

线 性 代 数(2)

学 习 指 导

许甫华 编著

清 华 大 学 出 版 社
北 京

内 容 简 介

本书是学习和使用线性代数(高等代数)较深内容的辅导书,与《理工科代数基础》(俞正光等著,1998年清华大学出版社出版)相配套.包括代数系统、多项式、线性空间、线性变换、欧几里得空间、酉空间及二次型共7章内容.每章安排:理论脉络、典型题详解讲评、考试题详解评析3部分内容.书中的内容融入了作者在中国科学技术大学和清华大学长期任教的经验和心得.适用于各类高校学生学习线性代数、高等代数、矩阵论等课程时复习参考,也可供使用代数知识的科技人员参考.

图书在版编目(CIP)数据

线性代数(2)学习指导/许甫华编著. —北京:清华大学出版社,2003
ISBN 7-302-06562-4

I .线… II .许… III .线性代数 - 高等学校 - 教学参考资料
IV .O151 .2

中国版本图书馆 CIP 数据核字(2003)第 027644 号

出 版 者:清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.com.cn>

责任编辑:刘 颖

封面设计:

版式设计:韩爱君

印 刷 者:

发 行 者:新华书店总店北京发行所

开 本:850×1168 1/32 印张:8 字数:200千字

版 次:2003年6月第1版 2003年6月第1次印刷

书 号:ISBN 7-302-06562-4/O·295

印 数:0001~ 000

定 价:0.00元

引 言

目前,清华大学理工科(非数学系)的线性代数课程安排为两学期.第一学期为基础内容,课程名称为“几何与代数(1)”;第二学期为进一步的内容,课程名称为“几何与代数(2)”.本书就是针对第二学期的内容而编写的辅导参考书,所以书的名字取为《线性代数(2)学习指导》.书中章节的顺序是按《理工科代数基础》(俞正光等编著,清华大学出版社1998年出版)一书组织的.

本书内容包括:二次型、代数系统、多项式形式、线性空间、线性变换、欧几里得空间、酉空间共7章.每章首先总结理论脉络;然后给出典型题的详尽解答和分析与讲评;最后对往期的期末考试题作详细分析解答.本书对例题的选择,主要是考虑问题的典型性和代表性,力图使每道题说明一个或几个问题和方法.读者会发现,有的问题的解答,需要融会理解定理和定义;有的解题方法,需要有巧妙的构思.本书不少问题(包括一些难题)的解答有独特的思路和方法,融入了作者在中国科学技术大学数学系、少年班、非数学系和清华大学全校性实验班,理工科各系的长期的教学经验和积累.期望能通过本书,启发学生更好理解代数的理论,掌握解决问题的思路和方法,克服在代数的学习、复习、迎考和应用中遇到的困难.也希望由此更能培养学习代数的兴趣,增进对代数理论的深入理解.

每章均设计如下3个板块.

理论脉络 华罗庚教授特别提倡“由薄到厚,由厚变薄”的读书方法.浅尝即止的“薄”固不可取,食而不化的“厚”也非佳境.薄—厚—薄是一个辩证的过程,是否定之否定的过程.最后的“薄”是要融会贯通.知识要自然联系.本版块就是力图体现这一“读书

法则”,帮助读者理出各章节主干,进而从知识的总体联系上把握理论脉络,理解概念和定理结论.同时结合各章内容,也总结了解决某类问题的方法.

典型题讲评 选取了《理工科代数基础》中的部分问题,也从各类代数参考书中选取一些典型问题,作者也自行设计一些典型问题.对这些问题给出详细解答并分析解题思路 and 进行讲评.可以说这些题的选取都是为了说明一个或几个理论问题,强调有关知识和方法.读者会发现许多新奇的,不同于以往的解题方法和思路.期望能启发读者掌握解决问题的思路和方法,体会到代数课学习和解题的乐趣.

考试题评析 精心挑选能涵盖本章基本理论、基本概念及解题技巧的以往期末考试题,给出标准解答;同时分析所用的知识点、联想方法、思考方向等.并适当指出以往考生易出错的地方,提醒读者正确理解和掌握代数的基础理论,提高抽象思维的能力和严谨的逻辑推理能力.

多年来,在教学中,一直有许多学生反映,希望有相应的学习辅导书.特别是近年教改以后,课内学时(和讲课内容)压缩较大,学生课外自己支配的自学时间增多.学生对学习辅导书的要求更强烈.这促使我们下决心编写本书.

为了学好线性代数这样比较抽象的课程,理解理论和演练习题二者缺一不可.抽象的理论和方法在具体的例题中,在不断的解题实践中,会变得生动直观,印象深刻.而只有深入系统地学习和理解基本理论,并在典型问题的解决中不断分析、思考、总结,才能不断提高分析和解决问题的能力.建议读者在使用本书时,先自己动手独立寻找解题思路和方法,再看解答,最后看讲评(或评析).这样既可考查读者对知识的掌握程度和解题能力,增强自信心和学习兴趣,又可对一些方法印象深刻,感悟更深.

本书适用于各类高校本科生或研究生学习“线性代数”、“高等代数”、“矩阵论”等课程时作为参考书;也可供使用代数知识的科学技术人员参考.不足之处,请读者指正.

作 者

2003 年春于清华园

目 录

引言	I
第 1 章 二次型和实对称矩阵的相合	1
1.1 理论脉络	1
1.2 典型题讲评	8
1.3 考试题评析	36
第 2 章 代数系统	48
2.1 理论脉络	48
2.2 典型题讲评	52
2.3 考试题评析	62
第 3 章 多项式形式	66
3.1 理论脉络	66
3.2 典型题讲评	73
3.3 考试题评析	82
第 4 章 线性空间	89
4.1 理论脉络	89
4.2 典型题讲评	94
4.3 考试题评析	117
第 5 章 线性变换	129
5.1 理论脉络	129
5.2 典型题讲评	137
5.3 考试题评析	165

第 6 章 欧几里得空间·····	183
6.1 理论脉络 ·····	183
6.2 典型题讲评 ·····	189
6.3 考试题评析 ·····	211
第 7 章 酉空间·····	218
7.1 理论脉络 ·····	218
7.2 典型题讲评 ·····	222
7.3 考试题评析 ·····	240
参考文献·····	247

第 1 章 二次型和实对称矩阵的相合

1.1 理论脉络

本章内容可概括为以下 4 个问题:

1 二次型的定义, 二次型和实对称矩阵的 1-1 对应关系

定义 1 一个系数在数域 F 中的 n 个变数 $x_1, x_2, \dots, x_n$ 的二次齐次多项式

$$Q(x_1, x_2, \dots, x_n) = a_{11} x_1^2 + 2a_{12} x_1 x_2 + \dots + 2a_{1n} x_1 x_n + \\ a_{22} x_2^2 + \dots + 2a_{2n} x_2 x_n + \\ \dots + a_{nn} x_n^2$$

称为数域 F 上的一个 n 元二次型.

令 $a_{ij} = a_{ji}$, 因为 $x_i x_j = x_j x_i$, 故可以把上述定义中的上三角形形式写成矩阵形式(即把非平方项分一半到左下角). 记 $A = (a_{ij})$, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, 则可用矩阵乘法把二次齐次多项式表示出来, 即有

$$Q(x_1, x_2, \dots, x_n) = a_{11} x_1^2 + a_{12} x_1 x_2 + \dots + a_{1n} x_1 x_n + \\ a_{21} x_2 x_1 + a_{22} x_2^2 + \dots + a_{2n} x_2 x_n + \\ \dots + \\ a_{n1} x_n x_1 + a_{n2} x_n x_2 + \dots + a_{nn} x_n^2 \\ = \sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j = \mathbf{x}^T \mathbf{A} \mathbf{x}. \quad (1)$$

由 $a_{ij} = a_{ji}$, 知 $A^T = A$, 故 A 是对称矩阵. 当 $F = \mathbb{R}$ 时, $Q(x_1, x_2, \dots, x_n)$ 是实二次型, 此时 A 是实对称矩阵, 我们主要讨论实二次型. A 称为二次型的矩阵, 它是由二次型中所有系数排成的, 显然它是由

二次型惟一决定的.

2 二次型化平方和的方法

从本质上说,二次型是 n 维线性空间 V 到 $\mathbb{R}$ 的映射.把 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 看成是 V 中向量 α 在某组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的坐标,则有 $Q(\alpha) = Q(x_1, x_2, \dots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.于是,与空间解析几何中求有心二次曲面的标准方程类似,要再选一组基 $\eta_1, \eta_2, \dots, \eta_n$, 使二次型在此基下的表示为平方和.设从基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 到基 $\eta_1, \eta_2, \dots, \eta_n$ 的过渡矩阵为 $\mathbf{P}$, 则 $\mathbf{P}$ 为可逆矩阵,且可形式地写为

$$\eta_1, \eta_2, \dots, \eta_n = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \mathbf{P}.$$

记 α 在这两组基下的坐标分别为 $\mathbf{x}, \mathbf{y}$, 则它们之间有关系

$$\mathbf{x} = \mathbf{P} \mathbf{y}. \quad (2)$$

把可逆线性代换 $\mathbf{x} = \mathbf{P} \mathbf{y}$ 代入(1)式中,得

$$\begin{aligned} Q(\alpha) &= Q(x_1, x_2, \dots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x} \Big|_{\mathbf{x} = \mathbf{P} \mathbf{y}} = (\mathbf{P} \mathbf{y})^T \mathbf{A} (\mathbf{P} \mathbf{y}) \\ &= \mathbf{y}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{y} = \mathbf{y}^T \mathbf{B} \mathbf{y} \\ &= Q(y_1, y_2, \dots, y_n). \end{aligned}$$

这里 $\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}$, 它是二次型在新基 $\eta_1, \eta_2, \dots, \eta_n$ 下的矩阵,它与二次型在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的矩阵 $\mathbf{A}$ 是相合的.于是,二次型在可逆线性代换下化为平方和的问题,就等价于实对称矩阵 $\mathbf{A}$ 相合于对角矩阵的问题.即寻找可逆矩阵 $\mathbf{P}$ 和对角矩阵 $\mathbf{D}$ 使得

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{D}.$$

既然这两个问题等价,故在以下介绍的 3 个方法中,有的用“矩阵的语言”叙述,有的用“二次型的语言”叙述.

方法 1 主轴化方法

定理 1 对任意 n 阶实对称矩阵 $\mathbf{A}$, 总存在实正交矩阵 $\mathbf{Q}$, 使得 $\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q}$ 是对角矩阵.

这里由 $\mathbf{Q}^{-1} = \mathbf{Q}^T$, 所以是正交相合对角矩阵, 记 $\mathbf{Q} = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n)$, 及

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

则 t 满足: $At = \lambda t$ ($i = 1, 2, \dots, n$), 即 Q 的列都是 A 的特征向量 (相互正交). 由此方法化平方和, 先写出二次型 $Q(\alpha)$ 的矩阵 A , 然后: (1) 求 A 的特征值; (2) 求 A 的属于每个特征值 λ 的特征向量; (3) 属于同一特征值的特征向量要施密特 (Schmidt) 正交化.

注 主轴化方法是大块配方法和初等变换法不能替代的, 比如在二次曲面方程的化简中只能用主轴化方法. 因为是从直角坐标系到直角坐标系的过渡矩阵, 它一定是正交矩阵, 见 1.3 节例 3.

方法 2 大块配方法

定理 2 每个二次型恒可通过可逆线性代换 $x = Py$ 化为平方和, 即

$$Q(x_1, x_2, \dots, x_n) \Big|_{x=Py} = \mu_1 y_1^2 + \dots + \mu_r y_r^2 - \mu_{r+1} y_{r+1}^2 - \dots - \mu_{r+s} y_{r+s}^2,$$

其中 $\mu_i > 0$ ($i = 1, 2, \dots, r+s$). 这种表示称为二次型的标准形.

这个定理的证明是对变量的个数作归纳法. 讨论 n 个变数的二次型时要分 4 种情况讨论, 而其中最主要的有两种:

(1) 若 $a_{11} \neq 0$, 则第一次把含 x_1 的全配方进去, 余下的是 $n-1$ 个变量的二次型, 再继续做下去, 即有

$$Q(\alpha) = a_{11} x_1^2 + \frac{1}{a_{11}} (a_{12} x_2 + \dots + a_{1n} x_n)^2 -$$

$$\frac{1}{a_{11}} \sum_{j=2}^n a_{jj} x_j^2 + \sum_{i,j=2}^n a_{ij} x_i x_j.$$

$$\text{令 } y_1 = x_1 + \frac{a_{12}}{a_{11}} x_2 + \dots + \frac{a_{1n}}{a_{11}} x_n,$$

$$y_i = x_i, \quad i = 2, \dots, n.$$

也就是做可逆线性代换

$$x_i = y_i, \quad i = 2, \cdots, n.$$

记 $\mathbf{x} = \mathbf{P}_1 \mathbf{y}$, 则有

$$\mathbf{P} = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1\ n-1}}{a_{11}} & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \text{W} & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

且有

$$\mathbf{P}_1^T \mathbf{A} \mathbf{P}_1 = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{n-1} \end{bmatrix}.$$

(2) 若 $a_{11} = 0$, 但 $a_{12} \neq 0$, 则利用中学讲的平方差公式的原理使之出现 y_1^2 项 即令

$$\begin{aligned} x_1 &= y_1 + y_2, \\ x_2 &= y_1 - y_2, \\ x_i &= y_i, \quad i = 3, \dots, n. \end{aligned}$$

即

$$\mathbf{P} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

则有

$$\begin{aligned} Q(\mathbf{a}) \Big|_{\mathbf{x} = \mathbf{P}_1 \mathbf{y}} &= 2 \mathbf{a}_2 (y_1 + y_2)(y_1 - y_2) + \cdots \\ &= 2 \mathbf{a}_2 y_1^2 - 2 \mathbf{a}_2 y_2^2 + \cdots \end{aligned}$$

个充分必要条件 .

两个 n 阶实对称矩阵 A, B 相合

^{def} 存在可逆矩阵 P , 使得 $B = P^T A P$

$r_1 = r_2$, 且 $s_1 = s_2$ (正、负惯性指数分别相等)

$\text{rank} A = \text{rank} B$, $r_1 - s_1 = r_2 - s_2$ (秩和符号差分别相等) .

其中 r_i 和 s_i ($i = 1, 2$) 分别为实对称阵的正、负惯性指数, $r_i - s_i$ 称为符号差 .

由于“相合”是一种等价关系(具有自返性、对称性、传递性), 所以 n 阶实对称矩阵的全体可以按“相合”进行等价分类, 一共可以分成多少类呢? 只要看全体 n 阶实对称矩阵的“规范型”一共有多少种即可 .

定理 4 对任意 n 阶实对称矩阵 A , 总存在可逆矩阵 P , 使得

$$P^T A P = \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{pmatrix} .$$

此式称为矩阵 A 的规范型, 其中 r, s 分别称为 A 的正、负项个数, 即正、负惯性指数 . 由此知:

当 $r = 0$ 时, s 可以取 $0, 1, \cdots, n$, 故共有 $n + 1$ 类;

当 $r = 1$ 时, s 可以取 $0, 1, \cdots, n - 1$, 共 n 类;

.....

当 $r = n$ 时, $s = 0$, 只 1 类 .

所以相合分类共有

$$1 + 2 + \cdots + n + 1 = \frac{(n+1)(n+2)}{2} \text{ (类) } .$$

在这些“类”中有 4 类是重要的, 即

$$\begin{array}{llll} r = n, & r < n, & r = 0, & r = 0, \\ s = 0; & s = 0; & s = n; & s < n . \end{array}$$

它们分别称为正定、半正定、负定、半负定矩阵.其中以正定矩阵为重点.

4 正定二次型的定义及判断二次型是否正定的充分必要条件

定义 2 若对任意 $\alpha \neq 0$, 都有 $Q(\alpha) > 0$, 则称二次型 $Q(\alpha)$ 是正定的.

定义 2' 若二次型 $Q(\alpha) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ 正定, 则称实对称矩阵 $\mathbf{A}$ 正定, 记为 $\mathbf{A} > 0$.

引理 设二次型经可逆线性代换 $\mathbf{x} = \mathbf{P} \mathbf{y}$ 化为平方和, 即

$$Q(x_1, x_2, \dots, x_n) \Big|_{\mathbf{x} = \mathbf{P} \mathbf{y}} = c_1 y_1^2 + \dots + c_n y_n^2,$$

则 $Q(x_1, x_2, \dots, x_n)$ 正定的充分必要条件是 $c_i > 0 (i = 1, 2, \dots, n)$.

设 $\mathbf{A}$ 是 n 阶实对称矩阵, 则:

$\mathbf{A}$ 正定 $\stackrel{\text{def}}{=} Q(x_1, \dots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$, 对任意 $\mathbf{x} = \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \neq \mathbf{0}$ 成立

正惯性指数是 n , 负惯性指数是 0

$\mathbf{A}$ 的特征值全大于 0

存在可逆矩阵 $\mathbf{P}$, 使 $\mathbf{A} = \mathbf{P}^T \mathbf{P}$ (等价于 $\mathbf{A}$ 相合于 $\mathbf{I}_n$)

$\mathbf{A}$ 的各阶顺序主子式 > 0

$\mathbf{A}$ 的所有主子式 > 0 .

实对称矩阵正定的定义和这 5 个充分必要条件在证明题中都各自发挥作用, 而若要判断一个给定的二次型是否正定, 通常用充要条件中的第 2 和第 4 条, 特别是用第 4 条, 即 $\mathbf{A}$ 的各阶顺序主子式 > 0 是比较方便的.

关于半正定、负定、半负定实对称矩阵的相应定理读者可自行总结或见《高等代数学》.

定理 5 设 $\mathbf{A}, \mathbf{B}$ 是 n 阶实对称矩阵, 且 $\mathbf{A}$ 正定, 则存在可逆矩阵 $\mathbf{P}$, 使得

$$P^T A P = I_n, \quad P^T B P = D (\text{对角矩阵}).$$

定理 6 设 A, B 均为 n 阶实对称矩阵, 且 $AB = BA$, 则存在正交矩阵 Q 和对角矩阵 D_1 和 D_2 , 使得

$$Q^T A Q = D_1, \quad Q^T B Q = D_2.$$

我们将在例题中给出这两个定理的证明, 同时也对具体矩阵实践一下对角化的过程, 即把证明题变成计算题. 这里我们要特别强调的是定理 5 和定理 6 的不同. 从本质上来说定理 5 说的是两个实对称矩阵同时相合于对角阵的问题, 其条件是“ A 正定”; 定理 6 是两个实对称矩阵同时相似于对角阵的问题, 其条件是“ $AB = BA$ ”, 即 A 与 B 乘法可交换, 千万不要把它们混为一谈.

1.2 典型题讲评

例 1 写出二次型 $Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^{n-2} (x_i - x_{i+2})^2$ 的矩阵.

解 方法 1 由 $(x_i - x_{i+2})^2 = x_i^2 + x_{i+2}^2 - 2x_i x_{i+2}$, 知有

$$\begin{aligned} Q(x_1, x_2, \dots, x_n) &= \sum_{i=1}^{n-2} (x_i - x_{i+2})^2 \\ &= \sum_{i=1}^{n-2} x_i^2 + \sum_{i=1}^{n-2} x_{i+2}^2 - 2 \sum_{i=1}^{n-2} x_i x_{i+2} \\ &= x_1^2 + x_2^2 + \sum_{i=3}^{n-2} 2x_i^2 + x_{n-1}^2 + \\ &\quad x_n^2 - 2 \sum_{i=1}^{n-2} x_i x_{i+2}. \end{aligned}$$

故易得

$$\begin{aligned}
 & \begin{matrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{matrix} \begin{matrix} W \\ W \\ W \end{matrix} \\
 \mathbf{A} = & \begin{matrix} W & W & W & W & W \\ & W & W & 2 & 0 & -1 \\ & & W & 0 & 1 & 0 \\ & & & -1 & 0 & 1 \end{matrix} .
 \end{aligned}$$

方法2 做可逆线性代换

$$y_i = x_i - x_{i+2}, \quad i = 1, \cdots, n-2,$$

$$y_i = x_i, \quad i = n-1, n.$$

$$\text{则有 } Q(x_1, x_2, \cdots, x_n) \Big|_{x=Py} = \mathbf{y}^T \begin{pmatrix} \mathbf{I}_{n-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{y}, \quad \text{即有}$$

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{pmatrix} \mathbf{I}_{n-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

而由 $\mathbf{y} = \mathbf{P}^{-1} \mathbf{x}$, 知

$$\begin{aligned}
 & \begin{matrix} 1 & 0 & -1 \\ & W & W & W \\ & & 1 & 0 & -1 \\ & & & 1 & 0 \\ & & & & 1 \end{matrix} \\
 \mathbf{P}^{-1} = & \begin{matrix} 1 \\ 0 & W \\ -1 & W \\ & W & 1 \\ & & 0 & 1 \\ & & & -1 & 0 & 1 \end{matrix} .
 \end{aligned}$$

故可算出

$$\begin{aligned}
 A &= P^{-1} \begin{pmatrix} I_{n-2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} P^{-1} \\
 &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} W & W \\ W & W \\ W & W \end{pmatrix} \\
 &= \begin{pmatrix} W & W & 2 & 0 & -1 \\ W & W & 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.
 \end{aligned}$$

讲评 显然方法 1 比较好. 有的同学看到题目中已配成平方和, 就顺势做了可逆线性代换, 事实上因阶数高 P^{-1} 较繁, 故方法 2 加大了计算量, 还是直接计算出系数较好.

例 2 求正交线性替换 $x = Py$, 把二次型

$$Q(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 2x_2x_3$$

化为标准形.

解 由题目要求, 要用主轴化方法. 因

$$|\lambda I - A| = \begin{vmatrix} \lambda & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} = \lambda^3 - 2\lambda - 3 = (\lambda + 1)^2(\lambda - 2),$$

故 $\lambda_1 = 2, \lambda_2 = \lambda_3 = -1$.

对 $\lambda_1 = 2$, 解 $(2I - A)x = 0$, 得 $x = \alpha(1, 1, 1)^T$. 取

$$\eta_1 = \frac{1}{3}(1, 1, 1)^T.$$

对 $\lambda_{2,3} = -1$, 解 $(I + A)x = 0$, 得 $x = \alpha(1, -1, 0)^T + \beta(1, 1, -2)^T$. 取 $\eta_2 = \frac{1}{2}(1, -1, 0)^T, \eta_3 = \frac{1}{6}(1, 1, -2)^T$.

令 $P = (\eta_1, \eta_2, \eta_3)$, 则有

$$P^T AP = \text{diag}\{2, -1, -1\},$$

即有

$$Q(x_1, x_2, x_3) \Big|_{x=Py} = 2y_1^2 - y_2^2 - y_3^2.$$

讲评 以上所做的可逆线性代换中, P 是正交矩阵, 这是题目的要求. 有的同学喜欢“初等变换法”, 以为只要化标准形都可以用该法. 应该指出的是二次型化标准形, 其平方和的系数是不惟一的 (当然, 正、负项的个数惟一), 因 P 的不同而异. 而正交线性代换得出的平方和是惟一的——由系数矩阵 A 的特征值构成. 于是, 有的同学又提出: “那我先求特征值, 然后用初等变换法化成对角线上都是 A 的特征值, 得到的 P 是否为正交矩阵呢?” 下面以本题为例说明之.

$$\begin{aligned}
 (A, I) = & \begin{array}{cccccc|cccccc}
 0 & 1 & 1 & 1 & 0 & 0 & & 2 & 1 & 2 & 1 & 1 & 0 \\
 1 & 0 & 1 & 0 & 1 & 0 & r_2 + r_2 & 1 & 0 & 1 & 0 & 1 & 0 \\
 1 & 1 & 0 & 0 & 0 & 1 & c_2 + c_1 & 2 & 1 & 0 & 0 & 0 & 1 \\
 -r_1 + r_3 & & & & & & & & & & & & \\
 -\frac{1}{2}r_1 + r_2 & & 2 & & 0 & & & 1 & & 1 & & 0 & \\
 -c_1 + c_3 & & 0 & & -\frac{1}{2} & & & 0 & & -\frac{1}{2} & & \frac{1}{2} & 0 \\
 -\frac{1}{2}c_1 + c_2 & & 0 & & 0 & & -2 & -1 & & -1 & & 1 & \\
 2r_2 & & 2 & & 0 & & 0 & 1 & & 1 & & 0 & \\
 \frac{1}{2}r_3 & & 0 & & -1 & & 0 & -\frac{1}{2} & & \frac{1}{2} & & 0 & \\
 2c_2 & & & & & & & & & & & & \\
 \frac{1}{2}c_3 & & 0 & & 0 & & -1 & -\frac{1}{2} & & -\frac{1}{2} & & \frac{1}{2} &
 \end{array}, \\
 & \begin{array}{cccccc}
 1 & -\frac{1}{2} & -\frac{1}{2} \\
 1 & \frac{1}{2} & -\frac{1}{2} \\
 0 & 0 & \frac{1}{2}
 \end{array}
 \end{aligned}$$

故得

$$P_1 = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

显见 P_1 不是正交矩阵, 它的列不是由 A 的属于相应特征值的特征向量构成. 从理论上来说, 初等变换法所得到的 P_1 满足

$$P_1^T A P_1 = D$$

是相合变换. 改写上式得 $A P_1 = P_1^{-1 T} D$; 而主轴化方法得到的 P 满足

$$P^T A P = D \quad P^{-1} A P = D,$$

是既相合又相似, 即 P 满足 $A P = P D$. 故其列是 A 的特征向量, 故只能走“求特征值, 再求相应的特征向量”这条路.

例 3 用配方法把二次型

$$Q(x_1, x_2, x_3) = 2x_1x_2 - 2x_2^2 + 2x_2x_3 + x_1x_3$$

化为平方和, 并写出所做的可逆线性替换.

解 因 $a_{11} = 0$, 而 $a_{22} = -2 \neq 0$, 故先做可逆线性代换

$$y_1 = x_2,$$

$$y_2 = x_1,$$

$$y_3 = x_3.$$

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

即

$$\mathbf{x} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{y} = P_1 \mathbf{y},$$

$$1$$

从而得

$$\begin{aligned} Q(x_1, x_2, x_3) \Big|_{\mathbf{x} = P_1 \mathbf{y}} &= -2y_1^2 + 2y_1y_2 + 2y_1y_3 + y_2y_3 \\ &= -2y_1 - \frac{1}{2}(y_2 + y_3)^2 + \frac{1}{2}y_2^2 + \frac{1}{2}y_3^2 + 2y_2y_3 \\ &= -2y_1 - \frac{1}{2}(y_2 + y_3)^2 + \frac{1}{2}(y_2 + 2y_3)^2 - \frac{3}{2}y_3^2 \\ &= -2z_1^2 + \frac{1}{2}z_2^2 - \frac{3}{2}z_3^2. \end{aligned}$$

上面最后一个等号是做了可逆线性代换

即

$$z_1 = y_1 - \frac{1}{2}y_2 - \frac{1}{2}y_3,$$

$$z_2 = y_2 + 2y_3,$$

$$z_3 = y_3.$$

$$y_1 = z_1 + \frac{1}{2}z_2 - \frac{1}{2}z_3,$$

$$y_2 = z_2 - 2z_3,$$

$$y_3 = z_3.$$

记 $y = Pz$, 则 $x = P_1 y = P_1 P z = Pz$, 故知

$$P = P_1 P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2 \\ 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix},$$

且有

$$Q(x_1, x_2, x_3) \Big|_{x=Py} = -2z_1^2 + \frac{1}{2}z_2^2 - \frac{3}{2}z_3^2.$$

讲评 此题虽满足 $a_{11} = 0$, 而 $a_{12} \neq 0$, 按一般的程序可令

$$x_1 = y_1 + y_2,$$

$$x_2 = y_1 - y_2,$$

$$x_3 = y_3.$$

使其出现 y_1^2 的系数不为 0, 但对此题, 上法行不通: 因把上述可逆线性代换代入原二次型中得

$$\begin{aligned} Q(x_1, x_2, x_3) &= 2(y_1 + y_2)(y_1 - y_2) - 2(y_1 - y_2)^2 + \\ &\quad 2(y_1 - y_2)y_3 + (y_1 + y_2)y_3 \\ &= -4y_2^2 + 4y_1y_2 + 3y_1y_3 - y_2y_3, \end{aligned}$$

y_1^2 的系数仍为 0, 这是由 a_{12} 和 a_{22} 系数的特殊关系造成的. 故遇到这种情况, 要先做 1, 2 行和 1, 2 列的对换, 使 (1, 1) 位的元素不等于 0, (即用变换使 x_1, x_2 互换位置) 再进行配方. 同理, 若 $a_{11} = 0$, 而 $a_{1j} \neq 0$, 且 $a_{jj} \neq 0$ 若 $2a_{1j} + a_{jj} = 0$, 也得先做 1, j 行(和列)的对换才行.

例 4 用配方法将二次型

$$Q(x_1, x_2, x_3, x_4) = x_1^2 + 3x_2^2 + 4x_4^2 + 4x_1x_2 - 2x_1x_4 - 2x_2x_3 - 6x_2x_4 + 2x_3x_4$$

化为标准形, 并用矩阵的相合验证结果.

解 用大块配方法, 即第一次把含 x_1 的项全拿来一起配方, 有

$$\begin{aligned} Q(x_1, x_2, x_3, x_4) &= (x_1 + 2x_2 - x_4)^2 - x_2^2 + 3x_4^2 - 2x_2x_3 - 2x_2x_4 + 2x_3x_4 \\ &= (x_1 + 2x_2 - x_4)^2 - (x_2 + x_3 + x_4)^2 + x_3^2 + 4x_4^2 + 4x_3x_4 \\ &= (x_1 + 2x_2 - x_4)^2 - (x_2 + x_3 + x_4)^2 + (x_3 + 2x_4)^2, \end{aligned}$$

故做可逆线性代换

$$\begin{aligned} y_1 &= x_1 + 2x_2 - x_4, \\ y_2 &= x_2 + x_3 + x_4, \\ y_3 &= x_3 + 2x_4, \\ y_4 &= x_4. \end{aligned}$$

也就是

$$\begin{aligned} x_1 &= y_1 - 2y_2 + 2y_3 - y_4, \\ x_2 &= y_2 - y_3 + y_4, \\ x_3 &= y_3 - 2y_4, \\ x_4 &= y_4. \end{aligned}$$

即

$$P = \begin{pmatrix} 1 & -2 & 2 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

则有 $Q(x_1, x_2, x_3, x_4) \Big|_{x=Py} = y_1^2 - y_2^2 + y_3^2$, 且易验证

$$P^T A P = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 2 & 0 & -1 \\ -2 & 1 & 0 & 0 & 2 & 3 & -1 & -3 \\ 2 & -1 & 1 & 0 & 0 & -1 & 0 & 1 \\ -1 & 1 & -2 & 1 & -1 & -3 & 1 & 4 \end{pmatrix}.$$

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

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

$$= \text{diag}\{1, -1, 1, 0\}.$$

讲评 这里用的是大块配方法, 即若 $a_{11} \neq 0$, 则第一次把含 x_1 的项全拿来一起配方, 再若 $a'_{22} \neq 0$, 则把含 y_2 的项全配方, 依此类推, 这样能保证所做的是可逆线性代换, 即 $x = Py$ 中的 P 是可逆矩阵. 这是与中学的配方不同的, 同学们一定要注意, 否则会出现错误. 下面就是一个反例: 有人用如下的配方做本题得到

$$Q(x_1, x_2, x_3, x_4) = 2(x_1 + x_2)^2 - (x_1 + x_4)^2 + (x_2 - x_3 - 3x_4)^2 - (x_3 + 2x_4)^2,$$

于是令

$$\begin{aligned}
 y_1 &= x_1 + x_2, \\
 y_2 &= x_1 \quad \quad \quad + x_4, \\
 y_3 &= \quad x_2 - x_3 - 3x_4, \\
 y_4 &= \quad \quad \quad x_3 + 2x_4.
 \end{aligned}$$

写成矩阵形式为

$$\mathbf{y} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & 2 \end{pmatrix} \mathbf{x} = \mathbf{P}^{-1} \mathbf{x}.$$

显然,以上代换的系数矩阵各行之间满足 $r_1 - r_2 - r_3 = r_4$, 故其不可逆, $\mathbf{P}^{-1}$ 只是摆样子而已, 无法求出 $\mathbf{P}$. 由此错误导致二次型的标准形误为 $2y_1^2 - y_2^2 + y_3^2 - y_4^2$, 即出现两正, 两负. 而由正确解法知它应为两正, 一负才对. 我们在 1.1 节中已说过, $\mathbf{x} = \mathbf{P}\mathbf{y}$ 相当于一个向量在不同基下的坐标, $\mathbf{P}$ 是两组基之间的过渡矩阵, 故它必须是可逆的, 由这个反例可以看出如不可逆就改变了二次型的类型. 这是绝不容许的.

例 5 设 $\mathbf{A}$ 是一个 n 阶实对称矩阵, 如果对任意 n 维向量 $\mathbf{x}$, 有 $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$, 那么 $\mathbf{A} = \mathbf{0}$.

证 方法 1 记 $\mathbf{A} = (a_{ij})_{n \times n}$, 由于对任意 $\mathbf{x}$, 有 $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$, 于是, 分别取

$$\mathbf{x} = (0, \cdots, 0, 1, 0, \cdots, 0)^T, \quad i = 1, \cdots, n,$$

得 $\mathbf{x}^T \mathbf{A} \mathbf{x} = a_{ii} = 0$, 故 $a_{ii} = 0 (i = 1, \cdots, n)$; 再分别取

$$\mathbf{x}_{ij} = (0, \cdots, 0, 1, 0, \cdots, 0, 1, 0, \cdots, 0)^T, \quad 1 \leq i < j \leq n,$$

得 $\mathbf{x}_{ij}^T \mathbf{A} \mathbf{x}_{ij} = a_{ii} + a_{jj} + 2a_{ij} = 0$, 于是得 $a_{ij} = 0 (i \neq j)$. 得证.

方法 2 先把二次型化为平方和, 再选特殊的变量, 即有

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \Big|_{\mathbf{x} = \mathbf{P}\mathbf{y}} = \mathbf{y}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{y} = d_1 y_1^2 + \cdots + d_n y_n^2.$$

由 $\mathbf{P}$ 可逆, 故对任意 $\mathbf{x}$, 有 $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ 转化为对任意 $\mathbf{y}$ 有 $\mathbf{y}^T \mathbf{B} \mathbf{y} = 0$.

分别取

$$\mathbf{y}_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T, \quad i = 1, \cdots, n,$$

得 $0 = \mathbf{y}_i^T \mathbf{B} \mathbf{y}_i = d_i$, 知 $d_i = 0 (i = 1, \cdots, n)$. 于是得 $\mathbf{B} = \mathbf{0}$, 从而有

$$\mathbf{A} = \mathbf{P}^{-1} \mathbf{P}^T \mathbf{0} \mathbf{P}^{-1} = \mathbf{0}.$$

讲评 上面两个证法都可以, 繁易相当. 这道题的本质是: 若一个二次型的值永远为 0, 那么其对应的矩阵为零矩阵. 有的同学采用如下的证法: “因对任意 $\mathbf{x}$, 有 $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$, 故 $\mathbf{A} \mathbf{x} = \mathbf{0}$. 又此式对任意 $\mathbf{x}$ 成立, 故得 $\mathbf{A} = \mathbf{0}$.” 这等于没证, 重复原命题. 不错, 从线性方程组的理论看: 若对任意 $\mathbf{x}$ 有

$$\mathbf{A} \mathbf{x} = \mathbf{0} \quad \text{或} \quad \mathbf{x}^T \mathbf{B} = \mathbf{0}$$

都可以得到 $\mathbf{A} = \mathbf{0}$ 或 $\mathbf{B} = \mathbf{0}$ (因解空间为 n 维, 故其系数矩阵的秩为 0) 也就是说分别取 $\mathbf{x}_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T (i = 1, \cdots, n)$ 代入方程中计算, 就可得 $\mathbf{A}$ 的列或 $\mathbf{B}$ 的行均为零. 但对任意 $\mathbf{x}$, 由

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \quad \text{不能推出} \quad \mathbf{A} \mathbf{x} = \mathbf{0}.$$

因当 $\mathbf{x}^T$ 变时 $\mathbf{A} \mathbf{x}$ 也在变, 故不能用线性方程组的理论. 事实上, 由证法 1 知, 当 $\mathbf{x}_i = (0, \cdots, 0, 1, 0, \cdots, 0)^T$ 时, 只能得到 $a_{ii} = 0$. 故该证法不可取.

例 6 已知二次型 $Q(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + 3x_3^2 + 2tx_2x_3$ ($t > 0$), 通过正交线性替换化为 $Q(\alpha) = 2y_1^2 + y_2^2 + 5y_3^2$.

(1) 求 t 及正交线性替换的正交矩阵;

(2) 证明在条件 $x_1^2 + x_2^2 + x_3^2 = 1$ 下, $Q(\alpha)$ 的最大值为 5.

解 (1) 由已知得: 2, 1, 5 是二次型的矩阵 $\mathbf{A}$ 的 3 个不同的特征值, 而

$$\mathbf{A} = \begin{pmatrix} 2 & & \\ & 3 & t \\ & t & 3 \end{pmatrix} = \begin{pmatrix} 2 & & \\ & \mathbf{A}_2 & \\ & & 5 \end{pmatrix}.$$

于是知 1, 5 应为 $\mathbf{A}_2$ 的特征值, 即是 $|\lambda \mathbf{I} - \mathbf{A}_2| = 0$ 的根. 由

$$|I - A| = \begin{vmatrix} -2 & -t \\ -t & -2 \end{vmatrix} = 4 - t^2 = 0,$$

得 $t = \pm 2$. 而要求 $t > 0$, 故 $t = 2$ 为所求. 由此知

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.$$

解 $(I - A)x = 0$, 得 $x = c_1(1, -1)^T$, 取 $x_1 = \frac{1}{2}(1, -1)^T$;

解 $(5I - A)x = 0$, 得 $x = c_2(1, 1)^T$, 取 $x_2 = \frac{1}{2}(1, 1)^T$.

令 $Q = [x_1, x_2]$, 则有 $Q^T A Q = \begin{pmatrix} 1 & \\ & 5 \end{pmatrix}$, 于是令

$$P = \begin{pmatrix} 1 & & \\ & 1 & 0 \\ & 0 & 1 \end{pmatrix} Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

则有

$$Q(x_1, x_2, x_3) \Big|_{x=Py} = 2y_1^2 + y_2^2 + 5y_3^2.$$

(2) 当 $x_1^2 + x_2^2 + x_3^2 = 1$ 时, 有

$$y_1^2 + y_2^2 + y_3^2 = y^T y = y^T P^T P y = x^T x = x_1^2 + x_2^2 + x_3^2 = 1,$$

故得

$$Q(\alpha) \Big|_{x=Py} = 2y_1^2 + y_2^2 + 5y_3^2 \leq 5(y_1^2 + y_2^2 + y_3^2) = 5,$$

等号当且仅当 $y_1 = y_2 = 0, y_3 = \pm 1$ 时成立, 即

$$x = Py = \pm \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}^T$$

时二次型达到最大值 5.

讲评 在问题(1)中我们采用了一些小技巧, ①没有去计算 A 的特征多项式再定 t 的值, 而是充分利用 A 中元素的特点, 大大简

化了计算过程;②也没有去算 A 的特征向量,通过算 A 的特征向量,找出了正交矩阵 P .另外,通过证明本题的结论(2),读者可以思考这样两件事:①二次型的最小值是什么?在 x 为何变量时达到此最小值;②对任意二次型是否都有这样的结论: A 的最小,最大特征值分别为二次型的最小、最大值?如果这样去思考,并去解决了,那么做一道题就等于做了 3 道题,对问题的理解也就深刻了.

例 7 设 A, B 均为 n 阶实对称矩阵,其中 A 正定,证明:当实数 t 充分大时, $tA + B$ 亦正定.

证 首先,对任意 t , 均有 $(tA + B)^T = tA + B$, 故 $tA + B$ 为实对称矩阵. 又由 A 正定, 故存在可逆矩阵 P , 使

$$P^T A P = I, \quad P^T B P = D = \text{diag}\{d_1, \cdots, d_n\}.$$

于是有

$P^T (tA + B) P = tP^T A P + P^T B P = \text{diag}\{t + d_1, \cdots, t + d_n\}$, 所以当 $t > \max |d_i|$ 时, $t + d_i > 0 (i = 1, \cdots, n)$, 从而有 $tA + B$ 正定.

讲评 此证法主要是用定理 5, 即 A, B 同时相合于对角阵. 由证明过程, 还可以得到当 $t < -\max |d_i|$ 时, $t + d_i < 0 (i = 1, \cdots, n)$, 于是得 t 充分小时, $tA + B$ 负定. 又当 $A = I$ 时, 设 B 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则当 $t < -\max |\lambda_i|$ 时, $tI + B$ 负定.

例 8 设 A, B 均为 n 阶正定实对称矩阵, 且 $AB = BA$, 证明 AB 也是正定实对称矩阵.

证明 方法 1 由 $AB = BA$, 得 $(AB)^T = B^T A^T = BA = AB$, 故知 AB 是实对称矩阵. 又由定理 6, 知 A, B 可同时正交相似于对角矩阵, 即存在正交矩阵 Q , 使

$$\begin{aligned} Q^T A Q &= \text{diag}\{\lambda_1, \cdots, \lambda_n\}, & \text{其中 } \lambda_i > 0, \\ Q^T B Q &= \text{diag}\{\mu_1, \cdots, \mu_n\}, & \text{其中 } \mu_i > 0. \end{aligned}$$

于是有

$$Q^T A B Q = Q^T A Q Q^T B Q = \text{diag}\{\lambda_1 \mu_1, \cdots, \lambda_n \mu_n\},$$

且有 $\lambda_i \mu_i > 0 (i = 1, \cdots, n)$, 故 AB 的特征值均正, 所以 AB 正定.

方法 2 设 λ 为 AB 的特征值, x 为相应的特征向量, 即有

$$ABx = \lambda x.$$

左乘 A^{-1} 得

$$Bx = \lambda A^{-1} x;$$

再左乘 x^T , 有

$$x^T Bx = \lambda x^T A^{-1} x. \quad (1)$$

由 $A > 0$, 知 $A^{-1} > 0$, 又 $B > 0$, 故对任意 $x \neq 0$, 均有

$$x^T Bx > 0, \quad x^T A^{-1} x > 0.$$

代入(1)式得 $\lambda > 0$, 又 $(AB)^T = AB$, 即 AB 为对称矩阵, 故知 AB 为正定实对称矩阵.

方法 3 AB 为实对称矩阵同前证明. 以下用两个思路证 AB 的特征根均正.

思路 1 由 $A > 0, B > 0$, 知存在可逆矩阵 P, Q 使

$$A = P^T P, \quad B = Q^T Q.$$

于是有

$$AB = P^T P Q^T Q = Q^{-1} Q P^T P Q^T Q = Q^{-1} Q P^T P Q^T Q,$$

也就是说 AB 相似于 $QP^T PQ^T$, 而后者可写为

$$QP^T PQ^T = PQ^{T^T} PQ^T.$$

又 PQ^T 为可逆矩阵, 故 $QP^T PQ^T$ 是正定实对称矩阵, 其特征值皆大于 0. 因相似的矩阵有相同的特征值, 所以知 AB 的特征值全为正, 故 $AB > 0$.

思路 2 由于对任意 n 阶矩阵 M, N 有

$$\det \lambda I_n - MN = \det \lambda I_n - NM,$$

(用分块矩阵的初等变换可证, 类似于 $\det(I_m - AB) = \det(I_n - BA)$ 的证明) 这里取 $M = P^T, N = PQ^T Q$, 即有

$$|\lambda I_n - AB| = |\lambda I_n - P^T PQ^T Q| = |\lambda I_n - PQ^T QP^T|,$$

故知 AB 与 $PQ^T QP^T = QP^{T^T} QP^T$ 的特征值相同. 而后者为正定,

其特征值全为正,故 AB 为正定实对称矩阵.

讲评 (1) 此题中证 AB 为实对称矩阵的步骤不能少,这由条件 $AB=BA$ 可以充分地证出,有的同学没用上这个条件就得出结论了,肯定哪一步出了问题.

(2) 若没有 $AB=BA$ 的条件,不能保证 AB 为对称矩阵,从而也谈不上正定.如

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix},$$

显然有 $A > 0, B > 0$, 但 $AB = \begin{pmatrix} 1 & 1 \\ 2 & 8 \end{pmatrix}$, 不是对称矩阵,故非正定矩阵.

(3) 由证法 2, 3 知有结论:“若 $A > 0, B > 0$, 则 AB 的特征值必大于 0.”

例 9 设 A 为 n 阶实对称矩阵,且 $\det A < 0$, 证明存在实非零向量 x_0 , 使 $x_0^T A x_0 < 0$.

证 方法 1 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为 A 的特征值,则由

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n < 0,$$

(λ_i 为实数) 知必有某个 $\lambda_i < 0$. 不妨设 $\lambda_n < 0$, 且取 x_0 为 A 的属于特征值 λ_n 的单位化特征向量, 则有

$$x_0^T A x_0 = x_0^T \lambda_n x_0 = \lambda_n x_0^T x_0 = \lambda_n < 0.$$

方法 2 (反证法) 若对所有的实非零向量 x , 都有 $x^T A x \geq 0$, 则 A 是半正定的. 于是知 A 的所有主子式 ≥ 0 , 与已知 $\det A < 0$ 矛盾, 故必存在 x_0 , 使 $x_0^T A x_0 < 0$.

方法 3 因 $\det A < 0$, 故 A 可逆, 即 $\text{rank } A = n$. 由定理 4, 知存在可逆矩阵 P , 使得

$$P^T A P = \begin{pmatrix} I_r & & \\ & -I_{n-r} & \\ & & \end{pmatrix}.$$

若 $r = n$, 则 $A > 0$, 于是得 $\det A > 0$, 与已知 $|A| < 0$ 矛盾, 故必 $r <$

n . 令 $\mathbf{y}_0 = (0, \cdots, 0, 1)^T$, 即

取 $\mathbf{x}_0 = \mathbf{P}\mathbf{y}_0 = (p_{1n}, \cdots, p_{nn})^T$, 则有

$$\mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 = \mathbf{y}_0^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{y}_0 = \mathbf{y}_0^T \begin{matrix} I_r \\ -I_{n-r} \end{matrix} \mathbf{y}_0 = -1 < 0.$$

讲评 由已知条件 $\mathbf{A}$ 为实对称矩阵, 且 $|\mathbf{A}| < 0$, 可以得到(或联想到)以下信息:

- (1) 由 $\det \mathbf{A} = \prod_{i=1}^n \lambda_i < 0$, 知必有 i 使 $\lambda_i < 0$ (因 λ_i 均为实数);
- (2) $\mathbf{A}$ 非正定 (因 $|\mathbf{A}| < 0$), 故不可能对所有的非零 $\mathbf{x}$, $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$;
- (3) $r(\mathbf{A}) = n$, 故 $\mathbf{A}$ 的规范型中正, 负项数之和为 n (即对角线上元素均非 0).

把这些思路理清, 写完全就分别得到上述的 3 个方法.

例 10 设 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 是一个 n 元实二次型, 若有 n 维实向量 $\mathbf{x}_1$ 和 $\mathbf{x}_2$, 使得 $\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 > 0$, $\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 < 0$, 试证明存在 n 维实向量 $\mathbf{x}_0$, 使得 $\mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 = 0$.

证 方法 1 由已知存在向量 $\mathbf{x}_1, \mathbf{x}_2$ 使 $\mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 > 0$, $\mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 < 0$, 知二次型的规范型中既不可能全是正项, 也不可能全是负项, 故存在可逆线性代换 $\mathbf{x} = \mathbf{P}\mathbf{y}$, 使

$$\mathbf{x}^T \mathbf{A} \mathbf{x} \Big|_{\mathbf{x} = \mathbf{P}\mathbf{y}} = y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2,$$

且 $r \neq 0, s \neq 0$. 令 $\mathbf{y}_0 = (1, 0, \cdots, 0, 1, 0, \cdots, 0)^T$, 则

$$\mathbf{x}_0 = \mathbf{P}\mathbf{y}_0 = (p_{11} + p_{1r+1}, \cdots, p_{n1} + p_{nr+1})^T \neq \mathbf{0}$$

(因 $\mathbf{P}$ 是可逆矩阵, 它的第 1 列和第 $r+1$ 列不成比例), 且有

$$\mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 = \mathbf{P}\mathbf{y}_0^T \mathbf{A} \mathbf{P}\mathbf{y}_0 = \mathbf{y}_0^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{y}_0 = 0.$$

方法 2 由于对任意 $\lambda \in \mathbb{R}$ 有

$$\mathbf{x}_1 + \lambda \mathbf{x}_2^T \mathbf{A} (\mathbf{x}_1 + \lambda \mathbf{x}_2) = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + 2\lambda \mathbf{x}_1^T \mathbf{A} \mathbf{x}_2 + \lambda^2 \mathbf{x}_2^T \mathbf{A} \mathbf{x}_2. \quad (2)$$

显然, 右边是 λ 的二次三项式, 其判别式是

$$\Delta = 4 (\mathbf{x}_1^T \mathbf{A} \mathbf{x}_2)^2 - 4 \mathbf{x}_2^T \mathbf{A} \mathbf{x}_2 \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 > 0,$$

所以,该二次三项式有两个不等的实根 λ_1, λ_2 , 由此知

$$\mathbf{x}_1 + \lambda_1 \mathbf{x}_2 \neq \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \quad (\text{注意 } \mathbf{x} \neq \mathbf{0}).$$

取 $\mathbf{x}_0$ 为它们中不等于 $\mathbf{0}$ 的一个向量, 如取 $\mathbf{x}_0 = \mathbf{x}_1 + \lambda_1 \mathbf{x}_2$, 则由(2)式知 $\mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 = 0$.

方法3 显见二次型 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 是 n 个变数 $x_1, x_2, \dots, x_n$ 的连续函数, 又在“两点”上函数值异号 因 $\frac{\mathbf{x}_1^T}{\|\mathbf{x}_1\|} \mathbf{A} \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} > 0$, $\frac{\mathbf{x}_2^T}{\|\mathbf{x}_2\|} \mathbf{A} \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} < 0$, 而 $\|\mathbf{x}\| = 1 = \mathbf{x}^T \mathbf{x} = x_1^2 + x_2^2 + \dots + x_n^2$ 是一个 n 维闭球, 故由中值定理知, 必有 $\mathbf{x}_0$ 且 $\|\mathbf{x}_0\| = 1$, 使 $\mathbf{x}_0^T \mathbf{A} \mathbf{x}_0 = 0$.

讲评 由二次型值可正, 可负, 知其不可能是正定、半正定、负定、半负定二次型. 故方法1是容易想到的. 方法2是比较巧的, 主要用到二次三项式的知识. 方法3是用到微积分的结论: “闭区间上的连续函数 $f(x)$, 若有 $f(a) > 0, f(b) < 0$, 则存在 $c \in [a, b]$, 使 $f(c) = 0$.” 是这一结论在 n 维空间的推广.

例11 设 $\mathbf{A}$ 为 n 阶实对称矩阵, 若 $\mathbf{A}^2 = \mathbf{I}$, 证明: $\mathbf{A} + \mathbf{I}$ 为半正定矩阵或正定矩阵.

证 **方法1** 设 λ 为 $\mathbf{A}$ 的特征值, 则由 $\mathbf{A}^2 = \mathbf{I}$, 可知 $\lambda^2 = 1$, 故 $\lambda = \pm 1$ (因 $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$, 所以 $\mathbf{x} = \mathbf{A}^2 \mathbf{x} = \mathbf{A} \lambda \mathbf{x} = \lambda^2 \mathbf{x} \quad \lambda^2 = 1$). 又由 $\mathbf{A}$ 为实对称矩阵, 故可正交相合于对角矩阵, 即存在正交矩阵 $\mathbf{P}$, 使

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \text{diag}\{1, \dots, 1, -1, \dots, -1\}.$$

于是有

$$\mathbf{P}^T (\mathbf{A} + \mathbf{I}) \mathbf{P} = \mathbf{P}^T \mathbf{A} \mathbf{P} + \mathbf{P}^T \mathbf{P} = \text{diag}\{2, \dots, 2, 0, \dots, 0\}.$$

当 $r < n$ 时, $\mathbf{A} + \mathbf{I}$ 为半正定矩阵; 当 $r = n$ 时, $\mathbf{A} + \mathbf{I}$ 为正定矩阵 (其中 r 为 $\mathbf{A}$ 的正惯性指数).

方法2 因 $(\mathbf{A} + \mathbf{I})(\mathbf{A} + \mathbf{I}) = \mathbf{A}^2 + 2\mathbf{A} + \mathbf{I} = 2(\mathbf{A} + \mathbf{I})$, 且 $(\mathbf{A} + \mathbf{I})^T = \mathbf{A}^T + \mathbf{I} = \mathbf{A} + \mathbf{I}$, 故有

$$(\mathbf{A} + \mathbf{I}) = \frac{1}{2} (\mathbf{A} + \mathbf{I})^T \frac{1}{2} (\mathbf{A} + \mathbf{I}) = \mathbf{Q}^T \mathbf{Q},$$

于是得 $(A + I)$ 为半正定或正定矩阵, 当 $\det Q \neq 0$, 即 $\det(A + I) \neq 0$ 时, $(A + I)$ 为正定矩阵.

讲评 方法 1 主要有两个要点: (1) A 是实对称矩阵, 故 A 可正交相合(相似)于对角矩阵, 对角矩阵的对角元由 A 的特征值组成; 那么接下来的问题就是 A 的特征值是什么? (2) 由 $A^2 = I$ 知 $\lambda^2 = 1$, 可以推广证明, 设 $Ax = \lambda x$, 则对任意 $f(x) \in F[x]$, 有

$$f(A)x = f(\lambda)x.$$

方法 2 用的是实对称矩阵 M 半正定(正定)的充要条件是存在(可逆)矩阵 Q , 使 $M = Q^T Q$. 若不用这个定理, 直接用定义证明也行, 即对任意 $x \neq 0$, 有

$$x^T (A + I) x = x^T \frac{1}{2} (A + I)^T \frac{1}{2} (A + I) x = y^T y \geq 0.$$

例 12 设 A 为 n 阶实对称正定矩阵, 如果 $A - I$ 为正定矩阵, 证明: $I - A^{-1}$ 亦为正定矩阵.

证 由 $A > 0$, 知 A 的特征值 $\lambda_i > 0 (i = 1, \cdots, n)$, 且存在正交矩阵 P , 使得

$$P^T A P = \text{diag}\{\lambda_1, \cdots, \lambda_n\}.$$

于是有

$$\begin{aligned} P^T (A - I) P &= P^{-1} (A - I) P = P^{-1} A P - I \\ &= \text{diag}\{\lambda_1 - 1, \cdots, \lambda_n - 1\}. \end{aligned}$$

因 $A - I > 0$, 故 $\lambda_i - 1 > 0$, 即 $\lambda_i > 1 (i = 1, \cdots, n)$. 又 A^{-1} 的特征值为 $\frac{1}{\lambda_i}$, 由上得 $\frac{1}{\lambda_i} < 1 (i = 1, \cdots, n)$. 而显然有

$$P^{-1} (I - A^{-1}) P = I - P^{-1} A^{-1} P = \text{diag} \left\{ 1 - \frac{1}{\lambda_1}, \cdots, 1 - \frac{1}{\lambda_n} \right\}.$$

由 $1 - \frac{1}{\lambda_i} > 0$, 得 $I - A^{-1}$ 必为正定矩阵.

讲评 (1) 此题主要用到 A 是正定矩阵, 则 A 的特征值 $\lambda_i > 0$,

A^{-1} 也是正定矩阵, 且 A^{-1} 的特征值为 $\frac{1}{\lambda}$ 由 $P^{-1}AP = \text{diag}\{\lambda_1, \dots, \lambda_n\}$, 得 $P^{-1}A^{-1}P = \text{diag}\{\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\}$. (2) 此题可推广为把 I 换为任意正定实对称矩阵 B , 即: 设 A, B 均为 n 阶正定实对称矩阵, 若 $A - B$ 为正定矩阵, 则 $B^{-1} - A^{-1}$ 亦为正定矩阵. 其证明过程类似, 叙述如下:

因 B 为正定矩阵, 故由定理 5, 存在可逆矩阵 P , 使得

$$P^T B P = I, \quad P^T A P = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}. \quad (1)$$

由 $A > 0$, 知 $d_i > 0 (i = 1, \dots, n)$. 又

$$P^T (A - B) P = \text{diag}\{d_1 - 1, \dots, d_n - 1\},$$

再因 $A - B > 0$ 知, 应有 $d_i - 1 > 0$, 即 $d_i > 1 (i = 1, \dots, n)$.

由(1)式可得

$$P^{-1} B^{-1} P^{-1T} = I, \quad P^{-1} A^{-1} P^{-1T} = \begin{pmatrix} \frac{1}{d_1} & & \\ & \ddots & \\ & & \frac{1}{d_n} \end{pmatrix}.$$

令 $Q = P^{-1T}$, 则有

$$Q^T (B^{-1} - A^{-1}) Q = \text{diag}\{1 - \frac{1}{d_1}, \dots, 1 - \frac{1}{d_n}\}.$$

因 $d_i > 1$, 故 $\frac{1}{d_i} < 1$, 即 $1 - \frac{1}{d_i} > 0 (i = 1, \dots, n)$. 于是知 $B^{-1} - A^{-1}$ 为正定矩阵(其正惯性指数为 n).

注 推广的命题与原习题的区别,前者用的是正交矩阵,因 I 与 A 同时相合于对角阵.要保持 I 不变,必须用正交矩阵;后者是 A, B 同时相合于对角矩阵(要把 B 化为 I ,必用可逆矩阵)用的是可逆矩阵.

例 13 设 A 是 $m \times n$ 阶实矩阵, $m < n$, 证明: AA^T 为正定矩阵的充要条件是 $\text{rank}(A) = m$.

证 (充分性)由已知 AA^T 是 m 阶实矩阵,且有

$$AA^{TT} = AA^T,$$

故 AA^T 是实对称矩阵.

又对任意 $x \neq 0$, $A^T x \neq 0$ (因 $r(A) = r(A^T) = m$, 故 A^T 的列向量线性无关, 所以不存在不全为 0 的数 $x_1, x_2, \dots, x_n$ 使 $x_1 a_1 + x_2 a_2 + \dots + x_n a_n = 0$, 其中 a_i 是 A^T 的第 i 列), 故有

$$x^T AA^T x = A^T x^T A^T x > 0 \quad (1)$$

即知 AA^T 是正定矩阵.

(必要性)由 AA^T 是正定矩阵, 知对任意 $x (\neq 0) \in \mathbb{R}^n$ 均有

$$x^T (AA^T) x = (A^T x)^T (A^T x) > 0. \quad (2)$$

此时 $A^T x \neq 0$ (因若 $A^T x = 0$, 则上式值为 0), 即 $A^T x = 0$ 只有零解, 故 $r(A^T) = m \rightarrow r(A) = m$.

讲评 此题的本质是, 对任意 $x \in \mathbb{R}^n$, 均有

$$x^T (AA^T) x = (A^T x)^T (A^T x) \geq 0,$$

且

等号成立 $x = 0$ (看(1)式, 也即等价于 $AA^T > 0$)

$A^T x = 0$ (看(2)式, 是向量 $A^T x$ 的模平方),

故 $AA^T > 0 \iff A^T x = 0 \iff x = 0 \iff r(A^T) = m$.

例 14 设 A 为 n 阶实对称正定矩阵, 证明存在上三角矩阵 R , 使

$$A = R^T R.$$

证 方法 1 因 $A > 0$, 故存在可逆矩阵 P , 使 $A = P^T P$. 又由于

对任意可逆矩阵都有 QR 分解, 即存在正交矩阵 Q 和上三角矩阵 R , 使得 $P = QR$. 于是得

$$A = P^T P = (QR)^T (QR) = R^T R \quad (\text{因 } Q^T Q = I_n).$$

方法 2(归纳法) 对 A 的阶数用归纳法. 当 $n = 2$ 时, 设

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \text{其中 } a_{12} = a_{21},$$

则有

$$\begin{pmatrix} 1 & 0 \\ -\frac{a_{12}}{a_{11}} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} - \frac{a_{12} a_{21}}{a_{11}} \end{pmatrix},$$

故得

$$A = \begin{pmatrix} 1 & 0 \\ \frac{a_{12}}{a_{11}} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} - \frac{a_{12}^2}{a_{11}} \end{pmatrix} = R^T R,$$

其中

$$R = \begin{pmatrix} a_{11} & 0 \\ \frac{a_{12} a_{22} - a_{12}^2}{a_{11}} & 1 \end{pmatrix} \quad \text{是上三角矩阵.}$$

假设对 $n - 1$ 阶矩阵结论成立, 则当 A 为 n 阶矩阵时, 设

$$A = \begin{pmatrix} A_{n-1} & \beta \\ \beta^T & a_{nn} \end{pmatrix},$$

其中 A_{n-1} 为 $n - 1$ 阶正定实对称矩阵. 由归纳假设, 存在上三角矩阵 R_{n-1} , 使

$$A_{n-1} = R_{n-1}^T R_{n-1}.$$

取

$$P = \begin{pmatrix} I_{n-1} & -A_{n-1}^{-1} \beta \\ 0 & 1 \end{pmatrix},$$

则

$$P^T A P = \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - \beta^T A_{n-1}^{-1} \beta \end{pmatrix}.$$

记 $b = a_{nn} - \beta^T A_{n-1}^{-1} \beta$, 则 $b > 0$, 且有

$$\begin{aligned} A &= P^{-1} \begin{pmatrix} R_{n-1}^T R_{n-1} & 0 \\ 0 & b \end{pmatrix} P^{-1} \\ &= P^{-1} \begin{pmatrix} R_{n-1}^T & 0 & R_{n-1} & 0 \\ 0 & b & 0 & b \end{pmatrix} P^{-1} = R^T R, \end{aligned}$$

其中

$$R = \begin{pmatrix} R_{n-1} & 0 \\ 0 & b \end{pmatrix} P^{-1} = \begin{pmatrix} R_{n-1} & R_{n-1} A_{n-1}^{-1} \beta \\ 0 & b \end{pmatrix}$$

是上三角矩阵.

讲评 从理论上来说方法 1 很漂亮, 它是两个定理的联合使用. 但正因为它用了两个定理的结果, 而忽视了它们丰富的内涵. 在这里不利于理解这个命题的本质, 也不易具体计算: 给一个具体的正定实对称矩阵, 怎么去找 R , 使 $A = R^T R$. 方法 2 虽然书写较长, 但恰恰可以给人以启迪从本质上去理解这个命题, 且给出计算方法. 进一步的分析见 1.3 节例 1.

例 15 设 $\alpha, \alpha, \dots, \alpha_m$ 为 n 维向量, 定义 $(\alpha_i, \alpha_j) = \alpha_i^T \alpha_j$, 证明矩阵

$$G = \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_m) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_m) \\ \cdots & \cdots & \cdots & \cdots \\ (\alpha_m, \alpha_1) & (\alpha_m, \alpha_2) & \cdots & (\alpha_m, \alpha_m) \end{pmatrix}$$

为正定实对称矩阵的充分必要条件是 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性无关.

证 方法 1 由已知条件中 $(\alpha_i, \alpha_j) = \alpha_i^T \alpha_j$, 知这里 $\alpha_i \in R^n$. 记

$$A^T = (\alpha_1, \cdots, \alpha_m), \quad \text{即 } A = \begin{matrix} & \alpha_1^T \\ \cdots & \\ & \alpha_m^T \end{matrix},$$

则有 $G = AA^T$, 于是由例 13 知

$$G > 0 \quad r(A) = r(A^T) = m \quad \alpha_1, \cdots, \alpha_m \text{ 线性无关}.$$

方法 2 (必要性) 用反证法. 当 $G > 0$ 时, 有 $\det G > 0$. 此时若有 $\alpha_1, \cdots, \alpha_m$ 线性相关, 则其中必有某个 $\alpha_i (1 \leq i \leq m)$ 可由其余 $m-1$ 个向量线性表出, 于是由内积的线性性, 知 G 的第 i 行也可以由其余 $m-1$ 行线性表出, 即 G 的行线性相关, 故 $\det G = 0$, 这与 $\det G > 0$ 矛盾, 故必有 $\alpha_1, \cdots, \alpha_m$ 线性无关.

(充分性) 当 $\alpha_1, \cdots, \alpha_m$ 线性无关时, 把它扩充为 $V = \mathbb{R}^n$ 的一组基

$$\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n,$$

则由基向量的内积排成的矩阵

$$M = \begin{matrix} (\alpha_1, \alpha_1) & \cdots & (\alpha_1, \alpha_n) \\ \cdots & & \cdots \\ (\alpha_n, \alpha_1) & \cdots & (\alpha_n, \alpha_n) \end{matrix} = \begin{matrix} G & D \\ D & B \end{matrix}$$

是内积的度量矩阵(在基 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$ 下), 故它是正定矩阵, 于是 M 的各阶顺序主子式均正, 从而知矩阵 G 的各阶顺序主子式均大于 0, 故知 $G > 0$.

讲评 (1) 因本题的已知条件中给出了 $\alpha_i \in \mathbb{R}^n$, 故很容易将 G 写成 $G = A^T A$ 的形式并联想到例 13. 利用 13 题的结论或仿 13 题的证明过程都很容易得出结论, 所以方法 1 是指用这些内容.

(2) 方法 2 也适合于 $\alpha_i \in V_n(F)$, 而 V 是任意欧几里得空间的情况. 因其没用到内积的表达式, 只是用到定义. 故把条件 $(\alpha_i, \alpha_j) = \alpha_i^T \alpha_j$ 去掉结论仍成立.

(3) 若把内积的定义改为 $(\alpha_i, \alpha_j) = \alpha_i^T A \alpha_j$ (即 A 是内积在自然基下的度量矩阵, 则 A 是正定矩阵), 且记 $B = (\alpha_1, \alpha_2, \cdots, \alpha_m)$, 则有

$$G = \begin{pmatrix} \alpha_1^T \\ \cdots \\ A(\alpha_1, \alpha_2, \cdots, \alpha_m) \\ \alpha_m^T \end{pmatrix} = B^T A B.$$

于是结论变为

$G = B^T A B$ 为正定矩阵 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性无关 $r(B) = m$.

也就是说有: 当 A 为正定实对称矩阵时, 有

$B^T A B$ 为正定矩阵 是 $r(B) = m$.

这是 1999 年考研题中数学试卷(一)中一道 6 分题. 由上它至少有两种解法.

例 16 设 A 为 n 阶实对称正定矩阵, x 为任意 n 维向量, 证明:

$$\begin{vmatrix} 0 & x^T \\ x & A \end{vmatrix}$$

为负定二次型.

证 取 $P = \begin{pmatrix} 1 & -x^T A^{-1} \\ 0 & I_n \end{pmatrix}$, 则有

$$\begin{pmatrix} 1 & -x^T A^{-1} & 0 & x^T \\ 0 & I_n & x & A \end{pmatrix} = \begin{pmatrix} -x^T A^{-1} x & 0 \\ x & A \end{pmatrix}.$$

上式等号两边取行列式, 因 $\det P = 1$, 故得

$$Q(x_1, x_2, \cdots, x_n) = \begin{vmatrix} 0 & x^T \\ x & A \end{vmatrix} = - (x^T A^{-1} x) |A| < 0$$

对任意 $x \neq 0$ 成立(由 A 为正定矩阵知 A^{-1} 也是正定矩阵, $|A| > 0$, $x^T A^{-1} x > 0$), 故 $Q(x_1, x_2, \cdots, x_n)$ 是负定二次型.

讲评 (1) 二次型 $Q(x_1, x_2, \cdots, x_n)$ 由一个矩阵的行列式的形式给出, 这是一个分块矩阵, 这个二次型的对应实对称矩阵是什么要经过计算才知道, 由以上过程知该矩阵为 $-|A| A^{-1}$, 所以千万不能不动脑筋, 上来就认为 A 是它的对应矩阵.

(2) 若 A 和 0 对换位置, 则结论仍成立, 即

$$f(x_1, x_2, \dots, x_n) = \begin{vmatrix} \mathbf{A} & \mathbf{x} \\ \mathbf{x}^T & 0 \end{vmatrix} < 0.$$

(3) 关于解题思路,在看清已知的二次型是用一个分块矩阵的行列式给出时,要算出这个行列式,就要联想到行列式的计算方法,用初等变换化为上(下)三角矩阵、准对角矩阵都可以容易算出行列式值.这里是化为准下三角矩阵.

例 17 证明定理 5.

证 由 $\mathbf{A} > 0$, 故存在可逆矩阵 $\mathbf{P}$, 使 $\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{I}_n$, 记 $\mathbf{B} = \mathbf{P}^T \mathbf{B} \mathbf{P}$, 则因 $\mathbf{B}$ 仍是实对称矩阵, 所以必存在正交矩阵 $\mathbf{Q}$, 使得

$$\mathbf{Q}^T \mathbf{B} \mathbf{Q} = \text{diag}\{\mu_1, \mu_2, \dots, \mu_n\} = \mathbf{D},$$

故取 $\mathbf{P} = \mathbf{P} \mathbf{Q}$, 则有

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{Q}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{Q} = \mathbf{Q}^T \mathbf{I}_n \mathbf{Q} = \mathbf{I}_n,$$

$$\mathbf{P}^T \mathbf{B} \mathbf{P} = \mathbf{Q}^T \mathbf{B} \mathbf{Q} = \mathbf{D}.$$

讲评 这个定理的证明过程是可“操作”的,即给出具体的 $\mathbf{A}$, $\mathbf{B}$ 只要满足题设条件,就可按上面的“操作”过程同时相合于对角矩阵.

例 18 设矩阵

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

试求可逆矩阵 $\mathbf{P}$, 使得 $\mathbf{P}^T \mathbf{A} \mathbf{P}$, $\mathbf{P}^T \mathbf{B} \mathbf{P}$ 同时为对角矩阵.

解 观察 $\mathbf{A}$ 的各阶顺序主子式有

$$d_1 = 1 > 0, \quad d_2 = 1 > 0, \quad d_3 = 1 > 0,$$

故 $\mathbf{A}$ 是正定矩阵 (于是由定理 5, 知 $\mathbf{A}$, $\mathbf{B}$ 可同时相合于对角矩阵).

对 $(\mathbf{A}, \mathbf{I})$ 做初等行变换, 得

$$(\mathbf{A}, \mathbf{I}) = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{array}{rcl}
 -r_1 + r_2 & 1 & 0 & 0 & 1 & 0 & 0 \\
 -r_1 + r_3 & & & & & & \\
 -c_1 + c_i & 0 & 1 & 0 & -1 & 1 & 0 \\
 i = 2, 3 & 0 & 0 & 1 & -1 & 1 & 1
 \end{array},$$

所以

$$\mathbf{P} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ 且 } \mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{I}.$$

而

$$\mathbf{B} = \mathbf{P}^T \mathbf{B} \mathbf{P} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

由 $|\lambda \mathbf{I} - \mathbf{B}| = (\lambda - 2)(\lambda^2 - 1)$, 得 $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$.

对 $\lambda_1 = 1$, 解 $(\lambda \mathbf{I} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 得单位化特征向量 $\eta_1 = \frac{1}{2}(1, 0, 1)^T$; 对 $\lambda_2 = 2$, 解 $(\lambda \mathbf{I} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 得单位化特征向量 $\eta_2 = (0, 1, 0)^T$; 对 $\lambda_3 = -1$, 解 $(\lambda \mathbf{I} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 得单位化特征向量 $\eta_3 = \frac{1}{2}(1, 0, -1)^T$. 取 $\mathbf{Q} = (\eta_1, \eta_2, \eta_3)$, 则有

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}\{1, 2, -1\}.$$

于是令

$$\mathbf{P} = \mathbf{P} \mathbf{Q} = \begin{pmatrix} 0 & -1 & 2 \\ 0 & 1 & 0 \\ \frac{2}{2} & 0 & -\frac{2}{2} \end{pmatrix},$$

则有

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{I}, \quad \mathbf{P}^T \mathbf{B} \mathbf{P} = \begin{pmatrix} 1 & & \\ & 2 & \\ & & -1 \end{pmatrix}.$$

讲评 (1) 以上在把 B 化为对角矩阵时,用的是正交矩阵,这样才能保证已化好的 I 不变.若化 B 时,用可逆矩阵 P_2 ,则 $P_2^T (P_1^T A P_1) P_2 = P_2^T P_2$ 不一定是单位矩阵,也不一定是对角矩阵,但此题可以把 A 化为对角矩阵.因 B 对应的二次型为

$$Q(x_1, x_2, x_3) = 2x_1x_3 + 2x_2^2,$$

故令

$$x_1 = y_1 + y_3,$$

$$x_2 = y_2,$$

$$x_3 = y_1 - y_3.$$

即

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = P_2 y,$$

则 P_2 是可逆矩阵,且有

$$P_2^T B P_2 = \text{diag}\{2, 2, -2\}.$$

故记 $P = P_1 P_2$, 则有

$$P^T A P = \text{diag}\{2, 1, 2\},$$

$$P^T B P = \text{diag}\{2, 2, -2\}.$$

(2) 同学问的第2个问题是各特征向量不单位化行不行?显然,这时,令 $Q = (\eta_1, \eta_2, \eta_3)$ 不是正交矩阵,故不满足 $Q^T Q = I$.但因其各列都是 A 的特征向量,所以可以保证 $Q^T B Q$ 为对角矩阵.而此时 $Q^T Q$ 也是对角矩阵.如此题中,若取 $\xi_1 = (1, 0, 1)^T$, $\xi_2 = (0, 1, 0)^T$, $\xi_3 = (1, 0, -1)^T$, 且令 $P = (\xi_1, \xi_2, \xi_3)$, 则有

$$P^T A P = P_2^T I P_2 = \text{diag}\{2, 1, 2\},$$

$$P^T B P = P_2^T B P_2 = \text{diag}\{1, 2, -1\}.$$

例 19 证明定理 6.

证 因 A 为实对称矩阵,所以存在 n 阶正交矩阵 T , 使

$$T^{-1} A T = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_s I_{n_s} \end{pmatrix}, \quad \lambda_i \neq \lambda_j (i \neq j).$$

由 $AB = BA$, 得

$$T^{-1}AT \quad T^{-1}BT = T^{-1}BT \quad T^{-1}AT. \quad (*)$$

记 $T^{-1}BT = (B_{ij})$, 则由 $(*)$ 式知有

$$\lambda_i B_{ij} = B_{ij} \lambda_j.$$

因当 $i \neq j$ 时, $\lambda_i \neq \lambda_j$, 故 $B_{ij} = 0$ ($i \neq j$), 即有

$$T^{-1}BT = \begin{pmatrix} B_{11} & & \\ & \ddots & \\ & & B_{ss} \end{pmatrix},$$

其中 B_{ii} 为 n_i 阶矩阵. 又由 $T^{-1}BT$ 仍为实对称矩阵 (因 $T^{-1} = T^T$), 知 B_{ii} 均为实对称矩阵 ($i = 1, \dots, s$), 于是对每个 B_{ii} 存在 n_i 阶正交矩阵 Q_i 使

$$Q_i^{-1} B_{ii} Q_i = \text{diag}\{\mu_{i1}, \dots, \mu_{in_i}\}, \quad i = 1, \dots, s.$$

取

$$Q = \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_s \end{pmatrix}, \quad P = TQ,$$

则 Q 为正交矩阵, P 亦为正交矩阵, 且有

$$\begin{aligned} P^{-1}AP &= Q^{-1} \text{diag}\{\lambda_1 I_{n_1}, \dots, \lambda_s I_{n_s}\} Q \\ &= \text{diag} \lambda_1 Q^{-1} Q, \dots, \lambda_s Q_s^{-1} Q_s \\ &= \text{diag} \lambda_1 I_{n_1}, \dots, \lambda_s I_{n_s}, \\ P^{-1}BP &= Q^{-1} \text{diag} B_{11}, \dots, B_{ss} Q \\ &= \text{diag} Q_1^{-1} B_{11} Q_1, \dots, Q_s^{-1} B_{ss} Q_s \\ &= \text{diag} \mu_{11}, \dots, \mu_{1n_1}, \dots, \mu_{s1}, \dots, \mu_{sn_s}. \end{aligned}$$

注意, 这里 P, Q, T 均为正交矩阵, 故 $P^{-1} = P^T, Q^{-1} = Q^T, T^{-1} = T^T$.

例 20 设 n 阶矩阵

$$\begin{aligned}
 \mathbf{A} &= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \\
 \mathbf{B} &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -1 & -1 \\ -\frac{1}{2} & \frac{1}{2} & -1 & -1 \\ -1 & -1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & -1 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.
 \end{aligned}$$

试求正交矩阵 $\mathbf{P}$ 和对角矩阵 $\mathbf{D}_1, \mathbf{D}_2$, 使得

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \mathbf{D}_1, \quad \mathbf{P}^{-1} \mathbf{B} \mathbf{P} = \mathbf{D}_2.$$

解 由

$$|\lambda \mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda & 1 \\ 1 & 1 \end{vmatrix}^2 = (\lambda - 1)^2 (\lambda + 1)^2 = 0,$$

知 $\lambda_{1,2} = 1, \lambda_{3,4} = -1$.

对 $\lambda = 1$, 解 $(\lambda \mathbf{I} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 得 $\mathbf{x} = \alpha (1, 0, 0, -1)^T + \beta (0, 1, -1, 0)^T$, $\alpha, \beta \in \mathbb{R}$, 取

$$\boldsymbol{\eta}_1 = \frac{1}{2}(1, 0, 0, -1)^T, \quad \boldsymbol{\eta}_2 = \frac{1}{2}(0, 1, -1, 0)^T;$$

对 $\lambda = -1$, 解 $(\lambda \mathbf{I} - \mathbf{A}) \mathbf{x} = \mathbf{0}$, 得

$$\mathbf{x} = \gamma (1, 0, 0, 1)^T + \delta (0, 1, 1, 0)^T, \quad \gamma, \delta \in \mathbb{R},$$

取 $\boldsymbol{\eta}_3 = \frac{1}{2}(1, 0, 0, 1)^T, \quad \boldsymbol{\eta}_4 = \frac{1}{2}(0, 1, 1, 0)^T$.

令 $\mathbf{T} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3, \boldsymbol{\eta}_4)$, 则有

$$\mathbf{T}^{-1} \mathbf{A} \mathbf{T} = \text{diag}\{1, 1, -1, -1\}.$$

此时有

$$T^{-1}BT = \frac{1}{2} \begin{pmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & -1 & -3 \\ 0 & 0 & -3 & -1 \end{pmatrix} = \begin{pmatrix} B_1 & \\ & B_2 \end{pmatrix}.$$

对 B_1 , 解 $|\mu I - B_1| = 0$, 得特征值为 $\mu_1 = 1, \mu_2 = 2$, 再求 B_1 的特征向量可得正交矩阵

$$Q = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix},$$

使 $Q^{-1}B_1Q = \text{diag}\{1, 2\}$. 对 B_2 , 解 $|\mu I - B_2| = 0$, 得特征值 $\mu_1 = 1, \mu_2 = -2$, 再求 B_2 的特征向量, 得正交矩阵

$$Q = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix},$$

使 $Q^{-1}B_2Q = \text{diag}\{1, -2\}$. 于是取

$$Q = \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix},$$

并令

$$P = TQ = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix},$$

则有

$$P^{-1}AP = \text{diag}\{1, 1, -1, -1\} = D,$$

$$P^{-1}BP = \text{diag}\{1, 2, 1, -2\} = D_2.$$

1.3 考试题评析

例 1(开卷考试题) 设三阶矩阵

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

(1) 试求实上三角矩阵 R , 使得 $A = R^T R$;

(2) 由问题(1)的计算结果, 你对 1.2 节例 14 的证法 1 有何改进? 你的结论是什么? 证明你的结论;

(3) 设想一下, 例 14(或你的)结论是否可以推广到其他类型的实对称矩阵.

解 (1) 对 (A, I) 做初等行变换, 有

$$\begin{aligned} (A, I) = & \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{pmatrix} \\ & \begin{matrix} -r_1 + r_i \\ i=2,3 \end{matrix} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix}, \\ & \begin{matrix} -c_1 + c_j \\ j=2,3 \end{matrix} \end{aligned}$$

故

$$P = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ 使 } P^T A P = I.$$

令

$$R = P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ 有 } A = R^T R.$$

(2) 由(1)的计算过程, 我们看到把 A 化为 I 的变换矩阵 P 就是对角线上元素均为 1 的上三角矩阵, 且只要 $a_{11} \neq 0$ 即可 (进一步计算要求 $a_{22}' \neq 0, \dots$). 证法 1 中的 QR 分解是不必要做的. 于是有结论: 设 A 为 n 阶正定实对称矩阵, 则存在对角线元素均为 1 的上三角矩阵 P , 使 $P^T A P$ 为对角矩阵. 以下证明之.

证法 1 由于 A 是正定矩阵, 所以 $a_{11} \neq 0$, 于是取

$$P_1 = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ & 1 & \cdots & 0 \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

则有

$$P_1^T A P_1 = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & & \cdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}.$$

又因 $a'_{22} > 0$, 故再取

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ & 1 & -\frac{a'_{23}}{a'_{22}} & \cdots & -\frac{a'_{2n}}{a'_{22}} \\ & & 1 & \cdots & 0 \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

则有

$$P_2^T P_1^T A P_1 P_2 = \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a'_{22} & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & * \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & * & \cdots & * \end{pmatrix}.$$

继续做下去, 得

$$P_s^T \cdots P_1^T A P_1 \cdots P_s = \text{diag}\{a_{11}, a_{22}^{(1)}, \cdots, a_{nn}^{(s)}\},$$

其中 $s \leq n-1$, 且 $P_1, \cdots, P_s$ 均为对角线上元素为 1 的上三角矩阵, 其积也是同类的上三角矩阵. 令 $P = P_1 \cdots P_s$, 则有

$$P^T A P = \text{diag}\{a_{11}, a_{22}^{(1)}, \cdots, a_{nn}^{(s)}\}.$$

证法 2(数学归纳法) 与例 14 方法 2 类似, 现简要写出.

当 $n=2$ 时, 取 $P = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} \\ 0 & 1 \end{pmatrix}$ 即可.

对 n 阶矩阵 A , 分块如下:

$$A = \begin{pmatrix} A_{n-1} & \beta \\ \beta^T & a_{nn} \end{pmatrix}.$$

取

$$P_1 = \begin{pmatrix} I_{n-1} & -A_{n-1}^{-1}\beta \\ 0 & 1 \end{pmatrix},$$

则

$$P_1^T A P_1 = \begin{pmatrix} A_{n-1} & 0 \\ 0 & a_{nn} - \beta^T A_{n-1}^{-1} \beta \end{pmatrix},$$

由归纳假设, 对 A_{n-1} 存在对角线上元素为 1 的上三角矩阵 P_{n-1} , 使

$$P_{n-1}^T A_{n-1} P_{n-1} = \text{diag}\{d_1, \cdots, d_{n-1}\}.$$

取

$$P_2 = \begin{pmatrix} P_{n-1} & \\ & 1 \end{pmatrix}, \quad P = P_1 P_2,$$

则 P 仍为对角线上元素为 1 的上三角矩阵, 且有

$$\begin{aligned} P^T A P &= P_2^T P_1^T A P_1 P_2 = \begin{pmatrix} P_{n-1}^T & & \\ & A_{n-1} & \\ & & P_{n-1} \end{pmatrix} \begin{pmatrix} 1 & & \\ & b & \\ & & 1 \end{pmatrix} \\ &= \text{diag}\{d_1, \cdots, d_{n-1}, b\}, \end{aligned}$$

其中

$$b = a_{nn} - \beta^T A_{n-1}^{-1} \beta.$$

(3) 可以推广到半正定实对称矩阵.

证 设 B 为半正定实对称矩阵, 则 B 的对角线元素 ≥ 0 . 记 $B = (b_{ij})$, 于是, 若 $b_{11} > 0$, 则仍可如(2)选取 P_1 , 再若 $b'_{22} > 0$, 又可选取 $P_2, \cdots$, 若所有的 $b_{11}, b'_{22}, \cdots, b_{n-1, n-1}^{(n-2)}$ 均为正, 则可一直如(2)做法. 下面要说明的是若做到某一次出现 $b_{ii}^{(i-1)} = 0$ 时怎么办, 为了书写

方便,不妨设 $i=2$,即取

$$P = \begin{pmatrix} 1 & -\frac{b_2}{b_1} & \cdots & -\frac{b_n}{b_1} \\ & 1 & & 0 \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

使

$$P^T B P_1 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & a'_{23} & \cdots & a'_n \\ 0 & a'_{32} & * & \cdots & * \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & a'_{n2} & * & \cdots & a'_{nn} \end{pmatrix}.$$

此时,由 B 的半正定性,必有 $a'_{2j} = 0 (j = 3, \cdots, n)$,从而 $a'_{j2} = 0$. 证明如下 若有 j 使 $a'_{2j} \neq 0$,于是二阶主子式

$$\begin{vmatrix} a'_{22} & a'_{2j} \\ a'_{j2} & a'_{jj} \end{vmatrix} = \begin{vmatrix} 0 & a'_{2j} \\ a'_{j2} & a'_{jj} \end{vmatrix} = - (a'_{2j})^2 < 0,$$

与 B 为半正定矩阵($P_1^T B P_1$ 半正定)矛盾,故当 $b'_2 = 0$ 时 必有

$$P^T B P_1 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & a'_{33} & \cdots & a'_{3n} \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & a'_{n3} & \cdots & a'_{nn} \end{pmatrix}.$$

同理当 $a'_{ii} = 0$ 时 必有 $a'_{ij} = 0 (i < j \leq n)$ 结论得证 .

评析 (1) 注意,这里 P, R 都是实方阵 .由 $P^T A P = D \quad A = R^T R$,故要求 D 的对角元 ≥ 0 ,若为负,则得不到实的方阵 R .

(2) 对正定矩阵,因其对角线上元素永远为正,故可以仅用第 2 种初等变换化为对角矩阵 .即永远是大块配方法中的第 1 种情况,故变换矩阵是上三角矩阵 .

例 2 设

$$A = \begin{pmatrix} & 1 & \\ & & 2 \\ 1 & & \end{pmatrix}, \quad B = \begin{pmatrix} & 1 & \\ & & \\ 1 & & -2 \end{pmatrix}.$$

(1) 求可逆矩阵 P , 使得 $P^T A P = B$;

(2) 当 λ 为何值时, 有 $\lambda I + A$ 为正定实对称矩阵.

解 (1) 因为以 A 为系数的二次型为

$$Q(x_1, x_2, x_3) = 2x_1x_3 + x_2^2,$$

$$x_1 = y_1 + y_3,$$

故令

$$x_2 = y_2,$$

$$x_3 = y_1 - y_3.$$

即

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & & 1 \\ & 1 & \\ 1 & & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

则有 $Q(x_1, x_2, x_3) \Big|_{x=By} = 2y_1^2 + y_2^2 - 2y_3^2$ 即取

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

$$\text{则 } P^T A P = B.$$

$$$$

(2) 由 $\lambda I + A$ 的各阶顺序主子式

$$d_1 = \lambda > 0, \quad d_2 = \lambda(\lambda + 1) > 0, \quad d_3 = (\lambda + 1)(\lambda^2 - 1) > 0,$$

故当 $\lambda > 1$ 时, $\lambda I + A$ 正定.

评析 (1) 是 $a_{11} = 0$, 而 $a_{33} \neq 0$ 的情况, 仿 $a_{12} \neq 0$ 做, 即可.

(2) 事实上只要看 $\lambda I + A$ 的各阶顺序主子式何时大于 0 即可 (由 $d_1 > 0 \Rightarrow \lambda > 0$; 再由 $d_2 > 0$, 得 $\lambda + 1 > 0$; 代入 $d_3 > 0$ 知 $\lambda^2 - 1 > 0$, 故 $\lambda > 1$). 可有的同学却用 $\lambda I + B$ 是正定矩阵代替, 故得出 $\lambda > 2$ 的结论. 这是错误的. 其问题的本质是: 当 $P^T A P = B$ 时, 有

$$P^T (\lambda I + A) P = \lambda P^T P + P^T A P = \lambda P^T P + B \neq \lambda I + B,$$

即 $\lambda \mathbf{I} + \mathbf{A}$ 与 $\lambda \mathbf{I} + \mathbf{B}$ 不相合, 要实践一下, 不能想当然。

例 3 用正交变换与坐标平移将二次曲面方程

$$2x^2 + 3y^2 + 3z^2 + 4yz + 4x - y + z + \frac{3}{2} = 0$$

化为标准方程, 并判断是什么曲面。

解 因二次曲面方程的二次项部分的矩阵为

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & & \\ & \mathbf{A}_2 & \end{pmatrix},$$

故知 $\mathbf{A}$ 有一个特征值为 2, 另两个满足

$$\left| \lambda \mathbf{I} - \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \right| = (\lambda - 3)^2 - 4 = (\lambda - 1)(\lambda - 5) = 0.$$

对 $\lambda = 1$, 解 $(\lambda \mathbf{I} - \mathbf{A}_2) \mathbf{x} = \mathbf{0}$, 得 $\boldsymbol{\eta} = \frac{1}{2}(-1, 1)^T$,

对 $\lambda = 5$, 解 $(\lambda \mathbf{I} - \mathbf{A}_2) \mathbf{x} = \mathbf{0}$, 得 $\boldsymbol{\eta} = \frac{1}{2}(1, 1)^T$ 。

令 $\mathbf{Q} = (\boldsymbol{\eta}, \boldsymbol{\eta})$, 则有 $\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \begin{pmatrix} 1 & \\ & 5 \end{pmatrix}$ 故令

$$\mathbf{P} = \begin{pmatrix} 1 & & \\ & \mathbf{Q} & \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

则有 $\mathbf{P}^T \mathbf{A} \mathbf{P} = \text{diag}\{2, 1, 5\}$ 。

记 $\mathbf{x} = (x, y, z)^T$, $\mathbf{b} = (4, -1, 1)$, $\mathbf{y} = (x_1, y_1, z_1)^T$, 则原二次曲面方程可以写为

$$Q(x, y, z) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b} \mathbf{x} + \frac{3}{2} \Big|_{\mathbf{x} = \mathbf{P} \mathbf{y}}$$

$$\begin{aligned}
 &= 2x_1^2 + y_1^2 + 5z_1^2 + 4x_1 + 2y_1 + \frac{3}{2} \\
 &= 2(x_1 + 1)^2 + y_1^2 + \frac{2}{2} + 5z_1^2 - 1 = 0,
 \end{aligned}$$

故再做平移变换得

$$2x_2^2 + y_2^2 + 5z_2^2 = 1.$$

从而得曲面是椭球面.

评析 (1) 求二次曲面的标准方程只能是: ①先做正交变换消去交叉项, 即把 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 化为平方和; ②再做平移, 即可得标准形. 如上过程 这里先做平移是没用的.

(2) 把 $\mathbf{x}^T \mathbf{A} \mathbf{x}$ 化为平方和这一步一定要用主轴化方法, 即求特征值、特征向量这一条路, 不能用初等变换法 如有的同学做成

$$\begin{aligned}
 &\begin{pmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 0 & 1 & 0 \\ 0 & 2 & 3 & 0 & 0 & 1 \end{pmatrix} \\
 &\begin{aligned} &- \frac{2}{3}r_2 + r_3 \end{aligned} \begin{pmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{5}{3} & 0 & -\frac{2}{3} & 1 \end{pmatrix} \\
 &\begin{aligned} &- \frac{2}{3}c_2 + c_3 \end{aligned}
 \end{aligned}$$

于是得

$$2x_2^2 + 3y_2^2 + \frac{5}{3}z_2^2 = 1.$$

虽然也是椭球面, 但显然有两个轴的大小发生了变化. 相当做了一个仿射变换, 其中

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}$$

是可逆矩阵, 但非正交矩阵.

例 4 把二次型

$$Q(\alpha) = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 2 \\ 1 & -1 & 2 & 1 \end{pmatrix} \mathbf{x}$$

化为标准形,并求可逆线性代换 $\mathbf{x} = \mathbf{P}\mathbf{y}$.

解 设二次型的矩阵为 $\mathbf{A}$,用初等变换法有

$$(\mathbf{A}, \mathbf{I}) = \begin{pmatrix} 0 & 1 & -1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 2 & 0 & 0 & 1 & 0 \\ 1 & -1 & 2 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \end{matrix}$$
$$\begin{matrix} r_2 + r_1, c_2 + c_1 \\ -\frac{1}{2}r_1 + r_2 \end{matrix} \begin{pmatrix} 2 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -1 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & -1 & 2 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \end{matrix}$$
$$\begin{matrix} 2r_2 + r_3, -2r_2 + r_4 \end{matrix} \begin{pmatrix} 2 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -1 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 & -1 & 0 & 1 \end{pmatrix},$$

故所求的可逆线性代换为

$$\mathbf{x} = \mathbf{P}\mathbf{y} = \begin{pmatrix} 1 & -\frac{1}{2} & -1 & 1 \\ 1 & \frac{1}{2} & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{y},$$

它使

$$Q(\alpha) \Big|_{x=Py} = y^T \left(-\frac{1}{2} \right) y.$$

下面是一个错误的做法,请读者看看错在哪里?

$$\begin{array}{l}
 (A, I) \quad \begin{array}{l} r_2 \setminus r_1 \\ r_1 + r_3 \\ -r_1 + r_4 \end{array} \begin{array}{ccccccc}
 1 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\
 0 & 1 & -1 & 1 & 1 & 0 & 0 & 0 \\
 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
 0 & -1 & 1 & 2 & 0 & -1 & 0 & 1
 \end{array} \\
 \begin{array}{l} -r_2 + r_3 \\ r_2 + r_4 \end{array} \begin{array}{ccccccc}
 1 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\
 0 & 1 & -1 & 1 & 1 & 0 & 0 & 0 \\
 0 & 0 & 2 & 0 & -1 & 1 & 1 & 0 \\
 0 & 0 & 0 & 3 & 1 & -1 & 0 & 1
 \end{array}
 \end{array}$$

所以

$$P = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P^T A P = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 2 & \\ & & & 3 \end{pmatrix}.$$

评析 在前面的正确解答中,虽然也是化为上三角矩阵,但再做那些列变换时,不会影响对角线上的结果.即少做的都是某列的多少倍加到另一列上去如 $-\frac{1}{2}c_1 + c_2$ 等,那不影响已化好的对角线元.很容易补做那些列变换,保持对角线不动而化为对角矩阵.在错误解法中少做了一个列变换是 $c_1 \setminus c_2$,再补做时,就破坏了现有的上三角形式,得重新考虑.显见它改变了二次型的属性.由三个正项一个负项变为全是正项.故我们强调列、行变换是一对一对

的,可以先连续做几个行变换,但千万不要忘记还要再做相应的列变换.

例 5 两个 n 阶实对称矩阵相合的条件是什么? 试判断矩阵

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -3 \\ 0 & -3 & 2 \end{pmatrix}$$

是否相合.

解 两个实对称矩阵相合的条件(见 1.1 节(3))是:

(1) 存在可逆矩阵 P , 使得 $P^T A P = B$;

(2) A 与 B 的正负惯性指数相同.

下面用(2)来判断 A, B 是否相合.

(分析: 因显然 A 的正、负惯性指数分别为 2, 1, 故只要看看 B 的正、负惯性指数是什么即可. 而 B 是准对角矩阵, 故只要看右下角的二阶块的正、负惯性指数即可, 这由求特征值立刻得到.)

$$\text{由 } \begin{vmatrix} \lambda I - \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix} \end{vmatrix} = (\lambda - 2)^2 - 9 = (\lambda - 5)(\lambda + 1) = 0, \text{ 知}$$

$\lambda = 5, \lambda = -1$, 故 B 的正、负惯性指数分别为 2, 1, 与 A 的惯性指数相同. 所以, A 与 B 相合.

讲评 (1) 在求矩阵 $\begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}$ 的正、负惯性指数时, 当然也

可以用初等变换法

$$\begin{pmatrix} 2 & -3 & 1 & 0 \\ -3 & 2 & 0 & 1 \end{pmatrix} \xrightarrow{\frac{3}{2}r_1 + r_2} \begin{pmatrix} 2 & -3 & 1 & 0 \\ 0 & -\frac{5}{2} & \frac{3}{2} & 1 \end{pmatrix},$$

故知其正、负项数均为 1. 于是 B 的正、负惯性指数分别为 2, 1. 这里有一个技巧, 就是把三阶的问题化为二阶的. 书写要简单许多.

(2) 有的同学用条件(1)做, 即找 P 使 $P^T A P = B$, 本题为

$$(\mathbf{B}, \mathbf{I}) = \begin{array}{cccccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -3 & 0 & 1 & 0 \\ 0 & -3 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccccc} & 1 & 0 & 0 & 1 & 0 & 0 \\ \frac{3}{2}r_2 + r_3 & 0 & 2 & 0 & 0 & 1 & 0 \end{array}$$

$$\begin{array}{cccccc} \frac{3}{2}c_2 + c_3 & 0 & 0 & -\frac{5}{2} & 0 & \frac{3}{2} & 1 \end{array}$$

$$\begin{array}{cccccc} & 2 & 0 & 0 & 2 & 0 & 0 \\ 2r_1, \frac{4}{2}r_2 & 0 & 3 & 0 & 0 & \frac{4}{2}3 & 0 \\ \frac{2}{5} \cdot \frac{4}{5}3r_3 & 0 & 0 & -3 & 0 & \frac{3}{10}3 & \frac{4}{5}12 \end{array} .$$

$$\text{于是得 } \mathbf{P} = \begin{array}{ccc} 2 & 0 & 0 \\ 0 & \frac{4}{2}3 & \frac{3}{10}3 \\ 0 & 0 & \frac{4}{5}12 \end{array} , \text{ 使 } \mathbf{P}^T \mathbf{B} \mathbf{P} = \mathbf{A}, \text{ 故 } \mathbf{A} \text{ 与 } \mathbf{B} \text{ 相合} .$$

显然,此法计算量大多了.因题目中没要求 $\mathbf{P}$,故做第一个初等变换后,已可看出 $\mathbf{B}$ 与 $\mathbf{A}$ 的正、负惯性指数相同,故已可以结束了.

第2章 代数系统

2.1 理论脉络

这一章内容比较抽象,学习要求是如下.

1 了解群、环、域的基本概念,特别是数域的各种性质

定义1(代数系统) 带运算的集合称为一个代数系统。一般是自身上的运算,有时,也和其他集合作运算.

如线性空间是一个代数系统,其中定义的数乘和加法两种运算分别是

$$F \times V \rightarrow V \quad (F \times V \text{ 到 } V \text{ 的映射}),$$

$$V \times V \rightarrow V \quad (V \times V \text{ 到 } V \text{ 的映射}).$$

定义2(群) 设 G 是非空集合,若在 G 中定义了一个二元运算 $*$ (即对 G 中任意元素 a, b 有 G 中惟一元素(记为 $a * b$)与之对应),且满足如下规律:

(1) 封闭性 对任意 $a, b \in G$, 总有 $a * b \in G$;

(2) 结合律 $a * (b * c) = (a * b) * c$, 对任意 $a, b, c \in G$ 成立;

(3) 有恒元 存在 $e \in G$, 使 $e * a = a$, 对任意 $a \in G$ 成立;

(4) 逆元 对任意 $a \in G$, 都存在 $b \in G$, 使 $b * a = e$.

则称 G (连同其上的运算) 为一个群 (group), 记为 $(G; *)$. 如果还有

$$a * b = b * a, \quad \text{对任意 } a, b \in G,$$

则称 G 为 Abel 群或交换群. (4) 中的 b 称为 a 的逆元, 记为 $b = a^{-1}$.

按此定义,知 $(\mathbb{Z}; +)$ 是 Abel 群;线性空间是一个加法群;

$$GL_n(F) = \{A \in M_n(F) \mid \det A \neq 0\}$$

(按通常矩阵乘法)是乘法群,称为一般线性群;

$$SL_n(F) = \{A \in M_n(F) \mid \det A = 1\}$$

也是乘法群,称为特殊线性群.

可以证明当 $(G; *)$ 是群时,也有 $a * a^{-1} = e, a * e = a$.

定义 3(环) 设 R 是一个集合,在 R 上定义了两个二元运算,分别记为加法 $(+)$ 和乘法 $(\cdot)$ 且满足:

- (1) $(R; +)$ 是一个 Abel 群;
- (2) $(R; \cdot)$ 是一个半群,即满足封闭性和结合律;
- (3) 乘法对加法有分配律,对任意 $a, b, c \in R$, 有

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

则称 R 为一个环(ring),记为 $(R; +, \cdot)$.加法恒元常记为 0 .

容易验证 $(\mathbb{Z}; +, \cdot)$ 是环,称为整数环.

n 阶矩阵的全体,按通常矩阵的加法和乘法是环.记为 $M_n(F)$.

定义 4(域) 设 F 是有两个二元运算 $(+)$ 和 $(\cdot)$ 的集合,且满足:

- (1) $(F; +)$ 是 Abel 群;
- (2) $(F^*; \cdot)$ 是 Abel 群, F^* 指 F 的非 0 元全体(即去掉加法恒元);

- (3) 有乘法对加法的分配律,对任意 $a, b, c \in F$, 有

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

则称 F 是一个域(field),记为 $(F; +, \cdot)$.乘法恒元常称单位元,记为 1 .

由以上定义知: $(\mathbb{Z}; +, \times)$ 不是域, $\mathbb{Q}$ 是最小的数域, $\mathbb{R}, \mathbb{C}$ 都是数域;此外形如 $F = \{a + b\sqrt{5} \mid a, b \in \mathbb{Q}\}$ 的集合(按通常数的加法和乘法)是一个数域.常记为 $\mathbb{Q}(\sqrt{5})$.这些都是无限域.下面举

一个有限域的例子:

设 $F_2 = \{\bar{0}, \bar{1}\}$, 定义加法和乘法如下:

$$\bar{0} + \bar{0} = \bar{0}, \quad \bar{0} + \bar{1} = \bar{1} + \bar{0} = \bar{1}, \quad \bar{1} + \bar{1} = \bar{0},$$

$$\bar{0} \cdot \bar{0} = \bar{0}, \quad \bar{0} \cdot \bar{1} = \bar{1} \cdot \bar{0} = \bar{0}, \quad \bar{1} \cdot \bar{1} = \bar{1}.$$

则可以验证 $(F_2; +, \cdot)$ 是域, 称为二元域. $\bar{0}$ 为加法恒元, $\bar{1}$ 为乘法恒元.

2 掌握等价关系, 集合的分类, 商集的定义及应用

定义 5 设 A 和 B 为两个集合, 称集合

$$\{(a, b) \mid a \in A, b \in B\}$$

为集合 A 与集合 B 的笛卡儿 (Cartesian) 积, 记作 $A \times B$.

注意, 没有交换律, 即通常 $A \times B \neq B \times A$. (a, b) 是个有序“对”.

定义 6 设 A, B 为两个集合, 则 $A \times B$ 的子集 R 称为从 A 到 B 的一个二元关系. 若 $(a, b) \in R$, 称为 a 与 b 具有关系 R , 记为 aRb ; 若 $(a, b) \notin R$, 称 a 与 b 不具有关系 R , 记作 $a \bar{R} b$.

当 $R = \emptyset$ (空集) 时, 称为空关系;

当 $R = A \times A$ 时, 称为全关系;

当 $A = B$ 时, $A \times A$ 的子集 R 称为 A 上的一个二元关系.

定义 7 设 R 是集合 A 上的一个二元关系.

(1) 若对任意 $a \in A$, 均有 $(a, a) \in R$, 则称 R 具有自返性.

(2) 对任意 $a, b \in A$, 若当 $(a, b) \in R$ 时, 也有 $(b, a) \in R$, 则称 R 具有对称性.

(3) 对任意 $a, b \in A$, 若当 $(a, b), (b, a) \in R$ 时, 有 $b = a$, 则称 R 具有反对称性.

(4) 对任意 $a, b, c \in A$, 若当 $(a, b), (b, c) \in R$ 时, 有 $(a, c) \in R$, 则称 R 具有传递性.

显然, $M_{m \times n}(R)$ 上的相抵关系, $M_n(R)$ 上的相似关系, $SM_n(R)$ (实对称矩阵的全体) 上的相合关系, 都具有自反性、对称

性、传递性 .

定义 8 具有自反性、对称性、传递性的 A 上的二元关系 R , 称为 A 上的等价关系 .

定义 9 A 的关于 R 的所有等价类组成的一个集合记为 A/R , 即

$$A/R = \{[a]_R \mid a \in A\},$$

称集合 A/R 为 A 关于 R 的商集 ($[a]_R$ 是 A 中与 a 等价的元素全体, 即 $[a]_R = \{b \mid b \in A \text{ 且 } (a, b) \in R\}$, 称为 a 关于 R 的等价类) .

定理 1 设 R 是 A 上的一个等价关系, 则商集 A/R 恰好是 A 的一个分类 .

定义 10 设 A 是非空集合, 若 A 的子集族 $\{A_\alpha\}_{\alpha \in I}$ (I 为指标集) 满足:

- (1) 对任意 $\alpha \in I$, 均有 $A_\alpha \neq \emptyset$;
- (2) 对任意 $\alpha, \beta \in I$, 若 $\alpha \neq \beta$, 则 $A_\alpha \cap A_\beta = \emptyset$;
- (3) $A = \bigcup_{\alpha \in I} A_\alpha$.

则称这个子集族为 A 的一个分类 .

定理 2 集合 A 上的任何一个分类, 都可以确定 A 上的一个等价关系 .

定理 1, 定理 2 读者可自己补证一下 . 前者只需证商集 A/R 中的元素 (集合) 恰满足定义 10 中的 3 条要求; 后者只要定义 A 上的二元关系 R : 对任意 $a, b \in A$,

$$aRb \iff \text{存在 } \alpha \in I, \text{ 使得 } a, b \in A_\alpha .$$

再证 R 是等价关系即可 .

总的来说要求: 给定集合 A , 会用 $A \times A$ 的子集写出 A 上的二元关系 R_1 , 特别是等价关系 R , 知道商集 A/R 是什么及 R 把 A 分成多少类 . 另外还要求会判断一个集合及其上的运算是否构成群 .

2.2 典型题讲评

例1 设 $|A| = m, |B| = n$, 由 A 到 B 有多少种二元关系?

解 由于集合 A 有 m 个元素, 集合 B 有 n 个元素, 故

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

共有 mn 个元素.

而每个 $A \times B$ 的子集都是 A 到 B 的一种二元关系. 所以只要看 $A \times B$ 的子集的个数, 有

$$|P(A \times B)| = 1 + C_{mn}^1 + C_{mn}^2 + \cdots + C_{mn}^{mn-1} + 1 = 2^{mn}.$$

故共有 2^{mn} 种二元关系.

讲评 (1) 这是一个考察基本概念的题目, 首先是“什么是二元关系”, 它是 $A \times B$ 的子集, 那么接下来就是“ $A \times B$ 即集合 A 到集合 B 的笛卡儿积是什么?”, 即本题主要用到定义 5, 定义 6. 这两个概念清楚, 问题就迎刃而解了.

(2) 注意这里 $|A|$ 表示集合 A 的元素个数, $P(A)$ 是集合 A 的全体子集所组成的集合, 称为 A 的幂集.

例2 对 $A = \{1, 2, 3\}$ 给出以下 5 个二元关系, 试分别判断它们是否具有自返性、对称性、反对称性、传递性.

(1) $R_1 = \{(1, 1), (1, 2), (1, 3), (3, 3)\};$

(2) $R_2 = \{(1, 1), (1, 2), (2, 2), (2, 3)\};$

(3) $R_3 = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3)\};$

(4) $R_4 =$;

(5) $R_5 = A \times A$.

解 (1) 由于 $(2, 2) \notin R_1$, 故无自返性. 又 $(1, 2) \in R_1$, 但 $(2, 1) \notin R_1$, 故 R_1 无对称性. 所以 R_1 具有反对称性和传递性.

(2) 由于 $(3, 3) \notin R_2$; $(1, 2) \in R_2$ 而 $(2, 1) \notin R_2$; $(1, 2) \in R_2$, $(2, 3) \in R_2$, 但 $(1, 3) \notin R_2$. 从而知 R_2 无自返性、对称性、传递性.

故 R_2 仅有反对称性.

(3) 由于 $(1, 2) \in R_3, (2, 1) \in R_3$, 但 $1 \neq 2$, 知 R_3 无反对称性. 而对任意 $a \in A$, 有 $(a, a) \in R_3$; 当 $(a, b) \in R_3$ 时, 有 $(b, a) \in R_3$. 所以 R_3 具有自返性、对称性, 同时它也具有传递性(因为若 $(a, b) \in R_3, (b, c) \in R_3$, 则 $(a, c) \in R_3$ 成立).

(4) 因为对任意 $a \in A, (a, a) \notin R_4$, 故 R_4 无自返性, 但它有对称性(与(3)比较就可理解)、反对称性、传递性.

(5) 由于 $(1, 2) \in R_5, (2, 1) \in R_5$, 但 $1 \neq 2$, 故 R_5 无反对称性. 显然, 由定义 7 知其有自返性、对称性、传递性.

讲评 (1) 注意定义 7 中自反性是“无条件要求”, 即对任意 $a \in A$, 都有 $(a, a) \in R$, 故判断任一个 R 是否有此性质, 就看是否对每个元素都有 aRa . 如每个 $m \times n$ 矩阵都和自己相抵($AD \ A$); 每个 n 阶矩阵都和自己相似($A \sim A$); 每个实对称矩阵都和自己相合($A \ddot{u} \ A$). 故它们都有自返性.

(2) 对称性、反对称性、传递性是“有条件要求”, 即若 $(a, b) \in R$ 时, 要求 $(b, a) \in R$, 满足此要求, 就称有对称性, 如 R_3 中有 $(1, 2) \in R_3, (2, 1) \in R_3$, 而且其中没有另外的 $(a, b) \in R$, 但 $(b, a) \notin R$, 故有对称性. 注意这里因没有 $(1, 3) \in R$, 故不必去看 $(3, 1) \in R$ 是否存在. 由此可理解 $R_4 =$ 为什么有对称性. 同样可理解反对称性和传递性.

例 3 判断下列二元关系 R 是否为等价关系? 若是, 写出关于 R 的商集, 并说明商集中元素的意义.

(1) 设 R 是 $Z \times Z$ 上的二元关系:

$$((a, b), (c, d)) \in R \quad a - c = b - d;$$

(2) R 是 $N \times N$ 上的二元关系:

$$((a, b), (c, d)) \in R \quad ac = bd;$$

(3) 设 $A = R \times R$ 是平面上所有点的集合, 二元关系 R 如下规定: 任意 $a, b \in A$, 有

$(a, b) \in R$ a 与 b 在同一水平直线上;

(4) 设 $L = \{ l \mid l \text{ 为平面上的直线} \}$, L 上二元关系 R 如下规定: 任 $l, l' \in L$, 有

$$(l, l') \in R \quad l \perp l';$$

(5) $M_n(R)$ 上的二元关系 R 如下规定:

$$(A, B) \in R \quad \det A = \det B.$$

分析 要正确地得出结果, 主要需弄清楚: 集合 A 是什么? A 上的二元关系 R 是什么, 当 R 是等价关系时, 商集 A/R 是什么?

解 (1) R 是等价关系. 集合 $A = \{ (a, b) \mid a, b \in \mathbb{Z} \}$, 即平面上坐标为整数的点的全体. $R: ((a, b), (c, d)) \in R \quad a - b = c - d$, 即两个整数点有关系的充要条件是其横坐标与纵坐标之差相等. 可以验证 R 满足自返性、对称性、传递性, 故 R 是等价关系. 商集

$$A/R = \{ [(a, b)]_R \mid (a, b) \in \mathbb{Z} \times \mathbb{Z} \}.$$

这是一个无限集, 因为 $(a - b) = 0, 1, 2, \dots, n, \dots$. 故商集也可以写成

$$A/R = \{ [(a, a)]_R, [(a+1, a)]_R, [(a+2, a)]_R, \dots, [(a+n, a)]_R, \dots \},$$

其中 $a \in \mathbb{Z}$. $[(a+n, a)]_R$ 表示横坐标与纵坐标差为 n 的所有整数点.

(2) R 不是等价关系. 集合 $A = \{ (a, b) \mid a, b \in \mathbb{N} \}$, 即平面上坐标为自然数的点的全体. $R: (a, b) R (c, d) \quad ac = bd$. 因当 $a \neq b$ 时, $a^2 \neq b^2$, 故 $((a, b), (a, b)) \notin R$, 所以 R 无自返性. 由此得 R 不是等价关系.

(3) R 是等价关系. 集合 $A = \{ a = (x, y) \mid x, y \in \mathbb{R} \}$,

二元关系 $R: aRb$ 两点 a 与 b 在同一水平线上.

显然, R 满足自返性(每个点 a 都与自己在同一水平线上)、对称性(当 a 与 b 在同一水平线上时, b 与 a 也在同一水平线上)、传递性

(当 a 与 b 在同一水平线上, b 与 c 在同一水平线上, 则 a 与 c 也在同一水平线上), 故 R 是等价关系. 商集

$$\begin{aligned} A/R &= \{[a]_R \mid a \in R \times R\} \\ &= \{[(x_1, 0)]_R, [(x_2, 1)]_R, [(x_3, 2)]_R, \cdots\}, \end{aligned}$$

其中 $x_i \in R$. 显然 A/R 是无限集, 每个元素 $[a]_R$ 表示过点 $a = (x, y)$ 的水平直线.

(4) R 不是等价关系. 集合 $A = L = \{l \mid l \text{ 为平面上的直线}\}$,

$$R: l R l' \iff l \perp l'.$$

因 $l \perp l$ 不成立, 故 R 无自返性. 所以 R 不是等价关系.

(5) R 是等价关系. 集合 $M_n(R) = \{A = (a_{ij}) \mid a_{ij} \in R\}$,

$$R: \text{任意 } A, B \in M_n(R), A R B \iff \det A = \det B.$$

显然 R 具有自返性、对称性、传递性, 故 R 是等价关系. 商集

$$M_n(R)/R = \{[A]_R \mid A \in M_n(R)\},$$

其中 $[A]_R$ 表示与 A 的行列式值相等的方阵的全体. 这里商集也是个无限集.

讲评 正如本题解前分析的那样, 做这样的题, 首先要看清楚已知集合是什么? 如(1), (2)中 A 是平面上一些特殊的点. 它们是用坐标形式给出的, 即 (a, b) 表示一个点, 为 A 中一个元素, 而不是表示 a 与 b 之间有关系 R . 两点之间的关系是用 $(a, b) R (c, d) \iff a - c = b - d$ (如(1)), 或 $(a, b) R (c, d) \iff ac = bd$ (如(2))表示的. 这样就不会把集合 A 与关系 R 混淆了. 另外商集是由 A 关于 R 的等价类组成的集合, 即是集合的集合, 故 A/R 中的元素一定是集合(某个等价类), 而不会是 A 中的某一个元素.

例 4 设 $A = \{1, 2, 3, 4, 5\}$, 试构造出 A 上的一个等价关系, 并说明按照这个等价关系可以把 A 分成多少类?

讲评 要构造出 A 上的等价关系 R , 首先要知道 R 是一个二元关系, 故它应是 $A \times A$ 的子集. 其次它应有自返性, 故任意 $a \in A$, $(a, a) \in R$; 故 R 中要包含

$$(1,1), (2,2), (3,3), (4,4), (5,5).$$

再次,若 $(1,2) \in R$, 则 $(2,1)$ 必属于 R , 这样才满足对称性. 又若要 $(1,2) \in R, (2,3) \in R$, 则 $(1,3)$ 必属于 R , 才满足传递性的要求. 经过这样的分析, 就不难写出想要的等价关系了.

解 如 $R_1 = \{(1,1), (2,2), (3,3), (4,4), (5,5)\}$ 是 A 上的等价关系, 它把 A 分为 5 类. 商集

$$A/R_1 = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$$

是 A 的一个分类, 即每个元素在一个类里.

再如 $R_2 = \{(1,1), (2,2), (3,3), (4,4), (5,5), (1,2), (2,1)\}$ 也是 A 上的等价关系, 它把 A 分为 4 类. 商集

$$A/R_2 = \{\{1,2\}, \{3\}, \{4\}, \{5\}\}$$

是 A 的一个分类, 即 1, 2 在一个类里, 3, 4, 5 各在一个类里.

又如 $R_3 = \{(1,1), (2,2), (3,3), (4,4), (5,5), (1,2), (2,1), (1,3), (3,1), (2,3), (3,2), (4,5), (5,4)\}$ 是 A 上的一个等价关系, 它把 A 分为两类. 商集 $A/R_3 = \{\{1,2,3\}, \{4,5\}\}$ 是 A 的一个分类, 即 1, 2, 3 组成一个等价类, 4, 5 组成另一个等价类.

例 5 设 $|A| = 3$, 求 A 上的所有等价关系.

分析 由于 A 上的二元关系都是 $A \times A$ 的子集, 而 $|A| = 3$, 所以很容易如例 4 那样分析, 写出全部等价关系; 其次可由定理 1 和定理 2, 看 A 有多少种不同的分类, 从而写出等价关系.

解 设 $A = \{a, b, c\}$

方法 1 由于 A 的元素个数少, 故可直接写出全部等价关系. 首先由于 R 具有自返性, 故

$$R_1 = \{(a, a), (b, b), (c, c)\} \text{ 是一个等价关系.}$$

又 $(a, b) \in R$ 时, 有 $(b, a) \in R$, 故有

$$R_2 = \{(a, a), (b, b), (c, c), (a, b), (b, a)\},$$

$$R_3 = \{(a, a), (b, b), (c, c), (a, c), (c, a)\},$$

$$R_4 = \{(a, b), (b, b), (c, c), (b, c), (c, b)\},$$

都有对称性、自返性(且也满足传递性).故都是等价关系.下面要 R_1 的元素更多,即 $(a, b) \in R, (b, c) \in R$, 则由传递性知须有 $(a, c) \in R$.再按对称性的要求,故有

$$R_5 = A \times A,$$

所以, A 上的等价关系共有以上 5 种.

方法 2 由定理 1 和 2 知道,每个等价关系都与一个分类对应,故看 A 有多少种分类方式,就可以写出相应的等价关系.如 3 个元素都在一个等价类中,对应的等价关系就是 R_5 ;两个在一类,而另一个独立一类有 3 种情况,对应的等价关系分别是上面的 R_2, R_3, R_4 ;最后每个元素自己在一类,即 A 被分为 3 类,对应的等价关系就是 R_1 .

讲评 由于 A 中元素个数少,用这两种方法找全部等价关系都可以.但对初学者来说,方法 1 抽象些;方法 2 比较直观.3 个元素有多少种分类方式,很容易想明白.希望读者能学会用方法 1 的思维方式.

例 6 设 $*$ 是定义在 $M_n(R)$ 上或是从 $M_n(R)$ 到 R 上的二元运算,判断下列运算是否满足交换律、结合律.

$$(1) A * B = \max(\det A, \det B);$$

$$(2) A * B = 2A + 3B;$$

$$(3) A * B = \min(r(A), r(B));$$

$$(4) A * B = AB + A.$$

解 (1) 显然 $A * B = B * A$, 故满足交换律.又因为对任意 $A, B, C \in M_n(R)$,

$$(A * B) * C = \max(\det A, \det B) * C$$

无意义,故不满足结合律.

(2) 因 $A * B = 2A + 3B \neq 2B + 3A = B * A$, 故不满足交换律,且

$$(A * B) * C = (2A + 3B) * C = 2(2A + 3B) + 3C,$$

$$A * (B * C) = 2A + 3(B * C) = 2A + 3(2B + 3C),$$

故也不满足结合律.

(3) 因 $A * B = B * A$, 故满足交换律. 但因 $(A * B) * C = \min(r(A), r(B)) * C$ 无意义, 故不满足结合律.

(4) 由于 $A * B = AB + A \neq BA + B = B * A$, 且

$$(A * B) * C = (AB + A) * C = (AB + A)C + AB + A,$$

$$A * (B * C) = A * (BC + B) = A(BC + B) + A,$$

故 $A * (B * C) \neq (A * B) * C$, 所以不满足交换律和结合律.

讲评 只要严格按定义做, 不难得出结果. 有的同学凭想当然, 以为(1)中 $(A * B) * C = \max(\det A, \det B, \det C)$ (仔细审题, 就可以知道, $A * B$ 要先算, 算出来已是一个数 $(\in M_n(R))$, 故下一步无法计算), 故得出“满足结合律”的错误结论来.

例 7 在 $GL_n(R)$ 上定义的二元运算 $*$ 为: " $A, B \in GL_n(R), A * B = ABA^{-1}$ ". 问 $*$ 是否满足结合律、交换律? $GL_n(R)$ 关于 $*$ 是否存在单位元? $GL_n(R)$ 中每一元素是否可逆? 可逆元的逆元是什么?

解 $GL_n(R)$ 是 n 阶可逆矩阵的全体, 因 $A * B = ABA^{-1} \neq BAB^{-1} = B * A$, 故不满足交换律. 又 $(A * B) * C = (ABA^{-1}) * C = (ABA^{-1})C(ABA^{-1})^{-1}$, 而

$$A * (B * C) = A(B * C)A^{-1} = A(BCB^{-1})A^{-1},$$

所以也不满足结合律.

关于单位元, 若有左单位元 C , 则对任意 $B \in GL_n(R)$, 应有

$$C * B = CBC^{-1} = B,$$

即只要 $CB = BC$, 就是说 C 要与一切 $B \in GL_n(R)$ 乘法可交换, 故 $C = \lambda I$ (与一切可逆矩阵均乘法可交换的矩阵必为纯量矩阵), 所以有左单位元, 且不惟一 (任一个纯量矩阵都是左单位元).

再来看哪些元素有左逆元. 由

$$\lambda I = A * B = ABA^{-1} \quad AB = \lambda I \quad B = \lambda I,$$

故只有纯量矩阵有左逆,其左逆不惟一(由上知,任意 A 均为其逆).

又若有右单位元 D ,则应有对任意 $A \in GL_n(R)$,都有

$$A * D = ADA^{-1} = A \quad DA^{-1} = I \quad D = A,$$

故不存在右单位元.于是无单位元,也谈不上逆元.

讲评 在 $GL_n(R)$ 中,若是通常的矩阵乘法,则单位矩阵 I 是单位元.每个 A^{-1} 是 A 的逆元.但这里给的运算“ $*$ ”不是通常的矩阵乘法(当然也不是加法),而是 $A * B = ABA^{-1}$.在此运算下,一切都变了,由论证的结果,既无单位元,也无逆元.显然,运算“是什么”起了决定的作用,理解了这点,就不会问:“可逆矩阵怎么会没有逆元呢?”这样的问题了.

例 8 设 $(A; *)$ 是一代数系统, $*$ 满足如下规律: " $a, b \in A$, 有

$$(a * b) * a = a, \quad (a * b) * b = (b * a) * a.$$

求证: 1) $a * (a * b) = a * b$;

$$(2) \quad a * a = (a * b) * (a * b) = b * b;$$

$$(3) \quad a * b = b * a \quad a = b.$$

解 (1) 记题设中的两个已知公式分别为①, ②, 则有

$$a * (a * b) \xrightarrow{\text{用①}} [(a * b) * a] * (a * b) \xrightarrow{\text{再用①}} (a * b).$$

第 2 个等号是在公式①中用 $a * b$ 代 a ; 用 a 代 b , 即得.

$$(2) \quad a * a \xrightarrow{\text{由①}} [(a * b) * a] * a \xrightarrow{\text{用②}} [a * (a * b)] * (a * b) \\ \xrightarrow{\text{由(1)}} (a * b) * (a * b). \quad (*)$$

第 2 个等号是在公式②式的右边取 $b = a * b$, $a = a$, 即得左边应为上述结果. 又由 $(*)$ 式, a, b 互换位置, 即得

$$b * b = (b * a) * (b * a) \\ \xrightarrow{\text{用}(*)} [(b * a) * a] * [(b * a) * a] \\ \xrightarrow{\text{用②}} [(a * b) * b] * [(a * b) * b] \\ \xrightarrow{\text{由}(*)} (a * b) * (a * b).$$

第2个等号是这样用 $(*)$ 式的: 令 $a = b * a, b = a$; 第3个等号是从右边往左边用②式; 最后一个等号是从右边往左用 $(*)$ 式: 即用 $a * b$ 代替 a , 即得.

(3) 充分性显然 ($a = b$ 时, 有 $a * b = b * a$). 再证必要性. 当 $a * b = b * a$ 时, 有

$$\begin{aligned} & a \xrightarrow{\text{用①}} (a * b) * a \xrightarrow[\text{代入}]{\text{条件}} (b * a) * a \xrightarrow{\text{用②}} (a * b) * b \\ & \xrightarrow{\text{用条件}} (b * a) * b \xrightarrow{\text{用①}} b. \end{aligned}$$

讲评 本题没什么大难度, 只是在已知的两个公式中巧妙地选取 a, b , 如(1)中先用 $a = (a * b) * a$ 代入, 即从右到左用公式①, 再用 $(a * b)$ 代 a, a 代 b 从左到右用公式①, 就得结果. 以下几个公式类似.

例9 设 $|A| = m, |B| = n \quad (m > n)$.

(1) 从 A 到 B 的不同映射共有多少?

(2) A 上的二元运算共有多少?

解 (1) 设 $A = \{a_1, \dots, a_m\}, B = \{b_1, \dots, b_n\} \quad (m > n)$, 则因为 a_i 可以被分别映为 $b_1, \dots, b_n$, 故共有 n 种, 而 $i = 1, \dots, m$, 所以, 共有 n^m 种映射.

(2) A 上的二元运算可由下表给出:

A	a_1	a_2	$\dots$	a_m
a_1	c_{11}	c_{12}	$\dots$	c_{1m}
a_2	c_{21}	c_{22}	$\dots$	c_{2m}
$\dots$	$\dots$	$\dots$		$\dots$
a_m	c_{m1}	c_{m2}	$\dots$	c_{mm}

因为 $a_i * a_j = c_{ij} \quad (i = 1, \dots, m, j = 1, \dots, m)$, 所以其运算结果共 m^2 个数. 而这 m^2 个位置都可有 m 种选取, 所以共有 m^{m^2} 种 (每个位

置有 m 种不同的选择,共有 m^2 个位置)。

讲评 因 A 上的二元运算是 $A \times A$ 到 A 的映射,所以每个位置(即 c_{ij})都可以独立选取 A 中任一个元素.若要求每行的元素是 $a_1, \dots, a_m$ 的排列,或要求运算有某些性质,则运算的个数就会大大减少。

例 10 设 F_1, F_2 为任意两个数域,问 $F_1 \cap F_2, F_1 \cup F_2$ 是否还为数域?若是,证明之;若不是,举反例。

解 因为有理数域 Q 是最小的数域,故 $Q \subseteq F_1, Q \subseteq F_2$, 于是知 $Q \subseteq F_1 \cap F_2$. 所以加法恒元 0 和乘法单位元 1 均属于 $F_1 \cap F_2$. 由此也得加法逆元和乘法逆元均属于 $F_1 \cap F_2$. 任意 $a \in F_1 \cap F_2$, $a \in F_1$ 且 $a \in F_2$, 因 $b + a = 0 \Rightarrow b = 0 - a \in F_1$, 且 $b = 0 - a \in F_2$, 故 $b \in F_1 \cap F_2$. 同理 $\frac{1}{a} \in F_1 \cap F_2$. 又对任意 $a, b, c \in F_1 \cap F_2$, 它们同时也是 F_1 (和 F_2) 中的元素,于是由已知 F_1 (和 F_2) 是数域,故满足封闭性、结合律(即 $a + b \in F_1$ 且 $a + b \in F_2 \Rightarrow a + b \in F_1 \cap F_2$, 同理 $ab \in F_1 \cap F_2$)及分配律等. 故得:

- (1) $(F_1 \cap F_2; +)$ 是 Abel 群;
- (2) $(F_1 \cap F_2; \cdot)$ 除 0 外是 Abel 群;
- (3) 乘法对加法有分配律。

所以 $F_1 \cap F_2$ 是数域. 而 $F_1 \cup F_2$ 不一定是数域. 如

$$F_1 = \{a + b\sqrt{2} \mid a, b \in Q\}, \quad F_2 = \{c + d\sqrt{5} \mid c, d \in Q\},$$

则 $F_1 \cup F_2$ 对加法不封闭. 如 $a = 1 + \sqrt{2}, b = \sqrt{2} + \sqrt{5}$, 则 $a + b \notin F_1 \cup F_2$. 故 $F_1 \cup F_2$ 不是数域, 当 $F_1 \not\subseteq F_2$ 或 $F_2 \not\subseteq F_1$ 时, $F_1 \cup F_2$ 显然仍为数域。

讲评 对任意一个非空数集,可以用它对加、减、乘、除(分母不为 0)是否封闭来验证是否为数域. 我们这里是用的一般域的定义来证明的。

2.3 考试题评析

例 1 设集合 $A = \{1, 2, 3, 4\}$, A 中的二元运算 $*$ 和 $\circ$ 分别定义为:

$*$	1	2	3	4
1	2	1	3	4
2	4	2	3	1
3	1	3	2	4
4	3	4	1	2

	1	2	3	4
1	3	4	1	2
2	4	3	2	1
3	1	2	3	4
4	2	1	4	3

则代数系统 $(A; *)$ 与 $(A; \circ)$ 中哪一个成为群? 成为群的代数系统中单位元是什么? 又 1 的逆元是什么?

解 $(A; \circ)$ 是群, 单位元是 3, 1 的逆元是 1.

评析 由定义 2, 要判断一个代数系统是否为群. 要检验它满足封闭性、结合律, 还要检验它有恒元、有逆元. 封闭性一眼就可看出, 两个都满足, 但要检查结合律就太麻烦了. 要验证 C_4^3 ! 那么多个运算, 但因群有性质“左恒元等于右恒元”及“左逆元等于右逆元”, 所以可用此否定掉不是群的那个. 如本题 $(A; *)$ 无左恒元, (右恒元等于 2), 故它不是群. 于是 $(A; \circ)$ 应是群. 由 $3 \circ a = a$, 对任意 $a \in A$, $a \circ 3 = a$, 故 3 是恒元, 且对每个 $a \in A$, 均有 $a \circ a = 3$, 故 $a = a^{-1}$.

例 2 令 G 是所有秩不大于 $r (\leq n)$ 的 n 阶复矩阵的集合, 则在通常的矩阵乘法运算下, G 是否构成群? 说明理由.

解 G 不构成群. 因为若 $r < n$, 则 $I \notin G$, 即 G 中无恒元; 而当 $r = n$ 时, 虽然 $I \in G$, 但 G 中包含所有秩 $< r$ 的 n 阶矩阵, 这些矩阵是不可逆的, 所以, 无论如何 G 都不是群.

评析 对于比较抽象的数学题, 认真审题是很重要的. 有的同学得出“当 $r = n$ 时 G 是群, $r < n$ 时 G 不是群”的结论, 就是没有懂

得题中的“ G 中包含所有秩小于 r 的矩阵”这句话.若看到这点,无论 $r=n$ 还是 $r<n$,这些秩小于 r 的矩阵都无逆.故不用讨论,即知 $(G; *)$ 不是群.

例3 试问相合关系将全体三阶实对称矩阵 $SM_3(\mathbb{R})$ 分成多少类?并写出各等价类的代表元.

解 共分成10个等价类.各类的代表元为:

$$\begin{array}{cccc}
 0 & & 1 & & -1 & & 1 \\
 & 0 & & 0 & & 0 & & 1 \\
 & & 0 & & 0 & & 0 & & 0 \\
 1 & & -1 & & 1 & & 1 \\
 & -1 & & -1 & & 1 & & 1 \\
 & & 0 & & 0 & & 1 & & -1 \\
 & & 1 & & -1 \\
 & & & -1 & & -1 \\
 & & & & -1 & & -1
 \end{array}$$

评析 (1) 由等价类的个数为 $\frac{(n+1)(n+2)}{2} = \frac{4 \times 5}{2} = 10$,可知分10类.也可直接求代表元知有10类(这就要求代表元不能错,故若先按计算公式算出10类,也可校验代表元的正误).

(2) 代表元怎么找,当然是找最简单的矩阵——即在实规范型中找.按一定规律找就不会漏掉,比如上面是按秩为0,秩为1,秩为2,3的顺序,再按秩确定时先排全正,再有一负、两负,……零矩阵往往被忽视,回答成9类的.多半是缺了零矩阵.凭心而论,这是一个基本题.同学之所以答错,还是基本概念不清楚.

例4 在集合 $M_2(\mathbb{Z})$ ($\mathbb{Z}$ 为整数集)中,定义等价关系 R : 任给 $A, B \in M_2(\mathbb{Z})$, $ARB \iff \det A \equiv \det B \pmod{2}$, 则它把 $M_2(\mathbb{Z})$ 分成

多少类?各等价类的代表元是什么?又矩阵 $\begin{pmatrix} 3 & 6 \\ -4 & -2 \end{pmatrix}$ 属于其

中的哪一个等价类?

解 此等价关系把 $M_2(\mathbb{Z})$ 分成 2 类, 其代表元分别为:

$$\begin{matrix} & \begin{matrix} 1 & & 1 \\ & 0 & \text{和} & 1 \\ & 3 & 6 \\ -4 & -2 \end{matrix} & \begin{matrix} \\ \\ \\ \end{matrix} \end{matrix} \in \begin{matrix} \overline{1} \\ \\ 0 \end{matrix}.$$

矩阵

评析 由题设, 等价关系 R 是计算两个矩阵的行列式, 且其值满足模 2 同余. 由方阵的元素为整数, 知 $\det A$ 为整数, 而整数模 2 同余, 只有两类, 即余 0 和余 1. 故 $M_2(\mathbb{Z})$ 按此分类只有两种情况, 即行列式值分别为偶数与奇数. 这两类中任一个矩阵都有资格当代表元, 通常应写最简单的. 看起来一目了然, 不必再需计算, 故

$$\text{选 } \begin{matrix} 1 & & 1 \\ & 0 & \text{和} & 1 \\ & & & \end{matrix}, \text{ 即两个等价类为 } \begin{matrix} \overline{1} & \overline{1} \\ 0 & 1 \end{matrix}.$$

前者是行列式值为偶数 ($\equiv 0 \pmod{2}$) 的方阵全体, 后者是行列式值为奇数 ($\equiv 1 \pmod{2}$) 的方阵全体.

$$\text{又计算 } \begin{vmatrix} 3 & 6 \\ -4 & -2 \end{vmatrix} = -6 + (24) = 18 \equiv 0 \pmod{2}, \text{ 故 } \begin{vmatrix} 3 & 6 \\ -4 & -2 \end{vmatrix} \in \begin{matrix} \overline{1} \\ \\ 0 \end{matrix}.$$

如果答 $\begin{vmatrix} 3 & 6 \\ -4 & -2 \end{vmatrix} \in \begin{vmatrix} 3 & 6 \\ -4 & -2 \end{vmatrix}$ 也不能算错, 但依题意应理解为需要回答它与前面得出的代表元中哪一个属于同一类.

例 5 设集合 $A = \{1, 3, 5, 7, 9\}$, 试写出 A 上的一个等价关系 (要求用 $A \times A$ 的子集写出) R , 并说明按此等价关系把 A 分成几类?

解 $R = \{(1, 1), (3, 3), (5, 5), (7, 7), (9, 9)\}$, 它把 A 分成 5 类.

评析 此题答案不惟一, 上面写出的是最简单的一个, 仿 2.2

节例 4 可以写出更多的等价关系. 对这样比较抽象的题, 不要心存恐惧, 从定义出发, 一步步分析思路就清晰了. 如等价关系的第一个要求是要有自返性, 即对任意 $a \in A$, 都有 $(a, a) \in R$, 故知 R 中必含全部的 (a, a) (任意 $a \in A$), 已知 A 中有 5 个元素, 故知 R 中包含了至少 5 个元素 (a, a) 即 R 至少为

$$R_1 = \{(1, 1), (3, 3), (5, 5), (7, 7), (9, 9)\}.$$

接下来是考虑对称性和传递性的要求. 首先可以看出 R_1 同时也满足对称性和传递性, 故 R_1 已是一个等价关系. 由此等价关系, 每个 A 中的元素只与自己有关系, 与其他元素无关, 即商集为

$$A/R_1 = \{\{1\}, \{3\}, \{5\}, \{7\}, \{9\}\}.$$

所以, 这个分类把 A 分为 5 类.

如果读者觉得 R_1 太简单, 或希望 5 个元素之间有点关系, 那么一旦写入 $(1, 3) \in R$, 则必须有 $(3, 1) \in R$, 即 R 至少为

$$R_2 = \{(1, 1), (3, 3), (5, 5), (7, 7), (9, 9), (1, 3), (3, 1)\}.$$

若觉得还不够复杂, 还要写入 $(5, 7) \in R$, 则必须有 $(7, 5) \in R$ 才满足对称性的要求, 即 R 成为

$$R_3 = \{(1, 1), (3, 3), (5, 5), (7, 7), (9, 9), (1, 3), (3, 1), (5, 7), (7, 5)\}.$$

显然, R_2, R_3 同时也都满足传递性, 故它们都是等价关系.

但若想写入 $(1, 3) \in R, (3, 5) \in R$, 则麻烦事来了, 除了对称性要求必有 $(3, 1) \in R, (5, 3) \in R$ 外, 还要有传递性, 即必须有 $(1, 5) \in R$, 再由对称性, 得 $(5, 1) \in R$, 于是 R 必须含更多的元素

$$R_4 = \{(1, 1), (3, 3), (5, 5), (7, 7), (9, 9), (1, 3), (3, 5), (3, 1), (5, 3), (1, 5), (5, 1)\}.$$

此时 $A/R_4 = \{\{1, 3, 5\}, \{7\}, \{9\}\}$, A 被分成 3 类.

第 3 章 多项式形式

3.1 理论脉络

这一章可按以下 3 个脉络理解和掌握多项式形式的基本内容.

1 会带余除法, 了解整除的简单性质

定义 1 (1) 设 F 为一个域, x 是不属于 F 的任意一个符号, 则形如

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in F$$

的表达式称为域 F 上的一个多项式形式. 不等于 0 的最高次项的系数称为首项系数, 首项系数为 1 的多项式称为首一多项式.

(2) 两个多项式形式相等. 定义为它们的次数和各同次系数均相等.

定理 1 域 F 上 x 的多项式形式全体(记为 $F[x]$)按如下定义加法和乘法成为交换环(称为多项式形式环).

$$\begin{aligned} \sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i &= \sum_{i=0}^{\infty} (a_i + b_i) x^i, \\ \sum_{i=0}^{\infty} a_i x^i \sum_{j=0}^{\infty} b_j x^j &= \sum_{k=0}^{\infty} \sum_{i+j=k} a_i b_j x^k. \end{aligned}$$

注意, 用无限求和的形式给出, 只是为书写方便(不用讨论两个多项式 f 和 g 的次数), 事实上, 我们有个约定, 和号 $\sum_{i=0}^{\infty}$ 中只有有限项的系数非零.

定理 2 $F[x]$ 满足乘法消去律: 即对任意 $f, g, h \in F[x]$, 若有 $fg = fh$, 且 $f \neq 0$, 则 $g = h$.

定义 2 有消去律的含么交换环,称为整环(domain).如 $\mathbb{Z}$ 和 $F[x]$ 均为整环.

定理 3(带余除法) 对任意 $f, g \in F[x]$, 若 $g(x) \neq 0$, 则总存在 $q, r \in F[x]$, 使

$$f = gq + r, \quad \text{其中 } \deg r < \deg g \text{ 或 } r = 0,$$

且 q 和 r 由(多项式形式) f 和 g 惟一决定.

定义 3 有带余除法的环称为欧几里得(Euclid)环.如 $\mathbb{Z}$ 和 $F[x]$ 都是欧几里得环.

当定理 3 中的 $r = 0$ 时,称 g 整除 f ,记为 $g|f$,整除有传递性、反身性、互伴性.即:

- (1) 若 $g|f, f|h$, 则有 $g|h$;
- (2) 对任意 $g \in F[x]$, 均有 $g|g$;
- (3) 若 $g|f$ 且 $f|g$, 则必有 $f = cg$, 其中 $c \in F^*$.

定义 4 (1) 设 $f(x) = a_n x^n + \cdots + a_1 x + a_0$ 是一个 F 上的多项式形式, $c \in F$, 则 $f(c) = a_n c^n + \cdots + a_1 c + a_0$ 称为 $f(x)$ 在 c 点的值.

(2) 当 $f(c) = 0$ 时,称 c 为 $f(x)$ 在 F 中的根或零点,也称 c 为 $f(x) = 0$ 的解或根.

关于 $f(x)$ 在一点的值和 $f(x)$ 的根有如下的著名定理.

定理 4 (1) 余数定理 $f(x) = (x - c)q(x) + f(c)$.

(2) 零点定理 $f(c) = 0 \quad (x - c) | f(x)$.

余数定理说明 $f(x)$ 除以 $x - c$ 的余式(数)是 $f(c)$;而零点定理说明 c 为 $f(x)$ 的根 $f(x) = (x - c)q(x)$, 即 $f(x)$ 有因式 $x - c$.

由此可以证明以下重要的结论.

定理 5 域 F 上的 n 次多项式 $f(x)$ 在 F 中最多有 n 个根(重根按重数计入).

推论 设 $f(x) \in F[x], \deg f(x) = n$, 如果 $f(x)$ 在 F 中有多于 n 个不同的根, 则必有 $f(x) = 0$.

从定理 5 和它的推论就得到判断两个多项式形式相等的一个方法.

定理 6 设 $f(x), g(x) \in F[x]$, $\deg f, \deg g < n$. 若在 n 个不同的点 $\alpha_1, \dots, \alpha_n$ 上, 有 $f(\alpha_i) = g(\alpha_i)$, 则必有 $f(x) = g(x)$.

此定理的思想可以扩展在“函数逼近”(读者在数值分析课程中将会学到). 注意定理 6 的条件比定理 5 的结论要强, 事实上, 这只是一中给出 n 个点上的值的方式: (1) 设这 n 个点不同, 这时只要说 $f(x)$ 与 $g(x)$ 在这 n 个点上的函数值相等; (2) n 个点中有“重根”, 于是只能说 $f(x)$ 与 $g(x)$ 在其中 m 个点上的函数值相等, 怎么表示“重根”呢? 若某一个点 α 是 $f(x) - g(x)$ 的二重根, 则 $f(x)$ 与 $g(x)$ 在 α 处不但函数值相等而且导函数值也相等, 即条件变成

$$f(\alpha) = g(\alpha), \text{ 及 } f'(\alpha) = g'(\alpha),$$

因此若只想用函数值给条件, 则必须说那些点不同.

2 掌握求最大公因子的辗转相除法, 理解 Bezout 等式的本质, 给定两个多项式 f, g , 会求 Bezout 等式

定义 5 设 $f, g \in F[x]$.

(1) 若 $h(x) \in F[x]$ 满足 $h \mid f, h \mid g$, 则称 $h(x)$ 为 $f(x)$ 与 $g(x)$ 的公因式.

(2) 若 $d(x) \in F[x]$ 是 f 和 g 的公因式, 且是 f 和 g 的任意一个公因式的倍式, 则称 $d(x)$ 为 $f(x)$ 与 $g(x)$ 的最大公因式.

f 与 g 的首一最大公因式记为 (f, g) .

引理 1 若 $f = gq + r$, 则 $(f, g) = (r, g)$, 其中 $f, g, r \in F[x]$.

由此引理得到求最大公因式的辗转相除法:

$$f = gq_1 + r_1, \quad \deg r_1 < \deg g,$$

$$g = r_1q_2 + r_2, \quad \deg r_2 < \deg r_1,$$

$$r_1 = r_2q_3 + r_3, \quad \deg r_3 < \deg r_2,$$

...

$$r_{s-2} = r_{s-1} q_s + r_s \quad \deg r_s < \deg r_{s-1},$$

$$r_{s-1} = r_s q_{s+1}.$$

所以 $(f, g) = (g, r_1) = \cdots = (r_{s-1}, r_s) = cr_s$.

注意, 因 (f, g) 是首一多项式, 而 r_s 不一定首一, 故其前有一个修正值 c (见 3.3 节例 1).

定理 7 设 $f, g \in F[x]$, 则 $(f, g) = d(x)$ 存在且惟一, 而且存在 $u, v \in F[x]$, 使

$$uf + vg = d \quad (\text{Bezout 等式}).$$

关于 Bezout 等式有几点说明, 请读者留意.

(1) 求 Bezout 等式的方法 应是先找公式, 再代入 q (辗转相除各步得到的商式) 算出 u, v . 这里以辗转相除 3 次得到最大公因式为例说明之.

$$\begin{aligned} \text{由 } (f, g) &= c(r_3) = c(r_1 - q_3 r_2) \\ &= c[(f - q_3 g) - q_3 (g - q_2 r_1)] \\ &= c[f - q_3 g - q_3 (g - q_2 (f - q_1 g))] \\ &= c(1 + q_1 q_3) f - c(q_3 + q_1 q_2 q_3 + q_2) g, \end{aligned}$$

故 $u = c(1 + q_1 q_3), v = -c(q_3 + q_1 q_2 q_3 + q_2)$.

再把 $q_i (i=1, 2, 3)$ 代入, 就可求出 u, v . 千万不要在一开始“回代”时就用具体的 q_i 参加运算, 那样算起来眼花缭乱, 容易出错.

(2) Bezout 等式中的 u, v 是不惟一的. 事实上设 u, v 是满足条件的一组值, 即有恒等式

$$uf + vg = (f, g).$$

上式左边加 fg , 再减去 fg , 值不变. 于是有

$$(u + g) f + (v - f) g = (f, g).$$

所以 $u = u + g, v = v - f$ 也满足要求, 故计算过程要头脑清楚, 找出最简单的 u, v 即可. 不要心不在焉地算, 人为地加大计算量 (见 3.3 节例 5).

(3) Bezout 等式是必要条件, 即若 f, g 有首一的最大公因式

$d(x)$, 则一定有 u, v , 使

$$uf + vg = d(x). \quad (\text{Bezout 等式成立}). \quad (*)$$

但若有 $u_1 f + v_1 g = h(x)$, 则不能断定 $h(x) = (f, g)$. 事实上, 由 $(*)$ 两边乘任意多项式 $s(x)$, 得

$$usf + vsg = d(x)s(x).$$

显然不能判定 $d(x)s(x)$ 仍为 f, g 的最大公因式.

定义 6 设 $f, g \in F[x]$, 若 $(f, g) = 1$, 则称 $f(x)$ 与 $g(x)$ 互素.

定理 8 设 $f, g \in F[x]$, 则

$$(f, g) = 1 \quad \text{存在 } u, v \in F[x], \text{ 使 } uf + vg = 1.$$

上式右边的等式也称为(关于互素的)Bezout 等式. 与定理 7 不同的是对于互素的两个多项式 f, g , 其 Bezout 等式是充分必要条件. 当然, 这里的 u, v 也不惟一.

关于整除和互素有如下性质:

定理 9 (1) 若 $f | gh$, 且 $(f, g) = 1$, 则 $f | h$.

(2) 若 $f_1 | g, f_2 | g$, 且 $(f_1, f_2) = 1$, 则 $f_1 f_2 | g$.

(3) 若 $(f, g) = 1$, 且 $(f, h) = 1$, 则 $(f, gh) = 1$.

3 理解惟一析因定理, 掌握 $C[x]$ 和 $R[x]$ 上的分解理论并会应用

定理 10 $F[x]$ 是惟一析因整环, 即任意一非常数多项式 $f \in F[x]$ 均可表为一些不可约的多项式的乘积

$$f = p_1 p_2 \cdots p_s,$$

且若不计常数倍及 p_i 的次序, 分解是惟一的. 或写为

$$f = cp_1^{n_1} \cdots p_s^{n_s}, \quad \text{其中每个 } p_i \text{ 不可约且两两互异}.$$

定理 11 复系数多项式 $f(x)$ ($\deg f \geq 1$) 在复数域 C 上总可以惟一分解为一次因式的乘积, 即

$$f(x) = c(x - \alpha)^{n_1} (x - \alpha)^{n_2} \cdots (x - \alpha)^{n_s} \\ (c, \alpha_i \in C, \sum_{i=1}^s n_i = n).$$

引理 2 实系数多项式 $f(x)$ 的复根总是成共轭对出现. 即当 α 是 $f(x)$ 的根时, 则 $\bar{\alpha}$ 也是 $f(x)$ 的根.

定理 12 实系数多项式 $f(x)$ ($\deg f \geq 1$) 在实数域上总可以惟一分解为一次和二次不可约因子之积, 即

$$f(x) = a(x - \alpha_1)^{n_1} \cdots (x - \alpha_s)^{n_s} \cdot (x^2 - b_1x + c_1)^{e_1} \cdots (x^2 - b_tx + c_t)^{e_t}.$$

在这里看, 定理 11 和定理 12 是关于复(实)系数多项式的分解(理论)问题, 把它用在具体的多项式上, 就是实践问题了. 如用在求方阵的特征值上得:

(1) 实方阵的特征多项式在实数域里的分解必是一次因式和二次因式之积, 故实方阵不一定在实数域里相似对角阵(从空间分解的角度看, 就是若 σ 在复数域 C 中可对角化, 则 V 在实数域里只能分解为一维和二维不变子空间的直和), 如实正交矩阵;

(2) 实方阵的特征多项式是实系数多项式, 故其复特征值成共轭对出现, 如实正交矩阵的特征值除 $+1$, -1 外, 还有模为 1 的复数, 如含 $e^{i\theta}$, $e^{-i\theta}$ 这样一对一对的复根.

定理 13 (Vieta 定理) 设

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \\ &= a_n (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n), \end{aligned}$$

则

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = -\frac{a_{n-1}}{a_n},$$

$$\alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \cdots + \alpha_{n-1} \alpha_n = \frac{a_{n-2}}{a_n},$$

.....

$$\alpha_1 \cdots \alpha_{n-1} + \cdots + \alpha_{n-1} \cdots \alpha_n = (-1)^{n-1} \frac{a_1}{a_n},$$

$$\alpha_1 \alpha_2 \cdots \alpha_n = (-1)^n \frac{a_0}{a_n}.$$

这里 $\alpha, \alpha, \dots, \alpha$ 是 $f(x)$ 的 n 个根, Vieta 定理讲的是多项式的根与系数的关系, 是中学讲的二次三项式的 Vieta 定理在 n 维空间的推广.

当 $f(\lambda) = |\lambda I - A|$ 是 A 的特征多项式时, 有 $a_n = 1, a_{n-1} = -(\alpha_{11} + \dots + \alpha_{nn}) = -\operatorname{tr} A, a_0 = (-1)^n \det A$. 记 A 的特征根为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则由上述根与系数的关系有

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = -a_{n-1} = \sum_{i=1}^n \alpha_{ii} = \operatorname{tr} A,$$

$$\lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n = (-1)^n a_0 = \det A.$$

定理 14 设 $f(x) \in Z[x]$,

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$

若 $f\left(\frac{r}{s}\right) = 0$, 其中 $r, s \in Z$, 且 $(r, s) = 1$, 则 $r | a_0, s | a_n$.

这个定理给出: 若 $f(x)$ 有有理根 $\frac{r}{s}$, 则此分数的分子应能整除多项式的常数项, 而分母整除首项.

定理 13 和定理 14 可用在求三次以上特征多项式的根上, 可以先确定根的选取范围, 求出 1 或 2 个根后, 化为二次多项式时, 就可全部解出来了. 如 A 为 4 阶矩阵, $|\lambda I - A| = \lambda^4 + \lambda^3 - 7\lambda^2 - \lambda + 6$, 则由首项分解为 ± 1 , 常数项分解为 $\pm 1, \pm 2, \pm 3, \pm 6$, 故根有 8 种可能, 它们分别为 $\pm 1, \pm 2, \pm 3, \pm 6$, 把 ± 1 代入得 $|\lambda I - A| = 0$, 故特征多项式中有因子 $(\lambda^2 - 1)$, 于是

$$\begin{aligned} |\lambda I - A| &= \lambda^2(\lambda^2 - 1) + \lambda(\lambda^2 - 1) - 6(\lambda^2 - 1) \\ &= (\lambda^2 - 1)(\lambda^2 + \lambda - 6) \\ &= (\lambda^2 - 1)(\lambda + 3)(\lambda - 2). \end{aligned}$$

3.2 典型题讲评

例 1 设 $g(x) = x^2 - 2ax + 2$ 整除 $f(x) = x^4 + 3x^2 + ax + b$, 求 a 和 b , 其中 $a, b \in \mathbb{R}$.

解 做带余除法

$$\begin{array}{r|l}
 g(x) & f(x) \\
 x^2 - 2ax + 2 & \begin{array}{l} x^4 \qquad \qquad + 3x^2 + ax + b \\ x^4 - 2ax^3 \qquad + 2x^2 \\ \hline 2ax^3 \qquad + x^2 + ax + b \\ 2ax^3 - 4a^2x^2 + 4ax \end{array} \\
 \hline
 \end{array}$$

$$(1 + 4a^2)x^2 - 3ax + b$$

$$(1 + 4a^2)x^2 - (2a + 8a^3)x + 2 + 8a^2$$

第 1 次做除法后, 得余式 $2ax^3 + x^2 + ax + b$. 此时若 $a = 0$, 则变为要求 $x^2 + 2 \mid x^2 + b$, 故 $b = 2$; 否则, 又经 3 次做除法后, 要求 $g \mid f$, 则必有

$$-3a = -2a - 8a^3,$$

$$b = 2 + 8a^2.$$

由第 1 个方程得 $8a^2 = 1$, 故 $a = \pm \frac{1}{2}$, 代入第 2 个方程得 $b = 3$.

$$\begin{array}{ll}
 \text{故得} & a = 0, \quad \text{或} \quad a = \pm \frac{1}{2}, \\
 & b = 2. \quad \quad \quad b = 3.
 \end{array}$$

讲评 这样的题, 从理论上说没什么难度, 但要得出正确结果, 就需要仔细计算, 认真对齐上、下各项, 否则也会变成难题.

例 2 设 m, n 和 p 是正整数, 证明 $g(x) = x^2 + x + 1$ 整除 $f(x) = x^{3m} + x^{3n+1} + x^{3p+2}$.

解 方法1 由于

$$(x^3 - 1) = (x - 1)(x^2 + x + 1),$$

$$(x^{3k} - 1) = ((x^3)^k - 1)$$

$$= (x^3 - 1)((x^3)^{k-1} + (x^3)^{k-2} + \cdots + x^3 + 1),$$

故知 $(1 + x + x^2) | (x^3 - 1)$, $(x^3 - 1) | (x^{3k} - 1)$ (对任意整数 k). 于是知 $g = (1 + x + x^2) | x^{3m} - 1$, $g | x^{3n} - 1$, $g | x^{3p} - 1$. 又已知多项式可变形为

$$f(x) = (x^{3m} - 1) + x(x^{3n} - 1) + x^2(x^{3p} - 1) + (1 + x + x^2),$$

由 $(1 + x + x^2)$ 可以整除右边每一项(共4项) 故

$$(1 + x + x^2) | f(x).$$

讲评 此题主要用到两个因式分解的公式: 立方差公式和 k 方差公式, 即

$$(a^3 - b^3) = (a - b)(a^2 + ab + b^2),$$

$$(a^k - b^k) = (a - b)(a^{k-1} + a^{k-2}b + \cdots + ab^{k-2} + b^{k-1}).$$

在第1个公式中取 $a = x$, $b = 1$, 在第2个公式中取 $a = x^3$, $b = 1^3$, 即得到我们要用的公式, 从而证出结论. 这需要对这些公式很熟, 看到 $1 + x + x^2$ 就会联想到它是 $x^3 - 1$ 的因式; 看到 $x^{3m} - 1$ 就会想到它含有 $x^3 - 1$ 这个因式, 等等. 此题也可用多项式的根来证(即用零点定理).

方法2 由于 $g(x)$ 的两个根为 $\omega = \frac{-1 + 3i}{2}$, $\bar{\omega} = \frac{-1 - 3i}{2}$,

且 $\omega = \omega^3$, 又由

$$x^3 - 1 = (x - 1)(x^2 + x + 1) = (x - 1)(x - \omega)(x - \bar{\omega}),$$

故 $\omega (i=1, 2)$ 都是3次单位根. 即有 $\omega^3 = 1$. 而

$$f(\omega) = \omega^{3m} + \omega^{3n+1} + \omega^{3p+2} = 1 + \omega + \omega = 0,$$

$$f(\bar{\omega}) = f(\omega^2) = 1 + \omega^2 + \omega = 0.$$

故知 ω 与 $\bar{\omega}$ 都是 $f(x)$ 的根, 由零点定理知

$$\begin{aligned}
 f(x) &= (x - \omega)(x - \omega)h(x) \\
 &= (x^2 - (\omega + \omega)x + \omega\omega)h(x) \\
 &= (x^2 + x + 1)h(x),
 \end{aligned}$$

所以 $1 + x + x^2 \mid f(x)$.

例 3 设 $f(x) = 2x^5 - 5x^3 + 8x$, 用综合除法求 $f(-3)$.

解 由余数定理(定理 4)知 $f(-3)$ 等于 $f(x)$ 除以 $(x+3)$ 得到的余式. 因除式为一次式 $x - a$ 的形式, 故可用综合除法简化求之:

$$\begin{array}{r|rrrrrr}
 -3 & 2 & 0 & -5 & 0 & 8 & 0 \\
 & & -6 & 18 & -39 & 117 & -375 \\
 \hline
 & & -6 & 13 & -39 & 125 & -375
 \end{array}$$

故

$$f(x) = (x+3)(2x^4 - 6x^3 + 13x^2 - 39x + 125) - 375,$$

而 $f(-3) = r(x) = -375$.

讲评 综合除法仅用在除式为 $x - a$ 时, 省略 x 后除数写为 a . 本题要求 $f(-3)$, 故知除式应为 $x - (-3) = x + 3$. 这一点要理解清楚, 读者可从做带余除法的算式中分析、总结、理解. 数学是理解地记忆, 不用死记硬背.

例 4 用综合除法把 $f(x) = x^4 - 2x^3 + 3$ 表成 $x - 2$ 的方幂.

解 $f(x) = x^3(x - 2) + 3$, 而由综合除法知

$$\begin{aligned}
 x^3 &= (x^2 + 2x + 4)(x - 2) + 8, \\
 x^2 + 2x + 4 &= (x + 4)(x - 2) + 12, \\
 (x + 4) &= (x - 2) + 6,
 \end{aligned}$$

故有

$$\begin{aligned}
 f(x) &= (x - 2)^4 + 6(x - 2)^3 + 12(x - 2)^2 + \\
 &\quad 8(x - 2) + 3.
 \end{aligned}$$

讲评 上述综合除法的算式可以连续进行, 注意每次除后得

到的最后一个数字为余数(竖式中用虚线割开) .

$$\begin{array}{r|rrrr}
 2 & 1 & -2 & 0 & 0 & 3 \\
 & 1 & \frac{2}{0} & \frac{0}{0} & \frac{0}{0} & \frac{3}{3} \\
 & 1 & \frac{2}{2} & \frac{4}{4} & \frac{8}{8} & \\
 & 1 & \frac{2}{4} & \frac{8}{12} & & \\
 & & \frac{2}{6} & & &
 \end{array}$$

于是知

$$\begin{aligned}
 f(x) &= (x-2)q(x) + 3, \\
 q(x) &= (x-2)q(x) + 8, \\
 q(x) &= (x-2)q(x) + 12, \\
 q(x) &= (x-2) + 6,
 \end{aligned}$$

故得

$$f(x) = (x-2)^4 + 6(x-2)^3 + 12(x-2)^2 + 8(x-2) + 3.$$

例 5 求次数最低的多项式 $f(x)$, 使得 $f(x)$ 被 $(x-1)^2$ 除时余式为 $2x$, 被 $(x-2)^3$ 除时, 余式为 $3x$.

解 由题意知

$$\begin{aligned}
 f(x) &= q(x-1)^2 + 2x = q(x-2)^3 + 3x \\
 &= q[(x-1)-1]^3 + 3x \\
 &= q[(x-1)^3 - 3(x-1)^2 + 3(x-1) - 1] + 3x.
 \end{aligned}$$

显然, 最后一个式子中前两项都可以被 $(x-1)^2$ 整除, 故若要求 $(x-1)^2$ 除 $f(x)$ 余 $2x$, 必有 $(x-1)^2 \mid [q(3(x-1)-1) + x]$, 由此知 $\deg q \geq 1$.

设 $q = (ax+b)$, 则应有

$$(3x - 4)(ax + b) + x = 3a(x - 1)^2.$$

比较两边各项的系数(主要是一次项和常数项,二次项我们已令其相等)得

$$\begin{aligned} -4a + 3b + 1 &= -6a, & \text{即} & a = 4, \\ 3a &= 4b, & & b = -3. \end{aligned}$$

故知

$$\begin{aligned} f(x) &= (x - 1)^3(4x - 3) + 3x \\ &= 4x^4 - 27x^3 + 66x^2 - 65x + 24. \end{aligned}$$

讲评 (1) 这个题是人人会做的,因为由题设很容易写出第一个式子.那么,无论对 q , 或对 q 用待定系数法都可以做出来.所以说不会的同学,实质上是“怕繁”,从心理上不想做.故我们常说不是会不会的问题,而是“懒不懒”的问题,只要肯动手,是可以做出来的.

(2) 上述解法用了个小技巧,即用中学学的“差立方”公式把 $(x - 2)^3$ 写成 $[(x - 1) - 1]^3$, 与前一个带余除法互应,使问题得以化简,故我们建议同学,遇到计算题,第一要“不怕繁”,耐心、认真地做,争取一遍做对;第二要“多动脑筋”,看看有没有“巧”思路,简化计算,使繁变简.

例 6 设非零实系数多项式 $f(x)$ 满足 $f(f(x)) = f^k(x)$, 其中 k 是给定正整数,求多项式 $f(x)$.

解 设

$$f(x) = a_n x^n + \cdots + a_1 x + a_0,$$

则

$$f(f(x)) = a_n f^n(x) + a_{n-1} f^{n-1}(x) + \cdots + a_1 f(x) + a_0.$$

而

$$f^k(x) = (a_n x^n + \cdots + a_1 x + a_0)^k,$$

由已知条件 $f(f(x)) = f^k(x)$, 故应有(等号两边的多项式的次数相等)

$$\deg f^k(x) = \deg f(f(x)),$$

即得 $nk = n^2$, 所以 $n = 0$ 或 $n = k$.

当 $n = 0$ 时, 有

$$f(x) = a_0, \quad f(f(x)) = a_0, \quad f^k(x) = a_0^k.$$

故当 $k \neq 1$ 时, 有 $a_0 = 1$, 即 $f(x) = 1$.

当 $n = k$ 时, 则 $f(x) = a_k x^k + \cdots + a_1 x + a_0$,

$$f^k(x) = (a_k x^k + \cdots + a_1 x + a_0)^k, \quad (1)$$

$$f(f(x)) = a_k f^k(x) + a_{k-1} f^{k-1}(x) + \cdots + a_1 f(x) + a_0. \quad (2)$$

比较 $f^k(x) = f(f(x))$ 两边的首项, 知有

$$a_k^k = a_k^{k+1}, \quad \text{故 } a_k = 1.$$

代回(1), (2)两式中, 知有

$$f(f(x)) = f^k(x) + a_{k-1} f^{k-1}(x) + \cdots + a_1 f(x) + a_0,$$

故得

$$a_{k-1} f^{k-1}(x) + \cdots + a_1 f(x) + a_0 = 0. \quad (3)$$

上式算出来应是 x 的多项式, 其值恒为 0, 故首项应为 0, 得

$$a_{k-1} = 0.$$

由此又知有 $a_{k-2} f^{k-2}(x) + \cdots + a_1 f(x) + a_0 = 0$.

继续做下去(不断讨论首项的情况), 得

$$a_{k-1} = 0, \quad a_{k-2} = 0, \quad \cdots, \quad a_1 = 0, \quad a_0 = 0.$$

于是知

$$f(x) = x^k.$$

讲评 (1) 此题由 $f(f(x)) = f^k(x)$ 建立关系, 乍一看有点“神”, 也许以为计算量将非常大(用待定系数法定系数), 但当一“走出”设 $f(x) = a_n x^n + \cdots + a_1 x + a_0$ 这一步后, 其实思路就有了, 接下来就会看到“走起来”比“只想”要容易得多.

(2) 用“首项”讨论问题是多项式形式这一章问题的一个特点. 对所有的系数, 不仅“繁”且“不好说”时, 用“首项”说事就方便许多. 本题可以说是很典型的一个题. 多次用“首项”解决问题.(或说主要用首项解决问题.)

(3) 当一个多项式 $g(x) = b_m x^m + \cdots + b_1 x + b_0 = 0$ 时,立刻得出所有的系数 $b_i = 0 (0 \leq i \leq m)$,但对(3)式,因其中 $f(x), f^2(x), \cdots, f^{k-1}(x)$ 这些项都是 x 的多项式,它们展开后得到一个关于 x 的多项式.显然除首项系数外,其余的 $a_{k-2}, \cdots, a_1, a_0$ 不是该多项式的系数,故我们“不能”由(3)式直接推得 $a_i = 0 (0 \leq i \leq k-1)$.要一次一次地只讨论首项(必为 0)来得出结论.

例 7 设 $f(x) = x^4 + 2x^3 - x^2 - 4x - 2, g(x) = x^4 + x^3 - x^2 - 2x - 2$, 求 f, g 的首一最大公因式 (f, g) 及 u, v , 使得

$$uf + vg = (f, g).$$

解 用辗转相除法求 (f, g) :

$$q = x + 1 \left| \begin{array}{r|l} g(x) & f(x) \\ \hline x^4 + x^3 - x^2 - 2x - 2 & x^4 + 2x^3 - x^2 - 4x - 2 \\ \hline \frac{x^4}{x^4} - \frac{2x^2}{x^2} & \frac{x^4 + x^3 - x^2 - 2x - 2}{x^4 + 2x^3 - x^2 - 2x} \\ \hline x^3 + x^2 - 2x - 2 & r_1 = x^3 - 2x \\ \hline \frac{x^3}{x^3} - \frac{2x}{x} & \\ \hline r_2 = x^2 - 2 & = x(x^2 - 2) \end{array} \right| q = 1$$

故 $(f, g) = x^2 - 2 = r_2$ (注意当算出 $r_2 = x^2 - 2$ 时,再看 $r_1 = x r_2$, 则辗转相除到此结束,不用再除下一步).由

$$f = gq_1 + r_1, \quad g = r_1 q + r_2,$$

知

$$\begin{aligned} (f, g) &= r_2 = (g - r_1 q) = (g - (f - gq_1)q) \\ &= -qf + (1 + q q_1)g, \end{aligned}$$

故得

$$u = -q = -x - 1,$$

$$v = (1 + q q_1) = 1 + x + 1 = x + 2.$$

讲评 (1) 这道题即使不要求计算 (f, g) 或者说 (f, g) 可由因式分解直接看出,则为了求 Bezout 等式,也要用辗转相除法去找各个 q_i ,从而计算 u, v .因为 u, v 恰由商式的代数运算得到.因此,读者就不必去用别的方法找 (f, g) ,反正 q 总是要用到的,故

干脆从开始就认真做辗转相除.

(2) 计算时注意降幂排列, 若缺少某个幂次, 要空出位置, 以便上下对齐, 便于计算. 应该说, 这些题的设计总是由原作者精心策划的, 哪一步一旦计算失误, 计算量将加大, 会“越算越繁”, 务请读者注意.

(3) u, v 是不惟一的, 如上步骤(给出的解答)得出的 u, v 应是次数最低的, 也是最简单的.

例 8 设 $f(x) = x^3 + (1+t)x^2 + 2x + 2s$, $g(x) = x^3 + tx^2 + s$ 的最大公因式是一个二次多项式, 求 t, s 的值.

解 由 $f(x) - g(x) = x^2 + 2x + s$, 所以 $x^2 + 2x + s$ 是 f, g 的最大公因式(设 $(f, g) = d(x)$, 则 $d(x) | f, d(x) | g$, 故 $d(x) | x^2 + 2x + s$). 故以下只要看当 t, s 为何值时, $x^2 + 2x + s | f$ 或 $x^2 + 2x + s | g$ 即可. 因

$$x^2 + 2x + s \left| \begin{array}{r} g \\ x^3 + tx^2 \\ x^3 + 2x^2 + sx \\ \hline (t-2)x^2 - sx + s \\ (t-2)x^2 + 2(t-2)x + s(t-2) \end{array} \right| \quad q = x + (t-2)$$

故要 $x^2 + 2x + s | g(x)$, 必有

$$\begin{array}{llll} -s = 2(t-2), & \text{即} & s = 0, & \text{或} & s = -2, \\ s = s(t-2). & & t = 2. & & t = 3. \end{array}$$

讲评 (1) 从理论来说, 这样的题可用辗转相除法来做. 当某个余式 r_i 是二次多项式时, 应有 $r_i | r_{i-1}$. 从而可定出 t, s 的值. 本题解法采用了一个技巧, 即通过直观看 $f - g$ 就是一个二次多项式, 于是改用带余除法定 t, s 的值, 简便了许多.

(2) 在完成第一次除法运算后, 可讨论: 若余式中 x^2 的系数为 0, 则其余两项系数也应为 0, 故有 $t = 2$ 时, $s = 0$, 这个结果已包

含在最后的两个结果之中,故不讨论也可以(没有丢解)。

例 9 设 $f(x), g(x) \in F[x]$, $a, b, c, d \in F$. 若 $ad - bc \neq 0$, 证明 $(af(x) + bg(x), cf(x) + dg(x)) = (f(x), g(x))$.

证 记 $af + bg = f_1(x)$, $cf(x) + dg(x) = g_1(x)$, 则由

$$(f, g) \mid f, \quad (f, g) \mid g,$$

得 $(f, g) \mid af + bg = f_1, \quad (f, g) \mid cf + dg = f_2,$

故 $(f, g) \mid (f_1, g_1).$

又线性方程组

$$\begin{aligned} af + bg &= f_1, \\ cf + dg &= g_1, \end{aligned} \tag{1}$$

的系数行列式不为 0(因 $ad - bc \neq 0$), 故有惟一解, 设为

$$f = a_1 f_1 + b_1 g_1, \quad g = c_1 f_1 + d_1 g_1.$$

于是同上讨论有

$$(f_1, g_1) \mid f, \quad \text{且} (f_1, g_1) \mid g, \quad \text{故} (f_1, g_1) \mid (f, g).$$

故得 $(f_1, g_1) = (f, g).$

讲评 本题关键的一步, 是会利用条件 $ad - bc \neq 0$, 把 f, g 用 $f_1 = af + bg$ 和 $g_1 = cf + dg$ 表示. 因为大部分人会做前半部, 即可以看出 $(f, g) \mid f_1, (f, g) \mid g_1$. 接下来怎么办? 显然, f_1 和 g_1 都是 f 和 g 的线性组合(函数), 故应该想到能否用 f_1 和 g_1 的线性(组合)函数去表示 f, g , 这时就会联想到已知条件中 $ad - bc \neq 0$, 这恰是保证线性方程组(1)有解的条件, 这样问题就解决了。

例 10 设 $f(x) = d(x) f_1(x)$, $g(x) = d(x) g_1(x)$, 且 $f(x), g(x)$ 不全为零, 证明 $d(x)$ 是 $f(x), g(x)$ 的一个最大公因式当且仅当 $(f_1(x), g_1(x)) = 1$.

证 (必要性) 当 $d(x)$ 是 f, g 的最大公因式时, 有 u, v 使

$$uf + vg = d(x). \tag{1}$$

由 f, g 不全为 0, 故 $d(x) \neq 0$, 且由 $f = d(x) f_1, g(x) = d(x) g_1$, 于是(1)式两边同除 $d(x)$, 得

$$uf_1 + vg_1 = 1 \quad \text{由定理 8} \quad (f_1, g_1) = 1.$$

(充分性) 由 $(f_1, g_1) = 1$, 知存在 $u(x), v(x)$, 使

$$uf_1 + vg_1 = 1.$$

上式两边同乘 $d(x)$, 得

$$u(x)f + v(x)g = d(x). \quad (2)$$

由题设知 $d(x) | f, d(x) | g$, 所以 $d(x)$ 是 f, g 的公因式. 下面要证它是最大公因式, 即是任意公因式的倍式. 若有 $d_1(x)$ 满足 $d_1(x) | f, d_1(x) | g$, 则由(2)式知 $d_1(x) | d(x)$, 故 $d(x)$ 确为 f 和 g 的最大公因式.

讲评 命题的必要条件证明较容易, 而充分性的证明要避免的错误是: 不能由公式(2)直接得结论, 正如我们在理论部分强调的 Bezout 等式(定理 7)不是充要条件. 故由公式(2)之后必须要证以下两点: (1) $d | f, d | g$, 即 $d(x)$ 确为 f, g 的公因式;

(2) 若有 $d_1 | f, d_1 | g$, 则 $d_1 | d$, 即 $d(x)$ 为一切公因式的倍式.

3.3 考试题评析

例 1 设 $f(x) = \frac{1}{4}x^4 + \frac{1}{4}x^2 - \frac{1}{2}, g(x) = \frac{1}{4}x^3 + \frac{1}{4}x^2 - \frac{1}{2}x - \frac{1}{2}$, 则其首一最大公因式 $(f(x), g(x))$ 等于什么? 并求 $u(x), v(x)$, 使得 $u(x)f(x) + v(x)g(x) = (f(x), g(x))$. 又问 $u(x), v(x)$ 互素吗?

解 用辗转相除法求 (f, g) , 有

$$\begin{array}{c}
 q_2 = \\
 \frac{1}{4}x + \frac{1}{4}
 \end{array}
 \left| \begin{array}{c}
 g(x) \\
 \frac{1}{4}x^3 + \frac{1}{4}x^2 - \frac{1}{2}x - \frac{1}{2} \\
 \hline
 \frac{1}{4}x^3 \qquad - \frac{1}{4}x \\
 \hline
 \frac{1}{4}x^2 - \frac{1}{4}x - \frac{1}{2} \\
 \hline
 \frac{1}{4}x^2 \qquad - \frac{1}{4} \\
 \hline
 r_2 = -\frac{1}{4}x - \frac{1}{4} \\
 = -\frac{1}{4}(x+1)
 \end{array} \right|
 \begin{array}{c}
 f(x) \\
 \frac{1}{4}x^4 \qquad + \frac{1}{4}x^2 - \frac{1}{2} \\
 \hline
 \frac{1}{4}x^4 + \frac{1}{4}x^3 - \frac{1}{2}x^2 - \frac{1}{2}x \\
 \hline
 -\frac{1}{4}x^3 + \frac{3}{4}x^2 + \frac{1}{2}x - \frac{1}{2} \\
 \hline
 -\frac{1}{4}x^3 - \frac{1}{4}x^2 + \frac{1}{2}x + \frac{1}{2} \\
 \hline
 r_1 = x^2 - 1 \\
 = (x+1)(x-1)
 \end{array}
 \left| \begin{array}{c}
 q_1 = x - 1
 \end{array} \right.$$

所以 $(f, g) = (x+1) = -4r_2$.再由

$$f = gq_1 + r_1, \quad g = r_1q + r_2,$$

得

$$\begin{aligned}
 r_2 &= g - r_1q = g - (f - gq_1)q, \\
 (f, g) &= -4r_2 = -4qf - 4(1 + qq)g.
 \end{aligned}$$

故知

$$\begin{aligned}
 u(x) &= -4q = x + 1, \\
 v(x) &= -4(1 + qq) = -x^2 - 3.
 \end{aligned}$$

又显然 $(u, v) = 1$, 故 u, v 互素 .

评析 (1) 此题中给定的两个多项式的系数都是分数, 故使计算变得麻烦, 但用心做, 避免错误发生, 只要辗转相除两次也就得出结果. $r_2 \mid r_1$, 这一步直观看出就可以了, 不用再进行除法算式.

(2) 再强调一遍求 u, v 的方法. 就是先导出 Bezout 等式的公式, 或者说先导出 u, v 的计算公式(用商式的代数运算给出), 再把具体的 q_i 值代入, 算出 $u(x), v(x)$.

(3) 这里 u, v 是具体多项式, 可以直接看 $(u, v) = ?$, 事实上, 从理论上可以证明: 对任意 $f, g \in F[x]$, 若有 $uf + vg = (f, g)$, 则 $(u, v) = 1$.

例 2(思考题) 设 $d(x) = (f(x), g(x))$, 试证明在 Bezout 等

式 $uf + vg = d(x)$ 中可以选取 u, v 使得

$$\deg u < \deg g(x), \deg v < \deg f(x).$$

证 注意 Bezout 等式中的 u, v 是不惟一的. 由辗转相除法知总可以找到 u, v 使

$$uf + vg = d(x). \quad (1)$$

若(1)式中的 u 满足 $\deg u > \deg g(x)$, 则由带余除法知, 存在 q, r 使 $u = qg + r$, 其中 $\deg r < \deg g(x)$. 代入(1)式有

$$(qg + r)f + vg = rf + (v + qf)g = d(x).$$

由于 $\deg(rf) < \deg(fg)$, 故 $\deg(v + gf)g < \deg(fg)$, 于是知此时也有 $\deg(v + gf) < \deg f$. 故取

$$u_1 = r(x), \quad v_1 = v + gf,$$

就有

$$u_1 f + v_1 g = d(x), \quad \text{且} \quad \deg u_1 < \deg g, \quad \deg v_1 < \deg f.$$

评析 此题主要理解 Bezout 等式中 u, v 的选取不惟一. 因为对于含 fg 的这一项, 它即可合于前(f 中)也可合于后(g 中). 证明过程用的是讨论的方法(或反设法), 即只要找到一个不满足条件的 u (或 v), 利用带余除法就可以得到一个满足条件的 u (或 v)来. 故思路并不难想到.

例3 已知 $f(x), g(x) \in F[x]$, 且不全为零. 若存在 $u(x), v(x) \in F[x]$, 使得 $uf + vg = (f, g)$, 试证 $u(x), v(x)$ 必互素.

证 由题设有

$$uf + vg = (f, g).$$

由 f, g 不全为零, 故 $(f, g) \neq 0$. 上式两边同除 (f, g) 得

$$\frac{f}{(f, g)}u + \frac{g}{(f, g)}v = 1,$$

又 $(f, g) \mid f, (f, g) \mid g$, 故上式中 u, v 前的系数都是多项式. 于是, 由定理8知, $(u, v) = 1$, 即 $u(x), v(x)$ 互素.

评析 像这样的题, 大部分考生可以证出来, 但怎样用简洁的

语言,清晰的思路写出来是重要的.上述证明中最主要的一步是从 Bezout 等式出发,两边同除以 (f, g) .想到这一步就“大功告成”了,其他只要进行讨论、论证,于是就从用定理 7 转化为用定理 8,这就是思路.这里不提倡用“反证法”,明明是 $(u, v) = 1$,却要假设它不互素,再证矛盾.从思维上来说有点“绕弯子”.把简单的事,变得复杂.

例 4(思考题) 试求七次多项式,使 $f(x) + 1$ 能被 $(x - 1)^4$ 整除,而 $f(x) - 1$ 能被 $(x + 1)^4$ 整除.

解 由 $x = 1$ 是 $f(x) + 1$ 的 4 重根,所以 $x = 1$ 是 $(f(x) + 1)' = f'(x)$ 的 3 重根.同理知 $x = -1$ 是 $(f(x) - 1)' = f'(x)$ 的 3 重根.又因为 $\deg f(x) = 7$,故 $\deg f'(x) = 6$.于是,可设

$$f'(x) = a(x - 1)^3(x + 1)^3 = a(x^6 - 3x^4 + 3x^2 - 1) \quad (a \text{ 待定}).$$

积分得

$$f(x) = a \left(\frac{1}{7} x^7 - \frac{3}{5} x^5 + x^3 - x + b \right). \quad (1)$$

又由题设知 $f(1) + 1 = 0$, $f(-1) - 1 = 0$,故知 $f(1) = -1$, $f(-1) = 1$.

代入(1)式中,得关于 a, b 的二元一次方程组

$$\begin{aligned} a \left(\frac{1}{7} - \frac{3}{5} + b \right) &= -1, \\ a \left(-\frac{1}{7} + \frac{3}{5} + b \right) &= 1, \end{aligned} \quad \text{从而得} \quad \begin{aligned} a &= \frac{35}{16}, \\ b &= 0. \end{aligned}$$

于是知

$$f(x) = \frac{5}{16} x^7 - \frac{21}{16} x^5 + \frac{35}{16} x^3 - \frac{35}{16} x.$$

评析 此解法巧妙地利用“重根”的性质:(1)若 α 是 $g(x)$ 的 m 重根,则 α 是 $g'(x)$ 的 $(m - 1)$ 重根;(2)若 α 是 $g(x)$ 的 m 重根,则有 $g(x) = (x - \alpha)^m h(x)$.使得计算大大简化(只需定两个未知量 a, b).

若设 $f(x) = (x-1)^4 q_1 - 1 = (x+1)^4 q_2 + 1$, 再待定 q_1 (或 q_2) 也可做, 但显然计算量要大许多 (q_1 和 q_2 都是三次多项式, 故有 4 个待定系数).

例 5 设 $f(x) = x^5 + 2x^4 - 3x^2 - x + 1$, $g(x) = x^3 + 2x^2 - x - 2$, 则其首一最大公因式 $(f, g) = \underline{\hspace{2cm}}$, 且有 $u(x) = \underline{\hspace{2cm}}$, $v(x) = \underline{\hspace{2cm}}$, 使得

$$u(x)f(x) + v(x)g(x) = (f(x), g(x)).$$

解 由辗转相除法, 知

$$q_2 = -\frac{1}{3}x - \frac{2}{3} \left| \begin{array}{r} g(x) \\ x^3 + 2x^2 - x - 2 \\ \hline x^3 - x \\ \hline 2x^2 - 2 \\ \hline 2x^2 - 2 \\ \hline r_2 = 0 \end{array} \right| \quad \begin{array}{r} f(x) \\ x^5 + 2x^4 - 3x^2 - x + 1 \\ \hline x^5 + 2x^4 - x^3 - 2x^2 \\ \hline x^3 - x^2 - x + 1 \\ \hline x^3 + 2x^2 - x - 2 \\ \hline r_1 = -3x^2 + 3 \end{array} \quad q_1 = x^2 + 1$$

故

$$\begin{aligned} (f, g) &= (x^2 - 1) = -\frac{1}{3}r_1 = -\frac{1}{3}(f - gq_1) \\ &= -\frac{1}{3}f + \frac{1}{3}gq_1. \end{aligned}$$

于是得

$$u = -\frac{1}{3}, \quad v = \frac{1}{3}(x^2 + 1).$$

评析 应该说这是几年来类似的考题中最容易的一道题, 原本想作为填空题, 计算量不要太大, 故设计了题中的多项式的系数, 可没想到许多同学在这道题上却花了许多功夫, 得出许多不同的答案

如: 1) $u = \frac{1}{6}x, \quad v = -\frac{1}{6}x^3 - \frac{1}{6}x + \frac{1}{2};$

(2) $u = -\frac{1}{3}x - 1, \quad v = \frac{1}{3}(x^3 + 3x^2 + x);$

(3) $u = x + \frac{5}{3}, \quad v = -x^3 - \frac{5}{3}x^2 - x + \frac{4}{3}.$

经验证,这些答案都是对的,都满足 $uf + vg = (f, g)$. 而另一些答案是错的,如:

$$\textcircled{1} \quad u = \frac{1}{3}x, \quad v = -\frac{1}{3}x^3 - \frac{1}{3}x + 1, \quad \text{使 } uf + vg = 2(f, g);$$

$$\textcircled{2} \quad u = \frac{1}{6}x + \frac{1}{3}, \quad v = -\frac{1}{6}x^3 - \frac{1}{3}x^2 - \frac{1}{6}x + \frac{1}{6}, \quad \text{使 } uf + vg = 0;$$

$$\textcircled{3} \quad u = -\frac{1}{9}(x+2), \quad v = \frac{1}{9}(x^3 + 2x^2 + x - 1), \quad \text{使 } uf + vg = 0.$$

等等. 为什么这些同学在得出 $(f, g) = -\frac{1}{3}f + \frac{1}{3}g$ 后, 不直接写出答案, 而要另想办法找 u, v 呢? 我猜想他们认为 $u = -\frac{1}{3}$ 太简单了, 不像个多项式.

这里我说一下那么多答案是怎么出来的: 由于

$$f(x) = (x^2 - 1)(x^3 + 2x^2 + x - 1) \quad \text{记为} \quad d(x)f_1(x),$$

$$g(x) = (x^2 - 1)(x + 2) \quad \text{记为} \quad d(x)g_1(x),$$

则在 Bezout 等式

$$-\frac{1}{3}f + \frac{1}{3}g = d(x) = (f, g) \quad (*)$$

的两边加上任意倍的 $(d(x)f_1(x)g_1(x) - d(x)f_1(x)g_1(x))$, 其值均不变. 如加上 $\frac{1}{6}df_1g_1 - \frac{1}{6}df_1g_1$ 得 (注意 $df_1 = f, dg_1 = g$)

$$-\frac{1}{3} + \frac{1}{6}g_1 \quad f + \frac{1}{3}g_1 - \frac{1}{6}f_1 \quad g = (f, g).$$

这就得到(1)中的

$$u = -\frac{1}{3} + \frac{1}{6}(x+2) = \frac{1}{6}x,$$

$$v = \frac{1}{3}(x^2 + 1) - \frac{1}{6}f_1 = -\frac{1}{6}x^3 - \frac{1}{6}x + \frac{1}{2}.$$

若(*)式两边加上 $-\frac{1}{3}df_1g_1 + \frac{1}{3}df_1g_1$, 得

$$-\frac{1}{3} - \frac{1}{3}(x+2)f +$$

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

就得到(2)中的 u, v . 同理若在(*)式两边加上 $df_1g_1 - df_1g_1$ 得

$$-\frac{1}{3}f + df_1g_1 - df_1g_1 + \frac{1}{3}qg = (f, g),$$

即有 $-\frac{1}{3} + g_1f + \frac{1}{3}q_1 - f_1g = (f, g).$

故 $u = -\frac{1}{3} + g_1, v = \frac{1}{3}q_1 - f_1$ 这就是(3)中 u, v 值.

事实上还有, 对任意 $h(x) \in F[x]$

$$-\frac{1}{3}f + h(x)d(x)f_1(x)g_1(x) - h(x)d(x)f_1(x)g_1(x) +$$

$$\frac{1}{3}qg = d(x)$$

成立. 于是得(左边第1与第2项合并, 第3与第4项合并)

$$-\frac{1}{3} + h(x)g_1(x)f + \frac{1}{3}q_1 - h(x)f_1(x)g = d(x),$$

即 $u(x) = -\frac{1}{3} + hg_1, v = \frac{1}{3}q_1 - hf_1$. 由此大家可以看到 u, v 的选取是多么的广泛(不惟一).

第4章 线性空间

4.1 理论脉络

这里所说的线性空间不包含读者已在课程“几何与代数(1)”中学过的线性空间的内容(称为线性空间(1)):即线性空间的定义及性质;线性空间中向量的线性相关及线性无关;线性空间的维数,基(向量)以及向量的坐标等(相当于《理工科代数基础》的第3章前3节).对于线性空间(1)的内容我们这里只强调一下:

(1) 关于形式记法

设 $\varepsilon_1, \dots, \varepsilon_n; \eta_1, \dots, \eta_r$ 是 $V_n(F)$ 的两组基, 则从基 $\varepsilon_1, \dots, \varepsilon_n$ 到基 $\eta_1, \dots, \eta_r$ 的过渡矩阵 $C = (c_{ij})$ 是这样定义的: $C = (c_{ij})$ 的第 j 列是 η_j 在基 $\varepsilon_1, \dots, \varepsilon_n$ 上的坐标列, 即 $\eta_j = \sum_{i=1}^n c_{ij} \varepsilon_i$. 形式地记为

$$(\eta_1, \dots, \eta_r) = (\varepsilon_1, \dots, \varepsilon_n) C. \quad (*)$$

这种记法满足:

- ① $[(a_1, \dots, a_n)A]B = (a_1, \dots, a_n)AB$;
- ② $(a_1, \dots, a_n)A + (e'_1, \dots, e'_n)A = (a_1 + e'_1, \dots, a_n + e'_n)A$;
- ③ $(a_1, \dots, a_n)(A+B) = (a_1, \dots, a_n)A + (a_1, \dots, a_n)B$.

其证明见北京大学编写的《高等代数》. 由于 ε, η 是任意线性空间中的向量, 故 $(*)$ 不能当成矩阵乘法用, 除非 ε, η 是 n 维数组向量. 形式记法只能在满足上述3条性质时使用.

(2) 关于过渡矩阵

① 过渡矩阵是可逆矩阵的证明: 记 $C = (c_1, \dots, c_n)$. 显然, $\det C \neq 0$ C 的列向量线性无关.

$$\begin{aligned}
 & \text{而 } k_1 \eta_1 + \cdots + k_n \eta_n = (\eta_1, \cdots, \eta_n) \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} \\
 & = ((\varepsilon_1, \cdots, \varepsilon_n) C) \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = (\varepsilon_1, \cdots, \varepsilon_n) C \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix},
 \end{aligned}$$

所以得

$$k_1 \eta_1 + \cdots + k_n \eta_n = 0 \quad C \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = 0.$$

上式即是说: $\eta_1, \cdots, \eta_n$ 线性无关的充分必要条件是 C 的列线性无关. 而 $\eta_1, \cdots, \eta_n$ 是 $V_n(F)$ 的基, 它们线性无关, 故 $a_1, \cdots, a_n$ 亦线性无关, 于是知 C 可逆.

② 当 C 是从基 $\varepsilon_1, \cdots, \varepsilon_n$ 到基 $\eta_1, \cdots, \eta_n$ 的过渡矩阵时, 则从基 $\eta_1, \cdots, \eta_n$ 到基 $\varepsilon_1, \cdots, \varepsilon_n$ 的过渡矩阵是 C^{-1} . 这只要在 $(*)$ 式两边同乘 C^{-1} , 即得

$$\begin{aligned}
 (\eta_1, \cdots, \eta_n) C^{-1} &= [(\varepsilon_1, \cdots, \varepsilon_n) C] C^{-1} = (\varepsilon_1, \cdots, \varepsilon_n) (C C^{-1}) \\
 &= (\varepsilon_1, \cdots, \varepsilon_n),
 \end{aligned}$$

其中第 2 个等号是利用形式记法的第 1 条性质(见 4.2 节例 11).

(3) 介绍求 $A^{-1}B$ 的方法: 由于

$$\begin{aligned}
 A^{-1}(A \quad I) &= (I \quad A^{-1}), \\
 A^{-1}(A \quad B) &= (I \quad A^{-1}B).
 \end{aligned}$$

故与“用初等变换法求方阵的逆”类似, 有“用初等变换法求方阵 $A^{-1}B$ 的方法为: 对 (AB) 同时做初等行变换, 当 A 的位置变为 I 时, B 的位置就变为 $A^{-1}B$ (见 4.2 节例 2).

线性空间(2)(本章)的内容可概括为以下 4 个要点.

1 子空间的定义及子空间的运算

定理 1 设 W 是线性空间 V 的非空子集, 那么 W 是 V 的子空间的充要条件是 W 对 V 中定义加法和数乘运算封闭.

三个重要的子空间.

(1) A 的化零空间 (null space): $Ax = 0$ 的全体解构成. 这个空间是 R^n 的子空间, 它在 A 的作用下化为 0 , 记为 $N(A)$ 即

$$N(A) = \{x \in R^n \mid Ax = 0\}.$$

当 $r(A) = r$ 时, $\dim N(A) = n - r$, $Ax = 0$ 的基础解系就构成它的一组基.

(2) 向量 $\alpha_1, \dots, \alpha_m$ 的生成空间: 向量 $\alpha_1, \dots, \alpha_m$ 的所有可能的线性组合

$$W = \sum_{i=1}^m k_i \alpha_i \mid k_i \in F, i = 1, 2, \dots, m$$

是 $V_n(F)$ 的子空间, 称为 $\alpha_1, \alpha_2, \dots, \alpha_m$ 的生成空间. 记为 $L(\alpha_1, \alpha_2, \dots, \alpha_m)$ 或 $\langle \alpha_1, \alpha_2, \dots, \alpha_m \rangle$.

(3) A 的列空间 $R(A)$: 由 A 的列向量生成的空间. 设 $A = (\alpha_1, \dots, \alpha_n)$, 则

$$R(A) = L(\alpha_1, \dots, \alpha_n) = \{Ax \mid x \in R^n\}.$$

$\dim R(A) = r(A)$, A 的列向量组的极大线性无关组就是 $R(A)$ 的一组基.

定义 1 设 W_1, W_2 是 $V_n(F)$ 的两个子空间, 则集合

$$W_1 \cap W_2 = \{\alpha \mid \alpha \in W_1, \text{ 且 } \alpha \in W_2\},$$

$$W_1 + W_2 = \{\alpha = \alpha_1 + \alpha_2 \mid \alpha_1 \in W_1, \alpha_2 \in W_2\},$$

分别称为 W_1 与 W_2 的交与和.

关于怎样求子空间的交与和有如下的命题.

命题 1 设 W_1 的基为 $\alpha_1, \dots, \alpha_s$, W_2 的基为 $\beta_1, \dots, \beta_t$, 则任意 $\alpha \in W_1 \cap W_2$, 有 $\alpha \in W_1$, 且 $\alpha \in W_2$. 故可设

$$\alpha = k_1 \alpha_1 + \dots + k_s \alpha_s = \mu_1 \beta_1 + \dots + \mu_t \beta_t.$$

解齐次线性方程组

$$k_1 \alpha + \cdots + k_s \alpha = \mu \beta + \cdots + \mu \beta,$$

即

$$k_1 \alpha + \cdots + k_s \alpha - \mu \beta - \cdots - \mu \beta = 0,$$

就可以求出 $k_1, \cdots, k_s$; 或 $\mu, \cdots, \mu$ 来(事实上, 只要有一组值即可得到 α)(见 4.2 节例 4) .

命题 2 设 $W_1 = \langle \alpha, \cdots, \alpha \rangle, W_2 = \langle \beta, \cdots, \beta \rangle$, 则

$$W_1 + W_2 = L(\alpha, \cdots, \alpha, \beta, \cdots, \beta).$$

也就是说, 找出 W_1 的生成元, 再找出 W_2 的生成元, 则它们放在一块就是 $W_1 + W_2$ 的生成元, 这组向量的极大线性无关组就是 $W_1 + W_2$ 的基(见 4.2 节例 4) .

定理 2(维数公式)

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

2 子空间的直和与空间直和分解

定义 2 设 W_1, W_2 是 $V_n(F)$ 的子空间, 如果 $W_1 \cap W_2 = \{0\}$, 则称 $W_1 + W_2$ 为子空间 W_1 与 W_2 的直和, 记为 $W_1 \oplus W_2$.

定理 3 设 W_1, W_2 是 $V_n(F)$ 的子空间, $W = W_1 + W_2$. 则以下 4 个命题等价:

$$(1) W_1 \cap W_2 = \{0\};$$

$$(2) \dim(W) = \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2);$$

(3) 任意 $\alpha \in W, \alpha = \alpha + \alpha$, 其中 $\alpha \in W_1, \alpha \in W_2$ 的分解式惟一;

$$(4) 0 \text{ 表为 } W_1 \text{ 与 } W_2 \text{ 中元素和的方法惟一, 即 } 0 = 0 + 0.$$

设 V 是 n 维线性空间, W 是 V 的子空间, 求补子空间 W_3 , 使得 $V = W \oplus W_3$ 的方法由以下定理给出(即求 W_3 的维数和一组基) .

定理 4 设 V 是 n 维线性空间, W 是 V 的 m 维子空间, 则

$$\dim W_3 = n - m.$$

在 W 中任选一组基 $\alpha_1, \dots, \alpha_n$, 把它扩充为 V 的基为 $\alpha_1, \dots, \alpha_n, \beta_{n+1}, \dots, \beta_r$. 则 $\beta_{n+1}, \beta_{n+2}, \dots, \beta_r$ 是补子空间 W_3 的一组基.

3 线性空间的同构

定义 3 设 V_1 与 V_2 是域 F 上的两个线性空间, 如果存在从 V_1 到 V_2 的双射 ϕ , 满足:

- $$\begin{aligned} (1) \quad \phi(\alpha + \beta) &= \phi(\alpha) + \phi(\beta), & " \alpha, \beta \in V_1, \\ (2) \quad \phi(k\alpha) &= k\phi(\alpha), & " \alpha \in V_1, k \in F. \end{aligned}$$

则称 ϕ 是同构映射, 建立了同构映射的两个线性空间称为是同构的.

命题 3

- $$\begin{aligned} (1) \quad & \text{任意 } V_n (n \text{ 维线性空间}) \text{ 都同构于 } F^n (n \text{ 维数组空间}); \\ (2) \quad & (\text{两个有限维空间}) V_1 \text{ 与 } V_2 \text{ 同构 } \dim V_1 = \dim V_2. \end{aligned}$$

4 商空间

定义 4 设 W 是 V 的子空间, 对向量 $\alpha, \beta \in V$, 若有 $\alpha - \beta \in W$, 则称 α 与 β 模 W 同余, 记作

$$\alpha \equiv \beta \pmod{W}.$$

由于模 W 同余关系具有自反性、对称性、传递性, 所以它是 V 上的等价关系.

定义 5 设 W 是 V 的子空间, 任意 $\alpha \in V$, V 的子集合

$$\alpha + W = \{\alpha + \beta \mid \beta \in W\}$$

称为模 W 的一个同余类, 而 α 称为 (α 所在的) 这个同余类的代表元.

定义 6 线性空间 V 上按模 W 同余关系所得等价类的集合

$$\overline{V} = \{\alpha + W \mid \alpha \in V\}$$

是 V 的商集. 在商集 $\overline{V}$ 中定义加法和数乘

$$\begin{aligned} (\alpha + W) + (\beta + W) &= (\alpha + \beta) + W, \\ c(\alpha + W) &= c\alpha + W. \end{aligned}$$

则 $\overline{V}$ 对于所定义的加法和数乘构成域 F 上的线性空间, 称为 V 的

商空间,记为 V/W .

定理 5 设 V 是 n 维线性空间, W 是 V 的 m 维子空间, 则商空间的维数

$$\dim V/W = n - m.$$

设 $\alpha, \cdots, \alpha_m$ 是 W 的一组基, 扩充为 V 的基为 $\alpha, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_n$. 则 $\beta_{m+1} + W, \beta_{m+2} + W, \cdots, \beta_n + W$ 是商空间 V/W 的一组基.

4.2 典型题讲评

例 1 证明实数域上所有形如 $f(A)$ 的矩阵组成的集合, 对于矩阵加法和数乘运算构成 $M_3(\mathbb{R})$ 的子空间. 其中 $f(x) \in \mathbb{R}[x]$, 而

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad \omega = \frac{-1 + \sqrt{3}i}{2}.$$

试求其维数及一组基.

证 记

$$W = \{ f(A) \mid f(x) \in \mathbb{R}[x], A \text{ 如题设} \},$$

则显然 W 是 $M_3(\mathbb{R})$ 的子集. 由于 $A \in W$, 故 W 非空. 又任意 $A_1, A_2 \in W$, 必有 $f_1(x), f_2(x) \in \mathbb{R}[x]$, 使

$$A_1 = f_1(A), \quad A_2 = f_2(A).$$

于是

$$A_1 + A_2 = f_1(A) + f_2(A) = g(A) \in W,$$

其中 $g(x) = f_1(x) + f_2(x)$. 又显然 $\lambda A_1 \in W$, 故 W 对加法, 数乘运算封闭. 所以是 $M_3(\mathbb{R})$ 的子空间.

又由于所有形如 $f(A)$ 的矩阵都是

$$I, A, A^2, A^3, \cdots, A^n, \cdots$$

的线性组合, 故 W 的维数和基由这组‘向量’决定. 首先, A 与 I 线性无关 (因 $A \neq \lambda I$), 又

$$A^2 = \frac{1}{\omega^2} \neq k_1 I + k_2 A, \quad k_1, k_2 \in \mathbb{R},$$

$$\omega$$

而

$$A^3 = I = 1 \cdot I + 0 \cdot A + 0 \cdot A^2,$$

即知 I, A, A^2 线性无关, A^3 (从而 $A^k, k \in \mathbb{N}$) 可由 I, A, A^2 线性表出. 故

$\dim W = 3$, I, A, A^2 是 W 的基.

讲评 (1) 若单独判断 W 是线性空间, 要用线性空间的定义. 即要验证加法满足 4 条性质, 数乘满足 4 条性质. 比较麻烦. 这里是用证明 W 是 $M_3(\mathbb{R})$ 的子空间的方法证它是线性空间, 于是只要证明 3 条: 非空子集、加法封闭、数乘封闭.

(2) 要找 W 的基和算出它的维数关键的一步是任意 $B \in W$ 都有 $f(x)$, 使

$$B = f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I,$$

故 B 是 $I, A, A^2, \cdots, A^n$ 的线性组合. 这样就可从这组“向量”中找“极大线性无关组”, 从而确定基和维数.

例 2 设

$$\begin{aligned} F_{11} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, F_{12} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, F_{13} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, F_{14} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}; \\ G_1 &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, G_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, G_3 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, G_4 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

(1) 证明 $\{F_{1i}\}, \{G_j\}$ 都是 $M_2(\mathbb{R})$ 的基;

(2) 求由基 $\{F_{1i}\}$ 到基 $\{G_j\}$ 的过渡矩阵;

(3) 求矩阵 $A = \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix}$ 在两组基下的坐标.

解 (1) 记 $(F_{11}, F_{12}, F_{13}, F_{14}) = (E_{11}, E_{12}, E_{21}, E_{22})B, \quad (*_1)$

$(G_1, G_2, G_3, G_4) = (E_{11}, E_{12}, E_{21}, E_{22})C, \quad (*_2)$

则因为

$$\det \mathbf{B} = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} \neq 0, \quad \det \mathbf{C} = \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix} \neq 0,$$

知 $\{\mathbf{F}_{1i}\}, \{\mathbf{G}_j\}$ 是两组线性无关的向量,所以均是 $M_2(\mathbb{R})$ 的基.

(2) 设由基 $\{\mathbf{F}_{1i}\}$ 到基 $\{\mathbf{G}_j\}$ 的过渡矩阵为 $\mathbf{D}$,则有形式写法

$$(\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{G}_4) = (\mathbf{F}_{11}, \mathbf{F}_{12}, \mathbf{F}_{13}, \mathbf{F}_{14}) \mathbf{D}, \quad (*_3)$$

把两组基在自然基 $\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}$ 下的坐标列代入上式有

$$\mathbf{C} = \mathbf{B}\mathbf{D}, \quad \text{所以 } \mathbf{D} = \mathbf{B}^{-1}\mathbf{C}.$$

以下用初等变换法求 $\mathbf{B}^{-1}\mathbf{C}$:

$$(\mathbf{B}, \mathbf{C}) = \begin{array}{cccccccc} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ & & & 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ -r_2 + r_1 & & 0 & 1 & 0 & 0 & 0 & 1 & & 1 & 1 \\ -r_4 + r_3 & & 0 & 0 & 1 & 0 & 1 & 1 & & -1 & -1 \\ & & & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ & & & & 1 & 0 & & -1 & & 0 \end{array},$$

所以

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & & 1 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

$$\begin{aligned} (3) \quad \mathbf{A} &= \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix} = \mathbf{E}_{11} - 2\mathbf{E}_{12} + 3\mathbf{E}_{21} \\ &= \mathbf{F}_{11} - 2(\mathbf{F}_{12} - \mathbf{F}_{11}) + 3\mathbf{F}_{13} = 3\mathbf{F}_{11} - 2\mathbf{F}_{12} + 3\mathbf{F}_{13}. \end{aligned}$$

又

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix} = \mathbf{E}_{11} - 2\mathbf{E}_{12} + 3\mathbf{E}_{21}$$

$$\begin{aligned}
 &= (\mathbf{G}_4 - \mathbf{G}_3) - 2(\mathbf{G}_2 - \mathbf{G}_1) + 3(\mathbf{G}_1 + \mathbf{G}_3 - \mathbf{G}_4) \\
 &= 5\mathbf{G}_1 - 2\mathbf{G}_2 + 2\mathbf{G}_3 - 2\mathbf{G}_4,
 \end{aligned}$$

所以 A 在两组基下的坐标分别为

$$\begin{aligned}
 \mathbf{x} &= (3, -2, 3, 0)^T, \\
 \mathbf{y} &= (5, -2, 2, -2)^T.
 \end{aligned}$$

讲评 (1) $\mathbf{E}_{ij} (1 \leq i, j \leq 2)$ 是 (i, j) 位元素为 1, 其余元素为零的

二阶矩阵, 它是 $M_2(\mathbb{R})$ 的自然基, 因为任意 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ 有

$$M = a\mathbf{E}_{11} + b\mathbf{E}_{12} + c\mathbf{E}_{21} + d\mathbf{E}_{22} = (\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}) \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix},$$

即 M 在此自然基上的坐标, 就是把 M 的元素按行排起来.

(2) $D = B^{-1}C$ 也可以这样得到, 把 $(*)_1$ 式, $(*)_2$ 式代入 $(*)_3$ 式中有

$$\begin{aligned}
 (\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22})C &= [(\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22})B]D \\
 &= (\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22})(BD),
 \end{aligned}$$

故得 $C = BD$.

(3) 求 $D = B^{-1}C$ 时, 用初等变换法直接得出解答, 比起先求 B^{-1} , 再做矩阵乘法要简单许多.

$$\text{例 3 设 } A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & w & \cdots \\ \cdots & \cdots & w & 0 \\ 1 & \cdots & \cdots & 1 \end{pmatrix} \text{ 是 } n \text{ 阶矩阵.}$$

(1) 证明所有与 A 可交换的矩阵 (记为 $T(A)$) 的全体是 $M_n(\mathbb{R})$ 的子空间;

(2) 求 $T(A)$ 的维数和一组基.

解 (1) 因为 $AI = IA$, 所以 $I \in T(A)$, 故 $T(A)$ 非空.

" $B, C \in T(A)$, 有 $AB = BA, AC = CA$, 于是, $A(B + C) = (B + C)A$, 所以 $B + C \in T(A)$, 又对任 $k \in F$, 均有 $(kB)A = A(kB)$, 即 $kB \in T(A)$, 所以 $T(A)$ 对加法和数乘封闭. 由 4.1 节定理 1, $T(A)$ 是 $M_n(R)$ 的子空间.

(2) 因为 $A = I + J + J^2 + \cdots + J^{n-1}$, 所以 $A^{-1} = I - J$, 其中

$$J = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & W & W & W & \cdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

又因为 $AB = BA$ $BA^{-1} = A^{-1}B$ $B(I - J) = (I - J)B$ $BJ = JB$. 于是知, 与 A 乘法可变换的全体 $B = (b_{ij})$ 须满足

$$\begin{pmatrix} b_{12} & b_{13} & \cdots & b_{1n} & 0 & 0 & 0 & \cdots & 0 \\ b_{22} & b_{23} & \cdots & b_{2n} & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ b_{32} & b_{33} & \cdots & b_{3n} & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n2} & b_{n3} & \cdots & b_{nn} & 0 & b_{n-11} & b_{n-12} & \cdots & b_{n-1n} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

比较两边元素知有 $b_{12} = \cdots = b_{1n} = 0 \rightarrow b_{23} = \cdots = b_{2n} = 0 \rightarrow b_{34} = \cdots = b_{3n} = 0 \rightarrow \cdots \rightarrow b_{n-1n} = 0$ (即 $b_{ij} = 0$ ($j > i$ 时)); 且 $b_{11} = b_{22} = \cdots = b_{nn}$, $b_{11} = b_{22} = \cdots = b_{n-1n}$, $b_{11} = b_{22} = \cdots = b_{n-2n}$, $\cdots$, $b_{n-1n} = b_{n2}$ (因为比较两边 (i, j) 位元素有 $b_{ij+1} = b_{i-1j}$) . 即

$$B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{11} & b_{11} & W & \cdots \\ \cdots & W & W & 0 \\ b_{11} & \cdots & b_{11} & b_{11} \end{pmatrix} = b_{11} I + b_{11} J + \cdots + b_{11} J^{n-1}.$$

所以 $T(A)$ 是 n 维的, $I, J, J^2, \cdots, J^{n-1}$ 为其中一组基.

讲评 本解法巧妙地使用了以下知识:

$$(1) AB = BA \quad A^{-1}B = BA^{-1};$$

(2) 若 $J^k = 0$, 则 $(I - J)^{-1} = I + J + \cdots + J^{k-1}$. 这是许多代数书上都有的习题 (原题用 A 记 J). 这里用的是 $(I + J + \cdots + J^{k-1})^{-1} = I - J$. 从而把求满足 $AB = BA$ 的 B 转化成求满足

$$JB = BJ$$

的矩阵 B 的问题, 大大简化了计算.

例 4 设 $W_1 = L(\alpha, \alpha, \alpha)$, $W_2 = L(\beta, \beta, \beta)$, 其中 $\alpha = (1, 1, 1, 0)^T$, $\alpha = (1, 2, 3, 4)^T$, $\alpha = (2, 1, 3, 4)^T$, $\beta = (5, 6, 7, 8)^T$, $\beta = (3, 4, 5, 6)^T$, $\beta = (2, 3, 4, 5)^T$; 试求 $W_1 \cap W_2$, $W_1 + W_2$ 的基与维数.

解 由命题 2 知

$$W_1 + W_2 = L(\alpha, \alpha, \alpha, \beta, \beta, \beta).$$

为求 $W_1 + W_2$ 的基和维数, 对 $\alpha, \alpha, \alpha, \beta, \beta, \beta$ 排成的矩阵做初等行变换

$\alpha, \alpha, \alpha, \beta, \beta, \beta$

$$\begin{array}{cccccc|cccccc} 1 & 1 & 2 & 5 & 3 & 2 & & 1 & 1 & 2 & 5 & 3 & 2 \\ 1 & 2 & 1 & 6 & 4 & 3 & -r_1 + r_2 & 0 & 1 & -1 & 1 & 1 & 1 \\ 1 & 3 & 3 & 7 & 5 & 4 & -r_1 + r_3 & 0 & 2 & 1 & 2 & 2 & 2 \\ 0 & 4 & 1 & 8 & 6 & 5 & & 0 & 4 & 1 & 8 & 6 & 5 \\ & & & & & & & & & & & & \\ & & & 1 & 1 & 2 & 5 & 3 & 2 & & 1 & 1 & 2 & 5 & 3 & 2 \\ -2r_2 + r_3 & & & 0 & 1 & -1 & 1 & 1 & 1 & & 0 & 1 & -1 & 1 & 1 & 1 \\ -4r_2 + r_4 & & & 0 & 0 & 3 & 0 & 0 & 0 & & 0 & 0 & 3 & 0 & 0 & 0 \\ & & & 0 & 0 & 5 & 4 & 2 & 1 & & 0 & 0 & 0 & 4 & 2 & 1 \end{array} \quad (*)$$

由于最后的阶梯型矩阵的秩为 4, 且前 4 列线性无关, 故 $\alpha, \alpha, \alpha, \beta$ 是 $W_1 + W_2$ 的基, $\dim(W_1 + W_2) = 4$, 而且由 (*) 式还可以看出 $\dim W_1 = 3$, $\dim W_2 = 2$. 所以

$$\dim W_1 \cap W_2 = \dim W_1 + \dim W_2 - \dim(W_1 + W_2) = 1.$$

下面求 $W_1 \cap W_2$ 的基. 因为任意 $\alpha \in W_1 \cap W_2$, 有 $\alpha \in W_1$ 且 $\alpha \in$

W_2 , 故 α 既可由 $\alpha_1, \alpha_2, \alpha_3$ 线性表出, 又可由 β_1, β_2 线性表出.

令 $\alpha = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \lambda_3 \alpha_3 = -\mu_1 \beta_1 - \mu_2 \beta_2$,
也即是

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \lambda_3 \alpha_3 + \mu_1 \beta_1 + \mu_2 \beta_2 = 0.$$

把 $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$ 代入上式得 5 个变量 ($\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2$ 为变量) 4 个方程的齐次线性方程组. 利用 (*) 式的最后一个矩阵的前 5 列可解出 $\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2$, 事实上只要解出他们中的一组 ($\lambda_1, \lambda_2, \lambda_3$ 或 μ_1, μ_2) 即可.

取 $\mu_1 = -2$, 得 $\mu_2 = 1$. 从而得 $\alpha = -\beta_1 + 2\beta_2 = (1, 2, 3, 4) = \alpha_4$.

讲评 解题过程采用先求 $W_1 + W_2$ 的思路, 计算量较少. (若先求 W_1 和 W_2 的基和维数, 则多做不少工作). 记

$$B = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2),$$

则对 B 做初等行变换化为阶梯型矩阵后, 不仅得出 $W_1 + W_2$ 的基和维数. 而且, 由 (*) 式的前 3 列知 $\alpha_1, \alpha_2, \alpha_3$ 线性无关, 故 $\dim W_1 = 3$, 看后 3 列: 显然任意两列不成比例, 故秩至少为 2. 又这后 3 列的第 1 行 = 第 2 行 + 第 4 行, 所以 3 阶子式为 0 (其余的 3 阶子式均含一行 0, 值自然为 0), 故秩为 2. 于是知 $\dim W_2 = 2$. 另外, 因 B 的前 5 列恰为求 $W_1 \cap W_2$ 中元素 α 应满足的齐次线性方程组的系数, 故利用已化好的阶梯型矩阵, 顺手得来结果, 几乎没什么计算量 (因由 (*) 式的最后一行有 $4\mu_1 + 2\mu_2 = 0$, 故取 $\mu_1 = -2$, 则 $\mu_2 = 1$).

例 5 已知

$$\begin{aligned} W_1 = & \begin{pmatrix} a & 0 & b \\ c & 0 & b \\ 0 & c & d \end{pmatrix} \left| \begin{array}{l} a, b, c, d \in \mathbb{R} \end{array} \right., \\ W_2 = & \begin{pmatrix} x & 0 & y \\ 0 & z & 0 \\ 0 & 0 & 0 \end{pmatrix} \left| \begin{array}{l} x, y, z \in \mathbb{R} \end{array} \right., \end{aligned}$$

求 $W_1 + W_2$ 和 $W_1 \cap W_2$ 的维数和基.

解 记 E_{ij} 为 (i, j) 位元素为 1, 而其余元素均为 0 的三阶矩阵, 则

$$\begin{aligned} W_1 &= a \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= a E_{11} + b(E_{13} + E_{23}) + c(E_{21} + E_{32}) + d E_{33}. \end{aligned}$$

因 $E_{11}, E_{13} + E_{23}, E_{21} + E_{32}, E_{33}$ 线性无关, 所以 $\dim W_1 = 4$.

而 $W_2 = x E_{11} + y E_{13} + z E_{22}$,
所以 $\dim W_2 = 3$.

又

$$\begin{aligned} W_1 + W_2 &= L(E_{11}, E_{13} + E_{23}, E_{21} + E_{32}, E_{33}, E_{11}, E_{13}, E_{22}) \\ &= L(E_{11}, E_{13}, E_{21} + E_{32}, E_{22}, E_{23}, E_{33}), \end{aligned}$$

所以 $\dim(W_1 + W_2) = 6$, 上式中 6 个生成元就是它的一组基, 于是 $\dim W_1 \cap W_2 = 4 + 3 - 6 = 1$, 易知 E_{11} 为 $W_1 \cap W_2$ 的基.

讲评 (1) 对于 $M_3(\mathbb{R})$ 中的矩阵, 当然也可以考虑他们在自然基 $\{E_{ij}\} (1 \leq i, j \leq 3)$ 上的坐标列, 类似于例 4 那样做. 但 9 个分量的列向量写起来太长、太啰嗦, 故我们不用“初等变换法”去化阶梯型矩阵, 而是用矩阵的几何直观“占位子”的方法找生成元, 看生成元的极大线性无关组: 如, 设有线性组合

$$k_1 E_{11} + k_2 (E_{13} + E_{23}) + k_3 (E_{21} + E_{32}) + k_4 E_{33} = \mathbf{0},$$

则得

$$\begin{pmatrix} k_1 & 0 & k_2 \\ k_3 & 0 & k_2 \\ 0 & k_3 & k_4 \end{pmatrix} = \mathbf{0},$$

故得 $k_1 = 0, k_2 = 0, k_3 = 0, k_4 = 0$ (看 9 个位子, 事实上就是 9 个坐标分量).

(2) 在求 $W_1 \cap W_2$ 时, 因先求出 $\dim W_1 \cap W_2 = 1$, 又因 E_{11} 既

是 W_1 也是 W_2 的生成元,故 $W_1 \cap W_2 = L(E_{11})$. 若不是先知道交的维数是 1, 只看出 E_{11} 是 $W_1 \cap W_2$ 中的一个向量, 不能下断言基中只此一个向量 E_{11} , 还要按定义去找, 看还有没有其他的生成元.

例 6 设 W_1, W_2 是线性空间 V 的两个非平凡子空间, 证明: 存在 $\alpha \in V$ 使 $\alpha \notin W_1$ 且 $\alpha \notin W_2$.

证 因为 W_1, W_2 是 V 的非平凡子空间, 所以存在 $\alpha \notin W_1$, $\beta \notin W_2$. 若 $\alpha \notin W_2$, 则 α 为所求; 若 $\beta \notin W_1$, 则 β 为所求. 否则便有

$$\alpha \notin W_1, \quad \beta \in W_1; \quad \beta \notin W_2, \quad \alpha \in W_2,$$

于是令 $\gamma = \alpha + \beta$, 则 $\gamma \notin W_1$, 且 $\gamma \notin W_2$.

讲评 (1) 此题无论用怎样的语言(或思路)最后都得达到这一步, 即有向量 α, β 满足: $\alpha \notin W_1$, 但 $\alpha \in W_2$; $\beta \notin W_2$, 但 $\beta \in W_1$. 于是 $\gamma = \alpha + \beta$ 为所求.

(2) $\gamma \notin W_1$, 且 $\gamma \notin W_2$ 是显然的. 否则若 $\gamma \in W_1$, 则由 $\beta \in W_1$ 可得 $\alpha = \gamma - \beta \in W_1$ 与 $\alpha \notin W_1$ 矛盾; 若 $\gamma \in W_2$, 则由 $\alpha \in W_2$, 可得 $\beta = \gamma - \alpha \in W_2$, 与 $\beta \notin W_2$ 矛盾.

(3) 下面的证法, 错在哪里?

设 $\dim V = n$, 因为 W_1, W_2 是 V 的非平凡子空间, 所以

$$0 < \dim W_1, \quad \dim W_2 < n.$$

设 V 的一组基为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 则取 $\alpha = k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n$, 其中 $k_1, k_2, \dots, k_n \neq 0$, 有 $\alpha \in V$, 且 $\alpha \notin W_1, \alpha \notin W_2$.

错在: 他以为子空间总是由 V 中基的一部分向量生成的. 事实上, 若先在 V 中选基; 子空间 W 的基不一定是上述基中的向量. 请看下面的例子.

例: $V_3(\mathbb{R})$ 中, 取 i, j, k 为 $V_3(\mathbb{R})$ 的一组基, W_1 为过 $(1, 1, 1)$ 点的直线, $W_2 = z$ 轴, 则 $\alpha = i + j + k \in W_1, \alpha \notin W_2$.

要想让子空间 W 的基是空间 V 中基的一部分. 只能从子空间 W 中取基并扩充为 V 的基.

(4) 此题可推广到有限个子空间的情况, 见 4.3 节例 7.

例 7 证明: 所有 n 阶矩阵空间是线性子空间 L_1 和 L_2 的直和, 其中 L_1 是对称矩阵子空间, L_2 是斜对称矩阵子空间.

证 方法 1 记 $M_n(F)$ 是 n 阶矩阵空间, 任意 $A \in M_n(F)$, 有

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) = B + C,$$

其中 $B \in L_1$, $C \in L_2$. 又显然 $L_1 + L_2 = M_n(F)$, 所以

$$M_n(F) = L_1 + L_2, \text{ 且 } L_1 \cap L_2 = \{0\}$$

(若 $M \in L_1 \cap L_2$, 有 $M^T = M$, $M^T = -M$, 故 $M = -M$, 所以 $M = 0$).

故得

$$M_n(F) = L_1 \oplus L_2.$$

方法 2 因为 $L_1 = M_n(F)$, $L_2 = M_n(F)$, 所以 $L_1 + L_2 = M_n(F)$. 又因为 $L_1 \cap L_2 = \{0\}$, 所以 $L_1 + L_2$ 是直和.

用 E_{ij} 记 (i, j) 位元素为 1, 其余元素为 0 的 n 阶矩阵, 则由于

$$\{(E_{ij} + E_{ji}), E_{ii}\}, 1 \leq i < j \leq n \text{ 是 } L_1 \text{ 的基,}$$

$$\{(E_{ij} - E_{ji})\}, 1 \leq i < j \leq n \text{ 是 } L_2 \text{ 的基.}$$

$$\text{从而知 } \dim L_1 = \frac{1}{2}n(n+1), \quad \dim L_2 = \frac{1}{2}n(n-1),$$

$$\text{所以 } \dim(L_1 + L_2) = \dim L_1 + \dim L_2 = n^2 = \dim M_n(F),$$

故

$$M_n(F) = L_1 \oplus L_2.$$

讲评 证空间分解式 $W = W_1 \oplus W_2$, 需要证明两点:

(1) 证等号成立; (2) 证“ $\oplus$ ”, 即和是直和. 要证等号成立就要

① 证明任意 $\alpha \in W$ 有 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in W_1$, $\alpha_2 \in W_2$, 故 $\alpha \in W_1 + W_2$, 从而有 $W = W_1 + W_2$ (因 W_1, W_2 都是 W 的子空间, 故 $W_1 + W_2$ 仍是 W 的子空间, 故 $W_1 + W_2 = W$ 显然); 或者 ② 证明

$$\dim W = \dim W_1 + \dim W_2.$$

要证“ $\oplus$ ”由定理 3 有 4 个等价命题: 故可证 ① $W_1 \cap W_2 = \{0\}$;

② $\dim W_1 + \dim W_2 = \dim W$; ③ 向量分解式惟一.

这里方法1用的是①和①';方法2用的是②和①'.

例8 设 W 是二阶上三角矩阵所构成的线性空间, 证明 W 与 $SM_2(F)$ (二阶实对称矩阵) 同构, 并给出一个同构映射.

解 由于

$$\begin{aligned} W &= \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a, b, c \in F \right\} \\ &= a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

故 $\dim W = 3$; 又

$$\begin{aligned} SM_2(F) &= \left\{ \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mid a, b, c \in F \right\} \\ &= a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

所以 $\dim SM_2(F) = 3 = \dim W$. 由命题3知, W 与 $SM_2(F)$ 同构. 定义

$$\begin{aligned} \phi: W &\rightarrow SM_2(F) \\ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} &\mapsto \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \end{aligned}$$

则 ϕ 是一个双射: 因为对任意

$$\mathbf{A} = \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix} \neq \mathbf{A}' = \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix},$$

有

$$\phi(\mathbf{A}) = \begin{pmatrix} a_1 & b_1 \\ b_1 & c_1 \end{pmatrix} \neq \begin{pmatrix} a_2 & b_2 \\ b_2 & c_2 \end{pmatrix} = \phi(\mathbf{A}'),$$

故 ϕ 是单射. 又对任意 $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in SM_2(F)$, 有

$$\mathbf{A} = \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{pmatrix} \in W, \text{ 使 } \phi(\mathbf{A}) = \mathbf{B},$$

故 ϕ 是满射. 综上得 ϕ 是双射, 且对任意 $A_1, A_2 \in W, \lambda \in F$ 有

$$\begin{aligned}\phi(A_1 + A_2) &= \phi \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} + \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \\ &= \phi \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ 0 & a_1 + a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ b_1 + b_2 & a_1 + a_2 \end{pmatrix} \\ &= \phi(A_1) + \phi(A_2), \\ \phi(\lambda A_1) &= \phi \begin{pmatrix} \lambda a_1 & \lambda b_1 \\ 0 & \lambda a_1 \end{pmatrix} = \begin{pmatrix} \lambda a_1 & \lambda b_1 \\ \lambda b_1 & \lambda a_1 \end{pmatrix} = \lambda \phi(A_1),\end{aligned}$$

即 ϕ 保加法, 保数乘. 由定义 3, 知 ϕ 为同构映射.

讲评 证明题的思路从哪里来? 我们认为首先要看清“已知”, “求证”; 其次“头脑中要有知识”. 比如这道题, 要证两个线性空间同构 (这是看清了题目), 就要到大脑中“搜索”与同构相关的知识; 于是找到了命题 3 中 (2) 两个线性空间同构 维数相同. 到这一步思路有了, 问题也变得非常简单, 因为 W 与 $SM_2(F)$ 的维数非常好算. 其次, 要建立同构映射, 又在头脑中找到同构映射的定义, 于是按定义 3, 证 ϕ 是双射, 保线性运算等, 思路非常清晰 (理科学生比较善于记一些“知识”, 而工科学生往往以将来可以“查到知识用”为借口, 不愿记“知识”, 事实上, 工科学生若能在学数学时, 有意加强数学思维的训练, 将来用数学时会展现其优势). 希望同学们在学习“线代”时, 头脑中有一棵“知识的树”, 主干、支干一目了然, 这样做起题目就快 (将来也会发挥更大的作用).

例 9 设 σ 是线性空间 V_1 到 V_2 的同构映射, W 是 V_1 的子空间, 证明 $\sigma(W)$ 是 V_2 的子空间.

解 因为 $0 \in W, \sigma(0) = 0 \in \sigma W$, 故 σW 非空.

又 " $\alpha, \beta \in \sigma W, \forall \xi, \eta \in W$ 使 $\alpha = \sigma\xi, \beta = \sigma\eta$, 且由于 W 是 V_1 的子空间, 故 $\xi + \eta \in W, k\xi \in W$.

于是有

$$\alpha + \beta = \sigma\xi + \sigma\eta = \sigma(\xi + \eta) \in \sigma W,$$

$$k\alpha = k\sigma\xi = \sigma(k\xi) \in \sigma W.$$

故得 σW 为 V_2 的子空间.

讲评 要证明某一个子集是子空间, 由定理 1 知, 需证 3 条: (1) 非空; (2) 对加法封闭; (3) 对数乘封闭. 现在要证的是 $\sigma(W)$, 即 W 中元素象的全体是 V_2 的子空间, 故一切要对 $\sigma(W)$ 说事. 即证 $\sigma(W)$ 满足上述 3 条. 证加法封闭, 是说 $\sigma(W)$ 对加法封闭. 故要说: 任取 $\alpha, \beta \in \sigma W, \dots\dots$ 而不能说任取 $\xi, \eta \in W, \sigma(\eta + \eta) = \sigma\xi + \sigma\eta = \sigma\eta \in \sigma W$ (那样写思路不清楚).

例 10 设 W_1 是齐次线性方程组

$$x_1 + x_2 + x_3 = 0,$$

$$x_2 - x_3 = 0$$

的解空间, 求商空间 R^3/W_1 的维数和基.

解 求解已知的齐次线性方程组, 得

$$W_1 = \{ k(-2, 1, 1)^T \}, \quad k \in R,$$

于是知 $(-2, 1, 1)^T$ 是 W_1 的基, 把它扩充为 R^3 的基, 得

$$(-2, 1, 1)^T, \quad (0, 1, 0)^T, \quad (0, 0, 1)^T.$$

由定理 5 知 R^3/W_1 的基为

$$(0, 1, 0)^T + W_1, \quad (0, 0, 1)^T + W_2,$$

且

$$\dim R^3/W_1 = 2.$$

讲评 商空间是“线性代数”中比较抽象的一个内容, 对这样的东西要弄清楚它的定义, 然后是理解它, 再掌握怎样求一个空间 V 的商空间 V/W 的方法, 这些概括为定义 4~定义 6 及定理 5 (给出求商空间的方法). 死记硬背是困难的, 深入理解它们, 则不

难掌握.

例 11 设

$$\begin{aligned} \mathbf{E}_{11} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{E}_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \mathbf{E}_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \mathbf{E}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}; \\ \mathbf{G}_{11} &= \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \mathbf{G}_{12} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \mathbf{G}_{21} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \mathbf{G}_{22} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

证明 $\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_1, \mathbf{G}_2$ 是 $M_2(\mathbb{R})$ 的一组基, 并求它到 $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_{21}, \mathbf{E}_{22}$ 的过渡矩阵. 再求 $A = \begin{pmatrix} 0 & 1 \\ 2 & -3 \end{pmatrix}$ 在这两组基下的坐标.

解 (1) 设有线性组合

$$k_1 \mathbf{G}_{11} + k_2 \mathbf{G}_{12} + k_3 \mathbf{G}_{21} + k_4 \mathbf{G}_{22} = \mathbf{0},$$

$$k_2 + k_3 + k_4 = 0,$$

$$k_1 + k_3 + k_4 = 0,$$

$$k_1 + k_2 + k_4 = 0,$$

$$k_1 + k_2 + k_3 = 0.$$

即得

因其系数行列式不等于 0, 故上述齐次线性方程只有零解, $k_1 = k_2 = k_3 = k_4 = 0$, 故 $\mathbf{G}_{11}, \mathbf{G}_{12}, \mathbf{G}_{21}, \mathbf{G}_{22}$ 线性无关.

(2) 记 $\{\mathbf{G}_{ij}\}$ 在自然基下的坐标列排成的矩阵为 P , 则有

$$(\mathbf{G}_{11}, \mathbf{G}_{12}, \mathbf{G}_{21}, \mathbf{G}_{22}) = (\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}) P.$$

上式两边同乘 P^{-1} , 得

$$(\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}) = (\mathbf{G}_{11}, \mathbf{G}_{12}, \mathbf{G}_{21}, \mathbf{G}_{22}) P^{-1},$$

故 P^{-1} 为所要求的从基 $\mathbf{G}_{11}, \mathbf{G}_{12}, \mathbf{G}_{21}, \mathbf{G}_{22}$ 到基 $\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}$ 的过渡矩阵. 可用初等变换法求得

$$P^{-1} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \xrightarrow{-1} \begin{pmatrix} -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \\ 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \\ 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 \end{pmatrix}.$$

(3) 因为 $\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22}$ 是自然基, 故 A 在其上的坐标就是

A 的元素按行排成一行, 即为 $x = (0, 1, 2, -3)^T$. 而 A 在 $G_{11}, G_{12}, G_{21}, G_{22}$ 下的坐标为

$$y = P^{-1}x = \frac{1}{3} \begin{pmatrix} -2 & 1 & 1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 1 & 1 & -1 \\ 1 & 1 & -2 & 1 & 2 & -2 \\ 1 & 1 & 1 & -2 & -3 & 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

讲评 (1) 首先, 判断 $G_{11}, G_{12}, G_{21}, G_{22}$ 线性无关等价于判断它们在自然基下的坐标列线性无关, 即判断其 4 个坐标列构成的方阵的行列式

$$\Delta = \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix}$$

不等于 0. 故也可用直接计算上述行列式不等于 0 来证 $\{G_{ij}\}$ 为基.

(2) Δ 可用下法计算, 记 $\alpha = (1, 1, 1, 1)^T$, 则

$$\Delta = |\alpha\alpha^T - I_4| = |I_4 - \alpha\alpha^T| = |1 - \alpha^T\alpha| = -3.$$

(3) 介绍一个计算 P^{-1} 的方法

$$\begin{aligned} P^{-1} &= (\alpha\alpha^T - I_4)^{-1} = - (I_4 - \alpha\alpha^T)^{-1} \stackrel{\text{def}}{=} - \frac{1}{(I_4 - \alpha\alpha^T)} \\ &= - (I + \alpha\alpha^T + (\alpha\alpha^T)^2 + (\alpha\alpha^T)^3 + \cdots + (\alpha\alpha^T)^n + \cdots) \\ &= - [I + \alpha\alpha^T (1 + 4 + 4^2 + \cdots + 4^{n-1} + \cdots)] \\ &= - I + \frac{\alpha\alpha^T}{1 - 4} \end{aligned}$$

$$= \frac{1}{3} \alpha\alpha^T - 3I = \frac{1}{3} \begin{pmatrix} -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \\ 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 \end{pmatrix}.$$

其中第 3 个等号是把它(前式)看成 $\frac{1}{1-x}$, 第 4 个等号是做“形式

展开”,第5个等号是由于 $(\alpha\alpha^T)^2 = \alpha(\alpha^T\alpha)\alpha^T = 4\alpha\alpha^T$,同理 $(\alpha\alpha^T)^n = 4^{n-1}\alpha\alpha^T$.第6个等号是“形式求和”.这里“形式展开”与“形式求和”是用到“形式幂级数的理论”,不考虑收敛性.这个方法适用于一类特殊方阵的求逆.整个过程只用到一个公式,就是微积分中级数展开式 $\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$.若不用形式幂级数的理论,请读者用分块矩阵的初等变换和求逆,证明:对任意 $\alpha, \beta \in F^n$,有

$$\begin{aligned} |I_n - \alpha\beta^T| &= |1 - \beta^T\alpha|, \\ (I_n - \alpha\beta^T)^{-1} &= I_n + (1 - \beta^T\alpha)^{-1}\alpha\beta^T. \end{aligned}$$

前几年,我们曾把它留做思考题(开卷考试题)给学生.可喜的是,许多学生做出了结论.

例 12 设 W 是二阶对角矩阵所构成的线性空间,它是 $SM_2(\mathbb{R})$ 的子空间,求商空间 $SM_2(\mathbb{R})/W$ 的维数及一组基.

解 由题设得

$$W = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, a, d \in \mathbb{R},$$

$$SM_2(\mathbb{R}) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, a, b, c \in \mathbb{R},$$

故

$$\begin{aligned} W &= a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ SM_2(\mathbb{R}) &= a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

由于 W 是 $SM_2(\mathbb{R})$ 的非空子集,且对加法和数乘封闭,故是

$SM_2(\mathbb{R})$ 的子空间.又由上面知 E_{11}, E_{22} 是 W 的基,扩充 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} =$

$E_{12} + E_{21}$ 后得到 $SM_2(\mathbb{R})$ 的基, $E_{11}, E_{22}, E_{12} + E_{21}$.故由定理 5 知

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + W$$

是商空间 $SM_2(\mathbb{R})/W$ 的基, 这是一个由 $SM_2(\mathbb{R})$ 中向量构成的集合. 由于商空间只有一个基向量, 故 $\dim SM_2(\mathbb{R})/W = 1$.

讲评 商空间是集合的集合, 即商空间的元素是按照“某一分类法则”把 V 中元素分类得到的集合, 而不再是某个具体向量. 这就好比某个系的一个年级中有 100 个同学, 把他们分成 3 个班, 那么年级主管再说事时, 只对某某班说, 而不对 100 个同学个人. 商集对分班来说, 它中间只有 3 个元素: 一班、二班、三班. 故上述商空间中的基

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + W$$

是一个集合, 绝不能说成商空间的元素是形如 $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ 的矩阵.

例 13 证明: 线性子空间 L_1 和 L_2 的和 S 是直和当且仅当至少有一个向量 $x \in S$ 惟一地表示为形式 $x = x_1 + x_2$, 其中 $x_1 \in L_1, x_2 \in L_2$.

证 (必要性) 由于 $S = L_1 + L_2$ 是直和, 所以 " $x \in S$ 关于 L_1, L_2 的分解式是惟一的, 故至少有一个 x 表示式惟一.

(充分性) 若存在一个 x , 表示式 $x = x_1 + x_2, x_1 \in L_1, x_2 \in L_2$ 惟一, 要证 $S = L_1 + L_2$ 是直和. 用反证法, 若不是直和, 由定理 3 知, 必有 $L_1 \cap L_2 \neq \{0\}$, 故有非零向量 $\alpha \in L_1 \cap L_2$, 又因为 $L_1 \cap L_2$ 为子空间, 所以 $-\alpha \in L_1 \cap L_2$, 于是有

$$x = x_1 + x_2 = (x_1 + \alpha) + (x_2 - \alpha), \quad x_1 + \alpha \in L_1, \quad x_2 - \alpha \in L_2.$$

这与已知 x 的分解式惟一矛盾, 所以必 $L_1 \cap L_2 = \{0\}$, 故必为直和.

讲评 本题可以帮助读者深入理解定理 3, 其中的第 3 个等价命题可以改为只对一个向量的分解式是惟一的. 从上述证明过程中可以看到: 分解式是否惟一 $L_1 \cap L_2$ 是否为 0. 即有一个向量的分解式不惟一, 则必有 $L_1 \cap L_2 \neq \{0\}$. 若 $L_1 \cap L_2 \neq \{0\}$ (只需

其中有一个非零向量), 则一切向量的分解式都不惟一.

例 14 求由向量组 α, α, α 和 β, β, β 所张成的两个线性子空间的和与交的基.

$$(1) \alpha = (1, 2, 1), \quad \alpha = (1, 1, -1), \quad \alpha = (1, 3, 3);$$

$$\beta = (2, 3, -1), \quad \beta = (1, 2, 2), \quad \beta = (1, 1, -3).$$

解 记 $W_1 = \langle \alpha, \alpha, \alpha \rangle, W_2 = \langle \beta, \beta, \beta \rangle$, 则

$$W_1 + W_2 = \langle \alpha, \alpha, \alpha, \beta, \beta, \beta \rangle.$$

因为 $W_1 + W_2$ 的基就是向量组 $\alpha, \alpha, \alpha, \beta, \beta, \beta$ 的极大线性无关组, 为求极大线性无关组, 只需把这 6 个向量的坐标列排成 3×6 矩阵, 并对该矩阵进行初等行变换化为阶梯型矩阵, 得

$$\begin{array}{ccccccccc} 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \\ 2 & 1 & 3 & 3 & 2 & 1 & 0 & -1 & 1 & -1 & 0 & -1 \\ 1 & -1 & 3 & -1 & 2 & -3 & 0 & 0 & 0 & -1 & 1 & -2 \end{array} \rightarrow \begin{array}{ccccccccc} 1 & 1 & 1 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 \\ 0 & -1 & 1 & -1 & 0 & -1 & 0 & 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & -1 & 1 & -2 & 0 & 0 & 0 & -1 & 1 & -2 \end{array}.$$

(*)

由于阶梯型矩阵的 1, 2, 4 列线性无关, 所以 α, α, β 是 $W_1 + W_2$ 的一组基; 又对任意 $\alpha \in W_1 \cap W_2$, 有

$$\alpha = k_1 \alpha + k_2 \alpha + k_3 \alpha = \mu_1 \beta + \mu_2 \beta + \mu_3 \beta,$$

而由(*)式知 α 可由 α, α 线性表出, β 可由 β, β 线性表出, 所以可取 $k_3 = \mu_3 = 0$, 即只要解线性方程组

$$\begin{pmatrix} \alpha^T & \alpha^T & \beta^T & \beta^T \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ -\mu_1 \\ -\mu_2 \end{pmatrix} = \mathbf{0}.$$

利用(*)式(划去第 4, 6 列)易知解为

$$c(2, 1, -1, -1)^T, \quad c \in \mathbb{R},$$

所以 $W_1 \cap W_2$ 是 1 维的, 基为

$$2\alpha + \alpha = \beta + \beta = (3, 5, 1).$$

$$(2) \alpha = (1, 2, 1, -2), \alpha = (2, 3, 1, 0), \alpha = (1, 2, 2, -3);$$

$\beta = (1, 1, 1, 1), \beta = (1, 0, 1, -1), \beta = (1, 3, 0, -4)$.

解 同(1)有 $W_1 + W_2 = \langle \alpha, \alpha, \alpha, \beta, \beta, \beta \rangle$, 由于

$$\begin{aligned}
 (\alpha^T, \alpha^T, \alpha^T, \beta^T, \beta^T, \beta^T) &= \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 & 0 & 3 \\ 1 & 1 & 2 & 1 & 1 & 0 \\ -2 & 0 & -3 & 1 & -1 & -4 \\ 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & -1 & -2 & 1 \\ 0 & 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & -5 & 0 \end{pmatrix} \\
 &\rightarrow \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & -1 & -2 & 1 \\ 0 & 0 & 1 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & -5 & 0 \end{pmatrix},
 \end{aligned}$$

故知 $\alpha, \alpha, \alpha, \beta$ 是 $W_1 + W_2$ 的一组基.

为求 $W_1 \cap W_2$ 的基, 需解齐次线性方程组

$$\begin{pmatrix} \alpha^T, \alpha^T, \alpha^T, \beta^T, \beta^T, \beta^T \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ -\mu \\ -\mu \\ -\mu \end{pmatrix} = \mathbf{0},$$

其基础解系为

$$\eta_1 = (2, -1, -1, 1, 0, 0)^T, \quad \eta_2 = (-5, 1, 2, 0, 0, 1)^T.$$

于是得 $\dim W_1 \cap W_2 = 2$, 其基为

$$\alpha = 2\alpha - \alpha - \alpha = -\beta = (-1, -1, -1, -1),$$

$$\alpha = -5\alpha + \alpha + 2\alpha = -\beta = (-1, -3, 0, 4).$$

也就是 β, β 为 $W_1 \cap W_2$ 的基.

$$(3) \quad \alpha = (1, 1, 0, 0), \quad \alpha = (0, 1, 1, 0), \quad \alpha = (0, 0, 1, 1);$$

$$\beta = (1, 0, 1, 0), \quad \beta = (0, 2, 1, 1), \quad \beta = (1, 2, 1, 2).$$

解 因为

$$\begin{array}{cccccc}
 1 & 0 & 0 & 1 & 0 & 1 \\
 1 & 1 & 0 & 0 & 2 & 2 \\
 0 & 1 & 1 & 1 & 1 & 1 \\
 0 & 0 & 1 & 0 & 1 & 2
 \end{array}
 \begin{array}{l}
 \\
 \text{经初等} \\
 \text{行变换} \\
 \\
 \end{array}
 \begin{array}{cccccc}
 1 & 0 & 0 & 1 & 0 & 1 \\
 0 & 1 & 0 & -1 & 2 & 1 \\
 0 & 0 & 1 & 2 & -1 & 0 \\
 0 & 0 & 0 & -2 & 2 & 2
 \end{array},$$

所以 $W_1 + W_2 = \langle \alpha, \alpha, \alpha, \beta, \beta, \beta \rangle$ 的基是 $\alpha, \alpha, \alpha, \beta$ (不惟一) .

又齐次线性方程组

$$\begin{pmatrix} \alpha^T, \alpha^T, \alpha^T, \beta^T, \beta^T, \beta^T \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ -\mu_1 \\ -\mu_2 \\ -\mu_3 \end{pmatrix} = \mathbf{0}$$

的基础解系为

$$\eta_1 = (1, 1, 1, -1, -1, 0)^T, \quad \eta_2 = (2, 0, 1, -1, 0, -1)^T.$$

所以 $\dim W_1 \cap W_2 = 2$, $\mathbf{e}, \mathbf{e}$ 是 $W_1 \cap W_2$ 的一组基, 其中

$$\mathbf{e} = \alpha + \alpha + \alpha = \beta + \beta = (1, 2, 2, 1),$$

$$\mathbf{e} = 2\alpha + 2\alpha = \beta + \beta = (2, 2, 2, 2).$$

讲评 以上 3 个求子空间的交与和的实例, 从不同角度回答了同学常问的一些问题 .

(1) 由于初等行变换不改变矩阵的秩, 特别不改变矩阵列之间的线性相关(无关)性, 故(1)中由(*)式, 即最后的阶梯型矩阵, 可以看出 α, α 线性无关, 而 α_3 可由 α, α 线性表出; β, β 线性无关, 而 β 可由 β, β 线性表出. 故下面在求 $W_1 \cap W_2$ 时, 任意 $\alpha \in W_1 \cap W_2$ 有

$$\alpha = k_1 \alpha + k_2 \alpha = \mu_1 \beta + \mu_2 \beta.$$

即只用 W_1 或 W_2 的基线性表出, 这时 k_1, k_2 (或 μ_1, μ_2) 相当于 α 在 α, α 上的坐标(或 β, β 上的坐标), 它是惟一的. 如果把 α, β 的线性组合也写上, 则求出的系数 k_1, k_2, k_3 (或 μ_1, μ_2, μ_3) 可能

会出现这种情况,即 $k_1 \neq 0, k_2 \neq 0, k_3 \neq 0$, 但 $\alpha = k_1 \alpha + k_2 \alpha + k_3 \alpha = 0\alpha + 0\alpha + 0\alpha$. 也就是说 (k_1, k_2, k_3) 与 α 不是 1-1 对应的.

(2) 由于(2)和(3)中都有 $\dim W_1 = 3, \dim W_2 = 3$, 故在求 $W_1 \cap W_2$ 时, 两组系数都是含 3 个数. 于是都是求解‘含 6 个变量, 4 个方程’的齐次线性方程组, 由求 $W_1 + W_2$ 时得到的阶梯型矩阵的秩为 4, 故知该齐次线性方程组的解空间是两维的. 这决定了

$$\dim W_1 \cap W_2 = 2,$$

即齐次线性方程组 $k_1 \alpha^T + k_2 \alpha^T + k_3 \alpha^T - \mu \beta^T - \mu \beta^T - \mu \beta^T = 0$ 的解空间的维数等于 $\dim W_1 \cap W_2$ 的充要条件是 α, α, α 是 W_1 的基, β, β, β 是 W_2 的基.

例 15 证明: 每一个 n 维线性空间都可以表示成 n 个一维子空间的直和.

证 设 $\alpha, \alpha, \dots, \alpha$ 是 $V_n(F)$ 的一组基, 则

$$V_n(F) = \langle \alpha, \dots, \alpha \rangle = \langle \alpha \rangle + \langle \alpha \rangle + \dots + \langle \alpha \rangle,$$

其中 $\langle \alpha \rangle$ ($i = 1, 2, \dots, n$) 均是一维子空间, 又任意 $\alpha \in V_n(F)$, 有

$$\alpha = a_1 \alpha + a_2 \alpha + \dots + a_n \alpha, \quad \text{其中 } a_i \alpha \in \langle \alpha \rangle.$$

由于坐标 $a_1, a_2, \dots, a_n$ 惟一, 所以 $a_i \alpha$ 惟一, 即 α 关于子空间 $\langle \alpha \rangle, \langle \alpha \rangle, \dots, \langle \alpha \rangle$ 的分解式惟一, 所以

$$V_n(F) = \langle \alpha \rangle \oplus \langle \alpha \rangle \oplus \dots \oplus \langle \alpha \rangle.$$

讲评 这道题本身是容易的. 选在此是为了加深读者印象, 用时, 能从头脑中迅速调出来. 其次, 发散思维一下, 是否应有: 若 $\alpha, \alpha, \dots, \alpha$ 是 V 的基, 则

$$\begin{aligned} V &= \langle \alpha, \dots, \alpha \rangle \oplus \langle \alpha_{k+1}, \dots, \alpha_k \rangle \\ &= \langle \alpha, \dots, \alpha \rangle \oplus \langle \alpha_{k+1}, \dots, \alpha_{k+k} \rangle \oplus \langle \alpha_{k+k+1}, \dots, \alpha_k \rangle. \end{aligned}$$

即把基向量任意分组, 得到不同的子空间, 都有空间直和分解式.

例 16 证明: 和 $\sum_{i=1}^s V_i$ 是直和的充分必要条件是

$$V_i \cap \sum_{j=1}^{i-1} V_j = \{0\} \quad (i=2, \cdots, s).$$

证 方法 1(归纳法) 当 $s=2$ 时, 显然成立(定理 5), 假设 $s-1$ 个子空间时命题成立, 则当 s 个子空间时, 由

$$\begin{aligned} \sum_{i=1}^s V_i \text{ 为直和} \quad V_s + \sum_{i=1}^{s-1} V_i \text{ 为直和} \\ V_s \cap \sum_{i=1}^{s-1} V_i = \{0\}, \text{ 及 } V_i \cap \sum_{j=1}^{i-1} V_j = \{0\} \\ (i=2, \cdots, s-1) \end{aligned}$$

(由归纳假设,

$$\sum_{i=1}^{s-1} V_i \text{ 是直和} \quad V_i \cap \sum_{j=1}^{i-1} V_j = \{0\}, i=2, \cdots, s-1), \text{ 得证.}$$

方法 2 若有 i , 使得 $(V_1 + \cdots + V_{i-1}) \cap V_i \neq \{0\}$, 则必 $\forall \alpha \neq 0$, 使 $\alpha, -\alpha \in V_i$, 且 $\alpha, -\alpha \in V_1 + \cdots + V_{i-1}$. 设 α 有分解式:

$$\alpha = \alpha_1 + \cdots + \alpha_{i-1}, \text{ 其中 } \alpha_k \in V_k, (k=1, \cdots, i-1),$$

则 α_k 不全为 0, 于是有

$$0 = \alpha_1 + \cdots + \alpha_{i-1} + (-\alpha) = 0 + \cdots + 0.$$

即零向量的分解不惟一, 与已知 $\sum_{i=1}^s V_i$ 为直和矛盾, 所以必有

$$V_i \cap \sum_{j=1}^{i-1} V_j = \{0\}.$$

如果 $\sum_{i=1}^s V_i$ 不是直和, 则零向量表示不惟一, 即

$$0 = 0 + \cdots + 0 = \alpha + \alpha + \cdots + \alpha,$$

其中 α 不全为零. 设从后往前数第一个非零的为 α , 则有

$$\alpha = -(\alpha + \alpha + \cdots + \alpha_{i-1}),$$

又 $\alpha, \cdots, \alpha_{i-1}$ 不能全为 0 (否则 α 为零), 而

$$\alpha \in V_t, -(\alpha + \cdots + \alpha_{-1}) \in \sum_{j=1}^{t-1} V_j,$$

于是得
$$V_t \cap \sum_{j=1}^{t-1} V_j \neq \{0\}.$$

这与已知矛盾,所以必有 $\sum_{i=1}^s V_i$ 是直和.

讲评 这是把两个子空间的直和推广到有限个子空间的直和. 本题的充要条件 $V_i \cap \sum_{j=1}^{i-1} V_j = \{0\} (i=2, \cdots, s)$, 比起教材中的条件

$$V_i \cap \sum_{j \neq i} V_j = \{0\}, \quad i=1, \cdots, s$$

要弱. 故是更精细的结果. 我们可以从三维空间(几何上)去理解一下这个命题. 设 $W_1 = \langle \alpha_1 \rangle, W_2 = \langle \alpha_2 \rangle, W_3 = \langle \alpha_3 \rangle$, 则若

$$W_2 \cap W_1 = \{0\},$$

说明 $W_1 + W_2$ 是直和, 向量 α_1, α_2 生成一张过原点的平面 π . 又若

$$W_3 \cap (W_1 + W_2) = \{0\},$$

知 $\alpha_3 \notin \pi$, 故 $\alpha_1, \alpha_2, \alpha_3$ 为 $V_3(\mathbb{R})$ 的基, 从而有

$$V_3(\mathbb{R}) = \langle \alpha_1 \rangle \oplus \langle \alpha_2 \rangle \oplus \langle \alpha_3 \rangle = W_1 \oplus W_2 \oplus W_3.$$

例 17 令 V 是 x 的次数 $\leq n (n \geq 1)$ 的实系数多项式全体所构成的线性空间.

(1) 证明: V 中有给定实根 c 的全体多项式的集合 L 是 V 的一个子空间;

(2) 求 L 的维数.

证 (1) L 中加法封闭: " $f, g \in L, f + g \in L$."

事实上, 由

$$\begin{aligned} f(c) &= 0 \\ g(c) &= 0 \end{aligned} \quad (f + g)(c) = 0;$$

数乘封闭: " $\lambda \in \mathbb{R}, f \in L$, 由 $f(c) = 0 \quad \lambda f(c) = 0$."

所以 L 是子空间.

(2) 由 L 中有基

$$(x-c), (x-c)^2, \cdots, (x-c)^n.$$

它们都有实根 c , 且线性无关, 又任意 $f(x) \in L$, 有

$$f(x) = a_1(x-c) + a_2(x-c)^2 + \cdots + a_n(x-c)^n,$$

故

$$\dim L = n.$$

(事实上, 由 c 为 L 中多项式的根 故任意 $f(x) \in L$, 有

$$f(x) = (x-c)g(x), \quad \deg g(x) \leq n-1.$$

由泰勒 (Taylor) 展开, 有 $g(x) = a_1 + a_2(x-c) + \cdots + a_n(x-c)^{n-1}$, 于是知 $(x-c), (x-c)^2, \cdots, (x-c)^n$ 是 L 的生成元.)

讲评 由已知 V 是 $n+1$ 维的线性空间. 任意 $f(x) \in V$, 仿 3.2 节例 4 有 (或由泰勒展开)

$$\begin{aligned} f(x) &= a_n(x-c)^n + \cdots + a_1(x-c) + a_0 \\ &= (x-c)g(x) + f(c). \end{aligned}$$

令 $(x-c)g(x) = f_1(x)$, $f(c) = f_2(x)$, 故有

$$f(x) = f_1(x) + f_2(x) \quad f_1(x) \in L, \quad f_2(x) \in R.$$

从而有

$$V = L \oplus R.$$

即空间 V (次数 $\leq n$ 的多项式的全体) 分解为 (以 c 为根的多项式的全体) L 与 R 的直和 (R 的基 1 和 L 的基 $(x-c), (x-c)^2, \cdots, (x-c)^n$) 放在一起就是 V 的基

$$1, (x-c), (x-c)^2, \cdots, (x-c)^n.$$

4.3 考试题评析

例 1 设 W_1 是齐次线性方程组 $x_1 + x_2 + x_3 + x_4 + x_5 = 0$ 的解空间, W_3 是 W_1 在 $\mathbb{R}^5$ 中的补子空间, W_2 是齐次线性方程组

$$x_1 + x_2 + x_3 + x_4 = 0$$

$$x_3 + x_4 + x_5 = 0$$

的解空间 则:

$\dim W_1 \cap W_2 =$ _____, $W_1 \cap W_2$ 的基是 _____;

$\dim(W_1 + W_2) =$ _____, $W_1 + W_2$ 的基是 _____;

W_3 的基是 _____, $R^5 / W_2 =$ _____.

解 由于 $W_1 \cap W_2$ 恰是齐次线性方程组

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 0, \\ x_1 + x_2 + x_3 + x_4 &= 0, \\ x_3 + x_4 + x_5 &= 0. \end{aligned} \quad (1)$$

的解空间. 解此方程组(为了化阶梯形方便, 把第 1 个方程移到最后一行)

$$\begin{array}{ccccc|ccccc} 1 & 1 & 1 & 1 & 0 & & 1 & 1 & 1 & 1 & 0 \\ & & & & & -r_1 + r_3 & & & & & \\ 0 & 0 & 1 & 1 & 1 & & & 1 & 1 & 0 & , \\ & & & & & -r_3 + r_2 & & & & & \\ 1 & 1 & 1 & 1 & 1 & & 0 & 0 & 0 & 0 & 1 \end{array}$$

故由最后的阶梯型矩阵知 $x_5 = 0$, 自由未知量为 x_2, x_4 , 分别取

$$\begin{array}{l} x_2 \\ x_4 \end{array} = \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \text{ 和}$$

得基础解系

$$\eta_1 = (-1, 1, 0, 0, 0)^T, \quad \eta_2 = (0, 0, -1, 1, 0)^T.$$

η_1, η_2 就是 $W_1 \cap W_2$ 的基, 故 $\dim(W_1 \cap W_2) = 2$.

又由 $\dim W_1 = (5 - 1) = 4$, $\dim W_2 = (5 - 2) = 3$, 故由维数公式(定理 2), 知

$$\dim(W_1 + W_2) = 5.$$

因为 W_1 的基为

$$\mathbf{e}_1 = (1, -1, 0, 0, 0)^T, \quad \mathbf{e}_2 = (1, 0, -1, 0, 0)^T,$$

$$\mathbf{e}_3 = (1, 0, 0, -1, 0)^T, \quad \mathbf{e}_4 = (1, 0, 0, 0, -1)^T,$$

即 $W_1 = \langle \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4 \rangle$; 而 $W_2 = \langle \mathbf{e}_5, \mathbf{e}_6, \mathbf{e}_7 \rangle$, 其中

$$\mathbf{e}_5 = (1, -1, 0, 0, 0)^T = \mathbf{e}_1, \quad \mathbf{e}_6 = (1, 0, -1, 1)^T,$$

$$\mathbf{e}_7 = (0, 0, -1, 1, 0)^T,$$

故(由命题 2)

$$\begin{aligned} W_1 + W_2 &= L(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5, \boldsymbol{e}_6, \boldsymbol{e}_7) \\ &= \langle \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5 \rangle . \end{aligned}$$

最后一个等号是因为 $\boldsymbol{e}_6 = \boldsymbol{e}_1, \boldsymbol{e}_7 = \boldsymbol{e}_2 - \boldsymbol{e}_3$.

把 W_1 的基扩充为 R^5 的基

$$\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4, (1, 0, 0, 0, 0)^T ,$$

则由定理 4 知

$$W_3 = k(1, 0, 0, 0, 0)^T, \quad k \in R .$$

又把 W_2 的基扩充为 R^5 的基

$$\boldsymbol{e}_6, \boldsymbol{e}_7, \boldsymbol{e}_8, \boldsymbol{e}_9, \boldsymbol{e}_{10},$$

其中 $\boldsymbol{e}_8 = (0, 1, 0, 0, 0)^T, \boldsymbol{e}_9 = (0, 0, 0, 0, 1)^T$.由定理 5 得

$$R^5 / W_2 = L(\boldsymbol{e}_8 + W_2, \boldsymbol{e}_9 + W_2) .$$

评析 (1) 命题 1 给出了求两个子空间的交 $W_1 \cap W_2$ 的方法,但用在此处计算量会比较大和长(因向量有 5 个分量).此题解法根据 W_1, W_2 均是齐次线性方程组的解这一特点:任意 $\alpha \in W_1 \cap W_2$, 则 $\alpha \in W_1$, 即 α 满足第一个方程组, 又 $\alpha \in W_2$, 则 α 应满足后两个方程.故 α 满足这 3 个方程.有了这个思路, 交的问题一下子就解决了.若此题没要求 $W_1 + W_2$ 的话, 连 W_1 和 W_2 的生成元都不用去找.

(2) 若此题改为只求维数不求基的话, 计算量更小.因由已知易看出 $\dim W_1 = 4$ (1 个方程, 5 个未知量), $\dim W_2 = 3$ (5 个未知量, 且系数矩阵的秩为 2), $\dim W_1 \cap W_2 = 2$ (因方程组(1)中系数矩阵秩为 3, 未知量有 5 个), 于是知(由维数公式)

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim W_1 \cap W_2 = 5 .$$

再由定理 4 和定理 5 知

$$\begin{aligned} \dim W_3 &= \dim R^5 - \dim W_1 = 5 - 4 = 1, \\ \dim R^5 / W_2 &= \dim R^5 - \dim W_2 = 5 - 3 = 2 . \end{aligned}$$

例2 设 $N = \begin{pmatrix} 0 & 1 & & & \\ & W & W & & \\ & & W & 1 & \\ & & & & 0 \end{pmatrix}$ 为五阶矩阵, $A = N^2$, 则 A 的化

零空间 $N(A) = \underline{\hspace{2cm}}$, A 的列空间 $R(A) = \underline{\hspace{2cm}}$.

解 由

$$A = N^2 = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & & \\ & 0 & 0 & 1 & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix},$$

的元素特点及 $N(A)$ 是 $Ax = 0$ 的解空间, 知 $x_3 = x_4 = x_5 = 0$, x_1, x_2 任取. 故

$$N(A) = \langle (1, 0, 0, 0, 0)^T, (0, 1, 0, 0, 0)^T \rangle.$$

而 $R(A)$ 是由 A 的列向量生成的空间, 因 A 的前两列为零向量, 后3列线性无关, 是 A 的列的极大线性无关组, 故

$$R(A) = \langle (1, 0, 0, 0, 0)^T, (0, 1, 0, 0, 0)^T, (0, 0, 1, 0, 0)^T \rangle.$$

评析 这是一道基本题. 方阵 N 是一个很重要的矩阵, 在 Jordan 标准形的理论中发挥很大的作用, 它左乘在一个方阵 B 上, 相当于把 B 的各行向上移一行, 第 n 行换为零向量, 它右乘在一个方阵 B 上, 相当于把 B 的各列向后移一列, 第1列变为零向量. 现在 N^2 相当于 N 自己乘自己, 无论从行看, 从列看, 它的效果都是把斜排1向“右上角”移一格 ($N^5 = 0$, 故 N 是幂零方阵). 所以做这道题第一步 $A = N^2$ 要算对 (矩阵乘法运算熟的话, 看一眼, 就可以得出结果, 不用死记硬背“往哪个方向移?”). 接下来是两个基本概念“化零空间”和“列空间”, 这是要掌握的.

例3 设 A 是 n 阶实对称矩阵, $S = \{x | x^T Ax = 0, x \in \mathbb{R}^n\}$.

(1) 试给出 S 为 $\mathbb{R}^n$ 的子空间的充分必要条件, 并加以证明.

(2) 当 S 是子空间且 $r(\mathbf{A}) = r < n$ 时, 试求 $\dim S$.

解 因 $\mathbf{A}$ 是实对称矩阵, 故存在可逆阵 $\mathbf{P}$, 使

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_p & & \\ & -\mathbf{I}_q & \\ & & \mathbf{0} \end{pmatrix} \mathbf{P},$$

则

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{P}^T \begin{pmatrix} \mathbf{I}_p & & \\ & -\mathbf{I}_q & \\ & & \mathbf{0} \end{pmatrix} \mathbf{P} \mathbf{x} = \mathbf{y}^T \begin{pmatrix} \mathbf{I}_p & & \\ & -\mathbf{I}_q & \\ & & \mathbf{0} \end{pmatrix} \mathbf{y},$$

其中 $\mathbf{y} = \mathbf{P} \mathbf{x}$, 故

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \quad \mathbf{y}^T \begin{pmatrix} \mathbf{I}_p & & \\ & -\mathbf{I}_q & \\ & & \mathbf{0} \end{pmatrix} \mathbf{y} = 0$$

$$y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_{p+q}^2 = 0,$$

以下分情况讨论:

① $q = 0$ 时, 有

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \quad y_1^2 + \cdots + y_p^2 = 0$$

$$y_1 = \cdots = y_p = 0$$

$$\mathbf{y} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ * \\ \vdots \\ * \end{pmatrix} \quad \mathbf{x} = \mathbf{P}^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ * \\ \vdots \\ * \end{pmatrix},$$

故 S 是子空间, $\dim S = n - p = n - r$.

② $p = 0$ 时, 同上知 S 是子空间, $\dim S = n - q = n - r$.

(因为 $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0 \quad y_{p+1}^2 + \cdots + y_{p+q}^2 = 0 \quad y_{p+1} = \cdots = y_{p+q} = 0$

$$\mathbf{y} = (y_1, \cdots, y_p, 0, \cdots, 0, y_{p+q+1}, \cdots, y_n)^T$$

$$\mathbf{x} = \mathbf{P}^{-1}(y_1, \cdots, y_p, 0, \cdots, 0, y_{p+q+1}, \cdots, y_n)^T).$$

③ p 和 q 均非零时, S 不是子空间, 因为 S 对加法不封闭 如设

$$\alpha = \mathbf{P}^{-1}(1, 0, \cdots, 0, 1, 0, \cdots, 0)^T \in S,$$

$$\beta = \mathbf{P}^{-1}(1, 0, \cdots, 0, -1, 0, \cdots, 0)^T \in S,$$

以上 α 中的第 2 个 1 及 β 中的 -1 都是第 $p+1$ 位分量. 显然 $\alpha^T A \alpha = 0, \beta^T A \beta = 0$, 但 $\alpha + \beta \notin S$, 因为

$$(\alpha + \beta)^T A (\alpha + \beta) = \begin{pmatrix} 2 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 4 \neq 0.$$

于是得结论:

(1) S 为 $\mathbb{R}^n$ 的子空间的充要条件是 A 是半定的 (即半正定或半负定);

(2) 当 S 是子空间时, $\dim S = n - r(A) = n - r$.

评析 (1) 这是一道比较难的题. 思路从哪里来? 由 A 是实对称矩阵, 故 $\mathbf{x}^T A \mathbf{x}$ 是实二次型. 题设中要求 $\mathbf{x}^T A \mathbf{x} = 0$, 故可联想到二次型中一些结论: 若对任意 $\mathbf{x}$, 有 $\mathbf{x}^T A \mathbf{x} = 0$, 则 $A = 0$ (1.2 节例 5); 若有 $\mathbf{x}^T A \mathbf{x} > 0, \mathbf{x}^T A \mathbf{x} < 0$, 则必有 $\mathbf{x}$ 使 $\mathbf{x}^T A \mathbf{x} = 0$ (1.2 节例 10). 现在要求的不仅是存在 $\mathbf{x}$, 使 $\mathbf{x}^T A \mathbf{x} = 0$, 而且要求这样的 $\mathbf{x}$ 全体构成一个线性空间 (是 $\mathbb{R}^n$ 的子空间). 由以上这些联想, 将会有下一步的思路, 即把二次型化为规范型. 走出这一步基本已上路, 以下就是讨论规范型为哪几种情况时, 满足 $\mathbf{x}^T A \mathbf{x} = 0$ 的 $\mathbf{x}$ 的全体, 构成线性空间. 故只要讨论 3 种情况 (因 A 为正定、负定矩阵时, S 均为零空间. 去掉这两种情况, 就是半正定、半负定和不定二次型.), 看 A 为哪种类型时, S 对加法和数乘封闭, 则得“充要条件” A 是哪种类型的二次型.

(2) 由于 $A\mathbf{x} = \mathbf{0}$ 的解空间是 $n - r(A)$ 维的, 且使 $A\mathbf{x} = \mathbf{0}$ 的 $\mathbf{x}$,

也使 $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$, 故当 S 构成子空间时, $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$ 与 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 同解, 即

$$S = W_A (\mathbf{A}\mathbf{x} = \mathbf{0} \text{ 的解的全体}).$$

而当 S 不是子空间时, $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$ 的解不一定是 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 的解 如

$$\mathbf{x} = (1, 1, 0)^T, \quad \mathbf{A} = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix},$$

则有 $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$, 而 $\mathbf{A}\mathbf{x} = (1, -1, 0)^T \neq \mathbf{0}$.

以下的两个证明思路都不对:

① $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0 \quad \mathbf{A}\mathbf{x} = \mathbf{0}$; (评: 这恰是要证的, 不能一开始就断言);

② 对任意 $\mathbf{x}, \mathbf{y} \in S$, 由 $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0, \quad \mathbf{y}^T \mathbf{A}\mathbf{y} = 0$ 应有

$$(\mathbf{x} + \mathbf{y})^T \mathbf{A}(\mathbf{x} + \mathbf{y}) = 0 \quad (\mathbf{x}^T + \mathbf{y}^T) \mathbf{A}(\mathbf{x} + \mathbf{y}) = 0,$$

即有

$$\mathbf{x}^T \mathbf{A}\mathbf{x} + \mathbf{y}^T \mathbf{A}\mathbf{y} + \mathbf{x}^T \mathbf{A}\mathbf{y} + \mathbf{y}^T \mathbf{A}\mathbf{x} = 0.$$

又因为 $(\mathbf{x}^T \mathbf{A}\mathbf{y})^T = \mathbf{y}^T \mathbf{A}\mathbf{x}$, 故从上式得

$$2 \mathbf{x}^T \mathbf{A}\mathbf{y} = 0 \quad \mathbf{x}^T \mathbf{A}\mathbf{y} = 0.$$

因为对任意 $\mathbf{x}$ 成立, 故得 $\mathbf{A}\mathbf{y} = \mathbf{0}, \dots\dots$.

(评: 不是任意 $\mathbf{x}$, 而是满足 $\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$ 中的 $\mathbf{x}$, 若 $\mathbf{x}$ 是 $\mathbb{R}^n$ 中任意向量, 则得 $\mathbf{A} = \mathbf{0}$ (当然也有 $\mathbf{A}\mathbf{x} = \mathbf{0}$), 故错误.)

例 4 已知 $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R}), W_1 = L(\mathbf{A})$ 是由矩阵 $\mathbf{A}$

生成的 $M_2(\mathbb{R})$ 的子空间, 则商空间 $M_2(\mathbb{R}) / W_1$ 的维数是多少? 试求它的一组基.

解 因为 $\dim M_2(\mathbb{R}) = 4$, 而 $\dim W_1 = 1$, 故由定理 5 知

$$\dim M_2(\mathbb{R}) / W_1 = 4 - 1 = 3.$$

又 $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ 为 W_1 的基, 将其扩充为 $M_2(\mathbb{R})$ 的基

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

则商空间的基为(见定理 5)

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + W_1, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + W_1, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + W_1.$$

讲评 关于“商空间”,主要结论就是定理 5.这里首先要弄清 W_1 的维数和生成元,由于它是由一个矩阵 A 生成的,故是一维空间.全部向量为

$$kA = \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix}, \quad k \in \mathbb{R}.$$

把 A 扩充为 $M_2(\mathbb{R})$ 的基,则扩充基所在的模 W_1 同余类就是商空间的基,而其扩充基的个数就是商空间的维数.

有的同学把 $L(A)$ 理解成 $\begin{pmatrix} 0 & -a \\ b & 0 \end{pmatrix}$, 故得 $L(A)$ 是两维的,这

显然是对生成空间的概念没理解好. A 是 $M_2(\mathbb{R})$ 中的一个向量,其生成的只能是一维空间.

$$\begin{aligned} \text{例 5 设 } W_1 = & \begin{pmatrix} a & 0 & c \\ a & b & 0 \\ c & d & 0 \end{pmatrix} \Bigg| a, b, c, d \in \mathbb{R}, \\ W_2 = & \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & z & z \end{pmatrix} \Bigg| x, y, z \in \mathbb{R}, \end{aligned}$$

则 $W_1 + W_2 =$ _____; $W_1 \cap W_2 =$ _____.

解 因为

$$W_1 = a(\mathbf{E}_{11} + \mathbf{E}_{21}) + b\mathbf{E}_{22} + c(\mathbf{E}_{13} + \mathbf{E}_{31}) + d\mathbf{E}_{32},$$

$$W_2 = x\mathbf{E}_{11} + y\mathbf{E}_{22} + z(\mathbf{E}_{32} + \mathbf{E}_{33}),$$

$$\begin{aligned} \text{故 } W_1 + W_2 &= \langle \mathbf{E}_{11} + \mathbf{E}_{21}, \mathbf{E}_{22}, \mathbf{E}_{13} + \mathbf{E}_{31}, \mathbf{E}_{32}, \mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{32} + \mathbf{E}_{33} \rangle \\ &= \langle \mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{33}, \mathbf{E}_{21}, \mathbf{E}_{32}, \mathbf{E}_{13} + \mathbf{E}_{31} \rangle. \end{aligned}$$

又因为 $\dim W_1 = 4, \dim W_2 = 3$, 故

$$\dim W_1 \cap W_2 = 4 + 3 - 6 = 1.$$

而直观可见 $E_{22} \in W_1$, 且 $E_{22} \in W_2$, 故知

$$W_1 \cap W_2 = \langle E_{22} \rangle.$$

评析 (1) 类似的题, 读者都会遇到过, 也曾多次作为考题, 只是每次非零元的位置不同. 遗憾的是有的同学做了这道, 又不会做那道, 其根本原因是没从本质上理解它们. 做这样的题, 第一步是找 W_1 和 W_2 的生成元, 然后找出其中的极大线性无关组. 如

$$\begin{aligned} W_1 = & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\ & = aF_1 + bF_2 + cF_3 + dF_4. \end{aligned}$$

因为上式右边的 4 个矩阵中元素 1 都在不同的位置上, 故它们不能互相线性表出, 故线性无关.

也可以这样证: 设有线性组合

$$k_1 F_1 + k_2 F_2 + k_3 F_3 + k_4 F_4 = \mathbf{0},$$

$$\begin{aligned} & \begin{pmatrix} k_1 & 0 & k_3 \\ k_1 & k_2 & 0 \\ k_3 & k_4 & 0 \end{pmatrix} = \mathbf{0}. \end{aligned}$$

即

从而得到 $k_1 = k_2 = k_3 = k_4 = 0$, 故 F_1, F_2, F_3, F_4 线性无关.

(2) 利用命题 2 可以写出 $W_1 + W_2$ 的生成元, 共 7 个. E_{22} 是重复的, 故可以去掉一个, 为了简化生成元的表达式, 把 E_{11} 的 (-1) 倍加到第一个向量上, 把 E_{32} 的 (-1) 倍加到最后一个向量上, 就得到 $W_1 + W_2$ 的基为 $E_{21}, E_{22}, E_{13} + E_{31}, E_{32}, E_{11}, E_{33}$. 这样在正确选定 $W_1, W_2, W_1 + W_2$ 的生成元后 $W_1 \cap W_2$ 的维数也出来了.

下面用另法求 $W_1 \cap W_2$ (记 $W_2 = \langle \mathbf{F}_5, \mathbf{F}_6, \mathbf{F}_7 \rangle$).

(3) 任意 $\mathbf{G} \in W_1 \cap W_2$, 有

$$\mathbf{G} = k_1 \mathbf{F}_1 + k_2 \mathbf{F}_2 + k_3 \mathbf{F}_3 + k_4 \mathbf{F}_4 = k_5 \mathbf{F}_5 + k_6 \mathbf{F}_6 + k_7 \mathbf{F}_7,$$

即有

$$\begin{pmatrix} k_1 & 0 & k_3 & & k_5 \\ k_1 & k_2 & 0 & & k_6 \\ k_3 & k_4 & 0 & & k_7 & k_7 \end{pmatrix} = \begin{pmatrix} k_5 \\ k_6 \\ k_7 & k_7 \end{pmatrix}.$$

故 $k_3 = 0, k_7 = 0, k_4 = 0, k_1 = 0, k_5 = 0, k_2 = k_6$,

即 $\mathbf{G} = k\mathbf{F}_2 = k\mathbf{F}_6 = k\mathbf{E}_{22}$. 故 $W_1 \cap W_2 = k\mathbf{E}_{22} \quad (k \in \mathbb{R})$.

例 6 已知

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & 1 \\ 3 & 1 & 9 \\ 2 & -3 & -5 \end{pmatrix}, \quad \alpha = \begin{pmatrix} 3 \\ -1 \\ 0 \\ -1 \end{pmatrix}.$$

试求 $N(\mathbf{A})$, $R(\mathbf{A})$ 的维数及一组基, 并判断向量 α 是否在 $R(\mathbf{A})$ 中.

解 对 $\mathbf{A}$ 做初等行变换得

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 & & -1 & 1 & & -1 & 1 & 1 \\ -1 & 1 & 1 & & 0 & 1 & 3 & & 0 & 1 & 3 \\ 3 & 1 & 9 & & 0 & 4 & 12 & & 0 & 0 & 0 \\ 2 & -3 & -5 & & 0 & -1 & -3 & & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 1 & & -1 & 1 & & -1 & 1 & 1 \\ -1 & 1 & 1 & & 0 & 1 & 3 & & 0 & 1 & 3 \\ 3 & 1 & 9 & & 0 & 4 & 12 & & 0 & 0 & 0 \\ 2 & -3 & -5 & & 0 & -1 & -3 & & 0 & 0 & 0 \end{pmatrix}.$$

从而知 $r(\mathbf{A}) = 2$ (因最后一个阶梯型矩阵的秩为 2, 其任意两列线性无关). 取 $\mathbf{A}$ 的前两列为 $R(\mathbf{A})$ 的基, 则有

$$R(\mathbf{A}) = L((2, -1, 3, 2)^T, (-1, 1, 1, -3)^T).$$

又 $N(\mathbf{A})$ 是 $\mathbf{A}\mathbf{x} = \mathbf{0}$ 的解的全体, 由上述阶梯型矩阵可知

$$\dim N(\mathbf{A}) = 3 - r(\mathbf{A}) = 1,$$

$$N(\mathbf{A}) = L((2, 3, -1)^T).$$

由于 $\mathbf{A}\mathbf{x} = \alpha$ 无解, 故 $\alpha \notin R(\mathbf{A})$. 事实上

$$(\mathbf{A}, \alpha) = \begin{pmatrix} 2 & -1 & 1 & 3 \\ -1 & 1 & 1 & -1 \\ 3 & 1 & 9 & 0 \\ 2 & -3 & -5 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 1 & 3 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & -7 \end{pmatrix},$$

显然 $r(\mathbf{A}) \neq r(\mathbf{A}, \alpha)$, 故 $\mathbf{A}\mathbf{x} = \alpha$ 无解.

讲评 (1) 由定义 $R(\mathbf{A})$ 是 $\mathbf{A}$ 的列向量生成的空间 (称为 $\mathbf{A}$ 的列空间), 故求 $R(\mathbf{A})$ 就是求它的基——也就是 $\mathbf{A}$ 的列的极大线性无关组, 它的维数等于这个极大线性无关组中所含向量的个数. 用初等行变换是为了看 $\mathbf{A}$ 中列之间的关系 (初等行变换不改变矩阵列之间的线性关系, 而阶梯型矩阵比较直观地看出这种线性关系). 但是, 有的同学却取阶梯型矩阵的列作为 $R(\mathbf{A})$ 的基, 这就错了. 我们很容易看出 $(-1, 0, 0, 0)^T \notin R(\mathbf{A})$, 且 $(1, 1, 0, 0)^T \notin R(\mathbf{A})$. 从理论上来说, 有

$$\mathbf{P}\mathbf{A} = \mathbf{B},$$

即 $\mathbf{A}$ 经行变换得 $\mathbf{B}$, 相当于其前乘方阵 $\mathbf{P}$, 故 $\mathbf{B}$ 的行是 $\mathbf{A}$ 的行的线性组合, 而 $\mathbf{B}$ 的列不是 $\mathbf{A}$ 的列的线性组合, 故 $\mathbf{B}$ 的列不属于 $R(\mathbf{A})$.

(2) 判断一个向量是否在某个空间里, 只要看它是否可由那个空间的生成元线性表出, 即是否存在 $\mathbf{x} = (k_1, k_2, \dots, k_s)^T$, 使

$$\alpha = k_1 \alpha_1 + \dots + k_s \alpha_s.$$

本题中 $\mathbf{A}$ 只有 3 列, 即看

$$\mathbf{A}\mathbf{x} = k_1 \alpha_1 + k_2 \alpha_2 + k_3 \alpha_3 = (\alpha_1, \alpha_2, \alpha_3) \mathbf{x} = \alpha$$

是否有解. 若有解说明 $\alpha \in R(\mathbf{A})$, 若无解, 则 $\alpha \notin R(\mathbf{A})$.

例 7 设 $V_1, V_2, \dots, V_s$ 是线性空间 V 的 s 个非平凡子空间, 证明 V 中至少有一个向量不属于 $V_1, V_2, \dots, V_s$ 中任意一个.

证 对子空间的个数 s 用归纳法.

当 $s=2$ 时, 由 4.2 节例 6 知, 结论成立.

今设对 $s-1$ 个非平凡的子空间结论成立, 即在 V 中存在向量

α , 使

$$\alpha \in V_i, \quad i = 1, \cdots, s-1.$$

对第 s 个子空间 V_s , 若 $\alpha \in V_s$, 则 α 为所求, 结论已对; 若 $\alpha \notin V_s$, 则由于 V_s 是非平凡子空间, 故必存在向量 $\beta \in V_s$. 于是, 对任意数 k , 有

$$k\alpha + \beta \in V_s,$$

且对不同的 $k_1, k_2, k_1\alpha + \beta, k_2\alpha + \beta$ 不会同时属于一个 $V_i (1 \leq i \leq s-1)$, 否则便有

$$(k_1\alpha + \beta) - (k_2\alpha + \beta) = (k_1 - k_2)\alpha \in V_i.$$

这与 $\alpha \notin V_i$ 矛盾.

今取 s 个互不相同的数 $k_1, k_2, \cdots, k_s$, 则 s 个向量

$$k_1\alpha + \beta, k_2\alpha + \beta, \cdots, k_s\alpha + \beta,$$

中至少有一个不属于任何一个 $V_i (1 \leq i \leq s-1)$, 它即为所求.

评析 (1) 理解这道题的一个简单的例子是: 在三维几何空间中每个过 y 轴的平面 W_i 是 $V_3(\mathbb{R})$ 的一个子空间. 当 i 为有限数时, 如 $i = 360$, 即从 xOy 平面算起, 每隔 1° 有一张平面. 虽然这些平面密密麻麻充满 $V_3(\mathbb{R})$, 但在任意两个之间都能找到向量不属于所有这 360 个平面. 而且满足条件的向量有很多.

(2) 本题的基础是 $s = 2$ 的情况, 即 4.2 节例 6. 故对 s 个子空间时, 要找这样的向量 $\alpha \notin V_i (1 \leq i \leq s-1)$ 和 $\beta \in V_s$ (此时, 若 $\beta \in V_i$ 或 $\alpha \in V_s$, 则结论显然). 故接下来要考虑 α, β 的线性组合. 事实上 $k_1, k_2, \cdots, k_s$ 可取为 $1, 2, \cdots, s$ 这 s 个自然数, 即

$$\alpha + \beta, 2\alpha + \beta, 3\alpha + \beta, \cdots, s\alpha + \beta$$

中至少有一个不属于任意一个 $V_i (1 \leq i \leq s-1)$ (把 $V_1, \cdots, V_{s-1}$ 看成 $s-1$ 个房间的话, 上述 s 个向量不能有两个在一个房间内, 故总有一个留在外面).

第5章 线性变换

5.1 理论脉络

这里所说的线性变换不包括读者已在课程“几何与代数(1)”中学过的线性变换内容(称为线性变换(1)).这里先把线性变换(1)的内容概述如下.

线性变换的定义:

设 V 是数域 F 上的线性空间, σ 是 V 上的一个变换, 若 σ 满足

$$(1) \sigma(\alpha + \beta) = \sigma\alpha + \sigma\beta, \text{ 对任意 } \alpha, \beta \in V,$$

$$(2) \sigma(k\alpha) = k\sigma\alpha, \text{ 对任意 } \alpha \in V, k \in F,$$

则称 σ 是 V 上的线性变换.

线性变换的性质:

$$(1) \sigma(0) = 0;$$

$$(2) \sigma(-\alpha) = -\sigma(\alpha);$$

(3) 若

$$\beta = k_1\alpha + k_2\alpha + \cdots + k_s\alpha \quad \sigma\beta = k_1\sigma\alpha + k_2\sigma\alpha + \cdots + k_s\sigma\alpha,$$

$$k_1\alpha + k_2\alpha + \cdots + k_s\alpha = 0 \quad k_1\sigma\alpha + k_2\sigma\alpha + \cdots + k_s\sigma\alpha = 0;$$

(4) σ 把线性相关的向量组映成线性相关的向量组. 但是线性变换可能把线性无关的向量组映成线性相关的向量组.

线性变换的运算:

(1) 线性变换的加法和数乘 " $\sigma, \tau \in L(V)$, 定义

$$(\sigma + \tau)\alpha = \sigma(\alpha) + \tau(\alpha), \quad \text{对任意 } \alpha \in V,$$

$$(k\sigma)\alpha = k\sigma(\alpha), \quad \text{对任意 } \alpha \in V, k \in F.$$

$L(V)$ 对如上定义的加法和数乘运算构成 F 上的线性空间.

(2) 线性变换的乘法 " $\sigma, \tau \in L(V)$, 定义

$$(\sigma\tau)(\alpha) = \sigma(\tau(\alpha)), \text{ 对任意 } \alpha \in V.$$

$L(V)$ 是线性空间(在加法和数乘下), 是环(在如上定义的加法和乘法下), 且数乘和乘法可交换, 故是一个“代数”.

线性变换的矩阵:

在 V 中取定一组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, 则 $\sigma\varepsilon_j (1 \leq j \leq n)$ 在基上的坐标列排成的方阵, 称为线性变换的矩阵, 形式地记为

$$\sigma(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (\sigma\varepsilon_1, \sigma\varepsilon_2, \dots, \sigma\varepsilon_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)A.$$

由此, $L(V)$ 与 $M_n(F)$ 之间建立了 1-1 对应 $\phi: \phi(\sigma) = A$, 且映射 ϕ 满足: 保加法、保数乘、保乘法. 即:

(1) 线性变换的和对应于矩阵的和 设 $\phi(\sigma) = A, \phi(\tau) = B$, 则有

$$(\sigma + \tau)(\varepsilon_1, \dots, \varepsilon_n) = (\varepsilon_1, \dots, \varepsilon_n)(A + B), \text{ 即 } \phi(\sigma + \tau) = A + B;$$

(2) 线性变换的数乘对应于矩阵的数乘

$$k\sigma(\varepsilon_1, \dots, \varepsilon_n) = (\varepsilon_1, \dots, \varepsilon_n)kA, \text{ 即 } \phi(k\sigma) = kA;$$

(3) 线性变换的乘积对应于矩阵的乘积

$$\sigma\tau(\varepsilon_1, \dots, \varepsilon_n) = (\varepsilon_1, \dots, \varepsilon_n)AB, \text{ 即 } \phi(\sigma\tau) = AB;$$

(4) 可逆的线性变换与可逆矩阵对应, 且逆变换对应逆矩阵

$$\sigma\tau = \tau\sigma = \varepsilon \text{ 与 } AB = BA = I_n \text{ 相对应, 当 } \sigma \text{ 可逆时, } A \text{ 可逆, } \tau =$$

σ^{-1} 与 $B = A^{-1}$ 对应, 故 $L(V) \xrightarrow{\phi} M_n(F)$ 是两个“线性代数”之间的同构.

线性变换(2)从《理工科代数基础》的第4章第3节开始. 这一章的内容很丰富, 是学习的重点, 也是难点, 特整理出以下7个问题.

1 线性变换的核与值域

定义1 设 σ 是 $V_n(F)$ 的线性变换, V 中向量在 σ 的作用下全

体象的集合称为 σ 的值域, 即

$$\text{Im}\sigma = \{\alpha\alpha \mid \alpha \in V\} = \sigma V.$$

$\dim(\text{Im}\sigma)$ 称为线性变换 σ 的秩.

定理 1 (1) $\text{Im}\sigma$ 是 $V_n(F)$ 的子空间.

(2) $\text{Im}\sigma = L(\sigma\varepsilon_1, \dots, \sigma\varepsilon_n)$, 其中 $\varepsilon_1, \dots, \varepsilon_n$ 是 V 的基. 故

$$\dim(\sigma V) = \text{rank}(\sigma\varepsilon_1, \dots, \sigma\varepsilon_n).$$

(3) $\dim(\text{Im}\sigma) = \text{rank} A$, 其中 A 是 σ 在基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵.

此定理的(2)给出怎样求值域 $\text{Im}\sigma$ 的方法(见 5.2 节例 7).

定义 2 设 σ 是 $V_n(F)$ 的线性变换, 所有被 σ 映成零向量的向量集合称为 σ 的核, 即

$$\text{Ker}\sigma = \{\alpha \in V \mid \alpha\alpha = 0\} = \sigma^{-1}(0).$$

$\dim(\text{Ker}\sigma)$ 称为线性变换 σ 的零度.

定理 2 (1) $\text{Ker}\sigma$ 是 $V_n(F)$ 的子空间.

(2) 设 σ 在 $V_n(F)$ 的基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵为 A , " $\alpha \in V$, 记 α 的坐标为 x , 则

$$\alpha\alpha = 0 \quad Ax = 0.$$

(3) " $\beta \in V$, 若 $\forall \alpha \in V$, 使 $\alpha\alpha = \beta$, 则

$$\sigma^{-1}\beta = \alpha + \text{Ker}\sigma = \{(\alpha + \alpha) \mid \alpha\alpha = 0, \alpha \in V, \alpha \in V\}.$$

此定理的(2)给出求 $\text{Ker}\sigma$ 的方法: 设 $x_1, \dots, x_{n-r}$ 为 $Ax = 0$ 的基础解系, 记 $\alpha = (\varepsilon_1, \dots, \varepsilon_n)x_i$, 则

$$\text{Ker}\sigma = L(\alpha_1, \alpha_2, \dots, \alpha_{n-r}).$$

由此定理的(3)得: σ 是单射 $\text{Ker}\sigma = \{0\}$.

定理 3 设 σ 是 $V_n(F)$ 上的线性变换, 则

$$\dim V = \dim(\text{Ker}\sigma) + \dim(\text{Im}\sigma).$$

推论 设 σ 是 $V_n(F)$ 上的线性变换, 则, σ 是单射 σ 是满射.

2 不变子空间

定义 3 设 σ 是 $V_n(F)$ 上的线性变换, W 是 V 的子空间, 如果对任意 $\alpha \in W$, 有 $\alpha\alpha \in W$, 则称 W 为 σ 的不变子空间.

几个重要的不变子空间的例子:

(1) $\text{Ker}\sigma$ 和 $\text{Im}\sigma$ 均为 σ 的不变子空间.

(2) 设 σ, τ 是 $V_n(F)$ 的线性变换, 若 $\sigma\tau = \tau\sigma$, 则 $\text{Ker}\tau, \text{Im}\tau$ 是 σ 的不变子空间.

(3) 由于 $f(\sigma)\sigma = \sigma f(\sigma)$, 故 $\text{Ker} f(\sigma)$ 与 $\text{Im} f(\sigma)$ 都是 σ 的不变子空间.

定理 4 设 σ 是 $V_n(F)$ 上的线性变换, W 是 V 的子空间, $\varepsilon_1, \dots, \varepsilon_k$ 为 W 的基, 并扩展为 V 的基 $\varepsilon_1, \dots, \varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_n$, 则

(1) W 是 σ 的不变子空间 σ 在此基下的方阵为

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}.$$

(2) $A_2 = 0$ $W_2 = \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$ 也是 σ 的不变子空间.

3 线性变换在其不变子空间上的限制

定义 4 设 σ 是 V 上的线性变换, W 是 σ 的不变子空间, 则

$$\sigma|_W : W \rightarrow W, \alpha \mapsto \sigma\alpha = \alpha\sigma$$

称为 σ 在 W 上的限制, 记为 $\sigma|_W$.

定义 5 设 σ 是 $V_n(F)$ 上的线性变换, W 是 σ 的不变子空间, 在 V/W 上定义变换

$$\tau: V/W \rightarrow V/W$$

$$\alpha + W \mapsto \tau(\alpha + W) = \sigma\alpha + W, \quad \alpha \in V,$$

称 τ 为 σ 诱导的变换.

4 线性变换的特征值和特征向量

定义 6 设 σ 是 $V_n(F)$ 上的线性变换, 若对 $\lambda \in F$, 存在非零向量 ξ , 使得 $\sigma\xi = \lambda\xi$, 则称 λ 是线性变换 σ 的一个特征值, ξ 是 σ 的属于特征值 λ 的特征向量.

定义 7 $V_\lambda = \{\xi \in V \mid \sigma\xi = \lambda\xi\}$ 称为线性变换 σ 的属于特征值 λ 的特征子空间. 特征子空间 V_λ 的维数叫做特征值 λ 的几何重数.

如何求 σ 的特征值和特征向量:

设 σ 在 V 中一组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的矩阵为 A , " $\alpha \in V_\lambda$, 有 $\sigma\alpha = \lambda\alpha$."

设 α 的坐标是 x , 则 $\sigma\alpha = \lambda\alpha$ 的坐标形式是 $Ax = \lambda x$. 于是, 求线性变换的特征值和特征向量问题转化为求矩阵的特征值、特征向量问题. 或 $\sigma\alpha = \lambda\alpha \quad (\sigma - \lambda\varepsilon)\alpha = 0$, 所以 $V_\lambda = \text{Ker}(\sigma - \lambda\varepsilon)$, 即

$$(\sigma - \lambda\varepsilon)\alpha = 0 \quad (A - \lambda I)x = 0.$$

定义 8 多项式 $f_A(\lambda) = \det(A - \lambda I)$ 称为 σ 的特征多项式, 它的根称为 σ 的特征根.

命题 1 设 W 是 σ 的一维不变子空间, 则 W 中任一非零向量是 σ 的特征向量; 反之, 每一个特征向量生成一个 σ 的一维不变子空间.

5 极小多项式

定义 9 设 $A \in M_n(F)$, $\sigma \in L(V)$, $f(x) = a_m x^m + \dots + a_1 x + a_0 \in F[x]$ 若

$$f(\sigma) = a_m \sigma^m + \dots + a_1 \sigma + a_0 \varepsilon = 0,$$

$$f(A) = a_m A^m + \dots + a_1 A + a_0 I = 0,$$

则称 $f(x)$ 是 σ (或 A) 的化零多项式.

定理 5 (Hamilton-Cayley 定理) 设 $f_A(\lambda) = \det(\lambda I - A)$, 则 $f_A(A) = 0$.

推论 设 σ 是 V 上的线性变换, $f_A(\lambda)$ 是 σ 的特征多项式, 则 $f(\sigma) = 0$.

即每个方阵 A 的特征多项式都是它自己的化零多项式 (也是它对应的线性变换的化零多项式).

定义 10 σ (或 A) 的化零多项式中次数最低的首一多项式称为 σ (或 A) 的极小多项式, 记为 $m_A(\lambda)$ 或 $m_\sigma(\lambda)$.

定理 6 对任意一个线性变换 σ , 极小多项式存在且惟一; 又 " $f(x) \in F[x]$, $f(\sigma) = 0 \quad m_A(x) \mid f(x)$."

推论 σ 的极小多项式是特征多项式的因子.

定理 7 设 A, B 均为 n 阶矩阵, 若 A 相似 B , 则

$$m_A(x) = m_B(x).$$

求极小多项式的方法:

(1) 利用极小多项式存在且惟一的证明过程, 有

若存在 d , 使 $\varepsilon, \sigma, \sigma^2, \dots, \sigma^{d-1}$ 线性无关, 而且 σ^d 可由前 d 个线性表出, 即

$$\sigma^d = a_0 \varepsilon + a_1 \sigma + a_2 \sigma^2 + \dots + a_{d-1} \sigma^{d-1},$$

则

$$m_A(x) = x^d - a_{d-1} x^{d-1} - \dots - a_1 x - a_0.$$

(2) 利用极小多项式是特征多项式的因式, 设

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s},$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 σ 的互异特征根, 则极小多项式应为

$$m_A(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s},$$

其中 $m_i \leq n_i$. 可以利用从低往高选取 $f(\lambda)$ 的因式的方式来确定 $m_i (i = 1, \dots, s)$. 首先令

$$f_1(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_s).$$

若 $f_1(A) = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_s I) = 0$,

则 $m_A(x) = f_1(x)$. 否则, 再令

$$f_2(\lambda) = (\lambda - \lambda_1)^2 (\lambda - \lambda_2) \cdots (\lambda - \lambda_s).$$

.....

(见 5.2 节例 21).

(3) 利用 Jordan 标准形. 设特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \quad \lambda_i \neq \lambda_j,$$

极小多项式为

$$m_A(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s}, \quad \text{其中 } m_i \leq n_i.$$

则 m_i 可由 A 的 Jordan 标准形来确定. 若

$$A \sim \text{diag}\{J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \dots, J_{n_s}(\lambda_s)\},$$

则 $m_i = J_{n_i}(\lambda_i)$ 中最大 Jordan 块的阶数.

6 线性变换可对角化的条件

σ 可对角化定义为在 V 中存在一组基, 使 σ 在此基下的矩阵是对角矩阵.

定理 8 σ 可对角化 $\iff \sigma$ 有 n 个线性无关的特征向量
 V 可分解成一维不变子空间的直和
 $\dim V_{\lambda_i} = n_i$ (几何重数 = 代数重数)
 V 可分解成 σ 的特征子空间的直和
 σ 的极小多项式无重根.

最后一条意为: 设特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \quad \lambda_i \neq \lambda_j,$$

则

$$\sigma \text{ 可对角化} \iff m_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}.$$

定义 11 形如
$$\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}$$
 的矩阵称为 Jordan 块. 由若

若干个 Jordan 块构成的准对角矩阵 $J_{n_i}(\lambda_i)$ 称为 Jordan 大块 (相对于特征值 λ_i 的), 由 Jordan 大块 $J_{n_i}(\lambda_i)$ 构成的准对角矩阵

$$\text{diag}\{J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \cdots, J_{n_s}(\lambda_s)\}$$

称为方阵 A 的 Jordan 标准形.

7 Jordan 标准形的计算

(1) 首先求 A 的特征多项式, 解出特征值.

$$f_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \quad \lambda_i \neq \lambda_j, \quad \sum_{i=1}^s n_i = n.$$

由 A 的特征多项式, 知

$$A \sim \text{diag}\{J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \cdots, J_{n_s}(\lambda_s)\},$$

即 Jordan 标准形是有 s 个大块的准对角矩阵, 每个大块对应于一个特征值 λ_i , 其阶数等于特征值 λ_i 的代数重数.

(2) 确定每个 $J_{n_i}(\lambda_i)$ 分两步:

① 求 $J_{n_i}(\lambda_i)$ 中 Jordan 块的个数, 它等于

$$\dim V_{\lambda_i} = n - r(\mathbf{A} - \lambda_i \mathbf{I})$$

= 属于特征值 λ_i 的线性无关的特征向量的个数;

② 求 k 阶 Jordan 块的个数

$$d_k = r(\mathbf{A} - \lambda_i \mathbf{I})^{k-1} - r(\mathbf{A} - \lambda_i \mathbf{I})^k - 2r(\mathbf{A} - \lambda_i \mathbf{I})^k,$$

或

$$d_k = [r(\mathbf{A} - \lambda_i \mathbf{I})^{k-1} - r(\mathbf{A} - \lambda_i \mathbf{I})^k] - [r(\mathbf{A} - \lambda_i \mathbf{I})^k - r(\mathbf{A} - \lambda_i \mathbf{I})^{k+1}];$$

$\geq k$ 阶块的个数 $> k$ 阶块的个数

(3) 写出 $\mathbf{A}$ 的 Jordan 标准形;

(4) 求可逆矩阵 $\mathbf{P}$

由 $\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \mathbf{J}$ $\mathbf{A} \mathbf{P} = \mathbf{P} \mathbf{J}$, 可解出 $\mathbf{P}$ 的各列(是 $\mathbf{A}$ 的特征向量和广义特征向量构成)(见 5.3 节例 2) .

定义 12 设 σ 是 $V_n(F)$ 上的线性变换, 若存在自然数 m , 使得 $\sigma^m = 0$ 而 $\sigma^{m-1} \neq 0$, 则称 σ 为幂零变换, m 称为 σ 的幂零次数 .

定义 13 设 σ 是 V 上的线性变换, $\alpha (\neq 0) \in V$, 若 $\alpha, \sigma\alpha, \dots, \sigma^{n-1}\alpha$ 构成 V 的一组基(称作循环基), 则称 σ 是一个循环变换. 此时称 V 是循环空间, α 为 σ 的循环向量. 当 $\sigma^n \alpha = 0$ 时, σ 在循环基 $\sigma^{n-1}\alpha, \dots, \sigma\alpha, \alpha$ 下的矩阵是

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

若对某个向量 β 存在 $k (< n)$, 使

$$\beta, \sigma\beta, \dots, \sigma^{k-1}\beta$$

线性无关, 而 $\beta, \sigma\beta, \dots, \sigma^{k-1}\beta, \sigma^k\beta$ 线性相关, 则称子空间

$$W = L(\beta, \sigma\beta, \dots, \sigma^{k-1}\beta)$$

是由 β 生成的循环子空间, 它是 σ 的不变子空间, $\sigma|_W$ 是 W 上的循

环变换.

定理 9 设 σ 是 $V_n(\mathbb{C})$ 上的线性变换, 则 V 中存在一组基, 使得 σ 在这组基下的矩阵是上三角矩阵.

推论 (Schur 定理) 任意 n 阶复矩阵总相似于上三角矩阵 (证明见 5.2 节例 15).

5.2 典型题讲评

例 1 设 M, N 是 V 的子空间, $V = M \oplus N$, 又设 $\alpha \in V, \alpha = \alpha_M + \alpha_N, \alpha_M \in M, \alpha_N \in N$, 定义 V 上的变换

$$\pi: \pi\alpha = \alpha_M.$$

试证明 π 是 V 上的线性变换, 这个变换 π 称为 V 对子空间 M 的投影变换.

证 对任意 $\alpha, \beta \in V, k \in F$, 设 $\alpha = \alpha_M + \alpha_N, \beta = \beta_M + \beta_N$, 则

$$\pi(\alpha + \beta) = (\alpha + \beta)_M = \alpha_M + \beta_M = \pi(\alpha) + \pi(\beta);$$

$$\pi(k\alpha) = (k\alpha)_M = k\alpha_M = k\pi(\alpha).$$

讲评 (1) 因为 $\alpha + \beta = \alpha_M + \alpha_N + \beta_M + \beta_N = (\alpha_M + \beta_M) + (\alpha_N + \beta_N)$, 又因 M, N 均为子空间, 故 $\alpha_M + \beta_M \in M, \alpha_N + \beta_N \in N$, 故 $(\alpha + \beta)_M = \alpha_M + \beta_M$.

(2) 由题中定义: 是先有空间分解, 后有变换. 这样的 π , 使任意 $\alpha \in M$, 有 $\pi\alpha = \alpha$, 任意 $\beta \in N$, $\pi\beta = 0$. 由此 (回答提问: 任意一个线性变换 σ 是否为 V 到 σV 的投影?) 知任意一个 V 上的线性变换 σ 不一定是 V 到 σV 的投影变换. 因为, 虽然仍有空间分解式 $V = \sigma V \oplus \sigma V^\perp$, 任取 $\alpha \in V$, 有

$$\alpha = \beta + \gamma, \quad \text{其中 } \alpha \in \sigma V, \gamma \in \sigma V^\perp.$$

但不一定有 $\sigma\alpha = \beta$! 特别地, 若 $\alpha \in \sigma V, \alpha = \alpha + 0$, 虽 $\sigma\alpha \in \sigma V$, 却不一定有 $\sigma\alpha = \alpha$. 请看下例.

在 $F_{n+1}[x]$ 中, σ 是微分变换, $\{1, x, \dots, x^{n-1}, x^n\}$ 为 $F_{n+1}[x]$

的一组基, 则 $F_{n+1}[x] = L(1, x, \cdots, x^{n-1}, x^n)$. 微分变换后得 $F_n[x]$, 故

$$\sigma V = L(1, x, \cdots, x^{n-1}),$$

$$F_{n+1}[x] = V = \sigma V \oplus \sigma V^{\perp} = F_n[x] \oplus L(x^n).$$

任取 $f = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} + a_n x^n \in V$,

有 $f = (a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}) + a_n x^n$
 $= \beta + \gamma,$

其中 $\beta = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} \in \sigma V, \gamma = a_n x^n \in \sigma V^{\perp}$. 但

$$\sigma f = a_1 + \cdots + (n-1)a_{n-1} x^{n-1} + n a_n x^{n-1} \in \sigma V, \sigma f \neq \beta.$$

故 σ 不是 V 到 σV 的投影变换.

例 2 在 $F[x]$ 中, $\sigma f(x) = f'(x), \tau f(x) = x f(x)$, 证明:

$$\sigma\tau - \tau\sigma = \varepsilon.$$

证 由于任取 $f(x) \in F[x]$, 有

$$\begin{aligned} (\sigma\tau - \tau\sigma) f(x) &= \sigma\tau f(x) - \tau\sigma f(x) = \sigma(x f(x)) - \tau(f'(x)) \\ &= f(x) + x f'(x) - x f'(x) = f(x), \end{aligned}$$

故得

$$\sigma\tau - \tau\sigma = \varepsilon.$$

讲评 对于无限维线性空间 $F[x]$, 其上定义的两个特殊线性变换 σ, τ , 我们证明了有关系式 $\sigma\tau - \tau\sigma = \varepsilon$. 这个结论对于有限维线性空间不成立. 即, 对任意 n 维线性空间 V 中的任意两个线性变换 A, B , 都有

$$AB - BA \neq \varepsilon.$$

证 设 $\varepsilon_1, \cdots, \varepsilon_n$ 为 V 的一组基, A, B 在此基下的矩阵分别为 A, B , 则只要证明

$$AB - BA \neq I.$$

由于 $\text{tr}(AB) = \text{tr}(BA)$, 故 $\text{tr}(AB - BA) = 0$, 而 $\text{tr} I = n$, 故

$$\text{tr}(AB - BA) \neq \text{tr} I, \text{从而得 } AB - BA \neq I.$$

例 3 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 V 的一组基, σ 是 V 上的线性变换. 证明 σ 可逆当且仅当 $\sigma\varepsilon_1, \sigma\varepsilon_2, \cdots, \sigma\varepsilon_n$ 线性无关.

证 因为

σ 可逆 A 可逆 (因 σ 的秩 = A 的秩)

A 的列线性无关 (因 $r(A) = A$ 的列秩)

$\sigma \varepsilon_1, \dots, \sigma \varepsilon_n$ 线性无关 (因 A 的列为 $\sigma \varepsilon_i$ 的坐标) .

讲评 这道题不难, 大部分人都会做, 这里是想说明怎样用最简单、最清晰的思路写出证明. 本题不用分开证充分性与必要性, 一起证就可以了. 最后一个“ ”是因 V 同构于 F^n , 故一组向量 $\sigma \varepsilon_1, \dots, \sigma \varepsilon_n$ 是否线性无关, 就看它们的坐标列 $a_1, a_2, \dots, a_n$ 是否线性无关.

例 4 设 σ 是线性空间 V 上的线性变换. 如果 $\sigma^{k-1} \xi \neq 0$, $\sigma^k \xi = 0$, 求证: $\xi, \sigma \xi, \dots, \sigma^{k-1} \xi (k > 0)$ 线性无关.

证 设有线性组合

$$\mu \xi + \mu_1 \sigma \xi + \dots + \mu_k \sigma^{k-1} \xi = 0.$$

上式两边作用 σ^{k-1} , 因 $\sigma^k \xi = \sigma^{k+1} \xi = \dots = \sigma^{2k-2} \xi = 0$, 故得

$$\mu \sigma^{k-1} \xi = 0,$$

从而得 $\mu = 0$. 同理, 若继续作用 σ^{k-2} , 得 $\mu_1 = 0$. 以此类推, 最后得

$$\mu_k \sigma^{k-1} \xi = 0,$$

从而得 $\mu_k = 0$ (因 $\sigma^{k-1} \xi \neq 0$). 即有 $\mu = \mu_1 = \dots = \mu_k = 0$, 所以 $\xi, \sigma \xi, \dots, \sigma^{k-1} \xi$ 线性无关.

讲评 由 $\sigma^{k-1} \xi \neq 0, \sigma^k \xi = 0$, 知有

$$\sigma(\sigma^{k-1} \xi) = 0,$$

故 $\lambda = 0$ 是 σ 的特征值, 而 $\sigma^{k-1} \xi$ 恰为 σ 属于特征值 0 的特征向量. 又

$$\sigma(\sigma^{k-2} \xi) = \sigma^{k-1} \xi,$$

所以 $\sigma^{k-2} \xi$ 是 σ 的广义特征向量. 同理, $\xi, \sigma \xi, \dots, \sigma^{k-3} \xi$ 均是广义特征向量, 故

$$\xi, \sigma \xi, \dots, \sigma^{k-2} \xi, \sigma^{k-1} \xi$$

恰为一个 Jordan 链. 也就是说我们证明了: 先解特征向量, 然后

由特征向量做方程组的常数项解得广义特征向量,再由这个广义特征向量去解另一个广义特征向量,得到的这一串向量肯定是线性无关的.因为 $\lambda=0$,故相当于解 $Ax=0$,得 x_1 ,再解

$$Ax_i = x_{i-1} \quad (i=2, \cdots, k),$$

得全部广义特征向量,则

$$x_k, x_{k-1}, \cdots, x_2, x_1 \text{ 线性无关.}$$

若 $\lambda \neq 0$,则特征向量满足 $(A - \lambda I)x = 0$,广义特征向量满足

$$(A - \lambda I)x_i = x_{i-1} \quad (i=2, \cdots, k).$$

这是说从解特征向量开始,解出 x_1 ,再代入上式解出 x_2 ,继续解非齐次线性方程组,……最后得 x_k .若这些向量都用 x_k 表示,即是

$$x_k, (A - \lambda I)x_k, (A - \lambda I)^2 x_k, \cdots, (A - \lambda I)^{k-1} x_k.$$

同学曾问:怎样保证这样解出的向量都线性无关.本题恰是一个严格的理论证明!

题设中的表示方式,也说明 Jordan 链是一个循环基.记

$$W = L(\xi, \sigma\xi, \cdots, \sigma^{k-1}\xi),$$

则 ξ 是 W 的循环向量, W 称为 V 的循环子空间.又记

$$\alpha = \sigma|_W,$$

则 $\alpha^k = 0$ ($\alpha^{k-1} \neq 0$, 因 $\forall \xi$ 使 $\alpha^{k-1}\xi \neq 0$),故 α 在 W 上是 (k 次)幂零变换.

例 5 设 σ, τ 是线性变换, $\sigma^2 = \sigma, \tau^2 = \tau$, 试证明:

(1) 如果 $(\sigma + \tau)^2 = \sigma + \tau$, 那么 $\sigma\tau = 0$;

(2) 如果 $\sigma\tau = \tau\sigma$, 那么 $(\sigma + \tau - \sigma\tau)^2 = \sigma + \tau - \sigma\tau$.

证 (1) 由 $(\sigma + \tau)^2 = \sigma + \tau$, 知

$$\text{左} = \sigma^2 + \tau^2 + \sigma\tau + \tau\sigma = \sigma + \tau + \sigma\tau + \tau\sigma = \text{右},$$

故有 $\sigma\tau + \tau\sigma = 0$, 从而 $\sigma\tau = -\tau\sigma$.

又 $\sigma\tau = \sigma^2\tau = \sigma(\sigma\tau) = -\sigma(\tau\sigma) = -(\sigma\tau)\sigma$

$$= \tau\sigma^2 = \tau\sigma, \quad (*)$$

代回 $\sigma\tau + \tau\sigma = 0$ 中, 有 $2\sigma\tau = 0$, 故 $\sigma\tau = 0$.

(2) 把 $\sigma\tau = \tau\sigma$ 代入下面的展开式中, 有

$$\begin{aligned}(\sigma + \tau - \sigma\tau)^2 &= \sigma^2 + \tau^2 + \sigma^2\tau^2 + 2\sigma\tau - 2\sigma^2\tau - 2\tau\sigma\tau \\&= \sigma + \tau + \sigma\tau + 2\sigma\tau - 2\sigma\tau - 2\sigma\tau \\&= \sigma + \tau - \sigma\tau.\end{aligned}$$

讲评 (1) 这两道小题目从本质上说没用到太多知识, 故大家可以会做. 事实上, 在推导过程中只是反复地用 $\sigma^2 = \sigma$, $\tau^2 = \tau$ 及 $\sigma\tau = -\tau\sigma$ 或 $\sigma\tau = \tau\sigma$.

那么有的读者为什么会感到难呢? 比如问题(1)做到

$$\sigma\tau + \tau\sigma = 0$$

就做不下去了. 我们认为根本原因是忘了“目标是什么?”要证 $\sigma\tau = 0$, 而由上式只能得到 $\sigma\tau = -\tau\sigma$ (你需要的是 $\sigma\tau = \tau\sigma$, 因若有此等式则上式可写成 $2\sigma\tau = 0$, 从而 $\sigma\tau = 0$), 那么当然想到. 能不能证出 $\sigma\tau = \tau\sigma$, 或说利用 $\sigma^2 = \sigma$, $\tau^2 = \tau$, $\sigma\tau = -\tau\sigma$ 能不能推出 $\sigma\tau = \tau\sigma$. 于是自然有了(*)式的过程. 这就是思路.

(2) 由此题的证明得到的结论是什么呢?

① 当 $\sigma, \tau, \sigma + \tau$ 均为幂等变换时, $\sigma\tau$ 必是零变换;

② 当 σ, τ 是两个乘法可变换的幂等变换时, $\sigma + \tau - \sigma\tau$ 也是幂等变换.

例6 设 V 是数域 F 上的 n 维线性空间, 证明由 V 的全体线性变换组成的线性空间 $L(V)$ 是 n^2 维的. 以二维线性空间为例, 写出 $L(V)$ 的一组基.

解 在 $V_n(F)$ 中取基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, 任取 $\sigma \in L(V)$, 设 σ 在此基下的矩阵表示为 A , 即

$$\sigma(\varepsilon_1, \dots, \varepsilon_n) = (\varepsilon_1, \dots, \varepsilon_n)A,$$

则映射

$$\Phi: L(V) \rightarrow M_n(F)$$

$$\sigma \mapsto A$$

是同构映射, 故 $L(V)$ 与 $M_n(F)$ 是两个同构的线性空间. 由 3.1 节命题 3, 知

$$\dim L(V) = \dim M_n(F) = n^2.$$

当 $n=2$ 时, 取 $V_2(F)$ 的基为 $\varepsilon_1, \varepsilon_2$. 由以上的同构映射知, 只要在 $M_2(F)$ 中取一组基(例如取自然基), 则这组基所对应的变换必然线性无关, 且是 $L(V)$ 的基. 故取

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

则分别以这 4 个矩阵为“矩阵表示”的线性变换 A_1, A_2, A_3, A_4 把基向量变为

$$A_1 \varepsilon_1 = \varepsilon_1, \quad A_2 \varepsilon_1 = 0, \quad A_3 \varepsilon_1 = \varepsilon_2, \quad A_4 \varepsilon_1 = 0,$$

$$A_1 \varepsilon_2 = 0; \quad A_2 \varepsilon_2 = \varepsilon_1; \quad A_3 \varepsilon_2 = 0; \quad A_4 \varepsilon_2 = \varepsilon_2.$$

因 $\{E_{ij}\}$ 是 $M_2(F)$ 的自然基, 故 A_1, A_2, A_3, A_4 是 $L(V)$ 的基.

讲评 (1) 这是一道很好的习题, $L(V)$ 是比较抽象的, 因为线性变换本身是抽象的. 但 $M_n(F)$ 相应地较直观, 通过两个空间的同构, 知道它们是维数相同的线性空间(事实上, 它们是两个“线性代数”同构). 这样很多问题可以用 $M_n(F)$ 来理解 $L(V)$ 的构造.

(2) 既然 $L(V)$ 是线性空间, 就可以选基. 怎么选? 选什么样的一组变换, 它们线性无关呢? (这个问题读者会感到抽象、困难.)

由同构可以通过找 $M_n(F)$ 中的基, 再转而找它们对应的 $L(V)$ 中的元素. 本题以二阶矩阵为例, 做了这件事. 推广来看, 对 n 维空间自然是选 $M_n(F)$ 的基 $\{E_{ij}\} (i, j=1, \cdots, n)$. 它们所对应的变换满足

$$A_{ij}(\varepsilon_1, \cdots, \varepsilon_n) = (\varepsilon_1, \cdots, \varepsilon_n) E_{ij}.$$

则 $\{A_{ij}\} (i, j=1, \cdots, n)$ 必为 $L(V)$ 的基. 那么 A_{ij} 是怎样对基向量起作用呢? 如 A_{12} , 有

$$A_{12} \varepsilon_1 = \varepsilon_2,$$

$$A_{12} \varepsilon_j = 0, \quad j \neq 2.$$

例 7 设 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 是四维线性空间的一组基, 已知线性变

换 σ 在这组基下的矩阵为

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 0 & -2 & -1 \\ 2 & 5 & 1 & 3 \\ 1 & 3 & 1 & 2 \end{pmatrix}.$$

(1) 试求 σ 在基 $\eta_1 = \varepsilon_1, \eta_2 = \varepsilon_1 + \varepsilon_2, \eta_3 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3, \eta_4 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4$ 下的矩阵;

(2) 求 σ 的核与值域;

(3) 在 σ 的核中, 选一组基, 把它扩充为 V 的一组基, 并求 σ 在这组基下的矩阵;

(4) 在 σ 的值域中选一组基, 把它扩充为 V 的基, 并求 σ 在这组基下的矩阵.

解 (1) 设 σ 在基 $\eta_1, \eta_2, \eta_3, \eta_4$ 下的矩阵为 B , 记从基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 到基 $\eta_1, \eta_2, \eta_3, \eta_4$ 的过渡矩阵为 T , 则由过渡矩阵的定义知

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(T 的列是新基在旧基上的坐标) 于是有

$$B = T^{-1}AT$$

$$= \begin{pmatrix} 1 & -1 & 0 & 0 & 1 & 2 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 & -2 & -1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 2 & 5 & 1 & 3 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 3 & 1 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 2 & 4 & 6 \\ -1 & -6 & -9 & -13 \\ 1 & 3 & 3 & 4 \\ 1 & 4 & 5 & 7 \end{pmatrix}.$$

(2) 由于任意 $\alpha \in \text{Ker}\sigma$, 有 $\sigma\alpha = 0$. 又设 α 的坐标为 x , 则 $\sigma\alpha$ 的坐标为 Ax , 故 $\sigma\alpha = 0 \iff Ax = 0$. 解齐次线性方程组 $Ax = 0$, 因

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 0 & -2 & -1 \\ 2 & 5 & 1 & 3 \\ 1 & 3 & 1 & 2 \end{pmatrix} \xrightarrow{\text{初等行变换}} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad *$$

得 $x = \alpha(2, -1, 1, 0)^T + \alpha(1, -1, 0, 1)^T$. 记

$$\alpha = 2\varepsilon_1 - \varepsilon_2 + \varepsilon_3, \quad \alpha = \varepsilon_1 - \varepsilon_2 + \varepsilon_4,$$

则有

$$\text{Ker}\sigma = L(\alpha, \alpha).$$

再利用(*)式知 $r(A) = 2$, 且 A 的任意两列线性无关. 故

$$\begin{aligned} \text{Im}\sigma &= L(\sigma\varepsilon_1, \sigma\varepsilon_2, \sigma\varepsilon_3, \sigma\varepsilon_4) \\ &= L(\sigma\varepsilon_1, \sigma\varepsilon_2) = L(\beta_1, \beta_2), \end{aligned}$$

其中 $\beta_1 = \sigma\varepsilon_1 = \varepsilon_1 + \varepsilon_2 + 2\varepsilon_3 + \varepsilon_4$, $\beta_2 = \sigma\varepsilon_2 = -2\varepsilon_1 + \varepsilon_2 + \varepsilon_4$.

(3) 在 $\text{Ker}\sigma$ 中取基 α, α , 扩充成 V 的基 $\alpha, \alpha, \varepsilon_1, \varepsilon_2$, 即有

$$(\alpha, \alpha, \varepsilon_1, \varepsilon_2) = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) \begin{pmatrix} 2 & 1 & 1 & 0 \\ -1 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\text{记为} \quad (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)P,$$

(显然 $|P| \neq 0$, 所以这样选取的 $\varepsilon_1, \varepsilon_2$ 使 $\alpha, \alpha, \varepsilon_1, \varepsilon_2$ 线性无关, 故可为 V 的基). 于是 σ 在基 $\alpha, \alpha, \varepsilon_1, \varepsilon_2$ 下的矩阵表示为

$$C = P^{-1}AP = \begin{pmatrix} 2 & 5 & 1 & 3 \\ 1 & 3 & 1 & 2 \\ -4 & -11 & -3 & -7 \\ 4 & 8 & 0 & 4 \end{pmatrix} P$$

$$= \begin{pmatrix} 0 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -4 & -11 \\ 0 & 0 & 4 & 8 \end{pmatrix}.$$

(4) 在 $\text{Im}\sigma$ 中取基 β, β , 扩充为 V 的基 $\beta, \beta, \varepsilon, \varepsilon$, 有

$$(\beta, \beta, \varepsilon, \varepsilon) = (\varepsilon, \varepsilon, \varepsilon, \varepsilon) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

由 $\beta, \beta, \varepsilon, \varepsilon$ 在基 $\varepsilon, \varepsilon, \varepsilon, \varepsilon$ 上的坐标列线性无关(显然, 上式右边的矩阵行列式非 0, 是准下三角阵), 故是基, 且上式中右边的矩阵就是两组基之间的过渡矩阵, 记为 Q , 则 σ 在新基 $\beta, \beta, \varepsilon, \varepsilon$ 下的矩阵表示为

$$D = Q^{-1} A Q = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad Q = \begin{pmatrix} 4 & -3 & 0 & 1 \\ 4 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

讲评 (1) 在理论证明中, 时常用到“在子空间 W 中取基, 并扩充为 V 的基”这样的叙述. 遇到具体问题时, 已知子空间的基, 怎样扩充为空间的基? 本题(3), (4)都回答了这个问题: 如果知道子空间的基的坐标列, 则将要扩充的向量的坐标需满足, 使得所有这些向量的坐标列排成的矩阵的行列式不为 0(于是, 该矩阵可逆, 列向量线性无关), 且应尽量简单、一目了然. 如(3)中 α, α 的两个坐标列后两个分量构成的二阶子式不为 0(一眼看出, 不用算), 故扩充的两个向量的前两个分量组成的二阶子式不等于 0,

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

$$\det \mathbf{P} = \begin{vmatrix} \mathbf{B}_1 & \mathbf{L} \\ \mathbf{L} & \mathbf{0} \end{vmatrix} = \det \mathbf{L} = 1 \neq 0,$$

故 $\alpha, \alpha, \varepsilon, \varepsilon$ 必线性无关. 又由于 α, α 的前两个坐标分量构成的二阶子式 $\begin{vmatrix} 2 & 1 \\ -1 & -1 \end{vmatrix} \neq 0$, 故也可扩充成 $\alpha, \alpha, \varepsilon, \varepsilon$, 即它们的坐标列排成的矩阵为

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

再进一步看, 因上述矩阵的前两列的任意二阶子式不为 0, 故, 后两列的两个 1 还可在其他的位置上. 所以, 可以总结为:

若前两列有一个二阶子式不为 0, 那么, 在这个二阶子式以外的行上分别取 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 和 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, 则得到的向量即为所求. 由此不难得出 n 维空间中 s 维子空间的处理方法.

(2) 在(3)中无论怎样扩充基向量 α, α , 线性变换 σ 在基 $\alpha, \alpha, \alpha, \alpha$ 上的矩阵表示的前两列总是零向量, 这是因为前两列分别为 $\sigma\alpha, \sigma\alpha$ 的坐标列, 而 $\sigma\alpha = \sigma\alpha = 0$, 故坐标均为 0 (因 $\alpha, \alpha \in \text{Ker}\sigma$).

而(4)中无论怎样扩充基向量 β, β , 线性变换 σ 在基 $\beta, \beta, \beta, \beta$ 上的矩阵表示的后两行的元素均为 0. 这是由于 $\sigma\beta, \sigma\beta, \sigma\beta, \sigma\beta \in \text{Im}\sigma$, 故它们均可由 β, β 线性表出. 所以都只是 β, β 的线性组合.

例 8 设 A 为线性空间 V 上的非奇异线性变换, 证明 A 的不变子空间均为 A^{-1} 的不变子空间.

证 方法 1 设 W 是 A 的不变子空间, 在 W 中取基 $\alpha, \dots, \alpha$, 并扩充为 V 的基 $\alpha, \dots, \alpha, \alpha_{s+1}, \dots, \alpha_n$, 则 A 在此基上的方阵表

示为 $A = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}$, 其中 A_1, A_2 均可逆(因 A 可逆).

又 A^{-1} 在上述基上的方阵表示为 $A^{-1} = \begin{pmatrix} A_1^{-1} & * \\ 0 & A_2^{-1} \end{pmatrix}$, 即仍为准上三角矩阵, 所以 W 是 A^{-1} 的不变子空间.

方法 2 设 W 是 A 的不变子空间, 在 W 中取基 $\alpha, \dots, \alpha$, 则因 A 可逆, 所以 $A\alpha, \dots, A\alpha$ 亦线性无关且是 W 的基. 于是任取 $\beta \in W$, 有

$$\beta = k_1 A\alpha + \dots + k_r A\alpha = A(k_1 \alpha + \dots + k_r \alpha),$$

所以 $A^{-1}\beta = k_1 \alpha + \dots + k_r \alpha \in W$.

方法 3 因为 A 可逆, W 为 A 的不变子空间, 所以 $A|_W = A_1$ 仍为可逆, 所以 $W = A_1 W$, 两边用 A^{-1} 作用得 $A^{-1}W = A^{-1}A_1 W = A^{-1}A W = W$, 故 W 是 A^{-1} 的不变子空间.

讲评 对任意 $W \subseteq V$, 要证 W 是某个变换 σ 的不变子空间, 只要证明对任意 $\alpha \in W$, 有 $\sigma\alpha \in W$ 即可. 故以上给出的 3 个证法都紧扣这一主题. 又本题中 A 是可逆线性变换, 故当 W 是 A 的不变子空间时, 事实上有 $W = A W$. 证法 1 的思路是用 5.1 节定理 4, 即 V 中必存在基使 A 的矩阵表示为准上三角矩阵. 证法 2 的思路是在题设条件下有 $W = \langle \alpha, \dots, \alpha \rangle = \langle A\alpha, \dots, A\alpha \rangle = A W$. 证法 3 是用 $A^{-1}W = W$ 来证明(当 W 是 A 的不变子空间时) W 是 A^{-1} 的不变子空间.

例 9 设 σ 是 $V_n(F)$ 上的线性变换, 证明:

- (1) 在 $F[x]$ 中有一次数 $\leq n^2$ 的多项式 $f(x)$, 使得 $f(\sigma) = 0$;
- (2) 如果 $f(\sigma) = 0, g(\sigma) = 0$, 那么对于 $f(x)$ 与 $g(x)$ 的最大公因式 $d(x)$, 有 $d(\sigma) = 0$;
- (3) σ 可逆的充分必要条件是有一常数项不为 0 的多项式 $f(x)$, 使得 $f(\sigma) = 0$.

解 (1) **方法 1** 因为 $\dim L(V) = n^2$, 故 $L(V)$ 中 $n^2 + 1$ 个向

量 $\varepsilon, \sigma, \sigma^2, \dots, \sigma^{n^2}$ 必线性相关, 即存在不全为零的数 $a_0, a_1, \dots, a_{n^2}$, 使

$$a_{n^2} \sigma^{n^2} + a_{n^2-1} \sigma^{n^2-1} + \dots + a_1 \sigma + a_0 \varepsilon = 0.$$

于是, 令

$$f(x) = a_{n^2} x^{n^2} + a_{n^2-1} x^{n^2-1} + \dots + a_1 x + a_0,$$

则有 $f(\sigma) = 0$.

方法 2 由于 σ 的特征多项式 $f_\sigma(\lambda)$ 是 σ 的化零多项式, 而 $\deg f_\sigma(\lambda) = n \leq n^2$ ($n > 1$), 故得证 (即取 $f(x) = f_\sigma(x)$).

(2) 当 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式时, 则存在 $u(x), v(x)$, 使得

$$u(x)f(x) + v(x)g(x) = d(x).$$

把 σ 代入上式, 并利用 $f(\sigma) = 0, g(\sigma) = 0$, 得

$$d(\sigma) = u(\sigma)f(\sigma) + v(\sigma)g(\sigma) = 0.$$

(3) 由(1)的结论, 存在 $f(x)$, 使得 $f(\sigma) = 0$.

设 $f(\sigma) = \sigma^m + a_{m-1}\sigma^{m-1} + \dots + a_1\sigma + a_0\varepsilon = 0$, 若 $a_0 \neq 0$, 则把常数项移到右边, 并提取 σ 得

$$\sigma(\sigma^{m-1} + a_{m-1}\sigma^{m-2} + \dots + a_1) = -a_0\varepsilon,$$

故知 σ 可逆.

方法 1 当 σ 可逆时, 设此多项式从低次往高次数, 第一个不为零的项为 $a_k\sigma^k$, 则由

$$\sigma^m + a_{m-1}\sigma^{m-1} + \dots + a_k\sigma^k = 0$$

左乘 $(\sigma^{-1})^k$, 得

$$\sigma^{m-k} + a_{m-1}\sigma^{m-1-k} + \dots + a_k\varepsilon = 0.$$

故令

$$g(x) = x^{m-k} + a_{m-1}x^{m-1-k} + \dots + a_k,$$

则 $g(x)$ 为所求常数项不为零的多项式.

方法 2 设 σ 的特征多项式为

$$f_\sigma(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0\varepsilon, \text{ 其中 } a_0 = \det A,$$

则 $f_{\sigma}(\sigma) = 0$ 且当 σ 可逆时, $a = \det A \neq 0$.

讲评 (1) 本题设计者的原意是考查“ $V_n(F)$ 上的线性变换的全体 $L(V)$ 是一个 n^2 维的线性空间”, 故应该是不用特征多项式是化零多项式这一结论也可以做出来. 再者 Hamilton-Cayley 定理本身的证明也很长(繁), 故不用为好.

(2) 要证明一个线性变换 σ 可逆, 依定义只要证存在 τ , 使 $\sigma\tau = \varepsilon$ (恒等变换), 故(3)的充分性就是用的 σ 可逆的这一定义, 对 σ 找到了

$$\tau = -\frac{1}{a}(\sigma^{m-1} + \cdots + a_1).$$

(3) 注意问题(3)的必要性可用“特征多项式是化零多项式”这一点来证, 只要证出 $f_A(\lambda)$ 的常数项不为零也满足题目要求. 但充分性的证明, 要注意不能用“设特征多项式的常数项不为零”这与原题条件不符(它只说“存在一常数项不为零的多项式 $f(x)$, 使 $f(\sigma) = 0$ ”, 没说 $f(x)$ 是特征多项式).

例 10 设 σ, τ 是线性空间 V 上的线性变换, 且满足 $\sigma^2 = \sigma$, $\tau^2 = \tau$, 试证:

(1) $\text{Im}\sigma = \text{Im}\tau$ 的充要条件是 $\sigma\tau = \tau, \tau\sigma = \sigma$;

(2) $\text{Ker}\sigma = \text{Ker}\tau$ 的充要条件是 $\sigma\tau = \sigma, \tau\sigma = \tau$.

解 方法 1

(1) 由 $\text{Im}\sigma = \text{Im}\tau$, 知任取 $\xi \in V, \eta \in V$, 使 $\sigma\xi = \tau\eta$, 两边作用 τ , 有

$$\tau\sigma\xi = \tau^2\eta = \tau\eta = \sigma\xi, \text{ 故 } \tau\sigma = \sigma;$$

同理, 在 $\sigma\xi = \tau\eta$ 两边作用 σ , 得 $\sigma\tau = \tau$.

因 $\sigma\tau = \tau, \tau\sigma = \sigma$, 有

$$\text{Im}\tau = \tau V = \sigma\tau V = \sigma V = \text{Im}\sigma,$$

$$\text{Im}\sigma = \sigma V = \tau\sigma V = \tau V = \text{Im}\tau,$$

故得 $\text{Im}\sigma = \text{Im}\tau$.

(2) 由题设 $\sigma^2 = \sigma, \tau^2 = \tau$, 故有空间分解

$$\begin{aligned} V &= \text{Ker}\sigma \oplus \text{Im}\sigma \\ &= \text{Ker}\tau \oplus \text{Im}\tau. \end{aligned}$$

则对任意 $\alpha \in V$, 有 $\alpha = \alpha + \alpha = \beta + \beta$, 其中 $\alpha, \beta \in \text{Ker}\sigma = \text{Ker}\tau$, $\alpha \in \text{Im}\sigma, \beta \in \text{Im}\tau$, 且有

$$\sigma\alpha = \alpha, \tau\beta = \beta.$$

于是得

$$\sigma\tau\alpha = \sigma(\tau\alpha) = \sigma(\beta) = \alpha(\text{因 } \sigma\beta = 0);$$

$$\tau\sigma\alpha = \tau(\sigma\alpha) = \tau(\alpha) = \alpha(\text{因 } \tau\alpha = 0).$$

故得 $\sigma\tau = \sigma, \tau\sigma = \tau$.

由 $\sigma\tau = \sigma$, 得 $\text{Ker}\sigma = \text{Ker}\sigma\tau \supset \text{Ker}\tau$. 又因 $\tau\sigma = \tau$, 故得

$$\text{Ker}\tau = \text{Ker}\tau\sigma \supset \text{Ker}\sigma, \text{故 } \text{Ker}\sigma = \text{Ker}\tau.$$

方法2 设线性变换 σ, τ 在线性空间 V 的基 $\varepsilon_1, \dots, \varepsilon_n$ 上的矩阵分别为 A, B .

(1) 当 $\text{Im}\sigma = \text{Im}\tau$ 时, 即有 $L(\sigma\varepsilon_1, \dots, \sigma\varepsilon_n) = L(\tau\varepsilon_1, \dots, \tau\varepsilon_n)$, 于是知 $R(A) = R(B)$, 故存在可逆矩阵 P , 使得 $A = BP$, 或 $B = AP^{-1}$, 则有

$$AB = AAP^{-1} = AP^{-1} = B \rightarrow \sigma\tau = \tau;$$

$$BA = BBP = BP = A \rightarrow \tau\sigma = \sigma.$$

当 $\sigma\tau = \tau, \tau\sigma = \sigma$ 时, 即有 $AB = B, BA = A$, 故得

$$R(B) = R(AB) \subset R(A),$$

$$R(A) = R(BA) \subset R(B).$$

从而 $R(A) = R(B)$, 即 $\text{Im}\sigma = \text{Im}\tau$.

(2) 由 $\text{Ker}\sigma = \text{Ker}\tau$, 知线性方程组 $Ax = 0$ 与 $Bx = 0$ 同解, 故存在可逆方阵 P, Q , 使得 $A = PB$, 或 $B = QA$, 于是有

$$AB = PBB = PB = A \rightarrow \sigma\tau = \sigma,$$

$$BA = QAA = QA = B \rightarrow \tau\sigma = \tau.$$

当 $\sigma\tau = \sigma, \tau\sigma = \tau$ 时, 则有 $AB = A, BA = B$.

又因 $Bx = 0$ 的解都是 $ABx = 0$ 的解, 故也是 $Ax = 0$ 的解 (因 $AB = A$) .

同理, $Ax = 0$ 的解都是 $BAx = 0$ 的解, 也即是 $Bx = 0$ 的解 (因 $BA = B$) . 故得

$$\text{null } A = \text{null } B, \text{ 即 } \text{Ker } \sigma = \text{Ker } \tau .$$

讲评 (1) 以上给出的两种证法分别用“线性变换的语言”和“ n 阶矩阵的语言”来叙述、证明. 通过这两种思路, 读者也把求 $\text{Ker } \sigma$ 和 $\text{Im } \sigma$ 等知识复习和联想了一遍 (因 $\alpha = 0 \Rightarrow Ax = 0$, 即 $\text{Ker } \sigma$ 与齐次线性方程组 $Ax = 0$ 的解的关系; $\text{Im } \sigma = (\varepsilon_1, \cdots, \varepsilon_n) R(A)$, 即 $\text{Im } \sigma$ 与 A 的列空间的关系) .

方法 2 (用矩阵语言的证明) 由清华大学生物系周喆同学给出 .

(2) 关于问题 (2) 的必要性的几点说明:

① 方法 1 中利用了《理工科代数基础》例 4.6 的结论: 若 $\sigma^2 = \sigma$, 则有:

(i) $V = \text{Im } \sigma \oplus \text{Ker } \sigma$;

(ii) 任取 $\alpha \in \text{Im } \sigma$, 有 $\sigma\alpha = \alpha$ (实际上因为 $\forall \xi \in V$, 使 $\alpha = \sigma\xi$, 于是有 $\sigma\alpha = \sigma^2\xi = \sigma\xi = \alpha$) .

② 还可以有证法 3 .

因 $\sigma^2 = \sigma$, 故对任意 $\alpha \in V$, 有 $(\sigma^2 - \sigma)\alpha = 0$, 即 $\sigma(\sigma - \varepsilon)\alpha = 0$. 从而得

$$(\sigma - \varepsilon)\alpha \in \text{Ker } \sigma = \text{Ker } \tau,$$

故有 $\tau(\sigma - \varepsilon)\alpha = 0$, 即 $(\tau\sigma - \tau)\alpha = 0$, 于是 $\tau\sigma = \tau$.

同理, 由 $\tau^2 = \tau$, 对于任意 $\alpha \in V$ 有

$$(\tau^2 - \tau)\alpha = \tau(\tau - \varepsilon)\alpha = 0, \text{ 即 } (\tau - \varepsilon)\alpha \in \text{Ker } \tau = \text{Ker } \sigma,$$

故有 $\sigma(\tau - \varepsilon)\alpha = 0$. 从而得 $\sigma\tau - \sigma = 0$, 即 $\sigma\tau = \sigma$.

由证法 3 可以看出, 因 $\sigma^2 = \sigma$, 及 $\tau^2 = \tau$, 故对任意 $\alpha \in V$, 均有

$$(\sigma^2 - \sigma)\alpha = 0, \quad (\tau^2 - \tau)\alpha = 0,$$

即

$$\sigma(\sigma - \varepsilon)\alpha = 0, \quad \tau(\tau - \varepsilon)\alpha = 0. \quad (*)$$

由于 $\text{Ker}\sigma = \text{Ker}\tau$, 故把上面两式 σ 和 τ 后面的向量 $(\sigma - \varepsilon)\alpha$ 和 $(\tau - \varepsilon)\alpha$ 互换位置, $(*)$ 式仍成立, 即

$$\sigma(\tau - \varepsilon)\alpha = 0, \quad \tau(\sigma - \varepsilon)\alpha = 0.$$

所以有 $\sigma\tau - \sigma = 0, \quad \tau\sigma - \tau = 0.$

事实上, 证法 1 中也是用了这点, 如 $\sigma\beta = 0, \tau\alpha = 0$ 等.

③ 学生易出错的地方: (i) “当 $\text{Ker}\sigma = \text{Ker}\tau$ 时, 必有 $\text{Im}\sigma = \text{Im}\tau$ ” 这是由于对直和的定义没弄清楚, 事实是 $\text{Im}\sigma$ 不一定等于 $\text{Im}\sigma$, 故不要把 $\text{Im}\sigma$ 扯进来. 如方法 3, 没提到 $\text{Im}\sigma$ 和 $\text{Im}\tau$, 方法 1 中虽提到, 但只是借用一下而已.

(ii) “任取 $\alpha \in \text{Ker}\sigma$, 有 $\sigma\tau\alpha = \sigma 0 = 0$. 又 $\alpha\alpha = 0$, 故 $\sigma\tau = \sigma$ ” 这是只对 $\text{Ker}\sigma$ 中的向量 α 证出有 $\sigma\tau = \sigma$. 而要证明这一关系在 V 上成立, 应对任意 $\alpha \in V$, 相同的关系式成立才行. 请读者注意刚才这段括号中证明没用到幂等变换的条件, 故肯定不对.

例 11 设 σ 是线性空间 $V_n(F)$ 上的线性变换, $f(x) = g(x)h(x)$ 是 F 上的多项式, 且有 $f(\sigma) = 0, (g(x), h(x)) = 1$, 证明

$$V = \text{Ker}g(\sigma) + \text{Ker}h(\sigma).$$

证 由 $(g, h) = 1$, 知存在 $u, v \in F[x]$, 使得 $ug + vh = 1$. 把 σ 代入得

$$u(\sigma)g(\sigma) + v(\sigma)h(\sigma) = \varepsilon. \quad (*)$$

" $\alpha \in V$, 上式两边对 α 作用, 得

$$\alpha = \varepsilon\alpha = u(\sigma)g(\sigma)\alpha + v(\sigma)h(\sigma)\alpha = \alpha + \alpha.$$

因

$$h(\sigma)\alpha = h(ug\alpha) = u(hg)\alpha = uf(\sigma)\alpha = 0,$$

$$g(\sigma)\alpha = g(vh\alpha) = v(gh)\alpha = vf(\sigma)\alpha = 0,$$

故

$$\alpha = u(\sigma)g(\sigma)\alpha \in \text{Ker}h(\sigma), \quad \alpha = v(\sigma)h(\sigma)\alpha \in \text{Ker}g(\sigma).$$

所以 $\alpha \in \text{Ker}h(\sigma) + \text{Ker}g(\sigma)$, 即 $\text{Ker}g(\sigma) + \text{Ker}h(\sigma) = V$.

又对任意 $\beta \in \text{Ker } g(\sigma) \cap \text{Ker } h(\sigma)$, 把式(*) 两边对 β 作用, 有

$$\beta = \varepsilon\beta = u(\sigma)g(\sigma)\beta + v(\sigma)h(\sigma)\beta = 0,$$

故 $\text{Ker } g(\sigma) \cap \text{Ker } h(\sigma) = \{0\}$, 即知 $V = \text{Ker } g(\sigma) \oplus \text{Ker } h(\sigma)$.

讲评 (1) 有不少同学说“证明题没有思路”. 现在以此题为例, 说一下思路. 先看清目标是什么? 要证明 $V = V_1 \oplus V_2$ (其中 $V_1 = \text{Ker } h(\sigma)$, $V_2 = \text{Ker } g(\sigma)$), 即要证, 对于任意 $\alpha \in V$, 有 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in V_1$, $\alpha_2 \in V_2$. 故关键一步是要找出分解式 ($\alpha = \alpha_1 + \alpha_2$). 再看题设中给的条件, 哪一个可以给出一个恒等式呢? 由互素可以有 Bezout 等式, 故从这里开始, 即由 $(g, h) = 1$, $v \mid u$, $v \in f[x]$, 使

$$ug + vh = 1.$$

用 σ 代替上式中的不定元 x , 有 $\varepsilon = u(\sigma)g(\sigma) + v(\sigma)h(\sigma)$. 两边对 α 作用, 立得

$$\alpha = \varepsilon\alpha = u(\sigma)g(\sigma)\alpha + v(\sigma)h(\sigma)\alpha = \alpha_1 + \alpha_2.$$

写到这一步, 我想大部分人都会看到“胜利的曙光”. 接下来要做 3 件事:

① 证 $\alpha_1 \in \text{Ker } h(\sigma)$; ② 证 $\alpha_2 \in \text{Ker } g(\sigma)$; ③ 证“ $+$ ”是“ $\oplus$ ”(即和是直和).

这道题确实内容丰富. 不仅要用 3.1 节定理 7 (互素的 Bezout 等式), 还要用到 4.1 节定理 3 (和是直和的定义是交等于 0)、4.1 节定义 1 (和的定义) 等. 故不仅要求看清已知条件, 要证明的结论, 还要能把相关的知识从大脑中调出来. 这样有思路, 有步骤, 才能“走到”目的地.

(2) 由 $f(\sigma) = 0$ 及 $f(x) = g(x)h(x)$, 得 $f(\sigma) = g(\sigma)h(\sigma) = 0$, 因此 $f(\sigma)$ 是 V 上的零变换. 事实上有 $V = \text{Ker } f(\sigma)$. 但要注意:

① 由 $g(\sigma)h(\sigma) = 0$, 不能推出 $g(\sigma) = 0$ 或 $h(\sigma) = 0$!.

因为设 σ 在 V 中基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵为 A , 则

$$g(\sigma)h(\sigma) = 0 \quad g(A)h(A) = 0.$$

而 $g(A)h(A) = \mathbf{0}$ 不能说必有 $g(A) = \mathbf{0}$ 或 $h(A) = \mathbf{0}$ (矩阵乘法“有零因子”, 即两个非零方阵相乘可以为 $\mathbf{0}$).

② $V = \text{Ker } f(\sigma) = \text{Ker } g(\sigma) \cap \text{Ker } h(\sigma)$, 说的是空间分解, 不是集合的并集 $V_1 \cup V_2$. 故不能说任取 $\alpha \in V$, 若 $\alpha \in \text{Ker } g(\sigma)$, 如何如何; 若 $\alpha \in \text{Ker } h(\sigma)$, 必有 $\alpha \in \text{Ker } g(\sigma)$. 事实上, 有许多 $\alpha \in V$, 但 α 既不属于 V_1 也不属于 V_2 . 它只是可以写成 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in V_1, \alpha_2 \in V_2$ 这种样子.

例 12 在 $F_n[x]$ 中 ($n > 1$), 求微分变换 σ 的特征多项式, 并证明 σ 在任何一组基下的矩阵都不可能是对角阵. 求 σ 的极小多项式 $m_\sigma(x)$.

证 方法 1 在 $F_n[x]$ 中取基 $1, x, x^2, \dots, x^{n-1}$, 则微分变换 σ 在此基下的矩阵 A 满足

$$\sigma(1, x, x^2, \dots, x^{n-1}) = (1, x, x^2, \dots, x^{n-1}) \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix},$$

故 $f_\sigma(\lambda) = f_A(\lambda) = \lambda^n$, 所以 $\lambda = 0$ 为 n 重根.

因 $V_\lambda = \text{Ker}(\sigma - \lambda E) = \text{Ker } \sigma$, 故

$\dim V_\lambda = n - \text{rank } A = n - (n - 1) = 1 \neq n$ (几何重数 $\neq$ 代数重数), 由 5.1 节定理 8 知 σ 不可对角化.

由 $A^{n-1} \neq \mathbf{0}, A^n = \mathbf{0}$, 故 $m_\sigma(x) = x^n$.

方法 2 若 σ 可对角化, 因 A 的特征值都为 0, 则存在可逆阵 P 使 $P^{-1}AP = \mathbf{0}$, 故 $A = \mathbf{0}$, 矛盾, 所以 σ 不可对角化.

讲评 (1) $F_n[x]$ 上的微分变换 σ 是幂零变换, 由 $\sigma^{n-1} \neq 0, \sigma^n = 0$, 故 σ 是幂零变换, 其幂零次数是 n .

(2) 由证明中知, σ 在基 $1, x, \dots, x^{n-1}$ 下的矩阵表示是对角线上元素为 0 的上三角矩阵 (对角线上应是 σ 的特征值), 记为 A . 则显然 $A \neq \mathbf{0}$. 那么, 本题恰好证明了

“非零的幂零方阵不可相似于对角阵”.

对变换来说,即是

“非零的幂零变换不可对角化.”

(3) 因 A 每乘方一次秩减少 1, 故 $r(A^{n-1}) = 1, r(A^n) = 0$.

所以极小多项式为 A^n .

例 13 如果 $\sigma, \dots, \sigma$ 是 $V_n(F)$ 的 s 个两两不同的线性变换, 那么在 V 中存在向量 α , 使 $\sigma\alpha, \sigma\alpha, \dots, \sigma\alpha$ 也两两不同.

证 由 $\sigma, \dots, \sigma$ 两两不同, 故 $\sigma - \sigma \neq 0 (1 \leq i < j \leq s)$, 即它们全都不是零变换. $\text{Ker}(\sigma - \sigma) \neq V$ 都是 V 的真子空间 (除可能有零空间外, 都是不平凡子空间) 这些子空间共有

$$s-1 + s-2 + \dots + 2 + 1 = \frac{s(s-1)}{2} \uparrow.$$

其中非平凡子空间的个数也是有限个. 故由 4.3 节例 7 知, V 中存在向量 α , 使得

$$\alpha \notin \text{Ker}(\sigma - \sigma) \quad (1 \leq i < j \leq s),$$

也就是说 $(\sigma - \sigma)\alpha \neq 0 \quad (1 \leq i < j \leq s).$

故得

$\sigma\alpha \neq \sigma\alpha (1 \leq i < j \leq s)$, 即 $\sigma\alpha, \sigma\alpha, \dots, \sigma\alpha$ 两两不同.

讲评 此题的思路是这样得到的. 要证 $\sigma\alpha, \sigma\alpha, \dots, \sigma\alpha$ 两两不同. 即证 $\sigma\alpha - \sigma\alpha \neq 0$, 进而 $(\sigma - \sigma)\alpha \neq 0$ (要证 $(\sigma - \sigma)\alpha = 0$ 成立, 只要证 $\alpha \in \text{Ker}(\sigma - \sigma)$, 现要证 $(\sigma - \sigma)\alpha \neq 0$, 就是说 $\alpha \notin \text{Ker}(\sigma - \sigma)$). 故问题转化为要证: 存在 $\alpha \in V$, 使 $\alpha \notin \text{Ker}(\sigma - \sigma) (1 \leq i < j \leq s)$. 由此联想到用 4.3 节例 7. 就可以解决问题.

例 14 设 σ 是 n 维线性空间 V 上的线性变换, W 是 V 的子空间, σW 是 W 的象空间, 证明

$$\dim(\sigma W) + \dim(\text{Ker}\sigma \cap W) = \dim W.$$

证 设 $\dim(\text{Ker}\sigma \cap W) = r, \dim W = m$, 在 $\text{Ker}\sigma \cap W$ 中选基

$\alpha_1, \dots, \alpha_r$, 并扩充为 W 的基

$$\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_m,$$

则有

$$\begin{aligned}\sigma W &= L(\sigma \alpha_1, \dots, \sigma \alpha_r, \sigma \alpha_{r+1}, \dots, \sigma \alpha_m) \\ &= L(\sigma \alpha_{r+1}, \dots, \sigma \alpha_m).\end{aligned}$$

最后一个等号是由于 $\sigma \alpha_1 = \dots = \sigma \alpha_r = 0$. 设有线性组合

$$\mu_{r+1} \sigma \alpha_{r+1} + \dots + \mu_m \sigma \alpha_m = 0,$$

即

$$\sigma(\mu_{r+1} \alpha_{r+1} + \dots + \mu_m \alpha_m) = 0,$$

故 $\mu_{r+1} \alpha_{r+1} + \dots + \mu_m \alpha_m \in \text{Ker} \sigma \cap W$, 它可由 $\alpha_1, \dots, \alpha_r$ 线性表出, 即

$$\mu_{r+1} \alpha_{r+1} + \dots + \mu_m \alpha_m = \mu_1 \alpha_1 + \dots + \mu_r \alpha_r,$$

也就是有

$$\mu_1 \alpha_1 + \dots + \mu_r \alpha_r - \mu_{r+1} \alpha_{r+1} - \dots - \mu_m \alpha_m = 0.$$

由于 $\alpha_1, \dots, \alpha_m$ 为 W 的基, 故所有 $\mu_i = 0$, 特别 $\mu_{r+1} = \dots = \mu_m = 0$, 故得 $\sigma \alpha_{r+1}, \dots, \sigma \alpha_m$ 线性无关, 它们是 σW 的基, 所以 $\dim \sigma W = m - r$,

即得

$$\dim \sigma W + \dim(\text{Ker} \sigma \cap W) = \dim W.$$

讲评 (1) 此题证明过程是仿本章定理 3 的证明, 而不能直接用本章定理 3 的结论, 即不能把公式 $\dim V = \dim \text{Im} \sigma + \dim \text{Ker} \sigma$ 用到子空间 W 上, 因本题的目的, 就是要说明, 对子空间 $W \subset V$ 有类似的维数公式.

(2) 同本章定理 3 一样, 这仅仅是一个空间维数之间的关系式, 不能保证有直和分解: 如 $V = F_5[x]$, $W = F_4[x]$, σ 是微分变换, 即对任意 $f \in F_5[x]$, $\sigma f(x) = f'(x)$ 则有

$$\sigma W = F_4[x], \text{Ker} \sigma \cap W = F,$$

故有

$$\dim W = 5 = \dim \sigma W + \dim(\text{Ker} \sigma \cap W) = 4 + 1.$$

而

$$\dim(\sigma W + \text{Ker} \sigma \cap W) = 4,$$

故

$$W \neq \sigma W + \text{Ker} \sigma \cap W.$$

(3) 此题证明不能用“在 W 中取基 $\alpha_1, \dots, \alpha_m$, 并设其中几个向量属于子空间 $\text{Ker} \sigma \cap W$ ”. 因不能保证 $\alpha_1, \dots, \alpha_m$ 中有向量在

$\text{Ker}\sigma \cap W$ 中. 注意, 在考虑子空间和空间的关系问题时, 必先在子空间中选基, 再扩充为空间的基. 这里 $\text{Ker}\sigma \cap W$ 是 W 的子空间.

例 15 证明: 对任意 n 阶复矩阵 A 总存在可逆矩阵 P , 使得

$$P^{-1}AP = \Gamma, \text{ 其中 } \Gamma = (\gamma_{ij}), \text{ 满足当 } i > j \text{ 时, } \gamma_{ij} = 0.$$

证 对 A 的阶数 n 用归纳法.

当 $n = 2$ 时, 设 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, 在 $\mathbb{C}$ 中 A 有两个特征值 λ_1, λ_2 ,

故至少有一个特征向量 x_1 , 使 $Ax_1 = \lambda_1 x_1$, 于是找 x_2 使得 x_1, x_2 线性无关, 且令 $P = (x_1, x_2)$, 则有

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & a'_{12} \\ 0 & \lambda_2 \end{pmatrix} \text{ 为上三角矩阵.}$$

设对 $n-1$ 阶矩阵结论成立. 而当 A 为 n 阶矩阵时, 它的特征多项式 $f_A(\lambda)$ 在复数域 $\mathbb{C}$ 上有 n 个根, 即 A 有 n 个特征值. 任取其中一个记为 λ_1 , 记 A 的属于特征值 λ_1 的特征向量为 x_1 , 扩充 $x_2, \dots, x_n$ 使 $x_1, x_2, \dots, x_n$ 为 $\mathbb{C}^n$ 的基, 即找 $x_2, \dots, x_n$, 使得

$$P = (x_1, x_2, \dots, x_n) \text{ 是可逆矩阵.}$$

于是有

$$A_1 = P^{-1}AP_1 = \begin{pmatrix} \lambda_1 & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \beta \\ 0 & A_{n-1} \end{pmatrix}$$

由归纳假设, 存在可逆矩阵 P_{n-1} , 使得

$$P_{n-1}^{-1}A_{n-1}P_{n-1} = \begin{pmatrix} \lambda_2 & * & \cdots & * \\ & \lambda_3 & \cdots & * \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

令 $P_2 = \begin{pmatrix} 1 \\ & P_{n-1} \end{pmatrix}$, 则 $P_2^{-1} = \begin{pmatrix} 1 & \\ & P_{n-1}^{-1} \end{pmatrix}$. 记 $P = P_2 P_1$, 则有

$$P^{-1}AP = P^{-1}A P = \begin{pmatrix} \lambda & \beta P_{n-1} \\ 0 & P_{n-1}^{-1} A_{n-1} P_{n-1} \end{pmatrix} = \begin{pmatrix} \lambda & \beta P_{n-1} \\ & \Gamma_{n-1} \end{pmatrix} = \Gamma_n,$$

即 A 相似于上三角矩阵 Γ_n , Γ_n 的对角元全是 A 的特征值.

讲评 (1) 这个证明是“构造性”的, 或称“可操作”的, 即给定一个具体矩阵, 可按定理的程序把它化为上三角矩阵. 又定理证明中用的是归纳法, 首次只找一个特征向量. 事实上, 在具体对一个矩阵操作时, 首次求得几个线性无关的特征向量, 就用几个. 即可有 A 相似于

$$B = \begin{pmatrix} \lambda_1 & & & & \\ & W & & & \\ & & \lambda_r & & \\ & & & \lambda_{r+1} & * & \dots & * \\ & & & & \lambda_{r+2} & \dots & * \\ & & & & & W & \dots \\ & & & & & & \lambda_n \end{pmatrix}.$$

前面为对角矩阵, 后面是上三角矩阵. 另外因 λ 可能为复数, 特征向量就会是复的, 故 P 为复矩阵.

(2) 如果每次只找一个特征向量(即严格化为上三角矩阵), 可以取单位向量, 扩充的线性无关向量可以取标准正交向量组, 即所找的

$$P = (x_1, x_2, \dots, x_n)$$

可以是酉矩阵(在实数域中称为正交矩阵, P 的列是两两正交且模为 1 的向量组).

例 16 设 $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 由 $x \mapsto Ax$ 定义了 $\mathbb{R}^2$ 上的线性变换

A . 求 A 的不变子空间.

解 事实上 A 是 A 在 $\mathbb{R}^2$ 的自然基下的矩阵表示. 自然基

$\mathbf{e} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. 任取 $\mathbf{x} \in \mathbb{R}^2$, 设 $\mathbf{x} = (x_1, x_2)^T$, 则

$$\mathbf{Ax} = \begin{pmatrix} 0 & -1 & x_1 \\ 1 & 0 & x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

所以当 $\mathbf{x} = \mathbf{0}$ 时, $\mathbf{Ax} = \mathbf{0}$; 当 $\mathbf{x} \in \mathbb{R}^2$ 时, $\mathbf{Ax} \in \mathbb{R}^2$, 故 $\mathbf{0}, \mathbb{R}^2$ 为 A 的不变子空间.

因为一维不变子空间是由特征向量生成的, 即满足 $A\xi = \lambda\xi$ 的非零向量, 为此先求 A 的特征值. 因

$$|\lambda \mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1 = 0, \text{ 得 } \lambda = \pm i,$$

因为 A 无实特征值, 所以无实特征向量, 故在 $\mathbb{R}^2$ 中 A 无一维不变子空间.

讲评 这是一个二维空间求不变子空间的计算题, 比较简单. 由 $\mathbb{R}^2$ 的子空间只能是零维, 一维, 二维, 一一验证是否为不变子空间即可. 对于维数高于二维的空间, 由命题 1, 线性变换 σ 的一维不变子空间都由特征向量生成, σ 的任意 r 维的不变子空间可利用 5.2 节例 15 化为上三角矩阵后求得.

例 17 设

$$\mathbf{A} = \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix},$$

A 在 $\mathbb{R}$ 上是否相似于对角形矩阵? 在 $\mathbb{C}$ 上呢?

解

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda - 6 & 3 & 2 \\ -4 & \lambda + 1 & 2 \\ -10 & 5 & \lambda + 3 \end{vmatrix} = (\lambda - 2)(\lambda^2 + 1) = 0,$$

所以 $\lambda = 2, \lambda_1 = i, \lambda_2 = -i$.

因为 A 有两个复特征根, 所以 A 没有 3 个线性无关的实特征向

量,所以在 $\mathbb{R}$ 上 A 不能相似于对角矩阵,在 $\mathbb{C}$ 上 A 相似于

$$\begin{pmatrix} 2 & & \\ & i & \\ & & -i \end{pmatrix}.$$

讲评 (1) 由本题看出, σ 可对角化的条件是指在复数域上. 如从上述过程看出 A 有3个不同特征值,这是它相似对角矩阵的充分条件,但因特征值中有复的,故只能在复数域中相似于对角矩阵.

(2) 在实数域中 A 只能相似于

$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

例 18 设 V 是复数域 $\mathbb{C}$ 上的 n 维线性空间, A 是 V 的线性变换,则 A 有任意 r 维的不变子空间($1 \leq r \leq n$).

证 因为 A 在 $\mathbb{C}$ 上有 n 个特征根 $\lambda_1, \dots, \lambda_n$ (重根计入),于是由5.2节例15,在 V 中存在一组基 $\varepsilon_1, \dots, \varepsilon_n$,使 A 在此基上的矩阵表示为上三角矩阵

$$\begin{pmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ & \lambda_2 & w & \cdots \\ & & w & b_{n-1n} \\ & & & \lambda_n \end{pmatrix},$$

则 $W_r = \langle \varepsilon_1, \dots, \varepsilon_r \rangle$ ($1 \leq r \leq n$) 是 A 的 r 维不变子空间.事实上,

任取 $\alpha \in W_r$,有 $\alpha = \sum_{i=1}^r x_i \varepsilon_i$,则 $A\alpha = \sum_{i=1}^r x_i A\varepsilon_i$,其中

$$A\varepsilon_i = b_{i1}\varepsilon_1 + \cdots + b_{i-1,i}\varepsilon_{i-1} + \lambda_i\varepsilon_i \in W_r \quad (1 \leq i \leq r).$$

所以 $A\alpha \in W_r$,故得证.

讲评 (1) 通过本题的证明过程,读者应学会: 对一个具体变换 A ,怎样求它的 $1, 2, \dots, n-1$ 维的不变子空间.另外对一些特

殊的矩阵,如准对角矩阵等可一眼看出它的 r 维不变子空间时,就不必再化成上三角矩阵.

(2) 这里说有 r 维不变子空间.证存在性同时也给出计算 1 个 r 维($1 \leq r \leq n$) 不变子空间的方法(各维找一个的方法),而要求全部的一维,二维, $\cdots$, $n-1$ 维不变子空间是一件非常复杂的事.

例 19 设 R^n 的线性变换 A 在基 $\alpha_1, \cdots, \alpha_n$ 下的矩阵表示为对角形矩阵,且对角线上元素互异,求 A 的所有不变子空间,共多少个?

证 由已知 $\alpha_1, \cdots, \alpha_n$ 均为 A 的分别属于特征值 $\lambda_1, \cdots, \lambda_n$ 的特征向量.对任意 k , $W = \langle \alpha_{i_1}, \cdots, \alpha_{i_k} \rangle$ 均是 A 的不变子空间(由 18 题),其中 $1 \leq i_1 < \cdots < i_k \leq n$.

又若 $W_1 \subset R^n$ 是 A 的不变子空间,则 $A_1 = A|_{W_1}$ 是 W_1 上的线性变换,其特征值是 A 的特征值的一部分,不妨设为 $\lambda_1, \cdots, \lambda_r$,且 $r = \dim W_1$,故存在 $\beta_1, \cdots, \beta_r \in W_1$,使 $A_1 \beta_1 = \lambda_1 \beta_1, \cdots, A_1 \beta_r = \lambda_r \beta_r$,由于 $A_1 \beta = A \beta$,所以 β 是 A 的特征向量,故 $\beta = k \alpha_i, k \neq 0$,所以 $W_1 = \langle \beta_1, \cdots, \beta_r \rangle = \langle \alpha_{i_1}, \cdots, \alpha_{i_r} \rangle$,即所有不变子空间均可由特征向量张成,共有

$$1 + C_n^1 + C_n^2 + \cdots + C_n^{n-1} + 1 = (1+1)^n = 2^n \text{ 个}.$$

讲评 (1) 由对角线上元素互异,知 A 有 n 个不同的特征值,故每个特征子空间 V_{λ_i} 均是一维的.故一维不变子空间共 n 个,而二维空间是由其中任意选两个 V_{λ_i} 和 V_{λ_j} 中各取一个向量,故共有 C_n^2 个, $\cdots$ 下面证明对任意 k , $W = \langle \alpha_{i_1}, \cdots, \alpha_{i_k} \rangle$ 是 A 的不变子空间.

因对任意 $\alpha \in W$, 有 $\alpha = \mu_1 \alpha_{i_1} + \cdots + \mu_k \alpha_{i_k}$, 则

$$\begin{aligned} \sigma \alpha &= \mu_1 \sigma \alpha_{i_1} + \cdots + \mu_k \sigma \alpha_{i_k} \\ &= \mu_1 \lambda_{i_1} \alpha_{i_1} + \cdots + \mu_k \lambda_{i_k} \alpha_{i_k} \in W. \end{aligned}$$

(2) 对角线上元素只要有二个相同,情况就发生了变化.不妨设 $\lambda_1 = \lambda_2$, 故 $V_{\lambda_1} = \mu_1 \alpha_1 + \mu_2 \alpha_2$. 此时 V_{λ_1} 中任意一个非零向量都

生成一个一维不变子空间,故 A 有无穷个一维不变子空间,由此二维不变子空间也是无穷多个.因除 V_{λ_1} 是二维的, V_{λ_1} 中任意一个非零向量与其余特征子空间中一个非零向量都组成二维不变子空间.三维, $\dots$, $n-1$ 维都如此,故特征值有重根时,情况比较复杂.因此当特征值有重根时,也不会让学生去计算全部的 r 维不变子空间.

例 20 证明: 复线性空间的任何两个可交换的线性变换必有公共的特征向量.

证 记这样的两个线性变换分别为 A, B , 由已知有 $AB = BA$. 又设 λ_0 是 A 的一个特征值(特征多项式在复数域上总有根), 相应的特征子空间记为 V_{λ_0} , 即 $V_{\lambda_0} = \text{Ker}(A - \lambda_0 E)$. 下证 V_{λ_0} 是 B 的不变子空间: 任取 $\alpha \in V_{\lambda_0}$, 有

$$(A - \lambda_0 E)B\alpha = B(A - \lambda_0 E)\alpha = B0 = 0,$$

所以 $B\alpha \in \text{Ker}(A - \lambda_0 E) = V_{\lambda_0}$, 故得证.

又记 $B_1 = B|_{V_{\lambda_0}}$, 则 B_1 是 V_{λ_0} 上的线性变换, 在 $\mathbb{C}$ 上有特征值 μ , 于是有特征向量 β , 使 $B_1\beta = \mu\beta$ ($\beta \neq 0, \beta \in V_{\lambda_0}$), 此即 $B\beta = \mu\beta$ 而 $A\beta = \lambda_0\beta$, 所以 β 是 A, B 的公共特征向量.

讲评 (1) 本题要证两个变换有公共的特征向量, 那么先从一个变换着手, 看看它有没有特征向量. 因是复线性空间, 故 A 所对应的矩阵在复数域上有特征值, 从而有特征向量(由此知, 复线性空间的条件不能改. 若改为实线性空间, 这一步就走不通了). 接下来自然是看 V_{λ_A} 中是否有 B 的特征向量, 于是要用到线性变换在子空间上的限制(在定义“限制”之前要证 V_{λ_A} 是 B 的不变子空间. 因只有 $BW \subset W$, 才能定义 W 上的线性变换). 此题不用 B_1 也行不通, 因 B 虽有特征值、特征向量, 但不能保证 B 的特征向量属于 V_{λ_A} .

(2) 可交换的条件在这里起了重要的作用. 进一步, 若 A, B 都可对角化的话, 则 A, B 有 n 个公共的线性无关的特性向量. 用“矩

阵的语言”来说,就是存在可逆矩阵 P , 使 $P^{-1}AP, P^{-1}BP$ 同时为对角矩阵. 把它用在实对称矩阵上, 就是 1.1 节定理 6 (即 A, B 均为 n 阶实对称矩阵, 且 $AB = BA$, 则存在正交矩阵 Q , 使得 $Q^{-1}AQ, Q^{-1}BQ$ 同时为对称矩阵, 其证明见 1.1 节例 19) .

例 21 求方阵 A 的最小多项式和相似标准形 .

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

解 由

$$A^2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \end{pmatrix} = 4A,$$

所以 $m(\lambda) = \lambda^3 - 4\lambda = \lambda(\lambda + 2)(\lambda - 2)$.

又因为 $r(A) = 2$, 所以 A 相似于 $\text{diag}\{0, 0, 2, -2\}$, 此为其 Jordan 标准形 .

讲评 (1) 这里用的是“求极小多项式的方法 1” . 因 I, A 线性无关, 又 A^2 不能用 I, A 线性表出, A^3 可由 I, A, A^2 线性表出 ($A^3 = 4A$), 即

$$A^3 - 4A = 0 \text{ 故得 } m_A(x) = x^3 - 4x.$$

(2) 也可以用其他两个方法求极小多项式 . 比如可以用方法 3 . 由

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda & -1 & 0 & -1 \\ -1 & \lambda & -1 & 0 \\ 0 & -1 & \lambda & -1 \\ -1 & 0 & -1 & \lambda \end{vmatrix} = \lambda(\lambda^3 - 2\lambda) - 2\lambda^2 \\ &= \lambda^2(\lambda + 2)(\lambda - 2), \end{aligned}$$

故特征值为 $\lambda = 0$ (二重根), $\lambda_2 = 2, \lambda_3 = -2$.

又因为 $r(A) = 2$, 故 $r(J) = 2$. 于是知 A 的 Jordan 标准形为

$$J = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 2 \\ & & & -2 \end{pmatrix},$$

极小多项式为

$$m_A(x) = x(x^2 - 4).$$

例 22 找一个三阶矩阵 A , 其最小多项式为 x^2 .

解 由 $m(x) = x^2$, 知 $A^2 = 0$, 所以特征值为 0, 且最大若当块为二阶的, 于是立得

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

由此易知形如

$$\begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix} \quad \text{或} \quad \begin{pmatrix} 0 & & \\ & 0 & 1 \\ & & 0 \end{pmatrix}$$

$P \quad \quad \quad P^{-1} \text{ 或 } P \quad \quad \quad P^{-1}$

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

(其中 P 为任意可逆矩阵) 的矩阵均满足要求.

讲评 (1) 因 $m_A(x) = x^2$, 故特征多项式 $P_A(\lambda) = \lambda^3$ (因 $m_A(x) \mid P_A(x)$, 且 A 是三阶矩阵), 即 A 的特征值全为 0. 又最大 Jordan 块为 2, 故 Jordan 标准形中一个一阶块, 一个二阶块.

(2) 本题不能推广到四阶矩阵. 即对四阶矩阵, 只知道其极小多项式为 $m_A(x) = x^2$, 不能确定它的 Jordan 标准形. 因

$$J_1 = \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \quad \text{和} \quad J_2 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}$$

都满足条件.当然若只要找一个的话, J_1, J_2 都满足要求.

例 23 设 A, B 为域 F 上 n 阶矩阵, 已知 AB 和 BA 的特征多项式相同. 它们的最小多项式是否相同?

解 $p_{AB}(\lambda) = \det(\lambda I - AB) = \det(\lambda I - BA) = p_{BA}(\lambda)$, 但 AB 和 BA 的极小多项式不一定相同. 如

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \text{ 则 } AB = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = BA,$$

$$p_{AB}(\lambda) = p_{BA}(\lambda) = \lambda^2, \text{ 而 } m_{AB}(\lambda) = \lambda^2, m_{BA}(\lambda) = \lambda.$$

讲评 (1) 事实上, 因极小多项式是特征多项式的因式, 故当特征多项式确定后, 它的因子有多种选择, 所以极小多项式就会不相同.

(2) 关于公式 $|\lambda I_n - AB| = |\lambda I_n - BA|$, 用分块矩阵的初等变换可证得.

5.3 考试题评析

例 1 设 V 是 n 维线性空间, $\sigma \in L(V)$, 且满足

$$\text{秩}(\sigma + \varepsilon) + \text{秩}(\sigma - \varepsilon) = n.$$

求证: σ 必可对角化 (即 V 中必有基, 使 σ 的矩阵表示为对角矩阵).

解 设 σ 在 V 中基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵表示为 A , 则

$$\text{秩}(\sigma + \varepsilon) = r(A + I), \text{秩}(\sigma - \varepsilon) = r(A - I),$$

故由题设得

$$r(A + I) + r(A - I) = n. \quad (*)$$

设 $r(A + I) = r$, 则 $(A + I)x = 0$ 的解空间为 $n - r$ 维的, 即 σ 的属于特征值 -1 的特征子空间 $V_{\lambda=-1}$ 是 $n - r$ 维的.

又由 $(*)$ 式知有 $r(A - I) = n - r$, 故线性方程组 $(A - I)x = 0$ 的解空间为 r 维, 即 $\dim V_{\lambda=1} = n - (n - r) = r$, 而

$$V_{\lambda=1} \cap V_{\lambda=-1} = \{0\},$$

故 $\dim(V_{\lambda=1} + V_{\lambda=-1}) = \dim V_{\lambda=1} + \dim V_{\lambda=-1} = n = \dim V$,
即

$$V = V_{\lambda=1} \oplus V_{\lambda=-1}.$$

在 $V_{\lambda=1}$ 中选基 $\eta_1, \dots, \eta_r$, 在 $V_{\lambda=-1}$ 中选基 $\eta_{r+1}, \dots, \eta_n$, 则 σ 在基 $\eta_1, \dots, \eta_r, \eta_{r+1}, \dots, \eta_n$ 下的矩阵表示为

$$D = \begin{pmatrix} I_r & \\ & -I_{n-r} \end{pmatrix},$$

即 σ 可对角化.

评析 (1) 此题要证明在题设条件下, σ 可对角化, 故立刻会联想到 5.1 节定理 8, 那里给了 5 个充要条件. 在这儿可以用两个, 也就有了两条思路: ①证 σ 有 n 个线性无关的特征向量; ②证 V 可分解为 σ 的特征子空间的直和 (这里是用的思路②). 但无论用哪个思路都得去考虑特征值和特征向量. 怎么求? 没有一个明显的式子可利用来先算特征值. 题设中只有一个条件, 先把它转化成矩阵形式得到 $r(A+I) + r(A-I) = n$ (由于 σ 的特征值和特征向量都是利用 σ 在某基下的矩阵求得的, 故当然想到这一步). 两个方阵的秩的和为 n , 则以这两个方阵为系数的线性方程组的解空间维数之和也是 n . 眼前豁然开朗, A 可能有 n 个线性无关的特征向量. 下面是最关键的一步: 一旦走出设 $r(A+I) = r$ 这一步, 立刻会想到线性方程组 $(A+I)x=0$ 的解空间是 $n-r$ 维的. 这说明什么? 若 $r < n$, 则 $(A+I)x=0$ 有非零解, 于是 $|A+I|=0$, 故 -1 为 A 的 (从而也是 σ 的) 特征值, 于是 $(A+I)x=0$ 的解空间恰为特征子空间. 这实际上是在考察学生对特征值、特征向量的定义理解的深度. 如果给一个具体数字的矩阵, 人人都会求它的特征值和特征向量: 使 $|\lambda I - A| = 0$ 的 λ 是特征值, 使 $(\lambda I - A)x=0$ 的非零向量 x 为特征向量, 而本题恰是逆向用这个基本概念: 若 $(\lambda I - A)x=0$ 有非零解, 则 $|\lambda I - A| = 0$, 即 λ 必为特征值.

(2) 若题中条件改为秩 $\sigma + \text{秩}(\sigma - \varepsilon) = n$, 则仍有 σ 可对角化, 证明完全相同, 只不过特征值变为 0 和 1.

(3) 若用“ σ 有 n 个线性无关的特征向量”这个思路证, 要用“属于不同特征值的特征向量线性无关”的命题来说明 $\eta, \cdots, \eta, \eta_{+1}, \cdots, \eta$ 线性无关.

$$\text{例 2 已知 } A = \begin{pmatrix} 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 1 \\ 0 & -2 & 0 & 2 \end{pmatrix},$$

(1) 试求可逆矩阵 P 及 Jordan 标准形 J 使得 $P^{-1}AP = J$;

(2) 试求 A 的极小多项式 $m_A(x)$.

解 (1) 由

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda - 2 & -1 & 0 & 1 \\ 0 & \lambda & 0 & 0 \\ 2 & 0 & \lambda & -1 \\ 0 & 2 & 0 & \lambda - 2 \end{vmatrix} = \lambda \begin{vmatrix} \lambda - 2 & 0 & 1 \\ 2 & \lambda & -1 \\ 0 & 0 & \lambda - 2 \end{vmatrix} \\ &= \lambda^2 (\lambda - 2)^2, \end{aligned}$$

故 A 的特征值为 $\lambda_1 = 2$ (二重), $\lambda_2 = 0$ (二重) (故知 A 的 Jordan 标准形有两个 Jordan 大块, 每个都是二阶矩阵).

由 $r(2I - A) = 3$, 知 $\dim V_{\lambda=2} = 1$; 由 $r(0I - A) = 3$, 知 $\dim V_{\lambda=0} = 1$. 故相应于特征值 2 和 0 的 Jordan 块均只有一个, 从而有

$$J = \begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.$$

对 $\lambda = 2$, 解 $(A - 2I)x = 0$, 对系数矩阵做初等变换得

$$\begin{array}{cccc|cccc} 0 & 1 & 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & -2 & 0 & 0 & 0 & 1 & 0 & 0 \\ -2 & 0 & -2 & 1 & 0 & 0 & 0 & -1 \\ 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \rightarrow$$

$$x_2 = 0,$$

故得

$$x_4 = 0,$$

$$x_1 = -x_3.$$

故全部解为 $c(1, 0, -1, 0)^T$ 取 $\mathbf{x} = (1, 0, -1, 0)^T$, 再解 $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{x}_1$, 得 $\mathbf{x} = (0, 0, 0, 1)^T + c(1, 0, -1, 0)^T$ 取 $\mathbf{x} = (0, 0, 0, 1)^T$.

对 $\lambda = 0$, 解 $(\mathbf{A} - 0\mathbf{I})\mathbf{x} = \mathbf{0}$, 由

$$\begin{array}{cccc|cccc} 2 & 1 & 0 & -1 & -2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & -2 & 0 & 2 & 0 & 0 & 0 & 0 \end{array} \rightarrow$$

$$x_4 = 0,$$

故得

$$x_2 = 0,$$

$$x_1 = 0,$$

$$x_3 = k.$$

故全部解为 $k(0, 0, 1, 0)^T$, 取 $\mathbf{x} = (0, 0, 1, 0)^T$, 再解 $(\mathbf{A} - 0\mathbf{I})\mathbf{x} = \mathbf{x}_3$, 得 $\mathbf{x} = (0, 1, 0, 1)^T + k(0, 0, 1, 0)^T$ 取 $\mathbf{x}_4 = (0, 1, 0, 1)^T$.

令

$$\mathbf{P} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

则有 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$.

(2) 因为 $\mathbf{A}$ 与 $\mathbf{J}$ 有相同的极小多项式、特征多项式. 而

$$\begin{aligned}
 J^2 (J - 2I)^2 &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\
 &= \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\
 &= \mathbf{0},
 \end{aligned}$$

故极小多项式为 $m_A(x) = x^2(x-2)^2$.

评析 (1) 问题(1)是按计算 Jordan 标准的程序给出的标准程序(模式),学生在试卷上也应有这些叙述.考试中看到有的学生缺少过程叙述.比如系数矩阵的初等行变换不写在答题纸上,而写在草稿纸上,而有的人又因是在草稿纸上运算,故计算马虎,因而导致结果错误.事实上,解线性方程组是线性代数最基本的,最简单的知识,没有人会承认自己不会做.可会做又解错(不能算真会),最终特征向量错、广义特征向量错,于是 P 也错,非常可惜可叹.希望读者加强基本功的训练,认真做方阵的初等行变换,认真求方程组的解.

(2) 关于求极小多项式,5.1 节中给出 3 种方法.因本题已求出 Jordan 标准形.故用方法 3 最省事.由于特征多项式为

$$f_A(\lambda) = \lambda^2(\lambda-2)^2,$$

而极小多项式为 $m_A(x) = (\lambda)^{m_1}(\lambda-2)^{m_2}$, 其中 m_1, m_2 分别是对应特征值 $\lambda=0$ 和 $\lambda=2$ 的最大 Jordan 块的阶数.又对应于特征值 0 和 2 的最大 Jordan 块都是 2 阶的,故

$$m_A(x) = x^2(x-2)^2.$$

问题(2)的解题过程中用了 $m_A(x) = m_J(x)$ 这一事实,而计算 J 的极小多项式要简单许多,3 种方法都可用.如用方法 2 有:记 $f_1(x) = x(x-2)$, $f_2(x) = x^2(x-2)$, $f_3(x) = x(x-2)^2$, 则因

有 $f_1(J) \neq 0$, $f_2(J) \neq 0$, $f_3(J) \neq 0$, 所以它们都不是 J 的化零多项式, 故

$$m_A(x) = m_J(x) = f_A(x) = x^2(x-2)^2.$$

(3) 当特征值求错时, 该值使特征多项式不为 0, 故齐次线性方程组 $(A - \lambda I)x = 0$ 只有零解. 这时, 应想到可能特征值出错了. 也就是说, 求特征向量时, 对特征值是个校验. 可个别学生居然求出特征向量为零向量, 而没有觉得有问题. 还把这个零向量写入可逆矩阵 P 的某列中. 这里有几个概念问题. ① 特征向量是非零向量; ② 可逆矩阵的某列不可能全为 0. 这些知识要清楚, 就不会出问题了.

例 3 设线性空间 V 上的线性变换 σ 在基 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ 下的矩阵是

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

则: 1) 计算 $\text{Ker}\sigma$;

(2) 试写出一个 σ 的一维不变子空间;

(3) 问由向量 $\varepsilon_1, \varepsilon_2$ 生成的子空间 $\langle \varepsilon_1, \varepsilon_2 \rangle$ 是不是 σ 的二维不变子空间;

(4) 求 σ 的包含向量 ε_4 的最小不变子空间.

解 (1) 任取 $\alpha \in \text{Ker}\sigma$, 设

$$\alpha = x_1 \varepsilon_1 + x_2 \varepsilon_2 + x_3 \varepsilon_3 + x_4 \varepsilon_4 = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)x,$$

则

$$\sigma\alpha = 0 \quad Ax = 0.$$

解齐次线性方程组 $Ax = 0$, 得全部解为 $x = c(1, 0, -2, 0)^T$, 故

$$\text{Ker}\sigma = \langle \varepsilon_1 - 2\varepsilon_3 \rangle.$$

(2) 由问题(1)的结论知 $\text{Ker}\sigma$ 是 σ 的属于特征值 0 的特征子空间, 因为任意线性变换的核都是它自己的不变子空间, 故 $\text{Ker}\sigma$

是 σ 的不变子空间,且由问题(1)的结论知 $\dim \text{Ker}\sigma = 1$,故 $\text{Ker}\sigma$ 为所求.另外,由题设矩阵 A 的元素可知, ε 也是特征向量.因有

$$\sigma\varepsilon = 2\varepsilon.$$

故 $\langle\varepsilon\rangle$ 和 $\langle\varepsilon - 2\varepsilon\rangle$ 都是一维不变子空间.

(3) 因 $\varepsilon \in \langle\varepsilon, \varepsilon\rangle, \varepsilon \in \langle\varepsilon, \varepsilon\rangle$,且

$$\sigma\varepsilon = 2\varepsilon \in \langle\varepsilon, \varepsilon\rangle, \text{而 } \sigma\varepsilon = \varepsilon - 2\varepsilon \notin \langle\varepsilon, \varepsilon\rangle,$$

故 $\langle\varepsilon, \varepsilon\rangle$ 不是 σ 的二维不变子空间.

(4) 由已知矩阵 A ,可知 $\langle\varepsilon, \varepsilon, \varepsilon\rangle$ 是含 ε 的一个三维不变子空间(因为 A 为准对角矩阵).下面看它是不是最小的.设所求最小不变子空间为 W ,则 $\varepsilon \in W$.依定义应有 $\sigma\varepsilon = 2\varepsilon \in W$,故 $\varepsilon \in W$.又 $\sigma^2\varepsilon = 2\sigma\varepsilon = 2\varepsilon - 4\varepsilon \in W$,故 $\varepsilon \in W$.所以, σ 的包含向量 ε 的最小不变子空间为

$$\langle\varepsilon, \varepsilon, \varepsilon\rangle.$$

讲评 这是一个考查学生多个基本概念的题目.涉及到线性变换的核,不变子空间等知识.

(1) 怎样求 $\text{Ker}\sigma$?重要的步骤是 $\alpha = 0 \quad Ax = 0$,故只要齐次线性方程组 $Ax = 0$ 求解准确. α 就可被定出,要强调的是 $Ax = 0$ 的解是 α 的坐标列,还不是 $\text{Ker}\sigma$ 中向量.

(2) 其他3个问题都是与不变子空间有关.而 W 是 σ 的不变子空间.任取 $\alpha \in W$,有 $\sigma\alpha \in W$.故紧紧围绕这一定义就可解决问题.有这样相关的几个结论可用在此处.

- ① $\text{Ker}\sigma$ 与 $\text{Im}\sigma$ 都是 σ 的不变子空间;
- ② 每个特征向量都生成 σ 的一维不变子空间;
- ③ 每个特征子空间 V_λ 都是 σ 的不变子空间;
- ④ 若 σ 在基 $\varepsilon_1, \dots, \varepsilon_r, \varepsilon_{r+1}, \dots, \varepsilon_n$ 下的矩阵为准对角矩阵

$B = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix}$, 则 $\langle\varepsilon_1, \dots, \varepsilon_r\rangle$ 和 $\langle\varepsilon_{r+1}, \dots, \varepsilon_n\rangle$ 都是 σ 的不变子空间.

⑤ 若 $\varepsilon, \sigma\varepsilon, \dots, \sigma^{k-1}\varepsilon$ 线性无关, 而 $\sigma^k\varepsilon$ 可由 $\varepsilon, \sigma\varepsilon, \dots, \sigma^{k-1}\varepsilon$ 线性表出, 则 $\langle \varepsilon, \sigma\varepsilon, \dots, \sigma^{k-1}\varepsilon \rangle$ 是 σ 的含向量 ε 的最小不变子空间. 这些都不用死记硬背, 用不变子空间的概念很容易一一证明之, 在深刻理解的基础上, 很容易记住, 做题时就可运用自如了. 在以上解答的问题(4)中给出的方法是用的④和不变子空间的定义(任意 $\alpha \in W$, 有 $\sigma\alpha \in W$); 下面用⑤给出另一个解答问题(4)的方法. 因为 $\sigma\varepsilon = 2\varepsilon$, 故 $\varepsilon, \sigma\varepsilon$ 线性无关; 又 $\sigma(\sigma\varepsilon) = 2(\varepsilon - 2\varepsilon) = 2\varepsilon - 4\varepsilon$, 不是 $\varepsilon, \sigma\varepsilon$ 的线性组合, 即

$$\varepsilon, \sigma\varepsilon, \sigma^2\varepsilon \text{ 线性无关.}$$

而 $\sigma^3(\varepsilon) = \sigma(2\varepsilon - 4\varepsilon) = 4\varepsilon - 4\varepsilon = 0$ 可由 $\varepsilon, \sigma\varepsilon, \sigma^2\varepsilon$ 线性表出, 故 σ 的含 ε 的最小不变子空间为

$$W = L(\varepsilon, \sigma\varepsilon, \sigma^2\varepsilon) = \langle \varepsilon, 2\varepsilon, 2\varepsilon - 4\varepsilon \rangle = \langle \varepsilon, \varepsilon, \varepsilon \rangle.$$

若只写 $L(\varepsilon, \sigma\varepsilon, \sigma^2\varepsilon)$ 虽也对, 应视为没有计算出结果. 有点欠妥.

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

例4 设 $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ 是线性空间 V 上的线性变换 σ 在基

$$\varepsilon, \sigma\varepsilon, \sigma^2\varepsilon$$

下的矩阵表示, 试求:

- (1) $\text{Im}\sigma$;
- (2) σ 的全部非平凡不变子空间;
- (3) σ 的包含向量 α 的最小不变子空间 W , 其中 $\alpha = \varepsilon + \sigma\varepsilon$;
- (4) σ 的极小多项式 $m_A(x)$.

解 (1) 因 A 的列为 $\sigma\varepsilon, \sigma^2\varepsilon, \sigma^3\varepsilon$ 的坐标, 故有

$$\text{Im}\sigma = L(\sigma\varepsilon, \sigma^2\varepsilon, \sigma^3\varepsilon) = L(2\varepsilon, \varepsilon + \sigma\varepsilon) = \langle \varepsilon, \sigma\varepsilon \rangle.$$

(2) 因题设矩阵 A 是一个 Jordan 标准形, 很容易看出其特征值为

$$\lambda = 0, \quad \lambda = 2(\text{二重}),$$

且 $\sigma\varepsilon = 0, \sigma^2\varepsilon = 2\varepsilon$, 故 $\langle \varepsilon \rangle, \langle \sigma\varepsilon \rangle$ 均为 σ 的一维不变子空间. 又 A

为准对角矩阵, 故 $\langle \varepsilon, \varepsilon \rangle$ 是 σ 的二维不变子空间 (事实上, 由 Jordan 链组成, 是循环子空间). 再者, 两个特征子空间的直和, $\langle \varepsilon, \varepsilon \rangle$ 也是 σ 的二维不变子空间 (把 A 看成准上三角矩阵, 左上角为二阶矩阵, 也得 $\langle \varepsilon, \varepsilon \rangle$ 为 σ 的不变子空间). 所以 σ 的所有不变子空间 (非平凡的) 为

$$\langle \varepsilon \rangle, \langle \varepsilon \rangle, \langle \varepsilon, \varepsilon \rangle, \langle \varepsilon, \varepsilon \rangle.$$

(3) 由 $\alpha \in W$ 知, $\alpha_1 = \sigma(\varepsilon + \varepsilon) = \alpha\varepsilon_1 + \alpha\varepsilon_2 = 2\varepsilon \in W$. 而 $\alpha = \varepsilon + \varepsilon, \alpha_1 = 2\varepsilon$ 线性无关, 又 $\sigma^2\alpha = \sigma(2\varepsilon) = 4\varepsilon$ 可由 α, α_1 线性表出, 故

$$W = \langle \alpha, \alpha_1 \rangle = \langle \varepsilon + \varepsilon, 2\varepsilon \rangle = \langle \varepsilon, \varepsilon \rangle.$$

(4) 因 A 的特征多项式为 $p_A(\lambda) = \lambda(\lambda - 2)^2$, 记 $f_1(x) = x(x - 2)$, 有 $f_1(A) = A(A - 2I) \neq 0$, 故

$$m_A(x) = x(x - 2)^2.$$

讲评 (1) 关于如何求 $\text{Im}\sigma$ (值域) 见 5.1 节定理 1. 由 $\text{Im}\sigma = \langle \sigma\varepsilon, \dots, \sigma\varepsilon_n \rangle$, 知问题转化为求 $\sigma\varepsilon_1, \dots, \sigma\varepsilon_n$ 的极大线性无关组, 其间要借助于 $R(A)$ (即 A 的列空间, 同求 $\text{Ker}\sigma$ 类似的一个问题是向量与坐标的关系, A 的列的极大线性无关组是 $\text{Im}\sigma$ 的基的坐标列! 本题因 A 很简单, 故计算量较小).

(2) 关于不变子空间仍要紧扣定义: 任取 $\alpha \in W$, 若 $\sigma\alpha \in W$, 则 W 是 σ 的不变子空间. 本题用到:

① 每个特征向量生成一个 σ 的一维不变子空间;

② $\text{Im}\sigma$ 是 σ 的不变子空间;

③ 每个 k 阶 Jordan 块提供一个 k 维不变子空间 (这里 $\langle \varepsilon \rangle$ 是一阶 Jordan 块提供, $\langle \varepsilon, \varepsilon \rangle$ 是二阶块提供的), 每个 Jordan 块中前 r 阶子矩阵提供一个 r 阶不变子空间 (这里 $\langle \varepsilon \rangle$ 是二阶 Jordan 块

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

的一阶子矩阵提供的, 事实上, 它对应的基恰为 σ 的特

征向量) .

(3) 关于极小多项式有 3 种求法 .因 A 有许多零元素,故用哪种方法做都不难,但有的学生写成 $m_A(x) = (x - 2)^2$, 即把 x 抹去,我们认为还应从什么是化零多项式,什么是极小多项式的概念入手,彻底弄清楚基本概念,做题时,就不会犯糊涂 .

例 5 设 n 阶矩阵 A 的特征多项式为

$$f_A(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \lambda_i \neq \lambda_j, \sum_{i=1}^s n_i = n .$$

$$J_{n_1}(\lambda_1)$$

$$J_{n_2}(\lambda_2)$$

A 的 Jordan 标准形为 $J =$

W

$$J_{n_s}(\lambda_s)$$

记极小多项式为

$$m_A(x) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s}, m_i \leq n_i .$$

试证: $m_i = J_{n_i}(\lambda_i)$ 中最大 Jordan 块的阶数 .

证 因为 $A \sim J$, 故 $m_A(x) = m_J(x)$.故只要看当 m_i 至少为何值时,使得

$$m_A(J) = (J - \lambda_1 I)^{m_1} (J - \lambda_2 I)^{m_2} \cdots (J - \lambda_s I)^{m_s} = 0 .$$

由于

$$m_A(J) = (J - \lambda_1 I)^{m_1} (J - \lambda_2 I)^{m_2} \cdots (J - \lambda_s I)^{m_s}$$

$$J_{n_1}(\lambda_1) - \lambda_1 I_{n_1} \quad m_1$$

$$J_{n_2}(\lambda_2) - \lambda_2 I_{n_2}$$

=

W

$$J_{n_s}(\lambda_s) - \lambda_s I_{n_s}$$

$$\begin{array}{ccc}
 J_1(\lambda) - \lambda I_1 & & m_2 \\
 J_2(\lambda) - \lambda I_2 & & \dots \\
 & W & \\
 J_s(\lambda) - \lambda I_s & & m_s \\
 J_1(\lambda) - \lambda I_1 & & \\
 J_2(\lambda) - \lambda I_2 & & , \\
 & W & \\
 J_s(\lambda) - \lambda I_s & &
 \end{array}$$

等号右边是 s 个矩阵相乘, 每个矩阵都是 (有 s 个对角块的) 准对角矩阵. 则有: ① 每个矩阵的幂 m_i 可以写到准对角矩阵的每个对角块上; ② 这 s 个准对角矩阵相乘结果仍为准对角矩阵, 记为 Q , 则 Q 的第 i 个对角元 (是个方阵) 等于前述 s 个方阵的第 i 个对角元之积, 即

$$Q_i = (J_{n_i}(\lambda) - \lambda I_{n_i})^{m_1} (J_{n_i}(\lambda) - \lambda I_{n_i})^{m_2} \cdots (J_{n_i}(\lambda) - \lambda I_{n_i})^{m_s},$$

其中有一个方阵的对角元为 0, 而其余的方阵对角元均非零, 从而都是可逆矩阵. 而

$$Q = \text{diag}\{Q_1, Q_2, \cdots, Q_s\}.$$

故要 $Q = 0$ $Q_i = 0 (i = 1, 2, \cdots, s)$

$$(J_{n_i}(\lambda) - \lambda I_{n_i})^{m_i} = 0 (i = 1, 2, \cdots, s).$$

又

$$(J_{n_i}(\lambda_i) - \lambda_i I_{n_i})^{m_i} =$$

01

W W

01

0

01

W W

01

0

01

W W

01

0

m_i

$$=$$

$N_{i1}^{m_i}$

$N_{i2}^{m_i}$

$N_{ik_i}^{m_i}$

W

其中 N_{ij} 是形如

$$N_{ij} =$$

01

W W

01

0 $\times l$

的幂零方阵,满足 $N_{ij}^l = 0$.故

$$(J_{n_i}(\lambda_i) - \lambda_i I_{n_i})^{m_i} = 0 \quad N_{ij}^{m_i} = 0 (j = 1, \cdots, k_i) .$$

从而 m_i 等于 $N_{i1}, N_{i2}, \cdots, N_{ik_i}$ 中阶数最大者 .当满足此条件时,有

$$m_A(J) =$$

0

$*$

0

$*$

$*$

$*$

$*$

$*$

W

W

$$\begin{pmatrix} * \\ & \ddots & \\ & & * \end{pmatrix} W \begin{pmatrix} * \\ & \ddots & \\ & & * \end{pmatrix} = \mathbf{0}.$$

$$\mathbf{0}$$

第 1 个等号右边的 s 个矩阵中每一个都有一块为 $\mathbf{0}$ (而其余块为可逆矩阵), 故乘积为零矩阵.

评析 本证法没有用定理:“准对角矩阵的最小多项式是准对角线上各矩阵的极小多项式的最小公倍式”这一现成结论, 而是用 Jordan 标准形给出一个具有几何直观性的理解和证明. 如果读者能深刻理解其本质, 那就不用抽象地背定理, 而是“用眼一看”, “用心一想”结论就出来了.

例 6 设 σ 是 V 上的线性变换, 证明 σ 把 V 中某线性无关的向量组变换成线性相关的向量组的充分必要条件是 σ 把某非零向量变换成零向量.

证 设有向量组 $\alpha, \cdots, \alpha$ 线性无关, 而 $\alpha_1, \cdots, \alpha_s$ 线性相关, 则必存在不全为 0 的数 $k_1, \cdots, k_s$, 使

$$k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_s \alpha_s = 0,$$

即有 $\sigma(k_1 \alpha + k_2 \alpha + \cdots + k_s \alpha) = 0$.

令 $\alpha = k_1 \alpha + k_2 \alpha + \cdots + k_s \alpha_s$, 则有 $\alpha \neq 0$ (因 $\alpha, \cdots, \alpha$ 线性无关), 且 $\sigma \alpha = 0$.

若有向量 $\beta \neq 0$, 使 $\sigma \beta = 0$. 在 V 中取基 $\varepsilon, \cdots, \varepsilon_n$. 设

$$\beta = x_1 \varepsilon + x_2 \varepsilon + \cdots + x_n \varepsilon_n \quad (x_i \text{ 不全为 } 0),$$

由 $0 = \sigma \beta = x_1 \sigma \varepsilon_1 + x_2 \sigma \varepsilon_2 + \cdots + x_n \sigma \varepsilon_n$, 且 x_i 不全为 0, 知 $\sigma \varepsilon_1, \sigma \varepsilon_2, \cdots, \sigma \varepsilon_n$ 线性相关.

评析 众所周知, 线性变换把线性相关的向量组映为线性相关的向量组; 也可能把线性无关的向量组变换为线性相关的向量组. 本题恰是证明当且仅当 $\text{Ker} \sigma \neq \{0\}$ 时, σ 会把线性无关的向量

组变换为线性相关的向量组. 换句话说, 当 $\text{Ker}\sigma = \{0\}$ 时, σ 是单射也是满射, 即是双射. 此时, 任意线性无关的向量组均被映为线性无关的向量组.

例7 设 $A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & -2 & 3 & 0 \\ 4 & 0 & 2 & 3 \end{pmatrix}$, 求可逆矩阵 P 及 Jordan 标准

形 J , 使 $P^{-1}AP = J$.

解 因

$$p_A(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 3 & -1 & 0 & 0 \\ -2 & \lambda - 3 & -1 & 0 \\ 0 & 2 & \lambda - 3 & 0 \\ -4 & 0 & -2 & \lambda - 3 \end{vmatrix} = (\lambda - 3)^4 = 0,$$

故 $\lambda = 3$ 为四重根, 所以 A 的 Jordan 标准形 J 中只有一大块. 对 $\lambda = 3$, 解齐次线性方程组 $(A - 3I)x = 0$. 因

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 0 \\ 4 & 0 & 2 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

得 $V_{\lambda=3} = \alpha_1 x^{(1)} + \alpha_2 x^{(2)} = \alpha_1 (1, 0, -2, 0)^T + \alpha_2 (0, 0, 0, 1)^T$.

由 $\dim V_{\lambda=3} = 2$, 故 J 有两个 Jordan 块. 又因

$$r(A - 3I) = 2, \quad r(A - 3I)^2 = 1,$$

知 $r(J - 3I) = 2, r(J - 3I)^2 = 1$, 所以 J 中有一个三阶块. 故

$$J = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{且 } P^{-1}AP = J.$$

3

解 $(A - 3I)x = x^{(1)}$ (注意, 此处 $x^{(1)} \in V_{\lambda=3}$, 且使非齐次方程有

解).由

$$\begin{array}{ccccc} 0 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 & -2 \\ 4 & 0 & 2 & 0 & 0 \end{array} \rightarrow \begin{array}{ccccc} 0 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array},$$

得特解为 $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, 故全部解为

$$\mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

为使 $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{x}^{(2)}$ 有解, 需取

$\mathbf{x}^{(2)} = (0, 1, 0, 2)^T$, 即在通解中取 $c_1 = 0, c_2 = 2$.

解 $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{x}^{(2)}$, 得 $\mathbf{x}^{(3)} = \frac{1}{2}, 0, 0, 0^T$.

故令

$$\mathbf{P} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \mathbf{x}^{(1)}) = \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix},$$

则有 $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{J}$.

评析 (1) 当写出 Jordan 标准形

$$\mathbf{J} = \begin{pmatrix} 3 & 1 \\ & 3 & 1 \\ & & 3 \end{pmatrix}$$

时,就已经决定了可逆矩阵 $P = (p_1, p_2, p_3, p_4)$ 中 p_1, p_4 为特征向量,且 p_1, p_2, p_3 为一个 Jordan 链 即 p_1, p_4 满足 $(A - 3I)x = 0$, p_2, p_3 满足: $(A - 3I)p_2 = p_1, (A - 3I)p_3 = p_2$. 2 节例 4 证明了这样算出的 p_1, p_2, p_3 一定线性无关.它们之间的关系又可以写为

$$(A - 3I)^2 p_3, (A - 3I)p_3, p_3,$$

故 p_3 是子空间 $\langle p_1, p_2, p_3 \rangle$ 中的循环向量,它生成了循环子空间 $\langle p_1, p_2, p_3 \rangle$.

(2) 当一个 Jordan 大块中有多个 Jordan 块时,需求多个 Jordan 链.第一步求特征向量时,因是解齐次线性方程组,故选择基础解系时,一般找比较简单的向量,而求 Jordan 链时,要求选择的特征向量使得非齐次线性方程组 $(A - \lambda I)x = x_i$ 有解,故应先把线性方程组的基础解系求好、写对.再判断基础解系中的向量哪个满足要求.如本题, $x^{(1)}$ 满足有解的要求.若基础解系中的向量均不满足有解的要求,则需分析、讨论,查看它们什么样的线性组合,使得非齐次线性方程组有解(即要在特征子空间 V_λ 中另选向量做 Jordan 链的第一个向量).在求 $x^{(2)}$ 中也是这样.特解要找最简单的,再分析,每个解的第 2 分量为 1,要使下一步有解,第 4 个分量必为 2,故依此要求去选取 α, β . p_3 的选取也是不惟一的,因不用再解下一个非齐次方程组,故对它没有更多的要求(只要它是解就可以),以简单为好.

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

例 8 设 $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, 试求它的极小多项式.

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

解 方法 1 因 I, A 线性无关 ($A \neq kI$), 又

$$A^2 = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix} \neq k_1 I + k_2 A,$$

而 A 是三阶方阵, 它的特征多项式是三次多项式, 故

$$m_A(x) = p_A(x) = |xI - A| = x(x-2)^2.$$

方法2 因

$$p_A(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 1 & & -1 \\ & \lambda - 1 & -1 \\ & & \lambda - 2 \end{vmatrix} = \lambda(\lambda - 2)^2.$$

记 $f_1(\lambda) = \lambda(\lambda - 2)$, 而

$$f_1(A) = A(A - 2I) = \begin{pmatrix} 1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \neq \mathbf{0},$$

故 $f_1(\lambda)$ 不是化零多项式, 所以特征多项式是次数最低的化零多项式, 即

$$m_A(x) = p_A(x) = x(x-2)^2.$$

方法3 因 $p_A(\lambda) = |\lambda I - A| = \lambda(\lambda - 2)^2$, 故 A 的 Jordan 标准形由两大块构成 (有两个不同特征值). 而 $\lambda = 0$ 是单根, 故 $J_{n_1}(\lambda)$ 是一阶块, 而 $\lambda = 2$ 是重根, 故需确定 $J_{n_2}(\lambda)$ 是两个一阶块还是一个二阶块.

因

$$r(A - 2I) = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 2,$$

故 $r(J - 2I) = 2$, 即

$$J - 2I = \begin{pmatrix} -2 & & 0 \\ 0 & 1 & \\ 0 & & 2 \end{pmatrix}, \text{ 从而得 } J = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}.$$

又因 $m_A(x) = m_J(x)$, 故

$$m_A(x) = x(x-2)^2.$$

评析 (1) 以上用求极小多项式的3种方法分别给出了解答, 希望通过此例的实践, 加深对3种方法本质的理解.

(2) 考场上不知为何有一些同学把 x 去掉, 写成

$$m_A(x) = (x - 2)^2 .$$

一定是死记硬背了什么, 又记错了. 特征根 $\lambda = 0$, 与其他非零特征值地位是相同的. 事实上, 稍微动一下脑筋, 看一下

$$(A - 2I)^2 = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 2 & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \neq \mathbf{0},$$

即只算一个数, 就知道 $(x - 2)^2$ 不是 A 的化零多项式. 用 Jordan 标准形算更容易看出

$$(J - 2I)^2 = \begin{pmatrix} -2 & & \\ & 0 & 1 \\ & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 4 & & \\ & 0 & 0 \\ & & 0 \end{pmatrix} \neq \mathbf{0} .$$

第 6 章 欧几里得空间

6.1 理论脉络

这里所说的欧几里得(Euclid)空间,不包含读者已在课程“几何与代数(1)”中学过的欧几里得空间的内容(简称欧几里得空间(1)).欧几里得空间(1)的内容概述如下.

内积的定义: $V_n(\mathbb{R})$ 中内积是一个对称、正定、双线性函数.

$$\begin{aligned}\Phi: V(\mathbb{R}) \times V(\mathbb{R}) &\rightarrow \mathbb{R}, \\ \alpha, \beta &\mapsto (\alpha, \beta),\end{aligned}$$

且满足:

- (1) $(\alpha, \beta) = (\beta, \alpha)$;
- (2) $(k\alpha, \beta) = k(\alpha, \beta)$; (2') $(\alpha, k\beta) = k(\alpha, \beta)$;
- (3) $(\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$; (3') $(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$;
- (4) $(\alpha, \alpha) \geq 0$, 且 $(\alpha, \alpha) = 0 \iff \alpha = 0$.

这 4 条性质恰是三维几何空间定义的内积

$$\alpha \cdot \beta = \|\alpha\| \|\beta\| \cos \phi$$

所具有的代数性质.故可以说,是从代数性质上把内积的概念推广到 n 维空间.有了内积后,又可以把“向量的长度”,“两个向量的夹角”,“向量的正交”等度量概念推广到 n 维空间,即定义

$$\|\alpha\| = \sqrt{(\alpha, \alpha)}; \quad \phi = \arccos \frac{(\alpha, \beta)}{\|\alpha\| \|\beta\|}, \quad 0 \leq \phi \leq \pi;$$

$$\alpha \perp \beta \stackrel{\text{def}}{=} (\alpha, \beta) = 0.$$

另外,其余内容中有 5 个问题值得强调一下:

- (1) Cauchy-Schwarz 不等式: $|(\alpha, \beta)| \leq \|\alpha\| \|\beta\|$.

它在两个特殊的线性空间上的结果非常有用

$$|a_1 b_1 + \cdots + a_n b_n| \leq \sqrt{a_1^2 + \cdots + a_n^2} \sqrt{b_1^2 + \cdots + b_n^2},$$

$$\left| \int_a^b f(x) g(x) dx \right| \leq \sqrt{\int_a^b f^2(x) dx} \sqrt{\int_a^b g^2(x) dx}.$$

(2) 标准正交基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的概念, 它们两两正交, 且 $\|\varepsilon_i\| = 1 (i=1, \cdots, n)$.

(3) 施密特正交化程序. 设 $\alpha_1, \cdots, \alpha_n$ 线性无关, 令

$$\beta_1 = \alpha_1, \quad e_1 = \frac{\beta_1}{\|\beta_1\|} = \frac{\alpha_1}{\|\alpha_1\|},$$

$$\beta_2 = \alpha_2 - (\alpha_2, e_1)e_1, \quad e_2 = \frac{\beta_2}{\|\beta_2\|},$$

...

$$\beta_n = \alpha_n - (\alpha_n, e_1)e_1 - (\alpha_n, e_2)e_2 \cdots - (\alpha_n, e_{n-1})e_{n-1},$$

$$e_n = \frac{\beta_n}{\|\beta_n\|}.$$

(4) 可逆矩阵的 QR 分解.

从上面的施密特正交化程序中反解出 $\alpha_1, \cdots, \alpha_n$ 得

$$\alpha_1 = \beta_1 = \|\alpha_1\| e_1,$$

$$\alpha_2 = (\alpha_2, e_1)e_1 + \|\beta_2\| e_2,$$

...

$$\alpha_n = (\alpha_n, e_1)e_1 + \cdots + (\alpha_n, e_{n-1})e_{n-1} + \|\beta_n\| e_n.$$

写成矩阵乘法就是

$$(\alpha_1, \cdots, \alpha_n) = (e_1, \cdots, e_n) \begin{pmatrix} \|\beta_1\| & * & \cdots & * \\ & W & W & \cdots \\ & & W & * \end{pmatrix}.$$

当 $\alpha_1, \cdots, \alpha_n \in F^n$ 时, 上面等式的左面 $(\alpha_1, \cdots, \alpha_n)$ 是可逆矩阵, 而右面

$$\begin{pmatrix} \parallel \beta \parallel & * & \dots & * \\ & W & W & \dots \\ & & W & * \end{pmatrix}$$
 是上三角矩阵.

$(e_1, \dots, e_n)$ 是正交矩阵,

$$\parallel \beta \parallel$$

即把可逆矩阵 P 的列看成 F^n 中的一组基, 对它做施密特正交化, 得到的标准正交基构成正交矩阵 Q , 并如上述公式所示: 用 Q 的列表出 P 的列, 则那些表出系数排成的矩阵就是上三角矩阵.

(5) 内积的度量矩阵 $A = (\varepsilon_i, \varepsilon_j)_{n \times n}$ 由基向量的内积排成.

首先, 度量矩阵是实对称矩阵; 其次它是正定的; 又内积在不同基下的度量矩阵是相合的.

欧几里得空间(2)从《理工科代数基础》的第5章第2.3节开始. 主要有以下3个学习的重点.

1 正交补与直和分解

定义1 设 V 是欧几里得空间, $W \subseteq V, \alpha \in V$, 若对 $\beta \in W$, 都有 $(\alpha, \beta) = 0$, 则称向量 α 与子空间 W 正交, 记为 $\alpha \perp W$.

定义2 设 $W_1, W_2 \subseteq V$ (欧几里得空间), 若对 $\alpha \in W_1, \beta \in W_2$, 都有 $\alpha \perp \beta$, 则称两个子空间 W_1 与 W_2 正交, 记为 $W_1 \perp W_2$.

定义3 设 $W \subseteq V$ (欧几里得空间), $\alpha \in V, \beta \in W$, 若 $\alpha - \beta \perp W$, 则称 β 是 α 在 W 上的投影, 记为 $(\alpha)_W = \beta$.

定理1 设 W 是欧几里得空间 $E_n(F)$ 的子空间, 任取 W 的标准正交基 $\alpha_1, \dots, \alpha_m$, 则 $E_n(F)$ 中任意向量 α 在子空间 W 上的投影向量惟一, 且

$$\beta = (\alpha)_W = \sum_{i=1}^m (\alpha, \alpha_i) \alpha_i.$$

定义4 设 W 是欧几里得空间 $E_n(F)$ 的子空间, 则集合

$$W^\perp = \{\alpha \mid \alpha \in E_n(F), \alpha \perp W\}$$

称为子空间 W 的正交补.

定理2 设 W 是欧几里得空间 $E_n(F)$ 的子空间, 则有空间直

和分解

$$E_n(F) = W \oplus W^\perp.$$

关于求 $W^\perp$ (即求 $W^\perp$ 的维数和一组基)的方法:

设 V 是 n 维欧几里得线性空间, W 是 V 的 m 维子空间, 则 $\dim W^\perp = n - m$. 在 W 中任选一组标准正交基 $\varepsilon_1, \dots, \varepsilon_m$, 把它扩充为 V 的标准正交基

$$\varepsilon_1, \dots, \varepsilon_m, \varepsilon_{m+1}, \dots, \varepsilon_n$$

则 $\varepsilon_{m+1}, \dots, \varepsilon_n$ 是 $W^\perp$ 的一组基.

2 正交变换和正交矩阵

定义 5 设 σ 是 n 维欧几里得空间 V 中的一个线性变换, 若对任 $\alpha, \beta \in V$, 都有

$$(\sigma\alpha, \sigma\beta) = (\alpha, \beta),$$

则称 σ 是欧几里得空间 V 的一个正交变换.

定理 3 设 σ 是欧几里得空间 $E_n(\mathbb{R})$ 上的线性变换, 则以下命题等价:

- (1) σ 是正交变换;
- (2) " $\alpha \in E_n(\mathbb{R})$, 有 $\|\sigma\alpha\| = \|\alpha\|$;
- (3) σ 把一组标准正交基变为另一组标准正交基;
- (4) σ 在任意一组标准正交基下的矩阵是正交矩阵.

设正交变换 σ 在某组标准正交基下的矩阵为 A . 由 $A^T A = I$, 得 $\det A = \pm 1$. 当 $\det A = 1$ 时, 称 σ 为第一类正交变换; 当 $\det A = -1$ 时称 σ 为第二类正交变换.

定理 4 已知 σ 是 n 维欧几里得空间 V 的正交变换, W 是 σ 的不变子空间, 则 W 的正交补 $W^\perp$ 也是 σ 的不变子空间.

定义 6 设 $V = W_1 \oplus W_2$, 则 " $\alpha \in V$, 有惟一分解式 $\alpha = \alpha_1 + \alpha_2$, 其中 $\alpha_1 \in W_1, \alpha_2 \in W_2$. 若 $\sigma\alpha = \alpha_1$, 则称 σ 是 V 到 W_1 的投影变换; 若 $\sigma\alpha = \alpha_2$, 则称 σ 是 V 到 W_2 的投影变换.

定义 7 设 $W \subset V, V = W \oplus W^\perp$, 任取 $\alpha \in V$, 有 $\alpha = \beta + \gamma$, 其中

$\beta \in W, \gamma \in W^\perp$, 若 $\sigma\alpha = \beta$, 则称 σ 为 V 到 W 上的正交投影(变换)。

定理 5 设 σ 是欧几里得空间 V 上的线性变换, 则

σ 是 V 到 σV 上的正交投影 $\sigma V \perp (\varepsilon - \sigma)V$ 。

推论 1 正交投影变换都是幂等变换, 即满足 $\sigma^2 = \sigma$ 。

3 对称变换和对称矩阵

定义 8 设 σ 是欧几里得空间 $E_n(\mathbb{R})$ 中的线性变换, 如对 $\alpha, \beta \in E_n(\mathbb{R})$ 有

$$(\sigma\alpha, \beta) = (\alpha, \sigma\beta),$$

则称 σ 是一个对称变换。

定理 6 欧几里得空间 $E_n(\mathbb{R})$ 中线性变换 σ 是对称变换的充分必要条件是 σ 在标准正交基下的矩阵 A 是实对称矩阵。

定理 7 对称变换 σ 的属于不同特征值的特征向量必相互正交。

定理 8 设 σ 是欧几里得空间 $E_n(\mathbb{R})$ 上的对称变换, 有 $\sigma\xi = \lambda\xi$, 令

$$W = \{\alpha \mid \alpha \in V, (\alpha, \xi) = 0\},$$

则 W 是 σ 的不变子空间, 且 $\dim W = n - 1$ 。

定理 9 设 σ 是 $E_n(\mathbb{R})$ 上的对称变换, 则在 $E_n(\mathbb{R})$ 中存在标准正交基, 使 σ 在这组基下的矩阵是对角矩阵。

推论 2 设 A 是 n 阶实对称矩阵, 则存在 n 阶正交矩阵 T , 使 $T^{-1}AT = T^T AT = D$ 为对角矩阵。

* 关于最小二乘法与广义逆

(1) 一个变量的最小二乘法: 考虑线性方程 $ax = b$ 的解, 其中

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

当它无解时, 在 $R(\mathbf{a})$ 中找一个向量 $\mathbf{p}$ 使 $\|\mathbf{b} - \mathbf{p}\|$ 最小. 因 $\mathbf{p}$ 可以

写为 a 的倍, 设 $p = \overline{ax}$, 则把 $\overline{x}$ 作为 $ax = b$ 的近似解, 有

$$\overline{x} = \frac{a^T b}{a^T a}, \quad p = \overline{ax} = \frac{a^T b}{a^T a} a \text{ (称为 } b \text{ 在 } a \text{ 上的投影)}.$$

(2) 多变元的最小二乘法: 设 A 是 $m \times n$ 矩阵, 记 $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$, 因

$$R(A) = L(\alpha_1, \alpha_2, \dots, \alpha_n) = \{Ax \mid x \in \mathbb{R}^n\},$$

故当 $Ax = b$ 无解时, 有 $b \notin R(A)$.

记 b 在 $R(A)$ 的投影 $(b)_{R(A)} = p = A\overline{x} \in R(A)$, 这时 $(b - p) \perp R(A)$.

" $\alpha \in R(A)$, 设 $\alpha = Ay$, 有

$$0 = (\alpha, b - p) = (\alpha, b - A\overline{x}) = (Ay, b - A\overline{x}) = y^T A^T (b - A\overline{x}).$$

由于 y 是任意的, 故有 $A^T (b - A\overline{x}) = 0$, 即 $A^T Ax = A^T b$. 把 $A^T Ax = A^T b$ 称为正则方程, 它的解是 $Ax = b$ 的最小二乘解. 当 $r(A) = n$ (即 A 列满秩) 时, 有 $r(A^T A) = r(A) = n$, 则 $A^T A$ 可逆. 于是正则方程 $A^T A\overline{x} = A^T b$ 有惟一解

$$\overline{x} = (A^T A)^{-1} (A^T b),$$

它是方程 $Ax = b$ 的惟一最小二乘解. 而向量 b 在 $R(A)$ 的投影是

$$p = A\overline{x} = A(A^T A)^{-1} (A^T b).$$

当 A 不是列满秩, 正则方程 $A^T Ax = A^T b$, 即 $A\overline{x} = p$ 的解不惟一, 把这些解中长度最短的称为 $Ax = b$ 的最优最小二乘解.

定义 9 矩阵 A 的 Moore-Penrose 广义逆 A^+ 是指同时满足下列等式的矩阵 X :

$$AXA = A; \quad XAX = X; \quad (\overline{AX})^T = AX; \quad (\overline{XA})^T = XA.$$

即 A, X 互为广义逆. 在实数域中 AX, XA 为对称矩阵.

定理 10 任意矩阵 A 的 Moore-Penrose 广义逆存在且惟一. 事实上, 若 $A = LU$ (其中 L, U 为列, 行独立阵) 则

$$A^+ = \overline{U}^T (U \overline{U}^T)^{-1} (\overline{L}^T L)^{-1} \overline{L}^T.$$

关于 $A = LU$ 的证明非常简单. 因由相抵标准形有

$$A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = LU,$$

其中 $L = P \begin{pmatrix} I_r \\ 0 \end{pmatrix}$ 是列满秩矩阵, $U = (I_r \ 0) Q$ 是行满秩矩阵.

定理 11 对于线性方程组 $Ax = b$, 当它无解且 A 不是列满秩矩阵时, 有

$$\bar{x} = A^+ b = \bar{U}^T (U \bar{U}^T)^{-1} (\bar{L}^T L)^{-1} \bar{L}^T b$$

是它的最优最小二乘解. 此时

$$p = A \bar{x} = L U \bar{U}^T (U \bar{U}^T)^{-1} (\bar{L}^T L)^{-1} \bar{L}^T b = L (\bar{L}^T L)^{-1} \bar{L}^T b$$

是 b 在列空间 $R(A)$ 上的投影.

从几何上看齐次线性方程组 $Ax = 0$, 知解向量与 A 的每一行向量都正交, 故 $\text{null}(A) \perp R(A^T)$, 即 A 的零空间 $\perp A^T$ 的列空间. 设 $A_{m \times n}$, 则有 $\dim N(A) + \dim R(A^T) = n = \dim R^n$.

6.2 典型题讲评

例 1 设内积在 R^3 中的基 α, α, α 下的度量为

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}.$$

试把 α, α, α 施密特正交化得出 R^3 的标准正交基.

解 方法 1 因 $(\alpha, \alpha_j) = a_{ij}$, 故令

$$\beta = \alpha, \quad e = \frac{\beta}{\|\beta\|} = \frac{\alpha}{\|\alpha\|} = \alpha;$$

$$\beta = \alpha - (\alpha, e)e = \alpha - \alpha, \quad e = \frac{\beta}{\|\beta\|} = \alpha - \alpha$$

$$\beta = \alpha - (\alpha, e)e - (\alpha, e)e = \alpha + \alpha - \alpha,$$

$$(\beta, \beta) = (\alpha + \alpha - \alpha, \alpha + \alpha - \alpha) = 2, \text{ 故 } e = \frac{2}{2}(\alpha + \alpha - \alpha).$$

方法 2 设所求标准正交基为 $\varepsilon, \varepsilon, \varepsilon$, 则内积在此基下的度

量矩阵为单位矩阵 I . 设从基 α, α, α 到基 $\varepsilon, \varepsilon, \varepsilon$ 的过渡矩阵为 P , 则有

$$P^T A P = I \text{ (内积在不同基下的度量矩阵相合) }.$$

用初等变换法求 P 如下:

$$\begin{array}{l}
 (A, I) = \begin{array}{cccccc|cccccc}
 1 & 1 & 0 & 1 & 0 & 0 & & 1 & 0 & 0 & 1 & 0 & 0 \\
 1 & 2 & -1 & 0 & 1 & 0 & -r_1 + r_2 & 0 & 1 & -1 & -1 & 1 & 0 \\
 0 & -1 & 3 & 0 & 0 & 1 & -c_1 + c_2 & 0 & -1 & 3 & 0 & 0 & 1 \\
 \\
 & 1 & 0 & 0 & 1 & 0 & 0 & \frac{1}{2}r_3 & 1 & 0 & 0 & 1 & 0 & 0 \\
 r_2 + r_3 & 0 & 1 & 0 & -1 & 1 & 0 & & 0 & 1 & 0 & -1 & 1 & 0 \\
 c_2 + c_3 & 0 & 0 & 2 & -1 & 1 & 1 & \frac{1}{2}c_3 & 0 & 0 & 1 & -\frac{2}{2} & \frac{2}{2} & \frac{2}{2}
 \end{array}, \\
 \end{array}$$

故

$$P = \begin{pmatrix} 1 & -1 & -\frac{2}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad (\varepsilon, \varepsilon, \varepsilon) = (\alpha, \alpha, \alpha) P,$$

所以, $\varepsilon = \alpha, \varepsilon = \alpha - \alpha, \varepsilon = \frac{2}{2}(\alpha + \alpha - \alpha)$.

讲评 (1) 简单地说, 本题就是要从(线性无关的)基出发经施密特正交化程序构成“标准正交”基. 虽然向量 α, α, α 未知, 但告诉了 $(\alpha, \alpha)(i, j = 1, 2, 3)$, 而正交化程序中仅需要这些量. 故可以构造出用 α, α, α 表出的标准正交基 e, e, e . 不能用特殊的 α (有的同学自己设 $\alpha = (1, 0, 0)^T, \alpha = (1, 1, 0, 0)^T$ 等等, 不知道从 A 中获取需要的信息.

(2) 关于内积在不同基下的度量矩阵是相合的可以用形式写法证明如下: 设从 α, α, α 到标准正交基的过渡矩阵为 P , 则有

$$(\varepsilon, \varepsilon, \varepsilon) = (\alpha, \alpha, \alpha)P. \quad (*)$$

因 $\varepsilon, \alpha \in \mathbb{R}^3 (i=1, 2, 3)$, 故上述适合于方阵的各种运算. 其两边转置得

$$(\varepsilon, \varepsilon, \varepsilon)^T = P^T (\alpha, \alpha, \alpha)^T.$$

此式两边左乘 $(*)$ 式, 有

$$(\varepsilon, \varepsilon, \varepsilon)^T (\varepsilon, \varepsilon, \varepsilon) = P^T (\alpha, \alpha, \alpha)^T (\alpha, \alpha, \alpha)P$$

左边为 I . 因 $\varepsilon, \varepsilon, \varepsilon$ 为标准正交基, 故 $(\varepsilon_i, \varepsilon_j) = \delta_{ij}$; 右边的 $(\alpha, \alpha, \alpha)^T (\alpha, \alpha, \alpha) = A$ 故得

$$I = P^T A P.$$

例 2 设 $V_n(\mathbb{R})$ 是 n 维欧几里得空间, $\alpha \neq 0 \in V$. 证明

$$W = \{\beta \mid \beta \in V, (\beta, \alpha) = 0\}$$

是 V 的子空间, 且维数为 $n-1$.

证 先证 W 是子空间. 任取 $\beta, \beta \in W, k \in \mathbb{R}$, 有

$$(\beta, \alpha) = 0, (\beta, \alpha) = 0.$$

于是得

$$(\beta + \beta, \alpha) = (\beta, \alpha) + (\beta, \alpha) = 0 + 0 = 0,$$

$$(k\beta, \alpha) = k(\beta, \alpha) = 0,$$

故 $\beta + \beta \in W, k\beta \in W$, 所以 W 是子空间.

又因 $\alpha \neq 0$, 令 $\varepsilon = \frac{\alpha}{\|\alpha\|}$, 则 ε 为单位向量. 把 ε 扩充为 V 的标准正交基 $\varepsilon_1, \varepsilon, \dots, \varepsilon_n$, 则有

$$V = \langle \varepsilon \rangle \oplus \langle \varepsilon, \dots, \varepsilon_n \rangle,$$

且因 $(\varepsilon, \varepsilon_j) = 0 (j=2, \dots, n)$, 故 $\varepsilon_j \in W (j=2, \dots, n)$, 任取 $\beta \in \langle \varepsilon, \dots, \varepsilon_n \rangle$, 设 $\beta = k_2 \varepsilon + \dots + k_n \varepsilon_n$, 有

$$(\beta, \varepsilon) = \sum_{j=2}^n k_j (\varepsilon, \varepsilon_j) = 0,$$

故 $\beta \in W$. 所以 $\langle \varepsilon, \dots, \varepsilon_n \rangle \subset W$. 反之, 任取 $\gamma \in W$, 记 $\gamma = x_1 \varepsilon + \dots + x_n \varepsilon_n$, 其中 $x_i = (\gamma, \varepsilon_i) (i=1, \dots, n)$, 所以当 γ 与 ε 正交时, 有 $x_1 =$

$(\gamma, \varepsilon) = 0$, 即

$$\gamma = x_2 \varepsilon + \cdots + x_n \varepsilon_n \in \langle \varepsilon, \cdots, \varepsilon_n \rangle,$$

故有

$$W = \langle \varepsilon, \cdots, \varepsilon_n \rangle.$$

所以 $W = \langle \varepsilon, \cdots, \varepsilon_n \rangle = \langle \varepsilon \rangle^\perp$, 故 $\dim W = n - 1$.

讲评 (1) 从 α 单位化为 ε 并扩充为 V 的标准正交基, 可先扩充为线性无关的基(这只要它们的坐标列构成的矩阵行列式不为 0), 再经施密特正交化程序得标准正交基.

(2) 扩充的标准正交向量的生成空间 $\langle \varepsilon, \cdots, \varepsilon_n \rangle$ 实际上就是 $\langle \varepsilon \rangle$ 的正交补, 若利用定理 2, 那么已得 $W = \langle \varepsilon, \cdots, \varepsilon_n \rangle$. 但若把它们看成等价命题的话(相当于定理 2 中 $\dim W = 1$), 则必须直接证明

$$\langle \varepsilon, \cdots, \varepsilon_n \rangle = W,$$

即证: ① $\langle \varepsilon, \cdots, \varepsilon_n \rangle$ 中每个向量都与 ε 正交(都属于 W);

② 任意与 ε 正交的向量($\in W$)都包含于 $\langle \varepsilon, \cdots, \varepsilon_n \rangle$.

例 3 已知向量 $\alpha = (1, 1, 0, 0)^T$, $\alpha = (0, 1, 1, 0)^T$, 并记 $W = L(\alpha, \alpha)$, 试在标准内积下, 求 W 在 $\mathbb{R}^4$ 中的正交补 $W^\perp$ 的一组标准正交基.

解 方法 1 设与 α, α 均正交的向量为 $x = (x_1, x_2, x_3, x_4)^T$, 即有

$(\alpha, x) = 0, (\alpha, x) = 0$, 用坐标写出就是

$$x_1 + x_2 = 0,$$

$$x_2 + x_3 = 0.$$

$$x_1 = k_1,$$

$$x_2 = -k_1,$$

$$x_3 = k_1,$$

$$x_4 = k_2.$$

$$k_1, k_2 \in \mathbb{R}.$$

从而得

故全部解为 $x = k_1 (1, -1, 1, 0)^T + k_2 (0, 0, 0, 1)^T$, 所以 $W^\perp$ 的标准正交基为 $\frac{1}{3}(1, -1, 1, 0)^T, (0, 0, 0, 1)^T$.

方法2 把 α, α 扩充为 R^4 的基,再施密特正交化得标准正交基.因 α, α 的前两个分量组成的二阶子式不为0,故扩充 α, α 使它们的后两个分量组成的子行列式不为0.故令

$$P = (\alpha, \alpha, \alpha, \alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

则 P 可逆(因 $|P| = |I_2| \cdot |I_2| = 1$).把 $\alpha, \alpha, \alpha, \alpha$ 施密特正交化,得

$$\beta = \alpha, \quad e = \frac{\beta}{\|\beta\|} = \frac{1}{2}\alpha = \frac{1}{2}(1, 1, 0, 0)^T;$$

$$\beta = \alpha - (\alpha, e)e = \alpha - \frac{1}{2}\alpha,$$

$$e = \frac{\beta}{\|\beta\|} = \frac{1}{6}(-1, 1, 2, 0)^T;$$

$$\beta = \alpha - (\alpha, e)e - (\alpha, e)e = \alpha - \frac{1}{3}\beta$$

$$= \frac{1}{3}(1, -1, 1, 0)^T,$$

$$e = \frac{1}{3}(1, -1, 1, 0)^T;$$

显然 $e = \alpha$.故

$$W^\perp = \langle e, e \rangle = \langle (1, -1, 1, 0)^T, (0, 0, 0, 1)^T \rangle.$$

讲评 (1) 方法1可以有两种理解: ① 利用齐次线性方程组的几何意义: A 的零空间 $\perp A^T$ 的列空间 ($N(A) \perp R(A^T)$).把已知向量 α, α 看成 A^T 的列,则 Ax 的解空间(即 $N(A) = R(A^T)^\perp$)就是 $W^\perp$. ② 利用待定系数法,设任意 $x \in W^\perp$,则应有 $(\alpha, x) = 0$, $(\alpha, x) = 0$.解法是按此思路写出的.

(2) 把 α, α 扩充为 R^4 的基,扩充的两个向量是不惟一的,比

如若看 α, α 的第 2, 3 分量的子式也不为 0, 所以可以选择 α, α

的第 1, 4 分量组成二阶子式 $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$, 即令

$$P = (\alpha, \alpha, \alpha, \alpha) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

此时, e, e 不变, 且仍有 $e = \alpha - \frac{1}{2}\beta + \frac{1}{6}(-1, 1, 2, 0)^T =$

$$\frac{1}{3}(1, -1, 1, 0)^T = \beta. \quad e = \alpha = \alpha.$$

(3) 若选 $\alpha = (1, 0, 0, -\frac{1}{3})^T, \alpha = (0, 0, 0, 1)^T$, 则它们的第

1, 4 分量构成的二阶子式 $\begin{vmatrix} 1 & 0 \\ -\frac{1}{3} & 1 \end{vmatrix} = 1 \neq 0$, 从而得标准正交基

e, e 同前. 而

$$e = \frac{1}{2}(1, -1, 1, -1)^T, \quad e = \frac{1}{2 \cdot 3}(1, -1, 1, 3)^T,$$

即从 e, e 扩充的标准正交基可以不同(不惟一)有 e, e 或 e, e 等等. 但

$$\langle e, e \rangle = \langle e, e \rangle = W^\perp.$$

即无论扩充的基向量怎样选取, 总有 $W^\perp$ 是惟一的(事实上显然有 $e = e + e, e = e + 3e$).

例 4 设 $\alpha, \beta \in V$, 且 $\|\alpha\| = \|\beta\|$, 试证必有正交变换 σ , 使 $\sigma\alpha = \beta$.

证 若 $\|\alpha\| = \|\beta\| = 0$, 显然任意正交变换均可.

若 $\|\alpha\| = \|\beta\| \neq 0$, 记

$$\varepsilon = \frac{\alpha}{\|\alpha\|}, \quad \eta = \frac{\beta}{\|\beta\|}.$$

由 ε, η 分别扩充成 V 的两组标准正交基

$$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n \text{ 与 } \eta_1, \eta_2, \dots, \eta_n.$$

对 " $\gamma \in V$, 记 $\gamma = x_1 \varepsilon_1 + x_2 \varepsilon_2 + \dots + x_n \varepsilon_n$, 定义变换 σ 为

$$\sigma \gamma \stackrel{\text{def}}{=} x_1 \eta_1 + x_2 \eta_2 + \dots + x_n \eta_n,$$

则, 对 " $\gamma_1, \gamma_2 \in V$, 设 $\gamma_1 = \sum_{i=1}^n x_i \varepsilon_i, \gamma_2 = \sum_{i=1}^n y_i \varepsilon_i$, 有

$$\begin{aligned} \sigma(\gamma_1 + \gamma_2) &= \sigma\left(\sum_{i=1}^n x_i \varepsilon_i + \sum_{i=1}^n y_i \varepsilon_i\right) = \sigma\sum_{i=1}^n (x_i + y_i) \varepsilon_i \\ &= \sum_{i=1}^n (x_i + y_i) \eta_i = \sum_{i=1}^n x_i \eta_i + \sum_{i=1}^n y_i \eta_i = \sigma\gamma_1 + \sigma\gamma_2, \end{aligned}$$

$$\sigma(k\gamma_1) = \sigma\sum_{i=1}^n kx_i \varepsilon_i = \sum_{i=1}^n (kx_i) \eta_i = k\sum_{i=1}^n x_i \eta_i = k\sigma\gamma_1.$$

所以 σ 是线性变换, 且 $\sigma(\varepsilon_i) = \eta_i (i = 1, \dots, n)$.

又 σ 把标准正交基变成标准正交基, 故 σ 是正交变换, 且

$$\sigma(\alpha) = \sigma(\|\alpha\| \varepsilon_1) = \|\alpha\| \eta_1 = \|\beta\| \eta_1 = \beta.$$

讲评 (1) 本题证明最关键的一步是写出这个变换的“功能”, 即任取 $\gamma \in V$,

$$\sigma\gamma = x_1 \eta_1 + x_2 \eta_2 + \dots + x_n \eta_n,$$

其中 $x_1, x_2, \dots, x_n$ 是 γ 在基 $\varepsilon_1, \dots, \varepsilon_n$ 下的坐标. 即给出了 $\sigma\gamma$ 的表达式. 证明中不能只说“存在”一个变换这样的话.

(2) 证明的结论部分包含 4 层意思: ① 给出一个 V 上的变换 σ ; ② 要求 σ 是一个线性变换; ③ 要求它是正交变换; ④ 满足 $\sigma(\alpha) = \beta$.

例 5 设 $\alpha, \dots, \alpha_m$ 和 $\beta_1, \dots, \beta_m$ 是 n 维欧几里得空间的两个向量组, 试证存在一个正交变换 σ , 使得 $\sigma(\alpha_i) = \beta_i (i = 1, \dots, m)$ 的充分必要条件是

$$(\alpha_i, \alpha_j) = (\beta_i, \beta_j) \quad (i, j = 1, 2, \dots, m).$$

证 (必要性) 显然. 因为正交变换保持内积不变

$$(\alpha, \alpha) = (\sigma\alpha, \sigma\alpha) = (\beta, \beta).$$

$$(\text{充分性}) \text{ 当 } (\alpha, \alpha) = (\beta, \beta) \quad (i, j = 1, \cdots, m). \quad (1)$$

令 $V_1 = L(\alpha_1, \cdots, \alpha_m)$, $V_2 = L(\beta_1, \cdots, \beta_m)$, 且有空间分解式

$$V = V_1 \quad V_1^\perp = V_2 \quad V_2^\perp. \quad (2)$$

现作 V_1 到 V_2 的映射 ϕ 如下: 任取 $\alpha \in V_1$, 设 $\alpha = k_1 \alpha_1 + \cdots + k_m \alpha_m$, 定义

$$\phi(\alpha) = \phi(k_1 \alpha_1 + \cdots + k_m \alpha_m) = k_1 \beta_1 + \cdots + k_m \beta_m = \beta \in V_2, \quad (3)$$

则根据(1)可推出 ϕ 为 V_1 到 V_2 的一个同构映射^[1].

由(2)知 $\dim V_1^\perp = \dim V_2^\perp$, 所以 $V_1^\perp$ 与 $V_2^\perp$ 之间可建立同构映射 ϕ 则 " $\gamma \in V$, 有 $\gamma = \gamma_1 + \gamma_1'$, 其中 $\gamma_1 \in V_1, \gamma_1' \in V_1^\perp$, 令

$$\sigma\gamma = \phi \gamma_1 + \phi \gamma_1'.$$

可以验证 σ 是 V 的一个线性变换^[2], 且保持向量内积^[3], 故 σ 是正交变换.

又因 $\alpha = \alpha + \theta$ (θ 为零向量), 所以

$$\sigma(\alpha) = \phi \alpha + \phi \theta = \phi \alpha = \beta \quad (i = 1, \cdots, m).$$

注 (1) 显然 ϕ 为满射; 下证 ϕ 为单射, 即证 $\text{Ker} \phi = \{\theta\}$.

设 $\alpha \in \text{Ker} \phi \quad V_1$, 则 $\alpha = \sum_{i=1}^m k_i \alpha_i$. 于是 $\theta = \phi(\alpha) = \sum_{i=1}^m k_i \beta_i$.

又因 $(\beta, \beta) = (\alpha, \alpha)$, 故

$$0 = \sum_{i=1}^m k_i \beta_i, \sum_{j=1}^m k_j \beta_j = \sum_{i=1}^m k_i \sum_{j=1}^m k_j (\alpha, \alpha) = (\alpha, \alpha) \rightarrow \alpha = \theta.$$

所以 ϕ 是双射, 又由(3)知, ϕ 保加法, 保数乘. 因

$$\phi(\alpha + \beta) = \phi \alpha + \phi \beta, \quad \phi(k\alpha) = k\phi(\alpha),$$

故 ϕ 是同构映射.

(2) " $a, b \in V$, 设 $a = a_1 + a_2, b = b_1 + b_2$, 其中 $a_1, b_1 \in V_1, a_2, b_2 \in V_1^\perp$, 有

$$\begin{aligned} \sigma(a + b) &= \sigma(a_1 + a_2 + b_1 + b_2) = \phi(a_1 + b_1) + \phi(a_2 + b_2) \\ &= \phi(a_1) + \phi(b_1) + \phi(a_2) + \phi(b_2) = \sigma(a) + \sigma(b), \end{aligned}$$

$$\begin{aligned}\sigma(ka) &= \sigma(ka_1 + ka_2) = \Phi(ka_1) + \Phi(ka_2) \\ &= k(\Phi a_1 + \Phi a_2) = k\sigma a,\end{aligned}$$

所以 σ 是线性变换.

$$\begin{aligned}(3) \text{ 因 } (\sigma a, \sigma b) &= (\Phi a_1 + \Phi a_2, \Phi b_1 + \Phi b_2) \\ &= (\Phi a_1, \Phi b_1) + (\Phi a_2, \Phi b_1) + (\Phi a_1, \Phi b_2) + \\ &\quad (\Phi a_2, \Phi b_2).\end{aligned}$$

由 $(\alpha, \alpha) = (\beta, \beta)$, 知有 $(\Phi(a_1), \Phi(b_1)) = (a_1, b_1)$; 因 Φ 是同构映射, 可得 $(\Phi(a_2), \Phi(b_2)) = (a_2, b_2)$.

$$\begin{aligned}\text{又} \quad (\Phi(a_1), \Phi(b_2)) &= 0 = (a_1, b_2), \\ (\Phi(a_2), \Phi(b_1)) &= 0 = (a_2, b_1).\end{aligned}$$

$$\begin{aligned}\text{所以有} \quad (\sigma a, \sigma b) &= (a_1, b_1) + (a_2, b_2) + (a_1, b_2) + (a_2, b_1) \\ &= (a_1 + a_2, b_1 + b_2) = (a, b).\end{aligned}$$

讲评 (1) 若已知两个向量组都是线性无关的, 可以有另一个简单的证法:

将 $\alpha, \dots, \alpha_m$ 施密特正交化为标准正交向量组 $\varepsilon_1, \dots, \varepsilon_m$, 且设

$$(\varepsilon_1, \dots, \varepsilon_m) = (\alpha_1, \dots, \alpha_m) A,$$

其中

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix}.$$

$$\text{令} \quad (e_1, \dots, e_m) = (\beta_1, \dots, \beta_m) A,$$

$$\begin{aligned}\text{则} \quad (e_i, e_j) &= \sum_{k=1}^m a_{ki} \beta_k, \sum_{l=1}^m a_{lj} \beta_l = \sum_{k=1}^m \sum_{l=1}^m a_{ki} a_{lj} (\beta_k, \beta_l) \\ &= \sum_{k=1}^m \sum_{l=1}^m a_{ki} a_{lj} (\alpha_k, \alpha_l) = (\varepsilon_i, \varepsilon_j) = \delta_{ij}.\end{aligned}$$

故当 $\varepsilon_1, \dots, \varepsilon_m$ 为标准正交向量组时, $e_1, \dots, e_m$ 也是标准正交向量组. 把它们扩充为空间两组标准正交基

$$\varepsilon_1, \cdots, \varepsilon_m, \cdots, \varepsilon_n; \quad e_1, \cdots, e_m, \cdots, e_n.$$

任取 $\alpha \in V$, 设 $\alpha = \sum_{i=1}^n x_i \varepsilon_i$, 定义 $\alpha = x_1 e_1 + \cdots + x_n e_n$, 则可证 σ 为线性变换. 又把标准正交基映为标准正交基, 故是正交变换, 且有 $\sigma(\alpha) = \beta$.

(2) 当不加“线性无关”的条件时, 若走(1)的思路要讨论, 则证明太长. 目前没有找到比题解中更好的办法.

例 6 设 η 是 n 维欧几里得空间的一个单位向量, 定义 $\sigma(\alpha) = \alpha - 2(\eta, \alpha)\eta$, 试证:

(1) σ 是正交变换 (这样的变换称为镜面反射);

(2) σ 是第二类的;

(3) 若正交变换 τ 以 1 作为一个特征值, 且属于特征值 1 的特征子空间 V_1 的维数是 $n - 1$, 那么 τ 是镜面反射.

解 (1) 因对 $\alpha, \beta \in V$, 有

$$\begin{aligned} \sigma(\alpha + \beta) &= (\alpha + \beta) - 2(\eta, \alpha + \beta)\eta \\ &= \alpha + \beta - 2(\eta, \alpha)\eta - 2(\eta, \beta)\eta \\ &= \alpha + \beta, \end{aligned}$$

$$\sigma(k\alpha) = k\alpha - 2(\eta, k\alpha)\eta = k[\alpha - 2(\eta, \alpha)\eta] = k\sigma\alpha,$$

故 σ 是线性变换.

又

$$(\sigma\alpha, \sigma\beta) = (\alpha - 2(\eta, \alpha)\eta, \beta - 2(\eta, \beta)\eta) = (\alpha, \beta),$$

所以 σ 是一个正交变换.

(2) 将 η 扩充为 V 的一组标准正交基 $\eta, \eta_2, \cdots, \eta_n$, 则由于

$$\sigma\eta = \eta - 2(\eta, \eta)\eta = -\eta, \quad \sigma\eta_i = \eta_i \quad (i = 2, \cdots, n),$$

知 σ 在此基下的矩阵为 $A = \text{diag}(-1, 1, \cdots, 1)$, 显然, $\det A = -1$, 故 σ 是第二类的.

(3) 取 V_1 的一组标准正交基 $\varepsilon_1, \cdots, \varepsilon_{n-1}$, 则由于 $V_1^\perp$ 是 τ 的不变子空间, 故任取 $\varepsilon_n \in V_1^\perp$, 则有 $\tau\varepsilon_n = \lambda\varepsilon_n$. 又 τ 的实特征值为 ± 1 ,

而 τ 恰有 $n-1$ 个属于 $\lambda=1$ 的线性无关的特征向量, 所以 $\lambda=-1$, 从而 τ 在基 $\varepsilon_1, \dots, \varepsilon_{n-1}, \varepsilon_n$ 下的矩阵为 $A = \text{diag}\{1, \dots, 1, -1\}$.

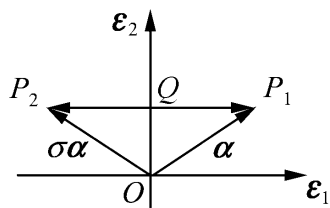
" $\alpha = x_1 \varepsilon_1 + x_2 \varepsilon_2 + \dots + x_n \varepsilon_n \in V$, 有

$$\begin{aligned}\tau(\alpha) &= \tau(x_1 \varepsilon_1 + x_2 \varepsilon_2 + \dots + x_n \varepsilon_n) = x_1 \varepsilon_1 + \dots + x_{n-1} \varepsilon_{n-1} - x_n \varepsilon_n \\ &= x_1 \varepsilon_1 + x_2 \varepsilon_2 + \dots + x_n \varepsilon_n - 2x_n \varepsilon_n = \alpha - 2(\alpha, \varepsilon_n) \varepsilon_n,\end{aligned}$$

故 τ 是镜面反射.

讲评 (1) 题设中线性变换 σ 的几何

意义: 在平面 V 中取两个正交的向量 $\varepsilon_1, \varepsilon_2$, 令 $\eta = \varepsilon_2$, 则对任意 $\alpha \in V$, $\sigma\alpha = \alpha - 2(\eta, \alpha)\eta$ 表达的是如下的向量之间的关系



$$\sigma\alpha = \overline{OP_2} = \overline{OP_1} + \overline{P_1P_2} = \overline{OP_1} - \overline{P_2P_1} = \overline{OP_1} - 2\overline{QP_1}.$$

故 σ 是(对 ε_2 向量)的“反演”, 即照镜子, 故称为镜面反射.

(2) 问题(3)的证明中最后的推导不能少. 只有证出 τ 满足 " $\alpha \in V$, 有 $\tau(\alpha) = \alpha - 2(\alpha, \varepsilon_n) \varepsilon_n$, 才能说明 τ 是镜面反射.

例 7 向量 $\alpha = (1, 0, 2, 1)$, $\alpha = (2, 1, 2, 3)$, $\alpha = (0, 1, -2, 1)$ 张成的子空间为 W , 求 W 的正交补 $W^\perp$ 的标准正交基.

解 设

$$A = \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 3 \\ 0 & 1 & -2 & 1 \end{pmatrix},$$

则 W 同构于 $R(A^T)$, 即 A^T 的列空间, 故 $W^\perp$ 同构于齐次线性方程组 $Ax = 0$ 的解空间(也就是 $N(A)$). 由于

$$A \rightarrow \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \text{所以 } \eta_1 = \begin{pmatrix} -2 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad \eta_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

又因为 $(\eta_1, \eta_2) = 0$, 所以 η_1, η_2 正交, 故

$$\boldsymbol{\varepsilon}_1 = \frac{1}{3} \boldsymbol{\eta}_1 = \frac{1}{3}(-2, 2, 1, 0)^T, \quad \boldsymbol{\varepsilon}_2 = \frac{1}{3} \boldsymbol{\eta}_2 = \frac{1}{3}(1, 1, 0, 1)^T$$

是 $N(\mathbf{A})$ 的标准正交基, 从而

$$\boldsymbol{\varepsilon}_1^T = \frac{1}{3}(-2, 2, 1, 0), \quad \boldsymbol{\varepsilon}_2^T = \frac{1}{3}(1, 1, 0, 1)$$

是 $W^\perp$ 的标准正交基.

讲评 此题主要利用齐次线性方程组的几何意义: 因 $\mathbf{Ax} = \mathbf{0}$, 故 $R(\mathbf{A}^T) \perp N(\mathbf{A})$, 即 $R(\mathbf{A}^T) = N(\mathbf{A})^\perp$, 它们都是 $\mathbb{R}^4$ 的子空间. 而题设中给的是行向量, 即 $\mathbf{A}^T$ 的列可看成已知行向量在自然基的坐标列. 故求出 $R(\mathbf{A}^T)^\perp$ 后, 可以看成得到的是 $W^\perp$ 中向量的坐标构成的 $\mathbb{R}^4$ 的子空间, 回到行向量空间即得 $W^\perp$ 的标准正交基.

例 8 线性子空间 W 用下列方程组给出

$$\begin{aligned} 2x_1 + x_2 + 3x_3 - x_4 &= 0, \\ 3x_1 + 2x_2 - 2x_4 &= 0, \\ 3x_1 + x_2 + 9x_3 - x_4 &= 0. \end{aligned}$$

求给出正交补 $W^\perp$ 的诸方程.

解 用消元法解齐次线性方程组得基础解系

$$\boldsymbol{\eta}_1 = \begin{pmatrix} -6 \\ 9 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\eta}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad W = L(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2),$$

则 $W^\perp$ 由方程组

$$\begin{aligned} \boldsymbol{\eta}_1^T \mathbf{x} &= 0, & \text{即} & & 6x_1 - 9x_2 - x_3 &= 0, \\ \boldsymbol{\eta}_2^T \mathbf{x} &= 0, & & & x_2 + x_4 &= 0. \end{aligned}$$

给出.

讲评 设题设方程组为 $\mathbf{Ax} = \mathbf{0}$, 则 W 是此齐次线性方程组的解空间, 即 $W = N(\mathbf{A})$. 而 $\mathbf{A}^T$ 的列空间 $R(\mathbf{A}^T) = N(\mathbf{A})^\perp = W^\perp$. 但此题不是要求出 $W^\perp$ 的维数和基, 而是要求出能给出 $W^\perp$ 的方程

组,即找 $Bx=0$, 使 $W^\perp = N(B)$. 由 $R(B^T) \perp N(B)$ 知只要选 B^T 的列是 W 的基就可以了. 故令

$$B^T = (\eta_1, \eta_2), \quad \text{即 } B = \begin{pmatrix} \eta_1^T \\ \eta_2^T \end{pmatrix}.$$

例 9 设在欧几里得空间 V 中, 正交线性变换 A 在某标准正交基 $\alpha, \dots, \alpha_n$ 下的矩阵表示为以下方阵 A . 求 V 的标准正交基 $e_1, \dots, e_n$ 使 V 分解为一维和二维不变子空间直和, A 在二维子空间上作用为旋转.

$$(1) A = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}; \quad (2) A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

解 (1) ① 求 A 的特征值

$$f(\lambda) = \det(\lambda I - A) = (\lambda - 1)(\lambda^2 - \lambda + 1),$$

所以 $\lambda_1 = 1, \quad \lambda_2 = \frac{1 + 3i}{2}, \quad \lambda_3 = \frac{1 - 3i}{2}.$

② 分别求 A 的属于特征值 λ_i 的特征向量得

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -\frac{1+3i}{2} \\ \frac{3i-1}{2} \\ 1 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} \frac{3i-1}{2} \\ -\frac{1+3i}{2} \\ 1 \end{pmatrix},$$

故 A 实正交相合于

$$A \sim \begin{pmatrix} 1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \frac{1}{2} & -\frac{3}{2} \\ & \frac{3}{2} & \frac{1}{2} \end{pmatrix} = B.$$

③ 把 $\mathbf{x}$ 单位化得 ε , 取 $\mathbf{x}, (\mathbf{x})$ 的实部、虚部得两个向量, 它们构成二维子空间 $L(\mathbf{x}_2, \mathbf{x}_3)$ 的两个实的基向量, 再把它们单位化得 ε, ε , 取

$$Q = (\varepsilon, \varepsilon, \varepsilon) = \begin{pmatrix} \frac{1}{3} & -\frac{1}{6} & -\frac{1}{2} \\ \frac{1}{3} & -\frac{1}{6} & +\frac{1}{2} \\ \frac{1}{3} & \frac{2}{6} & 0 \end{pmatrix},$$

则有 $Q^T A Q = Q^{-1} A Q = B$.

因为 Q 的列就是所求标准正交基 $e_1, \dots, e_n$ 在 $\alpha_1, \dots, \alpha_n$ 下的坐标列, 即

$$(e_1, \dots, e_n) = (\alpha_1, \dots, \alpha_n) Q.$$

又 Q 的第 j 列为 ε_j , 由此得

$$e_j = (\alpha_1, \dots, \alpha_n) \varepsilon_j, \quad j = 1, 2, \dots, n.$$

故(此题 $n=3$)

$$e_1 = (\alpha_1, \alpha_2, \alpha_3) \varepsilon_1 = \frac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3);$$

$$e_2 = (\alpha_1, \alpha_2, \alpha_3) \varepsilon_2 = \frac{1}{6}(-\alpha_1 - \alpha_2 + 2\alpha_3);$$

$$e_3 = (\alpha_1, \alpha_2, \alpha_3) \varepsilon_3 = \frac{1}{2}(-\alpha_1 + \alpha_2);$$

$$V = \langle e_1 \rangle + \langle e_2, e_3 \rangle.$$

且 A 在不变子空间 $\langle e, e \rangle$ 上的作用为旋转 因为 $\cos\theta = \frac{1}{2}$, $\sin\theta = \frac{\sqrt{3}}{2}$, 所以 $\theta = \frac{\pi}{3}$, 即是 $\langle e, e \rangle$ 绕 e 转 $\frac{\pi}{3}$. 或说成是 yOz 平面绕 x 轴旋转 $\frac{\pi}{3}$.

(2) A 的特征值为 $1, -1, i, -i$, 特征向量分别取为

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -i \\ i \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} i \\ -i \\ 1 \\ 1 \end{pmatrix}.$$

把 $\mathbf{x}_1, \mathbf{x}_2$ 单位化得 $\mathbf{e}_1, \mathbf{e}_2$, 取 $\mathbf{x}_3$ 的实部和虚部并单位化得 $\mathbf{e}_3, \mathbf{e}_4$. 令 $Q = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$, 即有

$$Q = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \text{则 } Q^{-1}AQ = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & i & \\ & & & -i \end{pmatrix}.$$

标准正交基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ 满足

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)Q,$$

故得

$$\mathbf{e}_1 = \frac{1}{2}(\alpha_1 + \alpha_2); \quad \mathbf{e}_2 = \frac{1}{2}(-\alpha_1 + \alpha_2);$$

$$\mathbf{e}_3 = \frac{1}{2}(\alpha_3 + \alpha_4); \quad \mathbf{e}_4 = \frac{1}{2}(-\alpha_3 + \alpha_4);$$

$$V = \langle \mathbf{e}_1 \rangle + \langle \mathbf{e}_2 \rangle + \langle \mathbf{e}_3, \mathbf{e}_4 \rangle.$$

A 在 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ 上的矩阵表示为

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \cos\theta & \sin\theta \\ & & -\sin\theta & \cos\theta \end{pmatrix}, \quad \text{其中 } \theta = \frac{\pi}{2}.$$

讲评 (1) 题设中给出的两个矩阵都是正交矩阵, 它们的实特征值为 ± 1 , 复特征值是模为 1 的复数.

(2) 这两个正交矩阵都有 n 不同的特征值(这里 $n=3$, 或 4), 故它们都能相似于对角矩阵.

(3) 因它们都有一对共轭复特征值, 从而特征向量也是复数域上的向量, 故它们只能在复数域上相似于对角矩阵.

(4) 它们在实数域上均不能相似于对角矩阵, 即 V 不能分解为 A 的一维不变子空间的直和, V 只能分解为 A 的一维(由实特征值, 实特征向量提供)和二维(由复特征值和特征向量提供)不变子空间的直和.

(5) A 的二维不变子空间求法的理论根据(证明), 设

$$Ax = \lambda x, \text{ 则两边取共轭得 } A\bar{x} = \bar{\lambda}\bar{x},$$

即当 x 为 A 的属于特征值 λ 的特征向量时, $\bar{x}$ 是 A 的属于特征值 $\bar{\lambda}$ 的特征向量, 记 $\lambda = \lambda_{\mathbb{R}} + i\lambda_{\mathbb{I}}$, $x = x_{\mathbb{R}} + ix_{\mathbb{I}}$ ($\bar{\lambda} = \lambda_{\mathbb{R}} - i\lambda_{\mathbb{I}}$, $\bar{x} = x_{\mathbb{R}} - ix_{\mathbb{I}}$), 则有

$$A(x + \bar{x}) = \lambda x + \bar{\lambda}\bar{x}, \quad (1)$$

$$A(x - \bar{x}) = \lambda x - \bar{\lambda}\bar{x}. \quad (2)$$

由(1)式得

$$2Ax_{\mathbb{R}} = 2\lambda_{\mathbb{R}}x_{\mathbb{R}} - \lambda_{\mathbb{I}}x_{\mathbb{I}}, \text{ 即 } Ax_{\mathbb{R}} = \lambda_{\mathbb{R}}x_{\mathbb{R}} - \lambda_{\mathbb{I}}x_{\mathbb{I}}.$$

由(2)式得

$$Ax_{\mathbb{I}} = \lambda_{\mathbb{R}}x_{\mathbb{I}} + \lambda_{\mathbb{I}}x_{\mathbb{R}}.$$

故有

$$A(x_{\mathbb{R}}, x_{\mathbb{I}}) = (x_{\mathbb{R}}, x_{\mathbb{I}}) \begin{pmatrix} \lambda_{\mathbb{R}} & \lambda_{\mathbb{I}} \\ -\lambda_{\mathbb{I}} & \lambda_{\mathbb{R}} \end{pmatrix}.$$

在问题(1)和问题(2)中 $\begin{pmatrix} \lambda_{\mathbb{R}} & \lambda_{\mathbb{I}} \\ -\lambda_{\mathbb{I}} & \lambda_{\mathbb{R}} \end{pmatrix}$ 分别为

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

例 10 非奇异二次型可以用正交变换化为规范型(即元素为 ± 1 的对角形)当且仅当它的矩阵是正交矩阵.

证 设二次型为 $Q(x_1, \dots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, 且有 $\mathbf{A}^T = \mathbf{A}$, $\det \mathbf{A} \neq 0$, 则原命题就是要证: 存在正交矩阵 $\mathbf{Q}$, 使

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(1, \dots, 1, -1, \dots, -1) \quad \mathbf{A}^T \mathbf{A} = \mathbf{I}.$$

由 $\mathbf{A}^T \mathbf{A} = \mathbf{I}$, 又 $\mathbf{A}^T = \mathbf{A}$, 所以 $\mathbf{A}^2 = \mathbf{I}$, 设 $\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$, 则有

$$\mathbf{x} = \mathbf{I} \mathbf{x} = \mathbf{A}^2 \mathbf{x} = \mathbf{A} \lambda \mathbf{x} = \lambda^2 \mathbf{x}, \text{ 故得 } \lambda^2 = 1, \text{ 即 } \lambda = \pm 1.$$

故由 1.1 节定理 1 可知, 存在正交矩阵 $\mathbf{Q}$, 使

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \begin{pmatrix} \mathbf{I}_r & \\ & -\mathbf{I}_{n-r} \end{pmatrix}.$$

$$\text{由} \quad \mathbf{Q}^T \mathbf{A} \mathbf{Q} = \begin{pmatrix} \mathbf{I}_r & \\ & -\mathbf{I}_{n-r} \end{pmatrix},$$

得

$$\mathbf{A} = \mathbf{Q} \begin{pmatrix} \mathbf{I}_r & \\ & -\mathbf{I}_{n-r} \end{pmatrix} \mathbf{Q}^T,$$

故

$$\mathbf{A}^2 = \mathbf{I},$$

又因为 $\mathbf{A}^T = \mathbf{A}$, 故 $\mathbf{A}^T \mathbf{A} = \mathbf{A}^2 = \mathbf{I}$, 所以 $\mathbf{A}$ 是正交矩阵.

讲评 (1) 证明题要看清题目条件, 明确要证的结论是什么, 从而向“目标”前进, 本题经分析、联想知是在条件 $\mathbf{A}^T = \mathbf{A}$ 下, 证明存在正交矩阵 $\mathbf{Q}$ 使

$$\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} = \begin{pmatrix} \mathbf{I} & \\ & -\mathbf{I}_{n-r} \end{pmatrix} \quad \mathbf{A}^T \mathbf{A} = \mathbf{I}.$$

而左边又等价于 $\mathbf{A}^2 = \mathbf{I}$. 故事实上是证在 $\mathbf{A}^T = \mathbf{A}$ 条件下

$$\mathbf{A}^2 = \mathbf{I} \quad \mathbf{A}^T \mathbf{A} = \mathbf{I}.$$

必要性是由 $\mathbf{A}^T = \mathbf{A}$, $\mathbf{A}^2 = \mathbf{I}$ 推出 $\mathbf{A}^T \mathbf{A} = \mathbf{I}$; 而充分性是由 $\mathbf{A}^T =$

$A, A^T A = I$ 推出 $A^2 = I$.

(2) 前面已一再强调“在条件 $A^T = A$ 下”,这一条件是暗含于题设中的(实二次型的矩阵是实对称矩阵),故同学不能由此题得出“正交矩阵的特征值都是 ± 1 ”这样的错误结论.应是既对称又正交的矩阵它的特征值才是 ± 1 .

例 11 设 K 是 n 阶反对称矩阵, 并设 λ 是它的纯虚特征根, α 是属于特征根 λ 的 K 的特征向量, 记 $\alpha = \beta + i\gamma$, 试证: β 和 γ 正交且长度相等.

证 方法 1 设 $\lambda = \mu i$ 是 K 的特征根, 则有

$$K(\beta + i\gamma) = \mu i(\beta + i\gamma),$$

展开得

$$K\beta = -\mu\gamma, \quad (1)$$

$$K\gamma = -\mu\beta. \quad (2)$$

用 γ^T 左乘(1)式得 $\gamma^T K\beta = -\mu\gamma^T \gamma$, (2)式转置后右乘 β 得 $\gamma^T K\beta = \mu\beta^T \beta$. 这两式左边相等, 故有 $\mu(\beta^T \beta - \gamma^T \gamma) = 0$, 所以 $\beta^T \beta = \gamma^T \gamma$, 即 β 与 γ 长度相等.

又因为 K 是反对称的, 所以对任 x , 恒有 $x^T Kx = 0$, (1)式左乘 β^T 得 $\beta^T K\beta = -\mu\beta^T \gamma = 0$, 所以 $\beta^T \gamma = 0$, 故 β 与 γ 正交.

方法 2 因为实系数多项式的复根共轭出现, 所以当 λ 为 K 的特征根时, $\bar{\lambda} = -i\mu$ 也是 K 的特征根. 故由

$$K\alpha = \lambda\alpha \text{ 知 } K\bar{\alpha} = \bar{\lambda}\bar{\alpha}.$$

又由 K 的属于不同特征值的特征向量是正交的, 于是有

$$\begin{aligned} 0 &= (\beta + i\gamma, \beta - i\gamma) = (\beta, \beta) + (i\gamma, \beta) + (\beta, -i\gamma) + (i\gamma, -i\gamma) \\ &= (\beta, \beta) - (\gamma, \gamma) - i(\gamma, \beta) - i(\beta, \gamma). \end{aligned}$$

由实部为 0, 虚部为 0, 并利用 $(\beta, \gamma) = (\gamma, \beta)$ 知

$$(\beta, \beta) = (\gamma, \gamma), \quad (\beta, \gamma) = 0.$$

讲评 (1) K 是实的反对称矩阵, 故它的特征多项式为实系

数多项式. 于是由 3.1 节定理 12 知它的复特征值成共轭对出现.

(2) 解答中给出的两个方法各有特色, 方法 1 的思路是大部分人可以想到的, 只不过有的人嫌繁, 有的人算错, 才做不出结果. 方法 2 计算量小, 思路比较巧, 用了两条理论知识: ① 当 $K\alpha = \lambda\alpha$ 时, 也有 $K\bar{\alpha} = \bar{\lambda}\bar{\alpha}$; ② K 的属于不同特征值的特征向量相互正交 (对规范矩阵都对), 缺少任何一条也证不出来. 所以, 学过的知识在头脑中有条理地“存储”, 做题时就会有巧思路.

例 12 试证: 正交矩阵的任一子方阵的特征根的模小于或等于 1.

证 设正交矩阵

$$Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_1 & Q_2 \end{pmatrix},$$

则由 $Q^T Q = I$ 可得关系式

$$Q_1^T Q_1 + Q_2^T Q_2 = I.$$

对 Q 的子方阵 Q_1 , 设 λ 为其特征值, x 为相应的特征向量, 于是有

$$Q_1 x = \lambda x.$$

两边取共轭转置后再与原式相乘有

$$\bar{\lambda} \bar{x}^T x = \bar{x}^T Q_1^T Q_1 x = \bar{x}^T (I - Q_2^T Q_2) x,$$

所以

$$(1 - |\lambda|^2) \bar{x}^T x = \bar{x}^T Q_2^T Q_2 x \geq 0$$

(因为 $Q_2 x$ 是一个列向量, 右边是这个向量的模平方, 所以非负).

又 $\bar{x}^T x > 0$, 所以 $|\lambda| \leq 1$.

又对 Q 的任一子方阵, 可通过交换行、列把它换到“左上角”, 这可通过方阵两边同乘第一种初等矩阵完成, 且后者仍是正交矩阵 (单位矩阵经交换两行后得到的矩阵称为第一种初等矩阵).

讲评 乍一看此题, 可能不少同学会觉得无从下手. 但仔细分析题设条件. 说 Q 是正交矩阵, 故有 $Q^T Q = I$, 要讨论它的子矩阵的特征值, 就得联想到特征值的定义, 先讨论主子矩阵, 再考虑一般的子矩阵, 于是有 $Q_1 x = \lambda x$. 再有要讨论 λ 的模, 而 $|\lambda|^2 = \lambda \bar{\lambda}$, 故要把 $Q_1 x = \lambda x$ 两边取共轭转置后相乘, 这样一分析、联想, 就有

了如上的解法 .

例 13 如果 A 和 B 都是正交矩阵, 且 $\det A = -\det B$, 则矩阵 $A+B$ 是奇异矩阵 .

证 方法 1 因为 A, B 均是正交矩阵, 所以 $A^T B$ 仍为正交矩阵, 且 $\det(A^T B) = -1$. 又复根共轭出现, 特征根模均为 1, 行列式等于特征值之积, 故 -1 为 $A^T B$ 的特征根, 所以

$$|-I - A^T B| = 0, \text{ 即 } |I + A^T B| = 0.$$

故 $|A+B| = |A| |I + A^T B| = 0,$

于是得 $A+B$ 是奇异矩阵 .

方法 2 由 $A^T A = B^T B = I, \det A \neq 0$, 可知

$$\begin{aligned} \det(A+B) &= \det A \det(I + A^T B) & (*) \\ &= \det B \cdot \det(I + B^T A) = -\det A \cdot \det(I + B^T A). \end{aligned}$$

于是得

$$\det(I + A^T B) = -\det(I + B^T A) = -\det(I + A^T B)$$

(因 $(I + B^T A)^T = I + A^T B$, 故最后一个等号成立). 所以 $\det(I + A^T B) = 0$. 由 $(*)$ 式即知有

$$\det(A+B) = 0, \text{ 故 } A+B \text{ 是奇异矩阵.}$$

讲评 要证 $A+B$ 不可逆, 只要证 $|A+B| = 0$, 左边经变形就得 $|A| |I + A^T B| = 0$, 故问题转化成要证 $|I + A^T B| = 0$, 证此式有两条思路: ①证 -1 为 $A^T B$ 的特征值; ②通过对 $|A+B|$ 的各种恒等变形而得 $|I + A^T B| = -|I + A^T B|$. 它们分别构成方法 1 和方法 2, 此题绝不能用 $|A+B| = |A| + |B| = 0$ 的思想写, 因行列式的运算绝无此性质. 这是为凑结论而犯的错误!

例 14 已知 σ 是 n 维欧几里得空间 V 的一个正交变换, W 是 σ 的不变子空间, 证明, W 的正交补 $W^\perp$ 也是 σ 的不变子空间 .

证 设 $\alpha, \alpha, \dots, \alpha_n$ 是 W 的一组标准正交基, 把它扩充为 V 的标准正交基

$$\alpha, \alpha, \dots, \alpha_n, \beta_{n+1}, \dots, \beta_n,$$

则 $L(\beta_{n+1}, \dots, \beta_n) = W^\perp$. 因为 σ 是正交变换, 故

$$\alpha_1, \dots, \alpha_m, \sigma\beta_{n+1}, \dots, \sigma\beta_n$$

亦是 V 的标准正交基, 所以

$$L(\sigma\beta_{n+1}, \dots, \sigma\beta_n) = (L(\alpha_1, \dots, \alpha_m))^\perp.$$

又因 W 是 σ 的不变子空间, 故 $\alpha_i \in W (i = 1, 2, \dots, m)$, 知 $\alpha_1, \dots, \alpha_m$ 是 W 的标准正交基, 所以有

$$L(\alpha_1, \dots, \alpha_m) = W,$$

于是

$$L(\sigma\beta_{n+1}, \dots, \sigma\beta_n) = W^\perp.$$

任取 $\gamma \in W^\perp$, 设 $\gamma = x_{m+1}\beta_{n+1} + x_{m+2}\beta_{n+2} + \dots + x_n\beta_n$, 则

$$\begin{aligned} \sigma\gamma &= \sigma(x_{m+1}\beta_{n+1} + x_{m+2}\beta_{n+2} + \dots + x_n\beta_n) \\ &= x_{m+1}\sigma\beta_{n+1} + x_{m+2}\sigma\beta_{n+2} + \dots + x_n\sigma\beta_n \\ &\in L(\sigma\beta_{n+1}, \dots, \sigma\beta_n) = W^\perp. \end{aligned}$$

故, $W^\perp$ 也是 σ 的不变子空间.

讲评 (1) 证明的思路要紧扣主题(主要目标): 即证 $W^\perp$ 是 σ 的不变子空间, 依定义, 任取 $\gamma \in W^\perp$, 要证 $\sigma\gamma \in W^\perp$ 即可.

为此需要给出 γ 的线性表示式, 故需要 $W^\perp$ 的基 $\beta_{n+1}, \dots, \beta_n$. 由于 σ 是线性变换, 所以 $\sigma\gamma$ 是 $\sigma\beta_{n+1}, \dots, \sigma\beta_n$ 的线性组合, 故需要先给出 $L(\sigma\beta_{n+1}, \dots, \sigma\beta_n)$ 与 $W^\perp$ 之间的关系. 把需要的东西先叙述清楚(先准备好), 再按正向叙述方向表述出来就行了. 一般证明题基本上是先分析, 有了思路, 再理顺. 主线要一目了然, 不能上来就写, 需要什么现证什么(主线中夹验证).

(2) 这里只能从 W 中选基, 再扩充为空间的基, 不能这样叙述“在 V 中选标准正交基 $\alpha_1, \dots, \alpha_n$, 则它们中的一部分是 W 的基, 不妨设为 $\alpha_1, \dots, \alpha_m$ ”.

例 15 证明第二类正交变换一定以 -1 作为它的一个特征值.

证 要证 -1 为 A 的特征值, 即要证 $|A + I| = 0$, 因

$$|A + I| = |A + A^T A| = |I + A^T| |A| = -|A + I|,$$

故得

$$|\mathbf{A} + \mathbf{I}| = 0.$$

讲评 (1) 要证数 λ 是某矩阵 $\mathbf{A}$ 的特征值, 都是证 $|\lambda \mathbf{I} - \mathbf{A}| = 0$, 上面证明中, 用到 $|\mathbf{A}| = -1$ (因 $\mathbf{A}$ 是第二类正交变换的矩阵) 及 $|\mathbf{I} + \mathbf{A}^T| = |\mathbf{I} + \mathbf{A}|$ (行列式转置后值不变).

(2) 也可用 $\det \mathbf{A} = \lambda_1 \cdots \lambda_n$ (特征值之积) 及正交矩阵的特征值是模为 1 的复数, 复根成共轭对出现 (而 $\lambda \bar{\lambda} = |\lambda|^2 = 1$), 故由 $\lambda_1 \lambda_2 \cdots \lambda_n = -1$ 知必有 $\lambda = -1$ 这样一个思路写出方法 2.

例 16 设矩阵

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

试求 $\mathbf{A}$ 的广义逆 $\mathbf{A}^+$ 和方程组 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 的最小二乘解 $\bar{\mathbf{x}}$, 将 $\mathbf{b}$ 投影到 $R(\mathbf{A})$ 上得到 $\mathbf{p}$, 验证 $\mathbf{A}\bar{\mathbf{x}} = \mathbf{p}$.

解 由于 $\mathbf{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} (1, 1, 1) = \mathbf{L}\mathbf{U}$, 故

$$\mathbf{A}^+ = \mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{1}{3} \frac{1}{2} (1, 1) = \frac{1}{6} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

$$\bar{\mathbf{x}} = \mathbf{A}^+ \mathbf{b} = \frac{1}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{A}\bar{\mathbf{x}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

又因

$$R(\mathbf{A}) = k \begin{pmatrix} 1 \\ 1 \end{pmatrix} = R(\mathbf{a}); \quad \mathbf{p} = \frac{\mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}} \mathbf{a} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

故知, 确有 $\mathbf{p} = \mathbf{A}\bar{\mathbf{x}}$ 为 $\mathbf{b}$ 在 $R(\mathbf{A})$ 上的投影.

讲评 (1) 关于广义逆的求解过程见《高等代数学》.

(2) 因

$$\mathbf{P}\mathbf{A}\mathbf{Q} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\text{故 } A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{1} (1, 1, 1) = LU.$$

6.3 考试题评析

例 1 设 σ 为欧氏空间 V 中的对称变换, 且有 $\sigma^2 = \sigma$, 试证: 在 V 中必存在子空间 W , 使 σ 是 V 到 W 上的一个正交投影(变换). (注: 设 W 是 V 的子空间, 且 $V = W \bar{\cup} W^\perp$, 任 $\alpha \in V$ 有 $\alpha = \beta + \gamma$, 其中 $\beta \in W, \gamma \in W^\perp$, 若 V 上的线性变换 σ 满足 $\sigma\alpha = \beta$, 则称 σ 是 V 到 W 上的一个正交投影(变换).)

证 方法 1 由 $\sigma^2 = \sigma$, 知 σ 是幂等变换, 任取 $\beta \in \text{Im}\sigma, \forall \alpha \in V$, 使 $\beta = \sigma\alpha$, 于是

$$\sigma\beta = \sigma(\sigma\alpha) = \sigma^2\alpha = \sigma\alpha = \beta,$$

故当 $\beta \neq 0$ 时, $\sigma\beta \neq 0$, 即 $\beta \notin \text{Ker}\sigma$, 所以知

$$\text{Im}\sigma \cap \text{Ker}\sigma = \{0\}.$$

故得

$$n = \dim\text{Im}\sigma + \dim\text{Ker}\sigma = \dim(\text{Im}\sigma + \text{Ker}\sigma),$$

所以有空间分解式

$$V = \text{Im}\sigma \bar{\cup} \text{Ker}\sigma.$$

又任取 $\beta \in \text{Im}\sigma (\forall \alpha \in V, \text{使 } \sigma\alpha = \beta), \gamma \in \text{Ker}\sigma$, 有

$$(\beta, \gamma) = (\sigma\alpha, \gamma) = (\alpha, \sigma\gamma) = (\alpha, 0) = 0,$$

其中第 2 个等号是由于 σ 为对称变换, 故得

$$\text{Im}\sigma \perp \text{Ker}\sigma.$$

于是取 $W = \text{Im}\sigma$, 则 $\text{Ker}\sigma = W^\perp$, 且有

$$V = W \bar{\cup} W^\perp = \text{Im}\sigma \bar{\cup} \text{Ker}\sigma$$

任取 $\alpha \in V$, 有 $\alpha = \beta + \gamma$, 其中 $\beta \in W = \text{Im}\sigma, \gamma \in W^\perp = \text{Ker}\sigma$, 且

$$\sigma\alpha = \sigma(\beta + \gamma) = \sigma\beta + \sigma\gamma = \beta + 0 = \beta.$$

故 σ 是 V 到 $\text{Im} \sigma(W)$ 的正交投影.

方法 2 由 $\sigma^2 = \sigma$, 可知 σ 的特征值 λ 满足 $\lambda^2 = \lambda$, 故 $\lambda = 1$ 或 0 , 又由 σ 是对称变换, 故 σ 可对角化 (由 6.1 节定理 9), 即在 V 中存在一组标准正交基 $\eta_1, \dots, \eta_r, \eta_{r+1}, \dots, \eta_n$, 使 σ 在此基上的矩阵表示为

$$D = \begin{pmatrix} I_r & \\ & 0 \end{pmatrix}.$$

令 $W = \langle \eta_1, \dots, \eta_r \rangle = V_{\lambda=1}$, $V_{\lambda=0} = \langle \eta_{r+1}, \dots, \eta_n \rangle$. 因 $\eta_1, \dots, \eta_n$ 是标准正交基, 故 $V_{\lambda=0} = W^\perp$, 即有

$$V = V_{\lambda=1} \bar{1} \quad V_{\lambda=0} = W + W^\perp.$$

又任取 $\beta \in V_{\lambda=1} = W$, 设 $\beta = k_1 \eta_1 + \dots + k_r \eta_r$, 则有

$$\sigma \beta = k_1 \sigma \eta_1 + \dots + k_r \sigma \eta_r = k_1 \eta_1 + \dots + k_r \eta_r = \beta.$$

任取 $\alpha \in V$, 有分解式 $\alpha = \beta + \gamma$, 其中 $\beta \in W = V_{\lambda=1}$, $\gamma \in W^\perp = V_{\lambda=0}$. 且有

$$\sigma \alpha = \sigma(\beta + \gamma) = \sigma \beta + \sigma \gamma = \beta + 0 = \beta.$$

故 σ 是 V 到 $W = V_{\lambda=1}$ 的正交投影.

评析 本题证明的关键是要: ①指出 (或给出) W 是什么, $W^\perp$ 是什么; ②写出 (或证明) 空间分解式 $V = W \bar{1} \quad W^\perp$ (对具体的 W 写出); ③任取 $\alpha \in V$, 有分解式 $\alpha = \beta + \gamma$, 要证明 $\sigma \alpha = \beta$. 这 3 个问题回答清楚, 证明思路就明确了.

方法 1 和 **方法 2** 恰是从不同的角度给出了 W 是什么, $W^\perp$ 是什么, 有的学生写了十几行文字, 老在那里说“存在……”, 始终没有回答 $W, W^\perp$ 都是什么, 故一切都建立在“虚设”的 W 上, 当然证不出来.

例 2 设 σ 为 n 维欧氏空间 V 中的正交变换, A 是 σ 在 V 中一组标准正交基下的矩阵表示, 证明: $|\text{tr} A| \leq n$.

证 因 σ 是正交变换, 由 6.1 节定理 3 知 A 是正交矩阵.

方法 1 先证正交矩阵的特征值是模为 1 的复数, 设 $Ax =$

$\lambda \mathbf{x}$, 两端取共轭转置得

$$\overline{\mathbf{x}}^T \mathbf{A}^T = \overline{\lambda \mathbf{x}}^T .$$

两式相乘得

$$\overline{\mathbf{x}}^T \mathbf{x} = \overline{\mathbf{x}}^T \mathbf{A}^T \mathbf{A} \mathbf{x} = \overline{\lambda} \lambda \overline{\mathbf{x}}^T \mathbf{x} .$$

由 $\overline{\mathbf{x}}^T \mathbf{x} = \|\mathbf{x}\|^2 \neq 0$, 故得 $|\lambda|^2 = \overline{\lambda} \lambda = 1$, $|\lambda| = 1$. 又由特征多项式根与系数的关系有

$$\lambda_1 + \cdots + \lambda_n = \sum_{i=1}^n a_{ii} .$$

故有

$$|\operatorname{tr} \mathbf{A}| = \left| \sum_{i=1}^n a_{ii} \right| = \left| \sum_{i=1}^n \lambda_i \right| \leq \sum_{i=1}^n |\lambda_i| = n .$$

方法 2 设 $\mathbf{A} = (a_{ij})$, 则由 $\mathbf{A} \mathbf{A}^T = \mathbf{I}$ 知

$$\sum_{k=1}^n a_{ik}^2 = 1 \quad (i = 1, \cdots, n) .$$

因 a_{ik} 均为实数, 故 a_{ik}^2 都是正实数. 因此有

$$a_{ii}^2 \leq 1, \text{ 即 } |a_{ii}| \leq 1 \quad (i = 1, \cdots, n) .$$

故得

$$|\operatorname{tr} \mathbf{A}| = \left| \sum_{i=1}^n a_{ii} \right| \leq \sum_{i=1}^n |a_{ii}| \leq n .$$

评析 (1) $\operatorname{tr} \mathbf{A}$ 是 $\mathbf{A}$ 的对角线元素之和. ①把这个和与特征值联系起来(利用多项式根与系数的关系)就是方法 1, 此时还需要知道特征值是模为 1 的复数, 若原来不知这个结论, 由要求证 $|\operatorname{tr} \mathbf{A}| \leq n$, 提示考生去证 $|\lambda| = 1$. 有的同学不去证 $|\lambda| = 1$, 只下断言 $\lambda = \pm 1$, 虽可证出结论, 但不能算全对. 因用简单的情况代替了模为 1 (也反映这个知识点掌握得不好).

②把这个和与 $\mathbf{A}$ 的行向量的模联系起来($\mathbf{A}$ 的行向量或列向量都是单位向量)就得方法 2, 如

$$a_{11}^2 + a_{12}^2 + \cdots + a_{1n}^2 = 1, \text{ 推出 } a_{11}^2 \leq 1 .$$

直接推出每个 $|a_{ii}| \leq 1$, 故得证.

(2) 也可以利用相似的矩阵有相同的迹, A 相似于 Jordan 标准形 J , J 的对角线元素都是特征值 λ_i , 知 $\operatorname{tr} A = \sum_{i=1}^n \lambda_i$.

例 3 设 A 为欧几里得空间 V 中的反对称变换, 即对任意 $\alpha, \beta \in V$, 都有 $(A\alpha, \beta) = -(\alpha, A\beta)$.

(1) 证明 A 在标准正交基下的矩阵是反对称矩阵;

(2) 若 W 是 A 的不变子空间, 证明 $W^\perp$ 也是 A 的不变子空间.

证 (1) 取 V 的标准正交基 $\varepsilon_1, \dots, \varepsilon_n$, 并设 A 在此基下的矩阵表示为 A . 又对任意 $\alpha, \beta \in V$, 设 $\alpha = \sum_{i=1}^n x_i \varepsilon_i$, $\beta = \sum_{j=1}^n y_j \varepsilon_j$, 于是有

$$(A\alpha, \beta) = -(\alpha, A\beta) \text{ 对一切 } \alpha, \beta \in V \text{ 成立.}$$

特别对基向量成立, 所以 $(A\varepsilon_i, \varepsilon_j) = -(\varepsilon_i, A\varepsilon_j)$. 又由 $A(\varepsilon_1, \dots, \varepsilon_n) =$

$$(\varepsilon_1, \dots, \varepsilon_n)A, \text{ 知 } A\varepsilon_i = \sum_{k=1}^n a_{ki} \varepsilon_k, A\varepsilon_j = \sum_{k=1}^n a_{kj} \varepsilon_k, \text{ 故}$$

$$(A\varepsilon_i, \varepsilon_j) = -(\varepsilon_i, A\varepsilon_j) \quad \sum_{k=1}^n a_{ki} \varepsilon_k, \varepsilon_j = -\varepsilon_i, \sum_{k=1}^n a_{kj} \varepsilon_k,$$

即

$$a_{ji} = -a_{ij} \quad A^T = -A.$$

(2) 由 $V = W \oplus W^\perp$, 又因为 W 是 A 的不变子空间, 所以任取 $\beta \in W$, 有 $A\beta \in W$. 于是

" $\alpha \in W^\perp$ (要证 $A\alpha \in W^\perp$), 有

$(A\alpha, \beta) = -(\alpha, A\beta) = 0$ (因为 $\alpha \in W^\perp, A\beta \in W$), 故 $A\alpha \in W^\perp$, 所以 $W^\perp$ 是 A 的不变子空间.

评析 (1) 事实上, 问题(1)可以写为 V 的线性变换 A 是反对称变换的充要条件是 A 在标准正交基下的矩阵为反对称矩阵 (我们以上给的是必要性的证明). 其证明可以简述如下: 对任意

$\alpha, \beta \in V$, 有

$$(A\alpha, \beta) = -(\alpha, A\beta) \quad (A\varepsilon, \varepsilon) = -(\varepsilon, A\varepsilon)$$

$$a_{ji} = -a_{ij} \quad A^T = -A.$$

(2) 本题两个问题的证明都与实对称变换的相应定理和命题类似. 故, 若对实对称变换的结论思路很清楚, 这里是不难有思路的. (1) 的证明主要是对内积的双线性性比较熟练, 从 $(A\varepsilon, \varepsilon) = -(\varepsilon, A\varepsilon)$, 立刻可得 $a_{ji} = -a_{ij}$. 而且这也提示了要把反对称的定义用在基向量上. (2) 的证明是紧扣不变子空间的定义, 任取 $\alpha \in W^\perp$, 证 $\sigma\alpha \in W^\perp$ 就可做出.

例 4 欧几里得空间 V 中保持向量内积不变的变换 σ 一定是线性变换, 从而是正交变换, 这里 σ 保持内积不变的意义为对任意 $\alpha, \beta \in V$ 均有

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta).$$

证 先证明对任意 α, β 均有 $\sigma(\alpha + \beta) = \sigma\alpha + \sigma\beta$. 由于

$$\begin{aligned} & (\sigma(\alpha + \beta) - \sigma\alpha - \sigma\beta, \sigma(\alpha + \beta) - \sigma\alpha - \sigma\beta) \\ &= (\sigma(\alpha + \beta), \sigma(\alpha + \beta)) - 2(\sigma(\alpha + \beta), \sigma\alpha) - 2(\sigma(\alpha + \beta), \sigma\beta) + \\ & \quad (\sigma\alpha, \sigma\alpha) + (\sigma\beta, \sigma\beta) + 2(\sigma\alpha, \sigma\beta) \\ &= (\alpha + \beta, \alpha + \beta) - 2(\alpha + \beta, \alpha) - 2(\alpha + \beta, \beta) + \\ & \quad (\alpha, \alpha) + (\beta, \beta) + 2(\alpha, \beta) = 0, \end{aligned}$$

故

$$\sigma(\alpha + \beta) - \sigma\alpha - \sigma\beta = 0,$$

即

$$\sigma(\alpha + \beta) = \sigma\alpha + \sigma\beta \quad (\alpha, \beta \in V).$$

又对 $k \in \mathbb{R}$, 有

$$\begin{aligned} & (\sigma(k\alpha) - k(\sigma\alpha), \sigma(k\alpha) - k(\sigma\alpha)) \\ &= (\sigma(k\alpha), \sigma(k\alpha)) - k(\sigma(k\alpha), \sigma\alpha) - k(\sigma\alpha, \sigma(k\alpha)) + k^2(\sigma\alpha, \sigma\alpha) \\ &= (k\alpha, k\alpha) - k(k\alpha, \alpha) - k(\alpha, k\alpha) + k^2(\alpha, \alpha) \\ &= k^2(\alpha, \alpha) - k^2(\alpha, \alpha) - k^2(\alpha, \alpha) + k^2(\alpha, \alpha) = 0, \end{aligned}$$

故知

$$\sigma(k\alpha) = k(\sigma\alpha) \quad (k \in \mathbb{R}, \alpha \in V).$$

故 σ 为线性变换且保持内积不变, 从而为正交变换.

评析 (1) 值得注意的是保持向量长度的变换不一定是正交变换, 例如若 $\mathbb{R}^n$ 的内积定义为 $(x, y) = x^T y$, 则变换

$$\sigma: (x_1, \dots, x_n)^T \rightarrow (|x_1|, \dots, |x_n|)^T$$

保持向量长度, 但不是线性变换.

(2) 正交变换的定义(6.1节定义5)及定理(6.1节定理3)都是在 σ 是线性变换的条件下再如何如何. 故我们做任何题目时, 要证明某个变换是正交变换, 先要证它是线性变换, 再证它保持向量的内积不变才能说它是正交变换. 本题恰说明, 若有保持内积不变, 这一条件已暗示着它是一个线性变换(因可证明之). 但(1)指出定理中第2个等价命题“任取 $\alpha \in V, |\sigma\alpha| = |\alpha|$ ”这条必在有“线性变换”的假设下才能保证 σ 是正交变换.

例5 设 A 为 n 阶实正交矩阵.

(1) 若 $\det A = -1$, 证明 -1 必为 A 的特征值;

(2) 若 $\lambda = \alpha + i\beta$ ($\beta \neq 0, \alpha, \beta$ 是实数) 是 A 的特征值, $x = x_1 + ix_2$ ($x_1, x_2 \in \mathbb{R}^n$) 是 A 的属于特征值 λ 的特征向量, 试证明: 向量 x_1 与 x_2 的模相等, 且相互正交.

证 (1) 因为 $\det A = -1, A^T A = I$, 故知

$$|-I - A| = |-A^T A - A| = |A| |-A^T - I| = -|-I - A|,$$

所以 $|-I - A| = 0$, 故 -1 为 A 的特征值.

(2) 因为 $\lambda = \alpha + i\beta$ 是 A 的特征值, A 为实矩阵, 故其特征多项式 $p_A(\lambda)$ 是实系数多项式, 其根成共轭对出现, 所以 $\bar{\lambda} = \alpha - i\beta$ 也是 A 的特征根. 由 $Ax = \lambda x$, 知 $A\bar{x} = \bar{\lambda}\bar{x}$ (前式两端取共轭即得), 又因为正交矩阵 A 的属于不同特征值的特征向量相互正交, 所以 x 与 $\bar{x}$ 正交, 即有

$$0 = (x, \bar{x}) = (x_1 + ix_2, x_1 - ix_2)$$

$$= (\mathbf{x}_1, \mathbf{x}_1) + i(\mathbf{x}_2, \mathbf{x}_1) + i(\mathbf{x}_1, \mathbf{x}_2) - (\mathbf{x}_2, \mathbf{x}_2).$$

由实部,虚部分别等于 0, 知

$$(\mathbf{x}_1, \mathbf{x}_1) = (\mathbf{x}_2, \mathbf{x}_2), \text{ 即 } |\mathbf{x}_1| = |\mathbf{x}_2|;$$

$$i(\mathbf{x}_2, \mathbf{x}_1) + i(\mathbf{x}_1, \mathbf{x}_2) = 2i(\mathbf{x}_1, \mathbf{x}_2) = 0, \text{ 即 } (\mathbf{x}_1, \mathbf{x}_2) = 0.$$

评析 (1) 问题(1)考察的是特征值的定义和计算公式, 满足 $|\lambda I - A| = 0$ 的 λ 是 A 的特征值, 故只要证出 $|(-1)I - A| = 0$ 即知 -1 为 A 的特征值.

(2) 由问题(2)的证明结果知, 实正交矩阵的复特征向量的实部和虚部正交(且模相等), 这正是实正交矩阵相似标准形的基础

取 $\mathbf{x}_1, \mathbf{x}_2$ 为基的一部分, 则在实数域上, σ 在子空间 $L(\mathbf{x}_1, \mathbf{x}_2)$ 上

的限制的矩阵表示为 $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. 与 6.2 节例 9 比较, 加强对实正

交矩阵的特点的理解和认识(如果孤立地看一个个定理和习题, 会觉得枯燥、抽象. 而从总体上去理解, 把各知识面联系起来, 就会发现数学是既深奥又有趣的).

第7章 酉空间

7.1 理论脉络

1 内积的定义、伴随变换的定义

定义1 在 $V_n(\mathbb{C})$ 中定义了一个二元复函数, 即对 " $\alpha, \beta \in V$ ", 确定一个复数, 记作 (α, β) , 它满足:

$$(1) (\alpha, \beta) = \overline{(\beta, \alpha)};$$

$$(2) (k\alpha, \beta) = k(\alpha, \beta); \quad (2') (\alpha, k\beta) = \overline{k}(\alpha, \beta);$$

$$(3) (\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma); \quad (3') (\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma);$$

$$(4) (\alpha, \alpha) \geq 0, \text{ 且 } (\alpha, \alpha) = 0 \iff \alpha = 0.$$

则称 (α, β) 是 α 与 β 的内积, 定义了内积的 $V_n(\mathbb{C})$ 称为酉空间, 记为 $U_n(\mathbb{C})$.

酉空间的讨论与欧氏空间的讨论很相似, 前面通过内积所做的关于欧氏空间的讨论, 都可以平行地在酉空间中进行. 这里强调一下不同.

(1) 内积的定义中, 代替“对称性”的是“共轭对称性”;

(2) 内积的定义中, 没有了双线性性; 因 $(\alpha, k\beta) = \overline{k}(\alpha, \beta)$;

(3) Cauchy-Schwarz 不等式仍成立, 但用内积写出来时要修改.

$$|(\alpha, \beta)| \leq |\alpha| |\beta|, \rightarrow (\alpha, \beta) \overline{(\alpha, \beta)} \leq (\alpha, \alpha) (\beta, \beta).$$

(4) 内积的度量矩阵 (不再是实对称阵) 而是埃尔米特 (Hermite) 阵. 设 $\alpha, \alpha, \dots, \alpha$ 为 $U_n(\mathbb{C})$ 中的一组基, 令 $G = ((\alpha, \alpha)) = (g_{ij})$, 则因 $(\alpha, \alpha) = \overline{(\alpha, \alpha)}$, 故 $\overline{G}^T = G$, 称为埃尔米特阵. 由于 $i = j$ 时, $g_{ii} = \overline{g_{ii}}$, 故埃尔米特阵的对角线上元素全为实数.

定义2 设 σ, τ 是 $U_n(\mathbb{C})$ 上的线性变换, 若对 " $\alpha, \beta \in U_n(\mathbb{C})$ ", 有

$$(\sigma\alpha, \beta) = (\alpha, \tau\beta),$$

则称 τ 是 σ 的伴随变换(或共轭变换), 记为 σ^* .

定理 1 设线性变换 σ 在 $U_n(\mathbb{C})$ 的标准正交基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 上的矩阵是 A , 则 σ 的伴随变换 σ^* 在同一组基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下的矩阵是 $\overline{A}^T$.

2 酉变换与酉矩阵的定义与主要性质

定义 3 若 n 阶矩阵 U 满足 $\overline{U}^T U = I$ (即 $\overline{U}^T = U^{-1}$), 则称 U 为酉矩阵.

定义 4 设 σ 是 $U_n(\mathbb{C})$ 上的一个线性变换, 若对任意 $\alpha, \beta \in U_n(\mathbb{C})$, 都有

$$(\sigma\alpha, \sigma\beta) = (\alpha, \beta),$$

则称 σ 是 $U_n(\mathbb{C})$ 上的一个酉变换.

定理 2 设 σ 是酉空间 $U_n(\mathbb{C})$ 上的线性变换, 则以下 4 个命题等价:

- (1) σ 是为酉变换;
- (2) " $\alpha \in U_n(\mathbb{C})$, 有 $\|\sigma\alpha\| = \|\alpha\|$;
- (3) σ 把标准正交基变为标准正交基;
- (4) σ 在任意一组标准正交基下的矩阵是酉矩阵.

酉矩阵有以下性质:

- (1) 行列式是模为 1 的复数, 即 $|\det U| = 1$.
- (2) 特征值是模为 1 的复数. 设 $\sigma\alpha = \lambda\alpha$, 则

$$(\alpha, \alpha) = (\sigma\alpha, \sigma\alpha) = (\lambda\alpha, \lambda\alpha) = \bar{\lambda}\lambda(\alpha, \alpha),$$

故 $\bar{\lambda}\lambda = 1$, 即 $|\lambda| = 1$.

- (3) 两个酉矩阵之积以及酉矩阵之逆仍为酉矩阵.

(4) 酉矩阵的行、列都是标准正交向量组, 即在内积 $(x, y) = x^T \overline{y}$ 下

$$\overline{U}^T U = I, \quad \text{得} \quad \sum_{k=1}^n \overline{u_{ki}} u_{kj} = \delta_{ij};$$

$$U \bar{U}^T = I, \quad \text{得} \quad \sum_{k=1}^n u_{ik} \bar{u}_{jk} = \delta_{ij}.$$

(5) 酉矩阵的属于不同特征值的特征向量是相互正交的.

3 规范变换与规范矩阵

定义 5 设 σ 是 $U_n(\mathbb{C})$ 上的线性变换, 若有

$$\sigma\sigma^* = \sigma^* \sigma,$$

则称 σ 是规范变换(正规变换).

定义 6 设 n 阶矩阵 A 满足 $A \bar{A}^T = \bar{A}^T A$, 则称 A 为规范矩阵.

命题 规范变换在标准正交基下的矩阵是规范矩阵.

定理 3 设 σ 是 $U_n(\mathbb{C})$ 上的规范变换, 若有 $\sigma\xi = \lambda\xi$, 则 $\sigma^*\xi = \bar{\lambda}\xi$.

定理 4 设 σ 是 $U_n(\mathbb{C})$ 上的规范变换, 则 σ 的属于不同特征值的特征向量相互正交.

定理 5 n 阶矩阵 A 可以酉相似于对角矩阵 A 是规范矩阵.

定理 6 两个可交换的规范矩阵, 可以同时酉对角化, 即当 $AB = BA$, 且 $\bar{A}^T A = A \bar{A}^T$, $\bar{B}^T B = B \bar{B}^T$ 时, 存在酉矩阵 U , 使

$$U^{-1}AU = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{pmatrix},$$

其中 n_i 为 A 的特征值 λ_i 的重数, $\lambda_i \neq \lambda_j (i \neq j)$, $\sum_{i=1}^s n_i = n$;

$$U^{-1}BU = \begin{pmatrix} \mu_{11} & & \\ & \ddots & \\ & & \mu_{n_1} \end{pmatrix} \oplus \begin{pmatrix} \mu_{21} & & \\ & \ddots & \\ & & \mu_{n_2} \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \mu_{s1} & & \\ & \ddots & \\ & & \mu_{n_s} \end{pmatrix},$$

其中 μ_{ij} 为 B 的特征值, $i = 1, \dots, s, j = 1, \dots, n_i$.

4 埃尔米特变换与埃尔米特矩阵、定义与主要定理、正定性

定义 7 若 n 阶矩阵 $H = (h_{ij})$ 满足 $\bar{H}^T = H$ (或 $h_{ij} = \bar{h}_{ji}$), 则称

H 为埃尔米特矩阵 .

定义 8 设 σ 为 $U_n(\mathbb{C})$ 上的线性变换, 若对 " $\alpha, \beta \in U_n(\mathbb{C})$, 有

$$(\alpha, \sigma\beta) = (\sigma\alpha, \beta),$$

则称 σ 为埃尔米特变换 .

定理 7 σ 为埃尔米特变换

σ 在标准正交基下的矩阵为埃尔米特矩阵 .

对埃尔米特矩阵和埃尔米特变换, 有:

(1) 特征值是实数 . 设 $\sigma\alpha = \lambda\alpha$, 由

$$(\sigma\alpha, \alpha) = (\alpha, \sigma\alpha), \text{ 得 } \lambda(\alpha, \alpha) = \bar{\lambda}(\alpha, \alpha).$$

因 $\alpha \neq 0$, 故 $\bar{\lambda} = \lambda$, 即知 λ 为实数 .

(2) 属于不同特征值的特征向量相互正交 .

(3) 对任意 $\alpha \in U_n(\mathbb{C})$, $(\sigma\alpha, \alpha)$ 是实数 . 由于 $(\sigma\alpha, \alpha) = (\alpha, \sigma\alpha) = \overline{(\sigma\alpha, \alpha)}$, 故得证 .

定义 9 n 个复变量 $x_1, \dots, x_n$ 的二次齐次函数

$$f = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \bar{x}_i x_j \quad (a_{ij} = \bar{a}_{ji}),$$

称为埃尔米特二次型 .

定义 10 设 σ 是 $U_n(\mathbb{C})$ 的埃尔米特变换, 若对 " $\alpha (\neq 0) \in U_n(\mathbb{C})$, 有 $(\sigma\alpha, \alpha) > 0$, 则称 σ 是正定埃尔米特变换 .

令 $A = (a_{ij})$, $a_{ij} = \bar{a}_{ji}$, $\mathbf{x} = (x_1, \dots, x_n)^T$,

则
$$f = \begin{pmatrix} \bar{x}_1 & \cdots & \bar{x}_n \end{pmatrix} \begin{matrix} A \\ \cdots \\ A \end{matrix} \begin{matrix} x_1 \\ \cdots \\ x_n \end{matrix} = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x}.$$

显然, $\bar{A}^T = A$. A 称为埃尔米特二次型的矩阵 .

设 σ 在 $U_n(\mathbb{C})$ 的标准正交基 $\varepsilon_1, \dots, \varepsilon_n$ 上的矩阵是 A , 任取 $\alpha \in U_n(\mathbb{C})$, 它的坐标记为 $\mathbf{x}$, 则有

$$(\sigma\alpha, \alpha) = \mathbf{x}^T A^T \bar{\mathbf{x}} = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x}.$$

这说明正定埃尔米特变换恰相当于说埃尔米特二次型正定 . 因由

上式知

$$(\alpha, \alpha) > 0 \quad \overline{x}^T A x > 0 .$$

这里矩阵 A 是线性变换 σ 的矩阵, 也是埃尔米特二次型的矩阵.

在实线性空间中, 我们讨论过正定二次型, 正定实对称矩阵, 当然也可以讨论正定实对称变换, 显然它在标准正交基下的矩阵是正定实对称矩阵.

在酉空间中关于正定性的讨论是以“变换的语言”来叙述的;

正定埃尔米特变换在标准正交基下的矩阵称为正定埃尔米特矩阵;

以正定埃尔米特矩阵为矩阵的埃尔米特二次型称为正定埃尔米特二次型.

定理 8 埃尔米特变换 σ 正定 σ 的特征值 > 0 .

又因 $\det A = \lambda_1 \lambda_2 \cdots \lambda_n > 0$, 故得到下面的结论.

推论 正定埃尔米特矩阵的行列式是正实数.

定理 9 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 $U_n(C)$ 的任意一组基, 记 $G = ((\alpha_i, \alpha_j))$, 则 $\det G > 0$.

定理 10 设 A 是埃尔米特矩阵, 则 A 正定 A 的顺序主子式 > 0 .

定理 11 设 A, B 是两个埃尔米特矩阵, 且 A 正定, 则存在可逆复矩阵 P , 使得

$$\overline{P}^T A P = I, \quad \overline{P}^T B P = \text{diag}(\alpha_1, \alpha_2, \cdots, \alpha_n).$$

7.2 典型题讲评

例 1 设 $A = \begin{pmatrix} 0 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$, 试求一个酉矩阵 U , 使得 $\overline{U}^T A U$

为上三角矩阵.

解 方法 1 首先求 A 的特征多项式, 因为 $p_A(\lambda) = |\lambda I - A| = \lambda^3 - 1$, 所以 $\lambda = 1$ 为 A 的实特征值. 对 $\lambda = 1$, 解得模为 1 的特征向量 $\mathbf{x} = \frac{1}{2}(1, 1, 0)^T$, 将其扩充为 C^3 的标准正交基, 取

$$\mathbf{x}_2 = \frac{1}{2}(1, -1, 0)^T, \mathbf{x}_3 = (0, 0, 1)^T.$$

令 $U_1 = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, 则 U_1 是酉矩阵, 且有

$$\overline{U_1}^T A U_1 = \begin{pmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & -1 & -\frac{1}{2} \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ \mathbf{0} & A \end{pmatrix},$$

其中 $\alpha = 0, -\frac{3}{2}$, A 为右下角二阶矩阵, 对 A 解得属于特征值

$\lambda_k = \frac{-1 + 3i}{2}$ 的特征向量 $\mathbf{y}_1 = -\frac{1}{3}, \frac{1+i}{6}$, 将其扩充为 C^2 的标准正交基, 得

$$\mathbf{y}_1 = \begin{pmatrix} -\frac{1}{3} \\ \frac{1+i}{6} \end{pmatrix}, \mathbf{y}_2 = \begin{pmatrix} \frac{1-i}{6} \\ \frac{1}{3} \end{pmatrix}.$$

令 $U_2 = (\mathbf{y}_1, \mathbf{y}_2)$, 则 U_2 是酉矩阵, 且有

$$\overline{U_2}^T A_2 U_2 = \begin{pmatrix} \frac{-1+i}{2} & -\frac{3}{6}i \\ 0 & \frac{-1-i}{2} \end{pmatrix}.$$

$$U = \begin{pmatrix} \mathbf{U}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{6} & \frac{1-i}{2} \\ \frac{1}{2} & \frac{1}{6} & \frac{-1+i}{2} \\ 0 & \frac{1+i}{6} & \frac{1}{3} \end{pmatrix},$$

则得

$$U^{-1}AU = \begin{pmatrix} 1 & \frac{-3-3i}{2} & -\frac{3}{6} \\ 0 & \frac{-1+i}{2} & -\frac{3i}{6} \\ 0 & 0 & \frac{-1-i}{2} \end{pmatrix} \quad (U^{-1} = \overline{U}^T).$$

方法2 由 $p_A(\lambda) = \lambda^3 - 1 = 0$, 得

$$\lambda_1 = 1, \quad \lambda_2 = \frac{-1+i}{2}, \quad \lambda_3 = \frac{-1-i}{2}.$$

对 λ , 分别解 $(\lambda I - A)\mathbf{x} = \mathbf{0}$, 得各相应的特征向量: 取

$$\mathbf{x}_1 = (1, 1, 0)^T;$$

$$\mathbf{x}_2 = \frac{1+i}{2}, 1, 1^T;$$

$$\mathbf{x}_3 = \frac{1-i}{2}, 1, 1^T.$$

令 $P = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, 则有

$$P^{-1}AP = \text{diag} \left(1, \frac{-1+i}{2}, \frac{-1-i}{2} \right),$$

又由对复矩阵 P 有 U, T 分解, 即

$$P = UT,$$

其中 P 可逆, U 为酉矩阵, T 为对角线元素为正数的上三角矩阵.

所以

$$(UT)^{-1}A(UT) = D, \text{ 即 } U^{-1}AU = TDT^{-1}.$$

因为上三角矩阵的逆为上三角矩阵, 上三角矩阵之积为上三角矩阵, 故 TDT^{-1} 为上三角矩阵, 且因 U 为酉矩阵, 故 $U^{-1} = \overline{U}^T$.

此方法要对 P 的列进行施密特正交化, 故计算量较大, 书写如下:

$$\mathbf{e}_1 = \frac{1}{2}\mathbf{x}_1, \text{ 记 } \boldsymbol{\beta}_1 = \mathbf{x}_1, \text{ 则 } \mathbf{e}_1 = \frac{1}{\|\boldsymbol{\beta}_1\|}\boldsymbol{\beta}_1,$$

$$\begin{aligned}\boldsymbol{\beta}_2 &= \mathbf{x}_2 - (\mathbf{x}_2, \mathbf{e}_1)\mathbf{e}_1 = \mathbf{x}_2 - \frac{3+i}{2} \frac{3}{2} \mathbf{e}_1 \\ &= \frac{-1+i}{4} \frac{3}{2}, \frac{1-i}{4} \frac{3}{2}, 1^T,\end{aligned}$$

$$\|\boldsymbol{\beta}_2\| = \frac{3}{2}, \text{ 所以 } \mathbf{e}_2 = \frac{1}{\|\boldsymbol{\beta}_2\|}\boldsymbol{\beta}_2,$$

$$\begin{aligned}\boldsymbol{\beta}_3 &= \mathbf{x}_3 - (\mathbf{x}_3, \mathbf{e}_1)\mathbf{e}_1 - (\mathbf{x}_3, \mathbf{e}_2)\mathbf{e}_2 \\ &= \mathbf{x}_3 - \frac{3-i}{2} \frac{3}{2} \mathbf{e}_1 - \frac{3+i}{2} \frac{3}{2} \mathbf{e}_2 \\ &= -\frac{i}{3}, \frac{i}{3}, \frac{3-i}{2}^T,\end{aligned}$$

$$\|\boldsymbol{\beta}_3\| = 1, \text{ 所以 } \mathbf{e}_3 = \boldsymbol{\beta}_3.$$

反解之, 得

$$\mathbf{x}_1 = \|\boldsymbol{\beta}_1\| \mathbf{e}_1 = 2\mathbf{e}_1,$$

$$\mathbf{x}_2 = \|\boldsymbol{\beta}_2\| \mathbf{e}_2 + (\mathbf{x}_2, \mathbf{e}_1)\mathbf{e}_1 = \frac{3}{2}\mathbf{e}_2 + \frac{3+i}{2} \frac{3}{2} \mathbf{e}_1,$$

$$\mathbf{x}_3 = \|\boldsymbol{\beta}_3\| \mathbf{e}_3 + (\mathbf{x}_3, \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x}_3, \mathbf{e}_2)\mathbf{e}_2 = \mathbf{e}_3 + \frac{3-i}{2} \frac{3}{2} \mathbf{e}_1 + \frac{3+i}{2} \frac{3}{2} \mathbf{e}_2,$$

即有

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} e \\ e \\ e \end{pmatrix} = \begin{pmatrix} 2 & \frac{3+i}{2} & \frac{3-i}{2} \\ 0 & \frac{3}{2} & \frac{3+i}{2} \\ 0 & 0 & 1 \end{pmatrix} = UT,$$

而

$$U = \begin{pmatrix} e \\ e \\ e \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{-1+i}{2} & -\frac{i}{3} \\ \frac{1}{2} & \frac{1-i}{2} & \frac{i}{3} \\ 0 & \frac{2}{6} & \frac{3-i}{2} \end{pmatrix},$$

使得 $\overline{U}^T A U$ 为上三角矩阵.

讲评 (1) 此题考察的理论依据是 5.1 节定理 9 推论, 我们在 5.2 节例 15 中已给出证明. 本题是一个具体实践, 故从这个本意上来看, 提倡用方法 1. 另外, 虽然 A 有 3 个不同的特征值 (从而有 3 个线性无关的特征向量), 故 A 相似于对角矩阵. 对角矩阵当然也包含在上角矩阵范围之内, 但看清题目要求, 还有一条“求一个酉矩阵”, 即找的特征向量线性无关还不行, 还要求它们是正交的. 所以若走求特征值、特征向量的路, 而算出的特征向量不正交的话, 就还得施密特正交化, 即如方法 2, 事实上计算量会比一步步求上三角矩阵还要大.

(2) 由本章定理 5 知, 当且仅当 A 为规范矩阵, A 酉相似于对角矩阵. 即如果你先验证出有 $A \overline{A}^T = \overline{A}^T A$, 则用方法 2 求得的 P 可以证它是酉矩阵, 因为规范矩阵属于不同特征值的特征向量相互正交, 故只要把各特征向量单位化一下即得酉矩阵 U .

(3) 两种方法算出的 U 不同, 说明 U 是不惟一. 当然上三角

矩阵也不惟一.

例 2 设 A, B 是正定埃尔米特矩阵, 若 AB 是埃尔米特矩阵, 证明 AB 正定.

证 方法 1 由已知 AB 是埃尔米特矩阵, 故

$$AB = (\overline{AB})^T = \overline{B}^T \overline{A}^T = BA,$$

即 A, B 乘法可交换. 由本章定理 6 知, 存在酉矩阵 U , 使

$$\overline{U}^T AU = \text{diag}\{\lambda_1, \cdots, \lambda_n\},$$

$$\overline{U}^T BU = \text{diag}\{\mu_1, \cdots, \mu_n\},$$

则有

$$\begin{aligned}\overline{U}^T ABU &= \overline{U}^T AU \overline{U}^T BU, \\ &= \text{diag}\{\lambda_1 \mu_1, \cdots, \lambda_n \mu_n\},\end{aligned}$$

因为 $\lambda_i \mu_i > 0 (i = 1, \cdots, n)$, 所以 AB 正定.

方法 2 因为 A, B 正定, 所以存在可逆矩阵 P, Q 使 $A = \overline{P}^T P, B = \overline{Q}^T Q$, 则

$$AB = \overline{P}^T P \overline{Q}^T Q = \overline{P}^T P \overline{Q}^T Q \overline{P}^T (\overline{P}^T)^{-1},$$

故知 AB 与 $P \overline{Q}^T Q \overline{P}^T$ 相似. 而 $P \overline{Q}^T Q \overline{P}^T = (\overline{Q \overline{P}^T})^T Q \overline{P}^T$ 是正定埃尔米特矩阵, 故其特征值全为正, 于是由相似的矩阵有相同的特征值得知: AB 的特征值皆为正. 又已知 AB 为埃尔米特矩阵, 所以 AB 正定.

方法 3 设 A, B 分别为 $U_n(C)$ 中的线性变换 A, B 在标准正交基下的矩阵, 则 A, B, AB 均为埃尔米特变换, 由

$$AB = (AB)^* = B^* A^* = BA,$$

知 A, B 乘法可交换. 故必存在 $U_n(C)$ 的标准正交基 $\eta_1, \eta_2, \cdots, \eta_n$ 使得

$$A\eta_i = \lambda_i \eta_i, \quad B\eta_i = \mu_i \eta_i, \quad i = 1, \cdots, n.$$

于是, 对任意 $\alpha \in V$, 设 $\alpha = \sum_{i=1}^n x_i \eta_i$, 有

$$(AB\alpha, \alpha) = AB \sum_{i=1}^n x_i \eta_i, \sum_{j=1}^n x_j \eta_j = \sum_{i=1}^n x_i \lambda_i \mu_i \eta_i, \sum_{j=1}^n x_j \eta_j$$

$$= \sum_{i=1}^n x_i \sum_{j=1}^n \overline{x_j} \lambda_i \mu_j (\eta_i, \eta_j) = \sum_{i=1}^n x_i \overline{x_i} \lambda_i \mu_i > 0$$

(这里用到因 A, B 正定, 所以 λ, μ 均为正) 故知 AB 为正定埃尔米特变换, 故 AB 是正定埃尔米特矩阵.

讲评 (1) 方法 3 与方法 1 用的理论是一个, 只不过一个用“变换的语言”, 一个用“矩阵的语言”叙述, 即“当 $AB = BA$ 时, A, B 可同时对角化”. 如果没有这一条, 方法 3 就证不出来. 换句话说, 如果你的证明中没用到这一条件, 肯定哪儿出了问题.

如有的同学这样写:

“任取 $\alpha \in V$, 有

$$\begin{aligned} (AB\alpha, \alpha) &= (A\mu\alpha, \alpha) = \mu(A\alpha, \alpha) \\ &= \lambda\mu(\alpha, \alpha). \end{aligned}$$

因 $(\alpha, \alpha) > 0, \lambda > 0, \mu > 0$, 故得

$$(AB\alpha, \alpha) > 0,$$

所以, AB 正定.”

这事实上假定了空间 $U_n(\mathbb{C})$ 中所有的向量都是 A, B, AB 的特征向量. 这显然是加了条件, 我们说:

① 方法 3 利用 $AB = BA$ 这一条件说明空间中可以找到一组基, 即 A 和 B 的公共的特征向量, 从而也是 AB 的特征向量;

② 想一想, 若任意 α 都是 A, B, AB 的特征向量, 这 3 个变换都是什么变换? 它们恰是

$$A = \lambda\varepsilon, \quad B = \mu\varepsilon, \quad AB = \lambda\mu\varepsilon.$$

这改动是不是太大了?!

例 3 设 A 是正定的埃尔米特矩阵, 证明若 A 是酉矩阵, 则 $A = I$.

证 因埃尔米特矩阵的特征值都是实数 (证明见 7.1 节), 且由 A 是埃尔米特矩阵, 知有 $\overline{A^T} = A$, 又由 A 是酉矩阵, 故有 $\overline{A^T} A = I$, 于是得

$$\overline{A^T} A = A^2 = I,$$

故得 $\lambda^2 = 1$, 即 $\lambda = \pm 1$ (由 $Ax = \lambda x$, 知 $Ix = A^2 x = A\lambda x = \lambda^2 x$, 故 $\lambda^2 = 1$). 由题设知 A 正定, 所以 $\lambda = 1$.

因为埃尔米特矩阵可酉相似于对角矩阵, 即存在酉矩阵 U , 使得

$$\overline{U}^T A U = I, \text{ 故 } A = U \overline{U}^T = I.$$

讲评 此题中条件给得很充分(强), 故证明并不难. 比如也可这样叙述: 由 A 是酉矩阵, 知其特征值模为 1; 又 A 是正定的埃尔米特矩阵, 特征值应为正实数, 故 A 的特征值全为 1. 由本章定理 5 故存在酉矩阵, 使

$$\overline{U}^T A U = I, \text{ 即 } A = I.$$

例 4 设 $A \in M_{m,n}(C)$, 证明 $\overline{A}^T A$ 和 $A \overline{A}^T$ 都是半正定埃尔米特矩阵.

证 因 $(\overline{A}^T A)^H = (\overline{A}^T A)$, 故 $\overline{A}^T A$ 和 $A \overline{A}^T$ 都是埃尔米特矩阵. (用定义证), 对任意 $x \in C$, 有

$$\overline{x}^T \overline{A}^T A x = (\overline{Ax})^T (Ax) \geq 0,$$

所以 $\overline{A}^T A$ 半正定 (当 $x \neq 0$ 时, $Ax \neq 0$ A 可逆), 故当 A 可逆时, $\overline{A}^T A$ 是正定的. 同理可证 $A \overline{A}^T$ 是半正定的.

讲评 事实上, 我们有定理: 埃尔米特矩阵 B 半正定的充要条件为存在复矩阵 A , 使得 $B = \overline{A}^T A$. 本题恰为该定理的充分性. 定理的必要性证明如下: 因 B 是埃尔米特矩阵, 由本章定理 5, 知存在酉矩阵 U , 使得

$$\begin{aligned} B &= \overline{U}^T \text{diag}\{\lambda_1, \dots, \lambda_n\} U \quad (\text{其中 } \lambda_i \geq 0). \\ &= \overline{U}^T \text{diag}\{\lambda_1, \dots, \lambda_n\} U \overline{U}^T \text{diag}\{\lambda_1, \dots, \lambda_n\} U \\ &= \overline{A}^T A, \end{aligned}$$

其中 $A = \overline{U}^T \text{diag}\{\lambda_1, \dots, \lambda_n\} U$.

例 5 设 V 是酉空间, H 是 V 上埃尔米特变换 (即自伴随变换). 证明:

(1) $\| \alpha + iH\alpha \| = \| \alpha - iH\alpha \|$ (对任意 $\alpha \in V$);

(2) $\alpha + iH\alpha = \beta + iH\beta$ 当且仅当 $\alpha = \beta$;

(3) $\varepsilon + iH$ 非奇异;

(4) $\varepsilon - iH$ 非奇异;

(5) $U = (\varepsilon - iH)(\varepsilon + iH)^{-1}$ 是酉变换(此处设 V 是有限维) .

证 (1) $\| \alpha + iH\alpha \|^2 = (\alpha + iH\alpha, \alpha + iH\alpha) = (\alpha, \alpha) + (H\alpha, H\alpha) - i(H\alpha, \alpha) + i(\alpha, H\alpha) = \| \alpha - iH\alpha \|^2$;

(2) 充分性显然 .

由已知等式得 $iH(\alpha - \beta) = -(\alpha - \beta)$, 故 $H(\alpha - \beta) = i(\alpha - \beta)$, 由 H 的特征值为实数, 所以 $\alpha - \beta = 0$, 即 $\alpha = \beta$ (事实上, 若 $\alpha \neq \beta$, 则 $\alpha - \beta$ 是 H 的属于特征值 i 的特征向量, 与 H 的特征值为实数矛盾).

(3) 因为 H 的特征值均为实数, 所以 $\det(I + iH) = \det i \left(\frac{1}{i}I + H \right) \neq 0$, 故 $I + iH$ 非奇异(这里 H 为 H 在 V 中标准正交基下的矩阵表示, H 是埃尔米特矩阵).

(4) 同(3), 因为 $|I - iH| = \left| i \left(\frac{1}{i}I - H \right) \right|$, $\frac{1}{i}$ 不是 H 的特征值, 所以 $\left| \frac{1}{i}I - H \right| \neq 0$.

(5) 因为 $(\varepsilon + iH)(\varepsilon - iH) = (\varepsilon - iH)(\varepsilon + iH)$, 前后均乘 $(\varepsilon - iH)^{-1}$ 得

$$(\varepsilon - iH)^{-1}(\varepsilon + iH) = (\varepsilon + iH)(\varepsilon - iH)^{-1} .$$

于是由

$$\begin{aligned} U^* &= ((\varepsilon + iH)^{-1})^* (\varepsilon - iH)^* = (\varepsilon^* - iH^*)^{-1} (\varepsilon + iH^*) \\ &= (\varepsilon - iH)^{-1} (\varepsilon + iH), \end{aligned}$$

利用前式可得 $U^* = (\varepsilon + iH)(\varepsilon - iH)^{-1}$,

所以

$$U^* U = U U^* = \varepsilon.$$

故知 U 是酉变换.

讲评 这组小题目的证明主要用了以下这些知识:

(1) 埃尔米特变换的定义: 对任意 $\alpha \in V$, 若有 $(H\alpha, \alpha) = (\alpha, H\alpha)$, 则称 H 为埃尔米特变换.

(2) 埃尔米特变换的特征值均为实数. 如把它用到问题(2)的证明中, 从式

$$H(\alpha - \beta) = i(\alpha - \beta),$$

知 $\alpha - \beta = 0$; 用到问题(3)的证明中, 知 $|I + iH| = \left| i \frac{1}{i} I + H \right| \neq 0$

因 $-\frac{1}{i}$ 不是 H 的特征值; 用到问题(4)中, $\frac{1}{i}$ 不是 H 的特征值,

故 $\left| \frac{1}{i} I - A \right| \neq 0$ 等等.

(3) 酉变换的定义: 满足 $U^* U = \varepsilon$ 或 $U U^* = \varepsilon$ 的 U 称为酉变换. 问题(5)就是用此定义证出, 当然直接按定义乘, 还不是所要的结论, 为此根据所需, 中间把 U^* 做了一次恒等变形.

例 6 设 θ 为实数, 证明以下矩阵酉相似

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad D = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

$$\text{解 } |\lambda I - A| = \begin{vmatrix} \lambda - \cos\theta & \sin\theta \\ -\sin\theta & \lambda - \cos\theta \end{vmatrix} = \lambda^2 - 2\lambda\cos\theta + 1,$$

所以

$$\lambda = \cos\theta \pm i \sin\theta = e^{\pm i\theta}.$$

对 $\lambda = \cos\theta + i \sin\theta$, 解得全部特征向量为 $x_1 = c_1 \begin{pmatrix} i \\ 1 \end{pmatrix}$, 取

$e = \frac{1}{2} \begin{pmatrix} i \\ 1 \end{pmatrix}$; 对 $\lambda = \cos\theta - i \sin\theta$, 解得特征向量为 $x = c_2 \begin{pmatrix} 1 \\ i \end{pmatrix}$, 取

$$\mathbf{e} = \frac{1}{2} \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

$$\text{令} \quad U = (\mathbf{e}, \mathbf{e}) = \frac{1}{2} \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix},$$

则 U 为酉矩阵, 且有

$$U^{-1} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} U = \begin{pmatrix} e^{i\theta} & \\ & e^{-i\theta} \end{pmatrix}.$$

讲评 事实上, 这里的矩阵 A 是一个实正交矩阵 (它也是一个规范矩阵), 因满足 $A^T A = AA^T = I$.

由本题的证明, 同学们可以看到实正交矩阵不一定实相似于对角矩阵, 因为它的特征值除了可以为 ± 1 (实根) 还可以是模为 1 的复数 (形如 $e^{i\theta}$). 当实正交矩阵有复特征值时, 它只能酉相似于对角矩阵 D . 若要在实数域中求它的相似标准形, 只能是

$$A \sim \text{diag } 1, \cdots, 1, -1, \cdots, -1 \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \cdots, \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

即共轭复根部分提供的是二阶子矩阵, 此时空间分解为一维和二维不变子空间的直和.

例 7 证明本章定理 5, 即 n 阶矩阵 A 可以酉相似于对角矩阵 A 是规范矩阵.

证 (必要性) 设存在酉矩阵 U , 使得

$$U^{-1} A U = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = W, \quad U^{-1} = \overline{U}^T,$$

则

$$A = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^{-1}, \quad \overline{A}^T = U \begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix} U^{-1},$$

所以

$$\begin{aligned} & \lambda \bar{\lambda} \\ & A \bar{A}^T = U \quad W \quad U^{-1} = \bar{A}^T A. \\ & \lambda \bar{\lambda}_n \end{aligned}$$

(充分性)对 A 的阶数用归纳法.当 $n=1$ 时,结论显然成立.今设 A 为 $n-1$ 阶时结论成立,则当 A 为 n 阶矩阵时,设 $Ax_1 = \lambda x_1$ ($\|x_1\|=1$),可扩充 $x_1, \dots, x_n$, 使

$$U_n = (x_1, x_2, \dots, x_n)$$

为酉矩阵,则有

$$U_n^{-1} A U_n = \begin{pmatrix} \lambda & c \\ 0 & A_{n-1} \end{pmatrix}, \quad \text{其中 } c \in \mathbb{C}^{1 \times (n-1)}.$$

于是得

$$\begin{aligned} A &= U_n \begin{pmatrix} \lambda & c \\ 0 & A_{n-1} \end{pmatrix} U_n^{-1}, \quad \bar{A}^T = U_n \begin{pmatrix} \bar{\lambda} & 0 \\ \bar{c}^T & \bar{A}_{n-1}^T \end{pmatrix} U_n^{-1}, \\ A \bar{A}^T &= U_n \begin{pmatrix} \lambda \bar{\lambda} + c \bar{c}^T & c \bar{A}_{n-1}^T \\ A_{n-1} \bar{c}^T & A_{n-1} \bar{A}_{n-1}^T \end{pmatrix} U_n^{-1}, \\ \bar{A}^T A &= U_n \begin{pmatrix} \bar{\lambda} \lambda & \lambda c \\ \bar{c}^T \lambda & \bar{c}^T c + \bar{A}_{n-1}^T A_{n-1} \end{pmatrix} U_n^{-1}. \end{aligned}$$

由 $A \bar{A}^T = \bar{A}^T A$, 得

$$\begin{aligned} \lambda \bar{\lambda} + c \bar{c}^T &= \bar{\lambda} \lambda, & \text{即 } c \bar{c}^T &= 0, \\ A_{n-1} \bar{A}_{n-1}^T &= \bar{c}^T c + \bar{A}_{n-1}^T A_{n-1}, & A_{n-1} \bar{A}_{n-1}^T &= \bar{A}_{n-1}^T A_{n-1}. \end{aligned}$$

所以 A_{n-1} 为规范矩阵.由归纳假设存在 $n-1$ 阶酉矩阵 U_{n-1} , 使

$$\begin{aligned} & \lambda \\ & U_{n-1}^{-1} A_{n-1} U_{n-1} = \quad W \quad \\ & \lambda \end{aligned}$$

$$\text{令 } U = U_n \begin{pmatrix} 1 & 0 \\ 0 & U_{n-1} \end{pmatrix}, \quad \text{则 } U^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & U_{n-1}^{-1} \end{pmatrix} U_n^{-1}.$$

因为两个酉矩阵之积仍为酉矩阵,故 U 为酉矩阵,且有

$$\begin{aligned}
 U^{-1}AU &= \begin{pmatrix} 1 & 0 \\ 0 & U_{n-1}^{-1} \end{pmatrix} \begin{pmatrix} U_n^{-1}AU_n & 0 \\ 0 & U_{n-1} \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & \lambda & 0 \\ 0 & U_{n-1}^{-1} & 0 & A_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & U_{n-1} \end{pmatrix} \\
 &= \begin{pmatrix} \lambda & 0 \\ 0 & U_{n-1}^{-1}A_{n-1}U_{n-1} \end{pmatrix} = \begin{pmatrix} \lambda & & \\ & W & \\ & & \lambda \end{pmatrix}.
 \end{aligned}$$

讲评 (1) 这个定理揭示了规范矩阵的本质: 即可以酉相似于对角矩阵的矩阵. 也可以这样认为: 前人恰是把“可以酉相似于对角矩阵的所有矩阵”称为规范矩阵, 而后发现它们都满足 $AA^T = A^T A$ (见证明的必要性). 现在则拿后者做规范矩阵的定义 (你会觉得既抽象又神秘, 他怎么会想到这个呢? 实际上, 可对角化这条性质在起决定作用).

(2) 实对称矩阵、实正交矩阵、酉矩阵, 埃尔米特矩阵都是规范矩阵, 那么规范矩阵是否仅由这 4 种特殊矩阵组成? 回答是否定的. 如, 形如

$$\overline{U}^T \begin{pmatrix} 3+5i & 0 \\ 0 & 2-i \end{pmatrix} U \quad (\text{其中 } U \text{ 为酉矩阵})$$

的矩阵是规范矩阵, 但它不是那 4 种矩阵之一. 理解了那 4 种矩阵特征值的特点, 和酉相似对角矩阵的本质, 你也可以随手举出一个例子来 (注意实对称矩阵和埃尔米特矩阵的特征值为实数, 实正交矩阵与酉矩阵的特征值是模为 1 的复数).

例 8 证明本章定理 6, 即两个可交换的规范矩阵, 可以同时酉对角化.

证 设 A, B 为 n 阶规范阵, 且有 $AB = BA$. 因 A 为规范矩阵, 故存在 n 阶酉矩阵 T , 使

$$T^{-1}AT = \begin{matrix} \lambda_1 I_{n_1} & & \\ & \lambda_k I_{n_k} & \\ & & W \\ & & & \lambda_s I_{n_s} \end{matrix}, \quad \lambda_i \neq \lambda_j (i \neq j), i, j = 1, \cdots, s.$$

由 $AB = BA$, 得

$$(T^{-1}AT)T^{-1}BT = T^{-1}BT(T^{-1}AT). \quad (1)$$

设 $T^{-1}BT = (B_{ij})$, 则由上式, 知有

$$\lambda_i (B_{ij}) = \lambda_j (B_{ji}), \quad i, j = 1, \cdots, n.$$

因当 $i \neq j$ 时, $\lambda_i \neq \lambda_j$, 故有 $B_{ji} = 0 (i \neq j)$, 即

$$T^{-1}BT = \begin{matrix} B_1 & & \\ & B_2 & \\ & & W \\ & & & B_s \end{matrix},$$

其中 B_i 为 n_i 阶矩阵, 又由 $\overline{B}^T B = B \overline{B}^T$, 知

$$\overline{B}_{ii}^T B_{ii} = B_{ii} \overline{B}_{ii}^T \quad (i = 1, \cdots, s), \quad (2)$$

即 $B_{ii} (i = 1, \cdots, s)$ 是规范矩阵, 故存在 n_i 阶酉矩阵 Q_i 使

$$Q_i^{-1} B_{ii} Q_i = \text{diag}\{\mu_1, \cdots, \mu_{n_i}\}.$$

取 $Q = \begin{matrix} Q_1 & & \\ & W & \\ & & Q_s \end{matrix}$, 它为酉矩阵, 令 $U = TQ$, U 亦为酉矩阵,

则有

$$\begin{aligned} U^{-1}AU &= Q^{-1} \text{diag}\{\lambda_1 I_{n_1}, \cdots, \lambda_s I_{n_s}\} Q \\ &= \text{diag}\{\lambda_1 I_{n_1}, \cdots, \lambda_s I_{n_s}\}; \\ U^{-1}BU &= Q^{-1} (T^{-1}BT) Q = Q^{-1} \text{diag}(B_1, \cdots, B_s) Q \\ &= \text{diag}\{\mu_1, \cdots, \mu_{n_1}, \cdots, \mu_{s1}, \cdots, \mu_{n_s}\}. \end{aligned}$$

讲评 (1) 这里说的酉对角化是指相似于对角矩阵, 且变换矩阵是酉矩阵. 同时是指找到公共的 U , 使 $U^{-1}AU$ 和 $U^{-1}BU$ 都是

对角矩阵.

(2) 证明中有两步是关键: ①由(1)式知 $T^{-1}BT$ 与对角矩阵 $T^{-1}AT$ 可交换. 马上可联想到矩阵乘法中的习题: 与 $T^{-1}AT$ 这样的对角矩阵可交换的矩阵必为准对角矩阵, 其分块与 $T^{-1}AT$ 的分块同(同时, 也可联想到另一个题: 与对角线上元素不相等的对角矩阵乘法可交换的矩阵必为对角矩阵). 事实上, 我们在解答中给出了证明. ②(2)式的得出是另一个关键, 这只要在前一个式子中把 B 解出并代入 $\overline{B}^T B = B \overline{B}^T$ 中即得. 于是在 $T^{-1}AT$ 为对角矩阵的前提下, 可继续把每个 B_{ii} 化为对角矩阵(且保持 $T^{-1}AT$ 的对角矩阵不变).

(3) 当把定理中的规范矩阵改为埃尔米特矩阵时, 有命题: 设 A, B 是两个 n 阶埃尔米特矩阵, 且满足 $AB = BA$, 则存在酉矩阵 U , 使得

$$\begin{aligned}\overline{U}^T AU &= \text{diag}\{\lambda_1, \dots, \lambda_n\}, \\ \overline{U}^T BU &= \text{diag}\{\mu_1, \dots, \mu_n\}.\end{aligned}$$

注意 此命题不同于本章定理 11, 从本质上说这里是相似对角化, 对角线上分别是 A, B 的特征值, 那里是相合对角化, 由 P 的不同, α 也会变化.

例 9 设 σ 是酉空间上的线性变换, $\varepsilon_1, \dots, \varepsilon_n$ 是 $U_n(\mathbb{C})$ 的标准正交基, 若对 $\alpha \in U_n(\mathbb{C})$, 有 $\|\sigma\alpha\| = \|\alpha\|$. 试证: $\sigma\varepsilon_1, \dots, \sigma\varepsilon_n$ 也是 $U_n(\mathbb{C})$ 的标准正交基.

证 取 $\alpha = \varepsilon_i + \varepsilon_j$, 由

$$(\sigma(\varepsilon_i + \varepsilon_j), \sigma(\varepsilon_i + \varepsilon_j)) = (\varepsilon_i + \varepsilon_j, \varepsilon_i + \varepsilon_j),$$

$$\begin{aligned}\text{得} \quad & (\sigma\varepsilon_i, \sigma\varepsilon_i) + (\sigma\varepsilon_i, \sigma\varepsilon_j) + (\sigma\varepsilon_j, \sigma\varepsilon_i) + (\sigma\varepsilon_j, \sigma\varepsilon_j) \\ &= (\varepsilon_i, \varepsilon_i) + (\varepsilon_i, \varepsilon_j) + (\varepsilon_j, \varepsilon_i) + (\varepsilon_j, \varepsilon_j).\end{aligned}$$

又因 $(\sigma\varepsilon_i, \sigma\varepsilon_i) = (\varepsilon_i, \varepsilon_i)$, $(\sigma\varepsilon_j, \sigma\varepsilon_j) = (\varepsilon_j, \varepsilon_j)$, 故有

$$\text{Re}(\sigma\varepsilon_i, \sigma\varepsilon_j) = \text{Re}(\varepsilon_i, \varepsilon_j).$$

再取 $\alpha = \varepsilon_i + i\varepsilon_j$, 由

$$(\sigma(\varepsilon + i\varepsilon_j), \sigma(\varepsilon + i\varepsilon_j)) = (\varepsilon + i\varepsilon_j, \varepsilon + i\varepsilon_j)$$

得

$$\begin{aligned} & (\sigma\varepsilon, \sigma\varepsilon) - i(\sigma\varepsilon, \sigma\varepsilon_j) + i(\sigma\varepsilon_j, \sigma\varepsilon) + (\sigma\varepsilon_j, \sigma\varepsilon_j) \\ &= (\varepsilon, \varepsilon) - i(\varepsilon, \varepsilon_j) + i(\varepsilon_j, \varepsilon) + (\varepsilon_j, \varepsilon_j) - i(\sigma\varepsilon, \sigma\varepsilon_j) + i(\sigma\varepsilon_j, \sigma\varepsilon) \\ &= -i(\varepsilon, \varepsilon_j) + i(\varepsilon_j, \varepsilon), \end{aligned}$$

$$\operatorname{Im}(\sigma\varepsilon, \sigma\varepsilon_j) = \operatorname{Im}(\varepsilon, \varepsilon_j).$$

于是有

$$(\sigma\varepsilon, \sigma\varepsilon_j) = (\varepsilon, \varepsilon_j) = \delta_{ij}.$$

故知 $\sigma\varepsilon, \dots, \sigma\varepsilon_n$ 亦是 $U_n(\mathbb{C})$ 的标准正交基.

讲评 (1) 在这里不同于欧几里得空间的是 $(\sigma\varepsilon_i, \sigma\varepsilon_i) = \overline{(\sigma\varepsilon_i, \sigma\varepsilon_i)}$, 故 $(\sigma\varepsilon_i, \sigma\varepsilon_j) + (\sigma\varepsilon_j, \sigma\varepsilon_i) = 2\operatorname{Re}(\sigma\varepsilon_i, \sigma\varepsilon_j)$. 所以第一步取 $\alpha = \varepsilon + \varepsilon_j$ 时, 只能得到 $(\sigma\varepsilon_i, \sigma\varepsilon_j)$ 与 $(\varepsilon, \varepsilon_j)$ 的实部相等, 而不是它们相等. 因此还必须选新的向量 α , 以利用 $\|\sigma\alpha\| = \|\alpha\|$ 证出前面关系式的虚部相等.

(2) 上述证明同样适合于证: σ 是酉变换 是对任意 $\alpha \in V$ 有 $\|\sigma\alpha\| = \|\alpha\|$ 的充分性的证明. 因为当有了 $(\sigma\varepsilon_i, \sigma\varepsilon_j) = (\varepsilon, \varepsilon_j)$ 后, 两边乘 $\overline{x_i y_j}$, 再求和, 得

$$\sum_{i=1}^n \sum_{j=1}^n \overline{x_i y_j} (\sigma\varepsilon_i, \sigma\varepsilon_j) = \sum_{i=1}^n \sum_{j=1}^n \overline{x_i y_j} (\varepsilon, \varepsilon_j),$$

即对任意 $\alpha, \beta \in V$, 有 $(\sigma\alpha, \sigma\beta) = (\alpha, \beta)$, 故 σ 为酉变换.

例 10 试证: 酉矩阵的特征值是模为 1 的复数.

证 设 $\sigma\alpha = \lambda\alpha$, 则有

$$(\alpha, \alpha) = (\sigma\alpha, \sigma\alpha) = (\lambda\alpha, \lambda\alpha) = \overline{\lambda}\lambda(\alpha, \alpha).$$

因 $(\alpha, \alpha) \neq 0$, 故有 $\overline{\lambda}\lambda = 1$, 即 $|\lambda| = 1$.

讲评 由上述证明, 知正交矩阵的特征值也是模为 1 的复数. 因为在不知道正交矩阵的特征值为实数时, $(\lambda\alpha, \lambda\alpha) \neq \lambda^2(\alpha, \alpha)$ 只能按复的处理, 只有在知其特征值为实的时, 才能用欧几里得空间的内积来算. 事实上, 如 7.2 节例 6, 实正交矩阵的特征值确实可

以为复数 .

例 11 证明: 酉矩阵的属于不同特征值的特征向量是相互正交的 .

证 设 $\sigma\xi = \lambda \xi, \sigma\xi = \lambda \xi, \lambda \neq \lambda$, 则有

$$(\xi, \xi) = (\sigma\xi, \sigma\xi) = (\lambda \xi, \lambda \xi) = \lambda \bar{\lambda} (\xi, \xi).$$

因 $\lambda \neq \lambda$, 故 $\lambda \bar{\lambda} \neq 1$, 所以必 $(\xi, \xi) = 0$.

讲评 (1) 设 $\lambda = e^{i\theta_1}, \lambda = e^{i\theta_2}$, 因 $\lambda \neq \lambda$, 故 $\theta \neq \theta$, 而 $\bar{\lambda} = e^{-i\theta_2}$, 所以

$$\lambda \bar{\lambda} = e^{i(\theta_1 - \theta_2)} \neq e^{i0} = 1.$$

(2) 实正交矩阵、埃尔米特矩阵、规范矩阵属于不同特征值的特征向量是相互正交的, 证明与这里的叙述类似, 请读者自己写出 .

例 12 试证 Cauchy-Schwarz 不等式, 即对 " $\alpha, \beta \in U_n(C)$, 有

$$(\alpha, \beta) \overline{(\alpha, \beta)} \leq (\alpha, \alpha) (\beta, \beta).$$

证 设 $\alpha \neq 0$, 对任 $\lambda \in C$ 有

$$\begin{aligned} 0 &\leq (\lambda\alpha + \beta, \lambda\alpha + \beta) = (\lambda\alpha, \lambda\alpha) + (\lambda\alpha, \beta) + (\beta, \lambda\alpha) + (\beta, \beta) \\ &= \lambda\bar{\lambda}(\alpha, \alpha) + \lambda(\alpha, \beta) + \bar{\lambda}(\beta, \alpha) + (\beta, \beta) \\ &= \bar{\lambda}(\lambda(\alpha, \alpha) + (\beta, \alpha)) + \lambda(\alpha, \beta) + (\beta, \beta). \end{aligned}$$

取 $\lambda = -\frac{(\beta, \alpha)}{(\alpha, \alpha)}, \bar{\lambda} = -\frac{(\alpha, \beta)}{(\alpha, \alpha)}$, 则有 $-\frac{(\alpha, \beta)(\beta, \alpha)}{(\alpha, \alpha)} + (\beta, \beta) \geq 0$,

所以有 $(\alpha, \beta) \overline{(\alpha, \beta)} \leq (\alpha, \alpha) (\beta, \beta)$.

讲评 在欧几里得空间, 证明这个不等式用的是二次三项式的理论, 这里不能用, 故选择特殊的 λ , 以完成证明 .

例 13 证明: 线性变换 A 是规范 (正规) 的当且仅当 $A = H_1 + iH_2$, 其中 H_1 和 H_2 是可交换的埃尔米特变换 .

证 先证对任意一线性变换 A , 均有 $A = H_1 + iH_2$, 其中 H_1, H_2 均为埃尔米特变换 . 因为

$$A = \frac{A + A^*}{2} + \frac{A - A^*}{2} = \frac{A + A^*}{2} + i \frac{A - A^*}{2i} = H_1 + iH_2,$$

$$\text{且有} \quad (H_1)^* = \frac{A + A^*}{2}^* = H_1; \\ (H_2)^* = \frac{A - A^*}{2i}^* = \frac{-(A^* - A)}{2i} = H_2.$$

下面只需证明

$$A \text{ 是规范的 } H_1 H_2 = H_2 H_1.$$

由 $A = H_1 + iH_2$, 得 $A^* = (H_1 + iH_2)^* = H_1^* - iH_2^* = H_1 - iH_2$, 所以

$$A^* A = (H_1 - iH_2)(H_1 + iH_2) = H_1^2 - iH_2 H_1 + iH_1 H_2 + H_2^2,$$

$$A A^* = (H_1 + iH_2)(H_1 - iH_2) = H_1^2 + iH_2 H_1 - iH_1 H_2 + H_2^2.$$

于是

$$A^* A = A A^* \quad 2iH_1 H_2 = 2iH_2 H_1 \quad H_1 H_2 = H_2 H_1,$$

故知

$$A \text{ 是规范的 } A^* A = A A^* \quad H_1 H_2 = H_2 H_1.$$

讲评 (1) 这道题给出规范变换的又一条性质, 即它可以写成两个埃尔米特变换的“和” $(H_1 + iH_2)$, 且 H_1 和 H_2 乘法可交换. 对规范矩阵当然也有类似的结论.

(2) 比较实矩阵的分解: 任意 n 阶矩阵

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2} = A_1 + A_2,$$

其中 A_1 为实对称矩阵, A_2 为反对称矩阵. 容易验证.

$$\overline{A A^T} = (A_1 + A_2)(A_1 - A_2) = A_1^2 + A_2 A_1 - A_1 A_2 - A_2^2,$$

$$\overline{A^T} A = (A_1 - A_2)(A_1 + A_2) = A_1^2 - A_2 A_1 + A_1 A_2 - A_2^2.$$

故得

$$\overline{A^T} A = A A^T \quad A A_2 = A_2 A_1.$$

(3) 设 A 为实矩阵, 按题的分解是

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2} = \frac{A + A^T}{2} + i \frac{-i(A - A^T)}{2} = H_1 + iH_2,$$

即 $H_2 = i(A^T - A)/2$, 则 H_1, H_2 均为埃尔米特矩阵, 于是由题解知有

$$\overline{A^T} A = A A^T \quad H_1 H_2 = H_2 H_1,$$

即按如此拆开成两个矩阵之和,则 A 为规范矩阵的充要条件是 H_1 与 H 乘法可交换.

注意,反对称矩阵是规范矩阵,但不是埃尔米特矩阵,实对称矩阵当然是埃尔米特矩阵.

7.3 考试题评析

例 1 设 $A = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

试求酉矩阵 U , 使 $U^{-1}AU$, $U^{-1}BU$ 同时为对角矩阵 (写出对角矩阵).

解 因 $A = I + B$, 故有 $AB = BA$, 所以 A, B 可同时酉对角化, 由

$$|\lambda I - B| = \begin{vmatrix} \lambda & -i \\ -i & \lambda \end{vmatrix} = \lambda^2 + 1 = 0,$$

知 $\lambda = \pm i$ 为 B 的特征值, 分别解 B 的属于特征值 $\pm i$ 的特征向量, 有

由 $(iI - B) = \begin{pmatrix} i & -i \\ -i & i \end{pmatrix}$, 得 $x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $e = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$;

由 $(-iI - B) = \begin{pmatrix} -i & -i \\ -i & -i \end{pmatrix}$, 得 $x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $e = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

取 $U = (e, e) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, 则有

$$U^{-1}AU = \begin{pmatrix} 1+i & \\ & 1-i \end{pmatrix}, U^{-1}BU = \begin{pmatrix} i & \\ & -i \end{pmatrix}.$$

评析 (1) 解这类问题, 若先验证 $AB = BA$, 且 $\bar{A}^T A = A \bar{A}^T$, $\bar{B}^T B = B \bar{B}^T$, 则从理论上可保证它们可同时对角化, 故当把一个对角化后, 比如求 B 的特征值、特征向量, 找到方阵 U , 使 $U^{-1}BU$ 为对角矩阵时, $U^{-1}AU$ 亦为对角矩阵, 直接做矩阵乘法, 算出对角矩阵是什么就行了.

(2) 若不验证, A, B 是否为规范矩阵及 $AB = BA$ 则可分别求 A, B 的特征值和特征向量, 看它们是否有公共的 n 个正交的特征向量? 若有, 排成矩阵 U , 则

$$\overline{U}^T A U = \text{diag}\{\lambda_1, \cdots, \lambda_n\},$$

$$\overline{U}^T B U = \text{diag}\{\mu_1, \cdots, \mu_n\}.$$

直接写出各自的对角矩阵就行了, 不用做矩阵乘法.

例 2 设 e_1, e_2, e_3 为欧几里得空间 R^3 的自然基, R^3 上的线性变换 σ 分别把向量 $\alpha_1 = (1, 1, 1)^T$, $\alpha_2 = (0, 1, 1)^T$, $\alpha_3 = (0, 0, 1)^T$, 变为向 $\beta_1 = (7, -1, 4)^T$, $\beta_2 = (3, 1, 2)^T$, $\beta_3 = (1, 2, 1)^T$.

试求 σ 和 σ^* (σ 的共轭变换) 在基 e_1, e_2, e_3 下的矩阵表示.

解 设 σ 在基 e_1, e_2, e_3 下的矩阵表示为 A , 则 σ^* 的矩阵为 $\overline{A}^T$, 即有

$$\sigma(e_1, e_2, e_3) = (e_1, e_2, e_3) A,$$

任取 $\alpha \in R^3$, 若 $\beta = \sigma \alpha$, 则它们的坐标(在自然基下)满足

$$y = Ax,$$

其中 x, y 分别为 α, β 的坐标, 因 α, β 在自然基下的坐标就是它们自身, 故有

$$\beta_1 = A\alpha_1, \quad \beta_2 = A\alpha_2, \quad \beta_3 = A\alpha_3,$$

$$\text{即} \quad (\beta_1, \beta_2, \beta_3) = A(\alpha_1, \alpha_2, \alpha_3).$$

$$\text{故} \quad A = (\beta_1, \beta_2, \beta_3)(\alpha_1, \alpha_2, \alpha_3)^{-1}$$

$$= \begin{pmatrix} 7 & 3 & 1 \\ -1 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 1 \\ -2 & -1 & 2 \\ 2 & 1 & 1 \end{pmatrix},$$

$$\text{所以} \quad \overline{A}^T = \begin{pmatrix} 4 & -2 & 2 \\ 2 & -1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

评析 (1) 这里用的是公式 $y = Ax$, 若已知 3 个向量的象, 则可构成一个三阶矩阵, 故可用矩阵乘法得出 A . 关键是弄清所有向

量的坐标及矩阵 A 都是在同一个基上的.

(2) 若设 σ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵为 B , 则从基 e_1, e_2, e_3 到基 $\alpha_1, \alpha_2, \alpha_3$ 的过渡矩阵为 $T = (\alpha_1, \alpha_2, \alpha_3)$, 则

$$B = T^{-1}AT, \text{ 即 } A = TB T^{-1},$$

且有 $\sigma(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3)B. \quad (1)$

故 $B = (\alpha_1, \alpha_2, \alpha_3)^{-1}(\beta_1, \beta_2, \beta_3).$

所以得

$$\begin{aligned} A &= T(\alpha_1, \alpha_2, \alpha_3)^{-1}(\beta_1, \beta_2, \beta_3)T^{-1} \\ &= (\beta_1, \beta_2, \beta_3)(\alpha_1, \alpha_2, \alpha_3)^{-1}. \end{aligned}$$

最后这个表达式同题解中公式. 这里主要公式是两个: ①公式 (1); ② σ 在两个不同的基上的矩阵是相似的.

例 3 设 σ, τ 都是西空间 V 的正定的埃尔米特变换, 证明: 线性变换 σ 的特征值均为正数.

解 方法 1 设 $\sigma\alpha = \lambda\alpha$, 并记 σ, τ 在基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵分别为 A, B , 向量 α 的坐标为 x , 则有 A, B 均正定, 且有

$$ABx = \lambda x, \text{ 故 } A^{-1}ABx = \lambda A^{-1}x, \text{ 即 } Bx = \lambda A^{-1}x.$$

在上边最后一个等式两边同乘 $\bar{x}^T$, 得

$$\bar{x}^T Bx = \lambda \bar{x}^T A^{-1}x.$$

当 $x \neq 0$ 时, 有 $\bar{x}^T Bx > 0, \bar{x}^T A^{-1}x > 0$, 故知 $\lambda > 0$, 得证.

方法 2 任取 $\alpha \in V$, 因 τ 是正定的埃尔米特变换, 故是可逆变换, 所以当 $\alpha \neq 0$ 时, $\tau\alpha \neq 0$, 于是由 σ 正定知有

$$(\sigma\alpha, \tau\alpha) > 0.$$

又 τ 是自伴随变换, 所以有

$$(\tau(\sigma\alpha), \alpha) = (\sigma\alpha, \tau\alpha) > 0.$$

今取 α 为 σ 的特征向量, 即有 $\sigma\alpha = \lambda\alpha$, 代入上式得

$$(\tau(\sigma\alpha), \alpha) = (\tau\lambda\alpha, \alpha) = \lambda(\tau\alpha, \alpha) > 0.$$

因 $(\tau\alpha, \alpha) > 0$, 故 $\lambda > 0$.

评析 无论用哪种方法, 要注意严谨. 如方法 2 中, 要说明

$\tau\alpha \neq 0$, 否则 $(\sigma\tau\alpha, \tau\alpha) \geq 0$. 方法 1 中, 要说明 A^{-1} 正定 (当 A 正定时) 等.

例 4 设 σ 是 n 维酉空间 V 上的线性变换, σ^* 是 σ 的共轭变换, $W \subseteq V$ 试证:

(1) 若 W 是 σ 的不变子空间, 则 $W^\perp$ 为 σ^* 的不变子空间;

(2) σ 具有任意 r 维不变子空间 ($r = 0, 1, \dots, n$).

证 (1) 任取 $\beta \in W$, 由 W 是 σ 的不变子空间, 知有 $\sigma\beta \in W$, 于是, 任取 $\alpha \in W^\perp$, 有

$$(\sigma^* \alpha, \beta) = (\alpha, \sigma\beta) = 0,$$

所以 $\sigma^* \alpha \in W^\perp$, 故 $W^\perp$ 是 σ^* 的不变子空间.

(2) 设 $\sigma^* \xi = \lambda \xi$, 则 $L(\xi)$ 是 σ^* 的不变子空间, 由问题 (1) 知 $L(\xi)^\perp$ 是 σ 的 $n-1$ 维不变子空间. 记 $\sigma|_{V_{n-1}} = \tau$, 则 τ^* 在 V_{n-1} 上有一维不变子空间, 其正交补是 τ 的 $n-2$ 维不变子空间, 而 σ, τ 在 V_{n-1} 上作用相同, 故 σ 有 $n-2$ 维不变子空间, 依此续行, 可知 σ 有任意 r 维的不变子空间.

评析 问题 (2) 也可用 Jordan 标准形的理论证明: 设 σ 在基 $\alpha, \dots, \alpha$ 下的矩阵为 Jordan 标准形 J , 则 $W_r = \langle \alpha, \dots, \alpha \rangle$ 是 σ 的 r 维不变子空间 ($r = 1, 2, \dots, n-1$). 因满足任取 $\alpha \in W_r, \sigma\alpha \in W_r$.

例 5 试证: 复线性空间上的任意两个可交换的线性变换必有公共的特征向量.

证 设 σ, τ 为复线性空间上的线性变换, σ 在复数域上必有特征值记为 λ_0 , 并记属于特征值 λ_0 的特征子空间为 V_{λ_0} . 则由 $\sigma\tau = \tau\sigma$ 知对 $\alpha \in V_{\lambda_0}$, 有

$$(\sigma - \lambda_0 \varepsilon)(\tau\alpha) = \tau(\sigma\alpha) - \lambda_0 \tau\alpha = \tau\lambda_0 \alpha - \lambda_0 \tau\alpha = 0,$$

故 $V_{\lambda_0} = \text{Ker}(\sigma - \lambda_0 \varepsilon)$ 是 τ 的不变子空间.

记 $\tau = \tau|_{V_{\lambda_0}}$, 则 τ 是 V_{λ_0} 上的线性变换. 在 $\mathbb{C}$ 上 τ 有特征值 μ , 故有特征向量 α , 使 $\tau\alpha = \mu\alpha$ ($\alpha \neq 0, \alpha \in V_{\lambda_0}$), 而 $\sigma\alpha = \lambda_0\alpha$, 所以 α 是 σ, τ 的公共特征向量 (因 $\tau\alpha = \mu\alpha = \tau\alpha$).

评析 (1) 这是由《理工科代数基础》上的一道习题改造而出的考题(当然也见于其他参考书) 如果对那道题理解深刻的话, 这道题就可看做“熟题”了. 第一步很重要, σ 在复数域上有特征值, 从而有特征向量, 有特征子空间, 接下来就与习题类似了.

(2) 由证明过程看, 把“复”改成“实”就做不下去了. 因不能保证每个线性变换都有实特征值, 从而有实特征向量. 只有当 σ, τ 均为实对称变换时, 若有 $\sigma\tau = \tau\sigma$, 则 σ, τ 必有公共特征向量. 事实上, 它们可以同时对角化.

例6 设 A 是 n 维酉空间 V 上的规范变换, $W \subset V$, 且 W 是 A 的不变子空间, $W^\perp$ 是 W 在 V 中的正交补, 求证: A 在 $W^\perp$ 上的限制 $A|_{W^\perp}$ 是规范变换.

证 在 W 中取标准正交基 $\alpha_1, \dots, \alpha_r$, 并扩充为 V 的标准正交基

$$\alpha_1, \dots, \alpha_r, \beta_{r+1}, \dots, \beta_n,$$

故有

$$W^\perp = L(\beta_{r+1}, \dots, \beta_n), \quad V = W \oplus W^\perp.$$

又因为 W 是 A 的不变子空间, 故 A 在基 $\alpha_1, \dots, \alpha_r, \beta_{r+1}, \dots, \beta_n$ 下的方阵表示为

$$M = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix},$$

则由 A 是规范变换, 知 M 是规范矩阵, 故有

$$\overline{M}^T M = M \overline{M}^T,$$

$$\text{即} \quad \begin{pmatrix} \overline{A_1}^T & 0 \\ \overline{C}^T & \overline{A_2}^T \end{pmatrix} \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix} = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} \overline{A_1}^T & 0 \\ \overline{C}^T & \overline{A_2}^T \end{pmatrix},$$

$$\text{即} \quad \begin{pmatrix} \overline{A_1}^T A_1 & \overline{A_1}^T C \\ \overline{C}^T A_1 & \overline{C}^T C + \overline{A_2}^T A_2 \end{pmatrix} = \begin{pmatrix} A_1 \overline{A_1}^T + C \overline{C}^T & C \overline{A_2}^T \\ A_2 \overline{C}^T & A_2 \overline{A_2}^T \end{pmatrix}.$$

$$\text{故有} \quad \overline{A_1}^T A_1 = A_1 \overline{A_1}^T + C \overline{C}^T, \quad (1)$$

$$\overline{C}^T C + \overline{A_2}^T A_2 = A_2 \overline{A_2}^T, \quad (2)$$

(2)式两边取迹(即对角线元素之和)得

$$\operatorname{tr}(\bar{\mathbf{C}}^T \mathbf{C}) + \operatorname{tr}(\bar{\mathbf{A}}_2^T \mathbf{A}_2) = \operatorname{tr}(\mathbf{A} \bar{\mathbf{A}}_2^T). \quad (3)$$

这是由于两个方阵相等,它们的对角线元素之和当然相等,以及对任意两个 n 阶矩阵 $\mathbf{A}, \mathbf{B}$, 有 $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B}$.

记 $\mathbf{C} = (\mathbf{c}_1, \cdots, \mathbf{c}_{n-r})$, $\bar{\mathbf{C}}^T \mathbf{C} = (d_{ij})$, 则 $d_{ii} = \bar{\mathbf{c}}_i^T \mathbf{c}_i$, 故

$$\operatorname{tr} \bar{\mathbf{C}}^T \mathbf{C} = \sum_{i=1}^n d_{ii} = \sum_{i=1}^n \bar{\mathbf{c}}_i^T \mathbf{c}_i = \sum_{i=1}^n \|\mathbf{c}_i\|^2 \geq 0,$$

等号仅当 $\mathbf{c}_i = \mathbf{0} (i=1, \cdots, n-r)$ 成立.

又 $\operatorname{tr}(\bar{\mathbf{A}}_2^T \mathbf{A}_2) = \operatorname{tr} \mathbf{A}_2 \bar{\mathbf{A}}_2^T$ (利用对任意 n 阶矩阵 $\mathbf{A}, \mathbf{B}$, 有 $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$, 其证明如下:

设 $\mathbf{A} = (a_{ij}), \mathbf{B} = (b_{ij}), \mathbf{AB} = (h_{ij}), \mathbf{BA} = (\text{璁}_{ij})$, 则

$$\operatorname{tr} \mathbf{AB} = \sum_{i=1}^n h_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki},$$

$$\operatorname{tr} \mathbf{BA} = \sum_{k=1}^n \text{璁}_{kk} = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ik} = \sum_{k=1}^n \sum_{i=1}^n a_{ik} b_{ki},$$

于是得 $\operatorname{tr} \mathbf{AB} = \operatorname{tr} \mathbf{BA}$. 故由(3)式知

$$\operatorname{tr} \bar{\mathbf{C}}^T \mathbf{C} = 0 \rightarrow \mathbf{c}_i = \mathbf{0}, \quad i=1, 2, \cdots, n-r.$$

故知 $\mathbf{C} = \mathbf{0}$, 再代回(2)式中即得

$$\bar{\mathbf{A}}_2^T \mathbf{A}_2 = \mathbf{A}_2 \bar{\mathbf{A}}_2^T,$$

即 $\mathbf{A}_2$ 为规范矩阵. 而此时有

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} \\ \mathbf{A}_2 \end{pmatrix},$$

故 $\mathbf{A}_2$ 就是 $\mathbf{A}|_{W^\perp}$ 在基 $\beta_{r+1}, \cdots, \beta_n$ 上的矩阵表示, 因 $\mathbf{A}_2$ 为规范矩阵, $\beta_{r+1}, \cdots, \beta_n$ 为标准正交基, 故 $\mathbf{A}|_{W^\perp}$ 为规范变换.

评析 (1) 此题要彻底证明出来, 必须对(2)式取迹, 如果想不到这一步, 就不好办了. 但怎么才能想到这一步呢. 注意“目标”是 $\mathbf{C} = \mathbf{0}$, 即 $\mathbf{C}$ 的各列为零向量, 只能从 $\bar{\mathbf{c}}_i^T \mathbf{c}_i = 0$ 中推出, 又联想知识 $\operatorname{tr} \mathbf{AB} = \operatorname{tr} \mathbf{BA}$, 所以用在(2)式上, 有 $\operatorname{tr} \bar{\mathbf{A}}_2^T \mathbf{A}_2 = \operatorname{tr} \mathbf{A}_2 \bar{\mathbf{A}}_2^T$. 于是有

$\text{tr } \bar{C}^T C = 0$. 需要的结果或称“目标”是“ $C = 0, \bar{A}_2^T A_2 = A_2 \bar{A}_2^T$.”通过取迹达到目的.

(2) 证明的前半部, 大部分人应该会的, 因由已知 W 是不变子空间, 由第 5 章定理 4, 知如解答中取基, 方阵表示为

$$\begin{pmatrix} A & C \\ 0 & A_2 \end{pmatrix} \quad (\text{如果这个定理也忘了, 就会感到“寸步难行”}). \text{ 下面的}$$

目标是证 $C = 0$ 和 A_2 是规范矩阵. 如果不知或“忘记” $\text{tr } AB = \text{tr } BA$ 这个知识, 就无法联想到在 (2) 式两边取迹. 故在理解的基础上, 多记一些知识, 做题时路子就多了.

例 7 设 σ 是 n 维西空间 V 上的线性变换, σ^* 是 σ 的共轭变换, 求证 $\text{Im } \sigma^* = (\text{Ker } \sigma)^\perp$.

证 因任取 $\alpha \in \text{Im } \sigma^*$ ($\forall \xi \in V$, 使 $\alpha = \sigma^* \xi$), $\beta \in \text{Ker } \sigma$, 有 $(\alpha, \beta) = (\sigma^* \xi, \beta) = (\xi, \sigma \beta) = (\xi, 0) = 0$, 故

$$\text{Im } \sigma^* \perp \text{Ker } \sigma$$

又 $\dim \text{Im } \sigma^* = r(\bar{A}^T) = r(A) = \dim(\text{Ker } \sigma)^\perp$,

所以有

$$\text{Im } \sigma^* = (\text{Ker } \sigma)^\perp.$$

评析 (1) 这个题的证明主要有两层意思: ①证 $\text{Im } \sigma^* \perp \text{Ker } \sigma$, 即证出 $\text{Im } \sigma^* \subset (\text{Ker } \sigma)^\perp$; ②证“相等”, 只要证维数相同. 关于维数的推导, 用了 (i) 当 σ 在 V 中基 $\varepsilon_1, \dots, \varepsilon_n$ 下的矩阵为 A 时, σ^* 在此基上的矩阵为 $\bar{A}^T$ 而方阵经共轭转置后秩不变. (ii) $\dim(\text{Ker } \sigma) = n - r(A)$, 即是 $Ax = 0$ 的解空间的维数. (iii) $\dim(\text{Ker } \sigma)^\perp = n - \dim(\text{Ker } \sigma) = n - (n - r(A)) = r(A)$.

(2) 也许有的同学看到这样的题, 开始有点不知如何下手, 但紧扣定义, 要证两个子空间相等, 一个要它们维数相同, 就有了一个思路. 另一个是证 $\text{Im } \sigma^* \perp \text{Ker } \sigma$, 这样就可得到 $\text{Im } \sigma^* \subset (\text{Ker } \sigma)^\perp$. 故见到证明题不必紧张, 仔细分析“目标”是什么?! 就可以找到思路.

参 考 文 献

- [1] 张贤科,许甫华. 高等代数学. 北京:清华大学出版社,1998
- [2] 许甫华,张贤科. 高等代数解题方法. 北京:清华大学出版社,2000
- [3] 许以超. 代数学引论(第二版). 上海:上海科学技术出版社,1992
- [4] 俞正光,李永乐,吕志. 理工科代数基础. 北京:清华大学出版社,1998
- [5] ipschutz S .Theory and Problems of Linear Algebra .Springer-Verlag,1991