现代数学基础

5

线性代数与矩阵论

（第二版）

■ 许以超　编著

高等教育出版社
HIGHER EDUCATION PRESS

内容简介

本书是将矩阵论和线性空间理论溶合在一起编写的。先以中学时熟悉的多项式为基础，将多项式理论交代清楚。接下去讲多元多项式。然后是矩阵论和线性空间理论的基本工具：行列式、矩阵以及线性方程组求解理论。从而引进线性空间、线性不等式和它上面的线性变换，以及求复方阵的 Jordan 标准形的代数理论和几何解释，Jordan 标准形的应用，它包含了方阵函数和方阵在复相似下的标准型理论。给出了线性函数和它的推广，即多重线性函数，Grassmann 代数以及张量场。接着转向内积空间（即实和复 Euclid 空间的结构和二次型的分类）。最后三章是广义逆矩阵的几何基础和矩阵处理，非负矩阵的基本性质和复矩阵偶在相抵下的标准形。

本书的特点是充分发挥矩阵技巧在矩阵论和线性空间理论中的应用，涉及面也比较广。本书的另一个特点是书中的例题和习题比较难一点，虽然本书的一些习题已经被一些作者选为例题，但是本书的目的是使同学有一个良好的严格训练环境，可以自由地选择这些习题来做。

本书可作为大学数学系高等代数或矩阵论的教科书或教学参考书，也可作为高年级学生考研的复习参考资料，同时希望本书能对科研工作者有较大的参考价值。

图书在版编目(CIP)数据

线性代数与矩阵论/许以超编著.—2版.—北京:高等教育出版社,2008.6
ISBN 978-7-04-024307-9

Ⅰ.线… Ⅱ.许… Ⅲ.①线性代数-高等学校-教材 ②矩阵-理论-高等学校-教材 Ⅳ.O151.2

中国版本图书馆 CIP 数据核字(2008)第 067064 号

策划编辑 王丽萍　**责任编辑** 高尚华　**封面设计** 张　楠
版式设计 王艳红　**责任校对** 胡晓琪　**责任印制** 韩　刚

出版发行 高等教育出版社
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
总　　机 010-58581000

经　　销 蓝色畅想图书发行有限公司
印　　刷 北京外文印刷厂

开　　本 787×1092　1/16
印　　张 32.75
字　　数 650 000

购书热线 010-58581118
免费咨询 800-810-0598
网　　址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.landraco.com
http://www.landraco.com.cn
畅想教育 http://www.widedu.com

版　　次 1992 年 5 月第 1 版
2008 年 6 月第 2 版
印　　次 2008 年 6 月第 1 次印刷
定　　价 59.00 元

物料号　24307-00

序言

本书起源于作者在 1960 年为北京大学数学系 1956 级 (即 1962 届, 这一届的学生是六年制) 单复变函数论及多复变函数论专门化讲授的一门副课时所写的讲义 (当时作者是华罗庚教授的多复变函数论研究生)。当时的目的是为单复变函数论及多复变函数论专门化的学生学习华罗庚教授的典型域理论打基础, 所以作者从华罗庚教授的典型群和典型域的论文以及许宝騄教授的矩阵论的论文中收集了许多矩阵论结果并写成讲义。

在 1958 年, 中国科学技术大学数学系第一届学生的基础课 (这里是指大学数学系一、二年级的所有数学课程) 是由当时任数学系主任的华罗庚教授用 “一条龙” 的教学方式包下来的。在 1960 级由吴文俊教授用 “一条龙” 的教学方式包下来后, 1961 级则由华罗庚教授指定龚升教授和作者 (作者在 1961 年研究生毕业不久) 来担任数学系一、二年级的所有数学课程的教学任务。龚升教授讲授与分析有关的内容, 史济怀教授当时担任助教。其他的数学课程由作者来讲授, 所以作者在 1961 年开始为中国科学技术大学数学系 1961 级和 1963 级两个年级授课, 时间长达两届, 共四年。讲授的内容包括平面和空间解析几何 (当时中学不教平面解析几何)、高等代数和线性代数、抽象代数 (当时大学数学系不教抽象代数) 等。这门课由李乔教授和李炯生教授当时分别担任一届助教。其后, 作者将讲义整理成《代数学引论》一书, 在华罗庚教授的推荐下, 于 1966 年在上海科学技术出版社出版。这本书, 在国内教材中第一

次从几何和代数的角度, 将解析几何、仿射几何、射影几何、辛几何和线性代数交叉地, 作为整体地编写, 而且充分地利用了矩阵工具, 将一些线性空间的问题化为代数问题。书中收录了大量难题, 成为“文革”后, 考研究生所必备的参考书, 并且影响了“文革” 后出版的很多高等代数教科书。

作者在 1986 年, 为了适应新的需要, 将《代数学引论》中的部分章节重新整理, 删去几何部分, 并加写了一些与应用数学有关的矩阵论内容, 将改写好的讲义为南开数学研究所和南开大学数学系合办的第一届 (陈省身) 数学实验班 1986 级讲授了一年半的高等代数。后来又在清华大学数学系为 1989 级本科生讲授了一年半高等代数。然后, 将讲义改写成《线性代数和矩阵论》一书, 在 1992 年由高等教育出版社出版。该书在 1996 年获得教育部国家优秀教材一等奖。随后, 在河南大学数学系为 2000 级本科生讲授了一年高等代数。从 2005 年下半年起, 作者应聘在浙江大学数学研究中心和数学系合办的丘成桐英才班从第一届 (2005 级) 开始讲授高等代数。另一方面, 高等教育出版社希望作者将《线性代数和矩阵论》一书重新整理后出版。藉此机会, 作者才有可能从教学角度重新整理和改写并出版此书。

众所周知, 数学分析、高等代数和空间解析几何学是数学系一年级学生学习数学的“启蒙课程”。在此基础上, 数学系的学生才开始进入数学的各个一级分支学科的基础理论课程的学习, 它们就是数学系学生的大量后继数学课程。例如, 抽象代数课才是代数学的“入门课程”。

高等代数课主要讲授线性代数理论。线性代数的重要性在于它考虑了一类简单的数学模型, 而大量的理论问题及应用问题, 可以通过“线性化”变成线性代数的问题。作为基础知识和基础训练, 也为了训练数学逻辑思维, 熟练地掌握线性代数的理论和技巧是很有必要的, 而且这也是我编写本书的出发点。

贯穿这本书, 讲述了线性代数理论。可以这样去理解: 由于引进了直角坐标系, 平面上的点便可以用实数对来表示。反之, 实数对的几何意义为: 它表示了平面上的点, 所谓平面解析几何学, 就是将平面上的一些几何问题化为代数问题。确切地说, 化为多项式的问题。和它完全类似, 在线性空间中取定一组基后, 线性映射可以用矩阵来表示, 二次型可以用对称方阵来表示。反之, 矩阵的几何意义为: 它表示了线性空间上的线性映射。所谓线性代数学, 就是或者直接研究线性空间的几何问题, 或者将线性空间的一些几何问题化为矩阵问题。所以线性空间理论和矩阵论实际上是相伴而生的。

本书的特点是偏重矩阵技巧。这是因为矩阵方法表达具体和明显, 并且具有一整套的标准计算技巧和处理办法。技巧性很强, 且具有实用价值。在数值分析、组合数学、概率论、数理统计和经济数学等方面, 都有重要的应用。由于矩阵方法着重在技巧, 所以初学者不能很快地适应和掌握它。这只有通过反复计算, 才能逐渐熟练。要深入地理解数学, “概念清楚”、“运算熟练” 和 “灵活运用” 这三个方面是缺一不可的。正因为如此, 书中例题着眼于介绍各种矩阵技巧, 并且在每节后面都附上了习

题。习题分两类, 一类是普通的习题, 即所谓的台阶题; 另一类是补充题, 其中包括了历届主要单位的研究生考题中有一定难度的题目, 我们将这部分题目放在每一章的最后。通过这些习题, 希望培养学生们如何运用已知理论和掌握的知识来学会和培养分析问题的能力和解决问题的能力。进入大学数学系一年级的同学们有一个通病, 他们不重视数学理论, 只关心数学技巧, 认为有了数学技巧, 题目就会做了, 什么都不怕了。他们用中学时学习数学的经验, 不明白数学技巧来源于数学理论, 不明白概念和理论的重要性, 因此, 他们非常需要改变中学时学习数学的方法。请注意, 本书的习题, 是本书的重要组成部分之一, 我们花了大量工作收集和整理。由于题解书不论对学生或老师, 都只会促使人们不去思考, 丧失通过做题来掌握和熟练数学理论和技巧的功能。实际上题解书毁了一本习题集, 所以我们不希望看到出现本书的题解书。

全书共分十四章。以第一章到第四章的多项式 (一元多项式和多元多项式)、行列式、矩阵以及线性方程组求解理论为基础, 从第五章到第十二章, 引进了线性空间和线性变换, 给出了 Jordan 标准形的代数理论和几何理论, 以及 Jordan 标准形的应用, 给出了线性函数和多重线性函数, 介绍了多重线性函数的张量积和外积、张量场, 讨论了实和复 Euclid 空间, 实和复二次型和 Hermite 型的分类。作为应用, 给出各种类型的矩阵标准形理论和它们的证明, 以及它们的应用。为了对矩阵标准形理论有深刻的了解, 我们给出了相等概念的扩充, 即引进等价关系的定义, 同时自然地引进了商空间的概念。另一方面, 讨论了线性不等式和广义逆矩阵。这些都是线性代数理论的主要部分。在此基础上, 引进了方阵函数, 给出了方阵在复相似下的标准型 (复相似的定义和标准型是许宝騄教授在 1955 年得到的, 他将这一工作发表在 1955 年的北京大学学报 (自然科学版) 上, 这一工作的后继工作也发表在 1955 年及 1957 年北京大学学报 (自然科学版) 上)。另外, 在第十三章介绍了非负矩阵的基本性质, 在第十四章处理了复矩阵偶在相抵下的标准形。

本书可作为大学一年级的高等代数课程的教科书 (书中打星号“*” 的部分可不作为教材的正式内容), 也可作为低年级高材生的参考书、高年级学生为考研究生的复习参考书, 以及高等代数和矩阵论的教学参考书。我们也希望此书能对科研工作者有较大的参考价值。全书中每章的定义、定理、引理和推论的编号是一起排的, 引理和定理证完的符号为 “□”。

作者在大学期间, 有幸得到导师段学复教授、聂灵沼教授、丁石孙教授和王萼芳教授以及许宝騄教授的教诲和指导; 在研究生期间, 在导师华罗庚教授指导下, 掌握了矩阵技巧, 在此对他们表示深深的感谢。在本书的编写和修改过程中, 按时间次序, 得到了李乔、李炯生、史济怀、白志东、沈忠民、石生明、顾沛、王天择、陈杰诚和李方等教授的具体帮助和支持, 在此对他们也表示衷心的感谢。另外, 高等教育出版社的张小萍、高尚华和王丽萍三位编辑对本书的编辑所作的辛勤劳动, 这些都对本书的出版起了很大的作用。最后, 浙江大学数学研究中心和浙江大学数学系大力

支持我的工作。在此, 一并表示衷心的感谢。由于作者水平有限, 书中错误和不妥之处在所难免, 还望读者多多提出宝贵意见。

许以超 2007 年 10 月 7 日
中国科学院数学研究院

目录

第一章　多项式理论

§1.1　一元多项式的代数运算

为了书写简便起见, 我们先引进**求和符号** $\sum$ 和**求积符号** $\prod$.

给定正整数 n 以及 n 个数 $a_1, a_2, \cdots, a_n$. 由于数的加法适合交换律和结合律, 所以 n 个数 $a_1, a_2, \cdots, a_n$ 的连加可以记作

$$a_1+a_2+\cdots+a_n=\sum_{j=1}^{n} a_j. \tag{1.1.1}$$

这个和式中的**足码** j 取遍 $1, 2, \cdots, n$. 即 j 只是表示求和是从 1 加到 n, 所以也可以用其他足码来代替. 例如, 易 j 为 k, 或易 j 为 i_j, 其中 $i_1 i_2 \cdots i_n$ 是 $1, 2, \cdots, n$ 的排列, 即

$$\sum_{j=1}^{n} a_j=\sum_{k=1}^{n} a_k=\sum_{j=1}^{n} a_{i_j} \tag{1.1.2}$$

等. 由此可见, 不管和式的足码用什么符号, 重要的是它实际上表示了 n 个数 $a_1, a_2, \cdots, a_n$ 的和.

用归纳法思想, 我们也可以引进**多重和号**

$$\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_m=1}^{n_m} a_{i_1 i_2 \cdots i_m}. \tag{1.1.3}$$

为了说明问题, 先从**二重和号**入手. 给定正整数 m 和 n, 将 mn 个数 a_{jk}, $1 \leqslant j \leqslant m$,

$1 \leqslant k \leqslant n$ 排成如下矩形表:

$$
\begin{array}{cccc}
a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\
\vdots & \vdots & & \vdots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{m,n-1} & a_{mn}
\end{array} \tag{1.1.4}
$$

其中横排为**行**, 第 i 行为 $a_{i1}, a_{i2}, \cdots, a_{i,n-1}, a_{in}$; 竖排为**列**, 第 j 列为 $a_{1j}, a_{2j}, \cdots, a_{m-1,j}, a_{mj}$. 所以有 m 个行, n 个列. 第 i 行, 第 j 列的交叉元素为 a_{ij}, 前指标表示行, 后指标表示列. 将这些数全部加起来, 总和记作 S. 由于加法有交换律和结合律, 所以可以不计先后和次序地相加. 下面利用不同的求和方法具体地将总和 S 表达出来. 例如, 先按行将 n 个数加起来, 它们分别为 $\sum\limits_{k=1}^{n} a_{ik}$, $1 \leqslant i \leqslant m$. 再将这 m 个数加起来, 所以有

$$
\begin{aligned}
S &= (a_{11} + a_{12} + \cdots + a_{1n}) + \cdots + (a_{m1} + a_{m2} + \cdots + a_{mn}) \\
&= \sum_{k=1}^{n} a_{1k} + \cdots + \sum_{k=1}^{n} a_{mk} = \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} a_{jk} \Big).
\end{aligned}
$$

也可以先按列将 m 个数加起来, 它们分别为 $\sum\limits_{i=1}^{n} a_{ik}$, $1 \leqslant k \leqslant n$. 再将这 n 个数加起来, 所以有

$$
\begin{aligned}
S &= (a_{11} + a_{21} + \cdots + a_{m1}) + \cdots + (a_{1n} + a_{2n} + \cdots + a_{mn}) \\
&= \sum_{j=1}^{m} a_{j1} + \cdots + \sum_{j=1}^{m} a_{jn} = \sum_{k=1}^{n} \Big(\sum_{j=1}^{m} a_{jk} \Big).
\end{aligned}
$$

最后, 也可以按对角线将数加起来, 它们分别为 a_{11}, $a_{12} + a_{21}$, $a_{13} + a_{22} + a_{31}$, $a_{14} + a_{23} + a_{32} + a_{41}, \cdots$. 由相加的规律可以看出, 上面每个数 a_{ij} 的足码的和 $i+j$ 分别为定值 $1+1 = 2$, $1+2 = 2+1 = 3$, $1+3 = 2+2 = 3+1 = 4$, $1+4 = 2+3 = 3+2 = 4+1$. 所以总和 S 为

$$
\begin{aligned}
S &= a_{11} + (a_{12} + a_{21}) + (a_{13} + a_{22} + a_{31}) + (a_{14} + a_{23} + a_{32} + a_{41}) + \cdots \\
&= \sum_{i+j=2} a_{ij} + \sum_{i+j=3} a_{ij} + \sum_{i+j=4} a_{ij} + \sum_{i+j=5} a_{ij} + \cdots,
\end{aligned}
$$

其中足码 i, j 有如下限制:

$$
1 \leqslant i \leqslant m, \quad 1 \leqslant j \leqslant n. \tag{1.1.5}
$$

所以, 利用同一对角线上两实数的双足码的和相等可知, 总和

$$
S = \sum_{k=2}^{m+n} \Big(\sum_{i+j=k} a_{ij} \Big).
$$

到现在为止, 我们给出了总和 S 的三种表达式, 这时括号被省略掉了. 即

$$S=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}=\sum_{k=2}^{m+n}\sum_{i+j=k}a_{ij}, \tag{1.1.6}$$

其中第三个和式中指标适合式 (1.1.5). 当然, 我们还可以按照其他的求和规则, 将 S 写成各种形式的二重和式. 不过上面三种求和法是最常用的.

由上一等式还给出了和号交换的规则, 即有

$$S=\sum_{j=1}^{m}\sum_{k=1}^{n}a_{jk}=\sum_{k=1}^{n}\sum_{j=1}^{m}a_{jk}. \tag{1.1.7}$$

特别, 当 $m=n$ 时可知: 这一个二重和号可以写成一个统一的形式, 即写成

$$S=\sum_{j,k=1}^{n}a_{jk}. \tag{1.1.8}$$

和上面一样, 我们可以用归纳法引进**多重和号**. 要注意的是, 应该从求和规则出发来了解二重和号. 例如 $\sum\limits_{j=1}^{m}\sum\limits_{k=1}^{n}a_{jk}$ 就是先将里面和号 $\sum\limits_{k=1}^{n}a_{jk}$ 看作一个确定的实数, 按照外面的和号 $\sum\limits_{j=1}^{m}$ 排成单项, 再将每一项 $\sum\limits_{k=1}^{n}a_{jk}$ 排成单项, 这样就将全部 a_{jk} 排成了单项. 利用这一原则, 不难定义 p 重和号了. 例如, 给定 $n_1n_2\cdots n_p$ 个实数 $a_{i_1i_2\cdots i_p}$, $1\leqslant i_j\leqslant n_j$, $1\leqslant j\leqslant p$. 记这 $n_1n_2\cdots n_p$ 个实数的总和为 S, 则 S 可以记成 **p 重和式**

$$S=\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{i_1i_2\cdots i_p}. \tag{1.1.9}$$

它按照关系

$$\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{i_1i_2\cdots i_p}=\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{1i_2\cdots i_p}+\cdots+\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{n_1i_2\cdots i_p} \tag{1.1.10}$$

来定义. 由于二重和号可以交换次序, 所以 p 重和号也可以交换次序. 即对 $1, 2, \cdots, p$ 的任一排列 $k_1k_2\cdots k_p$, 则

$$\sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_p=1}^{n_p}a_{i_1i_2\cdots i_p}=\sum_{i_{k_1}=1}^{n_{k_1}}\sum_{i_{k_2}=1}^{n_{k_2}}\cdots\sum_{i_{k_p}=1}^{n_{k_p}}a_{i_{k_1}i_{k_2}\cdots i_{k_p}}. \tag{1.1.11}$$

特别, 当 $n_1=n_2=\cdots=n_p=n$ 时, 由和号的交换性可知, p 重和号可以写成一个统一的形式, 即写成

$$\sum_{i_1,i_2,\cdots,i_p=1}^{n}a_{i_1i_2\cdots i_p}. \tag{1.1.12}$$

给定正整数 n 以及 n 个数 $a_1, a_2, \cdots, a_n$. 由于数的乘法也适合交换律和结合律, 所以 n 个数的连乘 $a_1, a_2, \cdots, a_n$ 也可以简单地记作

$$a_1a_2\cdots a_n = \prod_{j=1}^{n} a_j. \tag{1.1.13}$$

这里足码的意义对求和符号 $\sum$ 和求积符号 $\prod$ 是一样的.

和上面一样, 可以引进**二重乘积**, 我们有

$$\prod_{j=1}^{m}\prod_{k=1}^{n} a_{jk} = \prod_{k=1}^{n}\prod_{j=1}^{m} a_{jk} = \prod_{i=2}^{m+n}\prod_{j+k=i} a_{jk}. \tag{1.1.14}$$

当 $m=n$ 时, 我们有

$$\prod_{j=1}^{n}\prod_{k=1}^{n} a_{jk} = \prod_{k=1}^{n}\prod_{j=1}^{n} a_{jk} = \prod_{j,k=1}^{n} a_{jk}. \tag{1.1.15}$$

同样可以用归纳法来定义**多重乘积**

$$\prod_{i_1=1}^{n_1}\prod_{i_2=1}^{n_2}\cdots\prod_{i_p=1}^{n_p} a_{i_1i_2\cdots i_p} = \prod_{i_1=1}^{n_1}\Big(\prod_{i_2=1}^{n_2}\cdots\prod_{i_p=1}^{n_p} a_{i_1i_2\cdots i_p}\Big). \tag{1.1.16}$$

符号 $\prod$ 和符号 $\sum$ 的性质完全相同, 在此就不再仔细讨论了.

现在开始讨论一元多项式. 本书只在**复数域** $\mathbb{C}$, 或**实数域** $\mathbb{R}$, 或**有理数域** $\mathbb{Q}$, 或**整数环** $\mathbb{Z}$ 的范围内进行讨论, 这里 $\mathbb{Z}\subset\mathbb{Q}\subset\mathbb{R}\subset\mathbb{C}$. 为方便起见, 我们记 $\mathbb{F}$ 为 $\mathbb{C}$, 或 $\mathbb{R}$, 或 $\mathbb{Q}$ 之一, 称 $\mathbb{F}$ 为**域**, 而 $\mathbb{Z}$ 不改变符号. 即

$$\mathbb{F}=\mathbb{C}, \quad 或 \quad \mathbb{F}=\mathbb{R}, \quad 或 \quad \mathbb{F}=\mathbb{Q}.$$

定义 1.1.1 给定域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$. 给定正整数 n 及 $n+1$ 个数 $a_0, a_1, \cdots, a_n \in \mathbb{F}$ 或 $\mathbb{Z}$. 记 x 为**未知数**, 或称为**自变量**, 或称为**不定元**, 则**代数式**

$$f = f(x) = a_mx^m + a_{m-1}x^{m-1} + \cdots + a_1x + a_0 = \sum_{j=0}^{m} a_jx^j \tag{1.1.17}$$

称为**一元多项式**, 或称为**多项式**. $a_m, a_{m-1}, \cdots, a_0$ 称为**多项式的系数**. 多项式 $f(x)$ 称为**零多项式**, 如果 $a_m = a_{m-1} = \cdots = a_0 = 0$; 称为**非零多项式**, 如果 $a_m, a_{m-1}, \cdots, a_0$ 不全为零. 域 $\mathbb{F}$ 上多项式全体构成集合 $\mathbb{F}[x]$, 称为**域 $\mathbb{F}$ 上多项式环**; 整数环 $\mathbb{Z}$ 上多项式全体构成集合 $\mathbb{Z}[x]$, 称为 **整数环 $\mathbb{Z}$ 上多项式环**.

任给非零多项式 $f(x) = \sum\limits_{j=0}^{m} a_jx^j$. 由 $a_m, a_{m-1}, \cdots, a_0$ 不全为零, 对系数从左向右看, 总有第一个不等于零的数 a_n, 而 $a_m = \cdots = a_{n+1} = 0$. 我们约定 $0\cdot x^i = 0$,

$i=0, 1, \cdots$. 所以 $f(x)=\sum\limits_{j=0}^{n} a_j x^j$, $a_n \neq 0$. 我们称 a_n 为非零多项式 $f(x)$ 的**首项系数**, a_0 为 **常数项**, a_j 为 x^j 的系数, $0 \leqslant j \leqslant n$. 非负整数 n 称为非零多项式 $f(x)$ 的**多项式的次数**, 记作 $\deg(f(x))$. 它是一个有限数. 我们约定零多项式的首项系数 a_0 为零, 次数为 $-\infty$. 零次多项式 $f(x)=a_0$ 的首项系数 a_0 不为零. 即零次多项式为非零常数.

定义 1.1.2　当域 $\mathbb{F}=\mathbb{C}$ 时, 即 $f(x)$ 为复数域上的多项式时, $f(x)$ 称为**复多项式**; 当域 $\mathbb{F}=\mathbb{R}$ 时, 即 $f(x)$ 为实数域上的多项式时, $f(x)$ 称为**实多项式**; 当域 $\mathbb{F}=\mathbb{Q}$ 时, 即 $f(x)$ 为有理数域上的多项式时, $f(x)$ 称为**有理系数多项式**; 又整数环 $\mathbb{Z}$ 上的多项式称为**整系数多项式**.

域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$ 上两个多项式

$$\begin{aligned} f(x)&=a_n x^n+a_{n-1}x^{n-1}+\cdots+a_1 x+a_0, \quad a_n \neq 0, \\ g(x)&=b_m x^m+b_{m-1}x^{m-1}+\cdots+b_1 x+b_0, \quad b_m \neq 0 \end{aligned} \tag{1.1.18}$$

称为**相等**, 记作 $f(x)=g(x)$, 如果 $m=n$, 而且所有同类项系数相等, 即

$$a_j=b_j, \quad 0 \leqslant j \leqslant n. \tag{1.1.19}$$

下面在域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$ 上多项式之间引进代数运算:

I　加法和减法.

给定域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$ 上多项式

$$f(x)=\sum_{j=0}^{n} a_j x^j, \quad g(x)=\sum_{j=0}^{n} b_j x^j,$$

这里不要求 $a_n \neq 0$ 和 $b_n \neq 0$, 则 $f(x)$ 和 $g(x)$ 的**和**定义为

$$f(x)+g(x)=\sum_{j=0}^{n} a_j x^j+\sum_{j=0}^{n} b_j x^j=\sum_{j=0}^{n}(a_j+b_j)x^j. \tag{1.1.20}$$

由定义可知加法有下面重要性质

(1)　加法结合律:

$$(f(x)+g(x))+h(x)=f(x)+(g(x)+h(x));$$

(2)　加法交换律:

$$f(x)+g(x)=g(x)+f(x);$$

(3)　零多项式 0 有性质: 对任一多项式 $f(x)$,

$$f(x)+0=0+f(x)=f(x);$$

(4) 给定多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$, 定义多项式 $f(x)$ 的**负多项式**为

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

于是有

$$f(x)+(-f(x))=(-f(x))+f(x)=0.$$

因此可以引进多项式的**减法**: 多项式 $f(x)$ 和 $g(x)$ 的**差**定义为

$$f(x)-g(x)=f(x)+(-g(x)). \tag{1.1.21}$$

II 乘法

给定域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$ 上多项式

$$f(x)=\sum_{i=0}^{n}a_ix^i,\quad g(x)=\sum_{j=0}^{m}b_jx^j,\quad a_nb_m\neq 0,$$

则多项式 $f(x)$ 和 $g(x)$ 的**乘积**定义为

$$f(x)g(x)=\Big(\sum_{j=0}^{n}a_jx^j\Big)\Big(\sum_{k=0}^{m}b_kx^k\Big)=\sum_{j=0}^{n}\sum_{k=0}^{m}a_jb_kx^{j+k}=\sum_{i=0}^{m+n}c_ix^i, \tag{1.1.22}$$

其中

$$c_i=\sum_{j+k=i}a_jb_k,\quad 0\leqslant j\leqslant n,\quad 0\leqslant k\leqslant m,\quad 0\leqslant i\leqslant n+m. \tag{1.1.23}$$

由乘法的定义可知, 多项式的乘法有下面重要性质:

(1) **乘法结合律**:

$$(f(x)g(x))h(x)=f(x)(g(x)h(x));$$

(2) **乘法交换律**:

$$f(x)g(x)=g(x)f(x);$$

(3) 零多项式 0 和零次多项式 1 有 $f(x)\cdot 0=0\cdot f(x)=0$, $f(x)\cdot 1=1\cdot f(x)=f(x)$;

(4) 设 $f(x)\neq 0$, 则有**乘法消去律**: $f(x)g(x)=f(x)h(x)$ 当且仅当 $g(x)=h(x)$;

(5) 多项式 $f(x)$ 和 $g(x)$ 有 $f(x)g(x)=0$ 当且仅当 $f(x)=0$, 或 $g(x)=0$;

(6) **加乘分配律**:

$$f(x)(g(x)+h(x))=f(x)g(x)+f(x)h(x).$$

另外有下面次数关系.

引理 1.1.3 域 $\mathbb{F}$ 或整数环 $\mathbb{Z}$ 上多项式 $f(x)$ 和 $g(x)$ 有下面次数关系:

$$\deg(f(x)+g(x))\leqslant\max\left(\deg(f(x)),\deg(g(x))\right), \tag{1.1.24}$$

$$\deg(f(x)g(x)) = \deg(f(x)) + \deg(g(x)). \tag{1.1.25}$$

这包括了多项式 $f(x)$ 或 $g(x)$ 为零多项式的情形, 这时 $\deg(0) = -\infty$.

证 由多项式的加法定义, (1.1.23) 立即成立.

由多项式的乘法定义可知: 设 $\deg(f(x)) = n$, $\deg(g(x)) = m$, 则

$$f(x) = a_n x^n + \cdots + a_0, \quad g(x) = b_m x^m + \cdots + b_0,$$

其中 $a_n b_m \neq 0$. 而乘积

$$f(x)g(x) = (a_n x^n + \cdots + a_0)(b_m x^m + \cdots + b_0) = a_n b_m x^{n+m} + \cdots + a_0 b_0.$$

所以乘积 $f(x)g(x)$ 的首项为 $a_n b_m \neq 0$, 常数项为 $a_0 b_0$. 因此 $f(x)g(x)$ 的次数为 $n+m$. 所以有 $\deg(f(x)g(x)) = n+m = \deg(f(x)) + \deg(g(x))$. 这包括了多项式 $f(x)$ 或 $g(x)$ 为零多项式的情形, 这时 $\deg(0) = -\infty$. □

由上面性质可知, 多项式的加、减、乘和整数的加、减、乘有完全相同的性质, 而且它们都不能随便作除法. 因此研究可除性, 是多项式理论的重要组成部分.

习 题 1.1

1.1.1 试证: 域 $\mathbb{F}$ 和环 $\mathbb{Z}$ 上多项式的乘法结合律和乘法消去律成立.

§1.2 一元多项式的可除性理论

在这一节讨论域 $\mathbb{F}$ ($\mathbb{F}$ 为 $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$ 之一) 上多项式全体构成的域 $\mathbb{F}$ 上多项式环 $\mathbb{F}[x]$. 关于整数环上多项式环 $\mathbb{Z}[x]$, 将在 §1.4 中讨论.

定理 1.2.1(除法公式) 设 $g(x)$ 为域 $\mathbb{F}$ 上的非零多项式. 对任一多项式 $f(x)$, 则唯一存在域 $\mathbb{F}$ 上两多项式 $q(x)$ 和 $r(x)$, 使得

$$f(x) = q(x)g(x) + r(x), \tag{1.2.1}$$

其中 $r(x) = 0$, 或者 $r(x) \neq 0$, 且

$$\deg(r(x)) < \deg(g(x)). \tag{1.2.2}$$

$q(x)$ 称为**商式**, $r(x)$ 称为**余式**. 式 (1.2.1) 又可记作

$$f(x) \equiv r(x) \pmod{g(x)}. \tag{1.2.3}$$

证 先证存在性. 设 $f(x) = 0$, 则取 $q(x) = r(x) = 0$ 就行了. 设 $f(x) \neq 0$, $\deg(g(x)) > \deg(f(x))$, 则取 $q(x) = 0$, $r(x) = f(x)$ 就行了. 设 $f(x) \neq 0$, $\deg(g(x)) \leqslant$

$\deg(f(x))$. 记 $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_0$, $g(x) = b_mx^m + b_{m-1}x^{m-1} + \cdots + b_0$, 其中 $a_nb_m \neq 0$, 又 $n \geqslant m$. 则多项式

$$f_1(x) = f(x) - a_nb_m^{-1}x^{n-m}g(x)$$

的次数小于或等于 $n-1$. 由归纳法假设, 便证明了存在多项式 $q_1(x)$, $r(x)$, 使得

$$f_1(x) = q_1(x)g(x) + r(x),$$

其中 $r(x) = 0$, 或 $r(x) \neq 0$, $\deg(r(x)) < \deg(g(x))$. 因此,

$$f(x) = a_nb_m^{-1}x^{n-m}g(x) + f_1(x) = (a_nb_m^{-1}x^{n-m} + q_1(x))g(x) + r(x).$$

取 $q(x) = a_nb_m^{-1}x^{n-m} + q_1(x)$ 便证明了分解的存在性.

下面证唯一性. 今若 $f(x)$ 有两种分解

$$f(x) = q(x)g(x) + r(x) = q_0(x)g(x) + r_0(x),$$

其中 $r(x) = 0$ 或者 $\deg(r(x)) < \deg(g(x))$, 又 $r_0(x) = 0$ 或者 $\deg(r_0(x)) < \deg(g(x))$. 所以

$$r(x) - r_0(x) = (q_0(x) - q(x))g(x).$$

设 $q_0(x) = q(x)$, 则 $r_0(x) = r(x)$; 设 $q_0(x) \neq q(x)$, 则 $q_0(x) - q(x) \neq 0$, 因此有 $r_0(x) \neq r(x)$, 且

$$\deg(g(x)) > \deg(r(x) - r_0(x)) = \deg(q_0(x) - q(x)) + \deg(g(x)) \geqslant \deg(g(x)).$$

所以导出矛盾. 因此证明了 $r_0(x) = r(x)$, $q_0(x) = q(x)$. □

定义 1.2.2 设 $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上两多项式, 其中 $g(x) \neq 0$. 如果存在域 $\mathbb{F}$ 上多项式 $q(x)$, 使得

$$f(x) = q(x)g(x), \tag{1.2.4}$$

则 $g(x)$ 称为 $f(x)$ 的**因式**, $f(x)$ 称为 $g(x)$ 的**倍式**. 又称 $g(x)$ **除得尽** $f(x)$, 记作 $g(x) \mid f(x)$. 否则称 $g(x)$**除不尽** $f(x)$, 记作 $g(x) \nmid f(x)$.

由定义可知, 任取域 $\mathbb{F}$ 上多项式 $g(x) \neq 0$, 则 $g(x) \mid 0$.

引理 1.2.3 设 $f(x)$, $f_1(x)$, $f_2(x)$, $g(x)$, $h(x)$, $h_1(x)$, $h_2(x)$ 为域 $\mathbb{F}$ 上的多项式, 则除得尽关系有下面性质:

(1) 设 $g(x)h(x) \neq 0$, 且 $h(x) \mid g(x)$, $g(x) \mid f(x)$, 则有 $h(x) \mid f(x)$;

(2) 设 $g(x) \neq 0$, 且 $g(x) \mid f_1(x)$, $g(x) \mid f_2(x)$, 则对任意多项式 $h_1(x)$, $h_2(x)$ 有 $g(x) \mid (h_1(x)f_1(x) + h_2(x)f_2(x))$;

(3) 设 $f(x)g(x) \neq 0$, 且 $f(x) \mid g(x)$, $g(x) \mid f(x)$, 则存在非零常数 $c \in \mathbb{F}$, 使得 $g(x) = cf(x)$. 这时称非零多项式 $f(x)$ 和 $g(x)$ **相伴**;

(4) 设 $f(x) \neq 0$, λ, μ 为域 $\mathbb{F}$ 中非零常数, 则 $f(x)$ 有因式 λ 和 $\mu f(x)$.

证 由除得尽的定义可证 (1), (2), (4) 成立. 下面来证 (3) 成立.

因为 $f(x)g(x) \neq 0$, 且 $g(x) \mid f(x)$, $f(x) \mid g(x)$, 所以存在多项式 $q(x)$ 及 $q_1(x)$, 使得 $f(x) = q(x)g(x)$, $g(x) = q_1(x)f(x)$. 消去 $g(x)$, 有 $f(x) = q(x)q_1(x)f(x)$. 由消去律有 $q(x)q_1(x) = 1$. 由次数公式有 $\deg(q(x)) + \deg(q_1(x)) = 0$. 由于次数为非负整数, 所以 $\deg(q(x)) = \deg(q_1(x)) = 0$, 即 $q_1(x)$ 和 $q(x)$ 为域 $\mathbb{F}$ 中非零常数. 记 $q_1(x) = c$, 有 $g(x) = cf(x)$. □

为了给出域 $\mathbb{F}$ 上多项式的可除性理论, 最重要的是引进和计算两个不全为零的多项式的最高公因式.

定义 1.2.4 **(1)** 给定域 $\mathbb{F}$ 上的多项式 $f_1(x)$ 和 $f_2(x)$, 如果域 $\mathbb{F}$ 上的非零多项式 $g(x)$ 是 $f_1(x)$ 及 $f_2(x)$ 的因式, 则 $g(x)$ 称为 $f_1(x)$ 和 $f_2(x)$ 的**公因式**.

(2) 域 $\mathbb{F}$ 上的多项式 $f_1(x)$ 和 $f_2(x)$ 的公因式 $d(x)$ 称为**最高公因式**, 如果 $f_1(x)$ 和 $f_2(x)$ 的任一公因式都是 $d(x)$ 的因式. 首项系数为 1 的最高公因式记作 $(f_1(x), f_2(x))$, 或 $g.c.d.(f_1(x), f_2(x))$.

由最高公因式的定义可知, 设 $f_1(x)$ 是域 $\mathbb{F}$ 上的非零多项式, 且首项系数为 1, 则 $(f_1(x), 0) = f_1(x)$; $(f_1(x), c) = 1, \forall c \in \mathbb{F}^* = \mathbb{F} - \{0\}$.

引理 1.2.5 给定域 $\mathbb{F}$ 上不全为零的多项式 $f_1(x)$ 和 $f_2(x)$. 如果 $f_1(x)$ 和 $f_2(x)$ 的最高公因式存在, 则必唯一, 且有 $(f_1(x), f_2(x)) = (f_2(x), f_1(x))$.

证 由最高公因式的定义以及引理 1.2.3 的 (4) 立即可知. □

定理 1.2.6 给定域 $\mathbb{F}$ 上不全为零的多项式 $f_1(x)$ 和 $f_2(x)$, 则唯一存在 $f_1(x)$ 和 $f_2(x)$ 的最高公因式 $(f_1(x), f_2(x))$.

证 由引理 1.2.5, 下面只需证明存在性. 为此, 引进**辗转相除法**如下: 为方便起见, 无妨设 $f_2(x) \neq 0$. 于是有

$$
\begin{aligned}
f_1(x) &= q_1(x)f_2(x) + r_1(x), \\
f_2(x) &= q_2(x)r_1(x) + r_2(x), \\
r_1(x) &= q_3(x)r_2(x) + r_3(x), \\
&\ \ \vdots \\
r_t(x) &= q_{t+2}(x)r_{t+1}(x) + r_{t+2}(x), \\
&\ \ \vdots
\end{aligned}
$$

其中 $r_1(x), r_2(x), \cdots, r_t(x), \cdots$ 为非零多项式, 又

$$\deg(f_2(x)) > \deg(r_1(x)) > \deg(r_2(x)) > \cdots > \deg(r_t(x)) > \cdots.$$

但是 $\deg(f_2(x))$ 为有限数, 所以存在正整数 s, 使得下面辗转相除算式成立:

$$\begin{aligned}
f_1(x) &= q_1(x)f_2(x) + r_1(x),\\
f_2(x) &= q_2(x)r_1(x) + r_2(x),\\
r_1(x) &= q_3(x)r_2(x) + r_3(x),\\
&\vdots\\
r_{s-2}(x) &= q_s(x)r_{s-1}(x) + r_s(x),\\
r_{s-1}(x) &= q_{s+1}(x)r_s(x),
\end{aligned} \tag{1.2.5}$$

其中 $r_1(x), r_2(x), \cdots, r_s(x)$ 为非零多项式, 又

$$\deg(f_2(x)) > \deg(r_1(x)) > \deg(r_2(x)) > \cdots > \deg(r_s(x)) \geqslant 0. \tag{1.2.6}$$

记 $r_s(x)$ 的首项系数为 $a \neq 0$, 则有 $(f_1(x), f_2(x)) = a^{-1}r_s(x)$. 事实上, 从上面辗转相除的算式可以看出, 对 $f_1(x)$, $f_2(x)$ 的任一公因式 $h(x)$, 由上往下看, 便证明了 $h(x) \mid r_s(x)$. 再由下往上看, 可知 $r_s(x) \mid r_{s-1}(x), \cdots, r_s(x) \mid r_1(x), r_s(x) \mid f_2(x), r_s(x) \mid f_1(x)$, 即 $r_s(x)$ 为 $f_1(x)$ 和 $f_2(x)$ 的公因式. 由最高公因式的定义可知定理成立. □

辗转相除法还可以导出下面重要性质.

定理 1.2.7 给定域 $\mathbb{F}$ 上两个不全为零的多项式 $f_1(x)$ 和 $f_2(x)$, 则存在域 $\mathbb{F}$ 上的多项式 $u(x)$ 及 $v(x)$, 使得

$$u(x)f_1(x) + v(x)f_2(x) = (f_1(x), f_2(x)). \tag{1.2.7}$$

证 由定理 1.2.6 可知 $(f_1(x), f_2(x)) = a^{-1}r_s(x)$. 在辗转相除公式中, 从上往下依次消去 $r_1(x), r_2(x), \cdots, r_{s-1}(x)$, 便证明了定理. □

定义 1.2.8 **(1)** 给定域 $\mathbb{F}$ 上 m 个不全为零的多项式 $f_1(x), \cdots, f_m(x)$, 域 $\mathbb{F}$ 上的非零多项式 $d(x)$ 称为多项式 $f_1(x), \cdots, f_m(x)$ 的**公因式**, 如果 $d(x) \mid f_1(x), \cdots, d(x) \mid f_m(x)$;

(2) 域 $\mathbb{F}$ 上的非零多项式 $d(x)$ 称为多项式 $f_1(x), \cdots, f_m(x)$ 的**最高公因式**, 如果 $d(x)$ 是多项式 $f_1(x), \cdots, f_m(x)$ 的公因式, 而且任一多项式 $f_1(x), \cdots, f_m(x)$ 的公因式 $d_0(x)$, 有 $d_0(x) \mid d(x)$. 多项式 $f_1(x), \cdots, f_m(x)$ 首项系数为 1 的最高公因式记作 $(f_1(x), \cdots, f_m(x))$, 或记作 $g.c.d.(f_1(x), \cdots, f_m(x))$.

定理 1.2.9 给定域 $\mathbb{F}$ 上 m 个不全为零的多项式 $f_1(x), \cdots, f_m(x)$, 则它们的

首项系数为 1 的最高公因式唯一存在, 且有

$$(f_1(x),\cdots,f_m(x))=(\cdots((f_1(x),f_2(x)),f_3(x)),\cdots,f_m(x)). \tag{1.2.8}$$

证 用归纳法及最高公因式的定义立即可得. □

由定理 1.2.7 和 1.2.9 及归纳法立即有

定理 1.2.10 给定域 $\mathbb{F}$ 上 m 个不全为零的多项式 $f_1(x)$, $f_2(x)$, $\cdots$, $f_m(x)$, 则存在多项式 $u_1(x)$, $u_2(x)$, $\cdots$, $u_m(x)$, 使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_m(x)f_m(x)=(f_1(x),f_2(x),\cdots,f_m(x)). \tag{1.2.9}$$

习 题 1.2

1.2.1 试求下列各组复多项式的最高公因式:

(1) $x^3+(2+i)x^2+(3+2i)x+6,\quad x^5+ix^4+10x^3+28x+21i$;

(2) $x^6+2x^4-4x^3-3x^2+8x-5,\quad x^5+x^2-x+1$.

1.2.2 试利用辗转相除法, 求有理系数多项式 $u(x)$ 和 $v(x)$, 使得 $u(x)f(x)+v(x)g(x)=(f(x),g(x))$, 其中

(1) $f(x)=3x^3-2x^2+x+2,\quad g(x)=x^2-x+1$;

(2) $f(x)=x^4+2x^3-x^2-4x-2,\quad g(x)=x^4+x^3-x^2-2x-2$.

1.2.3 试求域 $\mathbb{F}$ 上一个次数最低的多项式, 使得它被 $x^4-2x^3-2x^2+10x-7$ 除后, 余式为 x^2+x+1; 被 $x^4-2x^3-3x^2+13x-10$ 除后余式为 $2x^2-3$.

1.2.4 试求域 $\mathbb{F}$ 上七次项式 $f(x)$, 使得 $(x-1)^4\mid(f(x)+1)$, $(x+1)^4\mid(f(x)-1)$.

1.2.5 试求域 $\mathbb{F}$ 上适合条件 $(x^2+1)\mid f(x)$, $(x^3+x^2+1)\mid(f(x)+1)$ 的次数最低的多项式 $f(x)$.

1.2.6 求正整数 n, 使得

(1) 域 $\mathbb{F}$ 上多项式 $1+x+x^2+\cdots+x^{n-1}$ 除得尽多项式 $1+x^2+\cdots+x^{2n-2}$;

(2) 域 $\mathbb{F}$ 上多项式 $1+x+x^2+\cdots+x^{n-1}$ 除得尽多项式 $1+x+x^2+\cdots+x^{2n-1}$.

1.2.7 给定域 $\mathbb{F}$ 上两个不全为零的多项式 $f(x)$ 和 $g(x)$, 记

$$F(x)=(x^2+1)f(x)+(x^2+x+1)g(x),\quad G(x)=xf(x)+(x+1)g(x),$$

试证: $(f(x),g(x))=1$ 当且仅当 $(F(x),G(x))=1$.

1.2.8 试求实多项式 $f(x)$ 和 $g(x)$, 使得 $\deg(f(x))\neq\deg(g(x))$, 而且 $f(x)^2-g(x)^2=x^4+x^3+x^2+x+1$.

1.2.9 试求域 $\mathbb{F}$ 上的多项式 $u(x)$ 及 $v(x)$, 使得 $x^m u(x)+(1-x)^n v(x)=1$.

1.2.10 设 $f(x)$, $g(x)$, $h(x)$ 为域 $\mathbb{F}$ 上的非零多项式. 试证: 由 $(f(x),g(x)h(x))=1$ 可推出 $(f(x),g(x))=1$, $(f(x),h(x))=1$.

1.2.11 设 $f(x)$, $g(x)$, $h(x)$ 为域 $\mathbb{F}$ 上的非零多项式, a 为 $h(x)$ 的首项系数. 试证: $(h(x)f(x),h(x)g(x))=a^{-1}h(x)(f(x),g(x))$.

1.2.12 设 $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上的非零多项式, 又 $(f(x),g(x))=1$. 试证: $\forall\, h(x)\in\mathbb{F}[x]$, 有 $(f(x)h(x),g(x))=(h(x),g(x))$.

1.2.13 设 $f(x)$, $g(x)$ 是域 $\mathbb{F}$ 上的非零多项式, 试证: $(f(x),g(x))$ 也是域 $\mathbb{F}$ 上的.

1.2.14 设 $f_1(x)$, $f_2(x)$ 为域 $\mathbb{F}$ 上的非零多项式, 试证: $f_1(x)$ 和 $f_2(x)$ 的最高公因式为 1 的条件等价于 $\forall\, r_1(x)$, $r_2(x)\in\mathbb{F}[x]$, 则存在 $q_1(x)$, $q_2(x)\in\mathbb{F}[x]$, 使得 $q_1(x)f_1(x)+r_1(x)=q_2(x)f_2(x)+r_2(x)$.

1.2.15 设域 $\mathbb{F}$ 上多项式 $p_i(x)$, $0\leqslant i\leqslant n$ 有 $(x-1)\mid p_n(x)$, $\sum\limits_{i=0}^{n-1}x^ip_i(x^{in})=p_n(x^n)$. 试求 $p_i(x)$, $0\leqslant i\leqslant n$.

1.2.16 设域 $\mathbb{F}$ 上多项式 $f(x)$ 被 $x-1$, $x-2$, $x-3$ 除后, 余式分别为 4, 8, 16. 试求 $f(x)$ 被 $(x-1)(x-2)(x-3)$ 除后的余式.

1.2.17 试分别求多项式 x^m-1 和 x^n-1 在有理数域, 实数域, 复数域上多项式环中的最高公因式, 它们间的关系如何?

1.2.18 给定正整数 n, 试证: 存在正整数 m, 使得域 $\mathbb{F}$ 上多项式

$$(1+x)(1+x^2)\cdots(1+x^{2^n})=1+x+x^2+\cdots+x^m.$$

1.2.19 给定域 $\mathbb{F}$ 上两个不全为零的多项式 $f(x)$ 和 $g(x)$. 作多项式集合

$$\mathfrak{S}=\{\,u(x)f(x)+v(x)g(x)\mid u(x),v(x)\in\mathbb{F}[x]\,\},$$

其中 $\mathbb{F}[x]$ 是域 $\mathbb{F}$ 上的多项式环. 试证: 这个集合 $\mathfrak{S}$ 中存在次数最低, 首项系数为 1 的非零多项式, 它就是 $f(x)$ 和 $g(x)$ 的最高公因式 $(f(x),g(x))$. 所以它也是最高公因式的等价定义.

1.2.20 给定域 $\mathbb{F}$ 上三个不全为零的多项式 $a(x)$, $b(x)$, $c(x)$. 试证: 存在域 $\mathbb{F}$ 上六个多项式 $f(x)$, $g(x)$, $h(x)$, $u(x)$, $v(x)$, $w(x)$, 使得三阶行列式

$$\det\begin{pmatrix} a(x) & b(x) & c(x)\\ f(x) & g(x) & h(x)\\ u(x) & v(x) & w(x)\end{pmatrix}=(a(x),b(x),c(x)).$$

(提示: 辗转相除法.)

§1.3 一元多项式的因式分解

在这一节, 我们引进域 $\mathbb{F}$ 上不可约多项式, 证明域 $\mathbb{F}$ 上多项式必可分解为域 $\mathbb{F}$ 上不可约多项式的乘积. 为此, 先引进互素的定义.

定义 1.3.1 域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 称为**互素的**, 如果它们的最大公因式 $(f(x),g(x))=1$. 否则称为**不互素的**.

定理 1.3.2 域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 互素当且仅当存在域 $\mathbb{F}$ 上多项式 $u(x)$ 和 $v(x)$, 使得 $u(x)f(x)+v(x)g(x)=1$.

证 由定理 1.2.7 及互素的定义可知, 当 $(f(x),g(x))=1$ 时, 必存在多项式 $u(x)$ 和 $v(x)$, 使得 $u(x)f(x)+v(x)g(x)=1$. 反之, 如果存在多项式 $u(x)$ 和 $v(x)$, 使得

$u(x)f(x)+v(x)g(x)=1$. 记 $d(x)=(f(x),g(x))$, 则有 $d(x)\mid(u(x)f(x)+v(x)g(x))$, 即 $d(x)\mid 1$, 这证明了 $d(x)=1$, 即 $(f(x),g(x))=1$. 所以 $f(x)$ 和 $g(x)$ 互素. □

引理 1.3.3 域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 及多项式 $h(x)$ 若有 $(f(x),g(x))=1$, $f(x)\mid g(x)h(x)$, 则有 $f(x)\mid h(x)$.

证 由定理 1.3.2, 存在多项式 $u(x)$ 和 $v(x)$, 使得 $u(x)f(x)+v(x)g(x)=1$. 于是 $h(x)=(u(x)h(x))f(x)+v(x)(g(x)h(x))$. 由条件 $f(x)\mid g(x)h(x)$, 便证明了 $f(x)\mid h(x)$. 所以引理成立. □

为了 §4.4 的需要, 现在给出

定理 1.3.4 域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 不互素的充分且必要条件为存在非零多项式 $u(x)$ 和 $v(x)$, 使得

$$u(x)f(x)=v(x)g(x), \tag{1.3.1}$$

其中

$$\deg(u(x))<\deg(g(x)),\quad \deg(v(x))<\deg(f(x)). \tag{1.3.2}$$

证 记 $(f(x),g(x))=d(x)$. 设多项式 $f(x)$ 和 $g(x)$ 不互素, 于是 $\deg(d(x))>0$. 记 $f(x)=d(x)v(x)$, $g(x)=d(x)u(x)$, 则有 $\deg(v(x))<\deg(f(x))$, $\deg(u(x))<\deg(g(x))$. 令 $u(x)f(x)=d(x)u(x)v(x)=v(x)g(x)$, 所以式 (1.3.1) 和 (1.3.2) 成立.

反之, 若存在非零多项式 $u(x)$ 和 $v(x)$, 使得

$$u(x)f(x)=v(x)g(x),\quad \deg(u(x))<\deg(g(x)),\quad \deg(v(x))<\deg(f(x)).$$

下面证 $f(x)$ 和 $g(x)$ 两多项式不互素. 我们用反证法. 若 $f(x)$ 和 $g(x)$ 互素, 即 $(f(x),g(x))=1$. 由 $f(x)\mid v(x)g(x)$ 及引理 1.3.3, 所以证明了 $f(x)\mid v(x)$, 这和 $\deg(v(x))<\deg(f(x))$ 矛盾. □

由引理 1.2.3, $f|g$, $g|f$ 当且仅当 f 和 g 相伴, 所以我们引进如下定义.

定义 1.3.5 域 $\mathbb{F}$ 上非零多项式 $p(x)$ 称为 **不可约多项式**, 如果除了因式 λ 和 $\mu p(x)$ 外, 无其他域 $\mathbb{F}$ 上因式, 其中 λ 和 μ 为域 $\mathbb{F}$ 中的非零数. 又, 不是不可约的多项式称为**可约多项式**.

注意 显然域 $\mathbb{F}$ 上零次多项式和一次多项式都是不可约的. 又域 $\mathbb{F}$ 上可约的多项式是两个域 $\mathbb{F}$ 上低次多项式的乘积. 为什么?

对域 $\mathbb{F}$ 上相伴的两非零多项式 $f(x)$ 和 $g(x)$, 利用引理 1.2.3 的**(4)**, 我们可以约定其中有一个非零多项式的首项系数为 1. 特别, 我们可以取不可约多项式 $p(x)$ 的首项系数为 1.

引理 1.3.6 设 $p(x)$ 为域 $\mathbb{F}$ 上的不可约多项式, 则对域 $\mathbb{F}$ 上任一多项式 $f(x)$,

有 $p(x) \mid f(x)$, 或者有 $(p(x), f(x)) = 1$. 特别, 任意两个首项系数为 1 的不可约多项式 $p(x)$ 和 $q(x)$, 必有 $p(x) = q(x)$, 或 $(p(x), q(x)) = 1$.

证 记 $(p(x), f(x)) = d(x)$. 设 $(p(x), f(x)) = 1$, 则不必证了. 设 $\deg(d(x)) > 0$, 则 $d(x) \mid p(x)$. 但是 $p(x)$ 是不可约多项式, 所以 $d(x) = \mu p(x)$, 其中 μ^{-1} 为 $p(x)$ 的首项系数. 因此由 $d(x) \mid f(x)$ 便证明了 $p(x) \mid f(x)$. □

引理 1.3.7 设 $p(x)$ 为域 $\mathbb{F}$ 上的不可约多项式, 则对域 $\mathbb{F}$ 上任两多项式 $f(x)$ 和 $g(x)$, 只要 $p(x) \mid f(x)g(x)$, 便有 $p(x) \mid f(x)$, 或者 $p(x) \mid g(x)$.

证 若 $p(x) \mid g(x)$, 则不必证了. 若 $p(x) \nmid g(x)$, 则 $(p(x), g(x)) = 1$. 由引理 1.3.6, 今 $p(x) \mid f(x)g(x)$, 所以 $p(x) \mid f(x)$. □

由归纳法立即有

引理 1.3.8 设 $p(x)$ 为域 $\mathbb{F}$ 上不可约多项式, $f_1(x)$, $f_2(x)$, $\cdots$, $f_m(x)$ 为域 $\mathbb{F}$ 上非零多项式. 设 $p(x) \mid f_1(x)f_2(x)\cdots f_m(x)$, 则 $p(x)$ 必除得尽 $f_1(x)$, $f_2(x)$, $\cdots$, $f_m(x)$ 中某个多项式.

定理 1.3.9 (唯一析因定理) 设 $f(x)$ 为域 $\mathbb{F}$ 上次数大于零的多项式, 则 $f(x)$ 有下列因式分解:

$$f(x) = a_0 p_1(x) p_2(x) \cdots p_r(x), \tag{1.3.3}$$

其中 a_0 是 $f(x)$ 的首项系数, $p_1(x)$, $p_2(x)$, $\cdots$, $p_r(x)$ 为域 $\mathbb{F}$ 上次数大于零, 且首项系数为 1 的不可约多项式. 若另有因式分解

$$f(x) = a_0 q_1(x) q_2(x) \cdots q_t(x), \tag{1.3.4}$$

其中 $q_1(x)$, $q_2(x)$, $\cdots$, $q_t(x)$ 为域 $\mathbb{F}$ 上次数大于零, 且首项系数为 1 的不可约多项式, 则 $t = r$, 且存在 1, 2, $\cdots$, r 的排列 $i_1 i_2 \cdots i_r$, 使得

$$q_j(x) = p_{i_j}(x), \quad 1 \leqslant j \leqslant r. \tag{1.3.5}$$

又, $f(x)$ 的次数大于零的因式必为

$$\lambda p_{j_1}(x) p_{j_2}(x) \cdots p_{j_s}(x), \tag{1.3.6}$$

其中 λ 为域 $\mathbb{F}$ 中非零数, $j_1, j_2, \cdots, j_s$ 为 1, 2, $\cdots$, r 中 s 个不同的数.

证 先证因式分解的存在性. 设 $f(x)$ 不可约, 则定理成立. 设 $f(x)$ 可约, 由定义可知 $f(x) = f_1(x)f_2(x)$, 其中 $f_1(x)$ 和 $f_2(x)$ 的次数都小于 $f(x)$ 的次数. 对次数作归纳法, 因此可对 $f_1(x)$ 和 $f_2(x)$ 用归纳法假设, 所以证明了因式分解的存在性.

再证唯一性. 由式 (1.3.3) 和 (1.3.4), 我们有

$$p_1(x) p_2(x) \cdots p_r(x) = q_1(x) q_2(x) \cdots q_t(x).$$

为讨论简便起见, 我们无妨设 $r \leqslant t$. 今由 $q_1(x) \mid p_1(x)p_2(x)\cdots p_r(x)$ 及引理 1.3.8, 所以存在 $p_{i_1}(x)$ 使得 $q_1(x) \mid p_{i_1}(x)$. 但是 $q_1(x)$ 和 $p_{i_1}(x)$ 都是次数大于零, 且首项系数为 1 的不可约多项式, 这证明了 $q_1(x) = p_{i_1}(x)$. 代回原式, 由消去律便得到

$$p_2(x)\cdots p_r(x) = q_1(x)q_2(x)\cdots q_{i_1-1}(x)q_{i_1+1}(x)\cdots q_t(x).$$

这样依次讨论下去, 便证明了存在 1, 2, $\cdots$, t 的排列 $i_1 i_2 \cdots i_t$, 使得 $p_j(x) = q_{i_j}(x)$, $1 \leqslant j \leqslant r$, 又

$$1 = q_{i_{r+1}}(x)\cdots q_{i_t}(x).$$

由于 $\deg(q_i(x)) > 0$, $1 \leqslant i \leqslant t$, 这证明了 $r = t$, 于是证明了分解的唯一性.

最后, 设 $g(x)$ 为域 $\mathbb{F}$ 上次数大于零, 且首项系数为 1 的多项式, 且有 $g(x) \mid f(x)$. 将 $g(x)$ 分解因式为

$$g(x) = r_1(x)r_2(x)\cdots r_s(x),$$

其中 $r_1(x)$, $r_2(x)$, $\cdots$, $r_s(x)$ 为次数大于零, 且首项系数为 1 的不可约多项式. 由 $r_1(x) \mid g(x)$, $g(x) \mid f(x)$, 所以 $r_1(x) \mid f(x)$, 即 $r_1(x) \mid a_0p_1(x)p_2(x)\cdots p_r(x)$. 和唯一性证明相同, 有 $r_1(x) = p_{j_1}(x)$. 再依次讨论下去, 最后便证明了

$$g(x) = p_{j_1}(x)p_{j_2}(x)\cdots p_{j_s}(x),$$

其中 j_1, j_2, $\cdots$, j_s 为 1, 2, $\cdots$, r 中 s 个不同数. □

下面讨论重因式和公倍式.

定义 1.3.10 域 $\mathbb{F}$ 上 n 次多项式

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \tag{1.3.7}$$

的导数定义为域 $\mathbb{F}$ 上多项式

$$f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1. \tag{1.3.8}$$

设 k 为正整数, 则用归纳法, 我们可以引进 k 次导数 $f^{(k)}(x)$. 它定义为:

$$f^{(k)}(x) = (f^{(k-1)})'(x), \quad k = 1, 2, \cdots, \quad f^{(0)}(x) = f(x). \tag{1.3.9}$$

显然 $f^{(k)}(x) = 0$, $n < k$. 由定义出发, 可以证明

引理 1.3.11 设 $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上多项式, 则导数有性质

(1) 设 $f(x)$ 为域 $\mathbb{F}$ 上多项式, 则 $f'(x) = 0$ 当且仅当 $f(x)$ 为域 $\mathbb{F}$ 中常数;

(2) $(f(x) + g(x))' = f'(x) + g'(x)$;

(3) $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

因此有

$$(f_1(x)f_2(x)\cdots f_m(x))'=\sum_{j=1}^{m}f_1(x)f_2(x)\cdots f_{j-1}(x)f_j'(x)f_{j+1}(x)\cdots f_m(x). \tag{1.3.10}$$

所以当 $m\geqslant 1$ 时,

$$(f(x)^m)'=mf(x)^{m-1}f'(x). \tag{1.3.11}$$

证 由导数的定义, 所以 (1) 和 (2) 成立. 下面证 (3) 成立. 事实上, 记 $f(x)=\sum\limits_{i=0}^{n}a_ix^i$, $g(x)=\sum\limits_{j=0}^{m}b_jx^j$, 则 $f(x)g(x)=\sum\limits_{k=0}^{n+m}(\sum\limits_{i+j=k}a_ib_j)x^k$, 因此

$$\begin{aligned}(f(x)g(x))'&=\sum_{k=0}^{n+m}\sum_{i+j=k}(i+j)a_ib_jx^{i+j-1}\\&=\sum_{k=0}^{n+m}\sum_{i+j=k}ia_ib_jx^{i+j-1}+\sum_{k=0}^{n+m}\sum_{i+j=k}ja_ib_jx^{i+j-1}\\&=\sum_{i=0}^{n}ia_ix^{i-1}\sum_{j=0}^{m}b_jx^j+\sum_{i=0}^{n}a_ix^i\sum_{j=0}^{m}jb_jx^{j-1}\\&=f'(x)g(x)+f(x)g'(x).\end{aligned}$$

用归纳法, 立即可证式 (1.3.10) 成立. 令 $f_1(x)=\cdots=f_m(x)=f(x)$, 则式 (1.3.11) 成立. □

定义 1.3.12 设 $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上次数大于零的多项式, m 为正整数. 如果 $g(x)^m\mid f(x)$, $g(x)^{m+1}\nmid f(x)$, 则称 $g(x)$ 为 $f(x)$ 的 m **重因式**, 简称为 $f(x)$ 有 m 重因式 $g(x)$. 一重因式也称为**单重因式**.

利用导数概念, 可以讨论多项式的重因式.

定理 1.3.13 域 $\mathbb{F}$ 上次数大于零的多项式 $f(x)$ 没有重因式的充分且必要条件为

$$(f(x),f'(x))=1. \tag{1.3.12}$$

证 由唯一析因定理可知

$$f(x)=a_0p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_s(x)^{e_s}, \tag{1.3.13}$$

其中 a_0 是 $f(x)$ 的首项系数, $p_1(x)$, $p_2(x)$, $\cdots$, $p_s(x)$ 为域 $\mathbb{F}$ 上次数大于零, 且首项系数为 1 的 s 个不同的不可约多项式, e_1, e_2, $\cdots$, e_s 为正整数. 由式 (1.3.10), 有

$$\begin{aligned}f'(x)&=a_0\sum_{j=1}^{s}e_jp_j'(x)p_1(x)^{e_1}\cdots p_{j-1}(x)^{e_{j-1}}p_j(x)^{e_j-1}p_{j+1}(x)^{e_{j+1}}\cdots p_s(x)^{e_s}\\&=a_0p_1^{e_1-1}(x)\cdots p_s(x)^{e_s-1}\sum_{j=1}^{s}e_jp_1(x)\cdots p_{j-1}(x)p_j'(x)p_{j+1}(x)\cdots p_s(x).\end{aligned} \tag{1.3.14}$$

今由 $\deg(p_j'(x)) = \deg(p_j(x)) - 1$, $1 \leqslant j \leqslant s$, 不难证明

$$(f(x), f'(x)) = p_1(x)^{e_1-1} \cdots p_s(x)^{e_s-1}. \tag{1.3.15}$$

由 $f(x)$ 的分解式可知, $f(x)$ 有重因式当且仅当存在 j, 使得 $e_j > 1$. 所以 $f(x)$ 无重因式当且仅当 $e_1 = e_2 = \cdots = e_s = 1$, 即 $(f(x), f'(x)) = 1$. □

定理 1.3.14 设 $f(x)$ 为域 $\mathbb{F}$ 上次数大于零的多项式, 则

$$g(x) = (f(x), f'(x))^{-1} f(x) \tag{1.3.16}$$

仍为域 $\mathbb{F}$ 上的次数大于零的多项式, 而且无重因式. 但是 $g(x)$ 和 $f(x)$ 有完全一样的不同不可约因式组.

证 由式 (1.3.15) 及 (1.3.16) 可知, $g(x)$ 无重因式, 而且 $g(x)$ 和 $f(x)$ 有完全一样的不同不可约因式组. 由于 $f(x)$, $f'(x) \in \mathbb{F}[x]$, 而最高公因式是由辗转相除法的算式算出来的. 所以 $(f(x), f'(x)) \in \mathbb{F}[x]$. 这证明了 $g(x) \in \mathbb{F}[x]$. □

设 $f(x)$ 为实多项式, 则导数还有性质

(a) 若实数 x_0 有 $f'(x_0) = 0$, $f''(x_0) > 0\,(< 0)$, 则曲线 $y = f(x)$ 在 $x = x_0$ 达到极小 (大) 值;

(b) 若在开区间 (a, b) 中有 $f'(x) > 0\,(< 0)$, 则 $f(x)$ 在区间 (a, b) 中严格单调递增 (递减);

(c) 若在开区间 (c, d) 中有两实根 $c < a < b < d$, 即 $f(a) = f(b) = 0$, 则 $f'(x)$ 在闭区间 $[a, b]$ 中必有实根.

这些性质的证明会在分析课中学到.

和最高公因式相类似的是最低公倍式的概念.

定义 1.3.15 **(1)** 域 $\mathbb{F}$ 上多项式 $m_0(x)$ 称为域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 的**公倍式**, 如果 $f(x) \mid m_0(x)$, $g(x) \mid m_0(x)$;

(2) 域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 的公倍式 $m(x)$ 称为 **最低公倍式**, 如果对 $f(x)$ 和 $g(x)$ 的任一公倍式 $m_0(x)$, 都有 $m(x) \mid m_0(x)$. 首项系数为 1 的最低公倍式记作 $[f(x), g(x)]$, 或记作 $l.c.m.[f(x), g(x)]$.

定理 1.3.16 给定域 $\mathbb{F}$ 上首项系数为 1 的非零多项式 $f(x)$ 和 $g(x)$, 则唯一存在最低公倍式 $[f(x), g(x)]$, 它有

$$f(x)g(x) = (f(x), g(x))[f(x), g(x)]. \tag{1.3.17}$$

证 记 $d(x) = (f(x), g(x))$, 于是

$$f(x) = d(x)f_1(x), \quad g(x) = d(x)g_1(x),$$

其中 $f_1(x)$ 和 $g_1(x)$ 为域 $\mathbb{F}$ 上首项系数为 1 的非零多项式, 它们有 $(f_1(x), g_1(x)) = 1$. 于是记

$$m(x) = \frac{f(x)g(x)}{(f(x), g(x))} = d(x)f_1(x)g_1(x) = f(x)g_1(x) = f_1(x)g(x),$$

则有 $f(x) \mid m(x)$, $g(x) \mid m(x)$, 所以 $m(x)$ 为 $f(x)$ 和 $g(x)$ 的公倍式, 且首项系数为 1. 另一方面, 设 $m_0(x)$ 为 $f(x)$ 和 $g(x)$ 的公倍式, 则 $m_0(x) = f(x)h_1(x) = g(x)h_2(x)$, 于是

$$m_0(x) = d(x)f_1(x)h_1(x) = d(x)g_1(x)h_2(x).$$

这证明了 $f_1(x)h_1(x) = g_1(x)h_2(x)$. 由 $(f_1(x), g_1(x)) = 1$ 和引理 1.3.3, 这就推出了 $f_1(x) \mid h_2(x)$, $g_1(x) \mid h_1(x)$. 所以 $d(x)f_1(x)g_1(x) \mid m_0(x)$. 由 $m(x) = d(x)f_1(x)g_1(x)$ 可知 $m(x) \mid m_0(x)$. 由定义可知 $m(x)$ 为 $f(x)$ 和 $g(x)$ 的最低公倍式, 且首项系数为 1. 即 $m(x) = [f(x), g(x)]$. □

定义 1.3.17 **(1)** 域 $\mathbb{F}$ 上非零多项式 $m_0(x)$ 称为域 $\mathbb{F}$ 上 s 个非零多项式 $f_1(x)$, $f_2(x), \cdots, f_s(x)$ 的**公倍式**, 如果 $f_i(x) \mid m_0(x)$, $1 \leqslant j \leqslant s$;

(2) 域 $\mathbb{F}$ 上 s 个非零多项式 $f_1(x), f_2(x), \cdots, f_s(x)$ 的公倍式 $m(x)$ 称为**最低公倍式**, 如果 $f_1(x), f_2(x), \cdots, f_s(x)$ 的任一公倍式 $m_0(x)$ 必为 $m(x)$ 的倍式. 首项系数为 1 的最低公倍式记作 $[f_1(x), f_2(x), \cdots, f_s(x)]$, 或记作 $l.c.m.[f_1(x), f_2(x), \cdots, f_s(x)]$.

计算域 $\mathbb{F}$ 上 s 个非零多项式 $f_1(x), f_2(x), \cdots, f_s(x)$ 的最低公倍式, 可用下面的定理.

定理 1.3.18 对域 $\mathbb{F}$ 上 s 个非零多项式 $f_1(x), f_2(x), \cdots, f_s(x)$, 则有

$$[f_1(x), \cdots, f_s(x)] = [\cdots[[f_1(x), f_2(x)], f_3(x)], \cdots, f_m(x)]. \tag{1.3.18}$$

证 用归纳法及最低公倍式的定义立即可证. □

习 题 1.3

1.3.1 设域 $\mathbb{F}$ 上非零多项式 $f(x)$ 和 $g(x)$ 互素. 试证: 唯一存在多项式 $u(x)$ 和 $v(x)$, 适合条件 $\deg(u(x)) < \deg(g(x))$ 或 $\deg(v(x)) < \deg(f(x))$, 使得 $u(x)f(x) + v(x)g(x) = 1$.

1.3.2 设 $p(x)$ 为域 $\mathbb{F}$ 上非零多项式, 试证下面条件互相等价:

(1) $p(x)$ 为不可约多项式;

(2) 由 $p(x) \mid f(x)g(x)$ 推出 $p(x) \mid f(x)$ 或 $p(x) \mid g(x)$, $\forall f(x), g(x) \in \mathbb{F}[x]$;

(3) 任取域 $\mathbb{F}$ 上多项式 $f(x)$, 则 $p(x) \mid f(x)$ 或 $(p(x), f(x)) = 1$.

1.3.3 设 $f(x)$, $g(x)$ 和 $h(x)$ 为域 $\mathbb{F}$ 上非零多项式, 试证:

$$[f(x), (g(x), h(x))] = ([f(x), g(x)], [f(x), h(x)]).$$

1.3.4 设 $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上两非零多项式. 试证: 任取正整数 m, 则 $(f(x),g(x))^m = (f(x)^m, g(x)^m)$.

1.3.5 任意给定正整数 $n > 1$. 设 $f_i(x) = a_i x + b_i$, $i = 1, 2, 3$ 为非零实多项式 (不要求 $a_i \neq 0$), 有 $f_1(x)^n + f_2(x)^n = f_3(x)^n$. 试证: 存在实多项式 $f(x) = ax + b$, 使得 $f_i(x) = c_i f(x)$, 其中 $c_1, c_2, c_3 \in \mathbb{R}$.

1.3.6 域 $\mathbb{F}$ 上多项式环 $\mathbb{F}[x]$ 到域 $\mathbb{F}$ 的映射 $f(x) \to \mathfrak{D}(f(x))$, $\forall f(x) \in \mathbb{F}[x]$, 它适合条件: 给定常数 $a \in \mathbb{F}$,

(1) $\mathfrak{D}(\lambda f(x) + \mu g(x)) = \lambda \mathfrak{D}(f(x)) + \mu \mathfrak{D}(g(x))$;

(2) $\mathfrak{D}(f(x)g(x)) = f(a)\mathfrak{D}(g(x)) + g(a)\mathfrak{D}(f(x))$,

其中 $f(x), g(x) \in \mathbb{F}[x]$, $\lambda, \mu \in \mathbb{F}$. 试证: 存在常数 $b \in \mathbb{F}$, 使得 $\mathfrak{D}(f(x)) = bf'(a)$.

§1.4 一元整系数多项式

对复多项式环 $\mathbb{C}[x]$, 实多项式环 $\mathbb{R}[x]$, 有理系数多项式环 $\mathbb{Q}[x]$, 它们和整系数多项式环 $\mathbb{Z}[x]$ 有本质上的不同, 差别在于不可约多项式的定义. 我们知道, 对域 $\mathbb{F}$ 上多项式环 $\mathbb{F}[x]$, 求平凡公因子的方法是: 设 $f(x)$ 和 $g(x)$ 为非零多项式, 而且有 $f(x) \mid g(x)$, $g(x) \mid f(x)$, 即 $f(x)$ 和 $g(x)$ 是相伴的, 这意味着 $f(x)$ 和 $g(x)$ 关于相除是等价的. 这时存在 $c \in \mathbb{F}$, 使得 $g(x) = cf(x)$. 但是对整系数多项式环 $\mathbb{Z}[x]$, 我们有

引理 1.4.1 设 $f(x)$ 和 $g(x)$ 为两个非零整系数多项式, 而且适合条件 $f(x) \mid g(x)$, $g(x) \mid f(x)$, 即 $f(x)$ 和 $g(x)$ 是相伴的, 则 $g(x) = f(x)$ 或 $g(x) = -f(x)$. 因此 $f(x)$ 和 $g(x)$ 是相伴的充分且必要条件为 $g(x) = \pm f(x)$.

证 令 $f(x) \mid g(x)$, $g(x) \mid f(x)$, 即 $g(x) = f(x)g_1(x)$, $f(x) = g(x)f_1(x)$, 其中 $f(x)$, $f_1(x)$, $g(x)$, $g_1(x) \in \mathbb{Z}[x]$. 所以 $f(x) = g(x)f_1(x) = (f(x)g_1(x))f_1(x)$. 由乘法消去律, 因此 $g_1(x)f_1(x) = 1$. 这证明了 $\deg(f_1(x)) + \deg(g_1(x)) = 0$. 所以有 $\deg(f_1(x)) = \deg(g_1(x)) = 0$. 因此存在整数 a 和 b, 使得 $f_1(x) = b$, $g_1(x) = a$, 又 $ab = 1$. 所以 $a = b = \pm 1$, 即 $g(x) = \pm f(x)$. □

定义 1.4.2 非零整系数多项式 $p(x)$ 称为**不可约多项式**, 如果除了 ± 1 和 $\pm p(x)$ 以外, 多项式 $p(x)$ 无其他因式. 又, 不是不可约的多项式称为**可约多项式**.

由定义可知, $\mathbb{Z}[x]$ 中零次多项式不可约当且仅当它是**素数**. 即**合数**是 $\mathbb{Z}[x]$ 中零次可约多项式.

仅管如此, 我们仍可藉助于有理系数多项式环 $\mathbb{Q}[x]$ 上的可除性理论, 来给出整系数多项式环 $\mathbb{Z}[x]$ 上的可除性理论, 这里 $\mathbb{Z}[x] \subset \mathbb{Q}[x]$.

定义 1.4.3 非零整系数多项式

$$g(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 \tag{1.4.1}$$

称为**本原多项式**, 如果 $(c_0, c_1, \cdots, c_n) = 1$, 即 $c_0, c_1, \cdots, c_n$ 互素.

用这个定义, 我们有: 不可约多项式必为本原多项式. 但是次数大于零的整系数可约多项式不一定是两个低次本原多项式的乘积, 所以整系数多项式的性质和域 $\mathbb{F}$ 上多项式的性质有些不同. 下面先给出有理系数多项式和本原多项式间的关系. 我们有

引理 1.4.4 任一非零有理系数多项式为一个有理数和一个本原多项式的乘积.

证 设 $f(x) \neq 0$ 为有理系数多项式, 将所有系数的分母通分, 便可写成

$$f(x) = \frac{a}{b} g(x), \tag{1.4.2}$$

其中 $g(x)$ 为非零整系数多项式, 且为本原多项式. 又, a 和 b 为整数, 且 $b > 0$, $(a, b) = 1$. 即 a/b 为既约分数. □

由引理 1.4.4 可知, 有理系数多项式的因式分解和整系数多项式的因式分解密切相关. 关于整系数多项式分解为整系数多项式乘积的因式分解理论中, 重要的是

引理 1.4.5 (Gauss 引理) 任两本原多项式的乘积仍为本原多项式.

证 给定两个本原多项式 $f(x) = \sum\limits_{i=0}^{n} a_i x^i$, $g(x) = \sum\limits_{j=0}^{m} b_j x^j$. 为了书写方便起见, 当 $q > n$ 时, 记 $a_q = 0$; 当 $u > m$ 时, 记 $b_u = 0$. 记

$$f(x)g(x) = \sum_{k=0}^{n+m} c_k x^k, \tag{1.4.3}$$

其中

$$c_k = \sum_{i+j=k} a_i b_j, \quad 0 \leqslant i \leqslant n, \quad 0 \leqslant j \leqslant m, \quad 0 \leqslant k \leqslant n+m. \tag{1.4.4}$$

下面要证明 $(c_0, \cdots, c_{n+m}) = 1$. 为此用反证法. 设存在素数 p, 使得 $p \mid c_k$, $0 \leqslant k \leqslant n+m$. 今 $c_0 = a_0 b_0$. 由 $p \mid c_0$, 所以 $p \mid a_0$ 或 b_0. 无妨设 $p \mid a_0$. 由 $(a_0, \cdots, a_n) = 1$, 所以存在指标 $r \in \{0, 1, \cdots, n-1\}$, 使得 $p \mid a_0, \cdots, p \mid a_r, p \nmid a_{r+1}$. 但是

$$c_{r+1} = a_{r+1} b_0 + a_r b_1 + \cdots + a_1 b_r + a_0 b_{r+1}.$$

今 $p \mid c_{r+1}$, $p \mid a_0, \cdots, p \mid a_r$, 所以 $p \mid a_{r+1} b_0$. 但是 $p \nmid a_{r+1}$, 因此 $p \mid b_0$. 由 $(b_0, \cdots, b_m) = 1$, 所以存在指标 $s \in \{0, 1, \cdots, m-1\}$, 使得 $p \mid b_0, \cdots, p \mid b_s, p \nmid b_{s+1}$. 由

$$c_{r+s+2} = a_{r+s+2} b_0 + a_{r+s+1} b_1 + \cdots + a_{r+2} b_s + a_{r+1} b_{s+1} + a_r b_{s+2} + \cdots + a_0 b_{r+s+2}.$$

今 $p \mid c_{r+s+2}$, $p \mid b_0$, $\cdots$, $p \mid b_s$, $p \mid a_0$, $\cdots$, $p \mid a_r$, 所以 $p \mid a_{r+1}b_{s+1}$. 这和 $p \nmid b_{s+1}$, $p \nmid a_{r+1}$ 矛盾. □

引理 1.4.6 任一整系数多项式必为一个整数和一个本原多项式的乘积. 所以次数大于零的整系数不可约多项式必为本原多项式, 次数等于零的整系数不可约多项式为 ± 1, 或为素数 $\pm p$.

证明 由引理 1.4.5 可知引理成立. □

定理 1.4.7 设整系数本原多项式的次数大于零. 则它不可约当且仅当它作为有理系数多项式仍不可约.

证 设本原多项式 $f(x)$ 作为有理系数多项式, 它不可约. 自然它作为整系数多项式也不可约. 反之, 若本原多项式 $f(x)$ 作为整系数多项式, 它不可约, 但作为有理系数多项式, 它可约. 于是存在有理系数多项式 $g(x)$ 和 $h(x)$, 使得 $\deg(g(x))$, $\deg(h(x)) < \deg(f(x))$, 又 $f(x) = g(x)h(x)$. 由引理 1.4.4, 存在有理数 λ 和 μ 以及本原多项式 $g_1(x)$ 和 $h_1(x)$, 使得 $g(x) = \lambda g_1(x)$, $h(x) = \mu h_1(x)$. 因此

$$f(x) = g(x)h(x) = \lambda\mu g_1(x)h_1(x).$$

由引理 1.4.5, $g_1(x)h_1(x)$ 仍为本原多项式. 今 $f(x)$ 为本原多项式, 这证明了 $\lambda\mu = \pm 1$. 所以作为整系数多项式, $f(x) = \pm g_1(x)h_1(x)$, 即 $f(x)$ 可约. 这和假设矛盾, 所以证明了定理. □

定理 1.4.8 (唯一析因定理) 设 $f(x)$ 为整数环 $\mathbb{Z}$ 上次数大于零, 且首项系数大于零的多项式, 则 $f(x)$ 有下列因式分解

$$f(x) = p_1(x)p_2(x)\cdots p_s(x), \tag{1.4.5}$$

其中 $p_1(x)$, $p_2(x)$, $\cdots$, $p_s(x)$ 为整系数多项式环 $\mathbb{Z}[x]$ 中首项系数大于零的不可约多项式. 若另有因式分解

$$f(x) = q_1(x)q_2(x)\cdots q_t(x), \tag{1.4.6}$$

其中 $q_1(x)$, $q_2(x)$, $\cdots$, $q_t(x)$ 为整系数多项式环 $\mathbb{Z}[x]$ 中首项系数大于零的不可约多项式, 则 $t = s$, 且存在 $1, 2, \cdots, s$ 的排列 $i_1 i_2 \cdots i_s$, 使得

$$q_j(x) = p_{i_j}(x), \quad 1 \leqslant j \leqslant s. \tag{1.4.7}$$

又 $f(x)$ 的因式必为

$$\pm p_{j_1}(x)p_{j_2}(x)\cdots p_{j_t}(x), \tag{1.4.8}$$

其中 $j_1, j_2, \cdots, j_t$ 为此 $1, 2, \cdots, s$ 中 t 个不同数, 所以 $1 \leqslant t \leqslant s$.

证 今整系数多项式环 $\mathbb{Z}[x] \subset \mathbb{Q}[x]$. 由定理 1.3.9, 存在次数大于零的有理系数不可约多项式 $\widetilde{p}_1(x)$, $\widetilde{p}_2(x)$, $\cdots$, $\widetilde{p}_r(x)$, 使得 $f(x) = \widetilde{p}_1(x)\widetilde{p}_2(x)\cdots\widetilde{p}_r(x)$. 由引理 1.4.4,

存在有理数 λ_j 和次数大于零的整系数不可约多项式 $p_j(x)$, 使得 $\widetilde{p}_j(x)=\lambda_j p_j(x)$, $1\leqslant j\leqslant r$. 代入有

$$f(x)=\lambda_1\cdots\lambda_r p_1(x)p_2(x)\cdots p_r(x).$$

由引理 1.4.5, $p_1(x)p_2(x)\cdots p_r(x)$ 仍为本原多项式, 由 $f(x)\in\mathbb{Z}[x]$, 所以有理数 $\lambda=\lambda_1\cdots\lambda_r\in\mathbb{Z}$, 而且 $\lambda>0$. 由整数环 $\mathbb{Z}$ 的唯一析因定理, 所以

$$\lambda=p_{r+1}(x)\cdots p_s(x),$$

其中 $p_j(x)\in\mathbb{Z}[x]$, 因此 $\deg(p_j(x))=0$, 所以 $p_j(x)$ 为素数, $r+1\leqslant j\leqslant s$.

$$f(x)=p_1(x)p_2(x)\cdots p_r(x)p_{r+1}(x)\cdots p_s(x),$$

其中 $p_k(x)$ 为首项系数大于零的整系数不可约多项式, $1\leqslant k\leqslant r$. 这证明了因式分解的存在性.

因式分解的唯一性及和因子的表达式的证明和定理 1.3.9 相同. □

注意从唯一析因定理看出, 从乘法角度来看, 研究多项式的问题化为研究不可约多项式. 在整系数多项式的情形, 判别一个多项式是否不可约就很重要了. 但是没有好的判别法. 下面给出

定理 1.4.9 (Esenstein 不可约判别法) 给定次数大于零的本原多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

如果存在素数 p, 使得

$$p\nmid a_n,\quad p\mid a_j,\quad 0\leqslant j\leqslant n-1,\quad p^2\nmid a_0,$$

则 $f(x)$ 为整系数不可约多项式.

证 用反证法. 若本原多项式 $f(x)$ 为整系数可约多项式, 则存在整系数多项式 $g(x)$ 和 $h(x)$, 使得 $\deg(g(x))>0$, $\deg(h(x))>0$, 且 $f(x)=g(x)h(x)$. 记 $g(x)=\sum\limits_{i=0}^{r}b_ix^i$, $h(x)=\sum\limits_{j=0}^{s}c_jx^j$, 则 $r+s=n$, 又

$$f(x)=g(x)h(x)=\sum_{k=0}^{n}a_kx^k,\quad a_n=b_rc_s\neq 0,$$

其中

$$a_k=\sum_{i+j=k}b_ic_j,\quad 0\leqslant i\leqslant r,\quad 0\leqslant j\leqslant s,\quad 0\leqslant k\leqslant n.$$

由于素数 p 有 $p\nmid a_n$, $p\mid a_k$, $0\leqslant k\leqslant n-1$, $p^2\nmid a_0$, 所以 $p\nmid b_rc_s$ 给出 $p\nmid b_r$, $p\nmid c_s$; $p\mid b_0c_0$, $p^2\nmid b_0c_0$ 给出 $p\mid b_0$, $p\nmid c_0$ 或 $p\mid c_0$, $p\nmid b_0$; 又有 $p\mid\sum\limits_{i+j=k}b_ic_j$, $1\leqslant k\leqslant n-1$.

为方便起见, 无妨设 $p \mid b_0$. 这时 $p \nmid b_r$, $p \nmid c_0$, $p \nmid c_s$. 因此存在正整数 $u < r$, 使得 $p \mid b_0, \cdots, p \mid b_u, p \nmid b_{u+1}$.

当 $j > r$ 或 $j < 0$, 记 $b_j = 0$, 则有

$$a_{u+1} = c_0 b_{u+1} + c_1 b_u + c_2 b_{u-1} + \cdots + c_s b_{u+1-s}.$$

因为 $p \mid (c_1 b_u + c_2 b_{u-1} + \cdots + c_s b_{u+1-s})$, 所以 $p \mid c_0 b_{u+1}$. 这和假设矛盾. $\square$

注意 在使用 Esenstein 不可约判别法时, 可以在 $f(0) \neq 0$ 时考虑 $g(x) = x^{-n} f(x)$, 其中 $\deg(f(x)) = n$; 也可以考虑 $g(x) = f(x - m)$, 其中 m 是任意取定的整数.

习 题 1.4

1.4.1 试证: 下列整系数多项式不可约:

$x^4 - 8x^3 + 12x^2 - 6x + 2, \quad x^4 - x^3 + 2x + 1, \quad x^4 + 1, \quad x^6 + x^3 + 1, \quad x^5 - x^2 + 1.$

1.4.2 设 $f(x) = x^3 + bx^2 + cx + d$ 为整系数多项式, 设 $(b+c)d$ 为奇数. 试证: $f(x)$ 在整系数范围内不可约.

1.4.3 试证: 本原多项式 $x^{10} \pm x^5 + 1$ 可约.

1.4.4 试证: $\sum\limits_{j=0}^{p-1} x^j = 1 + x + x^2 + \cdots + x^{p-1}$, 为整系数不可约多项式, 其中 p 为素数.

1.4.5 给定 r 个不同素数 $p_1 < p_2 < \cdots < p_r$, 试证: $\sqrt[m]{p_1 p_2 \cdots p_r}$ 为无理数, 其中 $m \geqslant 2$ 为正整数.

1.4.6 设 $n > 1$, 试证: $f(x) = x^n + 5x^{n-1} + 3$ 在整系数范围内不可约.

1.4.7 设 $n > 2$, $a_1, a_2, \cdots, a_n$ 是 n 个不同整数. 求证: $f(x)$ 是整系数不可约多项式, 其中

$$f(x) = (x - a_1)(x - a_2) \cdots (x - a_n) \pm 1.$$

1.4.8 给定 n 个不同整数 $a_1, a_2, \cdots, a_n$. 试证: $f(x)$ 是 $2n$ 次整系数不可约多项式, 其中

$$f(x) = \prod_{i=1}^{n} (x - a_i)^2 + 1.$$

1.4.9 设 p 为素数, 它的十进制表示为

$$p = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \cdots + a_0.$$

试证: 整系数多项式 $f(x)$ 在整系数范围内不可约, 其中

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

1.4.10 试证: 整系数多项式

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

在整系数范围内不可约, 其中 $|a_0| > \sum\limits_{k=1}^{n} |a_k|$, 又 a_0 为素数; 或 $\sqrt{|a_0|} > 1 + \sqrt{|a_n|}$.

§1.5 一元多项式的根

定义 1.5.1 给定域 $\mathbb{F}$ 上多项式 $f(x)$. 复数 α 称为多项式 $f(x)$ 的**根**, 也称为多项式 $f(x)$ 的**零点**, 如果 $f(\alpha)=0$. 当 α 为复数时, 称为多项式 $f(x)$ 的**复根**; 当 α 为实数时, 称为多项式 $f(x)$ 的**实根**; 当 α 为有理数时, 称为多项式 $f(x)$ 的**有理数根**; 当 α 为整数时, 称为多项式 $f(x)$ 的**整数根**.

显然, 域 $\mathbb{F}$ 中所有数都是零多项式的根, 又零次多项式无根. 下面讨论域 $\mathbb{F}$ 上次数大于零的多项式在域 $\mathbb{F}$ 中的根, 考虑根和因式分解之间的关系, 介绍一些求根的方法 —— 试验法.

定理 1.5.2 给定域 $\mathbb{F}$ 上多项式 $f(x)$, 域 $\mathbb{F}$ 中数 α 为多项式 $f(x)$ 的根当且仅当

$$(x-\alpha)\mid f(x). \tag{1.5.1}$$

证 由定理 1.2.1, 用 $x-\alpha$ 除 $f(x)$, 则有

$$f(x)=(x-\alpha)q(x)+c,$$

其中 c 为常数, $q(x)$ 为商式. 取 $x=\alpha$, 则有 $f(\alpha)=c$. 所以证明了

$$f(x)=(x-\alpha)q(x)+f(\alpha). \tag{1.5.2}$$

因此域 $\mathbb{F}$ 中数 α 为多项式 $f(x)$ 的根, 即 $f(\alpha)=0$ 当且仅当 $(x-\alpha)\mid f(x)$. □

定理 1.5.3 域 $\mathbb{F}$ 上 $n(\geqslant 1)$ 次多项式 $f(x)$ 在域 $\mathbb{F}$ 中至多有 n 个不同的根.

证 设 $f(x)$ 为域 $\mathbb{F}$ 上 n 次多项式. 设 $\alpha_1\in\mathbb{F}$ 为 $f(x)$ 的根. 于是 $f(x)=(x-\alpha_1)f_1(x)$, 其中 $f_1(x)$ 仍为域 $\mathbb{F}$ 上 $n-1$ 次多项式. 显然 $f_1(x)$ 在域 $\mathbb{F}$ 中的根仍为 $f(x)$ 在域 $\mathbb{F}$ 中的根. 用归纳法, 已知一次多项式恰好有一个根. 设 $n-1$ 次多项式 $f_1(x)$ 在域 $\mathbb{F}$ 中至多有 $n-1$ 个不同根. 由 $f(x)=(x-\alpha_1)f_1(x)$ 可知 n 次多项式在域 $\mathbb{F}$ 中至多有 n 个不同根. □

定理 1.5.4 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \tag{1.5.3}$$

为域 $\mathbb{F}$ 上多项式, 这里不要求 $a_n\neq 0$. 设存在域 $\mathbb{F}$ 中 $n+1$ 个不同的数 $\beta_1,\beta_2,\cdots,\beta_{n+1}$, 使得 $f(\beta_j)=0,\ 1\leqslant j\leqslant n+1$, 则 $f(x)$ 为零多项式, 即 $a_0=a_1=\cdots=a_n=0$.

证 用反证法. 假设 $f(x)$ 不是零多项式, 则 $\deg(f(x))\leqslant n$. 于是由假设, $f(x)$ 有 $n+1$ 个不同根 $\beta_j,\ 1\leqslant j\leqslant n+1$. 由定理 1.5.3 便导出矛盾. □

定理 1.5.5 设域 $\mathbb{F}$ 上两多项式 $f(x)$ 和 $g(x)$ 的次数都不超过 n, 则 $f(x)=g(x)$

当且仅当存在域 $\mathbb{F}$ 中 $n+1$ 个不同的数 $\beta_1,\ \beta_2,\ \cdots,\ \beta_{n+1}$，使得 $f(\beta_j) = g(\beta_j)$，$1 \leqslant j \leqslant n+1$.

证 考虑多项式 $G(x) = f(x) - g(x)$，由定理 1.5.4 便证明了定理. □

定理 1.5.6 (Lagrange 插值公式) 给定域 $\mathbb{F}$ 中 $n+1$ 个不同的数 $\alpha_1, \alpha_2, \cdots, \alpha_{n+1}$. 给定域 $\mathbb{F}$ 中另外 $n+1$ 个数 $\beta_1, \beta_2, \cdots, \beta_{n+1}$，则唯一存在域 $\mathbb{F}$ 中一个次数不超过 n 的多项式 $f(x)$，它有 $f(\alpha_j) = \beta_j$，$1 \leqslant j \leqslant n+1$，其中

$$f(x) = \sum_{i=1}^{n+1} \beta_i \frac{(x-\alpha_1)\cdots(x-\alpha_{i-1})(x-\alpha_{i+1})\cdots(x-\alpha_{n+1})}{(\alpha_i-\alpha_1)\cdots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\cdots(\alpha_i-\alpha_{n+1})}. \tag{1.5.4}$$

证 由式 (1.5.4) 给出的 $f(x)$，作为域 $\mathbb{F}$ 上的多项式，显然有 $f(\alpha_j) = \beta_j$，$1 \leqslant j \leqslant n+1$. 这证明了存在性. 由定理 1.5.5，便证明了唯一性. □

下面讨论多项式的重根.

定义 1.5.7 给定域 $\mathbb{F}$ 上多项式 $f(x)$. 设复数 α 为多项式 $f(x)$ 的根. 设正整数 e 有 $(x-\alpha)^e \mid f(x)$，但是 $(x-\alpha)^{e+1} \nmid f(x)$，则复数 α 称为多项式 $f(x)$ 的 e **重根**. 当 $e=1$ 时，复数 α 也称为多项式 $f(x)$ 的**单重根**.

定理 1.5.8 设域 $\mathbb{F}$ 上 n 次多项式 $f(x)$ 有 s 个不同根 $\alpha_1, \alpha_2, \cdots, \alpha_s$，它们的重数分别为 $e_1, e_2, \cdots, e_s$，而且 $e_1 + e_2 + \cdots + e_s = n$，则

$$f(x) = a(x-\alpha_1)^{e_1}(x-\alpha_2)^{e_2}\cdots(x-\alpha_s)^{e_s}, \tag{1.5.5}$$

其中 a 为 $f(x)$ 的首项系数. 于是多项式

$$g(x) = \frac{f(x)}{(f(x), f'(x))} = a(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_s) \in \mathbb{F}[x] \tag{1.5.6}$$

无重根，但是和多项式 $f(x)$ 有相同的不同根. 因此，$f(x)$ 无重根当且仅当

$$(f(x), f'(x)) = 1. \tag{1.5.7}$$

证 由定理 1.3.14 可知定理成立. □

下面分别在不同的数的范围内讨论根的存在性及求根方法. 这里要注意 $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

(I) 复多项式情形

定理 1.5.9 (代数基本定理) $n \geqslant 1$ 次复多项式必有复根.

证 * 设 $f(x)$ 为 $n \geqslant 1$ 次复多项式. 设 $f(0) = 0$，则 $f(x)$ 有复根 $x=0$. 设 $f(0) = a \neq 0$，则 $f(x)$ 可表为

$$f(x) = a + bx^m + x^{m+1}g(x),$$

其中 a, b 为非零复数. 当 $m = n$ 时, $g(x) = 0$; 当 $m < n$ 时,

$$g(x) = g_{n-m-1}x^{n-m-1} + \cdots + g_1 x + g_0, \quad g_{n-m-1} \neq 0$$

为 $n - m - 1$ 次复多项式.

今 $\lim\limits_{|x|\to+\infty} |f(x)| = +\infty$. 所以 $|f(x)|$ 在复数域 $\mathbb{C}$ 中一点 x_0 达到最小模, 即 $|f(x_0)| \leqslant |f(x)|, \forall x \in \mathbb{C}$ (这里用到连续函数的性质). 下面来证 $f(x_0) = 0$, 即 $n \geqslant 1$ 次复多项式必有复根.

用反证法. 若 $f(x_0) \neq 0$. 作平移 $x \to x - x_0$, 所以无妨设 $x_0 = 0$. 这时, $0 < a = |f(0)| \leqslant |f(x)|, \forall x \in \mathbb{C}$. 取 $0 < t < 1$, $x_1 = \sqrt[m]{-ab^{-1}}$, 则 $x_1^m = -ab^{-1}$. 今 $f(tx) = a + bt^m x^m + t^{m+1}x^{m+1}g(tx)$. 因此, 由 $0 < t < 1$, 有

$$\begin{aligned}|f(tx_1)| &= |a + bt^m x_1^m + t^m x_1^m (tx_1 g(tx_1))| = \left|a - at^m - at^m \frac{tx_1 g(tx_1)}{b}\right| \\ &\leqslant |a|\left(1 - t^m + t^m \left|\frac{tx_1 g(tx_1)}{b}\right|\right),\end{aligned}$$

其中

$$\begin{aligned}\left|\frac{tx_1 g(tx_1)}{b}\right| &= \frac{t|x_1|}{|b|}|g_{n-m-1}(tx_1)^{n-m-1} + \cdots + g_1(tx_1) + g_0| \\ &\leqslant \frac{t|x_1|}{|b|}(|g_{n-m-1}|(t|x_1|)^{n-m-1} + \cdots + |g_1|(t|x_1|) + |g_0|).\end{aligned}$$

令 $M = \max(|g_{n-m-1}||x_1|^{n-m-1}, \cdots, |g_1||x_1|, |g_0|)$. 由于 $0 < t < 1$, 因此

$$\left|\frac{tx_1 g(tx_1)}{b}\right| \leqslant (t^{n-m} + \cdots + t^2 + t)\frac{M|x_1|}{|b|} \leqslant \frac{(n-m)M|x_1|}{|b|}t.$$

所以存在 $t \in [0, 1]$, 使得 $0 < t << \min\left(1, \dfrac{|b|}{(n-m)M|x_1|}\right)$. 因此,

$$|f(tx_1)| < |a|(1 - t^m + t^m) = |a|.$$

这和点 x_0 的选取矛盾. 所以证明了如果 $|f(x)|$ 在复数域 $\mathbb{C}$ 中一点 x_0 达到最小模, 则 $f(x_0) = 0$, 即代数基本定理成立. □

这个定理的证明, 不能避开连续函数的中间值定理. 同学们也可以在数学系高年级复变函数课中学到代数基本定理的证明.

定理 1.5.10 设 $n \geqslant 1$, 则 n 次复多项式必可分解为 n 个一次因式的乘积. 于

是 n 次复多项式恰有 n 个复根 (不计重数).

证 由定理 1.5.2 及归纳法, 立即可证. □

定理 1.5.11 不可约复多项式是且只是零次复多项式及一次复多项式.

证 由不可约多项式的定义及定理 1.5.10 立即可证. □

(II) 实多项式情形

记

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \tag{1.5.8}$$

为复多项式, 于是 $f(x)$ 为实多项式当且仅当 $\overline{a_j} = a_j$, $0 \leqslant j \leqslant n$, 即

$$\overline{f(x)} = f(x), \tag{1.5.9}$$

其中 x 不参与取共轭.

定理 1.5.12 实多项式若有复根 α, 则必有**共轭复根** $\overline{\alpha}$. 若 α 不是实数, 则 $f(x)$ 有实因式

$$(x-\alpha)(x-\overline{\alpha}) = x^2 - (\alpha+\overline{\alpha})x + \mid \alpha \mid^2 . \tag{1.5.10}$$

证 由定理 1.5.2 和式 (1.5.9) 可证. □

显然, 当 α 不是实数, 即 $\overline{\alpha} \neq \alpha$ 时, $x^2 - (\alpha+\overline{\alpha})x + \mid \alpha \mid^2$ 为实不可约多项式, 这时它的判别式 $\Delta = (\alpha+\overline{\alpha})^2 - 4 \mid \alpha \mid^2 = (\alpha-\overline{\alpha})^2 < 0$. 结合定理 1.5.2 及定理 1.5.12 便证明了

定理 1.5.13 实不可约多项式是且只是零次、一次及二次实多项式. 在二次的情形, 还要求判别式小于零.

由于实多项式的复根和它的共轭复根成对出现, 所以也证明了

定理 1.5.14 当 $n \geqslant 1$ 时, n 次实多项式恰有 $2r$ 个复根 $\alpha_1, \cdots, \alpha_r, \overline{\alpha}_1, \cdots, \overline{\alpha}_r$, 其中 $\alpha_1, \cdots, \alpha_r$ 的虚部都不等于零, 且有 $n-2r$ 个实根 $c_{2r+1}, \cdots, c_n$. 所以 $f(x)$ 能分解为下面不可约因式的乘积:

$$f(x) = a[(x-\alpha_1)(x-\overline{\alpha}_1)]\cdots[(x-\alpha_r)(x-\overline{\alpha}_r)](x-c_{2r+1})\cdots(x-c_n). \tag{1.5.11}$$

关于求实多项式的实根的最好方法是 Sturm 定理, 这是下一节的内容.

(III) 有理系数多项式的情形

由引理 1.4.4, 有理系数多项式的有理数根为相应本原多项式的有理数根, 反之

亦然. 所以问题化为如何来求本原多项式的有理数根. 这见下面的内容.

(IV) 整系数多项式的情形

给定本原多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad (a_0, \cdots, a_n) = 1. \tag{1.5.12}$$

下面介绍求 $f(x)$ 的整数根及有理数根的试验法.

设 $a_0 = 0$, 则 $f(0) = 0$; 设 $a_0 \neq 0$, 设 m 是 $f(x)$ 的整数根, 即有

$$f(m) = a_n m^n + a_{n-1}m^{n-1} + \cdots + a_1 m + a_0 = 0. \tag{1.5.13}$$

于是 $m \mid a_0$. 因此可先求出 a_0 的所有因子, 将每个因子代入 $f(x)$ 中计算它何时为零, 从而可求出 $f(x)$ 的所有整数根.

设 m/q 为既约分数, 即有 $m, q \in \mathbb{Z}$, $q > 0$, $(m, q) = 1$. 设 m/q 为本原多项式 $f(x)$ 的有理数根, 即有

$$f\left(\frac{m}{q}\right) = a_n\left(\frac{m}{q}\right)^n + a_{n-1}\left(\frac{m}{q}\right)^{n-1} + \cdots + a_1\left(\frac{m}{q}\right) + a_0 = 0.$$

公式两边乘以 q^n, 则有

$$a_n m^n + a_{n-1}m^{n-1}q + \cdots + a_1 m q^{n-1} + a_0 q^n = 0. \tag{1.5.14}$$

由 $(m, q) = 1$, 于是 $q \mid a_n$, $m \mid a_0$. 因此可先求出 a_0 和 a_n 的所有因子, 分别记作 a_0' 和 a_n', 得到整数对 $\{a_0', a_n'\}$. 我们先筛去 $(a_0', a_n') \neq 1$ 的整数对, 余下的整数对可定义一个既约分数 $x_0 = a_0'/a_n'$, 再将每个有理数 x_0 代入 $f(x)$ 中, 计算它何时为零, 从而可求出 $f(x)$ 的所有有理数根.

在 16 世纪, 给出了三次及四次多项式的公式解法后, 在一, 二, 三, 四次多项式的情形, 已经知道根都可以从系数 $a_n, a_{n-1}, \cdots, a_1, a_0$ 按照确定的规律, 经过加, 减, 乘, 除和开任意次方这五种代数运算得到. 此后有二百年的时间, 很多数学家都致力于寻求 $n \geqslant 5$ 次多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad 5 \leqslant n$$

的根的表达形式, 这里系数是作为文字, 而不是一批给定的数字. 所以对 n 次多项式, 人们自然也希望寻找到一个这样的公式. 直到 1820 年, Abel (1802—1829) 才证明了在 $n \geqslant 5$ 时, 这是办不到的.

尽管对于文字系数的多项式, 不可能得到一个这样的公式, 然而, 对于具体数值系数的多项式, 是不是可能找到一个这样的公式呢? 换句话说, 它们的根是不是可以由系数经过加, 减, 乘, 除和开任意次方这五种代数运算求得呢? 这个问题在 1830 年

由 Galois (1811—1832) 彻底地解决了. 他给出了一种方法来判断: 哪些数值系数的多项式不可以, 哪些可以. 例如, 多项式 x^5-4x-2 的根就不能由系数经过加, 减, 乘, 除和开任意次方等运算得到. 由于这个问题的解决, 也就回答了为什么在初等几何中不能用圆规和直尺将一角三等分等问题.

到现在为止, 我们介绍了古典方程式论, 即关于域 $\mathbb{F}$ 上一元多项式环 $\mathbb{F}[x]$ 的唯一析因理论和根的理论. 这两个理论关系很密切. 在唯一析因理论中, 基本概念是除不除得尽, 可不可约, 有没有因子分解, 分解为不可约因子乘积时, 在什么意义下唯一. 而唯一析因理论告诉我们, 分解在不计次序时唯一. 所以问题便化为决定所有不可约多项式; 在后一理论中, 首先要确定根与因子分解的关系, 然后给出方法来考查根的分布, 最好是进一步算出根, 同时定出根的重数.

上面提到的一系列问题, 是 19 世纪前期及更早些时候的代数学者们研究的重点, 这方面的成果是极其丰富的, 这里仅仅介绍了一小部分.

习 题 1.5

1.5.1 试证: 下列域 $\mathbb{F}$ 上多项式都以 1 为三重根:

(1) $x^{2n}-nx^{n+1}+nx^{n+1}-1$;

(2) $(n-2m)x^n-nx^{n-m}+nx^m-(n-2m)$, $n>m$;

(3) $x^{2n+1}-(2n+1)x^{n+1}+(2n+1)x^n-1$.

1.5.2 试求正整数 n, 使得域 $\mathbb{F}$ 上多项式 $(x+1)^n-x^n-1$ 无重根.

1.5.3 试证: 域 $\mathbb{F}$ 上多项式

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

无重根.

1.5.4 设 x_0 为域 $\mathbb{F}$ 上多项式 $g(x)=f(x)g'(x)-f'(x)g(x)$ 的 k 重根, 记 $h(x)=f(x_0)g(x)-f(x)g(x_0)$ 为非零多项式. 试证: x_0 为域 $\mathbb{F}$ 上多项式 $h(x)$ 的 $k+1$ 重根. 反之亦然.

1.5.5 设 $f(x)$ 为实多项式, 由数学分析可知当 $f'(x_0)=0$, $f''(x)<0$ 时 $f(x)$ 在 $x=x_0$ 达极大值; 当 $f'(x_0)=0$, $f''(x)>0$ 时 $f(x)$ 在 $x=x_0$ 达极小值. 用这个结果证明: 设 $f(x)$ 为 n 次实多项式, 且对一切实数 x, 有 $f(x)>0$, 则对一切实数 x 有

$$f(x)+f'(x)+\cdots+f^{(n)}(x)\geqslant 0,$$

其中 $f^{(n)}(x)=\dfrac{\mathrm{d}\,f^{(n-1)}(x)}{\mathrm{d}\,x}$, $n=2,3,\cdots$, $f^{(1)}(x)=f'(x)$.

1.5.6 试证: 域 $\mathbb{F}$ 上多项式 $f'(x)\mid f(x)$ 当且仅当 $f(x)=a_n(x-a)^n$, 其中 a_n 为 $f(x)$ 的首项系数, a_n, $a\in\mathbb{F}$.

1.5.7 设 $f(x)$ 和 $p(x)$ 为域 $\mathbb{F}$ 上两个非零多项式, 且有公根 α. 试证: 当 $p(x)$ 为不可约多项式时有 $p(x)\mid f(x)$.

1.5.8 设实多项式 $f(x)=x^4+ax^3+bx+c$ 的根都是实数, 试证: $ab\leqslant 0$.

1.5.9 给定实数 a, b, c, 其中 $c \neq 0$. 试证: 实多项式 $f(x) = x^5 + ax^4 + bx^3 + c$ 至少有两个复根.

1.5.10 设 a, b 为实数, $2a^2 < 5b$. 试证: 实多项式 $f(x) = x^5 + ax^4 + bx^3 + cx^2 + dx + e$ 必有虚部不等于零的复根.

1.5.11 设 $f(x)$ 为整系数多项式, 且 $f(0)$ 和 $f(1)$ 都是奇数. 试证: $f(x)$ 无有理数根.

1.5.12 设

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{Z}[x],$$

其中 $a_{n-1}, \cdots, a_0$ 都是奇数, n 是偶数. 试证 $f(x)$ 无整数根.

1.5.13 设 $x^2 + px + q$ 和 $x^2 + rx + s$ 都是整系数多项式, 且它们有一个公根 α 不是整数. 试证 $p = r, q = s$.

1.5.14 试证: 当 $n > 1$ 时, 整系数多项式

$$f(x) = nx^{n-1} + (n-1)x^{n-2} + \cdots + 3x^2 + 2x + 1 - n^2$$

在整系数范围内可约, 即至少有一个有理数根.

1.5.15 设 $n \geqslant 1$, 求形如

$$x^n \pm x^{n-1} \pm \cdots \pm x \pm 1$$

的实多项式, 使得它无复根, 只有实根.

1.5.16 设实多项式 $f(x), g(x), h(x), k(x)$ 有

$$(x^2+1)h(x) + (x-1)f(x) + (x-2)g(x) = 0,$$
$$(x^2+1)k(x) + (x+1)f(x) + (x+2)g(x) = 0.$$

求证: $(x^2+1)|f(x)$, $(x^2+1)|g(x)$.

1.5.17 设三次方程 $x^3 - x^2 - x - 1 = 0$ 的根是复数 a, b, c. 试证:

(1) a, b, c 为三个不同复数;

(2) 对任意正整数 n,

$$\frac{a^n - b^n}{a - b} + \frac{b^n - c^n}{b - c} + \frac{c^n - a^n}{c - a} \in \mathbb{Z}.$$

1.5.18 设实多项式 $f(x), g(x), h(x), k(x)$ 有

$$f(x^5) + xg(x^5) + x^2h(x^5) = (x^4 + x^3 + x^2 + x + 1)k(x),$$

试证: 多项式 $f(x), g(x), h(x)$ 以 $x - 1$ 为公因式.

1.5.19 **(1)** 求方程 $\sqrt{x^2 - a} + 2\sqrt{x^2 - 1} = x$ 的所有实根, 其中 a 为实常数;

(2) 求解方程 $\sqrt{x + \sqrt{2x-1}} + \sqrt{x - \sqrt{2x-1}} = a > 0$.

1.5.20 (Newton 公式) 记 $f(x)$ 为复多项式,

$$f(x) = (x-a_1)(x-a_2)\cdots(x-a_n) = x^n - \sigma_1 x^{n-1} + \cdots + (-1)^j\sigma_j x^{n-j} + \cdots + (-1)^n\sigma_n,$$

$$s_k = a_1^k + a_2^k + \cdots + a_n^k, \quad k = 0, 1, 2, \cdots.$$

试证:

$$g_k(x)=x^{k+1}f'(x)-(s_0x^k+s_1x^{k-1}+\cdots+s_k)f(x),\quad k=0,1,2,\cdots$$

其中 $\deg(g_k(x))<\deg(f(x))$. 试用上式推出 Newton 公式

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\cdots+(-1)^{k-1}\sigma_{k-1}s_1+(-1)^k\sigma_k k=0,\quad 1\leqslant k<n;$$
$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\cdots+(-1)^n\sigma_ns_{k-n}=0,\quad k\geqslant n.$$

1.5.21　给定正整数 n, 试求适合条件

$$f(f(x))=f(x)^n+a_{n-1}f(x)^{n-1}+\cdots+a_1f(x)+a_0$$

的所有 n 次复多项式. (提示: 用代数基本定理.)

1.5.22　设实多项式 $f(x)$ 的所有根都是纯虚数. 试证: 它的微分 $f'(x)$ 的所有根, 除了一个例外, 也都是纯虚数.

1.5.23　设 $p(x)$ 是 $n-1$ 次多项式, 系数全为正实数. 记 a_n 为正实数. 试证: n 次多项式 $f(x)=a_nx^n-p(x)$ 有唯一的单重正实根.

1.5.24　设实多项式序列 $\{p_n(x)\}_{n=1}^{\infty}$ 有初值 $p_1(x)=x^2-2$, 又有

$$p_k(x)=p_1(p_{k-1}(x))=p_{k-1}(x)^2-2,\quad k=2,3,\cdots.$$

试证: 对一切正整数 n, 多项式 $p_n(x)-x$ 的根都是实数, 且都是单重根.

1.5.25　设 $f(x)=\sum\limits_{j=0}^{n}a_jx^j$ 是非零实多项式, 它的根也都是实根. 试证: 多项式 $g(x)=\sum\limits_{j=0}^{n}\binom{n}{j}a_jx^j$ 的根也都是实根.

1.5.26　设 $f(x)$ 为实多项式. 则 $f(x)$ 关于一切实数 x_0, 总有 $f(x_0)\geqslant 0$ 当且仅当存在实多项式 $g(x)$, $h(x)$, 使得 $f(x)=g(x)^2+h(x)^2$.

1.5.27　设 $f(x)$ 为实多项式. 试证: 如果 $f(x)$ 的根都是实根, 则 $f(x)^2$ 不能表成两个非零实多项式 $g(x)$, $h(x)$ 的平方和 $f(x)^2=g(x)^2+h(x)^2$, 其中 $f(x)$ 除不尽 $g(x)$. 反之如何?

1.5.28　任给实多项式

$$f(x)=x^{2n}+a_1x^{2n-1}+\cdots+a_{2n-1}x+1,$$

其中系数 a_1, a_2, $\cdots$, a_{2n-1} 用下法决定: 甲、乙两人依次轮流给定一个系数, 最后所得的多项式若无实根, 则甲胜, 若有实根, 则乙胜. 设由甲先取, 问谁胜. (答案为乙有必胜策略.)

1.5.29　设 $\alpha\in\mathbb{C}$, $|\alpha|=1$ 为适合条件

$$0<a_n\leqslant a_{n-1}\leqslant\cdots\leqslant a_1\leqslant a_0$$

的实多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+a_1x+a_0$$

的根. 试证: $\alpha^{n+1}=1$.

1.5.30　试求.

$$\sin\frac{\pi}{2n}\cdot\sin\frac{2\pi}{2n}\cdots\sin\frac{(n-1)\pi}{2n}.$$

1.5.31 设

$$f(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

为整系数多项式, 若有三个不同整数 a, b, c, 使得

$$|f(a)| = |f(b)| = |f(c)| = 1.$$

试证: $f(x)$ 没有整数根.

1.5.32 设 $f(x)$ 为整系数多项式, 且有 $q \geqslant 5$ 个互不相同的整数 a_1, a_2, $\cdots$, a_q, 使得 $f(a_i) = p$, $1 \leqslant i \leqslant q$, 其中 p 为素数. 试证: $f(x)$ 无整数根.

1.5.33 设 $f(x)$ 为有理系数不可约多项式, 且 $\deg(f(x)) = 2n+1$, $n \geqslant 1$. 试证: $f(x)$ 的任两不同的根的和, 差, 积, 商都不是有理数.

1.5.34 设存在奇数 $2m+1$ 和偶数 $2n$ 有 $f(2m+1)$, $f(2n)$ 都是奇数, 其中 $f(x)$ 为整数多项式. 试证:$f(x)$ 无整数根.

1.5.35 设 α 是非零复数. α 是非零有理系数多项式的根当且仅当存在有理系数多项式 $f(x)$, 使得 $f(\alpha) = \alpha^{-1}$.

1.5.36 设 $p(x)$ 为整系数多项式, 且 $\deg(p(x)) > 1$. 记适合条件 $p(m)^2 = 1$ 的不同整数 m 的总个数为 $n(p(x))$, 试证

$$n(p(x)) - \deg(p(x)) \leqslant 2.$$

1.5.37 一些一次实多项式构成的集合 S, 若适合条件

(1) 若 $f(x)$, $g(x) \in S$, 则 $f(x)(g(x)) \in S$;

(2) 若 $f(x) \in S$, 则 $f^{-1}(x) \in S$, 其中 $f(x) = ax + b$, $f^{-1}(x) = a^{-1}(x-b)$;

(3) $\forall f(x) \in S$, 则存在 $x_f \in \mathbb{R}$, 使得 $f(x_f) = x_f$.

试证: 存在实数 ξ, 使 $x_f = \xi$, $\forall f(x) \in S$.

1.5.38 设 $f(x)$ 为次数小于 n 的复多项式, N 为正整数, $N = m_1m_2\cdots m_n$, 其中 m_1, m_2, $\cdots$, m_n 为大于 1 的两两互素的正整数. 记

$$\omega = \cos\frac{2\pi}{N} + \sqrt{-1}\sin\frac{2\pi}{N}.$$

试证: 存在 0, 1, $\cdots$, $N-1$ 的排列 $i_0i_1\cdots i_{N-1}$, 使得 $\sum_{k=0}^{N-1} f(i_k)\omega^k = 0$.

1.5.39 试用题 **1.5.39** 来证明是否存在一个凸 n 边形, 它等角, 且 n 条边的长度恰好为 1^2, 2^2, $\cdots$, n^2 的排列.

§1.6 一元实多项式的 Sturm 定理*

在实际问题中, 会出现许多求实多项式的实根问题. 当然具体去求一个已知的实多项式的实根是较困难的事. 已经发现许多方法求近似根, 这些方法属于计算数学的一个分支.. 下面给出一种经典的办法来求实根的近似值, 这就是著名的 Sturm 定理. Sturm 利用了实多项式的一个特性: 设实多项式 $f(x)$ 有 $f(a)f(b) < 0$, 其中

$a<b$ 为实数, 则在区间 (a,b) 中存在实根 α. 他将这一事实和辗转相除公式结合起来, 给出了下面结果.

定义 1.6.1 给定次数大于零的实多项式 $f(x)$. 对多项式 $f_0(x)=f(x)$ 及 $f_1(x)=f'(x)$ 作辗转相除, 有

$$\begin{aligned}
f_0(x)&=q_1(x)f_1(x)-f_2(x),\\
f_1(x)&=q_2(x)f_2(x)-f_3(x),\\
f_2(x)&=q_3(x)f_3(x)-f_4(x),\\
&\vdots\\
f_{s-2}(x)&=q_{s-1}(x)f_{s-1}(x)-f_s(x),\\
f_{s-1}(x)&=q_s(x)f_s(x),
\end{aligned} \tag{1.6.1}$$

其中

$$\deg(f_1(x))>\deg(f_2(x))>\cdots>\deg(f_s(x))\geqslant 0. \tag{1.6.2}$$

则多项式序列

$$f_0(x),f_1(x),\cdots,f_s(x)$$

称为 **Sturm 序列**.

定义 1.6.2 给定 $s+1$ 个实数 $c_0,c_1,\cdots,c_s$. 从左向右观察, 如果 $c_i=0$, 则除去 c_i; 如果 $c_i\neq 0$, $c_{i+1}=0,\cdots,c_{j-1}=0$, $c_j\neq 0$. 这时考察 c_i 和 c_j 的正负号. 如果符号相同 (即同时为正数或同时为负数), 则称为**无变号**; 如果符号相反 (即其中有一个为正数, 另一个为负数), 则称为有一个变号. 从 c_0 观察到 c_s, 变号数的总和称为该序列的**变号总数**.

定义 1.6.3 给定次数大于零的实多项式 $f(x)$. 记它的 Sturm 序列为 $f_0(x)$, $f_1(x)$, $\cdots$, $f_s(x)$. 对任一实数 c, 记序列 $f_0(c)$, $f_1(c)$, $\cdots$, $f_s(c)$ 的总变号数为 $w(c)$ (它是非负整数). 于是得到实轴上的函数 $w(x)$, 它有 w: $\mathbb{R}\to\mathbb{Z}_+$, 称为 Sturm 序列 $f_0(x)$, $f_1(x)$, $\cdots$, $f_s(x)$ 的**变号函数**.

引理 1.6.4 设次数大于零的多项式 $f(x)$ 为无重根的实多项式, 即实多项式 $f(x)$ 有 $(f(x),f'(x))=1$. 则 $f(x)$ 的 Sturm 序列有性质:

(1) $f_s(x)=f_s(0)\neq 0$;

(2) 相邻的两多项式 $f_j(x)$ 和 $f_{j+1}(x)$ 无公共实根;

(3) 如果实数 c 有 $f_j(c)=0$, 则

$$f_{j-1}(c)=-f_{j+1}(c),\quad 1\leqslant j\leqslant s-1;$$

(4) 如果实数 c 有 $f_0(c)=0$, 则存在 $\varepsilon>0$, 使得

$$f_0(x)f_1(x)<0,\quad c-\varepsilon<x<c;\quad f_0(x)f_1(x)>0,\quad c<x<c+\varepsilon.$$

证 **(1)** 由 $(f(x), f'(x)) = 1$ 可知 $f_s(x)$ 为非零常数, 所以 **(1)** 成立. **(2)** 用反证法. 若存在 $j \in \{0, 1, \cdots, s-1\}$ 以及实数 c, 使得 $f_j(c) = f_{j+1}(c) = 0$. 由 Sturm 序列的定义 1.6.1 可知 $f_{j+2}(c) = 0, \cdots, f_s(c) = 0$. 这和 **(1)** 矛盾, 所以 **(2)** 成立. **(3)** 设 $j \in \{1, 2, \cdots, s-1\}$, 于是 $f_{j-1}(x) = q_j(x)f_j(x) - f_{j+1}(x)$. 取 $x = c$, 由 $f_j(c) = 0$, 则有 $f_{j-1}(c) = -f_{j+1}(c)$. 这证明了**(3)** 成立. **(4)** 最后, 如果实数 c 为 $f_0(x)$ 的根, 即有 $f_0(c) = 0$. 由 $(f(x), f'(x)) = 1$ 可知 $f_1(c) \neq 0$. 但是 $f_1(x)$ 为连续函数, 所以存在 $\varepsilon > 0$, 使得当 $c - \varepsilon < x < c + \varepsilon$ 时, 有 $f_1(x) = f'(x) \neq 0$, 即 $f(x)$ 在 $c - \varepsilon < x < c + \varepsilon$ 时为严格单调递增函数或为严格单调递减函数. 这分为两种情形: (4′) 设 $f'(x) > 0$, $|x - c| < \varepsilon$, 即 $f(x)$ 为严格单调递增函数. 由 $f(c) = 0$, 有 $f(x) > 0$, $c < x < c + \varepsilon$; 又 $f(x) < 0$, $c - \varepsilon < x$. 所以由 $f_0(x) = f(x)$, 有 $f_0(x)f_1(x) > 0$, $c < x < c + \varepsilon$; 又 $f_0(x)f_1(x) < 0$, $c - \varepsilon < x < c$. (4″) 设 $f'(x) < 0$, $|x - c| < \varepsilon$, 即 $f(x)$ 为严格单调递减函数. 由 $f(c) = 0$, 有 $f(x) < 0$, $c < x < c + \varepsilon$; 又 $f(x) > 0$, $c - \varepsilon < x < c$. 所以由 $f_0(x) = f(x)$, 有 $f_0(x)f_1(x) > 0$, $c < x < c + \varepsilon$; 又 $f_0(x)f_1(x) < 0$, $c - \varepsilon < x < c$.

上面两种情形推出了同样的结论: $f_0(x)f_1(x) > 0$, $c < x < c + \varepsilon$, 又 $f_0(x)f_1(x) < 0$, $c - \varepsilon < x < c$. 这证明了**(4)** 成立. □

定理 1.6.5 (Sturm 定理) 设 $f(x)$ 为无重根的非零实多项式, 即

$$(f(x), f'(x)) = 1.$$

设实数 a 和 b 都不是 $f(x)$ 的实根, 且 $a < b$. 则实多项式 $f(x)$ 在区间 (a, b) 中恰有 $w(b) - w(a)$ 个实根.

证 由于 Sturm 序列 $f_0(x), f_1(x), \cdots, f_s(x)$ 由多项式构成, 所以当 x 从 a 增大到 b 时, 如果不经过所有多项式 $f_0(x), f_1(x), \cdots, f_s(x)$ 的根时, 变号函数 $w(x)$ 的值不改变. 所以为了证明定理, 只要证明当 x 从 a 增大到 b 时, 如果不经过 $f_0(x)$ 的根时, 变号函数的值也不改变; 而经过 $f_0(x)$ 的一个根时, 变号函数的值减少 1, 则 Sturm 定理成立. 下面我们证明这两个事实.

(1) 注意到 $f(x)$ 和 $f'(x)$ 没有公根. 设实数 c 不是 $f_0(x)$ 的根, 而是某些 $f_j(x)$, $j > 0$ 的根. 由引理 1.6.4 的性质**(1)**, $j \in \{1, 2, \cdots, s-1\}$. 由引理 1.6.4 的性质**(3)**, $f_{j-1}(c) = -f_{j+1}(c)$. 由引理 1.6.4 的性质**(2)**, $f_{j-1}(c)f_{j+1}(c) \neq 0$. 所以当 $f_{j-1}(c) > 0$ 时有下表

	$f_{j-1}(x)$	$f_j(x)$	$f_{j+1}(x)$
$c - \varepsilon < x < c$	$+$	$\pm$	$-$
$x = c$	$+$	0	$-$
$c < x < c + \varepsilon$	$+$	$\mp$	$-$

因此 Sturm 序列的如上部分 $f_{j-1}(x)$, $f_j(x)$, $f_{j+1}(x)$ 当从 $c - \varepsilon$ 到 $c + \varepsilon$ 变化时, 变

号数始终为 1, 即变号数没有改变. 当 $f_{j-1}(c)<0$ 时有下表

	$f_{j-1}(x)$	$f_j(x)$	$f_{j+1}(x)$
$c-\varepsilon<x<c$	$-$	$\pm$	$+$
$x=c$	$-$	0	$+$
$c<x<c+\varepsilon$	$-$	$\mp$	$+$

因此 Sturm 序列的如上部分 $f_{j-1}(x)$, $f_j(x)$, $f_{j+1}(x)$ 当从 $c-\varepsilon$ 到 $c+\varepsilon$ 变化时, 变号数也始终为 1, 即变号数也没有改变. 所以证明了当区间 (a,b) 中没有 $f_0(x)$ 的根, 则在 x 从 a 增大到 b, 不经过 $f_0(x)$ 的根时, 变号函数 $w(x)$ 的值不改变.

(2) 余下讨论当 $c\in(a,b)$, 而且 c 为 $f_0(x)$ 的根, 则在 x 从 a 增大到 b 时, 变号函数 $w(x)$ 的值的改变如何?

由引理 1.6.4 的性质 (2), $f_1(c)\neq 0$. 所以存在 $\varepsilon>0$, 使当 $c-\varepsilon<x<c+\varepsilon$ 时 $f_1(x)\neq 0$. 于是由引理 1.6.4 的性质 (4), 有下表

	$f_0(x)=f(x)$	$f_1(x)=f'(x)$
$c-\varepsilon<x<c$	$\pm$	$\mp$
$x=c$	0	$\mp$
$c<x<c+\varepsilon$	$\mp$	$\mp$

因此当 x 从 a 增大到 b 的过程中, 在 $c-\varepsilon<x<c$ 时有一个变号数, 在 $c\leqslant x<c+\varepsilon$ 时没有变号数. 总之, 证明了当通过 $f(x)$ 的根时, 变号函数 $w(x)$ 的值减少一个 1. 这时 $f_1(x)\neq 0$, 于是对 Sturm 序列 $f_0(x), f_1(x), \cdots, f_s(x)$, 存在适当的 $\varepsilon>0$, 和**(1)** 同样讨论可知变号数仍不改变.

上面的讨论, 证明了 Sturm 序列 $f(x)=f_0(x)$, $f'(x)=f_1(x), \cdots, f_s(x)$ 的变号函数 $w(x)$ 只在经过 $f(x)$ 的根时减少了值 1. 因此如果 a 和 b 不是 $f(x)$ 的根, 则在区间 (a,b) 中有且只有 $w(b)-w(a)$ 个不同实根. 这证明了定理. □

注意 在使用 Sturm 定理时, 先考虑实多项式 $g(x)$ 及 $g'(x)$, 求出 $(g(x),g'(x))$. 再求实多项式 $f(x)=g(x)(g(x),g'(x))^{-1}$. 又在求出实多项式 $f(x)$ 的实根 c 后, 计算 $f'(c), f''(c), \cdots$, 从而可以算出实多项式 $g(x)$ 的实根 c 的重数.

习 题 1.6

1.6.1 试讨论下列实多项式的实根的分布:

(1) $x^3-\frac{7}{4}x^2+\frac{3}{4}x-\frac{3}{32}$; **(2)** x^n+px+q; **(3)** $nx^n-x^{n-1}-x^{n-2}-\cdots-x-1$.

§1.7 多元多项式和对称多项式*

给定域 $\mathbb{F}$ ($\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ 之一) 及域 $\mathbb{F}$ 上 n 个独立未知数 $x_1, x_2, \cdots, x_n$. 它们也称

为 n 个**不定元**. 给定 n 个非负整数 $k_1, k_2, \cdots, k_n$, 及数 $a_{k_1k_2\cdots k_n} \in \mathbb{F}$, 则

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

称为**单项式**. $a_{k_1k_2\cdots k_n}$ 称为这个单项式的**系数**. 而 $k_1+k_2+\cdots+k_n$ 称为这个单项式的**次数**. 单项式

$$a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

和

$$b_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$$

称为**同类型的项**, 或称为**同类项**.

定义 1.7.1 有限个单项式的和 $f(x_1,x_2,\cdots,x_n)$ 称为 $\boldsymbol{n}$ **元多项式**, 这时同类项已经合并了. n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 的**次数**为这有限个非零单项式的次数中的最大值, 记作 $\deg(f(x_1,x_2,\cdots,x_n))$.

我们约定**零多项式** $f(x_1,x_2,\cdots,x_n)=a_{0,0,\cdots,0}=0$ 的首项系数为零. 对零多项式, 定义次数为 $-\infty$.

零次多项式 $f(x_1,x_2,\cdots,x_n)=a_{0,0,\cdots,0}$ 的首项系数 $a_{0,0,\cdots,0}$ 不为零.

一般, n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 可以写为

$$f(x_1,x_2,\cdots,x_n)=\sum_{k_1=0}^{m_1}\sum_{k_2=0}^{m_2}\cdots\sum_{k_n=0}^{m_n}a_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}. \tag{1.7.1}$$

给定一个 n 元多项式 $f(x_1,x_2,\cdots,x_n)$, 可以有很多种不同的办法, 将它书写成单项式的和. 例如, 多项式 $f(x_1,x_2,x_3,x_4)$ 由三个项 $x_1^2x_2x_3^3x_4^3$, $x_1^3x_3^4$, $x_1x_3^4x_4^2$ 构成, 它可以写为下面三种表达式:

$$\begin{aligned} f(x_1,x_2,x_3,x_4) &= x_1^2x_2x_3^3x_4^3+x_1^3x_3^4+x_1x_3^4x_4^2 \\ &= x_1^3x_3^4+x_1x_3^4x_4^2+x_1^2x_2x_3^3x_4^3 \\ &= x_1x_3^4x_4^2+x_1^2x_2x_3^3x_4^3+x_1^3x_3^4 \end{aligned}$$

等. 这对讨论它的性质非常不方便. 为了使各项的次序确定起见, 我们用一种按照字典中字的排列方式, 把多元多项式 $f(x_1,x_2,\cdots,x_n)$ 的各项排成单项式, 即引进 **"字典排列法"**. 为此, 要约定两个不同类型的项

$$x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}, \quad x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n} \tag{1.7.2}$$

哪个在前, 哪个在后. 如果 $j_1=k_1, \cdots, j_r=k_r, j_{r+1}>k_{r+1}$, 那末就说 $x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 在 $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 之前, 这时丝毫不考虑 $j_{r+2}, \cdots, j_n$ 和 $k_{r+2}, \cdots, k_n$ 之间的关系. 例如, 对三个项 $x_1^2x_2x_3^3x_4^3$, $x_1^3x_3^4$, $x_1x_3^4x_4^2$ 按照字典排序法来排列, 那末次序为

$x_1^3x_3^4$, $x_1^2x_2x_3^3x_4^3$, $x_1x_3^4x_4^2$. 按照上面约定的次序, 可以将它写成 $f(x_1,x_2,x_3,x_4) = x_1^3x_3^4 + x_1^2x_2x_3^3x_4^3 + x_1x_3^4x_4^2$.

将 n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 按上面的字典排序法排起来. 这时有一个特点, 即次数最高的项不一定排在第一个位置, 也不一定排在最后一个位置. 例如, 在多项式 $f(x_1,x_2,x_3,x_4) = x_1^3x_3^4 + 2x_1^2x_2x_3^3x_4^3 + 7x_1x_3^4x_4^2$ 中, 次数最高的项为 $2x_1^2x_2x_3^3x_4^3$.

字典排列法不但确定了一种单项式的排列规则, 而且在对 n 个不定元 x_1, x_2, $\cdots$, x_n 的多项式作了字典排列法后, 再任取 $j=1, 2, \cdots, n$, 对 $n-j$ 个不定元 x_{j+1}, x_{j+2}, $\cdots$, x_n, 取这 $n-j$ 个不定元 $x_{j+1}=\cdots=x_n=1$, 那末得到 x_1, $\cdots$, x_j 的多项式, 按照原来次序仍然是字典排列法排列的.

当然, 单项式的和的排法不是唯一的. 有时为了讨论方便, 也可以采用下面一种排法. 为此引进

定义 1.7.2 如果一个 n 元多项式的各项次数相同, 则称为**齐次多项式**. 如果各项次数都是 k, 则称为 k 次齐次多项式.

显然设 $g(x) = g(x_1,x_2,\cdots,x_n)$ 为 k 次齐次多项式, 于是对一切参数 t, 则 $g(tx) = t^kg(x)$. 反之可证: 适合上式的多元多项式必为 k 次齐次多项式.

给定一个 m 次 n 元多项式 $f(x_1,x_2,\cdots,x_n)$, 将它的同为 k 次的单项式加在一起, 记作 $f_k(x_1,x_2,\cdots,x_n)$, $k=0, 1, \cdots, m$. 于是 $f(x_1,x_2,\cdots,x_n)$ 分解为 $m+1$ 个齐次多项式的和:

$$f(x_1,x_2,\cdots,x_n) = f_m(x_1,x_2,\cdots,x_n)+f_{m-1}(x_1,x_2,\cdots,x_n)+\cdots+f_0(x_1,x_2,\cdots,x_n), \tag{1.7.3}$$

其中 $f_m(x_1,x_2,\cdots,x_n) \neq 0$. 再对每个齐次多项式 $f_j(x_1,x_2,\cdots,x_n)$ 按照字典排列法排起来, 这样就把多项式 $f(x_1,x_2,\cdots,x_n)$ 的所有项按照一个确定的规律排成一个单项式的和, 而且每个单项式的次数也是按降幂排的.

显然, 零次齐次多项式为一个非零常数. 又零本身称为零多项式. 和一个未知数的情形相同, 我们可以很自然地引进两个 n 元多项式的加法、减法和乘法这三种代数运算, 所有这些运算的性质也和一元情形完全相同.

现在利用上面两种不同的排法, 给出多元多项式一些简单的运算性质.

引理 1.7.3 域 $\mathbb{F}$ 上 n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 和 $g(x_1,x_2,\cdots,x_n)$ 在按照字典排列法排成不同类单项式的和后, 它们的首项分别设为

$$a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}, \quad b_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}, \tag{1.7.4}$$

那末乘积 $f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)$ 在字典排列法排列后, 首项为

$$a_{j_1j_2\cdots j_n}b_{k_1k_2\cdots k_n}x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}. \tag{1.7.5}$$

换句话说, 按照字典排列法, 两个多项式乘积的首项等于两多项式的首项的乘积.

证 将多项式 $f(x_1,x_2,\cdots,x_n)$ 和 $g(x_1,x_2,\cdots,x_n)$ 按照字典排列法排好. 由于多项式

$$f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)$$

的任一项必由 $f(x_1,x_2,\cdots,x_n)$ 的项 $a_{r_1r_2\cdots r_n}x_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$ 和 $g(x_1,x_2,\cdots,x_n)$ 的项 $b_{s_1s_2\cdots s_n}x_1^{s_1}x_2^{s_2}\cdots x_n^{s_n}$ 作乘积, 则乘积是

$$a_{r_1r_2\cdots r_n}b_{s_1s_2\cdots s_n}x_1^{r_1+s_1}x_2^{r_2+s_2}\cdots x_n^{r_n+s_n}.$$

然后在这些乘积中取出同类型的项, 并相加而构成. 我们要证

$$a_{j_1j_2\cdots j_n}b_{k_1k_2\cdots k_n}x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}$$

是首项. 事实上, 显然 $j_1=r_1, \cdots, j_u=r_u, j_{u+1}>r_{u+1}$, 又 $k_1=s_1, \cdots, k_v=s_v, k_{v+1}>s_{v+1}$. 所以在 $u\geqslant v$ 时, 有 $j_1+k_1=r_1+s_1, \cdots, j_v+k_v=r_v+s_v, j_{v+1}+k_{v+1}>r_{v+1}+s_{v+1}$; 在 $u<v$ 时有 $j_1+k_1=r_1+s_1, \cdots, j_u+k_u=r_u+s_u, j_{u+1}+k_{u+1}>r_{u+1}+s_{u+1}$. 因此不管在哪种情形, $f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)$ 的任一项, 必然排在 $a_{j_1j_2\cdots j_n}b_{k_1k_2\cdots k_n}x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}$ 之后. 而且没有其他的同类项. 这就证明了引理. □

引理 1.7.4 设 n 元多项式 $f(x_1,x_2,\cdots,x_n)$ 和 $g(x_1,x_2,\cdots,x_n)$ 的次数分别为 m_1 和 m_2, 则乘积 $f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)$ 的次数为 m_1+m_2.

证 将 $f(x_1,x_2,\cdots,x_n)$ 和 $g(x_1,x_2,\cdots,x_n)$ 都分解为不同次数的齐次多项式的和, 即令

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)&=f_{m_1}(x_1,x_2,\cdots,x_n)+\cdots+f_0(x_1,x_2,\cdots,x_n),\\ g(x_1,x_2,\cdots,x_n)&=g_{m_2}(x_1,x_2,\cdots,x_n)+\cdots+g_0(x_1,x_2,\cdots,x_n),\end{aligned}$$

其中 $f_{m_1}(x_1,x_2,\cdots,x_n)g_{m_2}(x_1,x_2,\cdots,x_n)\neq 0$. 令

$$\begin{aligned}f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)&=\Big(\sum_{j=0}^{m_1}f_j(x_1,x_2,\cdots,x_n)\Big)\Big(\sum_{k=0}^{m_2}g_k(x_1,x_2,\cdots,x_n)\Big)\\&=\sum_{i=0}^{m_1+m_2}\sum_{j+k=i}f_j(x_1,x_2,\cdots,x_n)g_k(x_1,x_2,\cdots,x_n).\end{aligned}$$

自然, $f_j(x_1,x_2,\cdots,x_n)g_k(x_1,x_2,\cdots,x_n)$ 是 $j+k$ 次齐次多项式. 事实上, 令

$$\begin{aligned}f_j(x_1,x_2,\cdots,x_n)&=\sum_{j_1+\cdots+j_n=j}a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n},\\ g_k(x_1,x_2,\cdots,x_n)&=\sum_{k_1+\cdots+k_n=k}b_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n},\end{aligned}$$

于是

$$\begin{aligned}&f_j(x_1,x_2,\cdots,x_n)g_k(x_1,x_2,\cdots,x_n)\\=&\sum_{j_1+\cdots+j_n=j}\ \sum_{k_1+\cdots+k_n=k}a_{j_1j_2\cdots j_n}b_{k_1k_2\cdots k_n}x_1^{j_1+k_1}x_2^{j_2+k_2}\cdots x_n^{j_n+k_n}.\end{aligned}$$

所以多项式 $f_j(x_1,x_2,\cdots,x_n)g_k(x_1,x_2,\cdots,x_n)$ 的每一项的次数都是

$$j_1+k_1+j_2+k_2+\cdots+j_n+k_n=j+k.$$

故断言成立. 正因为如此, 所以齐次多项式 $f_{m_1}(x_1,x_2,\cdots,x_n)g_{m_2}(x_1,x_2,\cdots,x_n)$ 的次数为 m_1+m_2, 而且它比任一其他项 $f_j(x_1,x_2,\cdots,x_n)g_k(x_1,x_2,\cdots,x_n)$ 的次数都要大, 所以 $f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)$ 的次数为 m_1+m_2. □

现在来讨论一种极为有用的 n 元多项式. 先从一元多项式的根与系数关系的 Vieta 定理着手. 即引进不定元 x 及不定元 $x_1, x_2, \cdots, x_n$, 构作 x 的 n 次多项式

$$\begin{aligned}&(x-x_1)(x-x_2)\cdots(x-x_n)\\&=x^n+(-1)^1\sigma_1x^{n-1}+(-1)^2\sigma_2x^{n-2}\cdots+(-1)^{n-1}\sigma_{n-1}x+(-1)^n\sigma_n,\end{aligned}\tag{1.7.6}$$

因此有 **Vieta 定理**:

$$\begin{aligned}\sigma_1&=x_1+x_2+\cdots+x_n,\\\sigma_2&=x_1x_2+\cdots+x_1x_n+x_2x_3+\cdots+x_2x_n+\cdots+x_{n-1}x_n=\sum_{1\leqslant j_1<j_2\leqslant n}x_{j_1}x_{j_2},\\&\ \vdots\\\sigma_{n-1}&=\sum_{1\leqslant j_1<\cdots<j_{n-1}\leqslant n}x_{j_1}x_{j_2}\cdots x_{j_{n-1}},\\\sigma_n&=x_1x_2\cdots x_n.\end{aligned}\tag{1.7.7}$$

于是引进了 n 个特殊的 n 元多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$, 它们称为 n 元**初等对称多项式**. 由上面的表达式可以看出来, 将多项式的不定元 $x_1, x_2, \cdots, x_n$ 作任意方式的排列, 那末得到的多项式还是原来的多项式. 实际上, 这就相当于将乘积 $(x-x_1)(x-x_2)\cdots(x-x_n)$ 的次序重新排列一下. 将这个性质提炼出来, 便引进了

定义 1.7.5 n 个不定元 $x_1, x_2, \cdots, x_n$ 的多项式 $f(x_1,x_2,\cdots,x_n)$ 称为**对称多项式**, 如果

$$f(x_{i_1},x_{i_2},\cdots,x_{i_n})=f(x_1,x_2,\cdots,x_n),\tag{1.7.8}$$

其中 $i_1i_2\cdots i_n$ 遍历 $1, 2, \cdots, n$ 的所有排列. 换句话说, 对称多项式不因不定元的排法不同而不同.

例如, 初等对称多项式都是对称多项式; 又如齐次多项式

$$S_k = x_1^k + x_2^k + \cdots + x_n^k, \quad k = 1, 2, \cdots \tag{1.7.9}$$

也都是对称多项式, 它们称为 **Newton 对称幂和**, 或**等幂和**. 又如

$$x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 + x_1 + x_2 + x_3 + 4$$

是三个不定元 x_1, x_2, x_3 的对称多项式.

对称多项式有一个简单的性质: 将对称多项式 $f(x_1, x_2, \cdots, x_n)$ 按照齐次多项式的方式分解为

$$f(x_1, x_2, \cdots, x_n) = \sum_{k=0}^{m} f_k(x_1, x_2, \cdots, x_n), \tag{1.7.10}$$

则每个 k 次齐次多项式 $f_k(x_1, x_2, \cdots, x_n)$ 也都是对称多项式, 反之亦然. 另外, 显然, 对称多项式的和、差、积仍为对称多项式. 再, 由于排列是由标准排列 $12\cdots n$ 经过一系列对换 (即相邻两个文字的互换) 得到 (见 §2.1), 所以为了证明多项式 $f(x_1, x_2, \cdots, x_n)$ 是对称多项式, 只要证明在不定元 x_1, x_2, $\cdots$, x_n 的任一对换下多项式不改变就行了. 最后, 将初等对称多项式 σ_1, σ_2, $\cdots$, σ_n 看作 n 个不定元. 构作域 $\mathbb{F}$ 上 n 元多项式

$$g(\sigma_1, \sigma_2, \cdots, \sigma_n) = \sum_{k_1=0}^{m_1} \sum_{k_2=0}^{m_2} \cdots \sum_{k_n=0}^{m_n} a_{k_1k_2\cdots k_n} \sigma_1^{k_1} \sigma_2^{k_2} \cdots \sigma_n^{k_n}, \tag{1.7.11}$$

再将 σ_j 用不定元 $x_1, x_2, \cdots, x_n$ 代入, 于是 $g(\sigma_1, \sigma_2, \cdots, \sigma_n)$ 是不定元 $x_1, x_2, \cdots, x_n$ 的多项式, 记作 $f(x_1, x_2, \cdots, x_n)$. 显然它仍为对称多项式, 且 $g(\sigma_1, \sigma_2, \cdots, \sigma_n)$ 的每一项 $a_{k_1k_2\cdots k_n} \sigma_1^{k_1} \sigma_2^{k_2} \cdots \sigma_n^{k_n}$ 作为 x_1, x_2, $\cdots$, x_n 的对称多项式, 则是 $k_1 + 2k_2 + \cdots + nk_n$ 次齐次多项式, 且将它按照字典排列法排成单项后, 其首项为

$$a_{k_1k_2\cdots k_n} x_1^{k_1+k_2+\cdots+k_n} x_2^{k_2+k_3+\cdots+k_n} \cdots x_n^{k_n}. \tag{1.7.12}$$

关于对称多项式, 有

定理 1.7.6 (对称多项式基本定理) 任一对称多项式 $f(x_1, x_2, \cdots, x_n)$ 一定能够表示成初等对称多项式 σ_1, σ_2, $\cdots$, σ_n 的多项式, 且表示法唯一.

证 将对称多项式 $f(x_1, x_2, \cdots, x_n)$ 按照字典排列法排成单项式的和. 设首项为

$$a_{j_1j_2\cdots j_n} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}.$$

下面先来证明 $j_1 \geqslant j_2 \geqslant \cdots \geqslant j_n$. 设若不然, 则存在指标 $j_i < j_{i+1}$. 因为 $f(x_1, x_2, \cdots, x_n)$ 是对称多项式, 所以

$$f(x_1, \cdots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \cdots, x_n) = f(x_1, \cdots, x_n),$$

因此在 $f(x_1,\cdots,x_n)$ 中包含项

$$a_{j_1j_2\cdots j_n}x_1^{j_1}\cdots x_{i-1}^{j_{i-1}}x_i^{j_{i+1}}x_{i+1}^{j_i}x_{i+2}^{j_{i+2}}\cdots x_n^{j_n}.$$

然而, 按照字典排列法, 首项是 $a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$, 所以 $j_i\geqslant j_{i+1}$, 这和 $j_i<j_{i+1}$ 矛盾, 所以证明了 $j_1\geqslant j_2\geqslant\cdots\geqslant j_n$.

作对称多项式

$$\varphi_1(x_1,x_2,\cdots,x_n)=a_{j_1j_2\cdots j_n}\sigma_1^{j_1-j_2}\sigma_2^{j_2-j_3}\cdots\sigma_{n-1}^{j_{n-1}-j_n}\sigma_n^{j_n},$$

由引理 1.7.3 可知: 它的首项为

$$\begin{aligned}&a_{j_1j_2\cdots j_n}x_1^{j_1-j_2}(x_1x_2)^{j_2-j_3}\cdots(x_1x_2\cdots x_{n-1})^{j_{n-1}-j_n}(x_1x_2\cdots x_{n-1}x_n)^{j_n}\\=&\,a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_{n-1}^{j_{n-1}}x_n^{j_n}.\end{aligned}$$

所以 $f(x_1,x_2,\cdots,x_n)$ 和 $\varphi_1(x_1,x_2,\cdots,x_n)$ 有相同的首项. 因此对称多项式

$$f_1(x_1,x_2,\cdots,x_n)=f(x_1,x_2,\cdots,x_n)-\varphi_1(x_1,x_2,\cdots,x_n)$$

的项, 在按照字典排列法排成单项式的和后, 首项 $a'_{k_1k_2\cdots k_n}x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 必然在

$$a_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$$

之后, 且 $k_1\geqslant k_2\geqslant\cdots\geqslant k_n$. 因此

$$j_1=k_1,\quad\cdots,\quad j_r=k_r,\quad j_{r+1}>k_{r+1}\geqslant k_{r+2}\geqslant\cdots\geqslant k_n.$$

再作对称多项式

$$\varphi_2(x_1,x_2,\cdots,x_n)=a'_{k_1k_2\cdots k_n}\sigma_1^{k_1-k_2}\sigma_2^{k_2-k_3}\cdots\sigma_{n-1}^{k_{n-1}-k_n}\sigma_n^{k_n},$$

则对称多项式 $f-\varphi_1-\varphi_2$ 的首项为 $a''_{t_1t_2\cdots t_n}x_1^{t_1}x_2^{t_2}\cdots x_n^{t_n}$, 它也适合 $t_1\geqslant t_2\geqslant\cdots\geqslant t_n$, 又

$$k_1=t_1,\cdots,k_u=t_u,k_{u+1}>t_{u+1}\geqslant t_{u+2}\geqslant\cdots\geqslant t_n.$$

这样依次作下去, 最后得到一串对称多项式 $\varphi_1,\varphi_2,\cdots$, 而对称多项式

$$f,\quad f-\varphi_1,\quad f-\varphi_1-\varphi_2,\quad\cdots$$

的首项都不相同, 且这些项依次排好, 就是按照字典排列法排列的. 另一方面, 每个首项 $a^{(k)}_{p_1p_2\cdots p_n}x_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$ 又适合 $p_1\geqslant\cdots\geqslant p_n$ 及 $j_1\geqslant p_1$. 显然, 适合条件 $j_1\geqslant p_1\geqslant\cdots\geqslant p_n$ 的非负整数组 $(p_1,\cdots,p_n)$ 共有有限组. 所以经过有限步后, 便得到有限个以初等对称多项式 $\sigma_1,\sigma_2,\cdots,\sigma_n$ 为不定元的多项式 $\varphi_1,\varphi_2,\cdots,\varphi_s$, 使得

$$f=\varphi_1+\cdots+\varphi_s.$$

这就证明了任一对称多项式是初等对称多项式的多项式.

下面来证明表示法的唯一性. 即证明: 如果两个初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式 $g_1(\sigma_1,\cdots,\sigma_n)$ 和 $g_2(\sigma_1,\cdots,\sigma_n)$ 作为 $x_1, x_2, \cdots, x_n$ 的多项式是相等的, 那末 g_1 和 g_2 作为不定元 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式也是相等的. 换句话说, 只要证明: 如果多项式 $g(\sigma_1,\cdots,\sigma_n)$ 作为 $x_1, x_2, \cdots, x_n$ 的多项式等于零, 那末作为 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式也是等于零的. 我们用反证法来证明这点. 如果 $g(\sigma_1,\cdots,\sigma_n)$ 作为 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的多项式不等于零, 那末在 $g(\sigma_1,\cdots,\sigma_n)$ 中任取两个不同项 $a_{j_1j_2\cdots j_n}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}$ 和 $a_{k_1k_2\cdots k_n}\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n}$, 它们作为 $x_1, x_2, \cdots, x_n$ 的多项式, 首项各为

$$a_{j_1j_2\cdots j_n}x_1^{j_1}(x_1x_2)^{j_2}\cdots(x_1x_2\cdots x_n)^{j_n}, \quad a_{k_1k_2\cdots k_n}x_1^{k_1}(x_1x_2)^{k_2}\cdots(x_1x_2\cdots x_n)^{k_n},$$

显然这是不同类型的项, 因此 $g(\sigma_1,\cdots,\sigma_n)$ 的每一项, 作为 $x_1, x_2, \cdots, x_n$ 的多项式, 首项两两不同. 将这些首项按照字典排列法排成单项式的和后, 自然有一个首项. 这个首项就是 $g(\sigma_1,\cdots,\sigma_n)$ 作为 $x_1, x_2, \cdots, x_n$ 的多项式的首项. 因此 $g(\sigma_1,\cdots,\sigma_n)$ 作为 $x_1, x_2, \cdots, x_n$ 的多项式, 它不恒为零. □

上面的定理实际上也证明了, 元素 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 作为域 $\mathbb{F}$ 上的不定元是独立未知数. 而且上面定理的证明也给出了一种将对称多项式表为初等对称多项式的具体算法. 下面先叙述方法, 再举一例.

在下面的计算, 实际上用到了对称多项式的两个特点: 首先, 是将对称多项式分解为齐次多项式的和后, 每个齐次多项式仍为对称多项式. 因此在具体计算时, 只要对每个齐次对称多项式来计算就行了; 其次, 是将对称多项式 $f(x_1,x_2,\cdots,x_n)$ 表示成初等对称多项式的多项式后, 出现的每一项 $a_{j_1j_2\cdots j_n}\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}$, 作为 $x_1, x_2, \cdots, x_n$ 的多项式, 首项为

$$a_{j_1j_2\cdots j_n}x_1^{j_1+j_2+\cdots+j_n}x_2^{j_2+\cdots+j_n}\cdots x_n^{j_n}.$$

这些项都在 $f(x_1,x_2,\cdots,x_n)$ 的首项之后. 由于这两点, 所以计算方案为:

(I) 给定对称多项式 $f(x_1,x_2,\cdots,x_n)$ 后, 把它分解为齐次多项式的和, 然后对每个齐次部分计算. 所以下面只需要给出齐次对称多项式的计算方案;

(II) 设 $f(x_1,x_2,\cdots,x_n)$ 是 m 次齐次对称多项式, 它的首项为

$$a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},$$

则 $i_1 \geqslant i_2 \geqslant \cdots \geqslant i_n$, 且 $i_1+i_2+\cdots+i_n=m$. 写出所有排在首项以后的项 $a_{p_1p_2\cdots p_n}x_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$, 使得 $p_1 \geqslant p_2 \geqslant \cdots \geqslant p_n$, 又 $p_1+p_2+\cdots+p_n=m$;

(III) 每给一项 $a_{p_1p_2\cdots p_n}x_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$, 构作一个多项式

$$A_{p_1p_2\cdots p_n}\sigma_1^{p_1-p_2}\sigma_2^{p_2-p_3}\cdots\sigma_n^{p_n},$$

其中 $A_{p_1p_2\cdots p_n}$ 为待定系数.

(IV) 于是, $f(x_1,x_2,\cdots,x_n)$ 可写成

$$f(x_1,x_2,\cdots,x_n)=\sum A_{p_1p_2\cdots p_n}\sigma_1^{p_1-p_2}\sigma_2^{p_2-p_3}\cdots\sigma_n^{p_n}.$$

再来求系数 $A_{p_1p_2\cdots p_n}$. 求的方法是将 $x_1, x_2, \cdots, x_n$ 代以一些特殊的数值, 例如

$$x_1=x_2=\cdots=x_r=1,\quad x_{r+1}=x_{r+2}=\cdots=x_n=0$$

等, 便得到系数 $A_{p_1p_2\cdots p_n}$ 所适合的线性方程组, 求其解便得到所要的值.

例 1.7.1 将对称多项式

$$f(x_1,x_2,\cdots,x_n)=\sum_{1\leqslant j_1<j_2<j_3\leqslant n}(x_{j_1}^2x_{j_2}^2x_{j_3}+x_{j_1}^2x_{j_2}x_{j_3}^2+x_{j_1}x_{j_2}^2x_{j_3}^2)$$

表示成初等对称多项式的多项式.

解 已知 $f(x_1,x_2,\cdots,x_n)$ 是五次齐次多项式, 它的首项为 $x_1^2x_2^2x_3$. 于是在它以后的项有 $x_1^2x_2x_3x_4$, $x_1x_2x_3x_4x_5$, 即

$$\begin{aligned}g(\sigma_1,\sigma_2,\cdots,\sigma_n)&=\sigma_1^{2-2}\sigma_2^{2-1}\sigma_3^1+A\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^1+B\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-1}\sigma_5^1\\&=\sigma_2\sigma_3+A\sigma_1\sigma_4+B\sigma_5.\end{aligned}$$

取 $x_1=x_2=x_3=x_4=1$, $x_5=\cdots=x_n=0$, 于是 $\sigma_1=4$, $\sigma_2=\binom{4}{2}=6$, $\sigma_3=\binom{4}{3}=4$, $\sigma_4=1$. 又 $f(1,1,1,1,0,\cdots,0)=3\times4=12$, 故有 $12=24+4A$, 即 $A=-3$. 再取 $x_1=x_2=x_3=x_4=x_5=1$, $x_6=\cdots=x_n=0$, 于是 $\sigma_1=\binom{5}{1}=5$, $\sigma_2=\binom{5}{2}=10$, $\sigma_3=\binom{5}{3}=10$, $\sigma_4=\binom{5}{4}=5$, $\sigma_5=1$. 又 $f(1,1,1,1,1,0,\cdots,0)=3\times\binom{5}{3}=30$, 故有 $30=100-3\times25+B$, 即 $B=5$. 于是,

$$f(x_1,x_2,\cdots,x_n)=\sigma_2\sigma_3-3\sigma_1\sigma_4+5\sigma_5.$$

例 1.7.2 等幂和 $S_k=\sum\limits_{j=1}^n x_j^k$, $k=1,2,\cdots$ 和初等对称多项式 $\sigma_1,\cdots,\sigma_n$ 有如下的 Newton 公式:

$$\begin{aligned}&S_k-S_{k-1}\sigma_1+S_{k-2}\sigma_2-\cdots+(-1)^{k-1}S_1\sigma_{k-1}+(-1)^k k\sigma_k=0,\quad k\leqslant n,\\&S_k-S_{k-1}\sigma_1+S_{k-2}\sigma_2-\cdots+(-1)^nS_{k-n}\sigma_n=0,\quad k>n.\end{aligned}$$

利用这两个递推关系式, 可以将等幂和表示成初等对称多项式的多项式.

为了书写方便, 今后记 $S(x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n})$ (其中 $j_1\geqslant j_2\geqslant\cdots\geqslant j_n$) 是由对 $x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$ 作所有可能的不定元排列所得的齐次对称多项式. 例如:

$$S(x_1^2x_2^2)=\sum_{1\leqslant j_1<j_2\leqslant n}x_{j_1}^2x_{j_2}^2.$$

今显然有

$$\begin{aligned}
S_{k-1}\sigma_1 &= S_k + S(x_1^{k-1}x_2),\\
S_{k-2}\sigma_2 &= S(x_1^{k-1}x_2) + S(x_1^{k-2}x_2x_3),\\
&\ \ \vdots\\
S_{k-j}\sigma_j &= S(x_1^{k-j+1}x_2\cdots x_j) + S(x_1^{k-j}x_2\cdots x_jx_{j+1}),\\
&\ \ \vdots
\end{aligned}$$

及当 $k \leqslant n$ 时,

$$S_1\sigma_{k-1} = S(x_1^2x_2\cdots x_{k-1}),$$

当 $k > n$ 时,

$$S_{k-n}\sigma_n = S(x_1^{k-n+1}x_2\cdots x_n).$$

在上述关系中消去 $S(x_1^{k-j+1}x_2\cdots x_j)$, $j = 2,3,\cdots,n$, 即上述各式中依次乘以 $+1$ 及 -1, 然后相加, 便得到所要的公式.

习　题　1.7

1.7.1 试对自变量 (不定元) 的个数 n 作归纳法, 以证明域 $\mathbb{F}$ 上对称多项式基本定理中的表示法唯一性.

1.7.2 试用域 $\mathbb{F}$ 上初等对称多项式表示出下列对称多项式:

(1) $S(x_1^2x_2)$; **(2)** $S(x_1^2x_2^2)$; **(3)** $\sum\limits_{1\leqslant j<k\leqslant n}(x_j-x_k)^2$;

(4) $(x_1+x_2-x_3-x_4)(x_1-x_2+x_3-x_4)(x_1-x_2-x_3+x_4)$;

(5) $(-x_1+x_2+\cdots+x_n)(x_1-x_2+x_3+\cdots+x_n)\cdots(x_1+\cdots+x_{n-1}-x_n)$;

(6) $\sum\limits_{j=1}^{n}\dfrac{1}{x_j}$; **(7)** $\sum\limits_{j\neq k}\dfrac{x_j}{x_k}$; **(8)** $\sum\limits_{j\neq k,p}\sum\limits_{k<p}\dfrac{x_kx_p}{x_j}$.

1.7.3 试求: $\sum\limits_{j=1}^{n}\dfrac{\partial\sigma_k}{\partial x_j}$, $1\leqslant k\leqslant n$.

1.7.4 将域 $\mathbb{F}$ 上对称多项式 $f(x_1,x_2,\cdots,x_n)$ 及 $\sum\limits_{j=1}^{n}\dfrac{\partial f(x_1,x_2,\cdots,x_n)}{\partial x_j}$ 同时表示为初等对称多项式的多项式, 并求它们间的关系.

1.7.5 任意给定正整数 m. 试证:

$$(x_1+x_2+\cdots+x_n)^m = \sum_{i_1+i_2+\cdots+i_n=m}\frac{m!}{i_1!i_2!\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}.$$

1.7.6 试用方幂和 $S_k = x_1^k+\cdots+x_n^k$, $k=0,1,\cdots$ 表示出:

(1) $\sum\limits_{1\leqslant j<k\leqslant n}(x_j+x_k)^2$; **(2)** $\sum\limits_{1\leqslant j<k\leqslant n}(x_j-x_k)^2$.

1.7.7 试解方程组

$$
\begin{aligned}
S_1 &= x_1 + \cdots + x_n = a,\\
S_2 &= x_1^2 + \cdots + x_n^2 = a,\\
&\vdots\\
S_n &= x_1^n + \cdots + x_n^n = a
\end{aligned}
$$

并求 $S_{n+1} = x_1^{n+1} + \cdots + x_n^{n+1}$ 的值, 其中 $a \in \mathbb{C}$.

1.7.8 试计算复多项式

$$x^n + (a+b)x^{n-1} + (a^2+ab+b^2)x^{n-2} + \cdots + (a^n + a^{n-1}b + \cdots + ab^{n-1} + b^n)$$

的根的方幂和 S_1, S_2, $\cdots$, S_n.

1.7.9 设域 $\mathbb{F}$ 上多项式 $f(x_1, x_2, \cdots, x_n)$ 在不定元的偶排列下不变. 试证:

$$f(x_1, x_2, \cdots, x_n) = f_1(x_1, x_2, \cdots, x_n) + f_2(x_1, x_2, \cdots, x_n)V(x_1, x_2, \cdots, x_n),$$

其中 f_1 及 f_2 都是对称多项式, 而 $V(x_1, x_2, \cdots, x_n)$ 是不定元 x_1, x_2, $\cdots$, x_n 的 Vandermonde 行列式.

1.7.10 设域 $\mathbb{F}$ 上多项式 $f(x_1, x_2, \cdots, x_n)$ 在不定元的轮回排列, 即 n 个特殊的排列

$$12\cdots n, \quad n12\cdots(n-1), \quad \cdots, \quad 23\cdots n1$$

下不变. 试证:

$$f(x_1, x_2, \cdots, x_n) = \sum a_{m_0 m_1 \cdots m_{n-1}} g_0^{m_0} g_1^{m_1} \cdots g_{n-1}^{m_{n-1}},$$

其中和号遍历所有非负整数 m_0, m_1, $\cdots$, m_{n-1}, 使得 $m_1 + 2m_2 + \cdots + (n-1)m_{n-1}$ 能被 n 除得尽, $m_0 + m_1 + \cdots + m_{n-1} = \deg(f)$, 且 $g_j = x_1\varepsilon^j + x_2\varepsilon^{2j} + \cdots + x_n\varepsilon^{nj}$, $0 \leqslant j \leqslant n-1$. 又 $\varepsilon = \exp\left(\dfrac{2\pi i}{n}\right)$ 是多项式 $x^n - 1$ 的**本原单位根**.

1.7.11 试求实多元多项式

$$\frac{x_1^2}{2} + \frac{x_2^{2^2}}{2^2} + \cdots + \frac{x_n^{2^n}}{2^n} - x_1x_2\cdots x_n + \frac{1}{2^n}$$

的所有实根.

1.7.12 给定复多项式序列 $\{p_n(x,y,z)\}_{n=0}^{\infty}$, 其中 $p_0(x,y,z) = 1$. 如果有递推公式

$$p_n(x,y,z) = (x+z)(y+z)p_{n-1}(x,y,z+1) - z^2 p_{n-1}(x,y,z), \quad n = 1, 2, \cdots,$$

试证: $p_n(x,y,z)$, $n = 1$, 2, $\cdots$ 为对称多项式.

第二章　行列式理论

§2.1　排列

为了定义行列式, 需要引进**排列**的概念. 考虑 n 个正整数 $1, 2, \cdots, n$. 这 n 个数字可以按照各种各样方式来进行排列. 例如, 两个正整数 1, 2 可以排成下面两种形状: 12, 21; 又如三个正整数 1, 2, 3 可以排成下面六种形状: 123, 132, 213, 231, 312, 321. 一般地说, n 个数字 $1, 2, \cdots, n$ 的排列可以按照下面方法作出来: 即先在集合 $S_n = \{1, 2, \cdots, n\}$ 中取出一个正整数, 记作 i_1, 将它放在第一个位置. 由于 i_1 可以是 $1, 2, \cdots, n$ 中任一正整数, 所以选取 i_1 的方法有 n 种. 再在集合 S_n 中除去正整数 i_1, 在余下的 $n-1$ 个正整数中取出一个正整数, 记作 i_2, 将它放在第二个位置. 显然选取 i_2 的方法有 $n-1$ 种. 这样依次选下去, 经过 n 次选取, 便得到 $1, 2, \cdots, n$ 的一个排列 $i_1 i_2 \cdots i_n$. 注意: 这 n 个有顺序的数字需看作一个整体. 由排列的构造法可知, $i_1, i_2, \cdots, i_n$ 表示了 n 个不同的正整数. 而且 $i_1, i_2, \cdots, i_n$ 中每个正整数是集合 S_n 中的一个元素. 换句话说,

$$1 \to i_1, \quad 2 \to i_2, \cdots, n \to i_n \tag{2.1.1}$$

定义了集合 S_n 到自身上的一个一一对应. 这个一一对应又可以用符号

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \tag{2.1.2}$$

记之, 称为**置换**. 由于置换仅仅刻画了一一对应, 而上述一一对应又可以改写为

$$j_1 \to i_{j_1}, \quad j_2 \to i_{j_2}, \quad \cdots, \quad j_n \to i_{j_n}, \tag{2.1.3}$$

其中 $j_1j_2\cdots j_n$ 是 $1, 2, \cdots, n$ 的另一个排列, 所以这个一一对应也可以用符号

$$\begin{pmatrix} j_1 & j_2 & \cdots & j_n \\ i_{j_1} & i_{j_2} & \cdots & i_{j_n} \end{pmatrix} \tag{2.1.4}$$

记之. 因此对 $1, 2, \cdots, n$ 的任一个排列 $j_1j_2\cdots j_n$, 排列 $i_1i_2\cdots i_n$ 可记为

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \cdots & j_n \\ i_{j_1} & i_{j_2} & \cdots & i_{j_n} \end{pmatrix}. \tag{2.1.5}$$

在构作排列时, 因为各个位置的文字分别有 $n, n-1, \cdots, 1$ 种不同的选法. 所以总共有 n 个正整数 $1, 2, \cdots, n$ 的 $n!$ 个不同的排列, 亦即集合 S_n 总共有 $n!$ 个不同的自身到自身上的一一对应.

任取一个排列 $i_1i_2\cdots i_n$, 将其中两个相邻的正整数 i_{j-1}, i_j 对换一下, 便造出一个新的排列 $i_1i_2\cdots i_{j-2}i_ji_{j-1}i_{j+1}\cdots i_n$, 称为原来排列的**对换排列**. 这样一种步骤称为**对换**.

对任一排列 $i_1i_2\cdots i_n$, 若有两个前面的大于后面的相邻的正整数 i_{j-1}, i_j, 就作一次对换. 这样, 依次地对排列作下去, 就可以把任一排列 $i_1i_2\cdots i_n$ 经过若干次对换变成**标准排列** $12\cdots n$.

例如 $321 \to 312 \to 132 \to 123$. 但是将排列 321 变为标准排列 123 的对换方式不止一种. 例如 $321 \to 312 \to 321 \to 231 \to 213 \to 123$. 由此可见, 将排列 $i_1i_2\cdots i_n$ 变成标准排列 $12\cdots n$ 的对换方式, 也可以不同. 不过不管经过什么途径作对换, 在给定排列 $i_1i_2\cdots i_n$ 后, 关于对换的次数有下面的重要定理.

定理 2.1.1 如果有一种对换方式, 经过作用偶 (奇) 数次对换, 将排列 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$, 那末, 不管用哪种对换方式将排列 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$, 都需要作用偶 (奇) 数次对换. 换句话说, 对换次数的奇偶性和对换方式无关.

证 对 n 作归纳法. 当 $n=1$ 时定理显然成立. 设对 $n-1$ 个正整数 $1, 2, \cdots, n-1$ 的排列, 定理成立. 我们来证对 n 个正整数 $1, 2, \cdots, n$ 的排列, 定理也成立.

任取 $1, 2, \cdots, n$ 的一个排列 $i_1i_2\cdots i_n$, 其中必有一个正整数 $i_j = n$. 现在来考虑将 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$ 的某种对换方式. 将这种对换方式中所有对换构成的集合记作 $\mathfrak{S}$. 由于对换是相邻两个正整数改变次序, 两个正整数中有一个是 n 的对换全体组成 $\mathfrak{S}$ 的一个子集合, 记作 $\mathfrak{S}_n$. 又, 两个正整数都不是 n 的对换全体组成 $\mathfrak{S}$ 的另一子集合, 记作 $\mathfrak{S}_0$. 由于任一对换必为这两种对换之一, 所以 $\mathfrak{S} = \mathfrak{S}_0 \cup \mathfrak{S}_n$, $\mathfrak{S}_0 \cap \mathfrak{S}_n = \varnothing$.

先来考察集合 $\mathfrak{S}_0$. 假设 $\mathfrak{S}_0$ 中有 q 个元素. 由于 $i_j = n$, 这一类对换是变动 $n-1$ 个正整数 $i_1, \cdots, i_{j-1}, i_{j+1}, \cdots, i_n$ 的对换, 它们实际上是变动 $1, 2, \cdots, n-1$. 所以

对它们可以用归纳法假设. 经过这一系列对换, 其结果是将排列 $i_1\cdots i_{j-1}i_{j+1}\cdots i_n$ 变为标准排列 $12\cdots(n-1)$. 由归纳法假设, 对换次数 q 的奇偶性和对换方式无关.

再来考察集合 $\mathfrak{S}_n$. 假设 $\mathfrak{S}_n$ 中有 p 个元素, 那末只需证明数值 p 的奇偶性和对换方式无关. 为此计算数值 p. 因为集合 $\mathfrak{S}_n$ 中的对换全体构成了将 $i_j=n$ 变到第 n 个位置上去的对换过程, 在这个过程中 $i_j=n$ 必然要分别和第 $j+1, j+2, \cdots, n$ 个位置上的正整数改变次序, 所以一定包含 $n-j$ 次对换. 除了这些对换外, 其他的对换是由于正整数 n 在某个位置 (例如第 k 个位置) 往前跳造成的, 可是只要往前跳一个位置, 例如这时跳到第 $k-1$ 个位置, 那末后面一定有一个对换是将 n 从第 $k-1$ 个位置对换到第 k 个位置, 否则数 n 便不能到达第 n 个位置了. 假设往回跳的总对换次数为 m, 上面的讨论证明了 $p=n-j+2m$. 所以数 p 和 $n-j$ 的奇偶性相同. 然而数值 $n-j$ 只由排列 $i_1i_2\cdots i_n$ 本身决定, 和对换方式无关. 这就证明了定理. □

利用这个定理, 我们可以引进

定义 2.1.2 正整数 $1, 2, \cdots, n$ 的排列 $i_1i_2\cdots i_n$ 称为**偶 (奇) 排列**, 如果有一种对换方式, 经过偶 (奇) 数次对换后, 可以将排列 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$.

设排列 $i_1i_2\cdots i_n$ 经过 s 次对换后变成标准排列 $12\cdots n$, 则在 $i_1i_2\cdots i_n$ 为偶排列时数值为 $(-1)^s=+1$; 在 $i_1i_2\cdots i_n$ 为奇排列时数值为 $(-1)^s=-1$, 所以数值 $(-1)^s$ 和对换方式无关. 将它改写为

$$\delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}=(-1)^s=\begin{cases}1, & \text{当}\quad i_1i_2\cdots i_n\quad\text{是}\quad 1,2,\cdots,n\quad\text{的偶排列时;}\\ -1, & \text{当}\quad i_1i_2\cdots i_n\quad\text{是}\quad 1,2,\cdots,n\quad\text{的奇排列时.}\end{cases}\tag{2.1.6}$$

上面讨论了 n 个确定的正整数 $1, 2, \cdots, n$ 的排列. 它可以看作是集合 $\mathfrak{S}_n=\{1, 2, \cdots, n\}$ 到自身上的一个一一对应. 将这个概念推广, 任取 n 个元素的集合 $\mathfrak{S}_n=\{a_1, a_2, \cdots, a_n\}$, 对于集合 $\mathfrak{S}_n$ 到自身上的一一对应 $a_1\to a_{i_1}, a_2\to a_{i_2}, \cdots, a_n\to a_{i_n}$ 称为 $a_1, a_2, \cdots, a_n$ 的一个排列, 记作 $a_{i_1}a_{i_2}\cdots a_{i_n}$. 容易看出 $a_{i_1}a_{i_2}\cdots a_{i_n}$ 是 $a_1, a_2, \cdots, a_n$ 的排列的充分且必要条件为 $i_1i_2\cdots i_n$ 是 $1, 2, \cdots, n$ 的排列.

自然, $a_1, a_2, \cdots, a_n$ 的排列也共有 $n!$ 个. 又, 将排列 $a_{i_1}a_{i_2}\cdots a_{i_n}$ 变为标准排列 $a_1a_2\cdots a_n$ 的对换总次数的奇偶性也和对换的方式无关. 所以也可以定义偶排列及奇排列. 同样也可以引进符号

$$\delta^{a_1\ a_2\ \cdots\ a_n}_{a_{i_1}a_{i_2}\cdots a_{i_n}}=(-1)^s=\begin{cases}1, & \text{当}\quad i_1i_2\cdots i_n\quad\text{是}\quad 1,2,\cdots,n\quad\text{的偶排列时;}\\ -1, & \text{当}\quad i_1i_2\cdots i_n\quad\text{是}\quad 1,2,\cdots,n\quad\text{的奇排列时,}\end{cases}\tag{2.1.7}$$

其中 s 是某一种将 $a_{i_1}a_{i_2}\cdots a_{i_n}$ 变为标准排列 $a_1a_2\cdots a_n$ 的对换方式的对换总次数.

下面给出排列的一系列性质.

定理 2.1.3 将正整数 $1, 2, \cdots, n$ 的排列 $i_1i_2\cdots i_n$ 的第 j 个数 i_j 和第 k 个数 i_k (其中 $j<k$) 互换后, 所得的排列的奇偶性改变, 即

$$\delta^{1\ \cdots\ j\cdots\ k\cdots\ n}_{i_1\cdots i_k\cdots i_j\cdots i_n} = -\delta^{1\ \cdots\ j\cdots\ k\cdots\ n}_{i_1\cdots i_j\cdots i_k\cdots i_n}. \tag{2.1.8}$$

证 设排列 $i_1\cdots i_j\cdots i_k\cdots i_n$ 经过 s 次对换变为标准排列 $12\cdots n$. 现在将排列 $i_1\cdots i_{j-1}i_ki_{j+1}\cdots i_{k-1}i_ji_{k+1}\cdots i_n$ 按照下面的对换方式变为标准排列 $12\cdots n$. 先将 i_k 从第 j 个位置依次和 $i_{j+1}, \cdots, i_{k-1}, i_j$ 这 $k-j$ 个正整数作对换, 变到第 k 个位置, 于是变成排列 $i_1\cdots i_{j-1}i_{j+1}\cdots i_{k-1}i_ji_ki_{k+1}\cdots i_n$. 再将 i_j 从第 $k-1$ 个位置依次和 $i_{k-1}, \cdots, i_{j+1}$ 这 $k-j-1$ 个正整数作对换, 变到第 j 个位置, 于是变成排列 $i_1\cdots i_j\cdots i_k\cdots i_n$. 再经过 s 次对换便能变成标准排列 $12\cdots n$. 所以排列 $i_1\cdots i_k\cdots i_j\cdots i_n$ 可以经过 $s+2(k-j)-1$ 次对换变成标准排列 $12\cdots n$. 今

$$\delta^{1\ \cdots\ j\cdots\ k\cdots\ n}_{i_1\cdots i_j\cdots i_k\cdots i_n} = (-1)^s, \quad \delta^{1\ \cdots\ j\cdots\ k\cdots\ n}_{i_1\cdots i_k\cdots i_j\cdots i_n} = (-1)^{s+2(k-j)-1} = -(-1)^s,$$

所以定理成立. □

定理 2.1.4 设 $i_1i_2\cdots i_n$ 是 $1, 2, \cdots, n$ 的一个排列. 则 $12\cdots n$ 是 $i_1, i_2, \cdots, i_n$ 的一个排列, 且

$$\delta^{i_1i_2\cdots i_n}_{1\ 2\ \cdots\ n} = \delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}. \tag{2.1.9}$$

证 显然 $12\cdots n$ 是 $i_1, i_2, \cdots, i_n$ 的排列. 对任一将排列 $i_1i_2\cdots i_n$ 变为标准排列 $12\cdots n$ 的对换过程, 将它们依次全部逆过来, 得到一种将 $i_1, i_2, \cdots, i_n$ 的排列 $12\cdots n$ 变为标准排列 $i_1i_2\cdots i_n$ 的对换过程, 所以对换次数相同. □

定理 2.1.5 设 $j_1j_2\cdots j_n$ 和 $k_1k_2\cdots k_n$ 都是 $1, 2, \cdots, n$ 的排列, 则

$$\delta^{1\ 2\ \cdots\ n}_{j_1j_2\cdots j_n}\delta^{j_1j_2\cdots j_n}_{k_1k_2\cdots k_n} = \delta^{1\ 2\ \cdots\ n}_{k_1k_2\cdots k_n}. \tag{2.1.10}$$

证 显然 $k_1k_2\cdots k_n$ 也是 $j_1, j_2, \cdots, j_n$ 的排列. 又将 $k_1k_2\cdots k_n$ 变成 $j_1j_2\cdots j_n$, 再将 $j_1j_2\cdots j_n$ 变成 $12\cdots n$ 的对换过程合并, 便得到一种将 $k_1k_2\cdots k_n$ 变成 $12\cdots n$ 的对换过程. 设前两种对换过程分别经过 s 和 t 次对换, 即 $\delta^{j_1j_2\cdots j_n}_{k_1k_2\cdots k_n} = (-1)^s$, $\delta^{1\ 2\ \cdots\ n}_{j_1j_2\cdots j_n} = (-1)^t$. 那末, 合并的过程需要 $s+t$ 次对换, 即 $\delta^{1\ 2\ \cdots\ n}_{k_1k_2\cdots k_n} = (-1)^{s+t}$. □

定理 2.1.6 设 $j_1j_2\cdots j_r$ 是 $1, 2, \cdots, r$ 的排列, $j_{r+1}j_{r+2}\cdots j_n$ 是 $r+1, r+2, \cdots, n$ 的排列, 则 $j_1j_2\cdots j_rj_{r+1}\cdots j_n$ 是 $1, 2, \cdots, n$ 的排列, 且

$$\delta^{1\ 2\ \cdots\ r}_{j_1j_2\cdots j_r}\delta^{(r+1)(r+2)\cdots n}_{j_{r+1}\ \ j_{r+2}\cdots j_n} = \delta^{1\ 2\ \cdots\ n}_{j_1j_2\cdots j_n}. \tag{2.1.11}$$

证 显然, 将 $j_1j_2\cdots j_r$ 变为 $12\cdots r$, 将 $j_{r+1}j_{r+2}\cdots j_n$ 变为 $(r+1)(r+2)\cdots n$ 的对换过程合并, 便得到一种将 $j_1j_2\cdots j_n$ 变为 $12\cdots n$ 的对换过程. 所以这三种对

换过程分别作了 s, t, $s+t$ 次对换. 故定理成立. □

定理 2.1.7 设 1, 2, $\cdots$, n 的排列 $i_1i_2\cdots i_n$ 有性质: $1\leqslant i_1<\cdots<i_r\leqslant n$, $1\leqslant i_{r+1}<\cdots<i_n\leqslant n$. 则

$$\delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}=(-1)^{i_1+\cdots+i_r+r(r+1)/2}. \tag{2.1.12}$$

证 将排列 $12\cdots n$ 经过 i_1-1 次对换变成排列 $i_112\cdots(i_1-1)(i_1+1)\cdots n$, 其中 $i_2>i_1$. 所以经过 i_2-2 次对换变成

$$i_1i_212\cdots(i_1-1)(i_1+1)\cdots(i_2-1)(i_2+1)\cdots n.$$

这样依次作下去, 最后经过 i_r-r 次对换变成

$$i_1i_2\cdots i_r12\cdots(i_1-1)(i_1+1)\cdots(i_2-1)(i_2+1)\cdots(i_r-1)(i_r+1)\cdots n.$$

所以后面 $n-r$ 个正整数仍是按大小次序排好的. 因此, 它们必然为 $i_{r+1}i_{r+2}\cdots i_n$. 所以, 排列 $12\cdots n$ 经过 $(i_1-1)+(i_2-2)+\cdots+(i_r-r)$ 次对换可变成排列 $i_1i_2\cdots i_n$, 即

$$\delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n}=\delta^{i_1i_2\cdots i_n}_{1\ 2\ \cdots\ n}=(-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)}=(-1)^{i_1+\cdots+i_r+r(r+1)/2}. \qquad \square$$

由此定理可知, 若 1, 2, $\cdots$, n 的排列 $j_1j_2\cdots j_n$ 仍有 $1\leqslant j_1<\cdots<j_r\leqslant n$, $1\leqslant j_{r+1}<\cdots<j_n\leqslant n$, 则

$$\delta^{i_1i_2\cdots i_n}_{j_1j_2\cdots j_n}=(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}. \tag{2.1.13}$$

定理 2.1.8 在 $n>1$ 时, n 个正整数的所有排列中, 奇排列的总个数和偶排列的总个数都是 $n!/2$.

证 对 n 作归纳法. 在 $n=2$ 时定理显然成立. 设对任意 $n-1$ 个正整数时定理成立, 我们来证对 n 个正整数 1, 2, $\cdots$, n 时定理也成立.

由于 1, 2, $\cdots$, n 的排列的第 1 个位置可以放 1, 2, $\cdots$, n 中的任一正整数. 考虑第 1 个位置是 j, ($j\in\{1, 2, \cdots, n\}$) 的所有排列 $ji_2i_3\cdots i_n$, 这时 $i_2i_3\cdots i_n$ 自然取 1, 2, $\cdots$, $j-1$, $j+1$, $\cdots$, n 的任一排列. 由归纳法假设, 它们共有 $(n-1)!/2$ 个奇排列, $(n-1)!/2$ 个偶排列. 今若排列 $i_2i_3\cdots i_n$ 经过 s 次对换变成标准排列 $12\cdots(j-1)(j+1)\cdots n$. 于是同样的对换过程, 可使排列 $ji_2i_3\cdots i_n$ 变成排列 $j12\cdots(j-1)(j+1)\cdots n$. 再经过 $j-1$ 次对换便可变为标准排列 $12\cdots n$. 所以 $ji_2i_3\cdots i_n$ 经过 $s+j-1$ 次对换, 可变为标准排列 $12\cdots n$. 因此, 在 j 为奇数时, 由于 s 和 $s+j-1$ 的奇偶性相同, 故 $i_2i_3\cdots i_n$ 和 $ji_2i_3\cdots i_n$ 的奇偶性相同. 所以在所有形如 $ji_2i_3\cdots i_n$ 的排列中, 奇排列和偶排列的总数都是 $(n-1)!/2$. 在 j 为偶数时, 由于 s 和 $s+j-1$ 的奇偶性相反, 故 $i_2i_3\cdots i_n$ 和 $ji_2i_3\cdots i_n$ 的奇偶性相反. 所以在

所有形如 $ji_2i_3\cdots i_n$ 的排列中, 奇排列和偶排列的总数也都是 $(n-1)!/2$. 总之, 形如 $ji_2i_3\cdots i_n$ 的奇排列共有 $(n-1)!/2$ 个. 取 $j=1, 2, \cdots, n$, 这就证明了 $1, 2, \cdots, n$ 的奇排列共有 $n(n-1)!/2=n!/2$ 个. □

习 题 2.1

2.1.1 试通过一系列对换, 将排列 13852476 变为排列 72453816.

2.1.2 设 $1274j56k9$ 和 $1j25k4897$ 是 9 个正整数 1, 2, 3, 4, 5, 6, 7, 8, 9 的排列. 试决定整数 j 和 k, 使得前者是奇排列, 后者是偶排列.

2.1.3 试求下列排列的奇偶性, 即计算

(1) $\delta^{1\ 2\ \cdots(n-1)n}_{n(n-1)\cdots\ 2\ 1}$; **(2)** $\delta^{1234\cdots\ n\ (n+1)(n+2)\cdots(2n)}_{1357\cdots(2n-1)\ 2\ 4\ \cdots(2n)}$;

(3) $\delta^{123\cdots\ n(n+1)(n+2)\cdots(2n)}_{246\cdots(2n)1\ 3\cdots(2n-1)}$; **(4)** $\delta^{1234\cdots(2n-1)\ (2n)}_{2143\cdots(2n)\ (2n-1)}$;

(5) $\delta^{123\,4\cdots\ n(n+1)(n+2)(n+3)\cdots\ (2n)(2n+1)(2n+2)(2n+3)\cdots(3n)}_{369(12)\cdots(3n)\ 1\ 4\ 7\ \cdots(3n-2)\ 2\ 5\ 8\cdots(3n-1)}$.

§2.2 行列式

为了引进行列式, 且讨论行列式的性质, 下面再引进一些特殊的记号及和号.

给定域 $\mathbb{F}$ 中 n^n 个数 $a_{i_1i_2\cdots i_n}$, $1\leqslant i_1, i_2, \cdots, i_n\leqslant n$, 它们适合条件: 假设在足码 $i_1, i_2, \cdots, i_n$ 中有两个数字相同时有 $a_{i_1i_2\cdots i_n}=0$, 则我们可记

$$\sum_{i_1,i_2,\cdots,i_n=1}^{n} a_{i_1i_2\cdots i_n}=\sum_{\begin{pmatrix}1&2&\cdots&n\\ i_1&i_2&\cdots&i_n\end{pmatrix}} a_{i_1i_2\cdots i_n}, \tag{2.2.1}$$

其中第二个和式表示指标 $i_1i_2\cdots i_n$ 遍历 $1, 2, \cdots, n$ 的所有排列.

例如, 在 $n=1$ 时 $S=a_1$; 在 $n=2$ 时

$$S=\sum_{\begin{pmatrix}1&2\\ j_1&j_2\end{pmatrix}} a_{j_1j_2}=a_{12}+a_{21};$$

在 $n=3$ 时

$$S=\sum_{\begin{pmatrix}1&2&3\\ j_1&j_2&j_3\end{pmatrix}} a_{j_1j_2j_3}=a_{123}+a_{132}+a_{312}+a_{321}+a_{231}+a_{213}.$$

下面引进矩阵的概念.

定义 2.2.1 给定域 $\mathbb{F}$ 中 nm 个数 a_{jk}, 其中 $j=1, 2, \cdots, n$, $k=1, 2, \cdots, m$.

将它们排成矩形表:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}. \tag{2.2.2}$$

A 称为**矩阵**, 其中横排称为**行**, 第 j 行为

$$(a_{j1}, a_{j2}, \cdots, a_{jm}), \tag{2.2.3}$$

又竖排称为**列**, 第 k 列为

$$\begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}. \tag{2.2.4}$$

所以有 n 个行, m 个列, 称为 $n\times m$ 矩阵. 为了表明矩阵 A 的行数 n 和列数 m, 因此矩阵 A 可记作 $A_{n,m}$, 也可记作 $A^{(n,m)}$. 矩阵 A 的第 j 行, 第 k 列交叉位置的元素为 a_{jk}, 称为矩阵 A 的第 j 行, 第 k 列元素. 所以元素 a_{jk} 的前指标表示行数, 后指标表示列数. $n\times m$ 矩阵 A 一共有 n 个行 (横排的) 和 m 个列 (竖排的). 每一个元素 a_{jk} 有两个足标, 前足标表示它所在的行数, 后足标表示它所在的列数, 所以 a_{jk} 是第 j 行, 第 k 列交叉位置的元素. 因此, 矩阵 A 也可记作 $A=(a_{jk})$.

$n\times m$ 矩阵 A 的**转置矩阵**是 $m\times n$ 矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{nm} \end{pmatrix}, \tag{2.2.5}$$

记作 A', 或 A^T, 或 ${}^TA$. 它是将 A 的第 j 行变成第 j 列, 第 k 列变成第 k 行后所得到的矩阵.

$n\times m$ 复矩阵 A 的**共轭矩阵**是 $n\times m$ 矩阵

$$\begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{n1}} \\ \vdots & & \vdots \\ \overline{a_{1m}} & \cdots & \overline{a_{nm}} \end{pmatrix}, \tag{2.2.6}$$

记作 $\overline{A}$. 它是将 A 的第 j 行, 第 k 列交叉位置的元素 a_{jk} 变成它的共轭元素 $\overline{a_{jk}}$ 后所得到的矩阵.

由定义可知

$$(\overline{A})'=\overline{(A')}=\overline{A}'=A^*. \tag{2.2.7}$$

又, 复矩阵 A 是实矩阵当且仅当 $\overline{A}=A$. 因此, 这时 $A^*=A'$.

当 $n=m$ 时, 矩阵 A 也称为 **n 阶方阵**. n 阶方阵 A 也可记作 A_n. 方阵 A 中行列相同的元素 a_{ii} 称为方阵 A 的**对角元素**. n 阶方阵 A 有 n 个对角元素, 它们依次为 $a_{11}, a_{22}, \cdots, a_{nn}$. a_{ii} 称为方阵 A 的 **第 i 个对角元素**, $1\leqslant i\leqslant n$.

定义 2.2.2 给定域 $\mathbb{F}$ 上 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}. \tag{2.2.8}$$

取定 $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$, $1\leqslant j_1<j_2<\cdots<j_s\leqslant m$, 在 A 中取出第 i_1 行, $\cdots$, 第 i_r 行, 第 j_1 列, $\cdots$, 第 j_s 列的所有交叉元素, 按原来的次序, 便构成一个 $r\times s$ 矩阵, 称为矩阵 A 的**子矩阵**, 记作

$$A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_s \end{pmatrix}=\begin{pmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_s} \\ a_{i_2j_1} & a_{i_2j_2} & \cdots & a_{i_2j_s} \\ \vdots & \vdots & & \vdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_s} \end{pmatrix}. \tag{2.2.9}$$

关于矩阵的代数运算是第三章的内容. 在这一章, 我们只用它的符号. 重要的是引进 n 阶方阵的函数 —— 行列式, 它是自变量为 n 阶方阵的函数.

定义 2.2.3 域 $\mathbb{F}$ 上 n 阶方阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{2.2.10}$$

的行列式是域 $\mathbb{F}$ 中一数, 记作 $\det(A)$, 它定义为

$$\det(A)=\sum_{\begin{pmatrix} 1 & 2\cdots n \\ i_1 & i_2\cdots i_n \end{pmatrix}} \delta^{1\ 2\ \cdots\ n}_{i_1i_2\cdots i_n} a_{1i_1}a_{2i_2}\cdots a_{ni_n}. \tag{2.2.11}$$

行列式有下列几种常用的符号:

$$\det A=\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=|A|. \tag{2.2.12}$$

行列式是本书的主要工具之一, 而且是近代数学中不可缺少的代数工具之一. 本章将系统地给出行列式的基本理论.

由行列式的定义可知, **(1)** 它的展开式是 $n!$ 项的和; **(2)** 每一项的主要部分是方阵中 n 个元素 $a_{1i_1}, a_{2i_2}, \cdots, a_{ni_n}$ 的乘积. 这 n 个元素分别落在第 $1, 2, \cdots, n$ 行中, 又分别落在第 $i_1, i_2, \cdots, i_n$ 列中; **(3)** 由于 $i_1 i_2 \cdots i_n$ 是 $1, 2, \cdots, n$ 的排列, 所以这 n 个元素 $a_{1i_1}, a_{2i_2}, \cdots, a_{ni_n}$ 在方阵 A 的 n 个行中及 n 个列中分别出现一次, 且只出现一次; **(4)** 这 n 个元素的乘积的系数是 1 或 -1, 当排列 $i_1 i_2 \cdots i_n$ 是偶排列时, 系数是 1; 当排列 $i_1 i_2 \cdots i_n$ 是奇排列时, 系数是 -1. 由定理 2.1.8, 有 $n!/2$ 个项的系数为 1, 另外 $n!/2$ 个项的系数为 -1; **(5)** 当 $n=1$ 时, $\det(A) = \det(a_{11}) = a_{11}$.

由行列式的定义还可知, 利用定义直接来计算行列式是很困难的. 只有在阶数低时, 才可以直接用定义计算. 为了能够进行计算, 需要先导出行列式的若干基本性质, 再通过这些性质, 将复杂的行列式的计算化为简单的行列式的计算, 也可以将高阶行列式的计算化为低阶行列式的计算.

定理 2.2.4 设 A 是域 $\mathbb{F}$ 上 n 阶方阵, 则

$$\det(A') = \det(A). \tag{2.2.13}$$

证 记域 $\mathbb{F}$ 上 n 阶方阵 A 和它的转置方阵 A' 分别为

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}.$$

所以记 $b_{jk} = a_{kj}$, $1 \leqslant j, k \leqslant n$. 由定义可知:

$$\det(A') = \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ i_1 & i_2 \cdots i_n \end{pmatrix}} \delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} b_{1i_1} b_{2i_2} \cdots b_{ni_n} = \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ i_1 & i_2 \cdots i_n \end{pmatrix}} \delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}.$$

现在考虑每一项 $a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$. 由于 $i_1 i_2 \cdots i_n$ 是 $1, 2, \cdots, n$ 的排列, 所以在 $i_1, i_2, \cdots, i_n$ 中分别有数 $i_{j_1} = 1, i_{j_2} = 2, \cdots, i_{j_n} = n$. 自然, $j_1 j_2 \cdots j_n$ 也是 $1, 2, \cdots, n$ 的排列. 又,

$$a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} = a_{i_{j_1} j_1} a_{i_{j_2} j_2} \cdots a_{i_{j_n} j_n} = a_{1 j_1} a_{2 j_2} \cdots a_{n j_n},$$

且对任一将排列 $i_1 i_2 \cdots i_n$ 变为标准排列 $12 \cdots n$ 的对换过程, 作用在标准排列上, 变为排列 $j_1 j_2 \cdots j_n$. 而且记 S_n 为 $1, 2, \cdots, n$ 的所有排列构成的集合, 则上面这个等式给出了 S_n 到自身上的一个一一对应 $i_1 i_2 \cdots i_n \to j_1 j_2 \cdots j_n$. 特别, 当排列 $i_1 i_2 \cdots i_n$ 遍历所有排列时, 排列 $j_1 j_2 \cdots j_n$ 也遍历所有排列. 于是 $\delta^{i_1 i_2 \cdots i_n}_{1\,2\,\cdots\,n} = \delta^{1\,2\,\cdots\,n}_{j_1 j_2 \cdots j_n}$, 又

$$\det(A') = \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ j_1 & j_2 \cdots j_n \end{pmatrix}} \delta^{1\,2\,\cdots\,n}_{j_1 j_2 \cdots j_n} a_{1 j_1} a_{2 j_2} \cdots a_{n j_n} = \det(A).$$

这就证明了定理. □

这个定理告诉我们, 行列式的行和列的地位是平等的. 确切地说, 每一个关于行的性质, 对列必然成立; 反之亦然. 所以今后只需要导出关于行的一系列性质, 由这个定理就可以知道这一系列性质对列也成立.

定理 2.2.5 行列式的任两行 (或两列) 互换, 则行列式变号.

证 任取 $1 \leqslant j < k \leqslant n$, 将行列式 $\det(A)$ 的第 j 行和第 k 行互换, 于是得到行列式

$$\det(B) = \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$= \sum_{\left(\begin{smallmatrix}1 & 2 \cdots n \\ i_1 i_2 \cdots i_n\end{smallmatrix}\right)} \delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} \cdot a_{1i_1} \cdots a_{j-1,i_{j-1}} a_{ki_j} a_{j+1,i_{j+1}} \cdots a_{k-1 i_{k-1}} a_{ji_k} a_{k+1,i_{k+1}} \cdots a_{ni_n}$$

$$= \sum_{\left(\begin{smallmatrix}1 & 2 \cdots n \\ i_1 i_2 \cdots i_n\end{smallmatrix}\right)} \delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} \cdot a_{1i_1} \cdots a_{j-1,i_{j-1}} a_{ji_k} a_{j+1,i_{j+1}} \cdots a_{k-1 i_{k-1}} a_{ki_j} a_{k+1,i_{k+1}} \cdots a_{ni_n}.$$

将 $1, 2, \cdots, n$ 的排列 $i_1 i_2 \cdots i_n$ 经过 $2(k-j+1)+1$ 次对换, 可变为排列

$$i_1 \cdots i_{j-1} i_k i_{j+1} \cdots i_{k-1} i_j i_{k+1} \cdots i_n,$$

所以

$$-\delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} = \delta^{1\,\cdots(j-1)j(j+1)\cdots(k-1)k(k+1)\cdots n}_{i_1 \cdots\, i_{j-1} i_k i_{j+1}\, \cdots\, i_{k-1}\, i_j\, i_{k+1} \cdots i_n}.$$

显然置换 $\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ 到置换

$$\begin{pmatrix} 1 & \cdots & j-1 & j & j+1 & \cdots & k-1 & k & k+1 & \cdots & n \\ i_1 & \cdots & i_{j-1} & i_k & i_{j+1} & \cdots & i_{k-1} & i_j & i_{k+1} & \cdots & i_n \end{pmatrix}$$

上的对应是一一对应, 又当 $i_1 i_2 \cdots i_n$ 遍历 $1, 2, \cdots, n$ 的所有排列时, $i_1 \cdots\ i_{j-1} i_k i_{j+1} \cdots\ i_{k-1} i_j i_{k+1} \cdots i_n$ 也遍历 $1, 2, \cdots, n$ 的所有排列时. 所以行列式 $\det(B)$ 等于

$$\sum_{\left(\begin{smallmatrix}1 & 2 \cdots n \\ i_1 i_2 \cdots i_n\end{smallmatrix}\right)} \delta^{1\,2\,\cdots\,n}_{i_1 i_2 \cdots i_n} \cdot a_{1i_1} \cdots a_{j-1,i_{j-1}} a_{ji_k} a_{j+1,i_{j+1}} \cdots a_{k-1,i_{k-1}} a_{ki_j} a_{k+1,i_{k+1}} \cdots a_{ni_n}$$

$$= - \sum_{\begin{pmatrix} 1,\cdots,j-1,j,j+1,\cdots,k-1,k,k+1,\cdots,n \\ i_1,\cdots,i_{j-1},i_k,i_{j+1},\cdots,i_{k-1},i_j,i_{k+1},\cdots,i_n \end{pmatrix}} \delta^{1,\cdots,(j-1),j,(j+1),\cdots,(k-1),k,(k+1),\cdots,n}_{i_1,\cdots,i_{j-1},i_k,i_{j+1},\cdots,i_{k-1},i_j,i_{k+1},\cdots,i_n}$$

$$\cdot a_{1i_1}\cdots a_{j-1,i_{j-1}} a_{ji_k} a_{j+1,i_{j+1}} \cdots a_{k-1 i_{k-1}} a_{ki_j} a_{k+1,i_{k+1}} \cdots a_{ni_n}.$$

改记求和符号中的 i_j 为 i_k, i_k 为 i_j, 所以有

$$\det(B) = - \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ i_1 i_2 \cdots i_n \end{pmatrix}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2 \cdots i_n} a_{1i_1} \cdots a_{ni_n} = -\det(A). \qquad \square$$

定理 2.2.6 行列式的某两行 (或两列) 相同时, 行列式的值为零.

证 设行列式 $\det(A)$ 中方阵 A 的第 i 行和第 j 行相同. 将这两行互换后, 显然, 行列式的值不变. 但是由定理 2.2.5, 则行列式的值变号, 即有 $\det(A) = -\det(A)$, 所以 $\det(A) = 0$. $\square$

定理 2.2.7 用数 $\lambda \in \mathbb{F}$ 乘以任一行 (或列) 后, 所得到的行列式等于原来的行列式乘以 λ.

证 今

$$\det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{j1} & \lambda a_{j2} & \cdots & \lambda a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ i_1 i_2 \cdots i_n \end{pmatrix}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2 \cdots i_n} a_{1i_1} \cdots (\lambda a_{ji_j}) \cdots a_{ni_n}$$

$$= \lambda \sum_{\begin{pmatrix} 1 & 2 \cdots n \\ i_1 i_2 \cdots i_n \end{pmatrix}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2 \cdots i_n} a_{1i_1} \cdots a_{ji_j} \cdots a_{ni_n} = \lambda \det(A). \qquad \square$$

定理 2.2.8 行列式有两行 (或两列) 成比例, 则该行列式的值为零. 特别, 当行列式有一行 (或列) 全为零时, 行列式的值为零.

证 由定理 2.2.7 及定理 2.2.6 便能证明定理成立. 证完. $\square$

定理 2.2.9

$$\Delta = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ a_{j1} + b_{j1} & a_{j2} + b_{j2} & \cdots & a_{jn} + b_{jn} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$= \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} + \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j-1,1} & a_{j-1,2} & \cdots & a_{j-1,n} \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ a_{j+1,1} & a_{j+1,2} & \cdots & a_{j+1,n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{2.2.14}$$

对列也一样成立.

证 由行列式的定义可知,

$$\begin{aligned} \Delta = & \sum_{\binom{1\ 2\cdots n}{i_1 i_2\cdots i_n}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{j-1,i_{j-1}}(a_{ji_j}+b_{ji_j})a_{j+1,i_{j+1}}\cdots a_{ni_n} \\ = & \sum_{\binom{1\ 2\cdots n}{i_1 i_2\cdots i_n}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{j-1,i_{j-1}}a_{ji_j}a_{j+1,i_{j+1}}\cdots a_{ni_n} \\ & + \sum_{\binom{1\ 2\cdots n}{i_1 i_2\cdots i_n}} \delta^{1\ 2\ \cdots\ n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{j-1,i_{j-1}}b_{ji_j}a_{j+1,i_{j+1}}\cdots a_{ni_n}. \end{aligned}$$

由行列式的定义, 便证明了定理. □

定理 2.2.10 将行列式的任一行 (或任一列) 乘以实数, 再相应地加到另一行 (或另一列) 上去, 则行列式的值不变. 例如, 当 $j\neq 1$ 时,

$$\det\begin{pmatrix} a_{11}+\lambda a_{j1} & a_{12}+\lambda a_{j2} & \cdots & a_{1n}+\lambda a_{jn} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{2.2.15}$$

证 利用定理 2.2.9 及定理 2.2.8 便得证. □

下面给出一种方法来计算行列式. 首先, 不难证明:

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11}\det\begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{2.2.16}$$

事实上, 由行列式的定义可知

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \sum_{\binom{1\ 2\cdots n}{i_1 i_2\cdots i_n}} \delta^{1\,2\,\cdots\,n}_{i_1 i_2\cdots i_n} a_{1i_1}\cdots a_{ji_j}\cdots a_{ni_n},$$

其中 $a_{1j}=0,\ j>1$. 所以在 $i_1\neq 1$ 时, 相应的项为零, 即上式为

$$\det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11} \sum_{\binom{2\ 3\cdots n}{i_2 i_3\cdots i_n}} \delta^{2\,\cdots\,n}_{i_2\cdots i_n} a_{2i_2}\cdots a_{ji_j}\cdots a_{ni_n}.$$

因此断言成立. □

今任取行列式

$$\det(B)=\det\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}. \tag{2.2.17}$$

假设 $b_{ij}\neq 0$. 将第 i 行遍乘 $-b_{kj}b_{ij}^{-1}$ 加到第 k 行上去 (这里 $k\neq i$), 于是得到一个新的行列式, 其值仍为 $\det(B)$. 但是它的第 k 行, 第 j 列元素为零. 上面过程对 $k=1,$ $2,\cdots,i-1,i+1,\cdots,n$ 分别进行, 这样就把 $\det(B)$ 的计算化为一个新的行列式的计算, 它的第 j 列元素除了 b_{ij} 外都为零, 且第 i 行仍为 $b_{i1},\cdots,b_{in}$, 即

$$\det(B)=\det\begin{pmatrix} c_{11} & \cdots & c_{1,j-1} & 0 & c_{1,j+1} & \cdots & c_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{i-1,1} & \cdots & c_{i-1,j-1} & 0 & c_{i-1,j+1} & \cdots & c_{i-1,n} \\ b_{i1} & \cdots & b_{i,j-1} & b_{ij} & b_{i,j+1} & \cdots & b_{in} \\ c_{i+1,1} & \cdots & c_{i+1,j-1} & 0 & c_{i+1,j+1} & \cdots & c_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_{n1} & \cdots & c_{n,j-1} & 0 & c_{n,j+1} & \cdots & c_{nn} \end{pmatrix}.$$

将第 i 行依次和第 $i-1,\cdots,1$ 行作对换, 再将第 j 列依次和第 $j-1,\cdots,1$ 列作对

换, 于是

$$\det(B)=(-1)^{i+j}b_{ij}\det\begin{pmatrix} c_{11} & \cdots & c_{1,j-1} & c_{1,j+1} & \cdots & c_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{i-1,1} & \cdots & c_{i-1,j-1} & c_{i-1,j+1} & \cdots & c_{i-1,n} \\ c_{i+1,1} & \cdots & c_{i+1,j-1} & c_{i+1,j+1} & \cdots & c_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{n1} & \cdots & c_{n,j-1} & c_{n,j+1} & \cdots & c_{nn} \end{pmatrix}. \tag{2.2.18}$$

这样依次作下去, 总可以把行列式的值算出来. 但是, 这个方法并不是最好的方法, 读者应该根据具体的行列式, 灵活运用上面六个基本性质进行计算. 下一节还要介绍一种重要的性质, 即所谓 Laplace 展开, 这对行列式的计算也起很大作用.

习 题 2.2

2.2.1 任取 $1\leqslant j,k\leqslant n$, 引进 **Kronecker 符号**, 它定义为

$$\delta_{jk}=\begin{cases}1, & j=k;\\ 0, & j\neq k.\end{cases} \tag{2.2.19}$$

试证: $\displaystyle\sum_{j,k=1}^{n}\delta_{jk}=n,\quad \sum_{j=1}^{n}a_j\delta_{jk}=a_k,\quad \sum_{j=1}^{n}\delta_{jk}\delta_{ji}=\delta_{ik},\quad \sum_{j,k=1}^{n}a_{jk}\delta_{jk}=\sum_{j=1}^{n}a_{jj}.$

2.2.2 试证:

$$\det\begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}=\det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}=a_{11}a_{22}\cdots a_{nn}. \tag{2.2.20}$$

2.2.3 不必展开行列式, 试证明下列等式:

(1) $\det\begin{pmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ca \\ 1 & c & c^2-ab \end{pmatrix}=0;$

(2) $\det\begin{pmatrix} b+c & c+a & a+b \\ b_1+c_1 & c_1+a_1 & a_1+b_1 \\ b_2+c_2 & c_2+a_2 & a_2+b_2 \end{pmatrix}=2\det\begin{pmatrix} a & b & c \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$

2.2.4 试计算下列行列式:

(1) $\det\begin{pmatrix} 246 & 427 & 327 \\ 1014 & 543 & 443 \\ -342 & 721 & 666 \end{pmatrix}$; (2) $\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix}$;

(3) $\det\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{pmatrix}$; (4) $\det\begin{pmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1+y & 1 & 1 \\ 1 & 1 & 1+z & 1 \\ 1 & 1 & 1 & 1+w \end{pmatrix}$;

(5) $\det\begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}$; (6) $\det\begin{pmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{pmatrix}$;

(7) $\det\begin{pmatrix} \cos\varphi & \sin\varphi & \cos\varphi & \sin\varphi \\ \cos 2\varphi & \sin 2\varphi & 2\cos 2\varphi & 2\sin 2\varphi \\ \cos 3\varphi & \sin 3\varphi & 3\cos 3\varphi & 3\sin 3\varphi \\ \cos 4\varphi & \sin 4\varphi & 4\cos 4\varphi & 4\sin 4\varphi \end{pmatrix}$.

2.2.5 证明下列等式:

(1)

$$\det\begin{pmatrix} 1 & 1 & 1 \\ \tan A & \tan B & \tan C \\ \tan 2A & \tan 2B & \tan 2C \end{pmatrix} = 0, \quad 其中 \quad A+B+C=2\pi;$$

(2)

$$\det\begin{pmatrix} \cos(a-b) & \cos(b-c) & \cos(c-a) \\ \cos(a+b) & \cos(b+c) & \cos(c+a) \\ \sin(a+b) & \sin(b+c) & \sin(c+a) \end{pmatrix} = -2\sin(a-b)\sin(b-c)\sin(c-a);$$

(3)

$$\det\begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & -x_4 & x_3 \\ x_3 & x_4 & x_1 & -x_2 \\ x_4 & -x_3 & x_2 & x_1 \end{pmatrix} = (x_1^2+x_2^2+x_3^2+x_4^2)^2.$$

2.2.6 在域 F 中给定 n^2 个数, 将它们任意排成一个 n 阶方阵. 试证: 它们的行列式至多有 $(n^2)!/(n!)^2$ 个不同的数.

2.2.7 试证: 任一 n 阶方阵 (a_{ij}) 有 $\det(a_{ij}) = \det((-1)^{i+j}a_{ij})$.

§2.3 代数余子式及 Laplace 展开式

给定 n 阶方阵

$$A = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{2.3.1}$$

任取 $2p$ 个正整数 $1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_p \leqslant n$. 在方阵 A 中取出第 $i_1, i_2, \cdots, i_p$ 行, 按次序排成

$$\begin{matrix} a_{i_1 1} & \cdots & a_{i_1 n} \\ \vdots & & \vdots \\ a_{i_p 1} & \cdots & a_{i_p n}. \end{matrix}$$

再在其中取出第 $j_1, j_2, \cdots, j_p$ 列元素, 按次序排成 p 阶子方阵

$$A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} = \begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_p} \\ \vdots & & \vdots \\ a_{i_p j_1} & \cdots & a_{i_p j_p} \end{pmatrix}. \tag{2.3.2}$$

显然, 它是由方阵 A 的第 $i_1, i_2, \cdots, i_p$ 行, 第 $j_1, j_2, \cdots, j_p$ 列的交叉元素, 按原来的次序排成的 p 阶方阵. 这个方阵的行列式

$$\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} = \det \begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_p} \\ \vdots & & \vdots \\ a_{i_p j_1} & \cdots & a_{i_p j_p} \end{pmatrix} \tag{2.3.3}$$

称为方阵 A 的 p 阶子式. 注意, 符号

$$\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}$$

的上指标是行的指标, 下指标是列的指标.

设 A 是 n 阶方阵, 将行列式 $\det(A)$ 的第 i_1 行经过相邻两行互换的方法换到第 1 行, 要作 $i_1 - 1$ 次对换; 再将行列式 $\det(A)$ 的第 i_2 行经过相邻两行互换的方法换到第 2 行, 要作 $i_2 - 2$ 次对换; 这样依次作下去, 最后, 将行列式 $\det(A)$ 的第 i_p 行

经过相邻两行互换的方法换到第 p 行, 要作 i_p-p 次对换, 于是有

$$\det(A)=(-1)^{(i_1-1)+(i_2-2)+\cdots+(i_p-p)}\det\begin{pmatrix} a_{i_11} & a_{i_12} & \cdots & a_{i_1n} \\ \vdots & \vdots & & \vdots \\ a_{i_p1} & a_{i_p2} & \cdots & a_{i_pn} \\ a_{i_{p+1}1} & a_{i_{p+1}2} & \cdots & a_{i_{p+1}n} \\ \vdots & \vdots & & \vdots \\ a_{i_n1} & a_{i_n2} & \cdots & a_{i_nn} \end{pmatrix},$$

其中 $1\leqslant i_{p+1}<i_{p+2}<\cdots<i_n\leqslant n$, 且 $i_{p+1}, i_{p+2}, \cdots, i_n$ 是由 $1, 2, \cdots, n$ 中除去正整数 $i_1, i_2, \cdots, i_p$ 而得, 所以它由排列 $i_1\cdots i_p$ 唯一确定, 且 $i_1\cdots i_pi_{p+1}\cdots i_n$ 是 $1, 2, \cdots, n$ 的一个排列. 由定理 2.1.7, $\delta_{i_1i_2\cdots i_n}^{1\,2\,\cdots\,n}=(-1)^{(i_1-1)+\cdots+(i_p-p)}$.

同样, 经 $(j_1-1)+(j_2-2)+\cdots+(j_p-p)$ 次相邻两列的互换, 可将行列式 $\det(A)$ 的第 $j_1, j_2, \cdots, j_p$ 列依次变为第 $1, 2, \cdots, p$ 列, 而由排列 $j_1j_2\cdots j_p$ 唯一决定了 $1, 2, \cdots, n$ 的一个排列 $j_1j_2\cdots j_pj_{p+1}\cdots j_n$, 其中 $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$. 所以 $\delta_{j_1j_2\cdots j_n}^{1\,2\,\cdots\,n}=(-1)^{(j_1-1)+\cdots+(j_p-p)}$,

$$\det(A)=\delta_{j_1j_2\cdots j_n}^{i_1i_2\cdots i_n}\det\begin{pmatrix} a_{i_1j_1} & \cdots & a_{i_1j_p} & a_{i_1j_{p+1}} & \cdots & a_{i_1j_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i_pj_1} & \cdots & a_{i_pj_p} & a_{i_pj_{p+1}} & \cdots & a_{i_pj_n} \\ a_{i_{p+1}j_1} & \cdots & a_{i_{p+1}j_p} & a_{i_{p+1}j_{p+1}} & \cdots & a_{i_{p+1}j_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i_nj_1} & \cdots & a_{i_nj_p} & a_{i_nj_{p+1}} & \cdots & a_{i_nj_n} \end{pmatrix}.$$

新的行列式的前 p 行, 前 p 列决定的行列式为 $\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}$, 后 $n-p$ 行, 后 $n-p$ 列决定的行列式为 $\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}$.

定义 2.3.1 给定指标 $1\leqslant i_1<i_2<\cdots<i_p\leqslant n$, $1\leqslant j_1<j_2<\cdots<j_p\leqslant n$. 它们唯一地决定了 $1, 2, \cdots, n$ 的排列 $i_1i_2\cdots i_n$ 和 $j_1j_2\cdots j_n$, 其中 $1\leqslant i_{p+1}<\cdots<i_n\leqslant n$, $1\leqslant j_{p+1}<\cdots<j_n\leqslant n$. 这时

$$\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}, \qquad \delta_{j_1j_2\cdots j_p}^{i_1i_2\cdots i_p}\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix} \tag{2.3.4}$$

分别称为方阵 A 的 p 阶**子式** 和 p 阶**代数子式**;

$$\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}, \qquad \delta_{j_1j_2\cdots j_p}^{i_1i_2\cdots i_p}\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix} \tag{2.3.5}$$

分别称为方阵 A 的子式

$$\det A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix} \tag{2.3.6}$$

的**余子式**和**代数余子式**.

由于

$$\delta_{i_1 i_2 \cdots i_p}^{1\ 2\ \cdots\ p} = (-1)^{(i_1-1)+\cdots+(i_p-p)}, \quad \delta_{j_1 j_2 \cdots j_p}^{1\ 2\ \cdots\ p} = (-1)^{(j_1-1)+\cdots+(j_p-p)},$$

所以

$$\delta_{j_1 j_2 \cdots j_p}^{i_1 i_2 \cdots i_p} = \delta_{i_1 i_2 \cdots i_p}^{1\ 2\ \cdots\ p}\delta_{j_1 j_2 \cdots j_p}^{1\ 2\ \cdots\ p} = (-1)^{i_1+\cdots+i_p+j_1+\cdots+j_p}.$$

又由于

$$(i_1+\cdots+i_p)+(i_{p+1}+\cdots+i_n) = (j_1+\cdots+j_p)+(j_{p+1}+\cdots+j_n) = 1+\cdots+n,$$

所以

$$(-1)^{i_{p+1}+\cdots+i_n+j_{p+1}+\cdots+j_n} = (-1)^{i_1+\cdots+i_p+j_1+\cdots+j_p}.$$

因此 $\det A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix}$ 的代数子式和代数余子式可以分别改写为

$$\delta_{j_1 j_2 \cdots j_p}^{i_1 i_2 \cdots i_p}\det A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix} = (-1)^{i_1+\cdots+i_p+j_1+\cdots+j_p}\det A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix} \tag{2.3.7}$$

和

$$\delta_{j_{p+1} j_{p+2} \cdots j_n}^{i_{p+1} i_{p+2} \cdots i_n}\det A\begin{pmatrix} i_{p+1} \cdots i_n \\ j_{p+1} \cdots j_n \end{pmatrix}. \tag{2.3.8}$$

所以方阵 A 的子式的代数余子式等于子式的余子式的代数子式.

由子式及余子式的定义可知, 子式 $\det A\begin{pmatrix} i_1 \cdots i_p \\ j_1 \cdots j_p \end{pmatrix}$ 是由方阵 A 中第 i_1, i_2, $\cdots$, i_p 行, 第 j_1, j_2, $\cdots$, j_p 列的交叉元素按照原来的次序构成的 p 阶行列式.

余子式 $\det A\begin{pmatrix} i_{p+1} \cdots i_n \\ j_{p+1} \cdots j_n \end{pmatrix}$ 是由方阵 A 除去第 i_1, i_2, $\cdots$, i_p 行以及第 j_1, j_2, $\cdots$, j_p 列的所有元素后, 余下的元素, 按照原来的次序构成的 $n-p$ 阶行列式.

由此可见, 子式及其余子式是由方阵 A 中确定位置的元素构成, 因此子式作为行列式的值和余子式作为行列式的值是两个没有关系的实数. 例如两个方阵

$$A = \det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad B = \det\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

的元素除了 a_{1j} 和 b_{1j}, $j=1,2,\cdots,n$ 外完全相同. 容易看出, 这时子式 $\det A\begin{pmatrix}1\\j\end{pmatrix}=a_{1j}$ 和 $\det B\begin{pmatrix}1\\j\end{pmatrix}=b_{1j}$ 不同, 但是, 它们的余子式和代数余子式都相同.

由行列式的定义可知, 子式

$$\det A\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}=\det A\begin{pmatrix}a_{i_1j_1}&\cdots&a_{i_1j_p}\\\vdots&&\vdots\\a_{i_pj_1}&\cdots&a_{i_pj_p}\end{pmatrix}$$

有

$$\det A\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}=\sum_{\begin{pmatrix}1&2\cdots&p\\k_1&k_2\cdots&k_p\end{pmatrix}}\delta^{1\,2\,\cdots\,p}_{k_1k_2\cdots k_p}a_{i_1j_{k_1}}a_{i_2j_{k_2}}\cdots a_{i_pj_{k_p}}.\tag{2.3.9}$$

显然, 对于 $1,2,\cdots,p$ 的任一排列 $k_1k_2\cdots k_p$, 则 $j_{k_1}j_{k_2}\cdots j_{k_p}$ 是 $j_1,j_2,\cdots,j_p$ 的一个排列. 又

$$\delta^{1\,2\,\cdots\,p}_{k_1k_2\cdots k_p}=\delta^{j_1\,j_2\,\cdots\,j_p}_{j_{k_1}j_{k_2}\cdots j_{k_p}},$$

且当 $k_1k_2\cdots k_p$ 遍历 $1,2,\cdots,p$ 的所有排列时, $j_{k_1}j_{k_2}\cdots j_{k_p}$ 遍历 $j_1j_2\cdots j_p$ 的所有排列. 这证明了

$$\det A\begin{pmatrix}i_1&\cdots&i_p\\j_1&\cdots&j_p\end{pmatrix}=\sum_{\begin{pmatrix}j_1j_2\cdots j_p\\k_1k_2\cdots k_p\end{pmatrix}}\delta^{j_1j_2\cdots j_p}_{k_1k_2\cdots k_p}a_{i_1k_1}a_{i_2k_2}\cdots a_{i_pk_p}.\tag{2.3.10}$$

引理 2.3.2　给定正整数 n 和 p, $p<n$. 给定域 $\mathbb{F}$ 中 $n!$ 个数 $a_{i_1i_2\cdots i_n}$, 其中指标 $i_1i_2\cdots i_n$ 遍历 $1,2,\cdots,n$ 的所有排列. 在 n 个正整数 $1,2,\cdots,n$ 中任取 p 个正整数, 将它们按大小次序排起来, 记作 $j_1<j_2<\cdots<j_p$. 在 $1,2,\cdots,n$ 中再除去 j_1, $j_2,\cdots,j_p$, 余下的正整数再按大小次序排起来, 记作 $j_{p+1}<\cdots<j_n$. 则有

$$\sum_{\begin{pmatrix}1&2\cdots&n\\i_1&i_2\cdots&i_n\end{pmatrix}}a_{i_1i_2\cdots i_n}=\sum_{1\leqslant j_1<\cdots<j_p\leqslant n}\ \sum_{\begin{pmatrix}j_1j_2\cdots j_p\\i_1i_2\cdots i_p\end{pmatrix}}\ \sum_{\begin{pmatrix}j_{p+1}j_{p+2}\cdots j_n\\i_{p+1}i_{p+2}\cdots i_n\end{pmatrix}}a_{i_1i_2\cdots i_n}.\tag{2.3.11}$$

证　今任取排列 $i_1i_2\cdots i_n$, 它的前 p 个数 $i_1,\cdots,i_p$ 按大小次序排起来, 记作 $j_1<j_2<\cdots<j_p$. 它的后 $n-p$ 个数 $i_{p+1},\cdots,i_n$ 再按大小次序排起来, 记作 $j_{p+1}<j_{p+2}<\cdots<j_n$. 所以排列 $i_1i_2\cdots i_n$ 是排列 $j_1j_2\cdots j_n$ 的排列. 由三重和号的定义, 排列 $i_1i_2\cdots i_n$ 给出了三重和号中一项 $a_{i_1i_2\cdots i_n}$. 又三重和号中的不同项 $a_{i_1i_2\cdots i_n}$ 和 $a_{i'_1i'_2\cdots i'_n}$ 决定了两个不同的排列 $i_1i_2\cdots i_n$ 和 $i'_1i'_2\cdots i'_n$. 最后, 由于三重和号的第一个和号是 $\binom{n}{p}$ 个项的和, 第二个和号是 $p!$ 个项的和, 第三个和号是 $(n-p)!$

个项的和, 所以总共是 $p!(n-p)!\binom{n}{p}=n!$ 个项的和, 而 $1, 2, \cdots, n$ 的排列一共有 $n!$ 个. 这就证明了式 (2.3.11) 成立. □

定理 2.3.3 (Laplace 定理) 任取指标 $1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$, 则

$$\det(A) = \sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n} \det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}\left[\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}\right], \tag{2.3.12}$$

其中 $i_1 i_2\cdots i_n$ 和 $j_1 j_2\cdots j_n$ 是 $1, 2, \cdots, n$ 的排列, 且

$$1 \leqslant i_{p+1} < \cdots < i_n \leqslant n, \quad 1 \leqslant j_{p+1} < \cdots < j_n \leqslant n. \tag{2.3.13}$$

证 由引理 2.3.2 及行列式的基本性质, 可知

$$\det(A) = (-1)^{i_1+\cdots+i_p-1-\cdots-p}\det\begin{pmatrix} a_{i_1 1} & \cdots & a_{i_1 n} \\ \vdots & & \vdots \\ a_{i_n 1} & \cdots & a_{i_n n} \end{pmatrix}$$

$$= \delta^{1\,2\,\cdots\,n}_{i_1 i_2\cdots i_n}\sum_{\begin{pmatrix} 1 & 2\cdots & n \\ k_1 & k_2\cdots & k_n \end{pmatrix}} \delta^{1\,2\,\cdots\,n}_{k_1 k_2\cdots k_n} a_{i_1 k_1}a_{i_2 k_2}\cdots a_{i_n k_n}$$

$$= \sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n}\ \sum_{\begin{pmatrix} j_1 j_2\cdots j_p \\ k_1 k_2\cdots k_p \end{pmatrix}}\ \sum_{\begin{pmatrix} j_{p+1} j_{p+2}\cdots j_n \\ k_{p+1} k_{p+2}\cdots k_n \end{pmatrix}} \delta^{i_1 i_2\cdots i_n}_{k_1 k_2\cdots k_n} a_{i_1 k_1}a_{i_2 k_2}\cdots a_{i_n k_n}.$$

所以由

$$\delta^{i_1 i_2\cdots i_n}_{k_1 k_2\cdots k_n} = \delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\delta^{j_1 j_2\cdots j_n}_{k_1 k_2\cdots k_n} = \delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\delta^{j_1 j_2\cdots j_p}_{k_1 k_2\cdots k_p}\delta^{j_{p+1} j_{p+2}\cdots j_n}_{k_{p+1} k_{p+2}\cdots k_n},$$

有

$$\det(A) = \sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n}\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\left(\sum_{\begin{pmatrix} j_1 j_2\cdots j_p \\ k_1 k_2\cdots k_p \end{pmatrix}}\delta^{j_1 j_2\cdots j_p}_{k_1 k_2\cdots k_p} a_{i_1 k_1}a_{i_2 k_2}\cdots a_{i_p k_p}\right)$$

$$\cdot\left(\sum_{\begin{pmatrix} j_{p+1} j_{p+2}\cdots j_n \\ k_{p+1} k_{p+2}\cdots k_n \end{pmatrix}}\delta^{j_{p+1} j_{p+2}\cdots j_n}_{k_{p+1} k_{p+2}\cdots k_n} a_{i_{p+1} k_{p+1}}a_{i_{p+2} k_{p+2}}\cdots a_{i_n k_n}\right)$$

$$= \sum_{1\leqslant j_1<j_2<\cdots<j_p\leqslant n} \det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}\left[\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}\right].$$

这就证明了定理. □

由代数余子式的定义可知

$$\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\det A\begin{pmatrix} i_{p+1} & \cdots & i_n \\ j_{p+1} & \cdots & j_n \end{pmatrix}$$

是子式 $\det A\begin{pmatrix} i_1 & \cdots & i_p \\ j_1 & \cdots & j_p \end{pmatrix}$ 的代数余子式, 所以 Laplace 展开定理也可以叙述为:

任取指标 $1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$, 遍取由第 $i_1, i_2, \cdots, i_p$ 行决定的 p 阶子式 (共有 $\binom{n}{p}$ 个), 分别乘以其代数余子式, 则总和为 $\det(A)$.

所以, 这个定理也称为按照第 $i_1, i_2, \cdots, i_p$ 行的 **Laplace 展开**. 下面给出按照第 $i_1, i_2, \cdots, i_p$ 列的 Laplace 展开.

定理 2.3.4 (Laplace 定理) 任取指标 $1 \leqslant i_1 < i_2 < \cdots < i_p \leqslant n$, 则

$$\det(A) = \sum_{1 \leqslant j_1 < j_2 < \cdots < j_p \leqslant n} \det A\begin{pmatrix} j_1 & \cdots & j_p \\ i_1 & \cdots & i_p \end{pmatrix} \left[\delta^{i_1 i_2 \cdots i_n}_{j_1 j_2 \cdots j_n} \det A\begin{pmatrix} j_{p+1} & \cdots & j_n \\ i_{p+1} & \cdots & i_n \end{pmatrix}\right], \tag{2.3.14}$$

其中 $i_1 i_2 \cdots i_n$ 和 $j_1 j_2 \cdots j_n$ 是 $1, 2, \cdots, n$ 的排列, 且

$$1 \leqslant i_{p+1} < \cdots < i_n \leqslant n, \quad 1 \leqslant j_{p+1} < \cdots < j_n \leqslant n.$$

证 由于方阵 A 的转置方阵 A' 是将 A 的行列地位互换而得. 将 $\det(A')$ 按照第 $i_1, i_2, \cdots, i_p$ 行作 Laplace 展开, 我们得到 $\det(A)$ 按照第 $i_1, i_2, \cdots, i_p$ 列的 Laplace 展开式 (2.3.14). □

作为 Laplace 展开式的最简单应用, 有

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & 0 & \cdots & 0 \\ a_{p+1,1} & \cdots & a_{p+1,p} & a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} & a_{n,p+1} & \cdots & a_{nn} \end{pmatrix}$$

$$= \det\begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} \end{pmatrix} \det\begin{pmatrix} a_{p+1,p+1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots \\ a_{n,p+1} & \cdots & a_{nn} \end{pmatrix}. \tag{2.3.15}$$

通常, 很少应用一般的 Laplace 展开式来计算行列式. 今后主要是应用按一行及按一列的 Laplace 展开式:

$$\det(A) = \sum_{j=1}^{n} a_{ij} A_{ij}, \quad 1 \leqslant i \leqslant n \tag{2.3.16}$$

和

$$\det(A) = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad 1 \leqslant j \leqslant n, \tag{2.3.17}$$

其中 A_{ij} 是方阵 A 中子式 $\det A\binom{i}{j}=a_{ij}$ 的代数余子式. 亦即在方阵 A 中除去第 i 行, 第 j 列的所有元素后, 得到的一个 $n-1$ 阶子式, 它的代数子式就是 A_{ij}, 所以

$$A_{ij}=(-1)^{i+j}\det\begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}. \tag{2.3.18}$$

定理 2.3.5 对任一 n 阶方阵 A, 有

$$\sum_{j=1}^{n} a_{ij}A_{kj}=\delta_{ik}\det(A), \quad 1\leqslant i,k\leqslant n, \tag{2.3.19}$$

及

$$\sum_{j=1}^{n} a_{ji}A_{jk}=\delta_{ik}\det(A), \quad 1\leqslant i,k\leqslant n, \tag{2.3.20}$$

其中 δ_{ik} 为 Kronecker 符号, 定义见习题 2.2.1.

证 由于方法一样, 我们只证等式 (2.3.19) 成立. 在 $i=k$ 时, 利用按照第 i 行的 Laplace 展开可知定理成立. 在 $i\neq k$ 时, 构作方阵

$$B=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,n} \\ a_{i1} & \cdots & a_{in} \\ a_{k+1,1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} (\text{第 } k \text{ 行}).$$

它是由方阵 A 将第 k 行换上 A 的第 i 行而得. 再将 $\det(B)$ 按照第 k 行作 Laplace 展开, 便得

$$\det(B)=\sum_{j=1}^{n} a_{ij}A_{kj}.$$

另一方面, 由于 $i\neq k$, 所以方阵 B 中第 i 行和第 k 行相同, 因此 $\det(B)=0$. □

利用按照第 1 行的 Laplace 展开, 便得

$$\det(A)=\sum_{j=1}^{n} a_{1j}A_{1j}=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}. \tag{2.3.21}$$

右式是 n 个 $n-1$ 阶行列式的和, 所以这就把 n 阶行列式的计算化为 $n-1$ 阶行列式的计算. 继续作下去, 最终可将行列式的值算出来.

另一方面, 利用这个关系, 我们还可用归纳法来定义行列式.

定义 2.3.6 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{2.3.22}$$

的行列式是元素 $a_{11}, \cdots, a_{1n}, \cdots, a_{n1}, \cdots, a_{nn}$ 的函数, 记作 $\det(A)$. 它定义为:

(1) 当 $n=1$ 时 $\det(A)=a_{11}$;

(2) 假设对一切 $n-1$ 阶方阵, 其行列式都已定义好, 则 n 阶方阵 A 的行列式定义为

$$\det(A)=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}. \tag{2.3.23}$$

由归纳法原理, 便给出了行列式的定义. 这个定义和我们讲的定义 2.2.3 是等价的.

习 题 2.3

2.3.1 试将下列行列式按第二行作 Laplace 展开, 再按第二列作 Laplace 展开, 从而求行列式的值:

(1) $\det\begin{pmatrix} 1 & 5 & 3 \\ 2 & 3 & -1 \\ 3 & 2 & -4 \end{pmatrix}$; **(2)** $\det\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$; **(3)** $\det\begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$;

(4) $\det\begin{pmatrix} a & 1 & 1 & 1 \\ b & 0 & 1 & 1 \\ c & 1 & 0 & 1 \\ d & 1 & 1 & 0 \end{pmatrix}$; **(5)** $\det\begin{pmatrix} x & y & z & t \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 1 \end{pmatrix}$.

2.3.2 试将下列行列式按照第 2, 3 行作 Laplace 展开, 再求行列式的值.

(1) $\det\begin{pmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 & 1 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 1 & 6 \end{pmatrix}$; **(2)** $\det\begin{pmatrix} 5 & 6 & 0 & 0 & 0 \\ 1 & 5 & 6 & 0 & 0 \\ 0 & 1 & 5 & 6 & 0 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & 1 & 5 \end{pmatrix}$.

2.3.3 求行列式

$$\det\begin{pmatrix} 0 & 0 & 0 & a_4 \\ 0 & 0 & a_3 & 0 \\ 0 & a_2 & 0 & 0 \\ a_1 & 0 & 0 & 0 \end{pmatrix}$$

的一切元素的三阶代数余子式的和.

2.3.4 如果把行列式

$$\det(A)=\det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} & 1\\ \vdots & & \vdots & \vdots\\ a_{n1} & \cdots & a_{n,n-1} & 1\end{pmatrix}$$

的第 j 行易为 $x_1, x_2, \cdots, x_{n-1}, 1$, 得到的新行列式记作 $\det(A_j)$, $j=1, 2, \cdots, n$, 试证:

$$\det(A_1)+\det(A_2)+\cdots+\det(A_n)=\det(A).$$

2.3.5 给定域 $\mathbb{F}$ 上不全为零的多项式 $f_{11}(x), f_{12}(x), \cdots, f_{1n}(x)$, $n>3$. 试证: 存在域 $\mathbb{F}$ 上 $n(n-1)$ 个多项式 $f_{ij}(x)$, $2\leqslant i\leqslant n$, $1\leqslant j\leqslant n$, 使得 n 阶行列式

$$\det\begin{pmatrix} f_{11}(x) & \cdots & f_{1n}(x)\\ \vdots & & \vdots\\ f_{n1}(x) & \cdots & f_{nn}(x)\end{pmatrix}=(f_{11}(x),\cdots,f_{1n}(x)).$$

2.3.6 试证: 定义 2.3.6 所作的行列式定义和定义 2.2.3 所给出的行列式定义等价.

2.3.7 试从定义 2.3.6 所给出的行列式定义出发, 直接来证明定理 2.3.5.

2.3.8 设 n 阶方阵

$$A=A(x)=\begin{pmatrix} a_{11}(x) & \cdots & a_{1n}(x)\\ \vdots & & \vdots\\ a_{n1}(x) & \cdots & a_{nn}(x)\end{pmatrix}$$

的元素 $a_{ij}(x)$ 都是变数 x 的可微函数, 试证: $A(x)$ 的行列式 $\det(A(x))$ 有

$$\frac{\mathrm{d}\det(A(x))}{\mathrm{d}x}=\sum_{i,j=1}^{n}\frac{\mathrm{d}a_{ij}(x)}{\mathrm{d}x}A_{ij}(x),$$

其中 $A_{ij}(x)$ 是方阵 $A(x)$ 中子式 $a_{ij}(x)$ 的代数余子式, $1\leqslant i,j\leqslant n$.

2.3.9 设 $b_{jk}=a_{j1}+a_{j2}+\cdots+a_{jn}-\lambda a_{jk}$, 试证:

$$\det\begin{pmatrix} b_{11} & \cdots & b_{1n}\\ \vdots & & \vdots\\ b_{n1} & \cdots & b_{nn}\end{pmatrix}=(-\lambda)^{n-1}(n-\lambda)\det\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix},$$

其中 $\lambda=1, 2$.

2.3.10 设 n 阶方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}$$

有 $\sum\limits_{j=1}^{n}a_{ij}=\sum\limits_{j=1}^{n}a_{ji}=0$, $i=1,2,\cdots,n$. 试证: 方阵 A 的 n^2 个代数余子式 A_{ij} 的值都相等.

§2.4 行列式计算的一些技巧

在这一节将通过一些典型的例题, 介绍行列式计算的一些技巧. 希望读者举一反三, 灵活运用行列式的各种性质, 掌握一些技巧.

例 2.4.1 计算 **Vandermonde 行列式**

$$V_n(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix}1 & 1 & \cdots & 1\\ a_1 & a_2 & \cdots & a_n\\ a_1^2 & a_2^2 & \cdots & a_n^2\\ \vdots & \vdots & & \vdots\\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1}\end{pmatrix}.$$

解 将 Vandermonde 行列式的第 $n-1$ 行遍乘以 $-a_1$, 再加到第 n 行上去, 然后, 将它的第 $n-2$ 行遍乘以 $-a_1$, 再加到第 $n-1$ 行上去. 这样依次作下去, 最后, 将它的第 1 行乘以 $-a_1$, 再加到第 2 行上去, 便得

$$V_n(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix}1 & 1 & \cdots & 1\\ 0 & a_2-a_1 & \cdots & a_n-a_1\\ 0 & a_2(a_2-a_1) & \cdots & a_n(a_n-a_1)\\ \vdots & \vdots & & \vdots\\ 0 & a_2^{n-2}(a_2-a_1) & \cdots & a_n^{n-2}(a_n-a_1)\end{pmatrix}.$$

再按第一列作 Laplace 展开, 于是

$$V_n(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix}a_2-a_1 & \cdots & a_n-a_1\\ a_2(a_2-a_1) & \cdots & a_n(a_n-a_1)\\ \vdots & & \vdots\\ a_2^{n-2}(a_2-a_1) & \cdots & a_n^{n-2}(a_n-a_1)\end{pmatrix}.$$

所以

$$V_n(a_1,a_2,\cdots,a_n)=V_{n-1}(a_2,\cdots,a_n)\prod_{j=2}^{n}(a_j-a_1).$$

这样就把 n 阶 Vandermonde 行列式化为 $n-1$ 阶 Vandermonde 行列式. 即建立起递推公式. 有了这个递推公式, 利用归纳法, 便能算出 Vandermonde 行列式的值:

$$\begin{aligned}V_n(a_1,a_2,\cdots,a_n)&=\prod_{j=2}^{n}(a_j-a_1)V_{n-1}(a_2,\cdots,a_n)\\ &=\prod_{j=2}^{n}(a_j-a_1)\prod_{j=3}^{n}(a_j-a_2)V_{n-2}(a_3,\cdots,a_n)=\cdots\\ &=\prod_{j=2}^{n}(a_j-a_1)\prod_{j=3}^{n}(a_j-a_2)\cdots\prod_{j=n}^{n}(a_j-a_{n-1})=\prod_{1\leqslant i<j\leqslant n}(a_j-a_i).\end{aligned}$$

□

这个行列式的计算方法的原则大致是: 利用已给行列式的特性, 建立起同类型的 n 阶行列式和 $n-1$ 阶 (或更低阶) 行列式间的关系, 即导出递推公式. 然后, 利用归纳法的原理写出行列式的值. 这个原则很有效, 下面再举一个运用这个原则的例子.

例 2.4.2 计算**三对角方阵**的行列式

$$D_n = \det\begin{pmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{pmatrix}.$$

解 将行列式 D_n 按第一行作 Laplace 展开, 于是得到递推公式:

$$D_n = aD_{n-1} - bcD_{n-2}, \quad n = 3, 4, \cdots.$$

显然 $D_1 = a$, $D_2 = a^2 - bc$.

从这个递推公式出发, 下面引进一个特殊的技巧来求行列式 D_n.

设若 $bc = 0$, 则 $D_n = aD_{n-1}$, $D_1 = a$. 所以 $D_n = a^n$.

设若 $bc \neq 0$, 作二次多项式 $x^2 - ax + bc$, 它有两个根 u 和 v. 自然 $u + v = a$, $uv = bc$. 今由递推公式 $D_n = aD_{n-1} - bcD_{n-2}$ 可知:

$$U_n = D_n - uD_{n-1} = vD_{n-1} - uvD_{n-2} = vU_{n-1},$$
$$V_n = D_n - vD_{n-1} = uD_{n-1} - uvD_{n-2} = uV_{n-1},$$

且 $U_2 = D_2 - uD_1 = a^2 - bc - au = v^2$, $V_2 = D_2 - vD_1 = a^2 - bc - av = u^2$. 因此

$$U_n = D_n - uD_{n-1} = v^n, \quad V_n = D_n - vD_{n-1} = u^n, \quad n = 2, 3, \cdots.$$

设 $u \neq v$, 即 $a^2 \neq 4bc$, 则

$$D_n = \frac{u^{n+1} - v^{n+1}}{u - v}, \quad u = \frac{a + \sqrt{a^2 - 4bc}}{2}, \quad v = \frac{a - \sqrt{a^2 - 4bc}}{2}.$$

设 $u = v$, 即 $a^2 = 4bc$, 则 $D_n = uD_{n-1} + u^n$, $n = 2, 3, \cdots$, $D_1 = a = 2u$. 所以

$$D_n = (n+1)u^n = (n+1)\left(\frac{a}{2}\right)^n.$$

□

除了运用递推公式来求行列式的值外, 下面还介绍许多别的方法.

例 2.4.3 求证

$$d_n = \det\begin{pmatrix} a_{11}+x_1 & \cdots & a_{1n}+x_n \\ \vdots & & \vdots \\ a_{n1}+x_1 & \cdots & a_{nn}+x_n \end{pmatrix} = \det(A) + \sum_{j=1}^{n} x_j \sum_{k=1}^{n} A_{kj},$$

其中 A_{kj} 是方阵 $A=(a_{jk})$ 的第 k 行, 第 j 列位置的元素的代数余子式.

证 今 d_n 等于行列式

$$\det\begin{pmatrix} a_{11} & a_{12}+x_2 & \cdots & a_{1n}+x_n \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2}+x_2 & \cdots & a_{nn}+x_n \end{pmatrix} + \det\begin{pmatrix} x_1 & a_{12}+x_2 & \cdots & a_{1n}+x_n \\ \vdots & \vdots & & \vdots \\ x_1 & a_{n2}+x_2 & \cdots & a_{nn}+x_n \end{pmatrix}.$$

由于行列式有两列成比例, 则行列式的值为零, 所以上式的第二个行列式等于行列式

$$\det\begin{pmatrix} x_1 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ x_1 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

这样继续作下去, 所以有

$$d_n = \det(A) + \sum_{j=1}^{n} \det\begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & x_j & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & x_j & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}.$$

即

$$d_n = \det(A) + \sum_{j=1}^{n} x_j \sum_{k=1}^{n} A_{kj}. \qquad \square$$

例 2.4.4 求证

$$D_n = \det\begin{pmatrix} 1 & \cdots & 1 \\ a_{21}-a_{11} & \cdots & a_{2n}-a_{1n} \\ a_{31}-a_{21} & \cdots & a_{3n}-a_{2n} \\ \vdots & & \vdots \\ a_{n1}-a_{n-1,1} & \cdots & a_{nn}-a_{n-1,n} \end{pmatrix} = \sum_{j,k=1}^{n} A_{jk}.$$

证 今将第 $1, 2, \cdots, n-1$ 行全都加到第 n 行上去, 再将第 $1, 2, \cdots, n-2$ 行

全都加到第 $n-1$ 行上去. 这样依次作下去, 最后将第 1 行加到第 2 行上去, 便得

$$D_n = \det\begin{pmatrix} a_{11}+(1-a_{11}) & \cdots & a_{1n}+(1-a_{1n}) \\ a_{21}+(1-a_{11}) & \cdots & a_{2n}+(1-a_{1n}) \\ a_{31}+(1-a_{11}) & \cdots & a_{3n}+(1-a_{1n}) \\ \vdots & & \vdots \\ a_{n1}+(1-a_{11}) & \cdots & a_{nn}+(1-a_{1n}) \end{pmatrix}.$$

由例 2.4.3 可知, 此行列式的值为

$$\begin{aligned} D_n &= \det(A) + \sum_{j=1}^{n}(1-a_{1j})\sum_{k=1}^{n}A_{kj} = \det(A) + \sum_{j,k=1}^{n}A_{jk} - \sum_{j,k=1}^{n}a_{1j}A_{kj} \\ &= \det(A) + \sum_{j,k=1}^{n}A_{jk} - \sum_{k=1}^{n}\delta_{1k}\det(A) = \sum_{j,k=1}^{n}A_{jk}. \end{aligned}$$

□

由例 2.4.3 和例 2.4.4, 便建立了公式

$$\det(A) + x\sum_{j,k=1}^{n}A_{jk} = \det\begin{pmatrix} a_{11}+x & \cdots & a_{1n}+x \\ \vdots & & \vdots \\ a_{n1}+x & \cdots & a_{nn}+x \end{pmatrix}$$

$$= \det(A) + x\det\begin{pmatrix} 1 & \cdots & 1 \\ a_{21}-a_{11} & \cdots & a_{2n}-a_{1n} \\ a_{31}-a_{21} & \cdots & a_{3n}-a_{2n} \\ \vdots & & \vdots \\ a_{n1}-a_{n-1,1} & \cdots & a_{nn}-a_{n-1,n} \end{pmatrix}.$$

利用这一公式, 例如可以计算 n 阶行列式

$$\Delta_n = \det\begin{pmatrix} 1+x_1y_1 & 1+x_1y_2 & \cdots & 1+x_1y_n \\ 1+x_1y_1^2 & 1+x_1y_2^2 & \cdots & 1+x_1y_n^2 \\ \vdots & \vdots & & \vdots \\ 1+x_1y_1^n & 1+x_1y_2^n & \cdots & 1+x_1y_n^n \end{pmatrix},$$

其中 $y_1, y_2, \cdots, y_n \neq 1$. 读者试自证之.

例 2.4.5　计算 n 阶行列式

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix} 0 & a_1+a_2 & \cdots & a_1+a_n \\ a_2+a_1 & 0 & \cdots & a_2+a_n \\ a_3+a_1 & a_3+a_2 & \cdots & a_3+a_n \\ \vdots & \vdots & & \vdots \\ a_n+a_1 & a_n+a_2 & \cdots & 0 \end{pmatrix},$$

其中 $a_1a_2\cdots a_n \neq 0$.

解　将 n 阶行列式 $\Delta(a_1,a_2,\cdots,a_n)$ 镶边构成 $n+1$ 阶行列式, 使其值不变, 即作

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix} 1 & a_1 & a_2 & \cdots & a_n \\ 0 & 0 & a_1+a_2 & \cdots & a_1+a_n \\ 0 & a_2+a_1 & 0 & \cdots & a_2+a_n \\ 0 & a_3+a_1 & a_3+a_2 & \cdots & a_3+a_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_n+a_1 & a_n+a_2 & \cdots & 0 \end{pmatrix}.$$

再将第 $2, 3, \cdots, n+1$ 行分别减去第一行, 有

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix} 1 & a_1 & a_2 & \cdots & a_n \\ -1 & -a_1 & a_1 & \cdots & a_1 \\ -1 & a_2 & -a_2 & \cdots & a_2 \\ -1 & a_3 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & a_n & a_n & \cdots & -a_n \end{pmatrix}.$$

将这 $n+1$ 阶行列式再镶边, 构成 $n+2$ 阶行列式, 使其值也不变, 即

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & a_1 & a_2 & \cdots & a_n \\ a_1 & -1 & -a_1 & a_1 & \cdots & a_1 \\ a_2 & -1 & a_2 & -a_2 & \cdots & a_2 \\ a_3 & -1 & a_3 & a_3 & \cdots & a_3 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_n & -1 & a_n & a_n & \cdots & -a_n \end{pmatrix}.$$

再将第 $3, 4, \cdots, n+2$ 列分别减去第一列, 便得

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix}1&0&-1&-1&\cdots&-1\\0&1&a_1&a_2&\cdots&a_n\\a_1&-1&-2a_1&0&\cdots&0\\a_2&-1&0&-2a_2&\cdots&0\\\vdots&\vdots&\vdots&\vdots&&\vdots\\a_n&-1&0&0&\cdots&-2a_n\end{pmatrix}.$$

将第 $3, 4, \cdots, n+2$ 列乘以 $1/2$, 然后都加到第一列上去. 再将第 $3, 4, \cdots, n+2$ 列分别乘以 $-1/2a_1, -1/2a_2, \cdots, -1/2a_n$, 然后都加到第二列上去. 于是有

$$\Delta(a_1,a_2,\cdots,a_n)=\det\begin{pmatrix}1-\dfrac{n}{2}&\dfrac{1}{2}\displaystyle\sum_{j=1}^{n}\dfrac{1}{a_j}&-1&\cdots&-1\\\dfrac{1}{2}\displaystyle\sum_{j=1}^{n}a_j&1-\dfrac{n}{2}&a_1&\cdots&a_n\\0&0&-2a_1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&-2a_n\end{pmatrix}.$$

因此,

$$\Delta(a_1,a_2,\cdots,a_n)=(-2)^n(a_1a_2\cdots a_n)\det\begin{pmatrix}1-\dfrac{n}{2}&\dfrac{1}{2}\displaystyle\sum_{j=1}^{n}\dfrac{1}{a_j}\\\dfrac{1}{2}\displaystyle\sum_{j=1}^{n}a_j&1-\dfrac{n}{2}\end{pmatrix},$$

即

$$\Delta(a_1,a_2,\cdots,a_n)=(-2)^n(a_1a_2\cdots a_n)\Big[(2-n)^2-\sum_{j,k=1}^{n}\frac{a_j}{a_k}\Big].\qquad\square$$

这个例子告诉我们, 有时把行列式的阶增高, 反而容易求出行列式的值, 而且在行列式的计算中, 充分利用元素间的特性, 才能较简单地计算出行列式的值来.

例 2.4.6 最后, 给出五种方法来计算同一个行列式

$$d_n=\det\begin{pmatrix}x&a&a&\cdots&a\\a&x&a&\cdots&a\\a&a&x&\cdots&a\\\vdots&\vdots&\vdots&&\vdots\\a&a&a&\cdots&x\end{pmatrix}.$$

解法 1　将第 2, 3, $\cdots$, n 行都加到第一行上去, 得

$$d_n = \det\begin{pmatrix} x+(n-1)a & x+(n-1)a & \cdots & x+(n-1)a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{pmatrix},$$

因此,

$$d_n = [x+(n-1)a]\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{pmatrix}.$$

再将第一行遍乘 $-a$, 然后分别加到第 2, 3, $\cdots$, n 行上去, 得

$$d_n = [x+(n-1)a]\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ & x-a & & \\ & & \ddots & \\ & & & x-a \end{pmatrix} = (x-a)^{n-1}[x+(n-1)a].$$

□

解法 2　将第 2, 3, $\cdots$, n 行分别减去第一行, 得

$$d_n = \det\begin{pmatrix} x & a & a & \cdots & a \\ a-x & x-a & 0 & \cdots & 0 \\ a-x & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a-x & 0 & 0 & \cdots & x-a \end{pmatrix}.$$

再将第 2, 3, $\cdots$, n 列都加到第一列上去, 便有

$$d_n = \det\begin{pmatrix} x+(n-1)a & a & a & \cdots & a \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x-a \end{pmatrix} = (x-a)^{n-1}[x+(n-1)a].$$

□

解法 3 将 d_n 添加一行及一列, 构成 $n+1$ 阶行列式

$$d_n = \det\begin{pmatrix} 1 & a & a & \cdots & a \\ 0 & x & a & \cdots & a \\ 0 & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a & a & \cdots & x \end{pmatrix}.$$

再将第 $2, 3, \cdots, n+1$ 行分别减去第一行, 于是有

$$d_n = \det\begin{pmatrix} 1 & a & a & \cdots & a \\ -1 & x-a & 0 & \cdots & 0 \\ -1 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & x-a \end{pmatrix}.$$

在 $x = a$ 时, 显然有 $d_n = 0$; 在 $x \neq a$ 时, 则

$$d_n = (x-a)^n \det\begin{pmatrix} 1 & \dfrac{a}{x-a} & \dfrac{a}{x-a} & \cdots & \dfrac{a}{x-a} \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

所以

$$d_n = (x-a)^n \det\begin{pmatrix} 1 + \dfrac{na}{x-a} & \dfrac{a}{x-a} & \dfrac{a}{x-a} & \cdots & \dfrac{a}{x-a} \\ & 1 & & \cdots & \\ & & 1 & \cdots & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

即

$$d_n = (x-a)^{n-1}[x + (n-1)a]. \qquad \square$$

解法 4 令

$$d_n = \det\begin{pmatrix} x - a + a & a & \cdots & a \\ 0 + a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ 0 + a & a & \cdots & x \end{pmatrix}.$$

所以,

$$\det\begin{pmatrix} x-a & a & \cdots & a \\ 0 & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ 0 & a & \cdots & x \end{pmatrix} + \det\begin{pmatrix} a & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{pmatrix} = (x-a)d_{n-1} + a(x-a)^{n-1}.$$

因此导出递推公式

$$\begin{aligned} d_1 &= x, \\ d_2 &= (x-a)d_1 + a(x-a), \\ d_3 &= (x-a)d_2 + a(x-a)^2, \\ &\vdots \\ d_{n-1} &= (x-a)d_{n-2} + a(x-a)^{n-2}, \\ d_n &= (x-a)d_{n-1} + a(x-a)^{n-1}. \end{aligned}$$

依次乘以 $(x-a)^{n-1}$, $(x-a)^{n-2}$, $\cdots$, $x-a$, 1, 便有

$$\begin{aligned} d_n &= (x-a)d_{n-1} + a(x-a)^{n-1}, \\ (x-a)d_{n-1} &= (x-a)^2 d_{n-2} + a(x-a)^{n-1}, \\ &\vdots \\ (x-a)^{n-3}d_3 &= (x-a)^{n-2}d_2 + a(x-a)^{n-1}, \\ (x-a)^{n-2}d_2 &= (x-a)^{n-1}d_1 + a(x-a)^{n-1}, \\ (x-a)^{n-1}d_1 &= (x-a)^{n-1}x. \end{aligned}$$

再全体相加, 有

$$d_n = (x-a)^{n-1}x + a(x-a)^{n-1}(n-1) = [x+(n-1)a](x-a)^{n-1}.$$ □

解法 5 利用例 2.4.3 和例 2.4.4. 令

$$d_n = \det\begin{pmatrix} (x-a)+a & 0+a & \cdots & 0+a \\ 0+a & (x-a)+a & \cdots & 0+a \\ \vdots & \vdots & & \vdots \\ 0+a & 0+a & \cdots & (x-a)+a \end{pmatrix}.$$

所以

$$d_n = \det\big(\mathrm{diag}\,(x-a, x-a, \cdots, x-a)\big) + a\det\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ a-x & x-a & \cdots & 0 & 0 \\ 0 & a-x & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & x-a & 0 \\ 0 & 0 & \cdots & a-x & x-a \end{pmatrix}.$$

将右式中的第二个行列式的第 2, 3, $\cdots$, n 列全加到第一列上去, 再利用 Laplace 展开, 所以得

$$d_n = (x-a)^n + na\det\begin{pmatrix} x-a & \cdots & 0 & 0 \\ a-x & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & x-a & 0 \\ 0 & \cdots & a-x & x-a \end{pmatrix},$$

即

$$d_n = (x-a)^n + na(x-a)^{n-1} = (x-a)^{n-1}[x+(n-1)a]. \qquad \square$$

当然, 我们也可以按第一行作 Laplace 展开来求这个行列式. 另一方面, 这个例子告诉我们, 在演算一个问题时, 需要仔细分析已给的条件, 灵活运用已经知道的性质和掌握的技巧, 不要死套公式, 这样就能很快地求出答案.

习　　题　2.4

2.4.1 试计算 n 阶行列式:

$$\det\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ 2 & 3 & 4 & \cdots & n+1 \\ 3 & 4 & 5 & \cdots & n+2 \\ \vdots & \vdots & \vdots & & \vdots \\ n & n+1 & n+2 & \cdots & 2n-1 \end{pmatrix}, \det\begin{pmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \end{pmatrix}.$$

2.4.2 试计算 n 阶行列式:

$$\det\begin{pmatrix} \sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1 \\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2 \\ \vdots & \vdots & & \vdots \\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n \end{pmatrix}, \det\begin{pmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{pmatrix}.$$

2.4.3 试计算 n 阶行列式:

$$\det\begin{pmatrix} c_1 & b & b & \cdots & b \\ a & c_2 & b & \cdots & b \\ a & a & c_3 & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & c_n \end{pmatrix},\quad \det\begin{pmatrix} c_1 & b & b & \cdots & b & 1 \\ a & c_2 & b & \cdots & b & 1 \\ a & a & c_3 & \cdots & b & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a & a & a & \cdots & c_{n-1} & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}.$$

2.4.4 试计算 n 阶行列式:

$$\det\begin{pmatrix} a & 0 & 0 & \cdots & 0 & 1 \\ 0 & a & 0 & \cdots & 0 & 0 \\ 0 & 0 & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & 0 \\ 1 & 0 & 0 & \cdots & 0 & a \end{pmatrix},\ \det\begin{pmatrix} 1+x_1y_1 & 1+x_1^2y_2 & \cdots & 1+x_1^ny_n \\ 1+x_2y_1 & 1+x_2^2y_2 & \cdots & 1+x_2^ny_n \\ \vdots & \vdots & & \vdots \\ 1+x_ny_1 & 1+x_n^2y_2 & \cdots & 1+x_n^ny_n \end{pmatrix}.$$

2.4.5 试计算 n 阶行列式:

$$\det\begin{pmatrix} c_1 & a_2 & \cdots & a_n \\ a_1 & c_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & c_n \end{pmatrix},$$

且用它来计算例 2.4.6.

2.4.6 求 n 阶行列式的值:

$$\sum_{\binom{1\ 2\cdots n}{j_1 j_2\cdots j_n}} \det\begin{pmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_n} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_n} \\ \vdots & \vdots & & \vdots \\ a_{nj_1} & a_{nj_2} & \cdots & a_{nj_n} \end{pmatrix}.$$

2.4.7 试计算 n 阶行列式:

$$\det\begin{pmatrix} \dfrac{1}{a_1+b_1} & \cdots & \dfrac{1}{a_1+b_n} \\ \vdots & & \vdots \\ \dfrac{1}{a_n+b_1} & \cdots & \dfrac{1}{a_n+b_n} \end{pmatrix}.$$

2.4.8 试计算 n 阶行列式:

$$\det\begin{pmatrix} \dbinom{m}{0} & \dbinom{m}{1} & \cdots & \dbinom{m}{n-1} \\ \vdots & \vdots & & \vdots \\ \dbinom{m+n-1}{0} & \dbinom{m+n-1}{1} & \cdots & \dbinom{m+n-1}{n-1} \end{pmatrix}.$$

2.4.9 设 $a_{11} \neq 0$. 试证明: 当 $n \geqslant 2$ 时, n 阶方阵 $A = (a_{ij})$ 的行列式等于

$$a_{11}^{2-n}\det\begin{pmatrix} \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} & \det\begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} & \cdots & \det\begin{pmatrix} a_{11} & a_{1n} \\ a_{21} & a_{2n} \end{pmatrix} \\ \vdots & \vdots & & \vdots \\ \det\begin{pmatrix} a_{11} & a_{12} \\ a_{n1} & a_{n2} \end{pmatrix} & \det\begin{pmatrix} a_{11} & a_{13} \\ a_{n1} & a_{n3} \end{pmatrix} & \cdots & \det\begin{pmatrix} a_{11} & a_{1n} \\ a_{n1} & a_{nn} \end{pmatrix} \end{pmatrix}.$$

注意: 右边是 $n-1$ 阶行列式. 所以, 这也给出了一个公式, 它把 n 阶行列式的计算化为 $n-1$ 阶行列式的计算.

2.4.10 求 $n+1$ 阶行列式的值:

$$\det\begin{pmatrix} x & 1 & & & & & \\ -n & x-2 & 2 & & & & \\ & -(n-1) & x-4 & 3 & & & \\ & & \ddots & \ddots & & \ddots & \\ & & & -2 & x-2(n-1) & & n \\ & & & & -1 & & x-2n \end{pmatrix}.$$

2.4.11 试计算 n 阶行列式:

$$\det\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ -a_1 & a_2 & b_{23} & \cdots & b_{2n} \\ -a_1 & -a_2 & a_3 & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ -a_1 & -a_2 & -a_3 & \cdots & b_{n-1,n} \\ -a_1 & -a_2 & -a_3 & \cdots & a_n \end{pmatrix}.$$

2.4.12 设 $0 \leqslant k \leqslant n-2$, 求 n 阶缺项 Vandermonde 行列式的值:

$$\det\begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^k & a_1^{k+2} & \cdots & a_1^n \\ 1 & a_2 & a_2^2 & \cdots & a_2^k & a_2^{k+2} & \cdots & a_2^n \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^k & a_n^{k+2} & \cdots & a_n^n \end{pmatrix}.$$

2.4.13 设 a_{jk} 都是复数, $1 \leqslant j, k \leqslant n$. $A = (a_{jk})$ 为 n 阶方阵. 试证 **Lovy–Desplanques 定理**: 如果 $\det(A) = 0$, 则条件 $2|a_{jj}| > \sum\limits_{k=1}^{n} |a_{jk}|$, $1 \leqslant j \leqslant n$ 不全成立, 即 n 阶方阵 A 不是**对角线占优方阵**.

2.4.14 设 a_{jk} 都是实数, $1 \leqslant j, k \leqslant n$. $A = (a_{jk})$ 为 n 阶方阵. 如果 $a_{jj} > 0$, $1 \leqslant j \leqslant n$, $a_{jk} < 0$, $j \neq k$, $1 \leqslant j, k \leqslant n$. 又设 $\sum\limits_{k=1}^{n} a_{jk} > 0$, $1 \leqslant j \leqslant n$. 试证 **Minkowski 定理**: 行列式 $\det A > 0$.

2.4.15 设 a_{jk} 都是复数, $1 \leqslant j,k \leqslant n$. $A=(a_{jk})$ 为 n 阶方阵. 试证: 将 n 阶行列式

$$\det\begin{pmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{pmatrix}$$

展成 λ 的多项式, 则可表为

$$\lambda^n - a_1\lambda^{n-1} + a_2\lambda^{n-2} - \cdots + (-1)^{n-1}a_{n-1}\lambda + (-1)^n a_n,$$

其中 a_j 为行列式 $\det(A)$ 的所有 j 阶主子式的和, $j=1,2,\cdots,n$.

2.4.16 试用归纳法来证明**Burnside 定理**: 域 $\mathbb{F}$ 上反对称方阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \qquad a_{ji}=-a_{ij}, \qquad 1 \leqslant i \leqslant j \leqslant n$$

的行列式在 n 为奇数时值为零; 在 n 为偶数时, 记 $n=2m$, 则值为

$$\Bigg(\sum_{\begin{pmatrix} 1\,2\cdots\,2m \\ i_1 i_2\cdots i_{2m}\end{pmatrix};\ i_1<i_2,\ \cdots,\ i_{2m-1}<i_{2m}} \delta^{1\,2\,\cdots\,2m}_{i_1 i_2\cdots i_{2m}} a_{i_1 i_2} a_{i_3 i_4}\cdots a_{i_{2m-1}i_{2m}}\Bigg)^2.$$

§2.5 Cramer 法则

给定 mn 个数 a_{jk}, $1 \leqslant j \leqslant n$, $1 \leqslant k \leqslant m$ 及 n 个数 b_1, b_2, $\cdots$, b_n. 记 x_1, x_2, $\cdots$, x_m 为 m 个独立未知数. 则 n 个方程

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2, \\ \qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases} \tag{2.5.1}$$

构成**线性方程组**. $n \times 1$ 矩阵 $x_0=(x_1^{(0)}, x_2^{(0)}, \cdots, x_m^{(0)})'$ 称为线性方程组 (2.5.1) 的一组**解**, 如果它们同时适合这 n 个方程, 即

$$\sum_{k=1}^{m} a_{jk}x_k^{(0)} = b_j, \qquad j=1,2,\cdots,n. \tag{2.5.2}$$

线性方程组的理论就是求解理论. 它要解决下列问题: **(1)** 什么时候解才存在? **(2)** 如果解存在, 用什么办法可以求出所有解? **(3)** 什么时候解唯一存在? **(4)** 如果解不存在, 能不能求出最好的近似解 (最小二乘解).

关于一般线性方程组的求解理论, 放在第四章, 它回答了前三个问题. 在第十二章回答第四个问题. 在这一节, 我们讨论一种最简单的情形.

考虑由 n 个未知数 $x_1, x_2, \cdots, x_n$, n 个方程构成的线性方程组 (2.5.1). 记 n^2 个实数 a_{jk}, $1 \leqslant j, k \leqslant n$ 构成的 n 阶方阵为

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{2.5.3}$$

记它的行列式为

$$\Delta = \det(A). \tag{2.5.4}$$

定理 2.5.1 (Cramer 法则) 设线性方程组 (2.5.1) 中 $m = n$, 设式 (2.5.4) 给出的 n 阶行列式 $\Delta = \det(A) \neq 0$, 则线性方程组 (2.5.1) 有一组唯一的解

$$x_0 = (x_1^{(0)}, \cdots, x_n^{(0)})' = (\frac{\Delta_1}{\Delta}, \cdots, \frac{\Delta_n}{\Delta})', \tag{2.5.5}$$

其中 $\Delta_j = \sum\limits_{k=1}^{n} b_k A_{kj}$, 即

$$\Delta_j = \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}, \quad j = 1, 2, \cdots, n. \tag{2.5.6}$$

换句话说, Δ_j 是由 $\det(A)$ 中易第 j 列为常数项 $(b_1, b_2, \cdots, b_n)$ 所构成的新行列式. 而

$$\sum_{k=1}^{n} a_{jk} x_k^{(0)} = b_j, \qquad j = 1, 2, \cdots, n. \tag{2.5.7}$$

证 今用 $x_k = \Delta^{-1}\Delta_k$, $k = 1, 2, \cdots, n$ 代入线性方程组 (2.5.1), 则有

$$\sum_{k=1}^{n} a_{jk}\Delta_k = \sum_{k=1}^{n} a_{jk} \sum_{i=1}^{n} b_i A_{ik} = \sum_{i=1}^{n} b_i \sum_{k=1}^{n} a_{jk} A_{ik} = \Delta \sum_{i=1}^{n} b_i \delta_{ij} = \Delta b_j,$$

所以 $(\frac{\Delta_1}{\Delta}, \cdots, \frac{\Delta_n}{\Delta})$ 为一组解.

下面再证唯一性, 即任取一组解 $y_0 = (y_1^{(0)}, \cdots, y_n^{(0)})'$, 于是有 $\sum\limits_{k=1}^{n} a_{jk} y_k^{(0)} = b_j$, $j = 1, 2, \cdots, n$. 将第 j 个方程的两边遍乘行列式 $\det(A)$ 的第 j 行, 第 i 列位置元素的代数余子式 A_{ji}, 然后按照对指标 j 求和, 即有

$$\begin{aligned} \Delta_i &= \sum_{j=1}^{n} b_j A_{ji} = \sum_{j=1}^{n}\sum_{k=1}^{n} y_k^{(0)} a_{jk} A_{ji} = \sum_{k=1}^{n} y_k^{(0)} \sum_{j=1}^{n} a_{jk} A_{ji} \\ &= \sum_{k=1}^{n} y_k^{(0)} \delta_{ki} \det(A) = y_i^{(0)} \Delta. \end{aligned}$$

所以, 当 $\Delta \neq 0$ 时, 有 $y_i^{(0)} = \Delta_i \Delta^{-1}$, $i = 1, 2, \cdots, n$. 因此证明了: 线性方程组若有解, 则解必为 $(\Delta_1/\Delta, \cdots, \Delta_n/\Delta)'$. 这也证明了若有解, 则解必唯一. □

这个定理给出了一类特殊的线性方程组求解的完整理论. 在第四章中, 将要建立一般的线性方程组求解的完整理论. 在建立理论的过程中, 总希望把一般情形化为能运用 Cramer 法则的特殊情形, 所以 Cramer 法则在理论上也具有重要地位.

Cramer 法则具体给出了求解的规则, 这就是建立了解和已给出的系数 a_{jk}, b_j 之间的关系. 但是在具体求解时, 需要计算 $n+1$ 个 n 阶行列式. 一般说, 这是很困难的. 所以, 在实际计算时, 往往不用 Cramer 法则, 而是运用消去法. 关于消去法理论, 在第四章中将会详细讨论. 另外注意, 在求出了线性方程组的解后, 一定要进行验算. 这样可以发现在计算行列式过程中产生的错误.

定义 2.5.2 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0, \\ \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0 \end{cases} \tag{2.5.8}$$

称为**齐次线性方程组**.

由 Cramer 法则可以得到

定理 2.5.3 设 $m = n$, 如果齐次线性方程组 (2.5.8) 有非零解, 即如果存在 n 个不全为零的实数 $x_1^{(0)}, \cdots, x_n^{(0)}$ 是它的解, 则行列式 $\det(A) = \Delta = 0$.

证 用反证法. 设行列式 $\det(A) = \Delta \neq 0$. 如果齐次线性方程组有非零解, 由上面唯一性证明可知有唯一解, 它是 $x_j^{(0)} = \Delta_j/\Delta$, $1 \leqslant j \leqslant n$. 由 Δ_j 的定义式 (2.5.6) 以及 $b_1 = b_2 = \cdots = b_n = 0$ 可知 $\Delta_j = 0$, $1 \leqslant j \leqslant n$. 这就和非零解的假设矛盾. □

习 题 2.5

2.5.1 试用 Cramer 法则求下列线性方程组的通解:

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 5, \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 1, \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 1, \\ 4x_1 + 3x_2 + 2x_3 + x_4 = -5. \end{cases}$$

2.5.2 试求下列线性方程组的通解, 其中 $a_1, a_2, \cdots, a_n$ 为 n 个不同的非零数.

(1) $\sum\limits_{i=1}^{n} a_i^{k-1} x_i = b_k$, $k = 1, 2, \cdots, n$; **(2)** $\sum\limits_{k=1}^{n} a_i^{k-1} x_k = b_i$, $i = 1, 2, \cdots, n$.

第三章　矩阵

§3.1　矩阵的代数运算

在 §2.2, 引进了矩阵. 在这一节, 我们引进矩阵的代数运算.

定义 3.1.1　两个矩阵 A 和 B 称为 **相等的矩阵**, 如果它们的行数相同, 列数也相同, 而且对应每个同行、同列的交叉位置的元素也相等. 这时记作 $A = B$.

由定义可知, 如果

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pq} \end{pmatrix},$$

则 $p = n$, $q = m$, 又,

$$a_{jk} = b_{jk}, \quad j = 1, 2, \cdots, n, \quad k = 1, 2, \cdots, m.$$

现在在矩阵之间引进代数运算.

定义 3.1.2　两个 $n \times m$ 矩阵的**和**是一个 $n \times m$ 矩阵, 它定义为

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1m} + b_{1m} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nm} + b_{nm} \end{pmatrix}. \tag{3.1.1}$$

这样就引进了**加法运算**. 由定义可知, 只有同类型的矩阵才可以相加. 且不难验证, 加法运算具有和实数加法相同的性质:

(1) 加法交换律: 设 A 和 B 是任两 $n\times m$ 矩阵, 则

$$B+A=A+B. \tag{3.1.2}$$

(2) 加法结合律: 设 A, B 和 C 是三个 $n\times m$ 矩阵, 则

$$(A+B)+C=A+(B+C). \tag{3.1.3}$$

由 (1) 和 (2) 可知, 在任意有限多个 $n\times m$ 矩阵求和时, 可以不计它们的前后位置和先后次序.

(3) 存在 $n\times m$ **零矩阵**

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \tag{3.1.4}$$

它是元素全为零的 $n\times m$ 矩阵, 记为 0. 这个矩阵有性质: 对一切 $n\times m$ 矩阵 A, 有

$$A+0=0+A=A. \tag{3.1.5}$$

今后, 零矩阵总由数字 0 来表示, 这在行文中是不致引起混淆的.

(4) 对任一 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix},$$

存在 $n\times m$ 矩阵

$$-A=\begin{pmatrix} -a_{11} & \cdots & -a_{1m} \\ \vdots & & \vdots \\ -a_{n1} & \cdots & -a_{nm} \end{pmatrix}. \tag{3.1.6}$$

它有性质:

$$A+(-A)=(-A)+A=0.$$

反之, 若 $n\times m$ 矩阵 B 适合等式 $A+B=0$, 则不难证明 $B=-A$. 我们称矩阵 $-A$ 是矩阵 A 的**负矩阵**. 所以矩阵 A 有且只有一个负矩阵 $-A$.

(5) 对任两 $n\times m$ 矩阵 A 和 B, 则

$$(A+B)'=A'+B', \quad \overline{A+B}=\overline{A}+\overline{B}. \tag{3.1.7}$$

在引进负矩阵的概念后, 可以引进加法的逆运算 —— **减法**.

定义 3.1.3 两个 $n\times m$ 矩阵的**差**是一个 $n\times m$ 矩阵, 它定义为

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} - \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11}-b_{11} & \cdots & a_{1m}-b_{1m} \\ \vdots & & \vdots \\ a_{n1}-b_{n1} & \cdots & a_{nm}-b_{nm} \end{pmatrix}. \tag{3.1.8}$$

矩阵的减法和实数的减法的性质完全相同.

定义 3.1.4 域 $\mathbb{F}$ 中数 c 和 $n\times m$ 矩阵 A 的**纯量积**是一个 $n\times m$ 矩阵 B, 它定义为

$$B = cA = c\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1m} \\ \vdots & & \vdots \\ ca_{n1} & \cdots & ca_{nm} \end{pmatrix}. \tag{3.1.9}$$

纯量积有下面一些性质: 任取域 $\mathbb{F}$ 中数 c 和 d, 任取 $n\times m$ 矩阵 A 和 B, 则有

(1) $(c+d)A = cA + dA$;

(2) $c(A+B) = cA + cB$;

(3) $c(dA) = (cd)A$;

(4) $1\cdot A = A$, $(-1)A = -A$;

(5) $cA = 0$ 当且仅当 $c = 0$ 或 $A = 0$;

(6) $(cA)' = cA'$, $\overline{cA} = \overline{c}\,\overline{A}$.

现在引进 nm 个 $n\times m$ 矩阵 E_{jk}, 其中 E_{jk} 是由第 j 行, 第 k 列交叉元素为 1, 其他元素为零构成的方阵. (今后约定, 在矩阵中, 凡是空白处没有写出的元素都是零.) 利用矩阵的加法和纯量积的概念, 可知任一 $n\times m$ 矩阵 A 能唯一地表成:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \sum_{j=1}^{n}\sum_{k=1}^{m} a_{jk}E_{jk}. \tag{3.1.10}$$

所以矩阵的相加和纯量积可以写成: 记 $A = \sum\limits_{j=1}^{n}\sum\limits_{k=1}^{m} a_{jk}E_{jk}$, $B = \sum\limits_{j=1}^{n}\sum\limits_{k=1}^{m} b_{jk}E_{jk}$ 是 $n\times m$ 矩阵, 则

$$\begin{aligned} A+B &= \sum_{j=1}^{n}\sum_{k=1}^{m} a_{jk}E_{jk} + \sum_{j=1}^{n}\sum_{k=1}^{m} b_{jk}E_{jk} = \sum_{j=1}^{n}\sum_{k=1}^{m} (a_{jk}+b_{jk})E_{jk}, \\ cA &= \sum_{j=1}^{n}\sum_{k=1}^{m} (ca_{jk})E_{jk}. \end{aligned} \tag{3.1.11}$$

定义 3.1.5 给定域 $\mathbb{F}$ 上 $n\times m$ 矩阵 A, $m\times p$ 矩阵 B 和 $n\times p$ 矩阵 C:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mp} \end{pmatrix}, \quad C=\begin{pmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{np} \end{pmatrix}. \tag{3.1.12}$$

矩阵 A 和 B 的**乘积**是 $n\times p$ 矩阵 $C=AB$, 它定义为

$$c_{ik}=\sum_{j=1}^{m}a_{ij}b_{jk}, \quad i=1,2,\cdots,n, \quad k=1,2,\cdots,p. \tag{3.1.13}$$

相对于矩阵的乘积 AB, A 和 B 称为乘积 AB 的**因子**. 由矩阵 A 和 B 的乘积 AB 的定义可知, 只有域 $\mathbb{F}$ 上 $n\times m$ 矩阵和 $m\times p$ 矩阵 B 才能相乘, 乘积为 $n\times p$ 矩阵. 又乘积 AB 的第 i 行, 第 k 列交叉位置的元素是矩阵 A 的第 i 行和矩阵 B 的第 k 列的相应位置元素相乘, 然后全体加起来所得的值. 所以, 只在前一个因子 A 的列数和后一个因子 B 的行数相同时, 这两个矩阵才能相乘. 而且, 前一个因子 A 的第 i 行元素出现且只出现在乘积 AB 的第 i 行中, 后一个因子 B 的第 k 列元素出现且只出现在乘积 AB 的第 k 列中.

例 3.1.1

$$\begin{pmatrix} 1 & 3 & 2 & 5 \\ 4 & 8 & 7 & 0 \\ 3 & 2 & 1 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 & 1 \\ 1 & 3 & 1 & 0 \\ 1 & 4 & 0 & 1 \\ 1 & 5 & 1 & 0 \end{pmatrix}=\begin{pmatrix} 11 & 44 & 8 & 3 \\ 19 & 60 & 8 & 11 \\ 11 & 41 & 7 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 1\,3\,8\,2 \end{pmatrix}\begin{pmatrix} 7 \\ 0 \\ 4 \\ 0 \end{pmatrix}=39, \quad \begin{pmatrix} 0\,2\,1\,1 \\ -1\,1\,1\,1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \\ 2 \\ 4 \end{pmatrix}=\begin{pmatrix} 8 \\ 4 \end{pmatrix}.$$

下面从另一个角度写出矩阵的乘积. 利用直接计算可知 $n\times m$ 矩阵 $E_{ij}^{(n,m)}$ 和 $m\times p$ 矩阵 $E_{kq}^{(m,p)}$ 的乘积为

$$E_{ij}^{(n,m)}E_{kq}^{(m,p)}=\delta_{jk}E_{iq}^{(n,p)}, \quad 1\leqslant i\leqslant n, \quad 1\leqslant j,k\leqslant m, \quad 1\leqslant q\leqslant p. \tag{3.1.14}$$

所以记

$$A=\sum_{i=1}^{n}\sum_{j=1}^{m}a_{ij}E_{ij}^{(n,m)}, \quad B=\sum_{k=1}^{m}\sum_{q=1}^{p}b_{kq}E_{kq}^{(m,p)},$$

则

$$\begin{aligned}AB &= \Big(\sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}E_{ij}^{(n,m)}\Big)\Big(\sum_{k=1}^{m}\sum_{q=1}^{p} b_{kq}E_{kq}^{(m,p)}\Big)\\&= \sum_{i=1}^{n}\sum_{j,k=1}^{m}\sum_{q=1}^{p}(a_{ij}b_{kq})E_{ij}^{(n,m)}E_{kq}^{(m,p)} = \sum_{i=1}^{n}\sum_{j,k=1}^{m}\sum_{q=1}^{p}(\delta_{jk}a_{ij}b_{kq})E_{iq}^{(n,p)}\\&= \sum_{i=1}^{n}\sum_{q=1}^{p}\Big(\sum_{j=1}^{m}a_{ij}b_{jq}\Big)E_{iq}^{(n,p)} = \sum_{i=1}^{n}\sum_{q=1}^{p}c_{iq}E_{iq}^{(n,p)}.\end{aligned}$$

即 AB 的第 j 行, 第 p 列交叉位置的元素为 $c_{iq}=\sum\limits_{j=1}^{m}a_{ij}b_{jq}$. 这就是矩阵乘积的定义.

关于矩阵乘积, 有下面一些重要性质:

(一) 乘法结合律: 设 A, B, C 分别为 $n\times m$, $m\times p$, $p\times q$ 矩阵, 则

$$(AB)C = A(BC). \tag{3.1.15}$$

证 记

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nm}\end{pmatrix},\quad B=\begin{pmatrix}b_{11}&\cdots&b_{1p}\\\vdots&&\vdots\\b_{m1}&\cdots&b_{mp}\end{pmatrix},\quad C=\begin{pmatrix}c_{11}&\cdots&c_{1q}\\\vdots&&\vdots\\c_{p1}&\cdots&c_{pq}\end{pmatrix}.$$

记 $AB=(u_{ij})$, $BC=(v_{jk})$, 则

$$u_{ij}=\sum_{t=1}^{m}a_{it}b_{tj},\quad v_{jk}=\sum_{t=1}^{p}a_{jt}b_{tk}.$$

因此 $(AB)C$ 的第 i 行, 第 k 列交叉位置的元素为

$$\sum_{j=1}^{p}u_{ij}c_{jk}=\sum_{j=1}^{p}\sum_{t=1}^{m}a_{it}b_{tj}c_{jk},\quad 1\leqslant i\leqslant n,\quad 1\leqslant k\leqslant q.$$

而 $A(BC)$ 的第 i 行, 第 k 列交叉位置的元素为

$$\sum_{j=1}^{m}a_{ij}v_{jk}=\sum_{j=1}^{m}a_{ij}\sum_{t=1}^{p}b_{jt}c_{tk}=\sum_{t=1}^{m}\sum_{j=1}^{p}a_{it}b_{tj}c_{jk},\quad 1\leqslant i\leqslant n,\quad 1\leqslant k\leqslant q.$$

所以 $(AB)C$ 和 $A(BC)$ 的第 i 行, 第 k 列交叉位置的元素相等, 由矩阵相等的定义可知 $A(BC)=(AB)C$. □

由归纳法不难证明: 设 $A_1, A_2, \cdots, A_s$ 分别为 $n_0\times n_1$, $n_1\times n_2$, $\cdots$, $n_{s-1}\times n_s$ 矩阵, 则这 s 个矩阵可以连乘, 而且连乘时可以不计次序先后, 但是要计位置前后. 即乘积可以写成

$$A_0=A_1A_2\cdots A_s,$$

它是 $n_0 \times n_s$ 矩阵. 不难看出, 记 $A_i = \left(a_{jk}^{(i)}\right)$, $i = 0, 1, 2, \cdots, s$, 则

$$a_{ik}^{(0)} = \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_s=1}^{n_s} a_{ij_1}^{(1)} a_{j_1 j_2}^{(2)} \cdots a_{j_s k}^{(s)},$$

其中 $i = 1, 2, \cdots, n_0$, $k = 1, 2, \cdots, n_s$.

(二) n 阶方阵

$$E = E_n = E^{(n)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{3.1.16}$$

称为 n 阶**单位方阵**. 它的对角元素全是 1, 其他元素全是零. 今后, 总用大写拉丁字母 $E = E_n$ 表示单位方阵. 显然, 对任一 $n \times m$ 矩阵 A, 则有等式

$$E_n A = A E_m = A. \tag{3.1.17}$$

(三) **加乘分配律**: 设 A 和 B 都是 $n \times m$ 矩阵, 又设 C 和 D 分别为 $m \times p$ 矩阵和 $q \times n$ 矩阵, 则

$$(A + B)C = AC + BC, \quad D(A + B) = DA + DB. \tag{3.1.18}$$

(四) 设 $c \in \mathbb{F}$ 是数, A 和 B 分别为 $n \times m$ 矩阵及 $m \times p$ 矩阵, 则有

$$(cA)B = c(AB) = A(cB). \tag{3.1.19}$$

最后需要指出一点: 一般说来, 矩阵的乘法是**不可交换的**. 即矩阵的乘法不适合交换律. 事实上, 设 A 是 $n \times m$ 矩阵, B 是 $m \times p$ 矩阵, 则 AB 是 $n \times p$ 矩阵. 在 $p \neq n$ 时 B 和 A 不能相乘, 更谈不上 $BA = AB$, 所以这时乘法不可交换. 即使在 $p = n$ 时, 则 AB 是 n 阶方阵, 然而 BA 是 m 阶方阵, 所以在 $m \neq n$ 时 $AB \neq BA$, 即这时乘法也不可交换. 由此可见, 只有在 $n = m = p$ 时, AB 和 BA 的阶才相同. 但是在方阵的情形, 乘法也不一定可交换.

例 3.1.2 取

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

则

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

按相等的定义, $BA \neq AB$.

由于矩阵乘法的不可交换性, 所以在矩阵相乘时, 前后次序不能随便调换.

(五) 一般说来, 矩阵的乘法不适合**乘法消去律**. 即存在 $n \times m$ 矩阵 $A \neq 0$ 以及 $m \times p$ 矩阵 $B \neq 0$, 但是 $AB = 0$. (上面例 3.1.2.)

(六) 在前面已经定义了矩阵的转置矩阵和共轭矩阵的概念. 实际上, 这些也可以看作是施行在矩阵上的两种运算, 而转置运算在矩阵论中起着很大的作用.

设 A 和 B 分别为 $n \times m$ 矩阵及 $m \times p$ 矩阵, 从定义出发不难证明:

$$(AB)' = B'A', \quad \overline{AB} = \overline{A}\,\overline{B}. \tag{3.1.20}$$

上面这个性质特别重要. 由归纳法不难证明: 设 $A_1, A_2, \cdots, A_s$ 是 s 个可以依次连乘的矩阵, 即 A_{j-1} 的列数和 A_j 的行数相等, 则

$$(A_1A_2\cdots A_s)' = A_s'A_{s-1}'\cdots A_1'. \tag{3.1.21}$$

即矩阵乘积的转置矩阵等于每个因子分别转置后, 再将乘积的次序完全颠倒过来.

重要的矩阵是方阵. 记域 $\mathbb{F}$ 上所有 n 阶方阵构成的集合为 $\mathrm{gl}(n,\mathbb{F})$. 由矩阵的代数运算的定义可知: 在 $\mathrm{gl}(n,\mathbb{F})$ 中, 加法、纯量积和乘法都是封闭的. 这就是说: 两个 n 阶方阵的和仍为 n 阶方阵, 一个 n 阶方阵乘域 $\mathbb{F}$ 中一个数, 仍为 n 阶方阵. 两个 n 阶方阵必可相乘, 且乘积仍为 n 阶方阵. 最后, 不难看出, n 阶方阵的转置方阵仍为 n 阶方阵. n 阶方阵的共轭方阵也仍为 n 阶方阵. 取定 n 阶单位方阵 E_n, 于是对任一 n 阶方阵 A, 有

$$E_nA = AE_n = A. \tag{3.1.22}$$

还可以定义一些特殊的方阵, 例如:

(1) 如果 n 阶方阵 A 形如

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \tag{3.1.23}$$

则称 A 是**上三角方阵**. 这时由 Laplace 展开式可知

$$\det(A) = a_{11}a_{22}\cdots a_{nn}. \tag{3.1.24}$$

(2) 如果 n 阶方阵 A 形如

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \tag{3.1.25}$$

则称 A 是**下三角方阵**. 这时同样有

$$\det(A) = a_{11}a_{22}\cdots a_{nn}. \tag{3.1.26}$$

显然, 上 (下) 三角方阵的转置方阵为下 (上) 三角方阵.

(3) 如果 n 阶方阵 A 形如

$$A = \operatorname{diag}(a_{11}, a_{22}, \cdots, a_{nn}) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \tag{3.1.27}$$

则 A 称为**对角方阵**.

下面定义的对称方阵、反对称方阵、Hermite 方阵、反 Hermite 方阵都是以后常常会见到的重要方阵. 设 $i, j \in \{1, 2, \cdots, n\}$.

(4) 如果 n 阶方阵 A 适合 $A' = A$, 则 A 称为**对称方阵**. 这时 $a_{ij} = a_{ji}$;

(5) 如果 n 阶方阵 A 适合 $A' = -A$, 则 A 称为**反对称方阵**. 这时 $a_{ij} = -a_{ji}$;

(6) 如果 n 阶复方阵 A 适合 $\overline{A'} = A^* = A$, 则 A 称为**Hermite 方阵**. 这时 $\overline{a_{ij}} = a_{ji}$;

(7) 如果 n 阶复方阵 A 适合 $\overline{A'} = A^* = -A$, 则 A 称为**反 Hermite 方阵**. 这时 $\overline{a_{ij}} = -a_{ji}$.

在方阵的情形, 还可以引进方阵多项式的概念. 任取 n 阶方阵 A, 记

$$A^0 = E^{(n)}, \quad A^k = A^{k-1}A = \underbrace{AA\cdots A}_{k\text{个}}, \quad k = 1, 2, \cdots. \tag{3.1.28}$$

A^k 称为方阵 A 的 k 次**方幂**. 显然, 它们仍为 n 阶方阵. 再任取 $p+1$ 个数 $a_p \neq 0$, $a_{p-1}, \cdots, a_0$, 显然

$$f(A) = a_pA^p + a_{p-1}A^{p-1} + \cdots + a_0E_n \tag{3.1.29}$$

仍为 n 阶方阵.

今取 n^2 个独立未知量 x_{jk}, $j, k = 1, 2, \cdots, n$ 构作一个 n 阶方阵

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}. \tag{3.1.30}$$

对域 $\mathbb{F}$ 中任意 $p+1$ 个数 $a_p \neq 0, a_{p-1}, \cdots, a_0$, 则

$$p(X) = a_pX^p + a_{p-1}X^{p-1} + \cdots + a_0E_n = \sum_{j=0}^{p} a_jX^j \tag{3.1.31}$$

称为方阵 X 在域 $\mathbb{F}$ 上的 p 次**方阵多项式**. 显然, 在 1 阶方阵时, 就是普通的多项式概念.

n 阶方阵 A 称为方阵多项式 $p(X)$ 的**根**, 如果 $p(A)=0$. 即

$$p(A)=a_pA^p+a_{p-1}A^{p-1}+\cdots+a_0E_n=0, \tag{3.1.32}$$

其中 0 为 n 阶零方阵, E_n 为 n 阶单位方阵.

方阵多项式和普通多项式有些相同的性质, 下面给出两个重要的性质:

(1) 设 X 为 n 阶方阵, $p(X)$ 和 $q(X)$ 分别为 X 的 p 及 q 次方阵多项式:

$$\begin{aligned} p(X)&=a_pX^p+a_{p-1}X^{p-1}+\cdots+a_0E_n=\sum_{j=0}^{p}a_jX^j,\\ q(X)&=b_qX^q+b_{q-1}X^{q-1}+\cdots+b_0E_n=\sum_{k=0}^{q}b_kX^k, \end{aligned} \tag{3.1.33}$$

则 $p(X)q(X)$ 是 $p+q$ 次多项式:

$$\begin{aligned} p(X)q(X)&=\Big(\sum_{j=0}^{p}a_jX^j\Big)\Big(\sum_{k=0}^{q}b_kX^k\Big)=\sum_{j=0}^{p}\sum_{k=0}^{q}a_jb_kX^{j+k}\\ &=\sum_{i=0}^{p+q}\Big(\sum_{j+k=i}a_jb_k\Big)X^i. \end{aligned} \tag{3.1.34}$$

所以, 记 $p(X)q(X)=r(X)=\sum\limits_{i=0}^{p+q}c_iX^i$, 则

$$c_i=\sum_{j+k=i}a_jb_k,\quad 0\leqslant j\leqslant p,\quad 0\leqslant k\leqslant q.$$

由此可知:

(2) 设 X 为 n 阶方阵, 任两 X 的方阵多项式 $p(X)$ 及 $q(X)$ 的乘积是可交换的, 即

$$q(X)p(X)=p(X)q(X). \tag{3.1.35}$$

除了方阵多项式的概念外, 关于方阵, 还有两个重要的概念: 一个就是方阵的行列式的概念, 这在第二章中已详细讨论过; 另一个就是方阵的迹的概念. 它们都是方阵的函数.

定义 3.1.6 方阵 A 的 n 个对角元素的和称为方阵 A 的**迹**, 记作 $\operatorname{tr}(A)$. 即设

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}, \tag{3.1.36}$$

由定义可知

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{j=1}^{n} a_{jj}. \tag{3.1.37}$$

从定义出发, 不难证明下列基本性质:

(1) 任取两个 n 阶方阵 A 和 B, 则

$$\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B); \tag{3.1.38}$$

(2) 任取一个 n 阶方阵 A 和一个域 $\mathbb{F}$ 中的数 c, 则

$$\operatorname{tr}(cA) = c\operatorname{tr}(A); \tag{3.1.39}$$

(3) 任取两个 n 阶方阵 A 和 B, 则

$$\operatorname{tr}(AB) = \operatorname{tr}(BA); \tag{3.1.40}$$

(4) 任取一个 n 阶方阵 A, 则

$$\operatorname{tr}(A') = \operatorname{tr}(A), \quad \operatorname{tr}(\overline{A}) = \overline{\operatorname{tr}(A)}, \quad \operatorname{tr}(A^*) = \overline{\operatorname{tr}(A)}; \tag{3.1.41}$$

(5) $\operatorname{tr}(AA^*) = 0$ 当且仅当 $A = 0$.

例 3.1.3 设 A 为域 $\mathbb{F}$ 上 n 阶方阵, 如果 $A^2 = AA^*$, 则 $A = A^*$.

证 为了证 $A^* = A$, 只要证 $K = A - A^* = 0$ 就可以了. 由基本性质(5), 我们来证 $\operatorname{tr}(KK^*) = 0$.

利用迹的定义可知,

$$\begin{aligned}\operatorname{tr}(KK^*) &= \operatorname{tr}(A-A^*)(A-A^*)^* = \operatorname{tr}(AA^*) - \operatorname{tr}(A^2) - \operatorname{tr}(A^*)^2 + \operatorname{tr}(A^*A)\\ &= 2\operatorname{tr}(AA^* - A^2) = 0,\end{aligned}$$

于是 $A = A^*$. □

注意: 这是一道不容易做出来的题. 原因在于不会想到用迹的定义, 只会想到用方程组来求解.

最后给出矩阵计算的重要规则. 即在具体对矩阵作代数运算时, 往往整齐地把矩阵分成若干个小的子矩阵块. 对矩阵作了适当的分块后, 在进行代数运算时, 就可以把每一小块看做一个元素, 再作同样的代数运算.

进行**矩阵分块**的方式是用水平线和垂直线把矩阵分隔成一些子矩阵的小块. 例

如, 设 A 是 $n\times m$ 矩阵, 作分块为

$$\begin{array}{cc}\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix} & \begin{matrix} n_1\text{ 行} \\ n_2\text{ 行} \\ \vdots \\ n_p\text{ 行}\end{matrix}, \\ m_1\text{ 列}\ m_2\text{ 列}\cdots\ m_q\text{列} & \end{array} \tag{3.1.42}$$

其中 $n_1+n_2+\cdots+n_p=n$, $m_1+m_2+\cdots+m_q=m$, 又 A_{jk} 是 $n_j\times m_k$ 矩阵, 称为 A 的**子矩阵**.

(1) 加法: 设 A 和 B 都是 $n\times m$ 矩阵. 将 A 和 B 按照同样的规则进行分块, 即取

$$A=\begin{pmatrix} A_{11}\cdots A_{1q} \\ \vdots \quad\quad \vdots \\ A_{p1}\cdots A_{pq}\end{pmatrix},\quad B=\begin{pmatrix} B_{11}\cdots B_{1q} \\ \vdots \quad\quad \vdots \\ B_{p1}\cdots B_{pq}\end{pmatrix}, \tag{3.1.43}$$

其中 A_{jk} 和 B_{jk} 都是 $n_j\times m_k$ 矩阵. 显然有

$$\begin{pmatrix} A_{11}\cdots A_{1q} \\ \vdots \quad\quad \vdots \\ A_{p1}\cdots A_{pq}\end{pmatrix}+\begin{pmatrix} B_{11}\cdots B_{1q} \\ \vdots \quad\quad \vdots \\ B_{p1}\cdots B_{pq}\end{pmatrix}=\begin{pmatrix} A_{11}+B_{11}\cdots A_{1q}+B_{1q} \\ \vdots \quad\quad\quad\quad \vdots \\ A_{p1}+B_{p1}\cdots A_{pq}+B_{pq}\end{pmatrix}. \tag{3.1.44}$$

所以在作加法时, 可以先把每一小块看作一个元素作加法, 再把和的每一小块看作是两个矩阵的和.

(2) 纯量积: 设 c 为数, A 为 $n\times m$ 矩阵. 将 A 分块后, 显然有

$$cA=c\begin{pmatrix} A_{11}\cdots A_{1q} \\ \vdots \quad\quad \vdots \\ A_{p1}\cdots A_{pq}\end{pmatrix}=\begin{pmatrix} cA_{11}\cdots cA_{1q} \\ \vdots \quad\quad \vdots \\ cA_{p1}\cdots cA_{pq}\end{pmatrix}. \tag{3.1.45}$$

(3) 乘积:

引理 3.1.7 设 A 和 B 分别为 $n\times m$, $m\times s$ 矩阵. 将它们分块成

$$A=\begin{pmatrix} A_{11}\cdots A_{1q} \\ \vdots \quad\quad \vdots \\ A_{p1}\cdots A_{pq}\end{pmatrix},\quad B=\begin{pmatrix} B_{11}\cdots B_{1r} \\ \vdots \quad\quad \vdots \\ B_{q1}\cdots B_{qr}\end{pmatrix}, \tag{3.1.46}$$

其中 A_{ij} 的列数和 B_{jk} 的行数相同. 换句话说, 将对 A 的列的分块方式施行在 B 的

行上. 这样一来, 由乘法定义可知 A_{ij} 和 B_{jk} 可以相乘, 且乘积 AB 为

$$\begin{pmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{pmatrix}\begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{q1} & \cdots & B_{qr} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{q} A_{1j}B_{j1} & \cdots & \sum_{j=1}^{q} A_{1j}B_{jr} \\ \vdots & & \vdots \\ \sum_{j=1}^{q} A_{pj}B_{j1} & \cdots & \sum_{j=1}^{q} A_{pj}B_{jr} \end{pmatrix}. \tag{3.1.47}$$

证 先来考查矩阵乘积的定义. 令乘积

$$AB = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}\begin{pmatrix} b_{11} & \cdots & b_{1s} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{ms} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{m} a_{1j}b_{j1} & \cdots & \sum_{j=1}^{m} a_{1j}b_{js} \\ \vdots & & \vdots \\ \sum_{j=1}^{m} a_{nj}b_{j1} & \cdots & \sum_{j=1}^{m} a_{nj}b_{js} \end{pmatrix}$$

可以改写为

$$AB = \begin{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1m} \end{pmatrix} B \\ \vdots \\ \begin{pmatrix} a_{n1} & \cdots & a_{nm} \end{pmatrix} B \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pm} \end{pmatrix} B \\ \\ \begin{pmatrix} a_{p+1,1} & \cdots & a_{p+1,m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} B \end{pmatrix}$$

等, 同样可以改写为

$$\begin{aligned} AB &= \begin{pmatrix} A\begin{pmatrix} b_{11} \\ \vdots \\ b_{m1} \end{pmatrix} & \cdots & A\begin{pmatrix} b_{1s} \\ \vdots \\ b_{ms} \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} A\begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mq} \end{pmatrix} & A\begin{pmatrix} b_{1,q+1} & \cdots & b_{1s} \\ \vdots & & \vdots \\ b_{m,q+1} & \cdots & b_{ms} \end{pmatrix} \end{pmatrix} \end{aligned}$$

等. 所以, 一般对矩阵 A 的行作分块

$$A = \begin{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \end{pmatrix} \\ \begin{pmatrix} A_{21} & A_{22} & \cdots & A_{2q} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix} \end{pmatrix},$$

再对 B 的列作分块

$$B=\left(\begin{pmatrix}B_{11}\\B_{21}\\\vdots\\B_{q1}\end{pmatrix}\begin{pmatrix}B_{12}\\B_{22}\\\vdots\\B_{q2}\end{pmatrix}\cdots\begin{pmatrix}B_{1r}\\B_{2r}\\\vdots\\B_{qr}\end{pmatrix}\right),$$

则由

$$AB=\begin{pmatrix}\begin{pmatrix}A_{11}\cdots A_{1q}\end{pmatrix}B\\\begin{pmatrix}A_{21}\cdots A_{2q}\end{pmatrix}B\\\vdots\\\begin{pmatrix}A_{p1}\cdots A_{pq}\end{pmatrix}B\end{pmatrix},$$

有

$$AB=\begin{pmatrix}\begin{pmatrix}A_{11}\cdots A_{1q}\end{pmatrix}\begin{pmatrix}B_{11}\\\vdots\\B_{q1}\end{pmatrix}&\cdots&\begin{pmatrix}A_{11}\cdots A_{1q}\end{pmatrix}\begin{pmatrix}B_{1r}\\\vdots\\B_{qr}\end{pmatrix}\\\vdots&&\vdots\\\begin{pmatrix}A_{p1}\cdots A_{pq}\end{pmatrix}\begin{pmatrix}B_{11}\\\vdots\\B_{q1}\end{pmatrix}&\cdots&\begin{pmatrix}A_{p1}\cdots A_{pq}\end{pmatrix}\begin{pmatrix}B_{1r}\\\vdots\\B_{qr}\end{pmatrix}\end{pmatrix}.$$

因此问题化为证明

$$\begin{pmatrix}A_{i1}\cdots A_{iq}\end{pmatrix}\begin{pmatrix}B_{1k}\\\vdots\\B_{qk}\end{pmatrix}=\sum_{j=1}^{q}A_{ij}B_{jk}\tag{3.1.48}$$

成立就行了. 所以问题化为只要证明 $q=2$ 的情形就行了. 事实上, 设

$$\begin{pmatrix}C^{(n,u)}&D^{(n,v)}\end{pmatrix}\begin{pmatrix}H^{(u,m)}\\G^{(v,m)}\end{pmatrix}=CH+DG\tag{3.1.49}$$

成立, 记

$$(A_{i1}\cdots A_{iq})\begin{pmatrix}B_{1k}\\\vdots\\B_{qk}\end{pmatrix}=\Big((A_{i1}\cdots A_{i,q-1})\ A_{iq}\Big)\begin{pmatrix}\begin{pmatrix}B_{1k}\\\vdots\\B_{q-1,k}\end{pmatrix}\\\\B_{qk}\end{pmatrix},$$

我们有

$$(A_{i1}\cdots A_{iq})\begin{pmatrix}B_{1k}\\ \vdots\\ B_{qk}\end{pmatrix}=(A_{i1}\cdots A_{i,q-1})\begin{pmatrix}B_{1k}\\ \vdots\\ B_{q-1,k}\end{pmatrix}+A_{iq}B_{qk}.$$

用数学归纳法可知,

$$(A_{i1}\cdots A_{iq})\begin{pmatrix}B_{1k}\\ \vdots\\ B_{qk}\end{pmatrix}=\sum_{j=1}^{q}A_{ij}B_{jk}.$$

下面证明式 (3.1.49) 成立. 事实上, 记 C 及 D 的第 r 行分别为: $(c_{r1},\cdots,c_{ru})$ 和 $(d_{r1},\cdots,d_{rv})$. 记 H 及 G 的第 s 列分别为 $(h_{1s},\cdots,h_{us})'$ 和 $(g_{1s},\cdots,g_{vs})'$. 所以 CH 及 DG 的第 r 行、第 s 列元素分别为 $\sum\limits_{t=1}^{u}c_{rt}h_{ts}$ 及 $\sum\limits_{t=1}^{v}d_{rt}g_{ts}$. 因此 $CH+DG$ 的第 r 行、第 s 列元素为 $\sum\limits_{t=1}^{u}c_{rt}h_{ts}+\sum\limits_{t=1}^{v}d_{rt}g_{ts}$.

另一方面, 矩阵 $\begin{pmatrix}C^{(n,u)} & D^{(n,v)}\end{pmatrix}$ 的第 r 行为 $(c_{r1},\cdots,c_{ru},d_{r1},\cdots,d_{rv})$, 矩阵 $\begin{pmatrix}H\\ G\end{pmatrix}$ 的第 s 列为 $(h_{1s},\cdots,h_{us},g_{1s},\cdots,g_{vs})'$, 所以 $\begin{pmatrix}C^{(n,u)} & D^{(n,v)}\end{pmatrix}\begin{pmatrix}H\\ G\end{pmatrix}$ 的第 r 行、第 s 列元素也是 $\sum\limits_{t=1}^{u}c_{rt}h_{ts}+\sum\limits_{t=1}^{v}d_{rt}g_{ts}$. 这就证明了

$$\begin{pmatrix}C^{(n,u)} & D^{(n,v)}\end{pmatrix}\begin{pmatrix}H\\ G\end{pmatrix}=CH+DG. \qquad \square$$

(4) **转置**: 将 $n\times m$ 矩阵 A 作分块, 则

$$A'=\begin{pmatrix}A_{11}&\cdots&A_{1q}\\ \vdots&&\vdots\\ A_{p1}&\cdots&A_{pq}\end{pmatrix}'=\begin{pmatrix}A'_{11}&\cdots&A'_{p1}\\ \vdots&&\vdots\\ A'_{1q}&\cdots&A'_{pq}\end{pmatrix}. \tag{3.1.50}$$

即分块矩阵的转置, 不仅把每个小块看作元素后对矩阵作转置, 而且每个小块本身也要作转置.

(5) **共轭**: 将 $n\times m$ 矩阵 A 作分块, 则

$$\overline{A}=\overline{\begin{pmatrix}A_{11}&\cdots&A_{1q}\\ \vdots&&\vdots\\ A_{p1}&\cdots&A_{pq}\end{pmatrix}}=\begin{pmatrix}\overline{A_{11}}&\cdots&\overline{A_{1q}}\\ \vdots&&\vdots\\ \overline{A_{p1}}&\cdots&\overline{A_{pq}}\end{pmatrix}. \tag{3.1.51}$$

即分块矩阵的共轭矩阵为每个小块本身作共轭.

利用矩阵分块, 还可以定义一些特殊的矩阵. 例如:

(1) 将矩阵 A 分块后, 如果它形如

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ 0 & A_{22} & \cdots & A_{2p} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{pp} \end{pmatrix}, \tag{3.1.52}$$

则称 A 对此分块法是**准上三角矩阵**. 在 $A_{11}, \cdots, A_{pp}$ 都是方阵时, 显然 A 也是方阵. 这时由 Laplace 展开式可知

$$\det(A)=\det(A_{11})\det(A_{22})\cdots\det(A_{pp}). \tag{3.1.53}$$

(2) 将矩阵 A 分块后, 如果它形如

$$A=\begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pp} \end{pmatrix}, \tag{3.1.54}$$

则称 A 对此分块法是**准下三角矩阵**. 在 $A_{11}, \cdots, A_{pp}$ 都是方阵时, 显然 A 也是方阵. 这时同样有

$$\det(A)=\det(A_{11})\det(A_{22})\cdots\det(A_{pp}). \tag{3.1.55}$$

显然, 准上 (下) 三角矩阵的转置矩阵为准下 (上) 三角矩阵.

(3) 将矩阵 A 分块后, 如果它形如

$$A=\begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{pp} \end{pmatrix}, \tag{3.1.56}$$

则称为 A 对此分块法是**准对角矩阵**. 有时也用符号

$$\operatorname{diag}(A_{11}, A_{22}, \cdots, A_{pp}) \tag{3.1.57}$$

记之. 特别, 在 $A_{11}, A_{22}, \cdots, A_{pp}$ 都是一阶方阵, 即是域 $\mathbb{F}$ 上对角方阵.

下面举一例子, 说明 "准" 的概念和分法有关.

例 3.1.4 将矩阵分块为

$$\begin{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 8 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 0 & 7 \end{pmatrix} \end{pmatrix}.$$

再按下列方式分块:

$$\begin{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 8 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \end{pmatrix} & \begin{pmatrix} 2 \end{pmatrix} & \begin{pmatrix} 1 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 \end{pmatrix} & \begin{pmatrix} 7 \end{pmatrix} \end{pmatrix}, \quad \begin{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} & \begin{pmatrix} 0 \end{pmatrix} & \begin{pmatrix} 0 \end{pmatrix} \\ \begin{pmatrix} 3 & 8 \end{pmatrix} & \begin{pmatrix} 0 \end{pmatrix} & \begin{pmatrix} 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} & \begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix} \end{pmatrix}.$$

则它们分别为准对角矩阵、准上三角矩阵、准下三角矩阵.

由此可见, 在具体运用矩阵分块技巧时, 要充分考虑到在问题中选取何种分法最合适.

习 题 3.1

除非特别申明, 下面都是在域 $\mathbb{F}$ 上讨论.

3.1.1 试证: 对任一数 c 及任一 n 阶方阵 A, 成立着 $\det(cA) = c^n\det(A)$. 它等于 $c\det(A)$ 当且仅当 $c\det(A) = 0$ 或 $c^{n-1} = 1$.

3.1.2 试各举一例说明: 有两个 n 阶方阵 A 和 B, 使得等式**(1)** $AB = BA$; **(2)** $A \neq 0$, $B \neq 0$, $AB = 0$; **(3)** $\det(A+B) = \det(A) + \det(B)$; **(4)** $AB - BA = E_n$ 都不成立, 其中 E_n 为 n 阶单位方阵.

3.1.3 求下列各矩阵的乘积:

(1) $\begin{pmatrix} 2 & 5 & 3 \\ 3 & 1 & 2 \\ 0 & 4 & 1 \end{pmatrix}^4$; **(2)** $\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)^n$;

(3) $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 1 & 5 & 8 \end{pmatrix}\begin{pmatrix} 2 & 1 & 0 \\ 7 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 3 & 3 \\ 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$.

3.1.4 设方阵 A 的元素 a_{jk} 的代数余子式为 A_{jk}. 以 A_{kj} 为第 j 行, 第 k 列元素的 n 阶

方阵

$$A^{(*)} = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}' = \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix} \tag{3.1.58}$$

称为方阵 A 的**经典伴随方阵** 或**伴随方阵**. 注意前面定义了 A^*, 现在的记号是 $A^{(*)}$, 不是 A^*, 即不是 $\overline{A'}$. 试证下面的重要公式:

$$AA^{(*)} = A^{(*)}A = (\det A)E_n, \quad (AB)^{(*)} = B^{(*)}A^{(*)}, \tag{3.1.59}$$

其中 E_n 为 n 阶单位方阵.

3.1.5 设 n 阶方阵 A 和 B 可交换, 即 $AB = BA$. 试证下列等式成立:

(1) $(A+B)^2 = A^2 + 2AB + B^2$; **(2)** $(A+B)^m = \sum\limits_{k=0}^{m} \begin{pmatrix} m \\ k \end{pmatrix} A^k B^{m-k}$;

(3) $A^2 - B^2 = (A-B)(A+B)$; **(4)** $A^3 - B^3 = (A-B)(A^2 + AB + B^2)$.

3.1.6 试求与下列方阵可交换的方阵全体:

(1) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; **(2)** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$; **(3)** $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{pmatrix}$.

3.1.7 设 A 是 n 阶**幂等方阵**, 即 $A^2 = E_n$, 又 $\det(A) > 0$. 试证: $A + E_n$ 为 n 阶非异方阵 (参看下面的定义 3.3.3).

3.1.8 设 A 是 n 阶上三角方阵, 对角元素都是零. 试证: $A^n = 0$, 即 A 为 n 阶**幂零方阵**;

3.1.9 设 $A_1, A_2, \cdots, A_n$ 都是对角元素为零的 n 阶上三角方阵, 试证 $A_1A_2\cdots A_n = 0$.

3.1.10 试求矩阵方程 $X^2 = \pm E_2$ 的所有 2 阶实方阵解. 由此可知, 虽然不存在实数 x, 使 $x^2 = -1$, 但是存在实数方阵 X, 使 $X^2 = -E_2$.

3.1.11 试证: 对任一二阶方阵 A, 有 $A^2 - (\operatorname{tr}(A))A + (\det(A))E_2 = 0$.

3.1.12 给定 s 个不同数 $\lambda_1, \lambda_2, \cdots, \lambda_s$. 记 $E_1, E_2, \cdots, E_s$ 分别为 $n_1, n_2, \cdots, n_s$ 阶单位方阵. 记 $n = \sum\limits_{j=1}^{s} n_j$. 试证: 与 n 阶对角方阵 $\operatorname{diag}(\lambda_1E_1, \cdots, \lambda_sE_s)$ 可交换的 n 阶方阵必为 $\operatorname{diag}(B_1, \cdots, B_s)$, 其中 $B_1, \cdots, B_s$ 分别为 $n_1, n_2, \cdots, n_s$ 阶方阵.

3.1.13 试证: 与任意 n 阶方阵可交换的 n 阶方阵必为 cE_n, 其中 c 为任一常数, E_n 为 n 阶单位方阵, 这类方阵称为**纯量方阵**.

3.1.14 设 $f(A)$ 为 n 阶方阵 A 的函数, 即 f: $\mathrm{gl}(n, \mathbb{F}) \to \mathbb{F}$, 它有 $f(AB) = f(BA)$, 而且 $f(E_n) = n$. 试证: $f(A) = \operatorname{tr}(A)$. (这也是 n 阶方阵的迹的等价定义.)

3.1.15 设 A 和 B 为 n 阶实方阵. 试证:

$$\det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det(A+B)(A-B), \quad \det \begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \mid \det(A + \sqrt{-1}B) \mid^2 .$$

3.1.16 设 A 为 $n \times m$ 实矩阵, B 为 $n \times (n-m)$ 实矩阵. 试证:

$$\det \begin{pmatrix} A'A & A'B \\ B'A & B'B \end{pmatrix} \leqslant \det(A'A)\det(B'B).$$

3.1.17 求下列各矩阵的乘积:

(1) $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^n$; (2) $\begin{pmatrix} 1 & \dfrac{\lambda}{n} \\ -\dfrac{\lambda}{n} & 1 \end{pmatrix}^n$.

3.1.18 求 $(\lambda E_n + N_n)^m$, 其中 $\lambda \in \mathbb{F}$ 为数, E_n 为 n 阶单位方阵, N_n 为 n 阶幂零方阵, 它定义为 (从第六章起 N_n 就指下面这个 n 阶幂零方阵)

$$N_n = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{pmatrix}. \tag{3.1.60}$$

§3.2 Binet–Cauchy 公式

定理 3.2.1 (Binet-Cauchy 公式) 设 A 和 B 各为 $n \times m$ 和 $m \times n$ 矩阵, 则有

$$\det(AB) = \begin{cases} 0, & \text{当} n > m; \\ \det(A)\det(B), & \text{当} n = m; \\ \displaystyle\sum_{1 \leqslant j_1 < \cdots < j_n \leqslant m} \det A \begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix} \det B \begin{pmatrix} j_1 & j_2 & \cdots & j_n \\ 1 & 2 & \cdots & n \end{pmatrix}, & \text{当} n < m. \end{cases} \tag{3.2.1}$$

证 记 $A = (a_{jk})$ 和 $B = (b_{kj})$ 分别为 $n \times m$ 和 $m \times n$ 矩阵, 则

$$\begin{aligned} \det(AB) &= \det \begin{pmatrix} \displaystyle\sum_{k_1=1}^m a_{1k_1} b_{k_1 1} & \cdots & \displaystyle\sum_{k_n=1}^m a_{1k_n} b_{k_n n} \\ \vdots & & \vdots \\ \displaystyle\sum_{k_1=1}^m a_{nk_1} b_{k_1 1} & \cdots & \displaystyle\sum_{k_n=1}^m a_{nk_n} b_{k_n n} \end{pmatrix} \\ &= \sum_{k_1, k_2, \cdots, k_n = 1}^m \det \begin{pmatrix} a_{1k_1} b_{k_1 1} & \cdots & a_{1k_n} b_{k_n n} \\ \vdots & & \vdots \\ a_{nk_1} b_{k_1 1} & \cdots & a_{nk_n} b_{k_n n} \end{pmatrix} \\ &= \sum_{k_1, k_2, \cdots, k_n = 1}^m b_{k_1 1} b_{k_2 2} \cdots b_{k_n n} \det \begin{pmatrix} a_{1k_1} & \cdots & a_{1k_n} \\ \vdots & & \vdots \\ a_{nk_1} & \cdots & a_{nk_n} \end{pmatrix}. \end{aligned}$$

今若 $k_1, k_2, \cdots, k_n$ 中有两个指标相等, 那末行列式

$$\det\begin{pmatrix} a_{1k_1} & \cdots & a_{1k_n} \\ \vdots & & \vdots \\ a_{nk_1} & \cdots & a_{nk_n} \end{pmatrix}$$

中有两列相等, 所以它等于零. 因此 $\det(AB)$ 等于

$$\sum_{\substack{1\leqslant k_1,k_2,\cdots,k_n\leqslant m\\ k_1,k_2,\cdots,k_n\ \text{两两不等}}} b_{k_11}b_{k_22}\cdots b_{k_nn}\det\begin{pmatrix} a_{1k_1} & \cdots & a_{1k_n} \\ \vdots & & \vdots \\ a_{nk_1} & \cdots & a_{nk_n} \end{pmatrix}.$$

由和号可知, 指标 $k_1, k_2, \cdots, k_n$ 是集合 $\{1, 2, \cdots, m\}$ 中 n 个不同的元素.

设 $n > m$. 由于这时不存在 n 个指标 $k_1, k_2, \cdots, k_n$ 不同, 因此这时 $\det(AB) = 0$.

在 $n \leqslant m$ 时, 由式 (2.2.1) 可知

$$\sum_{\substack{1\leqslant k_1,k_2,\cdots,k_n\leqslant m\\ k_1,k_2,\cdots,k_n\ \text{两两不等}}} = \sum_{1\leqslant j_1<\cdots<j_n\leqslant m}\ \sum_{\binom{j_1j_2\cdots j_n}{k_1k_2\cdots k_n}}.$$

所以

$$\det(AB) = \sum_{1\leqslant j_1<\cdots<j_n\leqslant m}\ \sum_{\binom{j_1j_2\cdots j_n}{k_1k_2\cdots k_n}} b_{k_11}b_{k_22}\cdots b_{k_nn}\det\begin{pmatrix} a_{1k_1} & \cdots & a_{1k_n} \\ \vdots & & \vdots \\ a_{nk_1} & \cdots & a_{nk_n} \end{pmatrix}.$$

将 $j_1, j_2, \cdots, j_n$ 的排列 $k_1k_2\cdots k_n$ 经过 s 次对换变为标准排列 $j_1j_2\cdots j_n$, 用同样的对换方式对行列式

$$\det\begin{pmatrix} a_{1k_1} & \cdots & a_{1k_n} \\ \vdots & & \vdots \\ a_{nk_1} & \cdots & a_{nk_n} \end{pmatrix}$$

的列作 s 次相邻两列互换, 所以 $\det(AB)$ 等于

$$\sum_{1\leqslant j_1<\cdots<j_n\leqslant m}\Bigg(\sum_{\binom{j_1j_2\cdots j_n}{k_1k_2\cdots k_n}} \delta^{j_1j_2\cdots j_n}_{k_1k_2\cdots k_n} b_{k_11}b_{k_22}\cdots b_{k_nn}\Bigg)\det\begin{pmatrix} a_{1j_1} & \cdots & a_{1j_n} \\ \vdots & & \vdots \\ a_{nj_1} & \cdots & a_{nj_n} \end{pmatrix}$$

$$= \sum_{1\leqslant j_1<\cdots<j_n\leqslant m}\det\begin{pmatrix} a_{1j_1} & \cdots & a_{1j_n} \\ \vdots & & \vdots \\ a_{nj_1} & \cdots & a_{nj_n} \end{pmatrix}\det\begin{pmatrix} b_{j_11} & \cdots & b_{j_n1} \\ \vdots & & \vdots \\ b_{j_1n} & \cdots & b_{j_nn} \end{pmatrix}.$$

设 $n=m$, 这时 $1\leqslant j_1<\cdots<j_n\leqslant n$. 所以 $j_1=1,\cdots,j_n=n$. 因此 $\det(AB)=\det(A)\det(B)$.

设 $n<m$, 便得

$$\det AB=\sum_{1\leqslant j_1<\cdots<j_n\leqslant m}\det A\begin{pmatrix}1&2&\cdots&n\\j_1&j_2&\cdots&j_n\end{pmatrix}\det B\begin{pmatrix}j_1&j_2&\cdots&j_n\\1&2&\cdots&n\end{pmatrix}.\qquad\square$$

利用 Binet-Cauchy 定理, 可以计算矩阵乘积的子式, 即有

定理 3.2.2 设 A 和 B 分别为 $n\times m$ 及 $m\times s$ 矩阵, 于是 $n\times s$ 矩阵 $C=AB$ 的任一 r 阶子式为

(1) 当 $m<r$ 时,

$$\det C\begin{pmatrix}j_1j_2&\cdots&j_r\\k_1k_2&\cdots&k_r\end{pmatrix}=0;\tag{3.2.2}$$

(2) 当 $m\geqslant r$ 时, $\det C\begin{pmatrix}j_1j_2&\cdots&j_r\\k_1k_2&\cdots&k_r\end{pmatrix}$ 等于

$$\sum_{1\leqslant i_1<\cdots<i_r\leqslant m}\det A\begin{pmatrix}j_1j_2&\cdots&j_r\\i_1i_2&\cdots&i_r\end{pmatrix}\det B\begin{pmatrix}i_1i_2&\cdots&i_r\\k_1k_2&\cdots&k_r\end{pmatrix}.\tag{3.2.3}$$

特别, 当 $r=m$ 时, 有

$$\det C\begin{pmatrix}j_1j_2&\cdots&j_m\\k_1k_2&\cdots&k_m\end{pmatrix}=\det A\begin{pmatrix}j_1j_2&\cdots&j_m\\1\,2&\cdots&m\end{pmatrix}\det B\begin{pmatrix}1\,2&\cdots&m\\k_1k_2&\cdots&k_m\end{pmatrix}.\tag{3.2.4}$$

证 记

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nm}\end{pmatrix},B=\begin{pmatrix}b_{11}&\cdots&b_{1s}\\\vdots&&\vdots\\b_{m1}&\cdots&b_{ms}\end{pmatrix},C=\begin{pmatrix}c_{11}&\cdots&c_{1s}\\\vdots&&\vdots\\c_{n1}&\cdots&c_{ns}\end{pmatrix},$$

则 $c_{jk}=\sum\limits_{i=1}^{m}a_{ji}b_{ik}$, $1\leqslant j\leqslant n$, $1\leqslant k\leqslant m$. 于是子矩阵

$$\begin{aligned}C\begin{pmatrix}j_1&j_2&\cdots&j_r\\k_1&k_2&\cdots&k_r\end{pmatrix}&=\begin{pmatrix}\sum\limits_{i=1}^{m}a_{j_1i}b_{ik_1}&\cdots&\sum\limits_{i=1}^{m}a_{j_1i}b_{ik_r}\\\vdots&&\vdots\\\sum\limits_{i=1}^{m}a_{j_ri}b_{ik_1}&\cdots&\sum\limits_{i=1}^{m}a_{j_ri}b_{ik_r}\end{pmatrix}\\&=\begin{pmatrix}a_{j_11}&\cdots&a_{j_1m}\\\vdots&&\vdots\\a_{j_r1}&\cdots&a_{j_rm}\end{pmatrix}\cdot\begin{pmatrix}b_{1k_1}&\cdots&b_{1k_r}\\\vdots&&\vdots\\b_{mk_1}&\cdots&b_{mk_r}\end{pmatrix},\end{aligned}$$

即是 A 和 B 的子矩阵 $A\begin{pmatrix} j_1\, j_2 \cdots j_r \\ 1\ \ 2 \cdots\ r \end{pmatrix}$ 和 $B\begin{pmatrix} 1\ \ 2 \cdots\ r \\ k_1\, k_2 \cdots k_r \end{pmatrix}$ 的乘积. 所以可用 Binet-Cauchy 公式. 这就证明了定理. □

Binet-Cauchy 公式导出的最重要的事实, 首先是方阵的乘积的行列式等于方阵的行列式的乘积; 其次是能计算乘积的各阶子式. 利用前一事实, 还可以给出一种计算行列式的技巧. 它的原则可以叙述为: 为了求方阵 A 的行列式, 我们设法找出一个易求行列式的方阵 B, 使方阵 $C = AB$ 的行列式也易求出. 利用公式 $\det(C) = \det(A)\det(B)$. 在 $\det(B) \neq 0$ 时, 便得方阵 A 的行列式

$$\det(A) = \frac{\det(C)}{\det(B)}.$$

例 3.2.1 计算 n 阶轮回方阵

$$A = \begin{pmatrix} a_0 & a_1 & a_2 \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & a_1 \cdots & a_{n-3} & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 \cdots & a_{n-4} & a_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & a_3 \cdots & a_{n-1} & a_0 \end{pmatrix}.$$

的行列式.

解 对这个行列式, 用上一章的方法很难求出值来. 现在构作 n 阶 Vandermonde 方阵 $B = V(\omega^0, \cdots, \omega^{n-1})$, 它是

$$B = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{pmatrix},$$

其中

$$\omega = \exp\left(\frac{2i\pi}{n}\right) = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \quad i = \sqrt{-1}.$$

已知 $\det(B) = \prod\limits_{1\leqslant k<j\leqslant n} (\omega^j - \omega^k)$. 显然, 由于 $\omega^0, \omega^1, \cdots, \omega^{n-1}$ 是 n 个两两不等的复数, 所以 $\det(B) \neq 0$.

今由直接计算可知

$$AB=\begin{pmatrix}\sum_{j=0}^{n-1}a_j & \sum_{j=0}^{n-1}a_j\omega^j & \cdots & \sum_{j=0}^{n-1}a_j\omega^{j(n-1)}\\ \sum_{j=0}^{n-1}a_j & \sum_{j=0}^{n-1}a_j\omega^{j+1} & \cdots & \sum_{j=0}^{n-1}a_j\omega^{(j+1)(n-1)}\\ \vdots & \vdots & & \vdots\\ \sum_{j=0}^{n-1}a_j & \sum_{j=0}^{n-1}a_j\omega^{j+n-1} & \cdots & \sum_{j=0}^{n-1}a_j\omega^{(j+n-1)(n-1)}\end{pmatrix}.$$

令

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1},$$

则

$$\det(AB)=\det\begin{pmatrix}f(1) & f(\omega) & \cdots & f(\omega^{n-1})\\ f(1) & \omega f(\omega) & \cdots & \omega^{n-1}f(\omega^{n-1})\\ \vdots & \vdots & & \vdots\\ f(1) & \omega^{n-1}f(\omega) & \cdots & \omega^{(n-1)(n-1)}f(\omega^{n-1})\end{pmatrix}.$$

因此

$$\det(AB)=f(1)f(\omega)\cdots f(\omega^{n-1})\det(B).$$

由于 $\det(AB)=\det(A)\det(B)$, $\det(B)\neq 0$, 所以

$$\det(A)=\prod_{j=0}^{n-1}f(\omega^j)=\prod_{j=0}^{n-1}[a_0+a_1\omega^j+\cdots+a_{n-1}\omega^{(n-1)j}].$$

习 题 3.2

除非特别申明, 下面都是在域 $\mathbb{F}$ 上讨论.

3.2.1 试求 $\det(AB)$, 其中

(1) $A=\begin{pmatrix}2&0&0&1\\1&3&8&2\\5&1&0&0\end{pmatrix}$, $B=\begin{pmatrix}2&3&1\\5&1&1\\7&7&2\\4&2&3\end{pmatrix}$; **(2)** $A=\begin{pmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\end{pmatrix}$, $B=A'$.

3.2.2 试求矩阵 $C=AB$ 的子式 $\det C\begin{pmatrix}1\,3\,5\\2\,4\,6\end{pmatrix}$, 其中

$$A=\begin{pmatrix}0&1&1&1&1&1\\1&0&1&1&1&1\\1&1&0&1&1&1\\1&1&1&0&1&1\\1&1&1&1&0&1\end{pmatrix},\quad B=\begin{pmatrix}1&3&5&7&9&1\\2&4&6&8&10&1\\0&1&1&1&1&2\\1&0&1&1&1&2\\1&1&0&1&1&2\\1&1&1&0&1&2\end{pmatrix}.$$

3.2.3 试利用 Laplace 展开式来证明: 设 A 和 B 为 n 阶方阵, E_n 为 n 阶单位方阵, 则有 $\det\begin{pmatrix}A&0\\-E&B\end{pmatrix}=\det(A)\det(B)$. 再用行列式的性质来证明: $\det\begin{pmatrix}A&0\\-E&B\end{pmatrix}=\det(AB)$. 从而导出部分 Binet–Cauchy 公式: $\det(AB)=\det(A)\det(B)$.

3.2.4 设 $A^{(*)}$ 是 A 的伴随方阵 (定义见习题 3.1.4), 试证 $\det(A^{(*)})=(\det A)^{n-1}$.

3.2.5 设 $A=\begin{pmatrix}a_1\ a_2\cdots a_n\\b_1\ b_2\cdots b_n\end{pmatrix}$ 为 $2\times n$ 实矩阵, 试分别由直接计算及 Binet-Cauchy 公式这两种方法来求 $\det(AA')$, 从而导出等式

$$\Big(\sum_{j=1}^n a_j^2\Big)\Big(\sum_{j=1}^n b_j^2\Big)-\Big(\sum_{j=1}^n a_jb_j\Big)^2=\sum_{1\leqslant j<k\leqslant n}(a_jb_k-a_kb_j)^2.$$

由此可证 Cauchy 不等式

$$\Big(\sum_{j=1}^n a_j^2\Big)\Big(\sum_{j=1}^n b_j^2\Big)\geqslant\Big(\sum_{j=1}^n a_jb_j\Big)^2.$$

又可证上面不等式中等号成立的必要且充分条件为 $a_1:b_1=a_2:b_2=\cdots=a_n:b_n$.

3.2.6 利用 Binet-Cauchy 公式及 Cauchy 不等式, 试证: 对任两 $n\times m$ 实矩阵 A 和 B, 有不等式 $\det(AA')\det(BB')\geqslant(\det(AB'))^2$.

3.2.7 试证: 对任一 $n\times m$ 实矩阵 A, n 阶方阵 AA' 的任一主子式都是非负实数.

3.2.8 试证: 对任两 n 阶方阵 A 和 B, AB 和 BA 的 k 阶主子式的和相等. 利用这一事实及习题 2.4.15, 试证: $\det(\lambda E-AB)=\det(\lambda E-BA)$.

3.2.9 设 A 和 B 分别为 $n\times m$ 和 $m\times n$ 矩阵, 记

$$C=\begin{pmatrix}\lambda E_n&0\\0&E_m\end{pmatrix}-\begin{pmatrix}0&A\\B&0\end{pmatrix}.$$

试证: 存在 $n+m$ 阶方阵 P, 它适合 $\det(PC)=\lambda^m\det(\lambda E_n-AB)$, $\det(CP)=\lambda^n\det(\lambda E_m-BA)$. 由此可推出

$$\lambda^m\det(\lambda E_n-AB)=\lambda^n\det(\lambda E_m-BA).\tag{3.2.5}$$

3.2.10 设 $i<j$, $k<p$. n 阶方阵 A 的元素 a_{ik}, a_{ip}, a_{jk}, a_{jp} 的代数余子式分别记为 A_{ik}, A_{ip}, A_{jk}, A_{jp}. 试证公式

$$\det\begin{pmatrix}A_{ik}&A_{jk}\\A_{ip}&A_{jp}\end{pmatrix}=(-1)^{i+j+k+p}\det A\begin{pmatrix}1\cdots i-1,i+1\cdots j-1,j+1\cdots n\\1\cdots k-1,k+1\cdots p-1,p+1\cdots n\end{pmatrix}\det(A).$$

3.2.11 计算下列 n 阶行列式:

$$\det\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ \mu a_{n-1} & a_0 & a_1 & \cdots & a_{n-3} & a_{n-2} \\ \mu a_{n-2} & \mu a_{n-1} & a_0 & \cdots & a_{n-4} & a_{n-3} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \mu a_1 & \mu a_2 & \mu a_3 & \cdots & \mu a_{n-1} & a_0 \end{pmatrix};$$

$$\det\begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & & \vdots \\ s_{n-2} & s_{n-1} & \cdots & s_{2n-2} \\ s_{n-1} & s_n & \cdots & s_{2n-1} \end{pmatrix}, \quad s_j = \sum_{i=1}^{n} x_i^j, \quad j = 0,1,2,\cdots.$$

3.2.12 设 $A^{(*)}$ 是 A 的伴随方阵 (定义见习题 3.1.4), 试证

$$\begin{aligned} &\det A^{(*)}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \\ &= (\det A)^{k-1}\delta^{i_1 i_2\cdots i_n}_{j_1 j_2\cdots j_n}\det A\begin{pmatrix} i_{k+1} & i_{k+2} & \cdots & i_n \\ j_{k+1} & j_{k+2} & \cdots & j_n \end{pmatrix}, \end{aligned}$$

其中 $i_1i_2\cdots i_n$ 和 $j_1j_2\cdots j_n$ 都是 $1,2,\cdots,n$ 的排列, 且 $1 \leqslant i_1 < \cdots < i_k \leqslant n$, $1 \leqslant i_{k+1} < \cdots < i_n \leqslant n$; $1 \leqslant j_1 < \cdots < j_k \leqslant n$, $1 \leqslant j_{k+1} < \cdots < j_n \leqslant n$.

3.2.13 设 $A^{(*)}$ 是 A 的伴随方阵, 试证:

(1) $\operatorname{rank}(A^{(*)}) = n$ 的必要且充分条件为 $\operatorname{rank}(A) = n$;

(2) $\operatorname{rank}(A^{(*)}) = 1$ 的必要且充分条件为 $\operatorname{rank}(A) = n-1$;

(3) $\operatorname{rank}(A^{(*)}) = 0$ 的必要且充分条件为秩 $\operatorname{rank}(A) < n-1$;

(4) $(A^{(*)})^{(*)} = (\det(A))^{n-2}A$;

(5) $(AB)^{(*)} = B^{(*)}A^{(*)}$.

由此可见, n 阶方阵的伴随方阵的秩只有 0, 1, n 之一. (秩的定义见 §3.3)

§3.3 矩阵的逆方阵和秩

现在详细地考查 n 阶方阵.

定义 3.3.1 设 A 为域 $\mathbb{F}$ 上 n 阶方阵, 记为

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}. \tag{3.3.1}$$

记 A_{jk} 为元素 a_{jk} 的代数余子式. 用域 $\mathbb{F}$ 上 n^2 个数 A_{jk} 构作的 n 阶方阵

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}' = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{3.3.2}$$

称为方阵 A 的**伴随方阵**, 记作 $A^{(*)}$.

引理 3.3.2 设 $A^{(*)}$ 为域 $\mathbb{F}$ 上 n 阶方阵 A 的伴随方阵, 则有

$$AA^{(*)} = A^{(*)}A = (\det(A))E_n. \tag{3.3.3}$$

证 由定理 2.3.5 及矩阵乘法的定义可知式 (3.3.3) 成立. □

定义 3.3.3 域 $\mathbb{F}$ 上 n 阶方阵 A 称为 **非异的**, 如果

$$\det(A) \neq 0. \tag{3.3.4}$$

定义 3.3.4 域 $\mathbb{F}$ 上 n 阶方阵 A 称为 **可逆的**, 如果存在 n 阶方阵 B, 使得

$$AB = BA = E_n, \tag{3.3.5}$$

其中 E_n 为 n 阶单位方阵. n 阶方阵 B 称为可逆方阵 A 的**逆方阵**.

定理 3.3.5 n 阶方阵 A 是非异的当且仅当 A 是可逆的, 且当 A 可逆时, 它的逆方阵唯一存在, 记作 A^{-1}, 它是

$$A^{-1} = (\det(A))^{-1}A^{(*)}. \tag{3.3.6}$$

证 先证必要性. 今设 A 为非异方阵, 即 $\det(A) \neq 0$. 由式 (3.3.3), 令

$$B = (\det(A))^{-1}A^{(*)},$$

则有

$$AB = BA = E_n.$$

这证明了 A 为可逆方阵. 反之, 若存在方阵 B 使得 $AB = BA = E$, 由 Binet-Cauchy 公式, 两边取行列式, 得

$$1 = \det(AB) = \det(A)\det(B),$$

因此 $\det(A) \neq 0$. 这证明了 A 为非异方阵.

上面得到一个可逆方阵 A 的逆方阵 $B = (\det(A))^{-1}A^{(*)}$. 下面证明可逆方阵的逆方阵唯一存在.

事实上, 如果方阵 A 另有一个逆方阵 B_1, 则由定义有 $B_1A = AB_1 = E_n$. 所以

$$B_1 = E_nB_1 = (BA)B_1 = B(AB_1) = BE_n = B.$$

这就证明了唯一性. □

定理 3.3.5 证明了 "非异" 和 "可逆" 这两个概念是等价的. 关于逆方阵还有下列简单的性质:

(1) $(A^{-1})^{-1} = A$;

(2) $(AB)^{-1} = B^{-1}A^{-1}$;

(3) $(cA)^{-1} = c^{-1}A^{-1}$, $\forall c \in \mathbb{F}$, $c \neq 0$;

(4) $(A')^{-1} = (A^{-1})'$;

(5) $(\overline{A})^{-1} = \overline{A^{-1}}$. 由**(4)** 和**(5)**, 因此, $(A^*)^{-1} = (A^{-1})^*$;

(6) $\det(A^{-1}) = \det(A)^{-1} \neq 0$.

证 我们只证性质**(2)** 成立, 读者试自证其他的性质成立.

利用逆方阵的唯一性, 只要证明 $B^{-1}A^{-1}$ 是方阵 AB 的逆方阵就行了. 今

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = E_n,$$

又

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = E_n,$$

故断言成立. □

由性质**(2)**, 用数学归纳法, 很易证明: 设 n 阶方阵 $A_1, A_2, \cdots, A_p$ 都是可逆方阵, 则 $A_1A_2\cdots A_p$ 也是可逆方阵, 且其逆为

$$(A_1A_2\cdots A_{p-1}A_p)^{-1} = A_p^{-1}A_{p-1}^{-1}\cdots A_2^{-1}A_1^{-1}. \tag{3.3.7}$$

由此可见, 方阵的取逆运算和转置运算, 当作用在方阵的乘积上时, 具有类似的性质.

在 §2.5 中给出了一类特殊的线性方程组求解的 Cramer 法则. 实际上, Cramer 法则就相当于求方阵的逆.

考虑 n 个未知数 $x_1, x_2, \cdots, x_n$ 的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ \qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \tag{3.3.8}$$

记 n^2 个系数 a_{jk}, $1 \leqslant j, k \leqslant n$ 构成的方阵为

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}. \tag{3.3.9}$$

构作 $n\times 1$ 矩阵 $x=(x_1, x_2, \cdots, x_n)'$, 其中 $x'=(x_1, x_2, \cdots, x_n)$ 为 x 的转置矩阵, $\beta=(b_1, b_2, \cdots, b_n)'$. 按照矩阵乘法的定义, 线性方程组可以写成简洁的矩阵形式:

$$Ax=\beta. \tag{3.3.10}$$

假设 $\det(A)\neq 0$, 那末方阵 A 有逆方阵. 设线性方程组有一组解 $x^{(0)}=(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})'$, 其中 $x^{(0)}$ 为 $n\times 1$ 矩阵. 由定义, $Ax^{(0)}=\beta$. 所以

$$x^0=E_nx^{(0)}=A^{-1}Ax^{(0)}=A^{-1}\beta. \tag{3.3.11}$$

因此证明了线性方程组 $Ax=\beta$ 若有解, 则必为 $x^{(0)}=A^{-1}\beta$. 显然 $x^{(0)}=A^{-1}\beta$ 就是解. 事实上, 将它代入方程, 便有 $A(A^{-1}\beta)=E_n\beta=\beta$. 于是, 线性方程组 $Ax=b$ 有且只有一组解 $x^{(0)}=A^{-1}\beta$, 这组解就是定理 2.5.1 中给出的解. 事实上,

$$x^{(0)}=A^{-1}\beta=(\det(A))^{-1}A^{(*)}(b_1,b_2,\cdots,b_n)', \tag{3.3.12}$$

所以

$$x^{(0)}=(\det(A))^{-1}\Big(\sum_{k=1}^{n}b_kA_{k1},\cdots,\sum_{k=1}^{n}b_kA_{kn}\Big)',$$

即

$$x_j^{(0)}=(\det(A))^{-1}\sum_{k=1}^{n}b_kA_{kj}=\frac{\Delta_j}{\Delta},\quad j=1,2,\cdots,n. \tag{3.3.13}$$

所以, 利用矩阵的符号和性质, 重新给出了 Cramer 法则, 而且叙述得相当简洁. 由此也可以看出矩阵理论的作用.

利用可逆矩阵, 我们介绍矩阵计算中的最基本的技巧之一. 我们称之为 — **打洞技巧**.

设 A 为域 $\mathbb{F}$ 上 $n\times m$ 矩阵, 将它分成四块为

$$\begin{pmatrix} B^{(r)} & C^{(r,m-r)} \\ D^{(n-r,r)} & G^{(m-r)} \end{pmatrix}. \tag{3.3.14}$$

于是当 $\det(B)\neq 0$ 时有**Schur 公式**

$$\begin{pmatrix} E_r & 0 \\ -DB^{-1} & E_{n-r} \end{pmatrix}\begin{pmatrix} B & C \\ D & G \end{pmatrix}=\begin{pmatrix} B & C \\ 0 & G-DB^{-1}C \end{pmatrix}, \tag{3.3.15}$$

$$\begin{pmatrix} B & C \\ D & G \end{pmatrix}\begin{pmatrix} E_r & -B^{-1}C \\ 0 & E_{n-r} \end{pmatrix}=\begin{pmatrix} B & 0 \\ D & G-DB^{-1}C \end{pmatrix}. \tag{3.3.16}$$

用打洞技巧, 我们可以将高阶行列式的计算化为低阶行列式的计算.

定理 3.3.6 设 A 为域 $\mathbb{F}$ 上的 n 阶方阵. 按照 r, $n-r$ 行, 列作分块, 记作

$$A=\begin{pmatrix} B & C \\ D & G \end{pmatrix}. \tag{3.3.17}$$

设 $\det(B) \neq 0$, 则有

$$A = \begin{pmatrix} B & C \\ D & G \end{pmatrix} = \begin{pmatrix} E_r & 0 \\ DB^{-1} & E_{n-r} \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & G - DB^{-1}C \end{pmatrix} \begin{pmatrix} E_r & B^{-1}C \\ 0 & E_{n-r} \end{pmatrix}. \tag{3.3.18}$$

所以由 Laplace 展开定理可知

$$\det(A) = \det(B)\det(G - DB^{-1}C). \tag{3.3.19}$$

因此当 A 及 B 都是可逆方阵时, $G - DB^{-1}C$ 也可逆, 且

$$A^{-1} = \begin{pmatrix} E_r & -B^{-1}C \\ 0 & E_{n-r} \end{pmatrix} \begin{pmatrix} B^{-1} & 0 \\ 0 & (G - DB^{-1}C)^{-1} \end{pmatrix} \begin{pmatrix} E_r & 0 \\ -DB^{-1} & E_{n-r} \end{pmatrix}. \tag{3.3.20}$$

例 3.3.1 因为 $\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} \end{pmatrix}$, 所以记 $A = \begin{pmatrix} 1 & 2 & 3 & 7 \\ 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 2 \\ 2 & 1 & 1 & 1 \end{pmatrix}$, 则有

$$\det(A) = \det \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \det \left(\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} 3 & 7 \\ 1 & 0 \end{pmatrix} \right).$$

因此

$$\det(A) = \det \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \det \begin{pmatrix} -\frac{15}{4} & -\frac{17}{2} \\ -\frac{5}{4} & -\frac{5}{2} \end{pmatrix} = 5.$$

下面给出另一种矩阵技巧, 即所谓**摄动法**. 我们用一个例子来介绍摄动法如下:

例 3.3.2 将域 $\mathbb{F}$ 上 $2n$ 阶方阵 A 按 n, n 行, 列作分块, 记作

$$A = \begin{pmatrix} B & C \\ D & G \end{pmatrix}. \tag{3.3.21}$$

则当 $BC = CB$ 时有

$$\det(A) = \det(GB - DC). \tag{3.3.22}$$

证 引进域 $\mathbb{F}$ 上参数 t, 记 $B_t = tE_n + B$. 于是 $\det(B_t) = \det(tE_n + B)$ 为 t 的 n 次多项式, 它的首项系数为 1. 因此它至多有 n 个不同根. 于是存在无限序列 t_1, $t_2, \cdots$, 使得它们都不是 $\det(B_t)$ 的根, 且有 $t_k \to 0$. 考虑 $2n$ 阶方阵

$$A_t = \begin{pmatrix} B_t & C \\ D & G \end{pmatrix} = \begin{pmatrix} tE_n + B & C \\ D & G \end{pmatrix}. \tag{3.3.23}$$

由 $B_t=tE_n+B$ 及 $BC=CB$ 可知 $B_tC=CB_t$. 取 $t=t_k$, $\forall k$, 则有 $\det(B_{t_k})\neq 0$. 于是 $B_{t_k}^{-1}C=CB_{t_k}^{-1}$.

由定理 3.3.6, 有

$$\begin{aligned}\det(A_{t_k})&=\det(B_{t_k})\det(G-DB_{t_k}^{-1}C)=\det(G-DCB_{t_k}^{-1})\det(B_{t_k})\\&=\det(GB_{t_k}-DC)=\det(t_kG+GB-DC).\end{aligned}$$

两边都是 t_k 的多项式. 取 $t_k\to 0$, 便有 $\det(A)=\det(GB-DC)$. 故断言成立. □

下面引进矩阵的另一个重要概念 —— 秩.

定义 3.3.7 域 $\mathbb{F}$ 上 $n\times m$ 非零矩阵 A 的所有子式中必有一个阶数最大的非零子式. 其阶数称为矩阵 A 的**秩**, 记作**秩**(A), 或 $\operatorname{rank}(A)$. 零矩阵的秩定义为零.

由于矩阵 A 的子式的阶数不超过 A 的行数及列数, 所以有

引理 3.3.8 域 $\mathbb{F}$ 上 $n\times m$ 矩阵 A 的秩有

(1) $0\leqslant \operatorname{rank}(A)\leqslant \min(n,m)$;

(2) $\operatorname{rank}(cA)=\operatorname{rank}(A)$, $\forall\, 0\neq c\in\mathbb{F}$;

(3) $\operatorname{rank}(A')=\operatorname{rank}(A)$;

(4) $\operatorname{rank}(\overline{A})=\operatorname{rank}(A)$, $\quad\operatorname{rank}(A^*)=\operatorname{rank}(A)$.

定义 3.3.9 $n\times m$ 矩阵 A 称为**满秩的**, 如果

$$\operatorname{rank}(A)=\min(n,m). \tag{3.3.24}$$

定理 3.3.10 域 $\mathbb{F}$ 上 n 阶方阵 A 非异当且仅当 A 可逆, 当且仅当 A 满秩, 即对 n 阶方阵, 非异、可逆、满秩是三个互相等价的概念.

证 由秩的定义可知, n 阶方阵 A 的秩为 n (即满秩) 的必要且充分条件为 $\det(A)\neq 0$. 即对方阵而言,"秩为 n" 和 "行列式不等于零" 这两概念是等价的. 由定理 3.3.5, 就证明了定理. □

由上可知, n 阶方阵 A 的秩小于 n 的必要且充分条件为 $\det(A)=0$. 为了区别起见, 今后称适合条件 $\det(A)\neq 0$ 的方阵 A 为**非退化的**或**非异的**, 称适合条件 $\det(A)=0$ 的方阵 A 为**退化的**或**奇异的**.

下面给出秩的另一个等价定义.

定理 3.3.11 域 $\mathbb{F}$ 上 $n\times m$ 矩阵 A 的秩为 r 的必要且充分条件是: 在 A 中存在一个 r 阶子式不等于零, 且在 $r<\min(n,m)$ 时, 矩阵 A 的所有 $r+1$ 阶子式为零.

证 从秩的定义可知, 在矩阵 A 的秩为 r 时, 必然存在一个 r 阶子式不为零, 且所有阶数为 $r+1$ 的子式都等于零. 这证明了必要性. 下面来证充分性, 即设在 A 中存在一个 r 阶子式不等于零, 且在 $r<\min(n,m)$ 时, 矩阵 A 的所有 $r+1$ 阶子式为

零. 我们来证矩阵 A 的所有 $r+k$ 阶子式也都为零, $2 \leqslant k \leqslant \min(n, m)$. 事实上, 设 A 的子式 M 的阶为 $r_1 > r+1$. 将行列式 M 按照前 $r+1$ 行作 Laplace 展开, 由定理 2.3.3 可知, 它是 $\binom{r_1}{r+1}$ 项的和, 每一项是 M 的一个 $r+1$ 阶子式和域 $\mathbb{F}$ 中一数的乘积. 因为行列式 M 的子式也是矩阵 A 的子式, 所以行列式 M 的 Laplace 展开式中每一项都是零, 因此 $M=0$. □

这个定理有一个简单的推论: 设域 $\mathbb{F}$ 上矩阵 A 的秩为 r, 则在矩阵 A 中必分别存在一个 $1, 2, \cdots, r$ 阶不等于零的子式. 因此, 在具体求矩阵的秩时, 可以用这样一种技巧: 即如果矩阵 A 有一个 s 阶子式不为零, 则矩阵 A 的秩必大于或等于 s. 所以, 依次地去寻找 1, 2, $\cdots$ 阶不为零的子式. 如果在找到一个 r 阶不为零的子式后, 找不到 $r+1$ 阶不为零的子式, 换句话说, 所有 $r+1$ 阶子式全为零, 这时矩阵 A 的秩就是 r. 因此, 从原则上来讲, 是从低阶往高阶寻找不为零的子式的.

例 3.3.3 矩阵

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

有一个不为零的一阶子式, 而所有二阶子式都为零, 所以 $\operatorname{rank}(A)=1$.

例 3.3.4 矩阵

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

各有一个 $1,2,3$ 阶不为零的子式, 所以 A 的秩为 3.

定理 3.3.12 设 A, B, C 分别为域 $\mathbb{F}$ 上 $n \times m$, $p \times q$, $n \times q$ 矩阵, 则有

$$\operatorname{rank}\left(\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\right) \geqslant \operatorname{rank}\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right)=\operatorname{rank}(A)+\operatorname{rank}(B). \tag{3.3.25}$$

证 设 $r=\operatorname{rank}(A)$, $s=\operatorname{rank}(B)$. 于是在 A 中存在 r 阶子矩阵 A_0, 在 B 中存在 s 阶子矩阵 B_0, 使得 $\det(A_0) \neq 0$, $\det(B_0) \neq 0$. 从而在矩阵 $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ 中存在 $r+s$ 阶子矩阵 $\begin{pmatrix} A_0 & 0 \\ 0 & B_0 \end{pmatrix}$ 非异. 由 Laplace 展开定理可知在矩阵 $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ 中存在 $r+s$ 阶子矩阵 $\begin{pmatrix} A_0 & C_0 \\ 0 & B_0 \end{pmatrix}$ 非异. 这证明了

$$\operatorname{rank}\left(\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\right) \geqslant \operatorname{rank}(A)+\operatorname{rank}(B), \quad \operatorname{rank}\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) \geqslant \operatorname{rank}(A)+\operatorname{rank}(B).$$

所以问题化为证

$$\operatorname{rank}\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) \leqslant \operatorname{rank}(A) + \operatorname{rank}(B).$$

由定理 3.3.11, 考虑矩阵 $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ 的任意 $r+s+1$ 阶子方阵, 它必形如 $X = \begin{pmatrix} A_1 & 0 \\ 0 & B_1 \end{pmatrix}$, 其中 A_1 为 $n_1 \times m_1$ 矩阵, B_1 为 $p_1 \times q_1$ 矩阵, 而

$$n_1 + p_1 = m_1 + q_1 = r + s + 1.$$

下面证明 $\det(X) = 0$. 用反证法. 设若不然, 即 $\det(X) \neq 0$. 由 Laplace 展开式可证 $n_1 = m_1$, $p_1 = q_1$, 且 $\det(X) = \det(A_1)\det(B_1)$. 但是由 $n_1 + p_1 = r + s + 1$ 可知 $n_1 \geqslant r + 1$, 或者 $p_1 \geqslant s + 1$. 当 $n_1 \geqslant r + 1$, 由 $\operatorname{rank}(A) = r$ 可知 $\det(A_1) = 0$; 当 $p_1 \geqslant s + 1$, 由 $\operatorname{rank}(B) = s$ 可知 $\det(B_1) = 0$. 总之, 证明了 $\det(X) = 0$, 这导出矛盾. □

利用 Binet-Cauchy 公式, 可以讨论矩阵乘积的秩.

定理 3.3.13 设 A 为 $n \times m$ 矩阵, B 为 $m \times p$ 矩阵, 则有

$$\operatorname{rank}(AB) \leqslant \min(\operatorname{rank}(A), \operatorname{rank}(B)). \tag{3.3.26}$$

证 记 $\operatorname{rank}(AB) = r$, 任取 AB 的 r 阶子式, 由 Binet-Cauchy 公式可知

$$\det(AB)\begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix}$$

等于

$$\sum_{1 \leqslant k_1 \leqslant k_2 \leqslant \cdots \leqslant k_r \leqslant m} \det A \begin{pmatrix} i_1 i_2 \cdots i_r \\ k_1 k_2 \cdots k_r \end{pmatrix} \det B \begin{pmatrix} k_1 k_2 \cdots k_r \\ i_1 i_2 \cdots i_r \end{pmatrix}.$$

今 $\operatorname{rank}(AB) = r$, 所以存在一个 r 阶子式不等于零. 设为 $\det(AB)\begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix} \neq 0$, 因此存在 $1 \leqslant k_1 \leqslant k_2 \leqslant \cdots \leqslant k_r \leqslant m$, 使得

$$\det A \begin{pmatrix} i_1 i_2 \cdots i_r \\ k_1 k_2 \cdots k_r \end{pmatrix} \det B \begin{pmatrix} k_1 k_2 \cdots k_r \\ i_1 i_2 \cdots i_r \end{pmatrix} \neq 0.$$

于是 $\operatorname{rank}(A) \geqslant r$, $\operatorname{rank}(B) \geqslant r$. 这证明了定理. □

非异方阵有下面的重要性质, 即

定理 3.3.14 设 A 为 $n \times m$ 矩阵, P 为 n 阶非异方阵, Q 为 m 阶非异方阵. 则有

$$\operatorname{rank}(PA) = \operatorname{rank}(AQ) = \operatorname{rank}(A). \tag{3.3.27}$$

换句话说, 将非异方阵左乘或右乘在矩阵上, 其秩保持不变.

证 由定理 3.3.13,

$$\operatorname{rank}(PA) \leqslant \min(\operatorname{rank}(P), \operatorname{rank}(A)) \leqslant \operatorname{rank}(A).$$

由此可知 $\operatorname{rank}(A) = \operatorname{rank}(P^{-1}(PA)) \leqslant \operatorname{rank}(PA)$. 这证明了 $\operatorname{rank}(PA) = \operatorname{rank}(A)$. 而

$$\operatorname{rank}(AQ) = \operatorname{rank}(Q'A') = \operatorname{rank}(A') = \operatorname{rank}(A).$$

所以证明了定理. □

定理 3.3.15 设 A 为 $n \times m$ 矩阵, B 为 $m \times p$ 矩阵, 则有

$$\operatorname{rank}(A) + \operatorname{rank}(B) \leqslant \operatorname{rank}(AB) + m. \tag{3.3.28}$$

证 由定理 3.3.12, 有

$$\operatorname{rank}(A) + \operatorname{rank}(B) = \operatorname{rank}\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) \leqslant \operatorname{rank}\left(\begin{pmatrix} A & 0 \\ E_m & B \end{pmatrix}\right).$$

而

$$\begin{pmatrix} E_n & -A \\ 0 & E_m \end{pmatrix} \begin{pmatrix} A & 0 \\ E_m & B \end{pmatrix} \begin{pmatrix} E_m & -B \\ 0 & E_p \end{pmatrix} = \begin{pmatrix} 0 & -AB \\ E_m & 0 \end{pmatrix}.$$

由定理 3.3.14,

$$\operatorname{rank}\left(\begin{pmatrix} A & 0 \\ E_m & B \end{pmatrix}\right) = \operatorname{rank}\left(\begin{pmatrix} 0 & -AB \\ E_m & 0 \end{pmatrix}\right) = \operatorname{rank}\left(\begin{pmatrix} AB & 0 \\ 0 & E_m \end{pmatrix}\right)$$

由定理 3.3.12,

$$\operatorname{rank}(A) + \operatorname{rank}(B) \leqslant \operatorname{rank}\left(\begin{pmatrix} AB & 0 \\ 0 & E_m \end{pmatrix}\right) = \operatorname{rank}(AB) + \operatorname{rank}(E_m) = m + \operatorname{rank}(AB).$$

所以式 (3.3.28) 成立. □

这个定理给出了一种重要的矩阵技巧, 用来计算有关各种矩阵的秩的不等式.

习 题 3.3

除非特别申明, 下面都是在域 $\mathbb{F}$ 上讨论.

3.3.1 试计算下列矩阵的秩:

(1) $\begin{pmatrix} 14 & 12 \\ 6 & 104 \\ 2 & 1 \\ 5 & 0 \end{pmatrix}$; **(2)** $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$.

3.3.2 设 A 和 B 都是 $n\times m$ 矩阵. 试证 $\operatorname{rank}(A+B)\leqslant\operatorname{rank}((A,B))\leqslant\operatorname{rank}(A)+\operatorname{rank}(B)$.

3.3.3 设 A 为 $n\times m$ 矩阵, B 为 $m\times s$ 矩阵. 试证: $\operatorname{rank}(A)+\operatorname{rank}(B)-m\leqslant\operatorname{rank}(AB)\leqslant\min(\operatorname{rank}(A),\operatorname{rank}(B))$.

3.3.4 试证: 对任一 $n\times m$ 复矩阵 A, 有 $\operatorname{rank}(AA^*)=\operatorname{rank}(A^*A)=\operatorname{rank}(A)$.

3.3.5 设 A 是秩为 1 的 $n\times m$ 矩阵. 试证: 存在 $1\times n$ 矩阵 $\alpha=(a_1,a_2,\cdots,a_n)$ 及 $1\times m$ 矩阵 $\beta=(b_1,b_2,\cdots,b_m)$, 使得 $A=\alpha'\beta$.

3.3.6 试举一例说明: 存在 n 阶可逆方阵 A 和 B, 使得 $A+B$ 也可逆, 但是 $(A+B)^{-1}\neq A^{-1}+B^{-1}$.

3.3.7 设 λ 为数, A 和 B 为 $n\times 1$ 矩阵, 试证: $\det(E_n-\lambda AB')=1-\lambda B'A$. 什么时候 $E_n-\lambda AB'$ 有逆, 逆是什么?

3.3.8 设 n 阶方阵

$$A=\begin{pmatrix}P^{(r)} & Q\\ 0 & E_{n-r}\end{pmatrix}$$

有 $\det(E_r-P)\neq 0$, 试证:

$$A^m=\begin{pmatrix}P^m & (P^m-E_r)(P-E_r)^{-1}Q\\ 0 & E_{n-r}\end{pmatrix},\quad m=0,1,2,\cdots.$$

当 $\det(P)\neq 0$ 时, 试证: 上式对 $m=-1,m=-2,\cdots$ 也对.

3.3.9 将 n 阶方阵 A 按照前 r 行, 前 r 列分成四块为

$$A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}.$$

设 r 阶方阵 A_{11} 的秩为 r, 则 $\operatorname{rank}(A)=r$ 的充分且必要条件为 $A_{22}=A_{21}A_{11}^{-1}A_{12}$.

3.3.10 设 n 阶方阵 A 的秩为 r. 试证: 在其中任取 s 行构成 $s\times n$ 矩阵 B, 则 $\operatorname{rank}(B)\geqslant r+s-n$.

3.3.11 设 $n\times m$ 矩阵 A 的秩为 r. 试证: 在其中任取 s 行, t 列, 构成 $s\times t$ 矩阵 B, 则 $\operatorname{rank}(B)\geqslant r+s+t-m-n$.

3.3.12 设 n 阶方阵 A 至少有 n^2-n+1 个元素为零. 试证: $\operatorname{rank}(A)<n$. 并问最高的秩数是多少?

3.3.13 设 n 阶方阵 A 适合方程 $a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E_n=0$, 其中 $a_0a_m\neq 0$, 试证 A 非异.

3.3.14 设 A 是秩为 r 的对称方阵, 即适合条件 $S=S'$ 的方阵. 试证: S 至少有一个 r 阶主子式

$$\det S\begin{pmatrix}i_1\ i_2\ \cdots\ i_r\\ i_1\ i_2\ \cdots\ i_r\end{pmatrix}\neq 0.$$

3.3.15 设 n 阶方阵 A 适合 $A^k=0$, $A^{k-1}\neq 0$, 则 A 称为 n 阶 k 次**幂零方阵**. 这时试证: n 阶方阵 E_n-A 非异, 且 $(E_n-A)^{-1}=E_n+A+A^2+\cdots+A^{k-1}$.

3.3.16 求下列方阵的逆方阵:

(1) $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$; (2) $\begin{pmatrix} E_r & 0^{(r,s)} \\ P^{(s,r)} & E_s \end{pmatrix}$; (3) $\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 0 \end{pmatrix}$.

3.3.17 试证: $n+m$ 阶方阵 $\begin{pmatrix} B^{(n,m)} & E_n \\ 0 & A^{(m,n)} \end{pmatrix}$ 非异当且仅当 m 阶方阵 AB 可逆.

3.3.18 设 A 为 $n\times m$ 矩阵, B 为 $m\times p$ 矩阵. 设 $m=\operatorname{rank}(AB)$, 试证: $m=\operatorname{rank}(A)=\operatorname{rank}(B)$.

3.3.19 求 n 阶方阵

$$A=\begin{pmatrix} E_r & P^{(r,n-r)} \\ Q^{(n-r,r)} & E_{n-r} \end{pmatrix}$$

的逆方阵, 这里 $E_{n-r}-QP$ 可逆.

3.3.20 设 A 和 B 分别为 $n\times m$ 及 $m\times n$ 矩阵, 试证: E_n-AB 非异当且仅当 E_m-BA 非异. 试用 A, B 及 E_n-AB 的逆方阵来表达 E_m-BA 的逆方阵.

3.3.21 求 Vandermonde 方阵 $V_n(1,\omega,\omega^2,\cdots,\omega^{n-1})$ 的逆方阵, 其中

$$\omega=\exp\left(\frac{2\pi i}{n}\right)=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n},$$

3.3.22 设 A_{ii}, $i=1,2,\cdots,n$ 为 m 阶可逆方阵, $A_{i,i+1}$, $i=1,2,\cdots,n-1$ 为 m 阶方阵. 试求: nm 阶方阵 A 的逆方阵, 其中

$$A=\begin{pmatrix} A_{11} & A_{12} & & & \\ & A_{22} & A_{23} & & \\ & & \ddots & \ddots & \\ & & & A_{n-1,n-1} & A_{n-1,n} \\ & & & & A_{nn} \end{pmatrix}.$$

3.3.23 设 A 和 B 都是 n 阶方阵, 且有 $\operatorname{rank}(A)+\operatorname{rank}(B)=n$. 试求矩阵方程 $AXB=0$ 的通解.

3.3.24 设 A_1, A_2, $\cdots$, A_p 为 p 个 n 阶方阵, 试证:

$$\sum_{j=1}^{p}\operatorname{rank}(A_j)\leqslant n(p-1)+\operatorname{rank}(A_1A_2\cdots A_p),$$

且证等式可以达到.

3.3.25 设 A, B, C 分别为 $n\times m$, $m\times p$, $m\times q$ 矩阵, 且 $AB=0$, $AC=0$. $\operatorname{rank}(A)+\operatorname{rank}(B)=n$. 试证: 存在 $p\times q$ 矩阵 D, 使得 $C=BD$. 且证唯一存在 D 的必要且充分条件为 $\operatorname{rank}(B)=p$.

3.3.26 (Frobenious 不等式) 设 A, B, C 分别为 $n\times m$, $m\times p$, $p\times q$ 矩阵. 试证: $\operatorname{rank}(AB)+\operatorname{rank}(BC)\leqslant\operatorname{rank}(B)+\operatorname{rank}(ABC)$. 用此式来证: $\operatorname{rank}(A)+\operatorname{rank}(B)\leqslant m+\operatorname{rank}(AB)$ 以及当 $\operatorname{rank}(AB)=\operatorname{rank}(B)$ 时有 $\operatorname{rank}(BC)=\operatorname{rank}(ABC)$.

3.3.27 设 A 为 n 阶方阵, 且设存在正整数 $N>0$, 使得 $\operatorname{rank}(A^N)=\operatorname{rank}(A^{N+1})$. 试证: $\operatorname{rank}(A^N)=\operatorname{rank}(A^{N+1})=\operatorname{rank}(A^{N+2})=\cdots$.

§3.4 初等变换和矩阵的相抵

在这一节引进矩阵论的另一个重要概念: 初等变换. 下面首先给出三类非异方阵, 统称为**初等方阵**.

定义 3.4.1

(i) n 阶方阵

$$P_{jk}=\begin{pmatrix} E_{j-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & E_{k-j-1} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{n-k} \end{pmatrix}=\sum_{i\neq j,k}E_{ii}+E_{jk}+E_{kj},\quad 1\leqslant j<k\leqslant n \tag{3.4.1}$$

称为**置换方阵**, 或称为 **第一类初等方阵**. 一般, 一批置换方阵的乘积也称为置换方阵. 所以置换方阵的每一行 (列) 中有且只有一个 1, 其他元素为零.

(ii)

$$P_j(a)=\begin{pmatrix} E_{j-1} & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & E_{n-j} \end{pmatrix}=\sum_{i\neq j}E_{ii}+aE_{jj},\quad a\neq 0,\quad 1\leqslant j\leqslant n \tag{3.4.2}$$

称为**第二类初等方阵**.

(iii) $\forall\lambda\in\mathbb{F}$,

$$Q_{jk}(\lambda)=\begin{pmatrix} E_{j-1} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \lambda & 0 \\ 0 & 0 & E_{k-j-1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & E_{n-k} \end{pmatrix}=\sum_{i=1}^{n}E_{ii}+\lambda E_{jk},\ j\neq k,\ 1\leqslant j,k\leqslant n \tag{3.4.3}$$

称为**第三类初等方阵**.

显然, 方阵 P_{jk}, $P_j(a)$ 都是对称方阵, 又 P_{jk}, $P_j(a)$, $Q_{jk}(\lambda)$ 都是非异方阵, 它们的逆方阵分别为

$$P_{jk}^{-1}=P_{jk},\quad P_j(a)^{-1}=P_j(a^{-1}),\quad Q_{jk}(\lambda)^{-1}=Q_{jk}(-\lambda). \tag{3.4.4}$$

所以, 它们的逆是同类型的初等方阵. 由于初等方阵都非异, 由定理 3.3.14 可知: 在初等变换下, 矩阵的秩不改变. 又由直接计算不难证明: 给定 $n\times m$ 矩阵 A, 则有

(1) 将 A 的第 j 行 (列) 与第 k 行 (列) 互换, 等于矩阵 $P_{jk}^{(n)}A$ $(AP_{jk}^{(m)})$, 这时其他行 (列) 都未变动. 这称为对矩阵 A 的行 (列) 作了**第一类初等变换**;

(2) 将 A 的第 j 行 (列) 乘以域中非零常数 a, 等于矩阵 $P_j(a)^{(n)}A$ $(AP_j(a)^{(m)})$, 这时其他行 (列) 也都未变动. 这称为对矩阵 A 的行 (列) 作了**第二类初等变换**;

(3) 将 A 的第 k 行 (第 j 列) 遍乘以域中常数 λ, 再全部加到第 j 行 (第 k 列) 上, 等于矩阵 $P_{jk}(\lambda)^{(n)}A$ $(AP_{jk}(\lambda)^{(m)})$, 这时其他列都未变动, 特别第 k 行 (第 j 列) 未变动. 这称为对矩阵的行 (列) 作了**第三类初等变换**. 三类初等变换统称为**初等变换**.

关于初等变换, 有下面的重要引理.

引理 3.4.2 设 $n\times m$ 矩阵 A 的秩为 r, 则存在一系列初等变换, 它们将 A 变为 $n\times m$ 矩阵

$$\Lambda_r=\begin{pmatrix}E^{(r)} & 0^{(r,m-r)}\\ 0^{(n-r,r)} & 0^{(n-r,m-r)}\end{pmatrix}. \tag{3.4.5}$$

确切地说, 存在 u 个 n 阶初等方阵 $P_1, P_2, \cdots, P_u$ 及 v 个 m 阶初等方阵 $Q_1, Q_2, \cdots, Q_v$, 使得

$$P_uP_{u-1}\cdots P_2P_1AQ_1Q_2\cdots Q_{v-1}Q_v=\Lambda_r. \tag{3.4.6}$$

证 在 $A=0$ 时, 引理显然成立. 假设 $A\neq 0$, 即设 A 的秩 $r\geqslant 1$. 这时, 在 A 中一定有一个非零元素, 记作 a_{jk}. 将 A 的第一行和第 j 行互换, 然后再将第一列和第 k 列互换. 于是从矩阵 A, 经过第一类初等变换, 得到新的矩阵 B. 它的第一行, 第一列元素为 $a_{jk}\neq 0$. 所以, 在讨论中无妨设矩阵 A 的第一行, 第一列元素 $a_{11}\neq 0$. 这时将矩阵 A 的第一行遍乘实数 a_{11}^{-1}, 便将第一行, 第一列元素变成 1. 因此无妨设 $a_{11}=1$, 即

$$A=\begin{pmatrix}1 & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nm}\end{pmatrix}.$$

将方阵 A 的第一行遍乘 $-a_{j1}$, 再全部加到第 j 行上. 对 $j=2, 3, \cdots, n$, 一共施行 $n-1$ 次, 从而将 A 经过第三类初等变换变成矩阵

$$A_0=\begin{pmatrix}1 & a_{12} & \cdots & a_{1m}\\ 0 & b_{22} & \cdots & b_{2m}\\ \vdots & \vdots & & \vdots\\ 0 & b_{n2} & \cdots & b_{nm}\end{pmatrix}.$$

再将 A_0 的第一列遍乘 $-a_{1j}$, 再全部加到第 j 列上. 对 $j=2, 3, \cdots, m$, 一共施行了 $m-1$ 次, 从而将矩阵 A_0 经过初等变换变成矩阵

$$A_1=\begin{pmatrix}1 & 0^{(1,m-1)}\\ 0^{(n-1,1)} & B^{(n-1,m-1)}\end{pmatrix}.$$

显然, 对 $(n-1)\times(m-1)$ 矩阵 B 所施行的初等变换, 可以作为对矩阵 A_1 所施行的初等变换. 因此对 $(n-1)\times(m-1)$ 矩阵 B 用同样的方法进行处理, 可以一阶一阶地降下去, 最后便证明了可以经过一系列初等变换将矩阵 A 变为 $n\times m$ 矩阵

$$\Lambda_s = \begin{pmatrix} E^{(s)} & 0^{(s,m-s)} \\ 0^{(n-s,s)} & 0^{(n-s,m-s)} \end{pmatrix}.$$

因为在初等变换下矩阵的秩不改变, 所以证明了这个矩阵的秩为 $s=\operatorname{rank}(\Lambda_s)=r=\operatorname{rank}(A)$, 因此, $s=r$. 这就证明了引理. □

将这个引理应用于 n 阶非异方阵 A, 这时 $\Lambda_n=E_n$. 即存在 $s=u+v$ 个 n 阶初等方阵 $P_1, P_2, \cdots, P_u, Q_1, Q_2, \cdots, Q_v$, 使得

$$P_uP_{u-1}\cdots P_2P_1AQ_1Q_2\cdots Q_{v-1}Q_v=\Lambda_n=E_n, \tag{3.4.7}$$

因此

$$A=P_1^{-1}P_2^{-1}\cdots P_{u-1}^{-1}P_u^{-1}Q_v^{-1}Q_{v-1}^{-1}\cdots Q_2^{-1}Q_1^{-1}.$$

由于初等方阵的逆方阵仍为同类型的初等方阵, 所以证明了下述定理.

定理 3.4.3 任一 n 阶方阵 A 非异当且仅当 A 是有限个 n 阶初等方阵的连乘积.

因为初等方阵左乘在矩阵上, 引起了矩阵的行的变化, 右乘在矩阵上, 引起了矩阵的列的变化, 又知任一 n 阶非异方阵 A 是有限个初等方阵的乘积:

$$A=R_1R_2\cdots R_{s-1}R_s, \tag{3.4.8}$$

因此

$$R_s^{-1}R_{s-1}^{-1}\cdots R_2^{-1}R_1^{-1}A=AR_s^{-1}R_{s-1}^{-1}\cdots R_2^{-1}R_1^{-1}=E_n, \tag{3.4.9}$$

所以得到下述定理.

定理 3.4.4 任一非异方阵可以只经过行的初等变换变为单位方阵, 也可以只经过列的初等变换变为单位方阵.

利用等式

$$R_s^{-1}R_{s-1}^{-1}\cdots R_2^{-1}R_1^{-1}A=E_n, \quad R_s^{-1}R_{s-1}^{-1}\cdots R_2^{-1}R_1^{-1}=A^{-1}, \tag{3.4.10}$$

可以给出一种求逆方阵的方法, 它可以叙述为: 将 $n\times 2n$ 矩阵 (A,E_n) 经过一系列行的初等变换, 可化为 (E_n,B), 则 $B=A^{-1}$.

事实上, 对矩阵 (A, E_n) 作行的初等变换, 即左乘以 n 阶初等方阵的连乘积 R. 已知 $R(A, E_n) = (RA, R)$, 其中 $R = R_s^{-1} R_{s-1}^{-1} \cdots R_2^{-1} R_1^{-1}$. 所以

$$
\begin{aligned}
R_s^{-1} R_{s-1}^{-1} \cdots R_2^{-1} R_1^{-1} (A, E_n) &= (R_s^{-1} R_{s-1}^{-1} \cdots R_2^{-1} R_1^{-1} A, R_s^{-1} R_{s-1}^{-1} \cdots R_2^{-1} R_1^{-1}) \\
&= (E_n, A^{-1}).
\end{aligned} \tag{3.4.11}
$$

这一方法的优点在于, 不必计算方阵 A 的伴随方阵 A^* 和求 n 阶行列式 $\det(A)$.

例 3.4.1 令

$$
(A, E) = \left(\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{pmatrix},
$$

所以

$$
\begin{aligned}
&\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{Q_{21}(-2)} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -4 & -2 & 1 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 \end{pmatrix} \\
&\xrightarrow{Q_{31}(-1)} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -4 & -2 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{pmatrix} \xrightarrow{Q_{13}(-2)} \begin{pmatrix} 1 & 0 & 1 & 3 & 0 & -2 \\ 0 & -3 & -4 & -2 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\xrightarrow{Q_{23}(3)} \begin{pmatrix} 1 & 0 & 1 & 3 & 0 & -2 \\ 0 & 0 & -1 & -5 & 1 & 3 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{pmatrix} \xrightarrow{P_{23}} \begin{pmatrix} 1 & 0 & 1 & 3 & 0 & -2 \\ 0 & 1 & 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & -5 & 1 & 3 \end{pmatrix} \\
&\xrightarrow{Q_{13}(1)} \begin{pmatrix} 1 & 0 & 0 & -2 & 1 & 1 \\ 0 & 1 & 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & -5 & 1 & 3 \end{pmatrix} \xrightarrow{Q_{23}(1)} \begin{pmatrix} 1 & 0 & 0 & -2 & 1 & 1 \\ 0 & 1 & 0 & -6 & 1 & 4 \\ 0 & 0 & -1 & -5 & 1 & 3 \end{pmatrix} \\
&\xrightarrow{P_3(-1)} \begin{pmatrix} 1 & 0 & 0 & -2 & 1 & 1 \\ 0 & 1 & 0 & -6 & 1 & 4 \\ 0 & 0 & 1 & 5 & -1 & -3 \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 & 1 \\ -6 & 1 & 4 \\ 5 & -1 & -3 \end{pmatrix} \right),
\end{aligned}
$$

所以

$$
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 3 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 1 & 1 \\ -6 & 1 & 4 \\ 5 & -1 & -3 \end{pmatrix}.
$$

按照初等变换算出逆方阵后，还要利用定义来验证所求得的方阵确实是逆方阵. 在这一例中，经过直接验算也可知上式成立.

定义 3.4.5 两个 $n\times m$ 矩阵 A 和 B 称为**相抵的**，如果存在 n 阶非异方阵 P 和 m 阶非异方阵 Q，使得

$$B=PAQ. \tag{3.4.12}$$

显然，设 A 和 B 相抵，B 和 C 相抵，则 A 和 C 相抵. 今由定理 3.3.14 可知，两个相抵矩阵的秩相同. 又由定理 3.4.3 可知非异方阵是初等方阵的乘积，所以引理 3.4.2 可以改写为下述定理:

定理 3.4.6 设 $n\times m$ 矩阵 A 的秩为 r，则 A 相抵于 $n\times m$ 矩阵

$$\Lambda_r=\begin{pmatrix} E_r & 0\\ 0 & 0\end{pmatrix}. \tag{3.4.13}$$

于是有

定理 3.4.7 $n\times m$ 矩阵 A 和 B 相抵的充分且必要条件为它们有相同的秩.

证 必要性已证，现在证充分性. 今设 A 和 B 的秩都是 r，由定理 3.4.6 可知，存在 n 阶非异方阵 P_1,P_2 及 m 阶非异方阵 Q_1,Q_2，使得

$$P_1AQ_1=\Lambda_r,\quad P_2BQ_2=\Lambda_r.$$

所以 $B=(P_2^{-1}P_1)A(Q_1Q_2^{-1})$. 由定义可知矩阵 A 和 B 相抵. □

利用相抵的概念，可以解决一些矩阵问题，下面给出一个例子.

例 3.4.2 给定 $n\times m$ 矩阵 A，试求出下面矩阵方程的通解:

$$A'X=X'A.$$

解 设矩阵 A 的秩为 r. 已知存在 n 阶非异方阵 P 和 m 阶非异方阵 Q，使得

$$PAQ=\Lambda_r=\begin{pmatrix} E_r & 0^{(r,m-r)}\\ 0^{(n-r,r)} & 0^{(n-r,m-r)}\end{pmatrix}.$$

由条件可知 $A=P^{-1}\Lambda_rQ^{-1}$，所以 $(P^{-1}\Lambda_rQ^{-1})'X=X'(P^{-1}\Lambda_rQ^{-1})$，即

$$(Q')^{-1}\Lambda_r'(P')^{-1}X=X'P^{-1}\Lambda_rQ^{-1}.$$

等式双方左乘以 Q'，再右乘以 Q，于是等式变成

$$\Lambda_r'(P^{-1})'XQ=Q'X'P^{-1}\Lambda_r=((P^{-1})'XQ)'\Lambda_r.$$

利用矩阵的分块, 将 $n\times m$ 矩阵 $(P^{-1})'XQ$ 和 Λ_r 同法分块, 即记

$$(P^{-1})'XQ=\begin{pmatrix} Y_{11}^{(r)} & Y_{12}^{(r,m-r)} \\ Y_{21}^{(n-r,r)} & Y_{22}^{(n-r,m-r)} \end{pmatrix},$$

于是有

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}=\begin{pmatrix} Y_{11}' & Y_{21}' \\ Y_{12}' & Y_{22}' \end{pmatrix}\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

因此,

$$\begin{pmatrix} Y_{11} & Y_{12} \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} Y_{11}' & 0 \\ Y_{12}' & 0 \end{pmatrix},$$

即 $Y_{12}=0,\ Y_{11}=Y_{11}'$. 所以

$$(P^{-1})'XQ=\begin{pmatrix} Y_{11} & 0 \\ Y_{21} & Y_{22} \end{pmatrix},\quad Y_{11}'=Y_{11}.$$

这证明了所求的 $n\times m$ 矩阵 X 可表为

$$X=P'\begin{pmatrix} Y_{11} & 0 \\ Y_{21} & Y_{22} \end{pmatrix}Q^{-1},\quad Y_{11}'=Y_{11}.$$

反之, 任一上面形式的 $n\times m$ 矩阵 X, 只要 r 阶方阵适合条件 $Y_{11}'=Y_{11}$, 则 $X'A=A'X$. 故求出了矩阵方程 $A'X=X'A$ 的通解. □

从这个例子也可以看出, 运用矩阵分块的方法可以使得讨论既简单又明显. 这在以后的讨论中, 读者也可以发现这一点.

例 3.4.3 n 阶方阵 A 有

$$\operatorname{rank}(A)+\operatorname{rank}(E_n-A)=n+\operatorname{rank}(A(E_n-A)).$$

证 今

$$\operatorname{rank}(A)+\operatorname{rank}(E_n-A)=\operatorname{rank}\left(\begin{pmatrix} E_n-A & 0 \\ 0 & A \end{pmatrix}\right),$$

而

$$\begin{pmatrix} E_n & E_n \\ 0 & E_n \end{pmatrix}\begin{pmatrix} E_n-A & 0 \\ 0 & A \end{pmatrix}\begin{pmatrix} E_n & 0 \\ E_n & E_n \end{pmatrix}=\begin{pmatrix} E_n & A \\ A & A \end{pmatrix},$$

$$\begin{pmatrix} E_n & 0 \\ -A & E_n \end{pmatrix}\begin{pmatrix} E_n & A \\ A & A \end{pmatrix}\begin{pmatrix} E_n & -A \\ 0 & E_n \end{pmatrix}=\begin{pmatrix} E_n & 0 \\ 0 & A-A^2 \end{pmatrix}.$$

所以

$$\operatorname{rank}(A)+\operatorname{rank}(E_n-A)=\operatorname{rank}\left(\begin{pmatrix} E_n & 0 \\ 0 & A-A^2 \end{pmatrix}\right)=n+\operatorname{rank}(A-A^2).\qquad\square$$

例 3.4.4 n 阶方阵 A 称为**幂等方阵**, 如果 $A^2=A$. 试证:

(1) n 阶幂等方阵 A 有 $\operatorname{rank}(A)=\operatorname{tr}(A)$;

(2) A 为 n 阶幂等方阵当且仅当 $\operatorname{rank}(A)+\operatorname{rank}(E_n-A)=n$.

证 设 n 阶方阵 A 的秩为 r. 已知存在 n 阶非异方阵 P 和 Q, 使得

$$A=P\Lambda_rQ,\quad \Lambda=\Lambda_r=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}.$$

所以 $A^2=A$ 等价于 $P\Lambda QP\Lambda Q=P\Lambda Q$, 即 $A^2=A$ 等价于 $\Lambda T\Lambda=\Lambda$, 其中 $T=QP=\begin{pmatrix}T_{11}^{(r)} & T_{12}\\ T_{21} & T_{22}\end{pmatrix}$. 此即 $T_{11}=E_r$. 所以由 $Q=TP^{-1}$, 则 $A^2=A$ 等价于

$$A=P\Lambda Q=P\Lambda TP^{-1}=P\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}E_r & T_{12}\\ T_{21} & T_{22}\end{pmatrix}P^{-1}=P\begin{pmatrix}E_r & T_{12}\\ 0 & 0\end{pmatrix}P^{-1}.$$

今

$$\operatorname{tr}(A)=\operatorname{tr}(P\begin{pmatrix}E_r & T_{12}\\ 0 & 0\end{pmatrix}P^{-1})=\operatorname{tr}(\begin{pmatrix}E_r & T_{12}\\ 0 & 0\end{pmatrix})=r=\operatorname{rank}(\Lambda)=\operatorname{rank}(A).$$

这证明了**(1)** 成立. 由例 3.4.3, **(2)** 自然成立. □

例 3.4.5 试证: n 阶实对称方阵及 Hermite 方阵都有一个阶数等于秩的非零主子式.

证 设 n 阶实对称方阵 A 的秩为 r. 已知存在 n 阶非异方阵 P 和 Q, 使得

$$A=P\Lambda_rQ,\quad \Lambda=\Lambda_r=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}.$$

所以 $A'=A$ 等价于 $P\Lambda Q=Q'\Lambda P'$, 即 $T'\Lambda=\Lambda T$, 其中 $T=P'Q^{-1}$. 记

$$T=\begin{pmatrix}T_{11}^{(r)} & T_{12}\\ T_{21} & T_{22}\end{pmatrix},$$

则有

$$\begin{pmatrix}T_{11}' & T_{21}'\\ T_{12}' & T_{22}'\end{pmatrix}\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}T_{11} & T_{12}\\ T_{21} & T_{22}\end{pmatrix},$$

即

$$\begin{pmatrix}T_{11}' & 0\\ T_{12}' & 0\end{pmatrix}=\begin{pmatrix}T_{11} & T_{12}\\ 0 & 0\end{pmatrix}.$$

此即 $T'_{11}=T_{11}, T_{12}=0, P=Q'\begin{pmatrix} T'_{11} & T'_{21} \\ 0 & T'_{22} \end{pmatrix}$. 所以

$$A=P\Lambda Q=Q'\begin{pmatrix} T'_{11} & T'_{21} \\ 0 & T'_{22} \end{pmatrix}\Lambda Q=Q'\begin{pmatrix} T_{11} & 0 \\ 0 & 0 \end{pmatrix}Q, \quad T'_{11}=T_{11}.$$

即 $A=Q'SQ$, 其中 $S=\begin{pmatrix} T_{11} & 0 \\ 0 & 0 \end{pmatrix}$, $T_{11}=S\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}$ 为 r 阶实对称方阵. 由 $\operatorname{rank}(A)=r$, 所以 $\operatorname{rank}(T_{11})=r$, 即 $\det(T_{11})\neq 0$.

任取 n 阶实对称方阵 A 的 r 阶主子式 $\det A\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}$, $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$. 由 $A=Q'SQ$ 及 Binet–Cauchy 公式,

$$\det A\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}=\sum_{1\leqslant j_1<\cdots<j_r\leqslant n}\ \sum_{1\leqslant k_1<\cdots<k_r\leqslant n}\det Q'\begin{pmatrix} i_1 i_2\cdots i_r \\ j_1 j_2\cdots j_r \end{pmatrix}\cdot$$
$$\det S\begin{pmatrix} j_1 j_2\cdots j_r \\ k_1 k_2\cdots k_r \end{pmatrix}\det Q\begin{pmatrix} k_1 k_2\cdots k_r \\ i_1 i_2\cdots i_r \end{pmatrix}.$$

由 $S=\begin{pmatrix} T_{11} & 0 \\ 0 & 0 \end{pmatrix}$, 所以实对称方阵 S 只有一个 r 阶子式 $\det T_{11}\neq 0$. 因此,

$$\det A\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}=\left(\det Q\begin{pmatrix} 1\ 2\ \cdots\ r \\ i_1 i_2\cdots i_r \end{pmatrix}\right)^2\det S\begin{pmatrix} i_1 i_2\cdots i_r \\ i_1 i_2\cdots i_r \end{pmatrix}.$$

如果实对称方阵 A 的所有 r 阶主子式都等于零, 则推出 $\det Q\begin{pmatrix} 1\ 2\ \cdots\ r \\ i_1 i_2\cdots i_r \end{pmatrix}=0$, $1\leqslant i_1<i_2<\cdots<i_r\leqslant n$. 对行列式 $\det(Q)$ 的前 r 行作 Laplace 展开式可知 $\det(Q)=0$. 这和 Q 是非异方阵矛盾. □

习　　题　3.4

除非特别申明, 下面都是在域 $\mathbb{F}$ 上讨论.

3.4.1　试求下列矩阵的秩:

$$(1)\ \begin{pmatrix} 1 & -2 & 3 & -1 & -1 & -2 \\ 2 & -1 & 1 & 0 & -2 & -2 \\ -2 & -5 & 8 & -4 & 3 & -1 \\ 6 & 0 & -1 & 2 & -7 & -5 \\ -1 & -1 & 1 & -1 & 2 & 1 \end{pmatrix};\quad (2)\ \begin{pmatrix} 3 & 2 & -1 & 2 & 0 & 1 \\ 4 & 1 & 0 & -3 & 0 & 2 \\ 2 & -1 & -2 & 1 & 1 & -3 \\ 3 & 1 & 3 & -9 & -1 & 6 \\ 3 & -1 & -5 & 7 & 2 & -7 \end{pmatrix}.$$

3.4.2　试求下列非异方阵的逆方阵:

(1) $\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$; (2) $\begin{pmatrix}1&2&3&\cdots&n-1&n\\n&1&2&\cdots&n-2&n-1\\n-1&n&1&\cdots&n-3&n-2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\2&3&4&\cdots&n&1\end{pmatrix}$;

(3) $\begin{pmatrix}1&2&3&4\\2&3&1&2\\1&1&1&-1\\1&0&-2&-6\end{pmatrix}$; (4) $\begin{pmatrix}1&1&1&\cdots&1\\1&\omega&\omega^2&\cdots&\omega^{n-1}\\1&\omega^2&\omega^4&\cdots&\omega^{2(n-1)}\\\vdots&\vdots&\vdots&&\vdots\\1&\omega^{n-1}&\omega^{2(n-1)}&\cdots&\omega^{(n-1)(n-1)}\end{pmatrix}$,

其中 $\omega=\exp(2\pi i/n)=\cos(2\pi/n)+\sqrt{-1}\sin(2\pi/n)$.

3.4.3 仅对行作初等变换, 且对列作第一类初等变换, 试证: 可以将任一 $n\times m$ 矩阵 A 化为矩阵

$$\begin{pmatrix}E^{(r)}&B^{(r,m-r)}\\0^{(n-r,r)}&0^{(n-r,m-r)}\end{pmatrix}.$$

3.4.4 试证: 任一秩为 r 的矩阵都可以表示为 r 个秩为 1 的矩阵之和.

3.4.5 n 阶方阵 A 称为**对合的**, 如果 $A^2=E_n$. 试证: A 为对合方阵的必要且充分条件为 $\operatorname{rank}(E_n+A)+\operatorname{rank}(E_n-A)=n$.

3.4.6 设 A 为域 $\mathbb{F}$ 上 n 阶对合方阵. 试证: 存在 n 阶非异方阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}E_r&0\\0&-E_{n-r}\end{pmatrix}.$$

3.4.7 设 A 为域 $\mathbb{F}$ 上 n 阶幂等方阵, 即 $A^2=A$. 试证: 存在 n 阶非异方阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}E_r&0\\0&0\end{pmatrix}.$$

3.4.8 设 A 为域 $\mathbb{F}$ 上 n 阶幂零方阵, 即 $A^2=0$. 试证: 存在 n 阶非异方阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}0&E_r&0\\0&0&0\end{pmatrix}.$$

3.4.9 设 A 为域 $\mathbb{F}$ 上 $n\times m$ 矩阵, B 为 $m\times n$ 矩阵. 试证: $\operatorname{rank}(A)+\operatorname{rank}(B)=m+\operatorname{rank}(AB)$ 当且仅当矩阵方程 $XA-BY=E_m$ 有解. 在有解时, 试求通解.

3.4.10 试证: n 阶方阵 A 不可逆, 即 $\det(A)=0$ 当且仅当存在 n 阶非零方阵 B,C, 使得 $AB=0$, $CA=0$.

3.4.11 试证: 若约定只允许对 $n\times m$ 矩阵 A 的行作初等变换, 对 A 的列作第一类初等变换, 则可将 A 化为 $\begin{pmatrix}E_r&B\\0&0\end{pmatrix}$, 其中 $r=\operatorname{rank}(A)$.

3.4.12 设 A 为域 $\mathbb{F}$ 上 n 阶方阵. 试证: 存在置换方阵 (即第一类初等变换的连续作用) P, 使得 n 阶方阵 $A=PLU$, 其中 U 为上三角方阵, L 为下三角方阵, 对角元素都是 1.

3.4.13 设 A 为域 $\mathbb{F}$ 上 n 阶非异方阵. 试证: 唯一存在 n 阶上三角方阵 U, U_1 和 n 阶下三角方阵 L, L_1, 使得 L 的对角元素都是 1, 而且有分解 $A = LU = LDU_1$ 当且仅当 A 的 n 个顺序主子式:

$$\det S\begin{pmatrix} 12\cdots k \\ 12\cdots k \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}, \quad k = 1, 2, \cdots, n$$

都不等于零, 其中 D 为 n 阶对角方阵

$$D = \operatorname{diag}\left(d_1, \frac{d_2}{d_1}, \cdots, \frac{d_n}{d_{n-1}}\right).$$

3.4.14 设 $\left(\sum\limits_{i=0}^{n} a_i x^i\right)\left(\sum\limits_{j=0}^{n} b_j x^j\right) = 1$, 试证:

$$\det\begin{pmatrix} a_1 & a_0 & & & & \\ a_2 & a_1 & a_0 & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ a_{n-2} & a_{n-3} & a_{n-4} & \cdots & a_0 & \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & a_1 \end{pmatrix} = (-1)^n a_0^{n+1} b_n.$$

3.4.15 用题 1.5.20 的结论, 试证:

$$S_k = \det\begin{pmatrix} \sigma_1 & 1 & & & \\ 2\sigma_2 & \sigma_1 & 1 & & \\ 3\sigma_3 & \sigma_2 & \sigma_1 & \ddots & \\ \vdots & \vdots & \vdots & \ddots & 1 \\ k\sigma_k & \sigma_{k-1} & \cdots & & \sigma_1 \end{pmatrix}, \quad \sigma_k = \frac{1}{k!}\det\begin{pmatrix} S_1 & 1 & & & \\ S_2 & S_1 & 2 & & \\ S_3 & S_2 & S_1 & \ddots & \\ \vdots & \vdots & \vdots & \ddots & k-1 \\ S_k & S_{k-1} & \cdots & & S_1 \end{pmatrix}.$$

3.4.16 设 ω 为多项式 $x^n - 1$ 的本原单位根, $\omega_i = \omega^i$, $1 \leqslant i \leqslant n$. 试证:

$$\prod_{1\leqslant i<j\leqslant n} (\omega_i - \omega_j)^2 = (-1)^{\frac{(n-1)(n+2)}{2}} n^n.$$

3.4.17 设 A 和 B 分别为 $p\times q$ 及 $q\times p$ 矩阵. 试证: 对任一域 $\mathbb{F}$ 中的数 λ, 有 $\lambda^q \det(\lambda E_p - AB) = \lambda^p \det(\lambda E_q - BA)$. 特别, 在 $\lambda = 1$ 时, 有 $\det(E_p - AB) = \det(E_q - BA)$.

3.4.18 给定 $2n\times(m+1)$ 矩阵 $\begin{pmatrix} A^{(n,1)} & B^{(n,m)} \\ 0^{(n,1)} & C^{(n,m)} \end{pmatrix}$. 假设它的秩为 $2n$. 试证: 存在 n 阶非异方阵 P 及 m 阶非异方阵 Q, 使得 $\begin{pmatrix} PA & PBQ \\ 0 & PCQ \end{pmatrix}$ 为下列三情形之一:

(1) $\begin{pmatrix} \begin{pmatrix} 1 \\ 0^{(n-1,1)} \end{pmatrix} & E_n & 0 & 0 \\ 0^{(n,1)} & 0 & E_n & 0 \end{pmatrix}$;

$$\text{(2)}\quad \begin{pmatrix} \begin{pmatrix} 1 \\ 0^{(n-1,1)} \end{pmatrix} & \begin{pmatrix} 0^{(1,n-1)} & 0^{(1,n)} & 0 \\ E_{n-1} & 0^{(n-1,n)} & 0 \end{pmatrix} \\ 0^{(n,1)} & \begin{pmatrix} 0^{(n,n-1)} & E_n & 0 \end{pmatrix} \end{pmatrix};$$

$$\text{(3)}\quad \begin{pmatrix} \begin{pmatrix} 1 \\ 0^{(n-1,1)} \end{pmatrix} & \begin{pmatrix} 1 & 0^{(1,n-2)} & 0^{(1,n)} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & E^{(n-2)} & 0^{(n-2,n)} & 0 \end{pmatrix} \\ 0^{(n,1)} & \begin{pmatrix} 0^{(n,n-1)} & E_n & 0 \end{pmatrix} \end{pmatrix}.$$

§3.5 等价关系

我们知道, 有理数可以用许多不同形式的分数 m/n 表示, 而分数 m/n 又可以用一对整数 (m,n) 来表示, 其中 $n>0$. 记

$$\mathfrak{S} = \{\text{一对整数 } (m,n) \mid m,\ n\in\mathbb{Z},\ n>0\}.$$

两个分数 m_1/n_1 和 m_2/n_2 表示同一有理数的必要且充分条件是 $m_1n_2=m_2n_1$. 这时记作 $m_1/n_1=m_2/n_2$. 这就引出了分数相等的概念. 所有相等的分数表示了同一有理数. 将所有相等的分数放在一起, 构成了集合 $\mathfrak{S}$ 的一个子集合. 于是所有分数按照相等便分成许多子集合, 这些子集合中的不同元素表示了同一有理数, 而不同子集合中的元素表示了不同的有理数. 所以子集合和有理数之间建立了一个自然的一一对应关系.

再例如固定向量的相等, 矩阵的相等, 矩阵的相抵, 数的相等, 多项式的相等等, 都具有分数相等所导出的共性. 将这许许多多种 "相等" 概念抽象化, 便得到

定义 3.5.1 给定一个非空集合 $\mathfrak{S}$. 在集合 $\mathfrak{S}$ 中的元素间定义了一种关系, 记作 $\equiv$. 如果这种关系适合下列三条公理:

(1) 反身性 对集合 $\mathfrak{S}$ 中任一元素 a, 则有 $a\equiv a$;

(2) 对称性 对集合 $\mathfrak{S}$ 中任两元素 a 和 b, 如果 $a\equiv b$, 则有 $b\equiv a$;

(3) 传递性 对集合 $\mathfrak{S}$ 中任三元素 a, b 和 c, 如果 $a\equiv b$, $b\equiv c$, 则有 $a\equiv c$,

这时 $\equiv$ 称为**等价关系**. 在 $a\equiv b$ 时, 称元素 a 和 b **等价**.

由等价关系的定义可知, 集合 $\mathfrak{S}$ 中元素 a 和元素 b 等价, 则 b 和 a 也必等价. 由定义又可知, 上面提到的关系全是等价关系. 下面再举几个不是等价关系的例:

例 3.5.1 实数的不相等关系 $a\neq b$ 和不等关系 $a<b$ 都不适合公理(1); 不等关系 $a<b$ 和顺序关系 $a\leqslant b$ 都不适合公理(2); 不相等关系 $a\neq b$ 不适合公理(3). 这几个例子也证明了公理 (1), 公理 (2) 和公理 (3) 的互相独立性.

例 3.5.2 正实数 a 和 b 如果有 $ab<a+b^2$, 则它定义了 a 和 b 间的一种关系. 这个关系也不是等价关系. 因为公理 (1) 适合 $a^2<a+a^2$, 但不适合公理 (2) 和 (3).

事实上, $ab < a+b^2$ 推不出 $ba < b+a^2$. 例如取 $a=2, b=4$, 则 $8 < 2+4^2$, 但是 $8=4+2^2$. 又 $a=5, b=1, c=2$, 则 $5<5+1^2, 2<1+2^2$, 但是 $10>5+2^2$. 这个例子也证明了公理 (1) 和公理 (2), (3) 互相独立.

例 3.5.3 记 $\mathrm{gl}(n,\mathbb{F})$ 为域 $\mathbb{F}$ 上所有 n 阶方阵构成的集合. 两个方阵 A 和 B 如果有 $AB=BA$, 则它定义了 A 和 B 间的一种关系. 这个关系也不是等价关系. 因为公理(1) 和公理(2) 适合, 但不适合公理(3). 事实上, 对三个 3 阶方阵 A, B 和 C, 如果有 $AB=BA$, $BC=CB$, 但是推不出 $AC=CA$. 例如取 $A=\begin{pmatrix}0&0&1\\1&0&0\\0&0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix}$, $C=\begin{pmatrix}0&0&1\\2&0&0\\0&0&0\end{pmatrix}$, 则 $AB=BA=BC=CB=0$. 所以 A 和 B 有关系, B 和 C 有关系, 但是 $AC\neq CA$, 即 A 和 C 无关系. 这个例子也证明了公理(1), (2) 和公理(3) 互相独立.

等价关系是一种很重要的关系. 由等价关系出发, 可以得到下面一些重要的性质.

定义 3.5.2 在集合 $\mathfrak{S}$ 中引进等价关系 $\equiv$. 在集合 $\mathfrak{S}$ 中任取元素 a, 把和元素 a 等价的元素放在一起, 构成一个子集合, 称为元素 a 的**等价类**, 记作 $\mathfrak{S}_a$. a 称为等价类 $\mathfrak{S}_a$ 的**代表元素**.

引理 3.5.3 在集合 $\mathfrak{S}$ 中引进等价关系 $\equiv$. 以集合 $\mathfrak{S}$ 中任一元素 a 为代表元素的等价类 $\mathfrak{S}_a$ 有性质:

(1) 等价类 $\mathfrak{S}_a$ 中任两元素等价;

(2) 不在等价类 $\mathfrak{S}_a$ 中任一元素和等价类 $\mathfrak{S}_a$ 中任一元素决不等价.

证 在集合 $\mathfrak{S}$ 中任取两元素 b 和 c. 由等价类 $\mathfrak{S}_a$ 的定义可知 $b\equiv a, c\equiv a$. 由对称性, 有 $a\equiv c$. 由传递性, 有 $b\equiv c$. 所以性质(1) 成立.

我们用反证法来证明性质(2) 成立. 设若不然, 即存在集合 $\mathfrak{S}$ 中元素 b 和 c, 其中 $b\in\mathfrak{S}_a$, $c\notin\mathfrak{S}_a$, 又 $b\equiv c$. 由对称性, 有 $c\equiv b$. 由 $b\equiv a$ 及传递性, 有 $c\equiv a$. 由等价类的定义, 有 $c\in\mathfrak{S}_a$. 这和 $c\notin\mathfrak{S}_a$ 矛盾. □

由这个引理可知, 集合 $\mathfrak{S}$ 中元素 a 的等价类 $\mathfrak{S}_a$, 实际上是 $\mathfrak{S}_a$ 中任一元素为代表元素的等价类. 所以等价类的构成和代表元素的选取无关, 而是由等价关系本身完全确定的. 这样一来, 在集合中取出所有等价类, 那末不同的等价类不会相交. 而且所有不同等价类的并就是整个集合, 所以证明了下面的定理.

定理 3.5.4 在集合 $\mathfrak{S}$ 中给定等价关系 $\equiv$ 后, 集合 $\mathfrak{S}$ 便分解为两两不相交的

等价类的并集.

注意 这一分解的特点在于: 同一子集合中任两元素互相等价; 不同子集合中任两元素互不等价. 因此, 在每个等价类中任意取出一个元素, 都可以成为这个等价类的代表元素. 而且将这些元素 (每个等价类集中任意取出一个元素) 合并为另一个子集合, 则这个子集合具有性质: (1) 任两元素互不等价; (2) 集合 $\mathfrak{S}$ 中任一元素必和这个子集合中某个元素等价. 这个子集合称为**代表元素集**.

由上面的讨论可以看出来, 这个代表元集的选取, 即每个等价类中的代表元素的选取, 具有极大的任意性. 事实上, 等价类中任一元素都有资格充当代表元素. 所以在具体的集合中, 如果定义好等价关系, 我们便不满足于随便取出一个元素为代表元素, 而要求取出的元素在某种意义下具有最简单的形式. 为了说明这点, 下面详细讨论矩阵的情形.

在 §3.4 中我们引进了矩阵的相抵关系. 定义 3.4.5 给出了矩阵相抵的定义. 下面证明矩阵相抵的是等价关系.

定理 3.5.5 域 $\mathbb{F}$ 上 $n\times m$ 矩阵全体构成集合 $\mathrm{M}_{n,m}(\mathbb{F})$. 则在集合 $\mathrm{M}_{n,m}(\mathbb{F})$ 中任两个矩阵 A 和 B 的相抵关系是一种等价关系.

证 在集合 $\mathrm{M}_{n,m}(\mathbb{F})$ 中任取两元素 A 和 B. 由 $A=E_nAE_m$, 易证 A 和 A 自身相抵. 又设 A 和 B 相抵, 即存在 n 阶非异方阵 P 和 m 阶非异方阵 Q, 使得 $B=PAQ$. 所以 $A=P^{-1}BQ^{-1}$, 即 B 和 A 相抵. 再设 A 和 B 相抵, 则存在 n 阶非异方阵 P 和 m 阶非异方阵 Q, 使得 $B=PAQ$. 设 B 和 C 相抵, 则存在 n 阶非异方阵 U 和 m 阶非异方阵 V, 使得 $C=UBV$. 因此, $C=UBV=UPAQV$, 其中 UP 为 n 阶非异方阵, QV 为 m 阶非异方阵. 这证明了 A 和 C 相抵. 所以证明了相抵关系是一种等价关系. □

因此域 $\mathbb{F}$ 上 $n\times m$ 矩阵全体构成集合 $\mathrm{M}_{n,m}(\mathbb{F})$ 按照相抵关系分成等价类的并集. 由定理 3.4.7, 两矩阵在同一个等价类中当且仅当它们的秩相等. 所以实际上可以用秩数来完全刻画. 换句话说, 等价类由秩唯一确定. 所以等价类 $\mathrm{M}_{n,m}(\mathbb{F})_A$ 可记作 $\mathrm{M}_{n,m}(\mathbb{F})_r$, 其中 $r=\mathrm{rank}\,(A)$, $r=0, 1, \cdots, \min(n,m)$. 因此

$$\mathrm{M}_{n,m}(\mathbb{F})=\bigcup_{r=0}^{\min(n,m)}\mathrm{M}_{n,m}(\mathbb{F})_r, \tag{3.5.1}$$

为集合 $\mathrm{M}_{n,m}(\mathbb{F})$ 分解为所有等价类的并集公式.

另一方面, 由定理 3.4.6, 只要矩阵 A 的秩为 r, 则 A 必相抵于 $n\times m$ 矩阵

$$\Lambda_r=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}. \tag{3.5.2}$$

所以我们可以取 $n\times m$ 矩阵 Λ_r 为代表元素. 其好处在于矩阵 Λ_r 由秩完全决定, 且

最简单.

从这个例子也可以看出, 我们在等价类中寻找的代表元素具有这样的特点: 它由一批数量来确定, 而且这批数量是同一等价类中每个元素都具有的共性, 即这批数量在等价关系下是不变的. 所以其中每个量都是在等价关系下的**不变量**. 不变量的全体称为**全系不变量**. 例如, 在相抵关系下, 矩阵的秩就是全系不变量. 要注意的是: 单单不变量并不能区分等价类, 即不同的等价类中的元素可以具有相同的不变量. 但是不可能具有相同的全系不变量, 即全系不变量是能完全区分等价类的一批数量. 对矩阵或方阵集合而言, 在其中引进等价关系后, 具有最简形式的代表元素便称为**标准形**. 而矩阵标准形理论的目的就是建立寻找标准形的方法.

因此我们也证明了

定理 3.5.6 域 $\mathbb{F}$ 上 $n\times m$ 矩阵全体构成集合 $\mathrm{M}_{n,m}(\mathbb{F})$, 则其中任两个矩阵 A 和 B 的相抵关系是一种等价关系. 在相抵下的全系不变量是矩阵的秩, 集合 $\mathrm{M}_{n,m}(\mathbb{F})$ 的等价类可记作 $\mathrm{M}_{n,m}(\mathbb{F})_r$. 等价类的并集公式为式 (3.5.1). 而在每个等价类 $\mathrm{M}_{n,m}(\mathbb{F})_r$ 中可取代表元素为 $n\times m$ 矩阵 Λ_r.

下述各种等价关系下的标准形理论, 构成本书的主要内容之一. 我们分情形讨论.

(一) 方阵的情形. 记 $\mathrm{gl}\,(n,\mathbb{F})$ 为域 $\mathbb{F}$ 上所有 n 阶方阵构成的集合. 在其中引进关系 $A\to PAf(P)$, 其中函数 $f(P)$ 是将 n 阶非异方阵 P 用转置运算, 共轭运算和取逆运算作各种可能的组合. 这样可以导出八种关系:

$$PAP,\quad PAP',\quad PA\overline{P},\quad PA\overline{P}',\quad PAP^{-1},\quad PA(P^{-1})',\quad PA\overline{P^{-1}},\quad PA\overline{P^{-1}}'. \tag{3.5.3}$$

易证上面八种关系中, 四种关系 PAP, $PA\overline{P}$, $PA(P^{-1})'$, $PA\overline{P^{-1}}'$ 都不是等价关系, 它们都适合反身性和对称性, 但不适合传递性. 其他四种关系 PAP', $PA\overline{P}'$, PAP^{-1}, $PA\overline{P^{-1}}$ 都是等价关系, 它们分别称为**相合的**、**复相合的**、**相似的**、**复相似的**. 矩阵论的一个重要组成部分是研究这四种等价关系, 求出它们的标准形, 决定它们的全系不变量.

对 n 阶非异方阵 P 加上限制, 我们又可以得到一批新的等价关系. 一般条件为: 设 P 为域 $\mathbb{F}$ 上正交方阵, 它们全体构成集合 $O_n(\mathbb{F})=\{\,O\in\mathrm{gl}\,(n,\mathbb{F})\mid OO'=E_n\,\}$; 设 U 为复数域 $\mathbb{C}$ 上酉方阵, 它们全体构成集合 $U_n=\{\,U\in\mathrm{gl}\,(n,\mathbb{C})\mid U\overline{U}'=E_n\,\}$.

设 $P=U$ 为酉方阵时, 则等价关系 $A\to UA\overline{U}'$ 和等价关系 $A\to UAU^{-1}$ 相同, 又等价关系 $A\to UA\overline{U}^{-1}$ 和等价关系 $A\to UAU'$ 相同. 所以只有两种等价关系: $A\to UAU'$, $A\to UAU^{-1}$, 前者称为**酉相合的**, 后者称为**酉相似的**. 设 $P=O$ 为复正交方阵时, 则等价关系 $A\to OAO'$ 和等价关系 $A\to OAO^{-1}$ 相同, 又 $A\to OA\overline{O}^{-1}$ 和等价关系 $A\to OA\overline{O}'$ 相同, 所以有两种等价关系: $A\to OAO^{-1}$, $A\to OA\overline{O}'$, 前者称为**复正交相似的**, 后者称为**复正交复相合的**. 由于酉方阵和复正交方阵限制在实

的情形, 都是实正交方阵. 所以对 n 阶非异实方阵加上条件为实正交方阵, 改记为 $P=O$. 因此等价关系 $A\to OAO'$ 和 $A\to QAQ^{-1}$ 相同. 所以只有一种等价关系: $A\to OAO^{-1}$, 称为**实正交相似的**.

记 $\mathrm{gl}\,(n,\mathbb{R})$ 为所有 n 阶实方阵构成的集合. 则只有两种等价关系 $A\to PAP'$, $A\to PAP^{-1}$, 其中 P 为 n 阶非异实方阵, 前者称为**实相合的**, 后者称为**相似的**.

(二) 矩阵的情形. 设集合 $\mathrm{M}_{n,m}(\mathbb{C})$ 为所有 $n\times m$ 复矩阵构成的集合. 我们只考虑两种等价关系: $A\to PAQ$, $A\to UAV$, 其中 P,Q 分别为 n 阶及 m 阶复非异方阵, U,V 分别为 n 阶及 m 阶酉方阵. 前者称为**相抵的**, 后者称为**酉相抵的**. 设集合 $\mathrm{M}_{n,m}(\mathbb{R})$ 为所有 $n\times m$ 实矩阵构成的集合. 我们也只考虑两种等价关系: $A\to PAQ$, $A\to O_1AO_2$, 其中 P,Q 分别为 n 阶及 m 阶实非异方阵, O_1, O_2 分别为 n 阶及 m 阶实正交方阵, 前者称为**实相抵的**, 后者称为**实正交相抵的**.

本书的主要内容之一就是讨论矩阵和方阵在各种等价关系下的标准形理论.

习 题 3.5

3.5.1 下列关系是否为等价关系:

(1) 方阵的交换关系;

(2) 给定区间 $[a,b]$ 上的连续函数 $f(x)$. 区间 $[a,b]$ 中的实数 x_1 和 x_2 称为有关系的, 如果 $f(x_1)=f(x_2)$.

3.5.2 设集合 $\mathfrak{S}$ 中有一种关系 $\to$, 它具有对称性和传递性, 试证明下列推理的不正确性: "今由对称性, 从 $a\to b$ 可知 $b\to a$ 成立, 于是有 $a\to b$, $b\to a$. 再由传递性可知 $a\to a$. 所以反身性成立."

3.5.3 为什么说等价关系是数的相等概念的推广.

第四章　线性方程组理论

§4.1　非齐次线性方程组

m 个未知数 $x_1, x_2, \cdots, x_m$ 和 n 个线性方程构成的域 $\mathbb{F}$ 上线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2, \\ \qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases} \tag{4.1.1}$$

称为**非齐次线性方程组**. 当 $b_1 = b_2 = \cdots = b_n = 0$ 时, 称为**齐次线性方程组**. 它们统称为**线性方程组**.

由线性方程组 (4.1.1) 的 nm 个**系数** $a_{jk} \in \mathbb{F}$, n 个**常数项** $b_i \in \mathbb{F}$ 及 m 个 **未知数** x_j 构成下面四个重要的矩阵, 它们分别是域 F 上 $n \times 1$, $m \times 1$, $n \times m$, $n \times (m+1)$ 矩阵

$$\beta = (b_1, b_2, \cdots, b_n)', \quad x = (x_1, x_2, \cdots, x_m)', \tag{4.1.2}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}, \ B = \begin{pmatrix} A \, \beta \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{pmatrix}. \tag{4.1.3}$$

矩阵 A, β, x 和 B 分别称为线性方程组 (4.1.1) 的**系数矩阵**, **常向量**, **变向量**, **增广矩阵**. 由矩阵乘法的定义可知, 这时, 线性方程组 (4.1.1) 可简洁地写为矩阵形式

$$Ax = \beta. \tag{4.1.4}$$

定义 4.1.1 域 $\mathbb{F}$ 上 $m\times 1$ 矩阵 $x_0=(x_1^{(0)},x_2^{(0)},\cdots,x_m^{(0)})'$ 称为线性方程组 (4.1.1) 的一组**解**, 如果

$$Ax_0=\beta. \tag{4.1.5}$$

线性方程组 (4.1.1) 的所有解构成的集合称为线性方程组 (4.1.1) 的**通解**.

这一章的主要目的, 是要给出判别线性方程组 (4.1.1) 有解的必要且充分条件. 在有解时, 还要进一步求出它的通解. 所运用的方法就是**Gauss 消去法**. 在叙述 Gauss 消去法的原理以前, 先引进一个重要的概念.

定义 4.1.2 域 $\mathbb{F}$ 上 m 个有序的数 $x_1, x_2, \cdots, x_m$ 的两个线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=b_2,\\ \qquad\qquad\vdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m=b_n \end{cases} \tag{4.1.6}$$

和

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1m}x_m=c_1,\\ c_{21}x_1+c_{22}x_2+\cdots+c_{2m}x_m=c_2,\\ \qquad\qquad\vdots\\ c_{p1}x_1+c_{p2}x_2+\cdots+c_{pm}x_m=c_p \end{cases} \tag{4.1.7}$$

称为**等价的**, 如果它们的通解相同. 换句话说, $x=x_0$ 是线性方程组 (4.1.6) 的解当且仅当 $x=x_0$ 是线性方程组 (4.1.7) 的解.

又如果我们将线性方程组 (4.1.6) 的未知数 $x_1, x_2, \cdots, x_m$ 换一个次序重排为 $x_{i_1}, x_{i_2}, \cdots, x_{i_m}$, 其中 $i_1i_2\cdots i_m$ 是 1, 2, $\cdots$, m 的排列. 记 $y_1=x_{i_1}$, $y_2=x_{i_2}$, $\cdots$, $y_m=x_{i_m}$, 于是未知数 $y_1, y_2, \cdots, y_m$ 适合一个新的线性方程组, 它的增广矩阵是由线性方程组 (4.1.6) 的增广矩阵的前 m 列按照排列 $i_1i_2\cdots,i_m$ 重排而得, 这可以由对线性方程组 (4.1.6) 的增广矩阵 B 的前 m 列作一系列第一类列的初等变换而得. 自然, 这时新旧线性方程组的通解可以互相求出. 为方便起见, 这时也称这两个线性方程组等价.

显然线性方程组的等价是一种等价关系. 所以, 从求解的角度来研究线性方程组, 那末, 考虑一个给定的线性方程组, 便等于考虑和它等价的所有线性方程组. 而**Gauss 消去法**的原理就是: 从一个已给的线性方程组出发, 按照确定的规律, 即消去一个未知数的方法, 造出一个和已给线性方程组等价的线性方程组, 再考虑这个新的线性方程组的通解. 确切地说, 给定线性方程组 (4.1.6), 设 $a_{11}\neq 0$, 第一个方程遍乘以 $\dfrac{a_{21}}{a_{11}}$, 然后从第二个方程减去第一个方程, 因此得到一个不包含未知数 x_1 的

线性方程

$$a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2m}^{(1)}x_m = b_2^{(1)}.$$

用同样的方法处理第三个线性方程, $\cdots$, 第 n 个线性方程, 便造出一个新的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1, \\ \qquad\quad a_{22}^{(1)}x_2 + \cdots + a_{2m}^{(1)}x_m = b_2^{(1)}, \\ \qquad\qquad\qquad \vdots \\ \qquad\quad a_{n2}^{(1)}x_2 + \cdots + a_{nm}^{(1)}x_m = b_n^{(1)}, \end{cases} \tag{4.1.8}$$

其中

$$a_{jk}^{(1)} = a_{jk} - \frac{a_{j1}a_{1k}}{a_{11}}, \quad b_j^{(1)} = b_j - b_1\frac{a_{j1}}{a_{11}}, \quad j = 2, \cdots, n, \quad k = 2, \cdots, m. \tag{4.1.9}$$

从矩阵的角度来看, 式 (4.1.9) 就是对增广矩阵 $B = (A, \beta)$ 的第二行, 第三行, $\cdots$, 第 n 行作了行的第三类初等变换.

对于线性方程组 (4.1.8), 可以先求 $m-1$ 个未知数 $x_2, x_3, \cdots, x_m$ 和 $n-1$ 个方程线性构成的线性方程组

$$\begin{cases} a_{22}^{(1)}x_2 + \cdots + a_{2m}^{(1)}x_m = b_2^{(1)}, \\ \qquad\qquad \vdots \\ a_{n2}^{(1)}x_2 + \cdots + a_{nm}^{(1)}x_m = b_n^{(1)} \end{cases} \tag{4.1.10}$$

的通解, 再利用线性方程组 (4.1.8) 的第一个线性方程

$$x_1 = a_{11}^{-1}(b_1 - a_{12}x_2 - \cdots - a_{1m}x_m), \tag{4.1.11}$$

求出线性方程组 (4.1.8) 的通解. 假设 $a_{22}^{(1)} \neq 0$, 对于线性方程组 (4.1.10), 再运用 Gauss 消去法, 这样又可以消去一个变数, 得到和线性方程组 (4.1.8) 等价的线性方程组. 用同样的方法依次作下去, 最后得到一个比线性方程组 (4.1.1) 简单得多的线性方程组, 它们间彼此等价. 对于新的线性方程组可以直接求解, 从而求出原线性方程组 (4.1.1) 的通解.

为了进一步熟悉消去法, 下面给出两个例子.

例 4.1.1 求下面线性方程组的通解.

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 = 0, \\ x_1 + x_2 + 9x_3 = 2, \\ x_1 + 2x_2 - 6x_3 = 1. \end{cases}$$

将第二个线性方程减去第一个线性方程遍乘以 1/2, 于是消去了未知数 x_1, 从而得到线性方程 $-\dfrac{1}{2}x_2+7x_3=2$, 再将第三个线性方程减去第一个线性方程遍乘以 1/2, 于是得到线性方程 $\dfrac{1}{2}x_2-8x_3=1$. 因此求得等价的线性方程组

$$\begin{cases} 2x_1+3x_2+4x_3= & 0, \\ \quad -\dfrac{1}{2}x_2+7x_3= & 2, \\ \quad \dfrac{1}{2}x_2-8x_3= & 1. \end{cases}$$

再将第三个线性方程加上第二个线性方程, 于是消去了未知数 x_2, 从而得到等价的线性方程组

$$\begin{cases} 2x_1+3x_2+4x_3= & 0, \\ \quad -\dfrac{1}{2}x_2+7x_3= & 2, \\ \qquad\qquad -x_3= & 3. \end{cases}$$

这个线性方程组的任一组解 $(x_1^{(0)},x_2^{(0)},x_3^{(0)})'$ 首先要适合第三个线性方程, 即 $-x_3^{(0)}=3$, 因此 $x_3^{(0)}=-3$. 其次 $(x_1^{(0)},x_2^{(0)},x_3^{(0)})'$ 要适合第二个线性方程, 即 $-x_2^{(0)}+14x_3^{(0)}=4$, 因此 $x_2^{(0)}=14x_3^{(0)}-4=-46$; 最后 $(x_1^{(0)},x_2^{(0)},x_3^{(0)})'$ 要适合第一个线性方程, 即 $2x_1^{(0)}+3x_2^{(0)}+4x_3^{(0)}=0$. 因此 $x_1^{(0)}=-\dfrac{3}{2}x_2^{(0)}-2x_3^{(0)}=75$. 所以, 利用 Gauss 消去法, 求得例 4.1.1 的所有解, 它只有一组, 为

$$x_1^{(0)}=75, \quad x_2^{(0)}=-46, \quad x_3^{(0)}=-3. \qquad \square$$

例 4.1.2 求下面线性方程组的通解.

$$\begin{cases} 2x_1+3x_2+4x_3=0, \\ \quad x_1+x_2+9x_3=2, \\ 7x_1+11x_2+7x_3=0. \end{cases}$$

将第二个方程乘 2 减去第一个方程, 再将第三个方程乘 2, 第一个方程乘 7, 然后用前者减后者, 于是得到等价的方程组

$$\begin{cases} 2x_1+3x_2+4x_3=0, \\ \quad -x_2+14x_3=4, \\ \quad x_2-14x_3=0. \end{cases}$$

今若方程组有解 $x^{(0)}=(x_1^{(0)},x_2^{(0)},x_3^{(0)})'$, 则它必须适合第二个方程和第三个方程, 即

$$\begin{cases} -x_2^{(0)}+14x_3^{(0)} = 4, \\ x_2^{(0)}-14x_3^{(0)} = 0. \end{cases}$$

因此 $0=-4$. 这是不可能的, 所以证明了方程组没有解.

注意 由这两个例子可以看出, 我们可以用 Gauss 消去法判断方程组是否有解, 在有解时还可以求出它的通解. 另一方面, 在求得通解后, 需要将解代入原线性方程组, 验证它们是否确实是解. 这一步骤是必要的, 原因在于运用消去法求通解的过程中, 需要作一系列的运算, 这往往容易产生笔误. 经过验算才能避免产生这样的错误.

现在开始, 我们来建立线性方程组求解的严格的理论. 显然有

定理 4.1.3 给定域 $\mathbb{F}$ 上 $n\times m$ 矩阵 A, $n\times 1$ 矩阵 β, 记 x 为 m 个未知数 $x_1, x_2, \cdots, x_m$ 构成的 $m\times 1$ 变向量. 任取 $P\in \mathrm{gl}(n,\mathbb{F})$, Q 是 m 阶置换方阵, 则线性方程组 $Ax=\beta$ 和线性方程组 $PAQy=P\beta=\gamma$ 等价, 其中 $x=Qy$. 又线性方程组 $Ax=\beta$ 的增广矩阵 $B=(A,\beta)$ 和线性方程组 $PAQy=\gamma$ 的增广矩阵 $\widetilde{B}=(PAQ,\gamma)=(PAQ,P\beta)$ 有

$$\operatorname{rank}(A)=\operatorname{rank}(PAQ)=r,\quad \operatorname{rank}(B)=\operatorname{rank}(\widetilde{B})\geqslant r. \tag{4.1.12}$$

不难证明

引理 4.1.4 给定域 $\mathbb{F}$ 上秩为 r 的 $n\times m$ 矩阵 A 及 $n\times 1$ 矩阵 β, 则存在 $P\in \mathrm{gl}(n,\mathbb{F})$ 及 m 阶置换方阵 Q, 使得

$$PAQ=\begin{pmatrix} E_r & A_0 \\ 0 & 0 \end{pmatrix},\quad \gamma=P\beta=\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \tag{4.1.13}$$

其中 $\gamma_1=(c_1,\cdots,c_r)'$, $\gamma_2=(c_{r+1},\cdots,c_n)'$.

因此有

定理 4.1.5 线性方程组 $Ax=\beta$ 等价于线性方程组

$$\begin{pmatrix} E_r & A_0 \\ 0 & 0 \end{pmatrix} y=\gamma=\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}.$$

即记 $y=(y_1,\cdots,y_m)'$, $z=(y_1,\cdots,y_r)'$, $w=(y_{r+1},\cdots,y_m)'$, 则

$$z+A_0w=\gamma_1,\qquad 0=\gamma_2. \tag{4.1.14}$$

因此, 线性方程组 $Ax=\beta$ 有解当且仅当 $\gamma_2=0$. 这时通解为

$$y=\begin{pmatrix} z \\ w \end{pmatrix}=\begin{pmatrix} \gamma_1 \\ 0 \end{pmatrix}+\begin{pmatrix} -A_0t \\ t \end{pmatrix}=\begin{pmatrix} \gamma_1 \\ 0 \end{pmatrix}+\sum_{k=r+1}^{m} t_k\xi_k, \tag{4.1.15}$$

其中 $t=(t_{r+1},\cdots,t_m)'$, $\xi_k=(b_{k,1},\cdots,b_{kr},\delta_{k,r+1},\cdots,\delta_{km})'$, $k=r+1,r+2,\cdots,m$, δ_{jk} 为 Kronecker 符号.

定理 4.1.6 (线性方程组的相容性定理) 线性方程组 $Ax=\beta$ 有解的必要且充分条件为其系数矩阵 A 和增广矩阵 $B=(A,\beta)$ 的秩相等.

证 令线性方程组 $Ax=\beta$ 有解当且仅当线性方程组 $PAQy=\gamma$ 有解, 即 $\gamma_2=0$.

先证必要性. 设 $\gamma_2=0$, 则 $\operatorname{rank}(\widetilde{B})=r=\operatorname{rank}(\widetilde{A})$. 所以 $\operatorname{rank}(B)=\operatorname{rank}(A)=r$.

再证充分性. 即设 $\operatorname{rank}(A)=\operatorname{rank}(B)$, 所以 $\operatorname{rank}(\widetilde{B})=r=\operatorname{rank}(\widetilde{A})$. 因为 $\operatorname{rank}(\widetilde{A})=r$, 所以这时 $\widetilde{B}$ 的任一 $r+1$ 阶子式为零. 特别, $\det\begin{pmatrix}E_r & \gamma_1\\ 0 & c_j\end{pmatrix}=c_j=0$, $j=r+1,\cdots,n$. 即 $\gamma_2=0$. 所以线性方程组 $Ax=\beta$ 有解. □

由定理 4.1.5, 我们证明了

定理 4.1.7 设线性方程组 $Ax=\beta$ 有解, 则在 $m\geqslant r=\operatorname{rank}(A)$ 时, 它的通解为 (4.1.15), 它依赖于 $m-r$ 个独立的参数. 特别, 在 $m=r$ 时, 解唯一.

这样就建立了完整的线性方程组求解理论. 由整个讨论可以看出, 在具体处理一个线性方程组时, 并不需要先利用相容性定理去判断它是否有解, 然后再去求解. 而是从一开始就运用 Gauss 消去法, 到不能再消去未知数时, 实际上就能一眼看出它有没有解, 而且同时可以立刻写出它的通解. 这一事实反映了消去法的作用.

习　题　4.1

4.1.1 试判断下列线性方程组是否有解, 在有解时试求出它的通解:

(1) $\begin{cases} x_1+x_2-3x_3=-1,\\ 2x_1+x_2-2x_3=1,\\ x_1+x_2+x_3=3,\\ x_1+2x_2-3x_3=1;\end{cases}$ (2) $\begin{cases} \lambda x_1+x_2+x_3=1,\\ x_1+\lambda x_2+x_3=1,\\ x_1+x_2+\lambda x_3=1;\end{cases}$

(3) $\begin{cases} a_1x_1+a_2x_2+a_3x_3=c,\\ a_3x_2-a_2x_3=b_1,\\ -a_3x_1+a_1x_3=b_2,\\ a_2x_1-a_1x_2=b_3;\end{cases}$ (4) $\begin{cases} x_1+3x_2+5x_3-4x_4=1,\\ x_1+3x_2+2x_3-2x_4+x_5=-1,\\ x_1-2x_2+x_3-x_4-x_5=3,\\ x_1-4x_2+x_3+x_4-x_5=3,\\ x_1+2x_2+x_3-x_4+x_5=-1.\end{cases}$

4.1.2 在平面上引进直角坐标系, 试求:

(1) 直线 $a_1x+b_1y+c_1=0$ 和 $a_2x+b_2y+c_2=0$ 的交点;

(2) 试求 p 条直线 $a_jx+b_jy+c_j=0$, $1\leqslant j\leqslant p$ 有公共交点的必要且充分条件;

(3) 求四点 (a_i, b_i), $i = 1, 2, 3, 4$ 共圆的必要且充分条件.

4.1.3 在空间中引进直角坐标系. 试求 p 个平面 $A_j x + B_j y + C_j z + D_j = 0$, $1 \leqslant j \leqslant p$ 有公共交点的必要且充分条件, 再求有公共交线的必要且充分条件.

4.1.4 试求整数 a, 使得线性方程组

$$\sum_{k=1}^{5} k x_k = a, \quad \sum_{k=1}^{5} k^3 x_k = a^2, \quad \sum_{k=1}^{5} k^5 x_k = a^3$$

有非负实数解. 试求所有非负实数解.

4.1.5 方程组

$$x + y + z = a, \quad x^2 + y^2 + z^2 = b^2, \quad xy = z^2$$

中 a, b 适合什么条件时, 它的解必为三个不同正数.

4.1.6 试求 4 个实数 x_1, x_2, x_3, x_4 适合方程组

$$x_1 + x_2 x_3 x_4 = 2, \quad x_2 + x_3 x_4 x_1 = 2, \quad x_3 + x_4 x_1 x_2 = 2, \quad x_4 + x_1 x_2 x_3 = 2.$$

4.1.7 试解线性方程组

$$x_5 + x_2 = y x_1, \quad x_1 + x_3 = y x_2, \quad x_2 + x_4 = y x_3, \quad x_3 + x_5 = y x_4, \quad x_4 + x_1 = y x_5.$$

4.1.8 试证: 方程组

$$\begin{cases} \dfrac{x^2}{2^2 - 1^2} + \dfrac{y^2}{2^2 - 3^2} + \dfrac{z^2}{2^2 - 5^2} + \dfrac{w^2}{2^2 - 7^2} = 1, \\ \dfrac{x^2}{4^2 - 1^2} + \dfrac{y^2}{4^2 - 3^2} + \dfrac{z^2}{4^2 - 5^2} + \dfrac{w^2}{4^2 - 7^2} = 1, \\ \dfrac{x^2}{6^2 - 1^2} + \dfrac{y^2}{6^2 - 3^2} + \dfrac{z^2}{6^2 - 5^2} + \dfrac{w^2}{6^2 - 7^2} = 1, \\ \dfrac{x^2}{8^2 - 1^2} + \dfrac{y^2}{8^2 - 3^2} + \dfrac{z^2}{8^2 - 5^2} + \dfrac{w^2}{8^2 - 7^2} = 1 \end{cases}$$

的解 $x = \alpha$, $y = \beta$, $z = \gamma$, $w = \delta$ 必有 $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 36$.

4.1.9 试用线性方程组的矩阵表达式 (4.1.4) 以及 §3.4 的方法, 用矩阵语言, 给出线性方程组的相容性定理的证明以及通解的表达式.

4.1.10 设 A 为 $n \times (n+1)$ 常数矩阵, X 为 $(n+1) \times n$ 矩阵, 其元素由独立未知数构成. 试证: 矩阵方程 $AX = E_n$ 有解的必要且充分条件为 $\operatorname{rank}(A) = n$, 其中 E_n 为 n 阶单位方阵.

4.1.11 设 A 为 n 阶整数方阵, β 为 $n \times 1$ 整数矩阵. 试证: 对任意的 $n \times 1$ 整数矩阵 β, 矩阵方程 $Ax = \beta$ 都有解的必要且充分条件为 $\det(A) = \pm 1$.

4.1.12 试证: 域 $\mathbb{F}$ 上线性方程组 $Ax = \beta$ 有解的充分条件为矩阵 $\begin{pmatrix} A & \beta \\ \beta' & 0 \end{pmatrix}$ 和矩阵 A 的秩相等. 它是不是必要条件, 为什么?

4.1.13 设 A 和 B 分别为域 $\mathbb{F}$ 上 $n \times m$, $n \times p$ 矩阵, X 为由 mp 个独立未知数构成的 $m \times p$ 矩阵. 试证: 矩阵方程 $AX = B$ 有解的必要且充分条件为 $\operatorname{rank}(A) = \operatorname{rank}((A, B))$. 什么时候解唯一存在?

4.1.14 设

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

为 n 阶实方阵.

(1) 设 A 适合条件 $a_{ij} \geqslant 0$, $1 \leqslant i, j \leqslant n$, $\sum\limits_{j=1}^{n} a_{ij} < 1$, $1 \leqslant i \leqslant n$. 试证: 方阵 $E_n - A$ 非异, 而且 $(E - A)^{-1}$ 中每个元素非负;

(2) 设 A 适合条件 $2|a_{ii}| > \sum\limits_{j=1}^{n} |a_{ij}|$, $1 \leqslant i \leqslant n$. 试证: $\det(A) \neq 0$, 这时线性方程组 $Ax = \beta$ 的解唯一存在, 其中 $\beta = (b_1, b_2, \cdots, b_n)'$.

4.1.15 试证: 域 $\mathbb{F}$ 上线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2, \\ \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases}$$

有解的必要且充分条件为对任意 n 个数 $y_1, y_2, \cdots, y_n$, 只要 $\sum\limits_{j=1}^{n} y_j a_{jk} = 0$, $1 \leqslant k \leqslant m$, 便有 $\sum\limits_{j=1}^{n} y_j b_j = 0$.

4.1.16 设一组实数 $a_1, a_2, \cdots, a_n$ 有 $a_i + a_j \geqslant 0$, $i \neq j$. 设 $x_1 \geqslant 0, \cdots, x_n \geqslant 0$, $x_1 + \cdots + x_n = 1$, 则有 $\sum\limits_{i=1}^{n} a_i x_i \geqslant \sum\limits_{i=1}^{n} a_i x_i^2$.

4.1.17 任取正整数 n, 求解

$$\sum_{i=1}^{n} x_i = n, \quad \sum_{i=1}^{n} x_i^2 = n, \quad \cdots, \quad \sum_{i=1}^{n} x_i^n = n.$$

§4.2 齐次线性方程组

这一节考虑的是域 $\mathbb{F}$ 上 m 个未知数, n 个方程构成的齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0, \\ \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0 \end{cases} \tag{4.2.1}$$

的解. 这时, 系数矩阵 A 和增广矩阵 B 的差别在于多添加了一列零, 所以它们的秩总是相等的. 由相容性定理, 齐次线性方程组一定有解, 实际上, $x_1 = 0$, $x_2 = 0$, $\cdots$,

$x_m = 0$ 就是它的一组解, 这组解称为齐次线性方程组的**零解**. 若还有其他的解, 则其他的解称为齐次线性方程组的**非零解**.

在齐次线性方程组的情形, 重要在于求出所有非零解, 并研究解构成的集合. 即对这类线性方程组, 要考虑什么时候只有零解, 什么时候有非零解? 通解如何?

由式 (4.1.15) 可知, 设齐次线性方程组的系数矩阵 A 的秩为 r, 则齐次线性方程组 $Ax = 0$ 的通解可表为

$$y = \sum_{k=r+1}^{m} t_k \xi_k = \sum_{k=r+1}^{m} t_k (b_{k,1}, \cdots, b_{kr}, \delta_{k,r+1}, \cdots, \delta_{km})', \tag{4.2.2}$$

其中 $t_{r+1}, t_{r+2}, \cdots, t_m$ 是 $m-r$ 个独立参数, δ_{jk} 为 Kronecker 符号.

这就证明了

定理 4.2.1 (齐次线性方程组解的结构定理) m 个未知数 $x_1, x_2, \cdots, x_m$, n 个线性方程的齐次线性方程组 (4.2.1) 必有零解. 齐次线性方程组 (4.2.1) 有非零解的必要且充分条件为其系数矩阵 A 的秩 $r < m$, 而且通解有无穷多个解, 它依赖于 $m-r$ 个独立参数 $t_{r+1}, t_{r+2}, \cdots, t_m$, 使得通解可表为式 (4.2.2). 即我们记 $m \times 1$ 矩阵 $x_0 = (\, x_1^{(0)}, x_2^{(0)}, \cdots, x_m^{(0)})'$, 则通解可以写成矩阵的形式:

$$x_0 = t_{r+1}\beta_{r+1} + t_{r+2}\beta_{r+2} + \cdots + t_m\beta_m. \tag{4.2.3}$$

由此可见, 齐次线性方程组有 $m-r$ 组解, 它分别是 $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_m$. 而通解和这 $m-r$ 组解的关系由式 (4.2.3) 给出. 我们称这 $m-r$ 组解 $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_m$ 是齐次线性方程组 (4.2.1) 的**基础解系**.

定理 4.2.2 (非齐次线性方程组解的结构定理) 给定域 $\mathbb{F}$ 上 $n \times m$ 矩阵 A 及 $n \times 1$ 矩阵 β. 记 x 为 m 个未知数 $x_1, x_2, \cdots, x_m$ 构成的 $m \times 1$ 变向量. 则非齐次线性方程组 $Ax = \beta$ 有相伴的齐次线性方程组 $Ax = 0$. 它们之间的通解有下面密切的关系:

(1) 非齐次线性方程组 $Ax = \beta$ 的两解之差为它相伴的齐次线性方程组 $Ax = 0$ 的解;

(2) 非齐次线性方程组 $Ax = \beta$ 的解及它相伴的齐次线性方程组 $Ax = 0$ 的解之和是非齐次线性方程组 $Ax = \beta$ 的解;

(3) 非齐次线性方程组 $Ax = \beta$ 的通解由它相伴的齐次线性方程组 $Ax = 0$ 的通解加上非齐次线性方程组 $Ax = \beta$ 的一个**特解**构成.

证 记 $\mathfrak{S}$ 为非齐次线性方程组 $Ax = \beta$ 的所有解构成的集合, $\mathfrak{S}_0$ 为它相伴的齐次线性方程组 $Ax = 0$ 的所有解构成的集合, 目的是要证明: 任意取定 $\xi \in \mathfrak{S}$, 则

$$\mathfrak{S} = \xi + \mathfrak{S}_0 = \{\, \xi + \xi_0 \mid \forall \xi_0 \in \mathfrak{S}_0 \,\}.$$

任取 $\xi_1, \xi_2 \in \mathfrak{S}$, 则 $A\xi_1 = \beta$, $A\xi_2 = \beta$. 所以 $A(\xi_1 - \xi_2) = 0$, 即 $\xi_1 - \xi_2 \in \mathfrak{S}_0$. 这就证明了(1) 成立; 任取 $\xi \in \mathfrak{S}$, $\xi_0 \in \mathfrak{S}_0$, 则 $A\xi = \beta$, $A\xi_0 = 0$. 所以 $A(\xi + \xi_0) = \beta$, 即 $\xi + \xi_0 \in \mathfrak{S}$. 这就证明了(2) 成立; 至此也证明了 $\xi + \mathfrak{S}_0 \subset \mathfrak{S}$.

最后, 我们要证 $\mathfrak{S} \subset \xi + \mathfrak{S}_0$, 其中 $\xi \in \mathfrak{S}$. 由性质(1), 任取 $\xi_1 \in \mathfrak{S}$, 则 $\xi_1 - \xi \in \mathfrak{S}_0$, 即有 $\xi_0 \in \mathfrak{S}_0$, 使得 $\xi_1 - \xi = \xi_0$. 因此, $\xi_1 = \xi + \xi_0 \in \xi + \mathfrak{S}_0$. 这就证明了 $\mathfrak{S} \subset \xi + \mathfrak{S}_0$. 所以 $\mathfrak{S} = \xi + \mathfrak{S}_0$, 即(3) 成立. □

习 题 4.2

4.2.1 求下列齐次线性方程组的基础解系:

$$\begin{cases} x_1 + x_2 - 3x_4 - x_5 = 0, \\ x_1 - x_2 - 2x_3 - x_4 = 0, \\ 4x_1 - 2x_2 + 6x_3 + 3x_4 - 4x_5 = 0, \\ 2x_1 + 4x_2 - 2x_3 + 4x_4 - 7x_5 = 0; \end{cases} \qquad \begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 0, \\ 2x_1 + 2x_2 + x_3 + x_4 - 2x_5 = 0, \\ x_2 + 2x_3 + 2x_4 + x_5 = 0, \\ 5x_1 + 4x_2 - 3x_3 + 4x_4 - x_5 = 0. \end{cases}$$

4.2.2 求非齐次线性方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 4, \\ x_2 - x_3 + x_4 = -3, \\ x_1 + 3x_2 - 3x_4 = 1, \\ -7x_2 + 3x_3 + x_4 = -3 \end{cases}$$

的通解.

4.2.3 设齐次线性方程组 $\sum\limits_{k=1}^{n+1} a_{jk}x_k = 0$, $j = 1, 2, \cdots, n$ 的系数矩阵的秩为 n. 试证: 其通解为

$$x_k = t\det\begin{pmatrix} a_{11} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1,n+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{n,n+1} \end{pmatrix}, \quad k = 1, 2, \cdots, n+1,$$

其中 t 为参数.

4.2.4 在域 $\mathbb{F}$ 上取定 n 个不全为零的常数 $a_1, a_2, \cdots, a_n$, 求 n^2 个未知数 x_{jk}, $1 \leqslant j, k \leqslant n$ 的如下线性方程组的通解: $x_{jk}a_ia_p - x_{ip}a_ja_k = 0$, $i, j, k, p = 1, 2, \cdots, n$.

4.2.5 任取 $2n^2$ 个数 $a_{ij} \in \{0, 1, -1\}$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant 2n$. 试证: 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,2n}x_{2n} = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2,2n}x_{2n} = 0, \\ \qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{n,2n}x_{2n} = 0 \end{cases}$$

有不全为零的适合条件 $|c_j| \leqslant n$, $1 \leqslant j \leqslant 2n$ 的一组整数解 $c_1, c_2, \cdots, c_{2n}$.

§4.3　方阵的特征根

作为线性方程组理论的应用, 下面引进方阵的三个重要概念: 特征多项式, 特征根和特征向量. 这些概念和理论刻画了方阵的许多本质的特性, 而且在几何学、函数论、微分方程、应用数学、力学、天文学、物理学等方面都有着重要的应用.

定义 4.3.1　域 $\mathbb{F}$ 上 n 阶方阵

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

的**特征多项式**是以 λ 为未知数的 n 次多项式

$$\det(\lambda E_n - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n. \tag{4.3.1}$$

域 $\mathbb{F}$ 上多项式 $\det(\lambda E_n - A)$ 恰有 n 个复数根 $\lambda_1, \lambda_2, \cdots, \lambda_n$ (可能有相同的), 称为方阵 A 的**特征根**. 如果 A 的特征根 λ_j 是实数, 则称 λ_j 为方阵 A 的**实特征根**.

设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是方阵 A 的特征根 (可能是复数), 由定义可知

$$\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n). \tag{4.3.2}$$

再由根与系数的关系可知

$$\begin{aligned} &\lambda_1 + \lambda_2 + \cdots + \lambda_n = -a_1 = a_{11} + a_{22} + \cdots + a_{nn} = \mathrm{tr}(A), \\ &\lambda_1\lambda_2\cdots\lambda_n = (-1)^n a_n = (-1)^n \det(-A) = \det(A). \end{aligned} \tag{4.3.3}$$

如果 A 是 n 阶实方阵, 则 A 的特征多项式 $\det(\lambda E_n - A)$ 是 n 次实多项式. 因为实系数多项式的根或为实数, 或为复数, 且复根必和其共轭复根成对出现, 所以这时方阵 A 的特征根为

$$\lambda_j = \sigma_j + i\rho_j, \quad \overline{\lambda_j} = \sigma_j - i\rho_j, \quad j = 1, 2, \cdots, s, \quad \lambda_{2s+1}, \cdots, \lambda_n, \tag{4.3.4}$$

其中 $\sigma_j, 1 \leqslant j \leqslant s$, $\lambda_k, 2s+1 \leqslant k \leqslant n$ 都是实数, $\rho_j, 1 \leqslant j \leqslant s$ 都是非零实数.

由定义可知, 对方阵 A 的任一特征根 $\lambda_j, 1 \leqslant j \leqslant n$, 有

$$\det(\lambda_j E_n - A) = 0. \tag{4.3.5}$$

所以 n 阶方阵 $\lambda_j E_n - A$ 的秩小于 n. 设它的秩为 r_j, 于是 n 个未知数 $x_1, x_2, \cdots, x_n$ 的齐次线性方程组

$$\begin{cases} (\lambda_j - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n = 0, \\ -a_{21}x_1 + (\lambda_j - a_{22})x_2 - \cdots - a_{2n}x_n = 0, \\ \qquad \vdots \\ -a_{1n}x_1 - a_{2n}x_2 - \cdots + (\lambda_j - a_{nn})x_n = 0 \end{cases} \tag{4.3.6}$$

的通解依赖于 $n-r_j$ 个独立参数, 所在的域和特征根所在的域相同. 特别它有非零解.

设 $\alpha^{(j)}_{r_j+1}, \alpha^{(j)}_{r_j+2}, \cdots, \alpha^{(j)}_{n}$ 为它的一组基础解系, 则通解可表为

$$x=t_{r_j+1}\alpha^{(j)}_{r_j+1}+t_{r_j+2}\alpha^{(j)}_{r_j+2}+\cdots+t_n\alpha^{(j)}_{n}, \tag{4.3.7}$$

其中 $t_{r_j+1}, t_{r_j+2}, \cdots, t_n$ 为独立参数.

引进 $n\times 1$ 矩阵 $x=(x_1, x_2, \cdots, x_n)'$, 则齐次线性方程组 (4.3.6) 可以写成矩阵乘积的形式:

$$(\lambda_j E_n - A)x = 0. \tag{4.3.8}$$

利用矩阵乘法的分配律, 可改写为

$$Ax=\lambda_j x. \tag{4.3.9}$$

所以, 将齐次线性方程组 (4.3.6) 的非零解排成 $n\times 1$ 矩阵 $\alpha=(a_1, a_2, \cdots, a_n)'$, 则 α 适合关系:

$$A\alpha=\lambda_j\alpha. \tag{4.3.10}$$

由此可知, 如果 α 和 β 都适合线性方程组 (4.3.6), 则 $\alpha+\beta$, $b\alpha$ 都适合齐次线性方程组 (4.3.6), 其中 b 所在的域和特征根所在的域相同.

定义 4.3.2 给定域 $\mathbb{F}$ 上 n 阶方阵 A 的特征根 (可能是复数根) λ_0, 线性方程组 (4.3.6) 的非零解 (所在的域和特征根所在的域相同) 构成的 $n\times 1$ 非零矩阵

$$\alpha=(a_1,a_2,\cdots,a_n)'\neq 0 \tag{4.3.11}$$

称为方阵 A 的属于特征根 λ_0 的特征向量, 或者简称为方阵 A 的特征向量. 当 λ_0 是复 (实) 数时, α 是复 (实) 向量. 换句话说, 如果 $n\times 1$ 非零矩阵 α 适合关系

$$A\alpha=\lambda_0\alpha, \tag{4.3.12}$$

则 α 称为 A 的**特征向量**.

由此可知, 属于同一特征根 λ_j 的两特征向量的非零 $n\times 1$ 矩阵和, 以及非零纯量积都是属于特征根 λ_j 的特征向量. 另外, 也可知特征向量必然存在, 而且在实方阵的情形实特征根的特征向量一定是实向量. 又方阵 A 的特征向量必为非零向量.

例 4.3.1 下面举一例来计算方阵的特征多项式、特征根和特征向量. 今方阵

$$A=\begin{pmatrix} 5 & 6 & -3 \\ -1 & 0 & 1 \\ 1 & 2 & -1 \end{pmatrix}$$

的特征多项式为

$$\det(\lambda E_3 - A) = \det\begin{pmatrix} \lambda-5 & -6 & 3 \\ 1 & \lambda & -1 \\ -1 & -2 & \lambda+1 \end{pmatrix} = (\lambda-2)(\lambda-1-\sqrt{3})(\lambda-1+\sqrt{3}).$$

所以方阵 A 有三个不同特征根 $2, 1+\sqrt{3}, 1-\sqrt{3}$, 它们都是实数. 为方便起见, 分别记作 $\lambda_1, \lambda_2, \lambda_3$.

(1) 令特征根 $\lambda_1 = 2$. 作齐次线性方程组

$$\begin{cases} -3x_1 + x_2 - x_3 = 0, \\ -6x_1 + 2x_2 - 2x_3 = 0, \\ 3x_1 - x_2 + 3x_3 = 0. \end{cases}$$

由消去法求得其通解为 $x_1 = t_1$, $x_2 = 3t_1$, $x_3 = 0$, 所以属于特征根 $\lambda_1 = 2$ 的所有特征向量可表为 $t_1(1,3,0)'$, 其中 t_1 为任一非零实数.

(2) 令特征根 $\lambda_2 = 1+\sqrt{3}$. 作齐次线性方程组

$$\begin{cases} (-4+\sqrt{3})x_1 + x_2 - x_3 = 0, \\ -6x_1 + (1+\sqrt{3})x_2 - 2x_3 = 0, \\ 3x_1 - x_2 + (2+\sqrt{3})x_3 = 0. \end{cases}$$

由消去法可知其通解为 $x_1 = t_2$, $x_2 = 2t_2$, $x_3 = (-2+\sqrt{3})t_2$, 所以属于特征根 $\lambda_2 = 1+\sqrt{3}$ 的所有特征向量可表为 $t_2(1,2,-2+\sqrt{3})'$, 其中 t_2 为任一非零实数.

(3) 令特征根 $\lambda_3 = 1-\sqrt{3}$. 作齐次线性方程组

$$\begin{cases} (-4-\sqrt{3})x_1 + x_2 - x_3 = 0, \\ -6x_1 + (1-\sqrt{3})x_2 - 2x_3 = 0, \\ 3x_1 - x_2 + (2+\sqrt{3})x_3 = 0. \end{cases}$$

由消去法可知其通解为 $x_1 = t_3$, $x_2 = 2t_3$, $x_3 = -(2+\sqrt{3})t_3$, 所以属于特征根 $\lambda_3 = 1-\sqrt{3}$ 的所有特征向量可表为 $t_3(1,2,-2-\sqrt{3})'$, 其中 t_3 为任一非零实数.

最后, 给出特征多项式的下面重要定理.

定理 4.3.3 (Hamilton-Cayley 定理) 将域 $\mathbb{F}$ 上 n 阶方阵 A 的特征多项式

$$\det(\lambda E_n - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \tag{4.3.13}$$

的自变量 λ 写为方阵 A, 则有

$$A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nE_n = 0, \tag{4.3.14}$$

其中 0 为 n 阶零方阵.

证 记 $(\lambda E_n - A)^{(*)}$ 为方阵 $\lambda E_n - A$ 的伴随方阵 (由 (3.3.2) 定义), 它有性质

$$(\lambda E_n - A)(\lambda E_n - A)^{(*)} = (\det(\lambda E_n - A))E_n.$$

由于 $(\lambda E_n - A)^{(*)}$ 的每个元素都是 $\lambda E_n - A$ 的 $n-1$ 阶代数子式, 所以它是 λ 的多项式, 且次数最多是 $n-1$. 因此方阵 $(\lambda E_n - A)^{(*)}$ 的每个元素都是 λ 的多项式, 且次数最多是 $n-1$. 利用矩阵加法的定义, 有

$$(\lambda E_n - A)^{(*)} = \lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + \lambda B_{n-2} + B_{n-1},$$

其中 $B_0, B_1, \cdots, B_{n-1}$ 是和 λ 无关的 n 阶常数方阵. 所以推出方阵等式

$$\begin{aligned}&(\lambda E_n - A)(\lambda^{n-1}B_0 + \lambda^{n-2}B_1 + \cdots + \lambda B_{n-2} + B_{n-1})\\ =\ &\lambda^n E_n + a_1\lambda^{n-1}E_n + \cdots + a_{n-1}\lambda E_n + a_n E_n.\end{aligned}$$

即有

$$\begin{aligned}&\lambda^n B_0 + \lambda^{n-1}(B_1 - AB_0) + \cdots + \lambda(B_{n-1} - AB_{n-2}) - AB_{n-1}\\ =\ &\lambda^n E_n + a_1\lambda^{n-1}E_n + \cdots + a_{n-1}\lambda E_n + a_n E_n.\end{aligned}$$

由于 λ 为未知数, 所以相同项的系数相等, 这证明了

$$\begin{aligned}a_n E_n &= -AB_{n-1},\\ a_{n-1}E_n &= -AB_{n-2} + B_{n-1},\\ &\ \ \vdots\\ a_2 E_n &= -AB_1 + B_2,\\ a_1 E_n &= -AB_0 + B_1,\\ E_n &= B_0.\end{aligned}$$

于是有

$$\begin{aligned}B_0 &= E_n,\\ B_1 &= a_1E_n + AB_0 = a_1E_n + A,\\ B_2 &= a_2E_n + AB_1 = a_2E_n + A(a_1E_n + A) = A^2 + a_1A + a_2E_n,\\ &\ \ \vdots\\ B_{n-1} &= A^{n-1} + a_1A^{n-2} + a_2A^{n-3} + \cdots + a_{n-2}A + a_{n-1}E_n\end{aligned}$$

以及 $a_nE_n = -AB_{n-1}$. 因此

$$A(A^{n-1} + a_1A^{n-2} + \cdots + a_{n-2}A + a_{n-1}E_n) + a_nE_n = 0.$$

这就证明了定理. □

定义 4.3.4 给定域 $\mathbb{F}$ 上 n 阶方阵 A, 首项系数为 1 的多项式

$$m(\lambda)=m(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m \tag{4.3.15}$$

称为方阵 A 的**极小多项式**, 如果

$$m(A)=A^m+b_1A^{m-1}+\cdots+b_{m-1}A+b_mE_n=0, \tag{4.3.16}$$

且 $m(\lambda)$ 是适合这种条件的非零多项式中次数为最小, 首项系数为 1 的多项式.

定理 4.3.5 给定域 $\mathbb{F}$ 上 n 阶方阵 A, 则有

(1) 唯一存在方阵 A 的极小多项式 $m(\lambda)$;

(2) 设多项式

$$f(\lambda)=b_0\lambda^p+b_1\lambda^{p-1}+\cdots+b_{p-1}\lambda+b_p$$

有

$$f(A)=b_0A^p+b_1A^{p-1}+\cdots+b_{p-1}A+b_pE_n=0,$$

则 $m(\lambda)\mid f(\lambda)$;

(3) $m(\lambda)\mid\det(\lambda E_n-A)$. 所以方阵 A 的极小多项式的根必为方阵 A 的特征根;

(4) 反之, 方阵 A 的特征根, 即特征多项式 $\det(\lambda E_n-A)$ 的根必为极小多项式的根.

证 由 Hamilton-Cayley 定理可知存在多项式 $g(\lambda)=\det(\lambda E_n-A)$, 使得 $g(A)=0$. 于是多项式集合

$$\{\,h(\lambda)\in\mathbb{F}[\lambda]\ \mid\ h(A)=0\,\}$$

不是空集, 也不等于零. 因此在其中存在一个次数最小的, 首项系数为 1 的多项式. 这证明了极小多项式 $m(\lambda)$ 的存在. 下面先证**(2)** 成立, 再证极小多项式的唯一性.

今任取多项式 $f(\lambda)$, 使得 $f(A)=0$. 由 $m(\lambda)\neq 0$, 有

$$f(\lambda)=q(\lambda)m(\lambda)+r(\lambda),$$

其中 $r(\lambda)=0$ 或者 $r(\lambda)\neq 0$, $0\leqslant\deg(r(\lambda))<\deg(m(\lambda))$. 但是 $f(A)=0$, 此即

$$f(A)=q(A)m(A)+r(A)=0.$$

已知 $m(A)=0$, 所以 $r(A)=0$. 若 $r(\lambda)\neq 0$, 由极小多项式定义可知 $\deg(r(\lambda))\geqslant\deg(m(\lambda))$. 这导出矛盾. 所以证明了 $r(\lambda)=0$, 此即 $m(\lambda)\mid f(\lambda)$. 这证明了**(2)** 成立.

下面证(1) 中的极小多项式的唯一性. 今若非零多项式 $m_1(\lambda)$ 的首项系数为 1, 且 $m_1(A)=0$, 又 $m_1(\lambda)$ 是适合条件 $h(A)=0$ 的非零多项式 $h(\lambda)$ 中, 次数为最小的多项式. 由(2) 成立, 有 $m_1(\lambda)\mid m(\lambda)$. 另一方面, 由(2), 因此, $m(\lambda)\mid m_1(\lambda)$. 所以 $m_1(\lambda)=m(\lambda)$. 这证明了(1) 成立.

由 Hamilton-Cayley 定理和(2) 可知(3) 成立.

余下证(4) 成立. 今任取 A 的特征根 λ_0, 于是

$$\det(\lambda_0 E_n - A)=0.$$

所以记 $x=(x_1, x_2, \cdots, x_n)'$, 则齐次线性方程组 $(\lambda_0 E_n - A)x=0$ 有非零解 x_0, 即 $Ax_0=\lambda_0 x_0$. 记

$$m(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m,$$

则

$$\begin{aligned}&A^m x_0+b_1A^{m-1}x_0+\cdots+b_{m-1}Ax_0+b_mE_nx_0\\=&\lambda_0^m x_0+b_1\lambda_0^{m-1}x_0+\cdots+b_{m-1}\lambda_0x_0+b_mx_0\\=&m(\lambda_0)x_0=0.\end{aligned}$$

由 $x_0\neq 0$, 便证明了 $m(\lambda_0)=0$, 即 λ_0 为极小多项式 $m(\lambda)$ 的根. 这证明了(4) 成立. □

最后讨论特征根的粗略的界.

定理 4.3.6 给定 n 阶复方阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}, \tag{4.3.17}$$

记

$$\rho=\max_{1\leqslant i,j\leqslant n}|a_{ij}|,\ \sigma=\max_{1\leqslant i,j\leqslant n}\left|\frac{a_{ij}+\overline{a_{ji}}}{2}\right|,\ \tau=\max_{1\leqslant i,j\leqslant n}\left|\frac{a_{ij}-\overline{a_{ji}}}{2}\right|. \tag{4.3.18}$$

则复方阵 A 的特征根 λ_0 必有

$$|\lambda_0|\leqslant n\rho,\quad |\mathrm{Re}(\lambda_0)|\leqslant n\sigma,\quad |\mathrm{Im}(\lambda_0)|\leqslant n\tau. \tag{4.3.19}$$

证 记 α_0 为 n 阶方阵 A 的属于特征根 λ_0 的特征向量, 即有 $\alpha_0\neq 0$, 又 $A\alpha_0=\lambda_0\alpha_0$. 于是记 $\alpha_0=(a_1, \cdots, a_n)'$, 则有

$$\alpha_0'\alpha_0=\sum_{j=1}^n|a_j|^2>0,\quad \overline{\alpha}_0'A\alpha_0=\lambda_0(\overline{\alpha}_0'\alpha_0)=\lambda_0\sum_{j=1}^n|a_j|^2,\quad \overline{\alpha}_0'\overline{A'}\alpha_0=\overline{\lambda}_0\sum_{j=1}^n|a_j|^2, \tag{4.3.20}$$

所以,

$$\overline{\alpha}_0'(A+\overline{A'})\alpha_0=2\mathrm{Re}\,(\lambda_0)\sum_{j=1}^{n}|a_j|^2,\quad \overline{\alpha}_0'(A-\overline{A'})\alpha_0=2\sqrt{-1}\mathrm{Im}\,(\lambda_0)\sum_{j=1}^{n}|a_j|^2.\quad(4.3.21)$$

因此,

$$\begin{aligned}|\lambda_0|\sum_{j=1}^{n}|a_j|^2&=|\overline{\alpha}_0'A\alpha_0|=\Big|\sum_{i,j=1}^{n}a_{ij}\overline{a_i}a_j\Big|\\&\leqslant\sum_{i,j=1}^{n}|a_{ij}||a_i||a_j|\leqslant\rho(\sum_{i}|a_i|)^2\leqslant n\rho\sum_{i}|a_i|^2.\end{aligned}$$

所以 $|\lambda_0|\leqslant n\rho$. 同理, $|\mathrm{Re}\,(\lambda_0)|\leqslant n\sigma$, $|\mathrm{Im}\,(\lambda_0)|\leqslant n\tau$. □

引理 4.3.7 n 阶复方阵 H 为 Hermite 方阵当且仅当 $\sqrt{-1}H$ 为反 Hermite 方阵.

证 由定义可直接验证. □

定理 4.3.8 **(1)** Hermite 方阵的特征根都是实数, 所以实对称方阵的特征根也都是实数;

(2) 反 Hermite 方阵的非零特征根都是纯虚数, 所以实反对称方阵的非零特征根也都是纯虚数.

证 由引理 4.3.7 可知, 只要证 Hermite 方阵的特征根都是实数就够了.

证法一 任取方阵 A 的特征根 λ_0, 于是存在属于特征根 λ_0 的特征向量 α. 由定义 $A^*=\overline{A'}=A$ 及 $A\alpha=a\alpha$. 所以 $\overline{\alpha}'A=\overline{a}\,\overline{\alpha}'$. 令 $0\neq\alpha=(a_1,a_2,\cdots,a_n)'$, 因此 $b=\overline{\alpha}'\alpha=\sum\limits_{j=1}^{n}|a_j|^2>0$. 又 $\overline{\alpha}'A\alpha=\overline{a}\,\overline{\alpha}'\alpha=b\overline{a}$, 所以 $ba=\overline{\overline{\alpha}'A\alpha}'=\overline{\alpha}'A^*\alpha=\overline{\alpha}'A\alpha=b\overline{a}$. 因此证明了 a 为实数. □

证法二 令 A 是 Hermite 方阵, 即 $A^*=\overline{A'}=A$, 所以 $a_{ij}=\overline{a_{ji}}$, $1\leqslant i,j\leqslant n$. 由定理 4.3.6, 任取 Hermite 方阵 A 的特征根 λ_0, 则 $|\mathrm{Im}\,(\lambda_0)|\leqslant n\tau$, 其中 $2\tau=\max_{1\leqslant i,j\leqslant n}|a_{ij}-\overline{a_{ji}}|=0$. 这证明了 $|\mathrm{Im}\,(\lambda_0)|=0$, 即 $\lambda_0\in\mathbb{R}$. □

定理 4.3.9 (圆盘定理) 给定 n 阶复方阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix},\qquad(4.3.22)$$

记

$$\rho_j=\sum_{k=1}^{n}|a_{jk}|-|a_{jj}|,\quad \rho_j'=\sum_{k=1}^{n}|a_{kj}|-|a_{jj}|,\quad 1\leqslant j\leqslant n.\qquad(4.3.23)$$

在平面上作闭圆盘

$$\mathrm{C}_j = \{\, z \in \mathbb{C} \mid |z-a_{jj}| \leqslant \rho_j \,\}, \quad \mathrm{C}'_j = \{\, z \in \mathbb{C} \mid |z-a_{jj}| \leqslant \rho'_j \,\}, \quad 1 \leqslant j \leqslant n, \tag{4.3.24}$$

则方阵 A 的特征根必落在并集 $\cup_{1\leqslant j\leqslant n} \mathrm{C}_j$ 以及并集 $\cup_{1\leqslant j\leqslant n} \mathrm{C}'_j$ 的交

$$\Big(\bigcup_{1\leqslant j\leqslant n} \mathrm{C}_j\Big) \cap \Big(\bigcup_{1\leqslant j\leqslant n} \mathrm{C}'_k\Big) \tag{4.3.25}$$

中.

证 任取复方阵 A 的特征根 λ_0, 由 $\det(\lambda_0 E_n - A) = \det(\lambda_0 E_n - A')$ 可知 λ_0 也必为复方阵 A' 的特征根. 今若能证 $\lambda_0 \in \cup_{1\leqslant j\leqslant n} \mathrm{C}_j$. 考虑复方阵 A', 由 C'_j 的定义可知 $\lambda_0 \in \cup_{1\leqslant k\leqslant n} \mathrm{C}'_k$. 这证明了

$$\lambda_0 \in \left(\bigcup_{1\leqslant j\leqslant n} \mathrm{C}_j\right) \cap \left(\bigcup_{1\leqslant k\leqslant n} \mathrm{C}'_k\right).$$

下面证 $\lambda_0 \in \cup_{1\leqslant j\leqslant n} \mathrm{C}_j$. 由并集的定义可知, 要证存在指标 j, 使得 $\lambda_0 \in \mathrm{C}_j$.

记 α_0 为复方阵 A 的属于特征根 λ_0 的特征向量, 记 $\alpha_0 = (a_1, \cdots, a_n)'$. 由定义, $A\alpha_0 = \lambda_0\alpha_0$, 所以

$$\sum_{j=1}^{n} a_{ij}a_j = \lambda_0 a_i, \quad 1 \leqslant i \leqslant n.$$

此即

$$\sum_{j\neq i} a_{ij}a_j = (\lambda_0 - a_{ii})a_i, \quad 1 \leqslant i \leqslant n.$$

但是 $\alpha_0 \neq 0$, 所以在 n 个非负实数 $|a_1|, |a_2|, \cdots, |a_n|$ 中存在一个最大数, 记作 $|a_{i_0}|$. 于是

$$\max_{1\leqslant i\leqslant n} |a_i| = |a_{i_0}| > 0.$$

因此有

$$|\lambda_0 - a_{i_0i_0}||a_{i_0}| = |\sum_{j\neq i_0} a_{i_0j}a_j| \leqslant |a_{i_0}| \sum_{j\neq i_0} |a_{i_0j}| = \rho_{i_0}|a_{i_0}|.$$

由 $|a_{i_0}| > 0$, 便证明了 $|\lambda_0 - a_{i_0i_0}| \leqslant \rho_{i_0}$, 即 $\lambda_0 \in \mathrm{C}_{i_0} \subset \cup_{1\leqslant j\leqslant n} \mathrm{C}_j$. □

习 题 4.3

4.3.1 试求下列方阵的特征多项式, 特征根和每个特征根所对应的特征向量:

(1) $\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$; **(2)** $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$; **(3)** $\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$.

4.3.2　设方阵 A 的所有特征根 (共有 n 个) 为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 且设 $\det(A) \neq 0$, 试求其逆方阵 A^{-1} 的所有特征根.

4.3.3　设方阵 A 有特征根 λ_0. 试证: 方阵 $B = f(A) = b_0A^m + b_1A^{m-1} + \cdots + b_{m-1}A + b_mE_n$ 有特征根

$$\mu_0 = b_0\lambda_0^m + b_1\lambda_0^{m-1} + \cdots + b_{m-1}\lambda_0 + b_m.$$

特别, A^2 有特征根 λ_0^2, 则 A^k 有特征根 λ_0^k.

4.3.4　试求 n 阶方阵 $\beta\beta'$ 及 $\overline{\beta}\beta'$ 的特征根, 这里 $\beta = (b_1, b_2, \cdots, b_n)'$ 是 $n \times 1$ 矩阵.

4.3.5　试指出断言: “在特征多项式 $\det(\lambda E_n - A) = 0$ 中取 λ 为方阵 A, 由于 $A \cdot E_n - A = 0$ 及 $\det(A \cdot E_n - A) = \det(0) = 0$, 所以定理 4.3.3 成立” 的错误原因.

4.3.6　试证: 方阵 A 的任一特征根 λ_0 适合不等式

$$|\lambda_0| \leqslant \min\Big(\max_{1\leqslant i\leqslant n} \big(\sum_{k=1}^{n} |a_{ki}|\big),\quad \max_{1\leqslant i\leqslant n} \big(\sum_{k=1}^{n} |a_{ik}|\big)\Big).$$

4.3.7　试证: 幂等方阵 $A^2 = A$ 的特征根只有 0 和 1; 幂零方阵 A, $(A^n = 0)$ 的特征根都是零, 且 n 阶方阵 A 幂零当且仅当 $\mathrm{tr}(A^k) = 0$, $k = 1, 2, \cdots$.

4.3.8　试证: n 阶可逆方阵 A 的逆方阵 A^{-1} 为 A 的多项式.

4.3.9　设 K 为实反对称方阵, 则对任一正数 $a > 0$, 有 $\det(aE_n + K) > 0$.

4.3.10　设 n 阶方阵 $A = (a_{ij})$ 有 $\sum\limits_{i=1}^{n} a_{ij} = 1$, $j = 1, 2, \cdots, n$, 则 A 必有特征根 1. 设 A 有特征根 $\lambda_0 \neq 1$, 记 α_0 为 A 的属于特征根 λ_0 的特征向量, 则 $n \times 1$ 矩阵 α_0 的元素的和为零.

4.3.11　设 λ_0 为域 $\mathbb{F}$ 上 n 阶方阵

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

的特征根. 在和号 $s = \sum\limits_{j=1}^{n} \dfrac{\sum\limits_{k\neq j} |a_{jk}|^2}{|\lambda_0 - a_{jj}|^2}$ 中只要指标 j 使得 $a_{jj} = \lambda_0$, 则除去这项. 试证 $s \geqslant 0$.

4.3.12　记 $n_1, n_2, \cdots, n_s$ 阶域 $\mathbb{F}$ 上方阵 $A_1, A_2, \cdots, A_s$ 的极小多项式分别为 $m_1(\lambda)$, $m_2(\lambda), \cdots, m_s(\lambda)$. 试证: 准对角方阵 $\mathrm{diag}(A_1, A_2, \cdots, A_s)$ 的极小多项式为 $m_1(\lambda)$, $m_2(\lambda), \cdots, m_s(\lambda)$ 的最小公倍式 $[m_1(\lambda), m_2(\lambda), \cdots, m_s(\lambda)]$.

§4.4　结式和判别式 *

在这一节, 我们回到一元多项式的情形, 讨论两个一元多项式的公根.

给定域 $\mathbb{F}$ 上两个一元多项式

$$\begin{aligned} f(x) &= a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n, \\ g(x) &= b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m. \end{aligned} \tag{4.4.1}$$

假设它们的次数都大于零, 且 $a_0b_0 \neq 0$. 由定理 1.5.2 可知, 一元多项式 $f(x)$ 和 $g(x)$ 有**公根**当且仅当它们不互素. 由定理 1.3.4 可知一元多项式 $f(x)$ 和 $g(x)$ 不互素当且仅当存在两非零多项式 $u(x)$ 和 $v(x)$, 使得 $u(x)f(x) = v(x)g(x)$, 其中 $0 \leqslant \deg(u(x)) < m$, $0 \leqslant \deg(v(x)) < n$. 将 $u(x)$ 和 $v(x)$ 写成

$$\begin{aligned} u(x) &= x_{n+1}x^{m-1} + x_{n+2}x^{m-2} + \cdots + x_{n+m-1}x + x_{n+m}, \\ v(x) &= x_1x^{n-1} + x_2x^{n-2} + \cdots + x_{n-1}x + x_n, \end{aligned} \tag{4.4.2}$$

其中 $x_1, x_2, \cdots, x_n$ 不全为零, $x_{n+1}, x_{n+2}, \cdots, x_{n+m}$ 也不全为零.

引理 4.4.1 符号同上. 多项式 $f(x)$ 和 $g(x)$ 有公根当且仅当 $n+m$ 个未知数 $x_1, x_2, \cdots, x_{n+m}$ 的线性方程组

$$b_jx_1 + b_{j-1}x_2 + \cdots + b_{j-n+1}x_n = a_jx_{n+1} + a_{j-1}x_{n+2} + \cdots + a_{j-m+1}x_{n+m}, \tag{4.4.3}$$

$j = 0, 1, \cdots, n+m-1$ 有非零解, 其中: 当 $j < 0$ 或 $j > n$ 时, $a_j = 0$; 当 $j < 0$ 或 $j > m$ 时, $b_j = 0$.

证 由于多项式 $f(x)$ 和 $g(x)$ 有公根当且仅当存在 $x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)} \in \mathbb{F}$ 不全为零, 又存在 $x_{n+1}^{(0)}, x_{n+2}^{(0)}, \cdots, x_{n+m}^{(0)} \in \mathbb{F}$ 不全为零, 使得由 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_{n+m}^{(0)}) \in \mathbb{F}$ 按式 (4.4.2) 构造的多项式 $u(x)$ 和 $v(x)$ 有

$$\begin{aligned} 0 &= u(x)f(x) - v(x)g(x) \\ &= \Big(\sum_{i=0}^{n} a_ix^{n-i}\Big)\Big(\sum_{k=0}^{m-1} x_{n+k+1}^{(0)}x^{m-k-1}\Big) - \Big(\sum_{p=0}^{m} b_px^{m-p}\Big)\Big(\sum_{q=0}^{n-1} x_{q+1}^{(0)}x^{n-q-1}\Big) \\ &= \sum_{j=0}^{n+m-1}\Big(\sum_{i+k=j} a_ix_{n+k+1}^{(0)} - \sum_{p+q=j} b_px_{q+1}^{(0)}\Big)x^{n+m-j-1}, \end{aligned}$$

即齐次线性方程组

$$\sum_{i+k=j} a_ix_{n+k+1} - \sum_{i+k=j} b_ix_{k+1} = 0, \quad 0 \leqslant j \leqslant n+m-1 \tag{4.4.4}$$

的非零解为 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_{n+m}^{(0)}) \in \mathbb{F}$. 所以为了证明引理, 只要证明线性方程组 (4.4.4) 若有非零解 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_{n+m}^{(0)})$, 则 $x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}$ 不全为零, 且 $x_{n+1}^{(0)}, x_{n+2}^{(0)}, \cdots, x_{n+m}^{(0)}$ 也不全为零.

为此用反证法. 如果 $x_1^{(0)} = x_2^{(0)} = \cdots = x_n^{(0)} = 0$, 但是 $x_{n+1}^{(0)}, x_{n+2}^{(0)}, \cdots, x_{n+m}^{(0)}$ 不全为零. 则

$$a_jx_{n+1}^{(0)} + a_{j-1}x_{n+2}^{(0)} + \cdots + a_{j-m+1}x_{n+m}^{(0)} = 0, \quad j = 0, 1, \cdots, n+m-1.$$

取 $j = 0$, 有 $a_0x_{n+1}^{(0)} = 0$. 但是 $a_0 \neq 0$, 所以 $x_{n+1}^{(0)} = 0$. 再取 $j = 1$, 有 $x_{n+2}^{(0)} = 0$. 这样依次作下去, 最后证明了 $x_{n+1}^{(0)} = x_{n+2}^{(0)} = \cdots = x_{n+m}^{(0)} = 0$. 这导出了矛盾.

如果 $x_{n+1}^{(0)}=x_{n+2}^{(0)}=\cdots=x_{n+m}^{(0)}=0$, 但是 $x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}$ 不全为零, 则

$$b_jx_1^{(0)}+b_{j-1}x_2^{(0)}+\cdots+b_{j-n+1}x_n^{(0)}=0,\quad j=0,1,\cdots,n+m-1.$$

取 $j=0$, 有 $b_0x_1^{(0)}=0$. 但是 $b_0\neq 0$, 所以 $x_1^{(0)}=0$. 再取 $j=1$, 有 $x_2^{(0)}=0$. 这样依次作下去, 最后证明了 $x_1^{(0)}=x_2^{(0)}=\cdots=x_n^{(0)}=0$. 这又导出了矛盾. □

定义 4.4.2 两个非零多项式

$$\begin{aligned} f(x)&=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n,\\ g(x)&=b_0x^m+b_1x^{m-1}+\cdots+b_{m-1}x+b_m \end{aligned}\tag{4.4.5}$$

决定的 $n+m$ 阶行列式

$$R(f,g)=\det\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_n & & & \\ & a_0 & a_1 & a_2 & \cdots & a_n & & \\ & & \ddots & \ddots & \ddots & & \ddots & \\ & & & a_0 & a_1 & a_2 & \cdots & a_n \\ b_0 & b_1 & b_2 & \cdots & b_m & & & \\ & b_0 & b_1 & b_2 & \cdots & b_m & & \\ & & \ddots & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & b_2 & \cdots & b_m \end{pmatrix}\begin{matrix} \\ \\ \\ \text{第 } m \text{ 行} \\ \\ \\ \\ \\ \end{matrix}\tag{4.4.6}$$

称为多项式 $f(x)$ 和 $g(x)$ 的结式. 它是多项式 $f(x)$ 及 $g(x)$ 的系数的函数, 其中行列式 $R(f,g)$ 的前面 m 行由多项式 $f(x)$ 的系数给出, 后面 n 行由多项式 $g(x)$ 的系数给出.

引理 4.4.3 次数大于零的多项式 $f(x)$ 和 $g(x)$ 有公根当且仅当它们的结式 $R(f,g)=0$.

证 由引理 4.4.1 给出的线性方程组 (4.4.3) 的系数行列式等于 $(-1)^nR(f,g)$. 引理 4.4.1 证明了 $f(x)$ 和 $g(x)$ 有公根当且仅当线性方程组 (4.4.3) 有非零解. 由 §4.2 可知, 当且仅当线性方程组 (4.4.3) 的系数行列式等于零, 即 $R(f,g)=0$. □

定理 4.4.4 设次数大于零的多项式 $f(x)$ 和 $g(x)$ 分别有根 $x_1, x_2, \cdots, x_n$ 和 $y_1, y_2, \cdots, y_m$. 则多项式 $f(x)$ 和 $g(x)$ 的结式

$$R(f,g)=a_0^mb_0^n\prod_{j=1}^n\prod_{k=1}^m(x_j-y_k)=a_0^m\prod_{j=1}^n g(x_j)=(-1)^{nm}b_0^n\prod_{k=1}^m f(y_k),\tag{4.4.7}$$

其中 a_0 和 b_0 分别为多项式 $f(x)$ 和 $g(x)$ 的首项系数.

证 由结式的定义可知

$$R(\lambda f,\mu g)=\lambda^m\mu^nR(f,g),\quad \forall\,\lambda,\mu\in\mathbb{F}.$$

因此记

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$
$$g(x) = b_0x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m. \tag{4.4.8}$$

取 $\lambda = a_0^{-1}$, $\mu = b_0^{-1}$, 从而无妨设 $a_0 = b_0 = 1$. 即有

$$f(x) = (x - x_1)(x - x_2)\cdots(x - x_n) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n,$$
$$g(x) = (x - y_1)(x - y_2)\cdots(x - y_m) = x^m + b_1x^{m-1} + \cdots + b_{m-1}x + b_m,$$

记 x_1, x_2, $\cdots$, x_n 的初等对称多项式为 σ_1, σ_2, $\cdots$, σ_n. 由 Vieta 定理, 于是

$$a_j = (-1)^j\sigma_j, \quad j = 1, 2, \cdots, n;$$

记 y_1, y_2, $\cdots$, y_m 的初等对称多项式为 τ_1, τ_2, $\cdots$, τ_m. 由 Vieta 定理, 于是

$$b_k = (-1)^k\tau_k, \quad k = 1, 2, \cdots, m.$$

由行列式的定义可知结式 $R(f,g)$ 为初等对称多项式 σ_1, σ_2, $\cdots$, σ_n 的多项式, 也是初等对称多项式 τ_1, τ_2, $\cdots$, τ_m 的多项式. 注意到当 x_1, x_2, $\cdots$, x_n 和 y_1, y_2, $\cdots$, y_m 取定时, $R(f,g)$ 为 x_1 的多项式. 令 $x_1 = y_k$, $1 \leqslant k \leqslant m$, 则由于这时 f 和 g 有公根, 由引理 4.4.3, x_1 的多项式 $R(f,g)$ 之值为零, 即以 y_1, y_2, $\cdots$, y_m 为根. 于是有 $(x_1 - y_k) \mid R(f,g)$. 这证明了 $R(f,g) = \prod\limits_{k=1}^{m}(x_1 - y_k)h(x_1, \cdots, x_n; y_1, \cdots, y_m)$. 比较双方 x_1^m 的系数, 便证明了 $h(x_1, \cdots, x_n; y_1, \cdots, y_m)$ 为 x_2, $\cdots$, x_n, y_1, $\cdots$, y_m 的多项式. 对 x_2, $\cdots$, x_n 同样讨论, 最后证明了

$$R(f,g) = h_0\prod_{j=1}^{n}\prod_{k=1}^{m}(x_j - y_k),$$

其中 h_0 为常数.

为了确定常数 h_0, 取 $g(x) = x^m$, 即取 $b_0 = 1$, $b_1 = \cdots = b_m = 0$, 于是 y_1, $\cdots$, y_m 都为零. 因此

$$R(f,g) = h_0\prod_{j=1}^{n}x_j^m = h_0(x_1x_2\cdots x_n)^m = h_0\sigma_n^m.$$

由直接计算行列式 $R(f,g)$, 便有

$$R(f,g) = (-1)^{nm}a_n^m = [(-1)^na_n]^m = \sigma_n^m.$$

这证明了 $h_0 = 1$, 即

$$R(f,g) = \prod_{j=1}^{n}\prod_{k=1}^{m}(x_j - y_k). \tag{4.4.9}$$

□

定义 4.4.5　给定域 $\mathbb{F}$ 上次数大于零的多项式 $f(x)$, 记作

$$f(x)=a_0(x-x_1)(x-x_2)\cdots(x-x_n).$$

则 $x_1, x_2, \cdots, x_n$ 的对称多项式

$$\Delta(f)=(-1)^{\frac{n(n-1)}{2}}a_0^{2n-1}\prod_{j\neq k}(x_j-x_k)=a_0^{2n-1}\prod_{1\leqslant j<k\leqslant n}(x_j-x_k)^2 \tag{4.4.10}$$

称为多项式 $f(x)$ 的**判别式**.

显然有

定理 4.4.6　域 $\mathbb{F}$ 上次数大于零的多项式 $f(x)$ 有重根当且仅当 $f(x)$ 的判别式 $\Delta(f)=0$.

关于结式和判别式, 有下面重要的关系.

定理 4.4.7　对次数大于零的多项式 $f(x)$, 有

$$\Delta(f)=(-1)^{\frac{n(n-1)}{2}}R(f,f'). \tag{4.4.11}$$

证　记

$$f(x)=a_0(x-x_1)(x-x_2)\cdots(x-x_n),$$

则

$$f'(x)=a_0\sum_{k=1}^{n}(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n).$$

由定理 4.4.4,

$$R(f,f')=a_0^{n-1}\prod_{j=1}^{n}f'(x_j)=a_0^{2n-1}\prod_{j=1}^{n}\prod_{k\neq j}(x_j-x_k)=(-1)^{\frac{n(n-1)}{2}}\Delta(f).\qquad\square$$

习　题　4.4

4.4.1　试计算下列域 $\mathbb{F}$ 上多项式的结式:

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

(2)　$f(x)=2x^4-x^3+3$, $g(x)=3x^3-x^2+4$.

4.4.2　试计算域 $\mathbb{F}$ 上多项式 $f(x)=\dfrac{x^n-1}{x-1}$, $g(x)=\dfrac{x^m-1}{x-1}$ 的结式.

4.4.3　试计算下列域 $\mathbb{F}$ 上多项式的判别式:

(1)　$f(x)=x^n+a$;　**(2)**　$f(x)=x^n+px+q$;

(3)　$f(x)=a_0x^{n+m}+a_1x^m+a_2$;　**(4)**　$f(x)=\dfrac{x^n-1}{x-1}$;

(5)　$f(x)=x^n+ax^{n-1}+ax^{n-2}+\cdots+a$.

4.4.4　试证: 域 $\mathbb{F}$ 上结式有性质 $R(f,g_1g_2)=R(f,g_1)R(f,g_2)$.

4.4.5 试证: 域 $\mathbb{F}$ 上判别式有性质 $\Delta((x-a)f(x))=(f(a))^2\Delta(f(x))$.

4.4.6 在域 $\mathbb{F}$ 上求 $x^3+y^3+z^3-3xyz$ 的因子分解.

4.4.7 在域 $\mathbb{F}$ 上已知 $\Delta(f(x))$, 试求 $\Delta(f(x^m))$, $m=2,3,\cdots$.

4.4.8 试证: 域 $\mathbb{F}$ 上多项式 $f(g(x))$ 的判别式为

$$\Delta(f(g(x)))=\Delta(f(x))^m\prod_{1\leqslant j\leqslant n}\Delta(g(x)-x_j),$$

其中 n 是 $f(x)$ 的次数, m 是 $g(x)$ 的次数, x_1, x_2, $\cdots$, x_n 是 $f(x)$ 的所有根, 又 $f(x)$ 和 $g(x)$ 的首项系数都是 1.

第五章　线性空间

§5.1　线性空间

从几何的角度来看, 点空间和向量空间有这样的关系: 点空间可以看作是在空间中取定一点 (一般取原点), 再把向量空间安装在这点上构成的. 即以这点为起点的所有自由向量的终点充满整个空间. 由此可见, 点空间的许多性质, 例如点的共线、共面、直线和平面的平行、相交等, 都可以用向量空间的语言来刻画.

在许多实际问题中提炼出来的数学问题, 往往要求我们推广普通空间 (点空间及向量空间) 的概念, 创造新的数学模型. 然而, 点和有向线段都具有实在的几何内容, 所以很难从几何角度进行推广. 下面利用普通空间中点和向量的坐标来作推广, 不过我们尽可能保持了原有的几何关系.

定义 5.1.1　域 $\mathbb{F}$ 上 $n\times 1$ 矩阵 $(a_1, a_2, \cdots, a_n)'$ 称为一个 $\boldsymbol{n}$ **维点**, 其中 $(*)'$ 是 $1\times n$ 矩阵 $(*)$ 的转置矩阵. 域 $\mathbb{F}$ 上的数 a_j 称为点 A 的第 j 个**笛卡尔坐标**. n 维点的全体构成的集合称为 n 维**点空间**. 在 n 维点空间中任取一点 $C=(c_1,c_2,\cdots,c_n)'$, n 维点空间可记作 $(\mathbb{F}^n, C)$.

在 $n=1$ 时, 它就是通常的直线; 在 $n=2$ 时, 它就是通常的平面; 在 $n=3$ 时, 它就是通常的空间. 当然, 这些都理解为在引进笛卡尔坐标系后, 把点看作是数组. 在 $n>3$ 时就失去了几何的直观. 这时, 虽然我们不能从几何直观来进行想象, 不过还是可以用类比的方法, 借助于通常的空间来进行想象. 例如, 对 n 维点空间 $\mathbb{F}^n$ 中任两点 $P=(a_1, a_2, \cdots, a_n)'$ 和点 $Q=(b_1, b_2, \cdots, b_n)'$, 可以称点偶 (P,Q) 为**有向线段**, 或者称为**固定向量**, 记作 $\overrightarrow{PQ}$. 点 P 称为起点, 点 Q 称为终点. 显然, 任一固定向量 $\overrightarrow{PQ}$ 对应了一个 n 数组 $(b_1-a_1, b_2-a_2, \cdots, b_n-a_n)$, 称为固定向量 $\overrightarrow{PQ}$

的**坐标**. 两个固定向量的坐标若成比例, 则称为**平行**; 两个固定向量的坐标如果相同, 则称为**相等**. 把所有相等的固定向量看作一个东西, 这就导出自由向量的概念. 确切地说, 显然, 相等是等价关系, 等价类就是**自由向量**. 而过一点的固定向量代表了一个自由向量. 因此将所有自由向量安在同一点上, 那末终点必然充满整个点空间. 利用这些类比, 可以引进:

定义 5.1.2 域 $\mathbb{F}$ 上 $n\times 1$ 矩阵 $(a_1, a_2, \cdots, a_n)'$ 称为一个 n 维**向量**. 数值 a_j 称为第 j 个**分量**或第 j 个 **坐标**, $1\leqslant j\leqslant n$. 所有 n 维向量构成的集合称为 n 维**向量空间**, 用符号 $\mathbb{F}^n$ 记之.

如果两个 n 维向量的所有相同位置的分量相等, 则称这两向量**相等**.

显然, n 维向量又可以看作是 $n\times 1$ 矩阵. 而 $\mathbb{F}^n$ 是由所有域 $\mathbb{F}$ 上 $n\times 1$ 矩阵构成的集合. 已知 $n\times 1$ 矩阵间有加法和纯量积, 所以把 $n\times 1$ 矩阵看作是 n 维向量, 那末在 n 维向量空间 $\mathbb{F}^n$ 中也可以用矩阵的加法和纯量积来定义向量的加法和纯量积, 它们满足矩阵的加法和纯量积所满足的全部性质.

将 n 维向量空间 $\mathbb{F}^n$ 抽象化, 便引出线性空间的概念.

定义 5.1.3 给定集合 $\mathfrak{L}$. 如果在集合 $\mathfrak{L}$ 中可以定义两种代数运算: 加法和纯量积. 即集合 $\mathfrak{L}$ 中任一元素称为**向量**. 对任两向量 α 和 β, 按照一个确定的规律对应了集合 $\mathfrak{L}$ 中一个向量 ξ, 称为向量 α 和 β 的**和**, 记作 $\xi=\alpha+\beta$. 我们也称 α 和 β 作了**加法**; 又对集合 $\mathfrak{L}$ 中任一向量 α 及域 $\mathbb{F}$ 中任一数 b, 按照一个确定的规律对应了集合 $\mathfrak{L}$ 中一个向量 η, 称为向量 α 和数 b 的 **纯量积**, 记作 $\eta=b\alpha$. 且加法和纯量积适合下面一系列条件:

(一) 加法运算具有性质: 对任三向量 $\alpha, \beta, \gamma\in\mathfrak{L}$, 有

(1) **加法结合律**: $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$;

(2) **加法交换律**: $\alpha+\beta=\beta+\alpha$;

(3) 存在**零向量** 0, 它具有性质: $0+\alpha=\alpha+0=\alpha$;

(4) 对任一向量 α, 存在向量 β 使得: $\alpha+\beta=\beta+\alpha=0$, 记 β 为 $-\alpha$, 称为向量 α 的**负向量**. 利用负向量, 还可以引进**减法**: $\alpha-\beta=\alpha+(-\beta)$.

(二) 纯量积运算具有性质: 对一切数 $b, c\in\mathbb{F}$ 及一切元素 $\alpha, \beta\in\mathfrak{L}$ 有

(5) $b(\alpha+\beta)=b\alpha+b\beta$;

(6) $(b+c)\alpha=b\alpha+c\alpha$;

(7) $(bc)\alpha=b(c\alpha)$;

(8) $1\cdot\alpha=\alpha$.

这时, 集合 $\mathfrak{L}$ 称为域 $\mathbb{F}$ 上线性空间, 或简称为 **线性空间**. 集合 $\mathfrak{L}$ 中元素称为**向量**, 域 $\mathbb{F}$ 中数称为**纯量**.

下面给出域 $\mathbb{F}$ 上线性空间的若干例子.

例 5.1.1 分别由通常的直线、通常的平面、通常的空间中的所有自由向量构成

的集合都是实线性空间. 由解析几何课程可知, 在通常的空间中取定**坐标系**, 熟知每个自由向量有坐标 (a_1, a_2, a_3). 于是可以作一个 3×1 实矩阵 $(a_1, a_2, a_3)'$. 因此所有自由向量构成的线性空间, 可以自然地诱导了它们的坐标表达式: 由所有 3×1 实矩阵构成线性空间 $\mathbb{R}^3$. 这时的加法和纯量积就是 3×1 实矩阵的加法和纯量积.

例 5.1.2　域 $\mathbb{F}$ 上 n 维向量空间 $\mathbb{F}^n$ 就是一个域 $\mathbb{F}$ 上的线性空间.

例 5.1.3　记 $\mathrm{M}_{n,m}(\mathbb{F})$ 为由域 $\mathbb{F}$ 上所有 $n\times m$ 矩阵构成的集合. 它们在矩阵的加法和纯量积下, 构成域 $\mathbb{F}$ 上线性空间 $\mathrm{M}_{n,m}$. 它有 $\dim(\mathrm{M}_{n,m}) = nm$. 当 $n = m$ 时, $\mathrm{M}_{n,m}(\mathbb{F})$ 改记为 $\mathrm{gl}\,(n, \mathbb{F})$.

例 5.1.4　记 $\mathbb{F}_n[x]$ 为所有次数不超过 n 的域 $\mathbb{F}$ 上的多项式构成的集合, 所以 $\mathbb{F}_n[x]$ 中每一元素都可以写成

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0,\quad a_0,\ a_1,\ \cdots,\ a_n \in \mathbb{F}.$$

显然, 在多项式的加法及纯量积下, $\mathbb{F}_n[x]$ 构成域 $\mathbb{F}$ 上 $n+1$ 维线性空间.

例 5.1.5　记 $\mathbb{F}[x]$ 为所有域 $\mathbb{F}$ 上的多项式构成的集合. 按照例 5.1.4 所定义的多项式的加法和纯量积, 可证 $\mathbb{F}[x]$ 构成域 $\mathbb{F}$ 上无限维线性空间.

例 5.1.6　设 $\mathrm{C}((0,1))$ 为由所有在开区间 $(0,1)$ 上的实变数 x 的连续函数构成的集合. 显然, 两个连续函数的和仍为连续函数. 又, 一个连续函数和一个实数的乘积仍为连续函数. 可以验证, 在这样的加法和纯量积下, $\mathrm{C}((0,1))$ 构成无限维实线性空间.

在进一步展开线性空间理论以前, 先从定义出发给出关于运算的一些最基本的性质.

引理 5.1.4　域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 的加法和纯量积有下面性质:

(1)　任意有限个向量作加法时, 可以不计先后及次序.

(2)　**加法消去律**成立, 即由 $\alpha+\beta = \alpha+\gamma$ 可以推出 $\beta = \gamma$, $\forall\,\alpha,\ \beta,\ \gamma\in\mathfrak{L}$.

(3)　线性空间 $\mathfrak{L}$ 中有且只有一个零向量; 对任一向量有且只有一个负向量.

(4)　$b\alpha = 0$ 当且仅当 $b = 0$ 或 $\alpha = 0$, 其中 $b\in\mathbb{F}$, $\alpha\in\mathfrak{L}$. 因此若 u, v 和 $w\in\mathbb{F}$, 又 $u\neq 0$, $\alpha,\ \beta\in\mathfrak{L}$, 则由 $uv\alpha = uw\beta$ 可以推出 $v\alpha = w\beta$.

(5)　$(-b)\alpha = b(-\alpha) = -(b\alpha)$,　$\forall\, b\in\mathbb{F}$, $\alpha\in\mathfrak{L}$.

证　(1)　由加法结合律和归纳法可证有限个向量相加时可以不计先后. 由加法交换律和归纳法可证有限个向量相加时可以不计次序. 这证明了**(1)** 成立.

(2)　下面证加法消去律成立. 今由 $\alpha+\beta = \alpha+\gamma$, 由于等量加等量仍为等量, 所以 $(-\alpha)+(\alpha+\beta) = (-\alpha)+(\alpha+\gamma)$. 由加法结合律, 有 $((-\alpha)+\alpha)+\beta = ((-\alpha)+\alpha)+\gamma$. 由负向量的定义可知 $(-\alpha)+\alpha = 0$, 故有 $0+\beta = 0+\gamma$, 由零向量的定义可知 $\beta = \gamma$, 这证明了**(2)** 成立.

(3) 由线性空间的定义可知零向量 0 存在. 若另有一个零向量 0_1. 由定义, 则有 $0_1+\alpha=\alpha+0_1=\alpha$, $\forall\alpha\in\mathfrak{L}$. 特别, 由 $0\in\mathfrak{L}$, 所以有 $0_1+0=0+0_1=0$. 但是 0 为零向量, 故 $0_1+0=0+0_1=0_1$. 代入便证明了 $0_1=0$. 这证明了零向量的唯一性. 现在任给 $\alpha\in\mathfrak{L}$, 由线性空间的定义可知向量 α 的负向量 $-\alpha$ 存在. 若另有一个负向量 $\beta\in\mathfrak{L}$. 由定义, 则有 $\alpha+\beta=\beta+\alpha=0$. 但是已知 $(-\alpha)+\alpha=\alpha+(-\alpha)=0$, 所以有 $\alpha+\beta=0=\alpha+(-\alpha)$. 由加法消去律成立, 故证明了 $\beta=-\alpha$. 这证明了负向量的唯一性.

(4) 今 $b\alpha=0$. 若 $b=0$, 则由 $0+0=0$, 所以 $0\alpha=(0+0)\alpha$. 由纯量积的性质**(6)**, $(0+0)\alpha=0\alpha=0\alpha+0\alpha$. 由加法消去律成立, 故证明了 $0\alpha=0$. 若 $b\neq0$, 但是 $b\alpha=0$. 则由于存在 b^{-1}, 因此 $b^{-1}0=b^{-1}(b\alpha)$. 由纯量积的性质**(7)** 和**(8)**,

$$0=b^{-1}0=b^{-1}(b\alpha)=(b^{-1}b)\alpha=1\alpha=\alpha.$$

故证明了 $\alpha=0$. 因此证明了 $b\alpha=0$ 可以推出 $b=0$ 或 $\alpha=0$. 这证明了必要性. 反之. 已证了 $0\alpha=0$. 下面证若 $\alpha=0$, 则 $b\alpha=0$. 事实上, 由零向量的定义, 有 $a0=a(0+0)$, 由纯量积的性质**(5)** 可知 $a0=a0+a0$. 由加法消去律成立, 故证明了 $a0=0$, 所以充分性成立.

最后, 若 $u\neq0$, $v,w\in\mathbb{F}$, 又 $uv\alpha=uw\beta$, 则 $uv\alpha-uw\beta=0$, 因此, $u(v\alpha-w\beta)=0$. 今 $u\neq0$, 所以 $v\alpha-w\beta=0$, 即 $v\alpha=w\beta$. 这证明了**(4)** 成立.

(5) 由性质**(4)** 可知 $0=0\alpha=(b+(-b))\alpha$. 再由纯量积的性质**(6)** 可知 $0=b\alpha+(-b)\alpha$. 由负向量的定义, 有 $(-b)\alpha=-(b\alpha)$. 同理, $0=b0=b((-\alpha)+\alpha)$, 再由纯量积的性质**(5)** 可知 $0=b((-\alpha)+b\alpha$. 由负向量的定义, 有 $b(-\alpha)=-(b\alpha)$. 这证明了**(5)** 成立. □

注意 上面证明中, 每步都是根据定义条件, 或定义推出的性质出发, 推得所要的结论. 这就是所谓的 "说话要有根据".

从定义出发, 通过逻辑推理, 派生出整个数学模式是研究代数学的重要方法.

为了展开线性空间理论, 最基本, 最重要的概念是从加法和纯量积这两种代数运算所导出的三个概念: 线性相关、线性无关和线性组合.

定义 5.1.5 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中任意 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为**线性相关的**, 如果存在 m 个不全为零的数 $b_1,b_2,\cdots,b_m\in\mathbb{F}$, 使得

$$b_1\alpha_1+b_2\alpha_2+\cdots+b_m\alpha_m=0. \tag{5.1.1}$$

否则 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 称为**线性无关的**.

由定义可知, 包含零向量的任意有限个向量必线性相关. 特别零向量必线性相关. 证明如下: 给定 m 个向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$, 设 $\alpha_i=0$. 我们取 $a_i=1$, $a_j=0$, $j\neq i$, $1\leqslant j\leqslant m$. 则由引理 5.1.4 的**(4)**, 有 $a_1\alpha_1+\cdots+a_i\alpha_i+\cdots+a_m\alpha_m=0$. 由于

$a_1, \cdots, a_i, \cdots, a_m$ 不全为零, 这证明了 $\alpha_1, \cdots, \alpha_i, \cdots, \alpha_m$ 线性相关. 由定义又可知一个非零向量必线性无关. 用反证法来证明. 今若 $\alpha \neq 0$ 线性相关, 则存在数 $b \neq 0$, 有 $b\alpha = 0$. 这和引理 5.1.4 的(4) 矛盾. 所以证明了: 当 $\alpha \neq 0$, 则 α 线性无关. 同理, 向量组 $\{\alpha, -\alpha\}$, $\{\alpha, a\alpha\}$ 以及 $\{\alpha, a\alpha, \alpha_1, \alpha_2, \cdots, \alpha_m\}$ 等等也都线性相关. 一般有

定理 5.1.6 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 的任一向量组中, 如果有一部分向量线性相关, 那末整个向量组也必定线性相关.

证 设向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 中 r 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性相关, 那末存在 r 个不全为零的数 $b_1, b_2, \cdots, b_r$, 使得 $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_r\alpha_r = 0$. 于是

$$b_1\alpha_1 + b_2\alpha_2 + \cdots + b_r\alpha_r + 0 \cdot \alpha_{r+1} + \cdots + 0 \cdot \alpha_m = 0.$$

由于 m 个数 $b_1, b_2, \cdots, b_r, 0, \cdots, 0$ 仍不全为零, 由定义可知 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关. □

由定义又可知

定理 5.1.7 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性无关的必要且充分条件是: 如果有 m 个实数 $b_1, b_2, \cdots, b_m$, 使得 $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m = 0$, 那末必然有 $b_1 = b_2 = \cdots = b_m = 0$.

证 由线性相关和线性无关的定义可知, 向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 线性无关当且仅当不存在 m 个不全为零的数 $b_1, b_2, \cdots, b_m$, 使得

$$b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m = 0.$$

此即由 $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m = 0$ 可推出 $b_1 = b_2 = \cdots = b_m = 0$. □

由定理 5.1.6 很容易得到

定理 5.1.8 若域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中的向量组线性无关, 则它的任一部分也必定线性无关.

证 如若不然, 便有一部分向量线性相关. 由定理 5.1.6, 由部分向量线性相关可以推出整个向量组线性相关, 这和假设矛盾, 所以定理成立. □

定理 5.1.6 和定理 5.1.8 的逆命题都不成立. 即若向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关, 并不能保证它的任一部分也线性相关. 反之, 若向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的任一部分线性无关, 也不能保证向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性无关. 例如向量组 $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,1)$ 就是这样的例子.

从线性相关和线性无关的概念可知, 对向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 而言, 问题在于存不存在一组不全为零的数 $b_1, b_2, \cdots, b_m$, 使得 $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m = 0$. 由

此可见, 线性无关就是不存在这样的数组, 线性相关就是存在这样的数组. 而且, 在存在时, 只要能找到一组就够了.

定义 5.1.9 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中向量 β 称为是 m 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的**线性组合**, 如果存在 m 个数 $b_1, b_2, \cdots, b_m$, 使得

$$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m. \tag{5.1.2}$$

数 $b_1, b_2, \cdots, b_m$ 称为**组合系数**.

关于线性组合, 有下面性质:

定理 5.1.10 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关的必要且充分条件是在 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 中存在一个向量 α_j, 它是其余 $m-1$ 个向量的线性组合.

证 先证必要性. 今 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关, 故存在 m 个不全为零的数 $b_1, b_2, \cdots, b_m$, 使得 $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_m\alpha_m = 0$. 如果数 $b_j \neq 0$, 则这个等式可以改写为

$$\alpha_j = -\frac{b_1}{b_j}\alpha_1 - \cdots - \frac{b_{j-1}}{b_j}\alpha_{j-1} - \frac{b_{j+1}}{b_j}\alpha_{j+1} - \cdots - \frac{b_m}{b_j}\alpha_m.$$

所以必要性成立. 再证充分性. 设若 α_j 是其余 $m-1$ 个向量的线性组合, 亦即存在 $m-1$ 个数 $b_1, \cdots, b_{j-1}, b_{j+1}, \cdots, b_m$, 使得

$$\alpha_j = b_1\alpha_1 + \cdots + b_{j-1}\alpha_{j-1} + b_{j+1}\alpha_{j+1} + \cdots + b_m\alpha_m.$$

取 $b_j = -1$, 这就证明了 $b_1\alpha_1 + \cdots + b_m\alpha_m = 0$, 即 $\alpha_1, \cdots, \alpha_m$ 线性相关. □

利用线性相关、线性无关、线性组合这三个基本概念可以看出, 例 5.1.2 中有 n 个向量

$$e_1 = (1, 0, \cdots, 0)', \quad e_2 = (0, 1, 0, \cdots, 0)', \quad \cdots, \quad e_n = (0, \cdots, 0, 1)' \tag{5.1.3}$$

线性无关, 且任一向量是它们的线性组合, 即

$$(a_1, a_2, \cdots, a_n)' = a_1e_1 + a_2e_2 + \cdots + a_ne_n. \tag{5.1.4}$$

例 5.1.3 中有 nm 个向量 $E_{jk}^{(n,m)}$, $j = 1, 2, \cdots, n$, $k = 1, 2, \cdots, m$ 线性无关, 其中 E_{jk} 是第 j 行、第 k 列交叉位置的元素为 1, 其余位置元素为零的 $n \times m$ 矩阵. 且任一向量是它们的线性组合, 即

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \sum_{j=1}^{n}\sum_{k=1}^{m} a_{jk}E_{jk}. \tag{5.1.5}$$

例 5.1.4 中向量 $1, x, x^2, \cdots, x^n$ 线性无关, 又任一向量是它们的线性组合. 这些例子有一个特点, 即在 $\mathfrak{L}$ 中有有限个向量线性无关, 且任一向量是它们的线性组合. 然而例 5.1.5 和例 5.1.6 就没有这个性质, 这些不同点可以从 §5.2 的讨论中看出来.

习 题 5.1

5.1.1 试判别下列向量是否线性相关:

(1) $\alpha_1 = (1,3,2)'$, $\alpha_2 = (-1,1,3)'$;

(2) $\alpha_1 = (1,3,1)'$, $\alpha_2 = (-1,3,1)'$, $\alpha_3 = (-5,-7,3)'$;

(3) $\alpha_1 = (5,7,8)'$, $\alpha_2 = (3,7,10)'$, $\alpha_3 = (4,8,9)'$, $\alpha_4 = (10,11,13)'$.

5.1.2 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上线性空间. 试证:

(1) 若 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 中有一个零向量, 则它们一定线性相关;

(2) 设 $1 \leqslant k \leqslant n$. 若 k 维向量 $\alpha_j = (a_{j1}, a_{j2}, \cdots, a_{jk})'$, $1 \leqslant j \leqslant r$ 线性无关, 则 n 维向量 $\beta_i = (a_{i1}, a_{i2}, \cdots, a_{ik}, a_{i,k+1}, \cdots, a_{in})'$, $1 \leqslant i \leqslant r$ 也线性无关;

(3) 若 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性相关, 其中 $2 \leqslant r$. 则一定存在 r 个不全为零的数 $b_1, b_2, \cdots, b_r \in \mathbb{F}$, 使得对任一向量 α_{r+1}, 向量组 $\alpha_1 + b_1\alpha_{r+1}, \alpha_2 + b_2\alpha_{r+1}, \cdots, \alpha_r + b_r\alpha_{r+1}$ 线性相关.

5.1.3 试证: 域 $\mathbb{F}$ 上方阵 A 的属于不同特征根的特征向量线性无关.

5.1.4 如果 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是域 $\mathbb{F}$ 上方阵 A 的分别属于特征根 $b_1, b_2, \cdots, b_m$ 的特征向量, 且如果 $\alpha_1 + \alpha_2 + \cdots + \alpha_m$ 也是方阵 A 的特征根, 又 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的属于同一特征根的特征向量的和不等于零. 试证: $b_1 = \cdots = b_m$.

§5.2 基和基变换

利用上节的概念和性质, 我们来引进基和维数概念.

任给域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$. 如果 $\mathfrak{L}$ 只由一个元素构成, 那末这个元素必为零向量. 这时记作 $\mathfrak{L} = 0$, 称 $\mathfrak{L}$ 为**零维线性空间**.

如果在线性空间 $\mathfrak{L}$ 中不止一个元素, 即 $\mathfrak{L} \neq 0$, 于是在线性空间 $\mathfrak{L}$ 中有一个非零向量 α_1, 它自然线性无关. 因此有两种可能:

(I) 首先, 可能 $\mathfrak{L}$ 中任一向量 β 都和 α_1 线性相关. 即存在纯量 a, c 不全为零, 且 $a\beta + c\alpha_1 = 0$. 自然, $a \neq 0$. 因为若 $a = 0$, 则 $c\alpha_1 = 0$. 但是 $\alpha_1 \neq 0$, 所以 $c = 0$. 这和 a, c 不全为零矛盾. 因此证明了 $a \neq 0$, 即 $\beta = -(c/a)\alpha_1$, 所以 β 是 α_1 的纯量倍. 这证明了: 记

$$\mathfrak{L}_1 = \{\, a\alpha_1 \mid \forall a \in \mathbb{F} \,\} \subset \mathfrak{L}, \tag{5.2.1}$$

则 $\mathfrak{L} = \mathfrak{L}_1$. 这时 $\mathfrak{L}$ 称为**一维线性空间**, α_1 称为 $\mathfrak{L}$ 的一组**基**.

(II) 其次, 如果 $\mathfrak{L}_1 \subset \mathfrak{L}$, $\mathfrak{L}_1 \neq \mathfrak{L}$. 任取 $\alpha_2 \in \mathfrak{L}$, $\alpha_2 \notin \mathfrak{L}_1$, 则 α_2 和 α_1 线性无关. 因此得到两个线性无关的向量 α_1 和 α_2. 于是又出现两种情形: 一种情形为 $\mathfrak{L}$ 中任一向量 β 都和 α_1, α_2 线性相关. 因此, 存在纯量 a, c 和 d 不全为零, 且 $a\beta + c\alpha_1 + d\alpha_2 = 0$. 自然, $a \neq 0$. 因为若 $a = 0$, 则 $c\alpha_1 + d\alpha_2 = 0$. 但是 α_1

和 α_2 线性无关, 所以 $c = d = 0$. 这和 a, c, d 不全为零矛盾. 因此 $a \neq 0$, 即 $\beta = -(c/a)\alpha_1 - (d/a)\alpha_2$. 所以记

$$\mathfrak{L}_2 = \{\, c\alpha_1 + d\alpha_2 \mid \forall\, c, d \in \mathbb{F} \,\} \subset \mathfrak{L}, \tag{5.2.2}$$

则 $\mathfrak{L} = \mathfrak{L}_2$. 这时 $\mathfrak{L}$ 称为**二维线性空间**, α_1, α_2 称为 $\mathfrak{L}$ 的一组基; 另一种情形为 $\mathfrak{L}_2 \subset \mathfrak{L}$, $\mathfrak{L}_2 \neq \mathfrak{L}$. 这时任取 $\alpha_3 \in \mathfrak{L}$, $\alpha_3 \notin \mathfrak{L}_1$, 则 α_3 和 α_1, α_2 线性无关. 这样依次讨论下去, 用归纳法的思想, 便出现了两种情形:

(I′) 在线性空间 $\mathfrak{L}$ 中找不到有限个向量, 它们线性无关, 且在 $\mathfrak{L}$ 中任一向量和它们线性相关. 这种线性空间称为**无限维线性空间**. 关于域 $\mathbb{F}$ 上无限维线性空间的研究, 已超出这个课程的范围, 读者在泛函分析课中可以学到. 我们今后只讨论非无限维线性空间.

(II′) 在线性空间 $\mathfrak{L}$ 中找到有限个向量, 它们线性无关, 而且线性空间 $\mathfrak{L}$ 中任一向量和它们线性相关. 这类线性空间称为**有限维线性空间**. 它们有下面性质:

定义 5.2.1 如果在域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中找到 n 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关, 且 $\mathfrak{L}$ 中任一向量和它们线性相关, 那末 $\mathfrak{L}$ 称为 ***n* 维线性空间**. 非负整数 n 称为线性空间 $\mathfrak{L}$ 的**维数**, 记作 $\dim(\mathfrak{L})$. 这时 $\mathfrak{L}$ 称为域 $\mathbb{F}$ 上的 ***n* 维线性空间**. 又 α_1, α_2, $\cdots$, α_n 称为线性空间 $\mathfrak{L}$ 的一组 **基**. 这组基的向量 α_j 称为第 j 个**基向量**, $1 \leqslant j \leqslant n$.

定理 5.2.2 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 n 维线性空间 $\mathfrak{L}$ 的一组基, 则向量 β 可唯一地表为 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合:

$$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n. \tag{5.2.3}$$

将组合系数排成 $n \times 1$ 矩阵 $(b_1, \cdots, b_n)'$, 称为向量 β 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的 **坐标**. b_j 称为向量 β 在第 j 个方向的分量或**第 *j* 个分量**, 或 **第 *j* 个坐标**.

证 任取向量 β. 因为 β, α_1, α_2, $\cdots$, α_n 线性相关, 所以存在 $n+1$ 个不全为零的数 c_0, c_1, c_2, $\cdots$, c_n, 使得 $c_0\beta + c_1\alpha_1 + \cdots + c_n\alpha_n = 0$. 今若 $c_0 = 0$, 那末有 $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = 0$. 但是由条件 α_1, α_2, $\cdots$, α_n 线性无关, 因此 $c_1 = \cdots = c_n = 0$. 这和 $c_0, c_1, c_2, \cdots, c_n$ 不全为零的选取矛盾, 所以 $c_0 \neq 0$. 于是有

$$\beta = -c_0^{-1}c_1b_1\alpha_1 - c_0^{-1}c_2\alpha_2 + \cdots - c_0^{-1}c_n\alpha_n.$$

这就证明了 β 是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合.

如果 β 有两种线性组合的方式:

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = c_1'\alpha_1 + c_2'\alpha_2 + \cdots + c_n'\alpha_n,$$

则有 $\sum_{j=1}^{n}(c_j-c_j')\alpha_j=0$. 由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关, 所以 $c_j-c_j'=0$, 即 $c_j'=c_j$, $1\leqslant j\leqslant n$. 这就证明了向量 β 分解为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性组合时, 组合系数由 β 唯一确定. □

定理 5.2.3 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间. $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为线性空间 $\mathfrak{L}$ 的一组基. 则线性空间 $\mathfrak{L}$ 中任意 $n+1$ 个向量必定线性相关. 所以维数 n 和基的选取无关, 即线性空间 $\mathfrak{L}$ 中任意一组基都由 n 个线性无关的向量构成.

证 任取 $n+1$ 个向量 $\beta_0,\beta_1,\cdots,\beta_n$, 下面证它们必线性相关. 即证存在 $n+1$ 个不全为零的数 $x_0^{(0)},x_1^{(0)},\cdots,x_n^{(0)}$, 使得

$$\sum_{j=0}^{n}x_j^{(0)}\beta_j=0.$$

事实上, 由定理 5.2.2, 有

$$\beta_j=\sum_{k=1}^{n}a_{kj}\alpha_k,\quad 0\leqslant j\leqslant n.$$

引进 $n+1$ 个独立未知数 $x_0,x_1,\cdots,x_n$, 向量方程

$$0=\sum_{j=0}^{n}x_j\beta_j=\sum_{j=0}^{n}x_j\Big(\sum_{k=1}^{n}a_{kj}\alpha_k\Big)=\sum_{k=1}^{n}\Big(\sum_{j=0}^{n}a_{kj}x_j\Big)\alpha_k.$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关, 所以向量方程的坐标表达式为

$$\sum_{j=0}^{n}a_{kj}x_j=0,\quad 1\leqslant k\leqslant n.$$

它是 $n+1$ 个自变量 $x_0,x_1,\cdots,x_n$ 和 n 个方程的齐次线性方程组. 它的系数矩阵和变量矩阵分别为

$$A=\begin{pmatrix}a_{10}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{n0}&\cdots&a_{nn}\end{pmatrix},\quad x=\begin{pmatrix}x_0\\ x_1\\ \vdots\\ x_n\end{pmatrix},$$

矩阵表达式为 $Ax=0$. 注意到 $0\leqslant\operatorname{rank}(A)\leqslant n$, 所以它的通解依赖于 $n+1-\operatorname{rank}(A)\geqslant1$ 个独立参数. 因此证明了它必有非零解. 所以证明了 $\beta_0,\beta_1,\cdots,\beta_n$ 线性相关.

最后证明维数 n 和基的选取无关. 事实上, 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为线性空间 $\mathfrak{L}$ 的一组基. 由基的定义 5.2.1, 它们线性无关, 且 $\mathfrak{L}$ 中任意 $n+1$ 个向量线性相关. 设线性空间 $\mathfrak{L}$ 另有一组基 $\beta_1,\beta_2,\cdots,\beta_m$. 它们仍线性无关, 所以 $m<n+1$; 且 $\mathfrak{L}$ 中任

意 $m+1$ 个向量线性相关. 所以由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关, 有 $n < m+1$. 这证明了 $m \leqslant n, n \leqslant m$, 即 $m = n$. 因此维数 n 和基的选取无关. □

上面的结论在这一节和下一节都起着重要作用. 它们相当于在 n 维线性空间中引进笛卡尔坐标系.

定理 5.2.4 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$, 则**基变换公式**为

$$\beta_j = \sum_{i=1}^{n} a_{ij}\alpha_i, \quad 1 \leqslant j \leqslant n. \tag{5.2.4}$$

它所决定的 n 阶方阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{5.2.5}$$

非异. 反之, 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 任取 n 阶非异方阵 $A = (a_{ij})$ 如上, 构造 n 个向量

$$\beta_j = \sum_{i=1}^{n} a_{ij}\alpha_i, \quad 1 \leqslant j \leqslant n. \tag{5.2.6}$$

则 $\beta_1, \beta_2, \cdots, \beta_n$ 为 $\mathfrak{L}$ 中另一组基. 因此, $\mathfrak{L}$ 中的基和 n 阶非异方阵按照基变换公式, 给出了一个到上的一一对应. 而 n 阶方阵 A 称为基变换的**方阵表示**.

证 令线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$, 则有

$$\beta_j = \sum_{i=1}^{n} a_{ij}\alpha_i, \quad \alpha_j = \sum_{i=1}^{n} b_{ij}\beta_i, \quad 1 \leqslant j \leqslant n.$$

所以有

$$\alpha_j = \sum_{k=1}^{n}\Big(\sum_{i=1}^{n} a_{ki}b_{ij}\Big)\alpha_k, \quad 1 \leqslant j \leqslant n.$$

由于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 $\mathfrak{L}$ 的一组基, 即线性无关, 所以记

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

则有

$$AB = BA = E_n.$$

所以, $B = A^{-1}$, 即 A 非异, 而且 B 为 A 的逆方阵.

反之, 任取 n 阶非异方阵 $A=(a_{ij})$ 如上. 构造 n 个向量

$$\beta_j=\sum_{i=1}^{n}a_{ij}\alpha_i,\quad 1\leqslant j\leqslant n.$$

我们来证 $\beta_1,\beta_2,\cdots,\beta_n$ 为 $\mathfrak{L}$ 的一组基, 即线性无关. 事实上, 由定理 5.1.7, 我们来证由 $\sum\limits_{j=1}^{n}a_j\beta_j=0$ 可推出 $a_1=\cdots=a_n=0$. 因此, $\beta_1,\beta_2,\cdots,\beta_n$ 线性无关. 今

$$\sum_{j=1}^{n}a_j\beta_j=\sum_{i,j=1}^{n}a_ja_{ij}\alpha_i=\sum_{i=1}^{n}\Big(\sum_{j=1}^{n}a_ja_{ij}\Big)\alpha_i=0.$$

但是 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 $\mathfrak{L}$ 的一组基, 所以

$$\sum_{j=1}^{n}a_ja_{ij}=0,\quad 1\leqslant i\leqslant n.$$

今假设 A 是非异方阵, 由 Cramer 法则可知 $a_1=\cdots=a_n=0$. 所以证明了 $\beta_1,\beta_2,\cdots,\beta_n$ 线性无关. □

定理 5.2.5 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任取两组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$. 记基变换的方阵表示为 A. 在线性空间 $\mathfrak{L}$ 中取定向量 α, 设 α 关于基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 的坐标分别为 x 和 y, 则有如下 **坐标变换公式**

$$y=A^{-1}x,\quad x=Ay,\tag{5.2.7}$$

其中

$$x=(x_1,x_2,\cdots,x_n)',\quad y=(y_1,y_2,\cdots,y_n)'.\tag{5.2.8}$$

证 今基变换公式为

$$\beta_j=\sum_{i=1}^{n}a_{ij}\alpha_i,\quad 1\leqslant j\leqslant n.$$

而

$$\alpha=\sum_{i=1}^{n}x_i\alpha_i=\sum_{j=1}^{n}y_j\beta_j=\sum_{i=1}^{n}\Big(\sum_{j=1}^{n}y_ja_{ij}\Big)\alpha_i,$$

所以 $x_i=\sum\limits_{j=1}^{n}y_ja_{ij}$. 写成矩阵形式, 为 $Ay=x$, $y=A^{-1}x$. 所以, 在基变换下, 同一向量在不同基下的坐标变换公式为 $y=A^{-1}x$. □

例 5.2.1 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$, 试证: 存在 $1,2,\cdots,n$ 的排列 $i_1i_2\cdots i_n$, 使得

$$\alpha_1,\quad \alpha_2,\quad \cdots,\quad \alpha_{j-1},\quad \beta_{i_j},\quad \alpha_{j+1},\quad \cdots,\quad \alpha_n,\quad j=1,2,\cdots,n\tag{5.2.9}$$

都是 $\mathfrak{L}$ 的基.

证 首先, 我们将条件和结论用矩阵语言表达如下. 今基 $\beta_1, \beta_2, \cdots, \beta_n$ 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的基变换公式为

$$\beta_k = \sum_{p=1}^{n} b_{pk}\alpha_p, \quad 1 \leqslant k \leqslant n,$$

而基变换公式决定的非异方阵为

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

假设 (5.2.9) 是线性空间 $\mathfrak{L}$ 的基, 则有基变换公式:

$$\beta_{i_j} = \sum_{p=1}^{n} b_{pi_j}\alpha_p, \quad \alpha_k = \alpha_k, \quad k = 1, 2, \cdots, j-1, j+1, \cdots, n.$$

而基变换公式决定的非异方阵为

$$B_j = \begin{pmatrix} E_{j-1} & \xi & 0 \\ 0 & b_{ji_j} & 0 \\ 0 & \eta & E_{n-j} \end{pmatrix}, \quad 1 \leqslant j \leqslant n,$$

其中 $\xi = (b_{1,i_j}, \cdots, b_{j-1,i_j})'$ 是 $(j-1)\times 1$ 矩阵, $\eta = (b_{j+1,i_j}, \cdots, b_{n,i_j})'$, 是 $(n-j)\times 1$ 矩阵.

显然, 为了使得 B_j 为非异方阵, 我们必须有 $\det(B_j) = b_{j,i_j} \neq 0$, $1 \leqslant j \leqslant n$. 所以为了证明结论, 只要证明存在 $1, 2, \cdots, n$ 的排列 $i_1 i_2 \cdots i_n$, 使得 $b_{1i_1}b_{2i_2}\cdots b_{ni_n} \neq 0$ 就可以了.

回想行列式的定义, 我们知道 $\delta_{i_1 i_2 \cdots i_n}^{1\,2\,\cdots\,n} b_{1i_1}b_{2i_2}\cdots b_{ni_n}$ 是方阵 B 的行列式 $\det(B)$ 的某一项, 其中 $i_1 i_2 \cdots i_n$ 是 1, 2, $\cdots$, n 的排列. 由 $\det(B) \neq 0$ 可知至少存在一项不等于零. 这证明了结论. □

最后, 我们来考查不同的域上 n 维线性空间间的关系. 今 $\mathbb{R} \subset \mathbb{C}$.

定理 5.2.6 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 n 维实线性空间 $\mathfrak{L}$ 的一组基. 形式地引进集合

$$\mathfrak{L}^{\mathbb{C}} = \{ \sum_{k=1}^{n} c_k\alpha_k \mid c_k = a_k + \sqrt{-1}b_k \in \mathbb{C}, \quad 1 \leqslant k \leqslant n \}, \tag{5.2.10}$$

则 $\mathfrak{L}^{\mathbb{C}}$ 是复线性空间, 称为实线性空间 $\mathfrak{L}$ 的**复化**. 而且 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $\mathfrak{L}^{\mathbb{C}}$ 的一组基. 因此

$$\dim_{\mathbb{R}}(\mathfrak{L}) = \dim_{\mathbb{C}}(\mathfrak{L}^{\mathbb{C}}) = n, \tag{5.2.11}$$

其中 $\dim_{\mathbb{F}}(\mathfrak{L})$ 表示域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 的维数.

证 事实上, 我们先来证明集合 $\mathfrak{L}^{\mathbb{C}}$ 的定义和实线性空间 $\mathfrak{L}$ 的基的选取无关. 在 $\mathfrak{L}$ 中任取另一组基 $\beta_1, \beta_2, \cdots, \beta_n$, 因此有基变换公式

$$\beta_j = \sum_{i=1}^{n} a_{ij}\alpha_i, \quad a_{ij} \in \mathbb{R}, \quad 1 \leqslant i, j \leqslant n.$$

所以

$$\sum_{j=1}^{n} c_j\beta_j = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} c_j a_{ij}\Big)\alpha_i.$$

由 $c_j a_{ij} \in \mathbb{C}$, $1 \leqslant i, j \leqslant n$, 因此 $\sum\limits_{j=1}^{n} c_j\beta_j \in \mathfrak{L}^{\mathbb{C}}$. 这证明了用基 $\beta_1, \beta_2, \cdots, \beta_n$ 定义的复化就是用基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 定义的复化.

再证 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $\mathfrak{L}^{\mathbb{C}}$ 的一组基. 事实上, 如果存在复数 $c_1, c_2, \cdots, c_n$, 使得

$$\sum_{k=1}^{n} c_k\alpha_k = \sum_{k=1}^{n}(a_k + \sqrt{-1}b_k)\alpha_k = \sum_{k=1}^{n} a_k\alpha_k + \sqrt{-1}\sum_{k=1}^{n} b_k\alpha_k = 0,$$

则

$$\sum_{k=1}^{n} a_k\alpha_k = \sum_{k=1}^{n} b_k\alpha_k = 0.$$

由于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为实线性空间 $\mathfrak{L}$ 的一组基, 所以 $a_k = b_k = 0$, 即 $c_k = 0$, $1 \leqslant k \leqslant n$. □

定理 5.2.7 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为 n 维复线性空间 $\mathfrak{L}_c$ 的一组基, 形式地引进集合

$$\mathfrak{L}_c^{\mathbb{R}} = \{ \sum_{k=1}^{n} a_k\alpha_k \mid a_k \in \mathbb{R}, \quad 1 \leqslant k \leqslant n \}, \tag{5.2.12}$$

则 $\mathfrak{L}_c^{\mathbb{R}}$ 是实线性空间, 称为复线性空间 $\mathfrak{L}_c$ 的**实形式**. 而且 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $\mathfrak{L}_c^{\mathbb{R}}$ 的一组基, 因此

$$\dim_{\mathbb{R}}(\mathfrak{L}_c^{\mathbb{R}}) = \dim_{\mathbb{C}}(\mathfrak{L}_c) = n. \tag{5.2.13}$$

证 事实上, 由于实数为复数, 所以 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 实线性无关, 即为实线性空间 $\mathfrak{L}_c^{\mathbb{R}}$ 的一组基. □

注意 (1) 复线性空间 $\mathfrak{L}_c$ 的实形式 $\mathfrak{L}_c^{\mathbb{R}}$ 的定义和复线性空间 $\mathfrak{L}_c$ 的基的选取有关. 所以对不同的基, 有可能给出不同的实形式. 另一方面, 我们有: 设 $\mathfrak{L}$ 为 n 维实线性空间, 则 $\mathfrak{L}$ 是它的复化 $\mathfrak{L}^{\mathbb{C}}$ 的实形式. 设 $\mathfrak{L}_c$ 为 n 维复线性空间, 则 $\mathfrak{L}_c$ 的任一实形式 $\mathfrak{L}_c^{\mathbb{R}}$ 的复化 $(\mathfrak{L}_c^{\mathbb{R}})^{\mathbb{C}} = \mathfrak{L}_c$.

(2) 设 $\mathfrak{L}_c$ 为 n 维复线性空间, 而且 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $\mathfrak{L}_c$ 的一组基. 因此 $\mathfrak{L}_c$ 中任一向量

$$\alpha=\sum_{k=1}^{n}c_k\alpha_k=\sum_{k=1}^{n}(a_k+\sqrt{-1}b_k)\alpha_k=\sum_{k=1}^{n}a_k\alpha_k+\sum_{k=1}^{n}b_k\sqrt{-1}\alpha_k.$$

令 $\beta_k=\sqrt{-1}\alpha_k$, $1\leqslant k\leqslant n$, 则 $\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_n$ 实线性无关. 以它们作为一组基的实线性空间记作 $(\mathfrak{L}_c)_{\mathbb{R}}$. 因此, 设 $\mathfrak{L}$ 是以 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为一组实基的 n 维实线性空间, 则 $\sqrt{-1}\mathfrak{L}$ 仍为 n 维实线性空间, 它以 $\sqrt{-1}\alpha_1, \sqrt{-1}\alpha_2, \cdots, \sqrt{-1}\alpha_n$ 为一组实基. 而且作为集合, 有 $\mathfrak{L}^{\mathbb{C}}=\mathfrak{L}+\sqrt{-1}\mathfrak{L}$, 而作为实线性空间, $\mathfrak{L}+\sqrt{-1}\mathfrak{L}$ 是 $2n$ 维的.

习　题　5.2

5.2.1 求向量 $\alpha=(a_1, a_2, \cdots, a_n)'$ 在基 $\alpha_1=(1,1,\cdots,1)'$, $\alpha_2=(1,1,\cdots,1,0)'$, $\cdots$, $\alpha_n=(1,0,\cdots,0)'$ 下的坐标.

5.2.2 给定域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$. 设 $\mathfrak{L}$ 中有 n 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 它们适合条件:

(1) 对任一向量 α, 则 α 是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合, 即有分解式 $\alpha=\sum\limits_{j=1}^{n}a_j\alpha_j$;

(2) 设对某一固定的向量 α_0, 它关于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合的系数是唯一的. 试证: $\mathfrak{L}$ 是域 $\mathbb{F}$ 上 n 维线性空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成线性空间 $\mathfrak{L}$ 的一组基.

5.2.3 设 $\mathfrak{L}$ 是由所有次数不超过 n 的域 $\mathbb{F}$ 上的多项式构成的线性空间. 试证: 1, $x+1$, $(x+1)^2, \cdots, (x+1)^n$ 是 $\mathfrak{L}$ 的一组基. 试求 $\mathfrak{L}$ 中向量 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$ 在这组基下的坐标.

5.2.4 试证: 所有域 $\mathbb{F}$ 上 n 阶对称方阵构成域 $\mathbb{F}$ 上 $n(n+1)/2$ 维线性空间; 所有 n 阶反对称方阵构成域 $\mathbb{F}$ 上 $n(n-1)/2$ 维线性空间.

5.2.5 记 $\beta_1, \beta_2, \cdots, \beta_m$ 为 n 维线性空间 $\mathfrak{L}$ 中线性相关的向量组. 在其中任取 $m-1$ 个向量, 则必线性无关. 设 m 个纯量 $c_1, c_2, \cdots, c_m$ 有 $\sum\limits_{j=1}^{m}c_j\beta_j=0$, 则 $c_1c_2\cdots c_m\neq 0$ 或者 $c_1=c_2=\cdots=c_m=0$. 当 m 个纯量 $c_1c_2\cdots c_m\neq 0$ 时, 如果存在 m 个纯量 $d_1, d_2, \cdots, d_m$, 使得 $\sum\limits_{j=1}^{m}d_j\beta_j=0$, 则有 $c_1:d_1=c_2:d_2=\cdots=c_m:d_m$.

5.2.6 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为线性空间 $\mathfrak{L}$ 的一组基. 如果 $\beta\in\mathfrak{L}$, 而且 β 为 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 中任意 $n-1$ 个向量的线性组合, 试证: $\beta=0$.

§5.3 线性同构

由上节可知, 在域 $\mathbb{F}$ 上的 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则线性空间 $\mathfrak{L}$ 中任一向量 α 有分解式

$$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n. \tag{5.3.1}$$

即 α 对应了 n 维向量空间 $\mathbb{F}^n$ 中一向量 $(a_1, a_2, \cdots, a_n)'$, 其中 $(*)'$ 为矩阵 $(*)$ 的转置矩阵. 且由 α 的分解方式的唯一性可知这给出了 n 维线性空间 $\mathfrak{L}$ 到 $\mathbb{F}^n$ 上的一一

对应. 今设 $\alpha, \beta \in \mathfrak{L}$, $c, d \in \mathbb{F}$, 由

$$\alpha \longrightarrow (a_1, a_2, \cdots, a_n)', \qquad \beta \longrightarrow (b_1, b_2, \cdots, b_n)', \tag{5.3.2}$$

则由 $c\alpha + d\beta = c\sum_{j=1}^{n} a_j\alpha_j + d\sum_{j=1}^{n} b_j\alpha_j = \sum_{j=1}^{n}(ca_j + db_j)\alpha_j$ 可知

$$c\alpha + d\beta \longrightarrow (ca_1 + db_1, ca_2 + db_2, \cdots, ca_n + db_n)'. \tag{5.3.3}$$

换句话说, 在上述一一对应下, $\mathfrak{L}$ 中向量的和及纯量积对应 $\mathbb{F}^n$ 中向量的和及纯量积. 从这一点可以看出基的重要性. 将这个性质抽象化, 便引出到上的线性同构的概念.

定义 5.3.1 对于域 $\mathbb{F}$ 上两个线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$, 如果存在 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上 (内) 的一一映射 σ, 使得它适合条件:

(1) $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$, $\forall\, \alpha\ \beta \in \mathfrak{L}$;

(2) $\sigma(b\alpha) = b\sigma(\alpha)$, $\forall\, \alpha \in \mathfrak{L}$, $b \in \mathbb{F}$,

则映射 σ 称为线性空间 $\mathfrak{L}_1$ 到线性空间 $\mathfrak{L}_2$ 上 (内) 的**线性同构映射**, 此时 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 称为**线性同构**.

显然, 到上的线性同构映射的逆映射 σ^{-1} 是线性空间 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 上的线性同构映射.

定理 5.3.2 域 $\mathbb{F}$ 上线性空间之间到上的线性同构关系为等价关系.

证 **(1)** 作用恒等映射可知线性空间 $\mathfrak{L}$ 和自身线性同构, 这证明了反身性.

(2) 若 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的线性同构映射, 则它的逆映射 σ^{-1} 是线性空间 $\mathfrak{L}_2$ 到 $\mathfrak{L}_1$ 上的线性同构映射. 这证明了对称性.

(3) 最后, 若 σ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的线性同构映射, τ 为线性空间 $\mathfrak{L}_2$ 到 $\mathfrak{L}_3$ 上的线性同构映射, 则 $\tau \circ \sigma$ 为线性空间 $\mathfrak{L}_1$ 到 $\mathfrak{L}_3$ 上的线性同构映射. 这证明了传递性. 所以线性同构关系为等价关系. □

由定义可知, 对任意 s 个向量 $\beta_1, \beta_2, \cdots, \beta_s$ 及任意 s 个数 $a_1, a_2, \cdots, a_s$, 有

$$\sigma\Big(\sum_{j=1}^{s} a_j\beta_j\Big) = \sum_{j=1}^{s} a_j\sigma(\beta_j). \tag{5.3.4}$$

于是, 关于到上的线性同构映射, 有下面重要性质:

定理 5.3.3 设 σ 为域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}_1$ 到线性空间 $\mathfrak{L}_2$ 内的线性同构映射, 则 σ 把零向量变为零向量, 且线性同构映射保持线性关系不变.

证 先证 $\sigma(0) = 0$. 事实上, $\sigma(0) = \sigma(0 + 0) = \sigma(0) + \sigma(0)$. 所以 $\sigma(0) = 0$. 再证 σ 保持线性关系不变. 设若 $\mathfrak{L}_1$ 中的向量 $\beta_1, \beta_2, \cdots, \beta_s$ 线性相关, 则存在 s 个不全为零的实数 $a_1, a_2, \cdots, a_s$, 使得 $a_1\beta_1 + a_2\beta_2 + \cdots + a_s\beta_s = 0$. 因此

$$0 = \sigma(0) = \sigma(a_1\beta_1 + a_2\beta_2 + \cdots + a_s\beta_s) = a_1\sigma(\beta_1) + a_2\sigma(\beta_2) + \cdots + a_s\sigma(\beta_s).$$

由定义可知 $\mathfrak{L}_2$ 中向量 $\sigma(\beta_1), \sigma(\beta_2), \cdots, \sigma(\beta_s)$ 也线性相关. 所以, 在到内的线性同构映射 σ 下, 线性相关的向量组变为线性相关的向量组. 再设 $\mathfrak{L}_1$ 中向量 $\beta_1, \beta_2, \cdots, \beta_s$ 线性无关. 如果 $\sigma(\beta_1), \sigma(\beta_2), \cdots, \sigma(\beta_s)$ 线性相关, 则存在 s 个不全为零的实数 $a_1, a_2, \cdots, a_s$, 使得 $a_1\sigma(\beta_1)+a_2\sigma(\beta_2)+\cdots+a_s\sigma(\beta_s)=0$, 即 $\sigma(a_1\beta_1+a_2\beta_2+\cdots+a_s\beta_s)=0$. 由于 σ 是 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的一一对应, 且 $\sigma(0)=0$, 所以 $a_1\beta_1+a_2\beta_2+\cdots+a_s\beta_s=0$, 这和 $\beta_1, \beta_2, \cdots, \beta_s$ 线性无关矛盾. 因此, 在到内的线性同构映射 σ 下, 线性无关的向量组变为线性无关的向量组. □

从这个定理可知, 对于到上的线性同构的两个线性空间, 它们的性质完全一样 (我们所指的性质, 是指从线性空间的加法和纯量积这两种代数运算导出的性质). 因此, 很自然的, 我们希望在给定一个 n 维线性空间 $\mathfrak{L}$ 后, 找出和 $\mathfrak{L}$ 是到上的线性同构的所有线性空间. 这样, 这批线性空间的性质和 $\mathfrak{L}$ 的性质完全一样.

定理 5.3.4 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 对线性空间 $\mathfrak{L}$ 中任一向量 $\alpha=\sum\limits_{j=1}^{n} x_j\alpha_j$, 于是建立了线性空间 $\mathfrak{L}$ 到向量空间 $\mathbb{F}^n$ 上的线性同构映射

$$\sigma:\quad \alpha=\sum_{j=1}^{n} x_j\alpha_j \longrightarrow (x_1, x_2, \cdots, x_n)'. \tag{5.3.5}$$

因此任一域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 必线性同构于 $\mathbb{F}^n$. 所以在线性空间的到上的线性同构这个等价关系下, 等价类中可取代表元素为 n 维向量空间 $\mathbb{F}^n$.

证 由于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为基, 所以向量 α 的坐标表达唯一. 这证明了 σ 为到上的一一对应. 容易直接验证 σ 为到上的同构映射. □

定理 5.3.5 两个域 $\mathbb{F}$ 上有限维线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 是到上的线性同构的必要且充分条件是 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的维数相等. 所以线性空间在到上的线性同构关系下的全系不变量为它的维数.

证 先证必要性. 若 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 是到上的线性同构. 设 $\mathfrak{L}_1$ 的维数为 n. 由定理 5.3.3 可知, 在 $\mathfrak{L}_1$ 中有 n 个向量线性无关, 且任意 $n+1$ 个向量线性相关, 因此 $\mathfrak{L}_2$ 的维数也是 n, 所以 $\dim(\mathfrak{L}_1)=\dim(\mathfrak{L}_2)=n$. 再证充分性. 假设 $\dim(\mathfrak{L}_1)=\dim(\mathfrak{L}_2)=n$. 由定理 5.3.4 可知, 线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 都到上的线性同构于 n 维向量空间 $\mathbb{F}^n$. 因此, 线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 也是到上的线性同构. □

这两个定理告诉我们, 研究一般的 n 维线性空间和研究 n 维向量空间 $\mathbb{F}^n$ 是完全一样的, 因为它们必然到上的线性同构. 从这个角度来看, 尽管线性空间的定义比较抽象, 然而它实际上可以很具体地用 $\mathbb{F}^n$ 来理解. 但是在利用 $\mathbb{F}^n$ 来刻画 $\mathfrak{L}$ 时, 不要忘了是在 $\mathfrak{L}$ 中取定一组基的前提下刻画的.

为了方便起见, 今后在符号方面作如下的约定: 用大写德文花体字母 $\mathfrak{L}, \mathfrak{M}, \cdots$ 表示线性空间; 用小写希腊字母 $\alpha, \beta, \gamma, \cdots$ 表示向量, 例如 $\alpha_1, \alpha_2, \cdots$ 等表示向量.

用 $e_1, e_2, \cdots, e_n$ 表示 n 维向量空间 $\mathbb{F}^n$ 中一组确定的基, 其中 $e_j=(\delta_{1j},\delta_{2j},\cdots,\delta_{nj})$, $1\leqslant j\leqslant n$, δ_{jk} 为 Kronecker 符号; 用小写希腊字母 λ, μ, ν, ρ, $\cdots$ 表示纯量 (即域 **F** 中数); 用有足标的小写拉丁字母 a_1, a_2, $\cdots$, b_1, b_2, $\cdots$ 表示向量的分量或点的坐标, 它们也都是域 $\mathbb{F}$ 中的数. 又足标表示分量所在的位置; 用大写拉丁字母 A, B, C, $\cdots$ 表示矩阵; 用有两个足标的小写拉丁字母 a_{jk} 表示矩阵的元素, 足标表示该元素在矩阵中所在的行及列的位置. 另外, 0 既表示数 "零", 也表示零向量和零矩阵, 以及零维线性空间.

§5.4 子空间

定义 5.4.1 在域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中取定向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 和向量组 $\{\beta_1, \beta_2, \cdots, \beta_t\}$, 如果每个 β_j, $j=1,2,\cdots,t$ 都是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的线性组合, 则称**向量组 $\{\beta_1, \beta_2, \cdots, \beta_t\}$** 由向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ **线性表出**.

定理 5.4.2 (替换定理) 在域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中取定 m 个线性无关的向量 α_1, α_2, $\cdots$, α_m, 任给向量集 $\{\beta_1, \beta_2, \cdots, \beta_s\}$, 如果向量集 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 由向量集 $\{\beta_1, \beta_2, \cdots, \beta_s\}$ 线性表出, 则 $m\leqslant s$, 且在向量集 $\{\beta_1, \beta_2, \cdots, \beta_s\}$ 中存在子集合 $\{\beta_{i_1}, \beta_{i_2}, \cdots, \beta_{i_m}\}$, 使得在向量集 $\{\beta_1, \beta_2, \cdots, \beta_s\}$ 中除去子集合 $\{\beta_{i_1}, \beta_{i_2}, \cdots, \beta_{i_m}\}$ 后, 用 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 来替换, 这样得到的新的向量集和向量集 $\{\beta_1, \beta_2, \cdots, \beta_s\}$ 可以互相线性表出.

证 由条件

$$\alpha_j=\sum_{i=1}^{s}a_{ij}\beta_i, \quad j=1,2,\cdots,m,$$

于是有 $s\times m$ 矩阵

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1m}\\ \vdots&&\vdots\\ a_{s1}&\cdots&a_{sm}\end{pmatrix}.$$

引进 m 个独立未知数 x_1, x_2, $\cdots$, x_m. 于是由 α_1, α_2, $\cdots$, α_m 线性无关, 因此 $\sum\limits_{j=1}^{m}x_j\alpha_j=0$ 推出 $x_1=\cdots=x_m=0$. 今

$$\sum_{j=1}^{m}x_j\alpha_j=\sum_{i=1}^{s}\Big(\sum_{j=1}^{m}a_{ij}x_j\Big)\beta_i.$$

因此, 齐次线性方程组

$$\sum_{j=1}^{m}a_{ij}x_j=0, \quad 1\leqslant i\leqslant s$$

只有零解. 由定理 4.2.1, 有 $m-\operatorname{rank}(A)=0$. 因此证明了 $s\times m$ 矩阵 A 的秩为 m. 于是 $m\leqslant s$, 且在矩阵 A 中存在一个 m 阶子式不等于零. 如果 $s\times m$ 矩阵 A 的第

$i_1, i_2, \cdots, i_m$ 行决定的 m 阶子式不等于零, 它们和 $\beta_{i_1}, \beta_{i_2}, \cdots, \beta_{i_m}$ 有关. 为了方便起见, 我们改变指标如下: $i_j \to j$, $j = 1, 2, \cdots, m$. 即记 β_{i_j} 为 β_j, $1 \leqslant j \leqslant m$, 其他指标依次排下去. 因此, 我们无妨假设

$$\det(A_0) = \det\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} \neq 0.$$

我们来证明向量集

$$\mathfrak{S}_1 = \{\beta_1, \beta_2, \cdots, \beta_s\}, \qquad \mathfrak{S}_2 = \{\alpha_1, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_s\}$$

可以互相线性表出. 显然向量集 $\mathfrak{S}_2$ 可由向量集 $\mathfrak{S}_1$ 线性表出. 为了证向量集 $\mathfrak{S}_1$ 可由向量集 $\mathfrak{S}_2$ 线性表出, 只要证 β_j 可以由向量集 $\mathfrak{S}_2$ 线性表出就可以了.

记 $B = A_0^{-1} = (b_{ij})$, 则任取 $1 \leqslant i \leqslant m$, 对 β_i, 有

$$\begin{aligned} \beta_i &= \sum_{k=1}^{m} \delta_{ik}\beta_k = \sum_{k=1}^{m}\sum_{j=1}^{m} a_{kj}b_{ji}\beta_k \\ &= \sum_{j=1}^{m} b_{ji}\sum_{k=1}^{m} a_{kj}\beta_k = \sum_{j=1}^{m} b_{ji}\Big(\alpha_j - \sum_{k+1}^{s} a_{kj}\beta_k\Big). \end{aligned}$$

这证明了向量 β_i 由向量集 $\mathfrak{S}_2$ 线性表出, $1 \leqslant i \leqslant m$. 所以我们证明了定理. □

于是有

定理 5.4.3 在域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中取定两个线性无关的向量集 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 和 $\{\beta_1, \beta_2, \cdots, \beta_s\}$, 如果它们可以互相线性表出, 则有 $s = m$.

证 由定理 5.4.2, $m \leqslant s$, 同理, 有 $s \leqslant m$. 所以 $s = m$. □

定义 5.4.4 设 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中的子集合. $\mathfrak{S}_0$ 中有限个向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 称为子集合 $\mathfrak{S}_0$ 的一组 **极大线性无关向量组**, 如果 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是 $\mathfrak{S}_0$ 中一组线性无关的向量, 且 $\mathfrak{S}_0$ 中任一向量是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合, 即向量集 $\mathfrak{S}_0$ 由 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性表出.

假设 $\mathfrak{S}_0$ 只由零向量构成, 这时不存在一组极大线性无关向量组. 设 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中的非零子集合, 是否在 $\mathfrak{S}_0$ 中存在一组极大线性无关向量组? 这个回答是肯定的. 事实上, 我们有

定理 5.4.5 设 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中的非零子集合, 则子集合 $\mathfrak{S}_0$ 中必存在一组极大线性无关向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$, 且 $1 \leqslant r \leqslant n$, 又这时 $\mathfrak{S}_0$ 中任意 $r+1$ 个向量一定线性相关.

证 现在先给出一种在子集合 $\mathfrak{S}_0$ 中寻找极大线性无关向量组的方法.

假设 $\mathfrak{S}_0$ 中有非零向量, 任取一个非零向量为 α_1, 如果 $\mathfrak{S}_0$ 中任一向量是 α_1 的线性组合. 按照定义, $\{\alpha_1\}$ 就是 $\mathfrak{S}_0$ 的极大线性无关向量组, 否则, 在 $\mathfrak{S}_0$ 中一定存在向量 α_2, 使得 α_1, α_2 线性无关. 如果 $\mathfrak{S}_0$ 中任一向量是 α_1, α_2 的线性组合, 那末 $\{\alpha_1, \alpha_2\}$ 就是 $\mathfrak{S}_0$ 的极大线性无关向量组. 否则, 又存在和 α_1, α_2 线性无关的向量 α_3. 这样依次作下去, 由定理 5.2.3, $\mathfrak{S}_0$ 中任意 $n+1$ 个向量一定线性相关, 因此上述步骤至多到 n 步就不能再作下去了. 这样也就找到了一组向量 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$, 它是子集合 $\mathfrak{S}_0$ 的极大线性无关向量组. 这时 $\mathfrak{S}_0$ 中任意 $r+1$ 个向量一定线性相关. 事实上, 和定理 5.2.3 一样可证这一事实成立. □

任给子集合 $\mathfrak{S}_0$, 不但 $\mathfrak{S}_0$ 中存在极大线性无关向量组, 而且有

定理 5.4.6 设 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维实线性空间 $\mathfrak{L}$ 中的子集合. $\mathfrak{S}_0$ 中任意两个极大线性无关向量组中向量的个数相等.

证 假设子集合 $\mathfrak{S}_0$ 有两个极大线性无关向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ 和 $\{\beta_1, \beta_2, \cdots, \beta_s\}$. 由极大线性无关向量组的定义可知, 它们可以互相线性表出. 由定理 5.4.3, 所以 $r=s$. □

下面引进线性生成和子空间的概念.

定义 5.4.7 设 $\mathfrak{S}$ 是域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中的子集合. $\mathfrak{S}$ 称为线性子空间, 或简称为**子空间**, 如果它适合条件:

(1) 在子集合 $\mathfrak{S}$ 中任取向量 α 和 β, 则向量 $\alpha+\beta$ 也在 $\mathfrak{S}$ 中;

(2) 在子集合 $\mathfrak{S}$ 中任取向量 α, 再任取一数 $b\in\mathbb{F}$, 则 $b\alpha$ 也在 $\mathfrak{S}$ 中.

定理 5.4.8 域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 中的子空间 $\mathfrak{S}$ 关于 $\mathfrak{L}$ 的加法和纯量积仍封闭, 所以作为 $\mathfrak{S}$ 的加法和纯量积的定义, 则 $\mathfrak{S}$ 仍为线性空间.

证 由定义可知, 在子空间 $\mathfrak{S}$ 中按照 $\mathfrak{L}$ 的加法和纯量积也可以定义加法和纯量积, 而且加法交换律, 加法结合律以及线性空间定义中性质 (5), (6), (7), (8) 都适合. 所以只要验证性质 (3) 及 (4) 也适合就行了. 为此, 要证 $\mathfrak{L}$ 的零向量也在 $\mathfrak{S}$ 中, 又 $\mathfrak{S}$ 中任一向量的负向量也在 $\mathfrak{S}$ 中. 今在 $\mathfrak{S}$ 中任取向量 α, 则 $0\cdot\alpha=0$ 在 $\mathfrak{S}$ 中, 又 $(-1)\alpha=-\alpha$ 也在 $\mathfrak{S}$ 中. 所以证明了 $\mathfrak{S}$ 是 $\mathfrak{L}$ 的子空间. □

显然, 线性空间 $\mathfrak{L}$ 的子空间的子空间, 仍为线性空间 $\mathfrak{L}$ 的子空间.

我们知道, 域 $\mathbb{F}$ 上线性空间要末为有限维, 要末为无限维. 由维数定义可知, n 维线性空间 $\mathfrak{L}$ 的子空间也是有限维线性空间, 而且它的维数 r 不超过 n. 利用线性空间的一般理论可知, 子空间 $\mathfrak{S}$ 中任意一个极大线性无关向量组构成 $\mathfrak{S}$ 的一组基.

当 $r=0$ 时, $\mathfrak{S}$ 只由一个零向量构成, 这时 $\mathfrak{S}$ 称为**零维子空间**, 或简称为**零空间**, 记作 0. 当 $1\leqslant r\leqslant n-1$ 时, $\mathfrak{S}$ 称为**真子空间**, 当 $r=n$ 时, $\mathfrak{S}$ 就是整个空间 $\mathfrak{L}$.

另一方面, 子空间是由一组线性无关的向量, 作所有可能的线性组合构成. 一般有

定理 5.4.9 设 $\mathfrak{S}$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中的非零子空间, 则 $\mathfrak{S}$ 中向量组 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为子空间 $\mathfrak{S}$ 的一组基当且仅当 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为子空间 $\mathfrak{S}$ 的极大线性无关向量组. 这时, $0 < \dim(\mathfrak{S}) \leqslant n$.

定义 5.4.10 设 $\mathfrak{S}$ 和 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中的两个子集合, $\mathfrak{S}$ 称为由子集合 $\mathfrak{S}_0$ **线性生成**, 如果 $\mathfrak{S}$ 由 $\mathfrak{S}_0$ 中任意有限多个向量的任意线性组合构成. 这时记 $\mathfrak{S}$ 为 $[\mathfrak{S}_0]$.

定理 5.4.11 设 $\mathfrak{S}_0$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子集合, 则 $\mathfrak{S}_0$ 线性生成子空间 $\mathfrak{S} = [\mathfrak{S}_0]$. 特别, 子集合 $\mathfrak{S}_0$ 中任意一组极大线性无关向量组为此子空间的一组基.

证 显然 $\mathfrak{S}_0 \subset [\mathfrak{S}_0] = \mathfrak{S}$, 又因为任意有限多个向量 $\gamma_1, \gamma_2, \cdots, \gamma_s$ 的任意多个线性组合 $\xi_1, \xi_2, \cdots, \xi_s$ 的线性组合仍为 $\gamma_1, \gamma_2, \cdots, \gamma_s$ 的线性组合. 所以任取 α, $\beta \in \mathfrak{S}$, 则 $c\alpha + d\beta \in \mathfrak{S}, \forall c, d \in \mathbb{F}$, 即 $[\mathfrak{S}_0] = \mathfrak{S}$ 为子空间. 又由极大线性无关向量组的定义可知, $\mathfrak{S}_0$ 的极大线性无关向量组为子空间 $\mathfrak{S}$ 的一组基. □

现在给出基的一个重要性质.

定理 5.4.12 设 $\beta_1, \beta_2, \cdots, \beta_r$ 是 n 维线性空间 $\mathfrak{L}$ 中 r 个线性无关的向量, 于是 $r \leqslant n$. 且一定存在 $n-r$ 个向量 $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$, 使得 $\beta_1, \beta_2, \cdots, \beta_n$ 线性无关, 即构成 $\mathfrak{L}$ 的一组基.

证 如果 $r = n$, 定理不证自明. 假设 $r < n$, 如能证明存在向量 β_{r+1}, 使得 β_1, $\beta_2, \cdots, \beta_r, \beta_{r+1}$ 线性无关, 那末从 $\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}$ 出发, 用同法可找到 β_{r+2}, 这样依次作下去, 便能证明定理.

为了证明存在 β_{r+1}, 使得 $\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}$ 线性无关, 我们用反证法. 即若不存在, 那末 $\mathfrak{L}$ 中任一向量是 $\beta_1, \beta_2, \cdots, \beta_r$ 的线性组合. 由定理 5.2.3, $\mathfrak{L}$ 中任意 $r+1$ 个向量线性相关, 所以 $r+1 > n$. 即 $r \geqslant n$. 这和 $r < n$ 矛盾, 所以 β_{r+1} 存在. □

域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 和它的 r 维子空间 $\mathfrak{S}$ 间有很简单的关系: 在 $\mathfrak{S}$ 中任意取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_r$, 由定理 5.4.12, 在 $\mathfrak{L}$ 中一定存在 $n-r$ 个向量 α_{r+1}, $\alpha_{r+2}, \cdots, \alpha_n$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_n$ 是 $\mathfrak{L}$ 的一组基. 今子空间 $\mathfrak{S}$ 由 $\alpha_1, \cdots, \alpha_r$ 的一切线性组合构成, 所以任取 $\beta \in \mathfrak{S}$, 则 $\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$. 记作 $\beta = c_1\alpha_1 + \cdots + c_r\alpha_r + 0 \cdot \alpha_{r+1} + \cdots + 0 \cdot \alpha_n$, 则 β 作为 $\mathfrak{L}$ 的向量, 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的坐标为 $(c_1, \cdots, c_r, 0, \cdots, 0)$, 然而 β 作为 $\mathfrak{S}$ 的向量, 在基 $\{\alpha_1, \cdots, \alpha_r\}$ 下的坐标为 $(c_1, \cdots, c_r)$. 换句话说, 在 $\mathfrak{L}$ 中存在一组基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, 使得所有最后 $n-r$ 个分量为零的向量构成 $\mathfrak{L}$ 的 r 维子空间, 它的基是 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$.

关于子空间还可以引进运算: 子空间的交、和以及直接和.

定义 5.4.13　域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的两个子空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的**交**, 是由所有既在 $\mathfrak{L}_1$ 中, 也在 $\mathfrak{L}_2$ 中的向量构成的子集合, 记作 $\mathfrak{L}_1 \cap \mathfrak{L}_2 = \mathfrak{L}_2 \cap \mathfrak{L}_1$.

由定义可知, 子空间的交仍为子空间.

定义 5.4.14　域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的两个子空间 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 的**和**, 是由所有形如

$$\gamma = \alpha + \beta, \quad \alpha \in \mathfrak{L}_1, \quad \beta \in \mathfrak{L}_2 \tag{5.4.1}$$

的向量构成的子集合, 记作 $\mathfrak{L}_1 + \mathfrak{L}_2 = \mathfrak{L}_2 + \mathfrak{L}_1$.

由定义可知, 子空间的和仍为子空间, 且 $\mathfrak{L}_1 + \mathfrak{L}_2$ 以 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为其子空间. 又记 0 为由零向量构成的子空间, 则有 $\mathfrak{L}_1 + 0 = 0 + \mathfrak{L}_1 = \mathfrak{L}_1$.

关于子空间的交及和, 看作代数运算, 还有性质:

对三个子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \mathfrak{L}_3$, 则有

$$(\mathfrak{L}_1 \cap \mathfrak{L}_2) \cap \mathfrak{L}_3 = \mathfrak{L}_1 \cap (\mathfrak{L}_2 \cap \mathfrak{L}_3), \qquad \mathfrak{L}_1 + (\mathfrak{L}_2 + \mathfrak{L}_3) = (\mathfrak{L}_1 + \mathfrak{L}_2) + \mathfrak{L}_3. \tag{5.4.2}$$

所以括号可以取消. 因此, 可以引进有限个子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_t$ 的交及和, 它们分别记为

$$\mathfrak{L}_1 \cap \mathfrak{L}_2 \cap \cdots \cap \mathfrak{L}_t = \bigcap_{j=1}^{t} \mathfrak{L}_j, \quad \mathfrak{L}_1 + \mathfrak{L}_2 + \cdots + \mathfrak{L}_t = \sum_{j=1}^{t} \mathfrak{L}_j. \tag{5.4.3}$$

关于子空间的交及和的维数, 还有性质:

定理 5.4.15 (维数公式)　对域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任两子空间 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$, 有

$$\dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_2) = \dim(\mathfrak{L}_1 \cap \mathfrak{L}_2) + \dim(\mathfrak{L}_1 + \mathfrak{L}_2). \tag{5.4.4}$$

证　在子空间 $\mathfrak{L}_1 \cap \mathfrak{L}_2$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_r$. 它是子空间 $\mathfrak{L}_1$ 中 r 个线性无关的向量, 所以可以补足 s 个向量 $\beta_1, \beta_2, \cdots, \beta_s$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s$ 是 $\mathfrak{L}_1$ 的一组基. 同理, 可以补足 t 个向量 $\gamma_1, \gamma_2, \cdots, \gamma_t$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_r, \gamma_1, \gamma_2, \cdots, \gamma_t$ 是 $\mathfrak{L}_2$ 的一组基.

显然, 子空间 $\mathfrak{L}_1 + \mathfrak{L}_2$ 中任一向量是 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s, \gamma_1, \gamma_2, \cdots, \gamma_t$ 的线性组合. 我们来证, 它们是子空间 $\mathfrak{L}_1 + \mathfrak{L}_2$ 的一组基. 因此,

$$\dim(\mathfrak{L}_1 + \mathfrak{L}_2) = r + s + t = (r+s) + (r+t) - t = \dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_2) - \dim(\mathfrak{L}_1 \cap \mathfrak{L}_2).$$

即式 (5.4.4) 成立.

为了证 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s, \gamma_1, \gamma_2, \cdots, \gamma_t$ 线性无关, 考虑它们的线性组合

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

因此

$$\sum_{i=1}^{r} u_i\alpha_i + \sum_{j=1}^{s} v_j\beta_j = -\sum_{k=1}^{t} w_k\gamma_k \in \mathfrak{L}_1 \cap \mathfrak{L}_2.$$

但是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为 $\mathfrak{L}_1 \cap \mathfrak{L}_2$ 的基, 又 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s$ 为 $\mathfrak{L}_1$ 的基. 所以 $w_1 = \cdots = w_t = 0$, 而且 $\sum\limits_{i=1}^{r} u_i\alpha_i + \sum\limits_{j=1}^{s} v_j\beta_j = 0$. 由于 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s$ 为 $\mathfrak{L}_1$ 的基, 所以 $u_1 = \cdots = u_r = v_1 = \cdots = v_s = 0$. 这证明了定理. □

定义 5.4.16 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的 s 个子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_s$ 的和 $\mathfrak{S}$ 称为**直接和**, 如果 $\mathfrak{S}$ 中的任一向量能唯一地表成 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_s$ 中的向量的和. 即任取 $\alpha \in \mathfrak{S}$, 则

$$\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_s, \quad \alpha_j \in \mathfrak{L}_j, \quad j = 1, 2, \cdots, s, \tag{5.4.5}$$

且若

$$\alpha = \beta_1 + \beta_2 + \cdots + \beta_s, \quad \beta_j \in \mathfrak{L}_j, \quad j = 1, 2, \cdots, s, \tag{5.4.6}$$

则 $\beta_j = \alpha_j, j = 1, 2, \cdots, s$.

$\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_s$ 的直接和 $\mathfrak{S}$ 记作

$$\mathfrak{S} = \mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_s = \bigoplus_{j=1}^{s} \mathfrak{L}_j = {\sum_{j=1}^{s}}' \mathfrak{L}_j. \tag{5.4.7}$$

定理 5.4.17 关于域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的 s 个子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_s$ 的和 $\mathfrak{L}_0 = \sum\limits_{j=1}^{s} \mathfrak{L}_j$, 下面四个性质互相等价 (即, 直接和有下面四种互相等价的定义):

(1) $\dim(\mathfrak{L}_0) = \sum\limits_{j=1}^{s} \dim(\mathfrak{L}_j)$;

(2) 设 $\alpha_{j1}, \alpha_{j2}, \cdots, \alpha_{jn_j}$ 为 $\mathfrak{L}_j$ 的基, $j = 1, 2, \cdots, s$, 则

$$\alpha_{jk}, \quad 1 \leqslant k \leqslant n_j, \quad 1 \leqslant j \leqslant s$$

为 $\mathfrak{L}_0$ 的基;

(3) 若 $\sum\limits_{j=1}^{s} \alpha_j = 0$, 其中 $\alpha_j \in \mathfrak{L}_j, 1 \leqslant j \leqslant s$, 则有 $\alpha_j = 0, 1 \leqslant j \leqslant s$;

(4) 任取 $\alpha \in \mathfrak{L}_0$, 则 α 可唯一分解为 $\alpha = \sum\limits_{j=1}^{s} \alpha_j$, 其中 $\alpha_j \in \mathfrak{L}_j, 1 \leqslant j \leqslant s$. 即若 $\alpha = \sum\limits_{j=1}^{s} \beta_j$, 其中 $\beta_j \in \mathfrak{L}_j, 1 \leqslant j \leqslant s$, 便有 $\beta_j = \alpha_j, 1 \leqslant j \leqslant s$.

证 先证(1) 和(2) 等价.

从(2) 看到, 条件 $\dim(\mathfrak{L}_0) = \sum\limits_{j=1}^{s} \dim(\mathfrak{L}_j)$ 可写为 $\dim(\mathfrak{L}_0) = \sum\limits_{j=1}^{s} n_j$. 所以这个条件等价于向量集

$$\mathfrak{S} = \{\, \alpha_{jk} \mid 1 \leqslant k \leqslant n_j,\ 1 \leqslant j \leqslant s \,\}$$

为子空间 $\mathfrak{L}_0$ 的基. 所以(1) 和(2) 等价.

下面证由(2) 可推出(3), 由(3) 可推出(4), 由(4) 可推出(2). 于是, (2), (3) 和(4) 互相等价.

先由(2) 推(3). 由于(2) 成立, 即向量集 $\mathfrak{S}$ 为 $\mathfrak{L}_0$ 的基. 今由 $\alpha_j \in \mathfrak{L}_j$, $\sum\limits_{j=1}^{s} \alpha_j = 0$, $\alpha_j = \sum\limits_{k=1}^{n_j} c_{jk}\alpha_{jk}$, $1 \leqslant j \leqslant s$. 而 $0 = \sum\limits_{j=1}^{s} \alpha_j = \sum\limits_{j=1}^{s}\sum\limits_{k=1}^{n_j} c_{jk}\alpha_{jk}$. 所以 $c_{jk} = 0$, $1 \leqslant k \leqslant n_j$, $1 \leqslant j \leqslant s$. 这证明了(3) 成立.

再由(3) 推(4). 由于(3) 成立, 即若 $\sum\limits_{j=1}^{s} \gamma_j = 0$, $\gamma_j \in \mathfrak{L}_j$, $1 \leqslant j \leqslant s$, 则 $\gamma_j = 0$, $1 \leqslant j \leqslant s$. 今若 $\alpha_j, \beta_j \in \mathfrak{L}_j$, $1 \leqslant j \leqslant s$, 且 $\alpha = \sum\limits_{j=1}^{s} \alpha_j = \sum\limits_{j=1}^{s} \beta_j$, 则 $\sum\limits_{j=1}^{s} \gamma_j = 0$, 其中 $\gamma_j = \alpha_j - \beta_j \in \mathfrak{L}_j$, $1 \leqslant j \leqslant s$. 因此, $\gamma_j = 0$, $1 \leqslant j \leqslant s$. 此即 $\alpha_j = \beta_j$, $1 \leqslant j \leqslant s$. 这证明了分解的唯一性. 即(4) 成立.

最后由(4) 推(2). 已知任取 $\alpha \in \mathfrak{L}_0$, 则有 $\alpha = \sum\limits_{j=1}^{s} \alpha_j$, 且分解唯一, 其中 $\alpha_j \in \mathfrak{L}_j$, $1 \leqslant j \leqslant s$. 在 $\mathfrak{L}_j$ 中取基 α_{jk}, $1 \leqslant k \leqslant n_j$, 显然 α_{jk}, $1 \leqslant k \leqslant n_j$, $1 \leqslant j \leqslant s$ 线性生成 $\mathfrak{L}_0$. 因此问题化为证明它们是 $\mathfrak{L}_0$ 的一组基, 即证它们线性无关. 今若有数 c_{jk}, $1 \leqslant k \leqslant n_j$, $1 \leqslant j \leqslant s$, 使得 $\sum\limits_{j=1}^{s}\sum\limits_{k=1}^{n_j} c_{jk}\alpha_{jk} = 0$. 由于 $\sum\limits_{k=1}^{n_j} c_{jk}\alpha_{jk} = \alpha_j \in \mathfrak{L}_j$, 于是 $\alpha_1 + \alpha_2 + \cdots + \alpha_t = 0$. 由分解唯一性, 即(4) 成立, 所以有 $\sum\limits_{k=1}^{n_j} c_{jk}\alpha_{jk} = \alpha_j = 0$, $1 \leqslant j \leqslant s$. 又由 α_{jk}, $1 \leqslant k \leqslant n_j$ 为 $\mathfrak{L}_j$ 的一组基, 所以由 $\alpha_j = 0$ 推出 $c_{jk} = 0$, $1 \leqslant k \leqslant n_j$, $1 \leqslant j \leqslant s$. 这证明了 $\mathfrak{S}$ 为 $\mathfrak{L}_0$ 的一组基. □

推论 5.4.18 当 $s = 2$ 时, 下面关于直接和的两种定义互相等价:

$$\dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_2) = \dim(\mathfrak{L}_1 + \mathfrak{L}_2) \quad \text{当且仅当} \quad \mathfrak{L}_1 \cap \mathfrak{L}_2 = 0. \tag{5.4.8}$$

证 这是维数公式 (5.4.4) 的直接推论. □

在结束这一节以前, 给出子空间的一个重要应用. 它给出了矩阵的秩的一个等价定义.

定义 5.4.19 给定域 $\mathbb{F}$ 上 $n \times m$ 矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}. \tag{5.4.9}$$

它的 n 个行可以看作 m 维向量空间 $\mathbb{F}^m$ 中的 n 个向量,

$$\alpha_j = (a_{j1}, a_{j2}, \cdots, a_{jm})', \quad j = 1, 2, \cdots, n, \tag{5.4.10}$$

称为矩阵 A 的**行向量**; 同样, 它的 m 个列可以看作 n 维向量空间 $\mathbb{F}^n$ 中的 m 个向量

$$\beta_k = (a_{1k}, a_{2k}, \cdots, a_{nk})', \quad k = 1, 2, \cdots, m, \tag{5.4.11}$$

称为矩阵 A 的**列向量**, 其中 $(*)'$ 是矩阵 $(*)$ 的转置矩阵.

引理 5.4.20 给定域 $\mathbb{F}$ 上 $n \times m$ 矩阵 A. 矩阵 A 的行向量线性相关当且仅当 $\operatorname{rank}(A) < n$; 线性无关当且仅当 $\operatorname{rank}(A) = n$. 矩阵 A 的列向量线性相关当且仅当 $\operatorname{rank}(A) < m$; 线性无关当且仅当 $\operatorname{rank}(A) = m$.

证 符号和定义 5.4.19 相同. 注意到 $n \times m$ 矩阵 A 的转置矩阵 A' 是 $m \times n$ 矩阵. 而 A 的行向量是它的转置矩阵 A' 的列向量, 又 A 的列向量是它的转置矩阵 A' 的行向量, 所以只需对行向量来证明定理就可以了.

任取域 $\mathbb{F}$ 中 n 个数 $c_1, c_2, \cdots, c_n$, 使得 $c_1\alpha_1 + \cdots + c_n\alpha_n = 0$, 于是它可以改写为

$$\begin{cases} c_1a_{11} + c_2a_{21} + \cdots + c_na_{n1} = 0, \\ \qquad\qquad \vdots \\ c_1a_{1m} + c_2a_{2m} + \cdots + c_na_{nm} = 0. \end{cases}$$

视 $c_1, c_2, \cdots, c_n$ 为未知数, 则这是齐次线性方程组, 系数矩阵即原来的矩阵 A, 未知数个数为 n. 由 §4.2 可知, 它有非零解的必要且充分条件为 $\operatorname{rank}(A) < n$; 没有非零解的必要且充分条件为 $\operatorname{rank}(A) = n$. 这就证明了定理. □

定理 5.4.21 域 $\mathbb{F}$ 上的矩阵 A 的 n 个行向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的极大线性无关向量组的向量个数 s 等于 m 个列向量 $\beta_1, \beta_2, \cdots, \beta_m$ 的极大线性无关向量组的向量个数 t, 也等于矩阵 A 的秩 r.

证 由于域 $\mathbb{F}$ 上的矩阵 A 有 $\operatorname{rank}(A) = \operatorname{rank}(A')$, 而矩阵 A 的列向量为矩阵 A 的转置矩阵 A' 的行向量, 所以只要证明前一断言就够了.

记矩阵 A 的秩为 r. 由秩的定义, 有一个 r 阶子式不为零. 设为

$$A\begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix} \neq 0,$$

其中 $1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant m$. 现在来证明行向量 α_{i_1}, $\alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性无关, 而其他的行向量 α_k 是 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 的线性组合. 事实上, 由 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 的坐标决定的 $r \times m$ 矩阵为

$$\begin{pmatrix} a_{i_11} & a_{i_12} & \cdots & a_{i_1m} \\ a_{i_21} & a_{i_22} & \cdots & a_{i_2m} \\ \vdots & \vdots & & \vdots \\ a_{i_r1} & a_{i_r2} & \cdots & a_{i_rm} \end{pmatrix}.$$

它有一个 r 阶子式 $A\begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix} \neq 0$, 所以它的秩为 r. 由定理 5.4.20, 向量 α_{i_1}, $\alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性无关. 再由 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha_k$ 的坐标决定的 $(r+1)\times m$ 矩阵为

$$\begin{pmatrix} a_{i_1 1} & a_{i_1 2} & \cdots & a_{i_1 m} \\ a_{i_2 1} & a_{i_2 2} & \cdots & a_{i_2 m} \\ \vdots & \vdots & & \vdots \\ a_{i_r 1} & a_{i_r 2} & \cdots & a_{i_r m} \\ a_{k1} & a_{k2} & \cdots & a_{km} \end{pmatrix},$$

其中 $k \neq i_1, i_2, \cdots, i_r$. 因为 $i_1, i_2, \cdots, i_r, k$ 是 $r+1$ 个不同的指标, 所以它是矩阵 A 的子矩阵, 因此它的秩仍为 r. 由定理 5.4.20 可知向量 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha_k$ 线性相关. 由于 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 线性无关, 故 α_k 是 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 的线性组合, 这就证明了 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的极大线性无关向量组. 由定理 5.4.6, $s = r$. 同理可证 $t = r$. □

由定理 5.4.12, 则定理 5.4.21 可以改写为

定理 5.4.22 设矩阵 A 的秩为 r, 则矩阵 A 的行向量线性生成的子空间的维数等于 r; 矩阵 A 的列向量线性生成的子空间的维数也等于 r.

习 题 5.4

5.4.1 试判别 n 维实向量空间 $\mathbb{R}^n$ 中下列子集合是否构成子空间? 如果是, 试确定它的维数, 并找出一组基; 如果不是, 试写出由它线性生成的子空间, 并确定它的维数, 找出一组基:

(1) $\{(a_1, a_2, \cdots, a_r, a_{r+1}, \cdots, a_n) \in \mathbb{R}^n \mid a_1 + a_2 + \cdots + a_r = 0\}$, $1 \leqslant r \leqslant n$;

(2) n 个坐标同时大于零的所有向量构成的子集合;

(3) n 个坐标同时小于零的所有向量构成的子集合;

(4) $\{(a_1, a_2, \cdots, a_n) \mid a_1, a_2, \cdots, a_r$ 遍历所有正实数, $a_{r+1}, a_{r+2}, \cdots, a_n$ 遍历所有负实数$\}$, $1 \leqslant r \leqslant n-1$.

5.4.2 试证: 对任意两组向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 及 $\beta_1, \beta_2, \cdots, \beta_s$, 如果每个向量 α_j 都是向量组 $\beta_1, \beta_2, \cdots, \beta_s$ 的线性组合, 则前一向量组的极大线性无关向量组的向量个数不超过后一向量组的极大线性无关向量组的向量个数.

5.4.3 试证: n 维线性空间 $\mathfrak{L}$ 的子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \mathfrak{L}_3$ 适合

(1) $\mathfrak{L}_1 + \mathfrak{L}_1 + \cdots + \mathfrak{L}_1 = \mathfrak{L}_1$; (2) $\mathfrak{L}_1 \cap \mathfrak{L}_1 \cap \cdots \cap \mathfrak{L}_1 = \mathfrak{L}_1$;

(3) $\mathfrak{L}_1 + \mathfrak{L} = \mathfrak{L} + \mathfrak{L}_1 = \mathfrak{L}$; (4) $\mathfrak{L}_1 \cap \mathfrak{L} = \mathfrak{L} \cap \mathfrak{L}_1 = \mathfrak{L}_1$;

(5) $(\mathfrak{L}_1 \cap \mathfrak{L}_2) + \mathfrak{L}_3 \subset (\mathfrak{L}_1 + \mathfrak{L}_3) \cap (\mathfrak{L}_2 + \mathfrak{L}_3)$, 何时相等?

(6) $(\mathfrak{L}_1 \cap \mathfrak{L}_3) + (\mathfrak{L}_2 \cap \mathfrak{L}_3) \subset (\mathfrak{L}_1 + \mathfrak{L}_2) \cap \mathfrak{L}_3$, 何时相等?

(7) $\mathfrak{L}_1 \cap \mathfrak{L}_2 = \mathfrak{L}_1 + \mathfrak{L}_2$ 当且仅当 $\mathfrak{L}_1 = \mathfrak{L}_2$.

5.4.4 试证: 对域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任两子空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$, 它们的并 $\mathfrak{L}_1 \cup \mathfrak{L}_2$ 等于 $\mathfrak{L}_1 + \mathfrak{L}_2$ 的必要且充分条件为 $\mathfrak{L}_1 \subset \mathfrak{L}_2$ 或 $\mathfrak{L}_2 \subset \mathfrak{L}_1$.

5.4.5 对域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任一子空间 $\mathfrak{L}_1$, 试证: 必定存在子空间 $\mathfrak{L}_2$, 使得 $\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2$. 它唯一吗?

5.4.6 设 $\mathfrak{L}$ 是域 $\mathbb{F}$ 上 n 维线性空间. 试证: 由向量 $\beta_1, \beta_2, \cdots, \beta_m$ 线性生成的子空间 $\mathfrak{L}_1$ 是 n 维线性空间 $\mathfrak{L}$ 中包含向量 β_1, β_2, $\cdots$, β_m 的**最小子空间** (最小的意思是指包含向量 β_1, β_2, $\cdots$, β_m 的任一子空间 $\mathfrak{S}$ 必包含了 $\mathfrak{L}_1$).

5.4.7 设 $\mathfrak{L}_1$, $\mathfrak{L}_2$, $\mathfrak{L}_3$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的三个子空间, 而且 $\mathfrak{L}_1\cap\mathfrak{L}_2=0$, $(\mathfrak{L}_1+\mathfrak{L}_2)\cap\mathfrak{L}_3=0$. 设对 $\mathfrak{L}$ 中任一向量 α, 存在向量 $\alpha_j\in\mathfrak{L}_j$, $j=1,2,3$, 使得 $\alpha=\alpha_1+\alpha_2+\alpha_3$. 试证: $\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\mathfrak{L}_3$.

5.4.8 设 A 和 B 为域 $\mathbb{F}$ 上的 $n\times m$ 矩阵. 记 $\mathfrak{L}(A)$ 为矩阵 A 的 n 个列向量线性生成的子空间. 试证: $\mathfrak{L}(A)=\mathfrak{L}(AA^*)$. 再证

$$\dim(\mathfrak{L}(A+B))\leqslant\dim\left(\mathfrak{L}\begin{pmatrix}A\\B\end{pmatrix}\right)\leqslant\dim(\mathfrak{L}(A))+\dim(\mathfrak{L}(B)).$$

5.4.9 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathfrak{L}_1$, $\mathfrak{L}_2$ 为线性空间 $\mathfrak{L}$ 的两个非零真子空间. 试证: 存在 $\alpha\in\mathfrak{L}$, 又 $\alpha\notin\mathfrak{L}_1\cup\mathfrak{L}_2$.

5.4.10 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathfrak{L}_1$, $\mathfrak{L}_2$, $\cdots$, $\mathfrak{L}_s$ 为线性空间 $\mathfrak{L}$ 的非零真子空间. 试证:

(1) 存在 $\alpha\in\mathfrak{L}$, 又 $\alpha\notin\mathfrak{L}_j$, $1\leqslant j\leqslant s$;

(2) 设 $1\leqslant s\leqslant n$. 在 $\mathfrak{L}$ 中存在基 α_1, α_2, $\cdots$, α_n, 使得 $\alpha_j\notin\mathfrak{L}_j$, $1\leqslant j\leqslant s$.

5.4.11 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathfrak{L}_0$, $\mathfrak{L}_1$, $\mathfrak{L}_2$, $\cdots$, $\mathfrak{L}_s$ 为线性空间 $\mathfrak{L}$ 的非零子空间, 而且有 $\mathfrak{L}_0\subset\mathfrak{L}_1\cup\mathfrak{L}_2\cup\cdots\cup\mathfrak{L}_s$. 试证: 存在指标 $j\in\{1,2,\cdots,s\}$, 使得 $\mathfrak{L}_0\subset\mathfrak{L}_j$.

5.4.12 设 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_s$ 是 n 维线性空间 $\mathfrak{L}$ 的 s 个非零子空间, 试证下面三个条件等价:

(1) $\mathfrak{L}_0=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\cdots\oplus\mathfrak{L}_s$, 即 $\mathfrak{L}_1$, $\mathfrak{L}_2$, $\cdots$, $\mathfrak{L}_s$ 的和是直接和;

(2) $\mathfrak{L}_j\cap(\mathfrak{L}_1+\cdots+\mathfrak{L}_{j-1}+\mathfrak{L}_{j+1}+\cdots+\mathfrak{L}_s)=0$, $j=1,2,\cdots,s$;

(3) $\mathfrak{L}_1\cap\mathfrak{L}_2=0$, $(\mathfrak{L}_1+\mathfrak{L}_2)\cap\mathfrak{L}_3=0$, $\cdots$, $(\mathfrak{L}_1+\cdots+\mathfrak{L}_{s-1})\cap\mathfrak{L}_s=0$.

§5.5 线性方程组求解的几何理论

在这一节, 将利用向量的语言, 来讨论线性方程组的求解理论.

首先把 m 个未知数 x_1, x_2, $\cdots$, x_m, n 个方程构成的非齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=b_1,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=b_2,\\\qquad\vdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m=b_n\end{cases}\tag{5.5.1}$$

写成矩阵的形式

$$Ax=\beta,\tag{5.5.2}$$

其中 A 为系数矩阵, $\widetilde{A}=(A,\beta)$ 为增广矩阵, $x=(x_1,x_2,\cdots,x_m)'$. 再将矩阵形式 (5.5.2) 写成向量的形式. 即用 α_1, α_2, $\cdots$, α_m 表示系数矩阵 A 的 m 个列向量, 于是非齐次线性方程组 (5.5.2) 可以写成向量方程

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_m\alpha_m=\beta, \tag{5.5.3}$$

因此非齐次线性方程组 (5.5.1) 有解的必要且充分条件是向量 β 是向量 α_1, α_2, $\cdots$, α_m 的线性组合.

现在用向量语言给出定理 4.1.4 的另一个证明.

定理 5.5.1 (线性方程组的相容性定理) 非齐次线性方程组 (5.5.1) 有解的必要且充分条件是它的系数矩阵 A 和增广矩阵 $\widetilde{A}$ 的秩相等.

证 先证必要性. 设非齐次线性方程组有解. 从式 (5.5.3) 可以看出向量 β 是向量 α_1, α_2, $\cdots$, α_m 的线性组合, 于是向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 的极大线性无关向量组也是向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta\}$ 的极大线性无关向量组. 因为向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m$ 由系数矩阵 A 的列向量构成, 向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta\}$ 由增广矩阵 $\widetilde{A}$ 的列向量构成. 由定理 5.4.21, 所以系数矩阵 A 的秩等于增广矩阵 $\widetilde{A}$ 的秩.

再证充分性. 设系数矩阵 A 的秩等于增广矩阵 $\widetilde{A}$ 的秩. 由定理 5.4.22 可知, 系数矩阵 A 和增广矩阵 $\widetilde{A}$ 的列向量 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 和 $\{\alpha_1, \alpha_2, \cdots, \alpha_m, \beta\}$ 分别线性生成的两个子空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的维数相同. 今 $\mathfrak{L}_1$ 包含在 $\mathfrak{L}_2$ 中, 所以 $\mathfrak{L}_1$ 的基也是 $\mathfrak{L}_2$ 的基, 即 $\mathfrak{L}_1=\mathfrak{L}_2$. 特别, 由 $\beta\in\mathfrak{L}_2$ 可知 $\beta\in\mathfrak{L}_1$, 即 β 是 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 的线性组合. □

现在来叙述一个原则上和 §4.1 差不多的求解方法. 利用方程和未知数的重新编号, 无妨假设, 当增广矩阵 $\widetilde{A}$ 的秩等于系数矩阵 A 的秩 r 时, 它们的前 r 行和前 r 列决定的子式 $\det A(\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix})\neq 0$. 由定理 5.4.21, 增广矩阵 $\widetilde{A}$ 的第 $r+1$, $r+2,\cdots,n$ 行都是前 r 行的线性组合. 所以, 线性方程组 (5.5.1) 和它的前 r 个方程决定的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=b_2,\\ \qquad\vdots\\ a_{r1}x_1+a_{r2}x_2+\cdots+a_{rm}x_m=b_r \end{cases} \tag{5.5.4}$$

等价. 将方程组 (5.5.4) 改写为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r = & b_1 - a_{1,r+1}x_{r+1} - a_{1,r+2}x_{r+2} - \cdots - a_{1m}x_m, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2r}x_r = & b_2 - a_{2,r+1}x_{r+1} - a_{2,r+2}x_{r+2} - \cdots - a_{2m}x_m, \\ & \vdots \\ a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rr}x_r = & b_r - a_{r,r+1}x_{r+1} - a_{r,r+2}x_{r+2} - \cdots - a_{rm}x_m. \end{cases} \tag{5.5.5}$$

利用 Cramer 法则, 在域 $\mathbb{F}$ 中任意给定一组数 $x_{r+1}^{(0)}$, $x_{r+2}^{(0)}$, $\cdots$, $x_m^{(0)}$ 后, 便得到域 $\mathbb{F}$ 中一组唯一的解 $(x_1^{(0)}, x_2^{(0)}, \cdots, x_r^{(0)})'$. 所以得到线性方程组 (5.5.1) 的一组解 $(x_1^{(0)}, \cdots, x_m^{(0)})'$. 由此可见, 在有解时, 解依赖于 $m-r$ 个独立参数.

关于齐次线性方程组有

定理 5.5.2 (齐次线性方程组解的结构定理) 给定 m 个独立未知数, n 个方程构成的齐次线性方程组

$$Ax = 0 \tag{5.5.6}$$

一定有解. 设系数矩阵的秩为 r, 则解作为 m 维向量空间 $\mathbb{F}^m$ 中的向量, 它们全体构成 $\mathbb{F}^m$ 的 $m-r$ 维线性子空间, 称为齐次线性方程组 (5.5.6) 的**解空间**, 其中的基称为齐次线性方程组 (5.5.6) 的**基础解系**.

证 记 $n \times m$ 矩阵 A 的 m 个列向量为 α_1, α_2, $\cdots$, α_m, 则齐次线性方程组 $Ax = 0$ 可改写为

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_m\alpha_m = 0. \tag{5.5.7}$$

在 m 维向量空间 $\mathbb{F}^m$ 的向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 中任取极大线性无关向量组. 为讨论方便起见, 无妨设 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ 为极大线性无关向量组. 由定理 5.4.21, $r = \operatorname{rank}(A)$. 由极大线性无关向量组的定义可知

$$\alpha_k = \sum_{j=1}^{r} b_{jk}\alpha_j, \quad k = r+1, r+2, \cdots, m. \tag{5.5.8}$$

代回到式 (5.5.7), 有

$$0 = \sum_{j=1}^{m} x_j\alpha_j = \sum_{j=1}^{r} x_j\alpha_j + \sum_{k=r+1}^{m} x_k\alpha_k = \sum_{j=1}^{r}\Big(x_j + \sum_{k=r+1}^{m} b_{jk}x_k\Big)\alpha_j.$$

由于向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关, 所以有

$$x_j = -\sum_{k=r+1}^{m} b_{jk}x_k, \quad 1 \leqslant j \leqslant r.$$

记 $e_j=(\delta_{1j},\delta_{2j},\cdots,\delta_{mj})\in\mathbb{F}^m$, 其中 δ_{jk} 是 Kronecker 符号, 即 e_j 的第 j 个坐标为 1, 其余坐标为零, $1\leqslant j\leqslant m$. 于是 $m\times 1$ 矩阵 x 为

$$\begin{aligned}x&=\sum_{j=1}^{r}x_je_j+\sum_{k=r+1}^{m}x_ke_k=-\sum_{k=r+1}^{m}\sum_{j=1}^{r}b_{jk}x_ke_j+\sum_{k=r+1}^{m}x_ke_k\\&=\sum_{k=r+1}^{m}x_k\big(e_k-\sum_{j=1}^{r}b_{jk}e_j\big).\end{aligned}$$

记 $\beta_k=e_k-\sum\limits_{j=1}^{r}b_{jk}e_j$, $r+1\leqslant k\leqslant m$, 即有 $x=\sum\limits_{k=r+1}^{m}x_k\beta_k$. 由于 x_{r+1}, x_{r+2}, $\cdots$, x_m 可任取, 这证明了齐次线性方程组 (5.5.6) 的通解构成由 β_{r+1}, β_{r+2}, $\cdots$, β_m 线性生成的子空间. 于是 β_{r+1}, β_{r+2}, $\cdots$, β_m 是基础解系.

余下证 β_{r+1}, β_{r+2}, $\cdots$, β_m 线性无关. 事实上, 若存在常数 b_{r+1}, b_{r+2}, $\cdots$, b_m, 使得 $\sum\limits_{k=r+1}^{m}b_k\beta_k=0$, 此即

$$0=\sum_{k=r+1}^{m}b_k\beta_k=\sum_{k=r+1}^{m}b_k\Big(e_k-\sum_{j=1}^{r}b_{jk}e_j\Big)=\sum_{j=1}^{r}\Big(\sum_{k=r+1}^{m}(-b_{jk}b_k)\Big)e_j+\sum_{j=r+1}^{m}b_je_j.$$

由于 $e_1,e_2,\cdots,e_m$ 为 m 维向量空间 $\mathbb{F}^m$ 的基, 这证明了 $b_{r+1}=b_{r+2}=\cdots=b_m=0$, 所以 β_{r+1}, β_{r+2}, $\cdots$, β_m 线性无关. □

习 题 5.5

5.5.1 试用本节的方法求下列线性方程组的通解:

(1) $\begin{cases}2x_1+7x_2+3x_3=0,\\3x_1+4x_2+x_3=0;\end{cases}$ (2) $\begin{cases}2x_1+7x_2+3x_3=1,\\3x_1+4x_2+x_3=0.\end{cases}$

5.5.2 设 A, B, C 分别为 $m\times n$, $n\times p$, $p\times q$ 矩阵. 试用本节理论证明: 若 $\operatorname{rank}(B)=\operatorname{rank}(AB)$, 则有 $\operatorname{rank}(BC)=\operatorname{rank}(ABC)$.

5.5.3 设 A 和 B 分别为 $s\times n$ 及 $t\times n$ 矩阵. 记 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 分别为 n 维向量空间 $\mathbb{F}^n$ 中的子空间, 且 $\mathfrak{L}_1$ 为齐次线性方程组 $Ax=0$ 的解空间, $\mathfrak{L}_2$ 为齐次线性方程组 $Bx=0$ 的解空间. 试构造一个齐次线性方程组, 使得它的解空间为 n 维向量空间 $\mathbb{F}^n$ 中的子空间 $\mathfrak{L}_1\cap\mathfrak{L}_2$; 另外构造一个齐次线性方程组, 使得它的解空间为 n 维向量空间 $\mathbb{F}^n$ 中的子空间 $\mathfrak{L}_1+\mathfrak{L}_2$.

5.5.4 设 A,B,C 分别为 $m\times n$, $n\times p$, $p\times q$ 矩阵, 且 $\operatorname{rank}(A)+\operatorname{rank}(B)=n$, $AB=0$, $AC=0$. 试用本节理论来证明: 存在 $p\times q$ 矩阵 D, 使得 $C=BD$, 且唯一存在 $p\times q$ 矩阵 D 的必要且充分条件为 $\operatorname{rank}(B)=p$.

5.5.5 设 t 个 n 维向量 $\alpha_j=(a_{j1},\ a_{j2},\ \cdots,\ a_{jn})$, $j=1,\ 2,\ \cdots,\ t$ 适合条件 $2|a_{jj}|>\sum\limits_{k=1}^{2}|a_{jk}|$, $j=1,2,\cdots,t$. 试证: α_1, α_2, $\cdots$, α_t 线性无关.

5.5.6 设 A, B, C, D 为域 $\mathbb{F}$ 上 n 阶方阵, 适合 $A=BC$, $B=AD$. 试证: 存在 n 阶非异方阵 P, 使得 $B=AP$.

第六章　线性变换

§6.1　线性变换

设想平面是一张没有厚度和边的纸，它放在一张没有有边的桌子上，把桌上的点和纸上的点相重在一起. 在桌面上任取一点 P, 在纸上和它相重的点记作 $\widetilde{P}$. 现在将纸在桌面上按照某种规律变动一下. 在新的位置上, 纸上的点 $\widetilde{P}$ 不一定再和桌面上的点 P 相重了. 这时, 记和点 $\widetilde{P}$ 相重的点为 P'. 于是给出了桌面上每一点 P 到另一点 P' 的对应. 由于对应是通过可滑动的纸上的点 $\widetilde{P}$ 建立的, 所以这是桌面上点到桌面上另一点的一个一一对应, 即是一种运动, 称为变换.

再设想直线是一根可伸缩的橡皮筋, 在直线两端用力拉, 得到同一条直线. 但是直线上的点由于直线的伸长而变动了位置, 这样也就给出了直线上每一点移动到另点的变化规律. 显然这是一个直线到自身上的一一对应, 它也是一种运动, 也称为变换.

再如点关于轴的反射和面的反射, 它们分别是空间中点到空间自身内某一定直线或某一定平面上的单值对应, 但它们都是一一对应. 又如点关于轴的投影和面的投影, 它们分别是空间中点到空间自身内某一定直线或某一定平面上的单值对应, 但它们都不是一一对应.

将这些例子抽象化, 用集合的语言, 可以对上面这些现象下一个极其广泛的定义.

定义 6.1.1　集合 $\mathfrak{S}$ 到集合 $\mathfrak{S}_1$ 内的单值对应 σ 称为**映射**. 在集合 $\mathfrak{S}$ 中任给元素 a, 假设在单值对应 σ 下, 对应元素为 b, 则称 a 为 b 的**原象**, 称 b 为 a 的**象**, 用 $b=\sigma(a)$ 记之. 集合 $\mathfrak{S}$ 称为**原象集**, 集合 $\sigma(\mathfrak{S})(\subset \mathfrak{S}_1)$ 称为**象集**. 又集合 $\mathfrak{S}$ 到自身内的单值对应称为**变换**.

集合 $\mathfrak{S}$ 上两个映射 σ 和 τ 称为**相等的**, 如果 σ 和 τ 是 $\mathfrak{S}$ 上同一个单值对应. 即在集合 $\mathfrak{S}$ 中任取元素 a, 则有 $\sigma(a)=\tau(a)$. 这时记作 $\tau=\sigma$.

由两个映射的相等定义可知, 相等概念有性质: (1) 反身性: $\sigma=\sigma$; (2) 对称性: 由 $\sigma=\tau$ 可推知 $\sigma=\tau$; (3) 传递性: 由 $\sigma=\tau, \tau=\xi$ 可推知 $\sigma=\xi$. 所以映射的相等概念是等价关系. 今后, 凡相等的映射看作同一个映射.

除了前面所举的变换的例子外, 下面再举一些例. 给定正整数 $n>1$, 考虑由正整数 $1, 2, \cdots$ 构成的集合 $\mathbb{N}_+$, 在 $\mathbb{N}_+$ 中给出下列单值对应:

(1) $1\to 1, 2\to 2, \cdots, n\to n, \cdots$;

(2) $1\to i_1, 2\to i_2, \cdots, n\to i_n$, 其中 $i_1i_2\cdots i_n$ 是 $1, 2, \cdots, n$ 的任一排列;

(3) $1\to 2, 2\to 4, \cdots, n\to 2n, \cdots$;

(4) $1\to 1, 2\to 1, 3\to 1, \cdots, n\to 1$;

(5) 奇数 $\to 1$, 偶数 $\to 2$.

由定义可知(1), (2) 是 "到上" 的一一对应; (3) 是 "到内" 的一一对应. 这个情形很重要, 它表明: 一一对应不一定需要是集合 $\mathfrak{S}$ 到自身上的, 可以是集合 $\mathfrak{S}$ 到它的子集合 $\mathfrak{S}_1$ 上的. 当然, 在集合 $\mathfrak{S}$ 仅由有限个元素构成时, 没有这个性质. (4), (5) 是多一对应.

下面自然地引进一种代数运算, 即映射的乘积.

定义 6.1.2 给定集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的映射 σ: $a\to b=\sigma(a), \forall a\in\mathfrak{S}$ 和集合 $\mathfrak{S}_1$ 到 $\mathfrak{S}_2$ 内的映射 τ: $b\to c=\tau(b), \forall b\in\mathfrak{S}_1$. 映射 σ 和 τ 的依次作用 $\tau(\sigma(a))$: $a\to c=\tau(\sigma(a))$ 是一个新的映射. 这个新的映射将集合 $\mathfrak{S}$ 映到 $\mathfrak{S}_2$ 内, 称为映射 σ 和 τ 的**乘积**. 记作 $\tau\circ\sigma$. 即有

$$\tau(\sigma(a))=\tau\circ\sigma(a), \quad \forall a\in\mathfrak{S}. \tag{6.1.1}$$

由定义可知, 映射 σ 和 τ 的乘积是由依次作用变换 σ 及 τ 得到的. 映射的乘积, 有一个重要性质:

定理 6.1.3 给定集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的映射 ξ, 集合 $\mathfrak{S}_1$ 到 $\mathfrak{S}_2$ 内的映射 η, 集合 $\mathfrak{S}_2$ 到 $\mathfrak{S}_3$ 内的映射 ζ, 则映射的**乘法结合律**成立. 即有

$$(\zeta\circ\eta)\circ\xi=\zeta\circ(\eta\circ\xi). \tag{6.1.2}$$

证 在集合 $\mathfrak{S}$ 中任取一元 a. 由映射乘积的定义可知

$$((\zeta\circ\eta)\circ\xi)(a)=(\zeta\circ\eta)(\xi(a))=\zeta(\eta(\xi(a))),\ (\zeta\circ(\eta\circ\xi))(a)=\zeta((\eta\circ\xi)(a))=\zeta(\eta(\xi(a))).$$

所以式 (6.1.2) 成立. □

在变换理论中, 使集合 $\mathfrak{S}$ 中每一元素都不动的变换有特殊的作用, 它称为**恒等变换**. 恒等变换总用符号 id 来表示, 确切地可表为 $\mathrm{id}_{\mathfrak{S}}$. 由乘法定义可知, 对集合 $\mathfrak{S}$

到 $\mathfrak{S}_1$ 内的映射 σ, 则 $\sigma\circ \mathrm{id}_{\mathfrak{S}}=\mathrm{id}_{\mathfrak{S}_1}\circ\sigma=\sigma$.

定义 6.1.4 集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的映射 σ 称为**可逆映射**, 如果它是集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的一一对应.

定理 6.1.5 设 σ 为集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的映射. 映射 σ 为可逆映射当且仅当唯一存在象集 $\sigma(\mathfrak{S}_1)$ 到原象集 $\mathfrak{S}$ 上的映射 τ, 使得

$$\sigma\circ\tau=\mathrm{id}_{\mathfrak{S}_1},\quad \tau\circ\sigma=\mathrm{id}_{\mathfrak{S}}.$$

这时**逆映射** τ 记作 σ^{-1}, 所以有

$$\sigma\circ\sigma^{-1}=\mathrm{id}_{\mathfrak{S}_1},\quad \sigma^{-1}\circ\sigma=\mathrm{id}_{\mathfrak{S}}. \tag{6.1.3}$$

当 $\mathfrak{S}_1=\mathfrak{S}$ 时, 变换 σ 称为**可逆变换**, 这时变换 σ 的逆映射 σ^{-1} 称为变换 σ **的逆变换**.

证 设映射 σ 为集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 内的一一映射, 则映射 σ 为集合 $\mathfrak{S}$ 到象集 $\sigma(\mathfrak{S})$ 上的一一映射. 因此, 无妨设 σ 为集合 $\mathfrak{S}$ 到 $\mathfrak{S}_1$ 上的一一映射.

将象作为原象, 原象作为象; 象集作为原象集, 原象集作为象集. 定义了象集 $\mathfrak{S}_1$ 到原象集 $\mathfrak{S}$ 上的一个新的一一对应关系:

$$\tau:\quad b\to\tau(b)=a,\quad \forall\, b\in\mathfrak{S}_1. \tag{6.1.4}$$

它们有关系 $\sigma(\tau(b))=\sigma(a)=b$, $\forall\, b\in\mathfrak{S}_1$, $\tau(\sigma(a))=\tau(b)=a$, $\forall\, a\in\mathfrak{S}$. 因此可写为式 (6.1.3). 这时, 映射 τ 称为映射 σ 的逆映射, 记作 $\tau=\sigma^{-1}$.

反之, 下面来证 σ 为一一映射. 事实上, 任取 a, $b\in\mathfrak{S}$, 使得 $\sigma(a)=\sigma(b)$. 由于存在象集 $\sigma(\mathfrak{S}_1)$ 到原象集 $\mathfrak{S}$ 上的映射 τ, 使得 $\sigma\circ\tau=\mathrm{id}_{\mathfrak{S}_1}$, $\tau\circ\sigma=\mathrm{id}_{\mathfrak{S}}$. 于是 $\tau(\sigma(a))=\tau(\sigma(b))$, 即 $(\tau\circ\sigma)(a)=(\tau\circ\sigma)(b)$. 所以 $a=b$. 这证明了 σ 为一一映射.

下面证逆映射唯一. 今若有 $\mathfrak{S}_1$ 到 $\mathfrak{S}$ 上的映射 τ 和 τ_1, 它们适合 $\sigma\circ\tau=\sigma\circ\tau_1=\mathrm{id}_{\mathfrak{S}_1}$, $\tau\circ\sigma=\tau_1\circ\sigma=\mathrm{id}_{\mathfrak{S}}$. 由定理 6.1.3, 于是

$$\tau_1=(\mathrm{id}_{\mathfrak{S}})\circ\tau_1=(\tau\circ\sigma)\circ\tau_1=\tau\circ(\sigma\circ\tau_1)=\tau\circ\mathrm{id}_{\mathfrak{S}_1}=\tau.$$

这证明了 σ 的逆映射唯一存在. □

在有限维线性空间理论中, 最重要的映射是线性映射, 它定义为:

定义 6.1.6 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的映射. σ 称为**线性映射**, 也称为**线性算子**, 或称为**线性同态**, 如果它具有条件

(1) $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$, $\forall\,\alpha,\beta\in\mathfrak{L}$;

(2) $\sigma(c\alpha)=c\sigma(\alpha)$, $\forall\,\alpha\in\mathfrak{L}$, $c\in\mathbb{F}$.

设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射, 由定义可知下列性质成立.

(I) 线性映射将零向量映为零向量; 又 $\sigma(-\alpha) = -\sigma(\alpha), \forall\, \alpha \in \mathfrak{L}$;

证 今 $0+0=0$, 所以 $\sigma(0) = \sigma(0+0) = \sigma(0)+\sigma(0)$. 因此, $\sigma(0) = 0$. 又 $\sigma(-\alpha) = \sigma((-1)\alpha) = (-1)\sigma(\alpha) = -\sigma(\alpha)$. □

(II) 对 n 维线性空间 $\mathfrak{L}$ 中任意 p 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 及域 $\mathbb{F}$ 中 p 个数 $c_1, c_2, \cdots, c_p$, 有

$$\sigma(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_p\alpha_p) = c_1\sigma(\alpha_1) + c_2\sigma(\alpha_2) + \cdots + c_p\sigma(\alpha_p). \tag{6.1.5}$$

因此, 如果 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 是 n 维线性空间 $\mathfrak{L}$ 的一组基, 则象空间 $\sigma(\mathfrak{L})$ 由向量集 $\beta_1 = \sigma(\gamma_1), \beta_2 = \sigma(\gamma_2), \cdots, \beta_n = \sigma(\gamma_n)$ 线性生成.

证 由线性映射的定义, 作归纳法, 便可证明性质(II). □

(III) 如果 $\mathfrak{L}$ 中 p 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性相关, 则 $\mathfrak{L}_1$ 中 p 个向量 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_p)$ 也线性相关, 因此若 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_p)$ 线性无关, 则 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 也线性无关.

证 由线性相关和线性无关的定义和式 (6.1.5), 便证明了性质(III) 成立. □

注意: $\alpha_1, \alpha_2, \cdots, \alpha_p$ 线性无关, 推不出 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_p)$ 也线性无关; 又 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_p)$ 线性相关推不出 $\alpha_1, \alpha_2, \cdots, \alpha_p$ 也线性相关.

引理 6.1.7 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射. 则线性映射 σ 的**象空间** $\sigma(\mathfrak{L})$ 是线性空间 $\mathfrak{L}_1$ 的子空间; **核空间**

$$\ker(\sigma) = \{\, \alpha \in \mathfrak{L} \mid \sigma(\alpha) = 0 \,\} \tag{6.1.6}$$

是线性空间 $\mathfrak{L}$ 的子空间. 又每个**象元素** $\beta \in \mathfrak{L}_0$ 的**完全原象**

$$\sigma^{-1}(\beta) = \{\, \alpha \in \mathfrak{L} \mid \sigma(\alpha) = \beta \,\} \tag{6.1.7}$$

非空, 在其中任意取定向量 α_0, 则 $\sigma(\alpha_0) = \beta$, 而且

$$\sigma^{-1}(\beta) = \alpha_0 + \ker(\sigma) = \{\, \alpha_0 + \alpha \in \mathfrak{L} \mid \sigma(\alpha) = 0 \,\}. \tag{6.1.8}$$

证 由定义立即可证象空间 $\sigma(\mathfrak{L})$ 是线性空间 $\mathfrak{L}_1$ 的子空间; 核空间 $\ker(\sigma)$ 是线性空间 $\mathfrak{L}$ 的子空间.

下面来证式 (6.1.8) 成立. 记 $\mathfrak{S} = \{\, \alpha_0 + \alpha \in \mathfrak{L} \mid \sigma(\alpha) = 0 \,\}$. 为了证明 $\sigma^{-1}(\beta) = \mathfrak{S}$, 只要证明 $\mathfrak{S}$ 和 $\sigma^{-1}(\beta)$ 互相包含就可以了. 已知 $\sigma(\alpha_0) = \beta$. 任取 $\alpha \in \ker(\sigma)$, 则 $\sigma(\alpha_0 + \alpha) = \sigma(\alpha_0) + \sigma(\alpha) = \sigma(\alpha_0) = \beta$. 所以 $\alpha_0 + \alpha \in \sigma^{-1}(\beta)$, 即 $\mathfrak{S} \subset \sigma^{-1}(\beta)$. 反之, 任取 $\xi \in \sigma^{-1}(\beta)$, 则 $\sigma(\xi) = \beta$. 已知 $\sigma(\alpha_0) = \beta$, 所以

$\sigma(\xi)=\beta=\sigma(\alpha_0)$. 因此, $\sigma(\xi-\alpha_0)=0$. 这证明了 $\xi-\alpha_0\in\ker(\sigma)$. 于是 $\xi\in\mathfrak{S}$, 即 $\sigma^{-1}(\beta)\subset\mathfrak{S}$. 这证明了式 (6.1.8) 成立. □

定义 6.1.8 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射称为**零映射**, 如果它将 $\mathfrak{L}$ 中每一个向量都映为 $\mathfrak{L}_1$ 中的零向量. (今后, 零映射用通常的数 0 的符号来表示, 这样并不会产生混淆的.)

域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 到自身内的零映射称为**零变换**.

下面开始讨论线性映射的矩阵表示.

定理 6.1.9 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$. 在线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 在线性空间 $\mathfrak{L}_1$ 中任取 n 个向量 ξ_1, $\xi_2,\cdots,\xi_n$, 则存在且只存在一个线性映射 σ, 它将 α_j 映为 ξ_j, $1\leqslant j\leqslant n$. 因此, 所有线性映射和所有 $\mathfrak{L}_1$ 中 n 元有序向量集之间有一个自然的一一对应.

证 在线性空间 $\mathfrak{L}$ 中任取向量 $\alpha=\sum\limits_{j=1}^n x_j\alpha_j$, 定义单值对应 σ:

$$\alpha=\sum_{j=1}^n x_j\alpha_j\longrightarrow\sum_{j=1}^n x_j\xi_j=\xi=\sigma(\alpha),\tag{6.1.9}$$

于是有 $\sigma(\alpha_j)=\xi_j$, $1\leqslant j\leqslant n$, 而且映射 σ 是线性映射. 事实上, 设 $\sigma(\alpha)=\xi$, $\sigma(\beta)=\eta$, 其中 $\alpha=\sum\limits_{j=1}^n x_j\alpha_j$, $\widetilde{\alpha}=\sum\limits_{j=1}^n y_j\alpha_j$. 对域 $\mathbb{F}$ 中任意数 c 及 d, 则有

$$\begin{aligned}\sigma(c\alpha+d\beta)&=\sigma\Big(c\sum_{j=1}^n x_j\alpha_j+d\sum_{j=1}^n y_j\alpha_j\Big)=\sigma\Big(\sum_{j=1}^n(cx_j+dy_j)\alpha_j\Big)\\&=\sum_{j=1}^n(cx_j+dy_j)\xi_j=c\sum_{j=1}^n x_j\xi_j+d\sum_{j=1}^n y_j\xi_j=c\sigma(\alpha)+d\sigma(\beta).\end{aligned}$$

所以 σ 是线性映射. 再证唯一性. 若另有一个线性映射 τ, 它有 $\tau(\alpha_j)=\xi_j$, $j=1$, $2,\cdots,n$, 则对线性空间 $\mathfrak{L}$ 中任一向量 $\alpha=\sum\limits_{j=1}^n x_j\alpha_j$, 有

$$\tau(\alpha)=\tau\Big(\sum_{j=1}^n x_j\alpha_j\Big)=\sum_{j=1}^n x_j\tau(\alpha_j)=\sum_{j=1}^n x_j\xi_j=\sigma(\alpha),\quad\forall\,\alpha\in\mathfrak{L}.$$

由定义可知 $\tau=\sigma$, 所以唯一性成立. □

这个定理告诉我们, 线性映射和域 $\mathbb{F}$ 上 $m\times n$ 矩阵之间, 由式 (6.1.9) 建立了一个一一对应, 即我们有:

定理 6.1.10 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射, 在线性空间 $\mathfrak{L}$ 中取定基**(I)**: $\alpha_1,\alpha_2,\cdots,\alpha_n$; 在线性空间 $\mathfrak{L}_1$ 中取定基**(II)**:

$\beta_1, \beta_2, \cdots, \beta_m$, 于是有

$$\xi_i = \sigma(\alpha_i) = \sum_{j=1}^{m} a_{ji}\beta_j, \quad 1 \leqslant i \leqslant n. \tag{6.1.10}$$

因此在取定 $\mathfrak{L}$ 的基(I) 和 $\mathfrak{L}_1$ 的基(II) 后, 线性映射 σ 和 $m \times n$ 矩阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{6.1.11}$$

间有一个自然的到上的一一对应关系 $\sigma \to A$. 所以 $m \times n$ 矩阵 A 称为线性映射 σ 在 $\mathfrak{L}$ 的基(I) 和 $\mathfrak{L}_1$ 的基(II) 下基的**矩阵表示**, 其中 $m \times n$ 矩阵 A 的第 j 个列向量 $(a_{1j}, a_{2j}, \cdots, a_{mj})'$ 为向量 $\sigma(\alpha_j) = \xi_j$ 的坐标.

证 任取域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射 σ, 则 $\xi_1 = \sigma(\alpha_1)$, $\xi_2 = \sigma(\alpha_2)$, $\cdots$, $\xi_n = \sigma(\alpha_n)$ 为线性空间 $\mathfrak{L}_1$ 中的 n 个向量. 由定理 6.1.9, 在线性空间 $\mathfrak{L}$ 中取定一组基(I) 后, 线性映射 σ 和线性空间 $\mathfrak{L}_1$ 中向量组 $\{\xi_1, \xi_2, \cdots, \xi_n\}$ 之间有一个自然的一一对应: 在 m 维线性空间 $\mathfrak{L}_1$ 中取定一组基(II), 则有

$$\sigma(\alpha_k) = \xi_k = \sum_{j=1}^{m} a_{jk}\beta_j, \quad 1 \leqslant k \leqslant n. \tag{6.1.12}$$

这证明了 σ 及式 (6.1.10) 和 (6.1.11) 定义的 $m \times n$ 矩阵 A 之间有一个自然的一一对应, 其中 $m \times n$ 矩阵 A 的第 j 个列向量 $(a_{1j}, a_{2j}, \cdots, a_{mj})'$ 为向量 $\sigma(\alpha_j) = \xi_j$ 的坐标. □

这个定理也给出了 $m \times n$ 矩阵的一种几何解释. 而且可知线性映射的矩阵表示和基底的选取有关. 很自然地, 当选取不同基时, 线性映射的矩阵表示如何改变, 便是一个重要问题.

定理 6.1.11 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射, 在线性空间 $\mathfrak{L}$ 中取定基(I): $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和基$(\widetilde{I})$: $\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_n$; 在线性空间 $\mathfrak{L}_1$ 中取定基(II): $\beta_1, \beta_2, \cdots, \beta_m$ 和基$(\widetilde{II})$: $\widetilde{\beta}_1, \widetilde{\beta}_2, \cdots, \widetilde{\beta}_m$. 于是有基变换公式:

$$\widetilde{\alpha}_i = \sum_{j=1}^{n} p_{ji}\alpha_j, \quad 1 \leqslant i \leqslant n, \qquad \widetilde{\beta}_j = \sum_{k=1}^{m} q_{kj}\beta_k, \quad 1 \leqslant j \leqslant m. \tag{6.1.13}$$

它们决定的非异方阵表示分别为:

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}, \quad Q = \begin{pmatrix} q_{11} & \cdots & q_{1m} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mm} \end{pmatrix}. \tag{6.1.14}$$

线性空间 $\mathfrak{L}$ 到线性空间 $\mathfrak{L}_1$ 内的线性映射 σ 可表示为:

$$\sigma(\alpha_j)=\sum_{k=1}^{m}a_{kj}\beta_k,\quad 1\leqslant j\leqslant n,\qquad \sigma(\alpha_i')=\sum_{j=1}^{m}b_{ji}\beta_j',\quad 1\leqslant i\leqslant n,\tag{6.1.15}$$

其中

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix},\quad B=\begin{pmatrix}b_{11}&\cdots&q_{1n}\\ \vdots&&\vdots\\ q_{m1}&\cdots&q_{mn}\end{pmatrix}\tag{6.1.16}$$

分别为线性映射 σ 在基(I) 和基(II) 以及基($\widetilde{I}$) 和基($\widetilde{II}$) 下的矩阵表示. 则有

$$B=Q^{-1}AP.\tag{6.1.17}$$

证 今由式 (6.1.13) 和 (6.1.15), 有

$$\sigma(\widetilde{\alpha}_i)=\sum_{j=1}^{m}b_{ji}\widetilde{\beta}_j=\sum_{k=1}^{m}\Big(\sum_{j=1}^{m}q_{kj}b_{ji}\Big)\beta_k;$$

又有

$$\sigma(\widetilde{\alpha}_i)=\sum_{j=1}^{n}p_{ji}\sigma(\alpha_j)=\sum_{k=1}^{m}\Big(\sum_{j=1}^{n}a_{kj}p_{ji}\Big)\beta_k.$$

因为 $\beta_1,\beta_2,\cdots,\beta_m$ 为线性空间 $\mathfrak{L}_1$ 的基, 所以有

$$\sum_{j=1}^{m}q_{kj}b_{ji}=\sum_{j=1}^{n}a_{kj}p_{ji},\quad 1\leqslant i\leqslant n,\quad 1\leqslant k\leqslant m.$$

写成矩阵表达形式, 为 $QB=AP$, 即 $B=Q^{-1}AP$. □

由相抵下的标准形理论, 立即有

定理 6.1.12 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射, 则在线性空间 $\mathfrak{L}$ 中存在基(I): $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$; 在线性空间 $\mathfrak{L}_1$ 中存在基(II): $\{\beta_1,\beta_2,\cdots,\beta_m\}$, 使得线性映射 σ 关于这两组基的矩阵表示为相抵下的标准形

$$\begin{pmatrix}E_r&0\\0&0\end{pmatrix}.\tag{6.1.18}$$

因此有

$$\sigma(\alpha_i)=\beta_i,\quad 1\leqslant i\leqslant r,\quad \sigma(\alpha_j)=0,\quad r+1\leqslant j\leqslant n.\tag{6.1.19}$$

于是线性映射 σ 的象空间 $\sigma(\mathfrak{L})$ 是 $\mathfrak{L}_1$ 中以 $\beta_1,\beta_2,\cdots,\beta_r$ 为基的子空间. 线性映射 σ 的核空间是 $\mathfrak{L}$ 中以 $\alpha_{r+1},\alpha_{r+2},\cdots,\alpha_n$ 为基的子空间. 又有**维数公式**:

$$\dim(\mathfrak{L})=\dim(\sigma(\mathfrak{L}))+\dim(\ker(\sigma)).\tag{6.1.20}$$

即象空间 $\sigma(\mathfrak{L})$ 的维数加上核空间 $\ker(\sigma)$ 的维数等于原象空间 $\mathfrak{L}$ 的维数. 又象空间的维数 $\dim(\sigma(\mathfrak{L}))$ 为线性映射 σ 的任一矩阵表示的秩. 所以它也称为线性映射 σ **的秩**.

定理 6.1.13 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 内的线性映射, 则 σ 为线性同构当且仅当核空间 $\ker(\sigma)=0$, 即 $\dim(\mathfrak{L})=\dim(\sigma(\mathfrak{L}))$.

证 令线性映射 σ 为 $\mathfrak{L}$ 到 $\mathfrak{L}_1$ 上的线性同构, 由一一性, 所以 $\sigma(0)=0$ 推出 $\ker(\sigma)=0$. 反之, 设 $\ker(\sigma)=0$. 任取 $\alpha,\beta\in\mathfrak{L}$, 如果 $\sigma(\alpha)=\sigma(\beta)$, 则 $\sigma(\alpha-\beta)=0$. 所以 $\alpha-\beta\in\ker(\sigma)=0$. 这证明了 $\alpha=\beta$. 所以 σ 是线性同构. □

下面考虑线性映射的坐标表达式. 记 $x=(x_1,x_2,\cdots,x_n)'$ 为 $n\times 1$ 矩阵, $y=(y_1,y_2,\cdots,y_m)'$ 为 $m\times 1$ 矩阵. 我们有

定理 6.1.14 设 σ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 到域 $\mathbb{F}$ 上 m 维线性空间 $\mathfrak{L}_1$ 的线性映射. 在线性空间 $\mathfrak{L}$ 中取定基**(I)**: $\alpha_1,\alpha_2,\cdots,\alpha_n$; 在线性空间 $\mathfrak{L}_1$ 中取定基**(II)**: $\beta_1,\beta_2,\cdots,\beta_m$, 使得线性映射 σ 的矩阵表示为 $m\times n$ 矩阵 A. 则线性映射 σ 的**坐标表达式**为

$$y=Ax. \tag{6.1.21}$$

证 任取 $\alpha\in\mathfrak{L}$, 则 $\alpha=\sum\limits_{j=1}^{n}x_j\alpha_j$, $\sigma(\alpha)=\sum\limits_{i=1}^{m}y_i\beta_i$. 于是,

$$\sum_{j=1}^{m}y_j\beta_j=\sigma(\alpha)=\sigma\Big(\sum_{k=1}^{n}x_k\alpha_k\Big)=\sum_{k=1}^{n}x_k\sigma(\alpha_k)=\sum_{j=1}^{m}\sum_{k=1}^{n}a_{jk}x_k\beta_j.$$

由于 $\beta_1,\beta_2,\cdots,\beta_m$ 为线性空间 $\mathfrak{L}_1$ 中的基. 这证明了

$$y_j=\sum_{k=1}^{n}a_{jk}x_k,\quad 1\leqslant j\leqslant m. \tag{6.1.22}$$

写成矩阵表达形式, 则为 $y=Ax$. □

注意: 用线性映射的语言, 域 $\mathbb{F}$ 上 m 个未知数 $x_1,x_2,\cdots,x_m$ 和 n 个线性方程构成的非齐次线性方程组有矩阵形式 (4.1.3) $Ax=\beta$, 其中 A,x,β 分别为 $n\times m$, $m\times 1$ 和 $n\times 1$ 矩阵. 于是有域 $\mathbb{F}$ 上 m 维线性空间 $\mathbb{F}^m$ 到 n 维线性空间 $\mathbb{F}^n$ 内的线性映射 σ: $y=Ax$, 其中 y 为 $n\times 1$ 矩阵. 非齐次线性方程组 $Ax=\beta$ 有解的充分且必要条件就是: 何时 $\mathbb{F}^n$ 中向量 $\beta\in\sigma(\mathbb{F}^m)$? 而通解就是 $\sigma^{-1}(\beta)$. 特别, 齐次线性方程组 $Ax=0$ 的通解就是 $\ker(\sigma)$.

下面考虑线性变换, 即考虑线性空间 $\mathfrak{L}$ 到自身内的线性映射. 这时我们可以用下面办法来定义线性变换 $\mathscr{A}$ 的方阵表示. 我们用定理 6.1.10 和 6.1.14 的结果，立即有

定理 6.1.15 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. 在线性空间 $\mathfrak{L}$ 中

取定一组基(I): $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, 于是有

$$\sigma(\alpha_i) = \sum_{j=1}^{n} a_{ji}\alpha_j, \quad 1 \leqslant i \leqslant n. \tag{6.1.23}$$

因此在 $\mathfrak{L}$ 中取定的基(I) 后, 线性变换 $\mathscr{A}$ 的**方阵表示**为

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{6.1.24}$$

而且对应 $\mathscr{A} \to A$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的所有线性变换到域 $\mathbb{F}$ 上所有 n 阶方阵的一个自然的一一对应关系. 这时对线性空间 $\mathfrak{L}$ 中任一向量 $\alpha = \sum\limits_{i=1}^{n} x_i\alpha_i$, 记向量 $\mathscr{A}(\alpha) = \sum\limits_{j=1}^{n} y_j\alpha_j$, 且记 $x = (x_1, x_2, \cdots, x_n)'$, $y = (y_1, y_2, \cdots, y_n)'$. 则有

$$y = Ax. \tag{6.1.25}$$

所以式 (6.1.25) 称为线性变换 $\mathscr{A}$ 的**坐标表示式**.

用定理 6.1.11 的结果, 立即有

定理 6.1.16 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. 在线性空间 $\mathfrak{L}$ 中取定两组基(I): $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 和(II): $\{\beta_1, \beta_2, \cdots, \beta_n\}$. 记基变换公式为:

$$\beta_i = \sum_{j=1}^{n} p_{ji}\alpha_j, \quad 1 \leqslant i \leqslant n. \tag{6.1.26}$$

它对应了的 n 阶非异方阵

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}. \tag{6.1.27}$$

则线性空间 $\mathfrak{L}$ 上同一线性变换 $\mathscr{A}$ 在不同基下的方阵表示 A 和 B 间有关系

$$B = P^{-1}AP. \tag{6.1.28}$$

定义 6.1.17 域 $\mathbb{F}$ 上 n 阶方阵 A 和 B 称为**相似的**, 如果存在 n 阶非异方阵 P, 使得

$$B = P^{-1}AP. \tag{6.1.29}$$

由这个定义, 我们有矩阵论中重要概念之一: 相似概念.

定理 6.1.18　域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 上的线性变换在不同基下的方阵表示是相似的.

引理 6.1.19　域 $\mathbb{F}$ 上 n 阶方阵间的相似关系为等价关系.

证　记 A, B 和 C 为 n 阶方阵, P 和 Q 为 n 阶非异方阵. 显然 A 和 A 相似, 所以有反身性. 今若 A 和 B 相似, 即存在非异方阵 P, 使得 $B = P^{-1}AP$. 于是 $A = PBP^{-1}$. 这证明了 B 和 A 相似, 即有对称性. 最后, 若 A 和 B 相似, B 和 C 相似, 即存在 n 阶非异方阵 P 及 Q, 使得 $B = P^{-1}AP$, $C = Q^{-1}BQ$. 于是有 $C = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$. 这证明了 A 和 C 相似, 即证明了传递性. □

定理 6.1.20　给定集合 $\mathfrak{S}$, $\mathfrak{S}$ 上所有可逆变换构成集合 $\mathfrak{G} = \mathfrak{G}(\mathfrak{S})$, 它具有下列四个性质:

(1)　**乘法封闭性**: 对集合 $\mathfrak{G}$ 中任两可逆变换 σ 和 τ, 乘积 $\sigma \circ \tau$ 仍为可逆变换, 即其乘积仍在集合 $\mathfrak{G}$ 中, 而且有 $(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}$;

(2)　**结合律**成立: 对集合 $\mathfrak{G}$ 中任三可逆变换 ξ, η, ζ, 则有 $(\xi\eta)\zeta = \xi(\eta\zeta)$;

(3)　在集合 $\mathfrak{G}$ 中存在**恒等变换** $\mathrm{id} = \mathrm{id}_{\mathfrak{S}}$: $x \to x$, $\forall x \in \mathfrak{S}$. 它具有性质:

$$\mathrm{id} \circ \sigma = \sigma \circ \mathrm{id} = \sigma, \qquad \forall \sigma \in \mathfrak{G}; \tag{6.1.30}$$

(4)　对集合 $\mathfrak{G}$ 中任一可逆变换 σ, 存在唯一的**逆变换** σ^{-1}, 使得

$$\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{id}\,, \tag{6.1.31}$$

$\mathfrak{G} = \mathfrak{G}(\mathfrak{S})$ 称为集合 $\mathfrak{S}$ 上的**变换群**.

证　性质(1) 显然成立, 这时因为

$$\begin{aligned}(\tau^{-1} \circ \sigma^{-1}) \circ (\sigma \circ \tau) &= \tau^{-1} \circ (\sigma^{-1} \circ (\sigma \circ \tau)) = \tau^{-1} \circ ((\sigma^{-1} \circ \sigma) \circ \tau) \\ &= \tau^{-1} \circ (\mathrm{id} \circ \tau) = \tau^{-1} \circ \tau = \mathrm{id}\,.\end{aligned}$$

同理, $(\sigma \circ \tau) \circ (\tau^{-1} \circ \sigma^{-1}) = \mathrm{id}\,$. 由逆变换唯一存在, 所以 $(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}$.

由定理 6.1.3 可知性质(2) 成立. 性质(3) 显然成立. 由定理 6.1.5 可知性质(4) 成立. □

定理 6.1.21　设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换, 则下面性质互相等价:

(1)　线性变换 $\mathscr{A}$ 为线性同构, 即 $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 上的一一线性变换;

(2)　线性变换 $\mathscr{A}$ 的核空间 $\ker(\mathscr{A}) = 0$;

(3)　$\mathscr{A}(\mathfrak{L}) = \mathfrak{L}$, 即线性变换 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{L}$ 上的线性变换;

(4) 在线性空间 $\mathfrak{L}$ 中取定基(I) $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, 记 A 为线性变换 $\mathscr{A}$ 在基(I)下的方阵表示, 则 A 可逆.

证 先由(1) 推(2). 事实上, 由线性同构的定义可知, $\mathscr{A}(\alpha) = 0$, 则 $\alpha = 0$. 即 $\ker(\mathscr{A}) = 0$. 再由(2) 推(3). 这是因为由定理 6.1.12 的维数公式, 因此, $\dim(\mathfrak{L}) = \dim(\mathscr{A}(\mathfrak{L}))$. 所以 $\mathscr{A}(\mathfrak{L}) = \mathfrak{L}$. 再由(3) 推(4). 事实上, 由 $\mathfrak{L} = \mathscr{A}(\mathfrak{L})$, 所以 $\dim(\mathscr{A}(\mathfrak{L})) = n$. 而 $\dim(\mathscr{A}(\mathfrak{L}))$ 等于线性变换 $\mathscr{A}$ 的方阵表示的秩, 所以等于 $\dim(\mathfrak{L}) = n$. 即 $\mathscr{A}$ 在 $\mathfrak{L}$ 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的方阵表示 A 非异. 最后证由(4) 可推出(1). 今在 $\mathfrak{L}$ 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下, 线性变换 $\mathscr{A}$ 的方阵表示 A 为可逆方阵. 由 Cramer 法则可知, 线性方程组 $Ax = y$ 有唯一解 $x = A^{-1}y$. 而线性变换 $\mathscr{A}$ 的坐标表示式为 $y = Ax$. 这证明了 $\mathscr{A}$ 为线性同构. □

最后, 考虑线性变换的乘积, 再定义另外两种代数运算: 加法和纯量积, 并且给出代数运算的方阵表示.

定义 6.1.22 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间. 在线性空间 $\mathfrak{L}$ 上任给线性变换 $\mathscr{A}$ 及 $\mathscr{B}$, 以及纯量 $b \in \mathbb{F}$. 则有:

(1) 映射 $\alpha \to \mathscr{A}(\alpha) + \mathscr{B}(\alpha)$, $\forall \alpha$ 仍为 $\mathfrak{L}$ 上的线性变换, 称为线性变换 $\mathscr{A}$ 及 $\mathscr{B}$ 的**和**, 记作 $\mathscr{A} + \mathscr{B}$. 于是有

$$(\mathscr{A} + \mathscr{B})(\alpha) = \mathscr{A}(\alpha) + \mathscr{B}(\alpha), \quad \forall \alpha \in \mathfrak{L}; \tag{6.1.32}$$

(2) 映射 $\alpha \to \mathscr{A}(\mathscr{B}(\alpha))$, $\forall \alpha \in \mathfrak{L}$ 仍为 $\mathfrak{L}$ 上的线性变换, 称为线性变换 $\mathscr{A}$ 及 $\mathscr{B}$ 的 **乘积**, 记作 $\mathscr{A}\mathscr{B} = \mathscr{A} \circ \mathscr{B}$. 于是有

$$(\mathscr{A}\mathscr{B})(\alpha)) = \mathscr{A}(\mathscr{B}(\alpha)), \quad \forall \alpha \in \mathfrak{L}; \tag{6.1.33}$$

(3) 映射 $\alpha \to b(\mathscr{A}(\alpha))$, $\forall \alpha \in \mathfrak{L}$ 仍为 $\mathfrak{L}$ 上的线性变换, 称为线性变换 $\mathscr{A}$ 和纯量 b 的**纯量积**, 记作 $b\mathscr{A}$, 于是有

$$(b\mathscr{A})(\alpha) = b(\mathscr{A}(\alpha)), \quad \forall \alpha \in \mathfrak{L}. \tag{6.1.34}$$

定理 6.1.23 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基(I) $\alpha_1, \alpha_2, \cdots, \alpha_n$. 记 n 阶方阵 A 和 B 分别为线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 和 $\mathscr{B}$ 在基(I) 下的方阵表示. 记 $b \in \mathbb{F}$. 则线性变换 $\mathscr{A} + \mathscr{B}$, $\mathscr{A}\mathscr{B}$, $b\mathscr{A}$ 在基(I) 下的方阵表示分别为 $A + B$, AB, bA.

证 记 $A = (a_{ij})$, $B = (b_{ij})$, 于是

$$\mathscr{A}(\alpha_i) = \sum_{j=1}^{n} a_{ji}\alpha_j, \quad \mathscr{B}(\alpha_i) = \sum_{j=1}^{n} b_{ji}\alpha_j, \quad 1 \leqslant i \leqslant n.$$

因此

$$(\mathscr{A}+\mathscr{B})(\alpha_i)=\mathscr{A}(\alpha_i)+\mathscr{B}(\alpha_i)=\sum_{j=1}^{n}(a_{ji}+b_{ji})\alpha_j,$$

$$(\mathscr{A}\mathscr{B})(\alpha_i)=\mathscr{A}(\mathscr{B}(\alpha_i))=\mathscr{A}\Big(\sum_{j=1}^{n}b_{ji}\alpha_j\Big)=\sum_{k=1}^{n}\Big(\sum_{j=1}^{n}a_{kj}b_{ji}\Big)\alpha_k.$$

$$(b\mathscr{A})(\alpha_i)=\sum_{j=1}^{n}(ba_{ji})\alpha_j,\quad 1\leqslant i\leqslant n.$$

于是线性变换 $\mathscr{A}+\mathscr{B}$, $\mathscr{A}\mathscr{B}$, $b\mathscr{A}$ 在基(I) 下的方阵表示分别为 $A+B$, AB, bA. □

上面的定理告诉我们, 方阵的乘法、加法和纯量乘积的定义都是极其自然的, 它们有明确的几何意义.

习　题　6.1

6.1.1　试证: 平面上先作角度 θ 的旋转, 再作反射, 不等于先作反射, 再作角度 θ 的旋转. 由此说明, 一般来说, 变换的乘法不一定可交换.

6.1.2　设空间 $\mathbb{R}^3$ 中变换 σ 将平面变为平面, 试证: 它必将直线变为直线. 设空间 $\mathbb{R}^3$ 中变换 $\mathscr{A}$ 将任一对平行平面变为一对平行平面. 试证: 它必将一对平行的直线变为一对平行的直线.

6.1.3　试证: 相似的两个方阵有相同的特征多项式. 因此, 有相同的特征根, 但特征向量不一定相同. 试给出属于同一特征根的特征向量间的关系.

6.1.4　设域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的变换 σ 将点集 S_1 和 T_1 分别变为点集 S_2 和 T_2. 试证: σ 将交集 $S_1\cap T_1$ 变为交集 $S_2\cap T_2$ 的子集. 试给出 $\sigma(S_1\cap T_1)=S_2\cap T_2$ 的充分且必要条件.

6.1.5　设 σ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. 试证: 线性变换 σ 将子空间变为子空间, 非异线性变换 σ 将 r 维子空间变为 r 维子空间.

6.1.6　在域 $\mathbb{F}$ 上 n 阶方阵全体构成的线性空间 $\mathrm{gl}_n(\mathbb{F})$ 上引进映射 $\mathscr{A}$: $X\to AXB+CX+XD$, 其中 A, B, C, D 都是给定的 n 阶方阵. 试证: $\mathscr{A}$ 为 $\mathrm{gl}_n(\mathbb{F})$ 上线性变换. 设 $C=D=0$, 则 $\mathscr{A}$ 非异当且仅当 $\det(AB)\neq 0$. 设 $A=0$, $D=-C$, 试给出这个线性变换的方阵表示.

6.1.7　域 $\mathbb{F}$ 上次数不超过 n 的多项式全体构成的 $n+1$ 维线性空间 $\mathbb{F}_n[x]$. 在线性空间 $\mathbb{F}_n[x]$ 上引进映射 $f(x)\to f(x+1)-f(x)$, $\forall f(x)\in\mathbb{F}_n[x]$. 试回答下列问题: 这个映射是不是线性变换.

6.1.8　设域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 有 $\mathscr{A}^{n-1}\neq 0$, $\mathscr{A}^n=0$, 试证: 在 $\mathfrak{L}$ 中存在一组基, 使得在这组基下, 线性变换 $\mathscr{A}$ 的方阵表示为 N_n, 其中 N_n 由式 (3.1.60) 定义.

6.1.9　设 $\mathrm{gl}_n(\mathbb{F})$ 为域 $\mathbb{F}$ 上 n 阶方阵全体构成的线性空间. n 阶方阵 A 称为**半单的**, 如果 A 相似于对角方阵. 设 A 半单, 试证: 映射 f_A: $X\to XA-AX$, $\forall X\in\mathrm{gl}_n(\mathbb{F})$ 为 $\mathrm{gl}_n(\mathbb{F})$ 的线性变换, 而且 f_A 仍为半单的.

6.1.10　记 $\mathbb{R}_n[x,y]$ 为所有实多项式 $f(x,y)$ 构成的集合, 其中 $f(x,y)$ 为关于独立未知数 x 和 y 的多项式, 且次数都不超过 n. 试在 $\mathbb{R}_n[x,y]$ 中给出一组基, 并写出 $\mathbb{R}_n[x,y]$ 的维数. 记映射 σ: $f(x,y)\to(x^2+y^2)f(x,y)$, $\forall f(x,y)\in\mathbb{R}_n[x,y]$, 则有

$$\tau=\Big(\frac{\partial}{\partial x}\frac{\partial}{\partial y}+\frac{\partial}{\partial y}\frac{\partial}{\partial x}\Big)\sigma-\sigma\Big(\frac{\partial}{\partial x}\frac{\partial}{\partial y}+\frac{\partial}{\partial y}\frac{\partial}{\partial x}\Big)=\Big[\Big(\frac{\partial}{\partial x}\frac{\partial}{\partial y}+\frac{\partial}{\partial y}\frac{\partial}{\partial x}\Big),\sigma\Big]$$

为 $\mathbb{R}_n[x,y]$ 上的线性变换. 试求线性变换 τ 在基 $\dfrac{1}{j!k!}x^jy^k$, $0 \leqslant j,k \leqslant n$ 下的方阵表示.

6.1.11 设 $\mathfrak{L}_0$, $\mathfrak{L}_1$, $\cdots$, $\mathfrak{L}_m$, $\mathfrak{L}_{m+1}$ 都是域 $\mathbb{F}$ 上的线性空间, 其中 $\mathfrak{L}_0 = \mathfrak{L}_{m+1} = 0$. 设存在线性映射 $\mathscr{A}_i$: $\mathfrak{L}_i \to \mathfrak{L}_{i+1}$, $i = 0$, 1, $\cdots$, m, 它们有 $\ker(\mathscr{A}_{i+1}) = \mathscr{A}_i(\mathfrak{L}_i)$, $i = 0,1,\cdots,m$. 试证: $\sum\limits_{i=0}^{m+1}(-1)^i\dim(\mathfrak{L}_i) = 0$.

6.1.12 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的连续变换. 即在线性空间 $\mathfrak{L}$ 中任意取定一组基 α_1, α_2, $\cdots$, α_n, 任取 $\mathfrak{L}$ 中的向量 $\alpha = \sum\limits_{j=1}^{n} x_j\alpha_j$, 则 $\mathscr{A}(\alpha) = \sum\limits_{j=1}^{n} f_j(x_1,x_2,\cdots,x_n)\alpha_j$, 其中 f_1, f_2, $\cdots$, f_n 为 x_1, x_2, $\cdots$, x_n 的连续函数. 设 $\mathscr{A}$ 还有性质: 对 $\mathfrak{L}$ 上任两线性变换 $\mathscr{A}_1$ 和 $\mathscr{A}_2$, 则 $\mathscr{A}(\mathscr{A}_1 + \mathscr{A}_2) = \mathscr{A}\mathscr{A}_1 + \mathscr{A}\mathscr{A}_2$ 成立. 试证: $\mathscr{A}$ 也是线性变换.

6.1.13 设 $\mathscr{A}$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的幂零线性变换, $\mathscr{A}_1$, $\mathscr{A}_2$, $\cdots$, $\mathscr{A}_s$, $\mathscr{B}_1$, $\mathscr{B}_2$, $\cdots$, $\mathscr{B}_s$ 都是线性变换, 而且 $\mathscr{A}$ 和 $\mathscr{A}_k$ 可交换, $1 \leqslant k \leqslant s$, 又

$$\mathscr{A} = \sum_{k=1}^{s}[\mathscr{A}_k, \mathscr{B}_k] = \sum_{k=1}^{s}(\mathscr{A}_k\mathscr{B}_k - \mathscr{B}_k\mathscr{A}_k).$$

试证: $\mathscr{A}$ 是幂零线性变换.

6.1.14 设 $\mathfrak{S}$ 为齐次线性方程组 $ABx = 0$ 的解空间, 其中 x 是 $p \times 1$ 矩阵, A 是 $n \times m$ 矩阵, B 是 $m \times p$ 矩阵. 试证: m 维向量空间 $\mathbb{F}^m$ 中子集合 $\mathfrak{L}_1 = \{y = Bx \,|\, x \in \mathfrak{S}\}$ 是子空间, 它的维数等于 $\operatorname{rank}(B) - \operatorname{rank}(AB)$. 且有 $\operatorname{rank}(A) = \operatorname{rank}(AB)$ 当且仅当由 $ABx = 0$ 可推出 $Bx = 0$.

利用上述结论, 试证: 对任意三个 n 阶方阵 A, B, C, 有 $\operatorname{rank}(AB)+\operatorname{rank}(BC) \leqslant \operatorname{rank}(B)+\operatorname{rank}(ABC)$. 特别, 在 B 是单位方阵时, 有 $\operatorname{rank}(A) + \operatorname{rank}(C) \leqslant \operatorname{rank}(AC) + n$. 试另外用矩阵在相抵下标准形来证明这个不等式, 且比较何者简单.

6.1.15 设域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 的特征多项式为 $f(\lambda)$. 设 $f(\lambda) = g(\lambda)h(\lambda)$, 其中 $(g(\lambda),h(\lambda)) = 1$. 试证: 在 $\mathfrak{L}$ 中存在 $\mathscr{A}$ 的不变子空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$, 使得 $g(\mathscr{A}|_{\mathfrak{L}_1}) = 0$, $h(\mathscr{A}|_{\mathfrak{L}_2}) = 0$, 又 $\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2$ 为子空间直接和. (用这一事实, 试给出第一空间分解定理的新证明. 这个结论在 §7.1 中作.)

6.1.16 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间的线性变换, $f(x)$ 和 $g(x)$ 为域 $\mathbb{F}$ 上两个多项式. 试证:

(1) $\ker(f(\mathscr{A})) + \ker(g(\mathscr{A})) \subset \ker(f(\mathscr{A})g(\mathscr{A}))$;

(2) 当 $(f(x),g(x)) = 1$ 时, 则有下面子空间直接和分解 $\ker(f(\mathscr{A})g(\mathscr{A})) = \ker(f(\mathscr{A})) \oplus \ker(g(\mathscr{A}))$.

6.1.17 (Schur 引理) 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换集 $\mathfrak{S}$ 称为**不可约的**, 如果除了 $\mathfrak{L}$ 及零空间外, 没有其他在 $\mathfrak{S}$ 中所有线性变换作用下都不变的子空间.

设 $\mathfrak{S}_1$ 和 $\mathfrak{S}_2$ 分别为域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 上的不可约线性变换集. 设存在 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 内的线性映射 σ, 使得任取 $\mathscr{A} \in \mathfrak{S}_1$, 则存在 $\mathscr{B} \in \mathfrak{S}_2$, 使得 $\sigma \circ \mathscr{A} = \mathscr{B} \circ \sigma$, 且当 $\mathscr{A}$ 遍历 $\mathfrak{S}_1$ 时, $\mathscr{B}$ 也遍历 $\mathfrak{S}_2$. 试证: $\sigma = 0$ 或者 $\dim(\mathfrak{L}_1) = \dim(\mathfrak{L}_2)$, 且 σ 为 $\mathfrak{L}_1$ 到 $\mathfrak{L}_2$ 上的线性同构.

特别, 当 $\mathfrak{L}_1 = \mathfrak{L}_2$, 且 $\mathfrak{L}_1$ 为复线性空间, 又当 $\mathfrak{S}_1 = \mathfrak{S}_2$, $\sigma \circ \mathscr{A} = \mathscr{A} \circ \sigma$, $\forall \mathscr{A} \in \mathfrak{S}_1$ 时, 试证: σ 为 $\mathfrak{L}_1$ 上的纯量线性变换, 即存在复数 c, 使得 $\sigma = c(\mathrm{id})$.

6.1.18 给定 n^2 个 n 阶非零方阵 A_{jk}, $j,k = 1$, 2, $\cdots$, n. 设有 $A_{ij}A_{pq} = \delta_{jp}A_{iq}$, $i,j,p,q = 1,2,\cdots,n$. 试证: 存在 n 阶非异方阵 P, 使得 $P^{-1}A_{ij}P = E_{ij}$, $i,j = 1,2,\cdots,n$, 其

中 E_{ij} 为第 i 行, 第 j 列元素为 1, 其余元素为零的 n 阶方阵.

§6.2 商空间和不变子空间

上面一节实质上是讨论了线性变换和它的方阵表示之间的关系. 下面进一步讨论线性变换和子空间之间的关系.

下面类似于有理数和整数的关系, 利用线性空间和它的子空间, 引进一种等价关系, 称为"模相等"关系. 从而引进线性空间的商空间这个重要概念.

定义 6.2.1 设 $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间. 任取 $\alpha, \beta \in \mathfrak{L}$, 如果 $\alpha-\beta \in \mathfrak{L}_1$, 则称 α 和 β 是模子空间 $\mathfrak{L}_1$ **同余的**, 或称为模子空间 $\mathfrak{L}_1$ **相等的**, 记作

$$\alpha \equiv \beta, \quad (\operatorname{mod} \mathfrak{L}_1). \tag{6.2.1}$$

引理 6.2.2 设 $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间. 线性空间 $\mathfrak{L}$ 关于子空间 $\mathfrak{L}_1$ 的模相等关系是等价关系.

证 由 $\alpha-\alpha=0 \in \mathfrak{L}_1$ 可知 $\alpha \equiv \alpha$, $(\operatorname{mod} \mathfrak{L}_1)$, 即反身性成立. 再设 $\alpha-\beta \in \mathfrak{L}_1$, 显然 $\beta-\alpha \in \mathfrak{L}_1$, 所以由 $\alpha \equiv \beta$, $(\operatorname{mod} \mathfrak{L}_1)$ 可知 $\beta \equiv \alpha$, $(\operatorname{mod} \mathfrak{L}_1)$, 即对称性成立. 最后, 由 $\alpha-\beta \in \mathfrak{L}_1$, $\beta-\gamma \in \mathfrak{L}_1$, 则有 $\alpha-\gamma \in \mathfrak{L}_1$. 所以由 $\alpha \equiv \beta$, $(\operatorname{mod} \mathfrak{L}_1)$ 和 $\beta \equiv \gamma$, $(\operatorname{mod} \mathfrak{L}_1)$ 推出 $\alpha \equiv \gamma$, $(\operatorname{mod} \mathfrak{L}_1)$. 即传递性成立. □

定义 6.2.3 设 $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间. 在线性空间 $\mathfrak{L}$ 中关于子空间 $\mathfrak{L}_1$ 按模相等引进等价关系后, 线性空间 $\mathfrak{L}$ 便分为互不相交的等价类的并集, 等价类全体构成的集合记作 $\mathfrak{L}/\mathfrak{L}_1$, 称为线性空间 $\mathfrak{L}$ 模子空间 $\mathfrak{L}_1$ 的**商空间**.

现在在商空间 $\mathfrak{L}/\mathfrak{L}_1$ 中引进加法及纯量积如下:

定理 6.2.4 设 $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间. 在商空间 $\mathfrak{L}/\mathfrak{L}_1$ 中任取两个元素 $\alpha+\mathfrak{L}_1$ 和 $\beta+\mathfrak{L}_1$, 任取 $c \in \mathbb{F}$. **加法**定义为:

$$(\alpha+\mathfrak{L}_1)+(\beta+\mathfrak{L}_1)=(\alpha+\beta)+\mathfrak{L}_1. \tag{6.2.2}$$

纯量积定义为:

$$c(\alpha+\mathfrak{L}_1)=c\alpha+\mathfrak{L}_1. \tag{6.2.3}$$

则商空间 $\mathfrak{L}/\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上线性空间.

证 我们先来证明加法和纯量积的定义有意义. 因为加法和纯量积的定义依赖于代表元素的选取, 所以为了证明上述定义只是等价类之间的运算, 需要证明它们实际上与代表元素的选取无关. 事实上, 任取 $\widetilde{\alpha} \in \alpha+\mathfrak{L}_1$, $\widetilde{\beta} \in \beta+\mathfrak{L}_1$, 则存在 ξ, $\eta \in \mathfrak{L}_1$, 使得 $\widetilde{\alpha}=\alpha+\xi$, $\widetilde{\beta}=\beta+\eta$. 令 $\widetilde{\alpha}+\widetilde{\beta}=\alpha+\beta+(\xi+\eta)$, 其中 $\xi+\eta \in \mathfrak{L}_1$. 所以 $(\widetilde{\alpha}+\widetilde{\beta})+\mathfrak{L}_1=(\alpha+\beta)+\mathfrak{L}_1$. 这证明了当 $\alpha+\mathfrak{L}_1=\widetilde{\alpha}+\mathfrak{L}_1$, $\beta+\mathfrak{L}_1=\widetilde{\beta}+\mathfrak{L}_1$,

则 $(\alpha+\beta)+\mathfrak{L}_1=(\widetilde{\alpha}+\widetilde{\beta})+\mathfrak{L}_1$. 因此加法运算的定义与代表元选取无关. 再任取 $\widetilde{\alpha}\in\alpha+\mathfrak{L}_1$, 数 $c\in\mathbb{F}$, 则存在 $\xi\in\mathfrak{L}_1$, 使得 $\widetilde{\alpha}=\alpha+\xi$. 因此, $c\widetilde{\alpha}=c\alpha+c\xi$, 其中 $c\xi\in\mathfrak{L}_1$. 所以 $c\widetilde{\alpha}+\mathfrak{L}_1=c\alpha++c\xi+\mathfrak{L}_1=c\alpha+\mathfrak{L}_1$. 这证明了当 $\alpha+\mathfrak{L}_1=\widetilde{\alpha}+\mathfrak{L}_1$, 则 $c\widetilde{\alpha}+\mathfrak{L}_1=(c\alpha++c\xi)+\mathfrak{L}_1=c\alpha+\mathfrak{L}_1$. 因此纯量积运算的定义也与代表元素选取无关. 由线性空间的定义及直接验证可知商空间 $\mathfrak{L}/\mathfrak{L}_1$ 在上述加法及纯量积下构成线性空间. □

定理 6.2.5 设 $\mathfrak{L}_1$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间. 在子空间 $\mathfrak{L}_1$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_r$, 再在线性空间 $\mathfrak{L}$ 中补足 $n-r$ 个向量 $\alpha_{r+1}, \cdots, \alpha_n$, 使得 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$ 是 $\mathfrak{L}$ 的一组基. 则商空间 $\mathfrak{L}/\mathfrak{L}_1$ 有一组基 $\alpha_{r+1}+\mathfrak{L}_1, \alpha_{r+2}+\mathfrak{L}_1, \cdots, \alpha_n+\mathfrak{L}_1$.

证 任取 $\alpha\in\mathfrak{L}$, 则 $\alpha=\sum\limits_{j=1}^{n}c_j\alpha_j$. 于是 $\alpha+\mathfrak{L}_1=\sum\limits_{j=r+1}^{n}c_j(\alpha_j+\mathfrak{L}_1)$. 所以为了证明 $\alpha_{r+1}+\mathfrak{L}_1, \alpha_{r+2}+\mathfrak{L}_1, \cdots, \alpha_n+\mathfrak{L}_1$ 为商空间 $\mathfrak{L}/\mathfrak{L}_1$ 的基, 只要证它们线性无关. 今

$$\sum_{k=r+1}^{n}c_k(\alpha_k+\mathfrak{L}_1)=(\sum_{k=r+1}^{n}c_k\alpha_k)+\mathfrak{L}_1=0+\mathfrak{L}_1$$

当且仅当 $\sum\limits_{k=r+1}^{n}c_k\alpha_k\in\mathfrak{L}_1$. 由 $\mathfrak{L}_1$ 有基 $\alpha_1, \cdots, \alpha_r$, 所以 $c_{r+1}=\cdots=c_n=0$. 这证明了 $\alpha_{r+1}+\mathfrak{L}_1, \cdots, \alpha_n+\mathfrak{L}_1$ 线性无关. 所以它们是商空间 $\mathfrak{L}/\mathfrak{L}_1$ 的基. □

很自然地, 我们可以建立线性空间 $\mathfrak{L}$ 到商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的映射 π: $\alpha\to\alpha+\mathfrak{L}_1$. 显然它的核是 $\ker(\pi)=\pi^{-1}(0)=\mathfrak{L}_1$. 上面加法及纯量积的定义告诉我们有

$$\pi(\alpha)+\pi(\beta)=\pi(\alpha+\beta),\quad c\pi(\alpha)=\pi(c\alpha),\quad \forall c\in\mathbb{F},\quad \alpha,\ \beta\in\mathfrak{L}. \tag{6.2.4}$$

这证明了 π 是 $\mathfrak{L}$ 到商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的线性映射, 称为**自然映射**, 或称为**自然同态**.

由定理 6.1.12, 有维数公式:

$$\dim(\mathfrak{L})=\dim(\pi(\mathfrak{L}))+\dim(\ker(\pi))=\dim(\mathfrak{L}/\mathfrak{L}_1)+\dim(\mathfrak{L}_1),$$

因此有下面商空间的**维数公式**:

$$\dim(\mathfrak{L}/\mathfrak{L}_1)=\dim(\mathfrak{L})-\dim(\mathfrak{L}_1). \tag{6.2.5}$$

最后, 我们给出

定理 6.2.6 (同态基本定理) 设 σ 是线性空间 $\mathfrak{L}$ 到线性空间 $\mathfrak{L}_0$ 上的线性映射. 同态核 $\ker(\sigma)=\sigma^{-1}(0)=\mathfrak{L}_1$ 为 $\mathfrak{L}$ 的子空间. 记 π: $\mathfrak{L}\to\mathfrak{L}/\mathfrak{L}_1$ 为自然同态, 则存在商空间 $\mathfrak{L}/\mathfrak{L}_1$ 到线性空间 $\mathfrak{L}_0$ 上的同构映射 ρ, 使得

$$\sigma=\rho\circ\pi. \tag{6.2.6}$$

证 任取 $\alpha\in\mathfrak{L}$, 则 $\pi(\alpha)=\alpha+\mathfrak{L}_1$. 建立 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上的对应 ρ: $\alpha+\mathfrak{L}_1\to\sigma(\alpha)$, $\forall\alpha\in\mathfrak{L}$. 我们来证明 ρ 为 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上的线性同构, 且有 $\sigma=\rho\circ\pi$. 事实上, 任取 $\alpha\in\mathfrak{L}$, 则 $\rho\circ\pi(\alpha)=\rho(\pi(\alpha))=\rho(\alpha+\mathfrak{L}_1)=\sigma(\alpha)$. 这证明了 $\sigma=\rho\circ\pi$. 再证 ρ 为一一映射. 今若 $\alpha+\mathfrak{L}_1=\beta+\mathfrak{L}_1$, 于是 $\alpha-\beta\in\mathfrak{L}_1$. 所以由 $\sigma(\mathfrak{L}_1)=0$ 有 $\sigma(\alpha-\beta)=0$, 即 $\sigma(\alpha)=\sigma(\beta)$. 这证明了 ρ 是单值映射. 反之, 任取 $\alpha,\beta\in\mathfrak{L}$, 且 $\sigma(\alpha)=\sigma(\beta)$, 所以 $\sigma(\alpha-\beta)=0$, 即 $\alpha-\beta\in\mathfrak{L}_1$. 这证明了 $\alpha+\mathfrak{L}_1=\beta+\mathfrak{L}_1$, 即 ρ 为一一映射. 最后证 ρ 为线性同构. 事实上, 任取 $c,d\in\mathbb{F}$ 为数, $\alpha,\beta\in\mathfrak{L}$, 则

$$\begin{aligned}&\rho(c(\alpha+\mathfrak{L}_1)+d(\beta+\mathfrak{L}_1))=\rho(c\alpha+d\beta+\mathfrak{L}_1)=\rho(\pi(c\alpha+d\beta))\\&=\sigma(c\alpha+d\beta)=c\sigma(\alpha)+d\sigma(\beta)=c\rho(\alpha+\mathfrak{L}_1)+d\rho(\beta+\mathfrak{L}_1).\end{aligned}$$

这证明了 ρ 为 $\mathfrak{L}/\mathfrak{L}_1$ 到 $\mathfrak{L}_0$ 上的线性同构. □

下面开始引进线性空间的不变子空间这个重要概念.

定义 6.2.7 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的子空间 $\mathfrak{L}_1$ 称为 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 的**不变子空间**, 如果

$$\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1. \tag{6.2.7}$$

对不变子空间 $\mathfrak{L}_1$ 而言, 线性变换 $\mathscr{A}$ 可以看作是不变子空间 $\mathfrak{L}_1$ 到自身内的单值对应. 所以它诱导了 $\mathfrak{L}_1$ 上的线性变换. 这个线性变换记作 $\mathscr{A}_1|_{\mathfrak{L}_1}$, 称为线性变换 $\mathscr{A}$ 在不变子空间 $\mathfrak{L}_1$ 上的**限制**.

例如: $\mathfrak{L}$ 本身就是线性变换 $\mathscr{A}$ 的不变子空间; 线性变换 $\mathscr{A}$ 的象空间 $\mathscr{A}(\mathfrak{L})$ 及核空间 $\ker(\mathscr{A})$ 也都是线性变换 $\mathscr{A}$ 的不变子空间.

我们来考虑商空间和不变子空间之间的关系.

定理 6.2.8 设 $\mathscr{A}$ 是线性空间 $\mathfrak{L}$ 上的线性变换, $\mathfrak{L}_1$ 为线性变换 $\mathscr{A}$ 的不变子空间. 则线性变换 $\mathscr{A}$ 在商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上诱导了一个线性变换 $\widetilde{\mathscr{A}}$, 它定义为:

$$\widetilde{\mathscr{A}}(\alpha+\mathfrak{L}_1)=\mathscr{A}(\alpha)+\mathfrak{L}_1,\quad \forall\alpha\in\mathfrak{L}. \tag{6.2.8}$$

证 我们先来证明映射 $\widetilde{\mathscr{A}}$ 的定义有意义, 即证明映射 $\widetilde{\mathscr{A}}$ 和代表元素的选取无关, 所以它是单值映射. 事实上, 任取 $\widetilde{\alpha}\in\alpha+\mathfrak{L}_1$, 则存在 $\xi\in\mathfrak{L}_1$, 使得 $\widetilde{\alpha}=\alpha+\xi$. 今 $\mathscr{A}(\xi)\in\mathfrak{L}_1$, 所以

$$\begin{aligned}\widetilde{\mathscr{A}}(\widetilde{\alpha}+\mathfrak{L}_1)&=\mathscr{A}(\widetilde{\alpha})+\mathfrak{L}_1=\mathscr{A}(\alpha+\xi)+\mathfrak{L}_1=\mathscr{A}(\alpha)+\mathscr{A}(\xi)+\mathfrak{L}_1\\&=\mathscr{A}(\alpha)+\mathfrak{L}_1=\widetilde{\mathscr{A}}(\alpha+\mathfrak{L}_1).\end{aligned}$$

这证明了映射 $\widetilde{\mathscr{A}}$ 的定义与代表元素选取无关.

下面来证明映射 $\widetilde{\mathscr{A}}$ 是商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的线性变换. 事实上, 任取 $\alpha, \beta \in \mathfrak{L}$, c, $d \in \mathbb{F}$, 则

$$\begin{aligned}\widetilde{\mathscr{A}}(c(\alpha+\mathfrak{L}_1)+d(\beta+\mathfrak{L}_1)) &= \widetilde{\mathscr{A}}((c\alpha+d\beta)+\mathfrak{L}_1) = \mathscr{A}(c\alpha+d\beta)+\mathfrak{L}_1 \\ &= (c\mathscr{A}(\alpha)+d\mathscr{A}(\beta))+\mathfrak{L}_1 \\ &= c(\mathscr{A}(\alpha)+\mathfrak{L}_1)+d(\mathscr{A}(\beta)+\mathfrak{L}_1) \\ &= c\widetilde{\mathscr{A}}(\alpha+\mathfrak{L}_1)+d\widetilde{\mathscr{A}}(\beta+\mathfrak{L}_1).\end{aligned}$$

这证明了 $\widetilde{\mathscr{A}}$ 是商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的线性变换. □

我们有

引理 6.2.9 设 $\mathscr{A}$ 是线性空间 $\mathfrak{L}$ 的线性变换, $\mathfrak{L}_1$ 是线性变换 $\mathscr{A}$ 的不变子空间. 在 $\mathfrak{L}_1$ 中取定一组基(**I**): $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$, 再在线性空间 $\mathfrak{L}$ 中补足 $n-r$ 个向量, 使成 $\mathfrak{L}$ 的一组基(**II**): $\{\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n\}$. 则(**III**): $\{\alpha_{r+1}+\mathfrak{L}_1, \cdots, \alpha_n+\mathfrak{L}_1\}$ 为商空间 $\mathfrak{L}/\mathfrak{L}_1$ 的一组基 (由定理 6.2.5). 在上述基(**II**) 下, 线性变换 $\mathscr{A}$ 的方阵表示为

$$A=\begin{pmatrix} A_{11}^{(r)} & A_{12}^{(r,n-r)} \\ 0 & A_{22}^{(n-r)} \end{pmatrix}. \tag{6.2.9}$$

它是准上三角形矩阵, 其中

$$A_{11}=\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}, \tag{6.2.10}$$

$$A_{12}=\begin{pmatrix} a_{1,r+1} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{r,r+1} & \cdots & a_{rn} \end{pmatrix}, \quad A_{22}=\begin{pmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots \\ a_{n,r+1} & \cdots & a_{nn} \end{pmatrix}, \tag{6.2.11}$$

且由 $\mathscr{A}(\mathfrak{L}_1) \subset \mathfrak{L}_1$, 记 $\mathscr{A}$ 在 $\mathfrak{L}_1$ 上的限制为 $\mathscr{A}|_{\mathfrak{L}_1}$. 在上述基(**I**) 下, 线性变换 $\mathscr{A}|_{\mathfrak{L}_1}$ 的方阵表示为 A_{11}; 又 $\mathscr{A}$ 在商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上诱导了线性变换 $\widetilde{\mathscr{A}}$, 则线性变换 $\widetilde{\mathscr{A}}$ 在基(**III**) 下的方阵表示为 A_{22}.

证 今在基 $\{\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n\}$ 下, 线性变换 $\mathscr{A}$ 的方阵表示 $A=(a_{ij})$ 有 $\mathscr{A}(\alpha_i)=\sum\limits_{j=1}^{n} a_{ji}\alpha_j$, $1 \leqslant i \leqslant n$. 但是 $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ 为 $\mathfrak{L}_1$ 的基. 由 $\mathscr{A}(\mathfrak{L}_1) \subset \mathfrak{L}_1$, 所以 $\mathscr{A}(\alpha_1), \mathscr{A}(\alpha_2), \cdots, \mathscr{A}(\alpha_r)$ 是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合. 这证明了 $a_{ji}=0$, $1 \leqslant i \leqslant r$, $r+1 \leqslant j \leqslant n$. 且

$$\mathscr{A}(\alpha_i)=\sum_{k=1}^{r} a_{ki}\alpha_k, \quad 1 \leqslant i \leqslant r, \quad \mathscr{A}(\alpha_j)=\sum_{k=1}^{n} a_{kj}\alpha_k, \quad r+1 \leqslant j \leqslant n. \tag{6.2.12}$$

所以对这组基而言, 线性变换 $\mathscr{A}$ 的方阵表示为式 (6.2.9), 其中式 (6.2.10) 给出了线性变换 $\mathscr{A}|_{\mathfrak{L}_1}$ 的方阵表示.

为了给出 $n-r$ 阶方阵 A_{22} 的几何意义, 我们来看式 (6.2.9). 由定理 6.2.5 和定理 6.2.8, 商空间 $\mathfrak{L}/\mathfrak{L}_1$ 的线性变换 $\widetilde{\mathscr{A}}$ 有

$$\widetilde{\mathscr{A}}(\alpha_j+\mathfrak{L}_1)=\mathscr{A}(\alpha_j)+\mathfrak{L}_1=\sum_{k=r+1}^{n}a_{kj}\alpha_k+\mathfrak{L}_1=\sum_{k=r+1}^{n}a_{kj}(\alpha_k+\mathfrak{L}_1),\quad r+1\leqslant j\leqslant n. \tag{6.2.13}$$

所以商空间 $\mathfrak{L}/\mathfrak{L}_1$ 的线性变换 $\widetilde{\mathscr{A}}$ 在基 $\{\alpha_{r+1}+\mathfrak{L}_1,\cdots,\alpha_n+\mathfrak{L}_1\}$ 下的方阵表示为 A_{22}. □

反之, 显然有

引理 6.2.10 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 的线性变换. 设在 $\mathfrak{L}$ 中存在一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n\}$, 使得 $\mathscr{A}$ 在这组下有方阵表示

$$A=\begin{pmatrix}A_{11}^{(r)} & A_{12}^{(r,n-r)}\\ 0 & A_{22}^{(n-r)}\end{pmatrix}. \tag{6.2.14}$$

则以 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 为基的子空间 $\mathfrak{L}_1$ 为线性变换 $\mathscr{A}$ 的不变子空间. 且 $\mathscr{A}|_{\mathfrak{L}_1}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 下的方阵表示为 A_{11}. 又以 $\alpha_{r+1}+\mathfrak{L}_1,\cdots,\alpha_n+\mathfrak{L}_1$ 为基的商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上的线性变换 $\widetilde{\mathscr{A}}$: $\alpha+\mathfrak{L}_1\to\mathscr{A}(\alpha)+\mathfrak{L}_1,\forall\alpha\in\mathfrak{L}$ 在基 $\alpha_{r+1}+\mathfrak{L}_1,\cdots,\alpha_n+\mathfrak{L}_1$ 下的方阵表示为 A_{22}.

由于 n 维线性空间 $\mathfrak{L}$ 上线性变换在不同基下的方阵表示互相相似, 所以对 n 阶方阵 A, 找 n 阶非异方阵 P, 使得 A 相似于准上三角方阵 $P^{-1}AP$ 的问题, 变为首先在 $\mathfrak{L}$ 中寻找线性变换 $\mathscr{A}$ 的不变子空间的问题. 我们还有

引理 6.2.11 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 的线性变换, 而且 $\mathfrak{L}$ 是线性变换 $\mathscr{A}$ 的不变子空间 $\mathfrak{L}_1,\cdots,\mathfrak{L}_s$ 的子空间直接和: $\mathfrak{L}=\sum_{j=1}^{s}{}'\,\mathfrak{L}_j$ 当且仅当在线性空间 $\mathfrak{L}$ 中存在一组基, 使得在这组基下, 线性变换 $\mathscr{A}$ 的方阵表示为准对角形

$$A=\operatorname{diag}(A_{11},A_{22},\cdots,A_{ss}). \tag{6.2.15}$$

证 由归纳法, 下面只证 $s=2$ 的情形. 假设 $\mathfrak{L}$ 是线性变换 $\mathscr{A}$ 的不变子空间 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 的子空间直接和: $\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2$. 那末, 由 $\mathscr{A}(\mathfrak{L}_1)\subset\mathfrak{L}_1$, $\mathscr{A}(\mathfrak{L}_2)\subset\mathfrak{L}_2$, 在 $\mathfrak{L}_1$ 及 $\mathfrak{L}_2$ 中各取一组基 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 和 $\alpha_{r+1},\cdots,\alpha_n$. 于是, 线性变换 $\mathscr{A}$ 在这两组基下有

$$\mathscr{A}(\alpha_i)=\sum_{j=1}^{r}a_{ji}\alpha_j,\quad 1\leqslant i\leqslant r,\quad \mathscr{A}(\alpha_i)=\sum_{j=r+1}^{r}a_{ji}\alpha_j,\quad r+1\leqslant i\leqslant n. \tag{6.2.16}$$

因此 $\mathscr{A}$ 在线性空间 $\mathfrak{L}$ 的基 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$ 的方阵表示为准对角形

$$A = \operatorname{diag}(A_{11}^{(r)}, A_{22}^{(n-r)}).$$

反之, 如果线性变换 $\mathscr{A}$ 在线性空间 $\mathfrak{L}$ 的基 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$ 下的方阵表示为准对角形 $A = \operatorname{diag}(A_{11}^{(r)}, A_{22}^{(n-r)})$. 则有

$$\mathscr{A}(\alpha_i) = \sum_{j=1}^{r} a_{ji}\alpha_j, \quad 1 \leqslant i \leqslant r, \quad \mathscr{A}(\alpha_i) = \sum_{j=r+1}^{r} a_{ji}\alpha_j, \quad r+1 \leqslant i \leqslant n. \tag{6.2.17}$$

因此, 以 $\alpha_1, \cdots, \alpha_r$ 为基的子空间 $\mathfrak{L}_1$ 是 $\mathscr{A}$ 的不变子空间. 同理, 以 $\alpha_{r+1}, \cdots, \alpha_n$ 为基的子空间 $\mathfrak{L}_2$ 也是 $\mathscr{A}$ 的不变子空间. 而且 $\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2$. □

下面进一步讨论线性变换和子空间之间的关系. 我们特别注意一维不变子空间的重要性. 下面我们来求所有一维不变子空间.

定义 6.2.12 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为 $\mathfrak{L}$ 的线性变换. 在线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 记 $\mathscr{A}$ 在这组基下的方阵表示为 A, 则方阵 A 的特征多项式、特征根、极小多项式、也分别称为线性变换 $\mathscr{A}$ 的**特征多项式**、**特征根**、**极小多项式**.

定义 6.2.12 的合理性由下面引理给出:

引理 6.2.13 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为 $\mathfrak{L}$ 的线性变换. 则线性变换 $\mathscr{A}$ 的特征多项式、极小多项式和特征根都与基的选取无关.

证 由于在不同基下的线性变换的方阵表示是相似的, 所以只要证明特征多项式, 极小多项式和特征根都在相似下不改变就行了. 事实上, 设 P 为任一 n 阶非异方阵, 则任取 n 阶方阵 A, 有

$$\det(\lambda E_n - A) = \det P^{-1}(\lambda E_n - A)P = \det(\lambda E_n - P^{-1}AP).$$

于是 n 阶方阵 A 的特征多项式及特征根和 $P^{-1}AP$ 的特征多项式及特征根相同.

任取 n 阶方阵 A 的多项式 $g(A)$, 则有 $g(P^{-1}AP) = P^{-1}g(A)P$ 对任一 n 阶非异方阵 P 成立. 所以多项式集合

$$\mathfrak{S}_A = \{\, h(\lambda) \in \mathbb{F}[\lambda] \mid h(A) = 0 \,\} = \{\, h(\lambda) \in \mathbb{F}[\lambda] \mid h(P^{-1}AP) = 0 \,\} = \mathfrak{S}_{P^{-1}AP}.$$

因此多项式 $m(\lambda)$ 为方阵 A 的极小多项式当且仅当多项式 $m(\lambda)$ 为方阵 $P^{-1}AP$ 的极小多项式. 这证明了引理. □

定义 6.2.14 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为 $\mathfrak{L}$ 的线性变换. 设 $\lambda \in \mathbb{C}$ 为线性变换 $\mathscr{A}$ 的特征根. 对线性变换 $\mathscr{A}$, 如果存在非零向量 $\alpha \in \mathfrak{L}^{\mathbb{C}}$ (定义见引理 5.2.10), 使得

$$\mathscr{A}(\alpha) = \lambda\alpha, \tag{6.2.18}$$

则 λ 称为线性变换 $\mathscr{A}$ 的特征根, α 称为线性变换 $\mathscr{A}$ 的属于特征根 λ 的**特征向量**.

引理 6.2.15 设 $\mathfrak{L}_1$ 是线性变换 $\mathscr{A}$ 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中一维不变子空间, 则 $\mathfrak{L}_1$ 中任一非零向量 α 为特征向量, 且唯一存在数 $\lambda\in\mathbb{F}$, 使得

$$\mathscr{A}(\alpha)=\lambda\alpha. \tag{6.2.19}$$

再对 $\mathfrak{L}_1$ 中任一向量 β, 仍有

$$\mathscr{A}(\beta)=\lambda\beta. \tag{6.2.20}$$

所以 $\mathfrak{L}_1$ 中所有非零向量都是线性变换 $\mathscr{A}$ 的属于特征根 λ 的特征向量.

证 任取 $\beta\in\mathfrak{L}_1$, 由 $\dim(\mathfrak{L}_1)=1$ 可知 $\beta=c\alpha$, 其中 $c\in\mathbb{F}$. 所以 $\mathscr{A}(\beta)=\mathscr{A}(c\alpha)=c\mathscr{A}(\alpha)=c\lambda\alpha=\lambda\beta$. 因此, $\mathfrak{L}_1$ 中任一非零向量都是线性变换 $\mathscr{A}$ 的属于特征根 λ 的特征向量. $\square$

显然有

引理 6.2.16 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. $\lambda_0\in\mathbb{C}$ 为 $\mathscr{A}$ 的特征根. $\alpha_0\neq 0$ 为 $\mathfrak{L}^{\mathbb{C}}$ 中线性变换 $\mathscr{A}$ 的属于特征根 λ_0 的特征向量. 在 $\mathfrak{L}$ 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 记线性变换 $\mathscr{A}$ 在这组基下的方阵表示为 A, 特征向量 α_0 在这组基下的坐标表达式为 $n\times 1$ 矩阵 x_0, 则在这组基下, 有坐标表达式

$$Ax_0=\lambda_0x_0. \tag{6.2.21}$$

因此 λ_0 也是方阵表示 A 的特征根, 而 $n\times 1$ 非零矩阵 x_0 是方阵表示 A 的特征向量.

这也给出了在 §4.3 中定义的方阵的特征根和特征向量的几何解释.

由定义可知, 一维不变子空间中所有非零向量都是属于同一特征根的特征向量. 当然, 属于特征根 λ_0 的所有特征向量, 加上零向量, 构成的集合是线性变换 $\mathscr{A}$ 的不变子空间 $\mathscr{M}_{\lambda_0}$. 例如, 取定复数 c 和复方阵 $A=cE_n$, 有线性变换 $y=Ax$. 则线性变换 $y=Ax=cx$ 的不变子空间 $\mathscr{M}_c$ 就是整个空间 $\mathfrak{L}$.

定义 6.2.17 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为 $\mathfrak{L}$ 的线性变换, $\lambda_1,\lambda_2,\cdots,\lambda_s\in\mathbb{C}$ 为线性变换 $\mathscr{A}$ 的所有不同特征根. 线性变换 $\mathscr{A}$ 的所有属于特征根 λ_i 的特征向量, 加上零向量, 构成的集合 $\mathscr{M}_i\subset\mathfrak{L}^{\mathbb{C}}$ 是线性变换 $\mathscr{A}$ 的不变子空间, 称为线性变换 $\mathscr{A}$ 的属于特征根 λ_i 的**特征子空间**.

注意 定义 6.2.17 中的属于特征根 λ_i 的特征子空间 $\mathscr{M}_i$ 的并集 $\bigcup\limits_{i=1}^{s}\mathscr{M}_i$ 不一定是子空间, 它线性生成的子空间是线性变换 $\mathscr{A}$ 的不变子空间.

下面从另一个角度来讨论不变子空间. 即引进一般的循环子空间的概念.

定义 6.2.18 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. 在

$\mathfrak{L}$ 中任取非零向量 α, 则由线性空间 $\mathfrak{L}$ 的子集合 $\{\alpha, \mathscr{A}(\alpha), \mathscr{A}^2(\alpha), \cdots\}$ 线性生成的子空间 $\mathfrak{L}(\alpha)$ 称为线性变换 $\mathscr{A}$ 的由 α 生成的**循环子空间**, α 称为**循环向量**.

定理 6.2.19 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换, 则由非零向量 α 线性生成的 $\mathscr{A}$ 的循环子空间 $\mathfrak{L}(\alpha)$ 必为 $\mathscr{A}$ 的不变子空间. 且是 $\mathscr{A}$ 的包含非零向量 α 的最小不变子空间. 又存在最小正整数 r, 使得 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha)$ 为循环子空间 $\mathfrak{L}(\alpha)$ 的一组基, 称为**循环基**. 又

$$\mathscr{A}^r(\alpha) = c_{r-1}\mathscr{A}^{r-1}(\alpha) + c_{r-2}\mathscr{A}^{r-2}(\alpha) + \cdots + c_1\mathscr{A}(\alpha) + c_0\alpha, \tag{6.2.22}$$

其中 $c_0, c_1, \cdots, c_{r-1} \in \mathbb{F}$.

线性变换 $\mathscr{A}$ 在不变子空间 $\mathfrak{L}(\alpha)$ 上关于循环基的方阵表示为

$$\begin{pmatrix} 0 & & & & & c_0 \\ 1 & 0 & & & & c_1 \\ & 1 & 0 & & & c_2 \\ & & \ddots & \ddots & & \vdots \\ & & & 1 & 0 & c_{r-2} \\ & & & & 1 & c_{r-1} \end{pmatrix}, \tag{6.2.23}$$

称为线性变换 $\mathscr{A}|_{\mathfrak{L}(\alpha)}$ 的**自然标准形**.

证 现在考虑向量序列 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^t(\alpha), \cdots$. 因为 $\alpha \neq 0$, 又 $\dim(\mathfrak{L}) = n < \infty$, 所以存在最小正整数 r, 使得 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha)$ 线性无关. 但是, $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha), \mathscr{A}^r(\alpha)$ 线性相关. 即有式 (6.2.22). 由归纳法不难证明 $\mathscr{A}^i(\alpha)$, $i = 1, 2, \cdots$ 都是 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha)$ 的线性组合. 这证明了 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{r-1}(\alpha)$ 为循环子空间 $\mathfrak{L}(\alpha)$ 的基. 又由循环子空间 $\mathfrak{L}(\alpha)$ 的构造可知循环子空间 $\mathfrak{L}(\alpha)$ 是包含非零向量 α 的最小不变子空间.

现在考虑向量 $\alpha_j = \mathscr{A}^{j-1}(\alpha), j = 1, 2, \cdots, r$. 由式 (6.2.22), 则有

$$\mathscr{A}(\alpha_j) = \alpha_{j+1}, \quad j = 1, 2, \cdots, r-1, \quad \mathscr{A}(\alpha_r) = \mathscr{A}^r(\alpha) = c_0\alpha_1 + c_1\alpha_2 + \cdots + c_{r-1}\alpha_r.$$

所以 $\mathscr{A}$ 的方阵表示为式 (6.2.23). □

由定理 6.2.9, 显然有

引理 6.2.20 设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的幂零线性变换, 即存在最小正整数 m, 使得 $\mathscr{A}^m = 0$. 又线性变换 $\mathscr{A}$ 的由非零向量 α 生成的循环子空间 $\mathfrak{L}(\alpha)$ 有循环基 $\alpha, \mathscr{A}(\alpha), \cdots, \mathscr{A}^{q-1}(\alpha)$, 且 $\mathscr{A}^q(\alpha) = 0$, 其中 $q \leqslant m$.

习 题 6.2

6.2.1 n 维线性空间 $\mathfrak{L}$ 为它的线性变换 $\mathscr{A}$ 的循环空间 (即在 $\mathscr{A}$ 中存在向量 α, 使得 α,

$\mathscr{A}(\alpha), \cdots, \mathscr{A}^{n-1}(\alpha)$ 为 $\mathfrak{L}$ 的一组基) 当且仅当对 $\mathscr{A}$ 的任一特征根决定的特征向量全体线性生成一维子空间; 当且仅当 $\mathscr{A}$ 的特征多项式等于 $\mathscr{A}$ 的极小多项式.

6.2.2 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的一组基, 且它们都是线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 的特征向量, 又 $\mathscr{A}$ 的 n 个特征根都不相同. 试证: $\mathfrak{L}$ 为线性变换 $\mathscr{A}$ 的循环空间, 试找出这个循环空间的循环向量.

6.2.3 试给出一个例子说明 n 阶方阵 A 和 B 有: AB 和 BA 的特征多项式相同, 但是极小多项式不相同.

6.2.4 试证: n 阶幂等方阵 A (即适合 $A^2 = A$) 相似于对角形 $\operatorname{diag}(E_r, 0)$, 其中 $r = \operatorname{rank}(A)$. 用这一结果证明: n 阶幂等方阵 A 有 $\operatorname{tr}(A) = \operatorname{rank}(A)$.

6.2.5 试证: 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 和 $\mathscr{B}$ 若有 $\mathscr{A}^2 = \mathrm{id}$, $\mathscr{B}^2 = \mathrm{id}$, $\mathscr{A} \circ \mathscr{B} = \mathscr{B} \circ \mathscr{A}$. 则在 $\mathfrak{L}$ 中存在一组基, 使得 $\mathscr{A}$ 和 $\mathscr{B}$ 的方阵表示同时为对角方阵, 且对角元素为 1 或 -1.

6.2.6 试证: n 维线性空间 $\mathfrak{L}$ 上的线性变换 $\mathscr{A}$ 有 n 个线性无关的特征向量的必要且充分条件是存在一组基, 使得在这组基下线性变换 $\mathscr{A}$ 的方阵表示为对角形.

6.2.7 试证: **(1)** n 维复线性空间 $\mathfrak{L}$ 的两个两两可交换的线性变换有公共特征向量;

(2) n 维复线性空间 $\mathfrak{L}$ 上一组 (不一定可数多个) 两两可交换的线性变换有公共特征向量;

(3) 用**(2)** 证明: 一组两两可交换的 n 阶复方阵能用一个公共的 n 阶非异复方阵, 将它们同时相似于上三角方阵;

(4) 特别, 当**(3)** 中给出的那组两两可交换的 n 阶复方阵都相似于对角方阵时, 它们同时也都相似于对角方阵.

6.2.8 设 A 和 B 都是域 $\mathbb{F}$ 上的 n 阶方阵, 且 B 幂零, 又 $AB = BA$. 试证: $\det(\lambda E_n - (A+B)) = \det(\lambda E_n - A)$.

6.2.9 设 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中关于线性变换 $\mathscr{A}$ 的不变子空间, 什么时候 $\mathscr{A}(\mathfrak{L}_1 \cap \mathfrak{L}_2) = \mathscr{A}(\mathfrak{L}_1) \cap \mathscr{A}(\mathfrak{L}_2)$?

6.2.10 设 $\mathscr{A}$ 和 $\mathscr{B}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. 试证: $\dim(\ker(AB)) \leqslant \dim(\ker(A)) + \dim(\ker(B))$.

6.2.11 设 $\mathscr{A}$ 为 n 维复线性空间 $\mathfrak{L}$ 的线性变换. 试证: 在 $\mathfrak{L}$ 中存在一组基, 使得 $\mathscr{A}$ 的方阵表示为上三角方阵. 因此对任一正整数 r, $1 \leqslant r \leqslant n$, 则在复线性空间 $\mathfrak{L}$ 中存在一个 r 维不变子空间.

§6.3 λ 矩阵在相抵下的标准形

元素都是自变量 λ 的域 $\mathbb{F}$ 上的多项式的 $n \times m$ 矩阵称为域 $\mathbb{F}$ 上 $n \times m$ λ **矩阵**. 例如

$$A(\lambda) = \begin{pmatrix} \lambda^2+1 & 2\lambda+i & 2\lambda+3 \\ \lambda^2 & 1 & i\lambda^2+2\lambda \\ 3+4i\lambda & i+\lambda+i\lambda^2 & \sqrt{2}\lambda+1 \end{pmatrix}$$

就是一个 3×3 复 λ 矩阵, 其中 $i = \sqrt{-1}$.

因为 λ 矩阵实质上是以域 $\mathbb{F}$ 上矩阵为系数的 λ 的多项式, 例如上面的例子为

$$A(\lambda)=\lambda^2\begin{pmatrix}1&0&0\\1&0&i\\0&i&0\end{pmatrix}+\lambda\begin{pmatrix}0&2&2\\0&0&2\\4i&1&\sqrt{2}\end{pmatrix}+\begin{pmatrix}1&i&3\\0&1&0\\3&i&1\end{pmatrix},$$

又因为通常的域 $\mathbb{F}$ 中元为元素的矩阵的代数性质都是通过对矩阵的元素作加、减、乘、除等计算而得到的, 所以这些代数性质中的一大部分, 它们和除法的关系较少, 这些代数性质都可以推广到 λ 矩阵的情形. 具体地说, 矩阵的三种代数运算: 加、乘、纯量积, 以及方阵的行列式和它的性质, Laplace 展开, 伴随方阵, Binet–Cauchy 公式, 矩阵的秩, 满秩 (非异) 方阵, 可逆方阵, 以及矩阵被右乘或被左乘以可逆方阵后秩的不变性, 矩阵的初等变换, 相抵等等概念和性质, 都可以推广到 λ 矩阵的情形. 然而, 在 λ 矩阵的情形, 相对于在常数矩阵的情形, 有一个原则的差别. 那就是在常数方阵的情形中, 满秩, 非异和可逆这三个概念是等价的, 但是在 λ 方阵的情形, 满秩和非异是等价的, 而满秩 (非异) 和可逆是不等价的. 我们有

定义 6.3.1 域 $\mathbb{F}$ 上 n 阶 λ 方阵 $A(\lambda)$ 称为**可逆 λ 方阵**, 如果存在另外一个 n 阶 λ 方阵 $B(\lambda)$, 使得

$$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=E_n, \tag{6.3.1}$$

其中 E_n 为 n 阶单位方阵.

由定义可知, 域 $\mathbb{F}$ 上可逆 λ 方阵的行列式是 λ 的非零多项式, 所以是满秩的; 反之不然. 事实上, 有

定理 6.3.2 域 $\mathbb{F}$ 上 n 阶 λ 方阵 $A=A(\lambda)$ 可逆的必要且充分条件为 λ 方阵 $A(\lambda)$ 的行列式 $\det(A(\lambda))$ 的值为 $\mathbb{F}$ 中的非零数.

证 先证必要性. 今 $A(\lambda)$ 可逆, 由定义, 存在 λ 方阵 $B(\lambda)$, 使得 $A(\lambda)B(\lambda)=E_n$. 双方取行列式, 即有

$$\det(A(\lambda))\det(B(\lambda))=1,$$

而 $\det(A(\lambda))$ 及 $\det(B(\lambda))$ 都是 λ 的多项式. 因为多项式乘积的次数等于每个因式的次数的和, 所以 $\det(A(\lambda))$ 是零次多项式, 即为域 $\mathbb{F}$ 中非零数.

再证充分性. 设 λ 方阵 $A(\lambda)$ 的行列式 $\det(A(\lambda))$ 为域 $\mathbb{F}$ 中一非零常数. 作 λ 方阵 $A(\lambda)$ 的伴随方阵 $A(\lambda)^{(*)}$, 其中 $A(\lambda)^{(*)}$ 的第 j 行, 第 k 列元素是 $A(\lambda)$ 的第 k 行, 第 j 列元素的代数余子式, 所以仍为 λ 的多项式. 因此 $A(\lambda)^{(*)}$ 仍为 λ 方阵, 故

$$B(\lambda)=(\det(A(\lambda)))^{-1}A(\lambda)^{(*)} \tag{6.3.2}$$

也是 λ 方阵. 显然 $A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=E_n$, 即 $A(\lambda)$ 可逆. □

在这里引进域 $\mathbb{F}$ 上 $n\times m$ λ 矩阵的目的, 主要是为了研究域 $\mathbb{F}$ 上常数矩阵在某

些等价关系下的标准形理论. 在第六, 七, 十三和十四这四章中, 将要充分运用域 $\mathbb{F}$ 上 λ 矩阵以及 λ 矩阵在相抵下的标准形理论这个工具.

下面利用 §3.4 的全部思想和技巧. 在这一节, 对域 $\mathbb{F}$ 上 $n\times m$ λ 矩阵引进三类**初等变换**和三类初等 λ 方阵的概念, 其中**第一类初等 λ 方阵**和**第二类初等 λ 方阵**就是原来的第一类和第二类初等常数方阵 P_{jk} (见式 (3.4.1)) 和 $Q_j(a)$ (见式 (3.4.2)), $\forall a\in\mathbb{F}$, $a\neq 0$. 第一类初等 λ 方阵的乘积仍称为**置换方阵**. 第三类初等 λ 方阵定义为

定义 6.3.3 域 $\mathbb{F}$ 上 n 阶 λ 方阵

$$Q_{jk}(f(\lambda))=E_n+f(\lambda)E_{jk},\quad f(\lambda)\in\mathbb{F}[\lambda],\quad j\neq k,\quad 1\leqslant j,k\leqslant n \tag{6.3.3}$$

称为**第三类初等 λ 方阵**, 其中 E_{jk} 是第 j 行, 第 k 列位置元素是 1, 其余位置元素是零的 n 阶方阵.

将初等 λ 方阵左 (右) 乘以 λ 矩阵 $A(\lambda)$ 后, 引起 λ 矩阵 $A(\lambda)$ 的变化, 称为 λ 矩阵 $A(\lambda)$ 的**行 (列) 的初等变换**, 统称为**初等变换**. 为了确定起见, 用乘以第一, 二, 三类初等 λ 方阵所引起的初等变换, 分别称为**第一类初等变换**, **第二类初等变换**, **第三类初等变换**. 第一类初等变换的连续作用称为**置换变换**.

由直接计算可知

$$P_{jk}^{-1}=P_{jk},\quad P_j(a)^{-1}=P_j(a^{-1}),\quad Q_{jk}(f(\lambda))^{-1}=Q_{jk}(-f(\lambda)). \tag{6.3.4}$$

所以初等 λ 方阵都是可逆 λ 方阵, 且它们的逆方阵是同类型的初等 λ 方阵. 另外, 和 §3.4 相同, 不难证明, 对 λ 矩阵施行任一初等变换后, 秩并不改变.

定理 6.3.4 对域 $\mathbb{F}$ 上任一 λ 矩阵

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1m}(\lambda)\\ \vdots & & \vdots\\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda)\end{pmatrix}\neq 0, \tag{6.3.5}$$

则存在一系列初等变换, 它们将 $A(\lambda)$ 化为标准形

$$\begin{pmatrix} \operatorname{diag}(d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)) & 0^{(r,m-r)}\\ 0^{(n-r,r)} & 0^{(n-r,m-r)}\end{pmatrix}, \tag{6.3.6}$$

其中 r 是矩阵 $A(\lambda)$ 的秩, 又 $d_1(\lambda)$, $d_2(\lambda)$, $\cdots$, $d_r(\lambda)$ 都是首项系数为 1 的非零多项式, 它们依次地一个除得尽一个. 即有

$$d_1(\lambda)\mid d_2(\lambda),\quad d_2(\lambda)\mid d_3(\lambda),\quad \cdots,\quad d_{r-1}(\lambda)\mid d_r(\lambda). \tag{6.3.7}$$

证 令 λ 矩阵 $A(\lambda)\neq 0$. 所以在 $A(\lambda)$ 中必有一个元素 $a_{jk}(\lambda)\neq 0$, 因此, 对 λ 矩阵 $A(\lambda)$, 记数值

$$d=\min_{a_{ij}(\lambda)\neq 0,\ 1\leqslant i\leqslant n,\ 1\leqslant j\leqslant m}(\deg(a_{ij}(\lambda))).$$

则 d 为非负整数.

设 $\deg(a_{pq}(\lambda))=d$. 记 $a_{pq}(\lambda)$ 的首项系数为 a, 因此 λ 矩阵 $P_1(a^{-1})P_{1p}A(\lambda)P_{1q}$ 的第一行, 第一列位置元素为首项系数等于 1 的非零多项式, 而且它的次数最小.

上面讨论告诉我们, 总可以将次数为 d 的元素换到第一行, 第一列的位置上. 而且无妨设这个元素的首项系数等于 1.

我们对数值 d 作归纳法, 来证明下面事实: 对 λ 矩阵 $A(\lambda)$ 施行一系列初等变换后, 可以将 $A(\lambda)$ 化为 $n\times m$ λ 矩阵

$$B(\lambda)=\begin{pmatrix} b_1(\lambda) & 0\\ 0 & B_1(\lambda)\end{pmatrix}=\begin{pmatrix} b_1(\lambda) & 0\\ 0 & \begin{pmatrix} b_{22}(\lambda) & \cdots & b_{2m}(\lambda)\\ \vdots & & \vdots\\ b_{n2}(\lambda) & \cdots & b_{nm}(\lambda)\end{pmatrix}\end{pmatrix}, \tag{6.3.8}$$

其中 $b_1(\lambda)$ 为首项系数等于 1 的非零多项式, 又 $b_1(\lambda)$ 除得尽所有多项式 $b_{jk}(\lambda)$, $2\leqslant j\leqslant n$, $2\leqslant k\leqslant m$.

为了证明这一事实, 先从 $d=0$, 即是零次多项式时, 来证明这一事实.

无妨设 $a_{11}(\lambda)=1$. 令

$$P_1=\begin{pmatrix} 1 & & & \\ -a_{21}(\lambda) & 1 & & \\ -a_{31}(\lambda) & & 1 & \\ \vdots & & & \ddots & \\ -a_{n1}(\lambda) & & & & 1\end{pmatrix},\quad Q_1=\begin{pmatrix} 1 & -a_{12}(\lambda) & -a_{13}(\lambda) & \cdots & -a_{1m}(\lambda)\\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix}.$$

即

$$P_1=Q_{21}(-a_{21}(\lambda))Q_{31}(-a_{31}(\lambda))\cdots Q_{n1}(-a_{n1}(\lambda)),$$
$$Q_1=Q_{12}(-a_{12}(\lambda))Q_{13}(-a_{13}(\lambda))\cdots Q_{1m}(-a_{1m}(\lambda)).$$

由直接计算可知 $P_1A(\lambda)Q_1=B(\lambda)$, 其中 $a_{11}(\lambda)=1$. 所以式 (6.3.8) 中 $b_1(\lambda)=1$. 因此 $b_1(\lambda)\mid b_{jk}(\lambda)$, $2\leqslant j\leqslant n$, $2\leqslant k\leqslant m$, 即断言成立.

假设对任一 λ 矩阵 $A(\lambda)$, 只要数值 $d<N$ 时断言成立. 下面来证明当数值 $d=N$ 时断言也成立.

设 $\deg(a_{11}(\lambda))=d=N$, 而且无妨设 $a_{11}(\lambda)$ 的首项系数等于 1. 下面分三种情形讨论.

(1) 设在 λ 矩阵 $A(\lambda)$ 的第一列中有一个元素 $a_{j1}(\lambda)$, $j \geqslant 2$, 不能被 $a_{11}(\lambda)$ 除尽. 设

$$a_{j1}(\lambda) = q(\lambda)a_{11}(\lambda) + d(\lambda),$$

其中 $q(\lambda)$ 和 $d(\lambda)$ 都是域 $\mathbb{F}$ 上多项式. 且 $d(\lambda)$ 是次数小于 N 的非零多项式, 首项系数等于 a. 令

$$A_1(\lambda) = P_1(a^{-1})Q_{1j}(-q(\lambda))P_{1j}A(\lambda) = \begin{pmatrix} f(\lambda) & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

的第一行, 第一列元素是次数小于 N, 首项系数等于 1 的非零多项式 $f(\lambda)$. 对矩阵 $A_1(\lambda)$ 用归纳法假设, 故这时断言成立.

(2) 设在 λ 矩阵 $A(\lambda)$ 的第一行中有一个元素 $a_{1j}(\lambda)$, $j \geqslant 2$, 不能被 $a_{11}(\lambda)$ 除尽. 和**(1)** 一样讨论也可以证明断言成立.

(3) 设 λ 矩阵 $A(\lambda)$ 的第一行和第一列的所有元素都能被 $a_{11}(\lambda)$ 除尽. 令 $a_{1k}(\lambda) = q_{1k}(\lambda)a_{11}(\lambda)$, $1 < k \leqslant m$, $a_{j1}(\lambda) = q_{j1}(\lambda)a_{11}(\lambda)$, $1 < j \leqslant n$. 令

$$P_2 = Q_{21}(-q_{21}(\lambda))Q_{31}(-q_{31}(\lambda))\cdots Q_{n1}(-q_{n1}(\lambda)),$$
$$Q_2 = Q_{12}(-q_{12}(\lambda))Q_{13}(-q_{13}(\lambda))\cdots Q_{1m}(-q_{1m}(\lambda)).$$

由直接计算可知

$$P_2A(\lambda)Q_2 = \widetilde{B}(\lambda) == \begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & \widetilde{b}_{22}(\lambda) & \cdots & \widetilde{b}_{2m}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & \widetilde{b}_{n2}(\lambda) & \cdots & \widetilde{b}_{nm}(\lambda) \end{pmatrix}.$$

今若 $a_{11}(\lambda)$ 除得尽所有元素 $\widetilde{b}_{jk}(\lambda)$, $2 \leqslant j \leqslant n$, $2 \leqslant k \leqslant m$, 那末这就是式 (6.3.8), 故断言成立. 如果 $a_{11}(\lambda)$ 除不尽其中一个元素 $\widetilde{b}_{jk}(\lambda)$, 再作初等变换.

$$A_1(\lambda) = Q_{k1}(1)\widetilde{B}(\lambda) == \begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ \widetilde{b}_{2k}(\lambda) & \widetilde{b}_{22}(\lambda) & \cdots & \widetilde{b}_{2m}(\lambda) \\ \vdots & \vdots & & \vdots \\ \widetilde{b}_{nk}(\lambda) & \widetilde{b}_{n2}(\lambda) & \cdots & \widetilde{b}_{nm}(\lambda) \end{pmatrix}.$$

这就化为情形**(1)**, 所以断言成立.

到现在为止, 经过一系列初等变换, 将 λ 矩阵 $A(\lambda)$ 化为 λ 矩阵 $B(\lambda)$ 如式 (6.3.8). 显然对 $(n-1)\times(m-1)$ λ 矩阵 $B_1(\lambda)$ 作的初等变换可以看作是对 λ 矩阵 $A(\lambda)$ 作的初等变换, 而且将 λ 矩阵 $B_1(\lambda)$ 作了初等变换后, 每个元素仍能被多项式 $b_1(\lambda)$ 除得尽. 所以, 对 λ 矩阵 $B_1(\lambda)$ 可以进行同样的讨论. 这样继续下去, 因此对 λ

矩阵 $A(\lambda)$ 作一系列初等变换后, 可以将 λ 矩阵 $A(\lambda)$ 化为标准形 (6.3.6), 其中 r 是矩阵 $A(\lambda)$ 的秩. 又 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 都是首项系数为 1 的非零多项式, 它们依次地一个除尽一个. □

由上面定理, 可以得到下面的重要推论.

定理 6.3.5 域 $\mathbb{F}$ 上任一 n 阶可逆 λ 方阵是一些初等 λ 方阵的乘积.

证 给定 n 阶可逆 λ 方阵 $A(\lambda)$, 则 λ 矩阵 $A(\lambda)$ 的秩等于 n. 由定理 6.3.4 可知, 存在 n 阶初等 λ 方阵 $P_1(\lambda), P_2(\lambda), \cdots, P_s(\lambda)$ 和 $Q_1(\lambda), Q_2(\lambda), \cdots, Q_t(\lambda)$, 使得

$$\begin{aligned}A_0(\lambda) &= P_s(\lambda)P_{s-1}(\lambda)\cdots P_1(\lambda)A(\lambda)Q_1(\lambda)Q_2(\lambda)\cdots Q_t(\lambda)\\ &= \operatorname{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)),\end{aligned}$$

其中 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 都是首项系数为 1 的非零多项式, 它们依次地一个除尽一个. 因为可逆 λ 方阵的乘积仍为可逆 λ 方阵, 所以 $A_0(\lambda)$ 仍为可逆 λ 方阵. 由定理 6.3.4, $\det(A_0(\lambda)) = d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda) \in \mathbb{F}$, 所以 $d_1(\lambda) = d_2(\lambda) = \cdots = d_r(\lambda) = 1$, 即

$$P_s(\lambda)P_{s-1}(\lambda)\cdots P_1(\lambda)A(\lambda)Q_1(\lambda)Q_2(\lambda)\cdots Q_t(\lambda) = E_n,$$

所以

$$A(\lambda) = P_1(\lambda)^{-1}P_2(\lambda)^{-1}\cdots P_s(\lambda)^{-1}Q_t(\lambda)^{-1}Q_{t-1}(\lambda)^{-1}\cdots Q_1(\lambda)^{-1}.$$

由于初等方阵的逆方阵仍为初等方阵, 这证明了定理. □

根据这个定理, 和 §3.4 一样, 也可以利用 λ 矩阵的初等变换给出一种计算 λ 矩阵的秩的方法; 也可以给出一种计算可逆 λ 方阵的逆方阵的方法. 因为叙述和证明和 §3.4 一样, 所以不再详细叙述, 读者可以自己补出来.

定义 6.3.6 两个域 $\mathbb{F}$ 上 $n \times m$ λ 矩阵 $A(\lambda)$ 和 $B(\lambda)$ 称为**相抵的**, 如果存在 n 阶可逆 λ 方阵 $P(\lambda)$ 和 m 阶可逆 λ 方阵 $Q(\lambda)$, 使得

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda). \tag{6.3.9}$$

显然, λ 矩阵的相抵是一种等价关系. 在 λ 矩阵都是常数矩阵时, 这就是通常的相抵的概念. 在 §3.4 中, 我们是利用行和列的初等变换来求标准形和全系不变量的. 在这一节, 我们来求 λ 矩阵在相抵下的标准形和全系不变量. 按照定义 6.3.6, 显然, 定理 6.3.4 又可以改写为

定理 6.3.7 域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 相抵于标准形

$$\begin{pmatrix}\operatorname{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)) & 0^{(r,m-r)}\\ 0^{(n-r,r)} & 0^{(n-r,m-r)}\end{pmatrix}, \tag{6.3.10}$$

其中 r 是 λ 矩阵 $A(\lambda)$ 的秩. 又 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 都是首项系数为 1 的非零多项式, 它们依次地一个除得尽一个. $d_j(\lambda)$ 称为 λ 矩阵 $A(\lambda)$ 的第 j 个**不变因式**, $1 \leqslant j \leqslant r$. 它们全体组成 λ 矩阵 $A(\lambda)$ 的**不变因式组**.

这个定理告诉我们, λ 矩阵在相抵关系下的等价类中可以选取形如式 (6.3.10) 的 λ 矩阵为代表元素. 然而并没有找出代表元素集, 因此并没有解决分类问题. 为了解决标准形理论, 需要求 λ 矩阵在相抵下的全系不变量. 为此, 我们引进下面概念.

定义 6.3.8 设域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 的秩为 r. 任取正整数 j, $1 \leqslant j \leqslant r$, λ 矩阵 $A(\lambda)$ 的 j 阶**行列式因式** $D_j(\lambda)$ 是 $A(\lambda)$ 的所有 j 阶子式的最大公因式. 即 $D_j(\lambda)$ 是一个首项系数为 1 的 λ 的多项式, 它除得尽 $A(\lambda)$ 的所有 j 阶子式, 且任一除得尽 $A(\lambda)$ 的所有 j 阶子式的非零多项式必除得尽 $D_j(\lambda)$. $D_j(\lambda)$ 称为 λ 矩阵 $A(\lambda)$ 的第 j 个行列式因式, $1 \leqslant j \leqslant r$. 它们全体组成 λ 矩阵 $A(\lambda)$ 的**行列式因式组**.

显然, 设域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 的秩为 r, 那末它的行列式因式 $D_1(\lambda), \cdots, D_r(\lambda)$ 都是首项系数为 1 的非零多项式. 又由于 λ 矩阵 $A(\lambda)$ 的任一 j 阶子式按照第一行作 Laplace 展开后得到 j 个项的和, 每个项都有一个因式是 λ 矩阵 $A(\lambda)$ 的 $j-1$ 阶子式. 所以 $D_{j-1}(\lambda)$ 除得尽每个因式, 因此 $D_{j-1}(\lambda)$ 除得尽任一 j 阶子式. 这就证明了 $D_{j-1}(\lambda)$ 除得尽 $D_j(\lambda)$. 即 $D_1(\lambda), D_2(\lambda), \cdots, D_r(\lambda)$ 是依次地一个除得尽一个的多项式组.

特别, 考虑域 $\mathbb{F}$ 上 λ 矩阵

$$C(\lambda) = \begin{pmatrix} \operatorname{diag}(d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)) & 0^{(r,m-r)} \\ 0^{(n-r,r)} & 0^{(n-r,m-r)} \end{pmatrix}, \tag{6.3.11}$$

其中 r 是 λ 矩阵 $C(\lambda)$ 的秩, 又 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 都是首项系数为 1 的非零多项式, 它们依次地一个除尽一个. 显然, $C(\lambda)$ 的 t 阶行列式因式

$$D_t(\lambda) = d_1(\lambda)d_2(\lambda)\cdots d_t(\lambda), \quad 1 \leqslant t \leqslant r. \tag{6.3.12}$$

因此

$$d_t(\lambda) = \frac{D_t(\lambda)}{D_{t-1}(\lambda)}, \quad D_0(\lambda) = 1, \quad 1 \leqslant t \leqslant r. \tag{6.3.13}$$

由式 (6.3.12) 和 (6.3.13) 可知, λ 矩阵 $A(\lambda)$ 的不变因式组由 λ 矩阵 $A(\lambda)$ 的行列式因式组唯一决定, 反之亦然.

显然, 为了证明 λ 矩阵的不同的标准形互不相抵, 只要证明 λ 矩阵的行列式因式组在相抵下不变就可以了. 如果这点证明了, 那么因为标准形的不变因式组由行列式因式组唯一决定, 反之亦然. 所以不同的标准形有不同的不变因式组, 因此不同的标准形互不相抵. 换句话说, 矩阵的不变因式组构成它在相抵下的全系不变量. 因

此 λ 矩阵的行列式因式组也构成它在相抵下的全系不变量.

定理 6.3.9 域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 的 t 阶行列式因式在相抵下不变, $1 \leqslant t \leqslant r$, 其中 $r = \operatorname{rank}(A(\lambda))$. 因此, 域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 在相抵下的全系不变量为 λ 矩阵 $A(\lambda)$ 的行列式因式组. 这等价于全系不变量为不变因式组.

证 由定理 6.3.5 可知, λ 矩阵相抵是作用一系列初等变换. 所以下面分别就三类初等变换来证明 λ 矩阵的 t 阶行列式因式在初等变换下不变. 由于对行和列的初等变换, 证明是相仿的, 所以下面只对行的初等变换来证明就可以了.

对 λ 矩阵 $A(\lambda)$ 的行按照第一类初等方阵 P_{jk} 作第一类初等变换, 得到 λ 矩阵 $P_{jk}A(\lambda)$. 它是将 λ 矩阵 $A(\lambda)$ 的第 j 行和第 k 行互换, 而其余行不动. 所以 $P_{jk}A(\lambda)$ 的任一 t 阶子式或者是 $A(\lambda)$ 的 t 阶子式, 或者是 $A(\lambda)$ 的 t 阶子式乘以 -1. 即 $A(\lambda)$ 的所有 t 阶子式加上适当的正负号便是 $P_{jk}A(\lambda)$ 的所有 t 阶子式, 所以它们的最大公因式相等. 因此 λ 矩阵 $A(\lambda)$ 和 $P_{jk}A(\lambda)$ 的 t 阶行列式因式相等.

再对 λ 矩阵 $A(\lambda)$ 的行按照第二类初等方阵 $P_k(a)$ 作第二类初等变换, 其中 $0 \neq a \in \mathbb{F}$, 得到 λ 矩阵 $P_k(a)A(\lambda)$. 它是将 λ 矩阵 $A(\lambda)$ 的第 k 行遍乘以非零数 a, 而其余行不变. 所以 $P_k(a)A(\lambda)$ 的任一 t 阶子式, 或者是 $A(\lambda)$ 的 t 阶子式, 或者是 $A(\lambda)$ 的 t 阶子式乘以非零数 a. 由于 t 阶行列式因式的首项系数为 1, 所以 λ 矩阵 $A(\lambda)$ 和 $P_k(a)A(\lambda)$ 的 t 阶子式的最大公因式相等. 因此 λ 矩阵 $A(\lambda)$ 和 $P_k(a)A(\lambda)$ 的 t 阶行列式因式相等.

最后, 对 λ 矩阵 $A(\lambda)$ 的行按照第三类初等方阵 $Q_{jk}(f(z))$ 作第三类初等变换, 得到 λ 矩阵 $B(\lambda) = Q_{jk}(f(z))A(\lambda)$. 为了讨论简单起见, 取 $j = 1$, $k = 2$, 则

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & \cdots & a_{1m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix},$$

$$B(\lambda) = \begin{pmatrix} a_{11}(\lambda) + f(x)a_{21}(\lambda) & \cdots & a_{1n}(\lambda) + f(x)a_{2m}(\lambda) \\ a_{21}(\lambda) & \cdots & a_{2m}(\lambda) \\ \vdots & & \vdots \\ a_{n1}(\lambda) & \cdots & a_{nm}(\lambda) \end{pmatrix}.$$

记 λ 矩阵 $A(\lambda)$ 和 $B(\lambda)$ 的 t 阶行列式因式分别为 $D_t(\lambda)$ 和 $\widetilde{D}_t(\lambda)$. 任一相同位置的 t 阶子矩阵分别为 $A(\lambda)\begin{pmatrix} i_1\ i_2 \cdots i_t \\ j_1\ j_2 \cdots j_t \end{pmatrix}$, $B(\lambda)\begin{pmatrix} i_1\ i_2 \cdots i_t \\ j_1\ j_2 \cdots j_t \end{pmatrix}$, $1 \leqslant i_1 < i_2 < \cdots < i_t \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_t \leqslant m$. 因此有 λ 矩阵 $A(\lambda)$ 和 $B(\lambda)$ 的 t 阶子式间的如下关系:

(1) 设 $i_1 > 1$, 则

$$\det B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_t \\ j_1 & j_2 & \cdots & j_t \end{pmatrix} = \det A(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_t \\ j_1 & j_2 & \cdots & j_t \end{pmatrix};$$

(2) 设 $i_1 = 1, i_2 = 2$, 则

$$\det B(\lambda)\begin{pmatrix} 1 & 2 & i_3 & \cdots & i_t \\ j_1 & j_2 & j_3 & \cdots & j_t \end{pmatrix} = \det A(\lambda)\begin{pmatrix} 1 & 2 & i_3 & \cdots & i_t \\ j_1 & j_2 & j_3 & \cdots & j_t \end{pmatrix};$$

(3) 设 $i_1 = 1, i_2 > 2$, 则 $\det B(\lambda)\begin{pmatrix} 1 & i_2 & \cdots & i_t \\ j_1 & j_2 & \cdots & j_t \end{pmatrix}$ 等于

$$\det A(\lambda)\begin{pmatrix} 1 & i_2 & \cdots & i_t \\ j_1 & j_2 & \cdots & j_t \end{pmatrix} + f(x)\det A(\lambda)\begin{pmatrix} 2 & i_2 & \cdots & i_t \\ j_1 & j_2 & \cdots & j_t \end{pmatrix}.$$

总之, 由于 λ 矩阵 $A(\lambda)$ 的 t 阶行列式因式 $D_t(\lambda)$ 除得尽 $A(\lambda)$ 的所有 t 阶子式, 所以除得尽 λ 矩阵 $B(\lambda)$ 的所有 t 阶子式. 因此除得尽 λ 矩阵 $B(\lambda)$ 的 t 阶行列式因式. 即有 $D_t(\lambda) \mid \widetilde{D}_t(\lambda)$. 然而 $A(\lambda) = Q_{12}(-f(x))B(\lambda)$, 所以同样的讨论可证 $\widetilde{D}_t(\lambda)$ 除得尽 $D_t(\lambda)$. 今 $D_t(\lambda)$ 和 $\widetilde{D}_t(\lambda)$ 都是首项系数为 1 的多项式, 它们还互相可以除得尽, 所以 $D_t(\lambda) = \widetilde{D}_t(\lambda)$. 因此在 $B(\lambda) = Q_{12}(f(x))A(\lambda)$ 时证明了定理. 同理在 $B(\lambda) = Q_{jk}(f(x))A(\lambda)$ 时也证明了定理. □

至此, 我们给出了 λ 矩阵在相抵下的标准形理论. 这里要注意, 上面的讨论, 给出了求 λ 矩阵 $A(\lambda)$ 在相抵下的全系不变量的两种方法.

(一) 利用 λ 矩阵 $A(\lambda)$ 的元素的特性, 来求它的行列式因式组, 从而求出 λ 矩阵 $A(\lambda)$ 的不变因式组;

(二) 利用 λ 矩阵的初等变换将 λ 矩阵 $A(\lambda)$ 化为对角 λ 矩阵, 再将对角元素分别分解因式. 这种方法比之直接计算矩阵的行列式因式, 进而计算不变因式和初等因式, 往往有效些.

例 6.3.1 n 阶 λ 方阵

$$A = \begin{pmatrix} \lambda & -1 & & & \\ 0 & \lambda & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & -1 \\ & & & & \lambda \end{pmatrix} = \lambda E_n - N_n$$

有行列式为非零常数的子矩阵

$$A\begin{pmatrix}1\ 2 \cdots\ n-1\\ 2\ 3 \cdots\ \ n\end{pmatrix}=\begin{pmatrix}-1 & & & \\ \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \lambda & -1\end{pmatrix}=\lambda N'_{n-1}-E_{n-1}.$$

而 $\det(\lambda N'_{n-1}-E_{n-1})=(-1)^{n-1}$. 所以它的 n 个行列式因式分别为 $D_1(\lambda)=D_2(\lambda)=\cdots=D_{n-1}(\lambda)=1$, $D_n(\lambda)=\lambda^n$. 因此它的 n 个不变因式分别为 $d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=1$, $d_n(\lambda)=\lambda^n$.

例 6.3.2 求 λ 矩阵

$$A(\lambda)=\begin{pmatrix}\lambda+4 & -2 & -10\\ 4 & \lambda-3 & -7\\ 3 & -1 & \lambda-7\end{pmatrix}$$

的秩、行列式因式组和不变因式组.

解 作初等变换

$$A(\lambda)=\begin{pmatrix}\lambda+4 & -2 & -10\\ 4 & \lambda-3 & -7\\ 3 & -1 & \lambda-7\end{pmatrix}\xrightarrow{P_{12}}\begin{pmatrix}4 & \lambda-3 & -7\\ \lambda+4 & -2 & -10\\ 3 & -1 & \lambda-7\end{pmatrix}$$

$$\xrightarrow{Q_{21}(-\frac{\lambda}{4}-1)}\begin{pmatrix}4 & \lambda-3 & -7\\ 0 & -\dfrac{\lambda^2}{4}-\dfrac{\lambda}{4}+1 & \dfrac{7\lambda}{4}-3\\ 3 & -1 & \lambda-7\end{pmatrix}\xrightarrow{Q_{31}(-\frac{3}{4})}\begin{pmatrix}4 & \lambda-3 & -7\\ 0 & -\dfrac{\lambda^2}{4}-\dfrac{\lambda}{4}+1 & \dfrac{7\lambda}{4}-3\\ 0 & -\dfrac{3\lambda}{4}+\dfrac{5}{4} & \lambda-\dfrac{7}{4}\end{pmatrix}$$

$$\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda^2+\lambda-4 & 7\lambda-12\\ 0 & 3\lambda-5 & 4\lambda-7\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 7\lambda-12 & \lambda^2+\lambda-4\\ 0 & 4\lambda-7 & 3\lambda-5\end{pmatrix}$$

$$\xrightarrow{Q_{23}(-\frac{7}{4})}\begin{pmatrix}1 & 0 & 0\\ 0 & \dfrac{1}{4} & \lambda^2-\dfrac{17\lambda}{4}+\dfrac{19}{4}\\ 0 & 4\lambda-7 & 3\lambda-5\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 4\lambda^2-17\lambda+19\\ 0 & 4\lambda-7 & 3\lambda-5\end{pmatrix}$$

$$\xrightarrow{Q_{32}(7-4\lambda)}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 4\lambda^2-17\lambda+19\\ 0 & 0 & -16(\lambda-2)^3\end{pmatrix}\longrightarrow\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-2)^3\end{pmatrix},$$

所以 $A(\lambda)$ 的秩为 3, 行列式因式组为 $D_1(\lambda)=1$, $D_2(\lambda)=1$, $D_3(\lambda)=(\lambda-2)^3$, 不变因式组为 $d_1(\lambda)=1$, $d_2(\lambda)=1$, $d_3(\lambda)=(\lambda-2)^3$.

注意 在上面建立的理论中, 限制在有理数域 $\mathbb{F}=\mathbb{Q}$、实数域 $\mathbb{F}=\mathbb{R}$ 及复数域 $\mathbf{F}=\mathbf{C}$ 上的结果是完全相同的.

下面限制在复数域 $\mathbb{C}$ 上讨论. 由代数基本定理, 复系数多项式一定可以分解为一次因式的乘积, 所以可以给出下面定义.

定义 6.3.10 设 $d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda)$ 是秩为 r 的 λ 矩阵 $A(\lambda)$ 的不变因式组, 将它们分解为一次因式的乘积:

$$\left\{\begin{array}{l} d_1(\lambda)=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_s)^{e_{1s}}, \\ d_2(\lambda)=(\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_s)^{e_{2s}}, \\ \qquad\vdots \\ d_r(\lambda)=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdots(\lambda-\lambda_s)^{e_{rs}}, \end{array}\right. \tag{6.3.14}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 s 个不同的复数, e_{jk} 为非负整数, $1\leqslant j\leqslant r$, $1\leqslant k\leqslant s$. 再由如下性质: $d_i(\lambda)$ 除得尽 $d_{i+1}(\lambda)$, $1\leqslant i<r$, 所以指数 e_{jk} 有关系

$$\left\{\begin{array}{l} 0\leqslant e_{11}\leqslant e_{21}\leqslant\cdots\leqslant e_{r1}, \\ 0\leqslant e_{12}\leqslant e_{22}\leqslant\cdots\leqslant e_{r2}, \\ \qquad\vdots \\ 0\leqslant e_{1s}\leqslant e_{2s}\leqslant\cdots\leqslant e_{rs}. \end{array}\right. \tag{6.3.15}$$

我们称多项式集

$$\mathfrak{S}(A(\lambda))=\{\,(\lambda-\lambda_k)^{e_{jk}}\mid e_{jk}>0,\ j\in\{1,2,\cdots,r\},\ k\in\{1,2,\cdots,s\}\,\} \tag{6.3.16}$$

为 λ 矩阵 $A(\lambda)$ 的**初等因式组**, 每个多项式 $(\lambda-\lambda_k)^{e_{jk}}$, $e_{jk}>0$ 称为 λ 矩阵 $A(\lambda)$ 的**初等因式**.

注意 仅仅给定 λ 矩阵 $A(\lambda)$ 的初等因式组 $\mathfrak{S}(A(\lambda))$, 并不足以决定 λ 矩阵 $A(\lambda)$ 的不变因式组. 如果给定 λ 矩阵 $A(\lambda)$ 的初等因式组 $\mathfrak{S}(A(\lambda))$ 及秩 r, 那末我们能完全定出不变因式组.

例 6.3.3 设 n 阶 λ 矩阵 $A(\lambda)$ 的秩为 3, 初等因式组为

$$\mathfrak{S}(A(\lambda))=\{\,(\lambda-1)^3,\ (\lambda-1)^4,\ (\lambda-1)^5,\ (\lambda-2)^2,\ (\lambda-2)^3,\ (\lambda-3),\ (\lambda-4)\,\},$$

由于秩为 3, 所以有三个不变因式 $d_1(\lambda)$, $d_2(\lambda)$, $d_3(\lambda)$. 再利用指数关系, 必然有

$$
\begin{aligned}
d_1(\lambda) &= (\lambda-1)^3(\lambda-2)^0(\lambda-3)^0(\lambda-4)^0 = (\lambda-1)^3,\\
d_2(\lambda) &= (\lambda-1)^4(\lambda-2)^2(\lambda-3)^0(\lambda-4)^0 = (\lambda-1)^4(\lambda-2)^2,\\
d_3(\lambda) &= (\lambda-1)^5(\lambda-2)^3(\lambda-3)(\lambda-4).
\end{aligned}
$$

由此可见, 定理 6.3.9 可以改写为

定理 6.3.11 域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 在相抵下的全系不变量是它的行列式因式组, 这等价于全系不变量是它的不变因式组, 这也等价于全系不变量是 λ 矩阵 $A(\lambda)$ 的秩和它的初等因式组.

下面给出求 λ 矩阵 $A(\lambda)$ 的初等因式组的另一种方法.

定理 6.3.12 设域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 在相抵下等价于

$$
\begin{pmatrix} \operatorname{diag}(f_1(\lambda), f_2(\lambda), \cdots, f_r(\lambda)) & 0^{(r,m-r)} \\ 0^{(n-r,r)} & 0^{(n-r,m-r)} \end{pmatrix}, \tag{6.3.17}
$$

其中 r 是 λ 矩阵 $A(\lambda)$ 的秩, 又 $f_1(\lambda)$, $f_2(\lambda)$, $\cdots$, $f_r(\lambda)$ 都是首项系数为 1 的非零多项式. 不要求它们是依次地一个除得尽一个的多项式组. 将它们分解为一次因式的乘积:

$$
\begin{cases}
f_1(\lambda) = (\lambda-\lambda_1)^{f_{11}}(\lambda-\lambda_2)^{f_{12}}\cdots(\lambda-\lambda_s)^{f_{1s}},\\
f_2(\lambda) = (\lambda-\lambda_1)^{f_{21}}(\lambda-\lambda_2)^{f_{22}}\cdots(\lambda-\lambda_s)^{f_{2s}},\\
\qquad\vdots\\
f_r(\lambda) = (\lambda-\lambda_1)^{f_{r1}}(\lambda-\lambda_2)^{f_{r2}}\cdots(\lambda-\lambda_s)^{f_{rs}},
\end{cases} \tag{6.3.18}
$$

其中 λ_1, λ_2, $\cdots$, λ_s 是 s 个不同的复数, f_{jk} 为非负整数, $1 \leqslant j \leqslant r$, $1 \leqslant k \leqslant s$. 则 λ 矩阵 $A(\lambda)$ 的初等因式组为

$$
\mathfrak{S}(A(\lambda)) = \{\, (\lambda-\lambda_k)^{f_{jk}} \mid f_{jk} > 0,\ j \in \{1,2,\cdots,r\},\ k \in \{1,2,\cdots,s\}\,\}. \tag{6.3.19}
$$

证 我们来计算 λ 矩阵 $A(\lambda)$ 的行列式因式. 由定义可知

$$
\begin{aligned}
D_1(\lambda) &= (f_1(\lambda), f_2(\lambda), \cdots, f_r(\lambda)) = \prod_{j=1}^{s}(\lambda-\lambda_j)^{\min(f_{1j}, f_{2j}, \cdots, f_{rj})},\\
D_2(\lambda) &= \{\, f_i(\lambda)f_j(\lambda),\quad 1 \leqslant i < j \leqslant r \,\} \text{ 的最高公因式}\\
&= \prod_{j=1}^{s}(\lambda-\lambda_j)^{\min_{1\leqslant i<k\leqslant r}(f_{ij}+f_{kj})},
\end{aligned}
$$

于是一般取 $t=1,2,\cdots,r$, 则有

$$
\begin{aligned}
D_t(\lambda) &= \{\, f_{i_1}(\lambda)f_{i_2}(\lambda)\cdots f_{i_t}(\lambda),\quad 1\leqslant i_1<i_2<\cdots<i_t\leqslant r \,\}\ \text{的最高公因式}\\
&= \prod_{j=1}^{s}(\lambda-\lambda_j)^{\min_{1\leqslant i_1<i_2<\cdots<i_t\leqslant r}(f_{i_1j}+f_{i_2j}+\cdots+f_{i_tj})}.
\end{aligned}
$$

将 j 固定, 将 $f_{1j},f_{2j},\cdots,f_{rj}$ 按大小次序重新排为

$$0\leqslant e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{rj},$$

于是

$$\min_{1\leqslant i_1<i_2<\cdots<i_t\leqslant r}(f_{i_1j}+f_{i_2j}+\cdots+f_{i_tj})=e_{1j}+e_{2j}+\cdots+e_{tj},\quad 1\leqslant t\leqslant r.$$

取 $j=1,2,\cdots,s$, 便证明了

$$D_t(\lambda)=\prod_{j=1}^{s}(\lambda-\lambda_j)^{e_{1j}+e_{2j}+\cdots+e_{tj}},\quad 1\leqslant t\leqslant r.$$

因此 t 阶不变因式

$$d_t(\lambda)=\frac{D_t(\lambda)}{D_{t-1}(\lambda)}=\prod_{j=1}^{s}(\lambda-\lambda_j)^{e_{tj}},\quad 1\leqslant t\leqslant r.$$

这证明了域 $\mathbb{F}$ 上 λ 矩阵 $A(\lambda)$ 的初等因式组为

$$
\begin{aligned}
\mathfrak{S}(A(\lambda)) &= \{\,(\lambda-\lambda_k)^{e_{jk}}\mid e_{jk}>0,\ j\in\{1,2,\cdots,r,\}\ k\in\{1,2,\cdots,s,\}\,\}\\
&= \{\,(\lambda-\lambda_k)^{f_{jk}}\mid f_{jk}>0,\ j\in\{1,2,\cdots,r,\}\ k\in\{1,2,\cdots,s,\}\,\}.
\end{aligned}
\qquad\square
$$

定理 6.3.13 设 $A_1(\lambda)$ 是域 $\mathbb{F}$ 上 $n\times m$ λ 矩阵, $A_2(\lambda)$ 是域 $\mathbb{F}$ 上 $p\times q$ λ 矩阵, 设

$$A(\lambda)=\operatorname{diag}(A_1(\lambda),A_2(\lambda)),\tag{6.3.20}$$

则 λ 矩阵 $A(\lambda)$ 的全部初等因式由 λ 矩阵 $A_1(\lambda)$ 的全部初等因式和 λ 矩阵 $A_2(\lambda)$ 的全部初等因式合并而成.

证 由定理 6.3.4 和定理 6.3.11 可知, 存在可逆 λ 方阵 $P_i(\lambda)$ 和 $Q_i(\lambda)$, $i=1,2$, 它们分别将 $A_1(\lambda)$ 和 $A_2(\lambda)$ 化为标准形

$$P_1(\lambda)A_1(\lambda)Q_1(\lambda)=\begin{pmatrix}\operatorname{diag}(d_1(\lambda),\cdots,d_s(\lambda)) & 0\\ 0 & 0\end{pmatrix},$$

$$P_2(\lambda)A_2(\lambda)Q_2(\lambda)=\begin{pmatrix}\operatorname{diag}(d_{s+1}(\lambda),\cdots,d_{s+t}(\lambda)) & 0\\ 0 & 0\end{pmatrix},$$

其中 $s = \mathrm{rank}\,(A_1(\lambda))$, $t = \mathrm{rank}\,(A_2(\lambda))$.

显然 $P(\lambda) = \mathrm{diag}\,(P_1(\lambda), P_2(\lambda))$, $Q(\lambda) = \mathrm{diag}\,(Q_1(\lambda), Q_2(\lambda))$ 分别为 $n+p$ 阶及 $m+q$ 阶可逆 λ 方阵, 且

$$P(\lambda)A(\lambda)Q(\lambda) = \begin{pmatrix} \mathrm{diag}\,(d_1(\lambda),\cdots,d_s(\lambda)) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathrm{diag}\,(d_{s+1}(\lambda),\cdots,d_{s+t}(\lambda)) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

所以 λ 矩阵 $A(\lambda)$ 相抵于 λ 矩阵

$$P_1(\lambda)A(\lambda)Q_1(\lambda) = \begin{pmatrix} \mathrm{diag}\,(d_1(\lambda),\cdots,d_s(\lambda),d_{s+1}(\lambda),\cdots,d_{s+t}(\lambda)) & 0 \\ 0 & 0 \end{pmatrix},$$

其中 $P_1(\lambda) = P_0P(\lambda)$, $Q_1(\lambda) = Q(\lambda)Q_0$, P_0 和 Q_0 是置换方阵. 由定理 6.3.12 便证明了 λ 矩阵 $A(\lambda)$ 的初等因式组为 λ 矩阵 $A_1(\lambda)$ 的初等因式组加上 λ 矩阵 $A_2(\lambda)$ 的初等因式组而成. □

习 题 6.3

6.3.1 求下列 λ 矩阵的行列式因式、不变因式和初等因式. 并写出它们在相抵下的标准形.

$$\begin{pmatrix} \lambda+1 & 1 & 0 \\ 0 & \lambda+1 & 1 \\ 0 & 0 & \lambda+1 \end{pmatrix}, \quad \begin{pmatrix} \lambda+1 & 2 & 1 \\ 0 & \lambda+1 & 2 \\ 0 & 0 & \lambda+1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 & -1 \\ 0 & \lambda-1 & 0 \\ -1 & 0 & \lambda \end{pmatrix}.$$

6.3.2 设 H_1 和 H_2 为 n 阶 Hermite 方阵. 试证: λ 方阵 $H_1 + \lambda H_2$ 的行列式因式、不变因式和初等因式都是实多项式.

6.3.3 给定 n 个整数 $m_1, m_2, \cdots, m_n$. 设它们的最高公因式 $d = (m_1, m_2, \cdots, m_n)$. 试证: 存在一个行列式等于 d 的 n 阶整数方阵 A, 以整数 $m_1, m_2, \cdots, m_n$ 为指定的行或列.

6.3.4 以整数为元素的 $n \times m$ 矩阵称为**整数矩阵**. 试仿照 λ 矩阵的理论来建立整数矩阵在相抵下的标准形理论, 即定义行列式因式、不变因式和初等因式. 定义初等变换、相抵以及求相抵下的标准形和决定整数矩阵在相抵下的全系不变量. 这里要证明的是: 整数方阵可逆当且仅当它的行列式等于 ± 1.

§6.4 复方阵在相似下的 Jordan 标准形

定义 6.4.1 域 $\mathbb{F}$ 上两个 n 阶方阵 A 和 B 称为**相似的**, 如果存在域 $\mathbb{F}$ 上一个 n 阶非异方阵 P, 使得 $B = PAP^{-1}$.

本节的目的是寻求 n 阶方阵 A 和 B 在相似下的标准形. 下面先来解决域 $\mathbb{F} = \mathbb{C}$ 的情形, 在下一章, 我们利用复的情形来解决域 $\mathbb{F} = \mathbb{R}$ 的情形. 在这一节我们利用 λ

矩阵在相抵下的标准形理论, 给出复方阵在相似下的标准形的具体计算方法. 在下一章, 我们给出这种标准形的一种几何解释.

引理 6.4.2　设 D 为 n 阶常数复方阵, $T(\lambda)$ 为 n 阶 λ 方阵, 则唯一存在 n 阶 λ 方阵 $T_1(\lambda)$, $T_2(\lambda)$ 及 n 阶常数复方阵 W_1, W_2, 使得

$$T(\lambda)=(\lambda E_n+D)T_1(\lambda)+W_1=T_2(\lambda)(\lambda E_n+D)+W_2. \tag{6.4.1}$$

证　由于 λ 矩阵的每个元素都为 λ 的复多项式的方阵, 所以 λ 方阵实质上是以复方阵为系数的 λ 的多项式, 即

$$T(\lambda)=\lambda^tT_t+\lambda^{t-1}T_{t-1}+\cdots+\lambda T_1+T_0,$$

其中 $T_0, T_1, \cdots, T_t$ 为 n 阶常数复方阵, $T_t\neq 0$. 和通常多项式的情形一样, 所以

$$\begin{aligned}T(\lambda)&=\lambda^tT_t+\lambda^{t-1}T_{t-1}+\cdots+\lambda T_1+T_0\\&=(\lambda E_n+D)(\lambda^{t-1}A_{t-1})+\text{ 次数小于 } t \text{ 的项}.\end{aligned}$$

对 t 作归纳法, 便证明了存在 n 阶 λ 方阵 $T_1(\lambda)$ 及 n 阶常数复方阵 W_1, 使得 $T(\lambda)=(\lambda E_n+D)T_1(\lambda)+W_1$. 再证唯一性. 今若

$$T(\lambda)=(\lambda E_n+D)T_1(\lambda)+W_1=(\lambda E_n+D)\widetilde{T_1}(\lambda)+\widetilde{W_1},$$

其中 $T_1(\lambda)$, $\widetilde{T_1}(\lambda)$ 都是 n 阶 λ 方阵, W_1, $\widetilde{W_1}$ 都是 n 阶常数复方阵. 于是有

$$(\lambda E_n+D)(T_1(\lambda)-\widetilde{T_1}(\lambda))=\widetilde{W_1}-W_1,$$

设 $K(\lambda)=T_1(\lambda)-\widetilde{T_1}(\lambda)=\sum\limits_{j=0}^{s}\lambda^jB_j$, $B_s\neq 0$. 于是 $(\lambda E_n+D)(\lambda^sB_s+\cdots)=\widetilde{W_1}-W_1$. 双方比较 λ^{s+1} 的系数, 便推出矛盾. 所以证明了 $K(\lambda)=T_1(\lambda)-\widetilde{T_1}(\lambda)=0$. 因此 $\widetilde{W_1}=W_1$. 这证明了分解唯一性.

考虑 $T(\lambda)'$, 则有 n 阶 λ 方阵 $T_3(\lambda)$ 和 n 阶常数复方阵 W_2, 使得 $T(\lambda)'=(\lambda E_n+D')T_2(\lambda)'+W_2'$. 双方取转置, 便证明了 $T(\lambda)=T_2(\lambda)(\lambda E_n+D)+W_2$. □

定理 6.4.3　n 阶复方阵 A 和 B 相似的充分且必要条件是 n 阶 λ 方阵 λE_n-A 和 λE_n-B 相抵.

证　先证必要性. 假设存在 n 阶非异复方阵 P, 使得 $B=PAP^{-1}$, 那末

$$\lambda E_n-B=\lambda E_n-PAP^{-1}=P(\lambda E_n-A)P^{-1}.$$

显然常数非异复方阵 P 是可逆 λ 方阵, 所以 λ 方阵 λE_n-A 和 λE_n-B 相抵.

再证充分性. 设 λ 方阵 λE_n-A 和 λE_n-B 相抵, 所以存在 n 阶可逆 λ 方阵 $P(\lambda)$ 及 $Q(\lambda)$, 使得

$$P(\lambda)(\lambda E_n-A)=(\lambda E_n-B)Q(\lambda). \tag{6.4.2}$$

由引理 6.4.2 可知, 存在 n 阶 λ 方阵 $P_1(\lambda)$, $Q_1(\lambda)$ 及 n 阶常数复方阵 P, Q, 使得

$$P(\lambda)=(\lambda E_n-B)P_1(\lambda)+P,\quad Q(\lambda)=Q_1(\lambda)(\lambda E_n-A)+Q. \tag{6.4.3}$$

因此式 (6.4.2) 给出

$$\lambda(P-Q)+BQ-PA=(\lambda E_n-B)(Q_1(\lambda)-P_1(\lambda))(\lambda E_n-A). \tag{6.4.4}$$

我们断言 $Q_1(\lambda)=P_1(\lambda)$. 设若不然, 则

$$T(\lambda)=Q_1(\lambda)-P_1(\lambda)=\lambda^tT_t+\lambda^{t-1}T_{t-1}+\cdots+\lambda T_1+T_0,$$

其中 $T_0, T_1, \cdots, T_t$ 为 n 阶复方阵, 且 T_t 不是零方阵. 因此

$$(\lambda E_n-B)T(\lambda)(\lambda E_n-A)=\lambda^{t+2}T_t+\text{ 次数小于 }t+2\text{ 的项}.$$

今式 (6.4.4) 的右式关于 λ 的次数至少为 $t+2$, 然而式 (6.4.4) 的左式关于 λ 的次数至多为 1, 这就导出矛盾. 所以证明了断言. 而式 (6.4.4) 给出 $P=Q$, $BQ=PA$, 即 $BP=PA$. 所以为了要证明复方阵 A 和 B 相似, 只要证明 $\det(P)\neq 0$ 就行了.

注意到 $P(\lambda)$ 是可逆 λ 方阵, 所以存在 λ 方阵 $\widetilde{P}(\lambda)$, 使得 $P(\lambda)\widetilde{P}(\lambda)=E_n$. 同样, 不难证明有分解式

$$\widetilde{P}(\lambda)=(\lambda E_n-A)C(\lambda)+\widetilde{P}, \tag{6.4.5}$$

其中 $C(\lambda)$ 是 λ 方阵, $\widetilde{P}$ 是常数复方阵. 由式 (6.4.2) 和 (6.4.3), 有

$$\begin{aligned}E_n&=P(\lambda)\widetilde{P}(\lambda)=P(\lambda)(\lambda E_n-A)C(\lambda)+P(\lambda)\widetilde{P}\\&=(\lambda E_n-B)Q(\lambda)C(\lambda)+(\lambda E_n-B)P_1(\lambda)\widetilde{P}+P\widetilde{P}\\&=(\lambda E_n-B)(Q(\lambda)C(\lambda)+P_1(\lambda)\widetilde{P})+P\widetilde{P}.\end{aligned} \tag{6.4.6}$$

我们断言 $Q(\lambda)C(\lambda)+P_1(\lambda)\widetilde{P}=0$. 设若不然, 则

$$\widetilde{T}(\lambda)=Q(\lambda)C(\lambda)+P_1(\lambda)\widetilde{P}=\lambda^t\widetilde{T_t}+\lambda^{t-1}\widetilde{T_{t-1}}+\cdots+\lambda\widetilde{T_1}+\widetilde{T_0},$$

其中 $\widetilde{T_0}, \widetilde{T_1}, \cdots, \widetilde{T_t}$ 为 n 阶复方阵, 且 $\widetilde{T_t}$ 不是零方阵. 因此

$$(\lambda E_n-B)\widetilde{T}(\lambda)=\lambda^{t+1}\widetilde{T_t}+\text{ 次数小于 }t+1\text{ 的项}.$$

今式 (6.4.6) 的右边关于 λ 的次数至少为 $t+1$, 然而式 (6.4.6) 的左边为单位矩阵, 这就导出矛盾. 所以证明了断言. 因此 $E_n=P\widetilde{P}$, 所以 $\det(P)\neq 0$. □

这个定理的重要性在于, 它利用了定理 6.3.11, 便给出了

定理 6.4.4 n 阶复方阵 A 在相似下的全系不变量是非异 λ 方阵 λE_n-A 的全部初等因式.

为方便起见, 今后称 λ 方阵 λE_n-A 的初等因式为复方阵 A 的**初等因式**. λ 方阵 λE_n-A 的行列式因式组、不变因式组、初等因式组分别称为复方阵 A 的**行列式因式组**、**不变因式组**、**初等因式组**.

在前面, 凡讨论标准形理论, 总是先求标准形, 再定全系不变量. 在讨论相似下的标准形理论时, 恰好倒了过来. 由定理 6.4.4, 求 n 阶复方阵 A 在相似下的标准形的问题可以化为先计算出 λ 方阵 $\lambda E_n - A$ 的全部行列式因式, 再计算出全部不变因式, 从而算出全部初等因式. 或者用初等变换将 n 阶 λ 方阵 $\lambda E_n - A$ 化为对角形, 再将对角元素分解为一次因式的乘积, 从而定出了 A 的初等因式组. 当然, 我们求出了相似下的全系不变量, 然而并没有求出标准形. 不过, 我们已经知道了全系不变量以后, 就不再需要利用相似关系去寻找标准形. 即不需要求非异复方阵 P, 而是构作一个简单的复方阵, 使得这个复方阵的全部初等因式和已给的复方阵的全部初等因式一样, 那末利用定理 6.4.4, 便能断定这两个复方阵一定相似. 尽管这时我们不知道什么样的 n 阶非异复方阵使它们相似, 但是知道它一定存在.

另一方面, 我们并没有给出好办法来计算非异复方阵 P, 使得 $PAP^{-1} = J$, 其中 J 为相似下的标准形.

在下面我们将上面描述的过程, 严格地表达为:

为了构作复方阵在相似下的标准形, 即所谓 Jordan 标准形. 给定复方阵 A, 假设 A 的全部初等因式为

$$(\lambda - \lambda_1)^{e_1}, \quad (\lambda - \lambda_2)^{e_2}, \quad \cdots, \quad (\lambda - \lambda_s)^{e_s}, \tag{6.4.7}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 不要求都不相等. 又 $e_1, e_2, \cdots, e_s$ 为正整数. 它们有

$$e_1 + e_2 + \cdots + e_s = n. \tag{6.4.8}$$

利用定理 6.3.11, 找一个初等因式为 $(\lambda - \lambda_j)^{e_j}$ 的 e_j 阶复方阵 J_j, $j = 1, 2, \cdots, s$, 那末由式 (6.3.11), n 阶复方阵

$$J = \mathrm{diag}\,(J_1, J_2, \cdots, J_s) \tag{6.4.9}$$

的全部初等因式由复方阵 $J_1, J_2, \cdots, J_s$ 的全部初等因式构成, 所以是 (6.4.7). 即由复方阵 A 的全部初等因式构成. 这证明了复方阵 A 和复方阵 J 的全部初等因式相同. 故复方阵 A 和 J 相似.

今取 e_j 阶复方阵

$$J_j = \begin{pmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_j & 1 \\ & & & & \lambda_j \end{pmatrix} = \lambda_j E_{e_j} + N_{e_j}, \tag{6.4.10}$$

其中 E_{e_j} 为 e_j 阶单位方阵, N_{e_j} 为 e_j 阶幂零方阵

$$N_{e_j}=\begin{pmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&0&1\\&&&&0\end{pmatrix}. \tag{6.4.11}$$

所以 J_j 是 e_j 阶复方阵. 它的对角元素全是 λ_j, 对角元素上面一排全是 1, 其余位置元素全是零. 由直接计算可知, 这时方阵 J_j 的初等因式组只由一个初等因式 $(\lambda-\lambda_j)^{e_j}$ 组成. 事实上, e_j 阶 λ 方阵 $\lambda E_{e_j}-J_j=(\lambda-\lambda_j)E_{e_j}-N_{e_j}$ 的行列式因式为 $D_1=D_2=\cdots=D_{e_j-1}=1$, $D_{e_j}=(\lambda-\lambda_j)^{e_j}$. 所以不变因式为 $d_1=d_2=\cdots=d_{e_j-1}=1$, $d_{e_j}=(\lambda-\lambda_j)^{e_j}$, 因此方阵 J_j 的初等因式只有一个 $(\lambda-\lambda_j)^{e_j}$.

由于 e_j 阶复方阵 J_j 的阶数 e_j 和复方阵 J_j 的元素都由初等因式 $(\lambda-\lambda_j)^{e_j}$ 的次数 e_j 及特征根 λ_j 完全决定. 所以 J_j 称为属于初等因式的 Jordan 块, 或简称为**Jordan 块**.

定义 6.4.5 如果准对角方阵 J 的所有对角块 $J_1, J_2, \cdots, J_s$ 都是 Jordan 块, 即 $J_j=\lambda_j E_{e_j}+N_{e_j}$, $1\leqslant j\leqslant s$, 则 J 称为**Jordan 标准形**. 为了方便起见, 在表达 Jordan 标准形 J_j 时, 我们记 E_{e_j} 为 E_j; 记 N_{e_j} 为 N_j, 这里 E_j 不表示它是 j 阶单位复方阵, 表示它是 e_j 阶单位复方阵; N_j 不表示它是 j 阶方阵 (6.4.11), 表示它是 e_j 阶方阵 (6.4.11).

所以我们有

定理 6.4.6 设 n 阶复方阵 A 的所有初等因式为 (6.4.7), 那末 A 相似于 Jordan 标准形

$$\operatorname{diag}(J_1, J_2, \cdots, J_s), \tag{6.4.12}$$

其中 $J_j=\lambda_j E_j+N_j$ 是属于初等因式 $(\lambda-\lambda_j)^{e_j}$ 的 Jordan 块, $1\leqslant j\leqslant s$. 所以, 如果不计 Jordan 块的编排次序, 那末标准形是由初等因式组唯一确定的.

例 6.4.1 求方阵

$$A=\begin{pmatrix}13&16&16\\-5&-7&-6\\-5&-8&-7\end{pmatrix}$$

在相似下的标准形.

解 先求 λ 方阵

$$\lambda E_n - A = \begin{pmatrix} \lambda-13 & -16 & -16 \\ 5 & \lambda+7 & 6 \\ 5 & 8 & \lambda+7 \end{pmatrix}$$

的所有初等因式. 利用初等变换可知它们是 $(\lambda-2)$, $(\lambda-1)^2$, 于是方阵 A 相似于 Jordan 标准形

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

注意 1 关于 Jordan 标准形, 重要的是每块的阶数和每块的对角元素. 因为 λ 方阵 $\lambda E_n - A$ 的所有初等因式的连乘积等于方阵 A 的特征多项式, 所以每个初等因式的根都是方阵 A 的特征根. 换句话说, Jordan 标准形的 n 个对角元素是方阵 A 的所有特征根. 然而同一特征根可以对应几个初等因式, 即初等因式 $(\lambda-\lambda_j)^{e_j}$ 的次数 e_j 不是特征根 λ_j 的重数. 例如, 单位方阵的全部初等因式有 n 个, 它们都是 $\lambda-1$, 然而 1 是单位方阵的 n 重特征根. 这就是说, 虽然 Jordan 标准形为上三角方阵, 对角元素构成 n 阶复方阵 A 的所有特征根, 且同一特征根可能在几个 Jordan 块中出现. 所以这几个 Jordan 块的阶数不超过相应特征根的重数, 它们的和是特征根的重数.

注意 2 给定复方阵 A. A 可以有多种相似下的 Jordan 标准形, 即我们可以取下面四种形式的方阵作为初等因式 $(\lambda-\lambda_0)^e$ 的 e 阶 Jordan 块. 我们介绍如下:

$$J_0 = \lambda_0 E_n + N, \qquad J_0' = \lambda_0 E_n + N', \tag{6.4.13}$$

在 $\lambda_0 \neq 0$ 时, J_0 还可以相似于 e 阶方阵

$$\widetilde{J}_0 = \lambda_0 M^{(e)} = \lambda_0(E_n + N), \qquad \widetilde{J}_0' = \lambda_0 (M^{(e)})' = \lambda_0(E_n + N'). \tag{6.4.14}$$

这只要证明 $\lambda_0(E_n+N)$ 的初等因式是 $(\lambda-\lambda_0)^e$ 就可以了. 为此, 只要计算 λ 方阵 $\lambda E_n - \lambda_0(E_n+N)$, 即计算 λ 方阵

$$\begin{pmatrix} \lambda-\lambda_0 & -\lambda_0 & & & \\ & \lambda-\lambda_0 & -\lambda_0 & & \\ & & \ddots & \ddots & \\ & & & \lambda-\lambda_0 & -\lambda_0 \\ & & & & \lambda-\lambda_0 \end{pmatrix} \tag{6.4.15}$$

的行列式因式, 便能证明这个断言. 也可以计算 $P^{-1}(\lambda_0 E_n + N)P$, 其中

$$P = \operatorname{diag}(1, \lambda_0, \cdots, \lambda_0^{n-1}). \tag{6.4.16}$$

注意 3 设 λ 方阵 $\lambda E_n - A$ 的不变因式组为 $d_1(\lambda)$, $d_2(\lambda)$, $\cdots$, $d_n(\lambda)$, 则由式 (6.3.14) 和式 (6.3.15) 有

$$\begin{cases} d_1(\lambda) = (\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdots(\lambda-\lambda_t)^{e_{1t}}, \\ d_2(\lambda) = (\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdots(\lambda-\lambda_t)^{e_{2t}}, \\ \qquad\vdots \\ d_n(\lambda) = (\lambda-\lambda_1)^{e_{n1}}(\lambda-\lambda_2)^{e_{n2}}\cdots(\lambda-\lambda_t)^{e_{nt}}, \end{cases} \tag{6.4.17}$$

其中 λ_1, λ_2, $\cdots$, λ_t 是矩阵 A 的所有不同特征根, e_{jk} 为非负整数, $1 \leqslant j \leqslant n$, $1 \leqslant k \leqslant t$. 再由性质: $d_i(\lambda)$ 除得尽 $d_{i+1}(\lambda)$, $1 \leqslant i < n$, 所以指数 e_{jk} 有关系

$$\begin{cases} 0 = e_{11} = \cdots = e_{r_1,1} < e_{r_1+1,1} \leqslant \cdots \leqslant e_{n1}, \\ 0 = e_{12} = \cdots = e_{r_2,2} < e_{r_2+1,2} \leqslant \cdots \leqslant e_{n2}, \\ \qquad\vdots \\ 0 = e_{1t} = \cdots = e_{r_t,t} < e_{r_t+1,t} \leqslant \cdots \leqslant e_{nt}. \end{cases} \tag{6.4.18}$$

于是

$$\mathfrak{S}(A) = \{\, (\lambda-\lambda_k)^{e_{jk}} \;\mid\; e_{jk} > 0,\ j \in \{\, r_k+1, \cdots, n, \} \ k \in \{\, 1, 2, \cdots, t, \}\,\} \tag{6.4.19}$$

为矩阵 A 的初等因式组. 每个多项式 $(\lambda-\lambda_k)^{e_{jk}}$ 为矩阵 A 的初等因式.

另外, 虽然上面的理论告诉我们: 存在 n 阶非异复方阵 P, 使得 $PAP^{-1} = J$. 实际上也是能够按照定理 6.4.3 的充分性证明, 来给出下面 n 阶非异复方阵 P 的一种计算方法的.

定理 6.4.7 给定复方阵 A, 设复方阵 J 是复方阵 A 的 Jordan 标准形. 即存在一系列初等变换将 λ 方阵 $\lambda E_n - A$ 变为 λ 方阵 $\lambda E_n - J$ 方阵. 换句话说, 存在 n 阶可逆 λ 方阵 $P(\lambda)$ 及 $Q(\lambda)$, 使得

$$P(\lambda)(\lambda E_n - A)Q(\lambda)^{-1} = \lambda E_n - J, \tag{6.4.20}$$

其中 $P(\lambda)$ 是一系列行的初等变换决定的初等 λ 方阵的乘积, $Q(\lambda)$ 是一系列列的初等变换决定的初等 λ 方阵的乘积. 记

$$P(\lambda) = \lambda^t P_t + \lambda^{t-1}P_{t-1} + \cdots + \lambda P_1 + P_0, \tag{6.4.21}$$

其中 $P_0, P_1, \cdots, P_t$ 为 n 阶常数复方阵, 又 $P_t \neq 0$. 则有 $P(\lambda) = (\lambda E_n - J)P_1(\lambda) + P$, 其中 $P_1(\lambda)$ 为 n 阶 λ 方阵, P 为 n 阶常数复方阵. 又 $\det(P) \neq 0$, $PAP^{-1} = J$, 而且

$$P = J^t P_t + J^{t-1}P_{t-1} + \cdots + JP_1 + P_0. \tag{6.4.22}$$

证 今存在 n 阶 λ 方阵 $P_1(\lambda)$ 及 n 阶常数复方阵 P, 使得 $P(\lambda) = (\lambda E_n - J)P_1(\lambda) + P$. 由定理 6.4.3 的证明可知 $\det(P) \neq 0$, 且 $PAP^{-1} = J$. 所以问题化为给出 P 关于 J 的明显公式. 为此, 记

$$P_1(\lambda) = \lambda^{t-1}D_{t-1} + \lambda^{t-2}D_{t-2} + \cdots + \lambda D_1 + D_0,$$

其中 $D_0, D_1, \cdots, D_{t-1}$ 为 n 阶常数复方阵, 又 $D_{t-1} \neq 0$. 所以由式 (6.4.21) 及

$$\begin{aligned} P(\lambda) &= (\lambda E_n - J)P_1(\lambda) + P \\ &= (\lambda E_n - J)(\lambda^{t-1}D_{t-1} + \lambda^{t-2}D_{t-2} + \cdots + \lambda D_1 + D_0) + P \\ &= \lambda^t D_{t-1} + \lambda^{t-1}(D_{t-2} - JD_{t-1}) + \cdots + \lambda(D_0 - JD_1) + (P - JD_0). \end{aligned} \tag{6.4.23}$$

双方比较同类项系数, 因此式 (6.4.21) 和式 (6.4.22) 给出

$$\begin{aligned} P_t &= D_{t-1}, \\ P_{t-1} &= D_{t-2} - JD_{t-1}, \\ &\vdots \\ P_1 &= D_0 - JD_1, \\ P_0 &= P - JD_0. \end{aligned}$$

依次消去 $D_0, D_1, \cdots, D_{t-1}$, 便有

$$P = P_0 + JD_0 = P_0 + JP_1 + J^2D_1 = \cdots = P_0 + JP_1 + J^2P_2 + \cdots + J^tP_t. \qquad \square$$

定理 6.4.8 n 阶复方阵 A 的极小多项式为 λ 方阵 $\lambda E_n - A$ 的第 n 个不变因式 $d_n(\lambda)$.

证 由于在相似下极小多项式不改变, 所以我们只要证明 Jordan 标准形 $J = \operatorname{diag}(J_1, J_2, \cdots, J_s)$ 的极小多项式为 λ 方阵 $\lambda E_n - J$ 的第 n 个不变因式 $d_n(\lambda)$ 就行了.

由式 (6.4.17), (6.4.18), (6.4.19) 给出

$$d_n(\lambda) = (\lambda - \lambda_1)^{e_{n1}}(\lambda - \lambda_2)^{e_{n2}} \cdots (\lambda - \lambda_t)^{e_{nt}}, \tag{6.4.24}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是矩阵 A 的所有不同特征根. $e_{n1}, e_{n2}, \cdots, e_{nt}$ 为正整数. 又 $e_{jk} \leqslant e_{nk}$, $1 \leqslant j \leqslant n$, $1 \leqslant k \leqslant t$.

下面先证 $d_n(J)=0$. 由于 Jordan 标准形 J 是准对角方阵, 对应初等因式 $(\lambda-\lambda_k)^{e_{jk}}$ 的 Jordan 块为对角块 $J_{jk}=\lambda_k E_{jk}+N_{jk}$, 其中 E_{jk} 和 N_{jk} 都是 e_{jk} 阶方阵, 所以

$$\begin{aligned}d_n(J_{jk})&=(J_{jk}-\lambda_1 E_{jk})^{e_{n1}}\cdots(J_{jk}-\lambda_t E_{jk})^{e_{nt}}\\&=((\lambda_k-\lambda_1)E_{jk}+N_{jk})^{e_{n1}}\cdots((\lambda_k-\lambda_t)E_{jk}+N_{jk})^{e_{nt}},\end{aligned}$$

其中有因式 $((\lambda_k-\lambda_k)E_{uk}+N_{uk})^{e_{nk}}=N_{uk}^{e_{nk}}$. 由 $e_{uk}\leqslant e_{nk}$ 可知 $((\lambda_k-\lambda_k)E_{uk}+N_{uk})^{e_{nk}}=0$. 这证明了 $d_n(J)=0$.

记 Jordan 标准形 J 的极小多项式为 $m(\lambda)$. 由 $d_n(J)=0$, 有 $m(\lambda)\mid d_n(\lambda)$. 所以极小多项式

$$m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t},$$

其中 $m_1, m_2, \cdots, m_t$ 为非负整数, 而且 $m_j\leqslant e_{nj}$, $1\leqslant j\leqslant t$. 因此为了证 $d_n(\lambda)=m(\lambda)$, 只要证 $e_{nj}=m_j$ 就可以了. 今 $m(J)=0$, 所以有

$$m(J_{nj})=((\lambda_j-\lambda_1)E_{nj}+N_{nj})^{m_1}\cdots((\lambda_j-\lambda_t)E_{nj}+N_{nj})^{m_t}=0.$$

注意到 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 为 Jordan 标准形 J 的所有不同特征根. 当 $k\neq j$ 时, $\lambda_k\neq\lambda_j$. 因此, $((\lambda_j-\lambda_k)E_{nj}+N_{nj})^{m_k}$ 为非异复方阵. 所以为了 $m(J_{nj})=0$, 必须 $((\lambda_j-\lambda_j)E_{nj}+N_{nj})^{m_j}=N_{nj}^{m_j}=0$. 这证明了 $m_j\geqslant e_{nj}$, 于是 $e_{nj}=m_j$, $1\leqslant j\leqslant t$. 即 $d_n(\lambda)=m(\lambda)$. □

习 题 6.4

6.4.1 求下列方阵在相似下的 Jordan 标准形:

$$\begin{pmatrix}8&30&-14\\-6&-19&9\\-6&-23&11\end{pmatrix},\quad\begin{pmatrix}4&5&-2\\-2&-2&1\\-1&1&1\end{pmatrix},\quad\begin{pmatrix}1&2&3&4\\0&1&2&3\\0&0&1&2\\0&0&0&1\end{pmatrix},$$

$$\begin{pmatrix}3&1&0&0\\-4&-1&0&0\\7&1&2&1\\-17&-6&-1&0\end{pmatrix},\quad\begin{pmatrix}-5&-3&-2&4\\2&0&1&-1\\10&7&4&-9\\2&0&1&0\end{pmatrix}.$$

6.4.2 求方阵

$$\begin{pmatrix}&&&a_{1n}\\&&\cdot&\\&\cdot&&\\&a_{n-1,2}&&\\a_{n1}&&&\end{pmatrix}$$

在相似下的 Jordan 标准形.

6.4.3 求方阵

$$M=\begin{pmatrix}1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1\\ & & & & 1\end{pmatrix},\quad L=\begin{pmatrix}\cos(\theta)M & \sin(\theta)M\\ -\sin(\theta)M & \cos(\theta)M\end{pmatrix}$$

的 m 次方幂.

6.4.4 设 $f(x)$ 是域 $\mathbb{F}$ 上的多项式, 试证: 属于初等因式 $(\lambda-\lambda_0)^n$ 的 Jordan 块 J 有

$$f(J)=\begin{pmatrix}f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f''(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0)\\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0)\\ & & f(\lambda_0) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda_0)\\ & & & \ddots & \vdots\\ & & & & f(\lambda_0)\end{pmatrix},\tag{6.4.25}$$

其中 $f^{(k)}(\lambda)=\dfrac{\mathrm{d}f^{(k-1)}(\lambda)}{\mathrm{d}\lambda}$, $k=1,2,\cdots$, 又 $f^{(0)}(\lambda)=f(\lambda)$.

6.4.5 判断下列方阵是否相似:

$$\begin{pmatrix}3 & -3 & 1\\ -1 & -5 & -1\\ -4 & 6 & -2\end{pmatrix},\quad \begin{pmatrix}-6 & 4 & 16\\ -7 & 5 & 16\\ -2 & 1 & 6\end{pmatrix},\quad \begin{pmatrix}3 & -3 & 1\\ -2 & 4 & -1\\ -7 & 9 & -2\end{pmatrix},\quad \begin{pmatrix}-28 & 22 & -5\\ -60 & 46 & -10\\ -90 & 66 & -13\end{pmatrix}.$$

6.4.6 试证: 存在置换方阵 (即第一类初等变换的连续作用) P, 使得

$$P\begin{pmatrix}(a_{ij}^{(11)}) & \cdots & (a_{ij}^{(1m)})\\ \vdots & & \vdots\\ (a_{ij}^{(m1)}) & \cdots & (a_{ij}^{(mm)})\end{pmatrix}P^{-1}=\begin{pmatrix}(a_{11}^{(pq)}) & \cdots & (a_{1n}^{(pq)})\\ \vdots & & \vdots\\ (a_{n1}^{(pq)}) & \cdots & (a_{nn}^{(pq)})\end{pmatrix},$$

其中

$$(a_{ij}^{(uv)})=\begin{pmatrix}a_{ij}^{(11)} & \cdots & a_{ij}^{(1m)}\\ \vdots & & \vdots\\ a_{ij}^{(m1)} & \cdots & a_{ij}^{(mm)}\end{pmatrix},\quad (a_{uv}^{(pq)})=\begin{pmatrix}a_{11}^{(pq)} & \cdots & a_{1n}^{(pq)}\\ \vdots & & \vdots\\ a_{n1}^{(pq)} & \cdots & a_{nn}^{(pq)}\end{pmatrix},$$

$1\leqslant i,j\leqslant n$, $1\leqslant u,v\leqslant m$. 利用这一事实, 试证下列方阵互相相似:

(1) mn 阶方阵:

$$\begin{pmatrix} E_n & E_n & & & \\ & E_n & E_n & & \\ & & \ddots & \ddots & \\ & & & E_n & E_n \\ & & & & E_n \end{pmatrix}, \quad \operatorname{diag}(E_n+N,\cdots,E_n+N),$$

其中 $M_n=E_n+N_n$, N_n 为 n 阶方阵, 由式 (6.4.11) 定义;

(2) $2mn$ 阶方阵:

$$\begin{pmatrix} \begin{pmatrix} 0 & M_m \\ M_m & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & M_m \\ M_m & 0 \end{pmatrix} \end{pmatrix}, \quad \operatorname{diag}(R_{2n},R_{2n},\cdots,R_{2n}),$$

其中 R_{2n} 为 $2n$ 阶方阵, 定义为

$$R_{2n}=\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right).$$

(3) 设 $A_1, A_2, \cdots, A_m, B_1, B_2, \cdots, B_m, C_1, C_2, \cdots, C_m, D_1, D_2, \cdots, D_m$ 都是 n 阶方阵, 其中

$$\det\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}\neq 0.$$

试证:

$$\det\begin{pmatrix} \begin{pmatrix} A_m & \cdots & A_1 \\ \vdots & ⋰ & \\ A_1 & & \end{pmatrix} & \begin{pmatrix} B_m & \cdots & B_1 \\ \vdots & ⋰ & \\ B_1 & & \end{pmatrix} \\ \begin{pmatrix} C_m & \cdots & C_1 \\ \vdots & ⋰ & \\ C_1 & & \end{pmatrix} & \begin{pmatrix} D_m & \cdots & D_1 \\ \vdots & ⋰ & \\ D_1 & & \end{pmatrix} \end{pmatrix}\neq 0.$$

6.4.7 试证:n 阶方阵 A 相似于对角形当且仅当 A 的初等因式都是一次的.

6.4.8 试证: 适合条件 $A^m=E_n$ 的 n 阶方阵 A 都相似于对角形.

6.4.9 试证: 任一方阵和它的转置方阵都相似.

6.4.10 设 A 和 B 都是 n 阶方阵, 且 $\det(A)\neq 0$. 试证: AB 和 BA 相似.

6.4.11 设 n 阶方阵 A 的特征根都等于 1, 试证: A 和 A 的任意方幂都相似.

6.4.12 给定正整数 m, 试定出所有 n 阶方阵 A, 使得 A 和 A^m 相似.

6.4.13 设 A 和 B 分别为 n 阶及 m 阶方阵, X 为 nm 个独立未知数构成的 $n\times m$ 矩阵. 什么时候矩阵方程

$$AX=XB$$

有非零解?

6.4.14 试证: 存在一个 n 阶非异方阵 P, 使得它可以同时使所有轮回方阵

$$\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}$$

都相似于对角形.

6.4.15 设 $A_1, A_2, \cdots, A_m$ 为 n 阶方阵, 适合条件 $A_1+A_2+\cdots+A_m=E_n$. 试证: 下面三条件互相等价.

(1) $A_1, A_2, \cdots, A_m$ 都幂等;

(2) $\sum\limits_{j=1}^{m}\operatorname{rank}(A_j)=n$;

(3) $A_iA_j=0,\ i\neq j,\ 1\leqslant i,j\leqslant m$, 且这时存在 n 阶非异方阵 P, 使得

$$P^{-1}A_jP=\operatorname{diag}(0^{(n_1)},\cdots,0^{(n_{j-1})},E_{n_j},0^{(n_{j+1})},\cdots,0^{(n_m)}),\quad 1\leqslant j\leqslant m.$$

6.4.16 设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 为域 $\mathbb{F}$ 上 n 阶方阵 A 的 m 个不同特征根. 试证:

$$n(m-1)\leqslant\sum_{j=1}^{m}\operatorname{rank}(\lambda_j-A).$$

6.4.17 试证: 任一实方阵可写为两个实对称方阵的乘积.

6.4.18 设 $0\neq a\in\mathbb{F}$, 试求 n 阶复方阵

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

在相似下的 Jordan 标准形.

6.4.19 试求 N_n^m 的 Jordan 标准形, 其中 N_n 为 n 阶方阵, 由式 (6.4.11) 定义.

6.4.20 设 A 为 n 阶方阵, 且 $\operatorname{tr}(A)=0$. 试证: A 相似于 n 阶方阵 B, 其中 B 的对角元素都是零. 又存在 n 阶方阵 X, Y, 使得 $[X,Y]=XY-YX=A$.

6.4.21 设 A 为 n 阶方阵, 记 $A^m=(a_{ij}^{(m)})$, $m=1, 2, \cdots$. 设若存在正数 M, 使得 $|a_{ij}^{(m)}|<M$, $i,j=1,2,\cdots,n$, $m=1,2,\cdots$, 则称 A 为**幂有界方阵**. 试证: A 为幂有界方阵当且仅当 A 的特征根的模 $\leqslant 1$, 且模为 1 的特征根的初等因式都是一次的.

6.4.22 试证: 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 在一个不变子空间 $\mathfrak{L}_2$ 上诱导了非异线性变换; 在另外一个不变子空间 $\mathfrak{L}_1$ 上诱导了幂零线性变换 (即若干方幂后成为零线性变换, 换句话说, 它的方阵表示是幂零方阵), 且有 $\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2$ 为空间直接和.

利用这性质, 试证: 对任一 n 阶方阵 A, 存在 n 阶非异方阵 P, 使得 $P^{-1}AP = \mathrm{diag}\,(B, C)$, 其中 B 和 C 都是方阵, 前者幂零, 后者非异.

6.4.23 用归纳法定义 n 阶 $(s_0, s_1, \cdots, s_{m-1})$ **轮回方阵**如下: 当 $m = 1$ 时即为通常的 n 阶轮回方阵. 假设已定义好 $(s_0, s_1, \cdots, s_{m-2})$ 轮回方阵. 我们称 A 为 n 阶 $(s_0, s_1, \cdots, s_{m-1})$ 轮回方阵, 如果

$$A = \begin{pmatrix} B_0 & B_1 & B_2 \cdots & B_{s_{m-1}-1} \\ B_{s_{m-1}-1} & B_0 & B_1 \cdots & B_{s_{m-1}-2} \\ \vdots & \vdots & \vdots & \vdots \\ B_1 & B_2 & B_3 \cdots & B_0 \end{pmatrix},$$

其中 $B_0, B_1, \cdots, B_{s_{m-1}-1}$ 都是 $\dfrac{n}{s_{m-1}}$ 阶 $(s_0, s_1, \cdots, s_{m-2})$ 轮回方阵. 试证:

(1) A 为 $n = s_0 s_1 \cdots s_{m-1}$ 阶方阵;

(2) 存在 n 阶非异方阵 P, 使得它可同时使所有 n 阶 $(s_1, s_2, \cdots, s_m)$ 轮回方阵都相似于对角形.

6.4.24 域 $\mathbb{F}$ 上序列 $\{x_k\}_{k=1}^{\infty}$ 称为 n 阶**线性序列**, 如果存在 n 个常数 $a_1, a_2, \cdots, a_n \in \mathbb{F}$, 使得它适合**线性序列方程**:

$$x_{n+k} = a_1 x_{n+k-1} + a_2 x_{n+k-2} + \cdots + a_n x_k, \quad k = 1, 2, \cdots. \tag{$*$}$$

给定 n 个初值 $x_i = \lambda_i$, $1 \leqslant i \leqslant n$, 试将线性序列方程 $(*)$ 写成矩阵表达式, 并讨论如何用初值来求解线性序列. 当 $n = 2$ 时, 算出解的表达式. (在 $n > 2$ 时, 不用算出解的表达式.)

第七章 Jordan 标准形的应用 *

在这一章, 经常使用如下符号: 记 m 为正整数,

$$E^{(m)} = E_m, \quad N^{(m)} = N_m = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}. \tag{7.1}$$

即 E_m 为 m 阶单位方阵, N_m 为 m 阶幂零方阵. 因此, Jordan 标准形

$$J = \mathrm{diag}\,(J_1, J_2, \cdots, J_t), \quad J_k = \lambda_k E_{e_k} + N_{e_k}, \quad k = 1, 2, \cdots, t. \tag{7.2}$$

有时候, 我们更简单地将 Jordan 标准形记为

$$J = \mathrm{diag}\,(J_1, J_2, \cdots, J_t), \quad J_k = \lambda_k E_k + N_k, \quad k = 1, 2, \cdots, t, \tag{7.3}$$

这时下标不表示阶数, 默认 J_k 的阶数为 e_k.

§7.1 Jordan 标准形的几何意义 *

现在给出复方阵在相似下的 Jordan 标准形的几何解释. 考虑 n 维复线性空间 $\mathfrak{L}$. 在 $\mathfrak{L}$ 上给定一个线性变换 $\mathscr{A}$. 在 $\mathfrak{L}$ 中取定一组基后, 线性变换 $\mathscr{A}$ 便唯一地对应了一个方阵 A. 由引理 6.2.11, 线性变换 $\mathscr{A}$ 在不同基下的方阵表示互相相似. 反之, 任取一个和 n 阶方阵 A 相似的方阵 B, 则在线性空间 $\mathfrak{L}$ 中存在另外一组基, 使得线性变换 $\mathscr{A}$ 在新的基下对应的方阵为 B. 所以在相似关系下, n 阶复方阵的每个

等价类都由一个线性变换在所有基下的方阵表示构成. 而且我们可以取代表元素为 Jordan 标准形. 为了了解 Jordan 标准形的几何意义, 我们将定理 6.4.6 确切地写为

定理 7.1.1 设 n 阶复方阵 A 的所有不同特征根为 $\lambda_1, \lambda_2, \cdots, \lambda_t$, 所有初等因子为

$$(\lambda-\lambda_1)^{e_{1i_1}}, \quad i_1=1,2,\cdots,s_1, \quad \cdots, \quad (\lambda-\lambda_t)^{e_{ti_t}}, \quad i_t=1,2,\cdots,s_t, \tag{7.1.1}$$

其中

$$e_{11}\leqslant e_{12}\leqslant\cdots e_{1s_1}, \quad \cdots, \quad e_{t1}\leqslant e_{t2}\leqslant\cdots e_{ts_t}. \tag{7.1.2}$$

记 J_{ij} 为属于初等因子 $(\lambda-\lambda_i)^{e_{ij}}$ 的 e_{ij} 阶 Jordan 块

$$J_{ij}=\lambda_i E_{ij}+N_{ij}, \quad 1\leqslant j\leqslant s_i, \quad 1\leqslant i\leqslant t, \tag{7.1.3}$$

其中 E_{ij} 为 e_{ij} 阶单位方阵, $N_{ij}=N^{(e_{ij})}$ 为 e_{ij} 阶幂零方阵 (见式 (7.2)), 则 n 阶复方阵 A 相似于 Jordan 标准形

$$J=\operatorname{diag}(B_1,B_2,\cdots,B_t), \quad B_i=\operatorname{diag}(J_{i1},J_{i2},\cdots,J_{is_i}), \quad 1\leqslant i\leqslant t. \tag{7.1.4}$$

由定理 7.1.1, 则对 n 阶方阵 A, 存在 n 阶非异方阵 P, 使得 A 相似于 Jordan 标准形的结论, 可以用几何语言叙述为: 给定 n 维复线性空间 $\mathfrak{L}$ 及 $\mathfrak{L}$ 上的线性变换 $\mathscr{A}$, 则存在一组基

$$\alpha_{ijk}, \quad 1\leqslant k\leqslant e_{ij}, \quad 1\leqslant j\leqslant s_i, \quad 1\leqslant i\leqslant t, \tag{7.1.5}$$

使得在这组基下, 线性变换 $\mathscr{A}$ 的方阵表示为 Jordan 标准形 (7.1.4).

记由 n 维复线性空间 $\mathfrak{L}$ 中的 e_{ij} 个向量

$$\alpha_{ijk}, \quad 1\leqslant k\leqslant e_{ij} \tag{7.1.6}$$

为一组基的子空间是

$$\mathfrak{L}_{ij}, \qquad 1\leqslant j\leqslant s_i, \quad 1\leqslant i\leqslant t. \tag{7.1.7}$$

由引理 6.2.11, 则可以确切地写为

定理 7.1.2 设 $\mathfrak{L}$ 为 n 维复线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. 则 $\mathfrak{L}$ 是线性变换 $\mathscr{A}$ 的不变子空间 $\mathfrak{L}_{ij}$, $1\leqslant j\leqslant s_i$, $1\leqslant i\leqslant t$ 的子空间直接和:

$$\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_2\oplus\cdots\oplus\mathfrak{L}_t, \tag{7.1.8}$$

$$\mathfrak{L}_i=\mathfrak{L}_{i1}\oplus\mathfrak{L}_{i2}\oplus\cdots\oplus\mathfrak{L}_{is_i}, \quad i=1,2,\cdots,t, \tag{7.1.9}$$

其中不变子空间 $\mathfrak{L}_{ij}$ 有一组基 (7.1.6), 使得在这组基下,

$$\begin{aligned}(\mathscr{A}-\lambda_i(\mathrm{id}))(\alpha_{ij1})&=0,\\(\mathscr{A}-\lambda_i(\mathrm{id}))(\alpha_{ij2})&=\alpha_{ij1},\\&\vdots\\(\mathscr{A}-\lambda_i(\mathrm{id}))(\alpha_{ije_{ij}})&=\alpha_{ij,e_{ij}-1}.\end{aligned}\tag{7.1.10}$$

因此,

$$(\mathscr{A}-\lambda_i(\mathrm{id}))^k(\alpha_{ijk})=0,\quad 1\leqslant k\leqslant e_{ij}.\tag{7.1.11}$$

所以

$$(\mathscr{A}-\lambda_i(\mathrm{id}))^n(\alpha)=0,\quad \forall\,\alpha\in\mathfrak{L}_{ij}.\tag{7.1.12}$$

引理 7.1.3　符号同上. 设 $k\neq i$, 则线性变换 $\mathscr{A}-\lambda_k(\mathrm{id})$ 在不变子空间 $\mathfrak{L}_i$ 上非异.

证　用反证法. 假设线性变换 $\mathscr{A}-\lambda_k(\mathrm{id})$ 在不变子空间 $\mathfrak{L}_i$ 上奇异, 则存在 $\alpha\in\mathfrak{L}_i$, $\alpha\neq 0$, 使得 $(\mathscr{A}-\lambda_k(\mathrm{id}))(\alpha)=0$. 今

$$(\mathscr{A}-\lambda_i(\mathrm{id}))(\alpha)=(\lambda_k-\lambda_i)\alpha\neq 0.$$

因此, $(\mathscr{A}-\lambda_i(\mathrm{id}))^n(\alpha)=(\lambda_k-\lambda_i)^n\alpha\neq 0$. 这和 $\alpha\in\mathfrak{L}_i$ 矛盾. □

将特征向量概念推广, 我们有

定义 7.1.4　设 $\mathfrak{L}$ 为 n 维复线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换, λ_0 为线性变换 $\mathscr{A}$ 的特征根. $\mathfrak{L}$ 中的向量 α 称为线性变换 $\mathscr{A}$ 的属于特征根 λ_0 的**根向量**, 如果存在正整数 $m(\leqslant n)$, 它依赖于向量 α, 使得.

$$(\mathscr{A}-\lambda_0(\mathrm{id}))^m\alpha=0.\tag{7.1.13}$$

显然零向量是根向量, 又属于特征根的特征向量也是根向量. 一般有

定理 7.1.5　设 $\mathfrak{L}$ 为 n 维复线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. 任取复数 λ_0, 记

$$\mathfrak{L}_{\lambda_0}=\{\,\alpha\in\mathfrak{L}\mid \text{存在正整数 } m=m(\alpha),\text{ 使得 } (\mathscr{A}-\lambda_0(\mathrm{id}))^m(\alpha)=0\,\}.\tag{7.1.14}$$

则 $\mathfrak{L}_{\lambda_0}$ 是线性空间 $\mathfrak{L}$ 中线性变换 $\mathscr{A}$ 的不变子空间.

又不变子空间 $\mathfrak{L}_{\lambda_0}\neq 0$ 当且仅当 λ_0 是线性变换 $\mathscr{A}$ 的特征根. 所以当 λ_0 为线性变换 $\mathscr{A}$ 的特征根时, 不变子空间 $\mathfrak{L}_{\lambda_0}$ 称为属于特征根 λ_0 的**根子空间**.

证　任取 $\alpha,\beta\in\mathfrak{L}_{\lambda_0}$, 于是存在正整数 m, 使得

$$(\mathscr{A}-\lambda_0(\mathrm{id}))^m(\alpha)=(\mathscr{A}-\lambda_0(\mathrm{id}))^m(\beta)=0.$$

任取纯量 $c, d \in \mathbb{C}$, 于是

$$(\mathscr{A} - \lambda_0(\mathrm{id}))^m(c\alpha + d\beta) = c(\mathscr{A} - \lambda_0(\mathrm{id}))^m(\alpha) + d(\mathscr{A} - \lambda_0(\mathrm{id}))^m(\beta) = 0.$$

这证明了 $c\alpha + d\beta \in \mathfrak{L}_{\lambda_0}$, 即 $\mathfrak{L}_{\lambda_0}$ 为线性空间 $\mathfrak{L}$ 的子空间. 由 $\mathscr{A}(\mathscr{A} - \lambda_0(\mathrm{id}))^m = (\mathscr{A} - \lambda_0(\mathrm{id}))^m\mathscr{A}$ 可知 $\mathfrak{L}_{\lambda_0}$ 为线性空间 $\mathfrak{L}$ 的不变子空间.

今 $\mathfrak{L}_{\lambda_0} \neq 0$ 当且仅当存在 $\alpha \in \mathfrak{L}_{\lambda_0}$, $\alpha \neq 0$. 于是存在正整数 m, 使得 $(\mathscr{A} - \lambda_0(\mathrm{id}))^m(\alpha) = 0$. 因此存在正整数 p, 使得 $\beta = (\mathscr{A} - \lambda_0(\mathrm{id}))^{p-1}(\alpha) \neq 0$, $(\mathscr{A} - \lambda_0(\mathrm{id}))^p(\alpha) = 0$. 这证明了存在 $0 \neq \beta \in \mathfrak{L}_{\lambda_0}$, 而 $(\mathscr{A} - \lambda_0(\mathrm{id}))\beta = 0$, 即 $\mathscr{A}(\beta) = \lambda_0\beta$. 所以 β 是属于特征根 λ_0 的特征向量. 因此 λ_0 是线性变换 $\mathscr{A}$ 的特征根. □

由定理 7.1.2 和定理 7.1.5, 我们有

定理 7.1.6 (空间第一分解定理) 设 $\mathfrak{L}$ 为 n 维复线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. 记 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 为线性变换 $\mathscr{A}$ 的所有不同特征根, 则线性空间 $\mathfrak{L}$ 有根子空间直接和分解:

$$\mathfrak{L} = \mathfrak{L}_{\lambda_1} \oplus \mathfrak{L}_{\lambda_2} \oplus \cdots \oplus \mathfrak{L}_{\lambda_t}, \tag{7.1.15}$$

其中

$$\mathfrak{L}_{\lambda_i} = \mathfrak{L}_i, \quad 1 \leqslant i \leqslant t, \tag{7.1.16}$$

这里 $\mathfrak{L}_i$ 由定理 7.1.2 所定义. 且线性变换 $\mathscr{A}$ 在根子空间 $\mathfrak{L}_{\lambda_i}$ 上有且只有一个特征根 λ_i, $1 \leqslant i \leqslant t$. 又不计次序, 线性空间 $\mathfrak{L}$ 的根子空间直接和分解唯一存在.

证 先证 $\mathfrak{L}_{\lambda_i} = \mathfrak{L}_i$, $1 \leqslant i \leqslant t$. 事实上, 由式 (7.1.12) 可知 $\mathfrak{L}_i \subset \mathfrak{L}_{\lambda_i}$. 反之, 任取 $\alpha \in \mathfrak{L}_{\lambda_i}$, 则由

$$\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_t,$$

有 $\alpha = \sum\limits_{j=1}^{t} \alpha_j$, 其中 $\alpha_j \in \mathfrak{L}_j$. 今 $\mathfrak{L}_j$ 是 $\mathscr{A}$ 的不变子空间, 所以 $\mathfrak{L}_j$ 也是 $(\mathscr{A} - \lambda_i(\mathrm{id}))^n$ 的不变子空间. 于是 $(\mathscr{A} - \lambda_i(\mathrm{id}))^n(\alpha) = 0$ 推出 $(\mathscr{A} - \lambda_i(\mathrm{id}))^n(\alpha_j) = 0$, $1 \leqslant j \leqslant t$. 由引理 7.1.3, 假设 $j \neq i$, 则线性变换 $\mathscr{A} - \lambda_i(\mathrm{id})$ 在 $\mathfrak{L}_{\lambda_j}$ 上非异, 所以线性变换 $(\mathscr{A} - \lambda_i(\mathrm{id}))^n$ 在 $\mathfrak{L}_{\lambda_j}$ 上非异. 这证明了线性变换 $(\mathscr{A} - \lambda_i(\mathrm{id}))^n$ 在 $\mathfrak{L}_j$ 上非异. 因此 $\alpha_j = 0$, $j \neq i$, 即 $\alpha = \alpha_i \in \mathfrak{L}_i$, 所以 $\mathfrak{L}_{\lambda_i} \subset \mathfrak{L}_i$. 因此证明了线性空间 $\mathfrak{L}_{\lambda_i} = \mathfrak{L}_i$, $1 \leqslant i \leqslant t$. 所以也证明了线性空间 $\mathfrak{L}$ 是根子空间的直接和.

再证线性空间 $\mathfrak{L}$ 中属于特征根 λ_i 的特征向量 $\beta \in \mathfrak{L}_i$. 事实上, 设 $\mathscr{A}(\beta) = \lambda_i\beta$, $\beta \neq 0$. 今 $\beta = \sum\limits_{j=1}^{t} \beta_j$, 其中 $\beta_j \in \mathfrak{L}_j$. 于是 $(\mathscr{A} - \lambda_i(\mathrm{id}))(\beta) = \sum\limits_{j=1}^{t}(\mathscr{A} - \lambda_i(\mathrm{id}))(\beta_j) = 0$. 所以 $(\mathscr{A} - \lambda_i(\mathrm{id}))(\beta_j) = 0$, $1 \leqslant j \leqslant t$. 假设 $j \neq i$, 则线性变换 $\mathscr{A} - \lambda_i(\mathrm{id})$ 在 $\mathfrak{L}_j$ 上非异. 因此 $\beta_j = 0$, $j \neq i$. 这证明了 $\beta = \beta_i \in \mathfrak{L}_i$.

由 Jordan 标准形的唯一性和上面的构造过程便知, 线性空间 $\mathfrak{L}$ 的根子空间直接和分解, 不计次序是唯一的. □

回到定理 7.1.2. 在取定 $i \in \{1, 2, \cdots, t\}$ 后, 我们需要给出不变子空间 $\mathfrak{L}_i$ 的直**接和分解**:

$$\mathfrak{L}_{\lambda_i} = \mathfrak{L}_i = \mathfrak{L}_{i1} \oplus \mathfrak{L}_{i2} \oplus \cdots \oplus \mathfrak{L}_{is_i} \tag{7.1.17}$$

的几何解释. 由定理 7.1.5, 为了方便起见, 只考虑线性变换 $\mathscr{B} = \mathscr{A} - \lambda_i(\mathrm{id})$ 在不变子空间 $\mathfrak{L}_{\lambda_i}$ 上的作用就可以了. 这时不变子空间 $\mathfrak{L}_{\lambda_i}$ 可以表示为

$$\mathfrak{L}_{\lambda_i} = \{\, \alpha \in \mathfrak{L} \mid \text{存在正整数 } m = m(\alpha), \text{ 使得 } \mathscr{B}^m(\alpha) = 0 \,\}. \tag{7.1.18}$$

即 $\mathscr{B}$ 在线性空间 $\mathfrak{L}_i$ 上是幂零线性变换.

定义 7.1.7　设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. 线性变换 $\mathscr{A}$ 称为**幂零线性变换**, 如果存在正整数 m, 使得 $\mathscr{A}^m = 0$, $\mathscr{A}^{m-1} \neq 0$. 正整数 m 称为幂零线性变换的**幂零次数**.

显然, 零线性变换为幂零次数为零的幂零线性变换. 又若设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. $\mathfrak{L}_0$ 为线性变换 $\mathscr{A}$ 的关于特征根零的根子空间, 则 $\mathscr{A}$ 限制在 $\mathfrak{L}_0$ 上为幂零线性变换. 又显然, 线性变换 $\mathscr{A}$ 幂零当且仅当 $\mathscr{A}$ 只有零特征根.

定理 7.1.8 (空间第二分解定理)　设 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的幂零线性变换. 则线性空间 $\mathfrak{L}$ 可分解为循环子空间直接和, 且如不计次序, 则在线性同构意义下分解唯一.

证　由于幂零线性变换 $\mathscr{A}$ 的 Jordan 标准形为 $\mathrm{diag}\,(N_1, N_2, \cdots, N_s)$, 其中 N_j 为式 (7.1) 定义的 e_j 阶幂零方阵. 于是在线性空间 $\mathfrak{L}$ 中存在一组基

$$\alpha_{jk}, \quad 1 \leqslant k \leqslant e_j, \quad 1 \leqslant j \leqslant s, \tag{7.1.19}$$

使得在这组基下, 线性变换 $\mathscr{A}$ 的方阵表示为 Jordan 标准形 (7.1.4), 其中 $t = 1$, $s_1 = s$, $e_{1j} = e_j$, $J_{1j} = N_j$. 因此

$$\mathscr{A}(\alpha_{j1}) = 0, \quad \mathscr{A}(\alpha_{j2}) = \alpha_{j1}, \quad \cdots, \quad \mathscr{A}(\alpha_{je_j}) = \alpha_{j,e_j-1}. \tag{7.1.20}$$

所以以 $\alpha_{j1}, \alpha_{j2}, \cdots, \alpha_{je_j}$ 为基的子空间是由 α_{je_j} 生成的循环子空间 $\mathfrak{L}_j$. 而

$$\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_s. \tag{7.1.21}$$

因此也证明了 $\mathfrak{L}$ 是幂零线性变换 $\mathscr{A}$ 的循环子空间的直接和. 由 Jordan 标准形的唯一性和上面的构造过程便知, 线性空间 $\mathfrak{L}$ 的循环子空间直接和分解, 不计次序, 在线

性同构意义下分解是唯一的. □

将上面两个空间分解定理合在一起, 注意到线性变换 $\mathscr{A}-\lambda_i(\mathrm{id})$ 在根子空间 $\mathfrak{L}_{\lambda_i}$ 上为幂零线性变换, 所以证明了

定理 7.1.9 设 $\mathfrak{L}$ 为 n 维复线性空间, $\mathscr{A}$ 为线性空间 $\mathfrak{L}$ 的线性变换. $\lambda_1, \lambda_2, \cdots, \lambda_t$ 为线性变换 $\mathscr{A}$ 的全部不同特征根, 则 n 维复线性空间 $\mathfrak{L}$ 有根子空间直接和分解

$$\mathfrak{L}=\mathfrak{L}_{\lambda_1}\oplus\mathfrak{L}_{\lambda_2}\oplus\cdots\oplus\mathfrak{L}_{\lambda_t}, \tag{7.1.22}$$

分解不计次序唯一. 而每个根子空间 $\mathfrak{L}_{\lambda_i}$ 分解为关于幂零线性变换 $\mathscr{A}-\lambda_i(\mathrm{id})$ 的循环子空间直接和

$$\mathfrak{L}_{\lambda_i}=\mathfrak{L}_{i1}\oplus\mathfrak{L}_{i2}\oplus\cdots\oplus\mathfrak{L}_{is_i}, \tag{7.1.23}$$

且此分解若不计次序, 则在同构意义下唯一, $i=1, 2, \cdots, t$.

上面我们利用了 Jordan 标准形, 导出了定理 7.1.9. 我们也可以不利用初等因子, 而利用几何方法直接证明定理 7.1.9, 从而导出 Jordan 标准形. 证明的思路是先证定理 7.1.6, 再证定理 7.1.8, 从而得到定理 7.1.9. 再在每个循环子空间中取定一组循环基, 使得全体合并构成 n 维复线性空间 $\mathfrak{L}$ 的一组基. 于是在这组基下, 线性变换对应的方阵就是 Jordan 标准形. 但是, 用几何方法还不能给出如何计算 Jordan 标准形以及计算方阵在相似关系下的全系不变量. 求全系不变量一定要依靠 λ 矩阵在相抵下的标准形理论.

习　题　7.1

7.1.1 设 $\mathscr{A}$ 为域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性变换. λ_j 为线性变换 $\mathscr{A}$ 的 e_j 重特征根. 试证: 关于特征根 λ_j 的根子空间 $\mathfrak{L}_{\lambda_j}$ 可定义为

$$\mathfrak{L}_{\lambda_j}=\{\,\alpha\in\mathfrak{L}\mid(\mathscr{A}-\lambda_j(\mathrm{id}))^{e_j}(\alpha)=0\,\}.$$

7.1.2 试证: 两两可交换的 $2n+1$ 阶实方阵必有公共的实特征向量.

7.1.3 试给出下面一系列引理和定理的证明:

定理 1 设 $\mathscr{A}$ 为 n 维复线性空间 $\mathfrak{L}$ 的线性变换. 任取复数 λ_0, 记

$$\mathfrak{L}_{\lambda_0}=\{\,\alpha\in\mathfrak{L}\mid \text{存在正整数} m=m(\alpha),\ \text{使得}\ (\mathscr{A}-\lambda_0(\mathrm{id}))^m\alpha=0\,\}. \tag{7.1.24}$$

则有:

(1) $\mathfrak{L}_{\lambda_0}$ 为 $\mathfrak{L}$ 的子空间;

(2) $\mathscr{A}(\mathfrak{L}_{\lambda_0})\subset\mathfrak{L}_{\lambda_0}$, 即 $\mathfrak{L}_{\lambda_0}$ 为 $\mathscr{A}$ 的不变子空间;

(3) $\mathfrak{L}_{\lambda_0}\neq 0$ 当且仅当 λ_0 为 $\mathscr{A}$ 的特征根.

所以当 λ_0 为线性变换 $\mathscr{A}$ 的特征根时, 则 $\mathfrak{L}_{\lambda_0}$ 由所有属于特征根 λ_0 的根向量构成, 所以称为特征根 λ_0 的**根子空间**.

显然, 设 $\mathfrak{L}_{\lambda_0}$ 为 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 关于特征根 λ_0 的根子空间, 则线性变换 $\mathscr{A}-\lambda_0(\mathrm{id})$ 限制在 $\mathfrak{L}_{\lambda_0}$ 上为幂零线性变换.

引理 2 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 的线性变换, 则 $\mathfrak{L}$ 有不变子空间序列

$$\{0\} \subset \cdots \subset \mathscr{A}^2(\mathfrak{L}) \subset \mathscr{A}(\mathfrak{L}) \subset \mathfrak{L}. \tag{7.1.25}$$

所以存在正整数 q, 使得

$$\mathscr{A}^{q-1}(\mathfrak{L}) \neq \mathscr{A}^{q}(\mathfrak{L}) = \mathscr{A}^{q+1}(\mathfrak{L}) = \cdots = \bigcap_{m\geqslant 1} \mathscr{A}^{m}(\mathfrak{L}). \tag{7.1.26}$$

当 $\mathscr{A}$ 为幂零线性变换时, 则 q 为幂零次数; 当 $\mathscr{A}$ 不是幂零线性变换时, 则 $\mathscr{A}$ 限制在 $\mathscr{A}^q(\mathfrak{L})$ 上为非异线性变换.

引理 3 (Fitting 引理) 符号同上, 记

$$\widetilde{\mathfrak{L}} = \bigcap_{m\geqslant 1} \mathscr{A}^{m}(\mathfrak{L}), \qquad \mathfrak{L}_0 = \bigcup_{m\geqslant 1} \ker(\mathscr{A}^m). \tag{7.1.27}$$

则线性空间 $\mathfrak{L}$ 有 $\mathscr{A}$ 的不变子空间直和分解:

$$\mathfrak{L} = \widetilde{\mathfrak{L}} \oplus \mathfrak{L}_0. \tag{7.1.28}$$

当 $\mathfrak{L}_0 \neq 0$ 时, $\mathscr{A}|_{\mathfrak{L}_0}$ 为幂零线性变换; 当 $\widetilde{\mathfrak{L}} \neq 0$ 时, $\mathscr{A}|_{\widetilde{\mathfrak{L}}}$ 为非异线性变换.

定理 4 设 $\mathscr{A}$ 为 n 维复线性空间 $\mathfrak{L}$ 的线性变换. 记 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 为 $\mathscr{A}$ 的所有不同特征根, 则 $\mathfrak{L}$ 有根子空间直接和分解:

$$\mathfrak{L} = \mathfrak{L}_{\lambda_1} \oplus \mathfrak{L}_{\lambda_2} \oplus \cdots \oplus \mathfrak{L}_{\lambda_r}, \tag{7.1.29}$$

且不计次序, 分解唯一.

注意 可以有两种证法. 一是考虑 $\det(\lambda E_n - A) = (\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_r)^{m_r}$, 则 $(\lambda-\lambda_1)^{-m_j}\det(\lambda E_n - A)$, $j=1, 2, \cdots, r$ 互素; 另一是用 Fitting 引理.

定理 5 设 A 为 n 阶复方阵, $\lambda_1, \lambda_2, \cdots, \lambda_r$ 为 A 的所有不同特征根. 则 A 相似于准对角阵 $\mathrm{diag}(A_1, A_2, \cdots, A_r)$, 其中 $A_j - \lambda_j E$ 为幂零方阵, 所以 A_j 恰有一个特征根 λ_j, $j = 1, 2, \cdots, r$.

为了进一步将 n 阶复方阵 A 在相似下化简, 问题化为如何进一步在相似下化简复方阵 A_1, A_2, $\cdots$, A_r. 换句话说, 只要考虑线性空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$, 使得 $\mathscr{A}$ 在 $\mathfrak{L}$ 上恰有一个特征根 λ_0. 所以 $\mathscr{A}-\lambda_0(\mathrm{id})$ 在 $\mathfrak{L}$ 上幂零, 即化为考虑在 $\mathfrak{L}$ 上的幂零线性变换作用下, 如何将线性空间分解为若干个不变子空间的直接和. 为此, 先引进一般的循环子空间的概念.

定理 6 设 $\mathscr{A}$ 为 n 维线性空间 $\mathfrak{L}$ 的幂零线性变换, 则 $\mathfrak{L}$ 可分解为循环子空间直接和, 且如不计次序, 则在线性同构意义下分解唯一.

注意 证明可以用对线性空间 $\mathfrak{L}$ 的维数作归纳法, 仔细考虑循环基和 $\ker(\mathscr{A})$. 证明分解在同构意义下如不计次序则唯一, 仍用对线性空间 $\mathfrak{L}$ 的维数作归纳法.

将上面两个空间分解定理合在一起, 注意到 $\mathscr{A}-\lambda_j(\mathrm{id})$ 在根子空间 $\mathfrak{L}_{\lambda_j}$ 上为幂零线变换, 所以证明了

定理 7 设 $\mathscr{A}$ 为 n 维复线性空间 $\mathfrak{L}$ 的线性变换. 记 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 为 $\mathscr{A}$ 的所有不同特征根, 则 $\mathfrak{L}$ 有根子空间直接和分解式 (7.1.29), 而且分解不计次序则唯一. 而每个根子空间 $\mathfrak{L}_{\lambda_j}$ 又可分解为关于幂零线性变换 $\mathscr{A}-\lambda_j(\mathrm{id})$ 的循环子空间直接和:

$$\mathfrak{L}_{\lambda_j}=\mathfrak{L}_{j1}\oplus\mathfrak{L}_{j2}\oplus\cdots\oplus\mathfrak{L}_{js_j},\quad j=1,2,\cdots,r, \tag{7.1.30}$$

且此分解若不计次序, 则在同构意义下唯一.

进一步, 试给出上面定理的矩阵形式, 即给出复方阵在相似下的 Jordan 标准形.

§7.2 Jordan 标准形的应用 *

(一) 实方阵在实相似下的标准形

在上一章, 我们给出了复方阵在相似下的 Jordan 标准形. 现在我们给出实方阵在实相似下的标准形. 这时我们不能给出和复的情形一样的结论. 其原因在于实方阵的特征根不一定是实数. 如果它有复根, 那末 Jordan 块就是一个复方阵. 然而我们要求的标准形应该是实方阵. 尽管如此, 我们仍然可以借助于复的结果来给出实方阵在实相似下的标准形. 这是基于下面的引理 7.2.3.

定义 7.2.1 两个 n 阶实方阵 A 和 B 称为**实相似的**, 如果存在 n 阶非异实方阵 P, 使得

$$B=P^{-1}AP. \tag{7.2.1}$$

显然有

引理 7.2.2 实方阵的实相似关系为等价关系.

定理 7.2.3 n 阶实方阵 A 和 B 实相似的充分且必要条件是 A 和 B 相似.

证 由于实相似为相似, 所以只要证明当 A 和 B 相似, 则必实相似. 假设存在 n 阶非异复方阵 P_1, 使得 $B=P_1^{-1}AP_1$. 将 P_1 的实部和虚部分开, 即记 $P_1=P+\sqrt{-1}Q$, 其中 P 和 Q 都是 n 阶实方阵, 于是 $B=(P+\sqrt{-1}Q)^{-1}A(P+\sqrt{-1}Q)$. 即 $(P+\sqrt{-1}Q)B=A(P+\sqrt{-1}Q)$. 由于 A 和 B, P 和 Q 都是实方阵, 此即 $PB=AP$, $QB=AQ$. 所以任取实数 t, 则有 $(P+tQ)B=A(P+tQ)$. 这样, 问题就化为要证明存在实数 t_0, 使得 n 阶实方阵 $P+t_0Q$ 非异.

视 t 为实自变量, 则

$$\det(P+tQ)=a_0t^n+a_1t^{n-1}+\cdots+a_{n-1}t+a_n$$

为 t 的实多项式, 已知 $\det(P_1)=\det(P+\sqrt{-1}Q)\neq 0$, 所以证明了 $\det(P+tQ)$ 为非零实多项式, 次数不超过 n. 它至多有 n 个不同实根, 所以一定存在实数 t_0, 使得

t_0 不是非零多项式 $\det(P+tQ)$ 的根. 即有 $\det(P+t_0Q)\neq 0$, 所以 $P+t_0Q$ 为 n 阶非异实方阵, 且有 $A(P+t_0Q)=(P+t_0Q)B$. 这证明了实方阵 A 和 B 实相似. □

引理 7.2.4 设 A 为 n 阶实方阵, 则 A 的初等因式组中若有 $(\lambda-\lambda_0)^e$, 则必有 $(\lambda-\overline{\lambda_0})^e$.

证 令 A 为 n 阶实方阵, 则 $\lambda E_n - A$ 是由实系数多项式为元素构成的 λ 方阵. 对 λ 方阵 $\lambda E_n - A$ 作初等变换化为对角形 $\mathrm{diag}\,(d_1(\lambda), d_2(\lambda),\cdots,d_n(\lambda))$, 其中对角元素 $d_1(\lambda)$, $d_2(\lambda)$, $\cdots$, $d_n(\lambda)$ 为不变因式组. 由初等变换的作法可知, 当 A 为实方阵, 则不变因式组全是实系数多项式. 这证明了初等因式组中若有 $(\lambda-\lambda_0)^e$, 则必有 $(\lambda-\overline{\lambda_0})^e$. □

定理 7.2.5 设 A 为 n 阶实方阵, 则 A 的初等因式组为

$$\begin{aligned}&\lambda^{e_1},\ \cdots,\ \lambda^{e_s},\ (\lambda-\lambda_1)^{f_1},\ \cdots,\ (\lambda-\lambda_t)^{f_t},\\&(\lambda-\tau_1)^{g_1},\ \cdots,\ (\lambda-\tau_p)^{g_p},\ (\lambda-\overline{\tau_1})^{g_1},\ \cdots,\ (\lambda-\overline{\tau_p})^{g_p},\end{aligned}\tag{7.2.2}$$

其中 λ_1, λ_2, $\cdots$, λ_t 为 A 的非零实特征根, τ_1, $\overline{\tau_1}$, $\cdots$, τ_p, $\overline{\tau_p}$ 为 A 的复特征根, 又 τ_1, $\cdots$, τ_p 的虚部都大于零. 这时 A 有**实 Jordan 标准形**

$$\begin{aligned}J=\mathrm{diag}\,(&N_{e_1},\cdots,N_{e_s},\lambda_1E_{f_1}+N_{f_1},\cdots,\lambda_tE_{f_t}+N_{f_t},\\&r_1L(\theta_1)^{(2g_1)},\cdots,r_pL(\theta_p)^{(2g_p)}),\end{aligned}\tag{7.2.3}$$

其中 $\tau_j=r_j\exp(\sqrt{-1}\theta_j)$, $r_j=|\tau_j|$, $\theta_j=\arg(\tau_j)$, $1\leqslant j\leqslant p$, 又

$$r_jL(\theta_j)^{(2g_j)}=\begin{pmatrix}M_{g_j}\cos(\theta_j) & M_{g_j}\sin(\theta_j)\\ -M_{g_j}\sin(\theta_j) & M_{g_j}\cos(\theta_j)\end{pmatrix},\quad 1\leqslant j\leqslant p,\tag{7.2.4}$$

其中 $M_{g_j}=E_{g_j}+N_{g_j}$ 为 g_j 阶实方阵.

证 由引理 7.2.4, 实方阵 A 的初等因式组如定理中式 (7.2.2) 所述. 因此 A 相似于 Jordan 标准形

$$\begin{aligned}J=\mathrm{diag}\,(&N_{e_1},\cdots,N_{e_s},\lambda_1E_{f_1}+N_{f_1},\cdots,\lambda_tE_{f_t}+N_{f_t},\\&\tau_1E_{g_1}+N_{g_1},\overline{\tau_1}E_{g_1}+N_{g_1},\cdots,\tau_pE_{g_p}+N_{g_p},\overline{\tau_p}E_{g_p}+N_{g_p}),\end{aligned}$$

由定理 7.2.3, 我们只要证明任取 $\tau_0=r\exp(\sqrt{-1}\theta)$ 不是实数, 其中 $r=|\tau_0|$, $\theta=\arg(\tau_0)$, E 和 N 都由式 (7.1) 所定义的 g 阶实方阵, $L(\theta)$ 是 $2g$ 阶实方阵, 又 $\mathrm{diag}\,(\tau_0(E+N),\overline{\tau_0}(E+N))$ 和 $rL(\theta)$ 相似就行了. 事实上, 取

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

显然 $\det(P)\neq 0$, 而且 $rL(\theta)$ 实相似于 $P(rL(\theta))P^{-1}$, 其中

$$P(rL(\theta))P^{-1}=\operatorname{diag}(\tau_0 M_g,\overline{\tau_0}M_g),\quad M_g=E_g+N_g.$$

这是因为计算

$$r\begin{pmatrix} E_g & -\sqrt{-1}E_g\\ E_g & \sqrt{-1}E_g\end{pmatrix}\begin{pmatrix} M_g\cos(\theta) & M_g\sin(\theta)\\ -M_g\sin(\theta) & M_g\cos(\theta)\end{pmatrix}\begin{pmatrix} E_g & -\sqrt{-1}E_g\\ E_g & \sqrt{-1}E_g\end{pmatrix}^{-1}$$

即可知. □

(二) 和复方阵可交换的复方阵

由于下一节的需要, 下面我们定出所有和 n 阶复方阵 A 可交换的 n 阶复方阵. 假设 n 阶复方阵 A 和 B_0 可交换, 即 $AB_0=B_0A$. 已知 A 有 Jordan 标准形 J, 所以存在 n 阶非异复方阵 P, 使得 $A=P^{-1}JP$. 因此 $P^{-1}JPB_0=B_0P^{-1}JP$. 所以 $JPB_0P^{-1}=PB_0P^{-1}J$. 记 $B=PB_0P^{-1}$, 则有 $JB=BJ$. 这证明了为了定出所有和 n 阶复方阵 A 可交换的 n 阶复方阵 B_0, 问题化为定出所有和 Jordan 标准形 J 可交换的复方阵 B 就可以了.

定理 7.2.6 给定 Jordan 标准形

$$J=\operatorname{diag}(J_1,J_2,\cdots,J_t),\quad J_k=\lambda_k E_k^{(e_k)}+N_k^{(e_k)},\quad 1\leqslant k\leqslant t. \tag{7.2.5}$$

设 n 阶复方阵 B 和 Jordan 标准形 J 可交换. 将 B 和 J 一样分块为

$$B=\begin{pmatrix} B_{11} & \cdots & B_{1t}\\ \vdots & & \vdots\\ B_{t1} & \cdots & B_{tt}\end{pmatrix}, \tag{7.2.6}$$

则有

(1) 当 $\lambda_j\neq\lambda_k$ 时,

$$B_{jk}=0; \tag{7.2.7}$$

(2) 当 $\lambda_j=\lambda_k$, $e_j\geqslant e_k$ 时,

$$B_{jk}=\begin{pmatrix}\begin{pmatrix} b_0^{(jk)} & b_1^{(jk)} & \cdots & b_{e_k-1}^{(jk)}\\ & \ddots & \ddots & \vdots\\ & & b_0^{(jk)} & b_1^{(jk)}\\ & & & b_0^{(jk)}\end{pmatrix}\\ 0\end{pmatrix}=\begin{pmatrix}\sum\limits_{i=0}^{e_k-1} b_i^{(jk)}N_k^i\\ 0\end{pmatrix}, \tag{7.2.8}$$

其中 $b_0^{(jk)}, b_1^{(jk)}, \cdots, b_{e_k-1}^{(jk)}$ 是任意复数;

(3) 当 $\lambda_j = \lambda_k$, $e_j < e_k$ 时,

$$B_{jk} = \begin{pmatrix} 0 & \begin{pmatrix} b_0^{(jk)} & b_1^{(jk)} & \cdots & b_{e_j-1}^{(jk)} \\ & \ddots & \ddots & \vdots \\ & & b_0^{(jk)} & b_1^{(jk)} \\ & & & b_0^{(jk)} \end{pmatrix} \end{pmatrix} = \left(0 \quad \sum_{i=0}^{e_j-1} b_i^{(jk)} N_j^i \right), \tag{7.2.9}$$

其中 $b_0^{(jk)}, b_1^{(jk)}, \cdots, b_{e_j-1}^{(jk)}$ 是任意复数.

证　由 $JB = BJ$, 所以有 $J_j B_{jk} = B_{jk} J_k$, $1 \leqslant j, k \leqslant t$. 记

$$B_{jk} = \begin{pmatrix} a_{11} & \cdots & a_{1e_k} \\ \vdots & & \vdots \\ a_{e_j1} & \cdots & a_{e_je_k} \end{pmatrix},$$

则有

$$\begin{pmatrix} \lambda_j a_{11} + a_{21} & \cdots & \lambda_j a_{1e_k} + a_{2e_k} \\ \vdots & & \vdots \\ \lambda_j a_{e_j-1,1} + a_{e_j1} & \cdots & \lambda_j a_{e_j-1,e_k} + a_{e_je_k} \\ \lambda_j a_{e_j1} & \cdots & \lambda_j a_{e_je_k} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_k a_{11} & a_{11} + \lambda_k a_{12} & \cdots & a_{1,e_k-1} + \lambda_k a_{1e_k} \\ \vdots & \vdots & & \vdots \\ \lambda_k a_{e_j1} & a_{e_j1} + \lambda_k a_{e_j2} & \cdots & a_{e_j,e_k-1} + \lambda_k a_{e_je_k} \end{pmatrix}.$$

所以

$$\begin{pmatrix} (\lambda_j - \lambda_k)a_{11} + a_{21} & (\lambda_j - \lambda_k)a_{12} + a_{22} - a_{11} & \cdots & (\lambda_j - \lambda_k)a_{1e_k} + a_{2,e_k} - a_{1,e_k-1} \\ \vdots & \vdots & & \vdots \\ (\lambda_j - \lambda_k)a_{e_j-1,1} + a_{e_j1} & (\lambda_j - \lambda_k)a_{e_j-1,2} + a_{e_j2} - a_{e_j-1,1} & \cdots & (\lambda_j - \lambda_k)a_{e_j-1,e_k} + a_{e_j,e_k} - a_{e_j-1,e_k-1} \\ (\lambda_j - \lambda_k)a_{e_j1} & (\lambda_j - \lambda_k)a_{e_j2} - a_{e_j1} & \cdots & (\lambda_j - \lambda_k)a_{e_j,e_k} - a_{e_j,e_k-1} \end{pmatrix}$$

为零矩阵.

当 $\lambda_j \neq \lambda_k$, 看最后一行, 则有 $a_{e_j1} = 0$. 所以 $a_{e_j2} = 0$. 依次有 $a_{e_j1} = 0$, $a_{e_j2} = 0, \cdots, a_{e_je_k} = 0$. 再从最后一行依次考察到第一行, 便证明了 $B_{jk} = 0$.

当 $\lambda_j = \lambda_k$, 则有

$$\begin{pmatrix} a_{21} & a_{22} - a_{11} & \cdots & a_{2,e_k} - a_{1,e_k-1} \\ \vdots & \vdots & & \vdots \\ a_{e_j1} & a_{e_j2} - a_{e_j-1,1} & \cdots & a_{e_j,e_k} - a_{e_j-1,e_k-1} \\ 0 & -a_{e_j1} & \cdots & -a_{e_j,e_k-1} \end{pmatrix} = 0.$$

这时, 看第一列, 则有 $a_{21}=0, \cdots, a_{e_j1}=0$. 从第一列依次考察到最后一列. 于是有 $a_{11}=a_{22}, a_{32}=\cdots=a_{e_j2}=0, \cdots$.

再看最后一行, 则有 $a_{e_j1}=0, \cdots, a_{e_j,e_k-1}=0$. 再从最后一行依次考察到第一行, 则有 $a_{pq}=a_{p-1,q-1}, p=2,\cdots,e_j, q=2,\cdots,e_k$.

分别在 $e_j \geqslant e_k$ 和 $e_j<e_k$ 这两种情形来考察, 便证明了定理. □

引理 7.2.7 (广义 Lagrange 插值公式) 任给 rm 个复数 a_{jk}, $1\leqslant j\leqslant r$, $0\leqslant k<m$, 任给 r 个不同复数 $a_1, a_2, \cdots, a_r$, 则存在多项式 $g(x)$, 使得

$$g^{(k)}(a_j)=a_{jk}, \quad 1\leqslant j\leqslant r, \quad 0\leqslant k<m. \tag{7.2.10}$$

证 记

$$\Phi_j(\lambda)=(\lambda-a_1)^m\cdots(\lambda-a_{j-1})^m(\lambda-a_{j+1})^m\cdots(\lambda-a_r)^m. \tag{7.2.11}$$

令

$$\mu_{j0}=\frac{a_{j0}}{\Phi_j(a_j)}, \quad j=1,2,\cdots,r, \tag{7.2.12}$$

作递推公式:

$$\mu_{jk}\Phi_j(a_j)+\binom{k}{1}\mu_{j,k-1}\Phi_j^{(1)}(a_j)+\cdots+\binom{k}{k}\mu_{j,0}\Phi_j^{(k)}(a_j)=a_{jk}, \quad k=1,2,\cdots,m, \tag{7.2.13}$$

则决定了 $\mu_{j0}, \mu_{j1}, \mu_{j2}, \cdots, \mu_{jm}$. 构造多项式

$$\varphi_j(x)=\mu_{j0}+\frac{1}{1!}\mu_{j1}(x)(x-a_j)+\cdots+\frac{1}{(m-1)!}\mu_{j,m-1}(x-a_j)^{m-1}. \tag{7.2.14}$$

取 $1\leqslant i\leqslant r$, 令 $g(x)=\sum\limits_{j=1}^{r}\Phi_j(x)\varphi_j(x)$. 我们来证明多项式 $g(x)$ 适合引理的条件.

取 $1\leqslant i\leqslant r$, $0\leqslant k<m$, 则有

$$g^{(k)}(a_i)=\left.\frac{\mathrm{d}^k g(x)}{\mathrm{d}x^k}\right|_{x=a_i}=\sum_{j=1}^{r}\left.\frac{\mathrm{d}^k(\Phi_j(x)\varphi_j(x))}{\mathrm{d}x^k}\right|_{x=a_i}. \tag{7.2.15}$$

令取 $0\leqslant u\leqslant m$, 则

$$\left.\frac{\mathrm{d}^u\Phi_j(x)}{\mathrm{d}x^u}\right|_{x=a_i}=\left.\frac{\mathrm{d}^u\prod\limits_{v\neq j}(x-a_v)^m}{\mathrm{d}x^u}\right|_{x=a_i}=\left.\prod_{v\neq j}(x-a_v)^{m-u}\xi_{ju}(x)\right|_{x=a_i},$$

其中 $\xi_{ju}(x)$ 为 x 的多项式. 所以当 $i\neq j$, 则有

$$\left.\frac{\mathrm{d}^u\Phi_j(x)}{\mathrm{d}x^u}\right|_{x=a_i}=0, \quad 0\leqslant u\leqslant m.$$

所以

$$g^{(k)}(a_i)=\frac{\mathrm{d}^k\Phi_i(x)\varphi_i(x)}{\mathrm{d}x^k}|_{x=a_i}=\sum_{v=0}^{k}\binom{k}{v}\Phi_i^{(v)}(a_i)\varphi_i^{(k-v)}(a_i)=\sum_{v=0}^{k}\binom{k}{v}\mu_{i,k-v}\Phi_i^{(v)}(a_i).$$

由递推公式 (7.2.13), 所以 $g^{(k)}(a_i)=a_{ik}$. □

定理 7.2.8 和 n 阶复方阵 A 可交换的所有方阵可交换的方阵必为 A 的多项式.

证 考虑 n 阶 Jordan 标准形 $J=\mathrm{diag}\,(J_1,J_2,\cdots,J_t)$. 由定理 7.2.6, 则和 Jordan 标准形 J 可交换的方阵为

$$B=\begin{pmatrix}B_{11}&\cdots&B_{1t}\\ \vdots&&\vdots\\ B_{t1}&\cdots&B_{tt}\end{pmatrix}, \tag{7.2.16}$$

其中矩阵 B_{jk} 为 $e_j\times e_k$ 矩阵, $1\leqslant j,k\leqslant t$. 它由式 (7.2.7), (7.2.8) 和 (7.2.9) 给出, 即当 $\lambda_j\neq\lambda_k$ 时, $B_{jk}=0$; 当 $\lambda_j=\lambda_k$ 时,

$$B_{jk}=\sum_{i=0}^{\min(e_j,e_k)-1}b_i^{(jk)}\begin{pmatrix}0&N^i\\0&0\end{pmatrix}, \tag{7.2.17}$$

其中 N 为 $\min(e_j,e_k)$ 阶方阵 (见式 (7.1)), $b_i^{(jk)}$ 为任意常数, $1\leqslant j,k\leqslant t$.

设 n 阶方阵 C 和所有 n 阶方阵 B 可交换, 其中 n 阶方阵 B 的元素 $b_i^{(jk)}$ 由独立参数组成. 将 C 和 B 一样分块为

$$C=\begin{pmatrix}C_{11}&\cdots&C_{1t}\\ \vdots&&\vdots\\ C_{t1}&\cdots&C_{tt}\end{pmatrix},$$

则有

$$\sum_{j=1}^{t}C_{ij}B_{jk}=\sum_{j=1}^{t}B_{ij}C_{jk},\quad j,k=1,2,\cdots,t.$$

由于 B_{ij} 和 B_{jk} 中非零元素分别由独立参数 $b_u^{(ij)}$ 和 $b_v^{(jk)}$ 组成, 所以证明了

$$C_{ij}B_{jk}=0,\quad i\neq j,\quad B_{ij}C_{jk}=0,\quad j\neq k,\quad C_{jj}B_{jk}=B_{jk}C_{kk},\quad 1\leqslant j,k\leqslant t.$$

这证明了 $C_{ij}=0, i\neq j$, 而

$$C_{ii}=\sum_{k=0}^{e_i-1}\frac{1}{k!}a_{ik}N_i^k,\quad i=1,2,\cdots,t,$$

其中 $N_i = J_i$ 是式 (7.2) 中的 e_i 阶方阵.

无妨设 $e_i \leqslant e_j$. 设 $\lambda_i = \lambda_j$, 由 $B_{ij}C_{jj} = C_{ii}B_{ij}$ 可知

$$a_{ik} = a_{jk}, \quad k = 0, 1, \cdots, e_i - 1.$$

由广义 Lagrange 插值公式 (引理 7.2.7), 存在多项式 $g(x)$, 使得 $g^{(k)}(\lambda_i) = a_{ik}$, $1 \leqslant i \leqslant t$, $0 \leqslant k < e_i$. 所以

$$C_{ii} = \sum_{k=0}^{e_i-1} \frac{1}{k!} a_{ik} N_i^k = \sum_{k=0}^{e_i-1} \frac{1}{k!} g^{(k)}(\lambda_i) N_i^k = g(J_i), \quad i = 1, 2, \cdots, t.$$

因此,

$$C = \operatorname{diag}(C_{11}, C_{22}, \cdots, C_{tt}) = \operatorname{diag}(g(J_1), g(J_2), \cdots, g(J_t)) = g(J). \qquad \square$$

今任取 n 阶方阵 A. 设 P 为 n 阶非异复方阵, 使得 $J = P^{-1}AP$, 其中 J 为 Jordan 标准形. 今记 X 和 B 为 n 阶复方阵,

$$\mathfrak{S} = \{\, X \mid XA = AX \,\} = \{\, X \mid P^{-1}XPJ = JP^{-1}XP \,\}.$$

记 $\mathfrak{S}_J = \{B \mid BJ = JB\}$. 于是证明了

$$\mathfrak{S}_J = P^{-1}\mathfrak{S}\, P = \{\, P^{-1}XP \mid \forall X \in \mathfrak{S} \,\}.$$

若 n 阶方阵 Y 和 $\mathfrak{S}$ 中任一方阵可交换, 则 $P^{-1}YP$ 和 $\mathfrak{S}_J$ 中任一方阵可交换, 所以 $P^{-1}YP = g(J)$, 其中 $g(x)$ 为多项式. 因此

$$Y = P(g(J))P^{-1} = g(PJP^{-1}) = g(A). \qquad \square$$

(三) 复方阵的 Jordan 分解

定义 7.2.9 设 A 为 n 阶复方阵, A 称为**半单的**, 如果 A 相似于对角方阵.

定理 7.2.10 (Jordan 分解) 设 A 为 n 阶复方阵, 则有

$$A = S + N, \quad SN = NS, \tag{7.2.18}$$

其中 S 是 n 阶半单方阵, 即 A 相似于对角形; N 是 n 阶幂零方阵. 又 S 和 N 都是 A 的多项式, 其中常数项都等于零. 又若 $A = \widetilde{S} + \widetilde{N}$, 其中 $\widetilde{S}$ 相似于对角形, $\widetilde{N}$ 是幂零方阵, 而且 $\widetilde{S}\widetilde{N} = \widetilde{N}\widetilde{S}$, 则有 $\widetilde{S} = S$, $\widetilde{N} = N$, 即分解唯一.

证 下面先证 Jordan 分解的唯一性.

今若 $A = S + N = \widetilde{S} + \widetilde{N}$, 其中 S 和 $\widetilde{S}$ 相似于对角形, N 和 $\widetilde{N}$ 幂零, 又 $SN = NS$, $\widetilde{S}\widetilde{N} = \widetilde{N}\widetilde{S}$. 于是 $\widetilde{S}A = A\widetilde{S}$, $\widetilde{N}A = A\widetilde{N}$. 由于已证 S 和 N 为 A 的多项

式, 所以 $\widetilde{S}S=S\widetilde{S}$, $\widetilde{N}N=N\widetilde{N}$. 于是由 $S-\widetilde{S}=\widetilde{N}-N$ 可知 $S-\widetilde{S}$ 相似于对角形, 且 $\widetilde{N}-N$ 幂零. 这证明了 $S-\widetilde{S}$ 又半单, 又幂零, 因此, $\widetilde{S}=S$, 所以 $\widetilde{N}=N$. 这证明了唯一性.

下面证 Jordan 分解的存在性. 我们给出两种证法.

证法一　下面的证明充分依赖于 Jordan 标准形. 由于定理的条件和结论在相似下都不改变, 所以无妨设 A 为 Jordan 标准形

$$J=\operatorname{diag}(\lambda_1E_1+N_1,\lambda_2E_2+N_2,\cdots,\lambda_tE_t+N_t), \tag{7.2.19}$$

取

$$S=\operatorname{diag}(\lambda_1E_1,\lambda_2E_2,\cdots,\lambda_tE_t),\quad N=\operatorname{diag}(N_1,N_2,\cdots,N_t). \tag{7.2.20}$$

自然 S 为对角形, N 为幂零方阵, 又 $SN=NS$. 余下证 S 为 J 的常数项等于零的多项式 $g(J)$. 如果 $S=g(J)$, 于是 $N=J-S=J-g(J)$ 仍为 J 的常数项等于零的多项式.

今任取和 A 可交换的方阵 B, 将 B 和 A 一样分块, 有

$$B=\begin{pmatrix}B_{11}&\cdots&B_{1t}\\\vdots&&\vdots\\B_{t1}&\cdots&B_{tt}\end{pmatrix}.$$

由定理 7.2.6, 当 $\lambda_j\neq\lambda_k$, 则 $B_{jk}=0$. 于是

$$SB-BS=((\lambda_j-\lambda_k)B_{jk})=0.$$

由定理 7.2.8, S 为 A 的多项式, 记为 $h(\lambda)$, 则 $h(J)=S$.

如果 Jordan 标准形 J 有零特征根, 例如 $\lambda_1=0$. 则由 $h(J)=S$, 其中 S 由式 (7.2.20) 定义, 所以 $h(\lambda_1)=\lambda_1$, 即 $h(0)=0$. 这时令 $g(\lambda)=h(\lambda)$, 则 $g(0)=0$, 又 $g(J)=h(J)=S$;

如果 Jordan 标准形 J 的特征根都不等于零. 记 Jordan 标准形 J 的极小多项式为 $m(\lambda)$, 则 $m(0)\neq 0$. 于是令 $g(\lambda)=h(\lambda)-\dfrac{h(0)}{m(0)}m(\lambda)$, 则有 $g(0)=0$, 而且 $g(J)=h(J)-\dfrac{h(0)}{m(0)}m(J)=h(J)=S$. 至此证明了 Jordan 分解的存在性.

证法二　由于条件和结论都在相似下不变, 因此无妨设 n 阶复方阵 A 为 Jordan 标准形 $J=\operatorname{diag}(J_1,J_2,\cdots,J_t)$, 其中 J_k 为由初等因式 $(\lambda-\lambda_k)^{e_k}$ 决定的 Jordan 块, $1\leqslant k\leqslant t$.

记 A 的所有不同特征根为 $\mu_1,\mu_2,\cdots,\mu_s$, 于是特征多项式

$$\det(\lambda E_n-A)=\prod_{j=1}^{s}(\lambda-\mu_j)^{n_j}=\prod_{k=1}^{t}(\lambda-\lambda_k)^{e_k}. \tag{7.2.21}$$

当 $\mu_1, \mu_2, \cdots, \mu_s$ 中有一个为零时, 取 $g_j(\lambda) = \lambda^{\delta_{\mu_j,0}}(\lambda-\mu_j)^{-n_j}\det(\lambda E_n - A)$; 当 $\mu_1, \mu_2, \cdots, \mu_s$ 都不等于零时, 取 $g_j(\lambda) = \lambda(\lambda-\mu_j)^{-n_j}\det(\lambda E_n - A)$. 于是在 $\mu_j \neq 0$ 时, $\lambda \mid g_j(\lambda)$. 另一方面, 对 $i = 1, 2, \cdots, s$ 时有 $(g_i(\lambda), (\lambda-\mu_i)^{n_i}) = 1$, 所以存在多项式 $u_i(\lambda)$ 和 $v_i(\lambda)$, 使得

$$u_i(\lambda)g_i(\lambda) + v_i(\lambda)(\lambda-\mu_i)^{n_i} = 1. \tag{7.2.22}$$

令

$$f(\lambda) = \sum_{i=1}^{s} \mu_i u_i(\lambda) g_i(\lambda) = \sum_{i=1}^{s} \mu_i(1 - v_i(\lambda)(\lambda-\mu_i)^{n_i}). \tag{7.2.23}$$

由于当 $\mu_i = 0$ 时, $\mu_i u_i(\lambda)g_i(\lambda) = 0$; 当 $\mu_i \neq 0$ 时, $\lambda \mid g_i(\lambda)$, 所以 $\lambda \mid f(\lambda)$, 即 $f(\lambda)$ 的常数项为零, 因此, $\lambda - f(\lambda)$ 的常数项也为零. 又有

$$(\lambda-\mu_i)^{n_i} \mid (f(\lambda)-\mu_i), \quad i = 1, 2, \cdots, s. \tag{7.2.24}$$

令

$$S = f(A), \quad N = A - f(A) = A - S. \tag{7.2.25}$$

所以有

$$SN = NS, \quad A = S + N. \tag{7.2.26}$$

因此为了证明这是 Jordan 分解, 只需证明 S 相似于对角形, 且 N 幂零.

令

$$f(A) = f(\operatorname{diag}(J_1, J_2, \cdots, J_t)) = \operatorname{diag}(f(J_1), f(J_2), \cdots, f(J_t)). \tag{7.2.27}$$

又由式 (7.2.24), $f(\lambda) = \mu_i + f_i(\lambda)(\lambda-\mu_i)^{n_i}$, 其中 $f_i(\lambda)$ 仍为多项式, $i = 1, 2, \cdots, s$. 所以当 Jordan 块 J_k 有 $\lambda_k = \mu_i$, 则 $e_k \leqslant n_i$. 于是

$$f(J_k) = \mu_i E_k + f_i(J_k)N_k^{n_i} = \lambda_k E_k, \quad k = 1, 2, \cdots, t.$$

这证明了

$$\begin{aligned} S = f(A) &= \operatorname{diag}(\lambda_1 E_1, \lambda_2 E_2, \cdots, \lambda_t E_t) \\ &= \operatorname{diag}(J_1 - N_1, J_2 - N_2, \cdots, J_t - N_t) \\ &= A - \operatorname{diag}(N_1, N_2, \cdots, N_t). \end{aligned} \tag{7.2.28}$$

所以 S 为对角方阵, 而 $A - S = N = \operatorname{diag}(N_1, N_2, \cdots, N_t)$ 幂零. 至此证明了 Jordan 分解的存在性. □

注意 上面第二种证明是利用构造一些多项式, 从而造出 S 及 N. 因此, 如果 A 为 n 阶实方阵, 则 S 及 N 仍为 n 阶实方阵. 所以 Jordan 分解在实的情形也成

立. 证法一的证明中, 需要证 $g(\lambda)$ 为实多项式.

(四) 复方阵为两个复对称方阵的乘积

定理 7.2.11 任一 n 阶复方阵 A 能分解为两个复对称方阵的乘积, 且可指定其中一个为非异的.

证 给定 n 阶复方阵 A, 所以存在 n 阶非异复方阵 P, 使得 (见式 (7.2.4) 和式 (7.2.5))

$$P^{-1}AP = J = \operatorname{diag}(J_1, J_2, \cdots, J_t), \quad J_k = \lambda_k E_k + N_k, \quad k = 1, 2, \cdots, t. \tag{7.2.29}$$

记

$$S_k^{(e_k)} = \begin{pmatrix} & & & & 1 \\ & & & \cdot & \\ & & \cdot & & \\ & \cdot & & & \\ 1 & & & & \end{pmatrix}. \tag{7.2.30}$$

则 S_k 为 e_k 阶非异对称方阵, 且 $S_k^{-1} = S_k$. 又 $S_k J_k$ 仍为 e_k 阶复对称方阵. 记

$$S = \operatorname{diag}(S_1^{(e_1)}, S_2^{(e_2)}, \cdots, S_t^{(e_t)}), \tag{7.2.31}$$

则 S 为 n 阶非异对称方阵, 又 $T = SJ$ 仍为复对称方阵. 即有 $J = ST$. 所以有

$$A = P(P^{-1}AP)P^{-1} = PJP^{-1} = PSTP^{-1} = (PSP')((P^{-1})'TP^{-1}) = S_1S_2, \tag{7.2.32}$$

其中 $S_1 = PSP'$ 为 n 阶非异复对称方阵, $S_1 = (P^{-1})'TP^{-1}$ 为 n 阶复对称方阵.

考虑 n 阶复方阵 A', 则有 n 阶非异复对称方阵 S_3 和 n 阶复对称方阵 S_4, 使得 $A' = S_3S_4$. 即 $A = S_4S_3$. 这证明了 n 阶复方阵 A 可分解为两个 n 阶复对称方阵的乘积, 且可指定其中一个为非异的. □

§7.3 方阵幂级数和方阵函数 *

这一节, 也是 Jordan 标准形理论的应用. 我们只考虑一类特殊的方阵函数, 称为纯函数. 对这类函数的研究能够充分地利用方阵在相似下的 Jordan 标准形理论. 正因为要利用 Jordan 标准形, 所以这一节始终是以元素为复数的方阵作为讨论对象的.

为了引出方阵纯函数的概念, 我们先讨论方阵多项式函数. 给定 n^2 个独立未知

数 x_{ij}, $1 \leqslant i,j \leqslant n$. 它们组成 n 阶方阵

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}. \tag{7.3.1}$$

给定 $m+1$ 个复常数 $a_0, a_1, \cdots, a_m$, 则 n 阶方阵

$$f(X) = \sum_{k=0}^{m} a_k X^k = a_m X^m + a_{m-1} X^{m-1} + \cdots + a_1 X + a_0 E_n \tag{7.3.2}$$

称为方阵 X 的多项式, 简称为**方阵多项式**.

引理 7.3.1 设 X 为 n 阶方阵, $f(X)$ 为方阵多项式, 则有下面一些特性.

(1) 任取 n 阶非异方阵 P, 则有

$$f(PXP^{-1}) = Pf(X)P^{-1}; \tag{7.3.3}$$

(2) 当 X 为准对角方阵 $X = \operatorname{diag}(X_1, X_2, \cdots, X_t)$ 时, 有

$$f(\operatorname{diag}(X_1, X_2, \cdots, X_t)) = \operatorname{diag}(f(X_1), f(X_2), \cdots, f(X_t)). \tag{7.3.4}$$

因此, 设 n 阶方阵 A 相似于 Jordan 标准形

$$PAP^{-1} = J = \operatorname{diag}(J_1, J_2, \cdots, J_t), \tag{7.3.5}$$

则

$$\begin{aligned} f(A) &= f(P^{-1}JP) = P^{-1}f(J)P = P^{-1}f(\operatorname{diag}(J_1, J_2, \cdots, J_t))P \\ &= P^{-1}\operatorname{diag}(f(J_1), f(J_2), \cdots, f(J_t))P. \end{aligned} \tag{7.3.6}$$

所以在给定 n 阶方阵 A 后, 若知道它经过 n 阶非异方阵 P 化为 Jordan 标准形, 那末, 要计算多项式 $f(X)$ 在 $X = A$ 时的函数值, 只要知道多项式 $f(X)$ 对每个 e 阶 Jordan 块的函数值就行了.

(3) 当 $X = J_0 = \lambda_0 E_e + N_e$ 为 e 阶 Jordan 块时, 则

$$f(J_0) = f(\lambda_0 E_e + N_e) = \sum_{k=0}^{e-1} \frac{1}{k!} f^{(k)}(\lambda_0) N_e^k. \tag{7.3.7}$$

所以在 $X = J_0$ 时, 方阵多项式 $f(X)$ 的函数值 $f(J_0)$ 只和数值多项式

$$f(\lambda) = a_m \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_0 \tag{7.3.8}$$

以及它的若干次微商 $f(\lambda) = f^{(0)}(\lambda)$, $f'(\lambda) = f^{(1)}(\lambda)$, $f''(\lambda) = f^{(2)}(\lambda)$, $\cdots$, $f^{(n-1)}(\lambda)$ 在 $\lambda = \lambda_0$ 的值有关, 其中

$$f^{(k+1)}(\lambda) = \frac{\mathrm{d}f^{(k)}(\lambda)}{\mathrm{d}\lambda}, \quad k = 0, 1, \cdots. \tag{7.3.9}$$

证 由多项式函数的定义可知性质(1) 与(2) 显然成立. 为了证明引理, 只要证(3) 成立. 记 $f(X)=\sum\limits_{k=0}^{m}a_kX^{m-k}$, 于是

$$\begin{aligned}f(J_0)&=\sum_{k=0}^{m}a_kJ_0^{m-k}=\sum_{k=0}^{m}a_k(\lambda_0E+N)^{m-k}\\&=\sum_{k=0}^{m}a_k\sum_{j=0}^{m-k}\binom{m-k}{j}\lambda_0^{m-k-j}N^j=\sum_{j=0}^{m}\frac{1}{j!}\Big(\sum_{k=0}^{m-j}a_k\frac{(m-k)!}{(m-k-j)!}\lambda_0^{m-k-j}\Big)N^j.\end{aligned}$$

记 $f(x)=\sum\limits_{j=0}^{m}a_kx^{m-k}$ 为通常的多项式, 则

$$f^{(j)}(x)=\frac{\mathrm{d}^kf(x)}{\mathrm{d}x^k}=\sum_{k=0}^{m-j}a_k\frac{(m-k)!}{(m-k-j)!}\lambda_0^{m-k-j}.$$

因此证明了式 (7.3.7) 成立. □

由引理 7.3.1, 有下面自然的推论.

定理 7.3.2 给定 n 阶复方阵 A, 记

$$J=\operatorname{diag}(J_1,J_2,\cdots,J_t) \tag{7.3.10}$$

为 A 的 Jordan 标准形, 其中

$$J_k=\lambda_kE_k+N_k \tag{7.3.11}$$

为 e_k 阶 Jordan 块, $k=1,2,\cdots,t$. 这时存在 n 阶非异复方阵 P, 使得 $A=PJP^{-1}$, 且对 m 次多项式 $f(x)$, 有

$$f(A)=P\Big(\operatorname{diag}\Big(\sum_{k=0}^{e_1}\frac{1}{k!}f^{(k)}(\lambda_1)N_1^k,\cdots,\sum_{k=0}^{e_t}\frac{1}{k!}f^{(k)}(\lambda_t)N_t^k\Big)\Big)P^{-1}. \tag{7.3.12}$$

所以 $f(A)$ 的值只和非异复方阵 P 及 $f(x)$ 的各阶导数在特征根上的值有关.

在分析课中, 我们知道: 作为多项式的自然推广为幂级数. 所以, 我们可以同样讨论方阵幂级数. 为此先给出极限的定义.

定义 7.3.3 任给一个 $n\times m$ **矩阵序列**

$$A_k=\begin{pmatrix}a_{11}^{(k)}&\cdots&a_{1m}^{(k)}\\\vdots&&\vdots\\a_{n1}^{(k)}&\cdots&a_{nm}^{(k)}\end{pmatrix},\quad k=0,1,\cdots, \tag{7.3.13}$$

如果 nm 个复数序列 $\{a_{ij}^{(k)}\}_{k=0}^{\infty}$, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant m$ 都同时收敛, 且 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}$, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant m$, 那末称矩阵序列 $\{A_k\}_{k=0}^{\infty}$ **收敛**, 且收敛于 $n\times m$ 矩阵

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}. \tag{7.3.14}$$

用下面符号记之:

$$\lim_{k\to\infty} A_k=\lim_{k\to\infty}\begin{pmatrix} a_{11}^{(k)} & \cdots & a_{1n}^{(k)} \\ \vdots & & \vdots \\ a_{n1}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}=\begin{pmatrix} \lim\limits_{k\to\infty} a_{11}^{(k)} & \cdots & \lim\limits_{k\to\infty} a_{1n}^{(k)} \\ \vdots & & \vdots \\ \lim\limits_{k\to\infty} a_{n1}^{(k)} & \cdots & \lim\limits_{k\to\infty} a_{nn}^{(k)} \end{pmatrix}=A. \tag{7.3.15}$$

定义 7.3.4 任给一个 $n\times m$ 矩阵序列 $\{A_k\}_{k=0}^{\infty}$ 及一个复数序列 $\{a_k\}_{k=0}^{\infty}$, 作 $n\times m$ 矩阵序列

$$B_k=\sum_{k=0}^{p} a_k A_k, \quad k=0,1,\cdots. \tag{7.3.16}$$

若矩阵序列 $\{B_k\}_{k=0}^{\infty}$ 收敛, 且收敛于 $n\times m$ 矩阵 B, 那末称 $n\times m$ 矩阵级数 $\sum\limits_{k=0}^{\infty} a_k A_k$ **收敛** 于 $n\times m$ 矩阵 B. 按定义,

$$a_0A_0+a_1A_1+\cdots=\sum_{k=0}^{\infty} a_kA_k=B. \tag{7.3.17}$$

若 $n\times m$ 矩阵序列 $\{B_k\}_{k=0}^{\infty}$ 不收敛, 那末称 $n\times m$ 矩阵级数 (7.3.17) **发散**.

引理 7.3.5 设

$$A_k=\begin{pmatrix} a_{11}^{(k)} & \cdots & a_{1m}^{(k)} \\ \vdots & & \vdots \\ a_{n1}^{(k)} & \cdots & a_{nm}^{(k)} \end{pmatrix}, \quad k=0,1,\cdots \tag{7.3.18}$$

为 $n\times m$ 矩阵, $k=0, 1, 2, \cdots$, 又 $a_0, a_1, a_2, \cdots$ 为复数, 则 $n\times m$ 矩阵

$$\sum_{k=0}^{\infty} a_kA_k \tag{7.3.19}$$

收敛的必要且充分条件为 nm 个数值级数

$$\sum_{k=0}^{\infty} a_k a_{ij}^{(k)}, \quad i=1,2,\cdots,n, \quad j=0,1,2,\cdots,m \tag{7.3.20}$$

同时收敛. 方阵级数 $\sum\limits_{k=0}^{\infty} a_k A_k$ 发散的必要且充分条件为式 (7.3.20) 的 nm 个数值级数中至少有一个数值级数发散.

作为方阵多项式的推广, 现在可以推广数值幂级数的概念到方阵的情形了.

定义 7.3.6 对任一 n 阶方阵 A 及任一复数序列 $\{a_k\}_{k=0}^{\infty}$, 方阵级数

$$\sum_{k=0}^{\infty} a_k A^k = a_0 E_n + a_1 A + a_2 A^2 + \cdots \tag{7.3.21}$$

称为方阵 A 的**方阵幂级数**. 因此, 对方阵变量 X, 利用方阵 X 的幂级数便定义了一个**方阵函数**

$$Y = f(X) = \sum_{k=0}^{\infty} a_k X^k = a_0 E_n + a_1 X + a_2 X^2 + \cdots . \tag{7.3.22}$$

下面来证明, 方阵幂级数和方阵多项式具有同样的三个特性, 即有

定理 7.3.7 方阵 X 的幂级数 $f(X)$ 有性质:

(1) 对一切同阶非异方阵 P, 有

$$f(PXP^{-1}) = Pf(X)P^{-1}; \tag{7.3.23}$$

(2)

$$f(\operatorname{diag}(X_1, X_2, \cdots, X_s)) = \operatorname{diag}(f(X_1), f(X_2), \cdots, f(X_s)); \tag{7.3.24}$$

(3) 记复变数的幂级数为

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \tag{7.3.25}$$

则对任一 e 阶 Jordan 块 $J_0 = \lambda_0 E + N$, 当 $|\lambda_0| < \rho$, 其中 ρ 为幂级数 (7.3.25) 的收敛半径时, 有

$$f(J_0) = \sum_{k=0}^{\infty} a_k J_0^k = \begin{pmatrix} f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f''(\lambda_0) & \cdots & \frac{1}{(e-1)!}f^{(e-1)}(\lambda_0) \\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(e-2)!}f^{(e-2)}(\lambda_0) \\ & & f(\lambda_0) & \cdots & \frac{1}{(e-3)!}f^{(e-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_0) \end{pmatrix}, \tag{7.3.26}$$

即

$$f(J_0)=\sum_{k=0}^{\infty}a_kJ_0^k=\sum_{k=0}^{e-1}\frac{1}{k!}f^{(k)}(\lambda_0)N_e^k. \tag{7.3.27}$$

证 因为对任一非异复方阵 P 及任一方阵序列 $\{B_k\}_{k=0}^{\infty}$, 有

$$\lim_{k\to\infty}PB_kP^{-1}=P(\lim_{k\to\infty}B_k)P^{-1},$$

所以有

$$f(PXP^{-1})=\sum_{k=0}^{\infty}a_k(PXP^{-1})^k=P(\sum_{k=0}^{\infty}a_kX^k)P^{-1}=Pf(X)P^{-1},$$

即性质**(1)** 成立. 再由

$$(\operatorname{diag}(X_1,X_2,\cdots,X_s))^k=\operatorname{diag}(X_1^k,X_2^k,\cdots,X_s^k),$$

可知性质**(2)** 成立.

下面来证明性质**(3)** 成立. 记 $f_p(z)=\sum\limits_{k=0}^{p}a_kz^k$, 则 $\lim\limits_{p\to\infty}f_p(z)=f(z)$. 令 $B_p=\sum\limits_{k=0}^{p}a_kJ_0^k$, 则 $\lim\limits_{p\to\infty}B_p=f(J_0)$. 由于方阵多项式具有性质**(3)**, 所以

$$B_p=\sum_{k=0}^{e-1}\frac{1}{k!}f_p^{(k)}(\lambda_0)N_e^k.$$

当 $p\to+\infty$ 时, 由方阵序列极限的定义及数值级数的性质可知性质**(3)**成立. □

利用这个定理, 就可以把方阵 A 的幂级数的性质转化为当 A 是 Jordan 块时幂级数的性质. 这样讨论就方便了. 利用这点, 在下面先推广数值幂级数的收敛半径的概念, 然后引进几种重要的方阵级数, 诸如指数函数、对数函数、三角函数、二项展开式等等.

现在来证方阵幂级数的主要定理.

定理 7.3.8 给定一个复变数的幂级数 $\sum\limits_{k=0}^{\infty}a_kz^k$. 记它的**收敛半径** 为 ρ, 则 ρ 也称为**方阵幂级数** $\sum\limits_{k=0}^{\infty}a_kA^k$ 的收敛半径. 设 n 阶复方阵 A 的 n 个特征根为 λ_1, $\lambda_2,\cdots,\lambda_n$, 记

$$\lambda_0=\max_{1\leqslant j\leqslant n}|\lambda_j|\geqslant 0, \tag{7.3.28}$$

则方阵的幂级数有如下性质:

(1) 当 $\rho>\lambda_0$ 时, 方阵 A 的幂级数 $\sum\limits_{k=0}^{\infty}a_kA^k$ 收敛, 且对 e 阶 Jordan 块 $\lambda_*E_e+N_e$, 有

$$f(\lambda_*E_e+N_e)=\sum_{k=0}^{e-1}\frac{1}{k!}f^{(k)}(\lambda_*)N_e^k; \tag{7.3.29}$$

(2) 当 $\rho < \lambda_0$ 时, 方阵幂级数 $\sum\limits_{k=0}^{\infty} a_k A^k$ 发散;

(3) 当 $\rho = \lambda_0$ 时, 方阵幂级数 $\sum\limits_{k=0}^{\infty} a_k A^k$ 收敛的必要且充分条件是: 对每一绝对值为 ρ 的特征根 λ_j, 设方阵 A 的属于特征根 λ_j 的初等因子中最高方次为 n_j, 则 $\boldsymbol{n_j}$ 个数值幂级数

$$f^{(p)}(z) = \frac{\mathrm{d}^p f(z)}{\mathrm{d}z} = \sum_{k=0}^{\infty} k(k-1)\cdots(k-p+1)a_k z^{k-p}, \quad p = 0, 1, \cdots, n_j - 1 \tag{7.3.30}$$

都收敛;

(4) 在方阵幂级数 $\sum\limits_{k=0}^{\infty} a_k A^k$ 收敛时, 它的 n 个特征根是

$$\sum_{k=0}^{\infty} a_k \lambda_j^k, \quad j = 1, 2, \cdots, n. \tag{7.3.31}$$

证 设 A 相似于 Jordan 标准形 $\mathrm{diag}\,(J_1, J_2, \cdots, J_t)$, 其中 J_1, J_2, $\cdots$, J_t 都是 Jordan 块. 则判断方阵幂级数 $\sum\limits_{k=0}^{\infty} a_k A^k$ 收敛或发散的问题便化为: t 个方阵幂级数 $f(J_j) = \sum\limits_{k=0}^{\infty} a_k J_j^k$, $1 \leqslant j \leqslant t$ 何时同时收敛, 何时至少有一个发散. 由定理 7.3.7 可知, 设 $J_j = \lambda_j E_j + N_j$ 是 e_j 阶 Jordan 块, 则 $f(J_j)$ 收敛的必要且充分条件为数值幂级数 (7.3.31) 同时收敛; 发散的必要且充分条件为数值幂级数 (7.3.31) 中有一个发散.

已知复变幂级数 $f(z) = \sum\limits_{k=0}^{\infty} a_k z^k$ 的收敛半径为 ρ, 那末在 $|z| < \rho$ 时, $f(z)$ 是一个全纯函数. 因此, $f(z)$ 的任意次微商

$$f^{(p)}(z) = \frac{\mathrm{d}^p f(z)}{\mathrm{d}z} = \sum_{k=0}^{\infty} k(k-1)\cdots(k-p+1)a_k z^{k-p}, \quad p = 0, 1, \cdots \tag{7.3.32}$$

在 $|z| < \rho$ 也都收敛. 所以证明了: 当 $|\lambda_j| < \rho$ 时, $f(J_j)$ 收敛, 即断言**(1)** 成立. 当存在 j, 使得 $|\lambda_j| > \rho$, 已知数值幂级数 $\sum\limits_{k=0}^{\infty} a_k \lambda_j^k$ 发散, 所以断言**(2)** 也成立. 断言**(3)** 显然成立. 最后, 因为当 A 为 Jordan 块 J_0 时, $f(J_0) = \sum\limits_{k=0}^{\infty} a_k J_0^k$ 是上三角方阵, 对角元素全是 $f(\lambda_0) = \sum\limits_{k=0}^{\infty} a_k \lambda_0^k$. 已知它是方阵 $f(J_0)$ 的特征根, 所以断言**(4)** 成立. □

利用这个定理, 我们可以利用特殊的数值幂级数, 来引进一系列方阵幂级数, 它

们是**初等方阵函数**:

$$\begin{aligned}\exp(A) &= e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k; \\ \sin(A) &= \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{(2k-1)!} A^{2k-1}; \\ \cos(A) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} A^{2k}; \\ \ln(E_n + A) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k} A^k, \\ (E_n + A)^\alpha &= \sum_{k=0}^{\infty} \binom{\alpha}{k} A^k,\end{aligned} \tag{7.3.33}$$

其中 α 为任意确定的实数,

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}, \quad k = 0, 1, 2, \cdots. \tag{7.3.34}$$

函数 $\exp(A)$, $\sin(A)$, $\cos(A)$ 对一切方阵 A 都有定义, 原因是复变幂级数

$$e^z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k; \quad \sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^{k-1}}{(2k-1)!} z^{2k-1}; \quad \cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k} \tag{7.3.35}$$

的收敛半径都是无穷大. 另一方面, 因为复变幂级数

$$\ln(1+z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k} z^k, \quad (1+z)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k \tag{7.3.36}$$

的收敛半径都是 1, 所以方阵幂级数 $\ln(E_n + A)$ 及 $(E_n + A)^\alpha$ 对这样的方阵 A 才有定义: 即 A 的特征根的模都小于 1. 我们不再仔细追究方阵函数 $\ln(E_n + A)$ 及 $(E_n + A)^\alpha$ 的定义范围了. 读者学了复变函数理论后, 可以看出, 只要 A 没有特征根 -1, 那末 $\ln(E_n + A)$ 及 $(E_n + A)^\alpha$ 总是有定义的. 然而只有在 A 的特征根的模都小于 1 时才能用上面的幂级数表示.

最有用的方阵幂级数是 $\exp(A)$, $\ln(E_n + A)$ 及 $(E_n + A)^\alpha$. 不过, 在应用时要注意, 这些函数和普通的复变函数 e^z, $\ln(1+z)$ 及 $(1+z)^\alpha$ 具有完全不同的性质. 其主要原因出在复数相乘有交换律, 而方阵相乘却不一定可交换. 例如, 对应两复数 a 和 b, 则 $e^a e^b = e^b e^a$. 但是对任两方阵 A 和 B, 则方阵 $\exp(A+B)$, $\exp(A)\exp(B)$ 和 $\exp(B)\exp(A)$ 可以是三个完全不同的方阵. 下面便是一个很简单的例子.

例 7.3.1

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

则

$$\exp(A)=\begin{pmatrix} e & 0\\ 0 & 0\end{pmatrix},\quad \exp(B)=\begin{pmatrix} e & e\\ 0 & 0\end{pmatrix},$$

所以

$$\exp(A+B)=\begin{pmatrix} e^2 & \frac{1}{2}e^2\\ 0 & 0\end{pmatrix},\quad e^Ae^B=\begin{pmatrix} e^2 & e^2\\ 0 & 0\end{pmatrix},\quad e^Be^A=\begin{pmatrix} e^2 & 0\\ 0 & 0\end{pmatrix}.$$

这个例子告诉我们, 由于矩阵的乘法不适合交换律, 所以方阵幂级数也有一些和普通幂级数不同的性质. 因此在运用方阵幂级数时, 不能随便套上普通函数的性质.

定理 7.3.9 记 A 和 B 是一对互相可交换的 n 阶方阵, 则.

$$\exp(A)\exp(B)=\exp(B)\exp(A)=\exp(A+B). \tag{7.3.37}$$

证 设 $\exp(A+B)=\exp(A)\exp(B)$ 对任意 n 阶方阵 A 和 B 成立, 所以, $\exp(B+A)=\exp(B)\exp(A)$. 因此, 只要证 $\exp(A+B)=\exp(A)\exp(B)$ 就可以了.

显然, 由 $AB=BA$, 有

$$\begin{aligned}\exp(A+B)&=\sum_{q=0}^{\infty}\frac{1}{q!}(A+B)^q=\sum_{q=0}^{\infty}\frac{1}{q!}\sum_{k=0}^{q}\binom{q}{k}A^kB^{q-k}\\&=\sum_{q=0}^{\infty}\sum_{k=0}^{q}\frac{1}{k!(q-k)!}A^kB^{q-k}=\sum_{q=0}^{\infty}\sum_{j+k=q}\frac{1}{j!k!}A^jB^k,\end{aligned} \tag{7.3.38}$$

$$\exp(A)\exp(B)=\Big(\sum_{k=0}^{\infty}\frac{1}{k!}A^k\Big)\Big(\sum_{q=0}^{\infty}\frac{1}{q!}B^q\Big)=\sum_{j,k=0}^{\infty}\frac{1}{j!k!}A^jB^k.$$

记方阵

$$C_p=\sum_{j,k=0}^{p}\frac{1}{j!k!}A^jB^k-\sum_{q=0}^{p}\sum_{j+k=q}\frac{1}{j!k!}A^jB^k=\sum_{q=0}^{p}\sum_{\substack{j+k>p\\0\leqslant j,k\leqslant p}}\frac{1}{j!k!}A^jB^k.$$

为了证明 $\exp(A+B)=\exp(A)\exp(B)$, 只要证明当 p 趋于无穷时方阵 C_p 趋于零, 即方阵 C_p 的每个元素都趋于零. 记 $A=(a_{ir})$, $B=(b_{rq})$. 令

$$d=\max_{1\leqslant i,r\leqslant n}(|a_{ir}|,|b_{ir}|)+1.$$

并记 $A^j=(a_{ir}^{(j)})$, $B^k=(b_{ir}^{(k)})$, 用归纳法, 不难证明 $|a_{ir}^{(j)}|\leqslant(nd)^j$, $|b_{ir}^{(k)}|\leqslant(nd)^k$, $1\leqslant i,r\leqslant n$, $k=0,1,2,\cdots$. 而 C_p 的第 i 行第 u 列元素为

$$c_{iu}^{(p)}=\sum_{\substack{j+k>p\\0\leqslant j,k\leqslant p}}\frac{1}{j!k!}\sum_{r=1}^{n}a_{ir}^{(j)}b_{ru}^{(k)}.$$

由于 $d>1$, 所以

$$|c_{ij}^{(p)}| \leqslant \sum_{\substack{j+k>p\\0\leqslant j,k\leqslant p}} \frac{1}{j!k!}\sum_{r=1}^{n}|a_{ir}^{(j)}||b_{ru}^{(k)}| \leqslant \sum_{\substack{j+k>p\\0\leqslant j,k\leqslant p}} \frac{1}{j!k!}n(nd)^{j+k}$$
$$= n\sum_{j,k=0}^{p}\frac{1}{j!k!}(nd)^{j+k} - n\sum_{q=0}^{p}\sum_{j+k=q}\frac{1}{j!k!}(nd)^{j+k}.$$

当 $p\to\infty$, 则极限小于等于 $n\exp(2nd)-n\exp(2nd)=0$. 这证明了 $\lim\limits_{p\to\infty}C_p=0$. 所以证明了 $\exp(A+B)=\exp(A)\exp(B)$. □

下面再进一步考虑指数函数.

由定理 7.3.8 的(4) 可知, 设 n 阶方阵 B 的 n 个特征根为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 那末 n 阶方阵 $\exp(B)$ 的 n 个特征根为 $e^{\lambda_1}, e^{\lambda_2}, \cdots, e^{\lambda_n}$. 所以

$$\det(\exp(B)) = \prod_{k=1}^{n} e^{\lambda_k} = e^{\lambda_1+\lambda_2+\cdots+\lambda_n} = e^{\operatorname{tr}(B)} \neq 0. \tag{7.3.39}$$

因此对任一 n 阶方阵 B, 则 $\exp(B)$ 是 n 阶非异方阵. 反之, 有

定理 7.3.10 设 n 阶复方阵 A 非异, 则存在 n 阶复方阵 B, 使得

$$A=\exp(B). \tag{7.3.40}$$

即当 n 阶复方阵 A 非异时, 矩阵方程

$$\exp(X)=A \tag{7.3.41}$$

必有解.

证 令 A 是非异方阵. 设 A 的全部初等因子为 $(\lambda-\lambda_1)^{e_1}$, $(\lambda-\lambda_2)^{e_2}$, $\cdots$, $(\lambda-\lambda_t)^{e_t}$, 其中 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是 t 个非零复数. 因此存在 n 阶非异复方阵 P, 使得

$$PAP^{-1} = \operatorname{diag}(\lambda_1 M^{(e_1)}, \lambda_2 M^{(e_2)}, \cdots, \lambda_s M^{(e_s)}), \tag{7.3.42}$$

其中 $M^{(e)}=E_e+N_e$ 是特征根为 1 的 Jordan 块 (见 §6.4). 所以, 如果我们能解方阵方程

$$\exp(Y)=\lambda_0 M=\lambda_0(E+N), \quad \lambda_0\neq 0, \tag{7.3.43}$$

那末存在阶数各为 $e_1, e_2, \cdots, e_s$ 的方阵 $Y_1, Y_2, \cdots, Y_s$, 使得 $\exp(Y_j)=\lambda_j M$, $j=1, 2, \cdots, s$. 所以

$$\begin{aligned}PAP^{-1} &= \operatorname{diag}(\lambda_1 M^{(e_1)}, \lambda_2 M^{(e_2)}, \cdots, \lambda_s M^{(e_s)})\\ &= \operatorname{diag}(\exp(Y_1), \exp(Y_2), \cdots, \exp(Y_s)) = \exp(\operatorname{diag}(Y_1, Y_2, \cdots, Y_s)).\end{aligned}$$

即

$$A = P^{-1}\exp(\operatorname{diag}(Y_1, Y_2, \cdots, Y_s))P = \exp(P^{-1}\operatorname{diag}(Y_1, Y_2, \cdots, Y_s)P).$$

因此 $X = P^{-1}\operatorname{diag}(Y_1, Y_2, \cdots, Y_s)P$ 便是所求的解.

现在来求 n 阶方阵方程 $\exp(Y) = \lambda_0 M$ 的解. 由于 $\lambda_0 \neq 0$, 令 $Y = Z+(\ln(\lambda_0))E$, 则由定理 7.3.3, 有

$$\exp(Y) = \exp(Z + (\ln(\lambda_0))E) = \exp(\ln(\lambda_0)E)\exp(Z) = \lambda_0 \exp(A).$$

因此, 如果 $\exp(Y) = \lambda_0 M$, 则 $\exp(Z) = M$. 所以无妨设 $\lambda_0 = 1$.

令 e 阶方阵 $M = E + N$, 其中 E 是 e 阶单位方阵, N 是 e 阶幂零方阵, 它的特征根全是零. 利用通常的级数有 $\exp(\ln(1+z)) = 1 + z, \forall |z| < 1$. 因此,

$$\ln(M) = \ln(E + N) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} N^k = \sum_{k=1}^{e-1} \frac{(-1)^{k-1}}{k} N^k$$

有意义. 所以利用 $N^e = 0$ 及

$$\sum_{i=0}^{\infty} \frac{1}{i!} \Big[\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} z^k\Big]^i = 1 + z, \quad |z| < 1.$$

用 e 阶幂零方阵 N 代入, 易证

$$\sum_{i=0}^{\infty} \frac{1}{i!} \Big[\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} N^k\Big]^i = 1 + N.$$

这证明了

$$\exp(\ln(M)) = M. \tag{7.3.44}$$

所以矩阵方程 $\exp(Z) = M$ 有解 $\ln(M)$. □

注意　设 A 为 n 阶非异复方阵. 上面证明了矩阵方程 $\exp(X) = A$ 有解. 显然它的解不唯一. 事实上, 由于

$$e^{2k\pi\sqrt{-1}} = 1, \quad k = 0, \pm 1, \pm 2, \cdots, \tag{7.3.45}$$

所以若 n 阶复方阵 B 有 $\exp(B) = A$, 则

$$\exp(B + 2k\pi\sqrt{-1}E) = A, \quad k = 0, \pm 1, \pm 2, \cdots. \tag{7.3.46}$$

这个定理也告诉我们, 任一非异方阵都可以开任意次方, 即有

定理 7.3.11　设 A 为 n 阶非异复方阵, α 为非零复数. 则存在 n 阶复方阵 B, 使得 $B^\alpha = A$. 即矩阵方程

$$X^\alpha = A \tag{7.3.47}$$

有解.

证 令存在 n 阶复方阵 C, 使得 $A=\exp(C)$. 令

$$B=\exp\left(\frac{1}{\alpha}C\right),$$

则有 $B^\alpha=\Big(\exp\left(\frac{1}{\alpha}C\right)\Big)^\alpha=\exp(C)=A.$ □

注意 这里 B^α 理解为: 无妨设 B 的特征根 λ_j 有 $|\lambda_j-1|<1$, 则 $B^\alpha=\exp(\alpha\ln(B))$. 又这个推论只给出了矩阵方程的解的存在性. 实际上解并不唯一. 事实上, 若 $B^\alpha=A$, 则有

$$\Big(B\exp\left(\frac{2k\pi\sqrt{-1}}{\alpha}\right)\Big)^\alpha=B^\alpha\exp(2k\pi\sqrt{-1})=B^\alpha=A,\quad k=0,\pm1,\pm2,\cdots.$$

所以方阵方程 $X^\alpha=A$ 有很多解. 又对奇异的方阵, 一般是不可以开任意次方的. 例如, 二阶复方阵 $A=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, 就不存在方阵 X, 使得 $X^2=A$. 读者试自证之.

上面详细地考查了方阵多项式和方阵幂级数, 它们都具有定理 7.3.7 的特性, 即函数值和自变量在相似下的 Jordan 标准形有着密切的关系. 将上面的理论推广, 我们发现关键在于多项式函数和幂级数函数具有三条基本性质. 下面我们利用定理 7.3.7 的特性来引进一类方阵函数.

定义 7.3.12 给定数值函数 $f(z)$. n^2 个独立未知数 x_{ij} 构成的 n 阶方阵为自变量的函数 $Y=f(X)$ 称为**纯函数**, 如果它适合下面性质:

(1) 对一切同阶非异常数复方阵 P 有

$$f(PXP^{-1})=Pf(X)P^{-1};\tag{7.3.48}$$

(2) 对任一准对角方阵 $\operatorname{diag}(X_1,X_2,\cdots,X_s)$, 有

$$f(\operatorname{diag}(X_1,X_2,\cdots,X_s))=\operatorname{diag}(f(X_1),f(X_2),\cdots,f(X_s));\tag{7.3.49}$$

(3) 对任一 n 阶 Jordan 块 $J_0=\lambda_0E_n+N_n$, 当复变量函数 (即一阶复方阵为自变量时) 在 $z=\lambda_0$ 时有 1, 2, $\cdots$, $n-1$ 阶连续导数时, 方阵函数 $f(X)$ 在 $X=J_0$

时才有定义. 这时函数值定义为

$$f(J_0)=\begin{pmatrix} f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \frac{1}{2!}f''(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) \\ & f(\lambda_0) & \frac{1}{1!}f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0) \\ & & f(\lambda_0) & \cdots & \frac{1}{(n-3)!}f^{(n-3)}(\lambda_0) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_0) \end{pmatrix}. \tag{7.3.50}$$

即对角线上为 $f(\lambda_0)$, 其上一斜排为 $\frac{1}{1!}f'(\lambda_0)$, 最后一斜排为 $\frac{1}{(n-1)!}f^{(n-1)}(\lambda_0)$, 所以,

$$f(J_0)=\sum_{k=0}^{n-1}\frac{1}{k!}f^{(k)}(\lambda_0)N^k. \tag{7.3.51}$$

由定义可知, 方阵多项式和方阵幂级数在收敛半径 ρ 内都是纯函数. 纯函数也可以具体构造出来, 这可以由下一定理看出.

定理 7.3.13　方阵函数 $f(X)$ 为纯函数的必要且充分条件是: 如果非异方阵 P 将 X 化为 Jordan 标准形 $PXP^{-1}=\mathrm{diag}\,(J_1,J_2,\cdots,J_s)$, 那末

$$f(X)=P^{-1}\mathrm{diag}\,(f(J_1),f(J_2),\cdots,f(J_s))P, \tag{7.3.52}$$

其中 $f(J_1)$, $f(J_2)$, $\cdots$, $f(J_s)$ 由上面的纯函数定义中的式 (7.3.50) 给出. 因此, 纯函数的定义有意义.

证　先证必要性. 假设 $f(X)$ 为纯函数. 由定义

$$\begin{aligned} f(X)&=f(P^{-1}\mathrm{diag}\,(J_1,J_2,\cdots,J_s)P)=P^{-1}f(\mathrm{diag}\,(J_1,J_2,\cdots,J_s)P \\ &=P^{-1}\mathrm{diag}\,(f(J_1),f(J_2),\cdots,f(J_s))P, \end{aligned}$$

所以必要性成立. 再证充分性. 即证由式 (7.3.52) 定义的函数是纯函数. 先来证明由式 (7.3.52) 定义的函数是单值函数. 因为方阵 X 的标准形 (若不计对角块的次序) 是唯一的, 所以要证单值性, 即要证函数值 $f(X)$ 与非异方阵 P 的选取无关. 今任取 P 和 Q 为非异方阵, 即要证当

$$P^{-1}\mathrm{diag}\,(J_1,J_2,\cdots,J_s)P=Q^{-1}\mathrm{diag}\,(J_1,J_2,\cdots,J_s)Q \tag{7.3.53}$$

时有

$$f(X)=P^{-1}\mathrm{diag}\,(f(J_1),f(J_2),\cdots,f(J_s))P=Q^{-1}\mathrm{diag}\,(f(J_1),f(J_2),\cdots,f(J_s))Q \tag{7.3.54}$$

就可以了. 事实上, 非异方阵 P 和 Q 有关系

$$\operatorname{diag}(J_1, J_2, \cdots, J_s)PQ^{-1} = PQ^{-1}\operatorname{diag}(J_1, J_2, \cdots, J_s).$$

将非异方阵 $B = PQ^{-1}$ 按照 Jordan 标准形 $\operatorname{diag}(J_1, J_2, \cdots, J_s)$ 的分块方法分块为

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1s} \\ \vdots & & \vdots \\ B_{s1} & \cdots & B_{ss} \end{pmatrix},$$

则 $B_{jk}J_k = J_jB_{jk}$, $1 \leqslant j, k \leqslant s$. 由定理 7.2.7, 所以在 $\lambda_j \neq \lambda_k$ 时, $B_{jk} = 0$; 在 $\lambda_j = \lambda_k$ 时, $f(\lambda_j) = f(\lambda_k)$. 这时已知 $B_{jk}J_k = J_jB_{jk}$, 确切地说, 当 $\lambda_j = \lambda_k$, 则

$$B_{jk} = \begin{pmatrix} 0 & \sum\limits_{i=0}^{\min(e_j, e_k)-1} & b_i^{(jk)}N^i \\ 0 & & 0 \end{pmatrix},$$

其中 N 为 $\min(e_j, e_k)$ 阶方阵, 由式 (7.1) 所定义. 所以

$$B_{jk}f(J_k) = \sum_{i=0}^{e_k-1} \frac{1}{i!} f^{(i)}(\lambda_k) B_{jk} N^i.$$

当 $e_j \geqslant e_k$ 时有

$$B_{jk}N_{e_k}^i = \begin{pmatrix} \sum\limits_{p=0}^{e_k-1} b_p^{(jk)} N_{e_k}^p \\ 0 \end{pmatrix} N_{e_k}^i$$

$$= \begin{pmatrix} \sum\limits_{p=0}^{e_k-1} b_p^{(jk)} N_{e_k}^{p+i} \\ 0 \end{pmatrix} = N_{e_j}^i \begin{pmatrix} \sum\limits_{p=0}^{e_k-1} b_p^{(jk)} N_{e_k}^p \end{pmatrix} = N_{e_j}^i B_{jk}.$$

同理, 当 $e_j < e_k$ 时,

$$B_{jk}N_{e_k}^i = \begin{pmatrix} 0 & \sum\limits_{p=0}^{e_j-1} b_p^{(jk)} N_{e_j}^p \end{pmatrix} N_{e_k}^i = N_{e_j}^i \begin{pmatrix} 0 & \sum\limits_{p=0}^{e_j-1} b_p^{(jk)} N_{e_j}^p \end{pmatrix} = N_{e_j}^i B_{jk}.$$

所以证明了当 $\lambda_j = \lambda_k$ 时, 有 $B_{jk}f(J_k) = f(J_i)B_{jk}$. 总之, 证明了非异方阵 $B = PQ^{-1}$ 和 Jordan 标准形 $\operatorname{diag}(J_1, J_2, \cdots, J_s)$ 可交换. 即式 (7.3.54) 成立. 这就证明了 $f(X)$ 的单值性. 至于适不适合纯函数的定义, 这只需直接利用定义来验证就能证明. □

由定义立即有

定理 7.3.14 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 n 阶复方阵 A 的所有特征根. 设 $f(X)$ 为纯函数, 则 $f(A)$ 的所有特征根为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$.

由这个定理可以看出, 纯函数可以由一个具有相当多阶的连续可微的复变函数

$f(z)$ 造出来. 由于纯函数的定义依赖于所给的复函数 $f(z)$, 所以 $\sin(X)$, $\cos(X)$, $\exp(X)$ 都是纯函数, 而 $(E+X)^{\alpha}$ 和 $\ln(E+X)$ 都不是纯函数. 当然, 我们也可以扩充纯函数的概念, 使得"自变量"不能取任意方阵, 而是限制在一定条件下的一批方阵. 换句话说, 我们取一个具有定义域 $\mathfrak{D}$ 的全纯函数 $f(z)$, 只要方阵 A 的特征根必须落在定义域 $\mathfrak{D}$ 中, 那末矩阵函数 $f(X)$ 在 $X=A$ 时有意义. 在这意义下的纯函数, 也就包括了初等函数 $(E+X)^{\alpha}$ 和 $\ln(E+X)$.

纯函数的基本性质依赖于广义 Lagrange 插值公式 (引理 7.2.8).

定理 7.3.15 设 $f(X)$ 是纯函数, 且在 $X=A$ 时函数值 $f(A)$ 有意义, 则 $f(A)$ 是 A 的多项式.

证 令

$$A=P^{-1}\operatorname{diag}(J_1,J_2,\cdots,J_s)P,$$

则

$$f(A)=P^{-1}\operatorname{diag}(f(J_1),f(J_2),\cdots,f(J_s))P.$$

所以要证明 $f(A)$ 是 A 的多项式, 只要证明方阵 $\operatorname{diag}(f(J_1),f(J_2),\cdots,f(J_s))$ 是 Jordan 标准形 $\operatorname{diag}(J_1,J_2,\cdots,J_s)$ 的多项式就行了. 亦即要找一个数值多项式 $g(z)$, 使得

$$\begin{aligned}\operatorname{diag}(f(J_1),f(J_2),\cdots,f(J_s))&=g(\operatorname{diag}(J_1,J_2,\cdots,J_s))\\&=\operatorname{diag}(g(J_1),g(J_2),\cdots,g(J_s)).\end{aligned}$$

所以要找一个数值多项式 $g(z)$, 使得

$$\begin{pmatrix} f(\lambda_j) & \frac{1}{1!}f'(\lambda_j) & \frac{1}{2!}f''(\lambda_j) & \cdots & \frac{1}{(e_j-1)!}f^{(e_j-1)}(\lambda_j) \\ & f(\lambda_j) & \frac{1}{1!}f'(\lambda_j) & \cdots & \frac{1}{(e_j-2)!}f^{(e_j-2)}(\lambda_j) \\ & & f(\lambda_j) & \cdots & \frac{1}{(e_j-3)!}f^{(e_j-3)}(\lambda_j) \\ & & & \ddots & \vdots \\ & & & & f(\lambda_j) \end{pmatrix}$$

$$=\begin{pmatrix} g(\lambda_j) & \frac{1}{1!}g'(\lambda_j) & \frac{1}{2!}g''(\lambda_j) & \cdots & \frac{1}{(e_j-1)!}g^{(e_j-1)}(\lambda_j) \\ & g(\lambda_j) & \frac{1}{1!}g'(\lambda_j) & \cdots & \frac{1}{(e_j-2)!}g^{(e_j-2)}(\lambda_j) \\ & & g(\lambda_j) & \cdots & \frac{1}{(e_j-3)!}g^{(e_j-3)}(\lambda_j) \\ & & & \ddots & \vdots \\ & & & & g(\lambda_j) \end{pmatrix}$$

对一切 $j=1,2,\cdots,s$ 成立. 即要求多项式 $g(x)$, 使得

$$g^{(k)}(\lambda_j)=f^{(k)}(\lambda_j)=a_{jk},\quad 1\leqslant j\leqslant s,\quad 0\leqslant k<=\max_{1\leqslant j\leqslant s}\max\ (e_j).$$

这由引理 7.2.8 可知多项式 $g(x)$ 存在. □

最后, 利用定理 7.3.15 可以证明

定理 7.3.16 对任一非异方阵 A, 存在一个多项式 $g(\lambda)$, 使得 $g(A)^2=A$. 即任一非异方阵都可以开方, 而且开方后为方阵 A 的多项式.

证 考虑数值函数 $f(\lambda)=\sqrt{\lambda}$. 已知 $\sqrt{\lambda}$ 在点 $z=0$ 外为全纯函数. 设 A 非异, 所以特征根都不等于零. 因此纯函数 $f(X)$ 有意义. 所以, 对任一 Jordan 块 J_0 有

$$f(J_0)=f(\lambda_0)E+\frac{1}{1!}f'(\lambda_0)N+\frac{1}{2!}f''(\lambda_0)N^2+\cdots+\frac{1}{(e-1)!}f^{(e-1)}(\lambda_0)N^{e-1}.$$

由直接计算可知 $f(J_0)f(J_0)=J_0$. 所以对任一非异方阵 X, 由于它没有零特征根, 可以利用数值函数 $f(\lambda)=\sqrt{\lambda}$ 来构作一个纯函数 $f(X)$. 取 $X=A$, 则由定理 7.3.15 可知, 存在方阵多项式 $g(A)$, 使得 $g(A)=f(A)$.

设 $PAP^{-1}=\mathrm{diag}\,(J_1,J_2,\cdots,J_s)=J_0$ 为 A 的 Jordan 标准形, 则

$$\begin{aligned} g(A)^2=f(A)^2&=\left(P^{-1}\mathrm{diag}\,(f(J_1),f(J_2),\cdots,f(J_s))P\right)^2\\ &=P^{-1}\mathrm{diag}\,(f(J_1)^2,f(J_2)^2,\cdots,f(J_s)^2)P\\ &=P^{-1}\mathrm{diag}\,(J_1,J_2,\cdots,J_s)P=A.\end{aligned}$$

□

注意 这个定理实际上也告诉了我们如何去构作多项式 $g(A)$, 虽然计算起来并不太容易.

习 题 7.3

7.3.1 记 $X(t)=(x_{ij}(t))$ 为 n 阶方阵, 其中 $x_{ij}(t)$ 是 t 的可微函数. 记

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t}=\left(\frac{\mathrm{d}x_{ij}(t)}{\mathrm{d}t}\right).$$

试求:

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t}=X(t)A$$

的适合初值 $X(0)=A$ 的解, 其中 A 和 $X(t)$ 的阶数相同.

7.3.2 试证:

$$\exp\left(\begin{pmatrix}0 & \lambda\\ -\lambda & 0\end{pmatrix}\right)=\begin{pmatrix}\cos\lambda & \sin\lambda\\ -\sin\lambda & \cos\lambda\end{pmatrix}.$$

7.3.3 试证: 对任一 n 阶复方阵 A, 总有

$$\cos^2 A+\sin^2 A=E_n.$$

7.3.4 设 A 为特征根的模小于 1 的 n 阶复方阵. 试证:

$$\exp\left(\ln(E_n+A)\right)=E_n+A.$$

7.3.5 设 $f_1(x)$ 和 $f_2(x)$ 为光滑函数, X 为 n^2 个独立未知数构成的 n 阶方阵. 记 $f_1(X)$ 和 $f_2(X)$ 为纯函数. 试证: $f_1(X')=f_1(X)'$; 又记 $f_3(x)=f_1(x)+f_2(x)$, $f_4(x)=f_1(x)f_2(x)$, 则纯函数 $f_3(X)$ 和 $f_4(X)$ 仍有意义, 且 $f_3(X)=f_1(X)+f_2(X)$, $f_4(X)=f_1(X)f_2(X)$.

7.3.6 试证: 由 $f(x)=x^{-1}$ 定义的纯函数 $f(X)$ 只有在 n 阶方阵 $X=A$ 非异时才有意义. 这时 $f(A)=A^{-1}$. 又由 $f(x)=x^{\frac{a}{b}}$ 定义的纯函数 $f(X)$, 它在有意义的 n 阶方阵 $X=A$ 上, 总有 $f(A)^b=A^a$, 其中 a 和 b 是正整数.

7.3.7 试证: n 阶非异复方阵 A 的逆方阵 A^{-1} 为 A 的多项式. 又对任意非零整数 m, 则存在 A 的多项式 $g_m(x)$, 使得 $g_m(A)^m=A$. 特别, 当 A 是复对称方阵时, $S_m=g_m(A)$ 也复对称.

§7.4 方阵在复相似下的标准形 *

前面考虑了相似关系: $A\to PAP^{-1}$, 其中 P 是非异方阵. 可以看出, 它们都是形如 $A\to PAf(P)$ 的关系, 其中函数 $f(P)$ 是将非异方阵 P 取转置运算、共轭运算和取逆运算. 将这三种运算作各种可能的组合, 由式 (3.5.3) 可知, 它们可以导出八种关系, 其中只有四种关系: $A\to PAP^{-1}$, $A\to PA\overline{P}^{-1}$ $A\to PAP'$, $A\to PA\overline{P}'$ 是等价关系.

在这一节, 我们考虑等价关系 $A\to PA\overline{P}^{-1}$, 求出它的标准形和全系不变量. 这一结果是许宝騄教授在 1955 年获得的. 下面的讨论全在复数范围内进行.

定义 7.4.1 n 阶方阵 A 和 B 称为**复相似的**, 如果存在 n 阶非异方阵 P, 使得

$$B=PA\overline{P^{-1}}.\tag{7.4.1}$$

显然有

引理 7.4.2 两 n 阶复方阵的复相似关系是等阶关系.

引理 7.4.3 若 n 阶复方阵 A 和 B 复相似, 则 n 阶复方阵 $A\overline{A}$ 和 $B\overline{B}$ 相似. 因

此, n 阶复方阵 $A\overline{A}$ 的初等因子组是 A 在复相似下的不变量.

证 今存在非异方阵 P, 使得 $B = PA\overline{P^{-1}}$, 所以

$$B\overline{B} = (PA\overline{P^{-1}})\overline{(PA\overline{P^{-1}})} = PA\overline{P^{-1}}\ \overline{PA}P^{-1} = P(A\overline{A})P^{-1}.$$

这就证明了引理. □

这引理的逆是不成立的. 例如, 取

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

则

$$A\overline{A} = B\overline{B} = 0.$$

但是 A 和 B 的秩不同, 所以 A 和 B 不可能复相似. 不过, 对方阵 A 加上可逆的条件后, 逆定理便成立, 即有

定理 7.4.4 设 n 阶复方阵 A 和 B 都是非异复方阵, 则 A 和 B 复相似当且仅当 n 阶复方阵 $A\overline{A}$ 和 $B\overline{B}$ 相似. 因此, n 阶复方阵 $A\overline{A}$ 的初等因子组是 n 阶复方阵 A 在复相似下的全系不变量.

证 今存在 n 阶非异方阵 P, 使得 $PA\overline{A}P^{-1} = B\overline{B}$, 亦即 $PA\overline{A} = B\overline{B}P$. 我们的目的是要找一 n 阶非异方阵 Q, 使得 $QA = B\overline{Q}$. 事实上,

$$(B\overline{B}P)A = B(\overline{B}PA) = B\overline{B\overline{PA}}, \quad B\overline{PA}A = B(\overline{PA}A) = B\overline{PA\overline{A}}.$$

因此, 对任一实数 θ, 令 $i = \sqrt{-1}$,

$$Q(\theta) = (\exp(i\theta))B\overline{B}P + (\exp(-i\theta))B\overline{PA},$$

则

$$Q(\theta)A = (\exp(i\theta))B\overline{B}PA + (\exp(-i\theta))B\overline{PA}A,$$

$$\begin{aligned} B\overline{Q(\theta)} &= (\exp(i\theta))B\overline{B}PA + (\exp(-i\theta))B\overline{B}B\overline{P} \\ &= (\exp(i\theta))B\overline{B\overline{PA}} + (\exp(-i\theta))\overline{P(A\overline{A})} \\ &= \big((\exp(i\theta))B\overline{B}P + (\exp(-i\theta))B\overline{PA}\big)\,A = Q(\theta)\,A. \end{aligned}$$

而

$$\begin{aligned} \det(Q(\theta)) &= \det((\exp(i\theta))B\overline{B}P + (\exp(-i\theta))B\overline{PA}) \\ &= \det((\exp(-i\theta))B\overline{B}P)((\exp(2i\theta))E_n + P^{-1}\overline{B^{-1}}\ \overline{PA}). \end{aligned}$$

显然 $\det((\exp(i\theta))B\overline{B}P) \neq 0$. 因为 n 阶方阵 $P^{-1}\overline{B^{-1}}\ \overline{PA}$ 的不同特征根至多 n 个, 所以必定存在实数 θ_0, 使得 $\det(\exp(2i\theta_0)E_n + P^{-1}\overline{B^{-1}}\ \overline{PA}) \neq 0$, 即 $\det(Q(\theta_0)) \neq 0$. 这证明了 n 阶非异复方阵 A 和 B 复相似. □

由 n 阶复方阵 $A\overline{A}$ 和 $B\overline{B}$ 相似, 所以存在 n 阶非异复方阵 Q, 使得 $QA\overline{A}Q^{-1} = J = \operatorname{diag}(J_1, J_2, \cdots, J_t)$, 其中

$$J = \operatorname{diag}(N_{e_1}, N_{e_2}, \cdots, N_{e_s}, \lambda_{s+1}M^{(e_{s+1})}, \cdots, \lambda_t M^{(e_t)}), \tag{7.4.2}$$

M_j 由式 (7.1) 定义, 下面的方阵 M 都由式 (7.1) 定义. $\lambda_{s+1}, \cdots, \lambda_t$ 是非零复数. 这里约定 $\lambda_{s+1}, \cdots, \lambda_t$ 的排法是, 先排实特征根, 将具有相同实特征根的 Jordan 块排在一起; 再排复特征根, 将具有相同复特征根的 Jordan 块排在一起, 且在排了以 λ_j 为特征根的 Jordan 块后, 如果 $\overline{\lambda_j}$ 仍为特征根, 则接下去排以 $\overline{\lambda_j}$ 为特征根的 Jordan 块.

将方阵 $B = QA\overline{Q^{-1}}$ 按照 $B\overline{B} = QA\overline{A}Q^{-1} = J$ 的方式分块为

$$B = \begin{pmatrix} B_{11} & \cdots & B_{1t} \\ \vdots & & \vdots \\ B_{t1} & \cdots & B_{tt} \end{pmatrix}. \tag{7.4.3}$$

由 $B\overline{B} = QA\overline{A}Q^{-1} = J$ 以及 $(B\overline{B})B = B(\overline{B}B) = B(\overline{B\overline{B}})$ 可知 $JB = B\overline{J}$. 因此有

$$\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_t \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1t} \\ \vdots & & \vdots \\ B_{t1} & \cdots & B_{tt} \end{pmatrix} = \begin{pmatrix} B_{11} & \cdots & B_{1t} \\ \vdots & & \vdots \\ B_{t1} & \cdots & B_{tt} \end{pmatrix} \begin{pmatrix} \overline{J_1} & & \\ & \ddots & \\ & & \overline{J_t} \end{pmatrix}. \tag{7.4.4}$$

即

$$J_j B_{jk} = B_{jk}\overline{J_k} \quad j, k = 1, 2, \cdots, t. \tag{7.4.5}$$

由定理 7.2.7 可知, 当 J_j 和 $\overline{J_k}$ 的特征根不相同, 即 $\lambda_j \neq \overline{\lambda_k}$ 时, $B_{jk} = 0$; 当 J_j 和 $\overline{J_k}$ 的特征根相同, 即 $\lambda_j = \overline{\lambda_k}$, 而且当 $e_j \geqslant e_k$ 时, 有

$$B_{jk} = \begin{pmatrix} \sum\limits_{i=0}^{e_k-1} a_i^{(jk)} N_k^i \\ 0 \end{pmatrix}. \tag{7.4.6}$$

又当 $e_j < e_k$ 时, 有

$$B_{jk} = \begin{pmatrix} 0 & \sum\limits_{i=0}^{e_j-1} a_i^{(jk)} N_j^i \end{pmatrix}. \tag{7.4.7}$$

因此, 将 $J=QA\overline{A}Q^{-1}$ 的对角块按照上述约定排起来后, 相应的方阵 B 也必然分成准对角形. 确切地说, 问题化为在下列三种情形时, 求 n 阶方阵 B 在复相似下的标准形:

情形 (一)

$$B\overline{B}=\operatorname{diag}(N_{e_1},N_{e_2},\cdots,N_{e_s}), \tag{7.4.8}$$

其中 $e_1\geqslant e_2\cdots\geqslant e_s>0$.

情形 (二) 设 λ_0 为非零复数, 而且 $\operatorname{Im}(\lambda_0)\neq 0$,

$$B\overline{B}=\operatorname{diag}(\lambda_0M^{(e_1)},\overline{\lambda_0}M^{(f_1)},\cdots,\lambda_0M^{(e_s)},\overline{\lambda_0}M^{(f_s)}), \tag{7.4.9}$$

其中 $e_1\geqslant\cdots\geqslant e_s>0$, $f_1\geqslant\cdots\geqslant f_s\geqslant 0$.

情形 (三) 设 λ_0 为非零实数,

$$B\overline{B}=\lambda_0\operatorname{diag}(M^{(e_1)},\cdots,M^{(e_s)}), \tag{7.4.10}$$

其中 $e_1\geqslant\cdots\geqslant e_s>0$.

写成初等因子的形式, 即在下面三种情形下, 求方阵 B 在复相似下的标准形:

情形 (一) n 阶方阵 $B\overline{B}$ 的初等因子组为

$$\lambda^{e_1},\lambda^{e_2},\cdots,\lambda^{e_s}, \tag{7.4.11}$$

其中 $e_1\geqslant e_2\cdots\geqslant e_s>0$.

情形 (二) n 阶方阵 $B\overline{B}$ 的初等因子组为

$$(\lambda-\lambda_0)^{e_1},(\lambda-\overline{\lambda_0})^{f_1},\cdots,(\lambda-\lambda_0)^{e_s},(\lambda-\overline{\lambda_0})^{f_s}, \tag{7.4.12}$$

其中 λ_0 是虚部不为零的复数, 又 $e_1\geqslant\cdots\geqslant e_s>0$, $f_1\geqslant\cdots\geqslant f_s\geqslant 0$.

情形 (三) n 阶方阵 $B\overline{B}$ 的初等因子组为

$$(\lambda-\lambda_0)^{e_1},\cdots,(\lambda-\lambda_0)^{e_s}, \tag{7.4.13}$$

其中 λ_0 为非零实数, 又 $e_1\geqslant\cdots\geqslant e_s>0$.

下面分别来求方阵 B 在复相似下的标准形.

情形 (一) 下面先在第一种情形下求标准形.

定理 7.4.5 条件如上, 则 n 阶方阵 A 复相似于唯一的标准形

$$C=\operatorname{diag}(N_{g_1},N_{g_2},\cdots,N_{g_s}), \tag{7.4.14}$$

n 阶方阵 C 的初等因子组为

$$\lambda^{g_1}, \lambda^{g_2}, \cdots, \lambda^{g_s}, \tag{7.4.15}$$

其中 $g_1 \geqslant \cdots \geqslant g_t > 0$.

证 对方阵 A 的阶数 n 作归纳法. 在 $n=1$ 时 A 为零方阵, C 当然是零方阵, 故定理成立. 假设对一切 $n-1$ 阶方阵 A_1, 只要 $A_1\overline{A_1}$ 的特征根都是零时, 定理成立. 现在来证明对 n 阶方阵 A, 只要 $A\overline{A}$ 的特征根都为零时, 定理也成立.

首先, 由于 $A\overline{A}$ 的特征根全是零, 所以 n 阶复方阵 A 奇异. 因此存在 n 阶复方阵 A 的属于特征根零的特征向量 α_1, 它是 $n\times 1$ 复矩阵, 且有 $A\alpha_1=0$. 以 α_1 为第一个列向量, 作一个 n 阶非异复方阵 P_1, 记

$$P_1=(\alpha_1,\alpha_2,\cdots,\alpha_n),$$

则

$$B_0=\overline{P_1^{-1}}AP_1=\overline{P_1^{-1}}A(\alpha_1,\alpha_2,\cdots,\alpha_n)=\overline{P_1^{-1}}(0,\beta_2,\beta_3,\cdots,\beta_n).$$

即

$$B_0=\overline{P_1^{-1}}AP_1=\begin{pmatrix}0 & v^{(1,n-1)}\\ 0 & B_1^{(n-1)}\end{pmatrix}.$$

今

$$B_0\overline{B_0}=\overline{P_1^{-1}}A\overline{AP_1}=\begin{pmatrix}0 & v\\ 0 & B_1\end{pmatrix}\begin{pmatrix}0 & \overline{v}\\ 0 & \overline{B_1}\end{pmatrix}=\begin{pmatrix}0 & v\overline{B_1}\\ 0 & B_1\overline{B_1}\end{pmatrix}.$$

而 $A\overline{A}$ 的特征根皆为零, 所以 $B_0\overline{B_0}$ 的特征根也皆为零. 因此, $B_1\overline{B_1}$ 的特征根也皆为零. 由归纳法假设, 存在 $n-1$ 阶非异复方阵 P_2, 使得

$$\overline{P_2^{-1}}B_1P_2=\mathrm{diag}\,(N_{h_1},\cdots,N_{h_t}),$$

则复方阵 A 复相似于复方阵

$$B_2=\begin{pmatrix}1 & 0\\ 0 & \overline{P_2}\end{pmatrix}^{-1}\overline{P_1^{-1}}AP_1\begin{pmatrix}1 & 0\\ 0 & P_2\end{pmatrix}=\begin{pmatrix}0 & vP_2\\ 0 & \overline{P_2^{-1}}B_1P_2\end{pmatrix},$$

所以

$$B_2=\begin{pmatrix}0 & v_1 & \cdots & v_t\\ 0 & N_{h_1} & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & N_{h_t}\end{pmatrix},$$

其中 v_j 是 $1\times h_j$ 复矩阵, $j=1,2,\cdots,t$, 且 $(v_1,\cdots,v_t)=vP_2$. 再作 n 阶非异复方阵

$$P_3=\begin{pmatrix}1 & (\overline{v_1}N'_{h_1},\cdots,\overline{v_t}N'_{h_t})\\ 0 & E_{n-1}\end{pmatrix},$$

则

$$\overline{P_3^{-1}}\begin{pmatrix} 0 & v_1 & \cdots & v_t \\ 0 & N_{h_1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & N_{h_t} \end{pmatrix} P_3$$

$$= \begin{pmatrix} 0 & v_1(E_{h_1} - N'_{h_1}N_{h_1}) & \cdots & v_t(E_{h_t} - N'_{h_t}N_{h_t}) \\ 0 & N_{h_1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & N_{h_t} \end{pmatrix}.$$

由于 $N'_{h_j}N_{h_j} = \operatorname{diag}(0, E_{h_j})$, 所以

$$v_j(E_{h_j} - N'_{h_j}N_{h_j}) = a_j(1, 0, \cdots, 0) \in \mathbb{C}^{h_j}, \quad j = 1, 2, \cdots, t.$$

因此复方阵 A 复相似于复方阵

$$B_3 = \begin{pmatrix} 0 & a_1(1, 0, \cdots, 0) & \cdots & a_t(1, 0, \cdots, 0) \\ 0 & N_{h_1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & N_{h_t} \end{pmatrix}.$$

如果 $(a_1, a_2, \cdots, a_t) = 0$, 则复方阵 A 复相似于复方阵

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & N_{h_1} & & \\ \vdots & & \ddots & \\ 0 & & & N_{h_t} \end{pmatrix} = \operatorname{diag}(0, N_{h_1}, N_{h_2}, \cdots, N_{h_t}),$$

令 $s = t + 1$, $g_1 = 1$, $g_2 = h_1$, $g_3 = h_2$, $\cdots$, $g_s = h_{s-1}$, 所以定理成立.

如果 $a_1, \cdots, a_t$ 不全为零, 对 B_3 的行及列作同步的改变, 从而无妨设 $a_1 \cdots a_r \neq 0$, $a_{r+1} = \cdots = a_t = 0$, $h_1 \geqslant h_2 \geqslant \cdots \geqslant h_r$. 记

$$P_4 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \Lambda_1 & \Lambda_2 & \cdots & \Lambda_r & 0 \\ 0 & 0 & \Lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & 0 & 0 \\ 0 & 0 & 0 & \cdots & \Lambda_r & 0 \\ 0 & 0 & 0 & \cdots & 0 & E \end{pmatrix},$$

其中 $\Lambda_j = \mathrm{diag}\,(a_j^{-1}, \overline{a_j^{-1}}, a_j^{-1}, \overline{a_j^{-1}}, \cdots)$ 为 h_j 阶非异复方阵, 则由 $N_j\Lambda_j = \overline{\Lambda_j}N_j$ 可知, 复方阵 A 复相似于复方阵

$$B_4 = \overline{P_4^{-1}}B_3P_4 = \begin{pmatrix} 0 & w_1 & w_2 & \cdots & w_r & 0 \\ 0 & N_{h_1} & 0 & \cdots & 0 & 0 \\ 0 & 0 & N_{h_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N_{h_r} & 0 \\ 0 & 0 & 0 & \cdots & 0 & T \end{pmatrix},$$

其中 $T = \mathrm{diag}\,(N_{h_{r+1}}, \cdots, N_{h_t})$, $w_j = (1, 0, \cdots, 0)$ 是 $1 \times h_j$ 实方阵, $j = 1, \cdots, r$.

最后我们来作 n 阶非异实方阵,

$$P_5 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & E_1 & X_2 & \cdots & X_r & 0 \\ 0 & 0 & E_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & 0 & 0 \\ 0 & 0 & 0 & \cdots & E_r & 0 \\ 0 & 0 & 0 & \cdots & 0 & E \end{pmatrix},$$

其中 X_j 为 h_j 阶实方阵, $j = 2, \cdots, r$, 于是复方阵 A 复相似于实方阵

$$B_5 = \overline{P_5^{-1}}B_4P_5 = \begin{pmatrix} 0 & w_1 & w_1X_2 + w_2 & \cdots & w_1X_r + w_r & 0 \\ 0 & N_{h_1} & N_1X_2 - X_2N_2 & \cdots & N_1X_r - X_rN_r & 0 \\ 0 & 0 & N_{h_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N_{h_t} & 0 \\ 0 & 0 & 0 & \cdots & 0 & T \end{pmatrix},$$

其中 $T = \mathrm{diag}\,(N_{h_{r+1}}, \cdots, N_{h_t})$. 由 $h_1 \geqslant h_2 \geqslant \cdots \geqslant h_r$ 及定理 7.2.7, 可取

$$X_j = \begin{pmatrix} -E_{h_j} \\ 0 \end{pmatrix}, \quad 2 \leqslant j \leqslant r,$$

则 $N_{h_1}X_j = X_jN_{h_j}$, $w_1X_j + w_j = 0$, $2 \leqslant j \leqslant r$. 所以

$$B_5 = \overline{P_5^{-1}}B_4P_5 = \begin{pmatrix} 0 & (1, 0, \cdots, 0) & 0 & \cdots & 0 & 0 \\ 0 & N_{h_1} & 0 & \cdots & 0 & 0 \\ 0 & 0 & N_{h_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N_{h_t} & 0 \\ 0 & 0 & 0 & \cdots & 0 & T \end{pmatrix},$$

其中 $T=\mathrm{diag}\,(N_{h_{r+1}},\cdots,N_{h_t})$. 因此, 复方阵 A 复相似于实方阵

$$B_5=\mathrm{diag}\,(N_{h_1+1},N_{h_2},\cdots,N_{h_t}).$$

令 $s=t$, $g_1=h_1+1$, $g_2=h_2$, $g_3=h_3$, $\cdots$, $g_s=h_s$, 这就是标准形.

下面证明标准形的唯一性. 设 $f_1\geqslant f_2\geqslant\cdots\geqslant f_s$, $g_1\geqslant g_2\geqslant\cdots\geqslant g_t$, 且 n 阶方阵

$$D_1=\mathrm{diag}\,(N_{f_1},N_{f_2},\cdots,N_{f_s}),\quad D_2=\mathrm{diag}\,(N_{g_1},N_{g_2},\cdots,N_{g_t})$$

复相似. 所以存在 n 阶非异复方阵 $P=X+\sqrt{-1}Y$, 使得 $D_2=\overline{P^{-1}}D_1P$, 其中 X 和 Y 为 n 阶实方阵. 因此有

$$D_1X=XD_2,\quad D_1Y=-YD_2,$$

即

$$\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}\begin{pmatrix} D_1 & 0 \\ 0 & -D_1 \end{pmatrix}=\begin{pmatrix} D_2 & 0 \\ 0 & -D_2 \end{pmatrix}\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}.$$

今记

$$Q=\frac{1}{\sqrt{2}}\begin{pmatrix} E & -iE \\ E & iE \end{pmatrix}.$$

则

$$\det\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}=\det\left(Q\begin{pmatrix} X & Y \\ -Y & X \end{pmatrix}\overline{Q^{-1}}\right)=\det\begin{pmatrix} P & 0 \\ 0 & \overline{P} \end{pmatrix}=|\det(P)|^2>0.$$

这证明了 $\mathrm{diag}\,(D_1,-D_1)$ 和 $\mathrm{diag}\,(D_2,-D_2)$ 实相似. 因此, 它们的初等因式组 λ^{f_1}, $\cdots$, λ^{f_s}, λ^{f_1}, $\cdots$, λ^{f_s}, 和 λ^{g_1}, $\cdots$, λ^{g_t}, λ^{g_1}, $\cdots$, λ^{g_t} 为同一集合. 所以 $t=s$, $f_j=g_j$, $1\leqslant j\leqslant s$. 这也证明了 $D_1=D_2$. 所以标准形在复相似下唯一. □

这个定理说明了: 如果 A 复相似于标准形 $C=\mathrm{diag}\,(N_{g_1},\cdots,N_{g_s})$, 则正整数 $g_1\geqslant g_2\geqslant\cdots\geqslant g_s$ 是方阵 A 在复相似下的全系不变量. 但是上面的讨论并没有给出下列结论: 如何不通过将 A 化为标准形的途径而求出它在复相似下的全系不变量, 这问题留给读者考虑.

情形 (二) 现在在第二种情形下求标准形.

注意: 由定理 7.4.4 可知, 在第二, 三种情形, n 阶复方阵 A 都是非异复方阵. 因此, n 阶复方阵 $A\overline{A}$ 的初等因子组就是 n 阶复方阵 A 在复相似下的全系不变量. 而且我们只要具体构造一个 n 阶复方阵 C, 它具有最简单的形式, 使得 n 阶复方阵 $C\overline{C}$ 和 $A\overline{A}$ 相似, 即有相同的初等因式组, 则 n 阶复方阵 B 和 C 就复相似.

现在在第二种情形下求标准形. 设 λ_0 是非零复数, 而且 $\overline{\lambda_0} \neq \lambda_0$, 即 $\operatorname{Im}(\lambda_0) \neq 0$. 于是在复相似意义下可取 $A\overline{A}$ 为标准形:

$$A\overline{A} = \operatorname{diag}(\lambda_0 M^{(e_1)}, \cdots, \lambda_0 M^{(e_s)}, \overline{\lambda_0} M^{(f_1)}, \cdots, \overline{\lambda_0} M^{(f_t)}), \tag{7.4.16}$$

其中 $M^{(e)} = E_e + N_e$. 所以 n 阶复方阵 $A\overline{A}$ 的初等因式组为

$$(\lambda - \lambda_0)^{e_1}, (\lambda - \overline{\lambda_0})^{f_1}, \cdots, (\lambda - \lambda_0)^{e_s}, (\lambda - \overline{\lambda_0})^{f_t}. \tag{7.4.17}$$

定理 7.4.6 条件同上, 这时 n 阶复方阵 A 复相似于标准形

$$C = \operatorname{diag}\left(\begin{pmatrix} 0 & \sqrt{\lambda_0} M^{(e_1)} \\ \sqrt{\overline{\lambda_0}} M^{(e_1)} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & \sqrt{\lambda_0} M^{(e_s)} \\ \sqrt{\overline{\lambda_0}} M^{(e_s)} & 0 \end{pmatrix}\right), \tag{7.4.18}$$

所以初等因式组为

$$(\lambda - \lambda_0)^{e_1}, \ \cdots, (\lambda - \lambda_0)^{e_s}, \ (\lambda - \overline{\lambda_0})^{e_1}, \ \cdots, (\lambda - \overline{\lambda_0})^{e_s}. \tag{7.4.19}$$

证 令 $A\overline{A} = A(\overline{A}A)A^{-1}$, 所以 $A\overline{A}$ 和 $\overline{A}A$ 相似. 已知 $A\overline{A}$ 的初等因式组为 (7.4.17), 于是 $\overline{A}A = \overline{A\overline{A}}$ 的初等因式组为

$$(\lambda - \overline{\lambda_0})^{e_1}, (\lambda - \lambda_0)^{f_1}, \cdots, (\lambda - \overline{\lambda_0})^{e_s}, (\lambda - \lambda_0)^{f_t}.$$

由于 $\overline{\lambda_0} \neq \lambda_0$, 这证明了 $t = s$, $f_j = e_j$, $1 \leqslant j \leqslant s$. 即初等因式组为式 (7.4.19).

另一方面, 由式 (7.4.18),

$$C\overline{C} = \operatorname{diag}(\lambda_0 (M^{(e_1)})^2, \overline{\lambda_0}(M^{(e_1)})^2, \cdots \lambda_0 (M^{(e_s)})^2, \overline{\lambda_0}(M^{(e_s)})^2).$$

而记 $M^{(e_k)} = M_k$, 则 $M^2 = (E+N)^2 = E + 2N + N^2$ 的初等因式组为 $(\lambda - \lambda_0)^{e_k}$. 这证明了 $C\overline{C}$ 的初等因式组为式 (7.4.19). 所以 $C\overline{C}$ 和 $A\overline{A}$ 有相同的初等因式组. 由定理 7.4.4, A 和 C 复相似. □

情形 (三) 现在在第三种情形下求标准形.

现在在第三种情形下求标准形. 设 λ_0 是非零实数, 即 $\overline{\lambda_0} = \lambda_0$. 于是在复相似意义下可取 $A\overline{A}$ 为标准形:

$$A\overline{A} = \lambda_0 \operatorname{diag}(M^{(e_1)}, \cdots, M^{(e_s)}), \tag{7.4.20}$$

其中 $M^{(e)} = E_e + N_e$ 为 e 阶方阵, 所以 $\overline{M} = M$. 又 n 阶方阵 $A\overline{A}$ 的初等因式组为

$$(\lambda - \lambda_0)^{e_1}, \quad \cdots, \quad (\lambda - \lambda_0)^{e_s}. \tag{7.4.21}$$

这时分 $\lambda_0 > 0$ 及 $\lambda_0 < 0$ 这两种情况来讨论.

定理 7.4.7 条件同上. 设 $\overline{\lambda_0} = \lambda_0 > 0$, 这时 n 阶复方阵 A 复相似于标准形

$$C = \sqrt{\lambda_0} \operatorname{diag}(M^{(e_1)}, M^{(e_2)}, \cdots, M^{(e_s)}), \tag{7.4.22}$$

所以初等因式组为

$$(\lambda-\lambda_0)^{e_1}, \quad \cdots, \quad (\lambda-\lambda_0)^{e_s}, \tag{7.4.23}$$

其中 $e_1 \geqslant e_2 \geqslant \cdots \geqslant e_s$.

证 令

$$C\overline{C} = \lambda_0 \mathrm{diag}\,(M^{(e_1)}M^{(e_1)}, \cdots, M^{(e_s)}M^{(e_s)}),$$

其中 $M^{(e)} = E^{(e)} + N^{(e)}$ 为 e 阶方阵, 所以 $\overline{M} = M$, $M\overline{M} = M^2 = E + 2N + N^2$. 而 M^2 的初等因子组为 $(\lambda-1)^e$, 所以和 M 的初等因子组相同, 即 M^2 和 M 相似. 这证明了 $C\overline{C}$ 和 $A\overline{A}$ 相似, 因此 C 和 A 复相似. 再, 方阵 C 由 $A\overline{A}$ 的初等因子组完全确定, 所以 C 就是标准形. □

定理 7.4.8 条件同上. 设 $\overline{\lambda_0} = \lambda_0 < 0$, 则这时 $s = 2t$, $e_1 = 2f_1$, $e_2 = 2f_2$, $\cdots$, $e_s = 2f_s$ 为偶数, 而且 n 阶复方阵 A 复相似于标准形

$$C = \sqrt{-\lambda_0}\mathrm{diag}\,(\begin{pmatrix} 0 & M^{(f_1)} \\ -M^{(f_1)} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & M^{(f_t)} \\ -M^{(f_t)} & 0 \end{pmatrix}), \tag{7.4.24}$$

所以初等因式组为

$$(\lambda-\lambda_0)^{f_1}, (\lambda-\lambda_0)^{f_1} \quad \cdots, \quad (\lambda-\lambda_0)^{f_t}, (\lambda-\lambda_0)^{f_t}, \tag{7.4.25}$$

其中 $f_1 \geqslant f_2 \geqslant \cdots \geqslant f_t > 0$.

证 令

$$C\overline{C} = \lambda_0 \mathrm{diag}\,((M^{(f_1)})^2, (M^{(f_1)})^2, \cdots, (M^{(f_t)})^2, (M^{(f_t)})^2),$$

其中 $M^{(e)} = E^{(e)} + N^{(e)}$ 为 e 阶方阵. 所以 $\overline{M} = M$, $M\overline{M} = M^2 = E + 2N + N^2$. 而 M^2 的初等因子组为 $(\lambda-1)^e$, 所以 $C\overline{C}$ 的初等因子组为式 (7.4.25). 由定理 7.4.4, 为了证明定理, 只要证明 $e_1, e_2, \cdots, e_s$ 都是偶数就够了. 即设

$$e_1 \geqslant \cdots \geqslant e_p > e_{p+1} = e_{p+2} = \cdots = e_{p+q} > e_{p+q+1} \geqslant e_{p+q+2} \geqslant \cdots \geqslant e_s > 0.$$

则要证 q 为偶数.

事实上, 由 $A\overline{A} = \lambda_0 \mathrm{diag}\,(M^{(e_1)}, M^{(e_2)}, \cdots, M^{(e_s)})$. 将 A 和 $A\overline{A}$ 一样分块为

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{ss} \end{pmatrix}.$$

于是由 $A\overline{A}$ 是 n 阶实方阵, 所以 $A\overline{A} = \overline{A}A$, 因此 $(A\overline{A})A = A(\overline{A}A) = A(A\overline{A})$. 即有 $M^{(e_i)}A_{ij} = A_{ij}M^{(e_j)}$, $1 \leqslant i, j \leqslant s$. 但是 $M^{(e_k)} = E_{e_k} + N_{e_k}$, $1 \leqslant k \leqslant s$, 所以有

$N_{e_i}A_{ij}=A_{ij}N_{e_j}$, 其中 A_{ij} 是 $e_i\times e_j$ 矩阵, $1\leqslant i,j\leqslant s$. 由定理 7.2.7 有

$$A_{ij}=\sum_{k=0}^{\min(e_i,e_j)-1}a_{ij}^{(k)}\begin{pmatrix}0 & N_{\min(e_i,e_j)}^k\\ 0 & 0\end{pmatrix},\quad 1\leqslant i,j\leqslant s,$$

其中 $N_{\min(e_i,e_j)}$ 为 $\min(e_i,e_j)$ 阶幂零方阵. 但是由

$$A\overline{A}=\lambda_0\mathrm{diag}\,(M^{(e_1)},M^{(e_2)},\cdots,M^{(e_s)}),$$

所以有

$$\sum_{j=1}^{s}A_{ij}\overline{A_{jk}}=\delta_{ik}\lambda_0M^{(e_i)}=\delta_{ik}\lambda_0(E_{e_i}+N_{e_i}).$$

考虑 $i,k=p+1,p+2,\cdots,p+q$, 这时有 $e_i=e_k$. 当 $j\neq p+1,p+2,\cdots,p+q$ 时有 $e_j\neq e_i=e_k$. 又这时 $A_{ij}A_{jk}$ 的对角元素为零. 所以在和式 $\sum\limits_{j=1}^{s}A_{ij}\overline{A_{jk}}$ 中只要考虑和式 $\sum\limits_{j=p+1}^{p+q}A_{ij}\overline{A_{jk}}$ 的对角元素就可以了. 今

$$\sum_{j=p+1}^{p+q}A_{ij}\overline{A_{jk}}=\sum_{j=p+1}^{p+q}\sum_{u,v=0}^{e_{p+1}-1}a_{ij}^{(u)}a_{jk}^{(v)}N_{e_i}^{u+v},$$

它的对角元素为 $\sum\limits_{j=p+1}^{p+q}a_{ij}^{(0)}a_{jk}^{(0)}$. 这证明了 $\sum\limits_{j=p+1}^{p+q}a_{ij}^{(0)}a_{jk}^{(0)}=\delta_{ik}\lambda_0$, $i,k=p+1,p+2,\cdots,p+q$. 所以 q 阶方阵

$$T=\begin{pmatrix}a_{p+1,p+1}^{(0)} & \cdots & a_{p+1,p+q}^{(0)}\\ \vdots & & \vdots\\ a_{p+q,p+1}^{(0)} & \cdots & a_{p+q,p+q}^{(0)}\end{pmatrix}$$

有 $T\overline{T}=\lambda_0E_q$. 双方取行列式, 有 $\lambda_0^q=|\det(T)|^2>0$. 但是 $\lambda_0<0$, 这证明了 q 为偶数. □

最后给出复相似的一个应用.

定理 7.4.9　设 A 为 n 阶复方阵, 则存在 Hermite 方阵 H_1 和 H_2, 又存在 n 阶复对称方阵 S_1 和 S_2, 使得

$$A=H_1S_1=S_2H_2,\tag{7.4.26}$$

而且还可以事先规定哪一个是非异的.

证　显然只证 $A=H_1S_1$ 即可. 因为将 n 阶复方阵 A' 分解为 $A'=H_1S_1$, 则 $A=S_1H_1'$, 所以记 $S_2=S_1$, $H_2=H_1'$, 则 S_2 是复对称方阵, H_2 是 Hermite 方阵, 且 $A=S_2H_2$.

今若有分解 $A=H_1S_1$, 则对任一 n 阶复非异方阵 P 有

$$\overline{P^{-1}}AP=(\overline{P^{-1}}H_1(P^{-1})')(P'S_1P),$$

其中 $H_3=\overline{P^{-1}}H_1(P^{-1})'$ 仍为 Hermite 方阵, $S_3=P'S_1P$ 仍为复对称方阵. 所以问题化为对 n 阶复方阵 A 在复相似下的标准形来证明定理就可以了.

记 $E=E_n$, $N=N_n$, $M=E+N$,

$$\widetilde{M}=\begin{pmatrix} & & & & 1\\ & & & \cdot & \\ & & \cdot & & \\ & \cdot & & & \\ 1 & & & & \end{pmatrix},$$

则

$$N\widetilde{M}=\begin{pmatrix} & & & & 1 & 0\\ & & & \cdot & 0 & \\ & & \cdot & \cdot & & \\ & \cdot & \cdot & & & \\ 1 & \cdot & & & & \\ 0 & & & & & \end{pmatrix}.$$

当 $\lambda_0=0$, 时作 $N=(N\widetilde{M})\widetilde{M}$; 当 $\lambda_0>0$, 时作 $\sqrt{\lambda_0}M=\sqrt{\lambda_0}(M\widetilde{M})\widetilde{M}$; 当 $\lambda_0<0$, 时作

$$\sqrt{-\lambda_0}\begin{pmatrix} 0 & M\\ -M & 0\end{pmatrix}=\sqrt{-\lambda_0}\begin{pmatrix} 0 & M\widetilde{M}\\ -M\widetilde{M} & 0\end{pmatrix}\begin{pmatrix} \widetilde{M} & 0\\ 0 & \widetilde{M}\end{pmatrix};$$

当 $\lambda_0\neq\overline{\lambda_0}$ 时作

$$\begin{pmatrix} 0 & \sqrt{\lambda_0}M\\ \sqrt{\overline{\lambda_0}}M & 0\end{pmatrix}=\begin{pmatrix} 0 & \sqrt{\lambda_0}M\widetilde{M}\\ \sqrt{\overline{\lambda_0}}M\widetilde{M} & 0\end{pmatrix}\begin{pmatrix} \widetilde{M} & 0\\ 0 & \widetilde{M}\end{pmatrix}.$$

由直接验证, 便证明了定理. □

习　　题　7.4

7.4.1 试证: 两个 n 阶复方阵 A 和 B 复相似的必要且充分条件为 $2n$ 阶方阵

$$\begin{pmatrix} 0 & A\\ \overline{A} & 0\end{pmatrix},\quad\begin{pmatrix} 0 & B\\ \overline{B} & 0\end{pmatrix}$$

相似, 或者

$$\begin{pmatrix} 0 & A\\ -\overline{A} & 0\end{pmatrix},\quad\begin{pmatrix} 0 & B\\ -\overline{B} & 0\end{pmatrix}$$

相似. 由此给出两个 n 阶实方阵复相似的必要且充分条件.

7.4.2　试求一个 n 阶复方阵分解为一个 n 阶 Hermite 方阵和一个 n 阶反对称方阵的乘积的必要且充分条件.

7.4.3　试求一个 n 阶幂零复方阵 B 能表为一个 n 阶复方阵 A 和 $\overline{A}$ 的乘积 $B = A\overline{A}$ 的必要且充分条件.

7.4.4　设 n 阶复方阵 $A\overline{A}$ 的特征根全是零, 试直接算出方阵 A 在复相似下的全系不变量 $g_1, g_2, \cdots, g_s$.

7.4.5　设

$$A\overline{A} = B\overline{B} = \operatorname{diag}(N_{e_1}, N_{e_2}, \cdots, N_{e_s}), \quad e_1 \geqslant e_2 \geqslant \cdots \geqslant e_s > 0.$$

什么时候 A 和 B 复相似? 试用 $e_1, e_2, \cdots, e_s$ 来表示出 A 在复相似下的标准形

$$\operatorname{diag}(N_{g_1}, N_{g_2}, \cdots, N_{g_t})$$

中的数值 t 以及 $g_1, g_2, \cdots, g_t$.

第八章　线性函数和多重线性函数

§8.1　线性函数

在这一章, 主要考虑域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的一些简单的函数, 即考虑自变量是向量 α, 函数值是域 $\mathbb{F}$ 中元素的一些简单的函数. 这些函数和矩阵都有着密切的关系.

定义 8.1.1　给定域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的 $\mathbb{F}$ 值函数 $f(\alpha)$. $f(\alpha)$ 称为线性空间 $\mathfrak{L}$ 的**线性函数**, 如果任取 $\alpha, \beta \in \mathfrak{L}$, c, $d \in \mathbb{F}$, 则 $f(\alpha)$ 适合条件:

$$f(c\alpha + d\beta) = cf(\alpha) + df(\beta). \tag{8.1.1}$$

由定义立即可知, 对任意 s 个向量 $\xi_1, \xi_2, \cdots, \xi_s$ 和 s 个数 $a_1, a_2, \cdots, a_s \in \mathbb{F}$, 则有

$$f(\sum_{k=1}^{s} a_k\xi_k) = \sum_{k=1}^{s} a_k f(\xi_k). \tag{8.1.2}$$

现在来定出所有线性函数. 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 α_1, $\alpha_2, \cdots, \alpha_n$, 于是线性空间 $\mathfrak{L}$ 中任一向量

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i, \tag{8.1.3}$$

其中 $x_1, x_2, \cdots, x_n$ 是域 $\mathbb{F}$ 上 n 个独立自变量. 对线性空间 $\mathfrak{L}$ 的任意一个线性函数 $f(\alpha)$, 于是

$$f(\alpha) = f(\sum_{j=1}^{n} x_j\alpha_j) = \sum_{j=1}^{n} x_j f(\alpha_j) = \sum_{j=1}^{n} a_j x_j,$$

其中 $a_j = f(\alpha_j)$, $j = 1, 2, \cdots, n$. 因此, 只要知道线性函数 $f(\alpha)$ 在自变向量分别是基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 时, 函数值 $a_1 = f(\alpha_1), a_2 = f(\alpha_2), \cdots, a_n = f(\alpha_n)$, 那末, 线性函数 $f(\alpha)$ 在自变向量是 $\alpha = \sum\limits_{i=1}^{n} x_i\alpha_i$ 时, 函数值是

$$f(\alpha) = \sum_{j=1}^{n} x_j f(\alpha_j) = \sum_{j=1}^{n} a_j x_j = x'\xi_f, \tag{8.1.4}$$

其中

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \xi_f = \begin{pmatrix} f(\alpha_1) \\ f(\alpha_2) \\ \vdots \\ f(\alpha_n) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{8.1.5}$$

都是域 $\mathbb{F}$ 上 $n \times 1$ 矩阵, 即 $x, \xi_f \in \mathbb{F}^n$. 由此可见, 任给一个线性函数 $f(\alpha)$, 它就单值地对应了一个域 $\mathbb{F}$ 上 $n \times 1$ 矩阵.

反之, 任给域 $\mathbb{F}$ 上一个 $n \times 1$ 矩阵 ξ_f 如式 (8.1.5), 定义

$$f(\alpha) = f(\sum_{j=1}^{n} x_j\alpha_j) = x'\xi_f = \sum_{j=1}^{n} a_j x_j, \tag{8.1.6}$$

由定义不难验证 $f(\alpha)$ 是线性空间 $\mathfrak{L}$ 上线性函数, 它有 $a_j = f(\alpha_j)$, $j = 1, 2, \cdots, n$. 即域 $\mathbb{F}$ 上每个 $n \times 1$ 矩阵 ξ_f 单值地对应了一个线性函数 $f(\alpha)$. 所以我们证明了

引理 8.1.2 记域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的所有线性函数构成的集合是 $\mathfrak{L}^*$, 则线性空间 $\mathfrak{L}$ 的线性函数 $f(\alpha)$ 和 $n \times 1$ 矩阵 ξ_f 之间有一个到上的一一对应关系:

$$f(\alpha) \to \begin{pmatrix} f(\alpha_1) \\ f(\alpha_2) \\ \vdots \\ f(\alpha_n) \end{pmatrix} = \xi_f. \tag{8.1.7}$$

定理 8.1.3 设 f 和 g 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的线性函数, 则 $\alpha \to f(\alpha) + g(\alpha)$, $\forall\, \alpha \in \mathfrak{L}$ 定义了线性空间 $\mathfrak{L}$ 上的线性函数, 记作 $f + g$, 称为线性函数 f 以及 g 的**和**. 给定 $c \in \mathbb{F}$, 则 $\alpha \to cf(\alpha)$, $\forall\, \alpha \in \mathfrak{L}$ 定义了线性空间 $\mathfrak{L}$ 上的线性函数, 称为线性函数 f 关于纯量 c 的**纯量积**.

证 任取两线性函数 $f(\alpha)$ 和 $g(\alpha)$, 又任取常数 $a, b, c\ d \in \mathbb{F}$ 及两向量 α 和 β, 则有

$$\begin{aligned}(af + bg)(c\alpha + d\beta) &= af(c\alpha + d\beta) + bg(c\alpha + d\beta) \\ &= acf(\alpha) + adf(\beta) + bcg(\alpha) + bdg(\beta) \\ &= c(af(\alpha) + bg(\alpha)) + d(af(\beta) + bg(\beta)) \\ &= c(af + bg)(\alpha) + d(af + bg)(\beta).\end{aligned}$$

这证明了 $af+bd$ 仍是线性函数. □

由上面定理, 我们有

定义 8.1.4 记 $\mathfrak{L}^*$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的所有线性函数构成的集合, 则 $\mathfrak{L}^*$ 在上面定义的加法及纯量乘积下仍构成域 $\mathbb{F}$ 上线性空间. 称为线性空间 $\mathfrak{L}$ 的**对偶空间**, 或称为**共轭空间**.

引理 8.1.5 记 $\mathfrak{L}^*$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的对偶空间, 则式 (8.1.7) 给出的到上的一一对应关系:

$$f(\alpha)\to\xi_f \tag{8.1.8}$$

给出了对偶空间 $\mathfrak{L}^*$ 到域 $\mathbb{F}$ 上 n 维线性空间 $\mathbb{F}^n$ 的线性同构对应. 所以

$$\dim(\mathfrak{L}^*)=\dim(\mathfrak{L})=n.$$

证 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 任取两线性函数 $f(\alpha)$ 和 $g(\alpha)$, 则有

$$f(\alpha)\to\xi_f=\begin{pmatrix}f(\alpha_1)\\f(\alpha_2)\\\vdots\\f(\alpha_n)\end{pmatrix},\qquad g(\alpha)\to\xi_g=\begin{pmatrix}g(\alpha_1)\\g(\alpha_2)\\\vdots\\g(\alpha_n)\end{pmatrix}$$

又任取常数 $a,b\in\mathbb{F}$, 则有

$$(af+bg)(\alpha)\to\begin{pmatrix}(af+bg)(\alpha_1)\\(af+bg)(\alpha_2)\\\vdots\\(af+bg)(\alpha_n)\end{pmatrix}=\begin{pmatrix}af(\alpha_1)+bg(\alpha_1)\\af(\alpha_2)+bg(\alpha_2)\\\vdots\\af(\alpha_n)+bg(\alpha_n)\end{pmatrix}=a\xi_f+b\xi_g.$$

这证明了 $af+bg\to\xi_{af+bg}=a\xi_f+b\xi_g$. 即式 (8.1.8) 定义了线性空间 $\mathfrak{L}$ 到线性空间 $\mathbb{F}^n$ 上的线性同构. □

下面给定域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$: 我们来具体描述共轭空间, 即给出所有线性函数的坐标表达式.

定理 8.1.6 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 任取 $1\leqslant i\leqslant n$, 在线性空间 $\mathfrak{L}$ 上定义函数 f_i:

$$f_i(\alpha)=f(\sum_{j=1}^{n}x_j\alpha_j)=x_i, \tag{8.1.9}$$

即

$$f_i(\alpha_j)=\delta_{ij},\quad j=1,2,\cdots,n. \tag{8.1.10}$$

则 f_i 是线性空间 $\mathfrak{L}$ 上的线性函数, $i=1, 2, \cdots, n$. 又 n 个线性函数 $f_1, f_2, \cdots, f_n$ 是线性空间 $\mathfrak{L}^*$ 中一组基, 即任取 $f \in \mathfrak{L}^*$, 有

$$f=\sum_{i=1}^{n} f(\alpha_i) f_i. \tag{8.1.11}$$

证 显然, 式 (8.1.9) 定义了 n 个线性函数. 下面证它们构成线性空间 $\mathfrak{L}^*$ 中一组基. 设 $\sum\limits_{i=1}^{n} a_i f_i=0$. 任取 $1 \leqslant j \leqslant n$, 则 $\sum\limits_{i=1}^{n} a_i f_i(\alpha_j)=\sum\limits_{i=1}^{n} a_i \delta_{ij}=a_j=0$. 这证明了 $f_1, f_2, \cdots, f_n$ 线性无关. 今任取 $f \in \mathfrak{L}^*$, $\alpha=\sum\limits_{j=1}^{n} x_j \alpha_j$, 则

$$\begin{aligned}
\Big(f-\sum_{i=1}^{n} f(\alpha_i) f_i\Big)(\alpha) &= f(\alpha)-\sum_{i=1}^{n} f(\alpha_i) f_i(\alpha) \\
&= \sum_{j=1}^{n} x_j f(\alpha_j)-\sum_{i,j=1}^{n} f(\alpha_i) x_j f_i(\alpha_j) \\
&= \sum_{j=1}^{n}\Big(f(\alpha_j)-\sum_{i=1}^{n} \delta_{ij} f(\alpha_i)\Big) x_j=0 .
\end{aligned}$$

这证明了 $f=\sum\limits_{i=1}^{n} f(\alpha_i) f_i$. □

定义 8.1.7 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则对偶空间 $\mathfrak{L}^*$ 的适合条件

$$f_i(\alpha_j)=\delta_{ij}, \quad 1 \leqslant i, j \leqslant n$$

的基 $f_1, f_2, \cdots, f_n$ 称为关于线性空间 $\mathfrak{L}$ 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的**对偶基**.

上面的讨论指出: 在线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 后, 线性函数 f 对应了 $\mathbb{F}^n$ 中向量 $(f(\alpha_1), f(\alpha_2), \cdots, f(\alpha_n))'$, 这里, $(*)'$ 表示 $n \times 1$ 矩阵 $(*)$ 的转置矩阵. 如果在线性空间 $\mathfrak{L}$ 中另外再取定一组基 $\beta_1, \beta_2, \cdots, \beta_n$, 则同一线性函数 f 对应的向量是 $(f(\beta_1), f(\beta_2), \cdots, f(\beta_n))'$. 下面来求这两个向量间的关系. 即要证明

定理 8.1.8 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$, 假设基变换所对应的非异方阵是 P. 且在对偶空间 $\mathfrak{L}^*$ 中对应的对偶基分别是 $f_1, f_2, \cdots, f_n$ 和 $g_1, g_2, \cdots, g_n$. 则对偶基的基变换所对应的非异方阵便是 $(P')^{-1}$.

证 令 $f_i(\alpha_j)=g_i(\beta_j)=\delta_{ij}$, $1 \leqslant i, j \leqslant n$, 又

$$\beta_i=\sum_{j=1}^{n} p_{ji} \alpha_j, \quad i=1,2, \cdots, n,$$

其中 $P=(p_{ij})$ 是线性空间 $\mathfrak{L}$ 中基变换所对应的非异方阵. 记

$$g_i=\sum_{k=1}^{n}q_{ki}f_k,\quad i=1,2,\cdots,n,$$

其中 $Q=(q_{ij})$ 是对偶空间 $\mathfrak{L}^*$ 中基变换所对应的非异方阵. 而

$$\delta_{ik}=g_i(\beta_j)=\sum_{k=1}^{n}q_{ki}f_k(\sum_{m=1}^{n}p_{mj}\alpha_m)=\sum_{k,m=1}^{n}q_{ki}p_{mj}\delta_{km}=\sum_{k=1}^{n}q_{ki}p_{kj}.$$

这证明了 n 阶方阵 $Q=(q_{ij})$ 有 $Q'P=E_n$, 所以 $Q=(P^{-1})'$. □

另一方面, 对域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任一线性变换 $\mathscr{A}$ 及任一线性函数 $f(\alpha)$, 则 $f(\mathscr{A}(\alpha))$ 也是线性函数, 记作 $\mathscr{A}^*(f)$. 于是, 引进了对偶空间 $\mathfrak{L}^*$ 的变换 $f\to\mathscr{A}^*(f)$, 它记作 $\mathscr{A}^*$.

定理 8.1.9 设 $\mathscr{A}$ 是 n 维线性空间 $\mathfrak{L}$ 的线性变换. 任取 $\mathfrak{L}$ 的线性函数 f, 则

$$\alpha\to f(\mathscr{A}(\alpha)),\quad \forall\,\alpha\in\mathfrak{L}\tag{8.1.12}$$

也是 $\mathfrak{L}$ 的线性函数, 记作 $\mathscr{A}^*(f)$, 即有

$$(\mathscr{A}^*(f))(\alpha)=f(\mathscr{A}(\alpha)),\quad \forall\,\alpha\in\mathfrak{L}.\tag{8.1.13}$$

于是引进了线性空间 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 的一个映射 $\mathscr{A}^*$: $f\to\mathscr{A}^*(f)$, $\forall\,f\in\mathfrak{L}^*$. 则 $\mathscr{A}^*$ 是对偶空间 $\mathfrak{L}^*$ 的线性变换, 称为线性变换 $\mathscr{A}$ 的**对偶变换**或**共轭变换**. 在线性空间 $\mathfrak{L}$ 中任取一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 记线性变换 $\mathscr{A}$ 对应的方阵表示是 $A=(a_{ij})$. 再在对偶空间 $\mathfrak{L}^*$ 中取对偶基 $f_1,f_2,\cdots,f_n$, 则共轭变换 $\mathscr{A}^*$ 在对偶基 $f_1,f_2,\cdots,f_n$ 下对应的方阵表示是 A', 即方阵 A 的转置方阵.

证 显然, 由线性变换 $\mathscr{A}$ 以及线性函数的单值性和线性性, 所以 $f\to\mathscr{A}^*(f)$, $\forall\,f\in\mathfrak{L}^*$ 是单值映射. 再证 $\mathscr{A}^*$ 是线性变换. 今对域 $\mathbb{F}$ 中任两数 a 和 b 及任两线性函数 f 和 g, 有

$$\begin{aligned}(\mathscr{A}^*(af+bg))(\alpha)&=(af+bg)(\mathscr{A}(\alpha))=af(\mathscr{A}(\alpha))+bg(\mathscr{A}(\alpha))\\&=a(\mathscr{A}^*f)(\alpha)+b(\mathscr{A}^*g)(\alpha)=(a(\mathscr{A}^*f)+b(\mathscr{A}^*g))(\alpha),\end{aligned}$$

对一切向量 $\alpha\in\mathfrak{L}$ 成立. 所以 $\mathscr{A}^*(af+bg)=a(\mathscr{A}^*f)+b(\mathscr{A}^*g)$, 即 $\mathscr{A}^*$ 是线性变换.

最后, 我们来求共轭变换 $\mathscr{A}^*$ 在对偶空间 $\mathfrak{L}^*$ 的对偶基 $f_1,f_2,\cdots,f_n$ 下对应的方阵表示. 由假设, 设 $\mathscr{A}(\alpha_i)=\sum\limits_{k=1}^{n}a_{ki}\alpha_k$, $\mathscr{A}^*(f_i)=\sum\limits_{p=1}^{n}b_{pi}f_p$, $i=1,2,\cdots,n$, 其中 $B=(b_{ij})$ 是共轭变换 $\mathscr{A}^*$ 在对偶基 $f_1,f_2,\cdots,f_n$ 下对应的方阵表示. 今

$$\mathscr{A}^*(f_i)(\alpha_k)=f_i(\mathscr{A}(\alpha_k))=f_i(\sum_{p=1}^{n}a_{pk}\alpha_p)=\sum_{p=1}^{n}a_{pk}f_i(\alpha_p)=\sum_{p=1}^{n}a_{pk}\delta_{ip}=a_{ik},$$

又

$$\mathscr{A}^*(f_i)(\alpha_k)=\sum_{p=1}^{n}b_{pi}f_p(\alpha_k)=\sum_{p=1}^{n}b_{pi}\delta_{kp}=b_{ki}.$$

所以, $b_{ki}=a_{ik}$, 其中 $i,k=1,2,\cdots,n$. 这证明了 $B=A'$. □

§8.2 多重线性函数 *

从这一节开始, 我们引进在近代数学中十分有用的多重线性函数和张量的概念. 作为向量概念的推广, 张量的概念已经成为一个很基本的概念了. Einstein 的相对论就是以张量作为最基本的语言的.

在 §8.2 和 §8.3, 我们在给定的域 $\mathbb{F}$ 上讨论.

引理 8.2.1 设 $\mathfrak{L}$ 是 n 维线性空间, $\mathfrak{L}^*$ 是 $\mathfrak{L}$ 的对偶空间, $(\mathfrak{L}^*)^*$ 是 $\mathfrak{L}^*$ 的对偶空间, 则我们在

$$\alpha(f)=f(\alpha)\in\mathbb{F},\quad \forall\,\alpha\in\mathfrak{L},f\in\mathfrak{L}^* \tag{8.2.1}$$

的意义下可以理解为 $(\mathfrak{L}^*)^*=\mathfrak{L}$.

证 在线性空间 $\mathfrak{L}$ 中取基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 在对偶空间 $\mathfrak{L}^*$ 中取基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基 $f_1,f_2,\cdots,f_n$. 在对偶空间 $\mathfrak{L}^*$ 的对偶空间 $(\mathfrak{L}^*)^*$ 中取基 $f_1,f_2,\cdots,f_n$ 的对偶基 $F_1,F_2,\cdots,F_n$. 于是有 $\dim(\mathfrak{L})=\dim(\mathfrak{L}^*)=\dim((\mathfrak{L}^*)^*)$, 又

$$f_i(\alpha_j)=\delta_{ij},\quad F_i(f_j)=\delta_{ij},\quad 1\leqslant i,j\leqslant n.$$

作对应

$$\alpha_j\to F_j,\quad 1\leqslant j\leqslant n.$$

将它拓展为线性空间 $\mathfrak{L}$ 到线性空间 $(\mathfrak{L}^*)^*$ 内的对应

$$\alpha=\sum_{i=1}^{n}a_i\alpha_i\to\sum_{i=1}^{n}a_iF_i=F,\quad \forall\,a_1,a_2,\cdots,a_n\in\mathbb{F}.$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性空间 $\mathfrak{L}$ 的基, $F_1,F_2,\cdots,F_n$ 是线性空间 $(\mathfrak{L}^*)^*$ 的基. 所以上述对应是线性空间 $\mathfrak{L}$ 到线性空间 $(\mathfrak{L}^*)^*$ 上的线性同构. 今任取 $f=\sum\limits_{i=1}^{n}x_if_i\in\mathfrak{L}^*$, 则

$$F(f)=\sum_{i=1}^{n}a_iF_i(\sum_{j=1}^{n}x_jf_j)=\sum_{i,j=1}^{n}a_ix_jF_i(f_j)=\sum_{i,j=1}^{n}a_ix_j\delta_{ij}=\sum_{i=1}^{n}a_ix_i,$$

又

$$f(\alpha)=\sum_{i=1}^{n}x_if_i(\sum_{j=1}^{n}a_j\alpha_j)=\sum_{i,j=1}^{n}x_ia_jf_i(\alpha_j)=\sum_{i,j=1}^{n}x_ia_j\delta_{ij}=\sum_{i=1}^{n}a_ix_i.$$

于是有

$$F(f)=f(\alpha), \quad \forall f\in\mathfrak{L}^*.$$

由于 $\alpha\to F$ 给出 $\mathfrak{L}$ 到 $(\mathfrak{L}^*)^*$ 上的线性同构, 所以我们改记 F 为 α, 则有

$$\alpha(f)=f(\alpha), \quad \forall f\in\mathfrak{L}^*.$$

在这意义下, $(\mathfrak{L}^*)^*=\mathfrak{L}$. □

定义 8.2.2 设 $\mathfrak{L}$ 是域 $\mathbb{F}$ 上 n 维线性空间. 给定 s 个互相独立的变向量 ξ_1, ξ_2, $\cdots$, ξ_s, t 个互相独立的变线性函数 g_1, g_2, $\cdots$, g_t. 设这 $s+t$ 个变元的函数

$$f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t) \tag{8.2.2}$$

适合条件: 在其中任意固定 $s+t-1$ 个变元, 则 f 是余下的一个变元的线性函数. 这时 f 称为**多重线性函数**, 又称为 **(s,t) 型张量**, 即 s **阶共变**, t **阶反变张量**. 它们全体构成集合 $\mathfrak{D}_t^s$. 特别, 集合 $\mathfrak{D}_0^0$ 约定是所有纯量构成的集合 $\mathbb{F}$.

显然, 在集合 $\mathfrak{D}_t^s$ 中可以引进加法和纯量积如下:

(1) 加法 $\forall f,g\in\mathfrak{D}_t^s$,

$$(f+g)(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=f(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)+g(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t); \tag{8.2.3}$$

(2) 纯量积

$$(af)(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t)=af(\xi_1,\cdots,\xi_s;g_1,\cdots,g_t), \quad \forall a\in\mathbb{F}, \quad f\in\mathfrak{D}_t^s. \tag{8.2.4}$$

由上面的运算定义, 不难验证 $\mathfrak{D}_t^s$ 是域 $\mathbb{F}$ 上线性空间.

引理 8.2.3 线性空间 $\mathfrak{D}_0^0$ 有基 1, 它是一维线性空间. 线性空间 $\mathfrak{D}_1^0=\mathfrak{L}$, 它有基 α_1, α_2, $\cdots$, α_n. 线性空间 $\mathfrak{D}_0^1=\mathfrak{L}^*$, 它有基 f_1, f_2, $\cdots$, f_n. 线性空间 $\mathfrak{D}_t^s$ 是 n^{s+t} 维线性空间, 它有基 $f_{i_1,\cdots,i_s}^{j_1,\cdots,j_t}$, $1\leqslant i_1$, $\cdots$, i_s, j_1, $\cdots$, $j_t\leqslant n$. 基元素作为 (s,t) 型张量, 定义为

$$\begin{aligned} &f_{i_1,\cdots,i_s}^{j_1,\cdots,j_t}\Big(\sum_{k_1=1}^n x_{1k_1}\alpha_{k_1},\cdots,\sum_{k_s=1}^n x_{sk_s}\alpha_{k_s};\sum_{p_1=1}^n y_{1p_1}f_{p_1},\cdots,\sum_{p_t=1}^n y_{tp_t}f_{p_t}\Big) \\ &=x_{1i_1}\cdots x_{si_s}y_{1j_1}\cdots y_{tj_t}, \end{aligned} \tag{8.2.5}$$

其中 α_1, α_2, $\cdots$, α_n 是线性空间 $\mathfrak{L}$ 的基, f_1, f_2, $\cdots$, f_n 是对偶空间 $\mathfrak{L}^*$ 中 α_1, $\cdots$, α_n 的对偶基.

证 记 s 个变向量 $\xi_i=\sum\limits_{j=1}^n x_{ij}\alpha_j$, $1\leqslant i\leqslant s$, t 个变线性函数 $g_k=\sum\limits_{j=1}^n y_{kj}f_j$,

$1 \leqslant k \leqslant t$. 任取 $f \in \mathfrak{D}_t^s$, 则

$$\begin{aligned}&f(\xi_1, \cdots, \xi_s; g_1, \cdots, g_t)\\ =& \sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} x_{1i_1} \cdots x_{si_s} y_{1j_1} \cdots y_{tj_t} f(\alpha_{i_1}, \cdots, \alpha_{i_s}; f_{j_1}, \cdots, f_{j_t}),\end{aligned}$$

其中

$$f(\alpha_{i_1}, \cdots, \alpha_{i_s}; f_{j_1}, \cdots, f_{j_t}) = a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} \in \mathbb{F}.$$

利用式 (8.2.5) 定义的多重线性函数 $f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}$, 则有

$$f(\xi_1, \cdots, \xi_s; g_1, \cdots, g_t) = \sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}(\xi_1, \cdots, \xi_s; g_1, \cdots, g_t).$$

这证明了

$$f = \sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}.$$

所以 $\mathfrak{D}_t^s$ 由 $f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}$, $1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$ 线性生成. 余下证明它们线性无关.

今 $a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} \in \mathbb{F}$, $1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$ 有

$$\sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} = 0.$$

任取 $1 \leqslant k_1, \cdots, k_s, p_1, \cdots, p_t \leqslant n$, 则

$$f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}(\alpha_{k_1}, \cdots, \alpha_{k_s}; f_{p_1}, \cdots, f_{p_t}) = \delta_{i_1k_1} \cdots \delta_{i_sk_s} \delta_{j_1p_1} \cdots \delta_{j_tp_t}.$$

于是有

$$0 = \sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} a_{i_1, \cdots, i_s}^{j_1, \cdots, j_t} \delta_{i_1k_1} \cdots \delta_{i_sk_s} \delta_{j_1p_1} \cdots \delta_{j_tp_t} = a_{k_1, \cdots, k_s}^{p_1, \cdots, p_t}.$$

这证明了 $a_{k_1, \cdots, k_s}^{p_1, \cdots, p_t} = 0$, $1 \leqslant k_1, \cdots, k_s, p_1, \cdots, p_t \leqslant n$. 即 $f_{i_1, \cdots, i_s}^{j_1, \cdots, j_t}$, $1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$ 是 $\mathfrak{D}_t^s$ 的一组基. □

(s,t) 型张量 f 可以看作 n^{s+t} 个独立自变量 x_{ij}, $1 \leqslant i \leqslant s$, $1 \leqslant j \leqslant n$ 及 y_{ij}, $1 \leqslant i \leqslant t$, $1 \leqslant j \leqslant n$ 的函数. 我们称 n^{s+t} 个数

$$a_{i_1 i_2 \cdots i_s}^{j_1 j_2 \cdots j_t}, \quad 1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$$

为 (s,t) 型张量 f 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的**坐标**. 显然, 它唯一决定了 (s,t) 型张量 f. 反之, 任给 n^{s+t} 个数 $a_{i_1 i_2 \cdots i_s}^{j_1 j_2 \cdots j_t}$, $1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$, 考虑函数

$$\sum_{i_1, \cdots, i_s=1}^{n} \sum_{j_1, \cdots, j_t=1}^{n} a_{i_1 i_2 \cdots i_s}^{j_1 j_2 \cdots j_t} x_{1i_1} x_{2i_2} \cdots x_{si_s} y_{1j_1} y_{2j_2} \cdots y_{tj_t}. \tag{8.2.6}$$

在线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 后, 它定义了线性空间 $\mathfrak{L}$ 上一个 (s,t) 型张量. 所以 $\mathfrak{D}_t^s$ 的坐标表示为 $\mathbb{F}^{s+t}$.

作形式直接和

$$\mathfrak{D} = \bigoplus_{s,t=0}^{\infty} \mathfrak{D}_t^s = {\sum_{s,t=0}^{\infty}}' \mathfrak{D}_t^s. \tag{8.2.7}$$

这意味着线性空间 $\mathfrak{D}$ 中任一元 F 可唯一地表为

$$F = \sum_{s,t=0}^{\infty} F_t^s, \tag{8.2.8}$$

其中 $F_t^s \in \mathfrak{D}_t^s$. 注意, 所谓形式和, 指不考虑收敛性, 也不考虑 $F_t^s + F_v^u$ 是如何加起来的. 显然, 域 $\mathbb{F}$ 上线性空间 $\mathfrak{D}$ 为无限维线性空间, 其中加法及纯量积分别定义为

$$\sum_{s,t=0}^{\infty} F_t^s + \sum_{s,t=0}^{\infty} G_t^s = \sum_{s,t=0}^{\infty} (F_t^s + G_t^s), \quad a\sum_{s,t=0}^{\infty} F_t^s = \sum_{s,t=0}^{\infty} (aF_t^s), \tag{8.2.9}$$

其中 $a \in \mathbb{F}$.

重要的是引进张量积 $\otimes$ 的定义.

$$(\sum_{s,t=0}^{\infty} F_t^s) \otimes (\sum_{p,q=0}^{\infty} G_q^p) = \sum_{s,t,p,q=0}^{\infty} (F_t^s \otimes G_q^p) = \sum_{u,v=0}^{\infty} H_v^u, \tag{8.2.10}$$

其中

$$H_v^u = \sum_{s+p=u} \sum_{t+q=v} F_t^s \otimes G_q^p \in \mathfrak{D}_v^u. \tag{8.2.11}$$

而 $F_t^s \otimes G_q^p$ 定义为

$$\begin{aligned}&(F_t^s \otimes G_q^p)(\xi_1, \cdots, \xi_s, \xi_{s+1}, \cdots, \xi_{s+p}; g_1, \cdots, g_t, g_{t+1}, \cdots, g_{t+q}) \\ &= F_t^s(\xi_1, \cdots, \xi_s; g_1, \cdots, g_t) G_q^p(\xi_{s+1}, \cdots, \xi_{s+p}; g_{t+1}, \cdots, g_{t+q}).\end{aligned} \tag{8.2.12}$$

显然由式 (8.2.10)—(8.2.12) 给出的**张量积**定义保证了 $F_t^s \otimes G_q^p$ 仍为多重线性函数, 即为 $s+p$ 阶共变, $t+q$ 阶反变张量. 因此 $H_v^u \in \mathfrak{D}_v^u$, 且上面给出的线性空间 $\mathfrak{D}$ 中张量的和, 纯量积和张量积适合加法结合律和交换律, 乘法结合律, 但不适合乘法交换律, 又适合加、乘分配律. 且有单位元素 $1 \in \mathfrak{D}_0^0 = \mathbb{F}$ 和零元素 $0 \in \mathfrak{D}_0^0 = \mathbb{F}$.

定义 8.2.4 设 $\mathfrak{L}$ 是域 $\mathbb{F}$ 上线性空间, $\mathfrak{L}$ 称为**结合代数**, 如果在 $\mathfrak{L}$ 中还可定义乘法, 即 $\forall\, \alpha, \beta \in \mathfrak{L}$, 定义 $\alpha\beta \in \mathfrak{L}$. 设它适合下面条件

(1) 乘法结合律

$$\alpha(\beta\gamma) = (\alpha\beta)\gamma, \quad \forall\, \alpha, \beta, \gamma \in \mathfrak{L}; \tag{8.2.13}$$

(2) 加乘分配律

$$\alpha(\beta+\gamma) = \alpha\beta + \alpha\gamma, \quad (\beta+\gamma)\alpha = \beta\alpha + \gamma\alpha, \quad \forall\, \alpha, \beta, \gamma \in \mathfrak{L}; \tag{8.2.14}$$

(3) 纯量积和乘法间有关系

$$a(\alpha\beta) = (a\alpha)\beta = \alpha(a\beta), \quad \forall\, a \in \mathbb{F}, \quad \forall\, \alpha, \beta \in \mathfrak{L}. \tag{8.2.15}$$

$\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$, 如果任取 α, $\beta \in \mathfrak{L}_1$, 则 $\alpha\beta \in \mathfrak{L}_1$, 这时 $\mathfrak{L}_1$ 称为结合代数 $\mathfrak{L}$ 的**子代数**, 它仍为结合代数.

结合代数在代数学中是一种重要的 "代数". 例如域 $\mathbb{F}$ 上所有 n 阶方阵在矩阵的加法, 纯量乘积和乘法下构成 n^2 维结合代数. 关于结合代数的研究属于抽象代数这一学科. 我们引进这个概念, 是为了给出

定义 8.2.5 设 $\mathfrak{L}$ 是域 $\mathbb{F}$ 上线性空间, 式 (8.2.7) 定义的无限维线性空间 $\mathfrak{D}$ 在加法, 纯量积及张量积下构成结合代数, 称为域 $\mathbb{F}$ 上**张量代数**.

张量代数是近代微分几何的重要工具. 它首先是由 Einstein 在研究相对论时, 应用到物理学上去的. 从而也推动了近代微分几何的发展. 但是 Einstein 是用张量的坐标来表示张量, 而坐标和线性空间 $\mathfrak{L}$ 上的基底有关, 为此需要给出在不同基底下坐标间的关系.

利用张量积的定义, 我们有

定理 8.2.6 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是域 $\mathbb{F}$ 上线性空间 $\mathfrak{L}$ 的一组基, $f_1, f_2, \cdots, f_n$ 是线性空间 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 中的对偶基. 则式 (8.2.5) 给出的线性空间 $\mathfrak{D}_t^s$ 的基

$$f_{i_1,\cdots,i_s}^{j_1,\cdots,j_t} = f_{i_1} \otimes \cdots \otimes f_{i_s} \otimes \alpha_{j_1} \otimes \cdots \otimes \alpha_{j_t}, \quad 1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n. \tag{8.2.16}$$

由 $\mathfrak{D}_0^1 = \mathfrak{L}$ 以及 $\mathfrak{D}_1^0 = \mathfrak{L}^*$, 因此

$$\mathfrak{D}_t^s = (\mathfrak{L}^*)^s \otimes (\mathfrak{L})^t, \tag{8.2.17}$$

其中 $(\mathfrak{L}^*)^s = \mathfrak{L}^* \otimes \cdots \otimes \mathfrak{L}^*$ 是 s 个 $\mathfrak{L}^*$ 的张量积, $(\mathfrak{L})^t = \mathfrak{L} \otimes \cdots \otimes \mathfrak{L}$ 是 t 个 $\mathfrak{L}$ 的张量积. 这里我们约定线性空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的张量积 $\mathfrak{L}_1 \otimes \mathfrak{L}_2$ 是由集合

$$\{\, \xi_1 \otimes \xi_2 \mid \forall \xi_1 \in \mathfrak{L}_1, \quad \xi_2 \in \mathfrak{L}_2 \,\} \tag{8.2.18}$$

线性生成的线性空间.

证 由定义可知 $\mathfrak{D}_0^0 = \mathbb{F}$, $\mathfrak{D}_0^1 = \mathfrak{L}$, $\mathfrak{D}_1^0 = \mathfrak{L}^*$. 由引理 8.2.1, 下面来证式 (8.2.16)

成立. 今任取 $1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n$, 则

$$\begin{aligned}
&(f_{i_1} \otimes \cdots \otimes f_{i_s} \otimes \alpha_{j_1} \otimes \cdots \otimes \alpha_{j_t})\\
&(\sum_{p_1=1}^{n} x_{1p_1}\alpha_{p_1}, \cdots, \sum_{p_s=1}^{n} x_{sp_s}\alpha_{p_s}; \sum_{q_1=1}^{n} y_{1q_1} f_{q_1}, \cdots, \sum_{q_t=1}^{n} y_{tq_t} f_{q_t})\\
=& \sum_{p_1=1}^{n} \delta_{i_1p_1} x_{1p_1} \cdots \sum_{p_s=1}^{n} \delta_{i_sp_s} x_{sp_s} \sum_{q_1=1}^{n} \delta_{j_1q_1} y_{1q_1} \cdots \sum_{q_t=1}^{n} \delta_{j_tq_t} y_{tq_t}\\
=& x_{1i_1} \cdots x_{si_s} y_{1j_1} \cdots y_{tj_t}\\
=& f_{i_1,\cdots,i_s}^{j_1,\cdots,j_t}(\sum_{p_1=1}^{n} x_{1p_1}\alpha_{p_1}, \cdots, \sum_{p_s=1}^{n} x_{sp_s}\alpha_{p_s}; \sum_{q_1=1}^{n} y_{1q_1} f_{q_1}, \cdots, \sum_{q_t=1}^{n} y_{tq_t} f_{q_t}).
\end{aligned}$$

因此证明了式 (8.2.16) 成立. □

定理 8.2.7 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$, 假设基变换所对应的非异方阵是 P, 且在对偶空间 $\mathfrak{L}^*$ 中对应的对偶基分别是 $f_1, f_2, \cdots, f_n$ 和 $g_1, g_2, \cdots, g_n$, 对偶基的基变换所对应的非异方阵便是 $Q = (P')^{-1}$. 任取 $f \in \mathfrak{D}_t^s$, 设 f 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 下对应的坐标分别是

$$a_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}, \quad 1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n \tag{8.2.19}$$

和

$$b_{i_1i_2\cdots i_s}^{j_1j_2\cdots j_t}, \quad 1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n. \tag{8.2.20}$$

则它们之间有关系

$$a_{k_1k_2\cdots k_s}^{p_1p_2\cdots p_t} = \sum_{i_1,\cdots,i_s=1}^{n} \sum_{j_1,\cdots,j_t=1}^{n} b_{i_1\cdots i_s}^{j_1\cdots j_t} q_{k_1i_1} q_{k_2i_2} \cdots q_{k_si_s} p_{p_1j_1} p_{p_2j_2} \cdots p_{p_tj_t}, \tag{8.2.21}$$

其中 $1 \leqslant k_1, \cdots, k_s, p_1, \cdots, p_t \leqslant n$, $P = (p_{ij})$, $Q = (q_{ij}) = (P^{-1})'$.

证 由定理 8.1.8, 有 $P = (p_{ij})$, $Q = (q_{ij}) = (P^{-1})'$, 即

$$\beta_i = \sum_{j=1}^{n} p_{ji}\alpha_j, \quad g_i = \sum_{j=1}^{n} q_{ji} f_j, \quad 1 \leqslant i \leqslant n. \tag{8.2.22}$$

另一方面,

$$f = \sum_{k_1,\cdots,k_s=1}^{n} \sum_{p_1,\cdots,p_t=1}^{n} a_{k_1k_2\cdots k_s}^{p_1p_2\cdots p_t} f_{k_1} \otimes \cdots \otimes f_{k_s} \otimes \alpha_{p_1} \otimes \cdots \otimes \alpha_{p_t},$$

又

$$f = \sum_{i_1,\cdots,i_s=1}^{n} \sum_{j_1,\cdots,j_t=1}^{n} b_{i_1\cdots i_s}^{j_1\cdots j_t} g_{i_1} \otimes \cdots \otimes g_{i_s} \otimes \beta_{j_1} \otimes \cdots \otimes \beta_{j_t}.$$

由式 (8.2.22), 所以式 (8.2.21) 成立. □

下面考虑共变张量, 它们全体构成张量代数 $\mathfrak{D}$ 的子代数, 记作 $\mathfrak{D}_0$. 由定义可知

$$\mathfrak{D}_0=\sum_{s=0}^{\infty}\mathfrak{D}_0^s=\sum_{s=0}^{\infty}(\mathfrak{D}_0^1)^s=\sum_{s=0}^{\infty}(\mathfrak{L}^*)^s. \tag{8.2.23}$$

定义 8.2.8 n 维线性空间 $\mathfrak{L}$ 上 s 阶共变张量 f 称为**对称共变张量**, 如果对 1, 2, $\cdots$, s 的任一排列 $i_1i_2\cdots i_s$, 有

$$f(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_s})=f(\xi_1,\xi_2,\cdots,\xi_s). \tag{8.2.24}$$

s 阶共变张量 f 称为**反对称共变张量**, 如果

$$f(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_s})=\delta^{1\,2\,\cdots\,s}_{i_1i_2\cdots i_s}f(\xi_1,\xi_2,\cdots,\xi_s). \tag{8.2.25}$$

对称共变张量的结构简单. 反对称共变张量的结构复杂, 但很重要. 反对称共变张量的讨论为下一节的内容.

定理 8.2.9 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 α_1, α_2, $\cdots$, α_n. 设 f_1, f_2, $\cdots$, f_n 是 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 关于基 α_1, α_2, $\cdots$, α_n 的对偶基. 则 $\mathfrak{L}$ 的任一 s 阶对称共变张量

$$f=\sum_{1\leqslant i_1\leqslant\cdots\leqslant i_s\leqslant n}a_{i_1i_2\cdots i_s}\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s}. \tag{8.2.26}$$

因此 s 阶对称共变张量全体构成以

$$\sum_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s},\qquad 1\leqslant i_1\leqslant\cdots\leqslant i_s\leqslant n \tag{8.2.27}$$

为基的子空间, 这里求和号定义为: 当 $i_1<i_2<\cdots<i_s$ 时, $j_1j_2\cdots j_s$ 是 i_1, i_2, $\cdots$, i_s 的排列; 当 $i_1\leqslant i_2\leqslant\cdots\leqslant i_s$ 中有指标相同时, 我们仍看作不同对象, 在此意义下 $j_1j_2\cdots j_s$ 是 i_1, i_2, $\cdots$, i_s 的排列. 所以和号总共 $s!$ 项.

证 任取 s 阶对称共变张量 f, 则 $f=\sum\limits_{i_1,\cdots,i_s=1}^{n}a_{i_1\cdots i_s}f_{i_1}\otimes\cdots\otimes f_{i_s}$, 其中 $a_{i_1\cdots i_s}=f(\alpha_{i_1},\cdots,\alpha_{i_s})$. 视 i_1, i_2, $\cdots$, i_s 是 s 个不同符号. 取它的任意排列 $j_1j_2\cdots j_s$, 则有 $a_{j_1j_2\cdots j_s}=a_{i_1i_2\cdots i_s}$. 于是式 (8.2.26) 成立.

反之, 显然 $\sum\limits_{\binom{i_1i_2\cdots i_s}{j_1j_2\cdots j_s}}f_{j_1}\otimes\cdots\otimes f_{j_s}$ 是 s 阶对称共变张量. 又对所有 $1\leqslant i_1\leqslant i_2\leqslant\cdots\leqslant i_s\leqslant n$, 它们全体线性无关. 所以它们构成一组基. □

§8.3 Grassman 代数 *

定理 8.3.1 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 设 $f_1, f_2, \cdots, f_n$ 是 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的对偶基. 则 $\mathfrak{L}$ 中任一反对称共变张量

$$f = \sum_{1 \leqslant i_1 < \cdots < i_s \leqslant n} a_{i_1 i_2 \cdots i_s} \sum_{\binom{i_1 i_2 \cdots i_s}{j_1 j_2 \cdots j_s}} \delta^{i_1 i_2 \cdots i_s}_{j_1 j_2 \cdots j_s} f_{j_1} \otimes f_{j_2} \otimes \cdots \otimes f_{j_s}. \tag{8.3.1}$$

所以所有 s 阶反对称共变张量构成 $\binom{n}{s}$ 维线性空间 Λ^s, 它有基

$$F_{i_1 i_2 \cdots i_s} = \frac{1}{s!} \sum_{\binom{i_1 i_2 \cdots i_s}{j_1 j_2 \cdots j_s}} \delta^{i_1 i_2 \cdots i_s}_{j_1 j_2 \cdots j_s} f_{j_1} \otimes f_{j_2} \otimes \cdots \otimes f_{j_s}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_s \leqslant n. \tag{8.3.2}$$

特别 $\Lambda^0 = \mathfrak{D}^0_0 = \mathbb{F}$, $\Lambda^1 = \mathfrak{L}^*$, $\Lambda^{n+1} = \Lambda^{n+2} = \cdots = 0$. 因此张量代数 $\mathfrak{D}$ 中所有反对称共变张量构成 2^n 维线性空间

$$\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \Lambda^2 \oplus \cdots \oplus \Lambda^n. \tag{8.3.3}$$

证 任取 s 阶反对称共变张量 $f = \sum\limits_{i_1, \cdots, i_s = 1}^{n} a_{i_1 i_2 \cdots i_s} f_{i_1} \otimes f_{i_2} \otimes f_{i_s}$, 其中 $a_{i_1 i_2 \cdots i_s} = f(\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_s})$. 我们先证当存在 $1 \leqslant j < k \leqslant n$, $i_j = i_k$ 时, 有 $a_{i_1 i_2 \cdots i_s} = 0$. 这是因为

$$\begin{aligned} & f(\alpha_{i_1}, \cdots, \alpha_{i_{j-1}}, \alpha_{i_j}, \alpha_{i_{j+1}}, \cdots, \alpha_{i_{k-1}}, \alpha_{i_k}, \alpha_{i_{k+1}}, \cdots, \alpha_{i_s}) \\ = & -f(\alpha_{i_1}, \cdots, \alpha_{i_{j-1}}, \alpha_{i_k}, \alpha_{i_{j+1}}, \cdots, \alpha_{i_{k-1}}, \alpha_{i_j}, \alpha_{i_{k+1}}, \cdots, \alpha_{i_s}), \end{aligned}$$

所以由 $i_j = i_k$ 有 $a_{i_1 i_2 \cdots i_s} = 0$. 因此也证明了

$$f = \sum_{1 \leqslant i_1 < \cdots < i_s \leqslant n} \sum_{\binom{i_1 i_2 \cdots i_s}{j_1 j_2 \cdots j_s}} a_{j_1 j_2 \cdots j_s} f_{j_1} \otimes f_{j_2} \otimes \cdots \otimes f_{j_s}.$$

今

$$a_{j_1 j_2 \cdots j_s} = f(\alpha_{j_1}, \cdots, \alpha_{j_s}) = \delta^{i_1 i_2 \cdots i_s}_{j_1 j_2 \cdots j_s} f(\alpha_{i_1}, \cdots, \alpha_{i_s}) = \delta^{i_1 i_2 \cdots i_s}_{j_1 j_2 \cdots j_s} a_{i_1 i_2 \cdots i_s}.$$

所以式 (8.3.2) 给出的 $F_{i_1 i_2 \cdots i_s}$, $1 \leqslant i_1 < \cdots < i_s \leqslant n$ 线性生成子空间 Λ^s. 由指标 $1 \leqslant i_1 < \cdots < i_s \leqslant n$ 可知它们线性无关. 因此证明了式 (8.3.2) 给出了子空间 Λ^s 的一组基, 它有 $\binom{n}{s}$ 个基元素, 所以 $\dim(\Lambda^s) = \binom{n}{s}$. 因此, $\dim(\Lambda) = \sum\limits_{k=0}^{n} \dim(\Lambda^k) = \sum\limits_{k=0}^{n} \binom{n}{k} = 2^n$. □

利用反对称共变张量, 我们可以给出行列式的另一种定义.

定义 8.3.2 域 $\mathbb{F}$ 上线性空间 $\mathbb{F}^n$ 上 n 阶反对称共变张量 $n! F_{12 \cdots n}(\xi_1, \xi_2, \cdots, \xi_n)$

称为由 n 个列向量 $\xi_1, \xi_2, \cdots, \xi_n$ 决定的**行列式**.

为了给出上面的定义和第二章中行列式的定义一致. 我们给出更一般的结果.

定理 8.3.3 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 记 $f_1, f_2, \cdots, f_n$ 是线性空间 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 关于基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的对偶基. 记变向量

$$\xi_i = \sum_{j=1}^{n} x_{ij}\alpha_j, \quad 1 \leqslant i \leqslant s, \tag{8.3.4}$$

则 $\mathfrak{L}$ 上任一 s 阶反对称共变张量 $f = \sum\limits_{1\leqslant i_1<\cdots<i_s\leqslant n} a_{i_1 i_2\cdots i_s} F_{i_1 i_2\cdots i_s}$, 有

$$F_{i_1 i_2\cdots i_s}(\xi_1, \cdots, \xi_s) = \frac{1}{s!}\det\begin{pmatrix} x_{1i_1} & \cdots & x_{1i_s} \\ \vdots & & \vdots \\ x_{si_1} & \cdots & x_{si_s} \end{pmatrix}, \quad 1 \leqslant i_1 < \cdots < i_s \leqslant n. \tag{8.3.5}$$

特别

$$F_{12\cdots n}(\xi_1, \cdots, \xi_n) = \frac{1}{n!}\det\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}. \tag{8.3.6}$$

证 由式 (8.3.2), 有

$$\begin{aligned} F_{i_1 i_2\cdots i_s}(\xi_1, \cdots, \xi_s) &= \frac{1}{s!}\sum_{\begin{pmatrix} i_1 i_2 \cdots i_s \\ j_1 j_2 \cdots j_s \end{pmatrix}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s}(f_{j_1}\otimes f_{j_2}\otimes\cdots\otimes f_{j_s})(\xi_1, \xi_2, \cdots, \xi_s) \\ &= \frac{1}{s!}\sum_{\begin{pmatrix} i_1 i_2 \cdots i_s \\ j_1 j_2 \cdots j_s \end{pmatrix}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s} f_{j_1}(\xi_1) f_{j_2}(\xi_2)\cdots f_{j_s}(\xi_s) \\ &= \frac{1}{s!}\sum_{\begin{pmatrix} i_1 i_2 \cdots i_s \\ j_1 j_2 \cdots j_s \end{pmatrix}} \delta^{i_1 i_2\cdots i_s}_{j_1 j_2\cdots j_s} x_{1j_1} x_{2j_2}\cdots x_{sj_s}. \end{aligned}$$

所以式 (8.3.2) 成立. □

为了引进重要的 "外乘", 我们先证明下面引理.

引理 8.3.4 任取 $f \in \Lambda^s$, $g \in \Lambda^t$, $1 \leqslant s, t \leqslant n$. 记

$$h(\xi_1, \cdots, \xi_{s+t}) = \frac{1}{(s+t)!}\sum_{\begin{pmatrix} 1\ 2\ \cdots\ (s+t) \\ j_1\ j_2\ \cdots\ j_{s+t} \end{pmatrix}} \delta^{1\ 2\ \cdots (s+t)}_{i_1 i_2\cdots\ i_{s+t}} f(\xi_{i_1}, \cdots, \xi_{i_s}) g(\xi_{i_{s+1}}, \cdots, \xi_{i_{s+t}}), \tag{8.3.7}$$

则 $h \in \Lambda^{s+t}$. 因此当 $s+t>n$, 则 $h=0$.

证 为了证 $h \in \Lambda^{s+t}$, 只要证对 1, 2, $\cdots$, $s+t$ 的任一排列 $j_1j_2\cdots j_{s+t}$, 则有

$$h(\xi_{j_1},\cdots,\xi_{j_{s+t}}) = \delta^{1\,2\,\cdots\,(s+t)}_{j_1j_2\cdots j_{s+t}} h(\xi_1,\cdots,\xi_{s+t}).$$

由 h 的定义可知

$$h(\xi_{j_1},\cdots,\xi_{j_{s+t}}) = \frac{1}{(s+t)!} \sum_{\binom{j_1j_2\cdots j_{s+t}}{i_1i_2\cdots i_{s+t}}} \delta^{j_1j_2\cdots j_{s+t}}_{i_1i_2\cdots i_{s+t}} f(\xi_{i_1},\cdots,\xi_{i_s}) g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}}).$$

今 $j_1j_2\cdots j_{s+t}$ 是 1, 2, $\cdots$, $s+t$ 的排列, 而 $i_1i_2\cdots i_{s+t}$ 是 $j_1j_2\cdots j_{s+t}$ 的排列, 所以也是 1, 2, $\cdots$, $s+t$ 的排列. 因此 $i_1i_2\cdots i_{s+t}$ 遍历 $j_1j_2\cdots j_{s+t}$ 的所有排列, 可推出 $i_1i_2\cdots i_{s+t}$ 遍历 1, 2, $\cdots$, $(s+t)$ 的所有排列. 所以

$$\begin{aligned}
&\delta^{1\,2\,\cdots\,(s+t)}_{j_1j_2\cdots j_{s+t}} h(\xi_{j_1},\cdots,\xi_{j_{s+t}}) \\
&= \delta^{1\,2\,\cdots\,(s+t)}_{j_1j_2\cdots j_{s+t}} \frac{1}{(s+t)!} \sum_{\binom{j_1j_2\cdots j_{s+t}}{i_1i_2\cdots i_{s+t}}} \delta^{j_1j_2\cdots j_{s+t}}_{i_1i_2\cdots i_{s+t}} f(\xi_{i_1},\cdots,\xi_{i_s}) g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}}), \\
&= \frac{1}{(s+t)!} \sum_{\binom{12\cdots(s+t)}{i_1i_2\cdots\, i_{s+t}}} \delta^{1\,2\,\cdots\,(s+t)}_{i_1i_2\cdots i_{s+t}} f(\xi_{i_1},\cdots,\xi_{i_s}) g(\xi_{i_{s+1}},\cdots,\xi_{i_{s+t}}) = h(\xi_1,\cdots,\xi_{s+t}).
\end{aligned}$$

这证明了 $h \in \Lambda^{s+t}$. □

由引理 8.3.4, 我们定义了线性空间 Λ 到自身内的对应 $(f,g) \to h$. 用此关系式, 我们可以定义 "外乘" 运算如下.

定义 8.3.5 2^n 维线性空间 $\Lambda = \sum\limits_{i=0}^{n} \Lambda^i$ 中存在**外乘运算** "$\wedge$", 它定义为

$$f \wedge g = h, \quad \forall f \in \Lambda^s, \quad g \in \Lambda^t, \tag{8.3.8}$$

其中 $h \in \Lambda^{s+t}$ 由引理 8.3.4 所定义. 在一般情形, 定义

$$\Big(\sum_{s=0}^{n} f_s\Big) \wedge \Big(\sum_{t=0}^{n} g_t\Big) = \sum_{s,t=0}^{n} f_s \wedge g_t = \sum_{k=0}^{n} \Big(\sum_{s+t=k} f_s \wedge g_t\Big) = \sum_{k=0}^{n} h_k, \tag{8.3.9}$$

其中

$$h_k = \sum_{s+t=k} f_s \wedge g_t \in \Lambda^k. \tag{8.3.10}$$

特别

$$f_0 \wedge g_s = g_s \wedge f_0 = f_0 g_s, \quad \forall f_0 \in \Lambda^0 = \mathbb{F}, \quad g_s \in \Lambda^s. \tag{8.3.11}$$

按照上面的定义, 我们有如下一系列运算性质:

定理 8.3.6 符号同上. 我们有: 任取 $f, g, h \in \Lambda$, 则

(i) $f \wedge g$ 关于 f 线性, 关于 g 也线性;

(ii) **结合律**成立, 即

$$(f \wedge g) \wedge h = f \wedge (g \wedge h); \tag{8.3.12}$$

(iii) **加乘分配律**成立, 即

$$f \wedge (g+h) = f \wedge g + f \wedge h, \quad (g+h) \wedge f = g \wedge f + h \wedge f. \tag{8.3.13}$$

因此域 $\mathbb{F}$ 上 2^n 维线性空间 $\Lambda = \sum\limits_{s=0}^{n} \Lambda^s$ 在外乘 "$\wedge$" 下构成有限维结合代数, 称为 **Grassmann 代数**, 或简称为**外代数**.

证 由外乘的定义及直接计算立即可得. □

下面给出外代数 Λ 的一组基.

定理 8.3.7 设 $\mathfrak{L}$ 是 n 维线性空间, $\mathfrak{L}^*$ 是 $\mathfrak{L}$ 的对偶空间. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性空间 $\mathfrak{L}$ 的基, $f_1, f_2, \cdots, f_n$ 是 $\mathfrak{L}^*$ 中 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的对偶基, 则外代数 Λ 有基

$$1, \quad f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_s}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_s \leqslant n, \quad s = 1, 2, \cdots, n, \tag{8.3.14}$$

且有

$$f_i \wedge f_j = -f_j \wedge f_i, \quad 1 \leqslant i, j \leqslant n. \tag{8.3.15}$$

因此 $f_i \wedge f_i = 0$, $1 \leqslant i \leqslant n$. 所以

$$f_s \wedge g_t = (-1)^{st} g_t \wedge f_s, \quad \forall f_s \in \Lambda^s, \quad g_t \in \Lambda^t, \quad 1 \leqslant s, t \leqslant n. \tag{8.3.16}$$

证 由定理 8.3.1 和式 (8.3.2), 我们来证 (8.3.15), (8.3.16) 以及

$$F_{i_1 i_2 \cdots i_s} = f_{i_1} \wedge f_{i_2} \wedge \cdots \wedge f_{i_s}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_s \leqslant n, \quad s = 1, 2, \cdots, n. \tag{8.3.17}$$

今

$$\begin{aligned} (f_i \wedge f_j)(\xi_1, \xi_2) &= \frac{1}{2!} \sum_{\binom{1\,2}{i_1 i_2}} f_i(\xi_{i_1}) f_j(\xi_{i_2}) = \frac{1}{2}[f_i(\xi_1) f_j(\xi_2) - f_i(\xi_2) f_j(\xi_1)] \\ &= -\frac{1}{2}[f_j(\xi_1) f_i(\xi_2) - f_j(\xi_2) f_i(\xi_1)] = -(f_j \wedge f_i)(\xi_1, \xi_2). \end{aligned}$$

这证明了 $f_i \wedge f_j = -f_j \wedge f_i$. 特别有 $f_i \wedge f_i = -f_i \wedge f_i$, 所以 $f_i \wedge f_i = 0$.

下面证 (8.3.17) 成立. 为此对 s 作归纳法. 当 $s=1$ 时显然成立. 设 $s-1$ 时成立, 即我们有 $F_{i_1\cdots i_{s-1}}=f_{i_1}\wedge\cdots\wedge f_{i_{s-1}}$. 为此, 我们要证

$$F_{i_1\cdots i_s}=f_{i_1}\wedge\cdots\wedge f_{i_{s-1}}\wedge f_{i_s}=F_{i_1\cdots i_{s-1}}\wedge f_{i_s}.$$

今

$$(F_{i_1\cdots i_{s-1}}\wedge f_{i_s})(\xi_1,\cdots,\xi_s)=\frac{1}{s!}\sum_{\binom{1\,2\,\cdots\,s}{j_1j_2\cdots j_s}}\delta^{1\,2\,\cdots\,s}_{j_1j_2\cdots j_s}F_{i_1\cdots i_{s-1}}(\xi_{j_1},\cdots,\xi_{j_{s-1}})f_{i_s}(\xi_{j_s}).$$

记 $\xi_k=\sum\limits_{p_k=1}^{n}x_{kp_k}\alpha_{p_k}$, 所以定理 8.3.3 的式 (8.3.5) 给出

$$(F_{i_1\cdots i_{s-1}}\wedge f_{i_s})(\xi_1,\cdots,\xi_s)=\frac{1}{s!}\sum_{\binom{1\,2\,\cdots\,s}{j_1j_2\cdots j_s}}\delta^{1\,2\,\cdots\,s}_{j_1j_2\cdots j_s}\frac{x_{j_si_s}}{(s-1)!}\begin{pmatrix}x_{j_1i_1}&\cdots&x_{j_1i_{s-1}}\\ \vdots&&\vdots\\ x_{j_{s-1}i_1}&\cdots&x_{j_{s-1}i_{s-1}}\end{pmatrix}.$$

注意到 $j_1j_2\cdots j_s$ 遍历 $1,2,\cdots,s$ 的所有排列, 所以 j_s 跑遍 $1,2,\cdots,s$. 因此取定 $j_s=j$, 则排列 $j_1\cdots j_{s-1}$ 遍历 $1,2,\cdots,j-1,j+1,\cdots,s$ 的所有排列. 于是将排列 $j_1\cdots j_{s-1}$ 作一系列相邻两文字的对换, 变为标准排列 $12\cdots(j-1)(j+1)\cdots s$, 共作了 r 次. 于是 $\det\begin{pmatrix}x_{j_1i_1}&\cdots&x_{j_1i_{s-1}}\\ \vdots&&\vdots\\ x_{j_{s-1}i_1}&\cdots&x_{j_{s-1}i_{s-1}}\end{pmatrix}$ 的行相应作了一系列对换. 我们有

$$\det\begin{pmatrix}x_{j_1i_1}&\cdots&x_{j_1i_{s-1}}\\ \vdots&&\vdots\\ x_{j_{s-1}i_1}&\cdots&x_{j_{s-1}i_{s-1}}\end{pmatrix}=(-1)^r\Delta_j,$$

其中 $(-1)^r=\delta^{12\cdots(j-1)(j+1)\cdots s}_{j_1\cdots j_{s-1}}$, 又 $j_1,j_2,\cdots,j_{s-1}$ 遍历 $1,2,\cdots,j-1,j+1,\cdots,s$ 的所有排列,

$$\Delta_j=\det\begin{pmatrix}x_{1i_1}&\cdots&x_{1i_{s-1}}\\ \vdots&&\vdots\\ x_{j-1,i_1}&\cdots&x_{j-1,i_{s-1}}\\ x_{j+1,i_1}&\cdots&x_{j+1,i_{s-1}}\\ \vdots&&\vdots\\ x_{si_1}&\cdots&x_{si_{s-1}}\end{pmatrix},\quad 1\leqslant j\leqslant s.$$

所以

$$(F_{i_1\cdots i_{s-1}}\wedge f_{i_s})(\xi_1,\cdots,\xi_s)=\frac{1}{s!(s-1)!}\sum_{j=1}^{n}\sum_{\begin{pmatrix}1\ 2\ \cdots\ (s-1)\ s\\ j_1 j_2\cdots j_{s-1} j\end{pmatrix}}(-1)^{s-j}x_{ji_s}\Delta_j$$

$$=\frac{1}{s!}\sum_{j=1}^{n}(-1)^{s-j}x_{ji_s}\Delta_j.$$

而

$$\det\begin{pmatrix}x_{1i_1}&\cdots&x_{1i_{s-1}}&x_{1i_s}\\ \vdots&&\vdots&\vdots\\ x_{ji_1}&\cdots&x_{ji_{s-1}}&x_{ji_s}\\ \vdots&&\vdots&\vdots\\ x_{si_1}&\cdots&x_{si_{s-1}}&x_{si_s}\end{pmatrix}=\sum_{j=1}^{s}(-1)^{j+s}x_{ji_s}\Delta_j.$$

这证明了

$$(F_{i_1\cdots i_{s-1}}\wedge f_{i_s})(\xi_1,\cdots,\xi_s)=\frac{1}{s!}\det\begin{pmatrix}x_{1i_1}&\cdots&x_{1i_s}\\ \vdots&&\vdots\\ x_{si_1}&\cdots&x_{si_s}\end{pmatrix}=F_{i_1\cdots i_{s-1}i_s}(\xi_1,\cdots,\xi_s).$$

所以 $F_{i_1\cdots i_{s-1}}\wedge f_{i_s}=F_{i_1\cdots i_s}$. 即线性空间 Λ^s 的基是 $f_{i_1}\wedge\cdots\wedge f_{i_s}, 1\leqslant i_1<\cdots<i_s\leqslant n$.

最后, 我们来证 (8.3.16) 成立. 这只要对线性空间 Λ^s 和线性空间 Λ^t 的基来证明公式就可以了.

任取 $1\leqslant i_1<\cdots<i_s\leqslant n$, $1\leqslant j_1<\cdots<j_t\leqslant n$, 取

$$f_s=f_{i_1}\wedge\cdots\wedge f_{i_s},\quad g_t=f_{j_1}\wedge\cdots\wedge f_{j_t}.$$

由于 $f_i\wedge f_j=-f_j\wedge f_i$, $1\leqslant i,j\leqslant n$. 我们有 $f_s\wedge g_t=(-1)^{st}g_t\wedge f_s$. 这证明了式 (8.3.16) 成立. □

推论 8.3.8 符号同上. 我们有

$$\Lambda^s=\Lambda^1\wedge\Lambda^1\wedge\cdots\wedge\Lambda^1,\tag{8.3.18}$$

其中 Λ^s 的右上角指标 s 表示 Λ^1 外乘 s 次.

习 题 8.3

8.3.1 设 $f(\xi_1,\cdots,\xi_s)$ 是 n 维线性空间 $\mathbb{F}^n$ 的 s 重对称线性函数, 取 $\xi_1,\xi_2,\cdots,\xi_s$ 都相等, 即取函数 $f(x,x,\cdots,x)$ 是 $x_1,\cdots,x_n$ 的 s 次齐次多项式, 其中 $x=(x_1,\cdots,x_n)'$. 反之, 任给独立自变量 $x_1,x_2,\cdots,x_n$ 的 s 次齐次多项式 $F(x_1,\cdots,x_n)$, 则唯一存在一个 s 重对称线性函数 $f(\xi_1,\cdots,\xi_s)$, 使得 $f(x,x,\cdots,x)=F(x,\cdots,x)$.

§8.4 张量场 *

在这一节我们将 §8.3 推广到解析函数空间. 下面给出的定义和结论, 很容易地在讲了微分流形的定义后, 推广到比较弱的函数空间. 因为我们实际上考虑的是只有一个标架的流形, 它就是 $\mathbb{R}^n$.

在这一节中, 我们记 x_1, x_2, $\cdots$, x_n 是 n 个实独立自变量, $n\times 1$ 实矩阵 $x=(x_1,\cdots,x_n)'\in\mathbb{R}^n$.

设 $\{f_{ij}^{(m)}(x)\}$ 是 $\mathbb{R}^n$ 上收敛函数序列, 极限是 $f_{ij}(x)$, 即

$$\lim_{m\to\infty} f_{ij}^{(m)}(x)=f_{ij}(x),\quad 1\leqslant i,j\leqslant n. \tag{8.4.1}$$

于是我们可以定义 $n\times n$ 方阵序列 $A_m(x)$, 其中

$$A_m(x)=\begin{pmatrix} f_{11}^{(m)}(x) & \cdots & f_{1n}^{(m)}(x)\\ \vdots & & \vdots\\ f_{n1}^{(m)}(x) & \cdots & f_{nn}^{(m)}(x)\end{pmatrix},\quad m=1,2,\cdots. \tag{8.4.2}$$

定义

$$\lim_{m\to\infty} A_m(x)=A(x)=\begin{pmatrix} f_{11}(x) & \cdots & f_{1n}(x)\\ \vdots & & \vdots\\ f_{n1}(x) & \cdots & f_{nn}(x)\end{pmatrix}. \tag{8.4.3}$$

设 $f_{ij}(x)$ 是可微函数, 则又可定义

$$\frac{\mathrm{d}\,A(x)}{\mathrm{d}\,x}=\begin{pmatrix} \dfrac{\mathrm{d}\,f_{11}(x)}{\mathrm{d}\,x} & \cdots & \dfrac{\mathrm{d}\,f_{1n}(x)}{\mathrm{d}\,x}\\ \vdots & & \vdots\\ \dfrac{\mathrm{d}\,f_{n1}(x)}{\mathrm{d}\,x} & \cdots & \dfrac{\mathrm{d}\,f_{nn}(x)}{\mathrm{d}\,x}\end{pmatrix}, \tag{8.4.4}$$

$$\int A(x)\mathrm{d}\,x=\begin{pmatrix} \displaystyle\int f_{11}(x)\mathrm{d}\,x & \cdots & \displaystyle\int f_{1n}(x)\mathrm{d}\,x\\ \vdots & & \vdots\\ \displaystyle\int f_{n1}(x)\mathrm{d}\,x & \cdots & \displaystyle\int f_{nn}(x)\mathrm{d}\,x\end{pmatrix}. \tag{8.4.5}$$

定义 8.4.1 $\mathbb{R}^n$ 上函数 $f(x_1,\cdots,x_n)$ 称为 $\mathbb{R}^n$ 上**解析函数**, 如果

$$\begin{aligned} f(x)&=\sum_{i_1,\cdots,i_n=0}^{\infty} a_{i_1\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\\ &=\lim_{N_1\to\infty}\cdots\lim_{N_n\to\infty}\sum_{i_1=0}^{N_1}\cdots\sum_{i_n=0}^{N_n} a_{i_1\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\end{aligned} \tag{8.4.6}$$

是 $\mathbf{R}^n$ 上收敛幂级数. 这时

$$\frac{\partial f(x)}{\partial x_j}=\sum_{i_1,\cdots,i_n=0}^{\infty} i_j a_{i_1\cdots i_n} x_1^{i_1}\cdots x_{j-1}^{i_{j-1}} x_j^{i_j-1} x_{j+1}^{i_{j+1}}\cdots x_n^{i_n} \tag{8.4.7}$$

仍为 $\mathbb{R}^n$ 上的收敛幂级数. 所以 $\dfrac{\partial f(x)}{\partial x_j}$ 是 $\mathbb{R}^n$ 上解析函数, $1\leqslant j\leqslant n$.

$\mathbb{R}^n$ 上所有解析函数构成线性空间 $\mathrm{A}(\mathbb{R}^n)$, 它是条件最强的函数类, 称为**解析函数类**. 注意 $x_1,x_2,\cdots,x_n$ 的多元多项式是解析函数. 它们全体记作 $\mathbb{R}[x_1,x_2,\cdots,x_n]$, **称为实多元多项式环**. 又对一个变元 x 的函数 $\cos(x)$, $\sin(x)$, $\mathrm{e}^x=\exp(x)$ 都是解析函数, 但是 $\ln(x)$ 不是解析函数.

定义 8.4.2　无限维实线性空间 $\mathrm{A}(\mathbb{R}^n)$ 上线性算子 X 定义为 $X(\mathrm{A}(\mathbb{R}^n))\subset \mathrm{A}(\mathbb{R}^n)$, 又

$$X(af+bg)=aX(f)+bX(g),\quad \forall\, a,b\in\mathbb{R},\quad f,g\in\mathrm{A}(\mathbb{R}^n).$$

$\mathrm{A}(\mathbb{R}^n)$ 上线性算子 X 称为 $\mathbb{R}^n$ 上**向量场**, 如果它适合

$$X(fg)=fX(g)+gX(f),\quad \forall\, f,g\in\mathrm{A}(\mathbb{R}^n). \tag{8.4.8}$$

$\mathbb{R}^n$ 上所有向量场构成实线性空间 $\mathrm{T}(\mathbb{R}^n)$.

定义 8.4.3　设 $\mathfrak{L}$ 是域 $\mathbb{F}$ 上线性空间, $\mathfrak{L}$ 上有一种 "乘法", 称为**换位运算**, 如果它适合条件

(1)　反交换性

$$[Y,X]=-[X,Y],\quad \forall\, X,Y\in\mathfrak{L}; \tag{8.4.9}$$

(2)　双线性性

$$[aX+bY,Z]=a[X,Z]+b[Y,Z],\quad \forall\, a,b\in\mathbb{R},\quad X,Y,Z\in\mathfrak{L}; \tag{8.4.10}$$

(3)　Jacobi 恒等式

$$[[X,Y],Z]+[[Z,X],Y]+[[Y,Z],X]=0,\quad \forall\, X,Y,Z\in\mathfrak{L}, \tag{8.4.11}$$

则 $\mathfrak{L}$ 称为域 $\mathbb{F}$ **上Lie 代数**.

引理 8.4.4　实线性空间 $\mathrm{T}(\mathbb{R}^n)$ 在换位运算

$$[X,Y]=XY-YX,\quad \forall\, X,Y\in\mathrm{T}(\mathbb{R}^n) \tag{8.4.12}$$

下构成实 Lie 代数.

证　只要证 $[X,Y]\in\mathrm{T}(\mathbb{R}^n)$, $\forall\, X,\,Y\in\mathrm{T}(\mathbb{R}^n)$. 事实上, 显然 $[X,Y]=XY-YX$

是 $\mathrm{A}(\mathbb{R}^n)$ 上线性算子, 又任取 $f,g\in\mathrm{A}(\mathbb{R}^n)$, 则

$$\begin{aligned}(fg)&=XY(fg)-YX(fg)=X(fY(g)+gY(f))-Y(fX(g)+gX(f))\\&=fXY(g)+X(f)Y(g)+gXY(f)+X(g)Y(f)\\&\quad-fYX(g)-Y(f)X(g)-gYX(f)-Y(g)X(f)\\&=f[X,Y](g)+g[X,Y](f).\end{aligned}$$

这证明了 $[X,Y]\in\mathrm{T}(\mathbb{R}^n)$, $\forall X,Y\in\mathrm{T}(\mathbb{R}^n)$. 由直接计算可知实线性空间 $\mathrm{T}(\mathbb{R}^n)$ 在换位运算 $[X,Y]=XY-YX$ 下封闭, 且是实 Lie 代数. □

考虑 n 个特殊的向量场

$$\frac{\partial}{\partial x_1},\quad \frac{\partial}{\partial x_2},\quad \cdots,\quad \frac{\partial}{\partial x_n}. \tag{8.4.13}$$

今显然 $\dfrac{\partial}{\partial x_i}$ 是 $\mathrm{A}(\mathbb{R}^n)$ 上线性算子, 又

$$\frac{\partial}{\partial x_i}f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)=f\frac{\partial g}{\partial x_i}+g\frac{\partial f}{\partial x_i}.$$

所以 $\dfrac{\partial}{\partial x_1}, \dfrac{\partial}{\partial x_2}, \cdots, \dfrac{\partial}{\partial x_n}\in\mathrm{T}(\mathbb{R}^n)$. 下面来证 $\mathbb{R}^n$ 上向量场 X 有坐标表达式:

$$X=\sum_{i=1}^n\xi_i(x)\frac{\partial}{\partial x_i},\quad \xi_i(x)=X(x_i),\quad 1\leqslant i\leqslant n. \tag{8.4.14}$$

事实上,

$$\begin{aligned}X(f)&=X\Big(\sum_{i_1,\cdots,i_n=0}^{\infty}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\Big)\\&=\sum_{i_1,\cdots,i_n=0}^{\infty}a_{i_1i_2\cdots i_n}X(x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n})\\&=\sum_{i_1,\cdots,i_n=0}^{\infty}a_{i_1i_2\cdots i_n}\sum_{j=1}^n i_jx_1^{i_1}\cdots x_{j-1}^{i_{j-1}}x_j^{i_j-1}x_{j+1}^{i_{j+1}}\cdots x_n^{i_n}X(x_i)\in\mathrm{A}(\mathbb{R}^n).\end{aligned}$$

记 $\xi_i(x)=X(x_i)$, $1\leqslant i\leqslant n$, 则

$$X(f)=\sum_{j=1}^n\xi_j(x)\frac{\partial f}{\partial x_j}=\Big(\sum_{j=1}^n\xi_j(x)\frac{\partial}{\partial x_j}\Big)(f),\quad \forall f\in\mathrm{A}(\mathbb{R}^n).$$

所以证明了断言.

反之, 任取 $\xi_1(x), \xi_2(x), \cdots, \xi_n(x)\in\mathrm{A}(\mathbb{R}^n)$, 由直接计算可知

$$\sum_{i=1}^n\xi_i(x)\frac{\partial}{\partial x_i}\in\mathrm{T}(\mathbb{R}^n).$$

所以

$$\mathrm{T}(\mathbb{R}^n) = \{\, X = \sum_{i=1}^{n} \xi_i(x)\frac{\partial}{\partial x_i} \mid \forall \xi_1(x), \cdots, \xi_n(x) \in \mathrm{A}(\mathbb{R}^n)\,\}. \tag{8.4.15}$$

□

上面的讨论也证明了所有向量场构成的集合 $\mathrm{T}(\mathbb{R}^n)$ 是纯量取在 $\mathrm{A}(\mathbb{R}^n)$ 上的线性空间. 由于 $\mathrm{A}(\mathbb{R}^n)$ 中非零元不一定有逆元, 所以我们改称 $\mathrm{T}(\mathbb{R}^n)$ 为一个环 $\mathrm{A}(\mathbb{R}^n)$ 上的**模**. 对这个模而言,

$$\frac{\partial}{\partial x_1}, \quad \frac{\partial}{\partial x_2}, \quad \cdots, \quad \frac{\partial}{\partial x_n}$$

是一组 "基". 下面开始, 我们只考虑 $\mathrm{A}(\mathbb{R}^n)$- 模.

相同于普通线性空间的对偶空间, 我们定义 $\mathrm{A}(\mathbb{R}^n)$- 模 $\mathrm{T}(\mathbb{R}^n)$ 的对偶空间

$$\mathrm{T}(\mathbb{R}^n)^* = \Lambda^1(\mathbb{R}^n),$$

其中元称为 1- 形式.

下面从另外一个角度来看模. 我们知道 $\mathrm{A}(\mathbb{R}^n)$ 是解析函数类, 自变量取值在 $\mathbb{R}^n$ 中. 在 $\mathbb{R}^n$ 中取定一点 x_0, 则任一解析函数 $f(x)$ 的函数值是 $f(x_0) \in \mathbb{F}$. 所以对 $\mathbb{R}^n$ 中每个固定点 x_0, 所谓 $\mathrm{A}(\mathbb{R}^n)$- 模便成为通常的线性空间. 这个线性空间记作 $\mathfrak{L}_{x_0}$. 于是有线性空间簇 $\mathfrak{L}_x, \forall x \in \mathbb{R}^n$. 所以 $\mathrm{A}(\mathbb{R}^n)$- 模实际上就是在 $\mathbb{R}^n$ 中的每点, 装上一个普通的线性空间. 但是要求 $x \to \mathfrak{L}_x$ 在 $\mathbb{R}^n$ 上解析.

定义 8.4.5 $\mathrm{A}(\mathbb{R}^n)$- 模 $\mathrm{T}(\mathbb{R}^n)$ 上 $\mathrm{A}(\mathbb{R}^n)$ 线性函数 φ 定义为

$$\varphi(\sum_{i=1}^{n} \xi_i(x)\frac{\partial}{\partial x_i}) = \sum_{i=1}^{n} \xi_i(x)\varphi(\frac{\partial}{\partial x_i}) = \sum_{i=1}^{n} \xi_i(x) g_i(x), \tag{8.4.16}$$

其中

$$\varphi(\frac{\partial}{\partial x_i}) = g_i(x) \in \mathrm{A}(\mathbb{R}^n), \quad 1 \leqslant i \leqslant n. \tag{8.4.17}$$

所有 $\mathrm{A}(\mathbb{R}^n)$- 线性函数构成 $\mathrm{A}(\mathbb{R}^n)$- 模 $\Lambda(\mathbb{R}^n)$, 其中元称为 $\mathbb{R}^n$ 上**1- 形式**.

我们约定 $\mathrm{d}x_1, \mathrm{d}x_2, \cdots, \mathrm{d}x_n \in \Lambda(\mathbb{R}^n)$, 它们定义为

$$\mathrm{d}x_i(\frac{\partial}{\partial x_j}) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n, \tag{8.4.18}$$

其中 δ_{ij} 是 Kronecker 符号. 又任取

$$X = \sum_{i=1}^{n} \xi_i(x)\frac{\partial}{\partial x_i} \in \mathrm{T}(\mathbb{R}^n),$$

定义

$$\mathrm{d}x_i(X) = \mathrm{d}x_i(\sum_{j=1}^{n} \xi_j(x)\frac{\partial}{\partial x_j}) = \sum_{j=1}^{n} \xi_j(x)\mathrm{d}x_i(\frac{\partial}{\partial x_j}) = \xi_i(x), \quad 1 \leqslant i \leqslant n.$$

由定义可知 $\mathrm{d}\,x_1, \mathrm{d}\,x_2, \cdots, \mathrm{d}\,x_n \in \Lambda(\mathbb{R}^n)$.

而且任一 1- 形式为

$$\omega = \sum_{j=1}^{n} g_i(x)\mathrm{d}\,x_j, \quad g_j(x) = \omega(\frac{\partial}{\partial x_j}), \quad 1 \leqslant j \leqslant n. \tag{8.4.19}$$

所以所有 1- 形式构成 $\mathrm{A}(\mathbb{R}^n)$- 模

$$\Lambda^1(\mathbb{R}^n) = \{\, \omega = \sum_{i=1}^{n} g_i(x)\mathrm{d}\,x_i \mid \forall\, g_1(x), \cdots, g_n(x) \in \mathrm{A}(\mathbb{R}^n)\,\}. \tag{8.4.20}$$

事实上, 任取 $\omega \in \Lambda^1(\mathbb{R}^n)$, 则

$$\begin{aligned}
\omega(X) &= \omega(\sum_{i=1}^{n} \xi_i(x)\frac{\partial}{\partial x_i}) = \sum_{i=1}^{n} \xi_i(x)\omega(\frac{\partial}{\partial x_i}) = \sum_{i=1}^{n} \xi_i(x)g_i(x) \\
&= \sum_{i,j=1}^{n} \xi_j(x)g_i(x)\mathrm{d}\,x_i(\frac{\partial}{\partial x_j}) = (\sum_{i=1}^{n} g_i(x)\mathrm{d}\,x_i)(\sum_{j=1}^{n} \xi_j(x)\frac{\partial}{\partial x_j}) \\
&= (\sum_{i=1}^{n} g_i(x)\mathrm{d}\,x_i)(X), \quad \forall X \in \mathrm{T}(\mathbb{R}^n).
\end{aligned}$$

所以证明了 $\omega = \sum\limits_{i=1}^{n} g_i(x)\mathrm{d}\,x_i$, 其中 $g_i(x) = \omega(\dfrac{\partial}{\partial x_i})$, $1 \leqslant i \leqslant n$.

反之, 任取 $g_1(x), g_2(x), \cdots, g_n(x) \in \mathrm{A}(\mathbb{R}^n)$, 由 $\mathrm{d}\,x_i \in \Lambda^1(\mathbb{R}^n)$, 所以易证 $\sum\limits_{i=1}^{n} g_i(x)\mathrm{d}\,x_i \in \mathrm{A}(\mathbb{R}^n)$. □

和 §8.2 一样, 我们从 $\mathrm{A}(\mathbb{R}^n)$ 模 $\mathrm{T}(\mathbb{R}^n)$ 及其**对偶模** $\Lambda^1(\mathbb{R}^n)$ 出发, 定义 $\mathrm{A}(\mathbb{R}^n)$- 多重线性函数, 即 (s,t) 型张量场, 定义张量积, 从而定义张量代数. 这只要将线性空间 $\mathfrak{L}$ 改为 $\mathrm{A}(\mathbb{R}^n)$- 模. 对此而言, $\mathrm{T}(\mathbb{R}^n)$ 中有 $\mathrm{A}(\mathbb{R}^n)$- 模的基 $\dfrac{\partial}{\partial x_1}, \dfrac{\partial}{\partial x_2}, \cdots, \dfrac{\partial}{\partial x_n}$, $\Lambda^1(\mathbb{R}^n)$ 中有 $\mathrm{A}(\mathbb{R}^n)$- 模的对偶基

$$\mathrm{d}\,x_1, \quad \mathrm{d}\,x_2, \quad \cdots, \quad \mathrm{d}\,x_n.$$

于是所有讨论及结论成立. 特别我们有 §8.3 的结论.

$\mathrm{A}(\mathbb{R}^n)$- 多重线性函数 $f(\xi_1, \cdots, \xi_s; g_1, \cdots, g_t)$ 全体构成 $\mathrm{A}(\mathbb{R}^n)$- 模 $\mathfrak{D}_t^s$, 在其中可以定义加法及纯量 (是解析函数) 积, 又可定义**张量积**. 从而 $\mathfrak{D} = \sum\limits_{s,t=0}^{\infty} \mathfrak{D}_t^s$ 构成**张量场代数**. 又 $\mathfrak{D}_t^s$ 中有基

$$\mathrm{d}\,x_{i_1} \otimes \cdots \otimes \mathrm{d}\,x_{i_s} \otimes \frac{\partial}{\partial x_{j_1}} \otimes \cdots \otimes \frac{\partial}{\partial x_{j_t}}, \quad 1 \leqslant i_1, \cdots, i_s, j_1, \cdots, j_t \leqslant n. \tag{8.4.21}$$

又

$$\mathfrak{D}_0^0 = \mathrm{A}(\mathbb{R}^n), \tag{8.4.22}$$

$$\mathfrak{D}_t^s = (\Lambda^1(\mathbb{R}^n))^s \otimes (\mathrm{T}(\mathbb{R}^n))^t, \tag{8.4.23}$$

其中

$$\mathfrak{D}_0^1 = \Lambda^1(\mathbb{R}^n), \quad \mathfrak{D}_1^0 = \mathrm{T}(\mathbb{R}^n), \quad \mathfrak{D}_0^0 = \mathrm{A}(\mathbb{R}^n).$$

又

$$(\Lambda^1(\mathbb{R}^n))^s = \Lambda^1(\mathbb{R}^n) \otimes \cdots \otimes \Lambda^1(\mathbb{R}^n)$$

是 s 个 $\Lambda^1(\mathbb{R}^n)$ 的张量积, $(\mathrm{T}(\mathbb{R}^n))^t = \mathrm{T}(\mathbb{R}^n) \otimes \cdots \otimes \mathrm{T}(\mathbb{R}^n)$ 是 t 个 $\mathrm{T}(\mathbb{R}^n)$ 的张量积. 而 $\mathfrak{D}_t^s$ 中元称为 s 阶共变, t 阶反变张量场, 即**(s,t) 型张量场**.

同样可定义**对称共变张量场** 和**反对称共变张量场**. 所有反对称共变张量场全体构成 $\mathrm{A}(\mathbb{R}^n)$- 模

$$\Lambda = \Lambda^0 + \Lambda^1 + \cdots + \Lambda^n. \tag{8.4.24}$$

其中 Λ^s 是 $\mathrm{A}(\mathbb{R}^n)$- 子模. Λ^s 有基 $\mathrm{d}\,x_{i_1} \wedge \cdots \wedge \mathrm{d}\,x_{i_s}$, $1 \leqslant i_1 < \cdots < i_s \leqslant n$. Λ^s 中任一元是

$$\omega = \sum_{1\leqslant i_1 < \cdots < i_s \leqslant n} a_{i_1 i_2 \cdots i_s}(x) \mathrm{d}\,x_{i_1} \wedge \cdots \wedge \mathrm{d}\,x_{i_s}, \tag{8.4.25}$$

其中 $s = 1, 2, \cdots, n$. 又 Λ^0 有基 1.

定义 8.4.6 $\mathrm{A}(\mathbb{R}^n)$- 模 Λ 上有映射 d, 称为**外微分**, 它定义为: 任取 $\omega \in \Lambda^s$, 当 $s = 0$, 则 $\omega \in \Lambda^0 = \mathrm{A}(\mathbb{R}^n)$. 这时定义

$$\mathrm{d}\,\omega = \sum_{i=1}^n \frac{\partial \omega}{\partial x_i} \mathrm{d}\,x_i \in \Lambda^1. \tag{8.4.26}$$

当 $s > 0$, 则

$$\omega = \sum_{1\leqslant i_1 < \cdots < i_s \leqslant n} a_{i_1 i_2 \cdots i_s}(x) \mathrm{d}\,x_{i_1} \wedge \cdots \wedge \mathrm{d}\,x_{i_s} \in \Lambda^s. \tag{8.4.27}$$

定义

$$\begin{aligned} \mathrm{d}\,\omega &= \sum_{1\leqslant i_1 < \cdots < i_s \leqslant n} (\mathrm{d}\,a_{i_1 i_2 \cdots i_s}(x)) \wedge \mathrm{d}\,x_{i_1} \wedge \cdots \wedge \mathrm{d}\,x_{i_s} \\ &= \sum_{1\leqslant i_1 < \cdots < i_s \leqslant n} \sum_{j=1}^n \frac{\partial a_{i_1 i_2 \cdots i_s}(x)}{\partial x_j} \mathrm{d}\,x_j \wedge \mathrm{d}\,x_{i_1} \wedge \cdots \wedge \mathrm{d}\,x_{i_s} \in \Lambda^{s+1}. \end{aligned} \tag{8.4.28}$$

定理 8.4.7 外代数 Λ 上外微分映射 $\mathrm{d}\colon \Lambda^s \to \Lambda^{s+1}$, $0 \leqslant s \leqslant n$ 有性质:

(1) d 是线性空间 Λ 上线性变换;

(2) $\mathrm{d}(\mathbb{R}) = 0$, $\mathrm{d}\,\Lambda^n = 0$;

(3) $\mathrm{d}(f \wedge g) = (\mathrm{d}\,f) \wedge g + (-1)^s f \wedge (\mathrm{d}\,g)$, $\forall\, f \in \Lambda^s$, $g \in \Lambda$;

(4) $\mathrm{d}^2 = 0$.

证 **(1)** 和**(2)** 显然成立.

下面证(3) 成立. 任取

$$f=\sum_{1\leqslant i_1<\cdots<i_s\leqslant n} a_{i_1i_2\cdots i_s}(x)\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}, \tag{8.4.29}$$

$$g=\sum_{1\leqslant j_1<\cdots<j_t\leqslant n} b_{j_1j_2\cdots j_t}(x)\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_t}, \tag{8.4.30}$$

则

$$\begin{aligned}
\mathrm{d}(f\wedge g)=&\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}\sum_{1\leqslant j_1<\cdots<j_t\leqslant n}\sum_{p=1}^{n}\frac{\partial a_{i_1\cdots i_s}(x)b_{j_1\cdots j_t}(x)}{\partial x_p}\\
&\mathrm{d}x_p\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_t}\\
=&\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}\sum_{1\leqslant j_1<\cdots<j_t\leqslant n}\sum_{p}\Big(\frac{\partial a_{i_1\cdots i_s}(x)}{\partial x_p}b_{j_1\cdots j_t}(x)\\
&+a_{i_1\cdots i_s}(x)\frac{\partial b_{j_1\cdots j_t}(x)}{\partial x_p}\Big)\mathrm{d}x_p\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_t}\\
=&\Big(\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}\sum_{p}\frac{\partial a_{i_1\cdots i_s}(x)}{\partial x_p}\mathrm{d}x_p\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\Big)\wedge\\
&\Big(\sum_{1\leqslant j_1<\cdots<j_t\leqslant n} b_{j_1\cdots j_t}(x)\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_t}\Big)\\
&+(-1)^s\Big(\sum_{1\leqslant i_1<\cdots<i_s\leqslant n} a_{i_1\cdots i_s}(x)\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\Big)\wedge\\
&\Big(\sum_{1\leqslant j_1<\cdots<j_t\leqslant n}\sum_{p}\frac{\partial b_{j_1\cdots j_t}(x)}{\partial x_p}\mathrm{d}x_p\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_t}\Big)\\
=&(\mathrm{d}f)\wedge g+(-1)^s f(\mathrm{d}g).
\end{aligned}$$

这证明了(3) 成立.

最后证(4) 成立. 任取 $f=\sum\limits_{1\leqslant i_1<\cdots<i_s\leqslant n} a_{i_1i_2\cdots i_s}(x)\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\in\Lambda^s$, 则

$$\begin{aligned}
\mathrm{d}^2f=\mathrm{d}(\mathrm{d}f)&=\mathrm{d}\Big(\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}\sum_{p=1}^{n}\frac{\partial a_{i_1\cdots i_s}(x)}{\partial x_p}\mathrm{d}x_p\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}\Big)\\
&=\sum_{1\leqslant i_1<\cdots<i_s\leqslant n}\sum_{p,q=1}^{n}\frac{\partial^2 a_{i_1\cdots i_s}(x)}{\partial x_p\partial x_q}\mathrm{d}x_p\wedge\mathrm{d}x_q\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_s}.
\end{aligned}$$

今显然 $\sum\limits_{p,q=1}^{n}\dfrac{\partial^2 a_{i_1\cdots i_s}(x)}{\partial x_p\partial x_q}\mathrm{d}x_p\wedge\mathrm{d}x_q=0$. 因此 $\mathrm{d}^2f=0, \forall f\in\Lambda^s$. 所以 $\mathrm{d}^2=0$. □

利用定理 8.4.7, 我们有如下子空间的线性映射序列:

$$\mathbb{R}\xrightarrow{\mathrm{d}}\Lambda^0=\mathrm{A}(\mathbb{F}^n)\xrightarrow{\mathrm{d}_0}\Lambda^1\xrightarrow{\mathrm{d}_1}\Lambda^2\xrightarrow{\mathrm{d}_2}\cdots\xrightarrow{\mathrm{d}_{n-2}}\Lambda^{n-1}\xrightarrow{\mathrm{d}_{n-1}}\Lambda^n\xrightarrow{\mathrm{d}_n}0, \tag{8.4.31}$$

其中 $\mathrm{d}, \mathrm{d}_0, \mathrm{d}_1, \cdots, \mathrm{d}_n$ 都是上面定义的外微分映射 d.

定义 8.4.8　象集 $\mathrm{d}_{k-1}\Lambda^{k-1}$ 中元称为 ***k* 阶上边缘**, 核

$$\ker(\mathrm{d}_k) = \{ f \in \Lambda^k \mid \mathrm{d}_k f = 0 \} \tag{8.4.32}$$

中元称为 ***k* 阶上循环**.

引理 8.4.9　象集 $\mathrm{d}_{k-1}\Lambda^{k-1}$ 和核 $\ker(\mathrm{d}_k)$ 是实线性子空间, 它们有

$$\mathrm{d}_{k-1}\Lambda^{k-1} \subset \ker(\mathrm{d}_k) \subset \Lambda^k. \tag{8.4.33}$$

证　由 $\mathrm{d}_{k-1} = \mathrm{d}$, $\mathrm{d}_k = \mathrm{d}$. 任取 $f \in \Lambda^{k-1}$, 由 $\mathrm{d}^2 f = 0$ 可知 $\mathrm{d}f \in \Lambda^k$, 又 $\mathrm{d}^2 f = \mathrm{d}_k(\mathrm{d}_{k-1}f) = 0$, 即 $\mathrm{d}_{k-1}f \in \ker(\mathrm{d}_k)$, $\forall f \in \Lambda^{k-1}$. □

定义 8.4.10　商空间 $\ker(\mathrm{d}_k)/\mathrm{d}_{k-1}\Lambda^{k-1} = H^k(\Lambda, \mathrm{d})$ 称为外代数 Λ 的 ***k* 阶上同调群**, $0 \leqslant k < n$.

显然有 $\mathrm{d}\mathbb{R} = 0$, 因此,

$$\ker(\mathrm{d}_0) = \{ a \in \Lambda^0 \mid \mathrm{d}a(x) = \sum_{p=1}^{n} \frac{\partial a(x)}{\partial x_p}\mathrm{d}x_p = 0 \} = \mathbb{R}. \tag{8.4.34}$$

所以**零阶上同调群** $H^0(\Lambda, \mathrm{d}) = \mathbb{R}$.

对什么指标 $k \in \{0, 1, \cdots, n-1\}$, 上同调群 $H^k(\Lambda, \mathrm{d}) = 0$ 是至关重要的问题.

定义 8.4.11　序列

$$\Lambda^{k-1} \to \Lambda^k \to \Lambda^{k-1}$$

称为**正合序列**, 如果

$$\ker(\mathrm{d}_k) = \mathrm{d}_{k-1}\Lambda_{k-1},$$

即如果 $H^k(\Lambda, \mathrm{d}) = 0$.

习　题　8.4

8.4.1　通常的二重积分 $\iint_D F(x,y)\mathrm{d}x\mathrm{d}y$ 改记作 $\iint_D F(x,y)\mathrm{d}x \wedge \mathrm{d}y$. 试证: 对任一变数变换

$$\sigma: \qquad u = f(x,y), \quad v = g(x,y),$$

设它有逆

$$\sigma^{-1}: \qquad x = \tilde{f}(u,v), \quad y = \tilde{g}(u,v),$$

则

$$\iint_D F(x,y)\mathrm{d}x \wedge \mathrm{d}y = \iint_{\sigma(D)} F(\sigma^{-1}(u,v))J(u,v)\mathrm{d}u \wedge \mathrm{d}v,$$

其中 Jacobian 行列式是

$$J(u,v)=\det\begin{pmatrix}\dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v}\\[2ex] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v}\end{pmatrix},$$

又

$$J(u,v)^{-1}=J(x,y)=\det\begin{pmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y}\\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y}\end{pmatrix}.$$

第九章　实 Euclid 空间

§9.1　双线性函数

在这一节, 为了引进实和复 Euclid 空间, 以及实和复的二次型和 Hermite 型, 我们引进域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数, 这里域 $\mathbb{F}=\mathbb{Q}$, $\mathbb{R}$ 或 $\mathbb{C}$.

定义 9.1.1　域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的函数 $f(\alpha,\beta)$ 称为**双线性函数**, 如果有

$$f(c\alpha+d\beta,\gamma)=cf(\alpha,\gamma)+df(\beta,\gamma),\quad f(\alpha,c\beta+d\gamma)=cf(\alpha,\beta)+df(\alpha,\gamma),\tag{9.1.1}$$

其中 α, β, γ, $c,d\in\mathbb{F}$.

由定义可知, 给定双线性函数 $f(\alpha,\beta)$. 那末, 当 β 固定时, $f(\alpha,\beta)$ 便是 α 的线性函数; 又当 α 固定时, $f(\alpha,\beta)$ 便是 β 的线性函数, 所以 $f(\alpha,\beta)$ 称为双线性函数.

和线性函数的情形一样, 下面将定出所有的双线性函数.

定理 9.1.2　在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 α_1, α_2, $\cdots$, α_n. 设 f 是线性空间 $\mathfrak{L}$ 的双线性函数. 记 n 阶方阵

$$A=\begin{pmatrix} f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n)\\ \vdots & & \vdots\\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n)\end{pmatrix},\tag{9.1.2}$$

则方阵 A 称为双线性函数 f 在基 α_1, α_2, $\cdots$, α_n 下的**方阵表示**. 这时, 对线性空间 $\mathfrak{L}$ 中任两向量 $\alpha=\sum\limits_{i=1}^{n}x_i\alpha_i$, $\beta=\sum\limits_{j=1}^{n}y_j\alpha_j$, 有

$$f(\alpha,\beta)=x'Ay,\tag{9.1.3}$$

其中

$$x = (x_1, x_2, \cdots, x_n)', \quad y = (y_1, y_2, \cdots, y_n)' \tag{9.1.4}$$

是两个 $n \times 1$ 矩阵. 反之, 任取 n 阶方阵 A, 则

$$x'Ay = f(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j)$$

定义了线性空间 $\mathfrak{L}$ 的双线性函数. 所以在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定基后, 双线性函数 f 与域 $\mathbb{F}$ 上 n 阶方阵 A 之间有一个自然的一一对应: $f \to A$.

证 令

$$f(\alpha, \beta) = f(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j) = \sum_{i,j=1}^{n} x_i y_j f(\alpha_i, \alpha_j) = x'Ay.$$

反之, 任取 n 阶方阵 $A = (a_{ij})$. 由定义可知 $f(\sum\limits_{i=1}^{n} x_i\alpha_i, \sum\limits_{j=1}^{n} y_j\alpha_j) = x'Ay$ 是线性空间 $\mathfrak{L}$ 上双线性函数. 为了证明 $f \to A$ 是一一对应, 只要证线性空间 $\mathfrak{L}$ 的双线性函数 f 和 g 有: $f = g$ 当且仅当 $f(\alpha_i, \alpha_j) = g(\alpha_i, \alpha_j)$, $1 \leqslant i, j \leqslant n$.

事实上, 记 A 和 B 分别是双线性函数 f 和 g 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的方阵表示, 即 $f(\alpha, \beta) = x'Ay$, $g(\alpha, \beta) = x'By$, $\forall \alpha, \beta \in \mathfrak{L}$. 由此可知 n 阶方阵 $A = B$ 当且仅当双线性函数 $f = g$. □

上面所确定的一一对应, 是在给定一组基的前提下建立的. 因此, 在不同基下, 同一双线性函数对应的 n 阶方阵不同. 下面来建立它们间的关系.

定理 9.1.3 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$. 记基变换对应的 n 阶非异方阵是 $P = (p_{ij})$. 设 f 是线性空间 $\mathfrak{L}$ 的双线性函数, 它们在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 下对应的 n 阶方阵分别是 A 和 B, 则有

$$B = P'AP. \tag{9.1.5}$$

证 由 n 阶方阵 B 的定义可知, 它的第 i 行, 第 j 列元素是

$$f(\beta_i, \beta_j) = f(\sum_{k=1}^{n} p_{ki}\alpha_k, \sum_{s=1}^{n} p_{sj}\alpha_s) = \sum_{k,s=1}^{n} p_{ki} f(\alpha_k, \alpha_s) p_{sj},$$

其中 $f(\alpha_k, \alpha_s)$ 是 n 阶方阵 A 的第 k 行, 第 s 列元素. 将上式写成方阵乘积的形式, 则有 $B = P'AP$. □

于是, 自然可以引进

定义 9.1.4 域 $\mathbb{F}$ 上 n 阶方阵 A 和 B 称为**相合的**, 如果存在 n 阶非异方阵 P, 使得

$$B = P'AP. \tag{9.1.6}$$

定理 9.1.5 域 $\mathbb{F}$ 上 n 阶方阵 A 和 B 间的相合关系是等价关系.

证 在相合定义中取 P 是 n 阶单位方阵, 所以证明了 A 和 A 相合, 即证明了反身性. 设 A 和 B 相合, 即存在 n 阶非异方阵 P, 使得 $B = P'AP$. 于是 $A = (P^{-1})'B(P^{-1})$. 这证明了 B 和 A 相合, 所以证明了对称性. 最后, 若 n 阶方阵 A, B 和 C 有: A 和 B 相合; B 和 C 也相合, 即存在 n 阶非异方阵 P 和 Q, 使得 $B = P'AP$, $C = Q'BQ$. 于是 $C = Q'(P'AP)Q = (PQ)'A(PQ)$. 这证明了 A 和 C 相合, 所以证明了传递性. 至此证明了相合关系是等价关系. □

由相合定义可知

定理 9.1.6 域 $\mathbb{F}$ 上 n 阶方阵 A 和 B 相合当且仅当存在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数 f, 使得 A 和 B 是双线性函数 f 在不同基下对应的方阵. 所以所有互相相合的方阵构成的集合 (即等价类) 和双线性函数间有一个自然的一一对应关系. 换句话说, 方阵在相合的等价类表示了一个双线性函数.

因为双线性函数的定义和线性空间有关, 与基的选取无关. 上面的定理实际上给出了方阵相合这个概念的几何解释. 在下面, 我们引进一些重要的双线性函数.

定义 9.1.7 记 f 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数.

(1) f 称为**对称双线性函数**, 如果有

$$f(\alpha,\beta) = f(\beta,\alpha), \quad \forall\, \alpha,\beta \in \mathfrak{L}; \tag{9.1.7}$$

(2) f 称为**反对称双线性函数**, 如果有

$$f(\alpha,\beta) = -f(\beta,\alpha), \quad \forall\, \alpha,\beta \in \mathfrak{L}. \tag{9.1.8}$$

定理 9.1.8 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数是对称双线性函数当且仅当在 $\mathfrak{L}$ 中存在一组基, 使得在这组基下, 它对应的方阵表示是对称方阵; 双线性函数是反对称双线性函数当且仅当在 $\mathfrak{L}$ 中存在一组基, 使得在这组基下, 它对应的方阵表示是反对称方阵.

证 在线性空间 $\mathfrak{L}$ 中任取一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 设 f 是线性空间 $\mathfrak{L}$ 的对称双线性函数, 则有 $f(\alpha_i,\alpha_j) = f(\alpha_j,\alpha_i)$, $1 \leqslant i,j \leqslant n$. 于是, f 的方阵表示

$$A = \begin{pmatrix} f(\alpha_1,\alpha_1) & \cdots & f(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ f(\alpha_n,\alpha_1) & \cdots & f(\alpha_n,\alpha_n) \end{pmatrix}$$

是对称方阵. 反之, 若方阵 A 对称, 即有 $f(\alpha_i,\alpha_j) = f(\alpha_j,\alpha_i)$, $1 \leqslant i,j \leqslant n$. 自然 $f(\alpha,\beta) = f(\beta,\alpha)$, $\forall\, \alpha$, $\beta \in \mathfrak{L}$. 即 f 是对称双线性函数. 在反对称的情形, 证明是相

同的. □

下面的结论是比较重要的.

引理 9.1.9 域 $\mathbb{F}$ 上 n 阶方阵 A 可以唯一地分解为域 $\mathbb{F}$ 上 n 阶对称方阵 S (称为 A 的**对称部分**) 和域 $\mathbb{F}$ 上 n 阶反对称方阵 K (称为 A 的**反对称部分**) 的和, 其中

$$S=\frac{1}{2}(A+A'),\quad K=\frac{1}{2}(A-A'). \tag{9.1.9}$$

证 显然, 记 $S=\frac{1}{2}(A+A')$, $K=\frac{1}{2}(A-A')$, 则有 $S'=S$, $K'=-K$. 下面证唯一性. 今若 $A=S+K=S_1+K_1$, 其中 $S'=S$, $S_1'=S_1$, $K'=-K$, $K_1'=-K_1$, 于是 $S-S_1=K_1-K$, 又 $(S-S_1)'=S'-S_1'=S-S_1$, 所以 $(K_1-K)'=K_1-K$. 因此, $K_1-K=K_1'-K'=-K_1+K$. 这证明了 $K_1=K$, 所以 $S_1=S$. □

定理 9.1.10 域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数 $f(\alpha,\beta)$ 可唯一地分解为一个对称双线性函数 $s(\alpha,\beta)$ 和一个反对称双线性函数 $k(\alpha,\beta)$ 的和. 且在线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 后, 则双线性函数 $f(\alpha,\beta)$, $s(\alpha,\beta)$ 和 $k(\alpha,\beta)$ 分别对应于 n 阶方阵 A, $S=\frac{1}{2}(A+A')$ 和 $K=\frac{1}{2}(A-A')$.

证 记

$$s(\alpha,\beta)=\frac{1}{2}f(\alpha,\beta)+\frac{1}{2}f(\beta,\alpha),\quad k(\alpha,\beta)=\frac{1}{2}f(\alpha,\beta)-\frac{1}{2}f(\beta,\alpha). \tag{9.1.10}$$

不难证明 $s(\alpha,\beta)$ 是对称双线性函数, $k(\alpha,\beta)$ 是反对称双线性函数. 显然 $f(\alpha,\beta)=s(\alpha,\beta)+k(\alpha,\beta)$. 在线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 设 $f(\alpha,\beta)$ 对应的 n 阶方阵是 A, 则 $s(\alpha,\beta)$ 和 $k(\alpha,\beta)$ 分别对应的 n 阶方阵是 $\frac{1}{2}(A+A')=S$ 及 $\frac{1}{2}(A-A')=K$.

余下证分解的唯一性. 若 $f(\alpha,\beta)=s_1(\alpha,\beta)+k_1(\alpha,\beta)$, 其中 $s_1(\alpha,\beta)$ 是对称双线性函数, $k_1(\alpha,\beta)$ 是反对称双线性函数, 它们分别对应 n 阶方阵 A, S_1 和 K_1. 显然 S_1 和 K_1 分别是对称及反对称方阵, 又 $A=S_1+K_1$. 由引理 9.1.9, 所以 $S_1=S$, $K_1=K$. 这证明了 $s_1(\alpha,\beta)=s(\alpha,\beta)$, $k_1(\alpha,\beta)=k(\alpha,\beta)$, 即分解唯一. □

定义 9.1.11 设 $f(\alpha,\beta)$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数, 则

$$\varphi(\alpha)=f(\alpha,\alpha),\quad \forall\,\alpha\in\mathfrak{L} \tag{9.1.11}$$

是线性空间 $\mathfrak{L}$ 的函数, 称为**二次型**.

定理 9.1.12 在域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 则二次型 $\varphi(\alpha)$ 有坐标表达式

$$\varphi(\alpha)=x'Sx, \tag{9.1.12}$$

其中 S 是域 $\mathbb{F}$ 上的 n 阶对称方阵, 且上述表达式唯一, 称为二次型 $\varphi(\alpha)$ 的**方阵表示**.

证 今 $\alpha=\sum\limits_{j=1}^{n}x_j\alpha_j$, $\beta=\sum\limits_{j=1}^{n}y_j\alpha_j$, 则 $f(\alpha,\beta)=x'Ay$, 其中 A 是域 $\mathbb{F}$ 上的 n 阶

方阵. 于是

$$\varphi(\alpha)=f(\alpha,\alpha)=x'Ax=\frac{1}{2}(x'Ax+x'A'x)=x'(\frac{1}{2}(A+A'))x=x'Sx,$$

其中 S 是 n 阶方阵 A 的对称部分.

下面证唯一性. 今若 $x'Sx=x'S_1x$, 其中 S 和 S_1 都是域 $\mathbb{F}$ 上的 n 阶对称方阵, 于是 $x'(S-S_1)x=0$, 其中 $S-S_1$ 仍是对称方阵. 比较 $x_1, x_2, \cdots, x_n$ 的各项系数, 便证明了 $S-S_1=0$. □

上面的定理说明域 $\mathbb{F}$ 上的二次型是 n 个未知数 $x_1, x_2, \cdots, x_n$ 的 n 元二次齐次多项式, 而且只由域 $\mathbb{F}$ 上 n 阶对称方阵 S 唯一决定.

为了进一步展开二次型理论, 下面分别对域 $\mathbb{F}$ 是复数域和实数域进行讨论.

定义 9.1.13　设 $\varphi(\alpha)$ 是 n 维实线性空间 $\mathfrak{L}$ 的二次型. $\varphi(\alpha)$ 分别称为**正定的**、**半正定的**、**负定的**、**半负定的**, 如果分别有 $\varphi(\alpha)>0$, $\varphi(\alpha)\geqslant 0$, $\varphi(\alpha)<0$, $\varphi(\alpha)\leqslant 0$, $\forall\,\alpha\neq 0$, $\alpha\in\mathfrak{L}$.

设 S 是 n 阶实对称方阵. S 分别称为**正定的**、**半正定的**、**负定的**、**半负定的**, 如果由 n 阶实对称方阵 S 决定的二次型 $\varphi(\alpha)=x'Sx$ 分别是正定的、半正定的、负定的、半负定的. 正定实对称方阵 S 记为 $S>0$, 半正定实对称方阵 S 记为 $S\geqslant 0$, 负定实对称方阵 S 记为 $S<0$, 半负定实对称方阵 S 记为 $S\leqslant 0$.

注意　由于当 S 是 n 阶复对称方阵时, x 是 $n\times 1$ 复矩阵, 则 $x'Sx$ 仍是复数. 所以定义 9.1.13 失效. 因此, 在复的情形, 我们必须从下面角度推广. 即利用复数有共轭的概念, 我们来引进双线性函数的推广. 和引理 9.1.9 一样证明, 我们有

引理 9.1.14　设 A 是 n 阶复方阵, 则 $\frac{1}{2}(A+A^*)=H$, $\frac{1}{2}(A-A^*)=K$ 分别是 Hermite 方阵及反 Hermite 方阵, 且 $A=H+K$, 即 A 能分解为 Hermite 方阵 H (称为方阵 A 的**Hermite 部分**) 和反 Hermite 方阵 K (称为方阵 A 的**反 Hermite 部分**) 的和. 且上述分解唯一.

定义 9.1.15　n 维复线性空间 $\mathfrak{L}$ 上复值函数 $h(\alpha,\beta)$ 称为**Hermite 双线性函数**, 如果它适合条件:

(1)

$$h(c\alpha+d\beta,\gamma)=ch(\alpha,\gamma)+dh(\beta,\gamma);\tag{9.1.13}$$

(2)

$$h(\alpha,c\beta+d\gamma)=\bar{c}h(\alpha,\beta)+\bar{d}h(\alpha,\gamma),\tag{9.1.14}$$

其中 c, d 是复数, α, β, $\gamma\in\mathfrak{L}$.

和定理 9.1.8 一样证明, 我们有

定理 9.1.16　在 n 维复线性空间 $\mathfrak{L}$ 中任取一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 给定 Hermite

双线性函数 $h(\alpha,\beta)$, 则

$$h(\alpha,\beta)=x'A\overline{y}, \tag{9.1.15}$$

其中 $\alpha=\sum\limits_{j=1}^{n}x_j\alpha_j$, $\beta=\sum\limits_{k=1}^{n}y_k\alpha_k$, 又

$$x=(x_1,x_2,\cdots,x_n)',\quad y=(y_1,y_2,\cdots,y_n)' \tag{9.1.16}$$

是 $n\times 1$ 复矩阵, 而 A 是 n 阶复方阵:

$$A=\begin{pmatrix} h(\alpha_1,\alpha_1) & \cdots & h(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ h(\alpha_n,\alpha_1) & \cdots & h(\alpha_n,\alpha_n) \end{pmatrix}. \tag{9.1.17}$$

反之, 任给 n 阶复方阵 A, 则

$$h(\sum_{j=1}^{n}x_j\alpha_j,\sum_{k=1}^{n}y_k\alpha_k)=x'A\overline{y} \tag{9.1.18}$$

定义了 $\mathfrak{L}$ 上 Hermite 双线性函数, 所以 Hermite 双线性函数和 n 阶复方阵间有如下的到上的一一对应: $h\to A$.

和定理 9.1.3 一样证明, 我们有

定理 9.1.17 在 n 维复线性空间 $\mathfrak{L}$ 中任取两组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$. 记基变换对应的 n 阶非异方阵是 $P=\overline{(p_{ij})}$. 设 h 是复线性空间 $\mathfrak{L}$ 上 Hermite 双线性函数, 它在基 α_1, α_2, $\cdots$, α_n 和 β_1, β_2, $\cdots$, β_n 下对应的方阵分别是 n 阶复方阵 A 和 B, 则有

$$B=P^*AP. \tag{9.1.19}$$

定义 9.1.18 n 阶复方阵 A 和 B 称为**复相合的**, 如果存在 n 阶非异复方阵 P, 使得 $B=P^*AP$.

由此定义, 上面定理可以改写为

定理 9.1.19 n 维复线性空间上的 Hermite 双线性函数在不同基下对应的方阵是复相合的.

和定理 9.1.5 一样证明, 我们有

定理 9.1.20 复方阵的复相合关系是等价关系.

下面考虑两类特殊的 Hermite 双线性函数.

定义 9.1.21 n 维复线性空间 $\mathfrak{L}$ 上 Hermite 双线性函数 $f(\alpha,\beta)$ 称为**Hermite 函数**, 如果有

$$f(\alpha,\beta)=\overline{f(\beta,\alpha)},\quad \forall\,\alpha,\beta\in\mathfrak{L}; \tag{9.1.20}$$

称为**反 Hermite 函数**, 如果有

$$f(\alpha,\beta)=-\overline{f(\beta_1,\alpha)},\quad \forall\,\alpha,\beta\in\mathfrak{L}. \tag{9.1.21}$$

显然有

定理 9.1.22 在 n 维复线性空间 $\mathfrak{L}$ 中任取一组基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 则 Hermite 函数对应的方阵是 n 阶 Hermite 方阵, 反 Hermite 函数对应的方阵是 n 阶反 Hermite 方阵.

由定理 9.1.22 和引理 9.1.14, 有

定理 9.1.23 n 维复线性空间 $\mathfrak{L}$ 上 Hermite 双线性函数能唯一地表成一个 Hermite 函数和一个反 Hermite 函数的和. 换句话说, 任一 n 阶复方阵能唯一地表成一个 Hermite 方阵和一个反 Hermite 方阵的和.

定义 9.1.24 设 $h(\alpha,\beta)$ 是 n 维复线性空间 $\mathfrak{L}$ 上 Hermite 函数, 则

$$\varphi(\alpha)=h(\alpha,\alpha),\quad \forall\,\alpha\in\mathfrak{L} \tag{9.1.22}$$

是 $\mathfrak{L}$ 上的函数, 称为 $\mathfrak{L}$ 上 **Hermite 型**. 它是复线性空间 $\mathfrak{L}$ 上实值函数. 设分别有 $\varphi(\alpha)>0$、$\varphi(\alpha)\geqslant 0$、$\varphi(\alpha)<0$、$\varphi(\alpha)\leqslant 0$, $\forall\,\alpha\in\mathfrak{L}$, $\alpha\neq 0$, 则 $\varphi(\alpha)$ 分别称为复线性空间 $\mathfrak{L}$ 上**正定的**、**半正定的**、**负定的**、**半负定的** Hermite 型.

和定理 9.1.12 一样证明, 我们有

定理 9.1.25 在 n 维复线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 则 Hermite 型 $\varphi(\alpha)$ 有坐标表达式

$$\varphi(\alpha)=x'H\overline{x}, \tag{9.1.23}$$

其中 $\alpha=\sum\limits_{j=1}^{n}x_j\alpha_j$, 又 H 是 Hermite 方阵. 且上述表达式唯一, 即 Hermite 型 φ 和 Hermite 方阵 H 间在式 (9.1.23) 下是到上的一一对应.

定义 9.1.26 n 阶 Hermite 方阵 H 分别称为**正定的**、**半正定的**、**负定的**、**半负定的**, 如果它唯一决定的 Hermite 型 $x'H\overline{x}$ 分别是正定的、半正定的、负定的、半负定的, 即对一切 $0\neq x\in\mathbb{C}^n$, 分别有 $x'H\overline{x}>0$, $x'H\overline{x}\geqslant 0$, $x'H\overline{x}<0$, $x'H\overline{x}\leqslant 0$. 这时分别记作 $H>0$, $H\geqslant 0$, $H<0$, $H\leqslant 0$.

习　题　9.1

9.1.1 试证: 实线性空间上线性函数的平方是半正定二次型; 实线性空间上有限个线性函数的平方和也是半正定二次型. 利用这一性质, 试证下列方阵是正定对称方阵:

$$\begin{pmatrix} n & -1 & \cdots & -1 \\ -1 & n & \cdots & -1 \\ \vdots & \vdots & & \vdots \\ -1 & -1 & \cdots & n \end{pmatrix},\quad \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix},\quad \begin{pmatrix} 1 & \pm\frac{1}{2} & & & \\ \pm\frac{1}{2} & 1 & \pm\frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \pm\frac{1}{2} & 1 & \pm\frac{1}{2} \\ & & & \pm\frac{1}{2} & 1 \end{pmatrix}.$$

9.1.2 试证: 实线性空间 $\mathfrak{L}$ 上二次型 $\varphi(x)=x'Sx$ 在限制条件 $x'x=1$ 下的极值是对称方阵 S 的极大特征根和极小特征根. 且证达到极值; 再, 达到极大值的向量是对称方阵 S 的极大特征根所属的单位特征向量, 达到极小值的向量是对称方阵 S 的极小特征根所属的单位特征向量.

9.1.3 设 $f(\alpha,\beta)$ 是实线性空间 $\mathfrak{L}$ 的双线性函数, 记

$$\ker(f)=\{\,\alpha\in\mathfrak{L}\mid f(\alpha,\beta)=0,\ \forall\,\beta\in\mathfrak{L}\,\}.$$

试证: $\ker(f)$ 是 $\mathfrak{L}$ 的子空间, 称为双线性函数 f 的**核**. 当 $\ker(f)=0$, 则称 $\mathfrak{L}$ 的双线性函数 f 为**非退化的**. 试证: f 是非退化的当且仅当它对应的方阵是非异的. 在实线性空间 $\mathfrak{L}$ 中任取向量 α 和 β, 如果 $f(\alpha,\beta)=0$, 则称 α 和 β 关于双线性函数 f 是**正交的**. 试证: 在实线性空间 $\mathfrak{L}$ 中任给子空间 $\mathfrak{L}_1$, 则

$$\mathfrak{L}_1^{\perp}=\{\,\alpha\in\mathfrak{L}\mid f(\alpha,\beta)=0,\ \forall\,\beta\in\mathfrak{L}_1\,\}$$

是实线性空间 $\mathfrak{L}$ 的子空间, 称为 $\mathfrak{L}_1$ 关于 f 的**正交补**.

9.1.4 设 $h(\alpha,\beta)$ 是复线性空间 $\mathfrak{L}$ 上 Hermite 双线性函数, 和上面第 9.1.3 题一样, 可定义核、非退化、正交补. 试证: 上述性质也都成立.

9.1.5 记 $f(\alpha,\beta)$ 是实线性空间 $\mathfrak{L}$ 上对称双线性函数. 记

$$\mathfrak{N}_0=\{\,\alpha\in\mathfrak{L}\mid f(\alpha,\alpha)=0\,\},$$

则 $\mathfrak{N}_0$ 中子空间称为**迷向子空间**.

(i) 试证: 实线性空间 $\mathfrak{L}$ 中的子空间 $\mathfrak{M}$ 是迷向子空间当且仅当 $\mathfrak{M}\subset\mathfrak{M}^{\perp}$;

(ii) 设 $\mathfrak{M}_1$ 和 $\mathfrak{M}_2$ 是实线性空间 $\mathfrak{L}$ 中两迷向子空间, $\mathfrak{P}_1$ 和 $\mathfrak{P}_2$ 是实线性空间 $\mathfrak{L}$ 中子空间. 设 $\mathfrak{M}_i$ 有子空间直接和分解:

$$\mathfrak{M}_i=\mathfrak{P}_i\oplus(\mathfrak{M}_1\cap\mathfrak{M}_2),\quad i=1,2.$$

试证: $\mathfrak{M}_1\cap\mathfrak{M}_2^{\perp}$ 有子空间直接和分解

$$\mathfrak{M}_1\cap\mathfrak{M}_2^{\perp}=(\mathfrak{M}_1\cap\mathfrak{M}_2)\oplus(\mathfrak{P}_1\cap\mathfrak{P}_2^{\perp}).$$

§9.2 实 Euclid 空间

现在将解析几何中的直角坐标系推广到 n 维实线性空间.

设 $\mathfrak{L}$ 是 n 维实线性空间, 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 于是 $\mathfrak{L}$ 中任一向量 α 可表为 $\alpha = \sum\limits_{j=1}^{n} a_j\alpha_j$, 其中 $n\times 1$ 矩阵 $(a_1, a_2, \cdots, a_n)'$ 是在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标, 这里, $(*)'$ 表示矩阵 $(*)$ 的转置矩阵. 在这一节, 始终在取定了基后进行讨论, 有时仍旧用符号 α 记这个矩阵.

显然有

引理 9.2.1 设 $\mathfrak{L}$ 是 n 维实线性空间, 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 设 $\mathfrak{L}$ 中向量 α 和 β 在这组基下的坐标分别是 $(a_1, a_2, \cdots, a_n)'$ 和 $(b_1, b_2, \cdots, b_n)'$, 则

$$(\alpha, \beta) = \sum_{j=1}^{n} a_j b_j, \quad \forall\, \alpha, \beta \in \mathfrak{L} \tag{9.2.1}$$

定义了 n 维实线性空间上一个函数. 它有下面一些性质:

(1) (α, β) 是一个实正定对称双线性函数. 即 $(a\alpha + b\beta, \gamma) = a(\alpha, \gamma) + b(\beta, \gamma)$, $\forall\, a, b \in \mathbb{R}$, $\alpha, \beta, \gamma \in \mathfrak{L}$, 又 $(\alpha, \beta) = (\beta, \alpha)$, (α, α) 是非负实数. 而且 $(\alpha, \alpha) = 0$ 当且仅当 $\alpha = 0$;

(2) 利用归纳法和性质**(1)** 可知, 对任意 $p+q$ 个向量 $\beta_1, \beta_2, \cdots, \beta_p, \gamma_1, \gamma_2, \cdots, \gamma_q$ 及 $p+q$ 个实数 $b_1, b_2, \cdots, b_p, c_1, c_2, \cdots, c_q$, 有性质:

$$\left(\sum_{j=1}^{p} b_j\beta_j, \sum_{k=1}^{q} c_k\gamma_k\right) = \sum_{j=1}^{p}\sum_{k=1}^{q} b_j c_k (\beta_j, \gamma_k). \tag{9.2.2}$$

定义 9.2.2 n 维实线性空间 $\mathfrak{L}$ 上的实正定对称双线性函数 $f(\alpha, \beta)$ 称为向量 α 和 β 在 $\mathfrak{L}$ 上的**内积**. 取定内积 $f(\alpha, \beta)$ 后, 我们改记为 (α, β).

定义 9.2.3 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积 (α, β), 则 n 维实线性空间 $\mathfrak{L}$ 称为关于内积 (α, β) 的**实 Euclid 空间**或称为**实欧氏空间**. 也称为关于内积 (α, β) 的**实内积空间**.

引理 9.2.4 设 $\mathfrak{L}$ 是 n 维实线性空间, 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则在 $\mathfrak{L}$ 中存在和这组基有关的内积 (α, β), 它由式 (9.2.1) 定义.

证 由引理 9.2.1 立即可知. □

利用向量的内积, 可以定义向量的长度和两向量的夹角.

定义 9.2.5 设 $\mathfrak{L}$ 是 n 维实线性空间, 在 $\mathfrak{L}$ 中取定内积 (α, β). 非负实数

$$|\alpha| = \sqrt{(\alpha, \alpha)} \tag{9.2.3}$$

称为实线性空间 $\mathfrak{L}$ 中向量 α 的**长度**. 长度为 1 的向量称为**单位向量**.

由定义可知, 对线性空间 $\mathfrak{L}$ 中任一非零向量 α, 则

$$\beta = \frac{1}{|\alpha|}\alpha \tag{9.2.4}$$

是单位向量.

在定义夹角之前, 先证明一个有用的不等式.

引理 9.2.6 (Cauchy 不等式) 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积 (α,β), 则有

$$(\alpha,\beta)^2 \leqslant (\alpha,\alpha)(\beta,\beta), \quad \forall\, \alpha,\beta\in\mathfrak{L}, \tag{9.2.5}$$

且等式成立当且仅当 α 和 β 线性相关.

证 引进实参数 t, 则有 $\alpha+t\beta\in\mathfrak{L}$. 于是

$$(\alpha+t\beta,\alpha+t\beta)=(\alpha,\alpha)+2t(\alpha,\beta)+t^2(\beta,\beta)\geqslant 0, \quad \forall\, t\in\mathbb{R}.$$

当 $\beta=0$ 时, 不等式 $(\alpha,\beta)^2 \leqslant (\alpha,\alpha)(\beta,\beta)$ 自然成立. 当 $\beta\neq 0$ 时, 有 $(\beta,\beta)>0$. 视右边是 t 的二次多项式, 我们知道, 自变量 t 取一切实数值时, 若二次多项式不取负值, 则这个二次多项式的判别式 Δ 小于或等于零, 反之亦然. 这个条件写成向量形式, 即有

$$(\alpha,\beta)^2-(\alpha,\alpha)(\beta,\beta)\leqslant 0.$$

此即 Cauchy 不等式成立.

又等号成立当且仅当二次多项式的判别式等于零, 即它有重根. 所以等号成立的必要且充分条件是二次多项式有实根 t_0, 即 $(\alpha+t_0\beta,\alpha+t_0\beta)=0$. 由于内积是实正定对称双线性函数, 所以有 $\alpha+t_0\beta=0$, 此即 α 和 β 线性相关. □

引理 9.2.7 (Cauchy 不等式) 任取两组实数 $a_1, a_2, \cdots, a_n$ 和 $b_1, b_2, \cdots, b_n$, 则有

$$\Big(\sum_{j=1}^{n} a_jb_j\Big)^2 \leqslant \Big(\sum_{j=1}^{n} a_j^2\Big)\Big(\sum_{j=1}^{n} b_j^2\Big). \tag{9.2.6}$$

上式中等号成立的必要且充分条件是 $a_1, a_2, \cdots, a_n$ 和 $b_1, b_2, \cdots, b_n$ 成比例.

证 由引理 9.2.1 和引理 9.2.6 立即可知. □

由 Cauchy 不等式可知, 在 α 和 β 都不是零向量时有不等式

$$\left|\frac{(\alpha,\beta)}{|\alpha||\beta|}\right| \leqslant 1. \tag{9.2.7}$$

所以有

定义 9.2.8 在 n 维实线性空间 $\mathfrak{L}$ 上取定内积 (α,β). 当 α 或 β 是零向量时, 我们约定向量 α 和 β 的**夹角** 任意. 当 $\alpha\neq 0, \beta\neq 0$ 时, 数值 θ 记作 $\theta=<\alpha,\beta>$, 它定义为

$$\theta=<\alpha,\beta>=\cos^{-1}\frac{(\alpha,\beta)}{|\alpha||\beta|}, \tag{9.2.8}$$

称为向量 α 和 β 的**夹角**. 这里约定 $\cos^{-1}$ 取**主值**, 即有 $0 \leqslant \theta \leqslant \pi$. 夹角是直角时, 即当 $(\alpha, \beta) = 0$ 时, 我们称向量 α 和 β **互相正交**, 记作 $\alpha \perp \beta$. 显然, 只有零向量才和任何向量正交. 夹角 θ 也称为向量 α 和 β 的**角度**.

由定义可知, 不论在什么情形, 恒有关系

$$(\alpha, \beta) = \sqrt{(\alpha, \alpha)}\sqrt{(\beta, \beta)} \cos < \alpha, \beta >= |\alpha||\beta| \cos < \alpha, \beta > . \tag{9.2.9}$$

上面定义了长度和角度. 在 $n = 2$ 和 $n = 3$ 的情形, 即在平面和普通空间的情形, 在其中取定直角坐标系后, 内积的定义及其长度和角度的关系和这里的定义完全一样. 这便给出上面的定义的一种直观的解释. 但是, 这里的情况有些不一样. 例如在平面中取定笛卡尔坐标系 (不是直角坐标系), 用这一节的方式同样可以定义内积, 定义长度和夹角, 这时就和通常意义下的长度和夹角不一样了. 所以上面的定义不仅仅是一种形式推广.

定义 9.2.9 设 $\mathfrak{L}$ 是 n 维实 Euclid 空间, 内积是 (α, β). 实 Euclid 空间 $\mathfrak{L}$ 中的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 称为**标准正交基**, 如果

$$(\alpha_i, \alpha_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n, \tag{9.2.10}$$

即 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 个两两正交的单位向量.

在标准正交基的定义中, 要求这组基由一组两两正交的单位向量组成. 实际上, 后一条件保证了它们是一组基. 其实可以证明

引理 9.2.10 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任意 s 个两两正交的非零向量 $\beta_1, \beta_2, \cdots, \beta_s$ 必线性无关.

证 今若存在 s 个实数 $c_1, c_2, \cdots, c_s$, 使得 $\sum\limits_{j=1}^{s} c_j\beta_j = 0$, 则

$$0 = (0, \beta_k) = (\sum_{j=1}^{s} c_j\beta_j, \beta_k) = \sum_{j=1}^{s} c_j(\beta_j, \beta_k) = \sum_{j=1}^{s} c_j\delta_{jk}|\beta_j|^2 = c_k|\beta_k|^2,$$

其中 $1 \leqslant k \leqslant s$. 所以 $c_1 = \cdots = c_s = 0$, 即 $\beta_1, \beta_2, \cdots, \beta_s$ 线性无关. □

定义 9.2.11 在 n 维实向量空间 $\mathbb{R}^n$ 中取定一组基

$$e_j = (\delta_{j1}, \delta_{j2}, \cdots .\delta_{jn})', \quad j = 1, 2, \cdots, n, \tag{9.2.11}$$

其中 δ_{jk} 是**Kronecker 符号**, 即 $\delta_{jj} = 1$, $\delta_{jk} = 0$, $j \neq k$. 可以引进**标准内积** (α, β), 它定义为:

$$(e_i, e_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n. \tag{9.2.12}$$

换句话说, $e_1, e_2, \cdots, e_n$ 是两两正交的单位向量. 于是 $\mathbb{R}^n$ 是 n 维**实 Euclid 空间**.

今后凡是提到 $\mathbb{R}^n$, 即所有 $n\times 1$ 矩阵构成的 n 维实线性空间, 作为 n 维实 Euclid 空间, 我们理解为标准内积 (α,β), 它由定义 9.2.11 给出, 而不再加以说明了. 由定义 9.2.9 还可知, 在 n 维实 Euclid 空间 $\mathbb{R}^n$ 中事先取定的一组基 $e_1, e_2, \cdots, e_n$ 就是一组标准正交基, 它是直角坐标系的推广. 当然, 并不是 $\mathbb{R}^n$ 中所有的基都是标准正交基.

由定义 9.2.9 又可知, 标准正交基的特点是:

定理 9.2.12 设 $\mathfrak{L}$ 是关于内积 (α,β) 的 n 维实 Euclid 空间. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是实 Euclid 空间 $\mathfrak{L}$ 的标准正交基. 记 $\alpha=\sum\limits_{j=1}^{n}x_j\alpha_j$, $\beta=\sum\limits_{j=1}^{n}y_j\alpha_j$, 则有

$$(\alpha,\beta)=\sum_{j=1}^{n}x_jy_j. \tag{9.2.13}$$

即在标准正交基下, 内积 (α,β) 的方阵表示是单位方阵 E_n.

证 由 $(\alpha_i,\alpha_j)=\delta_{ij}$, $1\leqslant i,j\leqslant n$, 于是

$$(\alpha,\beta)=(\sum_{i=1}^{n}x_i\alpha_i,\sum_{j=1}^{n}y_j\alpha_j)=\sum_{i,j=1}^{n}x_iy_j(\alpha_i,\alpha_j)=\sum_{j=1}^{n}x_jy_j. \qquad \square$$

定义 9.2.13 n 维实 Euclid 空间 $\mathfrak{L}$ 中保持向量长度不变的线性变换称为**正交变换**.

引理 9.2.14 n 维实 Euclid 空间 $\mathfrak{L}$ 中线性变换 $\mathscr{A}$ 是正交变换当且仅当它保持向量内积不变.

证 显然, 线性变换 $\mathscr{A}$ 保持向量内积不变, 则它保持向量长度不变, 所以线性变换 $\mathscr{A}$ 是正交变换. 反之, 设线性变换 $\mathscr{A}$ 是正交变换, 则 $|\mathscr{A}(\alpha)|=\alpha$, $\forall\,\alpha\in\mathfrak{L}$. 于是 $(\mathscr{A}(\alpha+t\beta),\mathscr{A}(\alpha+t\beta))=(\alpha+t\beta,\alpha+t\beta)$, $\quad\forall\,t\in\mathbb{R}$. 因此,

$$t^2(\mathscr{A}(\beta),\mathscr{A}(\beta))+2t(\mathscr{A}(\alpha),\mathscr{A}(\beta))+(\mathscr{A}(\alpha),\mathscr{A}(\alpha))=t^2(\beta,\beta)+2t(\alpha,\beta)+(\alpha,\alpha),$$

$\forall\,t\in\mathbb{R}$. 所以

$$(\mathscr{A}(\alpha),\mathscr{A}(\beta))=(\alpha,\beta),\quad \forall\,\alpha,\beta\in\mathfrak{L}.$$

这证明了线性变换 $\mathscr{A}$ 保持向量内积不变. $\square$

定理 9.2.15 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 线性变换 $\mathscr{A}$ 在这组基下的方阵表示记为 A, 则线性变换 $\mathscr{A}$ 是正交变换的必要且充分条件是

$$A'A=E_n, \tag{9.2.14}$$

其中 E_n 是 n 阶单位方阵.

证 今任取向量 α, 它在线性变换 $\mathscr{A}$ 下的象是向量 $\beta = \mathscr{A}(\alpha)$. 记向量 α 的坐标是 $x = (x_1, x_2, \cdots, x_n)'$, 则象 β 的坐标是 $y = A(x_1, x_2, \cdots, x_n)'$. 于是 α 的长度是 $|x| = \sqrt{x'x}$. 向量 β 的长度是

$$|y| = \sqrt{(Ax)'(Ax)} = \sqrt{x'A'Ax}.$$

先证充分性. 今若 $A'A = E_n$, 则 $|y| = \sqrt{x'x} = |x|$, 所以线性变换 $\mathscr{A}$ 保持向量长度不变. 再证必要性. 今若线性变换 $\mathscr{A}$ 保持向量长度不变, 即方阵 A 有关系

$$y'y = (Ax)'(Ax) = x'A'Ax = x'x,$$

对一切向量 $x \in \mathbb{R}^n$ 成立. 记 $S = E - A'A$, 显然 $S' = S$. 再记 $S = (s_{ij})$, 则 $s_{ij} = s_{ji}$. 又上式可改写为

$$0 = x'(E - A'A)x = x'Sx = \sum_{i=1}^{n} s_{ii}x_i^2 + 2\sum_{1 \leqslant i < j \leqslant n} s_{ij}x_i x_j.$$

这是恒等式, 所以必须有 $s_{ij} = 0, 1 \leqslant i \leqslant j \leqslant n$. 因此 $S = 0$, 亦即 $A'A = E$. □

定理 9.2.16 n 维实 Euclid 空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 是正交变换的必要且充分条件是它将标准正交基变为标准正交基.

证 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 经过线性变换 $\mathscr{A}$ 的作用, 变成一组向量 $\beta_1 = \mathscr{A}(\alpha_1), \beta_2 = \mathscr{A}(\alpha_2), \cdots, \beta_n = \mathscr{A}(\alpha_n)$, 于是 $\beta_j = \sum\limits_{k=1}^{n} a_{kj}\alpha_k, 1 \leqslant j \leqslant n$.

先证必要性. 假设 $\mathscr{A}$ 是正交变换, 由引理 9.2.14, 线性变换 $\mathscr{A}$ 保持向量内积不变. 因此

$$(\beta_i, \beta_j) = (\mathscr{A}(\alpha_i), \mathscr{A}(\alpha_j)) = (\alpha_i, \alpha_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n,$$

由定义 9.2.10, $\beta_1, \beta_2, \cdots, \beta_n$ 是一组标准正交基.

再证充分性. 今已知 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是一组标准正交基, 假设 $\beta_1, \beta_2, \cdots, \beta_n$ 也是一组标准正交基, 下面来证明 $\mathscr{A}$ 是正交变换, 即证明 $\mathscr{A}$ 将实线性空间 $\mathfrak{L}$ 中任一向量的长度保持不变. 事实上, 任取 $\alpha \in \mathfrak{L}$, 由于 $\alpha = \sum\limits_{j=1}^{n} a_j\alpha_j$, 则 $\mathscr{A}(\alpha) = \sum\limits_{j=1}^{n} a_j\beta_j$, 所以

$$\begin{aligned}(\mathscr{A}(\alpha), \mathscr{A}(\alpha)) &= (\sum_{j=1}^{n} a_j\beta_j, \sum_{k=1}^{n} a_k\beta_k) = \sum_{j,k=1}^{n} a_j a_k(\beta_j, \beta_k) \\ &= \sum_{j,k=1}^{n} a_j a_k \delta_{jk} = \sum_{j,k=1}^{n} a_j a_k(\alpha_j, \alpha_k) \\ &= (\sum_{j=1}^{n} a_j\alpha_j, \sum_{k=1}^{n} a_k\alpha_k) = (\alpha, \alpha).\end{aligned}$$

这就证明了 $\mathscr{A}$ 是正交变换. □

定理 9.2.17 设 $\mathfrak{L}$ 是 n 维实 Euclid 空间, 内积是 (α,β). 任取标准正交基 α_1, α_2, $\cdots$, α_n. 再在 $\mathfrak{L}$ 中任取一组基 β_1, β_2, $\cdots$, β_n, 使得基变换公式是

$$\beta_i=\sum_{j=1}^{n}a_{ji}\alpha_j,\quad i=1,2,\cdots,n, \tag{9.2.15}$$

它对应的 n 阶非异实方阵是

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots & &\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}. \tag{9.2.16}$$

则 β_1, β_2, $\cdots$, β_n 是标准正交基当且仅当 n 阶非异实方阵 A 适合条件 $A'A=E_n$. 所以在 n 维实 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基后, 适合条件 $A'A=E_n$ 的 n 阶实方阵 A 和 $\mathfrak{L}$ 中标准正交基间便有一个自然的一一对应关系.

证 令 $\mathfrak{L}$ 中基 β_1, β_2, $\cdots$, β_n 是标准正交基, 即 $(\beta_i,\beta_j)=\delta_{ij}$, $i,j=1,2,\cdots,n$. 此即

$$\delta_{ij}=(\sum_{k=1}^{n}a_{ki}\alpha_k,\sum_{p=1}^{n}a_{pj}\alpha_p)=\sum_{k,p=1}^{n}a_{ki}a_{pj}(\alpha_k,\alpha_p).$$

由于已知 α_1, α_2, $\cdots$, α_n 是标准正交基, 即有 $(\alpha_k,\alpha_p)=\delta_{kp}$, $1\leqslant k,p\leqslant n$, 所以证明了 β_1, β_2, $\cdots$, β_n 是标准正交基当且仅当

$$\sum_{k=1}^{n}a_{ki}a_{kj}=\delta_{ij},\quad 1\leqslant i,j\leqslant n.$$

此即 $A'A=E_n$. □

上面初步讨论了正交变换. 为了进一步考虑正交变换的性质, 需要对正交变换在标准正交基下的方阵表示进行详细讨论.

定义 9.2.18 n 阶实方阵 O 称为**实正交方阵**, 如果

$$O'O=E_n. \tag{9.2.17}$$

所有 n 阶实正交方阵构成的集合记作 $\mathrm{O}(n,\mathbb{R})$, 或简记为 $\mathrm{O}(n)$, 称为 n 阶**实正交群**.

由定义可知, 实正交方阵有性质:

引理 9.2.19 n 阶实正交群 $\mathrm{O}(n)$ 有性质

(1) 任取 $O_1, O_2\in\mathrm{O}(n)$, 则 $O_1O_2\in\mathrm{O}(n)$;

(2) 实正交方阵的乘法适合结合律;

(3) n 阶单位方阵 $E_n \in \mathrm{O}(n)$;

(4) 任取 $O \in \mathrm{O}(n)$, 则 $O^{-1} = O' \in \mathrm{O}(n)$, 所以 $O'O = OO' = E_n$;

(5) 实正交方阵的行列式等于 1 或 -1.

证 今由 $O_1, O_2 \in \mathrm{O}(n)$, 即有 $O_1O_1' = O_1'O_1 = E_n$, $O_2O_2' = O_2'O_2 = E_n$. 于是

$$(O_1O_2)'(O_1O_2) = O_2'O_1'O_1O_2 = O_2'O_2 = E_n.$$

这证明了 $O_1O_2 \in \mathrm{O}(n)$. 再任取 $O \in \mathrm{O}(n)$, 于是 $O^{-1} = O'$, 因此 $OO' = OO^{-1} = E_n$, 所以 $O'O = OO' = E_n$. 又 $O' \in \mathrm{O}(n)$. 事实上, $(O')'O' = OO' = E_n$. 所以 $O' \in \mathrm{O}(n)$, 因此 $O^{-1} \in \mathrm{O}(n)$.

再者, 由实正交方阵定义可知 n 阶单位方阵 $E_n \in \mathrm{O}(n)$. 最后, 由 $O'O = E_n$, 双方取行列式, 所以有 $1 = \det(O'O) = \det(O')\det(O) = (\det O)^2$, 即有 $\det O = \pm 1$. □

在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 对 $\mathfrak{L}$ 中任两向量 α 和 β, 设它们在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标分别是 $(a_1, a_2, \cdots, a_n)'$ 和 $(b_1, b_2, \cdots, b_n)'$, 则有

$$(\alpha, \beta) = (\sum_{j=1}^{n} a_j\alpha_j, \sum_{k=1}^{n} b_k\alpha_k) = \sum_{j,k=1}^{n} a_jb_k(\alpha_j, \alpha_k) = \sum_{j,k=1}^{n} a_jb_k\delta_{jk} = \sum_{j=1}^{n} a_jb_j.$$

所以, 如果按照 $\mathfrak{L}$ 中标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 来定义向量 α 和 β 的内积, 那末它们的内积就是向量 α 和 β 的坐标, 作为实向量空间 $\mathbb{R}^n$ 中的向量, 按照标准正交基 $e_1, e_2, \cdots, e_n$ 所定义的标准内积. 换句话说, 我们证明了内积不依赖于标准正交基的选取; 特别, 向量的长度不依赖于标准正交基的选取.

将 n 阶实方阵 B 按照列分成 n 块:

$$B = \begin{pmatrix} \beta_1 \, \beta_2 \cdots \beta_n \end{pmatrix}, \tag{9.2.18}$$

其中 $\beta_1, \beta_2, \cdots, \beta_n$ 都是 $n \times 1$ 实矩阵. 所以 $\beta_1, \beta_2, \cdots, \beta_n$ 是 n 维实向量 $\mathbb{R}^n$ 中 n 个向量. 再由矩阵分块乘法的规则可知

$$B'B = \begin{pmatrix} \beta_1' \\ \beta_2' \\ \vdots \\ \beta_n' \end{pmatrix} \begin{pmatrix} \beta_1 \, \beta_2 \cdots \beta_n \end{pmatrix} = \begin{pmatrix} \beta_1'\beta_1 & \cdots & \beta_1'\beta_n \\ \vdots & & \vdots \\ \beta_n'\beta_1 & \cdots & \beta_n'\beta_n \end{pmatrix}.$$

所以 B 是实正交方阵的必要且充分条件是

$$\beta_i'\beta_j = \delta_{ij} = (\beta_i, \beta_j), \quad 1 \leqslant i.j \leqslant n. \tag{9.2.19}$$

即 n 阶实方阵 B 的列向量构成 $\mathbb{R}^n$ 的一组标准正交基. 所以有

定理 9.2.20 n 阶实方阵是实正交方阵的必要且充分条件是: 它的 n 个列向量构成一组标准正交基.

另一方面, 在 B 是实正交方阵的时候, B' 也是实正交方阵, 由于 B 的 n 个列向量构成一组标准正交基, 而 B' 的列是 B 的行, 所以有

定理 9.2.21 n 阶实方阵是实正交方阵的必要且充分条件是它的 n 个行向量构成一组标准正交基.

利用实正交方阵的基本性质, 可以证明正交变换有性质:

(1) 正交变换的乘积仍是正交变换;

(2) 正交变换的乘法适合结合律;

(3) 恒等变换是正交变换;

(4) 正交变换的逆变换 (等于它的共轭变换) 也是正交变换;

(5) 正交变换的行列式等于 1 或 -1.

定义 9.2.22 设 $\mathfrak{L}$ 是 n 维实 Euclid 空间, $\mathfrak{L}$ 上所有正交变换构成的集合称为**正交变换群**. 研究在所有正交变换下都不变的几何性质的学科称为**实 Euclid 几何学**.

在实 Euclid 空间 $\mathfrak{L}$ 中, 最基本的定理是给出从任取一组基, 构作标准正交基的方法, 即所谓 Schmidt 正交化.

定理 9.2.23 (Schmidt 正交化) 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取一组基 β_1, $\beta_2, \cdots, \beta_s$, 则可以按照一种确定的方法 (称为**Schmidt 正交化**) 造出一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 其中

$$\left\{\begin{aligned} \alpha_1 &= b_{11}\beta_1, \\ \alpha_2 &= b_{21}\beta_1 + b_{22}\beta_2, \\ &\ \vdots \\ \alpha_n &= b_{n1}\beta_1 + b_{n2}\beta_2 + \cdots + b_{nn}\beta_n, \end{aligned}\right. \tag{9.2.20}$$

又 $b_{11}, b_{22}, \cdots, b_{nn}$ 都是正实数.

证 今 $\beta_1 \neq 0$. 令

$$\alpha_1 = \frac{1}{|\beta_1|}\beta_1 = b_{11}\beta_1, \tag{9.2.21}$$

则 α_1 是单位向量. 又 $b_{11} > 0$. 再作向量

$$\gamma_2 = \beta_2 + \lambda_{21}\alpha_1 = \beta_2 + \frac{\lambda_{21}}{|\beta_1|}\beta_1, \tag{9.2.22}$$

由于 β_1 和 β_2 线性无关, 所以 $\gamma_2 \neq 0$. 现在来求数值 λ_{21}, 使得 γ_2 和 α_1 正交, 即使得

$$0 = (\gamma_2, \alpha_1) = (\beta_2 + \lambda_{21}\alpha_1, \alpha_1) = (\beta_2, \alpha_1) + \lambda_{21}. \tag{9.2.23}$$

所以取

$$\lambda_{21} = -(\beta_2, \alpha_1), \tag{9.2.24}$$

那末

$$\gamma_2 = \beta_2 + \lambda_{21}\alpha_1 = \beta_2 - \frac{(\beta_2\alpha_1)}{|\beta_1|}\beta_1 \tag{9.2.25}$$

便和 α_1 正交. 再由 $\gamma_2 \neq 0$, 令

$$\alpha_2 = \frac{1}{|\gamma_2|}\gamma_2 = \frac{1}{|\gamma_2|}\beta_2 - \frac{(\beta_2, \alpha_1)}{|\beta_1||\gamma_2|}\beta_1 = b_{21}\beta_1 + b_{22}\beta_2, \tag{9.2.26}$$

自然, α_1 和 α_2 都是单位向量, 且 α_2 和 α_1 正交. 又 $b_{22} > 0$. 于是按照上述程序找到了两个互相正交的单位向量 α_1 和 α_2.

现在用归纳法来证明定理. 假设按照这样的程序能够做到第 r 步, 即找到 r 个两两正交的单位向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$. 它们是

$$\begin{cases} \alpha_1 = b_{11}\beta_1, \\ \alpha_2 = b_{21}\beta_1 + b_{22}\beta_2, \\ \qquad\vdots \\ \alpha_r = b_{r1}\beta_1 + b_{r2}\beta_2 + \cdots + b_{rr}\beta_r, \end{cases} \tag{9.2.27}$$

其中 $b_{11}, b_{22}, \cdots, b_{rr}$ 都是正实数. 如果 $r = n$, 则不证自明了. 如果 $r < n$, 再作向量

$$\gamma_{r+1} = \beta_{r+1} + \lambda_{r+1,1}\alpha_1 + \cdots + \lambda_{r+1,r}\alpha_r. \tag{9.2.28}$$

今 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 只是 $\beta_1, \beta_2, \cdots, \beta_r$ 的线性组合, 所以向量 $\lambda_{r+1,1}\alpha_1 + \cdots + \lambda_{r+1,r}\alpha_r$ 也只是向量 $\beta_1, \beta_2, \cdots, \beta_r$ 的线性组合. 如果 $\gamma_{r+1} = 0$, 则 β_{r+1} 必是 $\beta_1, \beta_2, \cdots, \beta_r$ 的线性组合, 这和 $\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}$ 线性无关矛盾, 所以证明了 $\gamma_{r+1} \neq 0$. 现在再来确定数值 $\lambda_{r+1,1}, \lambda_{r+1,2}, \cdots, \lambda_{r+1,r}$, 使得 γ_{r+1} 和 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 都正交. 为此, 只要

$$\begin{aligned} 0 &= (\gamma_{r+1}, \alpha_j) = (\beta_{r+1} + \sum_{k=1}^{r}\lambda_{r+1,k}\alpha_k, \alpha_j) \\ &= (\beta_{r+1}, \alpha_j) + \sum_{k=1}^{r}\lambda_{r+1,k}(\alpha_k, \alpha_j) = (\beta_{r+1}, \alpha_j) + \sum_{k=1}^{r}\lambda_{r+1,k}\delta_{jk} \\ &= (\beta_{r+1}, \alpha_j) + \lambda_{r+1,j}, \quad 1 \leqslant j \leqslant r, \end{aligned}$$

即只要

$$\lambda_{r+1,j} = -(\beta_{r+1}, \alpha_j), \quad 1 \leqslant j \leqslant r \tag{9.2.29}$$

就可以了. 所以, 非零向量

$$\gamma_{r+1} = \beta_{r+1} - \sum_{j=1}^{r} (\beta_{r+1}, \alpha_j)\alpha_j \tag{9.2.30}$$

和向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 正交. 再令

$$\alpha_{r+1} = \frac{1}{|\gamma_{r+1}|}\gamma_{r+1} = b_{r+1,1}\beta_1 + \cdots + b_{r+1,r}\beta_r + b_{r+1,r+1}\beta_{r+1}, \tag{9.2.31}$$

于是 $b_{r+1,r+1} > 0$, 且 $\alpha_1, \alpha_2, \cdots, \alpha_{r+1}$ 是 $r+1$ 个两两正交的单位向量. 所以由归纳法, 证明了一定能够找到 n 个两两正交的单位向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得它们由式 (9.2.20) 给出. □

Schmidt 正交化有下列重要结论.

定理9.2.24 任一 n 阶非异实方阵 A 一定可以唯一地表成一个实上 (下) 三角方阵 T 和一个实正交方阵 O 的乘积,

$$A = OT, \tag{9.2.32}$$

其中 T 的对角元素都是正实数; 也一定可以唯一地表成一个实正交方阵O_1 和一个实下 (上) 三角方阵 T_1 的乘积

$$A = T_1O_1, \tag{9.2.33}$$

其中 T_1 的对角元素也都是正实数.

证 先证前一断言. 今 n 阶实方阵 A 非异, 故它的 n 个列向量 $\alpha_1 = \xi_n$, $\alpha_2 = \xi_{n-1}, \cdots, \alpha_n = \xi_1$ 线性无关, 所以构成 n 维向量空间 $\mathbb{R}^n$ 的一组基. 由定理 9.2.23 可知, 存在一组标准正交基 $\beta_1 = \eta_n, \beta_2 = \eta_{n-1}, \cdots, \beta_n = \eta_1$, 使得

$$\begin{cases} \beta_1 = b_{11}\alpha_1, \\ \beta_2 = b_{12}\alpha_1 + b_{22}\alpha_2, \\ \qquad \vdots \\ \beta_n = b_{1n}\alpha_1 + b_{2n}\alpha_2 + \cdots + b_{nn}\alpha_n, \end{cases} \tag{9.2.34}$$

其中 $b_{11} > 0, b_{22} > 0, \cdots, b_{nn} > 0$. 将 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 表成 $\beta_1, \beta_2, \cdots, \beta_n$ 的线性组合, 即有

$$\begin{cases} \alpha_1 = b_{11}^{-1}\beta_1 = t_{11}\beta_1, \\ \alpha_2 = b_{22}^{-1}(\beta_2 - b_{12}\alpha_1) = t_{12}\beta_1 + t_{22}\beta_2, \\ \qquad \vdots \\ \alpha_n = b_{nn}^{-1}(\beta_n - b_{1n}\alpha_1 + \cdots + b_{n-1,n}\alpha_{n-1}) = t_{1n}\beta_1 + t_{2n}\beta_2 + \cdots + t_{nn}\beta_n, \end{cases} \tag{9.2.35}$$

其中 $t_{jj}=b_{jj}^{-1}>0,\ 1\leqslant j\leqslant n$. 令 n 个列向量 $\beta_1,\beta_2,\cdots,\beta_n$ 构成的 n 阶方阵

$$O=\begin{pmatrix}\beta_1\ \beta_2\cdots\ \beta_n\end{pmatrix}\tag{9.2.36}$$

是正交方阵. 利用分块矩阵的乘法规则, 所以

$$A=\begin{pmatrix} t_{11}\beta_1 \\ t_{12}\beta_1+t_{22}\beta_2 \\ \vdots \\ t_{1n}\beta_1+t_{2n}\beta_2+\cdots+t_{nn}\beta_n\end{pmatrix}=O\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ & t_{22} & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & & & t_{nn}\end{pmatrix}=OT.$$

这就导出了分解式, 其中上三角方阵 T 的对角元素 $t_{11},\cdots,t_{nn}$ 恒大于零.

由 $\alpha_i=\xi_{n-i+1},\ \beta_i=\eta_{n-i+1},\ 1\leqslant i\leqslant n$, 式 (9.2.35) 可变成

$$\begin{cases}\xi_1=u_{11}\eta_1+u_{21}\eta_2+\cdots+u_{n1}\eta_n,\\ \xi_2=u_{22}\eta_2+u_{32}\eta_3+\cdots+u_{n2}\eta_n,\\ \qquad\vdots\\ \xi_n=u_{nn}\eta_n,\end{cases}\tag{9.2.35$'$}$$

其中 $u_{jj}=t_{n-j+1,n-j+1}>0,\ 1\leqslant j\leqslant n$. 令 n 个列向量$\eta_1,\eta_2,\cdots,\eta_n$ 构成的 n 阶方阵

$$O=\begin{pmatrix}\eta_1\ \eta_2\cdots\ \eta_n\end{pmatrix}\tag{9.2.36$'$}$$

是正交方阵. 利用分块矩阵的乘法规则, 所以

$$A=\begin{pmatrix} u_{11}\eta_1+u_{21}\eta_2+\cdots+u_{n1}\eta_n \\ u_{22}\eta_2+u_{32}\eta_3+\cdots+u_{n2}\eta_n \\ \vdots \\ u_{nn}\eta_n\end{pmatrix}=O\begin{pmatrix} u_{11} & & & \\ u_{21} & u_{22} & & \\ \vdots & \vdots & \ddots & \\ u_{n1} & u_{n2} & \cdots & u_{nn}\end{pmatrix}=OT_1.$$

这就导出了分解式, 其中下三角方阵 T_1 的对角元素 $u_{11},\cdots,u_{nn}$ 恒大于零.

再来证明唯一性. 今若方阵 A 有两种分解

$$A=OT=O_2T_2,\tag{9.2.37}$$

其中 O 和 O_2 都是正交方阵, T 和 T_2 都是上 (下) 三角方阵, 且对角元素恒大于零. 则有

$$TT_2^{-1}=O^{-1}O_2,\tag{9.2.38}$$

其中 $T_3 = TT_2^{-1}$ 仍是上 (下) 三角方阵, 且对角元素恒大于零; $T_3 = O_3 = O^{-1}O_2$ 仍是正交方阵. 所以 O_3 也是上 (下) 三角方阵, 且对角元素恒大于零. 即

$$O_3 = \begin{pmatrix} d_{11} & & \\ \vdots & \ddots & \\ d_{n1} & \cdots & d_{nn} \end{pmatrix}.$$

由 $O_3O_3' = E_n$, 因此 $d_{ij} = 0$, $1 \leqslant j < i \leqslant n$, $d_{11}^2 = d_{22}^2 = \cdots = d_{nn}^2 = 1$. 由 O_3 的对角元素恒大于零可知 $d_{11} = d_{22} = \cdots = d_{nn} = 1$, 即$O_3 = E_n$. 因此 $O_2 = O$, 所以 $T_2 = T$. 这就证明了唯一性.

再证后一断言. 设 A 非异, 则 A' 也是 n 阶非异方阵. 对方阵 A' 用前一断言, 所以唯一地存在一个上 (下) 三角方阵 T_1 和正交方阵 O_1, 使得 $A' = O_1'T_1'$, 其中 T_1' 的对角元素恒大于零. 这证明了$A = T_1O_1$. □

定理 9.2.25 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取 $m(\leqslant n)$ 个两两正交的单位向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$, 则存在 $n-m$ 个单位向量 $\alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成实 Euclid 空间 $\mathfrak{L}$ 的一组标准正交基.

证 今 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性无关, 所以存在 $n-m$ 个向量 $\beta_{m+1}, \beta_{m+2}, \cdots, \beta_n$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_{m+1}, \beta_{m+2}, \cdots, \beta_n$ 构成 $\mathfrak{L}$ 的一组基. 利用 Schmidt 正交化, 从这组基出发, 对 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 作 Schmidt 正交化, 所得到的向量仍是 $\alpha_1, \alpha_2, \cdots, \alpha_m$. 再依次作下去, 便能作出标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ □

这个定理的矩阵语言是:

定理 9.2.26 设 $m \leqslant n$. 任给一个 $n \times m$ 实矩阵

$$O_1 = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \tag{9.2.39}$$

它适合条件 $O_1'O_1 = E_m$, 则必有一个 $n \times (n-m)$ 实矩阵

$$O_2 = \begin{pmatrix} a_{1,m+1} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n,m+1} & \cdots & a_{nn} \end{pmatrix}, \tag{9.2.40}$$

使得 n 阶实方阵

$$O = \begin{pmatrix} O_1 & O_2 \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{9.2.41}$$

是一个实正交方阵.

证 令 O_1 的 m 个列向量

$$\alpha_1=(a_{11},\cdots,a_{n1})',\quad \cdots,\quad \alpha_m=(a_{1m},\cdots,a_{nm})'.$$

作为 n 维实 Euclid 空间 $\mathbb{R}^n$ 中的向量, 已知适合关系

$$O_1'O_1=\begin{pmatrix}\alpha_1 & \cdots & \alpha_m\end{pmatrix}\begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha_m'\end{pmatrix}=\begin{pmatrix}\alpha_1\alpha_1' & \cdots & \alpha_1\alpha_m'\\ \vdots & & \vdots\\ \alpha_m\alpha_1' & \cdots & \alpha_m\alpha_m'\end{pmatrix}=E_m.$$

所以

$$(\alpha_i,\alpha_j)=\alpha_i\alpha_j'=\delta_{ij},\quad 1\leqslant i,j\leqslant m.$$

即 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是 $\mathbb{R}^n$ 中一组两两正交的单位向量. 由定理 9.2.25, 存在 $n-m$ 个单位向量 $\alpha_{m+1},\alpha_{m+2},\cdots,\alpha_n$, 使得 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 构成 $\mathbb{R}^n$ 的一组标准正交基. 再以向量 $\alpha_{m+1},\alpha_{m+2},\cdots,\alpha_n$ 的坐标是列向量, 构作 $n\times n$ 矩阵 O, 那末 O 的列向量构成 $\mathbb{R}^n$ 的一组标准正交基, 所以 O 是正交方阵. □

上面两个定理在 $m=1$ 的情形是最常用的. 这时定理可以改写为

定理 9.2.27 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取单位向量 α, 则存在一组标准正交基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 使得 α 是指定的第 j 个基向量 α_j. 特别, 单位向量能成为一组标准正交基的第一个向量.

例 9.2.1 试用 Schmidt 正交化, 从基

$$\beta_1=(1,1,0)',\quad \beta_2=(1,-1,0)',\quad \beta_3=(0,1,2)'$$

造出一组 $\mathbb{R}^3$ 的标准正交基, 且将三阶非异方阵

$$B=\begin{pmatrix}\beta_1 & \beta_2 & \beta_3\end{pmatrix}=\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 1\\ 0 & 0 & 2\end{pmatrix}$$

表成 OT, 其中 T 是上三角方阵, 对角元素大于零, O 是正交方阵.

解 首先, 取

$$\alpha_1=\frac{1}{|\beta_1|}\beta_1=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0)',$$

再令

$$\gamma_2=\beta_2+\lambda_{12}\alpha_1=(1,-1,0)'+\lambda_{12}(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0),$$

使得 γ_2 和 α_1 正交, 即

$$0=(\gamma_2,\alpha_1)=(\beta_2,\alpha_1)+\lambda_{12}=\lambda_2.$$

所以取 $\gamma_2=\beta_2$, 则 γ_2 和 α_1 正交. 再将 γ_2 单位化, 即令

$$\alpha_2=\frac{1}{|\beta_2|}\beta_2=(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0)',$$

则 α_2 和 α_1 是互相正交的单位向量. 最后, 令

$$\gamma_3=\beta_3+\lambda_{13}\alpha_1+\lambda_{23}\alpha_2=(0,1,2)+\lambda_{13}(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0)'+\lambda_{23}(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0)',$$

使得 γ_3 和 α_1 及 α_2 都正交, 即

$$0=(\gamma_3,\alpha_1)=(\beta_3,\alpha_1)+\lambda_{13}=\frac{1}{\sqrt{2}}+\lambda_{13},$$
$$0=(\gamma_3,\alpha_2)=(\beta_3,\alpha_2)+\lambda_{23}=-\frac{1}{\sqrt{2}}+\lambda_{23}.$$

所以取

$$\gamma_3=\beta_3-\frac{1}{\sqrt{2}}\alpha_1+\frac{1}{\sqrt{2}}\alpha_2=(0,0,2)'.$$

则 γ_3 和 α_1 及 α_2 都正交. 再将 γ_3 单位化, 即令

$$\alpha_3=(0,0,1)'.$$

则

$$\begin{cases}\alpha_1=\dfrac{1}{\sqrt{2}}\beta_1,\\[2mm] \alpha_2=\dfrac{1}{\sqrt{2}}\beta_2,\\[2mm] \alpha_3=-\dfrac{1}{4}\beta_1+\dfrac{1}{4}\beta_2+\dfrac{1}{2}\beta_3\end{cases}$$

构成 $\mathbb{R}^3$ 的一组标准正交基. 且

$$O=\begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}.$$

令

$$\begin{cases}\beta_1=\sqrt{2}\alpha_1,\\ \beta_2=\sqrt{2}\alpha_2,\\ \beta_3=\dfrac{1}{\sqrt{2}}\alpha_1-\dfrac{1}{\sqrt{2}}\alpha_2+2\alpha_3,\end{cases}$$

所以三阶非异方阵 B 可表成:

$$B = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & \dfrac{1}{\sqrt{2}} \\ 0 & \sqrt{2} & -\dfrac{1}{\sqrt{2}} \\ 0 & 0 & 2 \end{pmatrix} = OT,$$

其中 T 是上三角方阵, 对角元素大于零, O 是正交方阵.

例 9.2.2 设 N 是实正交方阵 O 的一个 r 阶子式, M 是 N 的代数余子式. 试证:

$$N = M\det(O). \tag{9.2.42}$$

证 先证如下情形. 设 N 是 n 阶实正交方阵 O 的前 r 行, 前 r 列决定的子矩阵的行列式, 即记

$$O = \begin{pmatrix} O_{11}^{(r)} & O_{12} \\ O_{21} & O_{22} \end{pmatrix},$$

其中 $\det(O_{11}) = N$. 于是 $\det(O_{22}) = M$. 所以要证 $\det(O_{11}) = \det(O_{22})\det(O)$.

今 $OO' = E_n$, 即

$$O_{11}O_{11}' + O_{12}O_{12}' = E_r, \quad O_{11}O_{21}' + O_{12}O_{22}' = 0, \quad O_{21}O_{21}' + O_{22}O_{22}' = E_{n-r}.$$

因此

$$\begin{pmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{pmatrix} \begin{pmatrix} E_r & O_{21}' \\ 0 & O_{22}' \end{pmatrix} = \begin{pmatrix} O_{11} & 0 \\ O_{21} & E_{n-r} \end{pmatrix}.$$

两边取行列式, 便证明了 $\det(O)\det(O_{22}) = \det(O_{11})$. □

上面是在特殊情形时证明了例题. 现在来考虑一般情形. 设

$$N = \det O \begin{pmatrix} i_1 i_2 \cdots i_r \\ j_1 j_2 \cdots j_r \end{pmatrix},$$

其中 $1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n$, $1 \leqslant j_1 < j_2 < \cdots < j_r \leqslant n$, 符号的意义见 §2.2. 将实正交方阵 O 按照两行互换的方法及两列互换的方法变成另一方阵 O_1, 使得 O_1 的前 r 行, 前 r 列构成的子式 N_1 就是 N. 由行列式基本性质可知, 子式 N_1 在方阵 O_1 中的余子式 M_1 就是子式 N 在方阵 O 中的余子式 M. 所以

$$M = (-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r} M_1 = \delta_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r} M_1.$$

由于 O_1 仍是正交方阵, 这时我们已证 $N_1 = M_1 \det(O_1)$, 于是

$$N = N_1 = M_1 \det(O_1) = \delta_{j_1 j_2 \cdots j_r}^{i_1 i_2 \cdots i_r} M_1 \det(O) = M\det(O).$$ □

例 9.2.3 设 A 是 $2n$ 阶实方阵. 记 $2n$ 阶实方阵

$$J=\begin{pmatrix}0 & E_n\\ -E_n & 0\end{pmatrix}.$$

如果实方阵 A 适合条件

$$AJA'=J,$$

试证: $\det(A)=1$. (例 11.1.1 给出用复技巧的证明.)

证 将 A 按前 n 行, 前 n 列分为四块 $A=\begin{pmatrix}B & C\\ D & G\end{pmatrix}$, 则条件 $AJA'=J$ 可详细地写为 $BC'=CB'$, $GD'=DG'$, $BG'-CD'=E_n$. 下面用摄动法来证明 $\det(A)=1$. 为此, 先证明矩阵方程 $XC'=CX'$ 必有解 X, 其中 X 是 n 阶非异实方阵. 事实上, 存在 n 阶非异实方阵 P 和 Q, 使得

$$C=P\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}Q,$$

其中 $r=\operatorname{rank}(C)$. 代入方程, 有

$$XQ'\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}P'=P\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}QX'.$$

此即

$$Y\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}Y',$$

其中 $Y=P^{-1}XQ'$ 是 n 阶非异实方阵. 为了求解 X, 显然取 $Y=E_n$ 即可. 这时, $X=P(Q^{-1})'$, 它非异, 且有 $XC'=CX'$. 这证明了存在所需要的解 X.

现在引进实参数 $\varepsilon>0$. 作 $2n$ 阶实方阵

$$A_\varepsilon=\begin{pmatrix}B+\varepsilon X & C\\ D & G\end{pmatrix},$$

则

$$\det(B+\varepsilon X)=\det(\varepsilon E_n+BX^{-1})\det(X)=(\det(X))\varepsilon^n+\cdots$$

是 ε 的 n 次实多项式, 这里用到 $\det(X)\neq 0$. 所以至多有 n 个不同实根. 因此存在无限序列 $\{\varepsilon_j\}$, 使得

$$\varepsilon_1>\varepsilon_2>\cdots>\varepsilon_j>\cdots>0,\qquad \lim_{j\to\infty}\varepsilon_j=0,$$

且 $\det(\varepsilon_k X+B)\neq 0$, $k=1,2,\cdots$. 由

$$\begin{pmatrix}E_n & 0\\ -D(B+\varepsilon_jX)^{-1} & E_n\end{pmatrix}\begin{pmatrix}B+\varepsilon_jX & C\\ D & G\end{pmatrix}=\begin{pmatrix}B+\varepsilon_jX & C\\ 0 & G-D(B+\varepsilon_jX)^{-1}C\end{pmatrix}$$

注意到 $BC' = CB'$, $XC' = CX'$, 所以 $(B+\varepsilon_j X)C' = C(B+\varepsilon_j X)'$. 因为 $\det(B+\varepsilon_j X) \neq 0$, 所以有 $C'(B'+\varepsilon_j X')^{-1} = (B+\varepsilon_j X)^{-1}C$. 代入, 有

$$\begin{aligned}\det(A_{\varepsilon_j}) &= \det(B+\varepsilon_j X)\det(G - D(B+\varepsilon_j X)^{-1}C)\\ &= \det(G - DC'(B'+\varepsilon_j X')^{-1})\det(B'+\varepsilon_j X')\\ &= \det(\varepsilon_j GX' + GB' - DC'). = \det(\varepsilon_j GX' + E_n).\end{aligned}$$

双方取 $j \to \infty$, 便有 $\det(A) = \det(E_n) = 1$. □

下面考虑子集 $\mathfrak{S}$ 的正交补. 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中任取子集合 $\mathfrak{S}_1$ 和 $\mathfrak{S}_2$, 则 $(\mathfrak{S}_1, \mathfrak{S}_2) = 0$ 表示 $(\alpha, \beta) = 0, \forall \alpha \in \mathfrak{S}_1, \beta \in \mathfrak{S}_2$.

定义 9.2.28　设 $\mathfrak{S}$ 是 n 维实 Euclid 空间 $\mathfrak{L}$ 中的子集. 记

$$\mathfrak{S}^{\perp} = \{\, \alpha \in \mathfrak{L} \mid (\alpha, \mathfrak{S}) = 0 \,\}, \tag{9.2.43}$$

则 $\mathfrak{S}^{\perp}$ 称为子集 $\mathfrak{S}$ 的**正交补**.

定理 9.2.29　n 维实 Euclid 空间 $\mathfrak{L}$ 中子集 $\mathfrak{S}$ 的正交补 $\mathfrak{S}^{\perp}$ 是 $\mathfrak{L}$ 的子空间, 且

$$\mathfrak{S} \subset (\mathfrak{S}^{\perp})^{\perp}. \tag{9.2.44}$$

又

$$\mathfrak{L} = \mathfrak{S}^{\perp} \oplus (\mathfrak{S}^{\perp})^{\perp} \tag{9.2.45}$$

是子空间直接和. 特别, 当 $\mathfrak{S}$ 本身是子空间时, 则有

$$(\mathfrak{S}^{\perp})^{\perp} = \mathfrak{S}. \tag{9.2.46}$$

所以 $\mathfrak{L}$ 有子空间直接和分解

$$\mathfrak{L} = \mathfrak{S} \oplus \mathfrak{S}^{\perp}. \tag{9.2.47}$$

证　显然 $\mathfrak{S}^{\perp}$ 是子空间. 由 $(\mathfrak{S}^{\perp})^{\perp}$ 的定义可知 $\mathfrak{S} \subset (\mathfrak{S}^{\perp})^{\perp}$. 下面证 $\mathfrak{L} = \mathfrak{S}^{\perp} \oplus (\mathfrak{S}^{\perp})^{\perp}$ 是空间直接和.

任取 $\alpha \in \mathfrak{S}^{\perp} \cap (\mathfrak{S}^{\perp})^{\perp}$, 则 $\alpha \in \mathfrak{S}^{\perp}$, 且 $\alpha \in (\mathfrak{S}^{\perp})^{\perp}$. 即 $(\alpha, \alpha) = 0$. 由于内积是实正定对称双线性函数, 所以 $\alpha = 0$. 这证明了 $\mathfrak{S}^{\perp} \cap (\mathfrak{S}^{\perp})^{\perp} = 0$. 所以问题化为证明 $\mathfrak{S}^{\perp} + (\mathfrak{S}^{\perp})^{\perp} = \mathfrak{L}$.

在 $\mathfrak{S}^{\perp}$ 中取标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_r$, 于是在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$. 由于 $(\alpha_{r+j}, \alpha_k) = 0$, $j = 1, 2, \cdots, n-r$. 所以 $\mathfrak{S}^{\perp} + (\mathfrak{S}^{\perp})^{\perp} = \mathfrak{L}$. 由 $\mathfrak{S}^{\perp} \cap (\mathfrak{S}^{\perp})^{\perp} = 0$ 便证明了 $\mathfrak{L} = \mathfrak{S}^{\perp} + (\mathfrak{S}^{\perp})^{\perp}$ 是空间直接和. 特别, 当 $\mathfrak{S}$ 本身是子空间时, 在 $\mathfrak{S}$ 中取定标准正交基 $\beta_1, \cdots, \beta_r$, 由定理 9.2.25, 在 $\mathfrak{L}$ 中存在标准正交基 $\beta_1, \cdots, \beta_r, \beta_{r+1}, \cdots, \beta_n$. 由正交补的定义可知 $\beta_{r+1}, \cdots, \beta_n$ 在 $\mathfrak{S}^{\perp}$ 中, 所以

$\mathfrak{S}+\mathfrak{S}^{\perp}=\mathfrak{L}$. 再 $\mathfrak{S}\cap\mathfrak{S}^{\perp}\subset(\mathfrak{S}^{\perp})^{\perp}\cap\mathfrak{S}^{\perp}=0$, 这证明了 $\mathfrak{L}=\mathfrak{S}+\mathfrak{S}^{\perp}$ 是子空间直接和. □

定义 9.2.30 对 n 维实 Euclid 空间 $\mathfrak{L}$ 中子空间 $\mathfrak{S}$, 任取 $\alpha\in\mathfrak{L}$, 则唯一存在 $\beta\in\mathfrak{S}$, $\gamma\in\mathfrak{S}^{\perp}$, 使得 $\alpha=\beta+\gamma$. 于是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的对应 $\alpha\to\beta$ 是 $\mathfrak{L}$ 上的线性变换, 称为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的**正交投影**.

定理 9.2.31 设 $\mathscr{A}$ 是 n 维实 Eudlid 空间 $\mathfrak{L}$ 上的线性变换. 记 $\mathscr{A}$ 的象空间 $\mathscr{A}(\mathfrak{L})=\mathfrak{S}$. 则 $\mathscr{A}$ 是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影当且仅当

$$(\mathscr{A}(\mathfrak{L}),(\mathrm{id}-\mathscr{A})(\mathfrak{L}))=0. \tag{9.2.48}$$

这时有 $\mathscr{A}^2=\mathscr{A}$.

证 今 $\mathfrak{L}=\mathfrak{S}\oplus\mathfrak{S}^{\perp}$. 任取 $\alpha\in\mathfrak{L}$, 记 $\beta\in\mathfrak{S}$, $\gamma\in\mathfrak{S}^{\perp}$, 使得 $\alpha=\beta+\gamma$. 由定义, $\mathscr{A}$ 是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影当且仅当 $\mathscr{A}(\alpha)=\beta$. 显然 $\mathscr{A}(\alpha)=\beta$ 当且仅当 $(\mathrm{id}-\mathscr{A})(\alpha)=\alpha-\beta=\gamma$, 此即

$$(\mathrm{id}-\mathscr{A})(\mathfrak{L})=\mathfrak{S}^{\perp}.$$

由 $\mathscr{A}(\mathfrak{L})=\mathfrak{S}$, 所以 $\mathscr{A}$ 是正交投影当且仅当 $(\mathscr{A}(\mathfrak{L}),(\mathrm{id}-\mathscr{A})(\mathfrak{L})=0$.

设 $\mathscr{A}$ 是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影. 于是

$$\mathscr{A}^2(\alpha)=\mathscr{A}(\mathscr{A}(\alpha))=\mathscr{A}(\beta)=\beta=\mathscr{A}(\alpha),\quad\forall\alpha\in\mathfrak{L},$$

此即 $\mathscr{A}^2=\mathscr{A}$. □

习　题　9.2

9.2.1 试解释 Schmidt 正交化方法在通常的平面和空间中的几何意义.

9.2.2 试求边长是 1 的 n 维立方体的对角线的长度.

9.2.3 给出一个例子以说明: 如果 n 阶方阵 A 的 n 个行向量两两正交, 但是它的 n 个列向量并不一定是两两正交的.

9.2.4 如果一个实正交方阵 A 是准上 (下) 三角方阵, 则对角块都是实正交方阵, 且 A 是准对角方阵. 当 A 是上 (下) 三角方阵, 则此正交方阵必是对角方阵, 且其对角元素均是 ±1.

9.2.5 如果 A 和 B 都是 n 阶实正交方阵, 且 $\det(A)+\det(B)=0$, 则方阵 $A+B$ 是奇异方阵.

9.2.6 如果 O 是 n 阶实正交方阵, 且 $A=\mathrm{diag}(a_1,a_2,\cdots,a_n)$, 则方阵 OA 的特征根 λ_j 适合不等式

$$m\leqslant|\lambda_j|\leqslant M,\quad 1\leqslant j\leqslant n, \tag{9.2.49}$$

其中

$$m=\min_{1\leqslant j\leqslant n}|a_j|,\quad M=\max_{1\leqslant j\leqslant n}|a_j|. \tag{9.2.50}$$

9.2.7 试证: 实正交方阵的任一子方阵的特征根的模小于等于 1.

9.2.8　试证: n 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关的必要且充分条件是它的**Gram 方阵**

$$G = \begin{pmatrix} (\alpha_1, \alpha_1) & \cdots & (\alpha_1, \alpha_n) \\ \vdots & & \vdots \\ (\alpha_n, \alpha_1) & \cdots & (\alpha_n, \alpha_n) \end{pmatrix} \tag{9.2.51}$$

非异. 在 n 维实 Euclid 空间 $\mathbb{R}^n$ 中取定一组标准正交基 $\alpha_1 = (a_{11}, \cdots, a_{n1})', \cdots, \alpha_n = (a_{1n}, \cdots, a_{nn})'$ 后, 试证明:

$$\det(G) = \det(\alpha_1\, \alpha_2\, \cdots\, \alpha_n). \tag{9.2.52}$$

9.2.9　试证: 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中的两组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 经一正交变换能互变的必要且充分条件是它们的 Gram 方阵相等. 又证: 当 $1 \leqslant r \leqslant n$ 时, r 个向量 $\delta_1, \delta_2, \cdots, \delta_r$ 线性无关当且仅当 $\det(G(\delta_1, \delta_2, \cdots, \delta_r)) \neq 0$.

9.2.10　设 O 是 n 阶实正交方阵. 并设 λ 是它的特征根, α 是属于特征根 λ 的特征向量. 试证: $|\lambda| = 1$. 当特征根 λ 是复数时, 将特征向量 α 的实部及虚部分开, 即记 $\alpha = \beta + \sqrt{-1}\gamma$. 试证: β 和 γ 正交且长度相等.

9.2.11　设 K 是 n 阶实反对称方阵, 并设 λ 是它的非零纯虚特征根, α 是属于特征根 λ 的特征向量. 将 α 的实部及虚部分开, 即记 $\alpha = \beta + \sqrt{-1}\gamma$. 试证: β 和 γ 正交且长度相等.

9.2.12　设 n 阶实方阵 A 的秩是 r. 试证: 存在实正交方阵 O 及置换方阵 P, 使得

$$PAO = \begin{pmatrix} T_1 & 0 \\ T_2 & 0 \end{pmatrix},$$

其中 T_1 是 r 阶下三角方阵, 对角元素都大于零.

9.2.13　设 O 是三阶实正交方阵, 且 $\det(O) = 1$. 试证: 在区间 $[-1, 3]$ 中存在实数 λ_0, 使得 $O^3 - \lambda O^2 + \lambda_0 O - E_3 = 0$.

9.2.14　设 $\mathscr{A}$ 是 n 阶实方阵, 且 A 的特征根不等于 0, -1. 试证: A 和 $A + E$ 都非异, 且 A 实正交当且仅当 $(A + E)^{-1} + (A' + E)^{-1} = E$.

9.2.15　试证: 任一行列式是 1 的 n 阶实正交方阵, 是有限个形如

$$O_{jk} = \begin{pmatrix} E_{j-1} & 0 & 0 & 0 & 0 \\ 0 & \cos(\theta_{ij}) & 0 & \sin(\theta_{ij}) & 0 \\ 0 & 0 & E_{k-j} & 0 & 0 \\ 0 & -\sin(\theta_{ij}) & 0 & \cos(\theta_{ij}) & 0 \\ 0 & 0 & 0 & 0 & E_{n-k} \end{pmatrix},$$

$1 \leqslant j < k \leqslant n$ 的 n 阶实正交方阵的乘积, 其中 $0 \leqslant \theta_{ij} < 2\pi$, $1 \leqslant i < j \leqslant n$. 对于行列式是 -1 的实正交方阵, 还需要加上一个因子, 它是实正交方阵 $\operatorname{diag}(E_{n-1}, 1)$.

§9.3　实方阵在实正交相似下的标准形

定义 9.3.1　两个 n 阶实方阵 A 和 B 称为**实正交相似的**, 如果存在 n 阶实正交方阵 O, 使得

$$B = O^{-1}AO = O'AO. \tag{9.3.1}$$

不难验证, 实正交相似是所有 n 阶实方阵构成的集合 $\mathrm{gl}(n,\mathbb{R})$ 中的等价关系, 这个等价关系下的标准形理论没有解决. 下面考虑实正交相似下的某些特殊实方阵构成的子集合中的标准形理论.

定义 9.3.2 适合条件 $AA' = A'A$ 的实方阵称为**实规范方阵**; 适合条件 $A' = A$ 的实方阵称为**实对称方阵**; 适合条件 $A' = -A$ 的实方阵称为**实反对称方阵**.

由定义可知, 实正交方阵、实对称方阵和实反对称方阵都是实规范方阵. 又设 A, O_1, S, K 分别是实规范方阵、实正交方阵、实对称方阵、实反对称方阵, 则对任一实正交方阵 O, $O^{-1}AO, O^{-1}O_1O, O^{-1}SO, O^{-1}KO$ 仍分别是实规范方阵、实正交方阵、实对称方阵、实反对称方阵. 在讨论这四种实方阵在实正交相似下的标准形以前, 先给出下面的定理.

定理 9.3.3 任一 n 阶实方阵 A 必实正交相似于实准下 (上) 三角方阵,

$$B = \begin{pmatrix} A_{11}^{(2)} & & & & & \\ \vdots & \ddots & & & & \\ A_{t1}^{(2)} & \cdots & A_{tt}^{(2)} & & & \\ A_{2t+1,1}^{(1,2)} & \cdots & A_{2t+1,t}^{(1,2)} & \lambda_{2t+1} & & \\ \vdots & & \vdots & \vdots & \ddots & \\ A_{n1}^{(1,2)} & \cdots & A_{nt}^{(1,2)} & \lambda_{n,2t+1} & \cdots & \lambda_n \end{pmatrix}, \tag{9.3.2}$$

其中 $A_{11}, \cdots, A_{tt}$ 都是二阶实方阵. 它们有非实数特征根, 分别是 $\lambda_1, \overline{\lambda_1}, \cdots, \lambda_t, \overline{\lambda_t}$, 这些特征根也构成实方阵 A 的所有复特征根. 又 $\lambda_{2t+1}, \cdots, \lambda_n$ 是实方阵 A 的所有实特征根.

特别, 当实方阵 A 的特征根都是实数时, 实方阵 A 实正交相似于一个实下 (上) 三角方阵.

证 若证明了 n 阶实方阵 A 实正交相似于准下三角方阵, 对 A' 用此结果, 便证明了 A 实正交相似于准上三角方阵.

为了要证明存在实正交方阵 O, 使得 $O'AO = B$. 为此, 对方阵 A 的阶数作归纳法. 当 $n = 1$ 时, 定理显然成立. 假设对一切阶数 $\leqslant n-1$ 的实方阵, 定理成立. 现在来证对任一 n 阶实方阵 A, 定理也成立. 下面分两种情形讨论.

设 A 有一个实特征根 λ_n, 则属于特征根 λ_n 的特征向量 β_n 仍是实向量. 显然, 无妨设 β_n 是单位向量, 所以有 $\beta_n'\beta_n = 1$, $A\beta_n = \lambda_n\beta_n$. 由定理 9.2.25, 存在一个以 β_n 为第 n 个列向量的 n 阶实正交方阵 $O_1 = (\beta_1, \beta_2, \cdots, \beta_n)$. 令

$$O_1'AO_1 = \begin{pmatrix} \beta_1' \\ \vdots \\ \beta_n' \end{pmatrix} A \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} = \begin{pmatrix} \beta_1'A\beta_1 & \cdots & \beta_1'A\beta_n \\ \vdots & & \vdots \\ \beta_n'A\beta_1 & \cdots & \beta_n'A\beta_n \end{pmatrix}.$$

由 $\beta_i' A\beta_n = \lambda_n \beta_i'\beta_n = \delta_{in}\lambda_n$, $1 \leqslant i \leqslant n$, 所以

$$O_1' A O_1 = \begin{pmatrix} A_1 & 0 \\ A_2 & \lambda_n \end{pmatrix},$$

其中 A_1 是 $n-1$ 阶实方阵. 由归纳法假设, 存在 $n-1$ 阶实正交方阵 O_2, 使得

$$O_2' A_1 O_2 = \begin{pmatrix} A_{11}^{(2)} & & & & & \\ \vdots & \ddots & & & & \\ A_{t1}^{(2)} & \cdots & A_{tt}^{(2)} & & & \\ A_{2t+1,1}^{(1,2)} & \cdots & A_{2t+1,t}^{(1,2)} & \lambda_{2t+1} & & \\ \vdots & & \vdots & \vdots & \ddots & \\ A_{n-1,1}^{(1,2)} & \cdots & A_{n-1,t}^{(1,2)} & \lambda_{n-1,2t+1} & \cdots & \lambda_{n-1} \end{pmatrix},$$

其中 $A_{11}, \cdots, A_{tt}$ 都是二阶实方阵, 它们的特征根都不是实数.

再令 $O_3 = \mathrm{diag}\,(O_2, 1)$, 它仍旧是实正交方阵. 由直接计算可知, 记 $O_4 = O_1O_3$, 则

$$O_4' A O_4 = \begin{pmatrix} O_2' A_1 O_2 & 0 \\ A_2 O_2 & \lambda_n \end{pmatrix}.$$

因为实方阵 $O_4' = O_4^{-1}$, 所以实方阵 A 和实方阵 $O_4'AO_4$ 的特征多项式相同, 即实方阵 $O_4'AO_4$ 的实特征根是方阵 A 的实特征根 $\lambda_{t+1}, \cdots, \lambda_n$, 实方阵 $O_4'AO_4$ 的复特征根是实方阵 A 的复特征根 $\lambda_1, \overline{\lambda_1}, \cdots, \lambda_t, \overline{\lambda_t}$, 其中 λ_j 和 $\overline{\lambda_j}$ 是二阶实方阵 A_{jj} 的一对共轭复特征根, $1 \leqslant j \leqslant t$. 这就证明了定理.

设实方阵 A 没有实特征根. 这时任取一个复特征根, 将它分开实部和虚部: $\lambda_0 + \sqrt{-1}\mu_0$, 其中 λ_0, μ_0 都是实数, 自然, $\mu_0 \neq 0$. 将实方阵的属于特征根 $\lambda_0 + \sqrt{-1}\mu_0$ 的复特征向量也分开实部和虚部, 即令相应的特征向量是 $\sigma + \sqrt{-1}\tau$, 其中 σ 和 τ 都是实向量. 则 $A(\sigma + \sqrt{-1}\tau) = (\lambda_0 + \sqrt{-1}\mu_0)(\sigma + \sqrt{-1}\tau)$ 等价于 $A\sigma = \lambda_0\sigma - \mu_0\tau$, $A\tau = \mu_0\sigma + \lambda_0\tau$.

首先, 我们断言 σ 和 τ 线性无关. 用反证法. 设若线性相关, 则存在非零实向量 β, 使得 $\sigma + \sqrt{-1}\tau = a\beta$, 其中 a 是非零复常数. 代入有 $A(a\beta) = (\lambda_0 + \sqrt{-1}\mu_0)a\beta$. 由 $a \neq 0$, 有 $A\beta = (\lambda_0 + \sqrt{-1}\mu_0)\beta$, 但是 $A\beta$ 实, $\beta \neq 0$, 故有 $\mu_0 = 0$. 这和条件 $\mu_0 \neq 0$ 矛盾.

今 σ 和 τ 线性无关, 所以在 $\mathbb{R}^n$ 中可取基 $\sigma, \tau, \gamma_3, \cdots, \gamma_n$. 对它作 Schmidt 正交化, 便得到标准正交基 $\beta_n, \beta_{n-1}, \cdots, \beta_1$, 使得 $\beta_{n-1} = b\sigma$, $\beta_n = c\sigma + d\tau$, 其中 $b > 0$, $d > 0$, $c \in \mathbb{R}$. 因此 $\sigma = b^{-1}\beta_{n-1}$, $\tau = d^{-1}\beta_n - \frac{c}{bd}\beta_{n-1}$, 又 $\beta_j'\sigma = \beta_j'\tau = 0$, $j = 1, 2, \cdots, n-2$. 而 $A(\beta_{n-1}, \beta_n) = (A\beta_{n-1}, A\beta_n) = (bA\sigma, cA\sigma + dA\tau)$, 所以

$$A(\beta_{n-1}, \beta_n) = (\beta_{n-1}, \beta_n)\begin{pmatrix} b & c \\ 0 & d \end{pmatrix}^{-1}\begin{pmatrix} \lambda_0 & \mu_0 \\ -\mu_0 & \lambda_0 \end{pmatrix}\begin{pmatrix} b & c \\ 0 & d \end{pmatrix}.$$

又当 $j=1,2,\cdots,n-2$ 时, 有

$$\begin{aligned}\beta_j'A\beta_{n-1}&=b\beta_j'A\sigma=b\beta_j'(\lambda_0\sigma-\mu_0\tau)=0,\\ \beta_j'A\beta_n&=c\beta_j'A\sigma+d\beta_j'A\tau=d\beta_j'(\mu_0\sigma-\lambda_0\tau)=0.\end{aligned}$$

记 n 阶实正交方阵 $O_1=(\beta_1,\beta_2,\cdots,\beta_n)$, 则有

$$O_1'AO_1=\begin{pmatrix}\beta_1'A\beta_1 & \cdots & \beta_1'A\beta_n\\ \vdots & & \vdots\\ \beta_n'A\beta_1 & \cdots & \beta_n'A\beta_n\end{pmatrix}=\begin{pmatrix}A_1^{(n-2)} & 0\\ A_2 & \begin{pmatrix}\beta_{n-1}'A\beta_{n-1} & \beta_{n-1}'A\beta_n\\ \beta_n'A\beta_{n-1} & \beta_n'A\beta_n\end{pmatrix}\end{pmatrix},$$

所以

$$O_1'AO_1=\begin{pmatrix}A_1 & 0\\ A_2 & \begin{pmatrix}b & c\\ 0 & d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}\end{pmatrix}.$$

由于 A_1 的特征根仍是 A 的特征根, 所以 A_1 仍没有实特征根. 今

$$\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}^{-1}\begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}\begin{pmatrix}b & c\\ 0 & d\end{pmatrix}\quad 和\quad \begin{pmatrix}\lambda_0 & \mu_0\\ -\mu_0 & \lambda_0\end{pmatrix}$$

的特征根相同, 是 $\lambda_0+\sqrt{-1}\mu_0$ 和 $\lambda_0-\sqrt{-1}\mu_0$. 对 A_1 用归纳法假设, 便证明了定理成立. □

这个定理虽然未能解决实方阵在实正交相似下的标准形理论. 不过它说明了: 将实方阵在实正交相似下分成等价类后, 可以在每个等价类中取上列形式的上三角方阵, 这总比任取一个代表元素来得简单. 另一方面, 由于实正交相似必须相似, 所以同一等价类中任两方阵的特征根相同. 换句话说, 实方阵的特征根是它在实正交相似下的不变量, 但是它不是全系不变量.

另一方面, 由引理 9.1.9, n 阶实方阵 A 可以唯一地分解为 n 阶实对称方阵 S 和 n 阶实反对称方阵 K 的和. 所以这也说明了一般的实方阵在实正交相似下分类理论的困难性. 事实上, 任取实正交方阵 O, 则 $O'AO=O'SO+O'KO$, 其中 $S_0=O'SO$ 仍为 n 阶实对称方阵, $K_0=O'KO$ 仍为 n 阶实反对称方阵. 所以问题化为相当于 S 和 K 同时化为标准形. 记**实方阵偶** (S,K) 全体构成的集合是 $\mathfrak{S}$, 我们称实方阵偶 (S,K) 和 (S_0,K_0) 实正交相似. 不难证明实方阵偶的实正交相似是集合中的一种等价关系. 上面的讨论告诉我们, 一般实方阵在实正交相似下的标准形理论就是实方阵偶 (S,K) 在实正交相似下的标准形理论 (其中 S 是 n 阶实对称方阵, K 是 n 阶实反对称方阵). 关于矩阵偶的标准形理论中最简单的结果, 读者可以在第十四章中找到.

下面进一步考虑一些特殊的方阵在实正交相似下的标准形, 给出实规范方阵、实正交方阵、实对称方阵及实反对称方阵在实正交相似下的标准形, 并且给出标准形的几何解释.

定理 9.3.4 (Schur 不等式) 记 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 n 阶实方阵 A 的 n 个特征根, 则有不等式

$$\sum_{j=1}^{n} |\lambda_j|^2 \leqslant \operatorname{tr}(A'A), \tag{9.3.3}$$

且等号成立当且仅当实方阵 A 是实规范方阵.

证 用定理 9.3.3, 所以存在 n 阶实正交方阵 O, 使得 $O'AO = B$, 其中 B 由式 (9.3.2) 给出. 因此, 由 $O^{-1} = O'$, 有

$$\operatorname{tr}(A'A) = \operatorname{tr}(O'AO)'(O'AO) = \operatorname{tr}(B'B). \tag{9.3.4}$$

注意到 $\operatorname{tr}(X'X) = \sum\limits_{i,j=1}^{n} |x_{ij}|^2$, 其中 $X = (x_{ij})$, 所以

$$\operatorname{tr}(A'A) \geqslant \sum_{j=1}^{t} \operatorname{tr}(A'_{jj}A_{jj}) + \sum_{k=2t+1}^{n} |\lambda_k|^2,$$

而且等式成立当且仅当

$$A = \operatorname{diag}(A_{11}, \cdots, A_{tt}, \lambda_{2t+1}, \cdots, \lambda_n), \quad \text{其中}$$

$$A_{jj} = \begin{pmatrix} a_j & b_j \\ c_j & d_j \end{pmatrix}, \quad 1 \leqslant j \leqslant t. \tag{9.3.5}$$

下面证明 Schur 不等式成立, 而且证明等式成立当且仅当 B 为实规范方阵, 即 A_{jj} 为实规范方阵, $1 \leqslant j \leqslant t$.

事实上,任取 $j \in \{1, 2, \cdots, t\}$, 则由式 (9.3.5), 有

$$\operatorname{tr}(A'_{jj}A_{jj}) = a_j^2 + b_j^2 + c_j^2 + d_j^2 \geqslant 2(a_jd_j - c_jd_j) = 2\det(A_{jj}) = 2\lambda_j\overline{\lambda_j} = |\lambda_j|^2 + |\overline{\lambda_j}|^2.$$

所以 Schur 不等式成立. 又等式成立当且仅当任取 $j \in \{1, 2, \cdots, t\}$, 则$a_j^2 + b_j^2 + c_j^2 + d_j^2 = 2(a_jd_j - c_jd_j)$, 即 $(a_j - d_j)^2 + (b_j + c_j)^2 = 0$. 这证明了 $d_j = a_j$, $c_j = -b_j$, 即

$$A_{jj} = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}, \quad 1 \leqslant j \leqslant t. \tag{9.3.6}$$

另一方面, 由直接计算可知, 式 (9.3.5) 定义的方阵 A_{jj} 是实规范方阵当且仅当

$$a_j^2 + b_j^2 = a_j^2 + c_j^2, \quad a_jc_j + b_jd_j = a_jb_j + c_jd_j, \quad c_j^2 + d_j^2 = b_j^2 + d_j^2.$$

因此, $b_j^2 = c_j^2$, $(a_j - d_j)(c_j - b_j) = 0$. 但是 A_{jj} 的特征根是非实的, 所以 $(a_j - d_j)^2 + 4b_jc_j < 0$, 因此 $b_jc_j < 0$, 所以 $d_j = a_j$, $c_j = -b_j \neq 0$. 这证明了实规范方阵有且只有式 (9.3.6). □

上面定理的证明, 实际上也给出了

定理 9.3.5 n 阶实规范方阵 A 实正交相似于准对角形

$$B = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n\right), \tag{9.3.7}$$

其中 $b_1 > 0, \cdots, b_t > 0$, 且 A 的所有特征根是

$$a_1 + \sqrt{-1}b_1,\ a_1 - \sqrt{-1}b_1,\ \cdots,\ a_t + \sqrt{-1}b_t,\ a_t - \sqrt{-1}b_t,\ \lambda_{2t+1},\ \cdots,\ \lambda_n. \tag{9.3.8}$$

所以实规范方阵 A 在实正交相似下的全系不变量是 A 的所有特征根.

这个定理告诉我们, 在所有实规范方阵构成的集合中考虑实正交相似, 则在每个等价类中可以取定理 9.3.5 中给出的标准形是代表元素; 而且, 两个具有相同特征根的实规范方阵必互相等价, 所以实规范方阵的 n 个特征根是它在实正交相似下的全系不变量. 换句话说, 只要知道实规范方阵的 n 个特征根, 不需要寻找实正交方阵, 就能立刻写出它们的标准形来. 这里要注意的是, 标准形仅由 n 个特征根确定, 而表达方式不是唯一的.

今考虑第一类初等方阵 P_{ij} (见 §3.4). 显然, 它是实正交方阵. 且对任一实方阵 A, 则 $P'_{ij}AP_{ij}$ 是将实方阵 A 的第 i 行和第 j 行互换, 同时将它的第 i 列和第 j 列互换. 特别, 这样引起了第 i 个对角元素和第 j 个对角元素互换, 而非对角元素必和非对角元素互换. 有限个第一类初等方阵的乘积称为**置换方阵**. 显然, 置换方阵仍为实正交方阵. 设 P 是置换方阵, 则 $P'AP$ 是将 A 的行和列按同同样方式作了调换, 因此将对角元素作了排列. 正因是如此, 随便给出标准形 B, 那末存在置换方阵 P, 使得标准形 B 的对角块可以任意调换次序.

定理 9.3.6 n 阶实正交方阵 A 实正交相似于准对角形

$$B = \operatorname{diag}\left(\begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{pmatrix}, \cdots, \begin{pmatrix} \cos(\theta_t) & \sin(\theta_t) \\ -\sin(\theta_t) & \cos(\theta_t) \end{pmatrix}, E_u, -E_v\right), \tag{9.3.9}$$

其中 $2t + u + v = n$, $0 < \theta_1 \leqslant \cdots \leqslant \theta_t < \pi$. 所以实正交方阵的特征根的模是 1, 且实正交方阵 O 在实正交相似下的全系不变量是 O 的所有特征根.

证 由 O 实正交, 即有 $OO' = O'O = E_n$, 所以它是实规范方阵. 由定理 9.3.5, 它实正交相似于准对角形

$$B = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n\right),$$

由 $\lambda_j^2 = 1$, 所以 $\lambda_j = \pm 1$, $2t+1 \leqslant j \leqslant n$. 无妨取 $\lambda_{2t+1} = \cdots = \lambda_{2t+u} = 1$, $\lambda_{2t+u+1} = \cdots = \lambda_n = -1$. 记 $v = n-(2t+u)$. 又 $\begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}\begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, 此即 $a_j^2 + b_j^2 = 1$. 因此可取 $a_j = \cos(\theta_j)$, $b_j = \sin(\theta_j)$, $0 \leqslant \theta_j < \pi$. □

定理 9.3.7 n 阶实对称方阵 S 实正交相似于对角形

$$\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n), \tag{9.3.10}$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 且 λ_1, λ_2, $\cdots$, λ_n 是 S 的所有特征根. 所以实对称方阵的特征根必是实数, 且实对称方阵在实正交相似下的全系不变量是它的所有特征根.

证 由 S 实对称, 即有 $SS' = S'S$, 所以它是实规范方阵. 由定理 9.3.5, 存在实正交方阵 O, 使得

$$O'SO = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n\right), \tag{9.3.11}$$

由 S 对称可知 $O'SO$ 也对称, 所以 $b_1 = \cdots = b_t = 0$. 这和假设 $b_1 > 0$, $\cdots$, $bz_t > 0$ 矛盾. 所以证明了 t 不存在, 即 $O'SO = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$. 因此, 实对称方阵的所有特征根必是实数. □

定理 9.3.8 n 阶实反对称方阵实正交相似于准对角形:

$$\operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_t \\ -b_t & 0 \end{pmatrix}, 0\right), \tag{9.3.12}$$

其中 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_t > 0$, 且

$$\sqrt{-1}b_1,\ -\sqrt{-1}b_1,\ \cdots,\ \sqrt{-1}b_t,\ -\sqrt{-1}b_t,\ 0,\ \cdots,\ 0 \tag{9.3.13}$$

是 K 的所有特征根. 所以实反对称方阵的非零特征根必是纯虚数, 因此实反对称方阵的秩必是偶数. 又实反对称方阵在实正交相似下的全系不变量是它的所有特征根.

证 由 S 实反对称, 即有 $KK' = K'K$, 所以它是实规范方阵. 由定理 9.3.5, 所以存在实正交方阵 O, 使得

$$O'KO = \operatorname{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_t & b_t \\ -b_t & a_t \end{pmatrix}, \lambda_{2t+1}, \cdots, \lambda_n\right).$$

由 K 反对称, 所以 $O'KO$ 也是反对称的. 这证明了

$$a_1 = \cdots = a_t = \lambda_{2t+1} = \cdots = \lambda_n = 0.$$

即有

$$O'KO = \operatorname{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_t \\ -b_t & 0 \end{pmatrix}, 0, \cdots, 0\right),$$

其中 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_t > 0$. □

引理 9.3.9 设 $\mathfrak{L}$ 是 n 维实 Euclid 空间. $\mathscr{A}$ 是 $\mathfrak{L}$ 上线性变换, 则它在 $\mathfrak{L}$ 的不同标准正交基下对应的方阵表示互相实正交相似.

证 由于标准正交基间的基变换对应的是实正交方阵, 而 n 维实 Euclid 空间 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 在不同基下的方阵表示在基变换对应的方阵下相似, 这证明了它们互相实正交相似. □

在上面给出了一些特殊的矩阵的标准形, 整个讨论是利用矩阵方法进行的. 它们有下面的几何意义. 在实 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 在 $\mathfrak{L}$ 中取定线性变换 $\mathscr{A}$, 记 $\mathscr{A}$ 在这组基下的方阵表示是 A. 现在在 $\mathfrak{L}$ 中另外取一组标准正交基 $\beta_1, \beta_2, \cdots, \beta_n$. 记旧基到新基的基变换的矩阵表示是 O, 则 O 是实正交方阵, 而线性变换 $\mathscr{A}$ 在新的基下的方阵表示是 $B = O'AO$, 这就给出实正交相似的一种几何解释. 由此可见, 实方阵在实正交相似下分成等价类后, 同一个等价类中的实方阵, 是实 Euclid 空间 $\mathfrak{L}$ 中一个确定的线性变换在所有标准正交基下的方阵表示, 不同等价类表示了不同的线性变换.

下面定义共轭变换. 为此, 先给出下面重要定理.

定理 9.3.10 (Riesz 表示定理) 记 (α, β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 上的内积. 任取线性函数 f, 则在实 Euclid 空间 $\mathfrak{L}$ 中唯一存在向量 α_f, 使得

$$f(\beta) = (\beta, \alpha_f), \quad \forall \beta \in \mathfrak{L}, \tag{9.3.14}$$

且对应 $f \to \alpha_f$ 给出了实 Euclid 空间 $\mathfrak{L}$ 的对偶空间 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构.

证 在实 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 任取 $f \in \mathfrak{L}^*$, 令向量

$$\alpha_f = \sum_{j=1}^{n} f(\alpha_j)\alpha_j, \quad \forall f \in \mathfrak{L}^*,$$

则对任意的 $\beta = \sum\limits_{k=1}^{n} y_k\alpha_k \in \mathfrak{L}$, 有

$$(\beta, \alpha_f) = \sum_{j=1}^{n} f(\alpha_j)(\beta, \alpha_j) = \sum_{j,k=1}^{n} f(\alpha_j)y_k(\alpha_k, \alpha_j) = f(\sum_{j=1}^{n} y_j\alpha_j) = f(\beta).$$

这证明了向量 α_f 的存在性. 再证唯一性, 今若

$$f(\beta) = (\beta, \alpha_f) = (\beta, \xi), \quad \forall \beta \in \mathfrak{L}.$$

于是 $(\beta,\alpha_f-\xi)=0,\ \forall\beta\in\mathfrak{L}$. 此即 $\alpha_f-\xi\in\mathfrak{L}^{\perp}=0$, 即 $\xi=\alpha_f$. 这证明了唯一性.

因此对应 $f\to\alpha_f,\ \forall f\in\mathfrak{L}^*$ 是 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 内的一一映射. 下面证这是到上的一一映射. 今任取 $\alpha\in\mathfrak{L}$, 则显然 $(\beta,\alpha),\ \forall\beta\in\mathfrak{L}$ 是 $\mathfrak{L}$ 上的线性函数, 记作 f. 则有 $f(\beta)=(\beta,\alpha),\ \forall\beta\in\mathfrak{L}$. 下面证 $\alpha_f=\alpha$. 事实上, 已知存在 $\alpha_f\in\mathfrak{L}$, 使得 $f(\beta)=(\beta,\alpha_f),\ \forall\beta\in\mathfrak{L}$. 由唯一性可知 $\alpha_f=\alpha$. 这证明了对应 $f\to\alpha_f$ 是 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的一一对应.

最后证 $f\to\alpha_f$ 是线性同构. 事实上, 任取 $f,g\in\mathfrak{L}^*$, a, b 是数, 则

$$\begin{aligned}(\alpha_{af+bg},\beta)&=(af+bg)(\beta)=af(\beta)+bg(\beta)\\&=a(\alpha_f,\beta)+b(\alpha_g,\beta)=(a\alpha_f+b\alpha_g,\beta),\quad\forall\beta\in\mathfrak{L}.\end{aligned}$$

这证明了 $\alpha_{af+bg}=a\alpha_f+b\alpha_g$. 即 $f\to\alpha_f$ 是 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构. □

记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 上的内积. 设 $\mathscr{A}$ 是实 Euclid 空间 $\mathfrak{L}$ 的线性变换. 任给 $\beta\in\mathfrak{L}$, 易知 $f(\alpha)=(\mathscr{A}(\alpha),\beta),\ \forall\alpha\in\mathfrak{L}$ 是实 Euclid 空间 $\mathfrak{L}$ 上的线性函数. 由上面定理可知 $f(\alpha)$ 唯一决定了 $\mathfrak{L}$ 中一个向量 β_f, 使得

$$(\mathscr{A}(\alpha),\beta)=(\alpha,\beta_f),\quad\alpha\in\mathfrak{L}.\tag{9.3.15}$$

定理 9.3.11 符号同上. 则 $\beta\to\beta_f$ 是实 Euclid 空间 $\mathfrak{L}$ 上的线性变换, 记作 $\mathscr{A}^*$. 即有

$$(\mathscr{A}(\alpha),\beta)=(\alpha,\mathscr{A}^*(\beta)),\quad\forall\alpha,\beta\in\mathfrak{L}.\tag{9.3.16}$$

线性变换 $\mathscr{A}^*$ 称为线性变换 $\mathscr{A}$ 的**共轭变换**.

在实 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1,\alpha_2,\cdots,\alpha_n$. 记线性变换 $\mathscr{A}$ 对应的方阵表示是 A, 则共轭变换 $\mathscr{A}^*$ 对应的方阵表示是 $A^*=A'$.

证 先证 $\mathscr{A}^*$ 是线性变换. 任取 $\beta,\gamma\in\mathfrak{L}$, 任取 a, b 是纯量, 于是

$$\begin{aligned}(\mathscr{A}(\alpha),a\beta+b\gamma)&=a(\mathscr{A}(\alpha),\beta)+b(\mathscr{A}(\alpha),\gamma)=a(\alpha,\mathscr{A}^*(\beta))+b(\alpha,\mathscr{A}^*(\gamma))\\&=(\alpha,a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)),\quad\alpha\in\mathfrak{L}.\end{aligned}$$

另一方面,

$$(\mathscr{A}(\alpha),a\beta+b\gamma)=(\alpha,\mathscr{A}^*(a\beta+b\gamma)),\quad\alpha\in\mathfrak{L}.$$

因此有

$$(\alpha,\mathscr{A}^*(a\beta+b\gamma))=(\alpha,a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)),\quad\forall\alpha\in\mathfrak{L}.$$

这证明了 $\mathscr{A}^*(a\beta+b\gamma)=a\mathscr{A}^*(\beta)+b\mathscr{A}^*(\gamma)$, 即 $\mathscr{A}^*$ 是线性变换.

令 $\mathscr{A}(\alpha_i)=\sum\limits_{j=1}^{n}a_{ji}\alpha_j,\ 1\leqslant i\leqslant n$, 所以 $\mathscr{A}$ 在实 Euclid 空间 $\mathfrak{L}$ 的标准正交基 α_1, $\alpha_2,\cdots,\alpha_n$ 下的方阵表示是

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

于是记 $\mathscr{A}^*(\alpha_i)=\sum\limits_{j=1}^{n}b_{ji}\alpha_j,\ 1\leqslant i\leqslant n$. 则由

$$(\mathscr{A}(\alpha_i),\alpha_j)=(\alpha_i,\mathscr{A}^*(\alpha_j)),\quad i,j=1,2,\cdots,n,$$

有

$$(\sum_{k=1}^{n}a_{ki}\alpha_k,\alpha_j)=(\alpha_i,\sum_{p=1}^{n}b_{pj}\alpha_p).$$

这证明了 $a_{ji}=b_{ij},\ 1\leqslant i,j\leqslant n$. 所以 $\mathscr{A}^*$ 的方阵表示是 $B=A'=A^*$. □

定义 9.3.12 记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**规范的**, 如果有 $\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}$, 即有

$$(\mathscr{A}(\alpha),\mathscr{A}(\beta))=(\mathscr{A}^*(\alpha),\mathscr{A}^*(\beta)),\quad \forall\,\alpha,\beta\in\mathfrak{L}.\tag{9.3.17}$$

定义 9.3.13 记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**正交的**, 如果有 $\mathscr{A}\mathscr{A}^*=\mathscr{A}^*\mathscr{A}=\mathrm{id}$, 即有

$$(\mathscr{A}(\alpha),\mathscr{A}(\beta))=(\alpha,\beta),\quad \forall\,\alpha,\beta\in\mathfrak{L}.\tag{9.3.18}$$

定义 9.3.14 记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**对称的**, 如果有 $\mathscr{A}^*=\mathscr{A}$, 即有

$$(\mathscr{A}(\alpha),\beta)-(\alpha,\mathscr{A}(\beta))=0,\quad \forall\,\alpha,\beta\in\mathfrak{L}.\tag{9.3.19}$$

定义 9.3.15 记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称为**反对称的**, 如果有 $\mathscr{A}^*=-\mathscr{A}$, 即有

$$(\mathscr{A}(\alpha),\beta)+(\alpha,\mathscr{A}(\beta))=0,\quad \forall\,\alpha,\beta\in\mathfrak{L}.\tag{9.3.20}$$

由上面四个定义, 所以定理 9.3.5—9.3.8 可改写为

定理 9.3.16 记 (α,β) 是 n 维实 Euclid 空间 $\mathfrak{L}$ 的内积, $\mathscr{A}$ 是 $\mathfrak{L}$ 上的线性变换. 则有

(1) $\mathscr{A}$ 是规范变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得

$$\begin{cases} \mathscr{A}(\alpha_1) = a_1\alpha_1 - b_1\alpha_2, \\ \mathscr{A}(\alpha_2) = b_1\alpha_1 + a_1\alpha_2, \end{cases} \cdots, \begin{cases} \mathscr{A}(\alpha_{2t-1}) = a_t\alpha_{2t-1} - b_t\alpha_{2t}, \\ \mathscr{A}(\alpha_{2t}) = b_t\alpha_{2t-1} + a_t\alpha_{2t}, \end{cases}$$

$$\mathscr{A}(\alpha_{2t+1}) = \lambda_{2t+1}\alpha_{2t+1}, \quad \cdots, \quad \mathscr{A}(\alpha_n) = \lambda_n\alpha_n.$$

其中 $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_t > 0$. 因此 $\alpha_{2j-1}, \alpha_{2j}$ 线性生成了在 $\mathscr{A}$ 下不变的二维子空间 $\mathfrak{L}_j$, $1 \leqslant j \leqslant t$, α_k 线性生成了在 $\mathscr{A}$ 下不变的一维子空间 $\mathfrak{L}_k$, $2t+1 \leqslant k \leqslant n$, 它们两两正交, 且 $\mathfrak{L}$ 是它们的不变子空间的直接和.

(2) $\mathscr{A}$ 是实正交变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得

$$\begin{cases} \mathscr{A}(\alpha_1) = \cos(\theta_1)\alpha_1 - \sin(\theta_1)\alpha_2, \\ \mathscr{A}(\alpha_2) = \sin(\theta_1)\alpha_1 + \cos(\theta_1)\alpha_2, \end{cases}$$

$$\vdots$$

$$\begin{cases} \mathscr{A}(\alpha_{2t-1}) = \cos(\theta_t)\alpha_{2t-1} - \sin(\theta_t)\alpha_{2t}, \\ \mathscr{A}(\alpha_{2t}) = \sin(\theta_t)\alpha_{2t-1} + \cos(\theta_t)\alpha_{2t}, \end{cases}$$

$$\mathscr{A}(\alpha_{2t+1}) = \alpha_{2t+1}, \quad \cdots, \quad \mathscr{A}(\alpha_{2t+u}) = \alpha_{2t+u},$$

$$\mathscr{A}(\alpha_{2t+u+1}) = -\alpha_{2t+u+1}, \quad \cdots, \quad \mathscr{A}(\alpha_n) = -\alpha_n,$$

其中 $0 < \theta_1 \leqslant \cdots \leqslant \theta_t < \pi$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维及二维不变子空间的直接和.

(3) $\mathscr{A}$ 是实对称变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得

$$\mathscr{A}(\alpha_1) = \lambda_1\alpha_1, \cdots, \mathscr{A}(\alpha_n) = \lambda_n\alpha_n, \tag{9.3.21}$$

其中 $\lambda_1 \geqslant \cdots \geqslant \lambda_n > 0$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维不变子空间的直接和.

(4) $\mathscr{A}$ 是实反对称变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得

$$\begin{cases} \mathscr{A}(\alpha_1) = -b_1\alpha_2, \\ \mathscr{A}(\alpha_2) = b_1\alpha_1, \end{cases} \cdots, \begin{cases} \mathscr{A}(\alpha_{2t-1}) = b_t\alpha_{2t}, \\ \mathscr{A}(\alpha_{2t}) = b_t\alpha_{2t-1}, \end{cases}$$

$$\mathscr{A}(\alpha_{2t+1}) = 0, \quad \cdots, \quad \mathscr{A}(\alpha_n) = 0,$$

其中 $b_1 \geqslant \cdots \geqslant b_t > 0$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维及二维不变子空间的直接和. 所有一维不变子空间的直接和构成实反对称变换 $\mathscr{A}$ 的核 $\ker(\mathscr{A})$.

习 题 9.3

9.3.1 给定一组两两可交换的 n 阶实规范方阵 (不一定有限个). 试证: 存在一个公共的实正交方阵, 将它们同时化为标准形.

9.3.2 设 $S_1, S_2, \cdots, S_m$ 是 m 个 n 阶实对称方阵, 且

$$\sum_{j=1}^{m} S_j^2 = 0.$$

试证: $S_1 = S_2 = \cdots = S_m = 0$.

9.3.3 试证: 实对称方阵的对应不同特征根的特征向量互相正交. 因此, 存在一组标准正交基全由特征向量构成.

9.3.4 试证: 任给 n 阶实对称方阵 S, 则存在 n 阶实对称方阵 S_1, 使得 $S_1^3 = S$. 它唯一吗?

9.3.5 设 S 的秩等于 r, 又 S 是实对称 (实反对称) 方阵, 则存在 r 阶主子式不等于零.

9.3.6 设 A 和 B 都是 n 阶实规范方阵. 试证: AB 是 n 阶实规范方阵当且仅当 BA 也是实规范方阵.

9.3.7 试证: 若 $2n+1$ 阶实正交方阵 O 的行列式是 1, 则必有一个属于特征根 1 的特征向量.

9.3.8 设 O 是 n 阶实正交方阵, 特征根不等于 -1. 试证: n 阶实方阵 $E_n + O$ 非异, 且

$$K = (E_n - O)(E_n + O)^{-1}$$

是实反对称方阵.

反之, 任给实反对称方阵 K. 则 $E_n + K$ 非异, 且

$$O = (E_n - K)(E_n + K)^{-1}$$

是实正交方阵, 且 O 的特征根不等于 -1. 因此上面的关系给出了特征根不等于 -1 的实正交方阵和实反对称方阵间的一一对应, 称为**Cayley 变换**.

9.3.9 设 A 是实**幂等方阵**, 即 $A^2 = A$. 试证: A 只有特征根 0 和 1, 且 A 实正交相似于

$$\begin{pmatrix} E_r & 0^{(r,n-r)} \\ A_1^{(n-r,r)} & 0_{n-r} \end{pmatrix}.$$

由此证明: 存在非异方阵 P, 使得

$$P^{-1}AP = \begin{pmatrix} E_r & 0^{(r,n-r)} \\ 0^{(n-r,r)} & 0_{n-r} \end{pmatrix}.$$

试给出上述性质的几何解释.

9.3.10 设 A 是 n 阶**对合方阵**, 即 $A^2 = E_n$. 试证: 存在 n 阶实正交方阵 O, 使得

$$O'AO = \begin{pmatrix} E_r & 0^{(r,n-r)} \\ A_1^{(n-r,r)} & -E_{n-r} \end{pmatrix}.$$

因此存在 n 阶非异实方阵 P, 使得

$$P^{-1}AP = \begin{pmatrix} E_r & 0^{(r,n-r)} \\ 0^{(n-r,r)} & -E_{n-r} \end{pmatrix}.$$

试给出上述性质的几何解释.

9.3.11 n 阶实对称方阵 S 是幂等的当且仅当它只有特征根 0 和 1, 且记 $S=(a_{ij})$, 则有 $a_{jj}\geqslant 0,\ 1\leqslant j\leqslant n$. 又 $a_{jj}=0$ 当且仅当 $a_{jk}=a_{kj}=0,\ k=1,2,\cdots,n$.

9.3.12 设 A 和 B 是 n 阶实正交方阵, 试证: $n-\operatorname{rank}(A+B)$ 是偶数当且仅当 $\det(A)=\det(B)$.

9.3.13 设 A 是 n 阶实方阵. 试证: A 是对称方阵当且仅当 $A^2=AA'$, 能否给出四种不同的证明?

(1) 用 Cauchy 不等式于 A 的元素上; **(2)** 用迹; **(3)** 用化为非异方阵的情形; **(4)** 用先证明实正交相似于上三角方阵.

9.3.14 设实方阵的秩是 r. 试证: A 能表成 $A=PTO$, 其中 O 是 n 阶实正交方阵, T 是实下三角方阵:

$$T=\begin{pmatrix}\begin{pmatrix}t_{11} & & & \\ t_{21} & t_{22} & & \\ \vdots & \vdots & \ddots & \\ t_{r1} & t_{r2} & \cdots & t_{rr}\end{pmatrix} & 0^{(r,n-r)} \\ T_1^{(n-r,r)} & 0_{n-r}\end{pmatrix},$$

且其中对角元素 $t_{11},\ t_{22},\ \cdots,\ t_{rr}$ 恒大于零, 再 P 是置换方阵, 即是第一种初等方阵的乘积 (见 §3.3).

9.3.15 试证: 行列式等于 1 的三阶实正交方阵 A 必有

$$A=\begin{pmatrix}\cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\varphi & \sin\varphi\\ 0 & -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}\cos\psi & \sin\psi & 0\\ -\sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{pmatrix},$$

其中 $\theta,\ \varphi,\ \psi$ 称为**Euler 角**.

9.3.16 试求最大个数 r, 使得 n 维实 Euclid 空间 $\mathfrak{L}$ 中的 r 个向量 $\beta_1,\ \beta_2,\ \cdots,\ \beta_r$ 有 $(\beta_i,\beta_j)<0,\ 1\leqslant i<j\leqslant r$.

9.3.17 设 $\mathfrak{L}$ 是 n 维实 Euclid 空间. 试证:

(1) 在 $\mathfrak{L}$ 中给定非零向量 α, 则 $\mathfrak{L}$ 上线性函数 $f(\beta)=(\alpha,\beta)$ 连续, 即任取 $\varepsilon>0$, 则存在 $\eta>0$, 使得当 $\sigma\in\mathfrak{L},\ |\sigma-\beta|<\eta$ 时, 有 $|f(\sigma)-f(\beta)|<\varepsilon$;

(2) 对 $\mathfrak{L}$ 中的有限个向量 $\beta_1,\beta_2,\cdots,\beta_s$, 则存在向量 β_0, 使得 $(\beta_0,\beta_j)\neq 0,\ j=1,2,\cdots,s$;

(3) 设 $\mathscr{A}_1,\mathscr{A}_2,\cdots,\mathscr{A}_s$ 是 $\mathfrak{L}$ 上有限多个两两不同的线性变换, 则存在向量 β_0, 使得 $\mathscr{A}_1(\beta_0)$, $\mathscr{A}_2(\beta_0),\ \cdots,\ \mathscr{A}_s(\beta_0)$ 也两两不等.

9.3.18 在 n 维实 Euclid 空间 $\mathfrak{L}$ 中取定有限多个真子空间 $\mathfrak{L}_1,\ \mathfrak{L}_2,\ \cdots,\ \mathfrak{L}_r$. 试证: 存在一组标准正交基 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n$, 使得每个向量 α_i 都不在 $\mathfrak{L}_1\cup\mathfrak{L}_2\cup\cdots\cup\mathfrak{L}_r$ 中.

9.3.19 设 $\mathscr{A}$ 是以 $(*,*)_1$ 是内积的 n 维实 Euclid 空间 $\mathfrak{L}$ 上的线性变换, $\mathscr{B}$ 是以 $(*,*)_2$ 是内积的 n 维实 Euclid 空间 $\mathfrak{L}$ 到 m 维实 Euclid 空间 $\mathfrak{M}$ 上的线性映射. 令 π 是 $\mathfrak{L}$ 到 $\ker(\mathscr{B})$ 内的正交投影, 使得 $\pi\circ\mathscr{A}|_{\ker(\mathscr{B})}$ 是线性同构. 又存在常数 $\lambda_0>0$, 使得

$$\inf_{0\neq\beta\in\mathfrak{M}}\ \supset_{0\neq\alpha\in\mathfrak{L}}\cdot\ \frac{|(\beta,\mathscr{B}(\alpha))_2|}{|\alpha|_1|\beta|_2}\geqslant\lambda_0.$$

试证: $\forall \xi \in \mathfrak{L}$, $\eta \in \mathfrak{M}$, 则唯一存在 $\alpha \in \mathfrak{L}$, $\beta \in \mathfrak{M}$, 使得 $\mathscr{A}(\alpha) + \mathscr{B}^*(\beta) = \xi$, $\mathscr{B}(\alpha) = \eta$, 其中 $(\mathscr{B}^*(\beta), \alpha)_1 = (\beta, \mathscr{B}(\alpha))_2$.

9.3.20 试证: 实规范变换将不变子空间的正交补保持不变.

9.3.21 设 A 和 B 是 $n \times m$ 实矩阵. 试证有不等式:

$$\mathrm{tr}\,(AB' - BA')(AB' - BA')' \leqslant \mathrm{tr}\,(AA')\mathrm{tr}\,(BB'),$$

且等式成立当且仅当存在 n 阶实正交方阵 O_1 和 m 阶实正交方阵 O_2, 使得

$$O_1 A O_2 = \begin{pmatrix} \mathrm{diag}\,(a, b) & 0 \\ 0 & 0 \end{pmatrix}, \qquad O_1 B O_2 = \begin{pmatrix} \begin{pmatrix} 0 & -b \\ a & 0 \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},$$

其中 $a > 0$, $b \geqslant 0$.

§9.4 实对称方阵的特征根 *

由定理 9.3.7, 给定 n 阶实对称方阵 S, 则存在实正交方阵

$$O = (\beta_1, \beta_2, \cdots, \beta_n), \tag{9.4.1}$$

使得

$$SO = O\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n), \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n, \tag{9.4.2}$$

其中 $\beta_1, \beta_2, \cdots, \beta_n$ 是 n 个列向量. 所以有

$$S\beta_i = \lambda_i \beta_i, \quad i = 1, 2, \cdots, n. \tag{9.4.3}$$

即 $\beta_1, \beta_2, \cdots, \beta_n$ 是 n 阶实对称方阵 S 的 n 个实特征向量, 它们构成 $\mathbb{R}^n$ 的一组标准正交基.

定理 9.4.1 (Courant—Fisher 最小最大定理) 设

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \tag{9.4.4}$$

是 n 阶实对称方阵 S 的 n 个实特征根, 则

$$\lambda_j = \min_{v_1, \cdots, v_{j-1} \in \mathbb{R}^n} \max_{(x, v_1) = \cdots = (x, v_{j-1}) = 0,\ (x, x) = 1} x' S x, \tag{9.4.5}$$

其中 $j = 1, 2, \cdots, n$.

证 令存在实正交方阵 $O = (\beta_1, \beta_2, \cdots, \beta_n)$, 使得

$$O' S O = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n) = \Lambda, \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n.$$

记 $y=O'x$, $u_k=O'v_k$, $1\leqslant k<j$, 于是 $x'Sx=y'O'SOy=y'\Lambda y$. 又对 $\mathbb{R}^n$ 的标准内积 $(*,*)$, 有

$$\begin{aligned}\mu_j&=\min_{v_1,\cdots,v_{j-1}\in\mathbb{R}^n}\ \max_{(x,v_1)=\cdots=(x,v_{j-1})=0,\ (x,x)=1}x'Sx\\&=\min_{u_1,\cdots,u_{j-1}\in\mathbb{R}^n}\ \max_{(y,u_1)=\cdots=(y,u_{j-1})=0,\ (y,y)=1}y'\Lambda y.\end{aligned}$$

记 $y=(y_1,\cdots,y_n)'$, $y'y=\sum\limits_{i=1}^n y_i^2=1$, 则 $y'\Lambda y=\sum\limits_{i=1}^n\lambda_i y_i^2$,

$$\mu_j=\min_{u_1,\cdots,u_{j-1}\in\mathbb{R}^n}\ \max_{(y,u_1)=\cdots=(y,u_{j-1})=0,\ (y,y)=1}\sum_{p=1}^n\lambda_p y_p^2.$$

记 $e_j=(\delta_{j1},\delta_{j2},\cdots,\delta_{jn})'$, $1\leqslant j\leqslant n$, 其中 δ_{jk} 是 Kronecker 符号. 取 $u_1=e_1$, $u_2=e_2$, $\cdots$, $u_{j-1}=e_{j-1}$, 则

$$\mu_j\leqslant\max_{(y,e_1)=\cdots=(y,e_{j-1})=0}\sum_{p=1}^n\lambda_p y_p^2=\max_{y_j,\cdots,y_n\in\mathbb{R},\ \sum\limits_{i=j}^n|y_i|^2\leqslant1}\ \sum_{p=j}^n\lambda_p y_p^2\leqslant\lambda_j.$$

另一方面, 任取 $u_1,\cdots,u_{j-1}\in\mathbb{R}^n$. 下面给出 $y\in\mathbb{R}^n$, 其中 $(y,u_k)=0$, $k=1,2,\cdots,j-1$. 这是 $j-1$ 个方程, 而 $y=\sum\limits_{p=1}^j y_pe_p$, 即有 j 个独立未知数. 所以一定有非零解 $y=y^{(0)}=(y_1^{(0)},\cdots,y_n^{(0)})$, $(y^{(0)},y^{(0)})=1$, 于是

$$\begin{aligned}\mu_j&=\min_{u_1,\cdots,u_{j-1}\in\mathbb{R}^n}\ \max_{(y,u_1)=\cdots=(y,u_{j-1})=0,\ (y,y)=1}\sum_{p=1}^j\lambda_p y_p^2\\&\geqslant\min_{u_1,\cdots,u_{j-1}\in\mathbb{R}^n}\sum_{p=1}^j\lambda_p(y_p^{(0)})^2\geqslant\min_{u_1,\cdots,u_{j-1}\in\mathbb{R}^n}\lambda_j=\lambda_j.\end{aligned}$$

这证明了 $\mu_j=\lambda_j$. $j=1,2,\cdots,n$. □

定理 9.4.2 (Courant– Fisher 最大最小定理) 设

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\tag{9.4.6}$$

是 n 阶实对称方阵 S 的 n 个实特征根, 则

$$\lambda_j=\max_{v_1,\cdots,v_{n-j}\in\mathbb{R}^n}\ \min_{(x,v_1)=\cdots=(x,v_{n-j})=0,\ (x,x)=1}x'Sx,\tag{9.4.7}$$

其中 $j=1,2,\cdots,n$.

证 和定理 9.4.1 的证明相同. 我们有

$$\mu_j=\max_{u_1,\cdots,u_{n-j}\in\mathbb{R}^n}\ \min_{(y,u_1)=\cdots=(y,u_{n-j})=0,\ (y,y)=1}\sum_{p=1}^n\lambda_p y_p^2,$$

其中 $j=1,2,\cdots,n$. 取 $u_1=e_{j+1},\cdots,u_{n-j}=e_n$, $y=\sum\limits_{k=1}^{n}y_ke_k$, 则有

$$\mu_j \geqslant \min_{(y,e_{j+1})=\cdots=(y,e_n)=0,\ (y,y)=1} \sum_{p=1}^{n}\lambda_p y_p^2 \geqslant \min_{y_1,\cdots,y_j\in\mathbb{R},\ \sum\limits_{i=1}^{j}|y_j|^2\leqslant 1} \sum_{p=1}^{j}\lambda_p y_p^2 \geqslant \lambda_j.$$

另一方面, 取定 $u_1,\cdots,u_{n-j}$, 取 $y=\sum\limits_{p=j}^{n}y_pe_p$, 使得 $(y,u_k)=0$, $k=1,2,\cdots,n-j$. 它是 $n-j+1$ 个未知数, $n-j$ 个方程构成的线性方程组, 所以一定有非零解 $y=y^{(0)}=(y_1^{(0)},\cdots,y_n^{(0)})$, $(y^{(0)},y^{(0)})=1$. 于是

$$\mu_j \leqslant \max_{u_1,\cdots,u_{n-j}\in\mathbb{R}^n} \sum_{p=j}^{n}\lambda_p (y_p^{(0)})^2 \leqslant \lambda_j.$$

这证明了 $\mu_j=\lambda_j$. $j=1,2,\cdots,n$. □

定理 9.4.3 (单调性定理) 设 A 和 B 是 n 阶实对称方阵, 且 $B\geqslant 0$. 设 A 的特征根是 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$, $A+B$ 的特征根是 $\nu_1\geqslant\nu_2\geqslant\cdots\geqslant\nu_n$, 则有

$$\nu_j \geqslant \lambda_j,\quad j=1,2,\cdots,n. \tag{9.4.8}$$

证 令 $x'Bx\geqslant 0, \forall x\in\mathbb{R}^n$, $(x,x)=1$. 所以有 $x'(A+B)x=x'Ax+x'Bx\geqslant x'Ax$, $x\in\mathbb{R}^n$, $(x,x)=1$. 任取 $v_1,\cdots,v_{j-1}\in\mathbb{R}^n$, 使得 $(v_k,x)=0$, $k=1,2,\cdots,j-1$, 则有

$$\max_{(v_1,x)=\cdots=(v_{j-1},x)=0} x'(A+B)x \geqslant y'Ay,$$

其中 $\forall y\in\mathbb{R}^n$, $(y,y)=1$, $(v_1,y)=\cdots=(v_{j-1},y)=0$. 所以有

$$\max_{(v_1,x)=\cdots=(v_{j-1},x)=0} x'(A+B)x \geqslant \max_{(v_1,x)=\cdots=(v_{j-1},x)=0} x'Ax.$$

同理便证明了

$$\begin{aligned}&\min_{v_1,\cdots,v_{j-1}\in\mathbb{R}^n}\ \max_{(v_1,x)=\cdots=(v_{j-1},x)=0,\ (x,x)=1} x'(A+B)x\\ \geqslant &\min_{v_1,\cdots,v_{j-1}\in\mathbb{R}^n}\ \max_{(v_1,x)=\cdots=(v_{j-1},x)=0,\ (x,x)=1} x'Ax.\end{aligned}$$

此即 $\nu_j\geqslant\lambda_j$, $j=1,2,\cdots,n$. □

引理 9.4.4 (Rayleigh–Ritz 定理) 设 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$ 是 n 阶实对称方阵 S 的 n 个特征根, 则有

$$\lambda_1=\max_{x\in\mathbb{R}^n,\ (x,x)=1} x'Sx,\quad \lambda_n=\min_{x\in\mathbb{R}^n,\ (x,x)=1} x'Sx. \tag{9.4.9}$$

证 今存在 n 阶实正交方阵 O, 使得

$$O'SO = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n) = \Lambda, \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n.$$

于是记 $y = \sum\limits_{j=1}^{n} y_j e_j = O'x$, 则有

$$\begin{aligned}\mu &= \max_{x\in\mathbb{R}^n,\ (x,x)=1} x'Sx = \max_{y\in\mathbb{R}^n,\ (y,y)=1} y'\Lambda y \\ &= \max_{y\in\mathbb{R}^n,\ (y,y)=1} \sum_{p=1}^{n} \lambda_p y_p^2 \geqslant \sum_{p=1}^{n} \lambda_p y_p^2\Big|_{y=e_1} = \lambda_1.\end{aligned}$$

而

$$\max_{y\in\mathbb{R}^n,\ (y,y)=1} \sum_{p=1}^{n} \lambda_p y_p^2 \leqslant \lambda_1.$$

这证明了 $\mu = \lambda_1$.

又

$$\begin{aligned}\mu_0 &= \min_{x\in\mathbb{R}^n,\ (x,x)=1} x'Sx = \min_{y\in\mathbb{R}^n,\ (y,y)=1} y'\Lambda y \\ &= \min_{y\in\mathbb{R}^n,\ (y,y)=1} \sum_{p=1}^{n} \lambda_p y_p^2 \leqslant \sum_{k=1}^{n} \lambda_k y_k^2\Big|_{y=e_n} = \lambda_n.\end{aligned}$$

而

$$\min_{y\in\mathbb{R}^n,\ (y,y)=1} \sum_{k=1}^{n} \lambda_k y_k^2 \geqslant \lambda_n.$$

这证明了 $\mu_0 = \lambda_n$. □

引理 9.4.5 (Weyl 定理) 设 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$, $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ 分别是 n 阶实对称方阵 A 和 B 的 n 个特征根, $\nu_1 \geqslant \nu_2 \geqslant \cdots \geqslant \nu_n$ 是 $A+B$ 的 n 个特征根, 则有

$$\lambda_i + \mu_n \leqslant \nu_i \leqslant \lambda_i + \mu_1. \tag{9.4.10}$$

证 令

$$\frac{x'Ax}{x'x} + \min_{x\neq 0} \frac{x'Bx}{x'x} \leqslant \frac{x'(A+B)x}{x'x} \leqslant \frac{x'Ax}{x'x} + \max_{x\neq 0} \frac{x'Bx}{x'x}, \quad \forall x \in \mathbb{R}^n, \quad x \neq 0.$$

由引理 9.4.4 可知

$$\frac{x'Ax}{x'x} + \mu_n \leqslant \frac{x'(A+B)x}{x'x} \leqslant \frac{x'Ax}{x'x} + \mu_1.$$

由定理 9.4.1 和定理 9.4.2, 所以式 (9.4.10) 成立. □

定理 9.4.6 (严格单调性定理) 设 A 和 B 是 n 阶实对称方阵, 且 $B > 0$. 设 A 和 $A+B$ 的特征根分别是

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n, \quad \mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n, \tag{9.4.11}$$

则有

$$\lambda_j < \mu_j, \quad j = 1, 2, \cdots, n. \tag{9.4.12}$$

证 设正定对称方阵 B 的特征根是 $\rho_1 \geqslant \rho_2 \geqslant \cdots \geqslant \rho_n > 0$. 由引理 9.4.4, 任取 $x \in \mathbb{R}^n$, $(x, x) = 1$, 有

$$x'(A+B)x = x'Ax + x'Bx \geqslant x'Ax + \rho_n.$$

和定理 9.4.3 一样可证 $\mu_j \geqslant \lambda_j + \rho_j > \lambda_j$, $j = 1, 2, \cdots, n$. □

定理 9.4.7 (分离性定理) 设

$$S = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{9.4.13}$$

是 n 阶实对称方阵,

$$S_j = \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jj} \end{pmatrix} \tag{9.4.14}$$

是实对称方阵 S 的 j 阶主子矩阵. 设 S_j 的特征根是

$$\lambda_{j1} \geqslant \lambda_{j2} \geqslant \cdots \geqslant \lambda_{jj}, \quad j = 1, 2, \cdots n, \tag{9.4.15}$$

则有

$$\lambda_{j+1,k+1} \leqslant \lambda_{jk} \leqslant \lambda_{j+1,k}, \quad 1 \leqslant j, k \leqslant n-1. \tag{9.4.16}$$

证 记 $u_j = (x_1, \cdots, x_j)'$, $j = 1, 2, \cdots, n$. 于是 $u_{j+1} = \begin{pmatrix} u_j \\ x_{j+1} \end{pmatrix}$, $j = 1, 2, \cdots, n-1$. 且

$$\begin{pmatrix} u_j \\ 0 \end{pmatrix}' S \begin{pmatrix} u_j \\ 0 \end{pmatrix} = u_j' S_j u_j, \qquad j = 1, 2, \cdots, n.$$

由定理 9.4.1,

$$\lambda_{j+1,k} = \min_{v_1, \cdots, v_{k-1} \in \mathbb{R}^{j+1}} \quad \max_{(u_{j+1}, v_1) = \cdots = (u_{j+1}, v_{k-1}) = 0,\ (u_{j+1}, u_{j+1}) = 1} u_{j+1}' S_{j+1} u_{j+1}.$$

而取定 $v_1, \cdots, v_{k-1} \in \mathbb{R}^{j+1}$, 则有

$$\max_{(u_{j+1}, v_1) = \cdots = (u_{j+1}, v_{k-1}) = 0,\ (u_{j+1}, u_{j+1}) = 1} u_{j+1}' S_{j+1} u_{j+1}$$

$$\geqslant \max_{\left(\begin{pmatrix} u_j \\ 0 \end{pmatrix}, v_1\right) = \cdots = \left(\begin{pmatrix} u_j \\ 0 \end{pmatrix}, v_{k-1}\right) = 0,\ (u_j, u_j) = 1} u_j' S_j u_j.$$

记 $v_p = \begin{pmatrix} w_p \\ \widetilde{w_p} \end{pmatrix}$, $p = 1, 2, \cdots, k-1$, 于是证明了

$$\begin{aligned}
&\max_{(u_{j+1},v_1)=\cdots=(u_{j+1},v_{k-1})=0,\ (u_{j+1},u_{j+1})=1} u'_{j+1}S_{j+1}u_{j+1} \\
\geqslant &\max_{(u_j,w_1)=\cdots=(u_j,w_{k-1})=0,\ (u_j,u_j)=1} u'_jS_ju_j \geqslant \lambda_{jk}.
\end{aligned}$$

后一等式由定理 9.4.1 可知. 由于 $w_1, \cdots, w_{k-1}$ 可以任取, 所以证明了 $\lambda_{j+1,k} \geqslant \lambda_{j,k}$.

再用定理 9.4.2, 今

$$\lambda_{j+1,k+1} = \max_{v_1,\cdots,v_{j-k}\in\mathbb{R}^{j+1}} \min_{(u_{j+1},v_1)=\cdots=(u_{j+1},v_{j-k})=0,\ (u_{j+1},u_{j+1})=1} u'_{j+1}S_{j+1}u_{j+1},$$

而取定 $v_1, \cdots, v_{j-k} \in \mathbb{R}^{j+1}$, 由定理 9.4.2, 有

$$\begin{aligned}
&\min_{(u_{j+1},v_1)=\cdots=(u_{j+1},v_{j-k})=0,\ (u_{j+1},u_{j+1})=1} u'_{j+1}S_{j+1}u_{j+1} \\
\leqslant &\min_{(\begin{pmatrix} u_j \\ 0 \end{pmatrix},v_1)=\cdots=(\begin{pmatrix} u_j \\ 0 \end{pmatrix},v_{j-k})=0,\ (u_j,u_j)=1} u'_jS_ju_j \\
\leqslant &\min_{(u_j,w_1)=\cdots=(u_j,w_{j-k})=0,\ (u_j,u_j)=1} u'_jS_ju_j \leqslant \lambda_{jk}.
\end{aligned}$$

于是

$$\lambda_{j+1,k+1} = \max_{v_1,\cdots,v_{j-k}\in\mathbb{R}^{j+1}} \min_{(u_{j+1},v_1)=\cdots=(u_{j+1},v_{j-k})=0,(u_{j+1},u_{j+1})=1} u'_{j+1}S_{j+1}u_{j+1} \leqslant \lambda_{jk}.$$

□

习 题 9.4

9.4.1 设 n 阶实对称方阵 $S = (a_{ij}) \geqslant 0$, 且有 $\sum\limits_{j=1}^{n} a_{ij} = 0$, $1 \leqslant i \leqslant n$, 则有

(1) 设 S 的特征根是 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n = 0$, 则 $S - \lambda_{n-1}\left(E - \frac{1}{n}ee'\right) \geqslant 0$, 其中 $e = \sum\limits_{j=1}^{n} e_j = (1, 1, \cdots, 1)'$;

(2) 试证: $\lambda_{n-1} \leqslant \left(\frac{n}{n-1}\right)\min(a_{11}, \cdots, a_{nn})$.

§9.5 实线性不等式 *

在 m 维实线性空间 $\mathbb{R}^m$ 中取定标准正交基

$$e_j = (\delta_{j1}, \delta_{j2}, \cdots, \delta_{jm})', \quad 1 \leqslant j \leqslant m, \tag{9.5.1}$$

其中 δ_{jk} 是 Kronecker 符号. 标准内积 $(\alpha,\beta)_m$ 是

$$(\sum_{i=1}^{m} x_i e_i, \sum_{j=1}^{m} y_j e_j)_m = \sum_{k=1}^{m} x_k y_k. \tag{9.5.2}$$

在 n 维实向量空间 $\mathbb{R}^n$ 中取定标准正交基

$$f_j = (\delta_{j1}, \delta_{j2}, \cdots, \delta_{jn})', \quad 1 \leqslant j \leqslant n, \tag{9.5.3}$$

其中 δ_{jk} 是 Kronecker 符号. 标准内积 $(\alpha,\beta)_n$ 是

$$(\sum_{i=1}^{n} x_i f_i, \sum_{j=1}^{n} y_j f_j)_n = \sum_{k=1}^{n} x_k y_k. \tag{9.5.4}$$

则 $\mathbb{R}^m$ 和 $\mathbb{R}^n$ 都是实 Euclid 空间. 又 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 内的线性映射 $\mathscr{A}$ 和 $n\times m$ 实矩阵之间有一个自然的一一对应, 使得

$$\mathscr{A}(e_i) = \sum_{j=1}^{n} a_{ji} f_j, \quad 1 \leqslant i \leqslant m, \tag{9.5.5}$$

其中

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}. \tag{9.5.6}$$

在这一节中, 线性映射 $\mathscr{A}$ 和矩阵 A 间的关系就按上面方式约定. 因此, 有

$$\mathbb{R}^n = \mathscr{A}(\mathbb{R}^m) \oplus (\mathscr{A}(\mathbb{R}^m))^{\perp} \tag{9.5.7}$$

是空间直接和.

记

$$P_n = \{\, \alpha = (a_1, a_2, \cdots, a_n)' \in \mathbb{R}^n \mid a_1 \geqslant 0, \cdots, a_n \geqslant 0 \,\}, \tag{9.5.8}$$

$$P_n^0 = \{\, \alpha = (a_1, a_2, \cdots, a_n)' \in \mathbb{R}^n \mid a_1 > 0, \cdots, a_n > 0 \,\}, \tag{9.5.9}$$

则 P_n 是 $\mathbb{R}^n$ 中**第一卦限**, P_n^0 是 $\mathbb{R}^n$ 中的**内点集**.

定义 9.5.1 m 个未知数 $x_1, \cdots, x_m$ 的关系

$$\sum_{j=1}^{m} a_{ij} x_j - a_i \geqslant 0, \quad i = 1, 2, \cdots, n \tag{9.5.10}$$

称为**线性不等式**.

下面将线性不等式用矩阵语言来描述. 记

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}. \tag{9.5.11}$$

于是线性不等式可改写为

$$Ax - \alpha \in P_n. \tag{9.5.12}$$

线性不等式的问题是求通解.

定义 9.5.2 如果存在 $n \times 1$ 实矩阵 $x^{(0)} = (x_1^{(0)}, \cdots, x_n^{(0)})'$, 使得

$$\sum_{j=1}^{m} a_{ij} x_j^{(0)} - a_i \geqslant 0, \quad i = 1, 2, \cdots, n, \tag{9.5.13}$$

则 $x^{(0)}$ 称为线性不等式

$$Ax - \alpha \in P_n \tag{9.5.14}$$

的一组解.

定理 9.5.3 记 $\mathscr{A}$ 是 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 内的线性映射, 它的矩阵表示 A 是 $n \times m$ 实矩阵. 记 α 是 $n \times 1$ 实矩阵, x 是 m 个独立未知数构成的 $m \times 1$ 实矩阵. 则线性不等式

$$Ax - \alpha \in P_n \tag{9.5.15}$$

有解当且仅当

$$(P_n \cap (\mathscr{A}(\mathbb{R}^m))^{\perp}, \alpha)_n \leqslant 0. \tag{9.5.16}$$

证 设线性不等式 $Ax - \alpha \in P_n$ 有解 $x = x^{(0)}$, 则 $Ax^{(0)} - \alpha \in P_n$. 任取 $\beta \in P_n \cap (\mathscr{A}(\mathbb{R}^m))^{\perp}$, 则由 $\beta \in P_n$ 有 $(\beta, Ax^{(0)} - \alpha)_n \geqslant 0$. 因此 $(\beta, \alpha)_n \leqslant (\beta, Ax^{(0)})_n$. 由 $\beta \in (\mathscr{A}(\mathbb{R}^m))^{\perp}$, 所以 $(\beta, Ax^{(0)})_n = 0$. 这证明了 $(\beta, \alpha)_n \leqslant 0$. 由 β 任取, 所以 $(P_n \cap (\mathscr{A}(\mathbb{R}^m))^{\perp}, \alpha)_n \leqslant 0$.

反之, 用反证法. 即证若不存在 $x^{(0)} \in \mathbb{R}^m$, 使得 $Ax - \alpha \in P_n$, 则必存在 $\beta \in P_n \cap (\mathscr{A}(\mathbb{R}^m))^{\perp}$, 使得 $(\beta, \alpha)_n > 0$.

由条件, 不存在 $x^{(0)} \in \mathbb{R}^m$, 使得 $Ax^{(0)} - \alpha \in P_n$, 即

$$(\mathscr{A}(\mathbb{R}^m) - \alpha) \cap P_m = \emptyset.$$

在 $\mathbb{R}^n$ 的子空间 $(\mathscr{A}(\mathbb{R}^m))^{\perp}$ 中取基 $\beta_1, \beta_2, \cdots, \beta_r$, 在 $\mathbb{R}^r$ 中取子集合

$$\mathfrak{Q} = \{\, \xi = ((\gamma, \beta_1)_n, \cdots, (\gamma, \beta_r)_n)' \in \mathbb{R}^r \mid \forall \gamma \in P_n \,\},$$

于是 $\mathfrak{Q}$ 是以原点为顶点的闭凸锥. (定义: $\mathbb{R}^r$ 中子集 $\mathfrak{S}$ 称为**凸集**, 如果 $\xi, \eta \in \mathfrak{S}$, 则 $t\xi + (1-t)\eta \in \mathfrak{S}$, $1 \leqslant t \leqslant 1$; $\mathfrak{S}$ 称为以原点为顶点的**锥**, 如果 $\xi \in \mathfrak{S}$, 则 $t\xi \in \mathfrak{S}$, 对一切非负实数 t 成立; $\mathfrak{S}$ 称为**闭集**, 如果 $\mathfrak{S}$ 中一切收敛序列的极限点仍在 $\mathfrak{S}$ 中.) 证明如下:

今 $\xi, \eta \in \mathfrak{Q}$, 所以存在 $\gamma, \sigma \in P_n$, 使得

$$\xi = ((\gamma, \beta_1)_n, \cdots, (\gamma, \beta_r)_n)', \quad \eta = ((\sigma, \beta_1)_n, \cdots, (\sigma, \beta_r)_n)'.$$

于是

$$t\xi+(1-t)\eta=((t\gamma+(1-t)\sigma,\beta_1)_n,\cdots,(t\gamma+(1-t)\sigma,\beta_r)_n)'.$$

当 $0\leqslant t\leqslant 1$, 由 $\gamma,\sigma\in P_n$ 可知 $t\gamma+(1-t)\sigma\in P_n$. 所以 $t\xi+(1-t)\eta\in\Omega$. 这证明了 Ω 是凸集. 又当 $t\geqslant 0$,

$$t\xi=((t\gamma,\beta_1)_n,\cdots,(t\gamma,\beta_r)_n)'.$$

由 $\gamma\in P_n$ 可知 $t\gamma\in\Omega$. 这证明了 Ω 是以原点为顶点的锥. 最后由 P_n 是闭集, 所以 Ω 也是闭集, 至此证明了 Ω 是闭凸锥.

而条件 $(\mathscr{A}(\mathbb{R}^m)-\alpha)\cap P_n=\emptyset$ 推出

$$\xi=((-\alpha,\beta_1)_n,\cdots,(-\alpha,\beta_r)_n)'\notin\Omega.$$

事实上, 若 $\xi\in\Omega$, 即

$$((-\alpha,\beta_1)_n,\cdots,(-\alpha,\beta_r)_n)'\in\Omega,$$

则存在 $\gamma\in P_n$, 使得

$$((-\alpha,\beta_1)_n,\cdots,(-\alpha,\beta_r)_n)'=((\gamma,\beta_1)_n,\cdots,(\gamma,\beta_r)_n)',$$

即有

$$(\gamma+\alpha,\beta)_n=0,\quad j=1,2,\cdots,r.$$

由于 $\beta_1,\cdots,\beta_r$ 是 $(\mathscr{A}(\mathbb{R}^m))^{\perp}$ 的基, 所以 $\gamma+\alpha\in\mathscr{A}(\mathbb{R}^m)$. 因此, $\gamma\in(\mathscr{A}(\mathbb{R}^m)-\alpha)\cap P_n$, 这和 $(\mathscr{A}(\mathbb{R}^m)-\alpha)\cap P_n=\emptyset$ 矛盾. 所以证明了 $\xi\in\Omega$.

现在在 $\mathbb{R}^r$ 中有闭凸锥 Ω 及不在 Ω 中的点 ζ. 由图形可以看出: 在 $\mathbb{R}^r$ 中存在超平面:

$$\pi:\quad \sum_{j=1}^{r}c_jz_j=c>0,$$

使得 π 和 Ω 及 ξ 都不交, 且将它们隔开. 由于 $\mathbb{R}^r$ 中超平面将 $\mathbb{R}^r$ 分成两部分, 一部分使得 $\sum\limits_{j=1}^{r}c_jz_j-c>0$, 另一部分使得 $\sum\limits_{j=1}^{r}c_jz_j-c<0$. 又由于原点在闭凸锥 Ω 中, 而当 $z_1=\cdots=z_r=0$ 时, $\sum\limits_{j=1}^{r}c_jz_j-c<0$. 这证明了将 Ω 中点 $\zeta=((\gamma,\beta_1)_n,\cdots,(\gamma,\beta_r)_n)'$ 代入, 仍有

$$\sum_{j=1}^{r}c_j(\gamma,\beta_j)_n-c<0,\quad \forall\gamma\in P_n.$$

又 ζ 在超平面 π 的另一侧. 所以有

$$\sum_{j=1}^{r}c_j(-\alpha,\beta_j)_n-c>0.$$

记 $\beta=-\sum\limits_{j=1}^{r}c_j\beta_j$, 则有

$$(\gamma,\beta)_n+c>0,\quad (\alpha,\beta)_n>c>0,\quad \forall\,\gamma\in P_n.$$

这证明了 $(\alpha,\beta)_n>0$.

余下证 $\beta\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}$. 由于 $\beta_1,\cdots,\beta_r$ 是 $(\mathscr{A}(\mathbb{R}^m))^{\perp}$ 的基, 所以 $\beta\in(\mathscr{A}(\mathbb{R}^m))^{\perp}$. 又 $\beta\in P_n$. 事实上, 记 $\beta=(b_1,b_2,\cdots,b_n)'$, $\gamma_j=\lambda_j e_j$, $\lambda_j\geqslant 0$, $1\leqslant j\leqslant n$. 则 $\gamma_j\in P_n$, 而

$$0<c+(\gamma_j,\beta)_n=c+\lambda_j b_j,\quad \lambda_j\geqslant 0,\quad 1\leqslant j\leqslant n.$$

若某个 $b_{j_0}<0$, 取 $\lambda_j>\dfrac{c}{-b_{j_0}}$, 则 $c+\lambda_j b_{j_0}<c-c=0$, 这和条件 $c+\lambda_j b_{j_0}>0$ 矛盾. 所以证明了 β 的坐标都非负, 即 $\beta\in P_n$. 至此证明了存在 $\beta\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}$, 使得 $(\alpha,\beta)_n>0$. □

定理 9.5.4 设 A 是 $n\times m$ 实矩阵, x 是 m 个独立未知数构成的 $m\times 1$ 实矩阵, 则线性不等式

$$Ax-\alpha\in P_n^0 \tag{9.5.17}$$

有解当且仅当

$$(P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}-\{0\},\alpha)_n<0. \tag{9.5.18}$$

特别, 线性不等式

$$Ax\in P_n^0 \tag{9.5.19}$$

有解当且仅当

$$P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}=\{0\}. \tag{9.5.20}$$

证 若存在 $x^{(0)}\in\mathbb{R}^m$, 使得 $Ax^{(0)}-\alpha\in P_n^0$. 任取 $\beta\in P_n-\{0\}$, 则有

$$(Ax^{(0)}-\alpha,\beta)_n=(Ax^{(0)},\beta)_n-(\alpha,\beta)_n>0.$$

当 $\beta\in(P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp})-\{0\}$, 则有 $\beta\in\mathscr{A}(\mathbb{R}^m))^{\perp}$, 于是 $(Ax^{(0)},\beta)_n=0$. 因此证明了

$$(\alpha,\beta)_n<0,\quad \forall\,\beta\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}=\{0\}.$$

反之, 若 $Ax-\alpha\in P_n^0$ 无解, 记 $e=\sum\limits_{j=1}^{n}e_j$, 则 $Ax-\alpha-\dfrac{1}{p}e\in P_n$ 无解, 其中 $p=1,2,\cdots$. 用定理 9.5.3 的证明可知: 存在 $\beta^{(p)}\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}$, 使得 $(\beta^{(p)},\alpha+\dfrac{1}{p}e)_n>0$, $p=1,2,\cdots$. 显然, 不妨设 $\beta^{(p)}$ 是单位向量, 因此它们都在单位球面

$$S=\{\,\beta\in\mathbb{R}^n\mid \beta'\beta=1\,\}$$

上. 所以 S 上的序列 $\{\beta^{(p)}\}_{p=1}^{\infty}$ 中存在收敛子序列 $\{\beta^{(p_j)}\}_{p=1}^{\infty}$, 使得它的极限是 β_0. 即 β_0 仍在单位球面 S 上, 即是单位向量. 又 $\lim\limits_{j\to\infty}\beta^{(p_j)}=\beta_0$.

另一方面, 由 $\beta^{(p_j)}\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}$ 及 P_n, $(\mathscr{A}(\mathbb{R}^m))^{\perp}$ 都是 $\mathbb{R}^n$ 中的闭子集, 所以极限 $\beta_0\in P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}$. 而 $(\beta^{(p_j)},\alpha+\dfrac{1}{p}e)_n>0$, $j=1,2,\cdots$. 当 $j\to\infty$, 便有 $(\beta_0,\alpha)_n\geqslant 0$. 因此, 在 $P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}-\{0\}$ 中存在向量 β_0, 使得 $(\beta_0,\alpha)_n\geqslant 0$. 这证明了若 $Ax-\alpha\in P_n^0$ 无解, 则存在 β_0, 使得 $(\beta_0,\alpha)_n\geqslant 0$. 所以当 $P_n\cap(\mathscr{A}(\mathbb{R}^m))^{\perp}-\{0\},\alpha)_n<0$, 则线性不等式 $Ax-\alpha\in P_n^0$ 有解. □

定理 9.5.5 设 A 是 $n\times m$ 实矩阵, x 是由 m 个独立自变量构成的 $m\times 1$ 实矩阵. 则线性不等式

$$Ax\in P_n \tag{9.5.21}$$

在 $P_m-\{0\}$ 中无解当且仅当存在 $u\in P_n^0$, 使得

$$-A'u\in P_m^0. \tag{9.5.22}$$

证 设 $Ax\in P_n$ 在 $P_m-\{0\}$ 中有解 v. 且存在 $u\in P_n^0$, 使得 $-A'u\in P_m^0$. 于是 $v'(A'u)=u'(Av)\geqslant 0$. 但是, $-A'u\in P_m^0$, $v\in P_m-\{0\}$, 所以 $v'(A'u)=(-v)'(-A'u)<0$. 这导出矛盾. 因此当存在 $u\in P_n^0$, 使得 $-A'u\in P_m^0$, 则 $Ax\in P_n$ 在 $P_m-\{0\}$ 中必无解.

反之, 若不存在 $u\in P_n^0$, 使得 $-A'u\in P_m^0$, 则

$$\begin{pmatrix}-A'u\\ u\end{pmatrix}=\begin{pmatrix}-A'\\ E_n\end{pmatrix}u\in P_{n+m}^0$$

无解. 记 $\mathscr{B}$ 是 $\mathbb{R}^n$ 到 $\mathbb{R}^{n+m}$ 内的线性映射, 使得对 $\mathbb{R}^n$ 和 $\mathbb{R}^{n+m}$ 的标准基, $\mathscr{B}$ 所对应的矩阵表示是 $(n+m)\times n$ 矩阵 $\begin{pmatrix}-A'\\ E_n\end{pmatrix}$, 其中 E_n 是 n 阶单位方阵. 于是 $\mathscr{A}$ 的共轭变换 $\mathscr{A}^*$ 所对应的矩阵表示是 $n\times(n+m)$ 矩阵

$$\begin{pmatrix}-A'\\ E_n\end{pmatrix}'=\begin{pmatrix}-A & E_n\end{pmatrix}.$$

而任取 $\xi\in\mathbb{R}^{n+m}$, $\tau\in\mathbb{R}^n$, 则有

$$(\xi,\mathscr{A}(\tau))_{n+m}=(\mathscr{A}^*(\xi),\tau)_n.$$

由定理 9.5.4, $\begin{pmatrix}-A'\\ E_n\end{pmatrix}u\in P_{n+m}^0$ 无解当且仅当 $P_{n+m}\cap(\mathscr{A}(\mathbb{R}^n))^{\perp}\neq 0$. 所以存在 $0\neq v\in P_{n+m}\cap(\mathscr{A}(\mathbb{R}^n))^{\perp}$. 记

$$v=\begin{pmatrix}w_1\\ w_2\end{pmatrix}\neq 0,$$

其中 $w_1 \in P_m$, $w_2 \in P_n$, 则

$$\left(\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \mathscr{A}(\mathbb{R}^n)\right)_{n+m} = 0.$$

因此,

$$(\mathscr{A}^* \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \mathbb{R}^n)_n = 0.$$

所以 $\mathscr{A}^* \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 0$. 即

$$\begin{pmatrix} -A & E_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 0.$$

因此, $w_2 = Aw_1$. 由于

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \neq 0,$$

所以 $w_1 \neq 0$. 因为若 $w_1 = 0$, 则 $w_2 = Aw_1 = 0$, 这导出矛盾. 由 $w_1 \neq 0$ 可知

$$w_1 \in P_m - \{0\}, \quad w_2 = Aw_1 \in P_n.$$

这证明了线性不等式 $Ax \in P_n$ 在 $P_n - \{0\}$ 中有解, 所以这又导出矛盾. □

习　　题　9.5

9.5.1　设 A 是 n 阶整数方阵:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

整数方阵 A 称为**广义 Cartan 矩阵**, 如果它适合条件: **(1)** $a_{ii} = 2$, $1 \leqslant i \leqslant n$; **(2)** $a_{ij} \leqslant 0$, $i \neq j$, $1 \leqslant i, j \leqslant n$; **(3)** $a_{ij} = 0$ 当且仅当 $a_{ji} = 0$.

广义 Cartan 矩阵 A 称为**不可分解的**, 如果不存在置换方阵 P, 使得 $P'AP$ 是准对角方阵.

设 A 是 n 阶不可分解广义 Cartan 矩阵, 试证

(1)　由 $u \in \mathbb{R}^n$, $u \geqslant 0$, $Au \geqslant 0$, 则 $u = 0$ 或 $u > 0$;

(2)　记 $\mathfrak{K}_A = \{ v \in \mathbb{R}^n \mid Av \geqslant 0 \}$, 则 $\mathfrak{K}_A$ 是 $\mathbb{R}^n$ 中以原点是顶点的闭凸锥. 而且 $\mathfrak{K}_A \cap \partial P_n = \{0\}$, $\mathfrak{K}_A \cap P_n \subset P_n^0 \cup \{0\}$, 其中 $\partial P_n = P_n - P_n^0$. 当 $\mathfrak{K}_A \cap P_n^0 \neq \emptyset$, $\mathfrak{K}_A - P_n \neq \emptyset$, 则 $\dim(\mathfrak{K}_A) = 1$, 而且有

$$\begin{aligned} \mathfrak{K}_A &= \{ v \in \mathbb{R}^n \mid Av \geqslant 0 \} \\ &= \{ v \in \mathbb{R}^n \mid Av = 0, \text{其中 } v > 0, \text{ 或 } v = 0, \text{ 或 } v < 0 \}; \end{aligned}$$

(3)　A 必是下面三种情形之一:

(i) 有限型. 存在 $u \in \mathbb{R}^n$, $u > 0$, 使得 $Au > 0$. 这时 $\det(A) \neq 0$. 而且任取 $v \in \mathbb{R}^n$, 使得 $Av \geqslant 0$, 则必有 $v > 0$ 或 $v = 0$.

(ii) 仿射型. 存在 $u \in \mathbb{R}^n$, $u > 0$, 使得 $Au = 0$. 这时 $\det(A) = 0$, $\operatorname{rank}(A) = n - 1$. 而且任取 $v \in \mathbb{R}^n$, 使得 $Av \geqslant 0$, 则必有 $Av = 0$, 而且 $v > 0$ 或 $v = 0$, 或 $v < 0$.

(iii) 不定型. 存在 $u \in \mathbb{R}^n$, $u > 0$, 使得 $Au < 0$. 这时任取 $v \in \mathbb{R}^n$, $v \geqslant 0$, 使得 $Av \geqslant 0$, 则必有 $v = 0$.

第十章　二次型分类

§10.1　对称方阵在相合下的标准形

(一)　域 $\mathbb{F}$ 上二次型化标准形的配方法

例 10.1.1　试将二次型

$$\varphi(x) = x_1^2 + x_2^2 + x_3^2 - 2x_4^2 - 2x_1x_2 + 2x_1x_3 - 2x_1x_4 + 2x_2x_3 - 4x_2x_4$$

在域 $\mathbb{F} = \mathbb{R}$ 和 $\mathbb{F} = \mathbb{C}$ 的情形化为标准形.

解　将二次型 $\varphi(x)$ 看作 x_1 的二次多项式, 它可以改写为

$$\begin{aligned}\varphi(x) &= x_1^2 + 2x_1(-x_2 + x_3 - x_4) + (x_2^2 + x_3^2 - 2x_4^2 + 2x_2x_3 - 4x_2x_4)\\ &= [x_1 + (-x_2 + x_3 - x_4)]^2 - (-x_2 + x_3 - x_4)^2\\ &\quad + (x_2^2 + x_3^2 - 2x_4^2 + 2x_2x_3 - 4x_2x_4)\\ &= (x_1 - x_2 + x_3 - x_4)^2 - 3x_4^2 + 4x_2x_3 - 6x_2x_4 + 2x_3x_4,\end{aligned}$$

作坐标变换

$$\begin{cases} y_1 = x_1 - x_2 + x_3 - x_4, \\ y_2 = x_2, \\ y_3 = x_3, \\ y_4 = x_4, \end{cases}$$

于是二次型 $\varphi(x)$ 化为二次型

$$\varphi_1(y) = y_1^2 + (-3y_4^2 + 4y_2y_3 - 6y_2y_4 + 2y_3y_4),$$

括号内是不包含变数 y_1 的二次型. 对它施行同样的计算, 即再进行配方, 便有

$$\begin{aligned}\varphi_1(y) &= y_1^2 - 3[y_4^2 + 2(y_2 - \frac{1}{3}y_3)y_4] + 4y_2y_3\\ &= y_1^2 - 3[y_4^2 + 2(y_2 - \frac{1}{3}y_3)y_4 + (y_2 - \frac{1}{3}y_3)^2] + 4y_2y_3 + 3(y_2 - \frac{1}{3}y_3)^2\\ &= y_1^2 - 3(y_4 + y_2 - \frac{1}{3}y_3)^2 + 3(y_2 + \frac{1}{3}y_3)^2.\end{aligned}$$

再作坐标变换

$$\begin{cases} z_1 = y_1, \\ z_2 = \quad y_2 - \frac{1}{3}y_3 + y_4, \\ z_3 = \quad y_2 + \frac{1}{3}y_3, \\ z_4 = \qquad\qquad y_4, \end{cases}$$

于是二次型 $\varphi_1(y)$ 化为二次型

$$\varphi_2(z) = z_1^2 - 3z_2^2 + 3z_3^2.$$

最后, 作坐标变换

$$\begin{cases} u_1 = z_1, \\ u_2 = \quad \delta_2 z_3, \\ u_3 = \qquad \delta_3 z_2, \\ u_4 = \qquad\qquad z_4, \end{cases}$$

于是二次型化为标准二次型

$$\varphi_0(u) = u_1^2 + 3\delta_2^{-2}u_2^2 - 3\delta_3^{-2}u_3^2.$$

设域 $\mathbb{F} = \mathbb{C}$, 令 $\delta_2 = \sqrt{3}$, $\delta_3 = \sqrt{-3}$, 则标准二次型是

$$\varphi_0(u) = u_1^2 + u_2^2 + u_3^2.$$

设域 $\mathbb{F} = \mathbb{R}$, 令 $\delta_2 = \sqrt{3}$, $\delta_3 = \sqrt{3}$, 则标准二次型是

$$\varphi_0(u) = u_1^2 + u_2^2 - u_3^2.$$

设域 $\mathbb{F} = \mathbb{Q}$, 令 $\delta_2 = \delta_3 = 1$, 则标准二次型是

$$\varphi_0(u) = u_1^2 + 3u_2^2 - 3u_3^2.$$

例 10.1.2　再考虑二次型

$$\varphi(x) = x_1x_2 + x_2x_3.$$

在域 $\mathbb{F}$ 的情形化为标准形. 由于这个二次型对每个变数都是一次多项式, 所以上面的配方方法用不上.

先作坐标变换

$$\begin{cases} y_1 = \dfrac{1}{2}x_1 + \dfrac{1}{2}x_2, \\ y_2 = \dfrac{1}{2}x_1 - \dfrac{1}{2}x_2, \\ y_3 = \qquad\qquad x_3, \end{cases}$$

则有

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = \qquad y_3, \end{cases}$$

于是二次型 $\varphi(x)$ 化为二次型

$$\begin{aligned} \varphi_1(y) &= (y_1+y_2)(y_1-y_2)+(y_1-y_2)y_3 = (y_1^2+y_1y_3)-(y_2^2+y_2y_3) \\ &= (y_1+\frac{1}{2}y_3)^2-(y_2+\frac{1}{2}y_3)^2. \end{aligned}$$

所以, 作坐标变换

$$\begin{cases} z_1 = y_1 + \dfrac{1}{2}y_3, \\ z_2 = \qquad y_2 + \dfrac{1}{2}y_3, \\ z_3 = \qquad\qquad y_3, \end{cases} \tag{10.1.1}$$

由于它对应的方阵

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \tag{10.1.2}$$

是非异方阵, 所以这个变换是坐标变换. 经过这个坐标变换, 二次型 $\varphi_1(y)$ 化为标准形

$$\varphi_0(z) = z_1^2 - z_2^2. \tag{10.1.3}$$

由此可见, 平方项不出现的二次型, 在经过形如式 (10.1.1) 的坐标变换后, 就出现了平方项, 于是可以进行配方.

最后, 作坐标变换

$$\begin{cases} u_1 = z_1, \\ u_2 = \quad \delta_2 z_2, \\ u_3 = \qquad z_3, \end{cases}$$

于是二次型化为标准二次型

$$\varphi_0(u) = u_1^2 - \delta_2^{-2} u_2^2.$$

设域 $\mathbb{F} = \mathbb{C}$, 令 $\delta_2 = \sqrt{-1}$, 则标准二次型是 $\varphi_0(u) = u_1^2 + u_2^2$. 设域 $\mathbb{F} = \mathbb{R}$, 令 $\delta_2 = 1$, 则标准二次型是 $\varphi_0(u) = u_1^2 - u_2^2$. 设域 $\mathbb{F} = \mathbb{Q}$, 令 $\delta_2 = 1$, 则标准二次型是 $\varphi_0(u) = u_1^2 - u_2^2$.

现在来叙述配方法的一般原理:

任给域 $\mathbb{F}$ 上一个二次型 $\varphi(x)$. 假设 $\varphi(x) \neq 0$, 即记

$$\varphi(x) = \varphi(x_1, x_2, \cdots, x_n) = \sum_{j,k=1}^{n} a_{jk} x_j x_k = \sum_{j=1}^{n} a_{jj} x_j^2 + 2 \sum_{1 \leqslant j < k \leqslant n} a_{jk} x_j x_k = x' S x, \tag{10.1.4}$$

其中

$$S = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = S', \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{10.1.5}$$

第一步 如果 $a_{11} = a_{22} = \cdots = a_{nn} = 0$. 因为对称方阵 S 不是零, 所以存在元素 $a_{jk} \neq 0$. 例如 $a_{12} \neq 0$. 将二次型 $\varphi(x)$ 写成

$$\varphi(x) = 2a_{12} x_1 x_2 + 2 \sum_{k=3}^{n} (a_{1k} x_1 + a_{2k} x_2) x_k + 2 \sum_{3 \leqslant j < k \leqslant n} a_{jk} x_j x_k. \tag{10.1.6}$$

作坐标变换

$$\begin{cases} y_1 = \dfrac{1}{2} x_1 + \dfrac{1}{2} x_2, \\ y_2 = \dfrac{1}{2} x_1 - \dfrac{1}{2} x_2, \\ y_j = \qquad\qquad x_j, \qquad j = 3, 4, \cdots, n. \end{cases} \tag{10.1.7}$$

则二次型 $\varphi(x)$ 化为二次型

$$\begin{aligned} \varphi_1(y) &= 2a_{12}(y_1 + y_2)(y_1 - y_2) + 2 \sum_{k=3}^{n} [a_{1k}(y_1 + y_2) + a_{2k}(y_1 - y_2)] y_k \\ &\quad + 2 \sum_{3 \leqslant j < k \leqslant n} a_{jk} y_j y_k \\ &= 2a_{12} y_1^2 - 2a_{12} y_2^2 + 2 \sum_{k=3}^{n} [(a_{1k} + a_{2k}) y_1 y_k + 2 \sum_{k=3}^{n} [(a_{1k} - a_{2k}) y_2 y_k \\ &\quad + 2 \sum_{3 \leqslant j < k \leqslant n} a_{jk} y_j y_k. \end{aligned} \tag{10.1.8}$$

这就证明了可化为至少有一个平方项出现的二次型.

第二步 如果 $a_{11}, a_{22}, \cdots, a_{nn}$ 不全是零. 例如, $a_{11} \neq 0$. 将二次型 $\varphi(x)$ 改写为

$$\begin{aligned}\varphi(x) &= a_{11}x_1^2 + 2x_1\sum_{k=2}^{n} a_{1k}x_k + \sum_{j,k=2}^{n} a_{jk}x_jx_k \\ &= a_{11}\Big[x_1^2 + 2x_1\big(\sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k\big) + \big(\sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k\big)^2\Big] \\ &\quad - a_{11}\big(\sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k\big)^2 + \sum_{j,k=2}^{n} a_{jk}x_jx_k \\ &= a_{11}\big(x_1 + \sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k\big)^2 + \psi_1(x_2, \cdots, x_n),\end{aligned} \tag{10.1.9}$$

其中 $\psi(x_2, \cdots, x_n)$ 是 $n-1$ 个变量 $x_2, x_3, \cdots, x_n$ 的二次型. 作坐标变换

$$y_1 = x_1 + \sum_{k=2}^{n}\frac{a_{1k}}{a_{11}}x_k, \quad y_k = x_k, \quad k = 2, 3, \cdots, n, \tag{10.1.10}$$

则二次型 $\varphi(x)$ 化为二次型 $\varphi_1(y) = a_{11}y_1^2 + \psi_1(y_2, \cdots, y_n)$, 这样一来, 只需要将 $n-1$ 个变数的二次型 $\psi_1(y_2, \cdots, y_n)$ 化为标准形就行了. 对 $\psi_1(y_2, \cdots, y_n)$ 用上面的配方法, 依次作下去, 最后可以将二次型 $\varphi(x)$ 化为二次型

$$\widetilde{\varphi_0}(z) = \delta_1z_1^2 + \delta_2z_2^2 + \cdots + \delta_rz_r^2, \tag{10.1.11}$$

其中 $\delta_1, \delta_2, \cdots, \delta_r$ 是域 $\mathbb{F}$ 中非零元素.

设域 $\mathbb{F} = \mathbb{C}$, 作坐标变换 $u_j = \sqrt{\delta_j}z_j$, $1 \leqslant j \leqslant r$, $u_j = z_j$, $r+1 \leqslant j \leqslant n$. 则二次型 $\widetilde{\varphi_0}(z)$ 化为标准二次型

$$\varphi_0(u) = u_1^2 + u_2^2 + \cdots + u_r^2. \tag{10.1.12}$$

设域 $\mathbb{F} = \mathbb{R}$, 作使变量编号改变的坐标变换, 就可以将 $\varphi(x)$ 化为 z 的二次型

$$\widetilde{\varphi_0}(z) = \delta_1z_1^2 + \delta_2z_2^2 + \cdots + \delta_sz_s^2 - \delta_{s+1}z_{s+1}^2 - \delta_{s+2}z_{s+2}^2 - \cdots - \delta_rz_r^2,$$

其中 $\delta_1, \cdots, \delta_r$ 大于零. 作坐标变换 $u_j = \sqrt{\delta_j}z_j$, $1 \leqslant j \leqslant r$, $u_j = z_j$, $r+1 \leqslant j \leqslant n$. 则二次型 $\widetilde{\varphi_0}(z)$ 化为标准二次型

$$\varphi_0(u) = u_1^2 + u_2^2 + \cdots + u_s^2 - u_{s+1}^2 - u_{s+2}^2 + \cdots + u_r^2. \tag{10.1.13}$$

设域 $\mathbb{F} = \mathbb{Q}$, 则标准二次型为

$$\varphi_0(u) = \delta_1u_1^2 + \delta_2u_2^2 + \cdots + \delta_su_s^2 - \delta_{s+1}u_{s+1}^2 - \delta_{s+2}u_{s+2}^2 + \cdots + \delta_ru_r^2, \tag{10.1.14}$$

其中 $\delta_j=\frac{m_j}{n_j}$ 为正有理数, m_j 和 n_j 互素, 且无平方因子, $j=1,2,\cdots,r$.

(二) 域 $\mathbb{F}$ 上二次型化标准形的矩阵方法

注意到将二次型化为标准形和求对称方阵在相合下的标准形是同一回事. 在这一段介绍矩阵方法. 我们将配方法用矩阵语言来表达. 设 S 是域 $\mathbb{F}$ 上 n 阶对称方阵, 当 $S=0$ 时, 不必讨论; 当 $S\neq 0$ 时, 分两种情形讨论如下.

(1) 设 S 有一个对角元素 $a_{jj}\neq 0$, 作与 S 相合的对称方阵

$$S_1=P_{1j}'SP_{1j}=\begin{pmatrix}a_{jj} & S_{12}\\ S_{12}' & S_2\end{pmatrix}, \tag{10.1.15}$$

其中 P_{1j} 是第一类初等方阵. 于是对称方阵 S_1 的第一行, 第一列元素不是零.

(2) 设 S 的对角元素全等于零. 由 $S\neq 0$ 可知必有一非对角元素 $a_{jk}\neq 0$. 首先设 $a_{12}\neq 0$. 作与 S 相合的对称方阵

$$S_1=\begin{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} & 0^{(2,n-2)}\\ 0^{(n-2,2)} & E_{n-2}\end{pmatrix}' S\begin{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} & 0^{(2,n-2)}\\ 0^{(n-2,2)} & E_{n-2}\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}2a_{12} & 0\\ 0 & 2a_{12}\end{pmatrix} & S_{12}\\ S_{12}' & S_2\end{pmatrix}. \tag{10.1.16}$$

所以对称方阵 S_1 的第一行, 第一列元素不是零. 在 $a_{jk}\neq 0$ 时, 同法可知存在和 S 相合的对称方阵, 它的第 j 行第 j 列元素不是零, 再由(1) 可知存在和 S 相合的对称方阵, 它的第一行, 第一列元素不是零.

由上面讨论可知, 只要 $S\neq 0$, 便存在非异方阵 P_1, 使得

$$S_1=P_1SP_1'=\begin{pmatrix}b_1 & \beta'\\ \beta & S_2\end{pmatrix}, \tag{10.1.17}$$

其中 $b_1\neq 0$, 又 β 是 $(n-1)\times 1$ 矩阵, S_2 是 $n-1$ 阶对称方阵. 今

$$\begin{pmatrix}1 & 0\\ -b_1^{-1}\beta & E_{n-1}\end{pmatrix}\begin{pmatrix}b_1 & \beta'\\ \beta & S_2\end{pmatrix}\begin{pmatrix}1 & 0\\ -b_1^{-1}\beta & E_{n-1}\end{pmatrix}'=\begin{pmatrix}b_1 & 0\\ 0 & S_2-b_1^{-1}\beta\beta'\end{pmatrix}. \tag{10.1.18}$$

设域 $\mathbb{F}=\mathbb{C}$, 再作

$$\begin{pmatrix}\frac{1}{\sqrt{b_1}} & 0\\ 0 & E_{n-1}\end{pmatrix}\begin{pmatrix}b_{11} & 0\\ 0 & S_0-b_1^{-1}\beta\beta'\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{b_1}} & 0\\ 0 & E_{n-1}\end{pmatrix}'=\begin{pmatrix}1 & 0\\ 0 & S_0-b_1^{-1}\beta\beta'\end{pmatrix}. \tag{10.1.19}$$

自然, $S_0-b_1^{-1}\beta\beta'$ 是 $n-1$ 阶复对称方阵. 用同样办法处理, 这样继续作下去, 最后可以找到复非异方阵 P, 使得

$$PSP'=\operatorname{diag}(E_r,0^{(n-r)}), \tag{10.1.20}$$

其中 r 是 S 的秩. 这就得到复对称方阵在相合下的标准形. 可以看出, 标准形很简单.

设域 $\mathbb{F}=\mathbb{R}$, 再作

$$\begin{pmatrix}\frac{1}{\sqrt{|b_1|}} & 0\\ 0 & E_{n-1}\end{pmatrix}\begin{pmatrix}b_1 & 0\\ 0 & S_0-b_1^{-1}\beta\beta'\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{|b_1|}} & 0\\ 0 & E_{n-1}\end{pmatrix}'=\begin{pmatrix}\pm1 & 0\\ 0 & S_0-b_1^{-1}\beta\beta'\end{pmatrix}, \tag{10.1.21}$$

其中 $\pm b_1>0$. 自然, $S_0-b_1^{-1}\beta\beta'$ 是 n 阶实对称方阵. 用同样办法处理, 这样继续作下去, 最后可以找到非异实方阵 P, 再用一个置换方阵作相合, 将对角元素适当调动, 使得

$$PSP'=\Lambda=\operatorname{diag}(E_s,-E_{r-s},0,0,\cdots,0), \tag{10.1.22}$$

其中 r 是 S 的秩, $\delta(S)=2s-r=\operatorname{tr}(\Lambda)$ 是实对称方阵 S 的**符号差**. 这就得到实对称方阵在相合下的标准形.

例 10.1.3 试将二次型

$$\varphi(x)=x_1^2+x_2^2+x_3^2-2x_4^2-2x_1x_2+2x_1x_3-2x_1x_4+2x_2x_3-4x_2x_4$$

化为标准形.

解 将二次型 $\varphi(x)$ 写成矩阵的形式, 即

$$\varphi(x)=(x_1,x_2,x_3,x_4)\begin{pmatrix}1&-1&1&-1\\-1&1&1&-2\\1&1&1&0\\-1&-2&0&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}.$$

所以问题化为求对称方阵

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

在相合下的标准形. 令

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

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 2 & 0 & 1 \\ 0 & -3 & 1 & -3 \end{pmatrix}, P_{13} \begin{pmatrix} 0 & 2 & -3 \\ 2 & 0 & 1 \\ -3 & 1 & -3 \end{pmatrix} P_{13}' = \begin{pmatrix} -3 & 1 & -3 \\ 1 & 0 & 2 \\ -3 & 2 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 & 1 & -3 \\ 1 & 0 & 2 \\ -3 & 2 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{3} & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{3} & 1 \\ 0 & 1 & 3 \end{pmatrix},$$

$$\begin{pmatrix} \sqrt{3} & 0 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \sqrt{3} & -3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

所以取非异方阵

$$P_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} - \sqrt{3} & 0 & \sqrt{3} & \frac{1}{\sqrt{3}} \\ 2 & 1 & -3 & -2 \end{pmatrix},$$

则

$$P_0 S P_0' = \operatorname{diag}(1, -1, 1, 0),$$

故取

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} P_0,$$

则 $PSP' = \operatorname{diag}(1, 1, -1, 0)$. 所以, 在坐标变换 $y = (P^{-1})'x$ 下, 实二次型 $\varphi(x)$ 变为标准形 $y_1^2 + y_2^2 - y_3^2$.

(三) 由上面讨论可知, 由**(二)** 的配方法, 我们证明了在复的情形有

定理 10.1.1 任一 n 阶复对称方阵 S 必相合于对角方阵

$$\Lambda = \operatorname{diag}(E_r, 0), \tag{10.1.23}$$

其中 r 是复对称方阵 S 的秩. 即在所有 n 阶复对称方阵构成的集合中, 按照相合分成等价类之并后; 在每个等价类中可以取式 (10.1.23) 为代表元素, 而 n 阶复对称方阵在相合下的全系不变量是复对称方阵的秩.

对 n 维复线性空间 $\mathfrak{L}$ 中任一复二次型 φ, 在 $\mathfrak{L}$ 中存在一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得二次型 φ 在这组基下的坐标表达式具有最简单的形式:

$$\varphi(x) = x_1^2 + x_2^2 + \cdots + x_r^2, \tag{10.1.24}$$

其中正整数 r 是二次型 $\varphi(x)$ 的**秩**. 所以, 秩是复二次型在基变换下的全系不变量.

在实的情形, 我们有

定理 10.1.2 任一 n 阶实对称方阵 S 必相合于对角形

$$\Lambda = \operatorname{diag}(E_s, -E_{r-s}, 0), \tag{10.1.25}$$

其中 r 是实对称方阵 S 的秩, $\delta(S) = 2s - r = \operatorname{tr}(\Lambda)$ 是实对称方阵 S 的**符号差**, 它是标准形 Λ 的迹, 也是实对称方阵 S 的正特征根个数减负特征根的个数.

这个定理说明了: 在所有 n 阶实对称方阵构成的集合中, 按照相合分成等价类之并后, 在每个等价类中可以取式 (10.1.25) 为代表元素. 下面来求 n 阶实对称方阵在相合下的全系不变量. 为此, 先给出

定理 10.1.3 (Witt 引理) 设 S 是 n 阶非异实对称方阵, S_1 和 S_2 是两个 m 阶非异实对称方阵. 如果两个 $n+m$ 阶非异实对称方阵

$$\begin{pmatrix} S & 0 \\ 0 & S_1 \end{pmatrix}, \qquad \begin{pmatrix} S & 0 \\ 0 & S_2 \end{pmatrix} \tag{10.1.26}$$

相合, 则 S_1 和 S_2 也相合.

证 下面的简单证明是华罗庚给出的. 今存在 n 阶非异实方阵 P, 使得

$$S = P\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)P',$$

于是

$$\begin{aligned} \operatorname{diag}(S, S_i) &= \operatorname{diag}(P\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)P', S_i) \\ &= \operatorname{diag}(P, E_m)\operatorname{diag}(\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n), S_i)\operatorname{diag}(P, E_m)', \quad i = 1, 2. \end{aligned}$$

所以实对称方阵 $\operatorname{diag}(S, S_1)$ 和 $\operatorname{diag}(S, S_2)$ 分别相合于实对称方阵

$$\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n, S_1), \qquad \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n, S_2).$$

因此, 问题化为只要证明在 $n=1$ 的情形, Witt 引理成立就可以了.

事实上, S 非异等价于 $\lambda_j \neq 0$, $1 \leqslant j \leqslant n$, 又由 S_i 非异可知 $n-1+m$ 阶实对称方阵

$$\widetilde{S}_i = \operatorname{diag}(\lambda_2, \lambda_3, \cdots, \lambda_n, S_i)$$

非异, $i=1,2$. 设若实对称方阵

$$\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n,S_1)=\operatorname{diag}(\lambda_1,\widetilde{S_1}),\qquad \operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n,S_2)=\operatorname{diag}(\lambda_1,\widetilde{S_2})$$

相合, 则实对称方阵 $\widetilde{S_1}$ 和 $\widetilde{S_2}$ 相合. 由归纳法便证明了 Witt 引理.

所以下面证明: 由 $\operatorname{diag}(\lambda_1,S_1)$ 和 $\operatorname{diag}(\lambda_1,S_2)$ 相合, 则可推出 S_1 和 S_2 相合.

今存在非异实方阵 P, 将它按第一行及第一列分块, 即记

$$P=\begin{pmatrix} a & \beta' \\ \alpha & P_1\end{pmatrix},$$

使得

$$\begin{aligned}\begin{pmatrix} \lambda_1 & 0 \\ 0 & S_2\end{pmatrix}&=\begin{pmatrix} a & \beta' \\ \alpha & P_1\end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & S_1\end{pmatrix}\begin{pmatrix} a & \alpha' \\ \beta & P_1'\end{pmatrix}\\ &=\begin{pmatrix} a^2\lambda_1+\beta' S_1\beta & a\lambda_1\alpha'+\beta' S_1 P_1' \\ a\lambda_1\alpha+P_1S_1\beta & \lambda_1\alpha\alpha'+P_1S_1P_1'\end{pmatrix}.\end{aligned}$$

等价于下列关系成立, 它们是

$$\begin{cases} a^2\lambda_1+\beta' S_1\beta=\lambda_1,\\ a\lambda_1\alpha+P_1S_1\beta=0,\\ \lambda_1\alpha\alpha'+P_1S_1P_1'=S_2.\end{cases}$$

因为如果 $a=1$, 改取 P 为 $-P$, 同样讨论之, 这时 P 的第一行, 第一列元素是 -1. 所以无妨设 $a\neq 1$. 这时有

$$\begin{aligned}&[P_1+(1-a)^{-1}\alpha\beta']S_1[P_1+(1-a)^{-1}\alpha\beta']'\\ &=P_1S_1P_1'+(1-a)^{-1}P_1S_1\beta\alpha'+(1-a)^{-1}\alpha\beta' S_1P_1'+(1-a)^{-2}\beta' S_1\beta\alpha\alpha'\\ &=S_2+\left(\frac{1+a}{1-a}-1-2\frac{a}{1-a}\right)\lambda_1\alpha\alpha'=S_2.\end{aligned}$$

所以取

$$Q=P_1+(1-a)^{-1}\alpha\beta',$$

则 $S_2=QS_1Q'$. 两边取行列式, 便得 $\det(S_2)=\det(S_1)\det(Q)^2$. 所以由 $\det(S_2)\neq 0$, 有 $\det(Q)\neq 0$, 即 S_1 和 S_2 相合. □

由 Witt 引理, 便有

定理 10.1.4 (惯性定理) 如果两个 n 阶实对称方阵

$$\operatorname{diag}(E_p,-E_{r-p},0^{(n-r)}),\quad \operatorname{diag}(E_q,-E_{s-q},0^{(n-s)}) \tag{10.1.27}$$

相合, 则 $s=r$, $q=p$. 因此, 实对称方阵 S 在相合下的标准形是

$$\Lambda=\operatorname{diag}(E_s,-E_{r-s},0),$$

其中 r 是实对称方阵 S 的**秩**, 整数 $\delta(S)=2s-r=\operatorname{tr}(\Lambda)$ 称为实对称方阵 S 的**符号差**. 所以实对称方阵的秩及符号差是它在相合下的全系不变量.

证 今存在 n 阶非异实方阵 P, 使得

$$P\operatorname{diag}(E_p,-E_{r-p},0^{(n-r)})P'=\operatorname{diag}(E_q,-E_{s-q},0^{(n-s)}).$$

由于相合下秩不改变, 所以 $s=r$. 再, 将实方阵 P 按照前 r 行及前 r 列分成四块, 即记

$$\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix},$$

则有

$$\operatorname{diag}(E_q,-E_{r-q},0)=\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}\operatorname{diag}(E_p,-E_{r-p},0)\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}',$$

于是

$$\begin{pmatrix} E_q & 0 \\ 0 & -E_{r-q} \end{pmatrix}=P_{11}\begin{pmatrix} E_p & 0 \\ 0 & -E_{r-p} \end{pmatrix}P_{11}'.$$

两边取行列式, 可证 $\det(P_{11})\neq 0$. 即 $\operatorname{diag}(E_p,-E_{r-p})$ 和 $\operatorname{diag}(E_q,-E_{r-q})$ 相合. 假设 $p\neq q$, 无妨设 $p>q$. 由 Witt 引理可知 $-E_{r-q}$ 和 $\operatorname{diag}(E_{p-q},-E_{r-p})$ 相合. 即存在 $r-q$ 阶非异实方阵 Q, 使得 $\operatorname{diag}(-E_{p-q},E_{r-p})=QQ'$. 比较两边第一行、第一列元素. 由 Q 是实方阵可知, 实方阵 QQ' 的第一行、第一列元素是正实数, 这就导出矛盾, 所以只有 $p=q$. □

将实对称方阵在相合下的标准形理论应用到实二次型, 便得

定理 10.1.5 对 n 维实线性空间 $\mathfrak{L}$ 中任一二次型 φ, 则在 $\mathfrak{L}$ 中存在一组基 α_1, $\alpha_2,\cdots,\alpha_n$, 使得二次型在这组基下的坐标表达式具有最简单的形式:

$$\varphi(x)=x_1^2+x_2^2+\cdots+x_p^2-x_{p+1}^2-x_{p+2}^2-\cdots-x_r^2, \tag{10.1.28}$$

其中整数 r 及 $2p-q$ 分别称为二次型 $\varphi(x)$ 的**秩** 和**符号差**. 所以, 实二次型的秩和符号差是二次型在基变换下的全系不变量.

(四) 推广及例

在 **(二)** 中我们给出了二次型化标准形的矩阵方法. 下面再介绍一种在计算上很有用的矩阵技巧, 它是配方法的推广. 将域 $\mathbb{F}$ 上 n 阶对称方阵 S 按照前 k 行及前

k 列分块为

$$S=\begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}, \tag{10.1.29}$$

于是 S_{11} 和 S_{22} 分别是域 $\mathbb{F}$ 上 k 阶和 $n-k$ 阶对称方阵. 假设 $\det(S_{11})\neq 0$, 那末 S_{11} 是 k 阶可逆对称方阵. 因此, 可以对 S_{11} 作相合

$$\begin{aligned}&\begin{pmatrix} E_k & 0 \\ -S_{12}'S_{11}^{-1} & E_{n-k} \end{pmatrix}\begin{pmatrix} S_{11} & S_{12} \\ S_{12}' & S_{22} \end{pmatrix}\begin{pmatrix} E_k & 0 \\ -S_{12}'S_{11}^{-1} & E_{n-k} \end{pmatrix}' \\ &=\begin{pmatrix} S_{11} & 0 \\ 0 & S_{22}-S_{12}'S_{11}^{-1}S_{12} \end{pmatrix}.\end{aligned} \tag{10.1.30}$$

然后对低阶对称方阵 S_{11} 及 $S_{33}=S_{22}-S_{12}'S_{11}^{-1}S_{12}$ 再作相合. 上面这个方法实际上是(二) 的推广, 即是配方法的推广.

最后举出两个例子, 介绍运用标准形理论的方法.

例 10.1.4 试求函数

$$f(x)=x'Sx+2\beta'x+b$$

的极小值, 其中 S 是 n 阶实正定对称方阵, $\beta=(b_1,b_2,\cdots,b_n)'$, $x=(x_1,x_2,\cdots,x_n)'$, b 是实数.

解 将函数 $f(x)$ 改写为

$$f(x)=(x',1)\begin{pmatrix} S & \beta \\ \beta' & b \end{pmatrix}(x',1)'.$$

由于 $\det(S)\neq 0$, 故

$$\begin{pmatrix} E_n & 0 \\ -\beta'S^{-1} & 1 \end{pmatrix}\begin{pmatrix} S & \beta \\ \beta' & b \end{pmatrix}\begin{pmatrix} E_n & 0 \\ -\beta'S^{-1} & 1 \end{pmatrix}'=\begin{pmatrix} S & 0 \\ 0 & b-\beta'S^{-1}\beta \end{pmatrix},$$

因此,

$$\begin{pmatrix} S & \beta \\ \beta' & b \end{pmatrix}=\begin{pmatrix} E_n & 0 \\ \beta'S^{-1} & 1 \end{pmatrix}\begin{pmatrix} S & 0 \\ 0 & b-\beta'S^{-1}\beta \end{pmatrix}\begin{pmatrix} E_n & 0 \\ \beta'S^{-1} & 1 \end{pmatrix}'.$$

所以

$$f(x)=(x'+\beta'S^{-1},1)\begin{pmatrix} S & 0 \\ 0 & b-\beta'S^{-1}\beta \end{pmatrix}(x'+\beta'S^{-1},1)'.$$

即

$$f(x)=(x+S^{-1}\beta)'S(x+S^{-1}\beta)+b-\beta'S^{-1}\beta.$$

因为 $S>0$, 所以对任一 x,

$$(x+S^{-1}\beta)'S(x+S^{-1}\beta)\geqslant 0,$$

又, 等号成立当且仅当 $x=-S^{-1}\beta$. 这证明了

$$f(x)\geqslant b-\beta'S^{-1}\beta.$$

又, 当且仅当在 $x=-S^{-1}\beta$ 时达到此值, 即

$$f(-S^{-1}\beta)=b-\beta'S^{-1}\beta.$$

因此 $b-\beta'S^{-1}\beta$ 就是函数 $f(x)$ 的极小值.

注意 在这个例子中, 如果将 S 是正定的条件改成 S 是负定的条件, 那末有

$$f(x)\leqslant b-\beta'S^{-1}\beta.$$

又, 当且仅当在 $x=-S^{-1}\beta$ 时达到极大值.

在这个例子中, 如果 S 非异, 且既非正定, 也非负定. 这时 $f(x)$ 既无极大值, 也无极小值. 事实上, 存在非异方阵 P, 使得

$$(P')^{-1}SP^{-1}=\mathrm{diag}\,(E_p,-E_{n-p}).$$

记 $z=P(x+S^{-1}\beta)$, 于是,

$$z'\mathrm{diag}\,(E_p,-E_{n-p})z=(x'+\beta'S^{-1})S(x+S^{-1}\beta)=f(x)+\beta'S^{-1}\beta-b,$$

所以

$$f(x)=z'\mathrm{diag}\,(E_p,-E_{n-p})z+b-\beta'S^{-1}\beta=\sum_{j=1}^{p}z_j^2-\sum_{j=p+1}^{n}z_j^2+b-\beta'S^{-1}\beta.$$

取 $z=(z_1,0,\cdots,0)'$, 则当 $z_1\to+\infty$ 时, $f(x)\to+\infty$; 取 $z=(0,\cdots,0,,0,\cdots,0,z_n)'$, 则当 $z_n\to+\infty$ 时, $f(x)\to-\infty$, 所以 $f(x)$ 没有极大值, 也没有极小值.

建议读者同时用数学分析的方法求函数 $f(x)$ 的极值.

例 10.1.5 (华罗庚) 设 S 和 $S-\alpha\alpha'$ 都是 n 阶非异实对称方阵, 其中 α 是 $n\times 1$ 实矩阵. 试证: 当 $\alpha'S^{-1}\alpha>1$ 时, 符号差 $\delta(S)=\delta(S-\alpha\alpha')+2$. 当 $\alpha'S^{-1}\alpha<1$ 时, 符号差 $\delta(S)=\delta(S-\alpha\alpha')$.

证 考虑 $n+1$ 阶实对称方阵

$$\begin{pmatrix}1&\alpha'\\ \alpha&S\end{pmatrix}.$$

由

$$\begin{pmatrix} 1 & 0 \\ -\alpha & E_n \end{pmatrix} \begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\alpha & E_n \end{pmatrix}' = \begin{pmatrix} 1 & 0 \\ 0 & S-\alpha\alpha' \end{pmatrix},$$

$$\begin{pmatrix} 1 & -\alpha' S^{-1} \\ 0 & E_n \end{pmatrix} \begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix} \begin{pmatrix} 1 & -\alpha' S^{-1} \\ 0 & E_n \end{pmatrix}' = \begin{pmatrix} 1-\alpha' S^{-1}\alpha & 0 \\ 0 & S \end{pmatrix},$$

于是符号差有关系

$$\delta\left(\begin{pmatrix} 1 & 0 \\ 0 & S-\alpha\alpha' \end{pmatrix}\right) = \delta\left(\begin{pmatrix} 1-\alpha' S^{-1}\alpha & 0 \\ 0 & S \end{pmatrix}\right).$$

由于符号差是标准形中 1 的个数减去 -1 的个数, 即等于标准形的对角元素相加, 因此证明了: 当 $\alpha' S^{-1}\alpha \neq 1$, 则

$$1+\delta(S-\alpha\alpha') = \frac{1-\alpha' S^{-1}\alpha}{|1-\alpha' S^{-1}\alpha|} + \delta(S).$$

当 $\alpha' S^{-1}\alpha < 1$, 则有 $1+\delta(S-\alpha\alpha') = 1+\delta(S)$, 即 $\delta(S) = \delta(S-\alpha\alpha')$; 当 $\alpha' S^{-1}\alpha > 1$, 则有 $1+\delta(S-\alpha\alpha') = -1+\delta(S)$, 即 $\delta(S) = \delta(S-\alpha\alpha')+2$; 当 $\alpha' S^{-1}\alpha = 1$, 则有

$$\det\begin{pmatrix} 1 & \alpha' \\ \alpha & S \end{pmatrix} = \det(S-\alpha\alpha') = 0,$$

这和假设 $S-\alpha\alpha'$ 非异矛盾. □

习 题 10.1

10.1.1 试求下列实对称方阵在相合下的标准形:

$$\begin{pmatrix} 2 & 5 & 8 \\ 5 & 3 & 1 \\ 8 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0_n & E_n \\ E_n & 0_n \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix},$$

$$\begin{pmatrix} 0_n & E_n \\ E_n & 0_n \end{pmatrix}, \quad \begin{pmatrix} 0_n & E_n & 0 \\ E_n & 0_n & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & \frac{1}{2} & & \\ \frac{1}{2} & 1 & \ddots & \\ & \ddots & \ddots & \frac{1}{2} \\ & & \frac{1}{2} & 1 \end{pmatrix}.$$

10.1.2 试求下列实二次型的标准形:

(i) $2x_1^2+3x_2^2+10x_1x_2+16x_1x_3+2x_2x_3$; **(ii)** $x_1^2+2x_2^2+3x_3^2+4x_4^2+x_1x_3+x_2x_4$;
(iii) $\sum\limits_{1\leqslant j<k\leqslant n} x_jx_{j+1}$; **(iv)** $\sum\limits_{1\leqslant j<k\leqslant n} x_jx_k$; **(v)** $\sum\limits_{1\leqslant j<k\leqslant n} (-1)^{j+k}x_jx_k$;
(vi) $\sum\limits_{1\leqslant j<k\leqslant n} |j-k|x_jx_k$; **(vii)** $\sum\limits_{j=1}^{n}(S-x_j)^2$, 其中 $S=\dfrac{1}{n}\sum\limits_{k=1}^{n}x_k$.

10.1.3 设 A 和 B 都是 n 阶实对称方阵, 且 A 正定. 试证: 存在一个 n 阶非异实方阵 P, 使得 $P'AP$ 和 $P'BP$ 同时是对角形.

10.1.4 如果实二次型 $f(x)$ 等于零的必要且充分条件是 $x=0$, 那末 $f(x)$ 是正定或负定的.

10.1.5 试利用实对称双线性函数来证惯性定理 (即定理 10.1.4), 即若

$$\sum_{j=1}^{p} x_j^2 - \sum_{j=p+1}^{r} x_j^2, \quad \sum_{j=1}^{q} y_j^2 - \sum_{j=q+1}^{s} y_j^2$$

是同一对称双线性函数在不同基下的表达式, 则 $s=r$, $q=p$.

10.1.6 题 1 及 2 中的实对称方阵及实二次型如果都视为复对称方阵及复二次型, 它们的标准形各是什么?

10.1.7 试将下列复对称方阵及复二次型化为标准形:

(i) $\displaystyle\sum_{1\leqslant j<k\leqslant n} (j+ik)x_j x_k$;

(ii) $\begin{pmatrix} 1 & 1+i & 2+i & \cdots & (n-1)+i \\ 1+i & 1 & 0 & \cdots & 0 \\ 2+i & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ n-1+i & 0 & 0 & \cdots & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & \frac{i}{2} & & & \\ \frac{i}{2} & 1 & \ddots & & \\ & \ddots & \ddots & \frac{i}{2} & \\ & & \frac{i}{2} & 1 \end{pmatrix}$.

10.1.8 如果二次型 $f(x)$ 对应的方阵 S 的主子式:

$$\det(S_j) = \det S\begin{pmatrix} 12\cdots j \\ 12\cdots j \end{pmatrix} \neq 0, \quad 1 \leqslant j \leqslant n,$$

试证: $f(x)$ 的标准形是

$$\det(S_1)y_1^2 + \frac{\det(S_2)}{\det(S_1)}y_2^2 + \cdots + \frac{\det(S_n)}{\det(S_{n-1})}y_n^2.$$

10.1.9 试给出实二次型能分解为两个实线性函数的乘积的必要且充分条件.

10.1.10 设实二次型$f(x)=x'Sx$ 有性质

$$\det(S)\neq 0, \quad f(x_1, \cdots, x_r, 0, \cdots, 0)=0.$$

试证: 二次型$f(x)$可经过非异线性变换变为二次型

$$g(y) = y_1y_{r+1} + \cdots + y_ry_{2r} + h(y_{2r+1}, \cdots, y_n),$$

其中实二次型 h 的方阵表示为 $n-2r$ 阶非异实对称方阵, 它的符号差的绝对值不超过 $n-2r$.

§10.2 实正定对称方阵和实方阵的极分解

定义 10.2.1 n 阶实对称方阵 S 分别称为**正定的**、**半正定的**、**负定的**、**半负定的**, 如果对任一非零 $n\times 1$ 矩阵 α, 实数 $\alpha'S\alpha$ 分别有

$$\alpha S\alpha' > 0, \quad \alpha S\alpha' \geqslant 0, \quad \alpha S\alpha' < 0, \quad \alpha S\alpha' \leqslant 0, \tag{10.2.1}$$

这时分别用符号

$$S>0,\quad S\geqslant 0,\quad S<0,\quad S\leqslant 0 \tag{10.2.2}$$

记之.

由定义可知, $S<0$ 的必要且充分条件是 $-S>0$; $S\leqslant 0$ 的必要且充分条件是 $-S\geqslant 0$. 所以只需要研究正定对称方阵和半正定对称方阵就够了, 因为可以将每个性质相应地转换为负定对称方阵和半负定对称方阵的性质.

引理 10.2.2 如果 n 阶实对称方阵 $S>0$ ($S\geqslant 0$), 则对任一 n 阶非异实方阵 P, 恒有 $P'SP>0$ ($PSP'\geqslant 0$).

证 今 P 是非异方阵, 所以任取 $n\times 1$ 非零矩阵 β, 则 $n\times 1$ 矩阵 $\alpha=P\beta\neq 0$. 由条件 $S>0$, 所以 $\alpha'S\alpha>0$, 即 $\beta'(P'SP)\beta>0$. 再由 β 的任意性及定义可知 $P'SP>0$. 将整个讨论易 “$>$” 为 “$\geqslant$”, 则证明了对半正定的情形引理也成立. □

由引理 10.2.2 可知, 如果对任一实对称方阵 S, 存在一个非异实方阵 P, 使得 $P'SP>0$ ($P'SP\geqslant 0$), 那末,

$$S=(P^{-1})'(P'SP)P^{-1}>0\qquad (S\geqslant 0), \tag{10.2.3}$$

即 S 也正定 (S 半正定).

下面给出实对称方阵正定和半正定的若干必要且充分条件.

定理 10.2.3 设

$$S=\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix} \tag{10.2.4}$$

是 n 阶实对称方阵, 则下面条件互相等价:

(1) $S>0$, 即 n 阶实对称方阵 S 正定;

(2) S 的特征根全是正实数;

(3) 存在 n 阶非异实方阵 Q, 使得

$$S=Q'Q; \tag{10.2.5}$$

(4) S 的一切主子式

$$\det S\begin{pmatrix} i_1i_2\cdots i_k\\ i_1i_2\cdots i_k\end{pmatrix},\qquad 1\leqslant i_1<i_2<\cdots<i_r\leqslant n \tag{10.2.6}$$

都是正实数;

(5) S 的 n 个顺序主子式:

$$\det S\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}=\det\begin{pmatrix}a_{11}&\cdots&a_{1k}\\\vdots&&\vdots\\a_{k1}&\cdots&a_{kk}\end{pmatrix},\quad k=1,2,\cdots,n \tag{10.2.7}$$

都大于零.

证 我们来证(1) $\longrightarrow$ (2) $\longrightarrow$ (3) $\longrightarrow$ (4) $\longrightarrow$ (5) $\longrightarrow$ (1).

(1) $\longrightarrow$ (2). 设(1) 成立, 即 $S>0$. 由定理 9.3.7, 存在实正交方阵 O, 使得

$$O'SO=\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

其中 $\lambda_1\leqslant\cdots\leqslant\lambda_n$ 都是实数. 由引理 10.2.2, 所以 $O'SO>0$. 于是对任一 i, $1\leqslant i\leqslant n$, 有

$$e_i'\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)e_i=\lambda_i>0,\quad 1\leqslant i\leqslant n, \tag{10.2.8}$$

其中 $e_i=(\delta_{1i},\delta_{2i},\cdots,\delta_{ni})'$, δ_{ij} 是 Kronecker 符号, $1\leqslant i,j\leqslant n$. 这就证明了 (2) 成立.

(2) $\longrightarrow$ (3). 设(2) 成立, 即 S 的所有特征根 λ_1, λ_2, $\cdots$, λ_n 全是正实数. 于是无妨设 $0<\lambda_1\leqslant\cdots\leqslant\lambda_n$. 由定理 9.3.7 可知, 存在实正交方阵 O, 使得 $OSO'=\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)$. 所以

$$\begin{aligned}S&=O\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)O'\\&=O\mathrm{diag}\,(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})\mathrm{diag}\,(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})O'.\end{aligned} \tag{10.2.9}$$

记 $Q=\mathrm{diag}\,(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})O'$, 于是 $\det(Q)=\sqrt{\lambda_1\lambda_2\cdots\lambda_n}\det(O)\neq0$, 又 $S=Q'Q$. 这就证明了(3) 成立.

(3) $\longrightarrow$ (4). 设(3) 成立, 即存在 n 阶非异实方阵 Q, 使得 $S=Q'Q$. 则由 Binet–Cauchy 公式, 即定理 3.2.1, 知方阵 S 的任一 k 阶主子式

$$\Delta_k=\det S\begin{pmatrix}i_1i_2\cdots i_k\\i_1i_2\cdots i_k\end{pmatrix}=\det(Q'Q)\begin{pmatrix}i_1i_2\cdots i_k\\i_1i_2\cdots i_k\end{pmatrix}.$$

所以

$$\begin{aligned}\Delta_k&=\det(Q'Q)\begin{pmatrix}i_1i_2\cdots i_k\\i_1i_2\cdots i_k\end{pmatrix}\\&=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\det Q'\begin{pmatrix}i_1i_2\cdots i_k\\j_1j_2\cdots j_k\end{pmatrix}\det Q\begin{pmatrix}j_1j_2\cdots j_k\\i_1i_2\cdots i_k\end{pmatrix}\\&=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\det Q\begin{pmatrix}j_1j_2\cdots j_k\\i_1i_2\cdots i_k\end{pmatrix}^2\geqslant0.\end{aligned}$$

今 Q 非异, 所以 Q 的 n 列线性无关. 因此存在 k 阶子式 $\det Q\begin{pmatrix} j_1j_2\cdots j_k \\ i_1i_2\cdots i_k \end{pmatrix}$, 使得

$$\det Q\begin{pmatrix} j_1j_2\cdots j_k \\ i_1i_2\cdots i_k \end{pmatrix}^2 > 0.$$

这证明了

$$\Delta_k = \det S\begin{pmatrix} i_1i_2\cdots i_k \\ i_1i_2\cdots i_k \end{pmatrix} \neq 0.$$

所以**(4)** 成立.

(4) $\longrightarrow$ **(5)**. 显然成立.

(5) $\longrightarrow$ **(1)**. 设**(5)** 成立, 即设 n 阶实对称方阵 S 的 n 个顺序主子式都大于零. 下面来证明 $S > 0$. 为此对阶数 n 作归纳法. 在 $n = 1$ 时结论显然成立. 设对一切 $n-1$ 阶实对称方阵 S_0, 结论成立. 下面来证在 n 时, 结论也成立.

将 n 阶实对称方阵 S 按照前 $n-1$ 行及 $n-1$ 列分成四块, 即记

$$S = \begin{pmatrix} S_{n-1} & v \\ v' & a_{nn} \end{pmatrix}.$$

由于 S 的 $n-1$ 个主子式 $\det S\begin{pmatrix} 12\cdots k \\ 12\cdots k \end{pmatrix}$, $k = 1, 2, \cdots, n-1$ 是 S_{n-1} 的 $n-1$ 个顺序主子式, 它们都是正实数. 由归纳法假设可知 S_{n-1} 正定. 而 S 的第 n 个主子式是 S 的行列式, 所以问题化为证明在条件 $\det(S) > 0$ 及 $S_{n-1} > 0$ 下有 $S > 0$.

因为 $S_{n-1} > 0$, 故 S_{n-1} 有逆方阵. 于是

$$\begin{pmatrix} E_{n-1} & 0 \\ -v'S_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} S_{n-1} & v \\ v' & a_{nn} \end{pmatrix}\begin{pmatrix} E_{n-1} & 0 \\ -v'S_{n-1}^{-1} & 1 \end{pmatrix}' = \mathrm{diag}\,(S_{n-1}, a_{nn} - v'S_{n-1}^{-1}v).$$

记 $w = a_{nn} - v'S_{n-1}^{-1}v$, 两边取行列式, 得到 $\det(S) = w\det(S_{n-1})$. 再由 $\det(S) > 0$, $\det(S_{n-1}) > 0$ 可知 $w = a_{nn} - v'S_{n-1}^{-1}v > 0$. 因此 $\mathrm{diag}\,(S_{n-1}, a_{nn} - v'S_{n-1}^{-1}v) > 0$. 所以 $S > 0$. □

注意 虽然我们还可以给出一些其他的等价条件. 但是学生不理解哪几个等价条件最有用. 因此, 在定理中我们只给出重要的等价条件, 其他等价条件可以在习题中找到.

半正定对称方阵也有类似的判别条件:

定理 10.2.4 设

$$S = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{10.2.10}$$

是 n 阶实对称方阵, 则下面条件互相等价:

(1) $S \geqslant 0$, 即 n 阶实对称方阵 S 半正定;

(2) S 的特征根全是非负实数;

(3) 存在 n 阶实方阵 Q, 使得

$$S = Q'Q; \tag{10.2.11}$$

(4) S 的一切主子式

$$S\begin{pmatrix} i_1 i_2 \cdots i_k \\ i_1 i_2 \cdots i_k \end{pmatrix}, \quad 1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant n \tag{10.2.12}$$

都是非负实数.

证 我们来证**(1)** $\longrightarrow$ **(2)** $\longrightarrow$ **(3)** $\longrightarrow$ **(4)** $\longrightarrow$ **(1)**.

(1) $\longrightarrow$ **(2)** $\longrightarrow$ **(3)** $\longrightarrow$ **(4)** 的证明和定理 10.2.3 的证明相同.

(4) $\longrightarrow$ **(1)**. 设**(4)** 成立, 即设 n 阶实对称方阵 S 的所有主子式都非负. 下面来证明 $S \geqslant 0$.

首先, 我们来证实对称方阵 S 的特征根都是非负实数. 考虑实对称方阵 S 的特征多项式

$$\det(\lambda E_n - S) = \lambda^n - a_1\lambda^{n-1} + a_2\lambda^{n-2} - \cdots + (-1)^{n-1}a_{n-1}\lambda + (-1)^n a_n,$$

由于系数 a_j 是方阵 S 的所有 j 阶主子式之和. 由条件可知这些主子式都大于或等于零, 所以 $a_1 \geqslant 0$, $a_2 \geqslant 0$, $\cdots$, $a_n \geqslant 0$. 另一方面, 已知实对称方阵的特征根都是实数. 如果有一个负根 $-\lambda_0 < 0$, 则

$$(-1)^n \det(\lambda_0 E_n - S) = \lambda_0^n + a_1\lambda_0^{n-1} + a_2\lambda_0^{n-2} + \cdots + a_{n-1}\lambda_0 + a_n = 0.$$

因此, 由 $a_j\lambda_0^{n-j} \geqslant 0, j = 1, 2, \cdots, n$ 有 $a_j\lambda_0^{n-j} = 0, 1 \leqslant j \leqslant n$. 再由 $\lambda_0 > 0$, 因此 $a_1 = a_2 = \cdots = a_n = 0$, 所以实对称方阵 S 的特征多项式 $\det(\lambda E_n - S) = \lambda^n$. 这证明了方阵 S 的特征根都是零. 这和 S 有负特征根 $-\lambda_0 < 0$ 矛盾, 所以证明了 S 的特征根都是非负实数. 由定理 9.3.7, 存在实正交方阵 O, 使得 $O'SO = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 其中 $\lambda_j \geqslant 0$, $1 \leqslant j \leqslant n$. 任取 $\alpha \in \mathbb{R}^n$, 记 $(b_1, b_2, \cdots, b_n)' = \beta = O'\alpha$, 则

$$\alpha' S\alpha = \alpha' O \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) O'\alpha = \beta' \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\beta = \sum_{k=1}^{n} b_k^2 \lambda_k \geqslant 0.$$

这证明了 $\alpha' S\alpha \geqslant 0, \forall \alpha \in \mathbb{R}^n$. 所以 $S \geqslant 0$, 即**(1)** 成立. □

注意 定理 10.2.3 的**(5)** 不能直接推广到半定正的情形. 例如, 实对称方阵

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

的第 1, 2 个主子式

$$\det S\begin{pmatrix}1\\1\end{pmatrix}=0,\qquad \det S\begin{pmatrix}1\ 2\\1\ 2\end{pmatrix}=\det(S)=0,$$

但是它不是半正定的.

上面给出了正定实对称方阵及半正定实对称方阵的一些判别条件. 但是, 有时利用矩阵方法来考虑这类问题时还会感到困难. 下面举一个例子, 给出处理这类问题的另一种方法.

例 10.2.1 设两个 n 阶实方阵

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\cdots&a_{nn}\end{pmatrix},\qquad B=\begin{pmatrix}b_{11}&\cdots&b_{1n}\\\vdots&&\vdots\\b_{n1}&\cdots&b_{nn}\end{pmatrix}\tag{10.2.13}$$

都是半正定的实对称方阵. 试证: n 阶实方阵

$$C=\begin{pmatrix}a_{11}b_{11}&\cdots&a_{1n}b_{1n}\\\vdots&&\vdots\\a_{n1}b_{n1}&\cdots&a_{nn}b_{nn}\end{pmatrix}\tag{10.2.14}$$

也是半正定对称方阵.

证 显然 C 也是实对称方阵. 今任取 $n\times 1$ 矩阵 $x=(x_1,x_2,\cdots,x_n)'$, 则由 $A\geqslant 0$, $B\geqslant 0$ 可知

$$\sum_{j,k=1}^{n}a_{jk}x_jx_k\geqslant 0,\quad \sum_{j,k=1}^{n}b_{jk}x_jx_k\geqslant 0.$$

由半正定的定义可知, 需要证明

$$\sum_{j,k=1}^{n}a_{jk}b_{jk}x_jx_k\geqslant 0.$$

下面给出五种证法, 有些证法是大同小异的.

证一 今 $B\geqslant 0$, 所以存在 n 阶方阵 $Q=(q_{jk})$, 使得 $B=Q'Q$, 即

$$b_{jk}=\sum_{i=1}^{n}q_{ij}q_{ik}\quad j,k=1,2,\cdots,n.$$

所以

$$\sum_{j,k=1}^{n}a_{jk}b_{jk}x_jx_k=\sum_{j,k=1}^{n}a_{jk}\Big(\sum_{i=1}^{n}q_{ij}q_{ik}\Big)x_jx_k=\sum_{i=1}^{n}\sum_{j,k=1}^{n}a_{jk}(x_jq_{ij})(x_kq_{ik}),$$

因为对任意 $n\times 1$ 矩阵 $(x_1q_{i1}, x_2q_{i2}, \cdots, x_nq_{in})'$, 有

$$\sum_{j,k=1}^{n} a_{jk}(x_jq_{ij})(x_kq_{ik}) \geqslant 0,$$

所以

$$\sum_{j,k=1}^{n} a_{jk}b_{jk}x_jx_k \geqslant 0. \qquad \square$$

证二 令 $B\geqslant 0$, 所以存在实正交方阵 $O=(p_{jk})$, 使得

$$B = O'\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)O,$$

其中 $0\leqslant\lambda_1\leqslant\cdots\leqslant\lambda_n$. 所以

$$b_{jk} = \sum_{i=1}^{n} p_{ij}\lambda_i p_{ik} \quad j,k=1,2,\cdots,n.$$

因此

$$\sum_{j,k=1}^{n} a_{jk}b_{jk}x_jx_k = \sum_{i=1}^{n}\lambda_i \sum_{j,k=1}^{n} a_{jk}(x_jp_{ij})(x_kp_{ik}) \geqslant 0. \qquad \square$$

证三 令 $B = O\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)O'$. 设 B 的秩是 r, 则 $0=\lambda_1=\cdots=\lambda_{n-r}<\lambda_{n-r+1}\leqslant\cdots\leqslant\lambda_n$. 所以

$$B = \sum_{j=n-r+1}^{n} O\operatorname{diag}(0, \cdots, 0, \lambda_j, 0, \cdots, 0)O' = B_1+B_2+\cdots+B_r,$$

其中 B_1, B_2, $\cdots$, B_r 都是秩为 1 的实半正定对称方阵. 换句话说, 我们证明了: 任一实半正定对称方阵必定能分解为有限个秩为 1 的实半正定对称方阵之和.

今记 $B_t=(b_{jk}^{(t)})$, 则

$$b_{jk} = \sum_{t=1}^{r} b_{jk}^{(t)}.$$

从而记

$$C_t = \begin{pmatrix} a_{11}b_{11}^{(t)} & \cdots & a_{1n}b_{1n}^{(t)} \\ \vdots & & \vdots \\ a_{n1}b_{n1}^{(t)} & \cdots & a_{nn}b_{nn}^{(t)} \end{pmatrix}, \quad 1\leqslant t\leqslant r,$$

其中 C_t 是 n 阶实方阵, 它由半正定对称方阵 A 和 B_t 构成. 则 $C=\sum\limits_{t=1}^{r} C_t$. 所以, 如果能证 C_t 半正定, $1\leqslant t\leqslant r$, 那末再利用半正定对称方阵之和仍为半正定对称方阵, 可知方阵 C 也是半正定的. 换句话说, 问题化为证明: 在半正定对称方阵 B 的

秩是 1 时, C 也是半定正的. 事实上, 这时一定存在 $n\times 1$ 矩阵 $\beta=(b_1,b_2,\cdots,b_n)'$, 使得 $B=\beta\beta'$, 即 $b_{jk}=b_jb_k$. 所以

$$\sum_{j,k=1}^n a_{jk}b_{jk}x_jx_k=\sum_{j,k=1}^n a_{jk}b_jb_kx_jx_k=\sum_{j,k=1}^n a_{jk}(b_jx_j)(b_kx_k)\geqslant 0.$$

□

证四 由证三可知: 半正定对称方阵 A 和 B 都能分解为秩等于 1 的半正定对称方阵之和. 所以问题化为: 在 A 和 B 都是秩为 1 的半正定对称方阵时, 证 C 半正定. 这时 $A=\alpha\alpha'$, $B=\beta\beta'$, 其中 $\alpha=(a_1,a_2,\cdots,a_n)'$, $\beta=(b_1,b_2,\cdots,b_n)'$. 所以 $a_{jk}=a_ja_k$, $b_{jk}=b_jb_k$. 因此

$$\sum_{j,k=1}^n a_{jk}b_{jk}x_jx_k=\sum_{j,k=1}^n a_ja_kb_jb_kx_jx_k=\sum_{j,k=1}^n (a_jb_jx_j)(a_kb_kx_k)=\Big(\sum_{j,k=1}^n a_jb_jx_j\Big)^2\geqslant 0.$$

□

证五 今设

$$B=\begin{pmatrix} b_{11} & \cdots & b_{1n}\\ \vdots & & \vdots\\ b_{n1} & \cdots & b_{nn}\end{pmatrix}.$$

由

$$B_1=O'BO=\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n),\quad 0\leqslant\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n,$$

则由 $a_{jk}=a_{kj}$, $b_{jk}=b_{kj}$, $1\leqslant j,k\leqslant n$, 所以

$$\sum_{j,k=1}^n a_{jk}b_{jk}=\mathrm{tr}\,(AB)=\mathrm{tr}\,(O'ABO)=\mathrm{tr}\,(O'AO)(O'BO)=\mathrm{tr}\,(A_1B_1),$$

其中

$$A_1=O'AO=\begin{pmatrix} \widetilde{a}_{11} & \cdots & \widetilde{a}_{1n}\\ \vdots & & \vdots\\ \widetilde{a}_{n1} & \cdots & \widetilde{a}_{nn}\end{pmatrix}\geqslant 0,$$

$$B_1=O'BO=\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)\geqslant 0.$$

所以 $\widetilde{a}_{11}\geqslant 0,\cdots,\widetilde{a}_{nn}\geqslant 0$, 又 $\lambda_1\geqslant 0,\cdots,\lambda_n\geqslant 0$. 故

$$\sum_{j,k=1}^n a_{jk}b_{jk}=\mathrm{tr}\,(A_1B_1)=\sum_{j=1}^n\lambda_j\widetilde{a}_{jj}\geqslant 0.$$

因此证明了当 $A=(a_{jk})\geqslant 0$, $B=(b_{jk})\geqslant 0$ 时

$$\sum_{j,k=1}^n a_{jk}b_{jk}\geqslant 0.$$

今记

$$\sum_{j,k=1}^{n} a_{jk}b_{jk}x_jx_k = \sum_{j,k=1}^{n} a_{jk}(x_jb_{jk}x_k) = \sum_{j,k=1}^{n} a_{jk}\widetilde{b}_{jk},$$

其中

$$(\widetilde{b}_{jk}) = (x_jb_{jk}x_k) = \operatorname{diag}(x_1, x_2, \cdots, x_n)'B\operatorname{diag}(x_1, x_2, \cdots, x_n).$$

因此 n 阶实对称方阵 $(\widetilde{b}_{jk})$ 也半正定. 再由 $A \geqslant 0$, 所以

$$\sum_{j,k=1}^{2} a_{jk}\widetilde{b}_{jk} \geqslant 0,$$

即

$$\sum_{j,k=1}^{n} a_{jk}b_{jk}x_jx_k \geqslant 0. \qquad \square$$

下面给出一个 n 阶半正定实对称方阵的和的行列式的一个不等式.

定理 10.2.5 假设 A 和 B 是 n 阶半正定实对称方阵, 则存在 n 阶非异实方阵 P, 使得 $P'AP$ 和 $P'BP$ 同时是对角形

$$P'AP = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r, 0_{n-r}), \quad P'BP = \operatorname{diag}(1-\lambda_1, 1-\lambda_2, \cdots, 1-\lambda_r, 0_{n-r}), \tag{10.2.15}$$

其中 $r = \operatorname{rank}(A+B)$, $1 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r \geqslant 0$. 又有**行列式不等式**

$$\det(A+B) \geqslant \det(A) + \det(B), \tag{10.2.16}$$

而且等号成立的充分且必要条件是 $A = 0$ 或 $B = 0$.

证 设 $r = \operatorname{rank}(A+B)$. 由定理 10.2.4, 因为 $A+B$ 半正定, 所以存在 n 阶非异实方阵 P_1, 使得

$$P_1'(A+B)P_1 = \operatorname{diag}(E_r, 0_{n-r}).$$

将 $P_1'AP_1$ 和 $P_1'BP_1$ 按前 r 行、前 r 列分块为

$$P_1'AP_1 = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{pmatrix}, \qquad P_1'BP_1 = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{pmatrix}, \tag{10.2.17}$$

则 $A_{11} + B_{11} = E_r$, $A_{22} + B_{22} = 0$. 由 $A \geqslant 0$ 及 $B \geqslant 0$ 可知 $A_{22} \geqslant 0$, $B_{22} \geqslant 0$. 所以 $A_{22} + B_{22} = 0$ 推出 $A_{22} = B_{22} = 0$. 这时, 由 $A \geqslant 0$, 计算 $P_1'AP_1$ 的主子式 $\det A\begin{pmatrix} ij \\ ij \end{pmatrix}$, $1 \leqslant i \leqslant r < j \leqslant n$ 可知 $A_{12} = 0$. 同理, $B_{12} = 0$. 所以 $P_1'AP_1 = \operatorname{diag}(A_{11}, 0)$, $P_1'BP_1 = \operatorname{diag}(E_r - A_{11}, 0)$. 由式 (9.3.10), 存在 r 阶实正交方阵 P_2, 使得 $P_2'A_{11}P_2 = \operatorname{diag}(\lambda_1, \lambda_2, \cdots ; \lambda_r)$, 其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r \geqslant 0$. 取 $P = P_1\operatorname{diag}(P_2, E_{n-r})$, 便证明了式 (10.2.15) 成立.

现在考虑 $\det(A+B)$, $\det(A)$ 和 $\det(B)$. 显然行列式不等式 $\det(A+B) \geqslant \det(A)+\det(B)$ 等价于行列式不等式 $\det(P'(A+B)P) \geqslant \det(P'AP)+\det(P'BP)$. 因此, 问题化为无妨设 A 和 B 是对角方阵 (10.2.15). 所以当 $\operatorname{rank}(A+B)=r<n$ 时 $\det(A+B)=\det(A)=\det(B)=0$, 即等式成立. 当 $\operatorname{rank}(A+B)=r=n$ 时, 问题化为证明

$$1 \geqslant \prod_{j=1}^{n}\lambda_j+\prod_{j=1}^{n}(1-\lambda_j), \quad 1\geqslant \lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant 0,$$

且等号成立当且仅当 $\lambda_1=\cdots=\lambda_n=0$ 或 $\lambda_1=\cdots=\lambda_n=1$.

事实上,

$$1-\prod_{j=1}^{n}\lambda_j-\prod_{j=1}^{n}(1-\lambda_j)=\lambda_n(1-\prod_{j=1}^{n-1}\lambda_j)+(1-\lambda_n)\left[1-\prod_{j=1}^{n-1}(1-\lambda_j)\right]\geqslant 0,$$

且等号成立当且仅当 $\lambda_n(1-\prod_{j=1}^{n-1}\lambda_j)=0$, $(1-\lambda_n)\left[1-\prod_{j=1}^{n-1}(1-\lambda_j)\right]=0$. 所以当 $\lambda_n=0$ 时, $\prod_{j=1}^{n-1}(1-\lambda_j)=1$, 因此, $\lambda_1=\cdots=\lambda_{n-1}=0$, 即 $A=0$, $B=E_n$; 当 $\lambda_n\neq 0$ 时, $\prod_{j=1}^{n-1}\lambda_j=1$, 因此, $\lambda_1=\cdots=\lambda_{n-1}=1$, 又 $1-\lambda_n=0$, 即 $\lambda_n=1$. 因此, $A=E_n$, $B=0$. □

下面再给出实正定和半正定对称方阵的一些应用, 特别重要的是给出极分解式.

定理 10.2.6 设实对称方阵 S 半正定, 则存在且只存在一个实半正定对称方阵 S_1, 使得 $S_1^2=S$, 且任一和实对称方阵 S 可交换的方阵必和实对称方阵 S_1 可交换 (注意, 当 S 正定时 S_1 也正定).

证 今存在实正交方阵 O, 使得

$$S=O'\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)O, \tag{10.2.18}$$

其中 $0\leqslant\lambda_1\leqslant\cdots\leqslant\lambda_n$. 所以, 令

$$S_1=O'\operatorname{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})O, \tag{10.2.19}$$

则 S_1 也是半正定对称方阵, 且 $S_1^2=S$, 这就证明了存在性. 再证唯一性. 假设另有一半正定对称方阵 S_2 适合 $S_2^2=S$. 今存在实正交方阵 O_1, 使得

$$S_2=O_1'\operatorname{diag}(\mu_1,\mu_2,\cdots,\mu_n)O_1,$$

其中 $0\leqslant\mu_1\leqslant\cdots\leqslant\mu_n$. 由此可知

$$S=S_2^2=O_1'\operatorname{diag}(\mu_1^2,\mu_2^2,\cdots,\mu_n^2)O_1,$$

因此 $0 \leqslant \mu_1^2 \leqslant \cdots \leqslant \mu_n^2$ 是方阵 S 的特征根组, 所以 $\mu_1^2 = \lambda_1, \cdots, \mu_n^2 = \lambda_n$, 即 $\mu_j = \sqrt{\lambda_j}, j = 1, 2, \cdots, n$, 故

$$S_2 = O_1'\mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})O_1.$$

再由 $S = S_1^2 = S_2^2$ 可知

$$S = O'\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)O = O_1'\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)O_1.$$

所以记

$$OO_1' = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix},$$

则

$$(OO_1')\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n) = \mathrm{diag}\,(\lambda_1 \lambda_2, \cdots, \lambda_n)(OO_1'),$$

即

$$p_{jk}\lambda_k = p_{jk}\lambda_j, \quad j, k = 1, 2, \cdots, n.$$

所以, 在 $\lambda_j \neq \lambda_k$ 时 $p_{jk} = 0$. 因此恒有

$$p_{jk}\sqrt{\lambda_k} = p_{jk}\sqrt{\lambda_j}, \quad j, k = 1, 2, \cdots, n.$$

换句话说, 恒有

$$(OO_1')\mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}) = \mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})(OO_1'),$$

即

$$S_1 = O'\mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})O = O_1'\mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})O_1 = S_2.$$

这就证明了唯一性.

最后, 若方阵 A 和 S 可交换, 即 $AS = SA$, 则

$$A(O'\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)O) = (O'\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)O)A,$$

所以

$$(OAO')\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n) = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)(OAO').$$

和上面一样讨论, 可以证明

$$(OAO')\mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}) = \mathrm{diag}\,(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n})(OAO').$$

即 $AS_1 = S_1A$. 这样, 就证明了方阵 A 和 S_1 也可交换. □

注意 不能用 $\sqrt{S}$ 来表示 S_1. 而且, 一般来说, 如果不限定 S_1 是半正定对称方阵, 那末方阵 $S_1^2 = S$ 的解 S_1 不唯一.

定理 10.2.7 (极分解式) 对任一 n 阶非异实方阵 A, 存在实正交方阵 O 及两个实正定对称方阵 S_1 和 S_2, 使得

$$A = S_1O = OS_2, \tag{10.2.20}$$

且分解是唯一的.

证 令 A 非异, 故 $A'A$ 是实正定对称方阵. 于是存在实正定对称方阵 S_2, 使得 $A'A = S_2^2$. 因此

$$E_n = S_2^{-1}(S_2^2)S_2^{-1} = S_2^{-1}(A'A)S_2^{-1} = (AS_2^{-1})'(AS_2^{-1}).$$

所以 AS_2^{-1} 是实正交方阵, 记作 O. 于是 $A = OS_2$. 这就证明了极分解的存在性. 再证唯一性. 若另外存在实正交方阵 $\widetilde{O}$ 及正定对称方阵 $\widetilde{S}_2$, 使得 $A = OS_2 = \widetilde{O}\widetilde{S}_2$, 则由直接计算可知 $A'A = S_2^2 = \widetilde{S}_2^2$, 由定理 10.2.6 可知 $\widetilde{S}_2 = S_2$, 因此 $\widetilde{O} = O$, 故唯一性成立. 其次, 证 $A = S_1O$. 事实上, 由 $A = OS_2 = (OS_2O')O$, 记 $S_1 = OS_2O'$, 则由 S_2 半正定可知 S_1 也半正定, 故得极分解式 $A = S_1O$. 同法可证分解有唯一性. □

利用极分解式, 可以证明

定理 10.2.8 对任一 n 阶非异实方阵 A, 则存在实正交方阵 O_3 及 O_4, 使得

$$O_3AO_4 = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n), \tag{10.2.21}$$

其中 $0 < \lambda_n \leqslant \lambda_{n-1} \leqslant \cdots \leqslant \lambda_1$, 又 $\lambda_n^2, \lambda_{n-1}^2, \cdots, \lambda_1^2$ 是正定对称方阵 AA' 的所有特征根.

证 令 n 阶非异实方阵 A 有极分解 $A = S_1O_1$, 其中 $S_1^2 = AA'$. 将 S_1 化为标准形, 即存在实正交方阵 O_3, 使得

$$O_3S_1O_3' = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

于是 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$ 构成 $S_1^2 = AA'$ 的所有特征根. 再者, 极分解式变成

$$A = S_1O_1 = O_3'\mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)(O_3O_1).$$

所以, 记 $O_4 = (O_3O_1)'$, 便证明了定理. □

最后给出奇异方阵的极分解. 因为利用这类讨论, 实际上可以给出更多的内容, 所以下面先从矩阵讲起.

定义 10.2.9 两个 $n \times m$ 矩阵 A 和 B 称为**正交相抵的**, 如果存在一个 n 阶正

交方阵 O_1 和一个 m 阶正交方阵 O_2, 使得

$$B = O_1 A O_2. \tag{10.2.22}$$

显然, 正交相抵是一个等价关系. 下面给出在正交相抵下的标准形. 首先有

定理 10.2.10 对任一秩是 r 的 $n \times m$ 实矩阵 A, 存在 n 阶实正交方阵 O_1 和 m 阶实正交方阵 O_2, 使得

$$O_1 A O_2 = \begin{pmatrix} A_1^{(r)} & 0^{(r,m-r)} \\ 0^{(n-r,r)} & 0^{(n-r,m-r)} \end{pmatrix}, \tag{10.2.23}$$

其中 r 阶实方阵 A_1 有 $\det(A_1) \neq 0$.

证 先证存在 n 阶实正交方阵 O_1, 使得

$$A_0 = O_1 A = \begin{pmatrix} A_{11}^{(r,m)} \\ 0^{(n-r,m)} \end{pmatrix},$$

其中 A_{11} 的秩是 r. 为此, 对矩阵 A 的行数 n 作归纳法. 当 $n=1$ 时显然成立. 设在 $n-1$ 时断言成立, 今考虑行数是 n 的矩阵 A. 设 A 的秩是 n, 取 O_1 是单位方阵即成. 设 A 的秩 r 小于 n, 则 n 个未知数 $x_1, x_2, \cdots, x_n$ 的齐次线性方程组 $A'(x_1, x_2, \cdots, x_n)' = 0$ 有无穷多个非零解. 取一个单位长度的解向量 α, 它是 $n \times 1$ 矩阵, 而且有 $A'\alpha = 0$. 则存在 n 阶实正交方阵

$$O = \begin{pmatrix} V^{(n-1,n)} \\ \alpha' \end{pmatrix},$$

而

$$OA = \begin{pmatrix} V \\ \alpha' \end{pmatrix} A = \begin{pmatrix} VA \\ \alpha' A \end{pmatrix} = \begin{pmatrix} VA \\ 0 \end{pmatrix}.$$

自然, VA 是 $(n-1) \times m$ 矩阵, 它的秩和 A 的秩相等, 即秩是 r. 由归纳法假设可知, 存在 $n-1$ 阶实正交方阵 O_3, 使得

$$O_3(VA) = \begin{pmatrix} A_{11}^{(r,m)} \\ 0^{(n-r-1,m)} \end{pmatrix},$$

其中 A_{11} 的秩是 r. 因此, 取 n 阶实正交方阵

$$O_1 = \begin{pmatrix} O_3 & 0 \\ 0 & 1 \end{pmatrix} O,$$

便有

$$A_0 = O_1 A = \begin{pmatrix} O_3 & 0 \\ 0 & 1 \end{pmatrix} OA = \begin{pmatrix} O_3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} VA \\ 0 \end{pmatrix} = \begin{pmatrix} A_{11}^{(r,m)} \\ 0^{(n-r,m)} \end{pmatrix}.$$

所以由归纳法证明了断言成立. 再考虑 $m\times r$ 矩阵 A'_{11}, 已知它的秩是 r, 所以存在 m 阶实正交方阵 O'_2, 使得

$$O'_2A'_{11}=\begin{pmatrix}A'_1\\0^{(m-r,r)}\end{pmatrix},$$

其中 r 阶实方阵 A_1 的秩是 r, 即 $\det(A_1)\neq 0$. 另一方面,

$$O_1AO_2=\begin{pmatrix}A_{11}\\0\end{pmatrix}O_2=\begin{pmatrix}A_{11}O_2\\0\end{pmatrix}=\begin{pmatrix}A_1&0\\0&0\end{pmatrix}.$$ □

定理 10.2.11 设 $n\times m$ 矩阵 A 的秩是 r, 则 A 正交相抵于标准形

$$\begin{pmatrix}\Lambda&0^{(r,m-r)}\\0^{(n-r,r)}&0^{(n-r,m-r)}\end{pmatrix},\tag{10.2.24}$$

其中

$$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_r),\quad 0<\lambda_r\leqslant\cdots\leqslant\lambda_1.$$

又 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_r^2$ 是 n 阶半正定对称方阵 AA' 的所有非零特征根. 所以 n 阶方阵 AA' 的所有非零特征根是矩阵 A 在正交相抵下的全系不变量.

证 由定理 10.2.10, 存在实正交方阵 $O_1^{(n)}$ 及 $O_2^{(m)}$, 使得

$$O_1AO_2=\begin{pmatrix}A_1^{(r)}&0\\0&0\end{pmatrix}.$$

今 A_1 是 r 阶非异实方阵, 由定理 10.2.8, 存在 r 阶正交方阵 O_3 及 O_4, 使得

$$O_3A_1O_4=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_r),$$

其中 $0<\lambda_r\leqslant\cdots\leqslant\lambda_1$, 又 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_r^2$ 是正定对称方阵 $A_1A'_1$ 的所有特征根. 所以

$$\begin{aligned}&\begin{pmatrix}O_3^{(r)}&0\\0&E_{n-r}\end{pmatrix}O_1AO_2\begin{pmatrix}O_4^{(r)}&0\\0&E_{m-r}\end{pmatrix}\\&=\begin{pmatrix}O_3&0\\0&E\end{pmatrix}\begin{pmatrix}A_1&0\\0&0\end{pmatrix}\begin{pmatrix}O_4&0\\0&E\end{pmatrix}=\begin{pmatrix}\Lambda&0\\0&0\end{pmatrix},\end{aligned}$$

即存在实正交方阵 $O_5^{(n)}$ 及 $O_6^{(m)}$, 使得

$$O_5AO_6=\begin{pmatrix}\Lambda&0\\0&0\end{pmatrix}.$$

于是

$$AA'=O'_5\begin{pmatrix}\Lambda&0^{(r,m-r)}\\0^{(n-r,r)}&0^{(n-r,m-r)}\end{pmatrix}O'_6O_6\begin{pmatrix}\Lambda&0^{(r,n-r)}\\0^{(m-r,r)}&0^{(m-r,n-r)}\end{pmatrix}O_5$$

$$= O_5' \begin{pmatrix} \Lambda^2 & 0 \\ 0 & 0^{(n-r)} \end{pmatrix} O_5.$$

因为 O_5 是 n 阶实正交方阵, 所以 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_r^2$ 是方阵 AA' 的所有非零特征根. □

由定理 10.2.11 立即可得

推论 10.2.12　对任一 $n \times m$ 实矩阵 A, 则 n 阶半正定对称方阵 AA' 和 m 阶半正定对称方阵 $A'A$ 的非零特征根完全相同. 特别,

$$\mathrm{rank}\,(AA') = \mathrm{rank}\,(A'A) = \mathrm{rank}\,(A).$$

再者, 在方阵的情形, 立即得到

定理 10.2.13 (极分解式)　设 A 是 n 阶实方阵, 则存在 n 阶实正交方阵 O 和 n 阶半正定实对称方阵 S_1 及 S_2, 使得

$$A = S_1 O = O S_2. \tag{10.2.25}$$

证　由定理 10.2.11 可知, 存在 n 阶实正交方阵 P 及 Q, 使得

$$A = P\Lambda Q, \qquad \Lambda = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

其中 $0 \leqslant \lambda_n \leqslant \cdots \leqslant \lambda_1$, 又 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$ 是方阵 AA' 的特征根. 将 A 的分解式改写为

$$A = P\Lambda Q = P\Lambda P'(PQ) = (PQ)Q'\Lambda Q.$$

记 $S_1 = P\Lambda P'$, $S_2 = Q'\Lambda Q$, 则 S_1 和 S_2 都是半正定对称方阵. 再记 $O = PQ$, 则 O 是实正交方阵, 而且有 $A = S_1 O = O S_2$. □

注意　在方阵 A 奇异的时候, 极分解式不唯一 (为什么?).

习　题　10.2

10.2.1　试证: 实对称方阵 $S > 0$ 的必要且充分条件是 $S^{-1} > 0$.

10.2.2　试给出实对称方阵负定以及半负定的必要且充分条件.

10.2.3　设 S 是实对称方阵, 试证: 存在正实数 μ, 使得 $\mu E_n + S > 0$, $(-\mu)E_n + S < 0$.

10.2.4　设半正定实对称方阵 S_1 和 S_2 可交换, 试证: $S_1 S_2$ 仍为半正定实对称方阵.

10.2.5　设 S 是正定实对称方阵. 试证: 存在非异的上 (下) 三角实方阵 Q, 使得 $S = Q'Q$. 它唯一吗? 为什么.

10.2.6　试证: 实方阵 A 是实规范方阵的必要且充分条件是它有极分解式

$$A = SO = OS,$$

其中 S 为半正定实对称方阵, O 实正交方阵.

10.2.7 试证: 任一秩为 1 的半正定实对称方阵 S 必有 $S=\alpha\alpha'$, 其中 α 是 $n\times 1$ 非零实矩阵.

10.2.8 设 $A=(a_{jk})$ 和 $B=(b_{jk})$ 都是 n 阶正定实对称方阵, 试证: 实方阵

$$C=(a_{jk}b_{jk})$$

也是 n 阶正定实对称方阵.

10.2.9 设实对称方阵 S 的 $n-1$ 个主子式

$$\det S\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix},\quad k=1,2,\cdots,n-1$$

都是正实数, 又 $\det(S)\geqslant 0$, 试证 $S\geqslant 0$.

10.2.10 如果 $S=(a_{jk})$ 是正定实对称方阵, 则

$$\det(S)\leqslant a_{11}a_{22}\cdots a_{nn}.$$

又, 等号成立的必要且充分条件为 S 是实对角方阵.

10.2.11 试证明(**Hadamard 不等式**): 设 $A=(a_{jk})$ 是一个 n 阶实方阵, 则

$$\det\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \vdots& &\vdots\\ a_{n1}&\cdots&a_{nn}\end{pmatrix}\leqslant\Big(\sum_{j=1}^{n}a_{j1}^2\Big)^{\frac12}\cdots\Big(\sum_{j=1}^{n}a_{jn}^2\Big)^{\frac12}.$$

又, 等号成立的必要且充分条件是实方阵 A 的列向量两两正交.

10.2.12 设 S 和 T 都是半正定实对称方阵, 则

$$\det(S+T)\geqslant\det(S).$$

当 S 和 T 都是正定实对称方阵时, 不等号成立.

10.2.13 设 S 是半正定实对称方阵, 试证:

$$\det(S)\leqslant\det S\begin{pmatrix}12\cdots r\\12\cdots r\end{pmatrix}\det S\begin{pmatrix}(r+1)\cdots n\\(r+1)\cdots n\end{pmatrix},\quad 1\leqslant r\leqslant n.$$

又, 等号成立的必要且充分条件为 S 是实对角方阵.

10.2.14 试证: 实对称方阵 S 的特征根全部落在区间 $[a,b]$ 上的必要且充分条件是 $S-aE_n\geqslant 0$, $bE_n-S\geqslant 0$.

10.2.15 设实对称方阵 S_1 和 S_2 的特征根分别落在闭区间 $[a,b]$ 和 $[c,d]$ 内. 试证: 实对称方阵 S_1+S_2 的特征根落在闭区间 $[a+c,b+d]$ 内.

10.2.16 设 S_1 和 S_2 都是实对称方阵, 其中 $S_1\geqslant 0$. 设若 $\det(S_1+iS_2)=0$, 则存在 $n\times 1$ 非零实矩阵 α, 使得 $(S_1+iS_2)\alpha=0$.

10.2.17 设 S 是正定实对称方阵. 试证: 对任两实向量 x 和 y, 成立着

$$(xSy')^2\leqslant(xSx')(ySy').$$

又, 等式成立的必要且充分条件是 x 和 y 线性相关.

10.2.18 试证: 对一切 $n\times 1$ 非零实矩阵 $x=(x_1,x_2,\cdots,x_n)'$, 成立着

$$\sum_{j=1}^{n}x_j^2-\sum_{j=1}^{n-1}x_jx_{j+1}>0.$$

试求它的最小值, 并求它在限制条件 $x_n=1$ 下的最小值.

10.2.19 试求 n 元实函数

$$f(x_1,x_2,\cdots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j+2\sum_{j=1}^{n}b_jx_j+c$$

的极值, 其中 n 阶实方阵 S 是正定实对称方阵.

10.2.20 设 A 和 B 都是 n 阶正定实对称方阵, P 是 n 阶实方阵. 试证: $A>P'B^{-1}P$ 当且仅当 $B>P'A^{-1}P$.

10.2.21 设 A 是 n 阶正定实对称方阵, B 是 $n\times m$ 实矩阵. 设 $\operatorname{rank}(B)=m$, 试求实方阵

$$\begin{pmatrix} A^{(n)} & B^{(n,m)} \\ B' & 0^{(m)} \end{pmatrix}$$

的逆方阵.

10.2.22 设 A 是 n 阶正定实对称方阵. 试证:

$$A+A^{-1}-2E_n\geqslant 0.$$

又

$$(Ax,x)+(A^{-1}x,x)\geqslant 2(x,x).$$

何时等式成立?

10.2.23 设 A 是 n 阶实方阵, $S=\dfrac{1}{2}(A+A')>0$. 试证:

$$\det(A)\geqslant\det(S).$$

10.2.24 设 $S=(a_{ij})$ 是 n 阶正定实对称方阵, 试证:

$$\sum_{i,j\neq k}x_ix_j\det\begin{pmatrix} a_{kk} & a_{ik} \\ a_{kj} & a_{ij} \end{pmatrix}$$

是正定实二次型, $k=1,\ 2,\ \cdots,\ n$.

10.2.25 设 A 是 n 阶正定实对称方阵. 试证: n 维实空间 $\mathbb{R}^n$ 中由不等式 $x'Ax\leqslant 1$ 所定义的区域是有界的, 且它的体积

$$V=\int_{x'Ax\leqslant 1}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=\frac{\pi^{\frac{n}{2}}}{\Gamma\left(\dfrac{n}{2}+1\right)}(\det(A))^{-\frac{1}{2}}.$$

10.2.26 设 S 是 n 阶正定实对称方阵. 作方阵序列

$$X_0=E_n,\quad X_{k+1}=\frac{1}{2}(X_k+SX_k^{-1}),\quad k=1,2,\cdots.$$

试证: $\lim\limits_{k\to\infty} X_k$ 存在, 记作 B. 则 B 是 n 阶正定实对称方阵, 且 $B^2 = S$. 又由 $SP = PS$ 可推出 $BP = PB$, 其中 P 是 n 阶实方阵.

10.2.27 设 A 是 n 阶半正定实对称方阵, 且 $E \geqslant A$, 即 $E - A \geqslant 0$. 试证: 任取 n 阶实正交方阵 O, 则有

$$\det(E - AO) \geqslant \det(E - A).$$

10.2.28 试证: $\operatorname{rank}(A'AB) = \operatorname{rank}(AB)$, 其中 A 是 $m \times n$ 实矩阵, B 是 $n \times p$ 实矩阵.

10.2.29 设 A 和 B 是 n 阶实对称方阵. 试证: $AB = 0$ 当且仅当

$$\lambda^n \det(\lambda E_n - A - B) = \det(\lambda E_n - A)\det(\lambda E_n - B).$$

10.2.30 设 n 元函数 $f(x_1, x_2, \cdots, x_n)$ 在点 $p_0 = (x_1^0, x_2^0, \cdots, x_n^0)$ 及其邻域连续, 且有二阶连续偏微商. 记

$$\partial_x f = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right),$$

$$S = \partial_x' \partial_x f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

试证: 如果在 $x = x_0$ 时 $\partial_x f = 0$, 则当实对称方阵 S 在点 $x = x_0$ 时 $S > 0$, 点 $x = x_0$ 是极小值点; 当实对称方阵 S 在点 $x = x_0$ 时 $S < 0$, 点 $x = x_0$ 是极大值点; 当 $\operatorname{rank}(S) = n$, 且 S 不是正定也不是负定时, 点 $x = x_0$ 不是极值点.

10.2.31 (Hurwicz 矩阵方程) 给定 p 个 $n \times m$ 实矩阵 $A_1, A_2, \cdots, A_p$. 任取 n 阶实正交方阵 O_1 及 m 阶实正交方阵 O_2, 则矩阵组 $A_1, A_2, \cdots, A_p$ 和 $B_1 = O_1A_1O_2$, $B_2 = O_1A_2O_2, \cdots, B_p = O_1A_pO_2$ 称为**正交相抵的**.

假设实矩阵组 $A_1, A_2, \cdots, A_p$ 适合条件:

$$A_j'A_k + A_k'A_j = 2\delta_{jk}E_n, \quad j, k = 1, 2, \cdots, p.$$

显然, 实矩阵组 $B_1, B_2, \cdots, B_p$ 也适合此关系. 换句话说, 此关系在正交相抵下不变.

设 $n = m = p$, 试证: 这时必须 $n = 1, 2, 4, 8$. 试分别定出 $n = 1, 2, 4, 8$ 时, n 阶实方阵组 $A_1, A_2, \cdots, A_n$ 在正交相抵下的标准形.

(注: 在一般情形, 当 $p = 2$ 时已求得标准形, 当 $p > 2$ 时, 适合上面条件的实矩阵组 $A_1^{(n,m)}, \cdots, A_p^{(n,m)}$ 在正交相抵下的标准形, 迄今未求得. 即使下面的问题也没有得到解决: 给定 n 和 m 后, p 自然是 n 及 m 的函数, 试定出这个函数.)

§10.3 反对称方阵在相合下的标准形 *

定理 10.3.1 对域 $\mathbb{F}$ 上任一 n 阶反对称方阵 K, 存在一个 n 阶方阵 P, 使得 P 的元素是方阵 K 的元素的**有理函数**(即为两个多元多项式之商), 又 $\det(P) = 1$, 且

$$P'KP = \operatorname{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}, 0_{n-2s}\right), \tag{10.3.1}$$

其中 $a_1, a_2, \cdots, a_s$ 都是域 $\mathbb{F}$ 的非零数, 它们也是方阵 K 的元素的有理函数.

证 对阶数 n 用归纳法. 在 $n=1$ 时定理显然成立. 设在阶数 $\leqslant n-1$ 时定理成立, 下面来证在阶数是 n 时定理也成立.

当 $K=0$ 时, 取 $P=E_n$ 即成. 今设 $K \neq 0$, 于是, 存在元素 $a_{jk} \neq 0$, 其中 $j<k$. 作相合 $(P_{2k}P_{1j})K(P_{2k}P_{1j})'$, 于是, 后一反对称方阵的第一行、第二列元素是 $a_{jk} \neq 0$. 另一方面, 方阵 $P_{2k}P_{1j}$ 是行列式为 1 的常数方阵, 所以适合定理要求. 因此无妨设 $a_{12} \neq 0$, 即记

$$K=\begin{pmatrix} \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} & A^{(2,n-2)} \\ -A' & K_1^{(n-2)} \end{pmatrix}, \quad P_3=\begin{pmatrix} E_2 & 0 \\ A'\begin{pmatrix} 0 & -a_{12}^{-1} \\ a_{12}^{-1} & 0 \end{pmatrix} & E_{n-2} \end{pmatrix},$$

则

$$P_3\begin{pmatrix} \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} & A \\ -A' & K_1 \end{pmatrix}P_3'=\operatorname{diag}\left(\begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix}, K_0\right),$$

其中

$$K_0=K_1-A'\begin{pmatrix} 0 & a_{12}^{-1} \\ -a_{12}^{-1} & 0 \end{pmatrix}A.$$

显然, 方阵 P_3 适合定理要求. 且 K_0 仍是反对称方阵, 其中元素是方阵 K 的元素的有理函数.

对 $n-2$ 阶反对称方阵 K_0 用归纳法假设, 这就证明了: 存在 $n-2$ 阶方阵 P_2, 它适合 $\det(P_2)=1$, 且 P_2 的元素是方阵 K_0 的元素的有理函数. 由于有理函数的有理函数仍为有理函数, 所以方阵 P_2 的元素是反对称方阵 K 的元素的有理函数. 另一方面,

$$P_2K_0P_2'=\operatorname{diag}\left(\begin{pmatrix} 0 & a_2 \\ -a_2 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}, 0, \cdots, 0\right),$$

其中 $a_2, a_3, \cdots, a_s$ 都是方阵 K 的元素的有理函数. 令

$$P=\begin{pmatrix} E_2 & 0 \\ 0 & P_2 \end{pmatrix}\begin{pmatrix} E_2 & 0 \\ A'\begin{pmatrix} 0 & -a_{12}^{-1} \\ a_{12}^{-1} & 0 \end{pmatrix} & E_{n-2} \end{pmatrix}P_1.$$

显然, 记 a_{12} 是 a_1, 则方阵 P 适合定理要求, 这就证明了定理. □

定理 10.3.2 域 $\mathbb{F}$ 的任一 n 阶反对称方阵 K 必相合于标准形

$$\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0_{n-2s}\right), \tag{10.3.2}$$

其中 $r=2s$ 是方阵 K 的秩, 所以反对称方阵的秩是它在相合下的全系不变量.

证 由定理 10.3.1 可知, 存在域 $\mathbb{F}$ 的非异方阵 P, 使得

$$P'KP=\operatorname{diag}\left(\begin{pmatrix}0 & a_1\\ -a_1 & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & a_s\\ -a_s & 0\end{pmatrix},0_{n-2s}\right),$$

其中 $a_1, a_2, \cdots, a_s$ 都是域 $\mathbb{F}$ 的非零数. 因为方阵的秩在相合下不变, 所以方阵 K 的秩 r 等于 $2s$. 又由

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

所以令

$$Q=\operatorname{diag}\left(\begin{pmatrix}1 & -1\\ 0 & a_1^{-1}\end{pmatrix},\cdots,\begin{pmatrix}1 & -1\\ 0 & a_s^{-1}\end{pmatrix},E_{n-2s}\right)P,$$

则 QKQ' 就是标准形. □

从这个定理立即可得

推论 10.3.3 域 $\mathbb{F}$ 的反对称方阵的秩必是偶数.

定义 10.3.4 设 $f(\alpha,\beta)$ 是域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 的双线性函数. $\mathfrak{L}$ 的子空间

$$\ker(f)=\{\,\beta\in\mathfrak{L}\mid f(\alpha,\beta)=0,\ \forall\alpha\in\mathfrak{L}\,\} \tag{10.3.3}$$

称为双线性函数 $f(\alpha,\beta)$ 的**核**. 当核 $\ker(f)=0$ 时, 双线性函数 $f(\alpha,\beta)$ 称为**非退化的**.

域 $\mathbb{F}$ 上 n 维线性空间 $\mathfrak{L}$ 中任取一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得在这组基下, 双线性函数 $f(\alpha,\beta)$ 的方阵表示为 A, 则双线性函数 $f(\alpha,\beta)$ 非退化当且仅当 $\det(A)\neq 0$.

定理 10.3.5 对域 $\mathbb{F}$ 的任一 m 维线性空间 $\mathfrak{L}$ 的反对称双线性函数 $f(x,y)$, 则存在一组基, 使得在这组基下, 函数 $f(x,y)$ 的坐标表达式具有下面的标准形:

$$f(x,y)=(x_1y_2-x_2y_1)+\cdots+(x_{2s-1}y_{2s}-x_{2s}y_{2s-1}), \tag{10.3.4}$$

其中 $2s=r$ 称为反对称双线性函数 $f(x,y)$ 的秩, 它是反对称双线性函数在坐标变换下的全系不变量. 特别当反对称双线性函数 $f(x,y)$ 的秩 $r=m$ 时, 即反对称双线性函数 $f(x,y)$ 非退化时, m 是偶数 $2n$, 记

$$J=\begin{pmatrix}0 & E_n\\ -E_n & 0\end{pmatrix}, \tag{10.3.5}$$

则反对称双线性函数 $f(x,y)$ 的坐标表达式为

$$f(x,y)=x'Jy. \tag{10.3.6}$$

上面的定理10.3.1 及定理10.3.2 的证明, 实际上也给出了求反对称方阵或反对称双线性函数化标准形的方法.

下面介绍辛空间. 这部分内容是浙江大学数学系李方教授所写的讲义, 他同意提供我们使用的, 在此表示感谢.

定义 10.3.6　设 $f(\alpha,\beta)$ 是域 $\mathbb{F}$ 上 $2n$ 维线性空间 $\mathfrak{L}$ 的非退化反对称双线性函数, 则线性空间 $\mathfrak{L}$ 称为域 $\mathbb{F}$ 上的**辛空间**. 这时 $f(\alpha,\beta)$ 记作 $<\alpha,\beta>$. 设 $(\mathfrak{L},<\alpha,\beta>)$ 为辛空间. $\mathfrak{L}$ 中向量 α 和 β 称为**互相正交的**, 如果 $<\alpha,\beta>=0$. 设 $\mathfrak{L}_0$ 为辛空间 $\mathfrak{L}$ 的子空间, 则 $\mathfrak{L}$ 中子集

$$\mathfrak{L}_0^{\perp}=\{\,\alpha\in\mathfrak{L}\mid(\alpha,\mathfrak{L}_0)=0\,\}\tag{10.3.7}$$

称为子空间 $\mathfrak{L}_0$ 的**正交补**. 显然 $\mathfrak{L}_0^{\perp}$ 是 $\mathfrak{L}$ 的子空间.

定义 10.3.7　设 $(\mathfrak{L},<\alpha,\beta>)$ 为辛空间, $\mathfrak{L}$ 中的一组基 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n,\ \alpha_{-1},\ \alpha_{-2},\ \cdots,\ \alpha_{-n}$ 称为**辛正交基**, 如果非退化反对称双线性函数 $<\alpha,\beta>$ 在这组基下的方阵表示为 J. 即

$$<\alpha_i,\alpha_j>=<\alpha_{-i},\alpha_{-j}>=0,\quad <\alpha_i,\alpha_{-j}>=\delta_{ij},\quad 1\leqslant i,j\leqslant n.\tag{10.3.8}$$

定理10.3.8　设 $(\mathfrak{L},<\alpha,\beta>)$ 为辛空间, 则辛正交基到辛正交基的基变换的方阵表示 A 为辛方阵.

证　任取两组辛正交基, 基变换的方阵表示为 A. 由 $<\alpha,\beta>$ 在这两组基下的方阵表示都是 J, 则有 $A'JA=J$. 即 A 为辛方阵. □

定义 10.3.9　记 $\mathrm{Sp}\,(n,\mathbb{F})$ 为域 $\mathbb{F}$ 上 $2n$ 阶辛方阵全体构成的集合

$$\mathrm{Sp}\,(n,\mathbb{F})=\{\,A\in\mathrm{GL}\,(n,\mathbb{F})\ \mid\ A'JA=J\,\},\tag{10.3.9}$$

称为 $2n$ 阶**辛群**.

由定义易证

引理 10.3.10　域 $\mathbb{F}$ 上 $2n$ 阶辛群 $\mathrm{Sp}\,(n,\mathbb{F})$ 有性质:

(1)　任取 $A_1,\ A_2\in\mathrm{Sp}\,(n,\mathbb{F})$, 则 $A_1A_2\in\mathrm{Sp}\,(n,\mathbb{F})$;

(2)　辛方阵的乘法适合结合律;

(3)　$2n$ 阶单位方阵 $E_n\in\mathrm{Sp}\,(n,\mathbb{F})$;

(4)　任取 $A\in\mathrm{Sp}\,(n,\mathbb{F})$, 则 $A^{-1}\in\mathrm{Sp}\,(n,\mathbb{F}),\ A'\in\mathrm{Sp}\,(n,\mathbb{F}),\bar{A}\in\mathrm{Sp}\,(n,\mathbb{F})$;

(5)　辛方阵的行列式等于 1.

定义 10.3.11　域 $\mathbb{F}$ 上 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 的保持非退化反对称双线性函数 $<\alpha,\beta>$ 的线性变换称为**辛变换**.

定理 10.3.12　在域 $\mathbb{F}$ 上 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 中取定辛正交基 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n$. 线性变换 $\mathscr{A}$ 在这组基下的方阵表示记为 A, 则线性变换 $\mathscr{A}$ 是辛变

换的必要且充分条件是

$$A'JA = J. \tag{10.3.10}$$

证 今任取向量 α 和 β, 它在线性变换 $\mathscr{A}$ 下的象是向量 $\mathscr{A}(\alpha)$ 和 $\mathscr{A}(\beta)$. 由定义 $<\alpha,\beta>=<\mathscr{A}(\alpha),\mathscr{A}(\beta)>$. 写为坐标表达式, 有 $x'Jy=(Ax)'J(Ay)=x'(A'JA)y$. 因此 $A'JA=J$, 即 A 为辛方阵. □

利用辛方阵的基本性质, 可以证明辛变换有性质:

(1) 辛变换的乘积仍是辛变换;

(2) 辛变换的乘法适合结合律;

(3) 恒等变换是辛变换;

(4) 辛变换的逆变换也是辛变换;

(5) 辛变换的行列式等于 1.

由定理10.3.5, 在辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 中必存在辛正交基. 上面讨论也告诉我们, 域 $\mathbb{F}$ 上 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 必同构于域 $\mathbb{F}$ 上 $2n$ 维辛空间 $(\mathbb{F}^{2n},x'Jy)$. 所以域 $\mathbb{F}$ 上 $2n$ 维辛空间在同构意义下唯一.

定理 10.3.13 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 中子空间 $\mathfrak{L}_0$ 有

$$\dim(\mathfrak{L})=\dim(\mathfrak{L}_0)+\dim(\mathfrak{L}_0^{\perp}) \tag{10.3.11}$$

证 在子空间 $\mathfrak{L}_0$ 中任取基 $\alpha_1,\cdots,\alpha_r$, 于是在辛空间 $\mathfrak{L}$ 中存在基 $\alpha_1,\cdots,\alpha_r$, $\alpha_{r+1},\cdots,\alpha_{2n}$. 任取 $\alpha=\sum\limits_{j=1}^{r}x_j\alpha_j\in\mathfrak{L}_0$, $\beta=\sum\limits_{k=1}^{2n}y_k\alpha_k\in\mathfrak{L}_0^{\perp}$, 则有 $x'Ay=0$, 其中 $x=(x_1,\cdots,x_r,0,\cdots,0)'$, $y=(y_1,\cdots,y_{2n})'$, A 是 $2n$ 阶非异反对称方阵. 于是 $2n\times 1$ 矩阵 y 适合条件 $e_j'Ay=0$, $1\leqslant j\leqslant r$. 它是 r 个线性方程构成的齐次线性方程组, 系数矩阵

$$\begin{pmatrix} e_1'A \\ \vdots \\ e_r'A \end{pmatrix}$$

的秩为 r. 因此, 解空间的维数为 $2n-r$. 这证明了子空间 $\mathfrak{L}_0$ 的正交补 $\mathfrak{L}_0^{\perp}$ 有$\dim(\mathfrak{L}_0^{\perp})=2n-r=\dim(\mathfrak{L})-\dim(\mathfrak{L}_0)$. □

定理 10.3.14 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 中子空间 $\mathfrak{L}_1$ 和 $\mathfrak{L}_2$ 的正交补 $\mathfrak{L}_1^{\perp}$ 和 $\mathfrak{L}_2^{\perp}$ 有性质:

(a) 设 $\mathfrak{L}_1\subset\mathfrak{L}_2$, 则 $\mathfrak{L}_2^{\perp}\subset\mathfrak{L}_1^{\perp}$;

(b) $(\mathfrak{L}_1^{\perp})^{\perp}=\mathfrak{L}_1$.

证 设 $\mathfrak{L}_1\subset\mathfrak{L}_2$, 任取 $\alpha\in\mathfrak{L}_2^{\perp}$, 则有$<\alpha,\mathfrak{L}_2>=0$. 由 $\mathfrak{L}_1\subset\mathfrak{L}_2$, 所以 $<\alpha,\mathfrak{L}_1>=0$. 因此, $\alpha\in\mathfrak{L}_1^{\perp}$. 这证明了 (a).

又由 $\mathfrak{L}_1^{\perp}$ 的定义可知 $<\mathfrak{L}_1, \mathfrak{L}_1^{\perp}>=0$. 因此, $\mathfrak{L}_1 \subset (\mathfrak{L}_1^{\perp})^{\perp}$. 由定理10.3.13, 有 $\dim(\mathfrak{L}) = \dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_1^{\perp})$, $\dim(\mathfrak{L}) = \dim(\mathfrak{L}_1^{\perp}) + \dim((\mathfrak{L}_1^{\perp})^{\perp})$. 所以$\dim(\mathfrak{L}_1) = \dim((\mathfrak{L}_1^{\perp})^{\perp})$. 这证明了 (b). □

定义 10.3.15 设 $(\mathfrak{L}, <\alpha, \beta>)$ 为辛空间.

(1) $\mathfrak{L}$ 中子空间$\mathfrak{L}_1$ 称为**辛子空间**, 如果 $\mathfrak{L}_1 \cap \mathfrak{L}_1^{\perp} = 0$;

(2) $\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$ 称为**迷向子空间**, 如果 $<\mathfrak{L}_1, \mathfrak{L}_1>=0$, 即 $\mathfrak{L}_1 \subset \mathfrak{L}_1^{\perp}$, 称为**极大迷向子空间**, 如果 $\mathfrak{L}_1$ 为迷向子空间, 而且 $\mathfrak{L}$ 的子空间 $\mathfrak{L}_2$ 有 $\mathfrak{L}_1 \subset \mathfrak{L}_2$, $\mathfrak{L}_1 \neq \mathfrak{L}_2$, 则 $\mathfrak{L}_2$ 不是迷向子空间;

(3) $\mathfrak{L}$ 中子空间 $\mathfrak{L}_1$ 称为 **Lagrange 子空间**, 如果 $\mathfrak{L}_1 = \mathfrak{L}_1^{\perp}$.

由上面定义可知, 辛子空间和Lagrange 子空间是两种完全不同的子空间. 在一般情形, 子空间 $\mathfrak{L}_0$ 有 $\mathfrak{L}_0 \cap \mathfrak{L}_0^{\perp} \neq 0$, $\mathfrak{L}_0 \neq \mathfrak{L}_0^{\perp}$.

注意 由于 $f(\alpha, \alpha) = 0$, 因此, 一维子空间是迷向子空间. 所以在辛空间中不能引进长度和角度, 只能定义互相正交, 而且正交的性质和Euclid 空间有本质不同.

下面考虑辛空间 $(\mathfrak{L}, <\alpha, \beta>)$ 的辛子空间.

定理 10.3.16 设 $\mathfrak{L}_0$ 是辛空间 $(\mathfrak{L}, <\alpha, \beta>)$ 的辛子空间, 则式(9.3.11) 等价于

$$\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_1^{\perp}. \tag{10.3.12}$$

而且 $\mathfrak{L}_1^{\perp}$ 仍为辛子空间. 又非退化反对称双线性函数 $<\alpha, \beta>$ 限制在辛子空间 $\mathfrak{L}_1$ 上仍为非退化反对称双线性函数, 记作 $<\alpha, \beta>|_{\mathfrak{L}_1}$, 即 $(\mathfrak{L}_1, <\alpha, \beta>|_{\mathfrak{L}_1})$ 仍为辛空间.

证 由定义可知式(9.3.12) 成立. 所以在 $\mathfrak{L}_1$ 中取基 $\alpha_1, \cdots, \alpha_r$, 在 $\mathfrak{L}_1^{\perp}$ 中取基 $\alpha_{r+1}, \cdots, \alpha_{2n}$, 则由 $<\mathfrak{L}_1, \mathfrak{L}_1^{\perp}>=0$ 有 $<\alpha, \beta>$ 的方阵表示为 $A = \mathrm{diag}(A_1, A_2)$. 由 $\det(A) \neq 0$ 可知 A_1 和 A_2 都是非异反对称方阵. 因此, $r = 2s$, $2n - r = 2(n - s)$. 这证明了 $(\mathfrak{L}_1, <\alpha, \beta>|_{\mathfrak{L}_1})$ 为辛空间. □

由定理10.3.2, 定理10.3.13 和定理10.3.16, 有

定理 10.3.17 设 $\mathfrak{L}_1$ 是辛空间 $(\mathfrak{L}, <\alpha, \beta>)$ 的辛子空间, 则在 $\mathfrak{L}$ 中存在辛正交基 $\alpha_1, \cdots, \alpha_n, \alpha_{-1}, \cdots, \alpha_{-n}$, 使得 $\alpha_1, \cdots, \alpha_r, \alpha_{-1}, \cdots, \alpha_{-r}$ 为 $\mathfrak{L}_1$ 的辛正交基, $\alpha_{r+1}, \cdots, \alpha_n, \alpha_{-(r+1)}, \cdots, \alpha_{-n}$ 为 $\mathfrak{L}_1^{\perp}$ 的辛正交基.

下面考虑辛空间 $(\mathfrak{L}, <\alpha, \beta>)$ 的 Lagrange 子空间.

定理 10.3.18 设 $\mathfrak{L}_1$ 是 $2n$ 维辛空间 $(\mathfrak{L}, <\alpha, \beta>)$ 的迷向子空间, 则 $\dim(\mathfrak{L}_1) \leqslant n$.

证 由定理10.3.13, $2n = \dim(\mathfrak{L}) = \dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_1^{\perp})$. 由 $\mathfrak{L}_1 \subset \mathfrak{L}_1^{\perp}$, 所以

$2n \geqslant 2\dim(\mathfrak{L}_1)$. □

定理 10.3.19 设 $\mathfrak{L}_1$ 是辛空间 $(\mathfrak{L}, <\alpha,\beta>)$ 的迷向子空间, 如果 $\mathfrak{L}_1$ 不是 Lagrange 子空间, 则存在 $\dim(\mathfrak{L}_1)+1$ 维迷向子空间 $\mathfrak{L}_3 = \mathfrak{L}_1 \oplus \mathfrak{L}_2$, 其中 $\mathfrak{L}_2$ 为一维子空间, 基 $\alpha_0 \in \mathfrak{L}_1^{\perp} - \mathfrak{L}_1$.

证 由 $\mathfrak{L}_1$ 为迷向子空间, 即 $\mathfrak{L}_1 \subset \mathfrak{L}_1^{\perp}$. 但是 $\mathfrak{L}_1$ 不是 Lagrange 子空间, 因此, $\mathfrak{L}_1 \neq \mathfrak{L}_1^{\perp}$. 所以任取 $\alpha_0 \in \mathfrak{L}_1^{\perp} - \mathfrak{L}_1$, 则 $<\alpha_0,\alpha_0>= 0$, $<\mathfrak{L}_1,\alpha_0>= 0$. 又 $<\mathfrak{L}_1,\mathfrak{L}_1>= 0$. 所以 $<\mathfrak{L}_3,\mathfrak{L}_3>= 0$. 因此, $\mathfrak{L}_3 \subset \mathfrak{L}_3^{\perp}$. 即 $\mathfrak{L}_3$ 为 $\dim(\mathfrak{L}_1)+1$ 维迷向子空间. □

定理 10.3.20 $\mathfrak{L}_1$ 是 $2n$ 维辛空间 $(\mathfrak{L}, <\alpha,\beta>)$ 的极大迷向子空间当且仅当 $\mathfrak{L}_1$ 为 Lagrange 子空间, 即 $\mathfrak{L}_1 = \mathfrak{L}_1^{\perp}$. 而且 Lagrange 子空间的维数为$n$.

证 由定理10.3.19, 如果迷向子空间 $\mathfrak{L}_1$ 不是 Lagrange 子空间, 则 $\mathfrak{L}_1$ 不是极大迷向子空间. 反之, 设 $\mathfrak{L}_1$ 不是极大迷向子空间, 则存在迷向子空间 $\mathfrak{L}_2$, 使得 $\mathfrak{L}_1 \subset \mathfrak{L}_2 \subset \mathfrak{L}_2^{\perp}$, $\mathfrak{L}_1 \neq \mathfrak{L}_2$. 因此, $\mathfrak{L}_2 \subset \mathfrak{L}_2^{\perp} \subset \mathfrak{L}_1^{\perp}$. 这证明了 $\mathfrak{L}_1 \subset \mathfrak{L}_1^{\perp}$, $\mathfrak{L}_1 \neq \mathfrak{L}_1^{\perp}$. 所以 $\mathfrak{L}_1$ 不是 Lagrange 子空间.

最后, 由定理10.3.13, $2n = \dim(\mathfrak{L}) = \dim(\mathfrak{L}_1) + \dim(\mathfrak{L}_1^{\perp})$. 而 $\mathfrak{L}_1$ 为 Lagrange 子空间, 所以 $\mathfrak{L}_1 = \mathfrak{L}_1^{\perp}$. 因此, $\dim(\mathfrak{L}_1) = n$. □

定理 10.3.21 设 $\mathfrak{L}_+$ 是 $2n$ 维辛空间 $(\mathfrak{L}, <\alpha,\beta>)$ 的 Lagrange 子空间, 即 $\mathfrak{L}_+ = \mathfrak{L}_+^{\perp}$. 任取 $\mathfrak{L}_+$ 中一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则在 $\mathfrak{L}$ 中存在 n 个向量 $\alpha_{-1}, \alpha_{-2}, \cdots, \alpha_{-n}$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_n, \alpha_{-1}, \alpha_{-2}, \cdots, \alpha_{-n}$ 为辛空间 $(\mathfrak{L}, <\alpha,\beta>)$ 的辛正交基. 这时, 以 $\alpha_{-1}, \alpha_{-2}, \cdots, \alpha_{-n}$ 为基的子空间 $\mathfrak{L}_-$ 仍为 Lagrange 子空间.

证 任取 $j \in \{1,2,\cdots,n\}$, $\mathfrak{L}_+$ 有子空间 $\mathfrak{M}_j$, 使得 $\mathfrak{M}_j$ 有基 $\alpha_1, \cdots, \alpha_{j-1}, \alpha_{j+1}, \cdots, \alpha_n$. 由 $\mathfrak{M}_j \subset \mathfrak{L}_+$ 可知 $\mathfrak{L}_+ = \mathfrak{L}_+^{\perp} \subset \mathfrak{M}_j^{\perp}$, 而且 $\dim(\mathfrak{M}_j^{\perp}) = 2n - \dim(\mathfrak{M}_j) = n+1 > n = \dim(\mathfrak{L}_+)$. 所以存在向量 $\beta_j \in \mathfrak{M}_j^{\perp} - \mathfrak{L}_+$, 它有 $<\alpha_k,\beta_j>= 0$, $k \neq j$, 我们来证 $<\alpha_j,\beta_j>\neq 0$. 因此, 可取$\beta_j \in \mathfrak{M}_j^{\perp} - \mathfrak{L}_+$, 使得 $<\alpha_j,\beta_j>= 1$.

事实上, 设 $<\alpha_j,\beta_j>= 0$, 则 $<\beta_j,\mathfrak{L}_+>= 0$, 因此, $\beta_j \in \mathfrak{L}_+^{\perp} = \mathfrak{L}_+$, 这和 $\beta_j \notin \mathfrak{L}_+$ 矛盾.

上面证明了存在 $\beta_j \notin \mathfrak{L}_+$, 使得 $<\alpha_i,\beta_j>= \delta_{ij}$, $i, j = 1, 2, \cdots, n$. 下面求 $\alpha_{-1}, \alpha_{-2}, \cdots, \alpha_{-n} \notin \mathfrak{L}_+$, 使得 $\alpha_1,\cdots,\alpha_n,\alpha_{-1},\cdots,\alpha_{-n}$ 为 $\mathfrak{L}$ 的辛正交基.

为此, 令 $\alpha_{-j} = \beta_j + \sum\limits_{i=1}^{j-1} x_{ij}\alpha_i = \beta_j + \sum\limits_{i=1}^{n} x_{ij}\alpha_i$, $1 \leqslant j \leqslant n$, 于是由 $<\mathfrak{L}_+,\mathfrak{L}_+>= 0$,

有

$$\begin{aligned}<\alpha_i,\alpha_{-j}> &= <\alpha_i,\beta_j+\sum_{k=1}^{j-1}x_{kj}\alpha_k> \\ &= <\alpha_i,\beta_j>+\sum_{k=1}^{j-1}x_{kj}<\alpha_i,\alpha_k>=\delta_{ij},\quad 1\leqslant i,j\leqslant n.\end{aligned}$$

下面证存在 x_{ij}, 使得$\alpha_1,\cdots,\alpha_n,\alpha_{-1},\cdots,\alpha_{-n}$ 为 $2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 的一组辛正交基. 事实上, 设 $\sum\limits_{i=1}^{n}(a_i\alpha_i+a_{-i}\alpha_{-i})=0$, 则

$$0=<\alpha_j,\sum_{i=1}^{n}(a_i\alpha_i+a_{-i}\alpha_{-i})>=\sum_{i=1}^{n}a_{-i}\delta_{ij}=a_{-j},\quad \forall 1\leqslant j\leqslant n.$$

所以 $\sum\limits_{i=1}^{n}a_i\alpha_i=0$. 因此, $a_i=0,\ 1\leqslant i\leqslant n$. 这证明了 $\alpha_1,\cdots,\alpha_n,\alpha_{-1},\cdots,\alpha_{-n}$ 为$2n$ 维辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 的一组基.

余下证存在 $x_{ij},\ 1\leqslant i<j\leqslant n$, 使得 $<\alpha_{-i},\alpha_{-j}>=0,\ 1\leqslant i<j\leqslant n$. 注意到 $i<j$, 所以$x_{ji}=0$. 而

$$\begin{aligned}<\alpha_{-i},\alpha_{-j}> &= <\beta_i+\sum_{r=1}^{i-1}x_{ri}\alpha_r,\beta_j+\sum_{s=1}^{j-1}x_{sj}\alpha_s> \\ &= <\beta_i,\beta_j>+\sum_{r=1}^{i-1}x_{ri}<\alpha_r,\beta_j>+\sum_{s=1}^{j-1}x_{sj}<\beta_i,\alpha_s> \\ &= <\beta_i,\beta_j>+\sum_{r=1}^{i-1}x_{ri}\delta_{rj}+\sum_{s=1}^{j-1}x_{sj}\delta_{is}=<\beta_i,\beta_j>+x_{ij}.\end{aligned}$$

所以取 $x_{ij}=<\beta_j,\beta_i>,\ 1\leqslant i<j\leqslant n$, 便证明了 $\alpha_1,\cdots,\alpha_n,\alpha_{-1},\cdots,\alpha_{-n}$ 为辛正交基, 即

$$<\alpha_i,\alpha_j>=<\alpha_{-i},\alpha_{-j}>=0,\quad <\alpha_i,\alpha_{-j}>=\delta_{ij},\quad 1\leqslant i,j\leqslant n.\qquad\square$$

习 题 10.3

10.3.1 试利用辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 的所有辛正交基, 求辛空间 $(\mathfrak{L},<\alpha,\beta>)$ 的所有辛子空间, 迷向子空间和 Lagrange 子空间.

10.3.2 在复的情形, 可定义 Hermite 辛空间 $(\mathfrak{L},<\alpha,\beta>)$, 其中 $\mathfrak{L}$ 为 $2n$ 维复线性空间, $<\alpha,\beta>$ 为非退化反 Hermite 函数. 哪些可以推广? 哪些不可以推广? 它有什么特点?

第十一章　复 Euclid 空间

§11.1　复 Euclid 空间

本章将把第九章和第十章的主要结果推广到 n 维复线性空间中去, 即引进复 Euclid 空间的概念. 由于相当大一部分是形式推广, 所以在下面全面叙述结果; 也着重讨论和 "实" 的情形不同的性质, 这些性质体现了复 Euclid 空间和实 Euclid 空间的本质差别. 另一方面, 从下面讨论可以发现复的情形反而比实的情形简单多了, 原因在于复数域 $\mathbb{C}$ 是最大的数集.

本章考虑的数都是复数, 而且我们在这一章将 n 阶复方阵 A 的转置且共轭方阵 $\overline{A}'$ 记成 A^*.

设 $\mathfrak{L}$ 是 n 维复线性空间. 在线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 于是线性空间 $\mathfrak{L}$ 中任一向量 α 可表为 $\alpha = \sum\limits_{j=1}^{n} a_j\alpha_j$, 其中 $n\times 1$ 矩阵 $(a_1, a_2, \cdots, a_n)' \in \mathbb{C}^n$ 是向量 α 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标, 这里, $(*)'$ 表示矩阵 $(*)$ 的转置矩阵. 在这一节, 始终在取定了基后进行讨论. 有时仍旧用符号 α 记这个 $n\times 1$ 矩阵. 显然有

引理 11.1.1　设 $\mathfrak{L}$ 是 n 维复线性空间. 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 设线性空间 $\mathfrak{L}$ 中向量 α 和 β 在这组基下的坐标分别是 $(a_1, a_2, \cdots, a_n)'$ 和 $(b_1, b_2, \cdots, b_n)'$, 则

$$(\alpha, \beta) = \sum_{j=1}^{n} a_j\overline{b_j}, \quad \forall\, \alpha, \beta \in \mathfrak{L} \tag{11.1.1}$$

定义了 n 维复线性空间 $\mathfrak{L}$ 上一个正定 Hermite 双线性函数, 它的方阵表示是 n 阶单位方阵 E_n. 又任意 $p+q$ 个向量 $\beta_1, \beta_2, \cdots, \beta_p, \gamma_1, \gamma_2, \cdots, \gamma_q$, 及 $p+q$ 个复数 b_1,

$b_2, \cdots, b_p, c_1, c_2, \cdots, c_q$, 则有

$$(\sum_{j=1}^{p} b_j\beta_j, \sum_{k=1}^{q} c_k\gamma_k) = \sum_{j=1}^{p}\sum_{k=1}^{q} b_j\overline{c_k}(\beta_j, \gamma_k). \tag{11.1.2}$$

定义 11.1.2 n 维复线性空间 $\mathfrak{L}$ 上的正定 Hermite 双线性函数 $f(\alpha,\beta)$ 称为向量 α 和 β 在 $\mathfrak{L}$ 上的 **内积**. 取定内积 $f(\alpha,\beta)$ 后, 我们改记为 (α,β).

定义 11.1.3 在 n 维复线性空间 $\mathfrak{L}$ 上给定内积 $(\alpha,\beta), \forall \alpha, \beta \in \mathfrak{L}$, 则 n 维复线性空间 $\mathfrak{L}$ 称为关于内积 (α,β) 的 **复 Euclid 空间**, 或称为**复欧氏空间**, 也称为**内积空间**.

引理 11.1.4 设 $\mathfrak{L}$ 是 n 维复线性空间. 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则在 $\mathfrak{L}$ 中存在和这组基有关的内积 (α,β), 它由式 (11.1.1) 定义. 所以, n 维复线性空间必有内积, 因此, 它关于这个内积 (α,β) 是复 Euclid 空间.

注意 在从实向复推广过程中可以看出. 从实向复推广有两种方式. 一种如上所述; 另一种是: 设 $\mathfrak{L}$ 是 n 维复线性空间, 在 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 取定复向量 $\alpha = \sum\limits_{k=1}^{n} a_k\alpha_k$ 和 $\beta = \sum\limits_{k=1}^{n} b_k\alpha_k$. 定义它们的内积是复数 $(\alpha,\beta)_0 = \sum\limits_{k=1}^{n} a_i b_i$. 然而按照这种方式来下定义, 那末无法定义向量的长度, 也就无法定义保持长度不变的复线性变换了. 从这里可以看出, 用定义 11.1.2 定义内积的方式来推广是最自然的. 另一方面, 从矩阵论的角度, 从实向复推广也有两种方式. 例如, n 阶复方阵 A 如果适合条件 $A' = A$, 则称为**对称方阵**; 如果适合条件 $A' = -A$, 则称为**反对称方阵**; 如果适合条件 $O'O = OO' = E_n$, 则称为**正交方阵**等等. 在矩阵论中也可以对它们进行一系列讨论. 然而这类方阵的研究, 在复的情形就失去了几何意义. 但是, 从矩阵论的角度来看, 还是很重要的.

利用向量的内积, 可以定义向量的长度和两向量的夹角.

定义 11.1.5 设 $\mathfrak{L}$ 是 n 维复线性空间. 在 $\mathfrak{L}$ 中取定内积 (α,β). 非负实数

$$|\alpha| = \sqrt{(\alpha,\alpha)} \tag{11.1.3}$$

称为复线性空间 $\mathfrak{L}$ 中向量 α 的**长度**. 长度为 1 的向量称为**单位向量**.

由定义可知, 对复线性空间 $\mathfrak{L}$ 中任一非零向量 α, 则

$$\beta = \frac{1}{|\alpha|}\alpha \tag{11.1.4}$$

是单位向量.

在定义夹角之前, 先证明一个有用的不等式.

引理 11.1.6 (Cauchy 不等式) 在 n 维复线性空间 $\mathfrak{L}$ 上取定内积 (α,β), 则有

$$|(\alpha,\beta)|^2 \leqslant (\alpha,\alpha)(\beta,\beta), \quad \forall \alpha,\beta \in \mathfrak{L}, \tag{11.1.5}$$

且等式成立当且仅当 α 和 β 线性相关.

证 引进复参数 t, 则有 $\alpha+t\beta\in\mathfrak{L}$. 于是

$$(\alpha+t\beta,\alpha+t\beta)=(\alpha,\alpha)+t\overline{(\alpha,\beta)}+\bar{t}(\alpha,\beta)+|t|^2(\beta,\beta)\geqslant 0,\qquad \forall t\in\mathbb{C}.$$

当 $\beta=0$ 时, 不等式 $|(\alpha,\beta)|^2\leqslant(\alpha,\alpha)(\beta,\beta)$ 自然成立. 当 $\beta\neq 0$ 时有 $(\beta,\beta)>0$. 记 $t=x+\sqrt{-1}y$, $x,y\in\mathbb{R}$, 则有

$$(\alpha,\alpha)+2x\mathrm{Re}\,(\alpha,\beta)+2y\mathrm{Im}\,(\alpha,\beta)+(x^2+y^2)(\beta,\beta)\geqslant 0,$$

此即

$$(\beta,\beta)[x+\frac{\mathrm{Re}\,(\alpha,\beta)}{(\beta,\beta)}]^2+(\beta,\beta)[y+\frac{\mathrm{Im}\,(\alpha,\beta)}{(\beta,\beta)}]^2+(\alpha,\alpha)-\frac{|(\alpha,\beta)|^2}{(\beta,\beta)}\geqslant 0.$$

我们知道, 复自变量 t 取一切复数值时, 即实自变量 x 和 y 取一切实数值时, 上式成立的充分且必要条件是

$$(\alpha,\alpha)(\beta,\beta)\geqslant|(\alpha,\beta)|^2,$$

此即 Cauchy 不等式成立.

又等号成立的必要且充分条件是存在复数 $t_0=\dfrac{-(\alpha,\beta)}{(\beta,\beta)}$, 使得 $(\alpha+t_0\beta,\alpha+t_0\beta)=0$. 由于内积是正定 Hermite 双线性函数, 所以有 $\alpha+t_0\beta=0$, 此即 α 和 β 线性相关. □

引理 11.1.7 (Cauchy 不等式) 任取两组复数 $a_1,a_2,\cdots,a_n$ 和 $b_1,b_2,\cdots,b_n$, 则有

$$|\sum_{j=1}^{n}a_j\overline{b_j}|^2\leqslant\Big(\sum_{j=1}^{n}|a_j|^2\Big)\Big(\sum_{j=1}^{n}|b_j|^2\Big).\tag{11.1.6}$$

上式中等号成立的必要且充分条件是复数 $a_1,a_2,\cdots,a_n$ 和 $b_1,b_2,\cdots,b_n$ 成比例.

利用 Cauchy 不等式, 还可以引进夹角的概念. 事实上, 由 Cauchy 不等式 (11.1.5) 可知, 在 α 和 β 都不是零向量时有不等式

$$0\leqslant\frac{|(\alpha,\beta)|}{|\alpha||\beta|}\leqslant 1,\quad \forall\alpha,\beta\in\mathfrak{L}.\tag{11.1.7}$$

所以有

定义 11.1.8 在 n 维复线性空间 $\mathfrak{L}$ 上取定内积 (α,β). 当 α 或 β 是零向量时, 我们约定向量 α 和 β 的**夹角** 任意. 当 $\alpha\neq 0,\beta\neq 0$ 时, 数值 θ 记作 $\theta=<\alpha,\beta>$, 它定义为

$$\theta=<\alpha,\beta>=\cos^{-1}\frac{|(\alpha,\beta)|}{|\alpha||\beta|},\tag{11.1.8}$$

称为向量 α 和 β 的**夹角**. 这里约定 $0 \leqslant \theta \leqslant \frac{\pi}{2}$. 当夹角是直角, 即当 $(\alpha, \beta) = 0$ 时, 我们称向量 α 和 β **互相正交**, 记作 $\alpha \perp \beta$. 显然, 只有零向量才和任何向量正交. 夹角 θ 也称为向量 α 和 β 的**角度**.

由定义可知, 在复的情形, 角度的定义和实的情形是不一同的. 另一方面, 恒有关系

$$(\alpha, \beta) = \sqrt{(\alpha, \alpha)}\sqrt{(\beta, \beta)} \cos < \alpha, \beta >= |\alpha||\beta| \cos < \alpha, \beta > . \tag{11.1.9}$$

定义 11.1.9 设 $\mathfrak{L}$ 是 n 维复 Euclid 空间, 内积是 (α, β). 复 Euclid 空间 $\mathfrak{L}$ 中的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 称为**标准正交基**, 如果

$$(\alpha_i, \alpha_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n, \tag{11.1.10}$$

即 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 个两两正交的单位向量.

利用单位向量和正交的概念, 不难证明: 在定义内积 (11.1.1) 时所选取的基 α_1, $\alpha_2, \cdots, \alpha_n$ 是由一组两两正交的单位向量构成, 即 $(\alpha_j, \alpha_k) = \delta_{jk}$, j, $k = 1, 2, \cdots, n$. 一般有

引理 11.1.10 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任意 s 个两两正交的向量 $\beta_1, \beta_2, \cdots$, β_s 必线性无关.

定义 11.1.11 由所有 $n \times 1$ 复矩阵构成的 n 维复线性空间 $\mathbb{C}^n$ 称为**复向量空间**. 在 $\mathbb{C}^n$ 中取定一组基

$$e_j = (\delta_{j1}, \delta_{j2}, \cdots .\delta_{jn})', \quad j = 1, 2, \cdots, n, \tag{11.1.11}$$

其中 δ_{jk} 是 Kronecker 符号, 即 $\delta_{jj} = 0$, $\delta_{jk} = 0$, $j \neq k$. 所以可以引进标准内积 (α, β), 它定义为:

$$(e_i, e_j) = \delta_{ij}, \quad 1 \leqslant i, j \leqslant n. \tag{11.1.12}$$

换句话说, $e_1, e_2, \cdots, e_n$ 都是单位向量, 而且这 n 个向量两两正交. 于是 $\mathbb{C}^n$ 是 n 维复 Euclid 空间.

注意 今后凡是提到 $\mathbb{C}^n$, 作为 n 维复 Euclid 空间, 我们理解为标准内积 (α, β) 由定义 11.1.11 给出, 而不再加以说明了. 由定义 11.1.9 还可知, 在 n 维复 Euclid 空间 $\mathbb{C}^n$ 中取定的基 (11.1.11): $e_1, e_2, \cdots, e_n$ 是一组标准正交基, 它是直角坐标系的推广. 当然, 并不是 $\mathbb{C}^n$ 中所有的基都是标准正交基. 由定义 11.1.9 又可知, 标准正交基的特点为:

定理 11.1.12 设 $\mathfrak{L}$ 是关于内积 (α, β) 的 n 维复 Euclid 空间. $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是复 Euclid 空间 $\mathfrak{L}$ 的标准正交基. 记 $\alpha = \sum\limits_{j=1}^{n} x_j \alpha_j$, $\beta = \sum\limits_{j=1}^{n} y_j \alpha_j$, 则有

$$(\alpha, \beta) = \sum_{j=1}^{n} x_j \overline{y_j}. \tag{11.1.13}$$

即在标准正交基下, 内积 (α,β) 的方阵表示是单位方阵 E_n.

定义 11.1.13 n 维复 Euclid 空间 $\mathfrak{L}$ 中保持向量长度不变的线性变换称为**酉变换**.

引理 11.1.14 n 维复 Euclid 空间 $\mathfrak{L}$ 中线性变换 $\mathscr{A}$ 是酉变换当且仅当它保持向量内积不变.

定理 11.1.15 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 线性变换 $\mathscr{A}$ 在这组基下的方阵表示记为 A, 则线性变换 $\mathscr{A}$ 是酉变换的必要且充分条件是

$$A^*A = AA^* = E_n, \tag{11.1.14}$$

其中 E_n 是 n 阶单位方阵.

定理 11.1.16 n 维复 Euclid 空间 $\mathfrak{L}$ 的线性变换 $\mathscr{A}$ 是酉变换的必要且充分条件是它将标准正交基变为标准正交基.

定理 11.1.17 设 $\mathfrak{L}$ 是 n 维复 Euclid 空间, 内积是 (α,β). 任取标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 再在 $\mathfrak{L}$ 中任取一组基 $\beta_1, \beta_2, \cdots, \beta_n$, 使得基变换公式为

$$\beta_i = \sum_{j=1}^{n} a_{ji}\alpha_j, \quad i = 1, 2, \cdots, n, \tag{11.1.15}$$

它对应的 n 阶非异复方阵为

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{11.1.16}$$

则 $\beta_1, \beta_2, \cdots, \beta_n$ 是标准正交基当且仅当 n 阶非异复方阵 A 适合条件 (11.1.14).

上面初步讨论了酉变换. 为了进一步考虑酉变换的性质, 需要对酉变换在标准正交基下的方阵表示进行详细讨论.

定义 11.1.18 n 阶复方阵 U 称为**酉方阵**, 如果

$$U^*U = UU^* = E_n. \tag{11.1.17}$$

所有 n 阶酉方阵构成的集合记作 $\mathrm{U}(n)$, 称为 **n 阶酉群**.

由定义可知, 定理 11.1.17 证明了: 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基后, 酉方阵和 $\mathfrak{L}$ 中标准正交基间便有一个自然的一一对应关系.

另一方面, 实的酉方阵就是实正交方阵, 又酉方阵有性质:

引理 11.1.19 n 阶酉群 $\mathrm{U}(n)$ 有性质

(1) 任取 $U, V \in \mathrm{U}(n)$, 则 $UV \in \mathrm{U}(n)$;

(2) 酉方阵的乘法适合结合律;

(3) n 阶单位方阵 $E_n \in \mathrm{U}(n)$;

(4) 任取 $U \in \mathrm{U}(n)$, 则 $\overline{U}, U' \in \mathrm{U}(n)$, 又 $U^{-1} = U^* \in \mathrm{U}(n)$;

(5) 酉方阵的行列式等于模为 1 的复数, 即 $\det(U) = e^{i\theta}$, $\forall\, \theta \in \mathbb{R}$.

在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任取一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 对复 Euclid 空间 $\mathfrak{L}$ 中任两向量 α 和 β, 设它们在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标分别为 $(a_1, a_2, \cdots, a_n)' \in \mathbb{C}^n$ 和 $(b_1, b_2, \cdots, b_n)' \in \mathbb{C}^n$, 则有

$$(\alpha, \beta) = (\sum_{j=1}^{n} a_j \alpha_j, \sum_{k=1}^{n} b_k \alpha_k) = \sum_{j,k=1}^{n} a_j \overline{b_k} (\alpha_j, \alpha_k) = \sum_{j,k=1}^{n} a_j \overline{b_k} \delta_{jk} = \sum_{j=1}^{n} a_j \overline{b_j}.$$

所以, 如果按照标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 来定义向量 α 和 β 的内积, 那末它们的内积就是向量 α 和 β 的坐标按照复 Euclid 空间 $\mathbb{C}^n$ 的标准正交基 $e_1, e_2, \cdots, e_n$ 所定义的标准内积. 换句话说, 我们证明了内积不依赖于标准正交基的选取; 特别, 向量的长度不依赖于标准正交基的选取.

将 n 阶复方阵 B 按照列分成 n 块:

$$B = (\beta_1, \beta_2, \cdots, \beta_n), \tag{11.1.18}$$

其中 $\beta_1, \beta_2, \cdots, \beta_n$ 都是 $n \times 1$ 复矩阵. 所以 $\beta_1, \beta_2, \cdots, \beta_n$ 是 n 维复 Euclid 空间 $\mathbb{C}^n$ 中 n 个向量. 再由矩阵分块乘法的规则可知

$$B^*B = \begin{pmatrix} \beta_1^* \\ \beta_2^* \\ \vdots \\ \beta_n^* \end{pmatrix} \begin{pmatrix} \beta_1 \ \beta_2 \cdots \ \beta_n \end{pmatrix} = \begin{pmatrix} \beta_1^*\beta_1 & \cdots & \beta_1^*\beta_n \\ \vdots & & \vdots \\ \beta_n^*\beta_1 & \cdots & \beta_n^*\beta_n \end{pmatrix}.$$

所以 B 是酉方阵的必要且充分条件为

$$\beta_i^*\beta_j = \delta_{ij} = (\beta_i, \beta_j), \quad 1 \leqslant i, j \leqslant n. \tag{11.1.19}$$

即 n 阶复方阵 B 的列向量构成 $\mathbb{C}^n$ 的一组标准正交基. 所以有

定理 11.1.20 n 阶复方阵是酉方阵的必要且充分条件为: 它的 n 个列向量构成 $\mathbb{C}^n$ 的一组标准正交基.

定理 11.1.21 n 阶复方阵是酉方阵的必要且充分条件为它的 n 个行向量构成一组标准正交基.

利用酉方阵的基本性质, 可以证明酉变换有性质:

(1) 酉变换的乘积仍为酉变换;

(2) 酉变换的乘法适合结合律;

(3) 恒等变换是酉变换;

(4) 酉变换的逆变换 (等于它的共轭变换) 也是酉变换;

(5) 酉变换的行列式等于模为 1 的复数.

定义 11.1.22 设 $\mathfrak{L}$ 是 n 维复 Euclid 空间. $\mathfrak{L}$ 上所有酉变换构成的集合称为**酉变换群**. 研究在所有酉变换下都不变的几何性质的学科称为**复 Euclid 几何学**.

在复 Euclid 空间 $\mathfrak{L}$ 中, 最基本的定理是给出从任取一组基, 构作标准正交基的方法, 即所谓 Schmidt 正交化.

定理 11.1.23 (Schmidt 正交化) 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任取一组基 β_1, $\beta_2, \cdots, \beta_s$, 则存在一组标准正交基

$$\begin{cases} \alpha_1 = b_{11}\beta_1, \\ \alpha_2 = b_{21}\beta_1 + b_{22}\beta_2, \\ \qquad\vdots \\ \alpha_n = b_{n1}\beta_1 + b_{n2}\beta_2 + \cdots + b_{nn}\beta_n, \end{cases} \tag{11.1.20}$$

其中 $b_{11} > 0, b_{22} > 0, \cdots, b_{nn} > 0$.

Schmidt 正交化有下列重要结论.

定理 11.1.24 任一 n 阶非异复方阵 A 一定可以唯一地表成一个复上 (下) 三角方阵 T 和一个酉方阵 U 的乘积,

$$A = TU, \tag{11.1.21}$$

其中 T 的对角元素都是正实数; 也一定可以唯一地表成一个酉方阵 U_1 和一个复上 (下) 三角方阵 T_1 的乘积

$$A = U_1T_1, \tag{11.1.22}$$

其中 T_1 的对角元素也都是正实数.

定理 11.1.25 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任取 $m(\leqslant n)$ 个两两正交的单位向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$, 则存在 $n-m$ 个单位向量 $\alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n$, 使得 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 构成复 Euclid 空间 $\mathfrak{L}$ 的一组标准正交基.

这个定理的矩阵语言是:

定理 11.1.26 设 $m \leqslant n$. 任给一个 $n \times m$ 复矩阵

$$U_1 = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \tag{11.1.23}$$

它适合条件 $U_1^* U_1 = E_m$, 则必有一个 $n \times (n-m)$ 复矩阵

$$U_2 = \begin{pmatrix} a_{1,m+1} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n,m+1} & \cdots & a_{nn} \end{pmatrix}, \tag{11.1.24}$$

使得 n 阶复方阵

$$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{11.1.25}$$

是一个酉方阵.

上面两个定理在 $m=1$ 的情形是最常用的定理. 这时定理可以改写为

定理 11.1.27 在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任取单位向量 α, 则存在一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得 α 为指定的第 j 个基向量 α_j. 特别, 单位向量能成为一组标准正交基的第一个向量.

在 n 维复 Euclid 空间 $\mathfrak{L}$ 中任取子集合 $\mathfrak{S}_1$ 和 $\mathfrak{S}_2$, 则 $(\mathfrak{S}_1, \mathfrak{S}_2) = 0$ 表示 $(\alpha, \beta) = 0, \forall \alpha \in \mathfrak{S}_1, \beta \in \mathfrak{S}_2$.

定义 11.1.28 设 $\mathfrak{S}$ 是 n 维复 Euclid 空间 $\mathfrak{L}$ 中的子集. 记

$$\mathfrak{S}^\perp = \{ \alpha \in \mathfrak{L} \mid (\alpha, \mathfrak{S}) = 0 \}. \tag{11.1.26}$$

则 $\mathfrak{S}^\perp$ 称为子集 $\mathfrak{S}$ 的**正交补**.

定理 11.1.29 n 维复 Euclid 空间 $\mathfrak{L}$ 中子集 $\mathfrak{S}$ 的正交补 $\mathfrak{S}^\perp$ 为 $\mathfrak{L}$ 的子空间, 且

$$\mathfrak{S} \subset (\mathfrak{S}^\perp)^\perp. \tag{11.1.27}$$

又

$$\mathfrak{L} = \mathfrak{S}^\perp \oplus (\mathfrak{S}^\perp)^\perp \tag{11.1.28}$$

为空间直接和. 特别, 当 $\mathfrak{S}$ 本身为子空间时, 则有

$$(\mathfrak{S}^\perp)^\perp = \mathfrak{S}. \tag{11.1.29}$$

所以 $\mathfrak{L}$ 有子空间直接和分解

$$\mathfrak{L} = \mathfrak{S} \oplus \mathfrak{S}^\perp. \tag{11.1.30}$$

定义 11.1.30 对 n 维复 Euclid 空间 $\mathfrak{L}$ 中的子空间 $\mathfrak{S}$, 任取 $\alpha \in \mathfrak{L}$, 则唯一存在 $\beta \in \mathfrak{S}, \gamma \in \mathfrak{S}^\perp$, 使得 $\alpha = \beta + \gamma$. 于是 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的对应 $\alpha \to \beta$ 为 $\mathfrak{L}$ 上的线性变

换, 称为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的**正交投影**.

定理 11.1.31 设 $\mathscr{A}$ 是 n 维复 Euclid 空间 $\mathfrak{L}$ 上的线性变换. 记 $\mathscr{A}$ 的象空间 $\mathscr{A}(\mathfrak{L})=\mathfrak{S}$. 则 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{S}$ 上的正交投影当且仅当

$$(\mathscr{A}(\mathfrak{L}),(\mathrm{id}-\mathscr{A})(\mathfrak{L}))=0. \tag{11.1.31}$$

这时有 $\mathscr{A}^2=\mathscr{A}$.

例 11.1.1 设 A 是 $2n$ 阶实方阵. 记

$$J=\begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}. \tag{11.1.32}$$

如果实方阵 A 适合条件

$$AJA'=J, \tag{11.1.33}$$

试证: $\det(A)=1$.

证 下面的证明特点是用复方阵的技巧来处理实方阵.

证法一 令

$$A_0=AJ+JA=\begin{pmatrix} B & C \\ D & G \end{pmatrix}\begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}+\begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}\begin{pmatrix} B & C \\ D & G \end{pmatrix}.$$

即

$$A_0=\begin{pmatrix} D-C & B+G \\ -B-G & D-C \end{pmatrix}=\begin{pmatrix} P & Q \\ -Q & P \end{pmatrix}.$$

考虑 $2n$ 阶酉方阵

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix} E_n & \sqrt{-1}E_n \\ E_n & -\sqrt{-1}E_n \end{pmatrix},$$

则由 P 和 Q 为 n 阶实方阵, 又

$$UA_0U^*=\mathrm{diag}\,(P-\sqrt{-1}Q,P+\sqrt{-1}Q),$$

所以

$$\det(A_0)=\det(UA_0U^*)=|\det(P+\sqrt{-1}Q)|^2\geqslant 0.$$

而 $\det(J)=1$, $AA'>0$, 所以 $E_n+AA'>0$. 因此由 $A_0A'=(AJ+JA)A'=AJA'+JAA'=J(E_n+AA')$, 两边取行列式, 有

$$\det(A_0)\det(A)=\det(J)\det(E_n+AA')=\det(E_n+AA')>0.$$

所以 $\det(A_0) > 0$, 且 $\det(A) = (\det(A_0))^{-1}\det(E_n + AA') > 0$. 但是由 $AJA' = J$, 所以 $(\det(A))^2 = 1$. 这证明了 $\det(A) = 1$. □

证法二 令 U 为如上的 $2n$ 阶酉方阵, 而

$$T = UAU^* = \frac{1}{2}\begin{pmatrix} E_n & \sqrt{-1}E_n \\ E_n & -\sqrt{-1}E_n \end{pmatrix}\begin{pmatrix} B & C \\ D & G \end{pmatrix}\begin{pmatrix} E_n & E_n \\ -\sqrt{-1}E_n & \sqrt{-1}E_n \end{pmatrix},$$

所以

$$T = \frac{1}{2}\begin{pmatrix} B+G+\sqrt{-1}(D-C) & B-G+\sqrt{-1}(D+C) \\ B-G-\sqrt{-1}(D+C) & B+G-\sqrt{-1}(D-C) \end{pmatrix}.$$

由于 B, C, D, G 都是 n 阶实方阵, 所以有

$$T = UAU^* = \begin{pmatrix} P & Q \\ \overline{Q} & \overline{P} \end{pmatrix},$$

其中 $2P = B + G + \sqrt{-1}(D - C)$, $2Q = B - G + \sqrt{-1}(D + C)$. 由 $A = U^*TU$, $AJA' = J$, 所以有 $U^*TUJU^*T^*U = J$, 即有 $T(UJU^*)T^* = UJU^*$. 而 $UJU^* = -\sqrt{-1}\mathrm{diag}(E_n, -E_n)$, 所以有

$$\begin{pmatrix} P & Q \\ \overline{Q} & \overline{P} \end{pmatrix}\begin{pmatrix} E_n & 0 \\ 0 & -E_n \end{pmatrix}\begin{pmatrix} P^* & Q' \\ Q^* & P' \end{pmatrix} = \begin{pmatrix} E_n & 0 \\ 0 & -E_n \end{pmatrix}.$$

即 $PP^* - QQ^* = E_n$, $PQ' = QP'$. 于是 $PP^* = E_n + QQ^* > 0$. 即有 $\det(P) \neq 0$, 而

$$\det(A) = \det(T) = \det\begin{pmatrix} P & Q \\ \overline{Q} & \overline{P} \end{pmatrix} = \det\begin{pmatrix} E_n & 0 \\ -\overline{Q}P^{-1} & E_n \end{pmatrix}\begin{pmatrix} P & Q \\ \overline{Q} & \overline{P} \end{pmatrix}.$$

所以

$$\det(A) = \det\begin{pmatrix} P & Q \\ 0 & \overline{P} - \overline{Q}P^{-1}Q \end{pmatrix}.$$

由 $PQ' = QP'$, 有

$$\begin{aligned} \det(A) &= \det(P)\det(\overline{P} - \overline{Q}P^{-1}Q) = \det(\overline{P}P' - \overline{Q}P^{-1}QP') \\ &= \det(\overline{P}P' - \overline{Q}Q') = \det(E_n) = 1. \end{aligned}$$

□

下面一个例子也是一种标准技巧.

例 11.1.2 设 A 为 $2n$ 阶酉方阵, 且有 $AJA' = J$, 则 A 称为**辛方阵**. 所有 $2n$ 阶辛方阵构成集合 $\mathrm{Sp}(n)$, 称为**辛群**. 试证:

(1) 任取 A 和 $B \in \mathrm{Sp}(n)$, 则 $AB \in \mathrm{Sp}(n)$, $A^{-1} \in \mathrm{Sp}(n)$;

(2) 任取 $A\in \mathrm{Sp}\,(n)$, 则存在 $P\in \mathrm{Sp}\,(n)$, 使得

$$P^*AP=\mathrm{diag}\,(e^{\sqrt{-1}\theta_1},\cdots,e^{\sqrt{-1}\theta_n},e^{-\sqrt{-1}\theta_1},\cdots,e^{-\sqrt{-1}\theta_n}).$$

特别, $\det(A)=1$.

证 由定义, (1) 显然成立. 下面证 (2) 成立.

设 $A\in \mathrm{Sp}\,(n)$, 即有 $A^*A=E_{2n}$, $AJA'=J$. 设 a 为 A 的特征根, α 为 A 的属于特征根 a 的特征向量, 且 α 为单位向量, 即有 $A\alpha=a\alpha$, $|\alpha|=1$.

由 $AJA'=J$, 所以 $AJA'\overline{A}=J\overline{A}$, 即有 $AJ=J\overline{A}$. 记 $\beta=-J\overline{\alpha}$, 则有

$$A\beta=-AJ\overline{\alpha}=-J\overline{A\alpha}=-\overline{a}J\overline{\alpha}=\overline{a}\beta.$$

又

$$|\beta|=\sqrt{(\beta,\beta)}=\sqrt{\beta'\overline{\beta}}=\sqrt{\alpha^*J'J\alpha}=\sqrt{\alpha^*\alpha}=1.$$

所以 β 为属于特征根 $\overline{a}$ 的单位特征向量.

为方便起见, 仍用 J 及 A 表示 $\mathbb{C}^{2n}$ 上的线性变换, 使得它们在 $\mathbb{C}^{2n}$ 的标准正交基 $e_1,e_2,\cdots,e_{2n}$ 下对应的方阵表示分别为 J 及 A. 今 $\beta=-J\overline{\alpha}$, 则

$$(\alpha,\beta)=\alpha'\overline{\beta}=-\alpha'J\alpha=0.$$

所以 $\alpha\perp\beta$. 于是 α 和 β 为 $\mathbb{C}^{2n}$ 中二维子空间 $\mathfrak{R}$ 的标准正交基. 由于 $\mathbb{C}^{2n}=\mathfrak{R}\oplus\mathfrak{R}^{\perp}$, 下面证明 $\mathfrak{R}$ 及 $\mathfrak{R}^{\perp}$ 都在 $\mathbb{C}^{2n}$ 上的线性变换 A 作用下不变, 即有 $A(\mathfrak{R})\subset\mathfrak{R}$, $A(\mathfrak{R}^{\perp})\subset\mathfrak{R}^{\perp}$. 再证明 $J(\overline{\mathfrak{R}})\subset\mathfrak{R}$, $J(\overline{\mathfrak{R}^{\perp}})\subset\mathfrak{R}^{\perp}$. 事实上, 任取 $\mathfrak{R}$ 中向量 $\xi=c\alpha+d\beta$, 其中 c 和 d 为复数. 由 $J^2=-E_{2n}$, $JJ'=J'J=E_{2n}$, 于是有

$$A\xi=cA\alpha+dA\beta=ac\alpha+\overline{a}d\beta\in\mathfrak{R},\quad J\overline{\xi}=\overline{c}J\overline{\alpha}+\overline{d}\alpha=\overline{d}\alpha-\overline{c}\beta\in\mathfrak{R}.$$

这证明了 $A(\mathfrak{R})\subset\mathfrak{R}$, $J\overline{\mathfrak{R}}\subset\mathfrak{R}$.

再任取 $\eta\in\mathfrak{R}^{\perp}$, 于是 $(\alpha,\eta)=\eta'\overline{\alpha}=0$, $(\beta,\eta)=\eta'\overline{\beta}=-\eta'J\alpha=0$. 即有 $\eta^*\alpha=\eta'J\alpha=0$.

为了证 $A(\mathfrak{R}^{\perp})\subset\mathfrak{R}^{\perp}$, 只要证 $(A\eta)^*\alpha=0$, $(A\eta)'J\alpha=0$. 事实上, 由 $A^*=A^{-1}$ 及 $A^{-1}J=JA'$, 所以有

$$\overline{(A\eta)}'\alpha=\overline{\eta}'A^{-1}\alpha=a^{-1}\eta^*\alpha=0,$$

$$(A\eta)'J\alpha=\eta'A'J\alpha=\eta'\overline{(A^*J)}\alpha=\eta'JA^{-1}\alpha=a^{-1}\eta'J\alpha=0.$$

为了证 $J(\overline{\mathfrak{R}^{\perp}})\subset\mathfrak{R}^{\perp}$, 只要证 $(\overline{J\overline{\eta}})'\alpha=0$, $(J\overline{\eta})'J\alpha=0$. 事实上,

$$(\overline{J\overline{\eta}})'\alpha=\eta'J'\alpha=-\eta'J\alpha=0,\quad (J\overline{\eta})'J\alpha=\eta^*\alpha=0.$$

至此证明了断言.

现在对维数作归纳法, 因此证明了在 $\mathbb{C}^{2n}$ 中存在标准正交基 $\alpha_1, \cdots, \alpha_n$, $\beta_1 = -J\overline{\alpha_1}, \cdots, \beta_n = -J\overline{\alpha_n}$, 使得 $A\alpha_j = a_j\alpha_j$, $A\beta_j = \overline{a_j}\beta_j$, $1 \leqslant j \leqslant n$. 由于 A 为酉方阵, 特征根的模为 1, 所以 $a_j = e^{\sqrt{-1}\theta_j}$, $1 \leqslant j \leqslant n$.

作 $2n$ 阶酉方阵

$$P = (\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_n).$$

我们来证 $P \in \mathrm{Sp}\,(n)$. 事实上, 由 $J^2 = -E_{2n}$, 有

$$\begin{aligned}
PJP' &= (-\beta_1, \cdots, -\beta_n, \alpha_1, \cdots, \alpha_n)P' = (J^2\beta_1, \cdots, J^2\beta_n, -J^2\alpha_1, \cdots, -J^2\alpha_n)P' \\
&= J(J\beta_1, \cdots, J\beta_n, -J\alpha_1, \cdots, -J\alpha_n)P' = J(\overline{\alpha_1}, \cdots, \overline{\alpha_n}, \overline{\beta_1}, \cdots, \overline{\beta_n})P' \\
&= J\overline{P}P' = J.
\end{aligned}$$

此即 $P \in \mathrm{Sp}\,(n)$. 再

$$\begin{aligned}
P^*AP &= P^*(A\alpha_1, \cdots, A\alpha_n, A\beta_1, \cdots, A\beta_n) = P^*(a_1\alpha_1, \cdots, a_n\alpha_n, \overline{a_1}\beta_1, \cdots, \overline{a_n}\beta_n) \\
&= P^*P\mathrm{diag}\,(a_1, \cdots, a_n, \overline{a_1}, \cdots, \overline{a_n}) = \mathrm{diag}\,(a_1, \cdots, a_n, \overline{a_1}, \cdots, \overline{a_n}).
\end{aligned}$$

这证明了(2) 成立. 这时显然有 $\det(A) = 1$. □

习　　题　11.1

11.1.1 在习题 9.2 中, 哪些题目可推广到复的情形? 证明是否一样的?

11.1.2 试证: n 阶复方阵

$$\frac{1}{\sqrt{n}}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega & \cdots & \omega^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)^2} \end{pmatrix}$$

是酉方阵, 其中 $\omega = \exp\left(\dfrac{2\pi\sqrt{-1}}{n}\right)$.

11.1.3 试证: 任一二阶酉方阵 U 可以分解为

$$\begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{pmatrix}\begin{pmatrix} \cos(\tau) & \sin(\tau) \\ -\sin(\tau) & \cos(\tau) \end{pmatrix}\begin{pmatrix} e^{i\theta_3} & 0 \\ 0 & e^{i\theta_4} \end{pmatrix},$$

其中 θ_1, θ_2, θ_3, θ_4, τ 都是实数.

11.1.4 将复方阵 A 分为实部 $P = \mathrm{Re}\,(A) = \dfrac{1}{2}(A + \overline{A})$ 及虚部 $Q = \mathrm{Im}\,(A) = \dfrac{1}{2\sqrt{-1}}(A - \overline{A})$, 则 A 为酉方阵的必要且充分条件为 $P'Q$ 是实对称方阵, 又 $P'P + Q'Q = E_n$.

11.1.5 记 $\mathfrak{L}$ 为所有 $n \times m$ 复矩阵构成的 nm 维复线性空间. 任取 $n \times m$ 复矩阵 A 和 B. 试证: $(A, B) = \mathrm{tr}\,(AB^*)$ 为 $\mathfrak{L}$ 上的内积. 并求出一组标准正交基.

§11.2 复方阵在酉相似下的标准形

定义 11.2.1 两个 n 阶复方阵 A 和 B 称为**酉相似的**, 如果存在 n 阶酉方阵 U, 使得

$$B = U^{-1}AU = U^*AU. \tag{11.2.1}$$

不难验证, 酉相似是所有 n 阶复方阵构成的集合 $\mathrm{gl}(n,\mathbb{C})$ 中的等价关系. 这个等价关系下的标准形理论超出了本书的范围. 下面考虑酉相似在某些特殊复方阵构成的子集合中的标准形理论.

定义 11.2.2 适合条件 $AA^* = A^*A$ 的方阵称为**规范方阵**; 适合条件 $H^* = H$ 的方阵称为**Hermite 方阵**; 适合条件 $K^* = -K$ 的方阵称为**反 Hermite 方阵**. 记 Hermite 方阵

$$H = \begin{pmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & & \vdots \\ h_{n1} & \cdots & h_{nn} \end{pmatrix}, \tag{11.2.2}$$

则 H 是 Hermite 方阵的条件可以改写为

$$h_{ij} = \overline{h_{ji}}, \quad i,j = 1,2,\cdots,n. \tag{11.2.3}$$

特别, 对角元素都是实数. 记反 Hermite 方阵

$$K = \begin{pmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & & \vdots \\ k_{n1} & \cdots & k_{nn} \end{pmatrix}, \tag{11.2.4}$$

则 K 是反 Hermite 方阵的条件可以改写为

$$k_{ij} = -\overline{k_{ji}}, \quad i,j = 1,2,\cdots,n. \tag{11.2.5}$$

特别, 非零对角元素都是纯虚数.

今若 K 是反 Hermite 方阵, 则令 $H = \sqrt{-1}K$, 于是有

$$H^* = \overline{\sqrt{-1}K}' = -\sqrt{-1}K^* = -\sqrt{-1}(-K) = \sqrt{-1}K = H, \tag{11.2.6}$$

所以 $\sqrt{-1}K$ 是 Hermite 方阵. 反之, 若 H 是 Hermite 方阵. 令 $K = \sqrt{-1}H$, 于是有

$$K^* = \overline{\sqrt{-1}H}' = (-\sqrt{-1})H^* = -\sqrt{-1}H = -K, \tag{11.2.7}$$

即 $\sqrt{-1}H$ 为反 Hermite 方阵. 由此可见, Hermite 和反 Hermite 仅相差一个常数因子 $\sqrt{-1}$.

当 Hermite 方阵 H 的元素都是实数时, 它就是实对称方阵. 当反 Hermite 方阵 K 的元素都是实数时, 它就是实反对称方阵. 在 §9.3 中已经看出, 实对称和实反对称方阵具有本质上的差别, 但是按照上面方式一扩充到 "复" 的情况, 便没有本质差别了. 一般来说, 在这种扩充方式下, 很多问题在复数情形都变得简单些.

和实的情形一样, 任给 n 阶复方阵 A, 则

$$H = \frac{1}{2}(A + A^*), \qquad K = \frac{1}{2}(A - A^*) \tag{11.2.8}$$

分别为 Hermite 方阵及反 Hermite 方阵, 且

$$A = H + K; \tag{11.2.9}$$

再者, 不难证明分解是唯一的. 我们称 H 为 A 的 **Hermite 部分**, K 为 A 的**反 Hermite 部分**.

由定义可知, 酉方阵, Hermite 方阵和反 Hermite 方阵都是规范方阵. 又设 A, U_1, H, K 分别是规范方阵, 酉方阵, Hermite 方阵, 反 Hermite 方阵, 对任一酉方阵 U, 则 $U^{-1}AU$, $U^{-1}U_1U$, $U^{-1}HU$, $U^{-1}KU$ 仍分别为规范方阵, 酉方阵, Hermite 方阵, 反 Hermite 方阵. 在讨论这四种方阵在酉相似下的标准形以前, 先给出下面的定理.

定理 11.2.3 任一 n 阶复方阵 A 必酉相似于下 (上) 三角方阵,

$$B = \begin{pmatrix} a_{11} & & \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \tag{11.2.10}$$

其中 $a_{11}, \cdots, a_{nn}$ 是复方阵 A 的所有特征根.

证 和定理 9.2.3 的第一部分证明一样. 因为现在不需要区分实特征根和复特征根的情形来考虑, 所以证明简单多了. □

这个定理虽然未能解决复方阵在酉相似下的标准形理论. 不过它说明了: 将复方阵在酉相似下分成等价类后, 可以在每个等价类中取上列形式的下三角方阵, 这总比任取一个代表元素来得简单. 另一方面, 由于酉相似必须相似, 所以同一等价类中任两方阵的特征根相同. 换句话说, 复方阵的特征根是它在酉相似下的部分不变量.

另一方面, 由式 (11.2.9), n 阶复方阵 A 可以唯一地分解为 n 阶 Hermite 方阵 H 和 n 阶反 Hermite 方阵 $K = \sqrt{-1}H_1$ 的和, 其中 H_1 是 Hermite 方阵. 所以这也说明了一般的复方阵在酉相似下分类理论的困难性. 事实上, 任取酉方阵 U, 则 $U^*AU = U^*HU + \sqrt{-1}U^*H_1U$, 其中 U^*HU 和 U^*H_1U 仍为 n 阶 Hermite 方阵. 所以问题化为相当于 H 和 H_1 同时化为标准形. 记**Hermite 方阵偶** (H, H_1) 全体构成的集合为 $\mathfrak{S}$, 我们称 Hermite 方阵偶 (H, H_1) 和 (H_0, H_0) 酉相似. 不难证明 Hermite

方阵偶的酉相似是集合 $\mathfrak{S}$ 的一种等价关系. 上面的讨论告诉我们, 一般复方阵在酉相似下的标准形理论就是 Hermite 方阵偶 (H, H_1) 在酉相似下的标准形理论 (其中 H 和 H_1 为 n 阶 Hermite 方阵).

下面进一步考虑一些特殊的方阵在酉相似下的标准形, 给出规范方阵, 酉方阵, Hermite 方阵及反 Hermite 方阵在酉相似下的标准形, 并且给出标准形的几何解释.

定理 11.2.4 (Schur 不等式) 记 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 n 阶复方阵 A 的 n 个特征根, 则有不等式

$$\sum_{j=1}^{n} |\lambda_j|^2 \leqslant \operatorname{tr}(A^* A), \tag{11.2.11}$$

且等号成立当且仅当实方阵 A 是规范方阵.

上面定理的证明, 实际上也给出了

定理 11.2.5 n 阶规范方阵 A 酉相似于对角形

$$B = \operatorname{diag}(\lambda_1, \cdots, \lambda_n), \tag{11.2.12}$$

其中 $\lambda_1, \cdots, \lambda_n$ 为 A 的所有特征根. 所以规范方阵 A 在酉相似下的全系不变量是 A 的所有特征根.

证 和定理 9.3.5 的第一部分证明一样, 因为现在不需要区分实特征根和复特征根的情形来考虑, 所以证明简单多了. □

这个定理告诉我们, 在所有规范方阵构成的集合中考虑酉相似, 则在每个等价类中可以取定理 11.2.5 所给出的标准形为代表元素. 而且, 两个具有相同特征根的规范方阵必互相等价, 所以规范方阵的 n 个特征根是它在酉相似下的全系不变量. 换句话说, 只要知道规范方阵的 n 个特征根, 不需要寻找酉方阵, 就能立刻写出它们的标准形来. 这里要注意的是, 标准形仅由 n 个特征根确定, 而表达方式不是唯一的.

今考虑第一类初等方阵 P_{ij} (见 §3.4). 显然, 它是酉方阵, 且对任一复方阵 A, 则 $P'_{ij} A P_{ij}$ 是将复方阵 A 的第 i 行及第 j 行互换, 同时将它的第 i 列及第 j 列互换. 特别, 这样引起了第 i 个对角元素和第 j 个对角元素互换, 而非对角元素必和非对角元素互换. 有限个第一类初等方阵的乘积称为**置换方阵**. 显然, 置换方阵仍为酉方阵. 设 P 为置换方阵, 则 $P^* A P$ 是将 A 的行和列按同样方式作了调换, 因此将对角元素作了排列. 正因为如此, 随便给出标准形 B, 那末存在置换方阵 P, 使得标准形 B 的对角块可以任意调换次序.

正因为定理 11.2.5 比定理 9.3.5 简单, 所以寻找规范方阵、酉方阵、Hermite 方阵、反 Hermite 方阵等在酉相似下的标准形也就简单了. 因为证明和完全相同, 所以

只写下结果.

定理 11.2.6　n 阶酉方阵 A 酉相似于对角形

$$B = \operatorname{diag}(\exp(\sqrt{-1}\theta_1), \exp(\sqrt{-1}\theta_2), \cdots, \exp(\sqrt{-1}\theta_n)), \tag{11.2.13}$$

其中 $0 \leqslant \theta_1 \leqslant \cdots \leqslant \theta_n < 2\pi$. 所以酉方阵的特征根的模为 1, 且酉方阵 U 在酉相似下的全系不变量为 U 的所有特征根.

定理 11.2.7　n 阶 Hermite 方阵 H 酉相似于对角形

$$\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n), \tag{11.2.14}$$

其中 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 且 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 H 的特征根. 所以 Hermite 方阵的特征根必为实数, 且 Hermite 方阵在酉相似下的全系不变量为它的所有特征根.

定理 11.2.8　n 阶反 Hermite 方阵酉相似于对角形:

$$\operatorname{diag}(\sqrt{-1}\lambda_1, \sqrt{-1}\lambda_2, \cdots, \sqrt{-1}\lambda_n), \tag{11.2.15}$$

其中 $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$, 且 $\sqrt{-1}\lambda_1$, $\sqrt{-1}\lambda_2$, $\cdots$, $\sqrt{-1}\lambda_n$ 为 K 的特征根. 所以反 Hermite 方阵的非零特征根必为纯虚数. 因此反 Hermite 方阵在酉相似下的全系不变量为它的所有特征根.

引理 11.2.9　设 $\mathfrak{L}$ 为 n 维复 Euclid 空间. $\mathscr{A}$ 为 $\mathfrak{L}$ 的线性变换, 则它在 $\mathfrak{L}$ 的不同标准正交基下对应的方阵表示互相酉相似.

在上面给出了一些标准形. 整个讨论是利用矩阵方法进行的. 它们有下面的几何意义: 在复 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 在 $\mathfrak{L}$ 中取定线性变换 $\mathscr{A}$, 记 $\mathscr{A}$ 在这组基下的方阵表示为 A. 现在在 $\mathfrak{L}$ 中另外取一组标准正交基 β_1, $\beta_2, \cdots, \beta_n$. 记旧基到新基的基变换的矩阵表示为 $\overline{U}$, 则 U 为酉方阵. 而线性变换 $\mathscr{A}$ 在新的基下的方阵表示为 $B = U^*AU$, 这就给出酉相似一种几何解释. 由此可见, 复方阵在酉相似下分成等价类后, 同一个等价类中的复方阵是复 Euclid 空间 $\mathfrak{L}$ 中一个确定的线性变换在所有标准正交基下的方阵表示. 不同等价类表示了不同的线性变换.

下面定义共轭变换. 为此, 先给出下面重要定理.

定理 11.2.10 (Riesz 表示定理)　记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 上的内积. 任取线性函数 f, 则在复 Euclid 空间 $\mathfrak{L}$ 中唯一存在向量 β_f, 使得

$$f(\alpha) = (\alpha, \beta_f), \quad \forall\, \alpha \in \mathfrak{L}, \tag{11.2.16}$$

且对应 $f \to \overline{\beta_f}$ 给出了对偶空间 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构.

证　令 (α, β) 是线性空间 $\mathfrak{L}$ 上的正定 Hermite 双线性函数. 所以 (α, ξ) 是线性空间 $\mathfrak{L}$ 上的线性函数. 反之, 在线性空间 $\mathfrak{L}$ 中任取一组标准正交基 $\alpha_1, \alpha_2, \cdots,$

α_n. 设线性空间 $\mathfrak{L}$ 中向量 α 和 β 在这组基下的坐标分别是 $(a_1, a_2, \cdots, a_n)'$ 和 $(b_1, b_2, \cdots, b_n)'$, 则内积 $(\alpha, \beta) = \sum\limits_{j=1}^{n} a_i \overline{b_i}$. 任取线性函数 f, 则 $f(\alpha) = \sum\limits_{j=1}^{n} a_j f(\alpha_j)$. 所以取 $\beta_f = \sum\limits_{j=1}^{n} \overline{f(\alpha_j)} \alpha_j \in \mathfrak{L}$, 则

$$f(\alpha) = \sum_{j=1}^{n} a_j f(\alpha_j) = (\alpha, \beta_f), \quad \forall\, \alpha \in \mathfrak{L}.$$

唯一性证明和定理 9.2.10 相同.

下面证对应 σ: $f \to \overline{\beta_f}$ 给出对偶空间 $\mathfrak{L}^*$ 到 $\mathfrak{L}$ 上的线性同构. 事实上, $\beta_f = \overline{\sigma(f)}$. 任取 $f, g \in \mathfrak{L}^*$, $a, b \in \mathbb{C}^n$, 则

$$\begin{aligned}\sigma(af + bg) = \overline{\beta_{af+bg}} &= \sum_{j=1}^{n} (af + bg)(\alpha_j)\overline{\alpha_j} \\ &= a\sum_{j=1}^{n} f(\alpha_j)\overline{\alpha_j} + b\sum_{j=1}^{n} g(\alpha_j)\overline{\alpha_j} = a\sigma(f) + b\sigma(g).\end{aligned}$$

这证明了 σ 是到上的线性同构. □

注意 在复 Euclid 空间 $\mathfrak{L}$ 中唯一存在向量 α_f, 使得 $f(\beta) = (\alpha_f, \beta)$, $\forall\, \beta \in \mathfrak{L}$. 但是, (α_f, β) 关于 β 是共轭线性的, 即 f 不是线性函数.

记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 上的内积. 设 $\mathscr{A}$ 为复 Euclid 空间 $\mathfrak{L}$ 上线性变换. 任给 $\beta \in \mathfrak{L}$, 易知 $f(\alpha) = (\mathscr{A}(\alpha), \beta)$, $\forall\, \alpha \in \mathfrak{L}$ 为复 Euclid 空间 $\mathfrak{L}$ 上的线性函数. 由上面定理可知 $f(\alpha)$ 唯一决定了 $\mathfrak{L}$ 中一个向量 β_f, 使得

$$(\mathscr{A}(\alpha), \beta) = (\alpha, \beta_f) = (\alpha, \overline{\sigma(f)}). \tag{11.2.17}$$

定理 11.2.11 符号同上. 则 $\mathscr{A}^*$: $\beta \to \overline{\beta_f}$ 为复 Euclid 空间 $\mathfrak{L}$ 上的线性变换, 记作 $\mathscr{A}^*$. 即有

$$(\mathscr{A}(\alpha), \xi) = f(\alpha) = (\alpha, \overline{\beta_f}) = (\alpha, \mathscr{A}^*(\xi)), \quad \forall\, \alpha, \xi \in \mathfrak{L}. \tag{11.2.18}$$

线性变换 $\mathscr{A}^*$ 称为线性变换 $\mathscr{A}$ 的**共轭变换**.

证 今任取 $\xi_1, \xi_2 \in \mathfrak{L}^*$, $c, d \in \mathbb{C}$. 由式 (11.2.18),

$$\begin{aligned}(\mathscr{A}(\alpha), \xi_1) = f(\alpha) = (\alpha, \overline{\beta_f}) = (\alpha, \mathscr{A}^*(\xi_1)), \\ (\mathscr{A}(\alpha), \xi_2) = g(\alpha) = (\alpha, \overline{\beta_g}) = (\alpha, \mathscr{A}^*(\xi_2))\end{aligned}$$

以及由定理 11.2.10,

$$\beta_{cf+dg} = \overline{\sigma(cf + dg)} = \overline{c\sigma(f) + d\sigma(g)} = \overline{c}\overline{\sigma(f)} + \overline{d}\overline{\sigma(g)} = \overline{c}\beta_f + \overline{d}\beta_g.$$

所以

$$(\alpha, \mathscr{A}^*(c\xi_1 + d\xi_2)) = (\alpha, \overline{\beta_{cf+dg}}) = (\alpha, \overline{\overline{c}\beta_f + \overline{d}\beta_g}) = \overline{c}(\alpha, \overline{\beta_f}) + \overline{d}(\alpha, \overline{\beta_g})$$
$$= \overline{c}(\alpha, \mathscr{A}^*(\xi_1)) + \overline{d}(\alpha, \mathscr{A}^*(\xi_2)) = (\alpha, c\mathscr{A}^*(\xi_1) + d\mathscr{A}^*(\xi_2)).$$

因此, $\mathscr{A}^*$ 是线性变换. □

在复 Euclid 空间 $\mathfrak{L}$ 中取定标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 记线性变换 $\mathscr{A}$ 对应的方阵表示为 A, 则共轭变换 $\mathscr{A}^*$ 对应的方阵表示为 A^*.

定义 11.2.12 记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称**为规范变换**, 如果有 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A}$, 即有

$$(\mathscr{A}(\alpha), \mathscr{A}(\beta)) = (\mathscr{A}^*(\alpha), \mathscr{A}^*(\beta)), \quad \forall \alpha, \beta \in \mathfrak{L}. \tag{11.2.19}$$

定义 11.2.13 记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称**为酉变换**, 如果有 $\mathscr{A}\mathscr{A}^* = \mathscr{A}^*\mathscr{A} = \mathrm{id}$, 即有

$$(\mathscr{A}(\alpha), \mathscr{A}(\beta)) = (\alpha, \beta), \quad \forall \alpha, \beta \in \mathfrak{L}. \tag{11.2.20}$$

定义 11.2.14 记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称**为Hermite 变换**, 如果有 $\mathscr{A}^* = \mathscr{A}$, 即有

$$(\mathscr{A}(\alpha), \beta) - (\alpha, \mathscr{A}(\beta)) = 0, \quad \forall \alpha, \beta \in \mathfrak{L}. \tag{11.2.21}$$

定义 11.2.15 记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 的内积. $\mathfrak{L}$ 上线性变换 $\mathscr{A}$ 称**为反 Hermite 变换**, 如果有 $\mathscr{A}^* = -\mathscr{A}$, 即有

$$(\mathscr{A}(\alpha), \beta) + (\alpha + \mathscr{A}(\beta)) = 0, \quad \forall \alpha, \beta \in \mathfrak{L}. \tag{11.2.22}$$

由上面四个定义, 所以定理 11.2.5—11.2.8 可改写为

定理 11.2.16 记 (α, β) 为 n 维复 Euclid 空间 $\mathfrak{L}$ 的内积, $\mathscr{A}$ 为 $\mathfrak{L}$ 上的线性变换. 则有

(1) $\mathscr{A}$ 为规范变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得 $\mathscr{A}(\alpha_k) = \lambda_k\alpha_k$, $1 \leqslant k \leqslant n$. 因此, α_k 线性生成了在 $\mathscr{A}$ 下不变的一维子空间 $\mathfrak{L}_k$, $1 \leqslant k \leqslant n$, 它们两两正交, 且 $\mathfrak{L}$ 是它们的不变子空间的直接和.

(2) $\mathscr{A}$ 为酉变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得

$$\mathscr{A}(\alpha_k) = (\exp\sqrt{-1}\theta_k)\alpha_k, \quad 1 \leqslant k \leqslant n,$$

其中 $0 \leqslant \theta_1 \leqslant \cdots \leqslant \theta_n < 2\pi$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维不变子空间的直接和.

(3) $\mathscr{A}$ 为 Hermite 变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得 $\mathscr{A}(\alpha_k) = \lambda_k\alpha_k$, $1 \leqslant k \leqslant n$, 其中 $\lambda_1 \leqslant \cdots \leqslant \lambda_n$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维不变子空间的直接和.

(4) $\mathscr{A}$ 为反 Hermite 变换当且仅当在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得 $\mathscr{A}(\alpha_k) = \sqrt{-1}\lambda_k\alpha_k$, $1 \leqslant k \leqslant n$. 其中 $\lambda_1 \leqslant \cdots \leqslant \lambda_n$. 因此, $\mathfrak{L}$ 可分解为两两正交的一维不变子空间的直接和.

习　题　11.2

11.2.1 在 §9.3 的习题中, 哪些题目可以推广到"复" 的情形, 如何推广? 证明是不是更简单些?

11.2.2 设 n 阶酉方阵 U 的特征根不等于 -1, 那末 n 阶方阵 $E_n + U$ 非异, 且

$$H = \sqrt{-1}(E_n - U)(E_n + U)^{-1} \tag{11.2.23}$$

是 Hermite 方阵. 反之, 若 H 是 Hermite 方阵, 则 n 阶方阵 $E_n - \sqrt{-1}H$ 非异, 且

$$U = (E_n + \sqrt{-1}H)(E_n - \sqrt{-1}H)^{-1} \tag{11.2.24}$$

是酉方阵.

因此, 上列关系建立了不以 -1 为特征根的酉方阵全体和所有 Hermite 方阵间的一个一一对应, 这个对应称为**Cayley 变换**. 它表明下述事实: 利用这个对应, 有时可以将 Hermite 方阵和酉方阵的性质互相转化.

11.2.3 如 $S = S'$, $T = -T'$ 都是实方阵, 且 $\det(E_n - T - \sqrt{-1}S) \neq 0$, 则 $(E_n + T + \sqrt{-1}S)(E_n - T - \sqrt{-1}S)^{-1}$ 是酉方阵.

11.2.4 对任一 $n \times 1$ 复矩阵 α, 试证: 存在一个 n 阶实正交方阵 O, 使得

$$O\alpha = (\exp(\sqrt{-1}\theta))(\lambda_1, \pm\sqrt{-1}\lambda_2, 0, \cdots, 0), \quad \lambda_1 \geqslant \lambda_2 \geqslant 0. \tag{11.2.25}$$

11.2.5 设 A 为 n 阶规范方阵. 它的 n 个特征根 $\lambda_j = a_j + \sqrt{-1}b_j$, $1 \leqslant j \leqslant n$, 其中 $a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_n$ 为 $2n$ 个不同实数. 试证: 若 α 为方阵 A, $H = \dfrac{1}{2}(A + A^*)$, $K = \dfrac{1}{2\sqrt{-1}}(A - A^*)$ 中某一个方阵的特征向量, 则必存在复数 $\lambda = \mu + \sqrt{-1}\nu$, 其中 μ 和 ν 都是实数, 使得

$$A\alpha = \lambda\alpha, \quad H\alpha = \mu\alpha, \quad K\alpha = \nu\alpha.$$

11.2.6 设 A 为 n 阶规范方阵, A 的模最大的特征根为 λ_1, 模最小的特征根为 λ_n. 设 U 为 n 阶酉方阵, 且 $\det(\lambda_0 A - U) = 0$. 试证: $|\lambda_n| \leqslant |\lambda_0|^{-1} \leqslant |\lambda_1|$.

11.2.7 记 A 为 n 阶复方阵, $H = \dfrac{1}{2}(A + A^*)$, $K = \dfrac{1}{2\sqrt{-1}}(A - A^*)$. 记 a, h, k 分别为 A, H, K 中元素的最大模. 记 $\lambda = b + \sqrt{-1}c$ 为 A 的特征根, 其中 b 和 c 为实数. 试证: $|\lambda| \leqslant na$, $|b| \leqslant nh$, $|c| \leqslant nk$. 以此来证明 Hermite 方阵、反 Hermite 方阵及酉方阵的特征根分别为实数、纯虚数及模为 1 的复数. (注意: b 不一定是 H 的特征根, c 不一定是 K 的特征根.)

11.2.8 记 A 为 n 阶复方阵, $H=\frac{1}{2}(A+A^*)$, $K=\frac{1}{2\sqrt{-1}}(A-A^*)$. 记 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的 n 个特征根. 试证:

$$\sum_{k=1}^{n}|\lambda_k|^2\leqslant \operatorname{tr}(AA^*),\quad \sum_{k=1}^{n}(\operatorname{Re}(\lambda_k))^2\leqslant \operatorname{tr}(HH^*),\quad \sum_{k=1}^{n}(\operatorname{Im}(\lambda_k))^2\leqslant \operatorname{tr}(KK^*),$$

且其中有一个等式成立的充分且必要条件为三个等式同时成立. 这时 A 是规范方阵.

§11.3 Hermite 方阵在复相合下的标准形

本节将把第十章的主要结果推广到 n 维复线性空间中去. 由于相当大一部分是作形式推广, 所以在下面全面叙述结果, 也着重讨论和 "实" 的情形不同的性质, 这些性质体现实和复的差别. 另一方面, 从下面讨论可以发现复的情形反而比实的情形简单多了.

定义 11.3.1 n 阶复方阵 A 和 B 称为**复相合的**, 如果存在 n 阶可逆复方阵 P, 使得

$$B=P^*AP. \tag{11.3.1}$$

引理 11.3.2 复相合为等价关系.

显然 Hermite 方阵和反 Hermite 方阵在复相合下仍分别变为 Hermite 方阵和反 Hermite 方阵. 下面讨论 Hermite 型的标准形. 由 §9.1, 这就是讨论 Hermite 方阵在复相合下的标准形. 将 §10.1 关于实对称方阵在相合下的标准形理论的全部讨论, 易实数为复数, 且凡取转置的地方都易为同时取转置及共轭, 则所有推理和结论全部成立, 即有

定理 11.3.3 任一 n 阶 Hermite 方阵 H 必复相合于对角形

$$\Lambda=\operatorname{diag}(E_p,-E_{r-p},0), \tag{11.3.2}$$

其中 r 为 Hermite 方阵 H 的秩, $\delta(H)=2s-r=\operatorname{tr}(\Lambda)$ 为 Hermite 方阵 H 的**符号差**, 它是 Hermite 方阵 H 的正特征根个数减负特征根的个数.

这个定理说明了: 在所有 n 阶 Hermite 方阵构成的集合中, 按照复相合分成等价类的并后, 在每个等价类中可以取式 (11.3.2) 为代表元素. 下面来求 n 阶 Hermite 方阵在复相合下的全系不变量. 为此, 先给出

定理 11.3.4 (Witt 引理) 设 H 是 n 阶非异 Hermite 方阵, H_1 和 H_2 是两个 m 阶非异 Hermite 方阵. 如果两个 $n+m$ 阶非异 Hermite 方阵

$$\begin{pmatrix} S & 0\\ 0 & S_1\end{pmatrix},\qquad \begin{pmatrix} S & 0\\ 0 & S_2\end{pmatrix} \tag{11.3.3}$$

复相合, 则 H_1 和 H_2 也复相合.

由 Witt 引理, 便有

定理 11.3.5 (惯性定理) 如果两个 n 阶 Hermite 方阵

$$\Lambda = \operatorname{diag}(E_p, -E_{r-p}, 0^{(n-r)}), \quad \Lambda_0 = \operatorname{diag}(E_q, -E_{s-q}, 0^{(n-s)}) \tag{11.3.4}$$

复相合, 则 $\Lambda_0 = \Lambda$, 即 $s = r$, $q = p$.

给定 Hermite 方阵 H, 记 H 在复相合下的标准形为

$$\operatorname{diag}(E_p, -E_{r-p}, 0),$$

其中 r 为 Hermite 方阵 H 的秩, 整数 $\delta(H) = 2p - r$ 为 Hermite 方阵 H 的符号差. Hermite 方阵的秩及符号差是它在复相合下的全系不变量.

定理 11.3.6 对 n 维复线性空间 $\mathfrak{L}$ 中任一 Hermite 型 $\varphi(x)$, 在 $\mathfrak{L}$ 中一定存在一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 使得在这组基下, Hermite 型 $\varphi(x)$ 的坐标表达式具有最简单的形式:

$$\varphi(x) = |x_1|^2 + \cdots + |x_p|^2 - |x_{p+1}|^2 + \cdots + |x_r|^2, \tag{11.3.5}$$

其中 r 及 $\delta(f) = 2p - r$ 分别称为 Hermite 型 φ 的**秩** 及**符号差**. 所以, 秩及符号差是 Hermite 型在基变换下的全系不变量.

关于配方法, 即矩阵打洞技巧, 可以如下写出: 设 H 为 n 阶 Hermite 方阵. 将它按前 s 行、前 s 列分块为

$$\begin{pmatrix} H_{11}^{(s)} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}. \tag{11.3.6}$$

假设 $\det(H_{11}) \neq 0$, 即 H_{11} 为 s 阶非异 Hermite 方阵, 则有

$$\begin{aligned} &\begin{pmatrix} E_s & -H_{11}^{-1}H_{12} \\ 0 & E_{n-s} \end{pmatrix}^* \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} \begin{pmatrix} E_s & -H_{11}^{-1}H_{12} \\ 0 & E_{n-s} \end{pmatrix} \\ &= \operatorname{diag}(H_{11}, H_{22} - H_{12}^* H_{11}^{-1} H_{12}). \end{aligned} \tag{11.3.7}$$

定义 11.3.7 n 维复线性空间 $\mathfrak{L}$ 上的 Hermite 双线性函数 $h(\alpha, \beta)$ 称为**非退化的**, 如果由 $h(\alpha, \alpha) = 0$ 可推出 $\alpha = 0$.

下面给出非退化 Hermite 双线性函数的标准形. 由 §9.1 可知: 在 n 维复线性空间 $\mathfrak{L}$ 中取定基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则 Hermite 双线性函数有坐标表达式

$$h(\alpha, \beta) = x^* A y, \tag{11.3.8}$$

其中

$$A=\begin{pmatrix} h(\alpha_1,\alpha_1) & \cdots & h(\alpha_1,\alpha_n) \\ \vdots & & \vdots \\ h(\alpha_n,\alpha_1) & \cdots & h(\alpha_n,\alpha_n) \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\quad y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},\tag{11.3.9}$$

又

$$\alpha=\sum_{j=1}^{n}\overline{x_j}\alpha_j,\qquad \beta=\sum_{j=1}^{n}\overline{y_j}\alpha_j.\tag{11.3.10}$$

而 A 为 n 阶复方阵. 于是记

$$H_1=\frac{1}{2}(A+A^*),\qquad H_2=\frac{1}{2\sqrt{-1}}(A-A^*),\tag{11.3.11}$$

则 H_1 和 H_2 为 n 阶 Hermite 方阵, 且

$$A=H_1+\sqrt{-1}H_2.\tag{11.3.12}$$

又在 n 维复线性空间 $\mathfrak{L}$ 中另取一组基 $\beta_1, \beta_2, \cdots, \beta_n$, 记基变换公式为

$$\beta_j=\sum_{k=1}^{n}\overline{p_{kj}}\alpha_k,\quad 1\leqslant j\leqslant n,\tag{11.3.13}$$

则 $P=(p_{jk})$ 为 n 阶非异复方阵. 记

$$B=\begin{pmatrix} h(\beta_1,\beta_1) & \cdots & h(\beta_1,\beta_n) \\ \vdots & & \vdots \\ h(\beta_n,\beta_1) & \cdots & h(\beta_n,\beta_n) \end{pmatrix},\quad u=\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},\quad v=\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},\tag{11.3.14}$$

其中

$$\alpha=\sum_{j=1}^{n}\overline{u_j}\beta_j,\qquad \beta=\sum_{j=1}^{n}\overline{v_j}\beta_j,\tag{11.3.15}$$

则有

$$h(\alpha,\beta)=u^*Bv.\tag{11.3.16}$$

于是 $B+B^*=P^*(A+A^*)P$, $B-B^*=P^*(A-A^*)P$, 因此记 n 阶 Hermite 方阵

$$\widetilde{H_1}=\frac{1}{2}(B+B^*),\qquad \widetilde{H_2}=\frac{1}{2\sqrt{-1}}(B-B^*),\tag{11.3.17}$$

则有

$$\widetilde{H_1}=P^*H_1P,\quad \widetilde{H_2}=P^*H_2P.\tag{11.3.18}$$

而

$$B=\widetilde{H_1}+\sqrt{-1}\widetilde{H_2}=P^*(H_1+\sqrt{-1}H_2)P=P^*AP.\tag{11.3.19}$$

由此可见，所谓 Hermite 双线性函数化标准形问题, 即为: 给定 n 阶**Hermite 方阵偶** (H_1, H_2), 寻找 n 阶非异复方阵 P, 使得 (P^*H_1P, P^*H_2P) 为标准形.

引理 11.3.8 在 n 维复线性空间 $\mathfrak{L}$ 中取定基, 则 Hermite 双线性函数 $h(\alpha,\beta)=x^*Ay$ 非退化当且仅当它决定的 n 阶 Hermite 方阵偶 (H_1,H_2) 有: 由 $x^*H_1x=x^*H_2x=0$ 可推出 $x=0$.

证 令

$$f(\alpha,\beta)=x^*H_1y+\sqrt{-1}x^*H_2y.$$

由定义, $h(\alpha,\beta)$ 非退化当且仅当由 $h(\alpha,\alpha)=0$ 可推出 $\alpha=0$, 即由

$$x^*H_1x+\sqrt{-1}x^*H_2x=0$$

可推出 $x=0$. 由于 x^*H_1x 和 x^*H_2x 都是实数, 此即由 $x^*H_1x=x^*H_2x=0$ 可推出 $x=0$. □

定理 11.3.9 设 H_1 和 H_2 为 n 阶 Hermite 方阵, 适合条件: 设 $x_0\in\mathbb{C}^n$, 且由 $x_0^*H_1x_0=x_0^*H_2x_0=0$ 可推出 $x_0=0$. 于是存在 n 阶非异复方阵 P, 使得

$$P^*H_1P=\operatorname{diag}(E_p,-E_{r-p},0_{n-r}),\quad P^*H_2P=\operatorname{diag}(\Lambda_1,\Lambda_2,\pm E_{n-r}),\tag{11.3.20}$$

其中 $\Lambda_1=\operatorname{diag}(\lambda_1,\cdots,\lambda_p)$ 为 p 阶实方阵, $\Lambda_2=\operatorname{diag}(\lambda_{p+1},\cdots,\lambda_r)$ 为 $r-p$ 阶实方阵.

证 由定理 11.3.3, 存在 n 阶非异复方阵 P_1, 使得

$$P_1^*H_1P_1=\operatorname{diag}(E_p,-E_{r-p},0_{n-r}).$$

记

$$P_1^*H_2P_1=\begin{pmatrix}A^{(r)} & B\\ B^* & D\end{pmatrix},\quad A^*=A,\quad D^*=D.$$

任取 n 阶非异复方阵 $P_2=\operatorname{diag}(E_r,P_3)$, 则有

$$P_2^*P_1^*H_1P_1P_2=P_1^*H_1P_1,\quad P_2^*P_1^*H_2P_1P_2=\begin{pmatrix}A & BP_3\\ P_3^*B^* & P_3^*DP_3\end{pmatrix}.$$

由定理 11.3.3, 存在 $n-r$ 阶非异复方阵 P_3, 使得 $P_3^*DP_3=\operatorname{diag}(E_q,-E_{s-q},0)$. 于是在复相合意义下, 无妨设

$$H_1=\operatorname{diag}(E_p,-E_{r-p},0_{n-r}),\quad H_2=\begin{pmatrix}A & B_1 & B_2 & B_3\\ B_1^* & E_q & 0 & 0\\ B_2^* & 0 & -E_{s-q} & 0\\ B_3^* & 0 & 0 & 0\end{pmatrix}.$$

设 $s<n-r$, 取 $x_0=e_n$, 则有 $x_0^*H_1x_0=x_0^*H_2x_0=0$. 这和条件 $x_0\neq 0$ 矛盾. 所以证明了 $s=n-r$.

设 $1\leqslant q\leqslant n-r$, 于是 $0<s-q<n-r$. 这时任取 $\alpha\in\mathbb{C}^q$, $\beta\in\mathbb{C}^{s-q}$, $x_0^*=(0,\alpha^*,\beta^*)$. 则有 $x_0^*H_1x_0=0$, $x_0^*H_2x_0=\alpha^*\alpha-\beta^*bz$. 取 $\alpha=e_1\in\mathbb{C}^q$, $\beta=e_1\in\mathbb{C}^{s-q}$, 则有 $x_0^*H_1x_0=x_0^*H_2x_0=0$. 这又和条件 $x_0=0$ 矛盾. 因此证明了 $q=0$ 或 $q=n-r$. 即

$$H_1=\operatorname{diag}(E_p,-E_{r-p},0_{n-r}),\quad H_2=\begin{pmatrix}A & B\\ B^* & \pm E_{n-r}\end{pmatrix}.$$

作复相合

$$\begin{pmatrix}E_r & 0\\ \pm\overline{B}' & E_{n-r}\end{pmatrix}^* H_1\begin{pmatrix}E_r & 0\\ \pm\overline{B}' & E_{n-r}\end{pmatrix}=H_1,$$

$$\begin{pmatrix}E_r & 0\\ \pm\overline{B}' & E_{n-r}\end{pmatrix}^* H_2\begin{pmatrix}E_r & 0\\ \pm\overline{B}' & E_{n-r}\end{pmatrix}=\operatorname{diag}(A_1^{(r)},\pm E_{n-r}).$$

这样一来, 问题化为将 $\operatorname{diag}(E_p,-E_{r-p})$ 和 A_1 在同一复相合下化为对角形. 换句话说, 问题化为 $r=n$ 的情形, 即在复相合意义下无妨设 n 阶 Hermite 方阵 H_1 非异.

设 H_1 和 H_2 为 n 阶 Hermite 方阵, 且 $\det(H_1)\neq 0$. 考虑多项式

$$\det(\lambda H_1-H_2)=\det(H_1)\det(\lambda E_n-H_1^{-1}H_2).$$

任取一根 λ_0, 则存在 $0\neq\alpha_n\in\mathbb{C}^n$, 使得 $\lambda_0H_1\alpha_n=H_2\alpha_n$. 今由 $\det(H_1)\neq 0$ 可知 $H_1\alpha_n=\beta_n\neq 0$. 且无妨取 α_n, 使得 β_n 为单位向量. 于是在 $\mathbb{C}^n$ 中存在标准正交基 $\beta_1,\beta_2,\cdots,\beta_n$. 由 $\beta_j^*\beta_n=\beta_j^*H_1\alpha_n=0$, $1\leqslant j\leqslant n-1$. 所以这时 $\beta_1,\beta_2,\cdots,\beta_{n-1}$, α_n 线性无关. 事实上, 若线性相关, 则有 $\alpha_n=\sum\limits_{k=1}^{n-1}a_k\beta_k$. 于是,

$$\alpha_n^*H_1\alpha_n=\sum_{k=1}^{n-1}\overline{a_k}\beta_k^*H_1\alpha_n=0.$$

因此, $\alpha_n^*H_2\alpha_n=\lambda_0\alpha_n^*H_1\alpha_n=0$. 由条件可知 $\alpha_n=0$, 这和 $\alpha_n\neq 0$ 矛盾. 所以 $P_0=(\beta_1\beta_2\cdots\beta_{n-1}\alpha_n)$ 为 n 阶非异复方阵. 而

$$P^*H_kP=\begin{pmatrix}\beta_1^*H_k\beta_1 & \cdots & \beta_1^*H_k\beta_{n-1} & \beta_1^*H_k\alpha_n\\ \vdots & & \vdots & \vdots\\ \beta_{n-1}^*H_k\beta_1 & \cdots & \beta_{n-1}^*H_k\beta_{n-1} & \beta_{n-1}^*H_k\alpha_n\\ \alpha_n^*H_k\beta_1 & \cdots & \alpha_n^*H_k\beta_{n-1} & \alpha_n^*H_k\alpha_n\end{pmatrix}.$$

由 $\beta_j^*H_k\alpha_n=0$, $1\leqslant j\leqslant n-1$, $k=1,2$, 所以证明了

$$P^*H_kP=\operatorname{diag}(\widetilde{H_k^{(n-1)}},\alpha_n^*H_k\alpha_n),\quad k=1,2.$$

令 $\det(H_1)\neq 0$, 所以 $\det(\widetilde{H_1})\neq 0$. 由归纳法, 便证明了存在 n 阶非异复方阵 P, 使得

$$P^*H_1P=\operatorname{diag}(\mu_1,\cdots,\mu_n),\quad P^*H_2P=\operatorname{diag}(\nu_1,\cdots,\nu_n).$$

由于 P^*H_kP 是 Hermite 方阵, $k=1,2$, 所以 $\mu_1,\ \mu_2,\ \cdots,\ \mu_n,\ \nu_1,\ \nu_2,\ \cdots,\ \nu_n$ 都是实数. 由 $\det(H_1)\neq 0$, 再用一个置换方阵作复相合, 从而无妨设 $\mu_1\geqslant\cdots\geqslant\mu_p>0>\mu_{p+1}\geqslant\cdots\geqslant\mu_n$. 再用

$$\operatorname{diag}\left(\frac{1}{\sqrt{\mu_1}},\cdots,\frac{1}{\sqrt{\mu_p}},\frac{1}{\sqrt{-\mu_{p+1}}},\cdots,\frac{1}{\sqrt{-\mu_n}}\right)$$

作复相合, 便证明了当 H_1 非异时, Hermite 方阵偶 (H_1,H_2) 在同一复相合下的标准形为式 (11.3.20). □

定理 11.3.10 设 $h(\alpha,\beta)$ 为 n 维复线性空间 $\mathfrak{L}$ 上非退化 Hermite 双线性函数, 则在 $\mathfrak{L}$ 中存在基 $\beta_1,\ \beta_2,\ \cdots,\ \beta_n$, 使得 $h(\alpha,\beta)=h(\sum\limits_{j=1}^{n}x_j\beta_j,\sum\limits_{k=1}^{n}y_k\beta_k)$ 有

$$h(\alpha,\beta)=\sum_{j=1}^{p}(1+\sqrt{-1}\lambda_j)x_j\overline{y_j}+\sum_{j=p+1}^{r}(-1+\sqrt{-1}\lambda_j)x_j\overline{y_j}\pm\sqrt{-1}\sum_{j=r+1}^{n}x_j\overline{y_j}.\quad(11.3.21)$$

证 今在 $\mathfrak{L}$ 中任取一组基 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n$, 则

$$h(\alpha,\beta)=h(\sum_{j=1}^{n}\mu_j\alpha_j,\sum_{k=1}^{n}\nu_k\alpha_k)=\mu'(H_1+\sqrt{-1}H_2)\overline{\nu}.$$

由非退化条件, 有 $\mu'H_1\mu=0$, $\mu'H_2\mu=0$ 可推出 $\mu=0$. 由定理 11.3.9, 存在基 $\beta_1,\ \beta_2,\ \cdots,\ \beta_n$, 使得 $h(\alpha,\beta)=h(\sum\limits_{j=1}^{n}x_j\beta_j,\sum\limits_{k=1}^{n}y_k\beta_k)$ 有

$$h(\alpha,\beta)=x'\operatorname{diag}(E_p+\sqrt{-1}\Lambda_1,-E_{r-p}+\sqrt{-1}\Lambda_2,\pm\sqrt{-1}E_{n-r})\overline{y},$$

其中

$$\Lambda_1=\operatorname{diag}(\lambda_1,\cdots,\lambda_p),\quad \Lambda_2=\operatorname{diag}(\lambda_{p+1},\cdots,\lambda_r).$$

这证明了式 (11.3.21) 成立. □

习　题　11.3

11.3.1 试将下列 Hermite 方阵在复相合下化为标准形, 其中 $i=\sqrt{-1}$:

$$\begin{pmatrix}1&1+i&2+i&\cdots&n-1+i\\1-i&1&0&\cdots&0\\2-i&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\n-1-i&0&0&\cdots&1\end{pmatrix},\quad\begin{pmatrix}1&\frac{i}{2}&&&\\-\frac{i}{2}&1&\ddots&&\\&\ddots&\ddots&\frac{i}{2}\\&&-\frac{i}{2}&1\end{pmatrix}.$$

11.3.2 试将下列 Hermite 型化为标准形, 其中 $i=\sqrt{-1}$:

(i) $\sum\limits_{j=1}^{n-1}(x_j\overline{x_{j+1}}+x_{j+1}\overline{x_j})$; **(ii)** $\sum\limits_{j,k=1}^{n}|j-ik|x_j\overline{x}_k$;

(iii) $\sum\limits_{1\leqslant j<k\leqslant n}(j+ik)x_j\overline{x}_k+\sum\limits_{1\leqslant j<k\leqslant n}(j-ik)\overline{x}_jx_k$.

§11.4 正定 Hermite 方阵和复方阵的极分解

定义 11.4.1 n 阶 Hermite 方阵分别称为**正定的**、**半正定的**、**负定的**、**半负定的**, 如果对任一非零 $n\times 1$ 复矩阵 α, 实数 $\alpha^*H\alpha$ 分别适合

$$\alpha^*H\alpha>0,\quad \alpha^*H\alpha\geqslant 0,\quad \alpha^*H\alpha<0,\alpha^*H\alpha\leqslant 0. \tag{11.4.1}$$

这时, 分别用符号

$$H>0,\quad H\geqslant 0,\quad H<0,\quad H\leqslant 0 \tag{11.4.2}$$

记之.

不难看出, §9.4 的全部结果都可以原封不动地推广到 Hermite 方阵的情形. 只要把 "实正交方阵" 易为 "酉方阵", "实对称方阵" 易为 "Hermite 方阵", 实方阵 A 的 "转置" 方阵 A' 全部易为复方阵 A 的 "共轭且转置" 方阵 $\overline{A}'=A^*$ 就可以了. 为了查阅方便起见, 下面仅叙述这些结果:

定理 11.4.2 关于 n 阶 Hermite 方阵 H, 下列诸条件互相等价:

(1) $H>0$;

(2) $P^*HP>0$, 这里 P 是任意取定的一个 n 阶非异复方阵;

(3) H 的特征根都是正实数;

(4) $H=Q^*Q$, 这里 Q 是某一 n 阶非异复方阵;

(5) H 的所有主子式都是正实数;

(6) H 的 n 个顺序主子式

$$H\begin{pmatrix}12\cdots k\\12\cdots k\end{pmatrix}=\begin{vmatrix}h_{11}&\cdots&h_{1k}\\ \vdots&&\vdots\\ h_{k1}&\cdots&h_{kk}\end{vmatrix}>0,\quad k=1,2,\cdots,n \tag{11.4.3}$$

都是正实数.

定理 11.4.3 关于 n 阶 Hermite 方阵 H, 下列诸条件互相等价:

(1) $H\geqslant 0$;

(2) $P^*HP\geqslant 0$, 这里 P 是任意取定的一个 n 阶非异复方阵;

(3) H 的特征根都是非负实数;

(4) $H=Q^*Q$, 这里 Q 是和 H 的秩相同的某一 n 阶复方阵;

(5) H 的所有主子式都是非负实数.

下面给出一个半正定 Hermite 方阵的和的行列式的一个不等式.

定理 11.4.4 假设 A 和 B 是 n 阶半正定 Hermite 方阵, 则存在 n 阶非异复方阵 P, 使得 P^*AP 和 P^*BP 同时为对角形

$$P^*AP = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r, 0_{n-r}), \quad P^*BP = \operatorname{diag}(1-\lambda_1, 1-\lambda_2, \cdots, 1-\lambda_r, 0_{n-r}), \tag{11.4.4}$$

其中 $r = \operatorname{rank}(A+B)$, $1 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0$. 又**行列式不等式**

$$\det(A+B) \geqslant \det(A) + \det(B), \tag{11.4.5}$$

且等号成立的充分且必要条件为 $A = 0$ 或 $B = 0$.

下面给出极分解式

定理 11.4.5 设 Hermite 方阵 H 半正定 (正定), 则存在一个半正定 (正定) Hermite 方阵 H_1, 使得

$$H_1^2 = H, \tag{11.4.6}$$

且任一和方阵 H 可交换的复方阵必和 H_1 可交换. 而 H_1 在 $\det(H) \neq 0$ 时是唯一的, 在 $\det(H) = 0$ 时是不唯一的.

定理 11.4.6 (极分解式) 对任一 n 阶复方阵 A, 存在酉方阵 U 及两个半正定 (在 $\det(A) \neq 0$ 时是两个正定) Hermite 方阵 H_1 和 H_2, 使得

$$A = H_1U = UH_2, \tag{11.4.7}$$

而分解式在 $\det(A) \neq 0$ 时是唯一的, 在 $\det(A) = 0$ 时是不唯一的.

极分解的名称渊源于: 在 A 是一阶复方阵时, 定理 11.4.6 给出复数的极坐标形式: $a = r\exp(\sqrt{-1}\theta)$, 其中 $r \geqslant 0$, $0 \leqslant \theta < 2\pi$.

定义 11.4.7 两个 $n \times m$ 复矩阵 A 和 B 称为**酉相抵的**, 如果存在一个 n 阶酉方阵 U 及一个 m 阶酉方阵 V, 使得

$$B = UAV. \tag{11.4.8}$$

显然, 酉相抵是一个等价关系. 我们有

定理 11.4.8 设 $n \times m$ 复矩阵 A 的秩为 r, 则 A 酉相抵于标准形

$$\begin{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_r \end{pmatrix} & 0^{(r,m-r)} \\ 0^{(n-r,r)} & 0^{(n-r,m-r)} \end{pmatrix}, \tag{11.4.9}$$

其中 $0<\lambda_r\leqslant\cdots\leqslant\lambda_1$, 又 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_r^2$ 是 n 阶半正定 Hermite 方阵 A^*A 的所有非零特征根. 所以, n 阶 Hermite 方阵 A^*A 的非零特征根是复矩阵 A 在酉相抵下的全系不变量.

定理 11.4.9 设 A 是 $n\times m$ 复矩阵, 则

$$\operatorname{rank}(A)=\operatorname{rank}(A^*A)=\operatorname{rank}(AA^*),\tag{11.4.10}$$

且 A^*A 和 AA^* 的所有非零特征根相同.

上面一系列结论的证明都和 §10.2 相同.

习　题　11.4

11.4.1 在 §10.2 的习题中, 哪些题目可以推广到 "复" 的情形, 如何推广? 证明是不是更简单些?

11.4.2 设 $H=(h_{jk})$ 是正定 Hermite 方阵, 试证: $\det(H)\leqslant h_{11}h_{22}\cdots h_{nn}$.

11.4.3 设 A 为 n 阶复方阵, 试证: $|\det A|^2\leqslant\prod\limits_{j=1}^n\sum\limits_{k=1}^n|a_{jk}|^2$.

11.4.4 设 H 为 n 阶正定 Hermite 方阵, 试证:

$$\det(H)\leqslant h_{11}H\begin{pmatrix}23\cdots n\\23\cdots n\end{pmatrix}=h_{11}\det\begin{pmatrix}h_{22}&\cdots&h_{2n}\\\vdots&&\vdots\\h_{n2}&\cdots&h_{nn}\end{pmatrix}.$$

11.4.5 设 A 为 n 阶复方阵, 且设它的每个元素的模都等于 1, 试证: $|\det(A)|^2\leqslant n^n$.

11.4.6 试证: 对任一正定 Hermite 方阵 H 和任一 Hermite 方阵 H_1, 存在一非异方阵 P, 使得 $H=P^*P$, $H_1=P^*\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)P$, 其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是实数.

11.4.7 试证: 对任两正定 Hermite 方阵 H_1 和 H_2, 有

(1) 若 $H_2-H_1>0$, 则 $H_1^{-1}-H_2^{-1}>0$;

(2) $(\det(H_1))^{\frac{1}{n}}+(\det(H_2))^{\frac{1}{n}}\leqslant(\det(H_1+H_2))^{\frac{1}{n}}$, 又等号成立当且仅当 $H_2=aH_1$, 其中 a 为正实数.

11.4.8 设 $n\leqslant m$, 又 A 是 $n\times m$ 矩阵. 设 A 的秩为 n, B 为 $m\times p$ 矩阵, 试证:

$$\det[B^*(E_m-A^*(AA^*)^{-1}A)B]\leqslant\det(B^*B).$$

11.4.9 设 H 为半正定 Hermite 方阵, 记 A 和 B 分别是它的实部及虚部, 则 A 实对称, B 实反对称, 又 $\max(\operatorname{rank}(H),\operatorname{rank}(B))\leqslant\operatorname{rank}(A)$.

11.4.10 设 H 为 n 阶正定 Hermite 方阵, 试证:

$$\det(H)\leqslant\det H\begin{pmatrix}1\\1\end{pmatrix}\det H\begin{pmatrix}23\cdots n\\23\cdots n\end{pmatrix}\leqslant\cdots\leqslant\det H\begin{pmatrix}1\\1\end{pmatrix}\det H\begin{pmatrix}2\\2\end{pmatrix}\cdots\det H\begin{pmatrix}n\\n\end{pmatrix},$$

且有 $\det(H)\leqslant\det(\operatorname{Re}(H))$, 其中等号成立当且仅当 H 为实正定对称方阵. 又有

$$\det(\operatorname{Im}(H))\leqslant\det(\operatorname{Re}(H)).$$

11.4.11 设 H 为 $2n$ 阶正定 Hermite 方阵, 且 $HJ = J\overline{H}$, 其中

$$J = \begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}.$$

试证: 存在 $2n$ 阶非异复方阵 Q, 使得 $QJ = J\overline{Q}$, $H = Q^*Q$. 再证 Q 可取非异上三角方阵.

11.4.12 设 A 为 n 阶复方阵, 记作 $A = (a_{jk})$. 试证: $|\det(A)|^2 \leqslant \prod_{j=1}^{n}\sum_{k=1}^{n}|a_{jk}|^2$.

11.4.13 设 A 和 B 都是 Hermite 方阵, 且 A 正定. 试证: $\det(\lambda A + B)$ 的根都是实数.

11.4.14 设 A 和 B 都是 n 阶复对称方阵, $\det(B) \neq 0$. 设 λ 的多项式 $\det(A + \lambda B)$ 没有重根. 试证: A 和 B 可以同时相合于对角形.

11.4.15 n 阶复方阵 S 称为**酉对称的**, 如果 $S = S'$, 且 $SS^* = E_n$. 试证: 酉对称方阵有下列性质:

(1) S 实正交相似于对角方阵

$$\mathrm{diag}\,(\exp(\sqrt{-1}\theta_1), \exp(\sqrt{-1}\theta_2), \cdots, \exp(\sqrt{-1}\theta_n)), \quad 0 \leqslant \theta_1 \leqslant \cdots \leqslant \theta_n < 2\pi;$$

(2) 存在酉方阵 U, 使得 $S = UU'$;

(3) 存在实对称方阵 S_1, 使得

$$S = \exp(\sqrt{-1}S_1) = \lim_{j\to\infty}\sum_{k=0}^{j}\frac{(\sqrt{-1})^k}{k!}S_1^k = \sum_{k=0}^{\infty}\frac{(\sqrt{-1})^k}{k!}(S_1)^k,$$

这里矩阵序列的极限是指相同行列位置的元素的序列的极限构成的矩阵 (见 §7.3);

(4) 任一酉方阵 U 可以分解为一个实正交方阵 O 和一个酉对称方阵 S 的乘积. 试讨论它的唯一性;

(5) 利用(4), 试证: 对任一酉方阵 U, 存在两实正交方阵 O_1, O_2, 使得

$$O_1UO_2 = \mathrm{diag}\,(\exp(\sqrt{-1}\theta_1), \cdots, \exp(\sqrt{-1}\theta_n)), \quad 0 \leqslant \theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_n < 2\pi.$$

11.4.16 n 阶复方阵 O 称为 **正交 Hermite 方阵**, 如果 $OO' = O'O = E_n$, $O = O^*$. 试证: 正交 Hermite 方阵有下述性质:

(1) 正交 Hermite 方阵 O 实正交相似于标准形

$$\mathrm{diag}\left(\begin{pmatrix} a_1 & \sqrt{-1}b_1 \\ -\sqrt{-1}b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & \sqrt{-1}b_s \\ -\sqrt{-1}b_s & a_s \end{pmatrix}, E_u, -E_v\right),$$

其中 $a_1, \cdots, a_s$ 都是实数, $b_1, \cdots, b_s$ 都是正实数, 且 $a_j^2 - b_j^2 = 1$, $j = 1, 2, \cdots, s$;

(2) 正交 Hermite 方阵 O 必可分解为一个正交 Hermite 方阵 A 和一个实反对称方阵 K 定义的 n 阶方阵

$$\exp(\sqrt{-1}K) = \lim_{j\to\infty}\sum_{k=0}^{j}\frac{(\sqrt{-1})^k}{k!}K^k = \sum_{k=0}^{\infty}\frac{(\sqrt{-1})^k}{k!}K^k$$

的乘积 $O = A\exp(\sqrt{-1}K)$, 其中 $AK = KA$, $A^2 = E_n$. 试讨论分解的唯一性. 当 O 正定时可取 $A = E_n$, 即 $O = \exp(\sqrt{-1}K)$;

(3) 任一复正交方阵 O 可以表为 $O = A\exp(\sqrt{-1}K)$, 其中 A 为实正交方阵, K 为实反对称方阵. 试讨论分解的唯一性.

11.4.17 记 $2n$ 阶反对称方阵

$$J_0 = \mathrm{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right).$$

如果 Hermite 方阵 H 适合 $HJ_0 = J_0\overline{H}$, 则有

(1) 若 α 为 H 的属于特征根 λ_0 的特征向量, 则 $J\alpha$ 为 Hermite 方阵 H' 的属于特征根 λ_0 的特征向量;

(2) H 的特征根为重根, 记作 $\lambda_1, \lambda_1, \cdots, \lambda_n, \lambda_n$, 其中 $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$;

(3) 存在酉方阵 U, 使得 $J_0U = \overline{U}J_0$, 且 H 酉相似于对角形

$$\Lambda = \mathrm{diag}(\lambda_1 E_2, \lambda_2 E_2, \cdots, \lambda_n E_2),$$

又存在酉方阵 U, 使得 $H = U\Lambda U^*$;

(4) 当 $H > 0$ 时, 存在非异方阵 A, 使得

$$J_0 A = \overline{A} J_0,$$

且 $H = AA^*$;

(5) 设 $2n$ 阶复方阵 A 适合 $J_0A = \overline{A}J_0$, 则存在酉方阵 U, 使得 $J_0U = \overline{U}J_0$, 且 UA 为上三角方阵.

提示 设 α_1 是 H 的属于 λ_1 的特征向量, 试证: $J_0\alpha_1$ 是 $\overline{H}$ 的属于 λ_1 的特征向量.

11.4.18 设 $2n$ 阶复方阵 H 为正定 Hermite 方阵, 且 $J_0H = \overline{H}J_0$. 试证: 存在上三角方阵 B, 适合 $J_0B = \overline{B}J_0$, 且 $B^*B = H$.

11.4.19 (华罗庚不等式) 设 A 和 B 是 n 阶复方阵, 使得 $E_n - AA^* \geqslant 0$, $E_n - BB^* \geqslant 0$, 则有

$$|\det(E_n - AB^*)|^2 \geqslant \det(E_n - AA^*)\det(E_n - BB^*).$$

11.4.20 设 A 和 B' 都是 $n \times m$ 复矩阵. 试证: AB 和 BA 同时为规范方阵当且仅当 $A^*AB = BAA^*$, $B^*BA = ABB^*$.

11.4.21 给定 $n \times m$ 复矩阵 $A_1, A_2, \cdots, A_p, B_1, B_2, \cdots, B_p$. 如果存在 n 阶酉方阵 U 及 m 阶酉方阵 V, 使得

$$B_j = UA_jV, \quad j = 1, 2, \cdots, p, \tag{*}$$

则称矩阵组 $A_1, A_2, \cdots, A_p$ 和矩阵组 $B_1, B_2, \cdots, B_p$ 是**酉相抵的**. 显然, 酉相抵是等价关系. 假设矩阵组 $A_1, A_2, \cdots, A_p$ 还适合条件

$$A_j^*A_k + A_k^*A_j = 2\delta_{jk}E_m, \qquad j, k = 1, 2, \cdots, p, \tag{**}$$

则矩阵组 $B_1, B_2, \cdots, B_p$ 也满足此条件. 换句话说, 关系 (**) 在酉相抵下不变.

设 $n = m$, 试证 $2^{[\frac{p-1}{2}]}$ 除得尽 n, 这里 $[\frac{p-1}{2}]$ 是分数 $\frac{p-1}{2}$ 的整数部分, 即在 p 是奇数 $2q+1$ 时, $[\frac{p-1}{2}] = q$, 在 p 是偶数 $2q$ 时, $[\frac{p-1}{2}] = q - 1$.

设 $n=m$, 试定出 n 阶方阵组 $A_1, A_2, \cdots, A_p$ 在酉相抵下的标准形, 从而写出它们在酉相抵下的全系不变量.

(注: 在一般情形, 适合条件 (**) 的 $n \times m$ 矩阵组 $A_1, A_2, \cdots, A_p$ 在酉相抵下的标准形, 当 $p=1, 2$ 时已求得; 当 $p>2$ 时, 迄今未求得在酉相抵下的标准形.)

11.4.22 给定 $n \times m$ 复矩阵 $A_1, A_2, \cdots, A_p$ 和 $B_1, B_2, \cdots, B_p$, 如果存在 n 阶实正交方阵 O 及 m 阶酉方阵 U, 使得

$$B_j = OA_jU \quad j=1,2,\cdots,p, \tag{*}$$

则称矩阵组 $A_1, A_2, \cdots, A_p$ 和 $B_1, B_2, \cdots, B_p$ 是**实正交酉相抵的**. 显然, 这是等价关系. 假设矩阵组 $A_1, A_2, \cdots, A_p$ 还适合条件

$$A_j\overline{A_k}' + \overline{A_k}A_j' = 2\delta_{jk}E_n \quad j,k=1,2,\cdots,p. \tag{**}$$

则和矩阵组 $A_1, A_2, \cdots, A_p$ 实正交酉相抵的矩阵组 $B_1, B_2, \cdots, B_p$ 自然也满足此条件. 换句话说, 关系 (**) 在实正交酉相抵下不变.

试证: 在 $p=1$ 时, 矩阵 A_1 在实正交酉相抵下的标准形为 $\mathrm{diag}\,(P, Q)$, 这里

$$P = \mathrm{diag}\left(E_{e_0}, \begin{pmatrix} E_{e_1} & 0 \\ -\sqrt{-1}a_1E_{e_1} & \sqrt{1-a_1^2}E_{e_1} \end{pmatrix}, \cdots, \begin{pmatrix} E_{e_s} & 0 \\ -\sqrt{-1}a_sE_{e_s} & \sqrt{1-a_s^2}E_{e_s} \end{pmatrix}\right),$$

$$Q = \begin{pmatrix} E_e & 0 \\ -\sqrt{-1}E_e & 0 \end{pmatrix},$$

其中 $e_0, e_1, \cdots, e_s, e$ 都是非负整数, 且 $e_0 + 2\sum\limits_{j=1}^{s} e_j + e = n$. 又, $0 < a_1 < \cdots < a_s < 1$. 再, $e_0, e_1, \cdots, e_s, e$ 和实数 $a_1, a_2, \cdots, a_s$ 由 $\mathrm{Im}\,(QQ^*)$ 在实正交相似下的全系不变量 e_0 个 0, e_1 个 $\pm\sqrt{-1}a_1, \cdots, e_s$ 个 $\pm\sqrt{-1}a_s$, e 个 $\pm\dfrac{\sqrt{-1}}{2}$ 决定, 其中 $\mathrm{Im}\,(QQ^*)$ 表示 QQ^* 的虚部.

(注: 在一般情形 (包括方阵的情形), 适合条件 (**) 的 $n \times m$ 矩阵组 $A_1, A_2, \cdots, A_p$ 在实正交酉相抵下的标准形迄今未求得.)

§11.5　复方阵在酉相合下的标准形 *

定义 11.5.1 两个 n 阶复方阵 A 和 B 称为**酉相合**的, 如果存在一个 n 阶酉方阵 U, 使得

$$B = UAU'. \tag{11.5.1}$$

引理 11.5.2 n 阶复方阵的酉相合关系是一种等价关系.

下面来求复对称方阵及复反对称方阵在酉相合下的标准形和全系不变量. 为此, 先证明一个引理.

引理 11.5.3 对任一 n 阶复方阵 A, 设有一个酉方阵 U, 使得 $UA = AU'$, 则存

在酉方阵 V, 使得 $V^2 = U$, 且 $VA = AV'$.

证 由定理 11.2.6, 存在酉方阵 U_1, 使得

$$U_1^{-1}UU_1 = \mathrm{diag}\,(e^{2i\theta_1}, e^{2i\theta_2}, \cdots, e^{2i\theta_n}),$$

其中 $i = \sqrt{-1}$, $0 \leqslant \theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_n < \pi$. 令

$$V = U_1^{-1}\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U_1,$$

则 V 也是酉方阵, 且 $V^2 = U$. 余下要证 $VA = AV'$.

今已知 $UA = AU'$, 即令 $B = \overline{U_1}'A\overline{U_1}$, 则有

$$\mathrm{diag}\,(e^{2i\theta_1}, e^{2i\theta_2}, \cdots, e^{2i\theta_n})B = B\mathrm{diag}\,(e^{2i\theta_1}, e^{2i\theta_2}, \cdots, e^{2i\theta_n}).$$

于是记 $B = (b_{jk})$, 此即 $(e^{2i\theta_j} - e^{2i\theta_k})b_{jk} = 0$. 当 $b_{jk} \neq 0$ 时有 $e^{2i\theta_j} = e^{2i\theta_k}$, 因此 $e^{i\theta_j} = e^{i\theta_k}$, 所以证明了 $(e^{i\theta_j} - e^{i\theta_k})b_{jk} = 0$; 当 $b_{jk} = 0$ 时必有 $(e^{i\theta_j} - e^{i\theta_k})b_{jk} = 0$. 此即 $(e^{i\theta_j} - e^{i\theta_k})b_{jk} = 0$, $1 \leqslant j, k \leqslant n$, 即

$$\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})(\overline{U_1}'A\overline{U_1}) = (\overline{U_1}'A\overline{U_1})\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n}).$$

所以

$$U_1\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U_1^{-1}A = A[U_1\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U_1^{-1}]'.$$

故 $VA = AV'$, 即 V 适合要求. □

定理 11.5.4 任一 n 阶复对称方阵 S 必酉相合于标准形

$$\Lambda = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n), \tag{11.5.2}$$

其中 $\lambda_1 \geqslant \cdots \geqslant \lambda_n \geqslant 0$, 又 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$ 是半正定 Hermite 方阵 SS^* 的特征根. 所以, 半正定 Hermite 方阵 SS^* 的特征根的算术平方根是对称方阵 S 在酉相合下的全系不变量.

证 由定理 11.4.8 可知, 存在两个酉方阵 U_1 和 U_2, 使得

$$U_1SU_2 = \Lambda,$$

其中 $\lambda_1 \geqslant \cdots \geqslant \lambda_n \geqslant 0$, 又 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$ 是半正定 Hermite 方阵 SS^* 的特征根. 另一方面, 由 $\Lambda' = \Lambda$, 所以 U_1SU_2 是对称方阵, 即有 $(U_1SU_2)' = U_2'SU_1' = U_1SU_2$. 因此,

$$\overline{U_2}U_1S = SU_1'\overline{U_2}' = S(\overline{U_2}U_1)'.$$

记 $U=\overline{U_2}U_1$, 则 U 仍是酉方阵, 且 $US=SU'$. 由引理 11.5.3, 存在酉方阵 V, 使得 $V^2=U=\overline{U_2}U_1$, 又 $VS=SV'$. 今 $VSV'=V^2S=\overline{U_2}U_1S$, 所以

$$U_2'VSV'U_2=U_2'(\overline{U_2}U_1S)U_2=U_1SU_2=\Lambda.$$

取 $U=U_2'V$, 则 $USU'=\Lambda$. □

在求复反对称方阵在酉相合下的标准形之前, 还需要一个引理.

引理 11.5.5 对任一 n 阶非异复反对称方阵 K, 则 KK^* 的特征根必为重根.

证 今 K 非异, $K'=-K$, 所以 K 的秩是偶数, 即阶是偶数 $n=2s$.

为了证明引理, 只要证明特征多项式 $\det(\lambda E_n-KK^*)$ 是一个 s 次多项式的平方就够了. 今

$$\det(\lambda E_n-KK^*)=\det(\lambda E_n+K\overline{K})=\det(K)\det(\lambda K^{-1}+\overline{K}).$$

而 $(\lambda K^{-1}+\overline{K})'=-\lambda K^{-1}-\overline{K}=-(\lambda K^{-1}+\overline{K})$, 所以仍为反对称方阵. 由定理 10.3.1, 存在一个以方阵 $\lambda K^{-1}+\overline{K}$ 的元素的有理函数 (换句话说, 是 λ 的有理函数) 为元素的 n 阶非异方阵 P, 使得

$$P(\lambda K^{-1}+\overline{K})P'=\operatorname{diag}\left(\begin{pmatrix}0&a_1\\-a_1&0\end{pmatrix},\begin{pmatrix}0&a_2\\-a_2&0\end{pmatrix},\cdots,\begin{pmatrix}0&a_s\\-a_s&0\end{pmatrix}\right),$$

其中 $2s=n$. 又 $\det(P)=1$, 且 $a_1,a_2,\cdots,a_s$ 也是 λ 的有理函数. 然而 $\det(\lambda K^{-1}+\overline{K})=\det P(\lambda K^{-1}+\overline{K})P'$, 所以

$$\det(\lambda K^{-1}+\overline{K})=\det\operatorname{diag}\left(\begin{pmatrix}0&a_1\\-a_1&0\end{pmatrix},\cdots,\begin{pmatrix}0&a_s\\-a_s&0\end{pmatrix}\right)=(a_1a_2\cdots a_s)^2.$$

因此, $\det(\lambda E_n-KK^*)=\det(K)\det(\lambda K^{-1}+\overline{K})$ 是 λ 的有理函数的平方, 即存在多项式 $f(\lambda)$ 及 $g(\lambda)$, 使得

$$\det(\lambda E_n-KK^*)=\frac{f(\lambda)^2}{g(\lambda)^2},\quad (f(\lambda),g(\lambda))=1,$$

其中 $g(\lambda)\neq 0$, 且 $g(\lambda)$ 的首项系数为 1. 又因 $\det(\lambda E_n-KK^*)$ 是 λ 的多项式, 而 $(g(\lambda))^2\det(\lambda E_n-KK^*)=(f(\lambda))^2$. 所以多项式 $g(\lambda)^2\mid f(\lambda)^2$. 由 $(f(\lambda),g(\lambda))=1$ 可知 $g(\lambda)=1$. 这也证明了 $\det(\lambda E_n-KK^*)$ 是 λ 的多项式 $f(\lambda)$ 的平方. □

定理 11.5.6 任一 n 阶复反对称方阵 K 必酉相合于下面的标准形:

$$\operatorname{diag}\left(\begin{pmatrix}0&a_1\\-a_1&0\end{pmatrix},\cdots,\begin{pmatrix}0&a_s\\-a_s&0\end{pmatrix},0,\cdots,0\right),\tag{11.5.3}$$

其中 K 的秩是偶数 $r=2s$, $0<a_s\leqslant\cdots\leqslant a_1$, 又 $a_1^2, a_1^2, \cdots, a_s^2, a_s^2$ 是半正定 Hermite 方阵 KK^* 的全部非零特征根. 所以, KK^* 的全部特征根的算术平方根是复反对称方阵 K 在酉相合下的全系不变量.

证 考虑齐次线性方程组 $Kx=0$, 它的解空间为 $\mathbb{C}^n$ 中的 $n-r=n-2s$ 维子空间, 其中 $r=\operatorname{rank}(K)$. 在解空间中任取标准正交基 $\beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$, 于是在 $\mathbb{C}^n$ 中存在标准正交基 $\beta_1, \beta_2, \cdots, \beta_r, \beta_{r+1}, \beta_{r+2}, \cdots, \beta_n$. 记

$$U=(\beta_1,\beta_2,\cdots,\beta_r,\beta_{r+1},\beta_{r+2},\cdots,\beta_n),$$

则 U 为 n 阶酉方阵. 且由 $K\beta_{r+1}=K\beta_{r+2}=\cdots=K\beta_n=0$, 所以有 $KU=(K_1^{(n,r)}, 0^{(n,n-r)})$. 因此

$$U'KU=-U'K'U=(KU)'U=\begin{pmatrix} K_1'U \\ 0 \end{pmatrix}.$$

由于 K 为 n 阶复反对称方阵, 所以 $U'KU$ 也是 n 阶复反对称方阵. 因此有 $U'KU=\begin{pmatrix} K_0 & 0 \\ 0 & 0 \end{pmatrix}$, 其中 K_0 为 r 阶非异复反对称方阵.

因此问题化为在 K 是 n 阶非异复反对称方阵的情形来证明定理, 这时 $n=2s$. 由引理 11.5.5, 所以 KK^* 的特征根为

$$a_1^2=a_1^2\geqslant a_2^2=a_2^2\geqslant\cdots\geqslant a_s^2=a_s^2>0,\quad a_1\geqslant a_2\geqslant\cdots\geqslant a_s>0,$$

且存在 n 阶酉方阵 U_1 和 U_2, 使得 $U_1KU_2=\operatorname{diag}(a_1,a_1,a_2,a_2,\cdots,a_s,a_s)$. 取 n 阶酉方阵

$$U_3=\operatorname{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right),$$

则有 $\Lambda=U_3U_1KU_2$, 其中

$$\Lambda=\operatorname{diag}\left(\begin{pmatrix} 0 & a_1 \\ -a_1 & 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 & a_s \\ -a_s & 0 \end{pmatrix}\right).$$

因此 $U_3U_1KU_2=-(U_3U_1KU_2)'=U_2'KU_1'U_3'$, 所以 $\overline{U_2}U_3U_1K=K(\overline{U_2}U_3U_1)'$, 由引理 11.5.3, 存在 n 阶酉方阵 V, 使得 $V^2=\overline{U_2}U_3U_1$, 且 $VK=KV'$. 令 $U=V'U_2$, 则 U 仍为 n 阶酉方阵, 且有

$$U'KU=U_2'VKV'U_2=U_2'V^2KU_2=U_2'\overline{U_2}U_3U_1KU_2=U_3U_1KU_2=\Lambda.$$

这证明了 $U'KU$ 为标准形. □

由定理 11.5.4 及定理 11.5.6, 可以证明

定理 11.5.7 设 $f(x,y)$ 是 n 维复 Euclid 空间 $\mathfrak{L}$ 中的复对称或复反对称双线性函数, 则存在一组标准正交基, 使得它在这组基下的坐标表达式具有下面的标准形:

(i) $f(x,y)$ 是复对称双线性函数, 则

$$f(x,y)=\lambda_1x_1y_1+\lambda_2x_2y_2+\cdots+\lambda_rx_ry_r,$$

其中 $\lambda_1\geqslant\cdots\geqslant\lambda_r>0$;

(ii) $f(x,y)$ 是复反对称双线性函数, 则

$$f(x,y)=\lambda_1(x_1y_2-x_2y_1)+\lambda_2(x_3y_4-x_4y_3)+\cdots+\lambda_t(x_{2t-1}y_{2t}-x_{2t}y_{2t-1}),$$

其中 $\lambda_1\geqslant\cdots\geqslant\lambda_t>0$.

§11.6 复方阵在复正交相合下的标准形 *

定义 11.6.1 n 阶复方阵 O 称为**复正交方阵**, 如果

$$OO'=O'O=E_n. \tag{11.6.1}$$

这时有 $O^{-1}=O'$. 所有 n 阶复正交方阵构成集合 $O(n,\mathbb{C})$, 称为**复正交群**.

引理 11.6.2 n 阶复正交群 $\mathrm{O}(n,\mathbb{C})$ 有性质

(1) 任取 $O_1,O_2\in\mathrm{O}(n,\mathbb{C})$, 则 $O_1O_2\in\mathrm{O}(n,\mathbb{C})$;

(2) 复正交方阵的乘法适合结合律;

(3) n 阶单位方阵 $E_n\in\mathrm{O}(n,\mathbb{C})$;

(4) 任取 $O\in\mathrm{O}(n,\mathbb{C})$, 则 $\overline{O}$, $O'\in\mathrm{O}(n,\mathbb{C})$, 又 $O^{-1}=O'\in\mathrm{O}(n,\mathbb{C})$;

(5) 复正交方阵的行列式等于 ±1, 即 $\det(O)=\pm1$.

定义 11.6.3 n 阶复方阵 A 和 B 称为**复正交相似的**, 或称为**复正交相合的**, 如果存在复正交方阵 O, 使得 $B=O'AO$.

引理 11.6.4 复正交相合为等价关系.

定理 11.6.5 n 阶复对称 (复反对称、复正交) 方阵 A 如果相似于 n 阶复对称 (复反对称、复正交) 方阵 B, 则 A 和 B 必复正交相似.

证 今存在 n 阶非异复方阵 P, 使得

(a) 假设 $A'=A$ 和 $(PAP^{-1})'=PAP^{-1}$, 则

$$PAP^{-1}=(PAP^{-1})'=(P^{-1})'AP',$$

因此, $P'PA=AP'P$;

(b) 假设 $A'=-A$ 和 $(PAP^{-1})'=-PAP^{-1}$, 则

$$PAP^{-1}=-(PAP^{-1})'=(P^{-1})'AP',$$

因此, $P'PA=AP'P$;

(c) 假设 $AA' = E_n$ 和 $(PAP^{-1})(PAP^{-1})' = E_n$, 则

$$(PAP^{-1})(PAP^{-1})' = PAP^{-1}(P')^{-1}A^{-1}P' = E,$$

因此, $AP^{-1}(P')^{-1} = P^{-1}(P')^{-1}A$, 所以, $P'PA^{-1} = A^{-1}P'P$. 因此, $AP'P = P'PA$.

由 (a), (b), (c) 可知这三种情形都导出 $P'PA = AP'P$.

由定理 7.3.16 可知, 存在非异方阵 $P'P$ 的多项式 $S = g(P'P)$, 使得 $S^2 = P'P$. 自然, S 是复对称方阵, 且 $SA = AS$. 另一方面, 由 $(PS^{-1})'(PS^{-1}) = S^{-1}P'PS^{-1} = E_n$, 可知 $O = PS^{-1}$ 是复正交方阵. 所以 $P = OS$, 而

$$B = PAP^{-1} = OSAS^{-1}O' = OAO'. \qquad \square$$

从定理 11.6.5 的证明, 还告诉我们一种新形式的极分解式, 即有

定理 11.6.6 任一 n 阶复方阵 A 必可分解为

$$A = OS = S_1O, \tag{11.6.2}$$

其中 O 是 n 阶复正交方阵, S 和 S_1 是 n 阶复对称方阵, 又 $\operatorname{rank}(S) = \operatorname{rank}(S_1) = \operatorname{rank}(A)$.

证 设 A 为 n 阶非异复方阵. 下面证明定理成立. 今 AA' 为 n 阶非异复对称方阵. 由定理 7.3.16, 存在 AA' 的多项式 $p(AA') = S_1$, S_1 仍为复对称方阵, 且 $S_1^2 = AA'$. 记 $O = S_1^{-1}A$, 则

$$OO' = S_1^{-1}AA'S_1^{-1} = S_1^{-1}S_1^2S_1^{-1} = E_n.$$

所以 O 为复正交方阵, 且有 $A = S_1O = O(O'S_1O)$. 记 $S = O'S_1O$, 则 S 仍为复对称方阵. 所以有分解 $A = OS = S_1O$.

设 A 为 n 阶复方阵. 由定理 10.2.10 的证明在复正交相抵下也对. 所以记 $r = \operatorname{rank}(A)$, 则存在 n 阶复正交方阵 O_1 和 O_2, 使得

$$A = O_1\begin{pmatrix} A_1^{(r)} & 0 \\ 0 & 0 \end{pmatrix}O_2,$$

其中 A_1 为 r 阶非异复方阵, 因此, 存在 r 阶复正交方阵 O_3 和 r 阶非异复对称方阵 S_3, 使得 $A_1 = O_3S_3$. 于是

$$A = O_1\begin{pmatrix} O_3S_3 & 0 \\ 0 & 0 \end{pmatrix}O_2 = O_1\operatorname{diag}(O_3, E_{n-r})O_2O_2'\begin{pmatrix} S_3 & 0 \\ 0 & 0 \end{pmatrix}O_2.$$

记 $O = O_1\operatorname{diag}(O_3, E_{n-r})O_2$, 则 O 仍为复正交方阵. 记 $S = O_2'\begin{pmatrix} S_3 & 0 \\ 0 & 0 \end{pmatrix}O_2$, 则 S 仍

为复对称方阵, 且有 $A = OS$. 又 $A = OS = (OSO')O$, 则 $S_1 = OSO'$ 仍为复对称方阵.

□

下面有一个有趣的应用, 它在理论物理中有应用.

定理 11.6.7 设 H 为 n 阶正定 Hermite 方阵, 则存在复正交方阵 O, 使得 H 复正交复相合于

$$OHO^* = \operatorname{diag}(\lambda_1 E_{m_1}, \lambda_2 E_{m_2}, \cdots, \lambda_t E_{m_t}),$$

其中 $m_1 + m_2 + \cdots + m_t = n$, $\lambda_1 > \lambda_2 > \cdots > \lambda_t > 0$ 都是正实数, 且 m_1 个 λ_1^2, m_2 个 $\lambda_2^2, \cdots$, m_t 个 λ_t^2 是 $H\overline{H} = HH'$ 的全部特征根. 所以 HH' 的全部特征根的算术平方根是正定 Hermite 方阵 H 在复正交方阵的复相合下的全系不变量.

证 今 $H > 0$, 所以存在 n 阶非异复方阵 Q, 使得 $H = Q'\overline{Q}$. 今 $QHQ^* > 0$, 所以存在 n 阶酉方阵 U, 使得

$$UQHQ^*U^* = \operatorname{diag}(\mu_1 E_{m_1}, \mu_2 E_{m_2}, \cdots, \mu_t E_{m_t}) = \Lambda,$$

其中 $\mu_1 > \mu_2 > \cdots > \mu_t$, $m_1 + \cdots + m_t = n$. 因此

$$H\overline{H} = (UQ)^{-1}\Lambda U(Q^*)^{-1}\overline{H} = (UQ)^{-1}\Lambda U(Q^*)^{-1}(Q^*Q) = (UQ)^{-1}\Lambda UQ.$$

所以 $H\overline{H}$ 和 Λ 相似. 今 $\Lambda' = \Lambda$, $(H\overline{H})' = H\overline{H}$, 即 Λ 和 $H\overline{H}$ 都是复对称方阵. 由定理 11.6.5, 存在复正交方阵 O_0, 使得 $H\overline{H} = O_0'\Lambda O_0$. 于是

$$\Lambda = O_0 H\overline{H}O_0' = (O_0 H\overline{O_0}')(\overline{O_0 H}O_0') = (O_0 H O_0^*)(O_0 H O_0^*)'.$$

记 $H_0 = O_0 H O_0^*$, 则 H_0 仍为正定 Hermite 方阵, 且 $H_0 H_0' = \Lambda$. 所以 $H_0' = H_0^{-1}\Lambda$ 是 Hermite 方阵, 即有 $H_0^{-1}\Lambda = \Lambda H_0^{-1}$. 即 $\Lambda H_0 = H_0 \Lambda$. 于是将 H_0 和 Λ 同样分块, 则有

$$H_0 = \operatorname{diag}(H_1^{(m_1)}, H_2^{(m_2)}, \cdots, H_t^{(m_t)}).$$

由 $\Lambda = H_0 H_0'$, 有 $H_j H_j' = \mu_j E_{m_j}$, $1 \leqslant j \leqslant t$. 于是 $O_j = \dfrac{1}{\sqrt{\mu_j}} H_j$ 为 m_j 阶复正交方阵, 且为 m_j 阶正定 Hermite 方阵.

将 O_j 的实部和虚部分开: $O_j = S_j + \sqrt{-1}K_j$, 则 S_j 是 m_j 阶实对称方阵, K_j 是 m_j 阶实反对称方阵, 且

$$E_{m_j} = O_j O_j' = (S_j + \sqrt{-1}K_j)(S_j - \sqrt{-1}K_j) = S_j^2 + K_j^2 + \sqrt{-1}(K_j S_j - S_j K_j).$$

即有

$$S_j^2 + K_j^2 = E, \qquad K_j S_j = S_j K_j.$$

而 $O_j > 0$ 等价于实对称方阵

$$\begin{pmatrix} S_j & K_j \\ -K_j & S_j \end{pmatrix} > 0.$$

于是不难证明, 存在 m_j 阶实正交方阵 P_j, 使得

$$O_j = P_j' \mathrm{diag}\,(T_{1j}, T_{2j}, \cdots, T_{n_j j}) P_j,$$

其中 $2n_j = m_j$,

$$T_{kj} = \begin{pmatrix} \mathrm{ch}\,(2\theta_{kj}) & \sqrt{-1}\mathrm{sh}\,(2\theta_{kj}) \\ -\sqrt{-1}\mathrm{sh}\,(2\theta_{kj}) & \mathrm{ch}\,(2\theta_{kj}) \end{pmatrix}, \quad 1 \leqslant k \leqslant n_j.$$

取

$$Q_j = P_j' \mathrm{diag}\,(L_{1j}, L_{2j}, \cdots, L_{n_j j}) P_j,$$

其中

$$L_{kj} = \begin{pmatrix} \mathrm{ch}\,(\theta_{kj}) & \sqrt{-1}\mathrm{sh}\,(\theta_{kj}) \\ -\sqrt{-1}\mathrm{sh}\,(\theta_{kj}) & \mathrm{ch}\,(\theta_{kj}) \end{pmatrix}, \quad 1 \leqslant k \leqslant n_j.$$

则 $L_{kj}\overline{L_{kj}} = (\mathrm{ch}^2(\theta_{kj}) - \mathrm{sh}^2(\theta_{kj}))E_2 = E_2$. 又 Q_j 仍为正定 Hermite 方阵, 且 Q_j 复正交, 又 $Q_j^2 = O_j$. 因此,

$$Q_j' O_j \overline{Q_j} = Q_j' Q_j^2 \overline{Q_j} = Q_j \overline{Q_j} = E_{m_j}, \quad j = 1, 2, \cdots, t.$$

取

$$O = \mathrm{diag}\,(Q_1', Q_2', \cdots, Q_t') O_0,$$

于是

$$\begin{aligned} OHO^* &= \mathrm{diag}\,(Q_1', Q_2', \cdots, Q_t') O_0 H O_0^* \mathrm{diag}\,(\overline{Q_1}, \overline{Q_2}, \cdots, \overline{Q_t}) \\ &= \mathrm{diag}\,(Q_1', Q_2', \cdots, Q_t') H_0 \mathrm{diag}\,(\overline{Q_1}, \overline{Q_2}, \cdots, \overline{Q_t}) \\ &= \mathrm{diag}\,(Q_1' H_1 \overline{Q_1}, Q_2' H_2 \overline{Q_2}, \cdots, Q_t' H_t \overline{Q_t}) \\ &= \mathrm{diag}\,(\sqrt{\mu_1} Q_1' O_1 \overline{Q_1}, \cdots, \sqrt{\mu_t} Q_t' O_t \overline{Q_t}) \\ &= \mathrm{diag}\,(\sqrt{\mu_1} E_{m_1}, \cdots, \sqrt{\mu_t} E_{m_t}). \end{aligned}$$

记 $\lambda_j = \sqrt{\mu_j}$, $j = 1, 2, \cdots, t$, 就证明了定理. □

可以指出, 当 H 是半正定方阵时, 不一定存在复正交方阵, 使得 H 复正交复相合于对角形. 例如 $H = \begin{pmatrix} 1 & -\sqrt{-1} \\ \sqrt{-1} & 1 \end{pmatrix}$ 便是. 弄清楚哪些行, 哪些不行, 是一个很好的习题, 读者可以考虑考虑.

习　　题　11.6

11.6.1 证明: A 为反 Hermite 方阵当且仅当 $\exp(tA)$ 为酉方阵, $\forall t \in \mathbb{R}$; 又证: A 为复反对称方阵当且仅当 $\exp(tA)$ 为复正交方阵, $\forall t \in \mathbb{R}$.

11.6.2 设 O 为 n 阶复正交正定 Hermite 方阵, 试证: 存在实反对称方阵 K, 使得 $O = \exp(\sqrt{-1}K)$.

11.6.3 记 n 阶复方阵

$$S_n = \begin{pmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 1 \\ & & & & 1 & 0 \end{pmatrix} + \sqrt{-1}\begin{pmatrix} & & & & 1 & 0 \\ & & & 1 & 0 & -1 \\ & & & \cdot\ \cdot\ \cdot & & \\ & & \cdot\ \cdot\ \cdot & & & \\ & \cdot\ \cdot\ \cdot & & & & \\ 1 & 0 & -1 & & & \\ 0 & -1 & & & & \end{pmatrix}.$$

试证: n 阶复方阵 $\lambda_0 E_n + S_n$ 的初等因式组为 $\{(\lambda - \lambda_0)^n\}$, 所以相似于 Jordan 块 $\lambda_0 E_n + N_n$, 因此有

(1) 任一 n 阶复方阵相似于复对称方阵;

(2) 任一 n 阶复对称方阵复正交相似于

$$\mathrm{diag}\,(\lambda_1 E_{e_1} + S_{e_1}, \lambda_2 E_{e_2} + S_{e_2}, \cdots, \lambda_t E_{e_t} + S_{e_t}).$$

第十二章　广义逆矩阵

§12.1　线性方程组的最小二乘解 *

为了从几何学角度来说明为什么要引进广义逆矩阵. 下面先引进域 $\mathbb{F}$ 上 n 维点空间 $(\mathbb{F}^n, O)$, 这里 O 为原点, 它是一个固定点. 再引进 n 维点空间 $(\mathbb{F}^n, O)$ 的解析几何. 当 $\mathbb{F}=\mathbb{C}$ 时, 称为 ***n* 维复点空间**; 当 $\mathbb{F}=\mathbb{R}$ 时, 称为 ***n* 维实点空间**.

熟知平面解析几何学是研究如何将平面的几何性质化为代数问题, 然后利用代数工具来解决这些代数问题. 办法是在普通平面中引进直角坐标系, 使得平面中的点有唯一的实数组 (x, y) 为此点的坐标. 因此, 平面为二维实点空间, 记作 $(\mathbb{R}^2, O)$, 其中 O 为原点. 对平面中任两点 P 和 Q, 坐标分别为 $P=(a_1, a_2)$ 和 $Q=(b_1, b_2)$. 以 P, Q 为端点的线段记作 $\overline{PQ}$. 以 P 为起点, Q 为终点的有向线段, 称为**固定向量**, 记作 $\overrightarrow{PQ}$. 于是固定向量 $\overrightarrow{PQ}$ 有坐标 (b_1-a_1, b_2-a_2). 而线段 $\overline{PQ}$ 的长度等于固定向量 $\overrightarrow{PQ}$ 的长度, 它定义为

$$\mathrm{d}(P, Q)=|\overline{PQ}|=|\overrightarrow{PQ}|=\sqrt{(a_1-b_1)^2+(a_2-b_2)^2}. \tag{12.1.1}$$

这里要注意, 坐标原点为点 O, 它的坐标为 $(0,0)$. 两个坐标轴上分别有坐标为 $e_1=(1,0)$ 及 $e_2=(0,1)$ 的单位向量, 它们互相正交.

平面上两向量 $\overrightarrow{PQ}=(q_1-p_1, q_2-p_2)$ 和 $\overrightarrow{ST}=(t_1-s_1, t_2-s_2)$ 称为**相等的**, 如果 $t_i-s_i=q_i-p_i$, $i=1,2$. 显然这是等价关系. 所以平面上所有固定向量构成的集合按此等价关系分成了等价类的并集. 它们两两不相交. 每个等价类看作一个向量, 即将所有相等的向量看作一个**自由向量**.

由于固定向量 $\overrightarrow{PQ}$ 所在的等价类 (自由向量), 可用固定向量 $\overrightarrow{PQ}$ 的坐标 (q_1-p_1, q_2-p_2) 来表示, 所以 (q_1-p_1, q_2-p_2) 也称为这个自由向量的坐标. 由于 q_1, q_2,

p_1, p_2 可任取, 所以证明了所有自由向量构成一个二维实线性空间 $\mathbb{R}^2$. 它的任意一个一维子空间 $\mathfrak{L}$ 中的向量, 都按在平面上一个固定点 P 上, 即都以 P 点为起点. 则终点构成平面中子集合, 称为**点子空间**, 记作 $(\mathfrak{L}, P)$, 它就是直线. 平面称为**二维实点空间**, 或**二维实 Euclid 点空间**. 此平面记作 $(\mathbb{R}^2, O)$, 其中 O 为原点. 在平面中任意取定一点 P为起点, 此平面又可记作 $(\mathbb{R}^2, P)$.

上面关于平面的讨论也就启发我们可以推广到普通的空间的情形. 因此也可以推广到抽象的 n 维情形. 即引进域 $\mathbb{F} = \mathbb{R}$ 或 $\mathbb{C}$ 上 n 维点空间, 它也称为**域 $\mathbb{F}$ 上 n 维 Euclid 点空间**. 由于和普通情形不一样, 没有一个几何实体, 但是我们可以用点坐标 (即 n 个复数或 n 个实数) 来代替.

定义 12.1.1 域 $\mathbb{F}$ 上所有 $n \times 1$ 矩阵作为**点**, 全体构成域 $\mathbb{F}$ 上 **n 维点空间**, 记作 $(\mathbb{F}^n, O)$, 其中 O 为固定点. $(\mathbb{F}^n, O)$ 中点 P 的**坐标**为 $P = (p_1, \cdots, p_n)'$. 坐标为 $(0, 0, \cdots, 0)'$ 的点称为 **原点**. 坐标为 $e_i = (\delta_{i1}, \cdots, \delta_{in})'$ 的点记作 P_i, 其中 δ_{ij} 为 Kronecker 符号. 则 $\overrightarrow{OP_i}$ 为单位向量, 由点 O 及 P_i 决定的直线构成第 i 个**坐标轴**, $i = 1, 2, \cdots, n$. 又单位向量 $\overrightarrow{OP_1}, \overrightarrow{OP_2}, \cdots, \overrightarrow{OP_n}$ 两两正交.

定义 12.1.2 设 $(\mathbb{F}^n, O)$ 为域 $\mathbb{F}$ 上 n 维点空间. 在 $(\mathbb{F}^n, O)$ 中任取两点 $P = (p_1, \cdots, p_n)'$ 和 $Q = (q_1, \cdots, q_n)'$, 则以点 P 为起点, 以点 Q 为终点的有向线段 $\overrightarrow{PQ}$ 称为**固定向量**. 它有坐标 $\overrightarrow{PQ} = (q_1 - p_1, \cdots, q_n - p_n)'$.

两固定向量 $\overrightarrow{PQ} = (q_1 - p_1, \cdots, q_n - p_n)'$ 和 $\overrightarrow{ST} = (t_1 - s_1, \cdots, t_n - s_n)'$ 称为**相等**, 如果它们的坐标相等, 即 $t_i - s_i = q_i - p_i$, $1 \leqslant i \leqslant n$. 相等是等价关系, 等价类定义为**自由向量**. 包含固定向量 $\overrightarrow{PQ}$ 的自由向量有坐标 $(q_1 - p_1, \cdots, q_n - p_n)'$, 即和固定向量 $\overrightarrow{PQ}$ 有相同的坐标表示. 点空间 $(\mathbb{F}^n, O)$ 中所有自由向量构成域 $\mathbb{F}$ 上 n 维线性空间 $\mathbb{F}^n$. 因此任取 r 维子空间 $\mathfrak{L}$, 在 n 维点空间 $(\mathbb{F}^n, O)$ 中任取一点 P, 将 $\mathfrak{L}$ 中所有自由向量都按在起点 P 上, 则所有终点构成 r 维**点子空间**, 记作 $(\mathfrak{L}, P)$. 显然另取点子空间 $(\mathfrak{L}, P)$ 中一点 Q, 则 $(\mathfrak{L}, P) = (\mathfrak{L}, Q)$.

定义 12.1.3 在域 $\mathbb{F}^n$ 上 n 维线性空间 $\mathbb{F}^n$ 中引进内积如下: 当域 $\mathbb{F} = \mathbb{R}$ 时, 自由向量 $x = (x_1, \cdots, x_n)'$ 和 $y = (y_1, \cdots, y_n)'$ 的**内积**为 $(x, y) = \sum\limits_{i=1}^{n} x_i y_i$; 当域 $\mathbb{F} = \mathbb{C}$ 时, 自由向量 x 和 y 的**内积**为 $(x, y) = \sum\limits_{i=1}^{n} x_i \overline{y}_i$. 于是在 $\mathbb{F}^n$ 中有标准正交基 e_1, $e_2, \cdots, e_n$, 其中 $e_i = (\delta_{1i}, \cdots, \delta_{ni})$, $1 \leqslant i \leqslant n$. 又 δ_{ij} 是 Kronecker 符号, $1 \leqslant i, j \leqslant n$. 而任一自由向量可表为

$$x = \sum_{i=1}^{n} x_i e_i, \qquad \forall x_1, \cdots, x_n \in \mathbb{F}. \tag{12.1.2}$$

这时 n 维线性空间 $\mathbb{F}^n$ 为**域 F 上 n 维Euclid 空间**. 当 $\mathbb{F} = \mathbb{R}$ 时, $\mathbb{R}^n$ 为**实 Euclid 空间**; 当 $\mathbb{F} = \mathbb{C}$ 时, $\mathbb{C}^n$ **复 Euclid 空间**.

定义 12.1.4 设 $(\mathbb{F}^n, O)$ 为域 $\mathbb{F}$ 上 n 维点空间. 对 $(\mathbb{F}^n, O)$ 中任两点 P 和 Q

(见定义 12.1.2), 则线段 $\overline{PQ}$ 的**长度**为

$$|\overline{PQ}|=|\overrightarrow{PQ}|=|\sum_{k=1}^{n}(q_k-p_k)e_k|=\sqrt{\sum_{i=1}^{n}|q_i-p_i|^2}=\sqrt{(\overrightarrow{PQ},\overrightarrow{PQ})}. \tag{12.1.3}$$

线段 $\overline{PQ}$ 的长度也称为点 P 和 Q 的**距离**, 改记作 $\mathrm{d}\,(P,Q)$.

定义 12.1.5 在域 $\mathbb{F}$ 上 n 维点空间 $(\mathbb{F}^n,O)$ 中任取两线段 $\overline{PQ}$ 和 $\overline{ST}$. 我们说 $\overline{PQ}\perp\overline{ST}$, 即线段 $\overline{PQ}$ 和 $\overline{ST}$ **互相正交**, 如果固定向量 $\overrightarrow{PQ}$ 和 $\overrightarrow{ST}$ 所代表的自由向量互相正交.

我们说线段 $\overline{PQ}$ 和 $\overline{RS}$ **互相平行**, 如果固定向量 $\overrightarrow{PQ}$ 和 $\overrightarrow{ST}$ 所代表的自由向量互相平行, 即差一个纯量倍数.

引理 12.1.6 在域 $\mathbb{F}$ 上 n 维点空间 $(\mathbb{F}^n,O)$ 中任取三点 P, Q, R, 则距离函数有性质

(1) $\mathrm{d}\,(P,Q)\geqslant 0$, 且等号成立当且仅当 $P=Q$;

(2) $\mathrm{d}\,(P,Q)=\mathrm{d}\,(Q,P)$;

(3) **三角不等式**成立, 即

$$\mathrm{d}\,(P,Q)+\mathrm{d}\,(Q,R)\geqslant\mathrm{d}\,(P,R). \tag{12.1.4}$$

证 前两性质显然成立. 下面证第三个性质成立. 今由固定向量相加的平行四边形法则, 有 $\overrightarrow{PQ}+\overrightarrow{QR}=\overrightarrow{PR}$. 所以记 $\alpha=\overrightarrow{PQ}$, $\beta=\overrightarrow{QR}$, 则 $\alpha+\beta=\overrightarrow{PR}$. 而

$$\begin{aligned}\mathrm{d}\,(P,R)^2=|\overrightarrow{PR}|^2&=|\alpha+\beta|^2=(\alpha+\beta,\alpha+\beta)\\&=(\alpha,\alpha)+(\alpha,\beta)+(\beta,\alpha)+(\beta,\beta)\\&=|\alpha|^2+|\beta|^2+2\mathrm{Re}\,((\alpha,\beta)).\end{aligned}$$

由 Cauchy 公式, 有 $\mathrm{Re}\,((\alpha,\beta))\leqslant|\mathrm{Re}\,((\alpha,\beta))|\leqslant|(\alpha,\beta)|\leqslant|\alpha||\beta|$. 因此

$$\mathrm{d}\,(P,R)^2\leqslant|\alpha|^2+|\beta|^2+2|\alpha||\beta|=(|\alpha|+|\beta|)^2.$$

这证明了

$$\mathrm{d}\,(P,R)\leqslant|\alpha|+|\beta|=|\overrightarrow{PQ}|+|\overrightarrow{QR}|=\mathrm{d}\,(P,Q)+\mathrm{d}\,(Q,R).$$ □

定义 12.1.7 在域 $\mathbb{F}$ 上 n 维点空间 $(\mathbb{F}^n,O)$ 中任取一点 Q 和一点子空间 $(\mathfrak{L},P)$, 则数值

$$\mathrm{d}\,(Q,(\mathfrak{L},P))=\min_{T\in(\mathfrak{L},P)}\mathrm{d}\,(Q,T) \tag{12.1.5}$$

称为点 Q 到点子空间 $(\mathfrak{L},P)$ 的**距离**.

下面给出普通空间中 "任一点到任一平面有且只有一条法线, 即垂直于此平面的线段, 且垂点在平面上" 这一几何事实的推广.

引理 12.1.8 在域 $\mathbb{F}$ 上 n 维点空间 $(\mathbb{F}^n,O)$ 中任取一点 Q 和一点子空间 $(\mathfrak{L},P)$,

使得 $Q \notin (\mathfrak{L}, P)$, 则在点子空间 $(\mathfrak{L}, P)$ 中唯一存在一点 S, 使得线段 $\overline{PS}$ 垂直于点子空间 $(\mathfrak{L}, P)$ 中任一线段, 且有

$$\min_{T\in(\mathfrak{L},P)} \mathrm{d}(T, Q) = \mathrm{d}(Q, S). \tag{12.1.6}$$

证 今 $\mathfrak{L}$ 为域 $\mathbb{F}$ 上 n 维 Euclid 空间 $\mathbb{F}^n$ 的子空间. 于是 $\mathfrak{L}$ 有正交补

$$L^{\perp} = \{\, \gamma \in \mathbb{F}^n \mid (\gamma, \mathfrak{L}) = 0 \,\},$$

且 $\mathbb{F}^n = \mathfrak{L} \oplus \mathfrak{L}^{\perp}$ 为子空间直接和.

在点子空间 $(\mathfrak{L}, P)$ 中任取一点 T, 则有固定向量 $\overrightarrow{QT}$. 以 $\overrightarrow{QT}$ 为代表的自由向量记作 α, 则 $\alpha \in \mathbb{F}^n$. 于是唯一存在 $\beta \in \mathfrak{L}$, $\gamma \in \mathfrak{L}^{\perp}$, 使得 $\alpha = \beta + \gamma$.

将自由向量 $-\beta \in \mathfrak{L}$ 以 T 为起点, 则在点子空间 $(\mathfrak{L}, P)$ 中存在点 S, 使得固定向量 $\overrightarrow{TS}$ 为代表的自由向量就是 $-\beta$. 作固定向量 $\overrightarrow{SQ}$, 它代表的自由向量记作 $\overrightarrow{SQ} = \gamma_1$. 由平行四边形法则, 有 $\gamma_1 = \overrightarrow{SQ} = \overrightarrow{ST} + \overrightarrow{TQ} = \beta - \alpha = \beta - (\beta + \gamma) = -\gamma$. 所以 $\gamma = \overrightarrow{QS} \in \mathfrak{L}^{\perp}$. 注意到 $\gamma \in \mathfrak{L}^{\perp}$, 因此 $(\gamma, \mathfrak{L}) = 0$. 这证明了点子空间 $(\mathfrak{L}, P)$ 中任一线段和线段 $\overrightarrow{QS}$ 垂直. 这也证明了存在点 Q 到子点空间 $(\mathfrak{L}, P)$ 的垂线, 垂足为 $S \in (\mathfrak{L}, P)$.

下面证从点 Q 到点子空间 $(\mathfrak{L}, P)$ 只能作一条垂线. 事实上, 若对点子空间 $(\mathfrak{L}, P)$ 中另取一点 T_1, 同上作出一点 $S_1 \in (\mathfrak{L}, P)$, 且点子空间 $(\mathfrak{L}, P)$ 中任一线段和线段 $\overline{QS_1}$ 垂直, 那末对点子空间 $(\mathfrak{L}, P)$ 中线段 $\overline{SS_1}$, $\overline{QS} \perp \overline{SS_1}$, $\overline{QS_1} \perp \overline{SS_1}$. 注意到当 $S \neq S_1$ 时, 点空间 $\mathbb{F}^n$ 中三点 Q, S, S_1 不共线. 它们决定了一平面. 而三角形 QSS_1 有两个内角为 $\dfrac{\pi}{2}$. 这推出另一内角为零. 所以证明了 $S = S_1$. 即点子空间 $(\mathfrak{L}, P)$ 外任一点 Q 能且只能作一条垂线, 垂直于点子空间 $(\mathfrak{L}, P)$.

最后, 我们来证线段 $\overline{QS}$ 的距离 $\mathrm{d}(Q, S)$ 为点 Q 到点子空间 $(\mathfrak{L}, P)$ 的距离. 事实上, 任取 $T \in (\mathfrak{L}, P)$, 则由 $(\beta, \gamma) = 0$, 有

$$|\overline{QT}|^2 = |\alpha|^2 = |\beta + \gamma|^2 = |\beta|^2 + |\gamma|^2 \geqslant |\gamma|^2 = |\overline{QS}|^2.$$

所以

$$\min_{T\in(\mathfrak{L},P)} \mathrm{d}(Q, T) = \min_{T\in(\mathfrak{L},P)} |\overline{QT}| \geqslant |\overline{QS}| = \mathrm{d}(Q, S).$$

由 $S \in (\mathfrak{L}, P)$ 可知 $\mathrm{d}(Q, S)$ 为点 Q 到点子空间 $(\mathfrak{L}, P)$ 的距离. □

为了给出最小二乘解的定义, 我们先说明它的起源. 在实际测量中, 不可避免地要出现各种误差. 例如人为观察造成的误差, 仪器测量时产生的误差, 以及用四舍五入法处理数据引起的误差等. 于是由观测数据导出的线性方程组

$$\sum_{k=1}^{m} a_{jk} x_k = b_j, \quad 1 \leqslant j \leqslant n \tag{12.1.7}$$

中的数值 a_{jk} 和 b_j 实际上都是近似值. 因此或者出现这个线性方程组不相容的情形; 或者虽然相容, 可是在系数值和常数项值已经改变的情况下, 所得的解并不见得能准确地反映客观实际的数据. 很自然地, 我们希望寻找误差最小的解, 也就是说, 从数理统计学的角度来处理数据. 最自然和最通用的方法就是所谓的最小二乘法.

定义 12.1.9 对域 $\mathbb{F}$ 上 $n\times m$ 矩阵 A, $m\times 1$ 矩阵 x 及 $n\times 1$ 矩阵 q, 记

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm}\end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ \vdots \\ x_m\end{pmatrix},\quad q=\begin{pmatrix} q_1 \\ \vdots \\ q_n\end{pmatrix}. \tag{12.1.8}$$

线性方程组

$$Ax=q \tag{12.1.9}$$

的**最小二乘解**为 $m\times 1$ 矩阵 x_0, 使得

$$\min_{x\in\mathbb{F}^m}|Ax-q|=|Ax_0-q|. \tag{12.1.10}$$

适合上述条件的解 $x_0\in\mathbb{F}^m$ 不唯一, 其中模最小的解称为**模最小解**.

下面先给出最小二乘解的几何意义.

我们考虑域 $\mathbb{F}$ 上点空间 $(\mathbb{F}^m,O)$ 和点空间 $(\mathbb{F}^n,O)$. 它们分别决定了域 $\mathbb{F}$ 上 Euclid 空间 $\mathbb{F}^m$ 和 $\mathbb{F}^n$. 而矩阵 A 定义了域 $\mathbb{F}$ 上 Euclid 空间 $\mathbb{F}^m$ 到 $\mathbb{F}^n$ 内的线性映射 $\mathscr{A}$, 象空间 $\mathscr{A}(\mathbb{F}^m)\subset\mathbb{F}^n$. 于是

$$\mathbb{F}^n=\mathscr{A}(\mathbb{F}^m)\oplus(\mathscr{A}(\mathbb{F}^m))^{\perp} \tag{12.1.11}$$

为子空间的直接和分解.

定理 12.1.10 设 A 为域 $\mathbb{F}$ 上 $n\times m$ 矩阵. 则矩阵 A 定义了域 $\mathbb{F}$ 上点空间 $(\mathbb{F}^m,O)$ 到点空间 $(\mathbb{F}^n,O)$ 内的映射 $x\to Ax=y$. 又 $x_0\in\mathbb{F}^m$ 为线性方程组 $Ax=q$ 的最小二乘解当且仅当 $Ax_0=y_0\in\mathbb{F}^n$ 为点子空间 $(A(\mathbb{F}^m),O)$ 中点 S 的坐标, 使得线段 $\overline{QS}$ 和点子空间 $(A(\mathbb{F}^m),O)$ 中任一线段正交. 这里 $x_0\in\mathbb{F}^m$ 不唯一, 但是 $y_0=Ax_0$ 唯一.

证 由引理 12.1.8, 记点 Q 为点空间 $(\mathbb{F}^n,O)$ 中坐标为 $q=(q_1,q_2,\cdots,q_n)'$ 的点. 则在点子空间 $(A(\mathbb{F}^m),O)$ 中唯一存在一点 S, 使得

$$\mathrm{d}\,(Q,A(\mathbb{F}^m))=|\overline{QS}|=\min_{T\in A(\mathbb{F}^m)}|\overline{QT}|.$$

今 $S\in A(\mathbb{F}^m)$, 所以 S 的坐标为 Ax_0, 其中 $x_0\in\mathbb{F}^m$. 注意 x_0 不一定唯一, 但是 Ax_0 唯一. 而

$$|\overrightarrow{QS}|=|Ax_0-q|,\quad |\overrightarrow{QT}|=|Ax-q|.$$

当 T 遍历点子空间 $(A(\mathbb{F}^m), O)$ 中点, 则 x 遍历点空间 $(\mathbb{F}^m, O)$ 中点. 而

$$|Ax_0 - q| = \min_{x \in \mathbb{F}^m} |Ax - q|.$$

这证明了 x_0 为 $Ax = q$ 的最小二乘解当且仅当点 Q 到点子空间 $(A(\mathbb{F}^m), O)$ 的垂线的垂足 S 有坐标 Ax_0. □

上面给出了最小二乘解一个很明确的几何意义. 下面给出最小二乘解的代数计算方法.

定理 12.1.11 设 A 为域 $\mathbb{F}$ 上 $n \times m$ 矩阵, x_0 为 $m \times 1$ 矩阵, q 为 $n \times 1$ 矩阵. 则 $x_0 \in \mathbb{F}^m$ 为域 $\mathbb{F}$ 上线性方程组 $Ax = q$ 的最小二乘解当且仅当 x_0 为相容的线性方程组 $A^*Ax = A^*q$ 的解.

证 先证线性方程组 $A^*Ax = A^*q$ 必有解. 事实上, 增广矩阵 (A^*A, A^*q) 的秩等于系数矩阵 A^*A 的秩. 这是因为

$$\begin{aligned}\operatorname{rank}(A^*A, A^*q) &= \operatorname{rank}(A^*(A, q)) \leqslant \min(\operatorname{rank}(A^*), \operatorname{rank}(A, q)) \\ &= \min(\operatorname{rank}(A), \operatorname{rank}(A, q)) = \operatorname{rank}(A) = \operatorname{rank}(A^*A).\end{aligned}$$

这证明了断言.

再证 x_0 为 $Ax = q$ 的最小二乘解当且仅当 x_0 为 $A^*Ax = A^*q$ 的解. 事实上, 由引理 12.1.10, x_0 为 $Ax = q$ 的最小二乘解当且仅当记 $Ax_0 = y_0$ 为点子空间 $(A(\mathbb{F}^m), O)$ 中点 S 的坐标, 则线段 $\overline{QS}$ 和点子空间 $(A(\mathbb{F}^m), O)$ 中任一线段正交, 即 $(\overrightarrow{QS}, \overrightarrow{OT}) = 0$, 其中点 Q 的坐标为 q, 点 S 的坐标为 Ax_0, 点 O 的坐标为 0, 点 T 的坐标为 Ax, $\forall x \in \mathbb{F}^m$. 即有 $(Ax_0 - q, Ax) = 0$, $\forall x \in \mathbb{F}^m$. 于是对 A 的共轭矩阵 $A^* = \overline{A}'$, 有 $(A^*(Ax_0 - q), x) = 0$, $\forall x \in \mathbb{F}^m$. 注意到 $Ax_0 - q \in \mathbb{F}^n$, $A^*(\mathbb{F}^n) \subset \mathbb{F}^m$, 于是 $A^*(Ax_0 - q) = 0$, $\forall x \in \mathbb{F}^m$. 这证明了 x_0 为 $Ax = q$ 的最小二乘解当且仅当 $A^*Ax_0 = A^*q$. □

§12.2 强广义逆矩阵

定义 12.2.1 给定 $n \times m$ 复矩阵 A, 则存在 n 阶酉方阵 U 及 m 阶酉方阵 V, 使得复矩阵 A 在酉相抵下有

$$A = U \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} V, \quad \Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_r), \quad \lambda_1 \geqslant \cdots \geqslant \lambda_r > 0, \tag{12.2.1}$$

其中 $\lambda_1^2, \lambda_2^2, \cdots, \lambda_r^2$ 为方阵 A^*A 的所有非零特征根. 则 $m \times n$ 复矩阵

$$V^{-1} \begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^{-1} \tag{12.2.2}$$

称为 $n\times m$ 复矩阵 A 的**强广义逆矩阵**, 记作 A^+.

显然, 当 A 为 n 阶非异方阵时, 强广义逆矩阵就是普通的逆矩阵. 从这个角度来说, 强广义逆矩阵是把逆方阵的概念推广到奇异方阵及一般矩阵的情形.

另一方面, 上面的定义依赖于酉方阵 U 及 V 的选取. 但是 U 和 V 并不唯一, 这样定义的强广义逆矩阵是否确定呢? 我们的回答是肯定的, 即有

定理 12.2.2 给定 $n\times m$ 复矩阵 A, 则 A 的强广义逆矩阵 A^+ 由 A 唯一确定. 即设存在 n 阶酉方阵 U 及 U_0, m 阶酉方阵 V 及 V_0, 使得复矩阵 A 在酉相抵下有

$$A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V=U_0\begin{pmatrix}\Lambda_0 & 0\\ 0 & 0\end{pmatrix}V_0, \tag{12.2.3}$$

其中

$$\Lambda=\mathrm{diag}\,(\lambda_1,\cdots,\lambda_r),\quad \lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>0,$$
$$\Lambda_0=\mathrm{diag}\,(\mu_1,\cdots,\mu_s),\quad \mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_s>0.$$

则 $s=r$, $\mu_i=\lambda_i$, $1\leqslant i\leqslant r$, 即 $\Lambda_0=\Lambda$. 又

$$A^+=V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-1}=V_0^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U_0^{-1}, \tag{12.2.4}$$

证 令

$$A^*A=V^{-1}\begin{pmatrix}\Lambda^2 & 0\\ 0 & 0\end{pmatrix}V=V_0^{-1}\begin{pmatrix}\Lambda_0^2 & 0\\ 0 & 0\end{pmatrix}V_0.$$

这证明了 $\lambda_1^2,\lambda_2^2,\cdots,\lambda_r^2$ 是 A^*A 的所有非零特征根, $\mu_1^2,\mu_2^2,\cdots,\mu_s^2$ 也是 A^*A 的所有非零特征根. 所以 $s=r$, 且 $\mu_i=\lambda_i$, $1\leqslant i\leqslant r$, 即 $\Lambda_0=\Lambda$. 因此式 (12.2.3) 为

$$\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}VV_0^{-1}=U^{-1}U_0\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}.$$

将 n 阶酉方阵 $U^{-1}U_0$ 按前 r 行, 前 r 列分块, m 阶酉方阵 VV_0^{-1} 也按前 r 行, 前 r 列分块为

$$U^{-1}U_0=\begin{pmatrix}U_{11} & U_{12}\\ U_{21} & U_{22}\end{pmatrix},\quad VV_0^{-1}=\begin{pmatrix}V_{11} & V_{12}\\ V_{21} & V_{22}\end{pmatrix}.$$

于是有

$$\begin{pmatrix}U_{11}\Lambda & 0\\ U_{21}\Lambda & 0\end{pmatrix}=\begin{pmatrix}\Lambda V_{11} & \Lambda V_{12}\\ 0 & 0\end{pmatrix}.$$

因此 $U_{11}\Lambda=\Lambda V_{11}$, $V_{12}=0$, $U_{21}=0$. 又由 U 和 V 都是酉方阵, 所以 $V_{21}=0$, $U_{12}=0$. 即 $U_0=U\mathrm{diag}\,(U_{11},U_{22})$, $V=\mathrm{diag}\,(V_{11},V_{22})V_0$, 又 $U_{11}\Lambda=\Lambda V_{11}$. 于是由 U_{11} 和 V_{11}

都是 r 阶酉方阵, 因此 $E_r = V_{11}V_{11}^* = \Lambda^{-1}U_{11}\Lambda^2U_{11}^*\Lambda^{-1}$, 即 $\Lambda^2U_{11} = U_{11}\Lambda^2$. 记 $\Lambda = \mathrm{diag}\,(\rho_1E_{r_1}), \cdots, \rho_tE_{r_t}), \rho_1 > \cdots > \rho_t > 0$. 于是

$$U_{11} = \mathrm{diag}\,(W_1^{(r_1)}, \cdots, W_t^{(r_t)}). \tag{12.2.5}$$

由 $U_{11}\Lambda = \Lambda V_{11}$, 于是 $V_{11} = U_{11}$. 因此, $U_0 = U\mathrm{diag}\,(U_{11}, U_{22})$, $V = \mathrm{diag}\,(U_{11}, V_{22})V_0$, 其中 U_{11} 由式 (12.2.5) 定义.

下面证强广义逆矩阵的表达式唯一, 即证式 (12.2.4) 成立. 事实上,

$$\begin{aligned} V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-1} &= V_0^{-1}\begin{pmatrix}U_{11}^{-1} & 0\\ 0 & V_{22}^{-1}\end{pmatrix}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}U_{11} & 0\\ 0 & U_{22}\end{pmatrix}U_0^{-1}\\ &= V_0^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U_0^{-1}. \end{aligned}$$ □

除非特殊要求, 在这节和下节中, 我们都用上面同样的符号.

定理 12.2.3 设 A 为 $n \times m$ 复矩阵, q 为 $n \times 1$ 复矩阵, x 为独立复变量构成的 $m \times 1$ 矩阵. 则线性方程组

$$Ax = q \tag{12.2.6}$$

的最小二乘解的通解为

$$x = A^+q + (E_m - A^+A)\xi, \quad \forall \xi \in \mathbb{C}^m. \tag{12.2.7}$$

所以模最小解为

$$x = A^+q. \tag{12.2.8}$$

证 由定理 12.1.11, 线性方程组 (12.2.6) 的最小二乘解为线性方程组

$$A^*Ax = A^*q \tag{12.2.9}$$

的解. 因此问题化为求线性方程组 (12.2.9) 的通解.

今存在 n 阶酉方阵 U 及 m 阶酉方阵 V, 使得

$$A = U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V, \quad \Lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_r), \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_r > 0.$$

于是

$$A^*A = V^{-1}\mathrm{diag}\,(\Lambda^2, 0)V, \quad A^*q = V^{-1}\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}U^{-1}q.$$

因此线性方程组 (12.2.9) 为

$$\mathrm{diag}\,(\Lambda^2, 0)Vx = \begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}U^{-1}q.$$

将 $y = Vx$ 按前 r 行分块为 $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, 将 $U^{-1}q$ 按前 r 行分块为 $U^{-1}q = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. 因此线性方程组 (12.2.9) 变为 $\Lambda^2 y_1 = \Lambda v_1$, 即 $y_1 = \Lambda^{-1} v_1$. 所以通解为 $y = \begin{pmatrix} \Lambda^{-1} v_1 \\ y_2 \end{pmatrix}$, 即

$$x = V^{-1}y = V^{-1}\begin{pmatrix} \Lambda^{-1} v_1 \\ y_2 \end{pmatrix} = V^{-1}\begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + V^{-1}\begin{pmatrix} 0 \\ y_2 \end{pmatrix}.$$

因此,

$$x = V^{-1}\begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^{-1}q + V^{-1}\begin{pmatrix} 0 & 0 \\ 0 & E \end{pmatrix} y = A^+q + V^{-1}\begin{pmatrix} 0 & 0 \\ 0 & E \end{pmatrix} Vx.$$

另一方面,

$$E_m - A^+A = E_m - V^{-1}\begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^{-1}U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} V.$$

因此,

$$E_m - A^+A = E_m - V^{-1}\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} V = V^{-1}\begin{pmatrix} 0 & 0 \\ 0 & E_{m-r} \end{pmatrix} V.$$

这证明了通解为 $x = A^+q + (E - A^+A)\xi$, $\forall \xi \in \mathbb{C}^m$.

下面计算通解的模. 今

$$(A^+q, (E - A^+A)\xi) = (A^+q)'\overline{(E - A^+A)\xi} = q'(A^+)'(E - \overline{A^+A})\overline{\xi}.$$

而

$$(A^+)'(E - \overline{A^+A}) = \overline{U}\begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix} \overline{V}V'\begin{pmatrix} 0 & 0 \\ 0 & E \end{pmatrix} \overline{V} = 0.$$

所以 $(A^+q, (E - A^+A)\xi) = 0$. 这证明了

$$(x, x) = (A^+q + (E - A^+A)\xi, A^+q + (E - A^+A)\xi) = |A^+q|^2 + |(E - A^+A)\xi|^2 \geqslant |A^+q|^2,$$

$\xi \in \mathbb{C}^m$. 所以 $x = A^+q$ 为模最小解. □

最后, 我们来证明当线性方程组 $Ax = q$ 相容时, 最小二乘解就是通常意义下的解. 由此可见, 最小二乘解是不相容线性方程组 $Ax = q$ 的最佳近似解.

定理 12.2.4　线性方程组 $Ax = q$ 相容当且仅当 $AA^+q = q$. 这时最小二乘解就是通常的解.

证　今线性方程组 $Ax = q$ 可写为 $U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} Vx = q$, 即 $\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} Vx = U^{-1}q$. 记 $Vx = y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ 为按前 r 行分块. 将 $U^{-1}q$ 也按前 r 行分块为 $U^{-1}q = \begin{pmatrix} u \\ v \end{pmatrix}$. 于是

线性方程组 $Ax=q$ 可改写为 $y_1=\Lambda^{-1}u$, $0=v$, y_2 任意. 所以线性方程组 $Ax=q$ 有解当且仅当 $v=0$.

今

$$E_n-AA^+=E_n-U\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}U^*=U\begin{pmatrix}0 & 0\\ 0 & E_{n-r}\end{pmatrix}U^*.$$

所以

$$(E_n-AA^+)q=U\begin{pmatrix}0 & 0\\ 0 & E_{n-r}\end{pmatrix}U^{-1}q=U\begin{pmatrix}0\\ v\end{pmatrix}.$$

因此 $v=0$ 当且仅当 $(E_n-AA^+)q=0$, 即 $A(A^+q)=q$. 这证明了线性方程组 $Ax=q$ 相容当且仅当 $AA^+q=q$.

下面证在相容时, 即在条件 $AA^+q=q$ 时, 最小二乘解就是通常的解. 事实上, 线性方程组 $Ax=q$ 有解 $x=A^+q$. 且通解为

$$x=V^{-1}y=V^{-1}\begin{pmatrix}y_1\\ y_2\end{pmatrix}=V^{-1}\begin{pmatrix}\Lambda^{-1}u\\ y_2\end{pmatrix}=V^{-1}\begin{pmatrix}\Lambda^{-1}u\\ 0\end{pmatrix}+V^{-1}\begin{pmatrix}0\\ y_2\end{pmatrix}.$$

即

$$x=V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}+V^{-1}\begin{pmatrix}0 & 0\\ 0 & E\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}=A^+q+(E-A^+A)\xi,\quad \forall\xi\in\mathbb{C}^m.$$

由定理 12.2.3, 便证明了线性方程组 $Ax=q$ 的通解为最小二乘解. □

下面给出强广义逆矩阵的两种等价定义

定理 12.2.5 设 A 为 $n\times m$ 复矩阵, X 为由 mn 个独立复变量构成的 $m\times n$ 矩阵. 则矩阵方程组

$$AXA=A, \tag{a}$$

$$XAX=X, \tag{b}$$

$$(AX)^*=AX, \tag{c}$$

$$(XA)^*=XA \tag{d}$$

有唯一解, 称为 $n\times m$ 复矩阵 A 的 **强广义逆矩阵**, 记作 A^+.

证 今 $A=UBV$, 其中

$$B=\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix},$$

符号 Λ 的意义同上. 于是方程组变为

$$B(VXU)B=B,\quad (VXU)B(VXU)=(VXU),$$
$$(VXU)^*B^*=B(VXU),\quad B^*(VXU)^*=(VXU)B.$$

将 VXU 按前 r 行及前 r 列分成四块, 记作

$$VXU = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}.$$

于是上面矩阵方程组可改写为 $\Lambda P\Lambda = \Lambda$, $P\Lambda P = P$, $P\Lambda Q = Q$, $R\Lambda P = R$, $R\Lambda Q = S$, $P^*\Lambda = \Lambda P$, $\Lambda Q = 0$, $Q^*\Lambda = 0$, $\Lambda P^* = P\Lambda$, $\Lambda R^* = 0$, $R\Lambda = 0$. 此即 $Q = 0$, $S = 0$, $R = 0$, $P = \Lambda^{-1}$. 所以解唯一存在, 它是

$$X = V^{-1}(VXU)U^{-1} = V^{-1}\begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix}U^{-1} = A^+. \qquad \square$$

上面定理给出可以用矩阵方程组的解来定义强广义逆矩阵. 下面用几何语言, 给出强广义逆矩阵的几何定义.

引理 12.2.6　给定 m 维复 Euclid 空间 $\mathfrak{L}$ 和 n 维复 Euclid 空间 $\mathfrak{R}$. 设 $\mathscr{A}$ 为 $\mathfrak{L}$ 到 $\mathfrak{R}$ 内的线性映射, 则唯一存在 $\mathfrak{R}$ 到 $\mathfrak{L}$ 内的线性映射 $\mathscr{A}^*$, 它有

$$(\mathscr{A}(u), v)_{\mathfrak{R}} = (u, \mathscr{A}^*(v))_{\mathfrak{L}}, \quad \forall\, u \in \mathfrak{L}, \quad v \in \mathfrak{R}. \tag{12.2.10}$$

这时 $\mathscr{A}^*$ 称为线性映射 $\mathscr{A}$ 的**共轭映射**.

证　在复 Euclid 空间 $\mathfrak{L}$ 及 $\mathfrak{R}$ 中分别取标准正交基 $e_1, e_2, \cdots, e_m$ 及 $f_1, f_2, \cdots, f_n$. 于是

$$\mathscr{A}(e_i) = \sum_{j=1}^{n} a_{ji} f_j, \quad i = 1, 2, \cdots, m.$$

它决定了一个 $n \times m$ 矩阵 $A = (a_{ij})$, 称为 $\mathscr{A}$ 的矩阵表示. 若 $\mathscr{A}^*$ 存在, 则有

$$\mathscr{A}^*(f_j) = \sum_{i=1}^{m} b_{ij} e_i, \quad j = 1, 2, \cdots, n.$$

由定义条件, $(\mathscr{A}(e_i), f_j) = (e_j, \mathscr{A}^*(f_j))$, 即

$$(\sum_{k=1}^{n} a_{ki} f_k, f_j) = (e_i, \sum_{p=1}^{m} b_{pj} e_p).$$

但是复 Euclid 空间的内积为正定 Hermite 双线性函数, 所以有

$$a_{ji} = \overline{b_{ij}}, \quad i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, n.$$

因此当 $\mathscr{A}$ 的矩阵表示为 A 时, $\mathscr{A}^*$ 的矩阵表示必为 $A^* = \overline{A}'$. 这也证明了对任一线性映射 $\mathscr{A}$: $\mathfrak{L} \to \mathfrak{R}$, 则唯一存在它的共轭映射 $\mathscr{A}^*$: $\mathfrak{R} \to \mathfrak{L}$. $\square$

引理 12.2.7 复 Euclid 空间 $\mathfrak{L}$ 上线性变换 $\mathscr{P}$: $\mathfrak{L}\to\mathfrak{L}$ 为**正交投影**当且仅当 $\mathscr{P}$ 在象空间 $\mathfrak{L}_1=\mathscr{P}(\mathfrak{L})$ 上为恒等变换 id, 在象空间 $\mathfrak{L}_1$ 的正交补 $\mathfrak{L}_1^{\perp}$ 上为零变换; 当且仅当

$$\mathscr{P}^2=\mathscr{P},\quad \mathscr{P}^*=\mathscr{P}. \tag{12.2.11}$$

证 今 $\mathscr{P}$ 为正交投影, 即存在子空间 $\mathfrak{L}_1$. 对直接和 $\mathfrak{L}=\mathfrak{L}_1\oplus\mathfrak{L}_1^{\perp}$, 任取 $\alpha\in\mathfrak{L}$, 则有唯一分解 $\alpha=\beta+\gamma$, 其中 $\beta\in\mathfrak{L}_1$, $\gamma\in\mathfrak{L}_1^{\perp}$. 而 $\mathscr{P}$: $\alpha\to\beta$. 这证明了 $\mathscr{P}(\beta)=\beta$, $\mathscr{P}(\gamma)=0$. 即 $\mathscr{P}$ 在 $\mathfrak{L}_1$ 上为恒等变换, 在 $\mathfrak{L}_1^{\perp}$ 上为零变换. 反之, 则

$$\mathscr{P}(\alpha)=\mathscr{P}(\beta+\gamma)=\mathscr{P}(\beta)+\mathscr{P}(\gamma)=\beta,\quad \forall\,\alpha\in\mathfrak{L}.$$

所以 $\mathscr{P}$ 为正交投影.

再若 $\mathscr{P}$ 为正交投影, 自然

$$\mathscr{P}^2(\alpha)=(\mathscr{P}\circ\mathscr{P})(\alpha)=\mathscr{P}(\mathscr{P}(\alpha))=\mathscr{P}(\beta)=\beta=\mathscr{P}(\alpha),\quad \forall\,\alpha\in\mathfrak{L},$$

其中 $\alpha=\beta+\gamma$, $\beta\in\mathfrak{L}_1$, $\gamma\in\mathfrak{L}_1^{\perp}$. 于是 $\mathscr{P}^2=\mathscr{P}$. 又任取 $\alpha=\beta+\gamma$, $\xi=\eta+\zeta$, 其中 $\alpha,\ \xi\in\mathfrak{L}$, $\beta,\ \eta\in\mathfrak{L}_1$, $\gamma,\ \zeta\in\mathfrak{L}_1^{\perp}$. 于是由共轭变换的定义, 有 $(\mathscr{P}(\alpha),\xi)=(\alpha,\mathscr{P}^*(\xi))$. 因此记 $\mathscr{P}^*(\xi)=\lambda+\mu$, 其中 $\lambda\in\mathfrak{L}_1$, $\mu\in\mathfrak{L}_1^{\perp}$. 则有 $(\beta,\xi)=(\alpha,\lambda+\mu)$, 于是 $(\beta,\eta)=(\beta,\lambda)+(\gamma,\mu)$. 由 α 任取, 我们取 $\alpha\in\mathfrak{L}_1$, 则 $\alpha=\beta$, $\gamma=0$. 所以 $(\beta,\eta-\lambda)=0$. 这证明了 $\eta=\lambda$. 因此对一般的 $\alpha\in\mathfrak{L}_1$, 则可任取 $\gamma\in\mathfrak{L}_1^{\perp}$, 使得 $(\gamma,\mu)=0$, 其中 $\gamma,\ \mu\in\mathfrak{L}_1^{\perp}$. 这证明了 $\mu=0$. 所以 $\mathscr{P}^*(\xi)=\mathscr{P}^*(\eta+\xi)=\lambda$. 这证明了 $\mathscr{P}^*(\alpha)=\beta$, $\forall\,\alpha\in\mathfrak{L}$, 即 $\mathscr{P}^*=\mathscr{P}$.

反之, 若 $\mathfrak{L}$ 上线性变换 $\mathscr{P}$ 有 $\mathscr{P}^2=\mathscr{P}$, $\mathscr{P}^*=\mathscr{P}$. 由共轭线性变换的定义可知, 在 $\mathfrak{L}$ 中存在标准正交基 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n$, 使得 $\mathscr{P}$ 的方阵表示 P 有 $P^2=P$, $P^*=P^*=P$. 所以 P 酉相似于对角形 $\mathrm{diag}\,(\lambda_1,\cdots,\lambda_n)$, 且有 $\lambda_j^2=\lambda_j$, $1\leqslant j\leqslant n$. 即 $\lambda_j=0$ 或 $\lambda_j=1$. 因此, 证明了在 $\mathfrak{L}$ 中存在一组标准正交基 $\beta_1,\ \beta_2,\ \cdots,\ \beta_n$, 使得 $\mathscr{P}$ 的方阵表示为 $\begin{pmatrix}E_r & 0\\ 0 & 0\end{pmatrix}$, 即 $\mathscr{P}(e_i)=e_i$, $1\leqslant i\leqslant r$; $\mathscr{P}(e_j)=0$, $r+1\leqslant j\leqslant n$. 记以 $e_1,\cdots,e_r$ 为基的子空间为 $\mathfrak{L}_1$, 则 $\mathscr{P}$ 在 $\mathfrak{L}_1$ 上为恒等映射, 在 $\mathfrak{L}_1^{\perp}$ 上为零映射. 所以 $\mathscr{P}$ 为正交投影. □

设 $\mathscr{A}$ 为 m 维复 Euclid 空间 $\mathfrak{L}$ 到 n 维复 Euclid 空间 $\mathfrak{R}$ 内的线性映射, 记 $\mathscr{P}_{\mathscr{A}}$ 为 $\mathfrak{R}$ 中关于 $\mathscr{A}(\mathfrak{L})$ 的正交投影. 于是有

引理 12.2.8 在 m 维复 Euclid 空间 $\mathfrak{L}$ 中任取标准正交基 $e_1,\ e_2,\ \cdots,\ e_m$, 在 n 维复 Euclid 空间 $\mathfrak{R}$ 中任取标准正交基 $f_1,\ f_2,\ \cdots,\ f_n$. 设 $\mathfrak{L}$ 到 $\mathfrak{R}$ 内的线性映射 $\mathscr{A}$ 有矩阵表示

$$A=(a_{ij})=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V, \tag{12.2.12}$$

其中 U 为 n 阶酉方阵, V 为 m 阶酉方阵,

$$\Lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_r), \qquad \lambda_1 \geqslant \cdots \geqslant \lambda_r > 0.$$

于是线性映射 $\mathscr{A}^*$: $\mathfrak{R} \to \mathfrak{L}$; $\mathscr{P}_{\mathscr{A}}$: $\mathfrak{R} \to \mathfrak{R}$; $\mathscr{P}_{\mathscr{A}^*}$: $\mathfrak{L} \to \mathfrak{L}$ 分别有矩阵表示

$$A^* = V^* \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} U^*, \quad P_{\mathscr{A}} = U \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} U^*, \quad P_{\mathscr{A}^*} = V^* \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} V. \qquad (12.2.13)$$

证 今存在 n 阶酉方阵 U 及 m 阶酉方阵 V, 使得

$$A = U \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} V, \quad \Lambda = \mathrm{diag}\,(\lambda_1 \cdots, \lambda_r), \quad \lambda_1 \geqslant \cdots \geqslant \lambda_r > 0.$$

记 $U = (u_{ij})$, $V = (v_{ij})$. 作标准正交基到标准正交基的基变换

$$e_i' = \sum_{j=1}^{m} \overline{v_{ij}} e_j, \quad 1 \leqslant i \leqslant m; \quad f_i' = \sum_{j=1}^{n} u_{ji} f_j, \quad 1 \leqslant i \leqslant n,$$

则有

$$f_i = \sum_{j=1}^{n} \overline{u_{ij}} f_j', \quad 1 \leqslant i \leqslant n,$$

且

$$\mathscr{A}(e_i') = \sum_{j=1}^{m} \overline{v_{ij}} \mathscr{A}(e_j) = \sum_{j=1}^{m} \sum_{k=1}^{n} \overline{v_{ij}} a_{kj} f_k = \sum_{j=1}^{m} \sum_{k=1}^{n} \sum_{p=1}^{m} \overline{v_{ij}} a_{kj} \overline{u_{kp}} f_p', \quad 1 \leqslant i \leqslant m.$$

这证明了线性映射 $\mathscr{A}$: $\mathfrak{L} \to \mathfrak{R}$ 在 $\mathfrak{L}$ 的标准正交基 $e_1', e_2', \cdots, e_m'$ 及 $\mathfrak{R}$ 的标准正交基 $f_1', f_2', \cdots, f_n'$ 下的矩阵表示为 $C = (c_{ij})$, 其中 $\mathscr{A}(e_i') = \sum\limits_{p=1}^{n} c_{pi} f_p'$, $1 \leqslant i \leqslant m$, 有 $c_{pi} = \sum\limits_{j=1}^{m} \sum\limits_{k=1}^{n} \overline{v_{ij}} a_{kj} \overline{u_{kp}}$, 即

$$C = U^* A V^* = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}.$$

此即

$$\mathscr{A}(e_i') = \begin{cases} \lambda_i f_i', & 1 \leqslant i \leqslant r, \\ 0, & r+1 \leqslant j \leqslant m. \end{cases}$$

于是 $\mathscr{A}(\mathfrak{L})$ 是以 $f_1', f_2', \cdots, f_r'$ 为标准正交基的子空间. 由引理 12.2.7, 在 $\mathfrak{R}$ 的标准正交基 $f_1', f_2', \cdots, f_n'$ 下, $\mathscr{P}_{\mathscr{A}}$ 有方阵表示

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

于是在 $\mathfrak{R}$ 的标准正交基 $f_1, f_2, \cdots, f_n$ 下, $\mathscr{P}_{\mathscr{A}}$ 有方阵表示

$$P_{\mathscr{A}} = U\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}U^{-1}.$$

同理, $\mathscr{A}^*$ 有矩阵表示

$$A^* = V^{-1}\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}U^{-1},$$

于是 $\mathscr{P}_{\mathscr{A}^*}$ 有矩阵表示

$$P_{\mathscr{A}^*} = V^{-1}\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}V.$$

□

定义 12.2.9 给定的 m 维复 Euclid 空间 $\mathfrak{L}$ 到 n 维复 Euclid 空间 $\mathfrak{R}$ 内的线性映射 $\mathscr{A}$ 及 $\mathfrak{R}$ 到 $\mathfrak{L}$ 内的线性映射 $\mathscr{L}$. 如果有

(1) $\operatorname{rank}(\mathscr{A}) = \operatorname{rank}(\mathscr{L})$;

(2) $\mathscr{A}\circ\mathscr{L} = \mathscr{P}_{\mathscr{A}}, \mathscr{L}\circ\mathscr{A} = \mathscr{P}_{\mathscr{A}^*}$,

则 $\mathscr{L}$ 称为 $\mathscr{A}$ 的**强广义逆映射**, 改记 $\mathscr{L}$ 为 $\mathscr{A}^+$.

引理 12.2.10 符号同引理 12.2.8. 记 $\mathscr{L}$ 有矩阵表示 B, 则 $\mathscr{L}$ 为 $\mathscr{A}$ 的强广义逆映射当且仅当 $B = A^+$.

证 设 $\mathscr{L}$ 为 $\mathscr{A}$ 的强广义逆映射, 即 $\operatorname{rank}(\mathscr{L}) = \operatorname{rank}(\mathscr{A})$, 且 $\mathscr{A}\circ\mathscr{L} = \mathscr{P}_{\mathscr{A}}$, $\mathscr{L}\circ\mathscr{A} = \mathscr{P}_{\mathscr{A}^*}$. 由引理 12.2.8, 所以矩阵关系

$$\operatorname{rank}(A) = \operatorname{rank}(B), \quad AB = P_{\mathscr{A}}, \quad BA = P_{\mathscr{A}^*} \tag{12.2.14}$$

成立当且仅当 $\operatorname{rank}(B) = r$,

$$U\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}VB = U\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}U^*, \quad BU\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}V = V^*\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}V. \tag{12.2.15}$$

记

$$B_1 = VBU = \begin{pmatrix} C^{(r)} & D \\ F & G \end{pmatrix},$$

则式 (12.2.14) 成立当且仅当 $\operatorname{rank}(B_1) = r$,

$$\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & D \\ F & G \end{pmatrix} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} C & D \\ F & G \end{pmatrix}\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

此即 $C=\Lambda^{-1}$, $D=0$, $F=0$. 于是

$$B_1=VBU=\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & G\end{pmatrix}.$$

由于 $\operatorname{rank}(B_1)=r$, 所以 $G=0$. 这证明了式 (12.2.14) 成立当且仅当

$$B=V^{-1}\begin{pmatrix}\Lambda^{-1} & 0\\ 0 & 0\end{pmatrix}U^{-1}=A^{+}. \qquad \square$$

注意 将上面的理论全部限制在实的情形, 即改复 Euclid 空间为实 Euclid 空间, 酉方阵为实正交方阵, 则全部结论也都成立.

习 题 12.2

下面只在复的情形计算.

12.2.1 设 A 为 $n\times m$ 矩阵. 试证: $(A^{+})^{+}=A$, $(A^{*})^{+}=(\overline{A}^{+})'$, $A^{*}(A^{*}AA^{*})^{+}A^{*}=A^{+}$. 又当 A 为 n 阶非异方阵, 则 $A^{+}=A^{-1}$.

12.2.2 设 A 为 $n\times m$ 矩阵, P 为 n 阶酉方阵, Q 为 m 阶酉方阵, 则 $(PAQ)^{+}=Q^{-1}A^{+}P^{-1}$.

12.2.3 设 A 为 $n\times m$ 矩阵, B 为 $m\times q$ 矩阵, 且 $AB=0$. 试证: $A(B^{*})^{+}=0$, $(A^{*})^{+}B=0$.

12.2.4 试举反例来说明: 当 A 为 $n\times m$ 矩阵, B 为 $m\times q$ 矩阵, 则一般来说, $(AB)^{+}=B^{+}A^{+}$ 不成立.

12.2.5 试给出 $n\times 1$ 矩阵及 $1\times m$ 矩阵的强广义逆矩阵的明显表达式.

12.2.6 设 A 为 $n\times m$ 矩阵. 试证: 存在 $n\times r$ 矩阵 B 及 $r\times m$ 矩阵 C, 使得 $\operatorname{rank}(B)=\operatorname{rank}(C)=r$, 其中 $r=\operatorname{rank}(A)$, 又 $A=BC$, 且 $A^{+}=C^{*}(CC^{*})^{-1}(B^{*}B)^{-1}B^{*}=C^{+}B^{+}$.

§12.3 广义逆矩阵

定义 12.3.1 设 A 为 $n\times m$ 复矩阵, X 为由 mn 个独立复变量构成的 $m\times n$ 矩阵. 则矩阵方程

$$AXA=A \tag{a}$$

的通解称为 $n\times m$ 复矩阵 A 的**广义逆矩阵**, 记作 A^{-}.

定理 12.3.2 设 A 为 $n\times m$ 复矩阵, 且

$$A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V,\quad \Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_r),\quad \lambda_1\geqslant\cdots\geqslant\lambda_r>0, \tag{12.3.1}$$

其中 U 及 V 分别为 n 阶及 m 阶酉方阵. 则 A 的广义逆矩阵为

$$A^{-}=V^{*}\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}U^{*}=A^{+}+V^{*}\begin{pmatrix}0 & X_{12}\\ X_{21} & X_{22}\end{pmatrix}U^{*}, \tag{12.3.2}$$

其中 X_{12}, X_{21} 和 X_{22} 分别由 $r(m-r)$, $r(n-r)$ 和 $(m-r)(n-r)$ 个独立复参数构成. 所以共有 $mn-r^2$ 个独立复参数. 因此, 广义逆矩阵不唯一. 实际上它们全体构成一个集合, 而 A^- 只表示这个集合中的任意一个元素.

证 令 $A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V$, 记

$$VXU=\begin{pmatrix}X_{11} & X_{12}\\ X_{21} & X_{22}\end{pmatrix},$$

则矩阵方程 (a) 等价于

$$\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}VXU\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}.$$

则 $X_{11}=\Lambda^{-1}$. 所以通解为 (12.3.2). □

注意 矩阵方程 $AXA=A$ 的解不唯一, 所以 A^- 是泛指其中一个解.

定理 12.3.3 设 U 和 $U_1\in \mathrm{U}(n)$, V, $V_1\in \mathrm{U}(m)$, 且

$$A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V=U_1\begin{pmatrix}\Lambda_1 & 0\\ 0 & 0\end{pmatrix}V_1, \tag{12.3.3}$$

其中 $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$, $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>0$, $\Lambda_1=\mathrm{diag}(\mu_1,\cdots,\mu_s)$, $\mu_1\geqslant\mu_2\geqslant\cdots\geqslant\mu_s>0$, 则 $s=r$, $\mu_i=\lambda_i$, $1\leqslant i\leqslant r$, 即 $\Lambda_1=\Lambda$. 又

$$V^{-1}\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}U^{-1}=V_1^{-1}\begin{pmatrix}\Lambda^{-1} & Y_{12}\\ Y_{21} & Y_{22}\end{pmatrix}U_1^{-1}, \tag{12.3.4}$$

其中

$$Y_{12}=U_{11}^{-1}X_{12}U_{22}^{-1},\quad Y_{21}=V_{22}^{-1}X_{21}U_{11},\quad Y_{22}=V_{22}^{-1}X_{22}U_{22},$$

U_{11}, U_{22}, V_{22} 为酉方阵.

证 令

$$A^*A=V^{-1}\begin{pmatrix}\Lambda^2 & 0\\ 0 & 0\end{pmatrix}V=V_1^{-1}\begin{pmatrix}\Lambda_1^2 & 0\\ 0 & 0\end{pmatrix}V_1.$$

这证明了 $\lambda_1^2,\cdots,\lambda_r^2$ 为 A^*A 的所有非零特征根, $\mu_1^2,\cdots,\mu_s^2$ 也是 A^*A 的所有非零特征根. 因此, $s=r$, 且 $\mu_i=\lambda_i$, $1\leqslant i\leqslant r$, 即 $\Lambda_1=\Lambda$. 所以式 (12.3.3) 为

$$\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}VV_1^{-1}=U^{-1}U_1\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}.$$

将 n 阶酉方阵 $U^{-1}U_1$ 按前 r 行, 前 r 列分块, m 阶酉方阵 VV_1^{-1} 也按前 r 行, 前 r 列分块为

$$U^{-1}U_1 = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \quad VV_1^{-1} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

于是有

$$\begin{pmatrix} U_{11}\Lambda & 0 \\ U_{21}\Lambda & 0 \end{pmatrix} = \begin{pmatrix} \Lambda V_{11} & \Lambda V_{12} \\ 0 & 0 \end{pmatrix},$$

因此 $U_{11}\Lambda = \Lambda V_{11}$, $V_{12} = 0$, $U_{21} = 0$. 又由 U 和 V 都是酉方阵, 所以 $V_{21} = 0$, $U_{12} = 0$. 即 $U_1 = U\mathrm{diag}\,(U_{11}, U_{22})$, $V = \mathrm{diag}\,(V_{11}, V_{22})V_1$. 又 $U_{11}\Lambda = \Lambda V_{11}$, U_{11} 和 V_{11} 都是 r 阶酉方阵. 因此, $E_r = V_{11}V_{11}^* = \Lambda^{-1}U_{11}\Lambda^2U_{11}^*\Lambda^{-1}$. 即 $\Lambda^2U_{11} = U_{11}\Lambda^2$. 记 $\Lambda = \mathrm{diag}\,(\rho_1E_{r_1}), \cdots, \rho_tE_{r_t})$, $\rho_1 > \cdots > \rho_t > 0$. 于是

$$U_{11} = \mathrm{diag}\,(W_1^{(r_1)}, \cdots, W_t^{(r_t)}). \tag{12.3.5}$$

由 $U_{11}\Lambda = \Lambda V_{11}$, 于是 $V_{11} = U_{11}$. 因此, $U_1 = U\mathrm{diag}\,(U_{11}, U_{22})$, $V = \mathrm{diag}\,(U_{11}, V_{22})V_1$, 其中 U_{11} 由式 (12.3.5) 定义.

下面计算 (12.3.4). 今由 $U_1 = U\mathrm{diag}\,(U_{11}, U_{22})$, $V = \mathrm{diag}\,(U_{11}, V_{22})V_1$, 有

$$Y_{12} = U_{11}^{-1}X_{12}U_{22}, \quad Y_{21} = X_{21}U_{11}, \quad Y_{22} = V_{22}^{-1}X_{22}U_{22}.$$ □

定义 12.3.4 设 A 为 $n \times m$ 复矩阵, X 为由 mn 个独立复变量构成的 $m \times n$ 矩阵, 适合矩阵方程 (a) 以及余下三个方程 (b), (c) 和 (d) 中若干个, 组成的矩阵方程组的通解称为**其他类型的广义逆矩阵**.

例如矩阵方程 (a), (b) 和 (c) 组成的矩阵方程组的通解称为 (1,2,3) 型的广义逆矩阵, 记为 $A^{(123)}$. 因此, $A^- = A^{(1)}$, $A^+ = A^{(1234)}$.

注意 没有后三个矩阵方程时, 得到最一般的广义逆矩阵 $A^- = A^{(1)}$. 包含后三个矩阵方程时, 得到强广义逆矩阵 $A^+ = A^{(1234)}$. 其他情形, 共有六种广义逆矩阵. 这六种广义逆矩阵对解各种矩阵方程和讨论各种矩阵不等式, 以及在应用数学中有很多用处.

定理 12.3.5 设 A 为 $n \times m$ 复矩阵, β 为 $n \times 1$ 复矩阵, x 为由 m 个独立未知数构成的 $m \times 1$ 复矩阵. 则线性方程组

$$Ax = \beta \tag{12.3.6}$$

相容的必要且充分条件为

$$AA^-\beta = \beta. \tag{12.3.7}$$

这时通解为

$$x = A^-\beta + (I - A^-A)\xi, \quad \forall \xi \in \mathbb{C}^m. \tag{12.3.8}$$

证 用定理 12.3.2 的符号, 于是 $A=U\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}V$. 因此 $Ax=\beta$ 可改变为

$$\begin{pmatrix}\Lambda & 0\\ 0 & 0\end{pmatrix}Vx=U^*\beta.$$

记 $Vx=y=\begin{pmatrix}u\\ v\end{pmatrix}$, 其中 u 为 $r\times 1$ 矩阵. 记 $U^*\beta=\begin{pmatrix}\sigma\\ \tau\end{pmatrix}$, $\sigma\in\mathbb{C}^r$, $\tau\in\mathbb{C}^{m-r}$. 于是有 $\Lambda u=\sigma$, $\tau=0$. 所以通解为

$$x=V^{-1}y=V^{-1}\begin{pmatrix}u\\ v\end{pmatrix}=V^{-1}\begin{pmatrix}\Lambda^{-1}\sigma\\ v\end{pmatrix}.$$

又 $Ax=\beta$ 相容的必要且充分条件为 $\tau=0$, 即 $\beta=U\begin{pmatrix}\sigma\\ 0\end{pmatrix}$.

另一方面, 由定理 12.3.2 有

$$A^-=V^{-1}\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}U^{-1}.$$

于是

$$E_n-A^-A=V^{-1}\begin{pmatrix}0 & 0\\ -X_{21}\Lambda & E\end{pmatrix}V,$$

$$I-AA^-=U\begin{pmatrix}0 & -\Lambda X_{12}\\ 0 & E\end{pmatrix}U^{-1}.$$

因此,

$$(I-AA^-)\beta=U\begin{pmatrix}0 & -\Lambda X_{12}\\ 0 & E\end{pmatrix}\begin{pmatrix}\sigma\\ \tau\end{pmatrix}=U\begin{pmatrix}-\Lambda X_{12}\tau\\ \tau\end{pmatrix}.$$

所以 $\tau=0$ 当且仅当 $(I-AA^-)\beta=0$, 即 $AA^-\beta=\beta$. 又这时通解为

$$\begin{aligned}x&=V^*\begin{pmatrix}\Lambda^{-1}\sigma\\ v\end{pmatrix}=V^*\begin{pmatrix}\Lambda^{-1}\sigma\\ 0\end{pmatrix}+V^*\begin{pmatrix}0\\ v\end{pmatrix}\\ &=V^*\begin{pmatrix}\Lambda^{-1} & X_{12}\\ X_{21} & X_{22}\end{pmatrix}\begin{pmatrix}\sigma\\ \tau\end{pmatrix}-V^*\begin{pmatrix}0\\ X_{21}\sigma\end{pmatrix}+V^*\begin{pmatrix}0\\ v\end{pmatrix}\\ &=A^-\beta+(I-A^-A)\xi,\forall\xi\in\mathbb{C}^m.\end{aligned}$$

□

定理 12.3.6 线性函数组 $f(x)=Ax-\beta$ 的最小二乘通解为

$$x=(A^*A)^-A^*\beta+\big(E_n-(A^*A)^-(A^*A)\big)\xi,\quad\forall\xi\in\mathbb{C}^m.\tag{12.3.9}$$

证　由引理 12.1.11, $f(x)=Ax-\beta$ 的最小二乘通解为相容线性方程组 $A^*Ax=A^*\beta$ 的通解. 由定理 12.3.5, 便证明了定理. □

广义逆矩阵可用来求某些矩阵方程的通解. 为了给出证明, 我们先引进 Kronecker 乘积的定义.

定义 12.3.7　给定 $n\times m$ 和 $p\times q$ 矩阵 A 和 B. $np\times mq$ 矩阵

$$\begin{pmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{pmatrix} \tag{12.3.10}$$

称为矩阵 A 和 B 的Kronecker 乘积, 记作 $A\otimes B$.

由定义出发, 不难证明

(1)　$(A\otimes B)'=A'\otimes B'$;

(2)　$\overline{(A\otimes B)}=\overline{A}\otimes\overline{B}$;

(3)　$(A_1+A_2)\otimes B=A_1\otimes B+A_2\otimes B$;

(4)　$A\otimes(B_1+B_2)=A\otimes B_1+A\otimes B_2$;

(5)　$(\lambda A)\otimes(\mu B)=\lambda\mu(A\otimes B), \forall\,\lambda,\,\mu\in\mathbb{C}$;

(6)　$(A_1\otimes B_1)(A_2\otimes B_2)=(A_1A_2)\otimes(B_1B_2)$;

(7)　$(A\otimes B)^+=A^+\otimes B^+$;

(8)　$n\times m$ 矩阵 A 的任一广义逆矩阵 A^- 和 $u\times v$ 矩阵 B 的任一广义逆矩阵 B^- 的 Kronecker 乘积 $A^-\otimes B^-$ 必为 $(A\otimes B)$ 的广义逆矩阵. 反之不一定.

定理 12.3.8　给定 $n\times m$, $m\times p$, $p\times q$, $n\times q$ 矩阵 A, X, B, C, 其中 A, B, C 为常数矩阵, X 为由 mp 个独立未知数构成的矩阵. 则矩阵方程

$$AXB=C \tag{12.3.11}$$

有解当且仅当

$$AA^-CB^-B=C. \tag{12.3.12}$$

这时通解为

$$X=A^-CB^-+Z-A^-AZBB^-, \tag{12.3.13}$$

其中 Z 取任一 $m\times p$ 复矩阵.

证　记 $A=(a_{ij})$, $X=(x_{jk})$, $B=(b_{ks})$, $C=(c_{is})$. 于是矩阵方程可写为

$$\sum_{j=1}^{m}\sum_{k=1}^{p}a_{ij}x_{jk}b_{ks}=c_{is},\quad 1\leqslant i\leqslant n,\quad 1\leqslant s\leqslant q. \tag{12.3.14}$$

将 X 和 C 中元素按行依次排为 $mp \times 1$ 和 $nq \times 1$ 矩阵

$$x = (x_{11}, \cdots, x_{1p}, \cdots, x_{m1}, \cdots, x_{mp})', \tag{12.3.15}$$

$$c = (c_{11}, \cdots, c_{1q}, \cdots, c_{n1}, \cdots, c_{mq})'. \tag{12.3.16}$$

将上面矩阵方程用线性方程组的通常方式来表达, 即有

$$(A \otimes B')x = c, \tag{12.3.17}$$

其中 $A \otimes B'$ 为矩阵 A 和 B' 的 Kronecker 乘积. 由定理 12.3.5, 它有解的必要且充分条件为对 $A \otimes B'$ 的任意一个广义逆矩阵 $(A \otimes B')^-$, 则有

$$(A \otimes B')(A \otimes B')^- c = c. \tag{12.3.18}$$

这时通解为

$$x = (A \otimes B')^- c + \big(E - (A \otimes B')^-(A \otimes B')\big)\xi. \tag{12.3.19}$$

由性质 (8), $A^- \otimes (B')^- = A^- \otimes (B^-)'$ 为 $A \otimes B'$ 的广义逆矩阵. 于是取 $(A \otimes B')^-$ 为 $A^- \otimes (B')^-$, 则有

$$(A \otimes B')(A^- \otimes (B^-)')c = (AA^- \otimes (B^-B)')c = c, \tag{12.3.20}$$

且通解为

$$x = (A^- \otimes (B^-)')c + \big(E - (A^- \otimes (B^-)')(A \otimes B')\big)\xi. \tag{12.3.21}$$

写回矩阵形式, 即有解的充分且必要条件为

$$AA^-CB^-B = C, \tag{12.3.22}$$

又通解为

$$X = A^-CB^- + Z - A^-AZBB^-, \tag{12.3.23}$$

其中 Z 为 ξ 排成的 $m \times p$ 矩阵. □

第十三章　非负方阵 *

§13.1　不可分拆非负方阵的特征根 *

定义 13.1.1　n 阶方阵称为**非负方阵**, 如果它的元素全部为非负实数. n 阶非负方阵 A 称为**可分拆的**, 如果存在置换方阵 P, 使得

$$PAP' = \begin{pmatrix} A_1^{(s)} & A_2^{(s,n-s)} \\ 0 & A_3^{(n-s)} \end{pmatrix}, \tag{13.1.1}$$

其中 s 有 $1 \leqslant s \leqslant n-1$. 否则, 称为**不可分拆的**.

为了讨论非负方阵, 我们引进

定义 13.1.2　n 阶非负方阵 A 若非异, 且逆方阵 A^{-1} 仍为非负方阵, 则 A 称为**广义初等方阵**.

作为例子可知: 置换方阵 P_{jk} 为广义初等方阵, 又第二类初等方阵 $P_j(a)$, 其中 $a>0$, 也都是广义初等方阵. 但是第三类初等方阵不是广义初等方阵.

显然, 非负方阵在广义初等方阵的相抵下仍为非负方阵. 又广义初等方阵的乘积仍为广义初等方阵.

定理 13.1.3　广义初等方阵 P 能唯一地分解为

$$P = Q\Lambda_1 = \Lambda_2 Q, \tag{13.1.2}$$

其中 Q 为置换方阵, Λ_1, Λ_2 为对角元素大于零的对角方阵. 因此 Q, Λ_1, Λ_2 也都是

广义初等方阵.

证 记

$$P=\begin{pmatrix}p_{11}&\cdots&p_{1n}\\ \vdots&&\vdots\\ p_{n1}&\cdots&p_{nn}\end{pmatrix},\quad P^{-1}=\begin{pmatrix}q_{11}&\cdots&q_{1n}\\ \vdots&&\vdots\\ q_{n1}&\cdots&q_{nn}\end{pmatrix}. \tag{13.1.3}$$

由于 P 为广义初等方阵, 所以 p_{ij}, q_{ij} 也都是非负实数. 今

$$\sum_{k=1}^{n}p_{ik}q_{kj}=\delta_{ij},\quad i,j=1,2,\cdots,n.$$

所以当 $i\neq j$, 则有 $p_{ik}q_{kj}=0$, $k=1,\ 2,\ \cdots,\ n$. 但是 P^{-1} 非异, 所以它的第 k 行中必有一个元素 $q_{kj_0}\neq 0$. 因此 $p_{ik}=0$, $i=1,\ 2,\ \cdots,\ j_0-1,\ j_0+1,\ \cdots,\ n$. 即 P 的第 k 列中恰有一个元素不等于零. 由 k 任取, 所以证明了 P 的每一列恰有一个元素不等于零. 由 $\det(P)\neq 0$, 所以 P 的每一行也恰有一个元素不等于零. 将 P 中非零元素换成 1, 便得到一个置换方阵 Q, 且 PQ 及 QP 都是对角方阵, 分别记作 Λ_1, Λ_2. 它们的对角元素都大于零. 由 $Q^{-1}=Q$, 这证明了存在如定理要求的分解式 $P=Q\Lambda_1=\Lambda_2 Q$. 由于 Q 为实正交方阵, Λ_i 为正定实对称方阵, 所以上述分解实际上是极分解. 因此分解唯一. □.

引理 13.1.4 非负方阵的不可分拆性在广义初等方阵的相似下不改变.

证 设 A 为不可分拆非负方阵, P 为广义初等方阵. 由定理 13.1.3, $P=\Lambda_2 Q$, 其中 Q 为置换方阵, Λ_2 为对角元素大于零的对角方阵. 于是

$$P^{-1}AP=Q(\Lambda_2^{-1}A\Lambda_2)Q. \tag{13.1.4}$$

由于 A 和 $\Lambda_2^{-1}A\Lambda_2$ 这两个非负方阵的同位置元素同时等于零或同时不等于零, 这证明了 A 不可分拆当且仅当 $P^{-1}AP$ 不可分拆. □

定理 13.1.5 设 A 为 n 阶非负方阵, 则存在置换方阵 P, 使得

$$P^{-1}AP=\begin{pmatrix}\begin{pmatrix}A_{11}&A_{12}&\cdots&A_{1p}\\ &A_{22}&\cdots&A_{2p}\\ &&\ddots&\vdots\\ &&&A_{pp}\end{pmatrix}&\begin{pmatrix}A_{1,p+1}&\cdots&A_{1m}\\ A_{2,p+1}&\cdots&A_{2m}\\ \vdots&&\vdots\\ A_{p,p+1}&\cdots&A_{pm}\end{pmatrix}\\ 0&\begin{pmatrix}A_{p+1,p+1}&&\\ &\ddots&\\ &&A_{mm}\end{pmatrix}\end{pmatrix}, \tag{13.1.5}$$

其中 $A_{11}, A_{22}, \cdots, A_{mm}$ 为不可分拆非负方阵, 且 $A_{j,j+1}, \cdots, A_{jm}, 1 \leqslant j \leqslant p$ 中必有一个非零矩阵. 又不计子块的次序, 则标准形唯一.

证　如果 A 可分拆, 则存在置换方阵 P_1, 使得

$$P_1^{-1}AP_1 = \begin{pmatrix} B_1^{(q)} & B_2 \\ 0 & B_3 \end{pmatrix},$$

对置换方阵 $P_2^{(q)}, P_3^{(n-q)}$, 则

$$\begin{pmatrix} P_2 & 0 \\ 0 & P_3 \end{pmatrix}^{-1} P_1^{-1}AP_1 \begin{pmatrix} P_2 & 0 \\ 0 & P_3 \end{pmatrix} = \begin{pmatrix} P_2^{-1}B_1P_2 & P_2^{-1}B_2P_3 \\ 0 & P_3^{-1}B_3P_3 \end{pmatrix}.$$

所以若 B_1, B_3 不可分拆, 则已经证明了定理. 若有一个可分拆, 则变为

$$\begin{pmatrix} C_1 & C_2 & C_3 \\ 0 & C_4 & C_5 \\ 0 & 0 & C_6 \end{pmatrix}.$$

例如当 $C_2 = 0, C_3 = 0$, 则作

$$\begin{pmatrix} 0 & E & 0 \\ E & 0 & 0 \\ 0 & 0 & E \end{pmatrix}^{-1} \begin{pmatrix} C_1 & 0 & 0 \\ 0 & C_4 & C_5 \\ 0 & 0 & C_6 \end{pmatrix} \begin{pmatrix} 0 & E & 0 \\ E & 0 & 0 \\ 0 & 0 & E \end{pmatrix} = \begin{pmatrix} C_4 & 0 & C_5 \\ 0 & C_1 & 0 \\ 0 & 0 & C_6 \end{pmatrix}.$$

按上面的思路, 用归纳法便证明了标准形的存在性.

下面证唯一性. 为此用线性空间语言. 在 n 维线性空间 $\mathfrak{L}$ 中取定一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 于是存在 $\mathfrak{L}$ 上线性变换 $\mathscr{A}$, 使得它在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的方阵表示为 A. 对方阵 A 用置换方阵作相似, 就等于对基的顺序作排列. 所以用线性空间语言来描述这个定理, 即要证明在 $\mathfrak{L}$ 中存在子空间 $\mathfrak{L}_1, \mathfrak{L}_2, \cdots, \mathfrak{L}_m$, 使得 $\mathfrak{L}$ 有子空间直接和

$$\mathfrak{L} = \mathfrak{L}_1 \oplus \mathfrak{L}_2 \oplus \cdots \oplus \mathfrak{L}_m,$$

且有 $\mathscr{A}(\mathfrak{L}_1) \subset \mathfrak{L}_1$, $\mathscr{A}(\mathfrak{L}_2) \subset \mathfrak{L}_1 + \mathfrak{L}_2, \cdots, \mathscr{A}(\mathfrak{L}_p) \subset \mathfrak{L}_1 + \cdots + \mathfrak{L}_p$, $\mathscr{A}(\mathfrak{L}_{p+1}) \subset \mathfrak{L}_{p+1} + \mathfrak{L}_1 + \cdots + \mathfrak{L}_p, \cdots, \mathscr{A}(\mathfrak{L}_m) \subset \mathfrak{L}_m + \mathfrak{L}_1 + \cdots + \mathfrak{L}_p$. 而条件 $A_{11}, \cdots, A_{mm}$ 不可分拆等价于线性变换 $\mathscr{A}$ 在 $\mathfrak{L}_j$ 中没有非零真子空间 $\mathfrak{M}_j$, 使得 $\mathscr{A}(\mathfrak{M}_j) \subset \mathfrak{M}_j + \mathfrak{L}_{j-1} + \cdots + \mathfrak{L}_1$, $j = 1, 2, \cdots, m$.

所谓有两种标准形, 即存在 $\mathfrak{L}$ 的另一种子空间直接和分解:

$$\mathfrak{L} = \mathfrak{S}_1 \oplus \cdots \oplus \mathfrak{S}_q,$$

使得 $\mathscr{A}(\mathfrak{S}_1) \subset \mathfrak{S}_1$, $\mathscr{A}(\mathfrak{S}_2) \subset \mathfrak{S}_1 + \mathfrak{S}_2, \cdots, \mathscr{A}(\mathfrak{S}_s) \subset \mathfrak{S}_1 + \mathfrak{S}_2 + \cdots + \mathfrak{S}_s$, $\mathscr{A}(\mathfrak{S}_{s+1}) \subset \mathfrak{S}_{s+1} + \mathfrak{S}_1 + \cdots + \mathfrak{S}_s, \cdots, \mathscr{A}(\mathfrak{S}_q) \subset \mathfrak{S}_q + \mathfrak{S}_1 + \cdots + \mathfrak{S}_s$, 且线性变换 $\mathscr{A}$ 在 $\mathfrak{S}_j$ 中没

有非零真子空间 $\mathfrak{R}_j$, 使得

$$\mathscr{A}(\mathfrak{R}_j) \subset \mathfrak{R}_j + \mathfrak{S}_{j-1} + \cdots + \mathfrak{S}_1, \quad j = 1, 2, \cdots, q.$$

为了证明唯一性, 我们要证明 $q = m$, 且 $\mathfrak{S}_1, \cdots, \mathfrak{S}_m$ 是 $\mathfrak{L}_1, \cdots, \mathfrak{L}_m$ 的一个排列.

注意到在 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 中可取一些元素, 使得它们是 $\mathfrak{L}_j$ 的基, 且在 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 中可取另一些元素, 使得它们是 $\mathfrak{S}_k$ 的基, 这里 $j = 1, 2, \cdots, m$, $k = 1, 2, \cdots, q$. 所以存在 t 使得 $\mathfrak{S}_1 \cap \mathfrak{L}_1 = 0$, $\mathfrak{S}_1 \cap \mathfrak{L}_2 = 0, \cdots, \mathfrak{S}_1 \cap \mathfrak{L}_{t-1} = 0$, $\mathfrak{S}_1 \cap \mathfrak{L}_t \neq 0$. 今 $\mathscr{A}(\mathfrak{S}_1) \subset \mathfrak{S}_1$, $\mathscr{A}(\mathfrak{L}_1 + \cdots + \mathfrak{L}_t) \subset \mathfrak{L}_1 + \cdots + \mathfrak{L}_t$. 所以由基底的选取可知

$$\mathscr{A}(\mathfrak{S}_1 \cap \mathfrak{L}_t) \subset \mathfrak{S}_1 \cap (\mathfrak{L}_1 + \cdots + \mathfrak{L}_t) = \mathfrak{S}_1 \cap \mathfrak{L}_t.$$

由条件 $\mathfrak{S}_1 \cap \mathfrak{L}_t \neq 0$, 又 $\mathfrak{S}_1$ 中无 $\mathscr{A}$ 的不变子空间, 所以证明了

$$\mathfrak{S}_1 \cap \mathfrak{L}_t = \mathfrak{S}_1.$$

即 $\mathfrak{S}_1 \subset \mathfrak{L}_t$. 这证明了在 $\mathfrak{L}_t$ 中有非零子空间 $\mathfrak{S}_1$, 使得

$$\mathscr{A}(\mathfrak{S}_1) \subset \mathfrak{S}_1 + \mathfrak{L}_{t-1} + \cdots + \mathfrak{L}_1.$$

由条件便证明了 $\mathfrak{S}_1 = \mathfrak{L}_t$.

由于在 $\mathfrak{S}_1, \cdots, \mathfrak{S}_q$ 中各存在基, 它们拼成 $\{\alpha_1, \cdots, \alpha_n\}$, 在 $\mathfrak{L}_1, \cdots, \mathfrak{L}_m$ 中也各存在基, 它们也拼成 $\{\alpha_1, \cdots, \alpha_n\}$. 所以证明了

$$\mathfrak{S}_2 + \cdots + \mathfrak{S}_q = \mathfrak{L}_1 + \cdots + \mathfrak{L}_{t-1} + \mathfrak{L}_{t+1} + \cdots + \mathfrak{L}_m.$$

另一方面, 由于 $\mathfrak{L}_t$ 为 $\mathscr{A}$ 的不变子空间, 所以

$$\mathfrak{L} = \mathfrak{L}_t \oplus \mathfrak{L}_1 \oplus \cdots \oplus \mathfrak{L}_{t-1} \oplus \mathfrak{L}_{t+1} \oplus \cdots \oplus \mathfrak{L}_m$$

仍适合同样条件. 所以无妨设 $\mathfrak{S}_1 = \mathfrak{L}_1$.

考虑商空间 $\mathfrak{L}/\mathfrak{L}_1$. 由于 $\mathscr{A}(\mathfrak{L}_1) \subset \mathfrak{L}_1$, 所以 $\mathscr{A}$ 诱导了商空间 $\mathfrak{L}/\mathfrak{L}_1$ 上线性变换 $\widetilde{\mathscr{A}}$, 且有

$$\mathfrak{L}/\mathfrak{L}_1 = \mathfrak{L}_2/\mathfrak{L}_1 \oplus \cdots \oplus \mathfrak{L}_p/\mathfrak{L}_1 \oplus \mathfrak{L}_{p+1}/\mathfrak{L}_1 \oplus \cdots \oplus \mathfrak{L}_m/\mathfrak{L}_1,$$

$$\mathfrak{L}/\mathfrak{L}_1 = \mathfrak{L}/\mathfrak{S}_1 = \mathfrak{S}_2/\mathfrak{S}_1 \oplus \cdots \oplus \mathfrak{S}_s/\mathfrak{S}_1 \oplus \mathfrak{S}_{s+1}/\mathfrak{S}_1 \oplus \cdots \oplus \mathfrak{S}_q/\mathfrak{S}_1.$$

对线性变换 $\widetilde{\mathscr{A}}$ 而言, 它们仍然适合定理条件. 且 $\mathfrak{L}_2/\mathfrak{L}_1, \cdots, \mathfrak{L}_m/\mathfrak{L}_1$ 中有基, 它们构成集合 $\{\alpha_1 + \mathfrak{L}_1, \cdots, \alpha_n + \mathfrak{L}_1\}$. 同理, $\mathfrak{S}_2/\mathfrak{S}_1, \cdots, \mathfrak{S}_q/\mathfrak{S}_1$ 中有基, 它们构成集合 $\{\alpha_1 + \mathfrak{L}_1, \cdots, \alpha_n + \mathfrak{L}_1\}$. 所以对维数作归纳法, 便证明了在差一个排列意义下, 无妨设 $\mathfrak{S}_j/\mathfrak{S}_1 = \mathfrak{L}_j/\mathfrak{L}_1$, $j = 1, 2, \cdots, m$, 且 $q = m$. 由于 $\mathfrak{S}_1 = \mathfrak{L}_1$, 且 $\mathfrak{S}_1, \cdots, \mathfrak{S}_m$ 中

也有基拼成 $\{\alpha_1, \cdots, \alpha_n\}$. $\mathfrak{L}_1, \cdots, \mathfrak{L}_m$ 中有基拼成 $\{\alpha_1, \cdots, \alpha_n\}$, 这证明了 $q = m$, $\mathfrak{S}_j = \mathfrak{L}_j$, $j = 1, 2, \cdots, m$. □

给定 n 阶非负方阵 A, 定理 13.1.5 给出了标准形. 于是有

$$\det(\lambda E_n - A) = \det(\lambda E_n - P^{-1}AP) = \prod_{j=1}^{m} \det(\lambda E - A_j). \tag{13.1.6}$$

因此求非负方阵的特征根的问题化为求不可分拆非负方阵的特征根. 下面来讨论不可分拆非负方阵的特征根. 记一切 n 阶广义初等方阵构成的集合为 $\mathfrak{P}_n$.

定理 13.1.6 设 A 为非零方阵, 且为不可分拆非负方阵. 则 A 有正特征根 ρ 及属于 ρ 的特征向量 $(d_1, \cdots, d_n)'$, 其中 $d_1 > 0, \cdots, d_n > 0$. 且若另有正特征根 ρ_0 及相应的特征向量 $(f_1, \cdots, f_n)'$, 则有 $\rho_0 = \rho$, 且 $f_j = ad_j$, $1 \leqslant j \leqslant n$, 其中 a 为正实数.

证 取定 n 阶非负方阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \neq 0.$$

记

$$\rho(A) = \max_{1 \leqslant i \leqslant n} \left(\sum_{j=1}^{n} a_{ij} \right) > 0, \quad \rho = \inf_{P \in \mathfrak{P}_n} \rho(P^{-1}AP) \geqslant 0.$$

我们来证 $\rho > 0$, 且 ρ 为 A 的正特征根. 并求出一个属于特征根 ρ 的特征向量 $(d_1, \cdots, d_n)'$, 其中 $d_1 > 0, \cdots, d_n > 0$.

今任取广义初等方阵 $P \in \mathfrak{P}_n$. 由定理 13.1.3, $P = \Lambda Q$, 其中 Q 为置换方阵, Λ 为对角元素大于零的对角方阵, 记作 $\Lambda = \operatorname{diag}(d_1, \cdots, d_n)$. 于是 $\rho(P^{-1}AP) = \rho(Q^{-1}\Lambda^{-1}A\Lambda Q)$. 记由 Q 决定的 $1, 2, \cdots, n$ 的排列为 $i_1 i_2 \cdots i_n$, 则

$$Q^{-1}\Lambda^{-1}A\Lambda Q = \begin{pmatrix} a_{i_1 i_1} d_{i_1}^{-1} d_{i_1} & \cdots & a_{i_1 i_n} d_{i_1}^{-1} d_{i_n} \\ \vdots & & \vdots \\ a_{i_n i_1} d_{i_n}^{-1} d_{i_1} & \cdots & a_{i_n i_n} d_{i_n}^{-1} d_{i_n} \end{pmatrix}.$$

于是

$$\rho(P^{-1}AP) = \max_{1 \leqslant k \leqslant n} \sum_{p=1}^{n} a_{i_k i_p} d_{i_k}^{-1} d_{i_p} = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} a_{ij} d_i^{-1} d_j = \rho(\Lambda^{-1}A\Lambda).$$

因此

$$\rho = \inf_{\substack{\Lambda = \operatorname{diag}(d_1, \cdots, d_n) \\ d_1 > 0, \cdots, d_n > 0}} \rho(\Lambda^{-1}A\Lambda) = \inf_{\substack{\Lambda = \operatorname{diag}(d_1, \cdots, d_n) \\ d_1 > 0, \cdots, d_n > 0, d_1 + \cdots + d_n = 1}} \rho(\Lambda^{-1}A\Lambda).$$

由下确界的定义, 所以存在

$$\Lambda_k = \mathrm{diag}\,(d_{1k}, \cdots, d_{nk}), \quad d_{1k} > 0, \ \cdots, \ d_{nk} > 0, \ d_{1k} + \cdots + d_{nk} = 1,$$

使得 $\lim\limits_{k\to\infty} \rho(\Lambda_k^{-1} A \Lambda_k) = \rho$. 今 $d_{i1}, d_{i2}, \cdots, i = 1, 2, \cdots, n$ 是 n 个正项有界序列, 所以有收敛子序列. 因此无妨设每个序列 $d_{i1}, d_{i2}, \cdots$ 收敛于 $d_i, i = 1, 2, \cdots, n$. 自然 $d_1 \geqslant 0, \cdots, d_n \geqslant 0$.

今对 A 的行及列同步地作对换, 它诱导了对 $d_1, \cdots, d_n$ 作排列. 所以无妨设.

$$d_1 \geqslant \cdots \geqslant d_r > d_{r+1} = \cdots = d_n = 0, \quad d_1 + \cdots + d_n = 1.$$

又记 $\Lambda_0 = \mathrm{diag}\,(d_1, \cdots, d_n)$, 则有 $\lim\limits_{k\to\infty} \Lambda_k = \Lambda_0$, 且当 k 充分大, 有

$$\rho \leqslant \rho(\Lambda_k^{-1} A \Lambda_k) < \rho + \frac{1}{k}.$$

我们先来证明 $r = n$. 设若 $r < n$, 按前 r 行列分块, 有

$$A = \begin{pmatrix} A_1^{(r)} & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad \Lambda_k = \begin{pmatrix} D_k & 0 \\ 0 & G_k \end{pmatrix}, \quad \Lambda_0 = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix}.$$

因此 $D_k \to D_0$, $G_k \to 0$. 又

$$\Lambda_k^{-1} A \Lambda_k = \begin{pmatrix} D_k^{-1} A_1 D_k & D_k^{-1} A_2 G_k \\ G_k^{-1} A_3 D_k & G_k^{-1} A_4 G_k \end{pmatrix}.$$

今 $\rho(\Lambda_k^{-1} A \Lambda_k) < \rho + \dfrac{1}{k}$, 所以 $\Lambda_k^{-1} A \Lambda_k$ 的各行的和都小于 $\rho + \dfrac{1}{k}$. 特别, $G_k^{-1} A_3 D_k$ 的各行的和都小于 $\rho + \dfrac{1}{k}$. 这证明了

$$\sum_{j=1}^{r} a_{ij} d_{ik}^{-1} d_{jk} < \rho + \frac{1}{k}, \quad i = r+1, \cdots, n.$$

即

$$\sum_{j=1}^{r} a_{ij} d_{jk} < \left(\rho + \frac{1}{k}\right) d_{ik}, \quad i = r+1, \cdots, n.$$

当 $k \to \infty$, 便有 $\sum\limits_{j=1}^{r} a_{ij} d_j = 0$, 所以 $a_{ij} d_j = 0$, $r+1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant r$. 由 $d_1 > 0, \cdots, d_r > 0$, 便证明了 $a_{ij} = 0$, $1 \leqslant j \leqslant r$, $r+1 \leqslant i \leqslant n$, 即 $A_3 = 0$. 这和 A 不可分拆的假设矛盾. 所以证明了 $r = n$.

于是 $\Lambda_k \to \Lambda_0$, $\Lambda_k^{-1} \to \Lambda_0^{-1}$. 因此由 $\rho \leqslant \rho(\Lambda_k^{-1} A \Lambda_k) < \rho + \frac{1}{k}$, 则有 $\rho(\Lambda_0^{-1} A \Lambda_0) = \rho > 0$.

记 $B=\Lambda_0^{-1}A\Lambda_0=(b_{ij})$. 于是 $b_{ij}=a_{ij}d_i^{-1}d_j$, $1\leqslant i,j\leqslant n$. 因此令

$$\rho_i=\sum_{j=1}^n b_{ij}=\sum_{j=1}^n a_{ij}d_i^{-1}d_j,\quad 1\leqslant i\leqslant n.$$

则 $\rho=\rho(\Lambda_0^{-1}A\Lambda_0)=\max\limits_{1\leqslant i\leqslant n}\rho_i$. 我们无妨设

$$\rho=\rho_1=\cdots=\rho_s>\rho_{s+1}\geqslant\cdots\geqslant\rho_n>0.$$

下面来证明 $s=n$. 设若不然, 即 $s<n$. 将 B 按前 s 行列分块为

$$B=\begin{pmatrix}B_1 & B_2\\ B_3 & B_4\end{pmatrix}.$$

由于 A 不可分拆, 所以 B 也不可分拆, 即有 $B_2\neq 0$. 将 B 的前 s 行及前 s 列, 后 $n-s$ 行及后 $n-s$ 列作同步的置换, 从而无妨设 B_2 的第 1 行, 第 n 列元素 $b_{1n}\neq 0$. 作

$$C_1=P_n(a)^{-1}BP_n(a)=\begin{pmatrix}b_{11} & \cdots & b_{1,n-1} & ab_{1n}\\ \vdots & & \vdots & \vdots\\ b_{n-1,1} & \cdots & b_{n-1,n-1} & ab_{n-1,n}\\ a^{-1}b_{n1} & \cdots & a^{-1}b_{n,n-1} & b_{nn}\end{pmatrix},$$

于是各行和分别为

$$\rho_i'=\rho_i+(a-1)b_{in},\quad 1\leqslant i\leqslant n-1,\quad \rho_n'=a^{-1}\rho_n+(1-a^{-1})b_{nn}.$$

当 $0<a<1$ 时, 便有 $\rho_1'<\rho_1$, $\rho_i'\leqslant\rho_i$, $2\leqslant i<n$. 由 $b_{1n}>0$, 所以可取 a, 使得

$$0<a<1,\quad a^{-1}\rho_n+(1-a^{-1})b_{nn}<\rho_1+(a-1)b_{1n}.$$

此即 $\rho_n'<\rho_1'$. 所以

$$\rho_1'<\rho_1,\quad \rho_i'\leqslant\rho_i,\quad 2\leqslant i\leqslant n-1,\quad \rho_n'<\rho_1'<\rho_1.$$

于是由 $\rho(B)=\inf\rho(P^{-1}AP)$, 便有

$$\rho_1=\max_{1\leqslant i\leqslant n}\rho_i\geqslant\max_{1\leqslant i\leqslant n}\rho_i'=\rho(C_1)\geqslant\rho(B)=\rho_1.$$

这证明了存在 $\rho_{i_0}'=\rho_1$, 且 $\rho(C_1)=\rho(B)$, 这里 $i_0\in\{2,\cdots,s\}$. 这证明了至多有 $s-1$ 个 ρ_j' 等于 ρ.

对 C_1 按上法讨论, 所以至多 s 步, 便得到非负方阵序列 C_1, C_2, $\cdots$, C_q, 使得它们都由对 B 作广义初等方阵的相似后所得到非负方阵, 且 $\rho=\rho(B)=\rho(C_1)$

$= \cdots = \rho(C_q)$, 但是 C_q 的每个行和小于 ρ. 这和 $\rho = \inf \rho(P^{-1}AP)$ 矛盾. 所以证明了 $s = n$, 即有

$$\rho = \rho_1 = \cdots = \rho_n.$$

因此

$$A\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{1j}d_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}d_j \end{pmatrix} = \begin{pmatrix} d_1\rho_1 \\ \vdots \\ d_n\rho_n \end{pmatrix} = \rho\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}.$$

所以 ρ 为 A 的正特征根, 而 $(d_1, \cdots, d_n)'$ 为 A 的属于特征根 ρ 的特征向量, 其中 $d_1 > 0, \cdots, d_n > 0$.

下面证明唯一性. 若 A 有一个正特征根 ρ_0, 且 A 的属于特征根 ρ_0 的特征向量为 $(f_1, \cdots, f_n)'$, 使得 $f_1 > 0, \cdots, f_n > 0$. 我们来证 $\rho_0 = \rho$, 且存在正实数 a, 使得 $f_j = ad_j$, $j = 1, 2, \cdots, n$. 事实上, 记 $D = \operatorname{diag}(d_1, \cdots, d_n)$, $F = \operatorname{diag}(f_1, \cdots, f_n)$. 令 $B = D^{-1}AD = (b_{ij})$, $C = F^{-1}AF = (c_{ij})$. 记 $e = (1, 1, \cdots, 1)'$. 则由 $De = (d_1, \cdots, d_n)'$, $Fe = (f_1, \cdots, f_n)'$, 所以 $ADe = \rho De$, $AFe = \rho_0 Fe$, 此即 $Be = \rho e$, $Ce = \rho_0 e$. 但是 $A = DBD^{-1} = FCF^{-1}$, 即 $C = (F^{-1}D)B(F^{-1}D)^{-1}$, 其中

$$\Lambda = F^{-1}D = \operatorname{diag}(f_1^{-1}d_1, \cdots, f_n^{-1}d_n).$$

记 $\lambda_j^{-1} = f_j^{-1}d_j$, $1 \leqslant j \leqslant n$. 显然可无妨设

$$\lambda_1 = \cdots = \lambda_t > \lambda_{t+1} \geqslant \cdots \geqslant \lambda_n > 0.$$

令 $(F^{-1}D)^{-1}C = B(F^{-1}D)^{-1}$, 所以有 $\lambda_i c_{ij} = b_{ij}\lambda_j$. 由 $Be = \rho e$, $Ce = \rho_0 e$, 所以有 $\sum\limits_{j=1}^{n} b_{ij} = \rho$, $\sum\limits_{j=1}^{n} c_{ij} = \rho_0$. 因此

$$\rho_0 = \sum_{j=1}^{n} c_{ij} = \sum_{j=1}^{n} \lambda_i^{-1}\lambda_j b_{ij}, \quad i = 1, 2, \cdots, n.$$

于是

$$\rho \min_{1 \leqslant k \leqslant n} \lambda_k \leqslant \sum_{j=1}^{n} b_{ij}\lambda_j \leqslant \rho \max_{1 \leqslant k \leqslant n} \lambda_k, \quad 1 \leqslant i \leqslant n.$$

因此由 $\sum\limits_{j=1}^{n} b_{ij}\lambda_j = \lambda_i\rho_0$, 便有

$$\lambda_n\rho \leqslant \lambda_i\rho_0 \leqslant \lambda_1\rho_j, \quad 1 \leqslant i \leqslant n.$$

取 $i=n$, 有 $\rho\leqslant\rho_0$; 取 $i=1$, 有 $\rho_0\leqslant\rho$. 这证明了 $\rho_0=\rho$. 因此有

$$\sum_{j=1}^n b_{ij}\lambda_j=\lambda_i\rho_0=\lambda_i\rho=\sum_{j=1}^n\lambda_i b_{ij},$$

即

$$\sum_{j=1}^n(\lambda_i-\lambda_j)b_{ij}=0,\quad 1\leqslant i\leqslant n.$$

但是 $\lambda_1=\cdots=\lambda_t>\lambda_{t+1}\geqslant\cdots\geqslant\lambda_n>0$, 取 $i=1,\ 2,\ \cdots,\ t$, 便证明了 $\sum\limits_{j=t+1}^n(\lambda_1-\lambda_j)b_{ij}=0$, 其中 $\lambda_1-\lambda_j>0$, 所以 $b_{ij}=0,\ 1\leqslant i\leqslant t,\ t+1\leqslant j\leqslant n$, 即证明了当 $n>t$ 时 B 可分拆, 这和假设 B 不可分拆矛盾. 所以证明了 $n=t$, 即 $\Lambda=F^{-1}D=\lambda_1^{-1}E_n$. 即 $D=\lambda_1^{-1}F$, 所以 $f_j=\lambda_1 d_j,\ 1\leqslant j\leqslant n$, 其中 λ_1 为正实数.□

定理 13.1.7 (比较定理) 设 A 为 n 阶不可分拆非负方阵, 且

$$Ae=\rho e. \tag{13.1.7}$$

设 B 为 n 阶复方阵, 记 $B=(b_{ij})$, $A=(a_{ij})$. 假设

$$|b_{ij}|\leqslant a_{ij},\quad 1\leqslant i,j\leqslant n. \tag{13.1.8}$$

则 B 的任一特征根 λ_0 适合条件

$$|\lambda_0|\leqslant\rho, \tag{13.1.9}$$

且 B 有形如 $\lambda_0=\rho e^{\sqrt{-1}\theta}$ 的特征根当且仅当 λ_0 的特征向量为 $(\exp(-\sqrt{-1}\Lambda))e$, 又

$$B=(\exp(\sqrt{-1}\theta))(\exp(-\sqrt{-1}\Lambda))A(\exp(\sqrt{-1}\Lambda)),\quad \Lambda=\operatorname{diag}(\theta_1,\cdots,\theta_n), \tag{13.1.10}$$

其中 $0\leqslant\theta_1\leqslant\cdots\leqslant\theta_n<2\pi$.

证 设 λ_0 为 B 的特征根, $\alpha=(u_1,u_2,\cdots,u_n)'$ 为相应的特征向量. 由于 $B\alpha=\lambda_0\alpha$, 即有

$$\sum_{j=1}^n b_{kj}u_j=\lambda_0 u_k,\quad 1\leqslant k\leqslant n.$$

记 $u_0=\max\limits_{1\leqslant k\leqslant n}|u_k|$, 于是存在 k_0, 使得 $0\neq u_0=|u_{k_0}|$. 且由 $|b_{ij}|\leqslant a_{ij},\ 1\leqslant i,j\leqslant n$, 有

$$|\lambda_0|u_0=|\lambda_0 u_{k_0}|=\Big|\sum_{j=1}^n b_{k_0j}u_j\Big|\leqslant\sum_{j=1}^n|b_{k_0j}|u_0\leqslant\sum_{j=1}^n a_{k_0j}u_0=\rho u_0.$$

因此证明了 $|\lambda_0|\leqslant\rho$.

设 B 有特征根 $\lambda_0=\rho\exp(\sqrt{-1}\theta)$. 显然无妨设它的特征向量 α 有

$$|u_1|=\cdots=|u_t|>|u_{t+1}|\geqslant\cdots\geqslant|u_n|\geqslant 0.$$

我们来证 $t=n$. 事实上, 若 $t<n$, 将 A 按前 t 行列分块为

$$A=\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}.$$

由 A 不可分拆, $1\leqslant t<n$, 所以 $A_2\neq 0$. 对 A 的前 t 行列, 对 A 的后 $n-t$ 行列分别作同步的置换, 于是无妨设 $a_{1n}\neq 0$. 而

$$\rho|u_1|=\left|\rho e^{\sqrt{-1}\theta}u_1\right|=|\lambda_0 u_1|=\Big|\sum_{j=1}^n b_{1j}u_j\Big|\leqslant\sum_{j=1}^n a_{1j}|u_j|.$$

由于 $a_{1n}|u_n|<a_{1n}|u_1|$, 所以

$$\rho|u_1|<\sum_{j=1}^n a_{1j}|u_1|=\rho|u_1|,$$

这导出矛盾. 因此证明了 $t=n$, 即有

$$|u_1|=\cdots=|u_n|=a>0,$$

即

$$u_j=ae^{-\sqrt{-1}\theta_j},\quad j=1,2,\cdots,n.$$

从特征向量的角度可知, 无妨设 $a=1$. 所以特征向量

$$\alpha=\begin{pmatrix} u_1\\ \vdots \\ u_n\end{pmatrix}=\begin{pmatrix} e^{-\sqrt{-1}\theta_1}\\ \vdots \\ e^{-\sqrt{-1}\theta_n}\end{pmatrix}=\operatorname{diag}(e^{-\sqrt{-1}\theta_1},\cdots,e^{-\sqrt{-1}\theta_n})e=(\exp(-\sqrt{-1}\Lambda))e,$$

其中 $\Lambda=\operatorname{diag}(\theta_1,\cdots,\theta_n)$. 显然无妨设 $0\leqslant\theta_1\leqslant\cdots\leqslant\theta_n<2\pi$.

现在来计算 n 阶复方阵 B. 今

$$B\alpha=\lambda_0\alpha=\rho e^{\sqrt{-1}\theta}e^{-\sqrt{-1}\Lambda}e,$$

故有

$$B(\exp(-\sqrt{-1}\Lambda))e=\rho(\exp(\sqrt{-1}(\theta E-\Lambda)))e,$$

即有

$$\sum_{j=1}^n b_{ij}e^{-\sqrt{-1}\theta_j}=\rho e^{\sqrt{-1}(\theta-\theta_i)},\quad i=1,2,\cdots,n.$$

此即

$$\rho=\sum_{j=1}^n b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}=\Big|\sum_{j=1}^n b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}\Big|\leqslant\sum_{j=1}^n|b_{ij}|\leqslant\sum_{j=1}^n a_{ij}.$$

而 $Ae=\rho e$, 即有 $\rho=\sum\limits_{j=1}^{n}a_{ij}$, 这证明了

$$|b_{ij}|=a_{ij},\quad i,j=1,2,\cdots,n.$$

所以

$$b_{ij}=a_{ij}e^{\sqrt{-1}\xi_{ij}},\quad i,j=1,2,\cdots,n.$$

于是

$$\rho=\sum_{j=1}^{n}a_{ij}=\sum_{j=1}^{n}b_{ij}e^{\sqrt{-1}(\theta_i-\theta-\theta_j)}=\sum_{j=1}^{n}a_{ij}e^{\sqrt{-1}(\xi_{ij}+\theta_i-\theta-\theta_j)},\quad j=1,2,\cdots,n.$$

所以

$$\begin{aligned}\sum_{j=1}^{n}a_{ij}&=\sum_{j=1}^{n}a_{ij}\cos(\xi_{ij}+\theta_i-\theta-\theta_j)\\&=\sum_{j=1}^{n}a_{ij}-2\sum_{j=1}^{n}a_{ij}\sin^2\frac{1}{2}(\xi_{ij}+\theta_i-\theta-\theta_j).\end{aligned}$$

这证明了当 $a_{ij}>0$, 有 $\sin\frac{1}{2}(\xi_{ij}+\theta_i-\theta-\theta_j)=0$, 即有

$$\xi_{ij}=-\theta_i+\theta+\theta_j+2k_{ij}\pi,$$

其中 k_{ij} 为整数. 在 $a_{ij}=0$ 时, 约定 $\xi_{ij}=-\theta_i+\theta+\theta_j$. 总之, 有

$$e^{\sqrt{-1}\xi_{ij}}=e^{\sqrt{-1}(\theta+\theta_j-\theta_i)},\quad i,j=1,2,\cdots,n.$$

于是

$$B=(b_{ij})=(a_{ij}e^{\sqrt{-1}\xi_{ij}})=(a_{ij}e^{\sqrt{-1}(\theta+\theta_j-\theta_i)})=e^{\sqrt{-1}\theta}(\exp(\sqrt{-1}\Lambda))^{-1}A\exp(\sqrt{-1}\Lambda).$$

反之, 若

$$B=(\exp(\sqrt{-1}\theta))(\exp(\sqrt{-1}\Lambda))^{-1}A\exp(\sqrt{-1}\Lambda)$$

有特征根 λ_0, 相应特征向量为 $(\exp(\sqrt{-1}\Lambda))^{-1}e$, 我们来证 $\lambda_0=\rho e^{\sqrt{-1}\theta}$. 事实上, 由

$$B(\exp(\sqrt{-1}\Lambda))^{-1}e=\lambda_0(\exp(\sqrt{-1}\Lambda))^{-1}e,$$

所以

$$\lambda_0e=(\exp(\sqrt{-1}\Lambda))B(\exp(\sqrt{-1}\Lambda))^{-1}e=e^{\sqrt{-1}\theta}Ae=\rho e^{\sqrt{-1}\theta}e.$$

这证明了 $\lambda_0=\rho e^{\sqrt{-1}\theta}$. □

定理 13.1.8 n 阶不可分拆非负方阵 A 的任一特征根 λ_0 有

$$|\lambda_0|\leqslant\rho=\inf_{P\in\mathfrak{P}_n}\rho(P^{-1}AP),\tag{13.1.11}$$

且 A 的正特征根 ρ 为单重根. 又 $\rho E_n - A$ 的伴随方阵 $A(\rho)^{(*)}$ 的元素都是正实数. 记 $A=(a_{ij})$, 则有

$$\min_{1\leqslant i\leqslant n}\sum_{j=1}^{n}a_{ij}\leqslant\rho\leqslant\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}a_{ij},\tag{13.1.12}$$

且有

$$\rho=\min_{1\leqslant i\leqslant n}\sum_{j=1}^{n}a_{ij}\quad\text{当且仅当}\quad\rho=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}a_{ij}.\tag{13.1.13}$$

证 将 A 的第 k 行及第 k 列元素全换成零, 得到的非负方阵, 记作 A_k, 将 A_k 的第 k 行及第 k 列元素除去, 得到 A 的子矩阵, 记作 B_k, $k=1,2,\cdots,n$. 于是

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\det(\lambda E_n-A)=\sum_{k=1}^{n}\det(\lambda E_{n-1}-B_k).$$

由定理 13.1.6, A 有特征根 ρ 及属于特征根 ρ 的特征向量 $(d_1,\cdots,d_n)'$. 记 $D=\mathrm{diag}(d_1,\cdots,d_n)$, 则由

$$A\begin{pmatrix}d_1\\ \vdots\\ d_n\end{pmatrix}=\rho\begin{pmatrix}d_1\\ \vdots\\ d_n\end{pmatrix},$$

有 $ADe=\rho De$. 记 $\widetilde{A}=D^{-1}AD$, 则 $\widetilde{A}$ 仍为不可分拆非负方阵. 且 $\widetilde{A}e=\rho e$. 在定理 13.1.7 中取 A_k 为 B, 由 A_k 的定义可知定理条件适合. 因此证明了 A_k 的特征根 λ_0 有 $|\lambda_0|\leqslant\rho$, 且若 A_k 有特征根 $\lambda_0=\rho e^{i\theta}$, 对应的特征向量为 $(\exp(-\sqrt{-1}\Lambda))e$, 又

$$A_k=e^{\sqrt{-1}\theta}(\exp(-\sqrt{-1}\Lambda))A\exp(\sqrt{-1}\Lambda),$$

这里

$$\Lambda=\mathrm{diag}(\theta_1,\cdots,\theta_n).$$

但是由于 A_k 的第 k 行和第 k 列全等于零, 所以 A_k 可分拆, 这证明了 A 可分拆, 因此导出矛盾. 所以 $\widetilde{A_k}=D^{-1}A_kD$ 的所有特征根的模小于 ρ.

在 $\widetilde{A_k}$ 中划去第 k 行和第 k 列, 记作 $\widetilde{B_k}$. 由

$$\det(\lambda E_n-\widetilde{A_k})=\lambda\det(\lambda E_{n-1}-\widetilde{B_k}).$$

取 $\lambda=\rho$, 则由 $\det(\rho E_n-\widetilde{A_k})>0$, 所以 $\det(\rho E_{n-1}-\widetilde{B_k})>0$. 由

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\det(\lambda E_n-A)=\frac{\mathrm{d}}{\mathrm{d}\lambda}\det(\lambda E_n-\widetilde{A})=\sum_{k=1}^{n}\det(\lambda E_{n-1}-\widetilde{B_k}).$$

这证明了

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\det(\lambda E_n-A)\big|_{\lambda=\rho}>0.$$

所以 $\det(\lambda E_n - A)$ 以 ρ 为单重根.

今 $\det(\rho E_n - A) = 0$. 对 $\rho E_n - A$ 的伴随方阵 $A(\rho)^{(*)}$, 由 Laplace 展开定理可知

$$(\rho E_n - A)A(\rho)^{(*)} = \det(\rho E_n - A) \cdot E_n = 0.$$

记 $\boldsymbol{e_j} = (\delta_{j1}, \delta_{j2}, \cdots, \delta_{jn})' \in \mathbb{R}^n$, $1 \leqslant j \leqslant n$, 其中 δ_{jk} 为 Kronecker 符号. 又记 $\boldsymbol{\alpha_j} = A(\rho)^{(*)}e_j$, 即为 $A(\rho)^{(*)}$ 的第 j 列. 于是有 $A\alpha_j = \rho\alpha_j$, $j = 1, 2, \cdots, n$. 由于 A 以 ρ 为单重特征根, 即特征向量生成一维子空间. 所以对 A 的属于 ρ 的特征向量 $\boldsymbol{\alpha} = (d_1, \cdots, d_n)'$, $d_1 > 0, \cdots, d_n > 0$, 则有

$$A(\rho)^{(*)}e_j = \alpha_j = c_j\alpha, \quad j = 1, 2, \cdots, n,$$

其中 $c_1, \cdots, c_n$ 为实数. 所以

$$A(\rho)^{(*)} = (c_1\alpha \cdots c_n\alpha).$$

下面证 $A(\rho)^{(*)}$ 的元素全为正实数, 即证

$$c_1 > 0, \quad c_2 > 0, \quad \cdots, \quad c_n > 0.$$

事实上, 由于 $\det(\rho E_{n-1} - B_k)$ 为 $A(\rho)^{(*)}$ 的第 k 个对角元素, 而已证 $\det(\rho E_{n-1} - \widetilde{B_k}) > 0$. 又 $\widetilde{B_k}$ 和 B_k 相似, 所以证明了 $A(\rho)^{(*)}$ 的第 k 个对角元素为 $c_k d_k > 0$, $k = 1, 2, \cdots, n$. 已知 $d_1 > 0, \cdots, d_n > 0$, 这证明了 $c_1 > 0, \cdots, c_n > 0$. 所以 $A(\rho)^{(*)}$ 由正实数构成.

记 $A = (a_{ij})$, $\rho_i = \sum\limits_{j=1}^{n} a_{ij}$, $i = 1, 2, \cdots, n$. 下面证

$$\min_{1 \leqslant i \leqslant n} \rho_i \leqslant \rho \leqslant \max_{1 \leqslant i \leqslant n} \rho_i.$$

由

$$\rho = \inf \rho(P^{-1}AP) \leqslant \rho(A) = \max_{1 \leqslant i \leqslant n} \rho_i,$$

所以只要证 $\min\limits_{1 \leqslant i \leqslant n} \rho_i \leqslant \rho$. 令

$$0 = \det(\rho E_n - A) = \det \begin{pmatrix} \rho - a_{11} & \cdots & -a_{1n} \\ \vdots & & \vdots \\ -a_{n1} & \cdots & \rho - a_{nn} \end{pmatrix}$$

$$= \det \begin{pmatrix} \rho - a_{11} & \cdots & -a_{1,n-1} & \rho - \rho_1 \\ \vdots & & \vdots & \vdots \\ -a_{n-1,1} & \cdots & \rho - a_{n-1,n-1} & \rho - \rho_{n-1} \\ -a_{n1} & \cdots & -a_{n,n-1} & \rho - \rho_n \end{pmatrix}$$

按最后一列作 Laplace 展开, 则有

$$\sum_{j=1}^{n}(\rho-\rho_j)A_{jn}(\rho)=0,$$

其中 $A_{jn}(\rho)$ 为 $A(\rho)^{(*)}$ 的第 n 行, 第 j 列元素, 所以都是正实数. 这证明了 $\rho-\rho_1$, $\cdots$, $\rho-\rho_n$ 中若有元素大于零, 则也有元素小于零, 反之亦然. 所以

$$\min_{1\leqslant i\leqslant n}\rho_i\leqslant\rho\leqslant\max_{1\leqslant i\leqslant n}\rho_i.$$

最后, 设 $\rho=\min\limits_{1\leqslant i\leqslant n}\rho_i$, 则 $\rho\leqslant\rho_i$, $1\leqslant i\leqslant n$. 所以 $\rho-\rho_i\leqslant 0$, $i=1,2,\cdots,n$. 由 $\sum\limits_{j=1}^{n}(\rho-\rho_j)A_{jn}(\rho)=0$ 可知 $\rho_i=\rho$, $1\leqslant i\leqslant n$. 因此 $\rho=\max\limits_{1\leqslant i\leqslant n}\rho_i$. 同理, 设 $\rho=\max\limits_{1\leqslant i\leqslant n}\rho_i$, 则 $\rho\geqslant\rho_i$, $1\leqslant i\leqslant n$. 所以 $\rho_i=\rho$, $1\leqslant i\leqslant n$. 因此 $\rho=\min\limits_{1\leqslant i\leqslant n}\rho_i$. □

定理 13.1.9 设 A 为 n 阶不可分拆非负方阵, ρ 为 A 的最大正特征根. 则模等于 ρ 的特征根都是单重特征根, 它们是

$$\varepsilon_1\rho\quad,\varepsilon_2\rho\quad,\cdots,\quad\varepsilon_m\rho,\tag{13.1.14}$$

其中 m 为正整数,

$$\varepsilon_j=\exp\left(\frac{2\sqrt{-1}(j-1)\pi}{m}\right),\quad j=1,2,\cdots,m.\tag{13.1.15}$$

它们对应的特征向量分别为

$$(\exp(\sqrt{-1}\Lambda_1))^{-1}\alpha,\quad(\exp(\sqrt{-1}\Lambda_2))^{-1}\alpha,\quad\cdots,\quad(\exp(\sqrt{-1}\Lambda_m))^{-1}\alpha,\tag{13.1.16}$$

其中

$$\Lambda_j=\mathrm{diag}\,(\theta_{j1},\cdots,\theta_{jn}),\quad 0\leqslant\theta_{j1},\cdots,\theta_{jn}<2\pi,\quad j=1,2,\cdots,m.\tag{13.1.17}$$

又

$$\alpha=(d_1,\cdots,d_n)',\quad d_1>0,\quad\cdots,\quad d_n>0.\tag{13.1.18}$$

再者,

$$A=\left(\exp\left(\frac{2\sqrt{-1}\pi}{m}\right)\right)(\exp(\sqrt{-1}\Lambda_1))^{-1}A\exp(\sqrt{-1}\Lambda_1).\tag{13.1.19}$$

另一方面, 存在置换方阵 P, 使得

$$P^{-1}(\exp(\sqrt{-1}\Lambda_1))P=\mathrm{diag}\,(\varepsilon_1E_{n_1},\cdots,\varepsilon_mE_{n_m}),\tag{13.1.20}$$

$$P^{-1}AP=\begin{pmatrix} 0 & \cdots & 0 & A_{1m} \\ A_{21} & \ddots & \vdots & \vdots \\ & \ddots & 0 & 0 \\ & & A_{m,m-1} & 0 \end{pmatrix}. \qquad (13.1.21)$$

证 设 A 有特征根 $\rho\exp(i\theta)\neq\rho$, 其中 $0\leqslant\theta<2\pi$. 在这批特征根中取这样的特征根, 使得 θ 最小. 由定理 13.1.8, 存在正实数构成的对角方阵 D, 使得 $(D^{-1}AD)e=\rho e$. 由定理 13.1.7, 有

$$D^{-1}AD=(\exp(\sqrt{-1}\theta))(\exp(\sqrt{-1}\Lambda_1))^{-1}D^{-1}AD(\exp(\sqrt{-1}\Lambda_1)),$$

其中 $\Lambda_1=\mathrm{diag}(\theta_{11},\cdots,\theta_{1n})$, $0\leqslant\theta_{11},\cdots,\theta_{1n}<2\pi$. 由于 $\exp(\sqrt{-1}\Lambda_1)$ 为对角方阵, 所以 $D(\exp(\sqrt{-1}\Lambda_1))=(\exp(\sqrt{-1}\Lambda_1))D$. 这证明了

$$A=(\exp(\sqrt{-1}\theta))(\exp(\sqrt{-1}\Lambda_1))^{-1}A(\exp(\sqrt{-1}\Lambda_1)).$$

由 A 和 $(\exp(\sqrt{-1}\theta))A$ 相似, 所以 A 有特征根 ρ, $\rho\exp(\sqrt{-1}\theta)$, $\rho\exp(2\sqrt{-1}\theta),\cdots$. 但是 A 的特征根的个数有限, 所以存在最小自然数 m, 使得 $\exp(\sqrt{-1}m\theta)=1$. 即 $\exp(\sqrt{-1}\theta)$ 为 $x^m=1$ 的本原根, 即 $\theta=\dfrac{2\pi}{m}$. 于是记 $\varepsilon=\exp(\dfrac{2\pi\sqrt{-1}}{m})$, 则 A 有特征根 $\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$. 这时若 A 另有特征根 $\rho(\exp(\sqrt{-1}\varphi))$, 其中 $0<\varphi<2\pi$, 又 $\varphi\neq\dfrac{2j\pi\sqrt{-1}}{m}$, $j=1,2,\cdots,m$. 由 θ 的选取可知 $\dfrac{2\pi}{m}<\varphi$. 因此存在自然数 p, 使得 $\dfrac{2p\pi}{m}<\varphi<\dfrac{2(p+1)\pi}{m}$. 此即 $0<\varphi-\dfrac{2p\pi}{m}<\dfrac{2\pi}{m}$ 由于 A 和 $(\exp(\dfrac{-2p\pi\sqrt{-1}}{m}))A$ 相似, 所以从 A 有特征根 $\rho\exp(\sqrt{-1}\varphi)$ 可推出 A 有特征根 $\rho\exp(\sqrt{-1}(\varphi-\dfrac{2p\pi}{m}))$, 这和 θ 的选取矛盾. 所以形如 $\rho\exp(\sqrt{-1}\theta)$ 的特征根只有 $\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$. 由定理 13.1.8, A 有单重根 ρ. 由 A 和 $(\exp(\sqrt{-1}\theta))A$ 相似, 所以 $\varepsilon\rho,\varepsilon^2\rho,\cdots,\varepsilon^m\rho$ 都是 A 的单重根.

由定理 13.1.7, 所以 $D^{-1}AD$ 有属于特征根 $\varepsilon_j\rho$ 的特征向量 $(\exp(\sqrt{-1}\Lambda_j))^{-1}e$, 其中

$$\Lambda_j=\mathrm{diag}(\theta_{j1},\cdots,\theta_{jn}),\quad 0\leqslant\theta_{j1},\quad \theta_{j2},\cdots,\theta_{jn}<2\pi.$$

今 $D^{-1}AD(\exp(\sqrt{-1}\Lambda_j))^{-1}e=\varepsilon_j\rho(\exp(\sqrt{-1}\Lambda_j))^{-1}e$, 所以

$$A[(\exp(\sqrt{-1}\Lambda_j))^{-1}De]=(\varepsilon_j\rho)[(\exp(\sqrt{-1}\Lambda_j))^{-1}De].$$

因此, A 的属于特征根 $\varepsilon_j\rho$ 的特征向量为 $(\exp(\sqrt{-1}\Lambda_j))^{-1}\alpha$, 其中 $\alpha=(d_1,\cdots,d_n)'$, $j=1,2,\cdots,m$.

今 $0\leqslant\theta_{11},\cdots,\theta_{1n}<2\pi$. 于是存在置换方阵 P, 使得

$$P^{-1}\Lambda_1P=\mathrm{diag}(\mu_1E_{n_1},\cdots,\mu_sE_{n_s}),$$

其中 $\mu_1, \cdots, \mu_s$ 为区间 $[0,2\pi]$ 中不同数, 且无妨设 $\mu_1=0$. 将 A 和 Λ 一样分块, 则有

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1s} \\ \vdots & & \vdots \\ A_{s1} & \cdots & A_{ss} \end{pmatrix}.$$

由

$$A=(\exp(\frac{2\sqrt{-1}\pi}{m}))(\exp(\sqrt{-1}\Lambda_1))^{-1}A(\exp(\sqrt{-1}\Lambda_1)),$$

所以有

$$(\exp(\sqrt{-1}\mu_i)-\exp(\sqrt{-1}(\frac{2\pi}{m}+\mu_j)))A_{ij}=0, \quad i,j=1,2,\cdots,s.$$

取 $j=1$, 于是由 $\mu_1=0$, 有 $(\exp(\sqrt{-1}\mu_i)-\exp(\frac{2\sqrt{-1}\pi}{m}))A_{i1}=0$. 由 A 不可分拆, 所以 $A_{11}, A_{21}, \cdots, A_{s1}$ 中至少有一块不等于零, 而 $A_{11}=0$. 显然, 无妨设 $A_{21}\neq 0$, 即 $\mu_2=\frac{2\pi}{m}$, 于是 $A_{31}=0, \cdots, A_{s1}=0$. 取 $j=2$, 于是由 $\mu_1=0$, $\mu_2=\frac{2\pi}{m}$, 有 $(\exp(\sqrt{-1}\mu_i)-\exp(\frac{4\sqrt{-1}\pi}{m}))A_{i2}=0$. 所以 $A_{12}=0$, $A_{22}=0$. 由 A 不可分拆, 所以 $A_{32}, \cdots, A_{s2}$ 中至少有一块不为零. 显然, 无妨设 $A_{32}\neq 0$, 即 $\mu_3=\frac{4\pi}{m}$. 这样依次讨论下去, 便有 $\mu_j=\frac{2(j-1)\pi}{m}$, $j=1, 2, \cdots, s$, $s\leqslant m$, 且

$$A=\begin{pmatrix} 0 & 0 & \cdots & 0 & A_{1s} \\ A_{21} & 0 & \cdots & 0 & A_{2s} \\ 0 & A_{32} & \cdots & 0 & A_{3s} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & A_{s,s-1} & A_{ss} \end{pmatrix},$$

其中 $A_{21}\neq 0, \cdots, A_{s,s-1}\neq 0$. 而 $(\exp(\sqrt{-1}\mu_j)-\exp(\frac{2s\sqrt{-1}\pi}{m}))A_{is}=0$, $1\leqslant i\leqslant s$. 由 A 不可分拆可知 $s=m$, 且 $A_{2m}=0, \cdots, A_{mm}=0$. 所以证明了定理. □

习 题 13.1

13.1.1 设 A 为 n 阶不可分拆非负方阵, 试证 $(E+A)^{n-1}$ 仍为非负方阵, 其中元素全由正实数构成.

13.1.2 设 $A_1, A_2, \cdots$ 为不可分拆非负方阵序列. 记 $A_k=(a_{ij})^{(k)}$, $k=1, 2, \cdots$. 假设序列 $\{a_{ij}^{(k)}\}$ 为单调递增序列, $1\leqslant i,j\leqslant n$, 则 A_k 的最大正特征根 $\rho^{(k)}$ 也单调递增.

13.1.3 设不可分拆非负方阵 A 的最大正特征根 $\rho<1$, B 为非负方阵. 试证: 存在唯一的正数 b, 使得 $A+bB$ 为非负方阵, 且 $A+bB$ 的最大实特征根为 1.

13.1.4 设 ρ 为 n 阶不可分拆非负方阵 A 的最大正特征根, 记 $\rho_i(x)=\frac{e_i'Ax}{e_i'x}$, $\forall e_i'x>0$, 试证: $\min\limits_{x_1>0,\cdots,x_n>0}\max\limits_{1\leqslant j\leqslant n}\rho_j(x)=\rho=\max\limits_{x_1>0,\cdots,x_n>0}\min\limits_{1\leqslant j\leqslant n}\rho_j(x)$.

13.1.5 设 ρ 为不可分拆非负方阵 A 的最大正特征根. 记 α 为单位向量, 其坐标大于或等于零. 设 $A\alpha \leqslant \rho\alpha$ 或 $A\alpha \geqslant \rho\alpha$, 试证: $A\alpha = \rho\alpha$.

13.1.6 试严格证明定理 13.1.9 中的 $\varepsilon_1\rho, \cdots, \varepsilon_m\rho$ 都是不可分拆非负方阵 A 的单重特征根.

§13.2 非负方阵 *

设 A 是 n 阶非负方阵, 记 $n\times 1$ 矩阵 $e=(1,\cdots,1)'$. 任取正整数 k, 记

$$A_k = A + \frac{1}{k}ee' = A + \frac{1}{k}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}, \tag{13.2.1}$$

则 A_k 为不可分拆非负方阵. 所以 A_k 有最大正特征根 ρ_k, 对应的单位特征向量 α_k 由正实数构成. 今序列 $\{\alpha_k\}$ 有收敛子序列, 极限为 α_0, 即 $\lim\limits_{k\to\infty}\alpha_{i_k}=\alpha_0$. 它仍为单位向量, 元素由非负实数构成. 由 $A_{i_k}\alpha_{i_k}=\rho_{i_k}\alpha_{i_k}$, 双方取极限, 便证明了序列 $\{\rho_{i_k}\}$ 有极限 ρ_0, 且

$$A\alpha_0 = \rho_0\alpha_0. \tag{13.2.2}$$

今 $\rho_{i_k}>0$, 所以 $\rho_0\geqslant 0$. 又由 $\det(\rho_k E_n - A_k)=0$, 取极限, 有 $\det(\rho_0 E_n - A)=0$. 所以 ρ_0 为 A 的特征根. 另一方面, $\det(\lambda E_n - A_k)\to\det(\lambda E_n - A)$, 所以当 k 充分大时, 由 $\det(\lambda E_n - A_k)$ 和 $\det(\lambda E_n - A)$ 这两个多项式决定的曲线的纵坐标充分接近. 由于多项式函数的特性, 所以零点也充分接近. 这证明了 A_k 的最大实特征根 ρ_k 和 A 的最大实特征根充分接近. 由 $\rho_{i_k}\to\rho_0$, 这证明了 ρ_0 为 A 的最大正特征根, 所以 $\rho_0>0$. 由定理 13.1.7 可知 $\rho_k E - A_k$ 的伴随方阵 $A(\rho_k)^{(*)}$ 由正实数构成. 由于 $\lim\limits_{k\to\infty}A(\rho_{i_k})^{(*)}=A(\rho_0)^{(*)}$, 所以 $A(\rho_0)^{(*)}$ 为非负方阵.

引理 13.2.1 设 ρ_0 为非负方阵 A 的最大实特征根, 则当 $\lambda\geqslant\rho_0$ 时, $A(\lambda)^{(*)}$ 和 $\dfrac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)}$ 都是非负方阵.

证 今 $(\lambda E_n - A)A(\lambda)^{(*)}=\det(\lambda E_n - A)\cdot E_n$. 记 $A(\lambda)^{(*)}$ 的第 i 行, 第 j 列元素为 $A_{ij}(\lambda)$. 于是由

$$A(\lambda)^{(*)} + (\lambda E_n - A)\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)} = \mathrm{tr}\,(A(\lambda)^{(*)})E_n,$$

有

$$(\lambda E_n - A)^{-1}A(\lambda)^{(*)} + \frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)} = \mathrm{tr}\,(A(\lambda)^{(*)})(\lambda E_n - A)^{-1} = \sum_{j=1}^{n}A_{jj}(\lambda))(\lambda E_n - A)^{-1}.$$

所以由 $(\lambda E_n - A)^{-1}=(\det(\lambda E_n - A))^{-1}A(\lambda)^{(*)}$, 有

$$(\det(\lambda E_n - A))^{-1}(A(\lambda)^{(*)})^2 + \frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)} = (\det(\lambda E_n - A))^{-1}(\mathrm{tr}\,(A(\lambda)^{(*)}))A(\lambda)^{(*)}.$$

即有

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}A_{ij}(\lambda)=(\det(\lambda E_n-A))^{-1}\Big(\sum_{p=1}^{n}A_{pp}(\lambda)A_{ij}(\lambda)-\sum_{p=1}^{n}A_{ip}(\lambda)A_{pj}(\lambda)\Big)$$
$$=(\det(\lambda E_n-A))^{-1}\sum_{p=1}^{n}\det\begin{pmatrix}A_{pp}(\lambda) & A_{ip}(\lambda)\\ A_{pj}(\lambda) & A_{ij}(\lambda)\end{pmatrix}.$$

记 $n-1$ 阶主子方阵

$$\lambda E_{n-1}-A_p=\lambda E_{n-1}-A\begin{pmatrix}1\cdots(p-1)(p+1)\cdots n\\ 1\cdots(p-1)(p+1)\cdots n\end{pmatrix}$$

的伴随方阵的第 i 行, 第 j 列元素为 $A_{ij,p}(\lambda)$, 则易证

$$\det\begin{pmatrix}A_{pp}(\lambda) & A_{ip}(\lambda)\\ A_{pj}(\lambda) & A_{ij}(\lambda)\end{pmatrix}=(\det(\lambda E_n-A))A_{ij,p}(\lambda).$$

这证明了

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}A_{ij}(\lambda)=\sum_{p\neq i,j}A_{ij,p}(\lambda),\quad i,j=1,2,\cdots,n.$$

下面用对阶数作归纳法来证明引理. 当 $n=1$, 则 $A(\lambda)^{(*)}\equiv 1$, $\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)}\equiv 0$. 所以 $A(\lambda)^{(*)}$ 及 $\frac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)}$ 都是非负方阵. 设对阶数 $\leqslant n-1$ 的非负方阵, 相应的伴随方阵及其导数在 $\lambda\geqslant\rho_0$ 时都是非负方阵. 记

$$A_p=A\begin{pmatrix}12\cdots(p-1)(p+1)\cdots n\\ 12\cdots(p-1)(p+1)\cdots n\end{pmatrix}$$

的最大实特征根为 ρ_p. 今

$$\det(\lambda E_n-A)=(\lambda-a_{nn})A_{nn}(\lambda)-\sum_{j,k=1}^{n-1}a_{nj}a_{kn}A_{jk,n}(\lambda).$$

对 A_n 用归纳法假设, 所以 $A_{jk,n}$ 是非负实数, 因此

$$\det(\rho_nE_n-A)=(\rho_n-a_{nn})A_{nn}(\rho_n)-\sum_{j,k=1}^{n-1}a_{nj}a_{kn}A_{jk,n}(\rho_n).$$

由 $A_{nn}(\rho_n)=\det(\rho_nE_n-A_n)=0$, 所以 $\det(\rho_nE_n-A)\leqslant 0$. 但是 ρ_0 为 A 的最大实特征根, 即当 $\lambda_0>\rho_0$, 有 $\det(\lambda_0E_n-A)>0$. 这证明了 $\rho_n\leqslant\rho_0$. 同理可证 $\rho_i\leqslant\rho_0$, $i=1, 2, \cdots, n$. 所以 $\max\limits_{1\leqslant i\leqslant n}\rho_i\leqslant\rho_0$.

另一方面, 由归纳法假设, 当 $\lambda \geqslant \rho_i$ 时有 $A_i(\lambda)$ 非负. 这证明了当 $\lambda \geqslant \rho_0$ 时有

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}A_{ij}(\lambda) = \sum_{p\neq i,j} A_{ij,p}(\lambda) \geqslant 0.$$

所以 $\dfrac{\mathrm{d}}{\mathrm{d}\lambda}A(\lambda)^{(*)}$ 非负. 又由 $\dfrac{\mathrm{d}}{\mathrm{d}\lambda}A_{ij}(\lambda) \geqslant 0$, 所以 $A_{ij}(\lambda)$ 单调递增. 因此当 $\lambda \geqslant \rho_0$ 时, 有 $A_{ij}(\lambda) \geqslant A_{ij}(\rho_0) \geqslant 0$, $1 \leqslant i, j \leqslant n$. 这证明了 $A(\lambda)^{(*)}$ 当 $\lambda \geqslant \rho_0$ 时为非负方阵. □

推论 13.2.2 设 ρ_0 为不可分拆非负方阵 A 的最大实特征根, 则当 $\lambda \geqslant \rho_0$ 时, $A(\lambda)^{(*)}$ 由正实数构成.

证 由定理 13.1.7, $A(\rho_0)^{(*)}$ 中元素全为正实数. 由 $A_{ij}(\lambda) \geqslant A_{ij}(\rho_0) > 0$, $i, j = 1, 2, \cdots, n$, 所以 $A(\lambda)^{(*)}$ 中元素全为正实数. □

推论 13.2.3 设 ρ_0 为非负方阵 A 的最大实特征根, 则 $\det(\lambda E_n - A) > 0$, $\forall \lambda > \rho_0$. 因此 $(\lambda E_n - A)^{-1}$ 为非负方阵, $\forall \lambda > \rho_0$. 特别当 A 为不可分拆非负方阵时, $(\lambda E_n - A)^{-1}$ 由正实数构成, $\forall \lambda > \rho_0$.

证 令 ρ_0 为非负方阵 A 的最大实特征根, 所以 $\det(\lambda E_n - A)$ 在区间 $(\rho_0, +\infty)$ 中大于零. 由引理 13.2.1, $(\det(\lambda E_n - A))^{-1}A(\lambda)^{(*)}$ 为非负方阵, 所以 $(\lambda E_n - A)^{-1}$ 为非负方阵. 当 A 不可分拆时, 由推论 13.2.2 可知 $(\lambda E_n - A)^{-1}$ 由正实数构成. □

定理 13.2.4 设 ρ_0 为 n 阶非负方阵 A 的最大实特征根, 则 A 不可分拆当且仅当 $\rho_0 E_n - A$ 的伴随方阵 $A(\rho_0)^{(*)}$ 的对角元素 $A_{jj}(\rho_0) > 0$, $j = 1, 2, \cdots, n$.

证 设 A 不可分拆, 由于推论 13.2.2, $A(\rho_0)^{(*)}$ 的对角元素全是正实数. 反之, 若 $A(\rho_0)^{(*)}$ 的对角元素全是正实数, 记 A_p 为 A 中划去第 p 行、第 p 列后构成的 $n-1$ 阶子方阵, 则

$$A_{pp}(\rho_0) = \det(\rho_0 E_{n-1} - A_p) > 0.$$

所以 $n-1$ 阶方阵 A_p 的最大实特征根 $\rho_p < \rho_0$, $p = 1, 2, \cdots, n$. 如果 n 阶方阵 A 可分拆, 则无妨设

$$A = \begin{pmatrix} B^{(s)} & C \\ 0 & D \end{pmatrix},$$

其中 $0 < s < n$. 又无妨设 B 的最大实特征根为 ρ_0, 所以 A 中除去最后一行, 最后一列所得的子矩阵

$$A_n = \begin{pmatrix} B & C_1 \\ 0 & D_1 \end{pmatrix}$$

的最大实特征根 $\rho_n < \rho_0$. 但是 $\rho_0 \leqslant \rho_n$. 这导出矛盾. □

定义 13.2.5 设 A 为 n 阶不可分拆非负方阵, 若 A 的最大实特征根 ρ_0 大于所

有其他特征根的模, 则 A 称为 **本原方阵**, 否则称为**非本原方阵**.

引理 13.2.6 元素都是正数的方阵为本原方阵.

证 由定理 13.1.8, 设 A 的元素都是正数. 记 ρ 为 A 的最大正特征根, 则模等于 ρ 的特征根为 $\varepsilon_1\rho_1, \cdots, \varepsilon_m\rho$, 其中 $\varepsilon_j = \exp\dfrac{2\sqrt{-1}(j-1)\pi}{m}$, $j = 1, 2, \cdots, m$, 且

$$A = (\exp(\frac{2\pi\sqrt{-1}}{m}))(\exp(\sqrt{-1}\Lambda_1))^{-1}A(\exp(\sqrt{-1}\Lambda_1)),$$

其中 $\Lambda_1 = \mathrm{diag}(\theta_{11}, \cdots, \theta_{1n})$, $0 \leqslant \theta_{11}, \cdots, \theta_{1n} < 2\pi$. 记 $A = (a_{ij})$, 则有

$$(\exp(\sqrt{-1}\theta_{1i}))a_{ij} = a_{ij}\exp(\sqrt{-1}(\frac{2\pi}{m} + \theta_{1j})), \quad i, j = 1, 2, \cdots, n.$$

由 $a_{ij} > 0$, $i, j = 1, 2, \cdots, n$, 所以 $(\exp(\sqrt{-1}\theta_{1i})) = \exp(\sqrt{-1}(\dfrac{2\pi}{m} + \theta_{1i}))$. 这证明了 $m = 1$. 所以 A 的其他特征根 λ_0 必有 $|\lambda_0| < \rho$. □

引理 13.2.7 设 n 阶方阵 A 有 s 个不同特征根 $\lambda_1, \lambda_2, \cdots, \lambda_s$, 任取正整数 m, 则有

$$A^m = \sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda E_n - A)^{-1}]_{\lambda=\lambda_j},$$

其中 $m(\lambda) = (\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_s)^{e_s}$ 为 A 的极小多项式.

证 由于所求公式在相似下不改变, 所以我们可取 $A = \mathrm{diag}(J_1^{(f_1)}, \cdots, J_t^{(f_t)})$ 为 Jordan 标准形, 其中 $J_k^{(f_k)} = \mu_k E_{f_k} + N_{f_k}$, $k = 1, 2, \cdots, t$, 又 $\mu_1, \cdots, \mu_t$ 中不同数为 $\lambda_1, \cdots, \lambda_s$. 由于 $A^m = \mathrm{diag}(J_1^m, \cdots, J_t^m)$, 又

$$\begin{aligned}&\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda E_n - A)^{-1}]_{\lambda=\lambda_j}\\ =&\mathrm{diag}(\cdots, \sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda E_{f_k} - J_k)^{-1}]_{\lambda=\lambda_j}, \cdots),\end{aligned}$$

所以问题化为计算

$$\Omega_k = \sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}((\lambda-\mu_k)E_{f_k} - N_{f_k})^{-1}]_{\lambda=\lambda_j}.$$

注意到当 $\mu_k \neq \lambda_j$ 时, 由该项中有因式 $(\lambda-\lambda_j)^{e_j}$, 所以当 $\lambda = \lambda_j$ 时取值为零, 因此

存在指标 j, 使得 $\mu_k=\lambda_j$, 且

$$\begin{aligned}\Omega_k &= \frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}\lambda^m(\lambda-\lambda_j)^{e_j}[(\lambda-\lambda_j)E_{f_k}-N_{f_k}]^{-1}_{\lambda=\lambda_j}\\ &= \frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}\lambda^m\left[\sum_{p=0}^{f_k-1}(\lambda-\lambda_j)^{e_j-p-1}N_{f_k}^p\right]_{\lambda=\lambda_j}\\ &= \frac{1}{(e_j-1)!}\sum_{p=1}^{f_k-1}\left[\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j-p-1}]_{\lambda=\lambda_j}\right]N_{f_k}^p\\ &= \frac{1}{(e_j-1)!}\sum_{p=0}^{f_k-1}\frac{m!(e_j-1)!}{(m-p)!p!}\lambda_j^{m-p}N_{f_k}^p=\sum_{p=0}^{f_k-1}\binom{m}{p}\lambda_j^{m-p}N_{f_k}^p\\ &= \sum_{p=0}^{m}\binom{m}{p}\lambda_j^{m-p}N_{f_k}^p=(\lambda_jE_{f_k}-N_{f_k})^m=J_k^m,\quad k=1,2,\cdots,t.\end{aligned}$$

□

定理 13.2.8 n 阶不可分拆非负方阵 A 为本原方阵当且仅当存在正整数 m, 使得 A^m 的元素都是正数.

证 设 A 为本原方阵, $\lambda_1,\cdots,\lambda_s$ 为 A 的不同特征根, 其中 $\lambda_1>|\lambda_2|\geqslant\cdots\geqslant|\lambda_s|$. 由引理 13.2.7, 对任意正整数 m, 有

$$A^m=\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda E_n-A)^{-1}]_{\lambda=\lambda_j}.$$

今 A 不可分拆, A 的最大正特征根 λ_1 为单重根, 所以 $e_1=1$. 因此

$$A^m=[\lambda^m(\lambda-\lambda_1)(\lambda E_n-A)^{-1}]_{\lambda=\lambda_1}+\sum_{j=2}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[\lambda^m(\lambda-\lambda_j)^{e_j}(\lambda E_n-A)^{-1}]_{\lambda=\lambda_j}.$$

所以

$$\begin{aligned}\lambda_1^{-m}A^m=&[(\lambda-\lambda_1)(\lambda E_n-A)^{-1}]_{\lambda=\lambda_1}\\ &+\sum_{j=2}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}[(\frac{\lambda}{\lambda_1})^m(\lambda-\lambda_j)^{e_j}(\lambda E_n-A)^{-1}]_{\lambda=\lambda_j}.\end{aligned}$$

取 $m\to\infty$, 有

$$\lim_{m\to\infty}\frac{1}{\lambda_1^m}A^m=[(\lambda-\lambda_1)(\lambda E_n-A)^{-1}]_{\lambda=\lambda_1}=[(\lambda-\lambda_1)\det(\lambda E_n-A)^{-1}A(\lambda)^{(*)}]_{\lambda=\lambda_1}.$$

今

$$\det(\lambda E_n-A)=(\lambda-\lambda_1)(\lambda-\lambda_2)^{k_2}\cdots(\lambda_1-\lambda_s)^{k_s},$$

所以

$$\lim_{m\to\infty}\lambda_1^{-m}A^m=(\lambda_1-\lambda_2)^{-k_2}\cdots(\lambda_1-\lambda_s)^{-k_s}A(\lambda_1)^{(*)}.$$

由于 $\det(\lambda E_n-A)$ 为实多项式, 所以复根和它的共轭复根成对出现, 而实根都小于 λ_1, 所以证明了

$$(\lambda_1-\lambda_2)^{-k_2}\cdots(\lambda_1-\lambda_s)^{-k_s}>0.$$

由引理 13.2.1, $A(\lambda_1)^{(*)}$ 中元素都是正实数. 所以证明了 $\lim\limits_{m\to\infty}\lambda_1^{-m}A^m$ 为 n 阶方阵, 元素都是正实数. 所以存在充分大的数 m, 使得 $\lambda_1^{-m}A^m$ 中元素都是正实数, 即 A^m 中元素都是正实数.

反之, 若 A 为不可分拆非负方阵, 且存在正整数 m, 使得 A^m 中元素都是正实数. 由引理 13.2.6, 所以 A^m 为本原方阵. 因此 A^m 的不同特征根 $\mu_1, \cdots, \mu_s$ 有 $\mu_1>|\mu_2|\geqslant\cdots\geqslant|\mu_s|$. 由于 A 的 n 个特征根为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 所以 A^m 的 n 个特征根为 $\lambda_1^m, \lambda_2^m, \cdots, \lambda_n^m$. 又 A 有一个最大正特征根 λ_1, 所以 $\lambda_1=\sqrt[m]{\mu_1}$. 因此 A 的任意特征根如果有 $\lambda_j\neq\lambda_1$, 则有 $|\lambda_j|=\sqrt[m]{|\mu_k|}<\sqrt[m]{\mu_1}=\lambda_1$. 由定义可知 A 为本原方阵. □

定理 13.2.9 设 ρ 为不可分拆非负方阵 A 的最大正特征根, 则 A 为非本原的当且仅当存在置换方阵 P 及一个最大的正整数 m, 其中 $m\geqslant 2$, 且

$$P^{-1}A^mP=\operatorname{diag}(A_1,A_2,\cdots,A_m), \tag{13.2.3}$$

$A_1, A_2, \cdots, A_m$ 都是本原方阵, 且它们的最大正特征根都是 ρ_m, 又 A 恰有 m 个模等于 ρ 的特征根

$$\varepsilon_1\rho,\quad \varepsilon_2\rho,\quad \cdots,\quad \varepsilon_m\rho,$$

其中

$$\varepsilon_j=\exp\left(\frac{2(j-1)\pi\sqrt{-1}}{m}\right),\quad j=1,2,\cdots,m. \tag{13.2.4}$$

证 设 $P^{-1}A^mP=\operatorname{diag}(A_1,A_2,\cdots,A_m)$, 其中 $A_1, A_2, \cdots, A_m$ 本原, 且 $m\geqslant 2$. 所以 A^m 可分拆, 因此对任意正整数 q, A^{mq} 可分拆. 所以不存在正整数 p, 使得 A^p 中元素全是正数. 否则由 A^p 中元素全是正数, 则 A^{mp} 中元素也全是正数, 这和 A^{mp} 可分拆矛盾. 由定理 13.2.8, 所以 A 非本原.

反之, 若 A 为非本原的不可分拆非负方阵. 所以 A 有特征根 $\rho\exp(\sqrt{-1}\theta)\neq\rho$, 其中 ρ 为 A 的最大正特征根. 由定理 13.1.8, A 的模等于 ρ 的特征根为 $\varepsilon_1\rho, \varepsilon_2\rho, \cdots, \varepsilon_m\rho$, 其中 $m\geqslant 2$ 为正整数, $\varepsilon_j=\dfrac{2(j-1)\pi\sqrt{-1}}{m}$, $j=1, 2, \cdots, m$. 且存在置换方阵 P, 使得

$$P^{-1}(\exp(\sqrt{-1}\Lambda_1))P=\operatorname{diag}(\varepsilon_1E_{n_1},\cdots,\varepsilon_mE_{n_m}),$$

$$P^{-1}AP = \begin{pmatrix} 0 & \cdots & 0 & A_{1m} \\ A_{21} & \cdots & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & A_{m,m-1} & 0 \end{pmatrix}.$$

于是 $P^{-1}A^mP = \mathrm{diag}\,(A_1, \cdots, A_m)$. 下面分别求 A_1, A_2, $\cdots$, A_m 的最大正特征根. 今 A 不可分拆, 所以 $P^{-1}AP$ 不可分拆. 因此存在向量 $\alpha = (d_1, \cdots, d_n)'$ 为属于特征根 ρ 的特征向量, 其中 $d_1 > 0$, $\cdots$, $d_n > 0$. 将 α 和 $P^{-1}AP$ 一样分块, 则有 $\alpha' = (\alpha_1', \cdots, \alpha_m')$. 由 $P^{-1}AP\alpha = \rho\alpha$, 便有 $P^{-1}A^mP\alpha = \rho^m\alpha$, 所以有 $A_j\alpha_j = \rho^m\alpha_j$, $j = 1, 2, \cdots, m$. 这证明了 ρ^m 为 A_j 的特征根, 又对 A 的任一特征根 μ_0, 则有 $|\mu_0| \leqslant \rho$. 所以 $P^{-1}A^mP$ 有特征根 μ_0^m, 而 $|\mu_0^m| \leqslant \rho^m$. 这证明了 ρ^m 为 A_j 的最大正特征根. 设 $|\mu_0| = \rho$, 即

$$\mu_0 = \exp\left(\frac{2(k-1)\pi\sqrt{-1}}{m}\right)\rho,$$

其中 $k \in \{1, 2, \cdots, m\}$. 所以 $\mu_0^m = \rho^m$. 这证明了 A_j 中模为 ρ^m 的特征根 μ_0^m 只有 ρ^m. 因此 A_j 中其他特征根的模都小于 ρ^m. 由定义可知, 为了证 A_j 为本原方阵, 只要证 A_j 不可分拆就够了, 这里 $j = 1, 2, \cdots, m$.

今 A 不可分拆, 若 A_1, $\cdots$, A_m 中至少有一个可分拆. 由定理 13.1.3, 于是存在置换方阵 P_1, 使得非负方阵

$$P_1^{-1}P^{-1}A^mPP_1 = \begin{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1p} \\ & \ddots & \vdots \\ & & B_{pp} \end{pmatrix} & \begin{pmatrix} B_{1,p+1} & \cdots & B_{1q} \\ \vdots & & \vdots \\ B_{p,p+1} & \cdots & B_{pq} \end{pmatrix} \\ & \begin{pmatrix} B_{p+1,p+1} & & \\ & \ddots & \\ & & B_{qq} \end{pmatrix} \end{pmatrix},$$

其中 $B_{11}, \cdots, B_{qq}$ 不可分拆, 所以 $q > m$. 记 $B = P_1^{-1}P^{-1}APP_1$, 则由 A 不可分拆可知 B 及 B' 也不可分拆, 它们的最大正特征根都是 ρ. 于是存在属于特征根 ρ 的特征向量 α, β, 它们都由正实数构成. 由 $B\alpha = \rho\alpha$, $B'\beta = \rho\beta$. 将 α 和 β 按 B^m 一样分块方式分块为 $\alpha = (\alpha_1', \cdots, \alpha_q')'$, $\beta = (\beta_1', \cdots, \beta_q')'$. 由 $B^m\alpha = \rho^m\alpha$, $\beta'B^m = \rho^m\beta'$. 所以有 $B_{qq}\alpha_q = \rho^m\alpha_q$, $\beta_q'B_{qq} + \sum\limits_{j=1}^{p}\beta_j'B_{jq} = \rho^m\beta_q'$. 因此,

$$\rho^m\beta_q'\alpha_q = \beta_q'B_{qq}\alpha_q + \sum_{j=1}^{p}\beta_j'B_{jq}\alpha_q.$$

由 $B_{qq}\alpha_q = \rho^m\alpha_q$, 即有 $\sum\limits_{j=1}^{p}\beta_j' B_{jq}\alpha_q = 0$. 由于 B_{jq} 为非负方阵, β_j, α_q 的坐标都大于零, 所以有 $\beta_j' B_{jq}\alpha_q = 0$, $1 \leqslant j \leqslant p$. 即有 $B_{1q} = 0, \cdots, B_{pq} = 0$. 依次讨论下去, 于是同理可证

$$B^m = \operatorname{diag}\left(\begin{pmatrix} B_{11} & \cdots & B_{1p} \\ & \ddots & \vdots \\ & & B_{pp} \end{pmatrix}, B_{p+1,p+1}, \cdots, B_{qq}\right).$$

于是

$$\sum_{k=j}^{q} B_{jk}\alpha_k = \rho^m\alpha_j, \quad \sum_{k=1}^{j}\beta_k' B_{kj} = \rho^m\beta_j', \quad 1 \leqslant j \leqslant p.$$

比较 $j = 1$, 有

$$\rho^m\alpha_1 = \sum_{k=1}^{p} B_{1k}\alpha_k, \quad \rho^m\beta_1' = \beta_1' B_{11}.$$

所以 $\rho^m\beta_1'\alpha_1 = \beta_1' B_{11}\alpha_1 = \sum\limits_{k=1}^{p}\beta_1' B_{1k}\alpha_k$. 这证明了 $\sum\limits_{k=2}^{p}\beta_1' B_{1k}\alpha_k = 0$. 同理证明了 $B_{1k} = 0$, $k = 2, \cdots, p$. 依次讨论下去, 于是同理可证 $B_{jk} = 0$, $1 \leqslant j < k \leqslant p$. 即有 $B^m = \operatorname{diag}(B_{11}, \cdots, B_{qq})$, 其中 $B_{11}, \cdots, B_{qq}$ 都不可分拆, 且 $B_{jj}\alpha_j = \rho^m\alpha_j$, 即每个 B_{jj} 都以 ρ^m 为最大正特征根. 因此都是 B_{jj} 的单重特征根, 所以 B^m 有 q 重特征根 ρ^m.

另一方面, A 的模为 ρ 的特征根都是单重根, 且为 $\varepsilon_1\rho, \varepsilon_2\rho, \cdots, \varepsilon_m\rho$. 所以 B^m 的模为 ρ^m 的特征根为 $\varepsilon_1^m\rho_m, \cdots, \varepsilon_m^m\rho^m$, 即 B^m 有 m 重特征根 ρ^m, 这证明了 $q = m$. 它和 $q > m$ 矛盾, 所以证明了 $A_1, \cdots, A_m$ 都不可分拆, 即本原. □

习 题 13.2

13.2.1 设 ρ 为 n 阶非负方阵 A 的最大实特征根. 试证: $\lambda_0 > \rho$ 当且仅当 $\lambda_0 E_n - A$ 的所有主子式都大于零.

13.2.2 试用归纳法证明: 设 A 为 n 阶实方阵, a 为实数, $aE_n + A$ 为非负方阵, 设 A 的顺序主子式

$$\det A\begin{pmatrix} 12\cdots k \\ 12\cdots k \end{pmatrix}$$

有 $(-1)^k \det A\begin{pmatrix} 12\cdots k \\ 12\cdots k \end{pmatrix} > 0, \quad k = 1, 2, \cdots, n$, 试证: A 的一切主子式

$$\det A\begin{pmatrix} i_1 i_2\cdots i_k \\ i_1 i_2\cdots i_k \end{pmatrix}$$

有 $(-1)^k \det A\begin{pmatrix} i_1 i_2\cdots i_k \\ i_1 i_2\cdots i_k \end{pmatrix} > 0, \quad 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$, 由此推出第一题的条件可改为一切顺序主子式都大于零.

13.2.3 设 ρ 为非负方阵 A 的最大实特征根. 设 A 的特征根 ρ_0 有 $|\rho_0| < \rho$. 试证: 属于 ρ_0 的初等因式也都是一次因式.

13.2.4 记 ρ 为非负方阵 A 的最大实特征根. 设 ρ 为 $\det(\lambda E_n - A)$ 的单重根, 且 A 和 A' 的属于 ρ 的特征向量的坐标都是正数. 试证: A 为不可分拆非负方阵.

13.2.5 设 A 为不可分拆非负方阵. 记 $\det(\lambda E_n - A) = \lambda^n + a_1\lambda^{n_1} + \cdots + a_s\lambda^{n_s}$, 其中 $n > n_1 > \cdots > n_s \geqslant 0$, $a_1a_2\cdots a_s \neq 0$. 记 m 为 $n - n_1, \cdots, n_{s-1} - n_s$ 的最大公因式. 试证: A 恰有 m 个模等于 A 的最大正特征根的特征根, 它们也都是单重特征根.

§13.3 随机方阵 *

定义 13.3.1 记 A 为 n 阶非负方阵, $e = (1, \cdots, 1)'$. 如果 $Ae = e$, 即 A 有特征根 1, 对应的特征向量为 e, 则 A 称为 n 阶**随机方阵**.

引理 13.3.2 不可分拆随机方阵的最大正特征根 $\rho = 1$. 又若 A 为不可分拆非负方阵, 其中 ρ 为 A 的最大正特征根. 记 $\alpha = (d_1, \cdots, d_n)'$ 为 A 的属于 ρ 的特征向量, 其中 $d_1 > 0, \cdots, d_n > 0$. 记

$$\Lambda = \operatorname{diag}(d_1, \cdots, d_n). \tag{13.3.1}$$

则 $\rho^{-1}\Lambda^{-1}A\Lambda$ 为随机方阵.

证 设 ρ 为不可分拆随机方阵 A 的最大正特征根. 由定理 13.1.7, 所以

$$\min_{1\leqslant i\leqslant n} \sum_{j=1}^{n} a_{ij} \leqslant \rho \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} a_{ij},$$

其中 $A = (a_{ij})$. 由定义, $Ae = e$, 即有 $\sum\limits_{j=1}^{n} a_{ij} = 1$. 这证明了 $\rho = 1$.

又设 A 为不可分拆非负方阵. $A\alpha = \rho\alpha$, 即 $\sum\limits_{j=1}^{n} a_{ij}d_j = \rho d_i$, $i = 1, 2, \cdots, n$. 于是

$$A\Lambda e = A\alpha = \rho\alpha = \rho\Lambda,$$

即有 $(\Lambda^{-1}A\Lambda)e = \rho e$. 由于 $\Lambda^{-1}A\Lambda$ 仍为非负方阵. 由定义, $\rho^{-1}\Lambda^{-1}A\Lambda$ 为随机方阵.

□

显然, 设 A 为随机方阵, 对 A 用置换方阵作相似, 则仍为随机方阵.

定理 13.3.3 随机方阵 A 有标准形

$$B=\begin{pmatrix}\begin{pmatrix}A_{11} & \cdots & A_{1p}\\ & \ddots & \vdots\\ & & A_{pp}\end{pmatrix} & \begin{pmatrix}A_{1,p+1} & \cdots & A_{1m}\\ \vdots & & \vdots\\ A_{p,p+1} & \cdots & A_{pm}\end{pmatrix}\\ & \begin{pmatrix}A_{p+1,p+1} & & \\ & \ddots & \\ & & A_{mm}\end{pmatrix}\end{pmatrix}, \tag{13.3.2}$$

其中对角块 $A_{11}, \cdots, A_{mm}$ 都是不可分拆非负方阵, $A_{p+1,p+1}, \cdots, A_{mm}$ 为不可分拆随机方阵. 设 $A_{11}, \cdots, A_{pp}$ 的最大正特征根分别记作 $\rho_1, \cdots, \rho_p$, 则有 $\rho_1<1, \cdots, \rho_p<1$.

证 由 B 为随机方阵. 将 $e=(1,\cdots,1)'$ 和 B 一样分块, 即记 $e'=(\alpha_1',\cdots,\alpha_m')$, 则由 $Be=e$ 有 $A_{jj}\alpha_j=\alpha_j$, $j=p+1, \cdots, m$. 由于 $\alpha_j=(1,1,\cdots,1)'$, 这证明了 $A_{p+1,p+1}, \cdots, A_{mm}$ 为不可分拆随机方阵. 下面取 $j\in\{1,2,\cdots,p\}$. 设 A_{jj} 为 t 阶不可分拆非负方阵, 记作 $A_{jj}=(a_{kp})$. 由 $(A_{j,j+1}, \cdots, A_{jm})\neq 0$, 以及

$$\sum_{u=j}^{m}A_{ju}\alpha_u=\alpha_j,$$

所以 $A_{jj}\alpha_j$ 的坐标 $\leqslant 1$. 且至少有一个坐标 <1. 因此

$$\min_{1\leqslant k\leqslant t}\sum_{p=1}^{t}a_{kp}=a<1, \qquad \max_{1\leqslant k\leqslant t}\sum_{p=1}^{t}a_{kp}=b\leqslant 1.$$

由定理 13.1.7, $a\leqslant\rho_j\leqslant b$, 其中 ρ_j 为 A_{jj} 的最大正特征根.

设 $\rho_j=1$, 于是 $b=1$, $a<\rho_j$. 但是定理 13.1.7 证明了这时必须有 $\rho_j=a$. 这导出矛盾. 所以证明了 $\rho_j<1$. □

定理 13.3.4 随机方阵 A 的属于特征根 1 的初等因式都是一次的.

证 由定理 13.3.3, 随机方阵 A 的特征根 1 都在对角块 $A_{p+1,p+1}, \cdots, A_{mm}$ 中出现. 由于 $A_{p+1,p+1}, \cdots, A_{mm}$ 都是不可分拆非负方阵, 它们的最大正特征根都等于 1, 由定理 13.1.7 可知 1 是单重特征根. 这证明了定理. □

引理 13.3.5 n 阶复方阵 A 构成的序列 $\{A^q\}$ 有极限 A_∞ 当且仅当 A 的特征根 λ_0 的模 $|\lambda_0|\leqslant 1$. 当 $|\lambda_0|=1$ 时必有 $\lambda_0=1$, 且这时 $\lambda_0=1$ 的初等因式都是一次的.

证　显然无妨设 A 为 Jordan 标准形. 因此问题化为对 Jordan 块 $\lambda_0 E_n + N$ 来讨论. 今

$$(\lambda_0 E_n + N)^q = \sum_{p=0}^{q} \binom{q}{p} \lambda_0^{q-p} N^p,$$

其中 $N^{n-1} \neq 0$, $N^n = 0$.

设 $\{(\lambda_0 E_n + N)^q\}$ 的极限 A_∞ 存在, 于是对角元素 λ_0^q 的极限存在. 因此 $|\lambda_0| \leqslant 1$. 当 $|\lambda_0| < 1$ 时, λ_0^q 的极限为 0; 当 $|\lambda_0| = 1$ 时, λ_0^q 的极限存在当且仅当 $\lambda_0 = 1$. 这时 $(\lambda_0 E_n + N)^q = (E_n + N)^q = \sum\limits_{p=0}^{q} \binom{q}{p} N^p$. 当 N 的阶数 $n > 1$, 则 N 的系数 q 构成的序列的极限不存在. 这证明了 $n = 1$, 即 $\lambda_0 E_n + N$ 的初等因式 $(\lambda - \lambda_0)^n$ 为一次的.

反之, 若 $A = \lambda_0 E_n + N$ 的特征根 λ_0 有 $|\lambda_0| \leqslant 1$. 当 $|\lambda_0| = 1$ 时有 $\lambda_0 = 1$, 这时 $n = 1$. 显然, $\{(\lambda_0 E_n + N)^q\}$ 收敛. □

定理 13.3.6　设随机方阵 A 为本原方阵, 则序列 $\{A^q\}$ 的极限存在, 且极限为

$$\lim_{q\to\infty} A^q = \lim_{\lambda\to 1} (\lambda - 1)(\lambda E_n - A)^{-1} = A_\infty. \tag{13.3.3}$$

证　由引理 13.3.5 及定理 13.2.9,

$$\lim_{q\to\infty} A^q = \lim_{q\to\infty} \sum_{j=1}^{s} \frac{1}{(e_j-1)!} \frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}} [\lambda^q (\lambda - \lambda_j)^{e_j} (\lambda E_n - A)^{-1}]_{\lambda=\lambda_j},$$

其中 $\lambda_1, \cdots, \lambda_s$ 为不同复数, 又 $m(\lambda) = (\lambda - \lambda_1)^{e_1} (\lambda - \lambda_2)^{e_2} \cdots (\lambda - \lambda_s)^{e_s}$ 为极小多项式. 且 $|\lambda_j| \leqslant 1$. 当 $|\lambda_j| = 1$ 时 $\lambda_j = 1$, 且初等因式都是一次的.

今由 A 为随机方阵, 所以有 $Ae = e$, 因此存在特征根 1. 所以无妨设 $\lambda_1 = 1$. 由 A 为本原方阵, 所以 $|\lambda_p| < 1$, $2 \leqslant p \leqslant s$. 今 A 相似于 Jordan 标准形 $J = \mathrm{diag}\,(J_1^{(f_1)}, \cdots, J_t^{(f_t)})$, 即存在非异复方阵 Q, 使得 $Q^{-1}AQ = J$, 其中 J_k 的特征根为 $\lambda_1, \cdots, \lambda_s$ 之一. 设 $J_k^{(f_k)}$ 的特征根为 λ_p, 其中 $p \in \{2, \cdots, s\}$. 则由引理 13.2.7 的证明可知, 当 $q \to \infty$ 时有

$$\sum_{j=1}^{s} \frac{1}{(e_j-1)!} \frac{\mathrm{d}^{d_j-1}}{\mathrm{d}\lambda^{e_j-1}} [\lambda^q (\lambda - \lambda_j)^{e_j} (\lambda E_{f_k} - J_k)^{-1}]_{\lambda=\lambda_j} = \sum_{p=0}^{f_k-1} \binom{q}{p} \lambda_p^{q-p} N^p \to 0.$$

设 $J_k^{(f_k)}$ 的特征根为 1, 由于初等因式为一次的, 所以 $f_k = 1$, 从而 $J_k = 1$. 故

$$\begin{aligned}
&\sum_{j=1}^{s} \frac{1}{(e_j-1)!} \frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}} [\lambda^q (\lambda - \lambda_j)^{e_j} (\lambda E_{f_k} - J_k)^{-1}]_{\lambda=\lambda_j} \\
= &\sum_{j=1}^{s} \frac{1}{(e_j-1)!} \frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}} [\lambda^q (\lambda - \lambda_j)^{e_j} (\lambda - 1)^{-1}]_{\lambda=\lambda_j} \\
= &\frac{1}{(e_1-1)!} \frac{\mathrm{d}^{e_1-1}}{\mathrm{d}\lambda^{e_1-1}} [\lambda^q (\lambda - 1)^{e_1-1}]_{\lambda=1} = 1.
\end{aligned}$$

所以

$$\lim_{q\to\infty} A^q = Q\Big(\lim_{q\to\infty}\sum_{j=1}^{s}\frac{1}{(e_j-1)!}\frac{\mathrm{d}^{e_j-1}}{\mathrm{d}\lambda^{e_j-1}}\lambda^q(\lambda-\lambda_j)^{e_j}$$
$$\mathrm{diag}\,((\lambda E_{f_1}-J_1)^{-1},\cdots,(\lambda E_{f_t}-J_t)^{-1})_{\lambda=\lambda_j}\Big)Q^{-1}.$$

为方便起见, 无妨设 $J_1, \cdots, J_r$ 的特征根为 1, $J_{r+1}, \cdots, J_t$ 的特征根不是 1. 于是

$$\lim_{q\to\infty} A^q = Q\begin{pmatrix} E_r & 0 \\ 0 & 0\end{pmatrix}Q^{-1}.$$

另一方面,

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda E_n - A)^{-1} = Q\lim_{\lambda\to 1}(\lambda-1)\mathrm{diag}\,((\lambda E_{f_1}-J_1)^{-1},\cdots,(\lambda E_{f_t}-J_t)^{-1})Q^{-1}.$$

由于 $J_1=1, \cdots, J_t=1$, 所以

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda E_n - A)^{-1} = Q\begin{pmatrix} E_r & 0 \\ 0 & 0\end{pmatrix}Q^{-1}. \qquad \square$$

定理 13.3.7 设标准形

$$A = \begin{pmatrix} \begin{pmatrix} A_{11} & \cdots & A_{1p} \\ & \ddots & \vdots \\ & & A_{pp}\end{pmatrix} & \begin{pmatrix} A_{1,p+1} & \cdots & A_{1,m} \\ \vdots & & \vdots \\ A_{p,p+1} & \cdots & A_{pm}\end{pmatrix} \\ & \begin{pmatrix} A_{p+1,p+1} & & \\ & \ddots & \\ & & A_{mm}\end{pmatrix}\end{pmatrix} = \begin{pmatrix} B & C \\ 0 & D\end{pmatrix} \tag{13.3.4}$$

为随机方阵, 即有 $Ae=e$. 设 $A_{p+1,p+1}, \cdots, A_{mm}$ 为本原方阵, 则

$$\lim_{q\to\infty} A^q = \begin{pmatrix} 0 & (E-B)^{-1}C \\ 0 & E\end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & Q_\infty\end{pmatrix}, \tag{13.3.5}$$

其中

$$Q_\infty = \mathrm{diag}\,(Q_{p+1},\cdots,Q_m), \tag{13.3.6}$$

这里 $Q_{p+1}, \cdots, Q_m$ 分别定义为

$$\lim_{q\to\infty} A_{kk}^q = Q_k, \quad k=p+1,\cdots,m. \tag{13.3.7}$$

记 $\lambda E - A_{kk}$ 的伴随方阵为 $A_{kk}(\lambda)^{(*)}$, 则

$$Q_k = A_{kk}(1)^{(*)}\lim_{\lambda\to 1}(\lambda-1)\det(\lambda E - A_{kk})^{-1} = e\beta_k', \tag{13.3.8}$$

其中 $e'\beta_k = 1$, 且 β_k 由常数构成, $k = p+1, \cdots, m$.

证 由定理 13.3.6 及定理假设 $A_{p+1,p+1}, \cdots, A_{mm}$ 为随机本原方阵, 所以

$$\lim_{q\to\infty} A_{kk}^q = \lim_{\lambda\to 1}(\lambda-1)(\lambda E - A_{kk})^{-1} = Q_k, \quad k = p+1, \cdots, m.$$

由于 A 为随机方阵, 且 $A_{p+1,p+1}, \cdots, A_{mm}$ 为本原方阵. 由定理 13.3.3, 所以 A 为本原方阵. 由定理 13.3.6, 所以

$$\lim_{q\to\infty} A^q = \lim_{\lambda\to 1}(\lambda-1)(\lambda E_n - A)^{-1}.$$

今

$$(\lambda E_n - A) = \begin{pmatrix} \lambda E - B & -C \\ 0 & \lambda E - D \end{pmatrix},$$

且 B 的特征根的模小于 1, 所以在 $\lambda = 1$ 附近 $\lambda E - B$ 有逆. 因此

$$(\lambda E - A)^{-1} = \begin{pmatrix} (\lambda E - B)^{-1} & (\lambda E - B)^{-1}C(\lambda E - D)^{-1} \\ 0 & (\lambda E - D)^{-1} \end{pmatrix}.$$

由于

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda E - B)^{-1} = 0,$$

$$\lim_{\lambda\to 1}(\lambda-1)(\lambda E - D)^{-1} = \operatorname{diag}(Q_{p+1}, \cdots, Q_m) = Q_\infty.$$

所以

$$\lim_{q\to\infty} A^q = \begin{pmatrix} 0 & (E-B)^{-1}CQ_\infty \\ 0 & Q_\infty \end{pmatrix} = \begin{pmatrix} 0 & (E-B)^{-1}C \\ 0 & E \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & Q_\infty \end{pmatrix}.$$

于是问题化为证明: 设 A 为 n 阶随机本原方阵, 则

$$\lim_{q\to\infty} A^q = A(1)^{(*)} \lim_{\lambda\to 1}(\lambda-1)\det(\lambda E_n - A)^{-1} = e\beta',$$

其中 $\beta \in \mathbb{R}^n$, 且 $\beta' e = 1$, 又 β 的坐标全为正实数.

事实上, 由定理 13.3.6,

$$\lim_{q\to\infty} A^q = \lim_{\lambda\to 1}(\lambda-1)(\lambda E_n - A)^{-1}.$$

由于记 $\lambda E_n - A$ 的伴随方阵为 $A(\lambda)^{(*)}$, 则有

$$(\lambda E_n - A)^{-1} = (\det(\lambda E_n - A))^{-1}A(\lambda)^{(*)}.$$

所以

$$\begin{aligned} A_\infty = \lim_{q\to\infty} A^q &= \lim_{\lambda\to 1}(\lambda-1)(\det(\lambda E_n - A))^{-1}A(\lambda)^{(*)} \\ &= A(1)^{(*)} \lim_{\lambda\to 1}(\lambda-1)\det(\lambda E_n - A)^{-1}. \end{aligned}$$

另一方面, A 为本原方阵, 所以它的特征多项式

$$\det(\lambda E_n - A) = (\lambda-1)(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_s)^{e_s},$$

其中 $|\lambda_j| < 1$, $j = 2, \cdots, s$, 又 $1, \lambda_2, \cdots, \lambda_s$ 为 s 个不同的复数. 由于 $A^{q+1} = AA^q$, 所以取 $q \to \infty$, 有 $A_\infty = AA_\infty$. 因此由 $a = \lim\limits_{\lambda\to 1}(\lambda-1)\det(\lambda E_n - A)^{-1} > 0$ 及 $A(1)^{(*)}$ 的元素为正实数, 所以 A_∞ 的元素都是正实数. 记 A_∞ 的列向量分别为 $\beta_1, \cdots, \beta_n$. 所以有 $A_\infty = (\beta_1 \cdots \beta_n)$. 因此有

$$A\beta_j = \beta_j, \quad j = 1, 2, \cdots, n.$$

所以 $\beta_1, \cdots, \beta_n$ 都是方阵 A 的属于特征根 1 的特征向量. 由于 A 为本原随机方阵, 所以 A 的特征根 1 为单重根, 且 e 为特征向量. 这证明了 $\beta_j = a_j e$, $j = 1, 2, \cdots, n$. 其中 $a_1, \cdots, a_n$ 为正实数. 记 $\beta = (a_1 \cdots a_n)'$, 则

$$A_\infty = (\beta_1 \cdots \beta_n) = (a_1 e \cdots a_n e) = e\beta'.$$

由 $Ae = e$, 于是 $A^q e = e$. 因此 $A_\infty e = e$, 所以 $e\beta' e = e$. 这证明了 $\beta' e = 1$. □

习　　题　13.3

13.3.1 试证: 任一非负方阵必在广义初等方阵下相似于随机方阵.

13.3.2 设 A 为不可分拆非负方阵, ρ 为 A 的最大正特征根. 设 A 的模为 ρ 的不同特征根恰有 m 个. 试证: 序列 $\{A^q\}$ 中有 m 个收敛子序列, 它们的极限分别为 $A_\infty, A_\infty A, \cdots, A_\infty A^{N-1}$. 试用这个结论来证明序列 $\{\dfrac{1}{m}\sum\limits_{q=0}^{m} A^q\}$ 的极限存在.

13.3.3 设 A 和 A' 都是随机方阵, 则 A 称为**双随机方阵**. 试证:

(i) A 为双随机方阵当且仅当 A' 为双随机方阵;

(ii) 置换方阵为双随机方阵;

(iii) 设 A 为双随机方阵, 且 A 中不等于零的元素的个数不超过 $n+3$, 则存在置换方阵 P_1 和 P_2, 使得 $A = \lambda P_1 + (1-\lambda)P_2$, 其中 $\lambda \in [0, 1]$.

第十四章　矩阵偶的标准形理论 *

§14.1　矩阵偶在相抵下的标准形 *

在这一章, 我们考虑一对矩阵同时化标准形的问题. 为此，我们称一对同型矩阵 A 和 B 为**矩阵偶**.

在 §9.4 中已经提到, 方阵在实正交相似下的标准形理论, 等于实对称方阵和实反对称方阵构成的方阵偶在实正交相似下同时化为标准形的理论; 同样, 在 §11.2 中也已经提到, (复和实) 方阵在相合下的标准形理论, 也等于对称方阵和反对称方阵构成的方阵偶在相合下同时化为标准形的理论. 所以, 矩阵偶标准形理论的建立是很自然的. 在这一章只给出两种复矩阵偶的标准形理论.

定义 14.1.1　两个 $n\times m$ 矩阵偶 (A,B) 和 (C,D) 称为**相抵的**, 如果存在一个 n 阶非异方阵 P 及 m 阶非异方阵 Q, 使得

$$C=PAQ,\quad D=PBQ. \tag{14.1.1}$$

在这一节, 我们将求出矩阵偶在相抵下的标准形和全系不变量. 这一理论是属于 Kronecker 的. 处理矩阵偶的标准形理论的特点是引进参数 λ, 考虑 λ 矩阵, 用 §6.3 的相抵定义. 为了区别起见, 给出

定义 14.1.2　两个 $n\times m$ λ 矩阵 $A(\lambda)$ 和 $C(\lambda)$ 称为**严格相抵的**, 如果存在 n 阶非异常数方阵 P 及 m 阶非异常数方阵 Q, 使得

$$PA(\lambda)Q=C(\lambda). \tag{14.1.2}$$

结合定义 14.1.1 和定义 14.1.2, 我们有

引理 14.1.3　两个 $n\times m$ 矩阵偶 (A,B) 和 (C,D) 相抵当且仅当 λ 矩阵 $A+\lambda B$

和 $C+\lambda D$ 严格相抵. 所以矩阵偶 (A,B) 及 (C,D) 在相抵下的标准形理论就变成 λ 矩阵 $A+\lambda B$ 及 $C+\lambda D$ 在严格相抵下的标准形理论.

设 $n\times m$ λ 矩阵 $A+\lambda B$ 的秩为 r, 于是 $r\leqslant\min(n,m)$. 在下面, 我们分为三种情形考虑: **(1)** $n\geqslant r$, $m>r$; **(2)** $n>r$, $m\geqslant r$; **(3)** $n=m=r$.

注意到考虑 $m\times n$ λ 矩阵 $(A+\lambda B)'=A'+\lambda B'$, 则情形**(2)** 化为情形**(1)**. 所以实际上只考虑情形**(1)** 和情形**(3)** 就可以了.

设 $n\geqslant r$, $m>r$. 这时考虑线性方程组

$$(A+\lambda B)x=0. \tag{14.1.3}$$

视 λ 为任意参数, 则由齐次线性方程组理论可知它有非零解. 任取一个非零解

$$x=\Big(\frac{f_1(\lambda)}{g_1(\lambda)},\cdots,\frac{f_m(\lambda)}{g_m(\lambda)}\Big)', \tag{14.1.4}$$

其中 $f_1(\lambda)$, $f_2(\lambda)$, $\cdots$, $f_m(\lambda)$, $g_1(\lambda)$, $g_2(\lambda)$, $\cdots$, $g_m(\lambda)$ 都是 λ 的多项式, 且 $g_1(\lambda)$, $g_2(\lambda)$, $\cdots$, $g_m(\lambda)$ 都是 λ 的非零多项式. 由于齐次线性方程组的解乘以 λ 的非零多项式后仍为线性方程组 $(A+\lambda B)x=0$ 的非零解. 所以经过通分, 我们总可取线性方程组 $(A+\lambda B)x=0$ 的非零解为

$$x=(f_1(\lambda),f_2(\lambda),\cdots,f_m(\lambda))', \tag{14.1.5}$$

其中 $f_1(\lambda)$, $f_2(\lambda)$, $\cdots$, $f_m(\lambda)$ 为 m 个不全为零的多项式, 且我们约定 $f_1(\lambda)$, $f_2(\lambda)$, $\cdots$, $f_m(\lambda)$ 互素. 记

$$d=\max(\deg(f_1(\lambda)),\deg(f_2(\lambda)),\cdots,\deg(f_m(\lambda))), \tag{14.1.6}$$

则 d 称为这个解 x 的**次数**.

为了今后讨论方便, 我们称上面由 m 个不全为零的多项式构成的解为齐次线性方程组 $(A+\lambda B)x=0$ 的**解**.

定义 14.1.4 设 $n\times m$ 矩阵 $A+\lambda B$ 的秩 $r<m$, 则齐次线性方程组 $(A+\lambda B)x=0$ 的所有非零解中存在一个次数最小的非零解. 这个次数 ε (为非负整数) 称为 λ 矩阵 $A+\lambda B$ 的**最小列指标**. 设 $n\times m$ λ 矩阵 $A'+\lambda B'$ 的秩 $r<n$, 则它的最小列指标称为 $n\times m$ λ 矩阵 $A+\lambda B$ 的**最小行指标**.

引理 14.1.5 最小列 (行) 指标都在严格相抵下不改变.

证 设 $C+\lambda D=P(A+\lambda B)Q$, 其中 P 和 Q 是非异常数方阵. 则齐次线性方阵组 $(A+\lambda B)x=0$ 可改写为 $P^{-1}(C+\lambda D)Q^{-1}x=0$. 即 $(C+\lambda D)y=0$, 其中 $y=Q^{-1}x$. 这证明了 $C+\lambda D$ 的最小列指标 $d'\leqslant d$. 由于 $x=Qy$, 同理可证 $d\leqslant d'$,

所以证明了 $d' = d$. 对最小行指标的证明相同. □

引理 14.1.6 设 ε 为 $n \times m$ λ 矩阵 $A + \lambda B$ 的最小列指标, 则存在 $m \times 1$ λ 矩阵

$$\alpha = \alpha_0 + \lambda\alpha_1 + \lambda^2\alpha_2 + \cdots + \lambda^{\varepsilon}\alpha_{\varepsilon}, \tag{14.1.7}$$

其中 $\alpha_0, \alpha_1, \cdots, \alpha_{\varepsilon}$ 都是 $m \times 1$ 复常数矩阵, 且 $\alpha_0 \neq 0, \cdots, \alpha_{\varepsilon} \neq 0$, 又 $\alpha_0, \alpha_1, \cdots, \alpha_{\varepsilon}$ 线性无关, 且 $A\alpha_1, A\alpha_2, \cdots, A\alpha_{\varepsilon} \in \mathbb{C}^m$ 也线性无关.

证 今齐次线性方阵组 $(A + \lambda B)x = 0$ 有非零解

$$\alpha = (f_1(\lambda), f_2(\lambda), \cdots, f_m(\lambda))' = \alpha_0 + \lambda\alpha_1 + \lambda^2\alpha_2 + \cdots + \lambda^{\varepsilon}\alpha_{\varepsilon},$$

其中 $\alpha_{\varepsilon} \neq 0$. 由 $f_1(\lambda), f_2(\lambda), \cdots, f_m(\lambda)$ 互素可知 $\alpha \neq 0$. 今 $(A + \lambda B)\alpha = 0$, 所以有

$$A\alpha_0 = 0, \quad A\alpha_1 + B\alpha_0 = 0, \quad \cdots, \quad A\alpha_{\varepsilon} + B\alpha_{\varepsilon-1} = 0, \quad B\alpha_{\varepsilon} = 0.$$

先来证当 $\varepsilon > 0$, 则有 $A\alpha_1, A\alpha_2, \cdots, A\alpha_{\varepsilon} \in \mathbb{C}^m$ 线性无关. 今若 $A\alpha_1 = 0$, 于是 $B\alpha_0 = 0$. 因此 $(A + \lambda B)\alpha_0 = 0$. 但是 $\alpha_0 \neq 0$, 这证明了最小列指标 $\varepsilon = 0$. 这和条件 $\varepsilon > 0$ 矛盾. 所以 $A\alpha_1 \neq 0$, 即 $A\alpha_1$ 线性无关. 设 $A\alpha_1, A\alpha_2, \cdots, A\alpha_t$ 线性无关, 且 $t < \varepsilon$. 我们来证 $A\alpha_1, A\alpha_2, \cdots, A\alpha_t, A\alpha_{t+1}$ 也线性无关. 设若不然, 则 $A\alpha_{t+1}$ 为 $A\alpha_1, A\alpha_2, \cdots, A\alpha_t$ 的线性组合, 所以存在常数 $b_0, b_1, \cdots, b_{t-1}$, 使得

$$A\alpha_{t+1} + b_{t-1}A\alpha_t + \cdots + b_1A\alpha_2 + b_0A\alpha_1 = 0.$$

由于 $B\alpha_{j-1} = -A\alpha_j$, $j = 1, 2, \cdots, \varepsilon$, 代入, 有

$$B\beta_t = 0, \quad \beta_t = b_0\alpha_0 + b_1\alpha_1 + \cdots + b_{t-1}\alpha_{t-1} + \alpha_t.$$

又

$$\begin{aligned} A\beta_t &= b_0A\alpha_0 + b_1A\alpha_1 + \cdots + b_{t-1}A\alpha_{t-1} + A\alpha_t \\ &= -b_1B\alpha_0 - b_2B\alpha_1 - \cdots - b_{t-1}B\alpha_{t-2} - B\alpha_{t-1} \\ &= -B\beta_{t-1}, \end{aligned}$$

其中

$$\beta_{t-1} = b_1\alpha_0 + b_2\alpha_1 + \cdots + b_{t-1}\alpha_{t-2} + \alpha_{t-1}.$$

这样依次作下去, 即令

$$\beta_j = b_{t-j}\alpha_0 + b_{t-j+1}\alpha_1 + \cdots + b_{t-1}\alpha_{j-1} + \alpha_j, \quad j = 0, 1, \cdots, t,$$

则由归纳法易证

$$B\beta_t = 0, \quad B\beta_{t-1} + A\beta_t = 0, \quad \cdots, \quad B\beta_0 + A\beta_1 = 0, \quad A\beta_0 = 0.$$

作 λ 矩阵

$$\beta=\beta_0+\lambda\beta_1+\cdots+\lambda^t\beta_t,$$

则有 $(A+\lambda B)\beta=0$. 注意到 $\beta_0=\alpha_0\neq 0$, 所以 β 为齐次线性方程组 $(A+\lambda B)x=0$ 的非零解, 且 β 由 m 个不全为零的多项式构成, 它的最大次数 $\leqslant t<\varepsilon$. 这和最小列指标的假设矛盾. 所以我们证明了 $A\alpha_1, A\alpha_2, \cdots, A\alpha_t$ 线性无关.

再证 $\alpha_0, \alpha_1, \cdots, \alpha_\varepsilon$ 线性无关. 如果它们线性相关, 则有一组不全为零的常数 $b_0, b_1, \cdots, b_\varepsilon$, 使得 $\sum\limits_{p=0}^{\varepsilon} b_p\alpha_p=0$. 于是由 $A\alpha_0=0$, 则有 $\sum\limits_{p=1}^{\varepsilon} b_pA\alpha_p=0$. 上面已证了 $A\alpha_1, A\alpha_2, \cdots, A\alpha_t$ 线性无关, 所以证明了 $b_1=b_2=\cdots=b_\varepsilon=0$. 因此 $b_0\alpha_0=0$, 但是 $\alpha_0\neq 0$, 所以 $b_0=0$. 这就导出矛盾, 所以证明了 $\alpha_0, \alpha_1, \cdots, \alpha_\varepsilon$ 线性无关. □

引进标准块, 即引进 $k\times(k+1)$ 矩阵

$$L_k=\begin{pmatrix}\lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & -1\end{pmatrix}=\lambda\left(E_k\ 0^{(k,1)}\right)+\left(0^{(k,1)}\ E_k\right), \tag{14.1.8}$$

则有

引理 14.1.7 设 ε 为 $n\times m$ λ 矩阵 $A+\lambda B$ 的最小列指标, 且 $\varepsilon>0$, 则存在 n 阶非异常数方阵 P_1 及 m 阶非异常数方阵 Q_1, 使得

$$P_1(A+\lambda B)Q_1=\operatorname{diag}(L_\varepsilon, A_1+\lambda B_1), \tag{14.1.9}$$

其中 $A_1+\lambda B_1$ 是 $(n-\varepsilon)\times(m-\varepsilon-1)$ λ 矩阵. 又如果 $A_1+\lambda B_1$ 有最小行指标 τ, 那末最小行指标 $\tau\geqslant\varepsilon$.

证 由 ε 的定义, 所以存在齐次线性方程组 $(A+\lambda B)x=0$ 的非零解,

$$\alpha=\alpha_0+\lambda\alpha_1+\lambda^2\alpha_2+\cdots+\lambda^\varepsilon\alpha_\varepsilon,$$

其中 $\alpha_0, \alpha_1, \cdots, \alpha_\varepsilon$ 都是 $m\times 1$ 复常数矩阵, 且 $\alpha_0\neq 0, \cdots, \alpha_\varepsilon\neq 0$.

今 $(A+\lambda B)\alpha=0$ 等价于

$$A\alpha_0=0,\quad A\alpha_1+B\alpha_0=0,\quad \cdots,\quad A\alpha_\varepsilon+B\alpha_{\varepsilon-1}=0,\quad B\alpha_\varepsilon=0.$$

记

$$\widetilde{\alpha}=\begin{pmatrix}\alpha_0\\ \alpha_1\\ \vdots\\ \alpha_\varepsilon\end{pmatrix}\in\mathbb{C}^{(\varepsilon+1)m},$$

则 $(A+\lambda B)\alpha=0$ 等价于

$$\begin{pmatrix} A & & & \\ B & A & & \\ & B & \ddots & \\ & & \ddots & A \\ & & & B \end{pmatrix}\widetilde{\alpha}=0.$$

记 $(k+2)n\times(k+1)m$ 常数矩阵

$$M_k(A+\lambda B)=\begin{pmatrix} A & & & \\ B & A & & \\ & B & \ddots & \\ & & \ddots & A \\ & & & B \end{pmatrix},\quad k=0,1,\cdots,\varepsilon,$$

所以 $(A+\lambda B)\alpha=0$ 等价于

$$M_\varepsilon(A+\lambda B)\widetilde{\alpha}=0.$$

因此, $M_\varepsilon(A+\lambda B)$ 的秩小于 $(\varepsilon+1)m$. 由最小列指标的定义可知 $M_{\varepsilon-1}(A+\lambda B)$ 的秩等于 εm. 所以 $M_k(A+\lambda B)$ 的秩为 $(k+1)m$, $k=0,1,\cdots,\varepsilon-1$.

由此可见, 我们可以这样来计算最小列指标 ε. 先求 $A+\lambda B$ 的秩, 当 $r<m$, 再依次构造 $M_0(A+\lambda B)$, $M_1(A+\lambda B)$, $\cdots$, 使它们的列都是满秩的. 直到第一个 $M_\varepsilon(A+\lambda B)$, 它的列不满秩, 那末 ε 就是 λ 矩阵 $A+\lambda B$ 的最小列指标.

现在开始证明引理. 由引理 14.1.3, α_0, α_1, $\cdots$, $\alpha_\varepsilon\in\mathbb{C}^m$ 线性无关, 所以存在 m 阶常数非异方阵 $Q_2=(\alpha_0,\alpha_1,\cdots,\alpha_\varepsilon,*)$, 又由 $A\alpha_1$, $A\alpha_2$, $\cdots$, $A\alpha_\varepsilon$ 线性无关, 所以又可构造 n 阶常数非异方阵 $P_2^{-1}=-(A\alpha_1,A\alpha_2,\cdots,A\alpha_\varepsilon,*)$. 令

$$\begin{aligned}P_2(A+\lambda B)Q_2&=P_2((A+\lambda B)\alpha_0,(A+\lambda B)\alpha_1,\cdots,(A+\lambda B)\alpha_\varepsilon*)\\&=P_2(-\lambda A\alpha_1,A\alpha_1-\lambda A\alpha_2,\cdots,A\alpha_{\varepsilon-1}-\lambda A\alpha_\varepsilon,A\alpha_\varepsilon,*)\\&=P_2(-\lambda A\alpha_1,-A\alpha_2,\cdots,-A\alpha_\varepsilon,*)\begin{pmatrix}L_\varepsilon & * \\ 0 & *\end{pmatrix}\\&=\begin{pmatrix}L_\varepsilon & A_2+\lambda B_2\\ 0 & A_1+\lambda B_1\end{pmatrix}.\end{aligned}$$

现在来证明: 若 $A_1+\lambda B_1$ 有最小列指标 τ, 则 $\tau\geqslant\varepsilon$.

今无妨设

$$A+\lambda B=\begin{pmatrix}L_\varepsilon & A_2+\lambda B_2\\ 0 & A_1+\lambda B_1\end{pmatrix},$$

于是

$$A=\begin{pmatrix}\begin{pmatrix}0 & -E\end{pmatrix} & A_2\\ 0 & A_1\end{pmatrix},\quad B=\begin{pmatrix}\begin{pmatrix}E & 0\end{pmatrix} & B_2\\ 0 & B_1\end{pmatrix},$$

这里矩阵 A_1 和 B_1 是 $(n-\varepsilon)\times(m-\varepsilon-1)$ 矩阵. 为了证 $\tau\geqslant\varepsilon$, 我们只要证明 $(\varepsilon+1)(n-\varepsilon)\times\varepsilon(m-\varepsilon-1)$ 矩阵 $M_{\varepsilon-1}(A_1+\lambda B_1)$ 的列为满秩的.

今已知矩阵 $(A+\lambda B)$ 的最小列指标为 ε, 即 $M_{\varepsilon-1}(A+\lambda B)$ 的列是满秩的. 由于矩阵 A 和 B 的分块方式相同, 由 $M_{\varepsilon-1}(A+\lambda B)$ 的定义, 对 $M_{\varepsilon-1}(A+\lambda B)$ 用置换方阵作相抵, 则可证

$$M_{\varepsilon-1}(A+\lambda B)\quad 相抵于\quad \begin{pmatrix}M_{\varepsilon-1}(L_\varepsilon) & M_{\varepsilon-1}(A_2+\lambda B_2)\\ 0 & M_{\varepsilon-1}(A_1+\lambda B_1)\end{pmatrix}.$$

今

$$\operatorname{rank}(M_{\varepsilon-1}(A+\lambda B))=\varepsilon m,\qquad \operatorname{rank}(M_{\varepsilon-1}(L_\varepsilon))=\varepsilon(\varepsilon+1),$$

且 $M_{\varepsilon-1}(L_\varepsilon)$ 为 $\varepsilon(\varepsilon+1)$ 阶方阵, 所以它非异. 因此, $\operatorname{rank}(M_{\varepsilon-1}(A_1+\lambda B_1))=\varepsilon(m-\varepsilon-1)$. 由于 $M_{\varepsilon-1}(A_1+\lambda B_1)$ 是 $(n-\varepsilon)(\varepsilon+1)\times\varepsilon(m-\varepsilon-1)$ 矩阵, 所以 $M_{\varepsilon-1}(A_1+\lambda B_1)$ 的列是满秩的. 这证明了 $\tau\geqslant\varepsilon$.

所以余下要证: 存在 $\varepsilon\times(m-\varepsilon)$ 常数矩阵 X 及 $(\varepsilon+1)\times(m-\varepsilon-1)$ 常数矩阵 Y, 使得

$$\begin{pmatrix}E_\varepsilon & X\\ 0 & E_{m-\varepsilon}\end{pmatrix}\begin{pmatrix}L_\varepsilon & A_2+\lambda B_2\\ 0 & A_1+\lambda B_1\end{pmatrix}\begin{pmatrix}E_{\varepsilon+1} & Y\\ 0 & E_{m-\varepsilon-1}\end{pmatrix}=\begin{pmatrix}L_\varepsilon & 0\\ 0 & A_1+\lambda B_1\end{pmatrix}.$$

换句话说, 要证明存在矩阵 X 和 Y, 使得

$$L_\varepsilon Y+(A_2+\lambda B_2)+X(A_1+\lambda B_1)=0.$$

即

$$(0,-E)Y+A_2+XA_1=0,\quad (E,0)Y+B_2+XB_1=0.$$

将 X,Y,A_2,B_2 按照行分块, 即记

$$X=\begin{pmatrix}u_1\\ \vdots\\ u_\varepsilon\end{pmatrix},\quad Y=\begin{pmatrix}v_1\\ \vdots\\ v_{\varepsilon+1}\end{pmatrix},\quad A_2=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_\varepsilon\end{pmatrix},\quad B_2=\begin{pmatrix}\beta_1\\ \vdots\\ \beta_\varepsilon\end{pmatrix},$$

则有

$$v_{j+1}=\alpha_j+u_jA_1,\quad v_j=-\beta_j-u_jB_1,\quad j=1,2,\cdots,\varepsilon.$$

所以有 $v_1=-\beta_1-u_1B_1$, $v_j=\alpha_{j-1}+u_{j-1}A_j=-\beta_j-u_jB_1$, $j=2,3,\cdots,\varepsilon$, $v_{\varepsilon+1}=\alpha_\varepsilon+u_\varepsilon A_1$. 即有

$$(u_1,\cdots,u_\varepsilon)M_{\varepsilon-2}(A_1+\lambda B_1)=-(\alpha_1+\beta_2,\cdots,\alpha_{\varepsilon-1}+\beta_\varepsilon).$$

今 $M_{\varepsilon-2}(A_1+\lambda B_1)$ 为 $\varepsilon(n-\varepsilon)\times(\varepsilon-1)(m-\varepsilon-1)$ 矩阵, 视 $(u_1,u_2,\cdots,u_\varepsilon)$ 为 $\varepsilon(m-\varepsilon)$ 个独立未知数构成的向量. 所以这是 $\varepsilon(m-\varepsilon)$ 个未知数, $(\varepsilon-1)(m-\varepsilon-1)$ 个方程构成的非齐次线性方程组. 而系数矩阵的秩和方程个数相同, 所以一定有解. 任取一组解, 便求出了 $u_1, u_2, \cdots, u_\varepsilon$. 所以求出了 X. 再由 $v_j=-\beta_j-u_jB_1$, $1\leqslant j\leqslant\varepsilon$, $v_{\varepsilon+1}=\alpha_\varepsilon+u_\varepsilon A_1$, 便求出了 $v_1, v_2, \cdots, v_{\varepsilon+1}$, 所以求出了 Y. 至此证明了引理. □

由此引理立即有.

定理 14.1.8 给定 $n\times m$ 矩阵偶 A 和 B, 则 λ 矩阵 $A+\lambda B$ 严格相抵于标准形

$$\operatorname{diag}(0^{(\varepsilon_0,\eta_0)},L_{\varepsilon_1},\cdots,L_{\varepsilon_p},L'_{\eta_1},\cdots,L'_{\eta_q},A_0+\lambda B_0), \tag{14.1.10}$$

其中 $\varepsilon_0=0$, $0^{(0,\eta_0)}$ 表示它的前 η_0 列全为零, $\eta_0=0$, $0^{(\varepsilon_0,0)}$ 表示它的前 ε_0 行全为零. 又

$$\varepsilon_0,\eta_0\geqslant 0,\quad 0<\varepsilon_1\leqslant\varepsilon_2\leqslant\cdots\leqslant\varepsilon_p,\quad 0<\eta_1\leqslant\eta_2\leqslant\cdots\leqslant\eta_q, \tag{14.1.11}$$

又 $A_0+\lambda B_0$ 为 n_0 阶 λ 方阵, 且 $\det(A_0+\lambda B_0)\neq 0$.

证 由引理 14.1.7, 依次作出 $0^{(\varepsilon_0,\eta_0)}$, L_{ε_1}, L_{ε_2}, $\cdots$, L_{ε_p}, 于是余下的 λ 矩形无最小列指标. 这时考虑它的转置矩阵, 即考虑最小行指标. 于是最后所得的 λ 矩阵 $A_0+\lambda_0B_0$ 无最小行指标, 也无最小列指标. 所以 $A_0+\lambda B_0$ 的秩等于行数, 也等于列数, 即为非异方阵. □

这样一来, 求矩阵偶 (A,B) 在相抵下的标准形的问题化为求非异 λ 方阵 $A_0+\lambda B_0$ 在严格相抵下的标准形.

定理 14.1.9 设 $A_0+\lambda B_0$ 为 n_0 阶非异 λ 方阵, 则 $A_0+\lambda B_0$ 严格相抵于标准形:

$$\Omega=\operatorname{diag}(E_{\delta_1}+\lambda N_{\delta_1},\cdots,E_{\delta_s}+\lambda N_{\delta_s},(\lambda+\lambda_1)E_{\tau_1}+N_{\tau_1},\cdots,(\lambda+\lambda_u)E_{\tau_u}+N_{\tau_u}), \tag{14.1.12}$$

其中 $0<\delta_1\leqslant\cdots\leqslant\delta_s$, $0<\tau_1\leqslant\cdots\leqslant\tau_u$, 又

$$N_n=\begin{pmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&0&1\\&&&&0\end{pmatrix}. \tag{14.1.13}$$

证 今 $\det(A_0+\lambda B_0)\neq 0$, 即为 λ 的非零多项式, 所以存在复数 λ_0, 使得 $\det(A_0+\lambda_0B_0)\neq 0$. 而

$$A_0+\lambda B_0=A_0+\lambda_0B_0+(\lambda-\lambda_0)B_0=(A_0+\lambda_0B_0)(E+(\lambda-\lambda_0)(A_0+\lambda_0B_0)^{-1}B_0).$$

将常数方阵 $(A_0+\lambda_0 B_0)^{-1}B_0$ 相似于 Jordan 标准形 J, 这证明了 $A_0+\lambda B_0$ 严格相抵于 $E+(\lambda-\lambda_0)J$. 将 J 中属于特征根零的 Jordan 块拼在一块, 记作 J_0. 将 J 中不属于特征根零的 Jordan 块拼在一起, 记作 J_1. 于是 $J=\mathrm{diag}\,(J_0,J_1)$, 而

$$E+(\lambda-\lambda_0)J=\mathrm{diag}\,(E-\lambda_0 J_0+\lambda J_0, E-\lambda_0 J_1+\lambda J_1).$$

由 $\det(E-\lambda J_0)=1$, 所以

$$E-\lambda_0 J_0+\lambda J_0=(E-\lambda_0 J_0)[E+\lambda(E-\lambda_0 J_0)^{-1}J_0].$$

记 $M=(E-\lambda_0 J_0)^{-1}J_0$, 则 M 是幂零方阵, 所以相似于 Jordan 标准形 $\widetilde{J_0}=\mathrm{diag}\,(N_{\delta_1},\cdots,N_{\delta_s})$.

由 $\det(J_1)\neq 0$, 所以 $(E-\lambda_0 J_1)J_1^{-1}$ 相似于 Jordan 标准形

$$\widetilde{J_1}=\mathrm{diag}\,(\lambda_1 E_{\tau_1}+N_{\tau_1},\cdots,\lambda_u E_{\tau_u}+N_{\tau_u}).$$

所以 $E-\lambda_0 J+\lambda J$ 严格相抵于标准形 (14.1.12). □

于是我们有

定理 14.1.10 给定 $n\times m$ 矩阵偶 A 和 B, 则 λ 矩阵 $A+\lambda B$ 严格相抵于标准形

$$\mathrm{diag}\,(0^{(\varepsilon_0,\eta_0)},L_{\varepsilon_1},\cdots,L_{\varepsilon_p},L'_{\eta_1},\cdots,L'_{\eta_q},\Omega), \tag{14.1.14}$$

其中 $\varepsilon_0=0$, $0^{(0,\eta_0)}$ 表示它的前 η_0 列全为零, $\eta_0=0$, $0^{(\varepsilon_0,0)}$ 表示它的前 ε_0 行全为零. 又

$$\varepsilon_0,\eta_0\geqslant 0,\quad 0<\varepsilon_1\leqslant\varepsilon_2\leqslant\cdots\leqslant\varepsilon_p,\quad 0<\eta_1\leqslant\eta_2\leqslant\cdots\leqslant\eta_q, \tag{14.1.15}$$

又 $A_0+\lambda B_0$ 为 n_0 阶 λ 方阵, 且

$$\begin{aligned}\Omega=\mathrm{diag}\,(&E_{\delta_1}+\lambda N_{\delta_1},\cdots,E_{\delta_s}+\lambda N_{\delta_s},\\&(\lambda+\lambda_1)E_{\tau_1}+N_{\tau_1},\cdots,(\lambda+\lambda_u)E_{\tau_u}+N_{\tau_u}),\\&0<\delta_1\leqslant\cdots\leqslant\delta_s,\quad 0<\tau_1\leqslant\cdots\leqslant\tau_u.\end{aligned} \tag{14.1.16}$$

定理 14.1.11 $n\times m$ λ 矩阵 $A+\lambda B$ 的全部行指标、列指标以及初等因式是它在严格相抵下的全系不变量, 且标准形中对角块

$$(\lambda+\lambda_j)E_{\tau_j}+N_{\tau_j},\quad j=1,2,\cdots,u,$$

若不计次序, 则标准形唯一.

证 记

$$G_1=\mathrm{diag}\,(L_{\varepsilon_1},\cdots,L_{\varepsilon_p}),\qquad G_2=\mathrm{diag}\,(L'_{\eta_1},\cdots,L'_{\eta_q}),\qquad G_3=A_0+\lambda B_0. \tag{14.1.17}$$

则 $A+\lambda B$ 在严格相抵下的标准形为

$$D=\operatorname{diag}(0^{(\varepsilon_0,\tau_0)},G_1,G_2,G_3)=\operatorname{diag}(0^{(\varepsilon_0,\tau_0)},D_1).$$

如果 $A+\lambda B$ 另有标准形

$$\widetilde{D}=\operatorname{diag}(0^{(\widetilde{\varepsilon_0},\widetilde{\tau_0})},\widetilde{G_1},\widetilde{G_2},\widetilde{G_3})=\operatorname{diag}(0^{(\widetilde{\varepsilon_0},\widetilde{\tau_0})},\widetilde{D_1}).$$

下面我们依次来证明: **(i)** $\varepsilon_0=\widetilde{\varepsilon_0}$, $\tau_0=\widetilde{\tau_0}$; **(ii)** $G_1=\widetilde{G_1}$, $G_2=\widetilde{G_2}$; **(iii)** $\widetilde{G_3}$ 由重排 G_3 的对角块次序而得到.

(i) 证 $\varepsilon_0=\widetilde{\varepsilon_0}$, $\tau_0=\widetilde{\tau_0}$.

由于 D 及 $\widetilde{D}$ 都是 $A+\lambda B$ 在严格相抵下的标准形, 即存在 n 阶常数非异方阵 P 及 m 阶常数非异方阵 Q, 使得 $\widetilde{D}=PDQ$. 于是

$$\begin{aligned}\dim\{\,\beta\in\mathbb{C}^m\mid\widetilde{D}\beta=0\,\}&=\dim\{\,\beta\in\mathbb{C}^m\mid PDQ\beta=0\,\}\\&=\dim\{\,\alpha=Q\beta\in\mathbb{C}^m\mid D\alpha=0\,\}\\&=\dim\{\,\alpha\in\mathbb{C}^m\mid D\alpha=0\,\}.\end{aligned}$$

所以 $\widetilde{\varepsilon_0}=\varepsilon_0$. 同理可证 $\widetilde{\tau_0}=\tau_0$.

(ii) 证 D 和 $\widetilde{D}$ 严格相抵.

将 D 及 $\widetilde{D}$ 分成四块, 有

$$\begin{pmatrix}P_1^{(\varepsilon_0)}&P_2\\P_3&P_4\end{pmatrix}\begin{pmatrix}0^{(\varepsilon_0,\tau_0)}&0\\0&D_1\end{pmatrix}\begin{pmatrix}Q_1^{(\tau_0)}&Q_2\\Q_3&Q_4\end{pmatrix}=\begin{pmatrix}0^{(\widetilde{\varepsilon_0},\widetilde{\tau_0})}&0\\0&\widetilde{D_1}\end{pmatrix}.$$

即

$$P_2D_1Q_3=0,\quad P_2D_1Q_4=0,\quad P_4D_1Q_3=0,\quad P_4D_1Q_4=\widetilde{D_1}.$$

如果 $\det(P_4)=0$, 则存在非零向量 $\alpha\in\mathbb{C}^{n-\varepsilon_0}$, 使得 $\alpha'P_4=0$. 于是 $\alpha'\widetilde{D_1}=0$. 即 $\widetilde{D_1'}\alpha=0$. 由 ε_0 和 τ_0 的定义可知 $\alpha=0$. 这导出矛盾, 所以 $\det(P_4)\neq0$. 同理可证 $\det(Q_4)\neq0$. 所以 D_1 和 $\widetilde{D_1}$ 严格相抵.

(iii) 证 $\widetilde{G_1}=G_1$, 而且 $\operatorname{diag}(G_2,G_3)$ 和 $\operatorname{diag}(\widetilde{G_2},\widetilde{G_3})$ 严格相抵.

由于最小列指标在严格相抵下不变. 所以 $\widetilde{\varepsilon_1}=\varepsilon_1$. 设 $\varepsilon_1=\cdots=\varepsilon_{n_1}<\varepsilon_{n_1+1}$, $\widetilde{\varepsilon_1}=\cdots=\widetilde{\varepsilon_{m_1}}<\widetilde{\varepsilon_{m_1+1}}$. 无妨设 $m_1\leqslant n_1$. 所以可以记

$$D_1=\operatorname{diag}(L_{\varepsilon_1},\cdots,L_{\varepsilon_{n_1}},D_2),\qquad\widetilde{D_1}=\operatorname{diag}(L_{\widetilde{\varepsilon_1}},\cdots,L_{\widetilde{\varepsilon_{m_1}}},\widetilde{D_2}),$$

下面证 D_2 和 $\widetilde{D_2}$ 严格相抵. 再用归纳法, 便可证明断言.

为此, 将 P 及 Q^{-1} 分成四块. 由 $PD_1=D_1Q^{-1}$, 有

$$\begin{pmatrix}P_1&P_2\\P_3&P_4\end{pmatrix}\operatorname{diag}(L_{\varepsilon_1},\cdots,L_{\varepsilon_{m_1}};D_2)=\operatorname{diag}(L_{\widetilde{\varepsilon_1}},\cdots,L_{\widetilde{\varepsilon_{m_1}}};\widetilde{D_2})\begin{pmatrix}Q_1&Q_2\\Q_3&Q_4\end{pmatrix}.$$

所以有

$$P_1\mathrm{diag}\,(L_{\varepsilon_1},\cdots,L_{\varepsilon_{m_1}})=\mathrm{diag}\,(L_{\widetilde{\varepsilon_1}},\cdots,L_{\widetilde{\varepsilon_{m_1}}})Q_1,\quad P_4D_2=\widetilde{D_2}Q_4$$

$$P_2D_2=\mathrm{diag}\,(L_{\widetilde{\varepsilon_1}},\cdots,L_{\widetilde{\varepsilon_{m_1}}})Q_2,\qquad P_3\mathrm{diag}\,(L_{\varepsilon_1},\cdots,L_{\varepsilon_{m_1}})=\widetilde{D_2}Q_3.$$

作 ε_1+1 维复向量 α, 使得 $\alpha=(1,\lambda,\cdots,\lambda^{\varepsilon_1})'$, 则 $L_{\varepsilon_1}\alpha=0$. 取 m_1 个独立自变量 $x_1, x_2, \cdots, x_{m_1}$, 则

$$\mathrm{diag}\,(L_{\varepsilon_1},\cdots,L_{\varepsilon_{m_1}})(x_1\alpha',\cdots,x_{m_1}\alpha')'=0.$$

即有

$$\widetilde{D_2}Q_3\begin{pmatrix}x_1\alpha\\ \vdots\\ x_{m_1}\alpha\end{pmatrix}=0.$$

但是 $\widetilde{D_2}$ 的最小列指标 $\widetilde{\varepsilon_{m_1+1}}>\varepsilon_1$, 而 $Q_3\begin{pmatrix}x_1\alpha\\ \vdots\\ x_{m_1}\alpha\end{pmatrix}$ 的次数小于等于 ε_1. 这证明了它等于零. 注意到 λ 为独立参数, 而 Q_3 为常数矩阵. 所以证明了 $Q_3=0$. 于是 $P_3\mathrm{diag}\,(L_{\varepsilon_1},\cdots,L_{\varepsilon_{m_1}})=0$. 因此证明了 $P_3=0$. 所以 $\det(P_4)\det(Q_4)\neq 0$. 即 D_2 和 $\widetilde{D_2}$ 严格相抵.

(iv) 证 $\widetilde{G_2}=G_2$, 而且 $\widetilde{G_3}$ 和 G_3 严格相抵.

事实上, 由 **(iii)**, $\mathrm{diag}\,(G_2,G_3)$ 和 $\mathrm{diag}\,(\widetilde{G_2},\widetilde{G_3})$ 是严格相抵的. 所以 $\mathrm{diag}\,(G_2',G_3')$ 和 $\mathrm{diag}\,(\widetilde{G_2'},\widetilde{G_3'})$ 也是严格相抵的. 由 **(iii)** 便证明了 $\widetilde{G_2'}=G_2'$, 且 $\widetilde{G_3'}$ 和 G_3' 严格相抵. 因此 $\widetilde{G_2}=G_2$, 且 $\widetilde{G_3}$ 和 G_3 严格相抵.

(v) 设非异 λ 方阵 $A_0+\lambda B_0$ 和 $C_0+\lambda D_0$ 严格相抵, 则标准形不计对角块次序唯一.

最后, 设非异 λ 方阵 $A_0+\lambda B_0$ 和 $C_0+\lambda D_0$ 严格相抵. 由 $\det(A_0+\lambda B_0)\det(C_0+\lambda D_0)\neq 0$, 所以存在复数 λ_0, 使得 $det(A_0+\lambda_0 B_0)\det(C_0+\lambda_0 D_0)\neq 0$. 记 $\tau=\lambda-\lambda_0$, 于是

$$E+\tau B_0(A_0+\lambda_0B_0)^{-1}=E+\tau G_0,\quad E+\tau D_0(C_0+\lambda_0D_0)^{-1}=E+\tau H_0$$

严格相抵. 我们来证 G_0 和 H_0 相似. 事实上, 存在非异常数复方阵 P_0 和 Q_0, 使得 $E+\tau H_0=P_0(E+\tau G_0)Q_0$. 于是 $P_0Q_0=E$, 即 $Q_0=P_0^{-1}$. 又 $H_0=P_0G_0Q_0=P_0G_0P_0^{-1}$., 即 H_0 和 G_0 相似. 由定理 14.1.9 可知, $A_0+\lambda B_0$ 和 $C_0+\lambda D_0$ 分别由 G_0 及 H_0 的 Jordan 标准形派生出来的. 由 G_0 和 H_0 相似, 可知它们的 Jordan 标准形相同 (在特征根非零的 Jordan 块不计次序的意义下). 这证明了定理. □

习　　题　14.1

14.1.1 试求下列矩阵偶在相抵下的标准形:

$$A=\begin{pmatrix}1&2&3&4&5\\2&4&6&8&10\\0&1&0&1&0\\1&0&1&0&1\end{pmatrix},\qquad B=\begin{pmatrix}0&1&0&1&0\\1&0&1&0&1\\5&4&3&2&1\\10&8&6&4&2\end{pmatrix};$$

$$A=\begin{pmatrix}1&&&&\\&1&&&\\&&1&&\\&&&1&\\&&&&1\end{pmatrix},\qquad B=\begin{pmatrix}&&&&1\\&&&1&\\&&1&&\\&1&&&\\1&&&&\end{pmatrix}.$$

14.1.2 试求矩阵偶 (A,B) 在相抵下的标准形, 其中 $A=xy'$, $B=xz'$, y 和 z 都是 $n\times 1$ 矩阵, 而 x 是 $m\times 1$ 矩阵.

14.1.3 试求矩阵偶 (A,B) 在相抵下的标准形, 其中 $A=E_n$,

$$B=\begin{pmatrix}&&&1\\&&\cdot&\\&\cdot&&\\1&&&\end{pmatrix}.$$

§14.2 复对称及反对称方阵偶在相合下的标准形 *

定义 14.2.1 域 $\mathbb{F}$ 上两个 n 阶方阵偶 (A,B) 和 (C,D) 称为**相合的**, 如果存在域 $\mathbb{F}$ 上一个 n 阶非异方阵 P, 使得

$$C=PAP',\quad D=PBP';\tag{14.2.1}$$

称为**复相合的**, 如果

$$C=PAP^*,\quad D=PBP^*.\tag{14.2.2}$$

对两个 n 阶实方阵偶 (A,B) 和 (C,D), 复相合即为相合.

在这一节考虑如下情形, 即同时为对称方阵或同时为反对称方阵的方阵偶, 它们在相合下的标准形.

定理 14.2.2 对称 (反对称) 方阵偶 (A,B) 和 (C,D) 相合的必要且充分条件为对称 (反对称) 方阵严格相抵.

证 今 (A,B) 和 (C,D) 相合, 即存在非异复方阵 P, 使得 $C=P'AP$, $D=P'BP$. 所以, $C+\lambda D=P'(A+\lambda B)P$. 即 $A+\lambda B$ 和 $C+\lambda D$ 严格相抵.

反之, 若存在非异复方阵 P 和 Q, 使得 $P(A+\lambda B)Q=C+\lambda D$. 双方取转置, 有 $Q'(A+\lambda B)P'=C+\lambda D$. 所以 $P(A+\lambda B)Q=Q'(A+\lambda B)P'$. 取 $T=P^{-1}Q'$, 即有 $(A+\lambda B)T'=T(A+\lambda B)$. 由 §7.3 可知, 存在多项式 $p(\lambda)$, 使得 $p(T)^2=T$, 又 $(A+\lambda B)p(T)'=p(T)(A+\lambda B)$. 即 $P(T)^{-1}(A+\lambda B)=(A+\lambda B)p(T')^{-1}$. 今 $\det(T)=\det(P)^{-1}\det(Q)\neq 0$, 所以 $\det(p(T))\neq 0$. 令 $T_1=Q'p(T)^{-1}$, 则有

$$\begin{aligned}C+\lambda D&=P(A+\lambda B)Q=Q'T^{-1}(A+\lambda B)Q=Q'p(T)^{-2}(A+\lambda B)Q\\&=Q'p(T)^{-1}(A+\lambda B)(p(T)')^{-1}Q=T_1(A+\lambda B)T_1'.\end{aligned}$$

由 $\det(T_1)\neq 0$, 便证明了定理. □

由这个定理可知, 对称 (反对称) 方阵偶 (A,B) 在相合下的全系不变量为方阵偶 (A,B) 在相抵下的全系不变量. 利用这点, 所以问题化为先求出对称 (反对称) 方阵偶的全系不变量, 再寻找一个最简单的对称 (反对称) 方阵偶, 它具有相同的全系不变量. 这样做就建立了对称 (反对称) 方阵偶在相合下的标准形理论.

引理 14.2.3 设 λ 方阵 $A+\lambda B$ 对称 (反对称), 则它在严格相抵下的标准形为

$$\begin{aligned}&\mathrm{diag}\,(0_{\varepsilon_0},L_{\varepsilon_1},\cdots,L_{\varepsilon_p},L'_{\varepsilon_1},\cdots,L'_{\varepsilon_p},\\&(E+\lambda N)^{(\delta_1)},\cdots,(E+\lambda N)^{(\delta_s)},(\lambda+\lambda_1)E_{\tau_1}+N,\cdots,(\lambda+\lambda_u)E_{\tau_u}+N).\end{aligned}\tag{14.2.3}$$

证 由定理 14.1.8, $A+\lambda B$ 以及 $(A+\lambda B)'=\pm(A+\lambda B)$ 分别严格相抵于

$$\mathrm{diag}\,(0^{(\varepsilon_0,\eta_0)},L_{\varepsilon_1},\cdots,L_{\varepsilon_p},L_{\eta_1},\cdots,L_{\eta_q},A_0'+\lambda B_0'),\tag{14.2.4}$$

$$\mathrm{diag}\,(0^{(\varepsilon_0,\eta_0)},L'_{\varepsilon_1},\cdots,L'_{\varepsilon_p},L'_{\eta_1},\cdots,L'_{\eta_q},A_0+\lambda B_0),\tag{14.2.5}$$

这证明了

$$\eta_0=\varepsilon_0,\eta_1=\varepsilon_1,\cdots,\eta_p=\varepsilon_p,q=p.\tag{14.2.6}$$

□

定理 14.2.4 对称 λ 方阵 $A+\lambda B$ 在相合下的标准形为

$$\begin{aligned}\Omega=\mathrm{diag}\,&\left(0_{\varepsilon_0},\begin{pmatrix}0&L_{\varepsilon_1}\\L'_{\varepsilon_1}&0\end{pmatrix},\cdots,\begin{pmatrix}0&L_{\varepsilon_p}\\L'_{\varepsilon_p}&0\end{pmatrix},(E+\lambda N)^{(\delta_1)}S^{(\delta_1)},\cdots,\right.\\&\left.(E+\lambda N)^{(\delta_s)}S^{(\delta_s)},((\lambda+\lambda_1)E_{\tau_1}+N)S^{(\tau_1)},\cdots,((\lambda+\lambda_u)E_{\tau_u}+N)S^{(\tau_u)}\right),\end{aligned}$$

其中

$$S=\begin{pmatrix}&&&1\\&&\cdot&\\&\cdot&&\\1&&&\end{pmatrix}.\tag{14.2.7}$$

证 由引理 14.2.3, 所以 $A+\lambda B$ 严格相抵于

$$\operatorname{diag}\left(0_{\varepsilon_0},\begin{pmatrix}0 & L_{\varepsilon_1}\\ L'_{\varepsilon_1} & 0\end{pmatrix},\cdots,\begin{pmatrix}0 & L_{\varepsilon_p}\\ L'_{\varepsilon_p} & 0\end{pmatrix},A_0+\lambda B_0\right),\tag{14.2.8}$$

其中 $A_0+\lambda B_0$ 仍对称. 由定理 14.2.2, 所以 $A+\lambda B$ 严格相抵于对称 λ 方阵 Ω. 由定理 14.2.2 和 Ω 相合. □

引理 14.2.5 给定 Jordan 标准形 $J=\operatorname{diag}(J_1^{(e_1)},\cdots,J_t^{(e_t)})$, 其中 $J_k^{(e_k)}=\lambda_k E_{e_k}+N_{e_k}$, $k=1,2,\cdots,t$. 如果存在非异反对称方阵 K, 使得 $KJ'=JK$, 则 J 的初等因式成对出现.

证 将 K 和 J 一样分块, 记作

$$K=\begin{pmatrix}K_{11} & \cdots & K_{1t}\\ \vdots & & \vdots\\ K_{t1} & \cdots & K_{tt}\end{pmatrix},\qquad K_{ij}=\begin{pmatrix}b_{11}^{(ij)} & \cdots & b_{1v}^{(ij)}\\ \vdots & & \vdots\\ b_{u1}^{(ij)} & \cdots & b_{uv}^{(ij)}\end{pmatrix},$$

于是有

$$J_pK_{pq}=K_{pq}J'_q,\quad p,q=1,2,\cdots,t.$$

今若

$$(\lambda_pE_p+N_p)\begin{pmatrix}b_{11}^{(pq)} & \cdots & b_{1v}^{(pq)}\\ \vdots & & \vdots\\ b_{u1}^{(pq)} & \cdots & b_{uv}^{(pq)}\end{pmatrix}=\begin{pmatrix}b_{11}^{(pq)} & \cdots & b_{1v}^{(pq)}\\ \vdots & & \vdots\\ b_{u1}^{(pq)} & \cdots & b_{uv}^{(pq)}\end{pmatrix}(\lambda_qE_q+N'_q),$$

即有

$$\begin{aligned}
&\lambda_pb_{i-1,j}+b_{ij}=\lambda_qb_{i-1,j}+b_{i-1,j+1},\quad 2\leqslant i\leqslant p,\quad 1\leqslant j\leqslant q-1,\\
&\lambda_pb_{iq}+b_{i+1,q}=\lambda_qb_{iq},\quad 1\leqslant i\leqslant p-1,\\
&\lambda_pb_{pj}=\lambda_qb_{pj}+b_{p,j+1},\quad 1\leqslant j\leqslant q-1,\\
&(\lambda_p-\lambda_q)b_{pq}=0.
\end{aligned}$$

当 $\lambda_p\neq\lambda_q$, 则证明了 $b_{kj}=0$, $1\leqslant k\leqslant p$, $1\leqslant j\leqslant q$; 当 $\lambda_p=\lambda_q$, 则证明了 $b_{p2}=\cdots=b_{pq}=0$, $b_{2q}=\cdots=b_{pq}=0$, $b_{ij}=b_{i-1,j+1}$, $2\leqslant i\leqslant p$, $1\leqslant j\leqslant q-1$. 所以证明了当 $\lambda_p\neq\lambda_q$ 时有 $K_{pq}=0$; 当 $\lambda_p=\lambda_q$ 时有

$$K_{pq}=\begin{pmatrix}\begin{pmatrix}b_{pq}^{(1)} & & b_{pq}^{(\min(e_p,e_q))}\\ \vdots & \cdot^{\cdot^{\cdot}} & \\ b_{pq}^{(\min(e_p,e_q))} & & \end{pmatrix} & 0\\ 0 & 0\end{pmatrix}.$$

所以将特征根相同的 Jordan 块放在一起, 从而无妨设 J **有相同的特征根, 即有** $\boldsymbol{\lambda_1}$ $=\cdots=\lambda_t=\lambda_0$. 这时无妨设 $e_1\geqslant\cdots\geqslant e_t>0$. 特别 K_{jj} 为对称方阵. 由 $K=-K'$, 所以 $K_{jj}=0$. 今若 $e_1>e_2$, 则 $(K_{11},\cdots,K_{1t})$ 的第 e_1 行全为零, 这和非异矛盾. 所以 $e_1=e_2$. 设

$$e_1=\cdots=e_r>e_{r+1}\geqslant\cdots\geqslant e_t>0.$$

考虑 $(K_{11},\cdots,K_{1r}),(K_{1,r+1},\cdots,K_{1t})$. 由 $K_{11}=0$, $(K_{1,r+1},\cdots,K_{1t})$ 的第 e_1 行为零, 所以 $(K_{12},\cdots,K_{1r})$ 的第 e_1 行不等于零. 因此存在 K_{1j} 非异. 对 Jordan 块整块作初等变换, 所以无妨设 K_{12} 非异. 因此 $2e_1$ 阶反对称方阵 $\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}$ 非异. 记

$$K=\begin{pmatrix}\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix} & U\\ -U' & K_1\end{pmatrix},\quad J_0=\begin{pmatrix}N_{e_1} & 0\\ 0 & N_{e_1}\end{pmatrix},$$

$$\widetilde{J}=J_0=\begin{pmatrix}N_{e_2} & & \\ & \ddots & \\ & & N_{e_t}\end{pmatrix}.$$

则由 $JK=KJ'$, 有

$$\operatorname{diag}(N_{e_1},\cdots,N_{e_t})K=K\operatorname{diag}(N_{e_1},\cdots,N_{e_t})',$$

所以有

$$J_0\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}=\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}J_0',\quad J_0U=U\widetilde{J}',\quad \widetilde{J}K_1=K_1\widetilde{J}'.$$

而

$$\det(K)=\det\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}\det(\widetilde{K}),$$

其中

$$\widetilde{K}=K_1+U'\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}^{-1}U.$$

所以 $\widetilde{K}$ 为非异反对称方阵. 又

$$\begin{aligned}\widetilde{J}\widetilde{K}&=\widetilde{J}K_1+\widetilde{J}U'\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}^{-1}U\\ &=K_1\widetilde{J}'+U'\begin{pmatrix}0 & K_{12}\\ -K_{12}' & 0\end{pmatrix}^{-1}U\widetilde{J}'=\widetilde{K}\widetilde{J}.\end{aligned}$$

对 Jordan 标准形 $\tilde{J}$ 和非异反对称方阵同样讨论. 便证明了 $e_3 = e_1$. 所以由归纳法便证明了引理. □

定理 14.2.6 反对称 λ 方阵 $A+\lambda B$ 在相合下的标准形为

$$\begin{aligned}\Omega = \operatorname{diag}\Big(&0_{\varepsilon_0}, \begin{pmatrix} 0 & L_{\varepsilon_1} \\ L'_{\varepsilon_1} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & L_{\varepsilon_p} \\ L'_{\varepsilon_p} & 0 \end{pmatrix},\\ &\begin{pmatrix} 0 & (E+\lambda N)S^{(\delta_1)} \\ -(E+\lambda N)S^{(\delta_1)} & 0 \end{pmatrix}, \cdots,\\ &\begin{pmatrix} 0 & (E+\lambda N)S^{(\delta_s)} \\ -(E+\lambda N)S^{(\delta_s)} & 0 \end{pmatrix},\\ &\begin{pmatrix} 0 & ((\lambda+\lambda_1)E+N)S^{(\tau_1)} \\ -((\lambda+\lambda_1)E+N)S^{(\tau_1)} & 0 \end{pmatrix}, \cdots,\\ &\begin{pmatrix} 0 & ((\lambda+\lambda_u)E+N)S^{(\tau_u)} \\ -((\lambda+\lambda_u)E+N)S^{(\tau_u)} & 0 \end{pmatrix}\Big).\end{aligned}$$

证 显然, $A+\lambda B$ 严格相抵于

$$\operatorname{diag}\left(0_{\varepsilon_0}, \begin{pmatrix} 0 & L_{\varepsilon_1} \\ -L'_{\varepsilon_1} & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & L_{\varepsilon_p} \\ -L'_{\varepsilon_p} & 0 \end{pmatrix}, A_0+\lambda B_0\right),$$

其中 $A_0+\lambda B_0$ 为非异反对称方阵. 取复数 λ_0, 使得 $\det(A_0+\lambda_0 B_0) \neq 0$, 则 $A_0+\lambda B_0$ 严格相抵于 $E+(\lambda-\lambda_0)J$, 其中 J 为 Jordan 标准形. 即存在非异常数复方阵 P_0, Q_0, 使得 $A_0+\lambda B_0 = P_0(E+(\lambda-\lambda_0)J)Q_0$. 由 $(A_0+\lambda B_0)' = -(A_0+\lambda B_0)$, 即有 $P_0(E+(\lambda-\lambda_0)J))Q_0 = -Q_0'(E+(\lambda-\lambda_0)J)P_0'$. 因此, $P_0Q_0 = -Q_0'P_0'$, $P_0JQ_0 = -Q_0'J'P_0'$. 所以 $J[Q_0(P_0^{-1})'] = -P_0^{-1}Q_0'J' = [Q_0(P_0^{-1})']J'$. 令 $K = Q_0(P_0')^{-1}$ 反对称, 即有 $JK = KJ'$. 由引理 14.2.5, 所以 J 的初等因式成对出现. 这证明了 $A+\lambda B$ 和 Ω 严格相抵. 由定理 14.2.2 及 $A+\lambda B$ 和 Ω 同为反对称 λ 方阵, 所以 $A+\lambda B$ 和 Ω 相合. □

上面的定理, 是在复数范围内讨论的. 在实的情形, 由于一个实方阵 A 不一定能表成实多项式 $p(A)$ 的平方. 所以定理 14.2.2 不能形式推广. 因此对实的对称方阵偶及反对称方阵偶, 在实相合下的标准形比较复杂. 同样, 对 Hermite 方阵偶在复相合下的标准形, 也因为复方阵 A 虽可开方为 $p(A)$, 但是 $\overline{p(A)}$ 一般不等于 $p(\overline{A})$. 因此定理 14.2.2 仍不能直接推广. 这些标准形理论是经典结果. 但已超出本书范围. 下面仅叙述结果.

定理 14.2.7 实对称 λ 方阵 $A+\lambda B$ (或 Hermite λ 方阵) $A+\lambda B$ 的行指标和

列指标相等. 又初等因式组为

$$\lambda^{\delta_1},\cdots,\lambda^{\delta_s},\quad (\lambda+\lambda_1)^{\tau_1},\cdots,(\lambda+\lambda_u)^{\tau_u},$$
$$(\lambda+\mu_1)^{\rho_1},\cdots,(\lambda+\mu_t)^{\rho_t},(\lambda+\overline{\mu_1})^{\rho_1},\cdots,(\lambda+\overline{\mu_t})^{\rho_t},$$

其中 $\lambda_1,\cdots,\lambda_u$ 实, $\mu_1,\cdots,\mu_t$ 复, 且虚部不等于零.

定理 14.2.8 实对称方阵偶 (A,B) 决定的 λ 方阵 $A+\lambda B$ 在实相合下的标准形为实对称 λ 方阵, 它是准对角形. 对角块由下面五种标准块构成:

$$0_{\varepsilon_0},\quad \begin{pmatrix} 0 & L_\eta \\ L_\eta' & 0 \end{pmatrix},\quad \varepsilon(E+\lambda N)S^{(\delta)},\quad \varepsilon((\lambda+\lambda_0)E+N)S^{(\xi)},$$

$$\varepsilon\rho\begin{pmatrix} (\lambda E+M\cos(\theta))S^{(\tau)} & -(M\sin(\theta))S^{(\tau)} \\ -(M\sin(\theta))S^{(\tau)} & -(\lambda E+M\cos(\theta))S^{(\tau)} \end{pmatrix},$$

其中 λ_0 为实特征根, $\rho\exp(\sqrt{-1}\theta)$ 为复特征根, $\rho>0$, $0\leqslant\theta\leqslant 2\pi$, 又 ε 取 ±1 两值之一, 称为**标签**,

$$M=\begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}=E+N.$$

于是实对称方阵偶的行指标组以及附以标签的初等因式组为它在实相合下的全系不变量.

定理 14.2.9 Hermite 方阵偶 (A,B) 决定的 λ 方阵 $A+\lambda B$ 在复相合下的标准形为 Hermite λ 方阵, 它是准对角形. 对角块由下面五种标准块构成:

$$0_{\varepsilon_0},\quad \begin{pmatrix} 0 & L_\eta \\ L_\eta' & 0 \end{pmatrix},\quad \varepsilon(E+\lambda N)S^{(\delta)},\quad \varepsilon((\lambda+\lambda_0)E+N)S^{(\xi)},$$

$$\varepsilon\rho\begin{pmatrix} 0 & ((\lambda+\mu_0)E+N)S^{(\tau)} \\ ((\lambda+\overline{\mu_0})E+N)S^{(\tau)} & 0 \end{pmatrix},$$

其中 λ_0 为实特征根, μ_0 为复特征根, ε 取 ±1 两值之一, 称为**标签**,

$$S=\begin{pmatrix} & & & & 1 \\ & & & \cdot & \\ & & \cdot & & \\ & \cdot & & & \\ 1 & & & & \end{pmatrix},$$

于是 Hermite 方阵偶的行指标组以及附以标签的初等因式组为它在复相合下的全系不变量.

习　　题　14.2

试证: 两个 n 阶复方阵 A 和 B 复相合的充分且必要条件为复方阵偶 $\left(\frac{1}{2}(A+A'), \frac{1}{2}(A-A')\right)$ 和 $\left(\frac{1}{2}(B+B'), \frac{1}{2}(B-B')\right)$ 相抵.

名词索引

D

E

F

G