# 高等代数(II)习题课教案

version: 0.2

xiaxueqaq@stu.pku.edu.cn

August 2022

# 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35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30)并不单独出现, 而是作为中国剩余定理

(练习29)的推论产生, 并进一步给出Hermite插值 (练习31). 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

# 第一章　习题课 - 多项式I

## Exercise 1

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

## Solution 1

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

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

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:

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

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

$$\begin{aligned}\implies\quad &(x^2+2x+3)=(x-2)[(x^3-2)-(x-2)(x^2+2x+3)]+11\\ \implies\quad &(x^2+4x+5)g(x)-(x-2)f(x)=11\end{aligned}$$

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41.

## Exercise 2

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 2

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 3

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 3

(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2一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 4

$\mathbb{C}$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 4

$f'(x)=3x^2+12x+3a$, 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 1.** *设$A = a_0x^d + \cdots + a_d$, $B = b_0x^e + \cdots + b_e$为一整环$R$上的单变元多项式. 定义$A$和$B$的**Sylvester**结式为:*

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

*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$\operatorname{res}(A,B) = 0$, 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$\operatorname{res}(A,B) = 0$.*

*证明.* 记$P_n$为$R$上次数小于$n$的多项式集合, 则映射

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

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

回到本题, $f(x)$和$f'(x)$的Sylvester结式为:

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

因式分解得到$\operatorname{res}(f,f') = 108(-4+a)^2(5+a)$. 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.

xiaxueqaq 2022

## Exercise 5

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

## Solution 5

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

## Exercise 6

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

## Solution 6

不妨设$n > m$, 记$n$除以$m$的余数为$n\%m$: 则$\gcd(x^n-1, x^m-1) = \gcd(x^n-1-x^{n-m}(x^m-1), x^m-1) = \gcd(x^{n-m}-1, x^m-1) = \cdots = \gcd(x^{n\%m}-1, x^m-1)$. 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 7

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 7

让我们先看$A$可逆的情形:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \left| \begin{pmatrix} I & O \\ -CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \right| = \begin{vmatrix} A & B \\ O & D - CA^{-1}B \end{vmatrix}$$

$$= |A||D - CA^{-1}B| = |AD - ACA^{-1}B| = |AD - CAA^{-1}B| = |AD - CB|$$

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$\lambda$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$\lambda$的不恒为零的多项式, 因此至多只有有限个$\lambda$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 8

若方阵$A$为幂零阵, $A^m = O$, 证明$I + A$可逆.

### Solution 8

由$1/(1 + x)$的Taylor展开:

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

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

于是:

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

**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 9

设$a_1$, …, $a_n$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 9

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$a_1,\ldots,a_n$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, $a_1$, …, $a_n$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: $a_1,\ldots,a_n$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.

# 第二章　习题课 - 多项式II

## Exercise 10

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 10

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:

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

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

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

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4中的定理1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. 例如计算本例中的$f$和$g$的主子结式链得到 $\{-2t^4u - t^3u^2 - 4t^3u + 2t^2u^2 + 12t^2u - 4tu^2 - 14tu - u^3 - 7u^2 + 8u,\ t^3 + 3t^2 - tu - 3t - u + 4,\ -t - 1,\ 1\}$.

使得$f$和$g$的最大公因式恰为二次多项式的条件是前两个多项式为0而第三个多项式不为0. 解

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

得

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

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 11

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 11

显然$x^2+x+1$的两根为$\omega = \frac{-1+\sqrt{3}i}{2}$和$\omega^2 = \overline{\omega} = \frac{-1-\sqrt{3}i}{2}$. 由$(x^2+x+1) \mid [f_1(x^3)+xf_2(x^3)]$知$\omega$和$\omega^2$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$\omega$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 12

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 12

完全类似练习11. $x^{n-1}+\cdots+x+1=0$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 13

若$a$, $b$, $c$为实数, 证明:

$$a,b,c>0\Leftrightarrow a+b+c>0,\ ab+ac+bc>0,\ abc>0.$$

## Solution 13

$\Rightarrow$: 平凡的.

$\Leftarrow$: 记$\sigma_1=a+b+c$, $\sigma_2=ab+ac+bc$, $\sigma_3=abc$. 则$(x-a)(x-b)(x-c)=x^3-\sigma_1x^2+\sigma_2x-\sigma_3=0$有三个实根$a$, $b$, $c$. 若某根小于等于零则所有项都小于等于零且常数项小于零, 方程不能成立! 因此所有根都是正根: $a,b,c>0$.

练习: 推广本题到$n$个实数情形.

注记: 事实上有算法可以机械地计算一个实系数多项式的实根个数, 可以参考《高等代数》(丘维声著)中的Sturm定理或者维基百科页面.

## Exercise 14

若三次方程$x^3+px+q=0$ $(q\neq 0)$的三个根分别为$a$, $b$, $c$. 求另一多项式方程使得其三根分别为$\frac{b+c}{a^2}$, $\frac{c+a}{b^2}$, $\frac{a+b}{c^2}$.

## Solution 14

由韦达定理: $a+b+c=0$, 所以$\frac{b+c}{a^2}=-\frac{1}{a}$, $\frac{c+a}{b^2}=-\frac{1}{b}$, $\frac{a+b}{c^2}=-\frac{1}{c}$. 即三根分别为$a$, $b$, $c$的负倒数. 将$z=-\frac{1}{x}\Leftrightarrow x=-\frac{1}{z}$中得到

$$(-\frac{1}{z})^3+p(-\frac{1}{z})+q=0.$$

两边同乘$z^3$知所求方程为$-1-pz^2+qz^3=0$.

注记: 类似地, 若求一方程使得其根为原多项式根之倒数, 则只需将原多项式系数全部颠倒即可, 颠倒系数后的多项式称为原多项式的互反多项式.

## Exercise 15

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 15

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)!

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

最后$a^3b^3c^3=(abc)^3$. 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1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:

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

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 16

$a_1, a_2, \ldots, a_n$两两不同, 求证: 关于$x_1, \ldots, x_n$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 16

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

$$x_1+x_2a_i+x_3a_i^2+\cdots+x_na_i^{n-1}+a_i^n=0$$

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

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

由韦达定理展开比较系数知:

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

其中$\sigma_1$, ..., $\sigma_n$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.

## Exercise 17 (中国剩余定理)

(1) 求最小的正整数$x$,使得$x \equiv 4 \pmod 5$, $x \equiv 1 \pmod 7$;

(2) 若数域$K$上一元多项式$g$, $h \in K[x]$满足$\gcd(g,h) = 1$, 给定$r_1$, $r_2 \in K[x]$, 求$f \in K[x]$使得$f \equiv r_1 \pmod g$, $f \equiv r_2 \pmod h$.

## Solution 17

(1) 由扩展欧几里得算法知: $\gcd(5,7) = 1$, $5\times3-7\times2 = 1$. 所以$5\times3\times1 \equiv 1 \pmod 7$, $-7\times2\times4 \equiv 4 \pmod 5$.

$$\Longrightarrow \begin{cases} 5\times3\times1-7\times2\times4 &\equiv 1 \pmod 7 \\ 5\times3\times1-7\times2\times4 &\equiv 4 \pmod 5 \end{cases}$$

而$5\times3\times1-7\times2\times4 = -41 \equiv 29 \pmod{35}$. 故欲求的$x = 29$.

(2) $\gcd(g,h) = 1 \implies \exists u, v \in K[x]$ s.t. $ug + vh = 1$.

即$ug \equiv 1 \pmod h$, $vh \equiv 1 \pmod g$. 故$ugr_2 \equiv r_2 \pmod h$, $vhr_1 \equiv r_1 \pmod g$.

因此$f = ugr_2 + vhr_1$满足$\begin{cases} f &\equiv r_1 \pmod g \\ f &\equiv r_2 \pmod h \end{cases}$

思考: 这样的$f$唯一吗? 如何添加类似(1)中的"最小"限制?

# 第三章 习题课 - Jordan标准形I

## Exercise 18

设$f, g, \varphi, \psi \in K[\lambda]$, 且$f, g$分别与$\varphi, \psi$互素. 求证:

$$\begin{pmatrix} f\varphi & \\ & g\psi \end{pmatrix} \sim \begin{pmatrix} g\varphi & \\ & f\psi \end{pmatrix}$$

## Solution 18

分别计算行列式因子, 对第一个$\lambda$-矩阵:

$$\begin{aligned} D_1 &= \gcd(f\varphi, g\psi) = \gcd(f, g\psi)\gcd(\varphi, g\psi) = \gcd(f, g)\gcd(\varphi, g)\gcd(f, \psi)\gcd(\varphi, \psi) \\ &= \gcd(f, g)\gcd(\varphi, \psi), \\ D_2 &= fg\varphi\psi \end{aligned}$$

对第二个$\lambda$-矩阵有:

$$\begin{aligned} D_1 &= \gcd(g\varphi, f\psi) = \gcd(g, f\psi)\gcd(\varphi, f\psi) = \gcd(g, f)\gcd(g, \psi)\gcd(\varphi, f)\gcd(\varphi, \psi) \\ &= \gcd(f, g)\gcd(\varphi, \psi), \\ D_2 &= fg\varphi\psi \end{aligned}$$

由$\lambda$-矩阵相抵当且仅当具有相同的秩和行列式因子知两矩阵相抵.

注记: 也可通过相抵操作一步步转化过去, 相较这里展示的做法稍显繁琐.

## Exercise 19

$F \subseteq K$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 19

$\Rightarrow$: 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 20

设$F$为一数域, 证明$A$与$A^T$相似.

## Solution 20

令$S = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}$, 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$\mathrm{diag}(J_1, \ldots, J_r)$, 其中每个$J_k$为一Jordan块. 则容易发现$A^T$相似于$\mathrm{diag}(J_1^T, \ldots, J_r^T)$. 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19知$A$和$A^T$在$F$上相似.

而

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

于是直接计算立即有$S^{-1}J_kS = J_k^T$, 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 21

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

## Solution 21

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

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

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 22

如何计算Jordan标准形?

## Solution 22

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$J_1$, …, $J_r$对角线上为$\lambda$, $J_{r+1}$, …, $J_s$对角线上不为$\lambda$. 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: $n-\mathrm{rank}(A-\lambda I)=r$, 于是从$\mathrm{rank}(A-\lambda I)$可以计算出从属于$\lambda$的Jordan块个数 (即至少为一阶的$\lambda-$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21知道平方后所有大于等于2阶的$\lambda-$Jordan块秩减少1, 而一阶的Jordan块秩仍是0. 因此$\mathrm{rank}(A-\lambda I)-\mathrm{rank}(A-\lambda I)^2$是至少2阶的$\lambda-$Jordan块数量.

一般地, $\mathrm{rank}(A-\lambda I)^{t-1}-\mathrm{rank}(A-\lambda I)^t$是至少$t$阶的$\lambda-$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$\lambda_1$, …, $\lambda_m$;

Step 3. 对每个特征值$\lambda_i$,

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

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

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

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

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

计算

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

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 23

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$D:\begin{array}{ccc}\mathbb{C}[x]_n & \to & \mathbb{C}[x]_n\\ f & \mapsto & f'\end{array}$ 为求导算子.

具体写出$D$在单项式基$1,x,\ldots,x^n$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 23

(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$\operatorname{rank} M = n = (n+1)-1$, 且0就是所有特征值, 因此由练习22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$e_1$, ..., $e_{n+1}$为自然基, 则$e_{n+1} = (0, \ldots, 0, 1)^T$经过$M$反复作用后:

$$e_{n+1} \overset{M}{\mapsto} ne_n \overset{M}{\mapsto} n(n-1)e_{n-1} \overset{M}{\mapsto} \cdots \overset{M}{\mapsto} n!e_1,$$

于是$P = (n!e_1, n!/1e_2, \ldots, e_{n+1})$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$f, f', f'', \ldots, f^{(p)} \in W$, 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$\cdots\cdots$ 这样$f, f', f'', \ldots, f^{(p)} \in W$就与$x^p, x^{p-1}, \ldots, 1$张成相同的线性空间. 因此$W = \mathbb{C} \oplus \mathbb{C}x \oplus \cdots \oplus \mathbb{C}x^p = C[x]_p$.

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

# 第四章　习题课 - Jordan标准形II

### Exercise 24

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 24

令自然基为$e_1, \ldots, e_n$, 则$Ae_1 = e_2$, $Ae_2 = e_3$, $\cdots$, $Ae_{n-1} = e_n$, $Ae_n = \sum_{k=0}^{n-1} -a_k e_{k+1}$. 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Caylay-Hamilton定理知$A$的特征多项式$\varphi_A$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 25

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

## Solution 25

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21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

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

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 26

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 26

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$\varphi_A$. 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$e_1 = (1, 0, \ldots, 0)^T$, 则$v$, $Av$, $A^2v$, $\ldots$, $A^{n-1}v$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2.

## Exercise 27 (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 27

令

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

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

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

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$Y(x):\mathbb{R}\to\mathbb{R}^n$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

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

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: $\sin A$, $\cos A$.

## Exercise 28 (*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 28

(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27: 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.

## Exercise 29 (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$R$为含幺环, $N_1, \ldots, N_r \trianglelefteq R$为$R$的非平凡理想且两两互素 ($N_i + N_j = (1)$). 令$\sigma_i$表示自然同态$\sigma_i : R \to R/N_i$ ($i = 1 \ldots r$), 则映射

$$\sigma : \begin{array}{ccc} R & \to & R/N_1 \oplus \cdots \oplus R/N_r \\ x & \mapsto & (\sigma_1(x), \ldots, \sigma_r(x)) \end{array}$$

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$K[x]$, $a_1$, …, $a_r\in R$且两两互素. 再给定$b_1$, …, $b_r\in R$. 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$\bmod a_1a_2\cdots a_r$的意义下是唯一的: 即若$x_1$, $x_2\in R$都满足该方程组, 则$x_1-x_2\equiv 0 \pmod{a_1\ldots a_r}$.

证明该定理.

## Solution 29

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

$$u_2a_1+v_2a_2=u_3a_1+v_3a_3=\cdots=u_ra_1+v_ra_r=1$$

全部乘起来得到$\prod_{k=2}^r(u_ka_1+v_ka_k)=1$, 展开并合并全部含$a_1$的项得$a_1u+\prod_{k=2}^r(v_ka_k)=1$. 因此$a_1$与$a_2\cdots a_r$互素: $x_1a_1+y_1(a_2\cdots a_r)=1$. 同理有:

$a_2$与$a_1a_3\cdots a_r$互素: $x_2a_2+y_2a_1a_3\cdots a_r=1$

…

$a_r$与$a_1\cdots a_{r-1}$互素: $x_ra_r+y_ra_1\cdots a_{r-1}=1$

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

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

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 30 (*Lagrange*插值)

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

## Solution 30

此即

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

由中国剩余定理 (练习29)立得.

## Exercise 31 (*Hermite*插值)

在Lagrange插值30的要求上, 我们还额外要求前$s$个点处导数有特定值: $f'(a_1) = d_1, \ldots, f'(a_s) = d_s$. 求$f$.

## Solution 31

此即

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

仍由中国剩余定理 (练习29)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.

## Exercise 32 (子空间的并)

设$V$是无限域$F$上的有限维线性空间, $V_1, \ldots, V_s$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$e_1, \ldots, e_n$均不落在$\bigcup_{k=1}^{s} V_k$中.

## Solution 32

(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: $\exists \alpha \notin V_1 \cup \cdots \cup V_{s-1}$. 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$c_1$, $c_2$. …, $c_s$ ($F$是无限域)和$\beta \notin V_s$. 我们断言: 在$s$个向量

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

中必有一者不落在$V_1 \cup \cdots \cup V_{s-1}$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$\alpha$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$V_1 \cup \cdots \cup V_{s-1}$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: $c_i\alpha + \beta \notin V_1 \cup \cdots \cup V_{s-1} \cup V_s$.

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

# 第五章　习题课 - 线性变换

## Exercise 33 *(Fitting Lemma)*

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

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

## Solution 33

先来证明两个链条件:

$$\begin{aligned} \exists m \in \mathbb{N}_+ \quad \text{s.t.} \quad & \ker \varphi^m = \ker \varphi^{m+1} = \ker \varphi^{m+2} = \cdots \\ & \ker \varphi^m = \ker \varphi^{m+1} = \ker \varphi^{m+2} = \cdots \end{aligned} \tag{5.1}$$

事实上我们总是有：

$$\begin{aligned} \ker \varphi \subseteq \ker \varphi^2 \subseteq \ker \varphi^3 \subseteq \cdots \\ \operatorname{im} \varphi \supseteq \operatorname{im} \varphi^2 \supseteq \operatorname{im} \varphi^3 \supseteq \cdots \end{aligned}$$

于是有

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

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5.1成立.

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

1. 说明直和:$\ker \varphi^m \cap \operatorname{im} \varphi = 0$ 若$x \in \ker \varphi^m \cap \operatorname{im} \varphi^m$, 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$\ker\varphi^m = \varphi^{2m}$, $y$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: $\forall x \in V:\ \exists y \in \ker\varphi^m,\ z \in \operatorname{im}\varphi^m$ s.t. $x = y + z$.

   注意到$\varphi^m(x)\operatorname{im}\varphi^m = \operatorname{im}\varphi^{2m} : \exists z \in \operatorname{im}\varphi^m$ 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 34 *(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 34

$\Rightarrow$: 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$\mathrm{Spec}(A)=\{\lambda_1,\dots,\lambda_s\}$. $V$为全空间, $V_{\lambda_i}$为从属于特征值$\lambda_i$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$\lambda_i$的$A$的特征向量, 它落在$V_{\lambda_i}$中. 因此$V_{\lambda_i}$为$B$-不变子空间.

而$B$可对角化, 所以它在$V_{\lambda_i}$上的限制$B\mid_{V_{\lambda_i}}$也可对角化 (想一想, 为什么?). 于是$V_{\lambda_i}$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.

## Exercise 35

$K$为一数域, 映射$F:K^{n\times n}\to K^{n\times n}(n\geq 2)$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$\operatorname{tr}F=-3$, 求$F$的Jordan标准形.

## Solution 35

(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$\lambda$只能为$\pm 1$.

   (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) $\operatorname{tr} F$为全体特征值记重数之和:

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

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

## Exercise 36

令$V = \mathbb{R}^3$, $S \subset V$为$V$中由$2x - 2y + z = 0$定义的子空间, $P : V \to S \subset V$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 36

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

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

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

$$\begin{array}{rcccc}\forall v\in V: x= & (x-Ax) & + & Ax \\ & \pitchfork & & \pitchfork \\ & \ker P & & \operatorname{im} P\end{array}$$

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 37 (脑筋急转弯)

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

(1) 求$P$的极小多项式;

(2) 求$L$的极小多项式.

## Solution 37

(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 38

设$A,B\in\mathbb{C}^{n\times n}$, $A$的特征多项式为$\varphi_A$, 求证:
$\varphi_A(B)$可逆$\Leftrightarrow$ $A$和$B$没有公共特征值.

## Solution 38

$\Leftarrow$: 由$A$, $B$没有公共特征值以及代数基本定理可知, $\gcd(\varphi_A, \varphi_B) = 1$ ($\phi_B$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$0 \notin \mathrm{Spec}(\varphi_A(B))$. 但是由谱映射定理知: $\mathrm{Spec}(\varphi_A(B)) = \varphi_A(\mathrm{Spec}(B))$.

$$\begin{tikzcd} B \arrow[rr, "\varphi_A"] \arrow[d, "\mathrm{Spec}"] & & \varphi_A(B) \arrow[d, "\mathrm{Spec}"] \\ \mathrm{Spec}(B) \arrow[r, "\varphi_A"] & \varphi_A(\mathrm{Spec}(B)) = & \mathrm{Spec}(\varphi_A(B)) \end{tikzcd}$$

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

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

## Exercise 39

对给定域$F$上$n$阶方阵$A \in F^{n\times n}$, 定义$L : \begin{array}{ccc} F^{n\times n} & \to & F^{n\times n} \\ X & \mapsto & AX - XA \end{array}$. 若$A$可对角化, 求证$L$可对角化.

## Solution 39

将$A$对角化: $A = PDP^{-1}$, 其中$D = \mathrm{diag}(d_1, \dots, d_n)$为一对角阵. 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 40

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 40

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

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 41 (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 41

(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$1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$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$c_0, \ldots, c_{n-1}$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$1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$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.

# 第六章　习题课 - 可对角化和可交换专题

### Exercise 42

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

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

(1) 证明$S$是$\mathrm{Hom}(V, V)$的子空间;

(2) 求$\dim S$.

### Solution 42

(1) 显然, 若$V, W$都是$A_1, A_2 \in \mathrm{Hom}(V, V)$的不变子空间. 则$U, W$也是$k_1A_1 + k_2A_2$ $(k_1, k_2 \in F)$的不变子空间. 于是$S$是$\mathrm{Hom}(V, V)$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$\begin{array}{cccc} & m & & \\ m & \left(\begin{matrix}\boxed{*} \\ O\end{matrix}\right. & \left.\begin{matrix} O \\ \boxed{*}\end{matrix}\right) & r \\ & & r & \end{array}$. 此时$\dim S = m^2 + r^2$.

再看一般的情况, 选取$U \cap W$的一组基$e_1$, …, $e_d$, 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: $\begin{pmatrix} * & O & O \\ * & * & * \\ O & O & * \end{pmatrix}$.

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 43

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 43

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:

$$0 \xleftarrow{N^k} e_1 \xleftarrow{N^k} e_{k+1} \xleftarrow{N^k} \cdots$$

$$0 \xleftarrow{N^k} e_2 \xleftarrow{N^k} e_{k+2} \xleftarrow{N^k} \cdots$$

$$\vdots$$

$$0 \xleftarrow{N^k} e_k \xleftarrow{N^k} e_{2k} \xleftarrow{N^k} \cdots$$

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:

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

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 44

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 44

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

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

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$\lambda$的特征子空间$V_\lambda$的一组基为$x_1, \ldots, x_r$, 则$Bx_1, \ldots, Bx_r$为$BA$的一组线性无关的关于$\lambda$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$c_1 = \cdots = c_r = 0$, $Bx_1, \ldots, B_r$的确线性无关. 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$\lambda$的特征向量. 记$\mathrm{geo.mult.}_\lambda(M)$表示$M$关于$\lambda$的几何重数. 则我们已经知道

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

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$\mathrm{geo.mult.}_\lambda(AB) = \mathrm{geo.mult.}_\lambda(BA)$. 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:

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

由$TU$可对角化知, $\mathrm{rank}(TU) = \mathrm{rank}(TU)^2 = \mathrm{rank}(TU)^3$. 于是$\mathrm{rank}(UT)^2 = \mathrm{rank}(TU)^2$, 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:

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

其中$\mathrm{alg.mult.}_\lambda(M)$表示$M$关于$\lambda$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.

xiaxueqaq 2022

## Exercise 45

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 45

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

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

其中$\lambda_1, \dots \lambda_s$两两不同. 则$f(\lambda) = m_{A^2}(\lambda^2) = (\lambda^2 - \lambda_1) \cdot \dots \cdot (\lambda^2 - \lambda_s) = (\lambda - \sqrt{\lambda_1})(\lambda + \sqrt{\lambda_1}) \cdot \dots \cdot (\lambda - \sqrt{\lambda_s})(\lambda + \sqrt{\lambda_s})$零化$A$. $m_A(\lambda)|f(\lambda)$至多含有一个二次因子$\lambda^2$, 其余因子都是一次因子. 因此$A$的Jordan块中除了可能的$\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}$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 46

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 46

$\Rightarrow$: 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:

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

于是$P = P^T, S_1 = P^2 = P^T P$.

这样

$$A = S_1S_2 = P^2S_2 \sim P^{-1}P^2S_2P = PS_2P = P^TS_2P$$

而$S_2$为正定对称阵, $P^TS_2P$为正定二次型, 其特征值全为正数. 所以

$$A \sim P^TS_2P \sim \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} > 0$$

$\Leftarrow$: 将$A$正交相似到对角阵, 并从中写出乘积分解:

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

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 47

设$\lambda_1, \ldots, \lambda_t$为线性变换$A$的不同特征值. $\alpha_1$, ..., $\alpha_t$是这些特征值的特征向量. 求证$\beta := \alpha_1 + \cdots + \alpha_t$生成的循环子空间$U = F\alpha_1 \oplus \cdots \oplus F\alpha_t$.

## Solution 47

$F\alpha_1$, ..., $F\alpha_t$都是$A$的不变子空间, 因此$F\alpha_1 \oplus \cdots \oplus F\alpha_t$也是$A$的不变子空间. 所以$\beta \in F\alpha_1 \oplus \cdots \oplus F\alpha_t$生成的循环子空间$U \subseteq F\alpha_1 \oplus \cdots \oplus F\alpha_t$.

反过来, 令$f_i(\lambda) = \prod\limits_{1\le k\le t,\ k\ne i}(\lambda-\lambda_k)$. 则$f_i(\lambda_i)\ne 0$, $f_i(\lambda_j)=0$ $(j\ne i)$. 那么

$$\begin{aligned}
f_i(\lambda_i)^{-1}f_i(A)\beta &= f_i(\lambda_i)^{-1}f_i(A)\textstyle\sum_{j=1}^t\alpha_j\\
&= f_i(\lambda_i)^{-1}\textstyle\sum_{j=1}^t f_i(A)\alpha_j\\
&= f_i(\lambda_i)^{-1}\textstyle\sum_{j=1}^t f_i(\lambda_j)\alpha_j\\
&= f_i(\lambda_i)^{-1}f_i(\lambda_i)\alpha_i\\
&= \alpha_i.
\end{aligned}$$

所以$\alpha_i\in U$, $F\alpha_1\oplus\cdots\oplus F\alpha_t\subseteq U$. 这样最终得到$F\alpha_1\oplus\cdots\oplus F\alpha_t=U$.

注记: 本题从另一个角度印证了从初等因子求不变因子的方法: 把所有初等因子按所含不可约因子分类并按次数从高到低排列, 每次从每一类中选取不可约因子中次数最高的一个初等因子, 全部相乘得到一个不变因子. 这就对应本题中一些不同初等因子对应的不变子空间可以组装成一个更大的不变子空间.

练习: 设线性空间$V$上一线性变换为$A$. $\alpha$, $\beta\in V$. $f,g\in F[x]$分别为使$f(A)\alpha=0$, $g(A)\beta=0$的次数最低的多项式. 若$f$, $g$互素, 证明$\alpha$生成的循环子空间$F[A]\alpha$与$\beta$生成的循环子空间$F[A]\beta$直和为$\alpha+\beta$生成的循环子空间$F[A](\alpha+\beta)$.

## Exercise 48

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: $\operatorname{rank} S\le n^2-n$.

## Solution 48

注意到$S(X)=0\Leftrightarrow AX=XA$, 于是只要说明$\dim C(A)\ge n$, 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$J=\operatorname{diag}(J_{m_1}(\lambda_1),\ldots,J_{m_s}(\lambda_s))$且$P^{-1}AP=J$. 其中$J_{m_i}(\lambda_i)$表示特征值为$\lambda_i$, 大小为$m_i$的Jordan块. 显然我们有$\sum_{i=1}^s m_i=n$

令$X_{i,j}=P\operatorname{diag}(0,0,\ldots,0,[J_{m_i}(\lambda_i)]^j,0,\ldots,0)P^{-1}$, 特别地$j=0$时$X_{i,0}=$

$P\,\mathrm{diag}(0,0,\ldots,0,I_{m_i},0,\ldots,0)P^{-1}$. 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$X_{i,j}$ $(i = 1,\ldots,s,\ j = 0,\ldots,m_i - 1)$是线性无关的. 因此$\dim C(A) \geq n$, $\mathrm{rank}\, S \leq n^2 - n$.

**另解:** 事实上我们有著名的Cecioni-Frobenius定理:

**Theorem 2** (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$d_1$, …, $d_s$ $(d_i|d_{i+1})$. 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: $B \in \mathrm{End}_{\mathbb{F}[\lambda]}(\mathbb{F}^n) = \mathrm{Hom}_{\mathbb{F}[\lambda]}(\mathbb{F}^n, \mathbb{F}^n)$ (对比: 一般的矩阵$B \in \mathrm{End}_{\mathbb{F}}(\mathbb{F}^n) = \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$). 因此

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

于是

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

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 3** (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38)*

## Exercise 49

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

## Solution 49

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43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2知:

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

# 第七章　习题课 - 内积I

### Exercise 50

设$U = \operatorname{span}(\alpha_1, \alpha_2)$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 50

(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$\min\limits_{v\in U}\langle\alpha-v,\ \alpha-v\rangle$. 但这是直接的:

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

所以最短距离为$\frac{2}{\sqrt{3}}$.

## Exercise 51

设实线性空间$V$上的双线性函数$f(\alpha,\beta)$在$V$的基$\alpha_1,\alpha_2,\alpha_3$下的矩阵为$G=\begin{pmatrix}1&0&0\\0&3/2&-1/2\\0&-1/2&3/2\end{pmatrix}$.

(1) 证明$f(\alpha,\beta)$构成$V$上的内积;

(2) 求内积$f$下的一组标准正交基$\beta_1,\ \beta_2,\ \beta_3$使得基中每个元素都是$G$的特征向量;

(3) 问在内积$f$下, 是否存在正交变换$A$使得$A\alpha_1=\alpha_1,\ A\alpha_2=\alpha_3$. 若存在, 写出$A$在$\beta_1,\ \beta_2,\ \beta_3$下的矩阵.

## Solution 51

(1) 双线性性和对称性是显然的, 下证正定性. 由顺序主子式$D_1=1,\ D_2=\frac{3}{2},\ D_3=2$知度量矩阵正定. $f$是内积;

(2) 复习求标准正交基的算法:

(Step 1) 求特征值: $\lambda_1=\lambda_2=1,\ \lambda_3=2$.

(Step 2) 求每个特征值的特征子空间的一组基: $V_1=\mathrm{span}(\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix})$, $V_2=\begin{pmatrix}0\\-1\\1\end{pmatrix}$.

(Step 3) 将每个特征子空间做Gram-Schmidt正交化. $V_1 = \mathrm{span}(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix})$,

$V_2 = \begin{pmatrix} 0 \\ -1/2 \\ 1/2 \end{pmatrix}$.

(Step 4) 合起来就是标准正交基.

(3) 由$\alpha_1$, $\alpha_2$经$A$的像知$A$在以$\alpha_1$, $\alpha_2$, $\alpha_3$为基的矩阵表示为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{或} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -\frac{2}{3} \end{pmatrix}.$$

其中前两列是由$\alpha_1$, $\alpha_2$的像得到, 最后一列是因为$A$是正交变换, 设$A\alpha_3 = u\alpha_1 + v\alpha_2 + w\alpha_3$, 则有:

$$\begin{cases} 0 & = & \langle \alpha_1,\ \alpha_3 \rangle = \langle A\alpha_1,\ A\alpha_3 \rangle = \langle \alpha_1,\ A\alpha_3 \rangle = u \\ -\frac{1}{2} & = & \langle \alpha_2,\ \alpha_3 \rangle = \langle A\alpha_2,\ A\alpha_3 \rangle = \langle \alpha_3,\ A\alpha_3 \rangle = -\frac{1}{2}v + \frac{3}{2}w \\ \frac{3}{2} & = & \langle \alpha_3,\ \alpha_3 \rangle = \langle A\alpha_3,\ A\alpha_3 \rangle = u^2 + \frac{3}{2}v^2 + \frac{3}{2}w^2 - vw. \end{cases}$$

解得

$$\begin{cases} u & = & 0 \\ v & = & 1 \\ w & = & 0 \end{cases} \text{或} \begin{cases} u & = & 0 \\ v & = & -1 \\ w & = & -\frac{2}{3} \end{cases}$$

因此$A$在$\beta_1$, $\beta_2$, $\beta_3$下矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{或} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & -\frac{2\sqrt{2}}{3} \\ 0 & \frac{2\sqrt{2}}{3} & -\frac{1}{3} \end{pmatrix}.$$

注记: 标准正交基不是唯一的, 因此第二小问答案不唯一.

另外, 今后在微分流形的学习中我们将会看到全体正交矩阵$O(n)$构成一个$\mathrm{GL}_n(\mathbb{R})$中的维数为$\frac{n(n-1)}{2}$的子流形, 因此**粗糙地来说**, 只要指定3个自由变量就可以确定一个正交变换, 在本题中指定了$A\alpha_1 = \alpha_1$, $A\alpha_2 = \alpha_3$就是指定了三个自由量 (一共六个数, 有三个正交条件).

## Exercise 52

设实线性空间$V$上的双线性函数$f(\alpha,\beta)$在$V$的基$\alpha_1,\alpha_2,\alpha_3$下的矩阵为$G=\begin{pmatrix} 4/3 & 1/3 & 1/3 \\ 1/3 & 4/3 & 1/3 \\ 1/3 & 1/3 & 4/3 \end{pmatrix}$.

(1) 证明$f(\alpha,\beta)$构成$V$上的内积;

(2) 求内积$f$下的一组标准正交基$\beta_1$, $\beta_2$, $\beta_3$使得基中每个元素都是$G$的特征向量;

(3) 问在内积$f$下, 是否存在正交变换$A$使得$A\alpha_1=\alpha_1$, $A\alpha_2=\alpha_3$. 若存在, 写出$A$在$\beta_1$, $\beta_2$, $\beta_3$下的矩阵.

## Solution 52

(1) 双线性性和对称性是显然的, 下证正定性. 由顺序主子式$D_1=\frac{4}{3}$, $D_2=\frac{5}{3}$, $D_3=2$知度量矩阵正定. $f$是内积;

(2) 复习求标准正交基的算法:

(Step 1) 求特征值: $\lambda_1=\lambda_2=1$, $\lambda_3=2$.

(Step 2) 求每个特征值的特征子空间的一组基: $V_1=\mathrm{span}(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix})$, $V_2=\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

(Step 3) 将每个特征子空间做Gram-Schmidt正交化. $V_1=\mathrm{span}(\begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{pmatrix},\begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix})$, $V_2=\begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}$.

(Step 4) 合起来就是标准正交基.

(3) 由$\alpha_1$, $\alpha_2$经$A$的像知$A$在以$\alpha_1$, $\alpha_2$, $\alpha_3$为基的矩阵表示为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{或} \begin{pmatrix} 1 & 0 & \frac{2}{5} \\ 0 & 0 & -1 \\ 0 & 1 & \frac{2}{5} \end{pmatrix}.$$

其中前两列是由$\alpha_1, \alpha_2$的像得到, 最后一列是因为$A$是正交变换, 设$A\alpha_3 = u\alpha_1 + v\alpha_2 + w\alpha_3$, 则有:

$$\begin{cases} \frac{1}{3} &= \langle \alpha_1,\ \alpha_3\rangle = \langle A\alpha_1,\ A\alpha_3\rangle = \langle \alpha_1,\ A\alpha_3\rangle = \frac{4}{3}u + \frac{1}{3}v + \frac{1}{3}w \\ \frac{1}{3} &= \langle \alpha_2,\ \alpha_3\rangle = \langle A\alpha_2,\ A\alpha_3\rangle = \langle \alpha_3,\ A\alpha_3\rangle = \frac{1}{3}u + \frac{1}{3}v + \frac{4}{3}w \\ \frac{4}{3} &= \langle \alpha_3,\ \alpha_3\rangle = \langle A\alpha_3,\ A\alpha_3\rangle = \frac{4}{3}(u^2+v^2+w^2) + \frac{2}{3}(uv+uw+vw). \end{cases}$$

解得

$$\begin{cases} u &= 0 \\ v &= 1 \\ w &= 0 \end{cases} \text{或} \begin{cases} u &= \frac{2}{5} \\ v &= -1 \\ w &= \frac{2}{5} \end{cases}$$

因此$A$在$\beta_1, \beta_2, \beta_3$下矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \text{或} \begin{pmatrix} \frac{3}{5} & -\frac{2\sqrt{3}}{5} & \frac{2}{5} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{4}{5} & -\frac{3\sqrt{3}}{10} & \frac{3}{10} \end{pmatrix}.$$

注记: 标准正交基不是唯一的, 因此第二小问答案不唯一.

今后在微分流形的学习中我们将会看到全体正交矩阵$O(n)$构成一个$\mathrm{GL}_n(\mathbb{R})$中的维数为$\frac{n(n-1)}{2}$的子流形, 因此**粗糙地来说**, 只要指定3个自由变量就可以确定一个正交变换, 在本题中指定了$A\alpha_1 = \alpha_1$, $A\alpha_2 = \alpha_3$就是指定了三个自由量 (一共六个数, 有三个正交条件).

## Exercise 53

设$A : X \mapsto AX$是带标准内积的欧氏空间$\mathbb{R}^4 \to \mathbb{R}^3$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 53

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

$$\max_{\|x\|=1} \|AX\|^2 = \max_{\|x\|=1} \langle Ax,\ Ax\rangle = \max_{\|x\|=1} x^TA^TAx = \max_{\|x\|=1} x^TQ^TDQx$$

$$= \max_{\|x\|=1}(Qx)^T D(Qx) = \max_{\|x\|=1} x^T Dx = \lambda_{\max}(A^T A)$$

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$\min\limits_{\|x\|=1} \|Ax\|^2 = \lambda_{\min}(A^T A)$. 取到最大值和最小值的位置分别为$\lambda_{\max}(A^T A)$和$\lambda_{\min}$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 54

设$A$是规范方阵($AA^* = A^*A$). $\lambda$和$\mu$是$A$的两个不同特征值. $\alpha$, $\beta$分别是从属于$\lambda$和$\mu$的特征向量. 求证$\langle \alpha,\ \beta \rangle = 0$.

## Solution 54

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, 规范方阵$A$关于$\lambda$的特征向量$\alpha$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$A - \lambda I$也是规范方阵.

于是:

$$\begin{tikzcd}
\langle \alpha,\ A\beta \rangle \arrow[r, equal] \arrow[d, equal] & \langle \alpha,\ \mu\beta \rangle \arrow[r, equal] & \mu \langle \alpha,\ \beta \rangle \\
\langle A^*\alpha,\ \beta \rangle \arrow[r, equal] & \langle \bar{\lambda}\alpha,\ \beta \rangle \arrow[r, equal] & \lambda \langle \alpha,\ \beta \rangle
\end{tikzcd}$$

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

## Exercise 55

设$A$是$n$阶实规范方阵. $(a+bi)$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$\alpha$, $\beta$正交, 长度相等.

## Solution 55

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54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$\alpha$, $\beta$正交, 长度相等.

## Exercise 56

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

## Solution 56

由$f(\alpha,\alpha)=\alpha^TA\alpha=0$知$(\alpha^TA\alpha)^T=\alpha^TA^T\alpha=0$. 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 57 (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 57

$n=2$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).

(a) n=2 (b) n=3

图 7.1: $n=2,3$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$\alpha_1,\ \ldots,\ \alpha_m$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$x_1,\ \ldots,\ x_{m-1}$使得$x_1\alpha_1+\cdots+x_{m-1}\alpha_{m-1}=0$. 经过适当调换顺序, 不妨假设$x_1\geq\cdots\geq x_r\geq 0,\ 0>x_{r+1}\geq\cdots\geq x_{m-1}$. 这样就有

$$x_1\alpha_1+\ldots+x_r\alpha_r=-x_{r+1}\alpha_{r+1}-\cdots-x_{m-1}\alpha_{m-1}$$

再与$\alpha_m$做内积得到:

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

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

但是我们又有:

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

这显然矛盾. 于是$\alpha_1,\ \ldots,\ \alpha_{m-1}$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:

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

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).

# 第八章　习题课 - 内积II

## Exercise 58

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

## Solution 58

$\Leftarrow$: 这是容易的, $A^*A = f(A)A = Af(A) = AA^*$, 因此$A$是规范方阵.

$\Rightarrow$: 由$A$是规范方阵知, $A$可以酉相似对角化: 存在酉方阵$U$使得 $U^{-1}AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 于是$(U^{-1}AU)^* = U^{-1}A^*U = \mathrm{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})$. 问题转化为了构造多项式$f$使得$f(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)) = \mathrm{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})$. 事实上这只需要$f(\lambda_1) = \overline{\lambda_1}$, ..., $f(\lambda_n) = \overline{\lambda_n}$. 由Lagrange插值多项式(见练习30)知这样的$f$存在. 于是$U^{-1}f(A)U = f(U^{-1}AU) = U^{-1}A^*U$, 即$f(A) = A^*$.

## Exercise 59

若$A$是实规范方阵, 且对实方阵$B$有$AB = BA$. 证明也有$A^TB = BA^T$

## Solution 59

由练习58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

$$A^TB = f(A)B = \sum_{k=0}^n c_kA^kB = \sum_{k=0}^n c_kBA^k = Bf(A) = BA^T.$$

## Exercise 60

设$\mathbb{F}$为任意特征不为2的域 (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 60

$\Rightarrow$: 设$P^TAP = D_1$, $P^TBP = D_2$, $D_1$, $D_2$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$S_1 = P^TAP$, $S_2 = P^TBP$, 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$S_1^{-1}S_2 = P^{-1}A^{-1}BP = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 于是$S_2 = S_1D$, 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$D = \mathrm{diag}(\lambda_1 I, \ldots, \lambda_s I)$, 其中$\lambda_1$, ..., $\lambda_s$两两不同. 则$S_1$和$S_2$分块对角: $S_1 = \mathrm{diag}(M_1, \ldots, M_s)$, $S_2 = \mathrm{diag}(N_1, \ldots, N_s)$, 且$N_i = \lambda_i M_i$ $(i = 1, \ldots, s)$. 设可逆阵$Q_i$将$M_i$相合到对角阵$D_i$ (由于char $\mathbb{F} \neq 2$, 这总是能做到). 令$Q = \mathrm{diag}(Q_1, \ldots, Q_s)$: 则我们有

$$Q^TS_1Q = \mathrm{diag}(Q_1^T, \ldots, Q_s^T)\cdot\mathrm{diag}(M_1, \ldots, M_s)\cdot\mathrm{diag}(Q_1, \ldots, Q_s) = \mathrm{diag}(D_1, \ldots, D_s),$$

$$Q^TS_2Q = \mathrm{diag}(Q_1^T, \ldots, Q_s^T)\cdot\mathrm{diag}(N_1, \ldots, N_s)\cdot\mathrm{diag}(Q_1, \ldots, Q_s) = \mathrm{diag}(\lambda_1 D_1, \ldots, \lambda_s D_s).$$

于是可逆阵$Q$将$A$, $B$同时相合到对角阵.

## Exercise 61

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 61

这里与练习34是完全类似的. 注意到练习34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 62 ($SVD$分解)

设$A \in \mathbb{C}^{n\times n}$, 求证存在$m$阶酉方阵$P$和$n$阶酉方阵$Q$使得

$$PAQ = \begin{pmatrix} \mu_1 & & & \\ & \ddots & & \\ & & \mu_r & \\ & & & O \end{pmatrix}$$

其中$r = \operatorname{rank} A$, $\mu_1,\ \ldots,\ \mu_r$是$A^*A$的全体非零特征值$\lambda_1,\ \ldots,\ \lambda_r$的算术平方根. 称为矩阵$A$的奇异值,

## Solution 62

首先$A^*A$是半正定Herimite方阵, 存在酉方阵$Q$使得$Q^*A^*AQ = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & O \end{pmatrix}$.

取$D = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_r} & \\ & & & I \end{pmatrix}$, 则:

$$D^{-1*}Q^*A^*AQD^{-1} = \begin{pmatrix} 1/\sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & 1/\sqrt{\lambda_r} & \\ & & & I \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & O \end{pmatrix} \begin{pmatrix} 1/\sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & 1/\sqrt{\lambda_r} & \\ & & & I \end{pmatrix} = \begin{pmatrix} I_r & \\ & O_{n-r} \end{pmatrix}.$$

记$B = AQD^{-1} \in \mathbb{C}^{m\times n}$, 则$B^*B = \begin{pmatrix} I_r & \\ & O_{n-r} \end{pmatrix}$. 再令$\beta_i$为$B$的第$i$列, 上式说明:

$$\langle \beta_i,\ \beta_j \rangle = 0\ (\forall i \neq j),\quad \langle \beta_i,\ \beta_i \rangle = \begin{cases} 1 & 1 \leq i \leq r \\ 0 & r+1 \leq i \leq n \end{cases}$$

这样$\beta_{r+1} = \cdots = \beta_n = 0$, $\beta_1,\ \ldots,\ \beta_r$为$\mathbb{C}^m$中一组两两正交的单位向量, 将其扩充成一组标准正交基$\beta_1, \ldots, \beta_r, \gamma_{r+1}, \ldots, \gamma_m$. 令$P = (\beta_1, \ldots, \beta_r, \gamma_{r+1}, \ldots, \gamma_m)^*$, 则$P$是酉方阵, 且$PB = (\beta_1, \ldots, \beta_r, \gamma_{r+1}, \ldots, \gamma_m)^*(\beta_1, \ldots, \beta_r, 0, \ldots, 0) = \begin{pmatrix} I_r & \\ & O_{n-r} \end{pmatrix}$.

于是$PAQ = PBD = \begin{pmatrix} I_r & \\ & O_{n-r} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \ddots & & \\ & & \sqrt{\lambda_r} & \\ & & & I \end{pmatrix} = \begin{pmatrix} \mu_1 & & & \\ & \ddots & & \\ & & \mu_r & \\ & & & O \end{pmatrix}$,
如所欲证.

**Exercise 63** (*Cochran*分解)

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

(i) $\forall k \in \{1, 2, \ldots, s\} : A_k^2 = A_k$;

(ii) $\sum_{k=1}^s \operatorname{rank} A_k = n$;

(iii) $\forall i, j \in \{1, 2, \ldots, s\}(i \neq j) : A_i A_j = O$.

## Solution 63

$(i) \implies (ii)$: 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$\operatorname{rank} A_k = \operatorname{tr} A_k$. 所以

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

$(ii) \implies (iii)$: 因为$n = \operatorname{rank} \sum_{k=1}^s A_k \leq \sum_{k=1}^s \operatorname{rank} A_k = n$, 所以$\mathbb{F}^n = \bigoplus_{k=1}^s \operatorname{im} A_k$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:

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

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$w \in \operatorname{im} A_j$, $w$的直和分解式中$i \neq j$处都是0, 所以$A_i w = 0$. 于是$\operatorname{im} A_j \subseteq \ker A_i$, 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.

# 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, 许埈珥把目光转向了Rota猜想——这是一种Read猜想的推广. 2015年, 许埈珥和Karim Adiprasito以及Eric Katz一起解决了Rota猜想, 这项工作最终使许埈珥获得了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."

xiaxueqaq 2022