

高等代数(II)习题课教案

xiaxueq@stu.pku.edu.cn

June 2023

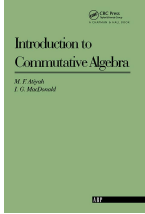
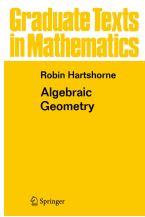

什么是线性代数	形式纯粹主义 必须是线性的	形式中立主义 在一个环上也行	形式自由主义 能加就行
内容纯粹主义 必须研究代数	 <p>高等代数是线性代数</p>	 <p>交换代数是线性代数</p>	 <p>椭圆曲线是线性代数</p>
内容中立主义 也可研究其它数学对象	 <p>泛函分析是线性代数</p>	 <p>代数几何是线性代数</p>	 <p>代数拓扑是线性代数</p>
内容自由主义 不必和数学有关	 <p>在线性感发牌也是线性代数</p>	 <p>指环王也是线性代数</p>	 <p>手抓饼难道不是线性代数吗?</p>



图 1: 高等代数习题课教员正在进行教学反思

修订版前言

2023年2月到6月, 笔者重新修订了这份习题课讲义. 一方面是为了与李文威老师的高等代数II课程相适应, 另一方面也是因为笔者深感初版讲义过于粗糙, 不堪入目. 所以在第二年讲授高等代数II习题课的过程中进行了重新编排和大幅度修订. 初版讲义按照讲授顺序编排, 而新版讲义按照课程内容进行了重新编排. 同时新版讲义为与今年正课课程内容相容, 新添加了许多内容, 如群论初步以及张量积部分都是此次新增的. 旧版讲义中一部分笔者认为不太具有启发性的习题在此次修订后也被删去了.

群论部分习题60-64以及张量积章节的所有习题来自李文威老师的平时作业. 多道习题来源于李尚志老师的各种著作. 习题57来自Michael Artin所著的 *Algebra* 一书. 部分习题解答参考了Nigel Hitchin所著的 *Projective Geometry* 和Keith Conrad所编写的讲义 *Simplicity of $PSL(n, F)$* . 在讲义的编写过程中笔者还参考了聂灵沼老师和丁石孙老师合著的《代数学引论》和Serge Lang的 *Algebra*. 此外在讲义的使用过程中吴永彤同学和艾心玥同学指出了几处错误, 在此一并表示感谢.

封二的白猫叫大宝. 它在新太阳地下书店学习过高等代数.

若您在阅读时感到不适, 请立刻停止阅读并及时向您的同伴/恋人/朋友寻求帮助. 必要时可以穿戴安全帽以避免头部受伤.

下雪

二零二三年夏 于 燕园

Disclaimer: dear collaborators, my work may contain the following

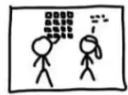
$g_{gh}\Gamma_{ef}^i$ $=\langle \nabla_{e_e} e_f, e_g \rangle$ <p>reused letters</p>	$\operatorname{Spec} \bigoplus_{i=1}^{+\infty} K_i$ $= \bigcup_{i=1}^{+\infty} \operatorname{Spec} K_i$ <p>false generalization</p>	$X^3 + Y^3 + Z^3 = 3\mu XYZ$ \Downarrow $y^2 = x^3 - \frac{2187\mu^4 + 17496\mu}{(\mu^3 - 1)^4} x + \dots$ <p>excessive concreteness</p>	<p>Prove that:</p> $\# \text{Primes} = +\infty$ $\prod \frac{1}{1-p^{-3}} = \zeta(3) \notin \mathbb{Q}$ <p>tempest in a teapot</p>
$\text{Any } X \rightarrow S$ $\operatorname{Spec} K \rightarrow X$ $\downarrow \quad \quad \downarrow$ $\operatorname{Spec} R \rightarrow S$ <p>ignored condition</p>	$c\operatorname{Prj}(Cnd)$ $= \operatorname{ExDsc}.$ <p>lack of explanation</p>	$\Xi \in \mathbb{C}^x, \text{ then}$ $\left \frac{\Xi}{\Xi} \right = 1$ <p>inconsiderate notation</p>	$R = R_{jkl}^i du^i \otimes \frac{\partial}{\partial u^j}$ $\otimes du^k \otimes du^l$ $R_{ikl,h}^j = \frac{\partial R_{ikl}^j}{\partial u^h}$ $- R_{pkl}^i \Gamma_{ih}^p + R_{ikl}^p \Gamma_{ph}^i$ <p>brutal computation</p>
$\lim_{\overline{I}} \lim_{\overline{J}} = \lim_{\overline{J}} \lim_{\overline{I}}$ $ I < +\infty, J \text{ filtered}$ <p>reversed arrows</p>	$\forall x, y \in \mathbb{R}, n \in \mathbb{Z}_+$ $x > (n-1)y', \forall y' \in \mathbb{R}$ $y' = \frac{n}{n-1}y \Rightarrow x > ny.$ <p>induction at will</p>	$\bigcap_{i=1}^{+\infty} \bigcup_{j=1}^{+\infty} U_{ij}$ $= \bigcup_{j=1}^{+\infty} \bigcap_{i=1}^{+\infty} U_{ij}$ <p>careless commutation</p>	<p>For quasi compact quasi separated, proper, finite type geometrically reduced projective variety / \mathbb{C}.</p> <p>unnecessary embellishment</p>
<p><u>THEOREM.</u> given</p> $\frac{\{r^{\wedge 2}\}}{\{r^{\wedge 2}\}}$ <p>handwritten LaTeX</p>	<p><u>LEMMA.</u> $\int \liminf_{n \rightarrow +\infty} f_n$</p> $\geq \liminf_{n \rightarrow +\infty} \int f_n.$ <p>reversed inequality</p>	<p>By Freyd-Mitchell, we can embed a small sub Abelian Cat of Mod_R into some $\operatorname{Mod}_{R'}$</p> <p>abstract nonsense</p>	 <p>fig 1.</p> <p>inline xkcd</p>

图 2: 讲义作者对阅读过程中可能造成的精神损伤概不负责

前言

本讲义为作者在2022年2月-6月期间在北京大学赵玉凤老师主讲的高等代数II所配套习题课上使用的讲义. 主要面向对象为数院和信科的大一同学. 内容包括多项式, Jordan标准形与有理标准形, 线性变换以及内积空间的相关习题.

高等代数作为数院大一三门基础课之一, 是每个数学系学生必须掌握的基本功, 也是开启代数方向学习必不可少的一门课程. 而练习则是每一门课程学习中不可或缺的重要组成部分, 几乎可以肯定的是: 如果一个学生没有做过充分多的练习, 那么他就不可能掌握高等代数. 然而练习的量只是一方面, 更为重要的是练习的“质”. 可以说, 大量重复而低质量的习题对学习只有百害而无一利, 例如: 做成百上千道求极限习题对学好数学分析毫无帮助反而只会产生虚假的满足感. 然而从初学者的角度而言, 没有人能指望一个新手可以一开始就判断出一道题目的价值高低, 这样挑选优质习题的任务就落到了习题课助教的身上.

出于这个原因, 作者在习题课上挑选了六十余道习题, 虽然不敢说一定都是好题, 但总体上遵循以下几条原则: 一. 题目新颖, 根据同学们的反馈, 许多习题他们之前并未见过, 这在不少人“刷丘砖”的贵校是一件不容易做到的事; 二. 重视知识间的联系, 往往同一堂习题课上, 前一道习题的结论立刻可以用在后一道习题上; 三. 紧扣正课大纲, 确保在完成每一道练习题中都能加深对正课所学知识的理解. 例如练习77将转置作为线性变换, 要求其Jordan标准形, 通常习惯了线性变换作为矩阵写出的学生面对这道题也许会大受震撼, 但是阅读了解答之后他们又能感受到线性变换概念独立于矩阵的合理性, 这对于破除线性映射等于矩阵的刻板印象大有好处. 再例如, Lagrange插值多项式 (练习35)并不单独出现, 而是作为中国剩余定理 (练习34)的推论产生, 并进一步给出Hermite插值 (练习36). 这样就将不同知识点串联在了一起, 使学生不至于产生“学了没用”的消极想法.

当然在“内卷”趋势愈演愈烈的当下, 也许在现在还十分新颖的选材, 几年后

由于教研的进步就变得略显陈旧;今天的好题,由于信息的传播,明天可能就变成了常见套路.因此作者也不能保证未来的读者在阅读本讲义时仍能同意上面几条原则.但作者相信任何时候都不会缺乏精妙的习题,到那时自然会有新的思想火花出现.

在讲义的编写过程中,作者参考了许多国内外的优质教材并选取了其中部分习题,如李尚志老师的《线性代数(数学专业用)》,丘维声老师的《高等代数》等等.赵玉凤老师布置的课后作业也是习题的重要组成部分,另外还有许多同学朋友提供了重要的习题素材.此外在这一学期的教学过程中,有同学指出了讲义中的几处笔误.在此向所有在讲义编写中提供帮助的老师同学表示衷心的感谢!另外作者还想特别感谢本科阶段遇到的几位特别认真负责的助教学长学姐,细致而耐心的你们是我学习的榜样.

囿于编写时间仓促以及作者能力水平所限,讲义中仍可能有错误或疏漏,望各界师生指出,以便及时修订.

下雪

二零二二年夏 于 燕园

目录

修订版前言	iii
前言	v
第一章 线性空间和线性变换	1
第二章 内积空间	11
第三章 多项式	27
第四章 群论初步	45
第五章 相似标准型理论	59
第六章 张量积	79
写在后面	85

第一章 线性空间和线性变换

Exercise 1 (子空间的并)

设 V 是无限域 F 上的有限维线性空间, V_1, \dots, V_s 是 V 的 s 个真子空间. 求证:

- (1) 存在 $\alpha \notin \bigcup_{k=1}^s V_k$;
- (2) 存在 V 的一组基 e_1, \dots, e_n 均不落在 $\bigcup_{k=1}^s V_k$ 中.

Solution 1

- (1) 对子空间个数做归纳:

当 $s = 1$ 时, 结论是显然的.

假设我们的结论已经对 $s - 1$ 成立: $\exists \alpha \notin V_1 \cup \dots \cup V_{s-1}$. 若 $\alpha \notin V_s$ 则无需再证, 因此接下来假设 $\alpha \in V_s$. 选取 F 中 s 个不同元素 c_1, c_2, \dots, c_s (F 是无限域)和 $\beta \notin V_s$. 我们断言: 在 s 个向量

$$c_1\alpha + \beta, \dots, c_s\alpha + \beta$$

中必有一者不落在 $V_1 \cup \dots \cup V_{s-1}$ 中. 假设我们的断言不成立, 那么由抽屉原理, s 个向量全部落在 $s - 1$ 个子空间的并中, 必定有一个子空间至少有两个向量 $c_i\alpha + \beta$ 和 $c_j\alpha + \beta$. 这样它们之差 $(c_i - c_j)\alpha$ 就落在这个子空间中, 与我们对 α 的选取矛盾. 于是我们的断言成立.

令 $c_i\alpha + \beta$ 为断言中不落在 $V_1 \cup \dots \cup V_{s-1}$ 中的向量, 由 $\alpha \in V_s$ 但 $\beta \notin V_s$ 知, $c_i\alpha + \beta \notin V_s$. 于是我们最终得到归纳假设对 s 也成立: $c_i\alpha + \beta \notin V_1 \cup \dots \cup V_{s-1} \cup V_s$.

(2) 反复利用(1)即可, 先用(1)取出 e_1 , 然后对 $V_1, \dots, V_s, V_{s+1} := \text{span}(e_1)$ 利用(1)取出 e_2 , 再对 $V_1, \dots, V_s, V_{s+1} := \text{span}(e_1, e_2)$ 利用(1)取出 e_3 . 以此类推直到 V_{s+1} 张成整个全空间 V .

Exercise 2 (*Fitting Lemma*)

设 V 是有限维线性空间, $\varphi : V \rightarrow V$ 为其上一线性变换, 证明:

$$\exists n \in \mathbb{N}_+ \text{ s.t. } V = \ker \varphi^n \oplus \text{im } \varphi^n.$$

Solution 2

先来证明两个链条件:

$$\begin{aligned} \exists m \in \mathbb{N}_+ \text{ s.t. } \ker \varphi^m = \ker \varphi^{m+1} = \ker \varphi^{m+2} = \dots \\ \ker \varphi^m = \ker \varphi^{m+1} = \ker \varphi^{m+2} = \dots \end{aligned} \quad (1.1)$$

事实上我们总是有:

$$\begin{aligned} \ker \varphi \subseteq \ker \varphi^2 \subseteq \ker \varphi^3 \subseteq \dots \\ \text{im } \varphi \supseteq \text{im } \varphi^2 \supseteq \text{im } \varphi^3 \supseteq \dots \end{aligned}$$

于是有

$$\begin{aligned} \dim \ker \varphi \leq \dim \ker \varphi^2 \leq \dim \ker \varphi^3 \leq \dots \leq n \\ \dim \text{im } \varphi \geq \dim \text{im } \varphi^2 \geq \dim \text{im } \varphi^3 \geq \dots \geq 0 \end{aligned}$$

最后的不等号是由于链中出现的线性空间都是 V 的子空间, 因此它们的维数总是大于等于0, 小于等于 n . 这迫使以上两个不等式在 m 充分大后总是取到等号. 因此我们总是有链条件1.1成立.

再证明直和式 $V = \ker \varphi^m \oplus \text{im } \varphi^m$ 成立, 为此:

1. 说明直和: $\ker \varphi^m \cap \text{im } \varphi^m = 0$ 若 $x \in \ker \varphi^m \cap \text{im } \varphi^m$, 则 $\varphi^m(x) = 0$, 存在 $y \in V$ 使得 $\varphi^m(y) = x$, 这样就有 $\varphi^{2m}(y) = 0$. 但是由于链条件 $\ker \varphi^m = \ker \varphi^{2m}$, y 也属于 $\ker \varphi^m = \ker \varphi^{2m}$, 这样就迫使 $x = \varphi^m(y) = 0$.

2. 再来找到直和分解: $\forall x \in V : \exists y \in \ker \varphi^m, z \in \text{im } \varphi^m \text{ s.t. } x = y + z$.

注意到 $\varphi^m(x) \in \text{im } \varphi^m = \text{im } \varphi^{2m} : \exists z \in \text{im } \varphi^m \text{ s.t. } \varphi^m(z) = \varphi^m(x)$

$$\therefore x = (x - z) + z$$

其中 $x - z \in \ker \varphi^m$, $z \in \operatorname{im} \varphi^m$.

注记1: 证明中我们没有用到维数公式, 事实上这是模论中Fitting Lemma的特例, 原条件为 V 为一Noetherian且Artinian模, 对应我们一开始证明的两个链条件.

注记2: 也可借助Jordan块和Jordan标准形处理. 考虑0-Jordan块和非0-Jordan块即可.

Exercise 3

令 $V = \mathbb{R}^3$, $S \subset V$ 为 V 中由 $2x - 2y + z = 0$ 定义的子空间, $P: V \rightarrow S \subset V$ 为 V 到 S 的投影映射.

(1) 求 P 在标准基 e_1, e_2, e_3 下的矩阵 A

(2) 计算 A^2

(3) 求证 $V = \ker P \oplus \operatorname{im} P$

Solution 3

(1) 令 $\alpha = (2, -2, 1)^T$. 注意到 $S = \{x \in V \mid \alpha^T x = 0\}$, 即 α 为 S 的法向量. 则投影方向由 α 确定, 投影后向量为 $Pe_i = e_i - \lambda_i \alpha$ 满足:

$$\alpha^T(e_i - \lambda_i \alpha) = 0 \implies \lambda_i = \frac{\alpha^T e_i}{\alpha^T \alpha}$$

求得 $\lambda_1 = \frac{2}{9}$, $\lambda_2 = -\frac{2}{9}$, $\lambda_3 = \frac{1}{9}$.

于是 $Pe_1 = (\frac{5}{9}, \frac{4}{9}, -\frac{2}{9})^T$, $Pe_2 = (\frac{4}{9}, \frac{5}{9}, \frac{2}{9})^T$, $Pe_3 = (-\frac{2}{9}, \frac{2}{9}, \frac{8}{9})^T$. 矩阵 A 为 $\begin{pmatrix} \frac{5}{9} & \frac{4}{9} & -\frac{2}{9} \\ \frac{4}{9} & \frac{5}{9} & \frac{2}{9} \\ -\frac{2}{9} & \frac{2}{9} & \frac{8}{9} \end{pmatrix}$.

(2) 直接计算得 $A^2 = A$.

(3) 由 $A^2 = A$ 知是直和, 且

$$\forall v \in V : x = \underbrace{(x - Ax)}_{\ker P} + \underbrace{Ax}_{\operatorname{im} P}$$

注记: 事实上, 投影算子就是由幂等定义的: $P^2 = P$, 这很容易想象: 所谓投影, 就是把高维空间中的物体(如牛奶盒)一脚踩扁踩到低维空间中去(踩瘪了的牛奶盒), 那当然踩一脚和踩两脚没有什么区别.

Exercise 4 (脑筋急转弯)

$$\text{令 } P = \begin{pmatrix} 1 & -2 \\ -2 & -1 \end{pmatrix}, \text{ 定义 } L: \begin{array}{ccc} \mathbb{C}^{n \times n} & \rightarrow & \mathbb{C}^{n \times n} \\ A & \mapsto & P^{-1}AP \end{array}$$

(1) 求 P 的极小多项式;

(2) 求 L 的极小多项式.

Solution 4

(1) 直接计算 $P^2 = \begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix}$, 所以 $m_P(x) = x^2 + 3$.

(2) 同样直接计算 $L^2(A) = P^{-1}(P^{-1}AP)P = P^{-2}AP^2 = (-\frac{1}{3}I)A(-3I) = A$, 所以 $m_L(x) = x^2 - 1$.

Exercise 5

设 $A \in \mathbb{C}^{n \times n}$ 的全体特征值为 $\lambda_1, \dots, \lambda_n$. $A = (a_{ij})$. 求证

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij} a_{ji}$$

Solution 5

由谱映射定理知: $\sum_{i=1}^n \lambda_i^2 = \text{tr } A^2 = \sum_{i=1}^n (A^2)_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} a_{ji}$.

注记: 这个做法是我在习题课上从同学们那学来的, 原本的做法相对复杂.

Exercise 6 (域扩张)

令 $h(x) \in \mathbb{Q}[x]$ 为 \mathbb{Q} 上一不可约多项式 ($\deg h = n$), $\alpha \in \mathbb{C}$ 满足 $h(\alpha) = 0$.
 $F = \{f(\alpha) | f \in \mathbb{Q}[x]\}$.

- (1) 求证 F 为一域;
- (2) 求证 F 为 \mathbb{Q} 上线性空间, $\dim F = n$.

Solution 6

- (1) 显然 F 关于加法乘法是封闭的, 我们只要说明非零元在 F 中都有乘法逆元即可. 若 $f \in \mathbb{Q}[x]$ 满足 $f(\alpha) \neq 0$, 则 $\gcd(f, h) = 1$ (h 不可约). 因此存在 $u, g \in \mathbb{Q}[x]$ 使得 $f(x)g(x) + u(x)h(x) = 1$. 则 $g(\alpha)$ 即为所欲求的乘法逆元.
- (2) 由 F 的定义显然可以知道 F 为 \mathbb{Q} 上的线性空间, 并且 $1, \alpha, \alpha^2, \dots, \alpha^{n-1} \in F$. 只要再说明它们构成一组基即可.

首先 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 的确张成整个 F , 为观察到这一点, 只需要注意到任何高于 n 次的多项式 $f \in \mathbb{Q}[x]$ 与 f 模掉 h 产生的余式在 F 中产生相同的像. 即 $f = q \cdot h + r$ ($\deg r < \deg h$) 蕴含 $f(\alpha) = r(\alpha)$.

再来说明线性无关性, 假设存在 c_0, \dots, c_{n-1} 使得

$$\sum_{k=0}^{n-1} c_k \alpha^k = 0,$$

则 $f(x) = \sum_{k=0}^{n-1} c_k x^k$ 满足 $f(\alpha) = 0$, 但 $\deg f \leq \deg h$, h 不可约, 这样就只能有 $f = 0$. 因此 $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$ 的确构成 F 的一组基.

注记: F 是 \mathbb{Q} 的一个扩张, F/\mathbb{Q} 称为域扩张, F 作为 \mathbb{Q} -线性空间的维数 $\dim_{\mathbb{Q}}(F)$ 称为 F/\mathbb{Q} 的扩张次数.

Exercise 7

设 V 是 n 维 F -线性空间, $U, W \subseteq V$ 分别为 V 的 m, r 维子空间, 且满足条件 $U + W = V$. 记

$$S = \{A \in \text{Hom}(V, V) | A(U) \subseteq U, A(W) \subseteq W\}$$

(1) 证明 S 是 $\text{Hom}(V, V)$ 的子空间;

(2) 求 $\dim S$.

Solution 7

(1) 显然, 若 V, W 都是 $A_1, A_2 \in \text{Hom}(V, V)$ 的不变子空间. 则 U, W 也是 $k_1 A_1 + k_2 A_2$ ($k_1, k_2 \in F$)的不变子空间. 于是 S 是 $\text{Hom}(V, V)$ 的子空间.

(2) 先来看简单的情形, 如果 $V = U \oplus W$. 那么分别选取 U 和 W 的一组基合并成 V 的一组基. 由于 U 和 W 都是 A -不变的, 在这组基下 $A \in S$ 的矩阵形如

$$\begin{pmatrix} \overset{m}{\boxed{\begin{smallmatrix} * & O \\ * & O \end{smallmatrix}}} \overset{r}{\boxed{\begin{smallmatrix} * & \\ & * \end{smallmatrix}}} \end{pmatrix}. \text{ 此时 } \dim S = m^2 + r^2.$$

再看一般的情况, 选取 $U \cap W$ 的一组基 e_1, \dots, e_d , 由维数公式知 $d = m + r - n$. 分别扩充成一组 U 的基和一组 W 的基, 合并成一组 V 的基. 由于 $U, W, U \cap W$ 都是 A -不变的, 此时 A 的矩阵形如:

$$\begin{pmatrix} \overset{m}{\boxed{\begin{smallmatrix} * & O & O \\ * & \boxed{\begin{smallmatrix} * & d \\ O & O \end{smallmatrix}} & d * \\ O & O & * \end{smallmatrix}}} \overset{r}{\boxed{\begin{smallmatrix} * & \\ & * \end{smallmatrix}}} \end{pmatrix}.$$

由小学知识计算*部分面积知:

$$\dim S = m^2 + r^2 - d \cdot n = m^2 + r^2 + n(n - m - r).$$

Exercise 8

设 $A \in \mathbb{R}^{n \times n}$. 对 $\alpha, \beta \in \mathbb{R}^n$, 定义 $f(\alpha, \beta) = \alpha^T A \beta$. 若 $\forall \alpha \in \mathbb{R}^n : f(\alpha, \alpha) = 0$. 求 A 满足的条件.

Solution 8

由 $f(\alpha, \alpha) = \alpha^T A \alpha = 0$ 知 $(\alpha^T A \alpha)^T = \alpha^T A^T \alpha = 0$. 相加得到 $\alpha^T (A^T + A) \alpha = 0$ ($\forall \alpha$). 但是 $A + A^T$ 是实对称方阵, 故 $A + A^T$ 定义的二次型为 0. 于是 $A + A^T = 0$, 即 $A = -A^T$, A 反对称.

反之由反对称方阵 A 定义的 f 一定满足 $\forall \alpha \in \mathbb{R}^n: f(\alpha, \alpha) = 0$.

注记: 当 A 可逆的时候这样的 f 定义了所谓的”辛内积” (这个条件对维数 n 有什么要求吗?).

Exercise 9 (Cochran 分解)

若 s 个 n 阶方阵 A_1, \dots, A_s 满足 $\sum_{k=1}^s A_k = I_n$. 求证以下三条等价:

- (i) $\forall k \in \{1, 2, \dots, s\}: A_k^2 = A_k$;
- (ii) $\sum_{k=1}^s \text{rank } A_k = n$;
- (iii) $\forall i, j \in \{1, 2, \dots, s\} (i \neq j): A_i A_j = O$.

Solution 9

(i) \implies (ii): 由 $A_k^2 = A_k$ 知 A_k 的最小多项式整除 $x^2 - x$, 因此 A_k 可对角化, 特别地 $\text{rank } A_k = \text{tr } A_k$. 所以

$$\sum_{k=1}^s \text{rank } A_k = \sum_{k=1}^s \text{tr } A_k = \text{tr } \sum_{k=1}^s A_k = \text{tr } I_n = n.$$

(ii) \implies (iii): 因为 $n = \text{rank } \sum_{k=1}^s A_k \leq \sum_{k=1}^s \text{rank } A_k = n$, 所以 $\mathbb{F}^n = \bigoplus_{k=1}^s \text{im } A_k$ 为直和 (思考: 为什么?). 于是对任意 $v \in \mathbb{F}^n$ 有唯一直和分解:

$$v = \sum_{k=1}^s v_k \quad (v_k \in \text{im } A_k)$$

但是 $I_n v = v = \sum_{k=1}^s A_k v$ 也是直和分解. 这迫使 $v_k = A_k v$.

而对任意 $w \in \text{im } A_j$, w 的直和分解式中 $i \neq j$ 处都是 0, 所以 $A_i w = 0$. 于是 $\text{im } A_j \subseteq \ker A_i$, 也就是 $A_i A_j = O$.

(iii) \implies (i):

$$A_k = A_k I_n = A_k \sum_{j=1}^s A_j = A_k^2.$$

注记: 在未来的学习中同学们还将见到许多形式与本例类似的定理: 如统计中的Cochran定理, 交换代数中的Artin环结构定理, 微分流形中的单位分解以及群表示论等等.

Exercise 10 (同时可对角与可交换)

若域 F 上 n 阶方阵 A 和 B 可对角化, 证明:

$\exists P \in GL_n(F)$ s.t. $D_1 := P^{-1}AP, D_2 := P^{-1}BP$ 均为对角阵 $\Leftrightarrow AB = BA$.

Solution 10

\Rightarrow : 这是简单的一边:

$$AB = (PD_1P^{-1})(PD_2P^{-1}) = PD_1D_2P^{-1} = PD_2D_1P^{-1} = (PD_2P^{-1})(PD_1P^{-1}) = BA$$

因为对角阵乘法可交换.

\Leftarrow : 首先回忆可对角化的含义:

$$\begin{aligned} A \text{ 可对角化} &\Leftrightarrow A \text{ 的全体特征子空间直和为全空间} \\ &\Leftrightarrow \text{存在一组 } A \text{ 的特征向量构成全空间的基} \\ &\Leftrightarrow A \text{ 的最小多项式无重根} \end{aligned}$$

记 A 的全体特征值为 $\text{Spec}(A) = \{\lambda_1, \dots, \lambda_s\}$. V 为全空间, V_{λ_i} 为从属于特征值 λ_i 的特征子空间, 则:

$$V = \bigoplus_{i=1}^s V_{\lambda_i}$$

注意到 $\forall v \in V_{\lambda_i}$:

$$A(Bv) = B(Av) = B(\lambda_i v) = \lambda_i(Bv) \quad (v \in V_{\lambda_i})$$

从而 Bv 也是一个从属于 λ_i 的 A 的特征向量, 它落在 V_{λ_i} 中. 因此 V_{λ_i} 为 B -不变子空间.

而 B 可对角化, 所以它在 V_{λ_i} 上的限制 $B|_{V_{\lambda_i}}$ 也可对角化 (想一想, 为什么?). 于是 V_{λ_i} 中存在一组基 $\beta_{i,1}, \dots, \beta_{i,\dim V_{\lambda_i}}$ 为 B 的特征向量. 显然它们也是 A 的特征向量. 这样

$$\beta_{1,1}, \dots, \beta_{1,\dim V_{\lambda_1}}, \dots, \beta_{i,1}, \dots, \beta_{i,\dim V_{\lambda_i}}, \dots, \beta_{s,1}, \dots, \beta_{s,\dim V_{\lambda_s}}$$

就构成了 V 的一组基, 将它们排成矩阵 P 即可同时对角化 A 和 B .

注记: 也可采用矩阵证法, 但相当繁琐: 先将 A 对角化到

$$\begin{pmatrix} \lambda_1 I & & \\ & \lambda_2 I & \\ & & \ddots \\ & & & \lambda_s I \end{pmatrix},$$

这样将 B 过渡到分块对角阵

$$\begin{pmatrix} * & & \\ & * & \\ & & \ddots \\ & & & * \end{pmatrix},$$

再分别对角化每一块. 这么做的背后

实质仍是我们上面的线性映射观点.

第二章 内积空间

Exercise 11 (钝角)

在 n 维欧氏空间 \mathbb{R}^n 中两两成钝角的向量最多有多少个? 叙述并证明.

Solution 11

$n = 2$ 时容易证明最多为3个, $n = 3$ 时也容易给出4个两两成钝角的构造 (正四面体中心分别向四个顶点连线).

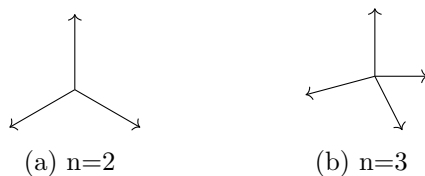


图 2.1: $n = 2, 3$ 时示意图

因此我们猜想: 在 n 维欧氏空间中至多有 $n + 1$ 个向量两两成钝角.

翻译: 两个向量成钝角的意思就是它们的内积小于0. 我们先来证明, 如果 $\alpha_1, \dots, \alpha_m$ 为 m 个两两成钝角的向量, 则前 $m - 1$ 个向量线性无关: 用反证法, 假设存在不全为零的 x_1, \dots, x_{m-1} 使得 $x_1\alpha_1 + \dots + x_{m-1}\alpha_{m-1} = 0$. 经过适当调换顺序, 不妨假设 $x_1 \geq \dots \geq x_r \geq 0, 0 > x_{r+1} \geq \dots \geq x_{m-1}$. 这样就有

$$\beta := x_1\alpha_1 + \dots + x_r\alpha_r = -x_{r+1}\alpha_{r+1} - \dots - x_{m-1}\alpha_{m-1}$$

再与 α_m 做内积得到:

$$\langle \beta, \alpha_m \rangle = \sum_{i=1}^r x_i \langle \alpha_i, \alpha_m \rangle = \sum_{j=r+1}^m -x_j \langle \alpha_j, \alpha_m \rangle < 0$$

严格的不等号是因为 x_1, \dots, x_{m-1} 不全为零, 所以 β 不为零向量.

但是我们又有:

$$0 < \langle \beta, \beta \rangle = \left\langle \sum_{i=1}^r x_i \alpha_i, \sum_{j=r+1}^{m-1} -x_j \alpha_j \right\rangle = \sum_{i=1}^r \sum_{j=r+1}^{m-1} -x_i x_j \langle \alpha_i, \alpha_j \rangle \leq 0.$$

这显然矛盾. 于是 $\alpha_1, \dots, \alpha_{m-1}$ 线性无关. 因此 $m-1 \leq n$. 这样我们就说明了至多只有 $n+1$ 个两两成钝角的向量.

而确实存在 $n+1$ 个两两成钝角的向量:

$$\begin{aligned} x_1 &= (-1, 0, 0, \dots, 0)^T \\ x_2 &= (1, -2, 0, \dots, 0)^T \\ &\vdots \\ x_k &= (1, 2, \dots, 2^{k-2}, -2^{k-1}, 0, \dots, 0)^T \\ &\vdots \\ x_n &= (1, 2, \dots, 2^{n-2}, -2^{n-1})^T \\ x_{n+1} &= (1, 2, \dots, 2^{n-2}, 2^{n-1})^T \end{aligned}$$

这样我们就完成了证明.

注记, 如果将钝角改成大于特定角度则问题将变得十分复杂, 关于这方面, 参见 Fejes Tóth's Problem (目前未解决).

Exercise 12

证明 \mathbb{R}^3 中方程

$$2x^2 + 4y^2 + 8z^2 - 2xy + 4xz + 6yz - 20 = 0$$

的图像是椭球面, 并求出这个椭球面所围成的立体体积. (已知椭球面 $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ 所围成的立体体积为 $\frac{4}{3}\pi abc$.)

Solution 12

将二次项整理成二次型对应的矩阵:

$$S = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 4 & 3 \\ 2 & 3 & 8 \end{pmatrix}.$$

其三个顺序主子式 $S_1 = 2$, $S_2 = 7$, $S_3 = \det S = 10$. 因此 S 正定, 存在正交方阵 P 使得 $P^{-1}SP = D = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$, 并且 $\lambda_1, \lambda_2, \lambda_3 > 0$, $\lambda_1\lambda_2\lambda_3 = \det S = 10$. 通过 P 建立新直角坐标系, 则在新坐标系下方程化为 $\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2 = 20$. 于是方程的图像是椭球面.

进一步整理方程得到 $\frac{(x')^2}{\frac{20}{\lambda_1}} + \frac{(y')^2}{\frac{20}{\lambda_2}} + \frac{(z')^2}{\frac{20}{\lambda_3}} = 1$. 所以椭球体积为 $V = \frac{4}{3}\pi\sqrt{\frac{20}{\lambda_1} \cdot \frac{20}{\lambda_2} \cdot \frac{20}{\lambda_3}} = \frac{4}{3}\pi\sqrt{800} = \frac{80}{3}\pi\sqrt{2}$.

Exercise 13 (*Min-Max* 原理变式)

设 $A: X \mapsto AX$ 是带标准内积的欧氏空间 $\mathbb{R}^4 \rightarrow \mathbb{R}^3$ 的线性映射. 其中 $A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{pmatrix}$. 求在条件 $\|x\| = 1$ 下, $\|Ax\|$ 能取到的最大值和最小值, 并确定在何处取到.

Solution 13

由 $A^T A$ 为半正定实对称矩阵知 $A^T A$ 可以正交相似到对角阵 $A^T A = Q^T D Q$, D 为一对角阵, 主对角线上元素均 ≥ 0 . 因此

$$\begin{aligned} \max_{\|x\|=1} \|AX\|^2 &= \max_{\|x\|=1} \langle Ax, Ax \rangle = \max_{\|x\|=1} x^T A^T A x = \max_{\|x\|=1} x^T Q^T D Q x \\ &= \max_{\|x\|=1} (Qx)^T D (Qx) = \max_{\|x\|=1} x^T D x = \lambda_{\max}(A^T A) \end{aligned}$$

其中倒数第二个等号是因为 Q 是正交阵, 它在单位球面 $\|x\| = 1$ 上是双射, $\lambda_{\max}(A^T A)$ 表示 $A^T A$ 的最大特征值.

同理我们有 $\min_{\|x\|=1} \|Ax\|^2 = \lambda_{\min}(A^T A)$. 取到最大值和最小值的位置分别为 $\lambda_{\max}(A^T A)$ 和 $\lambda_{\min}(A^T A)$ 的单位特征向量.

本题中 $\lambda_{\max}(A^T A) = 10$, $\lambda_{\min}(A^T A) = 0$. 对应最大值点和最小值点为 $\pm \frac{1}{\sqrt{90}} \begin{pmatrix} -1 \\ -2 \\ 2 \\ 9 \end{pmatrix}$ 和 $\pm \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 2 \\ -2 \\ 1 \end{pmatrix}$.

注记: 最大值点和最小值点可能有多个, 不要遗漏!

Exercise 14 (同时相合对角化)

设 \mathbb{F} 为任意特征不为 2 的域 ($\text{char } \mathbb{F} \neq 2$). $A, B \in \mathbb{F}^{n \times n}$ 为对称阵, A 可逆. 证明:

存在可逆阵 P 将 A, B 同时相合到对角阵 $\Leftrightarrow A^{-1}B$ 可对角化.

Solution 14

\Rightarrow : 设 $P^T AP = D_1$, $P^T BP = D_2$, D_1, D_2 为对角阵.

则 $P^{-1}A^{-1}P^{T^{-1}} = D_1^{-1} \implies A^{-1} = PD_1^{-1}P^T$, 同时 $B = P^{T^{-1}}D_2P^{-1}$. 因此 $A^{-1}B = PD_1^{-1}D_2P^{-1}$ 可对角化.

\Leftarrow : 反之 $A^{-1}B$ 可对角化: $P^{-1}A^{-1}BP = D$, D 为对角阵. 令 $S_1 = P^T AP$, $S_2 = P^T BP$, 只要证明 S_1, S_2 可以同时被一可逆阵相合到对角阵即可. 直接计算有 $S_1^{-1}S_2 = P^{-1}A^{-1}BP = D = \text{diag}(\lambda_1, \dots, \lambda_n)$. 于是 $S_2 = S_1 D$, 记 S_1 的 (i, j) -元为 s_{ij} , 由 S_1 和 S_2 对称知 $s_{ij} = s_{ji}$, $\lambda_j s_{ij} = \lambda_i s_{ji}$ ($\forall i, j$). 这样要么有 $\lambda_i = \lambda_j$, 要么有 $s_{ij} = s_{ji} = 0$. 于是不妨令 $D = \text{diag}(\lambda_1 I, \dots, \lambda_s I)$, 其中 $\lambda_1, \dots, \lambda_s$ 两两不同. 则 S_1 和 S_2 分块对角: $S_1 = \text{diag}(M_1, \dots, M_s)$, $S_2 = \text{diag}(N_1, \dots, N_s)$, 且 $N_i = \lambda_i M_i$ ($i = 1, \dots, s$). 设可逆阵 Q_i 将 M_i 相合到对角阵 D_i (由于 $\text{char } \mathbb{F} \neq 2$, 这总是能做到). 令 $Q = \text{diag}(Q_1, \dots, Q_s)$: 则我们有

$$Q^T S_1 Q = \text{diag}(Q_1^T, \dots, Q_s^T) \cdot \text{diag}(M_1, \dots, M_s) \cdot \text{diag}(Q_1, \dots, Q_s) = \text{diag}(D_1, \dots, D_s),$$

$$Q^T S_2 Q = \text{diag}(Q_1^T, \dots, Q_s^T) \cdot \text{diag}(N_1, \dots, N_s) \cdot \text{diag}(Q_1, \dots, Q_s) = \text{diag}(\lambda_1 D_1, \dots, \lambda_s D_s).$$

于是可逆阵 Q 将 A, B 同时相合到对角阵.

Exercise 15

设 $U = \text{span}(\alpha_1, \alpha_2)$ 为 \mathbb{R}^4 的子空间(带标准内积), 其中 $\alpha_1 = (1, 0, 1, 0)^T$, $\alpha_2 = (1, 1, 0, 0)^T$.

- (1) 求 U^\perp 的维数和它的一组正交基;
- (2) 求 $\alpha = (1, 1, 1, 1)^T$ 在 U 上的正交投影;
- (3) 求点 $(1, 1, 1, 1)$ 到 U 的最短距离.

Solution 15

(1) 易知 $\dim U^\perp = 2$. 由

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0$$

得 $U^\perp = \{a_1(-1, 1, 1, 0)^T + a_2(0, 0, 0, 1)^T | a_1, a_2 \in \mathbb{R}\}$, 这也是一组正交基.

(2) 设 $\alpha = u + u^\perp = u + a_1(-1, 1, 1, 0)^T + a_2(0, 0, 0, 1)^T$ ($u \in U$, $u^\perp \in U^\perp$, $a_1, a_2 \in \mathbb{R}$).

则 $\left\langle \alpha, \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 3a_1 = 1$, $\left\langle \alpha, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = a_2 = 1$. 于是 $a_1 = \frac{1}{3}$, $a_2 = 1$.

从而 $u = \alpha - \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4/3 \\ 2/3 \\ 2/3 \\ 0 \end{pmatrix}$.

(3) 即求 $\min_{v \in U} \langle \alpha - v, \alpha - v \rangle$. 但这是直接的:

$$\begin{aligned} & \min_{v \in U} \langle \alpha - v, \alpha - v \rangle \\ &= \min_{v \in U} \left\langle u - v + \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, u - v + \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle \\ &= \min_{v \in U} \langle u - v, u - v \rangle + \frac{1}{3} + 1 \\ &= \frac{4}{3} \end{aligned}$$

所以最短距离为 $\frac{2}{\sqrt{3}}$.

Exercise 16

给定矩阵 $A \in \mathbb{R}^{n \times n}$, 证明:

A 为两个 $\mathbb{R}^{n \times n}$ 中正定对称阵乘积 $\Leftrightarrow A$ 在 \mathbb{R} 上可对角化, 且特征值均为正数.

Solution 16

\Rightarrow : 设 $A = S_1 S_2$ 为两个正定对称阵的乘积, 则存在正交阵 $Q \in O(n)$ 将 S_1 正交相似到对角阵: $S_1 = Q^{-1} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q$. 令 $P = Q^{-1} \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} Q$. 那么:

$$P^T = Q^T \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} Q^{-1T} = Q^{-1} \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} Q = P$$

于是 $P = P^T$, $S_1 = P^2 = P^T P$.

这样

$$A = S_1 S_2 = P^2 S_2 \sim P^{-1} P^2 S_2 P = P S_2 P = P^T S_2 P$$

而 S_2 为正定对称阵, $P^T S_2 P$ 为正定二次型, 其特征值全为正数. 所以

$$A \sim P^T S_2 P \sim \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} > 0$$

\Leftarrow : 将 A 正交相似到对角阵, 并从中写出乘积分解:

$$\begin{aligned} A &= Q^{-1} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q \\ &= Q^{-1} (Q^{T^{-1}} Q^T) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q \\ &= (Q^{-1} Q^{-1T}) \left(Q^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q \right) \end{aligned}$$

注记: 前半部分证明中出现的 P 称为正定对称阵 S_1 的平方根, 它也是正定对称的.

Exercise 17 (Witt 扩张定理)

令 V 为带标准内积的有限维实线性空间. 给定 V 中两组向量 $S = \{\alpha_1, \dots, \alpha_m\}$ 和 $T = \{\beta_1, \dots, \beta_m\}$, 且满足 $\langle \alpha_i, \alpha_j \rangle = \langle \beta_i, \beta_j \rangle$ ($\forall 1 \leq i, j \leq m$), 则存在 V 上正交变换将 α_i 映到 β_i .

Solution 17

设 $\dim V = n$. 并令 $A = (\alpha_1, \dots, \alpha_m)$, $B = (\beta_1, \beta_m)$. 则由题设条件知 $A^T A = B^T B$. 又由它们是半正定实对称阵知存在正交方阵 P 使得

$$P^T A^T A P = P^T B^T B P = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0).$$

令 $D = \text{diag}(1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_r}, 1, \dots, 1)$, 则

$$(APD)^T(APD) = (BPD)^T(BPD) = \text{diag}(I_r, O_{n-r}).$$

记 $A_1 = APD, B_1 = BPD$. 上式说明 A_1 (B_1 同理) 的前 r 列为单位长度的互相正交的向量, 后 $n-r$ 列为 0. 分别将 A_1 和 B_1 的列向量组扩充为 V 的一组标准正交基并排成方阵 R 和 S . 则 R 和 S 为正交阵, $Q = SR^{-1}$ 仍为正交阵. Q 将 R 的前 r 列变到 S 的前 r 列, 即变 A_1 的前 r 列到 B_1 的前 r 列. 但 A_1, B_1 的后 $n-r$ 列均为 0. 故 $QA_1 = B_1$. 展开得 $QAPD = BPD$, 因为 P 和 D 均可逆, 所以最终有 $QA = B$.

Exercise 18

设 A 是正规方阵 ($AA^* = A^*A$). λ 和 μ 是 A 的两个不同特征值. α, β 分别是属于 λ 和 μ 的特征向量. 求证 $\langle \alpha, \beta \rangle = 0$.

Solution 18

先说明一个引理: $\ker A = \ker A^*$. 这是因为:

$$x \in \ker A \Leftrightarrow Ax = 0 \Leftrightarrow \langle Ax, Ax \rangle = 0 \Leftrightarrow x^* A^* Ax = 0 \Leftrightarrow x^* AA^* x = 0$$

$$\Leftrightarrow \langle A^* x, A^* x \rangle = 0 \Leftrightarrow A^* x = 0 \Leftrightarrow x \in \ker A^*$$

进一步得到推论, 正规方阵 A 关于 λ 的特征向量 α 也是 A^* 关于 $\bar{\lambda}$ 的特征向量:

$$A\alpha = \lambda\alpha \Leftrightarrow \alpha \in \ker(A - \lambda I) \Leftrightarrow \alpha \in \ker(A^* - \bar{\lambda} I) \Leftrightarrow A^*\alpha = \bar{\lambda}\alpha.$$

这是因为 $A - \lambda I$ 也是正规方阵.

于是:

$$\begin{aligned} \langle \alpha, A\beta \rangle &= \langle \alpha, \mu\beta \rangle = \mu \langle \alpha, \beta \rangle \\ &\parallel \\ \langle A^*\alpha, \beta \rangle &= \langle \bar{\lambda}\alpha, \beta \rangle = \bar{\lambda} \langle \alpha, \beta \rangle \end{aligned}$$

$$\text{因此 } (\lambda - \mu) \langle \alpha, \beta \rangle = 0 \implies \langle \alpha, \beta \rangle = 0.$$

Exercise 19

设 A 是 n 阶实正规方阵. $(a + bi)$ 是 A 的一个特征值 ($a, b \in \mathbb{R}, b \neq 0$). $\alpha + \beta i$ ($\alpha, \beta \in \mathbb{R}^n$)为对应的特征向量. 求证 α, β 正交, 长度相等.

Solution 19

显然 $(a - bi)$ 也是 A 的一个特征值 (实系数多项式的非实根成对出现). 且 $(\alpha - \beta i)$ 为对应的特征向量. 由 $b \neq 0$ 以及练习18知 $\langle \alpha + \beta i, \alpha - \beta i \rangle = 0$.

由双线性性展开得: $(\alpha^T \alpha - \beta^T \beta) + (-2\alpha^T \beta)i = 0$. 于是 $\alpha^T \alpha = \beta^T \beta, \alpha^T \beta = 0$. 此即 α, β 正交, 长度相等.

Exercise 20

证明: 复方阵 A 是正规方阵当且仅当存在复系数多项式 $f(\lambda)$ 使得 $A^* = f(A)$.

Solution 20

\Leftarrow : 这是容易的, $A^* A = f(A) A = A f(A) = A A^*$, 因此 A 是正规方阵.

\Rightarrow : 由 A 是正规方阵知, A 可以酉相似对角化: 存在酉方阵 U 使得 $U^{-1} A U = \text{diag}(\lambda_1, \dots, \lambda_n)$. 于是 $(U^{-1} A U)^* = U^{-1} A^* U = \text{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$. 问题转化为了构造多项式 f 使得 $f(\text{diag}(\lambda_1, \dots, \lambda_n)) = \text{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$. 事实上这只需要 $f(\lambda_1) = \overline{\lambda_1}, \dots, f(\lambda_n) = \overline{\lambda_n}$. 由Lagrange插值多项式(见练习35)知这样的 f 存在. 于是 $U^{-1} f(A) U = f(U^{-1} A U) = U^{-1} A^* U$, 即 $f(A) = A^*$.

Exercise 21

若 A 是实正规方阵, 且对实方阵 B 有 $AB = BA$. 证明也有 $A^T B = B A^T$

Solution 21

由练习20知存在多项式 $f(\lambda)$ 使得 $A^T = f(A)$, $f(\lambda) = \sum_{k=0}^n c_k \lambda^k$. 于是

$$A^T B = f(A) B = \sum_{k=0}^n c_k A^k B = \sum_{k=0}^n c_k B A^k = B f(A) = B A^T.$$

Exercise 22 (*Schur*上三角化)

证明任何复方阵都可以经酉矩阵相似到一个上三角阵.

Solution 22

对方阵阶数 n 做归纳. $n = 1$ 时结论显然成立. 若任何 $n - 1$ 阶复方阵都可以经酉矩阵相似到上三角阵. 任取 λ 为 A^* 的特征值, v 为一个 λ 对应的特征向量, 通过归一化不妨假设 $\|v\| = 1$. 取 v 对应的正交补

$$W = \{w \in \mathbb{C}^n \mid \langle w, v \rangle = 0\}.$$

则 W 是 A 的不变子空间: 任取 $w \in W$, $\langle Aw, v \rangle = \langle w, A^*v \rangle = \langle w, \lambda v \rangle = \lambda \langle w, v \rangle = 0$. 且 $\dim W = n - 1$. 由归纳假设知道 $A|_W$ 可以由酉矩阵相似到上三角阵, 翻译成线性映射的语言就是存在 W 的一组标准正交基 $\{w_1, \dots, w_{n-1}\}$ 使得 $A|_W$ 在这组基下满足 $A|_W w_i \in \text{span}(w_i, \dots, w_{n-1})$.

这样就有 $\{v, w_1, \dots, w_{n-1}\}$ 构成 \mathbb{C}^n 的一组标准正交基. 且显然 A 在这组基下满足 $Aw_i \in \text{span}(w_i, \dots, w_{n-1})$, $Av \in \text{span}(v, w_1, \dots, w_{n-1})$. 于是这组标准正交基对应的酉矩阵将 A 相似到上三角阵.

注记: 当 A 正规时 v 也是 A 的特征值 (见练习18中证明), 于是这事实上给出了正规算子可以酉对角化的另一个证明.

Exercise 23 (*Schur*不等式)

设 A 为 n 阶复方阵, $\lambda_1, \dots, \lambda_n$ 是 A 的全体特征值. 证明

$$\text{tr } A^* A \geq \sum_{k=1}^n |\lambda_k|^2,$$

其中等号成立当且仅当 A 为正规方阵.

Solution 23

利用Schur酉上三角化 (练习22), 将 A 写成 $U^* T U$ 的形式, 其中 U 为一酉矩阵, T 为上三角矩阵. 所以

$$\text{tr } A^* A = \text{tr } U^* T^* U U^* T U = \text{tr } U^* (T^* T) U = \text{tr } (U U^*) (T^* T) = \text{tr } T^* T.$$

记

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & t_{nn} \end{pmatrix},$$

显然 t_{11}, \dots, t_{nn} 为 A 的全体特征值. 直接计算得知

$$\operatorname{tr} T^* T = \sum_{i=1}^n \sum_{j=1}^i \overline{t_{ji}} t_{ji} = \sum_{i=1}^n \sum_{j=1}^i |t_{ji}|^2 \geq \sum_{i=1}^n |t_{ii}|^2 = \sum_{k=1}^n |\lambda_k|^2,$$

且等号成立当且仅当所有对角线以上元素 t_{ij} ($i < j$) 为零, 即 T 为对角阵. 于是等号成立时确实有 A 可以酉相似到对角阵, 此即 A 为正规方阵. 反之 A 为正规方阵时当然可以酉相似到对角阵, 必然有 $\operatorname{tr} A^* A = \sum_{k=1}^n |\lambda_k|^2$.

Exercise 24

设复方阵 A 满足 $A^* = -A$, 说明 $I \pm A$ 可逆, 而且 $(I - A)(I + A)^{-1}$ 是酉矩阵.

Solution 24

若向量 $v \in \mathbb{C}^n$ 满足 $(I + A)v = 0$, 即 $v = -Av$, 则

$$0 \leq \langle v, v \rangle = \langle v, -Av \rangle = \langle v, A^* v \rangle = \langle Av, v \rangle = \langle -v, v \rangle = -\langle v, v \rangle \leq 0,$$

这迫使 $\langle v, v \rangle = 0$, 由 Hermite 内积正定性知 $v = 0$. 于是 $(I + A)$ 可逆.

同理若 $v = Av$, 则

$$0 \leq \langle v, v \rangle = \langle Av, v \rangle = \langle v, A^* v \rangle = \langle v, -Av \rangle = -\langle v, v \rangle \leq 0,$$

仍然有 $v = 0$. 于是 $(I - A)$ 也可逆.

这样直接计算 $((I - A)(I + A)^{-1})^*((I - A)(I + A)^{-1})$ 得

$$\begin{aligned} & ((I - A)(I + A)^{-1})^*((I - A)(I + A)^{-1}) \\ &= (I + A)^{-1*}(I - A)^*(I - A)(I + A)^{-1} \\ &= (I - A)^{-1}(I + A)(I - A)(I + A)^{-1} \\ &= (I - A)^{-1}(I - A)(I + A)(I + A)^{-1} \\ &= I. \end{aligned}$$

注记: 法国人名“Hermite”的“H”不发音. 关于法语有这样一个笑话:

How do Chinese laugh?

- Hahahaha.

How do the French laugh?

- Aaaaaaa... Because the “H” is silent!

Exercise 25

若规范方阵 A, B 交换, 则它们可以由同一个酉方阵对角化.

Solution 25

这里与练习10是完全类似的. 注意到练习10中 A, B 可对角化的条件在本题中已经被 A, B 是规范方阵所满足. 而本题所要求的用酉方阵对角化无非就是在挑选 A, B 的公共特征向量成为全空间一组基的基础上增加这组基还是一组标准正交基的要求. 这可以通过Gram-Schmidt正交化得到. 细节留给读者自证.

Exercise 26

设 $n \geq 2$. Jacobi矩阵是具有以下形式的 $n \times n$ 实矩阵:

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \\ & c_2 & a_3 & b_3 & \\ & & \ddots & \ddots & \ddots \\ & & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{pmatrix}, \quad \forall i, b_i c_i > 0.$$

(1) 证明Jacobi矩阵总能在 \mathbb{R} 上对角化.

(2) 证明Jacobi矩阵有 n 个相异实特征值.

Solution 26

- (1) 首先我们证明 A 可以实相似到实对称阵. 注意到 $b_1 c_1 > 0$, 因此通过初等行变换第二行乘以 $\sqrt{\frac{b_1}{c_1}}$ 再通过初等列变换第二列乘以 $\sqrt{\frac{c_1}{b_1}}$, 得到

$$A' = \begin{pmatrix} a_1 & b_1 \sqrt{\frac{c_1}{b_1}} & & & & \\ c_1 \sqrt{\frac{b_1}{c_1}} & a_2 & b_2 \sqrt{\frac{b_1}{c_1}} & & & \\ & c_2 \sqrt{\frac{c_1}{b_1}} & a_3 & b_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & & c_{n-1} & a_n \end{pmatrix}.$$

这样就有 $b_1 \sqrt{\frac{c_1}{b_1}} = c_1 \sqrt{\frac{b_1}{c_1}}$. 而上述行变换为左乘矩阵 $\begin{pmatrix} 1 & & & \\ & \sqrt{\frac{b_1}{c_1}} & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$, 列变

换为右乘矩阵 $\begin{pmatrix} 1 & & & \\ & \sqrt{\frac{c_1}{b_1}} & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$, 两个矩阵恰巧互为对方的逆. 因此 A 和 A' 相

似. 且 A' 仍满足与 A 相同的条件: 即上对角线和下对角线对应元素乘积大于零. 我们可以重复上述操作, 将第三列和第三行分别乘以 $\sqrt{\frac{c_2}{b_2}}$ 和 $\sqrt{\frac{b_2}{c_2}}$, 第四列和第四行分别乘以 $\sqrt{\frac{c_3}{b_3}}$ 和 $\sqrt{\frac{b_3}{c_3}}$... 以此类推, 我们通过一系列相似变换将 A 逐渐使其上对角线和下对角线的对应元素相等. 这样就证明了 A 的确可以实相似到实对称阵, 由熟知的实对称阵可以在 \mathbb{R} 上对角化知 A 也可以在 \mathbb{R} 上对角化.

- (2) 若 $\lambda \in \mathbb{R}$ 为 A 的特征值, 我们来证明 λ 对应的特征子空间只能是一维的. 考虑方程 $(A - \lambda I)x = 0$. 令 x_i 为 x 的第 i 个坐标分量, 则 x 完全由 x_1 决定. 论证如下:

由 $a_1 x_1 + b_1 x_2 = 0$ 知 $x_2 = -\frac{a_1}{b_1} x_1$ (显然由 $b_1 c_1 > 0$ 知 $b_1 \neq 0$), 于是 x_2 完全由 x_1 确定.

再由 $c_1 x_1 + a_2 x_2 + b_2 x_3 = 0$ 知 $x_3 = -\frac{c_1}{b_2} x_1 - \frac{a_2}{b_2} x_2$, 于是 x_3 也完全由 x_1 确定.

以此类推, $c_{k-1} x_{k-1} + a_k x_k + b_k x_{k+1} = 0 \implies x_{k+1} = -\frac{a_k}{b_k} x_k - \frac{c_{k-1}}{b_k} x_{k-1}$. 这

样 x_2, x_3, \dots, x_n 都完全由 x_1 确定. 所以

$$1 \leq \dim \ker(A - \lambda I) \leq 1.$$

λ 对应的特征子空间只能是一维的. 由前一问可对角化已经知道 A 有 n 个实特征值, 且所有特征子空间维数之和为 n . 这迫使所有特征值各不相同 (否则 A 只有少于 n 个相异特征值且每一个特征值对应特征子空间都是一维的, 它们的维数之和不能等于 n .)

另解: 我们介绍实代数几何中的一个基本事实:

Theorem 1 (Baby version of Sturm-Sylvester Theorem).

一个 $\mathbb{R}[x]$ 上的多项式序列 $\{f_0, \dots, f_m\}$ 被称为 f_0 的一个**伪 Sturm 序列**, 如果它满足以下条件

ST 1. f_m 为一个非零常数.

ST 2. 若 $x_0 \in \mathbb{R}$ 是某个 f_i ($1 \leq i \leq m-1$) 的根, 则 $f_{i-1}(x_0) \cdot f_{i+1}(x_0) < 0$.

若实数 u, v ($u < v$) 不是任何 f_i 的根, 那么 f_0 在区间 $[u, v]$ 内 (不计重数) 的实根个数满足

$$\#\{x \in [u, v] | f_0(x) = 0\} \geq W_S(u) - W_S(v).$$

其中 $W_S(u)$ 表示序列 $\{f_0(u), f_1(u), \dots, f_m(u)\}$ 的符号变化次数 (如 $\{+1, -1, -1, +1, +1\}$ 的符号变化次数为2). 我们称 $W_S(u)$ 为伪 Sturm 序列 $\{f_0, \dots, f_m\}$ 在 u 处的**变号数**.

证明. 令 $\alpha_1 < \alpha_2 < \dots < \alpha_r$ 为所有 f_i 在 $[u, v]$ 上的根. 则显然 $W_S(x)$ 在任何一个开区间 (α_j, α_{j+1}) 上不会变化 (这是因为多项式是连续函数, 由介值定理知道每一个 f_i 在 (α_j, α_{j+1}) 上保持定号, 所以 $\{f_0(x), \dots, f_m(x)\}$ 在该开区间上保持相同的符号序列, 自然也保持变号数 $W_S(x)$ 不变.) 这样只要考虑 $W_S(x)$ 在跨过某个 f_i 的根 α_j 时的变化即可. 我们将证明 $W_S(x)$ 在跨过 f_1, \dots, f_{m-1} 的根时不变.

若 α_j 是某个 f_i ($1 \leq i \leq m-1$) 的根, 则由**ST 2**知道 $f_{i-1}(\alpha_j)f_{i+1}(\alpha_j) < 0$. 因此无论 $f_i(x)$ 的符号在跨过 α_j 时如何变化, $W_S(x)$ 都不会改变 (要么 $f_i(x)$ 和 $f_{i-1}(x)$ 异号, 要么 $f_i(x)$ 和 $f_{i+1}(x)$ 异号, 总之 $\{f_{i-1}(x), f_i(x), f_{i+1}(x)\}$ 的变号数是1.) 因此 $W_S(x)$ 在跨过 f_1, \dots, f_{m-1} 的根时不变.

下面再来分析 $W_S(x)$ 跨过 $f_0(x)$ 的根时的变化情况, 若 α_j 是 f_0 的根. 那么 $\{f_0(x), f_1(x)\}$ 只有四种变化可能:

- (a) 从同号变成同号, 变号数 $W_S(x)$ 不变.
- (b) 从同号变成异号, 变号数 $W_S(x)$ 加一.
- (c) 从异号变成同号, 变号数 $W_S(x)$ 减一.
- (d) 从异号变成异号, 变号数 $W_S(x)$ 不变.

无论如何变化, $W_S(x)$ 在跨过 f_0 的根时至多增加或减少1. 总结上述讨论, 我们知道 f_i ($1 \leq i \leq m-1$)的根对 $W_S(x)$ 没有贡献, f_0 的每个根至多会使 $W_S(x)$ 变化 ± 1 . 所以我们有 $\#\{x \in [u, v] | f_0(x) = 0\} \geq W_S(u) - W_S(v)$. \square

回到本题, 令 $D_i(\lambda)$ 为 $(\lambda I - A)$ 的第 i 个顺序主子式, 特别地令 $D_0 = 1$. 则由行列式展开知

$$\begin{aligned} D_0 &= 1; \\ D_1 &= \lambda - a_1; \\ D_2 &= (\lambda - a_2)D_1 - b_1c_1D_0; \\ &\vdots \\ D_n &= (\lambda - a_n)D_{n-1} - b_{n-1}c_{n-1}D_{n-2}. \end{aligned}$$

于是 $\{D_n, D_{n-1}, \dots, D_1, D_0\}$ 构成一个伪Sturm序列: 由行列式展开知 $D_{i+1} + b_i c_i D_{i-1} = (\lambda - a_{i+1})D_i$, 因此若 x_0 是 D_i 的根, 则要么 x_0 也同时是 D_{i+1} 和 D_{i-1} 的根, 要么 D_{i+1} 和 D_{i-1} 在 x_0 处异号. 但前者不可能, 这是因为 $D_i + b_{i-1}c_{i-1}D_{i-2} = (\lambda - a_i)D_{i-1}$ 迫使 x_0 也是 D_{i-2} 的根, 以此类推 x_0 是常数多项式 D_0 的根得到矛盾.

现在利用上述Sturm定理证明 D_n 确实有 n 个相异实根: 由 $D_i(\lambda)$ 的首项为 λ^i 知 $W_S(-\infty) = n$, $W_S(+\infty) = 0$. 所以

$$n \geq \#\{\lambda_0 \in \mathbb{R} | D_n(\lambda_0) = 0\} \geq W_S(-\infty) - W_S(+\infty) = n.$$

于是 D_n 确实有 n 个相异实根, 这证明了 A 有 n 个相异实特征值, 这也蕴含了 A 可以在 \mathbb{R} 上对角化.

下附完整版本的Sturm-Sylvester定理:

Theorem 2 (Sturm-Sylvester). 设 $f, g \in \mathbb{R}[x]$. 又令 $u, v \in \mathbb{R}$ ($u < v$), 且 u, v 都不是 f 的根. 则定义 f, g 的**Sturm序列**为:

$$\begin{aligned} f_0 &= f, f_1 = f'g, \\ f_i &= f_{i-1}q_i - f_{i-2}, \text{ 其中 } q_i \in \mathbb{R}[x], \deg f_i < \deg f_{i-1}, \\ f_m &= \gcd(f_0, f_1). \end{aligned}$$

即 $-f_i$ 为 f_{i-2} 除以 f_{i-1} 的余式 (注意负号). 则 (u, v) 上使 g 为正的 f 实根数减去使 g 为负的 f 实根数恰为Sturm序列两端变号数之差:

$$\begin{aligned} & \#\{x_0 \in (u, v) | f(x_0) = 0, g(x_0) > 0\} - \#\{x_0 \in (u, v) | f(x_0) = 0, g(x_0) < 0\} \\ & = W_S(u) - W_S(v). \end{aligned}$$

特别地, 令 $g = 1$, 则有 (u, v) 内 $f(x)$ 的实根个数恰好等于

$$W_S(u) - W_S(v).$$

证明是类似的, 只是需要对 $W_S(x)$ 在跨过 $f_0 = f$ 的实根时做更加细致的分析, 这部分留作练习. 我们特别指出, 通过Sturm-Sylvester定理和二分法, 我们可以求出一元多项式方程的实解. 更进一步的我们能利用Sturm-Sylvester定理证明实代数几何中的Tarski-Seidenberg原理, 它断言实代数几何中所有一阶命题都可以通过算法判定真伪.

Exercise 27

设 H_1, H_2 都是 n 阶正定Hermite方阵, 且 $H_1 - H_2$ 正定, 求证: $H_2^{-1} - H_1^{-1}$ 正定.

Solution 27

设可逆方阵 P 满足 $P^*P = H_1$, 则

$$H_1 - H_2 = P^*P - H_2 = P^*(I - (P^*)^{-1}H_2P^{-1})P > 0.$$

令酉方阵 Q 将 $(P^*)^{-1}H_2P^{-1}$ 相似到对角阵 $D = \text{diag}(d_1, \dots, d_n) > 0$:

$$(P^*)^{-1}H_2P^{-1} = Q^*DQ,$$

则 $H_1 - H_2 = P^*Q^*(I - D)QP > 0$. 从而对任意 i : $d_i \in (0, 1)$. 这样

$$\begin{aligned} H_2^{-1} - H_1^{-1} &= P^{-1}Q^*D^{-1}Q(P^*)^{-1} - P^{-1}(P^*)^{-1} \\ &= P^{-1}Q^*(D^{-1} - I)Q(P^*)^{-1} > 0. \end{aligned}$$

注记: 证明的实质是同时相合对角化. 结果也可以通过练习14和练习16得到.

练习: 设 H_1, H_2 都是 n 阶正定Hermite方阵, 且 H_1 正定. 求证 $H_1 + H_2$ 正定的充分必要条件是 $H_1^{-1}H_2$ 的特征值都大于 -1 .

第三章 多项式

Exercise 28

分母有理化 $\frac{1}{3+2\sqrt[3]{2}+\sqrt[3]{4}}$.

Solution 28

首先 $f(x) = x^3 - 2$ 满足 $f(\sqrt[3]{2}) = 0$, 再令 $g(x) = x^2 + 2x + 3$. 那么我们要求的就是

$$\frac{1}{g(x)} \Big|_{x=\sqrt[3]{2}} = \frac{u(x)}{u(x)g(x)} \Big|_{x=\sqrt[3]{2}} = \frac{u(x)}{u(x)g(x) + v(x)f(x)} \Big|_{x=\sqrt[3]{2}}.$$

于是问题转化为是否存在 $u(x), v(x)$ 使得 $ug + vf \in \mathbb{Q}$. 由 $ug + vf$ 的形式立即想到 $d(x) = \gcd(f(x), g(x))$ 也有相同的形式, 于是计算 $d(x)$. 由扩展欧几里得算法:

$$x^3 - 2 = (x - 2)(x^2 + 2x + 3) + (x + 4)$$

$$x^2 + 2x + 3 = (x - 2)(x + 4) + 11$$

$$\implies (x^2 + 2x + 3) = (x - 2)[(x^3 - 2) - (x - 2)(x^2 + 2x + 3)] + 11$$

$$\implies (x^2 + 4x + 5)g(x) - (x - 2)f(x) = 11$$

即 $d(x) = 11$, $u(x) = x^2 - 4x + 5$, $v(x) = -(x - 2)$.

因此:

$$\frac{1}{3 + 2\sqrt[3]{2} + \sqrt[3]{4}} = \frac{\sqrt[3]{4} - 4\sqrt[3]{2} + 5}{u(\sqrt[3]{2})g(\sqrt[3]{2}) + v(\sqrt[3]{2})f(\sqrt[3]{2})} = \frac{\sqrt[3]{4} - 4\sqrt[3]{2} + 5}{11}.$$

注记: 事实上 $\mathbb{Q}(\sqrt[3]{2}) = \mathbb{Q}[\sqrt[3]{2}] \cong \mathbb{Q}[x]/(x^3 - 2)$ 构成一数域, 见练习6.

Exercise 29

能否在 $\mathbb{R}[x]$ 中找到非平凡多项式 $f(x), g(x), h(x)$ 使得 $f^2(x) = x(g^2(x) + h^2(x))$?
若将 $\mathbb{R}[x]$ 换成 $\mathbb{C}[x]$ 又如何?

Solution 29

在 \mathbb{R} 上: 因为 $\deg f^2 = 2 \deg f$. 而且 g^2 和 h^2 的首项系数都是正数, 因此它们的首项不可能互相抵消, 所以 $\deg(g^2 + h^2) = \max\{2 \deg g, 2 \deg h\}$. 这样 $2 \deg f$ 是一个偶数, $\deg(x(g^2 + h^2))$ 是一个奇数, 它们不可能相等.

而在 \mathbb{C} 上情况则有所不同, 此时 g^2 和 h^2 的首项可能互相抵消, 如: $g(x) = (\frac{1}{2}x + \frac{1}{2})$, $h(x) = (\frac{1}{2}ix - \frac{1}{2}i)$. 那么 $g^2(x) + h^2(x) = x$, 从而有 $f(x) = x$ 满足 $f^2 = x(g^2 + h^2)$.

Exercise 30

$$\text{令 } f(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}.$$

- (1) 证明 $f(x)$ 没有重根;
- (2) $f(x)$ 在 \mathbb{R} 上有多少个根 (不记重数)?
- (3) $f(x)$ 在 \mathbb{C} 上有多少个根 (不记重数)?

Solution 30

- (1) 首先计算 $f'(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{n-1}}{(n-1)!}$, 而

$$\gcd(f(x), f'(x)) = \gcd(f(x) - f'(x), f'(x)) = (\frac{1}{n!}x^n, f'(x)).$$

注意到 $\frac{1}{n!}x^n$ 的因式里只有 x 的幂次, 但显然 $x \nmid f'(x)$, 所以 $\gcd(f(x), f'(x)) = 1$.
 $f(x)$ 没有重根.

- (2) 显然 $n = 0$ 时没有根, $n = 1$ 时有一个根. 下证明 $2 \mid n$ 时 $f(x)$ 在 \mathbb{R} 上无根, $2 \nmid n$ 时恰有一个根.

假设命题已对 $n < 2k$ 时成立, 我们证明对于 $n = 2k, n = 2k + 1$ 也成立 ($k \in \mathbb{N}_+$):

$n = 2k$ 时, $f(x)$ 为偶次数多项式, 它在 \mathbb{R} 上有最小值 $f(x_0)$, 且 $f'(x_0) = 0$. 注意到 $x_0 \neq 0$ ($f'(0) = 1 \neq 0$), 因此最小值 $f(x_0) = f'(x_0) + \frac{x_0^n}{n!} > 0$. $f(x) > 0$, $f(x)$ 在 \mathbb{R} 上无根.

$n = 2k + 1$ 时, 上面已证 $f'(x) > 0$, 因此 $f(x)$ 在 \mathbb{R} 上严格单调递增, 显然存在唯一的 $x_0 \in \mathbb{R}$ 使得 $f(x_0) = 0$. $f(x)$ 在 \mathbb{R} 上有一个根.

- (3) 根据代数基本定理, $f(x)$ 在 \mathbb{C} 上有 n 个根 (记重数). 而由(1)知道 $f(x)$ 没有重根, 因此 $f(x)$ 在 \mathbb{C} 上不记重数地也有 n 个根.

注记: 本练习和练习29一起说明了改变讨论的基域, 不但会影响多项式的分解和可约性, 还会影响多项式的解集.

Exercise 31

令 K 为一域, 且 K 非代数闭. 以下将分步证明 K 上的方程组都可以用一个方程表示:

- (1) 存在多项式 $f \in K[x, y]$ 使得 $f(x, y) = 0$ 当且仅当 $x = y = 0$;
- (2) 进一步说明对于任意正整数 n , 存在多项式 $g \in K[x_1, \dots, x_n]$ 使得 $g(x_1, \dots, x_n) = 0$ 当且仅当 $x_1 = \dots = x_n = 0$;
- (3) 证明对于任意一组方程组 $f_1(x_1, \dots, x_n) = \dots = f_s(x_1, \dots, x_n) = 0$, 都存在一个多项式 $h(x_1, \dots, x_n)$ 使得它们的解相同.

Solution 31

- (1) 因为 K 不是代数闭域, $K[x]$ 上存在多项式 $f_0(x) = \sum_{k=0}^m a_k x^k$ ($a_k \in K, a_m \neq 0$) 在 K 上无根. 将其齐次化, 令

$$f(x, y) = \sum_{k=0}^m a_k x^k y^{m-k}.$$

则显然 $x = y = 0$ 为 $f(x, y) = 0$ 的平凡解. 若还有非平凡解 (x_0, y_0) , 则 y_0 必不为0 (否则代入 f 后得到 $a_m x_0^m = 0$, 迫使 x_0 为零). 于是 $f(x_0, y_0)/y_0^m = \sum_{k=0}^m a_k (x_0/y_0)^k$, 即 $\frac{x_0}{y_0}$ 为 f_0 在 K 上的根, 这与 f_0 的选取矛盾. 因此 $f(x, y) = 0$ 当且仅当 $x = y = 0$.

- (2) 用归纳法, 若命题对 $n-1$ 成立: 记 $n-1$ 情形的多项式为 $g_{n-1}(x_1, \dots, x_{n-1})$, 则 $g(x_1, \dots, x_n) := f(g_{n-1}(x_1, \dots, x_{n-1}), x_n)$ 满足要求.
- (3) 令 $h(x_1, \dots, x_n) = g(f_1(x_1, \dots, x_n), \dots, f_s(x_1, \dots, x_n))$ 即可, 其中 g 为第二问中多项式.

注记: 本练习推广了 \mathbb{R} 上 $f_1 = \dots = f_s = 0 \Leftrightarrow f_1^2 + \dots + f_s^2 = 0$ 的结果. 这进一步告诉我们在非代数闭域中讨论方程的解是很糟糕的, 因此将来我们在解方程时如没有特殊说明, 一律在代数闭域上求解 (如 \mathbb{C}).

Exercise 32

\mathbb{C} 上有多项式 $f(x) = x^3 + 6x^2 + 3ax + 8$, 问 a 取何值时 $f(x)$ 有重根. 并求出此时 $f(x)$ 的根.

Solution 32

$f'(x) = 3x^2 + 12x + 3a$, 于是带余除法得: $f(x) = \frac{x+2}{3}f'(x) + (2a-8)(x-1)$

Case 1. $a = 4$, 从而 $f(x) = x^3 + 6x^2 + 12x + 8 = (x+2)^3$, $x_1 = x_2 = x_3 = -2$.

Case 2. $a \neq 4$, 为使 $(f(x), f'(x)) \neq 1$, 只能有 $(x-1)|f(x)$. 故 $3a+15=0$, $a=-5$, 此时 $f(x) = x^3 + 6x^2 - 15x + 8 = (x-1)^2(x+8)$, $x_1 = x_2 = 1$, $x_3 = -8$.

另解: 事实上我们有关于结式的定理:

Theorem 3. 设 $A = a_0x^d + \dots + a_d$, $B = b_0x^e + \dots + b_e$ 为一整环 R 上的单变元多项式. 定义 A 和 B 的 **Sylvester 结式** 为:

$$\text{Res}(A, B) = \begin{vmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdots & 0 \\ a_2 & a_1 & \ddots & 0 & b_2 & b_1 & \ddots & 0 \\ \vdots & \vdots & \ddots & a_0 & \vdots & \vdots & \ddots & b_0 \\ a_d & a_{d-1} & \cdots & \vdots & b_e & b_{e-1} & \cdots & \vdots \\ 0 & a_d & \ddots & \vdots & 0 & b_e & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{d-1} & \vdots & \vdots & \ddots & b_{e-1} \\ 0 & 0 & \cdots & a_d & 0 & 0 & \cdots & b_e \end{vmatrix}$$

即前 e 列为 A 的系数错位排列, 后 d 列为 B 的系数错位排列.

则 A 和 B 有非常数公因式当且仅当 $\text{Res}(A, B) = 0$, 特别地 A 和 B 在一个包含 R 的代数闭域上有公共根当且仅当 $\text{Res}(A, B) = 0$.

证明. 记 P_n 为 $\text{Frac } R$ 上次数小于 n 的多项式集合, 则映射

$$\varphi: \begin{array}{ccc} P_e \times P_d & \rightarrow & P_{d+e} \\ (u, v) & \mapsto & uA + vB \end{array}$$

的矩阵行列式恰为 $\text{Res}(A, B)$. 因此 A, B 有公因子当且仅当 $\ker \varphi \neq 0$, 也就当且仅当 $\text{Res}(A, B) = 0$. \square

回到本题, $f(x)$ 和 $f'(x)$ 的Sylvester结式为:

$$\text{Res}(f, f') = \begin{vmatrix} 1 & 0 & 3 & 0 & 0 \\ 6 & 1 & 12 & 3 & 0 \\ 3a & 6 & 3a & 12 & 3 \\ 8 & 3a & 0 & 3a & 12 \\ 0 & 8 & 0 & 0 & 3a \end{vmatrix} = 3(2880 - 864a - 108a^2 + 36a^3)$$

因式分解得到 $\text{Res}(f, f') = 108(-4 + a)^2(5 + a)$. 因此 f 有重根当且仅当 $a = 4$ 或 $a = -5$. 这是纯粹机械的计算, 无需动脑.

Exercise 33

设 $A, B \in \mathbb{R}^{n \times n}$ 为两实系数对称矩阵, A 正定. 试证明对充分大的 t , $tA + B$ 正定.

Solution 33

由 A 正定知存在可逆方阵 $P \in \mathbb{R}^{n \times n}$ 使得 $P^T A P = I$. 因此有相合关系: $tI + P^T B P = tP^T A P + P^T B P = P^T (tA + B) P$. 由于相合保持正定性, 因此只要证明 $tI + P^T B P$ 当 t 充分大时正定即可. 考虑 $tI + P^T B P$ 的顺序主子式 $D_1(t), \dots, D_n(t)$, 它们都是以 t 为变元的多项式, 且 $D_k(t)$ 的首项为 t^k . 所以当 t 充分大时对所有 $k = 1, \dots, n$ 都有 $D_k(t) > 0$. 此即 $tI + P^T B P$ 正定.

注记: 事实上, 关于多项式的根, 我们有著名的柯西界 (Cauchy's Bound). 因此我们不但可以证明 $tA + B$ 在 t 充分大时正定, 还可以计算得到一个 M 使得对于任意 $t > M$ 都有 $tA + B$ 正定.

Theorem 4 (柯西界). 若 x_0 满足多项式方程 $x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$, 则:

$$|x_0| \leq 1 + \max\{|a_0|, \dots, |a_{n-1}|\}.$$

证明. 若 $|x_0| \leq 1$ 则结论显然. 以下假设 $|x_0| > 1$. 因为三角不等式, 我们有

$$|x_0|^n = \left| \sum_{k=0}^{n-1} a_k x_0^k \right| \leq \sum_{k=0}^{n-1} |a_k| |x_0|^k \leq \max_{0 \leq k \leq n-1} |a_k| \cdot \sum_{k=0}^{n-1} |x_0|^k = \max_{0 \leq k \leq n-1} |a_k| \cdot \frac{|x_0|^n - 1}{|x_0| - 1}.$$

又因为 $|x_0| > 1$, 进一步放缩得到:

$$|x_0|^n \leq \max_{0 \leq k \leq n-1} |a_k| \cdot \frac{|x_0|^n - 1}{|x_0| - 1} \leq \max_{0 \leq k \leq n-1} |a_k| \cdot \frac{|x_0|^n}{|x_0| - 1}.$$

两边约去 $|x_0|^n$, 整理得到:

$$|x_0| \leq 1 + \max_{0 \leq k \leq n-1} |a_k|,$$

如所欲证. □

Exercise 34 (再论中国剩余定理)

我们先给出中国剩余定理的最一般形式: 设 R 为含么环, $N_1, \dots, N_r \trianglelefteq R$ 为 R 的非平凡理想且两两互素 ($N_i + N_j = (1)$). 令 σ_i 表示自然同态 $\sigma_i : R \rightarrow R/N_i$ ($i = 1 \dots r$), 则映射

$$\begin{aligned} \sigma : R &\rightarrow R/N_1 \oplus \cdots \oplus R/N_r \\ x &\mapsto (\sigma_1(x), \dots, \sigma_r(x)) \end{aligned}$$

为满同态, 且 $\ker \sigma = N_1 \cap \cdots \cap N_r$.

特别地 $R/(N_1 \cap \cdots \cap N_r) = R/N_1 \oplus \cdots \oplus R/N_r$.

我们主要关心的情形为 $R = \mathbb{Z}$ 或者 $K[x]$, 此时可以重新叙述定理为:

令环 $R = \mathbb{Z}$ 或 $K[x]$, $a_1, \dots, a_r \in R$ 且两两互素. 再给定 $b_1, \dots, b_r \in R$. 则同余方程组

$$\begin{cases} x \equiv b_1 \pmod{a_1} \\ x \equiv b_2 \pmod{a_2} \\ \vdots \\ x \equiv b_r \pmod{a_r} \end{cases}$$

在 R 内恒有解, 且这个解在 $\text{mod } a_1 a_2 \cdots a_r$ 的意义下是唯一的: 即若 $x_1, x_2 \in R$ 都满足该方程组, 则 $x_1 - x_2 \equiv 0 \pmod{a_1 \cdots a_r}$.

证明该定理.

Solution 34

由 $a_1, \dots, a_r \in R$ 两两互素知: a_1 与 $a_2 \cdots a_r$ 互素. 这是因为 a_1 与 a_2, \dots, a_r 分别互素:

$$u_2 a_1 + v_2 a_2 = u_3 a_1 + v_3 a_3 = \cdots = u_r a_1 + v_r a_r = 1$$

全部乘起来得到 $\prod_{k=2}^r (u_k a_1 + v_k a_k) = 1$, 展开并合并全部含 a_1 的项得 $a_1 u + \prod_{k=2}^r (v_k a_k) =$

1. 因此 a_1 与 $a_2 \cdots a_r$ 互素: $x_1 a_1 + y_1 (a_2 \cdots a_r) = 1$. 同理有:

$$a_2 \text{ 与 } a_1 a_3 \cdots a_r \text{ 互素: } x_2 a_2 + y_2 a_1 a_3 \cdots a_r = 1$$

...

$$a_r \text{ 与 } a_1 \cdots a_{r-1} \text{ 互素: } x_r a_r + y_r a_1 \cdots a_{r-1} = 1$$

令 $x = \sum_{i=1}^r b_i y_i \prod_{j \neq i} a_j$, 就有 x 满足:

$$x \equiv b_i y_i \prod_{j \neq i} a_j \equiv b_i \pmod{a_i}.$$

因此同余方程组恒有解, 而若 x_1, x_2 都满足该方程, 则 $x_1 - x_2 \equiv 0 \pmod{a_i}$. 所以 $x_1 - x_2 \equiv 0 \pmod{a_1 \cdots a_r}$. 这就证明了唯一性.

Exercise 35 (Lagrange 插值)

给定 $a_1, \dots, a_r, b_1, \dots, b_r \in K$. 求 $f \in K[x]$ 满足: $f(a_1) = b_1, \dots, f(a_r) = b_r$.

Solution 35

此即

$$\begin{cases} f \equiv b_1 \pmod{(x-a_1)} \\ \vdots \\ f \equiv b_r \pmod{(x-a_r)} \end{cases}$$

由中国剩余定理 (练习34)立得.

Exercise 36 (*Hermite插值*)

在Lagrange插值35的要求上, 我们还额外要求前 s 个点处导数有特定值: $f'(a_1) = d_1, \dots, f'(a_s) = d_s$. 求 f .

Solution 36

此即

$$\begin{cases} f \equiv b_1 + d_1(x-a_1) \pmod{(x-a_1)^2} \\ \vdots \\ f \equiv b_s + d_s(x-a_s) \pmod{(x-a_s)^2} \\ f \equiv b_{s+1} \pmod{(x-a_{s+1})} \\ \vdots \\ f \equiv b_r \pmod{(x-a_r)} \end{cases}$$

仍由中国剩余定理 (练习34)立得.

注记: 事实上还可以推广到要求一点上函数值直到某高阶导数满足一定条件, 请读者自行思考此时应该如何写出 f 满足的方程 (提示: 考虑 f 在某点处Taylor展开和它的各阶导数关系).

Exercise 37 (*Another Freshman's Dream*)

设 A, B, C, D 为数域 F 上 n 阶方阵, 且 $AC = CA$, 求证:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|$$

Solution 37

让我们先看 A 可逆的情形:

$$\begin{aligned} \begin{vmatrix} A & B \\ C & D \end{vmatrix} &= \begin{vmatrix} I & O \\ -CA^{-1} & I \end{vmatrix} \begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} A & B \\ O & D - CA^{-1}B \end{vmatrix} \\ &= |A||D - CA^{-1}B| = |AD - ACA^{-1}B| = |AD - CAA^{-1}B| = |AD - CB| \end{aligned}$$

这给了我们充分的信心. 注意到 $\forall \lambda \in F: (A + \lambda I)C = C(A + \lambda I)$. 因此当 $(A + \lambda I)$ 可逆时同样有

$$\begin{vmatrix} A + \lambda I & B \\ C & D \end{vmatrix} = |(A + \lambda I)D - CB|$$

令 $f(\lambda) = \begin{vmatrix} A + \lambda I & B \\ C & D \end{vmatrix}$, $g(\lambda) = |(A + \lambda I)D - CB|$. 则 f 和 g 都是关于 λ 的多项式 (这是因为行列式的定义中只出现加法和乘法).

由上面的讨论知道, 当 $A + \lambda I$ 可逆时, $f(\lambda) = g(\lambda)$. 再观察到: $A + \lambda I$ 不可逆当且仅当 $|A + \lambda I| = 0$, 而 $|A + \lambda I|$ 又是一个关于 λ 的不恒为零的多项式, 因此至多只有有限个 λ 使得 $|A + \lambda I| = 0$ 成立. 所以有无穷个 $\lambda \in F$ 使得 $f(\lambda) = g(\lambda)$, 这迫使两个多项式相等: $f = g$. 特别地, $f(0) = g(0)$, 即 $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD - CB|$.

注记: 先在一个“稠密”的集合上证明某种性质, 再推广到全集, 这是一种常用的证明手法.

Exercise 38

令 $A, B \in \mathbb{R}^{n \times n}$, $C = A + iB$ (i 为虚数单位). 求证:

$$\det \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = |\det C|^2.$$

Solution 38

$$\begin{aligned}
|\det C|^2 &= \det C \cdot \overline{\det C} = \det C \cdot \det \overline{C} = \det \begin{pmatrix} C & \\ & \overline{C} \end{pmatrix} \\
&= \det \begin{pmatrix} A+iB & \\ & A-iB \end{pmatrix} = \det \begin{pmatrix} A+iB & iA-B \\ & A-iB \end{pmatrix} \\
&= \det \begin{pmatrix} A+iB & iA-B \\ -iA+B & 2A \end{pmatrix} = \det \begin{pmatrix} \frac{A}{2} + i\frac{B}{2} & -B \\ -iA+B & 2A \end{pmatrix} \\
&= \det \begin{pmatrix} \frac{A}{2} & -B \\ B & 2A \end{pmatrix} = \det \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.
\end{aligned}$$

Exercise 39

若两个实系数方阵在复数域上相似, 则它们在实数域上也相似.

Solution 39

设 $B = P^{-1}AP$, 则将 P 拆成 $R + iS$ 的形式, 其中 R 和 S 都是实系数方阵, 那么由 $B = P^{-1}AP$ 得到 $(R + iS)B = PB = AP = A(R + iS)$. 比较实部虚部知

$$AR = RB, AS = SB.$$

令 $Q = R + \lambda S$, 则 $f(\lambda) := \det Q$ 为一多项式, 由 $f(i) = \det P \neq 0$ 知 $f(\lambda) \neq 0$. 因此 f 只有有限个根, 我们可以选取 $\lambda_0 \in \mathbb{R}$ 使得 $R + \lambda_0 S$ 可逆. 由 $(R + \lambda_0 S)B = A(R + \lambda_0 S)$ 以及 $R + \lambda_0 S$ 可逆知, A 和 B 在实数域上也相似.

注记: 这里利用多项式的理论证明了特殊情况下的相似对域扩张的不变性. 事实上这一命题对一般的域扩张都广泛成立: 即若 K/k 是一域扩张, 则两 k 系数方阵在 k 上相似当且仅当它们在 K 上相似, 我们将会在相似标准形理论部分见到这一事实.

Exercise 40

证明: 若 $(x-1) \mid f(x^n)$, 则 $(x^n-1) \mid f(x^n)$.

Solution 40

$$\begin{aligned}
& (x-1) \mid f(x^n) \\
\implies & 0 = f(1^n) = f(1) \\
\implies & (y-1) \mid f(y) \\
\implies & f(y) = q(y)(y-1) \\
\implies & f(x^n) = q(x^n)(x^n-1) \\
\implies & (x^n-1) \mid f(x^n).
\end{aligned}$$

Exercise 41

证明 $\gcd(x^n - 1, x^m - 1) = x^{\gcd(n, m)} - 1$.

Solution 41

不妨设 $n > m$, 记 n 除以 m 的余数为 $n \% m$: 则 $\gcd(x^n - 1, x^m - 1) = \gcd(x^n - 1 - x^{n-m}(x^m - 1), x^m - 1) = \gcd(x^{n-m} - 1, x^m - 1) = \dots = \gcd(x^{n \% m} - 1, x^m - 1)$. 由于这就是更相减损术求两个整数最大公因数的过程, 因此 $\gcd(x^n - 1, x^m - 1) = x^{\gcd(n, m)} - 1$.

Exercise 42

若方阵 A 为幂零阵, $A^m = O$, 证明 $I + A$ 可逆.

Solution 42

由 $1/(1+x)$ 的 Taylor 展开:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots$$

而 $x^m = 0$ 时 $x^m, x^{m+1}, x^{m+2} \dots$ 都不计入求和.

于是:

$$(I - A + A^2 - A^3 + \dots + (-1)^{m-1} A^{m-1})(I + A) = I + (-1)^{m-1} A^m = I$$

另解: 由于 $\gcd(x^m, x+1) = 1$, 所以存在 $u, v \in F[x]$ 使得 $x^m u + (1+x)v = 1$, 带入 $x = A$ 就有

$$I = u(A)A^m + v(A)(I+A) = v(A)(I+A).$$

于是 $v(A)$ 就是所欲求的 $I+A$ 的逆.

Exercise 43

设 a_1, \dots, a_n 为两两不同的整数. 求证: $(x-a_1)(x-a_2)\cdots(x-a_n)-1$ 在 \mathbb{Q} 上不可约.

Solution 43

设 $f(x) = (x-a_1)(x-a_2)\cdots(x-a_n)-1$, 若存在 $p(x), q(x) \in \mathbb{Q}[x]$ 使得 $f(x) = p(x)q(x)$, 由 Gauss 引理, 不妨假设 $p(x)$ 和 $q(x)$ 都是整系数多项式, 且 $p(x), q(x)$ 均为首一多项式. 由于对任意 a_i 我们总有 $f(a_i) = -1$, $p(a_i), q(a_i) \in \mathbb{Z}$. 于是下列两种情况有且仅有一种为真:

(i) $p(a_i) = 1, q(a_i) = -1$;

(ii) $p(a_i) = -1, q(a_i) = 1$.

无论是哪种情况都有 $p(a_i) + q(a_i) = 0$. 所以 $p+q$ 在 a_1, \dots, a_n 处都为 0. 由于 p, q 都是首一多项式, $p+q$ 非零. $p+q$ 至少有 n 个根. 所以 $\deg(p+q) \geq n$, $\deg p \geq n$ 或 $\deg q \geq n$. 这样 $f = p \cdot q$ 就不能是 f 的一个非平凡分解. 故 \mathbb{Q} 上 f 不可约.

练习: 若 n 是奇数, a_1, \dots, a_n 为两两不同的整数. 求证: $(x-a_1)(x-a_2)\cdots(x-a_n)+1$ 在 \mathbb{Q} 上不可约.

更多的练习: a_1, \dots, a_n 为两两不同的整数. 求证: $(x-a_1)^2(x-a_2)^2\cdots(x-a_n)^2+1$ 在 \mathbb{Q} 上不可约.

Exercise 44

设 $f(x) = x^3 + (1+t)x^2 + 2x + 2u$, $g(x) = x^3 + tx + u$ 的最大公因式为二次多项式. 求 t, u 的值

Solution 44

首先注意到 $f(x) - g(x) = (1+t)x^2 + (2-t)x + u$ 也被 $\gcd(f, g)$ 整除. 而 $\deg \gcd(f, g) = 2$. 故 $t \neq -1$, 且 $f(x) - g(x)$ 正是 $\gcd(f, g)$. 于是选取 g 进一步计算:

$$x^3 + tx + u = [(1+t)x^2 + (2-t)x + u] \left(\frac{1}{1+t}x + c \right)$$

其中 c 依赖于 u : 若 $u \neq 0$, 则 $c = 1$, 否则还需进一步讨论.

Case 1, $u \neq 0$: $x^3 + tx + u = [(1+t)x^2 + (2-t)x + u] \left(\frac{1}{1+t}x + 1 \right)$, 展开得:

$$\begin{cases} \frac{2-t}{1+t} + t + 1 = 0 \\ \frac{u}{1+t} + 2 - t = t \end{cases} \implies \begin{cases} t = \frac{-1 \pm \sqrt{11}i}{2} \\ u = -7 \mp \sqrt{11}i \end{cases}$$

Case 2, $u = 0$: 则直接解得最大公因式的根为 $x = 0$, $\frac{t-2}{t+1}$. 于是 $(\frac{t-2}{t+1})^2 + t = 0$, 解得 $t_1 = -4$, $t_{2,3} = \frac{1 \pm \sqrt{3}i}{2}$.

注记: 与练习32中的定理3类似, 我们有子结式的概念可以用于机械地计算两个多项式最大公因式恰好为某个次数的条件.

Theorem 5 (子结式). 设 $F = a_0x^d + \cdots + a_d$, $G = b_0x^e + \cdots + b_e$ 分别为整环 D 上的 d 和 e 次多项式. P_n 表示 $K = \text{Frac } D$ 上小于 n 次的多项式集合. 定义映射

$$\varphi_j: \begin{matrix} P_{e-j} \oplus P_{d-j} & \rightarrow & P_{d+e-j} \\ (u, v) & \mapsto & uF + vG \end{matrix}.$$

则 φ_j 是一个线性映射. 其在多项式环的自然基下的矩阵为

$$\begin{pmatrix} a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdots & 0 \\ a_2 & a_1 & \ddots & 0 & b_2 & b_1 & \ddots & 0 \\ \vdots & \vdots & \ddots & a_0 & \vdots & \vdots & \ddots & b_0 \\ a_d & a_{d-1} & \cdots & \vdots & b_e & b_{e-1} & \cdots & \vdots \\ 0 & a_d & \ddots & \vdots & 0 & b_e & \ddots & \vdots \\ \vdots & \vdots & \ddots & a_{d-1} & \vdots & \vdots & \ddots & b_{e-1} \\ 0 & 0 & \cdots & a_d & 0 & 0 & \cdots & b_e \end{pmatrix}$$

这是一个 $(d + e - j) \times (d + e - 2j)$ 的矩阵, 定义 $\text{sRes}_j(F, G)$ 为截取其前 $d + e - 2j$ 行的行列式. 由定义直接看出 $\text{sRes}_j(F, G) = 0$ 当且仅当存在非零多项式 $u \in P_{e-j}, v \in P_{d-j}$ 使得 $\deg uF + vG < j$ (回顾线性方程组理论!)

通过主子结式理论, 我们给出两个多项式最大公因式次数的等价刻画:

$\deg \gcd(F, G) \geq j$ 当且仅当前 j 个子结式为零: $\text{sRes}_0(F, G) = \cdots = \text{sRes}_{j-1}(F, G) = 0$.

关于结式的定理 3 即本定理中 $j = 1$ 情形.

证明. 若 $\deg \gcd(F, G) \geq j$, 则

$$\deg \text{lcm}(F, G) = \deg \frac{FG}{\gcd(F, G)} \leq d + e - j.$$

因此存在非零多项式 $u \in P_{e-j+1}, v \in P_{d-j+1}$ 使得 $\deg \text{lcm}(F, G) = uF = -vG$. 所以对于任意 $0 \leq k \leq j - 1$ 有 $\text{sRes}_k(F, G) = 0$.

对相反的方向, 我们使用归纳法证明. 首先反向对 $j = 1$ 成立: $\text{sRes}_0(F, G) = 0$ 蕴含存在非零的 $u, v \in K[x]$, 次数分别小于 e 和 d , 使得 $uF + vG = 0$. 假设命题已经对 $j - 1$ 成立, 由归纳假设知 $\text{sRes}_0(F, G) = \cdots = \text{sRes}_{j-2}(F, G) = 0$ 蕴含 $\deg \gcd(F, G) \geq j - 1$, 又知道 $\text{sRes}_{j-1}(F, G) = 0$ 蕴含存在非零 $u \in P_{e-j+1}, v \in P_{d-j+1}$ 使得 $\deg uF + vG < j - 1$, 这迫使 $uF + vG = 0$. 于是 $uF = -vG$ 为 F 和 G 的公倍式, $\deg \text{lcm}(F, G) \leq d + e - j$. 最终我们得到 $\deg \gcd(F, G) \geq j$. \square

作为定理的直接推论:

Corollary 6. $\deg \gcd(F, G) = j$ 当且仅当前 j 个子结式为零而第 $j + 1$ 个子结式不为零: $\text{sRes}_0(F, G) = \cdots = \text{sRes}_{j-1}(F, G) = 0, \text{sRes}_j(F, G) \neq 0$.

本题中的 $\varphi_0, \varphi_1, \varphi_2$ 分别有矩阵表示:

$$\left(\begin{array}{cccccc} 1 & & & 1 & & \\ 1+t & 1 & & 0 & 1 & \\ 2 & 1+t & 1 & t & 0 & 1 \\ 2u & 2 & 1+t & u & t & 0 \\ & 2u & 2 & & u & t \\ & & 2u & & & u \end{array} \right), \left(\begin{array}{cccc} 1 & & 1 & \\ 1+t & 1 & 0 & 1 \\ 2 & 1+t & t & 0 \\ 2u & 2 & u & t \\ \hline & 2u & & u \end{array} \right), \left(\begin{array}{cc} 1 & 1 \\ 1+t & 0 \\ \hline 2 & t \\ 2u & u \end{array} \right).$$

$\underbrace{\hspace{10em}}_{e-j \text{ 列}}$
 $\underbrace{\hspace{10em}}_{d-j \text{ 列}}$

截取横线上方子矩阵并计算行列式得到:

$$\text{sRes}_0(f, g) = -2t^4u - t^3u^2 - 4t^3u + 2t^2u^2 + 12t^2u - 4tu^2 - 14tu - u^3 - 7u^2 + 8u,$$

$$\text{sRes}_1(f, g) = t^3 + 3t^2 - tu - 3t - u + 4,$$

$$\text{sRes}_2(f, g) = -t - 1.$$

解

$$\begin{cases} -2t^4u - t^3u^2 - 4t^3u + 2t^2u^2 + 12t^2u - 4tu^2 - 14tu - u^3 - 7u^2 + 8u & = 0 \\ t^3 + 3t^2 - tu - 3t - u + 4 & = 0 \\ t & \neq -1 \end{cases}$$

得

$$\begin{cases} t = \frac{-1 \pm \sqrt{11}i}{2} \\ u = -7 \mp \sqrt{11}i \end{cases} \text{ 或 } \begin{cases} t = -4 \\ u = 0 \end{cases} \text{ 或 } \begin{cases} t = \frac{1 \pm \sqrt{3}i}{2} \\ u = 0 \end{cases}$$

这依然是纯粹机械的计算, 无需动脑.

关于主子结式, 可以参考《符号计算选讲》(王东明等著)一书或是维基百科. 限于篇幅限制在此处不再展开.

Exercise 45

证明: 如果 $(x^2 + x + 1) \mid [f_1(x^3) + xf_2(x^3)]$, 那么 $(x-1) \mid f_1(x)$, $(x-1) \mid f_2(x)$.

Solution 45

显然 $x^2 + x + 1$ 的两根为 $\omega = \frac{-1+\sqrt{3}i}{2}$ 和 $\omega^2 = \bar{\omega} = \frac{-1-\sqrt{3}i}{2}$. 由 $(x^2 + x + 1) \mid [f_1(x^3) + xf_2(x^3)]$ 知 ω 和 ω^2 也是 $[f_1(x^3) + xf_2(x^3)]$ 的根, 即

$$f_1(\omega^3) + \omega f_2(\omega^3) = f_1(\omega^6) + \omega^2 f_2(\omega^6) = 0.$$

但是注意到 ω 是三次单位根 ($\omega^3 = 1$), 于是 $f_1(1) + \omega f_2(1) = f_1(1) + \omega^2 f_2(1) = 0$. 写成线性方程组的形式就是:

$$\begin{pmatrix} 1 & \omega \\ 1 & \omega^2 \end{pmatrix} \begin{pmatrix} f_1(1) \\ f_2(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

由系数矩阵是 Vandermonde 矩阵知行列式不为 0, 该方程组只有零解. 所以 $f_1(1) = f_2(1) = 0$. 即 $(x-1) \mid f_1(x)$, $(x-1) \mid f_2(x)$

Exercise 46

证明: 如果 $(x^{n-1} + \cdots + x + 1) \mid [f_1(x^n) + xf_2(x^n) + \cdots + x^{n-2}f_{n-1}(x^n)]$, 那么 $(x-1) \mid f_i(x)$, $i = 1..n-1$.

Solution 46

完全类似练习45. $x^{n-1} + \cdots + x + 1 = 0$ 的根为除1外的全体 n 次单位根:

$$\omega_n = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}, \omega_n^2, \dots, \omega_n^{n-1}.$$

由整除关系, 它们带入 $f_1(x^n) + xf_2(x^n) + \cdots + x^{n-2}f_{n-1}(x^n)$ 后为0. 这样就有系数矩阵为Vandermonde矩阵的线性方程组

$$\begin{pmatrix} 1 & \omega_n & \cdots & \omega_n^{n-2} \\ 1 & \omega_n^2 & \cdots & \omega_n^{2(n-2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \cdots & \omega_n^{(n-1)(n-2)} \end{pmatrix} \begin{pmatrix} f_1(1) \\ f_2(1) \\ \vdots \\ f_{n-1}(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

所以该方程只有平凡解 $f_1(1) = \cdots = f_{n-1}(1) = 0$. 也就是 $\forall i: (x-1) \mid f_i(x)$.

Exercise 47

已知三次方程 $x^3 + px^2 + qx + r = 0$, 求另一多项式方程使得其三根分别为前一方程三根的立方.

Solution 47

不妨设原多项式方程的三根为 a, b, c . 我们要求一个方程使其根为 a^3, b^3, c^3 . 由韦达定理, 新方程的系数为 $-(a^3 + b^3 + c^3)$, $a^3b^3 + a^3c^3 + b^3c^3$, $-a^3b^3c^3$. 现在的问题是: 如何将这些系数用 p, q, r 表示出来? 答案是利用对称多项式和基本对称多项式的关系!

先计算 $a^3 + b^3 + c^3 = (a + b + c)^3 - 3(a + b + c)(ab + ac + bc) + 3abc$, 这可以通过由化对称多项式为基本对称多项式的组合的算法得到, 或是直接利用牛顿恒等式.

再来计算 $a^3b^3 + a^3c^3 + b^3c^3$, 这里注意我们可以利用前一步的结果 (不要浪费人生的宝贵时间在多余的计算上)!

$$\begin{aligned} & a^3b^3 + a^3c^3 + b^3c^3 \\ &= (ab)^3 + (ac)^3 + (bc)^3 \\ &= (ab + ac + bc)^3 - 3(ab + ac + bc)(a^2bc + ab^2c + abc^2) + 3a^2b^2c^2 \\ &= (ab + ac + bc)^3 - 3(ab + ac + bc)(a + b + c)(abc) + 3(abc)^2 \end{aligned}$$

$$\text{最后 } a^3b^3c^3 = (abc)^3. \text{ 故 } \begin{cases} a^3 + b^3 + c^3 &= -p^3 + 3pq - 3r \\ a^3b^3 + a^3c^3 + b^3c^3 &= q^3 - 3pqr + 3r^2 \\ a^3b^3c^3 &= -r^3 \end{cases}.$$

所以所要求的方程是 $z^3 + (p^3 - 3pq + 3r)z^2 + (q^3 - 3pqr + 3r^2)z + r^3 = 0$

另解: 也可通过定理3直接计算得到结果, 计算 $x^3 + px^2 + qx + r$ 与 $x^3 - z$ 关于 x 的结式:

$$\text{Res}(x^3 + px^2 + qx + r, x^3 - z) = -r^3 + (-q^3 + 3pqr - 3r^2)z + (-p^3 + 3pq - 3r)z^2 - z^3$$

这与我们之前的计算结果是一样的. 这体现了结式的另一作用: 从一组多元多项式中消去一个变元.

Exercise 48

a_1, a_2, \dots, a_n 两两不同, 求证: 关于 x_1, \dots, x_n 的线性方程组

$$\begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -a_1^n \\ -a_2^n \\ \vdots \\ -a_n^n \end{pmatrix}$$

有唯一解, 并求出这组解来.

Solution 48

由Vandermonde矩阵性质立即知道该方程组有唯一解. 为求出解: 将等号右边的常数项挪到等号左边, 第 i 行变成:

$$x_1 + x_2 a_i + x_3 a_i^2 + \cdots + x_n a_i^{n-1} + a_i^n = 0$$

令 $f(z) = z^n + \sum_{k=1}^n x_k z^{k-1}$, 则 $\forall i: f(a_i) = 0$. 即 a_1, \dots, a_n 为 n 次多项式 $f(z)$ 的全部 n 个根. 于是

$$f(z) = (z - a_1) \cdots (z - a_n) = z^n + \sum_{k=1}^n x_k z^{k-1}$$

由韦达定理展开比较系数知:

$$\begin{cases} x_1 &= (-1)^n \sigma_n(a_1, \dots, a_n) \\ x_2 &= (-1)^{n-1} \sigma_{n-1}(a_1, \dots, a_n) \\ &\vdots \\ x_n &= -\sigma_1(a_1, \dots, a_n) \end{cases},$$

其中 $\sigma_1, \dots, \sigma_n$ 是基本对称多项式.

注记: 事实上也可通过Cramer法则暴力计算出每一个 x_i , 这涉及到计算缺项Vandermonde行列式, 可以通过加边完成计算, 有兴趣的读者可以自行尝试.

第四章 群论初步

Exercise 49 (有限阶元素)

怎样的2阶整数方阵 A 的有限次正整数次幂 A^n 等于单位阵?

Solution 49

设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, A 具有特征值 ω_1, ω_2 . 由 A 有限阶知: A 的极小多项式 $m(\lambda)$ 整除 $\lambda^n - 1$, $m(\lambda)$ 无重根, A 在复数域上可对角化, $|\omega_1| = |\omega_2| = 1$, $\det A = \pm 1$. 又设 $\varphi(\lambda) = (\lambda - \omega_1)(\lambda - \omega_2)$ 为 A 的特征多项式, 则由根与系数关系以及 $\det A = \pm 1$ 知 $\omega_1\omega_2 = \pm 1$, $\omega_1 + \omega_2$ 为一整数.

若 $\omega_1\omega_2 = 1$, 设 $\omega_1 = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$, 则有 $\omega_2 = \overline{\omega_1} = \cos \frac{2k\pi}{n} - i \sin \frac{2k\pi}{n}$. $\omega_1 + \omega_2 = 2 \cos \frac{2k\pi}{n}$ 为一整数, 且 $|\omega_1 + \omega_2| \leq 2$. 这有五种可能:

- (i) $2 \cos \frac{2k\pi}{n} = 2$: 则 $\omega_1 = \omega_2 = 1$, A 相似到单位阵 I , 因此 $A = I$. $a = d = 1$, $b = c = 0$;
- (ii) $2 \cos \frac{2k\pi}{n} = 1$: 则 $k = 1, n = 6$, $a + d = 1$, $ad - bc = 1$, 例如 $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$;
- (iii) $2 \cos \frac{2k\pi}{n} = 0$: 则 $k = 1, n = 4$, $a + d = 0$, $ad - bc = 1$, 例如 $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$;
- (iv) $2 \cos \frac{2k\pi}{n} = -1$: 则 $k = 1, n = 3$, $a + d = -1$, $ad - bc = 1$, 例如 $\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$;
- (v) $2 \cos \frac{2k\pi}{n} = -2$: 则 $\omega_1 = \omega_2 = -1$, A 相似到 $-I$, 因此 $A = -I$, $a = d = -1$, $a + d = 0$.

若 $\omega_1\omega_2 = -1$, 则 $\omega_2 = -\overline{\omega_1}$, 这迫使 $\omega_1 + \omega_2 = 0$. 于是 $\omega_1 = 1, \omega_2 = -1$. 这样就有 $a + d = 0, ad - bc = -1$, 例如 $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

而上面所给六种条件又是充分的, 因此我们总结 $GL_2(\mathbb{Z})$ 中有限阶元素如下:

- (i) 阶为一的元素: 单位阵 $I, a = d = 1, b = c = 0$;
- (ii) 阶为二的元素: 这有两种可能, 一种是负单位阵 $-I, a = d = -1, b = c = 0$, 另一种具有特征值 ± 1 , 满足 $a + d = 0, ad - bc = -1$;
- (iii) 阶为三的元素: $a + d = -1, ad - bc = 1$;
- (iv) 阶为四的元素: $a + d = 0, ad - bc = 1$;
- (v) 阶为六的元素: $a + d = 1, ad - bc = 1$.

Exercise 50 (置换群的表示)

求0, 1组成的整系数方阵 A, B 使得它们在整系数可逆方阵群中的阶都为2, 乘积 AB 的阶是3.

Solution 50

考虑置换群 S_n 到可逆矩阵群 $GL_n(\mathbb{Z})$ 的嵌入群同态:

$$\sigma: \begin{matrix} S_n & \rightarrow & GL_n(\mathbb{Z}) \\ (i \ j) & \mapsto & P_{ij} \end{matrix},$$

其中 P_{ij} 为交换单位矩阵第 i 行和第 j 行的第一类初等方阵, 则 σ 良定义.

$$\text{取 } n = 3, \text{ 则 } A = \sigma((1 \ 2)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \sigma((1 \ 3)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \text{ 因}$$

$$\text{此 } BA = \sigma((1 \ 2)(1 \ 3)) = \sigma((1 \ 3 \ 2)) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

这样就有 A, B 均为2阶元素, 但 BA 是3阶元素.

注记: 从有限群 G 到域 F 上的一般线性群 $GL_n(F)$ 的群同态称为 G 的群表示(将群 G 的元素表示成矩阵形式).

Exercise 51

令 $F = \mathbb{Z}/p\mathbb{Z}$ 为 p 阶有限域(p 为一素数). 求 F 上全体 n 阶可逆方阵的个数 $|GL_n(F)|$.

Solution 51

这是纯粹的线性代数问题. 让我们来考虑如何选取一个可逆方阵 $A \in GL_n(F)$: 首先, A 的第一列 a_1 可以是任意非零向量, 这有 $p^n - 1$ 种选取方式. 其次, A 的第二列 a_2 必须且只需选自 $F^n \setminus \langle a_1 \rangle$ (因为 a_2 与 a_1 线性无关当且仅当 a_2 不落在 a_1 生成的线性子空间中), 所以 a_2 有 $p^n - p$ 种选法. 之后每一列 a_k 的选取都必须且只需满足 $a_k \notin \langle a_1, \dots, a_{k-1} \rangle$, 于是 a_k 有 $p^n - p^{k-1}$ 种选法.

因此由乘法原理, $|GL_n(F)| = \prod_{i=1}^n (p^n - p^{i-1})$.

Exercise 52 (万变不离其宗)

求一个正整数, 使得 $\mathbb{Z}/5\mathbb{Z}$ 上的3阶上三角可逆方阵 A 都满足 $A^n = I$.

Solution 52

答案不唯一, 以下给出一个可能的答案. 由Lagrange定理, 群中每个元素的阶数都整除群的阶数. 而上三角可逆矩阵关于矩阵乘法的确构成一个群 $T_n(F)$. 因此只需要考虑计算 $\mathbb{Z}/5\mathbb{Z}$ 上的3阶上三角可逆方阵群的阶数即可. 而注意到 A 是上三角可逆方阵当且仅当 A 的对角线上元素均为 $\mathbb{Z}/5\mathbb{Z}$ 中可逆元素, 对于对角线上方元素没有任何限制. 因此 $|T_3(\mathbb{Z}/5\mathbb{Z})| = 4^3 \cdot 5^6 = 1000000$. 对任何 $A \in T_3(\mathbb{Z}/5\mathbb{Z})$: $A^{1000000} = I$.

注记: 也可利用练习51计算结果.

Exercise 53

令 n 为一正整数.

- (1) 从0到 $10n - 1$ 中等概率地随机产生四个整数 a, b, c, d , 求 $ad - bc$ 为奇数的概率;
- (2) 从0到 $10n - 1$ 中等概率地随机产生四个整数 a, b, c, d , 求 $ad - bc$ 模5余2的概率.

Solution 53

- (1) 注意到 $ad - bc$ 的奇偶性只与 a, b, c, d 的奇偶性有关, 而每个元素为奇数和偶数是等可能的, 在模2意义下, $ad - bc$ 为奇数等价于方阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 可逆. 因此我们立刻约化到 $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ 上二阶可逆矩阵群阶数计算上. 由练习51知 $|GL_2(\mathbb{F}_2)| = 3 \cdot 2 = 6$. 而 $|M_2(\mathbb{F}_2)| = 2^4 = 16$. 因此 $ad - bc$ 为奇数的概率是 $\frac{6}{16} = \frac{3}{8}$.
- (2) 同上, 我们约化到 $\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$ 的计算上, 注意到 $\det : GL_2(\mathbb{F}_5) \rightarrow (\mathbb{F}_5)^\times$ 是乘法群同态, 同态的核为 $\ker \det = SL_2(\mathbb{F}_5)$. 因此 $ad - bc$ 模5余2的集合是 $SL_2(\mathbb{F}_5)$ 的一个陪集, 且这样的陪集恰有4个 (事实上由同态基本定理: $GL_2(\mathbb{F}_5)/SL_2(\mathbb{F}_5) \cong (\mathbb{F}_5)^\times \cong \mathbb{Z}/4\mathbb{Z}$). 所以 $ad - bc$ 模5余2的集合大小为 $\frac{1}{4}|GL_2(\mathbb{F}_5)| = \frac{24 \cdot 20}{4} = 120$, 所求概率为 $\frac{120}{5^4} = \frac{24}{125}$.

Exercise 54

为庆祝某校数学科学学院成立110周年, 概率系发行了一种彩票. 这种彩票在 3×3 的方格上印有取值自 $\{0, 1, 2\}$ 的九个数. 这9个数满足排成的行列式一定不是3的倍数且在此条件下等概率分布. 购买一张彩票1元钱, 规定若对角线下方的三个数都是0则可以获得五十元大奖, 问购买彩票是否划算?

Solution 54

这事实上要计算有限域 \mathbb{F}_3 上一般线性群 $GL_3(\mathbb{F}_3)$ 和可逆上三角阵群 $T_3(\mathbb{F}_3)$ 的大小. 这是经典的计算. 对可逆上三角阵群, 对角线上为可逆元, 对角线上方可以任选: $|T_3(\mathbb{F}_3)| = (3 - 1)^3 \cdot 3^3$. 对一般线性群, 第一列可以任取非零向量, 第二列可以选第一列张成线性子空间外任意元素, 第三列可以选第一第二列张成线性子空间外任意元素, 所以 $|GL_3(\mathbb{F}_3)| = (3^3 - 1) \cdot (3^3 - 3) \cdot (3^3 - 3^2)$. 因此中奖概率为

$$\frac{2^3 \cdot 3^3}{26 \times 24 \times 18} = \frac{1}{52}.$$

计算期望知 $\frac{1}{52} \times 50 - 1 < 0$, 购买彩票不划算.

Exercise 55

证明 p^2 阶群皆交换 (p 为素数).

Solution 55

令 G 为一 p^2 阶群. 考虑 G 对自身的共轭作用: $G \times G \rightarrow G, g, x \mapsto g^{-1}xg$, 则共轭作用将 G 划分为若干轨道. 由轨道-稳定化子公式

$$|G| = |\text{Orb}(x)| \cdot |\text{Stab}_G(x)|$$

知每个共轭类的大小都整除群的阶数 p^2 , 因此每个共轭类的大小都是 $1, p, p^2$ 之一. 单位元 e 显然单独成一轨道, 因此 G 的类方程 (即将群阶数写为各轨道大小之和) 为

$$p^2 = 1 + \cdots$$

记群的中心为 $Z(G)$, 则 $x \in Z(G)$ 当且仅当 x 单独成一轨道, 当且仅当 $p \nmid |\text{Orb}(x)|$. 因此 $|Z(G)|$ 一定被 p 整除 (否则类方程右侧不是 p 的倍数), 又群的中心非空, 因此 G 的中心非平凡.

注意到 $Z(G) \trianglelefteq G$, 由 Lagrange 定理以及上面讨论知 $|Z(G)| = p$ 或 p^2 . 若 $|Z(G)| = p$, 则存在非中心元素 x , 考虑 x 的稳定化子 $\text{Stab}_G(x)$ (即与 x 交换的元素), 显然 x 本身和 $Z(G)$ 都落在稳定化子中, 于是 $p < |\text{Stab}_G(x)| \leq p^2$, Lagrange 定理迫使 $\text{Stab}_G(x) = G$, 即 x 与全群交换, 这与 x 是非中心元素矛盾. 所以群的中心只能是全群, 即 G 是交换群.

推论: 由主理想整环上有限生成模结构定理 (特别地, 有限阿贝尔群结构定理) 知这样的群在同构意义下只有两种, $\mathbb{Z}/p^2\mathbb{Z}$ 和 $\mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}$.

Exercise 56

设 G 为一群, $H \leq G$ 为一子群. 则 H 是正规子群当且仅当 H 可以写成若干共轭类轨道的 (无交) 并.

Solution 56

这几乎是翻译定义. 若 H 是正规子群, 则 $\forall h \in H, \forall g \in G: g^{-1}hg \in H$, 即 H 中所有元素的共轭类都完全落在 H 中, 于是 H 是若干共轭类的并. 反之若子群 H 是若干共轭类的并, 则显然有 $\forall g \in G: g^{-1}Hg = H$.

Exercise 57

若群 G 的类方程为 $20 = 1 + 4 + 5 + 5 + 5$.

- (1) 群 G 有没有5阶子群? 若有, 它是不是正规子群?
- (2) 群 G 有没有4阶子群? 若有, 它是不是正规子群?

Solution 57

- (1) 考虑大小为4的共轭类 C 中任意元素 x , 令 $H = \text{Stab}_G(x)$ 为全体和 x 交换元素构成的子群. 由轨道-稳定化子公式知: $|H| = |G|/|\text{Orb}(x)| = 20/4 = 5$, 因此 G 有5阶子群 H , 存在性得证. 而 $e, x \in H$ ($e \neq x$), 由Lagrange定理知 $\langle x \rangle = H$, 即 H 是 x 生成的循环群, x 的阶为5.

注意到 H 中除单位元外的元素阶都为5, 而大小为4的共轭类 C 中恰好有4个5阶元. 现在要说明 $H = \{e\} \cup C$, 这是因为若 $x^i \notin C$ ($i \in \{1, 2, 3, 4\}$), 那么 x^i 落在某个大小为5的共轭类中, 类似上面的讨论知道 $\langle x^i \rangle \leq C_G(x^i) = \text{Stab}_G(x^i)$, 但左边的阶数为5, 右边的阶数为4, 这不可能.

这样就有 H 是共轭类 $\{e\}$ 和 C 的并, 由练习56知 H 是正规子群.

- (2) 类似的, 考虑大小为5的共轭类中元素的稳定化子, 这是一个大小为4的子群. 但任何大小为4的子群都不可能是正规子群 (因为唯一大小之和是4的一组共轭类不包含单位元 e , 不能构成群, 见练习56).

注记: 存在性可由Sylow定理立得.

Exercise 58 (Burnside引理)

令 G 是一有限群, X 为一有限集合, G 在 X 上有一作用. 记 $|X/G|$ 为 X 的轨道个数, $\text{Fix}_X(g) := \{x \in X | gx = x\}$ 为 g 固定的集合元素. 证明以下等式:

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |\text{Fix}_X(g)|.$$

Solution 58

$$\begin{aligned} |X/G| &= \sum_{x \in X} \frac{1}{|\text{Orb}(x)|} \\ &= \sum_{x \in X} \frac{|\text{Stab}_G(x)|}{|G|} \\ &= \frac{1}{|G|} \sum_{x \in X} |\text{Stab}_G(x)| \\ &= \frac{1}{|G|} \sum_{x \in X} \sum_{g \in G, gx=x} 1 \\ &= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X, gx=x} 1 \\ &= \frac{1}{|G|} \sum_{g \in G} |\text{Fix}_X(g)|. \end{aligned}$$

证明的精髓在于改变求和指标.

Exercise 59 (立方体染色计数)

给定一个正立方体, 现在要用红色和黑色两种颜色给立方体的六个面染上颜色, 规定如果染色方案 A 经过立方体旋转后能得到染色方案 B 则将两种染色方案视作本质相同的染色 (例如 A : 给前面和后面染上红色, B : 给左边和右边染上红色).

问有多少种本质不同的染色?

Solution 59

我们将应用上面证明的Burnside引理 (练习58). 记 G 为立方体的旋转变换群, X 为全体染色方案, 则本质不同的染色数等于 G 在 X 上作用的轨道数.

容易证明立方体的旋转对称群大小为24 (想一想, 为什么?), 其中有恒同变换1个, 绕面中心旋转90度6个, 绕面中心旋转180度3个, 绕体对角线旋转120度8个, 固定相对两条棱的旋转6个. 考虑每一个旋转变换固定多少个染色方案:

- (i) 恒同变换固定全部 2^6 种染色方案;
- (ii) 绕面中心旋转90度固定 2^3 种染色方案;
- (iii) 绕面中心旋转180度固定 2^4 种染色方案;
- (iv) 绕体对角线旋转120度固定 2^2 种染色方案;
- (v) 固定相对两条棱的旋转固定 2^3 种染色方案.

因此应用Burnside引理得:

$$|X/G| = (2^6 + 2^3 \times 6 + 2^4 \times 3 + 2^2 \times 8 + 2^3 \times 6)/24 = 10.$$

注记: 可以证明立方体的旋转对称群同构于 S_4 .

练习: 七夕情人节小C想送给小X一串手串, 手串由八个珠子串成一圈, 每个珠子可以从黑色珠子和白色珠子中选择一个. 若两串手串经过旋转和翻转之后相同则视为同一种手串, 问一共有多少种手串?

Exercise 60

如果群 G 的子群 H 对于所有自同构 $\varphi: G \rightarrow G$ 都满足 $\varphi(H) = H$, 则称之为特征子群.

- (i) 证明特征子群总是正规子群, 特征子群的特征子群仍是特征子群.
- (ii) 说明群的中心 Z_G 是特征子群, 而且群 G 的导出子群

$$G' := \langle aba^{-1}b^{-1} | a, b \in G \rangle$$

也是特征子群.

- (iii) 证明 $(\mathrm{GL}_n(F))' = \mathrm{SL}_n(F)$, 其中 F 为一个至少包含3个元素的域.

Solution 60

- (i) 注意到对任意群元素 g , 我们有共轭作用引起的内自同构 $\varphi_g : G \rightarrow G, x \mapsto g^{-1}xg$, 因此 $\forall g \in G : \varphi_g(H) = H$ 蕴含 H 正规.

而任何 G 的自同构限制到特征子群 H 上仍是子群 H 的自同构. 因此若 K 是 H 的特征子群, 那么 G 的自同构仍然将 K 映到 K , 所以 K 是 G 的特征子群.

- (ii) 考虑自同构 φ , 要说明 $\forall z \in Z_G, \varphi(z) \in Z_G$, 则只要说明 $\forall g \in G : \varphi(z)g = g\varphi(z)$. 由 φ 是自同构, 于是只要说明 $\varphi^{-1}(\varphi(z)g) = \varphi^{-1}(g\varphi(z))$, 即 $z\varphi^{-1}(g) = \varphi^{-1}(g)z$. 这由 $z \in Z_G$ 是显然的.

而导出子群当然是特征子群, 因为生成元 $aba^{-1}b^{-1}$ 在 φ 下的像仍是另一个生成元 $\varphi(a)\varphi(b)\varphi(a)^{-1}\varphi(b)^{-1}$.

- (iii) 首先注意到 $(\mathrm{GL}_n(F))'$ 的生成元行列式都为1, 因此 $(\mathrm{GL}_n(F))' \subseteq \mathrm{SL}_n(F)$.

反过来, 记 $E_{ij}(\lambda)$ ($i \neq j$)为对角线上全为1, 第 i 行第 j 列为 λ 的第一类初等方阵. 那么

Lemma 1. 设 F 为任意域, $\{E_{ij}(\lambda) | 1 \leq i, j \leq n, i \neq j, \lambda \in F\}$ 生成整个 $\mathrm{SL}_n(F)$.

证明. 当 $n = 2$ 时, 若 $a \neq 0$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c+1}{a} & 1 \end{pmatrix} \begin{pmatrix} 1 & 1-a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1+\frac{b-1}{a} \\ 0 & 1 \end{pmatrix}.$$

这是因为 $ad - bc = 1$, 否则 $b \neq 0$:

用 $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ 和 $\begin{pmatrix} b & -a \\ d & -c \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ 化归到前一情形.

当 $n \geq 3$ 时, 若 $A \in \mathrm{SL}_n(F)$ 第一列除了 a_{11} 外有任意的非零元素 $a_{i1} = a$, 则左乘 $E_{1i}(\frac{1-a_{11}}{a})$ 可以将 a_{11} 变成1, 于是再通过第一类初等行变换和第一类初等列变换可以将 A 消成 $\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{B} \end{pmatrix}$ 的形式, 这样就化归到 $n-1$ 的情形; 若 A 第一列除了 a_{11} 全为0, 那么将 A 的第一行加到第二行就又化成了第一列除 a_{11} 外有非零元素的情形. \square

通过上面的引理我们知道, 只要说明 $E_{ij}(\lambda)$ 可以写成 $aba^{-1}b^{-1}$ 的形式就可以完成反方向的证明.

先看 $n = 2$ 而 $|F| \geq 3$ 时: 选择 $\mu \in F \setminus \{0, 1\}$:

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \mu & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\lambda}{\mu-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \frac{\lambda}{\mu-1} \\ 0 & 1 \end{pmatrix}^{-1}$$

给出 $E_{12}(\lambda)$, 而转置给出 $E_{21}(\lambda)$.

再看 $n \geq 3$ 时: $E_{ij}(\lambda) = E_{il}(\lambda)E_{lj}(1)E_{il}(-\lambda)E_{lj}(-1)$, 其中 l 可以任取不同于 i, j 的1到 n 中的整数.

所以无论哪种情形, 我们都由引理有 $SL_n(F) \subseteq (GL_n(F))'$. 因此最终有 $SL_n(F) = (GL_n(F))'$.

Exercise 61

群的交换化定义为 $G_{ab} = G/G'$.

- (i) 说明 G_{ab} 交换, 而且对于任何交换群 A 和同态 $f: G \rightarrow A$, 存在唯一的同态 $\bar{f}: G_{ab} \rightarrow A$ 使得 f 通过 G_{ab} (即: f 分解为 $G \xrightarrow{f'} G_{ab} \xrightarrow{\bar{f}} A$).
- (ii) 确定 $S_3, Q_8, D_{2n}, GL_n(F)$ 的交换化, 其中 F 为任意域.
- (iii) 对于群 $G = GL_n(F)$, 其中 F 为含至少三个元素的域, 试将商同态 $G \rightarrow G_{ab}$ 等同于行列式.

Solution 61

- (i) G_{ab} 显然交换, 这是因为所有的“非交换部分” $aba^{-1}b^{-1}$ 都被模掉了, 而 f 通过 G_{ab} 是因为对于任意 $a, b \in G$, $f(a)f(b)f(a)^{-1}f(b)^{-1} = 1$ (A 交换). 于是所有 $aba^{-1}b^{-1}$ 都落在 $\ker f$ 中, 这导致 $G' \subseteq \ker f$. 因此 f 诱导 $\bar{f}: G_{ab} \rightarrow A$. 显然 \bar{f} 是唯一的.
- (ii) 直接计算有 $(S_3)_{ab} = \mathbb{Z}/2\mathbb{Z}$, $(Q_8)_{ab} = (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z})$. 而

$$(D_{2n})_{ab} = \begin{cases} \mathbb{Z}/2\mathbb{Z} & 2 \nmid n \\ (\mathbb{Z}/2\mathbb{Z}) \oplus (\mathbb{Z}/2\mathbb{Z}) & 2 \mid n \end{cases},$$

这是因为 $\sigma\tau\sigma\tau^{-1} = \tau^{-2}$ (σ 为反射, τ 为旋转).

$$\mathrm{GL}_n(F) = \begin{cases} \mathbb{Z}/2\mathbb{Z} & F = \mathbb{Z}/2\mathbb{Z}, n = 2 \\ F^\times & \text{otherwise} \end{cases},$$

这是因为 $(\mathrm{GL}_2(\mathbb{Z}/2\mathbb{Z})) \cong S_3$, 对剩下的情况用练习60, 以及 $\mathrm{GL}_n(F)/\mathrm{SL}_n(F) \cong F^\times$.

(iii) 由 $\mathrm{SL}_n(F) = (\mathrm{GL}_n(F))'$ 以及 $\det : \mathrm{GL}_n(F) \rightarrow F^\times$ 给出 $\mathrm{GL}_n(F)/\mathrm{SL}_n(F) \cong F^\times$ 立得.

Exercise 62 (*Iwasawa* 判据)

设群 G 作用在集合 X 上, $|X| \geq 2$, 若对所有 $(x, y), (x', y') \in X^2$, 其中 $x \neq y$ 且 $x' \neq y'$, 皆存在 $g \in G$ 使得 $gx = x'$ 而 $gy = y'$, 则称作用为双传递 (doubly transitive) 的. 以下设 G 双传递地作用于 X , 并且记 $H_x = \mathrm{Stab}_G(x)$.

- (i) 证明 H_x 对于所有 $x \in X$ 都是 G 的极大真子群: 换言之, 包含 H_x 的子群只有 G 和 H_x 本身.
- (ii) 说明任何正规子群 $N \triangleleft G$ 在 X 上的作用或者平凡, 或者传递.
- (iii) (*Iwasawa*) 假设 G 作用忠实, $G = G'$, 而且存在 x 使得 H_x 有正规交换子群 U , 而 U 在 G 中的所有共轭生成 G . 证明 G 为单群.

Solution 62

(i) 对于 $g \notin H_x$, 我们证明 $G = H_x \cup H_x g H_x$.

若另一个 $g' \notin H_x$, 我们来说明 $g' \in H_x g H_x$. 依 g, g' 的选取方式, $gx, g'x \neq x$. 因此由双传递性, 存在 $g'' \in G$ 将 (x, gx) 映到 $(x, g'x)$. 由 $g''x = x$ 知 $g'' \in H_x$. 而 $g''(gx) = g'x$, 所以 $g' \in g''gH_x \subseteq H_x g H_x$. 于是确实有 $G = H_x \cup H_x g H_x$.

而 H_x 显然是真子群 (G 传递, H_x 固定 x). 若还有子群 K 包含 H_x , 则选取 $g \in K \setminus H_x$. 由前面所说, $G = H_x \cup H_x g H_x$, 但 $H_x, H_x g H_x \subseteq K$, 因此 $G \subseteq K$, 即 $G = K$, 因此 H 极大.

(ii) 假设 N 作用非平凡: $\exists n \neq 1, x \in X$ s.t. $nx \neq x$. 对任意 $y, y' \in X$ ($y \neq y'$), 由双传递性, 存在 $g \in G$ 将 (x, nx) 映到 (y, y') , 这样 $y' = (gng^{-1})y$ 且 $gng^{-1} \in N$. 于是 N 的作用是传递的.

(iii) 若 $N \trianglelefteq G$ 是 G 的一个正规子群, 令 $H = \text{Stab}_G(x)$. 那么 $NH = \{nh | n \in N, h \in H\}$ 为一个包含 H 的 G 的子群 (N 正规). 由前面所证知道 $NH = H$ 或 $NH = G$, 且 N 的作用或者平凡或者传递.

如果 $NH = H$, 那么 $N \subseteq H$, 于是 N 固定 x 不动, N 的作用只能是平凡的. 但 G 的作用忠实, 这迫使 $N = \{1\}$.

而如果 $NH = G$, 令 U 为 H 的正规交换子群 (注意: 一个交换子群未必是正规的!), 且 U 在 G 中的共轭生成 G . 那么由 $U \trianglelefteq H$ 知道 $NU \trianglelefteq NH = G$. 因此对于任意 $g \in G$: $gUg^{-1} \subseteq g(NU)g^{-1} = NU$, 于是 NU 包含所有 U 在 G 中的共轭, 这样就有 $NU = G$.

所以 $G/N = (NU)/N \cong U/(N \cap U)$ (“Diamond Theorem”). 而 U 是交换的, 这一同构告诉我们 G/N 是交换群, 由交换化部分知识知 $G' \subseteq N$. 但 $G = G'$, 因此 $N = G$ 是平凡的. 总之 G 没有非平凡正规子群.

Exercise 63 ($\text{PSL}_2(F)$ 的单性, $|F| \geq 4$)

设 F 为域, 记 Z 为 $\text{SL}_n(F)$ 的中心, 按照以下步骤证明 $n = 2$ 而 $|F| \geq 4$ 时 $\text{PSL}_2(F) := \text{SL}_2(F)/Z$ 为单群.

- (i) 说明 $|F| \geq 4$ 时 $\text{SL}_2(F) = (\text{SL}_2(F))'$.
- (ii) 让 $\text{PSL}_2(F)$ 以显然方式作用在 $\mathbb{P}^1(F)$ 上. 记 $(x, y) \in F^2 \setminus \{0\}$ 生成的子空间为 $(x : y)$. 说明这是双传递作用, 然后写下 $(1 : 0)$ 的稳定化子群 H .
- (iii) 代入 Iwasawa 判据 (练习 62), 推导 $|F| \geq 4$ 时 $\text{PSL}_2(F)$ 是单群.

Solution 63

(i) 当 $|F| \geq 4$ 时, 存在 $a \in F \setminus \{0, 1, -1\}$, 因此 $a^2 \neq 1$. 而

$$\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix}^{-1} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & b(a^2 - 1) \\ 0 & 1 \end{pmatrix},$$

所以令 b 跑遍整个 F 就能给出 $E_{12}(\lambda)$ ($\lambda \in F$). 同理 $E_{21}(\lambda)$ 也能写成换位子 $aba^{-1}b^{-1}$ 的形式. 由引理1知 $\mathrm{SL}_2(F) = (\mathrm{SL}_2(F))'$.

- (ii) 为说明 $\mathrm{PSL}_2(F)$ 的作用是双传递作用, 只要说明任何 $(v, w) \in \mathbb{P}^1 \times \mathbb{P}^1$ ($v \neq w$)都可以被映射到 $((1:0), (0:1))$ 即可. 令 $v = (a:c)$, $w = (b:d)$, 则 $D := ad - bc \neq 0$. 令 $A = \begin{pmatrix} a & b/D \\ c & d/D \end{pmatrix}$, 则 $A \in \mathrm{SL}_2(F)$. 取 \bar{A} 为 A 在 $\mathrm{PSL}_2(F)$ 中的像, 则 $\bar{A}v = (a:c)$, $\bar{A}w = (b/D:d/D) = (b:d)$. $\mathrm{PSL}_2(F)$ 的作用是双传递的. 而 $(1:0)$ 的稳定化子群直接计算可知为

$$H := \left\{ \overline{M} \mid M = \begin{pmatrix} \lambda & \mu \\ 0 & 1/\lambda \end{pmatrix}, \lambda \neq 0 \right\}.$$

- (iii) 显然 $\mathrm{PSL}_2(F)$ 作用忠实, 而直接计算可以知道 $U = \left\{ \overline{M} \mid M = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} \right\} \cong F$ 是 H 的正规交换子群, 并且 U 的共轭生成整个 $\mathrm{PSL}_2(F)$:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix},$$

(回忆引理1). 结合第一部分 (蕴含 $\mathrm{PSL}_2(F) = (\mathrm{PSL}_2(F))'$)以及Iwasawa判据 (练习62): $|F| \geq 4$ 时 $\mathrm{PSL}_2(F)$ 是单群.

注记: $\mathrm{PSL}_2(\mathbb{Z}/2\mathbb{Z}) \cong S_3$, $\mathrm{PSL}_2(\mathbb{Z}/3\mathbb{Z}) \cong A_4$ 都不是单群.

Exercise 64 ($\mathrm{PSL}_n(F)$ 的单性, $n \geq 3$)

承上题, 设 F 为任意域. 按以下步骤证明 $n \geq 3$ 时 $\mathrm{PSL}_n(F) := \mathrm{SL}_n(F)/Z$ 是单群.

- (i) 说明 $\mathrm{SL}_n(F) = (\mathrm{SL}_n(F))'$.
- (ii) 让 $\mathrm{PSL}_n(F)$ 以显然方式作用在 $\mathbb{P}^{n-1}(F)$ 上. 说明这是双传递作用, 然后写下 $(1:0:\cdots:0)$ 的稳定化子群 H .
- (iii) 代入Iwasawa判据 (练习62), 推导 $\mathrm{PSL}_n(F)$ 单.

Solution 64

- (i) 注意到 $E_{ij}(\lambda) = E_{il}(\lambda)E_{lj}(1)E_{il}(\lambda)^{-1}E_{lj}(1)^{-1}$ 是换位子, 而引理1说明 $E_{ij}(\lambda)$ 生成 $\mathrm{SL}_n(F)$. 于是 $\mathrm{SL}_n(F) = (\mathrm{SL}_n(F))'$.
- (ii) 类似练习63中的做法可以构造线性映射将 (v, w) 映射到 (e_1, e_2) , 适当调整该映射可以使得行列式为1, 所以 $\mathrm{PSL}_n(F)$ 在 \mathbb{P}^{n-1} 上是双传递作用. 而稳定化子群 H 中元素形如 $\begin{pmatrix} a & * \\ \mathbf{0} & \mathbf{M} \end{pmatrix}$, 其中 $a \in F^\times$, $M \in \mathrm{GL}_{n-1}(F)$, $\det M = 1/a$.
- (iii) 上面的 H 有自然的到 $\mathrm{PGL}_{n-1}(F)$ 的映射: $H \rightarrow \mathrm{PGL}_{n-1}(F), \begin{pmatrix} a & * \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \mapsto M$, 映射的核为 $U = \left\{ \begin{pmatrix} 1 & * \\ \mathbf{0} & \mathbf{I}_{n-1} \end{pmatrix} \right\} \cong F^{n-1}$. 这是一个 H 的正规交换子群. 由第一部分知 $\mathrm{PSL}_n(F) = (\mathrm{PSL}_n(F))'$ 且显然 $\mathrm{PSL}_n(F)$ 作用忠实. U 包含所有 $E_{il}(\lambda)$, 经共轭作用后可以得到所有 $E_{li}(\lambda)$. 所以 U 的共轭生成整个 $\mathrm{PSL}_n(F)$. 由Iwasawa判据知 $\mathrm{PSL}_n(F)$ 单.

第五章 相似标准型理论

Exercise 65

$$\text{令 } N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{F}^{n \times n}. \text{ 求 } N^2, N^3, \dots, N^n.$$

Solution 65

令自然基为 e_1, e_2, \dots, e_n . 则显然: $Ne_n = e_{n-1}, Ne_{n-1} = e_{n-2}, \dots, Ne_2 = e_1, Ne_1 = 0$. 而 $N = (0, e_1, e_2, \dots, e_{n-1})$. 因此

$$\begin{aligned} N^2 &= N \cdot N = N(0, e_1, e_2, \dots, e_{n-1}) = (0, 0, e_1, \dots, e_{n-2}) \\ N^3 &= N \cdot N^2 = N(0, 0, e_1, \dots, e_{n-2}) = (0, 0, 0, \dots, e_{n-3}) \\ &\vdots \\ N^{n-1} &= N \cdot N^{n-2} = N(0, 0, \dots, e_1, e_2) = (0, 0, \dots, 0, e_1) \\ N^n &= N \cdot N^{n-1} = N(0, 0, \dots, 0, e_1) = (0, 0, \dots, 0, 0) = O \end{aligned}$$

注记: 将 N^k 矩阵具体写出来就知道, 每乘一个 N , 主对角线上方的一排1就向右上角移动一位.

Exercise 66

设 $A \in \mathbb{C}^{n \times n}$, A 的最小多项式为 x^n . 求 A^k 的 Jordan 标准形.

Solution 66

由 A 的最小多项式为 x^n 立知

$$A \sim N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

以下只需考虑 N , 自然基在 N 的作用下有如下箭头图:

$$0 \xleftarrow{N} e_1 \xleftarrow{N} e_2 \xleftarrow{N} \cdots \xleftarrow{N} e_n$$

于是自然基在 N^k 的作用下的箭头图为:

$$\begin{aligned} 0 &\xleftarrow{N^k} e_1 \xleftarrow{N^k} e_{k+1} \xleftarrow{N^k} \cdots \\ 0 &\xleftarrow{N^k} e_2 \xleftarrow{N^k} e_{k+2} \xleftarrow{N^k} \cdots \\ &\vdots \\ 0 &\xleftarrow{N^k} e_k \xleftarrow{N^k} e_{2k} \xleftarrow{N^k} \cdots \end{aligned}$$

一直延长上面的箭头图, 延长到哪里会结束呢?

令 $n = q \cdot k + r$ 为 n 除以 k 的带余除法. 则:

$$\begin{array}{cccccccc} 0 & \xleftarrow{N^k} & e_1 & \xleftarrow{N^k} & e_{k+1} & \xleftarrow{N^k} & \cdots & \xleftarrow{N^k} & e_{(q-1)k+1} & \xleftarrow{N^k} & e_{qk+1} \\ 0 & \xleftarrow{N^k} & e_2 & \xleftarrow{N^k} & e_{k+2} & \xleftarrow{N^k} & \cdots & \xleftarrow{N^k} & e_{(q-1)k+2} & \xleftarrow{N^k} & e_{qk+2} \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ 0 & \xleftarrow{N^k} & e_r & \xleftarrow{N^k} & e_{k+r} & \xleftarrow{N^k} & \cdots & \xleftarrow{N^k} & e_{(q-1)k+r} & \xleftarrow{N^k} & e_n \\ 0 & \xleftarrow{N^k} & e_{r+1} & \xleftarrow{N^k} & e_{k+r+1} & \xleftarrow{N^k} & \cdots & \xleftarrow{N^k} & e_{(q-1)k+r+1} & & \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \\ 0 & \xleftarrow{N^k} & e_k & \xleftarrow{N^k} & e_{2k} & \xleftarrow{N^k} & \cdots & \xleftarrow{N^k} & e_{qk} & & \end{array}$$

每一行箭头图都是一个循环子空间, 所有行对应循环子空间的直和是 \mathbb{C}^n . 因此每个循环子空间对应一个 N^k 的Jordan块. 每一行箭头图的长度就是这一循环子空间的维数, 也就是这一个Jordan块的大小, 因此 N^k 的Jordan标准形有 r 个大小为 $q+1$ 的0-Jordan块, $k-r$ 个大小为 q 的0-Jordan块.

Exercise 67

如何计算Jordan标准形?

Solution 67

课上已经学过先计算Smith标准形, 再由不变因子计算初等因子和Jordan块的方法, 还有别的方法吗?

设 n 阶方阵 $A \sim \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$, 不妨设前 r 个Jordan块 J_1, \dots, J_r 对角线上为 λ , J_{r+1}, \dots, J_s 对角线上不为 λ . 则 $A - \lambda I \sim \begin{pmatrix} J_1 - \lambda I & & \\ & \ddots & \\ & & J_s - \lambda I \end{pmatrix}$. 由相似的矩阵秩相同以及分块对角阵秩等于各块秩之和知: $n - \text{rank}(A - \lambda I) = r$, 于是从 $\text{rank}(A - \lambda I)$ 可以计算出从属于 λ 的Jordan块个数 (即至少为一阶的 λ -Jordan块数量).

进一步地, $(A - \lambda I)^2 \sim \begin{pmatrix} (J_1 - \lambda I)^2 & & \\ & \ddots & \\ & & (J_s - \lambda I)^2 \end{pmatrix}$, 再由练习65知道平方后所有大于等于2阶的 λ -Jordan块秩减少1, 而一阶的Jordan块秩仍是0. 因此 $\text{rank}(A - \lambda I) - \text{rank}(A - \lambda I)^2$ 是至少2阶的 λ -Jordan块数量.

一般地, $\text{rank}(A - \lambda I)^{t-1} - \text{rank}(A - \lambda I)^t$ 是至少 t 阶的 λ -Jordan块数量. 于是将至少 t 阶的Jordan块数量与至少 $t + 1$ 阶的Jordan块数量作差就得到恰好 t 阶的Jordan块数量.

这样我们就得到了计算 A 的Jordan标准形的算法:

Step 1. 计算 A 的特征多项式 $\det(\lambda I - A)$;

Step 2. 解方程 $\det(\lambda I - A) = 0$ 得到全体特征值 $\lambda_1, \dots, \lambda_m$;

Step 3. 对每个特征值 λ_i ,

(3a). 计算秩: $r_0 = n, r_1 = \text{rank}(A - \lambda_i I), r_2 = \text{rank}(A - \lambda_i I)^2, \dots$;

(3b). 计算至少 k 阶的Jordan块数量: $d_1 = r_0 - r_1, d_2 = r_1 - r_2, \dots$;

(3c). 计算恰好 k 阶的Jordan块数量: $c_1 = d_1 - d_2, c_2 = r_2 - r_3, \dots$.

例如, 令

$$A = \begin{pmatrix} 1 & 1 & 0 & -1 & -1 & 1 \\ 0 & 2 & 0 & -1 & -1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 2 & -1 \\ 0 & -2 & 0 & 1 & 2 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & & & & & \\ & 1 & 1 & & & \\ & & 1 & & & \\ & & & 1 & 1 & \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix}$$

计算

$$A - I = \begin{pmatrix} 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 & -1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & -1 \\ 0 & -2 & 0 & 3 & 2 & -2 \end{pmatrix}, \quad \text{rank}(A - I) = 3$$

计算

$$(A - I)^2 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \end{pmatrix}, \quad \text{rank}(A - I)^2 = 1$$

而 $(A - I)^3 = O$.

于是 $(r_0, r_1, r_2, r_3) = (6, 3, 1, 0)$, $(d_1, d_2, d_3) = (3, 2, 1)$, $(c_1, c_2, c_3) = (1, 1, 1)$. 分别对应一块1, 2, 3阶Jordan块.

Exercise 68

$F \subseteq K$ 为两个域, $A, B \in F^{n \times n}$, 求证:

A, B 在 F 上相似 $\Leftrightarrow A, B$ 在 K 上相似.

Solution 68

\Rightarrow : 显然的.

\Leftarrow : A, B 在 K 上相似, 于是 $\lambda I - A$ 和 $\lambda I - B$ 在 K 上相抵, 又注意到我们在计算 Smith 标准形时每一步计算都不会离开原来的域, 即: $F[\lambda]$ 系数矩阵 $\lambda I - A$ 在每一步等价变化后得到的仍是 $F[\lambda]$ 系数矩阵. 所以最终得到的 Smith 标准形也是 $F[\lambda]$ 系数矩阵: $\lambda I - A$ 的不变因子都是 $F[\lambda]$ 中多项式. 同理 $\lambda I - B$ 的不变因子都是 $F[\lambda]$ 中多项式. 又由两者在 K 上相抵得到它们的秩和不变因子相等, 于是 $\lambda I - A$ 和 $\lambda I - B$ 在 F 上也相抵, A 和 B 在 F 上相似.

特别地有: 若 F 为一数域, 则:

A, B 在 F 上相似当且仅当它们在 \mathbb{C} 上相似.

Exercise 69

设 F 为一数域, 证明 A 与 A^T 相似.

Solution 69

令 $S = \begin{pmatrix} & & 1 \\ & \ddots & \\ 1 & & \end{pmatrix}$, 则 $S^{-1} = S$, 且

$$S^{-1} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} S = S \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} S = \begin{pmatrix} a_{nn} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{11} \end{pmatrix}.$$

令 A 在 \mathbb{C} 上的 Jordan 标准形为 $\text{diag}(J_1, \dots, J_r)$, 其中每个 J_k 为一 Jordan 块. 则容易发现 A^T 相似于 $\text{diag}(J_1^T, \dots, J_r^T)$. 因此我们只要证明对于每个 Jordan 块 J_k 而言有 $J_k \sim J_k^T$ 就有 A 与 A^T 在 \mathbb{C} 上相似, 再由练习 68 知 A 和 A^T 在 F 上相似.

而

$$J_k = \begin{pmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{pmatrix}, \quad J_k^T = \begin{pmatrix} \lambda_k & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_k \end{pmatrix}.$$

于是直接计算立即有 $S^{-1} J_k S = J_k^T$, 即 $J_i \sim J_i^T$.

注记: 通过Jordan标准形将问题约化到Jordan块的情形是一种应该掌握的常用技巧. 事实上本题也可以通过计算两者的行列式因子直接比较得到结论, 我们这么做是为了展示更多的思路.

Exercise 70

令 $\mathbb{C}[x]_n$ 为全体不超过 n 次的复系数多项式组成的集合.

- (1) 证明 $\mathbb{C}[x]_n$ 是一个 \mathbb{C} -线性空间.
- (2) 记 $D: \mathbb{C}[x]_n \rightarrow \mathbb{C}[x]_n$ 为求导算子.
 $f \mapsto f'$

具体写出 D 在单项式基 $1, x, \dots, x^n$ 下的矩阵 M , 并求 M 的Jordan标准形. 更进一步地, 求矩阵 P 将 M 过渡到Jordan标准形.

- (3) 问 $\mathbb{C}[x]_n$ 在 D 下的所有不变子空间是什么.

Solution 70

- (1) 这是显然的.
- (2) 容易直接写出矩阵

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

由于 $\text{rank } M = n = (n+1) - 1$, 且0就是所有特征值, 因此由练习67中给出的算法知 M 的Jordan标准形中只有一个Jordan块, 即 $M \sim \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$.

进一步地, 令 e_1, \dots, e_{n+1} 为自然基, 则 $e_{n+1} = (0, \dots, 0, 1)^T$ 经过 M 反复作用后:

$$e_{n+1} \xrightarrow{M} ne_n \xrightarrow{M} n(n-1)e_{n-1} \xrightarrow{M} \dots \xrightarrow{M} n!e_1,$$

于是 $P = (n!e_1, n!/1e_2, \dots, e_{n+1})$ 将 M 过渡到 Jordan 标准形:

$$\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} = \begin{pmatrix} n! & & & \\ & n!/1! & & \\ & & n!/2! & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 1 & & \\ & 0 & 2 & \\ & & \ddots & \ddots \\ & & & \ddots & n \\ & & & & 0 \end{pmatrix} \begin{pmatrix} n! & & & \\ & n!/1! & & \\ & & n!/2! & \\ & & & \ddots \\ & & & & 1 \end{pmatrix}.$$

- (3) 设 W 为一个 $\mathbb{C}[x]_n$ 的 D -不变子空间, $f \in W$ 为 W 中次数最高的多项式, $\deg f = p$, 则 $f, f', f'', \dots, f^{(p)} \in W$, 由于 $f^{(p)}$ 是常数, 它可以消去其它多项式的所有常数项, 类似地, $f^{(p-1)}$ 和 $f^{(p)}$ 的线性组合可以消去其它多项式的所有一次项和常数项. 这样 $f, f', f'', \dots, f^{(p)} \in W$ 就与 $x^p, x^{p-1}, \dots, 1$ 张成相同的线性空间. 因此 $W = \mathbb{C} \oplus \mathbb{C}x \oplus \dots \oplus \mathbb{C}x^p = \mathbb{C}[x]_p$.

所以全部的 D -不变子空间为 $0, \mathbb{C}, \mathbb{C}[x]_1, \dots, \mathbb{C}[x]_n$.

Exercise 71

设 F 是一无限域, V 是一有限维 F -线性空间, T 为 V 上一线性变换. 证明 T 只有有限个不变子空间当且仅当 T 的极小多项式等于特征多项式.

Solution 71

T 的极小多项式等于特征多项式等价于 V 作为 $F[x]$ -模循环, 同构于 $F[x]/(f)$. 则 $F[x]/(f)$ 显然只有有限个子模, 每个子模由 f 的一个因子生成. 反之若不变子空间有限, 那么考虑 $V = \bigcup_{v \in V} F[x]v$, 即所有 V 中向量 v 生成的循环模, 由于不变子空间有限, 这个并实际上是有限并. 但是我们知道无限域上有限个真子空间不能覆盖全空间 (回忆练习1), 因此必须有某个 v 生成的循环模是全空间, 即 V 是循环模.

注记: 与练习70比较.

Exercise 72

证明:

$$A = \begin{pmatrix} & & -a_0 \\ 1 & & -a_1 \\ & \ddots & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}$$

的极小多项式为 $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$.

Solution 72

令自然基为 e_1, \dots, e_n , 则 $Ae_1 = e_2, Ae_2 = e_3, \dots, Ae_{n-1} = e_n, Ae_n = \sum_{k=0}^{n-1} -a_k e_{k+1}$. 因此对于任意次数小于 n 的多项式 $g(x) = \sum_{k=0}^m b_k x^k$:

$$g(A)e_1 = \sum_{k=0}^m b_k A^k e_1 = \sum_{k=0}^m b_k e_{k+1} \neq 0$$

任何次数小于 n 的多项式都不能零化 A . 又由 Cayley-Hamilton 定理知 A 的特征多项式 φ_A 零化 A , 因此 A 的极小多项式 m_A 次数恰好为 n ($m_A \mid \varphi_A \implies \deg m_A \leq n$). 而 $f(A)e_1 = 0$, $f(x)$ 是零化 e_1 的次数最低的多项式, 所以 $f \mid m_A$, 但 $\deg f = \deg m_A = n$ 又迫使 $m_A = f$, 命题得证.

Exercise 73

求递推数列 $a_n = 3a_{n-2} + 2a_{n-3}$ 的通项公式.

Solution 73

首先我们发现递推公式可以写成矩阵乘法的形式:

$$\begin{pmatrix} a_n \\ a_{n-1} \\ a_{n-2} \end{pmatrix} = \begin{pmatrix} 3a_{n-2} + 2a_{n-3} \\ a_{n-1} \\ a_{n-2} \end{pmatrix} = \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ a_{n-3} \end{pmatrix} = \cdots = \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^{n-3} \begin{pmatrix} a_3 \\ a_2 \\ a_1 \end{pmatrix}$$

于是问题就转化为: 如何计算一个矩阵的高次幂? 此时我们可以借助 Jordan 标准形, 令 $A = \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, 计算其 Jordan 标准形得到 $J = \begin{pmatrix} 2 & & \\ & -1 & 1 \\ & & -1 \end{pmatrix}$, 且 $P = \begin{pmatrix} 4 & 1 & 1 \\ 2 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ 满

足 $J = P^{-1}AP$. 这样 $A^k = (PJP^{-1})^k = PJ^kP^{-1}$, 问题化归到计算Jordan块的 k 次幂上. 由练习65知道0-Jordan块是幂零的, 因此我们可以采用二项式展开计算Jordan块的幂次:

$$\begin{aligned} \begin{pmatrix} -1 & 1 \\ & -1 \end{pmatrix}^k &= \left(\begin{pmatrix} -1 & \\ & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} \right)^k \\ &= \begin{pmatrix} -1 & \\ & -1 \end{pmatrix}^{-k} + k \begin{pmatrix} -1 & \\ & -1 \end{pmatrix}^{k-1} \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} \\ &= \begin{pmatrix} (-1)^k & k(-1)^{k-1} \\ & (-1)^k \end{pmatrix} \end{aligned}$$

所以

$$\begin{aligned} A^{n-3} &= P \begin{pmatrix} 2^{n-3} & & \\ & (-1)^{n-3} & (-1)^{n-2}(n-3) \\ & & (-1)^{n-3} \end{pmatrix} P^{-1} \\ &= \frac{1}{9} \begin{pmatrix} 2^{n-1} + (-1)^{n-3}(3n-4) & 2^n + (-1)^{n-3}(-3n+1) & 2^{n-1} + (-1)^{n-3}(-6n+14) \\ * & * & * \\ * & * & * \end{pmatrix} \end{aligned}$$

这就是: $a_n = \frac{1}{9}((2^{n-1} + (-1)^{n-3}(3n-4))a_3 + (2^n + (-1)^{n-3}(-3n+1))a_2 + (2^{n-1} + (-1)^{n-3}(-6n+14))a_1)$.

Exercise 74

设 $A \in \mathbb{C}^{n \times n}$ 满足最小多项式 $m_A(\lambda)$ 等于特征多项式 $\varphi_A(\lambda)$. 求证与 A 交换的每个方阵 B 都可以写成 A 的一个多项式: $f(A) = B$.

Solution 74

由于最小多项式 m_A 等于最大的不变因子, 所有的不变因子乘积为特征多项式 φ_A . 因此 A 的特征多项式等于最小多项式说明 A 只有一个不变因子 m_A , 它的有理标准型只有一块. 设 $m_A(\lambda) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$, 则 A 相似于:

$$S = \begin{pmatrix} & & -a_0 \\ 1 & & -a_1 \\ & \ddots & \vdots \\ & & 1 & a_{n-1} \end{pmatrix} = P^{-1}AP$$

令 $v = Pe_1$, 其中 $e_1 = (1, 0, \dots, 0)^T$, 则 $v, Av, A^2v, \dots, A^{n-1}v$ 构成 \mathbb{C} 一组基 (想一想, 为什么?)

考虑 v 在 B 下的像在这组基下的坐标: $Bv = \sum_{k=0}^{n-1} c_k A^k v$, 则令多项式 $f(\lambda) = \sum_{k=0}^{n-1} c_k \lambda^k$. 那么 $Bv = f(A)v$. 于是对于任意 $w \in \mathbb{C}^n$: $w = \sum_{k=0}^{n-1} d_k A^k v$, 记 $g(\lambda) = \sum_{k=0}^{n-1} d_k \lambda^k$, 则 $w = g(A)v$. 于是

$$Bw = Bg(A)v = g(A)Bv = g(A)f(A)v = f(A)g(A)v = f(A)w$$

对任意 w 成立, 这迫使 $B = f(A)$.

注记: 若 $m_A \neq \varphi_A$, 则还有其它不为 A 的多项式的矩阵与 A 交换, 见定理7.

Exercise 75 (矩阵指数与矩阵对数)

如何对复数域上方阵 A 定义 e^A 和 $\ln A$?

Solution 75

令

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k$$

可以证明上式对任意复系数方阵都收敛, $\exp: \begin{matrix} \mathbb{C}^{n \times n} & \rightarrow & \mathbb{C}^{n \times n} \\ A & \mapsto & e^A \end{matrix}$ 良定义.

现在的问题是如何具体的计算出 e^A ?

首先注意到一个事实: $P^{-1}e^AP = e^{P^{-1}AP}$. 这是因为

$$P^{-1}e^AP = P^{-1} \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \right) P = \sum_{k=0}^{\infty} \frac{1}{k!} (P^{-1}AP)^k = e^{P^{-1}AP}.$$

于是 Jordan 标准形再一次发挥作用: 一切计算都可以化归到 Jordan 标准形的矩阵指数计算上, 由从矩阵指数的定义式中可以看出, 对分块对角阵计算矩阵指数只需要分别对每块计算矩阵指数即可, 因此问题再一次简化到对 Jordan 块 $\lambda I + N$ 的矩阵指数计算上:

设 $\lambda I + N$ 为 $m \times m$ 矩阵:

$$\begin{aligned}
 e^{\lambda I + N} &= \sum_{k=0}^{\infty} \frac{1}{k!} (\lambda I + N)^k \\
 &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} \lambda^{k-j} N^j \\
 &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{\min\{k, m-1\}} \binom{k}{j} \lambda^{k-j} N^j \\
 &= \sum_{j=0}^{m-1} N^j \sum_{k=j}^{+\infty} \frac{1}{k!} \binom{k}{j} \lambda^{k-j} \\
 &= \sum_{j=0}^{m-1} \frac{1}{j!} N^j \sum_{k=j}^{+\infty} \frac{1}{(k-j)!} \lambda^{k-j} \\
 &= \sum_{j=0}^{m-1} \frac{e^{\lambda}}{j!} N^j
 \end{aligned}$$

矩阵指数的一个应用是给出常系数线性常微分方程组的解: $y'(x) = ay(x)$ 的解为 $y(x) = e^{ax}y(0)$. 相应地, 向量值函数 $Y(x) : \mathbb{R} \rightarrow \mathbb{R}^n$ 若满足 $Y'(x) = AY(x)$, 其中 A 为一 $n \times n$ 矩阵, 则 $Y(x) = e^{Ax}Y(0)$

接下来再看矩阵对数: 给定 $A \in \mathbb{C}^{n \times n}$, 若存在 $L \in \mathbb{C}^{n \times n}$ 使得 $e^L = A$, 则称 L 为 A 的矩阵对数.

矩阵对数未必存在, 如 O 显然就没有矩阵对数. 那么什么样的矩阵有对数呢?

仍然由 $P^{-1}e^AP = e^{P^{-1}AP}$ 知: 只要 A 的 Jordan 标准形 J 有矩阵对数 $e^L = J$, 那么 A 也有矩阵对数 $A = P^{-1}JP = P^{-1}e^LP = e^{P^{-1}LP}$.

再一次问题转化为了给 Jordan 块求矩阵对数, 由 $\ln(1+x) = \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k} x^k$ 类比: 若 $\lambda \neq 0$, 则

$$\begin{aligned}
 \ln(\lambda I + N) &= \ln(\lambda(I + \lambda^{-1}N)) \\
 &= (\ln \lambda)I + \ln(I + \lambda^{-1}N) \\
 &= (\ln \lambda)I + \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k} (\lambda^{-1}N)^k \\
 &= (\ln \lambda)I + \sum_{k=1}^{m-1} \frac{(-1)^{k+1} N^k}{k \lambda^k}
 \end{aligned}$$

直接验证知 $L = (\ln \lambda)I + \sum_{k=1}^{m-1} \frac{(-1)^{k+1} N^k}{k \lambda^k}$ 满足 $e^L = \lambda I + N$.

若 $\lambda = 0$, 则 $N = 0I + N$ 不存在矩阵对数. 原因: 矩阵指数必为可逆阵: $e^A \cdot e^{-A} = e^{A-A} = e^O = I$.

于是我们得到 Jordan 块 $\lambda I + N$ 有矩阵指数当且仅当 $\lambda \neq 0$.

进一步地: 方阵 A 有矩阵对数当且仅当 A 可逆.

练习: 尝试写出矩阵三角函数的表达式: $\sin A$, $\cos A$.

更多的练习: 若矩阵 A 可逆, 尝试写出矩阵平方根的表达式 \sqrt{A} .

Exercise 76 (*Jordan-Chevalley*分解)

设 $A \in \mathbb{C}^{n \times n}$.

- (1) 证明 A 可以写成 $D + N$ 的形式, 其中 D 可对角化, N 幂零;
- (2) 若 A 可逆, 则 A 可以写成 BC 的形式, 其中 B 可对角化, C 的特征值全为 1.

Solution 76

- (1) 由 A 的 Jordan 标准形可以写成主对角线和次对角线之和立得: $J = P^{-1}AP = D_0 + N_0$, 于是 $A = PD_0P^{-1} + PN_0P^{-1}$ 满足条件.
- (2) 由练习 75: 存在复系数方阵 L 使得 $A = e^L$, 再由 (1) 知 $L = D + N$, 又可以验证 $DN = ND$, 于是 $A = e^{D+N} = e^D \cdot e^N$. 注意到 $P^{-1}e^DP = e^{P^{-1}DP}$, 所以 $B := e^D$ 可对角化. 而 N 是幂零阵, 由矩阵指数算法可知 $C := e^N$ 特征值全为 1. 所以 $A = BC = e^De^N$ 为所求分解.

注记: 事实上我们还有 $DN = ND$, $BC = CB$, 并且可以证明这样的分解是唯一的.

若 $A = D + N$, D 可对角化, N 幂零, $DN = ND$. 我们来说明 A 完全确定 D 和 N . 设 D 的相异特征值分别为 $\lambda_1, \dots, \lambda_s$, 那么 $V = \mathbb{C}^n$ 分解为 D 的特征子空间直和 $\bigoplus_{i=1}^s W_i$, 其中 $W_i = \ker(D - \lambda_i I)$. 对任意 $x \in W_i$ 有 $(D - \lambda_i I)Nx = N(D - \lambda_i I)x = 0$, 所以 W_i 也是 N 的不变子空间. 考虑分别将 $A - \lambda_i I$ 和 $A - D = N$ 限制到 W_i 上, 因为在 W_i 上 D 的作用等同于 λ_i , 所以 $(A - \lambda_i I)|_{W_i} = N|_{W_i}$, $(A - \lambda_i I)|_{W_i}$ 幂零. 所以 A 关于 λ_i 的广义特征子空间 V_{λ_i} 包含 W_i . 所以有 $\bigoplus_i W_i \subset \bigoplus_i V_{\lambda_i}$. 比较两边维数, 左边等于 n , 右边小于等于 n , 所以必须有 $W_i = V_{\lambda_i}$. 这样就证明了 D 的特征子空间一定是 A 的广义特征子空间, 被 A 唯一确定. 并且 D 在 A 关于 λ_i 的广义特征子空间上的作用等同于 $\lambda_i I$, 这样进一步唯一确定了 D , 于是 $N = A - D$ 也被唯一确定.

类似的讨论可以说明 $BC = CB$.

更多的注记: 事实上基域是 \mathbb{C} 的假设不是必须的, 甚至基域可以不是代数闭域. 事实上只要 K 满足 \bar{K}/K 是可分扩张 (例如 K 是完美域的情形, 更特殊地, K 特征零或者是有限域), 那么 K 上的方阵总能分解成两个 K 系数方阵的和, 其

中一个是半单的 (semi-simple, 在代数封闭的语境下等同于可对角化), 另一个是幂零的. 这是因为 \overline{K}/K 是正规可分扩张, 即 Galois 扩张. 在 \overline{K} 上进行 Jordan-Chevalley 分解后, 发现对于所有 \overline{K} 的 K -自同构 $\alpha \in \text{Gal}(\overline{K}/K)$ 都有 $A = \alpha(A) = \alpha(D + N) = \alpha(D) + \alpha(N)$, 且 $\alpha(D)$ 可对角化, $\alpha(N)$ 幂零, $\alpha(D)\alpha(N) = \alpha(DN) = \alpha(ND) = \alpha(N)\alpha(D)$. 但由于上面所说的分解唯一性只能有 $N = \alpha(N)$, $D = \alpha(D)$. 再由 Galois 群的固定域 $\overline{K}^{\text{Gal}(\overline{K}/K)} = K$ 知道被所有 K -自同构固定的 \overline{K} 元素只能是 K 中元素, 故实际上 D, N 都是 K -系数方阵.

练习: 证明 D 和 N 以及 B 和 C 实际上是 A 的多项式 (提示: 使用中国剩余定理构造 A 的多项式使得其作用在各个广义特征子空间上作用等同于对应特征值的标量变换).

Exercise 77

K 为一数域, 映射 $F: K^{n \times n} \rightarrow K^{n \times n} (n \geq 2)$ 定义为 $F(A) = -A^T$.

- (1) 求 F 的极小多项式;
- (2) 求 F 的所有特征值以及其对应的特征子空间;
- (3) 若 $\text{tr } F = -3$, 求 F 的 Jordan 标准形.

Solution 77

- (1) 一眼看出 $m_F(\lambda) = \lambda^2 - 1$! 且 F 可对角化.
- (2) 给我翻译翻译, 什么叫特征向量: $-A^T = F(A) = \lambda A$, 显然 λ 只能为 ± 1 .
 - (a) 当 $\lambda = 1$ 时: $A = -A^T$, A 为反对称矩阵, 即全体反对称矩阵构成 $\lambda = 1$ 的特征子空间, 维数为 $\frac{n(n-1)}{2}$;
 - (b) 当 $\lambda = -1$ 时: $A = A^T$, A 为对称矩阵, 即全体对称矩阵构成 $\lambda = -1$ 的特征子空间, 维数为 $\frac{n(n+1)}{2}$.
- (3) $\text{tr } F$ 为全体特征值记重数之和:

$$\text{tr } F = \frac{n(n-1)}{2} - \frac{n(n+1)}{2} = -3$$

因此 $n = 3$. $F \sim \text{diag}(1, 1, 1, -1, -1, -1, -1, -1, -1)$

Exercise 78

设 $T, U \in \mathbb{C}^{n \times n}$. $T \cdot U$ 可对角化.

- (1) 若 T 或 U 可逆, 求证 $(U \cdot T)$ 也可对角化.
- (2) 一般地, 即使 T 和 U 都不可逆, 证明仍有 $(UT)^2$ 可对角化.

Solution 78

1. 这是显然的, 不妨假设 T 可逆, 则 $UT \sim T(UT)T^{-1} = TU$. 而 TU 由题设可对角化, 因此 UT 也可对角化.
2. 我们分三步证明.

第一步先证明无限域上 AB 和 BA 具有相同的特征值: 由 $\det(I - AB) = \det(I - BA)$ 知

$$\forall \lambda \neq 0: \det(\lambda I - AB) = \lambda^n \det(I - \lambda^{-1}AB) = \lambda^n \det(I - \lambda^{-1}BA) = \det(\lambda I - BA)$$

而 \mathbb{C} 为无限域, 这样就必须有特征多项式相等: $\varphi_{AB} = \varphi_{BA}$. 特别地有 $(TU)^2 = (TUT)U$ 和 $(UT)^2 = U(TUT)$ 具有相同特征值和相同代数重数.

第二步说明 AB 和 BA 关于非零特征值的几何重数相同. 设 $\lambda \neq 0$ 为 AB 的特征值, AB 的关于 λ 的特征子空间 V_λ 的一组基为 x_1, \dots, x_r , 则 Bx_1, \dots, Bx_r 为 BA 的一组线性无关的关于 λ 的特征向量: 若 $\sum_{k=1}^r c_k Bx_k = 0$, 则左乘 A 得: $0 = \sum_{k=1}^r c_k ABx_k = \lambda \sum_{k=1}^r c_k x_k$, 于是 $\sum_{k=1}^r c_k x_k = 0$, 这迫使 $c_1 = \dots = c_r = 0$, Bx_1, \dots, Bx_r 的确是线性无关. 同时 $(BA)(Bx_k) = B(ABx_k) = B(\lambda x_k) = \lambda(Bx_k)$, 这说明 Bx_k 的确是 BA 关于 λ 的特征向量. 记 $\text{geo.mult.}_\lambda(M)$ 表示 M 关于 λ 的几何重数. 则我们已经知道

$$\text{geo.mult.}_\lambda(AB) \leq \text{geo.mult.}_\lambda(BA).$$

但是由于 A 和 B 的地位完全相同. 对称地我们可以得到反方向不等式, 这就是 $\text{geo.mult.}_\lambda(AB) = \text{geo.mult.}_\lambda(BA)$. 特别地 $(TU)^2$ 和 $(UT)^2$ 在非零特征值上具有相同几何重数.

最后再来说明 $(TU)^2$ 和 $(UT)^2$ 在 $\lambda = 0$ 时仍有几何重数相同. 由秩不等式:

$$\text{rank}(TU) \geq \text{rank}(UTUT) \geq \text{rank}(TUTUTU)$$

由 TU 可对角化知, $\text{rank}(TU) = \text{rank}(TU)^2 = \text{rank}(TU)^3$. 于是 $\text{rank}(UT)^2 = \text{rank}(TU)^2$, 由维数公式知 $\dim \ker(UT)^2 = \dim \ker(TU)^2$. 这就是两者的几何重数相等.

由前三步以及 (TU) 可对角化知:

$$\begin{array}{ccc} \text{geo.mult.}_\lambda((UT)^2) & & \text{alg.mult.}_\lambda((UT)^2) \\ \parallel_{(2)(3)} & & \parallel_{(1)} \\ \text{geo.mult.}_\lambda((TU)^2) & \xlongequal{\text{可对角化}} & \text{alg.mult.}_\lambda((TU)^2) \end{array}$$

其中 $\text{alg.mult.}_\lambda(M)$ 表示 M 关于 λ 的代数重数. 由上述等式我们最终知道 $(UT)^2$ 的几何重数和代数重数相同, 也就是 $(UT)^2$ 可对角化.

Exercise 79

若 $A \in \mathbb{C}^{n \times n}$ 有 A^2 可对角化, 证明 A^3 也可对角化.

Solution 79

由 A^2 可对角化知 A^2 的最小多项式 $m_{A^2}(\lambda)$ 无重根. 不妨设

$$m_{A^2}(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_s).$$

其中 $\lambda_1, \dots, \lambda_s$ 两两不同. 则 $f(\lambda) = m_{A^2}(\lambda^2) = (\lambda^2 - \lambda_1) \cdots (\lambda^2 - \lambda_s) = (\lambda - \sqrt{\lambda_1})(\lambda + \sqrt{\lambda_1}) \cdots (\lambda - \sqrt{\lambda_s})(\lambda + \sqrt{\lambda_s})$ 零化 A . $m_A(\lambda) | f(\lambda)$ 至多含有一个二次因子 λ^2 , 其余因子都是一次因子. 因此 A 的Jordan块中除了可能的 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 之外都是一阶的, 于是 A^3 的Jordan块都是一阶的, A^3 可对角化.

Exercise 80

设 $A, B \in \mathbb{C}^{n \times n}$, A 的特征多项式为 φ_A , 求证:

$\varphi_A(B)$ 可逆 $\Leftrightarrow A$ 和 B 没有公共特征值.

Solution 80

\Leftarrow : 由 A, B 没有公共特征值以及代数基本定理可知, $\gcd(\varphi_A, \varphi_B) = 1$ (φ_B 为 B 的特征多项式). 于是存在 $u, v \in \mathbb{C}[x]$ 使得 $u\varphi_A + v\varphi_B = 1$, 因此由 Cayley-Hamilton 定理: $u(B)\varphi_A(B) = u(B)\varphi_A(B) + v(B)\varphi_B(B) = I$. 于是 $\varphi_A(B)$ 可逆.

\Rightarrow : $\varphi_A(B)$ 可逆, 则 $0 \notin \text{Spec}(\varphi_A(B))$. 但是由谱映射定理知: $\text{Spec}(\varphi_A(B)) = \varphi_A(\text{Spec}(B))$.

$$\begin{array}{ccc} B & \xrightarrow{\varphi_A} & \varphi_A(B) \\ \downarrow \text{Spec} & & \downarrow \text{Spec} \\ \text{Spec}(B) & \xrightarrow{\varphi_A} & \varphi_A(\text{Spec}(B)) = \text{Spec}(\varphi_A(B)) \end{array}$$

因此对于任意 $\mu \in \text{Spec}(B)$: $0 \neq \varphi_A(\mu) \in \varphi_A(\text{Spec}(B))$.

于是 $\gcd(\varphi_A, \varphi_B) = 1$, $\text{Spec}(A) \cap \text{Spec}(B) = \emptyset$.

Exercise 81

对给定域 F 上 n 阶方阵 $A \in F^{n \times n}$, 定义 $L: F^{n \times n} \rightarrow F^{n \times n}$. 若 A 可对角化, 求证 L 可对角化.

Solution 81

将 A 对角化: $A = PDP^{-1}$, 其中 $D = \text{diag}(d_1, \dots, d_n)$ 为 1 阶对角阵. 由 $\{E_{ij}\}_{i,j=1}^n$ 为 $F^{n \times n}$ 的一组基知 $\{PE_{ij}P^{-1}\}_{i,j=1}^n$ 也是 $F^{n \times n}$ 的一组基 ($X \mapsto PXP^{-1}$ 为可逆线性变换).

注意到:

$$\begin{aligned} L(PE_{ij}P^{-1}) &= (PDP^{-1}) \cdot (PE_{ij}P^{-1}) - (PE_{ij}P^{-1}) \cdot (PDP^{-1}) \\ &= PDE_{ij}P^{-1} - PE_{ij}DP^{-1} \\ &= d_iPE_{ij}P^{-1} - d_jPE_{ij}P^{-1} \\ &= (d_i - d_j)PE_{ij}P^{-1}. \end{aligned}$$

于是所有 $\{PE_{ij}P^{-1}\}_{i,j=1}^n$ 都是 L 的特征向量, 构成 $F^{n \times n}$ 的一组基. 因此 L 可对角化.

练习: 若 A 是实数域上的一正定对称阵, 求证方程 $AX + XA = B$ 对任意实系数方阵 B 有唯一解.

Exercise 82

设 A 是一个 n 级复矩阵. $S : X \mapsto AX - XA$ 是 $\mathbb{C}^{n \times n}$ 上的线性变换. 证明: $\text{rank } S \leq n^2 - n$.

Solution 82

注意到 $S(X) = 0 \Leftrightarrow AX = XA$, 于是只要说明 $\dim C(A) \geq n$, 其中 $C(A)$ 为全体与 A 交换的复方阵即可.

设 A 的Jordan标准形为 $J = \text{diag}(J_{m_1}(\lambda_1), \dots, J_{m_s}(\lambda_s))$ 且 $P^{-1}AP = J$. 其中 $J_{m_i}(\lambda_i)$ 表示特征值为 λ_i , 大小为 m_i 的Jordan块. 显然我们有 $\sum_{i=1}^s m_i = n$

令 $X_{i,j} = P \text{diag}(0, 0, \dots, 0, [J_{m_i}(\lambda_i)]^j, 0, \dots, 0)P^{-1}$, 特别地 $j = 0$ 时 $X_{i,0} = P \text{diag}(0, 0, \dots, 0, I_{m_i}, 0, \dots, 0)P^{-1}$. 则

$$\begin{aligned} S(X_{i,j}) &= AX_{i,j} - X_{i,j}A \\ &= PJX_{i,j}P^{-1} - PX_{i,j}JP^{-1} \\ &= P \left(\begin{pmatrix} J_{m_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{m_i}(\lambda_i) & \\ & & & \ddots \\ & & & & J_{m_s}(\lambda_s) \end{pmatrix} \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & [J_{m_i}(\lambda_i)]^j & \\ & & & \ddots \\ & & & & 0 \end{pmatrix} \right. \\ &\quad \left. - \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & [J_{m_i}(\lambda_i)]^j & \\ & & & \ddots \\ & & & & 0 \end{pmatrix} \begin{pmatrix} J_{m_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{m_i}(\lambda_i) & \\ & & & \ddots \\ & & & & J_{m_s}(\lambda_s) \end{pmatrix} \right) P^{-1} \\ &= 0 \end{aligned}$$

因此 $X_{i,j} \in C(A)$. 又显然可以注意到 $X_{i,j}$ ($i = 1, \dots, s, j = 0, \dots, m_i - 1$)是线性无关的. 因此 $\dim C(A) \geq n, \text{rank } S \leq n^2 - n$.

另解: 事实上我们有著名的Cecioni-Frobenius定理:

Theorem 7 (Cecioni-Frobenius).

$\dim C(A) \geq n$, 且等号取到当且仅当 A 的最小多项式 $m_A(\lambda)$ 等于特征多项式 $\varphi_A(\lambda)$. 事实上我们可以计算 $\dim C(A)$: 令 A 的不变因子为 d_1, \dots, d_s ($d_i | d_{i+1}$).

则

$$\dim C(A) = \sum_{i=1}^s (2s - 2i + 1) \deg d_i$$

证明. 由不变因子分解 (实际上是PID上有限生成模结构定理), \mathbb{F}^n 作为 $F[\lambda]$ -模的结构为

$$\mathbb{F}^n \cong \bigoplus_{i=1}^s \mathbb{F}[\lambda]/(d_i)$$

于是与 A 交换的矩阵 B 是 $\mathbb{F}[\lambda]$ -模同态: $B \in \text{End}_{\mathbb{F}[\lambda]}(\mathbb{F}^n) = \text{Hom}_{\mathbb{F}[\lambda]}(\mathbb{F}^n, \mathbb{F}^n)$ (对比: 一般的矩阵 $B \in \text{End}_{\mathbb{F}}(\mathbb{F}^n) = \text{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$). 因此

$$C(A) = \text{End}_{\mathbb{F}[\lambda]}(\mathbb{F}^n) \cong \text{End}_{\mathbb{F}[\lambda]}(\bigoplus_{i=1}^s \mathbb{F}[\lambda]/(d_i)) \cong \bigoplus_{1 \leq i, j \leq s} \text{Hom}_{\mathbb{F}[\lambda]}(\mathbb{F}[\lambda]/(d_i), \mathbb{F}[\lambda]/(d_j)).$$

于是

$$\begin{aligned} \dim C(A) &= \dim \left(\bigoplus_{1 \leq i, j \leq s} \text{Hom}_{\mathbb{F}[\lambda]}(\mathbb{F}[\lambda]/(d_i), \mathbb{F}[\lambda]/(d_j)) \right) \\ &= \sum_{1 \leq i, j \leq s} \dim_{\mathbb{F}}(\text{Hom}_{\mathbb{F}[\lambda]}(\mathbb{F}[\lambda]/(d_i), \mathbb{F}[\lambda]/(d_j))) \\ &= \sum_{1 \leq i, j \leq s} \deg \gcd(d_i, d_j) \\ &= \sum_{1 \leq i, j \leq s} \min\{\deg(d_i), \deg(d_j)\} \\ &= \sum_{i=1}^s \deg d_i + 2 \sum_{1 \leq i < j \leq s} \deg d_i \\ &= \sum_{i=1}^s \deg d_i + \sum_{i=1}^s 2(s-i) \deg d_i \\ &\geq \sum_{i=1}^s \deg d_i = n. \end{aligned}$$

且等号当且仅当 $s = 1$ 时取得. □

练习: 将上面的定理推广到求方程 $AX = XB$ 的解空间维数上. 特别地证明以下推论:

Theorem 8 (Sylvester Equation).

复数域上矩阵方程 $AX = XB$ 有非平凡解当且仅当 A 与 B 有公共特征值. (也见练习 80)

Exercise 83

设 V 是 n 维 F -线性空间. $T \in \text{Hom}(V, V)$ 循环幂零. 求 $\text{Hom}(V, V)$ 的子空间 $M = \{U \in \text{Hom}(V, V) | T^2 U = UT^2\}$ 的维数.

Solution 83

由 T 幂零知所有特征值为0, 由 T 循环知每个特征值只有一个Jordan块. 因此 T 相似于 $N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$. 由练习66知 T^2 的不变因子为

1. $2 \mid n$ 时: $x^{n/2}, x^{n/2}$;
2. $2 \nmid n$ 时: $x^{(n-1)/2}, x^{(n+1)/2}$.

再由定理7知:

$$\dim M = \begin{cases} (4-2+1)\frac{n}{2} + (4-4+1)\frac{n}{2} = 2n & (2 \mid n) \\ (4-2+1)\frac{n-1}{2} + (4-4+1)\frac{n+1}{2} = 2n-1 & (2 \nmid n) \end{cases}$$

第六章 张量积

Exercise 84

设 V 是有限维向量空间, $T \in \text{End}(V)$. 证明:

$$\det(\lambda I - T) = \sum_{k=0}^n (-1)^k \text{tr}(\bigwedge^k T) \lambda^{n-k}.$$

Solution 84

选定一组基 e_1, \dots, e_n , 将 T 等同于这组基下的矩阵, 只要证明 T 所有的 k 阶主子式之和等于 $\text{tr} \bigwedge^k T$ 即可 (因为按定义展开 $\det(\lambda I - T)$, 出现在 λ^{n-k} 系数上正是 $(-1)^k$ 倍的所有 k 阶主子式之和). 记指标集

$$\Lambda = \{\alpha = (\alpha_1, \dots, \alpha_k) | \alpha_i \in \{1, \dots, n\}, \alpha_1 < \dots < \alpha_k\},$$

P_α 为向子空间 $\langle e_i \rangle_{i \in \alpha}$ 投影的矩阵, 而 $T_\alpha = T$ 表示 T 中 α 对应行列构成的主子式, 则要证明 $\text{tr} \bigwedge^k T = \sum_{\alpha \in \Lambda} \det T_\alpha$.

计算迹只需要知道每一个基元素 $e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}$ 在线性变换下的像在对应方向上的分量是什么 (即线性变换对应矩阵的对角线元素). 注意到 $\bigwedge^k(V)$ 中向 $e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}$ 方向投影的映射正是 $\bigwedge^k P_\alpha$. 于是计算

$$\left(\bigwedge^k P_\alpha \bigwedge^k T \right) e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k} = \left(\bigwedge^k P_\alpha \bigwedge^k T \bigwedge^k P_\alpha \right) e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k} = \left(\bigwedge^k (P_\alpha T P_\alpha) \right) e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}.$$

而最右边的式子等于 $\det T_\alpha e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}$ (将 $\bigwedge^n T = \det T \cdot \text{id}_{\bigwedge^n V}$ 应用到子空间 $\langle e_i \rangle_{i \in \alpha}$ 上并将 $P_\alpha T P_\alpha$ 视作子空间上的线性变换 T_α). 这样就说明了 $\text{tr} \bigwedge^k T = \sum_{\alpha \in \Lambda} \det T_\alpha$, 证毕.

注记: 若使用一些代数几何的方法则本题比较容易证明.

首先过渡到代数闭域 \bar{F} (显然域扩张不影响行列式和迹). 然后考虑当 T 可对角化的情形, 选取一组 T 的特征向量 e_1, \dots, e_n 作为 V 的基, 而 λ_i 是对应 e_i 的特征向量. 则 $\bigwedge^k T(e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}) = (\prod_{i=1}^k \lambda_{\alpha_i}) e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}$. 于是自然基 $e_{\alpha_1} \wedge \dots \wedge e_{\alpha_k}$ 全部都是 $\bigwedge^k T$ 的特征向量, 对应特征值为 $\prod_{i=1}^k \lambda_{\alpha_i}$. 这样直接对所以特征值求和计算得到 $\text{tr} \bigwedge^k T = \sigma_k(\lambda_1, \dots, \lambda_n)$, 其中 σ_k 为基本对称多项式. 再由韦达定理 (根与系数关系) 知道相差一个 $(-1)^k$ 的意义下这就是 T 特征多项式中 $n-k$ 项的系数. 于是对可对角化的 T 我们有欲证明的等式成立.

将 $\text{End}(V)$ 等同于 \bar{F}^{n^2} , 赋予其Zariski拓扑(即满足所有的闭集为多项式方程组的零点集的拓扑). 在此拓扑下全体可对角矩阵是稠密的 (全体可对角矩阵至少包含特征多项式无重根的全体矩阵, 而特征多项式无重根当且仅当特征多项式的判别式不为零, 于是特征多项式无重根的全体矩阵在Zariski拓扑下构成开集, 因此稠密.) 于是存在一个 \bar{F}^{n^2} 的稠密集合满足 $\det(\lambda I - T) = \sum_{k=0}^n (-1)^k \text{tr}(\bigwedge^k T) \lambda^{n-k}$ 左右 λ 对应系数相等的多项式方程组. 这样的多项式方程组必须对于全体 \bar{F}^{n^2} 也成立, 于是欲证等式成立.

Exercise 85

设 $\omega \in \bigwedge^p(V) \setminus \{0\}$, 其中 $p \leq n := \dim V$. 若存在 $v_1, \dots, v_p \in V$ 使得 $\omega = v_1 \wedge \dots \wedge v_p$, 则称 ω 是可分解的.

- (i) 定义 $\text{ann}(\omega) := \{v \in V \mid \omega \wedge v = 0\}$. 说明它是 V 的子空间, 而且维数 $\leq p$.
- (ii) 设 $x_1, \dots, x_p \in V$ 线性无关, 对 $\omega := x_1 \wedge \dots \wedge x_p$ 证明 $\text{ann}(\omega) = \langle x_1, \dots, x_p \rangle$.
- (iii) 选定 ω 和 $\text{ann}(\omega)$ 的基 e_1, \dots, e_r , 说明存在 $\eta \in \bigwedge^{p-r}(V)$ 使得

$$\omega = e_1 \wedge \dots \wedge e_r \wedge \eta.$$

以此说明 $r = p$ 当且仅当 ω 可分解.

Solution 85

- (i) 显然 $\text{ann}(\omega)$ 满足对加法和数乘封闭, 它是 V 的子空间. 选取 $\text{ann}(\omega)$ 的一组基 e_1, \dots, e_r ($r = \text{rank ann}(\omega)$)并扩充 e_{r+1}, \dots, e_n 成为 V 的一组基. 记指标集

$$\Lambda = \{\alpha = (\alpha_1, \dots, \alpha_p) \mid \alpha_i \in \{1, \dots, n\}, \alpha_1 < \dots < \alpha_p\},$$

则 ω 可以表示为 $\sum_{\alpha \in \Lambda} c_\alpha e_{\alpha_1} \wedge \cdots \wedge e_{\alpha_p}$.

考虑计算 $\omega \wedge e_1 = \sum_{\alpha \in \Lambda} c_\alpha e_{\alpha_1} \wedge \cdots \wedge e_{\alpha_p} \wedge e_1$, 由定义知计算结果为0. 而 $\{e_{\alpha_1} \wedge \cdots \wedge e_{\alpha_p} \wedge e_1 | \alpha \in \Lambda, \alpha_1 \neq 1\}$ 在 $\bigwedge^{p+1}(V)$ 中线性无关. 所以这迫使所有 $\alpha_1 \neq 1$ 的 c_α 为0.

下面继续计算 $\omega \wedge e_2$, 由 ω 的展开式以及前一步知 $\omega \wedge e_2 = \sum_{\alpha \in \Lambda, \alpha_1=1} c_\alpha e_{\alpha_1} \wedge e_{\alpha_p} \wedge e_2 = 0$. 类似地我们知道所有满足 $\alpha_2 \neq 2$ 的 c_α 都为0.

继续计算 $\omega \wedge e_3, \omega \wedge e_4, \dots$, 我们将得到 $\alpha_3 = 3, \alpha_4 = 4, \dots$. 所以如果 $r > p$, 就必须有 $\omega = c e_1 \wedge e_2 \wedge \cdots \wedge e_p$. 但 $e_{p+1} \in \text{ann}(\omega)$, 这显然不可能.

(ii) 将 x_1, \dots, x_p 扩充成 V 的一组基 x_1, \dots, x_n . 对 V 中任意向量 $v = \sum_{i=1}^n c_i x_i$, 我们有 $\omega \wedge v = \sum_{i=p+1}^n c_i \omega \wedge x_i$. 所以 $\omega \wedge v = 0$ 当且仅当 $c_{p+1} = \cdots = c_n = 0$, 如所欲证.

(iii) 继续第一部分的讨论, 我们已经知道 $\alpha_1 = 1, \alpha_2 = 2, \dots, \alpha_r = r$. 这就是要证明的存在 $\eta \in \bigwedge^{p-r}(V)$ 使得 $\omega = e_1 \wedge \cdots \wedge e_r \wedge \eta$. 若 $r = p$, 这显然说明 ω 可分解. 而反之由第二部分知道对可分解的 $0 \neq \omega \in \bigwedge^p(V)$ 一定有 $\dim \text{ann}(\omega) = p$, 证毕.

Exercise 86

承上题, 证明所有 $\omega \in \bigwedge^{n-1}(V)$ 都是可分解的 ($n := \dim V$).

Solution 86

记基域为 F . 注意到 $\bigwedge^n(V) = \{c e_1 \wedge \cdots \wedge e_n | c \in F\} \cong F$, 而我们有线性映射

$$A: \begin{array}{ccc} V & \rightarrow & \bigwedge^n(V) \cong F \\ v & \mapsto & \omega \wedge v \end{array}.$$

显然 $\ker A = \text{ann}(\omega)$, 且由维数公式得 $\dim \ker A = \dim V - \dim \text{im } A \geq n - 1$. 结合练习85第一部分知 $\dim A = n - 1 = p$, 再由该练习的第三部分得 ω 可分解.

Exercise 87

设 U 和 W 都是 V 的 p 维子空间, U 有基 x_1, \dots, x_p 而 W 有基 y_1, \dots, y_p . 证明 $U = W$ 当且仅当 $x_1 \wedge \dots \wedge x_p$ 和 $y_1 \wedge \dots \wedge y_p$ 成比例.

Solution 87

若 $U = W$, 则将 y_i 写成 x_j 的线性组合并展开 $y_1 \wedge \dots \wedge y_p$, 则由外积性质显然结果将是 $x_1 \wedge \dots \wedge x_p$ 的常数倍, 又因为 $y_1 \wedge \dots \wedge y_p$ 线性无关, 因此结果不为零, 故 $y_1 \wedge \dots \wedge y_p$ 是 $x_1 \wedge \dots \wedge x_p$ 的非零常数倍, 即两者成比例.

反之, 若 $x_1 \wedge \dots \wedge x_p$ 和 $y_1 \wedge \dots \wedge y_p$ 成比例 (相差 $\lambda \neq 0$ 倍) 但 $U \neq W$, 则必存在 $y_i \notin U$. 考虑

$$(x_1 \wedge \dots \wedge x_p) \wedge y_i = \lambda(y_1 \wedge \dots \wedge y_p) \wedge y_i,$$

式子左边所含向量线性无关, 因此外积非零, 而右边含有重复的向量 y_i , 外积为零, 矛盾! 因此 $U = W$.

Exercise 88

若基域 F 特征不为2, 设 $\dim V \geq 2$ 而 $\omega \in \bigwedge^2(V)$, 证明 ω 可分解当且仅当 $\omega \wedge \omega = 0$.

Solution 88

若 ω 可分解, 显然有 $\omega \wedge \omega = 0$, 只要证明反方向.

对维数用归纳法, 当 $\dim V = 2$ 时 $\dim \bigwedge^2(V) = 1$, 显然所有元素都是 $e_1 \wedge e_2$ 的倍数, 当然可分解.

下面单独考虑 $\dim V = 3$ 情形, 给定 $\omega \in \bigwedge^2(V)$, 定义

$$A: \begin{array}{ccc} V & \rightarrow & \bigwedge^3(V) \\ v & \mapsto & \omega \wedge v. \end{array}$$

因为 $\dim \bigwedge^3(V) = 1$, 由维数公式可知 $\dim \ker A \geq 2$, 选取 $\ker A$ 的一组基 u_1, u_2 , 再添加 $u_3 \in V$ 扩充成 V 的一组基. 将 ω 用这组基的外积表示:

$$\omega = \lambda_1 u_2 \wedge u_3 + \lambda_2 u_3 \wedge u_1 + \lambda_3 u_1 \wedge u_2.$$

由 u_1, u_2 的选取方式知 $\omega \wedge u_1 = \omega \wedge u_2 = 0$, 进而知道 $\lambda_1 = \lambda_2 = 0$. 于是 $\omega = \lambda_3 u_1 \wedge u_2$ 是可分解的.

假设定理对于 $\dim V \leq n-1$ 的所有情形已经成立, 再来考虑 $\dim V = n$ 的情形. 选取 V 的一组基 v_1, \dots, v_n , 同样将 ω 用 $v_i \wedge v_j$ 表示:

$$\begin{aligned}\omega &= \sum_{1 \leq i < j \leq n} a_{ij} v_i \wedge v_j \\ &= \sum_{i=1}^{n-1} a_{in} v_i \wedge v_n + \sum_{1 \leq i < j \leq n-1} a_{ij} v_i \wedge v_j \\ &= u \wedge v_n + \eta\end{aligned}$$

其中 $u = \sum_{i=1}^{n-1} a_{in} v_i \in U := \langle v_1, \dots, v_{n-1} \rangle$, 而 $\eta \in \wedge^2(U)$.
注意到

$$0 = \omega \wedge \omega = (u \wedge v_n + \eta) \wedge (u \wedge v_n + \eta) = 2u \wedge \eta \wedge v_n + \eta \wedge \eta$$

这是因为 $u \wedge v_n \wedge \eta = \eta \wedge u \wedge v_n$ (它们相差一个偶置换, 想一想为什么). 而 v_n 不出现在 u 和 η 中, 因此 $u \wedge \eta = 0, \eta \wedge \eta = 0$.

由归纳假设知 η 可分解: $\eta = u_1 \wedge u_2$ ($u_1, u_2 \in U$). 于是 $u \wedge u_1 \wedge u_2 = 0$. 但是我们知道, 若干个向量外积为零当且仅当它们线性相关. 所以存在不全为零的 λ, μ_1, μ_2 使得

$$\lambda u + \mu_1 u_1 = \mu_2 u_2 = 0.$$

若 $\lambda = 0$, 则 u_1, u_2 线性相关, $\eta = u_1 \wedge u_2 = 0$. $\omega = u \wedge v_n$ 可分解.

若 $\lambda \neq 0$, 则 u 为 u_1 和 u_2 的线性组合. 这样 ω 就是 $u_1 \wedge v_n, u_2 \wedge v_n$ 和 $u_1 \wedge u_2$ 的线性组合, 由 $\dim V = 3$ 情形的讨论知道 ω 也是可分解的.

Exercise 89

说明向量空间 V 的对称代数 $\text{Sym}(V)$ 连同线性映射 $\iota : V = \text{Sym}^1(V) \hookrightarrow \text{Sym}(V)$ 具有以下泛性质: 对于所有交换代数 A ,

$$\begin{aligned}\{\text{代数同态 } \text{Sym}(V) \rightarrow A\} &\rightarrow \text{Hom}(V, A) \\ f &\mapsto f \circ \iota\end{aligned}$$

是1-1映射. 按此说明一旦 V 有限维, 选定 V 的基 v_1, \dots, v_n , 则有代数同构 $\text{Sym}(V) \cong F[X_1, \dots, X_n]$ 映 v_i 为 X_i .

Solution 89

$f \mapsto f \circ \iota$ 显然是单的, 因为一个代数同态 $\text{Sym}(V) \rightarrow A$ 完全由 V 的像决定. 而 $f \mapsto f \circ \iota$ 也是满的, 因为一旦给定线性映射 $f_0 : V \rightarrow A$, 它就自然地通过 $\text{Sym}(V)$ 上的乘法扩展到整个 $\text{Sym}(V)$ 上从而给出一个 $f : \text{Sym}(V) \rightarrow A$ 的代数同态.

当 V 有限维时更是显然有 $\text{Sym}(V) \cong F[X_1, \dots, X_n]$, 因为张量代数 $T(V) = \bigoplus_{n \geq 0} V^{\otimes n}$ 此时正是 v_1, \dots, v_n 生成的自由代数, 取交换化之后得到多项式代数.

写在后面

到这里, 我们一学期的高等代数II习题课就全部讲完了. 不过, 对于一个数院的同学来说, 他的数学学习才刚刚开始. 在学期末, 一个不可避免的挑战就是期末考试. 如果你在考试中取得了优异的成绩, 那我自然要祝贺你, 但是如果不巧 (我是说如果), 你没有取得理想的成绩呢?

我想这也并不是一件什么了不得的事. 就在我准备这份讲义的同时, 2022年菲尔兹奖结果揭晓: 39岁的韩裔数学家许埈珥因为他在组合数学方面引入代数几何工具所做出的优秀结果获得了当年的菲尔兹奖. 然而回首他的学术生涯其实并非一帆风顺: 在他刚进入首尔国立大学开始本科阶段的学习时, 他的志向是成为一名科学记者, 而他的专业是物理与天文. 然而事实证明他的兴趣和长处并不在此, 经常翘课导致他不得不重上了好几门课程. 许埈珥说:“当时我感到迷茫”.

事情的转机出现在他本科的第六年. 在这一年, 日本代数几何的领军人物, 菲尔兹奖得主Hironaka到首尔国立大学访问并开设一门为期一年的代数几何课程, 许埈珥想: 也许他可以通过听Hironaka的课与这位著名数学家混熟, 这样Hironaka就可以成为他作为科学记者的第一个采访对象. 他总是与Hironaka共进午餐, 从这时开始, 他才真正发现了自己的天赋所在: 数学. 在Hironaka的指导下他完成了在首尔国立大学的硕士. 在申请博士时, 几乎所有的大学都因为他的背景拒绝了他: 本科专业不是数学, 而且成绩单也并不出彩, 只有伊利诺斯大学香槟分校接受了他. 在这里的第一年, 许埈珥的数学生涯一飞冲天: 他在这里解决了悬而未决四十多年的Read猜想: 一个图的染色多项式系数绝对值总是对数凹的. 密歇根大学邀请他去做一场关于Read猜想的报告, 报告厅里坐满了一年前曾拒绝他的申请的教授, 一名教授极力建议一名博士后参加这场报告, 而理由是: “三十年后你可以骄傲地告诉你的孙辈们, 你在许出名之前听过他的报告”. 这场报告无疑是成功的, 密歇根大学在报告后马上邀请许埈珥转学到他们那. 在那

里, 许埭珥把目光转向了Rota猜想——这是一种Read猜想的推广. 2015年, 许埭珥和Karim Adiprasito以及Eric Katz一起解决了Rota猜想, 这项工作最终使许埭珥获得了2022年的菲尔兹奖.

我分享这个故事是想告诉同学们, 考试成绩并不能贬低一个人的能力. 在进入北大之前, 你们所有人都证明了自己至少在某方面拥有不平凡的能力, 偶尔的失利并不会抹杀掉这种能力. 在2018年北京大学毕业生晚会上, 当年的中文男足球队队长曹直说过这样一番话: “谁说十八岁的成功就不是成功? 既然站上过巅峰, 还怕什么深渊无穷——退一寸有退一寸的欢喜”. 一次考试没有成功不算什么, 人生还有很长的路要走.

所以请允许我用克林克兹的一段台词来结束这份讲义——

“与其感慨路难行, 不如马上出发.”