots,\widehat{k_j},\cdots,k_n=1}^n A_{1,k_1}\cdots\widehat{A_{j,k_j}}\cdots A_{n,k_n}f_{k_1}\wedge\cdots\wedge\widehat{f_{k_j}}\wedge\cdots\wedge f_{k_n}.\end{aligned}$$

再考虑 $\beta_j\coloneqq(\tau_V^{-1}\circ\phi^t\circ\tau_{\Lambda^{n-1}(V)})(T^{(n-1)}(L_j))\in V$, 即 $\tau_V(\beta_j)=(\phi^t\circ\tau_{\Lambda^{n-1}(V)})(T^{(n-1)}(L_j))$, 则在上述同构下 $\forall 1\le i\le n,\ f_i(\beta_j)=\tau_{\Lambda^{n-1}(V)}(T^{(n-1)}(L_j))(f_i\wedge(-))=f_i\wedge T^{(n-1)}(L_j)$, 代入上式知

$$\begin{aligned}f_i(\beta_j)&=(-1)^{j-1}\lambda^{-1}\sum_{k_1,\cdots,\widehat{k_j},\cdots,k_n=1}^n A_{1,k_1}\cdots\widehat{A_{j,k_j}}\cdots A_{n,k_n}f_i\wedge f_{k_1}\wedge\cdots\wedge\widehat{f_{k_j}}\wedge\cdots\wedge f_{k_n}\\&=(-1)^{j-1}\lambda^{-1}\sum_{\sigma\colon\{1,\cdots,n\}\setminus\{j\}\to\{1,\cdots,n\}\setminus\{i\}\text{为双射}}A_{1,\sigma(1)}\cdots\widehat{A_{j,\sigma(j)}}\cdots A_{n,\sigma(n)}f_i\wedge f_{\sigma(1)}\wedge\cdots\wedge\widehat{f_{\sigma(j)}}\wedge\cdots\wedge f_{\sigma(n)}\\&=(-1)^{j+i}\lambda^{-1}\sum_{\sigma\colon\{1,\cdots,n\}\setminus\{j\}\to\{1,\cdots,n\}\setminus\{i\}\text{为双射}}\mathrm{sgn}(\sigma)A_{1,\sigma(1)}\cdots\widehat{A_{j,\sigma(j)}}\cdots A_{n,\sigma(n)}f_1\wedge\cdots\wedge f_n\\&=\lambda^{-1}\mathrm{adj}(A)_{ij}f_1\wedge\cdots\wedge f_n,\end{aligned}$$

故 $f_i(\beta_j)=\mathrm{adj}(A)_{ij}$, 即 $\beta_j=\sum\limits_{i=1}^n\mathrm{adj}(A)_{ij}\alpha_i$. 综上可知, $\forall 1\le j\le n,\ \mathrm{adj}(T)(\alpha_j)=\sum\limits_{i=1}^n\mathrm{adj}(A)_{ij}\alpha_i$.

4.3.5 交错方阵的行列式

本节将以初等的方式讨论交错方阵的行列式, 它的表达式在微分几何中起着很重要的作用.

定义 4.3.2 (交错方阵) 设 R 为一个环, $A\in R^{n\times n}$, 若 $A+A^t=0$ 且 A 的对角分量均为 0, 则称 A 为**交错阵** (alternating matrix).

注:

(1) 记

$$\begin{aligned}\mathrm{Alt}(n,R)&\coloneqq\{A\in R^{n\times n}\colon A+A^t=0;\ \forall 1\le i\le n,\ A_{ii}=0\},\\\mathrm{Skew}(n,R)&\coloneqq\{A\in R^{n\times n}\colon A+A^t=0\},\\\mathrm{Sym}(n,R)&\coloneqq\{A\in R^{n\times n}\colon A=A^t\},\end{aligned}$$

分别为环 R 上的交错、反对称、对称方阵构成的模. 注意当 R 为交换环时, $[\mathrm{Skew}(n,R),\mathrm{Skew}(n,R)]\subseteq\mathrm{Alt}(n,R)$, 故当 $R=F$ 为一个域时, 前两者在括积 $[\cdot,\cdot]$ 下还构成域 F 上的 Lie 代数, 而第三者关于括积 $[\cdot,\cdot]$ 一般不封闭.

(2) 当 $\mathrm{char}(F)\ne 2$ 时, $\mathrm{Alt}(n,F)=\mathrm{Skew}(n,F)$, 且存在线性空间分解 $F^{n\times n}=\!=\mathrm{Skew}(n,F)\oplus\mathrm{Sym}(n,F)$;

$$A\longmapsto\left(\frac{A-A^t}{2},\frac{A+A^t}{2}\right)$$

当 $\mathrm{char}(F)=2$ 时, $\mathrm{Alt}(n,F)\subsetneqq\mathrm{Skew}(n,F)=\mathrm{Sym}(n,F)$, 此时 $I_n\in\mathrm{Skew}(n,F)\setminus\mathrm{Alt}(n,F)$.

现在开始计算交错方阵的行列式, 以下这个引理是讨论的基础.

引理 4.3.11 设 R 为一个交换环, 记 $\xi_n \coloneqq \{\sigma \in S_n \colon \sigma$可分解成若干个不相交的长为偶数的轮换的复合$\}$ (注意这里将任意 S_n 中元都分解成若干个不相交的轮换 (包括长度为 1 的轮换) 的复合, 此表达式在不计复合顺序的意义下存在唯一), 则 $\forall A \in \mathrm{Alt}(n,R)$, $\det(A) = \sum\limits_{\sigma\in\xi_n} \mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}$.

证明: 设 $\sigma \in S_n \setminus \xi_n$, 则在 σ 分解成若干个不相交的轮换的复合表达式中, 存在长为奇数的轮换. 若此式中存在长为 1 的轮换, 即 $\exists 1 \le i \le n,\ s.t.\ \sigma(i) = i$, 则由 $A_{ii} = 0$ 知, $\mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} = 0$. 现考虑集合 $\{\sigma \in S_n \setminus \xi_n \colon$
$\forall 1 \le i \le n,\ \sigma(i) \ne i\}$, 并将其中元唯一地记为 $\sigma = \sigma_1 \circ \cdots \circ \sigma_s$, 这里 $\sigma_1, \cdots, \sigma_s$ 为若干个不相交的轮换, 且 $i < j \iff \sigma_i$ 中涉及的最小数小于 σ_j 中涉及的最小数; 再记 $i_\sigma \coloneqq \min\{1 \le i \le s \colon \sigma_i$为长为奇数的轮换$\}$. 于是此集合中元可两两配对: $\sigma \sim \sigma' \iff \sigma' = \sigma_1 \circ \cdots \circ \sigma_{i_\sigma - 1} \circ \sigma_{i_\sigma}^{-1} \circ \sigma_{i_\sigma + 1} \circ \cdots \circ \sigma_s$, 此时 $\sigma \ne \sigma'$, 且 $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma')$, $A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} = -A_{1,\sigma'(1)} \cdots A_{n,\sigma'(n)}$, 故 $\mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} + \mathrm{sgn}(\sigma') A_{1,\sigma'(1)} \cdots A_{n,\sigma'(n)} = 0$. 综上, $\sum\limits_{\sigma \in S_n \setminus \xi_n} \mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} = 0$, 因此 $\det(A) = \sum\limits_{\sigma\in\xi_n} \mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)}$. □

推论 4.3.12 设 R 为一个交换环, $n \in \mathbb{N}^*$ 为奇数, 则 $\forall A \in \mathrm{Alt}(n,R)$, $\det(A) = 0$.

证明: 由 $n \in \mathbb{N}^*$ 为奇数知, $\xi_n = \varnothing$, 故由上述引理即知. □

推论 4.3.13 设 F 为一个域, $n \in \mathbb{N}^*$, 则 $\forall A \in \mathrm{Alt}(n,F)$, $r(A)$ 为偶数.

证明: 记 $r = r(A)$. 若 $r = 0$, 则结论显然. 现设 $1 \le r \le n$, 则可取 A 的 r 个线性无关行, 行指标为 $i_1, \cdots, i_r$; 由 A 的反对称性知, A 的第 $i_1, \cdots, i_r$ 列也线性无关. 取这 r 行与 r 列交点处的子矩阵 $A_0 \in F^{r\times r}$, 则 $A_0 \in \mathrm{Alt}(r,F)$ 且 $A_0 \in \mathrm{GL}(r,F)$, 故由上述推论知, r 为偶数. □

更不平凡的计算在于偶数阶交错方阵的行列式, 它与所谓的 Pfaffian 表达式有关. Pfaffian 的一般定义与计算依赖于以下引理.

引理 4.3.14 设 R 为一个交换环, $n \in \mathbb{N}^*$ 为偶数, $\sigma \in S_n$, $A \in \mathrm{Alt}(n,R)$. 记 $M_\sigma \coloneqq \{\{\sigma(2i-1), \sigma(2i)\} \colon 1 \le i \le n/2\}$, $w(A, M_\sigma) \coloneqq \mathrm{sgn}(\sigma) \prod\limits_{i=1}^{n/2} A_{\sigma(2i-1),\sigma(2i)}$, 则当 $M_\sigma = M_\tau$ 时, $w(A,M_\sigma) = w(A,M_\tau)$.

证明: 设 $M_\sigma = M_\tau$, 记 $\tau = \eta \circ \sigma$, 则 η 为若干个以下两种置换的复合: ① 对换 $(\sigma(2i-1), \sigma(2i))\,(1 \le i \le n/2)$; ② 两个对换的复合 $(\sigma(2i-1), \sigma(2j-1)) \circ (\sigma(2i), \sigma(2j))\,(1 \le i < j \le n/2)$. 对于 $w(A, \mathscr{F}_\sigma)$ 的表达式而言, 第一种置换将 $\mathrm{sgn}(\sigma)$ 变为 $-\mathrm{sgn}(\sigma)$, 且将 $A_{\sigma(2i-1),\sigma(2i)}$ 变为 $A_{\sigma(2i),\sigma(2i-1)} = -A_{\sigma(2i-1),\sigma(2i)}$, 故不改变表达式的值; 第二种置换将 $\mathrm{sgn}(\sigma)$ 仍变为 $\mathrm{sgn}(\sigma)$, 且将 $A_{\sigma(2i-1),\sigma(2i)} A_{\sigma(2j-1),\sigma(2j)}$ 变为 $A_{\sigma(2j-1),\sigma(2j)} A_{\sigma(2i-1),\sigma(2i)}$, 故不改变表达式的值. 综上, $w(A,M_\sigma) = w(A,M_\tau)$. □

定义 4.3.3 (Pfaffian) 设 R 为一个交换环, $n \in \mathbb{N}^*$ 为偶数, $A \in \mathrm{Alt}(n,R)$. 记 $\mathscr{M}_n \coloneqq \{M_\sigma \colon \sigma \in S_n\}$ (已去掉重复项), 则 $\mathrm{Pf}(A) \coloneqq \sum\limits_{M\in\mathscr{M}_n} w(A,M) \in R$ 称为交错方阵 A 的 Pfaffian.

注: 由于每个 $M \in \mathscr{M}_n$ 恰对应 $2^{n/2}\cdot(n/2)!$ 个 M_σ $(\sigma \in S_n)$, 故当 $2^{n/2}\cdot(n/2)! \in R^\times$ 时, $\mathrm{Pf}(A) = \dfrac{1}{2^{n/2}\cdot(n/2)!} \sum\limits_{\sigma\in S_n} w(A,M_\sigma)$.

引理 4.3.15 设 $n \in \mathbb{N}^*$ 为偶数, 则存在双射 $\xi_n \to \mathscr{M}_n \times \mathscr{M}_n$.

证明: 任取 $\sigma \in \xi_n$, 记 $\sigma = \sigma_1 \circ \cdots \circ \sigma_s$, 这里 $\sigma_1, \cdots, \sigma_s$ 为若干个不相交的长为偶数的轮换, 且 $i < j \iff \sigma_i$ 中涉及的最小数小于 σ_j 中涉及的最小数; 再记 σ_i 中涉及的最小数为 n_i, σ_i 的长度为 l_i. 令

$$M \coloneqq \{\{n_1, \sigma_1(n_1)\}, \cdots, \{\sigma_1^{l_1-2}(n_1), \sigma_1^{l_1-1}(n_1)\}, \cdots, \{n_s, \sigma_s(n_s)\}, \cdots, \{\sigma_s^{l_s-2}(n_s), \sigma_s^{l_s-1}(n_s)\}\},$$

$$M' \coloneqq \{\{\sigma_1(n_1), \sigma_1^2(n_1)\}, \cdots, \{\sigma_1^{l_1-1}(n_1), \sigma_1^{l_1}(n_1)\}, \cdots, \{\sigma_s(n_s), \sigma_s^2(n_s)\}, \cdots, \{\sigma_s^{l_s-1}(n_s), \sigma_s^{l_s}(n_s)\}\},$$

则 M, M' 均为由 $\dfrac{l_1 + \cdots + l_s}{2} = \dfrac{n}{2}$ 个不相交的二元集合构成的集族, 即 $(M, M') \in \mathscr{M}_n \times \mathscr{M}_n$. 由此过程可逆知, $\sigma \mapsto (M, M')$ 给出了双射 $\xi_n \to \mathscr{M}_n \times \mathscr{M}_n$. □

推论 4.3.16 设 R 为一个交换环, $n \in \mathbb{N}^*$ 为偶数, 则 $\forall A \in \mathrm{Alt}(n,R)$, $\det(A) = \mathrm{Pf}(A)^2$.

证明: 由以上引理知, 只需证明: 对于 $\sigma \in \xi_n$ 以及对应的 $(M, M') \in \mathscr{M}_n \times \mathscr{M}_n$,

$$\mathrm{sgn}(\sigma) A_{1,\sigma(1)} \cdots A_{n,\sigma(n)} = w(A,M) w(A,M').$$

事实上, 取 $\tau,\tau'\in S_n$ 满足 $M_\tau=M$, $M_{\tau'}=M'$, 则 $\mathrm{sgn}(\tau')=(-1)^s\mathrm{sgn}(\tau)$, $w(A,M)\coloneqq\mathrm{sgn}(\tau)\prod\limits_{i=1}^{n/2}A_{\tau(2i-1),\tau(2i)}$, $w(A,M')\coloneqq\mathrm{sgn}(\tau')\prod\limits_{i=1}^{n/2}A_{\tau'(2i-1),\tau'(2i)}$, 且 $\prod\limits_{i=1}^{n/2}A_{\tau(2i-1),\tau(2i)}\prod\limits_{i=1}^{n/2}A_{\tau'(2i-1),\tau'(2i)}=\prod\limits_{i=1}^{n/2}A_{\sigma(2i-1),\sigma(2i)}$. □

最后我们从交错双线性函数的角度给出 Pfaffian 的另一种解释.

设 F 为一个域, $A\in F^{n\times n}$, 则它决定了一个双线性函数 $\omega_A\colon\ F^{n\times1}\times F^{n\times1}\longrightarrow F\quad;\ (\alpha,\beta)\longmapsto\alpha^tA\beta$ 此双线性函数 ω_A 是交错的 (或反对称的, 或对称的), 当且仅当方阵 A 是交错的 (或反对称的, 或对称的). 现设 $n\in\mathbb{N}^*$ 为偶数, 且 $A\in\mathrm{Alt}(n,F)$, 则 $\omega_A\in\Lambda^2(F^{n\times1})$, 故 $\bigwedge^{n/2}\omega_A\in\Lambda^n(F^{n\times1})$. 注意 $\dim_F(\Lambda^n(F^{n\times1}))=1$, 则 $\bigwedge^{n/2}\omega_A$ 为某个 "标准的" 体积形式的常数倍. 取 $\{\alpha_i\}_{i=1}^n$ 为 $F^{n\times1}$ 的标准基, 即 α_i 为第 i 个分量为 1, 其余分量均为 0 的列向量, 以及 $\{f_i\}_{i=1}^n\subseteq(F^{n\times1})^*$ 为它的对偶基, 则 $\exists!c=c(A)\in F,\ s.t.\ \bigwedge^{n/2}\omega_A=cf_1\wedge\cdots\wedge f_n$. 断言: $c(A)=(n/2)!\mathrm{Pf}(A)$. 这是因为:

$$\begin{aligned}c(A)&=\bigwedge^{n/2}\omega_A(\alpha_1,\cdots,\alpha_n)\\&=\sum_{\sigma\in\mathrm{Sh}(2,\cdots,2)}\mathrm{sgn}(\sigma)\omega_A(\alpha_{\sigma(1)},\alpha_{\sigma(2)})\cdots\omega_A(\alpha_{\sigma(n-1)},\alpha_{\sigma(n)})\\&=\sum_{\sigma\in\mathrm{Sh}(2,\cdots,2)}\mathrm{sgn}(\sigma)A_{\sigma(1),\sigma(2)}\cdots A_{\sigma(n-1),\sigma(n)}\\&=(n/2)!\sum_{M\in\mathscr{M}_n}w(A,M).\end{aligned}$$

这种解释在微分几何中有很深刻的背景. 具体地说, 设 (M,g) 为定向闭 Riemann 流形, 维数 n 为偶数. 取流形 M 的局部坐标覆盖 $\{U_\alpha\}_{\alpha\in\mathcal{A}}$, 在 U_α 上的曲率形式为 $\Omega_\alpha=((\Omega_\alpha)^i_j)_{n\times n}\in\mathrm{Skew}(n,\Omega^2(U_\alpha))$, 其中 $\Omega^r(U_\alpha)\coloneqq\Gamma(\Lambda^r(TU_\alpha))$. 令 $\mathrm{Pf}(\Omega)_\alpha=\dfrac{1}{2^{n/2}(n/2)!}\sum\limits_{\sigma\in S_n}\mathrm{sgn}(\sigma)(\Omega_\alpha)^{\sigma(1)}_{\sigma(2)}\wedge\cdots\wedge(\Omega_\alpha)^{\sigma(n-1)}_{\sigma(n)}\in\Omega^n(U_\alpha)$, 可以验证 $\{\mathrm{Pf}(\Omega)_\alpha\}_{\alpha\in\mathcal{A}}$ 给出了流形 M 上整体定义的 n 形式 $\mathrm{Pf}(\Omega)$, 称为曲率形式 Ω 的 Pfaffian. 著名的 Gauss-Bonnet-Chern 公式即 $\int_M\mathrm{Pf}(\Omega)=(2\pi)^{n/2}\chi(M)$, 其中 $\chi(M)\coloneqq\sum\limits_{r=0}^n(-1)^r\dim_{\mathbb{R}}H^r_{\mathrm{dR}}(M)$ 为流形 M 的 Euler 示性数, 它的意义在于将流形的曲率与拓扑联系在一起.

习题 4.3 设 F 为一个域, $n\in\mathbb{N}^*$ 为偶数, $A,B\in F^{n\times n}$, 证明: $\mathrm{Pf}(BAB^t)=\det(B)\mathrm{Pf}(A)$.

参考文献与补注 4.3

(1) 关于常微分方程的基本理论, 可以参考 V. Arnold"Ordinary Differential Equations".

(2) 关于 Riemann 流形的基本理论, 可以参考梅加强 "流形与几何初步".

§4.4 行列式的计算

本节介绍若干行列式的计算方法, 并由此总结线性代数中部分矩阵运算的技巧.

第一种计算方法是利用行列式的多项式性质, 它与方阵的特征值理论密切相关.

例 4.4.1 (首一多项式的友矩阵) 设 F 为一个域, $f(X)=X^n+a_{n-1}X^{n-1}+\cdots+a_1X+a_0\in F[X]$ 为首一多项式, 则存在 $A\in F^{n\times n}$, 使得 $f(X)=\det(X\cdot I_n-A)$.

证明: 取 $A\coloneqq\begin{pmatrix}0&&&&-a_0\\1&0&&&-a_1\\&\ddots&\ddots&&\vdots\\&&\ddots&0&-a_{n-2}\\&&&1&-a_{n-1}\end{pmatrix}$, 通过比较系数或按最后一列展开即知结论. □

注: 上述证明中的方阵 $A\in F^{n\times n}$ 称为首一多项式 $f(X)\in F[X]$ 的**友矩阵** (companion matrix), 它在矩阵的有理标准形理论中起着很关键的作用.

定义 4.4.1 (方阵的特征多项式) 设 F 为一个域, $A \in F^{n\times n}$, 则 $f_A(X) \coloneqq \det(X\cdot I_n - A) \in F[X]$ 是首一多项式, 称为方阵 A 的**特征多项式** (characteristic polynomial); $\sigma(A) \coloneqq \{c \in F\colon f_A(c) = 0\}$ 称为方阵 A 的**谱集** (spectrum).

注:

(1) 上例表明任意首一多项式均为某个方阵的特征多项式; 但不同的方阵可能具有相同的特征多项式, 比如在相似变换下方阵的特征多项式不变.

(2) 显然 $|\sigma(A)| \le \deg(f_A(X))$; 另一方面, 由于方阵的特征多项式未必在系数域上分裂, 故方阵的谱集可能为空.

(3) $\sigma(A)$ 中点称为方阵 A 在域 F 中的**特征值** (eigenvalue). 设 $c \in F$, 则 $c \in \sigma(A) \iff \ker(cI_n - A) \neq \{0\}$, 此时 $\ker(cI_n - A)$ 中的非零向量称为方阵 A 的属于特征值 $c \in F$ 的**特征向量** (eigenvector).

引理 4.4.1 *同一方阵的属于不同特征值的特征向量线性无关.*

证明: 设 $A \in F^{n\times n}$, $\lambda_1, \cdots, \lambda_s \in \sigma(A)$ 两两不同, $\alpha_i \in \ker(\lambda_i I_n - A)\setminus\{0\}\,(1 \le i \le s)$. 假设 $\exists c_1, \cdots, c_s \in F$, $s.t.\ \sum\limits_{i=1}^{s} c_i\alpha_i = 0$, 则 $\forall 0 \le j \le s-1$, $0 = \sum\limits_{i=1}^{s} c_i A_j(\alpha_i) = \sum\limits_{i=1}^{s} c_i\lambda_i^j\alpha_i$, 即 $(0,\cdots,0) = (c_1\alpha_1,\cdots,c_s\alpha_s)\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \lambda_s & \cdots & \lambda_s^{s-1} \end{pmatrix}$.
由于右端方阵是可逆的 Vandermonde 阵, 故 $\forall 1 \le i \le s$, $c_i\alpha_i = 0$, 即 $c_i = 0$, 则 $\{\alpha_1, \cdots, \alpha_s\}$ 线性无关. □

推论 4.4.2 *设 $A \in F^{n\times n}$, 则 $|\sigma(A)| \le n$; 且 $|\sigma(A)| = n \iff A$ 可相似于一个对角分量两两不同的对角阵.*

当然, 行列式也可以视为关于方阵中分量的多元多项式, 这一观点常能收获意想不到的效果.

例 4.4.2 *设 F 为一个域, $A \in F^{n\times n}$, $P \in \mathrm{GL}(n,F)$ 为若干 $P(i,j)$(即交换 i,j 行) 型初等方阵的乘积, $\widetilde{A} = PAP^t$, 则 A_{ij} 在 A 中的代数余子式等于 $\widetilde{A}_{kl} = A_{ij}$ 在 $\widetilde{A}$ 中的代数余子式.*

证明: 考虑将 A 的行列式按第 i 行展开 $\det(A) = \sum\limits_{\substack{j'=1\\ j'\neq j}}^{n} A_{ij'}C_{ij'} + A_{ij}C_{ij}$, 以及将 $\widetilde{A}$ 的行列式按第 k 行展开 $\det(\widetilde{A}) = \sum\limits_{\substack{l'=1\\ l'\neq l}}^{n} \widetilde{A}_{kl'}\widetilde{C}_{kl'} + \widetilde{A}_{kl}\widetilde{C}_{kl}$, 其中 $A_{ij} = \widetilde{A}_{kl}$, 以下只需证明 $C_{ij} = \widetilde{C}_{kl}$. 由于 $\det(P) = \pm 1$, 则 $\det(A) = \det(\widetilde{A})$, 即它们的按行展开式相等. 现将 A_{ij} 视为未定元 X_{ij}, 将 $A = (A_{ij})_{1\le i,j\le n}$ 视为多元多项式环 $F[X_{ij}]_{1\le i,j\le n}$ 上的方阵, 则上述按行展开式相等仍成立. 注意交换行列的变换不改变方阵中分量是否在同一行或同一列, 故在按行展开式中仅有 A_{ij} 项与 $\widetilde{A}_{kl}$ 含未定元 X_{ij}. 再将等式两端关于 X_{ij} 求导知, $C_{ij} = \widetilde{C}_{kl}$. □

第二种计算方法是利用行列式的组合定义式, 它常可给出方阵行列式的定性判断.

例 4.4.3 *设 $A \in F^{n\times n}$, 若在第 $i_1 < \cdots < i_k$ 行与第 $j_1 < \cdots < j_l$ 列的交叉位置上 A 的分量均为 0, 且 $k + l > n$, 则 $\det(A) = 0$.*

证明: 将 A 的行列式按第 $\{i_1, \cdots, i_k\}$ 行展开, 则每个子式都为 0. □

例 4.4.4 *设 $\mathrm{char}(F) \neq 2$, 则 $\max\{k \in \mathbb{N}\colon 2^k \mid \det(A),\ \forall A \in \{\pm 1\}^{n\times n}\} = n - 1$.*

证明: 一方面, 任取 $A \in \{\pm 1\}^{n\times n}$, 通过将 A 的第一行的 ± 1 倍加往其余行, 可使得其余行的第一个分量均为 0, 此时其余分量均在 $\{0, \pm 2\}$ 中. 这表明可通过初等行变换将 A 变为 $\begin{pmatrix} \pm 1 & * \\ 0 & A_1 \end{pmatrix}$, 其中 $A_1 \in \{0, \pm 2\}^{n\times n}$, 则由组合定义式知, $2^{n-1} \mid \det(A_1) = \pm\det(A)$.

另一方面, 对 $n \ge 1$ 归纳证明: $\exists A \in \{\pm 1\}^{n\times n}$, $s.t.\ 2^n \nmid \det(A)$. 当 $n = 1$ 时, 取 $A = (1)$ 即可; 现设 $n \ge 2$ 且当 $(n-1)$ 时结论成立, 即 $\exists A_1 \in \{\pm 1\}^{(n-1)\times(n-1)}$, $s.t.\ 2^{n-1} \nmid \det(A_1)$. 记 A_1 的第一行为 $\alpha_1 \in F^{1\times(n-1)}$, 取 $A = \begin{pmatrix} 1 & \alpha_1 \\ \beta_1 & A_1 \end{pmatrix} \in \{\pm 1\}^{n\times n}$, 其中 $\beta_1 = (-1, \cdots, -1)^t \in F^{(n-1)\times 1}$, 则将 A 的行列式按第一列展开知, $\det(A) = 2\det(A_1)$, 故 $2^n \nmid \det(A)$. □

注: 历史上, 求 $\max\{|\det(A)| \in \mathbb{N} : A \in \{\pm 1\}^{n\times n}\}$ 是一个公开的问题, 称为 Hadamard 最大行列式问题. Hadamard 在文章 "Resolution d'une question relative aux determinants"(1893) 中给出的上界为 $n^{n/2}$.

例 4.4.5 (循环矩阵的秩与行列式) 设 $A=\begin{pmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1} \\ a_{n-1} & \cdots & a_1 & a_0 \end{pmatrix}\in\mathbb{C}^{n\times n}$, 求 $r(A)$ 与 $\det(A)$.

证明: 记 $J=\begin{pmatrix} 0 & & & 1 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}$, 以及 $f(X)=a_0+a_1X+\cdots+a_{n-1}X^{n-1}$, 则 $A=f(J)$. 由于 $J^n=I_n$, 则 $f_J(X)=X^n-1$, 故 $\sigma(J)=\{\omega\in\mathbb{C}\colon \omega^n=1\}$; 特别地, $|\sigma(J)|=n$. 由下述引理知, $\sigma(f(J))=\{f(\omega)\in\mathbb{C}\colon \omega^n=1\}$, 则 $r(f(J))=n-|\{\omega\in\mathbb{C}\colon \omega^n=1,\ f(\omega)=0\}|=n-\deg(\gcd(X^n-1,f(X)))$; $\det(f(J))=\prod\limits_{\substack{\omega\in\mathbb{C}\\ \omega^n=1}} f(\omega)$. □

引理 4.4.3 设 $A\in F^{n\times n}$ 且 A 在域 F 上可上三角化, 则 $\forall f(X)\in F[X]$, $\sigma(f(A))=f(\sigma(A))$.

证明: 由 A 在域 F 上可上三角化知, $\exists P\in\mathrm{GL}(n,F)$, $\lambda_1,\cdots,\lambda_n\in F$, $s.t.$ $P^{-1}AP=\begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ & & \lambda_n \end{pmatrix}$, 则 $\forall f(X)\in F[X]$, $P^{-1}f(A)P=f(P^{-1}AP)=\begin{pmatrix} f(\lambda_1) & * & * \\ & \ddots & * \\ & & f(\lambda_n) \end{pmatrix}$, 故 $\sigma(f(A))=\{f(\lambda_1),\cdots,f(\lambda_n)\}=f(\sigma(A))$.

□

推论 4.4.4 设 $a_0,\cdots,a_{n-1}\in\mathbb{Z}$, 则 $\prod\limits_{\substack{\omega\in\mathbb{C}\\ \omega^n=1}}\sum\limits_{i=0}^{n-1} a_i\omega^i\in\mathbb{Z}$.

推论 4.4.5 设 $A(c)=\begin{pmatrix} a_0 & ca_{n-1} & \cdots & ca_1 \\ a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & ca_{n-1} \\ a_{n-1} & \cdots & a_1 & a_0 \end{pmatrix}\in\mathbb{C}^{n\times n}$, 求 $r(A(c))$ 与 $\det(A(c))$.

证明: 当 $c=0$ 时, 显然 $r(A(c))=n-\min\{i\in\mathbb{N}^*\colon a_i\neq 0\}$; $\det(A(c))=(1-\delta_{0,a_0})a_0^n$.

当 $c\neq 0$ 时, 固定 $\sqrt[n]{c}\in\mathbb{C}$, 以及 $\{\omega\in\mathbb{C}\colon \omega^n=1\}=\{\omega_i\}_{i=1}^n$. 记 $P(c)=\begin{pmatrix} 1 & \sqrt[n]{c}\omega_1 & \cdots & (\sqrt[n]{c}\omega_1)^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \sqrt[n]{c}\omega_n & \cdots & (\sqrt[n]{c}\omega_n)^{n-1} \end{pmatrix}$, 以及 $f(X)=a_0+a_1X+\cdots+a_{n-1}X^{n-1}$, 则 $P(c)A(c)=\mathrm{diag}(f(\sqrt[n]{c}\omega_1),\cdots,f(\sqrt[n]{c}\omega_n))P(c)$. 由于 $P(c)$ 为可逆的 Vandermonde 阵, 则 $r(A(c))=n-\deg(\gcd(X^n-1,f(\sqrt[n]{c}X)))$; $\det(A(c))=\prod\limits_{\substack{\omega\in\mathbb{C}\\ \omega^n=1}} f(\sqrt[n]{c}\omega)$. □

第三种计算方法是利用行列式关于行 (或列) 向量组的多线性与交错性, 它可将行列式简化为标准的箭形.

例 4.4.6 (箭形行列式) 求 $\det\begin{pmatrix} a_1 & c_2 & \cdots & c_n \\ b_2 & a_2 & & \\ \vdots & & \ddots & \\ b_n & & & a_n \end{pmatrix}$.

证明: 先设 $a_i\neq 0\,(2\leq i\leq n)$, 则 $\det\begin{pmatrix} a_1 & c_2 & \cdots & c_n \\ b_2 & a_2 & & \\ \vdots & & \ddots & \\ b_n & & & a_n \end{pmatrix}\xrightarrow[\text{加往第 1 行}]{\text{将第 } i\geq 2 \text{ 行的 } -\frac{c_i}{a_i} \text{ 倍}}\det\begin{pmatrix} a_1-\sum\limits_{i=2}^n\frac{b_ic_i}{a_i} & 0 & \cdots & 0 \\ b_2 & a_2 & & \\ \vdots & & \ddots & \\ b_n & & & a_n \end{pmatrix}$

$=\prod\limits_{i=1}^n a_i-\sum\limits_{i=2}^n b_ic_i\cdot\prod\limits_{\substack{j\geq 2\\ j\neq i}}^n a_j$. 由扰动法可知一般情形成立. □

例 4.4.7 求 $\det\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{pmatrix}$.

证明: $\det\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{pmatrix} \xlongequal[\text{加往第 } i \text{ 行 } (i\geq 1)]{\text{将第 } (i+1) \text{ 行的 } (-1) \text{ 倍}} \det\begin{pmatrix} -a_1 & a_2 & & & \\ & -a_2 & a_3 & & \\ & & \ddots & \ddots & \\ & & & -a_{n-1} & a_n \\ x_1 & x_2 & \cdots & x_{n-1} & x_n-a_n \end{pmatrix}$

$\xlongequal[\text{加往第 } i \text{ 行 } (i\geq 1)]{\text{将第 } i+1,\cdots,n-1 \text{ 行}} \det\begin{pmatrix} -a_1 & & & & a_n \\ & -a_2 & & & a_n \\ & & \ddots & & \vdots \\ & & & -a_{n-1} & a_n \\ x_1 & x_2 & \cdots & x_{n-1} & x_n-a_n \end{pmatrix} \xlongequal{\text{箭形}} (-1)^n\prod_{i=1}^{n}a_i+(-1)^{n-1}\sum_{i=1}^{n}\left(\prod_{\substack{j\geq 1\\ i\neq j}}^{n}a_j\right)x_i.$ □

注: 上述行列式也可逐行拆行证明.

第四种计算方法是利用数学归纳法, 它可将有规律的行列式写成一阶或二阶递推数列进而求解.

例 4.4.8 (一阶递推关系) 记 $A_n=\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{pmatrix}$ $(n\geq 1)$, 求 $\det(A_n)$.

证明:

$$\det(A_n)\xlongequal{\text{拆最后一行}}\det\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}+\det\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -a_n \end{pmatrix}$$

$$\xlongequal[\text{(2) 按最后一行展开}]{\text{(1) 将最后一行的 } (-1) \text{ 倍 加往第 } i \text{ 行 } (1\leq i\leq n-1)}\det\begin{pmatrix} -a_1 & & & & \\ & -a_2 & & & \\ & & \ddots & & \\ & & & -a_{n-1} & \\ x_1 & x_2 & \cdots & x_{n-1} & x_n \end{pmatrix}+(-a_n)\det(A_{n-1})$$

$$=(-1)^{n-1}a_1\cdots a_{n-1}x_n+(-a_n)\det(A_{n-1})\,(n\geq 1).$$

再由 $\det(A_1)=x_1-a_1$, 一阶递推即知 $\det(A_n)=(-1)^n\prod_{i=1}^{n}a_i+(-1)^{n-1}\sum_{i=1}^{n}\left(\prod_{\substack{j\geq 1\\ i\neq j}}^{n}a_j\right)x_i$. □

例 4.4.9 (Vandermonde 行列式) 记 $V_n=\begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{pmatrix}$ $(n\geq 1)$, 证明: $\det(V_n)=\prod_{1\leq i<j\leq n}(a_j-a_i)$.

证明:

$$\det(V_n)\xlongequal{\substack{\text{将第 } i \text{ 列的 } (-a_1) \text{ 倍}\\ \text{加往第 } (i+1) \text{ 列 } (1\le i\le n-1)}}\det\begin{pmatrix}1 & 0 & \cdots & 0\\ 1 & a_2-a_1 & \cdots & a_2^{n-2}(a_2-a_1)\\ \vdots & \vdots & \cdots & \vdots\\ 1 & a_n-a_1 & \cdots & a_n^{n-2}(a_n-a_1)\end{pmatrix}$$

$$\xlongequal{\text{按第 1 列展开}}\det\begin{pmatrix}a_2-a_1 & \cdots & a_2^{n-2}(a_2-a_1)\\ \vdots & \cdots & \vdots\\ a_n-a_1 & \cdots & a_n^{n-2}(a_n-a_1)\end{pmatrix}\xlongequal{\text{提取各行的公因式}}\prod_{j=2}^{n}(a_j-a_1)\cdot\det(V_{n-1})\,(n\ge 1).$$

再由 $\det(V_1)=1$, 一阶递推即知 $\det(V_n)=\prod\limits_{1\le i<j\le n}(a_j-a_i)$. □

推论 4.4.6 (缺项 Vandermonde 行列式) 记 $A_n=\begin{pmatrix}1 & a_1 & \cdots & a_1^{k-1} & a_1^{k+1} & \cdots & a_1^n\\ \vdots & \vdots & \ldots & \vdots & \vdots & \ldots & \vdots\\ 1 & a_n & \cdots & a_n^{k-1} & a_n^{k+1} & \cdots & a_n^{n-1}\end{pmatrix}(1\le k\le n)$, 求 $\det(A_n)$.

证明: 记 $\widetilde{A_n}(X)=\begin{pmatrix}1 & a_1 & \cdots & a_1^{k-1} & a_1^k & a_1^{k+1} & \cdots & a_1^n\\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots\\ 1 & a_n & \cdots & a_n^{k-1} & a_n^k & a_n^{k+1} & \cdots & a_n^n\\ 1 & X & \cdots & X^{k-1} & X^k & X^{k+1} & \cdots & X^n\end{pmatrix}$, 则由 Vandermonde 行列式知,

$\det(\widetilde{A_n}(X))=\prod\limits_{1\le i<j\le n}(a_j-a_i)\cdot\prod\limits_{i=1}^{n}(X-a_i)$. 注意左端按最后一行展开, 关于 X^k 的系数为 $(-1)^{(n+1)+(k+1)}\det(A_n)$; 右端由韦达定理, 关于 X^k 的系数为 $\prod\limits_{1\le i<j\le n}(a_j-a_i)\cdot\sum\limits_{1\le i_1<\cdots<i_{n-k}\le n}(-1)^{n-k}a_{i_1}\cdots a_{i_{n-k}}$. 因此比较系数知,

$\det(A_n)=\prod\limits_{1\le i<j\le n}(a_j-a_i)\cdot\sum\limits_{1\le i_1<\cdots<i_{n-k}\le n}a_{i_1}\cdots a_{i_{n-k}}$. □

例 4.4.10 (Cauchy 行列式) 记 $A_n=\left(\dfrac{1}{a_i+b_j}\right)_{1\le i,j\le n}(n\ge 1)$, 证明: $\det(A_n)=\dfrac{\prod\limits_{1\le i<j\le n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\le i,j\le n}(a_i+b_j)}$.

证明:

$$\det\left(\frac{1}{a_i+b_j}\right)_{1\le i,j\le n}\xlongequal{\substack{\text{将第 1 行的 } -1 \text{ 倍}\\ \text{加往第 } (i+1) \text{ 行 } (1\le i\le n-1)}}\det\begin{pmatrix}\frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n}\\ \frac{a_2-a_1}{(a_1+b_1)(a_2+b_1)} & \cdots & \frac{a_2-a_1}{(a_1+b_n)(a_2+b_n)}\\ \vdots & \vdots & \vdots\\ \frac{a_n-a_1}{(a_1+b_1)(a_n+b_1)} & \cdots & \frac{a_n-a_1}{(a_1+b_n)(a_n+b_n)}\end{pmatrix}$$

$$\xlongequal{\text{提取各行、列的公因式}}\frac{\prod\limits_{j=2}^{n}(a_j-a_1)}{\prod\limits_{j=1}^{n}(a_1+b_j)}\det\begin{pmatrix}1 & \cdots & 1\\ \frac{1}{a_2+b_1} & \cdots & \frac{1}{a_2+b_n}\\ \vdots & \vdots & \vdots\\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n}\end{pmatrix}.$$

$$\xlongequal{\substack{\text{将第 1 列的 } -1 \text{ 倍}\\ \text{加往第 } (i+1) \text{ 列 } (1\le i\le n-1)}}\frac{\prod\limits_{j=2}^{n}(a_j-a_1)}{\prod\limits_{j=1}^{n}(a_1+b_j)}\det\begin{pmatrix}1 & 0 & \cdots & 0\\ \frac{1}{a_2+b_1} & \frac{b_2-b_1}{(a_2+b_1)(a_2+b_2)} & \cdots & \frac{b_n-b_1}{(a_2+b_1)(a_2+b_n)}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{1}{a_n+b_1} & \frac{b_2-b_1}{(a_n+b_1)(a_n+b_n)} & \cdots & \frac{b_n-b_1}{(a_n+b_1)(a_n+b_n)}\end{pmatrix}$$

$$\xlongequal{\text{按第一行展开}} \frac{\prod\limits_{j=2}^{n}(a_j-a_1)}{\prod\limits_{j=1}^{n}(a_1+b_j)} \det\begin{pmatrix} \frac{b_2-b_1}{(a_2+b_1)(a_2+b_2)} & \cdots & \frac{b_n-b_1}{(a_2+b_1)(a_2+b_n)} \\ \vdots & \vdots & \vdots \\ \frac{b_2-b_1}{(a_n+b_1)(a_n+b_2)} & \cdots & \frac{b_n-b_1}{(a_n+b_1)(a_n+b_n)} \end{pmatrix}$$

$$\xlongequal{\text{提取各行、列的公因式}} \frac{\prod\limits_{j=2}^{n}(a_j-a_1)(b_j-b_1)}{\prod\limits_{j=1}^{n}(a_1+b_j)(a_j+b_1)} \det\left(\frac{1}{a_i+b_j}\right)_{2\le i,j\le n}.$$

再由 $\det\left(\frac{1}{a_n+b_n}\right)=\frac{1}{a_n+b_n}$, 一阶递推即知结论 $\det\left(\frac{1}{a_i+b_j}\right)_{1\le i,j\le n}=\frac{\prod\limits_{1\le i<j\le n}(a_j-a_i)(b_j-b_i)}{\prod\limits_{1\le i,j\le n}(a_i+b_j)}$. □

推论 4.4.7 (Cauchy 方阵的逆) 记 $A_n=\left(\frac{1}{a_i+b_j}\right)_{1\le i,j\le n}$ $(n\ge 1)$, 证明: $A_n\in \mathrm{GL}(n,F)\iff \forall 1\le i,j\le n$, $a_i\ne a_j$, $b_i\ne b_j$; 且此时 $(A_n^{-1})_{ij}=(a_j+b_i)\prod\limits_{\substack{l=1\\l\ne j}}^{n}\frac{a_l+b_i}{a_j-a_l}\prod\limits_{\substack{k=1\\k\ne i}}^{n}\frac{a_j+b_k}{b_i-b_k}$.

例 4.4.11 (二阶递推关系) 记 $A_n=\begin{pmatrix} \alpha+\beta & \alpha\beta & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & \alpha\beta \\ & & 1 & \alpha+\beta \end{pmatrix}$, 求 $\det(A_n)$.

证明: 按最后一行展开知, $\det(A_n)=(\alpha+\beta)\det(A_{n-1})-\alpha\beta\det(A_{n-2})\,(n\ge 2)$. 再由 $\det(A_1)=\alpha+\beta$, $\det(A_2)=(\alpha+\beta)^2-\alpha\beta$, 二阶递推即知 $\det(A_n)=\sum\limits_{i=0}^{n}\alpha^i\beta^{n-i}$. □

例 4.4.12 (连分数的部分商) 设 F 为一个域, $\{a_n\}_{n\in\mathbb{N}}\subseteq F$. 记 $[a_0,a_1,\cdots,a_n]=a_0+\cfrac{1}{a_1+\cfrac{1}{\ddots+\cfrac{\ddots}{a_{n-1}+\cfrac{1}{a_n}}}}$,

$p_n=p_n(a_0,\cdots,a_n)=\det\begin{pmatrix} a_0 & -1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & 1 & a_n \end{pmatrix}$, $q_n=q_n(a_0,\cdots,a_n)=\det\begin{pmatrix} a_1 & -1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & 1 & a_n \end{pmatrix}$. 证明:

(1) $\begin{pmatrix} p_n \\ q_n \end{pmatrix}=\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$;

(2) $[a_0,a_1,\cdots,a_n]=\frac{p_n}{q_n}$.

证明: (1) 对 n 归纳证明: 当 $n=0$ 时, $p_0=a_0$, $q_0=1$; 当 $n=1$ 时, $p_1=a_0a_1+1$, $q_1=a_1$. 现设 $n\ge 2$ 且 $<n$ 时结论均成立. 记 $p'_n=p_n(a_1,\cdots,a_{n+1})$, $q'_n=q_n(a_1,\cdots,a_{n+1})$, 则由行列式按第一行展开知, $\begin{cases} p_n=a_0p'_{n-1}+q'_{n-1} \\ q_n=p'_{n-1} \end{cases}$, 故 $\begin{pmatrix} p_n \\ q_n \end{pmatrix}=\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} p'_{n-1} \\ q'_{n-1} \end{pmatrix}=\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

(2) 对 n 归纳证明: 当 $n=0$ 时, $[a_0]=a_0=\frac{a_0}{1}=\frac{p_0}{q_0}$; 当 $n=1$ 时, $[a_0,a_1]=a_0+\frac{1}{a_1}=\frac{a_0a_1+1}{a_1}=\frac{p_1}{q_1}$. 现设 $n\ge 2$ 且当 $<n$ 时结论均成立, 则 $[a_0,\cdots,a_n]=a_0+\frac{1}{[a_1,\cdots,a_n]}=a_0+\frac{q'_{n-1}}{p'_{n-1}}=\frac{a_0p'_{n-1}+q'_{n-1}}{p'_{n-1}}=\frac{p_n}{q_n}$. □

注: 一般地, 记三对角行列式 $\Delta_n=\det\begin{pmatrix} a_1 & b_1 & & \\ c_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n \end{pmatrix}$, 以及 $b_0,b_n,c_0,c_n\in F$ 为任意数, 则

(1) $\Delta_n=\begin{pmatrix}1&0\end{pmatrix}\begin{pmatrix}a_1&b_1\\-c_0&0\end{pmatrix}\begin{pmatrix}a_2&b_2\\-c_1&0\end{pmatrix}\cdots\begin{pmatrix}a_n&b_n\\-c_{n-1}&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$.

(2) $\Delta_n=\begin{pmatrix}1&0\end{pmatrix}\begin{pmatrix}a_n&b_{n-1}\\-c_n&0\end{pmatrix}\begin{pmatrix}a_{n-1}&b_{n-2}\\-c_{n-1}&0\end{pmatrix}\cdots\begin{pmatrix}a_1&b_0\\-c_1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}$.

第五种方法是利用矩阵乘法与分块零化技巧, 它常可将复杂的行列式化为若干已解决的情形.

例 4.4.13 设 $f_i(X)=\sum\limits_{k=0}^{n-1}c_{i,k+1}X^k\in F[X]\,(1\leq i\leq n)$, $a_1,\cdots,a_n\in F$, 求 $\det(f_i(a_j))_{1\leq i,j\leq n}$.

证明: 由 $f_i(a_j)=\sum\limits_{k=0}^{n-1}c_{i,k+1}a_j^k$ 知, $(f_i(a_j))_{1\leq i,j\leq n}=(c_{ij})_{1\leq i,j\leq n}\cdot(a_j^{i-1})_{1\leq i,j\leq n}$, 故由 Vandermonde 行列式知, $\det(f_i(a_j))_{1\leq i,j\leq n}=\det(c_{ij})_{1\leq i,j\leq n}\cdot\prod\limits_{1\leq i<j\leq n}(a_j-a_i)$. □

例 4.4.14 记 $s_i=\sum\limits_{k=1}^{n}a_k^i\,(0\leq i\leq 2n-2)$, 求 $\det(s_{i+j-1})_{1\leq i,j\leq n}$.

证明: 由 $s_{i+j-2}=\sum\limits_{k=1}^{n}a_k^{i-1}a_k^{j-1}$ 知, $(s_{i+j-2})_{1\leq i,j\leq n}=(a_j^{i-1})_{1\leq i,j\leq n}\cdot(a_k^{l-1})_{1\leq k,l\leq n}$, 故由 Vandermonde 行列式知, $\det(s_{i+j-2})_{1\leq i,j\leq n}=\det(a_j^{i-1})_{1\leq i,j\leq n}^2=\prod\limits_{1\leq i<j\leq n}(a_j-a_i)^2$. □

例 4.4.15 设 F 为一个域, $A\in F^{m\times n},B\in F^{n\times m}$, X 为不定元, 证明: $X^n\det(XI_m-AB)=X^m\det(XI_n-BA)$.

证明:

$$X^n\det(XI_m-AB)\xlongequal{\text{准对角块的行列式}}\det\begin{pmatrix}XI_n&0\\0&XI_m-AB\end{pmatrix}\xlongequal{\text{准上三角块的行列式}}\det\begin{pmatrix}XI_n&B\\0&XI_m-AB\end{pmatrix}$$
$$\xlongequal{\substack{\text{将第 1 行的左 } A \text{ 倍}\\\text{加往第 2 行}}}\det\begin{pmatrix}XI_n&B\\XA&XI_m\end{pmatrix}\xlongequal{\text{提取第 1 列的公因式}}X^n\det\begin{pmatrix}I_n&B\\A&XI_m\end{pmatrix};$$

$$X^m\det(XI_n-BA)\xlongequal{\text{准对角块的行列式}}\det\begin{pmatrix}XI_n-BA&0\\0&XI_m\end{pmatrix}\xlongequal{\text{准下三角块的行列式}}\det\begin{pmatrix}XI_n-BA&0\\A&XI_m\end{pmatrix}$$
$$\xlongequal{\substack{\text{将第 2 行的左 } B \text{ 倍}\\\text{加往第 1 行}}}\det\begin{pmatrix}XI_n&XB\\A&XI_m\end{pmatrix}\xlongequal{\text{提取第 1 行的公因式}}X^n\det\begin{pmatrix}I_n&B\\A&XI_m\end{pmatrix}.$$

因此 $X^n\det(XI_m-AB)=X^m\det(XI_n-BA)$. □

推论 4.4.8 设 F 为一个域, $A\in F^{m\times n},B\in F^{n\times m}$, 则

(1) $X^n\cdot f_{AB}(X)=X^m\cdot f_{BA}(X)$. 特别地, 当 $m=n$ 时, $f_{AB}(X)=f_{BA}(X)$.

(2) $\sigma(AB)\setminus\{0\}=\sigma(BA)\setminus\{0\}$. 特别地, 当 $m=n$ 时, $\sigma(AB)=\sigma(BA)$.

注:

(1) 上述推论 (1) 也可由相似性直接说明: 注意到

$$\begin{pmatrix}I_m&-A\\0&I_n\end{pmatrix}\begin{pmatrix}AB&0\\B&0_n\end{pmatrix}\begin{pmatrix}I_m&A\\0&I_n\end{pmatrix}=\begin{pmatrix}0_m&0\\B&BA\end{pmatrix},$$

则 $\begin{pmatrix}AB&0\\B&0_n\end{pmatrix}$ 与 $\begin{pmatrix}0_m&0\\B&BA\end{pmatrix}$ 的特征多项式相同, 即 $X^n\cdot f_{AB}(X)=X^m\cdot f_{BA}(X)$.

(2) 上述推论 (2) 也可不依赖行列式说明: 注意到 $(I_n-BA)(I_n+B(I_m-AB)^{-1}A)=I_n$, 且反之亦然, 故 $I_n-BA\in\mathrm{GL}(n,F)\iff I_m-AB\in\mathrm{GL}(m,F)$, 因此 $\forall c\in F^*$, $cI_n-BA\in\mathrm{GL}(n,F)\iff cI_m-AB\in\mathrm{GL}(m,F)$, 即 $\sigma(AB)\setminus\{0\}=\sigma(BA)\setminus\{0\}$. 特别地, 当 $m=n$ 时, $AB\in\mathrm{GL}(n,F)\iff A,B\in\mathrm{GL}(n,F)\iff BA\in\mathrm{GL}(n,F)$, 故 $\sigma(AB)=\sigma(BA)$.

推论 4.4.9 设 F 为一个域, $\mathrm{char}(F)\nmid n$, $A,B\in F^{n\times n}$, 则 $AB-BA\neq I_n$.

证明: 假设 $AB-BA=I_n$, 则 $f_{AB}(X)=f_{BA}(X-1)$. 又 $f_{AB}(X)=f_{BA}(X)$, 故 $f_{BA}(X-1)=f_{BA}(X)$. 由 $\mathrm{char}(F)\nmid n$ 知, $f_{BA}(X)-f_{BA}(0)$ 在域 F 中有至少 $(n+1)$ 个根, 则 $f_{BA}(X)$ 必为常数, 矛盾! □

推论 4.4.10 (Cauchy-Binet) 设 F 为一个域, $A\in F^{m\times n}$, $B\in F^{n\times s}$, 记 $C=A\cdot B\in F^{m\times s}$, 则

$$\det(C_{\{i_1,\cdots,i_r\},\{j_1,\cdots,j_r\}})=\begin{cases}0, & r>n\\ \sum\limits_{1\le k_1<\cdots<k_r\le n}\det(A_{\{i_1,\cdots,i_r\},\{k_1,\cdots,k_r\}})\cdot\det(B_{\{k_1,\cdots,k_r\},\{j_1,\cdots,j_r\}}), & r\le n\end{cases}.$$

证明: 由矩阵乘法 $C_{\{i_1,\cdots,i_r\},\{j_1,\cdots,j_r\}}=A_{\{i_1,\cdots,i_r\},\{1,\cdots,n\}}\cdot B_{\{1,\cdots,n\},\{j_1,\cdots,j_r\}}$ 知, 可不妨设 $m=r=s$. 以下考虑等式 $X^n\det(XI_m-AB)=X^m\det(XI_n-BA)$: 左端关于 X^n 的系数为 $(-1)^m\det(AB)$; 右端由行列式按 $(n-m)$ 行展开, 关于 X^n 的系数为 $\begin{cases}0, & m>n\\ (-1)^m\sum\limits_{1\le i_1<\cdots<i_m\le n}\det((BA)_{\{k_1,\cdots,k_m\},\{k_1,\cdots,k_m\}}), & m\le n\end{cases}$. 因此比较系数知,

$$\begin{aligned}\det(AB)&=\begin{cases}0, & m>n\\ \sum\limits_{1\le i_1<\cdots<i_m\le n}\det((BA)_{\{k_1,\cdots,k_m\},\{k_1,\cdots,k_m\}}), & m\le n\end{cases}\\ &=\begin{cases}0, & m>n\\ \sum\limits_{1\le i_1<\cdots<i_m\le n}\det(B_{\{k_1,\cdots,k_m\},\{1,\cdots,m\}})\det(A_{\{1,\cdots,m\},\{k_1,\cdots,k_m\}}), & m\le n\end{cases}.\end{aligned}$$

□

例 4.4.16 (秩一扰动的行列式) 设 F 为一个域, $A\in F^{n\times n}$, $\alpha\in F^{n\times 1}$, $\beta\in F^{1\times n}$, 求 $\det(A-\alpha\beta)\in F$.

证明: (1) 当 $A\in\mathrm{GL}(n,F)$ 时, $\det(A-\alpha\beta)=\det(A)\det(I_n-(A^{-1}\alpha)\beta)=\det(A)(1-\beta A^{-1}\alpha)$.

(2) 当 $A\notin\mathrm{GL}(n,F)$ 时, $r(A)\le n-1$. 又 $r(\alpha\beta)\le 1$, 则 $r(A-\alpha\beta)\le r(A)+r(\alpha\beta)\le n$, 且

第一个 "=" 成立 $\iff\begin{cases}\mathrm{row}(A)\cap\mathrm{row}(\alpha\beta)=\{0\}\\ \mathrm{column}(A)\cap\mathrm{column}(\alpha\beta)=\{0\}\end{cases}\iff\alpha=0$ 或 $\beta=0$ 或 $\begin{cases}\beta\notin\mathrm{row}(A)\\ \alpha\notin\mathrm{column}(A)\end{cases}$.

因此 $\det(A-\alpha\beta)\ne 0\iff r(A-\alpha\beta)=n\iff\begin{cases}\mathrm{Span}_F(\{\beta\})\oplus\mathrm{row}(A)=F^{1\times n}\\ \mathrm{Span}_F(\{\alpha\})\oplus\mathrm{column}(A)=F^{n\times 1}\end{cases}$. □

推论 4.4.11 $\det(1+a_ib_j)_{1\le i,j\le n}=\begin{cases}1+a_1b_1, & n=1\\ (a_1-a_2)(b_1-b_2), & n=2\\ 0, & n\ge 3\end{cases}$; $\det(a_i+b_j)_{1\le i,j\le n}=\begin{cases}a_1+b_1, & n=1\\ -(a_1-a_2)(b_1-b_2), & n=2\\ 0, & n\ge 3\end{cases}$.

参考文献与补注 4.4

(1) 关于方阵的特征值理论, 可以参考教材 Hoffman, Kunze"Linear Algebra" 的后续章节.

(2) 关于行列式的更多习题, 可以参考李炯生, 查建国, 王新茂 "线性代数"(第 2 版).

第 5 章　多项式

本章介绍代数学最基本的一个概念: 多项式 (polynomial). 与初等数学中的形式化讨论不同, 这里的语言和方法更具有环论或代数理论的背景.

§5.1　一元多项式代数

设 F 是一个域, 回忆 $\coprod_{n=0}^{\infty} F$ 在规定的加法、数乘与乘法下构成一个 F-结合代数, 称为域 F 上的一元多项式代数. 它作为线性空间的性质已在前述章节中阐明; 本节将更多着眼于它作为环的性质.

定义 5.1.1 (整环) 设 $(R,0,+;1,\cdot)$ 是一个环, 记 $R^*=R\backslash\{0\}$, 若 $(R^*,\cdot,1)$ 是一个交换的幺半群, 则称 $(R,0,+;1,\cdot)$ 是一个**整环** (integral domain).

注: 特别地, 整环非零环. 在整环中, 以下的乘法消去律成立: 若 $a\in R^*$, $b,c\in R$ 满足 $a\cdot b=a\cdot c$, 则 $b=c$.

例 5.1.1

(1) $(\mathbb{Z},0,+;1,\cdot)$ 是一个整环; 设 $n\in\mathbb{N}^*$, 则 $(\mathbb{Z}/n\mathbb{Z},\overline{0},+;\overline{1},\cdot)$ 是一个整环 $\Leftrightarrow(\mathbb{Z}/n\mathbb{Z},\overline{0},+;\overline{1},\cdot)$ 是一个域 $\Leftrightarrow n$ 为素数;

(2) 任意域 $(F,0,+;1,\cdot)$ 都是整环; 反之, 任意满足 Artin 性质的整环都是域;

(3) 设 $(R,0,+;1,\cdot)$ 是一个整环, 则环上的一元多项式环 $(R[X],0,+;1,\cdot)$ 是一个整环. (事实上, 若干多项式乘积的最高次项系数等于它们的最高次项系数的乘积, 故由整环无非零的零因子性质即知.)

定义 5.1.2 (Euclid 整环) 设 $(R,0,+;1,\cdot)$ 是一个整环. 若函数 $f\colon R^*\to\mathbb{N}$ 满足以下的带余除法性质: $\forall a\in R$, $\forall b\in R^*,\exists q,r\in R$, $s.t.\ a=b\cdot q+r$, 且 $r=0$ 或 $f(r)<f(b)$, 则称它为环 $(R,0,+;1,\cdot)$ 上的一个 **Euclid 函数**. 具有至少一个 Euclid 函数的整环称为 **Euclid 整环** (Euclidean domain).

注:

(1) 为方便起见, 有时我们将 Euclid 函数在 $0\in R$ 处的取值定义为 $-\infty$;

(2) 在给定 Euclid 函数后, 上述带余除法中的商与余数一般不唯一确定;

(3) 许多文献还要求 Euclid 函数 f 额外满足以下的性质: $\forall a,b\in R^*$, $f(a)\leq f(a\cdot b)$. 事实上, 这可通过调整 Euclid 函数的选取得到: 设 f 为一个 Euclid 函数, 令 $g\colon R^*\longrightarrow\mathbb{N}$, $a\longmapsto\min\limits_{b\in R^*}f(a\cdot b)$ 即为满足要求的 Euclid 函数.

例 5.1.2

(1) $(\mathbb{Z},0,+;1,\cdot)$ 是一个 Euclid 整环, 可选取 Euclid 函数为 $|\cdot|$;

(2) 任意域 $(F,0,+;1,\cdot)$ 都是 Euclid 整环, 可选取 Euclid 函数为常值 1;

(3) 设 $(R,0,+;1,\cdot)$ 为一个整环, 则 $(R[X],0,+;1,\cdot)$ 是一个 Euclid 整环 $\Longleftrightarrow(R,0,+;1,\cdot)$ 是一个域, 可选取 Euclid 函数为 deg, 即最高次项的次数. 这个性质的验证需要引入环中理想的工具.

定义 5.1.3 (主理想整环) 设 $(R,0,+;1,\cdot)$ 是一个整环, 若环 R 的理想都由一个元生成, 则称 $(R,0,+;1,\cdot)$ 是一个**主理想整环** (principal ideal domain).

命题 5.1.1 *任意Euclid整环都是主理想整环.*

证明: 设 $(R,0,+;1,\cdot)$ 是一个 Euclid 整环, 取 Euclid 函数为 f. 任取 $\mathfrak{a}\subseteq R$ 为理想, 若 $\mathfrak{a}=\{0\}$, 则 $\mathfrak{a}$ 可由一个元 0 生成; 若 $\mathfrak{a}\neq\{0\}$, 则可取 $b\in\mathfrak{a}\setminus\{0\}$, 满足 $f(b)=\min\limits_{a\in\mathfrak{a}\setminus\{0\}}f(a)$, 断言: $\mathfrak{a}=b\cdot R$. 这是因为, 显然 $\mathfrak{a}\supseteq b\cdot R$; 另一方面, 任取 $a\in\mathfrak{a}$, 由带余除法知, $\exists q,r\in R$, $s.t.\ a=b\cdot q+r$, 且 $r=0$ 或 $f(r)<f(b)$. 但由 $r=a-b\cdot q\in\mathfrak{a}$ 以及 b 的选取知, 只能有 $r=0$, 即 $a=b\cdot q\in b\cdot R$. □

注: 在域上的一元多项式环 $F[X]$ 中, 每个非零理想都有唯一的首一多项式作为生成元, 它可视为该理想中除零多项式外的次数最低者.

推论 5.1.2 设 $(R,0,+;1,\cdot)$ 是一个环, 则以下条件等价:

(1) $(R,0,+;1,\cdot)$ 是一个域;

(2) $(R[X],0,+;1,\cdot)$ 是一个 Euclid 整环;

(3) $(R[X],0,+;1,\cdot)$ 是一个主理想整环.

证明: (1)⇒(2)⇒(3) 显然; 现证 (3)⇒(1): 设 $(R[X],0,+;1,\cdot)$ 是一个主理想整环, 则 $(R,0,+;1,\cdot)$ 是一个整环. 任取 $a\in R^*$. 由于 $\mathfrak{a}=a\cdot R[X]+X\cdot R[X]$ 是环 $R[X]$ 中的理想, 则它可由一个元 $b(X)\in R[X]$ 生成, 故 $\exists c(X)\in R[X]$,

$s.t.\ a=b(X)\cdot c(X)$. 两端取 deg 知, $b(X)=b\in R^*$. 又由于 $\exists d(X)\in R[X],\ s.t.\ X=b\cdot d(X)$, 两端比较最高次项系数知, $b\in R^\times=\{R$ 中的乘法可逆元 $\}$, 故 b 在 $R[X]$ 中生成的理想为 $R[X]$, 即 $\mathfrak{a}=R[X]$. 特别地, $\exists f(X),g(X)\in R[X],\ s.t.\ 1=a\cdot f(X)+X\cdot g(X)$. 两端比较常数项知, $a\in R^\times$. 因此 $(R,0,+;1,\cdot)$ 是一个域. □

例 5.1.3 (线性变换的多项式代数) 设 V 是域 F 上的线性空间, $T\in L(V)$, 则存在代数同态 $\begin{array}{ccc}F[X]&\longrightarrow&L(V)\\ f(X)&\longmapsto&f(T)\end{array}$,

这里 $f(T)\coloneqq a_mT^m+\cdots+a_1T+a_0\mathrm{id}_V$, 若 $f(X)=a_mX^m+\cdots+a_1X+a_0$. 记 $M_T=\{f(X)\in F[X]\colon f(T)=0\}$, 则 M_T 为 $F[X]$ 的理想. 由 $F[X]$ 为主理想整环知, M_T 可由一个元生成. 若 $M_T=\{0\}$, 则它的唯一生成元为 0; 若 $M_T\neq\{0\}$, 则可取它的唯一的首一生成元 $p_T(X)$, 称为 $T\in L(V)$ 的**最小多项式** (minimal polynomial).

注意上述代数同态可分解为 $\begin{array}{ccccc}F[X]&\longrightarrow&F[X]/M_T&\xrightarrow{\cong}&F[T]\subseteq L(V)\\ f(X)&\longmapsto&f(X)+M_T&\longmapsto&f(T)\end{array}$. 于是当固定 $T\in L(V)$ 时, F-线性空间 V 可视为一元多项式环 $F[X]$ 的左模, 故此时可由主理想整环上模的分解性质研究线性变换的性质.

定义 5.1.4 (素理想与极大理想) 设 $(R,0,+;1,\cdot)$ 是一个交换环, $\mathfrak{a}\subsetneq R$ 为一个理想.

(1) 若 $\forall a,b\in R\ (a\cdot b\in\mathfrak{a}\Rightarrow a\in\mathfrak{a}$ 或 $b\in\mathfrak{a})$, 则称 $\mathfrak{a}$ 为 R 中的一个**素理想** (prime ideal).

(2) 若 $\mathfrak{a}$ 是 R 中真理想在包含关系下的极大者, 则称 $\mathfrak{a}$ 为 R 中的一个**极大理想** (maximal ideal).

注:

(1) 设 $\mathfrak{a}\subsetneq R$ 为一个理想, 则 $\mathfrak{a}$ 为一个素理想 $\iff R/\mathfrak{a}$ 是一个整环; $\mathfrak{a}$ 为一个极大理想 $\iff R/\mathfrak{a}$ 是一个域; 特别地, 任意极大理想必为素理想;

(2) 与 Zorn 引理等价的是 Krull 定理: 任意非零交换环中的真理想必包含于某个极大理想.

定义 5.1.5 (素元与不可约元) 设 $(R,0,+;1,\cdot)$ 是一个交换环, $a\in R^*\setminus R^\times$.

(1) 若 $\forall b,c\in R\ (a\mid b\cdot c\Rightarrow a\mid b$ 或 $a\mid c)$, 则称 a 为 R 中的一个**素元** (prime element).

(2) 若 $\forall b,c\in R\ (a=b\cdot c\Rightarrow a\mid b$ 或 $a\mid c)$, 则称 a 为 R 中的一个**不可约元** (irreducible element).

注:

(1) 设 $a\in R^*\setminus R^\times$, 则由定义知, $a\cdot R$ 为一个极大理想 $\Longrightarrow a\cdot R$ 为一个素理想 $\iff a$ 为一个素元 $\Longrightarrow a$ 为一个不可约元. 以下我们将仔细探究上述逆箭头成立的条件.

(2) 设 $a\in R^*\setminus R^\times$, 则以下条件 (a)⇒(b)⇒(c)⇒(d); 且当 $(R,0,+;1,\cdot)$ 是一个整环时, (d)⇒(a):

(a) $\forall b,c\in R\ (a=b\cdot c\Rightarrow b\in R^\times$ 或 $c\in R^\times)$;

(b) $\forall b,c\in R\ (a=b\cdot c\Rightarrow a\in b\cdot R^\times$ 或 $a\in c\cdot R^\times)$;

(c) $a\cdot R$ 是 R 中真主理想在包含关系下的极大者;

(d) a 是一个不可约元.

定义 5.1.6 (唯一分解整环) 设 $(R,0,+;1,\cdot)$ 是一个整环, 若 R 中每个非零元都可写成一个乘法可逆元与若干不可约元的乘积, 且表达式在不计顺序与乘法可逆元的意义下唯一, 则称 $(R,0,+;1,\cdot)$ 是一个**唯一分解整环** (unique factorization domain).

注: 上述唯一分解整环的定义可分成两部分, 第一部分是不可约分解的存在性, 它略弱于环中主理想的升链条件 (ACCP); 第二部分是不可约分解的唯一性, 它略弱于环中不可约元均为素元.

引理 5.1.3 设 $(R,0,+;1,\cdot)$ 是一个交换环, 若 R 中不存在无穷长的严格主理想升链, 则 R 中每个非零元都可写成一个乘法可逆元与若干不可约元的乘积.

证明: 记 $S=\{a\in R^*:a$ 不能写成一个乘法可逆元与若干不可逆元的乘积 $\}$, 假设 $S\neq\varnothing$. 取 $a\in S$, 则 $a\in R^*\setminus R^\times$, 且不为不可约元, 故 $\exists a_1,a_2\in R,\ s.t.\ a=a_1\cdot a_2$ 且 $a\nmid a_1,\ a\nmid a_2$. 注意 $a_1\notin S$ 或 $a_2\notin S$, 不妨设 $b_1=a_1\notin S$, 则 $b_1\in R^*\setminus R^\times$, 且不为不可约元, 故 $\exists a_{11},a_{12}\in R,\ s.t.\ b_1=a_{11}\cdot a_{12}$ 且 $b_1\nmid a_{11},\ b_1\nmid a_{12}$. 注意 $a_{11}\notin S$ 或 $a_{12}\notin S$, 不妨设 $b_2=a_{11}\notin S$, 则 $b_2\in R^*\setminus R^\times$, 且不为不可约元. 如此继续知, R 中存在无穷长的严格主理想升链 $\{0\}\subsetneqq b_1\cdot R\subsetneqq b_2\cdot R\subsetneqq\cdots\subsetneqq R$. □

注: 上述引理的逆命题未必成立, 反例可见 A. Grams "Atomic rings and the ascending chain condition for principal ideals "(1973).

引理 5.1.4 设 $(R,0,+;1,\cdot)$ 是一个整环, 若 R 中不可约元均为素元, 则 R 中非零元可写成一个乘法可逆元与若干不可约元的乘积的表达式在不计顺序与乘法可逆元的意义下唯一.

证明: 设 $a\in R^*$ 可写成 $a=u\cdot p_1\cdots p_n=u'\cdot p_1'\cdots p_m'$, 其中 $u,u'\in R^\times$, $p_i(1\leq i\leq n)$, $p_j'(1\leq j\leq m)$ 为不可约元, 则也为素元. 由 $p_1'\mid p_1\cdots p_n$ 知, $\exists 1\leq i_1\leq n,\ s.t.\ p_1'\mid p_{i_1}$. 由 p_{i_1} 的不可约性与无非零的零因子性知, $p_{i_1}=u_1\cdot p_1'$, 其中 $u_1\in R^\times$. 通过重排, 不妨设 $i_1=1$, 则表达式化为 $u\cdot u_1\cdot p_2\cdots p_n=u'\cdot p_2'\cdots p_m'$. 以下对不可约元的数量归纳可知, $n=m$, 且 $\exists\{i_1,\cdots,i_n\}=\{1,\cdots,n\},\ s.t.\ p_{i_j}=u_j\cdot p_j'$, 其中 $u_j\in R^\times\ (1\leq j\leq n)$. □

命题 5.1.5 设 $(R,0,+;1,\cdot)$ 是一个整环, 则以下条件等价:

(1) R 中不存在无穷长的严格主理想升链, 且 R 中不可约元均为素元;

(2) $(R,0,+;1,\cdot)$ 是一个唯一分解整环.

证明: "(1)⇒(2)": 由上述引理可知;

"(2)⇒(1)": 设 $(R,0,+;1,\cdot)$ 是一个唯一分解整环. 为证明 R 中不存在无穷长的严格主理想升链, 只需证明: 任取 $a\in R^*\setminus R^\times$, 则 R 中严格包含 $a\cdot R$ 的主理想只有有限多个. 事实上, 设 $b\cdot R\supsetneqq a\cdot R$, 即 $b\mid a$ 且 $a\nmid b$. 由唯一分解的性质知, b 的不可约因子集严格包含于 a 的不可约因子集, 故 $b\cdot R$ 只有有限种可能.

设 a 为一个不可约元, 任取 $b,c\in R$ 满足 $a\mid b\cdot c$, 即 $\exists d\in R,\ s.t.\ b\cdot c=a\cdot d$. 将等式两端分别展开为乘法可逆元与不可约元的乘积, 由唯一分解的性质知, a 必出现在 b 的表达式或 c 的表达式中, 即 $a\mid b$ 或 $a\mid c$, 故 a 为一个素元. □

命题 5.1.6 设 $(R,0,+;1,\cdot)$ 是一个整环, 则以下条件等价:

(1) $(R,0,+;1,\cdot)$ 是一个主理想整环;

(2) $(R,0,+;1,\cdot)$ 是一个唯一分解整环, 且 R 中非零素理想均为极大理想.

证明: "(1)⇒(2)": 设 $(R,0,+;1,\cdot)$ 是一个主理想整环. 假设 R 中存在无穷长的严格主理想升链 $\{0\}\subsetneqq a_1\cdot R\subsetneqq a_2\cdot R\subsetneqq\cdots\subsetneqq R$, 则可直接验证 $\mathfrak{a}=\bigcup\limits_{i=1}^{+\infty}a_i\cdot R$ 也为 R 的理想, 故 $\exists a\in R,\ s.t.\ \mathfrak{a}=a\cdot R$. 注意 $\exists i\geq 1,\ s.t.\ a\in a_i\cdot R$, 则 $a_{i+1}\in\mathfrak{a}=b\cdot R\subseteq a_i\cdot R$, 矛盾!

设 $a\in R^*\setminus R^\times$ 为一个不可约元, 则 $a\cdot R$ 是 R 中真主理想在包含关系下的极大者. 又 R 中任意理想均为主理想, 故 $\mathfrak{a}=a\cdot R$ 为极大理想, 也为素理想, 即 a 为一个素元. 因此由上述命题知, $(R,0,+;1,\cdot)$ 是一个唯一分解整环.

现设 $\{0\}\subsetneqq\mathfrak{a}\subsetneqq R$ 为一个素理想, 则 $\exists a\in R,\ s.t.\ \mathfrak{a}=a\cdot R$, 且 a 为一个素元, 也为不可约元, 故 $a\cdot R$ 是 R 中真主理想在包含关系下的极大者. 又 R 中任意理想均为主理想, 故 $\mathfrak{a}=a\cdot R$ 为极大理想.

"(2)⇒(1)": 设 $(R,0,+;1,\cdot)$ 是一个唯一分解整环, 且 R 中非零素理想均为极大理想. 先断言: R 中非零素理想均为主理想. 这是因为, 设 $\{0\}\subsetneqq\mathfrak{a}\subsetneqq R$ 为一个素理想, 任取 $a\in\mathfrak{a}\setminus\{0\}$, 记 $a=u\cdot p_1\cdots p_n$, 其中 $u\in R^\times$, $p_1,\cdots,p_n\in R^*\setminus R^\times$ 为不可约元. 由素性知, $\exists 1\leq i\leq n,\ s.t.\ p_i\in\mathfrak{a}$, 即 $p_i\cdot R\subseteq\mathfrak{a}$. 而由 $p_i\in R$ 也为素元知, $p_i\cdot R$ 为非零素理想, 故由条件知 $p_i\cdot R$ 也为极大理想, 因此 $p_i\cdot R=\mathfrak{a}$.

以下考虑理想族 $\mathscr{F}=\{\mathfrak{a}\subseteq R:\mathfrak{a}$是非主理想$\}$ 及其上的包含偏序, 假设 $\mathscr{F}\neq\varnothing$. 现设 $\{\mathfrak{a}_i\}_{i\in I}$ 为 $\mathscr{F}$ 中的一条链, 取 $\mathfrak{a}=\bigcup\limits_{i\in I}\mathfrak{a}_i\subseteq R$, 则可直接验证 $\mathfrak{a}$ 也为 R 的非主理想, 即 $\mathfrak{a}$ 为 $\{\mathfrak{a}_i\}_{i\in I}$ 在 $\mathscr{F}$ 中的上界. 由 Zorn 引理知, $\mathscr{F}$ 有极大元, 记为 $\mathfrak{a}$. 再断言: $\mathfrak{a}$ 是一个素理想. 这是因为, 由 $\mathfrak{a}$ 为非主理想知, $\{0\}\subsetneqq\mathfrak{a}\subsetneqq R$. 假设 $\exists a,b\in R\setminus\mathfrak{a},\ s.t.\ a\cdot b\in\mathfrak{a}$, 则 $\mathfrak{a}\subsetneqq\mathfrak{a}+a\cdot R\subseteq R$ 为理想. 由 $\mathfrak{a}$ 在 $\mathscr{F}$ 中的极大性知, $\mathfrak{a}+a\cdot R$ 是主理想, 即

$\exists c \in R,\ s.t.\ \mathfrak{a} + a \cdot R = c \cdot R$. 又 $\mathfrak{a} \subsetneqq \mathfrak{a} + b \cdot R \subseteq (\mathfrak{a} : a) \subseteq R$ 为理想, 由 $\mathfrak{a}$ 在 $\mathscr{F}$ 中的极大性知, $(\mathfrak{a} : a)$ 是主理想, 即 $\exists d \in R,\ s.t.\ (\mathfrak{a} : a) = d \cdot R$.

最后断言: $\mathfrak{a} = c \cdot d \cdot R$, 于是这与 $\mathfrak{a}$ 是非主理想矛盾. 这是因为, 一方面, $\forall e \in \mathfrak{a},\ \exists r \in R,\ s.t.\ e = c \cdot r$, 则 $a \cdot r = r \cdot a \in r \cdot a \cdot R \subseteq r \cdot c \cdot R = c \cdot r \cdot R = e \cdot R \subseteq \mathfrak{a}$, 即 $r \in (\mathfrak{a} : a) = d \cdot R$, 故 $e \in c \cdot d \cdot R$, 从而 $\mathfrak{a} \subseteq c \cdot d \cdot R$; 另一方面, 由 $d \in (\mathfrak{a} : a)$ 知, $d \cdot a = a \cdot d \in \mathfrak{a}$, 则 $c \cdot d \cdot R = d \cdot c \cdot R = d \cdot (\mathfrak{a} + a \cdot R) \subseteq \mathfrak{a} + d \cdot a \cdot R \subseteq \mathfrak{a}$. □

以下我们研究一元多项式环作为唯一分解整环的性质, 其中的关键在于不可约多项式的判别法.

引理 5.1.7 设 $(R, 0, +; 1, \cdot)$ 是一个环, 则 $R^\times = R[X]^\times$; $\{R\text{中不可约元}\} = R \cap \{R[X]\text{中不可约元}\}$. 因此记 $\mathrm{UF}(R) \coloneqq \{R\text{中可写成一个乘法可逆元与若干不可约元的乘积, 且表达式在不计顺序与乘法可逆元的意义下唯一的元}\}$, 则 $\mathrm{UF}(R) = R \cap \mathrm{UF}(R[X])$.

证明: 由比较多项式的次数知显然. □

定义 5.1.7 (多项式的容度与本原性) 设 $(R, 0, +; 1, \cdot)$ 是一个交换环, $f(X) \in R[X]$, 记 $\mathrm{cont}(f(X))$ 为 R 中由 $f(X)$ 的所有系数生成的理想, 则 $\mathrm{cont}(f(X))$ 称为 $f(X)$ 的**容度** (content). 若 $\mathrm{cont}(f(X)) = R$, 则 $f(X)$ 称为**本原的** (primitive).

引理 5.1.8 设 $(R, 0, +; 1, \cdot)$ 是一个交换环, $f(X), g(X) \in R[X]$, 则:

(1) $\mathrm{cont}(f(X) \cdot g(X)) \subseteq \mathrm{cont}(f(X)) \cdot \mathrm{cont}(g(X)) \subseteq \sqrt{\mathrm{cont}(f(X) \cdot g(X))}$; 特别地, $f(X), g(X)$ 均为本原的 $\iff f(X) \cdot g(X)$ 为本原的;

(2) 设 $(R, 0, +; 1, \cdot)$ 是一个唯一分解整环, 对于 R 中有限生成的理想 $\mathfrak{a}$, 记 $\gcd(\mathfrak{a})$ 为 R 中最小的包含 $\mathfrak{a}$ 的主理想, 则 $\gcd(\mathrm{cont}(f(X) \cdot g(X))) = \gcd(\mathrm{cont}(f(X))) \cdot \gcd(\mathrm{cont}(g(X)))$.

证明: 可见 D. Eisenbud"Commutative algebra"(1995), 或维基百科. □

定理 5.1.9 (Gauss 引理) 设 $(R, 0, +; 1, \cdot)$ 是一个唯一分解整环, $F = \mathrm{Frac}(R)$ 为它的分式域, $f(X) \in R[X] \setminus R$, 则以下条件等价:

(1) $f(X)$ 为 $R[X]$ 中的不可约多项式;

(2) $f(X)$ 为 $F[X]$ 中的不可约多项式, 且 $\gcd(\mathrm{cont}(f(X))) = R$.

证明: "(1)⇒(2)": 设 $f(X)$ 为 $R[X]$ 中的不可约多项式. 由 $f(X) \notin R$ 知, $f(X) \neq 0$. 假设 $\gcd(\mathrm{cont}(f(X))) \subsetneqq R$, 即 $f(X)$ 的所有系数的最大公因子 $d \in R^* \setminus R^\times$, 则 $f(X) = d \cdot \dfrac{f(X)}{d}$, 但 $f(X) \nmid d$ 且 $f(X) \nmid \dfrac{f(X)}{d}$, 这与 $f(X)$ 在 $R[X]$ 中的不可约性矛盾! 因此 $\gcd(\mathrm{cont}(f(X))) = R$.

假设 $f(X)$ 不为 $F[X]$ 中的不可约多项式, 即 $\exists g(X), h(X) \in F[X],\ s.t.\ f(X) = g(X) \cdot h(X)$, 且在 $F[X]$ 中 $f(X) \nmid g(X), f(X) \nmid h(X)$. 通过通分, 可设 $g(X) = \dfrac{a}{b} g_1(X),\ h(X) = \dfrac{c}{d} h_1(X)$, 其中 $a, b, c, d \in R^*,\ g_1(X), h_1(X) \in R[X]$ 为本原的, 则 $b \cdot d \cdot f(X) = a \cdot c \cdot g_1(X) \cdot h_1(X)$. 由引理, 两边取 $\gcd(\mathrm{cont}(\cdot))$ 知, $b \cdot d \cdot R = a \cdot c \cdot R$, 即 $\dfrac{a \cdot c}{b \cdot d} \in R^\times$, 故 $f(X) = \dfrac{a \cdot c}{b \cdot d} g_1(X) \cdot h_1(X)$. 由 $f(X)$ 为 $R[X]$ 中的不可约多项式知, 在 $R[X]$ 中 $f(X) \mid g_1(X)$ 或 $f(X) \mid h_1(X)$, 因此 $\deg(f(X)) \leq \min\{\deg(g_1(X)), \deg(h_1(X))\} = \min\{\deg(g(X)), \deg(h(X))\} \leq \deg(f(X))$. 由该不等式取等知, $\deg(f(X)) = \deg(g(X)) = \deg(h(X)) = 0$, 即 $f \in R$, 矛盾!

"(2)⇒(1)": 设 $f(X)$ 为 $F[X]$ 中的不可约多项式, 且 $\gcd(\mathrm{cont}(f(X))) = R$. 假设 $f(X)$ 不为 $R[X]$ 中的不可约多项式, 即 $\exists g(X), h(X) \in R[X],\ s.t.\ f(X) = g(X) \cdot h(X)$, 且在 $R[X]$ 中 $f(X) \nmid g(X),\ f(X) \nmid h(X)$. 又由 $f(X)$ 为 $F[X]$ 中的不可约多项式知, 在 $F[X]$ 中 $f(X) \mid g(X)$ 或 $f(X) \mid h(X)$. 不妨设在 $F[X]$ 中 $f(X) \mid g(X)$, 则 $g(X) = \dfrac{a}{b} f(X)$, 其中 $a \in R^*,\ b \in R^* \setminus R^\times$, 则由引理, 两边取 $\gcd(\mathrm{cont}(\cdot))$ 知, $b \cdot \gcd(\mathrm{cont}(g(X))) = a \cdot \gcd(\mathrm{cont}(f(X))) = a \cdot R$, 故 $a \in b \cdot R$, 这与在 $R[X]$ 中 $f(X) \nmid g(X)$ 矛盾! □

推论 5.1.10 设 $(R, 0, +; 1, \cdot)$ 是一个环, 则以下条件等价:

(1) $(R, 0, +; 1, \cdot)$ 是一个唯一分解整环;

(2) $(R[X], 0, +; 1, \cdot)$ 是一个唯一分解整环.

证明: "(1)⇒(2)": 设 $(R, 0, +; 1, \cdot)$ 是一个唯一分解整环, 记 $F = \mathrm{Frac}(R)$. 任取 $f(X) \in R[X]^*$, 由 $R[X]^* \subseteq F[X]^*$ 以及 $F[X]$ 为唯一分解整环知, $f(X) = c \cdot p_1(X) \cdots p_n(X)$, 其中 $c \in F^*,\ p_1(X), \cdots, p_n(X) \in F[X]$ 为不可约多项式. 通过通分, 可设 $f(X) = \widetilde{c} \cdot \widetilde{p_1}(X) \cdots \widetilde{p_n}(X)$, 其中 $\widetilde{c} \in F^*,\ \widetilde{p_1}(X), \cdots, \widetilde{p_n}(X) \in R[X]$,

满足 $\widetilde{p_1}(X),\cdots,\widetilde{p_n}(X)$ 为 $F[X]$ 中的不可约多项式, 且 $\gcd(\mathrm{cont}(\widetilde{p_1}(X)))=\cdots=\gcd(\mathrm{cont}(\widetilde{p_n}(X)))=R$.
由 Gauss 引理知, $\widetilde{p_1}(X),\cdots,\widetilde{p_n}(X)$ 为 $R[X]$ 中的不可约多项式. 由 $\gcd(\mathrm{cont}(f(X)))=\widetilde{c}\cdot\gcd(\mathrm{cont}(\widetilde{p_1}(X)))\cdots\gcd(\mathrm{cont}(\widetilde{p_n}(X$
$\widetilde{c}\cdot R$ 知, $\widetilde{c}\in R^*$.
最后证明唯一性: 假设 $f(X)=d\cdot q_1(X)\cdots q_m(X)$, 其中 $d\in R^\times$, $q_1(X),\cdots,q_m(X)\in R[X]$ 为不可约多项式, 则
由 Gauss 引理知, $q_1(X),\cdots,q_m(X)$ 也为 $F[X]$ 中的不可约多项式, 且 $\gcd(\mathrm{cont}(q_1(X)))=\cdots=\gcd(\mathrm{cont}(q_m(X)))$
$=R$. 由 $F[X]$ 中的唯一分解性知, $n=m$, 且 $\exists\{i_1,\cdots,i_n\}=\{1,\cdots,n\}$, $s.t.$ $\widetilde{p_{i_j}}(X)=\dfrac{a_j}{b_j}\cdot q_j(X)$, 其中
$a_j,b_j\in R^*$
$(1\le j\le n)$. 两边取 $\gcd(\mathrm{cont}(\cdot))$ 知, $b_j\cdot R=b_j\cdot\gcd(\mathrm{cont}(\widetilde{p_{i_j}}(X)))=a_j\cdot\gcd(\mathrm{cont}(q_j(X)))=a_j\cdot R$, 则
$\dfrac{a_j}{b_j}\in R^\times$.
"(2)⇒(1)": 由 $\mathrm{UF}(R)=R\cap\mathrm{UF}(R[X])$ 即知. □

最后我们讨论一元多项式环的分式域上的赋值, 它可视为前述环性质的一个应用.

定义 5.1.8 (域上的赋值) 设 F 为一个域, $(\Gamma,0,+;\le)$ 为一个全序 Abel 群. 引入 ∞ 满足以下性质:
(1) $\forall\alpha\in\Gamma$, $\alpha+\infty=\infty+\alpha=\infty+\infty=\infty$; (2) $\forall\alpha\in\Gamma$, $\alpha\le\infty$.
若映射 $v\colon F\to\Gamma\cup\{\infty\}$ 满足以下性质: $\forall a,b\in F$,
(1) $v(a)=\infty\iff a=0$; (2) $v(ab)=v(a)+v(b)$; (3) $v(a+b)\ge\min\{v(a),v(b)\}$, 且当 $v(a)\ne v(b)$ 时取等号,
则称 v 为域 F 上的一个**赋值** (valuation).

注: 设 $v\colon F\to\Gamma\cup\{\infty\}$ 为域 F 上的一个赋值,
(1) 若 $\forall a\in F^*$, $v(a)=0$, 则称 v 是平凡的;
(2) 记 $\Gamma_v\coloneqq v(F^*)\le\Gamma$ 为 v 的赋值群. 若 $\Gamma_v\subseteq\mathbb{Z}$, 则称 v 是离散的;
(3) 记 $R_v\coloneqq\{a\in F\colon v(a)\ge0\}$ 为 v 的赋值环, 则它具有唯一的极大理想 $\mathfrak{m}_v\coloneqq\{a\in F\colon v(a)>0\}$, 它的乘法可逆群为 $R_v^\times=R_v\setminus\mathfrak{m}_v=\{a\in F^*\colon v(a)=0\}$. 记 $k_v\coloneqq R_v/\mathfrak{m}_v$ 为 v 的剩余类域.

例 5.1.4 ($\mathbb{Q}$ 上的赋值) 设 p 为素数, 则映射 $v_p\colon$ $\begin{array}{ccc}\mathbb{Q}&\longrightarrow&\mathbb{Z}\cup\{\infty\}\\0&\longmapsto&\infty\end{array}$ 是
$$\frac{a}{b}\,(a,b\in\mathbb{Z}^*)\longmapsto\max\{e\in\mathbb{N}\colon p^e\mid a\}-\max\{e\in\mathbb{N}\colon p^e\mid b\}$$
$\mathbb{Q}$ 上的一个赋值, 称为 $\mathbb{Q}$ 上的 p-adic 赋值. 它的赋值环为 $\mathbb{Z}_{(p)}\coloneqq\left\{\dfrac{a}{b}\in\mathbb{Q}\colon a,b\in\mathbb{Z},\ \gcd(a,b)=1,\ p\nmid b\right\}$, 其中唯一的极大理想为 $\mathfrak{m}_p=p\cdot\mathbb{Z}_{(p)}$, 则剩余类域为 $k_p=\mathbb{Z}_{(p)}/p\cdot\mathbb{Z}_{(p)}\cong\mathbb{Z}/p\cdot\mathbb{Z}$.

事实上, $\mathbb{Q}$ 上的任意非平凡赋值均形如某个 v_p 的正数倍. 这是因为, 设 $v\colon\mathbb{Q}\to\Gamma\cup\{\infty\}$ 为一个赋值, 则 $\mathfrak{m}_v\cap\mathbb{Z}$ 是 $\mathbb{Z}$ 中的一个素理想. 若 $\mathfrak{m}_v\cap\mathbb{Z}=\{0\}$, 则可验证 v 是平凡的; 若 $\mathfrak{m}_v\cap\mathbb{Z}=p\cdot\mathbb{Z}$, 其中 p 为素数, 则
$\forall e\in\mathbb{N}$,
$\forall k\in\mathbb{Z}\setminus p\cdot\mathbb{Z}$, $v(p^e\cdot k)=\underbrace{v(p)+\cdots+v(p)}_{e}+v(k)=e\cdot v(p)$, 故 $v=v(p)\cdot v_p$, 其中 $v(p)>0$.

例 5.1.5 (分式域 $F(X)$ 上的赋值) 设 F 为一个域, $p(X)\in F[X]$ 为不可约多项式, 则映射
$$v_{p(X)}\colon\ \begin{array}{ccc}F(X)&\longrightarrow&\mathbb{Z}\cup\{\infty\}\\0&\longmapsto&\infty\end{array}$$
$$\frac{f(X)}{g(X)}\,(f(X),g(X)\in F[X]^*)\longmapsto\max\{e\in\mathbb{N}\colon p(X)^e\mid f(X)\}-\max\{e\in\mathbb{N}\colon p(X)^e\mid g(X)\}$$
是 $F(X)$ 上的一个赋值, 它衡量了有理多项式关于 $p(X)$ 的阶数. 特别地, 当 $p(X)=X-c$ $(c\in F)$ 时, $v_{p(X)}$ 衡量了有理多项式在 $X=a$ 处的零点 (为正) 或极点 (为负) 的阶数. 它的赋值环为
$$F[X]_{(p(X))}\coloneqq\left\{\frac{f(X)}{g(X)}\in F(X)\colon f(X),g(X)\in F[X],\ \gcd(f(X),g(X))=1,\ p(X)\nmid g(X)\right\},$$
其中唯一的极大理想为 $\mathfrak{m}_{p(X)}=p(X)\cdot F[X]_{(p(X))}$, 则剩余类域为
$$k_{p(X)}=F[X]_{(p(X))}/p(X)\cdot F[X]_{(p(X))}\cong F[X]/p(X)\cdot F[X].$$

值得注意的是, 分式域 $F(X)$ 上还有更多的赋值: 例如映射

$$\begin{aligned} v_\infty\colon \quad & F(X) \longrightarrow \mathbb{Z}\cup\{\infty\} \\ & 0 \longmapsto \infty \end{aligned},$$

$$\frac{f(X)}{g(X)}\,(f(X),g(X)\in F[X]^*) \longmapsto \deg(g(X))-\deg(f(X))$$

它衡量了有理多项式在 ∞ 处的零点 (为正) 或极点 (为负) 的阶数.

更进一步地, 若 $v\colon F\to\Gamma\cup\{\infty\}$ 为一个赋值, 任取 $\alpha\in\Gamma$, 记映射

$$\begin{aligned} v_\alpha\colon \quad & F(X) \longrightarrow \Gamma\cup\{\infty\} \\ & 0 \longmapsto \infty \end{aligned},$$

$$\frac{f(X)}{g(X)}\,(f(X),g(X)\in F[X]^*) \longmapsto v_\alpha(f(X))-v_\alpha(g(X))$$

其中 $v_\alpha(f(X)) \coloneqq \min\{v(a_i)+i\alpha\colon 0\le i\le n\}$, 若 $f(X)=a_nX^n+\cdots+a_1X+a_0$. 于是可直接验证 v_α 也为 $F(X)$ 上的一个赋值, 它可视为域 F 上赋值 v 到分式域 $F(X)$ 上的延拓.

命题 5.1.11 设 F 为一个域, 则

$$\begin{aligned} &\{v : v\text{为}F(X)\text{上的非平凡赋值, 且}v(F^*)=\{0\}\} \\ =&\{\alpha\cdot v_{p(X)}\colon \alpha>0, p(X)\in F[X]\text{为不可约多项式}\}\cup\{\alpha\cdot v_\infty\colon \alpha>0\}. \end{aligned}$$

证明: "$\supseteq$" 显然; 下证 "$\subseteq$": 设 $v\colon F(X)\to\Gamma\cup\{\infty\}$ 为一个赋值, 且 $v(F^*)=\{0\}$. 若 $v(X)<0$, 则 $v(X^{-1})>0$, 故 $\forall g(X^{-1})=b_kX^{-k}+b_{k+1}X^{-k-1}+\cdots+b_mX^{-m}\in F[X^{-1}]$ $(k\in\mathbb{N}, b_k\neq 0)$, $v(g(X^{-1}))=k\cdot v(X^{-1})$. 注意到 $\forall f(X)\in F[X]^*$, $X^{-\deg(f(X))}\cdot f(X)\in F[X^{-1}]$ 且它的常数项非零, 则由前知 $v(X^{-\deg(f(X))}\cdot f(X))=0$, 故 $v(f(X))=-v(X^{-\deg(f(X))})=-\deg(f(X))\cdot v(X^{-1})$, 因此 $v=v(X^{-1})\cdot v_\infty$, 其中 $v(X^{-1})>0$.

若 $v(X)\ge 0$, 则 $\forall f(X)\in F[X]$, $v(f(X))\ge 0$, 即 $F[X]\subseteq R_v$, 故 $\mathfrak{m}_v\cap F[X]$ 是 $F[X]$ 中的一个素理想. 若 $\mathfrak{m}_v\cap F[X]=\{0\}$, 则可验证 v 是平凡的; 若 $\mathfrak{m}_v\cap F[X]=p(X)\cdot F[X]$, 其中 $p(X)\in F[X]$ 是不可约多项式, 则 $\forall e\in\mathbb{N}$, $\forall g(X)\in F[X]\setminus p(X)\cdot F[X]$, $v(p(X)^e\cdot g(X))=\underbrace{v(p(X))+\cdots+v(p(X))}_{e}+v(g(X))=e\cdot v(p(X))$, 故 $v=v(p(X))\cdot v_{p(X)}$, 其中 $v(p(X))>0$. □

通常一般域上的赋值也可以指数化, 得到一般域上的绝对值.

定义 5.1.9 (域上的绝对值) 设 F 为一个域, 若映射 $|\cdot|\colon F\to\mathbb{R}$ 满足以下性质:
(1) $\forall a\in F$, $|a|\ge 0$, 且 $|a|=0\iff a=0$; (2) $\forall a,b\in F$, $|a\cdot b|=|a|\cdot|b|$; (3) $\forall a,b\in F$, $|a+b|\le|a|+|b|$,
则称 $|\cdot|$ 为域 F 上的一个**绝对值** (absolute value).

注: 设 $|\cdot|$ 为域 F 上的一个绝对值,
(1) 若 $\forall a\in F^*$, $|a|=1$, 则称 $|\cdot|$ 是平凡的;
(2) 若 $\{|n|\in\mathbb{R}\colon n\in\mathbb{Z}\}$ 是有界集, 则称 $|\cdot|$ 是非 Archimedean; 否则称 $|\cdot|$ 是 Archimedean.

引理 5.1.12 设 $|\cdot|$ 为域 F 上的一个绝对值, 则以下条件等价:
(1) $|\cdot|$ 是非 Archimedean;
(2) $\forall a,b\in F$, $|a+b|\le\max\{|a|,|b|\}$, 且当 $|a|\neq|b|$ 时取等号;
(3) 记 $-\log 0=\infty$, 则 $-\log|\cdot|\colon F\to\mathbb{R}\cup\{\infty\}$ 是域 F 上的一个赋值.

证明: "(1)⇒(2)": 设 $|\cdot|$ 是非 Archimedean, 即 $\exists M>0$, $\forall n\in\mathbb{Z}$, $|n|\le M$. 任取 $a,b\in F$, 不妨设 $|a|\ge|b|$, 则 $\forall 0\le i\le n$, $|a|^{n-i}|b|^i\le|a|^n$, 故 $|a+b|^n=|(a+b)^n|=\left|\sum\limits_{i=0}^{n}\binom{n}{i}a^{n-i}b^i\right|\le\sum\limits_{i=0}^{n}|\binom{n}{i}||a|^{n-i}|b|^i\le(n+1)M|a|^n$, 因此 $|a+b|\le(n+1)^{\frac{1}{n}}M^{\frac{1}{n}}|a|$. 令 $n\to+\infty$ 知, $|a+b|\le|a|=\max\{|a|,|b|\}$. 现设 $|a|>|b|$, 由 $|1|=|-1|=1$ 知, $|a|=|(a+b)+(-b)|\le\max\{|a+b|,|-b|\}=\max\{|a+b|,|b|\}$, 则 $|a|=|a+b|$.
"(2)⇒(1)": 设 $\forall a,b\in F$, $|a+b|\le\max\{|a|,|b|\}$, 则 $\forall n\in\mathbb{Z}$, $|n|\le\max\{\underbrace{|1|,\cdots,|1|}_{n}\}=|1|$, 故 $|\cdot|$ 是非 Archimedean;
"(2)⇔(3)": 显然. □

例 5.1.6 ($\mathbb{Q}$ 上的绝对值)

(1) $\{\mathbb{Q}$上的非平凡非 Archimedean 绝对值$\} = \{q^{-v_p(\cdot)}\colon q>1, p$为素数$\}$;

(2) $\{\mathbb{Q}$上的 Archimedean 绝对值$\} = \{|\cdot|_\infty^s\colon s\in(0,1]\}$, 其中 $|\cdot|_\infty$ 为通常的绝对值.

证明: (1) 由前例与上述引理即知;

(2) 显然 $|\cdot|_\infty^s\,(s\in(0,1])$ 是 $\mathbb{Q}$ 上的 Archimedean 绝对值; 反之, 设 $|\cdot|$ 是 $\mathbb{Q}$ 上的 Archimedean 绝对值, 先断言: $\forall n\in\mathbb{N}^*(n\geq 2)$, $|n|>1$. (这是因为: 假设 $\exists n\in\mathbb{N}^*(n\geq 2)$, $s.t.$ $|n|\leq 1$, 则任取 $m\in\mathbb{N}^*$, 记 $m=a_0+a_1n+\cdots+a_rn^r$, 其中 $a_0,\cdots,a_r\in\{0,1,\cdots,n-1\}$, 且 $m\geq n^r$, 则 $|a_i|\leq a_i|1|\leq n-1$, 且 $r\leq\dfrac{\log m}{\log n}$, 故 $|m|\leq\sum\limits_{i=0}^{r}|a_i||n|^i$ $\leq(r+1)(n-1)\leq\left(\dfrac{\log m}{\log n}+1\right)(n-1)$. 特别地, 将 m 换成 m^k 并取 k 次方根得, $|m|\leq\left(\dfrac{k\log m}{\log n}+1\right)^{\frac{1}{k}}(n-1)^{\frac{1}{k}}$; 再令 $k\to+\infty$, 则 $|m|\leq 1$, 这与 $|\cdot|$ 是 Archimedean 绝对值矛盾!)

再断言: $\forall m,n\in\mathbb{N}^*(m,n\geq 2)$, $|m|^{\frac{1}{\log m}}=|n|^{\frac{1}{\log n}}$. (这是因为: 记 $m=a_0+a_1n+\cdots+a_rn^r$, 其中 $a_0,\cdots,a_r$ $\in\{0,1,\cdots,n-1\}$, 且 $m\geq n^r$, 则 $|a_i|\leq a_i|1|\leq n-1$, 且 $r\leq\dfrac{\log m}{\log n}$, 故 $|m|\leq\sum\limits_{i=0}^{r}|a_i||n|^i\leq(r+1)(n-1)|n|^r$ $\leq\left(\dfrac{\log m}{\log n}+1\right)(n-1)|n|^{\frac{\log m}{\log n}}$. 特别地, 将 m 换成 m^k 并取 k 次方根得, $|m|\leq\left(\dfrac{k\log m}{\log n}+1\right)^{\frac{1}{k}}(n-1)^{\frac{1}{k}}|n|^{\frac{\log m}{\log n}}$; 再令 $k\to+\infty$, 则 $|m|\leq|n|^{\frac{\log m}{\log n}}$. 同理可知 $|n|\leq|m|^{\frac{\log n}{\log m}}$, 故 $|m|^{\frac{1}{\log m}}=|n|^{\frac{1}{\log n}}$.)

因此取 $s=\dfrac{\log|2|}{\log 2}\in(0,1]$, 则 $\forall n\in\mathbb{N}^*(n\geq 2)$, $\dfrac{\log|n|}{\log n}=s$, 即 $|n|=n^s$, 故 $|\cdot|=|\cdot|_\infty^s$. □

例 5.1.7 ($F(X)$ 上的绝对值) 设 F 为一个域, 则

$$\begin{aligned}&\{|\cdot|:|\cdot|\text{为}F(X)\text{上的非平凡绝对值, 且}|F^*|=\{1\}\}\\=&\{|\cdot|:|\cdot|\text{为}F(X)\text{上的非平凡非 Archimedean 绝对值, 且}|F^*|=\{1\}\}\\=&\{q^{-v_{p(X)}(\cdot)}\colon q>1, p(X)\in F[X]\text{为不可约多项式}\}\cup\{q^{-v_\infty(\cdot)}\colon q>1\}.\end{aligned}$$

证明: 由 $|F^*|=\{1\}$ 知, $|\cdot|$ 非 Archimedean, 故由前例与上述引理即知结论. □

命题 5.1.13 (绝对值的乘积公式)

(1) 记 $|\cdot|_p=p^{-v_p(\cdot)}$ (p 为素数), 则 $\prod\limits_{p\text{为素数}}|\cdot|_p\cdot|\cdot|_\infty=1$ 在 $\mathbb{Q}^*$ 上成立;

(2) 设 F 为一个域, 固定 $q>1$, 记 $|\cdot|_{p(X)}=q^{-\deg(p(X))\cdot v_{p(X)}(\cdot)}$ ($p(X)\in F[X]$ 为不可约多项式), $|\cdot|_\infty=q^{-v_\infty(\cdot)}$, 则 $\prod\limits_{p(X)\in F[X]\text{为不可约多项式}}|\cdot|_{p(X)}\cdot|\cdot|_\infty=1$ 在 $F[X]^*$ 上成立.

证明: 由 $\mathbb{Q}$ 与 $F[X]$ 为唯一分解整环即知. □

参考文献与补注 5.1

(1) 关于整环等定义定理的部分, 可以参考 N. Jacobson“Basic Algebra I”.

(2) 关于域上的赋值与绝对值的部分, 可以参考 J. Neukirch“Algebraic Number Theory”.

§5.2 多项式的根与形式导数

5.2.1 多项式的根与重根

设 R 为一个环, $f(X)=\sum\limits_{k=0}^{n}a_kX^k\in R[X]$, 则多项式的取值映射给出了环 R 上的多项式函数 $f_R\colon\begin{array}{ccc}R&\longrightarrow&R\\r&\longmapsto&\sum\limits_{k=0}^{n}a_kr^k\end{array}$.

若 $r\in R$ 满足 $f_R(r)=0$, 则称 r 为 $f(X)$ 的一个**根** (root). 关于一元多项式与它的根的关系, 一个基本的结果如下:

定理 5.2.1 (Talyor 公式) *设 R 为一个交换环, $f(X)\in R[X]$, $n=\deg(f(X))\in\mathbb{N}$, 则存在 $\{c_i(X)\}_{i=0}^n\subseteq R[X]$, 满足:*

(1) $\forall 0\leq i\leq n$, $\deg(c_i(X))=n-i$;

(2) $\forall r \in R,\ f(X) = \sum\limits_{i=0}^{n} (c_i)_R(r) \cdot (X-r)^i$.

证明: 记 $f(X) = \sum\limits_{k=0}^{n} a_k X^k$. 由二项式定理知, $\forall 0 \le k \le n, \forall r \in R,\ X^k = (r+(X-r))^k = \sum\limits_{i=0}^{k} \binom{k}{i} r^{k-i}(X-r)^i$, 则 $f(X) = \sum\limits_{k=0}^{n} a_k \sum\limits_{i=0}^{k} \binom{k}{i} r^{k-i}(X-r)^i = \sum\limits_{i=0}^{n} \left(\sum\limits_{k=i}^{n} a_k \binom{k}{i} r^{k-i} \right) (X-r)^i$. 记 $c_i(X) = \sum\limits_{k=i}^{n} a_k \binom{k}{i} X^{k-i}\ (0 \le i \le n)$ 即可. □

推论 5.2.2 *设 R 为一个交换环, $f(X) \in R[X]$, 则 $r \in R$ 为 $f(X)$ 的一个根 $\iff (X-r) \mid f(X)$.*

记 $m(r, f(X)) \coloneqq \max\{k \in \mathbb{N} \colon (X-r)^k \mid f(X)\}$ 为 $r \in R$ 作为 $f(X)$ 的根的**重数** (multiplicity).

推论 5.2.3 *设 R 为一个交换环, $f(X) \in R[X]$, $\{c_i(X)\}_{i=0}^{\deg(f(X))} \subseteq R[X]$ 为 Taylor 公式中所述, $m \in \mathbb{N}$, 则*

(1) *$r \in R$ 为 $f(X)$ 的 m 重根 $\iff r$ 为 $\gcd(c_0(X), \cdots, c_{m-1}(X))$ 的根, 且不为 $c_m(X)$ 的根.*

(2) *特别地, $f(X)$ 有重数 ≥ 2 的根 $\iff \gcd(c_0(X), c_1(X))$ 有根.*

注: 事实上, 注意到 $c_0(X) = f(X)$; $c_1(X) = \sum\limits_{k=1}^{n} k a_k X^{k-1}$ 为 $f(X)$ 的形式导数.

另外, 一元多项式的形式导数也可由以下的代数性质抽象刻画:

命题 5.2.4 *设 R 是一个交换环, $D \in L(R[X]; R[X])$, 满足:*

(1) *$D(1) = 0,\ D(X) = 1$;*

(2) *(Leibniz) $D(f(X) \cdot g(X)) = D(f(X)) \cdot g(X) + f(X) \cdot D(g(X)),\ \forall f(X), g(X) \in R[X]$;*

则 D 为多项式的形式求导.

证明: 直接归纳证明 $D(X^n) = nX^{n-1},\ \forall n \in \mathbb{N}^*$, 再由 D 的线性性即知. □

5.2.2 代数学基本定理

回忆代数学最初的主要任务是解代数方程, 即判断一个多项式是否有根并求出它的所有根. 关于一元多项式的次数和它的根的关系, 一个基本的结果如下:

命题 5.2.5 *设 R 为一个整环, $f(X) \in R[X]$, $n = \deg(f(X)) \in \mathbb{N}$, 则 $f(X)$ 在 R 中至多有 n 个不同的根.*

证明: 记 $f(X) = \sum\limits_{k=0}^{n} a_k X^k$. 假设 $f(X)$ 在 R 中有 $(n+1)$ 个不同的根 $r_1, \cdots, r_{n+1}$, 则 $f_R(r_i) = 0\,(1 \le i \le n+1)$ 等价于 $\begin{pmatrix} 1 & r_1 & \cdots & r_1^n \\ \vdots & \vdots & \dots & \vdots \\ 1 & r_{n+1} & \cdots & r_{n+1}^n \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = 0$. 记 $V_{n+1} = \begin{pmatrix} 1 & r_1 & \cdots & r_1^n \\ \vdots & \vdots & \dots & \vdots \\ 1 & r_{n+1} & \cdots & r_{n+1}^n \end{pmatrix} \in R^{(n+1)\times(n+1)}$ 为 Vandermonde 阵, 则 $\operatorname{adj}(V_{n+1}) \cdot V_{n+1} = \det(V_{n+1}) \cdot I_{n+1}$, 且 $\det(V_{n+1}) \in R^*$. 将原式两端左乘 $\operatorname{adj}(V_{n+1})$ 知, $\det(V_{n+1}) \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = 0$; 再由 R 的无非零的零因子性知, $a_0 = \cdots = a_n = 0$, 即 $f(X) = 0$, 矛盾! □

注: 上述命题对于一般交换环上的多项式不成立, 反例如 $R = \mathbb{Z}/8\mathbb{Z}$, $f(X) = X^2 - \overline{1} \in R[X]$ 在 R 中有 4 个不同的根: $\overline{1}, \overline{3}, \overline{5}, \overline{7}$.

设 R 为一个交换环, $f(X) \in R[X]$ 且 $\deg(f(X)) \in \mathbb{N}^*$, 有时 $f(X)$ 在 R 中未必有根. 一个简单的办法是, 考虑商环 $S = R[X]/f(X) \cdot R[X]$ 以及环嵌入 $\begin{array}{ccc} R & \hookrightarrow & S \\ r & \longmapsto & r + f(X) \cdot F[X] \end{array}$, 则 $f(X) \in S[X]$ 在 S 中必有根 $X + f(X) \cdot F[X]$. 为简单起见, 以下只考虑域 F 上一元多项式的根.

定义 5.2.1 (多项式的分裂域) 设 F 为一个域, $f(X) \in F[X]$, 若域扩张 E/F 满足 $f(X) \in E[X]$ 可分解成 $E[X]$ 中若干一次因式的乘积, 则称 $f(X)$ 在域 E 上**分裂** (split). 此时使得 $f(X)$ 在域 E 上分裂的最小域扩张 E/F 称为 $f(X)$ 在 F 上的一个**分裂域** (splitting field).

注:

(1) 域上多项式的分裂域必存在. 这是因为, 对多项式的次数归纳知, 只需证明域上不可约多项式必在某个域扩张上有根; 而这由上述构造显然. 特别地, 域上 n 次多项式的分裂域的扩张次数必整除 $n!$.

(2) 域上多项式的分裂域在同构意义下是唯一的, 这是域论中同构延拓定理的结果.

进一步地, 由 Zorn 引理可证明, 域上任意一族非常数多项式的分裂域必存在唯一. 关于分裂域次数的有限性, 一个基本的观察如下:

引理 5.2.6 *设 E/F 为域扩张, 则以下条件等价:*

(1) *$[E:F]<+\infty$;*

(2) *$E=F(\alpha_1,\cdots,\alpha_k)$, 其中 $\alpha_1,\cdots,\alpha_k\in E$ 均为域 F 上某个多项式的根.*

证明: "(1)⇒(2)": 设 $[E:F]<+\infty$, 即 $\dim_F(E)<+\infty$, 可取 E 的一组 F-基为 $\{\alpha_i\}_{i=1}^k$, 则 $E=F(\alpha_1,\cdots,\alpha_k)$. 由 $\dim_F(E)<+\infty$ 知, $\forall 1\leq i\leq k$, $\{\alpha_i^j\}_{j=0}^{\dim_F(E)}\subseteq E$ 是 F-线性相关的, 即 α_i 均为域 F 上某个多项式的根.
"(2)⇒(1)": 设 $E=F(\alpha_1,\cdots,\alpha_k)$, 其中 $\alpha_i\in E$ 为 $f_i(X)\in F[X]$ 的根, 则 $[F(\alpha_i):F]\leq\deg(f_i(X))<+\infty$, 故 $[E:F]=\prod\limits_{i=1}^k[F(\alpha_1,\cdots,\alpha_i):F(\alpha_1,\cdots,\alpha_{i-1})]\leq\prod\limits_{i=1}^k[F(\alpha_i):F]<+\infty$. □

命题 5.2.7 *设 E/F 为域扩张, 若域 E 为 $F[X]\setminus F$ 中多项式的分裂域, 则域 E 为 $E[X]\setminus E$ 中多项式的分裂域.*

证明: 任取 $f(X)\in E[X]\setminus E$, 设 $f(X)=\sum\limits_{k=0}^n a_kX^k$, 其中 $n\geq 1$, $a_k\in E$. 记 $K=F(a_0,\cdots,a_n)$, 则 $F\subseteq K\subseteq E$, 且 $f(X)\in K[X]\setminus K$. 由上述引理知, $[K:F]<+\infty$. 另一方面, 任取 α 为 $f(X)$ 的一个根, 则 $[K(\alpha):K]<+\infty$, 故 $[K(\alpha):F]=[K(\alpha):K][K:F]<+\infty$. 由上述引理知, α 也为域 F 上某个多项式的根, 故 $\alpha\in E$, 因此 $f(X)$ 在 E 上分裂. □

注:

(1) 在上述引理中, 域 E 称为域 F 的**代数闭包** (algebraic closure), 于是代数闭包必为代数闭域.

(2) 上述引理的逆命题不成立, 反例如 $\overline{\mathbb{Q}}^{\mathrm{alg}}\subsetneqq\mathbb{C}$ 都是域 $\mathbb{Q}$ 上的代数闭域.

定理 5.2.8 (代数学基本定理) *$\mathbb{C}$ 为代数闭域.*

证明: 一个纯线性代数的证明可见 H. Derkson"The Fundamental Theorem of Algebra and Linear Algebra"(2003). □

例 5.2.1 (素谱空间) 设 F 为一个代数闭域, 则 $F[X]$ 中的不可约多项式都是一次的, 即 $F[X]$ 中的极大理想都形如 $(X-c)\cdot F[X]\,(c\in F)$, 也即 $F[X]$ 中的素理想都形如 $\{0\}$ 或 $(X-c)\cdot F[X]\,(c\in F)$. 记

$$\mathrm{Spec}(F[X])\coloneqq\{F[X]\text{中的素理想}\}=\{\{0\}\}\cup\{(X-c)\cdot F[X]:c\in F\};$$

$$V(\mathfrak{a})\coloneqq\{F[X]\text{中包含}\mathfrak{a}\text{的素理想}\},\ \text{其中}\mathfrak{a}\subseteq F[X]\text{为理想}.$$

考虑在 $\mathrm{Spec}(F[X])$ 上赋予拓扑如下: 子集 $V\subseteq\mathrm{Spec}(F[X])$ 为闭集 $\Longleftrightarrow V$ 形如 $V(\mathfrak{a})$, 其中 $\mathfrak{a}\subseteq F[X]$ 为理想, 则可直接验证此定义符合闭集公理, 于是此拓扑称为 $\mathrm{Spec}(F[X])$ 上的 Zariski 拓扑, 此拓扑空间称为一元多项式环上的 **素谱空间** (spectrum space). 注意存在集合嵌入 $\begin{aligned}F&\lhook\joinrel\longrightarrow\mathrm{Spec}(F[X])\\c&\longmapsto(X-c)\cdot F[X]\end{aligned}$, 故 $\mathrm{Spec}(F[X])$ 上的 Zariski 拓扑在子集 F 上的限制可给出 F 上的 Zariski 拓扑. 具体地说, 子集 $V\subseteq F$ 为 Zariski-闭集 $\Longleftrightarrow V$ 为某个 $f(X)\in F[X]$ 的根集 $\Longleftrightarrow V$ 为有限集或 $F[X]$.

注: 上述素谱空间的概念也可推广到代数闭域上的多元多项式环上, 这是代数几何中的基本模型.

5.2.3 多项式的形式导数与结式

本节首先利用一元多项式的形式导数, 给出重根和重公因式的进一步判别法.

命题 5.2.9 *设 F 为一个域, $f(X),g(X)\in F[X]$, 则以下条件等价:*

(1) *$\gcd(f(X),g(X))=1$;*

(2) *$f(X)$ 与 $g(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无公共根;*

(3) $f(X)$ 与 $g(X)$ 在 $F[X]$ 内无不可约公因式.

证明: "(1)⇒(2)": 设 $\gcd(f(X),g(X))=1$, 由 Bezout 定理知, $\exists u(X),v(X)\in F[X],\ s.t.\ u(X)f(X)+v(X)g(X)=1$. 假设 $f(X)$ 与 $g(X)$ 存在公共根 $x_0\in\overline{F}^{\mathrm{alg}}$, 代入上式知 $0=1$, 矛盾!
"(2)⇒(3)": 设 $f(X)$ 与 $g(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无公共根, 假设 $f(X)$ 与 $g(X)$ 在 $F[X]$ 内存在不可约公因式 $p(X)$, 则 $p(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上的根为 $f(X)$ 与 $g(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上的公共根, 矛盾!
"(3)⇒(1)": 显然. □

推论 5.2.10 设 F 为一个域, $f(X)\in F[X]$, 则以下条件等价:

(1) $\gcd(f(X),f'(X))=1$;

(2) $f(X)$ 与 $f'(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无公共根;

(3) $f(X)$ 与 $f'(X)$ 在 $F[X]$ 内无不可约公因式;

(4) $f(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无重根.

证明: "(1)⇔(2)⇔(3)": 由上述命题即知; "(4)⇔(2)": 由第一小节命题即知. □

注: 显然 "(4)⇒(5) $f(X)$ 在 $F[X]$ 内无重不可约公因式"; 但 "(5)⇒(4)" 未必成立, 见以下的引理.

引理 5.2.11 设 F 是一个域, $p(X)\in F[X]$ 为不可约多项式,

(1) 若 $\mathrm{char}(F)=0$, 则 $\gcd(p(X),p'(X))=1$;

(2) 若 $\mathrm{char}(F)=p$ 为素数, 则 $\gcd(p(X),p'(X))\neq 1\iff p'(X)=0\iff p(X)=q(X^p)$, 其中 $q(X)\in F[X]$ 也为不可约多项式.

证明: (1) 设 $\mathrm{char}(F)=0$, 由 $p(X)\in F[X]$ 为不可约多项式知, $\deg(p(X))\in\mathbb{N}^*$, 则 $\deg(p'(X))=\deg(p(X))-1\in\mathbb{N}$, 故 $p'(X)\neq 0$, 因此 $\gcd(p(X),p'(X))=1$.
(2) 设 $\mathrm{char}(F)=p$ 为素数, 由 $p(X)\in F[X]$ 为不可约多项式知, $\gcd(p(X),p'(X))\neq 1\iff\gcd(p(X),p'(X))=p(X)$ $\iff p'(X)=0$. 进一步地, $p(X)$ 中非零单项式 a_kX^k 求导为 $0\iff p\mid k$, 故 $p'(X)=0\iff p(X)=q(X^p)$, 其中 $q(X)\in F[X]$, 再由 $p(X)$ 不可约知 $q(X)$ 也不可约. □

注:

(1) 若 $\mathrm{char}(F)=0$ 或 $F=\overline{F}^{\mathrm{alg}}$, 则任意不可约多项式 $p(X)\in F[X]$ 均满足 $\gcd(p(X),p'(X))=1$, 由上述推论知 $p(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无重根. 考虑 $f(X)\in F[X]$ 的不可约分解, 若 $f(X)$ 在 $F[X]$ 内无重不可约公因式, 则 $f(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上无重根, 即此时 "(5)⇒(4)" 成立.

(2) 若 $\mathrm{char}(F)=p$ 为素数, 取 $p(X)\in F[X]$ 为不可约多项式且 $p'(X)=0$, 则由上述推论知, $p(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上有重根, 即此时 "(5)⇒(4)" 不成立. 一个具体的例子是 $F=\mathbb{F}_p(t)$, $p(X)=X^p-t$.

接着我们补充一元多项式的结式的概念, 它可通过多项式的系数定量地刻画重根.

定义 5.2.2 (多项式的结式) 设 F 为一个域, 记 $f(X)=\sum\limits_{k=0}^{n}a_kX^k$, $g(X)=\sum\limits_{l=0}^{m}b_lX^l\in F[X]$, 其中 $n=\deg(f(X))$,
$m=\deg(g(X))\in\mathbb{N}$, 以及 $V_i:=\{h(X)\in F[X]\colon \deg(h(X))<i\}\,(i\in\mathbb{N})$. 考虑线性映射 (警告: 不是双线性映射!)

$$\begin{aligned} T\colon\quad V_m\coprod V_n &\longrightarrow V_{m+n} \\ (h_1(X),h_2(X))&\longmapsto f(X)\cdot h_1(X)+g(X)\cdot h_2(X)\end{aligned},$$

它在有序基 $\{(X^{m-l},0)\}_{l=1}^{m}\cup\{(0,X^{n-k})\}_{k=1}^{n}$ 与 $\{X^{m+n-i}\}_{i=1}^{m+n}$ 下的矩阵表示称为 $f(X)$ 与 $g(X)$ 的 **Sylvester 方阵**:

$$\begin{pmatrix} a_n & 0 & \cdots & 0 & b_m & 0 & \cdots & 0\\ a_{n-1} & a_n & \ddots & 0 & b_{m-1} & b_m & \ddots & 0\\ a_{n-2} & a_{n-1} & \ddots & 0 & b_{m-2} & b_{m-1} & \ddots & 0\\ \vdots & \vdots & \ddots & a_n & \vdots & \vdots & \ddots & b_m\\ a_0 & a_1 & \cdots & \vdots & b_0 & b_1 & \cdots & \ddots\\ 0 & a_0 & \ddots & \vdots & 0 & b_0 & \ddots & \vdots\\ \vdots & \vdots & \ddots & a_1 & \vdots & \vdots & \ddots & b_1\\ 0 & 0 & \cdots & a_0 & 0 & 0 & \ddots & b_0 \end{pmatrix},$$

该方阵的行列式 $\det(T)=:\mathrm{Res}(f(X),g(X))\in F$ 称为 $f(X)$ 与 $g(X)$ 的**结式** (resultant).

注:

(1) 记 $f(X)=a_n\cdot\prod\limits_{k=1}^{n}(X-x_k)$, $g(X)=b_m\cdot\prod\limits_{l=1}^{m}(X-y_l)\in\overline{F}^{\mathrm{alg}}[X]$, 通过对根 $x_k\,(1\le k\le n)$, $y_l\,(1\le l\le m)$ 作代数扰动, 可证明 $\mathrm{Res}(f(X),g(X))=a_n^m b_m^n\cdot\prod\limits_{\substack{1\le k\le n\\1\le l\le m}}(x_k-y_l)$, 故 $\mathrm{Res}(f(X),g(X))=0\iff f(X)$ 与 $g(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上有公共根.

(2) 特别地, 由于 $f'(X)=a_n\cdot\sum\limits_{j=1}^{n}\prod\limits_{\substack{k=1\\k\neq j}}^{n}(X-x_k)$, 则 $\mathrm{Res}(f(X),f'(X))=a_n^{n-1}\cdot\prod\limits_{k=1}^{n}f'_F(x_k)=a_n^{2n-1}\cdot\prod\limits_{k\neq k'}(x_k-x_{k'})$.

记 $\mathrm{disc}(f(X))\coloneqq\dfrac{\mathrm{Res}(f(X),f'(X))}{a_n\cdot(-1)^{\frac{n(n-1)}{2}}}=a_n^{2n-2}\cdot\prod\limits_{1\le k<k'\le n}(x_k-x_{k'})^2\in F$ 为 $f(X)$ 的**判别式** (discriminant), 则

$\mathrm{disc}(f(X))=0\iff f(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 上有重根.

例 5.2.2 (二元高次方程组的解) 设 F 为一个域, $f(X,Y),g(X,Y)\in F[X,Y]$, 考虑二元高次方程组 $\begin{cases}f(X,Y)=0\\g(X,Y)=0\end{cases}$,

记为 $\begin{cases}a_n(Y)X^n+\cdots+a_1(Y)X+a_0(Y)=0\\b_m(Y)X^m+\cdots+b_1(Y)X+b_0(Y)=0\end{cases}$. 注意关于 X 的结式为 $\mathrm{Res}_X(f(X,Y),g(X,Y))\in F[Y]$, 则

$y_0\in F$ 为 $\mathrm{Res}_X(f(X,Y),g(X,Y))=0$ 的根 $\iff\begin{cases}f(X,y_0)=0\\g(X,y_0)=0\end{cases}$ 在 $\overline{F}^{\mathrm{alg}}$ 上有解, 或 $a_n(y_0)b_m(y_0)=0$.

最后我们引入 Wronski 行列式, 以展示多项式的形式导数的广泛应用.

定义 5.2.3 (Wronski 行列式) 设 F 为一个域, $f_1(X),\cdots,f_n(X)\in F[X]$, 则 $f_1(X),\cdots,f_n(X)$ 的 Wronski 行列式为 $W(f_1,\cdots,f_n)(X)\coloneqq\det\begin{pmatrix}f_1(X) & f_2(X) & \cdots & f_n(X)\\ f_1'(X) & f_2'(X) & \cdots & f_n'(X)\\ \vdots & \vdots & \vdots & \vdots\\ f_1^{(n-1)}(X) & f_2^{(n-1)}(X) & \cdots & f_n^{(n-1)}(X)\end{pmatrix}\in F[X]$.

命题 5.2.12 *设 F 为一个域, 且 $\mathrm{char}(F)=0$, 则 $f_1(X),\cdots,f_n(X)\in F[X]$ 线性无关 $\iff W(f_1,\cdots,f_n)(X)\neq 0$.* **证明:** "⇒": 设 $f_1(X),\cdots,f_n(X)\in F[X]$ 线性无关, 则通过适当的线性组合, 即存在 $A\in\mathrm{GL}(n,F)$, 可使得 $(g_1(X),\cdots,g_n(X))=(f_1(X),\cdots,f_n(X))\cdot A$ 满足 $g_1(X),\cdots,g_n(X)\neq 0$, 且 $g_1(X),\cdots,g_n(X)$ 在 0 处的重数两两不同, 记为 $m_1,\cdots,m_n\in\mathbb{N}$. 此时 $W(g_1,\cdots,g_n)(X)=W(f_1,\cdots,f_n)(X)\cdot\det(A)$. 由于 $\forall 1\le j\le n$, $g_j(X)\equiv a_jX^{m_j}\pmod{X^{m_j+1}}$, 则 $\forall 1\le i\le n$, $g_j^{(i-1)}(X)\equiv a_j(m_j)_{i-1}X^{m_j-i+1}\pmod{X^{m_j-i+2}}$, 其中 $(m_j)_{i-1}\coloneqq\dfrac{m_j!}{(m_j-i+1)!}$,

故

$$\begin{aligned} W((g_1,\cdots,g_n)(X)) &= \det(g_j^{(i-1)}(X))_{1\le i,j\le n} \\ &\equiv \prod_{j=1}^{n} a_j \cdot X^{\sum_{j=1}^{n} m_j-\binom{n}{2}} \cdot \det((m_j)_{i-1})_{1\le i,j\le n} \pmod{X^{\sum_{j=1}^{n} m_j-\binom{n}{2}+n}} \\ &= \prod_{j=1}^{n} a_j \cdot X^{\sum_{j=1}^{n} m_j-\binom{n}{2}} \cdot \det(m_j^{i-1})_{1\le i,j\le n} \pmod{X^{\sum_{j=1}^{n} m_j-\binom{n}{2}+n}}, \end{aligned}$$

其中 $\det(m_j^{i-1})_{1\le i,j\le n} = \prod\limits_{1\le i<j\le n}(m_j-m_i) \neq 0$, 因此 $W((g_1,\cdots,g_n)(X)) \not\equiv 0$, 即 $W((f_1,\cdots,f_n)(X)) \not\equiv 0$.
"⇐": 设 $f_1(X),\cdots,f_n(X) \in F[X]$ 线性相关, 则显然 $W(f_1,\cdots,f_n)(X) \not\equiv 0$. □

注: 若 $\mathrm{char}(F)=p$ 为素数, 则上述命题未必成立, 反例如 $f_1(X)=1,\ f_2(X)=X^p$.

定理 5.2.13 (Mason-Stothers) *设 $a(X),b(X),c(X) \in F[X]$ 为两两互素的多项式, 满足 $a(X)+b(X)+c(X)=0$, 且 $a'(X),b'(X),c'(X)$ 不全为零多项式, 则 $\max\{\deg(a(X)),\deg(b(X)),\deg(c(X))\} \le N(a(X)\cdot b(X)\cdot c(X))-1$, 其中 $N(f(X))$ 为 $f(X)\in F[X]$ 在 $\overline{F}^{\mathrm{alg}}$ 中不同根的数量.*
证明: 由 $a(X)+b(X)+c(X)=0$ 知, $W(a,b)(X)=W(b,c)(X)=W(c,a)(X)=:W(X)$. 先断言: $W(X)\not\equiv 0$. (这是因为, 假设 $W(X)\equiv 0$, 即 $W(a,b)(X)=a(X)\cdot b'(X)-b(X)\cdot a'(X)\equiv 0$, 则由 $a(X),b(X)$ 互素知 $a(X)\mid a'(X)$, 故 $a'(X)\equiv 0$. 同理知 $b'(X)\equiv c'(X)\equiv 0$, 矛盾!) 于是 $\gcd(a(X),a'(X)),\gcd(b(X),b'(X)),\gcd(c(X),c'(X)) \mid W(X)$. 又由于 $a(X),b(X),c(X)$ 两两互素, 则 $\gcd(a(X),a'(X))\cdot\gcd(b(X),b'(X))\cdot\gcd(c(X),c'(X)) \mid W(X)$, 故 $\deg(\gcd(a(X),a'(X)))+\deg(\gcd(b(X),b'(X)))+\deg(\gcd(c(X),c'(X))) \le \deg(W(X))$. 另一方面, 注意到 $\deg(\gcd(f(X),f'(X))) \ge \deg(f(X))-N(f(X))$, 以及 $N(a(X)\cdot b(X)\cdot c(X))=N(a(X))+N(b(X))+N(c(X))$, 代入知 $\deg(a(X))+\deg(b(X))+\deg(c(X)) \le N(a(X)\cdot b(X)\cdot c(X))+\deg(W(X))$. 又由定义知 $\deg(W(X)) \le \min\{\deg(a(X))+\deg(b(X)),\deg(b(X))+\deg(c(X)),\deg(c(X))+\deg(a(X))\}-1$, 因此 $\max\{\deg(a(X)),\deg(b(X)),\deg(c(X))\} \le N(a(X)\cdot b(X)\cdot c(X))-1$. □

注:

(1) 上述定理中的不等号可取等号. 例如当 $\mathrm{char}(F)=0$ 时, 取 $a(X)=-(X+1)^n,\ b(X)=X^n,\ c(X)=(X+1)^n-X^n$, 则 $\max\{\deg(a(X)),\deg(b(X)),\deg(c(X))\}=n=1+1+(n-1)-1=N(a(X)\cdot b(X)\cdot c(X))-1$.

(2) 上述定理关于多项式的形式导数 (或视为关于域特征) 的要求是必要的, 反例如当 $\mathrm{char}(F)=p$ 为素数时, 取
$a(X)=-(X+1)^p,\ b(X)=X^p,\ c(X)=1$, 则 $\max\{\deg(a(X)),\deg(b(X)),\deg(c(X))\}=p>1+1+0-1$
$=N(a(X)\cdot b(X)\cdot c(X))-1$.

推论 5.2.14 (Fermat 大定理的多项式版本) *设 F 为一个域, 且 $\mathrm{char}(F)=0$, $n\ge 3$, 则方程 $f(X)^n+g(X)^n=h(X)^n$ 无不全为常数的两两互素的多项式解.*
证明: 假设方程 $f(X)^n+g(X)^n=h(X)^n$ 存在不全为常数的两两互素的多项式解, 记 $a(X)=f(X)^n,\ b(X)=g(X)^n$, $c(X)=-h(X)^n$, 则由 Mason-Stothers 定理知, $n\cdot\max\{\deg(f(X)),\deg(g(X)),\deg(h(X))\} \le N(f(X)\cdot g(X)\cdot h(X))-1$. 特别地, 对上式左端各项求和知, $n(\deg(f(X))+\deg(g(X))+\deg(h(X))) \le 3(N(f(X)\cdot g(X)\cdot h(X))-1)$
$\le 3(\deg(f(X))+\deg(g(X))+\deg(h(X))-1)$, 故 $n<3$, 矛盾! □

参考文献与补注 5.2

(1) 关于多项式的分裂域的部分, 可以参考 P. Morandi"Field and Galois Theory".

(2) 关于素谱空间的几何的部分, 可以参考 R. Hartshorne"Algebraic Geometry".

第 6 章　线性变换的标准形

本章介绍有限维线性空间上线性变换的两种分解方式: 准素分解 (primary decomposition) 与循环分解 (cyclic decomposition). 抽象地说, 它们反映了一元多项式环的性质如何影响其上模的分解性质.

§6.1　线性变换的准素分解

为了将有限维线性空间上的线性变换写成尽可能简单的形式, 我们将它的最小多项式的不可约分解对应于线性空间的扭模分解, 这就是所谓线性变换的准素分解.

6.1.1　特征多项式与最小多项式

定义 6.1.1 (线性变换的特征多项式) 设 V 为域 F 上有限维线性空间, $n=\dim_F(V)$, $T\in L(V)$, 考虑拉回映射

$$\begin{aligned} T^{(i)}\colon\ \Lambda^i(V) &\longrightarrow \Lambda^i(V) \\ L &\longmapsto (T^{(i)}(L)\colon (\alpha_1,\cdots,\alpha_i)\mapsto L(T(\alpha_1),\cdots,T(\alpha_i))) \end{aligned},$$

则 $f_T(X)\coloneqq\sum\limits_{i=0}^{n}(-1)^i\mathrm{tr}(T^{(i)})X^{n-i}\in F[X]$ 称为 T 的**特征多项式** (characteristic polynomial).

注: 上述定义的优点在于, 它显式给出了特征多项式的每项系数, 从而最符合抽象观点 $F[X]\coloneqq\coprod\limits_{n=0}^{+\infty}F$.

引理 6.1.1 *设 V 为域 F 上有限维线性空间, $T\in L(V)$ 在基 B 下的矩阵表示为 $A\in F^{n\times n}$, 则 $f_T(X)=f_A(X)$.*

证明: 设 B 的对偶基为 $\{f_j\}_{j=1}^n$, 则 $\Lambda^i(V)$ 的基为 $\{f_{j_1}\wedge\cdots\wedge f_{j_i}:1\le j_1<\cdots<j_i\le n\}$. 考虑拉回映射

$$\begin{aligned} T^{(i)}\colon\ \Lambda^i(V) &\longrightarrow \Lambda^i(V) \\ L &\longmapsto (T^{(i)}(L)\colon (\alpha_1,\cdots,\alpha_i)\mapsto L(T(\alpha_1),\cdots,T(\alpha_i))) \end{aligned},$$

则 $\forall 1\le j_1<\cdots<j_i\le n$,

$$\begin{aligned} T^{(i)}(f_{j_1}\wedge\cdots\wedge f_{j_i}) &= (f_{j_1}\circ T)\wedge\cdots\wedge(f_{j_i}\circ T) \\ &= \left(\sum_{k_1=1}^{n}A_{j_1,k_1}f_{k_1}\right)\wedge\cdots\wedge\left(\sum_{k_i=1}^{n}A_{j_i,k_i}f_{k_i}\right) \\ &= \sum_{1\le l_1<\cdots<l_i\le n}\left(\sum_{\{k_1,\cdots,k_i\}=\{l_1,\cdots,l_i\}}(-1)^{\tau(k_1,\cdots,k_i)}A_{j_1,k_1}\cdots A_{j_i,k_i}\right)\cdot f_{l_1}\wedge\cdots\wedge f_{l_i} \\ &= \sum_{1\le l_1<\cdots<l_i\le n}\det(A_{\{j_1,\cdots,j_i\},\{l_1,\cdots,l_i\}})\cdot f_{l_1}\wedge\cdots\wedge f_{l_i}, \end{aligned}$$

故 $\mathrm{tr}(T^{(i)})=\sum\limits_{1\le j_1<\cdots<j_i\le n}\det(A_{\{j_1,\cdots,j_i\},\{j_1,\cdots,j_i\}})$. 另一方面, 由行列式的组合定义式知, $\det(X\cdot I_n-A)$ 中 X^{n-i} 项的系数为 $(-1)^i\cdot\sum\limits_{1\le j_1<\cdots<j_i\le n}\det(A_{\{j_1,\cdots,j_i\},\{j_1,\cdots,j_i\}})$. 因此 $f_T(X)=\sum\limits_{i=0}^{n}(-1)^i\mathrm{tr}(T^{(i)})X^{n-i}=\det(X\cdot I_n-A)=f_A(X)$. □

注: 在此意义下, 关于方阵特征多项式的命题也可由上述抽象观点证明.

命题 6.1.2 *设 V,W 为域 F 上有限维线性空间, $T\in L(V,W)$, $S\in L(W,V)$, 则 $X^n\cdot f_{T\circ S}(X)=X^m\cdot f_{S\circ T}(X)$.*

证明: 设 $n=\dim_F(V)$, $m=\dim_F(W)\in\mathbb{N}$, 则

$$X^n \cdot f_{T\circ S}(X) = X^n \cdot \sum_{i=0}^{m}(-1)^i \mathrm{tr}((T\circ S)^{(i)})X^{m-i} = \sum_{i=0}^{+\infty}(-1)^i \mathrm{tr}(T^{(i)}\circ S^{(i)})X^{n+m-i};$$
$$X^m \cdot f_{S\circ T}(X) = X^m \cdot \sum_{i=0}^{n}(-1)^i \mathrm{tr}((S\circ T)^{(i)})X^{n-i} = \sum_{i=0}^{\infty}(-1)^i \mathrm{tr}(S^{(i)}\circ T^{(i)})X^{m+n-i}.$$

注意到 $\forall i \geq 0$, $\mathrm{tr}(T^{(i)}\circ S^{(i)}) = \mathrm{tr}(S^{(i)}\circ T^{(i)})$, 故 $X^n \cdot f_{T\circ S}(X) = X^m \cdot f_{S\circ T}(X)$. □

推论 6.1.3 *设 V 为域 F 上有限维线性空间, $T, S \in L(V)$, 则 $f_{T\circ S}(X) = f_{S\circ T}(X)$.*

关于线性变换的特征多项式, 最基本的定理是 Cayley-Hamilton 定理. 为揭露其证明本质, 我们先考虑一个更强而实用的引理:

引理 6.1.4 *设 F 为一个域, $A \in F^{n\times n}$, 则 $\exists B \in F[X]^{n\times n}$, s.t. $p_A(X)\cdot I_n = (X\cdot I_n - A)\cdot B$.*
证明: 记 $p_A(X) = \sum_{i=0}^{d} a_i X^i$, 则由 $p_A(A) = 0$ 知, $p_A(X)\cdot I_n = p_A(X)\cdot I_n - p_A(A) = \sum_{i=0}^{d} a_i(X^i\cdot I_n - A^i)$
$= (X\cdot I_n - A)\cdot \sum_{i=1}^{d}\sum_{j=0}^{i-1} a_i X^j A^{i-1-j}$. 取 $B = \sum_{i=1}^{d}\sum_{j=0}^{i-1} a_i X^j A^{i-1-j} \in F[X]^{n\times n}$ 即可. □

推论 6.1.5 *设 F 为一个域, $A \in F^{n\times n}$, $f(X) \in F[X]$, 则 $f(A) = 0 \iff \exists B \in F[X]^{n\times n}$, s.t. $f(X)\cdot I_n = (X\cdot I_n - A)\cdot B$.*
证明: "⇒": 由 $f(A) = 0 \iff p_A(X) \mid f(X)$ 与上述引理即知.
"⇐": 设 $\exists B \in F[X]^{n\times n}$, s.t. $f(X)\cdot I_n = (X\cdot I_n - A)\cdot B$. 记 $B(X) = B_0 + X\cdot B_1 + \cdots + X^m\cdot B_m$, 其中 $B_i \in F^{n\times n}$, 以及 $f(X) = a_0 + a_1 X + \cdots + a_d X^d$, 其中 $a_i \in F$. 不妨设 $m \geq d$, 则比较系数知 $a_i I_n = B_{i-1} - AB_i$, $\forall 0 \leq i \leq d$; $0 = B_{i-1} - AB_i$, $\forall d+1 \leq i \leq m+1$, 因此 $f(A) = \sum_{i=0}^{d} a_i A^i = \sum_{i=0}^{d} A^i(B_{i-1} - AB_i) = 0$. □

定理 6.1.6 (Cayley-Hamilton) *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则 $f_T(T) = 0$.*
证明: 为证明 $f_T(T) = 0$, 即 $f_A(A) = 0$, 只需求 $B \in F[X]^{n\times n}$ 满足 $(X\cdot I_n - A)\cdot B = f_A(X)\cdot I_n$ 即可. 于是注意到 $B = \mathrm{adj}(X\cdot I_n - A)$ 满足条件. □

推论 6.1.7 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则 $p_T(X) \mid f_T(X) \mid p_T(X)^{\dim_F(V)}$.*
证明: 一方面, 由 $f_T(T) = 0$ 即知 $p_T(X) \mid f_T(X)$; 另一方面, 记 $n = \dim_F(V)$, 则由引理知, $\exists B \in F[X]^{n\times n}$, s.t. $p_A(X)\cdot I_n = (X\cdot I_n - A)\cdot B$, 故两边取行列式知, $f_A(X) = \det(X\cdot I_n - A) \mid \det(p_A(X)\cdot I_n) = p_A(X)^n$. □

推论 6.1.8 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则 $p_T(X)$ 与 $f_T(X)$ 在 $F[X]$ 内有相同的不可约因子集, 即 $\{p(X) \in F[X]: p(X)$为不可约多项式, 且$\ker(p(T)) \neq \{0\}\}$. 特别地, 它们在 $\overline{F}^{\mathrm{alg}}$ 上有相同的根集.*
证明: 由 $p_T(X) \mid f_T(X) \mid p_T(X)^{\dim_F(V)}$ 知, $p_T(X)$ 与 $f_T(X)$ 在 $F[X]$ 内有相同的不可约因子集. 现设 $p(X) \in F[X]$ 为不可约多项式. 一方面, 若 $p(X) \nmid p_T(X)$, 则由 $p(X)$ 的不可约性知 $\gcd(p(X), p_T(X)) = 1$; 由 Bezout 定理知,
$\exists u(X), v(X) \in F[X]$, s.t. $u(X)\cdot p(X) + v(X)\cdot p_T(X) = 1$, 故 $u(T)\circ p(T) = \mathrm{id}_V$, 即 $\ker(p(T)) = \{0\}$. 另一方面, 若 $p(X) \mid p_T(X)$, 即 $\exists g(X) \in F[X]$, s.t. $p_T(X) = p(X)\cdot g(X)$, 则 $0 = p(T)\circ g(T)$. 假设 $\ker(p(T)) = \{0\}$, 则 $g(T) = 0$, 故 $p_T(X) \mid g(X)$, 这与 $\deg(p_T(X)) > \deg(g(X)) \geq 0$ 矛盾! □

6.1.2 中国剩余定理

为了将线性变换的最小多项式的不可约分解对应于线性空间的扭模分解, 我们引入关键的中国剩余定理:

定理 6.1.9 (中国剩余定理) *设 $(R, 0, +; 1, \cdot)$ 为一个交换环, $\mathfrak{a}_1, \cdots \mathfrak{a}_k \subseteq R$ 为两两互素的理想, 即 $\forall 1 \leq i < j \leq k$,*
$\mathfrak{a}_i + \mathfrak{a}_j = R$, 则存在 R-代数同构 $R/\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k \xrightarrow{\cong} R/\mathfrak{a}_1 \times \cdots \times R/\mathfrak{a}_k$.

$$r + \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k \longmapsto (r + \mathfrak{a}_1, \cdots, r + \mathfrak{a}_k)$$

证明: 显然上述映射为定义良好的 R-代数同态, 且为单射. 下证它也为满射, 只需证明 $\forall 1 \leq i \leq k$, $\exists r_i \in R$, s.t.

$\begin{cases} r_i + \mathfrak{a}_i = 1 + \mathfrak{a}_i \\ r_i + \mathfrak{a}_j = 0 + \mathfrak{a}_j, \quad j \neq i \end{cases}$. 事实上, 由于 $\forall 1 \leq i \neq j \leq k,\ \mathfrak{a}_i + \mathfrak{a}_j = R$, 则 $\exists s_{ij} \in \mathfrak{a}_i,\ s_j \in \mathfrak{a}_j,\ s.t.\ s_{ij} + s_j = 1$, 故

$r_i \coloneqq \prod\limits_{j \neq i} s_j = \prod\limits_{j \neq i} (1 - s_{ij})$ 满足条件. □

注: 上述证明中定义的 $\overline{r_i} \coloneqq r_i + \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k\ (1 \leq i \leq k)$ 是商代数 $R/\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_k$ 中两两正交的中心幂等元, 即它们满足 $\overline{r_i} \cdot \overline{r} = \overline{r} \cdot \overline{r_i},\ \forall r \in R;\ \overline{r_i}^2 = \overline{1};\ \overline{r_i} \cdot \overline{r_j} = 0, \forall i \neq j$. 并且, $\sum\limits_{i=1}^{k} \overline{r_i} = \overline{1}$.

推论 6.1.10 设 F 为一个域, $f(X) \in F[X]$, 记 $f(X) = \prod\limits_{i=1}^{k} p_i(X)^{r_i}$ 为不可约因子的幂次的乘积, 则存在 $F[X]$-代数同构 $F[X]/f(X) \cdot F[X] \xrightarrow{\cong} F[X]/p_1^{r_1}(X) \cdot F[X] \times \cdots \times F[X]/p_k^{r_k}(X) \cdot F[X]$. 进一步地, 可取

$$g(X) + f(X) \cdot F[X] \longmapsto (g(X) + p_1^{r_1}(X) \cdot F[X], \cdots, g(X) + p_k^{r_k}(X) \cdot F[X])$$

$e_i(X) \in F[X]$

$(1 \leq i \leq k)$ 满足 $\begin{cases} p_i^{r_i}(X) \mid e_i(X) - 1 \\ p_j^{r_j}(X) \mid e_i(X),\ j \neq i \end{cases}$, 则 $\overline{e_i(X)} \coloneqq e_i(X) + f(X) \cdot F[X]$ 满足 $\begin{cases} \overline{e_i(X)}^2 = \overline{e_i(X)} \\ \overline{e_i(X)} \cdot \overline{e_j(X)} = 0,\ \forall i \neq j \\ \sum\limits_{i=1}^{k} \overline{e_i(X)} = \overline{1} \end{cases}$.

注: 在上述推论中, 由 $\overline{e_i(X)} \in F[X]/f(X) \cdot F[X]$ 生成的理想为

$$\frac{\gcd(e_i(X), f(X)) \cdot F[X]}{f(X) \cdot F[X]} = \frac{\prod\limits_{\substack{j=1 \\ j \neq i}}^{k} p_j^{r_j}(X) \cdot F[X]}{f(X) \cdot F[X]} \xrightarrow{\cong} \frac{F[X]}{p_i^{r_i}(X) \cdot F[X]},$$

对应于右端乘积代数中的第 i 分量. 这是一个**可约** (reducible) 但**不可分解** (indecomposable) 的 $F[X]$-模. 事实上, 它的所有 $F[X]$-子模构成了一条链 $\dfrac{0 \cdot F[X]}{p_i^{r_i}(X) \cdot F[X]} \subsetneqq \dfrac{p_i^{r_i - 1}(X) \cdot F[X]}{p_i^{r_i}(X) \cdot F[X]} \subsetneqq \cdots \subsetneqq \dfrac{p_i(X) \cdot F[X]}{p_i^{r_i}(X) \cdot F[X]} \subsetneqq \dfrac{F[X]}{p_i^{r_i}(X) \cdot F[X]}$.

引理 6.1.11 (投影算子与空间分解) 设 V 为域 F 上的线性空间, $W_i \subseteq V\ (1 \leq i \leq k)$ 为线性子空间, 则 $V = \bigoplus\limits_{i=1}^{k} W_i \Longleftrightarrow \exists E_i \in L(V)\ (1 \leq i \leq k), s.t.$ (1) $E_i^2 = E_i$; (2) $E_i \circ E_j = 0,\ \forall i \neq j$; (3) $\sum\limits_{i=1}^{k} E_i = \mathrm{id}_V$; (4) $\mathrm{Im}(E_i) = W_i$.
进一步地, 设 $T \in L(V)$, 则上述 $W_i\ (1 \leq i \leq k)$ 都是 T-不变子空间 $\Longleftrightarrow \forall 1 \leq i \leq k,\ T \circ E_i = E_i \circ T$.

定理 6.1.12 (线性变换的准素分解) 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 记 $p_T(X) = \prod\limits_{i=1}^{k} p_i(X)^{r_i}$ 为不可约因子的幂次的乘积, 则 $V = \bigoplus\limits_{i=1}^{k} W_i$, 其中 $W_i = \ker(p_i^{r_i}(T)) \neq \{0\}$, 且 $p_{T|_{W_i}}(X) = p_i^{r_i}(X)$.

证明: 考虑 $F[X]$-模同态 $F[X]/p_T(X) \cdot F[X] \hookrightarrow L(V)$, 记上述推论中 $\overline{e_i(X)}$ 的像为 $E_i\ (1 \leq i \leq k)$, 则 $\begin{cases} E_i^2 = E_i \\ E_i \circ E_j = 0,\ \forall i \neq j \\ \sum\limits_{i=1}^{k} E_i = \mathrm{id}_V \end{cases}$. 记 $W_i = \mathrm{Im}(E_i)$, 则由上述引理可知 $V = \bigoplus\limits_{i=1}^{k} W_i$. 断言: $\forall 1 \leq i \leq k,\ W_i = \ker(p_i^{r_i}(T))$.
(这是因为, 一方面, 由 $\overline{p_i^{r_i}(X)} \mid \overline{e_i(X)} - \overline{1}$ 知, $\ker(p_i^{r_i}(T)) \subseteq \ker(E_i - \mathrm{id}_V) = \mathrm{Im}(E_i)$. 另一方面, 由 $p_j^{r_j}(X) \mid e_i(X),\ \forall j \neq i$ 知,
$p_T(X) = \prod\limits_{i=1}^{k} p_i(X)^{r_i} \mid p_i(X)^{r_i} \cdot e_i(X)$, 即 $\overline{0} = \overline{p_i(X)^{r_i}} \cdot \overline{e_i(X)}$, 故 $\mathrm{Im}(E_i) \subseteq \ker(p_i^{r_i}(T))$.) 最后,
由 $p_{T|_{W_i}}(X) \mid p_i^{r_i}(X)$ 且 $\prod\limits_{i=1}^{k} p_{T|_{W_i}}(X) = p_T(X) = \prod\limits_{i=1}^{k} p_i(X)^{r_i}$ 知, $p_{T|_{W_i}}(X) = p_i^{r_i}(X),\ \forall 1 \leq i \leq k$. □

注:

(1) 上述证明中定义的 $E_i \in L(V)\ (1 \leq i \leq k)$ 称为准素投影算子, 它们都是 T 的多项式.

(2) 事实上, 由上述证明知, 若将 $p_T(X)$ 换成任意与 $p_T(X)$ 在 $F[X]$ 内有相同不可约因子集的零化多项式, 则该准素分解仍成立, 故 $W_i = \ker(p_i^{e_i}(T)) = \{\alpha \in V : \exists e \in \mathbb{N}^*,\ s.t.\ p_i^e(T)(\alpha) = 0\},\ \forall e_i \geq r_i$. 另一方面, 由 $p_{T|_{W_i}}(X) = p_i^{r_i}(X)$ 可知, $W_i = \ker(p_i^{r_i}(T)) \supsetneqq \ker(p_i^{r_i - 1}(T)) \supsetneqq \cdots \supsetneqq \ker(p_i(T)) \supsetneqq \{0\}$.

(3) 警告: 这里的准素子模 $W_i\,(1\le i\le k)$ 未必是不可分解的, 它可能为若干准素子子模 $W_{ij}\,(1\le j\le l_i)$ 的直和, 其中 $p_{T|_{W_{ij}}}(X)\mid p_i^{r_i}(X)$, 且在某个 j 处取等号. 但这种更精细的分解无法由最小多项式的不可约分解得到.

(4) 上述线性变换的准素分解可推广至主理想整环上: 设 R 为一个 P.I.D., M 为一个 R-扭模. 记 M 的非零零化子 $a\in R$ 的唯一分解式为 $a=u\cdot p_1^{e_1}\cdots p_k^{e_k}$, 其中 $u\in R^\times$, $p_i\in R^*\setminus R^\times\,(1\le i\le k)$ 为不同的不可约元, 则 $M=\bigoplus\limits_{i=1}^{k}N_i$, 其中 $N_i\coloneqq\{x\in M\colon p_i^{e_i}x=0\}=\{x\in M\colon \exists e\in\mathbb{N}^*,\ s.t.\ p_i^{e}x=0\}\,(1\le i\le k)$.

推论 6.1.13 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 记 $f_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{d_i}$ 为不可约因子的幂次的乘积, 则 $V=\bigoplus\limits_{i=1}^{k}W_i$, 其中 $W_i=\ker(p_i^{d_i}(T))\neq\{0\}$, 且 $f_{T|_{W_i}}(X)=p_i^{d_i}(X)$. 特别地, $\dim_F(W_i)=d_i\cdot\deg(p_i(X))$.*

证明: 由于 $p_T(X)$ 与 $f_T(X)$ 在 $F[X]$ 内有相同的不可约因子, 记 $p_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{r_i}$ 为不可约因子的幂次的乘积, 其中 $0<r_i\le d_i$, 则由准素分解定理知, $V=\bigoplus\limits_{i=1}^{k}W_i$, 其中 $W_i=\ker(p_i^{r_i}(T))\neq\{0\}$, 且 $p_{T|_{W_i}}(X)=p_i^{r_i}(X)$. 由于 $p_{T|_{W_i}}(X)=p_i^{r_i}(X)$ 与 $f_{T|_{W_i}}(X)$ 在 $F[X]$ 内有相同的不可约因子, 且 $\prod\limits_{i=1}^{k}f_{T|_{W_i}}(X)=f_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{d_i}$, 则 $f_{T|_{W_i}}(X)=p_i^{d_i}(X),\ \forall 1\le i\le k$. □

进一步地, 线性变换的准素分解可由它的不变子空间继承:

命题 6.1.14 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的准素分解为 $V=\bigoplus\limits_{i=1}^{k}W_i$, $W\subseteq V$ 为 T-不变子空间, 则 $W=\bigoplus\limits_{i=1}^{k}(W\cap W_i)$.*

证明: 显然 $W\supseteq\bigoplus\limits_{i=1}^{k}(W\cap W_i)$; 另一方面, 记 $E_i\,(1\le i\le k)$ 为准素投影算子, 则由 $\mathrm{id}_V=\sum\limits_{i=1}^{k}E_i$, W 的 T-不变性以及 E_i 为 T 的多项式知, $W\subseteq\bigoplus\limits_{i=1}^{k}\mathrm{Im}(E_i\big|_W)\subseteq\bigoplus\limits_{i=1}^{k}(W\cap W_i)$. □

推论 6.1.15 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 记 $p_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{r_i}$ 为不可约因子的幂次的乘积, $f(X)=\prod\limits_{i=1}^{k}p_i(X)^{e_i}$, $g(X)=\prod\limits_{i=1}^{k}p_i(X)^{f_i}\in F[X]$, 则 $\ker(f(T))\subseteq\ker(g(T))\iff\forall 1\le i\le k,\ \min\{e_i,r_i\}\le\min\{f_i,r_i\}$.*

证明: 记 $T\in L(V)$ 的准素分解为 $V=\bigoplus\limits_{i=1}^{k}\ker(p_i^{r_i}(T))$. 由于 $\ker(f(T))\subseteq V$ 为 T-不变子空间, 则由上述命题知,
$\ker(f(T))=\bigoplus\limits_{i=1}^{k}(\ker(f(T))\cap\ker(p_i^{r_i}(T)))=\bigoplus\limits_{i=1}^{k}\ker(\gcd(f(T),p_i^{r_i}(T))=\bigoplus\limits_{i=1}^{k}\ker(p_i^{\min\{e_i,r_i\}}(T))$, 故由准素分解定理后的注知, $\ker(f(T))\subseteq\ker(g(T))\iff\forall 1\le i\le k,\ \ker(p_i^{\min\{e_i,r_i\}}(T))\subseteq\ker(p_i^{\min\{f_i,r_i\}}(T))\iff\forall 1\le i\le k,\ \min\{e_i,r_i\}\le\min\{f_i,r_i\}$. □

最后我们讨论与 T 可交换的线性变换的不变子空间的性质.

命题 6.1.16 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的准素分解为 $V=\bigoplus\limits_{i=1}^{k}W_i$, $S\in L(V)$ 满足 $T\circ S=S\circ T$, 则每个 $W_i\,(1\le i\le k)$ 都是 S 的不变子空间.*

证明: 记 $W_i=\ker(p_i^{r_i}(T))\,(1\le i\le k)$, 其中 $p_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{r_i}$ 为不可约分解. 由 $T\circ S=S\circ T$ 知, $p_i^{r_i}(T)\circ S=S\circ p_i^{r_i}(T)$, 则 $\ker(p_i^{r_i}(T))$ 是 S-不变子空间. □

注: 警告: 在上述命题中, 一般 S 的不变子空间无法继承 T 的准素分解, 简单的反例如 $S=\mathrm{id}_V$; 但 S 的核、像、准素子空间都是 T 的不变子空间, 则可以继承 T 的准素分解.

6.1.3 可对角化的 Zariski 稠密性

本节考虑形式上最简单的一种线性变换, 即可对角化的线性变换, 并由它们在全体线性变换中的 Zariski 稠密性再次证明 Cayley-Hamilton 定理.

命题 6.1.17 *设 V 为域 F 上的有限维线性变换, $T \in L(V)$, 记 $V_c \coloneqq \ker(T - c \cdot \mathrm{id}_V)\,(c \in \sigma(T))$, 则以下条件等价:*

(1) *T 可对角化;*

(2) *存在由 T 的特征向量组成的 V 的基;*

(3) *V 可分解为若干一维 T-不变子空间的直和;*

(4) $V = \bigoplus\limits_{c\in\sigma(T)} V_c$;

(5) $\dim_F(V) = \sum\limits_{c\in\sigma(T)} \dim_F(V_c)$;

(6) $f_T(X) = \prod\limits_{c\in\sigma(T)} (X-c)^{\dim_F(V_c)}$;

(7) *$f_T(X)$ 可分解为一次式的乘积, 且 $\forall c \in \sigma(T)$, $m(c, f_T(X)) = \dim_F(V_c)$;*

(8) $p_T(X) = \prod\limits_{c\in\sigma(T)} (X-c)$;

(9) *$p_T(X)$ 可分解为不同的一次式的乘积.*

注: 由上述命题知, $T \in L(V)$ 不可对角化的原因分为以下两类:

(1) $f_T(X)$ 不在 F 上分裂, 例如取 $F = \mathbb{R}$, T 的矩阵表示为 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ $(\theta \in \mathbb{R} \setminus \mathbb{Z}\cdot\pi)$.

(2) $\exists c \in \sigma(T)$, *s.t.* $m(c, f_T(X)) > \dim_F(V_c)$, 例如取 T 的矩阵表示为 $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ $(\lambda \in F^*)$.

通过考虑 $f_T(X) \in F[X]$ 在 F 上的分裂域 E/F (或甚至 $\overline{F}^{\mathrm{alg}}/F$), 我们总可使得 $f_{T\otimes_F \mathrm{id}_E}(X) = f_T(X)$ 在 E 上分裂, 但此时第二个不可对角化的原因仍存在. 对于 $c \in \sigma(T)$, 我们将 $m(c, f_T(X)) \in \mathbb{N}^*$ 称为它的**代数重数** (algebraic multiplicity); $\dim_F(V_c) \in \mathbb{N}^*$ 称为它的**几何重数** (geometric multiplicity). 显然 $m(c, f_T(X)) \geq \dim_F(V_c)$. 以下考虑其中取等号条件的本质含义.

命题 6.1.18 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则下述 "(1) ⇒ (2) ⇒ (3)" 成立:*

(1) *T 可对角化;*

(2) *T 为半单变换, 即任意 T-不变子空间均存在 T-不变的直和补空间;*

(3) *T 的任意特征值是半单的, 即它的代数重数等于几何重数.*

证明: "(1)⇒(2)": 设 T 可对角化, 则由上述命题知, $V = \bigoplus\limits_{c\in\sigma(T)} V_c$. 任取 $W \subseteq V$ 为 T-不变子空间, 断言: $W = \bigoplus\limits_{c\in\sigma(T)} (W \cap V_c)$. (这是因为: 显然 $W \supseteq \bigoplus\limits_{c\in\sigma(T)} (W \cap V_c)$; 现设 $\alpha \in W$, 由 $\alpha \in V = \bigoplus\limits_{c\in\sigma(T)} V_c$ 知, 可记 $\alpha = \sum\limits_{i=1}^{k} \alpha_i$, 其中 $\alpha_i \in V_{c_i}\,(1 \leq i \leq k)$, 且 $c_1, \cdots, c_k \in \sigma(T)$ 两两不同. 由于 $\forall 0 \leq j \leq k-1$, $T^j(\alpha) = \sum\limits_{i=1}^{k} T^j(\alpha_i) = \sum\limits_{i=1}^{k} c_i^j(\alpha_i)$, 其中 $(c_i^j)_{\substack{1\leq i\leq k \\ 0\leq j\leq k-1}}$ 为可逆的 Vandermonde 阵, 则 $(\alpha, T(\alpha), \cdots, T^{k-1}(\alpha)) \cdot (c_i^j)^{-1}_{\substack{1\leq i\leq k \\ 0\leq j\leq k-1}} = (\alpha_1, \cdots, \alpha_k)$, 故由 $T^j(\alpha) \in W\,(1 \leq j \leq k-1)$ 知, $\alpha_1, \cdots, \alpha_k \in W$, 因此 $\alpha \in \bigoplus\limits_{c\in\sigma(T)} (W \cap V_c)$, 即 $W \subseteq \bigoplus\limits_{c\in\sigma(T)} (W \cap V_c)$.) 于是任取 $W \cap V_c$ 在 V_c 中的直和补空间 U_c, 则 U_c 必为 T-不变子空间. 记 $U = \bigoplus\limits_{c\in\sigma(T)} U_c$, 则 $V = W \oplus U$, 且 U 为 T-不变子空间.

"(2)⇒(3)": 设 T 为半单变换, 任取 $c \in \sigma(T)$, 则 $V_c \subseteq V$ 为 T-不变子空间, 可取 $W \subseteq V$ 为 T-不变子空间, 满足 $V = V_c \oplus W$, 故 $f_T(X) = (X-c)^{\dim_F(V_c)} \cdot f_{T|_W}(X)$. 断言: $f_{T|_W}(c) \neq 0$. (这是因为, 假设 $f_{T|_W}(c) = 0$, 即 $c \in \sigma(T\big|_W)$, 则 $V_c \cap W \neq \{0\}$, 矛盾!) 因此 $m(c, f_T(X)) = \dim_F(V_c)$. □

注:

(1) 若 $F = \overline{F}^{\mathrm{alg}}$, 则 "(3)⇒(1)" 也成立, 故此时 "(1) ⇔ (2) ⇔ (3)";

(2) 若 $F \subsetneqq \overline{F}^{\mathrm{alg}}$, 则 "(2)⇒(1)","(3)⇒(2)" 均不成立, 反例如下: 若取 $F=\mathbb{R}$, T 的矩阵表示为 $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ ($\theta \in \mathbb{R}\setminus\mathbb{Z}\cdot\pi$), 则 T 满足 (2) 但不满足 (1); 若取 T 的矩阵表示为 $f(X)\in F[X]$ 的友矩阵, 其中 $f(X)$ 在 $F[X]$ 中的不可约因子的次数均 ≥ 2, 且存在重数 ≥ 2 的不可约因式, 则 T 满足 (3) 但不满足 (2).

(3) 事实上, (2) T 为半单变换 $\Longleftrightarrow$ (2') $p_T(X)$ 的不可约因子重数均为 1. 一个线性代数的证明可见教材 Hoffman, Kunze "Linear Algebra". 另外, 利用环论可给出一个更简洁的证明如下: 一方面, 注意 T 为半单变换 $\Longleftrightarrow F[T]\subseteq L(V)$ 为半单结合代数; 另一方面, 考虑 F-线性代数同构 $F[T]\cong F[X]/p_T(X)\cdot F[X]$ 以及中国剩余定理 $F[X]/p_T(X)\cdot F[X]\cong \bigoplus_{i=1}^{k} F[X]/p_i^{r_i}(X)\cdot F[X]$, 后者为若干不可分解的理想直和, 故 $F[T]$ 为半单结合代数 $\Longleftrightarrow \forall 1\leq i\leq k$, $F[X]/p_i^{r_i}(X)\cdot F[X]$ 为极小单理想 $\Longleftrightarrow \forall 1\leq i\leq k$, $r_i=1$.

注意对于可对角化的线性变换, Cayley-Hamilton 定理是显然的: 这是因为, 设 $T\in L(V)$ 可对角化, 则由 $V=\bigoplus_{c\in\sigma(T)}\ker(T-c\cdot\mathrm{id}_V)$ 以及 $f_T(X)=\prod_{c\in\sigma(T)}(X-c)^{\dim_F(V_c)}$ 知, $f_T(T)=\prod_{c\in\sigma(T)}(T-c\cdot\mathrm{id}_V)^{\dim_F(V_c)}=0$. 以下试图通过可对角化的线性变换在全体线性变换中的 Zariski 稠密性, 证明一般的 Cayley-Hamilton 定理.

引理 6.1.19 *设 F 为代数闭域, 则 $\{A\in F^{n\times n}\colon \forall c\in\sigma(A),\ m(c,f_A(X))=1\}\subseteq F^{n\times n}$ 在 Zariski 拓扑下是稠密的开集.*

证明: 注意到 $F^{n\times n}\setminus\{A\in F^{n\times n}\colon \forall c\in\sigma(A),\ m(c,f_A(X))=1\}=\{A\in F^{n\times n}\colon f_A(X)$在$F=\overline{F}^{\mathrm{alg}}$上有重根$\}$ $=\{A\in F^{n\times n}\colon \mathrm{disc}(f_A(X))=0\}$, 其中 $\mathrm{disc}(f(X))=(-1)^{\frac{n(n-1)}{2}}\cdot\mathrm{Res}(f_A(X),f_A'(X))$ 是关于 $A_{ij}\,(1\leq i,j\leq n)$ 的非零多元多项式, 则 $\{A\in F^{n\times n}\colon \mathrm{disc}(f_A(X))=0\}\subsetneqq F^{n\times n}$ 为 Zariski 闭的真子集, 故它的补集为非空的 Zariski 开集, 特别地也是 Zariski 稠密的. □

推论 6.1.20 *设 F 为一个域, $A\in F^{n\times n}$, 则 $f_A(A)=0$.*

证明: 由于 $A\in F^{n\times n}\subseteq(\overline{F}^{\mathrm{alg}})^{n\times n}$ 不改变 $f_A(A)$, 故可不妨设 F 为代数闭域. 注意到映射 $\begin{aligned}F^{n\times n}&\longrightarrow F^{n\times n}\\ A&\longmapsto f_A(A)\end{aligned}$ 的每个分量均为多元多项式, 故为 Zariski 拓扑下的连续映射; 又由上述引理知, 它在某个 Zariski 稠密集上取值为 0, 则恒为 0. □

习题 6.1 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $f(X),g(X)\in F[X]$, 则*

(1) $\ker(\gcd(f,g)(T))=\ker(f(T))\cap\ker(g(T))$;

(2) $\mathrm{Im}(\gcd(f,g)(T))=\mathrm{Im}(f(T))+\mathrm{Im}(g(T))$;

(3) $\ker(\mathrm{lcm}(f,g)(T))=\ker(f(T))+\ker(g(T))$;

(4) $\mathrm{Im}(\mathrm{lcm}(f,g)(T))=\mathrm{Im}(f(T))\cap\mathrm{Im}(g(T))$.

参考文献与补注 6.1

(1) 关于中国剩余定理的部分, 可以参考 N. Jacobson"Basic Algebra I".

(2) 关于一般主理想整环上扭模的准素分解定理部分, 可以参考 Dummit, Foote"Abstract Algebra".

§6.2 线性变换的循环分解

为了进一步将有限维线性空间上的线性变换写成尽可能简单的形式, 我们从最少生成元的角度考虑线性空间的循环子模分解, 这就是所谓线性变换的循环分解.

6.2.1 循环模及其子模

本节考虑线性空间作为 R-模只有一个生成元的情形, 并由此讨论它的子模的基本性质. 为方便起见, 对于 $T\in L(V)$ 以及 $\alpha\in V$, 记 $p_\alpha(X)\in F[X]$ 为 $M_\alpha\coloneqq\{f(X)\in F[X]\colon f(T)(\alpha)=0\}\,(\alpha\in V)$ 的首一生成元. 我们先回顾一些常用的循环向量的判别法:

命题 6.2.1 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $\alpha\in V$, 记 $R=F[X]$, 则以下条件等价:*

(1) $V = F[T]\cdot\alpha$;

(2) 存在 R-模同构 $\quad F[X]/p_T(X)\cdot F[X] \xrightarrow{\cong} V \quad ;$

$$g(X)+p_T(X)\cdot F[X] \longmapsto g(T)(\alpha)$$

(3) $\dim_F(F[T]\cdot\alpha) = \dim_F(V)$;

(4) $\deg(p_\alpha(X)) = \dim_F(V)$.

证明: "(1)⇔(2)⇔(3)": 显然 (2) 中映射是定义良好的 R-模单同态, 则它为 R-模同构 $\iff V = F[T]\cdot\alpha \iff \dim_F(V) = \dim_F(F[T]\cdot\alpha)$.

"(3)⇔(4)": 由 $\dim_F(F[T]\cdot\alpha) = \deg(p_\alpha(X))$ 即知; □

为了将循环模的理论应用于一般的循环分解, 我们需要对每个向量及其零化子做更精细的讨论, 同时也可以得到更多循环向量的判别法.

引理 6.2.2 设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $p(X)\in F[X]$, 则 $p(X)\mid p_T(X) \iff \exists\alpha\in V,\ s.t.\ p(X)=p_\alpha(X)$.

证明: "⇐": 由 $\forall\alpha\in V,\ p_\alpha(X)\mid p_T(X)$ 即知;

"⇒": 设 $p(X)\mid p_T(X)$, 则记 $p_T(X)=\prod\limits_{i=1}^{k}p_i(X)^{r_i}$, $p(X)=\prod\limits_{i=1}^{k}p_i(X)^{e_i}$ $(0\le e_i\le r_i)$ 为不可约分解. 对于 $1\le i\le k$, 若 $e_i\ge 1$, 则取 $\alpha_i\in\ker(p_i(T)^{e_i})\setminus\ker(p_i(T)^{e_i-1})$; 若 $e_i=0$, 则取 $\alpha_i=0$; 故总有 $p_{\alpha_i}(X)=p_i(X)^{e_i}$. 记 $\alpha=\sum\limits_{i=1}^{k}\alpha_i\in\bigoplus\limits_{i=1}^{k}\ker(p_i(T)^{e_i})$, 则可直接验证 $p_\alpha(X)=p(X)$. □

引理 6.2.3 设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $\alpha\in V$, $p(X)\in F[X]$ 满足 $p(X)\mid p_\alpha(X)$, 则以下条件等价:

(1) $\ker(p(T))\subseteq F[T]\cdot\alpha$;

(2) $\dim_F(\ker(p(T))) = \deg(p(X))$.

证明: 任取 $p(X)\mid p_\alpha(X)$, 记 $p_\alpha(X)=p(X)\cdot h(X)$, 其中 $h(X)\in F[X]$. 考虑映射 $\quad F[X]/p(X)\cdot F[X]\longrightarrow\ker(p(T)) \quad ,$

$$g(X)+p(X)\cdot F[X]\longmapsto g(T)h(T)(\alpha)$$

可直接验证这是定义良好的 R-模单同态.

"(1)⇒(2)": 任取 $\beta\in\ker(p(T))$, 由 $\ker(p(T))\subseteq F[T]\cdot\alpha$ 知, $\exists f(X)\in F[X],\ s.t.\ \beta=f(T)(\alpha)$, 则 $p(T)f(T)(\alpha)=0$, 故 $p_\alpha(X)\mid p(X)\cdot f(X)$; 而 $p_\alpha(X)=p(X)\cdot h(X)$, 则 $h(X)\mid f(X)$, 即 $\exists g(X)\in F[X],\ s.t.\ f(X)=g(X)\cdot h(X)$, 故 $\beta=g(T)h(T)(\alpha)$, 因此上述映射也为满射. 由比较维数知, $\deg(p(X)) = \dim_F(F[X]/p(X)\cdot F[X]) = \dim_F(\ker(p(T)))$. "(2)⇒(1)": 由 (2) 知 $\dim_F(F[X]/p(X)\cdot F[X]) = \deg(p(X)) = \dim_F(\ker(p(T)))$, 则上述映射为 R-模同构, 故 $\ker(p(T))\subseteq F[T]\cdot\alpha$. □

注: 上述引理是许多后续讨论的基础, 在于它将几何条件 (1) 与代数条件 (2) 联系起来.

引理 6.2.4 设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $\alpha\in V$, 则以下条件等价:

(1) 任取 $p(X)\mid p_T(X)$ 为不可约因子, $\ker(p(T))\subseteq F[T]\cdot\alpha$;

(2) $\forall\beta\in V\setminus\{0\}$, $F[T]\cdot\alpha\cap F[T]\cdot\beta\neq\{0\}$.

证明: "(1)⇒(2)": 任取 $\beta\in V\setminus\{0\}$, 则 $\deg(p_\beta(X))\ge 1$. 记 $p_\beta(X)=p(X)\cdot g(X)$, 其中 $p(X)\in F[X]$ 为不可约多项式. 由 $p(X)\mid p_\beta(X)\mid p_T(X)$ 以及 (1) 知, $\ker(p(T))\subseteq F[T]\cdot\alpha$, 故 $0\neq g(T)(\beta)\in\ker(p(T))\subseteq F[T]\cdot\alpha$, 因此 $F[T]\cdot\alpha\cap F[T]\cdot\beta\neq\{0\}$.

"(2)⇒(1)": 任取 $p(X)\mid p_T(X)$ 为不可约因子, 以及 $\beta\in\ker(p(T))\setminus\{0\}$, 则 $\deg(p_\beta(X))\ge 1$. 又 $p_\beta(X)\mid p(X)$, 则 $p_\beta(X)=p(X)$. 断言: $F[T]\cdot\beta$ 无非平凡的 T-不变子空间. (这是因为, 任取 $\gamma\in F[T]\cdot\beta\setminus\{0\}\subseteq\ker(p(T))\setminus\{0\}$, 由前知 $p_\gamma(X)=p(X)$, 则 $\dim_F(F[T]\cdot\gamma)=\deg(p_\gamma(X))=\deg(p_\beta(X))=\dim_F(F[T]\cdot\beta)$, 即 $F[T]\cdot\gamma=F[T]\cdot\beta$.) 由 (2) 知 $F[T]\cdot\alpha\cap F[T]\cdot\beta\neq\{0\}$, 则 $F[T]\cdot\beta\subseteq F[T]\cdot\alpha$, 故 $\ker(p(T))=\sum\limits_{\beta\in\ker(p(T))\setminus\{0\}}F[T]\cdot\beta\subseteq F[T]\cdot\alpha$. □

注: 上述引理的条件略显奇怪, 它们严格弱于 $V=F[T]\cdot\alpha$. 例如取 T 的矩阵表示为 $\begin{pmatrix}1&\lambda\\0&1\end{pmatrix}$ $(\lambda\in F^*)$, $\alpha\in V_1\setminus\{0\}$. 但将其稍微加强后即可得以下循环向量的判别法.

命题 6.2.5 设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $\alpha\in V$, 则以下条件等价:

(0) $V = F[T]\cdot\alpha$;
(1) $\forall p(X) \mid p_T(X)$, $\ker(p(T)) \subseteq F[T]\cdot\alpha$;
(2) $\forall \beta \in V\setminus\{0\}$, $\deg(p_\beta(X)) \le \deg(p_\alpha(X))$ 且 $F[T]\cdot\alpha \cap F[T]\cdot\beta \neq \{0\}$.

证明: "(0)⇒(1)","(0)⇒(2)": 显然;
"(1)⇒(0)": 取 $p(X) = p_T(X)$ 即可;
"(2)⇒(0)": 由 $\deg(p_\alpha(X)) = \max\limits_{\beta\in V\setminus\{0\}} \deg(p_\beta(X))$ 知, $p_\alpha(X) = p_T(X)$, 则由安师讲义循环分解定理中的引理知, $\forall L \in V/F[T]\cdot\alpha$, $\exists \beta \in L$, $s.t.\ p_\beta(X) = p_L(X)$. 断言: $F[T]\cdot\alpha \cap F[T]\cdot\beta = \{0\}$. (这是因为: 假设 $g_1(T)(\alpha) = g_2(T)(\beta)$, 其中 $g_1(X), g_2(X) \in F[X]$, 则投影到商空间 $V/F[T]\cdot\alpha$ 知, $g_2(T)(L) = 0$, 即 $p_L(X) \mid g_2(X)$, 故 $p_\beta(X) \mid g_2(X)$, 即 $g_2(T)(\beta) = 0$.) 因此由 (2) 知 $\beta = 0$, 则 $L = 0$, 即 $V = F[T]\cdot\alpha$. □

推论 6.2.6 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, $\alpha \in V$ 满足 $p_\alpha(X) = p_T(X)$, 则以下条件等价:*
(0) $V = F[T]\cdot\alpha$;
(1) $\forall p(X) \mid p_\alpha(X)$, $\ker(p(T)) \subseteq F[T]\cdot\alpha$;
(2) *任取* $p(X) \mid p_\alpha(X)$ *为不可约因子*, $\ker(p(T)) \subseteq F[T]\cdot\alpha$;
(3) $\forall \beta \in V\setminus\{0\}$, $F[T]\cdot\alpha \cap F[T]\cdot\beta \neq \{0\}$;
(4) $\forall p(X) \mid p_\alpha(X)$, $\dim_F(\ker(p(T))) = \deg(p(X))$;
(5) *任取* $p(X) \mid p_\alpha(X)$ *为不可约因子*, $\dim_F(\ker(p(T))) = \deg(p(X))$.

有时我们不关心具体的循环向量, 而更强调线性变换的整体性质, 则有以下循环模的判别法:

命题 6.2.7 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 记 $R = F[X]$, 则以下条件等价:*
(1) V *为循环 R-模*;
(2) $\deg(p_T(X)) = \dim_F(V)$;
(3) $p_T(X) = f_T(X)$;
(4) T *在某基下的矩阵表示为 $f_T(X)$ 的友矩阵*;
(5) $\forall p(X) \mid p_T(X)$, $\dim_F(\ker(p(T))) = \deg(p(X))$;
(6) *任取* $p(X) \mid p_T(X)$ *为不可约因子*, $\dim_F(\ker(p(T))) = \deg(p(X))$.

对于循环模的情形, 我们可以很方便地刻画它的子模; 同样也可通过子模的信息给出循环模的更多判别法:

引理 6.2.8 *设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 若 V 是 T-循环的, 则存在以下集合之间的保序一一对应:*
(1) $\{V$*的*T*-不变子空间*$\}$;
(2) $\{F[X]/p_T(X)\cdot F[X]$*的理想*$\}$;
(3) $\{F[X]$*中包含*$p_T(X)$*的理想*$\}$;
(4) $\{p_T(X)$*的首一因子*$\}$.

证明: "(1)↔(2)": 由 V 为 T-循环模知, 存在 R-模同构 $V = F[T]\cdot\alpha \longrightarrow F[X]/p_T(X)\cdot F[X]$, 故 V 的 T-不变子空间保序地一一对应于 $F[X]/p_T(X)\cdot F[X]$ 的理想.

$$g(T)(\alpha) \longmapsto g(X) + p_T(X)\cdot F[X]$$

"(2)↔(3)": 由环同态第一定理即知;
"(3)↔(4)": 由 $F[X]$ 为 P.I.D. 即知, □

注: 我们可以显式给出 "(4)↔(1)": 考虑对应 $\{p_T(X)$的首一因子$\} \leftrightarrow \{V$的T-不变子空间$\}$, 先证:

$$p(X) \longmapsto \ker(p(T))$$
$$p_{T|_W}(X) \longleftarrow\!\shortmid\ W$$

$p_{T|_{\ker(p(T))}}(X) = p(X)$; $\ker(p_{T|_W}(T)) = W$. 由上述命题, 一方面, 由 R-模同构 $F[X]/p(X)\cdot F[X] \cong \ker(p(T))$ 知, $p(X) = p_{T|_{\ker(p(T))}}(X)$; 另一方面, 由 $\ker(p_{T|_W}(T)) \supseteq W$ 且 $\dim_F(\ker(p_{T|_W}(T))) = \deg(p_{T|_W}(X)) \le \dim_F(W)$ 知, $\ker(p_{T|_W}(X)) = W$. 最后, 由准素分解的推论知, $p_1(X) \mid p_2(X) \mid p_T(X) \iff \ker(p_1(T)) \subseteq \ker(p_2(T))$.

推论 6.2.9 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则以下条件等价:*

(1) *V 是 T-循环的;*

(2) *$\{V$中T-不变子空间$\}=\{\ker(p(T)):p(X)\mid p_T(X)\}$;*

(3) *$\{V$中不可约的T-不变子空间$\}=\{\ker(p(T)):p(X)\mid p_T(X)$为不可约因子$\}$;*

(4) *$\{V$中不可分解的T-不变子空间$\}=\{\ker(p(T)):p(X)\mid p_T(X)$为不可约因子的幂$\}$.*

证明: "(1)⇒(2)⇒(3)": 由上述引理即知;

"(3)⇒(1)": 任取 $p(X)\mid p_T(X)$ 为不可约因子, 则由 (3) 知 $\ker(p(T))\subseteq V$ 为不可约的 T-不变子空间, 故 $\forall\alpha\in\ker(p(T))\setminus\{0\}$, $F[T]\cdot\alpha=\ker(p(T))$, 因此 $\deg(p_\alpha(X))=\dim_F(F[T]\cdot\alpha)=\dim_F(\ker(p(T)))$. 又由 $p_\alpha(X)\mid p(X)$ 且 $p(X)$ 不可约知, $p_\alpha(X)=p(X)$, 则 $\deg(p(X))=\dim_F(\ker(p(T)))$, 故由前述命题知 V 是 T-循环的.

"(2)⇔(4)": 由上述引理以及准素分解即知. □

推论 6.2.10 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则 V 是 T-循环的 $\iff$ V 中 T-不变子空间都是 T-循环的.*

推论 6.2.11 *设 F 为无限域, V 为域 F 上的有限维线性空间, $T\in L(V)$, 记 $R=F[X]$, 则以下条件等价:*

(1) *V 为循环 R-模;*

(2) *$|\{V$中T-不变子空间$\}|<+\infty$;*

(2) *$|\{V$中T-不变子空间$\}|<|F|$.*

证明: "(1)⇒(2)": 由上述推论即知;

"(2)⇒(3)": 由 F 为无限域即知;

"(3)⇒(1)": 任取 $\alpha_1,\alpha_2\in V$, 考虑 V 的子空间族 $\mathscr{F}:=\{F[T]\cdot(\alpha_1+c\alpha_2)\subseteq V:c\in F\}$, 由 $|\{V$中T-不变子空间$\}|<|F|$ 知, 映射 $F\longrightarrow\mathscr{F}$, $c\longmapsto F[T]\cdot(\alpha_1+c\alpha_2)$ 不为单射, 即 $\exists c_1\neq c_2\in F$, *s.t.* $F[T]\cdot(\alpha_1+c_1\alpha_2)=F[T]\cdot(\alpha_1+c_2\alpha_2)=:W$, 故 $(c_1-c_2)\alpha_2=(\alpha_1+c_1\alpha_2)-(\alpha_1+c_2\alpha_2)\in W$, 即 $\alpha_2\in W$, 因此 $\alpha_1=(\alpha_1+c_1\alpha_2)-c\alpha_2\in W$. 综上知 $\exists c_2\in F$, *s.t.* $F[T]\cdot\alpha_1+F[T]\cdot\alpha_2=F[T]\cdot(\alpha_1+c_2\alpha_2)$. 继续归纳可知 $\forall\alpha_1,\cdots,\alpha_n\in V$, $\exists c_2,\cdots,c_n\in F$, *s.t.* $F[T]\cdot\alpha_1+\cdots+F[T]\cdot\alpha_n=F[T]\cdot(\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n)$. 特别地, 取 $\{\alpha_i\}_{i=1}^n$ 为 V 的基, 则 $V=\bigoplus\limits_{i=1}^n F\alpha_i=\sum\limits_{i=1}^n F[T]\cdot\alpha_i=F[T]\cdot(\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n)$, 即 V 是循环 R-模. □

更多地, T 的循环性还可通过计算 T 的中心化子得到:

命题 6.2.12 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 记 $C(T):=\{S\in L(V):T\circ S=S\circ T\}$, 则 $\dim_F(C(T))\geq\dim_F(V)$, 且以下条件等价:*

(1) *V 为 T-循环的;*

(2) *$C(T)=F[T]$;*

(3) *$\dim_F(C(T))=\dim_F(V)$;*

(4) *任取 $W\subseteq V$ 为 T-不变子空间, 以及 $S\in C(T)$, 则 W 也是 S-不变的.*

证明: 对 $\dim_F(V)\in\mathbb{N}$ 归纳证明: 当 $\dim_F(V)=0,1$ 时结论显然; 当 $\dim_F(V)\geq 2$ 时, 若 V 可分解为两个非平凡的 T-不变子空间的直和, 则由归纳假设即知结论; 现设 V 为不可分解的, 则由循环分解的存在性知, V 为循环 R-模, 即 $\deg(p_T(X))=\dim_F(V)$, 故由 $C(T)\supseteq F[T]$ 知, $\dim_F(C(T))\geq\dim_F(F[T])=\deg(p_T(X))=\dim_F(V)$.

现证等号取到的等价条件:

"(1)⇒(2)": 显然 $C(T)\supseteq F[T]$; 另一方面, 任取 $S\in C(T)$, 设 $V=F[T]\cdot\alpha$, 则 $\exists f(X)\in F[X]$, *s.t.* $S(\alpha)=f(T)(\alpha)$; 再由 $V=F[T]\cdot\alpha$ 以及 $F[T]$ 的交换性即知 $S=f(T)\in F[T]$, 故 $C(T)\subseteq F[T]$.

"(2)⇒(3)": 由 $\dim_F(C(T))\geq\dim_F(V)\geq\deg(p_T(X))=\dim_F(F[T])$ 与 $C(T)=F[T]$ 知, $\dim_F(C(T))=\dim_F(V)$.

"(3)⇒(1)": 考虑 V 的循环分解 $V=\bigoplus\limits_{i=1}^r F[T]\cdot\alpha_i$, 其中 $p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)$, 则 $\dim_F(V)=\sum\limits_{i=1}^r\deg(p_{\alpha_i}(X))$. 由 $C(T)\supseteq\bigoplus\limits_{i=1}^r C\left(T\big|_{F[T]\cdot\alpha_i}\right)$ 知, $\dim_F(C(T))\geq\sum\limits_{i=1}^r\dim_F\left(C\left(T\big|_{F[T]\cdot\alpha_i}\right)\right)\geq\sum\limits_{i=1}^r\dim_F(F[T]\cdot\alpha_i)=\sum\limits_{i=1}^r\deg(p_{\alpha_i}(X))$. 因此由 (3) 知上式均取等号, 则 $C(T)=\bigoplus\limits_{i=1}^r C\left(T\big|_{F[T]\cdot\alpha_i}\right)$. 另一方面, 假设 $r\geq 2$, 则 S: $V\longrightarrow V$, $\sum\limits_{i=1}^r g_i(T)(\alpha_i)\longmapsto g_1(T)(\alpha_2)$

是定义良好的线性变换, 但 $S \in C(T) \setminus \bigoplus_{i=1}^{r} C\left(T\big|_{F[T]\cdot\alpha_i}\right)$, 矛盾!

"(2)⇒(4)": 显然;

"(4)⇒(1)": 考虑 V 的循环分解 $V = \bigoplus_{i=1}^{r} F[T]\cdot\alpha_i$, 其中 $p_{\alpha_r}(X) \mid \cdots \mid p_{\alpha_1}(X)$. 假设 $r \geq 2$, 取 $W = F[T]\cdot\alpha_1 \subseteq V$ 为 T-不变子空间, S: $V \longrightarrow V$ 是定义良好的线性变换, 且 $S \in C(T)$, 但 $S(\alpha_1) = \alpha_2 \notin W$,

$$\sum_{i=1}^{r} g_i(T)(\alpha_i) \longmapsto g_1(T)(\alpha_2)$$

这与 (4) 矛盾! □

注: 下一小节利用循环分解, 我们将计算一般线性变换的中心化子.

6.2.2 有限生成扭模的表现

为了从最少生成元的角度考虑线性空间的循环子模分解, 我们引入主理想整环上有限生成扭模的表现:

定义 6.2.1 (有限生成模) 设 $(R, 0, +; 1, \cdot)$ 是一个环, $(M, +, 0)$ 是一个 R-模, $S \subseteq M$ 为子集, 若 $M = \sum_{\alpha\in S} R\cdot\alpha$, 则称 M 在 R 上可由 S 生成, S 称为 M 的一组**生成元** (generating elements). 特别地,

(1) 若 M 在 R 上可由某个有限集生成, 则称 M 为 R-**有限生成模** (finitely-generated module);

(2) 若 M 在 R 上可由某个一元集生成, 则称 M 为 R-**循环模** (cyclic module).

注:

(1) 设 $(M, +, 0)$ 是一个 R-模, $S \subseteq M$ 为子集, 则 M 可在 R 上由 S 生成 $\Longleftrightarrow$ 存在 R-模满同态 $R^{(S)} \coloneqq \bigoplus_{\alpha\in S} R \twoheadrightarrow M$; 该 R-模满同态的核 $\ker(R^{(S)} \twoheadrightarrow M)$ 作为 R-模的一组生成元称为 M 的一组**生成关系** (generating relations); 生成元和生成关系组成了 M 的一个**表现** (presentation).

(2) 一般地, 有限生成模的子模未必是有限生成模, 甚至循环模的子模也未必是有限生成模, 反例如取 $R = F[X_i]_{i\in\mathbb{N}}$, R 是 R-循环模, 但 $M = \bigoplus_{i\in\mathbb{N}} R\cdot X_i \subseteq R$ 不是 R-有限生成模. 另外由定义知, 有限生成模的商模是有限生成模, 循环模的商模也是循环模.

定义 6.2.2 (自由模与扭模) 设 $(R, 0, +; 1, \cdot)$ 是一个环, $(M, +, 0)$ 是一个 R-模,

(1) 若存在子集 $S \subseteq M$, 满足 M 在 R 上可由 S 生成, 且 S 是 R-线性无关的, 则称 M 是一个 R-**自由模** (free module).

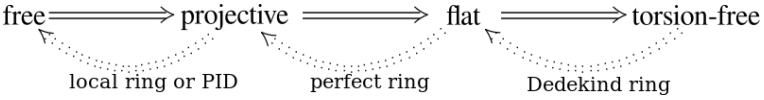
(2) 设 $\alpha \in M$, 若存在非零除子 $r \in R$, 满足 $r\cdot\alpha = 0$, 则称 α 为一个**扭元** (torsion element). 若 M 中元均为扭元, 则称 M 为一个 R-**扭模** (torsion module); 若 M 中元无扭元, 则称 M 为一个 R-**无扭模** (torsion-free module).

注:

(1) 一般地, 记 $tM \coloneqq \sum_{\alpha\in M\text{为扭元}} R\cdot\alpha \subseteq M$, 则 tM 未必是 R-模 (特别地, 当 R 为交换环时, tM 是一个 R-模).

(2) 上述关于扭元的定义值得特别注意. 在部分教材中, 扭元 $\alpha \in M$ 的定义语句为 "存在 $r \in R^*$ 满足 $r\cdot\alpha = 0$". 这两种定义当 R 为整环时是相同的; 但当 R 为一般的交换环时, 按我们的定义 M/tM 总是 R-无扭模, 按部分教材的定义 M/tM 未必是 R-无扭模, 这就容易引起混乱.

(3) 由定义知, 任意 R-自由模必为无扭模; 反之则需要更多的条件, 例如:



关于环上自由模与扭模的两个非平凡而实用的结果如下:

引理 6.2.13 *设 $(R, +, 0; \cdot, 1)$ 为一个非零交换环, 则以下条件等价:*

(1) *任意 R-自由模的子模均为 R-自由模;*

(2) *任意 R-有限生成自由模的子模均为 R-自由模;*

(3) *任意 R 中理想均为自由模;*

(4) *R 为主理想环, 即任意 R 中理想均为主理想.*

证明: "(1)⇒(2)⇒(3)": 显然;

"(3)⇒(4)": 任取 R 中理想 $\mathfrak{a}$. 若 $\mathfrak{a}=\{0\}$, 则 $\mathfrak{a}$ 显然为主理想. 若 $\mathfrak{a}\neq\{0\}$, 由 $\mathfrak{a}$ 为 R-自由模知, $r(\mathfrak{a})\geq 1$; 另一方面, 任取 $a\neq b\in\mathfrak{a}$, 由 $b\cdot a+(-a)\cdot b=0$ 知, $\{a,b\}$ 是 R-线性相关的, 故 $r(\mathfrak{a})\leq 1$, 因此 $r(\mathfrak{a})=1$, 即 $\mathfrak{a}=R\cdot a$ 为主理想.

"(4)⇒(1)": 任取 R-自由模 M, 记 M 的一组 R-基为 $\{\alpha_i\}_{i\in I}$, 并取 I 上的一个良序 $\preccurlyeq$, 则 $M_i\coloneqq\bigoplus\limits_{j\preccurlyeq i}R\cdot\alpha_j$, $M_i'\coloneqq\bigoplus\limits_{j\prec i}R\cdot\alpha_j$ 均为 M 的 R-子模, 且 $M_i=M_i'\oplus R\cdot\alpha_i$. 现任取 M 的 R-子模 N, 记 $N_i\coloneqq N\cap M_i$, $N_i'\coloneqq N\cap M_i'=N_i\cap M_i'$, 则 $N_i/N_i'=N_i/N_i\cap M_i'\cong N_i+M_i'/M_i'\subseteq M_i/M_i'\cong R\cdot\alpha_i\cong R$, 故 N_i/N_i' 作为 R-模同构于 R 的一个理想. 由于 R 为主理想环, 故 $N_i/N_i'=\{0\}$ 或 $N_i/N_i'\cong R$, 即 $\exists\beta_i\in N_i$, *s.t.* $N_i=N_i'\oplus R\cdot\beta_i$, 其中 $\beta_i=0$ 或 $R\cdot\beta_i\cong R$. 以下可超限归纳证明: $\{\beta_i\}_{i\in I}\setminus\{0\}$ 是 N 的一组 R-基, 详细可见 J. J. Rotman"Advanced Modern Algebra". □

引理 6.2.14 *设 $(R,+,0;\cdot,1)$ 为一个整环, 则以下条件等价:*

(0) *任意 R-有限生成无扭模都是自由模;*

(1) *任意 R-自由模的有限生成子模都是自由模;*

(2) *任意 R-有限生成自由模的有限生成子模都是自由模;*

(3) *任意 R 中有限生成理想都是自由模;*

(4) *R 是 Bézout 环, 即任意 R 中有限生成理想都是主理想.*

证明: "(0)⇒(1)⇒(2)⇒(3)": 显然;

"(3)⇒(4)⇒(1)": 由上述引理 "(3)⇒(4)⇒(1)" 的证明即知;

"(1)⇒(0)": 任取 R-有限生成无扭模 M, 记 M 的一组 R-生成元为 $\{x_i\}_{i=1}^m$, 不妨设 $\{x_i\}_{i=1}^r$ 为其中的极大 R-线性无关组. 记 $N\coloneqq\bigoplus\limits_{i=1}^{r}R\cdot x_i\subseteq M$, 则 N 为 R-有限生成自由模. 若 $M=N$, 则 M 为 R-自由模, 结论成立; 若 $M\neq N$, 则考虑 R-线性关系 $\sum\limits_{i=1}^{r}a_{ji}\cdot x_i+b_j\cdot x_j=0\,(r+1\leq j\leq m)$, 其中 $a_{ji}\in R$, $b_j\in R^*$. 由 R 的无零因子性知, $b\coloneqq\prod\limits_{j=r+1}^{m}r_j\in R^*$. 由于 $\forall r+1\leq j\leq m$, $b\cdot x_j=-\prod\limits_{\substack{k=r+1\\k\neq j}}^{m}b_k\cdot\sum\limits_{i=1}^{r}a_{ji}\cdot x_i\in N$, 则 $b\cdot M\subseteq N$. 由 (1) 知 $b\cdot M$ 为 R-自由模. 又由 M 无扭且 b 非零因子知, $M\cong b\cdot M$ 也为 R-自由模. □

注:

(1) 上述引理的证明仅在 "(1)⇒(0)" 部分用到了 R 的无零因子性.

(2) 在上述引理的 "(1)⇒(0)" 部分中, 有限生成性是必要的, 反例如 $R=\mathbb{Z}$ 为 P.I.D., 但 $\mathbb{Q}$ 作为 $\mathbb{Z}$-模无扭而不自由. 事实上, 在前述条件链中, $\mathbb{Q}$ 作为 $\mathbb{Z}$-模平坦而不投射.

推论 6.2.15 (主理想整环上有限生成模的结构) *设 $(R,+,0;\cdot,1)$ 是一个主理想整环, M 是一个 R-有限生成模, 则作为 R-模, M 同构于一个 R-扭模与一个 R-自由模的直和.*

证明: 考虑 R-模正合列 $\{0\}\to tM\to M\to M/tM\to\{0\}$, 其中 tM 为 R-扭模, M/tM 为 R-有限生成无扭模. 由上述引理知, M/tM 是自由模. 再通过 R-基的提升知, 该 R-模正合列可裂, 故 $M\cong tM\oplus M/tM$. □

命题 6.2.16 (环的直和作为自由模) *设 $(R,0,+;1,\cdot)$ 是一个环, $(M,+,0)$ 是一个 R-模, 则 M 是一个自由 R-模 $\Longleftrightarrow$ 存在 R-模同构 $R^{(S)}\cong M$. 特别地, 当 R 为非零交换环时, $|S|$ 由 M 唯一决定, 称为 M 的秩 (rank).*

证明: 由定义知该等价刻画显然, 其中 S 可取为 M 的一组 R-基. 现设 R 为一个非零交换环, 则由 Krull 定理可取它的极大理想 $\mathfrak{m}$, 即 $k=R/\mathfrak{m}$ 是一个域. 若存在 R-模同构 $R^{(S_1)}\cong M\cong R^{(S_2)}$, 则将上式两端作用 $-\underset{R}{\otimes}k$ 知, 存在 k-线性空间同构 $k^{(S_1)}\cong k^{(S_2)}$, 故由线性空间的维数良定义知, $|S_1|=|S_2|$. □

注:

(1) 一般地, 若对于任意的 $m,n\in\mathbb{N}^*$, 由 R-模同构 $R^m\cong R^n$ 总可推出 $m=n$, 则称环 R 具有 **IBN 性质** (invariant basis number property). 这是一个非平凡的性质, 反例如

$$S=R^{(\mathbb{N})\times\mathbb{N}}\coloneqq\{A\in R^{\mathbb{N}\times\mathbb{N}}\colon\forall j\in\mathbb{N},\ |\{i\in\mathbb{N}\colon A_{ij}\neq 0\}|<+\infty\},$$

则对于任意的 $n\in\mathbb{N}^*$, 存在 S-模同构 $S\longrightarrow \underbrace{S\times\cdots\times S}_{n}$, 其中 A_i 为 A 的诸 $\{i+kn\}_{k\in\mathbb{N}}$ 列 $(0\le i\le n-1)$.

$$A\longmapsto (A_0,\cdots,A_{n-1})$$

(2) 上述证明依赖于与 Zorn 引理等价的 Krull 定理. 事实上, 利用 Cayley-Hamilton 定理也可证明任意非零交换环具有 IBN 性质, 见以下的引理.

引理 6.2.17 *设 $(R,0,+;1,\cdot)$ 是一个非零交换环, $T\colon R^m\to R^n$ 为单的 R-模同态, 则 $m\le n$.*

证明: 假设 $m>n$, 考虑 R-模单同态 $i\colon R^n\longrightarrow R^m$, 则 $i\circ T\colon R^m\to R^m$ 也为 R-模单同态, 记它的矩阵表示

$$\alpha\longmapsto(\alpha,0)$$

为 $A\in R^{m\times m}$. 由 $(X\cdot I_m-A)\cdot\mathrm{adj}(X\cdot I_m-A)=\det(X\cdot I_m-A)\cdot I_m$ 知, 记 $f_A(X)\coloneqq\det(X\cdot I_m-A)\in F[X]^*$, 则 $f_A(A)=0$, 故 $M_A\coloneqq\{f(X)\in R[X]\colon f(A)=0\}$ 为 $R[X]$ 中的非零理想. 特别地, 取 $p(X)\in M_A$ 满足 $\deg(p(X))=\min\{\deg(f(X))\in\mathbb{N}\colon f(X)\in M_A\setminus\{0\}\}$, 则 $p(X)$ 的常数项非零, 记为 $p(X)=\sum\limits_{j=0}^{d}a_jX^j$ $(a_0\in R^*)$. 注意 $p(A)(0,\cdots,0,1)=p(i\circ T)(0,\cdots,0,1)=\sum\limits_{j=1}^{d}(i\circ T)^j(0,\cdots,0,1)+a_0\cdot(0,\cdots,0,1)=(*,\underbrace{0\cdots,0,a_0}_{m-n})$, 但 $p(A)=0$, 矛盾! □

命题 6.2.18 (线性空间作为有限表现的扭模) *设 V 为域 F 上的有限维线性空间, 基为 $B=\{\alpha_1,\cdots,\alpha_n\}$, $T\in L(V)$ 在基 B 下的矩阵表示为 $A\in F^{n\times n}$. 记 $R=F[X]$, 考虑 R-模同态 π:*

$$R^n\longrightarrow V,$$

则:

$$\sum_{i=1}^{n}g_i(X)\cdot\epsilon_i\longmapsto\sum_{i=1}^{n}g_i(T)(\alpha_i)$$

(1) *$f_T(T)=0$, 故 V 为 R-扭模;*

(2) *π 为满射, 故 V 在 R 上是有限生成的;*

(3) *$K\coloneqq\ker(\pi)\subseteq R^n$ 为 R-自由模, 且 $r(K)\le n$, 故 V 在 R 上是有限表现的;*

(4) *$(X\cdot I_n-A)\in R^{n\times n}$ 的列向量组是 K 的 R-基, 故 $r(K)=n$.*

证明: (1) 是 Cayley-Hamilton 定理; (2) 显然; (3) 由上述引理即知; 以下证明 (4): 注意 $(X\cdot I_n-A)$ 的第 j 个列向量为 $y_j(X)\coloneqq\begin{pmatrix}-A_{1j}\\ \vdots\\ X-A_{jj}\\ \vdots\\ -A_{nj}\end{pmatrix}=X\cdot\epsilon_j-\sum\limits_{i=1}^{n}A_{ij}\epsilon_i\in R^n$, 满足 $\pi(y_j(X))=T(\alpha_j)-\sum\limits_{i=1}^{n}A_{ij}\alpha_i=0$, $\forall 1\le j\le n$, 即 $\{y_1(X),\cdots,y_n(X)\}\subseteq K$. 欲证 $\{y_1(X),\cdots,y_n(X)\}$ 是 K 的 R-基, 即证它们 R-线性无关, 且在 R 上生成 K.

一方面, 假设 $\exists h_1(X),\cdots,h_n(X)\in R$ (不全为 0), *s.t.* $\sum\limits_{j=1}^{n}h_j(X)\cdot y_j(X)=0$, 则可取 $j_0\in\{1,\cdots,n\}$, *s.t.* $\deg(h_{j_0}(X))=\max\limits_{1\le j\le n}\deg(h_j(X))\ge 0$. 考虑 $\sum\limits_{j=1}^{n}h_j(X)\cdot\begin{pmatrix}-A_{1j}\\ \vdots\\ X-A_{jj}\\ \vdots\\ -A_{nj}\end{pmatrix}=0$ 的第 j_0 分量知, $h_{j_0}(X)\cdot X=\sum\limits_{j=1}^{n}h_j(X)\cdot A_{j_0,j}$, 两边比较次数即知矛盾! 因此 $\{y_1(X),\cdots,y_n(X)\}$ 是 R-线性无关的.

另一方面, 记 $K'\coloneqq\bigoplus\limits_{j=1}^{n}R\cdot y_j(X)\subseteq K$ 为 R-子模, $M\coloneqq K'+F^n\subseteq R^n$ 为 F-线性空间. 又由等式 $X\cdot\epsilon_j=y_j(X)+\sum\limits_{i=1}^{n}A_{ij}\epsilon_i$ 可知, M 也为 R-模, 则 $M\supseteq\bigoplus\limits_{j=1}^{n}R\cdot\epsilon_j=R^n$, 故 $M=R^n$. 现任取 $\beta(X)\in K$, 由 $K\subseteq R^n=M$ 知, $\exists h_1(X),\cdots,h_n(X)\in R$, $c_1,\cdots,c_n\in F$, *s.t.* $\beta(X)=\sum\limits_{j=1}^{n}h_j(X)\cdot y_j(X)+\sum\limits_{j=1}^{n}c_j\epsilon_j$, 则 $0=\pi(\beta(X))=\pi\left(\sum\limits_{j=1}^{n}h_j(X)\cdot y_j(X)\right)+\pi\left(\sum\limits_{j=1}^{n}c_j\epsilon_j\right)=\sum\limits_{j=1}^{n}c_j\alpha_j$. 而 $\{\alpha_1,\cdots,\alpha_n\}$ 线性无关, 故 $c_1=\cdots=c_n=0$, 即 $\beta(X)=\sum\limits_{j=1}^{n}h_j(X)\cdot y_j(X)\in K'$, 因此 $K\subseteq K'$, 即 $\{y_1(X),\cdots,y_n(X)\}$ 在 R 上生成 K. □

注: 在上述证明中, 由 $\mathrm{column}_R(X\cdot I_n-A)\subseteq K$ 均为 R-自由模且 $r(K)\leq n\leq r(\mathrm{column}_R(X\cdot I_n-A))$ 无法直接推出 $\mathrm{column}_R(X\cdot I_n-A)=K$, 反例如 $(2\mathbb{Z})^n\subsetneqq\mathbb{Z}^n$ 均为秩为 n 的 Abel 群.

推论 6.2.19 *设 V 为域 F 上的有限维线性空间, 基为 $B=\{\alpha_1,\cdots,\alpha_n\}$, $T\in L(V)$ 在基 B 下的矩阵表示为 $A\in F^{n\times n}$. 记 $R=F[X]$, $g_1(X),\cdots,g_n(X)\in R$, 则 $\sum\limits_{i=1}^{n}g_i(T)(\alpha_i)=0\Longleftrightarrow\exists\beta\in R^n$, s.t. $(X\cdot I_n-A)\cdot\beta=\begin{pmatrix}g_1(X)\\ \vdots\\ g_n(X)\end{pmatrix}$. 特别地, 记 $\{g_{ij}(X)\}_{1\leq i,j\leq n}\subseteq R$, 则 $\sum\limits_{i=1}^{n}g_{ij}(T)(\alpha_i)=0$, $\forall 1\leq j\leq n\Longleftrightarrow\exists B\in R^{n\times n}$, s.t. $(X\cdot I_n-A)\cdot B=(g_{ij}($*

注: 此推论是 Cayley-Hamilton 定理前引理的一般形式.

综上, 上述命题中的 R-模同态 π 诱导了 R-模同构 $\overline{\pi}\colon R^n/K\xrightarrow{\cong}V$, 其中 $K=\mathrm{column}_R(X\cdot I_n-A)$, 且 $r(K)=n$. 此式称为 V 作为有限生成 R-扭模的**表现** (presentation): $\{\epsilon_i\}_{i=1}^n\subseteq R^n$ 为生成元, $(X\cdot I_n-A)$ 的列向量组为生成关系. 为了进一步研究 V 的分解性质, 我们考虑对 $(X\cdot I_n-A)$ 做相抵变换 $P(X)\cdot(X\cdot I_n-A)\cdot Q(X)$, 其中 $P(X),Q(X)\in\mathrm{GL}(n,R)$, 则存在 R-模同构

$$R^n/\mathrm{column}_R(X\cdot I_n-A)\xrightarrow{\cong}R^n/\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A))=\!=R^n/\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A)\cdot Q(X))\,,$$

$$\sum_{i=1}^{n}g_i(X)\cdot\overline{\epsilon_i}\longmapsto\sum_{i,j=1}^{n}P_{ij}(X)g_i(X)\cdot\overline{\epsilon_i}\qquad\text{(右乘可逆阵不改变列空间)}$$

故可由 $(X\cdot I_n-A)$ 的相抵标准形给出 R^n/K 在 R-模同构意义下的简化形式.

引理 6.2.20 (Smith 标准形) *设 $(R,+,0;1,\cdot)$ 为一个主理想整环, 则 $\forall A\in R^{m\times n}$, $\exists P\in\mathrm{GL}(m,R)$, $Q\in\mathrm{GL}(n,R)$,*
$d_1\mid\cdots\mid d_r\in R^*$, s.t. $P\cdot A\cdot Q=\mathrm{diag}(d_1,\cdots,d_r,0,\cdots,0)$, 其中 $r=r(A)$, d_i 在不计乘法可逆元的意义下唯一.

证明: 利用 Bezout 等式归纳证明即可. □

推论 6.2.21 *设 F 为一个域, $A\in F^{n\times n}$. 记 $R=F[X]$, 则 $\exists P(X),Q(X)\in\mathrm{GL}(n,R)$, $p_r(X)\mid\cdots\mid p_1(X)\in R\setminus F$,*
s.t. $P(X)\cdot(X\cdot I_n-A)\cdot Q(X)=\mathrm{diag}(1,\cdots,1,p_r(X),\cdots,p_1(X))$, 其中 $r\in\mathbb{N}$ 唯一, $p_i(X)$ 在首一的意义下唯一.

注: 在上述推论中, $p_1(X),\cdots,p_r(X)$ 称为方阵 A 的**不变因子** (invariant factor), 我们将会看到它们唯一决定了方阵的相似标准形.

定理 6.2.22 (线性变换的循环分解) *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则存在 $\alpha_1,\cdots,\alpha_r\in V\setminus\{0\}$, 满足 $V=\bigoplus\limits_{i=1}^{r}F[T]\cdot\alpha_i$, 且 $p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)$. 记 $p_i(X)=p_{\alpha_i}(X)\,(1\leq i\leq r)$, 则 $r\in\mathbb{N}$ 与 $p_i(X)$ 均唯一, 且 $p_T(X)=p_1(X)$, $f_T(X)=\prod\limits_{i=1}^{r}p_i(X)$.*

证明: 记 $R=F[X]$. 由上述讨论与推论知, 存在 $p_r(X)\mid\cdots\mid p_1(X)\in R\setminus F$, 以及 R-模同构

$$\begin{aligned}V&\cong R^n/\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A)\cdot Q(X))\\&=R^n/\mathrm{column}_R(\mathrm{diag}(1,\cdots,1,p_r(X),\cdots,p_1(X)))\\&\cong R^n/\left(\bigoplus_{i=1}^{n-r}R\cdot\epsilon_i\oplus\bigoplus_{i=n-r+1}^{n}R\cdot p_{n-i+1}(X)\cdot\epsilon_i\right)\\&\cong\bigoplus_{i=1}^{r}R/R\cdot p_i(X),\end{aligned}$$

即存在 T-不变子空间的直和分解 $V=\bigoplus\limits_{i=1}^{r}V_i$, 且存在 R-模同构 $V_i\cong R/R\cdot p_i(X)$. 记 $\alpha_i\in V_i$ 为 $X+R\cdot p_i(X)$ 在此同构下的原像, 则 $V_i=F[T]\cdot\alpha_i$, 且 $p_{\alpha_i}(X)=p_i(X)$, 故 $p_T(X)=\mathrm{lcm}(p_1(X),\cdots,p_r(X))=p_1(X)$, 以及 $f_T(X)=\prod\limits_{i=1}^{r}f_{T|_{F[T]\cdot\alpha_i}}(X)=\prod\limits_{i=1}^{r}p_{T|_{F[T]\cdot\alpha_i}}(X)=\prod\limits_{i=1}^{r}p_i(X)$. 唯一性由上述 R-模同构的唯一性即知. □

注: 上述线性变换的循环分解可推广至主理想整环上: 设 R 为一个 P.I.D., M 为一个 R-有限生成的扭模, 则 $\exists r\in\mathbb{N},\ p_r\mid\cdots\mid p_1\in R^*\setminus R^\times,\ s.t.\ M\cong\bigoplus\limits_{i=1}^{r}R/R\cdot p_i$, 其中 $r\in\mathbb{N}$ 唯一, p_i 在不计乘法可逆元的意义下唯一.

6.2.3 循环分解的应用

值得注意的是, 线性变换的循环分解中循环向量组未必唯一:

例 6.2.1 (循环分解的不唯一性) *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的循环分解为 $V=\bigoplus\limits_{i=1}^{r}F[T]\cdot\alpha_i$, 其中 $p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)$. 记 $p_i(X)=p_{\alpha_i}(X)\,(1\le i\le r)$. 注意: 虽然 $r\in\mathbb{N}$ 与 $p_i(X)$ 都由 T 唯一决定, 但 $\{\alpha_i\}_{i=1}^r$ 未必唯一, 它们不唯一的原因可分为以下两类:*

(1) *在每个循环子空间 $F[T]\cdot\alpha_i$ 内, 可选取不同的循环向量;*

(2) *当 $r\ge 2$ 时, 循环子空间序列 $\{F[T]\cdot\alpha_i\}_{1\le i\le r}$ 可能完全不同.*

引理 6.2.23 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $\alpha\in V$ 满足 $V=F[T]\cdot\alpha$. 若 $|F|>\dim_F(V)\ge 2$, 则可取 $\beta\in V\setminus F\cdot\alpha$, 满足 $V=F[T]\cdot\beta$.*

证明: 任取 $\beta\in V\setminus F\cdot\alpha$. 由 $V=F[T]\cdot\alpha$ 知, $\exists f(X)\in F[X],\ s.t.\ \deg(f(X))\ge 1$, 且 $\beta=f(T)(\alpha)$. 注意 $(c+f(X))$ $(c\in F)$ 两两互素, 则由 $|F|>\dim_F(V)=\deg(p_\alpha(X))$ 知, $\exists c\in F,\ s.t.\ \gcd(c+f(X),p_\alpha(X))=1$. 记 $\beta'=c\alpha+\beta$, 则 $\beta'\in V\setminus F\cdot\alpha$, 且 $p_{\beta'}(X)=p_\alpha(X)$, 故 $V=F[T]\cdot\beta'$. □

注: 上述引理中关于域的要求是必要的, 反例如当 $|F|=\dim_F(V)=2$ 时, 设 V 的一组基为 $\{\alpha,\beta\}$, $T\in L(V)$ 满足 $(T-1)(\alpha)=\beta$, $T(\beta)=0$, 则 $p_T(X)=p_\alpha(X)=X(X-1)$, $p_\beta(X)=X$, $p_{\alpha+\beta}(X)=X-1$, 故 $V\setminus F\cdot\alpha$ 中元均不为循环向量.

引理 6.2.24 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的循环分解为 $V=\bigoplus\limits_{i=1}^{2}F[T]\cdot\alpha_i$, 其中 $p_{\alpha_2}(X)\mid p_{\alpha_1}(X)$. 若 $\max\{|F|,\dim_F(V)\}>2$, 则可取 $\beta_1,\beta_2\in V\setminus(F[T]\cdot\alpha_1\cup F[T]\cdot\alpha_2)$, 满足 $V=\bigoplus\limits_{i=1}^{2}F[T]\cdot\beta_i$ 也为循环分解.*

证明: 记 $p(X)=\dfrac{p_{\alpha_1}(X)}{p_{\alpha_2}(X)}$, $\beta_2=\alpha_2+p(T)(\alpha_1)$, 则可直接验证 $\beta_2\in V\setminus(F[T]\cdot\alpha_1\cup F[T]\cdot\alpha_2)$, 且 $p_{\beta_2}(X)=p_{\alpha_2}(X)$, 故 $V=F[T]\cdot\alpha_1\oplus F[T]\cdot\beta_2$. 由 $\max\{|F|,\dim_F(V)\}>2$ 知, 可取 $f(X)\in F[X]$, 满足 $p_{\alpha_1}(X)\nmid(1+f(X)p(X))$, 且 $p_{\alpha_2}(X)\nmid f(X)$. 记 $\beta_1=\alpha_1+f(T)(\beta_2)$, 则可直接验证 $\beta_1\in V\setminus(F[T]\cdot\alpha_1\cup F[T]\cdot\alpha_2)$, 且 $p_{\beta_1}(X)=p_{\alpha_1}(X)$, 故 $V=F[T]\cdot\beta_1\oplus F[T]\cdot\beta_2$. □

注: 上述引理中关于域的要求是必要的, 反例如当 $|F|=\dim_F(V)=2$ 时, 设 V 的一组基为 $\{\alpha_1,\alpha_2\}$, $T\in L(V)$ 满足 $T(\alpha_1)=\alpha_1$, $T(\alpha_2)=\alpha_2$, 则 $V\setminus(F[T]\cdot\alpha_1\cup F[T]\cdot\alpha_2)$ 中只有一个元 $\alpha_1+\alpha_2$.

命题 6.2.25 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则存在集合之间的一一对应:*

$$C(T)\cap\mathrm{GL}(V)\leftrightarrow\{(\alpha_1,\cdots,\alpha_r)\in V^{(r)}\colon V=\bigoplus_{i=1}^{r}F[T]\cdot\alpha_i\text{为循环分解, 其中}p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)\}.$$

证明: 固定 $T\in L(V)$ 的某个循环分解 $V=\bigoplus\limits_{i=1}^{r}F[T]\cdot\alpha_i$, 其中 $p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)$. 考虑对应

$$S\in C(T)\cap\mathrm{GL}(V)\mapsto(S(\alpha_1),\cdots,S(\alpha_r))\in V^{(r)},$$

则可直接验证 $V=\bigoplus\limits_{i=1}^{r}F[T]\cdot S(\alpha_i)$ 仍为循环分解, 其中 $p_{S(\alpha_i)}(X)=p_{\alpha_i}(X)$. 进一步地, 上述对应既单又满. □

另一值得注意的是, 线性变换的循环分解只可由它的部分不变子空间继承:

引理 6.2.26 *设 V 为域 F 上的线性空间, $T\in L(V)$, $\{W_i\}_{i\in I}$ 为 V 的一族 T-不变子空间, 满足 $V=\bigoplus\limits_{i\in I}W_i$, 则 $\forall f(X)\in F[X]$,*

(1) $\ker(f(T))=\bigoplus\limits_{i\in I}(\ker(f(T))\cap W_i)=\bigoplus\limits_{i\in I}\ker\left(f(T)\big|_{W_i}\right)$.

(2) $\mathrm{Im}(f(T))=\bigoplus\limits_{i\in I}(\mathrm{Im}(f(T))\cap W_i)=\bigoplus\limits_{i\in I}\mathrm{Im}\left(f(T)\big|_{W_i}\right)$.

证明: (1) 任取 $f(X)\in F[X]$, 由于 T-不变子空间也是 $f(T)$-不变子空间, 故通过将 T 换成 $f(T)$, 可不妨设 $f(X)=1$. 对于 $\alpha\in V$, 记 $\alpha=\sum\limits_{j=1}^{k}\alpha_{i_j}$, 其中 $\alpha_{i_j}\in W_{i_j}$, 则 $T(\alpha)=\sum\limits_{j=1}^{k}T(\alpha_{i_j})$, 其中 $T(\alpha_{i_j})\in W_{i_j}$, 故 $\alpha\in\ker(T)\Longleftrightarrow\forall 1\le j\le k$,
$\alpha_{i_j}\in\ker(T)$, 即 $\ker(T)=\bigoplus\limits_{i\in I}(\ker(T)\cap W_i)$.
(2) 完全同理. □

注:

(1) 警告: 一般的 T-不变子空间无法继承一般的 T-不变子空间分解; 简单的反例如 $T=\mathrm{id}_V$, 此时任意线性子空间都是 T-不变子空间, 但子空间的交与和未必满足分配律.

(2) 警告: 对于一般的 T-不变子空间 $W\subseteq V$, 条件 "$\mathrm{Im}(f(T))\cap W=\mathrm{Im}\left(f(T)\big|_W\right),\forall f(X)\in F[X]$" 未必成立, 满足此条件的不变子空间 W 称为 T-admissible. 事实上, T-不变子空间 $W\subseteq V$ 为 T-admissible $\Longleftrightarrow$ W 存在 T-不变的直和补空间, 可见 Kunze, Hoffman"Linear Algebra".

推论 6.2.27 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的循环分解为 $V=\bigoplus\limits_{i=1}^{r}F[T]\cdot\alpha_i$, 其中 $p_{\alpha_r}(X)\mid\cdots\mid p_{\alpha_1}(X)$, 则 $\forall f(X)\in F[X]$,*

(1) $\ker(f(T))=\bigoplus\limits_{i=1}^{r}\ker\left(f(T)\big|_{F[T]\cdot\alpha_i}\right)$, *不变因子为* $\gcd(f(X),p_{\alpha_r}(X))\mid\cdots\mid\gcd(f(X),p_{\alpha_1}(X))$ *中去掉 1.*

(2) $\mathrm{Im}(f(T))=\bigoplus\limits_{i=1}^{r}\mathrm{Im}\left(f(T)\big|_{F[T]\cdot\alpha_i}\right)$, *不变因子为* $\dfrac{p_{\alpha_r}(X)}{\gcd(f(X),p_{\alpha_r}(X))}\mid\cdots\mid\dfrac{p_{\alpha_1}(X)}{\gcd(f(X),p_{\alpha_1}(X))}$ *中去掉 1.*

反之, 不变子空间的循环分解未必可扩充为大空间的循环分解, 但不变因子之间总存在对应的整除关系:

例 6.2.2 *设 $V=F^4$, $T\in L(V)$ 在标准基下的矩阵表示为 $A=\mathrm{diag}(J_3(0),J_1(0))$, $W=F[T]\cdot(\epsilon_2+\epsilon_4)$, 则不存在 V 的循环分解使得 W 包含于其中某个循环子空间.*

证明: 先断言: V 中任意三维的循环子空间都不可能包含 W. 这是因为, 假设存在 $\alpha\in V$, 满足 $p_\alpha(X)=X^3$ 且 $W\subseteq F[T]\cdot\alpha$. 由于 $V=F[T]\cdot\epsilon_1\oplus F[T]\cdot\epsilon_4$, 记 $\alpha=f_1(T)(\epsilon_1)+f_2(T)(\epsilon_4)$; 又 $p_\alpha(X)=p_{\epsilon_1}(X)=X^3$ 且 $p_{\epsilon_4}(X)=X$, 则 $X\nmid f_1(X)$. 再由 $\epsilon_2+\epsilon_4\in F[T]\cdot\alpha$ 知, $\exists g(X)\in F[X]$, $s.t.\ \epsilon_2+\epsilon_4=g(T)(\alpha)=g(T)f_1(T)(\epsilon_1)+g(T)f_2(T)(\epsilon_4)$, 则 $T(\epsilon_1)=\epsilon_2=g(T)f_1(T)(\epsilon_1)$, $\epsilon_4=g(T)f_2(T)(\epsilon_4)$, 故 $X^3=p_{\epsilon_1}(X)\mid X-g(X)f_1(X)$, $X=p_{\epsilon_4}(X)\mid 1-g(X$
综上, 由 $X\nmid f_1(X)$ 且 $X^3\mid X-g(X)f_1(X)$ 知 $X\mid g(X)$; 但由 $X\mid 1-g(X)f_2(X)$ 知 $X\nmid g(X)$, 矛盾!

现任取 V 的循环分解 $V=F[T]\cdot\alpha_1\oplus F[T]\cdot\alpha_2$, 其中 $p_{\alpha_1}(X)=X^3$, $p_{\alpha_2}(X)=X$, 则 $F[T]\cdot\alpha_1$ 或 $F[T]\cdot\alpha_2$ 都不可能包含 W. □

注: 对于一般的 T-不变子空间 $W\subseteq V$, 是否可以给出性质 "存在 W 的某个循环分解可扩充为 V 的循环分解" 的等价刻画?

命题 6.2.28 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $W\subseteq V$ 为 T-不变子空间. 记 T 的不变因子为 $p_{r_V}(X)\mid\cdots\mid p_1(X)$, $T\big|_W$ 的不变因子为 $q_{r_W}(X)\mid\cdots\mid q_1(X)$, 则 $r_W\le r_V$, 且 $\forall 1\le i\le r_W$, $q_i(X)\mid p_i(X)$.*

证明: 考虑 V 的循环分解 $V=\bigoplus\limits_{i=1}^{r_V}F[T]\cdot\alpha_i$ 与 W 的循环分解 $V=\bigoplus\limits_{j=1}^{r_W}F[T]\cdot\beta_j$. 记 $n=\dim_F(V)$; 当 $r_V+1\le i\le n$ 与 $r_W+1\le j\le n$ 时, 约定 $\alpha_i=\beta_j=0$, 以及 $p_i(X)=q_j(X)=1$. 固定 $p(X)\in F[X]$ 为不可约多项式. 对于 $f(X)\in F[X]$, 记 $v(f(X))\coloneqq\max\{e\in\mathbb{N}\colon p(X)^e\mid f(X)\}$. 断言: $\forall 1\le i\le n$, $v(q_i(X))\le v(p_i(X))$.

事实上, 假设 $\exists 1\le t\le n$, $s.t.\ v(q_t(X))>v(p_t(X))$, 记 $k=v(q_t(X))\ge 1$, 则由不变因子之间的整除关系知, $|\{1\le i\le n\colon v(p_i(X))\ge k\}|\le t-1<t\le|\{1\le i\le n\colon v(q_i(X))\ge k\}|$. 由下述命题知,

$$\begin{aligned}\dim_F(\ker(p(T)^k))-\dim_F(\ker(p(T)^{k-1}))&=\sum_{i=1}^{n}(\deg\left(\gcd(p(X)^k,p_i(X))\right)-\deg\left(\gcd(p(X)^{k-1},p_i(X))\right))\\&=\deg(p(X))\cdot\sum_{i=1}^{n}(\min\{k,v(p_i(X))\}-\min\{k-1,v(p_i(X))\})\\&=\deg(p(X))\cdot|\{1\le i\le n\colon v(p_i(X))\ge k\}|;\end{aligned}$$

同理

$$\dim_F(\ker(p(T)\big|_W^k))-\dim_F(\ker(p(T)\big|_W^{k-1}))=\deg(p(X))\cdot|\{1\le i\le n\colon v(q_i(X))\ge k\}|;$$

因此

$$\dim_F(\ker(p(T)^k)) - \dim_F(\ker(p(T)^{k-1})) < \dim_F(\ker(p(T)|_W^k)) - \dim_F(\ker(p(T)|_W^{k-1})).$$

然而 $\forall k \geq 1$, $\dfrac{\ker(p(T)|_W^k)}{\ker(p(T)|_W^{k-1})} \longrightarrow \dfrac{\ker(p(T)^k)}{\ker(p(T)^{k-1})}$ 总是定义良好的线性单射, 故比较维数即知矛盾.

$$\alpha + \ker(p(T)|_W^{k-1}) \longmapsto \alpha + \ker(p(T)^{k-1})$$

因此由 $F[X]$ 中的唯一分解定理以及断言可知, $\forall 1 \leq i \leq n$, $q_i(X) \mid p_i(X)$. 特别地, $r_W \leq r_V$. □

以下我们说明线性变换的循环分解与普通循环子空间直和分解的不同之处, 从而给出循环分解中不变因子个数 $r \in \mathbb{N}$ 的几何意义.

命题 6.2.29 设 V 为域 F 上的有限维线性空间, $T \in L(V)$. 任取 V 的循环子空间直和分解 $V = \bigoplus_{i=1}^{s} F[T] \cdot \beta_i$, 则 $\forall f(X) \in F[X]$, $\dim_F(\ker(f(T))) = \sum_{i=1}^{s} \deg(\gcd(f(X), p_{\beta_i}(X)))$.

证明: 任取 $f(X) \in F[X]$. 由于 $\ker(f(T)) \subseteq V$ 可以继承 T 的任意不变子空间分解, 则取 V 的上述循环子空间直和分解知, $\ker(f(T)) = \bigoplus_{i=1}^{s} \ker\left(f(T)|_{F[T]\cdot\beta_i}\right) = \bigoplus_{i=1}^{s} \ker\left(\gcd(f, p_{\beta_i})(T)|_{F[T]\cdot\beta_i}\right)$, 故由循环模的等价刻画知, $\dim_F(\ker(f(T))) = \sum_{i=1}^{r} \dim_F\left(\ker\left(\gcd(f, p_{\beta_i})(T)|_{F[T]\cdot\beta_i}\right)\right) = \sum_{i=1}^{r} \deg(\gcd(f(X), p_{\beta_i}(X)))$. □

推论 6.2.30 设 V 为域 F 上的有限维线性空间, $T \in L(V)$. 任取 V 的循环子空间直和分解 $V = \bigoplus_{i=1}^{s} F[T] \cdot \beta_i$, 以及 $p(X) \in F[X]$ 为不可约多项式, $\dim_F(\ker(p(T))) = |\{1 \leq i \leq s : p(X) \mid p_{\beta_i}(X)\}| \cdot \deg(p(X))$. 特别地, $\dfrac{\dim_F(\ker(p(T)))}{\deg(p(X))} \in \{1, \cdots, s\}$.

推论 6.2.31 设 V 为域 F 上的有限维线性空间, $T \in L(V)$. 任取 V 的循环子空间直和分解 $V = \bigoplus_{i=1}^{s} F[T] \cdot \beta_i$, 则 $\max\{1 \leq k \leq s : \exists 1 \leq i_1 < \cdots < i_k \leq s,\ s.t.\ \deg(\gcd(p_{\beta_1}(X), \cdots, p_{\beta_k}(X))) > 1\} = \max\limits_{\substack{p(X)\in F[X] \\ \text{为不可约多项式}}} \dfrac{\dim_F(\ker(p(T)))}{\deg(p(X))}$.

特别地, 记 T 的循环分解为 $V = \bigoplus_{i=1}^{r} F[T] \cdot \alpha_i$, 其中 $p_{\alpha_r}(X) \mid \cdots \mid p_{\alpha_1}(X)$, 则

$r = \max\limits_{\substack{p(X)\in F[X] \\ \text{为不可约多项式}}} \dfrac{\dim_F(\ker(p(T)))}{\deg(p(X))} = \dfrac{\dim_F(\ker(p_0(T)))}{\deg(p_0(X))}$, 其中 $p_0(X) \in F[X]$ 为 $p_{\alpha_r}(X)$ 的任意不可约因子.

推论 6.2.32 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, $W \subseteq V$ 为 T-不变子空间. 记 r_V 为 $T \in L(V)$ 的循环分解中不变因子的个数, r_W 为 $T|_W \in L(W)$ 的循环分解中不变因子的个数, 则 $r_W \leq r_V$.

推论 6.2.33 设 V 为域 F 上的有限维线性空间, $T \in L(V)$ 的不变因子个数为 r, 则

$$r = \min\{s \in \mathbb{N} : \exists \beta_1, \cdots, \beta_s \in V,\ s.t.\ V = \bigoplus_{i=1}^{s} F[T] \cdot \beta_i\}.$$

注: 记 $R = F[X]$, 则此推论说明 V 作为 R-模分解成若干循环子模直和的最少项数为不变因子的个数. 于是由有限生成模的定义知, V 作为 R-模的最少生成元数量不超过不变因子的个数. 事实上, 以下我们将证明这二者恰好相等.

引理 6.2.34 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, $W_1, \cdots, W_k \subseteq V$ 为 T-不变子空间, 满足 $V = \sum_{i=1}^{k} W_i$, 则 $\forall f(X) \in F[X]$, $\dim_F(\ker(f(T))) \leq \sum_{i=1}^{k} \dim_F(\ker(f(T)) \cap W_i)$.

证明: 我们将此证明推迟至不变子空间理论一节中解释. □

推论 6.2.35 设 V 为域 F 上的有限维线性空间, $T \in L(V)$ 的不变因子个数为 r, 则

$$r = \min\{s \in \mathbb{N} : \exists \beta_1, \cdots, \beta_s \in V,\ s.t.\ V = \sum_{i=1}^{s} F[T] \cdot \beta_i\}.$$

证明: 假设 $\exists \beta_1, \cdots, \beta_s \in V,\ s.t.\ V = \sum\limits_{i=1}^{s} F[T]\cdot\beta_i$, 则由上述引理与前述推论的证明知, 任取 $p(X)\in F[X]$ 为不可约多项式, $\dim_F(\ker(p(T))) \leq \sum\limits_{i=1}^{s} \dim_F\left(\ker\left(p(T)\big|_{F[T]\cdot\beta_i}\right)\right) = |\{1\leq i\leq s\colon p(X)\mid p_{\beta_i}(X)\}|\cdot\deg(p(X))$, 故 $r = \max\limits_{\substack{p(X)\in F[X]\\ \text{为不可约多项式}}} \dfrac{\dim_F(\ker(p(T)))}{\deg(p(X))} \leq s.$ □

最后利用循环分解与不变因子, 我们可以讨论 T 的中心化子的 R-模结构:

定理 6.2.36 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的不变因子为 $p_r(X)\mid\cdots\mid p_1(X)$, 记 $R=F[X]$, 则:*

(1) *存在 R-模同构 $C(T)\cong (R/R\cdot p_{\max\{i,j\}}(X))_{1\leq i,j\leq n}$. 特别地, $\dim_F(C(T)) = \sum\limits_{i=1}^{r}(2i-1)\deg(p_i(X))$.*

(2) *进一步地, 记 $\{q_i(X)\}_{i=1}^{k}\subseteq F[X]$ 为 $p_T(X)$ 的所有不可约因子, 以及 $p_j(X) = \prod\limits_{i=1}^{k} q_i(X)^{e_{ij}}\ (1\leq j\leq r)$, 则存在 R-模同构 $C(T)\cong \bigoplus\limits_{i=1}^{k}\bigoplus\limits_{j,j'=1}^{r} R/R\cdot q_i(X)^{\min\{e_{i,j},e_{i,j'}\}}$. 特别地, $\dim_F(C(T)) = \sum\limits_{i=1}^{k}\deg(q_i(X))\sum\limits_{j,j'=1}^{r}\min\{e_{i,j},e_{i,j'}\}$.*

证明: (1) 取 $T\in L(V)$ 的循环分解为 $V=\bigoplus\limits_{i=1}^{r} F[T]\cdot\alpha_i$, 其中 $p_{\alpha_i}(X)=p_i(X)\,(1\leq i\leq r)$. 设 $S\in C(T)$, 考虑 $S(\alpha_j) = \sum\limits_{i=1}^{n} g_{ij}(T)(\alpha_i)\,(1\leq j\leq r)$, 其中 $g_{ij}(X)\in F[X]$ 且 $\deg(g_{ij}(X))<\deg(p_i(X))$. 注意 $T\circ S = S\circ T$, 则 $\forall 1\leq j\leq r,\ 0 = S(p_j(T)(\alpha_j)) = p_j(T)(S(\alpha_j)) = \sum\limits_{i=1}^{r} p_j(T)g_{ij}(T)(\alpha_j)$, 故 $\forall 1\leq i,j\leq r,\ p_i(X)\mid p_j(X)\cdot g_{ij}(X)$. 此式当 $i\geq j$ 时显然成立; 当 $i<j$ 时, 记 $g_{ij}(X) = \dfrac{p_i(X)}{p_j(X)}\cdot h_{ij}(X)$, 其中 $h_{ij}(X)\in F[X]$ 且 $\deg(h_{ij}(X))<\deg(p_j(X))$, 则 S 在 V 的 R-基 $\{\alpha_i\}_{i=1}^{r}$ 下的矩阵表示 $(g_{ij}(X))_{1\leq i,j\leq r}$ 恰好对应 $(R/R\cdot p_{\max\{i,j\}}(X))_{1\leq i,j\leq n}$ 中的方阵. 反之, 每个 $(R/R\cdot p_{\max\{i,j\}}(X))_{1\leq i,j\leq n}$ 中的方阵都唯一决定了一个 $S\in C(T)$. 又此对应关于 R-作用是等变的, 故存在 R-模同构 $C(T)\cong (R/R\cdot p_{\max\{i,j\}}(X))_{1\leq i,j\leq n}$.

(2) 由 T 的循环分解知, 存在 R-模同构 $V\cong\bigoplus\limits_{j=1}^{r} R/R\cdot p_j(X)$; 由中国剩余定理知, 任取 $1\leq j\leq r$, 存在 R-模同构 $R/R\cdot p_j(X)\cong\bigoplus\limits_{i=1}^{k} R/R\cdot q_i(X)^{e_{ij}}$, 则存在 R-模同构 $V\cong\bigoplus\limits_{j=1}^{r}\bigoplus\limits_{i=1}^{k} R/R\cdot q_i(X)^{e_{ij}}$. 注意作为 R-模, $C(T)\cong\mathrm{Hom}_R(V,V)$, 其中 $\mathrm{Hom}_R(V,V)\cong\bigoplus\limits_{j,j'=1}^{r}\bigoplus\limits_{i,i'=1}^{k}\mathrm{Hom}_R(R/R\cdot q_i(X)^{e_{ij}}, R/R\cdot q_{i'}(X)^{e_{i',j'}})\cong\bigoplus\limits_{j,j'=1}^{r}\bigoplus\limits_{i=1}^{k} R/R\cdot q_i(X)^{\min\{e_{i,j},e_{i,j'}\}}$. □

推论 6.2.37 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则 $\{S\in L(V)\colon C(T)\subseteq C(S)\} = C(C(T)) = F[T]$.*

证明: 第一个等号由定义即知; 第二个等号由定理 6.2.36的 (1) 中方阵计算即知. □

推论 6.2.38 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则 $\dim_F(V)\leq\dim_F(C(T))\leq\dim_F(V)^2$, 且前者取等号当且仅当 V 为 T-循环的, 后者取等号当且仅当 $T=c\mathrm{id}_V\,(c\in F)$. 进一步地, 若 $T\neq c\mathrm{id}_V\,(c\in F)$, 则 $\dim_F(C(T))\leq\dim_F(V)^2-2\dim_F(V)+2$, 且等号取到当且仅当 T 的矩阵表示相似于 $aI_n+bE_{12}\,(a,b\in F)$ 或 $\mathrm{diag}(a,b,\cdots,b)\,(a\neq b\in F)$.*

6.2.4 不变因子的求法

现在我们试图给出线性变换的循环分解中循环向量和不变因子的求法, 它们依赖于相抵变换过程中的可逆阵; 更具效率的算法可见初等因子的求法一节.

回忆设 V 为域 F 上的有限维线性空间, 基为 $B=\{\alpha_1,\cdots,\alpha_n\}$, $T\in L(V)$ 在基 B 下的矩阵表示为 $A\in F^{n\times n}$. 记 $R=F[X]$, 则有 R-模同构

$$V\xrightarrow{\cong} R^n/\mathrm{column}_R(X\cdot I_n-A)\xrightarrow{\cong} R^n/\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A))$$

$$\sum_{i=1}^{n} g_i(T)(\alpha_i)\longmapsto\sum_{i=1}^{n} g_i(X)\cdot\overline{\epsilon_i}\longmapsto\sum_{i,j=1}^{n} P_{ij}(X)g_i(X)\cdot\overline{\epsilon_i},$$

其中

$$\begin{aligned}&\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A))\\=&\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A)\cdot Q(X))\\=&\mathrm{column}_R(\mathrm{diag}(1,\cdots,1,p_r(X),\cdots,p_1(X)))\\=&\bigoplus_{i=1}^{n-r}R\cdot\epsilon_i\oplus\bigoplus_{i=n-r+1}^{n}R\cdot p_{n-i+1}(X)\cdot\epsilon_i,\end{aligned}$$

故

$$R^n/\mathrm{column}_R(P(X)\cdot(X\cdot I_n-A))\xrightarrow{\quad\cong\quad}\bigoplus_{i=1}^{r}R/R\cdot p_i(X)$$

$$\sum_{i=1}^{n}h_i(X)\cdot\overline{\epsilon_i}\longmapsto(h_n(X)+R\cdot p_1(X),\cdots,h_{n-r+1}(X)+R\cdot p_r(X)).$$

由于 $\forall 1\le i\le r$, $R/R\cdot p_i(X)$ 的一组基为 $\{X^j+R\cdot p_i(X)\colon 0\le j\le \deg(p_i(X))-1\}$, 则对应 V 的一组基为 $\{P^{-1}(T)T^j(\alpha_{n-i+1})\colon 1\le i\le r,\ 0\le j\le\deg(p_i(X))-1\}$, 其中 $\beta_i\coloneqq P^{-1}(T)(\alpha_{n-i+1})$ $(1\le i\le r)$ 为循环向量, 且 $p_{\beta_i}(X)=p_i(X)$.

当然, 我们可以从 $(X\cdot I_n-A)$ 的相抵标准形中直接读出 $p_r(X)\mid\cdots\mid p_1(X)$, 也可以通过 $(X\cdot I_n-A)$ 的非零子式性质间接写出 $p_r(X)\mid\cdots\mid p_1(X)$:

引理 6.2.39 *设 $(R,+,0;1,\cdot)$ 为一个交换环, $A\in R^{n\times n}$, 记 $D_i\,(1\le i\le r(A))$ 为 A 的所有 i 级非零子式的首一最大公因式, 则对 A 做任意的相抵变换都不改变 $D_i\,(1\le i\le r(A))$.*

证明: 分类讨论行、列初等变换即可. □

推论 6.2.40 *设 F 为一个域, $A\in F^{n\times n}$. 记 $D_i(X)\,(1\le i\le n)$ 为 $(X\cdot I_n-A)$ 的第 i 个行列式因子, 则 A 的不变因子为 $p_i(X)=\dfrac{D_{n-i+1}(X)}{D_{n-i}(X)}\,(1\le i\le r)$, 其中 $r=\max\{1\le i\le n\colon D_{n-i+1}(X)\ne D_{n-i}(X)\}$.*

参考文献与补注 6.2

(1) 关于有限生成扭模的表现的部分, 可以参考 J. J. Rotman“Advanced Modern Algebra”.

(2) 关于一般主理想整环上扭模的循环分解定理的部分, 可以参考 Dummit, Foote“Abstract Algebra”.

(3) 关于线性变换的中心化子结构的部分, 可以参考 N. Jacobson“Lectures in Abstract Algebra II: Linear Algebra”.

§6.3 准素循环分解与 Jordan 标准形

为了最终将有限维线性空间上的线性变换写成尽可能简单的形式, 我们将它的准素分解与循环分解结合起来, 得到若干不可分解的准素循环子模, 这就是所谓线性变换的准素循环分解, 此时它在某组基下的矩阵表示就是准素有理标准形或 Jordan 标准形.

6.3.1 不可分解子模与相似标准形

命题 6.3.1 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $W\subseteq V$ 为 T-不变子空间, 则以下条件等价:*

(0) *W 是不可分解的;*

(1) *W 是准素循环子空间;*

(2) *W 是某个准素子空间的循环子空间;*

(3) *W 是某个循环子空间的准素子空间.*

注: 显然取定准素分解和循环分解后, 每个准素分量与每个循环分量的交都是不可分解的, 但不可分解的子空间未必形如某个准素分量与某个循环分量的交. 反例如取 T 的矩阵表示为 $\begin{pmatrix}1&\lambda\\0&1\end{pmatrix}$ $(\lambda\in F^*)$, 此时 V 已是不可分解即准素循环的, 但 $\ker(T-\mathrm{id}_V)\subsetneq V$ 也是不可分解的.

于是为了显式地将线性空间分解为若干不可分解的不变子空间的直和, 我们可以采取以下三种方式之一:

(1) 同时取准素分解 $V=\bigoplus\limits_{i=1}^{k}V_i$ 与循环分解 $V=\bigoplus\limits_{j=1}^{r}W_i$, 则 $V=\bigoplus\limits_{i=1}^{k}\bigoplus\limits_{j=1}^{r}(V_i\cap W_j)$ 是准素循环分解;

(2) 先取准素分解 $V=\bigoplus\limits_{i=1}^{k}V_i$, 再对每个 V_i 取循环分解 $V_i=\bigoplus\limits_{j=1}^{r_i}W_{ij}$, 则 $V=\bigoplus\limits_{i=1}^{k}\bigoplus\limits_{j=1}^{r_i}W_{ij}$ 是准素循环分解;

(3) 先取循环分解 $V=\bigoplus\limits_{j=1}^{r}W_j$, 再对每个 W_j 取准素分解 $W_j=\bigoplus\limits_{i=1}^{k_j}V_{ij}$, 则 $V=\bigoplus\limits_{j=1}^{r}\bigoplus\limits_{i=1}^{k_j}V_{ij}$ 是准素循环分解.

事实上, 它们在形式上是统一的: 显然 (1) 已具有 (2),(3) 中的形式; 反之由以下引理即知.

引理 6.3.2 *设 V 为域 F 上的有限维线性空间, $\{W_i\}_{i=1}^{k}$ 为一族循环子空间, 满足 $p_{T|_{W_i}}(X)\,(1\leq i\leq k)$ 两两互素, 则 $\sum\limits_{i=1}^{k}W_i=\bigoplus\limits_{i=1}^{k}W_i$ 也为循环子空间.*

证明: 记 $W=\sum\limits_{i=1}^{k}W_i$. 由 $p_{T|_{W_i}}(X)\,(1\leq i\leq k)$ 两两互素知, $p_{T|_W}(X)=\prod\limits_{i=1}^{k}p_{T|_{W_i}}(X)$, 则对 W 准素分解知, $W=\bigoplus\limits_{i=1}^{k}W_i$. 于是 $f_{T|_W}(X)=\prod\limits_{i=1}^{k}f_{T|_{W_i}}(X)=\prod\limits_{i=1}^{k}p_{T|_{W_i}}(X)=p_{T|_W}(X)$, 即 W 也是循环子空间. (进一步地, 若记 $W_i=F[T]\cdot\alpha_i\,(1\leq i\leq k)$, $\alpha=\sum\limits_{i=1}^{k}\alpha_i$, 则 $W=F[T]\cdot\alpha$.) □

注: 此引理说明分解方式 (2) 也具有 (1) 的形式.

引理 6.3.3 *设 V 为域 F 上的有限维线性空间, $\{V_j\}_{j=1}^{r}$ 为一族准素子空间, 满足 $p_{T|_{V_j}}(X)\,(1\leq j\leq r)$ 是同一个不可约多项式的幂次, 则 $\sum\limits_{j=1}^{r}V_j$ 也为准素子空间.*

证明: 显然. □

注: 此引理说明分解方式 (3) 也具有 (1) 的形式.

综上, 我们可以将三种分解方式写成同一个表格, 其中第 (i,j) 个位置是第 i 个准素分量与第 j 个循环分量之交及其最小多项式:

$$\begin{pmatrix}(V_1\cap W_1,p_1^{e_{11}}(X)) & \cdots & (V_1\cap W_r,p_1^{e_{1r}}(X))\\ \vdots & \vdots & \vdots\\ (V_k\cap W_1,p_k^{e_{k1}}(X)) & \cdots & (V_k\cap W_r,p_k^{e_{kr}}(X))\end{pmatrix},$$

严格地说, 我们需要去掉此表中 $V_i\cap W_j=\{0\}$ (即 $e_{ij}=0$) 的位置, 剩下的 $p_i^{e_{ij}}(X)$ 在不计顺序的意义下由 T 唯一决定, 称为 T 的**初等因子** (elementary divisor).

类比线性变换的不变因子, 以下我们说明线性变换的初等因子的几何意义:

命题 6.3.4 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$ 的初等因子个数为 s, 则*

$$s=\max\{r\in\mathbb{N}\colon \exists\alpha_1,\cdots,\alpha_r\in V,\ s.t.\ V=\bigoplus_{j=1}^{r}F[T]\cdot\alpha_j\}.$$

证明: 假设 $\exists\alpha_1,\cdots,\alpha_r\in V,\ s.t.\ V=\bigoplus\limits_{j=1}^{r}F[T]\cdot\alpha_j$, 则对每个 $F[T]\cdot\alpha_j$ 取准素分解 $F[T]\cdot\alpha_j=\bigoplus\limits_{i=1}^{k_j}V_{ij}$, 可得 V 的准素循环分解 $V=\bigoplus\limits_{j=1}^{r}\bigoplus\limits_{i=1}^{k_j}V_{ij}$, 故 $s=\sum\limits_{j=1}^{r}k_j\geq r$. □

利用线性变换的初等因子, 我们可以给出它在某组基下矩阵表示的标准形. 例如, 由于 $V_i\cap W_j$ 是准素循环的, 记它的循环向量为 α_{ij}, 则可选取它的两组有序基如下:

(1) $\{\alpha_{ij},T(\alpha_{ij}),\cdots,T^{\deg(p_i^{e_{ij}}(X))-1}(\alpha_{ij})\}$;

(2) $\{\alpha_{ij},T(\alpha_{ij}),\cdots,T^{\deg(p_i(X))-1}(\alpha_{ij});p_i(T)(\alpha_{ij}),Tp_i(T)(\alpha_{ij}),\cdots,T^{\deg(p_i(X))-1}p_i(T)(\alpha_{ij});\cdots;$
$p_i^{e_{ij}-1}(T)(\alpha_{ij}),Tp_i^{e_{ij}-1}(T)(\alpha_{ij}),\cdots,T^{\deg(p_i(X))-1}p_i^{e_{ij}-1}(T)(\alpha_{ij})\}$.

在第一组有序基下, $T\big|_{V_i\cap W_j}$ 的矩阵表示为 $p_i^{e_{ij}}(X)$ 的友矩阵, 则 $T\big|_V$ 的矩阵表示为由初等因子的友矩阵组成的准对角阵, 称为 T 的**准素有理标准形** (primary rational canonical form).

在第二组有序基下, $T\big|_{V_i\cap W_j}$ 的矩阵表示为 $p_i^{e_{ij}}(X)$ 的 Jordan 块

$$J(p_i^{e_{ij}}(X)) \coloneqq \begin{pmatrix} C_{p_i(X)} & & & \\ N & C_{p_i(X)} & & \\ & \ddots & \ddots & \\ & & N & C_{p_i(X)} \end{pmatrix},$$

其中 $C_{p_i(X)}$ 为 $p_i(X)$ 的友矩阵, $N = E_{1,\deg(p_i(X))}$, $J(p_i^{e_{ij}}(X))$ 是 $e_{ij}\times e_{ij}$ 的分块阵, 则 $T\big|_V$ 的矩阵表示为由初等因子的 Jordan 块组成的准对角阵, 称为 T 的 **Jordan 标准形** (Jordan canonical form).

在有些情形中我们需要调整 Jordan 标准形中次对角线上的矩阵. 例如固定 $\epsilon\in F^*$, 上述 Jordan 块也可变形为

$$J(p_i^{e_{ij}}(X),\epsilon) \coloneqq \begin{pmatrix} C_{p_i(X)} & & & \\ \epsilon N & C_{p_i(X)} & & \\ & \ddots & \ddots & \\ & & \epsilon N & C_{p_i(X)} \end{pmatrix}.$$

这是因为, 取 $D_\epsilon \coloneqq \operatorname{diag}(I_{\deg(p_i(X))}, \epsilon I_{\deg(p_i(X))}, \cdots, \epsilon^{e_{ij}-1} I_{\deg(p_i(X))})$, 则 $D_\epsilon \cdot J(p_i^{e_{ij}}(X)) \cdot D_\epsilon^{-1} = J(p_i^{e_{ij}}(X),\epsilon)$.

6.3.2 复 Jordan 块与实 Jordan 块

例 6.3.1 (复 Jordan 块) *注意复数域上的不可约多项式都是一次的, 则初等因子 $(X-c)^e$ $(c\in\mathbb{C})$ 的 Jordan 块为*

$$J_e(c) \coloneqq \begin{pmatrix} c & & & \\ 1 & c & & \\ & \ddots & \ddots & \\ & & 1 & c \end{pmatrix}_{e\times e}.$$

例 6.3.2 (实 Jordan 块) *注意实数域上的不可约多项式都是不超过二次的, 则初等因子 $(X-c)^e$, $(X^2+aX+b)^e$ $(a,b,c\in\mathbb{R})$ 的 Jordan 块分别为*

$$\begin{pmatrix} c & & & \\ 1 & c & & \\ & \ddots & \ddots & \\ & & 1 & c \end{pmatrix}_{e\times e}, \quad \begin{pmatrix} 0 & -b & & & & & & \\ 1 & -a & & & & & & \\ 0 & 1 & 0 & -b & & & & \\ 0 & 0 & 1 & -a & & & & \\ & & \ddots & & \ddots & & & \\ & & & \ddots & & \ddots & & \\ & & & & 0 & 1 & 0 & -b \\ & & & & 0 & 0 & 1 & -a \end{pmatrix}_{2e\times 2e}.$$

在实际应用中, 记 X^2+aX+b 的一对共轭复根为 $r\pm si$, 则 $(X^2+aX+b)^e$ 的 Jordan 块可相似于

$$C_e(r,s) \coloneqq \begin{pmatrix} r & -s & & & & & & \\ s & r & & & & & & \\ 1 & 0 & r & -s & & & & \\ 0 & 1 & s & r & & & & \\ & & \ddots & & \ddots & & & \\ & & & \ddots & & \ddots & & \\ & & & & 1 & 0 & r & -s \\ & & & & 0 & 1 & s & r \end{pmatrix}_{2e\times 2e}.$$

注意这种变形的 *Jordan* 块并不是唯一确定的, 其中可将 s 换成 $-s$.

利用上述变形的实 Jordan 块, 我们可以得到复方阵相似于实方阵的判别条件.

命题 6.3.5 设 $A\in\mathbb{C}^{n\times n}$, 则以下条件等价:

(1) A 相似于实方阵;

(2) A 相似于 $\overline{A}$;

(3) $\forall c\in\mathbb{C}$, $\forall e\geq 1$, A 的复 *Jordan* 标准形中 $J_e(c)$ 与 $J_e(\overline{c})$ 的个数相同;

(4) $\forall c\in\mathbb{C}$, $\forall e\geq 1$, $r((cI_n-A)^e)=r((\overline{c}I_n-A)^e)$;

(5) $\forall c\in\mathbb{C}$, $\forall e\geq 1$, $r((cI_n-A)^e)=r((cI_n-\overline{A})^e)$;

证明: "(1)⇒(2)": 设 A 相似于 $B\in\mathbb{R}^{n\times n}$, 则 $\overline{A}$ 相似于 $\overline{B}=B$, 故 A 相似于 $\overline{A}$.

"(2)⇒(3)": 由 A 相似于 $\overline{A}$ 知, 在 A 的复 Jordan 标准形中, 共轭的特征值对应的同阶 Jordan 块个数相同.

"(3)⇔(4)": 任取 $c\in\mathbb{C}$, $e\geq 1$, 则 $r((cI_n-A)^e)=n-\dim_{\mathbb{C}}(\ker((cI_n-A)^e))=n-\sum\limits_{i=1}^{e}$ (A 的复 Jordan 标准形中 $J_{\geq i}(c)$ 的个数), 故 $r((cI_n-A)^{e-1})-r((cI_n-A)^e)=$($A$ 的复 Jordan 标准形中 $J_{\geq e}(c)$ 的个数), 因此 $r((cI_n-A)^{e-1})-2r((cI_n-A)^e)+r((cI_n-A)^{e+1})=$($A$ 的复 Jordan 标准形中 $J_e(c)$ 的个数), 于是结论显然.

"(4)⇔(5)": 注意共轭不改变矩阵的秩即可.

"(3)⇒(1)": 由 (3) 知, 对于 $c\in\mathbb{C}\setminus\mathbb{R}$ 以及 $e\geq 1$, A 的复 Jordan 标准形中 $J_e(c)$ 与 $J_e(\overline{c})$ 可两两配对. 记 $c=r+si\,(r,s\in\mathbb{R})$, 则 $\mathrm{diag}(J_e(c),\,J_e(\overline{c}))$ 可相似于 $C_e(r,s)$. 因此 A 的复 Jordan 标准形可相似于实 Jordan 标准形. □

推论 6.3.6 设 $A=\begin{pmatrix}A_{11}&A_{12}\\0&0\end{pmatrix}\in\mathbb{C}^{n\times n}$, 且 A_{11} 相似于实方阵, 则 A 也相似于实方阵.

证明: 任取 $c\in\mathbb{C}\setminus\mathbb{R}$, 以及 $e\geq 1$, 则 $r((cI_n-A)^e)=r\begin{pmatrix}(cI_k-A_{11})^e&*\\0&c^eI_{n-k}\end{pmatrix}=r((cI_k-A_{11})^e)+n-k$; 同理 $r((\overline{c}I_n-A)^e)=r((\overline{c}I_k-A_{11})^e)+n-k$. 由 A_{11} 相似于实方阵以及命题 6.3.5知, $r((cI_k-A_{11})^e)=r((\overline{c}I_k-A_{11})^e)$, 因此 $r((cI_n-A)^e)=r((\overline{c}I_n-A)^e)$. 再由命题 6.3.5知 A 也相似于实方阵. □

注: 上述推论对于一般的准对角块相似于实方阵的准上三角复方阵不成立, 例如取 $A=\begin{pmatrix}i&0&2i&2i\\0&-i&0&0\\0&0&i&0\\0&0&0&-i\end{pmatrix}$, 其中准对角块 $\begin{pmatrix}i&0\\0&-i\end{pmatrix}$ 相似于实方阵 $\begin{pmatrix}0&-1\\1&0\end{pmatrix}$, 但 $r(iI_4-A)=3\neq 2=r(-iI_4-A)$, 故 A 不相似于实方阵.

推论 6.3.7 设 $A\in\mathbb{C}^{n\times n}$, 则 $A\overline{A}$ 相似于 $\overline{A}A$, 也相似于实方阵.

证明: 考虑 $A\overline{A}$ 与 $\overline{A}A$ 的 Jordan 标准形: 由推论 9.5.8知, 它们的可逆部分相同; 由 $\forall k\geq 1$, $r((A\overline{A})^k)=r((\overline{A}A)^k)$ 知, 它们的幂零部分相同. 因此 $A\overline{A}$ 相似于 $\overline{A}A$. 由命题 6.3.5知, $A\overline{A}$ 也相似于实方阵. □

回忆通过不变因子或初等因子的观点, 我们已知任意域上的方阵总相似于它的转置. 以下试图借助 Jordan 标准形给出复数域或实数域上这一事实的直观证明.

命题 6.3.8 设 F 为一个域, 满足 $F[X]$ 中的不可约多项式都是不超过二次的, $A\in F^{n\times n}$, 则 A 相似于 A^t.

证明: 通过相似, 可不妨设 A 已为 Jordan 标准形; 再考虑准对角块的相似, 可不妨设 A 为一个 Jordan 块, 对应的初等因子为 $(X-c)^e$ 或 $(X^2+aX+b)^e$ $(a,b,c\in F)$. 记 $J_e=E_{1e}+E_{2,e-1}+\cdots+E_{e1}$, 则 $J_e^2=I_e$, 且 J_eAJ_e 为 A 的中心对称阵. 取例 6.3.2中变形的 Jordan 块知 A 的中心对称阵即 A^t. □

6.3.3 特征多项式与最小多项式的特例

本小节收集特征多项式与最小多项式的特殊性质对线性变换的影响.

命题 6.3.9 设 V 为域 F 上的有限维线性空间, $T\in L(V)$, 则以下条件等价:

(1) $p_T(X) = f_T(X)$;
(2) V 是 T-循环的;
(3) V 的准素分解已是不可分解的 T-不变子空间的直和分解;
(4) 任意 T-不变子空间的 T-不变直和补空间若存在必唯一.

命题 6.3.10 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则以下条件等价:
(1) $p_T(X)$ 的不可约因子重数均为 1;
(2) T 为半单变换, 即不可分解的 T-不变子空间必为不可约的;
(3) 任意 T-不变子空间 $W \subseteq V$ 都是 T-*admissible*, 即 $\forall f(X) \in F[X]$, $\mathrm{Im}(f(T)) \cap W = \mathrm{Im}\left(f(T)\big|_W\right)$;
(4) 任取 T-不变的子空间升链 $W_1 \subseteq \cdots \subseteq W_k$, 存在 T-不变的子空间降链 $U_1 \supseteq \cdots \supseteq U_k$, 满足 $W_i \oplus U_i = V$, $\forall 1 \leq i \leq k$.

推论 6.3.11 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则以下条件等价:
(1) $f_T(X)$ 的不可约因子重数均为 1;
(2) $p_T(X) = f_T(X)$ 且 T 为半单变换;
(3) V 的准素分解已是不可约的 T-不变子空间的直和分解;
(4) 任意 T-不变子空间 W 均形如 $W = \bigoplus_{i=1}^{k} W_i$, 其中 W_i 为不可约的 T-不变子空间, 且 $p_{T|_W}(X)$ 为 k 个不同的不可约因子的乘积;
(5) 任意 T-不变子空间都存在唯一的 T-不变直和补空间.

命题 6.3.12 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则以下条件等价:
(1) $p_T(X)$ 为某个不可约多项式的幂;
(2) $f_T(X)$ 为某个不可约多项式的幂;
(3) V 的循环分解已是不可分解的 T-不变子空间的直和分解;
(4)

命题 6.3.13 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则以下条件等价:
(1) $p_T(X)$ 为不可约多项式;
(2) $p_T(X) \cdot F[X] = \{f(X) \in F[X] \colon \ker(f(T)) \neq \{0\}\}$;
(3) $F[T] = F[X]/p_T(X) \cdot F[X]$ 是 F 的扩域;
(4) V 的循环分解已是不可约的 T-不变子空间的直和分解;
(5) $\forall \alpha, \beta \in V$, $\exists S \in C(T) \cap \mathrm{GL}(V)$, $s.t.\ \beta = S(\alpha)$.

证明:

□

推论 6.3.14 设 V 为域 F 上的有限维线性空间, $T \in L(V)$, 则以下条件等价:
(1) $f_T(X)$ 为不可约多项式;
(2) $f_T(X) \cdot F[X] = \{f(X) \in F[X] \colon \ker(f(T)) \neq \{0\}\}$;
(3) V 是不可约的 T-不变子空间;
(4) $\forall \alpha \in V \setminus \{0\}$, $V = F[T] \cdot \alpha$.

6.3.4 初等因子的求法

现在我们总结线性变换的准素循环分解中循环向量和初等因子的求法. 一个经典的算法是由 D. Dummit, R. Foote 给出的; 最近安师等人给出了一种更具效率的算法, 可以参考 J. An, K. Lin, Y. Sun " Construction of bases for rational and Jordan canonical forms"(2022).

参考文献与补注 6.3
(1)
(2)
(3)

§6.4 不变子空间理论

本节专门总结在线性变换的准素分解和循环分解中不变子空间思想的意义与应用.

6.4.1 同时对角化与同时上三角化

命题 6.4.1 *设 V 为域 F 上的有限维线性空间, $\mathscr{F}\subseteq L(V)$, 则:*

(1) *$\mathscr{F}$ 中元可同时对角化 $\Longleftrightarrow$ $\mathscr{F}$ 中元两两可交换, 且每个元均可对角化;*

(2) *$\mathscr{F}$ 中元可同时上三角化 $\Longleftarrow$ $\mathscr{F}$ 中元两两可交换, 且每个元均可上三角化;*

注:

(1) 上述命题证明的核心是, 对于一族两两交换的线性映射, 可以找它们的公共不变子空间, 从而利用归纳法即知结论.

(2) 注意 (2) 的 "⇒" 部分并不成立, 反例如当 $n\geq 3$ 时, $F^{n\times n}$ 中上三角阵就不是两两可交换的. 一般地,

$\mathscr{F}$中元可同时上三角化 $\Longleftrightarrow$$\mathscr{F}$中每个元均可上三角化, 且$\forall m\geq 1$, $\forall 1\leq i<j\leq m$,

$$\forall g(X_1,\cdots,X_m)\in F[X_1,\cdots,X_m]\ (\text{这里}X_iX_j\neq X_jX_i),$$

$$\forall A_1,\cdots,A_m\in\mathscr{F},\ g(A_1,\cdots,A_m)\cdot(A_iA_j-A_jA_i)\text{为幂零阵}.$$

M. P. Drazin, J. W. Dungey, K. W. Gruenberg "Some Theorems on Commutative Matrices"(1950) 给出了一个纯线性代数的证明.

推论 6.4.2 *设 F 为代数闭域, V 为域 F 上的有限维线性空间, $\mathscr{F}\subseteq L(V)$ 中元两两可交换, 则 $\mathscr{F}$ 中元可同时上三角化.*

推论 6.4.3 *设 V 为有限维实线性空间, $\mathscr{F}\subseteq L(V)$ 中元两两可交换, 则 $\mathscr{F}$ 中元可同时准上三角化, 其中对角块为至多二阶的.*

证明: 考虑 $V_{\mathbb{C}}\coloneqq V\underset{\mathbb{R}}{\otimes}\mathbb{C}$ 为有限维实线性空间, $\mathscr{F}\otimes 1\subseteq L(V_{\mathbb{C}})$ 中元两两可交换, 则由推论 6.4.2知, $\mathscr{F}\otimes 1$ 中元存在公共特征向量 $\alpha_{\mathbb{C}}\in V_{\mathbb{C}}$. 记 $\alpha_{\mathbb{C}}=\alpha_1+\sqrt{-1}\alpha_2$, 其中 $\alpha_1,\alpha_2\in V$. 若 $\{\alpha_1,\alpha_2\}$ 实线性相关, 则存在不全为零的 $c_1,c_2\in\mathbb{R}$, 以及 $\beta\in V$, 满足 $\alpha_1=c_1\beta$, $\alpha_2=c_2\beta$, 此时 $\beta\in V$ 也为 $\mathscr{F}\otimes 1$ 中元的公共特征向量, 且对应的特征值均为实数. 若 $\{\alpha_1,\alpha_2\}$ 实线性无关, 则它生成的实线性子空间是 $\mathscr{F}\otimes 1$ 中元的公共不变子空间, 且对应的矩阵表示均为二阶实矩阵. 最后再在商空间 $V/\mathrm{Span}_{\mathbb{R}}(\{\alpha_1,\alpha_2\})$ 上对维数归纳即知结论. □

对于 F 为代数闭域且 $|\mathscr{F}|=2$ 的情形, 以下我们给出一些可同时上三角化的充分条件, 但它们都不是必要的. 这里的技术都是证明其中一个方阵的核或像是另一个方阵的不变子空间.

命题 6.4.4 *设 F 为一个代数闭域, $A,B\in F^{n\times n}$, 满足条件 $AB=0$ 或 $r(AB-BA)\leq 1$ 之一, 则 A,B 可同时上三角化.*

证明: (1) 若 $AB=0$, 则 $A(\mathrm{Im}(B))=\mathrm{Im}(AB)=\{0\}\subseteq\mathrm{Im}(B)$, 故 $\mathrm{Im}(B)$ 为 A-不变子空间. 若 $\mathrm{Im}(B)=\{0\}$, 即 $B=0$, 则只需将 A 上三角化即可; 若 $\mathrm{Im}(B)=F^{n\times 1}$, 则 $A=0$, 只需将 B 上三角化即可; 现设 $\{0\}\subsetneqq\mathrm{Im}(B)\subsetneqq F^{n\times 1}$, 这是 A,B 的公共不变子空间, 则 A,B 可同时分别相似于相同分块方阵 $\widetilde{A}=\begin{pmatrix}\widetilde{A}_{11}&\widetilde{A}_{12}\\0&\widetilde{A}_{22}\end{pmatrix}$, $\widetilde{B}=\begin{pmatrix}\widetilde{B}_{11}&\widetilde{B}_{12}\\0&\widetilde{B}_{22}\end{pmatrix}$, 且 $\widetilde{A}\widetilde{B}=0$. 再由 $\widetilde{A}_{11}\widetilde{B}_{11}=0$, $\widetilde{A}_{22}\widetilde{B}_{22}=0$, 对 n 归纳即知结论.

(2) 若 $r(AB-BA)=0$, 即 $AB=BA$, 则由推论 6.4.2即知结论; 现设 $r(AB-BA)=1$, 断言: 至少 $\ker(B)$ 与 $\mathrm{Im}(B)$ 之一为 A-不变子空间. (这是因为: 假设 $\ker(B)$ 不是 A-不变子空间, 即 $\exists\alpha\in\ker(B)$, $s.t.$ $A\alpha\notin\ker(B)$, 则

$0\neq -BA\alpha=(AB-BA)\alpha\in\mathrm{Im}(AB-BA)$. 又 $\dim_F(\mathrm{Im}(AB-BA))=1$, 则 $\mathrm{Span}_F(BA\alpha)=\mathrm{Im}(AB-BA)$, 故 $\mathrm{Im}(AB)\subseteq\mathrm{Im}(AB-BA)+\mathrm{Im}(BA)\subseteq\mathrm{Im}(BA)\subseteq\mathrm{Im}(B)$, 即 $\mathrm{Im}(B)$ 为 A-不变子空间.)

取 $c\in\sigma(B)$. 若 B 相似于 cI_n, 则可先将 A,B 同时分别相似于 $\widetilde{A},cI_n$, 再将 $\widetilde{A}$ 上三角化即可; 若 B 不相似于 cI_n, 则至少 $\{0\}\subsetneqq\ker(B-cI_n),\mathrm{Im}(B-cI_n)\subsetneqq F^{n\times 1}$ 之一为 A,B 的公共不变子空间, 以下过程与 (1) 类似. □

反之, 我们指出任意方阵的不变子空间总可写成某个与它交换的方阵的核或像.

命题 6.4.5 *设 F 为一个域, $A \in F^{n\times n}$, $W \subseteq F^{n\times 1}$ 为 A-不变子空间, 则 $\exists B, C \in F^{n\times n}$, s.t. $AB = BA$, $AC = CA$, $BC = CB = 0$, $W = \ker(B) = \mathrm{Im}(C)$.*

证明: 通过相似, 可不妨设 $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$, 其中 $A_{11} \in F^{k\times k}$, $W = F^{k\times 1}$. 由于方阵总相似于它的转置, 故可取 $P \in \mathrm{GL}(n, F)$, $P_1 \in \mathrm{GL}(k, F)$, $P_2 \in \mathrm{GL}(n-k, F)$, s.t. $P^{-1}AP = A^t$, $P_1^{-1}A_{11}P_1 = A_{11}^t$, $P_2^{-1}A_{22}P_2 = A_{22}^t$. 可直接验证 $A\cdot\begin{pmatrix} P_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} P_1 & 0 \\ 0 & 0 \end{pmatrix}\cdot A^t$; $A^t\cdot\begin{pmatrix} 0 & 0 \\ 0 & P_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & P_2 \end{pmatrix}\cdot A$; 令 $B = P\cdot\begin{pmatrix} 0 & 0 \\ 0 & P_2 \end{pmatrix}$, $C = \begin{pmatrix} P_1 & 0 \\ 0 & 0 \end{pmatrix}\cdot P^{-1}$ 即可. □

最后, 当一族线性变换没有非平凡的公共不变子空间时, 我们希望证明它具有尽可能大的线性无关性. 为方便起见, 以下先引入一些术语:

设 V 为域 F 上的线性空间, $\mathscr{A} \subseteq L(V)$ 为线性子代数 (未必含幺元 id_V). 若 $\mathscr{A}$ 没有非平凡的公共不变子空间, 则称 $\mathscr{A}$ 在 V 上的作用是**不可约的** (irreducible); 若 $\forall \alpha \in V \setminus \{0\}$, $\mathscr{A}\alpha := \{T(\alpha) : T \in \mathscr{A}\} = V$, 则称 $\mathscr{A}$ 在 V 上的作用是**传递的** (transitive).

引理 6.4.6 *设 V 为域 F 上的线性空间, $\mathscr{A} \subseteq L(V)$ 为线性子代数. 对于 $\mathscr{A}$ 在 V 上的作用:*

(1) *若 $\mathscr{A}$ 是传递的, 则 $\mathscr{A}$ 是不可约的;*

(2) *若 $\mathscr{A}$ 是不可约的, 且 $(\dim_F(V), \dim_F(\mathscr{A})) \neq (1, 0)$, 则 $\mathscr{A}$ 是传递的.*

证明: 当 $V = \{0\}$ 时, $L(V) = \{0\}$, 此时 $\mathscr{A} = \{0\}$ 是不可约且传递的. 当 $\dim_F(V) = 1$ 时, $L(V) = \mathrm{Span}_F(\{\mathrm{id}_V\})$, 此时 $\mathscr{A} = \{0\}$ 是不可约但非传递的, $\mathscr{A} = \mathrm{Span}_F(\{\mathrm{id}_V\})$ 是不可约且传递的. 以下设 $\dim_F(V) \geq 2$.

(1) 设 $\mathscr{A}$ 是传递的, $W \subseteq V$ 为 $\mathscr{A}$ 的公共不变子空间. 若 $W \neq \{0\}$, 则取 $\alpha \in W \setminus \{0\}$, 由 $V = \mathscr{A}\alpha \subseteq W \subseteq V$ 知 $W = V$, 故 $\mathscr{A}$ 是不可约的.

(2) 设 $\mathscr{A}$ 是不可约的, 固定 $\alpha \in V \setminus \{0\}$, 由 $\mathscr{A}\alpha \subseteq V$ 为 $\mathscr{A}$ 的公共不变子空间知, $\mathscr{A}\alpha = \{0\}$ 或 V. 假设 $\mathscr{A}\alpha = \{0\}$, 则 $\{0\} \subsetneqq \mathrm{Span}_F(\{\alpha\}) \subsetneqq V$ 为 $\mathscr{A}$ 的非平凡公共不变子空间, 矛盾! 因此 $\forall \alpha \in V \setminus \{0\}$, $\mathscr{A}\alpha = V$, 即 $\mathscr{A}$ 是传递的. □

引理 6.4.7 *设 F 为代数闭域, V 为域 F 上的有限维线性空间, $\{0\} \neq \mathscr{A} \subseteq L(V)$ 为传递的线性子代数, 则 $\mathscr{A}$ 中包含某个秩一的线性变换.*

证明: 对 $\dim_F(V) \in \mathbb{N}^*$ 归纳证明: 当 $\dim_F(V) = 1$ 时, $L(V) = \mathrm{Span}_F(\{\mathrm{id}_V\})$, 由 $\{0\} \neq \mathscr{A}$ 知 $\mathscr{A} = \mathrm{Span}_F(\{\mathrm{id}_V\})$, 结论显然. 现设 $\dim_F(V) = n \geq 2$, 且当 $1 \leq \dim_F(V) \leq n-1$ 时结论成立. 先断言: $\mathscr{A} \not\subseteq \mathrm{GL}(n, F) \cup \{0\}$. 事实上, 由 $\dim_F(V) \geq 2$ 知, $\mathrm{Span}_F(\{\mathrm{id}_V\})$ 在 V 上的作用不是传递的, 则 $\mathscr{A} \not\subseteq \mathrm{Span}_F(\{\mathrm{id}_V\})$. 取 $T \in \mathscr{A} \setminus \mathrm{Span}_F(\{\mathrm{id}_V\})$. 若 $T \notin \mathrm{GL}(n, F)$, 则断言已成立; 若 $T \in \mathrm{GL}(n, F)$, 取 $c \in \sigma(T)$, 记 $S = cT - T^2 = T \circ (c\mathrm{id}_V - T)$, 则 $S \in \mathscr{A}$ 且 $S \notin \mathrm{GL}(n, F) \cup \{0\}$, 故断言也成立.

于是由断言可取 $S \in \mathscr{A} \setminus (\mathrm{GL}(n, F) \cup \{0\})$, 则 $\mathrm{Im}(S) \subseteq V$ 为非平凡的线性子空间. 再断言: $S \circ \mathscr{A}\big|_{\mathrm{Im}(S)}$ 在 $\mathrm{Im}(S)$ 上的作用传递. 事实上, 任取 $\alpha \in \mathrm{Im}(S) \setminus \{0\}$, 由 $\mathscr{A}$ 在 V 上的作用传递知, $\mathscr{A}\big|_{\mathrm{Im}(S)}\alpha = \mathscr{A}\alpha = V$, 则 $S \circ \mathscr{A}\big|_{\mathrm{Im}(S)}\alpha = \mathrm{Im}(S)$. 因此由归纳假设知, $S \circ \mathscr{A}\big|_{\mathrm{Im}(S)}$ 包含某个秩一的线性变换, 即 $\exists T_0 \in \mathscr{A}$, s.t. $r(S \circ T_0\big|_{\mathrm{Im}(S)}) = 1$, 故 $S \circ T_0 \circ S \in \mathscr{A}$ 且 $r(S \circ T_0 \circ S) = 1$. □

引理 6.4.8 *设 V 为域 F 上的有限维线性空间, $\mathscr{A} \subseteq L(V)$ 为传递的线性子代数, 且包含某个秩一的线性变换, 则*

$\mathscr{A} = L(V)$.

证明: 设 $\mathscr{A}$ 包含秩一的线性变换 $\alpha_0 \otimes f_0 : V \longrightarrow V$, $\beta \longmapsto f_0(\beta)\alpha_0$, 其中 $\alpha_0 \in V$, $f_0 \in V^*$. 一方面, 注意 $\forall T \in \mathscr{A}$, $T(\alpha_0) \otimes f_0 = T \circ (\alpha_0 \otimes f_0) \in \mathscr{A}$, 则由 $\mathscr{A}$ 在 V 上的作用传递知, $\forall \alpha \in V$, $\alpha \otimes f_0 \in \mathscr{A}$. 另一方面, 由引理 6.4.6知, $\mathscr{A}$ 在 V 上的作用不可约. 断言: $\mathscr{A}^t$ 在 V^* 上的作用也不可约. 事实上, 任取 $M \subseteq V^*$ 为 $\mathscr{A}^t$ 的公共不变子空间, 由命题 3.2.8知, $\forall T \in \mathscr{A}$, $T^t(M) = (T^{-1}(M^\diamondsuit))^0 \subseteq M = (M^\diamondsuit)^0$, 即 $T^{-1}(M^\diamondsuit) \supseteq M^\diamondsuit$, 也即 $M^\diamondsuit \supseteq T(M^\diamondsuit)$, 故 $M^\diamondsuit \subseteq V$ 为 $\mathscr{A}$ 的公共不变子空间. 由 $\mathscr{A}$ 在 V 上的作用不可约知, $M^\diamondsuit = \{0\}$ 或 V, 即 $M = V^*$ 或 $\{0\}$.

再由引理 6.4.6知, $\mathscr{A}^t$ 在 V^* 上的作用传递. 注意 $\forall T\in\mathscr{A}$, $\alpha_0\otimes T^t(f_0)=(\alpha_0\otimes f_0)\circ T\in\mathscr{A}$, 则 $\forall f\in V^*$, $\alpha_0\otimes f\in\mathscr{A}$. 综上可知, $\forall\alpha\in V$, $\forall f\in V^*$, $\alpha\otimes f\in\mathscr{A}$. 又 $L(V)=\mathrm{Span}_F(\{\alpha\otimes f\colon \alpha\in V,\ f\in V^*\})$, 则 $L(V)=\mathscr{A}$.

□

定理 6.4.9 (Burnside) *设 F 为代数闭域, V 为域 F 上的有限维线性空间, $\mathscr{A}\subseteq L(V)$ 为不可约的线性子代数, 且 $(\dim_F(V),\dim_F(\mathscr{A}))\neq(1,0)$, 则 $\mathscr{A}=L(V)$.*

证明: 由引理 6.4.6, 引理 6.4.7与引理 6.4.8 即知. □

推论 6.4.10 *设 F 为代数闭域, V 为域 F 上的有限维线性空间, $G\subseteq L(V)$ 为不可约子群, 则 $\mathrm{Span}_F(G)=L(V)$.*

注:

(1) 上述 Burnside 定理对于一般域上的线性空间未必成立, 反例如取 $F=\mathbb{R}$, $V=\mathbb{R}^2$, $\mathscr{A}=\mathrm{Span}_{\mathbb{R}}(\{\mathrm{id}_{\mathbb{R}^2},R\})$, 其中 $R\in L(\mathbb{R}^2)$ 为逆时针旋转 $\pi/2$, 则 $\mathscr{A}$ 为不可约的, 且 $\dim_{\mathbb{R}}(\mathscr{A})=2<4=\dim_{\mathbb{R}}(L(\mathbb{R}^2))$.

(2) 上述 Burnside 定理对于无穷维复线性空间未必成立, 反例如取 $\mathscr{A}\subsetneqq L(V)$ 为有限秩算子代数, 则由 $\mathscr{A}$ 传递以及引理 6.4.6可知 $\mathscr{A}$ 不可约. 更多的讨论可以参考 V. Lomonosov“An extension of Burnside's Theorem to infinite dimensional spaces”.

6.4.2 不变子空间的分解与生成

在循环分解的讨论中, 我们已经发现线性变换的核与像总能继承大空间的任意不变子空间直和分解. 以下的讨论将推广这一事实.

引理 6.4.11 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $W\subseteq V$ 为 T-不变子空间. 记 $\overline{T}$: $V/W\longrightarrow V/W$, $\alpha+W\longmapsto T(\alpha)+W$ 为诱导映射, 则 $\max\{\dim_F(\ker(T|_W)),\dim_F(\ker(\overline{T}))\}\leq\dim_F(\ker(T))\leq\dim_F(\ker(T|_W))+\dim_F(\ker(\overline{T}))$.*

证明: 一方面, 显然 $\ker(T|_W)\subseteq\ker(T)$; 由线性代数基本定理知, $T^{-1}(W)/\ker(T|_{T^{-1}(W)})\cong W\cap\mathrm{Im}(T)$, 则 $\dim_F(T^{-1}(W))=\dim_F(\ker(T|_{T^{-1}(W)}))+\dim_F(W\cap\mathrm{Im}(T))\leq\dim_F(\ker(T))+\dim_F(W)$. 注意 $\ker(\overline{T})=T^{-1}(W)/W$, 故 $\dim_F(\overline{T})\leq\dim_F(T)$. 另一方面, 由第二同构定理知, $\ker(T)/\ker(T|_W)\cong(\ker(T)+W)/W\subseteq\ker(\overline{T})$, 则 $\dim_F(\ker(T))\leq\dim_F(\ker(T|_W))+\dim_F(\ker(\overline{T}))$. □

注: 上述引理的本质是同调论中的 “蛇引理”(Snake Lemma). 考虑线性空间正合列的交换图:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & W & \longrightarrow & V & \longrightarrow & V/W & \longrightarrow & 0\,,\\
 & & \downarrow{\scriptstyle T|_W} & & \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle \overline{T}} & & \\
0 & \longrightarrow & W & \longrightarrow & V & \longrightarrow & V/W & \longrightarrow & 0
\end{array}$$

则存在线性空间的长正合列 $0\to\ker(T|_W)\to\ker(T)\to\ker(\overline{T})\xrightarrow{\delta}\mathrm{coker}(T|_W)\to\mathrm{coker}(T)\to\mathrm{coker}(\overline{T})\to0$, 即

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \ker(T|_W) & \longrightarrow & \ker(T) & \longrightarrow & \ker(\overline{T}) & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & W & \longrightarrow & V & \longrightarrow & V/W & \longrightarrow & 0\\
 & & \downarrow{\scriptstyle T|_W} & & \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle \overline{T}} & & \\
0 & \longrightarrow & W & \longrightarrow & V & \longrightarrow & V/W & \longrightarrow & 0\\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & \mathrm{coker}(T|_W) & \longrightarrow & \mathrm{coker}(T) & \longrightarrow & \mathrm{coker}(\overline{T}) & \longrightarrow & 0
\end{array}\quad(\delta)\,.$$

特别地, 截取该正合列的前三、四、五项即得上述引理.

推论 6.4.12 *设 V,W 为域 F 上的有限维线性空间, $p\in L(V,W)$ 为满射, $T\in L(V),S\in L(W)$ 满足 $p\circ T=S\circ p$, 则*

$$\max\{\dim_F\left(\ker\left(T|_{\ker(p)}\right)\right),\dim_F(\ker(S))\}\leq\dim_F(\ker(T))\leq\dim_F\left(\ker\left(T|_{\ker(p)}\right)\right)+\dim_F(\ker(S)).$$

证明: 由 $p\circ T=S\circ p$ 知, $\ker(p)\subseteq V$ 为 T-不变子空间. 记 $\overline{T}$: $V/\ker(p)\longrightarrow V/\ker(p)$ 为诱导映射, 则

$$\alpha+\ker(p)\longmapsto T(\alpha)+\ker(p)$$

由线性代数基本定理知, 存在交换图 $\begin{array}{ccc} V/\ker(p) & \xrightarrow{\overline{T}} & V/\ker(p) \\ \downarrow\overline{p} & & \downarrow\overline{p} \\ W & \xrightarrow{S} & W \end{array}$, 其中 $\overline{p}$: $V/\ker(p)\longrightarrow W$ 为线性同构, 故

$$\alpha+\ker(p)\longmapsto p(\alpha)$$

$\dim_F(\ker(\overline{T}))=\dim_F(\ker(S))$. 因此由上述引理即知结论. □

推论 6.4.13 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$, $W_1,\cdots,W_k\subseteq V$ 为 T-不变子空间, 满足 $V=\sum\limits_{i=1}^{k}W_i$, 则 $\forall f(X)\in F[X]$, $\dim_F(\ker(f(T)))\leq\sum\limits_{i=1}^{k}\dim_F(\ker(f(T))\cap W_i)$.*

证明: 任取 $f(X)\in F[X]$, 由于 T-不变子空间也是 $f(T)$-不变子空间, 故通过将 T 换成 $f(T)$, 可不妨设 $f(X)=1$.

现考虑线性空间的外直和 $\coprod\limits_{i=1}^{k}W_i$, 则存在 p: $\coprod\limits_{i=1}^{k}W_i\longrightarrow\sum\limits_{i=1}^{k}W_i$ 为线性满射, 以及 $\coprod\limits_{i=1}^{k}T\big|_{W_i}\in L\left(\coprod\limits_{i=1}^{k}W_i\right)$,

$$(\alpha_1,\cdots,\alpha_k)\longmapsto\sum_{i=1}^{k}\alpha_i$$

满足 $p\circ\coprod\limits_{i=1}^{k}T\big|_{W_i}=T\circ p$, 故由上述推论知, $\dim_F(\ker(T))\leq\dim_F\left(\ker\left(\coprod\limits_{i=1}^{k}T\big|_{W_i}\right)\right)=\sum\limits_{i=1}^{k}\dim_F(\ker(T)\cap W_i)$. □

注: 在此推论的条件下, 显然 $\forall f(X)\in F[X]$, $\mathrm{Im}(f(T))=\sum\limits_{i=1}^{k}\mathrm{Im}\left(f(T)\big|_{W_i}\right)=\sum\limits_{i=1}^{k}\mathrm{Im}(f(T))\cap W_i$.

记 $R=F[X]$. 回忆在上述推论的基础上, 我们已经看到 V 作为 R-模的最少生成元数量恰好等于不变因子的个数. 以下我们将从对偶的观点再次审视这一非平凡的事实.

引理 6.4.14 *设 V 为域 F 上的线性空间, $W\subseteq V$ 为线性子空间, $T\in L(V)$, 则*

(1) *W 包含的最大 T-不变子空间为 $\bigcap\limits_{i\in\mathbb{N}}T^{-i}(W)=\bigcap\limits_{i=0}^{\deg(p_T(X))-1}T^{-i}(W)$;*

(2) *包含 W 的最小 T-不变子空间为 $\sum\limits_{i\in\mathbb{N}}T^{i}(W)=\sum\limits_{i=0}^{\deg(p_T(X))-1}T^{i}(W)$.*

证明: 显然. □

推论 6.4.15 *设 V 为域 F 上的线性空间, $W\subseteq V$ 为线性子空间, $T\in L(V)$.*

(1) *设 $W'\subseteq W$, 则 W' 是 W 包含的最大 T-不变子空间 $\Longleftrightarrow$ $(W')^0$ 是包含 W^0 的最小 T^t-不变子空间;*

(2) *设 $W'\supseteq W$, 则 W' 是包含 W 的最小 T-不变子空间 $\Longleftrightarrow$ $(W')^0$ 是 W^0 包含的最大 T^t-不变子空间.*

证明: (1) 注意 $\left(\bigcap\limits_{i\in\mathbb{N}}T^{-i}(W)\right)^0=\sum\limits_{i\in\mathbb{N}}(T^{-i}(W))^0=\bigcap\limits_{i\in\mathbb{N}}(T^t)^i(W^0)$, 其中交与和由上述引理知都是有限的.

(2) 完全同理. □

推论 6.4.16 *设 V 为域 F 上的有限维线性空间, $T\in L(V)$. 记 T 的不变因子个数为 r, 则*

(1) *$\min\{d\in\mathbb{N}\colon V\text{的任意}d\text{维子空间均包含非零的}T\text{-不变子空间}\}=\dim_F(V)-r+1$;*

(2) *$\max\{d\in\mathbb{N}\colon V\text{的任意}d\text{维子空间均包含于真的}T\text{-不变子空间}\}=r-1$.*

证明: (1)

$$\begin{aligned}
&\min\{d\in\mathbb{N}\colon V\text{的任意}d\text{维子空间均包含非零的}T\text{-不变子空间}\}\\
=&\min\{d\in\mathbb{N}\colon\forall W\subseteq V\,(\dim_F(W)=d),W\text{包含的最大}T\text{-不变子空间}\neq\{0\}\}\\
=&\dim_F(V)-\max\{d'\in\mathbb{N}\colon\forall M\subseteq V^*\,(\dim_F(M)=d'),\text{包含}M\text{的最小}T^t\text{-不变子空间}\neq V^*\}\\
=&\dim_F(V)-(\min\{d'\in\mathbb{N}\colon\exists M\subseteq V^*\,(\dim_F(M)=d'),\text{包含}M\text{的最小}T^t\text{-不变子空间}=V^*\}-1)\\
=&\dim_F(V)-(V^*\text{作为}R\text{-模的最少生成元数量}-1)\\
=&\dim_F(V)-(T^t\text{的不变因子个数}-1)\\
=&\dim_F(V)-(T\text{的不变因子个数}-1).
\end{aligned}$$

(2)

$$\begin{aligned}
&\max\{d \in \mathbb{N}\colon V\text{的任意}d\text{维子空间均包含于真的}T\text{-不变子空间}\}\\
=&\max\{d \in \mathbb{N}\colon \forall W \subseteq V\,(\dim_F(W) = d), \text{包含}W\text{的最小}T\text{-不变子空间} \neq V\}\\
=&\min\{d \in \mathbb{N}\colon \exists W \subseteq V\,(\dim_F(W) = d), \text{包含}W\text{的最小}T\text{-不变子空间} = V\} - 1\\
=&T\text{的不变因子个数} - 1.
\end{aligned}$$

□

参考文献与补注 6.4

(1) 关于同时上三角化的充分条件的部分, 可以参考 H. Radjavi, P. Rosenthal, “Simultaneous Triangularization”.

(2) 关于同调论中蛇引理的部分, 可以参考 J. J. Rotman“An Introduction to Homological Algebra”.

(3) 关于不变子空间的对偶性的部分, 可以参考 I. Gohberg, P. Lancaster, L. Rodman “Invariant Subspaces of Matrices with Applications”.

第 7 章　内积空间与正规算子

本章初步介绍实复线性空间上的几何结构: 内积 (inner product) 与关于内积的伴随算子 (adjoint operator). 它们将赋予线性映射更多有趣而实用的性质, 以及将人们引向泛函分析中的谱理论.

§7.1　内积与内积空间

7.1.1　实复内积的联系

设 $F=\mathbb{R}$ 或 $\mathbb{C}$, V 为域 F 上的线性空间. 回忆 V 上的一个内积是指一个二元函数 $V\times V\to F$, 满足以下性质: (1) 关于第一分量线性; (2) Hermite 对称性; (3) 正定性. 特别地, 它关于第二分量是共轭线性的 (即保持加法但数乘作用为共轭).

以下命题将实复内积与线性空间的实形式或复结构联系起来. 它们的证明是定义的直接验证, 但其中蕴含的观点是不可忽略的.

命题 7.1.1 (实形式) *设 V 为域 $\mathbb{C}$ 上的线性空间.*

(1) *若 V 上存在共轭 σ (即 $\sigma\colon V\to V$ 为共轭线性变换且 $\sigma^2=\mathrm{id}_V$), 则 $V_1\coloneqq\ker(\sigma-\mathrm{id}_V)$, $V_{-1}\coloneqq\ker(\sigma+\mathrm{id}_V)\subseteq V$ 均为实线性子空间, 满足 $V_1\xrightarrow{\cong}V_{-1}$ $\alpha\longmapsto\sqrt{-1}\alpha$ 为线性同构, 且 $V=V_1\oplus V_{-1}=V_1\oplus\sqrt{-1}V_1$. 特别地, 此时任取 V 上的一个复内积 $\langle\cdot,\cdot\rangle_V$, 则 $\langle\cdot,\cdot\rangle_V\big|_{V_1\times V_1}$ 是 V_1 上的实内积 $\iff\langle\sigma(\alpha_1),\sigma(\alpha_2)\rangle_V=\langle\alpha_2,\alpha_1\rangle_V$, $\forall\alpha_1,\alpha_2\in V$.*

(2) *反之, 若存在 $W\subseteq V$ 为实线性子空间, 满足 $V=W\oplus\sqrt{-1}W$, 则可定义 σ: $V\longrightarrow V$, $\alpha+\sqrt{-1}\beta\longmapsto\alpha-\sqrt{-1}\beta$ 为共轭, 且 $W=\ker(\sigma-\mathrm{id}_V)$. 特别地, 此时 W 上的任意实内积可唯一地延拓为 V 上的复内积. 具体地说, 记 $\langle\cdot,\cdot\rangle_W$ 为 W 上的一个实内积, 则可定义 V 上的一个复内积 $\langle\cdot,\cdot\rangle_V\colon V\times V\to\mathbb{C}$ 如下:*

$$\langle\alpha_1,\alpha_2\rangle_V\coloneqq\langle\Re(\alpha_1),\Re(\alpha_2)\rangle_W-\sqrt{-1}\,\langle\Re(\alpha_1),\Im(\alpha_2)\rangle_W+\sqrt{-1}\,\langle\Im(\alpha_1),\Re(\alpha_2)\rangle_W+\langle\Im(\alpha_1),\Im(\alpha_2)\rangle_W,$$

满足 $\langle\cdot,\cdot\rangle_V\big|_{W\times W}=\langle\cdot,\cdot\rangle_W$, 且 $\langle\sigma(\alpha_1),\sigma(\alpha_2)\rangle_V=\langle\alpha_2,\alpha_1\rangle_V$, $\forall\alpha_1,\alpha_2\in V$.

注: 抽象地说, 考虑从实线性空间范畴到复线性空间范畴的复化函子 $-\otimes_{\mathbb{R}}\mathbb{C}\colon\mathbb{R}\text{-}\mathbf{Mod}\to\mathbb{C}\text{-}\mathbf{Mod}$, 任意给定 $V\in\mathrm{obj}(\mathbb{C}\text{-}\mathbf{Mod})$, 则它在该函子下的纤维 (即实形式) 与其上的共轭一一对应. 进一步地, V 上关于某个共轭反变的复内积与相应实形式上的实内积一一对应.

命题 7.1.2 (复结构)

(1) *设 V 为域 $\mathbb{R}$ 上的线性空间. 若存在 $J\in L(V)$ 满足 $J^2=-\mathrm{id}_V$, 则 V 上可赋予 $\mathbb{C}$ 的数乘作用 $\mathbb{C}\times V\longrightarrow V$, $(c,\alpha)\longmapsto(\Re(c)+\Im(c)J)(\alpha)$, 从而成为复线性空间. 特别地, 若 $\langle\cdot,\cdot\rangle_{\mathbb{R}}$ 为 V 上的一个实内积, 满足 $\langle J(\alpha),J(\beta)\rangle_{\mathbb{R}}=\langle\alpha,\beta\rangle_{\mathbb{R}}$, $\forall\alpha,\beta\in V$, 则可唯一确定 V 上的一个复内积 $\langle\cdot,\cdot\rangle_{\mathbb{C}}$: $V\times V\longrightarrow\mathbb{C}$, $(\alpha_1,\alpha_2)\longmapsto\langle\alpha_1,\alpha_2\rangle_{\mathbb{R}}+\sqrt{-1}\,\langle\alpha_1,J(\alpha_2)\rangle_{\mathbb{R}}$ 满足 $\Re(\langle\cdot,\cdot\rangle_{\mathbb{C}})=\langle\cdot,\cdot\rangle_{\mathbb{R}}$.*

(2) *反之, 设 V 为域 $\mathbb{C}$ 上的线性空间. 若忘记 $\sqrt{-1}$ 的数乘作用而保留 $\mathbb{R}$ 的数乘作用, 则 V 可视为实线性空间, 且 $\dim_{\mathbb{R}}V=2\dim_{\mathbb{C}}V$. 取 V 的 $\mathbb{R}$-基 $\{\alpha_i\}_{i=1}^{2n}$, 记 J: $V\longrightarrow V$ $(1\leqslant i\leqslant n)$, $\alpha_i\longmapsto\alpha_{n+i}$, $\alpha_{n+i}\longmapsto-\alpha_i$, 则 $J\in L(V)$ 且 $J^2=-\mathrm{id}_V$. 特别地, 此时 V 上的任意复内积的实部都是 V 上的一个实内积, 且它关于 J 是不变的.*

注: 抽象地说, 考虑从复线性空间范畴到实线性空间范畴的忘记函子 $\mathscr{G}\colon \mathbb{C}\text{-}\mathbf{Mod} \to \mathbb{R}\text{-}\mathbf{Mod}$, 任意给定 $W \in \mathrm{obj}(\mathbb{R}\text{-}\mathbf{Mod})$, 则它在该函子下的纤维 (即复结构) 与其上的平方 -1 变换一一对应. 进一步地, W 上关于某个平方 -1 变换不变的实内积与相应复结构上的复内积一一对应.

更多地, 上述命题中涉及的两个函子实际上是伴随的: $-\underset{\mathbb{R}}{\otimes}\mathbb{C}\colon \mathbb{R}\text{-}\mathbf{Mod} \rightleftarrows \mathbb{C}\text{-}\mathbf{Mod}\colon \mathscr{G}$. 类似地, 根据域 $\mathbb{R}$ 上的有限维可除结合代数分类, 我们还可以给出这样的一对伴随函子: $-\underset{\mathbb{C}}{\otimes}\mathbb{H}\colon \mathbb{C}\text{-}\mathbf{Mod} \rightleftarrows \mathbb{H}\text{-}\mathbf{Mod}\colon \mathscr{G}$. 一个自然的问题是: 上述命题能否类比到 $\mathbb{C}$ 与 $\mathbb{H}$ 的情形?

7.1.2 内积与范数

内积的几何意义在于它可诱导线性空间上的长度、夹角等概念, 从而给线性空间赋予拓扑结构. 为一般起见, 我们先引入一般线性空间上的 (半) 范数, 再说明它与内积的联系.

定义 7.1.1 ((半) 范数) 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, V 为域 F 上的线性空间, $p\colon V \to \mathbb{R}$ 为一个函数, 若 p 满足

(1) (次可加性) $\forall \alpha, \beta \in V$, $p(\alpha+\beta) \leq p(\alpha) + p(\beta)$;

(2) (绝对齐次性) $\forall c \in F$, $\forall \alpha \in V$, $p(c\alpha) = |c|p(\alpha)$;

(3) (非负性) $\forall \alpha \in V$, $p(\alpha) \geq 0$;

则称 p 为 V 上的一个**半范** (seminorm). 若 V 上的一个半范 p 还满足

(4) (正定性) $\forall \alpha \in V$ $(p(\alpha) = 0 \Rightarrow \alpha = 0)$,

则称 p 为 V 上的一个**范数** (norm), 或**向量范数** (vector norm).

注: 在 (半) 范数的定义中, 非负性条件 (3) 可由 (1),(2) 推出: 事实上, 任取 $\alpha \in V$, 由 (2) 知 $p(0) = 0$ 且 $p(\alpha) = p(-\alpha)$, 则由 (3) 知 $p(\alpha) = \dfrac{p(\alpha)+p(-\alpha)}{2} \geq \dfrac{p(0)}{2} = 0$. 但正定性条件 (4) 不可由 (1),(2) 推出: 例如 $p \equiv 0$.

引理 7.1.3 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, (V,p) 是域 F 上的赋范线性空间, 则以下条件等价:*

(1) *存在 V 上的内积 $\langle\cdot,\cdot\rangle$, 满足 $p(\alpha) = \sqrt{\langle\alpha,\alpha\rangle}$, $\forall \alpha \in V$;*

(2) *(平行四边形等式) $\forall \alpha, \beta \in V$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 = 2p(\alpha)^2 + 2p(\beta)^2$;*

(3) *(平行四边形不等式 I) $\forall \alpha, \beta \in V$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 \leq 2p(\alpha)^2 + 2p(\beta)^2$;*

(4) *(平行四边形不等式 II) $\forall \alpha, \beta \in V$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 \geq 2p(\alpha)^2 + 2p(\beta)^2$;*

证明: "(1)⇒(2)": 由内积的一次半线性性直接验证即可;

"(2)⇒(1)": 由极化恒等式定义内积, 再验证条件即可;

"(2)⇔(3)⇔(4)": 将 α 替换为 $\dfrac{\alpha+\beta}{2}$, β 替换为 $\dfrac{\alpha-\beta}{2}$ 即可. □

推论 7.1.4 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, (V,p) 是域 F 上的赋范线性空间, 则以下条件等价:*

(1) *存在 V 上的内积 $\langle\cdot,\cdot\rangle_V$, 满足 $p(\alpha) = \sqrt{\langle\alpha,\alpha\rangle_V}$, $\forall \alpha \in V$;*

(2) *任取 V 的二维子空间 W, 都存在 W 上的内积 $\langle\cdot,\cdot\rangle_W$, 满足 $p(\alpha) = \sqrt{\langle\alpha,\alpha\rangle_W}$, $\forall \alpha \in W$.*

注: 此推论的意义在于将整体的内积条件约化为局部的内积条件, 从而可以利用平面几何给出更多的内积判别法.

引理 7.1.5 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, (V,p) 是域 F 上的赋范线性空间, 则以下条件等价:*

(1) *存在 V 上的内积 $\langle\cdot,\cdot\rangle$, 满足 $p(\alpha) = \sqrt{\langle\alpha,\alpha\rangle}$, $\forall \alpha \in V$;*

(2) *$\forall \alpha, \beta \in V$ $(p(\alpha) = p(\beta) = 1)$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 = 4$;*

(3) *$\forall \alpha, \beta \in V$ $(p(\alpha) = p(\beta) = 1)$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 \leq 4$;*

(4) *$\forall \alpha, \beta \in V$ $(p(\alpha) = p(\beta) = 1)$, $p(\alpha+\beta)^2 + p(\alpha-\beta)^2 \geq 4$;*

(5) *(Ptolemy 不等式) $\forall \alpha, \beta, \gamma \in V$, $p(\alpha-\gamma)p(\beta) + p(\beta-\gamma)p(\alpha) \geq p(\alpha-\beta)p(\gamma)$.*

证明: 这是上述推论与平面几何的结果, 可以参考 P. Jordan, J. von Neumann"On inner products in linear metric spaces"(1935); M. M. Day"Some characterizations of inner-product spaces"(1947). □

关于内积 $\langle\cdot,\cdot\rangle$ 与其诱导的长度 $\|\cdot\|$, 最基本的不等式是以下的 Cauchy-Schwarz 不等式. 我们回忆一个初等的证明方法:

定理 7.1.6 (Cauchy-Schwarz 不等式) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, 则*

$$\forall\alpha,\beta\in V,\ |\langle\alpha,\beta\rangle|\le\|\alpha\|\cdot\|\beta\|,$$

且 "=" 取到当且仅当 α,β 线性相关.

证明: 固定 $c\in F$ 满足 $|c|=1$ 且 $c\cdot\langle\alpha,\beta\rangle=|\langle\alpha,\beta\rangle|$. 由内积的半正定性知, $\forall t\in\mathbb{R}$, $\langle tc\alpha-\beta,tc\alpha-\beta\rangle\ge0$; 再由一次半线性性与 Hermite 对称性展开知 $\langle\alpha,\alpha\rangle\cdot t^2-2|\langle\alpha,\beta\rangle|\cdot t+\langle\beta,\beta\rangle\ge0$. 若 $\langle\alpha,\alpha\rangle>0$, 则由二次方程的判别法知, $|\langle\alpha,\beta\rangle|^2\le\langle\alpha,\alpha\rangle\cdot\langle\beta,\beta\rangle$, 且 "=" 取到当且仅当 $\exists t\in\mathbb{R}$, $s.t.\ tc\alpha-\beta=0$. 若 $\langle\alpha,\alpha\rangle=0$, 则由一次函数的性质知 $\langle\alpha,\beta\rangle=0$, 此时 "=" 也取到, 且由内积的正定性知 $\alpha=0$. □

注:

(1) 事实上, 由上述证明可知, 对于实复线性空间上 Hermite 对称的半正定一次半线性形式, Cauchy-Schwarz 不等式仍成立, 但此时向量线性相关只是 "=" 取到的充分不必要条件.

(2) 警告: 对于一般的正定一次半线性形式, Cauchy-Schwarz 不等式未必成立. 一个著名的反例如下:

例 7.1.1 (Cauchy-Schwarz 不等式的反例) *设 $\mathbb{R}^n$ 上标准内积 $\langle\cdot,\cdot\rangle$ 诱导的长度为 $\|\cdot\|$. 固定 $K\in\mathbb{R}^{n\times n}\setminus\{0\}$ 为反对称阵, 以及 $\beta_0\in\mathbb{R}^n$, 满足 $\alpha_0:=-K\beta_0\ne0$. 任取 $\lambda\in\left(0,\dfrac{\|\alpha_0\|}{\|\beta_0\|}\right)$, 记 $A_\lambda=\lambda I_n-K\in F^{n\times n}$, $\varphi_\lambda\colon\ \mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}$, $(\alpha,\beta)\longmapsto\langle\alpha,A_\lambda\beta\rangle$*

证明:

(1) *φ_λ 是双线性函数, 且 $\forall\alpha\in\mathbb{R}^n\setminus\{0\}$, $\varphi_\lambda(\alpha,\alpha)>0$;*

(2) *$\varphi_\lambda(\alpha_0,\beta_0)>\sqrt{\varphi_\lambda(\alpha_0,\alpha_0)\cdot\varphi_\lambda(\beta_0,\beta_0)}$.*

证明: (1)φ_λ 的双线性性显然; 现任取 $\alpha\in\mathbb{R}^n\setminus\{0\}$, 则 $\varphi_\lambda(\alpha,\alpha)=\langle\alpha,A_\lambda\alpha\rangle=\lambda\|\alpha\|^2-\langle\alpha,K\alpha\rangle=\lambda\|\alpha\|^2>0$.
(2) $\varphi_\lambda(\alpha_0,\beta_0)=\lambda\langle\alpha_0,\beta_0\rangle-\langle\alpha_0,K\beta_0\rangle=-\lambda\langle K\beta_0,\beta_0\rangle+\|\alpha_0\|^2=\|\alpha_0\|^2>\sqrt{\lambda\|\alpha_0\|^2\cdot\lambda\|\beta_0\|^2}=\sqrt{\varphi_\lambda(\alpha_0,\alpha_0)\cdot\varphi_\lambda(\beta_0,\beta_0)}$.
□

注: 事实上, 由上述证明可知 $\lim\limits_{\lambda\to0^+}\dfrac{\varphi_\lambda(\alpha_0,\beta_0)}{\sqrt{\varphi_\lambda(\alpha_0,\alpha_0)\cdot\varphi_\lambda(\beta_0,\beta_0)}}=\lim\limits_{\lambda\to0^+}\dfrac{\|\alpha_0\|}{\lambda\|\beta_0\|}=+\infty$, 即使 $\{A_\lambda\}_{\lambda\to0^+}\subseteq\mathbb{R}^{n\times n}$ 在标准内积下是一致有界的.

以下我们考虑线性空间上由内积或范数诱导的拓扑结构. 具体地说, 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, (V,p) 是域 F 上的赋范线性空间. 记 $B_r(\alpha):=\{\beta\in V\colon p(\beta-\alpha)<r\}\,(\alpha\in V,r>0)$, 则 $\mathscr{B}:=\{B_r(\alpha)\colon\alpha\in V,r>0\}$ 可以生成 V 上的一个拓扑, 并成为该拓扑的拓扑基. 此拓扑称为线性空间上由范数诱导的拓扑. 类似地, 线性空间上的内积可给出范数从而诱导拓扑, 称为线性空间上由内积诱导的拓扑.

引理 7.1.7 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, (V,p) 是域 F 上的赋范线性空间, 则 $+\colon V\times V\to V$, $\cdot\colon F\times V\to V$, 以及 $p\colon V\to\mathbb{R}$ 都是连续映射.*

证明: 先证明 $+\colon V\times V\to V$ 连续, 由拓扑基知只需证明: $\forall B_\epsilon(\beta)\subseteq V$, $\forall\alpha_1,\alpha_2\in V\,(\alpha_1+\alpha_2\in B_\epsilon(\beta))$, $\exists\delta>0$, $s.t.\ B_\delta(\alpha_1)+B_\delta(\alpha_2)\subseteq B_\epsilon(\beta)$. 事实上, 由范数的次可加性, 取 $\delta=\dfrac{\epsilon-p(\alpha_1+\alpha_2-\beta)}{2}$ 即可.

再证明 $\cdot\colon F\times V\to V$ 连续, 由拓扑基知只需证明: $\forall B_\epsilon(\beta)\subseteq V$, $\forall c\in F,\alpha\in V\,(c\cdot\alpha\in B_\epsilon(\beta))$, $\exists\delta>0$, $s.t.\ B_\delta(c)\cdot B_\delta(\alpha)\subseteq B_\epsilon(\beta)$. 事实上, 由范数的次可加性, 取 $\delta>0$ 满足 $\delta(c+\delta+p(\alpha))=\epsilon-p(c\cdot\alpha-\beta)$ 即可.

最后, 为证明 $p\colon V\to\mathbb{R}$ 连续, 只需证明 $\forall(c-\epsilon,c+\epsilon)\subseteq\mathbb{R}$, $\forall\alpha\in V\,(p(\alpha)\in(c-\epsilon,c+\epsilon))$, $\exists\delta>0$, $s.t.\ p(B_\delta(\alpha))\subseteq(c-\epsilon,c+\epsilon)$. 事实上, 由范数的次可加性, 取 $\delta=\epsilon-|p(\alpha)-c|$ 即可. □

引理 7.1.8 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 是域 F 上的内积空间, 则 $+\colon V\times V\to V$, $\cdot\colon F\times V\to V$, 以及 $\langle\cdot,\cdot\rangle:V\times V\to F$ 都是连续映射.*

证明: 加法与数乘的连续性由上述引理即知. 为证明 $\langle\cdot,\cdot\rangle:V\times V\to F$ 连续, 只需证明 $\forall B_\epsilon(c)\subseteq F$, $\forall\alpha_1,\alpha_2\in V\,(\langle\alpha_1,\alpha_2\rangle\in B_\epsilon(c))$, $\exists\delta>0$, $s.t.\ \langle B_\delta(\alpha_1),B_\delta(\alpha_2)\rangle\subseteq B_\epsilon(c)$. 事实上, 由 Cauchy-Schwarz 不等式知, 取 $\delta>0$ 满足

$\delta(\delta+\|\alpha_1\|+\|\alpha_2\|)=\epsilon-|\langle\alpha_1,\alpha_2\rangle-c|$ 即可. □

注: 上述两个引理说明, 域 $\mathbb{R}$ 或 $\mathbb{C}$ 上的赋范线性空间或内积空间都是 **拓扑线性空间** (topological vector space), 即带有拓扑结构的线性空间, 且满足加法与数乘均为连续映射.

最后我们讨论内积空间中的正交性质. 一个平凡的观察是 Pythagorean 定理, 即有限个相互正交的向量的长度平方和等于向量和的长度平方. 进一步地, 以单位正交的向量组为骨架, 我们可以得到一般向量的长度与各方向上投影长度的关系.

命题 7.1.9 (Bessel 不等式) *设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $\{\alpha_i\}_{i\in I}$ 为标准正交集, 则 $\forall\beta\in V$, $\|\beta\|^2\geq\sum\limits_{i\in I}|\langle\beta,\alpha_i\rangle|^2$, 当且仅当 $\beta\in\overline{\mathrm{Span}_F(\{\alpha_i\}_{i\in I})}$ 时取等号.*

证明: 任取 $\beta\in V$. 由 $\sum\limits_{i\in I}|\langle\beta,\alpha_i\rangle|^2\coloneqq\lim\limits_{\substack{J\subseteq I\\|J|<+\infty}}\sum\limits_{j\in J}|\langle\beta,\alpha_j\rangle|^2$, 只需证明 $\forall J\subseteq I\,(|J|<+\infty)$, $\|\beta\|^2\geq\sum\limits_{j\in J}|\langle\beta,\alpha_j\rangle|^2$. 事实上, 由内积的一次半线性性与 Hermite 性展开知, $0\leq\|\beta-\sum\limits_{j\in J}\langle\beta,\alpha_j\rangle\,\alpha_j\|^2=\|\beta\|^2-\sum\limits_{j\in J}|\langle\beta,\alpha_j\rangle|^2$. 特别地, 由 $\sum\limits_{i\in I}|\langle\beta,\alpha_i\rangle|^2<+\infty$ 知, $|\{i\in I\colon\langle\beta,\alpha_i\rangle\neq0\}|$ 为至多可数集. 由此可知等号取到的条件为 $\beta\in\overline{\mathrm{Span}_F(\{\alpha_i\}_{i\in I})}$. □

推论 7.1.10 (Parseval 等式) *设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $\{\alpha_i\}_{i\in I}$ 为标准正交集, 满足 $V=\overline{\mathrm{Span}_F(\{\alpha_i\}_{i\in I})}$, 则 $\forall\beta\in V$, $\|\beta\|^2=\sum\limits_{i\in I}|\langle\beta,\alpha_i\rangle|^2$.*

例 7.1.2 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $V=C^0([0,1];F)$ 为 $[0,1]\subseteq\mathbb{R}$ 上的所有 F-值连续函数构成的线性空间, 其中内积定义为 $\langle\cdot,\cdot\rangle_{L^2}\colon V\times V\longrightarrow F$ $\quad(f,g)\longmapsto\int_0^1 f(t)\cdot\overline{g(t)}\mathrm{d}t$. 当 $F=\mathbb{R}$ 时, $\{t\mapsto\sqrt{2}\sin(2\pi kt)\}_{k\geq1}\cup\{t\mapsto\sqrt{2}\cos(2\pi kt)\}_{k\geq0}$ 是 $(V,\langle\cdot,\cdot\rangle)$ 中的标准正交集; 当 $F=\mathbb{C}$ 时, $\{t\mapsto e^{2\pi\sqrt{-1}kt}\}_{k\in\mathbb{Z}}$ 是 $(V,\langle\cdot,\cdot\rangle)$ 中的标准正交集. 进一步地, 由 Fourier 分析可知, 它们生成的线性子空间都是 V 的稠子空间. 因此由 Parseval 等式可知, $\forall f\in C^0([0,1];F)$,*

$$\int_0^1|f(t)|^2\mathrm{d}t=\sum_{k\in\mathbb{Z}}\left|\int_0^1 f(t)\cdot e^{-2\pi\sqrt{-1}kt}\mathrm{d}t\right|^2.$$

特别地, 这给出了等距线性映射 $(C^0([0,1];F),\langle\cdot,\cdot\rangle_{L^2})\longrightarrow(\ell^2(\mathbb{Z}),\langle\cdot,\cdot\rangle_{\ell^2})$ $\quad f\longmapsto\left(\int_0^1 f(t)\cdot e^{-2\pi\sqrt{-1}kt}\mathrm{d}t\right)_{k\in\mathbb{Z}}$. 然而此映射并不是满射: 例如取 $c_k=\begin{cases}\dfrac{1}{2}, & k=0\\ \dfrac{(-1)^k-1}{2\pi\sqrt{-1}k}, & k\in\mathbb{Z}\setminus\{0\}\end{cases}$, 则 $(c_k)_{k\in\mathbb{Z}}\in\ell^2(\mathbb{Z})$ 不在此映射的像中. 事实上, 这是因为 $(C^0([0,1];F),\langle\cdot,\cdot\rangle_{L^2})$ 不是完备的内积空间, 而它的完备化空间为 $(L^2([0,1];F),\langle\cdot,\cdot\rangle_{L^2})$, 此时上述映射延拓为等距线性同构 $(L^2([0,1];F),\langle\cdot,\cdot\rangle_{L^2})\xrightarrow{\cong}(\ell^2(\mathbb{Z}),\langle\cdot,\cdot\rangle_{\ell^2})$ $\quad f\longmapsto\left(\int_0^1 f(t)\cdot e^{-2\pi\sqrt{-1}kt}\mathrm{d}t\right)_{k\in\mathbb{Z}}$.

7.1.3 完备的内积空间

本小节研究拓扑线性空间的完备性, 并由此将有限维内积空间的许多几何性质推广至 Hilbert 空间. 为一般起见, 我们先讨论内积空间中线性子空间及其正交补的基本性质.

引理 7.1.11 *设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $W\subseteq V$ 为线性子空间, 则 $W^\perp=\overline{W}^\perp$ 为闭线性子空间.*

证明: 由内积的连续性即知结论. □

引理 7.1.12 *设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $W_1,W_2\subseteq V$ 为线性子空间, 则 $W_1^\perp\cap W_2^\perp=(\overline{W_1}+\overline{W_2})^\perp$.*

证明: 由定义与内积的连续性直接验证即可. □

引理 7.1.13 (正交分解的唯一性) 设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $W_1,W_2\subseteq V$ 为线性子空间, 满足 $V=W_1\oplus W_2$, 且 $W_1\perp W_2$, 则 $W_1=W_2^\perp$; $W_2=W_1^\perp$. 特别地, W_1,W_2 均为闭线性子空间.

证明: 一方面, 显然 $W_1\subseteq W_2^\perp$; 另一方面, 任取 $\alpha\in W_2^\perp$. 由 $V=W_1\oplus W_2$ 知, $\exists\alpha_1\in W_1,\alpha_2\in W_2$, $s.t.$ $\alpha=\alpha_1+\alpha_2$. 又 $W_1\perp W_2$, 则 $0=\langle\alpha,\alpha_2\rangle=\langle\alpha_1+\alpha_2,\alpha_2\rangle=\langle\alpha_2,\alpha_2\rangle$, 故 $\alpha_2=0$, 即 $\alpha=\alpha_1\in W_1$. 因此 $W_1\supseteq W_2^\perp$. □

引理 7.1.14 (正交分解的存在性) 设 $(V,\langle\cdot,\cdot\rangle)$ 为内积空间, $W\subseteq V$ 为有限维线性子空间, 则 $V=W\oplus W^\perp$.

证明: 设 $W\subseteq V$ 的一组标准正交基为 $\{\alpha_i\}_{i=1}^m$, 则 $W^\perp=\bigcap\limits_{i=1}^m\ker(\langle\cdot,\alpha_i\rangle)$. 由内积的正定性即知 $W\cap W^\perp=\{0\}$. 另外, 由于 $\forall\alpha\in V$, $\alpha-\sum\limits_{i=1}^m\langle\alpha,\alpha_i\rangle\cdot\alpha_i\in W^\perp$, 则 $\alpha=\sum\limits_{i=1}^m\langle\alpha,\alpha_i\rangle\cdot\alpha_i+\left(\alpha-\sum\limits_{i=1}^m\langle\alpha,\alpha_i\rangle\cdot\alpha_i\right)\in W+W^\perp$, 故 $V=W+W^\perp$. □

注: 显然上述引理对于一般的线性子空间未必成立, 反例如取 $W\subsetneqq V$ 为稠线性子空间, 则 $W^\perp=V^\perp=\{0\}$. 进一步地, 即使 $W\subsetneqq V$ 为闭线性子空间, 也可能出现 $W^\perp=\{0\}$ 的情形, 见以下的反例.

例 7.1.3 考虑 Hilbert 空间 $(\ell^2,\langle\cdot,\cdot\rangle)$ 及其非完备的子内积空间 $(F^{(\omega)},\langle\cdot,\cdot\rangle)$, 记 f: $\begin{array}{ccc}\ell^2&\longrightarrow&F\\(a_n)_{n\geq0}&\longmapsto&\sum\limits_{n\geq0}\dfrac{a_n}{n+1}\end{array}$.

由 Cauchy-Schwarz 不等式知, f 是定义良好的连续线性函数. 现记 $W\coloneqq\ker(f)\cap F^{(\omega)}$, 则 $W\subsetneqq F^{(\omega)}$ 为闭线性子空间. 断言: $W^\perp=\{0\}$(这是在 $F^{(\omega)}$ 中取正交补). (事实上, 假设 $(a_n)_{n\geq0}\in W^\perp$, 由于 $\forall n\geq1$, $\epsilon_0-(n+1)\epsilon_n\in W$, 则 $\forall n\geq1$, $a_0-(n+1)a_n=0$. 但 $|\{n\geq0\colon a_n\neq0\}|<+\infty$, 则 $(a_n)_{n\geq0}=0$.) 因此在 $F^{(\omega)}$ 中 $W\subsetneqq(W^\perp)^\perp$.

定义 7.1.2 (完备性) 设 V 为拓扑线性空间, 若 V 中的任意 Cauchy 网均收敛, 则称 V 为**完备的** (complete). 完备的赋范线性空间称为 Banach 空间. 完备的内积空间称为 Hilbert 空间.

注:

(1) 一般拓扑线性空间的完备性需要考虑 "Cauchy 网 "的收敛性; 但当拓扑线性空间满足第一可数性时, 则只需要考虑通常 "Cauchy 序列" 的收敛性.

(2) 由定义可知, 一个拓扑线性空间是完备的当且仅当它的所有闭线性子空间是完备的. 于是在拓扑线性空间中, 我们更多考虑的是闭线性子空间或者稠线性子空间.

Hilbert 空间的一个重要特征是其中任意闭线性子空间必存在正交补. 为深入理解正交分解, 我们先讨论 Hilbert 空间中的最佳逼近性质.

命题 7.1.15 (最佳逼近) 设 $(V,\langle\cdot,\cdot\rangle)$ 为 *Hilbert* 空间, $\varnothing\neq K\subseteq V$ 为一个闭凸集, $\alpha\in V$, 则存在唯一的 $\beta\in K$, 满足 $\|\beta-\alpha\|=\inf\limits_{\beta'\in K}\|\beta'-\alpha\|$. 进一步地, β 可由性质 "$\beta\in K$ 且 $\mathfrak{R}(\langle\beta-\alpha,\beta-\beta'\rangle)\leq0$, $\forall\beta'\in K$" 唯一决定.

证明: 记 $d=\inf\limits_{\beta'\in K}\|\beta'-\alpha\|\geq0$, 则 $\exists\{\beta_n\}_{n=1}^{+\infty}\subseteq K$, $s.t.$ $d=\lim\limits_{n\to+\infty}\|\beta_n-\alpha\|$. 由于 K 为凸集, 则 $\dfrac{\beta_n+\beta_m}{2}\in K$, 故 $\left\|\dfrac{\beta_n+\beta_m}{2}-\alpha\right\|\geq d$. 考虑平行四边形等式

$$\|(\beta_n-\alpha)+(\beta_m-\alpha)\|^2+\|(\beta_n-\alpha)-(\beta_m-\alpha)\|^2=2\|\beta_n-\alpha\|^2+2\|\beta_m-\alpha\|^2,$$

则 $\|\beta_n-\beta_m\|^2=2(\|\beta_n-\alpha\|^2+\|\beta_m-\alpha\|^2)-4\left\|\dfrac{\beta_n+\beta_m}{2}-\alpha\right\|^2\leq2(\|\beta_n-\alpha\|^2+\|\beta_m-\alpha\|^2)-4d^2\to0$ $(m,n\to+\infty)$, 即 $\{\beta_n\}_{n=1}^{+\infty}$ 为 Cauchy 列. 由 $(V,\langle\cdot,\cdot\rangle)$ 的完备性知, $\lim\limits_{n\to+\infty}\beta_n=:\beta\in V$ 存在; 再由 K 的闭性知 $\beta\in K$. 此时 $\|\beta-\alpha\|=\lim\limits_{n\to+\infty}\|\beta_n-\alpha\|=d$ 满足要求. 假设 $\beta,\beta'\in K$ 均满足 $\|\beta-\alpha\|=\|\beta'-\alpha\|=d$, 则考虑 $\{\beta_n\}_{n=1}^{+\infty}\subseteq K$ 满足 $\beta_{2n-1}\coloneqq\beta$, $\beta_{2n}\coloneqq\beta'$, $\forall n\geq1$, 由以上过程即知 $\beta=\beta'$.

现设 $\beta\in K$ 满足 $\|\beta-\alpha\|=\inf\limits_{\beta'\in K}\|\beta'-\alpha\|$. 任取 $\beta'\in K$, 以及 $t\in(0,1]$, 由 K 的凸性知 $(1-t)\beta+t\beta'\in K$, 则 $\|((1-t)\beta+t\beta')-\alpha\|\geq\|\beta-\alpha\|$. 将不等式两端平方展开知, $t^2\|\beta-\beta'\|^2-2t\mathfrak{R}(\langle\beta-\alpha,\beta-\beta'\rangle)\geq0$. 约去 $t>0$ 再令 $t\to0^+$ 得, $-\mathfrak{R}(\langle\beta-\alpha,\beta-\beta'\rangle)\geq0$. 反之, 设 $\beta\in K$ 满足 $\mathfrak{R}(\langle\beta-\alpha,\beta-\beta'\rangle)\leq0$, $\forall\beta'\in K$, 则由内积的一次半线性展开知, $\forall\beta'\in K$, $\|\beta-\alpha\|^2-\|\beta'-\alpha\|^2=2\mathfrak{R}(\langle\beta-\alpha,\beta-\beta'\rangle)-\|\beta-\beta'\|^2\leq0$, 即 $\|\beta-\alpha\|=\inf\limits_{\beta'\in K}\|\beta'-\alpha\|$. □

注:

(1) 上述最佳逼近的性质对于一般的 Banach 空间未必成立. 事实上, 设 (V,p) 为 Banach 空间, 则 (V,p) 是自反的当且仅当 V 中任意点到任意非空闭凸子集都存在最佳逼近; (V,p) 是严格凸的当且仅当 V 中任意点到任意非空闭凸子集的最佳逼近若存在必唯一. 可以参考 R. E. Megginson"An introduction to Banach space theory".

(2) 设 $\varnothing \neq K \subseteq V$ 为一个闭凸集, 记 $P_K(\alpha) \in K$ 为 $\alpha \in V$ 到 K 的唯一最佳逼近, 则 $P_K : V \to K$ 是一个映射, 且 $P_K\big|_K = \mathrm{id}_K$. 事实上, P_K 还满足 $\|P_K(\alpha_1) - P_K(\alpha_2)\| \leq \|\alpha_1 - \alpha_2\|$, $\forall \alpha_1, \alpha_2 \in V$. 这是因为, 由最佳逼近点的等价刻画知, $\mathfrak{R}(\langle P_K(\alpha_1) - \alpha_1, P_K(\alpha_1) - P_K(\alpha_2)\rangle) \leq 0$; $\mathfrak{R}(\langle P_K(\alpha_2) - \alpha_2, P_K(\alpha_2) - P_K(\alpha_1)\rangle) \leq 0$. 两式相加得, $\|P_K(\alpha_1) - P_K(\alpha_2)\|^2 - \mathfrak{R}(\langle \alpha_1 - \alpha_2, P_K(\alpha_1) - P_K(\alpha_2)\rangle) \leq 0$, 再由 Cauchy-Schwarz 不等式即知 $\|P_K(\alpha_1) - P_K(\alpha_2)\| \leq \|\alpha_1 - \alpha_2\|$.

命题 7.1.16 (正交分解的存在性) *设 $(V, \langle\cdot,\cdot\rangle)$ 为 Hilbert 空间, $W \subseteq V$ 为线性子空间, 记 $P_{\overline{W}} : V \to \overline{W}$ 为最佳逼近映射, 则*

(1) *$P_{\overline{W}}$ 是一个线性映射, 且为投影;*

(2) *$\ker(P_{\overline{W}}) = \overline{W}^{\perp} = W^{\perp}$ 为闭线性子空间; $\mathrm{Im}(P_{\overline{W}}) = \overline{W}$;*

(3) *$V = \overline{W} \oplus W^{\perp}$, 且 $(W^{\perp})^{\perp} = \overline{W}$.*

证明: 显然 $\overline{W} \subseteq V$ 是线性子空间. 由最佳逼近点的等价刻画以及 $\overline{W}$ 的线性性知, 对于 $\alpha \in V$, $\beta = P_{\overline{W}}(\alpha)$ 可由性质 "$\beta \in \overline{W}$ 且 $(\beta - \alpha) \perp \overline{W}$" 唯一决定. 由此即知 $P_{\overline{W}}$ 是线性的, 且 $\ker(P_{\overline{W}}) = \overline{W}^{\perp}$, $\mathrm{Im}(P_{\overline{W}}) = \overline{W}$, 以及 $P_{\overline{W}}^2 = P_{\overline{W}}$. 由投影映射的性质知, $V = \mathrm{Im}(P_{\overline{W}}) \oplus \ker(P_{\overline{W}}) = \overline{W} \oplus W^{\perp}$. 再由正交补的唯一性即知 $\overline{W} = (W^{\perp})^{\perp}$. □

反之, 利用线性子空间的双重正交补性质也可刻画 Hilbert 空间:

引理 7.1.17 *设 $(V, \langle\cdot,\cdot\rangle)$ 为内积空间, $\{\alpha_n\}_{n=1}^{+\infty} \subseteq V$ 为 Cauchy 列, 则 $f : V \longrightarrow F$, $\beta \longmapsto \lim\limits_{n\to+\infty} \langle \beta, \alpha_n\rangle$ 为定义良好的连续线性函数. 进一步地, $f \equiv 0 \Longleftrightarrow \lim\limits_{n\to+\infty} \alpha_n = 0$.*

证明: 任取 $\beta \in V$. 由 $\{\alpha_n\}_{n=1}^{+\infty} \subseteq V$ 为 Cauchy 列以及 Cauchy-Schwarz 不等式知, $\{\langle \beta, \alpha_n\rangle\}_{n=1}^{+\infty} \subseteq F$ 也为 Cauchy 列. 由 $F = \mathbb{R}$ 或 $\mathbb{C}$ 的完备性知, $\lim\limits_{n\to+\infty} \langle \beta, \alpha_n\rangle \in F$ 存在, 因此 f 定义良好. 显然 f 为线性函数.

为证明 f 连续, 由线性性知只需证明 f 在 0 处连续. 由 Cauchy 列必有界知, $\exists M \geq 1$, $s.t.$ $\|\alpha_n\| \leq M$, $\forall n \geq 1$. 由 Cauchy-Schwarz 不等式知, $\forall \epsilon > 0$, $\exists \delta = \epsilon/2M > 0$, $\forall \beta \in B_\delta(0)$, $\forall n \geq 1$, $|\langle \beta, \alpha_n\rangle| \leq \|\beta\| \cdot \|\alpha_n\| < \delta \cdot M = \epsilon/2$, 故 $|f(\beta)| = \lim\limits_{n\to+\infty} |\langle \beta, \alpha_n\rangle| \leq \epsilon/2 < \epsilon$.

显然当 $\lim\limits_{n\to+\infty} \alpha_n = 0$ 时 $f \equiv 0$. 反之, 现设 $f \equiv 0$. 由 Cauchy 列的定义知, $\forall \epsilon > 0$, $\exists N_1 \geq 1$, $\forall n \geq N_1$, $\|\alpha_n - \alpha_{N_1}\| < \epsilon/2M$. 由条件知 $\lim\limits_{n\to+\infty} \langle \alpha_{N_1}, \alpha_n\rangle = 0$, 即 $\forall \epsilon > 0$, $\exists N_2 \geq 1$, $\forall n \geq N_2$, $|\langle \alpha_{N_1}, \alpha_n\rangle| < \epsilon/2$. 因此由 Cauchy-Schwarz 不等式知, $\forall n \geq \max\{N_1, N_2\}$,

$$\langle \alpha_n, \alpha_n\rangle = \langle \alpha_n - \alpha_{N_1}, \alpha_n\rangle + \langle \alpha_{N_1}, \alpha_n\rangle \leq \|\alpha_n - \alpha_{N_1}\| \cdot \|\alpha_n\| + |\langle \alpha_{N_1}, \alpha_n\rangle| < (\epsilon/2M) \cdot M + \epsilon/2 = \epsilon,$$

即 $\lim\limits_{n\to+\infty} \alpha_n = 0$. □

命题 7.1.18 *设 $(V, \langle\cdot,\cdot\rangle)$ 为内积空间, 若任意线性子空间 $W \subseteq V$ 均满足 $\overline{W} = (W^{\perp})^{\perp}$, 则 $(V, \langle\cdot,\cdot\rangle)$ 为 Hilbert 空间.*

证明: 任取 $(V, \langle\cdot,\cdot\rangle)$ 中的 Cauchy 列 $\{\alpha_n\}_{n=1}^{+\infty}$. 由引理知, $f : V \longrightarrow F$, $\beta \longmapsto \lim\limits_{n\to+\infty} \langle \beta, \alpha_n\rangle$ 为定义良好的连续线性函数, 则 $W = \ker(f) \subseteq V$ 为闭线性子空间. 若 $W = V$, 即 $f \equiv 0$, 故 $\lim\limits_{n\to+\infty} \alpha_n = 0$. 若 $W \subsetneqq V$, 则由 $W = (W^{\perp})^{\perp}$ 知, $W^{\perp} \neq \{0\}$. 取 $\alpha \in W^{\perp} \setminus \{0\}$, 此时 $g : V \longrightarrow F$, $\beta \longmapsto \langle \beta, \alpha\rangle$ 也为连续线性函数, 且 $W \subseteq \ker(g)$. 断言: $\exists c \in F \setminus \{0\}$, $s.t.$ $g = c \cdot f$. (这是因为, 由 $W \neq V$ 知, $f \not\equiv 0$. 取 $\beta_0 \in V$ 满足 $f(\beta_0) = 1$, 则 $\forall \beta \in V$, $\beta - f(\beta)\beta_0 \in \ker(f) \subseteq \ker(g)$, 即 $g(\beta) = f(\beta)g(\beta_0)$, 故 $g = g(\beta_0) \cdot f$. 又 $g(\alpha) = \langle \alpha, \alpha\rangle \neq 0$, 则 $g \not\equiv 0$.) 因此 $f - \dfrac{1}{c} g \equiv 0$. 再由引理知, $\lim\limits_{n\to+\infty} \alpha_n = \dfrac{1}{c}\alpha$. □

最后我们讨论 Hilbert 空间中 (闭) 线性子空间的交与和的关系; 特别地, 我们需要指出两个闭线性子空间的和未必是闭的.

推论 7.1.19 设 $(V,\langle\cdot,\cdot\rangle)$ 为 *Hilbert* 空间, $W_1,W_2\subseteq V$ 为线性子空间, 则

(1) $(W_1^\perp\cap W_2^\perp)^\perp=\overline{W_1+W_2}$;

(2) $\overline{W_1}\cap\overline{W_2}=(W_1^\perp+W_2^\perp)^\perp$;

(3) $(\overline{W_1}\cap\overline{W_2})^\perp=\overline{W_1^\perp+W_2^\perp}$.

证明: 由本节开头的引理与 $(W^\perp)^\perp=\overline{W}$ 直接验证即可. 特别地, 在一般内积空间中, 由满足 $W\subsetneqq(W^\perp)^\perp$ 的闭线性子空间可构造上述结论的反例. □

例 7.1.4 考虑 *Hilbert* 空间 $(\ell^2,\langle\cdot,\cdot\rangle)$, 记

$$W_1=\{(a_n)_{n\geq0}\in\ell^2\colon a_{2n}=0,\ \forall n\geq0\},$$

$$W_2=\{(a_n)_{n\geq0}\in\ell^2\colon a_{2n+1}=na_{2n},\ \forall n\geq0\},$$

则 $W_1,W_2\subseteq\ell^2$ 均为闭线性子空间, 且 $W_1\cap W_2=\{0\}$, $W_1+W_2\supseteq F^{(\omega)}$, 故 $W_1+W_2\subseteq\ell^2$ 为稠线性子空间. 但 $\sum\limits_{n\geq0}\dfrac{\epsilon_n}{n+1}\notin W_1+W_2$, 故 $W_1+W_2\neq\ell^2$.

7.1.4 内积空间的基与维数

鉴于内积空间中的长度与夹角概念, 我们常常希望将其中的一组线性无关向量转化为标准正交集, 且两者生成的子空间相同. 在至多可数维内积空间的情形, Gram-Schmidt 正交化过程保证了这一要求总可行, 于是我们可以由线性空间的维数完全分类至多可数维内积空间:

命题 7.1.20 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的至多可数维内积空间, 则 $(V,\langle\cdot,\cdot\rangle)$ 保内积同构于 $F^{(I)}$ 上的标准内积空间, 其中 $|I|=\dim_F(V)$. 特别地, 域 F 上的两个至多可数维内积空间保内积同构当且仅当它们的维数相等.

同理, 对于具有标准正交 (Hamel) 基的不可数 (Hamel) 维内积空间, 上述命题仍成立:

命题 7.1.21 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, 且具有一组标准正交 *(Hamel)* 基. 则 $(V,\langle\cdot,\cdot\rangle)$ 保内积同构于 $F^{(I)}$ 上的标准内积空间, 其中 $|I|=\dim_F(V)$. 特别地, 域 F 上的两个具有标准正交 *(Hamel)* 基的内积空间保内积同构当且仅当它们的 *(Hamel)* 维数相等.

然而, 存在大量的不可数 (Hamel) 维内积空间并不具有标准正交 (Hamel) 基, 也即 Gram-Schmidt 正交化过程对于不可数个线性无关向量一般不成立. 以下我们试图借助拓扑给出一些经典的反例.

引理 7.1.22 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的可分内积空间, 则 V 中任意标准正交集都是至多可数集.

证明: 任取 V 中的两个标准正交向量 α,β, 由 Pythagorean 定理知, $\|\alpha-\beta\|=\sqrt{\|\alpha\|^2+\|\beta\|^2}=\sqrt{2}$. 现设 $D\subseteq V$ 为可数稠密子集, $S\subseteq V$ 为标准正交集, 则 $\forall\alpha\in S$, $\exists\alpha'\in D$, $s.t.\ \|\alpha-\alpha'\|<\dfrac{\sqrt{2}}{2}$. 由选择公理知, 存在由此条件定义的映射 $\begin{array}{ccc}S&\longrightarrow&D\\ \alpha&\longmapsto&\alpha'\end{array}$; 再由范数的次可加性知, 此映射为单射, 故 S 也为至多可数集. □

引理 7.1.23 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的无限维完备内积空间, 则 V 中不存在标准正交 *(Hamel)* 基.

证明: 假设 V 中存在标准正交 (Hamel) 基 $B=\{\alpha_i\}_{i\in I}$, 则由 $|I|$ 非有限知, 可取 B 的一个可数无限子集 $\{\alpha_{i_j}\}_{j=0}^{+\infty}$. 考虑 V 中的绝对收敛级数 $\sum\limits_{j=0}^{+\infty}\dfrac{\alpha_{i_j}}{2^j}$, 则由 V 的完备性知, 它的极限为 $\beta\in V$. 由于 β 为 (Hamel) 基 B 中有限个元的线性组合, 则 $|\{i\in I\colon\langle\beta,\alpha_i\rangle\neq0\}|<+\infty$. 但 $\forall j\geq0$, $\left\langle\beta,\alpha_{i_j}\right\rangle=2^{-j}>0$, 矛盾! □

推论 7.1.24 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 *Hilbert* 空间, 则 $\dim_F(V)$ 有限或不可数.

上述两个引理都表明, 不可数 (Hamel) 维的可分或完备内积空间并不具有标准正交 (Hamel) 基. 这里我们回忆一个具体的例子:

$$\ell^2 \coloneqq \{(a_n)_{n\geq 0} \in F^\omega \colon \sum_{n=0}^{+\infty} |a_n|^2 < +\infty\},$$

其中内积定义为 $\langle (a_n)_{n\geq 0}, (b_n)_{n\geq 0}\rangle \coloneqq \sum\limits_{n=0}^{+\infty} a_n\overline{b_n}$. 容易验证 $(\ell^2, \langle\cdot,\cdot\rangle)$ 是一个不可数 (Hamel) 维的可分 Hilbert 空间. 它有一个可数维的可分非完备子内积空间 $F^{(\omega)}$, 以及一个极大的标准正交集 $\{\epsilon_i\}_{i=0}^{+\infty}$, 满足 $F^{(\omega)} = \mathrm{Span}(\{\epsilon_i\}_{i=0}^{+\infty})$. 进一步地, $(F^{(\omega)}, \langle\cdot,\cdot\rangle) \subseteq (\ell^2, \langle\cdot,\cdot\rangle)$ 还是稠密子空间.

更进一步地, 如果不考虑标准正交 (Hamel) 基, 仅仅考虑线性空间的维数是否可能分类不可数 (Hamel) 维的内积空间? 这里我们借助 $(\ell^2, \langle\cdot,\cdot\rangle)$ 给出两个具体的反例.

引理 7.1.25 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, (V,τ) 为域 F 上的拓扑线性空间, $W_1, W_2 \subseteq V$ 满足 $\mathrm{codim}_F(W_1) = \mathrm{codim}_F(W_2) < +\infty$, 则 $(W_1, \tau\big|_{W_1})$ 与 $(W_2, \tau\big|_{W_2})$ 保拓扑线性同构.*

证明: 由第二同构定理知, $W_1/(W_1\cap W_2) \cong (W_1+W_2)/W_2 \subseteq V/W_2$, 由 $\mathrm{codim}_F(W_2) < +\infty$ 知, $\dim_F(W_1/(W_1\cap W_2)) < +\infty$, 故存在有限维线性子空间 $U_1 \subseteq W_1$, 满足 $W_1 = U_1 \oplus (W_1\cap W_2)$. 同理存在有限维线性子空间 $U_2 \subseteq W_2$, 满足 $W_2 = U_2 \oplus (W_1\cap W_2)$. 又由 $\mathrm{codim}_F(W_1) = \mathrm{codim}_F(W_2)$ 可知 $\dim_F(U_1) = \dim_F(U_2)$, 则作为有限维拓扑线性空间, $(U_1, \tau\big|_{U_1})$ 与 $(U_2, \tau\big|_{U_2})$ 可保拓扑线性同构. 再在 $W_1\cap W_2$ 上取恒同映射, 则 $(W_1, \tau\big|_{W_1})$ 与 $(W_2, \tau\big|_{W_2})$ 保拓扑线性同构. □

例 7.1.5 (保拓扑线性同构但不保内积同构的内积空间) *考虑 Hilbert 空间 $(\ell^2, \langle\cdot,\cdot\rangle)$, $f_1\colon \ell^2 \to F$ 为非零连续线性函数, $f_2\colon \ell^2 \to F$ 为不连续线性函数, 则 $W_1 = \ker(f_1) \subseteq \ell^2$ 为余一维的闭线性子空间, $W_2 = \ker(f_2) \subseteq \ell^2$ 为余一维的稠线性子空间. 由上述引理知, 它们保拓扑线性同构; 但由完备性知, 它们不可能保内积同构.*

例 7.1.6 ((Hamel) 维数相同但不保内积同构的 Hilbert 空间) *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, I 为任意集合, 记*

$$\ell^2(I) \coloneqq \{(a_i)_{i\in I} \in F^I \colon \sum_{i\in I} |a_i|^2 \coloneqq \sup_{\substack{J\subseteq I\\ |J|<+\infty}} \sum_{j\in J} |a_j|^2 < +\infty\},$$

其中内积定义为 $\langle (a_i)_{i\in I}, (b_i)_{i\in I}\rangle \coloneqq \sum\limits_{i\in I} a_i\overline{b_i}$. (注意由 $\sum\limits_{i\in I} |a_i|^2 < +\infty$ 知, $\{i\in I\colon a_i \neq 0\}$ 为至多可数集, 故由 Cauchy-Schwarz 不等式估计余项知, 这里定义内积的级数绝对收敛.) 证明: $\dim_F(\ell^2(\mathbb{N})) = \dim_F(\ell^2(\mathbb{R})) = 2^{\aleph_0}$, 但 $(\ell^2(\mathbb{N}), \langle\cdot,\cdot\rangle)$ 与 $(\ell^2(\mathbb{R}), \langle\cdot,\cdot\rangle)$ 不保内积同构.

证明: 事实上, 由于 $(\ell^2(I), \langle\cdot,\cdot\rangle)$ 的一个极大标准正交集为 $\{\epsilon_i\}_{i\in I}$, 而 $|\mathbb{N}| < |\mathbb{R}|$, 故 $(\ell^2(\mathbb{N}), \langle\cdot,\cdot\rangle)$ 与 $(\ell^2(\mathbb{R}), \langle\cdot,\cdot\rangle)$ 的极大标准正交集不等势, 因此它们不可能保内积同构 (具体原因见 Schauder 维数).

一方面, 任取 $\mathbb{Q}$ 的一个排列 $\{r_n\}_{n\in\mathbb{N}}$, 记 $x_t \coloneqq \left(\dfrac{\chi_{(-\infty, r_n)}(t)}{n}\right)_{n\in\mathbb{N}}$ $(t\in\mathbb{R})$, 则可直接验证 $\{x_t\}_{t\in\mathbb{R}} \subseteq \ell^2(\mathbb{N})$ 为线性无关集, 故 $\dim_F(\ell^2(\mathbb{N})) \geq |\mathbb{R}| = 2^{\aleph_0}$. 另一方面, 由于 $\forall (a_i)_{i\in I} \in \ell^2(I)$, $\{i\in I\colon a_i\neq 0\}$ 为至多可数集, 则 $|\ell^2(I)| \leq |F|\cdot|\{J\subseteq I\colon J\text{为至多可数集}\}| = 2^{\aleph_0}\cdot |I|^{\aleph_0}$. 特别地, $\dim_F(\ell^2(\mathbb{R})) \leq |\ell^2(\mathbb{R})| \leq 2^{\aleph_0}\cdot(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}$. 因此 $\dim_F(\ell^2(\mathbb{N})) = \dim_F(\ell^2(\mathbb{R})) = 2^{\aleph_0}$. □

综上可知, 对于一般的内积空间, 尤其是不可数 (Hamel) 维的内积空间, 传统的标准正交 (Hamel) 基与 (Hamel) 维数并不能很好地体现其性质, 原因在于此时 "有限线性组合" 的条件过强. 因此在泛函分析中, 我们应该引入 "允许极限存在" 的标准正交 Schauder 基与 Schauder 维数.

定义 7.1.3 (Schauder 基) 设 V 为域 F 上的拓扑线性空间, $B\subseteq V$ 为子集, 若满足 $\forall \beta\in V$, $\exists! c_\beta\colon B\to F$, *s.t.* $\beta = \sum\limits_{\alpha\in B} c_\beta(\alpha)\cdot\alpha \coloneqq \lim\limits_{\substack{B'\subseteq B\\ |B'|<+\infty}} \sum\limits_{\alpha\in B'} c_\beta(\alpha)\cdot\alpha$, 则称 B 为 V 的一个 **Schauder 基** (Schauder basis).

注:

(1) 在上述 Schauder 基的定义中, 符号 $\lim\limits_{\substack{B'\subseteq B\\ |B'|<+\infty}}$ 指关于 B 的所有有限子集构成的集合族的直极限. 具体地说, 语句 "$\beta = \lim\limits_{\substack{B'\subseteq B\\ |B'|<+\infty}} \sum\limits_{\alpha\in B'} c_\beta(\alpha)\cdot\alpha$" 的含义为: "任取 $\beta\in V$ 的开邻域 Ω, 都存在有限子集 $B'\subseteq B$, 使得对于任意有限子集 $B'\subseteq B''\subseteq B$, 都有 $\beta - \sum\limits_{\alpha\in B''} c_\beta(\alpha)\cdot\alpha \in \Omega$".

(2) 上述 Schauder 基 $B \subseteq V$ 的定义语句还可写成: "$\forall \beta \in V$, $\exists! c_\beta \colon B \to F$, $s.t.$ $B_\beta \coloneqq \{\alpha \in B \colon c_\beta(\alpha) \neq 0\}$ 为可数集, 且 $\beta = \sum\limits_{\alpha \in B_\beta} c_\beta(\alpha) \cdot \alpha$ 无条件收敛".

(3) 一般的拓扑线性空间, 即使是 Banach 空间, 也未必具有上述定义的 Schauder 基. 具体的反例与更多的讨论可见 I. Singer"Bases in Banach Spaces", 在那里上述定义的 Schauder 基叫做 "extended unconditional basis".

然而, 一般的拓扑线性空间的任两组 Schauder 基 (若存在) 一定等势, 该基数称为拓扑线性空间的 Schauder 维数. 这里采用集合论中的证明办法, 可见 J. W. Evans, R. A. Tapia"Hamel Versus Schauder Dimension"(1970).

引理 7.1.26 *拓扑线性空间的任两组 Schauder 基等势.*

证明: 设 V 为域 F 上的拓扑线性空间, $B_1, B_2 \subseteq V$ 为两组 Schauder 基. 当 $|B_1|, |B_2|$ 之一有限时, V 的 Hamel 维数也有限, 此时 Schauder 基就是 Hamel 基, 故 $|B_1| = |B_2|$. 当 $|B_1|, |B_2|$ 均非有限时, 假设 $|B_1| < |B_2|$. 由 Schauder 基的定义知, $\forall \alpha \in B_1$, $\exists! c_\alpha \colon B_2 \to F$, $s.t.$ $B_{2,\alpha} \coloneqq \{\beta \in B_2 \colon c_\alpha(\beta) \neq 0\}$ 为可数集, 且 $\alpha \in \overline{\mathrm{Span}_F(B_{2,\alpha})}$. 记 $B_2' = \bigcup\limits_{\alpha \in$

则 $\overline{\mathrm{Span}_F(B_2')} \supseteq \overline{\mathrm{Span}_F(B_1)} = V$, 故 B_2' 也为 V 的一组 Schauder 基. 但 $B_2' \subseteq B_2$ 且 $|B_2'| \leq \aleph_0 \cdot |B_1| = |B_1| < |B_2|$, 这与 B_2 为 V 的一组 Schauder 基矛盾! 因此 $|B_1| \geq |B_2|$. 同理知 $|B_1| \leq |B_2|$, 故由 Schroder-Bernstein 定理知, $|B_1| = |B_2|$. □

以下我们关注 Hilbert 空间的标准正交 Schauder 基与 Schauder 维数, 并由此分类所有的 Hilbert 空间.

命题 7.1.27 *任意 Hilbert 空间都具有一组标准正交的 Schauder 基, 且任意两组标准正交的 Schauder 基等势.*

证明: 设 $(V, \langle \cdot, \cdot \rangle)$ 为 Hilbert 空间. 考虑集合族 $\mathscr{F} = \{S \subseteq V \colon S\text{为标准正交集}\}$ 及其上的包含偏序, 由 $\varnothing \in \mathscr{F}$ 知

$\mathscr{F} \neq \varnothing$. 现设 $\{S_i\}_{i \in I}$ 为 $\mathscr{F}$ 中的一条链, 取 $S = \bigcup\limits_{i \in I} S_i \subseteq V$, 显然 S 也为标准正交集, 则 S 为 $\{S_i\}_{i \in I}$ 在 $\mathscr{F}$ 中的上界. 由 Zorn 引理知, $\mathscr{F}$ 有极大元, 记为 B. 下证 $V = \overline{\mathrm{Span}_F(B)}$. 事实上, 由正交分解定理知, $V = \overline{\mathrm{Span}_F(B)} \oplus \overline{\mathrm{Span}_F(B)}^\perp$. 假设 $\overline{\mathrm{Span}_F(B)}^\perp \neq \{0\}$, 取 $\alpha \in \overline{\mathrm{Span}_F(B)}^\perp$ 满足 $\|\alpha\| = 1$, 则 $\langle \alpha, B \rangle = 0$, 故 $B \cup \{\alpha\}$ 也为标准正交集, 这与 B 的极大性矛盾! 因此 $V = \overline{\mathrm{Span}_F(B)}$. 由此可直接验证 B 为 $(V, \langle \cdot, \cdot \rangle)$ 的一组标准正交的 Schauder 基. □

命题 7.1.28 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $(V, \langle \cdot, \cdot \rangle)$ 为域 F 上的 Hilbert 空间, 则 $(V, \langle \cdot, \cdot \rangle)$ 保内积同构于 $\ell^2(I)$ 上的标准内积空间, 其中 $|I|$ 为 $(V, \langle \cdot, \cdot \rangle)$ 的 Schauder 维数. 特别地, 域 F 上的两个 Hilbert 空间保内积同构当且仅当它们的 Schauder 维数相等.*

证明: 由上述命题知, 可取 $(V, \langle \cdot, \cdot \rangle)$ 的一组标准正交 Schauder 基 $\{\alpha_i\}_{i \in I}$. 考虑线性映射 $T \colon V \longrightarrow \ell^2(I)$, $\alpha \longmapsto (\langle \alpha, \alpha_i \rangle)_{i \in I}$

由 Parseval 等式知 $\|\alpha\|^2 = \sum\limits_{i \in I} |\langle \alpha, \alpha_i \rangle|^2 = \|T(\alpha)\|^2$, 故 T 为定义良好的等距. 以下只需证明 T 为满射: 任取 $(a_i)_{i \in I} \in \ell^2(I)$, 由 $\sum\limits_{i \in I} |a_i|^2 < +\infty$ 以及 $(V, \langle \cdot, \cdot \rangle)$ 的完备性知, $\alpha \coloneqq \sum\limits_{i \in I} a_i \alpha_i \in V$, 且 $\langle \alpha, \alpha_i \rangle = a_i$, 即 $T(\alpha) = (a_i)_{i \in I}$. □

参考文献与补注 7.1

(1)

(2)

(3)

§7.2 自伴算子与酉算子

内积作为非退化的一次半线性形式, 天然提供了一种联系线性空间及其对偶的方式, 从而线性变换关于内积也可做伴随变换. 这赋予了内积空间上的线性变换更丰富的性质.

7.2.1 关于内积的伴随

为定义关于内积的伴随变换, 我们先引入基本的 Riesz-Frechet 表示定理: 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, 则映射 $\begin{array}{ccc} V & \longrightarrow & V^* \\ \beta & \longmapsto & \langle\cdot,\beta\rangle \end{array}$ 为共轭线性同构. 回忆教材或者安师讲义上的证明: 显然上述映射是共轭线性的, 于是为证明同构在有限维情形只需证明单性或满性之一即可. 这里单性是内积的非退化性; 满性是利用 V 中的标准正交 (Hamel) 基且 $\dim_F(V)<+\infty$. 容易看出上述结论对于无穷维情形不成立, 这是因为当 $\dim_F(V)$ 非有限时总有 $\dim_F(V)<\dim_F(V^*)$, 即此时原定义的 "(线性代数) 对偶空间" 过大了.

于是为解决此问题, 我们应该利用无穷维内积空间上的拓扑, 即考虑 V^* 中所有连续线性函数构成的子空间, 称为**连续对偶空间** (continuous dual space), 记为 V'. 并且, 在 V' 上可以赋予标准的范数 $\|\cdot\|'$: $\begin{array}{ccc} V' & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \sup\limits_{\alpha\in V\setminus\{0\}}\dfrac{|f(\alpha)|}{\|\alpha\|} \end{array}$,

容易证明此范数是定义良好的, 则 $(V',\|\cdot\|')$ 也为赋范线性空间. 此时由内积的连续性知, 上述映射可以写成 $\begin{array}{ccc} V & \longrightarrow & V' \\ \beta & \longmapsto & \langle\cdot,\beta\rangle \end{array}$.

进一步地, 由于 $(V,\|\cdot\|)$, $(V',\|\cdot\|')$ 为两个赋范线性空间, 我们也可以讨论它们之间映射的连续性. 一般的结果可总结为以下的引理:

引理 7.2.1 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, (V_1,p_1), (V_2,p_2) 为域 F 上的两个赋范线性空间, $T\in L(V_1,V_2)$, 则以下条件等价:*

(1) *T 为连续的;*

(2) *T 在 0 处连续;*

(3) $\sup\limits_{\alpha\in V_1\setminus\{0\}}\dfrac{p_2(T(\alpha))}{p_1(\alpha)}<+\infty$.

证明: "(1)⇒(2)": 显然;

"(2)⇒(3)": 由 T 在 0 处连续知, $\forall\epsilon>0$, $\exists\delta>0$, *s.t.* $T(B_\delta^{V_1}(0))\subseteq B_\epsilon^{V_2}(0)$. 固定 $0<r<\delta$, 则由范数的绝对齐次性知, $\sup\limits_{\alpha\in V_1\setminus\{0\}}\dfrac{p_2(T(\alpha))}{p_1(\alpha)}=\sup\limits_{\alpha\in\partial B_r^{V_1}(0)}\dfrac{p_2(T(\alpha))}{p_1(\alpha)}\le\dfrac{\epsilon}{r}<+\infty$.

"(3)⇒(1)": 记 $\sup\limits_{\alpha\in V_1\setminus\{0\}}\dfrac{p_2(T(\alpha))}{p_1(\alpha)}=M<+\infty$. 由拓扑基只需证明: $\forall B_\epsilon^{V_2}(\beta)\subseteq V_2$, $\forall\alpha\in V_1\,(T(\alpha)\in B_\epsilon^{V_2}(\beta))$, $\exists\delta>0$, *s.t.* $T(B_\delta^{V_1}(\alpha))\subseteq B_\epsilon^{V_2}(\beta)$. 事实上, 由范数的次可加性知, 取 $\delta=\dfrac{\epsilon-p_2(T(\alpha)-\beta)}{M}$ 即可. □

注: 设 $T\in L(V_1,V_2)$ 为连续的, 则 $\|T\|_{\mathrm{op}}\coloneqq\sup\limits_{\alpha\in V_1\setminus\{0\}}\dfrac{p_2(T(\alpha))}{p_1(\alpha)}<+\infty$ 称为 T 的**算子范数** (operator norm).

特别地, 由 Cauchy-Schwarz 不等式知, 映射 $\begin{array}{ccc} V & \longrightarrow & V' \\ \beta & \longmapsto & \langle\cdot,\beta\rangle \end{array}$ 为连续的, 且它的算子范数为 1. 然而即使如此, 此映射仍然未必是满射. 考虑 $(F^{(\omega)},\langle\cdot,\cdot\rangle)\subseteq(\ell^2,\langle\cdot,\cdot\rangle)$, 记 f: $\begin{array}{ccc} F^{(\omega)} & \longrightarrow & F \\ (a_n)_{n\ge0} & \longmapsto & \sum\limits_{n=0}^{+\infty}\dfrac{a_n}{n+1} \end{array}$, 则显然 f 为线性的, 且由 Cauchy-Schwarz 不等式知它为连续的. 但由 $\forall(b_n)_{n\ge0}\in F^{(\omega)}$, $|\{n\in\mathbb{N}:b_n\neq0\}|<+\infty$ 知, 不存在 $(b_n)_{n\ge0}\in F^{(\omega)}$, 满足 $f(\cdot)=\langle\cdot,(b_n)_{n\ge0}\rangle$. 因此我们还应该考虑完备的内积空间, 即 Hilbert 空间.

定理 7.2.2 (Riesz-Frechet) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, 则映射 $\begin{array}{ccc} V & \longrightarrow & V' \\ \beta & \longmapsto & \langle\cdot,\beta\rangle \end{array}$ 为连续的共轭线性同构.*

证明: 显然上述映射为连续共轭线性的. 由内积的非退化性知, 上述映射为单射. 下证它为满射: 任取 $f\in V'$, 则 $\ker(f)\subseteq V$ 为闭线性子空间. 由正交分解定理知, $V=\ker(f)\oplus\ker(f)^\perp$. 若 $\ker(f)^\perp=\{0\}$, 则 $V=\ker(f)$, 即 $f\equiv0$, 取 $\beta=0$ 即可; 若 $\ker(f)^\perp\neq\{0\}$, 则可取 $\beta_0\in\ker(f)^\perp\setminus\{0\}$, 满足 $f(\beta_0)=1$, 故

$\forall\alpha\in V,\ \alpha-f(\alpha)\beta_0\in\ker(f)$, 特别地 $\langle\alpha-f(\alpha)\beta_0,\beta_0\rangle=0$. 因此 $\forall\alpha\in V,\ f(\alpha)=\left\langle\alpha,\beta_0/\|\beta_0\|^2\right\rangle$, 即 $f(\cdot)=\left\langle\cdot,\beta_0/\|\beta_0\|^2\right\rangle$. □

注: 利用 Riesz-Frechet 表示定理, 我们可以在 V' 上赋予内积 $\langle\cdot,\cdot\rangle'$: $\begin{array}{ccc} V'\times V' & \longrightarrow & F \\ (f_1,f_2) & \longmapsto & \langle\beta_1,\beta_2\rangle \end{array}$, 其中 $\beta_i\in V\ (i=1,2)$

由

$f_i(\cdot)=\langle\cdot,\beta_i\rangle$ 唯一决定. 此时映射 $\begin{array}{ccc} V & \longrightarrow & V' \\ \beta & \longmapsto & \langle\cdot,\beta\rangle \end{array}$ 为保内积的共轭线性同构, 故 $(V',\langle\cdot,\cdot\rangle')$ 也为 Hilbert 空间.

定义 7.2.1 (关于内积的伴随变换) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, $T\in L(V)$ 为连续的, 若存在

$T^*\in L(V)$ 为连续的, 满足 $\langle T(\alpha),\beta\rangle=\langle\alpha,T^*(\beta)\rangle$, $\forall\alpha,\beta\in V$, 则称 T^* 为 T 关于内积的**伴随变换** (adjoint operator).

注: 由内积的非退化性可知, 关于内积的伴随变换若存在必唯一.

在有限维内积空间的情形, 由 Riesz-Frechet 表示定理即可定义任意线性变换关于内积的伴随变换. 在无穷维内积空间的情形, 即使是连续线性变换也未必存在关于内积的伴随变换. 考虑 $(F^{(\omega)},\langle\cdot,\cdot\rangle)\subseteq(\ell^2,\langle\cdot,\cdot\rangle)$, 记 $T\colon\ \begin{array}{ccc} F^{(\omega)} & \longrightarrow & F^{(\omega)} \\ (a_n)_{n\geq0} & \longmapsto & \sum\limits_{n\geq0}\dfrac{a_n}{n+1}\cdot\epsilon_0 \end{array}$, 则显然 T 为线性的, 且由 Cauchy-Schwarz 不等式知它为连续的. 但由

$\forall(c_n)_{n\geq0}\in F^{(\omega)}$,

$|\{n\in\mathbb{N}:c_n\neq0\}|<+\infty$ 知, 当 $b_0\neq0$ 时, 不存在 $(c_n)_{n\geq0}\in F^{(\omega)}$, 满足 $\langle T(\cdot),b_0\epsilon_0\rangle=\langle\cdot,(c_n)_{n\geq0}\rangle$. 因此我们还应该考虑完备的内积空间, 即 Hilbert 空间.

命题 7.2.3 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 则 T 关于内积存在唯一的伴随变换.*

证明: 任取 $\beta\in V$, 由 $T\in L(V)$ 以及内积的连续性知, $f_\beta\colon\ \begin{array}{ccc} V & \longrightarrow & F \\ \alpha & \longmapsto & \langle T(\alpha),\beta\rangle \end{array}$ 为连续线性函数, 则 $\begin{array}{ccc} V & \longrightarrow & V' \\ \beta & \longmapsto & f_\beta \end{array}$ 为共轭线性映射. 再由 Cauchy-Schwarz 不等式知, 此映射也为连续的. 现将其复合 Riesz-Frechet 表示定理给出的保内积共轭线性同构 $\begin{array}{ccc} V' & \longrightarrow & V \\ \langle\cdot,\beta\rangle & \longmapsto & \beta \end{array}$, 记为 $T^*\in L(V)$, 则 T^* 也为连续的, 且满足 $\langle T(\alpha),\beta\rangle=\langle\alpha,T^*(\beta)\rangle$, $\forall\alpha,\beta\in V$. □

注: 在实际问题中, 给定 $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, 线性变换 T 可能只在某个稠线性子空间 $D(T)\subseteq V$ 上有定义, 称为**稠定算子** (densely defined operator). 此时我们也可以定义 T 关于内积的伴随变换 T^*: 它的定义域为

$D(T^*)\coloneqq\left\{\beta\in V:\sup\limits_{\alpha\in D(T)\setminus\{0\}}\dfrac{|\langle T(\alpha),\beta\rangle|}{\|\alpha\|}<+\infty\right\}$, 从而在 $D(T^*)$ 上可利用上述命题的证明过程定义 T^*, 且由

$D(T)\subseteq V$ 稠知 T^* 是唯一定义的. 注意 T^* 未必是稠定算子!

引理 7.2.4 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 则*

(1) $\ker(T)=\mathrm{Im}(T^*)^\perp$;

(2) $\ker(T^*)=\mathrm{Im}(T)^\perp$;

(3) $\ker(T)^\perp=\overline{\mathrm{Im}(T^*)}$;

(4) $\ker(T^*)^\perp=\overline{\mathrm{Im}(T)}$.

进一步地, $\mathrm{Im}(T)\subseteq V$ 为闭线性子空间 $\Longleftrightarrow\mathrm{Im}(T^*)\subseteq V$ 为闭线性子空间.

证明: (1) 由定义直接验证即可; (2) 由 (1) 与 $(T^*)^*=T$ 即知; (3),(4) 分别由 (1),(2) 以及 $\overline{W}=(W^\perp)^\perp$ 即知. 最后,

"$\mathrm{Im}(T)\subseteq V$ 为闭线性子空间 $\Longleftrightarrow\mathrm{Im}(T^*)\subseteq V$ 为闭线性子空间." 是一个非平凡的结果, 可以参考 H. Brezis "Functional Analysis, Sobolev Spaces and Partial Differential Equations" 或者 W. Rudin "Functional Analysis". □

引理 7.2.5 (闭值域刻画) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 则以下条件等价:*

(1) *T^* 为满射;*

(2) $\inf\limits_{\alpha\in V\setminus\{0\}}\dfrac{\|T(\alpha)\|}{\|\alpha\|}>0$;

(3) *T 为单射且 $\mathrm{Im}(T)\subseteq V$ 为闭线性子空间.*

证明: "(1)⇒(2)":

"(2)⇒(3)": 设 $\inf\limits_{\alpha\in V\setminus\{0\}}\dfrac{\|T(\alpha)\|}{\|\alpha\|}>0$, 即 $\exists C>0$, $\forall\alpha\in V$, $\|T(\alpha)\|\geq C\|\alpha\|$, 则显然 T 为单射. 现设 $\{T(\alpha_n)\}_{n=1}^{+\infty}\subseteq\mathrm{Im}(T)$ 且 $\lim\limits_{n\to+\infty}T(\alpha_n)=\beta\in V$, 则 $\{T(\alpha_n)\}_{n=1}^{+\infty}\subseteq V$ 为 Cauchy 序列. 由条件可知, $\{\alpha_n\}_{n=1}^{+\infty}\subseteq V$ 也为 Cauchy 序列, 故由 V 的完备性知, $\exists\alpha\in V$, *s.t.* $\lim\limits_{n\to+\infty}\alpha_n=\alpha$; 再由 T 的连续性知, $\beta=\lim\limits_{n\to+\infty}T(\alpha_n)=T(\alpha)\in\mathrm{Im}(T)$, 因此 $\mathrm{Im}(T)$ 为闭线性子空间.

"(3)⇒(1)": 设 T 为单射且 $\mathrm{Im}(T)\subseteq V$ 为闭线性子空间, 则由闭值域刻画知, $\mathrm{Im}(T^*)\subseteq V$ 也为闭线性子空间, 故 $\mathrm{Im}(T^*)=\ker(T)^\perp=\{0\}^\perp=V$. □

7.2.2 自伴算子

定义 7.2.2 (自伴算子) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 若 $T=T^*$, 则称 T 为**自伴算子** (self-adjoint operator).

注:

(1) 由引理 7.2.4知, Hilbert 空间上的自伴算子诱导了正交分解 $V=\ker(T)\oplus\overline{\mathrm{Im}(T)}$;

(2) 由自伴定义 $\forall\alpha,\beta\in V$, $\langle T(\alpha),\beta\rangle=\langle\alpha,T(\beta)\rangle$ 知, 自伴算子 T 的特征值必为实数;

(3) 另外, 自伴算子的一个重要性质是: 它的不变子空间的正交补仍为不变子空间.

利用上述观察, 我们很容易推出有限维内积空间上自伴算子的正交相似 (或酉相似) 标准形.

命题 7.2.6 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$, 则以下条件等价:*

(1) *T 为自伴算子;*

(2) *T 在标准正交基下的矩阵表示可正交相似 (或酉相似) 于对角形, 且对角分量均为实数.*

证明: "(1)⇒(2)": 由于自伴算子一定有一维不变子空间 (特征值为实数), 且它的不变子空间的正交补仍为不变子空间, 故对维数归纳即知结论.

"(2)⇒(1)": 显然. □

注: 记有限维内积空间上的自伴算子为 $T=\sum\limits_{i=1}^{k}c_iE_i$, 其中 E_i 是到第 i 个特征子空间上的投影算子, 则 $\mathrm{id}_V=\sum\limits_{i=1}^{k}E_i$, 且 $\mathrm{Im}(E_i)\perp\mathrm{Im}(E_j)$, $\forall 1\leq i\neq j\leq k$. 这里投影算子族 $\{E_i\}_{i=1}^k$ 由 T 唯一决定, 称为 T 的谱投影算子族. 事实上, 它们都是 T 的多项式: 取 $f_i(X)\in\mathbb{R}[X]$ 满足 $f_i(c_j)=\delta_{ij}$, $\forall 1\leq i,j\leq k$, 则 $E_i=f_i(T)$, $\forall 1\leq i\leq k$. 特别地, $E_i\,(1\leq i\leq k)$ 也为自伴算子.

一类最简单的自伴算子是到闭线性子空间的最佳逼近映射, 称为**正交投影** (orthogonal projection). 利用谱理论, 我们将看到到特征子空间上的正交投影算子类是构成一个自伴算子的基石, 即对于 Hilbert 空间上的自伴算子, 可以构造相应的谱族, 使得它写成关于谱测度的积分形式.

引理 7.2.7 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 满足 $T^2=T$, 则以下条件等价:*

(1) *T 为正交投影;*

(2) *T 为自伴算子;*

(3) $\forall\alpha\in V$, $\|T(\alpha)\|\leq\|\alpha\|$.

证明: "(1)⇒(2)": 设 T 为正交投影, 即存在 $W\subseteq V$ 为闭线性子空间, 满足 $T=P_W$. 由正交分解 $V=W\oplus W^\perp$ 知, $\forall\alpha,\beta\in V$, $\langle P_W(\alpha),\beta-P_W(\beta)\rangle=0$, 则 $\langle P_W(\alpha),\beta\rangle=\langle P_W(\alpha),P_W(\beta)\rangle=\langle\alpha,P_W(\beta)\rangle$, 即 $P_W=P_W^*$.

"(2)⇒(3)": 设 T 为自伴算子, 则存在正交分解 $V = \ker(T) \oplus \overline{\mathrm{Im}(T)}$. 由 $T^2 = T$ 以及 T 的连续性知, $T|_{\overline{\mathrm{Im}(T)}} = \mathrm{id}_{\overline{\mathrm{Im}(T)}}$. 记 $P_{\overline{\mathrm{Im}(T)}}$ 为到 $\overline{\mathrm{Im}(T)}$ 上的正交投影, 则 $\forall \alpha \in V$, $T(\alpha) = (T \circ P_{\overline{\mathrm{Im}(T)}})(\alpha) = P_{\overline{\mathrm{Im}(T)}}(\alpha)$, 故 $\|T(\alpha)\| = \|P_{\overline{\mathrm{Im}(T)}}(\alpha)\| \leq \|\alpha\|$. "(3)⇒(1)": 由 $T^2 = T$ 以及准素分解知, $V = \ker(T) \oplus \ker(T - \mathrm{id}_V)$. 断言: $\ker(T) \perp \ker(T - \mathrm{id}_V)$. (事实上, 任取 $\alpha \in \ker(T), \beta \in \ker(T - \mathrm{id}_V)$, 以及 $\epsilon \in F \setminus \{0\}$ 满足 $\langle \alpha, \epsilon \cdot \beta \rangle \in \mathbb{R}$, $t \in \mathbb{R}$, 则 $T(\alpha + t\epsilon \cdot \beta) = t\epsilon \cdot \beta$. 由 (3) 知 $\|t\epsilon \cdot \beta\| = \|T(\alpha + t\epsilon \cdot \beta)\| \leq \|\alpha + t\epsilon \cdot \beta\|$, 平方展开得 $0 \leq \|\alpha\|^2 + 2t \cdot \langle \alpha, \epsilon \cdot \beta \rangle$. 由于此式对于 $t \in \mathbb{R}$ 恒成立, 则 $\langle \alpha, \epsilon \cdot \beta \rangle = 0$, 即 $\langle \alpha, \beta \rangle = 0$.) 因此 $\forall \alpha \in V$, $T(\alpha) \in \ker(T - \mathrm{id}_V)$ 且 $(T(\alpha) - \alpha) \perp \ker(T - \mathrm{id}_V)$, 则由最佳逼近点的刻画知, $T(\alpha) = P_{\ker(T - \mathrm{id}_V)}(\alpha)$, 即 $T = P_{\ker(T - \mathrm{id}_V)}$. □

注: 上述引理表明, 到闭线性子空间上的正交投影可由投影算子类中的范数最小性唯一刻画, 即设 $W \subseteq V$ 为闭线性子空间, 则 P_W 是泛函 $\{P \in L(V) \colon P\text{连续}, P^2 = P, \mathrm{Im}(P) = W\} \longrightarrow \mathbb{R}$ 的唯一最小值点.

$$P \longmapsto \|P\|_{\mathrm{op}}$$

引理 7.2.8 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $(V, \langle \cdot, \cdot \rangle)$ 为域 F 上的 Hilbert 空间, $T_1, T_2 \in L(V)$ 为正交投影, 则:*

(1) *$T_1 \circ T_2$ 为正交投影 $\iff T_2 \circ T_1$ 为正交投影 $\iff T_1 \circ T_2 = T_2 \circ T_1 \iff T_1 + T_2 \circ T_1 \circ T_2$ 为正交投影;*

(2) *$T_1 + T_2$ 为正交投影 $\iff T_1 \circ T_2 = 0 \iff T_2 \circ T_1 = 0 \iff \mathrm{Im}(T_1) \perp \mathrm{Im}(T_2)$;*

(3) *$T_1 - T_2$ 为正交投影 $\iff T_1 \circ T_2 = T_2 \iff T_2 \circ T_1 = T_2 \iff \mathrm{Im}(T_2) \subseteq \mathrm{Im}(T_1) \iff \forall \alpha \in V$, $\|T_2(\alpha)\| \leq \|T_1(\alpha)\|$;*

(4) *若 $\|T_1 - T_2\|_{\mathrm{op}} < 1$, 则 $\dim_F(\ker(T_1)) = \dim_F(\ker(T_2))$; $\dim_F(\mathrm{Im}(T_1)) = \dim_F(\mathrm{Im}(T_2))$.*

证明: (1)(2)(3) 都是简单的验证; 以下证明 (4):

□

7.2.3 酉算子

定义 7.2.3 (酉算子) 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $(V, \langle \cdot, \cdot \rangle)$ 为域 F 上的 Hilbert 空间, $T \in L(V)$ 为连续的, 若 $T^* \circ T = T \circ T^* = \mathrm{id}_V$, 则称 T 为**酉算子** (unitary operator).

注: 当 $\dim_F(V) < +\infty$ 时, $T \in L(V)$ 自动连续, 且保内积条件 "$T^* \circ T = \mathrm{id}_V$" 已蕴涵 "$T \circ T^* = \mathrm{id}_V$", 于是此时酉算子的定义可简单地写成 "保内积的线性变换". 但当 $\dim_F(V)$ 非有限时, 上述条件缺一不可.

命题 7.2.9 *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $(V, \langle \cdot, \cdot \rangle)$ 为域 F 上的 Hilbert 空间, $T \in L(V)$ 为连续的, 则以下条件等价:*

(1) *T 为酉算子;*

(2) *T 为满射且保内积;*

(2) *$\mathrm{Im}(T) \subseteq V$ 为稠子空间, 且 T 保内积.*

证明: "(1)⇒(2)⇒(3)": 显然;
"(3)⇒(2)": 由 T 保距以及闭值域刻画知, $\mathrm{Im}(T)$ 为闭线性子空间, 故 $\mathrm{Im}(T) \subseteq V$ 为稠子空间 $\iff T$ 为满射.
"(2)⇒(1)": 由 T 保内积知, $T^* \circ T = \mathrm{id}_V$. 由 T 为满射知, $\forall \beta \in V$, $\exists \alpha \in V$, *s.t.* $\beta = T(\alpha)$, 则 $T^*(\beta) = (T^* \circ T)(\alpha) = \alpha$, 故 $(T \circ T^*)(\beta) = T(\alpha) = \beta$, 即 $T \circ T^* = \mathrm{id}_V$. 因此 T 为酉算子. □

注: 显然 $(\ell^2(\mathbb{N}), \langle \cdot, \cdot \rangle)$ 上的右移算子是保内积的连续线性变换, 但不是满射, 故不为酉算子. 但是 $(\ell^2(\mathbb{Z}), \langle \cdot, \cdot \rangle)$ 上的右移算子却是酉算子.

值得注意的是, 复 Hilbert 空间上的酉算子和自伴算子可通过 Cayley 变换联系起来:

引理 7.2.10 *设 $(V, \langle \cdot, \cdot \rangle)$ 为复 Hilbert 空间, $T \in L(V)$ 为自伴算子, 则:*

(1) *$\|(\mathrm{id}_V \pm \sqrt{-1}T)\alpha\|^2 = \|\alpha\|^2 + \|T(\alpha)\|^2$, $\forall \alpha \in V$;*

(2) *$\mathrm{id}_V \pm \sqrt{-1}T \in L(V)$ 均为线性同构.*

证明: (1) 任取 $\alpha \in V$, 则由 T 自伴知 $\langle \alpha, T(\alpha) \rangle = \langle T(\alpha), \alpha \rangle$, 故展开知 $\|(\mathrm{id}_V \pm \sqrt{-1}T)\alpha\|^2 = \|\alpha\|^2 + \|T(\alpha)\|^2$.
(2) 由 (1) 知 $(\mathrm{id}_V \pm \sqrt{-1}T)$ 为单射, 即 $\ker(\mathrm{id}_V \pm \sqrt{-1}T) = \{0\}$, 则由引理知 $\overline{\mathrm{Im}(\mathrm{id}_V \pm \sqrt{-1}T)} = \ker(\mathrm{id}_V \mp \sqrt{-1}T^*)^\perp$
$= \ker(\mathrm{id}_V \mp \sqrt{-1}T)^\perp = \{0\}^\perp = V$. 再由 (1) 知 $\|(\mathrm{id}_V \pm \sqrt{-1}T)\alpha\| \geq \|\alpha\|$, $\forall \alpha \in V$, 由闭值域刻画知 $\mathrm{Im}(\mathrm{id}_V \pm \sqrt{-1}T)$ 为闭线性子空间, 故 $(\mathrm{id}_V \pm \sqrt{-1}T)$ 为满射. □

命题 7.2.11 (Cayley 变换) 设 $(V,\langle\cdot,\cdot\rangle)$ 为复 *Hilbert* 空间, 则存在集合之间的一一对应:

$$\begin{aligned}\{T\in L(V)\colon T\text{为自伴算子}\}&\longrightarrow\{U\in L(V)\colon U\text{为酉算子, 且}1\notin\sigma(U)\}\\ T&\longmapsto(T-\sqrt{-1}\mathrm{id}_V)\circ(T+\sqrt{-1}\mathrm{id}_V)^{-1}\\ \sqrt{-1}(I+U)\circ(I-U)^{-1}&\longmapsfrom U\end{aligned}$$

证明: 设 T 为自伴算子, 则 $\mathrm{id}_V\pm\sqrt{-1}T\in L(V)$ 均为线性同构, 故 $U:=(T-\sqrt{-1}\mathrm{id}_V)\circ(T+\sqrt{-1}\mathrm{id}_V)^{-1}\in L(V)$ 为定义良好的线性同构. 任取 $\alpha\in V$, 记 $\alpha=(T+\sqrt{-1}\mathrm{id}_V)(\beta)$, 则 $\|U(\alpha)\|=\|(T-\sqrt{-1}\mathrm{id}_V)(\beta)\|=\|(T+\sqrt{-1}\mathrm{id}_V)(\beta)\|=\|\alpha\|$, 即 U 为等距算子. 因此由上述引理知, U 为酉算子. 再由 $\mathrm{id}_V-U=2\sqrt{-1}(T+\sqrt{-1}\mathrm{id}_V)^{-1}$ 知, $1\notin\sigma(U)$.

反之, 设 U 为酉算子且 $1\notin\sigma(U)$, 则 $T:=\sqrt{-1}(I+U)\circ(I-U)^{-1}\in L(V)$ 为定义良好的线性变换. 由定义 $U^*\circ U=U\circ U^*=\mathrm{id}_V$ 可直接验证 $T^*=T$. 显然上述对应是互逆的. □

另外, 对于有限维内积空间上的线性变换 T, 教材已证 T 为酉算子 $\Longleftrightarrow T$ 保内积 $\Longleftrightarrow T$ 保距离 $\Longleftrightarrow T$ 保长度. 以下设法推广这一观察或举一些反例.

命题 7.2.12 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 *Hilbert* 空间, $\{\alpha_i\}_{i=1}^m\cup\{\beta_i\}_{i=1}^m\subseteq V$, 则以下条件等价:

(1) $\forall 1\le i,j\le m$, $\langle\alpha_i,\alpha_j\rangle=\langle\beta_i,\beta_j\rangle$;

(2) 存在酉算子 $T\in L(V)$, 满足 $T(\alpha_i)=\beta_i$, $\forall 1\le i\le m$.

证明: "(1)⇒(2)": 记 $W_1=\mathrm{Span}_F(\{\alpha_i\}_{i=1}^m)$, $W_2=\mathrm{Span}_F(\{\beta_i\}_{i=1}^m)$. 任取 $c_1,\cdots,c_m\in F$, 则由条件知 $\left\|\sum\limits_{i=1}^m c_i\alpha_i\right\|^2=\sum\limits_{i,j=1}^m|c_i|^2\langle\alpha_i,\alpha_j\rangle=\sum\limits_{i,j=1}^m|c_i|^2\langle\beta_i,\beta_j\rangle=\left\|\sum\limits_{i=1}^m c_i\beta_i\right\|^2$, 故对于任意子集 $\{i_1,\cdots,i_k\}\subseteq\{1,\cdots,m\}$, $\{\alpha_{i_1},\cdots,\alpha_{i_k}\}$ 线性无关当且仅当 $\{\beta_{i_1},\cdots,\beta_{i_k}\}$ 线性无关, 因此 $\dim_F(W_1)=\dim_F(W_2)<+\infty$. 由正交分解的存在性知, $V=W_1\oplus W_1^\perp=W_2\oplus W_2^\perp$.

现定义酉算子 $T\in L(V)$ 如下: 先令 $T\big|_{W_1}:\quad W_1\longrightarrow W_2$, 由上述线性无关性的论证知这是定义

$$\sum_{i=1}^m c_i\alpha_i\longmapsto\sum_{i=1}^m c_i\beta_i$$

良好的, 且为保内积同构. 另外, 由第二同构定理知, $W_1^\perp/(W_1^\perp\cap W_2^\perp)\cong(W_1^\perp+W_2^\perp)/W_2^\perp\subseteq V/W_2^\perp$, 由 $\dim_F(V/W_2^\perp)<+\infty$
知, $\dim_F(W_1^\perp/(W_1^\perp\cap W_2^\perp))<+\infty$, 故由正交分解知, 存在有限维线性子空间 $U_1\subseteq W_1^\perp$, 满足 $W_1^\perp=U_1\oplus(W_1^\perp\cap W_2^\perp)$, 且 $U_1\perp(W_1^\perp\cap W_2^\perp)$. 同理, 存在有限维线性子空间 $U_2\subseteq W_2^\perp$, 满足 $W_2^\perp=U_2\oplus(W_1^\perp\cap W_2^\perp)$, 且 $U_2\perp(W_1^\perp\cap W_2^\perp)$. 再由 $\dim_F(V/W_1^\perp)=\dim_F(V/W_2^\perp)<+\infty$ 知, $\dim_F(U_1)=\dim_F(U_2)<+\infty$, 故可令 $T\big|_{U_1}:U_1\xrightarrow{\cong}U_2$ 为任意保内积同构. 最后, 令 $T\big|_{W_1^\perp\cap W_2^\perp}=\mathrm{id}_{W_1^\perp\cap W_2^\perp}$. 综上, $T=T\big|_{W_1}\oplus T\big|_{U_1}\oplus T\big|_{W_1^\perp\cap W_2^\perp}$ 为满足条件的酉算子.
"(2)⇒(1)": 显然. □

注: 上述命题对于无穷多个向量未必成立, 反例如取 $(V,\langle\cdot,\cdot\rangle)=(\ell^2(\mathbb{N}),\langle\cdot,\cdot\rangle)$, $\alpha_i=\epsilon_{2i}$, $\beta_i=\epsilon_i$, 则条件 (1) 成立, 但条件 (2) 不成立: 满足 $T(\epsilon_{2i})=\epsilon_i$, $\forall i\in\mathbb{N}$ 的连续线性变换为将某个真线性子空间映为全空间, 故不可能延拓为酉算子.

命题 7.2.13 设 $(V,\langle\cdot,\cdot\rangle)$ 为实内积空间, $T\colon V\to V$ 为保距变换 (不一定线性), 则 T 为一个平移与一个线性变换的复合.

证明: 通过平移, 可不妨设 $T(0)=0$. 任取 $\alpha,\beta\in V$, 由 T 保距知 $\|T(\alpha)-T(\beta)\|=\|\alpha-\beta\|$; 特别地, $\|T(\alpha)\|=\|\alpha\|$. 现由 $\|T(\alpha)-T(\beta)\|^2=\|\alpha-\beta\|^2$ 以及实内积的性质展开知, $\langle T(\alpha),T(\beta)\rangle=\langle\alpha,\beta\rangle$. 于是

$$\begin{aligned}\|T(\alpha+\beta)-T(\alpha)-T(\beta)\|^2&=\langle T(\alpha+\beta)-T(\alpha)-T(\beta),T(\alpha+\beta)-T(\alpha)-T(\beta)\rangle\\&=\langle(\alpha+\beta)-\alpha-\beta,\alpha+\beta-\alpha-\beta\rangle=0,\end{aligned}$$

则 $T(\alpha+\beta)=T(\alpha)+T(\beta)$. 同理 $\forall c\in\mathbb{R}$,

$$\|T(c\alpha)-c\cdot T(\alpha)\|^2=\langle T(c\alpha)-c\cdot T(\alpha),T(c\alpha)-c\cdot T(\alpha)\rangle=\langle c\alpha-c\alpha,c\alpha-c\alpha\rangle=0,$$

则 $T(c\alpha)=c\cdot T(\alpha)$. 因此 T 为线性变换. □

注:

(1) 上述命题可推广为: 实赋范线性空间上的保距满射一定是仿射的; 此即 Mazur-Ulam 定理.

(2) 上述命题对于复内积空间上的保距变换未必成立, 例如复共轭映射保零点但不是复线性的.
(3) 上述命题对于实复内积空间上的保长度变换都未必成立, 例如每分量取绝对值映射保零点但不是线性的.

以下我们关注实二三维内积空间上的旋转, 这是一种最特殊情形的酉算子.

例 7.2.1 设 $V=\mathbb{C}$ 视为实二维线性空间, 则 $\mathbb{C}$ 上标准复内积的实部是 V 上的实内积, 且 $\begin{aligned} V &\longrightarrow \mathbb{R}^2 \\ a+b\sqrt{-1} &\longmapsto (a,b)^t \end{aligned}$ 为保实内积的线性同构. 考虑实线性代数同态 $\begin{aligned} \mathbb{C} &\longrightarrow L(V) \\ c &\longmapsto (T_c\colon \alpha\mapsto c\alpha) \end{aligned}$, 注意 $T_{\bar{c}}=T_c^*$, $\forall c\in\mathbb{C}$, 则上述同态是一个 $*$-同态, 即保持二阶的共轭线性反自同构. 这里 T_c 自伴 $\Longleftrightarrow c\in\mathbb{R}$; T_c 酉 $\Longleftrightarrow c\in U(1)$; T_c 正定 $\Longleftrightarrow c>0$. 特别地, 上述同态限制到酉部分为群同态 $\begin{aligned} U(1) &\longrightarrow \mathrm{O}(2) \\ c &\longmapsto T_c \end{aligned}$, 这是一个单同态, 它的像为 $\mathrm{SO}(2)\subsetneq\mathrm{O}(2)$. 事实上, $\mathrm{O}(2)$ 关于指标为 2 的子群 $\mathrm{SO}(2)$ 可做陪集分解

$$\mathrm{O}(2)=\mathrm{SO}(2)\sqcup\mathrm{SO}(2)\cdot\mathrm{diag}(1,-1)=\left\{\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}:\theta\in[0,2\pi)\right\}\bigsqcup\left\{\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}:\theta\in[0,2\pi)\right\},$$

其中 $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ 为绕原点逆时针旋转 θ, $\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$ 为关于斜率为 $\tan(\theta/2)$ 的直线反射.

为方便描述实三维内积空间上的旋转, 我们先引入四元数的概念.

定义 7.2.4 (四元数) 记 $\mathbb{H}\coloneqq\{w+x\cdot i+y\cdot j+z\cdot k\colon w,x,y,z\in\mathbb{R}\}$ 是以 $\{1,i,j,k\}$ 为基的实四维线性空间. 现在按以下方式引入 $\mathbb{H}$ 上的乘法:

(1) 1 是乘法幺元;
(2) 实数与任意四元数乘法可交换;
(3) 记乘法表

$\cdot$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

, 并按照与加法分配的方式延拓为任意两个四元数的乘法.

于是任意非零四元数都有乘法逆元 $(w+x\cdot i+y\cdot j+z\cdot k)^{-1}=\dfrac{1}{w^2+x^2+y^2+z^2}(w-x\cdot i-y\cdot j-z\cdot k)$. 因此 $\mathbb{H}$ 关于加法、数乘与乘法构成了一个实线性体, 称为**四元数体** (quaternion).

注:

(1) $\mathbb{H}$ 中的乘法运算满足结合律, 但不满足交换律. 这是历史上数学家发现四元数体的主要困难. 1843 年 10 月 16 日, 当 W. M. Hamilton 与他的妻子沿着 Dublin 的皇家运河散步时, 他忽然意识到四元数的乘法可以由公式
$i^2=j^2=k^2=ijk=-1$ 定义, 于是将它刻在了 Brougham 桥下的石头上.
(2) 对于四元数 $q=w+x\cdot i+y\cdot j+z\cdot k\,(w,x,y,z\in\mathbb{R})$, 记它的共轭为 $\bar{q}=w-x\cdot i-y\cdot j-z\cdot k$, 长度为 $|q|=\sqrt{q\bar{q}}=\sqrt{w^2+x^2+y^2+z^2}$.
(3) 任意非零四元数都存在唯一分解 $q=|q|\cdot(\cos\theta+\sin\theta\cdot\alpha)$, 其中 $\cos\theta=\dfrac{w}{|q|}$, $\sin\theta=\dfrac{\sqrt{x^2+y^2+z^2}}{|q|}$, $\alpha=\dfrac{x\cdot i+y\cdot j+z\cdot k}{\sqrt{x^2+y^2+z^2}}$. 注意 $\theta\in[0,\pi]$, $\alpha\in\mathbb{H}$, $|\alpha|=1$, 且 α 无 1 分量.
(4) 记 $\mathbb{H}^0\coloneqq\mathrm{Span}_{\mathbb{R}}(\{i,j,k\})\subseteq\mathbb{H}$ 为实三维子空间, $\mathbb{H}^0$ 中元称为纯四元数. 在 $\mathbb{H}^0$ 上可定义内积与外积如下: $\begin{aligned}\langle\cdot,\cdot\rangle\colon\ \mathbb{H}^0\times\mathbb{H}^0 &\longrightarrow \mathbb{R}\\ (\alpha_1,\alpha_2) &\longmapsto -\frac{\alpha_1\alpha_2+\alpha_2\alpha_1}{2}\end{aligned}$, $\begin{aligned}\times\colon\ \mathbb{H}^0\times\mathbb{H}^0 &\longrightarrow \mathbb{H}^0\\ (\alpha_1,\alpha_2) &\longmapsto \frac{\alpha_1\alpha_2-\alpha_2\alpha_1}{2}\end{aligned}$. 注意在实线性同构 $\begin{aligned}\langle\cdot,\cdot\rangle\colon\ \mathbb{H}^0 &\longrightarrow \mathbb{R}^3\\ i &\longmapsto e_1\\ j &\longmapsto e_2\\ k &\longmapsto e_3\end{aligned}$ 下, 这对应了 $\mathbb{R}^3$ 上通常的内积与外积.

(5) 可直接验证 $Z(\mathbb{H}) = C_{\mathbb{H}}(\mathbb{H}^0) = \mathbb{R}$.

命题 7.2.14 *设 $\alpha \in \mathbb{H}^0$ 且 $|\alpha| = 1$, $T \in L(\mathbb{H}^0)$ 为绕 α 轴逆时针旋转 θ, 则 $\exists \delta \in \mathbb{H}$, s.t. $|\delta| = 1$ 且 $T(q) = \delta q \delta^{-1}$, $\forall q \in \mathbb{H}^0$.*

证明: 取 $\beta \in \mathbb{H}^0$ 且 $|\beta| = 1$, 满足 $\langle \alpha, \beta \rangle = 0$, 则 $\alpha\beta = -\beta\alpha$. 再取 $\gamma = \alpha\beta = \alpha \times \beta \in \mathbb{H}^0$, 则 $|\gamma| = |\alpha||\beta| = 1$, 且 $\alpha\gamma = \alpha^2\beta = -\alpha\beta\alpha = -\gamma\alpha$, 即 $\langle \alpha, \gamma \rangle = 0$; 同理知 $\langle \beta, \gamma \rangle = 0$. 因此 $\{\alpha, \beta, \gamma\}$ 是 $\mathbb{H}^0$ 的一组标准正交基, 且定向与 $\mathbb{R}^3$ 中的标准定向一致. 令 $\delta = \cos(\theta/2) + \sin(\theta/2) \cdot \alpha$, 则 $\delta \in \mathbb{H}$, $|\delta| = 1$, 且 $\delta^{-1} = \cos(\theta/2) - \sin(\theta/2) \cdot \alpha$. 考虑实线性映射 $\begin{array}{ccc} \mathbb{H}^0 & \longrightarrow & \mathbb{H}^0 \\ q & \longmapsto & \delta q \delta^{-1} \end{array}$, 可直接验证它在基 $\{\alpha, \beta, \gamma\}$ 下的矩阵表示为 $\begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}$, 即绕 α 轴逆时针旋转 θ. □

注:

(1) 上述命题表明存在群同态 $\{\delta \in \mathbb{H} \colon |\delta| = 1\} \xrightarrow{2:1} \{\mathbb{H}^0\text{中旋转}\}$, 即将 $\delta = \cos(\theta/2) + \sin(\theta/2) \cdot \alpha$ ($\theta \in [0, 2\pi]$, $\alpha \in \mathbb{H}^0$, $|\alpha| = 1$) 对应于绕 α 轴逆时针旋转 θ. 此对应的意义在于, 给定实三维空间中的两个绕轴旋转, 我们可以将它们转化为两个单位长度的四元数, 从而求这两个旋转复合后的旋转轴和夹角就转化为求对应的两个四元数的乘积. 这种方法比矩阵乘法或 Euler 角快捷许多.

(2) 利用上述命题, 可验证 Rodrigue 旋转公式: 设 $\alpha \in \mathbb{H}^0$ 且 $|\alpha| = 1$, $T \in L(\mathbb{H}^0)$ 为绕 α 轴逆时针旋转 θ, 则
$$\forall q \in \mathbb{H}^0,\ T(q) = \cos\theta \cdot q + \sin\theta \cdot (\alpha \times q) + \langle \alpha, q \rangle \cdot (1 - \cos\theta) \cdot \alpha.$$

引理 7.2.15 *存在实线性体的同构 $\mathbb{H} \xrightarrow{\cong} \left\{ \begin{pmatrix} u & -v \\ \overline{v} & \overline{u} \end{pmatrix} : u, v \in \mathbb{C} \right\}$, 其中*
$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ i \mapsto \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix},\ j \mapsto \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ k \mapsto \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}.$$

推论 7.2.16 *在上述实线性体的同构以及 $\mathbb{H}^0 \cong \mathbb{R}^3$ 的意义下, 上述命题中的群同态也可写为* $\begin{array}{ccc} \mathrm{SU}(2) & \xrightarrow{2:1} & \mathrm{SO}(3) \\ \downarrow\cong & & \downarrow\cong \\ S^3 & \xrightarrow{2:1} & \mathbb{R}P^3 \end{array}$.

这里我们插入关于 "三元数" 的一些讨论. 在之前的例子中, 若将 $\mathbb{C}$ 视为 $\mathbb{R}^2$, 将 $\mathbb{C}$ 上的乘法视为 $\mathbb{R}^2$ 上的乘法, 则我们自然地给 $\mathbb{R}^2$ 赋予了域结构; 同样地, 若将 $\mathbb{H}$ 视为 $\mathbb{R}^4$, 将 $\mathbb{H}$ 上的乘法视为 $\mathbb{R}^4$ 上的乘法, 则我们自然地给 $\mathbb{R}^4$ 赋予了体结构. 问题是: 能否在 $\mathbb{R}^3$ 上定义乘法, 使得它具有某种域 (或者体) 的结构? 历史上, W. M. Hamilton 受此问题困扰许久. 以下我们简单地说明这是不可能的.

命题 7.2.17 *考虑 $\mathbb{R}^3$ 上自然的加法, 则不存在乘法使得它具有某种域 (或者体) 的结构.*

证明: 假设 $\mathbb{R}^3$ 上存在某种乘法使得它具有域 (或者体) 的结构. 考虑实线性代数同态 $\begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & L(\mathbb{R}^3) \\ \alpha & \longmapsto & (L_\alpha \colon \beta \to \alpha\beta) \end{array}$, 这是一个单射, 它的像 $W \subseteq L(\mathbb{R}^3)$ 是一个线性子空间, 满足 $\mathrm{id}_{\mathbb{R}^3} \in W$, 且 W 中非零元均可逆. 断言: $W = \mathrm{Span}_{\mathbb{R}}(\{\mathrm{id}_{\mathbb{R}^3}\})$. (这是因为, 任取 $T \in W$, 由 $W \subseteq L(\mathbb{R}^3)$ 知, T 存在特征值 $c \in \mathbb{R}$, 则 $T - c\mathrm{id}_{\mathbb{R}^3} \in W$ 为不可逆的, 故 $T - c\mathrm{id}_{\mathbb{R}^3} = 0$, 即 $T = c\mathrm{id}_{\mathbb{R}^3} \in \mathrm{Span}_{\mathbb{R}}(\{\mathrm{id}_{\mathbb{R}^3}\})$.) 于是比较维数即知矛盾! □

最后我们给出一般实正交变换的正交相似标准形.

定理 7.2.18 (实正交阵的正交相似标准形) *任意实正交阵都正交相似于准对角阵 $\mathrm{diag}(A_1, \cdots, A_s)$, 其中 $A_i = \pm 1$ 或 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ ($\theta \in \mathbb{R} \setminus \mathbb{Z} \cdot \pi$).*

证明: 注意实方阵总有一维或二维不变子空间, 而正交变换的不变子空间的正交补空间仍为不变子空间, 故对维数归纳即知结论. □

注:

(1) 在实正交阵的正交相似标准形中, 对角分量 1 代表了不变方向, 对角分量 -1 代表了关于此方向反射, 对角块
$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ $(\theta\in\mathbb{R}\setminus\mathbb{Z}\cdot\pi)$ 代表了在此二维平面上旋转 θ.

(2) 特别地, 实正交阵若可正交相似于 $\mathrm{diag}(I_{n-1},-1)$, 则称为**镜面反射** (reflection). 取 $\eta\in\mathbb{R}^n$ 为特征值 -1 对应的单位特征向量, 则该镜面反射的表达式为 $T\colon \mathbb{R}^n\longrightarrow\mathbb{R}^n,\ \alpha\longmapsto\alpha-2\langle\eta,\alpha\rangle\cdot\eta$. 由此可直接验证, $\mathbb{R}^n$ 中的任意两个单位向量总可由某个镜面反射将一个映为另一个.

推论 7.2.19 (Cartan-Dieudonne) *任意实正交阵总可写成 $(n-k)$ 个镜面反射的乘积, 其中 k 为特征值 1 的重数.*

证明: 可不妨设该实正交阵已为正交相似标准形, 其中每个对角分量 -1 对应一个镜面反射, 每个对角块
$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\cdot\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$ $(\theta\in\mathbb{R}\setminus\mathbb{Z}\cdot\pi)$ 对应两个镜面反射的乘积. □

需要特别注意的是, 无穷维 Hilbert 空间上的酉算子未必满足条件 "它的不变子空间的正交补仍为不变子空间". 反例如 $(\ell^2(\mathbb{Z}),\langle\cdot,\cdot\rangle)$ 上的右移算子, 此时 $\ell^2(\mathbb{N})$ 为不变子空间, 但它的正交补 $\ell^2(\mathbb{Z}\setminus\mathbb{N})$ 不为不变子空间.

参考文献与补注 7.2

(1)

(2)

(3)

§7.3 正规算子与谱理论

为统一地研究 Hilbert 空间上的连续线性变换, 我们引入正规算子的概念和谱分解性质.

7.3.1 正规算子的算子性质

定义 7.3.1 (正规算子) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 若 $T\circ T^*=T^*\circ T$, 则称 T 为**正规算子** (normal operator).

由正规算子的定义, 我们可以提炼一些简单的判别准则, 并发现它的伴随变换的核与像恰等于自身的核与像.

引理 7.3.1 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 Hilbert 空间, $T\in L(V)$ 为连续的, 则以下条件等价:*

(1) *T 为正规算子;*

(2) *T^* 为正规算子;*

(3) $\forall\alpha\in V,\ \|T(\alpha)\|=\|T^*(\alpha)\|$.

证明: "(1)⇔(2)": 由 $(T^*)^*=T$ 知显然;

"(1)⇒(3)": 设 $T\circ T^*=T^*\circ T$, 则 $\forall\alpha\in V,\ \langle T^*(\alpha),T^*(\alpha)\rangle=\langle T(T^*(\alpha)),\alpha\rangle=\langle T^*(T(\alpha)),\alpha\rangle=\langle T^*(\alpha),T^*(\alpha)\rangle$, 即 $\|T^*(\alpha)\|=\|T(\alpha)\|$.

"(3)⇒(1)": 设 $\forall\alpha\in V,\ \|T(\alpha)\|=\|T^*(\alpha)\|$, 则平方展开知, $\langle(T\circ T^*-T^*\circ T)\alpha,\alpha\rangle=0$. 记 $S=T\circ T^*-T^*\circ T$, 则由 $\forall\alpha,\beta\in V,\ \langle S(\alpha+\beta),\alpha+\beta\rangle=0$ 展开知, $\langle S(\alpha),\beta\rangle+\langle\alpha,S(\beta)\rangle=0$, 即 $S=-S^*$. 但由定义知 $S=S^*$, 故 $S=0$, 即 $T\circ T^*=T^*\circ T$. □

注: 当 $\dim_F(V)<+\infty$ 时, 条件 (3) 还等价于 (3')"$\forall\alpha\in V,\ \|T(\alpha)\|\geq\|T^*(\alpha)\|$". (事实上, 记 $S=T\circ T^*-T^*\circ T$, 由 (3') 可知 $\forall\alpha\in V,\ \langle S(\alpha),\alpha\rangle\leq 0$, 故 S 的所有特征值非正. 再由 $\mathrm{tr}(S)=0$ 知 $\sigma(S)=\{0\}$. 由 S 自伴知 S 可对角化, 故为 0.) 但当 $\dim_F(V)$ 非有限时, 条件 (3) 严格强于 (3'), 反例如 $(\ell^2(\mathbb{N}),\langle\cdot,\cdot\rangle)$ 上的右移算子, 它的伴随为左移算子, 二者满足 (3') 但不满足 (3).

推论 7.3.2 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 *Hilbert* 空间, $T\in L(V)$ 为正规算子, 则 $\ker(T)=\ker(T^*)$; $\mathrm{Im}(T)=\mathrm{Im}(T^*)$.
证明: 由引理知, $\forall\alpha\in V$, $\|T(\alpha)\|=\|T^*(\alpha)\|$, 则 $T(\alpha)=0\iff T^*(\alpha)=0$, 即 $\ker(T)=\ker(T^*)$. 另外, 由此可知 $\overline{\mathrm{Im}(T)}=\ker(T^*)^\perp=\ker(T)^\perp=\overline{\mathrm{Im}(T^*)}$. 现考虑正交分解 $V=\ker(T)\oplus\overline{\mathrm{Im}(T^*)}=\ker(T^*)\oplus\overline{\mathrm{Im}(T)}$, 以及 $T|_{\overline{\mathrm{Im}(T)}}\in L(\overline{\mathrm{Im}(T)})$ 仍为正规算子. 由于 $\ker\left(T|_{\overline{\mathrm{Im}(T)}}\right)=\{0\}$, $\mathrm{Im}\left(T|_{\overline{\mathrm{Im}(T)}}\right)=\mathrm{Im}(T)$, 且 $\left(T|_{\overline{\mathrm{Im}(T)}}\right)^*=T^*|_{\overline{\mathrm{Im}(T)}}$, 则由 $T|_{\overline{\mathrm{Im}(T)}}$ 替代 T 知, 可不妨设 $\ker(T)=\ker(T^*)=\{0\}$, $\overline{\mathrm{Im}(T)}=\overline{\mathrm{Im}(T^*)}=V$.

此时令 $U\coloneqq T^*\circ T^{-1}\colon\mathrm{Im}(T)\to\mathrm{Im}(T^*)$, 则 U 为稠定的保距算子, 故可连续延拓为 V 上的酉算子, 仍记为 U. 由 $U\circ T=T^*$ 知, $T^*\circ U^*=T$, 即 $T^*=T\circ U$, 故 $\mathrm{Im}(T^*)=\mathrm{Im}(T)$. □

例 7.3.1 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 *Hilbert* 空间, $T_1,T_2\in L(V)$ 为正规算子, 则 $T_1\circ T_2=0\iff T_2\circ T_1=0$.
证明: 由引理 7.2.4与推论 7.3.2知, $T_1\circ T_2=0\iff\mathrm{Im}(T_2)\subseteq\ker(T_1)\iff\mathrm{Im}(T_2^*)\subseteq\ker(T_1^*)\iff\ker(T_2)^\perp\subseteq\mathrm{Im}(T_1)^\perp$ $\iff\overline{\mathrm{Im}(T_1)}\subseteq\ker(T_2)\iff\mathrm{Im}(T_1)\subseteq\ker(T_2)\iff T_2\circ T_1=0$. □

我们已经看到, 无穷维 Hilbert 空间上的正规算子 (甚至是酉算子) 未必满足条件 "它的不变子空间的正交补仍为不变子空间". 但在有限维复内积空间上, 正规算子恰好可由这一性质刻画.

引理 7.3.3 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T\in L(V)$, 则以下条件等价:
(1) T 为正规算子;
(2) 任意 T-不变子空间的正交补仍为 T-不变子空间;
(3) 任意 T-不变子空间也是 T^*-不变子空间;
(4) 若 $\alpha\in V$, $c\in\mathbb{C}$ 满足 $T(\alpha)=c\cdot\alpha$, 则 $T^*(\alpha)=\overline{c}\cdot\alpha$;
(5) T 在标准正交基下的矩阵表示可酉相似于对角形.

证明: "(1)⇒(2)⇒(3)": 安师讲义已证;
"(3)⇒(4)": 设 $\alpha\in V$, $c\in\mathbb{C}$ 满足 $T(\alpha)=c\cdot\alpha$. 若 $\alpha=0$, 则结论显然; 若 $\alpha\neq0$, 取 $W=\mathrm{Span}_{\mathbb{C}}(\{\alpha\})$ 为 T-不变子空间, 由 (3) 知 W 也为 T^*-不变子空间, 即 $\exists c'\in F$, $s.t.\ T^*(\alpha)=c'\cdot\alpha$. 注意 $c\|\alpha\|^2=\langle c\cdot\alpha,\alpha\rangle=\langle T(\alpha),\alpha\rangle=\langle\alpha,T^*(\alpha)\rangle$ $=\langle\alpha,c'\cdot\alpha\rangle=\overline{c'}\|\alpha\|^2$, 故 $c=\overline{c'}$.
"(4)⇒(5)": 当 $\dim_F(V)=0$ 时结论显然; 当 $\dim_F(V)=1$ 时, 取 T 的一个单位特征向量 α, 并延拓为 V 的一组标准正交基 B. 由条件知, $[T]_B$ 与 $[T^*]_B$ 的第一列都仅有第一个位置非零. 又 $[T]_B^*=[T^*]_B$, 则 $[T]_B$ 的第一行也仅有第一个位置非零. 因此 $\{\alpha\}^\perp$ 也为 T-不变子空间, 故由维数归纳即知结论.
"(5)⇒(1)": 显然. □

注: 上述引理对于实正规算子不成立, 反例如 T 在实二维空间的标准正交基下矩阵表示为 $\begin{pmatrix}0&-1\\2&0\end{pmatrix}$, 则 T 非正规算子, 且 T 无非平凡的不变子空间, 故条件 (2),(3),(4) 总成立. 但由 "(1)⇒(5)" 的过程, 完全类似可证明: 实正规阵可正交相似于准对角形, 且每个对角块至多为 2 阶的.

推论 7.3.4 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T,S\in L(V)$ 满足 $T\circ S=S\circ T$, 若 T 正规, 则 $T^*\circ S=S\circ T^*$.
证明: 由于复正规阵可酉相似于对角形, 故可不妨设 T 已对角. 记 T 的矩阵表示为 $\mathrm{diag}(a_1,\cdots,a_n)$, S 的矩阵表示为 $(b_{ij})_{1\leq i,j\leq n}$, 则由 $\forall1\leq i,j\leq n$, $a_ib_{ij}=b_{ij}a_j$ 知, $a_i=a_j$ 或 $b_{ij}=0$, 故 $\overline{a_i}b_{ij}=b_{ij}\overline{a_j}$ 总成立, 即 $T^*\circ S=S\circ T^*$. □

推论 7.3.5 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T_1,T_2,S\in L(V)$ 满足 $T_1\circ S=S\circ T_2$, 若 T_1,T_2 正规, 则 $T_1^*\circ S=S\circ T_2^*$.
证明: 记 $\widetilde{T}=\begin{pmatrix}T_1&0\\0&T_2\end{pmatrix}$, $\widetilde{S}=\begin{pmatrix}0&S\\0&0\end{pmatrix}$, 则 $\widetilde{T},\widetilde{S}\in L(V\oplus V)$ 满足 $\widetilde{T}\circ\widetilde{S}=\widetilde{S}\circ\widetilde{T}$, 且 $\widetilde{T}$ 正规, 故由推论 7.3.4知, $\widetilde{T}^*\circ\widetilde{S}=\widetilde{S}\circ\widetilde{T}^*$, 即 $T_1^*\circ S=S\circ T_2^*$. □

注: 上述推论称为 Fuglede-Putnam-Rosenblum 定理, 它对于复 Hilbert 空间上的正规算子也成立. 此时证明仅用到了复分析中的 Liouville 定理, 可以参考 M. Rosenblum"On a Theorem of Fuglede and Putnam"(1958).

例 7.3.2 (Fuglede 定理的错误版本) 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T_1,T_2\in L(V)$,

(1) *若 $S\in L(V)$ 为正规算子, 满足 $T_1\circ S=S\circ T_2$, 则未必有 $T_1\circ S^*=S^*\circ T_2$;*
(2) *若 $S\in L(V)$ 为正规算子, 满足 $T_1\circ S=S\circ T_2$ 且 $S\circ T_1=T_2\circ S$, 则 $T_1\circ S^*=S^*\circ T_2$ 且 $S^*\circ T_1=T_2\circ S^*$.*

证明: (1) 设在 V 的标准正交基下, T_1 的矩阵表示为 $\begin{pmatrix}0 & i\\ i+1 & 0\end{pmatrix}$, T_2 的矩阵表示为 $\begin{pmatrix}0 & i+1\\ i & 0\end{pmatrix}$, S 的矩阵表示为 $\begin{pmatrix}i & 0\\ 0 & i+1\end{pmatrix}$, 则此时 S 正规, $T_1\circ S=S\circ T_2$ 但 $T_1\circ S^*\neq S^*\circ T_2$.

(2) 记 $\widetilde{T}=\begin{pmatrix}T_1 & 0\\ 0 & T_2\end{pmatrix}$, $\widetilde{S}=\begin{pmatrix}0 & S\\ S & 0\end{pmatrix}$, 则 $\widetilde{T},\widetilde{S}\in L(V\oplus V)$ 满足 $\widetilde{T}\circ\widetilde{S}=\widetilde{S}\circ\widetilde{T}$, 且 $\widetilde{S}$ 正规, 故由推论 7.3.4知, $\widetilde{T}\circ\widetilde{S}^*=\widetilde{S}^*\circ\widetilde{T}$, 即 $T_1\circ S^*=S^*\circ T_2$ 且 $S^*\circ T_1=T_2\circ S^*$. □

推论 7.3.6 *设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T_1,T_2\in L(V)$ 满足 $T_1\circ T_2=T_2\circ T_1$, 若 T_1,T_2 正规, 则 $T_1\circ T_2$ 也正规.*

注: 由于上述推论对于复 Hilbert 空间上的正规算子也成立, 在泛函分析中, 固定 Hilbert 空间 $(V,\langle\cdot,\cdot\rangle)$ 上的一个正规算子 T, 记 $\mathscr{A}_T\coloneqq\overline{\{P(T,T^*)\colon P(X,Y)\in\mathbb{C}[X,Y]\}}$ 为由 T 生成的最小闭 C^*-代数, 其中 $*$ 取关于内积的伴随, 则 $\mathscr{A}_T$ 为交换代数, 且 $\mathscr{A}_T$ 中元均为正规算子.

7.3.2 正规算子的谱性质

对于正规算子而言, 最重要的性质就是所谓的谱分解, 这决定了它可以做算符演算. 为简单起见, 我们先对于复正规算子讨论这一特点.

引理 7.3.7 *设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维复内积空间, $T\in L(V)$, 则以下条件等价:*
(1) *T 为正规算子;*
(2) *$T=\sum\limits_{i=1}^{k}c_iE_i$, 其中 $\{c_i\}_{i=1}^k=\sigma(T)$, $\{E_i\}_{i=1}^k$ 为两两正交的正交投影, 且 $\sum\limits_{i=1}^{k}E_i=\mathrm{id}_V$;*
(3) *$\exists g(X)\in\mathbb{C}[X]$, $s.t.$ $\deg(T)\leq|\sigma(T)|-1$, $T^*=g(T)$;*

证明: "(1)⇒(2)": 由复正规阵可酉相似于对角形即知;
"(2)⇒(3)": 由 Lagrange 插值定理即知;
"(3)⇒(1)": 显然. □

注: 条件 (2) 中的正交投影算子族 $\{E_i\}_{i=1}^k$ 由 T 唯一决定, 称为 T 的谱投影算子族. 事实上, 它们都是 T 的多项式: 取 $f_i(X)\in\mathbb{C}[X]$ 满足 $f_i(c_j)=\delta_{ij}$, $\forall 1\leq i,j\leq k$, 则 $E_i=f_i(T)$, $\forall 1\leq i\leq k$. 一般地, 对于复 Hilbert 空间上的正规算子, 可以构造相应的谱族, 使得它写成关于谱测度的积分形式.

现设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$ 为自伴 (若 $F=\mathbb{R}$) 或正规 (若 $F=\mathbb{C}$) 算子. 记 T 的谱分解为 $T=\sum\limits_{i=1}^{k}c_iE_i$, 考虑对应

$$\begin{aligned}\{(S,\phi)\colon \sigma(T)\subseteq S\subseteq F\text{为集合},\ \phi\colon S\to F\text{为函数}\}&\longrightarrow L(V)\\ (S,\phi)&\longmapsto\phi(T)\coloneqq\sum_{i=1}^{k}\phi(c_i)E_i\end{aligned},$$

这称为算子 T 的**算符演算** (calculus of operators). 利用算符演算, 许多数域运算的性质可迁移为线性变换的性质. 例如, 对于 (S_1,ϕ_1), (S_2,ϕ_2) 如上, 以及 $c\in F$,
(1) 可定义 $(S_1,\phi_1)+(S_2,\phi_2)\coloneqq(S_1\cap S_2,\phi_1+\phi_2)$, 对应的算子为 $\phi_1(T)+\phi_2(T)=(\phi_1+\phi_2)(T)$;
(2) 可定义 $c\cdot(S_1,\phi_1)\coloneqq(S_1,c\cdot\phi_1)$, 对应的算子为 $c\cdot\phi_1(T)=(c\cdot\phi_1)(T)$;
(3) 可定义 $(S_1,\phi_1)\cdot(S_2,\phi_2)\coloneqq(S_1\cap S_2,\phi_1\cdot\phi_2)$, 对应的算子为 $\phi_1(T)\circ\phi_2(T)=(\phi_1\cdot\phi_2)(T)$.
另外, $\mathrm{id}_{\sigma(T)}(T)=T$; $\overline{\mathrm{id}_{\sigma(T)}}(T)=T^*$ ($\overline{\mathrm{id}_{\sigma(T)}}$ 为复共轭); $1(T)=\mathrm{id}_V$. 因此上述算符演算具有近似同态的性质.

推论 7.3.8 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$ 为自伴 (若 $F=\mathbb{R}$) 或正规 (若 $F=\mathbb{C}$) 算子. 若 $S\in L(V)$ 满足 $T\circ S=S\circ T$, 则任取 $\phi\colon\sigma(T)\to F$ 为函数, $\phi(T)\circ S=S\circ\phi(T)$.*

证明: 记 T 的谱分解为 $T=\sum\limits_{i=1}^{k}c_iE_i$, 其中 E_i $(1\leq i\leq k)$ 都是 T 的多项式, 则由 $T\circ S=S\circ T$ 知, $E_i\circ S=S\circ E_i$, $\forall 1\leq i\leq k$, 故由 $\phi(T)\coloneqq\sum\limits_{i=1}^{k}\phi(c_i)E_i$ 知, $\phi(T)\circ S=S\circ\phi(T)$. □

例 7.3.3 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$, 由极分解知存在酉算子 $U\in L(V)$, 以及非负自伴算子 $|T|\in L(V)$, 满足 $T=U\circ|T|$. 注意 $|T|=\sqrt{T^*\circ T}=\sqrt{\mathrm{id}_{\sigma(T)}\cdot\mathrm{id}_{\sigma(T)}}(T)$, 则由上述推论知, $C(T)\subseteq C(|T|)$. 进一步地, 若 T 可逆, 或 T 自伴 (若 $F=\mathbb{R}$), 或 T 正规 (若 $F=\mathbb{C}$), 则 (可适当选取)U 满足 $C(T)\subseteq C(U)$.

事实上, 若 T 可逆, 则 $|T|$ 也可逆, 此时 $U=T\circ|T|^{-1}$ 唯一, 显然 $C(T)\subseteq C(U)$. 若 T 自伴 (若 $F=\mathbb{R}$) 或正规 (若 $F=\mathbb{C}$), 则可令 ϕ: $\sigma(T)\longrightarrow F$, 选取 $U=\phi(T)$. 由 $\sigma(U)\subseteq S^1$ 知 U 为酉算子; 由 $\mathrm{id}_{\sigma(T)}=\phi\cdot|\cdot|$

$$c_i=0\longmapsto 1$$
$$c_i\neq 0\longmapsto c_i/|c_i|$$

知 $T=U\circ|T|$; 最后由上述推论知 $C(T)\subseteq C(U)$.

7.3.3 正交相似与酉相似

本节我们讨论方阵的正交相似与酉相似的性质, 主要思想是利用可逆阵的 QR 分解或极分解, 将普通相似的性质加强为正交相似或酉相似的性质. 首先我们给出正交相似或酉相似的一个必要条件.

命题 7.3.9 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$, 若 A,B 正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$), 则 $\mathrm{tr}(\overline{A}^tA)=\mathrm{tr}(\overline{B}^tB)$.

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全同理. 设 $O\in\mathrm{O}(n)$ 满足 $A=OBO^{-1}$, 则 $A^tA=(OBO^{-1})^t(OBO^{-1})=O(B^tB)O^{-1}$, 故 $\mathrm{tr}(A^tA)=\mathrm{tr}(B^tB)$. □

注:

(1) 对于 $A=(A_{ij})_{1\le i,j\le n}\in F^{n\times n}$, $\mathrm{tr}(\overline{A}^tA)=\sum\limits_{1\le i,j\le n}|A_{ij}|^2$ 的算术平方根称为 A 的 Frobenius 范数, 于是上述引理表明正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 不改变方阵的 Frobenius 范数. 由此可发现: 正交相似 (或酉相似) 严格强于相似且相合 (或 $*$-相合); 反例如 $\begin{pmatrix}0&1\\0&0\end{pmatrix}$ 与 $\begin{pmatrix}0&2\\0&0\end{pmatrix}$, 它们相似且相合 (或 $*$-相合), 但不正交相似 (或酉相似).

(2) 进一步地, 对于 $A,B\in F^{n\times n}$, A,B 正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) $\Longleftrightarrow\mathrm{tr}(W(A,\overline{A}^t))=\mathrm{tr}(W(B,\overline{B}^t))$,
$\forall W(X,Y)\in F[X,Y]$ (这里 $XY\neq YX$). 可以参考 W. Specht "Zur Theorie der Matrizen II"(1940).

推论 7.3.10 (Schur 不等式)

(1) 设 $A\in\mathbb{R}^{n\times n}$ 的实特征值为 $c_1,\cdots,c_k$, 则 $\mathrm{tr}(A^tA)\ge\sum\limits_{i=1}^k c_i^2$, 且 "=" 成立 $\Longleftrightarrow A$ 实自伴;

(2) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 的复特征值为 $c_1,\cdots,c_n$, 则 $\mathrm{tr}(\overline{A}^tA)\ge\sum\limits_{i=1}^n|c_i|^2$, 且 "=" 成立 $\Longleftrightarrow A$ 复正规.

证明: 先证明 (2) 中的不等式部分: 由命题 7.3.9知, 只需考虑 A 的酉相似阵即可. 由 Schur 上三角化知, A 酉相似于上三角形. 记 $A=(A_{ij})_{1\le i\le j\le n}$, 则 $\mathrm{tr}(\overline{A}^tA)=\sum\limits_{1\le i\le j\le n}\mathrm{tr}(\overline{A_{ij}}^tA_{ij})\ge\sum\limits_{i=1}^n\mathrm{tr}(\overline{A_{ii}}^tA_{ii})=\sum\limits_{i=1}^n|c_i|^2$. 特别地, (1) 中的不等式也成立.

注意 (2) 中 "=" 成立 $\Longleftrightarrow\forall 1\le i<j\le n$, $A_{ij}=0$, 即 A 酉相似于复对角形, 也即 A 复正规. 特别地, (1) 中 "=" 成立 $\Longleftrightarrow$ (2) 中 "=" 成立且 A 的复特征值均为实数, 即 A 酉相似于实对角形, 即 A 实自伴. □

推论 7.3.11 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为复正规阵, $B\in F^{n\times n}$, 满足 $f_A(X)=f_B(X)$, 且 $\mathrm{tr}(\overline{A}^tA)=\mathrm{tr}(\overline{B}^tB)$, 则 B 复正规, 且正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于 A.

证明: 记 A 的复特征值为 $c_1,\cdots,c_n$. 由 $f_A(X)=f_B(X)$ 知, B 的复特征值为 $c_1,\cdots,c_n$. 由 Schur 不等式的取等条件知, $\mathrm{tr}(\overline{B}^tB)=\mathrm{tr}(\overline{A}^tA)=\sum\limits_{i=1}^n|c_i|^2$, 故 B 复正规. 由于复正规阵可酉相似于对角形, 则 A 酉相似于 $\mathrm{diag}(c_1,\cdots,c_n)$, 也酉相似于 B. 特别地, 若 $A,B\in\mathbb{R}^{n\times n}$, 则由推论 7.3.18知, A 也正交相似于 B. □

例 7.3.4 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为半正定阵, $P\in\mathrm{O}(n)$(若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$(若 $F=\mathbb{C}$), 且 $f_{AP}(X)=f_A(X)$, 则 $AP=PA=A$.

证明: 记 $B=AP\in F^{n\times n}$, 则 $f_A(X)=f_B(X)$, 且 $\mathrm{tr}(\overline{A}^tA)=\mathrm{tr}(\overline{B}^tB)$. 由推论 7.3.11知, B 正交相似 (若

$F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于 A, 故 B 半正定. 考虑极分解 $B=AP\cdot I_n=A\cdot P$, 由唯一性知 $AP=A$. 同理 $PA=A$. □

例 7.3.5 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$ 为正规阵, 则 AB 为正规阵 $\Longleftrightarrow BA$ 为正规阵.*
证明: 注意到 $f_{AB}(X)=f_{BA}(X)$, 且 $\mathrm{tr}(\overline{AB}^t\cdot AB)=\mathrm{tr}(\overline{B}^t(\overline{A}^tA)B)=\mathrm{tr}((\overline{A}^tA)(B\overline{B}^t))=\mathrm{tr}((A\overline{A}^t)(\overline{B}^tB))$ $=\mathrm{tr}(\overline{A}^t(\overline{B}^tB)A)=\mathrm{tr}(\overline{BA}^t\cdot BA)$, 故由推论 7.3.11即知结论. □

利用 QR 分解, 我们可以将同时上三角化的命题加强为同时正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 上三角化的命题.

命题 7.3.12 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $\{A_i\}_{i\in I}\subseteq F^{n\times n}$ 两两可交换,*
(1) *若 $F=\mathbb{C}$, 则 $\{A_i\}_{i\in I}$ 可同时酉相似于上三角形;*
(2) *若 $F=\mathbb{R}$, 则 $\{A_i\}_{i\in I}$ 可同时正交相似于准上三角形, 且对角块为至多二阶的.*
证明: 由同时相似上三角化的命题与 QR 分解即知. □

类似地利用极分解, 我们也可以将同时对角化的命题转化为同时正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 对角化的命题.

引理 7.3.13 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $\{A_i\}_{i\in I}\cup\{B_i\}_{i\in I}\subseteq F^{n\times n}$,*
(1) *若 $F=\mathbb{R}$, 则 $\exists O\in\mathrm{O}(n)$, $s.t.\ O^{-1}A_iO=B_i$, $\forall i\in I\Longleftrightarrow\exists P\in\mathrm{GL}(n,\mathbb{R})$, $s.t.\begin{cases}P^{-1}A_iP=B_i\\P^{-1}A_i^tP=B_i^t\end{cases},\forall i\in I$;*
(2) *若 $F=\mathbb{C}$, 则 $\exists U\in\mathrm{U}(n)$, $s.t.\ U^{-1}A_iU=B_i$, $\forall i\in I\Longleftrightarrow\exists P\in\mathrm{GL}(n,\mathbb{C})$, $s.t.\begin{cases}P^{-1}A_iP=B_i\\P^{-1}\overline{A_i}^tP=\overline{B_i}^t\end{cases},\forall i\in I$.*
证明: 以 (1) 为例, (2) 完全同理. "⇒" 显然; 现证 "⇐": 由条件知, $\exists P\in\mathrm{GL}(n,\mathbb{R})$, $s.t.\ PB_iP^{-1}=(PB_i^tP^{-1})^t$, $\forall i\in I$, 即 $(P^tP)B_i=B_i(P^tP)$, $\forall i\in I$. 考虑 P 的极分解 $P=O\cdot N$, 其中 $O\in\mathrm{O}(n)$, N 为实正定阵, 则 $N^2B_i=B_iN^2$, $\forall i\in I$. 由算符演算的推论知, $NB_i=B_iN$, $\forall i\in I$, 故 $OB_iO^{-1}=PN^{-1}B_iNP^{-1}=PB_iP^{-1}=A_i$, $\forall i\in I$. □

推论 7.3.14 *两个实自伴阵若相似则正交相似; 两个复正规阵若相似则酉相似.*
证明: 第一个结论由引理 7.3.13即知; 第二个结论由 Fuglede 定理与引理 7.3.13即知. □

推论 7.3.15 *一族两两交换的实自伴阵可同时正交相似于对角形; 一族两两交换的复正规阵可同时酉相似于对角形.*
证明: 第一个结论由同时对角化的命题与引理 7.3.13即知; 第二个结论由 Fuglede 定理, 同时对角化的命题, 以及引理 7.3.13即知. □

进一步地, 利用行列式与多项式, 我们可以将实方阵的正交相似与酉相似联系起来.

引理 7.3.16 *设 $\{A_i\}_{i\in I}\cup\{B_i\}_{i\in I}\subseteq\mathbb{C}^{n\times n}$, 则以下条件等价:*
(1) *$\exists P\in\mathrm{GL}(n,\mathbb{R})$, $s.t.\ P^{-1}A_iP=B_i$, $\forall i\in I$;*
(2) *$\exists Q\in\mathrm{GL}(n,\mathbb{C})$, $s.t.\begin{cases}Q^{-1}A_iQ=B_i\\Q^{-1}\overline{A_i}Q=\overline{B_i}\end{cases},\forall i\in I$.*
证明: "⇒": 由取共轭知显然;
"⇐": 设 $Q\in\mathrm{GL}(n,\mathbb{C})$, 满足 $\begin{cases}A_iQ=QB_i\\A_i\overline{Q}=\overline{Q}B_i\end{cases},\forall i\in I$. 记 $Q=Q_1+\sqrt{-1}Q_2$, 其中 $Q_1,Q_2\in\mathbb{R}^{n\times n}$, 则

$\begin{cases}A_iQ_1=Q_1B_i\\A_iQ_2=Q_2B_i\end{cases},\forall i\in I$, 故 $\forall i\in I$, $\forall c\in\mathbb{R}$, $A_i(Q_1+cQ_2)=(Q_1+cQ_2)B_i$. 考虑多项式函数 $f\colon\begin{array}{ccc}\mathbb{C}&\longrightarrow&\mathbb{C}\\c&\longmapsto&\det(Q_1+cQ_2)\end{array}$,

由 $f(\sqrt{-1})\neq0$ 知 f 不为零多项式, 故 f 的零点数量有限; 特别地, $\exists c\in\mathbb{R}$, $s.t.\ f(c)\neq0$, 即 $Q_1+cQ_2\in\mathrm{GL}(n,\mathbb{R})$, 则 $P\coloneqq Q_1+cQ_2\in\mathrm{GL}(n,\mathbb{R})$ 满足 $P^{-1}A_iP=B$, $\forall i\in I$. □

推论 7.3.17 设 $\{A_i\}_{i\in I}\cup\{B_i\}_{i\in I}\subseteq\mathbb{R}^{n\times n}$, 则以下条件等价:

(1) $\exists P\in\mathrm{GL}(n,\mathbb{R})$, $s.t.$ $P^{-1}A_iP=B_i$, $\forall i\in I$;

(2) $\exists Q\in\mathrm{GL}(n,\mathbb{C})$, $s.t.$ $Q^{-1}A_iQ=B_i$, $\forall i\in I$.

推论 7.3.18 一族实方阵若可同时酉相似于另一族实方阵, 则也可同时正交相似于它们.

证明: 由引理 7.3.13与推论 7.3.17即知. □

推论 7.3.19 两个实正规阵若相似则正交相似.

证明: 由推论 7.3.14与推论 7.3.18即知. □

习题 7.3 设 $A,B\in\mathbb{C}^{n\times n}$, 且 $A\cdot\overline{A}=B\cdot\overline{B}=I_n$, 则

$$\exists P\in\mathrm{GL}(n,\mathbb{R}),\ s.t.\ P^{-1}AP=B\Longleftrightarrow\exists Q\in\mathrm{GL}(n,\mathbb{C}),\ s.t.\ Q^{-1}AQ=B.$$

参考文献与补注 7.3

(1)

(2)

(3)

§7.4 线性算子的分析性质

7.4.1 线性算子的数值范围

定义 7.4.1 (数值范围与数值半径) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, $T\in L(V)$ 为连续的, 则 $W(T)\coloneqq\{\langle T(\alpha),\alpha\rangle\in F\colon\ \alpha\in V,\ \|\alpha\|=1\}$ 称为 T 的**数值范围** (numerical range); $r(T)\coloneqq\sup\limits_{c\in W(T)}|c|$ 称为 T 的**数值半径** (numerical radius).

引理 7.4.1 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{2\times2}$ 且 $\mathrm{tr}(A)=0$, 则 A 可正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{R}$) 于对角分量均为 0 的方阵.

证明: 设 $A=\begin{pmatrix}a&b\\c&-a\end{pmatrix}$ $(a,b,c\in F)$. 若 $a=0$, 则结论显然; 若 $a\neq0$, 则可取 $\omega\in\mathbb{R}$ 满足 $\dfrac{ce^{i\omega}+be^{-i\omega}}{a}\in\mathbb{R}$, 再取 $\theta\in\mathbb{R}$ 满足 $a\cos(2\theta)+\dfrac{ce^{i\omega}+be^{-i\omega}}{2}\sin(2\theta)=0$. 记 $P=\begin{pmatrix}\cos(\theta)e^{i\omega}&-\sin(\theta)\\\sin(\theta)&\cos(\theta)e^{-i\omega}\end{pmatrix}$, 则 $P\in\mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), 且 $P^{-1}AP=\begin{pmatrix}0&*\\*&0\end{pmatrix}$. □

命题 7.4.2 (Hausdorff-Toeplitz) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, $T\in L(V)$, 则 $W(T)\subseteq F$ 为凸集.

证明: 任取 $c_1\neq c_2\in W(T)$, 则 $\exists\alpha_1,\alpha_2\in V$, $s.t.$ $\|\alpha_1\|=\|\alpha_2\|=1$, 且 $c_1=\langle T(\alpha_1),\alpha_1\rangle$, $c_2=\langle T(\alpha_2),\alpha_2\rangle$. 假设 $\{\alpha_1,\alpha_2\}$ 线性相关, 则 $\exists c\in F$, $s.t.$ $|c|=1$ 且 $\alpha_1=c\alpha_2$, 故 $c_1=\langle T(c\alpha_2),c\alpha_2\rangle=|c|^2\langle T(\alpha_2),\alpha_2\rangle=c_2$, 矛盾! 因此 $\{\alpha_1,\alpha_2\}$ 线性无关. 记 $W\coloneqq\mathrm{Span}_F(\{\alpha_1,\alpha_2\})\subseteq V$, 则 $\dim_F(W)=2$. 由引理 7.1.14知, 可取正交投影 $P_W\colon V\to W$, 则 $A\coloneqq P_W\circ T\big|_W\in L(W)$, 且 $c_1=\langle A(\alpha_1),\alpha_1\rangle$, $c_2=\langle A(\alpha_2),\alpha_2\rangle$. 显然 $W(A)\subseteq W(T)$, 故以下只需证明: $\forall\lambda\in(0,1)$, $(1-\lambda)c_1+\lambda c_2\in W(A)$, 也只需证明: $W(A)$ 为凸集.

注意 $W(A)=W(A-\dfrac{\mathrm{tr}(A)}{2}\mathrm{id}_W)+\dfrac{\mathrm{tr}(A)}{2}$, 则由平移不变性知, 可不妨设 $\mathrm{tr}(A)=0$. 由引理 7.4.1, 可取 W 的一组标准正交基, 使得 A 的矩阵表示为 $\begin{pmatrix}0&b\\c&0\end{pmatrix}$ $(b,c\in F)$, 则 $W(A)=\{bx_2\overline{x_1}+cx_1\overline{x_2}\colon x_1,x_2\in F,\ |x_1|^2+|x_2|^2=1\}$. 由于 $\{x_2\overline{x_1}\colon x_1,x_2\in F,\ |x_1|^2+|x_2|^2=1\}=\{x\in F\colon|x|\leq1/2\}$ 为凸集, 则 $W(A)=\{bx+c\overline{x}\colon x\in F,\ |x|\leq1/2\}$ 也为凸集. □

回忆引理 2.2.1及其推论 2.2.2, 利用内积我们试图将它加强为正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 的命题.

命题 7.4.3 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则域 F 上的迹为 0 的方阵总可正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于某个对角分量均为 0 的方阵.*

证明: 我们对 n 归纳证明结论. 当 $n=1$ 时, 迹为 0 的方阵为 (0), 结论显然; 假设 $n\geq 2$ 且当 $(n-1)$ 时结论成立, 设 $A\in F^{n\times n}$ 满足 $\mathrm{tr}(A)=0$. 断言: $\exists \alpha\in F^{n\times 1}\setminus\{0\}$, *s.t.* $\langle A\alpha,\alpha\rangle=0$. (这是因为, 由 Hausdorff-Toeplitz 定理知,

$W(A)$ 为凸集; 又 $\sigma(A)\subseteq W(A)$, 则 $\mathrm{conv}(\sigma(A))\subseteq W(A)$, 故 $0=\dfrac{\mathrm{tr}(A)}{n}\in\mathrm{conv}(\sigma(A))\subseteq W(A)$, 即 $\exists\alpha\in F^{n\times 1}\setminus\{0\}$, *s.t.* $\langle A\alpha,\alpha\rangle=0$.)

不妨设 $\|\alpha\|=1$. 现将 $\{\alpha\}$ 扩充为 $F^{n\times 1}$ 的标准正交基 B, 记 $P_0=(\alpha,\cdots)$ 为 B 中元以列向量方式排成的方阵, 则 $P_0\in \mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), 且 $P_0^{-1}AP_0=\begin{pmatrix}0 & *\\ * & A_1\end{pmatrix}$, 其中 $\mathrm{tr}(A_1)=\mathrm{tr}(P_0^{-1}AP_0)=\mathrm{tr}(A)=0$. 由归纳假设知, $\exists Q_0\in\mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), *s.t.* $Q_0^{-1}A_1Q_0=A_1'$, 这里 A_1' 的对角分量均为 0, 则 $\begin{pmatrix}1 & 0\\ 0 & Q_0\end{pmatrix}^{-1}\cdot P_0^{-1}AP_0\cdot\begin{pmatrix}1 & 0\\ 0 & Q_0\end{pmatrix}=\begin{pmatrix}0 & *'\\ *' & A_1'\end{pmatrix}=:A'$, 这里 A' 的对角分量均为 0. □

推论 7.4.4 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则域 F 上的方阵总可正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于某个对角分量均相等的方阵.*

命题 7.4.5 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$ 为实自伴 (若 $F=\mathbb{R}$) 或复正规 (若 $F=\mathbb{C}$)*

算子, 则 $W(T)=\mathrm{conv}(\sigma(T))$.

证明: 由实自伴或复正规算子的谱分解定理知, 可取 $(V,\langle\cdot,\cdot\rangle)$ 的标准正交基 $\{\alpha_i\}_{i=1}^n$, 满足 $T(\alpha_i)=c_i\alpha_i$, $\forall 1\leq i\leq n$, 则 $W(T)=\left\{\sum\limits_{i=1}^n c_i|x_i|^2\colon x_i\in F,\ \sum\limits_{i=1}^n|x_i|^2=1\right\}=\mathrm{conv}(\sigma(T))$. □

推论 7.4.6 (Rayleigh) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的有限维内积空间, $T\in L(V)$ 为自伴算子, 则* $\min W(T)=\min\sigma(T)$; $\max W(T)=\max\sigma(T)$.

命题 7.4.7 (Poincare) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 n 维内积空间, $T\in L(V)$ 为自伴算子, 记 T 的特征值为 $c_1\geq\cdots\geq c_n$, 则任取 $W\subseteq V$ 为 $1\leq k\leq n$ 维线性子空间, 存在 $\alpha,\beta\in W$, 满足 $\|\alpha\|=\|\beta\|=1$, 且 $\langle T(\alpha),\alpha\rangle\leq c_k$; $\langle T(\beta),\beta\rangle\geq c_{n+1-k}$.*

证明: 注意到 $-T$ 为自伴算子, 且 $-T$ 的特征值为 $-c_n\geq\cdots\geq -c_1$, 则通过由 $-T$ 代替 T, 只需证明 $\exists\alpha\in W$, *s.t.* $\|\alpha\|=1$, $\langle T(\alpha),\alpha\rangle\leq c_k$. 事实上, 由自伴算子的谱分解知, 存在 V 的一组标准正交基 $\{\alpha_i\}_{i=1}^n$, 满足 $T(\alpha_i)=c_i\alpha_i$, $\forall 1\leq i\leq n$. 记 $U:=\mathrm{Span}_F(\{\alpha_i\}_{i=k}^n)$, 则 $U\subseteq V$ 为 $(n-k+1)$ 维线性子空间, 故 $\dim_F(W\cap U)\geq 1$. 取 $\alpha\in W\cap U$ 满足 $\|\alpha\|=1$, 记 $\alpha=\sum\limits_{i=k}^n a_i\alpha_i$, 则 $\sum\limits_{i=k}^n|a_i|^2=1$, 故 $\langle T(\alpha),\alpha\rangle=\left\langle\sum\limits_{i=k}^n a_ic_i\alpha_i,\sum\limits_{i=k}^n a_i\alpha_i\right\rangle=\sum\limits_{i=k}^n|a_i|^2c_i\leq c_k$. □

推论 7.4.8 (Courant-Fischer) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 n 维内积空间, $T\in L(V)$ 为自伴算子, 记 T 的特征值为 $c_1\geq\cdots\geq c_n$, 则 $\forall 1\leq k\leq n$, $c_k=\max\limits_{\substack{W\subseteq V\\ \dim_F(W)=k}}\min\limits_{\substack{\alpha\in W\\ \|\alpha\|=1}}\langle T(\alpha),\alpha\rangle=\min\limits_{\substack{W\subseteq V\\ \dim_F(W)=n-k+1}}\max\limits_{\substack{\alpha\in W\\ \|\alpha\|=1}}\langle T(\alpha),\alpha\rangle$.*

证明: 注意到 $-T$ 为自伴算子, 且 $-T$ 的特征值为 $-c_n\geq\cdots\geq -c_1$, 则通过由 $-T$ 代替 T, 只需证明 $\forall 1\leq k\leq n$, $c_k=\max\limits_{\substack{W\subseteq V\\ \dim_F(W)=k}}\min\limits_{\substack{\alpha\in W\\ \|\alpha\|=1}}\langle T(\alpha),\alpha\rangle$. 事实上, 由命题 7.4.7即知 "≥" 部分; 另一方面, 取 $W=\mathrm{Span}_F(\{\alpha_i\}_{i=1}^k)\subseteq V$ 为 k 维线性子空间, 则 "=" 可取到. □

注: 若 $T\in L(V)$ 为连续的, 记 T 的奇异值为 $c_1\geq\cdots\geq c_n$, 则

$$\forall 1\leq k\leq n,\ c_k=\max_{\substack{W\subseteq V\\ \dim_F(W)=k}}\min_{\substack{\alpha\in W\\ \|\alpha\|=1}}\|T(\alpha)\|=\min_{\substack{W\subseteq V\\ \dim_F(W)=n-k+1}}\max_{\substack{\alpha\in W\\ \|\alpha\|=1}}\|T(\alpha)\|.$$

推论 7.4.9 (Weyl 不等式) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的 n 维内积空间, $T_1,T_2\in L(V)$ 为自伴算子. 记*

T_1,T_2,T_1+T_2 的特征值分别为 $c_1(T_1)\geq\cdots\geq c_n(T_1)$, $c_1(T_2)\geq\cdots\geq c_n(T_2)$, $c_1(T_1+T_2)\geq\cdots\geq c_n(T_1+T_2)$,

则 $\begin{cases}c_{i+j-1}(T_1+T_2)\leq c_i(T_1)+c_j(T_2), & \forall i,j\geq 1,\ i+j\leq n+1\\ c_{i+j-n}(T_1+T_2)\geq c_i(T_1)+c_j(T_2), & \forall i,j\geq 1,\ i+j\geq n+1\end{cases}$. 进一步地, 给定 $1\leq i,j\leq n$, 上述不等式取

等号当且仅当 T_1, T_2, T_1+T_2 对应的特征值存在公共特征向量.

证明: 通过由 $-T_1, -T_2, -T_1-T_2$ 分别代替 T_1, T_2, T_1+T_2, 只需证明第一个不等式. 事实上, 由自伴算子的谱分解知, 存在 V 的一组标准正交基 $\{\alpha_i\}_{i=1}^n$, 满足 $T_1(\alpha_i) = c_i(T_1)\alpha_i$, $\forall 1 \leq i \leq n$; 也存在 V 的一组标准正交基 $\{\beta_j\}_{j=1}^n$, 满足 $T_2(\beta_j) = c_j(T_2)\beta_j$, $\forall 1 \leq j \leq n$. 任取 $i, j \geq 1$ 满足 $i+j \leq n+1$, 记 $W_1 = \mathrm{Span}(\{\alpha_k\}_{k=i}^n)$, $W_2 = \mathrm{Span}(\{\beta_l\}_{l=j}^n)$, 则 $\dim_F(W_1) = n-i$, $\dim_F(W_2) = n-j$, 故 $\dim_F(W_1 \cap W_2) \geq \dim_F(W_1) + \dim$
取 $U \subseteq W_1 \cap W_2$ 为 $n-(i+j)$ 维线性子空间, 则

$$\forall \alpha \in U\,(\|\alpha\| = 1),\ \langle (T_1+T_2)(\alpha), \alpha\rangle = \langle T_1(\alpha), \alpha\rangle + \langle T_2(\alpha), \alpha\rangle \leq c_i(T_1) + c_j(T_2),$$

故由 Courant-Fischer 定理知, $c_{i+j-1}(T_1+T_2) = \min\limits_{\substack{U\subseteq V\\ \dim_F(U)=n-(i+j)}} \max\limits_{\substack{\alpha\in U\\ \|\alpha\|=1}} \langle (T_1+T_2)(\alpha), \alpha\rangle \leq c_i(T_1) + c_j(T_2)$. □

推论 7.4.10 (Cauchy Interlacing) 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $(V, \langle\cdot,\cdot\rangle)$ 为域 F 上的 n 维内积空间, $T \in L(V)$ 为自伴算子, $P_W \in L(V)$ 为到 m 维子空间 $W \subseteq V$ 上的正交投影, $S = (P_W^* \circ T \circ P_W)\big|_W \in L(W)$. 记 T 的特征值为 $c_1(T) \geq \cdots \geq c_n(T)$, S 的特征值为 $c_1(S) \geq \cdots \geq c_m(S)$, 则 $\forall 1 \leq k \leq m$, $c_k(T) \geq c_k(S) \geq c_{n-m+k}(T)$.

证明: 由 $T \in L(V)$ 自伴知, $S = (P_W^* \circ T \circ P_W)\big|_W \in L(W)$ 也自伴. 由自伴算子的谱分解知, 存在 W 的一组标准正交基 $\{\alpha_i\}_{i=1}^m$, 满足 $S(\alpha_i) = c_i(S)\alpha_i$, $\forall 1 \leq i \leq m$. 任取 $1 \leq k \leq m$, 记 $W_1 = \mathrm{Span}_F(\{\alpha_i\}_{i=1}^k) \subseteq W$ 为 k 维子空间, 则由 Courant-Fischer 定理知, $c_k(T) \geq \min\limits_{\substack{\alpha\in W_1\\ \|\alpha\|=1}} \langle T(\alpha), \alpha\rangle = \min\limits_{\substack{\alpha\in W_1\\ \|\alpha\|=1}} \langle (T\circ P_W)(\alpha), P_W(\alpha)\rangle = \min\limits_{\substack{\alpha\in W_1\\ \|\alpha\|=1}} \langle S(\alpha), \alpha\rangle = c_k(S)$.
另一方面, 记 $W_2 = \mathrm{Span}_F(\{\alpha_i\}_{i=k}^m) \subseteq W$ 为 $m-k+1$ 维子空间, 则由 Courant-Fischer 定理知,

$$c_{n-m+k}(T) \leq \min_{\substack{\alpha\in W_2\\ \|\alpha\|=1}} \langle T(\alpha), \alpha\rangle = \max_{\substack{\alpha\in W_2\\ \|\alpha\|=1}} \langle (T\circ P_W)(\alpha), P_W(\alpha)\rangle = \max_{\substack{\alpha\in W_2\\ \|\alpha\|=1}} \langle S(\alpha), \alpha\rangle = c_k(S).$$

□

7.4.2 矩阵范数与谱半径

在关于内积的伴随小节中, 我们已经由赋范线性空间的范数诱导了连续算子的算子范数. 事实上, 利用矩阵表示的特殊性, 在方阵代数上我们还可以引入更多的范数.

定义 7.4.2 (矩阵范数) 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $\wp$ 为 $F^{n\times n}$ 上的一个向量范数. 若 $\wp$ 满足 $\forall A, B \in F^{n\times n}$, $\wp(AB) \leq \wp(A)\cdot\wp(B)$, 则 $\wp$ 称为 $F^{n\times n}$ 上的一个 **矩阵范数** (matrix norm).

注: 注意矩阵范数未必是关于相似变换不变的, 因此本节仅限于对矩阵表示而非线性变换讨论范数.

例 7.4.1 (诱导范数) 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, p 为 $F^{n\times 1}$ 上的一个向量范数, 则 $\widetilde{p}$: $F^{n\times n} \longrightarrow \mathbb{R}$ 是

$$A \longmapsto \sup_{\alpha\in F^{n\times 1}\setminus\{0\}} \frac{p(A\alpha)}{p(\alpha)}$$

$F^{n\times n}$ 上的一个矩阵范数, 它就是由 p 诱导的算子范数.

(1) 取 p: $F^{n\times 1} \longrightarrow \mathbb{R}$ 为 ℓ^1 范数, 则 $\widetilde{p}$: $F^{n\times n} \longrightarrow \mathbb{R}$ 为最大列和范数.

$$(x_1, \cdots, x_n)^t \longmapsto \sum_{i=1}^n |x_i| \qquad A \longmapsto \max_{1\leq j\leq n} \sum_{i=1}^n |A_{ij}|$$

(2) 取 p: $F^{n\times 1} \longrightarrow \mathbb{R}$ 为 ℓ^2 范数, 则 $\widetilde{p}$: $F^{n\times n} \longrightarrow \mathbb{R}$ 为谱范数.

$$(x_1, \cdots, x_n)^t \longmapsto \left(\sum_{1\leq i\leq n} |x_i|^2\right)^{\frac{1}{2}} \qquad A \longmapsto A\text{的最大奇异值}$$

(3) 取 p: $F^{n\times 1} \longrightarrow \mathbb{R}$ 为 ℓ^∞ 范数, 则 $\widetilde{p}$: $F^{n\times n} \longrightarrow \mathbb{R}$ 为最大行和范数.

$$(x_1, \cdots, x_n)^t \longmapsto \max_{1\leq i\leq n} |x_i| \qquad A \longmapsto \max_{1\leq i\leq n} \sum_{j=1}^n |A_{ij}|$$

引理 7.4.11 设 $F = \mathbb{R}$ 或 $\mathbb{C}$,

(1) 若 $\widetilde{p}$ 为 $F^{n\times n}$ 上的一个诱导矩阵范数, 则 $\widetilde{p}(I_n) = 1$;

(2) 若 $\widetilde{p_1},\widetilde{p_2}$ 为 $F^{n\times n}$ 上的两个诱导矩阵范数, 它们分别由 $F^{n\times 1}$ 上的向量范数 p_1,p_2 诱导, 记

$$M=\sup_{\alpha\in F^{n\times 1}\setminus\{0\}}\frac{p_1(\alpha)}{p_2(\alpha)},\ m=\inf_{\alpha\in F^{n\times 1}\setminus\{0\}}\frac{p_1(\alpha)}{p_2(\alpha)};\ \widetilde{M}=\sup_{A\in F^{n\times n}\setminus\{0\}}\frac{\widetilde{p_1}(A)}{\widetilde{p_2}(A)},\ \widetilde{m}=\inf_{A\in F^{n\times n}\setminus\{0\}}\frac{\widetilde{p_1}(A)}{\widetilde{p_2}(A)},$$

则 $\widetilde{M}=\widetilde{m}^{-1}=Mm^{-1}$.

证明: (1) 由定义显然;

(2)

□

推论 7.4.12 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则 $F^{n\times n}$ 上的诱导矩阵范数是极小的诱导矩阵范数.

例 7.4.2 (非诱导范数) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则以下 $F^{n\times n}$ 上的矩阵范数都不是诱导范数:

(1) 取 $\wp\colon\ F^{n\times n}\longrightarrow\mathbb{R}$ 为 ℓ^1 范数, 则 $\wp$ 为矩阵范数且 $\wp(I_n)=n$.

$$A\longmapsto\sum_{i,j=1}^{n}|A_{ij}|$$

(2) 取 $\wp\colon\ F^{n\times n}\longrightarrow\mathbb{R}$ 为 ℓ^2 范数, 则 $\wp$ 为矩阵范数且 $\wp(I_n)=\sqrt{n}$.

$$A\longmapsto\left(\sum_{i,j=1}^{n}|A_{ij}|^2\right)^{\frac{1}{2}}$$

(3) 取 $\wp\colon\ F^{n\times n}\longrightarrow\mathbb{R}$ 为 ℓ^∞ 范数的 n 倍, 则 $\wp$ 为矩阵范数且 $\wp(I_n)=n$. 注意 ℓ^∞ 范数不是矩阵范数.

$$A\longmapsto n\cdot\max_{1\le i,j\le n}|A_{ij}|$$

引理 7.4.13 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $\wp$ 为 $F^{n\times n}$ 上的一个矩阵范数, $\alpha_0\in F^{n\times 1}\setminus\{0\}$.

(1) 记 $p_{\alpha_0}\colon\ F^{n\times 1}\longrightarrow\mathbb{R}$, 则 p_{α_0} 为 $F^{n\times 1}$ 上的向量范数, 且它诱导的矩阵范数 $\widetilde{p_{\alpha_0}}$ 满足 $\widetilde{p_{\alpha_0}}\le\wp$.

$$\alpha\longmapsto\wp(\alpha\cdot\overline{\alpha_0}^t)$$

(2) 任取 $\widetilde{p}$ 为 $F^{n\times n}$ 上的诱导矩阵范数, 则 $\wp\le\widetilde{p}\Longleftrightarrow\widetilde{p_{\alpha_0}}=\wp=\widetilde{p}$.

证明: (1) 可直接验证 p_{α_0} 满足向量范数的定义. 现任取 $A\in F^{n\times n}$, 则

$$\widetilde{p_{\alpha_0}}(A)=\sup_{\alpha\in F^{n\times 1}\setminus\{0\}}\frac{p_{\alpha_0}(A\alpha)}{p_{\alpha_0}(\alpha)}=\sup_{\alpha\in F^{n\times 1}\setminus\{0\}}\frac{\widetilde{p_0}(A\alpha\cdot\overline{\alpha_0}^t)}{\widetilde{p_0}(\alpha\cdot\overline{\alpha_0}^t)}\le\widetilde{p_0}(A).$$

(2) 由推论 7.4.12与 (1) 即知. □

推论 7.4.14 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 则 $F^{n\times n}$ 上的一个矩阵范数是诱导矩阵范数当且仅当它是极小的矩阵范数.

定义 7.4.3 (谱半径) 设 $A\in\mathbb{C}^{n\times n}$, 则 $\rho(A)\coloneqq\sup\limits_{c\in\sigma(A)}|c|$ 称为 A 的**谱半径** (spectral radius).

注: 方阵的谱半径并不是矩阵范数, 这是因为它不满足次可加性与次可乘性. 例如取 $A=E_{12},B=E_{21}\in\mathbb{C}^{2\times 2}$, 则 $\sigma(A)=\sigma(B)=\{0\}$, 但 $\sigma(A+B)=\{-1,1\}$, $\sigma(AB)=\{0,1\}$.

命题 7.4.15 设 $A\in\mathbb{C}^{n\times n}$, 则 $\rho(A)=\inf\{\wp(A)\colon\wp\text{为矩阵范数}\}=\inf\{\widetilde{p}(A)\colon\widetilde{p}\text{为诱导矩阵范数}\}$.

证明: 一方面, 任取 $\mathbb{C}^{n\times n}$ 上的一个矩阵范数 $\wp$, 由引理 7.4.13知, 存在 $\mathbb{C}^{n\times 1}$ 上的一个向量范数 p_0, 满足它诱导的矩阵范数 $\widetilde{p_0}\le\wp$, 即 $\forall\alpha\in\mathbb{C}^{n\times 1}$, $p_0(A\alpha)\le\wp(A)\cdot p_0(\alpha)$. 特别地, 取 $c\in\sigma(A)$ 满足 $|c|=\rho(A)$, $\alpha\in\mathbb{C}^{n\times 1}$ 为属于 c 的特征向量, 则由范数的绝对齐次性与正定性知, $\rho(A)\le\wp(A)$.

另一方面, 固定 $\epsilon>0$. 由 Jordan 标准形知, $\exists P\in\mathrm{GL}(n,\mathbb{C})$, $s.t.\ P^{-1}AP=J$, 其中 J 为每个 Jordan 块的次对角线上均为 ϵ 的 Jordan 标准形. 取 $\widetilde{p}_1$ 为 $\mathbb{C}^{n\times n}$ 上的最大列和范数, 以及 $\wp\colon\ F^{n\times n}\longrightarrow\mathbb{R}$, 则 $\wp$

$$B\longmapsto\widetilde{p}_1(P^{-1}BP)$$

也为 $F^{n\times n}$ 上的诱导矩阵范数, 且 $\wp(A)=\widetilde{p}_1(J)\le\rho(A)+\epsilon$.

综上可知 $\rho(A)=\inf\{\wp(A)\colon\wp\text{为矩阵范数}\}=\inf\{\widetilde{p}(A)\colon\widetilde{p}\text{为诱导矩阵范数}\}$. □

推论 7.4.16 设 $A\in\mathbb{C}^{n\times n}$, 则 $\lim\limits_{k\to+\infty}A^k=0\Longleftrightarrow\rho(A)<1$.

证明: "⇒": 设 $\lim\limits_{k\to+\infty}A^k=0$. 任取 A 的特征值 $c\in\mathbb{C}$ 以及对应的特征向量 $\alpha\in\mathbb{C}^{n\times 1}$, 则 $\forall k\ge 1$, $A^k\alpha=c^k\alpha$,

故由逐分量收敛于 0 知, $\lim\limits_{k\to+\infty} c^k = 0$, 即 $|c| < 1$. 因此 $\rho(A) < 1$.

"⇐": 设 $\rho(A) < 1$, 则由命题 7.4.15知, 存在 $\mathbb{C}^{n\times n}$ 上的矩阵范数 $\wp$, 满足 $\wp(A) < \rho(A) + (1-\rho(A)) = 1$, 则 $\lim\limits_{k\to+\infty} \wp(A^k) \leq \lim\limits_{k\to+\infty} \wp(A)^k = 0$, 故 $\lim\limits_{k\to+\infty} A^k = 0$. □

推论 7.4.17 设 $A \in \mathbb{C}^{n\times n}$, 则 $\forall \epsilon > 0$, $\exists C = C(A,\epsilon) > 0$, $\forall k \geq 1$, $\max\limits_{1\leq i,j\leq n} |(A^k)_{ij}| \leq C(\rho(A)+\epsilon)^k$.

证明: 任取 $\epsilon > 0$, 记 $\widetilde{A} = \dfrac{1}{\rho(A)+\epsilon} A$. 由于 $\rho(\widetilde{A}) < 1$, 则由推论 7.4.16知, $\lim\limits_{k\to+\infty} \widetilde{A}^k = 0$. 特别地, $\{\widetilde{A}^k\}_{k=1}^{+\infty} \subseteq \mathbb{C}^{n\times n}$ 有界, 则 $\exists C = C(A,\epsilon) > 0$, $\forall k \geq 1$, $\max\limits_{1\leq i,j\leq n} |(\widetilde{A}^k)_{ij}| \leq C$, 即 $\max\limits_{1\leq i,j\leq n} |(A^k)_{ij}| \leq C(\rho(A)+\epsilon)^k$. □

推论 7.4.18 (Gelfand) 设 $A \in \mathbb{C}^{n\times n}$, $\wp$ 为 $\mathbb{C}^{n\times n}$ 上的一个矩阵范数, 则 $\rho(A) = \lim\limits_{k\to+\infty} \wp(A^k)^{\frac{1}{k}}$.

证明: 一方面, 由命题 7.4.15知, $\rho(A) \leq \wp(A)$, 则 $\forall k \geq 1$, $\rho(A) = \rho(A^k)^{\frac{1}{k}} \leq \wp(A^k)^{\frac{1}{k}}$, 则 $\rho(A) \leq \liminf\limits_{k\to+\infty} \wp(A^k)^{\frac{1}{k}}$. 另一方面, 任取 $\epsilon > 0$, 由推论 7.4.17的证明知, $\exists C = C(A,\epsilon) > 0$, $\forall k \geq 1$, $\wp(A^k) \leq C(\rho(A)+\epsilon)^k$, 因此 $\limsup\limits_{k\to+\infty} \wp(A^k)^{\frac{1}{k}} \leq \limsup\limits_{k\to+\infty} C^{\frac{1}{k}}(\rho(A)+\epsilon) = \rho(A)+\epsilon$. 再令 $\epsilon \to 0^+$ 即知 $\limsup\limits_{k\to+\infty} \wp(A^k)^{\frac{1}{k}} \leq \rho(A)$. □

命题 7.4.19 设 $A \in \mathbb{C}^{n\times n}$, 幂级数 $\sum\limits_{k=0}^{+\infty} a_k z^k$ 的收敛半径为 R. 若 $\rho(A) < R$, 则 $\sum\limits_{k=0}^{+\infty} a_k A^k$ 收敛.

证明: 由 $\rho(A) < R$ 以及命题 7.4.15知, 存在 $\mathbb{C}^{n\times n}$ 上的一个矩阵范数 $\wp$, 满足 $\wp(A) < R$. 任取 $M \geq m \geq 0$, 则由矩阵范数的绝对齐次性、次可加性与次可乘性知, $\wp\left(\sum\limits_{k=m}^{M} a_k A^k\right) \leq \sum\limits_{k=m}^{M} |a_k| \wp(A)^k$; 再由幂级数的收敛性质知, $\limsup\limits_{m,M\to+\infty} \wp\left(\sum\limits_{k=m}^{M} a_k A^k\right) \leq \limsup\limits_{m,M\to+\infty} \sum\limits_{k=m}^{M} |a_k| \wp(A)^k = 0$, 则 $\lim\limits_{m,M\to+\infty} \left(\sum\limits_{k=m}^{M} a_k A^k\right) = 0$, 故由 Cauchy 收敛准则知, $\sum\limits_{k=0}^{+\infty} a_k A^k$ 收敛. □

例 7.4.3 设 $A \in \mathbb{C}^{n\times n}$, 则

(1) $\exp(A) \coloneqq \sum\limits_{k=0}^{+\infty} \dfrac{A^k}{k!}$ 总收敛;

(2) $\sin(A) \coloneqq \sum\limits_{k=0}^{+\infty} (-1)^k \dfrac{A^{2k+1}}{(2k+1)!}$ 总收敛; $\cos(A) \coloneqq \sum\limits_{k=0}^{+\infty} (-1)^k \dfrac{A^{2k}}{(2k)!}$ 总收敛;

(3) 当 $\rho(A) < 1$ 时, $\log(I_n - A) \coloneqq -\sum\limits_{k=1}^{+\infty} \dfrac{A^k}{k}$ 收敛; $(I_n - A)^{-1} = \sum\limits_{k=0}^{+\infty} A^k$ 收敛.

习题 7.4 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, 记 $\widetilde{p}_1, \widetilde{p}_2, \widetilde{p}_\infty$ 分别为 $F^{n\times n}$ 上的最大列和范数、谱范数与最大行和范数, $\wp_1, \wp_2, \wp_\infty$ 分别为 $F^{n\times n}$ 上的 ℓ^1 范数、ℓ^2 范数与 ℓ^∞ 范数的 n 倍, 验证以下表格给出了这些矩阵范数之间比值的最佳常数:

/	$\widetilde{p}_1$	$\widetilde{p}_2$	$\widetilde{p}_\infty$	$\wp_1$	$\wp_2$	$\wp_\infty$
$\widetilde{p}_1$	1	$\sqrt{n}$	n	1	$\sqrt{n}$	1
$\widetilde{p}_2$	$\sqrt{n}$	1	$\sqrt{n}$	1	1	1
$\widetilde{p}_\infty$	n	$\sqrt{n}$	1	1	$\sqrt{n}$	1
$\wp_1$	n	$n^{\frac{3}{2}}$	n	1	n	n
$\wp_2$	$\sqrt{n}$	$\sqrt{n}$	$\sqrt{n}$	1	1	1
$\wp_\infty$	n	n	n	n	n	1

参考文献与补注 7.4

(1)

(2)

(3)

第 8 章　1-$\frac{1}{2}$ 形式与双线性形式

本章介绍实复线性空间上的 1-$\frac{1}{2}$ 形式 (sesqui-linear form) 以及一般线性空间上的特殊双线性形式 (bilinear form). 它们的重要性不仅在于提供了空间的几何结构, 更在于联系了保持此结构的自同构群.

§8.1　1-$\frac{1}{2}$ 形式的正定性

设 $F=\mathbb{R}$ 或 $\mathbb{C}$, V 为域 F 上的线性空间. 回忆 V 上的一个 1-$\frac{1}{2}$ 形式是指一个二元函数 $V\times V\to F$, 满足以下性质: (1) 关于第一分量线性; (2) 关于第二分量共轭线性. 特别地, 正定的 Hermite 形式就是内积. 在有限维内积空间的情形, 若取定线性空间的一组基, 则一个 1-$\frac{1}{2}$ 形式的 Hermite 性 (或正定性) 与它在这组基下矩阵表示的 Hermite 性 (或正定性) 等价. 特别地, 内积在一组基下的矩阵表示称为这组基的 Gram 阵, 它总是 Hermite 正定的.

8.1.1　正定性的判别法

例 8.1.1 (Hilbert 阵的正定性) 记 $H=(H_{ij})_{1\leq i,j\leq n}\in\mathbb{R}^{n\times n}$, 其中 $H_{ij}=\dfrac{1}{i+j-1}$, 则 H 为正定阵.

证明: 任取 $x\in\mathbb{R}^n$, 则 $x^tHx=\sum\limits_{1\leq i,j\leq n}\dfrac{x_ix_j}{i+j-1}=\sum\limits_{1\leq i,j\leq n}x_ix_j\int_0^1t^{i+j-2}\mathrm{d}t=\int_0^1\left(\sum\limits_{i=1}^nx_it^{i-1}\right)^2\mathrm{d}t\geq 0$, 且 "=" 取到当且仅当 $\forall t\in[0,1]$, $\sum\limits_{i=1}^nx_it^{i-1}=0$, 当且仅当 $x=0\in\mathbb{R}^n$. □

注: 上述证明的本质是, Hilbert 阵 $H\in\mathbb{R}^{n\times n}$ 是 $\leq n-1$ 次实系数多项式空间 $\mathbb{R}[X]_{\leq n-1}$ 上实内积 $\langle f(X),g(X)\rangle=\int_0^1f(t)g(t)\mathrm{d}t$ 的 Gram 阵.

例 8.1.2 (Cauchy 阵的半正定性) 记 $C=(C_{ij})_{1\leq i,j\leq n}\in\mathbb{R}^{n\times n}$, 其中 $C_{ij}=\dfrac{1}{a_i+a_j}$, 且 $a_i>0$, 则 C 为半正定阵.

证明: (法一) 记 $D=\mathrm{diag}(a_1,\cdots,a_n)$, 则 $DC+CD=\epsilon\epsilon^t$, 其中 $\epsilon=(1,\cdots,1)^t$. 显然 C 对称. 现任取 C 的特征值 $c\in\mathbb{R}$ 与对应的特征向量 $\alpha\in\mathbb{R}^n\backslash\{0\}$, 即 $C\alpha=c\alpha$, 由 $\alpha^tDC\alpha+\alpha^tCD\alpha=\alpha^t\epsilon\epsilon^t\alpha$ 知, $2c\cdot\alpha^tD\alpha=(\epsilon^t\alpha)^t(\epsilon^t\alpha)$, 即 $c=\dfrac{\langle\epsilon,\alpha\rangle^2}{2\alpha^tD\alpha}\geq 0$. 因此 C 半正定.

(法二) 固定 $0<t<\min\limits_{1\leq i\leq n}a_i$, 则 $\dfrac{1}{a_i+a_j-t}=\dfrac{t}{a_ia_j}\cdot\dfrac{1}{1-\frac{(a_i-t)(a_j-t)}{a_ia_j}}=\dfrac{t}{a_ia_j}\cdot\sum\limits_{k=0}^{+\infty}\left(\dfrac{(a_i-t)(a_j-t)}{a_ia_j}\right)^k$. 记 $\alpha_k=\left(\dfrac{(a_1-t)^k}{a_1^{k+1}},\cdots,\dfrac{(a_n-t)^k}{a_n^{k+1}}\right)^t$, 则 $\left(\dfrac{1}{a_i+a_j-t}\right)_{1\leq i,j\leq n}=t\cdot\sum\limits_{k=0}^{+\infty}\alpha_k\alpha_k^t$ 为可数无穷个半正定阵之和, 它也半正定. 再令 $t\to 0^+$, 即知 $\left(\dfrac{1}{a_i+a_j}\right)_{1\leq i,j\leq n}$ 也半正定. 完全同理可证明 $\forall p>0$, $\left(\dfrac{1}{(a_i+a_j)^p}\right)_{1\leq i,j\leq n}$ 仍半正定. □

注: 利用对称 Cauchy 阵的行列式 $\det(C)=\dfrac{\prod\limits_{1\leq i<j\leq n}(a_j-a_i)^2}{\prod\limits_{1\leq i,j\leq n}(a_i+a_j)}$, 可归纳证明:

(1) C 为半正定阵 $\Longleftrightarrow\{a_i\}_{i=1}^n\subseteq(0,+\infty)$;

(2) C 为正定阵 $\Longleftrightarrow\{a_i\}_{i=1}^n\subseteq(0,+\infty)$ 两两不同;

(3) C 为全正定阵 (即所有子式均为正) $\Longleftrightarrow\{a_i\}_{i=1}^n\subseteq(0,+\infty)$ 严格单调.

为方便起见, 我们引入线性变换或方阵的主不变量的概念.

定义 8.1.1 (主不变量) 设 V 为域 F 上有限维线性空间, $n=\dim_F(V)$, $0\leq k\leq n$, $T\in L(V)$, 考虑拉回映射

$$\begin{aligned} T^{(k)}\colon\ \Lambda^k(V) &\longrightarrow \Lambda^k(V) \\ L &\longmapsto (T^{(k)}(L)\colon (\alpha_1,\cdots,\alpha_k)\mapsto L(T(\alpha_1),\cdots,T(\alpha_k))) \end{aligned},$$

则 $i_k(T):=\mathrm{tr}(T^{(k)})\in F$ 称为 T 的第 k 个**主不变量** (principal invariants).

注:

(1) 回忆线性变换的主不变量可由它的特征多项式系数唯一决定: $f_T(X)=\sum\limits_{k=0}^{n}(-1)^k i_k(T)X^{n-k}\in F[X]$. 于是若记 $f_T(X)$ 在 $\overline{F}^{\mathrm{alg}}$ 中的 n 个根为 $\{c_i(T)\}_{i=1}^n$, 则由 Vieta 定理知, $i_k(T)=\sum\limits_{1\leq i_1<\cdots<i_k\leq n} c_{i_1}(T)\cdots c_{i_k}(T)$.

(2) 对于方阵 $A\in F^{n\times n}$ 以及 $0\leq k\leq n$, 记 A 的第 k 个主不变量为 $i_k(A):=i_k(L_A)\in F$, 则可直接验证方阵的相似变换不改变主不变量. 特别地, 线性变换的主不变量等于它在任一基下矩阵表示的主不变量. 因此 $i_k(A)=\sum\limits_{1\leq i_1<\cdots<i_k\leq n}\det(A_{\{i_1,\cdots,i_k\},\{i_1,\cdots,i_k\}})$.

利用主不变量等记号, 我们总结若干判断 Hermite 阵正定 (或半正定) 性的方法:

命题 8.1.1 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为 Hermite 阵, 则以下条件等价:*

(1) *A 为正定阵;*
(2) *$\sigma(A)\subseteq(0,+\infty)$;*
(3) *$\exists P\in\mathrm{GL}(n,F)$, s.t. $A=\overline{P^t}P$;*
(4) *存在 P 为正定阵, 满足 $A=P^2$;*
(5) *A 的主子式均为正;*
(6) *A 的顺序主子式均为正;*
(7) *A 的主不变量均为正.*

证明: "(1)⇔(2)": 由 Hermite 阵可正交相似于对角阵即知;
"(2)⇒(7)": 显然;
"(7)⇒(2)": 设 A 的主不变量均为正, 特别地 $0<i_n(A)=\prod\limits_{c\in\sigma(A)}c$. 又由 A 的 Hermite 性知 $\sigma(A)\subseteq\mathbb{R}$, 则 $\sigma(A)\subseteq\mathbb{R}\setminus\{0\}$. 假设 $\exists c\in\sigma(A)\cap(-\infty,0)$, 则 $0=(-1)^n f_A(c)=\sum\limits_{k=0}^{n}i_k(A)(-c)^{n-k}>0$, 矛盾! 因此 $\sigma(A)\subseteq(0,+\infty)$.
"(1)⇒(3)": 由 A 的 Cholesky 分解即知;
"(3)⇒(4)": 设 $\exists P\in\mathrm{GL}(n,F)$, s.t. $A=\overline{P^t}P$, 则由定义知, $\overline{P^t}P$ 为正定阵, 故存在 $Q=\sqrt{\overline{P^t}P}$ 为正定阵, 满足 $\overline{P^t}P=Q^2$, 即 $A=Q^2$.
"(4)⇒(1)": 由正定性的定义即知;
"(1)⇒(5)": 由 A 正定知 A 的主子阵均正定; 由 "(1)⇒(3)" 知正定阵的行列式为正, 故 A 的主子式均为正;
"(5)⇒(6)": 显然;
"(6)⇒(1)": 安师讲义已证. □

命题 8.1.2 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为 Hermite 阵, 则以下条件等价:*

(1) *A 为半正定阵;*
(2) *$\sigma(A)\subseteq[0,+\infty)$;*
(3) *$\exists P\in F^{n\times n}$, s.t. $A=\overline{P^t}P$;*
(4) *存在 P 为半正定阵, 满足 $A=P^2$;*
(5) *A 的主子式均非负;*
(6) *A 的主不变量均非负.*

证明: "(1)⇔(2)": 由 Hermite 阵可正交相似于对角阵即知;
"(2)⇒(6)": 显然;
"(6)⇒(2)": 由 A 的 Hermite 性知 $\sigma(A)\subseteq\mathbb{R}$. 设 A 的主不变量均非负, 假设 $\exists c\in\sigma(A)\cap(-\infty,0)$, 则由 $i_0(A)=1$ 知 $0=(-1)^n f_A(c)=\sum\limits_{k=0}^{n}i_k(A)(-c)^{n-k}>0$, 矛盾! 因此 $\sigma(A)\subseteq[0,+\infty)$.

"(1)⇒(4)": 由 A 的极分解即知;
"(4)⇒(3)": 显然;
"(3)⇒(1)": 由半正定性的定义即知;
"(1)⇒(5)": 由 A 半正定知 A 的主子阵均半正定; 由 "(1)⇒(3)" 知半正定阵的行列式非负, 故 A 的主子式均非负;
"(5)⇒(6)": 显然. □

注: 警告 "A 为半正定阵"⇒"A 的顺序主子式均非负", 但 "⇐" 不成立, 反例如 $A=\begin{pmatrix}0&0\\0&-1\end{pmatrix}$. 以下我们给出一个由顺序主子式判定半正定性的充分不必要条件.

引理 8.1.3 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为 Hermite 阵, 若 A 的前 $(n-1)$ 个顺序主子式均为正, 且第 n 个顺序主子式非负, 则 A 为半正定阵.*

证明: 记 $A=\begin{pmatrix}A_1&\alpha\\\overline{\alpha^t}&a_{nn}\end{pmatrix}$, 其中 $A_1\in F^{(n-1)\times(n-1)}$ 为 Hermite 阵. 由 A_1 的顺序主子式均为正知, A_1 为正定阵, 故 $\exists P_1\in \mathrm{GL}(n-1,F)$, *s.t.* $A_1=\overline{P_1^t}P_1$. 取 $\beta=-A_1^{-1}\alpha\in F^{(n-1)\times 1}$, 以及 $P=\begin{pmatrix}P_1^{-1}&\beta\\0&1\end{pmatrix}\in \mathrm{GL}(n-1,F)$, 则可直接验证 $\overline{P^t}AP=\mathrm{diag}(I_{n-1},a'_{nn})$, 故 $a'_{nn}=\det(\overline{P^t}AP)=|\det(P)|^2\det(A)\geq 0$, 因此 $\overline{P^t}AP$ 为半正定阵, 从而 A 为半正定阵. □

设 $F=\mathbb{R}$ 或 $\mathbb{C}$, 在域 F 上的方阵代数 $F^{n\times n}$ 上, 我们引入以下的偏序关系: $A\geq B\iff A-B$ 为半正定阵; 当然, 也可引入以下的严格偏序关系: $A>B\iff A-B$ 为正定阵. 此序关系称为 Loewner 序. 容易验证, 此序可由加法与非负 (或正) 数的数乘保持, 但未必可由方阵的乘法保持: 对于 $A,B\in F^{n\times n}$ 且 $A,B\geq 0$, $AB\geq 0\iff AB=BA$. 以下我们说明此序与取逆的关系:

引理 8.1.4 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$ 为正定阵, 则 $A\geq(>)B\iff B^{-1}\geq(>)A^{-1}$.*

证明: 记 $P=B^{-\frac{1}{2}}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{-\frac{1}{2}}B^{-\frac{1}{2}}$, 则 $P\in\mathrm{GL}(n,F)$ 为 Hermite 阵, 且 $B^{-1}-A^{-1}=P(A-B)P$, 故结论成立. □

有时我们也可利用扰动法将半正定性的判别化为正定性的判别:

引理 8.1.5 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为 Hermite 阵, 则 A 半正定 $\iff\forall t>0$, (tI_n+A) 正定.*

证明: 由半正定与正定性的定义即知. □

例 8.1.3 (对角占优阵的半正定性) *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为 Hermite 阵,*

(1) *若满足严格对角占优条件 $\forall 1\leq i\leq n$, $|A_{ii}|>\sum\limits_{j\neq i}|A_{ij}|$ 且 $A_{ii}>0$, 则 A 为正定阵;*

(2) *若满足对角占优条件 $\forall 1\leq i\leq n$, $|A_{ii}|\geq\sum\limits_{j\neq i}|A_{ij}|$ 且 $A_{ii}\geq 0$, 则 A 为半正定阵.*

证明: (1) 由 A 的主子式仍满足上述条件以及 Levy-Desplanques 定理即知. (2) 由 (1) 与引理即知. □

注: 警告: 满足 (非严格) 对角占优条件与对角分量均正的 Hermite 阵未必是正定的, 例如 $\begin{pmatrix}1&1\\1&1\end{pmatrix}$.

8.1.2 Hermite 阵与正定阵的不等式

命题 8.1.6 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}\setminus\{0\}$, 则*

(1) $r(A)\geq\dfrac{|\mathrm{tr}(A)|^2}{\mathrm{tr}(\overline{A}^tA)}$, *等号取到当且仅当 A 酉相似于 $\mathrm{diag}(cI_{r(A)},0)\,(c\in\mathbb{C}^*)$.*

(2) *若 A 为 Hermite 阵, 则 $r(A)\geq\max\{r_1(A),r_2(A)\}\geq\dfrac{|\mathrm{tr}(A)|^2}{\mathrm{tr}(A^2)}$, 等号取到当且仅当 A 正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于 $\mathrm{diag}(cI_{r(A)},0)\,(c\in\mathbb{R}^*)$. 特别地, 若 $\mathrm{tr}(A)>0$ 且 $\dfrac{|\mathrm{tr}(A)|^2}{\mathrm{tr}(A^2)}>n-1$, 则 A 为正定阵.*

(3) 若 A 为正规阵, 记 $H(A)=\dfrac{\overline{A}^t+A}{2}$, 则 $r(A)\geq r(H(A))\geq\dfrac{|\mathrm{tr}(H(A))|^2}{\mathrm{tr}(H(A)^2)}$, 等号取到当且仅当 A 的复特征值不为非零的纯虚数, 且 $H(A)$ 正交相似 (若 $F=\mathbb{R}$) 或酉相似 (若 $F=\mathbb{C}$) 于 $\mathrm{diag}(cI_{r(A)},0)\,(c\in\mathbb{R}^*)$.

证明: (1) 由于不等式两端关于酉相似不变, 故可不妨设 A 为复上三角阵. 记 $A=(A_{ij})_{1\leq i\leq j\leq n}$, 则由 Cauchy 不等式知, $r(A)\geq|\{1\leq i\leq n\colon A_{ii}\neq 0\}|\geq\dfrac{\left|\sum\limits_{i=1}^n A_{ii}\right|^2}{\sum\limits_{i=1}^n|A_{ii}|^2}\geq\dfrac{\left|\sum\limits_{i=1}^n A_{ii}\right|^2}{\sum\limits_{1\leq i\leq j\leq n}|A_{ij}|^2}=\dfrac{|\mathrm{tr}(A)|^2}{\mathrm{tr}(\overline{A}^tA)}$, 等号取到当且仅当 A 酉相似于 $\mathrm{diag}(cI_{r(A)},0)\,(c\in\mathbb{C}^*)$.

(2) 由 A 为 Hermite 阵且不等式两端关于酉相似不变知, 可不妨设 A 为实对角阵. 记 $A=\mathrm{diag}(c_1,\cdots,c_{r(A)},0,\cdots,0)$, 其中 $c_i>0$, $\forall 1\leq i\leq r_1(A)$; $c_{r_1(A)+j}<0$, $\forall 1\leq j\leq r_2(A)$, 则类似 (1) 知 $r(A)=r_1(A)+r_2(A)\geq\dfrac{|\mathrm{tr}(A)|^2}{\mathrm{tr}(A^2)}$, 等号取到当且仅当 A 酉相似于 $\mathrm{diag}(cI_{r(A)},0)\,(c\in\mathbb{R}^*)$. 进一步地, 由推论 7.3.18知, 若 $A\in\mathbb{R}^{n\times n}$, 则上述酉相似可改为正交相似.

(3) 由 A 为正规阵知, A 与 $H(A)$ 可同时酉相似于对角形; 又不等式两端关于酉相似不变, 可不妨设 A 为复对角阵. 此时第一个不等号显然成立, 等号当且仅当 A 的复特征值不为非零的纯虚数; 第二个不等号及其取等条件由 (2) 即知. □

命题 8.1.7 (Hadamard 不等式) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$.

(1) 若 $A\in F^{n\times n}$ 为半正定阵, 则 $\det(A)\leq\prod\limits_{i=1}^n A_{ii}$, 等号成立当且仅当 $\exists 1\leq i\leq n$, $s.t.\ A_{ii}=0$, 或 A 为正对角阵.

(2) 若 $B\in F^{n\times n}$, 则 $\begin{cases}|\det(B)|\leq\prod\limits_{i=1}^n\left(\sum\limits_{j=1}^n|B_{ij}|^2\right)^{\frac{1}{2}}, & \text{等号成立当且仅当 } B \text{ 存在零行, 或 } B \text{ 的行向量两两正交.}\\ |\det(B)|\leq\prod\limits_{j=1}^n\left(\sum\limits_{i=1}^n|B_{ij}|^2\right)^{\frac{1}{2}}, & \text{等号成立当且仅当 } B \text{ 存在零列, 或 } B \text{ 的列向量两两正交.}\end{cases}$

证明: (1) 先设 A 正定. 考虑 A 的分块形式 $A=\begin{pmatrix}A_1&\alpha\\\overline{\alpha}^t&a_{nn}\end{pmatrix}$, 则由 Schur 公式知, $\det(A)=\det(A_1)\cdot(a_{nn}-\overline{\alpha}^tA_1^{-1}\alpha)$. 由于 A_1^{-1} 也正定, 则 $\overline{\alpha}^tA_1^{-1}\alpha\geq 0$, 故 $\det(A)\leq\det(A_1)\cdot a_{nn}$, 等号取到当且仅当 $\alpha=0$. 于是对 n 归纳即知不等式成立, 且等号取到当且仅当 A 为正对角阵. 再设 A 半正定且非正定, 则 $\det(A)=0$, 且 $\forall 1\leq i\leq n$, $A_{ii}\geq 0$, 故不等式显然成立, 且等号取到当且仅当 $\exists 1\leq i\leq n$, $s.t.\ A_{ii}=0$.

(2) 记 $A=B\overline{B}^t$ 或 $\overline{B}^tB$, 再由 (1) 即知结论. □

注: 注意一般方阵的行正交性与列正交性并不等价, 例如 $B=\begin{pmatrix}1&1\\0&0\end{pmatrix}$ 行正交但非列正交. 但是行向量组的单位正交性等价于列向量组的单位正交性.

推论 8.1.8 (Fischer 不等式) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $H=\begin{pmatrix}A&B\\\overline{B}^t&C\end{pmatrix}\in F^{n\times n}$ 为半正定阵, 其中 $A\in F^{k\times k}$, 则 $\det(H)\leq\det(A)\cdot\det(C)$, 等号取到当且仅当 $\exists 1\leq i\leq n$, $s.t.\ H_{ii}=0$, 或 H 为正对角阵.

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全类似. 由于 $A\in F^{k\times k}$, $C\in F^{(n-k)\times(n-k)}$ 均为半正定阵, 则它们可正交相似于非负对角阵, 故通过准对角块的正交相似, 可不妨设 A,C 均为非负对角阵. 由 Hadamard 不等式知, $\det(H)\leq\prod\limits_{i=1}^n H_{ii}$
$=\prod\limits_{i=1}^k A_{ii}\cdot\prod\limits_{i=1}^{n-k}C_{ii}=\det(A)\cdot\det(C)$, 等号取到当且仅当 $\exists 1\leq i\leq n$, $s.t.\ H_{ii}=0$, 或 H 为正对角阵. □

引理 8.1.9 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$ 为半正定阵, 则 $\det(A+B)\geq\det(A)+\det(B)$. 进一步地, 当 A 正定时, 等号取到当且仅当 $B=0$.

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全类似. 由扰动法与连续性知, 可不妨设 A 正定, 则 $A=(A^{\frac{1}{2}})^2$, 其中 $A^{\frac{1}{2}}$ 也正定. 通过由 $A^{-\frac{1}{2}}$ 正交相似, 可不妨设 $A=I_n$. 记 B 的特征值为 $c_1,\cdots,c_n\geq 0$, 则 I_n+B 的特征值为 $1+c_1,\cdots,1+c_n$, 故 $\det(I_n+B)=\prod\limits_{i=1}^n(1+c_i)\geq 1+\prod\limits_{i=1}^n c_i=\det(I_n)+\det(B)$, 等号取到当且仅当 $c_1=\cdots=c_n=0$, 即 $B=0$. □

推论 8.1.10 (Everitt 不等式) 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $H=\begin{pmatrix}A & B\\ \overline{B}^t & C\end{pmatrix}\in F^{2n\times 2n}$ 为半正定阵, 其中 $A\in F^{n\times n}$, 则 $\det(H)\leq \det(A)\cdot\det(C)-|\det(B)|^2$, 等号取到当且仅当以下条件之一成立:

(1) H 正定且 $B=0$;

(2) H 半正定且非正定, A,C 均正定, 且 $C=\overline{B}^tA^{-1}B$;

(2) H 以及至少 A,C 之一半正定且非正定.

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全类似. 先设 A 正定. 注意到

$$\begin{pmatrix}I_n & 0\\ -\overline{B}^tA^{-1} & I_n\end{pmatrix}\cdot\begin{pmatrix}A & B\\ \overline{B}^t & C\end{pmatrix}\cdot\begin{pmatrix}I_n & -A^{-1}B\\ 0 & I_n\end{pmatrix}=\begin{pmatrix}A & 0\\ 0 & C-\overline{B}^tA^{-1}B\end{pmatrix},$$

则由 $H=\begin{pmatrix}A & B\\ \overline{B}^t & C\end{pmatrix}$ 半正定知, $C-\overline{B}^tA^{-1}B$ 也半正定. 又由 A^{-1} 正定知, $\overline{B}^tA^{-1}B$ 半正定, 故由引理 8.1.9知, $\det(C)\geq\det(C-\overline{B}^tA^{-1}B)+\det(\overline{B}^tA^{-1}B)$. 因此

$\det(H)=\det(A)\cdot\det(C-\overline{B}^tA^{-1}B)\leq\det(A)\cdot(\det(C)-\det(\overline{B}^tA^{-1}B))=\det(H)\leq\det(A)\cdot\det(C)-|\det(B)|^2$.

综上可知等号取到的条件. □

注: 一般地, 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $H=(H_{ij})_{1\leq i,j\leq k}\in F^{kn\times kn}$ 半正定, 其中 $H_{ij}\in F^{n\times n}$, 则 $\det(H)\leq\det((\det(H_{ij}))_{1\leq i,j\leq k})$. 可以参考 R. C. Thompson "A Determinantal Inequality for Positive Definite Matrices"(1960). 更一般地, 在上述条件下, $\forall 1\leq r\leq n$, $\det(H)\leq\left(\frac{\det((\mathrm{tr}(H_{ij}^{(r)}))_{1\leq i,j\leq k})}{\binom{n}{r}^k}\right)^{\frac{n}{r}}$, 其中 $f_{H_{ij}}(X)=\sum_{i=0}^{n}(-1)^i\mathrm{tr}(H_{ij}^{(i)})X^{n-i}$. 可以参考 M. Lin, P. Zhang "Unifying a result of Thompson and a result of Fiedler and Markham on block positive definite matrices"(2017).

例 8.1.4 设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $(V,\langle\cdot,\cdot\rangle)$ 为域 F 上的内积空间, $\{\alpha_i\}_{i=1}^n\cup\{\beta_i\}_{i=1}^n\subseteq V$, 则

$$|\det(\langle\alpha_i,\beta_j\rangle)_{1\leq i,j\leq n}|^2\leq\det(\langle\alpha_i,\alpha_j\rangle)_{1\leq i,j\leq n}\cdot\det(\langle\beta_i,\beta_j\rangle)_{1\leq i,j\leq n}.$$

8.1.3

最后我们补充 Gram 阵的一处几何意义. 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维实内积空间, $\{\epsilon_i\}_{i=1}^n\subseteq V$ 为一组固定的标准正交基. 现任取 V 的 n 元子集 $\{\alpha_i\}_{i=1}^n$, 记 $(\alpha_1,\cdots,\alpha_n)=(\epsilon_1,\cdots,\epsilon_n)\cdot A$, 其中 $A\in\mathbb{R}^{n\times n}$, 则 $L\colon(\alpha_1,\cdots,\alpha_n)\mapsto\det(A)$ 为 n 重交错线性函数, 即 $L\in\Lambda^n(V)\backslash\{0\}$ 为一个体积形式. 特别地, 由 $(\langle\alpha_i,\alpha_j\rangle)_{1\leq i,j\leq n}=A^t\cdot(\langle\epsilon_i,\epsilon_j\rangle)_{1\leq i,j\leq n}\cdot A$ 知, $|L(\alpha_1,\cdots,\alpha_n)|=\sqrt{\det(A^tA)}=\sqrt{\det(\langle\alpha_i,\alpha_j\rangle)_{1\leq i,j\leq n}}$ 是广义平行多面体 $P(\alpha_1,\cdots,\alpha_n)$ 的绝对体积. 值得注意的是, 此绝对体积与标准正交基的选取无关, 只依赖于空间上的内积, 因此称为由内积定义的绝对体积.

一般地, 任取 V 的 m 元子集 $\{\alpha_i\}_{i=1}^m$, 记 $W=\mathrm{Span}_{\mathbb{R}}(\{\alpha_i\}_{i=1}^m)\subseteq V$, 则 $(W,\langle\cdot,\cdot\rangle\big|_{W\times W})$ 也为有限维实内积空间, 故也可由内积定义绝对体积. 特别地, 广义平行多面体 $P(\alpha_1,\cdots,\alpha_m)$ 的 m 维绝对体积为 $\sqrt{\det(\langle\alpha_i,\alpha_j\rangle)_{1\leq i,j\leq m}}$. 记 $(\alpha_1,\cdots,\alpha_m)=(\epsilon_1,\cdots,\epsilon_n)\cdot A$, 其中 $A\in\mathbb{R}^{n\times m}$, 则由 $(\langle\alpha_i,\alpha_j\rangle)_{1\leq i,j\leq m}=A^t\cdot(\langle\epsilon_i,\epsilon_j\rangle)_{1\leq i,j\leq n}\cdot A$ 知, 上述 m 维绝对体积也可写为 $\sqrt{\det(A^tA)}$.

命题 8.1.11 (Pythagorean 定理) 设 $(V,\langle\cdot,\cdot\rangle)$ 为有限维实内积空间, $\{\epsilon_i\}_{i=1}^n\subseteq V$ 为一组固定的标准正交基, $\{W_I=\mathrm{Span}_{\mathbb{R}}(\{\epsilon_i\}_{i\in I})\colon I=\{i_1<\cdots<i_m\}\}$ 为 $\binom{n}{m}$ 个 V 中的 m 维坐标平面, 则任取 V 的 m 元子集 $\{\alpha_i\}_{i=1}^m$, 广义平行多面体 $P(\alpha_1,\cdots,\alpha_m)$ 的 m 维绝对体积的平方等于它向上述 $\binom{n}{m}$ 个 m 维坐标平面上正交投影的 m 维绝对体积的平方和.

证明: 由 Cauchy-Binet 公式即知. □

§8.2 形式的非退化性

本节我们统一处理实复线性空间上 1-$\frac{1}{2}$ 形式与一般线性空间上双线性形式的非退化性, 并

定义 8.2.1 (σ-sesqui-线性形式) 设 V 为域 F 上的线性空间, σ 为域 F 的自同构, $\varphi\colon V\times V\to F$ 为映射, 若 φ 满足关于第一分量是线性的, 关于第二分量是 σ-线性的, 则称 φ 是 V 上的一个 **σ-sesqui-linear 形式**.

注:

(1) 特别地, 若 $F=\mathbb{R}$ 或 $\mathbb{C}$, σ 为复共轭映射, 则 ϕ 为实复线性空间上的 1-$\frac{1}{2}$ 形式; 若 F 为一般域, σ 为恒同映射, 则 φ 为一般线性空间上的双线性形式.

(2) 对于 V 上的 σ-sesqui-线性形式 φ, 考虑 σ^{-1}-线性映射 $\varphi_L\colon \begin{array}{ccc} V & \longrightarrow & V^* \\ \alpha & \longmapsto & \sigma^{-1}(\varphi(\alpha,\cdot)) \end{array}$, 则 $V^{\perp_L}\coloneqq\ker(\varphi_L)\subseteq V$ 称为 φ 的**左根** (left radical). 类似地, 考虑 σ-线性映射 $\varphi_R\colon \begin{array}{ccc} V & \longrightarrow & V^* \\ \beta & \longmapsto & \varphi(\cdot,\beta) \end{array}$, 则 $V^{\perp_R}\coloneqq\ker(\varphi_R)\subseteq V$ 称为 φ 的**右根** (right radical). 注意 $V^{\perp_L}$ 与 $V^{\perp_R}$ 都是 V 的线性子空间. 若 $V^{\perp_L}=\{0\}$, 则称 φ 是**左非退化的** (left non-degenerate); 若 $V^{\perp_R}=\{0\}$, 则称 φ 是**右非退化的** (right non-degenerate).

一般地, 对于无限维线性空间上的 σ-sesqui-线性形式, 它的左右非退化性未必等价. 在双线性函数的情形我们已有这样的反例. 另外, 对于有限维线性空间上的 σ-sesqui-线性形式, 以下引理说明它的左右非退化性等价.

引理 8.2.1 *设 V 为域 F 上的有限维线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式, 则以下条件等价:*

(1) *φ 是左非退化的;*

(2) *$\varphi_L\colon \begin{array}{ccc} V & \longrightarrow & V^* \\ \alpha & \longmapsto & \sigma^{-1}(\varphi(\alpha,\cdot)) \end{array}$ 为 σ^{-1}-线性同构;*

(3) *任取 V 的基 $\{\beta_i\}_{i=1}^n$, 存在 V 的基 $\{\alpha_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(4) *φ 是右非退化的;*

(5) *$\varphi_R\colon \begin{array}{ccc} V & \longrightarrow & V^* \\ \beta & \longmapsto & \varphi(\cdot,\beta) \end{array}$ 为 σ-线性同构;*

(6) *任取 V 的基 $\{\alpha_i\}_{i=1}^n$, 存在 V 的基 $\{\beta_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(7) *存在 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 $\{\beta_i\}_{i=1}^n$, 满足 $(\varphi(\alpha_i,\beta_j))_{n\times n}=I_n$;*

(8) *任取 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 $\{\beta_i\}_{i=1}^n$, 则 $(\varphi(\alpha_i,\beta_j))_{n\times n}\in\mathrm{GL}(n,F)$;*

(9) *任取 V 的基 $\{\alpha_i\}_{i=1}^n$, 则 $(\varphi(\alpha_i,\alpha_j))_{n\times n}\in\mathrm{GL}(n,F)$.*

现在我们考虑 σ-sesqui-线性形式的非退化性与线性子空间的关系. 一般地, 对于子空间 $W\subseteq V$, $i\colon W\hookrightarrow V$ 为含入映射, 记 $W^{\perp_L}\coloneqq\ker(i^t\circ\varphi_L)$, $W^{\perp_R}\coloneqq\ker(i^t\circ\varphi_R)$, 于是 $\perp_L$ 与 $\perp_R$ 定义了线性子空间格上的两个运算, 它们满足以下的性质:

引理 8.2.2 *设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式,*

(1) *若 $W\subseteq V$ 为线性子空间, 则 $(W^{\perp_L})^{\perp_R}\supseteq W$; $(W^{\perp_R})^{\perp_L}\supseteq W$;*

(2) *若 $W_1\subseteq W_2\subseteq V$ 为线性子空间, 则 $W_1^{\perp_L}\supseteq W_2^{\perp_L}$; $W_1^{\perp_R}\supseteq W_2^{\perp_R}$;*

(3) *若 $W_1,W_2\subseteq V$ 为线性子空间, 则 $(W_1+W_2)^{\perp_L}=W_1^{\perp_L}\cap W_2^{\perp_L}$; $(W_1+W_2)^{\perp_R}=W_1^{\perp_R}\cap W_2^{\perp_R}$; $(W_1\cap W_2)^{\perp_L}\supseteq W_1^{\perp_L}+W_2^{\perp_L}$; $(W_1\cap W_2)^{\perp_R}\supseteq W_1^{\perp_R}+W_2^{\perp_R}$.*

注:

(1) 在有限维线性空间的情形, 若 φ 是左 (或右) 退化的, 则上述结论 (1) 与 (3) 中的不等号未必取等. 例如, 考虑

$\varphi\colon \begin{array}{ccc} F^{2\times1}\times F^{2\times1} & \longrightarrow & F \\ ((a_1,a_2)^t,(b_1,b_2)^t) & \longmapsto & a_1b_2 \end{array}$, $W=\mathrm{Span}_F(\{\epsilon_2\})$, 则 $(W^{\perp_L})^{\perp_R}=F^{2\times1}$, $(W^{\perp_R})^{\perp_L}=W$. 再取

$W_1=\mathrm{Span}_F(\{\epsilon_1+\epsilon_2\})$, $W_2=W$, 则 $(W_1\cap W_2)^{\perp_L}=F^{2\times1}$, $W_1^{\perp_L}=W_2^{\perp_L}=W$.

(2) 在无限维线性空间的情形, 即使 φ 是左右非退化的, 上述结论 (1) 与 (3) 中的不等号也未必取等. 具体的反例可参考 Hilbert 空间一节中的讨论.

引理 8.2.3 *设 V 为域 F 上的有限维线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式, 若 $W \subseteq V$ 为线性子空间, 则 $\min\{\dim_F(W^{\perp_L}), \dim_F(W^{\perp_R})\} \geq \dim_F(V) - \dim_F(W)$. 特别地, 若 φ 为左 (或右) 非退化的, 则上述不等式左端两项相等, 且均等于右端.*

证明: 考虑 σ^{-1}-线性映射 $i^t \circ \varphi_L \colon V \longrightarrow W^*$, 由 ($\sigma^{-1}$-线性版本的) 线性代数基本定理知,

$$\alpha \longmapsto \sigma^{-1}(\varphi(\alpha, \cdot))\big|_W$$

$\dim_F(V) = \dim_F(\ker(i^t \circ \varphi_L)) + \dim_F(\mathrm{Im}(i^t \circ \varphi_L)) \leq \dim_F(W^{\perp_L}) + \dim_F(W^*)$. 又 $W \subseteq V$ 为有限维的, 则 $\dim_F(W^{\perp_L}) \geq \dim_F(V) - \dim_F(W)$. 同理知 $\dim_F(W^{\perp_R}) \geq \dim_F(V) - \dim_F(W)$.

现设 φ 为左 (或右) 非退化的, 则 φ_L 为 σ^{-1} 线性同构; 又 i^t 为满射, 则 $i^t \circ \varphi_L$ 也为满射, 故上述不等式取等号. 同理关于 $i^t \circ \varphi_R$ 的不等式也取等号. □

注: 对于一般的线性子空间 $W \subseteq V$, $\dim_F(W^{\perp_L})$ 与 $\dim_F(W^{\perp_R})$ 未必相等. 反例如 φ: $F^{2\times 1} \times F^{2\times 1} \longrightarrow F$, $((a_1, a_2)^t, (b_1, b_2)^t) \longmapsto a_1 b_2$

$W = \mathrm{Span}_F(\{\epsilon_2\})$, 则 $W^{\perp_L} = W$, $W^{\perp_R} = F^{2\times 1}$. 但以下的引理说明总有 $\dim_F(V^{\perp_L}) = \dim_F(V^{\perp_R})$.

引理 8.2.4 *设 V 为域 F 上的有限维线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式, 则 $\dim_F(V^{\perp_L}) = \dim_F(V^{\perp_R})$.*

证明: 记 $\overline{V}^L = V/V^{\perp_L}$, $\overline{V}^R = V/V^{\perp_R}$, 则 $\overline{\varphi}$: $\overline{V}^L \times \overline{V}^R \longrightarrow F$ 为定义良好的 σ-sesqui-线性形式, 且为左

$$(\overline{\alpha}, \overline{\beta}) \longmapsto \varphi(\alpha, \beta)$$

右非退化的. 由 $\overline{\varphi}_L$ 为单射知, $\dim_F(\overline{V}^L) \leq \dim_F((\overline{V}^R)^*)$; 由 $\overline{\varphi}_R$ 为单射知, $\dim_F(\overline{V}^R) \leq \dim_F((\overline{V}^L)^*)$. 因此由 $\overline{V}^L, \overline{V}^R$ 有限维知, $\dim_F(\overline{V}^L) = \dim_F(\overline{V}^R)$, 即 $\dim_F(V^{\perp_L}) = \dim_F(V^{\perp_R})$. □

注: 事实上, 任取 V 的基 $\{\alpha_i\}_{i=1}^n$ 与 $\{\beta_j\}_{j=1}^n$, 则 $\dim_F(V^{\perp_L}) = \dim_F(V^{\perp_R}) = n - r((\varphi(\alpha_i, \beta_j))_{n\times n})$.

推论 8.2.5 *设 V 为域 F 上的有限维线性空间, σ 为域 F 的自同构, φ 为 V 上的非退化 σ-sesqui-线性形式,*

(1) *若 $W \subseteq V$ 为线性子空间, 则 $(W^{\perp_L})^{\perp_R} = W$; $(W^{\perp_R})^{\perp_L} = W$;*

(2) *若 $W_1 \subseteq W_2 \subseteq V$ 为线性子空间, 则 $(W_1 \cap W_2)^{\perp_L} = W_1^{\perp_L} + W_2^{\perp_L}$; $(W_1 \cap W_2)^{\perp_R} = W_1^{\perp_R} + W_2^{\perp_R}$.*

对于 V 上的 σ-sesqui-线性形式 φ, 以及子集 $S_1, S_2 \subseteq V$, 若 $\forall \alpha \in S_1$, $\forall \beta \in S_2$, $\varphi(\alpha, \beta) = 0$, 则称 S_1 关于 φ 垂直于 S_2. 注意一般的垂直关系未必具有对称性. 以下我们讨论一种特殊的 σ-sesqui-线性形式, 它符合几何上垂直关系的对称性原理.

定义 8.2.2 (自反形式) 设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式, 若 $\forall \alpha, \beta \in V$,
$\varphi(\alpha, \beta) = 0 \Longleftrightarrow \varphi(\beta, \alpha) = 0$, 则称 φ 为**自反的** (reflexive).

注: 显然, 取 $\sigma = \mathrm{id}$, 则对称的或交错的双线性形式都是自反的; 取 σ 为域 F 的二阶自同构, 则 σ-Hermite 形式也是自反的. 一个不平凡的事实是, 自反的 σ-sesqui-线性形式均形如上述的常数倍.

定理 8.2.6 (Birkhoff-von Neumann) *设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上左 (或右) 非退化的 σ-sesqui-线性形式, 则 φ 为自反的当且仅当以下条件之一成立:*

(1) *φ 是对称或交错的双线性形式;*

(2) *φ 是某个 σ-Hermite 形式的常数倍.*

证明: 可以参考王杰 "典型群引论". □

为了赋予线性空间 "合理" 的几何结构, 我们假设 σ-sesqui-线性形式定义的垂直关系具有对称性, 此时线性空间及其形式构成的二元组 (V, φ) 称为一个**形式空间** (formed space). 具体地说, 若 φ 为对称的双线性形式, 则 (V, φ) 称为一个**正交空间** (orthogonal space); 若 φ 为交错的双线性形式, 则 (V, φ) 称为一个**正交空间** (symplectic space); 若 φ 为 σ-Hermite 形式, 则 (V, φ) 称为一个**酉空间** (unitary space).

特别地, 在形式空间 (V, φ) 中, 由于垂直关系具有对称性, 则对于任意线性子空间 $W \subseteq V$, 均有 $W^{\perp_L} = W^{\perp_R}$. 于是此时我们可以不区分左右, 记 $W^\perp := W^{\perp_L} = W^{\perp_R}$ 为 W 关于 φ 的**正交补** (orthogonal complement).

注意正交补空间一般不为直和补空间! 记 $\mathrm{rad}(W) \coloneqq W \cap W^{\perp}$ 为 W 关于 φ 的**根** (radical). 显然, $\mathrm{rad}(W) = \{0\}$ 当且仅当 $\varphi\big|_{W\times W}$ 是非退化的, 此时 W 称为关于 φ 的**非退化子空间** (non-degenerate subspace). 由于 $\mathrm{rad}(V) = V \cap V^{\perp} = V^{\perp} = V^{\perp_L} = V^{\perp_R}$, 故这里定义的根与非退化性与本节开头的定义相容.

命题 8.2.7 *设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式, $W \subseteq V$ 为线性子空间. 若三元组 (V, φ, W) 满足以下条件之一:*

(1) *(V, φ) 为形式空间;*

(2) *$W \subseteq V$ 为有限维子空间;*

则 $W \cap W^{\perp_L} = \{0\} \iff W \cap W^{\perp_R} = \{0\} \iff \varphi\big|_{W\times W}$ 是左右非退化的. 进一步地, 当条件 (2) 与 $\varphi\big|_{W\times W}$ 非退化成立时, $V = W \oplus W^{\perp_L} = W \oplus W^{\perp_R}$.

证明: 当条件 (1) 成立时, 上述等价性显然; 当条件 (2) 成立时, 由本节开头的引理即知等价性. 现设条件 (2) 与 $\varphi\big|_{W\times W}$ 非退化成立, 下证 $V = W + W^{\perp_L}$ ($V = W + W^{\perp_R}$ 完全类似). 事实上, 任取 W 的基 $\{\alpha_i\}_{i=1}^{k}$, 由 $\varphi\big|_{W\times W}$ 非退化知, $(\varphi(\alpha_i, \alpha_j))_{1\le i,j\le k} \in \mathrm{GL}(k, F)$. 记 $A = ((\varphi(\alpha_i, \alpha_j))^{-1}_{1\le i,j\le k})^t$, 令 $E\colon V \longrightarrow W$, $\beta \longmapsto \sum\limits_{1\le i,j\le k} A_{ij}\varphi(\beta, \alpha_j)\alpha_i$, 则

可直接验证 $\mathrm{Im}(\mathrm{id}_V - E) \subseteq W^{\perp_L}$, 故 $\forall \beta \in V$, $\beta = E(\beta) + (\mathrm{id}_V - E)(\beta) \in W + W^{\perp_L}$, 即 $V = W + W^{\perp_L}$. □

注: 当条件 (1) 与 $\varphi\big|_{W\times W}$ 非退化成立时, $V = W \oplus W^{\perp}$ 未必成立. 具体的反例可参考 Hilbert 空间一节的讨论.

这里我们再补充一些关于形式的正交分解中子空间非退化性的结论.

引理 8.2.8 *设 (V, φ) 为域 F 上的形式空间, $V = W \oplus U$ 为关于 φ 的正交分解, 则 $\mathrm{rad}(V) = \mathrm{rad}(W) \oplus \mathrm{rad}(U)$ 也为关于 φ 的正交分解. 特别地, φ 是非退化的 $\iff \varphi\big|_{W\times W}$ 与 $\varphi\big|_{U\times U}$ 都是非退化的.*

证明: 由定义直接验证即可. □

注: 警告: 即使 $\dim_F(V) < +\infty$, 若 V 上的 σ-sesqui-线性形式不是自反的, 则由 $V = W \oplus U$ 且 $W \perp U$ 无法推出 $V^{\perp_L} = W^{\perp_L} \oplus U^{\perp_L}$(或 $V^{\perp_R} = W^{\perp_R} \oplus U^{\perp_R}$). 例如 $\varphi\colon F^{2\times 1} \times F^{2\times 1} \longrightarrow F$, $((a_1, a_2)^t, (b_1, b_2)^t) \longmapsto a_1 b_2$, $U = \mathrm{Span}_F(\{\epsilon_1\})$, $W = \mathrm{Span}_F(\{\epsilon_2\})$, 则 $U^{\perp_L} = F^{2\times 1}$, $W^{\perp_L} = W$, 且 $(F^{2\times 1})^{\perp_L} = W$. 但我们总可证明以下较弱的结论.

引理 8.2.9 *设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的非退化 σ-sesqui-线性形式, $W \subseteq V$ 为线性子空间,*

(1) *若 $\varphi\big|_{W^{\perp_L}\times W^{\perp_L}}$ 是右非退化的, 则 $\varphi\big|_{W\times W}$ 是左非退化的;*

(2) *若 $\varphi\big|_{W^{\perp_R}\times W^{\perp_R}}$ 是左非退化的, 则 $\varphi\big|_{W\times W}$ 是右非退化的.*

证明: 由定义直接验证即可. □

注: 上述引理的逆命题不成立, 反例如当 (V, φ) 为形式空间时, 任取 $V^{\perp}$ 的直和补空间 U, 则 $\varphi\big|_{U\times U}$ 为非退化的; 但 $U^{\perp} = V^{\perp}$, 则 $\varphi\big|_{U^{\perp}\times U^{\perp}} \equiv 0$.

引理 8.2.10 *设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的非退化 σ-sesqui-线性形式, $V = W \oplus U$ 且 $W \perp U$,*

(1) *若 φ 是左非退化的, 则 $\varphi\big|_{W\times W}$ 是左非退化的;*

(2) *若 φ 是右非退化的, 则 $\varphi\big|_{U\times U}$ 是右非退化的;*

(3) *特别地, 若 $\dim_F(V) < +\infty$, 则 φ 是非退化的 $\iff \varphi\big|_{W\times W}$ 与 $\varphi\big|_{U\times U}$ 都是非退化的.*

证明: 由定义直接验证即可. □

一般地, 线性空间上的 σ-sesqui-线性形式可能是退化的, 此时我们往往希望能将空间写成 “完全退化部分”(即左右根) 与 “非退化部分”(即限制后非退化) 的直和. 以下说明这一操作是可行的.

引理 8.2.11 *设 (V,φ) 为域 F 上的形式空间, 则任取 $V^\perp$ 的直和补空间 U, $\varphi\big|_{U\times U}$ 为非退化的.*
证明: 由于 $V=V^\perp\oplus U$ 为关于 φ 的正交分解, 则由引理知 $\mathrm{rad}(V)=\mathrm{rad}(V^\perp)\oplus\mathrm{rad}(U)$. 又 $\mathrm{rad}(V)=\mathrm{rad}(V^\perp)=V^\perp$, 则 $\mathrm{rad}(U)=\{0\}$. □

注: 警告: 即使 $\dim_F(V)<+\infty$, 若 V 上的 σ-sesqui-线性形式不是自反的, 则 $V^{\perp_L}$(或 $V^{\perp_R}$) 也可能存在退化的直和补空间. 例如 φ: $\begin{array}{ccc} F^{2\times1}\times F^{2\times1} & \longrightarrow & F \\ ((a_1,a_2)^t,(b_1,b_2)^t) & \longmapsto & a_1b_2 \end{array}$, $U=\mathrm{Span}_F(\{\epsilon_1\})$, $W=\mathrm{Span}_F(\{\epsilon_2\})$, 则 $U\subseteq U^{\perp_L}=F^{2\times1}$, $W=(F^{2\times1})^{\perp_L}$, 且 $W\oplus U=F^{2\times1}$. 但以下的引理说明 $V^{\perp_L}$(或 $V^{\perp_R}$) 总存在非退化的直和补空间.

引理 8.2.12 *设 V 为域 F 上的有限维线性空间, $W_1,W_2\subseteq V$ 为线性子空间, 且 $\dim_F(W_1)=\dim_F(W_2)$, 则存在线性子空间 $U\subseteq V$, 满足 $V=W_1\oplus U=W_2\oplus U$.*
证明: 对 $r\coloneqq\mathrm{codim}_F(W_1)=\mathrm{codim}_F(W_2)$ 归纳证明: 当 $r=0$ 时, 取 $U=\{0\}$ 即可; 现设 $r\geq1$ 且当 $(r-1)$ 时结论成立, 由于 $W_1,W_2\subsetneq V$ 为真子空间, 则 $W_1\cup W_2\neq V$, 故可取 $\alpha\in V\setminus(W_1\cup W_2)$. 记 $W'_{1,2}\coloneqq W_{1,2}\oplus\mathrm{Span}_F(\{\alpha\})$. 注意 $\mathrm{codim}_F(W'_1)=\mathrm{codim}_F(W'_2)=r-1$, 由归纳假设知, 存在线性子空间 $U'\subseteq V$, 满足 $V=W'_1\oplus U'=W'_2\oplus U'$. 令 $U=U'\oplus\mathrm{Span}_F(\{\alpha\})$, 则 $V=W_1\oplus U=W_2\oplus U$. □

推论 8.2.13 *设 V 为域 F 上的有限维线性空间, σ 为域 F 的自同构, φ 为 V 上的非退化 σ-sesqui-线性形式, 则存在 $V^{\perp_L}$ 与 $V^{\perp_R}$ 的公共直和补空间 U, 使得 $\varphi\big|_{U\times U}$ 为非退化的.*
证明: 由于 $\dim_F(V^{\perp_L})=\dim_F(V^{\perp_R})$, 则 $V^{\perp_L}$ 与 $V^{\perp_R}$ 存在公共直和补空间 U. 任取 $\alpha\in U\cap U^{\perp_L}$, 则 $\alpha\perp U$. 又显然 $\alpha\perp V^{\perp_R}$, 则 $\alpha\perp(U+V^{\perp_R})$, 即 $\alpha\perp V$, 故 $\alpha\in U\cap V^{\perp_L}=\{0\}$. 因此 $U\cap U^{\perp_L}=\{0\}$, 即 $\varphi\big|_{U\times U}$ 左非退化. 同理 $\varphi\big|_{U\times U}$ 也右非退化. □

最后我们引入形式空间中向量与线性子空间的迷向性, 并陈述重要的 Witt 定理.

定义 8.2.3 (迷向性) 设 V 为域 F 上的线性空间, σ 为域 F 的自同构, φ 为 V 上的 σ-sesqui-线性形式,
(1) 对于 $\alpha\in V\setminus\{0\}$, 若 $\varphi(\alpha,\alpha)=0$, 则称 α 为**迷向的** (isotropic);
(2) 对于 $W\subseteq V$ 为线性子空间, 若 $\forall\alpha,\beta\in W$, $\varphi(\alpha,\beta)=0$, 则称 W 为**全迷向的** (totally isotropic); 若 $\forall\alpha\in W\setminus\{0\}$, $\varphi(\alpha,\alpha)\neq0$, 则称 W 为**非迷向的** (anisotropic).

注: 显然 $\{0\}$, $V^{\perp_L}$, $V^{\perp_R}$ 均为全迷向子空间; 一般地, 设 $W\subseteq V$ 为线性子空间, 则 W 为全迷向的 $\Longleftrightarrow W\subseteq W^{\perp_L}$ $\Longleftrightarrow W\subseteq W^{\perp_R}$.

定理 8.2.14 (Witt) *设 (V,φ) 为域 F 上的有限维非退化形式空间, 则 V 中所有极大全迷向子空间具有相同的维数, 且该维数 $\leq\dim_F(V)/2$.*
证明: 可以参考王杰 "典型群引论". □

§8.3 特殊的双线性形式

8.3.1 交错双线性形式

8.3.2 对称双线性形式

8.3.3 Hermite 形式

参考文献与补注 8.3
(1)
(2)
(3)

注: $D\in L(F[X];F[X])$ 为满射 $\Longleftrightarrow\mathrm{char}(F)=0$. 此时它存在右逆, 在相差常数的意义下即积分算子 U. 它们与 $F^{(\mathbb{N})}$ 上的左、右移算子存在密切的联系.

第 9 章　杂题

本章主要收录线性代数框架下的若干杂题 (miscellany). 由于它们往往涉及多种观点的复合应用, 而并不容易归于某一主题, 这里的整理仅仅是粗糙的.

§9.1　线性方程组与矩阵理论

9.1.1　LU 分解

为求解 n 元 n 行的线性方程组 $Ax=b$, 一个自然的想法是将系数方阵 A 分解为一个对角分量均为 1 的下三角阵 L 与一个上三角阵 U 的乘积, 此时只需求解线性方程组 $Ux=L^{-1}b$. 一方面, 由例 2.2.9知, L^{-1} 可以写成 L 的多项式, 故右端常数项容易求得; 另一方面, 由 U 的上三角结构, 可简单迭代求出 $x_n,\cdots,x_1$ 的可能值.

定义 9.1.1 (LU 分解) 设 F 是一个域, $A\in F^{n\times n}$, 则 A 的一个 **LU 分解**是指 $A=L\cdot U$, 其中 $L\in F^{n\times n}$ 为下三角阵, $U\in F^{n\times n}$ 为上三角阵.

注: 设 $A=L\cdot U$ 为一个 LU 分解, 则任取对角阵 $D\in\mathrm{GL}(n,F)$, $A=(LD)\cdot(D^{-1}U)$ 也为一个 LU 分解. 因此方阵的 LU 分解通常不唯一. 为确定起见, 有时我们会考虑方阵的 LDU 分解, 即 $A=L\cdot D\cdot U$, 其中 $L\in F^{n\times n}$ 为下三角阵, $D\in F^{n\times n}$ 为对角阵, $U\in F^{n\times n}$ 为上三角阵, 且 L 与 U 的对角分量均为 0 或 1.

命题 9.1.1 *设 F 是一个域, $A\in\mathrm{GL}(n,F)$, 则 A 的 LDU 分解若存在必唯一. 特别地, A 的 LU 分解若存在则在相差对角阵的意义下唯一.*

证明: 设 $A=L_1\cdot D_1\cdot U_1=L_2\cdot D_2\cdot U_2$ 为两个 LDU 分解, 由 $A\in\mathrm{GL}(n,F)$ 知, $L_i,D_i,U_i\in\mathrm{GL}(n,F)\,(i=1,2)$, 则 $L_i,U_i\,(i=1,2)$ 的对角分量均为 1. 注意到 $L_2^{-1}L_1=D_2U_2U_1^{-1}D_1^{-1}$ 的左端为对角分量均为 1 的下三角阵, 右端为上三角阵, 则必为 I_n, 即 $L_1=L_2$, $D_2^{-1}D_1=U_2U_1^{-1}$. 又此式左端为对角阵, 右端为对角分量均为 1 的上三角阵, 则必为 I_n, 即 $D_1=D_2$, $U_1=U_2$. □

注: 对于不可逆的方阵, LDU 分解未必唯一, 例如 $\begin{pmatrix}0&1\\0&0\end{pmatrix}=\underbrace{\begin{pmatrix}1&0\\a&1\end{pmatrix}}_{L}\cdot\underbrace{\begin{pmatrix}1&0\\0&-a\end{pmatrix}}_{D}\cdot\underbrace{\begin{pmatrix}0&1\\0&1\end{pmatrix}}_{U}\,(a\in F)$.

命题 9.1.2 *设 F 是一个域, $A\in F^{n\times n}$, 则:*

(1) *A 具有 LU 分解且其中 $L\in\mathrm{GL}(n,F)\iff A$ 具有行包含性质, 即*

$$\forall i\in\{1,\cdots,n-1\},\ A_{\{i+1\},\{1,\cdots,i\}}\in\mathrm{row}_F(A_{\{1,\cdots,i\},\{1,\cdots,i\}});$$

(2) *A 具有 LU 分解且 $U\in\mathrm{GL}(n,F)\iff A$ 具有列包含性质, 即*

$$\forall j\in\{1,\cdots,n-1\},\ A_{\{1,\cdots,j\},\{j+1\}}\in\mathrm{column}_F(A_{\{1,\cdots,j\},\{1,\cdots,j\}}).$$

证明: (1) 任取 $i\in\{1,\cdots,n-1\}$. 由 $L\in\mathrm{GL}(n,F)$ 知, $L_{\{1,\cdots,i\},\{1,\cdots,i\}}\in\mathrm{GL}(i,F)$, 则

$$\begin{aligned}A_{\{i+1\},\{1,\cdots,i\}}&=L_{\{i+1\},\{1,\cdots,i\}}\cdot U_{\{1,\cdots,i\},\{1,\cdots,i\}}\\&=(L_{\{i+1\},\{1,\cdots,i\}}\cdot L^{-1}_{\{1,\cdots,i\},\{1,\cdots,i\}})\cdot(L_{\{1,\cdots,i\},\{1,\cdots,i\}}\cdot U_{\{1,\cdots,i\},\{1,\cdots,i\}})\\&=(L_{\{i+1\},\{1,\cdots,i\}}\cdot L^{-1}_{\{1,\cdots,i\},\{1,\cdots,i\}})\cdot A_{\{1,\cdots,i\},\{1,\cdots,i\}}\in\mathrm{row}_F(A_{\{1,\cdots,i\},\{1,\cdots,i\}}).\end{aligned}$$

(2) 由 (1) 对于 A^t 成立即知. □

注: 具有 LU 分解的方阵未必具有行列包含性质. 例如, $\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}=\underbrace{\begin{pmatrix}0&0&0\\1&0&0\\0&1&1\end{pmatrix}}_{L}\cdot\underbrace{\begin{pmatrix}0&0&1\\0&1&0\\0&0&0\end{pmatrix}}_{U}$.

推论 9.1.3 *设 F 是一个域, $A\in F^{n\times n}$. 若 $\forall k\in\{1,\cdots,r(A)\}$, $A_{\{1,\cdots,k\},\{1,\cdots,k\}}\in \mathrm{GL}(k,F)$, 则 A 具有 LU 分解. 进一步地, L(或 U) 的对角分量可以均为 1.*

证明: 由条件知, $\dim_F(\mathrm{row}_F(A_{\{1,\cdots,i\},\{1,\cdots,i\}}))=\begin{cases} i, & 1\le i\le r(A)\\ r(A), & r(A)<i\le n\end{cases}$, 则

$$\mathrm{row}_F(A_{\{1,\cdots,i\},\{1,\cdots,i\}})=\begin{cases} F^{1\times i}, & 1\le i\le r(A)\\ \mathrm{row}_F(A_{\{1,\cdots,n\},\{1,\cdots,i\}}), & r(A)<i\le n\end{cases},$$

故 A 具有行包含性质. 同理 A 也具有列包含性质. 由命题 9.1.2知, A 具有 LU 分解, 且其中 L(或 U) 为可逆阵, 故它的对角分量可以均为 1. □

推论 9.1.4 *设 F 是一个域, $A\in\mathrm{GL}(n,F)$, 则 A 具有 LU 分解 $\iff\forall 1\le k\le n$, $A_{\{1,\cdots,k\},\{1,\cdots,k\}}\in\mathrm{GL}(k,F)$.*

证明: "⇐": 由推论 9.1.3即知;

"⇒": 设 $A=L\cdot U$, 其中 $L\in F^{n\times n}$ 为下三角阵, $U\in F^{n\times n}$ 为上三角阵, 则由 $A\in\mathrm{GL}(n,F)$ 知, $L,U\in\mathrm{GL}(n,F)$, 即 L,U 的对角分量均非零, 故 $\forall 1\le k\le n$, $A_{\{1,\cdots,k\},\{1,\cdots,k\}}=L_{\{1,\cdots,k\},\{1,\cdots,k\}}\cdot U_{\{1,\cdots,k\},\{1,\cdots,k\}}\in\mathrm{GL}(k,F)$. □

注: 一般地, $A\in F^{n\times n}$ 具有 LU 分解 $\iff\forall 1\le k\le n$, $r(A_{\{1,\cdots,k\},\{1,\cdots,k\}})+k\ge r(A_{\{1,\cdots,k\},\{1,\cdots,n\}})+r(A_{\{1,\cdots,n\},\{1,\cdots,k\}})$. 可以参考 P. Okunev, C. R. Johnson "Necessary and Sufficient Conditions for Existence of the LU Factorization of an Arbitrary Matrix"(2005).

利用 Gauss 消元法, 我们可以给出一般矩阵的 PLU 分解, 其中置换方阵 P 保证了此分解总存在.

命题 9.1.5 (PLU 分解) *设 F 为一个域, $A\in F^{m\times n}$, 则存在一个置换方阵 $P\in\mathrm{GL}(m,F)$ (即每行每列恰有一个分量为 1, 其余均为 0), 一个对角分量均为 1 的下三角阵 $L\in\mathrm{GL}(m,F)$, 以及一个行阶梯形矩阵 $U\in F^{m\times n}$, 满足 $P\cdot A=L\cdot U$.*

证明: 回忆域 F 上矩阵的 Gauss 消元法, 则存在置换方阵 $P_1,\cdots,P_r\in\mathrm{GL}(m,F)$, 其中左乘 P_i 为交换第 i 行与第 $i'\ge i$ 行, 对角分量均为 1 的下三角阵 $L_1,\cdots,L_r\in\mathrm{GL}(m,F)$, 其中 $(L_i)_{k,j}=0$, $\forall k>j\ne i$, 以及行阶梯形矩阵 $U\in F^{m\times n}$, 满足 $L_rP_r\cdots L_1P_1A=U$. 记 $P=P_r\cdots P_1$ 为置换矩阵, $L_i'=(P_r\cdots P_{i+1})L_i(P_r\cdots P_{i+1})^{-1}$ 为对角分量均为 1 的下三角阵, 则 $L=(L_r'\cdots L_1')^{-1}$ 仍为对角分量均为 1 的下三角阵, 且 $P\cdot A=L\cdot U$. □

注: 方阵的 PLU 分解未必唯一, 例如 $\begin{pmatrix}1&0\\1&1\end{pmatrix}=\underbrace{\begin{pmatrix}1&0\\1&1\end{pmatrix}}_{L}\cdot\underbrace{\begin{pmatrix}1&0\\0&1\end{pmatrix}}_{U}=\underbrace{\begin{pmatrix}0&1\\1&0\end{pmatrix}}_{P}\cdot\underbrace{\begin{pmatrix}1&0\\1&1\end{pmatrix}}_{L}\cdot\underbrace{\begin{pmatrix}1&1\\0&-1\end{pmatrix}}_{U}$.

类似地, 利用变形的 Gauss 消元法, 我们可以给出一般矩阵的 LPU 分解, 同样置换方阵 P 保证了此分解总存在.

命题 9.1.6 (LPU 分解) *设 F 为一个域, $A\in F^{m\times n}$, 则存在一个置换方阵 $P\in\mathrm{GL}(m,F)$, 一个对角分量均为 1 的下三角阵 $L\in\mathrm{GL}(m,F)$, 以及一个行阶梯形矩阵 $U\in F^{m\times n}$, 满足 $A=L\cdot P\cdot U$. 进一步地, 若 $A\in\mathrm{GL}(n,F)$, 则置换方阵 P 唯一.*

证明: 我们描述域 F 上矩阵的变形的 Gauss 消元法: 约定 $i_0=+\infty=\min\varnothing$, $L_0=I_m$. 对于 $1\le j\le n$, 归纳记

$i_j=\min\{1\le i\le m\colon j=\min\{1\le j'\le n\colon (L_{j-1}\cdots L_0A)_{i,j'}\ne 0\}\}$. 若 $i_j=+\infty$, 则取 $L_j=I_m$; 否则取对角分量均为 1 的下三角阵 $L_j\in\mathrm{GL}(m,F)$, 其中 $(L_j)_{k,l}=0$, $\forall k>l\ne i_j$, 满足 $\forall i_j<i\le m$, $(L_j\cdots L_0A)_{i,j}=0$. 最后可取 $\sigma\in S_m$ 满足 $(\sigma(i_1),\cdots,\sigma(i_n))$ 去掉其中 $\sigma(+\infty)$ 项后恰为 $(1,\cdots,m)$, 则置换方阵 $P=(\epsilon_{\sigma(1)},\cdots,\epsilon_{\sigma(m)})$ 满足 $P\cdot(L_n\cdots L_0A)$ 为行阶梯形矩阵.

现设 $A\in\mathrm{GL}(n,F)$, 则 $L,U\in\mathrm{GL}(n,F)$, 即 L,U 的对角分量均非零. 注意到

$$\forall 1\le i,j\le n,\ A_{\{1,\cdots,i\},\{1,\cdots,j\}}=L_{\{1,\cdots,i\},\{1,\cdots,i\}}\cdot P_{\{1,\cdots,i\},\{1,\cdots,j\}}\cdot U_{\{1,\cdots,j\},\{1,\cdots,j\}},$$

其中 $L_{\{1,\cdots,i\},\{1,\cdots,i\}}\in\mathrm{GL}(i,F)$, $U_{\{1,\cdots,j\},\{1,\cdots,j\}}\in\mathrm{GL}(j,F)$, 则 $r(A_{\{1,\cdots,i\},\{1,\cdots,j\}})=r(P_{\{1,\cdots,i\},\{1,\cdots,j\}})$. 又 P 为置换方阵, 则 P 唯一决定. □

定义 9.1.2 (三角等价性) 设 F 为一个域, $A,B\in F^{n\times n}$, 若存在下三角阵 $L\in \mathrm{GL}(n,F)$, 上三角阵 $U\in \mathrm{GL}(n,F)$, 满足 $A=L\cdot B\cdot U$, 则称 A,B 为**三角等价的** (triangularly equivalent).

注: 由命题 9.1.6知, 可逆方阵三角等价的标准形可取为 (唯一的) 置换方阵; 进一步地, $A,B\in\mathrm{GL}(n,F)$ 三角等价当且仅当 $\forall 1\leq i,j\leq n$, $r(A_{\{1,\cdots,i\},\{1,\cdots,j\}})=r(B_{\{1,\cdots,i\},\{1,\cdots,j\}})$.

推论 9.1.7 (LPDU 分解) *设 F 为一个域, $A\in\mathrm{GL}(n,F)$, 则存在唯一的置换方阵 $P\in\mathrm{GL}(n,F)$, 唯一的对角阵 $D\in\mathrm{GL}(n,F)$, 以及一个对角分量均为 1 的下三角阵 $L\in\mathrm{GL}(n,F)$, 一个对角分量均为 1 的上三角阵 $U\in\mathrm{GL}(n,F)$, 满足 $A=L\cdot P\cdot D\cdot U$.*

证明: 由命题 9.1.6知该分解存在, 且其中的置换方阵唯一. 现设 $A=L_1\cdot P\cdot D_1\cdot U_1=L_2\cdot P\cdot D_2\cdot U_2$ 为两种分解, 则 $(P^{-1}(L_2^{-1}L_1)P)\cdot D_1=D_2\cdot(U_2U_1^{-1})$. 注意到 $L_2^{-1}L_1$ 为对角分量均为 1 的下三角阵, 而 P 为置换方阵, 则 $P^{-1}(L_2^{-1}L_1)P$ 的对角分量也均为 1, 故 $(P^{-1}(L_2^{-1}L_1)P)\cdot D_1$ 的对角分量恰为 D_1 的对角分量; 由于 $U_2U_1^{-1}$ 为对角分量均为 1 的上三角阵, 则 $D_2\cdot(U_2U_1^{-1})$ 的对角分量恰为 D_2 的对角分量. 因此比较对角分量知, $D_1=D_2$. □

9.1.2 与交换子的交换性

引理 9.1.8 *设 F 为一个域且 $\mathrm{char}(F)\nmid n!$, $A,B\in F^{n\times n}$, 则 $f_A(X)=f_B(X)\Longleftrightarrow \mathrm{tr}(A^k)=\mathrm{tr}(B^k)$, $\forall 1\leq k\leq n$.*

证明: 设 A 在 $\overline{F}^{\mathrm{alg}}$ 上的特征值为 $c_1,\cdots,c_n$, 则由引理 4.4.3知, $\forall k\geq 1$, A^k 在 $\overline{F}^{\mathrm{alg}}$ 上的特征值为 $c_1^k,\cdots,c_n^k$, 故 $\mathrm{tr}(A^k)=\sum\limits_{i=1}^n c_i^k$. 同理, 记 B 在 $\overline{F}^{\mathrm{alg}}$ 上的特征值为 $d_1,\cdots,d_n$, 则 $\forall k\geq 1$, $\mathrm{tr}(B^k)=\sum\limits_{i=1}^n d_i^k$. 因此, 由 Newton 等式知, $f_A(X)=f_B(X)\Longleftrightarrow\sum\limits_{i=1}^n c_i^k=\sum\limits_{i=1}^n d_i^k$, $\forall 1\leq k\leq n\Longleftrightarrow \mathrm{tr}(A^k)=\mathrm{tr}(B^k)$, $\forall 1\leq k\leq n$. □

推论 9.1.9 *设 F 为一个域且 $\mathrm{char}(F)\nmid n!$, $A\in F^{n\times n}$, 则 A 幂零 $\Longleftrightarrow \mathrm{tr}(A^k)=0$, $\forall 1\leq k\leq n$.*

注: 当 $\mathrm{char}(F)\mid n!$ 时, 上述推论一般不成立, 反例如 $A=I_n$ 非幂零, 但 $\mathrm{tr}(I_n^k)=0$, $\forall k\geq 1$.

命题 9.1.10 *设 F 为一个域, $A,B\in F^{n\times n}$, 记 $C=AB-BA$.*

(1) *若 $r(C)\leq 1$, 则 C 幂零;*

以下设 $\mathrm{char}(F)\nmid n!$:

(2) *(Jacobson) 若 $AC=CA$ 或 $BC=CB$, 则 C 幂零;*

(3) *设 $n\geq 2$. 若 $AC=CA$ 且 $BC=CB$, 则 $C^{n-1}=0$;*

(4) *若 $AC=CA$, 且 $p_A(X)=\prod\limits_{i=1}^k p_i(X)^{r_i}$ 为不可约分解, 则 $C^{2\max\{r_1,\cdots,r_k\}-1}=0$.*

证明: (1) 由 $r(C)\leq 1$ 知, $m(0,f_C(X))\geq\dim_F(\ker(C))\geq n-1$, 即 $f_C(X)=X^{n-1}(X-c)\,(c\in F)$. 由 $\mathrm{tr}(C)=0$ 知, $c=0$, 即 $f_C(X)=X^n$, 故 C 幂零.

(2) 不妨设 $AC=CA$. 记 $\mathrm{ad}(A)\colon\begin{array}{ccc}F^{n\times n}&\longrightarrow&F^{n\times n}\\ M&\longmapsto&AM-MA\end{array}$, 则 $\mathrm{ad}(A)$ 为线性导子, 即满足线性性以及 Leibniz 法则 $\mathrm{ad}(A)(M_1\cdot M_2)=\mathrm{ad}(A)(M_1)\cdot M_2+M_1\cdot\mathrm{ad}(A)(M_2)$, $\forall M_1,M_2\in F^{n\times n}$. 注意 $\mathrm{ad}(A)^2(B)=0$, 则可归纳证明: $\forall k\geq 1$, $(\mathrm{ad}(A))^k(B^k)=k!\cdot C^k$, 故 $\forall 1\leq k\leq n$, $\mathrm{tr}(C^k)=\dfrac{1}{k!}\mathrm{tr}((\mathrm{ad}(A))^k(B^k))=0$. 因此由推论 9.1.9知 C 幂零.

(3) 由 (2) 知 C 幂零. 假设 $C^{n-1}\neq 0$, 则 $f_C(X)=p_C(X)=X^n$, 故由命题 6.2.12知, $A,B\in C(T)=F[T]$. 特别地, $C=AB-BA=0$, 这与 $p_C(X)=X^n\,(n\geq 2)$ 矛盾!

(4) 由条件 $\mathrm{ad}(A)^2(B)=0$, 可归纳证明以下事实:

① $\forall g(X)\in F[X]$, $\mathrm{ad}(g(A))(B)=g'(A)\cdot C=C\cdot g'(A)$; 特别地, $\mathrm{ad}(g(A))^2(B)=0$.

② $\forall g(X)\in F[X]$, $\ker(g(A))\subseteq\ker(g(A)^2\cdot B)$; 特别地, $\forall 1\leq i\leq k$, $\ker(p_i(A)^{r_i})$ 是 B-不变子空间.

记 $W_i=\ker(p_i(A)^{r_i})$, $\forall 1\leq i\leq k$, 则由 A 的准素分解知, $F^{n\times 1}=\bigoplus\limits_{i=1}^k W_i$, 且由②知每个 W_i 均为 C-不变子空间. 于是只需证明: $\forall 1\leq i\leq k$, $C^{2r_i-1}\big|_{W_i}=0$. 先断言: $\forall 1\leq i\leq k$, $p_i'(A)\big|_{W_i}$ 为可逆阵. (这是因为, 假设 $\exists 1\leq i\leq k$, $s.t.$ $p_i'(A)\big|_{W_i}$ 不可逆, 则在 $\overline{F}^{\mathrm{alg}}$ 中得 $0\in\sigma(p_i'(A)\big|_{W_i})=p_i'(\sigma(A\big|_{W_i}))$, 即

$\exists c\in\sigma(A\big|_{W_i}),\ s.t.\ p_i'(c)=0$. 又由 $p_i(A)\big|_{W_i}$ 幂零知, $p_i(c)\in p_i(\sigma(A\big|_{W_i}))=\sigma(p_i(A)\big|_{W_i})=\{0\}$, 因此 $c\in\overline{F}^{\mathrm{alg}}$ 为 $p_i'(X)$ 与 $p_i(X)$ 的公共根, 即 $\gcd(p_i(X),p_i'(X))\neq 1$. 由 $p_i(X)$ 的不可约性知, $p_i'(X)=0$, 这与 $\deg(p_i(X))\leq n$ 且 $\mathrm{char}(F)\nmid n!$ 矛盾!)

由①知 $\forall 1\leq i\leq k$, $\mathrm{ad}(p_i(A))(B)\big|_{W_i}=p_i'(A)\big|_{W_i}\cdot C\big|_{W_i}=C\big|_{W_i}\cdot p_i'(A)\big|_{W_i}$, 故由断言知只需证明: $\forall 1\leq i\leq k$, $\mathrm{ad}(p_i(A))(B)^{2r_i-1}\big|_{W_i}=0$. 事实上, 由 $p_i(A)^{r_i}\big|_{W_i}=0$ 知, $\mathrm{ad}(p_i(A)\big|_{W_i})^{2r_i-1}=0$, 故由 (2) 中归纳证明的结论知, $\mathrm{ad}(p_i(A))(B)^{2r_i-1}\big|_{W_i}=\dfrac{1}{(2r_i-1)!}\mathrm{ad}(p_i(A)\big|_{W_i})^{2r_i-1}(B^{2r_i-1}\big|_{W_i})=0$. □

推论 9.1.11 *设 F 为一个域, 且 $\mathrm{char}(F)\nmid n!$, $A,B\in F^{n\times n}$, 记 $D=AB-BA^t$.*

(1) *若 $AD=DA^t$, 则 $D\notin\mathrm{GL}(n,F)$;*

(2) *若 $AD=DA^t$ 且 $A^tD=DA$, 则 D 幂零;*

(3) *若 $AD=DA^t$ 且 A 半单, 则 $D=0$;*

(4) *若 $AD=DA^t$ 且 A 相似于某个多项式的友矩阵, 则 $r(D)\leq n-|\sigma(A)|$.*

证明: (1) 由 A 与 A^t 的不变因子相同知, $\exists P\in\mathrm{GL}(n,F),\ s.t.\ P^{-1}AP=A^t$, 则 $DP^{-1}=A(BP^{-1})-(BP^{-1})A$. 又 $AD=DA^t$, 则 $A(DP^{-1})=D(A^tP^{-1})=(DP^{-1})A$. 由 Jacobson 引理知, DP^{-1} 幂零; 特别地, $D\notin\mathrm{GL}(n,F)$.

(2) 由于 $D^2=ABD-BA^tD=A(BD)-(BD)A$, 且 $AD^2=(AD)D=D(A^tD)=D^2A$, 故由 Jacobson 引理知, D^2 幂零, 即 D 幂零.

(3) 不妨设 F 为代数闭域, 则 A 可对角化, 即 $\exists P\in\mathrm{GL}(n,F),\ s.t.\ P^{-1}AP=A'$, 其中 $A'=\mathrm{diag}(c_1I_{n_1},\cdots,c_kI_{n_k})$, 且 $c_1,\cdots,c_k$ 两两不同. 记 $B'=P^{-1}B(P^t)^{-1}$, $D'=P^{-1}D(P^t)^{-1}$, 则 $D'=A'B'-B'A'$, 且 $A'D'=D'A'$. 直接计算知, D' 具有与 A' 相同的准对角分块形式; 于是再直接计算知, B' 也具有与 A' 相同的准对角分块形式, 故 $D'=A'B'-B'A'=0$.

(4) 不妨设 F 为代数闭域, 则由条件知, $\exists P\in\mathrm{GL}(n,F),\ s.t.\ P^{-1}AP=A'$, 其中 $A'=\mathrm{diag}(J_{n_1}(c_1),\cdots,J_{n_k}(c_k))$, 且 $c_1,\cdots,c_k$ 两两不同. 记 $B'=P^{-1}B(P^t)^{-1}$, $D'=P^{-1}D(P^t)^{-1}$, 则 $D'=A'B'-B'(A')^t$, 且 $A'D'=D'(A')^t$. 直接计算知, D' 具有与 A' 相同的准对角分块形式; 于是再直接计算知, B' 也具有与 A' 相同的准对角分块形式. 记 $D'=\mathrm{diag}(D_1,\cdots,D_k)$, $B'=\mathrm{diag}(B_1,\cdots,B_k)$, 则 $\forall 1\leq i\leq k$, $D_i=J_{n_i}(c_i)B_i-B_iJ_{n_i}(c_i)^t$, $J_{n_i}(c_i)D_i=D_iJ_{n_i}(c_i)^t$. 对于 $i=1,\cdots,k$, 取 $P_i\in\mathrm{GL}(n_i,F)$ 满足 $P_i^{-1}J_{n_i}(c_i)P_i=J_{n_i}(c_i)^t$, 则 $D_iP_i^{-1}=J_{n_i}(c_i)(B_iP_i^{-1})-(B_iP_i^{-1})J_{n_i}(c_i)$, $J_{n_i}(c_i)(D_iP_i^{-1})=(D_iP_i^{-1})J_{n_i}(c_i)$, 故由 (1) 知 $D_iP_i^{-1}\notin\mathrm{GL}(n_i,F)$, 即 $D_i\notin\mathrm{GL}(n_i,F)$, 因此 $r(D_i)\leq n_i-1$. 综上, $r(D)=\sum\limits_{i=1}^{k}r(D_i)\leq\sum\limits_{i=1}^{k}(n_i-1)=n-k=n-|\sigma(A)|$. □

§9.2 线性空间与线性映射理论

9.2.1 方阵子空间的维数

以下命题说明了一族方阵的最大秩对线性无关性的限制:

命题 9.2.1 *设 F 为一个域, $V\subseteq F^{m\times n}$ 为子空间. 若 $|F|>r:=\max\{r(A)\in\mathbb{N}\colon A\in V\}$, 则 $\dim_F(V)\leq\max\{m,n\}\cdot r$.*

证明: 不妨设 $m\geq n$. 通过考虑子空间 $\widetilde{V}:=\{(A,0)\colon A\in V\}\subseteq F^{m\times m}$, 可不妨设 $m=n$. 取 $A_0\in V$ 使得 $r(A_0)=r$. 通过相抵变换, 可不妨设 $A_0=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}$. 断言: 任取 $A\in V$, 记 $A=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}$ 为如上分块, 则 $A_{21}\cdot A_{12}=0$, 且 $A_{22}=0$. (事实上, 任取 $A\in V$, 则 $\forall t\in F$, $A+tA_0\in V$, 故 $r(A+tA_0)\leq r$. 现任取 A_{21} 的行向量 $\alpha\in F^{1\times r}$, A_{12} 的列向量 $\beta\in F^{r\times 1}$, 该行该列对应的 A_{22} 中元 $c\in F$, 则 $\det\begin{pmatrix}A_{11}+tI_r&\beta\\\alpha&c\end{pmatrix}=0$. 这是一个关于 t 的 r 次多项式方程, 且有 $|F|>r$ 个根, 故该多项式为零多项式. 特别地, t^r 的系数 $c=0$; t^{r-1} 的系数 $-\alpha\cdot\beta=0$.) 因此, 任取 $A,B\in V$, 由 $A+B\in V$ 知, $(A_{21}+B_{21})(A_{12}+B_{12})=0$, 故 $A_{21}\cdot B_{12}+B_{21}\cdot A_{12}=0$.

以下考虑线性映射 $T\colon V\longrightarrow F^{r\times m}$. 一方面, 注意到 $\varphi\colon \ker(T)\longrightarrow(F^{r\times m})^*$

$A\longmapsto(A_{11},A_{12})$ $\qquad$ $A\longmapsto(\varphi(A)\colon(B_{11},B_{12})\mapsto\mathrm{tr}(A_{21}\cdot B_{12}))$

为单射, 则 $\dim_F(\ker(T))=\dim_F(\varphi(\ker(T)))$. 另一方面, 由于 $\forall A\in\ker(T)$, $\forall B\in V$, $A_{21}\cdot B_{12}=0$, 则

$\mathrm{Im}(T) \subseteq (\varphi(\ker(T)))^{\diamondsuit}$, 故 $\dim_F(\mathrm{Im}(T)) \leq mr - \dim_F(\varphi(\ker(T)))$. 因此, 由秩-零度定理知, $\dim_F(V) = \dim_F(\ker(T)) + \dim_F(\mathrm{Im}(T)) \leq mr$. □

注:

(1) 当上述命题结论中的不等式取等号时, 子空间 V 的形式可以确定, 可参考 H. Flanders"On spaces of linear transformations with bounded rank"(1960).

(2) 通过组合技巧, 上述命题中的条件 "$|F| > r$" 可去掉, 可参考 Roy Meshulam "On the maximal rank in a subspace of matrices"(1985).

以下命题说明了一族方阵的两两交换性对线性无关性的限制:

命题 9.2.2 *设 F 为一个域, 则* $\max\{k \geq 1\text{: 存在}A_1, \cdots, A_k \in F^{n\times n}\text{两两交换且线性无关}\} = \left\lfloor \dfrac{n^2}{4} \right\rfloor + 1$.

证明: 记 $\mathscr{F} = \{E_{ij} : 1 \leq i \leq \lfloor \frac{n}{2} \rfloor,\ \lfloor \frac{n}{2} \rfloor + 1 \leq j \leq n\} \cup \{I_n\}$, 则 $\mathscr{F}$ 中元两两交换且线性无关, 且 $|\mathscr{F}| = \left\lfloor \frac{n^2}{4} \right\rfloor + 1$. 以下证明最大值为 $\left\lfloor \frac{n^2}{4} \right\rfloor + 1$. 事实上, 可不妨设 F 为代数闭域, 此时一族两两可交换的方阵可同时上三角化, 且不改变其中的任意线性相关性, 故只需证明:

$$\max\{k \geq 1\text{: 存在}A_1, \cdots, A_k \in F^{n\times n}\text{上三角、两两交换且线性无关}\} \leq \left\lfloor \frac{n^2}{4} \right\rfloor + 1.$$

对 n 归纳证明: 当 $n = 1$ 时结论显然; 假设 $n \geq 2$ 且当 $(n-1)$ 时结论成立, 现证明当 n 时结论也成立. 记 $k = \lfloor \frac{n^2}{4} \rfloor + 2$. 若存在 $A_1, \cdots, A_k \in F^{n\times n}$ 上三角、两两交换且线性无关, 记 $V = \mathrm{Span}_F(\{A_1, \cdots, A_k\})$, 则 V 中方阵也是上三角且两两交换的. 考虑 $A_i' \coloneqq (A_i)_{\{2,\cdots,n\},\{2,\cdots,n\}} \in F^{(n-1)\times(n-1)}$ $(1 \leq i \leq k)$, 则 $A_1', \cdots, A_k'$ 上三角、两两交换; 记 $V' = \mathrm{Span}_F(\{A_1', \cdots, A_k'\})$, 则由归纳假设知 $\dim_F(V') =: k' \leq \lfloor \frac{(n-1)^2}{4} \rfloor + 1$. 不妨设 $\{A_1', \cdots, A_{k'}'\}$ 线性无关, 则 $\forall k'+1 \leq i \leq k$, $\exists c_{i,1}, \cdots, c_{i,k'} \in F$, $s.t.$ $A_i' = \sum\limits_{j=1}^{k'} c_{ij} A_j'$, 此时记 $B_i = A_i - \sum\limits_{j=1}^{k'} c_{ij} A_j = \begin{pmatrix} \beta_i \\ 0 \end{pmatrix}$, 其中 $\beta_i \in F^{1\times n}$. 由 $\{B_{k'+1}, \cdots, B_k\} \subseteq V$ 线性无关知, $\{\beta_{k'+1}, \cdots, \beta_k\} \subseteq F^{1\times n}$ 线性无关. 同理, 考虑 $A_i'' \coloneqq (A_i)_{\{1,\cdots,n-1\},\{1,\cdots,n-1\}}$ $\in F^{(n-1)\times(n-1)}$ $(1 \leq i \leq k)$, 则可取 $k'' \leq \lfloor \frac{(n-1)^2}{4} \rfloor + 1$, 以及 $\{\gamma_{k''+1}, \cdots, \gamma_k\} \subseteq F^{n\times 1}$ 线性无关, 使得 $C_i \coloneqq \begin{pmatrix} 0 & \gamma_i \end{pmatrix}$ 满足 $\{C_{k''+1}, \cdots, C_k\} \subseteq V$. 于是由 V 中方阵的交换性知, $\forall k'+1 \leq i \leq k$, $\forall k''+1 \leq j \leq k$, $B_i C_j = C_j B_i$, 即 $\begin{pmatrix} \beta_{k'+1} \\ \vdots \\ \beta_k \end{pmatrix} \cdot \begin{pmatrix} \gamma_{k''+1} & \cdots & \gamma_k \end{pmatrix} = 0$, 故由线性方程组理论知, $r \begin{pmatrix} \beta_{k'+1} \\ \vdots \\ \beta_k \end{pmatrix} + r \begin{pmatrix} \gamma_{k''+1} & \cdots & \gamma_k \end{pmatrix} \leq n$. 但另一方面, $r \begin{pmatrix} \beta_{k'+1} \\ \vdots \\ \beta_k \end{pmatrix} + r \begin{pmatrix} \gamma_{k''+1} & \cdots & \gamma_k \end{pmatrix} = (k-k') + (k-k'') \geq 2\left(\lfloor \frac{n^2}{4} \rfloor - \lfloor \frac{(n-1)^2}{4} \rfloor + 1\right) = 2\left(\lfloor \frac{n}{2} \rfloor + 1\right) > n$, 矛盾! □

特别地, Cayley-Hamilton 定理表明 $F^{n\times n}$ 中由一个方阵生成的子代数的维数不超过 n. 进一步地, 以下命题表明 $F^{n\times n}$ 中由两个交换的方阵生成的子代数的维数也不超过 n.

命题 9.2.3 (Gerstenhaber) *设 F 为一个域, $A, B \in F^{n\times n}$ 满足 $AB = BA$, 则* $\dim_F(\{A^i B^j : 0 \leq i, j \leq n-1\}) \leq n$.

证明: 可以参考 J. Barria, P. R. Halmos "Vector bases for two commuting matrices"(1990). □

§9.3 初等多项式理论

9.3.1 分圆多项式

9.3.2 Newton 多项式

§9.4 特征多项式与特征值理论

9.4.1 Sylvester 方程的解

命题 9.4.1 (Sylvester) *设 F 为一个域, $A\in F^{m\times m}$, $B\in F^{n\times n}$, 则以下条件等价:*

(1) $\forall C\in F^{m\times n}$, $\exists! X\in F^{m\times n}$, *s.t.* $AX-XB=C$;

(2) $\gcd(f_A(Z),f_B(Z))=1$;

(3) $\forall f(Z),g(Z)\in F[Z]$, $\exists h(Z)\in F[Z]$, *s.t* $f(A)=h(A)$, $g(B)=h(B)$.

证明: 考虑线性映射 $T\colon \begin{array}{ccc} F^{m\times n} & \longrightarrow & F^{m\times n} \\ X & \longmapsto & AX-XB \end{array}$, 则 (1)$\Longleftrightarrow T$ 为线性同构 $\Longleftrightarrow T$ 为线性单射, 以下只需考虑矩阵方程 $AX-XB=0$ 的解空间.

"(1)⇒(2)": 假设 $\gcd(f_A(Z),f_B(Z))\neq 1$, 则 $f_A(Z)$ 与 $f_B(Z)$ 在 $\overline{F}^{\mathrm{alg}}$ 上有公共根, 记为 $c\in\overline{F}^{\mathrm{alg}}$. 注意 $f_B(Z)=f_{B^t}(Z)$. 记 $\alpha\in(\overline{F}^{\mathrm{alg}})^{m\times 1}\setminus\{0\}$ 满足 $A\alpha=c\alpha$, $\beta\in(\overline{F}^{\mathrm{alg}})^{n\times 1}\setminus\{0\}$ 满足 $B^t\beta=c\beta$, 则 $X\coloneqq\alpha\beta^t\in(\overline{F}^{\mathrm{alg}})^{m\times n}\setminus\{0\}$ 满足 $AX-XB=(A\alpha)\beta^t-\alpha(\beta^tB)=(A\alpha)\beta^t-\alpha(B^t\beta)^t=(c\alpha)\beta^t-\alpha(c\beta)^t=0$, 即线性方程组 $AX-XB=0$ 在 $(\overline{F}^{\mathrm{alg}})^{m\times n}$ 上有非零解. 但它的系数都在 F 中, 则它在 $F^{m\times n}$ 上也有非零解, 这与 $AX-XB=0$ 的解空间平凡矛盾!

"(2)⇒(1)": 假设存在 $X\in F^{m\times n}\setminus\{0\}$ 满足 $AX=XB$, 则 $\forall f(Z)\in F[Z]$, $f(A)X=Xf(B)$; 特别地, 取 $f(Z)=f_B(Z)$, 则由 Cayley-Hamilton 定理知, $f_B(A)X=Xf_B(B)=0$, 故 $f_B(A)\notin\mathrm{GL}(m,F)$. 由 $f_B(Z)$ 的不可约因子分解知, 存在不可约多项式 $p(Z)\mid f_B(Z)$, 满足 $p(A)\notin\mathrm{GL}(m,F)$, 即 $\ker(p(A))\neq\{0\}$. 再由推论 6.1.8知, $p(Z)\mid f_A(Z)$, 故 $p(Z)\mid\gcd(f_A(Z),f_B(Z))$, 这与 $\gcd(f_A(Z),f_B(Z))=1$ 矛盾!

"(2)⇒(3)": 设 $\gcd(f_A(Z),f_B(Z))=1$, 则由 Bezout 定理知, $\exists u(Z),v(Z)\in F[Z]$, *s.t.* $u(Z)f_A(Z)+v(Z)f_B(Z)=1$. 任取 $f(Z),g(Z)\in F[Z]$, 记 $h(Z)\coloneqq u(Z)f_A(Z)g(Z)+v(Z)f_B(Z)f(Z)\in F[Z]$, 则由 Cayley-Hamilton 定理, 可知 $h(A)=u(A)f_A(A)g(A)+v(A)f_B(A)f(A)=(I_n-u(A)f_A(A))f(A)=f(A)$, 同理 $h(B)=g(B)$.

"(3)⇒(2)": 由 (3) 知, 取 $f(Z)\equiv 0$, $g(Z)\equiv 1$, 则 $\exists h(Z)\in F[Z]$, *s.t.* $0=h(A)$, $I_n=h(B)$. 一方面, 由 $h(A)=0$ 知 $p_A(Z)\mid h(Z)$, 则任取 $p_A(Z)$ 的不可约因子 $p(Z)$, $p(Z)\mid h(Z)$. 另一方面, 由 $h(B)=I_n$ 以及 $h(Z)$ 的不可约分解知, $p(B)\in\mathrm{GL}(n,F)$, 即 $\ker(p(B))=\{0\}$; 由推论 6.1.8知, $p(Z)$ 不为 $p_B(Z)$ 的不可约因子. 因此 $p_A(Z)$ 与 $p_B(Z)$ 无相同的不可约因子. 再由推论 6.1.8知, $f_A(Z)$ 与 $f_B(Z)$ 也无相同的不可约因子, 即 $\gcd(f_A(Z),f_B(Z))=1$. □

注: 事实上, 矩阵方程 $AX-XB=C\Longleftrightarrow(I_n\otimes A-B^t\otimes I_m)\mathrm{vec}(X)=\mathrm{vec}(C)$, 故它总存在唯一解当且仅当 $(I_n\otimes A-B^t\otimes I_m)\in\mathrm{GL}(mn,F)$. 记 A 在 $\overline{F}^{\mathrm{alg}}$ 中的所有特征值为 $c_i\,(1\leq i\leq m)$, B 在 $\overline{F}^{\mathrm{alg}}$ 中的所有特征值为 $d_j\,(1\leq j\leq n)$, 则 $(I_n\otimes A-B^t\otimes I_m)$ 在 $\overline{F}^{\mathrm{alg}}$ 中的所有特征值为 $c_i-d_j\,(1\leq i\leq m,\ 1\leq j\leq n)$, 故它为可逆阵当且仅当 $\gcd(f_A(Z),f_B(Z))=1$.

推论 9.4.2 *设 F 为一个域, $A\in F^{m\times m}$, $B\in F^{n\times n}$, 则以下条件等价:*

(1) $\forall C\in F^{m\times n}$, $\exists! X\in F^{m\times n}$, *s.t.* $AX-XB=C$;

(2) $\forall D\in F^{n\times m}$, $\exists! Y\in F^{n\times m}$, *s.t.* $BY-YA=D$.

利用齐次 Sylvester 方程只有零解的判别条件, 我们可以得到缠结矩阵方程 $AX=XB$ 的解的大致形式.

命题 9.4.3 *设 F 为一个域, $A,B\in F^{n\times n}$, 且它们具有相同的准对角分块形式, 记为 $A=\mathrm{diag}(A_1,\cdots,A_k)$, $B=\mathrm{diag}(B_1,\cdots,B_k)$. 若 $\gcd(f_{A_i}(Z),f_{B_j}(Z))=1$, $\forall 1\leq i\neq j\leq k$, 则缠结矩阵方程 $AX=XB$ 的解 $X\in F^{n\times n}$ 也具有上述相同的准对角分块形式.*

证明: 设 $X\in F^{n\times n}$ 与 A,B 具有相同的分块形式, 记为 $X=(X_{ij})_{1\leq i,j\leq k}$. 由 $AX=XB$ 知, $A_iX_{ij}=X_{ij}B_j$, $\forall 1\leq i,j\leq k$.

当 $i\neq j$ 时, 由 $\gcd(f_{A_i}(Z),f_{B_j}(Z))=1$ 以及命题 9.4.1知, $X_{ij}=0$. 因此 $X=\mathrm{diag}(X_1,\cdots,X_k)$ 也为准对角形. □

回忆线性变换的循环分解一节中关于中心化子的讨论, 我们可以计算缠结矩阵方程 $AX=XB$ 的解空间维数.

命题 9.4.4 (Cecioni-Frobenius) 设 F 为一个域, $A\in F^{m\times m}$, $B\in F^{n\times n}$. 记 $\{q_i(Z)\}_{i=1}^k$ 为 $\gcd(p_A(Z),p_B(Z))$ 的所有不可约因子, 以及 A 的 $q_i(Z)$-准素分量的不变因子为 $\cdots\mid q_i(Z)^{e_{ij}}\mid\cdots\mid q_i(Z)^{e_{i1}}$, B 的 $q_i(Z)$-准素分量的不变因子为 $\cdots\mid q_i(Z)^{f_{ij}}\mid\cdots\mid q_i(Z)^{f_{i1}}$, 则 $\dim_F(\{X\in F^{m\times n}\colon AX=XB\})=\sum\limits_{i=1}^{k}\deg(q_i(Z))\sum\limits_{j,j'}\min\{e_{i,j},f_{i,j'}\}$.

证明: 完全类似定理 6.2.36的 (2) 讨论即可. □

推论 9.4.5 (Byrnes-Gauger) 设 F 为一个域, $A\in F^{m\times m}$, $B\in F^{n\times n}$, 则

$$\dim_F(C(A))+\dim_F(C(B))\geq 2\dim_F(\{X\in F^{m\times n}\colon AX=XB\}),$$

且等号取到当且仅当 $m=n$ 且 A 相似于 B.

证明: 由命题 9.4.4, 只需证明: "设 $\{e_j\}_{j\geq 1}$, $\{f_j\}_{j\geq 1}\subseteq\mathbb{N}$ 均为有限项非零的递减数列, 则

$$\sum_{j,j'}(\min\{e_j,e_{j'}\}+\min\{f_j,f_{j'}\}-2\min\{e_j,f_{j'}\})\geq 0,$$

且等号取到当且仅当 $\{e_j\}_{j\geq 1}=\{f_j\}_{j\geq 1}$." 这是一个初等的组合结果. □

推论 9.4.6 设 F 为一个域, $A,B\in F^{n\times n}$, 则 A 相似于 B 当且仅当

$$\dim_F(C(A))=\dim_F(C(B))=\dim_F(\{X\in F^{m\times n}\colon AX=XB\}),$$

也当且仅当 $r(I_n\otimes A-A^t\otimes I_n)=r(I_n\otimes B-B^t\otimes I_n)=r(I_n\otimes A-B^t\otimes I_n)$.

最后, 我们补充非齐次 Sylvester 方程解的存在性判别准则. 为简单起见, 我们承认引理 2.3.5的以下多项式版本.

引理 9.4.7 设 F 为一个域, $A\in F[Z]^{m\times n},B\in F[Z]^{p\times q},C\in F[Z]^{p\times n}$, 则 $\begin{pmatrix}A&0\\0&B\end{pmatrix}$ 与 $\begin{pmatrix}A&0\\C&B\end{pmatrix}$ 相抵

$\iff\exists X\in F[Z]^{q\times n},Y\in F[Z]^{p\times m}$, $s.t.\ BX+YA=C$.

证明: 可以参考 W. E. Roth "The Equations $AX-YB=C$ and $AX-XB=C$ in Matrices"(1952). □

命题 9.4.8 (Roth) 设 F 为一个域, $A\in F^{m\times m}$, $B\in F^{n\times n}$, $C\in F^{m\times n}$, 则 $\begin{pmatrix}A&0\\0&B\end{pmatrix}$ 与 $\begin{pmatrix}A&C\\0&B\end{pmatrix}$ 相似

$\iff AX-XB=C$ 存在解 $X\in F^{m\times n}$.

证明: "⇐": 设 $AX-XB=C$ 存在解 $X\in F^{m\times n}$, 则 $\begin{pmatrix}I_m&X\\0&I_n\end{pmatrix}^{-1}\cdot\begin{pmatrix}A&0\\0&B\end{pmatrix}\cdot\begin{pmatrix}I_m&X\\0&I_n\end{pmatrix}=\begin{pmatrix}A&C\\0&B\end{pmatrix}$.

"⇒": 设 $\begin{pmatrix}A&0\\0&B\end{pmatrix}$ 与 $\begin{pmatrix}A&C\\0&B\end{pmatrix}$ 相似, 则 $\begin{pmatrix}ZI_m-A&0\\0&ZI_n-B\end{pmatrix}$ 与 $\begin{pmatrix}ZI_m-A&-C\\0&ZI_n-B\end{pmatrix}$ 相抵. 由引理 9.4.7知, $\exists X\in F[Z]^{m\times n}$, $Y\in F[Z]^{n\times m}$, $s.t.\ (ZI_m-A)X-Y(ZI_n-B)=-C$. 记 $X=X_0+X_1Z+\cdots+X_dZ^d$, $Y=Y_0+Y_1Z+\cdots+Y_dZ^d$, 则比较系数知, $AX_0-Y_0B=C$; $\forall 1\leq i\leq d+1$, $AX_i-Y_iB=X_{i-1}-Y_{i-1}$, 故

$$A\cdot\sum_{i=0}^{d+1}X_iB^i-\sum_{i=0}^{d+1}X_iB^i\cdot B=\sum_{i=0}^{d+1}(AX_i-Y_iB)B^i+(Y_i-X_i)B^{i+1}=C+\sum_{i=1}^{d+1}(X_{i-1}-Y_{i-1})B^i+\sum_{i=0}^{d}(Y_i-X_i)B^{i+1}=0.$$

□

9.4.2 方阵特征值的分布

命题 9.4.9 (三对角方阵的特征值) 设 $A\in\mathbb{C}^{n\times n}$ 为三对角方阵, 且 $\forall 1\leq i\leq n$, $A_{ii}\in\mathbb{R}$.

(1) 若 $\forall 1\leq i\leq n-1$, $A_{i,i+1}A_{i+1,i}>0$, 则 $\sigma(A)\subseteq\mathbb{R}$ 且 $|\sigma(A)|=n$;

(2) 若 $\forall 1\leq i\leq n-1$, $A_{i,i+1}A_{i+1,i}\geq 0$, 则 $\sigma(A)\subseteq\mathbb{R}$;

(3) 若 $\forall 1\leq i\leq n$, $A_{ii}=0$, 且 $\forall 1\leq i\leq n-1$, $A_{i,i+1}A_{i+1,i}<0$, 则 $\sigma(A)\subseteq\sqrt{-1}\mathbb{R}$ 且 $|\sigma(A)|=n$. 进一步地, 此时非零的纯虚特征值必正负成对出现.

证明: (1) 记 $A=\begin{pmatrix} a_1 & b_1 & & \\ c_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_{n-1} & a_n \end{pmatrix}$, 其中 $a_i,b_i,c_i\in\mathbb{C}$. 由 $\forall 1\le i\le n-1,\ b_ic_i>0$ 知, 可取 $\{d_i\}_{i=1}^n\subseteq\mathbb{C}^*$, 满足 $\forall 1\le i\le n-1,\ \left|\dfrac{d_{i+1}}{d_i}\right|^2\cdot b_i=\overline{c_i}$. 记 $D=\mathrm{diag}(d_1,\cdots,d_n)\in\mathrm{GL}(n,\mathbb{C})$, 则 $D^{-1}AD\in\mathbb{C}^{n\times n}$ 为 Hermite 阵, 故 $\sigma(A)\subseteq\mathbb{R}$ 且 A 可对角化. 注意到 $\forall c\in\mathbb{C},\ \det((cI_n-A)_{\{1,\cdots,n-1\},\{2,\cdots,n\}})=(-1)^{n-1}b_1\cdots b_{n-1}\neq 0$, 则 $r(cI_n-A)\ge n-1$; 特别地, $\forall c\in\sigma(A),\ r(cI_n-A)=n-1$, 即 $\dim_{\mathbb{C}}(\ker(cI_n-A))=1$, 故 $|\sigma(A)|=n$.
(2) 对于 $1\le i\le n-1$, 由于 $b_ic_i\ge 0$, 可取 $\theta_i\in\mathbb{R}$ 满足 $b_ie^{-\sqrt{-1}\theta_i}+c_ie^{\sqrt{-1}\theta_i}\ge 0$, 则 $\forall\epsilon>0,\ (b_i+\epsilon e^{\sqrt{-1}\theta_i})(c_i+\epsilon e^{-\sqrt{-1}\theta_i})>0$. 记 $A_\epsilon=A+\epsilon\cdot\sum\limits_{i=1}^{n-1}(e^{\sqrt{-1}\theta_i}E_{i,i+1}+e^{-\sqrt{-1}\theta_i}E_{i+1,i})\ (\epsilon>0)$, 则由 (1) 知 $\sigma(A_\epsilon)\subseteq\mathbb{R}$, 故由连续性知 $\sigma(A)\subseteq\mathbb{R}$.
(3) 由 (1) 知, $\sigma(\sqrt{-1}A)\subseteq\mathbb{R}$ 且 $|\sigma(\sqrt{-1}A)|=n$, 即 $\sigma(A)\subseteq\sqrt{-1}\mathbb{R}$ 且 $|\sigma(A)|=n$. 记 $D=\mathrm{diag}(-1,1,\cdots,(-1)^n)$, 则 $D^{-1}AD=-A$, 故 A 的非零的纯虚特征值必正负成对出现. □

一个经典的研究方阵特征值分布的方法是利用方阵的 Gersgorin 圆盘. 为方便起见, 对于 $A\in\mathbb{C}^{n\times n}$, 记 $R_i'(A)=\sum\limits_{j\neq i}|A_{ij}|\ (1\le i\le n),\ C_j'(A)=\sum\limits_{i\neq j}|A_{ij}|\ (1\le j\le n)$.

命题 9.4.10 (Gersgorin 圆盘) *设 $A\in\mathbb{C}^{n\times n}$, 则*

$$\sigma(A)\subseteq\left(\bigcup_{i=1}^n\{z\in\mathbb{C}\colon|z-A_{ii}|\le R_i'(A)\}\right)\cap\left(\bigcup_{j=1}^n\{z\in\mathbb{C}\colon|z-A_{jj}|\le C_j'(A)\}\right).$$

证明: 以前半部分包含为例, 后半部分完全类似. 设 $c\in\sigma(A)$, 即 $\det(cI_n-A)=0$. 由 Levy-Desplanques 定理知, $\exists 1\le i\le n,\ s.t.\ |c-A_{ii}|\le R_i'(A)$, 即 $c\in\bigcup\limits_{i=1}^n\{z\in\mathbb{C}\colon|z-A_{ii}|\le R_i'(A)\}$. □

推论 9.4.11 (对角占优方阵的特征值) *设 $A\in\mathbb{C}^{n\times n}$, 若 $\forall 1\le i\le n,\ |A_{ii}|>R_i'(A)$, 或 $\forall 1\le j\le n,\ |A_{jj}|>C_j'(A)$, 则:*

(1) *$0\notin\sigma(A)$;*

(2) *进一步设 $\forall 1\le i\le n,\ A_{ii}>0$, 则 $\sigma(A)\subseteq\{z\in\mathbb{C}\colon\Re(z)>0\}$.*

注: 上述推论存在定量版本如下: "设 $A\in\mathbb{C}^{n\times n}$, 则 A 的最小奇异值 $\ge\min\limits_{1\le i\le n}(|A_{ii}|-\dfrac{1}{2}(R_i'(A)+C_i'(A)))$." 可以参考 R. C. Johnson, R. Horn"Topics in Matrix Analysis" (1991).

事实上, 如果对 Gersgorin 圆盘稍做扰动, 我们可以得到稍弱的对角占优方阵的可逆性.

命题 9.4.12 *设 $A\in\mathbb{C}^{n\times n}$, 满足*

(1) *$\forall 1\le i\le n,\ A_{ii}\neq 0$;*

(2) *若 $\forall 1\le i\le n,\ |A_{ii}|\ge R_i'(A)$, 或 $\forall 1\le j\le n,\ |A_{jj}|\ge C_j'(A)$, 且其中至多一个不等式取等号;*

则 $0\notin\sigma(A)$.

证明: 由条件 (2) 知, $\exists 1\le k\le n,\ s.t.\ |A_{kk}|\ge R_k'(A)$, 且 $\forall i\neq k,\ |A_{ii}|>R_i'(A)$. 若 $|A_{kk}|>R_k'(A)$, 则由推论 9.4.11知, $0\notin\sigma(A)$. 以下设 $|A_{kk}|=R_k'(A)>0$. 固定 $\epsilon>0$, 记 $D=I_n+\epsilon E_{kk}$, 则 $R_k'(D^{-1}AD)=\dfrac{1}{1+\epsilon}\sum\limits_{j\neq k}|A_{kj}|$, 且 $\forall i\neq k,\ R_i'(D^{-1}AD)=R_i'(A)+\epsilon|A_{ik}|$, 故当 $0<\epsilon<\min\limits_{i\neq k}\dfrac{|A_{ii}|-R_i'(A)}{|A_{ik}|}$ 时, $D^{-1}AD$ 也为严格对角占优阵. 因此由推论 9.4.11知, $0\notin\sigma(D^{-1}AD)=\sigma(A)$. □

接着我们进一步讨论方阵的 Gersgorin 圆盘与谱集的紧密联系. 为方便起见, 对于 $A\in\mathbb{C}^{n\times n}$, 记 $G(A)=\bigcup\limits_{i=1}^n\{z\in\mathbb{C}\colon|z-A_{ii}|\le R_i'(A)\},\ G(A^t)=\bigcup\limits_{j=1}^n\{z\in\mathbb{C}\colon|z-A_{jj}|\le C_j'(A)\}$.

命题 9.4.13 *设 $A\in\mathbb{C}^{n\times n}$, 则 $\sigma(A)=\bigcap\limits_{P\in\mathrm{GL}(n,\mathbb{C})}G(P^{-1}AP)=\bigcap\limits_{P\in\mathrm{GL}(n,\mathbb{C})}G(P^{-1}A^tP)$.*

证明: 由于 $\sigma(A)=\sigma(A^t)$, 只需证明第一个等号. 一方面, 由 Levy-Desplanques 定理知, $\forall P\in\mathrm{GL}(n,\mathbb{C})$, $\sigma(A)=\sigma(P^{-1}AP)\subseteq G(P^{-1}AP)$, 则 $\sigma(A)\subseteq\bigcap\limits_{P\in\mathrm{GL}(n,\mathbb{C})}G(P^{-1}AP)$. 另一方面, 固定 $\epsilon>0$. 由 Jordan 标准形

知, $\exists P\in \mathrm{GL}(n,\mathbb{C})$, *s.t.* $P^{-1}AP=J$, 其中 J 为每个 Jordan 块的次对角线上均为 ϵ 的 Jordan 标准形, 则可直接验证

$G(P^{-1}AP)=G(J)\subseteq\{z\in\mathbb{C}\colon d(z,\sigma(A))\leq\epsilon\}$, 故 $\bigcap\limits_{P\in\mathrm{GL}(n,\mathbb{C})}G(P^{-1}AP)\subseteq\bigcap\limits_{\epsilon>0}\{z\in\mathbb{C}\colon d(z,\sigma(A))\leq\epsilon\}=\sigma(A)$.

□

注: 不平凡的是, 对于 $A\in\mathbb{C}^{n\times n}$ 且 $\deg(p_A(X))\leq 2$, 则 $\sigma(A)=\bigcap\limits_{U\in\mathrm{U}(n)}G(U^{-1}AU)=\bigcap\limits_{U\in\mathrm{U}(n)}G(U^{-1}A^tU)$. 可以参考 A. Z. Mitura, J. Zemanek "The Gerschgorin discs under unitary similarity"(1997).

命题 9.4.14 *设 $A\in\mathbb{C}^{n\times n}$. 若 $G(A)$ 中的 k 个圆盘之并 $G_k(A)$ 与剩下的 $n-k$ 个圆盘不相交, 则 $G_k(A)$ 恰好包含 k 个 A 的复特征值 (计代数重数); 类似的结论对于 $G(A^t)$ 也成立.*

证明: 通过由置换阵相似, 可不妨设 $G_k(A)=\bigcup\limits_{i=1}^{k}\{z\in\mathbb{C}\colon |z-A_{ii}|\leq R_i'(A)\}$, 则由条件知,

$G(A)\setminus G_k(A)=\bigcup\limits_{i=k+1}^{n}\{z\in\mathbb{C}\colon |z-A_{ii}|\leq R_i'(A)\}$. 记 $A=D+B$, 其中 $D=\mathrm{diag}(A_{11},\cdots,A_{nn})$, $B=A-D$. 固定 $\epsilon\in[0,1]$, 记 $A_\epsilon=D+\epsilon B$, 则 $\forall 1\leq i\leq n$, $R_i'(A_\epsilon)=R_i'(\epsilon B)=\epsilon R_i'(A)$, 故 A_ϵ 的每个 Gersgorin 圆盘都包含于 A 的每个对应的 Gersgorin 圆盘; 特别地, $G_k(A_\epsilon)\subseteq G_k(A)$, $G(A_\epsilon)\setminus G_k(A_\epsilon)\subseteq G(A)\setminus G_k(A)$. 现取 $\mathbb{C}$ 中的一条光滑的简单闭曲线 Γ, 满足 $G_k(A)$ 位于 Γ 围绕的有界区域内部, 且 $G(A)\setminus G_k(A)$ 位于 Γ 围绕的有界区域外部. 特别地, 对于 $\epsilon\in[0,1]$, $G_k(A_\epsilon)$ 也位于 Γ 围绕的有界区域内部, 且 $G(A_\epsilon)\setminus G_k(A_\epsilon)$ 位于 Γ 围绕的有界区域外部. 由 Gersgorin 圆盘定理知, $\sigma(A_\epsilon)\cap\Gamma=\varnothing$, 即 $f_{A_\epsilon}\big|_\Gamma$ 无零点. 由辐角原理知, f_{A_ϵ} 在 Γ 围绕的有界区域内部的零点个数为 $N(\epsilon)=\dfrac{1}{2\pi\sqrt{-1}}\oint_\Gamma\dfrac{f_{A_\epsilon}'(z)}{f_{A_\epsilon}(z)}\mathrm{d}z$. 注意到 $f_{A_\epsilon}'(X)$ 与 $f_{A_\epsilon}(X)$ 的系数都是关于 ϵ 的多项式, 则 $\dfrac{f_{A_\epsilon}'(z)}{f_{A_\epsilon}(z)}$ 是关于 z 与 ϵ 的有理函数, 故 $N(\cdot)$ 为连续函数; 而 $N(\cdot)$ 取非负整数值, 故 $N(\cdot)$ 只能为常数. 特别地, $N(1)=N(0)$. 由于 $f_{A_0}(X)=f_D(X)=\prod\limits_{i=1}^{n}(X-A_{ii})$ 在 Γ 围绕的有界区域内部的零点为 $A_{11},\cdots,A_{kk}$, 则 $f_A(X)=f_{A_1}(X)$ 在 Γ 围绕的有界区域内部也恰有 k 个零点, 即 Γ 围绕的有界区域内部恰好包含 A 的 k 个复特征值 (计代数重数). 再由 Γ 的选取以及 Gersgorin 圆盘定理知, $G_k(A)$ 恰好包含 A 的 k 个复特征值 (计代数重数). □

推论 9.4.15 *设 $A\in\mathbb{C}^{n\times n}$, 若 $\forall 1\leq i\neq j\leq n$, $|A_{ii}-A_{jj}|>R_i'(A)+R_j'(A)$, 则 A 的 n 个 Geršgorin 圆盘两两不交; 特别地, $|\sigma(A)|=n$. 进一步设 $\forall 1\leq i\leq n$, $A_{ii}\in\mathbb{R}$, 且 $f_A(X)\in\mathbb{R}[X]$, 则 A 具有 n 个互不相同的实特征值. 类似的结论对于 $G(A^t)$ 也成立.*

证明: 第一个结论由 Gersgorin 圆盘的定义即知. 由命题 9.4.14知, A 的每个 Gersgorin 圆盘恰好包含 1 个 A 的复特征值 (计代数重数); 特别地, $|\sigma(A)|=n$. 现设 $\forall 1\leq i\leq n$, $A_{ii}\in\mathbb{R}$, 则 A 的每个 Gersgorin 圆盘关于复共轭是不变的, 故它包含 $c\in\mathbb{C}\setminus\mathbb{R}$ 当且仅当它包含 $\overline{c}\in\mathbb{C}\setminus\mathbb{R}$. 又由 $f_A(X)\in\mathbb{R}[X]$ 知 $c\in\sigma(A)\setminus\mathbb{R}\Longleftrightarrow\overline{c}\in\sigma(A)\setminus\mathbb{R}$, 故 A 的每个 Gersgorin 圆盘与 $\sigma(A)\setminus\mathbb{R}$ 的交点必为偶数个, 因此只能为 0 个, 即 $\sigma(A)\subseteq\mathbb{R}$. □

最后, 我们给出预定方阵的部分分量对特征值的限制:

命题 9.4.16 (Mirky) *设 F 为一个域, $n\geq 2$, $\{a_i\}_{i=1}^n\cup\{c_i\}_{i=1}^n\subseteq F$, 则*

$$\exists A\in F^{n\times n},\ s.t.\ A_{ii}=a_i,\ \forall 1\leq i\leq n\ \&\ \sigma(A)=\{c_i\}_{i=1}^n\Longleftrightarrow\sum_{i=1}^{n}a_i=\sum_{i=1}^{n}c_i.$$

证明: "⇒": 由 $\sum\limits_{i=1}^{n}a_i=\mathrm{tr}(A)=\sum\limits_{i=1}^{n}c_i$ 即知;

"⇐": 由于相似不改变方阵的特征值, 只需证明: $J(c_1,\cdots,c_n):=\begin{pmatrix}c_1&&&\\1&c_2&&\\&\ddots&\ddots&\\&&1&c_n\end{pmatrix}$ 相似于某个对角线为 $\{a_i\}_{i=1}^n$ 的方阵. 以下对 $n\geq 2$ 归纳证明: 当 $n=2$ 时, 由于 $a_1+a_2=c_1+c_2$, 则

$$\begin{pmatrix}1&a_1-c_1\\0&1\end{pmatrix}\cdot\begin{pmatrix}c_1&0\\1&c_2\end{pmatrix}\cdot\begin{pmatrix}1&a_1-c_1\\0&1\end{pmatrix}^{-1}=\begin{pmatrix}a_1&*\\*&a_2\end{pmatrix}.$$

现设 $n\geq 3$ 且当 $(n-1)$ 时结论成立, 记 $UT(c)\coloneqq\begin{pmatrix}1 & c\\ 0 & 1\end{pmatrix}$, 则

$$\begin{pmatrix}UT(a_1-c_1) & 0\\ 0 & I_{n-2}\end{pmatrix}\cdot\begin{pmatrix}J(c_1,c_2) & 0\\ E_{12} & J(c_3,\cdots,c_n)\end{pmatrix}\cdot\begin{pmatrix}UT(a_1-c_1) & 0\\ 0 & I_{n-2}\end{pmatrix}^{-1}=\begin{pmatrix}a_1 & *\\ * & J(c_1+c_2-a_1,c_3,\cdots,c_n)\end{pmatrix}.$$

由于 $(c_1+c_2-a_1)+c_3+\cdots+c_n=\sum\limits_{i=2}^{n}a_i$, 则由归纳假设知 $\exists P\in\mathrm{GL}(n-1,F)$, *s.t.* $P\cdot J(c_1+c_2-a_1,c_3,\cdots,c_n)\cdot P^{-1}$ 的对角线为 $\{a_i\}_{i=2}^n$, 故 $\begin{pmatrix}1 & \\ & P\end{pmatrix}\cdot\begin{pmatrix}a_1 & *\\ * & J(c_1+c_2-a_1,c_3,\cdots,c_n)\end{pmatrix}\cdot\begin{pmatrix}1 & \\ & P\end{pmatrix}^{-1}$ 的对角线为 $\{a_i\}_{i=1}^n$. □

注: 上述命题存在它的补形式: "设 F 为代数闭域, $\{a_{ij}\}_{1\leq i\neq j\leq n}\cup\{c_i\}_{i=1}^n\subseteq F$, 则 $\exists A\in F^{n\times n}$, *s.t.* $A_{ij}=a_{ij}$, $\forall 1\leq i\neq j\leq n$, & $\sigma(A)=\{c_i\}_{i=1}^n$." 可以参考 S. Friedland "Matrices with prescribed off-diagonal elements" (1972).

命题 9.4.17 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $\{a_i\}_{i=1}^n\cup\{c_i\}_{i=1}^n\subseteq\mathbb{R}$ 均为递减数列, 则以下条件等价:*

(1) *存在 $A\in F^{n\times n}$ 为 Hermite 阵, 满足 $A_{ii}=a_i$, $\forall 1\leq i\leq n$, 且 $\sigma(A)=\{c_i\}_{i=1}^n$;*

(2) $\forall 1\leq k\leq n-1$, $\sum\limits_{i=1}^{k}a_i\leq\sum\limits_{i=1}^{k}c_i$, 且 $\sum\limits_{i=1}^{n}a_i=\sum\limits_{i=1}^{n}c_i$.

证明: "(1)⇒(2)": 固定 $1\leq k\leq n$, 记 $A_1=A_{\{1,\cdots,k\},\{1,\cdots,k\}}$. 由于 A 与 A_1 均为 Hermite 阵, 它们的特征值均为实数, 分别记为 $c_1(A)\geq\cdots\geq c_n(A)$, 与 $c_1(A_1)\geq\cdots\geq c_k(A_1)$. 由 Courant-Fischer 定理知, $\forall 1\leq i\leq k$, $c_i(A)\geq c_i(A_1)$, 则 $\sum\limits_{i=1}^{k}c_i(A)\geq\sum\limits_{i=1}^{k}c_i(A_1)=\mathrm{tr}(A_1)=\sum\limits_{i=1}^{k}a_i$. 显然当 $k=n$ 时上式取等号.
"(2)⇒(1)": 对 n 归纳证明: 当 $n=1$ 时结论显然; 当 $n=2$ 时, 由条件知 $c_1\geq a_1\geq a_2\geq c_2$. 若 $c_1=c_2$, 则 $c_1=a_1=a_2=c_2$, 取 $A=\mathrm{diag}(a_1,a_2)$ 即可; 若 $c_1>c_2$, 记 $D=\mathrm{diag}(c_1,c_2)$, $O=\dfrac{1}{\sqrt{c_1-c_2}}\begin{pmatrix}\sqrt{a_1-c_2} & -\sqrt{c_1-a_1}\\ \sqrt{c_1-a_1} & \sqrt{a_1-c_2}\end{pmatrix}$, 则可直接验证 $O\in\mathrm{O}(n)$, 故 $A\coloneqq O^{-1}DO$ 为 Hermite 阵, 满足 $A_{ii}=a_i$, $\forall 1\leq i\leq 2$ 且 $\sigma(A)=\{c_i\}_{i=1}^2$.

现设 $n\geq 3$ 且当 $(n-1)$ 时结论成立. 由条件知 $a_1\leq c_1$; 假设 $\forall 1\leq i\leq n-1$, $a_{i+1}<c_{i+1}$, 则 $\sum\limits_{i=1}^{n}a_i<\sum\limits_{i=1}^{n}c_i$, 与条件矛盾! 因此可取 $k=\min\{1\leq i\leq n-1\colon a_{i+1}\geq c_{i+1}\}$. 当 $k=1$ 时, 显然 $c_1\geq a_1\geq a_2\geq c_2$; 当 $2\leq k\leq n-1$ 时, 由选取知 $c_k>a_k\geq a_{k+1}\geq c_{k+1}$. 因此 $c_k\geq a_k\geq a_{k+1}\geq c_{k+1}$ 总成立. 记 $a'_{k+1}=c_k+c_{k+1}-a_k$, 则由 $n=2$ 的情形知, 可取 $O_0\in\mathrm{O}(2)$, 满足 $O_0^{-1}\cdot\mathrm{diag}(c_k,c_{k+1})\cdot O_0$ 的对角线为 $\{a_k,a'_{k+1}\}$. 取 $\widetilde{O_0}=\mathrm{diag}(I_{k-1},O_0,I_{n-k-1})\in\mathrm{O}(n)$, 则 $\widetilde{O_0}^{-1}\cdot\mathrm{diag}(c_1,\cdots,c_n)\cdot\widetilde{O_0}=$

$$\begin{pmatrix}c_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & c_{k-1} & & & & & & \\ & & & a_k & * & & & & \\ & & & * & a'_{k+1} & & & & \\ & & & & & c_{k+2} & & & \\ & & & & & & \ddots & & \\ & & & & & & & & c_n\end{pmatrix}.$$

考虑递减数列 $\{c_1,\cdots,c_{k-1},a'_{k+1},c_{k+2},\cdots,c_n\}$ 与 $\{a_1,\cdots,a_{k-1},a_{k+1},\cdots,a_n\}$, 它们也满足 (2) 中条件, 由归纳假设知, 可取 $O_1\in\mathrm{O}(n-1)$, 满足 $O_1^{-1}\cdot\mathrm{diag}(c_1,\cdots,c_{k-1},a'_{k+1},c_{k+2},\cdots,c_n)\cdot O_1$ 的对角线为 $\{a_1,\cdots,a_{k-1},a_{k+1},\cdots,a_n\}$. 再取 $\widetilde{O_1}\in\mathrm{O}(n)$, 满足 $(\widetilde{O_1})_{\{1,\cdots,n\}\setminus\{k\},\{1,\cdots,n\}\setminus\{k\}}=O_1$, 且 $(\widetilde{O_1})_{ik}=(\widetilde{O_1})_{ki}=\delta_{ik}$, 则可直接验证
$A\coloneqq\widetilde{O_1}^{-1}\cdot\widetilde{O_0}^{-1}\cdot\mathrm{diag}(c_1,\cdots,c_n)\cdot\widetilde{O_0}\cdot\widetilde{O_1}$ 的对角线为 $\{a_1,\cdots,a_{k-1},a_{k+1},\cdots,a_n\}$. □

§9.5 相似标准形理论

9.5.1 通过对称阵相似

命题 9.5.1 *设 F 为一个域, $A\in F^{n\times n}$, 则存在对称阵 $S\in \mathrm{GL}(n,F)$, 满足 $S^{-1}AS=A^t$.*

证明: 先设 A 为 $f(X)=X^n+a_{n-1}X^{n-1}+\cdots+a_1X+a_0\in F[X]$ 的友矩阵, 取 $S=\begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & 1\\ a_2 & \iddots & \iddots & 1 & \\ \vdots & \iddots & \iddots & & \\ a_{n-1} & 1 & & & \\ 1 & & & & \end{pmatrix}$, 则可直接验证 $S^{-1}AS=A^t$. 于是结论对于由多项式的友矩阵组成的准对角阵也成立. 现设 $A\in F^{n\times n}$, 由循环分解知, $\exists P\in \mathrm{GL}(n,F),\ s.t.\ P^{-1}AP=A_1$, 其中 A_1 为由多项式的友矩阵组成的准对角阵. 取对称阵 $S_1\in \mathrm{GL}(n,F)$, 满足 $S_1^{-1}A_1S_1=A_1^t$, 则 $S\coloneqq PS_1P^t\in \mathrm{GL}(n,F)$ 也为对称阵, 且 $S^{-1}AS=A^t$. □

推论 9.5.2 *设 F 为一个域, $A\in F^{n\times n}$, 则存在对称阵 $S_1\in \mathrm{GL}(n,F)$, $S_2\in F^{n\times n}$, 满足 $A=S_1\cdot S_2$.*

证明: 由命题 9.5.1知, 存在对称阵 $S\in \mathrm{GL}(n,F)$, 满足 $S^{-1}AS=A^t$, 则 $A=S\cdot(A^tS^{-1})$, 其中 $(A^tS^{-1})^t=S^{-1}A=A^tS^{-1}$. □

推论 9.5.3 *设 F 为一个域, $A\in F^{n\times n}$ 相似于某个域 F 上多项式的友矩阵, $S\in F^{n\times n}$ 满足 $AS=SA^t$, 则 S 为对称阵.*

证明: 由命题 9.5.1知, 存在对称阵 $S_1\in \mathrm{GL}(n,F)$, 满足 $S_1^{-1}AS_1=A^t$, 则由 $AS=SA^t$ 知, $A(SS_1^{-1})=(SS_1^{-1})A$. 再由 A 相似于某个域 F 上多项式的友矩阵以及命题 6.2.12知, $\exists g(X)\in F[X],\ s.t.\ SS_1^{-1}=g(A)$, 则 $S=g(A)S_1$, 故 $S^t=S_1g(A^t)=S_1g(S_1^{-1}AS_1)=g(A)S_1=S$. □

9.5.2 相似于对称阵

9.5.3 Jordan 标准形的零分量

在方阵的 Jordan 标准形中许多分量都是 0; 但在相似等价类中 Jordan 标准形未必含有最多的 0 分量. 反例如设 $f(X)=(X^2-1)^2\in F[X]$, $A=C_{f(X)}\in F^{4\times4}$, 则 A 有 11 个 0 分量, 但 A 的 Jordan 标准形 $\mathrm{diag}(J_2(1),J_2(-1))$ 只有 10 个 0 分量. 然而除对角分量外, A 有 7 个 0 分量, A 的 Jordan 标准形 $\mathrm{diag}(J_2(1),J_2(-1))$ 有 10 个 0 分量. 以下我们说明: 在相似等价类中, Jordan 标准形总在对角线外含有最多的 0 分量.

命题 9.5.4 *设 F 为一个域, $A\in F^{n\times n}$ 的初等因子个数为 s, 则 A 在对角线外的非 0 分量个数 $\geq n-s$.*

证明: 先断言: 若 $A\in F^{n\times n}$ 除对角分量外的非 0 分量个数 $<n-1$, 则 $\exists\sigma\in S_n,\ s.t.\ R_\sigma^{-1}AR_\sigma=A_1\oplus A_2$, 其中 $R_\sigma\coloneqq(\epsilon_{\sigma(1)},\cdots,\epsilon_{\sigma(n)})$, $A_1\in F^{n_1\times n_1}$, $A_2\in F^{n_2\times n_2}$, $n_1,n_2\geq1$. (这是因为, 考虑集合族

$$\mathscr{F}=\{I\subseteq\{1,\cdots,n\}\colon 1\in I;\ \forall i\in I,\forall j\in\{1,\cdots,n\}\setminus I,\ A_{ij}=A_{ji}=0\}$$

及其上的偏序关系, 则 $\mathscr{F}$ 中存在极小元, 记为 I_1. 再由 A 除对角分量外的非 0 分量个数 $<n-1$ 可知 $I_1\neq\{1,\cdots,n\}$. 记 $I_1=\{1=i_1,i_2,\cdots,i_{n_1}\}$, 以及 $\sigma\in S_n$ 满足 $\sigma(1)=1,\sigma(2)=i_2,\cdots,\sigma(n_1)=i_{n_1}$, 则可直接验证 σ 符合要求.)

注意到通过置换阵相似不改变方阵中对角线外的非 0 分量个数, 故可不妨设 $A=\mathrm{diag}(A_1,\cdots,A_r)$, 其中 $A_i\in F^{n_i\times n_i}$ 不可再通过置换阵相似分解为两个子矩阵的准对角形, $n_i\geq1$. 于是由断言知, A_i 除对角分量外的非 0 分量个数 $\geq n_i-1$, 故 A 除对角分量外的非 0 分量个数 $\geq\sum\limits_{i=1}^{r}(n_i-1)=n-r$. 又 $r\leq s$, 则 A 在对角线外的非 0 分量个数 $\geq n-s$. □

9.5.4 Weyr 示性数与 Segre 示性数

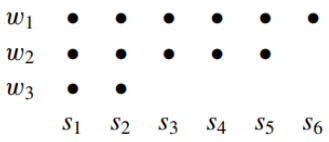
定义 9.5.1 (Weyr 示性数) 设 F 为代数闭域, $A\in F^{n\times n}$, $c\in F$, 记 $r_k(A,c)=r((cI_n-A)^k)\,(k\geq0)$, 则 $w_k(A,c)\coloneqq r_{k-1}(A,c)-r_k(A,c)\,(k\geq1)$ 称为 A 关于 c 的第 k 个 **Weyr 示性数**.

注: 由 Jordan 标准形可知, $r_k(A,c)=n-\dim_F(\ker((cI_n-A)^k))=n-\sum\limits_{i=1}^{k}$ (A 的 Jordan 标准形中 $J_{\geq i}(c)$ 的个数), 则 $w_k(A,c)=r_{k-1}(A,c)-r_k(A,c)=$($A$ 的 Jordan 标准形中 $J_{\geq k}(c)$ 的个数). 特别地, $r_k(A,c)$ 与 $w_k(A,c)$ 都是 A 的相似不变量. 进一步地, $w_k(A,c)-w_{k+1}(A,c)=$(A 的 Jordan 标准形中 $J_k(c)$ 的个数).

推论 9.5.5 *设 $A,B\in F^{n\times n}$, 则 A 相似于 $B\iff\forall c\in\overline{F}^{\mathrm{alg}}$, $\forall k\geq 1$, $w_k(A,c)-w_{k+1}(A,c)=w_k(B,c)-w_{k+1}(B,c)$.*

定义 9.5.2 (Segre 示性数) 设 F 为代数闭域, $A\in F^{n\times n}$, $c\in F$, 在 A 的 Jordan 标准形中关于 c 的 Jordan 块按阶数从大到小排列, 其中第 $k\geq 1$ 大的 Jordan 块阶数称为 A 关于 c 的第 k 个 **Segre 示性数**, 记为 $s_k(A,c)$.

一个重要的观察是, 方阵的 Weyr 示性数与 Segre 示性数可以通过点图 (dot diagram) 联系起来. 例如考虑 Jordan 标准形 $\mathrm{diag}(J_3(0),J_3(0),J_2(0),J_2(0),J_2(0),J_1(0))$, 它关于 0 的 Weyr 示性数列为 $(6,5,2,0,0,\cdots)$, 关于 0 的 Segre 示性数列为 $(3,3,2,2,2,1,0,0,\cdots)$, 总结为一张表格如下:



值得注意的是, 关于 0 的 Weyr 示性数列与 Segre 示性数列构成了 $n=13$ 的两个共轭分划, 因此通过点图可由其中一个决定另一个.

引理 9.5.6 *设 $p_1\geq p_2\geq\cdots$ 与 $p'_1\geq p'_2\geq\cdots$ 分别是 n 与 n' 的分划, 相应的共轭分划为 $q_1\geq q_2\geq\cdots$ 与 $q'_1\geq q'_2\geq\cdots$, 固定 $d\in\mathbb{N}$, 则* $\begin{cases}p'_k\geq p_{k+d}\\ p_k\geq p'_{k+d}\end{cases},\forall k\geq 1\iff|q_k-q'_k|\leq d,\ \forall k\geq 1.$

证明: "⇒": 任取 $k\geq 1$. 若 $q'_k\leq d$, 则显然 $q'_k\leq q_k+d$. 若 $q'_k>d$, 特别地 $q'_k\geq 1$, 则由共轭性知 $p'_{q'_k}\geq k>p_{q_k+1}$. 再由条件知, $p_{q'_k-d}\geq p'_{q'_k}$, 则 $p_{q'_k-d}>p_{q_k+1}$, 故由单调性知, $q'_k-d<q_k+1$, 即 $q'_k\leq q_k+d$. 因此总有 $q'_k\leq q_k+d$. 同理也总有 $q_k\leq q'_k+d$. 综上, $|q_k-q'_k|\leq d$.
"⇐": 任取 $k\geq 1$. 若 $p_{k+d}=0$, 则显然 $p'_k\geq p_{k+d}$. 若 $p_{k+d}\geq 1$, 则由条件知, $q'_{p_{k+d}}\geq q_{p_{k+d}}-d$; 再由共轭性知, $q_{p_{k+d}}-d\geq(k+d)-d=k>q'_{p'_k+1}$, 则 $q'_{p_{k+d}}>q'_{p'_k+1}$, 故由单调性知, $p_{k+d}<p'_k+1$, 即 $p_{k+d}\leq p'_k$. 因此总有 $p_{k+d}\leq p'_k$. 同理也总有 $p'_{k+d}\leq p_k$. 综上, $\begin{cases}p'_k\geq p_{k+d}\\ p_k\geq p'_{k+d}\end{cases}$. □

命题 9.5.7 *设 F 为代数闭域, $A\in F^{m\times n}$, $B\in F^{n\times m}$, $k\geq 1$, 则:*
(1) $\forall c\in F^*$, $w_k(AB,c)=w_k(BA,c)$;
(2) $w_k(AB,0)\geq w_{k+1}(BA,0)$; $w_k(BA,0)\geq w_{k+1}(AB,0)$.

证明: (1) 任取 $g(X)\in F[X]$ 满足 $g(0)\neq 0$. 注意到 $B\cdot g(AB)=g(BA)\cdot B$, 则 $\ker(g(AB))\longrightarrow\ker(g(BA))$ 是定义良好的线性映射 $\alpha\longmapsto B\alpha$. 断言: 它是单射. (这是因为: 假设 $\alpha\in\ker(g(AB))$ 满足 $B\alpha=0$, 则 $\alpha\in\ker(g(0)I_m)=\{0\}$.)
因此 $\dim_F(\ker(g(AB)))\leq\dim_F(\ker(g(BA)))$. 同理 $\geq$ 方向也成立. 综上, $\dim_F(\ker(g(AB)))=\dim_F(\ker(g(BA)))$. 特别地, 取 $g(X)=(c-X)^k$, 其中 $k\geq 1$, $c\in F^*$, 则 $\dim_F(\ker((cI_m-AB)^k))=\dim_F(\ker((cI_n-BA)^k))$, 即 $m-r_k(AB,c)=n-r_k(BA,c)$, 故 $w_k(AB,c)=r_{k-1}(AB,c)-r_k(AB,c)=r_{k-1}(BA,c)-r_k(BA,c)=w_k(BA,c)$.
(2) 一方面, 注意到 $\ker((AB)^{k+1})/\ker(B(AB)^k)\longrightarrow\ker((AB)^k)/\ker(B(AB)^{k-1})$ 是定义良好的线性单射, 则

$$\alpha+\ker(B(AB)^k)\longmapsto AB\alpha+\ker(B(AB)^{k-1})$$

比较维数知 $\dim_F(\ker((AB)^{k+1}))-\dim_F(\ker(B(AB)^k))\leq\dim_F(\ker((AB)^k))-\dim_F(\ker(B(AB)^{k-1}))$, 即 $\dim_F(\ker((AB)^{k+1}))-\dim_F(\ker((AB)^k))\leq\dim_F(\ker(B(AB)^k))-\dim_F(\ker(B(AB)^{k-1}))$.

另一方面, 注意到 $\ker((BA)^kB)/\ker((BA)^{k-1}B)\longrightarrow\ker((BA)^k)/\ker((BA)^{k-1})$ 是定义良好的线性单射,

$$\alpha+\ker((BA)^{k-1}B)\longmapsto B\alpha+\ker((BA)^{k-1})$$

则比较维数知 $\dim_F(\ker((BA)^kB))-\dim_F(\ker((BA)^{k-1}B))\leq\dim_F(\ker((BA)^k))-\dim_F(\ker((BA)^{k-1}))$.

又由 $B(AB)^k=(BA)^kB$ 知, $\dim_F(\ker((AB)^{k+1}))-\dim_F(\ker((AB)^k))\le\dim_F(\ker((BA)^k))-\dim_F(\ker((BA)^{k-1}))$, 则 $r_k(AB,0)-r_{k+1}(AB,0)\le r_{k-1}(BA,0)-r_k(BA,0)$, 即 $w_{k+1}(AB,0)\le w_k(BA,0)$. 同理可证 $w_{k+1}(BA,0)\le w_k(AB,0)$. □

推论 9.5.8 *设 F 为代数闭域, $A\in F^{m\times n}$, $B\in F^{n\times m}$, 则:*

(1) *$\forall c\in F^*$, $\forall k\ge1$, AB 与 BA 的 Jordan 标准形中 $J_k(c)$ 的个数相同, 且关于 c 的第 k 大的 Jordan 块的阶数相同.*

(2) *$\forall k\ge1$, AB 与 BA 的 Jordan 标准形中关于 0 的第 k 大的 Jordan 块的阶数至多相差 1.*

证明: (1) 由命题 9.5.7与点图即知; (2) 由命题 9.5.7与引理 9.5.6即知. □

注: 注意 AB 与 BA 的 Jordan 标准形中关于 0 的 Jordan 块可能不同, 例如取 $A=E_{11}$, $B=E_{21}\in F^{3\times3}$, 则 $AB=0=\mathrm{diag}(J_1(0),J_1(0),J_1(0))$, $BA=E_{21}=\mathrm{diag}(J_2(0),J_1(0))$.

以下我们将证明推论 9.5.8的逆命题也成立, 即由满足条件的两个分划数列构造相应的矩阵.

命题 9.5.9 *设 F 为一个域, $s_1\ge s_2\ge\cdots$ 与 $s_1'\ge s_2'\ge\cdots$ 分别是 m 与 n 的分划, 若 $|s_k-s_k'|\le1$, $\forall k\ge1$, 则 $\exists A\in F^{m\times n}$, $B\in F^{n\times m}$, $s.t.$ $s_k=s_k(AB,0)$, $s_k'=s_k(BA,0)$.*

证明: 对于 $k\ge1$ 满足 $\min\{s_k,s_k'\}\ge1$, 以下构造 $A_k\in F^{s_k\times s_k'}$, $B_k\in F^{s_k'\times s_k}$, 满足 $A_kB_k=J_{s_k}(0)$, $B_kA_k=J_{s_k'}(0)$. 事实上,

(1) 若 $s_k=s_k'$, 则令 $A_k=J_{s_k}(0)$, $B_k=I_{s_k}$;

(2) 若 $s_k=s_k'+1$, 则令 $A_k=\begin{pmatrix}0\\I_{s_k'}\end{pmatrix}$, $B_k=\begin{pmatrix}I_{s_k'}&0\end{pmatrix}$;

(3) 若 $s_k+1=s_k'$, 则令 $A_k=\begin{pmatrix}I_{s_k}&0\end{pmatrix}$, $B_k=\begin{pmatrix}0\\I_{s_k}\end{pmatrix}$.

最后令 $A=\mathrm{diag}(A_1,A_2,\cdots,0,0,\cdots)\in F^{m\times n}$, $B=\mathrm{diag}(B_1,B_2,\cdots,0,0,\cdots)\in F^{n\times m}$ 即可. □

推论 9.5.10 *设 F 为代数闭域, $C\in F^{m\times m}$, $D\in F^{n\times n}$ 满足 ① $\forall c\in F^*$, $\forall k\ge1$, $s_k(C,c)=s_k(D,c)$;*
② $\forall k\ge1$, $|s_k(C,0)-s_k(D,0)|\le1$, 则 $\exists A\in F^{m\times n}$, $B\in F^{n\times m}$, $s.t.$ $C=AB$, $D=BA$.

证明: 通过相似变换, 可不妨设 C,D 均为 Jordan 标准形; 再由条件①可不妨设 $C=\mathrm{diag}(P,C_0)$, $D=\mathrm{diag}(P,D_0)$, 其中 P 为可逆的 Jordan 标准形, C_0, D_0 均为幂零的 Jordan 标准形. 由条件②知 $\forall k\ge1$, $|s_k(C_0,0)-s_k(D_0,0)|$ 则由引理 9.5.9知, 可取合适的 "Jordan 标准形" 矩阵 A_0,B_0 满足 $s_k(C_0,0)=s_k(A_0B_0,0)$, $s_k(D_0,0)=s_k(B_0A_0,0)$. 最后令 $A=\mathrm{diag}(P,A_0)$, $B=\mathrm{diag}(I,B_0)$ 即可. □

§9.6 正规阵理论

9.6.1 QR 分解的应用

命题 9.6.1 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 为上三角阵且可对角化, 则 A 可通过上三角阵相似于对角形.*

证明: 由 A 可对角化知, 存在 $P\in\mathrm{GL}(n,F)$, $D\in F^{n\times n}$ 为对角阵, 满足 $PAP^{-1}=D$. 取 QR 分解 $P=Q\cdot R$, 其中 $Q\in\mathrm{O}(n)$(若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$(若 $F=\mathbb{C}$), $R\in\mathrm{GL}(n,F)$ 为上三角阵, 则 $RAR^{-1}=Q^{-1}DQ$. 注意此式左端为上三角阵, 右端为正规阵, 故由引理 7.3.3知, 它们均为对角阵, 即 A 可通过上三角阵相似于对角形. □

命题 9.6.2 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$.*

(1) *若 $F=\mathbb{R}$, 则 $\exists O_1,O_2\in\mathrm{O}(n)$, $s.t.$ O_1AO_2 为准上三角形, 其中对角块为至多二阶的, 且 O_1BO_2 为上三角形.*

(2) *若 $F=\mathbb{C}$, 则 $\exists U_1,U_2\in\mathrm{U}(n)$, $s.t.$ U_1AU_2,U_1BU_2 均为上三角形.*

证明: 以 (1) 为例, (2) 完全类似. 由于 $O(n)$ 为紧群, 故通过扰动, 可不妨设 $B\in\mathrm{GL}(n,\mathbb{R})$. 由 Jordan 标准形与 QR 分解知, 存在 $O_0\in\mathrm{O}(n)$, 以及 $J_0\in\mathrm{GL}(n,\mathbb{R})$ 为准上三角形, 其中对角块为至多二阶的, 满足 $O_0^{-1}(B^{-1}A)O_0=J_0$. 取 QR 分解 $BO_0=Q\cdot R$, 其中 $Q\in\mathrm{O}(n)$, $R\in\mathbb{R}^{n\times n}$ 为上三角形, 则 $A=BO_0J_0O_0^{-1}=Q(RJ_0)O_0^{-1}$, $B=QRO_0^{-1}$, 其中 RJ_0 为准上三角形, 且对角块为至多二阶的. □

9.6.2 奇异值分解的应用

命题 9.6.3 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$.*

(1) *若 $F=\mathbb{R}$, 则 A,B 可同时正交相抵于对角形 $\Longleftrightarrow$ AB^t, B^tA 均为实自伴阵;*

(2) *若 $F=\mathbb{C}$, 则 A,B 可同时酉相抵于对角形 $\Longleftrightarrow$ $A\overline{B}^t$, $\overline{B}^tA$ 均为复正规阵.*

证明: (1)"⇒": 设 $O_1,O_2\in\mathrm{O}(n)$ 满足 $A=O_1D_AO_2^{-1}$, $B=O_1D_BO_2^{-1}$, 其中 $D_A,D_B\in\mathbb{R}^{n\times n}$ 为对角阵, 则 $AB^t=O_1(D_AD_B)O_1^{-1}$, $B^tA=O_2(D_BD_A)O_2^{-1}$, 故 AB^t, B^tA 均为实自伴阵.

"⇐": 由奇异值分解, 可取 $O_1,O_2\in\mathrm{O}(n)$, 满足 O_1AO_2 为非负对角形, 此时 $O_1(AB^t)O_1^{-1}=(O_1AO_2)(O_1BO_2)^t$, $O_2(B^tA)O_2^{-1}=(O_1BO_2)^t(O_1AO_2)$ 仍为实自伴阵, 故通过由 O_1AO_2 替换 A, 由 O_1BO_2 替换 B, 可不妨设 A 已为非负对角形, 且 $A=\mathrm{diag}(c_1I_{n_1},\cdots,c_kI_{n_k})$, 其中 $c_1>\cdots>c_k\geq 0$. 注意 $A^2B=A(A^tB)=AB^tA=(BA^t)A=BA^2$, 则直接计算知 B 也具有与 A 相同的准对角分块形式, 记为 $B=\mathrm{diag}(B_1,\cdots,B_k)$. 注意 AB^t 为实自伴阵, 则 $\forall 1\leq i\leq k$,

$c_iB_i^t$ 为实自伴阵. 当 $c_i>0$ 时, B_i 为实自伴阵, 则 $c_iI_{n_i}$ 与 B_i 可同时正交相似于对角形; 当 $c_i=0$ 时, 取 B_i 的奇异值分解, 则 $0I_{n_i}$ 与 B_i 可同时正交相抵于对角形. 综上, A 与 B 可同时正交相抵于对角形.

(2)"⇒: 与 (1)"⇒" 完全类似;

"⇐: 先与 (1)"⇐" 类似, 可不妨设 A 已为对角形. 注意 $A\overline{B}^t$, $\overline{B}^tA$ 均为复正规阵, 且 $(A\overline{B}^t)A=A(\overline{B}^tA)$, 则由 Fuglede 定理知, $\overline{(A\overline{B}^t)}^tA=A\overline{(\overline{B}^tA)}^t$, 即 $B|A|^2=|A|^2B$. 其余过程与 (1)"⇐" 类似. □

注: 特别地, A,B 可同时正交相抵 (若 $F=\mathbb{R}$) 或酉相抵 (若 $F=\mathbb{C}$) 于非负对角形 $\Longleftrightarrow$ $A\overline{B}^t$, $\overline{B}^tA$ 均为半正定阵.

推论 9.6.4 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A\in F^{n\times n}$ 的 QR 分解为 $A=Q\cdot R$,*

(1) *若 $F=\mathbb{R}$, 则 Q,R^t 可同时正交相抵于对角形 $\Longleftrightarrow$ A 为实自伴阵;*

(2) *若 $F=\mathbb{C}$, 则 $Q,\overline{R}^t$ 可同时酉相抵于对角形 $\Longleftrightarrow$ A 为复正规阵.*

引理 9.6.5 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$, 则以下条件等价:*

(1) *$\overline{A}^tA=\overline{B}^tB$;*

(2) *$\exists P_0,P_1,P_2\in\mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), s.t. $P_1AP_0=P_2BP_0$ 为对角形;*

(3) *$\exists P\in\mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), s.t. $B=PA$.*

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全类似.

"(1)⇒(2)": 由奇异值分解, 可取 $P_0,P_1\in\mathrm{O}(n)$, 满足 $P_1AP_0=D$, 其中 $D\in\mathbb{R}^{n\times n}$ 为对角形. 由 $A^tA=B^tB$ 知, $(P_1AP_0)^t(P_1AP_0)=(P_1BP_0)^t(P_1BP_0)$, 即 $D^2=\sqrt{(P_1BP_0)^t(P_1BP_0)}^2$. 由算术平方根算子的唯一性知, $|D|=\sqrt{(P_1BP_0)^t(P_1BP_0)}$. 再考虑 D 与 P_1BP_0 的极分解, 则 $\exists O\in\mathrm{O}(n)$, $s.t.$ $D=OP_1BP_0$. 取 $P_2=OP_1$ 即可.

"(2)⇒(3)⇒(1)": 显然. □

命题 9.6.6 *设 $F=\mathbb{R}$ 或 $\mathbb{C}$, $A,B\in F^{n\times n}$, 则以下条件等价:*

(1) *$\exists X,Y\in F^{n\times n}$, s.t. $\begin{pmatrix}A&X\\Y&B\end{pmatrix}\in\mathrm{O}(2n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$);*

(2) *$\forall\alpha\in F^{n\times 1}$, $\|A\alpha\|\leq\|\alpha\|$, 且 $\exists P,Q\in\mathrm{O}(n)$ (若 $F=\mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F=\mathbb{C}$), s.t. $A=PBQ$;*

(3) *A 的奇异值都属于 $[0,1]$, 且 A 与 B 的奇异值序列仅相差一个排列.*

证明: 以 $F=\mathbb{R}$ 为例, $F=\mathbb{C}$ 完全类似.

"(1)⇒(2)": 设 $X,Y\in\mathbb{R}^{n\times n}$, 满足 $\begin{pmatrix}A&X\\Y&B\end{pmatrix}\in\mathrm{O}(2n)$, 则 $A^tA+Y^tY=YY^t+BB^t=I_n$. 任取 $\alpha\in\mathbb{R}^{n\times 1}$, 则 $\|A\alpha\|^2=\alpha^tA^tA\alpha=\alpha^t\alpha-\alpha^tY^tY\alpha=\|\alpha\|^2-\|Y\alpha\|^2\leq\|\alpha\|^2$, 即 $\|A\alpha\|\leq\|\alpha\|$. 对于 $Z\in\{A,B,Y\}$, 记 $Z=P_Z\sqrt{Z^tZ}$ 为极分解, 其中 $P_Z\in\mathrm{O}(n)$, 则 $ZZ^t=(P_Z\sqrt{Z^tZ})(P_Z\sqrt{Z^tZ})^t=P_Z(Z^tZ)P_Z^{-1}$, 故 $A^tA=I_n-Y^tY=I_n-P_Y^{-1}(YY^t)P_Y=I_n-P_Y^{-1}(I_n-BB^t)P_Y=P_Y^{-1}(P_BB^tBP_B^{-1})P_Y=(BP_B^{-1}P_Y)^t(BP_B^{-1}P_Y)$. 由引理 9.6.5知, $\exists P\in\mathrm{O}(n)$, $s.t.$ $A=PBP_B^{-1}P_Y$. 取 $Q=P_B^{-1}P_Y$ 即可.

"(2)⇒(1)": 记 $A=O_1CO_2$ 为奇异值分解, 其中 $O_1,O_2\in\mathrm{O}(n)$, $C\in\mathbb{R}^{n\times n}$ 的对角分量均非负. 由 $\forall\alpha\in\mathbb{R}^{n\times 1}$, $\|A\alpha\|\leq\|\alpha\|$ 知, C 的对角分量均 ≤ 1, 则可取对角阵 $S\in\mathbb{R}^{n\times n}$, 满足 $C^2+S^2=I_n$. 此时 $\begin{pmatrix}C&-S\\S&C\end{pmatrix}\in\mathrm{O}(2n)$.

再由 $B = P^{-1}AQ^{-1} = P^{-1}O_1CO_2Q^{-1}$ 可知, $\begin{pmatrix} O_1 & 0 \\ 0 & P^{-1}O_1 \end{pmatrix} \cdot \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \cdot \begin{pmatrix} O_2 & 0 \\ 0 & O_2Q^{-1} \end{pmatrix} = \begin{pmatrix} A & * \\ * & B \end{pmatrix} \in \mathrm{O}(2n)$.

"(2)⇔(3)": 由奇异值的定义即知 (2),(3) 的前半句等价; 由奇异值分解即知 (2),(3) 的后半句等价. □

命题 9.6.7 (CS 分解) *设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $M = \begin{pmatrix} A & X \\ Y & B \end{pmatrix} \in \mathrm{O}(2n)$ (若 $F = \mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F = \mathbb{C}$), 其中 $A \in F^{n\times n}$, 则 $\exists P_1, P_2, Q_1, Q_2 \in \mathrm{O}(n)$ (若 $F = \mathbb{R}$) 或 $\mathrm{U}(n)$ (若 $F = \mathbb{C}$), s.t. $\begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \cdot \begin{pmatrix} A & X \\ Y & B \end{pmatrix} \cdot \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix} = \begin{pmatrix} C & -S \\ S & C \end{pmatrix}$, 其中 $C = \mathrm{diag}(c_1, \cdots, c_n)$, 且 $(0 \le)c_1 \le \cdots \le c_n (\le 1)$ 为 A 的奇异值, $S = \mathrm{diag}(\sqrt{1-c_1^2}, \cdots, \sqrt{1-c_n^2})$.*

证明: 以 $F = \mathbb{R}$ 为例, $F = \mathbb{C}$ 完全类似. 由奇异值分解知, $\exists P_1, Q_1 \in \mathrm{O}(n)$, s.t. $P_1AQ_1 = C$, 其中 $C = \mathrm{diag}(c_1, \cdots, c_n)$, 且 $0 \le c_1 \le \cdots \le c_n$. 记 $\begin{pmatrix} P_1 & 0 \\ 0 & I_n \end{pmatrix} \cdot \begin{pmatrix} A & X \\ Y & B \end{pmatrix} \cdot \begin{pmatrix} Q_1 & 0 \\ 0 & I_n \end{pmatrix} = \begin{pmatrix} C & X_1 \\ Y_1 & B \end{pmatrix} \in \mathrm{O}(2n)$. 由正交性知 $C^2 + Y_1^tY_1 = C^2 + X_1X_1^t = I_n$, 则 $Y_1^tY_1 = X_1X_1^t = I_n - C^2 = S^2$, 其中 $S = \mathrm{diag}(\sqrt{1-c_1^2}, \cdots, \sqrt{1-c_n^2})$. 再考虑它们的极分解知, $\exists P_2', Q_2 \in \mathrm{O}(n)$, s.t. $P_2'Y_1 = S = -X_1Q_2$, 则记 $\begin{pmatrix} I_n & 0 \\ 0 & P_2' \end{pmatrix} \cdot \begin{pmatrix} C & X_1 \\ Y_1 & B \end{pmatrix} \cdot \begin{pmatrix} I_n & 0 \\ 0 & Q_2 \end{pmatrix} = \begin{pmatrix} C & -S \\ S & B_1 \end{pmatrix} \in \mathrm{O}(2n)$.

现记 $0 \le c_1 \le \cdots \le c_k < c_{k+1} = \cdots = 1$, 以及 $s_i = \sqrt{1-c_i^2}\ (1 \le i \le n)$, 则 $1 \ge s_1 \ge \cdots \ge s_k > s_{k+1} = \cdots = 0$. 由行列的正交关系知, $B_1 = \mathrm{diag}(c_1, \cdots, c_k, B_{11})$, 其中 $B_{11} \in \mathrm{O}(n-k)$. 记 $P_2'' = \mathrm{diag}(I_k, B_{11}^{-1}) \in \mathrm{O}(n)$, 则 $P_2''S = S$, $P_2''B_1 = C$, 故 $\begin{pmatrix} I_n & 0 \\ 0 & P_2'' \end{pmatrix} \cdot \begin{pmatrix} C & -S \\ S & B_1 \end{pmatrix} = \begin{pmatrix} C & -S \\ S & C \end{pmatrix}$. 取 $P_2 = P_2''P_2'$ 即可. □

9.6.3 通过 Hermite 阵相似

命题 9.6.8 *设 $A \in \mathrm{GL}(n, \mathbb{C})$, 则以下条件等价:*

(1) *A 相似于酉阵;*

(2) *存在正定阵 $H \in \mathrm{GL}(n, \mathbb{C})$, 满足 $H^{-1}A^{-1}H = \overline{A}^t$;*

(3) *$\exists P, Q \in \mathrm{GL}(n, \mathbb{C})$, s.t. $A = P^{-1}Q$, & $\overline{P}^tP = \overline{Q}^tQ$.*

证明: "(1)⇒(2)": 设 $\exists R \in \mathrm{GL}(n, \mathbb{C})$, $U \in \mathrm{U}(n)$, s.t. $R^{-1}AR = U$, 则 $I_n = U \cdot \overline{U}^t = (R^{-1}AR) \cdot (\overline{R}^t\overline{A}^t(\overline{R}^t)^{-1})$, 即 $A^{-1}(R\overline{R}^t) = (R\overline{R}^t)\overline{A}^t$. 记 $H \coloneqq R\overline{R}^t \in \mathrm{GL}(n, \mathbb{C})$ 为正定阵即可.

"(2)⇒(3)": 记 $P \coloneqq H^{-\frac{1}{2}} \in \mathrm{GL}(n, \mathbb{C})$ 为正定阵, $Q \coloneqq PA \in \mathrm{GL}(n, \mathbb{C})$, 则 $A = P^{-1}(PA) = P^{-1}Q$, 且 $\overline{P}^tP = H^{-1} = \overline{A}^tH^{-1}A = \overline{PA}^t \cdot (PA) = \overline{Q}^tQ$.

"(3)⇒(1)": 由 $P, Q \in \mathrm{GL}(n, \mathbb{C})$ 且 $\overline{P}^tP = \overline{Q}^tQ$ 知, $QP^{-1} \in \mathrm{U}(n)$, 则 $A = P^{-1}Q = P^{-1}(QP^{-1})P$. □

命题 9.6.9 *设 $A \in \mathrm{GL}(n, \mathbb{C})$, 则以下条件等价:*

(1) *A^{-1} 相似于 $\overline{A}^t$;*

(2) *存在 Hermite 阵 $H \in \mathrm{GL}(n, \mathbb{C})$, 满足 $H^{-1}A^{-1}H = \overline{A}^t$;*

(3) *$\exists P \in \mathrm{GL}(n, \mathbb{C})$, s.t. $A = P^{-1}\overline{P}^t$.*

证明: "(1)⇒(2)": 设 $\exists R \in \mathrm{GL}(n, \mathbb{C})$, s.t. $R^{-1}A^{-1}R = \overline{A}^t$, 即 $R = AR\overline{A}^t$, 则 $\overline{R}^t = A\overline{R}^t\overline{A}^t$, 故 $\forall c \in \mathbb{C}^*$, $(cR + \overline{c}\overline{R}^t) = A(cR + \overline{c}\overline{R}^t)\overline{A}^t$. 由于 $(cR + \overline{c}\overline{R}^t) = cR(1 + c^{-1}\overline{c}R^{-1}\overline{R}^t)$, 且 $\{c^{-1}\overline{c} \colon c \in \mathbb{C}^*\} = S^1$, 则 $\exists c \in \mathbb{C}^*$, s.t. $\det(I_n + c^{-1}\overline{c}R^{-1}\overline{R}^t) \ne 0$, 即 $cR + \overline{c}\overline{R}^t \in \mathrm{GL}(n, \mathbb{C})$. 记 $H \coloneqq cR + \overline{c}\overline{R}^t \in \mathrm{GL}(n, \mathbb{C})$ 为 Hermite 阵即可.

"(2)⇒(3)": 由于 $\{c^{-1}\overline{c} \colon c \in \mathbb{C}^*\} = S^1$, 则 $\exists c \in \mathbb{C}^*$, s.t. $\det(I_n + c^{-1}\overline{c}\overline{A}^t) \ne 0$, 即 $cI_n + \overline{c}\overline{A}^t \in \mathrm{GL}(n, \mathbb{C})$. 记 $P \coloneqq (cI_n + \overline{c}\overline{A}^t) \cdot H^{-1} \in \mathrm{GL}(n, \mathbb{C})$, 则 $PA = cH^{-1}A + \overline{c}\overline{A}^tH^{-1}A = cH^{-1}A + \overline{c}H^{-1} = \overline{P}^t$, 即 $A = P^{-1}\overline{P}^t$.

"(3)⇒(1)": 由 $P \in \mathrm{GL}(n, \mathbb{C})$ 且 $A = P^{-1}\overline{P}^t$ 知, $A^{-1} = (\overline{P}^t)^{-1}P = (\overline{P}^t)^{-1}(P(\overline{P}^t)^{-1})\overline{P}^t = (\overline{P}^t)^{-1}\overline{A}^t\overline{P}^t$. □

命题 9.6.10 *设 $A \in \mathbb{C}^{n\times n}$, 则以下条件等价:*

(1) A 相似于实方阵;

(2) A 相似于 $\overline{A}^t$;

(3) A 可通过 *Hermite* 可逆阵相似于 $\overline{A}^t$;

(4) $A=H_1H_2$, 其中 $H_1,H_2\in\mathbb{C}^{n\times n}$ 为 *Hermite* 阵, 且至少它们之一可逆;

(5) $A=H_1H_2$, 其中 $H_1,H_2\in\mathbb{C}^{n\times n}$ 为 *Hermite* 阵.

证明: "(1)⇔(2)": 由命题 6.3.5以及 $\overline{A}$ 相似于 $\overline{A}^t$ 知, A 相似于实方阵 $\iff$ A 相似于 $\overline{A}\iff$ A 相似于 $\overline{A}^t$.
"(2)⇒(3)": 设 $\exists R\in\mathrm{GL}(n,\mathbb{C})$, *s.t.* $R^{-1}AR=\overline{A}^t$, 即 $AR=R\overline{A}^t$, 则 $A\overline{R}^t=\overline{R}^t\overline{A}^t$, 故 $\forall c\in\mathbb{C}^*$, $A(cR+\overline{c}\overline{R}^t)=(cR+\overline{c}\overline{R}^t)\overline{A}^t$. 由于 $(cR+\overline{c}\overline{R}^t)=cR(1+c^{-1}\overline{c}R^{-1}\overline{R}^t)$, 且 $\{c^{-1}\overline{c}\colon c\in\mathbb{C}^*\}=S^1$, 则 $\exists c\in\mathbb{C}^*$, *s.t.* $\det(I_n+c^{-1}\overline{c}R^{-1}\overline{R}^t)\neq 0$, 即 $cR+\overline{c}\overline{R}^t\in\mathrm{GL}(n,\mathbb{C})$. 记 $H\coloneqq cR+\overline{c}\overline{R}^t\in\mathrm{GL}(n,\mathbb{C})$ 为 Hermite 阵即可.
"(3)⇒(4)": 设存在 Hermite 阵 $H\in\mathrm{GL}(n,\mathbb{C})$, 满足 $H^{-1}AH=\overline{A}^t$, 则 $A=H\cdot(\overline{A}^tH^{-1})$. 注意 $\overline{(\overline{A}^tH^{-1})}^t=H^{-1}A=\overline{A}^tH^{-1}$.
"(4)⇒(2)": 不妨设 $H_1\in\mathrm{GL}(n,\mathbb{C})$, 则 $H_1^{-1}AH_1=H_2H_1=\overline{A}^t$.
"(4)⇒(5)": 显然.
"(5)⇒(1)": 由于 Hermite 阵可酉相似于实对角阵, 故 $\exists U\in\mathrm{U}(n)$, *s.t.* $U^{-1}AU=(U^{-1}H_1U)(U^{-1}H_2U)$, 其中 $U^{-1}H_1U=\begin{pmatrix}D&0\\0&0\end{pmatrix}$, $U^{-1}H_2U=\begin{pmatrix}K&*\\*&*\end{pmatrix}$, 且 $D_1\in\mathrm{GL}(k,\mathbb{C})$ 为实对角阵, $K\in\mathbb{C}^{k\times k}$, 则 $U^{-1}AU=\begin{pmatrix}DK&*\\0&0\end{pmatrix}$.
由 "(4)⇒(2)⇒(1)" 知, DK 相似于实方阵; 再由推论 6.3.6知, $U^{-1}AU$ 相似于实方阵, 即 A 相似于实方阵. □

§9.7 伪正交群

以下我们考虑实四维线性空间上的 "广义旋转", 它与物理中的 Lorentz 变换有关.

例 9.7.1 设 $V=\mathbb{R}^4$ 为实四维线性空间. 受物理中 *Minkowski* 时空坐标系的影响, 这里我们不考虑 $\mathbb{R}^4$ 上的标准实内积, 而考虑 $\mathbb{R}^{1+3}$ 上的 *"Lorentz* 标量积*"*(固定 $c>0$*):*

$$\begin{array}{ccc}\mathbb{R}^{1+3}\times\mathbb{R}^{1+3}&\longrightarrow&\mathbb{R}\\((t_1,x_1,y_1,z_1)^t,(t_2,x_2,y_2,z_2)^t)&\longmapsto&c^2t_1t_2-x_1x_2-y_1y_2-z_1z_2\end{array}.$$

保持此 *Lorentz* 标量积的线性变换称为 *Lorentz* 变换. 例如, 一个常见的 *Lorentz* 变换为 (固定 $|v|<c$*):*

$$\begin{cases}t\mapsto\dfrac{t-vx/c^2}{\sqrt{1-v^2/c^2}},\\ x\mapsto\dfrac{x-vt}{\sqrt{1-v^2/c^2}},\\ y\mapsto y,\\ z\mapsto z,\end{cases}$$ 这解释了狭义相对论中尺缩、钟慢、同时的相对性等现象.

在线性代数中, 记 $H=\{A\in\mathbb{C}^{2\times2}\colon A=\overline{A}^t\}$, 则存在实线性空间同构 $U\colon\begin{array}{ccc}\mathbb{R}^{1+3}&\longrightarrow&H\\\begin{pmatrix}t\\x\\y\\z\end{pmatrix}&\longmapsto&\begin{pmatrix}ct+x&y+z\sqrt{-1}\\y-z\sqrt{-1}&ct-x\end{pmatrix}\end{array}$,

满足 $\|\alpha\|_L^2=\det(U(\alpha))$, $\forall\alpha\in\mathbb{R}^{1+3}$. 于是利用实线性代数同构 $\begin{array}{ccc}L(H)&\xrightarrow{\cong}&L(\mathbb{R}^{1+3})\\T&\longmapsto&U^{-1}\circ T\circ U\end{array}$, 求 $\mathbb{R}^{1+3}$ 上的 *Lorentz* 变换只需求 H 上保持行列式的线性变换. 例如, 考虑实线性代数同态 $\begin{array}{ccc}\mathbb{C}^{2\times2}&\longrightarrow&L(H)\\M&\longmapsto&(T_M\colon A\mapsto MA\overline{M}^t)\end{array}$, 它限制在保绝对体积部分为群同态 $\{M\in\mathbb{C}^{2\times2}\colon|\det(M)|=1\}\to\{T\in L(H)\colon\det(T(A))=\det(A),\ \forall A\in H\}$.

现在我们考虑上述实线性代数同态的复合 $\mathbb{C}^{2\times2}\to L(H)\xrightarrow{\cong}L(\mathbb{R}^{1+3})$, 它限制在保绝对体积部分为群同态

$$\begin{aligned}&\{M\in\mathbb{C}^{2\times2}\colon|\det(M)|=1\}\to\{T\in L(H)\colon\det(T(A))=\det(A),\ \forall A\in H\}\\&\xrightarrow{\cong}\{S\in L(\mathbb{R}^{1+3})\colon\|S(\alpha)\|_L^2=\|\alpha\|_L^2,\ \forall\alpha\in\mathbb{R}^{1+3}\}=:\mathrm{O}(1,3).\end{aligned}$$

容易发现此复合群同态的核为 $\{cI_2 \colon c = \pm 1, \pm\sqrt{-1}\}$; 重要的是它非满射, 以下研究它的像的群性质. 首先, 上述群同态的右端是 $\mathbb{R}^{1+3}$ 上的 *Lorentz* 变换群 $\mathrm{O}(1,3)$, 它的矩阵表示为 $\left\{O \in \mathbb{R}^{4\times 4} \colon O^t \cdot \mathrm{diag}(1, -I_3) \cdot O = \mathrm{diag}(1, -I_3)\right\}$. 显然它不是连通的, 且有一个指标为 2 的子群 $\mathrm{SO}(1,3) \coloneqq \mathrm{O}(1,3) \cap \mathrm{SL}(4,\mathbb{R})$. 需要注意的是, $\mathrm{SO}(1,3)$ 也不是连通的, 且有一个指标为 2 的子群 $\mathrm{SO}^+(1,3) \coloneqq \mathrm{SO}(1,3) \cap \{O \in \mathbb{R}^{4\times 4} \colon O_{11} > 0\}$. 这里 $\mathrm{SO}^+(1,3)$ 是连通的. 进一步地, 可直接验证上述复合群同态的像恰为 $\mathrm{SO}^+(1,3)$.

综上, 此复合群同态可写成 $\mathrm{SL}(2,\mathbb{C}) \xrightarrow{2:1} \mathrm{SO}^+(1,3)$, 即 $\mathrm{PSL}(2,\mathbb{C}) \xrightarrow{\cong} \mathrm{SO}^+(1,3)$.