线性代数与几何

赵连昌.刘晓东

高等教育出版社

内容提要

本书是教育部"高等教育面向21世纪教学内容和课程体系改革计划"的研究成果,是面向21世纪课程教材和教育部工科数学学科"九五"规划教材,本书共七章:向量代数与向量空间,空间解析几何,矩阵与行列式,线性方程组,线性变换,特征值、特征向量、矩阵对角化,二次型. 附录包括:群、环、域,应用实例——投入产出综合平衡的数学模型,Jordan标准形及四个定理证明.

本书可作为高等学校工科各专业本科的教科书,也可作为非数学类理科专业的教材或参考书.

图书在版编目(CIP)数据

线性代数与几何 赵连昌 刘晓东 .—北京: 高等教育出版社,2001

ISBN 7 - 04 - 009472 - X

Ⅰ. 线… Ⅱ. ①赵…②刘… Ⅲ. ①线性代数②几何 Ⅳ. 013

中国版本图书馆CIP数据核字(2001)第01103号

责任编辑 文小西 李 陶 封面设计 张 楠 责任绘图 黄建英
版式设计 马静如 责任校对 杨雪莲 责任印制

线性代数与几何
赵连昌 刘晓东

出版发行 高等教育出版社
社 址 北京市东城区沙滩后街55号 邮政编码 100009
电 话 010 - 64054588 传 真 010 - 64014048
网 址 http: www .hep .edu .cn
http: www .hep .com .cn
经 销 新华书店北京发行所
印 刷

开 本 787×960 1 16 版 次 年 月第 版
印 张 15.5 印 次 年 月第 次印刷
字 数 280 000 定 价 13.50元

前　言

在 17 世纪,笛卡儿及费马在几何空间中引入了坐标系,从而在几何与代数间建立了一座桥梁,用代数方法解决空间的几何问题,产生了解析几何.人们也注意到,对变量不多于三个的某些代数问题,如果将其解释为相应的几何问题,有助于代数问题的解决;当处理变量个数多于三个的问题时,直观的几何解释不再存在,但是数学家从几何学的经验中汲取直觉,把几何空间的向量运算规律抽象出来,形成了有限维向量空间理论,建立了空间基的概念,将坐标系的概念推广到抽象的线性空间中.历史上,几何与代数互为问题,互为方法,相互交融形成各自学科.在计算机广泛应用的今天,计算机图形学、计算机辅助设计、虚拟现实等技术都以几何与线性代数为其理论和算法基础的一部分.从历史渊源与现实需求两方面考虑,在工科数学中,将解析几何从高等数学中分离出来,与线性代数结合在一起构成一门课程,是工科数学教学内容与课程体系的改革趋势之一.本书在将解析几何与线性代数融为一体方面作了初步尝试,这有利于工科学生学习数学,对于正确应用数学方法也会有所启发.

关于本书内容说明如下.

1. 本书中线性代数与解析几何的基本内容,大致与“教学基本要求”相同.本书的内容是以向量空间与线性变换为主线展开的.

第一章向量代数与向量空间,介绍了向量代数与坐标系,为解析几何准备必要工具,也为抽象的线性空间的引入准备了直观背景.在讨论向量空间时,强调几何空间例子的引导作用,以化解在学习向量空间抽象概念时遇到的困难;同时也强调严格的数学证明,以提高学生的数学素养,帮助学生实现由“初等数学”到“高等数学”的过渡.

第二章空间解析几何,介绍了空间解析几何的基本内容,如果学生在高等数学中学习了向量代数及空间解析几何,可不学本书第一章第二节与第二章的内容,余下内容在逻辑上依然是连贯的.

第三章矩阵与行列式,以矩阵为主,行列式放在次要的位置.为了便于理解和节省学时,使用了归纳的方法定义行列式.在本书中,强调了分块矩阵在矩阵运算中的作用.

第四章线性方程组,除基本定理外,多少突出了消元法的重要性.作为选学内容,涉及一点主元消元法,以简略说明本课程与线性代数数值方法之间的联系;还涉及一点计算复杂性,以说明计算方法的重要性.

第五章线性变换,主要讲线性变换与其矩阵表示的关系,零空间、象空间与线性

变换的运算都作为选学内容.

第六章特征值、特征向量与矩阵对角化,主要讲特征值与特征向量的基本性质及对角化具体方法.

第七章二次型,主要讲二次型与正定矩阵,二次曲面分类是选学内容,在必学内容中讲述了将一个二次曲面方程化为标准形的方法.

附录,只供教师与学生参考.因为这本教材的体系与某些教材的体系有所不同,必须把一些定理证明补全,以形成完整内容,便于查阅.群、环、域,投入产出法与Jordan标准形供学生课外阅读.

2. 本书可供重点院校学时较少专业及非重点院校有关专业使用.为了适应不同专业、不同层次的教学要求,教材显现"模块式"结构,可以按学时不同分块使用:(1)没有星号的内容是基本内容,正常使用约需50学时左右,学生只选作习题A;(2)时间再多6学时左右,可讲带星号的内容,带星号的证明不讲,学生只选作习题A;(3)对学时更多专业,可选讲带星号证明;(4)习题B相当于或高于工科硕士研究生入学考试水平,仅供参考,习题B不属于期末考试内容.

3. 本书在阐述每个重要概念与定理前,常常用具体例子为先导,使学生从实例中了解问题由来,掌握解决问题的思路和算法步骤,以减少理解上的障碍,并可以节省学时.内容论述力求详细严谨,清楚易懂,益于自学.

4. 本书章节之间也常常穿插一些问题由来及思考方法的说明,这只是为了学生更好地理解教材内容,难免有片面之处,希望能起到抛砖引玉的作用.

本书是教育部《高等教育面向21世纪教学内容和课程体系改革计划》立项项目《工科数学系列课程教学内容与体系改革的研究与实践》的研究成果之一.在教材评审中,全国工科数学课程指导委员会主任马知恩教授及许多委员提出了很多宝贵意见,特别是主审人骆承钦教授和戴天时教授、高等教育出版社的文小西编审对教材体系及内容的改进提出一系列的具体建议.他们的建议对本书的形成和保证教材的质量起到十分重要的作用.在本书编写与试讲的几年之中,得到大连理工大学施光燕教授、孙丽华教授、王天明教授、夏尊铨教授、冯红副教授、吉林工业大学董加礼教授及大连水产学院的查健禄教授很多指导与帮助.在本书编写过程中,大连海事大学教务处予以立项并给予资助,教材科对本书试用本的三次印刷予以支持,数学教研室很多教师给予帮助,特别是杜祖缔教授、魏华副教授、王德强博士、张运杰副教授、邵方明副教授和卢玉贞副教授给予很多鼓励与帮助.在此向他们表示衷心的感谢!同时,我们也要感谢书后参考文献的作者们,他们的著作给予本书编者很多启发与借鉴.

由于编者水平有限,不妥与错误之处在所难免,敬请同行与读者批评指正.

编者

2000年8月于大连

目　录

第一章　向量代数与向量空间

本章研究向量代数与向量空间.所谓向量就是既有大小又有方向的量.很多物理量需要用向量表示,同时向量也是空间中基本的几何量.引入向量的加法、数与向量乘法之后,本书将以向量为工具建立坐标系.坐标系的引入是 R. Descartes(1596—1650)对数学的开创性的贡献.坐标系是把数学中不同分支联系起来的一座"桥梁".坐标系的建立使得向量与它的坐标建立一一对应关系,从而也使空间中的点与它的坐标建立一一对应关系,使得空间曲面和曲线可用方程或方程组表示,因而可用代数方法来研究几何问题;反过来,也可用几何方法来研究很多代数问题.几何与代数互为工具、相互融合是数学思想统一性的佐证.向量空间是以几何向量及其运算为直观背景作进一步抽象的结果,可应用它研究比几何向量更复杂的对象及其运算的规律.在讲述向量代数与向量空间之前,对集合及其上的映射略加陈述,以便了解必要术语与记号.

第一节　集合　二元关系　映射

1.1　集合

集合论的概念与方法是近代数学的基础.由于集合是数学的一个最基本的概念,所以不能用比它更简单的概念来定义,只能对它作一些解释.在此,仅引用数学名著 Van der Waerden 的《代数学》中关于集合的解释:"每个单个元素具有或者不具有的性质就定义一个集合;这个集合元素就是全体具有这个性质的对象".也即把集合理解为一些具有共同特征的事物的全体,组成集合的事物称为该集合的元素.例如"大于或等于零的整数"这个性质就定义了一个集合;每个大于或等于零的整数都是该集合的元素.

通常,用大写字母表示集合,小写字母表示集合的元素.当 a 是集合 A 的元素时,称 a 属于 A 或 A 包含 a,记为 $a\in A$.当 a 不是集合 A 的元素时,称 a 不属于 A 或 A 不包含 a,记为 $a\notin A$.不含元素的集合称为空集,记为 .

例如,所有实数,直线上所有点,平面上所有点等等,都分别构成集合.

由一个元素 a 构成的集合记作$\{a\}$.因而 a 和$\{a\}$是不同的概念.由有限个元素构成的集合称为有限集.例如由整数 1,2,3,4 构成的集合就是有限集.而由

无限个元素构成的集合称为无限集.全体自然数的集合 **N**、全体实数的集合 **R**、全体有理数的集合 **Q** 及全体复数的集合 **C** 都分别构成无限集.表示一个集合,通常有两种方式:一种是穷举法,即列举出集合中的全部元素;一种是描述法,即用集合中全部元素所具有的特性来表述集合.例如:$A=\{1,2,3,4\}$是穷举法.$\mathbf{Q}=\{q\mid q$ 为有理数$\}$和 $\mathbf{R}^{+}=\{r\mid r$ 为大于零的实数$\}$都为描述法.

设 S_1 和 S_2 是两个集合,如果 S_1 的元素都是 S_2 的元素,则称 S_1 是 S_2 的子集,且称 S_1 被包含在 S_2 中或称 S_2 包含 S_1,记为 $S_1\quad S_2$ 或 $S_2\quad S_1$.约定空集是任意一个集合 A 的子集,即 $\quad A$.如果组成 S_1 和 S_2 的元素完全相同,亦即 $S_1\quad S_2$ 和 $S_2\quad S_1$ 同时成立,则称 S_1 和 S_2 相等,记为 $S_1=S_2$.若 $S_1\quad S_2$,但是 $S_1\neq S_2$,则称 S_1 为 S_2 的真子集,记为 $S_1\quad S_2$.若存在 $x\in S_1$,但 $x|\quad S_2$,则称 S_2 不包含 S_1 或 S_1 不被 S_2 包含,记作 S_1ň S_2.

设 S_1 和 S_2 是两个集合,定义交集 $S_1\cap S_2$,并集 $S_1\cup S_2$,差集 $S_1\setminus S_2$,笛卡儿集(或称直积) $S_1\times S_2$ 如下:

$$S_1\cap S_2=\{a\mid a\in S_1 \text{ 且 } a\in S_2\},$$

$$S_1\cup S_2=\{a\mid a\in S_1 \text{ 或 } a\in S_2\},$$

$$S_1\setminus S_2=\{a\mid a\in S_1 \text{ 且 } a\,|\quad S_2\},$$

$$S_1\times S_2=\{(a,b)\mid a\in S_1 \text{ 且 } b\in S_2\}.$$

1.2 二元关系

有了集合的概念,就可以研究集合之间或同一集合中元素之间的各种关系及其分类等问题.例如研究平面内所有直线的集合中的直线之间的平行、垂直及相交关系.要用数学来研究这类问题,就要有相应的数学概念来描述这类关系.为此给出下列概念.

设 S_1 和 S_2 为两个集合,R 是涉及两个对象的一个规则.如果对于 S_1 的任意一个元素 a 与 S_2 的任意一个元素 b,都能确定它们适合 R(称 a 与 b 有关系 R,记为 aRb)或它们不适合 R(称 a 与 b 没有关系 R,记为 aRb),就称 R 为 S_1 到 S_2 的二元关系.

定义 1.1 设 S_1, S_2 为两个集合,如果 $R\quad S_1\times S_2$,则称 R 为 S_1 到 S_2 的一个二元关系,若$(a,b)\in R$,则称 a, b 具有关系 R,记作 aRb;否则称 a, b 无关系 R,记作 aRb.

设 S 为一集合,$R\quad S\times S$,即 R 为 S 到 S 的一个二元关系,也称 R 为 S 上的一个二元关系.如果 S 上的二元关系 R 满足下列条件,则称 R 为 S 上的一个等价关系:

(1) 反身性 对任意 $a\in S$, $(a,a)\in R$;

(2) 传递性　当 $(a,b)\in R$, $(b,c)\in R$ 时,有 $(a,c)\in R$;

(3) 对称性　当 $(a,b)\in R$ 时,有 $(b,a)\in R$.

设 R 为 S 上的一个等价关系, $a\in S$,称 $\bar{a}=\{b\mid b\in S,(a,b)\in R\}$ 为 a 关于等价关系 R 的一个等价类,即在等价关系 R 下所有与 a 具有等价关系 R 的元素构成的集合.

例 1　(1) 设集合 $A=\{$蛋,奶,谷$\}$,集合 $B=\{$奶牛,山羊,母鸡$\}$.定义一个由 A 到 B 的二元关系 R: $a\in A$, $b\in B$, $(a,b)\in R$ 当且仅当 b 能产生 a,即

$$R=\{(\text{蛋},\text{母鸡}),(\text{奶},\text{奶牛}),(\text{奶},\text{山羊})\}.$$

对于二元关系 R 来说,蛋 R 母鸡,奶 R 奶牛,奶 R 母鸡等等.

(2) 假定把有一段公共边界的两个国家称为邻国,那么"相邻于"便是地球上所有国家组成的集合上的一个二元关系 R.于是(中国,朝鲜),(朝鲜,中国) $\in R$,但(中国,英国) $\notin R$.

(3) 三角形相似是平面上所有三角形组成的集合上的一个二元关系 R.易验证 R 满足反身性、传递性和对称性,于是 R 是平面上所有三角形组成的集合上的一个等价关系,而所有正三角形就构成关于 R 的一个等价类.

命题 1.1　S 为一个集合, R 为 S 上的等价关系, $a,b\in S$.如果 $\bar{a}\cap\bar{b}\neq\varnothing$,则 $\bar{a}=\bar{b}$.

证　设 $c\in\bar{a}\cap\bar{b}$.有 $(a,c)\in R$, $(c,b)\in R$.由传递性知 $(a,b)\in R$.任意 $d\in\bar{b}$,即 $(b,d)\in R$.由传递性可知 $(a,d)\in R$,即 $d\in\bar{a}$.因此 $\bar{b}\subseteq\bar{a}$.同理可证 $\bar{a}\subseteq\bar{b}$.于是 $\bar{a}=\bar{b}$.□

命题 1.1 说明,集合 S 上的每一个等价关系都给出集合 S 中元素的一个分类,反之亦然.

设 S 为一个集合, R 为 S 上的等价关系.以 R 来划分 S,所有等价类作为元素(即每一个等价类作为一个元素)所构成的集合,称为 S 关于 R 的商集.记为 S/R.

例 2　设 $\mathbf{Z}$ 是所有整数构成的集合, k 是一个给定的正整数.设 $a,b\in\mathbf{Z}$,定义二元关系 R 如下: aRb 或 $(a,b)\in R$ 当且仅当 $a-b$ 可被 k 整除,即 a 和 b 除以 k 的余数相同,记作 $a\equiv b\pmod{k}$.称关系 R 为关于模 k 同余关系.易验证,二元关系 R 满足等价关系的定义条件(1)～(3).当 $k=2$ 时,把整数分为两类 $\bar{0}$, $\bar{1}$,即奇数类与偶数类,此时商集 $\mathbf{Z}/R=\{\bar{0},\bar{1}\}$.当 $k=n$ 时,商集 $\mathbf{Z}/R=\{\bar{0},\bar{1},\cdots,\overline{n-1}\}$,称为 $\mathbf{Z}$ 关于模 n 同余关系 R 的商集,通常记为 $\mathbf{Z}_n$.

1.3　映射

对集合 S_1 和 S_2,通过一个给定法则 φ,使得对 S_1 中任意一个元素 a,都有 S_2 中一个确定的元素 a' 与它对应,称 φ 为 S_1 到 S_2 的一个映射,记为 φ: $S_1\to$

S_2 或 $a'=\varphi(a)$，称 a'为 a 在映射 φ 下的象，称 a 为 a'在映射 φ 下的一个原象．如果映射 φ 使 S_1 中不同元素对应于 S_2 中的不同元素，则 φ 称为 S_1 到 S_2 的单射．如果任意 $a'\in S_2$，存在 $a\in S_1$ 使得 $a'=\varphi(a)$，则称 φ 为 S_1 到 S_2 上的满射，否则称为 S_1 到 S_2 内的映射．如果 φ 既是单射又是满射，则称 φ 为一一对应或双射．

例 3　设 $S_1=\{1,2,3\}$，$S_2=\{7,8,9,10\}$，定义 f_1：$f_1(1)=8$，$f_1(2)=9$，$f_1(3)=10$，则 f_1 是 S_1 到 S_2 的单射．

设 $S_1=\{1,2,3,4\}$，$S_2=\{7,8,9\}$，定义 f_2：$f_2(1)=f_2(2)=7$，$f_2(3)=8$，$f_2(4)=9$，则 f_2 是 S_1 到 S_2 的满射．

设 $S_1=\{1,2,3\}$，$S_2=\{7,8,9\}$，定义 f_3：$f_3(1)=7$，$f_3(2)=9$，$f_3(3)=8$，则 f_3 是 S_1 到 S_2 的一一对应．

对于每个映射 $f: S_1\to S_2$，存在 S_1 到 S_2 的一个二元关系 R，

$$R=\{(a, f(a)) \mid a\in S_1\}$$

与之对应，这表明映射 $f: S_1\to S_2$ 是一种特殊的 S_1 到 S_2 的二元关系．

第二节　向量代数

2.1　向量的概念

在工程技术和科学研究中，常会遇到像力、力矩、位移、速度、加速度等这样一类量，它们既有大小又有方向．这类量称为向量或矢量．

在数学上，用一条有向线段表示向量．其长度表示向量的大小，其方向表示向量的方向．以 A 为起点，B 为终点的有向线段所表示的向量，记为 $\overline{AB}$．也用粗体字母或用加箭头的字母来表示向量，例如可用 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ 或 $\bar{\alpha}$，$\bar{\beta}$，$\bar{\gamma}$ 等等表示向量(图 1)．

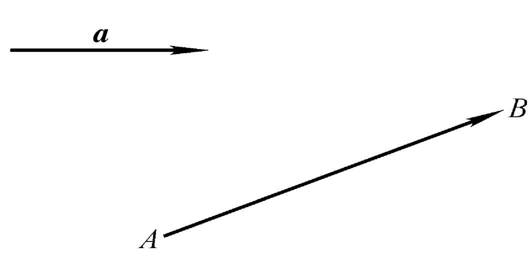


图 1

有些向量与其起点有关，有些向量与其起点无关．在数学上，将仅由长度与方向所定义的与起点无关的向量称为自由向量，以后本书所论及的向量总是指自由向量，且简称向量．由于只讨论自由向量，所以如果两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的大小相等，且方向相同，则可认为向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 相等，记作 $\boldsymbol{a}=\boldsymbol{b}$，也即向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 表示

同一个自由向量.

我们也可用等价关系来阐明自由向量.设 V 为一组给定向量组成的集合.$\boldsymbol{a},\boldsymbol{b}\in V$,在 V 中定义二元关系 R,$\boldsymbol{a}R\boldsymbol{b}$ 当且仅当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 大小相等且方向相同.容易验证,R 是一个等价关系.大小相等且方向相同的向量组成一个等价类,也即是一个自由向量,并且可用这个等价类中任意一个向量来表示这一类.

当一组向量的起点放在同一点时,如果它们的终点与公共起点在同一条直线上,则称这组向量共线.

向量的大小叫做向量的模.向量 $\boldsymbol{a}$ 的模记作 $|\boldsymbol{a}|$.模等于 1 的向量叫做单位向量.模等于零的向量叫做零向量.记作 $\boldsymbol{0}$.零向量没有确定的方向,其方向可以看作是任意的.由于零向量的方向可看作是任意的,因此可以认为零向量与任意一个向量共线.

当一组向量的起点放在同一点时,如果它们的终点与公共起点在同一平面上,就称这组向量共面.显然,任意两个向量必共面.

2.2　向量的加法、向量与数的乘积

1. 向量的加法

向量的运算来源于力学的研究.由力学知道,位移是向量.一个质点由 O 点出发,经过位移 $\boldsymbol{a}$ 到达 A 点,从 A 点经过位移 $\boldsymbol{b}$ 到达 B 点,其结果与从 O 点出发位移到 B 点是相同的,这个位移称为位移 $\boldsymbol{a}$ 与位移 $\boldsymbol{b}$ 的和.在数学上,用三角形法则定义两个向量的和.

设 $\boldsymbol{a},\boldsymbol{b}$ 为两个向量,因为它们是自由向量,可分别用以 O 为起点的向量 $\overline{OA}$ 和以 A 为起点的向量 $\overline{AB}$ 表示,即 $\boldsymbol{a}=\overline{OA}$,$\boldsymbol{b}=\overline{AB}$.则向量 $\overline{OB}=\boldsymbol{c}$ 称为向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的和,记为 $\boldsymbol{a}+\boldsymbol{b}$,即 $\boldsymbol{c}=\boldsymbol{a}+\boldsymbol{b}$.因为图 2 中 $\boldsymbol{a},\boldsymbol{b}$ 及 $\boldsymbol{a}+\boldsymbol{b}$ 构成三角形,所以上面求向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的和的方法称为向量加法的三角形法则.我们也可用平行四边形法则求两个向量的和:当向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 不平行时,作 $\overline{OA}=\boldsymbol{a}$,$\overline{OC}=\boldsymbol{b}$,以 OA 与 OC 为边作平行四边形 $OABC$,连对角线 OB,则向量 $\overline{OB}$ 即等于向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的和.显然,零向量 $\boldsymbol{0}$ 加任何向量 $\boldsymbol{a}$ 还等于向量 $\boldsymbol{a}$,即 $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$.由图 2 可知:

$$\boldsymbol{a}+\boldsymbol{b}=\overline{OA}+\overline{AB}=\overline{OB}=\boldsymbol{c},$$
$$\boldsymbol{b}+\boldsymbol{a}=\overline{OC}+\overline{CB}=\overline{OB}=\boldsymbol{c}.$$

即向量加法满足交换律.对于三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 相加(图 3),先作 $\boldsymbol{a}+\boldsymbol{b}$,再加 $\boldsymbol{c}$,即得和 $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$.先作 $\boldsymbol{b}+\boldsymbol{c}$,再作 $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$,即得和 $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$.图 3 表明 $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$ 与 $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$ 结果相同,可记为 $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$,因而加法满足结合律.

综上所述,对于任意向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$,其加法满足下列运算规律:

(1) 交换律:$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;

(2) 结合律:$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$.

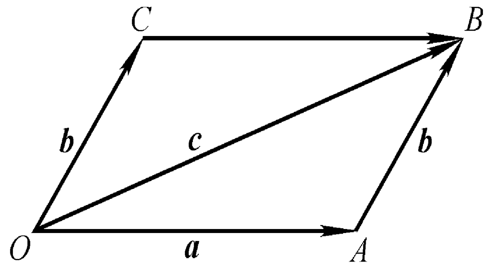


图 2

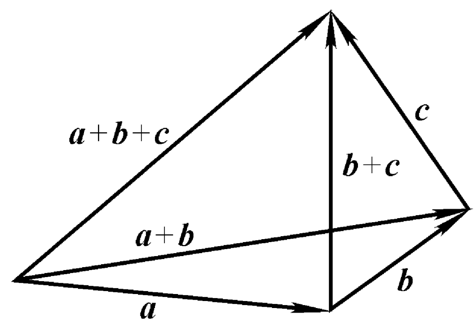


图 3

设 $\boldsymbol{a}$ 为一向量,与 $\boldsymbol{a}$ 的模相同而方向相反的向量叫做 $\boldsymbol{a}$ 的负向量,记作 $-\boldsymbol{a}$(图 4).定义两个向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的差 $\boldsymbol{a}-\boldsymbol{b}$ 等于 $\boldsymbol{a}+(-\boldsymbol{b})$.由三角形法则可知,将向量 $\boldsymbol{a}$ 与向量 $\boldsymbol{b}$ 的起点移到同一点,然后从向量 $\boldsymbol{b}$ 的终点向向量 $\boldsymbol{a}$ 的终点引一向量即为 $\boldsymbol{a}-\boldsymbol{b}$(图 5).

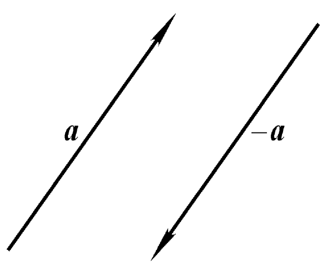


图 4

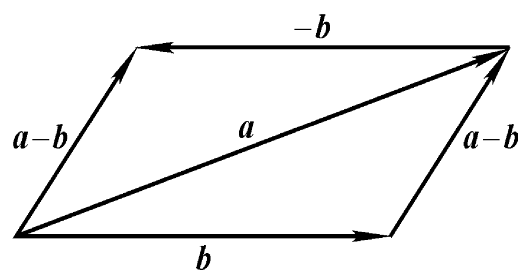


图 5

2. 向量与数的乘积(数乘)

设 λ 是一个实数,定义向量 $\boldsymbol{a}$ 与 λ 的乘积 $\lambda\boldsymbol{a}$ 为一个向量:当 $\lambda>0(\lambda<0)$ 时,$\lambda\boldsymbol{a}$ 表示一个方向与 $\boldsymbol{a}$ 的方向相同(相反)且模等于 $\lambda|\boldsymbol{a}|(|\lambda||\boldsymbol{a}|)$ 的向量;当 $\lambda=0$ 时,$\lambda\boldsymbol{a}$ 表示零向量,即 $\lambda\boldsymbol{a}=\boldsymbol{0}$.向量与数乘积,也简称数乘运算.

当向量 $\boldsymbol{a}$ 乘以 (-1) 时,得到 $(-1)\boldsymbol{a}$,它的模与 $\boldsymbol{a}$ 的模相等而方向与 $\boldsymbol{a}$ 的方向相反.所以有 $(-1)\boldsymbol{a}=-\boldsymbol{a}$.对于任意数 λ,μ 及任意向量 $\boldsymbol{a}$,向量 $\lambda(\mu\boldsymbol{a})$,$\mu(\lambda\boldsymbol{a})$,$(\lambda\mu)\boldsymbol{a}$ 方向都相同,而且

$$|\lambda(\mu\boldsymbol{a})|=|\mu(\lambda\boldsymbol{a})|=|(\lambda\mu)\boldsymbol{a}|=|(\lambda\mu)||\boldsymbol{a}|,$$

所以,

$$\lambda(\mu\boldsymbol{a})=\mu(\lambda\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}.$$

可直接验证向量与数的乘积有下列规律:

(1) $\lambda(\mu\boldsymbol{a})=\mu(\lambda\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$;

(2) $(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$;$\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$.

3. 向量共线、共面的充要条件

下面两个命题给出向量共线与共面的充要条件.

命题 1.2 两个向量 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 共线的充要条件是存在不全为零的实数 k_1 与 k_2,使得 $k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2=\boldsymbol{0}$.

证　必要性．设向量 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 共线，如果 $\boldsymbol{v}_1 \neq \mathbf{0}$，则 $|\boldsymbol{v}_1| \neq 0$，于是存在正实数 k，使 $|\boldsymbol{v}_2| = k|\boldsymbol{v}_1|$．当 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 方向相同时，$\boldsymbol{v}_2 = k\boldsymbol{v}_1$，于是，$\boldsymbol{v}_2 + (-k)\boldsymbol{v}_1 = \mathbf{0}$；当 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 方向相反时，$\boldsymbol{v}_2 = -k\boldsymbol{v}_1$，于是，$\boldsymbol{v}_2 + k\boldsymbol{v}_1 = \mathbf{0}$．当 $\boldsymbol{v}_1 = \mathbf{0}$，则有 $\boldsymbol{v}_1 + 0\boldsymbol{v}_2 = \mathbf{0}$．总之，适当选择 k_1 与 k_2，确实总存在不全为零的实数 k_1 与 k_2，使得 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 = \mathbf{0}$．

充分性．若存在不全为零的实数 k_1 与 k_2，使得 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 = \mathbf{0}$．不失一般性，可设 $k_1 \neq 0$，则 $\boldsymbol{v}_1 = -\dfrac{k_2}{k_1}\boldsymbol{v}_2$，即 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 共线．□

命题 1.3　三个向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 共面的充要条件是存在不全为零的实数 k_1, k_2, k_3，使得

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + k_3\boldsymbol{v}_3 = \mathbf{0}.$$

证　必要性．设向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 共面．如果 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 中有两个向量共线，不失一般性，可设 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 共线．由命题 1.2 知，存在不全为零的实数 k_1 与 k_2，使得 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 = \mathbf{0}$．因而存在不全为零的实数 $k_1, k_2, 0$，使得 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + 0\boldsymbol{v}_3 = \mathbf{0}$．

如果 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 互不共线，如图 6 所示，取从一点 O 出发的三个向量 $\overline{OA_1}$，$\overline{OA_2}$，$\overline{OA_3}$ 分别等于向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$，过 A_3 分别作平行于 OA_2 与 OA_1 的直线，交 OA_1 与 OA_2 于 B_1 与 B_2（图 6），则向量 $\overline{OB_1}$ 与 $\overline{OB_2}$ 分别与向量 $\overline{OA_1}$ 与 $\overline{OA_2}$ 共线，于是

$$\overline{OB_1} = k_1\overline{OA_1} = k_1\boldsymbol{v}_1,\ \overline{OB_2} = k_2\overline{OA_2} = k_2\boldsymbol{v}_2.$$

由平行四边形法则知

$$\boldsymbol{v}_3 = \overline{OA_3} = \overline{OB_1} + \overline{OB_2} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2.$$

于是有不全为零的实数 $k_1, k_2, -1$ 使 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + (-1)\boldsymbol{v}_3 = \mathbf{0}$．

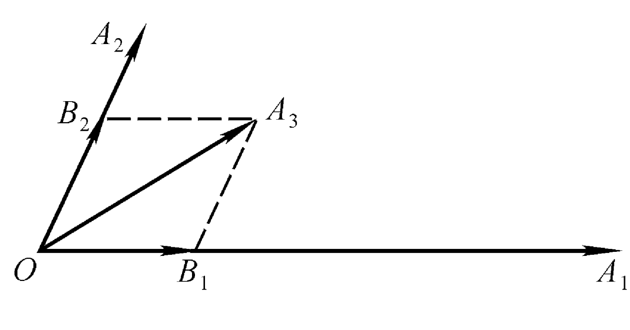


图 6

充分性．如果存在不全为零的实数 k_1, k_2, k_3，使 $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + k_3\boldsymbol{v}_3 = \mathbf{0}$．不失一般性，设 $k_1 \neq 0$，则 $\boldsymbol{v}_1 = -\dfrac{k_2}{k_1}\boldsymbol{v}_2 - \dfrac{k_3}{k_1}\boldsymbol{v}_3$，这说明 $\boldsymbol{v}_1$ 在由向量 $\boldsymbol{v}_2$、$\boldsymbol{v}_3$ 构成的平面内，即向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ 共面．□

命题 1.4　在一条直线上取定一个非零向量 $\boldsymbol{e}_1$，则该直线上的任一向量 $\boldsymbol{v}$

可唯一地表示为 $\boldsymbol{v}=x\boldsymbol{e}_1$,其中 x 为一个实数.

证 因为 $\boldsymbol{e}_1$ 与 $\boldsymbol{v}$ 是共线向量,由命题 1.2 可知,存在两个不全为零的实数 k_1 与 k_2,使 $k_1\boldsymbol{v}+k_2\boldsymbol{e}_1=\boldsymbol{0}$.若 $k_1=0$,由 $k_1\boldsymbol{v}+k_2\boldsymbol{e}_1=\boldsymbol{0}$,可推知 $k_2\boldsymbol{e}_1=\boldsymbol{0}$.因为 $\boldsymbol{e}_1$ 是非零向量,所以 $k_2=0$.这与 k_1 和 k_2 不全为零矛盾,因此 $k_1\neq 0$.由 $k_1\boldsymbol{v}+k_2\boldsymbol{e}_1=\boldsymbol{0}$ 知,$\boldsymbol{v}=-\dfrac{k_2}{k_1}\boldsymbol{e}_1$,令 $x=-\dfrac{k_2}{k_1}$,则 $\boldsymbol{v}=x\boldsymbol{e}_1$.设 $\boldsymbol{v}=x\boldsymbol{e}_1$,又 $\boldsymbol{v}=y\boldsymbol{e}_1$,则有

$$\boldsymbol{0}=\boldsymbol{v}-\boldsymbol{v}=x\boldsymbol{e}_1-y\boldsymbol{e}_1=(x-y)\boldsymbol{e}_1.$$

由于 $\boldsymbol{e}_1\neq\boldsymbol{0}$,所以 $x-y=0$, $x=y$,即 $\boldsymbol{v}$ 可唯一地表示为 $\boldsymbol{v}=x\boldsymbol{e}_1$.□

到现在为止,已定义向量的两种运算,即向量的加法,数与向量的乘积.向量的加法及向量与数的乘积统称为向量的线性运算.定义了向量的线性运算之后,就可研究向量间的代数关系.

例 1 设点 M 是 $\triangle ABC$ 的形心,AD 是 BC 边上的中线,用向量线性运算证明:$AM=\dfrac{2}{3}AD$(图 7).

证 因为向量 $\overline{AM}$, $\overline{AD}$ 在一条直线上,故可设 $\overline{AM}=x\,\overline{AD}$.又因为 D 是 BC 边的中点,由向量加法的平行四边形法则及图 7,有

$$\overline{AD}=\frac{1}{2}\overline{AF}=\frac{1}{2}(\overline{AB}+\overline{AC}),$$

因此,

$$\overline{AM}=x\,\overline{AD}=\frac{1}{2}x(\overline{AB}+\overline{AC}).$$

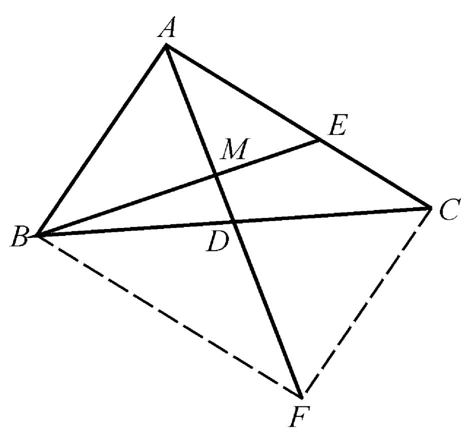


图 7

同理可设 $\overline{ME}=y\,\overline{BE}$.因为 BE 是 AC 边上中线,有

$$\overline{BE}=\overline{BA}+\overline{AE}=-\overline{AB}+\frac{1}{2}\overline{AC},$$

因此

$$\overline{ME}=y(-\overline{AB}+\frac{1}{2}\overline{AC}).$$

在△AME 中，$\overline{AM}+\overline{ME}+\overline{EA}=\mathbf{0}$，即

$$\frac{x}{2}\left(\overline{AB}+\overline{AC}\right)+y\left(-\overline{AB}+\frac{1}{2}\overline{AC}\right)-\frac{1}{2}\overline{AC}=\mathbf{0}.$$

由向量与数的乘积性质，加法的交换律和结合律有

$$\left(\frac{x}{2}-y\right)\overline{AB}+\frac{1}{2}(x+y-1)\overline{AC}=\mathbf{0}.$$

$$\left(\frac{x}{2}-y\right)\overline{AB}=-\frac{1}{2}(x+y-1)\overline{AC}=\boldsymbol{\gamma}.$$

若 $\boldsymbol{\gamma}\neq\mathbf{0}$，则$\overline{AB}$和$\overline{AC}$都平行于 $\boldsymbol{\gamma}$.这与点 A、B 和 C 构成三角形矛盾.因此 $\boldsymbol{\gamma}=\mathbf{0}$.又因为$\overline{AB}$和$\overline{AC}$都是非零向量，所以

$$\frac{x}{2}-y=0,$$

$$\frac{1}{2}(x+y-1)=0.$$

解此方程，得 $x=\frac{2}{3}$，即 $AM=\frac{2}{3}AD$.

在例 1 中，由向量$\overline{AB}$，$\overline{AC}$不共线这个几何关系，推导出 x 与 y 之间的代数关系，即二元一次方程组.解这个二元一次方程组，很容易地得到了 $AM=\frac{2}{3}AD$这个几何关系.借助代数与几何之间的内在联系，可以解决许多数学问题.

2.3　仿射坐标系　向量及其运算的坐标表示法

用向量方法处理问题的优点是比较直观，但向量运算不如数的运算方便.引入坐标系，在向量与有序数组之间建立一一对应关系，把几何问题转化成代数问题，亦可借助几何直观研究代数问题，这是 17 世纪数学的重大创新之一.坐标系有直角坐标系和仿射坐标系等等，首先介绍仿射坐标系.

1.仿射坐标系

根据命题 1.4，可建立任一直线 L 上的仿射坐标系.

在一直线 L 上取一个固定点 O 及非零向量 $\boldsymbol{e}_1$，构成一个仿射坐标系，记为$\{O;\boldsymbol{e}_1\}$，其中点 O 称为坐标原点，$\boldsymbol{e}_1$ 称为坐标向量或基，所在直线称为坐标轴.对于直线 L 上的任一向量 $\boldsymbol{v}$，有唯一的实数 x，使得 $\boldsymbol{v}=x\boldsymbol{e}_1$.$x$ 称为向量 $\boldsymbol{v}$ 在坐标系$\{O;\boldsymbol{e}_1\}$下的坐标.对于直线 L 上任一点 M，取向量$\overline{OM}$，称$\overline{OM}$为点 M 的向径，向量$\overline{OM}$的坐标 x 称为点 M 关于仿射坐标系$\{O;\boldsymbol{e}_1\}$的仿射坐标，点 M 可写为 $M(x)$.

命题 1.5　设 $\boldsymbol{e}_1$ 与 $\boldsymbol{e}_2$ 为一个平面上两个不共线的向量，则该平面上任一向量 $\boldsymbol{v}$ 可唯一地表示为 $\boldsymbol{v}=x\boldsymbol{e}_1+y\boldsymbol{e}_2$，其中 x，y 为两个实数.

类似命题 1.4,可根据命题 1.2 证明命题 1.5 .类似于直线上的仿射坐标系,可以建立平面仿射坐标系,留作练习 .

命题 1.6　设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 为空间中三个不共面的向量,$\boldsymbol{v}$ 为空间任一自由向量,则 $\boldsymbol{v}$ 可唯一地表示为 $\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$,其中 x, y, z 为三个实数 .

证　如图 8 所示,在空间任取一点 O,作向量 $\overline{OA_1} = \boldsymbol{e}_1$, $\overline{OA_2} = \boldsymbol{e}_2$, $\overline{OA_3} = \boldsymbol{e}_3$, $\overline{OA} = \boldsymbol{v}$.若 $\overline{OA}$ 与 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 中某两个共面,例如与 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 共面,根据命题 1.5, $\boldsymbol{v}$ 可唯一表示为 $\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + 0\boldsymbol{e}_3$,则命题成立 .

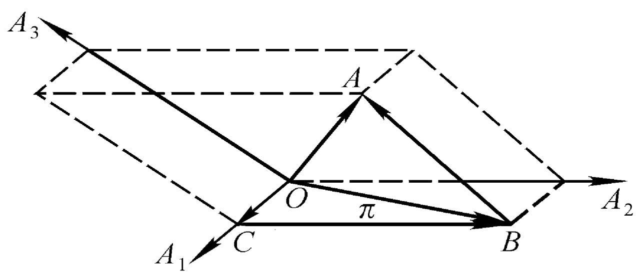


图 8

下面设 $\overline{OA}$ 与 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 中任意两个向量都不共面 .过 A 作平行于 $\overline{OA_3}$ 的直线交由 OA_1, OA_2 构成的平面 π 于 B .在平面 π 上过点 B 作平行于 $\overline{OA_2}$ 的直线交 OA_1 于 C .根据向量加法的三角形法则,有如下向量等式:

$$\overline{OA} = \overline{OB} + \overline{BA} = \overline{OC} + \overline{CB} + \overline{BA} .$$

向量 $\overline{OC}$, $\overline{CB}$, $\overline{BA}$ 分别与向量 $\overline{OA_1}$, $\overline{OA_2}$, $\overline{OA_3}$ 共线 .由命题 1.4,存在实数 x, y, z 使 .

$$\overline{OC} = x\boldsymbol{e}_1, \overline{CB} = y\boldsymbol{e}_2, \overline{BA} = z\boldsymbol{e}_3,$$

于是

$$\boldsymbol{v} = \overline{OA} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3 .$$

若 $\boldsymbol{v}$ 有两种不同的表示法,即

$$\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3, \boldsymbol{v} = x'\boldsymbol{e}_1 + y'\boldsymbol{e}_2 + z'\boldsymbol{e}_3,$$

则

$$\boldsymbol{0} = \boldsymbol{v} - \boldsymbol{v} = (x - x')\boldsymbol{e}_1 + (y - y')\boldsymbol{e}_2 + (z - z')\boldsymbol{e}_3 .$$

因为 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 不共面,由命题 1.3 知 $x - x' = 0$, $y - y' = 0$, $z - z' = 0$,这表明 $\boldsymbol{v}$ 的表示式是唯一的 .□

等式 $\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$ 称为向量 $\boldsymbol{v}$ 对 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的分解式 .在这个分解式中,仅使用了向量的线性运算,因而也称向量 $\boldsymbol{v}$ 是向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的线性组合 .

在空间中取一个定点及三个有顺序的不共面的向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 构成空间的一个仿射坐标系,记为 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$,其中点 O 称为坐标原点, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 称为坐标向量,它们所在直线分别称为 $\boldsymbol{x}$ 轴, $\boldsymbol{y}$ 轴, $\boldsymbol{z}$ 轴 .由两个坐标轴确定的平面都

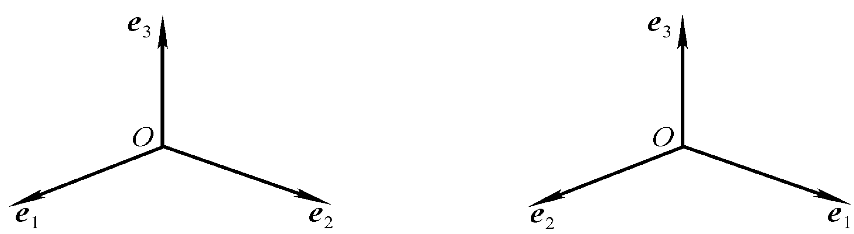
称为坐标面．由 x 轴与 y 轴构成的平面称为 xOy 面；由 y 轴与 z 轴构成的平面称为 yOz 面；由 z 轴与 x 轴构成的平面称为 zOx 面．由命题 1.6 知，在该仿射坐标系下，空间任意一个向量 $\boldsymbol{v}$ 可唯一地表示为

$$\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3,$$

数组 (x, y, z) 称为向量 $\boldsymbol{v}$ 在坐标系 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 下的坐标，并称 $\boldsymbol{v} = (x, y, z)$ 为向量 $\boldsymbol{v}$ 的坐标表示式．

在取定仿射坐标系 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 之后，对于空间中一点 M，向量 $\overline{OM}$ 称为点 M 的向径，向径 $\overline{OM}$ 在该坐标系下的坐标 (x, y, z) 称为点 M 在该坐标系下的坐标，记为 $M(x, y, z)$．

由于坐标向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的相对位置不同，可构成两类不同的坐标系：右手系（图 9(a)）与左手系（图 9(b)）．



(a)　　(b)

图 9

三个坐标平面把空间分为八个部分，每一部分称为一个卦限，卦限的名称由其间的点的坐标的正负决定，见下表：

卦限名称 / 坐标符号 / 点的坐标	Ⅰ	Ⅱ	Ⅲ	Ⅳ	Ⅴ	Ⅵ	Ⅶ	Ⅷ
x	+	-	-	+	+	-	-	+
y	+	+	-	-	+	+	-	-
z	+	+	+	+	-	-	-	-

有了仿射坐标系后，向量的线性运算便可以转化为坐标的运算．设向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 在仿射坐标系 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 中的坐标分解式分别为

$$\boldsymbol{a} = x_1\boldsymbol{e}_1 + y_1\boldsymbol{e}_2 + z_1\boldsymbol{e}_3,\quad \boldsymbol{b} = x_2\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + z_2\boldsymbol{e}_3,$$

λ 为任意实数．根据向量线性运算的规则，有

$$\begin{aligned}\boldsymbol{a} + \boldsymbol{b} &= (x_1\boldsymbol{e}_1 + y_1\boldsymbol{e}_2 + z_1\boldsymbol{e}_3) + (x_2\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + z_2\boldsymbol{e}_3)\\ &= (x_1 + x_2)\boldsymbol{e}_1 + (y_1 + y_2)\boldsymbol{e}_2 + (z_1 + z_2)\boldsymbol{e}_3.\end{aligned}$$

$$\lambda\boldsymbol{a} = \lambda(x_1\boldsymbol{e}_1 + y_1\boldsymbol{e}_2 + z_1\boldsymbol{e}_3) = (\lambda x_1)\boldsymbol{e}_1 + (\lambda y_1)\boldsymbol{e}_2 + (\lambda z_1)\boldsymbol{e}_3.$$

向量 $\boldsymbol{a}$、$\boldsymbol{b}$ 相应的坐标运算为

$$(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2),$$

$$\lambda(x_1, y_1, z_1) = (\lambda x_1, \lambda y_1, \lambda z_1).$$

上面两式称为向量线性运算的坐标表示式.这就把向量的几何运算转化为坐标的代数运算.

2．直角坐标系

在引入直角坐标系之前,首先引入两个向量夹角的概念.设有两个非零向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$,任取空间一点 O,作 $\overline{OA} = \boldsymbol{a}$, $\overline{OB} = \boldsymbol{b}$.规定不超过 π 的 $\angle AOB$ 为向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角,记为 $\langle \boldsymbol{a}, \boldsymbol{b}\rangle$.如果向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 中有一个为零向量,规定向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角可为 0 到 π 间的任意一个值.两个向量夹角等于 $\frac{\pi}{2}$ 时,称两个向量垂直且记为 $\boldsymbol{a} \perp \boldsymbol{b}$,并约定零向量与任何向量都垂直.

在以 O 为原点的仿射坐标系中,若三个坐标向量是互相垂直的单位向量,常将这三个坐标向量记为 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$,称仿射坐标系 $\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ 为直角坐标系.这样,空间中任一向量 $\boldsymbol{v}$ 均可唯一地表示为 $\boldsymbol{v} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$,有序数组 x, y, z 记为 (x, y, z),称为向量 $\boldsymbol{v}$ 在直角坐标系 $\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ 中的坐标.对空间中一点 M,依然将向径 $\overline{OM}$ 的坐标 (x, y, z) 定义为点 M 的坐标,也记为 (x, y, z).一点 M 的坐标 (x, y, z) 中的 x, y, z 分别称为点 M 的横坐标,纵坐标及竖坐标.

设向量 $\boldsymbol{v} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$,由图 10 有 $\overline{OP} = x\boldsymbol{i}$, $\overline{PN} = y\boldsymbol{j}$, $\overline{NM} = z\boldsymbol{k}$.按勾股定理,有

$$|OM|^2 = |ON|^2 + |NM|^2 = |OP|^2 + |PN|^2 + |NM|^2 = x^2 + y^2 + z^2,$$

即得

$$|\boldsymbol{v}| = \sqrt{x^2 + y^2 + z^2}.$$

这便是向量模的坐标表示式.利用此表示式,容易得出空间两点间的距离公式.

设 $M_1(x_1, y_1, z_1)$ 和 $M_2(x_2, y_2, z_2)$ 为空间中两点,则它们的向径

$$\overline{OM_1} = x_1\boldsymbol{i} + y_1\boldsymbol{j} + z_1\boldsymbol{k}, \quad \overline{OM_2} = x_2\boldsymbol{i} + y_2\boldsymbol{j} + z_2\boldsymbol{k},$$

从而

$$\overline{M_1M_2} = \overline{OM_2} - \overline{OM_1} = (x_2 - x_1)\boldsymbol{i} + (y_2 - y_1)\boldsymbol{j} + (z_2 - z_1)\boldsymbol{k},$$

于是 M_1 与 M_2 两点间的距离

$$|M_1M_2| = |\overline{M_1M_2}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

在直角坐标系 $\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ 中,一个向量 $\boldsymbol{v} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ 与向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 的夹角 α, β, γ 称为向量 $\boldsymbol{v}$ 的方向角,$\cos\alpha, \cos\beta, \cos\gamma$ 称为向量 $\boldsymbol{v}$ 的方向余弦.显然

$$\cos\alpha = \frac{x}{d}, \quad \cos\beta = \frac{y}{d}, \quad \cos\gamma = \frac{z}{d},$$

其中 $d = \sqrt{x^2 + y^2 + z^2}$ 为向量 $\boldsymbol{v}$ 的模,且 $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

在建立空间仿射坐标系 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ 时,既不要求坐标向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 相互

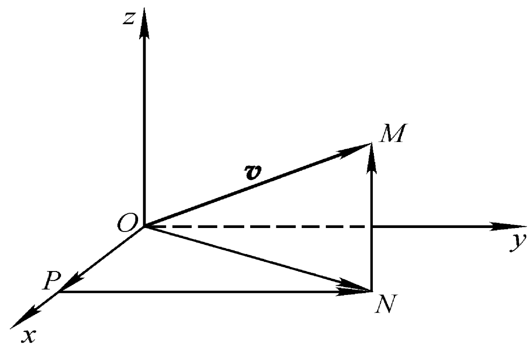


图 10

垂直,也不要求它们是单位向量.在建立直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$时,则要求坐标向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 是相互垂直的单位向量.正因为这样,用空间直角坐标系的坐标表示的空间中两点距离及方向余弦的公式才非常简单.

在建立直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$之后,空间一点 $M(x, y, z)$唯一地对应一个向量$\overline{OM} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$,当向量$\overline{OM}$的坐标为$(x, y, z)$时,则点 M 的坐标就是(x, y, z).反之,已知两点 $M_0(x_0, y_0, z_0)$, $M(x, y, z)$,那么向量

$$\overline{M_0M} = \overline{OM} - \overline{OM_0} = (x - x_0)\boldsymbol{i} + (y - y_0)\boldsymbol{j} + (z - z_0)\boldsymbol{k},$$

于是向量$\overline{M_0M}$的坐标为$(x - x_0, y - y_0, z - z_0)$.正因为在一个坐标系中,两个对象——(自由)向量和点都能用坐标表示,而且两者坐标又有上述的关联,才使向量代数成为解决空间几何问题的重要工具.在第二章的学习中将会更清楚地看到这一点.

2.4　数量积　向量积　混合积

1. 两向量的数量积

在研究物理问题时,除了需要向量加法及数乘构成的线性运算外,还需要向量的一些新的运算.首先考察一个例子.设一物体在常力 $\boldsymbol{F}$(向量)作用下沿直线从 O 点移动到 M 点.以向量 $\boldsymbol{s}$ 表示位移向量$\overline{OM}$.由物理学知道,力 $\boldsymbol{F}$ 所作的功为

$$W = |\boldsymbol{F}||\boldsymbol{s}|\cos\theta,$$

其中 θ 为 $\boldsymbol{F}$ 与 $\boldsymbol{s}$ 的夹角(图 11(a)).

于是,有如下定义:两个向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的数量积为一个实数,它等于 $\boldsymbol{a}, \boldsymbol{b}$ 的模$|\boldsymbol{a}|, |\boldsymbol{b}|$及它们的夹角 $\theta = \langle \boldsymbol{a}, \boldsymbol{b} \rangle$ 余弦的乘积,记作 $\boldsymbol{a} \cdot \boldsymbol{b}$(图 11(b)),即

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta.$$

当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 中有一个为零向量,定义 $\boldsymbol{a} \cdot \boldsymbol{b} = 0$.根据这个定义,力 $\boldsymbol{F}$ 所作的功 W 是力 $\boldsymbol{F}$ 与位移 $\boldsymbol{s}$ 的数量积,即 $W = \boldsymbol{F} \cdot \boldsymbol{s}$.

为了研究向量之间的关系,需要给出一个向量在另一个非零向量上的投影的概念.向量 $\boldsymbol{b}$ 在非零向量 $\boldsymbol{a}$ 上的投影,用 $\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{b}$ 来表示,定义

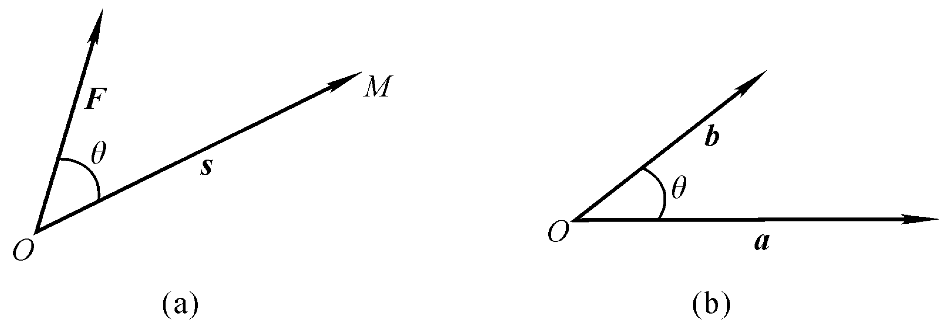


图 11

$$\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{b}=|\boldsymbol{b}|\cos\theta,$$

其中 $\boldsymbol{a}\neq\boldsymbol{0}$, $\theta=\langle\boldsymbol{a},\boldsymbol{b}\rangle$. 对于两个向量 $\boldsymbol{b},\boldsymbol{c}$ 的和向量在向量 $\boldsymbol{a}$ 上的投影(图 12). 由图 12 可知

$$\begin{aligned}\mathrm{Prj}_{\boldsymbol{a}}(\boldsymbol{b}+\boldsymbol{c})&=|\boldsymbol{b}+\boldsymbol{c}|\cos\alpha\\&=|\boldsymbol{b}|\cos\beta+|\boldsymbol{c}|\cos\gamma\\&=|\boldsymbol{b}|\cos\langle\boldsymbol{b},\boldsymbol{a}\rangle+|\boldsymbol{c}|\cos\langle\boldsymbol{c},\boldsymbol{a}\rangle\\&=\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{b}+\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{c}.\end{aligned}$$

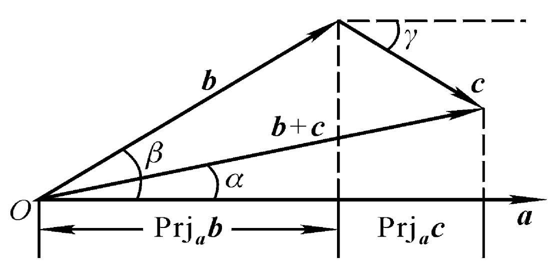


图 12

由于 $|\boldsymbol{b}|\cos\theta=\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{b}$, 便有 $\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}|\mathrm{Prj}_{\boldsymbol{a}}\boldsymbol{b}$. 同样也有 $\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{b}|\mathrm{Prj}_{\boldsymbol{b}}\boldsymbol{a}$.

命题 1.7 对于任意向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$, 下列等式成立:

(1) $\boldsymbol{a}\cdot\boldsymbol{a}=|\boldsymbol{a}|^2$;

(2) $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$(交换律);

(3) $\boldsymbol{a}\cdot\boldsymbol{b}=0$ 当且仅当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 互相垂直;

(4) $(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c}$(分配律);

(5) $(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=\lambda(\boldsymbol{a}\cdot\boldsymbol{b})$, λ 为任意数;

证 (1)—(2) 根据数量积定义可直接验证.

(3) 如果 $\boldsymbol{a}\cdot\boldsymbol{b}=0$, 当 $|\boldsymbol{a}|\neq0$, $|\boldsymbol{b}|\neq0$ 时, 有 $\cos\theta=0$, 从而 $\theta=\dfrac{\pi}{2}$, 即 $\boldsymbol{a}\perp\boldsymbol{b}$. 当 $\boldsymbol{a}$ 或 $\boldsymbol{b}$ 中有一个为零向量时, 因为约定零向量与任何向量垂直, 所以仍有 $\boldsymbol{a}\perp\boldsymbol{b}$. 反之, 如果 $\boldsymbol{a}\perp\boldsymbol{b}$, 则 $\theta=\dfrac{\pi}{2}$, 或 $|\boldsymbol{a}|=0$, 或 $|\boldsymbol{b}|=0$, 于是

$$\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta=0.$$

(4) 当 $\boldsymbol{c}=\boldsymbol{0}$时,

$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}=0=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c},$$

所以有

$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c}.$$

当 $\boldsymbol{c}\neq\boldsymbol{0}$ 时,由数量积的定义可知

$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}=|\boldsymbol{c}|\operatorname{Prj}_{\boldsymbol{c}}(\boldsymbol{a}+\boldsymbol{b}).$$

由投影性质,可知

$$\operatorname{Prj}_{\boldsymbol{c}}(\boldsymbol{a}+\boldsymbol{b})=\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{a}+\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{b},$$

所以

$$\begin{aligned}(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c}&=|\boldsymbol{c}|(\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{a}+\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{b})\\&=|\boldsymbol{c}|\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{a}+|\boldsymbol{c}|\operatorname{Prj}_{\boldsymbol{c}}\boldsymbol{b}\\&=\boldsymbol{a}\cdot\boldsymbol{c}+\boldsymbol{b}\cdot\boldsymbol{c}.\end{aligned}$$

(5) 当 $\lambda>0$ 时,$\lambda\boldsymbol{a}$ 与 $\boldsymbol{a}$ 的方向相同,$\lambda\boldsymbol{b}$ 与 $\boldsymbol{b}$ 的方向相同,所以

$$\langle\lambda\boldsymbol{a},\boldsymbol{b}\rangle=\langle\boldsymbol{a},\lambda\boldsymbol{b}\rangle=\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

按定义有

$$(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=|\lambda\boldsymbol{a}||\boldsymbol{b}|\cos\langle\lambda\boldsymbol{a},\boldsymbol{b}\rangle=\lambda|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

$$\lambda(\boldsymbol{a}\cdot\boldsymbol{b})=\lambda|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle.$$

从而

$$(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=\lambda(\boldsymbol{a}\cdot\boldsymbol{b}).$$

当 $\lambda=0$ 时,显然有

$$(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=\lambda(\boldsymbol{a}\cdot\boldsymbol{b})=0.$$

当 $\lambda<0$ 时,

$$\langle\lambda\boldsymbol{a},\boldsymbol{b}\rangle=\pi-\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

按定义有

$$\begin{aligned}(\lambda\boldsymbol{a})\cdot\boldsymbol{b}&=|\lambda\boldsymbol{a}||\boldsymbol{b}|\cos\langle\lambda\boldsymbol{a},\boldsymbol{b}\rangle=-\lambda|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\lambda\boldsymbol{a},\boldsymbol{b}\rangle\\&=\lambda|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle=\lambda(\boldsymbol{a}\cdot\boldsymbol{b}),\end{aligned}$$

总之,

$$(\lambda\boldsymbol{a})\cdot\boldsymbol{b}=\lambda(\boldsymbol{a}\cdot\boldsymbol{b}).\square$$

在一个直角坐标系$\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$中,设

$$\boldsymbol{a}=x_1\boldsymbol{i}+y_1\boldsymbol{j}+z_1\boldsymbol{k},\quad\boldsymbol{b}=x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k}$$

分别为向量 $\boldsymbol{a}$ 与向量 $\boldsymbol{b}$ 的坐标分解式.由命题 1.7,特别是分配律得到:

$$
\begin{aligned}
\boldsymbol{a}\cdot\boldsymbol{b} &= (x_1\boldsymbol{i}+y_1\boldsymbol{j}+z_1\boldsymbol{k})\cdot(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})\\
&= x_1\boldsymbol{i}\cdot(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})+y_1\boldsymbol{j}\cdot(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})+z_1\boldsymbol{k}\cdot(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})\\
&= x_1x_2\boldsymbol{i}\cdot\boldsymbol{i}+x_1y_2\boldsymbol{i}\cdot\boldsymbol{j}+x_1z_2\boldsymbol{i}\cdot\boldsymbol{k}+y_1x_2\boldsymbol{j}\cdot\boldsymbol{i}+y_1y_2\boldsymbol{j}\cdot\boldsymbol{j}+y_1z_2\boldsymbol{j}\cdot\boldsymbol{k}\\
&\quad +z_1x_2\boldsymbol{k}\cdot\boldsymbol{i}+z_1y_2\boldsymbol{k}\cdot\boldsymbol{j}+z_1z_2\boldsymbol{k}\cdot\boldsymbol{k}.
\end{aligned}
$$

由于 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 互相垂直,所以 $\boldsymbol{i}\cdot\boldsymbol{j}=\boldsymbol{j}\cdot\boldsymbol{i}=\boldsymbol{j}\cdot\boldsymbol{k}=\boldsymbol{k}\cdot\boldsymbol{j}=\boldsymbol{i}\cdot\boldsymbol{k}=\boldsymbol{k}\cdot\boldsymbol{i}=0$.又由于 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 的模均为 1,所以 $\boldsymbol{i}\cdot\boldsymbol{i}=\boldsymbol{j}\cdot\boldsymbol{j}=\boldsymbol{k}\cdot\boldsymbol{k}=1$,因而数量积的坐标表示式为

$$\boldsymbol{a}\cdot\boldsymbol{b}=x_1x_2+y_1y_2+z_1z_2. \tag{1.1}$$

由于 $\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle$,当 $\boldsymbol{a},\boldsymbol{b}$ 都不是零向量时,由(1.1)有

$$\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle=\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}.$$

将数量积及向量的坐标表示式代入上式,可得两向量夹角余弦的坐标表示式

$$\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle=\frac{x_1x_2+y_1y_2+z_1z_2}{\sqrt{x_1^2+y_1^2+z_1^2}\sqrt{x_2^2+y_2^2+z_2^2}}. \tag{1.2}$$

从该公式可知,两向量互相垂直当且仅当

$$x_1x_2+y_1y_2+z_1z_2=0.$$

例 2　四面体 $OABC$ 中,$OA\perp BC$,$OB\perp AC$(图 13),求证:$AB\perp OC$.

证　用向量代数的方法,取向量 $\overline{OA}=\alpha$,$\overline{OB}=\beta$,$\overline{OC}=\gamma$,则

$$\overline{BC}=\gamma-\beta,\ \overline{AC}=\gamma-\alpha,\ \overline{AB}=\beta-\alpha.$$

由 $OA\perp BC$,有 $\alpha\cdot(\gamma-\beta)=0$,即 $\alpha\cdot\gamma=\alpha\cdot\beta$.由 $OB\perp AC$,有 $\beta\cdot(\gamma-\alpha)=0$,即 $\beta\cdot\gamma=\alpha\cdot\beta$.因此

$$(\beta-\alpha)\cdot\gamma=\beta\cdot\gamma-\alpha\cdot\gamma=\alpha\cdot\beta-\alpha\cdot\beta=0,$$

也就是 $AB\perp OC$.

2. 两向量的向量积

在研究物体转动问题,例如开启一扇门的时候,就会发现,不但要考虑门所受的力,同时也要考虑力到门轴(转动轴)的距离,即要考虑力矩.下面考虑一个简化问题.

设 O 为一根杠杆 L 的支点,有一个力 $\boldsymbol{F}$ 作用于这杠杆上点 P 处.$\boldsymbol{F}$ 与向量 $\overline{OP}$ 的夹角为 θ(图 14).由力学知道,力 $\boldsymbol{F}$ 对支点 O 的力矩是一个向量 $\boldsymbol{M}$,记为 $\overline{OP}\times\boldsymbol{F}$,即 $\boldsymbol{M}=\overline{OP}\times\boldsymbol{F}$,其模 $|\boldsymbol{M}|=|OQ||\boldsymbol{F}|=|\overline{OP}||\boldsymbol{F}|\sin\theta$,其方向垂直于 $\overline{OP}$ 与 $\boldsymbol{F}$ 决定的平面,且使 $\overline{OP}$,$\boldsymbol{F}$ 及 $\boldsymbol{M}$ 组成右手系(图 15).由此,可以抽象出两个向量的向量积的概念.

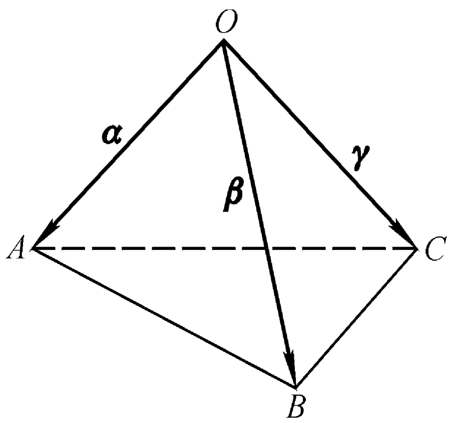


图 13

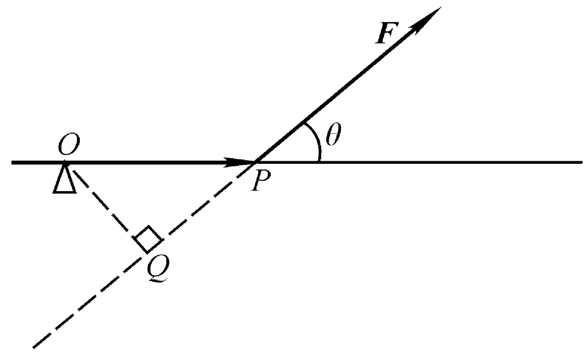


图 14

两个向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的向量积是一个向量 $\boldsymbol{c}$，记为 $\boldsymbol{a}\times\boldsymbol{b}$：$\boldsymbol{c}$ 的模 $|\boldsymbol{c}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\theta$，其中 θ 为 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 间的夹角；$\boldsymbol{c}$ 的方向垂直于 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 所决定的平面（$\boldsymbol{c}$ 既垂直于 $\boldsymbol{a}$，又垂直于 $\boldsymbol{b}$），且使 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}\times\boldsymbol{b}$ 组成右手系，即 $\boldsymbol{a}\times\boldsymbol{b}$ 的指向按右手规则从 $\boldsymbol{a}$ 转向 $\boldsymbol{b}$ 来确定（图 16）.

由向量积的定义可以直接推得：$\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$ 的充要条件是 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 共线，即 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 平行，应注意，由 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$ 不能推导出 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 必有一个为零向量，这与数的乘法不同.

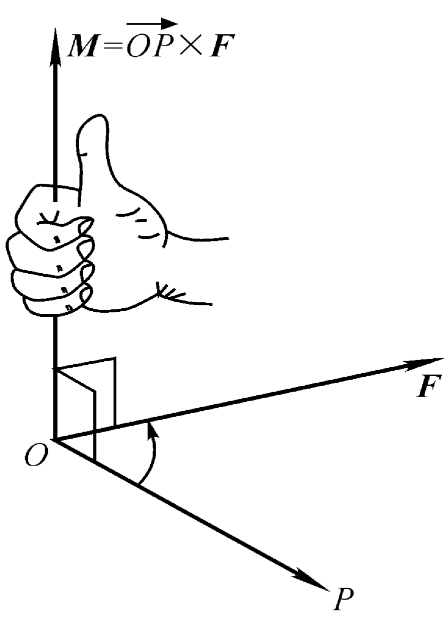


图 15

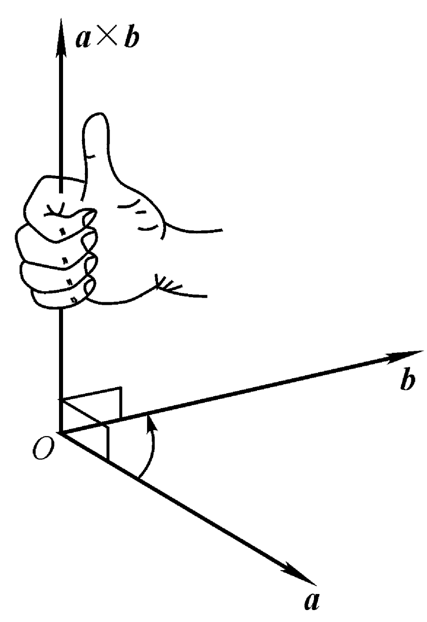


图 16

命题1.8　$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 为三个向量，则下列等式成立：

(1) $\boldsymbol{b}\times\boldsymbol{a}=-\boldsymbol{a}\times\boldsymbol{b}$；

(2) $(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}$（分配律）；

(3) $(\lambda\boldsymbol{a})\times\boldsymbol{b}=\lambda(\boldsymbol{a}\times\boldsymbol{b})$（$\lambda$ 为任意数）.

证　(1) 按右手规则从 $\boldsymbol{b}$ 转向 $\boldsymbol{a}$ 定出的方向与从 $\boldsymbol{a}$ 转向 $\boldsymbol{b}$ 定出的方向恰好相反.

(2) 与(3)的证明略.□

下面推导向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的向量积 $\boldsymbol{a}\times\boldsymbol{b}$ 在右手直角坐标系$\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$中的坐标表达式.设

$$\boldsymbol{a}=x_1\boldsymbol{i}+y_1\boldsymbol{j}+z_1\boldsymbol{k},$$

$$\boldsymbol{b}=x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k}$$

分别为向量 $\boldsymbol{a}$ 与向量 $\boldsymbol{b}$ 的坐标分解式.由命题 1.8 得到

$$\begin{aligned}\boldsymbol{a}\times\boldsymbol{b}&=(x_1\boldsymbol{i}+y_1\boldsymbol{j}+z_1\boldsymbol{k})\times(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})\\&=x_1\boldsymbol{i}\times(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})+y_1\boldsymbol{j}\times(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})+z_1\boldsymbol{k}\times(x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k})\\&=x_1x_2\boldsymbol{i}\times\boldsymbol{i}+x_1y_2\boldsymbol{i}\times\boldsymbol{j}+x_1z_2\boldsymbol{i}\times\boldsymbol{k}+y_1x_2\boldsymbol{j}\times\boldsymbol{i}+y_1y_2\boldsymbol{j}\times\boldsymbol{j}+y_1z_2\boldsymbol{j}\times\boldsymbol{k}\\&\quad+z_1x_2\boldsymbol{k}\times\boldsymbol{i}+z_1y_2\boldsymbol{k}\times\boldsymbol{j}+z_1z_2\boldsymbol{k}\times\boldsymbol{k}.\end{aligned}$$

由向量积定义可知,

$$\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k},\ \boldsymbol{j}\times\boldsymbol{i}=-\boldsymbol{k},\ \boldsymbol{j}\times\boldsymbol{k}=\boldsymbol{i},\ \boldsymbol{k}\times\boldsymbol{j}=-\boldsymbol{i},\ \boldsymbol{i}\times\boldsymbol{k}=-\boldsymbol{j},\ \boldsymbol{k}\times\boldsymbol{i}=\boldsymbol{j},$$

$$\boldsymbol{i}\times\boldsymbol{i}=\boldsymbol{j}\times\boldsymbol{j}=\boldsymbol{k}\times\boldsymbol{k}=\boldsymbol{0},$$

于是

$$\boldsymbol{a}\times\boldsymbol{b}=(y_1z_2-z_1y_2)\boldsymbol{i}+(z_1x_2-x_1z_2)\boldsymbol{j}+(x_1y_2-y_1x_2)\boldsymbol{k}.\qquad(1.3)$$

从公式(1.3)可知,两向量共线当且仅当

$$y_1z_2-z_1y_2=0,\ z_1x_2-x_1z_2=0,\ x_1y_2-y_1x_2=0.$$

例 3　在右手直角坐标系$\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$中,$\overrightarrow{AB}=\boldsymbol{i}-2\boldsymbol{j}+2\boldsymbol{k}$,$\overrightarrow{AC}=3\boldsymbol{i}-4\boldsymbol{k}$.试求△$ABC$ 的面积.

解　所求面积等于$|\overrightarrow{AB}\times\overrightarrow{AC}|\quad 2$.由公式(1.3),得

$$\frac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AC}|=\frac{1}{2}|8\boldsymbol{i}+10\boldsymbol{j}+6\boldsymbol{k}|=5\quad 2.$$

例 4　有一刚体绕一固定轴 l 以等角速率 ω 旋转,求该刚体上任意一点 P 处的速度向量 $\boldsymbol{v}$.

解　由物理学知道,点 P 处的速度向量 $\boldsymbol{v}$ 的方向与以 l 轴及点 P 所构成的平面垂直,速度的大小等于$|P_0P|\omega$,过 P 作平面垂直于轴 l,该平面与 l 交于 P_0,即 P_0 是 P 在 l 轴上的投影点.在旋转轴 l 上取一个向量 $\boldsymbol{\gamma}$,使得$|\boldsymbol{\gamma}|=\omega$,并且根据刚体的转动,按右手法则规定 $\boldsymbol{\gamma}$ 的正向,即$\overrightarrow{P_0P}$,$\boldsymbol{v}$ 与 $\boldsymbol{\gamma}$ 构成右手系(图 17).在轴 l 上任意取定一点 O,那么

$$|\boldsymbol{v}|=|\boldsymbol{\gamma}||\overrightarrow{P_0P}|=|\boldsymbol{\gamma}||\overrightarrow{OP}|\sin\langle\boldsymbol{\gamma},\overrightarrow{OP}\rangle,$$

其中$\langle\boldsymbol{\gamma},\overrightarrow{OP}\rangle$表示 $\boldsymbol{\gamma}$ 与$\overrightarrow{OP}$的夹角.按向量积的定义知

$$\boldsymbol{v}=\boldsymbol{\gamma}\times\overrightarrow{OP}.$$

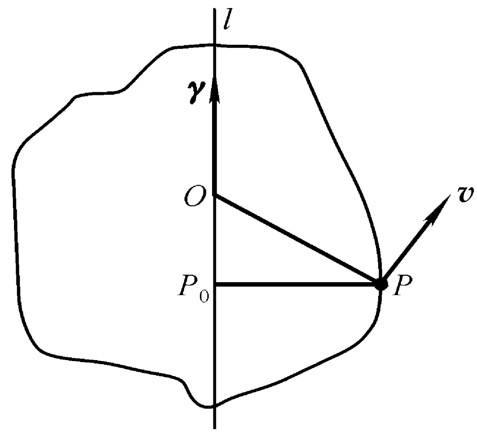


图 17

3．向量的混合积

已知三个向量 $\boldsymbol{a}, \boldsymbol{b}$ 和 $\boldsymbol{c}$, $(\boldsymbol{a}\times \boldsymbol{b})\cdot \boldsymbol{c}$ 叫做三向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 的混合积,记作 $[\boldsymbol{abc}]$.$(\boldsymbol{a}\times \boldsymbol{b})\cdot \boldsymbol{c}$ 表示先作向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的向量积 $\boldsymbol{a}\times \boldsymbol{b}$,再把向量 $\boldsymbol{a}\times \boldsymbol{b}$ 与第三个向量 $\boldsymbol{c}$ 作数量积,最终得到的结果是一个数量.

首先讨论混合积的符号.按向量积的定义,向量积 $\boldsymbol{a}\times \boldsymbol{b}=\boldsymbol{d}$ 是一个向量,其模等于以向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为边的平行四边形的面积,其方向垂直于该平行四边形所在的平面,且当向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 组成右(左)手系时,向量 $\boldsymbol{d}$ 与向量 $\boldsymbol{c}$ 朝向该平面的同(异)侧, $\boldsymbol{d}$ 与 $\boldsymbol{c}$ 的夹角 α 为锐(钝)角.因为

$$[\boldsymbol{abc}]=(\boldsymbol{a}\times \boldsymbol{b})\cdot \boldsymbol{c}=|\boldsymbol{a}\times \boldsymbol{b}||\boldsymbol{c}|\cos\alpha,$$

所以当 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 组成右(左)手系时,$[\boldsymbol{abc}]$为正(负).

其次讨论混合积的绝对值.以向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 为棱的平行六面体的底面上由向量 $\boldsymbol{a}, \boldsymbol{b}$ 构成的平行四边形的面积 A 在数值上等于 $|\boldsymbol{a}\times \boldsymbol{b}|$,平行六面体的高 h 等于向量 $\boldsymbol{c}$ 在向量 $\boldsymbol{d}$ 上的投影的绝对值,即 $h=\pm \mathrm{Prj}_{\boldsymbol{d}}\boldsymbol{c}=\pm|\boldsymbol{c}|\cos\alpha$,由于平行六面体的体积为正数,因此,下面公式中,当 α 为锐角时,取正号,当 α 为钝角时,取负号:

$$V=Ah=\pm|\boldsymbol{a}\times \boldsymbol{b}||\boldsymbol{c}|\cos\alpha=\pm[\boldsymbol{abc}].$$

向量的混合积$[\boldsymbol{abc}]=(\boldsymbol{a}\times \boldsymbol{b})\cdot \boldsymbol{c}$ 的几何意义是:$[\boldsymbol{abc}]$的绝对值等于以向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 为棱的平行六面体的体积.如果向量 $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ 组成右(左)手系,那么混合积的符号是正(负)的.

下面推导三向量的混合积的坐标表示式.在右手直角坐标系$\{O;\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$中,设 $\boldsymbol{a}=x_1\boldsymbol{i}+y_1\boldsymbol{j}+z_1\boldsymbol{k}$, $\boldsymbol{b}=x_2\boldsymbol{i}+y_2\boldsymbol{j}+z_2\boldsymbol{k}$, $\boldsymbol{c}=x_3\boldsymbol{i}+y_3\boldsymbol{j}+z_3\boldsymbol{k}$.因为

$$\boldsymbol{a}\times \boldsymbol{b}=(y_1z_2-z_1y_2)\boldsymbol{i}+(z_1x_2-x_1z_2)\boldsymbol{j}+(x_1y_2-y_1x_2)\boldsymbol{k}.$$

$$[\boldsymbol{abc}]=x_3(y_1z_2-z_1y_2)+y_3(z_1x_2-x_1z_2)+z_3(x_1y_2-y_1x_2)\quad(1.4)$$

例5　在右手直角坐标系$\{O;\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$中, $\boldsymbol{a}=\boldsymbol{i}-2\boldsymbol{j}+2\boldsymbol{k}$, $\boldsymbol{b}=3\boldsymbol{i}-4\boldsymbol{k}$, $\boldsymbol{c}=\boldsymbol{i}-\boldsymbol{j}+2\boldsymbol{k}$,试求以向量 $\boldsymbol{a}, \boldsymbol{b}$ 和 $\boldsymbol{c}$ 为棱围成的六面体的体积.

解　所求体积等于$|[\boldsymbol{abc}]|$,由公式1.3,得

$$\boldsymbol{a}\times\boldsymbol{b}=8\boldsymbol{i}+10\boldsymbol{j}+6\boldsymbol{k}.$$

所以

$$|[\boldsymbol{abc}]|=|(8\boldsymbol{i}+10\boldsymbol{j}+6\boldsymbol{k})\cdot(\boldsymbol{i}-\boldsymbol{j}+2\boldsymbol{k})|=10.$$

例6　空间三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面的充要条件是$[\boldsymbol{abc}]=0$.

解　假设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面.当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 共线时,$\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$,自然有$[\boldsymbol{abc}]=0$;当 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 不共线时,$\boldsymbol{a}\times\boldsymbol{b}$ 垂直于 $\boldsymbol{a},\boldsymbol{b}$ 所在的平面,因而 $\boldsymbol{a}\times\boldsymbol{b}$ 垂直于 $\boldsymbol{c}$,所以$[\boldsymbol{abc}]=0$.反之,假设$[\boldsymbol{abc}]=0$.当 $\boldsymbol{a}\times\boldsymbol{b}=\boldsymbol{0}$ 时,有 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 共线,故有 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面;当 $\boldsymbol{a}\times\boldsymbol{b}\neq\boldsymbol{0}$ 时,有 $\boldsymbol{a}\times\boldsymbol{b}$ 垂直于 $\boldsymbol{c}$,又因为 $\boldsymbol{a}\times\boldsymbol{b}$ 亦垂直于 $\boldsymbol{a}$ 及 $\boldsymbol{b}$,从而 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面.

第三节　向量空间

第二节以有向线段为直观背景建立了空间向量概念,研究了其运算性质.有许多问题的研究,如运筹学中的线性规划、控制论中的线性系统等都需要应用高维向量的概念.因而需要以空间向量为背景,将向量及其运算的本质抽象为一个更一般的代数结构——向量空间.向量空间、线性变换及与之相联系的矩阵理论是线性代数的中心内容,它们是研究线性问题的有力工具;同时由于很多非线性问题常常可转化为线性问题,因而在科学研究及工程计算中线性代数就显得特别重要.

本节内容比较抽象.在学习过程中,一定要注意思想的来源,回想所讨论的问题在平面和空间直角坐标系中的原型,看到抽象的概念和复杂的算式背后的几何直观,以便清楚地看到正在做什么及这样做的几何意义.如能这样,读者就容易克服抽象概念带来的困难,使学习更加轻松.

3.1　向量空间的定义及基本性质

1. 向量空间的定义

在第二节中,我们在空间所有向量的集合中定义了两种运算——向量加法,向量与数的乘积,即线性运算.令 V_3 为空间所有向量的集合.设 $\mathbf{R}$ 为实数集,任给 $l,k\in\mathbf{R}$,对空间中任意三个向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 及零向量 $\boldsymbol{0}$,线性运算显然满足下列运算规律:

(1) $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$(交换律);

(2) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})=\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$(结合律);

(3) $\boldsymbol{0}+\boldsymbol{a}=\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$;

(4) $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$;

(5) $1\boldsymbol{a}=\boldsymbol{a}$;

(6) $k(l\boldsymbol{a})=(kl)\boldsymbol{a}$;

(7) $k(\boldsymbol{a}+\boldsymbol{b})=k\boldsymbol{a}+k\boldsymbol{b}$;

(8) $(k+l)\boldsymbol{a}=k\boldsymbol{a}+l\boldsymbol{a}$.

事实上,空间向量的其他性质都可由这八条性质推导出来,但这八条性质中的任何一条都不能由其他条推导出来.如果把空间所有向量的集合 V_3 换成一个一般的集合 V,并且在 V 中也定义一个类似于空间向量加法的运算,用 V 和 **R** 定义一个类似于空间向量与数乘积的运算,那么这种抽象的代数系统是否也有与空间向量类似的性质呢?通过下面的讨论会发现,它有与空间向量几乎相同的性质.

定义 1.2 设 V 是非空集合,**R** 为实数域(即实数集及其上的加法和乘法运算),如果它们满足以下性质(这些性质称为公理),就称 V 是实数域 **R** 上的向量空间或线性空间,并称 V 的元素为实数域 **R** 上的向量:

封闭公理

(1) 加法闭合:对任意两个元 $\boldsymbol{u},\boldsymbol{v}\in V$,总有唯一的元 $\boldsymbol{w}\in V$ 与之对应,叫做 $\boldsymbol{u}$ 与 $\boldsymbol{v}$ 的和,记做 $\boldsymbol{w}=\boldsymbol{u}+\boldsymbol{v}$;

(2) 数乘闭合:对任意 $\boldsymbol{u}\in V$ 与任意实数 $a\in\mathbf{R}$,总有唯一的元 $\boldsymbol{w}\in V$ 与之对应,叫做 a 与 $\boldsymbol{u}$ 的积,记做 $\boldsymbol{w}=a\boldsymbol{u}$.

加法公理(V 在加法运算下为一个 Abel 群,参考附录 A.1 的 Abel 群定义)

(3) 交换性:对任意 $\boldsymbol{u},\boldsymbol{v}\in V$,有 $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$;

(4) 结合性:对任意 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V$,有 $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$;

(5) 零元的存在性:存在 $\mathbf{0}\in V$,对任意 $\boldsymbol{u}\in V$,有 $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$,$\mathbf{0}$ 称为零元;

(6) 负元的存在性:对任意 $\boldsymbol{u}\in V$,存在 $\boldsymbol{v}\in V$,使 $\boldsymbol{u}+\boldsymbol{v}=\mathbf{0}$,$\boldsymbol{v}$ 称为 $\boldsymbol{u}$ 的负元,记做 $-\boldsymbol{u}$;

数乘的公理

(7) 结合性:对任意 $\boldsymbol{u}\in V$ 和任意实数 $a,b\in\mathbf{R}$,有 $a(b\boldsymbol{u})=(ab)\boldsymbol{u}$;

(8) 对 V 中加法的分配性:对任意 $\boldsymbol{u},\boldsymbol{v}\in V$ 和任意实数 $a\in\mathbf{R}$,有 $a(\boldsymbol{u}+\boldsymbol{v})=a\boldsymbol{u}+a\boldsymbol{v}$;

(9) 对数的加法的分配性:对任意 $\boldsymbol{u}\in V$ 和任意数 $a,b\in\mathbf{R}$,有 $(a+b)\boldsymbol{u}=a\boldsymbol{u}+b\boldsymbol{u}$;

(10) 对实数 1 及任意 $\boldsymbol{u}\in V$,有 $1\boldsymbol{u}=\boldsymbol{u}$.

满足以上十条公理的运算称为线性运算,定义了线性运算的集合称为向量空间.

*注:将定义 1.2 中实数域 **R** 换为附录 A.1 中讲述的一般域 F,令 V 及 F 元素满足与定义 1.2 相同的十条公理,就可以定义一般域 F 上的线性空间 $V(F)$.同样,本章讲述的其他定义及所有命题与定理,只要将其中实数域 **R** 换

为域 F,逐字逐句重复,就可得到关于 $V(F)$的同样正确的结果 .

显然,以上十条都是比照前面第二节中的空间向量加法与数乘这两种运算的性质而抽象出来的 .从下面的例子就会看到,数学的很多对象,例如数组、函数等等,都可定义具有空间向量加法和数乘同样性质的两种运算,而且这些运算也具有与空间向量的线性运算相类似的性质 .既然如此,就可以用空间向量的集合及其运算为直观背景,提出向量空间这个抽象概念 .在研究抽象的向量空间时,人们不再关心这些集合中元素的特定属性,也不关心这些运算的具体形式,只要求运算满足以上性质 .在一个具体数学对象构成的集合 V 上规定了两种运算,只要它们满足定义 1.2 列举的十条公理,那么 V 就是向量空间,就可以将一般抽象向量空间得到的定理用到 V 上去 .

例 1　设 V_1 为平行于一条给定直线的所有向量构成的集合, V_2 为平行于一个给定平面的所有向量构成的集合, V_3 为空间所有向量构成的集合 .证明:按空间向量加法和数乘运算, V_1, V_2 和 V_3 分别构成实数域 $\mathbf{R}$ 上的向量空间 .本书称此类向量空间为几何空间,以区别一般的抽象的线性空间 .

证　任给向量 $\boldsymbol{a}, \boldsymbol{b} \in V_1$,任给 $k \in \mathbf{R}$,向量 $\boldsymbol{a}, \boldsymbol{b}$ 均平行给定直线 l,由几何空间向量加法定义可知,向量 $\boldsymbol{a}+\boldsymbol{b}$ 和向量 $k\boldsymbol{a}$ 还平行于直线 l .这说明 V_1 上向量加法及数乘满足定义 1.2(1),(2),易验证 V_1 上加法及数乘还满足定义 1.2 性质(3)～性质(10) .因此, V_1 是 $\mathbf{R}$ 上的向量空间 .类似地可证 V_2, V_3 也为 $\mathbf{R}$ 上的向量空间 .

例 2　令 $P[x]_n = \{ a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \mid a_n, a_{n-1}, \cdots, a_0 \in \mathbf{R} \}$,即所有次数不超过 n 的实系数多项式的集合 .任给

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0,\ b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0 \in P[x]_n,$$

定义加法如下:

$$\begin{aligned}&(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) + (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0)\\ &= (a_n + b_n) x^n + (a_{n-1} + b_{n-1}) x^{n-1} + \cdots + (a_0 + b_0) \in P[x]_n;\end{aligned}$$

对任意 $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \in V$,对任给 $r \in \mathbf{R}$,定义数乘如下:

$$r(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) = ra_n x^n + ra_{n-1} x^{n-1} + \cdots + ra_0 \in P[x]_n,$$

则定义的向量加法与数乘满足定义 1.2(1)～(10) .所以 $P[x]_n$ 为实数域 $\mathbf{R}$ 上的向量空间 .

例 3　设 $V = \{ f \mid f$ 为定义在$[a, b]$上的实函数$\}$ .V 中的加法定义如下:任给 $f, g \in V$, $(f+g)(x) = f(x) + g(x)$.任给 $f \in V$, $r \in \mathbf{R}$,定义数乘如下:$(rf)(x) = rf(x)$,其中 rf 的定义域为 f 的定义域 .易验证 V 为 $\mathbf{R}$ 上的向量空间 .

类似地,某一区间的所有连续函数、所有可微函数、所有可积函数都分别构

成实数域上的向量空间.

由 n 个实数 $a_1, a_2, \cdots, a_n$ 构成的有序数组,记为$(a_1, a_2, \cdots, a_n)$,叫做 **n** 元行向量或 **n** 元数组.全体 n 元数组构成的集合记做 $\mathbf{R}^n$.任给$(a_1, a_2, \cdots, a_n)$,$(b_1, b_2, \cdots, b_n) \in \mathbf{R}^n$,任意 $r \in \mathbf{R}$,在 $\mathbf{R}^n$ 中定义加法及数乘如下:

$$(a_1, a_2, \cdots, a_n) + (b_1, b_2, \cdots, b_n) = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n).$$

$$r(a_1, a_2, \cdots, a_n) = (ra_1, ra_2, \cdots, ra_n).$$

显然定义的加法和数乘满足定义 1.2(1)～(10),因此 $\mathbf{R}^n$ 也为 $\mathbf{R}$ 上的向量空间.称此向量空间为 **n** 元行向量空间.类似地称

$$\begin{matrix} a_1 \\ \cdots \\ a_n \end{matrix}$$

为 n 元列向量.全体 n 元列向量构成的集合也记做 $\mathbf{R}^n$.类似地可定义的 n 元列向量加法和数乘,$\mathbf{R}^n$ 也构成向量空间,称此向量空间为 **n** 元列向量空间.

例 4　设 V 是简谐运动的微分方程

$$y'' + \omega^2 y = 0 \quad (\omega \neq 0)$$

所有解的集合,即 $V = \{ y \mid y'' + \omega^2 y = 0 \}$.任给 $y_1, y_2 \in V$, $r \in \mathbf{R}$,运算 $y_1 + y_2$, ry_1 与例 3 类似.只要 V 对于加法和数乘封闭, V 的加法和数乘就满足定义 1.2 (1)～(10), V 就是 $\mathbf{R}$ 上的向量空间.下面验证 V 对于加法和数乘封闭.因为

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(y_1 + y_2) + \omega^2(y_1 + y_2) = \frac{\mathrm{d}^2 y_1}{\mathrm{d}t^2} + \omega^2 y_1 + \frac{\mathrm{d}^2 y_2}{\mathrm{d}t^2} + \omega^2 y_2 = 0,$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(ry_1) + \omega^2(ry_1) = r\frac{\mathrm{d}^2 y_1}{\mathrm{d}t^2} + \omega^2 y_1 = 0.$$

所以 $y_1 + y_2$, $ry_1 \in V$.因此 V 也为 $\mathbf{R}$ 上的向量空间.

2. 向量空间的性质

尽管下面两个命题的结论对几何空间是显然的,但对一般向量空间就必须进行证明.由于一般向量空间就是由定义 1.2 中十条性质规定的一个抽象对象,在以下定理的证明过程中不能直接使用定义 1.2(1)～(10)以外的任何结论.

命题 1.9　设 V 是 $\mathbf{R}$ 上的一个向量空间,则下面等式成立:

(1) 如果 $\boldsymbol{e} \in V$,任给 $\boldsymbol{u} \in V$, $\boldsymbol{e} + \boldsymbol{u} = \boldsymbol{u}$,则 $\boldsymbol{e} = \mathbf{0}$,即零向量唯一;

(2) $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$,如果 $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} = \mathbf{0}$,则 $\boldsymbol{v} = \boldsymbol{w} = -\boldsymbol{u}$,即负元唯一;

(3) 任给 $\boldsymbol{u} \in V$, $0\boldsymbol{u} = \mathbf{0}$;

(4) 任给 $k \in \mathbf{R}$, $k\mathbf{0} = \mathbf{0}$;

(5) $k \in \mathbf{R}$, $\boldsymbol{u} \in V$,当 $k \neq 0$ 且 $\boldsymbol{u} \neq \mathbf{0}$ 时, $k\boldsymbol{u} \neq \mathbf{0}$.

证　在证明前,应注意, $\mathbf{0}$ 与0具有完全不同的意义.$\mathbf{0}$ 是零向量.

(1) 由定义 1.2(5),有 $\boldsymbol{e}=\boldsymbol{e}+\boldsymbol{0}=\boldsymbol{0}$.

(2) 由定义 1.2(5),(4),有

$$\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{0}=\boldsymbol{v}+(\boldsymbol{u}+\boldsymbol{w})=(\boldsymbol{v}+\boldsymbol{u})+\boldsymbol{w}=\boldsymbol{0}+\boldsymbol{w}=\boldsymbol{w}.$$

(3) 任给 $k\in\mathbf{R}$,由定义 1.2(9),有 $k\boldsymbol{u}=(k+0)\boldsymbol{u}=k\boldsymbol{u}+0\boldsymbol{u}$.对于 $k\boldsymbol{u}$,由定义 1.2(6)有 $-(k\boldsymbol{u})$使得

$$\begin{aligned}\boldsymbol{0}&=k\boldsymbol{u}+(-(k\boldsymbol{u}))=(k\boldsymbol{u}+0\boldsymbol{u})+(-(k\boldsymbol{u}))=k\boldsymbol{u}+(-(k\boldsymbol{u})+0\boldsymbol{u})\\&=k\boldsymbol{u}+(-(k\boldsymbol{u}))+0\boldsymbol{u}=\boldsymbol{0}+0\boldsymbol{u}=0\boldsymbol{u}.\end{aligned}$$

(4) 由定义 1.2(5),(6),(8),有 $k\boldsymbol{0}=k(\boldsymbol{0}+\boldsymbol{0})=k\boldsymbol{0}+k\boldsymbol{0}$,且可推知

$$\boldsymbol{0}=k\boldsymbol{0}+(-(k\boldsymbol{0}))=(k\boldsymbol{0}+k\boldsymbol{0})+(-(k\boldsymbol{0}))=k\boldsymbol{0}+(k\boldsymbol{0}+(-(k\boldsymbol{0})))=k\boldsymbol{0}.$$

(5) 若 $k\neq0$, $\boldsymbol{u}\neq\boldsymbol{0}$,而 $k\boldsymbol{u}=\boldsymbol{0}$.由 1.2(7),(10)和本命题的(4)及 $k\boldsymbol{u}=\boldsymbol{0}$,有

$$\boldsymbol{0}=k^{-1}\boldsymbol{0}=k^{-1}(k\boldsymbol{u})=(k^{-1}k)\boldsymbol{u}=1\boldsymbol{u}=\boldsymbol{u}.$$

这与 $\boldsymbol{u}\neq\boldsymbol{0}$ 矛盾,故当 $k\neq0$ 且 $\boldsymbol{u}\neq\boldsymbol{0}$ 时, $k\boldsymbol{u}\neq\boldsymbol{0}$.□

命题 1.10 设 V 是 $\mathbf{R}$ 上的一个向量空间.任给 $k\in\mathbf{R}$, $\boldsymbol{u}\in V$,则

$$(-k)\boldsymbol{u}=k(-\boldsymbol{u})=-(k\boldsymbol{u}).$$

证 由定义 1.2(9)及命题 1.9(3),有 $k\boldsymbol{u}+(-k)\boldsymbol{u}=(k+(-k))\boldsymbol{u}=0\boldsymbol{u}=\boldsymbol{0}$.再由命题 1.9(2)负元唯一,有$(-k)\boldsymbol{u}=-(k\boldsymbol{u})$;又由定义 1.2(8)及命题 1.9(4),有 $k\boldsymbol{u}+k(-\boldsymbol{u})=k(\boldsymbol{u}+(-\boldsymbol{u}))=k\boldsymbol{0}=\boldsymbol{0}$.再由命题 1.9(2)负元唯一,有 $k(-\boldsymbol{u})=-(k\boldsymbol{u})$.□

对任意三个向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3$,有$(\boldsymbol{u}_1+\boldsymbol{u}_2)+\boldsymbol{u}_3=\boldsymbol{u}_1+(\boldsymbol{u}_2+\boldsymbol{u}_3)$,即相加结果与相加的顺序无关,故三个向量相加可表示为 $\boldsymbol{u}_1+\boldsymbol{u}_2+\boldsymbol{u}_3$.同理,多个向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 相加可表示为 $\boldsymbol{u}_1+\boldsymbol{u}_2+\cdots+\boldsymbol{u}_n$.为了简便,把 $\boldsymbol{u}+(-\boldsymbol{v})$表示为 $\boldsymbol{u}-\boldsymbol{v}$.根据命题 1.9,1.10,可把定义 1.2 中的(8),(9),(10)分别推广为:任给 $a,b\in\mathbf{R}$

(8′) $a(\boldsymbol{u}\pm\boldsymbol{v})=a\boldsymbol{u}\pm a\boldsymbol{v}$;

(9′) $(a\pm b)\boldsymbol{u}=a\boldsymbol{u}\pm b\boldsymbol{u}$;

(10′) $(\pm1)\boldsymbol{u}=\pm\boldsymbol{u}$;

证 (8′) $a(\boldsymbol{u}-\boldsymbol{v})=a(\boldsymbol{u}+(-\boldsymbol{v}))=a\boldsymbol{u}+a(-\boldsymbol{v})=a\boldsymbol{u}+(-(a\boldsymbol{v}))=a\boldsymbol{u}-a\boldsymbol{v}$.

(9′) $(a-b)\boldsymbol{u}=(a+(-b))\boldsymbol{u}=a\boldsymbol{u}+(-b)\boldsymbol{u}=a\boldsymbol{u}+(-(b\boldsymbol{u}))=a\boldsymbol{u}-b\boldsymbol{u}$.

(10′) 由$(-1)\boldsymbol{u}=1(-\boldsymbol{u})=-\boldsymbol{u}$(命题 1.10),知其成立.□

此外,用数学归纳法不难再把(8′)与(9′)分别推广为:

(8″) $a(\boldsymbol{u}_1\pm\boldsymbol{u}_2\pm\cdots\pm\boldsymbol{u}_n)=a\boldsymbol{u}_1\pm a\boldsymbol{u}_2\pm\cdots\pm a\boldsymbol{u}_n$;

(9″) $(a_1\pm a_2\pm\cdots\pm a_n)\boldsymbol{u}=a_1\boldsymbol{u}\pm a_2\boldsymbol{u}\pm\cdots\pm a_n\boldsymbol{u}$.

3.2　子空间

1. 子空间

3.1 小节例 1 中的几何空间 V_1, V_2 与 V_3 都是向量空间,向量加法与数乘运算都是相同的, V_1 与 V_2 都被包含在 V_3 中.研究这些空间彼此的关系对解决几何空间的许多问题有很大帮助.由此引出如下概念.

定义 1.3　设 V 是实数域 $\mathbf{R}$ 上的一个向量空间, S 是 V 的一个子集.如果 S 在 V 的加法和数乘下也是实数域 $\mathbf{R}$ 上的一个向量空间,则说 S 是向量空间 V 的一个子空间.

由定义 1.3 可知,$\mathbf{R}$ 上向量空间 V 本身也是 V 的一个子空间.由向量空间的定义可知,任何一个向量空间都有零向量 $\mathbf{0}$.在向量空间 V 中,仅由其零元素组成的子集 $S=\{\mathbf{0}\}$也是 V 的子空间,且是最小的子空间,即它被包含在 V 的所有子空间中.

例 5　$P[x]_n=\{a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0 \mid a_n, a_{n-1},\cdots,a_0\in\mathbf{R}\}$,即所有次数不超过 n 的实系数多项式为实数域上的向量空间, $P[x]_1$, $P[x]_2$, $\cdots$, $P[x\}_{n-1}$作为实数域 $\mathbf{R}$ 上的向量空间,都是 $P[x]_n$ 的子空间.

在学习下面一些定理的时候,如果首先能用几何空间向量的例子说明每个定理的几何涵义,以几何直觉引导我们思考,进而以定义 1.2 为根据来证明定理,那么定理就比较容易理解了.

定理 1.1　实数域 $\mathbf{R}$ 上的向量空间 V 的一个非空子集 S 为 V 的子空间的充要条件是

(1) 如果 $\boldsymbol{u}, \boldsymbol{v}\in S$,则 $\boldsymbol{u}+\boldsymbol{v}\in S$;

(2) 如果 $\boldsymbol{u}\in S$,对任意 $a\in\mathbf{R}$,则 $a\boldsymbol{u}\in S$.

证　条件的必要性是显然的.

下面证明充分性.假设 S 满足(1)与(2).仅验证定义 1.2(5),(6),其它各条都显然成立.若 $\boldsymbol{u}\in S$,由命题 1.9 和本定理条件(2),对 $0\in\mathbf{R}$,有零向量 $0\boldsymbol{u}=\mathbf{0}\in S$.由命题 1.10 及本定理条件(2)知,如果 $\boldsymbol{u}\in S$,则$(-1)\boldsymbol{u}=-\boldsymbol{u}\in S$,即负元存在.于是 S 满足定义 1.2(5),(6).□

定理 1.1 说明,若验证向量空间 V 的子集 S 为 V 的子空间,不需要逐条验证定义 1.2(1)～(10),而只需验证该定理的条件(1),(2)即可.

在几何空间 V_3 中,如果建立了空间仿射坐标系(注意,一般抽象向量空间现在还无坐标系的概念).V_3 中平行于各坐标轴、坐标面的全体向量分别构成几何空间 V_3 的子空间.

例 6　$V=\{(x_1,x_2,x_3)\in\mathbf{R}^3 \mid x_1+x_2+x_3=0\}$.证明: V 为三元行向量空间 $\mathbf{R}^3$ 的子空间.

证　任给$(x_1,x_2,x_3),(y_1,y_2,y_3)\in V$,任给$r\in\mathbf{R}$.

$$(x_1,x_2,x_3)+(y_1,y_2,y_3)=(x_1+y_1,x_2+y_2,x_3+y_3).$$

因为

$$x_1+y_1+x_2+y_2+x_3+y_3=(x_1+x_2+x_3)+(y_1+y_2+y_3)$$
$$=0+0=0,$$

所以$(x_1,x_2,x_3)+(y_1,y_2,y_3)\in V$.

$$r(x_1,x_2,x_3)=(rx_1,rx_2,rx_3).$$

因为

$$rx_1+rx_2+rx_3=r(x_1+x_2+x_3)=0,$$

所以$r(x_1,x_2,x_3)\in V$.由定理1.1知,V为向量空间$\mathbf{R}^3$的子空间.□

命题1.11　实数域$\mathbf{R}$上的向量空间V的两个子空间S_1与S_2的公共部分$S_1\cap S_2$必含$\mathbf{0}$.而且$S_1\cap S_2$也是V的一个子空间,叫做S_1与S_2的交空间.

证　由定理1.1知,V的任意子空间必含零向量$\mathbf{0}$,故$\mathbf{0}\in S_1\cap S_2$.欲证$S_1\cap S_2$为$V$的子空间,只要验证定理1.1条件(1)、(2)即可.若$\boldsymbol{u},\boldsymbol{v}\in S_1\cap S_2$,则$\boldsymbol{u},\boldsymbol{v}\in S_1$及$\boldsymbol{u},\boldsymbol{v}\in S_2$.因为$S_1,S_2$均为子空间,故$\boldsymbol{u}+\boldsymbol{v}\in S_1$及$\boldsymbol{u}+\boldsymbol{v}\in S_2$,从而$\boldsymbol{u}+\boldsymbol{v}\in S_1\cap S_2$,因此$S_1\cap S_2$满足定理1.1条件(1).再任取$\boldsymbol{u}\in S_1\cap S_2$及任意$a\in\mathbf{R}$,因为$\boldsymbol{u}\in S_1$,$\boldsymbol{u}\in S_2$及$S_1,S_2$均为子空间,故对任意$a\in\mathbf{R}$,$a\boldsymbol{u}\in S_1$,$a\boldsymbol{u}\in S_2$,因而$a\boldsymbol{u}\in S_1\cap S_2$,于是$S_1\cap S_2$满足条件定理1.1条件(2).□

在几何空间V_3的仿射坐标系中,若S_1和S_2分别为平行于xOy面及yOz面所有向量组成的集合,则S_1和S_2均为V_3的子空间,这时$S_1\cap S_2$是平行于y轴所有向量组成的集合,它也是V_3的一个子空间.

2.子空间的和

命题1.12　设S_1与S_2为实数域$\mathbf{R}$上的向量空间V的两个子空间.设

$$T=\{\boldsymbol{u}_1+\boldsymbol{u}_2\mid\boldsymbol{u}_1\in S_1,\boldsymbol{u}_2\in S_2\},$$

则T也是V的一个子空间,叫做S_1与S_2的和空间,记为S_1+S_2.

证　显然T为V的子集.若$\boldsymbol{u},\boldsymbol{v}\in T$,则必有$\boldsymbol{u}_1,\boldsymbol{v}_1\in S_1$,$\boldsymbol{u}_2,\boldsymbol{v}_2\in S_2$,使得$\boldsymbol{u}=\boldsymbol{u}_1+\boldsymbol{u}_2$,$\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$.因为$S_1$与$S_2$都是子空间,故$\boldsymbol{u}_1+\boldsymbol{v}_1\in S_1$,$\boldsymbol{u}_2+\boldsymbol{v}_2\in S_2$.再由

$$\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{u}_1+\boldsymbol{u}_2+\boldsymbol{v}_1+\boldsymbol{v}_2=(\boldsymbol{u}_1+\boldsymbol{v}_1)+(\boldsymbol{u}_2+\boldsymbol{v}_2),$$

可知$\boldsymbol{u}+\boldsymbol{v}\in T$.同理,对任意$a\in\mathbf{R}$,$a\boldsymbol{u}=a\boldsymbol{u}_1+a\boldsymbol{u}_2\in T$.由定理1.1可知,$T$为$V$的子空间.□

在几何空间 V_3 的仿射坐标系中,如 S_1 和 S_2 分别为平行于 x 轴及 y 轴所有向量组成的集合,它们均为 V_3 的子空间, S_1+S_2 是平行于 xOy 面所有向量组成的集合,它也是 V_3 的一个子空间.

不难把命题 1.11 与命题 1.12 推广到 n 个子空间的情形.

命题 1.13　设 $\boldsymbol{u}$ 为实数域 $\mathbf{R}$ 上的向量空间 V 中任意一个向量,令 $[\boldsymbol{u}]=\{a\boldsymbol{u}\mid$ 任意 $a\in\mathbf{R}\}$,则 $[\boldsymbol{u}]$ 是 V 的一个子空间.通常称 $[\boldsymbol{u}]$ 为 $\boldsymbol{u}$ 生成的子空间.

证　若 $\boldsymbol{w},\boldsymbol{v}\in[\boldsymbol{u}]$,则存在 $a,b\in\mathbf{R}$ 使得 $\boldsymbol{w}=a\boldsymbol{u}$, $\boldsymbol{v}=b\boldsymbol{u}$.于是

$$\boldsymbol{w}+\boldsymbol{v}=a\boldsymbol{u}+b\boldsymbol{u}=(a+b)\boldsymbol{u}\in[\boldsymbol{u}];$$

又任给 $r\in\mathbf{R}$,有 $r\boldsymbol{w}=r(a\boldsymbol{u})=(ra)\boldsymbol{u}\in[\boldsymbol{u}]$.由定理 1.1 可知, $[\boldsymbol{u}]$ 为 V 的一个子空间.□

例如,设 $\boldsymbol{u}$ 为第二节的几何空间 V_3 的一非零向量,则子空间 $[\boldsymbol{u}]$ 为与 $\boldsymbol{u}$ 共线的所有向量.

设 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 是 $\mathbf{R}$ 上的向量空间 V 中任意 n 个向量,由命题 1.12 及 1.13 可知, $[\boldsymbol{u}_1]+[\boldsymbol{u}_2]+\cdots+[\boldsymbol{u}_n]$ 也是 V 的子空间,称之为由 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 所生成的子空间,记为 $[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n]$.这个子空间是由所有形如 $a_1\boldsymbol{u}_1+a_2\boldsymbol{u}_2+\cdots+a_n\boldsymbol{u}_n$ 的向量构成的,其中 $a_1,a_2,\cdots,a_n\in\mathbf{R}$.容易验证 $[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n]$ 是包含 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 的 V 的最小的子空间,即 V 的包含 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 的子空间都包含 $[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n]$.在几何空间 V_3 中,任意找三个不共面的向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3$.对于空间任一向量 $\boldsymbol{u}$,存在 $a_1,a_2,a_3\in\mathbf{R}$ 使得

$$\boldsymbol{u}=a_1\boldsymbol{u}_1+a_2\boldsymbol{u}_2+a_3\boldsymbol{u}_3\in[\boldsymbol{u}_1]+[\boldsymbol{u}_2]+[\boldsymbol{u}_3].$$

命题 1.6 已证明了 a_1,a_2,a_3 是由向量 $\boldsymbol{u}$ 唯一确定的.若再增加一非零向量 $\boldsymbol{u}_4$,尽管也存在 $a_1,a_2,a_3,a_4\in\mathbf{R}$ 使得

$$\boldsymbol{u}=a_1\boldsymbol{u}_1+a_2\boldsymbol{u}_2+a_3\boldsymbol{u}_3+a_4\boldsymbol{u}_4\in[\boldsymbol{u}_1]+[\boldsymbol{u}_2]+[\boldsymbol{u}_3]+[\boldsymbol{u}_4],$$

但对 $\boldsymbol{u}$ 而言, a_1,a_2,a_3,a_4 不是唯一确定的,由此,有如下定义.

定义 1.4　设 S_1 与 S_2 均为 $\mathbf{R}$ 上的向量空间 V 的子空间, $S=S_1+S_2$ 为其和空间.如果 S 中每个向量表示为 $\boldsymbol{u}=\boldsymbol{u}_1+\boldsymbol{u}_2$ $(\boldsymbol{u}_1\in S_1,\boldsymbol{u}_2\in S_2)$ 的表示式是唯一的,则称 S 是 S_1 与 S_2 的直和,并且记为 $S=S_1\oplus S_2$ 或 $S=S_1\dotplus S_2$.

在几何空间 V_3 的仿射坐标系中,如 S_1 和 S_2 分别为平行于 xOy 面及 yOz 面的所有向量组成的集合,则 $V_3=S_1+S_2$.取平行于 y 轴的非零的向量 $\boldsymbol{a}$,因为 $\boldsymbol{a}$ 既平行于 xOy 面又平行于 yOz 坐标面,所以 $\boldsymbol{a}\in S_1\cap S_2$.既有 $\mathbf{0}=\boldsymbol{a}+(-\boldsymbol{a})$ 又有 $\mathbf{0}=\mathbf{0}+\mathbf{0}$.$\mathbf{0}$ 的表示式就不唯一,因此 S_1+S_2 不是直和.如果 S_1 和 S_2 分别为平行于 xOy 面及 z 轴的所有向量组成的集合,则 $V_3=S_1\oplus S_2$.下面是另外一个例子.

例 7　在三元行向量空间 $\mathbf{R}^3$ 中的子集 S_1,S_2,S_3 分别是由下面所有向量构成:

$$S_1 = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbf{R}\};$$
$$S_2 = \{(0, 0, x_3) \mid x_3 \in \mathbf{R}\};$$
$$S_3 = \{(0, x_2, x_3) \mid x_2, x_3 \in \mathbf{R}\}.$$

易验证它们都为 $\mathbf{R}^3$ 的子空间.显然 $\mathbf{R}^3 = S_1 + S_2$ 与 $\mathbf{R}^3 = S_1 + S_3$ 为同一个空间,即是由所有形如(x_1, x_2, x_3)(任意 $x_1, x_2, x_3 \in \mathbf{R}$)的向量构成的.但前者是直和,即 $\mathbf{R}^3 = S_1 \oplus S_2$;后者不是直和.因为 $\mathbf{R}^3$ 中的向量$(1,1,1)$既可表为 S_1 中的$(1,2,0)$与 S_3 中$(0,-1,1)$的和

$$(1,1,1) = (1,2,0) + (0,-1,1),$$

又可表为 S_1 中的$(1,0,0)$与 S_3 中的$(0,1,1)$的和

$$(1,1,1) = (1,0,0) + (0,1,1).$$

故 $\mathbf{R}^3 = S_1 + S_3$ 是 S_1 与 S_3 的和,但不是直和.虽然,$S_1 \oplus S_2 = S_1 + S_3$,但两边是不同的分解.

要检验 $S = S_1 + S_2$ 是直和,按定义需要检验每个向量的表示式都是唯一的,对于大多数情况,直接验证可能并不方便,下面的命题可简化这个过程.

命题 1.14　设 S_1 与 S_2 均为数域 $\mathbf{R}$ 上的向量空间 V 的子空间,且 $S = S_1 + S_2$.S 的每个向量的表示法是唯一的充要条件是零向量 $\mathbf{0}$ 的表示式是唯一的.

证　因为向量 $\mathbf{0}$ 自然是 S 中的一个向量,故条件显然是必要的.反之,设 $\mathbf{0}$ 的表示式是唯一的,即它只能表为 $\mathbf{0} = \mathbf{0} + \mathbf{0}$.今设 S 中的 $\boldsymbol{u}$ 既能表示为 $\boldsymbol{u} = \boldsymbol{u}_1 + \boldsymbol{u}_2$,又能表示为 $\boldsymbol{u} = \boldsymbol{v}_1 + \boldsymbol{v}_2$ ($\boldsymbol{u}_1, \boldsymbol{v}_1 \in S_1$, $\boldsymbol{u}_2, \boldsymbol{v}_2 \in S_2$),则由

$$-\boldsymbol{u} = -1\boldsymbol{u} = -1(\boldsymbol{v}_1 + \boldsymbol{v}_2) = (-1)\boldsymbol{v}_1 + (-1)\boldsymbol{v}_2 = -\boldsymbol{v}_1 - \boldsymbol{v}_2,$$

可得

$$\mathbf{0} = \boldsymbol{u} + (-\boldsymbol{u}) = \boldsymbol{u}_1 + \boldsymbol{u}_2 - \boldsymbol{v}_1 - \boldsymbol{v}_2 = (\boldsymbol{u}_1 - \boldsymbol{v}_1) + (\boldsymbol{u}_2 - \boldsymbol{v}_2).$$

由于 $\boldsymbol{u}_1 - \boldsymbol{v}_1 \in S_1$, $\boldsymbol{u}_2 - \boldsymbol{v}_2 \in S_2$,且 $\mathbf{0}$ 只能表为 $\mathbf{0} + \mathbf{0}$,故必有 $\boldsymbol{u}_1 - \boldsymbol{v}_1 = \mathbf{0}$, $\boldsymbol{u}_2 - \boldsymbol{v}_2 = \mathbf{0}$,即有 $\boldsymbol{u}_1 = \boldsymbol{v}_1$, $\boldsymbol{u}_2 = \boldsymbol{v}_2$.可见 $\boldsymbol{u}$ 的表示式是唯一的,条件的充分性得证.□

定义 1.4 与命题 1.14 均不难被推广到 n 个子空间的情形.

定理 1.2　设 S_1 与 S_2 均为实数域 $\mathbf{R}$ 上的向量空间 V 的子空间,$S = S_1 + S_2$.则 $S = S_1 \oplus S_2$ 当且仅当 $S_1 \cap S_2 = \{\mathbf{0}\}$.

证　如果 $S_1 \cap S_2$ 含有向量 $\boldsymbol{u} \neq \mathbf{0}$,则 $\mathbf{0}$ 既可表示为 $\mathbf{0} = \mathbf{0} + \mathbf{0}$,又可表示为 $\mathbf{0} = \boldsymbol{u} + (-\boldsymbol{u})$,$\mathbf{0}$ 的表示式就不是唯一的.由定义 1.4 知,S 不是 S_1 与 S_2 的直和,故条件是必要的.反之,设 $S_1 \cap S_2 = \{\mathbf{0}\}$.若 $\mathbf{0}$ 可表示为 $\mathbf{0} = \boldsymbol{u}_1 + \boldsymbol{u}_2$,其中 $\boldsymbol{u}_1 \in S_1$, $\boldsymbol{u}_2 \in S_2$,则由此式可知 $-\boldsymbol{u}_1 = \boldsymbol{u}_2$,即 $-\boldsymbol{u}_1 \in S_2$,从而必有 $\boldsymbol{u}_1 \in S_2$.于是 $\boldsymbol{u}_1 \in S_1 \cap S_2 = \{\mathbf{0}\}$,即知 $\boldsymbol{u}_1 = \mathbf{0}$,从而 $\boldsymbol{u}_2 = -\boldsymbol{u}_1 = \mathbf{0}$.可见,$\mathbf{0}$ 只能表示为 $\mathbf{0} + \mathbf{0}$,即 $\mathbf{0}$ 的表示式是唯一的.由命题 1.14 知,$S = S_1 \oplus S_2$.故条件为充分的.□

例 8　设 U 和 V 是三元行向量空间 $\mathbf{R}^3$ 的两个子集合,其中

$$U = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 0, x_2 + x_3 = 0\},$$
$$V = \{(x_1, x_2, x_3) \mid x_3 = 0\}.$$

试证明:

(1) U 和 V 都是 $\mathbf{R}^3$ 的子空间;

(2) $\mathbf{R}^3 = U \oplus V$.

证　(1) 任给 $\boldsymbol{a} = (a_1, a_2, a_3)$, $\boldsymbol{b} = (b_1, b_2, b_3) \in U$, 任给 $k \in \mathbf{R}$, 有 $k\boldsymbol{a} = (ka_1, ka_2, ka_3)$ 及

$$ka_1 + ka_2 + ka_3 = k(a_1 + a_2 + a_3) = 0, ka_2 + ka_3 = k(a_2 + a_3) = 0,$$

即 $k\boldsymbol{a} \in U$. 另外

$$\boldsymbol{a} + \boldsymbol{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

显然有

$$a_1 + b_1 + a_2 + b_2 + a_3 + b_3 = 0, a_2 + b_2 + a_3 + b_3 = 0,$$

即 $\boldsymbol{a} + \boldsymbol{b} \in U$. 因此 U 是 $\mathbf{R}^3$ 的子空间. 同理可证, V 是 $\mathbf{R}^3$ 的子空间.

(2) 任给 $\boldsymbol{a} = (a_1, a_2, a_3) \in U \cap V$, 有

$$\begin{aligned} a_1 + a_2 + a_3 &= 0, \\ a_2 + a_3 &= 0, \\ a_3 &= 0. \end{aligned}$$

显然 $a_1 = a_2 = a_3 = 0$, 即 $U \cap V = \{\mathbf{0}\}$ 这说明 $U + V$ 是直和. 任给 $\boldsymbol{a} = (a_1, a_2, a_3) \in \mathbf{R}^3$.

$$\boldsymbol{a} = (a_1, a_2, a_3) = (a_1, a_2 + a_3, 0) + (0, -a_3, a_3).$$

易验证 $(a_1, a_2 + a_3, 0) \in V$, $(0, -a_3, a_3) \in U$. 因此 $\mathbf{R}^3 = U \oplus V$. □

3.3　向量组的线性相关性

在几何空间 V_3 中任一向量 $\boldsymbol{a}$ 可表示为 $\boldsymbol{a} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$, 其中 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 为 V_3 中任意三个不共面的向量. 由命题 1.6 可知, 这个表示式是唯一的. 从子空间角度来看, $V_3 = [\boldsymbol{e}_1] \oplus [\boldsymbol{e}_2] \oplus [\boldsymbol{e}_3]$. V_3 中的任一向量都可表示为三个向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 的线性组合, 从而建立了空间的仿射坐标系. 下面将几何空间中这一重要性质推广到一般的向量空间中去. 首先引入两个重要的概念.

定义 1.5　设 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 是实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量, 如果存在 $a_1, a_2, \cdots, a_n \in \mathbf{R}$, 使得

$$\boldsymbol{u} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n,$$

则说向量 $\boldsymbol{u}$ 是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的一个线性组合, 或说 $\boldsymbol{u}$ 可由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性表示.

实际上,向量 $\boldsymbol{u}$ 是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的一个线性组合的充分必要条件是 $\boldsymbol{u} \in [\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n]$.下面的定义是命题 1.2,1.3 中两个向量共线、三个向量共面性质的抽象和推广 .因此在学习下面的定义时要注意与命题 1.2,1.3 对比 .

定义 1.6　设 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 是实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量,如果有不全为 0 的数 $a_1, a_2, \cdots, a_n \in \mathbf{R}$,使得

$$a_1 \boldsymbol{v}_1 + a_2 \boldsymbol{v}_2 + \cdots + a_n \boldsymbol{v}_n = \mathbf{0},$$

则说向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关,否则就说向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关 .

定义 1.6 指明,向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 不是线性相关的,就是线性无关的 .由此可知:(1)如果向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,且

$$a_1 \boldsymbol{v}_1 + a_2 \boldsymbol{v}_2 + \cdots + a_n \boldsymbol{v}_n = \mathbf{0},$$

则必有 $a_1 = a_2 = \cdots = a_n = 0$;(2)反之,若 $a_1, a_2, \cdots, a_n \in \mathbf{R}$,使得

$$a_1 \boldsymbol{v}_1 + a_2 \boldsymbol{v}_2 + \cdots + a_n \boldsymbol{v}_n = \mathbf{0},$$

就有 $a_1 = a_2 = \cdots = a_n = 0$,那么向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 是线性无关的 .

例 9　讨论三元行向量空间 $\mathbf{R}^3$ 中下列向量组的线性相关性:

(1) $(1,0,0),(0,1,0),(0,0,1)$;

(2) $(0,0,1),(0,1,1),(1,1,1)$;

(3) $(2,1,0),(1,-1,2),(0,3,-4)$.

解　(1) 设存在 $k_1, k_2, k_3 \in \mathbf{R}$,使

$$k_1(1,0,0) + k_2(0,1,0) + k_3(0,0,1) = (0,0,0),$$

则

$$(k_1,0,0) + (0,k_2,0) + (0,0,k_3) = (k_1,k_2,k_3) = (0,0,0).$$

因此 $k_1 = k_2 = k_3 = 0$,所以$(1,0,0),(0,1,0),(0,0,1)$线性无关 .

(2) 设存在 $k_1, k_2, k_3 \in \mathbf{R}$,使

$$k_1(0,0,1) + k_2(0,1,1) + k_3(1,1,1) = (0,0,0),$$

则

$$\begin{aligned}&(0,0,k_1) + (0,k_2,k_2) + (k_3,k_3,k_3)\\ =\ &(k_3, k_2 + k_3, k_1 + k_3 + k_3) = (0,0,0).\end{aligned}$$

即

$$\begin{aligned} k_3 &= 0\\ k_2 + k_3 &= 0\\ k_1 + k_2 + k_3 &= 0\end{aligned}$$

解之,得 $k_1 = k_2 = k_3 = 0$.所以$(0,0,1),(0,1,1),(1,1,1)$线性无关 .

(3) 因为

$$(-1)(2,1,0) + 2(1,-1,2) + (0,3,-4) = (0,0,0),$$

所以$(2,1,0)$,$(1,-1,2)$,$(0,3,-4)$线性相关.

例 10　设 $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 是实数域 **R** 上的向量空间 V 中的一组向量.试证:

(1) 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 中有某个 $\boldsymbol{v}_i = \mathbf{0}$,则 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关.

(2) 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关,则 $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 也线性相关.

(3) 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,则 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{n-1}$ 也线性无关.

(4) 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关,则对任给 $l_1, l_2, \cdots, l_n \in \mathbf{R}$, $l_1\boldsymbol{v}_1, l_2\boldsymbol{v}_2, \cdots, l_n\boldsymbol{v}_n$ 线性相关.

(5) 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,则对任给 n 个全不为零的数 $l_1, l_2, \cdots, l_n$, $l_1\boldsymbol{v}_1, l_2\boldsymbol{v}_2, \cdots, l_n\boldsymbol{v}_n$ 线性无关.

证　(1) 因为

$$0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \cdots + 0\boldsymbol{v}_{i-1} + 1\boldsymbol{v}_i + 0\boldsymbol{v}_{i+1} + \cdots + 0\boldsymbol{v}_n = \mathbf{0},$$

$\boldsymbol{v}_i$ 的系数为 $1 \neq 0$,所以 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关.

(2) 存在不全为零的数 $l_1, l_2, \cdots, l_n$,使得

$$l_1\boldsymbol{v}_1 + l_2\boldsymbol{v}_2 + \cdots + l_n\boldsymbol{v}_n = \mathbf{0}.$$

因此

$$0\boldsymbol{u} + l_1\boldsymbol{v}_1 + l_2\boldsymbol{v}_2 + \cdots + l_n\boldsymbol{v}_n = \mathbf{0},$$

则 $0, l_1, l_2, \cdots, l_n$ 也不全为零,故 $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 也线性相关.

(3) 反证,若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{n-1}$ 线性相关,则由(2)知,$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关.这与 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关矛盾.

(4) 如果 $l_1, l_2, \cdots, l_n$ 有一个为零,设 $l_i = 0$,则 $l_i\boldsymbol{v}_i = \mathbf{0}$,由(1)知,$l_1\boldsymbol{v}_1, l_2\boldsymbol{v}_2, \cdots, l_n\boldsymbol{v}_n$ 线性相关.下面假设,$l_1, l_2, \cdots, l_n$ 都不为零.因为 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关,所以存在不全为零的数 $k_1, k_2, \cdots, k_n$,使得

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n = \mathbf{0},$$

于是

$$(k_1 l_1^{-1}) l_1\boldsymbol{v}_1 + (k_2 l_2^{-1}) l_2\boldsymbol{v}_2 + \cdots + (k_n l_n^{-1}) l_n\boldsymbol{v}_n = \mathbf{0}.$$

因为 $k_1 l_1^{-1}, k_2 l_2^{-1}, \cdots, k_n l_n^{-1}$ 不全为零,故 $l_1\boldsymbol{v}_1, l_2\boldsymbol{v}_2, \cdots, l_n\boldsymbol{v}_n$ 线性相关.

(5) 设 $k_1, k_2, \cdots, k_n \in \mathbf{R}$,使得

$$k_1 l_1\boldsymbol{v}_1 + k_2 l_2\boldsymbol{v}_2 + \cdots + k_n l_n\boldsymbol{v}_n = \mathbf{0}.$$

因为 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,所以 $k_1 l_1 = k_2 l_2 = \cdots = k_n l_n = 0$.又因为 $l_1, l_2, \cdots, l_n$ 都不为零,所以 $k_1 = k_2 = \cdots = k_n = 0$,于是,$l_1\boldsymbol{v}_1, l_2\boldsymbol{v}_2, \cdots, l_n\boldsymbol{v}_n$ 线性无关.

由定义 1.6 可以看出,在几何空间 V_3 中,若三个向量线性相关,则它们共面;若三个向量线性无关,则它们不共面.若两个向量线性相关,则它们共线;若两个向量线性无关,则它们就不共线.由此可见线性相关是几何空间中向量共

线、共面概念在一般向量空间的推广.

下面,从另一个角度揭示线性相关的向量组与线性无关的向量组的本质区别.

命题 1.15　实数域 **R** 上的向量空间 V 中一组向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n (n \geq 2)$ 线性相关的充分必要条件是其中至少有一个向量可由其它向量线性表示.

证　必要性.设向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关,则根据定义 1.6,有不全为零的数 $k_1, k_2, \cdots, k_n$,使得

$$k_1 \boldsymbol{v}_1 + k_2 \boldsymbol{v}_2 + \cdots + k_n \boldsymbol{v}_n = \boldsymbol{0} .$$

设 $k_i \neq 0$,则得

$$\boldsymbol{v}_i = - k_i^{-1} k_1 \boldsymbol{v}_1 - \cdots - k_i^{-1} k_{i-1} \boldsymbol{v}_{i-1} - k_i^{-1} k_{i+1} \boldsymbol{v}_{i+1} - \cdots - k_i^{-1} k_n \boldsymbol{v}_n .$$

即 $\boldsymbol{v}_i$ 可由其余向量线性表示.

充分性.设向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 中有一个向量 $\boldsymbol{v}_j$ 可由其余向量线性表示,即

$$\boldsymbol{v}_j = l_1 \boldsymbol{v}_1 + \cdots + l_{j-1} \boldsymbol{v}_{j-1} + l_{j+1} \boldsymbol{v}_{j+1} + \cdots + l_n \boldsymbol{v}_n .$$

则

$$l_1 \boldsymbol{v}_1 + \cdots + l_{j-1} \boldsymbol{v}_{j-1} - \boldsymbol{v}_j + l_{j+1} \boldsymbol{v}_{j+1} + \cdots + l_n \boldsymbol{v}_n = \boldsymbol{0} .$$

向量 $\boldsymbol{v}_j$ 的系数为 -1 不为 0,因此向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关.□

命题 1.16　一个向量 $\boldsymbol{u}$ 线性相关当且仅当 $\boldsymbol{u} = \boldsymbol{0}$.

证　设 $\boldsymbol{u} = \boldsymbol{0}$,则 $1 \boldsymbol{u} = \boldsymbol{0}$,因而 $\boldsymbol{u} = 0$ 是线性相关的.反之,若 $\boldsymbol{u}$ 线性相关,则存在非零的数 l 使得 $l\boldsymbol{u} = \boldsymbol{0}$,因而 $l^{-1} l\boldsymbol{u} = l^{-1}\boldsymbol{0} = \boldsymbol{0}$,即 $1 \boldsymbol{u} = \boldsymbol{u} = \boldsymbol{0}$,所以 $\boldsymbol{u} = \boldsymbol{0}$.□

下面从子空间角度讨论向量组的线性相关性.设 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n (n \geq 2)$ 是实数域 **R** 上的向量空间 V 的非零向量.若存在 $i \in \{1, 2, \cdots, n\}$ 使得

$$[\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n] = [\boldsymbol{v}_1] + \cdots + [\boldsymbol{v}_{i-1}] + [\boldsymbol{v}_{i+1}] + \cdots + [\boldsymbol{v}_n],$$

则

$$\boldsymbol{v}_i \in [\boldsymbol{v}_i] \quad [\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n] = [\boldsymbol{v}_1] + \cdots + [\boldsymbol{v}_{i-1}] + [\boldsymbol{v}_{i+1}] + \cdots + [\boldsymbol{v}_n],$$

即

$$\boldsymbol{v}_i = l_1 \boldsymbol{v}_1 + \cdots + l_{i-1} \boldsymbol{v}_{i-1} + l_{i+1} \boldsymbol{v}_{i+1} + \cdots + l_n \boldsymbol{v}_n .$$

这说明 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性相关.因为

$$(-1) \boldsymbol{v}_i + l_1 \boldsymbol{v}_1 + \cdots + l_{i-1} \boldsymbol{v}_{i-1} + l_{i+1} \boldsymbol{v}_{i+1} + \cdots + l_n \boldsymbol{v}_n = \boldsymbol{0},$$

且 $\boldsymbol{v}_i \neq \boldsymbol{0}$,所以 $[\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n]$ 不是直和.假设 $[\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n]$ 是直和.若

$$\boldsymbol{0} = k_1 \boldsymbol{v}_1 + k_2 \boldsymbol{v}_2 + \cdots + k_n \boldsymbol{v}_n,$$

则有 $k_1 \boldsymbol{v}_1 = k_2 \boldsymbol{v}_2 = \cdots = k_n \boldsymbol{v}_n = \boldsymbol{0}$.由于 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n (n \geq 2)$ 都是非零向量,所以

$k_1 = k_2 = \cdots = k_n = 0$ 这说明 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关.同理,若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,则$[\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n]$是直和.

由此有下面命题 1.17.

命题 1.17　设 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n (n \geq 2)$ 是实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量, $\boldsymbol{v}_i \neq \boldsymbol{0} (i = 1, 2, \cdots, n)$.则向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关的充要条件是

$$[\boldsymbol{v}_1] + [\boldsymbol{v}_2] + \cdots + [\boldsymbol{v}_n] = [\boldsymbol{v}_1] \oplus [\boldsymbol{v}_2] \oplus \cdots \oplus [\boldsymbol{v}_n].$$

由命题 1.6 知,在几何空间 V_3 中任意向量 $\boldsymbol{a}$ 都可由三个不共面的非零向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 线性表示,即 $\boldsymbol{a} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$ 并且 x, y, z 是由 $\boldsymbol{a}$ 唯一确定的.下面这个命题就是上述性质在一般向量空间的推广.

命题 1.18　如果 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,而 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n, \boldsymbol{u}$ 线性相关,则 $\boldsymbol{u}$ 可由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性表示,而且表示式是唯一的.

证　因为 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n, \boldsymbol{u}$ 线性相关,有不全为 0 的数 $k, k_1, k_2, \cdots, k_n$,使

$$k\boldsymbol{u} + k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n = \boldsymbol{0}.$$

若 $k = 0$,则

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n = \boldsymbol{0}.$$

因 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,故 $k_1, k_2, \cdots, k_n$ 全为 0.这与数 $k, k_1, k_2, \cdots, k_n$ 不全为 0 矛盾.故必有 $k \neq 0$ 及

$$\boldsymbol{u} = -k^{-1}k_1\boldsymbol{v}_1 - k^{-1}k_2\boldsymbol{v}_2 - \cdots - k^{-1}k_n\boldsymbol{v}_n,$$

即 $\boldsymbol{u}$ 可由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性表示.

下面证表示式唯一.设 $\boldsymbol{u}$ 有两个表示式如下:

$$\boldsymbol{u} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n,\ \boldsymbol{u} = b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + \cdots + b_n\boldsymbol{v}_n.$$

两式相减得

$$(a_1 - b_1)\boldsymbol{v}_1 + (a_2 - b_2)\boldsymbol{v}_2 + \cdots + (a_n - b_n)\boldsymbol{v}_n = \boldsymbol{u} - \boldsymbol{u} = \boldsymbol{0}.$$

因为 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关,故必有 $a_1 - b_1 = a_2 - b_2 = \cdots = a_n - b_n = 0$,即 $a_1 = b_1, a_2 = b_2, \cdots, a_n = b_n$.可见 $\boldsymbol{u}$ 用 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性表示的表示式是唯一的.□

例 11　证明: n 元行向量空间 $\mathbf{R}^n$ 中 n 个向量

$$\boldsymbol{e}_1 = (1, 0, 0, \cdots, 0), \boldsymbol{e}_2 = (0, 1, 0, \cdots, 0), \cdots, \boldsymbol{e}_n = (0, 0, 0, \cdots, 1)$$

线性无关,并且 $\mathbf{R}^n$ 中任意向量都可由向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性表示.

证　设 $k_1, k_2, \cdots, k_n \in \mathbf{R}$,且 $k_1\boldsymbol{e}_1 + k_2\boldsymbol{e}_2 + \cdots + k_n\boldsymbol{e}_n = \boldsymbol{0}$,

$k_1(1, 0, 0, \cdots, 0) + k_2(0, 1, 0, \cdots, 0) + \cdots + k_n(0, 0, 0, \cdots, 1) = (k_1, k_2, \cdots, k_n) = \boldsymbol{0}$,

则 $k_1 = k_2 = \cdots = k_n = 0$.因此 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性无关.另外,任给 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n) \in \mathbf{R}^n$.显然有

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n,$$

即 $\mathbf{R}^n$ 中任意向量都可由向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性表示.由命题 1.18 可知 $\boldsymbol{\alpha}$ 的表示式是唯一的.

3.4 基与维数

有了向量组线性相关和线性无关的概念后,就可以研究实数域 $\mathbf{R}$ 上向量空间及其子空间的结构了.

在几何空间 V_3 中,取定三个不共面的非零向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$,则其中任一向量 $\boldsymbol{u}$ 都可以由 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 线性表示,即 $\boldsymbol{u} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$,并且表示式唯一.此时,$V_3 = [\boldsymbol{e}_1] \oplus [\boldsymbol{e}_2] \oplus [\boldsymbol{e}_3]$,在 V_3 中线性无关的向量个数最多为 3.这样,在几何空间 V_3 中,可以建立仿射坐标系 $\{O; \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$,其中 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ 称为 V_3 的基,(x, y, z) 为 $\boldsymbol{u}$ 在该仿射坐标系中的坐标.如果注意到,在几何空间 V_3 中,三个不共面的向量是线性无关的,那么就容易从几何直观受到启发,把“基”的概念推广到一般的向量空间,用它来描述空间的结构.为此需解决这样一些问题:对于给定的向量空间什么样向量可做“基”,这样的“基”有什么特点?对同一向量空间而言,基是否唯一?如果不唯一,不同的基所含向量个数是否一样多?在证明下面的命题 1.19 之前,先看一个例子.

例 12 若向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 中每个向量均为向量 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 的线性组合,则 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 线性相关.

证 因为 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 中每个向量均为向量 $\boldsymbol{v}_1, \boldsymbol{v}_2$ 的线性组合,可设

$$\boldsymbol{u}_1 = a_{11}\boldsymbol{v}_1 + a_{12}\boldsymbol{v}_2,$$
$$\boldsymbol{u}_2 = a_{21}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2,$$
$$\boldsymbol{u}_3 = a_{31}\boldsymbol{v}_1 + a_{32}\boldsymbol{v}_2.$$

若 $a_{11} = a_{12} = 0$,则 $\boldsymbol{u}_1 = \mathbf{0}$,那么 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 线性相关.反之,可以设 $a_{11} \neq 0$,于是

$$\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1 = (a_{22} - a_{11}^{-1}a_{21}a_{12})\boldsymbol{v}_2,$$
$$\boldsymbol{u}_3 - a_{11}^{-1}a_{31}\boldsymbol{u}_1 = (a_{32} - a_{11}^{-1}a_{31}a_{12})\boldsymbol{v}_2.$$

如果 $a_{22} - a_{11}^{-1}a_{21}a_{12} = 0$,则 $\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1 + 0\boldsymbol{u}_3 = \mathbf{0}$,因为 $\boldsymbol{u}_2$ 的系数为 1,所以 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 线性相关.若 $a_{22} - a_{11}^{-1}a_{21}a_{12} \neq 0$,可将上式简写为

$$\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1 = b_{22}\boldsymbol{v}_2,$$
$$\boldsymbol{u}_3 - a_{11}^{-1}a_{31}\boldsymbol{u}_1 = b_{32}\boldsymbol{v}_2.$$

其中 $b_{22} = a_{22} - a_{11}^{-1}a_{21}a_{12}$, $b_{32} = a_{32} - a_{11}^{-1}a_{31}a_{12}$.于是

$$\begin{aligned}\mathbf{0} &= b_{32}\boldsymbol{v}_2 - b_{32}b_{22}^{-1}b_{22}\boldsymbol{v}_2 \\ &= \boldsymbol{u}_3 - a_{11}^{-1}a_{31}\boldsymbol{u}_1 - b_{32}b_{22}^{-1}(\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1) \\ &= \boldsymbol{u}_3 - b_{22}^{-1}b_{32}\boldsymbol{u}_2 - (a_{11}^{-1}a_{31} - b_{22}^{-1}b_{32}a_{11}^{-1}a_{21})\boldsymbol{u}_1.\end{aligned}$$

因为 $\boldsymbol{u}_3$ 的系数为 1,所以 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ 线性相关 .

将例 12 推广,就会得到下面命题 .

命题 1.19　若向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 中每个向量都是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的线性组合,而 $n < m$,则 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 必线性相关 .

*证　由假设可知

$$\begin{aligned} \boldsymbol{u}_1 &= a_{11}\boldsymbol{v}_1 + a_{12}\boldsymbol{v}_2 + \cdots + a_{1n}\boldsymbol{v}_n, \\ \boldsymbol{u}_2 &= a_{21}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2 + \cdots + a_{2n}\boldsymbol{v}_n, \\ &\cdots \quad \cdots \qquad \cdots \qquad \cdots \\ \boldsymbol{u}_m &= a_{m1}\boldsymbol{v}_1 + a_{m2}\boldsymbol{v}_2 + \cdots + a_{mn}\boldsymbol{v}_n . \end{aligned}$$

考察 $a_{11}, a_{12}, \cdots, a_{1n}$,若它们都为零,则 $\boldsymbol{u}_1$ 为 $\boldsymbol{0}$,因而 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 线性相关,命题得证 .如果 $a_{11}, a_{12}, \cdots, a_{1n}$ 不都为零,通过调整 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的编号可使 $a_{11} \neq 0$.并且有

$$\begin{aligned} &\boldsymbol{u}_1 = a_{11}\boldsymbol{v}_1 + a_{12}\boldsymbol{v}_2 + \cdots + a_{1n}\boldsymbol{v}_n, \\ &\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1 = 0\boldsymbol{v}_1 + (a_{22} - a_{11}^{-1}a_{21}a_{12})\boldsymbol{v}_2 + \cdots + (a_{2n} - a_{11}^{-1}a_{21}a_{1n})\boldsymbol{v}_n, \\ &\boldsymbol{u}_3 - a_{11}^{-1}a_{31}\boldsymbol{u}_1 = 0\boldsymbol{v}_1 + (a_{32} - a_{11}^{-1}a_{31}a_{12})\boldsymbol{v}_2 + \cdots + (a_{3n} - a_{11}^{-1}a_{31}a_{1n})\boldsymbol{v}_n, \\ &\quad \cdots\cdots\cdots\cdots \\ &\boldsymbol{u}_m - a_{11}^{-1}a_{m1}\boldsymbol{u}_1 = 0\boldsymbol{v}_1 + (a_{m2} - a_{11}^{-1}a_{m1}a_{12})\boldsymbol{v}_2 + \cdots + (a_{mn} - a_{11}^{-1}a_{m1}a_{1n})\boldsymbol{v}_n . \end{aligned}$$

对上面第二个等式,考察 $a_{22} - a_{11}^{-1}a_{21}a_{12}, \cdots, a_{2n} - a_{11}^{-1}a_{21}a_{1n}$.若它们都为 0,则 $\boldsymbol{u}_2 - a_{11}^{-1}a_{21}\boldsymbol{u}_1 = \boldsymbol{0}$.因为 $\boldsymbol{u}_2$ 的系数为 1,所以 $\boldsymbol{u}_1, \boldsymbol{u}_2$ 线性相关,由例 10(2) 可知,$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 线性相关 .设 $a_{22} - a_{11}^{-1}a_{21}a_{12}, \cdots, a_{2n} - a_{11}^{-1}a_{21}a_{1n}$ 不都为 0 .对第 2～第 m 个等式再重复上面的过程 .第 r 次重复上面的过程($r \leq n$),已使第 $r+1$～第 m 个等式的右边向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_r$ 的系数都化为 0 .这样,最多重复 n 次,可使第 $n+1$～第 m 个等式的右边向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的系数都为 0 .不妨假设在第 r 次后($r \leq n$),使得第 $r+1$ 个等式右边向量 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 的系数都为 0 .这样必有

$$\boldsymbol{u}_r + k_1\boldsymbol{u}_{r-1} + \cdots + k_{r-1}\boldsymbol{u}_1 = \boldsymbol{0},$$

其中 $k_1, \cdots, k_{r-1}$ 是由上述计算过程所确定的一些实数 .因为 $\boldsymbol{u}_r$ 的系数为 1,所以 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ 线性相关,从而 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 线性相关 .□

证明命题 1.19 时,由于处理一般情况记号繁多,理解上可能有一定困难 .但有了例 1.12 作为先导,可能就容易理解命题 1.19 的证明了 .重视具体例子有助于理解一般定理,探讨具体例子也有助于发现新的数学定理 .

定义 1.7　设 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 为实数域 **R** 上的向量空间 V 的一个向量组,且该向量组中的 r 个向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_r$ 具有如下性质:

(1) $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_r$ 线性无关;

(2) $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 可由 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_r$ 线性表示.

则称向量组 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_r$ 为向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 的一个极大线性无关组.称 r 为向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 的秩.

读者可能问:对于给定的向量组,它的秩是唯一确定的吗?下面的定理将回答这个问题.

定理 1.3 若 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_l$ 与 $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_h$ 都是向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 的极大线性无关组,则 $l = h$.

证 用反证法.设 $l < h$.因 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_l$ 是 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 的极大线性无关组,所以任意 $\boldsymbol{u}_i$ 可由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_l$ 线性表示.特别地, $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_h$ 也可由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_l$ 线性表示.由命题 1.19 知 $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_h$ 必线性相关,这与 $\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_h$ 是向量组 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m$ 的极大线性无关组矛盾,因此 $h \leq l$.同理可证 $l \leq h$.故 $l = h$.□

定义 1.8 设 V 为实数域 $\mathbf{R}$ 上向量空间.V 中两个向量组

$A: \alpha_1, \alpha_2, \cdots, \alpha_r;$

$B: \beta_1, \beta_2, \cdots, \beta_s.$

如果向量组 A 中每个向量都能由向量组 B 中的向量线性表示,则称向量组 A 能由向量组 B 线性表示.如果向量组 A 与向量组 B 能相互表示,则称向量组 A 与向量组 B 等价.

由定义 1.7 和 1.8 可知,一个向量组与其极大线性无关组等价.

命题 1.20 如果向量组 A 能由向量组 B 线性表示,则向量组 A 的秩不大于向量组 B 的秩;如果向量组 A 与向量组 B 等价,则它们的秩相等.

证:设 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ 及 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_s$ 分别为向量组 A 及 B 中的极大线性无关组.因为向量组 A 能由向量组 B 线性表示,所以向量组 A 也能由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_s$ 线性表示.这样, $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ 也能由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_s$ 线性表示.若 $r > s$,由命题 1.19 可知 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r$ 线性相关,这与它们是向量组 A 的极大线性无关组矛盾.于是向量组 A 的秩 r 不大于向量组 B 的秩 s.如果向量组 A 与向量组 B 等价,由上面的证明可知 $r \leq s$ 且 $s \leq r$,因而 $r = s$.□

下面的定义是几何空间 V_3 中坐标向量概念在一般向量空间的推广.

定义 1.9 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量.如果 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性无关,而 V 中每个向量均可由它们线性表示,则说 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 构成 V 的一个基.

在几何空间 V_3 中,不共线的两个向量是线性无关的,但与这两个向量不共面的任意一个向量都不能用它们线性表示,因此它们不是 V_3 的基.不共面的三个向量是线性无关的, V_3 中所有向量都可用它们线性表示.因此, V_3 中任意三

个不共面的向量可作为 V_3 的一个基.而 V_3 中任意四个向量一定线性相关,因此 V_3 的基不可能含四个向量.总之,V_3 有无穷多个基,且每一个基所含向量个数都是 3.

命题 1.21　如果向量空间 V 有一个基是由 n 个向量构成的,则 V 中任意 n 个线性无关的向量均构成 V 的一个基,而且 V 的任意一个基必恰由 n 个线性无关的向量构成.

证　设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 V 的一个基,$\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 是 V 中 n 个线性无关的向量.对于任意 $\boldsymbol{u}\in V$,由于 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 V 的一个基,所以向量组 $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 可由向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 表示.这样 $n+1$ 个向量被 n 个向量表示,由命题 1.19 知,向量组 $\boldsymbol{u},\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 线性相关.再由命题 1.18 知,$\boldsymbol{u}$ 可由 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 线性表示.由定义 1.9 知 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n$ 也为 V 的一个基.设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 也是 V 的一个基,则 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 与 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 等价.由命题 1.19 可知,若 $r<n$,则 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 线性相关.矛盾,故 $r\geq n$.类似地,由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 线性无关知 $n\geq r$,因此 $r=n$.所以 V 的任意一个基必恰由 n 个线性无关的向量构成.
□

由命题 1.21 有如下定义

定义 1.10　实数域 $\mathbf{R}$ 上的向量空间 V 的基所含的向量个数称为 V 的维数,记作 $\dim(V)$.

如果向量空间 V 的基由 n 个向量组成,则说 V 是 n 维向量空间,只含零向量的向量空间叫做 0 维向量空间,n 维与 0 维的向量空间统称为有限维向量空间.如果 V 不存在有限个向量的基,则称 V 是无限维向量空间.

由定义 1.10 可知,几何空间 V_1,V_2,V_3 的维数分别为 1,2,3.在几何空间 V_3 中建立仿射坐标系之后,一个向量与三元有序数组——坐标一一对应.为了简明,本书把已建立了仿射坐标系的几何空间 V_3 称为几何空间 $\mathbf{R}^3$,而且有时把其坐标本身就称为向量.

命题 1.22　有限维向量空间中任意一组线性无关的向量均可扩充成一个基.

证　设 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 是有限维向量空间 V 中一组线性无关的向量组.若 $V\setminus[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]=$,则 $V=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]$,$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 是 V 的一个基.若 $V\setminus[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]\neq$,令 $\boldsymbol{u}\in V\setminus[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]$.由于 $\mathbf{0}\in[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]$,所以 $\boldsymbol{u}\neq\mathbf{0}$.若 $\boldsymbol{u},\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 线性相关,则 $\boldsymbol{u}$ 可由 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 线性表示,于是 $\boldsymbol{u}\in[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]$.这与 $\boldsymbol{u}\in V\setminus[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]$矛盾,因此 $\boldsymbol{u},\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 线性无关.重复上面的过程直到 $V\setminus[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r]=$,这样就可得到 V 的一个基.□

由命题 1.18 知,对于实数域 $\mathbf{R}$ 上任一个 n 维向量空间 V,当选定一个基

$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 后, V 中任一向量 $\boldsymbol{u}$ 都可唯一地表示为:

$$\boldsymbol{u} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n, \ a_1, a_2, \cdots, a_n \in \mathbf{R}.$$

称有序数组 $a_1, a_2, \cdots, a_n$ 为 $\boldsymbol{u}$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 下的坐标, 记为 $(a_1, a_2, \cdots, a_n)$. 这是三维几何空间中向量的坐标概念在 n 维向量空间中的推广.

在 n 维向量空间 V 中选定一组基之后, V 中任一向量就和它的坐标——有序数组 $(a_1, a_2, \cdots, a_n)$ 建立了一一对应的关系. 借此可给出向量空间 V 与 n 元行(列)向量空间 $\mathbf{R}^n$ 间一个一一对应的关系.

定义 1.11　对于实数域 $\mathbf{R}$ 上的两个向量空间 V_1, V_2, 如果存在 V_1 到 V_2 上的一一对应 f, 对于所有的向量 $\boldsymbol{a}, \boldsymbol{b} \in V_1$ 及任意 $\lambda \in \mathbf{R}$, 使得

(1) $f(\boldsymbol{a} + \boldsymbol{b}) = f(\boldsymbol{a}) + f(\boldsymbol{b})$,

(2) $f(\lambda\boldsymbol{a}) = \lambda f(\boldsymbol{a})$

都成立, 则一一对应 f 称为向量空间 V_1 到向量空间 V_2 上的同构映射, 且称向量空间 V_1 和 V_2 是同构的.

设 V 为实数域 $\mathbf{R}$ 上的 n 维向量空间, $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 V 的一个基. V 中任一向量

$$\boldsymbol{u} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n$$

与其坐标 $(a_1, a_2, \cdots, a_n)$(n 元数组) 恰好建立了一个一一对应, 即有映射

$$f: V \to \mathbf{R}^n, \ f(\boldsymbol{u}) = (a_1, a_2, \cdots, a_n) \in \mathbf{R}^n.$$

显然, 对于任意 $\boldsymbol{u}, \boldsymbol{v} \in V$, 任意 $k \in \mathbf{R}$, f 为一一对应且 $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$, $f(k\boldsymbol{u}) = kf(\boldsymbol{u})$ 这表明, 研究实数域 $\mathbf{R}$ 上的 n 维向量空间 V, 可以通过它的基及向量的坐标表示, 转化为研究 n 元行(列)向量空间 $\mathbf{R}^n$. 从这个角度来看, 尽管 n 维向量空间的定义比较抽象, 然而有了 V 到 $\mathbf{R}^n$ 的同构映射, 实际上它可具体地用 n 元行(列)向量空间 $\mathbf{R}^n$ 来理解. 在向量空间的定义中, 仅仅规定了空间的向量加法及数乘即线性运算应满足的性质, 对这两种运算如何具体实现及空间的向量具体含义都不作规定. 两个向量空间的同构映射是保持两个空间的线性运算一致的一一对应. 既然两个向量空间线性运算是一致的, 同构的向量空间实质上是相同的, 没有本质上的差别.

这里要特别指出, 把实数域 $\mathbf{R}$ 换成一般域 K(参考附录 A.1), 对于前面所讨论的关于实数域 $\mathbf{R}$ 上向量空间的所有结果(定义, 命题与定理) 也成立. 有限域上的向量空间在编码、计算机等工程技术领域有非常重要的应用. 事实上, 一般域 K 上的 n 维向量空间与 K^n 同构.

例 13　证明: 在 n 元行向量空间 $\mathbf{R}^n$ 中, n 个向量

$$\boldsymbol{e}_1 = (1, 0, 0, \cdots, 0), \ \boldsymbol{e}_2 = (0, 1, 0, \cdots, 0), \cdots, \boldsymbol{e}_n = (0, 0, 0, \cdots, 1)$$

是 $\mathbf{R}^n$ 的一个基, 并求 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n) \in \mathbf{R}^n$ 在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 下的坐标.

证　由例 11 及定义 1.9 知, $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是 $\mathbf{R}^n$ 的一组基 .因为

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n,$$

所以向量 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$在基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 下的坐标为$(a_1, a_2, \cdots, a_n)$.

例 14　证明:在 n 元行向量空间 $\mathbf{R}^n$ 中, n 个向量

$$\boldsymbol{u}_1 = (1,1,1,\cdots,1), \boldsymbol{u}_2 = (0,1,1,\cdots,1), \cdots, \boldsymbol{u}_n = (0,0,0,\cdots,1)$$

是 $\mathbf{R}^n$ 的一个基,并求 $\boldsymbol{b} = (b_1, b_2, \cdots, b_n) \in \mathbf{R}^n$ 在基 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 下的坐标 .

证　任给 $\boldsymbol{b} = (b_1, b_2, \cdots, b_n) \in \mathbf{R}^n$, 有

$b_1\boldsymbol{u}_1 + (b_2 - b_1)\boldsymbol{u}_2 + \cdots + (b_n - b_{n-1})\boldsymbol{u}_n = (b_1, b_1 + b_2 - b_1, \cdots, b_1 + b_2 - b_1 + \cdots + b_n - b_{n-1}) = \boldsymbol{b}$.

这说明 $\mathbf{R}^n$ 中任意向量可由 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 线性表示, 于是 $\mathbf{R}^n$ 的基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ (例 13)可被 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 线性表示 .若 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 线性相关,则其极大线性无关组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_r$ 的向量个数 $r < n$.由于 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 可被其极大线性无关组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_r$ 线性表示,所以 $\mathbf{R}^n$ 的基 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 可被 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_r$ 线性表示 .由命题 1.19 知 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性相关, 这与 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 线性无关矛盾 .于是 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 线性无关,即这组向量是 $\mathbf{R}^n$ 的基 .由 $\boldsymbol{b}$ 的线性表示可知,向量 $\boldsymbol{b}$ 在基 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 下的坐标为$(b_1, b_2 - b_1, \cdots, b_n - b_{n-1})$.

由例 13, 14 可看出, $\mathbf{R}^n$ 的向量 $\boldsymbol{b}$ 在不同的基下的坐标不同 .不同的基下向量的坐标之间的关系将在以后予以研究 .

第四节　欧 氏 空 间

4.1　内积　长度

在一般的向量空间中,没有像几何空间中向量的长度及向量的夹角等度量概念,但是在很多问题中需要引入这样的度量,因此有必要将这种度量推广到一般向量空间 .由第二节可知, 几何空间中两个向量 $\boldsymbol{a}, \boldsymbol{b}$ 的数量积可用两个向量长度及夹角表示为

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta .$$

如果在几何空间中建立了直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$,则向量 $\boldsymbol{a}, \boldsymbol{b}$ 的数量积可用两个向量的坐标表示为

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1b_1 + a_2b_2 + a_3b_3,$$

其中向量 $\boldsymbol{a}, \boldsymbol{b}$ 的坐标分别为$(a_1, a_2, a_3), (b_1, b_2, b_3)$.向量 $\boldsymbol{a}$ 的长度

$$|\boldsymbol{a}| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} = \sqrt{a_1^2 + a_2^2 + a_3^2} .$$

向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角 θ 的余弦为

$$\cos\theta=\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|},\quad 0\leq\theta\leq\pi.$$

这表明,对几何空间中的向量而言,也可用向量的数量积表示向量长度及两个向量的夹角等几何度量.

在一般向量空间中,不能像几何空间那样,用直观的几何意义来规定这些度量概念.可以先把几何空间中向量的数量积的运算满足的基本性质抽象出来,然后用数量积来定义向量长度和两个向量夹角等度量,这些基本性质对几何空间中向量是熟知的:对几何空间中的任何向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 及 $k\in\mathbf{R}$,

(1) $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$;

(2) $(\boldsymbol{a}+\boldsymbol{c})\cdot\boldsymbol{b}=\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{c}\cdot\boldsymbol{b}$;

(3) $(k\boldsymbol{a})\cdot\boldsymbol{b}=k(\boldsymbol{a}\cdot\boldsymbol{b})$;

(4) $\boldsymbol{a}\cdot\boldsymbol{a}\geq 0$;当且仅当 $\boldsymbol{a}=\mathbf{0}$ 时,$\boldsymbol{a}\cdot\boldsymbol{a}=0$.

如同定义一般向量空间中的向量加法及数乘两种运算一样,下面将以几何空间中的向量的数量积满足的性质为背景,在实数域上的一般向量空间中定义向量内积,再通过向量内积给出向量的长度及向量夹角的定义;如同在几何空间中建立直角坐标系一样建立一般实数域上向量空间的标准正交基,进而研究该向量空间的"几何"问题.

定义 1.12　设 V 为实数域 $\mathbf{R}$ 上的向量空间,对于 V 中任意两个向量 $\boldsymbol{a},\boldsymbol{b}$,如果有一实数 $(\boldsymbol{a},\boldsymbol{b})$ 与它们对应,并且对 V 中任意向量 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 及实数 $k\in\mathbf{R}$ 满足下列条件:

(1) $(\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{b},\boldsymbol{a})$;

(2) $(k\boldsymbol{a},\boldsymbol{b})=k(\boldsymbol{a},\boldsymbol{b})$;

(3) $(\boldsymbol{a}+\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{a},\boldsymbol{c})+(\boldsymbol{b},\boldsymbol{c})$;

(4) $(\boldsymbol{a},\boldsymbol{a})\geq 0$;$(\boldsymbol{a},\boldsymbol{a})=0$ 当且仅当 $\boldsymbol{a}=\mathbf{0}$.

则实数 $(\boldsymbol{a},\boldsymbol{b})$ 叫做向量 $\boldsymbol{a},\boldsymbol{b}$ 的内积.定义了内积的实数域上向量空间 V 称为(实)内积空间或欧氏空间.

例 1　在几何空间 $\mathbf{R}^3$ 中建立了直角坐标系 $\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$,证明:对任意两个向量 $\boldsymbol{a},\boldsymbol{b}$,若定义

$$(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\theta=a_1b_1+a_2b_2+a_3b_3,$$

其中向量 $\boldsymbol{a},\boldsymbol{b}$ 的坐标分别为 (a_1,a_2,a_3),(b_1,b_2,b_3),则 $\boldsymbol{a}\cdot\boldsymbol{b}$ 为内积且几何空间 $\mathbf{R}^3$ 为内积空间.

证　易验证 $\boldsymbol{a}\cdot\boldsymbol{b}$ 满足定义 1.12(1)～(4),因此 $\boldsymbol{a}\cdot\boldsymbol{b}$ 为内积且几何空间 $\mathbf{R}^3$ 为内积空间.为了简明,本书称该内积空间为 3 维欧氏几何空间 $\mathbf{R}^3$,其中坐标系为直角坐标系 $\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$.

例 2　在 n 维行向量空间 $\mathbf{R}^n$ 中,对任意两个向量 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$,$\boldsymbol{b}=$

$(b_1, b_2, \cdots, b_n) \in \mathbf{R}^n$,定义$(\boldsymbol{a}, \boldsymbol{b}) = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.容易验证,$(\boldsymbol{a}, \boldsymbol{b})$满足定义 1.12(1)～(4),因此$(\boldsymbol{a}, \boldsymbol{b})$为向量空间 $\mathbf{R}^n$ 的内积且向量空间 $\mathbf{R}^n$ 为一个欧氏空间.为了简明,本书称该内积空间为 n 维欧氏几何空间 $\mathbf{R}^n$.

定义 1.13　在欧氏空间 V 中,任意向量 $\boldsymbol{a}$ 的长度$|\boldsymbol{a}|$定义为 $(\boldsymbol{a}, \boldsymbol{a})$,两非零向量 $\boldsymbol{a}, \boldsymbol{b}(\boldsymbol{a} \neq \boldsymbol{0}, \boldsymbol{b} \neq \boldsymbol{0})$的夹角 θ的余弦定义为

$$\cos\theta = \frac{(\boldsymbol{a}, \boldsymbol{b})}{|\boldsymbol{a}|\,|\boldsymbol{b}|}, 0 \leq \theta \leq \pi.$$

如果$(\boldsymbol{a}, \boldsymbol{b}) = 0$,则 $\theta = \dfrac{\pi}{2}$,此时称向量 $\boldsymbol{a}, \boldsymbol{b}$ 正交,记为 $\boldsymbol{a} \perp \boldsymbol{b}$.规定零向量与任何向量正交.长度为 1 的向量叫做单位向量.为使定义 1.13 有意义,应证明

$$-1 \leq \frac{(\boldsymbol{a}, \boldsymbol{b})}{|\boldsymbol{a}|\,|\boldsymbol{b}|} \leq 1.$$

这正是以著名数学家 A.L.Cauchy(1789—1857)及 H.A.Schwarz(1843—1921)命名的不等式.

命题 1.23　(Cauchy-Schwarz 不等式)设 $\boldsymbol{a}, \boldsymbol{b}$ 为欧氏空间 V 中任意两个向量,则

$$(\boldsymbol{a}, \boldsymbol{b})^2 \leq (\boldsymbol{a}, \boldsymbol{a})(\boldsymbol{b}, \boldsymbol{b}). \tag{1.4}$$

证　用两种方法证明 Cauchy-Schwarz 不等式.

(1) 任给 $x \in \mathbf{R}$ 及任给 $\boldsymbol{a}, \boldsymbol{b} \in V$,由定义 1.12(4)有

$$(\boldsymbol{a} - x\boldsymbol{b}, \boldsymbol{a} - x\boldsymbol{b}) \geq 0.$$

由定义 1.12(1)、(2)和(3)及上式,任给 $x \in \mathbf{R}$,有

$$(\boldsymbol{b}, \boldsymbol{b}) x^2 - 2(\boldsymbol{a}, \boldsymbol{b}) x + (\boldsymbol{a}, \boldsymbol{a}) \geq 0. \tag{1.5}$$

当$(\boldsymbol{b}, \boldsymbol{b}) \neq 0$ 时,令 $x = \dfrac{(\boldsymbol{a}, \boldsymbol{b})}{(\boldsymbol{b}, \boldsymbol{b})}$,则

$$\frac{(\boldsymbol{a}, \boldsymbol{b})^2}{(\boldsymbol{b}, \boldsymbol{b})} - 2\frac{(\boldsymbol{a}, \boldsymbol{b})^2}{(\boldsymbol{b}, \boldsymbol{b})} + (\boldsymbol{a}, \boldsymbol{a}) \geq 0.$$

再用$(\boldsymbol{b}, \boldsymbol{b}) > 0$ 乘上式两端,就得不等式(1.4).

当$(\boldsymbol{b}, \boldsymbol{b}) = 0$ 时,即 $\boldsymbol{b} = \boldsymbol{0}$,不等式(1.4)自然成立.

(2) 由(1.5)式可知,相应的一元二次方程的判别式 $\Delta \leq 0$,即

$$4(\boldsymbol{a}, \boldsymbol{b})^2 - 4(\boldsymbol{a}, \boldsymbol{a})(\boldsymbol{b}, \boldsymbol{b}) \leq 0.$$

同样可得到不等式(1.4).□

例 3　在几何空间 $\mathbf{R}^3$ 中建立了直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$,$\boldsymbol{a}, \boldsymbol{b}$ 为 $\mathbf{R}^3$ 中任意两个向量,若定义

$$(\boldsymbol{a}, \boldsymbol{b}) = a_1 b_1 - a_2 b_1 - a_1 b_2 + 3 a_2 b_2 + a_3 b_3,$$

其中向量 $\boldsymbol{a}, \boldsymbol{b}$ 的坐标分别为(a_1, a_2, a_3),(b_1, b_2, b_3).证明:$(\boldsymbol{a}, \boldsymbol{b})$为 $\mathbf{R}^3$ 的内积运算,且 $\mathbf{R}^3$ 为一个欧氏空间.

证　任给向量 $\boldsymbol{a}=(a_1,a_2,a_3)$，$\boldsymbol{b}=(b_1,b_2,b_3)$，$\boldsymbol{c}=(c_1,c_2,c_3)\in\mathbf{R}^3$，显然，$a_1b_1-a_2b_1-a_1b_2+3a_2b_2+a_3b_3$ 为一实数．下面逐条验证定义 1.12 成立：

(1) $(\boldsymbol{a},\boldsymbol{b})=a_1b_1-a_2b_1-a_1b_2+3a_2b_2+a_3b_3=b_1a_1-b_2a_1-b_1a_2+3b_2a_2+b_3a_3=(\boldsymbol{b},\boldsymbol{a})$；

(2) $(k\boldsymbol{a},\boldsymbol{b})=ka_1b_1-ka_2b_1-ka_1b_2+3ka_2b_2+ka_3b_3$
$=k(a_1b_1-a_2b_1-a_1b_2+3a_2b_2+a_3b_3)=k(\boldsymbol{a},\boldsymbol{b})$，对于任意 $k\in\mathbf{R}$；

(3) $(\boldsymbol{a}+\boldsymbol{c},\boldsymbol{b})=(a_1+c_1)b_1-(a_2+c_2)b_1-(a_1+c_1)b_2+3(a_2+c_2)b_2+(a_3+c_3)b_3$
$=a_1b_1-a_2b_1-a_1b_2+3a_2b_2+a_3b_3+c_1b_1-c_2b_1-c_1b_2+3c_2b_2+c_3b_3$
$=(\boldsymbol{a},\boldsymbol{b})+(\boldsymbol{c},\boldsymbol{b})$；

(4) $(\boldsymbol{a},\boldsymbol{a})=a_1a_1-a_2a_1-a_1a_2+3a_2a_2+a_3a_3=(a_1-a_2)^2+2a_2^2+a_3^2\geq 0$，$(\boldsymbol{a},\boldsymbol{a})=(a_1-a_2)^2+2a_2^2+a_3^2=0$ 当且仅当 $(a_1-a_2)^2=0$，$2a_2^2=0$，$a_3^2=0$，即 $a_1=a_2=a_3=0$，于是 $\boldsymbol{a}=\mathbf{0}$．

这表明，$(\boldsymbol{a},\boldsymbol{b})$ 满足定义 1.12 的各条，因此运算 $(\boldsymbol{a},\boldsymbol{b})$ 为向量空间 $\mathbf{R}^3$ 的内积且向量空间 $\mathbf{R}^3$ 在该内积运算下为一个欧氏空间．

由例 3 可见，在同一个向量空间上，定义不同的内积就构造出不同的欧氏空间．由于例 3 和例 1 给出的内积运算不同，所以它们分别构成不同的欧氏空间．自然地，人们主要研究那些在代数、几何、分析及应用中具有重要意义的欧氏空间，如例 1 给出的 3 维欧氏几何空间和例 2 给出的 n 维欧氏几何空间．

例 4　设 $C[-1,1]$ 为区间 $[-1,1]$ 上所有连续函数构成的集合．显然，在普通意义的函数加法和数乘下，$C[-1,1]$ 为 $\mathbf{R}$ 上的向量空间．任给 $f,g\in C[-1,1]$，若定义

$$(f,g)=\int_{-1}^{1}f(x)g(x)\mathrm{d}x,$$

验证 (f,g) 为内积运算，且 $C[-1,1]$ 为一个欧氏空间．并求向量 $f_1=1$ 与另外两个向量 $f_2=x$，$f_3=x^2$ 的夹角 α,β 的余弦．

解　易验证 (f,g) 满足定义 1.12 的 (1)—(3)，下面验证 (4)，即

$$(f,f)=\int_{-1}^{1}f^2(x)\mathrm{d}x=0$$

当且仅当对任意 $x\in[-1,1]$，$f(x)=0$．当对任意 $x\in[-1,1]$，$f(x)=0$ 时，显然 $(f,f)=\int_{-1}^{1}f^2(x)\mathrm{d}x=0$．反之，假设 $(f,f)=\int_{-1}^{1}f^2(x)\mathrm{d}x=0$．若存在 $x_0\in(-1,1)$，$f(x_0)\neq 0$．由于 f 是连续函数，存在 δ 且 $0<\delta<\min\{x_0+1,1-$

$x_0\}$,使得对任意 $x\in(x_0-\delta,x_0+\delta)$, $f^2(x)>0$.因此$(f,f)=\int_{-1}^{1}f^2(x)\mathrm{d}x>0$,矛盾.故对任意 $x\in(-1,1)$, $f(x)=0$.同理可证 $f(1)=f(-1)=0$.

由定义 1.13,有

$$\cos\alpha=\frac{(1,x)}{\sqrt{(1,1)}\sqrt{(x,x)}}=\frac{\int_{-1}^{1}x\mathrm{d}x}{\sqrt{\int_{-1}^{1}\mathrm{d}x}\sqrt{\int_{-1}^{1}x^2\mathrm{d}x}}=0,$$

$$\cos\beta=\frac{(1,x^2)}{\sqrt{(1,1)}\sqrt{(x^2,x^2)}}=\frac{\int_{-1}^{1}x^2\mathrm{d}x}{\sqrt{\int_{-1}^{1}\mathrm{d}x}\sqrt{\int_{-1}^{1}x^4\mathrm{d}x}}=\frac{\sqrt{5}}{3}.$$

因为$(1,x)=0$,所以 $\alpha=\frac{\pi}{2}$,即在该欧氏空间内,向量 $f_1=1$ 与向量 $f_2=x$ 正交.

4.2　欧氏空间的标准正交基及正交补空间

1. 标准正交基

在 n 维欧氏空间 V 中,若取定基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$,则对于 V 中任意两向量 $\boldsymbol{a},\boldsymbol{b}$

$$\boldsymbol{a}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2+a_n\boldsymbol{e}_n=\sum_{i=1}^{n}a_i\boldsymbol{e}_i;\ \boldsymbol{b}=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2+\cdots+b_n\boldsymbol{e}_n=\sum_{i=1}^{n}b_i\boldsymbol{e}_i,$$

由定义 1.12,有

$$(\boldsymbol{a},\boldsymbol{b})=\sum_{i=1}^{n}a_i(\boldsymbol{e}_i,\sum_{j=1}^{n}b_j\boldsymbol{e}_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_ib_j(\boldsymbol{e}_i,\boldsymbol{e}_j)=\sum_{1\le i,j\le n}a_ib_j(\boldsymbol{e}_i,\boldsymbol{e}_j).$$

现在考虑,能否找到一个基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$,使得当 $i\neq j$ 时,$(\boldsymbol{e}_i,\boldsymbol{e}_j)=0$;当 $i=j$ 时,$(\boldsymbol{e}_i,\boldsymbol{e}_j)=1$.如果能找到这样一个基,就有

$$(\boldsymbol{a},\boldsymbol{b})=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

这可使内积运算大大简化,并且在此基下向量坐标的内积运算与 n 维欧氏几何空间 $\mathbf{R}^n$ 的内积运算的公式相同.

定义 1.14　假定 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 n 维欧氏空间 V 的一个基,如果它们都是单位向量,并且又两两正交,即

$$(\boldsymbol{e}_i,\boldsymbol{e}_j)=\delta_{ij}=\begin{cases}1 & i=j,\\ 0 & i\neq j.\end{cases}$$

其中 δ_{ij}为 Kronecker 符号(L.Kronecker,1823—1891),那么 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 就叫做 V 的一个标准正交基.

在 3 维欧氏几何空间 $\mathbf{R}^3$ 中,坐标向量

$$\boldsymbol{i}=(1,0,0),\ \boldsymbol{j}=(0,1,0),\ \boldsymbol{k}=(0,0,1),$$

就是一个标准正交基.显然,在 3 维欧氏几何空间 $\mathbf{R}^3$ 中有无数个标准正交基.

下面讨论一般的 n 维欧氏空间是否存在标准正交基底.

定理 1.4　欧氏空间 V 中任意 r 个两两正交的非零向量必线性无关.

证　设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_r$ 为欧氏空间 V 的 r 个两两正交的非零向量.若实数 $a_1,a_2,\cdots,a_r$ 使

$$a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_r\boldsymbol{v}_r=\mathbf{0},$$

则

$$\begin{aligned}0&=(\mathbf{0},\boldsymbol{v}_l)=(a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_r\boldsymbol{v}_r,\boldsymbol{v}_l)\\&=a_1(\boldsymbol{v}_1,\boldsymbol{v}_l)+a_2(\boldsymbol{v}_2,\boldsymbol{v}_l)+\cdots+a_r(\boldsymbol{v}_r,\boldsymbol{v}_l)=a_l(\boldsymbol{v}_l,\boldsymbol{v}_l),\end{aligned}$$

其中 $l=1,2,\cdots,r$.又由于 $\boldsymbol{v}_l\neq\mathbf{0}$,有$(\boldsymbol{v}_l,\boldsymbol{v}_l)\neq0$,故 $a_l=0$, $l=1,2,\cdots,r$.所以 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_r$ 线性无关.□

在 n 维欧氏空间中线性无关的向量个数不能超过 n 个,由定理 1.4 可知,该空间两两正交的非零向量个数也不能超过 n.在 $\mathbf{R}^2$ 与 $\mathbf{R}^3$ 中,两两正交的非零向量个数最多分别为 2 个和 3 个.

任意一个非零向量 $\boldsymbol{u}$,都可以对应一个与之夹角为 0 的单位向量$\dfrac{1}{|\boldsymbol{u}|}\boldsymbol{u}$.因此,由 r 个两两正交的非零向量,可得到 r 个两两正交的单位向量.

下面以一种构造性的方法证明:任意 n 维欧氏空间都存在标准正交基,即 n 个两两正交的单位向量,先考察 3 维欧氏几何空间 $\mathbf{R}^3$ 中的一个例子.

例 5　$\boldsymbol{v}_1=(1,0,0)$,$\boldsymbol{v}_2=(2,2,0)$,$\boldsymbol{v}_3=(3,3,3)$为 3 维欧氏几何空间 $\mathbf{R}^3$ 的一个基,由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 构造出一个标准正交基.

解　$\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 如图 1 所示,它们构成几何空间 $\mathbf{R}^3$ 的一个基.求出由它们产生的一个标准正交基,可分如下三个步骤:

(1) 因为 $\boldsymbol{v}_1\neq\mathbf{0}$,由 $\boldsymbol{v}_1$ 产生单位向量 $\boldsymbol{e}_1$,$\alpha_1=\boldsymbol{v}_1=(1,0,0)$,$\boldsymbol{e}_1=\dfrac{\alpha_1}{|\alpha_1|}=(1,0,0)$.

(2) 在 $\boldsymbol{v}_1$ 与 $\boldsymbol{v}_2$ 生成的平面内,向量$(\boldsymbol{v}_2,\boldsymbol{e}_1)\boldsymbol{e}_1$ 为向量 $\boldsymbol{v}_2$ 在 $\boldsymbol{e}_1$ 上的投影向量,下式中与 α_1 α_2 正交,且 $\boldsymbol{e}_2$ 是与 $\boldsymbol{e}_1$ 正交的单位向量,

$$\alpha_2=\boldsymbol{v}_2-(\boldsymbol{v}_2,\boldsymbol{e}_1)\boldsymbol{e}_1=\boldsymbol{v}_2-\left(\boldsymbol{v}_2,\frac{\alpha_1}{|\alpha_1|}\right)\frac{\alpha_1}{|\alpha_1|}=\boldsymbol{v}_2-\frac{(\boldsymbol{v}_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1=(0,2,0).$$

$$\boldsymbol{e}_2=\frac{\alpha_2}{|\alpha_2|}=(0,1,0).$$

(3) $\boldsymbol{v}_3$ 在 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 上的投影向量分别为$(\boldsymbol{v}_3,\boldsymbol{e}_1)\boldsymbol{e}_1$ 及$(\boldsymbol{v}_3,\boldsymbol{e}_2)\boldsymbol{e}_2$,因而下式 α_3 与 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 正交,且 $\boldsymbol{e}_3$ 为与 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 正交的单位向量,

$$\alpha_3=\boldsymbol{v}_3-(\boldsymbol{v}_3,\boldsymbol{e}_1)\boldsymbol{e}_1-(\boldsymbol{v}_3,\boldsymbol{e}_2)\boldsymbol{e}_2$$

$$= \boldsymbol{v}_3 - \left(\boldsymbol{v}_3, \frac{\alpha_1}{|\alpha_1|}\right)\frac{\alpha_1}{|\alpha_1|} - \left(\boldsymbol{v}_3, \frac{\alpha_2}{|\alpha_2|}\right)\frac{\alpha_2}{|\alpha_2|}$$

$$= \boldsymbol{v}_3 - \frac{(\boldsymbol{v}_3, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 - \frac{(\boldsymbol{v}_3, \alpha_2)}{(\alpha_2, \alpha_2)}\alpha_2 = (0 \quad 0 \quad 3) .$$

$$\boldsymbol{e}_3 = \frac{\alpha_3}{|\alpha_3|} = (0 \quad 0 \quad 1) .$$

于是,$\boldsymbol{e}_1 = (1,0,0)$, $\boldsymbol{e}_2 = (0,1,0)$, $\boldsymbol{e}_3 = (0,0,1)$为 3 维欧氏几何空间 $\mathbf{R}^3$ 的一个标准正交基 .

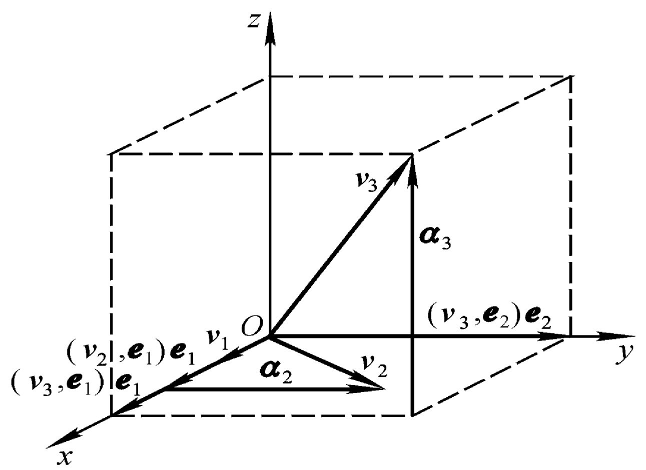


图 1

在上面这个例子中,用几何方法给出了,从 3 维欧氏几何空间 $\mathbf{R}^3$ 中给定的一个基构造一个标准正交基的一个算法 .可以用代数方法将这个具有直观的几何意义的算法推广到 n 维欧氏空间中,以处理更一般的问题 .对于一般 n 维欧氏空间 V,当已知一个基 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 时,如何构造一个标准正交基呢 ? 从例 5 的算法可以得到一个“合理的猜想”: 利用例 5 的公式作为一般公式计算前三个向量 $\alpha_1, \alpha_2, \alpha_3$ 的公式 .当由 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{i-1}$ 得到两两正交的向量 $\alpha_1, \alpha_2, \cdots, \alpha_{i-1}$后,可猜想:

$$\alpha_i = \boldsymbol{v}_i - \frac{(\boldsymbol{v}_i, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 - \frac{(\boldsymbol{v}_i, \alpha_2)}{(\alpha_2, \alpha_2)}\alpha_2 - \cdots - \frac{(\boldsymbol{v}_i, \alpha_{i-1})}{(\alpha_{i-1}, \alpha_{i-1})}\alpha_{i-1}, \qquad (1.6)$$

$i = 1,2,\cdots, n$,显然,α_k 是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ 的线性组合,$k = 1,2,\cdots, n$.于是可设

$$\alpha_1 = b_{11}\boldsymbol{v}_1,$$

$$\alpha_2 = b_{21}\boldsymbol{v}_1 + b_{22}\boldsymbol{v}_2,$$

$$\cdots\cdots\cdots\cdots$$

$$\alpha_k = b_{k1}\boldsymbol{v}_1 + b_{k2}\boldsymbol{v}_2 + \cdots + b_{kk}\boldsymbol{v}_k .$$

其中 $b_{ij} 1 \leq i, j \leq k$ 是实数 .如果 $\alpha_i = \mathbf{0}$,则 $\boldsymbol{v}_i$ 是 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{i-1}$的线性组合,这和 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 线性无关矛盾,所以 $\alpha_i \neq \mathbf{0}$.如果 α_i 与 $\alpha_1, \alpha_2, \cdots, \alpha_{i-1}$彼此正交,那么就得到彼此正交的向量 $\alpha_1, \alpha_2, \cdots, \alpha_{i-1}, \alpha_i$ 这样逐步得到 n 个彼此正

交的非零的向量.由定理1.4知道这 n 个向量线性无关,且为 n 维欧氏空间 V 的一个基.再将它们单位化就得到 V 的一个标准正交基.下面证明 α_j 与 $\alpha_i(1\leq j< i\leq n)$正交.

$$(\alpha_i,\alpha_j)=(\boldsymbol{v}_i,\alpha_j)-\frac{(\boldsymbol{v}_i,\alpha_1)}{(\alpha_1,\alpha_1)}(\alpha_1,\alpha_j)-\cdots-\frac{(\boldsymbol{v}_i,\alpha_{i-1})}{(\alpha_{i-1},\alpha_{i-1})}(\alpha_{i-1},\alpha_j).$$

因为 $\alpha_1,\alpha_2,\cdots,\alpha_{i-1}$彼此正交,注意到 $j< i$,所以上式可化为

$$(\alpha_i,\alpha_j)=(\boldsymbol{v}_i,\alpha_j)-(\boldsymbol{v}_i,\alpha_j)=0.$$

上面的讨论可总结为如下定理:

定理1.5　(Schmidt 正交化)在 n 维欧氏空间 V 中,任取一个基 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$,则可以按照一种确定的方法构造出一个标准正交基.

公式(1.6)给出了 Schmidt 正交化算法.

例6　在3维欧氏几何空间 $\mathbf{R}^3$ 中,已知 $\boldsymbol{v}_1=(1,1,1)$,$\boldsymbol{v}_2=(0,1,1)$,$\boldsymbol{v}_3=(0,0,1)$线性无关,它们是 $\mathbf{R}^3$ 的基.用 Schmidt 正交化将它们化为标准正交基.

解　由 Schmidt 正交化算法(1.6)有:令 $\alpha_1=\boldsymbol{v}_1=(1,1,1)$,

$$\alpha_2=\boldsymbol{v}_2-\frac{(\boldsymbol{v}_2,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1=\frac{1}{3}(-2,1,1),$$

$$\alpha_3=\boldsymbol{v}_3-\frac{(\boldsymbol{v}_3,\alpha_1)}{(\alpha_1,\alpha_1)}\alpha_1-\frac{(\boldsymbol{v}_3,\alpha_2)}{(\alpha_2,\alpha_2)}\alpha_2=\frac{1}{2}(0,-1,1).$$

为了简化计算,取如下与 $\alpha_1,\alpha_2,\alpha_3$ 共线的向量

$$\beta_1=(1,1,1),\ \beta_2=(-2,1,1),\ \beta_3=(0,-1,1),$$

再将它们化为单位向量,即得标准正交基

$$\boldsymbol{e}_1=\ \frac{1}{3},\frac{1}{3},\frac{1}{3}\ ,\ \boldsymbol{e}_2=\ -\frac{2}{6},\frac{1}{6},\frac{1}{6}\ ,\ \boldsymbol{e}_3=\ 0,-\frac{1}{2},\frac{1}{2}\ .$$

2.正交补空间

定义1.15　设 S_1,S_2 为欧氏空间 V 的两个子空间,如果任给 $\boldsymbol{a}\in S_1$,$\boldsymbol{b}\in S_2$ 都有$(\boldsymbol{a},\boldsymbol{b})=0$,则称 S_1 与 S_2 垂直,记作 $S_1\perp S_2$;若 $S_1\perp S_2$ 且 $V=S_1\oplus S_2$,则称 S_1 为 S_2 的正交补空间.

命题1.24　设 S 为 n 维欧氏空间 V 的一个子空间,则

$$S^{\perp}=\{\boldsymbol{a}\mid\boldsymbol{a}\in V,\text{任给 }\boldsymbol{b}\in S,(\boldsymbol{a},\boldsymbol{b})=0\}$$

为 S 的正交补空间,且 $\dim(S)+\dim(S^{\perp})=\dim(V)=n$.

证　任给 $\alpha,\beta\in S^{\perp}$,及任意实数 k,显然有 $\alpha+\beta\in S^{\perp}$ 及 $k\alpha\in S^{\perp}$.故 $S^{\perp}$ 为 V 的子空间.设 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k$ 为 S 的基,可以把它扩充为 V 的一个基 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k,\boldsymbol{v}_{k+1},\cdots,\boldsymbol{v}_n$.再用 Schmidt 正交化将它们化为标准正交基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$.由 Schmidt 正交化过程知,$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k$ 是 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k$ 的线性组合,故$[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k]\subseteq S$.因为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k$ 是两两正交的非零向量,所以它们

线性无关且也为 S 的一个基,即$[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k]=S$.于是对于任意 $\boldsymbol{u}\in S$,有 $\boldsymbol{u}=l_1\boldsymbol{e}_1+l_2\boldsymbol{e}_2+\cdots+l_k\boldsymbol{e}_k$.对于任意 $\boldsymbol{v}\in[\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n]$,有 $\boldsymbol{v}=h_1\boldsymbol{e}_{k+1}+\cdots+h_{n-k}\boldsymbol{e}_n$,因此

$$(\boldsymbol{u},\boldsymbol{v})=(l_1\boldsymbol{e}_1+l_2\boldsymbol{e}_2+\cdots+l_k\boldsymbol{e}_k,h_1\boldsymbol{e}_{k+1}+\cdots+h_{n-k}\boldsymbol{e}_n)=0.$$

所以$[\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n]\quad S^{\perp}$.综上有

$$V=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n]=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_k]+[\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n]\quad S+S^{\perp}\quad V.$$

即 $S+S^{\perp}=V$.若 $\boldsymbol{u}\in S\cap S^{\perp}$,则$(\boldsymbol{u},\boldsymbol{u})=0$,即 $\boldsymbol{u}=\boldsymbol{0}$.所以 $S\cap S^{\perp}=\{\boldsymbol{0}\}$,由定理 1.2 知 $V=S G\ S^{\perp}$.这表明 $S^{\perp}$为 S 的正交补空间,且 $\dim(S)+\dim(S^{\perp})=\dim(V)=n$.□

在 3 维欧氏几何空间 $\mathbf{R}^3$ 中,设所有与直线 l 共线的向量构成的子空间为 S,则 $S^{\perp}$是与直线 l 垂直的所有向量,即与以直线 l 上非零向量为法向量的平面共面的所有向量;若子空间 S 为所有与平面Ⅱ共面的向量,则 $S^{\perp}$为垂直于平面Ⅱ的所有向量.

例 7　设 S 为方程组

$$\begin{aligned}
a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=0,\\
a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=0,\\
&\cdots\cdots\\
a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n&=0.
\end{aligned}$$

的所有解的集合,其中 $\alpha_1=(a_{11},a_{12},\cdots,a_{1n})$,$\alpha_2=(a_{21},a_{22},\cdots,a_{2n})$,$\cdots$,$\alpha_m=(a_{m1},a_{m2},\cdots,a_{mn})\in\mathbf{R}^n$ 为已知向量.证明:

(1) S 为 n 维欧氏几何空间 $\mathbf{R}^n$ 的子空间且 $S=[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}$;

(2) $\dim(S)=n-\dim([\alpha_1,\alpha_2,\cdots,\alpha_m])$.

证　令 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$,即 $\boldsymbol{x}\in\mathbf{R}^n$,如果 $\boldsymbol{x}$ 是方程组的解,即 $\boldsymbol{x}\in S$,则

$$\begin{aligned}
(\alpha_1,\boldsymbol{x})&=0,\\
(\alpha_2,\boldsymbol{x})&=0,\\
&\cdots\cdots\\
(\alpha_m,\boldsymbol{x})&=0.
\end{aligned}\tag{1.7}$$

(1.7)式表明 $\boldsymbol{x}\in[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}$,因而 $S\quad[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}$.反之,如果 $\boldsymbol{x}\in[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}$,则 $\boldsymbol{x}$ 满足(1.7),$\boldsymbol{x}$ 为方程组的解,即 $\boldsymbol{x}\in S$,于是$[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}\quad S$.所以$[\alpha_1,\alpha_2,\cdots,\alpha_m]^{\perp}=S$.再由命题 1.23 知,$S$ 为 $\mathbf{R}^n$ 的子空间,且 $\dim(S)+\dim([\alpha_1,\alpha_2,\cdots,\alpha_m])=n$.其中 $\dim([\alpha_1,\alpha_2,\cdots,\alpha_m])$是向量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 生成的 $\mathbf{R}^n$ 的子空间的维数,也即向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的秩.

向量空间 V 作为一种代数结构,集合 V 的对象是抽象的,V 上的加法与数

乘两种运算是由它们满足的性质定义的.但向量空间的加法、数乘和基都是以几何空间中向量的加法、数乘及空间仿射坐标系为直觉图象发展起来的,使抽象有限维线性空间得到与几何空间仿射坐标系几乎同样的结构.n维欧氏空间可以说是以几何空间的直角坐标系为直觉图象发展起来的,甚至可以认为n维欧氏空间是像3维欧氏几何空间一样具体的几何对象,可以在其中谈论超平面以及展开多元微积分学.当熟知了该空间的性质之后,甚至进而可以谈论无穷维欧氏空间.物理学家汤川秀树(1907—1981,日本人,1949年获诺贝尔物理学奖)的一段话概括了科学中这一普遍的思维方式:“抽象不能单独起作用.在几何富有成果的科学思维中,直觉和抽象是交互为用的.不但某种本质性东西必须从我们丰富的而多少有点模糊的直觉图形中抽象出来,而且同样真实的是,作为人类抽象能力的成果而建立起来的某一个概念也常常在时间的进程中变成我们直觉图象的一部分.从这种新建立起来的直觉,人们可以继续作出进一步的抽象”(摘自汤川秀树.科学需要创造性思维.见:黎光耀.智慧的星光.北京:经济日报出版社,2000)

习　题　一

习题 A

1. 设 A, B, C 为集合.试证:

(1) $A\cup B = B\cup A$;

(2) $A\cap B = B\cap A$;

(3) $A\cup \quad = A$, $A\cap \quad = \quad$;

(4) $(A\cup B)\cup C = A\cup (B\cup C)$;

(5) $(A\cap B)\cap C = A\cap (B\cap C)$;

(6) $(A\cup B)\cap C = (A\cap C)\cup (B\cap C)$;

(7) $(A\cap B)\cup C = (A\cup C)\cap (B\cup C)$;

2. 设 $A \quad B$,试证: $A\cap B = A$, $A\cup B = B$.

3. 在空间直角坐标系(右手系)中,指出下列各点在哪个卦限?

$$A(1,-2,3);\ B(2,3,-4);\ C(2,-3,-4);\ D(-2,-3,1).$$

4. 在3维欧氏几何空间 $\mathbf{R}^3$ 中,自点 $P_0(x_0, y_0, z_0)$ 分别作各坐标面和各坐标轴的垂线,写出各垂足的坐标.

5. 求点 $M(4,-3,5)$ 到各坐标轴的距离.

6. 设 X 为一个集合, $X = \bigcup\limits_{1\le i\le n} X_i$, $X_i\cap X_j = \quad$, $i\neq j$, $i, j = 1, 2, \cdots, n$.在 X 上定义二元关系 R: $x, y\in X$, xRy 当且仅当存在 $i\in\{1, 2, \cdots, n\}$ 使得 $x, y\in X_i$.证明: R 为 X 上的一个等价关系.

7．设 S_1 与 S_2 为两集合，$f: S_1 \to S_2$，令 $B=\{(a,b) \mid a,b\in S_1, f(a)=f(b)\}$．证明：$B$ 为 S_1 上的一个等价关系．

8．设 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为几何空间 $\mathbf{R}^3$ 中向量且 $\boldsymbol{u}=\boldsymbol{a}-\boldsymbol{b}+2\boldsymbol{c}$，$\boldsymbol{v}=-\boldsymbol{a}+3\boldsymbol{b}-\boldsymbol{c}$．试用 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 表示 $2\boldsymbol{u}-3\boldsymbol{v}$．

9．设平面上一个四边形的对角线互相平分，用向量证明它是平行四边形．

10．在3维欧氏几何空间 $\mathbf{R}^3$ 中，设 $\boldsymbol{m}=3\boldsymbol{i}+5\boldsymbol{j}+8\boldsymbol{k}$，$\boldsymbol{n}=2\boldsymbol{i}-4\boldsymbol{j}-7\boldsymbol{k}$ 和 $\boldsymbol{p}=5\boldsymbol{i}+\boldsymbol{j}-4\boldsymbol{k}$，求向量 $\boldsymbol{a}=4\boldsymbol{m}+3\boldsymbol{n}-\boldsymbol{p}$ 在 x 轴上的投影及在 y 轴上的分向量．

11．在3维欧氏几何空间 $\mathbf{R}^3$ 中，求与向量 $\boldsymbol{a}=(6,7,-6)$ 共线的单位向量．

12．在3维欧氏几何空间 $\mathbf{R}^3$ 中，设 $\boldsymbol{a}=3\boldsymbol{i}-\boldsymbol{j}-2\boldsymbol{k}$，$\boldsymbol{b}=\boldsymbol{i}+2\boldsymbol{j}-\boldsymbol{k}$，求

(1) $\boldsymbol{a}\cdot\boldsymbol{b}$ 及 $\boldsymbol{a}\times\boldsymbol{b}$；

(2) $(-2\boldsymbol{a})\cdot 3\boldsymbol{b}$ 及 $\boldsymbol{a}\times 2\boldsymbol{b}$；

(3) $\boldsymbol{a},\boldsymbol{b}$ 的夹角的余弦．

13．已知 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为几何空间 $\mathbf{R}^3$ 中单位向量，且 $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}=\boldsymbol{0}$，计算 $\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{b}\cdot\boldsymbol{c}+\boldsymbol{a}\cdot\boldsymbol{c}$．

14．在几何空间 $\mathbf{R}^3$ 中，设 $\boldsymbol{a}=(3,5,-2)$，$\boldsymbol{b}=(2,1,4)$，问 λ 与 μ 有怎样的关系，能使得 $\lambda\boldsymbol{a}+\mu\boldsymbol{b}$ 与 z 轴垂直．

15．设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为几何空间 $\mathbf{R}^3$ 中向量且 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=2$．求 $[(\boldsymbol{a}+\boldsymbol{b})\times(\boldsymbol{b}+\boldsymbol{c})]\cdot(\boldsymbol{c}+\boldsymbol{a})$．

16．判别以下集合对于所指的运算是否构成实数域 $\mathbf{R}$ 上线性空间．

(1) 次数不大于 $n(n\geq 1)$ 的整系数多项式的全体，对于多项式的加法和实数与多项式的数乘．

(2) 所有在区间 $[0,1]$ 不连续的实函数，在如本章第三节例3定义加法和数乘下．

(3) V 为区间 $[0,1]$ 上所有不连续实函数的集合，任给 $f,g\in V$，任给 $a\in\mathbf{R}$，若定义：

$$(f+g)(x)=0,\quad (af)(x)=0,\quad \forall x\in[0,1].$$

(4) $V=\{f(t)=a_0+a_1\sin t+b_1\cos t+a_2\sin 2t+b_2\cos 2t+\cdots+a_n\sin nt+b_n\cos nt \mid a_0,a_1,b_1,a_2,b_2,\cdots,a_n,b_n\in\mathbf{R}\}$，在如本章第三节例3定义加法和数乘下．

(5) $V=\{f(t)=a_0+a_1\sin t+a_2\sin 2t+\cdots+a_n\sin nt \mid a_0,a_1,a_2,\cdots,a_n\in\mathbf{R}\}$，在如本章第三节例3定义加法和数乘下．

17．在 n 维行向量空间 $\mathbf{R}^n$ 中，分量满足下列条件的向量能否构成子空间？

(1) $\{(x_1,x_2,\cdots,x_n) \mid x_1+x_2+\cdots+x_n=0\}$；

(2) $\{(x_1,x_2,\cdots,x_n) \mid x_1+x_2+\cdots+x_n=1\}$；

(3) $\{(x_1,x_2,\cdots,x_n) \mid x_1\geq 0,\ x_1x_2=0,\ x_1+x_2=3x_3\}$；

(4) $\{(x_1, x_2, \cdots, x_n) \mid x_1 + 2x_2 + \cdots + nx_n = 0, x_1 + x_2 = 0\}$.

18．设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 是向量空间 V 中 n 个向量:

(1) 若 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性无关,能否判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$的线性相关性;

(2) 若 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$线性无关,能否判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$的线性相关性;

(3) 若 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性相关,能否判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$的线性相关性;

(4) 若 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$线性相关,能否判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$的线性相关性.

19．证明:向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$线性无关当且仅当向量组 $\boldsymbol{a}_1, \boldsymbol{a}_1 + \boldsymbol{a}_2, \cdots, \boldsymbol{a}_1 + \boldsymbol{a}_2 + \cdots + \boldsymbol{a}_n$ 线性无关.

20．设向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$和向量组 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_m$可以相互线性表示.证明:它们的秩相同.

21．试证:在四元行向量空间 $\mathbf{R}^4$ 中,由$(1,1,0,0)$,$(1,0,1,1)$生成的子空间与由$(2,-1,3,3)$,$(0,1,-1,-1)$生成的子空间是同一个子空间.

22．设 $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 是 n 维向量空间 V 中 n 个向量,若 V 中任何向量均可由它们线性表示,则 $\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_n$ 是 V 中的一个基.

23．证明:$\dim[\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m] \leq m$,且等号成立当且仅当 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$ 线性无关.

24．设 U, W 是向量空间 V 的两个非零维的子空间且 $U \not\subset W$, $W \not\subset U$.证明:

(1) 在 V 中存在向量 $\boldsymbol{a}$,使 $\boldsymbol{a} \notin U$, $\boldsymbol{a} \notin W$ 同时成立;

(2) $U \cup W$ 不是 V 的子空间.

25．U, W 是 V 的子空间.证明:

(1) $U^{\perp}$是 V 的子空间;

(2) $U \cap U^{\perp} = \{\mathbf{0}\}$;

(3) $(U + W)^{\perp} = U^{\perp} \cap W^{\perp}$;

(4) $\dim U + \dim U^{\perp} = \dim V$.

26．设

$$W_1 = \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid x_3 = x_2, x_1 = x_4\},$$

$$W_2 = \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid x_3 = -x_2, x_1 = -x_4\}.$$

证明:$\mathbf{R}^4 = W_1 \oplus W_2$.

27．证明:由下列向量组分别构造一个标准正交向量组.

(1) $(1,1,1)$,$(0,1,1)$,$(0,0,1)$;

(2) $(1,2,2,-1)$,$(1,1,-5,3)$,$(3,2,8,-7)$;

(3) $(1,1,-1,2)$,$(5,8,-2,-3)$,$(3,9,3,8)$;

(4) $(2,1,3,-1)$,$(7,4,3,-3)$,$(1,1,-6,0)$,$(5,7,7,8)$.

28．设 V 是 $\mathbf{R}$ 上所有多项式构成的向量空间,U 是 $\mathbf{R}$ 上所有常数项为零

的多项式所构成的向量空间.证明:U 是 V 的真子空间(即 V 中有向量不属于 U,而 U 又是 V 的子空间).

29.实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 线性相关的充分必要条件是:存在 $i\in\{1,2,\cdots,n\}$使得

$$[\boldsymbol{v}_1]+[\boldsymbol{v}_2]+\cdots+[\boldsymbol{v}_n]=[\boldsymbol{v}_1]+\cdots+[\boldsymbol{v}_{i-1}]+[\boldsymbol{v}_{i+1}]+\cdots+[\boldsymbol{v}_n],$$

即由 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 生成的向量空间等于除掉 $\boldsymbol{v}_i$ 后其余的向量所生成的空间.

30.$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 是实数域 $\mathbf{R}$ 上的向量空间 V 的一组向量且 $\boldsymbol{v}_i\neq\mathbf{0}$,$i=1,2,\cdots,n$.则下列命题等价:

(1) $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 线性无关;

(2) 任意 $i\in\{1,2,\cdots,n\}$,

$$[\boldsymbol{v}_1]+[\boldsymbol{v}_2]+\cdots+[\boldsymbol{v}_n]\neq[\boldsymbol{v}_1]+\cdots+[\boldsymbol{v}_{i-1}]+[\boldsymbol{v}_{i+1}]+\cdots+[\boldsymbol{v}_n],$$

即除掉任一向量 $\boldsymbol{v}_i$ 后其余的 $n-1$ 个向量所生成的空间为$[\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n]$的真子空间;

(3) $V=[\boldsymbol{v}_1]\oplus[\boldsymbol{v}_2]\oplus\cdots\oplus[\boldsymbol{v}_n]$.即 V 是子空间$[\boldsymbol{v}_1],[\boldsymbol{v}_2],\cdots,[\boldsymbol{v}_n]$的直和.

31.设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_s$ 和 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_t$ 是向量空间 V 的两个线性无关的向量组,证明:$\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_s,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_t$ 线性无关当且仅当

$$[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_s]\cap[\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_t]=\{\mathbf{0}\}.$$

32.假定 U,V,W 是一个线性空间的子空间,证明:

$$(U\cap V)+(W\cap U)\qquad U\cap(V+W).$$

33.$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是实数域上 n 维线性空间 V 的基,向量

$$\alpha=a_1\alpha_1+\cdots+a_n\alpha_n,\ \beta=b_1\alpha_1+\cdots+b_n\alpha_n$$

对应实数$(\alpha,\beta)=\sum_{k=1}^{n}ka_kb_k$.试问$(\alpha,\beta)$是否为向量空间 V 的内积运算?

习题 B

1.设向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性相关而 $\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{a}_4$ 线性无关.证明:(1) $\boldsymbol{a}_1$ 能被 $\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性表示;(2) $\boldsymbol{a}_4$ 不能被 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性表示.

2.设向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 线性无关.证明:向量组 $\boldsymbol{a}_1+\boldsymbol{a}_2,\boldsymbol{a}_2+\boldsymbol{a}_3,\cdots,\boldsymbol{a}_{n-1}+\boldsymbol{a}_n,\boldsymbol{a}_n+\boldsymbol{a}_1$,当 n 为偶数时线性相关,当 n 为奇数时线性无关.

3.U,W 是向量空间 V 的两个非零真子空间并且 $U\neq W$.证明:

(1) 任给 $\boldsymbol{a}\in V\setminus W$,$H=\{\boldsymbol{a}+\boldsymbol{b}\mid\boldsymbol{b}\in W\}$不是 V 的子空间;

(2) 当 $\dim(U)=\dim(V)$时,存在 V 的子空间 P 使得 $V=W\oplus P=U\oplus P$.

4.设 U,W 是向量空间 V 的两个子空间.证明:$U+W=U\cup W$ 的充分必

要条件为 $U \quad W$ **或** $W \quad U$.

5 . **设** G, H, K **为向量空间的子空间 .证明:**

(1) $G \cap (G \cap H + K) = G \cap H + G \cap K$;

(2) $(G + H) \cap (G + K) = G + (G + H) \cap K$.

6 . **设** U, W **是向量空间** V **的两个子空间 .证明:**

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) .$$

7 . **设** U, W **是向量空间** V **的两个子空间且**

$$\dim(U + W) = \dim(U \cap W) + 1 .$$

证明: $U \quad W$ **或** $W \quad U$.

8 . **设** $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_5\}$ **是欧氏空间** V **的一个标准正交基**, $W = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3]$, **其中**

$$\boldsymbol{\alpha}_1 = \boldsymbol{e}_1 - \boldsymbol{e}_5, \boldsymbol{\alpha}_2 = \boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_4, \boldsymbol{\alpha}_3 = 2\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 .$$

试求 W **的一个标准正交基** .

9 . **设** $S_1, S_2, \cdots, S_m$ **都是 R 上** n **维向量空间** V **的子空间,且维数均小于** n **.证明:存在** $\alpha \in V$, **对任意** $i = 1, 2, \cdots, m$, $\alpha | \quad S_i$.

10 . **在向量空间** $P[x]_3$ **(次数不超过 3 的所有实系数多项式的集合)定义两个向量** $f(x), g(x) \in P[x]_3$ **对应实数**

$$(f(x), g(x)) = \int_{-1}^{1} f(x) g(x) \mathrm{d} x,$$

证明: $(f(x), g(x))$ **是内积运算,且** $P[x]_3$ **在此内积运算下是一个欧氏空间 .并求它的一个标准正交基** .

第二章　空间解析几何

在解析几何中,坐标法把空间或平面上的点与它的坐标对应起来,从而使曲线和曲面与代数方程对应起来,这样既可以用代数方法来研究几何问题,也可以借助几何直观来研究代数问题,代数和几何常常是同一个数学问题的两个方面 .这个思想是伟大的数学家 R .Descartes(1596—1650)提出的 .他在 1637 年出版的《方法论》的附录《几何学》被后世作为解析几何的起点 .坐标法使数学研究的两类对象"形"与"数"统一起来,并在数学中引入"变量",完成数学史划时代的变革,为后来微积分奠定了基础 .不过"坐标"这个词一般认为是微积分创始人之一 G .W .Leibniz (1646—1716) 在 1694 年首次使用的 .在 18 世纪,数学家 A .Clairaut(1713—1765) 及 L .Euler (1707—1783) 发展了 3 维空间的解析几何, Clairaut 提出空间一个方程可表示曲面,而二个联立的方程可表示空间中一条曲线 在 1731 年他给出了诸如球面、椭球面等一些二次曲面的方程 .G .Monge (1746—1818) 更加扩展了三维空间的解析几何 .

第一节　平面及其方程

本章总假设在空间已建立了直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$(右手系),其中 O 为坐标原点, $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 为相互垂直的单位向量;并在该坐标系下讨论空间几何问题,也即在 3 维欧氏几何空间 $\mathbf{R}^3$ 中讨论问题 .

任何空间曲面都可看作点的几何轨迹,而曲面上每一个点在空间坐标系都有与之相对应的坐标,曲面上的点所具有的性质有时可用其点的坐标所满足的代数方程表达出来 .

定义 2.1　如果空间曲面 S 上任一点的坐标(x, y, z)都满足方程 $F(x, y, z) = 0$, 且坐标(x, y, z)满足方程 $F(x, y, z) = 0$ 的点均在曲面上,则该方程称为曲面 S 的方程,而曲面 S 称为该方程的图形 .

1.1　平面方程

1. 平面的点法式方程

因为过空间一点可以作且只能作一平面垂直一条已知的直线,所以当平面

Π上一点 $M_0(x_0, y_0, z_0)$和与其垂直的一条直线的方向为已知时,平面 Π的位置就完全确定了.下面求这个平面的方程.

如果一非零向量垂直于某一平面,这个向量就叫做该平面的法线向量,平面上任一向量均与该平面的法线向量垂直.设 M_0 是平面 Π上一个固定点.下面求过 M_0 且法线向量为 $\boldsymbol{n}$的平面方程.因为已建立了直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$,所以直角坐标系中每一点可用坐标表示,同时每个向量也可用坐标表示.设向量 $\boldsymbol{n}$的坐标为(A, B, C), $M_0(x_0, y_0, z_0)$为平面 Π上一个固定点, $M(x, y, z)$为平面 Π上一个动点,也即任意一点,点 $M(x, y, z)$在平面 Π上的充要条件是以 M_0 为起点 M 为终点的向量$\overline{M_0 M} = (x - x_0, y - y_0, z - z_0)$与平面的法线向量 $\boldsymbol{n}$ 垂直(图 1).$\overline{M_0 M}$与 $\boldsymbol{n}$ 垂直可用

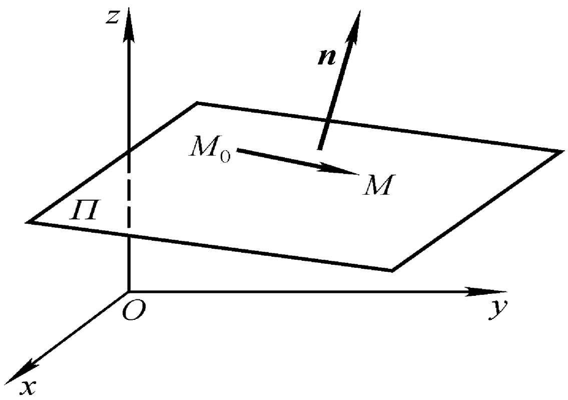


图 1

向量代数来表达,即它们数量积 $\boldsymbol{n}\cdot\overline{M_0 M} = 0$,由向量的数量积在直角坐标系$\{O; \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$中的坐标表达式,有

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \tag{2.1}$$

在平面 Π上的点的坐标满足方程(2.1);而坐标满足方程(2.1)的点都在平面 Π上,因此,方程(2.1)为平面 Π的方程,称为点法式方程.

求平面 Π的方程的过程说明,由于坐标法把几何问题转化为代数问题,所以可以用代数方法来解决几何问题.反过来,也使一些代数问题有了直观的几何解释,为代数问题的解决提供思路.

例 1 求过点$(5, -3, 0)$且以 $\boldsymbol{n} = (0, -2, 2)$为法线向量的平面的方程.

解 由方程(2.1)知,所求平面的方程为

$$0(x - 5) - 2(y + 3) + 2z = 0, \quad 即 \quad y - z + 3 = 0.$$

例 2 求过三点 $M_1(1, 2, 3)$、$M_2(3, 4, 6)$和 $M_3(4, 3, 3)$的平面的方程.

解 向量$\overline{M_1 M_2} = (2, 2, 3)$及$\overline{M_1 M_3} = (3, 1, 0)$在所求平面上,而

$$\boldsymbol{n} = \overline{M_1 M_2} \times \overline{M_1 M_3} = -3\boldsymbol{i} + 9\boldsymbol{j} - 4\boldsymbol{k}$$

为该平面的一个法线向量.由方程(2.1),所求平面的方程为

$$-3(x-1)+9(y-2)-4(z-3)=0,\quad \text{即}\quad -3x+9y-4z=3.$$

例 3　设平面与 x, y, z 轴的交点依次为 $P(a,0,0)$, $Q(0,b,0)$, $R(0,0,c)$ 三点(图 2),求这个平面的方程(其中 a, b, c 都不为零).

解　因为向量 $\overline{PQ}=(-a,b,0)$, $\overline{PR}=(-a,0,c)$ 在所求平面上,且

$$\overline{PQ}\times\overline{PR}=bc\boldsymbol{i}+ac\boldsymbol{j}+ab\boldsymbol{k}$$

为该平面法线向量,且 P 点在平面上.由方程(2.1)知,该平面方程为

$$bc(x-a)+acy+abz=0.$$

因为 a, b, c 都不为零,整理该方程便得如下的平面方程

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1.$$

该方程称为平面的截距式方程,而 a, b, c 分别称为该平面在 x, y, z 轴上的截距.

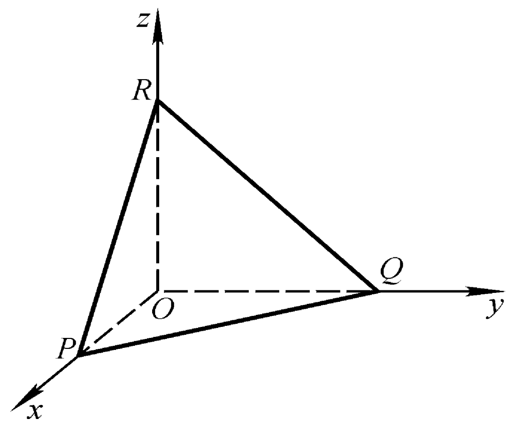


图 2

2. 平面的一般方程

既然任意一个平面可由其上一点及该平面的法线向量唯一确定,因而平面方程可用(2.1)式表示.这样,任一平面都可以用一个三元一次方程来表示.反过来,对任意一个三元一次方程

$$Ax+By+Cz+D=0, \tag{2.2}$$

任取满足该方程的一组数 x_0, y_0, z_0,即

$$Ax_0+By_0+Cz_0+D=0. \tag{2.3}$$

由(2.2)及(2.3)知

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0. \tag{2.4}$$

与(2.1)对比,可知方程(2.4)表示通过点 $M_0(x_0, y_0, z_0)$ 且以 $\boldsymbol{n}=(A, B, C)$ 为法线向量的平面.因此任一三元一次方程的图形都是一个平面.方程(2.2)称为平面的一般方程,其中系数 A, B, C 有明显的几何意义,即 $\boldsymbol{n}=(A, B, C)$ 为该平面的法线向量.

例 4 求过点 $M_1(2,-1,4)$, $M_2(-1,3,-2)$且垂直平面 $\Pi_1: -2x+3y-z=0$ 的平面方程.

解 设所求平面的一个法线向量为 $\boldsymbol{n}=(A, B, C)$.因为向量 $\overline{M_1M_2}=(-3,4,-6)$在所求的平面上,所以它必与 $\boldsymbol{n}$ 垂直.又因为所求的平面与平面 Π_1 垂直,所以平面 Π_1 的法线向量 $\boldsymbol{n}_1=(-2,3,-1)$必与 $\boldsymbol{n}$ 垂直.这样可取

$$\boldsymbol{n}=\overline{M_1M_2}\times\boldsymbol{n}_1=14\boldsymbol{i}+9\boldsymbol{j}-\boldsymbol{k}.$$

由平面的点法式方程(2.1)知,所求平面的方程为 $14(x-2)+9(y+1)-(z-4)=0$,即

$$14x+9y-z-15=0.$$

例 5 求通过 y 轴和点(8,3,-2)的平面的方程.

解 设所求平面 Π 的方程为

$$Ax+By+Cz+D=0. \tag{2.5}$$

因为平面通过 y 轴,其法线向量 $\boldsymbol{n}=(A, B, C)$垂直于 y 轴,即

$$0=\boldsymbol{n}\cdot\boldsymbol{j}=(A\boldsymbol{i}+B\boldsymbol{j}+C\boldsymbol{k})\cdot\boldsymbol{j}=B$$

又由于平面通过 y 轴,当然过原点,于是 $D=0$.所以可设这平面的方程为 $Ax+Cz=0$.又因为点(8,3,2)在该平面上,所以 $x=8$, $y=3$, $z=-2$ 满足方程 $Ax+Cz=0$,即 $8A-2C=0$,这样 $C=4A$,于是 $Ax+4Az=0$.由 $A\neq0$ 可知,所求的平面方程为 $x+4z=0$.

3. 两平面的夹角

两平面的法线向量的夹角(通常指锐角)称为两平面的夹角(图 3).设有平面

$$\Pi_1: A_1x+B_1y+C_1z+D_1=0,$$
$$\Pi_2: A_2x+B_2y+C_2z+D_2=0,$$

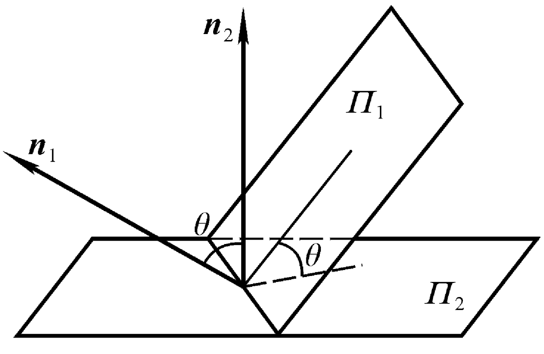


图 3

平面 Π_1 的法线向量 $\boldsymbol{n}_1=(A_1, B_1, C_1)$,平面 Π_2 的法线向量 $\boldsymbol{n}_2=(A_2, B_2, C_2)$,两平面的夹角 θ 为 $\langle\boldsymbol{n}_1, \boldsymbol{n}_2\rangle$ 和 $\langle-\boldsymbol{n}_1, \boldsymbol{n}_2\rangle$ 两者中的锐角,由第一章第二节知

$$\cos\theta=\frac{|A_1A_2+B_1B_2+C_1C_2|}{\sqrt{A_1^2+B_1^2+C_1^2}\sqrt{A_2^2+B_2^2+C_2^2}},\tag{2.6}$$

由公式(2.6)**知**, Π_1 **和** Π_2 **互相垂直当且仅当** $A_1A_2+B_1B_2+C_1C_2=0$; Π_1 **与** Π_2 **平行或重合当且仅当** $\boldsymbol{n}_1$ **与** $\boldsymbol{n}_2$ **线性相关**,**即**$(A_1,B_1,C_1)=k(A_2,B_2,C_2)$,**也即**$\dfrac{A_1}{A_2}=\dfrac{B_1}{B_2}=\dfrac{C_1}{C_2}$.

例 6　**求两平面**$2x-2y+4x-7=0$,$6x+3y+3z-15=0$**的夹角** θ.

解　**由公式**(2.6)**有**

$$\cos\theta=\frac{|2\times6+(-2)\times3+4\times3|}{\sqrt{2^2+(-2)^2+4^2}\sqrt{6^2+3^2+3^2}}=\frac{1}{2}.$$

因此,**两平面夹角** $\theta=\dfrac{\pi}{3}$.

例 7　**设** $P_0(x_0,y_0,z_0)$**是平面** Π: $Ax+By+Cz+D=0$ **外一个点**,**求** P_0 **到该平面的距离** d.

解　**在平面上任取一点** $M_1(x_1,y_1,z_1)$, $\boldsymbol{n}=(A,B,C)$**为平面** Π **的法线向量**,**由图 4 可知**, $d=|\mathrm{Prj}_{\boldsymbol{n}}\overline{M_1P_0}|$.**设** $\boldsymbol{e}_n$ **是向量** $\boldsymbol{n}$ **方向的单位向量,那么有**

$$d=|\mathrm{Prj}_{\boldsymbol{n}}\overline{M_1P_0}|=|\overline{M_1P_0}\cdot\boldsymbol{e}_n|,$$

$$\boldsymbol{e}_n=\frac{\boldsymbol{n}}{|\boldsymbol{n}|}=\frac{1}{|\boldsymbol{n}|}(A,B,C),\quad \text{其中}\quad |\boldsymbol{n}|=\sqrt{A^2+B^2+C^2},$$

$$\overline{M_1P_0}=(x_0-x_1,y_0-y_1,z_0-z_1),$$

$$d=|\mathrm{Prj}_{\boldsymbol{n}}\overline{M_1P_0}|=\frac{1}{|\boldsymbol{n}|}|A(x_0-x_1)+B(y_0-y_1)+C(z_0-z_1)|$$

$$=\frac{1}{|\boldsymbol{n}|}|Ax_0+By_0+Cz_0-(Ax_1+By_1+Cz_1)|.$$

又因 $Ax_1+By_1+Cz_1+D=0$,**所以**

$$d=|\mathrm{Prj}_{\boldsymbol{n}}\overline{M_1P_0}|=\frac{|Ax_0+By_0+Cz_0+D|}{|\boldsymbol{n}|}=\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}.\tag{2.7}$$

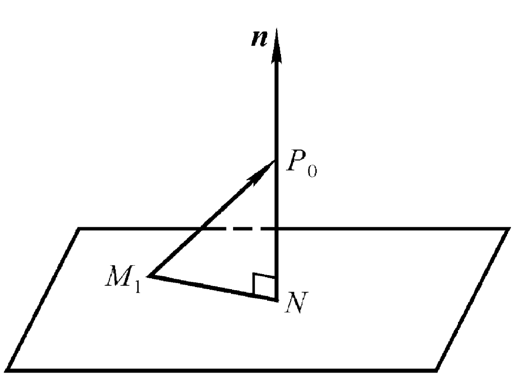


图 4

例如,求点(3,1,0)到平面 $x+y-2z+1=0$ 的距离,可利用公式(2.7),得

$$d=\frac{1\times 3+1\times 1-2\times 0+1}{1^2+1^2+(-2)^2}=\frac{5}{6}.$$

1.2　空间直线及其方程

1. 空间直线的参数方程与对称式方程

与一条已知直线平行的非零向量叫做这条直线的方向向量.在空间中,过空间一点能且只能作一条直线平行于一已知直线.

设 L 为一条直线, $M_0(x_0,y_0,z_0)$为直线 L 上一个固定点, L 的方向向量 $\boldsymbol{s}=(m,n,p)$.点 $M(x,y,z)$在直线 L 上的充要条件为向量

$$\overline{M_0M}=(x-x_0,y-y_0,z-z_0)$$

与 L 的方向向量 $\boldsymbol{s}$ 共线(图 5).向量 $\overline{M_0M}$ 与向量 $\boldsymbol{s}$ 共线的充要条件是

$$\overline{M_0M}=t\boldsymbol{s},$$

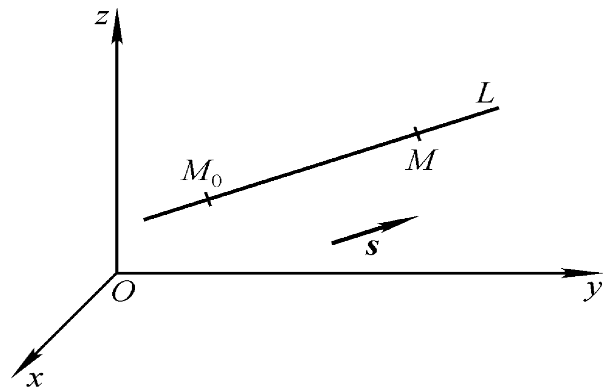


图 5

其中 t 为某一实数,也即

$$(x-x_0,y-y_0,z-z_0)=t(m,n,p)=(tm,tn,tp),\qquad(2.8)$$

$$x-x_0=tm,\quad y-y_0=tn,\ z-z_0=tp,$$

因此,当 m,n,p 都不为零时,有

$$\frac{x-x_0}{m}=\frac{y-y_0}{n}=\frac{z-z_0}{p}.\qquad(2.9)$$

方程组(2.9)叫做直线的对称式方程或点向式方程.(2.8)式可写为

$$\begin{aligned}x&=x_0+mt,\\ y&=y_0+nt,\\ z&=z_0+pt,\end{aligned}\qquad(2.10)$$

其中 t 为变数,称为参数,方程组(2.10)称为直线的参数方程.

由方程(2.10)可知,当 $t=0$ 时, $x=x_0$, $y=y_0$, $z=z_0$ 满足直线方程,这说

明点 $M_0(x_0, y_0, z_0)$在直线上.当 $t>0$ 时,随着 t 增加,动点 $M(x, y, z)$沿着 $\boldsymbol{s}=(m, n, p)$方向离点 M_0 越来越远;当 $t<0$ 时,随着$|t|$增加,动点 $M(x, y, z)$沿着 $\boldsymbol{s}=(m, n, p)$的反方向离点 M_0 越来越远(图 6).

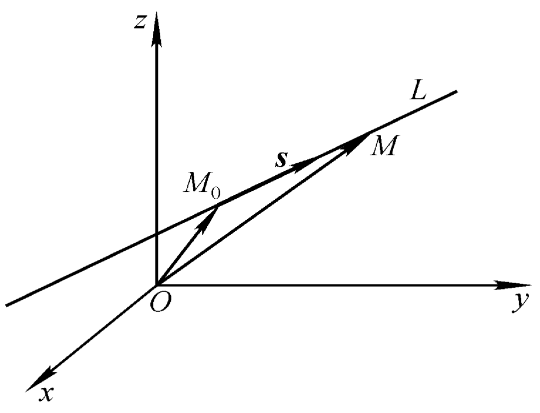


图 6

两直线的方向向量的夹角(通常指锐角)叫做两直线的夹角.可由向量的内积来确定两直线的夹角的余弦,公式推导留作练习.

例 8　求过点$(1, -2, 1)$且与平面 $3x+2y+4z+4=0$ 垂直的直线方程.

解　因为所求直线垂直于已知平面,可以取已知平面的法向量$(3, 2, 4)$作为所求直线的方向向量.由(2.9)式,得所求直线的方程为

$$\frac{x-1}{3}=\frac{y+2}{2}=\frac{z-1}{4}.$$

2. 空间直线的一般方程

由直线对称式方程(2.9)可知,当 $m\neq 0$ 时,(2.9)可写成

$$\frac{x-x_0}{m}=\frac{y-y_0}{n},$$

$$\frac{x-x_0}{m}=\frac{z-z_0}{p}.$$

第一个方程表示平行于 z 轴的平面,第二个方程表示平行于 y 轴的平面,即一个直线方程可以由两个相交平面方程组成的联立方程表示.

空间两平面平行或重合时它们的法线向量线性相关;相交时两平面的法线向量线性无关并且两平面的交集是一条空间直线.因此空间直线 L 可以看作是两个法线向量不平行的平面 Π_1 和 Π_2 的交线(图 7).设平面 Π_1 和 Π_2 法线向量不平行且它们的方程分别为

$$A_1x+B_1y+C_1z+D_1=0 \quad 和 \quad A_2x+B_2y+C_2z+D_2=0,$$

由于两个平面法线向量线性无关,所以下式不成立:

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}.$$

可设$\frac{A_1}{A_2}\neq\frac{B_1}{B_2}$,考察线性方程组

$$\begin{aligned}A_1x+B_1y+C_1z+D_1&=0,\\A_2x+B_2y+C_2z+D_2&=0.\end{aligned}\tag{2.11}$$

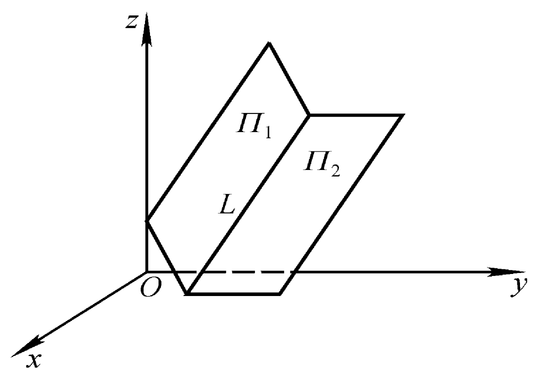


图 7

方程组(2.11)一定有如下形式的解:

$$x=mz+x_0,\ y=nz+y_0,\ z=z.$$

再令 $z=t$,于是平面 Π_1 与 Π_2 的交线方程为

$$\begin{aligned}x&=mt+x_0,\\y&=nt+y_0,\\z&=t+0.\end{aligned}$$

这表示平面 Π_1 与 Π_2 的交线是一条直线.因此相交的两个平面的方程构成的方程组(2.11)为直线 L 的方程.方程组(2.11)叫做空间直线的一般方程.

通过空间给定的直线 L 的平面有无限多个,因此直线 L 的形如(2.11)的方程组不惟一.只要在这些平面中任意选取不重合的两个平面,它们的联立方程组就表示空间直线 L.

例 9　求通过直线 L

$$\begin{aligned}A_1x+B_1y+C_1z+D_1&=0, &(\Pi_1)\\A_2x+B_2y+C_2z+D_2&=0, &(\Pi_2)\end{aligned}$$

的所有平面(通常称其为平面束)的方程.

解　设所求平面 Π 的方程为

$$Ax+By+Cz+D=0.$$

因为三平面 Π_1,Π_2 和 Π 通过同一直线,所以三个法线向量都垂直于 L,因此向量(A,B,C),(A_1,B_1,C_1),(A_2,B_2,C_2)线性相关.由于 Π_1 与 Π_2 的法线向量不平行,即(A_1,B_1,C_1),(A_2,B_2,C_2)线性无关,由命题 1.18 知$(A,B,C)=k_1(A_1,B_1,C_1)+k_2(A_2,B_2,C_2)$,实数 k_1,k_2 不能同时为零.直线 L 上的点的

坐标(x, y, z)满足方程组

$$A_1 x + B_1 y + C_1 z + D_1 = 0, \qquad (1)$$
$$A_2 x + B_2 y + C_2 z + D_2 = 0, \qquad (2)$$
$$Ax + By + Cz + D = 0. \qquad (3)$$

由$(3) - k_1(1) - k_2(2)$得 $0 = D - k_1 D_1 - k_2 D_2$,即 $D = k_1 D_1 + k_2 D_2$.注意到$(A, B, C) = k_1(A_1, B_1, C_1) + k_2(A_2, B_2, C_2)$,可得所求的平面方程为

$$\begin{aligned} Ax + By + Cz + D = {} & k_1(A_1 x + B_1 y + C_1 z + D_1) \\ & + k_2(A_2 x + B_2 y + C_2 z + D_2) = 0, \end{aligned}$$

其中实数 k_1, k_2 不能同时为零.

例 10 分别用对称式方程和参数方程表示直线

$$\begin{cases} x + y + z + 1 = 0, \\ x - y + 3z + 4 = 0. \end{cases}$$

解 令 z 等于参数 t 并解方程.如下四个方程组是彼此等价的,第二,四式分别为直线参数方程及对称式方程:

$$\begin{cases} x + y = -t - 1, \\ x - y = -3t - 4, \\ z = t; \end{cases} \qquad \begin{cases} x = -2t - \frac{5}{2}, \\ y = t + \frac{3}{2}, \\ z = t; \end{cases}$$

$$\begin{cases} \dfrac{x + \frac{5}{2}}{-2} = t, \\ \dfrac{y - \frac{3}{2}}{1} = t, \\ \dfrac{z}{1} = t; \end{cases} \qquad \frac{x + \frac{5}{2}}{-2} = \frac{y - \frac{3}{2}}{1} = \frac{z}{1}.$$

3. 直线与平面的夹角

一条直线 L 和 L 在平面 Π 上的投影直线的夹角 φ 称为直线 $\boldsymbol{L}$ 与平面 Π 的夹角.当 L 垂直 Π 时,L 在 Π 上的投影为一点,此时规定 L 与 Π 的夹角为$\dfrac{\pi}{2}$.

设平面 Π: $Ax + By + Cz + D = 0$,直线 L 的方程是

$$\frac{x - x_0}{m} = \frac{y - y_0}{n} = \frac{z - z_0}{p}.$$

因直线的方向向量 $\boldsymbol{s} = (m, n, p)$与平面的法线向量 $\boldsymbol{n} = (A, B, C)$的夹角为$\dfrac{\pi}{2} - \varphi$或$\dfrac{\pi}{2} + \varphi$(图 8).由两向量夹角余弦的坐标表示式有

$$\sin\varphi = |\cos\langle \boldsymbol{n}, \boldsymbol{s}\rangle| = \frac{|Am + Bn + Cp|}{\sqrt{A^2 + B^2 + C^2}\sqrt{m^2 + n^2 + p^2}}. \qquad (2.12)$$

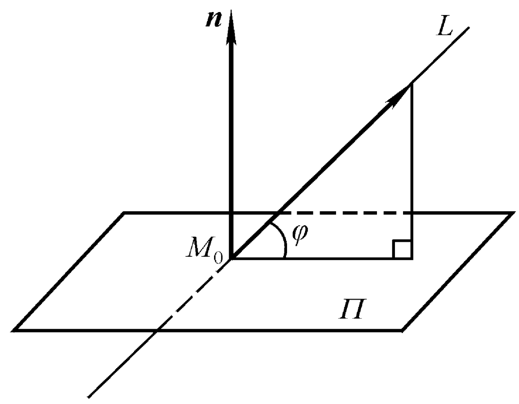


图 8

由公式(2.12)可以讨论直线 L 与平面 Π 垂直或平行的条件,留作练习.

例 11 设平面 Π 的方程为 $x+y+z=0$,直线 L 的方程为 $2x-1=2y-2=z-3$,求直线 L 与平面 Π 的夹角 φ.

解 平面 Π 的法线向量为 $(1,1,1)$,直线 L 的方向向量为 $(1,1,2)$,

$$\sin\varphi=\frac{1+1+2}{\sqrt{1^2+1^2+1^2}\sqrt{1^2+1^2+2^2}}=\frac{4}{\sqrt{3}\sqrt{6}}=\frac{4}{3\sqrt{2}}=\frac{2\sqrt{2}}{3},$$

于是 $\varphi=\arcsin\frac{2\sqrt{2}}{3}$.

第二节 曲面及其方程

2.1 几种常见曲面方程的求法

在空间解析几何及其应用中,关于曲面的研究经常遇到两个问题:(1)已知一曲面,建立这曲面的方程;(2)已知方程,研究曲面的形状.通常,无论已知曲面的方程来确立曲面,还是已知曲面确立曲面的方程都不是很容易的事情.应用专门软件,计算机可绘制出一个方程所表示的曲面形状.下面通过建立几个常见曲面的方程,来介绍几种确立曲面的方程的方法.

例 1 建立球心在点 $M_0(x_0, y_0, z_0)$,半径为 R 的球面的方程.

解 点 $M(x, y, z)$是球面上的点当且仅当 $|M_0M|=R$,即

$$R=|M_0M|=\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2},$$

或者

$$(x-x_0)^2+(y-y_0)^2+(z-z_0)^2=R^2. \tag{2.13}$$

方程(2.13)是以 $M_0(x_0, y_0, z_0)$为球心、R 为半径的球面方程.

特别地,当球心在原点(图 1),即 $x_0=y_0=z_0=0$,此时球面方程为

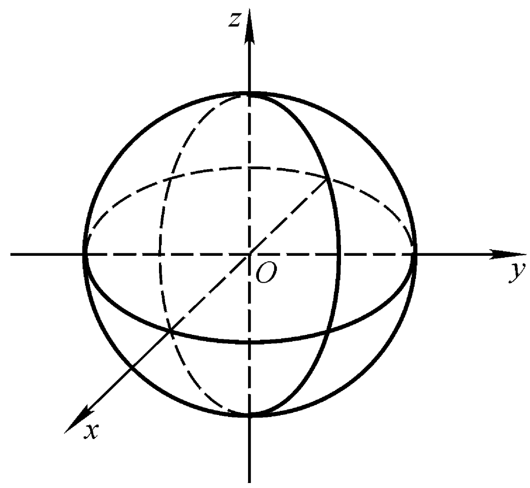


图 1

$$x^2 + y^2 + z^2 = R^2 .$$

例 2　**设** $A(2,1,0)$**和** $B(3,-1,2)$,**求与** A **及** B **两点距离相等点的轨迹的方程** .

解　**设** $M(x,y,z)$**为所求轨迹上的任何一点**,**应有** $|AM| = |BM|$,**即**

$$(x-2)^2 + (y-1)^2 + z^2 = \quad (x-3)^2 + (y+1)^2 + (z-2)^2 .$$

化简便得

$$2x - 4y + 4z - 9 = 0 .$$

由(2 2)**式知**,**所求轨迹为平面** .

例 3　**求** $x^2 + y^2 + z^2 + Dx + Ey + Fz + G = 0$ **表示的曲面**,**其中** D, E, F, G **为常数** .

解　**通过配方**,**有**

$$0 = x^2 + y^2 + z^2 + Dx + Ey + Fz + G$$
$$= (x + D\ 2)^2 + (y + E\ 2)^2 + (z + F\ 2)^2 + G - (D^2 + E^2 + F^2)\ 4 .$$
$$(x + D\ 2)^2 + (y + E\ 2)^2 + (z + F\ 2)^2 = (D^2 + E^2 + F^2)\ 4 - G$$

当 $(D^2 + E^2 + F^2)\ 4 - G > 0$ **时**,**方程表示的曲面为以** $(-D\ 2, -E\ 2, -F\ 2)$ **为球心**,　$(D^2 + E^2 + F^2)\ 4 - G$ **为半径的球面** .**当** $(D^2 + E^2 + F^2)\ 4 - G = 0$ **时**,**球面退化为一点** $(-D\ 2, -E\ 2, -F\ 2)$.**当** $(D^2 + E^2 + F^2)\ 4 - G < 0$ **时**,**在空间中不存在坐标满足方程的点**,**这时常称方程所表示的“ 曲面 ”为虚球面** .

2.2　旋转曲面

将一条平面曲线围绕其所在平面上的一条给定直线旋转一周所形成的曲面称为旋转曲面,**那条给定直线称为该旋转曲面的轴** .

设曲线 C **在** yOz **面上的方程为** $f(y,z) = 0$,**把曲线** C **绕** z **轴旋转一周**,**得到一个以** z **轴为轴的旋转曲面**(**图** 2) .**下面求旋转曲面的方程** .**设在** yOz **面的**

曲线 C 上的一点 $M_1(0, y_1, z_1)$绕 z 轴转动到 $M(x, y, z)$.这时,点 M 的 z 轴坐标等于 z_1,点 M 到 z 轴的距离 d 等于$|y_1|$,即

$$z_1 = z, \ |y_1| = d = \sqrt{x^2 + y^2}\left(y_1 = \pm\sqrt{x^2 + y^2}\right). \tag{2.14}$$

因为 $M_1(0, y_1, z_1)$在曲线 C 上,所以 $f(y_1, z_1) = 0$.由(2.14),有

$$f\left(\pm\sqrt{x^2 + y^2}, z\right) = 0. \tag{2.15}$$

该方程就是所求旋转曲面的方程.

同理,可求 yOz 面上的曲线 C 绕 y 轴旋转所成的旋转曲面的方程为

$$f\left(y, \pm\sqrt{x^2 + z^2}\right) = 0.$$

一条直线绕与其相交但不垂直的直线旋转一周,所得旋转曲面叫圆锥面.两直线的交点叫圆锥面的顶点,两直线所夹锐角 $\alpha\left(0 < \alpha < \dfrac{\pi}{2}\right)$叫圆锥面的半顶角(图 3).

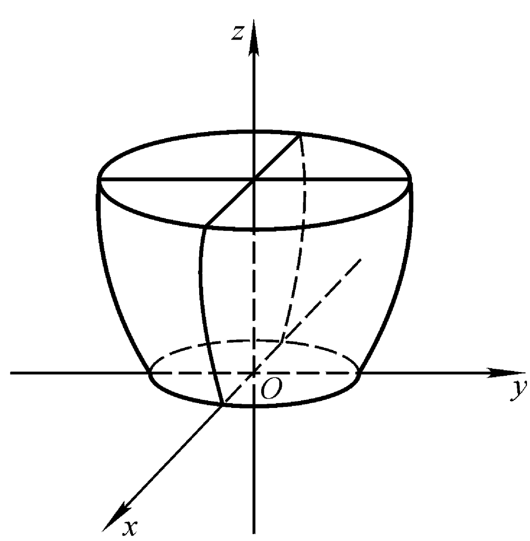


图 2

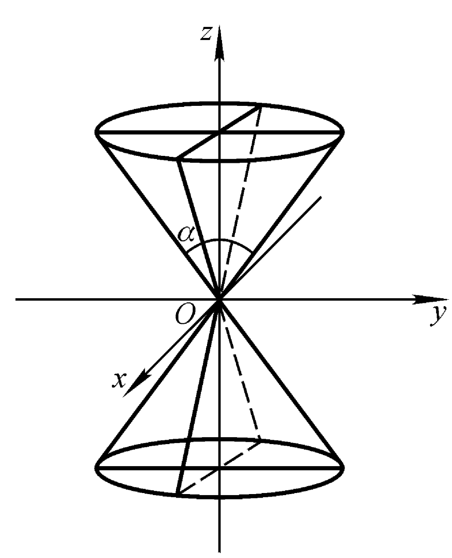


图 3

例 4　试求顶点在原点 O,旋转轴为 z 轴,半顶角为 $\alpha = \dfrac{\pi}{4}$的圆锥面的方程.

解　在 yOz 面上直线 L 的方程为 $z = y\cot\alpha$,因为以 z 轴为旋转轴,由公式(2.15),将方程 $z = y\cot\alpha$中的 y 改成$\pm\sqrt{x^2 + y^2}$,便得到该圆锥面的方程

$$z = \pm\sqrt{x^2 + y^2}\cot\frac{\pi}{4} \text{ 或 } z^2 = x^2 + y^2.$$

例 5　求将 xOz 面上的抛物线$\dfrac{x^2}{a^2} = \dfrac{z}{c}$绕 z 轴旋转一周所成的旋转曲面方程.

解　仿公式(2.15),只要将抛物线方程中的 x 代换为$\pm\sqrt{x^2 + y^2}$,就可得

到绕 z 轴旋转所生成的旋转曲面的方程为

$$\frac{x^2+y^2}{a^2}=\frac{z}{c}.$$

该曲面叫做旋转抛物面.如射电望远镜,包括跟踪设备的天线外形就是旋转抛物面的一部分.

2.3　柱面

表示空间一般曲面的方程 $F(x,y,z)=0$ 有三个变量.如果只有两个变量,那么它所描述的曲面有什么特点?例如 $F(x,y)=0$.对于空间点 $M(x,y,z)$,只要 x,y 满足 $F(x,y)=0$,则该点就在方程所描述的曲面上,而与 z 轴的坐标无关,先分析一个具体的例子.

例 6　讨论方程 $x^2+y^2=R^2$ 表示的曲面形状.

解　在 xOy 面上,方程 $x^2+y^2=R^2$ 表示圆心在原点 O 且半径为 R 的圆 C,在三维几何空间中通过圆 C 上的一点 $M_0(x,y,0)$,作平行于 z 轴的直线 L.因为方程 $x^2+y^2=R^2$ 不包含 z,所以直线 L 的任意点 $M(x,y,z)$ 都满足该方程,即 M 在所求曲面上.该曲面是平行于 z 轴的直线 L 绕圆 C 运动一周形成的轨迹,称其为圆柱面(图 4).

直线 L 平行于一条定直线并沿另一条给定曲线 C 移动而形成的轨迹叫做柱面,给定曲线 C 叫做柱面的准线,动直线 L 叫做柱面的母线.方程 $F(x,y)=0$,表示柱面,其母线平行于 z 轴,其准线是 xOy 面上由方程 $F(x,y)=0$ 表示的曲线 C(图 5).方程 $G(x,z)=0$ 和方程 $H(y,z)=0$ 分别表示母线平行于 y 轴和 x 轴的柱面.

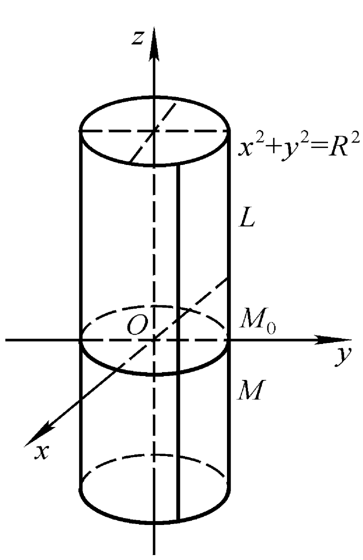


图 4

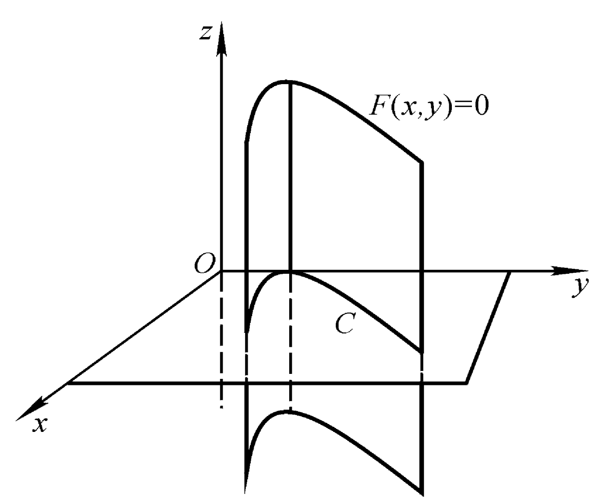


图 5

2.4　二次曲面

由平面解析几何知道,在平面直角坐标系中,二元二次方程通过坐标系的旋转及平移,可化为标准方程.由标准方程容易看出,有些方程没有图形(如 $x^2+y^2=-1$),有些方程的图形是一个点(如 $x^2+y^2=0$)、一条直线(如 $x^2=0$)、两条平行直线(如 $x^2=1$)、两条相交直线(如 $x^2-y^2=0$).还有三种方程的图形是曲线,它们是

(1) 椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$;

(2) 双曲线 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$;

(3) 抛物线 $ay=x^2$;

这三种曲线称为二次曲线.

类似地,在空间直角坐标系中,三元二次方程也可以通过坐标系的旋转及平移,化为标准方程(第七章将讨论把三元二次方程化为标准方程的问题).其中有九种方程的图形是曲面,称为二次曲面.以 xOy 面上的三种二次曲线为准线,母线平行于 z 轴的柱面都是二次曲面,分别称为椭圆柱面,双曲柱面和抛物柱面.除这三种外,还有六种二次曲面,下面介绍它们的标准方程和图形.

下面曲面方程中出现的 a, b, c 总是正数.

1. 椭球面

方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \tag{2.16}$$

的曲面称为椭球面,它与坐标轴的六个交点:$(\pm a,0,0)$,$(0,\pm b,0)$,$(0,0,\pm c)$.被称为椭球面的顶点,坐标原点到 x, y, z 轴正方向上的顶点的线段分别称为椭球半轴,a, b, c 三个正数称为半轴长.因为方程(2.16)只含 x, y, z 的平方项及常数项,如果 x, y, z 满足(2.16),则取 $\pm x, \pm y, \pm z$ 中所有可能的符号组合都满足(2.16),因此椭圆面关于坐标面、坐标轴、原点是对称的.三个坐标面、三个坐标轴及原点分别为椭球面的对称面、对称轴及对称中心.

为确定曲面的形状,一般采用平面截割法,可用一系列平行于坐标面的平面去截割曲面,得到一些截线.对二次曲面而言,这些截线一般为二次曲线、两条相交直线或点.根据这些截线的形状,可以确定曲面在空间中的图形.用平面 $z=z_0$ 截割椭球面得到截线是椭圆:

$$\begin{cases}\dfrac{x^2}{\left(a\sqrt{1-\dfrac{z_0^2}{c^2}}\right)^2}+\dfrac{y^2}{\left(b\sqrt{1-\dfrac{z_0^2}{c^2}}\right)^2}=1,\\ z=z_0.\end{cases}$$

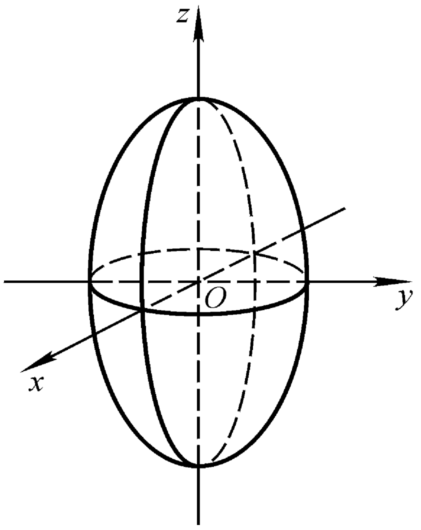


图 6

由这些截线方程可见,当$|z_0|>c$时,平面与椭球面的截线不存在;当$|z_0|=c$时,截线变为椭球面一个顶点;当$0\leq|z_0|<c$时,截线为椭圆,当$z=0$时,截出的椭圆半轴最大,分别为a与b.

用与其他两个坐标面平行的平面$y=y_0$或$z=z_0$截割椭球面与以上讨论情形相类似.椭球面形状如图6.当$a=b$时,椭球面是旋转曲面;而当$a=b=c$时,椭球面就是球面.

平面截割法是得到空间中一般曲面形状的基本方法.制作形如椭球面的灯笼,首先做成竹骨架,这些骨架常常就是该椭球面的截线.计算机图形学中也常采取这种方式给出曲面3维图形.

2. 双曲面

(1) 单叶双曲面

方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \tag{2.17}$$

的曲面称为单叶双曲面.

类似椭球面的情形,单叶双曲面是关于三个坐标面,三个坐标轴及原点对称的曲面.用平面$z=z_0$截割曲面,其截线的方程是

$$\begin{cases}\dfrac{x^2}{\left(a\sqrt{1+\dfrac{z_0^2}{c^2}}\right)^2}+\dfrac{y^2}{\left(b\sqrt{1+\dfrac{z_0^2}{c^2}}\right)^2}=1,\\ z=z_0.\end{cases}$$

截线是椭圆.当$|z_0|$从0递增,该椭圆的两半轴分别从a,b递增.

用平面$x=x_0$截割曲面,当$|x_0|<a$时,截线方程是

$$\begin{cases}\dfrac{y^2}{\left(b\sqrt{1-\dfrac{x_0^2}{a^2}}\right)^2}-\dfrac{z^2}{\left(c\sqrt{1-\dfrac{x_0^2}{a^2}}\right)^2}=1,\\ x=x_0.\end{cases}$$

截线是双曲线.当$|x_0|$从0递增趋于a,该双曲线的两半轴分别从b和c递减趋于零;当$|x_0|=a$时,截线是一对相交直线,其方程为

$$\frac{y^2}{b^2}-\frac{z^2}{c^2}=0,$$

$$|x|=a.$$

类似地,当$|x_0|>a$时,截线是半轴为

$$c\sqrt{\frac{x_0^2}{a^2}-1}\text{ 及 }b\sqrt{\frac{x_0^2}{a^2}-1}$$

的双曲线.当$|x_0|$从a递增,该双曲线的半轴从零递增.

平面$y=y_0$截割曲面的截线也是一组双曲线.单叶双曲面的形状如图7.当$a=b$时,曲面是旋转曲面,称为单叶旋转双曲面.发电厂水冷却塔的外形一般是该曲面的一部分.

(2) 双叶双曲面

方程为

$$\frac{z^2}{c^2}-\frac{y^2}{b^2}-\frac{x^2}{a^2}=1 \tag{2.18}$$

的曲面称为双叶双曲面.

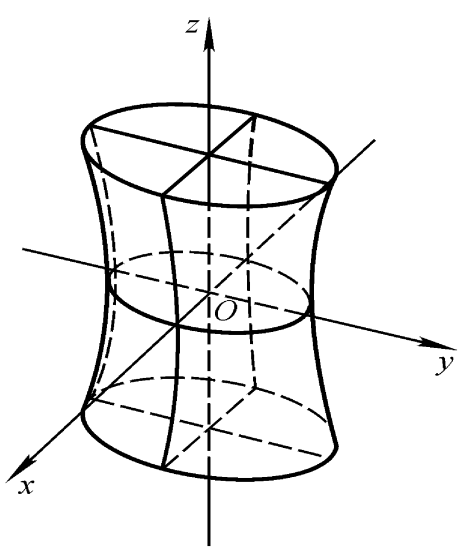


图 7

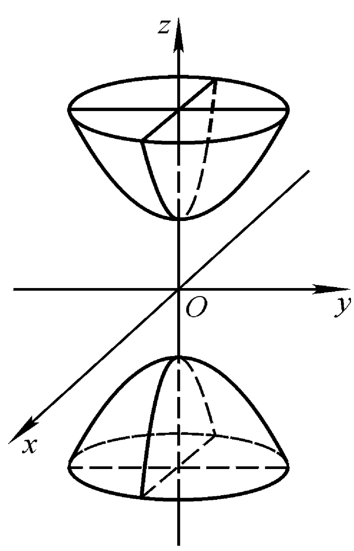


图 8

双叶双曲面是关于三个坐标面,三个坐标轴及坐标原点对称的曲面.用平面$z=z_0$截割该曲面,当$|z_0|<c$时,平面与该曲面不相交;当$|z_0|=c$时,平面与该曲面的交点仅为一点$(0,0,c)$或$(0,0,-c)$;当$|z_0|>c$时,平面$z=z_0$截割该曲面得到截线为椭圆,当$|z_0|$从c递增,两半轴从零递增.用平面$x=x_0$截割该曲面,截线为双曲线.当$|x_0|$从0递增,双曲线的半轴分别由c与b递增.用平面$y=y_0$截割该曲面,其截线也是双曲线.双叶双曲面的形状如图8所示.当$a=b$时,曲面为旋转曲面——双叶旋转双曲面.

3．锥面

方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \tag{2.19}$$

的曲面称为二次锥面.

二次锥面是关于三个坐标面,三个坐标轴及原点对称的曲面.用平面 $z=z_0$ 截割该曲面,截线是椭圆.$|z_0|$由 0 递增,椭圆的两半轴也由 0 递增.用平面 $x=x_0$ 截割该曲面,当$|x_0|=0$时,截线是两条相交直线;当$|x_0|\neq 0$时,截线是一组双曲线,当$|x_0|$从 0 递增,半轴也从 0 递增.用平面 $y=y_0$ 截割该曲面与 $x=x_0$ 情形相类似.当 $a=b$ 时,曲面为圆锥面,二次锥面图形如图 3 所示.

锥面(2.19)与双曲面(2.17)及(2.18)有密切关系:用 $z=z_0$ 截这三个曲面,如果三条截线都存在的话,截线均为椭圆.$z=z_0$ 与单叶双曲面、二次锥面、双叶双曲面相截的截线的椭圆半轴由大到小依次排列,当 z_0 递增,三个椭圆的半轴的差递减.因而称二次锥面(2.19)为双曲面(2.17)及(2.18)的渐近锥面.

4. 抛物面

(1) 椭圆抛物面

方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=z \tag{2.20}$$

的曲面称为椭圆抛物面.

由方程知,yOz 面与 zOx 面是椭圆抛物面的对称面,z 轴是对称轴,它没有对称中心.用平面 $z=z_0$ 截割该曲面,当 $z_0<0$ 时,$z=z_0$ 与该曲面不相交;当 $z=0$ 时,有一个交点为原点,称为椭圆抛物面的顶点;当 $z_0>0$ 时,截线是椭圆,其中心在 z 轴上,随着 z_0 由 0 递增,椭圆的半轴也由 0 递增.

用平面 $x=x_0$ 或 $y=y_0$ 截割该曲面,其截线是抛物线,且其开口向 z 轴正方向.椭圆抛物面的图形如图 9 所示.当 $a=b$ 时,曲面称为旋转抛物面.

(2) 双曲抛物面

方程为

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=z, \tag{2.21}$$

的曲面称为双曲抛物面.

由方程知,曲面是关于 yOz 面,zOx 面及 z 轴对称的曲面,用平面 $z=z_0$ 截割该曲面,其截线为一组双曲线.当 z_0 从负值递增到 0,双曲线的实轴平行于 y 轴,而虚轴平行于 x 轴,半轴从正值递减到 0;当 $z_0=0$ 时,截线是一对直线;当

z_0 从 0 递增,双曲线的实轴平行于 x 轴,而虚轴平行于 y 轴,半轴递增,用平面 $x = x_0$(或 $y = y_0$)截割该曲面,其截线是抛物线,其中与 yOz(zOx)面平行的截线是开口朝下(上)的抛物线.双曲抛物面的形状如图 10 所示,其形如马鞍,故又称鞍面.

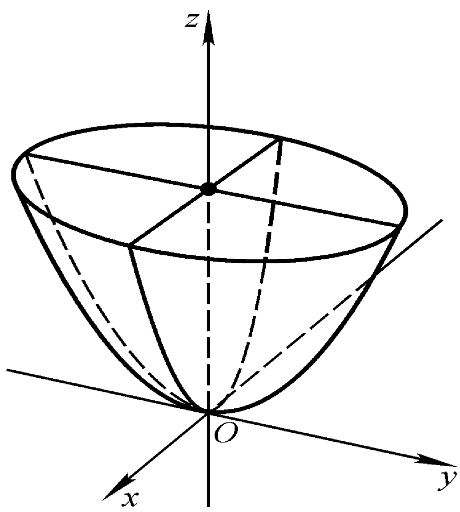


图 9

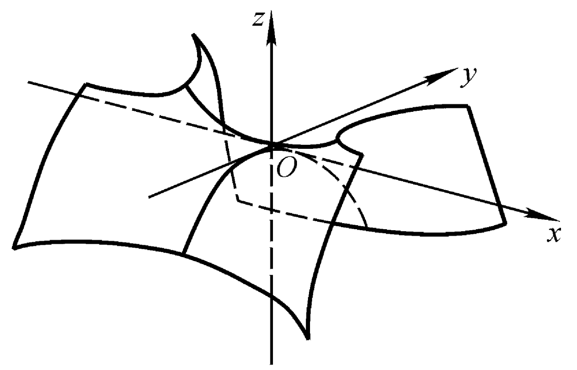


图 10

第三节　空间曲线及其方程

3.1　空间曲线的一般方程

任意一条空间直线可看作为两个平面的交线,而任意空间曲线可以看成两个曲面的交线.如果曲线 C 是由方程

$$F_1(x, y, z) = 0 \text{ 和 } F_2(x, y, z) = 0$$

表示的两个曲面的交线(图 1),则点 $M(x, y, z)$ 在曲线 C 上的充要条件是 M 的坐标满足方程组

$$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0. \end{cases} \tag{2.22}$$

方程组(2.22)叫做空间曲线 C 的一般方程.一般来说,要想像出两个曲面的交线,显然不是一件容易的事.这里讨论曲线的形状,只能作出示意图或大致轮廓,使用计算机可以画出曲面或曲线较精确的形状.

例 1　研究方程组

$$\begin{cases} z = \sqrt{a^2 - x^2 - y^2}, \\ \left(x - \dfrac{a}{2}\right)^2 + y^2 = \left(\dfrac{a}{2}\right)^2 \end{cases}$$

表示的曲线.

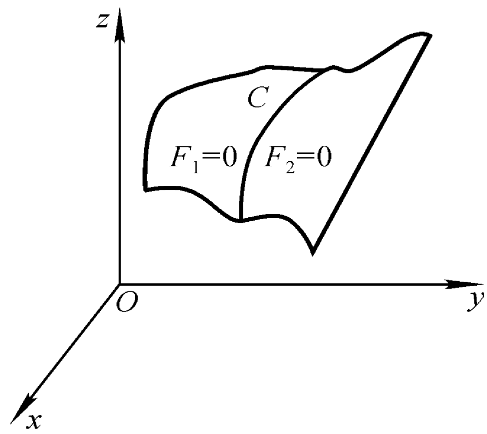


图 1

解　上面方程组中,第一个方程表示球心在坐标原点 O,半径为 a 的上半球面 .第二个方程表示圆柱面 .它的母线平行 z 轴,其准线是圆心在$(a\ 2,0)$的半径为 $a\ 2$ 的圆 .因而这个曲线通过$(a,0,0)$及$(0,0,a)$两点 .这个半球面与圆柱面的交线就是所求曲线(图 2) .

3.2　空间曲线的参数方程

如空间直线一样,空间曲线方程除了一般方程之外,还可以用参数 t 的三个函数给定:

$$
\begin{aligned}
x &= x(t),\\
y &= y(t),\\
z &= z(t).
\end{aligned}
\tag{2 23}
$$

该方程叫做空间曲线的参数方程 .任给 $t\in[a,b]$,可以确定曲线 C 上一点;随着 t 由 a 变到 b,动点 $M_t(x(t),y(t),z(t))$的轨迹就是曲线 C .

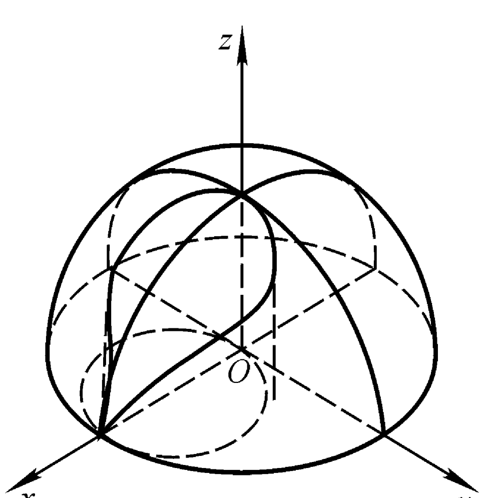


图 2

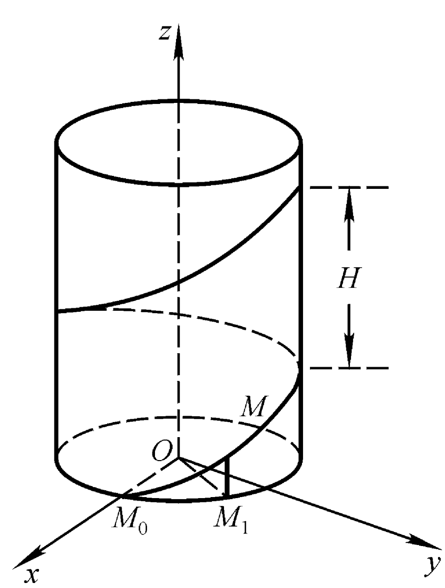


图 3

例 2　空间动点 M 由 $M_0(a,0,0)$出发,以等角速率 ω 绕 z 轴旋转作圆周运动,同时又以相等线速度 v 沿平行于 z 轴的正方向作直线上升运动.求动点 M 的运动轨迹.

解　取时间 t 为参数.设当 $t=0$ 时,开始于点 $M_0(a,0,0)$.经过时间 t 后,M_0 运动到 $M(x,y,z)$(图 3).M 点在 xOy 面投影为 M_1,从 x 轴正向到向量 $\overline{OM_1}$的转角 $\theta=\omega t$,而向量 $\overline{M_1M}=vt\boldsymbol{k}$,其中向量 $\boldsymbol{k}$ 为 z 轴方向单位向量,则向量

$$\overline{OM}=\overline{OM_1}+\overline{M_1M}=\boldsymbol{i}a\cos\omega t+\boldsymbol{j}a\sin\omega t+\boldsymbol{k}vt.$$

设向量 $\overline{OM}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$,则

$$x=a\cos\omega t,$$
$$y=a\sin\omega t,$$
$$z=vt.$$

该曲线称为螺旋线,它缠绕在圆柱面 $x^2+y^2=a^2$ 之上.$T=2\pi/\omega$ 是动点 M 的投影 M_1 的圆周运动的周期,当 M_1 运动一周,动点 M 在 z 轴方向上升的距离 $H=2\pi v/\omega$,称为螺距.

3.3　空间曲线在坐标面上的投影

一般空间曲线比较复杂,也很难想像和表示,空间曲线在各个坐标平面的投影有助于研究它的形状.在重积分、曲线积分和曲面积分的计算中,也往往需要确定一个空间曲线、曲面或立体在某一坐标面上的投影.

设曲线 C 的方程为

$$\begin{cases} F_1(x,y,z)=0, \\ F_2(x,y,z)=0. \end{cases} \tag{2.24}$$

由方程组(2.24)消去变量 z 后得到方程

$$G(x,y)=0. \tag{2.25}$$

由前面的讨论可知,方程(2.25)表示的曲面是母线平行于 z 轴的一个柱面,记为 A.如果点 $M(x,y,z)$的坐标满足方程组(2.24)时,则 M 的前两个坐标 x,y 必满足方程(2.25),这说明曲线 C 上的所有点都在柱面 A 上.以曲线 C 为准线、母线平行于 z 轴的柱面叫做曲线 C 关于 xOy 面的投影柱面,柱面 A 包含投影柱面.投影柱面与 xOy 面的交线叫做空间曲线 C 在 xOy 面上的投影曲线或简称投影,它被包含在柱面 A 与 xOy 面的交线之中,即方程为

$$\begin{cases} G(x,y)=0, \\ z=0 \end{cases}$$

的曲线包含曲线 C 在 xOy 面上的投影.类似地,可得到曲线 C 在 yOz 面或 zOx

面上的投影曲线方程 .

例 3　求曲线 C

$$\begin{cases} x^2 + z^2 = 1, \\ y^2 + z^2 = 1 \end{cases} \tag{2.26}$$

在 xOy 面上的投影方程 .

解　由(2.26)消去 z,得到母线平行于 z 轴的柱面 A

$$x^2 - y^2 = 0,$$

它包含方程为(2.26)的曲线.柱面 A 与 xOy 面的交线方程是

$$\begin{cases} x^2 - y^2 = 0, \\ z = 0 . \end{cases}$$

于是曲线 C 在 xOy 面的投影曲线方程为

$$\begin{cases} x + y = 0, \\ z = 0; \end{cases} \quad \text{或} \quad \begin{cases} x - y = 0, \\ z = 0 . \end{cases} \qquad (|x| \leq 1)$$

习　题　二

1 . 求过(1,1, - 1), (- 2, - 2,2)和(1, - 1,2)三点的平面方程 .

2 . 一平面过点(1,0, - 1)且平行于向量 $\boldsymbol{a} = (2,1,1)$和 $\boldsymbol{b} = (1, - 1,0)$,试求这平面方程 .

3 . 求满足下列条件的平面方程:

(1) 平行于 xOz 面且经过点(2, - 5,3);

(2) 通过 z 轴和点(- 3,1, - 2);

(3) 平行于 x 轴且经过两点(4,0, - 2)和(5,1,7) .

4 . 求过点 $M_0(2,9, - 6)$且与联结坐标原点 O 及点 M_0 的线段 OM_0 垂直的平面方程 .

5 . 指出下列各平面的特殊位置,并画出各平面:

(1) $x = 0$;　　(2) $3y - 1 = 0$;

(3) $2x - 3y - 6 = 0$;　　(4) $x - \sqrt{3}y = 0$;

(5) $y + z = 1$;　　(6) $x - 2z = 0$;

(7) $6x + 5y - z = 0$.

6 . 求三平面 $x + 3y + z = 1, 2x - y - z = 0, - x + 2y + 2z = 3$ 的交点 .

7 . 将下列曲线的一般方程化为参数方程:

(1) $\begin{cases} x^2 + y^2 + z^2 = 9, \\ y = x; \end{cases}$

(2) $\begin{cases} (x - 1)^2 + y^2 + (z + 1)^2 = 4, \\ z = 0 . \end{cases}$

8．求点$(1,2,1)$到平面 $x+2y+2z-10=0$ 的距离．

9．求过点$(4,-1,3)$且平行于直线$\frac{x-3}{2}=y=\frac{z-1}{5}$的直线方程．

10．求过两点$(3,-2,1)$和$(-1,0,2)$的直线方程．

11．求直线

$$\begin{cases} x+y+3z=0, \\ x-y-z=0, \end{cases}$$

与平面 $x-y-z+1=0$ 的夹角．

12．求过原点且与两直线

$$\begin{cases} x=1, \\ y=-1+t, \\ z=2+t. \end{cases} \qquad \frac{x+1}{1}=\frac{y+2}{2}=\frac{z-1}{1}$$

都平行的平面．

13．求平面 $2x-2y+z+5=0$ 与各坐标面的夹角的余弦．

14．求直线

$$\begin{cases} 2x-4y+z=0, \\ 3x-y-2z-9=0, \end{cases}$$

在平面 $4x-y+z=1$ 上的投影直线的方程．

15．过曲面 $z=4-x^2-y^2$ 上点 P 处的切平面平行于平面 $2x+2y+z-1=0$，求点 P 的坐标．

16．求过直线$\frac{x-1}{1}=\frac{y-2}{0}=\frac{z-3}{-1}$且平行于直线$\frac{x+2}{2}=\frac{y-1}{1}=\frac{z}{1}$的平面方程．

17．求直线$\frac{x-1}{1}=\frac{y}{1}=\frac{z-1}{-1}$在平面 $x-y+2z-1=0$ 上的投影直线的方程．

18．求点$(-1,2,0)$在平面 $x+2y-z+1=0$ 上的投影．

19．求点$(3,-1,2)$到直线

$$\begin{cases} x+y-z+1=0, \\ 2x-y+z-4=0, \end{cases}$$

的距离．

20．方程 $x^2+y^2+z^2-2x+4y+2z=0$ 表示什么曲线？

21．将 xOz 面上的抛物线 $z^2=5x$ 绕 x 轴旋转一周，求所生成的旋转曲面的方程．

22．将 xOy 面上的双曲线 $4x^2-9y^2=36$ 分别绕 x 轴及 y 轴旋转一周，求

所生成的旋转曲面的方程 .

23 . 分别求母线平行于 x 轴及 y 轴而且通过曲线

$$2x^2 + y^2 + z^2 = 16,$$
$$x^2 + z^2 - y^2 = 0,$$

的柱面方程 .

24 . 求球面 $x^2 + y^2 + z^2 = 9$ 与平面 $x + z = 1$ 的交线在 xOy 面上的投影的方程 .

25 . 一个立体由上半球面 $z = \sqrt{a^2 - x^2 - y^2}$, 柱面 $x^2 + y^2 - ax = 0$ 及平面 $z = 0$ 所围成, 求该立体在 xOy 面和 xOz 面上的投影 .

第三章　矩阵与行列式

第一节　矩阵及其运算

1.1　矩阵的概念

矩阵是线性代数的一个最基本的概念,矩阵的运算是线性代数的基本内容.矩阵的概念首先是由英国数学家 J.J.Sylvester(1814—1897)在 1850 年提出的.1850 年 A.Cayley(1821—1895)建立矩阵运算规则.从此,矩阵的理论逐渐成为数学的一个重要分支.矩阵把一组数,用一张表的形式联系在一起,视为一个整体,当作一个"量"来进行运算.它可以使大量相似的、重复的运算得到很大的简化,使问题的论述更简捷,更容易把握问题的整体和本质.在数学、自然科学、工程技术及生产实践中,有很多问题都可归结为矩阵的运算,可用矩阵的理论来处理.先看下面的例子.

例 1　城市 D 和 E 各有一个煤矿,一个发电厂和一条铁路.假设城市 D 和 E 每生产价值一元钱的煤,分别需要消耗 0.25 元和 0.20 元的电;为了把这一元钱的煤运出去,分别需要花 0.15 元和 0.25 元的运费.城市 D 和 E 每生产价值一元钱的电,都需要消耗 0.65 元的煤;为了运行电厂的辅助设备,都要消耗本身的 0.05 元的电;分别需要花 0.05 元和 0.15 元的运费.城市 D 和 E 每提供价值一元钱的铁路运输,都要消耗 0.55 元的煤,辅助设备都要消耗 0.10 元的电.城市 D 和 E 的煤矿、电厂和铁路局相互之间的消耗关系可以分别用表 1 与 2 表示如下:

表 1: D 城

	煤矿	电厂	铁路局
煤	0.00	0.65	0.55
电	0.25	0.05	0.10
运输	0.15	0.05	0

表 2: E 城

	煤矿	电厂	铁路局
煤	0.00	0.65	0.55
电	0.20	0.05	0.10
运输	0.25	0.15	0

城市 D 给出两个煤、电、运输量(单位万元)的生产计划 P_1 和 P_2,它们可以用表 3 表示.表 4 是执行生产计划 P_1 和 P_2 各自煤、电及运输的消耗量.

表 3

	P_1	P_2
煤产量	15	30
电产量	20	25
运输产量	30	20

表 4

	P_1	P_2
煤消耗量	c_{11}	c_{21}
电消耗量	c_{12}	c_{22}
运输消耗量	c_{13}	c_{23}

表 4 可由表 1 和表 3 计算得出,即表 4 中任意 c_{ij} 都可用表 1 和表 3 中的数据求得,如计划 P_2 的电消耗量

$$c_{22} = 0.25 \times 30 + 0.05 \times 25 + 0.10 \times 20 = 10.75,$$

即 c_{22} 等于向量$(0.25, 0.05, 0.10)$与向量$(30, 25, 20)$在 3 维欧氏几何空间 $\mathbf{R}^3$ 中做数量积.

从上面的例子可以看出:表 4 是表 1 和表 3 一种运算的结果,参加运算的对象是表,称之为矩阵.若把表 1 的每一行看成一个向量,若把表 3 的每一列看成一个向量,那么表 4 的第 l 行第 k 列的数是表 1 第 l 行向量与表 3 第 k 列向量(也看成行向量)的数量积.下面给出的矩阵及其运算的定义就是上面的表格及其运算的抽象.

定义 3.1　$m \times n$ 个实数 a_{ij} $(i = 1, 2, \cdots, m;\ j = 1, 2, \cdots, n)$排成 m 行 n 列的表

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \cdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

叫做 $\boldsymbol{m \times n}$ 矩阵,简记为 $\boldsymbol{A} = (a_{ij})$或 $\boldsymbol{A} = (a_{ij})_{m \times n}$,当 $m = n$ 时,称 $\boldsymbol{A}$ 为 $\boldsymbol{n}$ 阶矩阵,或 $\boldsymbol{n}$ 阶方阵.

当 a_{ij} 为实数时,称 $\boldsymbol{A}$ 为实数域 $\mathbf{R}$ 上的矩阵,简称实矩阵.一般地,当 a_{ij} 属于域 F 时,称 $\boldsymbol{A}$ 为域 F 上的矩阵,有关域的定义参看附录 A.1.若不特别说明,在本书中的矩阵总指实数域 $\mathbf{R}$ 上的矩阵.

按此定义,上面的表格 1, 2, 3, 4 分别对应矩阵

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0.65 & 0.55 \\ 0.25 & 0.05 & 0.1 \\ 0.15 & 0.05 & 0 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 0 & 0.65 & 0.55 \\ 0.20 & 0.05 & 0.10 \\ 0.25 & 0.15 & 0 \end{pmatrix},$$

$$\boldsymbol{C} = \begin{pmatrix} 15 & 30 \\ 20 & 25 \\ 30 & 20 \end{pmatrix}, \quad \boldsymbol{D} = (c_{ij})_{3\times 2}.$$

1.2　矩阵运算的性质

1. 矩阵的三种基本运算的定义

在定义矩阵运算之前,先定义两个矩阵 $\boldsymbol{A} = (a_{ij})_{m \times n}$ 和 $\boldsymbol{B} = (b_{ij})_{l \times k}$ 相等:

$\boldsymbol{A}=\boldsymbol{B}$当且仅当 $m=l, n=k, a_{ij}=b_{ij}(i=1,2,\cdots,m; j=1,2,\cdots,n)$,即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 完全一样.

定义 3.2 设 $\boldsymbol{A}=(a_{ij})_{m\times n}$和 $\boldsymbol{B}=(b_{ij})_{l\times k}$为两个矩阵.

(1) λ为任意数,定义 $\lambda\boldsymbol{A}=\boldsymbol{A}\lambda=(\lambda a_{ij})_{m\times n}$,称为数 λ 与矩阵 $\boldsymbol{A}$ 的数量乘积,简称数乘.

(2) 若 $m=l, n=k$,定义 $\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}$,称为矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的和.

(3) 若 $n=l$,定义 $\boldsymbol{AB}=\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)_{m\times k}$,称为矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的乘积.

当矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都为 1×1 矩阵时,矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和、乘积实际就是数的加法、乘法运算.从这个意义上看,矩阵的和、乘积是数的加法、乘法运算的某种推广.特别注意,只有当矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行数和列数都相等时,$\boldsymbol{A}+\boldsymbol{B}$ 才有定义;只有当矩阵 $\boldsymbol{A}$ 的列数等于矩阵 $\boldsymbol{B}$ 的行数时,$\boldsymbol{AB}$ 才有定义.

再来考察例 1,城市 D 和 E 的煤矿、电厂和铁路局每万元的产出的相互之间的消耗关系所对应的矩阵(简称消耗矩阵)等于

$$10\,000\boldsymbol{A}=\begin{pmatrix}0 & 6\,500 & 5\,500\\ 2\,500 & 500 & 1\,000\\ 1\,500 & 500 & 0\end{pmatrix},\quad 10\,000\boldsymbol{B}=\begin{pmatrix}0 & 6\,500 & 5\,500\\ 2\,000 & 500 & 1\,000\\ 2\,500 & 1\,500 & 0\end{pmatrix}$$

城市 D 和 E 的煤矿、电厂和铁路局每万元的产出的平均消耗矩阵等于

$$\frac{1}{2}\times10\,000(\boldsymbol{A}+\boldsymbol{B})=\begin{pmatrix}0 & 6\,500 & 5\,500\\ 2\,250 & 500 & 1\,000\\ 2\,000 & 1\,000 & 0\end{pmatrix}$$

城市 D 给出两个煤、电、运输量(单位万元)生产计划 P_1 和 P_2 中煤、电和运输消耗量所对应的矩阵等于

$$\boldsymbol{D}=(c_{ij})_{3\times2}=\boldsymbol{AC}=\begin{pmatrix}0 & 0.65 & 0.55\\ 0.25 & 0.05 & 0.1\\ 0.15 & 0.05 & 0\end{pmatrix}\begin{pmatrix}15 & 30\\ 20 & 25\\ 30 & 20\end{pmatrix}=\begin{pmatrix}29.5 & 27.25\\ 7.75 & 10.75\\ 3.25 & 5.75\end{pmatrix}.$$

2. 矩阵运算的性质

性质 3.1 任给矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{B}=(b_{ij})_{m\times n}, \boldsymbol{C}=(c_{ij})_{m\times n}$,则下列各式成立

(1) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$;

(2) $\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$;

(3) $\boldsymbol{O}+\boldsymbol{A}=\boldsymbol{A}$,其中 $\boldsymbol{O}=(o_{ij})_{m\times n}, o_{ij}=0, i=1,2,\cdots,m, j=1,\cdots,n$, $\boldsymbol{O}$ 称为零矩阵;

(4) $\boldsymbol{A}+(-1)\boldsymbol{A}=\boldsymbol{O}$.

$(-1)\boldsymbol{A}$ 为 $\boldsymbol{A}$ 的负矩阵，一般记 $(-1)\boldsymbol{A}$ 为 $-\boldsymbol{A}$.

性质 3.2　任给数 $\lambda,\mu\in\mathbf{R}$ 及 $\boldsymbol{A}=(a_{ij})_{m\times n}$，$\boldsymbol{B}=(b_{ij})_{m\times n}$，则有下列各式成立：

(1) $(\lambda+\mu)\boldsymbol{A}=\lambda\boldsymbol{A}+\mu\boldsymbol{A}$;

(2) $\lambda(\boldsymbol{A}+\boldsymbol{B})=\lambda\boldsymbol{A}+\lambda\boldsymbol{B}$;

(3) $\lambda(\mu\boldsymbol{A})=(\lambda\mu)\boldsymbol{A}$;

(4) $1\boldsymbol{A}=\boldsymbol{A}$，$(-1)\boldsymbol{A}=-\boldsymbol{A}$;

(5) $\lambda\boldsymbol{A}=\boldsymbol{O}$ 当且仅当 $\lambda=0$ 或 $\boldsymbol{A}=\boldsymbol{O}$.

性质 3.1，3.2 很易证明，证明略.

对照向量空间的定义，容易看出所有实数域 $\mathbf{R}$ 上 $m\times n$ 矩阵全体构成 $\mathbf{R}$ 上一个向量空间.

性质 3.3　设 $\boldsymbol{A}=(a_{ij})_{m\times l}$，$\boldsymbol{B}=(b_{ij})_{l\times k}$，$\boldsymbol{C}=(c_{ij})_{k\times n}$，$\boldsymbol{D}=(d_{ij})_{l\times k}$，则下列性质也成立：

(1) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{D})=\boldsymbol{AB}+\boldsymbol{AD}$，$(\boldsymbol{B}+\boldsymbol{D})\boldsymbol{C}=\boldsymbol{BC}+\boldsymbol{DC}$;

(2) $\boldsymbol{A}(\boldsymbol{BC})=(\boldsymbol{AB})\boldsymbol{C}$.

证　(1)的证明可直接由定义验证.

下面证明(2). 设 $\boldsymbol{A}(\boldsymbol{BC})=\boldsymbol{T}=(t_{ij})_{m\times n}$，$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{S}=(s_{ij})_{m\times n}$，$\boldsymbol{BC}=\boldsymbol{V}=(v_{ij})_{l\times n}$，$\boldsymbol{AB}=\boldsymbol{U}=(u_{ij})_{m\times k}$，那么

$$u_{ij}=\sum_{r=1}^{l}a_{ir}b_{rj},\quad v_{ij}=\sum_{s=1}^{k}b_{is}c_{sj}.$$

因此

$$s_{ij}=\sum_{s=1}^{k}u_{is}c_{sj}=\sum_{s=1}^{k}\left(\sum_{r=1}^{l}a_{ir}b_{rs}\right)c_{sj}=\sum_{s=1}^{k}\sum_{r=1}^{l}a_{ir}b_{rs}c_{sj},$$

$$t_{ij}=\sum_{r=1}^{l}a_{ir}v_{rj}=\sum_{r=1}^{l}a_{ir}\sum_{s=1}^{k}b_{rs}c_{sj}=\sum_{r=1}^{l}\sum_{s=1}^{k}a_{ir}b_{rs}c_{sj},$$

即 $s_{ij}=t_{ij}$，$\boldsymbol{S}=\boldsymbol{T}$. 所以(2)成立. 这就是说矩阵乘法适合结合律. 因此矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 的乘积可写成 $\boldsymbol{ABC}$ 而不加括号. □

(* 当 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 都为 n 阶方阵，性质 3.3 中(1)与(2)当然也成立. 故

$$M(\mathbf{R})_{n\times n}=\{\boldsymbol{A}=(a_{ij})_{n\times n}\mid a_{ij}\in\mathbf{R}\}$$

关于矩阵的和及矩阵的乘积为一环，一般称该环为实数域上的 n 阶矩阵环. 参考附录 A.1)

令

$$\boldsymbol{I}=(\delta_{ij})_{n\times n}=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\cdots&\cdots&&\cdots\\0&0&0&1\end{pmatrix},\quad \delta_{ij}=\begin{cases}1, & i=j,\\0, & i\neq j.\end{cases}$$

$\boldsymbol{I}$ 称为 n 阶单位矩阵,记作 $\boldsymbol{I}$ 或 $\boldsymbol{I}_n$,有时也记为 $\boldsymbol{E}$.对任意矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,有

$$\boldsymbol{I}_m\boldsymbol{A}=\left(\sum_{k=1}^{n}\delta_{ik}a_{kj}\right)_{m\times n}=(a_{ij})_{m\times n}=\boldsymbol{A},$$

$$\boldsymbol{A}\boldsymbol{I}_n=\left(\sum_{k=1}^{n}a_{ik}\delta_{kj}\right)_{m\times n}=(a_{ij})_{m\times n}=\boldsymbol{A}.$$

3. 矩阵乘积和幂的例子

对于数的加法和乘积上面的性质 3.1,3.3 自然成立,但是矩阵乘积与数的乘法性质不尽相同.首先,矩阵 $\boldsymbol{A},\boldsymbol{B}$ 的乘积 $\boldsymbol{AB}$ 有定义时,$\boldsymbol{BA}$ 却不一定有定义.这是因为矩阵乘积要求左面矩阵的列数等于右面矩阵的行数.对于方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$,即使 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 都有意义,它们也不一定相等.也就是说乘法的交换律不成立,即 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 不一定相等.

例 2　已知

$$\boldsymbol{A}=\begin{pmatrix}1&2\\3&4\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

计算 $\boldsymbol{AB},\boldsymbol{BA}$.

解

$$\boldsymbol{AB}=\begin{pmatrix}1&2\\3&4\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&3\\3&7\end{pmatrix}.\quad \boldsymbol{BA}=\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&2\\3&4\end{pmatrix}=\begin{pmatrix}4&6\\3&4\end{pmatrix}.$$

但如果

$$\boldsymbol{A}=\begin{pmatrix}7&-12\\-4&7\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}3&0\\0&3\end{pmatrix},$$

有

$$\boldsymbol{AB}=\begin{pmatrix}7&-12\\-4&7\end{pmatrix}\begin{pmatrix}3&0\\0&3\end{pmatrix}=\begin{pmatrix}21&-36\\-12&21\end{pmatrix},$$

$$\boldsymbol{BA}=\begin{pmatrix}3&0\\0&3\end{pmatrix}\begin{pmatrix}7&-12\\-4&7\end{pmatrix}=\begin{pmatrix}21&-36\\-12&21\end{pmatrix}.$$

由于 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相乘时,其结果与因子顺序有关,当用 $\boldsymbol{A}$ 乘 $\boldsymbol{B}$ 时,必须说明 $\boldsymbol{A}$ 是左乘 $\boldsymbol{B}$ 或右乘 $\boldsymbol{B}$.此外,更应注意的是,两个非零矩阵的乘积可能是零矩阵,即当 $\boldsymbol{A}\neq\boldsymbol{O}$ 与 $\boldsymbol{B}\neq\boldsymbol{O}$,$\boldsymbol{AB}$ 可能是零矩阵.例如

$$\begin{pmatrix}1&1\\-1&-1\end{pmatrix}\begin{pmatrix}-1&1\\1&-1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=\boldsymbol{O}.$$

$$\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ b_{31} & b_{32} & b_{32} \end{pmatrix} = \boldsymbol{O}.$$

还应注意：$\boldsymbol{AB} = \boldsymbol{AC}$，当 $\boldsymbol{A} \neq \boldsymbol{O}$ 时，不一定能得到 $\boldsymbol{B} = \boldsymbol{C}$，即矩阵乘法不一定适合消去律.

例 3　已知

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

计算 $\boldsymbol{AAA}$.

解

$$\boldsymbol{AAA} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{O}.$$

设 $\boldsymbol{A}$ 为 n 阶方阵，定义 $\boldsymbol{A}$ 的 n 次幂为 $\boldsymbol{A}^n = \underbrace{\boldsymbol{A}\cdots\boldsymbol{A}}_{n}$.令

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$$

易验证

$$\boldsymbol{A}^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & 0 \end{pmatrix},\cdots,\boldsymbol{A}^{n-1} = \begin{pmatrix} 0 & & & & 1 \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix},\ \boldsymbol{A}^n = \boldsymbol{O}$$

因此，当 $k \geq n$ 时，$A^k = \boldsymbol{O}$.

4．转置矩阵

把矩阵 $\boldsymbol{A} = (a_{ij})_{m\times l}$ 的行换成同序数的列得到的一个新的矩阵，叫做 $\boldsymbol{A}$ 的转置矩阵，记作 $\boldsymbol{A}^{\mathrm{T}}$ 或 $\boldsymbol{A}'$.例如，矩阵

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

$\boldsymbol{A}$ 的转置矩阵为

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

若 $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$,则称矩阵 $\boldsymbol{A}$ 为对称阵.如

$$\begin{pmatrix} 1 & 5 & 6 \\ 5 & 2 & 7 \\ 6 & 7 & 3 \end{pmatrix}, \quad \begin{pmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{pmatrix},$$

都为对称阵.

性质 3.4 假设下面的矩阵运算都有定义,则矩阵的转置满足下述运算规律:

(1) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$;

(2) $(\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$;

(3) $(\lambda\boldsymbol{A})^{\mathrm{T}} = \lambda\boldsymbol{A}^{\mathrm{T}}$;

(4) $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

证 (1)～(3)用定义可直接验证.下面只证明(4).设 $\boldsymbol{A} = (a_{ij})_{m\times s}$, $\boldsymbol{B} = (b_{ij})_{s\times n}$,记 $\boldsymbol{AB} = \boldsymbol{C} = (c_{ij})_{m\times n}$, $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{D} = (d_{ij})_{n\times m}$,于是有 $c_{ji} = \sum_{k=1}^{s} a_{jk}b_{ki}$.矩阵 $\boldsymbol{B}^{\mathrm{T}}$ 的第 i 行为 $(b_{1i}, b_{2i}, \cdots, b_{si})$, $\boldsymbol{A}^{\mathrm{T}}$ 的第 j 列为 $(a_{j1}, a_{j2}, \cdots, a_{js})^{\mathrm{T}}$,因此

$$d_{ij} = \sum_{k=1}^{s} b_{ki}a_{jk} = \sum_{k=1}^{s} a_{jk}b_{ki}.$$

这说明 $d_{ij} = c_{ji}\,(i = 1,2,\cdots,n;\ j = 1,2,\cdots,m)$,即 $\boldsymbol{D} = \boldsymbol{C}^{\mathrm{T}}$,亦即

$$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}. \square$$

例 4 已知

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 7 & -1 \\ 4 & 2 & 3 \\ 2 & 0 & 1 \end{pmatrix},$$

求 $(\boldsymbol{AB})^{\mathrm{T}}$.

解 1 因为

$$\boldsymbol{AB} = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 7 & -1 \\ 4 & 2 & 3 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 14 & -3 \\ 17 & 13 & 10 \end{pmatrix},$$

所以

$$(\boldsymbol{AB})^{\mathrm{T}} = \begin{pmatrix} 0 & 17 \\ 14 & 13 \\ -3 & 10 \end{pmatrix}.$$

解 2

$$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \begin{pmatrix} 1 & 4 & 2 \\ 7 & 2 & 0 \\ -1 & 3 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 17 \\ 14 & 13 \\ -3 & 10 \end{pmatrix}.$$

第二节　分块矩阵的乘法与初等变换

2.1　分块矩阵

1. 分块矩阵的定义

将一个矩阵分成一些小矩阵,即所谓矩阵的子块,原矩阵就构成分块矩阵.对于行数和列数较多的矩阵,用分块矩阵的技巧常常会收到事半功倍的效果,可解决许多困难的矩阵问题.由于不同的需要,同一个矩阵可以用不同的分块方法,构成不同的分块矩阵.

设

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}.$$

可对 A 用不同的方法分块构成下列不同分块阵:

(1) $$\boldsymbol{A} = \left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right) = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix},$$ 其中

$$\boldsymbol{A}_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \boldsymbol{A}_{12} = \begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix},$$

$$\boldsymbol{A}_{21} = (a_{31} \quad a_{32}), \quad \boldsymbol{A}_{22} = (a_{33} \quad a_{34}).$$

(2) $$\boldsymbol{A} = \left(\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right) = (\boldsymbol{A}_1 \quad \boldsymbol{A}_2 \quad \boldsymbol{A}_3 \quad \boldsymbol{A}_4),$$

其中

$$\boldsymbol{A}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \end{pmatrix}, \quad i = 1,2,3,4.$$

(3) $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34}\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_2\\ \boldsymbol{A}_3\end{pmatrix}$,

其中 $\boldsymbol{A}_j=(a_{j1}\quad a_{j2}\quad a_{j3}\quad a_{j4})$, $j=1,2,3$.

下面讨论分块矩阵的加法、数乘和乘法.

2. 分块矩阵的运算

(1) 加法

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$, $\boldsymbol{B}=(b_{ij})_{m\times n}$将 A 和 B 采用相同的分块方法,用若干条纵线和横线分成许多小矩阵,即分成一些子块如下:

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s}\\ \cdots & \cdots & & \cdots\\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1s}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2s}\\ \cdots & \cdots & & \cdots\\ \boldsymbol{B}_{r1} & \boldsymbol{B}_{r2} & \cdots & \boldsymbol{B}_{rs}\end{pmatrix},$$

其中子块 $\boldsymbol{A}_{ij}$与子块 $\boldsymbol{B}_{ij}$, $(i=1,2,\cdots,r,\ j=1,2,\cdots,s)$的行与列数均相同.这时,$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相加即是它们对应的子块相加.

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12} & \cdots & \boldsymbol{A}_{1s}+\boldsymbol{B}_{1s}\\ \boldsymbol{A}_{21}+\boldsymbol{B}_{21} & \boldsymbol{A}_{22}+\boldsymbol{B}_{22} & \cdots & \boldsymbol{A}_{2s}+\boldsymbol{B}_{2s}\\ \cdots & \cdots & & \cdots\\ \boldsymbol{A}_{r1}+\boldsymbol{B}_{r1} & \boldsymbol{A}_{r2}+\boldsymbol{B}_{r2} & \cdots & \boldsymbol{A}_{rs}+\boldsymbol{B}_{rs}\end{pmatrix}$$

(2) 数乘

数 k 乘分块阵 $\boldsymbol{A}$,等于用 k 乘 $\boldsymbol{A}$ 的每个子块,即

$$k\boldsymbol{A}=\begin{pmatrix}k\boldsymbol{A}_{11} & k\boldsymbol{A}_{12} & \cdots & k\boldsymbol{A}_{1s}\\ k\boldsymbol{A}_{21} & k\boldsymbol{A}_{22} & \cdots & k\boldsymbol{A}_{2s}\\ \cdots & \cdots & & \cdots\\ k\boldsymbol{A}_{r1} & k\boldsymbol{A}_{r2} & \cdots & k\boldsymbol{A}_{rs}\end{pmatrix}.$$

(3) 转置

$$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{A}_{11}^{\mathrm{T}} & \boldsymbol{A}_{21}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r1}^{\mathrm{T}}\\ \boldsymbol{A}_{12}^{\mathrm{T}} & \boldsymbol{A}_{22}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{r2}^{\mathrm{T}}\\ \cdots & \cdots & & \cdots\\ \boldsymbol{A}_{1s}^{\mathrm{T}} & \boldsymbol{A}_{2s}^{\mathrm{T}} & \cdots & \boldsymbol{A}_{rs}^{\mathrm{T}}\end{pmatrix}.$$

(4) 乘积

设 $\boldsymbol{A}=(a_{ij})_{m\times n}$, $\boldsymbol{B}=(b_{ij})_{n\times p}$,将 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 用若干条纵线和横线分成许多个小矩阵,即分成一些子块如下:

$$A = \begin{array}{c} \begin{array}{cccc} n_1 & n_2 & \cdots & n_s \end{array} \\ \left(\begin{array}{cccc} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \cdots & \cdots & & \cdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{array}\right) \end{array} \begin{array}{c} \\ m_1 \\ m_2 \\ \cdots \\ m_r \end{array}, \quad B = \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & \cdots & p_t \end{array} \\ \left(\begin{array}{cccc} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \cdots & \cdots & & \cdots \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{array}\right) \end{array} \begin{array}{c} \\ n_1 \\ n_2 \\ \cdots \\ n_s \end{array},$$

$$m_1 + m_2 + \cdots + m_r = m,$$

$$n_1 + n_2 + \cdots + n_s = n,$$

$$p_1 + p_2 + \cdots + p_t = p.$$

矩阵分块时,只作一点限制,对 A 的列的分法必须与对 B 的行的分法一致.矩阵右边的数 $m_1, m_2, \cdots, m_r, n_1, n_2, \cdots, n_s$ 分别表示 A 与 B 子块的行数.矩阵上面的数 $n_1, n_2, \cdots, n_s, p_1, p_2, \cdots, p_t$ 分别表示 A 与 B 子块的列数.这样分块以后,形式地把每个子块看成一个元素,从而 A 就形式地成为一个 $r\times s$ 矩阵, B 就形式地成为一个 $s\times t$ 矩阵.按矩阵的乘法规则,求出 A 和 B 的所谓形式乘积 C 如下:

$$C = \begin{array}{c} \begin{array}{cccc} p_1 & p_2 & \cdots & p_t \end{array} \\ \left(\begin{array}{cccc} C_{11} & C_{12} & \cdots & C_{1t} \\ C_{21} & C_{22} & \cdots & C_{2t} \\ \cdots & \cdots & & \cdots \\ C_{r1} & C_{r2} & \cdots & C_{rt} \end{array}\right) \end{array} \begin{array}{c} \\ m_1 \\ m_2 \\ \cdots \\ m_r \end{array}$$

其中 $C_{hk} = A_{h1}B_{1k} + A_{h2}B_{2k} + \cdots + A_{hs}B_{sk}(h = 1,2,\cdots, r; k = 1,2,\cdots, t)$.在矩阵分块时,由于已有 A 的列的分法必须与对 B 的行的分法一致的限制,所以这里的等式不仅是形式上的,而且子块 A_{hi} 与 B_{ik} 作为矩阵确实可以相乘,其积都是 $m_h\times p_k$ 矩阵$(i = 1,2,\cdots, s)$,故又可以相加,其和 C_{hk} 自然就是一个 $m_h\times p_k$ 矩阵,亦即这里的等式确实是矩阵的等式.需要特别注意,在作 AB 的分块乘法运算时,必须用 A 的子块左乘 B 的子块.从定义可以直接验证这些分块运算规则的正确性,证明略.

3. 矩阵分块运算的例子

例 1　设

$$A = \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 1 & 2 & 1 \\ 0 & 1 & 0 \end{array}\right) = \begin{pmatrix} I & O \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \left(\begin{array}{cc} 1 & 2 \\ 3 & 4 \\ \hline 5 & 6 \end{array}\right) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

用分块乘法计算 AB.

解

$$AB=\begin{pmatrix} I & O \\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix} B_1 \\ B_2\end{pmatrix}=\begin{pmatrix} IB_1+OB_2 \\ A_{21}B_1+A_{22}B_2\end{pmatrix}=\begin{pmatrix} B_1 \\ A_{21}B_1+A_{22}B_2\end{pmatrix}.$$

而

$$A_{21}B_1+A_{22}B_2=\begin{pmatrix}1&2\\0&1\end{pmatrix}\begin{pmatrix}1&2\\3&4\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}(5\quad 6)=\begin{pmatrix}7&10\\3&4\end{pmatrix}+\begin{pmatrix}5&6\\0&0\end{pmatrix}=\begin{pmatrix}12&16\\3&4\end{pmatrix}.$$

$$AB=\begin{pmatrix} B_1 \\ A_{21}B_1+A_{22}B_2\end{pmatrix}=\begin{pmatrix}1&2\\3&4\\12&16\\3&4\end{pmatrix}.$$

例 2 设 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times 1}$, $C=(c_{ij})_{1\times m}$. 计算 AB, CA.

解 对 A, B 进行如下分块:

$$\left(\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{array}\right)=(A_{11}\quad A_{12}\quad \cdots\quad A_{1n}),\quad \left(\begin{array}{c} b_{11} \\ \hline b_{21} \\ \hline \cdots \\ \hline b_{n1}\end{array}\right)=\begin{pmatrix} B_{11} \\ B_{21} \\ \cdots \\ B_{n1}\end{pmatrix}.$$

$$AB=A_{11}B_{11}+A_{12}B_{21}+\cdots+A_{1n}B_{n1}=\begin{pmatrix}a_{11}\\a_{21}\\\cdots\\a_{m1}\end{pmatrix}b_{11}+\begin{pmatrix}a_{12}\\a_{22}\\\cdots\\a_{m2}\end{pmatrix}b_{21}+\cdots+\begin{pmatrix}a_{1n}\\a_{2n}\\\cdots\\a_{mn}\end{pmatrix}b_{n1}.$$

当矩阵 B 为列向量, 若把矩阵 A 的 n 列看成是 n 个 m 维的列向量, 即 A_{11}, A_{12}, $\cdots$, $A_{1n}\in\mathbf{R}^m$, 则 AB 是一个可由这 n 个 m 维的列向量线性表示的列向量, 即

$$AB=b_{11}A_{11}+b_{21}A_{12}+\cdots+b_{n1}A_{1n}.$$

对 A 和 C 进行如下分块:

$$\left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline a_{21} & a_{22} & \cdots & a_{2n} \\ \hline \cdots & \cdots & & \cdots \\ \hline a_{m1} & a_{m2} & \cdots & a_{mn}\end{array}\right)=\begin{pmatrix} A_{11} \\ A_{21} \\ \cdots \\ A_{m1}\end{pmatrix},$$

$$(c_{11}\mid c_{12}\mid \cdots \mid c_{1m})=(C_{11}\quad C_{12}\quad \cdots\quad C_{1m}).$$

$$\begin{aligned} CA &= C_{11}A_{11}+C_{12}A_{21}+\cdots+C_{1m}A_{m1} \\ &= c_{11}(a_{11},a_{12},\cdots,a_{1n})+\cdots+c_{1m}(a_{m1},a_{m2},\cdots,a_{mn}).\end{aligned}$$

当矩阵 B 为行向量, 若把矩阵 A 的 m 行看成是 m 个 n 维的行向量, 即 A_{11},

$A_{21}, \cdots, A_{m1} \in \mathbf{R}^n$,则 CA 是一个可由这 m 个 n 维的行向量线性表示的行向量,即

$$CA = c_{11}A_{11} + c_{12}A_{21} + \cdots + c_{1m}A_{m1}.$$

对例 2 的情形而言,两个矩阵的乘积可应用向量及其运算表示.设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times l}$.对 A 和 B 进行如下分块:把 A 整体看成一块,把 B 的每一列看成一个子块,作 AB 的分块乘法:

$$AB = A\left(\begin{array}{c|c|c|c} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & b_{2l} \\ \cdots & \cdots & \ddots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nl} \end{array}\right) = A(B_{11} \quad B_{12} \quad \cdots \quad B_{1l})$$

$$= (AB_{11} \quad AB_{12} \quad \cdots \quad AB_{1l}).$$

上式说明,AB 的第 i 列等于 AB_{1i}, $i = 1, 2, \cdots, l$.由例 2 知,若把矩阵 A 的 n 列看成是 n 个 m 维列向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$,则 AB 的第 i 列作为一个列向量 γ_i,有

$$\gamma_i = b_{1i}\alpha_1 + b_{2i}\alpha_2 + \cdots + b_{ni}\alpha_n,$$

即 AB 的每一列作为一个列向量,都是 A 的列向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合.同理,AB 的每一行作为一个行向量,都是 B 的行向量的线性组合.

2.2　初等变换与初等矩阵

矩阵与线性方程组及其解法有密切关系.处理矩阵问题常用的一种基本方法,即对矩阵进行所谓的初等变换,也是从解线性方程组的消元法(Gauss 消元法)演变而来的.在算术中,每个整数可分解为若干个素数的乘积.自然会想到,矩阵是否也能进行类似的分解,即把一个复杂的矩阵分解为若干个简单的矩阵的乘积,以便更加深入地认识矩阵的性质.

1. 初等矩阵与矩阵的初等变换

定义 3.3　矩阵的下面三种变换,统称为矩阵的初等行(列)变换;

(1) 用任意不为零的数 k 去乘矩阵的第 i 行(列);

(2) 把矩阵的第 i 行(列)乘 k 加于第 j 行(列) $i \neq j$,其中 k 为任意实数;

(3) 互换矩阵的 i, j 两行(列).

上面三种变换依次称为倍乘变换、倍加变换及对换变换.

最简单的方阵是单位矩阵 I 和零矩阵 O.其次,比较简单的方阵就是对单位矩阵 I 做一次初等变换而得到的矩阵,这样的矩阵称为初等矩阵,共有以下三种类型初等矩阵:

(1) 初等倍乘矩阵——用不为零的数 k 去乘单位矩阵的第 i 行(列)而得到的矩阵

$$\boldsymbol{E}(kr_i)=\boldsymbol{E}(kc_i)=\begin{array}{c}\begin{pmatrix}1 & & & & 0\\ & \ddots & & & \\ & & k & & \\ & & & \ddots & \\ 0 & & & & 1\end{pmatrix}\begin{matrix}\\ \\ \leftarrow r_i;\\ \\ \\ \end{matrix}\\ \begin{matrix}\quad\quad\quad c_i\quad\quad\quad\end{matrix}\end{array}$$

(2) **初等倍加矩阵——把单位矩阵的第** i **行乘非零的数** k **加于第** j **行** $i\neq j$, **或把单位矩阵的第** j **列乘** k **加于第** i **列** $i\neq j$, **得到的矩阵**

$$\boldsymbol{E}(c_i+kc_j)=\boldsymbol{E}(r_j+kr_i)=\begin{array}{c}\begin{pmatrix}1 & & & & & & 0\\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \cdots & \ddots & & & \\ & & k & \cdots & 1 & & \\ & & & & & \ddots & \\ 0 & & & & & & 1\end{pmatrix}\begin{matrix}\\ \\ \leftarrow r_i\\ \\ \leftarrow r_j\\ \\ \\ \end{matrix};\\ \begin{matrix}\quad\quad c_i\quad\quad\quad c_j\quad\quad\end{matrix}\end{array}$$

(3) **初等对换矩阵——互换单位矩阵的** i, j **两行(列)而得到的矩阵**

$$\boldsymbol{E}(r_j\backslash r_i)=\boldsymbol{E}(c_j\backslash c_i)=\begin{array}{c}\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & & \\ & & \cdots & \ddots & \cdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}\begin{matrix}\\ \\ \leftarrow r_i\\ \\ \leftarrow r_j\\ \\ \\ \end{matrix}.\\ \begin{matrix}\quad\quad c_i\quad\quad\quad c_j\quad\quad\end{matrix}\end{array}$$

例 3　**设** $\boldsymbol{A}=(a_{ij})_{3\times 4}$, **分别用适当的初等矩阵左乘或右乘** $\boldsymbol{A}$ **并观察其结果**.

解

$$\boldsymbol{E}(kr_2)\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & k & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\end{pmatrix}=\begin{pmatrix}a_{11} & a_{12} & a_{13} & a_{14}\\ ka_{21} & ka_{22} & ka_{23} & ka_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\end{pmatrix}.$$

$$\boldsymbol{AE}(kc_1)=\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{pmatrix}\begin{pmatrix}k&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}=\begin{pmatrix}ka_{11}&a_{12}&a_{13}&a_{14}\\ka_{21}&a_{22}&a_{23}&a_{24}\\ka_{31}&a_{32}&a_{33}&a_{34}\end{pmatrix}.$$

$$\boldsymbol{E}(r_3+kr_1)\boldsymbol{A}=\begin{pmatrix}1&0&0\\0&1&0\\k&0&1\end{pmatrix}\boldsymbol{A}$$

$$=\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}+ka_{11}&a_{32}+ka_{12}&a_{33}+ka_{13}&a_{34}+ka_{14}\end{pmatrix}.$$

$$\boldsymbol{AE}(c_1+kc_4)=\boldsymbol{A}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\k&0&0&1\end{pmatrix}=\begin{pmatrix}a_{11}+ka_{14}&a_{12}&a_{13}&a_{14}\\a_{21}+ka_{24}&a_{22}&a_{23}&a_{24}\\a_{31}+ka_{34}&a_{32}&a_{33}&a_{34}\end{pmatrix}$$

$$\boldsymbol{AE}(c_2\backslash\ c_4)=\boldsymbol{A}\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{pmatrix}=\begin{pmatrix}a_{11}&a_{14}&a_{13}&a_{12}\\a_{21}&a_{24}&a_{23}&a_{22}\\a_{31}&a_{34}&a_{33}&a_{32}\end{pmatrix}.$$

通过观察,提出下面定理 .

定理 3.1　设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵 .对 $\boldsymbol{A}$ 进行某种初等行的变换,恰等于用同种初等矩阵去左乘 $\boldsymbol{A}$,即对单位矩阵做一次相同的初等行变换后再左乘 $\boldsymbol{A}$;对 $\boldsymbol{A}$ 进行某种初等列变换,恰等于用同种的初等矩阵去右乘 $\boldsymbol{A}$,即对单位矩阵做一次相同的初等列变换后再右乘 $\boldsymbol{A}$.

*证　这里仅对倍加变换情况给出证明 .其他情况,留作练习 .把 $\boldsymbol{I}_m$ 按行分块, $\boldsymbol{I}_n$ 按列分块,且分别记为:

$$\boldsymbol{I}_m=\begin{pmatrix}\boldsymbol{e}_1\\\boldsymbol{e}_2\\\cdots\\\boldsymbol{e}_m\end{pmatrix},\quad \boldsymbol{I}_n=(\boldsymbol{f}_1\quad\boldsymbol{f}_2\quad\cdots\quad\boldsymbol{f}_n).$$

初等倍加矩阵可分别记为:

$$\boldsymbol{E}(r_j+kr_i)=\begin{pmatrix}\boldsymbol{e}_1\\ \cdots\\ \boldsymbol{e}_i\\ \cdots\\ \boldsymbol{e}_j+k\boldsymbol{e}_i\\ \cdots\\ \boldsymbol{e}_m\end{pmatrix},\quad \boldsymbol{E}(c_j+kc_i)=(\boldsymbol{f}_1\ \cdots\ \boldsymbol{f}_i\ \cdots\ \boldsymbol{f}_j+k\boldsymbol{f}_i\ \cdots\ \boldsymbol{f}_n).$$

故知

$$\boldsymbol{E}(r_j+kr_i)\boldsymbol{A}=\begin{pmatrix}\boldsymbol{e}_1\boldsymbol{A}\\ \cdots\\ \boldsymbol{e}_i\boldsymbol{A}\\ \cdots\\ \boldsymbol{e}_j\boldsymbol{A}+k\boldsymbol{e}_i\boldsymbol{A}\\ \cdots\\ \boldsymbol{e}_m\boldsymbol{A}\end{pmatrix},$$

$$\boldsymbol{A}\boldsymbol{E}(c_j+kc_i)=(\boldsymbol{A}\boldsymbol{f}_1\ \cdots\ \boldsymbol{A}\boldsymbol{f}_i\ \cdots\ \boldsymbol{A}\boldsymbol{f}_j+k\boldsymbol{A}\boldsymbol{f}_i\ \cdots\ \boldsymbol{A}\boldsymbol{f}_n).$$

易验证: $\boldsymbol{e}_k\boldsymbol{A}$ 等于 $\boldsymbol{A}$ 的第 k 行, $\boldsymbol{A}\boldsymbol{f}_j$ 等于 $\boldsymbol{A}$ 的第 j 列, $k=1,2,\cdots,m$, $j=1,2,\cdots,n$.由此可知, $\boldsymbol{E}(r_j+kr_i)\boldsymbol{A}$ 恰好是把 $\boldsymbol{A}$ 的第 i 行乘 k 加于第 j 行而得的矩阵; $\boldsymbol{A}\boldsymbol{E}(c_j+kc_i)$恰好是把 A 的第 i 列乘 k 加于第 j 列而得的矩阵.□

2. 矩阵的初等矩阵分解

例 4 设

$$\boldsymbol{A}=\begin{pmatrix}1&1&0&1\\1&2&1&2\\2&3&1&3\end{pmatrix}.$$

证明:存在初等矩阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_l,\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_s$ 使得 $\boldsymbol{P}_1\boldsymbol{P}_2\cdots\boldsymbol{P}_l\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_s=\boldsymbol{U}$, 其中 $\boldsymbol{U}$ 是左上角为单位矩阵其他子块皆为零的矩阵.

证

$$\boldsymbol{A}=\begin{pmatrix}1&1&0&1\\1&2&1&2\\2&3&1&3\end{pmatrix}\xrightarrow{r_2+(-1)r_1}\begin{pmatrix}1&1&0&1\\0&1&1&1\\2&3&1&3\end{pmatrix}\xrightarrow{r_3+(-2)r_1}\begin{pmatrix}1&1&0&1\\0&1&1&1\\0&1&1&1\end{pmatrix}$$

$$\xrightarrow{c_2+(-1)c_1}\begin{pmatrix}1&0&0&1\\0&1&1&1\\0&1&1&1\end{pmatrix}\xrightarrow{c_4+(-1)c_1}\begin{pmatrix}1&0&0&0\\0&1&1&1\\0&1&1&1\end{pmatrix}$$

$$\xrightarrow{r_3+(-1)r_2}\begin{pmatrix}1&0&0&0\\0&1&1&1\\0&0&0&0\end{pmatrix}\xrightarrow[c_4+(-1)c_2]{c_3+(-1)c_2}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}.$$

写成初等矩阵乘积,有

$$\boldsymbol{E}(r_3-1r_2)\boldsymbol{E}(r_3-2r_1)\boldsymbol{E}(r_2-1r_1)\boldsymbol{A}E(c_2-1c_1)\boldsymbol{E}(c_4-1c_1)\boldsymbol{E}(c_3-1c_2)\boldsymbol{E}(c_4-1c_2)$$

$$=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_2&O\\O&O\end{pmatrix}.$$

将该例使用的方法推广到任意矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$,有如下定理.

定理 3.2　设 $\boldsymbol{A}$ 为任一 $m\times n$ 矩阵,则存在 m 阶初等矩阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_l$ 及 n 阶初等矩阵 $\boldsymbol{Q}_1,\boldsymbol{Q}_2,\cdots,\boldsymbol{Q}_k$ 使得

$$\boldsymbol{P}_l\boldsymbol{P}_{l-1}\cdots\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k=\begin{pmatrix}\boldsymbol{I}_r&O\\O&O\end{pmatrix}_{m\times n},$$

其中 $r\leqslant\min\{m,n\}$.

*证　若 $\boldsymbol{A}=\boldsymbol{O}$,令 $r=0$,定理成立.设 $\boldsymbol{A}=(a_{ij})\neq\boldsymbol{O}$ 则有 $a_{uv}\neq0$.将 1 行与 u 行互换,1 列与 v 列互换后,a_{uv}变到矩阵左上角,即存在初等矩阵

$$\boldsymbol{P}_1=\boldsymbol{E}(r_1\backslash r_u),\ \boldsymbol{Q}_1=\boldsymbol{E}(c_1\backslash c_v)$$

使得

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\boldsymbol{E}(r_1\backslash r_u)\boldsymbol{A}\boldsymbol{E}(c_1\backslash c_v)=\boldsymbol{A}_1=\begin{pmatrix}a_{uv}&\cdots&*\\\cdots&\ddots&\cdots\\ *&\cdots&*\end{pmatrix}.$$

对 A_1 作如下分块

$$\boldsymbol{A}_1=\left(\begin{array}{c|ccc}a_{uv}&\cdots&\cdots&*\\\hline\cdots&\cdots&\cdots&\cdots\\\cdots&\cdots&\ddots&*\\ *&\cdots&*&*\end{array}\right)=\begin{pmatrix}a_{uv}&\boldsymbol{A}_{12}\\\boldsymbol{A}_{21}&\boldsymbol{A}_{22}\end{pmatrix}.$$

由于 $a_{uv}\neq0$,有

$$\begin{pmatrix}a_{uv}^{-1}&\boldsymbol{O}\\-a_{uv}^{-1}\boldsymbol{A}_{21}&\boldsymbol{I}_{m-1}\end{pmatrix}\begin{pmatrix}a_{uv}&\boldsymbol{A}_{12}\\\boldsymbol{A}_{21}&\boldsymbol{A}_{22}\end{pmatrix}\begin{pmatrix}1&-a_{uv}^{-1}\boldsymbol{A}_{12}\\\boldsymbol{O}&\boldsymbol{I}_{n-1}\end{pmatrix}=\begin{pmatrix}1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{A}_{22}-a_{uv}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{12}\end{pmatrix}.$$

由于

$$\begin{pmatrix}a_{uv}^{-1}&\boldsymbol{O}\\-a_{uv}^{-1}\boldsymbol{A}_{21}&\boldsymbol{I}_{m-1}\end{pmatrix},\quad\begin{pmatrix}1&-a_{uv}^{-1}\boldsymbol{A}_{12}\\\boldsymbol{O}&\boldsymbol{I}_{n-1}\end{pmatrix},$$

是分别对单位矩阵 $\boldsymbol{I}_m$, $\boldsymbol{I}_n$ 进行了一系列的行、列的初等变换得来,所以存在初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_w, \boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_q$ 使得

$$\boldsymbol{P}_w \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{A} \boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_q = \begin{pmatrix} 1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_{22} - a_{uv}^{-1} \boldsymbol{A}_{21} \boldsymbol{A}_{12} \end{pmatrix}.$$

若 $\boldsymbol{A}_{22} - a_{uv}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{12}$ 为 $\boldsymbol{O}$ 或其行和列数为 1,定理得证 .记 $\boldsymbol{A}_{22} - a_{uv}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{12} = \boldsymbol{B}$, $\boldsymbol{B} \neq \boldsymbol{O}$ 且其行或列数大于 1,对 $\boldsymbol{B}$ 再重复上面的过程,最终一定可找到 m 阶的初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_l$ 及 n 阶的初等矩阵 $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_h$ 使得

$$\boldsymbol{P}_l \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{A} \boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_h = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}_{m \times n}.\square$$

学完第四节,我们将知道,定理 3.2 说明:任意矩阵,无论多么复杂,都可分解成三种类型的初等矩阵和一个左上角为单位矩阵且其余元素皆为零的矩阵的乘积 .这类似于,每一个整数都能分解成素数的乘积 .所不同的是矩阵的分解不唯一 .

第三节　行列式及其性质

用二阶行列式可简明表示某些二元一次方程组的解,同样用三阶行列式也可简明表示某些三元一次方程组的解 .是否也能用类似方法表示一般 n 元一次方程组的解呢 ? 回答是肯定的, n 阶行列式就可用来表示某些 n 元一次方程组的解 .另外,对矩阵的深入研究也需要行列式 .这节先简单介绍 n 阶行列的定义、性质及其应用 .这既是为了后续内容的需要,也是因为它在许多理论和应用问题中起着重要作用 .在线性代数中,行列式是一个基本工具,差不多从本世纪初以来,行列式理论已不是线性代数的中心,所以对它不拟多加论述 .

3.1　行列式的定义

1. 行列式的定义

求解二元一次方程组

$$a_{11}x_1 + a_{12}x_2 = b_1, \tag{3.1}$$

$$a_{21}x_1 + a_{22}x_2 = b_2. \tag{3.2}$$

无损一般性,可设 $a_{11} \neq 0$,方程(3.2)减去方程(3.1)的 a_{21}/a_{11}倍,可得

$$\left(a_{22} - \frac{a_{21}}{a_{11}} a_{12}\right) x_2 = \left(b_2 - \frac{a_{21}}{a_{11}} b_1\right),$$

若 $a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} \neq 0$,可解出 x_2,再将 x_2 代入方程(3.1)可求出 x_1,得

$$x_1 = \frac{b_1 a_{22} - b_2 a_{12}}{a_{11} a_{22} - a_{21} a_{12}},$$

$$x_2 = \frac{a_{11} b_1 - a_{21} b_2}{a_{11} a_{22} - a_{21} a_{12}}.$$

上面解方程组的方法称为消元法.对于任意二阶方阵 $\boldsymbol{A} = (a_{ij})_{2\times 2}$(为节省记号,用 $\boldsymbol{A}$ 来表示任意 2×2 矩阵,尽管在上面方程组中它只表示系数矩阵)定义相应的二阶行列式

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21}.$$

当 $D \neq 0$ 时,

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

对任意三阶矩阵 $\boldsymbol{A} = (a_{ij})_{3\times 3}$,定义相应的三阶行列式

$$\begin{aligned} D &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32} \\ &= (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{1+2} a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{1+3} a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{33} \end{vmatrix}. \end{aligned} \tag{3.3}$$

用消元法解三元一次方程组

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + a_{13} x_3 &= b_1, \\ a_{21} x_1 + a_{22} x_2 + a_{23} x_3 &= b_2, \\ a_{31} x_1 + a_{32} x_2 + a_{33} x_3 &= b_3, \end{aligned}$$

其解可用行列式简明表示为

$$x_1 = \frac{1}{D}\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad x_2 = \frac{1}{D}\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad x_3 = \frac{1}{D}\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

从(3.3)式可见,三阶行列式可用它的第一行元素与第二、三行元素组成的三个二阶行列式来表示,由此产生如下设想:能否模仿三阶行列式,用 $(n-1)$ 阶

行列式定义 n 阶行列式,最终使得 n 元一次线性方程组的解也能用 n 阶行列式简明地表示出来呢?回答是肯定的.首先,用归纳法来定义 n 阶行列式.然后再推导用行列式来表达线性方程组的解的公式.

定义 3.4　设 n 阶矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$, $\boldsymbol{A}$ 的 n 阶行列式

$$D=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

简记作 $\Delta(a_{ij})$, $\det(\boldsymbol{A})$ 或 $|\boldsymbol{A}|$,递推地定义如下:当 $n=1$ 时,定义 $|a|=a$(注意 $|a|$ 不表示 a 的绝对值,只表示一个元素构成的行列式);一般地,如果已知 $n-1$ 阶行列式的算法,则 n 阶($n\geq 2$)行列式

$$\Delta(a_{ij})=a_{11}M_{11}-a_{12}M_{12}+\cdots+(-1)^{1+n}a_{1n}M_{1n}=\sum_{i=1}^{n}a_{1i}(-1)^{1+i}M_{1i},\tag{3.4}$$

其中 M_{1j} 是划去 n 阶行列式 $\Delta(a_{ij})$ 第 1 行和第 j 列元素后得到的 $n-1$ 阶行列式,即 $n-1$ 阶子行列式或 $n-1$ 阶子式, M_{1j} 称为元素 a_{1j} 的余子式.

一般地说,在矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的行列式 D 中,取一个元素 a_{ij},去掉 D 的第 i 行及第 j 列后形成的 $n-1$ 阶子行列式称为 a_{ij} 的余子式,记为 M_{ij},而 $(-1)^{i+j}M_{ij}$ 称为 a_{ij} 的代数余子式.

公式(3.4)称为行列式按第一行元素展开的公式.若 a_{1j} 的代数余子式记为 A_{1j},则公式(3.4)变为

$$\Delta(a_{ij})=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}=\sum_{i=1}^{n}a_{1i}A_{1i}.\tag{3.5}$$

这样,定义 3.4 可重新叙述为:矩阵 A 的行列式等于其第一行元素与其代数余子式乘积之和.

例 1　用定义 3.4,计算 3 阶行列式

$$\Delta(a_{ij})=\begin{vmatrix} 1 & 0 & -1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix}.$$

解　由定义 3.4,有

$$\begin{aligned}\Delta(a_{ij}) &= a_{11}M_{11}-a_{12}M_{12}+a_{13}M_{13} \\ &= 1\times\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix}-0\times\begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix}+(-1)\times\begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix}.\end{aligned}$$

对于三个二阶行列式

$$\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix},$$

再由定义 3.4,分别有

$$\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 2 \times |1| - 1|1| = 1, \quad \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = 0 \times |1| - 1 \times |1| = -1,$$

$$\begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} = 0 \times |1| - 2 \times |1| = -2 .$$

$$\Delta(a_{ij}) = 1 \times \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} - 0 \times \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} + (-1) \times \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} = 1 \times 1 + (-1) \times (-2) = 3 .$$

由公式(3.4)可知,用定义 3.4 可求任意 n 阶行列式,但是随着行列式的阶数的增加,计算越来越复杂,当阶数较大时,将无法再用定义 3.4 来直接计算行列式(见第四章 1.3 小节的算法复杂性) .这样,必须研究行列式的性质,以便把一个复杂的行列式变成与之相等的一个相对简单的行列式 .那末,什么样的行列式是相对简单的呢 ? 请看下面的例子 .

例 2　计算下三角形行列式(主对角线以上的元素都为零的行列式)

$$\Delta(a_{ij}) = \begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \cdots & \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} .$$

解　因为 $a_{1j} = 0, j = 2, 3, \cdots, n$,所以

$$\Delta(a_{ij}) = a_{11} M_{11} - a_{12} M_{12} + \cdots + (-1)^{1+n} a_{1n} M_{1n} = a_{11} M_{11} . \quad (3.6)$$

同理,

$$M_{11} = a_{22}(M_{11})_{11} - a_{23}(M_{11})_{12} + \cdots + (-1)^{1+n} a_{2n}(M_{11})_{1n-1} = a_{22}(M_{11})_{11},$$

其中$(M_{11})_{1j}$为行列式 M_{11} 划去第 1 行和第 j 列后得到的 $n-2$ 阶行列式,即 M_{11}中第 1 行第 j 列元素的余子式 .将 M_{11}代入(3.6)有

$$\Delta(a_{ij}) = a_{11} M_{11} = a_{11} a_{22} (M_{11})_{11} .$$

如此下去,易知 $\Delta(a_{ij}) = a_{11} a_{22} \cdots a_{nn}$.

现在还不能直接用定义很简洁地证明上三角形行列式(对角线以下元素都为零)也等于对角线元素之积 .例 4 将应用行列式的一些性质简明地证明这个结论 .

应注意到,有了行列式的概念,向量积公式(1.3)和混合积公式(1.4)可形式

地用行列式表达,以方便记忆,见习题三 A 题 19,20 .

2. 按两行展开行列式

下面使用记号$(M_{1i})_{1k}$表示在行列式$\Delta(a_{ij})$的子式M_{1i}中划掉它的第 1 行第 k 列元素后产生的 $n-2$ 阶子式 .

例 3　证明:

$$\left|(a_{ij})_{4\times4}\right| = \sum_{i=2}^{4}\sum_{k=1}^{i-1}\begin{vmatrix} a_{1k} & a_{1i} \\ a_{2k} & a_{2i}\end{vmatrix}(-1)^{i+k+1+2}(M_{1i})_{1k} .$$

证　由定义 3 .4 知

$$\Delta(a_{ij}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44}\end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44}\end{vmatrix} -$$

$$a_{12}\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44}\end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44}\end{vmatrix} - a_{14}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43}\end{vmatrix} . \tag{3.7}$$

再用定义 3.4,将上面四个三阶行列式展成一些二阶行列式的和,而这些二阶行列式都是取于该四阶行列式的后两行,即

$$\begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42}\end{vmatrix} = (M_{14})_{13}, \quad \begin{vmatrix} a_{31} & a_{33} \\ a_{41} & a_{43}\end{vmatrix} = (M_{14})_{12}, \quad \begin{vmatrix} a_{31} & a_{34} \\ a_{41} & a_{44}\end{vmatrix} = (M_{13})_{12},$$

$$\begin{vmatrix} a_{32} & a_{33} \\ a_{42} & a_{43}\end{vmatrix} = (M_{14})_{11}, \quad \begin{vmatrix} a_{32} & a_{34} \\ a_{42} & a_{44}\end{vmatrix} = (M_{13})_{11}, \quad \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44}\end{vmatrix} = (M_{12})_{11} .$$

行列式 $\Delta(a_{ij})$等于上面六个二阶行列式各乘一个适当系数后再相加 .因此,只要算出每个二阶行列式所乘的系数,也就知道 $\Delta(a_{ij})$的值与上面六个二阶行列式的关系 .子式$(M_{14})_{13}$在(3.7)中出现两次(虚线框),将后两个三阶行列式展成二阶行列式和时,$(M_{14})_{13}$的系数为 $a_{13}a_{24}-a_{14}a_{23}$.用同样的方法可算出其他子式的系数 .上面六个子式的系数分别为

$$a_{13}a_{24} - a_{14}a_{23} = \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24}\end{vmatrix}(-1)^{1+3+4},$$

$$a_{14}a_{22} - a_{12}a_{24} = \begin{vmatrix} a_{12} & a_{14} \\ a_{22} & a_{24}\end{vmatrix}(-1)^{1+2+4},$$

$$a_{13}a_{23}-a_{12}a_{23}=\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix}(-1)^{1+2+2},$$

$$a_{11}a_{24}-a_{14}a_{21}=\begin{vmatrix}a_{11}&a_{14}\\a_{21}&a_{24}\end{vmatrix}(-1)^{1+1+4},$$

$$a_{13}a_{21}-a_{11}a_{23}=\begin{vmatrix}a_{11}&a_{13}\\a_{21}&a_{23}\end{vmatrix}(-1)^{1+1+3},$$

$$a_{11}a_{22}-a_{12}a_{21}=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}(-1)^{1+1+2}.$$

因此,有

$$\Delta(a_{ij})=\sum_{i=2}^{4}\sum_{k=1}^{i-1}\begin{vmatrix}a_{1k}&a_{1i}\\a_{2k}&a_{2i}\end{vmatrix}(-1)^{i+k+1}(M_{1i})_{1k}.$$

例 3 表明,可按前两行展开行列式.对一般情形,有下面定理 3.3.在研究行列式性质过程中,下面的公式起了关键的作用.n 阶行列式按前两行展开公式:

$$\begin{aligned}\Delta(a_{ij})&=\sum_{i=2}^{n}\sum_{k=1}^{i-1}(a_{1i}a_{2k}-a_{1k}a_{2i})(-1)^{i+k}(M_{1i})_{1k}\\&=\sum_{i=2}^{n}\sum_{k=1}^{i-1}\begin{vmatrix}a_{1k}&a_{1i}\\a_{2k}&a_{2i}\end{vmatrix}(-1)^{i+k+1}(M_{1i})_{1k},\end{aligned}\tag{3.8}$$

在公式(3.8)中,$(-1)^{i+k+1}(M_{1i})_{1k}=(-1)^{i+k+1+2}(M_{1i})_{1k}$;又因为 $k<i$,$(M_{1i})_{1k}$实际是在 $\Delta(a_{ij})$中去掉二阶行列式

$$\begin{vmatrix}a_{1k}&a_{1i}\\a_{2k}&a_{2i}\end{vmatrix}$$

所在行与列的元素之后所得的 $n-2$ 阶行列式,$(M_{1i})_{1k}$称为该二阶行列式的余子式,而$(-1)^{i+k+1+2}(M_{1i})_{1k}$称为该二阶行列式的代数余子式.这样,有如下定理

定理 3.3　在 n 阶行列式 $\Delta(a_{ij})$中,选定前两行,则 $\Delta(a_{ij})$等于含于这两行所有的二阶子式与其代数余子式乘积之和.

定理 3.3 的证明,即公式(3.8)的证明见附录 B.1.

3.2　行列式的性质

在本小节中,为了区别行列式 $\Delta(a_{ij})$的子式 M_{1i},$(M_{1i})_{1k}$和行列式 $\Delta(b_{ij})$的子式 M_{1i},$(M_{1i})_{1k}$,把它们依次记为 AM_{1i},$(AM_{1i})_{1k}$和 BM_{1i},$(BM_{1i})_{1k}$.

*定理 3.4　若互换行列式相邻两行,则行列式变号.

证　设 $\Delta(a_{ij})$是 n 阶行列式$(n\geq 2)$.将 $\Delta(a_{ij})$的第 i 行和第 $i+1$ 行$(i=1,2,\cdots,n-1)$的元素对换后变为 $\Delta(b_{ij})$.对 n 作数学归纳法.当 $n=2$ 时,由(3.4)有

$$\Delta(a_{ij}) = a_{11}a_{22} - a_{12}a_{21};\ \Delta(b_{ij}) = b_{11}b_{22} - b_{12}b_{21}.$$

因为将 $\Delta(a_{ij})$的第 i 行和第 $i+1$ 行(第 1 行与第 2 行)的数对换后变为 $\Delta(b_{ij})$,因而

$$b_{11} = a_{21},\ b_{12} = a_{22},\ b_{21} = a_{11},\ b_{22} = a_{12}$$

所以

$$\Delta(b_{ij}) = b_{11}b_{22} - b_{12}b_{21} = a_{12}a_{21} - a_{11}a_{22} = -\Delta(a_{ij}).$$

设 $n\leq m-1$ 时,定理成立;往证 $n=m$ 时,定理成立.

当 $i=1$ 时,即 $\Delta(a_{ij})$的第 1 行与第 2 行元素对换后 $\Delta(a_{ij})$变为 $\Delta(b_{ij})$,于是

$$b_{11} = a_{21},\ b_{12} = a_{22},\ \cdots,\ b_{1i} = a_{2i},\ \cdots,$$
$$b_{1m} = a_{2m},\ b_{21} = a_{11},\ b_{22} = a_{12},\ \cdots,\ b_{2m} = a_{1m}.$$

对其他行而言,$b_{ij} = a_{ij}\ (i=3,4,\cdots,m)$.由定义 3.4,有$(AM_{1i})_{1k} = (BM_{1i})_{1k}$ $(i=1,2,\cdots,m;\ k=1,2,\cdots,m-1)$.由定理 3.3,有

$$\begin{aligned}\Delta(b_{ij}) &= \sum_{i=2}^{m}\sum_{k=1}^{i-1}\begin{vmatrix} b_{1k} & b_{1i} \\ b_{2k} & b_{2i}\end{vmatrix}(-1)^{i+k+1}(BM_{1i})_{1k} \\ &= \sum_{i=2}^{m}\sum_{k=1}^{i-1}\begin{vmatrix} a_{2k} & a_{2i} \\ a_{1k} & a_{1i}\end{vmatrix}(-1)^{i+k+1}(AM_{1i})_{1k} \\ &= \sum_{i=2}^{m}\sum_{k=1}^{i-1}\begin{vmatrix} a_{1k} & a_{1i} \\ a_{2k} & a_{2i}\end{vmatrix}(-1)^{i+k+2}(AM_{1i})_{1k} = -\Delta(a_{ij}).\end{aligned}$$

当 $i\geq 2$ 时,将 $\Delta(a_{ij})$的第 i 行和第 $i+1$ 行对换后,它变为 $\Delta(b_{ij})$.$\Delta(b_{ij})$的第 1 行与 $\Delta(a_{ij})$完全一样,而 BM_{1i}是由 AM_{1i}对换第 $i-1$ 行与第 i 行后而得到的.由归纳假设,有 $BM_{1i} = -AM_{1i},\ i=1,2,\cdots,m-1$.由定义 3.4,有

$$\Delta(b_{ij}) = \sum_{i=1}^{m} b_{1i}(-1)^{i+1}BM_{1i} = -\sum_{i=1}^{m} a_{1i}(-1)^{i+1}AM_{1i} = -\Delta(a_{ij}).\square$$

推论 3.1　若互换行列式的任意两行(或两列),则行列式变号.

*证　设 $\Delta(a_{ij})$是 n 阶行列式$(n\geq 2)$.将 $\Delta(a_{ij})$的第 i 行和第 j 行$(i,j=1,2,\cdots,n,\ i\neq j)$对换后变为 $\Delta(b_{ij})$.不妨设 $j>i$,$\Delta(a_{ij})$经 $2(j-i)+1$ 次相邻行对换后变为 $\Delta(b_{ij})$.由定理 3.4 知

$$\Delta(a_{ij}) = (-1)^{2(j-i)+1}\Delta(b_{ij}) = -\Delta(b_{ij}).$$

另外,当第 i 列和第 j 列对换时,对 n 作数学归纳法.当 $n=2$ 时,显然推论成立.假设 $n\leq k$ 时,推论成立,往证 $n=k+1$ 时,推论成立.

不妨设 $l<h$.设 $\Delta(a_{ij})$第 l 列与第 h 列对换后为 $\Delta(b_{ij})$.由归纳假设知,当 $j\neq h$ 和 l 时, $BM_{1j}=-AM_{1j}$;当 $j=h$ 时, $BM_{1h}=(-1)^{h-l-1}AM_{1l}$;当 $j=l$ 时, $BM_{1l}=(-1)^{h-l-1}AM_{1h}$.由定义 3.4,有

$$\begin{aligned}\Delta(b_{ij}) &= \sum_{i=1}^{n}(-1)^{1+i}b_{1i}BM_{1i}\\ &= (-1)^{1+l}b_{1l}BM_{1l} + (-1)^{h+1}b_{1h}BM_{1h} + \sum_{i=1,i\neq h,i\neq l}^{n}(-1)^{1+i}b_{1i}BM_{1i}\\ &= (-1)^{h}a_{1h}AM_{1h} + (-1)^{l}a_{1l}AM_{1l} - \sum_{i=1,i\neq h,i\neq l}^{n}(-1)^{1+i}a_{1i}AM_{1i}\\ &= -\Delta(a_{ij}).\square\end{aligned}$$

在下面的定理及推论中,对行而言成立的行列式性质,对列而言也成立;在行列式性质的证明中,只针对行的情形予以证明.对于列的情形的证明留做作业.

推论 3.2　若行列式两行(或列)的对应的元素相等,则其值为零.

证　设 $\Delta(a_{ij})$是 n 阶行列式($n\geq 2$).若 $\Delta(a_{ij})$的第 i 行和第 j 行元素对应相等($i,j=1,2,\cdots,n,i\neq j$).则将 $\Delta(a_{ij})$的第 i 行和第 j 行($i,j=1,2,\cdots,n,i\neq j$)对换后仍为 $\Delta(a_{ij})$.由推论 3.1 知 $\Delta(a_{ij})=-\Delta(a_{ij})$,于是 $\Delta(a_{ij})=0$.$\square$

推论 3.3　若行列式的某一行(或列)的所有元素都乘上同一个数 k,则等于用数 k 乘该行列式.

证　设 $\Delta(a_{ij})$是 n 阶行列式($n\geq 2$).将 $\Delta(a_{ij})$的第 i 行每个元素都乘数 k ($i=1,2,\cdots,n$)后变为 $\Delta(b_{ij})$.若 $i=1$,由定义 3.4,显然有 $\Delta(b_{ij})=k\Delta(a_{ij})$.若 $i\geq 2$,设 $\Delta(b_{ij})$的第 i 行与第 1 行对换($i=1,2,\cdots,n$)后变为 $\Delta(c_{ij})$,则

$$\Delta(b_{ij}) = -\Delta(c_{ij}) = -k(-\Delta(a_{ij})) = k\Delta(a_{ij}).\square$$

推论 3.4　设 $\Delta(a_{ij})$是 n 阶行列式($n\geq 2$).若 n 阶行列式 $\Delta(b_{ij})$和 $\Delta(c_{ij})$满足,当 $i\neq l$, $j=1,2,\cdots,n$ 时, $b_{ij}=c_{ij}=a_{ij}$, $\Delta(a_{ij})$的第 l 行(第 l 列)元素 $a_{lk}=b_{lk}+c_{lk}$ ($k=1,2,\cdots,n$),则 $\Delta(a_{ij})=\Delta(b_{ij})+\Delta(c_{ij})$.

证　若 $l=1$,由定义 3.4 显然有 $\Delta(a_{ij})=\Delta(b_{ij})+\Delta(c_{ij})$.若 $l\geq 2$,用与推论 3.3 类似的证法:对换 $\Delta(a_{ij})$的第 l 行与第 1 行即可证.$\square$

推论 3.5　若把行列式的某一行(或列)的所有元素乘以同一个数,然后再

加到另外一行(或列)上去,则行列式值不变.

证 设 $\Delta(a_{ij})$是 n 阶行列式($n\geq 2$).将 $\Delta(a_{ij})$的第 i 行乘 k 加于第 j 行($i, j=1,2,\cdots,n, i\neq j$)后变为 $\Delta(b_{ij})$.用推论 3.4,将 $\Delta(b_{ij})$按第 j 行拆为两个行列式之和,即 $\Delta(b_{ij})=\Delta(a_{ij})+\Delta(c_{ij})$.对于 $\Delta(c_{ij})$的第 j 行应用推论 3.3,提出 k 后,其第 i 行的元素与第 j 行的元素对应相等,由推论 3.2 知 $\Delta(c_{ij})=0$.□

例 4 设 $\Delta(a_{ij})$是 n 阶上三角形行列式($n\geq 2$),即当 $j<i$ 时,$a_{ij}=0$.证明:

$$\Delta(a_{ij})=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}=a_{11}a_{22}\cdots a_{nn}.$$

证 对 n 作数学归纳法.当 $n=2$ 时,结论成立.假设 $n\leq k$ 时,结论成立;往证 $n=k+1$ 时,结论成立.当 $i=2,3,\cdots,n$ 时,M_{1i}的第一列所有元素都为零.由推论 3.3 知,当 $i=2,3,\cdots,n$ 时,$M_{1i}=0$.再由归纳假设,有 $M_{11}=a_{22}\cdots a_{nn}$.所以

$$\Delta(a_{ij})=a_{11}M_{11}-a_{12}M_{12}+\cdots+(-1)^{n+1}M_{1n}=a_{11}M_{11}=a_{11}a_{22}\cdots a_{nn}.\square$$

对于一般的高阶行列式几乎都不直接用定义 3.4 或定理 3.3 计算,而是用上面那些行列式性质来计算.为了简明.使用下面记号:$r_i\backslash r_j$ 表示第 i 行与第 j 行对换,kr_i 表示第 i 行乘 k;r_i+kr_j 表示第 j 行乘 k 加于第 i 行,对列有类似的记号,只是将字母 r 换为 c.

例 5 计算

$$D=\begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}.$$

解

$$D\overset{c_1\backslash c_2}{\underset{r_2-r_1,\, r_4+5r_1}{=}}-\begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & -8 & 4 & -6 \\ 0 & 2 & 1 & -1 \\ 0 & 16 & -2 & 7 \end{vmatrix}\overset{r_2\backslash r_3}{\underset{\substack{r_3+4r_2 \\ r_4-8r_2 \\ r_4+\frac{5}{4}r_3}}{=}}\begin{vmatrix} 1 & 3 & -1 & 2 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & 8 & -10 \\ 0 & 0 & 0 & \dfrac{5}{2} \end{vmatrix}=40.$$

例 5 说明,可利用行列式性质,即推论 3.1—推论 3.5,将任一行列式化为与之相等的上三角形行列式或下三角形行列式,这样可简化计算.

●

例 6　计算

$$D=\begin{vmatrix} c & a & d & b \\ a & c & d & b \\ a & c & b & d \\ c & a & b & d \end{vmatrix}.$$

解

$$D \xrightarrow[r_4-r_3]{r_1-r_2} \begin{vmatrix} c-a & a-c & 0 & 0 \\ a & c & d & b \\ a & c & b & d \\ c-a & a-c & 0 & 0 \end{vmatrix}=0.$$

例 7　设

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \\ 3 & 6 & 4 \end{pmatrix},$$

计算$|\boldsymbol{A}|$和$|\boldsymbol{A}^{\mathrm{T}}|$.

解

$$\begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 3 \\ 3 & 6 & 4 \end{vmatrix} \xrightarrow[r_3-2r_1]{r_2-2r_1} \begin{vmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 3 & 1 \end{vmatrix} \xrightarrow{c_2-3c_3} \begin{vmatrix} 1 & -2 & 1 \\ 0 & -5 & 1 \\ 0 & 0 & 1 \end{vmatrix}=-5.$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 1 & 0 & 6 \\ 1 & 3 & 4 \end{vmatrix} \xrightarrow[c_3-2c_1]{c_2-2c_1} \begin{vmatrix} 1 & 0 & 0 \\ 1 & -2 & 3 \\ 1 & 1 & 1 \end{vmatrix} \xrightarrow{r_2-3r_3} \begin{vmatrix} 1 & 0 & 0 \\ -2 & -5 & 0 \\ 1 & 1 & 1 \end{vmatrix}=-5.$$

定理 3.5　设 $\boldsymbol{A}=(a_{ij})_{n\times n}$.$\boldsymbol{A}^{\mathrm{T}}$为其转置矩阵,则$|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|$.

证　按某个顺序反复使用推论 3.1—推论 3.5,可将$|\boldsymbol{A}|$化为下三角形行列式(见例 2).对$|\boldsymbol{A}^{\mathrm{T}}|$,也按前述的顺序使用推论 3.1—推论 3.5,只不过对$|\boldsymbol{A}|$使用关于行(或列)运算,对$|\boldsymbol{A}^{\mathrm{T}}|$则使用相对应的列(或行)的相同运算(即若互换$\boldsymbol{A}$第 i 行与第 j 行,则相应地,互换$|\boldsymbol{A}^{\mathrm{T}}|$第 i 列与第 j 列,若将$|\boldsymbol{A}|$的第 i 行乘常数 k 后加于第 j 行,则相应地,将$|\boldsymbol{A}^{\mathrm{T}}|$的第 i 列乘常数 k 后加于第 j 列,等等).于是当将$|\boldsymbol{A}|$化为下三角形行列式时,$|\boldsymbol{A}^{\mathrm{T}}|$化为上三角形行列式(见例 6),因为此时$|\boldsymbol{A}|$所对应的下三角形行列式与$|\boldsymbol{A}^{\mathrm{T}}|$所对应的上三角形行列式对角线上的元素相同,所以$|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|$.□

称$|\boldsymbol{A}^{\mathrm{T}}|$为行列式$|\boldsymbol{A}|$的转置行列式.

例 8　证明:(1) $|(a_{ij})_{3\times3}|=a_{31}A_{31}+a_{32}A_{32}+a_{33}A_{33}$.(2) $a_{11}A_{31}+a_{12}A_{32}+a_{13}A_{33}=0$.

证 (1) 根据定理 3.4 及定义 3.4,有

$$\left|(a_{ij})_{3\times 3}\right| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^2 \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix}$$

$$= (-1)^2 \quad a_{31}(-1)^{1+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{32}(-1)^{1+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}(-1)^{1+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$= a_{31}(-1)^{1+3}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{32}(-1)^{3+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}(-1)^{3+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$= a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} .$$

(2) 考察下面的行列式,该行列式的第 1 行与第 3 行相等 .因此它为零 .于是应用(1)的结论有

$$0 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix}$$

$$= a_{11}(-1)^{3+1}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{12}(-1)^{3+2}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{13}(-1)^{3+3}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$= a_{11}A_{31} + a_{12}A_{32} + a_{13}A_{33} .$$

推广这个例子,有下面的定理 .

定理 3.6 设 $\Delta(a_{ij})$是 $n(n\geq 2)$阶行列式,则

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \delta_{ij}\Delta(a_{ij}), \tag{3.9}$$

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = \delta_{ij}\Delta(a_{ij}), \tag{3.10}$$

当 $i=j$ 时,$\delta_{ij}=1$;当 $i\neq j$ 时,$\delta_{ij}=0$, $A_{ij}=(-1)^{i+j}M_{ij}$是 a_{ij}的代数余子式 .

* **证** 设 $i=j$.当 $i=j=1$ 时,由定义 3.4,有

$$\Delta(a_{ij}) = a_{11}(-1)^{1+1}M_{11} + a_{12}(-1)^{1+2}M_{12} + \cdots + a_{1n}(-1)^{1+n}M_{1n}$$
$$= a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} .$$

当 $i=j\geq 2$ 时,将 $\Delta(a_{ij})$经 $i-1$ 次相邻行对换后变为 $\Delta(b_{ij})$.于是 $\Delta(b_{ij})$划掉第一行后与 $\Delta(a_{ij})$划掉第 i 行后完全一样 .这样就有 $AM_{jk} = BM_{1k}$, $b_{1k} = a_{ik}(k=1,2,\cdots,n)$及

$$(-1)^{i-1}\Delta(a_{ij}) = \Delta(b_{ij})$$
$$= (-1)^{1+1}b_{11}BM_{11} + (-1)^{1+2}b_{12}BM_{12} + \cdots + (-1)^{1+n}b_{1n}BM_{1n}$$
$$= (-1)^{1+1}a_{i1}AM_{i1} + (-1)^{1+2}a_{i2}AM_{i2} + \cdots + (-1)^{1+n}a_{in}AM_{in} .$$

整理后有 $a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \Delta(a_{ij})$.这个公式通常称为行列式按第 i 行展开.

当 $i \neq j$.将 $\Delta(a_{ij})$中第 j 行元素用其第 i 行元素代替,而其他行(包括第 i 行)元素不变得到 n 阶行列式 $\Delta(b_{ij})$,即 $\Delta(b_{ij})$划掉第 j 行后与 $\Delta(a_{ij})$划掉第 j 行后完全一样,且 $b_{jk} = a_{ik}(k = 1,2,\cdots,n)$.显然 $\Delta(b_{ij})$的第 j 行与第 i 行完全一样.由推论 3.2 有 $\Delta(b_{ij}) = 0$.再把 $\Delta(b_{ij})$按第 j 行展开有

$$\begin{aligned} 0 = \Delta(b_{ij}) &= b_{j1}(-1)^{j+1}BM_{j1} + b_{j2}(-1)^{j+2}BM_{j2} + \cdots + b_{jn}(-1)^{j+n}BM_{jn} \\ &= a_{i1}(-1)^{j+1}AM_{j1} + a_{i2}(-1)^{j+2}AM_{j2} + \cdots + a_{in}(-1)^{j+n}AM_{jn} \\ &= a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}. \end{aligned}$$

利用 Kronecker 符号 δ_{ij},上式与行列式按第 i 行展开可合并成一个公式

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = \delta_{ij}\Delta(a_{ij}).$$

对于列的情形,转置后用类似的方法可证.□

例 9　计算

$$D = \begin{vmatrix} -1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 2 & 1 \end{vmatrix}.$$

解　将第二行除第二行第三列元素 1 外其他元素都化为零,然后,再应用定理 3.6.

$$\begin{vmatrix} -1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 \\ 0 & 1 & 2 & 1 \end{vmatrix} \xrightarrow{c_1 + (-2)c_3} \begin{vmatrix} -3 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 \\ -4 & 1 & 2 & 1 \end{vmatrix}$$

$$= (-1)^{2+3}\begin{vmatrix} -3 & 1 & 1 \\ 1 & 1 & -1 \\ -4 & 1 & 1 \end{vmatrix} \xrightarrow[r_3 + r_2]{r_1 + r_2} - \begin{vmatrix} -2 & 2 & 0 \\ 1 & 1 & -1 \\ -3 & 2 & 0 \end{vmatrix}$$

$$= (-1)^{2+3}\begin{vmatrix} -2 & 2 \\ -3 & 2 \end{vmatrix} \xrightarrow{r_1 - r_2} - \begin{vmatrix} 1 & 0 \\ -3 & 2 \end{vmatrix} = -2.$$

例10　计算

$$D = \begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x & x_2 & x_3 \\ x_2 & x_2 & x & x_3 \\ x_3 & x_3 & x_3 & x \end{vmatrix}$$

解　用 $-x_1$, $-x_2$, $-x_3$ 分别乘第一行后加到第 2,3,4 行得

$$D=\begin{vmatrix}1&1&1&1\\0&x-x_1&x_2-x_1&x_3-x_1\\0&0&x-x_2&x_3-x_2\\0&0&0&x-x_3\end{vmatrix}=(x-x_1)(x-x_2)(x-x_3).$$

3.3　Cramer 法则

由前面的讨论可知, n 阶行列式是由二阶、三阶行列式归纳定义得来的.求解 n 个未知量的线性方程组时,用 n 阶行列式能否得到与二元一次方程组和三元一次方程组相类似的行列式表达的公式,还有待验证.下面证明可用 n 阶行列式来表示 n 个未知量的线性方程组的解,它说明了行列式定义是合理的.设 n 个未知量的线性方程组为

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=b_1,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=b_2,\\&\cdots\cdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n&=b_n.\end{aligned}\tag{3.11}$$

行列式

$$D=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}$$

称为线性方程组(3.11)的系数行列式.而 $D_i(i=1,2,\cdots,n)$是把 D 中第 i 列的元素 $a_{1i},a_{2i},\cdots,a_{ni}$分别换成常数项 $b_1,b_2,\cdots,b_n$ 而得到的行列式.如果(3.11)有解,即有一组数 $x_1,x_2,\cdots,x_n$ 满足(3.11),应用行列式的性质,有

$$\begin{aligned}Dx_1&=\begin{vmatrix}a_{11}x_1&a_{12}&\cdots&a_{1n}\\\cdots&\cdots&&\cdots\\a_{n1}x_1&a_{n2}&\cdots&a_{nn}\end{vmatrix}\\&=\begin{vmatrix}a_{11}x_1+\cdots+a_{1n}x_n&a_{12}&\cdots&a_{1n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}x_1+\cdots+a_{nn}x_n&a_{n2}&\cdots&a_{nn}\end{vmatrix}\\&=\begin{vmatrix}b_1&a_{12}&\cdots&a_{1n}\\\cdots&\cdots&&\cdots\\b_1&a_{n2}&\cdots&a_{nn}\end{vmatrix}=D_1.\end{aligned}$$

一般地,有 $Dx_i=D_i$, $i=1,2,\cdots,n$.当 $D\neq0$ 时,有

$$x_1 = \frac{D_1}{D},\ x_2 = \frac{D_2}{D},\cdots,\ x_n = \frac{D_n}{D},\tag{3.12}$$

即已证明:当 $D\neq 0$ 时,如果线性方程组(3.11)有解,则(3.12)是唯一解.下面证明:当 $D\neq 0$ 时,(3.12)是(3.11)的解.考察下面 $n+1$ 阶行列式,该行列式第1行与第 $i+1$ 行相同.并将它按第一行展开,

$$0 = \begin{vmatrix} b_i & a_{i1} & \cdots & a_{in} \\ b_1 & a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & & \cdots \\ b_n & a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$= b_i D + \sum_{k=1}^{n} a_{ik}(-1)^{1+k+1}\begin{vmatrix} b_1 & a_{11} & \cdots & a_{1,k-1} & a_{1,k+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & & \cdots \\ b_n & a_{n1} & \cdots & a_{n,k-1} & a_{n,k+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= b_i D - a_{i1}D_1 - \cdots - a_{in}D_n,\ i = 1,2,\cdots,n.$$

这说明,当 $D\neq 0$ 时,(3.12)的确是(3.11)的解.于是,得到下面重要法则:

定理3.7　[克拉默法则(G.Cramer,1704—1752)]如果线性方程组(3.11)的系数行列式 $D\neq 0$,那么它有唯一解

$$x_1 = \frac{D_1}{D},\ x_2 = \frac{D_2}{D},\cdots,\ x_n = \frac{D_n}{D}.$$

例 11　解方程组

$$\begin{aligned} x + 3y + 2z &= 0, \\ 2x - y + 3x &= 0, \\ 3x + 2y - z &= 0. \end{aligned}$$

解　因为其系数行列式 $D = 42\neq 0$,所以方程组有惟一的零解,即 $x = y = z = 0$.

例 12　解方程组

$$\begin{cases} 2x_1 + x_2 - 5x_3 + x_4 = 8, \\ x_1 - 3x_2 - 6x_4 = 9, \\ 2x_2 - x_3 + 2x_4 = -5, \\ x_1 + 4x_2 - 7x_3 + 6x_4 = 0. \end{cases}$$

解　用行列式性质,可计算如下各行列式的值:

$$D = \begin{vmatrix} 2 & 1 & -5 & 1 \\ 1 & -3 & 0 & -6 \\ 0 & 2 & -1 & 2 \\ 1 & 4 & -7 & 6 \end{vmatrix} = 27,\ D_1 = \begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix} = 81,$$

$$D_2=\begin{vmatrix}2&8&-5&1\\1&9&0&-6\\0&-5&-1&2\\1&0&-7&6\end{vmatrix}=-108,\ D_3=\begin{vmatrix}2&1&8&1\\1&-3&9&-6\\0&2&-5&2\\1&4&0&6\end{vmatrix}=-27,$$

$$D_4=\begin{vmatrix}2&1&-5&8\\1&-3&0&9\\0&2&-1&-5\\1&4&-7&0\end{vmatrix}=27.$$

由克拉默法则,求得唯一解

$$x_1=\frac{81}{27}=3,\ x_2=-\frac{108}{27}=-4,\ x_3=-\frac{27}{27}=-1,\ x_4=\frac{27}{27}=1.$$

第四节 方阵的逆及矩阵的秩

本节把方阵和行列式联系起来,并由此引出方阵的逆矩阵及一般矩阵的秩的概念,这两个概念对于矩阵理论的研究及应用都起着基本和关键的作用.

4.1 方阵的行列式

设方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$,其行列式为

$$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}$$

记为 $|\boldsymbol{A}|$ 或 $\det\boldsymbol{A}$.

设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$.$\boldsymbol{A}$ 的第 i_1 行,第 i_2 行,…,第 i_r 行和 $\boldsymbol{A}$ 的第 j_1 列,第 j_2 列,…,第 j_s 列的相交处的元素所排成的 $r\times s$ 矩阵,就叫做 $\boldsymbol{A}$ 的一个子块,其中 $i_1<i_2<\cdots<i_r$, $j_1<j_2<\cdots<j_s$,并记为

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

特别地,一个元素 a_{ij} 也可构成子块 $\boldsymbol{M}_{\boldsymbol{A}}\begin{pmatrix}i\\j\end{pmatrix}=(a_{ij})$.按矩阵乘积的定义可知,$\boldsymbol{AB}$ 的第 i 行第 j 列的元素是 $\boldsymbol{A}$ 的第 i 行做为 $\boldsymbol{A}$ 的子块与 $\boldsymbol{B}$ 的第 j 列做为 $\boldsymbol{B}$ 的子块

的矩阵乘积，即

$$\boldsymbol{M}_{AB}\begin{pmatrix} i \\ j \end{pmatrix} = \boldsymbol{M}_{A}\begin{pmatrix} i \\ 12\cdots n \end{pmatrix}\boldsymbol{M}_{B}\begin{pmatrix} 12\cdots n \\ j \end{pmatrix}.$$

当矩阵 $\boldsymbol{A}$ 的子块为 r 阶方阵

$$\boldsymbol{M}_{A}\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}$$

时，它的行列式叫做 $\boldsymbol{A}$ 的一个 r 阶子式，记为

$$\left|\boldsymbol{M}_{A}\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}\right| \text{或} \det \boldsymbol{M}_{A}\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}.$$

特别地，当 $\boldsymbol{A}$ 为方阵时，把 $\boldsymbol{A}$ 的第 $i_1, i_2, \cdots, i_r$ 行和第 $j_1, j_2, \cdots, j_r$ 列划掉后所得的 $\boldsymbol{A}$ 的 $n-r$ 阶子式和该子式乘以 $(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r}$ 分别称为

$$\boldsymbol{M}_{A}\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}$$

在 $\boldsymbol{A}$ 中的余子式和代数余子式，且分别记为

$$\text{余子式 } A\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}, \text{代数余子式 } A\begin{pmatrix} i_1\ i_2\cdots\ i_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix},$$

不过要注意，虽然这里的余子式与代数余子式用这个记号表示，但当 $\boldsymbol{A}$ 的阶数为 n 时，它们实际上都是 $n-r$ 阶行列式.

*定理 3.8　[Laplace(1749—1827)] 在 n 阶行列式 $|\boldsymbol{A}|$ 中，任取 $r(1\leq r\leq n-1)$ 个行(列)，由这 r 行(列)所组成的一切 r 阶子式与它们的代数余子式的乘积的和等于行列式 $|\boldsymbol{A}|$.用一个等式明确地表述为(只针对列的情形)：

$$|\boldsymbol{A}| = \sum_{k_1 k_2\cdots k_r} \det \boldsymbol{M}_{A}\begin{pmatrix} k_1\ k_2\cdots\ k_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}\cdot \text{代数余子式 } A\begin{pmatrix} k_1\ k_2\cdots\ k_r \\ j_1\ j_2\cdots\ j_r \end{pmatrix}.$$

定理的证明见许以超《代数学引论》(上海科学技术出版社，1966).

定理 3.3 为应用该定理将行列式按第一、二两行展开的特殊情况.

把分块乘法与初等变换结合起来，再用 Laplace 定理，就可以得出如下重要定理的简短的证明.这里，可看到分块乘法在矩阵运算中的简化作用.

定理 3.9　对任意两个 n 阶方阵 $\boldsymbol{A}, \boldsymbol{B}$，恒有 $|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$.

*证　若把 $\boldsymbol{B}$ 的每一列看成一个向量，由向量空间的理论可知，$\boldsymbol{B}$ 的每一列都能用 $\boldsymbol{I}_n$ 的 n 列线性表出.这样

$$\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}_n \end{pmatrix}$$

最多经过 n^2 次倍加变换就能变成

$$\begin{pmatrix} I_n & B \\ O & I_n \end{pmatrix}, 故 \begin{pmatrix} I_n & B \\ O & I_n \end{pmatrix}$$

是若干个初等倍加矩阵的乘积.

$$\begin{pmatrix} A & O \\ -I_n & B \end{pmatrix}\begin{pmatrix} I_n & B \\ O & I_n \end{pmatrix} = \begin{pmatrix} A & AB \\ -I_n & O \end{pmatrix}, \tag{3.13}$$

这表示(3.13)式左端第一个矩阵经过一系列倍加变换之后,变为(3.13)式右端的矩阵,由于倍加变换不改变行列式的值,故有

$$\begin{vmatrix} A & O \\ -I_n & B \end{vmatrix} = \begin{vmatrix} A & AB \\ -I_n & O \end{vmatrix}.$$

用 Laplace 定理将等号两边的行列式都按后 n 列展开,可证

$$|A||B| = |AB|(-1)^{1+2+\cdots+2n}|-I_n| = |AB|(-1)^{2(n^2+n)} = |AB|.\square$$

应特别注意,对任意 n 阶方阵 A 及实数 k, $|kA| = k^n|A|$;如果 $k\neq1$,则 $|kA|\neq k|A|$.

4.2 n 阶方阵的逆矩阵

到目前,已把数的加法和乘法运算“推广”到了矩阵运算.那么自然要问,如何把数的除法运算“推广”到矩阵运算?在数的运算中,每个数 a,只要不是 0,便有一个“逆”a^{-1},它使得 $aa^{-1}=1$.对于除法运算 $b\div a$ 可以用乘法运算 ba^{-1} 或 $a^{-1}b$ 实现.试问,是否每个矩阵 A,只要不为零矩阵,便会有一个矩阵 B,具有性质 $BA = AB = I_n$ 呢?首先,当 A 的行与列数不相等时,就不行;其次,当方阵 A 的第一行都是 0,不管其他行如何,则不存在矩阵 B,使 $BA = AB = I_n$.因为用任何矩阵右乘它,所得的矩阵的第一行总是 0,所以 $AB\neq I_n$.当然,也不是所有的 n 阶方阵都不行,例如 I_n 就行,它的逆就是它自己;当 $a\neq0$ 时, aI_n 的逆就是 $a^{-1}I_n$.与数的情况对照起来说,似乎这样的矩阵才是“正常的”.在应用中,这种矩阵很重要,于是给出下面的定义.

1. 逆矩阵的定义及存在条件

定义3.5 设 A 是一个 n 阶方阵.如果存在 n 阶方阵 B,使 $BA = AB = I_n$.则说 A 是可逆的或非奇异的,并说 B 是 A 的逆阵或逆.否则便说 A 是不可逆的或奇异矩阵.

若方阵 A 是可逆的,设 B, C 都是 A 的逆阵,则有

$$B = BI = B(AC) = (BA)C = IC = C,$$

即 A 的逆阵是唯一的.于是, A 的逆记作 A^{-1}.

例 1 试证明初等矩阵都可逆,并求它们的逆.

证 易验证

$$E(r_i+kr_j)E(r_i+(-k)r_j)=E(r_i+(-k)r_j)E(r_i+kr_j)=I,$$

$j,i=1,2,\cdots,n,i\neq j$.于是 $E(r_i+kr_j)^{-1}=E(r_i+(-k)r_j)$.同理可证

$$E(r_i\longleftrightarrow r_j)^{-1}=E(r_i\longleftrightarrow r_j),$$

$$E(kr_i)^{-1}=E(k^{-1}r_i).$$

设 $A=(a_{ij})$是一方阵,称 n 阶方阵

$$\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

为 A 的伴随阵,记作 A^*,其中 A_{ij}是 a_{ij}的代数余子式$(i,j=1,2,\cdots,n)$.由定理 3.6,易验证

$$AA^*=A^*A=(\delta_{ij}|A|)_{n\times n}=|A|I.$$

定理3.10　n 阶方阵 A 可逆的充要条件是$|A|\neq0$,且当$|A|\neq0$ 时,

$$A^{-1}=\frac{1}{|A|}A^*.$$

证　若 A 可逆,即有 $A^{-1}A=I_n$,则$|A^{-1}||A|=|I_n|=1$,所以$|A|\neq0$.

若$|A|\neq0$,由

$$AA^*=A^*A=|A|I$$

有

$$A\left(\frac{1}{|A|}A^*\right)=\frac{1}{|A|}(AA^*)=\frac{1}{|A|}(\delta_{ij}|A|)=\frac{1}{|A|}|A|I_n=I_n.$$

$$\left(\frac{1}{|A|}A^*\right)A=\frac{1}{|A|}(A^*A)=\frac{1}{|A|}(\delta_{ij}|A|)=\frac{1}{|A|}|A|I_n=I_n.$$

按逆阵的定义,有

$$A^{-1}=\frac{1}{|A|}A^*.\square$$

由定理 3.10,可得下述推论.

推论3.6　若 $AB=I$(或 $BA=I$),则 $B=A^{-1}$.

证　由 $AB=I$,知$|A||B|=|I|$,因而$|A|\neq0$,故 A^{-1}存在,且

$$B=IB=(A^{-1}A)B=A^{-1}(AB)=A^{-1}I_n=A^{-1}.\square$$

当 $A,B,\cdots,C$ 都可逆时,因为 $C^{-1}\cdots B^{-1}A^{-1}AB\cdots C=I$,所以

$$(AB\cdots C)^{-1}=C^{-1}\cdots B^{-1}A^{-1}.$$

例 2　用定理3.10求下列矩阵的逆

$$A=\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}.$$

解

$$A^{-1}=\frac{1}{|A|}A^{*}=\frac{1}{\cos^2\theta+\sin^2\theta}\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

如果一个矩阵其逆等于其转置矩阵,则称该矩阵为正交矩阵.例2中的矩阵 A 就是正交矩阵.

定理3.11 可逆矩阵恒为若干个初等矩阵之积.

证 由定理3.2知,对于任一 n 阶方阵 A,存在初等矩阵 $P_1, P_2, \cdots, P_r$, $Q_1, Q_2, \cdots, Q_s$ 使得

$$P_1 P_2\cdots P_r A Q_1 Q_2\cdots Q_s=\begin{pmatrix}I_k & O\\ O & O\end{pmatrix}_{n\times n}.$$

由于 A 可逆,初等矩阵 P_i, Q_i 也都可逆,所以 $|A|\neq 0$, $|P_i|\neq 0$, $|Q_i|\neq 0$,因此

$$\det\begin{pmatrix}I_k & O\\ O & O\end{pmatrix}_{n\times n}\neq 0,$$

推知 $k=n$,即

$$\begin{pmatrix}I_k & O\\ O & O\end{pmatrix}_{n\times n}=I_n,$$

于是

$$P_1 P_2\cdots P_r A Q_1 Q_2\cdots Q_s=I,$$

得

$$A=P_r^{-1}\cdots P_2^{-1}P_1^{-1}Q_s^{-1}\cdots Q_2^{-1}Q_1^{-1}.$$

由于初等矩阵的逆还是初等矩阵,故上式表明 A 为若干初等矩阵的积.□

该定理说明,除单位矩阵外,初等矩阵是最简单的可逆矩阵,可逆矩阵可分解为初等矩阵乘积.这很像每一个整数都能分解为素数的乘积,不同的是,可逆矩阵的分解不唯一.

2. 逆矩阵的求法

根据定理3.11,可得出一个求逆矩阵的方法.

设 A 可逆,则 A^{-1} 也可逆.对可逆阵 A^{-1},由定理3.11知,有初等矩阵 J, $K, \cdots, L$ 使得 $A^{-1}=JK\cdots L$.考察分块阵 $(A\,|\,I)$,有

$$JK\cdots L(A\,|\,I)=A^{-1}(A\,|\,I)=(A^{-1}A\,|\,A^{-1}I)=(I\,|\,A^{-1}).\tag{3.14}$$

(3.14)式表示:对 $n\times 2n$ 矩阵 $(A\,|\,I)$ 施以若干次初等行变换,当把 $(A\,|\,I)$ 化为 $(I\,|\,B)$ 之时,则 $B=A^{-1}$ 这样就得到一个求逆的方法:写出所给的 n 阶方阵 A,在 A 的右边写下 I_n 而成一个 $n\times 2n$ 矩阵 $(A\,|\,I)$ 然后,对矩阵 $(A\,|\,I)$ 做行的初等变换,其目的是把 A 化成 I_n.当此目的达到时,右边的子块 I_n 就化成 A^{-1}.注

意,只能对矩阵$(A|I)$做行的初等变换,类似地,也可用初等列变换求逆矩阵,即对 $\dfrac{A}{I}$ 实施初等列变换,当 $\dfrac{A}{I}$ 化成 $\dfrac{I}{B}$,则 $B=A^{-1}$.

例 3　求下列矩阵的逆矩阵

$$A=\begin{pmatrix}4&1&2\\3&2&1\\5&-3&2\end{pmatrix}.$$

解

$$\left(\begin{array}{ccc|ccc}4&1&2&1&0&0\\3&2&1&0&1&0\\5&-3&2&0&0&1\end{array}\right)\xrightarrow{r_1+(-1)r_2}\left(\begin{array}{ccc|ccc}1&-1&1&1&-1&0\\3&2&1&0&1&0\\5&-3&2&0&0&1\end{array}\right)$$

$$\xrightarrow[r_3+(-5)r_1]{r_2+(-3)r_1}\left(\begin{array}{ccc|ccc}1&-1&1&1&-1&0\\0&5&-2&-3&4&0\\0&2&-3&-5&5&1\end{array}\right)$$

$$\xrightarrow{r_2+(-2)r_3}\left(\begin{array}{ccc|ccc}1&-1&1&1&-1&0\\0&1&4&7&-6&-2\\0&2&-3&-5&5&1\end{array}\right)$$

$$\xrightarrow[\frac{r_3}{11}]{r_3+(-2)r_2}\left(\begin{array}{ccc|ccc}1&-1&1&1&-1&0\\0&1&4&7&-6&-2\\0&0&1&\frac{19}{11}&-\frac{17}{11}&-\frac{5}{11}\end{array}\right)$$

$$\xrightarrow[\substack{r_1+(-1)r_3\\ r_1+r_2}]{r_2+(-4)r_3}\left(\begin{array}{ccc|ccc}1&0&0&-\frac{7}{11}&\frac{8}{11}&\frac{3}{11}\\0&1&0&\frac{1}{11}&\frac{2}{11}&-\frac{2}{11}\\0&0&1&\frac{19}{11}&-\frac{17}{11}&-\frac{5}{11}\end{array}\right).$$

所以

$$A^{-1}=\begin{pmatrix}-\frac{7}{11}&\frac{8}{11}&\frac{3}{11}\\\frac{1}{11}&\frac{2}{11}&-\frac{2}{11}\\\frac{19}{11}&-\frac{17}{11}&-\frac{5}{11}\end{pmatrix}.$$

用分块矩阵求矩阵的逆矩阵,可将高阶矩阵求逆的问题转化为低阶矩阵的求逆,能较大地简化运算 .

例 4　设 m 阶方阵 A 和 n 阶方阵 B 均可逆,求下列矩阵的逆,其中 O 为零阵 .

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}.$$

解 对如下矩阵,实施初等行变换:

$$\left(\begin{array}{cc|cc} \boldsymbol{A} & \boldsymbol{C} & \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} & \boldsymbol{O} & \boldsymbol{I}_n \end{array}\right) \xrightarrow{r_1 + (-\boldsymbol{C}\boldsymbol{B}^{-1}) r_2} \left(\begin{array}{cc|cc} \boldsymbol{A} & \boldsymbol{O} & \boldsymbol{I}_m & -\boldsymbol{C}\boldsymbol{B}^{-1} \\ \boldsymbol{O} & \boldsymbol{B} & \boldsymbol{O} & \boldsymbol{I}_n \end{array}\right)$$

$$\xrightarrow{\boldsymbol{A}^{-1} r_1,\ \boldsymbol{B}^{-1} r_2} \left(\begin{array}{cc|cc} \boldsymbol{I}_m & \boldsymbol{O} & \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{B}^{-1} \\ \boldsymbol{O} & \boldsymbol{I}_n & \boldsymbol{O} & \boldsymbol{B}^{-1} \end{array}\right).$$

由上式可见

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{B}^{-1} \\ \boldsymbol{O} & \boldsymbol{B}^{-1} \end{pmatrix}.$$

用分块乘法验证

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{C}\boldsymbol{B}^{-1} \\ \boldsymbol{O} & \boldsymbol{B}^{-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I}_n \end{pmatrix}.$$

4.3 矩阵的秩

通过矩阵子块的行列式可以揭示矩阵的一个很重要的内在特性,即所谓矩阵的秩 .设 $\boldsymbol{A}$ 是任意一个矩阵(不一定是方阵),当 $\boldsymbol{A}$ 的所有元素都为零时, $\boldsymbol{A}$ 的任何子块的行列式,即子式,都必然是零;当 $\boldsymbol{A}$ 有不为零的元素时, $\boldsymbol{A}$ 中就至少有一个一阶子式非零;再往下看,如果 A 中还有一个 2 阶子式非零,则继续再看下去;最后,必到某一步,此时 A 中至少有一个 r 阶子式非零,但 $\boldsymbol{A}$ 中不再有 r 阶以上的子式非零 .这时 r 就反映了矩阵 A 的一个内在特性,即下面矩阵的秩的概念 .

定义3.6 如果矩阵 $\boldsymbol{A}$ 的所有元素都为零时,则说 $\boldsymbol{A}$ 的秩为零;如果 $\boldsymbol{A}$ 不是零矩阵, $\boldsymbol{A}$ 中非零的子式的最大阶数为 r,则说 $\boldsymbol{A}$ 的秩为 r . $\boldsymbol{A}$ 的秩记作 $R(\boldsymbol{A})$.

定义3.7 若 n 阶方阵的秩是 n,则称该矩阵为满秩的 .

命题3.1 n 阶方阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $R(\boldsymbol{A}) = n$,即 $\boldsymbol{A}$ 可逆的充要条件是 $\boldsymbol{A}$ 为满秩矩阵 .

可直接由定义证明 .

例 5 求下列矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩,其中

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 4 & 2 & 0 \end{pmatrix},\ \boldsymbol{B} = \begin{pmatrix} 5 & 4 & 0 & 1 & 3 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

解　在 $\boldsymbol{A}$ 中,容易看出它的二阶子式

$$\begin{vmatrix}1 & 2\\ 0 & 2\end{vmatrix} = 2 \neq 0 .$$

由于 $\boldsymbol{A}$ 的第三行与第一行的元素对应成比例,所以 $\boldsymbol{A}$ 的所有三阶子式都等于零,于是 $R(\boldsymbol{A}) = 2$.$\boldsymbol{B}$ 称为行阶梯形矩阵,其非零行有三行,即知其所有 4 阶子式全为零 .而三阶子式

$$\begin{vmatrix}5 & 4 & 1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{vmatrix} = 5 \neq 0,$$

因此 $R(\boldsymbol{B}) = 3$.

一般地,行阶梯形矩阵形如

$$\boldsymbol{A} = \begin{pmatrix}
0 & \cdots & 0 & a_{1i_1} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & \cdots & 0 & a_{2i_2} & \cdots & \cdots & \cdots & \cdots & \cdots & * \\
\cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & * \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & a_{ri_r} & \cdots & \cdots & \cdots & * \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0
\end{pmatrix},$$

其特点是:可以化出一条阶梯线,阶梯线下方的元素全为零,每个台阶为一行,台阶上方第一个元素为各非零行的非零首元素 .行阶梯形矩阵 $\boldsymbol{A}$ 的秩等于非零行的行数 r,这是因为 $\boldsymbol{A}$ 的所有阶数大于 r 的子式全为 0,而以各行非零首元素为对角线元素的 r 阶子式

$$\det \boldsymbol{M}_{\boldsymbol{A}}\begin{matrix}12\cdots r\\ i_1\ i_2\cdots i_r\end{matrix} = a_{1i_1}a_{2i_2}\cdots a_{ri_r} \neq 0 .$$

定理3.12　对矩阵 $\boldsymbol{A}$ 作任何一种初等变换其秩不变 .

证　将矩阵的某一行(列)乘一非零常数或对换某两行(列)后,其非零的子式的最大阶数不变,所以秩不变 .下面证明,将矩阵的某一行(列)乘一常数加到另外一行(列)后秩不变 .设 $R(\boldsymbol{A}) = r$,并设 r 阶子块

$$\boldsymbol{M}_{\boldsymbol{A}}\begin{matrix}i_1\ i_2\cdots i_r\\ j_1\ j_2\cdots j_r\end{matrix}$$

的行列式非零 .将 $\boldsymbol{A}$ 第 k 行乘常数 a 加于第 h 行($k \neq h$)变成 $\boldsymbol{B}$,则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 除第 h 行不同外其他行都相同 .若 $h \notin \{i_1, i_2, \cdots, i_r\}$,则

$$\det M_B\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} = \det M_A\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} \neq 0 .$$

若 $h, k\in\{i_1, i_2, \cdots, i_r\}$,将

$$M_A\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix}$$

第 k 行乘常数 a 加于第 h 行化为

$$M_B\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} .$$

由行列式性质,有

$$\det M_B\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} = \det M_A\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} \neq 0 .$$

若 $h\in\{i_1, i_2, \cdots, i_r\}$, $k\notin\{i_1, i_2, \cdots, i_r\}$,不妨设 $h = i_1$,由行列式性质,有

$$\det M_B\begin{pmatrix} h i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} = \det M_A\begin{pmatrix} h i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} + a\det M_A\begin{pmatrix} k i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} .$$

若 $\det M_B\begin{pmatrix} h i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} = 0$,由 $\det M_A\begin{pmatrix} i_1\ i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} \neq 0$,有 $\det M_A\begin{pmatrix} k i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} \neq 0$.又因为 $h\notin\{k, i_2, \cdots i_r\}$,所以

$$\det M_B\begin{pmatrix} k i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} = \det M_A\begin{pmatrix} k i_2\cdots i_r \\ j_1\ j_2\cdots j_r \end{pmatrix} \neq 0 .$$

综上所述,若 $\boldsymbol{A}$ 有一个非零的 r 阶子式,则 $\boldsymbol{B}$ 至少有一个非零的 r 阶子式 .由秩的定义有 $R(\boldsymbol{A})\leq R(\boldsymbol{B})$.将 $\boldsymbol{B}$ 第 k 行乘 $-a$ 加于第 h 行后 $\boldsymbol{B}$ 化成 $\boldsymbol{A}$,同理可推知 $R(\boldsymbol{B})\leq R(\boldsymbol{A})$,故 $R(\boldsymbol{B}) = R(\boldsymbol{A})$.对于列,用类似的方法也可证 $R(\boldsymbol{B}) = R(\boldsymbol{A})$.□

定理 3 .12 说明,有限次初等变换不改变矩阵的秩,即矩阵的秩是矩阵的初等变换下的不变量 .

例 6　求下列矩阵 $\boldsymbol{A}$ 的秩

$$\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ -1 & 2 & 1 & 3 & 6 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix} .$$

解　对 $\boldsymbol{A}$ 作初等行变换,将其化为行阶梯形矩阵 .

$$\begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ -1 & 2 & 1 & 3 & 6 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix} \xrightarrow{r_2 + r_1} \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix}$$

$$\xrightarrow[r_4+r_2]{r_3+(-1)r_1}\begin{pmatrix}1&-1&0&-1&-2\\0&1&1&2&4\\0&0&0&0&0\\0&0&0&3&6\end{pmatrix}\xrightarrow[r_3\leftrightarrow r_4]{\frac{1}{3}r_4}\begin{pmatrix}1&-1&0&-1&-2\\0&1&1&2&4\\0&0&0&1&2\\0&0&0&0&0\end{pmatrix}=\boldsymbol{B}.$$

由行阶梯形矩阵 $\boldsymbol{B}$ 有 3 个非零行,知 $R(\boldsymbol{B})=3$.由定理 3.12 知

$$R(\boldsymbol{A})=R(\boldsymbol{B})=3.$$

下面的定理,揭示了矩阵的秩与矩阵的行向量组和列向量组的秩之间的关系,它把表面似乎不相关的事联系起来.

定理3.13 $m\times n$ 矩阵 $\boldsymbol{A}$ 的行向量组的秩与列向量组的秩相等,且等于 $\boldsymbol{A}$ 的秩.

证 设 $R(\boldsymbol{A})=r$.作行初等变换把 $\boldsymbol{A}$ 变成行阶梯形矩阵 $\boldsymbol{B}$,则 $\boldsymbol{B}$ 恰有 r 个非零行,且此 r 个非零行向量线性无关.因为若不然,则此 r 行中必有一行是另 $r-1$ 行的线性组合,从而该行经倍加变换可化为零行,这与 $\boldsymbol{B}$ 是有 r 个非零行的阶梯形矩阵矛盾.因此,$\boldsymbol{B}$ 的行向量组的秩为 r,由于 $\boldsymbol{A}$ 的行向量组与 $\boldsymbol{B}$ 的行向量组等价,因此,$\boldsymbol{A}$ 的行向量组的秩也为 r.

又因为

$$R(\boldsymbol{A}^{\mathrm{T}})=R(\boldsymbol{A})=r,$$

由上证明知 $\boldsymbol{A}^{\mathrm{T}}$ 的行向量组的秩为 r,即 $\boldsymbol{A}$ 的列向量组的秩为 r.□

推论3.7 n 阶矩阵 $\boldsymbol{A}$ 的行(列)向量组线性无关的充要条件是 $|\boldsymbol{A}|\neq 0$.

可直接用定理 3.13 证明.

推论3.8 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别为 $m\times p$ 和 $p\times n$ 阶矩阵,则

$$R(\boldsymbol{AB})\leq\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}.$$

证 由本章第二节例 2 可知,$\boldsymbol{AB}$ 的每一列向量都是 $\boldsymbol{A}$ 的列向量的线性组合,由定理 3.13 可知,矩阵的秩等于其列向量的秩.由于 $\boldsymbol{AB}$ 的列向量可被 $\boldsymbol{A}$ 的列向量线性表示,所以,$R(\boldsymbol{AB})\leq R(\boldsymbol{A})$;同时 $\boldsymbol{AB}$ 的每一行向量都是 $\boldsymbol{B}$ 的行向量的线性组合,同理,可知 $R(\boldsymbol{AB})\leq R(\boldsymbol{B})$.故

$$R(\boldsymbol{AB})\leq\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}.$$

习 题 三

习题 A

1.计算下列矩阵乘积:

(1) $\begin{pmatrix}4&3&1&7\\1&-2&3&2\\5&7&0&1\end{pmatrix}$;

(2) $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$;

(3) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n$;

(4) $\begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}\begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}$;

(5) $\begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$;

(6) $\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$;

(7) $\begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

2．设 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$,问:

(1) $\boldsymbol{AB} = \boldsymbol{BA}$ 吗?

(2) $(\boldsymbol{A} + \boldsymbol{B})^2 = \boldsymbol{A}^2 + 2\boldsymbol{AB} + \boldsymbol{B}^2$ 吗?

(3) $(\boldsymbol{A} + \boldsymbol{B})(\boldsymbol{B} - \boldsymbol{A}) = \boldsymbol{A}^2 - \boldsymbol{B}^2$ 吗?

3．举例说明下列命题是错误的:

(1) 若 $\boldsymbol{A}^2 = \boldsymbol{O}$,则 $\boldsymbol{A} = \boldsymbol{O}$;

(2) 若 $\boldsymbol{A}^2 = \boldsymbol{A}$,则 $\boldsymbol{A} = \boldsymbol{O}$ 或 $\boldsymbol{A} = \boldsymbol{I}$;

(3) 若 $\boldsymbol{AX} = \boldsymbol{AY}$ 且 $\boldsymbol{A} \neq \boldsymbol{O}$,则 $\boldsymbol{X} = \boldsymbol{Y}$.

4．证明:若 $\boldsymbol{A}$, $\boldsymbol{B}$ 皆为 n 阶下三角形矩阵,则 $\boldsymbol{AB} = \boldsymbol{C}$ 也是下三角形矩阵,且 $c_{ii} = a_{ii}b_{ii}$, $i = 1, 2, \cdots, n$.

5．设 $\boldsymbol{A}$, $\boldsymbol{B}$ 都为 n 阶对称矩阵,证明:

(1) $\boldsymbol{A} + \boldsymbol{B}$, $\boldsymbol{A} - 2\boldsymbol{B}$ 也都为对称矩阵;

(2) $\boldsymbol{AB}$ 是对称矩阵当且仅当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换(即 $\boldsymbol{AB} = \boldsymbol{BA}$).

6．设 $\boldsymbol{A}$ 为 n 阶方阵,证明:

(1) $\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}$ 是对称矩阵, $\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}$ 是反对称矩阵(如果 $\boldsymbol{A} = -\boldsymbol{A}^{\mathrm{T}}$,则 $\boldsymbol{A}$ 称为

反对称矩阵);

(2) $\boldsymbol{A}$ 可表示为对称矩阵与反对称矩阵之和 .

7 . 矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的主对角元之和称为矩阵 $\boldsymbol{A}$ 的迹,记作

$$\operatorname{tr}(\boldsymbol{A})=\sum_{i=1}^{n}a_{ii}.$$

证明:(1) $\operatorname{tr}(\boldsymbol{A}_{n\times m}\boldsymbol{B}_{m\times n})=\operatorname{tr}(\boldsymbol{B}_{m\times n}\boldsymbol{A}_{n\times m})$.

(2) 对于任意 n 阶方阵 $\boldsymbol{A},\boldsymbol{B}$,恒有 $\boldsymbol{AB}-\boldsymbol{BA}\neq\boldsymbol{I}$.

8 . 如果方阵 $\boldsymbol{A}$ 非奇异,则 $\boldsymbol{A}^{\mathrm{T}}$ 亦然,且 $(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$.

9 . 对任意矩阵 $\boldsymbol{A}$,则 $\boldsymbol{AA}^{\mathrm{T}}$ 恒有意义且为对称矩阵 .

10 . 对任意实矩阵 $\boldsymbol{A}$,若 $\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{O}$,则 $\boldsymbol{A}=\boldsymbol{O}$.

11 . 证明:若 n 阶矩阵 $\boldsymbol{A}$ 具有性质 $\boldsymbol{A}^{k}=\boldsymbol{O}$,则 $\boldsymbol{I}-\boldsymbol{A}$ 必为非奇异矩阵,并求其逆 .(提示:模仿 $1-x^{k}$ 的因式分解) .

12 .设

$$\boldsymbol{A}=\begin{pmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{pmatrix},$$

求 $\boldsymbol{A}^{k}$, k 为正整数 .

13 . 计算下列行列式

(1) $\begin{vmatrix}4&1&2&4\\1&2&0&2\\10&5&2&0\\0&1&1&7\end{vmatrix}$;

(2) $\begin{vmatrix}0&1&0&0\\1&0&1&0\\0&1&0&1\\0&0&1&0\end{vmatrix}$;

(3) $\begin{vmatrix}x&y&x+y\\y&x+y&x\\x+y&x&y\end{vmatrix}$;

(4) $\begin{vmatrix}1&\omega&\omega^{2}\\\omega^{2}&1&\omega\\\omega&\omega^{2}&1\end{vmatrix}$, 其中 $\omega=\dfrac{-1+\ 3\,i}{2}$.

14 . 计算下列行列式

(1) $D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \cdots & \cdots & & \cdots \\ a & a & \cdots & x \end{vmatrix}$;

(2) $\begin{vmatrix} a & b & c & d \\ a & a+b & a+b+c & a+b+c+d \\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d \\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d \end{vmatrix}$;

(3) $D_{2n} = \begin{vmatrix} a_n & & & & & b_n \\ & \ddots & & & \iddots & \\ & & a_1 & b_1 & & \\ & & c_1 & d_1 & & \\ & \iddots & & & \ddots & \\ c_n & & & & & d_n \end{vmatrix}$,未标注的元均为 0;

(4) $D_n = \Delta(a_{ij})$,其中 $a_{ij} = |i - j|$;

(5) $D_n = \begin{vmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \cdots & 1 \\ \cdots & \cdots & & \cdots \\ 1 & 1 & \cdots & 1+a_n \end{vmatrix}$, $a_1 a_2 \cdots a_n \neq 0$.

15 . **证明**:

(1) $D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \cdots & \cdots & & \cdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{n \geq i > j \geq 1} (x_i - x_j)$,

该行列式称为 Vandermonde **行列式**;

(2) **设** $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$ **为** n **次多项式** .$x_1, x_2, \cdots, x_n$ **为** n **个两两不同的实数**, $y_1, y_2, \cdots, y_n$ **为** n **个实数** .**用** Vandermonde **行列式及** Cramer **法则证明**:**存在唯一的** n **次多项式** $p_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$, **使得** $y_i = p_n(x_i)$, $i = 1, 2, \cdots, n$;

(3) $D_n = \begin{vmatrix} x & -1 & 0 & \cdots & 0 \\ 0 & x & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & & \cdots \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_1 \end{vmatrix} = a_1 x^{n-1} + \cdots a_{n-1} x + a_n$;

(4) $D_n=\begin{vmatrix} a_1 & -1 & 0 & \cdots & 0 \\ a_2 & x & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & & \cdots \\ a_n & 0 & 0 & \cdots & x \end{vmatrix}=\sum_{k=1}^{n} a_k x^{n-k}$;

(5) $\begin{vmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{vmatrix}=n+1$;

16 . **用克拉默法则解方程组**

$$\begin{aligned} x_1+x_2+x_3+x_4&=5, \\ x_1+2x_2-x_3+4x_4&=-2, \\ 2x_1-3x_2-x_3-5x_4&=-2, \\ 3x_1+x_2+2x_3+11x_4&=0 . \end{aligned}$$

17 . **问 λ、μ 取何值时,下列齐次方程组有非零解**

$$\begin{aligned} \lambda x_1+x_2+x_3&=0, \\ x_1+\mu x_2+x_3&=0, \\ x_1+2\mu x_2+x_3&=0 . \end{aligned}$$

18 . **设** $A(x_1, y_2)$, $B(x_2, y_2)$**是平面上不同的点 .试证:过** A **和** B **的直线方程是**

$$\begin{vmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix}=0 .$$

19 . **在 3 维欧氏几何空间** $\boldsymbol{R}^3$ **中,向量** $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$, $\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k}$.**试证**

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} .$$

20 . **在 3 维欧氏几何空间** $\boldsymbol{R}^3$ **中**, $\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$, $\boldsymbol{b}=b_x\boldsymbol{i}+b_y\boldsymbol{j}+b_z\boldsymbol{k}$, $\boldsymbol{c}=c_x\boldsymbol{i}+c_y\boldsymbol{j}+c_z\boldsymbol{k}$.

试证:混合积

$$[\boldsymbol{abc}]=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}=\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} .$$

21．设 $L_1: \alpha x+\beta y+\gamma=0$，$L_2: \gamma x+\alpha y+\beta=0$，$L_3: \beta x+\gamma y+\alpha=0$ 是三条不同的直线，若 L_1，L_2，L_3 交于一点，试证：$\alpha+\beta+\gamma=0$．

22．对于任意 n 阶方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，试证：

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{vmatrix} = |\boldsymbol{A}+\boldsymbol{B}|\,|\boldsymbol{A}-\boldsymbol{B}|.$$

23．假定 $\boldsymbol{A}$，$\boldsymbol{B}$ 是两个 n 阶矩阵，试证 $R(\boldsymbol{AB})=R(\boldsymbol{B})$ 的充要条件是 $\boldsymbol{ABx}=\boldsymbol{O}$，$\boldsymbol{Bx}=\boldsymbol{O}$ 是同解方程．

24．假定 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 是 n 个线性无关的 n 维向量，$\boldsymbol{P}$ 为 n 阶方阵，证明：$\boldsymbol{Px}_1, \boldsymbol{Px}_2, \cdots, \boldsymbol{Px}_n$ 线性无关的充要条件是 $\boldsymbol{P}$ 是满秩矩阵．

25．证明：任意可逆阵都可分解为一个下三角形阵和一个正交阵之积（对角线以上的元都为零的方阵称为下三角形阵，满足 $\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{I}$ 的方阵 $\boldsymbol{A}$ 称为正交矩阵）．

26．求下列方阵的逆阵：

(1) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$；

(3) $\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$，$a_1 a_2 \cdots a_n \neq 0$．

27．设 $\boldsymbol{AB}=\boldsymbol{A}+2\boldsymbol{B}$，

(1) $\boldsymbol{A}=\begin{pmatrix} 3 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 4 \end{pmatrix}$，求 $\boldsymbol{B}$；

(2) $\boldsymbol{A}=\begin{pmatrix} 4 & 2 & 3 \\ 1 & 1 & 0 \\ -1 & 2 & 3 \end{pmatrix}$，求 $\boldsymbol{B}$．

28．用矩阵分块乘法及初等变换求下列方阵的逆，其中 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是可逆阵．

(1) $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$，$\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}$，$\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{X} \\ \boldsymbol{O} & \boldsymbol{I}_n \end{pmatrix}$，$\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{O} \\ \boldsymbol{Y} & \boldsymbol{I}_n \end{pmatrix}$，$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$；

(2) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 4 \end{pmatrix}$，$\begin{pmatrix} 5 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 8 & 3 \\ 0 & 0 & 5 & 2 \end{pmatrix}$，$\begin{pmatrix} 0 & 0 & 5 & 2 \\ 0 & 0 & 2 & 1 \\ 8 & 3 & 0 & 0 \\ 5 & 2 & 0 & 0 \end{pmatrix}$．

29．解下列矩阵方程：

(1) $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \boldsymbol{X} = \begin{pmatrix} 4 & -6 \\ 2 & 1 \end{pmatrix}$；

(2) $\boldsymbol{X} \begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 4 & 3 & 2 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & 4 \\ -1 & 2 \end{pmatrix} \boldsymbol{X} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 0 \\ 1 & 0 & 3 \end{pmatrix}$；

(4) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{X} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -4 & 3 \\ 2 & 0 & -1 \\ 1 & -2 & 0 \end{pmatrix}$．

30．利用逆阵解下列线性方程组：

(1) $\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ 2x_1 + 2x_2 + 5x_3 = 2, \\ 3x_1 + 5x_2 + x_3 = 3; \end{cases}$ (2) $\begin{cases} x_1 - x_2 - x_3 = 2, \\ 2x_1 - x_2 - 3x_3 = 1, \\ 3x_1 + 2x_2 - 5x_3 = 0. \end{cases}$

31．设

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \Lambda, \boldsymbol{P} = \begin{pmatrix} -1 & -4 \\ 1 & 1 \end{pmatrix}, \Lambda = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix},$$

求 $\boldsymbol{A}^{11}$．

32．设 m 次多项式 $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$，记 $f(\boldsymbol{A}) = a_0 \boldsymbol{I} + a_1 \boldsymbol{A} + a_2 \boldsymbol{A}^2 + \cdots + a_m \boldsymbol{A}^m$，$f(\boldsymbol{A})$ 称为方阵 $\boldsymbol{A}$ 的 m 次多项式．

(1) 设

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

证明：

$$f(\Lambda) = \begin{pmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{pmatrix};$$

(2) 设 $\boldsymbol{A} = \boldsymbol{PBP}^{-1}$，证明：$f(\boldsymbol{A}) = \boldsymbol{P}f(\boldsymbol{B})\boldsymbol{P}^{-1}$．

33．设 $\boldsymbol{A}^*$ 为 n 阶方阵 $\boldsymbol{A}$ 的伴随阵，证明：

(1) 若 $|\boldsymbol{A}| = 0$，则 $|\boldsymbol{A}^*| = 0$；

(2) $|\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}$．

34．设 n 阶非零实方阵 $\boldsymbol{A}$ 的伴随阵为 $\boldsymbol{A}^*$ 且 $\boldsymbol{A}^* = A^{\mathrm{T}}$，证明：$|\boldsymbol{A}| \neq 0$．

35．设

$$A = B = -C = D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

验证

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} \neq \begin{vmatrix} |A| & |B| \\ |C| & |D| \end{vmatrix}.$$

36．求下列矩阵的秩:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 & 0 & 5 \\ 3 & 4 & 2 & 0 & 0 \\ 0 & 0 & 7 & 0 & 9 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

37．$\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{B}$ 是 $n\times m$ 矩阵, $m>n$,证明: $|\boldsymbol{AB}|=0$.

38．$\boldsymbol{A}$ 是 n 阶方阵$(n\geq 2)$,证明:

$$R(\boldsymbol{A}^*) = \begin{cases} n, & R(\boldsymbol{A}) = n; \\ 0, & R(\boldsymbol{A}) \leq n-2; \\ 1, & R(\boldsymbol{A}) = n-1. \end{cases}$$

39．如果 $R(\boldsymbol{A})=r$, $R(\boldsymbol{B})=s$,则

$$R\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix} = r+s.$$

40．设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵.证明: $R(\boldsymbol{A})\leq r$ 当且仅当存在 $m\times r$ 矩阵 $\boldsymbol{B}$ 与 $r\times n$矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{A}=\boldsymbol{BC}$.

41．证明: $R(\boldsymbol{A}+\boldsymbol{B})\leq R(\boldsymbol{A})+R(\boldsymbol{B})$.

42．证明: $R(\boldsymbol{A},\boldsymbol{B})\leq R(\boldsymbol{A})+R(\boldsymbol{B})$.

习题 B

1．设 $\boldsymbol{A}$ 为一个 n 阶方阵,对于任一个 n 阶方阵 $\boldsymbol{B}$ 都有 $\boldsymbol{AB}=\boldsymbol{BA}$ 的充分必要条件是 $\boldsymbol{A}=a\boldsymbol{I}$(此类对角线上的所有元素都等于同一数 a 而对角线以外的其他元素都为零的矩阵称为纯量矩阵).

2．设 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$, $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$,且 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=2$,证明:

(1) $(\boldsymbol{xy}^{\mathrm{T}})^k=2^{k-1}\boldsymbol{xy}^{\mathrm{T}}$, $k\geq 2$;

(2) 如果 $\boldsymbol{A}=\boldsymbol{I}+\boldsymbol{xy}^{\mathrm{T}}$,则 $\boldsymbol{A}$ 可逆,并求其逆矩阵.

3．计算行列式

$$\begin{vmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \cdots & \cdots & & \cdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{vmatrix}.$$

4．证明：

$$\begin{vmatrix} \cos\theta & 1 & & & & \\ 1 & 2\cos\theta & 1 & & & \\ & 1 & 2\cos\theta & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 2\cos\theta & 1 \\ & & & & 1 & 2\cos\theta \end{vmatrix} = \cos\theta .$$

5．对列的情形证明推论 3.2，推论 3.3，推论 3.4，推论 3.5．

6．求满足 $\boldsymbol{A}+a\boldsymbol{B}=\begin{pmatrix}3&3\\2&4\end{pmatrix}$，$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{I}$，$\boldsymbol{AB}=\boldsymbol{O}$ 的数 a 及二阶矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$．并证明：$\boldsymbol{A}+a^n\boldsymbol{B}=(\boldsymbol{A}+a\boldsymbol{B})^n$．

7．假设 $\boldsymbol{A}$ 为对角形矩阵，它的对角线上元互不相等，$\boldsymbol{B}$ 是满足 $\boldsymbol{AB}=\boldsymbol{BA}$ 的任意矩阵，证明：$\boldsymbol{B}$ 是对角形矩阵．

8．设 $\boldsymbol{A}$ 为 n 阶方阵．当 $\boldsymbol{AA}^{\mathrm{T}}=\boldsymbol{I}$ 时，称 $\boldsymbol{A}$ 为正交矩阵．假定 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是正交矩阵，并且 $|\boldsymbol{A}|=-|\boldsymbol{B}|$．证明：$|\boldsymbol{A}+\boldsymbol{B}|=0$．

9．假定 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 n 阶矩阵，如果 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$，证明：$\boldsymbol{AB}=\boldsymbol{BA}$．

10．设 $a_n\boldsymbol{A}^n+a_{n-1}\boldsymbol{A}^{n-1}+\cdots+a_0\boldsymbol{I}=\boldsymbol{O}$，其中 $\boldsymbol{A}$ 为 n 阶方阵，$a_n, a_{n-1},\cdots, a_0$ 为实数．(1) 当 $a_0\neq 0$ 时，证明：$\boldsymbol{A}$ 可逆；(2) 设方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2-\boldsymbol{A}-2\boldsymbol{I}=\boldsymbol{O}$，证明：当 $k\neq 2$ 和 -1 时，$\boldsymbol{A}+k\boldsymbol{I}$ 及 $\boldsymbol{A}$ 都可逆，并求 $\boldsymbol{A}^{-1}$ 及 $(\boldsymbol{A}+k\boldsymbol{I})^{-1}$．

11．设 $\boldsymbol{a}$ 为 n 维非零列向量且 $\boldsymbol{A}=\boldsymbol{I}-\boldsymbol{aa}^{\mathrm{T}}$，证明：

(1) $\boldsymbol{A}^2=\boldsymbol{A}$ 的充要条件是 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}=1$；

(2) 当 $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}=1$ 时，$\boldsymbol{A}$ 不可逆．

12．设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都为 n 阶方阵．证明：

(1) 当 $\boldsymbol{A}$ 为非奇异矩阵时，$R(\boldsymbol{AB})=R(\boldsymbol{BA})=R(\boldsymbol{B})$；

(2) $R(\boldsymbol{AB})\geq R(\boldsymbol{A})+R(\boldsymbol{B})-n$．

提示：

$$R(\boldsymbol{A})+R(\boldsymbol{B})\leq R\begin{pmatrix}\boldsymbol{B}&\boldsymbol{I}\\\boldsymbol{O}&\boldsymbol{A}\end{pmatrix}=R\left[\begin{pmatrix}\boldsymbol{I}&\boldsymbol{O}\\-\boldsymbol{A}&\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{B}&\boldsymbol{I}\\\boldsymbol{O}&\boldsymbol{A}\end{pmatrix}\right]$$

$$=R\begin{pmatrix}\boldsymbol{B}&\boldsymbol{I}\\-\boldsymbol{AB}&\boldsymbol{O}\end{pmatrix}=R\left[\begin{pmatrix}\boldsymbol{B}&\boldsymbol{I}\\-\boldsymbol{AB}&\boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}&\boldsymbol{O}\\-\boldsymbol{B}&\boldsymbol{I}\end{pmatrix}\right].$$

13．设 n 阶矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{I}$．证明：$R(\boldsymbol{I}-\boldsymbol{A})+R(\boldsymbol{I}+\boldsymbol{A})=n$．(提示：由 $(\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{I}+\boldsymbol{A})=\boldsymbol{O}$ 和 $(\boldsymbol{I}-\boldsymbol{A})+(\boldsymbol{I}+\boldsymbol{A})=2\boldsymbol{I}$，再由习题 B12 可证明之)．

14．设 n 矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{A}$．证明：$R(\boldsymbol{I}-\boldsymbol{A})+R(\boldsymbol{A})=n$．

15．设矩阵 $\boldsymbol{A}=(a_{ij})$ 的元 a_{ij} 是 t 的可微函数，定义 $\boldsymbol{A}$ 的导数 $\dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}$ 是以 $\dfrac{\mathrm{d}a_{ij}}{\mathrm{d}t}$ 为

元的矩阵,即$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}=\left(\frac{\mathrm{d}a_{ij}}{\mathrm{d}t}\right)$.证明:下列微分规则成立:

(1) $\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}+\boldsymbol{B})=\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}+\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}$;

(2) $\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}\boldsymbol{B})=\left(\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\right)\boldsymbol{B}+\boldsymbol{A}\left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)$.

16. 设

$$\boldsymbol{A}=\begin{pmatrix} t^2 & t^3 \\ t & \sin t \end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix} \mathrm{e}^{2t} & 0 \\ \cos t & 4 \end{pmatrix},$$

求$\frac{\mathrm{d}(\boldsymbol{A}\boldsymbol{B})}{\mathrm{d}t}$.

17. 设矩阵 $\boldsymbol{A}=(a_{ij})$ 的元 a_{ij} 是 t 的可积函数,定义 $\boldsymbol{A}$ 的积分$\int \boldsymbol{A}\mathrm{d}t$是以$\int a_{ij}\mathrm{d}t$为元的矩阵,即$\int \boldsymbol{A}\mathrm{d}t=\left(\int a_{ij}\mathrm{d}t\right)$.证明下列积分性质成立:

(1) $\int(\boldsymbol{A}+\boldsymbol{B})\mathrm{d}t=\int \boldsymbol{A}\mathrm{d}t+\int \boldsymbol{B}\mathrm{d}t$.

(2) $\int_{t_1}^{t_2}(\boldsymbol{A}+\boldsymbol{B})\mathrm{d}t=\int_{t_1}^{t_2}\boldsymbol{A}\mathrm{d}t+\int_{t_1}^{t_2}\boldsymbol{B}\mathrm{d}t$.

18. 设

$$\boldsymbol{A}=\begin{pmatrix} t^3 & \mathrm{e}^{-t} \\ \sin t & \cos t \end{pmatrix};\ \boldsymbol{B}=\begin{pmatrix} 1 & \sin 2t \\ \mathrm{e}^{t} & \cos t \end{pmatrix},$$

计算$\int \boldsymbol{A}\mathrm{d}t$,$\int_0^1 \boldsymbol{B}\mathrm{d}t$.

第四章　线性方程组

解线性方程组是线性代数的主要内容之一, Gauss 消元法是解线性方程组的基本的有效方法 .这一章首先讲述 Gauss 消元法及其相关问题;其次讨论齐次线性方程组及非齐次线性方程组有解的充要条件及其解的结构 .

第一节　Gauss 消元法

大约 1820 年, 德国数学家、天文学家、物理学家 C.F.Gauss(1777—1855)为求解天文学和土地测量计算中出现的线性方程组,系统地提出了 Gauss 消元法 .应该特别指出的是,解线性方程组的消元法思想由来已久 .《九章算术》是中国古代一本很有名的算经,大约在秦朝之前就有这本书的原本, 后来由西汉(公元前 206 年—公元 25 年)张苍及耿其昌整理成书 .《九章算术》的第八章《方程》中, 给出了线性方程组的普遍解法,除符号与术语不同外, 实质和 Gauss 消元法是一样的(梁宗巨 .世界数学史简编 .沈阳: 辽宁人民出版社, 1980) .

1.1　Gauss 消元法

Gauss 消元法是解线性方程组一个重要的直接方法, 在不考虑舍入误差情况下, 经过有限步四则运算, 可求出准确解 .消元法的基本想法是用统一算法, 逐次消去方程组的未知数, 把原方程组化为同解的三角形方程组再求解 .Gauss 消元法, 特别是经过改进的主元素消元法, 是目前计算机解线性方程组常用的方法之一 .

首先讨论由 n 个未知数、n 个方程构成的线性方程组:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n . \end{aligned} \tag{4.1}$$

若用矩阵表示, 方程组(4.1)可写为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{4.2}$$

其中 $\boldsymbol{A} = (a_{ij})_{n\times n}$, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 是未知向量, $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$ 是(4.2)右端向量 .如果 $\boldsymbol{b}$ 是零向量, (4.1)称为齐次线性方程组; 否则, (4.1)称为

非齐次线性方程组.矩阵 A 与 $(A|b)$ 分别称为线性方程组(4.1)的系数矩阵与增广矩阵.若存在向量 x_0 满足线性方程组(4.2),即 $Ax_0=b$,则称 $x=x_0$ 为(4.2)的解向量.线性方程组可能没有解向量,有唯一解向量或有无穷多个解向量.

下面以一个实例说明 Gauss 消元法的步骤.这个方程组的方程个数与未知量个数相等,且方程组存在唯一的解向量.

例 1 解线性方程组

$$\begin{cases} 2x_1+x_2+x_3=3, & (r_1)\\ 4x_1+x_2+x_3=5, & (r_2)\\ -2x_1+x_2+3x_3=-3. & (r_3)\end{cases} \tag{4.3}$$

解 由于方程 r_1 中 x_1 的系数不为零,可利用方程 r_1,消去 r_2 及 r_3 中的 x_1:方程 r_2 减去方程 r_1 的 $\frac{4}{2}$ 倍,方程 r_3 减去方程 r_1 的 $\frac{(-2)}{2}$ 倍,依次记为 $r_2-\frac{4}{2}r_1$ 及 $r_3-\frac{(-2)}{2}r_1$,得到同解方程组

$$\begin{cases} 2x_1+x_2+x_3=3, & (r_1)\\ -x_2-x_3=-1, & (r_2)\\ 2x_2+4x_3=0. & (r_3)\end{cases}$$

上面方程组中第 2 与 3 个方程仍分别记为 r_2 与 r_3.方程 r_2 中 x_2 的系数不为零,利用方程 r_2,消去 r_3 中的 x_2,即 $r_3-\frac{2}{(-1)}r_2$,得到

$$\begin{cases} 2x_1+x_2+x_3=3, & (r_1)\\ -x_2-x_3=-1, & (r_2)\\ 2x_3=-2. & (r_3)\end{cases}$$

这是与(4.3)同解的三角形方程组.以上过程称为消元过程.

用方程 r_3 中 x_3 的系数除 r_3 两端,得到

$$\begin{cases} 2x_1+x_2+x_3=3, & (r_1)\\ -x_2-x_3=-1, & (r_2)\\ x_3=-1. & (r_3)\end{cases}$$

先消去方程 r_2 中的 x_3,即 $r_2-\frac{-1}{1}r_3$,再用得到的方程中 x_2 的系数除方程两端,得到

$$\begin{cases} 2x_1+x_2+x_3=3, & (r_1)\\ x_2=2, & (r_2)\\ x_3=-1. & (r_3)\end{cases}$$

最后消去方程 r_1 中的 x_2 及 x_3,即 $r_1 - r_2 - r_3$.再用得到方程中 x_1 的系数除方程两端,得到下面与原方程组(4.3)同解的最“简单”的方程组,这就是方程组(4.3)的解:

$$
\begin{aligned}
x_1 &= 1, & (r_1)\\
x_2 &= 2, & (r_2)\\
x_3 &= -1. & (r_3)
\end{aligned}
$$

以上称为回代过程 .

用消元法解线性方程组时,反复施行如下三种变换:

(1) 用一个非零数乘一个方程;

(2) 用一个数乘一个方程加到另一个方程上去;

(3) 交换两个方程在方程组中的次序 .

这些变换与方程组的增广矩阵的初等行变换相对应(见定义 3.3) .因此用消元法解方程组可以用方程组的增广矩阵的初等行变换来表示 .求解例 1 的过程可表示如下:

$$
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 4 & 1 & 1 & 5\\ -2 & 1 & 3 & -3 \end{array}\right)
\xrightarrow[r_3 - \frac{-2}{2} r_1]{r_2 - \frac{4}{2} r_1}
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 0 & -1 & -1 & -1\\ 0 & 2 & 4 & 0 \end{array}\right)
$$

$$
\xrightarrow{r_3 - \frac{2}{(-1)} r_2}
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 0 & -1 & -1 & -1\\ 0 & 0 & 2 & -2 \end{array}\right)
\xrightarrow{\frac{r_2}{2}}
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 0 & -1 & -1 & -1\\ 0 & 0 & 1 & -1 \end{array}\right)
$$

$$
\xrightarrow{r_2 - \frac{(-1)}{1} r_3}
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 0 & -1 & 0 & -2\\ 0 & 0 & 1 & -1 \end{array}\right)
\xrightarrow{\frac{r_2}{(-1)}}
\left(\begin{array}{ccc|c} 2 & 1 & 1 & 3\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -1 \end{array}\right)
$$

$$
\xrightarrow{r_1 - r_2 - r_3}
\left(\begin{array}{ccc|c} 2 & 0 & 0 & 2\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -1 \end{array}\right)
\xrightarrow{\frac{r_1}{2}}
\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1\\ 0 & 1 & 0 & 2\\ 0 & 0 & 1 & -1 \end{array}\right).
$$

从矩阵初等变换的角度看,用 Gauss 消元法解方程组是对其增广矩阵施行初等行变换:(1)消元过程,将方程组系数矩阵变换为三角形矩阵;(2)回代过程,将得到的三角形系数矩阵变换为单位矩阵 .

* **注:Gauss 消元法公式**

由例 1 可见,Gauss 消元法算法并不复杂,对一般线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 而言,为了便于设计计算机软件,上述两个过程可用公式表示如下:

(1) 消元过程

第一步:$(\boldsymbol{A} \mid \boldsymbol{b}) = (\boldsymbol{A}^{(1)} \mid \boldsymbol{b}^{(1)})$

$$
= \begin{array}{cccc|cc}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} & r_1^{(1)} \\
a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} & r_2^{(1)} \\
\cdots & \cdots & & \cdots & \cdots & \cdots \\
a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} & r_n^{(1)}
\end{array}
\quad \begin{array}{c} (\text{当 } a_{11}^{(1)} \neq 0) \\ r_i^{(1)} - \dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}} r_1^{(1)} \, (i = 2, 3, \cdots, n) \end{array}
$$

第二步: $(\boldsymbol{A}^{(2)} \mid \boldsymbol{b}^{(2)}) =$

$$
\begin{array}{cccc|cc}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} & r_1^{(1)} \\
0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} & r_2^{(2)} \\
\cdots & \cdots & & \cdots & \cdots & \cdots \\
0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} & r_n^{(2)}
\end{array}
\quad \begin{array}{c} (\text{当 } a_{22}^{(2)} \neq 0) \\ r_i^{(2)} - \dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}} r_2^{(2)} \, (i = 3, \cdots, n) \end{array}
$$

$$\cdots \ \cdots \ \cdots \ \cdots \ \cdots$$

第 k 步: $(\boldsymbol{A}^{(k)} \mid \boldsymbol{b}^{(k)}) =$

$$
\begin{array}{cccccc|cc}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} & r_1^{(1)} \\
0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & a_{2,k+1}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} & r_2^{(2)} \\
\cdots & \cdots & & \cdots & \cdots & & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} & r_k^{(k)} \\
0 & 0 & \cdots & a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} & b_{k+1}^{(k)} & r_{k+1}^{(k)} \\
\cdots & \cdots & & \cdots & \cdots & & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & a_{nk}^{(k)} & a_{n,k+1}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)} & r_n^{(k)}
\end{array}
\quad \begin{array}{c} (\text{当 } a_{kk}^{(k)} \neq 0) \\ r_i^{(k)} - \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}} r_k^{(k)} \, (i = k+1, \cdots, n) \end{array}
\tag{4.4}
$$

如此,

$$
\begin{aligned}
a_{ij}^{(k+1)} &= a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)} \, (i, j = k + 1, \cdots, n), \\
b_i^{(k+1)} &= b_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} b_k^{(k)}, \, (i = k + 1, \cdots, n) .
\end{aligned}
\tag{4.5}
$$

第 $k+1$ 步: $(\boldsymbol{A}^{(k+1)} \mid \boldsymbol{b}^{(k+1)}) =$

$$
\begin{array}{cccccc|cc}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} & r_1^{(1)} \\
0 & a_{22}^{(2)} & \cdots & a_{2k}^{(2)} & a_{2,k+1}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} & r_2^{(2)} \\
\cdots & \cdots & & \cdots & \cdots & & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} & r_k^{(k)} \\
0 & 0 & \cdots & 0 & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} & b_{k+1}^{(k+1)} & r_{k+1}^{(k+1)} \\
\cdots & \cdots & & \cdots & \cdots & & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & 0 & a_{n,k+1}^{(k+1)} & \cdots & a_{nn}^{(k+1)} & b_n^{(k+1)} & r_n^{(k+1)}
\end{array}
$$

……………

第 n 步:
$$(\boldsymbol{A}^{(n)} \mid \boldsymbol{b}^{(n)}) = \left(\begin{array}{cccc|cc} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} & r_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} & r_2^{(2)} \\ \cdots & \cdots & & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} & r_n^{(n)} \end{array}\right).$$

(2) **回代过程**

逐次回代,依次求出 $x_n, x_{n-1}, \cdots, x_1$:

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}},\ x_i = \frac{b_i^{(i)} - \sum_{j=i+1}^{n} a_{ij}^{(i)} x_j}{a_{ii}^{(i)}} (i = n-1, n-2, \cdots, 1), \qquad (4.6)$$

(4.5)**及**(4.6)**中的** $a_{ii}^{(i)}$ **称为约化主元.不难将上面这些公式变为计算机算法.要通过实例掌握算法具体过程,不用记住上面这个复杂公式.下面定理是以上内容的总结.**

定理 4.1　**对** n **元线性方程组** $\boldsymbol{Ax} = \boldsymbol{b}$, **其中** $\boldsymbol{A} = (a_{ij}^{(1)})_{n\times n}$, $\boldsymbol{b} = (b_1^{(1)}, b_2^{(1)}, \cdots, b_n^{(1)})^{\mathrm{T}}$,**如果约化主元** $a_{kk}^{(k)} \neq 0 (k = 1, 2, \cdots, n)$,**可以使用** Gauss **消元法求解,其消元过程和回代过程依次按公式**(4.5)**及**(4.6)**计算.**

*1.2　主元消元法

在本节 1.1 **中介绍的** Gauss **消元法,是按方程及未知数在方程组中的排列顺序进行的.当约化主元** $a_{kk}^{(k)} = 0$,**则第** k **步消元无法进行.如果** $a_{kk}^{(k)} \neq 0$,**但其绝对值很小,用它作除数计算** $\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} (i = k+1, \cdots, n)$,**计算机也可能发生溢出停机,使消元中断;即使不停机,也可能因为使用绝对值很小的约化主元,而无法保证解的精确度.下面的典型例子显示出约化主元选取的重要性.**

线性方程组

$$0.0001 x_1 + x_2 = 1,$$
$$x_1 + x_2 = 2,$$

其精确解为 $x_1 = \frac{10\ 000}{9\ 999}$, $x_2 = \frac{9\ 998}{9\ 999}$.**如果使用** 3 **位浮点十进制(计算机使用浮点计算时,每个数都保持一定位数.作为说明问题的简例,该例只保持十进制中** 3 **位数字,多于** 3 **位四舍五入),那么,**

$$x_1 \approx 0.100 \times 10,\ x_2 \approx 0.100 \times 10.$$

1. **按顺序消元**

$$\left(\begin{array}{cc|c} 0.100\times 10^{-3} & 0.100\times 10 & 0.100\times 10 \\ 0.100\times 10 & 0.100\times 10 & 0.200\times 10 \end{array}\right) \quad r_2 - \frac{0.100\times 10}{0.100\times 10^{-3}} r_1$$

$$\left(\begin{array}{cc|c} 0.100\times10^{-3} & 0.100\times10 & 0.100\times10 \\ 0 & -0.100\times10^{5} & -0.100\times10^{5} \end{array}\right) \xrightarrow{\frac{r_2}{-10^4}}$$

$$\left(\begin{array}{cc|c} 0.100\times10^{-3} & 0.100\times10 & 0.100\times10 \\ 0 & 0.100\times10 & 0.100\times10 \end{array}\right) \xrightarrow{r_1-r_2}$$

$$\left(\begin{array}{cc|c} 0.100\times10^{-3} & 0 & 0 \\ 0 & 0.100\times10 & 0.100\times10 \end{array}\right)$$

得到近似解:

$$x_1 = 0.000\times10 = 0,\ x_2 = 0.100\times10 .$$

x_1 **的近似值严重失真,主要原因是第一步的约化主元太小,舍入误差增大,有效数字消失 .**

2. 交换第 1 行与第 2 行,再消元:

$$\left(\begin{array}{cc|c} 0.100\times10 & 0.100\times10 & 0.200\times10 \\ 0.100\times10^{-3} & 0.100\times10 & 0.100\times10 \end{array}\right) \xrightarrow{r_2-\frac{0.100\times10^{-3}}{0.100\times10}r_1}$$

$$\left(\begin{array}{cc|c} 0.100\times10 & 0.100\times10 & 0.200\times10 \\ 0 & 0.100\times10 & 0.100\times10 \end{array}\right) \xrightarrow{r_1-r_2}$$

$$\left(\begin{array}{cc|c} 0.100\times10 & 0 & 0.100\times10 \\ 0 & 0.100\times10 & 0.100\times10 \end{array}\right).$$

得到近似解:

$$x_1 = 0.100\times10,\ x_2 = 0.100\times10 .$$

由上面可见,当(4.4)中的 $a_{kk}^{(k)}$ 为零或接近于零时,顺序消元法会使运算溢出停机或误差太大 .仔细观察上面的例子,就会醒悟到:每次消元前,要选择绝对值尽量大的元素作为约化主元,才会得到较精确的结果 .这样的消元法称为主元消元法 .

下面介绍两种主要的主元消元法 .

(1) 完全主元消元法:在(4.4)中的矩阵($\boldsymbol{A}^{(k)} \mid \boldsymbol{b}^{(k)}$)的子块

$$\boldsymbol{A}^{(k)} = \begin{pmatrix} a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{kn}^{(k)} \\ a_{k+1,k}^{(k)} & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ \cdots & \cdots & & \cdots \\ a_{nk}^{(k)} & a_{n,k+1}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}$$

中选取绝对值最大的元素作为主元进行消元,称为完全主元消元法 .

(2) 列主元消元法:在 $\boldsymbol{A}^{(k)}$ 中第一列选取绝对值最大元作为约化主元,称为列主元消元法,即选取

$$a_{i_k k}^{(k)} = \max\left\{ \left| a_{kk}^{(k)} \right|, \left| a_{k+1,k}^{(k)} \right|, \cdots, \left| a_{nk}^{(k)} \right| \right\} = \max_{k\le i\le n} \left| a_{ik}^{(k)} \right|$$

且将$(A^{(k)} \mid b^{(k)})$中第 k 行与第 i_k 行互换,再进行消元.在计算机计算时,实际并不交换行,只是当要进行行交换时,记下主元所在行的序号即可.

与列主元消元法相比,完全主元消元法解线性方程组计算精度高,但选择完全主元的程序较复杂,当方程组未知数较多时,计算速度较慢.

例 2 解线性方程组

$$\begin{aligned} 0.001000x_1 + 2.000x_2 + 3.000x_3 &= 1.000, \\ -2.346x_1 + 3.141x_2 + 8.647x_3 &= 2.000, \\ -4.000x_1 + 1.432x_2 + 7.643x_3 &= 3.000. \end{aligned}$$

保持四位有效数字(例如,$\pi = 3.14159\cdots$,保持四位有效数字,取 $\pi \approx 3.142$).

解 使用列主元消元法.在 x_1 系数中选取主元 -4.000,交换行(用主元上画框表示),得到

$$\left[\begin{array}{ccc|c} \boxed{-4.000} & 1.432 & 7.643 & 3.000 \\ 0.001000 & 2.000 & 3.000 & 1.000 \\ -2.346 & 3.141 & 8.647 & 2.000 \end{array}\right] \quad \begin{array}{l} r_2 + 0.001\ 4.000\ r_1 \\ r_3 - 2.346\ 4.000\ r_1 \end{array}$$

$$\left[\begin{array}{ccc|c} \boxed{-4.000} & 1.432 & 7.643 & 3.000 \\ & 2.000 & 3.002 & 1.001 \\ & \boxed{2.301} & 4.164 & 0.2405 \end{array}\right]$$

$$\begin{array}{l} r_3 \backslash\ r_2 \text{ 之后} \\ \text{再 } r_3 - \dfrac{2.000}{2.301} r_2 \end{array} \quad \left[\begin{array}{ccc|c} \boxed{-4.000} & 1.432 & 7.643 & 3.000 \\ & \boxed{2.301} & 4.164 & 0.2405 \\ & & -0.6173 & 0.7920 \end{array}\right]$$

进行回代,得:

$$x_1 = -2.333,\ x_2 = 2.426,\ x_3 = -1.283.$$

*1.3 Gauss 消元法 Cramer 法则与算法复杂性

尽管 Gauss 消元法与 Cramer 法则都可以用来求解线性方程组,但两种方法在实际计算中的效果并不相同.

显然,用 Gauss 消元法解 n 元一次线性方程组,所用的四则运算的次数与线性方程组未知量个数有关,将计算总次数记为 $f_1(n)$,粗略估算就会发现 $f_1(n) < n^3$.用 Cramer 法则解线性方程组首先要求出行列式的值.如何计算行列式的值呢?到目前为止,我们知道:(1)可以使用定义 3.4,即利用行列式的子式来计算,这时四则运算总次数 $f_2(n) > n!\ (> 2^n$,当 $n \geq 4)$.(2)可利用行列式的性质,将行列式变成与之相等的上三角形行列式(如第三章 第三节 例 5).第(2)种方法计算一个行列式与 Gauss 消元法计算量相当.下面谈到用 Cramer 法则解方

程时是指用第(1)种方法求行列式,再用 Cramer 法则求解方程组.如果使用 Gauss 消元法计算 20 个未知量的线性方程组,即便使用普通的计算机瞬间也能完成计算.但用 Cramer 法则解这个线性方程组至少要作 2.43×10^{18} 次四则运算;如果计算机每秒可计算 100 亿($=10^{10}$)次,一年最多计算 3.2×10^{17} 次,这样至少用七年时间才能完成计算!在实际应用中,解几十个,乃至成百上千个未知量的线性方程组也是常有的事情.在理论上有重要意义的 Cramer 法则,用于数值计算并不有效 这也说明,即便使用超高速计算机,依然需要找到有效的计算方法,才能解决实际计算问题.

为衡量解决问题的一个算法在时间上是否可行,计算机科学家引入了算法时间复杂性概念,以研究计算次数与问题规模的关系,所谓问题规模指的是用计算机解该问题所要输入变量的个数,或输入二进制数据的长度,以 n 表示.使用一个算法解一个问题,从数据输入计算机至得到结果需要的基本运算(算术运算,比较和转移指令等)次数 $f(n)$是问题规模 n 的函数,称为该算法的时间复杂性函数.如果一个算法时间复杂性函数 $f(n)$小于 n 的某一个多项式,称这个算法是有多项式上界的算法,或多项式复杂性算法.对一个问题而言,存在一个多项式复杂性算法,则称该问题为多项式时间可解的问题,简称 P 问题,也称该问题为易处理的,否则称为"难处理的".可以认为,所谓"难处理的问题"可能根本不存在多项式时间的算法,或者还未找到多项式时间的算法,当问题规模 n 较大时,在现有知识及技术条件下,在可行的时间内,难以算出问题的精确的结果.根据这个定义,解线性方程组、矩阵乘法及矩阵求逆都是 P 问题.尽管直接用行列式定义计算 n 阶行列式不是多项式算法,但利用行列式性质,可以在多项式时间计算行列式的值,因而行列式计算也是一个 P 问题.

对于"难处理的问题",尽管人们想出很多方法去处理其特殊情形或用近似算法处理一般问题,但仍然存在大量未解决的问题.

特别有趣的是,这样"难处理的问题"却有独特的用途.1976 年 W.Diffie 和 M.Hellman 首先提出公钥密码体制,这是密码学史上划时代事件,它为计算机网络安全提供了新理论和技术基础.他们提出某些类"难处理的问题"可以用来设计公钥密码体制的密码,破译这样的密码的难度相当于解决一个"难处理的问题".在数论中,求两个大素数(100～200 位十进数字)的积很容易;反过来,知道这个积,将其分解为两个大素数是一个"难处理的问题",利用这个道理,R.Rivest 等三人于 1978 年提出了 RSA 公钥密码,它已在计算机网络中得到了应用.

设计一个算法不只要考虑一个算法的时间复杂性,也要考虑计算过程占用计算机的多少存储单元,即考虑空间复杂性,只有这样才能设计出实用的有效算法.

还要再次强调，虽然 Cramer 法则在计算上并不有效，但在理论上有着重要的意义，因为它明确地揭示了线性方程组的解与系数之间的关系．

1.4　用消元法解一般线性方程组

1．例子

用消元法可以求解一般线性方程组．首先看一个例子．

例 3　解线性方程组

$$\begin{cases} x_1 + x_2 + 3x_3 + x_4 + x_5 = 2, \\ x_3 - x_5 = 1, \\ x_1 + x_2 + 5x_3 + x_4 - x_5 = 4. \end{cases} \tag{4.7}$$

解　用方程组增广矩阵的初等行变换表示 Gauss 消元法过程：

$$\left(\begin{array}{ccccc|c} 1 & 1 & 3 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 1 & 1 & 5 & 1 & -1 & 4 \end{array}\right) \xrightarrow{r_3 - r_1} \left(\begin{array}{ccccc|c} 1 & 1 & 3 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 2 & 0 & -2 & 2 \end{array}\right) \xrightarrow{r_3 - 2r_2}$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 3 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{r_1 - 3r_2} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 4 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

最后这个矩阵是行阶梯形矩阵，它的每行第一个非零元是 1 且其所在列的其余元已经全化为零，而且上一行从左开始的“0”的个数少于下一行从左开始的“0”的个数，这个矩阵称为原矩阵的行最简形．由上面的初等行变换会领悟到，当然也可以用归纳法证明：任何一个矩阵经过有限次初等行变换都能化为行最简形矩阵．方程组(4.7)与下列方程组同解：

$$\begin{cases} x_1 + x_2 + x_4 + 4x_5 = -1, \\ x_3 - x_5 = 1. \end{cases} \text{即} \begin{cases} x_1 = -x_2 - x_4 - 4x_5 - 1, \\ x_3 = x_5 + 1. \end{cases} \tag{4.8}$$

令 $x_2 = k_1,\ x_4 = k_2,\ x_5 = k_3$，得 $x_1 = -k_1 - k_2 - 4k_3 - 1,\ x_3 = k_3 + 1$，即

$$\begin{cases} x_1 = -k_1 - k_2 - 4k_3 - 1, \\ x_2 = k_1, \\ x_3 = k_3 + 1, \\ x_4 = k_2, \\ x_5 = k_3. \end{cases}$$

将解写成列向量形式：

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = k_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + k_3 \begin{pmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{4.9}$$

从上面消元过程可知,方程组(4.7)相应的齐次方程组

$$\begin{cases} x_1 + x_2 + 3x_3 + x_4 + x_5 = 0, \\ x_3 - x_5 = 0, \\ x_1 + x_2 + 5x_3 + x_4 - x_5 = 0, \end{cases}$$

与

$$\begin{cases} x_1 = -x_2 - x_4 - 4x_5, \\ x_3 = x_5, \end{cases} \tag{4.10}$$

同解,且解向量可表示为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = k_1 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + k_3 \begin{pmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

解向量被表示为三个线性无关列向量的线性组合,这三个列向量都是齐次方程组(4.10)的解.解向量(4.9)中$(-1,0,1,0,0)^T$却是非齐次方程组(4.7)的一个解;而(4.9)是方程组(4.7)的通解,即方程组(4.7)的任意一个解都可写成这种形式.

例 4 解线性方程组

$$\begin{cases} x_1 + x_2 + 3x_3 + x_4 + x_5 = 2, \\ x_3 - x_5 = 1, \\ x_1 + x_2 + 5x_3 + x_4 - x_5 = 6. \end{cases}$$

解 对方程组进行与例 3 类似的初等行变换,得到

$$\left(\begin{array}{ccccc|c} 1 & 1 & 3 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 1 & 1 & 5 & 1 & -1 & 6 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 4 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right). \tag{4.11}$$

最后矩阵第三行表示:$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = 2$,这显然矛盾.方程组无解.

2. 分析

从上面两个特殊实例可以看出,用消元法可以解一般的线性方程组,消元的

结果可表明方程是否有解,如果有解还可求出所有解(通解) .对于线性方程组有解的条件及解的结构,能从上面例子中归纳出什么一般结论呢 ?

例 4 的线性方程组之所以无解,是因为它的第三个方程左端是第一个方程左端和第二个方程左端 2 倍之和,但第三个方程的右端却不是第一个方程右端和第二个方程右端 2 倍之和,导致“ 0 = 2 ”这个矛盾,这反映在(4.11)中最后一个矩阵中系数矩阵的秩为 2,而增广矩阵的秩为 3;由于初等行变换并不改变矩阵的秩,因而原方程组系数矩阵的秩与增广矩阵的秩也不相同 .从例 3 可猜想,如果线性方程组的系数矩阵的秩与增广矩阵的秩相同,方程一定有解 .综合两例可能猜想:(A)线性方程组有解的充要条件是其系数矩阵的秩与增广矩阵的秩相同 观察例 3 及更多例子,可以发现:(B)非齐次线性方程组的一般解是由相应的齐次线性方程组的一般解及非齐次线性方程组的一个特殊解组成的 .例 3 中,齐次线性方程组有 5 个未知数,而方程组第三个方程是第一个方程加上第二个方程 2 倍组成的,即系数矩阵的秩为 2 .因而可以去掉第三个方程,并将两个未知数 x_1, x_3 留在第一、二个方程左端,其余三个未知数 x_2, x_4, x_5 移至右端 .令 x_2, x_4, x_5 等于任意常数会得到齐次方程组所有解,从(4.10)看出,当 $(x_2, x_4, x_5)^{\mathrm{T}} = (1,0,0)^{\mathrm{T}}, (0,1,0)^{\mathrm{T}}, (0,0,1)^{\mathrm{T}}$ 时,(4.10)所有的解都可写成三个线性无关解向量 $(-1,1,0,0,0)^{\mathrm{T}}, (-1,0,0,1,0)^{\mathrm{T}}, (-4,0,1,0,1)^{\mathrm{T}}$ 的线性组合 .(C)对 n 个未知数、系数矩阵的秩为 r 的线性方程组,我们大概会猜想:n 个未知数的齐次线性方程组,其系数矩阵的秩为 r,则其所有解向量均可表为 $n - r$ 个线性无关的解向量的线性组合 .猜想的结果不一定正确,也可能导致错误 .下面两节,将证明上面由观察得到的猜想(A)、(B)、(C)是正确的 .

数学家 L.Euler(1707—1783)说过:“ 我们认识到,在仍然很不完善的数论中,还得把最大希望寄托于观察之中;这些观察将导致我们继续获得以后尽力予以证明的新的性质……我们不要轻易地把观察所发现的和仅以归纳为旁证的关于数的那样一些性质信以为真……我们应该把这样一种发现当作一种机会,去更精确地研究所发现的性质,以便证明或推翻它;在这两种情况之中,我们都会学到一些有用的东西 ”(G.Polya .数学与猜想 .第 1 卷,李心灿,等译 .北京:科学出版社,1984) .上面这段话,谈的虽然是数论,其实对整个数学也是正确的 .

第二节　线性方程组解的结构

2.1　齐次线性方程组解的结构

上节使用 Gauss 消元法解线性方程组,从对方程组解的观察中得到一些猜

想,下面证明这些猜想是正确的.

设齐次线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}, \tag{4.12}$$

其中 $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{A}$ 的秩为 r, $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, (4.12)的解是 n 元向量,称之为解向量.显然, $\boldsymbol{x} = (0,0,\cdots,0)^{\mathrm{T}}$ 是方程组(4.12)的解.称为该齐次线性方程组的零解.而把其他的解称为齐次线性方程组的非零解.下面要研究齐次线性方程组满足什么条件才能有非零解,这些非零解有什么性质.

定理 4.2 若 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 是齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的两个解, k_1 与 k_2 是实数,则 $k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2$ 也是它的解.

证 由于

$$\boldsymbol{A}(k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2) = k_1\boldsymbol{A}\boldsymbol{x}_1 + k_2\boldsymbol{A}\boldsymbol{x}_2 = k_1\boldsymbol{0} + k_2\boldsymbol{0} = \boldsymbol{0},$$

所以 $k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2$ 也是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的解.□

定义 4.1 设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的解,如果 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 线性无关,且 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的任一解可由它们线性表示,则称 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 为 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的一个基础解系.

如果知道 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的基础解系 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$,则它的一般解可写成

$$\boldsymbol{x} = k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_s\boldsymbol{x}_s,$$

其中 $k_1, k_2, \cdots, k_s$ 为任意实数,上式称为它的通解.

定理 4.3 设 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$,其中 $\boldsymbol{A} = (a_{ij})_{m\times n}$,且 $\boldsymbol{A}$ 的秩为 $r(<n)$,则齐次线性方程组(4.12)存在 $n-r$ 个解向量组成的基础解系.若 $\boldsymbol{A}$ 的秩 $r = n$,则(4.12)只有零解.

证 下面用两种方法证明该定理.

方法一

证明的思路是沿着 Gauss 消元法展开的,即针对一般情况,将 Gauss 消元法予以严格的论述.首先证明, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 存在 $n-r$ 个线性无关的解向量.因为 $\boldsymbol{A}$ 的秩为 r,无损一般性,可设 $\boldsymbol{A}$ 的前 r 个列向量线性无关;为了简明,且可设对 $\boldsymbol{A}$ 进行初等行变换,最终可化为行最简形:

$$\boldsymbol{U} = \begin{matrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & c_{1,r+2} & \cdots & c_{1n} \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & c_{2,r+2} & \cdots & c_{2n} \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & c_{r,r+2} & \cdots & c_{rn} \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{matrix}$$

且 $\boldsymbol{Ax}=\mathbf{0}$与 $\boldsymbol{Ux}=\mathbf{0}$即如下方程组同解:

$$\begin{cases} x_1 + c_{1,r+1}x_{r+1} + c_{1,r+2}x_{r+2} + \cdots + c_{1n}x_n = 0, \\ x_2 + c_{2,r+1}x_{r+1} + c_{2,r+2}x_{r+2} + \cdots + c_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots \\ x_r + c_{r,r+1}x_{r+1} + c_{r,r+2}x_{r+2} + \cdots + c_{rn}x_n = 0 . \end{cases} \tag{4.13}$$

取 $x_{r+1}, x_{r+2}, \cdots, x_n$ 为自由变量,于是它们可取任意值,例如取如下的 $n-r$ 组值

$$\begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}. \tag{4.14}$$

显然这 $n-r$ 个列向量是线性无关的 .将(4.14)分别代入(4.13),得

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_r \end{pmatrix} = \begin{pmatrix} -c_{1,r+1} \\ -c_{2,r+1} \\ \cdots \\ -c_{r,r+1} \end{pmatrix}, \begin{pmatrix} -c_{1,r+2} \\ -c_{2,r+2} \\ \cdots \\ -c_{r,r+2} \end{pmatrix}, \cdots, \begin{pmatrix} -c_{1n} \\ -c_{2n} \\ \cdots \\ -c_{rn} \end{pmatrix}.$$

于是

$$\boldsymbol{x}_1 = \begin{pmatrix} -c_{1,r+1} \\ -c_{2,r+1} \\ \cdots \\ -c_{r,r+1} \\ 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \boldsymbol{x}_2 = \begin{pmatrix} -c_{1,r+2} \\ -c_{2,r+2} \\ \cdots \\ -c_{r,r+2} \\ 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix}, \boldsymbol{x}_{n-r} = \begin{pmatrix} -c_{1n} \\ -c_{2n} \\ \cdots \\ -c_{rn} \\ 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}$$

是方程组 $\boldsymbol{Ax}=\mathbf{0}$的解,且线性无关 .

下面证明, $\boldsymbol{Ax}=\mathbf{0}$的任一个解 $\boldsymbol{x}$ 可由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n-r}$线性表示 .设

$$\boldsymbol{x} = (k_1, k_2, \cdots, k_r, k_{r+1}, k_{r+2}, \cdots, k_n)^{\mathrm{T}}$$

为方程(4.13)的解,取

$$\boldsymbol{x}_0 = k_{r+1}\boldsymbol{x}_1 + k_{r+2}\boldsymbol{x}_2 + \cdots + k_n\boldsymbol{x}_{n-r},$$

则 $\boldsymbol{x}-\boldsymbol{x}_0$ 是 $\boldsymbol{Ax}=\mathbf{0}$的解,也是 $\boldsymbol{Ux}=\mathbf{0}$的解,即满足(4.13) .这时

$$\boldsymbol{x}-\boldsymbol{x}_0 = (c_1, c_2, \cdots, c_r, 0, \cdots, 0)^{\mathrm{T}} .$$

由于它满足(4.13),知 $\boldsymbol{x}-\boldsymbol{x}_0$ 的前 r 个分量 $c_1 = c_2 = \cdots = c_r = 0$.于是 $\boldsymbol{x}-\boldsymbol{x}_0 =$

$\mathbf{0}$,即

$$\boldsymbol{x} = k_{r+1}\boldsymbol{x}_1 + k_{r+2}\boldsymbol{x}_2 + \cdots + k_n\boldsymbol{x}_{n-r}.$$

这说明 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的任意一个解可用 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{n-r}$线性表示.又因为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{n-r}$是线性无关的,所以由定义 4.1 知,$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{n-r}$是 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的一个基础解系.

方法二

设

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \cdots \\ \boldsymbol{A}_m \end{pmatrix}$$

其中 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_m$ 为 $\boldsymbol{A}$ 的 m 个 n 维行向量.因为 $\boldsymbol{A}_i$ 为行向量,所以 $\boldsymbol{A}_i^{\mathrm{T}}$ 及 $\boldsymbol{x}$ 都可看成是 n 维欧氏几何空间 $\boldsymbol{R}^n$ 中的向量.方程 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$可写成

$$\boldsymbol{A}_i\boldsymbol{x} = (\boldsymbol{A}_i^{\mathrm{T}}, \boldsymbol{x}) = 0,\ i = 1,2,\cdots, m$$

于是,$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的解向量 $\boldsymbol{x}$ 就是与 m 个列向量 $\boldsymbol{A}_1^{\mathrm{T}},\boldsymbol{A}_2^{\mathrm{T}},\cdots,\boldsymbol{A}_m^{\mathrm{T}}$ 垂直的向量的全体.令 S 是由 $\boldsymbol{A}_1^{\mathrm{T}},\boldsymbol{A}_2^{\mathrm{T}},\cdots,\boldsymbol{A}_m^{\mathrm{T}}$ 生成的 $\boldsymbol{R}^n$ 的子空间,则 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$全体解向量就构成 S 的正交补空间 $S^{\perp}$.如果系数矩阵 $\boldsymbol{A}$ 的秩为 r,则子空间 S 的维数为 r,即 $\dim(S) = r$.由命题 1.23 知,$\dim(S^{\perp}) = n - r$ 且 $S^{\perp}$ 的基是 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的一个基础解系.当系数矩阵 $\boldsymbol{A}$ 的秩为 n 时,则 $\dim(S^{\perp}) = n - n = 0$,$S^{\perp} = \{\mathbf{0}\}$,这说明 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$只有零解.□

事实上,定理 4.3 的方法一给出了求齐次线性方程组通解的一般方法.由定理 4.2 及定理 4.3,可得如下推论.

推论 4.1　设 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$,其中 $\boldsymbol{A}=(a_{ij})_{m\times n}$为实矩阵,且 $\boldsymbol{A}$ 的秩为 r,则 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的解在 n 维向量空间 $\boldsymbol{R}^n$ 中构成一个 $n-r$ 维子空间,称其为 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的解空间.$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$的基础解系是其解空间的一个基.

例 1　解齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$,其中

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -4 \\ 1 & 2 & -7 \\ 1 & 3 & -10 \end{pmatrix}.$$

解　对矩阵 $\boldsymbol{A}$ 作初等行变换:

$$\begin{pmatrix} 1 & 1 & -4 \\ 1 & 2 & -7 \\ 1 & 3 & -10 \end{pmatrix} \xrightarrow[r_3-r_1]{r_2-r_1} \begin{pmatrix} 1 & 1 & -4 \\ 0 & 1 & -3 \\ 0 & 2 & -6 \end{pmatrix} \xrightarrow{r_3-2r_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}.$$

于是

$$\begin{cases} x_1 = x_3, \\ x_2 = 3x_3, \end{cases} \text{即} \begin{cases} x_1 = x_3, \\ x_2 = 3x_3, \\ x_3 = x_3 . \end{cases}$$

其中 x_3 为自由变量,令 $x_3 = k$,则方程组的通解

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} .$$

与定理 4.3 证法 2 相对照,在 3 维欧氏几何空间 $\boldsymbol{R}^3$ 中, S 是由向量$(1, 1, -4)^{\mathrm{T}}$及$(1, 2, -7)^{\mathrm{T}}$ 生成的 $\boldsymbol{R}^3$ 的子空间 .对 $\boldsymbol{Ax} = \boldsymbol{0}$而言,其解空间 $S^{\perp}$是 $\boldsymbol{R}^3$ 中一个 1 维子空间, $\boldsymbol{Ax} = \boldsymbol{0}$的基础解系 $\alpha = (1, 3, 1)^{\mathrm{T}}$ 是 $S^{\perp}$的一个基 .

例 2　求 a 与 b,使得齐次线性方程组

$$\begin{cases} ax_1 + x_2 + x_3 = 0, \\ x_1 + bx_2 + x_3 = 0, \\ x_1 + 2bx_2 + x_3 = 0, \end{cases}$$

有非零解,并求其通解 .

解　该齐次线性方程组的系数行列式

$$D = \begin{vmatrix} a & 1 & 1 \\ 1 & b & 1 \\ 1 & 2b & 1 \end{vmatrix} = -b(a-1) .$$

当 $D = 0$ 时,即 $b = 0$ 或 $a = 1$,方程组才有非零解 .

当 $b = 0$ 时,对方程组的系数矩阵作初等行变换有

$$D = \begin{pmatrix} a & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \xrightarrow{r_3 - r_2} \begin{pmatrix} a & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{r_1 - ar_2} \begin{pmatrix} 0 & 1 & 1-a \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix};$$

因此,方程组的解可写成

$$\begin{cases} x_1 = -x_3 \\ x_2 = (a-1)x_3, \\ x_3 = x_3, \end{cases}$$

即其通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} -1 \\ a-1 \\ 1 \end{pmatrix} .$$

当 $a = 1$ 时,对方程组系数矩阵作初等行变换有

$$\boldsymbol{D}=\begin{pmatrix}1&1&1\\1&b&1\\1&2b&1\end{pmatrix}\xrightarrow{r_2-r_1,r_3-r_1}\begin{pmatrix}1&1&1\\0&b-1&0\\0&2b-1&0\end{pmatrix}.$$

原方程与如下方程组同解

$$\begin{cases}x_1+x_2+x_3=0,\\(b-1)x_2=0,\\(2b-1)x_2=0.\end{cases}$$

因为 $b-1$ 与 $2b-1$ 至少有一个不为零,所以 $x_2=0$.因而方程组的通解为

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=k\begin{pmatrix}-1\\0\\1\end{pmatrix}.$$

2.2 非齐次线性方程组解的结构

考察非齐次线性方程组

$$\boldsymbol{Ax}=\boldsymbol{b},\tag{4.15}$$

其中 $\boldsymbol{A}=(a_{ij})_{m\times n}$, $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, $\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}\neq\boldsymbol{0}$,令 $\boldsymbol{A}$ 的 n 个 m 维列向量为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$,于是 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$,方程 $Ax=b$ 可写成

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{b}.\tag{4.16}$$

方程组(4.15)有解的充要条件是(4.16)有解,即向量 $\boldsymbol{b}$ 可由 $\boldsymbol{A}$ 的列向量线性表示.因为向量 $\boldsymbol{b}$ 可由矩阵 $\boldsymbol{A}$ 的列向量线性表示,所以矩阵 $\boldsymbol{A}$ 与矩阵 $(\boldsymbol{A}\mid\boldsymbol{b})$ 的秩相等;反过来,如果矩阵 $\boldsymbol{A}$ 与矩阵 $(\boldsymbol{A}\mid\boldsymbol{b})$ 的秩相等,那么,矩阵 $\boldsymbol{A}$ 的列向量的极大线性无关组可作为 $(\boldsymbol{A}\mid\boldsymbol{b})$ 的列向量的极大线性无关组,于是,向量 $\boldsymbol{b}$ 可由这个极大线性无关组线性表示,所以 $\boldsymbol{b}$ 可用 $\boldsymbol{A}$ 的列向量线性表示.综上所述,下面定理成立.

定理 4.4 对于非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$,下列条件等价:

(1) 方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解;

(2) 向量 $\boldsymbol{b}$ 可由 $\boldsymbol{A}$ 的列向量线性表示;

(3) 矩阵 $\boldsymbol{A}$ 的秩与其增广矩阵 $(\boldsymbol{A}\mid\boldsymbol{b})$ 的秩相等.

下面讨论 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解的结构.

定理 4.5 (1) 设 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的两个解,则 $\boldsymbol{x}_1-\boldsymbol{x}_2$ 是对应的齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解.

(2) 若 $\boldsymbol{x}_0$ 为非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解,$\boldsymbol{x}_H$ 为相应的齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解,则 $\boldsymbol{x}_0+\boldsymbol{x}_H$ 为 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解.

证 (1) 因为

$$\boldsymbol{A}(\boldsymbol{x}_1-\boldsymbol{x}_2)=\boldsymbol{A}\boldsymbol{x}_1-\boldsymbol{A}\boldsymbol{x}_2=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0},$$

故 $\boldsymbol{x}_1-\boldsymbol{x}_2$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$的解.

(2) 由 $\boldsymbol{A}(\boldsymbol{x}_0+\boldsymbol{x}_H)=\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{A}\boldsymbol{x}_H=\boldsymbol{b}+\boldsymbol{0}=\boldsymbol{b}$,即 $\boldsymbol{x}_0+\boldsymbol{x}_H$ 为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解.□

定理 4.6　设非齐次线性方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有解,其中 $\boldsymbol{A}=(a_{ij})_{m\times n}$,且 $R(\boldsymbol{A})=r$,$\boldsymbol{x}_0$ 是它的某一个解(习惯上,称为特解),$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$的一个基础解系为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{n-r}$,则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的一般解(通解)为

$$\boldsymbol{x}=\boldsymbol{x}_0+k_1\boldsymbol{x}_1+\cdots+k_{n-r}\boldsymbol{x}_{n-r},$$

其中 $k_1,\cdots,k_{n-r}$为任意实数.

证　设 $\boldsymbol{x}$ 为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解.由定理 4.5(1)知 $\boldsymbol{x}-\boldsymbol{x}_0$ 为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$的解.由定义 4.1 知,$\boldsymbol{x}-\boldsymbol{x}_0$ 总能写成 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{n-r}$的线性组合,即

$$\boldsymbol{x}-\boldsymbol{x}_0=k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_{n-r}\boldsymbol{x}_{n-r},$$

其中 $k_1,k_2,\cdots,k_{n-r}$为任意实数.即有

$$\boldsymbol{x}=\boldsymbol{x}_0+k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_{n-r}\boldsymbol{x}_{n-r}.$$

由定理 4.5(2)可知,对任意实数 $k_1,k_2,\cdots,k_{n-r}$,上式总是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解,故上式为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解.□

注:定理 4.2 表明,以秩为 r 的矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$为系数矩阵的齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$,它的解构成 n 元列向量空间 $\boldsymbol{R}^n$ 的一个 $n-r$ 维子空间.非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解不能构成 n 维向量空间 $\boldsymbol{R}^n$ 的子空间,因为若 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 分别是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解,$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$ 却不是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解.

例 3　设非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的增广矩阵

$$(\boldsymbol{A}\mid\boldsymbol{b})=\left(\begin{array}{ccccc|c}1&1&2&0&0&1\\1&2&-1&1&2&2\\2&3&1&1&2&3\\3&5&0&2&4&5\end{array}\right),$$

求 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解.

解　对$(\boldsymbol{A}\mid\boldsymbol{b})$作初等行变换

$$\left(\begin{array}{ccccc|c}1&1&2&0&0&1\\1&2&-1&1&2&2\\2&3&1&1&2&3\\3&5&0&2&4&5\end{array}\right)\xrightarrow[r_4-3r_1]{\substack{r_2-r_1\\r_3-2r_1}}\left(\begin{array}{ccccc|c}1&1&2&0&0&1\\0&1&-3&1&2&1\\0&1&-3&1&2&1\\0&2&-6&2&4&2\end{array}\right)\xrightarrow{\substack{r_3-r_2\\r_4-2r_2}}$$

$$\left(\begin{array}{ccccc|c}1&1&2&0&0&1\\0&1&-3&1&2&1\\0&0&0&0&0&0\\0&0&0&0&0&0\end{array}\right)\xrightarrow{r_1-r_2}\left(\begin{array}{ccccc|c}1&0&5&-1&-2&0\\0&1&-3&1&2&1\\0&0&0&0&0&0\\0&0&0&0&0&0\end{array}\right).$$

显然 $R(\boldsymbol{A})=R((\boldsymbol{A}\mid\boldsymbol{b}))=2$,此时方程有解.上面矩阵所对应的与原方程同解的方程组为

$$\begin{aligned}x_1+5x_3-x_4-2x_5&=0,\\x_2-3x_3+x_4+2x_5&=1.\end{aligned}$$

由于 $R(\boldsymbol{A})=2$,故可在未知数 x_1,x_2,x_3,x_4,x_5 中选取两个,使这两个未知数的系数组成的行列式不为零,可选取 x_1,x_2,它们可由其余三个未知数唯一地表示.得到下列与原方程同解的线性方程组

$$\begin{aligned}x_1&=-5x_3+x_4+2x_5,\\x_2&=3x_3-x_4-2x_5+1.\end{aligned}$$

即

$$\begin{aligned}x_1&=-5x_3+x_4+2x_5+0,\\x_2&=3x_3-x_4-2x_5+1,\\x_3&=x_3+0,\\x_4&=x_4+0,\\x_5&=x_5+0.\end{aligned}$$

改写为列向量形式

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=x_3\begin{pmatrix}-5\\3\\1\\0\\0\end{pmatrix}+x_4\begin{pmatrix}1\\-1\\0\\1\\0\end{pmatrix}+x_5\begin{pmatrix}2\\-2\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}.$$

令 $x_3=k_1,x_4=k_2,x_5=k_3$ 为任意实数.于是,得到 $\boldsymbol{Ax}=\boldsymbol{b}$ 的特解 $\boldsymbol{x}_0=(0,1,0,0,0)^{\mathrm{T}}$ 及 $\boldsymbol{Ax}=\boldsymbol{0}$的一个基础解系:

$$\boldsymbol{x}_1=(-5,3,1,0,0)^{\mathrm{T}},\boldsymbol{x}_2=(1,-1,0,1,0)^{\mathrm{T}},\boldsymbol{x}_3=(2,-2,0,0,1)^{\mathrm{T}}.$$

这样,$\boldsymbol{Ax}=\boldsymbol{b}$ 的通解为

$$\boldsymbol{x}=\boldsymbol{x}_0+k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+k_3\boldsymbol{x}_3,$$

其中 $k_1,k_2,k_3\in\mathbf{R}$.

例 4 设 $\boldsymbol{x}_0$ 是非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个特解,设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$ 为 $\boldsymbol{Ax}=\boldsymbol{0}$的基础解系.证明

$$\boldsymbol{x}_0,\boldsymbol{x}_0+\boldsymbol{x}_1,\boldsymbol{x}_0+\boldsymbol{x}_2,\cdots,\boldsymbol{x}_0+\boldsymbol{x}_s$$

线性无关,且 $\boldsymbol{Ax}=\boldsymbol{b}$ 的任意一个解可表示为

$$\boldsymbol{x}=k_0\boldsymbol{x}_0+k_1(\boldsymbol{x}_0+\boldsymbol{x}_1)+k_2(\boldsymbol{x}_0+\boldsymbol{x}_2)+\cdots+k_s(\boldsymbol{x}_0+\boldsymbol{x}_s),$$

其中 $k_0+k_1+\cdots+k_s=1$.

证　设有实数 $a_0, a_1, \cdots, a_s$ 使得

$$a_0\boldsymbol{x}_0 + a_1(\boldsymbol{x}_0 + \boldsymbol{x}_1) + a_2(\boldsymbol{x}_0 + \boldsymbol{x}_2) + \cdots + a_s(\boldsymbol{x}_0 + \boldsymbol{x}_s) = \boldsymbol{0},$$

改写为

$$(a_0 + a_1 + \cdots + a_s)\boldsymbol{x}_0 + a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_s\boldsymbol{x}_s = \boldsymbol{0},$$

如果 $c = a_0 + a_1 + \cdots + a_s \neq 0$，则

$$\boldsymbol{x}_0 = -c^{-1}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_s\boldsymbol{x}_s),$$

于是

$$\boldsymbol{A}\boldsymbol{x}_0 = -c^{-1}(a_1\boldsymbol{A}\boldsymbol{x}_1 + a_2\boldsymbol{A}\boldsymbol{x}_2 + \cdots + a_s\boldsymbol{A}\boldsymbol{x}_s) = \boldsymbol{0}.$$

这与 $\boldsymbol{x}_0$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解矛盾.因此 $c = 0$.于是 $a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_s\boldsymbol{x}_s = \boldsymbol{0}$.由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_s$ 线性无关知 $a_1 = a_2 = \cdots = a_s = 0$，又因为 $c = 0$，即 $a_0 + a_1 + \cdots + a_s = 0$，因而 $a_0 = 0$.这样证明了 $a_0 = a_1 = a_2 = \cdots = a_s = 0$，因此 $\boldsymbol{x}_0, \boldsymbol{x}_0 + \boldsymbol{x}_1, \boldsymbol{x}_0 + \boldsymbol{x}_2, \cdots, \boldsymbol{x}_0 + \boldsymbol{x}_s$ 线性无关.由定理 4.6 可知，$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的任意一解可表示为

$$\begin{aligned}\boldsymbol{x} &= \boldsymbol{x}_0 + k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_s\boldsymbol{x}_s, \\ &= (1 - (k_1 + \cdots + k_s))\boldsymbol{x}_0 + k_1(\boldsymbol{x}_0 + \boldsymbol{x}_1) + k_2(\boldsymbol{x}_0 + \boldsymbol{x}_2) + \cdots + k_s(\boldsymbol{x}_0 + \boldsymbol{x}_s).\end{aligned}$$

令 $k_0 = (1 - (k_1 + \cdots + k_s))$，则 $k_0 + k_1 + \cdots + k_s = 1$.□

例 5　在平面直角坐标系中，已知不在同一条直线的三点 $A(2,8)$，$B(7,3)$ 及 $C(5,7)$，求通过这三点的圆的方程.

解　由平面解析几何知，平面上圆的方程可写为

$$x^2 + y^2 + c_1x + c_2y + c_3 = 0.$$

因为 A, B, C 三点坐标满足上面方程，于是得到以 c_1, c_2, c_3 为未知数的线性方程组，

$$\begin{aligned}2c_1 + 8c_2 + c_3 &= -68, \\ 7c_1 + 3c_2 + c_3 &= -58, \\ 5c_1 + 7c_2 + c_3 &= -74.\end{aligned}$$

其解为 $c_1 = -4, c_2 = -6, c_3 = -12$，于是圆的方程为

$$x^2 + y^2 - 4x - 6y - 12 = 0 \text{ 即 } (x-2)^2 + (y-3)^2 = 5^2.$$

***例 6**　实验测得两个物理量 x, y 数据的样本 $(x_i, y_i)(i = 1, 2, \cdots, 8)$.

i	1	2	3	4	5	6	7	8
x_i	1.2	1.8	3.1	4.9	5.7	7.1	8.6	9.8
y_i	4.5	5.9	7.0	7.8	7.2	6.8	4.5	2.7

推测 y 与 x 间呈如下关系：

$$y = a + bx + cx^2.$$

如何选择参数 a, b, c，使得用上式预测 y 时产生的误差平方和最小？

解　对 $x_i(i=1,2,\cdots,8)$而言，y_i 与

$$\hat{y}_i = a + bx_i + cx_i^2$$

之间误差为

$$y_i - (a + bx_i + cx_i^2).$$

对于这 8 个数据样本而言，预测的误差平方和为

$$S = \sum_{i=1}^{8} (y_i - (a + bx_i + cx_i^2))^2.$$

由多元微分法知，使 S 达到极小值时，a, b, c 满足如下条件

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = \frac{\partial S}{\partial c} = 0.$$

因为

$$\frac{\partial S}{\partial a} = -2\sum_{i=1}^{8} [y_i - (a + bx_i + cx_i^2)],$$

$$\frac{\partial S}{\partial b} = -2\sum_{i=1}^{8} [y_i - (a + bx_i + cx_i^2)] x_i,$$

$$\frac{\partial S}{\partial c} = -2\sum_{i=1}^{8} [y_i - (a + bx_i + cx_i^2)] x_i^2.$$

所以 a, b, c 应满足线性方程组

$$\sum_{i=1}^{8} y_i = a8 + b\sum_{i=1}^{8} x_i + c\sum_{i=1}^{8} x_i^2,$$

$$\sum_{i=1}^{8} x_i y_i = a\sum_{i=1}^{8} x_i + b\sum_{i=1}^{8} x_i^2 + c\sum_{i=1}^{8} x_i^3,$$

$$\sum_{i=1}^{8} x_i^2 y_i = a\sum_{i=1}^{8} x_i^2 + b\sum_{i=1}^{8} x_i^3 + c\sum_{i=1}^{8} x_i^4.$$

代入数据样本$(x_i, y_i)(i=1,2,\cdots,8)$，则有

$$\begin{aligned} 8.00a + 42.20b + 291.20c &= 46.40, \\ 42.20a + 291.20b + 2275.35c &= 230.42, \\ 291.20a + 2275.35b + 18971.92c &= 1449.00. \end{aligned}$$

解之，得

$$a = 2.58,\ b = 2.06,\ c = -0.21,$$

即

$$y = 2.58 + 2.06x - 0.21x^2.$$

线性代数可以应用到科学技术与经济诸多领域，附录 A.2 给出了线性代数在经济学中应用的一个例子．

习 题 四

习题 A

1．用 Gauss 消元法解线性方程组:

(1)(《九章算术》)今有卖牛二,羊五,以买十三豕,有余钱一千(卖二头牛,五头羊,买 13 头猪还剩钱 1000);卖牛三,豕三,以买九羊,钱适足(钱正好不多不少);卖羊六,豕八,以买五牛,钱不足六百(还少钱 600),问牛、羊、豕各几何(每头多少钱)?

(2) $\begin{cases} 5x_1+3x_2+2x_3=4, \\ 3x_1+3x_2+2x_3=2, \\ x_2+x_3=5; \end{cases}$ (3) $\begin{cases} x+y+z=5, \\ x+y-4z=10, \\ -4x+y+z=0; \end{cases}$

(4) $\begin{cases} 2x_1+x_2+x_3=7, \\ -x_1+3x_2-2x_3=-5, \\ x_1+2x_2+3x_3=14; \end{cases}$ (5) $\begin{cases} x+y+z=3, \\ x+y-z=-1, \\ x+3y-8z=0. \end{cases}$

2．求下列齐次线性方程组的一个基础解系:

(1) $\begin{cases} x_1+x_2-3x_3-x_4=0, \\ 3x_1-x_2-3x_3+4x_4=0, \\ x_1+5x_2-9x_3-8x_4=0; \end{cases}$ (2) $\begin{cases} x_1-2x_2-x_3-2x_4=0, \\ 4x_1+x_2+2x_3+x_4=0, \\ 2x_1+5x_2+4x_3-x_4=0, \\ x_1+x_2+x_3+x_4=0; \end{cases}$

(3) $\begin{cases} x_1+x_2+x_3+x_4+x_5=0, \\ 3x_1+2x_2+x_3+x_4-3x_5=0, \\ x_2+2x_3+2x_4+6x_5=0, \\ 5x_1+4x_2+3x_3+3x_4-x_5=0; \end{cases}$ (4) $\begin{cases} x_1-2x_2+x_3-x_4+x_5=0, \\ 2x_1+x_2-x_3+2x_4-3x_5=0, \\ 3x_1-2x_2-x_3+x_4-2x_5=0, \\ 2x_1-5x_2+x_3-2x_4+2x_5=0. \end{cases}$

3．求下列非齐次线性方程组的通解:

(1) $\begin{cases} x_1+x_2+2x_3=5, \\ x_1+x_3=-2, \\ 2x_1+x_2+3x_3=3; \end{cases}$ (2) $\begin{cases} 2x_1+7x_2+3x_3+x_4=6, \\ 3x_1+5x_2+2x_3+2x_4=4, \\ -x_1+3x_2-x_3+2x_4=3, \\ 4x_1-7x_2-5x_4=-5; \end{cases}$

(3) $\begin{cases} x_1+3x_2+5x_3-4x_4=1, \\ x_1+3x_2+2x_3-2x_4+x_5=-1, \\ x_1-2x_2+x_3-x_4-x_5=3, \\ x_1-4x_2+x_3+x_4-x_5=3, \\ x_1+2x_2+x_3-x_4+x_5=-1; \end{cases}$ (4) $\begin{cases} x_1+x_2-3x_3-x_4=1, \\ 3x_1-x_2-3x_3+4x_4=4, \\ x_1+5x_2-9x_3-8x_4=0. \end{cases}$

4．讨论 λ 为何值下列方程组有解,并求解:

(1) $\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1, \\ x_1 + 2x_2 - x_3 + 4x_4 = 2, \\ x_1 + 7x_2 - 4x_3 + 11x_4 = \lambda; \end{cases}$　　(2) $\begin{cases} \lambda x_1 + x_2 + x_3 = 1, \\ x_1 + \lambda x_2 + x_3 = \lambda, \\ x_1 + x_2 + \lambda x_3 = \lambda^2 . \end{cases}$

5．平面直角坐标系中,三条不同直线

$$ax + by + c = 0,\ bx + cy + a = 0,\ cx + ay + b = 0,$$

其中 a, b, c 互不相同,证明:三直线交于一点的充要条件是 $a + b + c = 0$.

6．设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是 n 阶方阵,且 $\boldsymbol{AB} = \boldsymbol{0}$,求证 $R(\boldsymbol{A}) + R(\boldsymbol{B}) \leq n$.

7．在 $\mathbf{R}^4$ 中,V_1 是由 $\alpha_1 = (1,2,1,0)^{\mathrm{T}}$, $\alpha_2 = (-1,1,1,1)^{\mathrm{T}}$ 生成的子空间,V_2 是由 $\beta_1 = (2,-1,0,1)^{\mathrm{T}}$, $\beta_2 = (1,-1,3,7)^{\mathrm{T}}$ 生成的子空间,求 V_1 与 V_2 的交构成的子空间的一个基及其维数 .

8．在 n 元行向量空间 $\mathbf{R}^n$ 中, V_1 与 V_2 分别是齐次方程组

$$x_1 + x_2 + \cdots + x_n = 0 \text{ 及 } x_1 = x_2 = \cdots = x_n$$

的解空间,证明 $\mathbf{R}^n = V_1 \oplus V_2$.

9．证明:方程组

$$\begin{cases} x_1 - x_2 = a_1, \\ x_2 - x_3 = a_2, \\ x_3 - x_4 = a_3, \\ x_4 - x_5 = a_4, \\ x_5 - x_1 = a_5 \end{cases}$$

有解的充要条件是

$$a_1 + a_2 + a_3 + a_4 + a_5 = 0,$$

并求其通解 .

10．设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为 n 阶方阵 .证明 $R(\boldsymbol{AB}) = R(\boldsymbol{B})$ 的充要条件是 $(\boldsymbol{AB})\boldsymbol{x} = \boldsymbol{0}$ 与 $\boldsymbol{Bx} = \boldsymbol{0}$ 同解 .

11．求一个齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$,使其解空间由 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 构成,其中

$$\boldsymbol{x}_1 = (1,2,-1,4)^{\mathrm{T}},\ \boldsymbol{x}_2 = (1,0,-4,0)^{\mathrm{T}} .$$

12．对任意实方阵 $\boldsymbol{A}$,证明:方程 $\boldsymbol{AA}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$ 同解 .进一步说明 $R(\boldsymbol{A}) = R(\boldsymbol{AA}^{\mathrm{T}})$.

13．问 a, b 为何值,线性方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0, \\ x_2 + 2x_3 + 2x_4 = 1, \\ -x_2 + (a-3)x_3 - 2x_4 = b, \\ 3x_1 + 2x_2 + x_3 + ax_4 = -1 \end{cases}$$

无解,有唯一解,有无穷多解?并求有无穷多解时的通解.

14. 设 n 阶方阵 $\boldsymbol{A}$ 的各行元素之和均为零且 $R(\boldsymbol{A}) = n - 1$,求线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的通解.

15. 设线性方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,2n}x_{2n} &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2,2n}x_{2n} &= 0, \\ &\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{n,2n}x_{2n} &= 0 \end{aligned}$$

的一个基础解系为 $\boldsymbol{a}_1 = (b_{11}, b_{12}, \cdots, b_{1,2n})^{\mathrm{T}}$, $\boldsymbol{a}_2 = (b_{21}, b_{22}, \cdots, b_{2,2n})^{\mathrm{T}}$, $\cdots$, $\boldsymbol{a}_n = (b_{n1}, b_{n2}, \cdots, b_{n,2n})^{\mathrm{T}}$,求线性方程组

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1,2n}x_{2n} &= 0, \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2,2n}x_{2n} &= 0, \\ &\cdots\cdots\cdots\cdots \\ b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{n,2n}x_{2n} &= 0 \end{aligned}$$

的通解.

16. 设地球表面三点 A, B, C 相互接近,使得可以用通过三点平面近似表示三点附近地球表面,经度与纬度起着笛卡尔坐标 x, y 作用,从 $A(120°20'10'', 45°30'11'')$ 及 $B(120°25'59'', 45°31'10'')$ 测得从正北到 AC 与 BC 连线的夹角分别为 $\alpha = 60°$, $\beta = 50°$,求 C 的经纬度.

习题 B

1. 用 Gauss 列主元消元法解线性方程组(保持 4 位有效数字):

(1) $\begin{cases} 4.462x_1 + 1.620x_2 + 0.8764x_3 = 2.031, \\ 3.421x_1 + 4.620x_2 + 3.201x_3 = 2.708, \\ 0.3764x_1 + 0.8654x_2 + 0.2146x_3 = 0.3604; \end{cases}$

(2) $\begin{cases} 0.3846x_1 + 0.02160x_2 + 0.3620x_3 = 0.4341, \\ 0.1940x_1 + 0.1230x_2 + 0.1941x_3 = 0.2477, \\ 0.1997x_1 + 0.2000x_2 + 0.2010x_3 = 0.2729. \end{cases}$

2. 设 $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{b} = (b_1, \cdots, b_m)^{\mathrm{T}}$, $\boldsymbol{y} = (y_1, \cdots, y_n)^{\mathrm{T}}$, $\boldsymbol{x} = (x_1, \cdots, x_m)^{\mathrm{T}}$,证明:

(1) 若 $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ 有解,则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$ 的任意解都满足方程 $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$;

(2) 方程组 $\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b}$ 有解的充要条件是方程组

$$\begin{pmatrix} \boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \boldsymbol{0} \\ 1 \end{pmatrix}$$

无解．

3．设 $\boldsymbol{A}$ 为 n 阶方阵，k 为正整数，线性方程组 $\boldsymbol{A}^k\boldsymbol{x}=\boldsymbol{0}$ 有解向量 $\boldsymbol{a}$ 且 $\boldsymbol{A}^{k-1}\boldsymbol{a}\neq\boldsymbol{0}$，证明：向量组 $\boldsymbol{a},\boldsymbol{A}\boldsymbol{a},\cdots,\boldsymbol{A}^{k-1}\boldsymbol{a}$ 线性无关．

4．(阅读附录 A.2)某经济系统有三个主要部门 D_1,D_2,D_3，在一个生产周期内，各部门之间直接消耗系数及最终产值如下表，求各部门总产值及各部门中间产品分配．

产出 直接消耗系数 投入	D_1	D_2	D_3	最终产品	总产品
D_1	0.25	0.10	0.10	2.35	
D_2	0.20	0.20	0.10	1.25	
D_3	0.10	0.10	0.20	2.10	

5．(阅读附录 A.2)某经济系统在一个周期内的分配与消耗如下表：

产出 投入		中间产品			最终产品	总产品
		D_1	D_2	D_3		
生产资料补偿价值	D_1	1 000	250	300	y_1	4 000
	D_2	800	500	300	y_2	2 500
	D_3	400	250	600	y_3	3 000
	固定资产折旧	100	50	100		
新创造价值		u_1	u_2	u_3		
总产品(万元)		4 000	2 500	3 000		

求各部门最终产品价值 y_1,y_2,y_3 及各部门新创造价值 u_1,u_2,u_3．

6．求拟合如下两组点的最小二乘的直线 $y=a+bx$：

(1) $(0,0),(1,2),(2,7)$；　(2) $(0,1),(2,0),(3,1),(3,2)$．

7．求 $y=a_0+a_1x+a_2x^2$，使其与点 $(2,0),(3,-10),(5,-48),(6,-76)$ 的误差平方和最小．

第五章　线性变换

这章讨论一个 n 维线性空间 V_n 到其自身的一类最简单的映射,即线性变换.当一个空间的基确定之后,一个线性变换与一个矩阵一一对应.线性变换的诸多性质可用矩阵的语言等价地描述出来,这不仅具有理论意义,也为其在计算机图形学等诸多方面的应用提供了有用的工具.线性空间与线性变换是线性代数的核心内容之一.

第一节　线性变换及其矩阵表示

1.1　线性变换的定义与例子

线性变换的最简单的例子就是一元线性函数

$$y = f(x) = ax.$$

它是线性空间 $\mathbf{R}^1$ 到 $\mathbf{R}^1$ 的映射,这样映射的特点是,对任意向量 $x_1, x_2, x \in \mathbf{R}^1$ 及数 $k \in \mathbf{R}$,下式成立

$$f(x_1 + x_2) = f(x_1) + f(x_2), \tag{5.1}$$

$$f(kx) = kf(x). \tag{5.2}$$

保持性质(5.1)及(5.2),将这类映射的定义域与值域由 $\mathbf{R}^1$ 推广到线性空间 V_n,就得到一般线性空间的线性变换的定义.

定义 5.1　设 σ 是 n 维线性空间 V_n 到其自身的映射,即对任意 $\alpha \in V_n$,$\sigma(\alpha) \in V_n$.如果 σ 具有如下两个性质:

(1) 任给 $\alpha_1, \alpha_2 \in V_n$, $\sigma(\alpha_1 + \alpha_2) = \sigma(\alpha_1) + \sigma(\alpha_2)$,

(2) 任给 $\alpha \in V_n$ 及 $k \in \mathbf{R}$, $\sigma(k\alpha) = k\sigma(\alpha)$,

则称 σ 是 V_n 的一个线性变换,$\sigma(\alpha)$ 称为 α 在线性变换 σ 下的象,而 α 称为 $\sigma(\alpha)$ 在线性变换 σ 下的原象.

定义 5.2　设 σ 与 τ 都是 n 维线性空间 V_n 的线性变换.如果对任意 $\alpha \in V_n$,都有 $\sigma(\alpha) = \tau(\alpha)$,则称两个线性变换相等,记为 $\sigma = \tau$.

下面给出 n 维线性空间 V_n 的线性变换的几个例子.

例 1　线性空间 V_n 上定义如下变换:

(1) 恒等变换 I:任给 $\alpha\in V_n$, $I(\alpha)=\alpha$;

(2) 数乘变换 Λ:给定 $\lambda\in R$,任给 $\alpha\in V_n$, $\Lambda(\alpha)=\lambda\alpha$;

(3) 零变换 θ:任给 $\alpha\in V_n$, $\theta(\alpha)=\mathbf{0}$,其中 $\mathbf{0}$ 是 V_n 的零向量.则容易验证这些变换都是 V_n 的线性变换.

例 2 平面旋转变换:在 2 维欧氏几何空间 $\mathbf{R}^2$ 中,每个向量绕坐标原点 O 按逆时针旋转 θ 角的变换 R_θ 是 $\mathbf{R}^2$ 的一个线性变换.

证 *(1) 从几何可知,任给平面向量 $\alpha,\beta\in\mathbf{R}^2$ 及数 $k\in\mathbf{R}$,则向量 $\alpha,\beta,\alpha+\beta$ 构成一个平行四边形,而 $k\alpha$ 是将向量 α 沿 α 的方向或反方向延长或缩短 $|k|$ 倍.将这四个向量 $\alpha,\beta,\alpha+\beta$ 及 $k\alpha$ 保持长度不变绕原点逆时针旋转 θ 角(见图 1),得到 $R_\theta(\alpha)$, $R_\theta(\beta)$, $R_\theta(\alpha+\beta)$, $R_\theta(k\alpha)$.则旋转前后这四个向量相对位置没改变,因而运算关系没变,所以

$$R_\theta(\alpha+\beta)=R_\theta(\alpha)+R_\theta(\beta),$$

$$R_\theta(k\alpha)=kR_\theta(\alpha).$$

这表明 R_θ 是一个线性变换

(2) 在 2 维欧氏几何空间 $\mathbf{R}^2$ 中,每个向量都可用坐标表示,设

$$\alpha=(x_1,x_2)^{\mathrm{T}},\ R_\theta(\alpha)=(y_1,y_2)^{\mathrm{T}},$$

由平面解析几何知

$$\begin{pmatrix}y_1\\y_2\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix},\tag{5.3}$$

即 $R_\theta(\alpha)=\boldsymbol{A}\alpha$,其中

$$\boldsymbol{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

由矩阵运算性质有

$$R_\theta(\alpha+\beta)=\boldsymbol{A}(\alpha+\beta)=\boldsymbol{A}(\alpha)+\boldsymbol{A}(\beta)=R_\theta(\alpha)+R_\theta(\beta),$$

$$R_\theta(k\alpha)=\boldsymbol{A}(k\alpha)=k\boldsymbol{A}(\alpha)=kR_\theta(\alpha).$$

由此证明了 R_θ 为 $\mathbf{R}^2$ 的线性变换.

例 3 在3维欧氏几何空间 $\mathbf{R}^3$ 中,将向量投影到一个坐标轴及坐标平面上的变换都是 $\mathbf{R}^3$ 的线性变换.

证 把 $\mathbf{R}^3$ 中的向量投影到 $\mathbf{R}^3$ 的 1 维子空间 $V_1=[\boldsymbol{i}]$的变换及投影到 2 维子空间 $V_2=[\boldsymbol{i}]\oplus[\boldsymbol{j}]$的变换,分别将任意向量

$$\alpha=(x_1,x_2,x_3)^{\mathrm{T}}=x_1\boldsymbol{i}+x_2\boldsymbol{j}+x_3\boldsymbol{k}$$

映射为

$$P_1(\alpha)=P_1(x_1\boldsymbol{i}+x_2\boldsymbol{j}+x_3\boldsymbol{k})=x_1\boldsymbol{i},$$

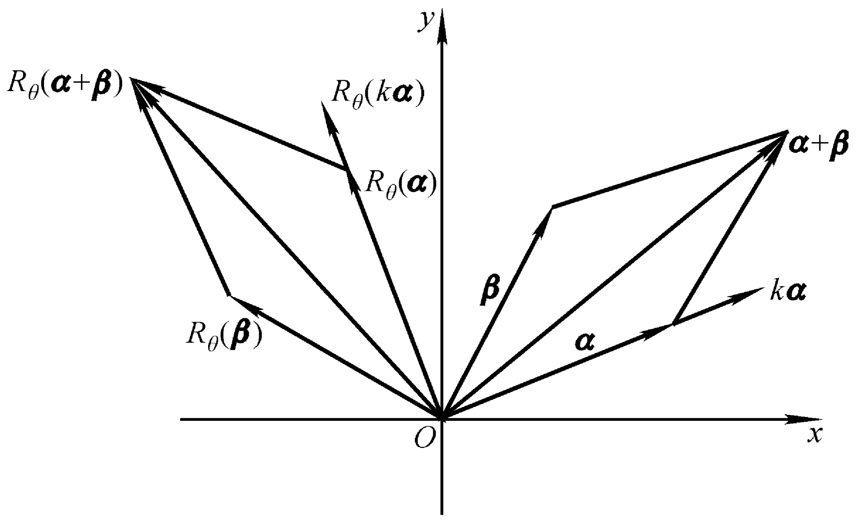


图 1

$$P_2(\alpha) = P_2(x_1 \boldsymbol{i} + x_2 \boldsymbol{j} + x_3 \boldsymbol{k}) = x_1 \boldsymbol{i} + x_2 \boldsymbol{j}.$$

很容易证明，P_1 及 P_2 均为 $\mathbf{R}^3$ 的线性变换．

1.2　线性变换的矩阵表示

1．线性变换在 V_n 的一个基下的矩阵表示

设 σ 是 n 维线性空间 V_n 的一个线性变换，$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$（这章的记号较多，为清楚起见，将空间的基记为集合形式）是 V_n 的一个基，σ 将 V_n 中向量 α 变换为 V_n 中向量 $\sigma(\alpha)$（简记为 $\sigma\alpha$），这时 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 被变换为 $\{\sigma\alpha_1, \sigma\alpha_2, \cdots, \sigma\alpha_n\}$．下面证明，一个 V_n 的线性变换完全被 V_n 的一个基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 及其象 $\{\sigma\alpha_1, \sigma\alpha_2, \cdots, \sigma\alpha_n\}$ 所确定．

定理 5.1　设 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 为 n 维线性空间 V_n 的一个基，σ 与 τ 为 V_n 的两个线性变换，且 $\sigma\alpha_i = \tau\alpha_i\ (i = 1, 2, \cdots, n)$，则 $\sigma = \tau$．

证　任给 $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \in V_n$，

$$\begin{aligned}\sigma\alpha &= \sigma(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) = x_1\sigma\alpha_1 + x_2\sigma\alpha_2 + \cdots + x_n\sigma\alpha_n \\ &= x_1\tau\alpha_1 + x_2\tau\alpha_2 + \cdots + x_n\tau\alpha_n = \tau(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) = \tau\alpha.\end{aligned}$$

由定义 5.2 知 $\sigma = \tau$．□

下面讨论线性变换 σ 在 n 维线性空间 V_n 的一个基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 下的矩阵表示．因为 σ 是 V_n 的线性变换，则 $\sigma\alpha_i \in V_n\ (i = 1, 2, \cdots, n)$，因而可将 $\sigma\alpha_i$ 也表示为 V_n 的基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的线性组合，即

$$\begin{cases}\sigma\alpha_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n, \\ \sigma\alpha_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \sigma\alpha_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n.\end{cases} \tag{5.4}$$

将$(\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_n)$记为$\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)$,于是

$$\begin{aligned}\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)&=(\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_n)\\&=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\\&=(\alpha_1,\alpha_2,\cdots,\alpha_m)\boldsymbol{A},\end{aligned}\tag{5.5}$$

其中$\boldsymbol{A}=(a_{ij})_{n\times n}$称为线性变换$\sigma$在$n$维线性空间$V_n$的基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵表示,或$\boldsymbol{A}$是$\sigma$在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下对应的矩阵.

例4 镜面反射变换R_e将2维欧氏几何空间$\mathbf{R}^2$中每个向量变换为关于一条过原点直线L对称的向量(图2).求R_e在基$\boldsymbol{i}=(1,0)^{\mathrm{T}}$,$\boldsymbol{j}=(0,1)^{\mathrm{T}}$下对应的矩阵.

解 设θ为$\boldsymbol{i}$所在坐标轴到L的转角,则L方向的单位向量为$\omega=\cos\theta\boldsymbol{i}+\sin\theta\boldsymbol{j}$,设$\alpha'=R_e\alpha$,

$$\alpha'=\alpha+2[(\alpha,\omega)\omega-\alpha]=-\alpha+2(\alpha,\omega)\omega,$$

其中(α,ω)为向量α与向量ω的内积.容易验证,R_e是线性变换.可直接利用上式来求R_e在基$\{\boldsymbol{i},\boldsymbol{j}\}$下对应的矩阵$\boldsymbol{A}$.令$\alpha=\boldsymbol{i},\boldsymbol{j}$,有

$$R_e\boldsymbol{i}=-\boldsymbol{i}+2(\boldsymbol{i},\omega)\omega=\boldsymbol{i}\cos2\theta+\boldsymbol{j}\sin2\theta,$$

$$R_e\boldsymbol{j}=-\boldsymbol{j}+2(\boldsymbol{j},\omega)\omega=\boldsymbol{i}\sin2\theta-\boldsymbol{j}\cos2\theta.$$

于是

$$(R_e\boldsymbol{i},R_e\boldsymbol{j},)=(\boldsymbol{i},\boldsymbol{j})\begin{pmatrix}\cos2\theta&\sin2\theta\\\sin2\theta&-\cos2\theta\end{pmatrix}=(\boldsymbol{i},\boldsymbol{j})\boldsymbol{A},$$

其中

$$\boldsymbol{A}=\begin{pmatrix}\cos2\theta&\sin2\theta\\\sin2\theta&-\cos2\theta\end{pmatrix}.$$

于是镜面反射变换在基$\{\boldsymbol{i},\boldsymbol{j}\}$下对应矩阵为$\boldsymbol{A}$.

注意,例2及例4都是讨论欧氏空间中的线性变换,其对应的矩阵$\boldsymbol{A}$都是正交矩阵,即$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$,它们对应的变换(旋转、镜面反射变换)都被称为正交变换.

以上讨论表明,n维线性空间V_n的任何一个线性变换,当V_n的基确定之后,都存在一个矩阵表示,即(5.5)成立.反过来,当V_n的基确定之后,任何一个实矩阵$\boldsymbol{A}=(a_{ij})_{n\times n}$能唯一地确定一个线性变换$\sigma$使(5.5)成立吗?回答是肯定的.(5.5)式可写成如下形式:

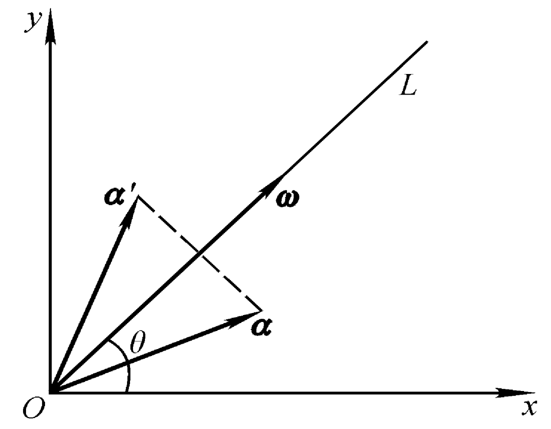


图 2

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{A} = (\gamma_1, \gamma_2, \cdots, \gamma_n). \qquad (5.6)$$

在 V_n 的基确定之后,任给矩阵 $\boldsymbol{A}$ 等价于任给 n 个向量 $\gamma_1, \gamma_2, \cdots, \gamma_n$.因此,上面提出的问题等价于如下问题:给定 V_n 一个基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$及 V_n 中 n 个向量 $\gamma_1, \gamma_2, \cdots, \gamma_n$,是否存在 V_n 的线性变换σ使$\sigma(\alpha_i) = \gamma_i (i = 1, 2, \cdots, n)$.下面定理回答了这个问题.

定理 5.2　设$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是 n 维线性空间 V_n 的一个基,在 V_n 中任意给定 n 个向量 $\gamma_1, \gamma_2, \cdots, \gamma_n$,则存在唯一的线性变换 σ,使得

$$\sigma(\alpha_i) = \gamma_i (i = 1, 2, \cdots, n).$$

证　先定义一个 V_n 的变换σ,它使(5.6)式成立,然后再证明 σ是满足该性质的唯一的线性变换.

设 $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \in V_n$,其中 $x_i \in \mathbf{R} (i = 1, 2, \cdots, n)$.定义变换 σ:

$$\sigma\alpha = x_1\gamma_1 + x_2\gamma_2 + \cdots + x_n\gamma_n.$$

特别地,对 α_i 有

$$\sigma(\alpha_i) = \gamma_i, (i = 1, 2, \cdots, n).$$

任给 V_n 中两个向量

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n,$$

$$\alpha' = x_1'\alpha_1 + x_2'\alpha_2 + \cdots + x_n'\alpha_n,$$

及数 $k \in \mathbf{R}$,可知

$$\begin{aligned}\sigma(\alpha + \alpha') &= \sigma((x_1 + x_1')\alpha_1 + (x_2 + x_2')\alpha_2 + \cdots + (x_n + x_n')\alpha_n) \\ &= (x_1 + x_1')\gamma_1 + (x_2 + x_2')\gamma_2 + \cdots + (x_n + x_n')\gamma_n \\ &= (x_1\gamma_1 + x_2\gamma_2 + \cdots + x_n\gamma_n) + (x_1'\gamma_1 + x_2'\gamma_2 + \cdots + x_n'\gamma_n) \\ &= \sigma\alpha + \sigma\alpha'.\end{aligned}$$

类似可证 $\sigma(k\alpha) = k\sigma\alpha$.于是,$\sigma$ 是 V_n 的一个线性变换.由定理 5.1 可知,σ是具

有所要求性质的唯一的线性变换.□

设$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$是 n 维线性空间 V_n 的一个基,σ在该基下的矩阵表示为 $\boldsymbol{A}$, V_n 中任意向量α在该基下的坐标为 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, $\sigma\alpha$在该基下的坐标为 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$,那么 $\boldsymbol{y}$ 如何用 $\boldsymbol{x}$ 表达呢?

已知

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n,$$

所以

$$\begin{aligned}
&(\alpha_1,\alpha_2,\cdots,\alpha_n)\boldsymbol{y}\\
&= y_1\alpha_1 + y_2\alpha_2 + \cdots + y_n\alpha_n = \sigma\alpha\\
&= \sigma(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) = x_1\sigma\alpha_1 + x_2\sigma\alpha_2 + \cdots + x_n\sigma\alpha_n\\
&= (\sigma\alpha_1,\sigma\alpha_2,\cdots,\sigma\alpha_n)\boldsymbol{x}\\
&= (\alpha_1,\alpha_2,\cdots,\alpha_n)\boldsymbol{A}\boldsymbol{x}.
\end{aligned}$$

由于 $\boldsymbol{y}$ 与 $\boldsymbol{A}\boldsymbol{x}$ 都是同一个向量α在同一个基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的坐标,因而

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

这样,可得如下定理:

定理 5.3　设$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$为 n 维线性空间 V_n 的一个基, $\boldsymbol{A}$ 是线性变换σ的矩阵表示,向量 α 的坐标为 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, $\sigma\alpha$ 坐标为 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$,则

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{5.7}$$

例 5　设σ为 n 维欧氏空间 V_n 的正交变换,即σ在 V_n 的一个标准正交基下的矩阵表示 $\boldsymbol{A}$ 是正交矩阵.证明:在正交变换σ下, V_n 中任意两向量夹角及每个向量的长度不变.

证　设 $\boldsymbol{a}$, $\boldsymbol{b}$ 为 V_n 中任意两个向量,在 V_n 的一个标准正交基中它们的坐标分别为 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$.由定理 5.3 知,$\sigma(\boldsymbol{x}_1)$及$\sigma(\boldsymbol{x}_2)$在同一基下的坐标为 $\boldsymbol{A}\boldsymbol{x}_1$, $\boldsymbol{A}\boldsymbol{x}_2$.因 $\boldsymbol{A}$ 为正交矩阵,即 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$,于是

$$(\boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{A}\boldsymbol{x}_2) = (\boldsymbol{A}\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{x}_2 = (\boldsymbol{x}_1,\boldsymbol{x}_2)$$

这说明在正交变换σ下, V_n 中任意两个向量内积不变,因此任意两个向量夹角及每个向量的长度也不变.

由例 5 可知,3 维欧氏几何空间 $\mathbf{R}^3$ 中,在正交变换下,任意两点距离不变,从而一个曲面或一条曲线的形状也不发生变化,这是正交变换的最重要的性质之一.

例 6　设 $\boldsymbol{i}=(1,0)^{\mathrm{T}}$, $\boldsymbol{j}=(0,1)^{\mathrm{T}}$ 为 $\mathbf{R}^2$ 中的标准正交基,线性变换σ使 $\boldsymbol{i}$ 逆时针方向旋转$\dfrac{\pi}{6}$变为 $\boldsymbol{i}'$,使 $\boldsymbol{j}$ 顺时针旋转$\dfrac{\pi}{6}$变为 $\boldsymbol{j}'$,求σ在基$\{\boldsymbol{i},\boldsymbol{j}\}$下的矩阵表

示.

解　因为

$$\sigma(\boldsymbol{i})=\boldsymbol{i}'=\cos\frac{\pi}{6}\boldsymbol{i}+\sin\frac{\pi}{6}\boldsymbol{j}=\frac{\sqrt{3}}{2}\boldsymbol{i}+\frac{1}{2}\boldsymbol{j},$$

$$\sigma(\boldsymbol{j})=\boldsymbol{j}'=\cos\frac{\pi}{3}\boldsymbol{i}+\sin\frac{\pi}{3}\boldsymbol{j}=\frac{1}{2}\boldsymbol{i}+\frac{\sqrt{3}}{2}\boldsymbol{j}.$$

于是

$$(\sigma(\boldsymbol{i}),\sigma(\boldsymbol{j}))=(\boldsymbol{i},\boldsymbol{j})\begin{pmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}=(\boldsymbol{i},\boldsymbol{j})\boldsymbol{A},$$

其中

$$\boldsymbol{A}=\begin{pmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}.$$

因此,线性变换 σ 在基 $\{\boldsymbol{i},\boldsymbol{j}\}$ 下的矩阵表示为 $\boldsymbol{A}$.

2. 线性变换在两个不同基下的矩阵表示之间的关系

设 n 维线性空间 V_n 中有两个基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 与 $\{\alpha_1',\alpha_2',\cdots,\alpha_n'\}$,把前者称为旧基,后者称为新基,且

$$(\alpha_1',\alpha_2',\cdots,\alpha_n')=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & & \cdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix},$$

则 $\boldsymbol{C}=(c_{ij})_{n\times n}$ 称为由旧基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 到新基 $\{\alpha_1',\alpha_2',\cdots,\alpha_n'\}$ 的过渡矩阵,简记为

$$(\alpha_1',\alpha_2',\cdots,\alpha_n')=(\alpha_1,\alpha_2,\cdots,\alpha_n)\boldsymbol{C}. \tag{5.8}$$

由于 $\alpha_1',\alpha_2',\cdots,\alpha_n'$ 也为 V_n 的基,所以存在 n 阶方阵 $\boldsymbol{Q}$,使得

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1',\alpha_2',\cdots,\alpha_n')\boldsymbol{Q}. \tag{5.9}$$

将(5.9)代入(5.8)有

$$(\alpha_1',\alpha_2',\cdots,\alpha_n')=(\alpha_1',\alpha_2',\cdots,\alpha_n')\boldsymbol{QC},$$

于是 $\boldsymbol{QC}=\boldsymbol{I}$,这表明旧基到新基的过渡矩阵 $\boldsymbol{C}$ 可逆.

如果 V_n 的一个线性变换 σ 在旧基和新基下的矩阵表示分别为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,那么 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有什么关系?

由(5.8)知

$$(\alpha_1', \alpha_2', \cdots, \alpha_n') = (\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{C},$$

$$\sigma(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{A},$$

$$\sigma(\alpha_1', \alpha_2', \cdots, \alpha_n') = (\alpha_1', \alpha_2', \cdots, \alpha_n')\boldsymbol{B}.$$

则

$$\begin{aligned}(\alpha_1', \alpha_2', \cdots, \alpha_n')\boldsymbol{B} &= \sigma(\alpha_1', \alpha_2', \cdots, \alpha_n') = \sigma((\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{C}) \\ &= (\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{AC} = (\alpha_1', \alpha_2', \cdots, \alpha_n')\boldsymbol{C}^{-1}\boldsymbol{AC}\end{aligned}$$

于是有下面定理:

定理 5.4　设 n 维线性空间 V_n 的线性变换 σ 在 V_n 的旧基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 与新基 $\{\alpha'_1, \alpha'_2, \cdots, \alpha'_n\}$ 下的矩阵表示分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 且旧基到新基的过渡矩阵为 $\boldsymbol{C}$, 即

$$(\alpha_1', \alpha_2', \cdots, \alpha_n') = (\alpha_1, \alpha_2, \cdots, \alpha_n)\boldsymbol{C},$$

则 $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{AC}$.

例 7　设3维欧氏几何空间 $\mathbf{R}^3$ 的线性变换 σ 把基

$$\alpha_1 = (1,0,1)^{\mathrm{T}}, \alpha_2 = (0,1,0)^{\mathrm{T}}, \alpha_3 = (0,0,1)^{\mathrm{T}}$$

变为

$$\beta_1 = (1,0,2)^{\mathrm{T}}, \beta_2 = (-1,2,-1)^{\mathrm{T}}, \beta_3 = (1,0,0)^{\mathrm{T}},$$

求 σ 对应基

$$\alpha_1' = (1,0,0)^{\mathrm{T}}, \alpha_2' = (0,1,0)^{\mathrm{T}}, \alpha_3' = (0,0,1)^{\mathrm{T}}$$

的矩阵.

解

$$\begin{aligned}\sigma\alpha_1 &= \beta_1 = (1,0,2)^{\mathrm{T}} = \alpha_1 + 0\alpha_2 + \alpha_3, \\ \sigma\alpha_2 &= \beta_2 = (-1,2,-1)^{\mathrm{T}} = -\alpha_1 + 2\alpha_2 + 0\alpha_3, \\ \sigma\alpha_3 &= \beta_3 = (1,0,0)^{\mathrm{T}} = \alpha_1 + 0\alpha_2 - \alpha_3.\end{aligned}$$

令

$$(\sigma(\alpha_1), \sigma(\alpha_2), \sigma(\alpha_3)) = \sigma(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3)\boldsymbol{A}$$

则 σ 在基 $\{\alpha_1, \alpha_2, \alpha_3\}$ 下的矩阵

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

又由

$$(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1', \alpha_2', \alpha_3')\boldsymbol{D}$$

其中

$$\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

新基$\{\alpha_1', \alpha_2', \alpha_3'\}$到旧基$\{\alpha_1, \alpha_2, \alpha_3\}$的过渡矩阵表示为 $\boldsymbol{C}$,即

$$(\alpha_1', \alpha_2', \alpha_3') = (\alpha_1, \alpha_2, \alpha_3)\boldsymbol{C},$$

则 $\boldsymbol{C} = \boldsymbol{D}^{-1}$.由定理 5.4 知 σ在基$\{\alpha_1', \alpha_2', \alpha_3'\}$下的矩阵

$$\begin{aligned} \boldsymbol{B} &= \boldsymbol{C}^{-1}\boldsymbol{AC} = \boldsymbol{DAD}^{-1} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 2 & 0 \\ 2 & -1 & 0 \end{pmatrix}. \end{aligned}$$

* 第二节　线性变换的象空间与零空间

对实数域上 n 维线性空间 V_n 中每个线性变换σ而言,有两个 V_n 的子空间相伴而生,即 σ的象空间及零空间,它们对深入研究线性变换有重要意义.

定义 5.3　设 σ是 n 维线性空间 V_n 上的线性变换,

$$R(\sigma) = \{\sigma(\alpha) \mid \alpha \in V_n\},$$

$$N(\sigma) = \{\alpha \mid \sigma(\alpha) = \boldsymbol{0}, \alpha \in V_n\},$$

即 $R(\sigma)$是 V_n 中所有向量在σ下的象构成的集合, $N(\sigma)$是 V_n 中零元的所有原象构成的集合.有时也将 $R(\sigma)$及 $N(\sigma)$分别记为 $\mathrm{Im}(\sigma)$及 $\mathrm{Ker}(\sigma)$.

下面证明 $R(\sigma)$及 $N(\sigma)$都是 V_n 的子空间.

定理 5.5　设σ为 n 维线性空间 V_n 上的一个线性变换,则 $R(\sigma)$及 $N(\sigma)$都是 V_n 的子空间.

证　1 $\sigma(\boldsymbol{0}) = \boldsymbol{0}$,所以 $\boldsymbol{0} \in R(\sigma)$,即 $R(\sigma)$是 V_n 的非空子集.令 $\beta_1, \beta_2 \in R(\sigma)$及 $k \in R$,存在 $\alpha_1, \alpha_2 \in V_n$,使 $\sigma(\alpha_1) = \beta_1$, $\sigma(\alpha_2) = \beta_2$,于是

$$\beta_1 + \beta_2 = \sigma(\alpha_1) + \sigma(\alpha_2) = \sigma(\alpha_1 + \alpha_2),$$

$$k\beta_1 = k\sigma(\alpha_1) = \sigma(k\alpha_1).$$

这表明 $\beta_1 + \beta_2 \in R(\sigma)$, $k\beta_1 \in R(\sigma)$,由定理 1.1 可知 $R(\sigma)$为 V_n 的一个子空间.

2 因为 $\sigma(\boldsymbol{0}) = \boldsymbol{0}$,所以 $\boldsymbol{0} \in N(\sigma)$,即 $N(\sigma)$是 V_n 的非空子集.令 $\alpha_1, \alpha_2 \in N(\sigma)$,即$\sigma(\alpha_1) = \sigma(\alpha_2) = \boldsymbol{0}$,则对任意 $k \in R$,

$$\sigma(\alpha_1 + \alpha_2) = \sigma(\alpha_1) + \sigma(\alpha_2) = \boldsymbol{0} + \boldsymbol{0},$$

$$\sigma(k\alpha_1) = k\sigma(\alpha_1) = k\boldsymbol{0} = \boldsymbol{0}.$$

这表明 $\alpha_1+\alpha_2\in N(\sigma)$，$k\alpha_1\in N(\sigma)$，由定理 1.1 可知 $N(\sigma)$为 V_n 的一个子空间.□

定义 5.4　设 σ为 n 维线性空间 V_n 上的一个线性变换，V_n 的子空间 $R(\sigma)$及 $N(\sigma)$分别称为线性变换 σ的象空间和零空间(或核空间)，它们的维数分别称为 σ的秩及 σ的零度，且分别记为 $\mathrm{rank}(\sigma)$及 $\mathrm{null}(\sigma)$.

关于线性变换的秩与零度有如下定理.

定理 5.6　设 σ为 n 维线性空间 V_n 上的一个线性变换，则 $\mathrm{rank}(\sigma)+\mathrm{null}(\sigma)=n$.

证　设 σ的零空间 $N(\sigma)$的一个基为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$，并将其扩充为 V_n 的一个基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$.任给 $\alpha\in V_n$，则 α可写成该基的线性组合，

$$\alpha=x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+\cdots+x_r\boldsymbol{e}_r+x_{r+1}\boldsymbol{e}_{r+1}+\cdots+x_n\boldsymbol{e}_n.$$

因为 $\sigma\boldsymbol{e}_i=\boldsymbol{0}$，$i=1,2,\cdots,r$，所以

$$\begin{aligned}\sigma\alpha&=x_1\sigma\boldsymbol{e}_1+x_2\sigma\boldsymbol{e}_2+\cdots+x_r\sigma\boldsymbol{e}_r+x_{r+1}\sigma\boldsymbol{e}_{r+1}+\cdots+x_n\sigma\boldsymbol{e}_n\\&=x_{r+1}\sigma\boldsymbol{e}_{r+1}+\cdots+x_n\sigma\boldsymbol{e}_n.\end{aligned}$$

这表明 $R(\sigma)$任意一个向量都是 $\sigma\boldsymbol{e}_{r+1},\cdots,\sigma\boldsymbol{e}_n$ 的线性组合.下面证明 $\sigma\boldsymbol{e}_{r+1},\cdots,\sigma\boldsymbol{e}_n$ 线性无关.令

$$k_{r+1}\sigma\boldsymbol{e}_{r+1}+\cdots+k_n\sigma\boldsymbol{e}_n=\boldsymbol{0}.$$

于是 $\sigma(k_{r+1}\boldsymbol{e}_{r+1}+\cdots+k_n\boldsymbol{e}_n)=\boldsymbol{0}$.这说明 $k_{r+1}\boldsymbol{e}_{r+1}+\cdots+k_n\boldsymbol{e}_n\in N(\sigma)$，所以它可以写成 $N(\sigma)$的基 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r$ 的线性组合，即

$$k_{r+1}\boldsymbol{e}_{r+1}+\cdots+k_n\boldsymbol{e}_n=l_1\boldsymbol{e}_1+l_2\boldsymbol{e}_2+\cdots+l_r\boldsymbol{e}_r.$$

也即

$$l_1\boldsymbol{e}_1+l_2\boldsymbol{e}_2+\cdots+l_r\boldsymbol{e}_r-k_{r+1}\boldsymbol{e}_{r+1}-\cdots-k_n\boldsymbol{e}_n=\boldsymbol{0}.$$

因为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r,\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n$ 为 n 维线性空间 V_n 的一个基，所以 $l_i=0$，$i=1,2,\cdots,r$，$k_j=0$，$j=r+1,r+2,\cdots,n$.这就证明了向量 $\sigma\boldsymbol{e}_{r+1},\cdots,\sigma\boldsymbol{e}_n$ 是线性无关的，它们构成了 $R(\sigma)$的一个基.因此 $\dim(N(\sigma))=r$，$\dim(R(\sigma))=n-r$，即

$$\dim(N(\sigma))+\dim(R(\sigma))=n.\square$$

当 $\mathrm{rank}(\sigma)=n$ 时，称 σ为满秩线性变换.当 σ为满秩线性变换，$R(\sigma)=V_n$，即 σ为满秩，$N(\sigma)=\{\boldsymbol{0}\}$，即 σ为单射，因此 σ为一一对应的线性变换.

例 1　三维几何空间 $\mathbf{R}^3$ 的线性变换 σ在标准正交基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$下的矩阵表示为

$$\boldsymbol{A}=\begin{pmatrix}1&2&3\\1&3&4\\2&5&7\end{pmatrix},$$

求 $R(\sigma)$, $N(\sigma)$, $\mathrm{rank}(\sigma)$及 $\mathrm{null}(\sigma)$.

解　设向量 α 在基$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$下的坐标 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$, $\sigma(\alpha)$的坐标为 $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}}$.由定理 5.3 知道, $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$.零空间 $N(\sigma)$的向量的坐标满足方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, 即

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 2 & 5 & 7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

其基础解系 $\boldsymbol{x}_1 = (1, 1, -1)^{\mathrm{T}}$, 因此, $\mathrm{null}(\sigma) = 1$ 且 $N(\sigma)$是由坐标为 $\boldsymbol{x}_1$ 的向量 $\boldsymbol{e}_1 + \boldsymbol{e}_2 - \boldsymbol{e}_3$ 生成的 $\mathbf{R}^3$ 的子空间 .

令 $\boldsymbol{A}$ 的列向量为 $\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3$, 对 $\boldsymbol{A}$ 进行初等列变换

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 2 & 5 & 7 \end{pmatrix} \xrightarrow{c_3 - c_1 - c_2} \begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 0 \\ 2 & 5 & 0 \end{pmatrix}.$$

这表明 $\boldsymbol{c}_1, \boldsymbol{c}_2$ 是 $\boldsymbol{A}$ 的列向量组的极大线性无关组 .公式 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 可写为

$$\boldsymbol{y} = x_1\boldsymbol{c}_1 + x_2\boldsymbol{c}_2 + x_3\boldsymbol{c}_3 = (x_1 + x_3)\boldsymbol{c}_1 + (x_2 + x_3)\boldsymbol{c}_2 .$$

于是 $R(\sigma)$中的向量的坐标 $\boldsymbol{y}$ 均可写为 $\boldsymbol{c}_1, \boldsymbol{c}_2$ 的线性组合, 且 $\boldsymbol{c}_1$ 与 $\boldsymbol{c}_2$ 线性无关 .因而 $\mathrm{rank}(\sigma) = 2$, 且 $R(\sigma)$是由坐标为 $\boldsymbol{c}_1$ 及 $\boldsymbol{c}_2$ 的两个向量, 即向量 $\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{e}_3$ 及 $2\boldsymbol{e}_1 + 3\boldsymbol{e}_2 + 5\boldsymbol{e}_3$, 生成的子空间 .

*第三节　线性变换的运算

本节主要研究 n 维线性空间 V_n 上线性变换的运算及其矩阵表示 .对定义在 V_n 上的线性变换, 可以定义加法与数乘两种运算, 并且 V_n 上所有线性变换形成 R 上的一个线性空间 .

3.1　线性变换的加法与数乘

设 V_n 是实数域上的 n 维线性空间, σ为 V_n 的一个线性变换 .如果对任意 $\alpha \in V_n$, $\sigma\alpha$ 都已知, 那么 σ也就唯一地确定了 .

定义 5.5　设 σ与 τ为实数域 $\mathbf{R}$ 上的 n 维线性空间 V_n 的两个线性变换, λ是一个实数 .定义 σ与 τ之和 $\sigma + \tau$, λ与σ乘积 $\lambda\sigma$ 如下: 任给 $\alpha \in V_n$,

$$(\sigma + \tau)\alpha = \sigma\alpha + \tau\alpha, (\lambda\sigma)\alpha = \lambda\sigma\alpha .$$

很容易验证, $\sigma + \tau$, $\lambda\sigma$ 都是 V_n 的线性变换 .

本章第一节已定义 V_n 的零变换 θ: 对任何 $\alpha \in V_n$, $\theta\alpha = \boldsymbol{0}$.

设 σ 为 V_n 的线性变换,定义 $-\sigma$:任给 $\alpha\in V$,$(-\sigma)\alpha=-\sigma\alpha$.于是 $\sigma+(-\sigma)=\theta$(零变换).

考虑 $\mathbf{R}$ 上线性空间 V_n 的所有线性变换构成的集合 $\mathrm{Hom}(V_n,V_n)$,有下面命题,证明留给读者.

命题 5.1 $\mathrm{Hom}(V_n,V_n)$是实数域 $\mathbf{R}$ 上的线性空间.

设 σ 和 τ 在线性空间 V_n 的一个基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下对应的矩阵分别为 $\boldsymbol{A}=(a_{ij})_{n\times n}$和 $\boldsymbol{B}=(b_{ij})_{n\times n}$,那么 $\sigma+\tau$ 及 $k\sigma$ 在该基下对应的矩阵是什么呢?因为

$$\sigma\alpha_j=\sum_{i=1}^{n}a_{ij}\alpha_i,\ \tau\alpha_j=\sum_{i=1}^{n}b_{ij}\alpha_i,\ j=1,2,\cdots,n.$$

所以

$$\begin{aligned}(\sigma+\tau)\alpha_j&=\sigma\alpha_j+\tau\alpha_j=\sum_{i=1}^{n}a_{ij}\alpha_i+\sum_{i=1}^{n}b_{ij}\alpha_i\\&=\sum_{i=1}^{n}(a_{ij}+b_{ij})\alpha_i,\ j=1,2,\cdots,n.\end{aligned}$$

这样 $\sigma+\tau$ 对应的矩阵为$(a_{ij}+b_{ij})_{n\times n}=\boldsymbol{A}+\boldsymbol{B}$.同理可证 $\lambda\sigma$ 对应的矩阵为$\lambda\boldsymbol{A}$.

3.2 线性变换的乘积

定义 5.6 设 σ,τ 是实数域 $\mathbf{R}$ 上的 n 维线性空间 V_n 的两个线性变换.对任意 $\alpha\in V_n$,定义$(\sigma\tau)\alpha=\sigma(\tau\alpha)$,称 $\sigma\tau$ 为 σ 与 τ 的乘积.

下面证明 $\sigma\tau$ 是 V_n 的线性变换,对任意 $\alpha,\beta\in V_n$,$k\in\mathbf{R}$,有

$$(\sigma\tau)(\alpha+\beta)=\sigma(\tau(\alpha+\beta))=\sigma(\tau\alpha+\tau\beta)=\sigma\tau\alpha+\sigma\tau\beta=(\sigma\tau)\alpha+(\sigma\tau)\beta,$$
$$(\sigma\tau)(k\alpha)=\sigma(\tau(k\alpha))=\sigma(k\tau\alpha)=k\sigma\tau\alpha=k(\sigma\tau)\alpha,$$

所以 $\sigma\tau$ 是 V_n 的线性变换.

设 I 为 V_n 上的恒等变换,即对任意 $\alpha\in V_n$,有 $I\alpha=\alpha$.于是对 V_n 上的任意线性变换σ,$I\sigma=\sigma$.一般说来,线性变换的积不满足交换律.例如,在几何空间 $\mathbf{R}^2$ 中,$\{\boldsymbol{i},\boldsymbol{j}\}$是一个标准正交基:

$$\boldsymbol{i}=(1,0)^{\mathrm{T}},\ \boldsymbol{j}=(0,1)^{\mathrm{T}}.$$

σ 是将 $\mathbf{R}^2$ 的向量绕原点反时针旋转$\dfrac{\pi}{2}$的线性变换,而 ρ 是将向量投影到横轴的线性变换.那么,

$$\sigma\rho\boldsymbol{i}=\sigma(\rho\boldsymbol{i})=\sigma\boldsymbol{i}=\boldsymbol{j},\ \rho\sigma\boldsymbol{i}=\rho(\sigma\boldsymbol{i})=\rho\boldsymbol{j}=(0,0)^{\mathrm{T}},$$

可见 $\sigma\rho\neq\rho\sigma$.

设 σ 和 τ 为 V_n 的两个线性变换,I 为 V_n 上恒等线性变换,如果 $\sigma\tau=\tau\sigma=$

I,称 σ 为可逆线性变换,τ 称为 σ 的逆变换,记为 σ^{-1}.

定理 5.7　设 σ,τ 为实数域 $\mathbf{R}$ 上线性空间 V_n 的两个线性变换,在 V_n 的一个基下,σ 与 τ 的矩阵表示分别为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$,则乘积 $\sigma\tau$ 在该基下矩阵表示为 $\boldsymbol{AB}$.

证　设 σ 和 τ 在 V_n 的一个基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵表示分别为 $\boldsymbol{A}=(a_{ij})_{n\times n}$,$\boldsymbol{B}=(b_{ij})_{n\times n}$,则

$$(\sigma\tau)\alpha_j=\sigma(\tau\alpha_j)=\sigma\left(\sum_{i=1}^n b_{ij}\alpha_i\right)=\sum_{i=1}^n b_{ij}\sigma(\alpha_i)$$

$$=\sum_{i=1}^n b_{ij}\left(\sum_{k=1}^n a_{ki}\alpha_k\right)=\sum_{k=1}^n\left(\sum_{i=1}^n a_{ki}b_{ij}\right)\alpha_k.$$

令 $c_{kj}=\sum_{i=1}^n a_{ki}b_{ij}$ 且 $\boldsymbol{C}=(c_{kj})_{n\times n}$,则 $\boldsymbol{AB}=\boldsymbol{C}$.上式表明,$\sigma\tau$ 在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下的矩阵表示为 $\boldsymbol{C}=\boldsymbol{AB}$.□

恒等变换在基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$下对应的矩阵表示为 n 阶单位矩阵,如果可逆线性变换 σ 在该基下对应的矩阵为 $\boldsymbol{A}$,由定理 5.7 可知,在该基下其逆变换 σ^{-1} 对应的矩阵为 $\boldsymbol{A}^{-1}$.

若 σ 是可逆线性变换,则 σ 是单射变换.这是因为:设 $\sigma\alpha=\sigma\beta$,则 $\sigma(\alpha-\beta)=\mathbf{0}$,因为 σ^{-1} 存在,所以

$$\alpha-\beta=\sigma^{-1}\sigma(\alpha-\beta)=\sigma^{-1}\mathbf{0}=\mathbf{0},$$

于是 $\alpha=\beta$;其次,σ 是一个满射:因为任给 $\alpha\in V_n$,由于 σ^{-1} 存在,令 $\beta=\sigma^{-1}\alpha\in V_n$,则 $\sigma\beta=\sigma\sigma^{-1}\alpha=\alpha$;因此 σ 是一一对应的线性变换.由上节讨论知,σ 是一一对应的线性变换的充要条件是 σ 为满秩变换.因此可得如下定理.

定理 5.8　设 σ 为 n 维线性空间 V_n 上的线性变换,下面命题等价:(1) σ 是可逆的线性变换;(2) σ 是满秩的线性变换;(3) σ 是一一对应的线性变换.

由定理 5.7 可知,满秩线性变换在 V_n 的任意一个基下的矩阵表示为一个可逆矩阵,即秩为 n 的矩阵.

例 1　三维欧氏几何空间 $\mathbf{R}^3$ 的线性变换 R_θ 将向量绕 z 轴旋转 θ 角(从 z 轴向原点方向看,将一个向量按逆时针旋转 θ 角),而线性变换 R_ϕ 将向量绕 y 轴旋转 ϕ 角(按逆时针旋转 ϕ 角),求 $R_\phi R_\theta$ 在 $\mathbf{R}^3$ 的基 $\boldsymbol{i}=(1,0,0)^{\mathrm{T}}$,$\boldsymbol{j}=(0,1,0)^{\mathrm{T}}$,$\boldsymbol{k}=(0,0,1)^{\mathrm{T}}$ 下的矩阵表示.

解　对 R_θ 而言

$$R_\theta(\boldsymbol{i})=\boldsymbol{i}\cos\theta+\boldsymbol{j}\sin\theta,$$

$$R_\theta(\boldsymbol{j})=-\boldsymbol{i}\sin\theta+\boldsymbol{j}\cos\theta,$$

$$R_\theta(\boldsymbol{k})=\boldsymbol{k}.$$

因此

$$R_\theta(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = (\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix}.$$

对 R_φ **而言**

$$R_\varphi(\boldsymbol{i}) = \boldsymbol{i}\cos\varphi - \boldsymbol{k}\sin\varphi,$$
$$R_\varphi(\boldsymbol{j}) = \boldsymbol{j},$$
$$R_\varphi(\boldsymbol{k}) = \boldsymbol{i}\sin\varphi + \boldsymbol{k}\cos\varphi.$$

因此

$$R_\varphi(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}) = (\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})\begin{pmatrix}\cos\varphi & 0 & \sin\varphi\\ 0 & 1 & 0\\ -\sin\varphi & 0 & \cos\varphi\end{pmatrix}.$$

由线性变换 R_φ **和** R_θ **在标准正交基**$\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$**下的矩阵表示,可得** $R_\varphi R_\theta$ **在该基下的矩阵表示是**

$$\begin{pmatrix}\cos\varphi & 0 & \sin\varphi\\ 0 & 1 & 0\\ -\sin\varphi & 0 & \cos\varphi\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{pmatrix} =$$
$$\begin{pmatrix}\cos\varphi\cos\theta & -\cos\varphi\sin\theta & \sin\varphi\\ \sin\theta & \cos\theta & 0\\ -\sin\varphi\cos\theta & \sin\varphi\sin\theta & \cos\varphi\end{pmatrix}.$$

例 2　**假设** n **维线性空间** V_n **是它的子空间** S_1 **与** S_2 **的直和,即** $V_n = S_1 \oplus S_2$**,任给** $\alpha \in V_n$**,有** $\alpha = \alpha_1 + \alpha_2$**,且** $\alpha_1 \in S_1$, $\alpha_2 \in S_2$**,试证** $P_1\alpha = \alpha_1$, $P_2\alpha = \alpha_2$ **是** V_n **上的线性变换,且** $P_1^2 = P_1$, $P_2^2 = P_2$, $P_1P_2 = P_2P_1 = \theta$**(零变换)**.

证　**对任意** $\alpha = \alpha_1 + \alpha_2$, $\beta = \beta_1 + \beta_2$**,其中** $\alpha_1, \beta_1 \in S_1$, $\alpha_2, \beta_2 \in S_2$, $k \in R$**,有**

$$P_1(\alpha + \beta) = P_1((\alpha_1 + \beta_1) + (\alpha_2 + \beta_2)) = \alpha_1 + \beta_1 = P_1\alpha + P_1\beta,$$
$$P_1(k\alpha) = P_1(k\alpha_1 + k\alpha_2) = k\alpha_1 = kP_1\alpha.$$

因此 P_1 **为** V_n **上的线性变换.同理可证,** P_2 **也为** V_n **上线性变换.还有**

$$P_1^2\alpha = P_1P_1(\alpha_1 + \alpha_2) = P_1\alpha_1 = \alpha_1 = P_1\alpha,$$
$$P_1P_2\alpha = P_1P_2(\alpha_1 + \alpha_2) = P_1\alpha_2 = \mathbf{0} = \theta\alpha.$$

所以 $P_1^2 = P_1$, $P_1P_2 = \theta$**.同理可证,** $P_2^2 = P_2$, $P_2P_1 = \theta$.□

满足例 2 所列举性质的 V_n **上线性变换称为投影变换.如果** $V_n = S_1 \oplus S_2$, S_1 **和** S_2 **维数大于 1,则投影变换不是可逆变换,即不是一一对应的变换.对** P_1 **而言,**

$$P_1\alpha = P_1(\alpha_1 + \alpha_2) = \alpha_1 = P_1(\alpha_1 + k\alpha_2),$$

其中 k 为任意数.这说明对任意实数 k, $\alpha_1 + k\alpha_2$ 的象相同,因而 P_1 就不是一一对应的变换.

例 3　设 σ 为 3 维欧氏几何空间 $\mathbf{R}^3$ 的一个可逆线性变换,证明:σ 将 $\mathbf{R}^3$ 的平面变为平面.

证　任给 $\mathbf{R}^3$ 中一个向量 α,其坐标为 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$,即 $\alpha = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$.设 σ 为 $\mathbf{R}^3$ 的一个线性变换,σ 在基下的矩阵表示为 $\boldsymbol{H}$,且将 $\boldsymbol{x}$ 变换为 $\boldsymbol{y}$.则 $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}$,其中 $\boldsymbol{H}$ 为三阶可逆矩阵.设 $\mathbf{R}^3$ 中平面 π 的方程为

$$Ax_1 + Bx_2 + Cx_3 + D = 0,$$

也即 $(A, B, C)\boldsymbol{x} + D = 0$,又因 $\boldsymbol{x} = \boldsymbol{H}^{-1}\boldsymbol{y}$,代入上面方程有

$$(A, B, C)\boldsymbol{H}^{-1}\boldsymbol{y} + D = 0,$$

令

$$(A, B, C)\boldsymbol{H}^{-1} = (A_1, B_1, C_1),$$

即变换后的点 $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}}$ 满足平面方程:

$$A_1y_1 + B_1y_2 + C_1y_3 + D = 0.$$

由例 3 可知,在 3 维欧氏几何空间 $\mathbf{R}^3$ 中,一个可逆线性变换将平面变换为平面,从而把作为两个平面交线的直线变为直线.

这章主要讲述了 n 维线性空间中线性变换定义与性质,是以线性变换与矩阵之间的关系为线索展开的.n 维线性空间中线性变换是该空间向量之间的对应关系,线性变换研究的是该空间的"几何"问题.当 n 维线性空间的基确定之后,一个线性变换与一个 n 阶方阵对应起来;线性变换的和运算对应矩阵的和运算,线性变换的积对应矩阵的积;在同一空间两个不同基下的一个线性变换的两个矩阵表示,$\boldsymbol{A}$, $\boldsymbol{B}$ 是相似的(即存在一个可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{A} = \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{C}$).一个线性变换的诸多性质之所以能用矩阵语言等价地表达是因为在 n 维线性空间中已建立了基的概念.在 n 维线性空间内引入基的概念,使得笛卡儿坐标方法在 n 维线性空间得以体现.线性变换可用矩阵研究;反过来,有很多矩阵问题也会从 n 维线性空间几何得到启发,找到"直观"的处理方法.

J.Dieudonne(1906—1992,法国数学家)在谈到代数与几何关系时说:"这种由语言相似性指引的从通常的'几何直观'到新的'抽象直观'的转移已被证明具有相当大的功效,不去使用会是很可惜的.用几何语言代替代数语言几乎能做到相当的简化,并使推理在一大堆错综复杂计算中未被察觉的性质显现出来",接着又说:"经典线性代数处理数组和矩阵,而在几何语言中它们成为向量和线性映射.然而现在这种语言已扩张到这种程度,其中代替数的'标量'是任何域甚至任何环的元素,它具有使得论证比较简短的功效,以至如今当我们面临例如线性

代数的最新推广——同调代数时,甚至不能看出没有这种‘空间’语言怎么能进行下去."(J.Dieudonne.当代数学——为了人类心智的荣耀.沈永欢译.上海:上海教育出版社,1999).

几何方法与代数方法的融合是数学自身的需要和数学统一性的体现,也是处理实际工程问题的有力手段.

习 题 五

习题 A

1. 在 3 元列向量空间 $\mathbf{R}^3$ 中,设 $\alpha=(x_1,x_2,x_3)^{\mathrm{T}}\in\mathbf{R}^3$,判断下列变换哪些是线性变换,其中哪些又是满秩线性变换:

(1) $\sigma\alpha=\alpha+\alpha_0$,其中 $\alpha_0\in\mathbf{R}^3$ 是固定向量;

(2) $\sigma\alpha=(x_1,x_1+x_2,x_1+x_2+x_3)^{\mathrm{T}}$;

(3) $\sigma\alpha=(x_1+x_2+x_3,x_1-x_2+x_3,x_1+x_3)^{\mathrm{T}}$;

(4) $\sigma\alpha=(0,x_1,x_2)^{\mathrm{T}}$;

(5) $\sigma\alpha=(x_1^2,x_2^2,x_3^2)^{\mathrm{T}}$.

2. 设 $M(R)_{n\times n}$为实数域 $\mathbf{R}$ 上全体 n 阶方阵构成的线性空间,给定 $\boldsymbol{C}\in M(R)_{n\times n}$,对任意 $\boldsymbol{X}\in M(R)_{n\times n}$,变换 $T(\boldsymbol{X})=\boldsymbol{CX}-\boldsymbol{XC}$.证明:$T$ 是 $M(R)_{n\times n}$的线性变换,且对任意 $\boldsymbol{A},\boldsymbol{B}\in M(R)_{n\times n}$,$T(\boldsymbol{AB})=(T\boldsymbol{A})\boldsymbol{B}+\boldsymbol{A}T(\boldsymbol{B})$.

3. 给定 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}\in M(R)_{n\times n}$,对任意 $\boldsymbol{X}\in M(R)_{n\times n}$,

(1) 定义 $T(\boldsymbol{X})=\boldsymbol{AXB}+\boldsymbol{CX}+\boldsymbol{XD}$,证明:$T$ 是 $M(R)_{n\times n}$的线性变换;

(2) 定义 $T(\boldsymbol{X})=\boldsymbol{AXB}$,证明:$T$ 是满秩线性变换当且仅当 $\boldsymbol{A},\boldsymbol{B}$ 可逆.

4. 求下列各线性变换对相应的欧氏几何空间的自然基的矩阵表示:

(1) $\sigma:R^2\to R^2$,其中 $\sigma((x_1,x_2))=(3x_1-x_2,2x_1+4x_2)$;

(2) $\sigma:R^3\to R^3$,其中 $\sigma((x_1,x_2,x_3))=(2x_1+3x_2-8x_3,x_1+x_2+x_3,4x_1-5x_3)$.

5. 设 σ 是 3 元列向量空间 $\mathbf{R}^3$ 的线性变换,设 $\boldsymbol{e}_1=(-1,0,2)^{\mathrm{T}}$,$\boldsymbol{e}_2=(0,1,1)^{\mathrm{T}}$,$\boldsymbol{e}_3=(3,-1,0)^{\mathrm{T}}$,且 $\sigma(\boldsymbol{e}_1)=(-5,0,3)^{\mathrm{T}}$,$\sigma(\boldsymbol{e}_2)=(0,-1,6)^{\mathrm{T}}$,$\sigma(\boldsymbol{e}_3)=(-5,-1,9)^{\mathrm{T}}$.

(1) 求 σ 在基底$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$下的矩阵;

(2) 求 σ 在自然基 $\varepsilon_1=(1,0,0)^{\mathrm{T}}$,$\varepsilon_2=(0,1,0)^{\mathrm{T}}$,$\varepsilon_3=(0,0,1)^{\mathrm{T}}$ 下的矩阵.

6. 在 3 元列向量空间 $\mathbf{R}^3$ 中有两组基:

$$\varepsilon_1=(1,0,1)^T,\ \varepsilon_2=(2,1,0)^T,\ \varepsilon_3=(1,1,1)^T;$$

$$\eta_1=(1,2,-1)^T,\ \eta_2=(2,2,-1)^T,\ \eta_3=(2,-1,-1)^T.$$

设 σ 为 $\mathbf{R}^3$ 中线性变换且 $\sigma(\varepsilon_i)=\eta_i(i=1,2,3)$，写出

(1) 由基$\{\varepsilon_1,\varepsilon_2,\varepsilon_3\}$到$\{\eta_1,\eta_2,\eta_3\}$的过渡矩阵;

(2) 写出 σ 在基$\{\varepsilon_1,\varepsilon_2,\varepsilon_3\}$下的矩阵表示.

习题 B

1. 设 T 为线性空间 V 的线性变换. $T^2=T$, $\mathrm{Im}\ T=U$, $\mathrm{Ker}\ T=W$. 证明:

(1) 若 $\boldsymbol{v}\in U$, 则 $T(\boldsymbol{v})=\boldsymbol{v}$;

(2) 若 T 不是恒等变换, 则 T 是降秩变换;

(3) $V=U\oplus W$, 即 V 是 U 与 W 的直和;

(4) $\mathrm{Ker}\ T=\{\boldsymbol{v}-T\boldsymbol{v}\mid \boldsymbol{v}\in V\}$.

2. 设 V_n 是实数域 $\mathbf{R}$ 上的 n 维线性空间, $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$为 V_n 的一个基. 令 T_{ij}为 V_n 上的线性变换: $T_{ij}\alpha_i=\alpha_j$, $T_{ij}\alpha_k=\mathbf{0}(k\neq i,\ i,j=1,2,\cdots,n)$. 证明: T_{ij}为 $\mathrm{Hom}(V_n,V_n)$的基.

3. 设 T 为 n 维欧氏空间 V_n 上的一个线性变换, 若对 V_n 中任意两个向量 $\boldsymbol{a}$、$\boldsymbol{b}$ 都有$(T\boldsymbol{a},T\boldsymbol{b})=(\boldsymbol{a},\boldsymbol{b})$, 则称 T 为保持内积不变的线性变换或正交变换. 证明: 下面三个条件是 T 为正交变换的充要条件:

(1) T 保持长度不变, 即任给 $\boldsymbol{a}\in V_n$, 有$|T\boldsymbol{a}|=\boldsymbol{a}$;

(2) 如果$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$为 V_n 中的标准正交基, 那么$\{T\alpha_1,T\alpha_2,\cdots,T\alpha_n\}$也为 V_n 中的标准正交基;

(3) $\boldsymbol{A}$ 是 T 在任意标准正交基下的矩阵, 则 $\boldsymbol{A}$ 是正交矩阵.

4. 设 σ 为 n 维线性空间 V_n 的线性变换, 如果 W 是 V_n 的子空间, 且 W 的所有向量在 σ 下的象仍在 W 中, 称 W 为 σ 的一个不变子空间. 证明: 若 σ 是线性空间 V_n 的一个正交变换, 则 σ 的不变子空间的正交补也是 σ 的不变子空间.

5. 设 σ 是三元列向量空间 $\mathbf{R}^3$ 的一个线性变换. 对任意 $\boldsymbol{x}=(x_1,x_2,x_3)^T\in\mathbf{R}^3$, 定义 $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$, 其中

$$(1)\ \boldsymbol{A}=\begin{pmatrix}1&-1&3\\5&6&-4\\7&4&2\end{pmatrix};\ (2)\ \boldsymbol{A}=\begin{pmatrix}2&0&-1\\4&0&-2\\6&0&-3\end{pmatrix};\ (3)\ \boldsymbol{A}=\begin{pmatrix}1&2&0\\0&1&2\\0&3&2\end{pmatrix}.$$

求 $\mathrm{Im}(T)$, $\mathrm{Ker}(T)$的维数与基.

6. 设 T 为 3 维欧氏几何空间 $\mathbf{R}^3$ 的一个线性变换. 对任意 $\boldsymbol{x}=(x_1,x_2,x_3)^T\in\mathbf{R}^3$, 定义 $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$, 其中

$$A=\begin{pmatrix}1 & 3 & 4\\ 3 & 4 & 7\\ -2 & 2 & 0\end{pmatrix}.$$

证明: $\mathrm{Im}(T)$, $\mathrm{Ker}(T)$**分别为过原点的平面和直线,并写出它们的方程**.

7. **设** T **为** 4 **元列向量空间** $\mathbf{R}^4$ **的一个线性变换**.**对任意** $\boldsymbol{x}=(x_1, x_2, x_3, x_4)^{\mathrm{T}}\in\mathbf{R}^4$,**定义** $T(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}$,**其中**

$$\boldsymbol{A}=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 3 & 4\\ 1 & 3 & 6 & 10\\ 1 & 4 & 10 & 19\end{pmatrix},\ \boldsymbol{b}=\begin{pmatrix}1\\ 5\\ 15\\ 31\end{pmatrix}.$$

求 $\mathrm{Im}(T)$, $\mathrm{Ker}(T)$**的维数与基,且求** $\boldsymbol{b}$ **的原象**.

8. **证明**: n **维欧氏空间** V_n **中正交变换之积仍为正交变换**.

第六章　特征值　特征向量　矩阵对角化

在理论与应用中,将一个矩阵化为最简单的矩阵有着重要的意义.由第五章知道,n 维线性空间 V_n 的一个线性变换可用矩阵来表示,线性变换在两个基下的矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 满足关系 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}$,这时称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是相似的.将一个线性变换用最简单矩阵表示,就是将一个矩阵化为与其相似的最简单矩阵问题.事实上,一个矩阵可化为与其相似的对角形矩阵或分块对角形矩阵(Jordan 标准形,参考附录 A.3).欲解决这个问题,首先须讨论矩阵的特征值与特征向量.本章以讨论这个问题开始,继而讨论矩阵可化为对角形矩阵的充要条件及方法.第四章讨论了线性方程组的解法,如果用矩阵语言表述,就是使用初等变换方法将线性方程组的系数矩阵化为行阶梯形矩阵;对于系数矩阵有逆的情形,甚至可以将矩阵化为对角形.应该注意的是,这章及第四章都是讨论将一个矩阵化为对角形矩阵的问题,但要求并不一样.这章的兴趣是,在将一个矩阵化为对角形矩阵过程中,要求其特征值不变;而第四章是要求矩阵行向量构成的线性空间——行向量空间不变.因此,这章不再使用初等变换的方法,而是从矩阵特征值与特征向量入手找到矩阵对角化的方法.

第一节　矩阵的特征值与特征向量

1.1　特征值与特征向量

本章讨论的矩阵均为方阵.给定 $\boldsymbol{A}=(a_{ij})_{n\times n}$,一个重要问题是在 n 维线性空间 V_n 中能否存在一个非零向量 $\boldsymbol{x}$,使 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$,其中 λ 为一个数.这类问题存在于很多应用科学之中,也存在于数学诸如几何、微分方程等很多分支之中.这节讨论一般矩阵 $\boldsymbol{A}$ 的特征值与特征向量问题,假设 $\boldsymbol{A}$ 为复数域 $\mathbf{C}$(见附录 A.1)上的 n 阶方阵,即假设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 中的元为复数.复数域 $\mathbf{C}$ 上的矩阵运算及方阵的行列式性质与实数域 $\mathbf{R}$ 上的矩阵运算及方阵的行列式性质相同.同样在第二章向量空间的定义中,把实数域 $\mathbf{R}$ 换成复数域 $\mathbf{C}$ 就是复数域 $\mathbf{C}$ 上的向量空间的定义.有关实数域 $\mathbf{R}$ 上的向量空间的定理、命题和推论对于复数域 $\mathbf{C}$ 上的

向量空间也成立.因此第二章,第三章关于向量空间,矩阵运算和行列式的定理、命题和推论都可直接应用于本章.本章所讨论的矩阵和向量如无特别声明都是指复数域 **C** 上的矩阵和向量.

定义 6.1　设 $\boldsymbol{A}$ 为复数域 **C** 上的 n 阶方阵,如果存在 $\lambda\in\mathbf{C}$ 和非零的 n 维向量 $\boldsymbol{x}$,使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \text{ 或 } (\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}, \tag{6.1}$$

就称 λ 为矩阵 $\boldsymbol{A}$ 的特征值,称 $\boldsymbol{x}$ 为矩阵 $\boldsymbol{A}$ 对应于特征值 λ 的特征向量.

从定义 6.1 可知,如果 $\boldsymbol{x}$ 是矩阵 $\boldsymbol{A}$ 对应于特征值 λ 的特征向量,对任意的非零数 k, $k\boldsymbol{x}$ 也是 $\boldsymbol{A}$ 对应于 λ 的特征向量.

由于齐次线性方程组(6.1)有非零解的充要条件是系数行列式

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0,$$

其中 $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$是 λ 的 n 次多项式,下面常记为 $|\lambda\boldsymbol{I} - \boldsymbol{A}|$.

定义 6.2　设 $\boldsymbol{A} = (a_{ij})_{n\times n}$为复数域 **C** 上的矩阵.称

$$f(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A}) \tag{6.2}$$

为矩阵 $\boldsymbol{A}$ 的特征多项式, $\lambda\boldsymbol{I} - \boldsymbol{A}$ 为 $\boldsymbol{A}$ 的特征矩阵, $f(\lambda) = 0$ 称为 $\boldsymbol{A}$ 的特征方程.

显然,矩阵 $\boldsymbol{A}$ 的特征方程的根即矩阵 $\boldsymbol{A}$ 的特征值,它的 k 重根称为 $\boldsymbol{A}$ 的 $\boldsymbol{k}$ 重特征值.

因为 $\boldsymbol{A}$ 是 n 阶方阵,所以 $\boldsymbol{A}$ 的特征方程是关于 λ 的一元 n 次方程.代数的基本定理可表述如下:系数为复数的一元 n 次方程在复数域 **C** 中有 n 个根;并且如果实系数的一元 n 次方程有复根 α,则其共轭复数 $\bar{\alpha}$ 也是该方程的根.由此可知.复数域上的 n 阶方阵一定有 n 个特征值;对实数域上的 n 阶方阵一定有 n 个复数特征值,如果复数 α 是其特征值,则其共轭复数 $\bar{\alpha}$ 也是它的特征值.

求矩阵 $\boldsymbol{A}$ 的特征值与特征向量并非是一件很容易的事情.对 $\boldsymbol{A} = (a_{ij})_{5\times 5}$情形, $\boldsymbol{A}$ 的特征方程 $f(\lambda) = 0$ 就是关于 λ 的 5 次方程.E.Galois(1811—1832)证明了高于四次的一般的代数方程的根已不能用其系数的有限次四则运算与开方根组成的公式来表示.因而,对高阶矩阵而言,求它的特征方程的根是困难的.一般说来,求解矩阵特征值与特征向量问题,比解线性方程组要困难得多.在实际应用中,矩阵的特征值与特征向量有广泛的应用,因而在数值计算中发展了不少行之有效的方法,例如 QR 方法, Jacobi 方法等.由于篇幅有限,本章并不讨论这些方法.对后面讨论的例题及习题来说一般,用简单的因式分解,就能求出矩阵的特征值.

例 1　试求矩阵

$$\boldsymbol{A} = \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix} \tag{6.3}$$

的特征值与特征向量 .

解　$\boldsymbol{A}$ 的特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 5 & 3 \\ 3 & \lambda - 5 \end{vmatrix} = (\lambda - 5)^2 - 9 = (\lambda - 2)(\lambda - 8),$$

因而 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = 2, \lambda_2 = 8$.对 $\lambda_1 = 2$,相应的特征向量满足的齐次线性方程组为

$$\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

其基础解系 $\boldsymbol{x}_1 = (1, 1)^{\mathrm{T}}$,对应 $\lambda_1 = 2$ 的任意特征向量为 $k_1 \boldsymbol{x}_1 (k_1 \neq 0)$.

当 $\lambda_2 = 8$,相应的特征向量满足的齐次线性方程组为

$$\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

其基础解系 $\boldsymbol{x}_2 = (-1, 1)^{\mathrm{T}}$,对应于 $\lambda_2 = 8$ 的任意特征向量为 $k_2 \boldsymbol{x}_2 (k_2 \neq 0)$.

注:矩阵的特征值与特征向量常具有明显的几何意义或物理意义 .

在例 1 中,矩阵 $\boldsymbol{A}$ 对应一个二次齐次多项式(或二次型)

$$f(x_1, x_2) = 5x_1^2 - 6x_1 x_2 + 5x_2^2,$$

因为可将 $f(x_1, x_2)$写为

$$f(x_1, x_2) = 5x_1^2 - 6x_1 x_2 + 5x_2^2 = (x_1, x_2) \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x},$$

其中 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$,上式中的实对称矩阵 $\boldsymbol{A}$ 与例 1 中的矩阵 $\boldsymbol{A}$ 是相同的 .如果考察方程 $5x_1^2 - 6x_1 x_2 + 5x_2^2 = 1$,在直角坐标系中画出这条曲线,它是中心在原点的椭圆(图 1) .这个椭圆长半轴长为 $a = \frac{1}{\lambda_1} = \frac{1}{2}$,与例 1 对比,矩阵 $\boldsymbol{A}$ 的对应于 λ_1 的单位特征向量

$$\boldsymbol{\varepsilon}_1 = \frac{\boldsymbol{x}_1}{2} = \left(\frac{1}{2}, \frac{1}{2}\right)^{\mathrm{T}}$$

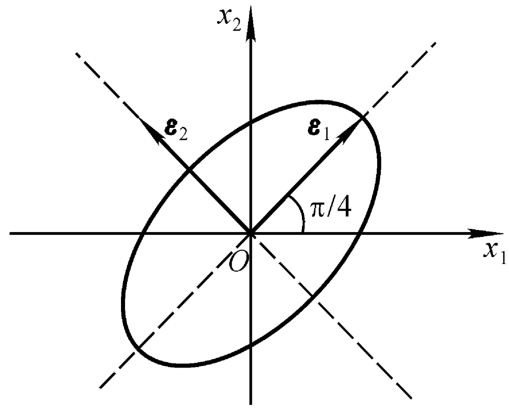


图 1

在椭圆长轴上;椭圆短半轴长为 $b = \frac{1}{\sqrt{\lambda_2}} = \frac{1}{2\sqrt{2}}$,矩阵 $\boldsymbol{A}$ 的对应于 λ_2 的单位特征向量

$$\boldsymbol{\varepsilon}_2 = \frac{\boldsymbol{x}_2}{\sqrt{2}} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^{\mathrm{T}}$$

在椭圆短轴上.可以使用坐标变换,使新的直角坐标系的坐标向量为 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2$.在新坐标系下椭圆的方程为标准形式.由此可见,实对称矩阵的特征值与特征向量常具有明显的几何意义.为把二次曲线化成标准形,关键是找到相应的二次型对应的矩阵的特征值及相应的特征向量.

例 2　试求矩阵

$$\boldsymbol{A} = \begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$$

的特征值与特征向量.

解　$\boldsymbol{A}$ 的特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 4 & -6 & 0 \\ 3 & \lambda + 5 & 0 \\ 3 & 6 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2 (\lambda + 2),$$

所以,$\boldsymbol{A}$ 有特征根 $\lambda_1 = -2, \lambda_2 = \lambda_3 = 1$.对于 $\lambda_1 = -2$,其特征向量满足齐次线性方程组

$$\begin{pmatrix} -6 & -6 & 0 \\ 3 & 3 & 0 \\ 3 & 6 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

有基础解系 $\boldsymbol{x}_1 = (-1, 1, 1)^{\mathrm{T}}$.因此,$\lambda_1 = -2$ 的任意特征向量为 $k_1 \boldsymbol{x}_1\ (k_1 \neq 0)$.

对于 $\lambda_2 = \lambda_3 = 1$,特征向量满足

$$\begin{pmatrix} -3 & -6 & 0 \\ 3 & 6 & 0 \\ 3 & 6 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

而其基础解系 $\boldsymbol{x}_2 = (-2, 1, 0)^{\mathrm{T}}, \boldsymbol{x}_3 = (0, 0, 1)^{\mathrm{T}}$.由齐次线性方程组的理论知道,对应 $\lambda_2 = \lambda_3 = 1$ 的任意特征向量为 $k_2 \boldsymbol{x}_2 + k_3 \boldsymbol{x}_3\ (k_2, k_3$ 不同时为 $0)$.这个例子中,$\lambda = 1$ 是 $\boldsymbol{A}$ 的特征方程的二重特征值,它对应二个线性无关的特征向量.

例 3　试求矩阵

$$\boldsymbol{A} = \begin{pmatrix} -1 & 1 \\ -9 & 5 \end{pmatrix}$$

的特征值与特征向量.

解　$\boldsymbol{A}$ 的特征多项式为

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda+1 & -1\\ 9 & \lambda-5\end{vmatrix}=(\lambda-2)^2,$$

$\boldsymbol{A}$ 有特征值 $\lambda_1=\lambda_2=2$.对于 $\lambda_1=\lambda_2=2$,特征向量满足齐次线性方程组

$$\begin{pmatrix}3 & -1\\ 9 & -3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix},$$

其基础解系为 $\boldsymbol{x}_1=(1,3)^{\mathrm{T}}$.因此,对应 $\lambda_1=\lambda_2=2$ 的任意特征向量为 $k_1\boldsymbol{x}_1(k_1\neq 0)$.应该注意的是,对于 $\boldsymbol{A}$ 的二重特征值 2,它只对应一个线性无关的特征向量.

下面讨论特征多项式 $f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的系数与矩阵 $\boldsymbol{A}$ 的子式之间的关系.

根据行列式性质,可将

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a_{11} & 0-a_{12} & \cdots & 0-a_{1n}\\ 0-a_{21} & \lambda-a_{22} & \cdots & 0-a_{2n}\\ \cdots & \cdots & & \cdots\\ 0-a_{n1} & 0-a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix},$$

分解为 2^n 个 n 阶行列式的和.例如,当 $n=3$ 时,$f(\lambda)$可分解 $2^3=8$ 个 3 阶行列式之和:

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a_{11} & 0-a_{12} & 0-a_{13}\\ 0-a_{21} & \lambda-a_{22} & 0-a_{23}\\ 0-a_{31} & 0-a_{32} & \lambda-a_{33}\end{vmatrix}$$

$$=\begin{vmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{vmatrix}+\begin{vmatrix}-a_{11} & 0 & 0\\ -a_{21} & \lambda & 0\\ -a_{31} & 0 & \lambda\end{vmatrix}+\begin{vmatrix}\lambda & -a_{12} & 0\\ 0 & -a_{22} & 0\\ 0 & -a_{32} & \lambda\end{vmatrix}+$$

$$\begin{vmatrix}\lambda & 0 & -a_{13}\\ 0 & \lambda & -a_{23}\\ 0 & 0 & -a_{33}\end{vmatrix}+\begin{vmatrix}-a_{11} & -a_{12} & 0\\ -a_{21} & -a_{22} & 0\\ -a_{31} & -a_{32} & \lambda\end{vmatrix}+\begin{vmatrix}-a_{11} & 0 & -a_{13}\\ -a_{21} & \lambda & -a_{23}\\ -a_{31} & 0 & -a_{33}\end{vmatrix}+$$

$$\begin{vmatrix}\lambda & -a_{12} & -a_{13}\\ 0 & -a_{22} & -a_{23}\\ 0 & -a_{32} & -a_{33}\end{vmatrix}+\begin{vmatrix}-a_{11} & -a_{12} & -a_{13}\\ -a_{21} & -a_{22} & -a_{23}\\ -a_{31} & -a_{32} & -a_{33}\end{vmatrix}$$

$$=\lambda^3-(a_{11}+a_{22}+a_{33})\lambda^2+\Bigg(\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}+$$

$$\left.\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}\right)\lambda - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$$

$\boldsymbol{A}$ 中行的序号与列的序号均相同的子式称为主子式.由上可见, $f(\lambda)$中 λ^3 的系数为 1, $\lambda^r(0\leq r\leq 2)$的系数 a_{3-r}等于 $\boldsymbol{A}$ 的所有 $3-r$ 阶主子式的和与$(-1)^{3-r}$的乘积.类似上面这种特殊情形,这一结论对于 n 阶方阵也成立.

定理 6.1　设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, $\boldsymbol{A}$ 的特征多项式 $f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$中 λ^n 的系数为 1.而 $\lambda^r(0\leq r\leq n-1)$的系数是 $\boldsymbol{A}$ 中一切 $n-r$ 阶主子式之和与$(-1)^{n-r}$的乘积.

证明从略.

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$,有 n 个特征值$\lambda_1,\lambda_2,\cdots,\lambda_n$,则

$$\begin{aligned} f(\lambda) &= |\lambda\boldsymbol{I}-\boldsymbol{A}| = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) \\ &= \lambda^n-(\lambda_1+\lambda_2+\cdots+\lambda_n)\lambda^{n-1}+\cdots+(-1)^n\lambda_1,\lambda_2,\cdots,\lambda_n. \end{aligned}$$

根据定理 6.1 知:(1) $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$,即 $\boldsymbol{A}$ 的特征值之和等于 $\boldsymbol{A}$ 的主对角元素之和,后者称为矩阵 $\boldsymbol{A}$ 的迹,记为 $\mathrm{tr}\boldsymbol{A}$;(2) $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$.

定理 6.2　若 λ为矩阵 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$是 $\boldsymbol{A}$ 的对应于λ的特征向量,则

(1) $k\lambda$是 $k\boldsymbol{A}$ 的特征值,其中 k 为任意常数;

(2) λ^m 是 $\boldsymbol{A}^m$ 的特征值,其中 m 为任意正整数;

(3) 当 $\boldsymbol{A}$ 可逆时, λ^{-1}是 $\boldsymbol{A}^{-1}$的特征值.

而且 $\boldsymbol{x}$还是 $k\boldsymbol{A}$, $\boldsymbol{A}^m$ 及 $\boldsymbol{A}^{-1}$的对应于特征值 $k\lambda$, λ^m 及λ^{-1}的特征向量.

证　只证明(3).由 $\boldsymbol{A}$ 可逆可知$|\boldsymbol{A}|\neq 0$.因为 $\boldsymbol{A}$ 的 n 个特征值之积$\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$,所以 $\boldsymbol{A}$ 的所有特征值不为零,即 $\lambda\neq 0$.由 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$,可得到

$$\boldsymbol{A}^{-1}\boldsymbol{x}=\boldsymbol{A}^{-1}\frac{1}{\lambda}(\boldsymbol{A}\boldsymbol{x})=\frac{1}{\lambda}\boldsymbol{A}^{-1}(\boldsymbol{A}\boldsymbol{x})=\frac{1}{\lambda}\boldsymbol{x}.$$

因此, $\dfrac{1}{\lambda}$为 $\boldsymbol{A}^{-1}$的特征值,且 $\boldsymbol{x}$ 为 A^{-1}对应于特征值$\dfrac{1}{\lambda}$的特征向量.□

定理 6.3　若 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 是 $\boldsymbol{A}$ 的对应于λ的特征向量, k_1, k_2 为任意常数,只要 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2\neq\boldsymbol{0}$,则它也是 $\boldsymbol{A}$ 的对应于λ的特征向量.

证　因为 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 是 $\boldsymbol{A}$ 对应于λ的特征向量,可知$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_1=\boldsymbol{0}$, $(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_2=\boldsymbol{0}$,于是$(\lambda\boldsymbol{I}-\boldsymbol{A})(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2)=\boldsymbol{0}$.由于$(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2)\neq\boldsymbol{0}$,可知 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2$ 为 $\boldsymbol{A}$ 对应于λ的特征向量.□

例 4　证明:若可逆矩阵 $\boldsymbol{A}$ 的全部特征值是$\lambda_1,\lambda_2,\cdots,\lambda_n$,则 $\boldsymbol{A}^{-1}$的全部特征值是 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$.

证　由假设知

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

由 $\boldsymbol{A}$ 是可逆矩阵及定理 6.1 知 $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|\neq 0$,所以 $\lambda_i\neq 0(i=1,2,\cdots,n)$.于是

$$\begin{aligned}|\lambda \boldsymbol{I}-\boldsymbol{A}^{-1}| &= |-\lambda\boldsymbol{A}^{-1}(\lambda^{-1}\boldsymbol{I}-\boldsymbol{A})| = |-\lambda\boldsymbol{A}^{-1}||\lambda^{-1}\boldsymbol{I}-\boldsymbol{A}|\\ &=(-1)^n\lambda^n|\boldsymbol{A}|^{-1}|\lambda^{-1}\boldsymbol{I}-\boldsymbol{A}|\\ &=(-1)^n\lambda^n\frac{1}{\lambda_1\lambda_2\cdots\lambda_n}\left(\frac{1}{\lambda}-\lambda_1\right)\left(\frac{1}{\lambda}-\lambda_2\right)\cdots\left(\frac{1}{\lambda}-\lambda_n\right)\\ &=\left(\lambda-\frac{1}{\lambda_1}\right)\left(\lambda-\frac{1}{\lambda_2}\right)\cdots\left(\lambda-\frac{1}{\lambda_n}\right).\end{aligned}$$

所以 $\boldsymbol{A}^{-1}$的全部特征值是 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$.□

在$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解空间中,除零向量外的全体解向量就是矩阵 $\boldsymbol{A}$ 对应于特征值λ的全体特征向量.因此把$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解空间称为矩阵 $\boldsymbol{A}$ 关于特征值λ的特征子空间.

注:当 $\boldsymbol{A}$ 为 n 阶实矩阵且特征值 λ 为实数时,$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 为实系数方程,必有实的基础解系,设为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$,记

$$S_1=\{\boldsymbol{x}=k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_s\boldsymbol{x}_s\mid k_1,k_2,\cdots,k_s\in\mathbf{C}\},$$

$$S_2=\{\boldsymbol{x}=k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_s\boldsymbol{x}_s\mid k_1,k_2,\cdots,k_s\in\mathbf{R}\},$$

则 S_1 称为矩阵 $\boldsymbol{A}$ 对应特征值λ在 $\mathbf{C}^n$ 中的特征子空间,而 S_2 称为矩阵 $\boldsymbol{A}$ 对应特征值 λ 在 $\mathbf{R}^n$ 中的特征子空间.当在实数域中讨论问题时,特征子空间指的是 S_2.

1.2　相似矩阵

定义 6.3　对于 n 阶矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,若存在可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$,则称 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$,记为 $\boldsymbol{A}\sim\boldsymbol{B}$.

显然,矩阵相似关系是矩阵间的一个等价关系.

从第五章知道,如果 n 维线性空间 V_n 的一个线性变换σ在 V_n 的旧基$\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$及新基$\{\alpha'_1,\alpha'_2,\cdots,\alpha'_n\}$下的矩阵分别为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$,且旧基到新基的过渡矩阵为 $\boldsymbol{C}$,则 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}$,即 $\boldsymbol{A}$ 相似 $\boldsymbol{B}$.下面定理给出相似矩阵的最重要的性质.

定理 6.4　相似矩阵有相同的特征多项式,因而有相同的特征值.

证　设 $\boldsymbol{A}\sim\boldsymbol{B}$,即存在可逆矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$.于是

$$\begin{aligned}|\lambda\boldsymbol{I}-\boldsymbol{B}| &= |\lambda\boldsymbol{I}-\boldsymbol{P}^{-1}\boldsymbol{AP}| = |\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}|\\ &=|\boldsymbol{P}^{-1}||(\lambda\boldsymbol{I}-\boldsymbol{A})||\boldsymbol{P}| = |\boldsymbol{P}^{-1}||\boldsymbol{P}||\lambda\boldsymbol{I}-\boldsymbol{A}| = |\lambda\boldsymbol{I}-\boldsymbol{A}|,\end{aligned}$$

表明 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有相同特征多项式,因而也有相同特征值 .□

例 5　设 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 分别为矩阵 $\boldsymbol{A}$ 对应于 λ_1, λ_2 的特征向量,且 $\lambda_1 \neq \lambda_2$, 则 $\boldsymbol{x}_1 + \boldsymbol{x}_2$ 不可能是 $\boldsymbol{A}$ 的一个特征向量 .

证　用反证法 .设 $\boldsymbol{x}_1 + \boldsymbol{x}_2$ 是 $\boldsymbol{A}$ 对应 λ 的特征向量,则 $\boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2)$,又因 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 分别为 $\boldsymbol{A}$ 对应于 λ_1, λ_2 的特征向量,所以 $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$, $\boldsymbol{A}\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2$.于是

$$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 = \boldsymbol{A}\boldsymbol{x}_1 + \boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2) .$$

即 $(\lambda_1 - \lambda)\boldsymbol{x}_1 + (\lambda_2 - \lambda)\boldsymbol{x}_2 = \boldsymbol{0}$.因为 $(\lambda_1 - \lambda) \neq (\lambda_2 - \lambda)$, 所以 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 线性相关 .不妨 $\boldsymbol{x}_1 = k\boldsymbol{x}_2$.有 $\lambda_1 \boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{A}k\boldsymbol{x}_2 = k\boldsymbol{A}\boldsymbol{x}_2 = \lambda_2 k\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_1$, 即 $(\lambda_1 - \lambda_2)\boldsymbol{x}_1 = \boldsymbol{0}$, 因此 $\lambda_1 = \lambda_2$, 这与 $\lambda_1 \neq \lambda_2$ 矛盾 .所以 $\boldsymbol{x}_1 + \boldsymbol{x}_2$ 不是 $\boldsymbol{A}$ 的特征向量 .

例 6　设 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ 为关于 x 的 n 次多项式,定义

$$f(\boldsymbol{A}) = a_n \boldsymbol{A}^n + a_{n-1} \boldsymbol{A}^{n-1} + \cdots + a_1 \boldsymbol{A} + a_0 \boldsymbol{I},$$

其中 $\boldsymbol{A} = (a_{ij})_{n\times n}$, 而 $\boldsymbol{I}$ 为 n 阶单位矩阵 .若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $f(\boldsymbol{A}) \sim f(\boldsymbol{B})$.

证　因 $\boldsymbol{A} \sim \boldsymbol{B}$, 存在可逆矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$.于是,对任意数 k 及正整数 m:

(1) $k\boldsymbol{B} = k\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{-1}(k\boldsymbol{A})\boldsymbol{P}$.所以 $k\boldsymbol{A} \sim k\boldsymbol{B}$;

(2) $\boldsymbol{B}^m = \boldsymbol{B}\boldsymbol{B}\cdots\boldsymbol{B} = (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\cdots(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}) = \boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}$, 所以 $\boldsymbol{A}^m \sim \boldsymbol{B}^m$;

(3) $f(\boldsymbol{B}) = a_n\boldsymbol{B}^n + a_{n-1}\boldsymbol{B}^{n-1} + \cdots + a_1\boldsymbol{B} + a_0\boldsymbol{I} = \boldsymbol{P}^{-1}(a_n\boldsymbol{A}^n + \cdots + a_1\boldsymbol{A} + a_0\boldsymbol{I})\boldsymbol{P} = \boldsymbol{P}^{-1} f(\boldsymbol{A})\boldsymbol{P}$.

所以 $f(\boldsymbol{A}) \sim f(\boldsymbol{B})$.

第二节　矩阵的对角化

2.1　矩阵可对角化的充要条件

所谓矩阵可对角化是指矩阵与对角形矩阵相似 .这节讨论矩阵与对角形矩阵相似的条件 .首先给出一个基本定理

定理 6.5　n 阶矩阵 $\boldsymbol{A}$ 与对角形矩阵相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 .

证　必要性 .设存在可逆矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \Lambda$, 其中 Λ 为对角形矩阵,即

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

令 $\boldsymbol{P}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$,其中 $\boldsymbol{x}_i(i=1,2,\cdots,n)$为 n 元列向量.因为 $\boldsymbol{P}$ 可逆,所以 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是线性无关的,因而 $\boldsymbol{x}_i\neq\boldsymbol{0}(i=1,2,\cdots,n)$.由 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$ 知,$\boldsymbol{AP}=\boldsymbol{P}\Lambda$,即

$$\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\operatorname{diag}(\lambda_1,\lambda_1,\cdots,\lambda_n).$$

因此,

$$(\boldsymbol{Ax}_1,\boldsymbol{Ax}_2,\cdots,\boldsymbol{Ax}_n)=(\lambda_1\boldsymbol{x}_1,\lambda_1\boldsymbol{x}_2,\cdots,\lambda_n\boldsymbol{x}_n).$$

于是 $\boldsymbol{Ax}_i=\lambda_i\boldsymbol{x}_i(i=1,2,\cdots,n)$.这样,$\boldsymbol{x}_i(i=1,2,\cdots,n)$是 $\boldsymbol{A}$ 对应于特征值 λ_i 的特征向量,且 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 线性无关.

充分性　必要性证明过程中的步骤是可逆的,因此充分性也成立.□

由定理 6.5 知,如果 $\boldsymbol{A}$ 有 n 个线性无关的特征向量,$\boldsymbol{A}$ 一定与对角形矩阵 Λ 相似.那么,$\boldsymbol{A}$ 满足什么条件才会有 n 个线性无关特征向量这是一个较复杂的问题.下面给出一些充分条件.

定理 6.6　设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是矩阵 $\boldsymbol{A}$ 的两两不相等的特征值,

$$B:\boldsymbol{x}_{11},\boldsymbol{x}_{12},\cdots,\boldsymbol{x}_{1l_1};\boldsymbol{x}_{21},\boldsymbol{x}_{22},\cdots,\boldsymbol{x}_{2l_2};\cdots;\boldsymbol{x}_{m1},\boldsymbol{x}_{m2},\cdots,\boldsymbol{x}_{ml_m}$$

是对应 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 的 m 组各自线性无关的特征向量,则 B 是线性无关向量组.

证　对 m 作数学归纳法.当 $m=1$ 时,由定理条件知向量组 $B:\boldsymbol{x}_{11},\boldsymbol{x}_{12},\cdots,\boldsymbol{x}_{1l_1}$是线性无关的.设 $m<r$ 时,结论成立.往证 $m=r$ 时,结论也成立.设实数 $b_{11},b_{12},\cdots,b_{1l_1};b_{21},b_{22},\cdots,b_{2l_2};\cdots;b_{r1},b_{r2},\cdots,b_{rl_r}$使得

$$b_{11}\boldsymbol{x}_{11}+\cdots+b_{1l_1}\boldsymbol{x}_{1l_1}+b_{21}\boldsymbol{x}_{21}+\cdots+b_{2l_2}\boldsymbol{x}_{2l_2}+\cdots+b_{r1}\boldsymbol{x}_{r1}+\cdots+b_{rl_r}\boldsymbol{x}_{rl_r}=\boldsymbol{0},\tag{6.4}$$

用 $\boldsymbol{A}$ 左乘(6.4)两端,因为 $\boldsymbol{Ax}_{ij}=\lambda_i\boldsymbol{x}_{ij}(1\leq j\leq l_i,i=1,2,\cdots,r)$,所以

$$\begin{aligned}&b_{11}\lambda_1\boldsymbol{x}_{11}+\cdots+b_{1l_1}\lambda_1\boldsymbol{x}_{1l_1}+b_{21}\lambda_2\boldsymbol{x}_{21}+\cdots+b_{2l_2}\lambda_2\boldsymbol{x}_{2l_2}\\&\quad+\cdots+b_{r1}\lambda_r\boldsymbol{x}_{r1}+\cdots+b_{rl_r}\lambda_r\boldsymbol{x}_{rl_r}=\boldsymbol{0}.\end{aligned}\tag{6.5}$$

当 $\lambda_1=0$ 时,有

$$b_{21}\lambda_2\boldsymbol{x}_{21}+\cdots+b_{2l_2}\lambda_2\boldsymbol{x}_{2l_2}+\cdots+b_{r1}\lambda_r\boldsymbol{x}_{r1}+\cdots+b_{rl_r}\lambda_r\boldsymbol{x}_{rl_r}=\boldsymbol{0}.$$

由归纳假设有 $\boldsymbol{x}_{21},\boldsymbol{x}_{22},\cdots,\boldsymbol{x}_{2l_2};\cdots;\boldsymbol{x}_{r1},\boldsymbol{x}_{r2},\cdots,\boldsymbol{x}_{rl_r}$ 线性无关.于是,$b_{ij}\lambda_i=0$ $(1\leq j\leq l_i,i=2,\cdots,r)$因为 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 是矩阵 A 的两两不相等的特征值,所以 $\lambda_2,\cdots,\lambda_r$ 都不为零,进而 $b_{ij}=0(1\leq j\leq l_i,i=2,\cdots,r)$;当 $\lambda_1\neq0$ 时,(6.5)减(6.4)乘 λ_1 得

$$b_{21}(\lambda_2-\lambda_1)\boldsymbol{x}_{21}+\cdots+b_{2l_2}(\lambda_2-\lambda_1)\boldsymbol{x}_{2l_2}+\cdots+b_{r1}(\lambda_r-\lambda_1)\boldsymbol{x}_{r1}$$

$$+\cdots+b_{rl_r}(\lambda_r-\lambda_1)\boldsymbol{x}_{rl_r}=\boldsymbol{0},$$

与 $\lambda_1=0$ 时同理,有 $b_{ij}(\lambda_i-\lambda_1)=0,(1\leq j\leq l_i,i=2,\cdots,r)$,进而 $b_{ij}=0(1\leq j\leq l_i,i=2,\cdots,r)$.综上,总有 $b_{ij}=0(1\leq j\leq l_i,i=2,\cdots,r)$,代入(6.4)知 $b_{11}\boldsymbol{x}_{11}+\cdots+b_{1l_1}\boldsymbol{x}_{1l_1}=\boldsymbol{0}$,但 $\boldsymbol{x}_{11},\cdots,\boldsymbol{x}_{1l_1}$ 线性无关,$b_{1j}=0(1\leq j\leq l_1)$,于是所有的 $b_{ij}=0(1\leq j\leq l_i,i=1,2,\cdots,r)$,因此向量组 B 线性无关.□

由定理 6.5 和定理 6.6 知,下面推论成立.

推论 假设 n 阶矩阵 $\boldsymbol{A}$ 有 n 个互不相同的特征值,则 $\boldsymbol{A}$ 与对角形相似.

推论的逆不成立,即 $\boldsymbol{A}$ 有 $k(<n)$ 个不同特征值,$\boldsymbol{A}$ 也可能与对角形相似,见下面的例 1.

从定理 6.5 知,n 阶矩阵 $\boldsymbol{A}$ 与对角形相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关特征向量,那么矩阵 $\boldsymbol{A}$ 具有什么性质,才能具有 n 个线性无关特征向量呢?下面定理回答了这个问题.

定理 6.7 n 阶矩阵 $\boldsymbol{A}$ 与对角形矩阵相似的充要条件是,对 $\boldsymbol{A}$ 的每个 k_i 重特征值 λ_i,特征矩阵 $\lambda_i I-\boldsymbol{A}$ 的秩为 $n-k_i$,即对应 λ_i,$\boldsymbol{A}$ 有 k_i 个线性无关的特征向量.

为证明这个定理,需先证明一个预备定理.

定理 6.8 设 λ_0 是 n 阶方阵 $\boldsymbol{A}$ 的一个 k 重特征值,对应于 λ_0 的线性无关的特征向量最多个数 l 小于等于 k.

证明见附录 B 2

定理 6.7 的证明

充分性.设 $\boldsymbol{A}$ 有 s 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$,它们的重数分别为 $k_1,k_2,\cdots,k_s$ 且分别对应 $k_1,k_2,\cdots,k_s$ 个线性无关特征向量.由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 为 n 次特征多项式的根,在复数域中,它们重数之和 $k_1+k_2+\cdots+k_n=n$.由定理 6.6 知,对应不同特征值的特征向量线性无关,因此 $\boldsymbol{A}$ 共有 $k_1+k_2+\cdots+k_s=n$ 个线性无关的特征向量.由定理 6.5 知,$\boldsymbol{A}$ 与对角形相似.

必要性.用反证法.设有一个特征值 $\lambda_i(i=1,2,\cdots,s)$ 对应的线性无关特征向量最多个数 l_i 不等于 λ_i 的重数 k_i,由定理 6.8,则 $l_i<k_i$,于是 $\boldsymbol{A}$ 的线性无关的特征向量个数小于 n.由定理 6.5 知,A 不能与对角形相似.□

例 1 矩阵

$$\boldsymbol{A}=\begin{pmatrix}3&-1&-2\\2&0&-4\\-1&1&4\end{pmatrix}$$

能否与对角形矩阵相似,若能,试求与其相似的对角形矩阵.

解　$\boldsymbol{A}$ 的特征多项式

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-3 & 1 & 2\\ -2 & \lambda & 4\\ 1 & -1 & \lambda-4\end{vmatrix}=(\lambda-3)(\lambda-2)^2,$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1=2$(二重根),$\lambda_2=3$.

对 $\lambda_1=2$,对应的特征向量满足方程组

$$\begin{pmatrix}-1 & 1 & 2\\ -2 & 2 & 4\\ 1 & -1 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

其基础解系

$$\boldsymbol{p}_1=(1,1,0)^{\mathrm{T}},\quad \boldsymbol{p}_2=(2,0,1)^{\mathrm{T}}.$$

对 $\lambda_2=3$,对应的特征向量满足方程组

$$\begin{pmatrix}0 & 1 & 2\\ -2 & 3 & 4\\ 1 & -1 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix},$$

其基础解系为 $\boldsymbol{p}_3=(1,2,-1)^{\mathrm{T}}$.由定理 6.7 知,$\boldsymbol{A}$ 与对角形相似.令

$$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix}1 & 2 & 1\\ 1 & 0 & 2\\ 0 & 1 & -1\end{pmatrix},$$

则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(2,2,3)$.

例 2　矩阵

$$\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$$

能否与对角阵相似?

解

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{vmatrix}=(\lambda-2)(\lambda-1)^2,$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1=1$(二重根),$\lambda_2=2$.

对重根 $\lambda_1=1$,对应的特征向量满足 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$,由于

$$\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}2 & -1 & 0\\ 4 & -2 & 0\\ -1 & 0 & -1\end{pmatrix}$$

的秩 $R(\boldsymbol{I}-\boldsymbol{A})=2$,故对应二重根 λ_1 只有 1 个线性无关的特征向量.由定理 6.7 知 $\boldsymbol{A}$ 不能与对角阵相似.

例 3　若 n 元向量空间 $\mathbf{R}^n$ 中任何非零向量均为 n 阶方阵 $\boldsymbol{A}$ 的特征向量,则 $\boldsymbol{A}=\lambda\boldsymbol{I}_n$($\lambda\boldsymbol{I}_n$ 称为数量矩阵).

证　因为 $\mathbf{R}^n$ 中任何非零向量均为 n 阶方阵 $\boldsymbol{A}$ 的特征向量,取 $\mathbf{R}^n$ 中自然基

$$\varepsilon_i=(0,\cdots,0,1,0,\cdots 0)\ (i=1,2,\cdots,n),$$

其中“1”位于该向量第 i 个分量上,令其为 $\boldsymbol{A}$ 的特征向量.令 $\boldsymbol{P}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\boldsymbol{I}_n$,则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,但 $\boldsymbol{P}=\boldsymbol{I}_n$,所以

$$\boldsymbol{A}=\boldsymbol{I}^{-1}\boldsymbol{A}\boldsymbol{I}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n),$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 对应于 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的特征值.下面证明,当 $i\neq j$ 时,$\lambda_i=\lambda_j$.用反证法,如果存在 $\lambda_i\neq\lambda_j\ (1\leq i,j\leq n)$.则 λ_i,λ_j 相应的特征向量 $\boldsymbol{x}_i,\boldsymbol{x}_j$ 是线性无关的,因而 $\boldsymbol{x}_i+\boldsymbol{x}_j\neq\boldsymbol{0}$,即 $\boldsymbol{x}_i+\boldsymbol{x}_j$ 是 $\mathbf{R}^n$ 的非零向量,由题设知 $\boldsymbol{x}_i+\boldsymbol{x}_j$ 也是 $\boldsymbol{A}$ 的特征向量.由例 5 知,$\boldsymbol{A}$ 对应不同特征值的特征向量之和不是 $\boldsymbol{A}$ 的特征向量,于是导致矛盾.所以 $\lambda_1=\lambda_2=\cdots=\lambda_n$.因而 $\boldsymbol{A}=\lambda\boldsymbol{I}_n$,即 $\boldsymbol{A}$ 是数量矩阵.

2.2　实对称矩阵的对角化

虽然一般的实矩阵不一定能对角化,但实对称矩阵一定能对角化,这是因为实对称矩阵的特征值与特征向量具有一些特殊的性质.实矩阵的特征多项式是实系数多项式,表面看来它的特征根有可能不是实数,相应的特征向量也有可能不是实向量.然而,实对称矩阵的特征值却都是实数,相应的特征向量也都可取实向量,且不同特征值对应的特征向量正交,自然也是线性无关的.下面来证明这些事实.

定理 6.9　实对称矩阵的特征值是实数.

证　设复数 λ 为实对称矩阵 $\boldsymbol{A}$ 的特征值,复向量 $\boldsymbol{x}$ 为其对应的特征向量,即 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x},\ \boldsymbol{x}\neq\boldsymbol{0}$,用 $\bar{\lambda}$ 表示 λ 的共轭复数,用 $\bar{\boldsymbol{x}}$ 表示 $\boldsymbol{x}$ 的共轭向量.

因为 $\boldsymbol{A}$ 是实对称的,所以 $\bar{\boldsymbol{A}}=\boldsymbol{A}$,有

$$\boldsymbol{A}\bar{\boldsymbol{x}}=\bar{\boldsymbol{A}}\bar{\boldsymbol{x}}=\overline{\boldsymbol{A}\boldsymbol{x}}=\overline{\lambda\boldsymbol{x}}=\bar{\lambda}\bar{\boldsymbol{x}}$$

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\bar{\boldsymbol{x}}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x})=\bar{\boldsymbol{x}}^{\mathrm{T}}\lambda\boldsymbol{x}=\lambda\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}\tag{6.6}$$

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}=(\boldsymbol{A}\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}=(\bar{\lambda}\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}=\bar{\lambda}\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}\tag{6.7}$$

(6.6)式减(6.7),有

$$(\lambda-\bar{\lambda})\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0},$$

由 $\boldsymbol{x}\neq\boldsymbol{0}$,$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}>0$,于是 $\lambda-\bar{\lambda}=0$,即 $\lambda=\bar{\lambda}$,这说明 λ 是实数.□

注:对于复数 $\lambda=a+bi$ 而言,$\bar{\lambda}=a-bi$ 为其共轭复数.设 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 为复矩阵,$\bar{a}_{ij}$ 为 a_{ij} 的共轭复数,称 $\bar{\boldsymbol{A}}=(\bar{a}_{ij})_{m\times n}$ 为 $\boldsymbol{A}$ 的共轭复矩阵,设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为

两个复矩阵,则下列性质成立:(1) $\overline{\boldsymbol{A}+\boldsymbol{B}}=\overline{\boldsymbol{A}}+\overline{\boldsymbol{B}}$;(2) $\overline{\lambda\boldsymbol{A}}=\overline{\lambda}\,\overline{\boldsymbol{A}}$;(3) $\overline{\boldsymbol{AB}}=\overline{\boldsymbol{A}}\,\overline{\boldsymbol{B}}$(当然,要求所列出的运算均有意义).

定理 6.10　在 n 维欧氏几何空间 $\mathbf{R}^n$ 中,实对称矩阵 $\boldsymbol{A}$ 对应于不同特征值的特征向量是正交的.

证　设 λ_1 和 λ_2 为实对称矩阵的两个不同特征值,$\mathbf{R}^n$ 中的向量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 分别是相应 λ_1,λ_2 的特征向量,于是 $\boldsymbol{A}\boldsymbol{x}_1=\lambda_1\boldsymbol{x}_1$, $\boldsymbol{A}\boldsymbol{x}_2=\lambda_2\boldsymbol{x}_2$.因为 A 是对称矩阵,即 $A=A^{\mathrm{T}}$,有

$$\lambda_1\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1=\boldsymbol{x}_2^{\mathrm{T}}\lambda_1\boldsymbol{x}_1=\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_1$$
$$=(\boldsymbol{A}\boldsymbol{x}_2)^{\mathrm{T}}\boldsymbol{x}_1=(\lambda_2\boldsymbol{x}_2)^{\mathrm{T}}\boldsymbol{x}_1=\lambda_2\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1$$

于是$(\lambda_1-\lambda_2)\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1=0$,因为 $\lambda_1\neq\lambda_2$,所以 $\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{x}_1=0$,即$(\boldsymbol{x}_2,\boldsymbol{x}_1)=0$,这表明 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 正交.□

定理 6.11　对任意 n 阶实对称矩阵 $\boldsymbol{A}$,存在 n 阶正交矩阵 $\boldsymbol{P}$,使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 为对角形矩阵.

证　对 n 作数学归纳法.当 $n=1$ 时,显然定理成立.设 $n\leq k$ 时,定理成立;往证 $n=k+1$ 时,定理成立.

由定理 6.9 知 $\boldsymbol{A}$ 的特征值都为实数.设 $\boldsymbol{p}_1$ 为对应于 A 的特征值 λ 的单位特征向量,即 $\boldsymbol{A}\boldsymbol{p}_1=\lambda\boldsymbol{p}_1$.将 $\boldsymbol{p}_1$ 扩充为 $k+1$ 维欧氏几何空间 $\mathbf{R}^{k+1}$($k+1$ 元列向量空间)的一组标准正交基 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_{k+1}$.令 $\boldsymbol{P}_0=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_{k+1})$,显然,$\boldsymbol{P}_0^{\mathrm{T}}\boldsymbol{P}_0=\boldsymbol{I}$,即 $\boldsymbol{P}_0$ 为正交矩阵.

$$\boldsymbol{P}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_0=\begin{pmatrix}\boldsymbol{p}_1^{\mathrm{T}}\\ \boldsymbol{p}_2^{\mathrm{T}}\\ \cdots\\ \boldsymbol{p}_{k+1}^{\mathrm{T}}\end{pmatrix}\boldsymbol{A}(\boldsymbol{p}_1\ \boldsymbol{p}_2\cdots\ \boldsymbol{p}_{k+1})=\begin{pmatrix}\boldsymbol{p}_1^{\mathrm{T}}\\ \boldsymbol{p}_2^{\mathrm{T}}\\ \cdots\\ \boldsymbol{p}_{k+1}^{\mathrm{T}}\end{pmatrix}(\boldsymbol{A}\boldsymbol{p}_1\ \boldsymbol{A}\boldsymbol{p}_2\cdots\ \boldsymbol{A}\boldsymbol{p}_{k+1})$$

$$=\begin{pmatrix}\boldsymbol{p}_1^{\mathrm{T}}\\ \boldsymbol{p}_2^{\mathrm{T}}\\ \cdots\\ \boldsymbol{p}_{k+1}^{\mathrm{T}}\end{pmatrix}(\lambda\boldsymbol{p}_1\ \boldsymbol{A}\boldsymbol{p}_2\cdots\ \boldsymbol{A}\boldsymbol{p}_{k+1})$$

$$=\begin{pmatrix}\lambda & \boldsymbol{p}_1^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{p}_2\cdots\ \boldsymbol{A}\boldsymbol{p}_{k+1})\\ 0 & \boldsymbol{p}_2^{\mathrm{T}}\\ \cdots & \cdots\ \ \boldsymbol{A}(\boldsymbol{p}_2\cdots\ \boldsymbol{p}_{k+1})\\ 0 & \boldsymbol{p}_{k+1}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\lambda & \boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{AB}\\ \boldsymbol{0} & \boldsymbol{B}^{\mathrm{T}}\boldsymbol{AB}\end{pmatrix},$$

其中 $\boldsymbol{B}=(\boldsymbol{p}_2,\cdots,\boldsymbol{p}_{k+1})$.由于 $\boldsymbol{P}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_0$ 为对称矩阵,所以 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{AB}=\boldsymbol{O}$,即

$$\boldsymbol{P}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_0 = \begin{pmatrix} \lambda & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{B} \end{pmatrix}.$$

而 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{B}$ 是 k 阶对称阵.由归纳假设存在 k 阶正交矩阵 $\boldsymbol{Q}$,使得 $\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{B})\boldsymbol{Q} = \Lambda$,其中 Λ 为对角矩阵.由于 $\boldsymbol{P}_0$ 和 $\boldsymbol{Q}$ 都是正交矩阵,所以

$$\boldsymbol{P} = \boldsymbol{P}_0\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{pmatrix},$$

也为正交矩阵.于是

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}^{\mathrm{T}} \end{pmatrix}\boldsymbol{P}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_0\begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q} \end{pmatrix} = \begin{pmatrix} \lambda & \boldsymbol{0} \\ \boldsymbol{0} & \Lambda \end{pmatrix}.\square$$

由定理 6.11 的证明可知,$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$,其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{A}$ 的特征值,$\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$,且 $\boldsymbol{p}_i$ 为 $\boldsymbol{A}$ 对应 λ_i 的特征向量.于是,下面定理成立.

定理 6.12 对于任一个 n 阶实对称矩阵 $\boldsymbol{A}$,存在 n 阶正交矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$,其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{A}$ 的特征根.

由定理 6.12 知,实对称矩阵一定能够对角化.求正交矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角形矩阵过程可归纳如下:

(1) 由实对称矩阵 $\boldsymbol{A}$ 的特征方程,求出 $\boldsymbol{A}$ 的所有特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$,其中 λ_i 的重数为 r_i.

(2) 由定理 6.7 知,r_i 重特征值 λ_i 对应 r_i 个线性无关特征向量,用 Schmidt 正交化方法将它们正交标准化,得到 r_i 个相互正交且单位化的特征向量.由定理 6.10 知,对应不同特征值的特征向量正交,因而,一共得到 $\sum_{i=1}^{m} r_i = n$ 个两两正交且单位化的特征向量(列向量) $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$.

(3) 若 $\boldsymbol{x}_i$ 对应特征值 $\lambda_i'(i = 1, 2, \cdots, n)$,$\lambda_i' \in \{\lambda_1, \lambda_2, \cdots, \lambda_m\}$.(注意:$\lambda_1', \lambda_2', \cdots, \lambda_n'$ 中有的可能相等).令 $\boldsymbol{P} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)$,则

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1', \lambda_2', \cdots, \lambda_n').$$

例 4 矩阵

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

求正交矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角形矩阵.

解 (1) 求特征值.

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

$\boldsymbol{A}$ 的特征值为 $\lambda_1=2,\lambda_2=-1$(二重根).

(2) 对 $\lambda_1=2$,特征向量满足方程组

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

其基础解系为 $\boldsymbol{x}_1=(1,1,1)^{\mathrm{T}}$.将其单位化,得 $\boldsymbol{p}_1=\frac{1}{3}(1,1,1)^{\mathrm{T}}$.

对 $\lambda_2=-1$,特征向量满足方程组

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

其基础解系为 $\boldsymbol{x}_2=(-1,1,0)^{\mathrm{T}},\boldsymbol{x}_3=(-1,0,1)^{\mathrm{T}}$.将其用 Schmidt 正交化公式正交化:

$$\boldsymbol{\xi}_2=\boldsymbol{x}_2=(-1,1,0)^{\mathrm{T}},\ \boldsymbol{\xi}_3=\boldsymbol{x}_3-\frac{(\boldsymbol{x}_3,\boldsymbol{x}_2)}{(\boldsymbol{x}_2,\boldsymbol{x}_2)}\boldsymbol{x}_2=\left(-\frac{1}{2},-\frac{1}{2},1\right)^{\mathrm{T}},$$

分别将它们单位化,有

$$\boldsymbol{p}_2=\frac{1}{2}(-1,1,0)^{\mathrm{T}},\ \boldsymbol{p}_3=\frac{1}{6}(-1,-1,2)^{\mathrm{T}}.$$

(3) 取正交矩阵

$$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix} \frac{1}{3} & -\frac{1}{2} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{2} & -\frac{1}{6} \\ \frac{1}{3} & 0 & \frac{2}{6} \end{pmatrix},$$

则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)=\mathrm{diag}(2,-1,-1)$.

当求出 $\boldsymbol{x}_2$ 之后,也可令 $\boldsymbol{x}_3=(1,1,a)^{\mathrm{T}}$,代入特征向量满足的方程组,得到 $a=-2$,即 $\boldsymbol{x}_3=(1,1,-2)^{\mathrm{T}}$,这时 $\boldsymbol{x}_2$ 与 $\boldsymbol{x}_3$ 是正交的,再将 $\boldsymbol{x}_2,\boldsymbol{x}_3$ 单位化,也可得到另外一组相互正交的单位向量 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$.

例 5　设 n 阶实对称阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是相似矩阵,证明存在正交矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

证　因为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是相似矩阵,因而有相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$,因为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 均为实对称矩阵,所以存在正交矩阵 $\boldsymbol{P}_1$,使 $\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,且存在正交矩阵 $\boldsymbol{P}_2$,使 $\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}_2=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$.这样,$\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1=\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}_2$,

因而 $P_2P_1^{\mathrm{T}}AP_1P_2^{\mathrm{T}}=B$.令 $P=P_1P_2^{\mathrm{T}}$,因为 P 是正交矩阵的乘积,所以 P 仍为正交矩阵,这时,$P^{\mathrm{T}}=(P_1P_2^{\mathrm{T}})^{\mathrm{T}}=P_2P_1^{\mathrm{T}}$,于是 $P^{\mathrm{T}}AP=B$.

习 题 六

习题 A

1.求下列矩阵的特征值与特征向量:

(1) $\begin{pmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 3 & 0 & -5 \\ \frac{1}{5} & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix}$;

(3) $\begin{pmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$; (4) $\begin{pmatrix} 5 & 0 & 1 \\ 1 & 1 & 0 \\ -7 & 1 & 0 \end{pmatrix}$;

2.设 λ 为 n 阶可逆矩阵 A 的特征值,证明:$\lambda^{-1}|A|$是 A^* 的特征值.

3.设 $A^2=I$,证明:A 的特征值只能是±1.

4.设 A 是奇数阶正交矩阵,如果$|A|=1$,试证 1 是 A 的特征值.

5.判断下列矩阵 A 能否与对角形矩阵 Λ 相似;若能,求出 Λ 及可逆矩阵 P,使得 $P^{-1}AP=\Lambda$:

(1) $\begin{pmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{pmatrix}$; (2) $\begin{pmatrix} -1 & 4 & -2 \\ -3 & 4 & 0 \\ -3 & 1 & 3 \end{pmatrix}$;

(3) $\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 5 & -5 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$; (4) $\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$.

6.已知下列矩阵 A,求 A^{100}:

(1) $\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} -1 & 7 & -1 \\ 0 & 1 & 0 \\ 0 & 15 & -2 \end{pmatrix}$.

7.对下列实对称矩阵 A,求出正交矩阵 P,使 $P^{\mathrm{T}}AP=\Lambda$ 为对角形矩阵:

(1) $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$; (2) $\begin{pmatrix} 6 & -2 \\ -2 & 3 \end{pmatrix}$;

(3) $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$;　　(4) $\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$;

(5) $\begin{pmatrix} -7 & 24 & 0 & 0 \\ 24 & 7 & 0 & 0 \\ 0 & 0 & -7 & 24 \\ 0 & 0 & 24 & 7 \end{pmatrix}$;　　(6) $\begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}$.

8 . 设 $\boldsymbol{A}$ 不是零方阵,证明:若 $\boldsymbol{A}^n = \boldsymbol{O}$,则 $\boldsymbol{A}$ 不能与对角形矩阵相似 .

9 . 设 n 阶方阵 $\boldsymbol{A}$ 的 n 个特征值互异, $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 有完全相同的特征值,证明:有可逆矩阵 $\boldsymbol{Q}$ 与另一矩阵 $\boldsymbol{R}$,使得 $\boldsymbol{A} = \boldsymbol{QR}$, $\boldsymbol{B} = \boldsymbol{RQ}$.

10 . 设 $\boldsymbol{A}$ 可逆且与对角形矩阵相似,证明: $\boldsymbol{A}^{-1}$ 也与对角形矩阵相似 .

11 . 设 $\boldsymbol{A}$ 是可逆矩阵且与对角形矩阵相似,证明: $\boldsymbol{A}^*$ 也与对角形矩阵相似 .

12 . 设 $\boldsymbol{A}$ 为实对称矩阵, $\boldsymbol{A}^2 = \boldsymbol{A}$,证明:存在正交矩阵 $\boldsymbol{P}$, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \mathrm{diag}(1,\cdots,1,0,\cdots,0)$.

13 . 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵, $\boldsymbol{A}^2 = \boldsymbol{I}_n$.证明:存在正交矩阵 $\boldsymbol{P}$,使

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I}_{n-r} \end{pmatrix},$$

其中 $\boldsymbol{I}_n$, $\boldsymbol{I}_r$, $\boldsymbol{I}_{n-r}$分别为 n, r, $n-r$ 阶单位矩阵 .

14 . 证明:对于任何矩阵 $\boldsymbol{A}_{s\times n}$, $\boldsymbol{B}_{n\times s}$,均有

(1) $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$;

(2) 若 $n = s$,则对任意 $a, b \in \mathbf{R}$,有 $\mathrm{tr}(a\boldsymbol{A} + b\boldsymbol{B}) = a\,\mathrm{tr}\,\boldsymbol{A} + b\,\mathrm{tr}\,\boldsymbol{B}$.

15 . 设 $\boldsymbol{I} - \boldsymbol{A}$ 的特征值之模均小于 1,证明 $0 < |\det \boldsymbol{A}| < 2^n$.

16 . 证明:方阵 $\boldsymbol{A}$ 的迹 $\mathrm{tr}\,\boldsymbol{A}$ 等于 $\boldsymbol{A}$ 的所有特征值之和 .

17 . 证明: n 阶实方阵 $\boldsymbol{A} = \boldsymbol{O}$ 的充要条件是对任意 n 阶方阵 $\boldsymbol{B}$,均有 $\mathrm{tr}(\boldsymbol{AB}) = 0$.

18 . 已知 $\boldsymbol{P}$ 为非奇异方阵,若有方阵 $\boldsymbol{A}$ 使 $\boldsymbol{B} = \boldsymbol{PAP}^{-1} - \boldsymbol{P}^{-1}\boldsymbol{AP}$,证明: $\boldsymbol{B}$ 的所有特征值之和为零 .

19 . 设 $\boldsymbol{A}$, $\boldsymbol{B}$ 均为 n 阶方阵,证明:矩阵

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{pmatrix}$$

的特征值是矩阵 $\boldsymbol{A} + \boldsymbol{B}$ 和 $\boldsymbol{A} - \boldsymbol{B}$ 的特征值之并集 .

20 . 设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$,与其对应的特征向量依次为 $\boldsymbol{\alpha}_1 = (1,1,1)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2 = (1,2,4)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3 = (1,3,9)^{\mathrm{T}}$,且 $\boldsymbol{b} = (1,1,3)^{\mathrm{T}}$.求 $\boldsymbol{A}^n\boldsymbol{b}$

(n 为自然数).

21. 设 3 阶实对称矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = -1, \lambda_2 = \lambda_3 = 1$ 且与 λ_1 对应的特征向量为 $\boldsymbol{a}_1 = (0,1,1)^{\mathrm{T}}$, 求 $\boldsymbol{A}$.

22. 设

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 2 \\ 5 & a & 3 \\ -1 & b & -2 \end{pmatrix}$$

的一个特征向量为 $\alpha = (1,1,-1)^{\mathrm{T}}$,

(1) 求 a, b 值及与特征向量 α 对应的 $\boldsymbol{A}$ 的特征值;

(2) $\boldsymbol{A}$ 与对角形矩阵是否相似? 说明理由.

23. 设 n 阶方阵 $\boldsymbol{A}$ 的伴随矩阵为 $\boldsymbol{A}^*$ 且 $|\boldsymbol{A}| \neq 0$, 若 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 求 $(\boldsymbol{A}^*)^2 + \boldsymbol{I}$ 的特征值.

24. 设

$$\boldsymbol{A} = \begin{pmatrix} a & -1 & c \\ 5 & b & 3 \\ 1-c & 0 & -a \end{pmatrix},$$

$|\boldsymbol{A}| = -1$, $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 有一个特征值 λ_0, 对应于 λ_0 的一个特征向量为 $\alpha = (-1, -1, 1)^{\mathrm{T}}$, 求 a, b, c, λ_0 的值.

习题 B

1. 设 $\boldsymbol{A}, \boldsymbol{B}$ 为任意 n 阶方阵, 证明: $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征多项式.

2. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 阶实对称矩阵. 证明: 若存在正交矩阵 $\boldsymbol{Q}$ 使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{AQ}$ 及 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{BQ}$ 都是对角阵, 则 $\boldsymbol{AB}$ 是实对称矩阵.

3. 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 n 阶实方阵, $\boldsymbol{AB} = \boldsymbol{BA}$, 并且有 n 个互异的特征值, 那么存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 和 $\boldsymbol{P}^{-1}\boldsymbol{BP}$ 同时为对角形矩阵.

4. 矩阵 $\boldsymbol{A}_{n\times n}$ 的所有元素均为 1. 证明: $\boldsymbol{A}$ 与对角矩阵相似并求该对角矩阵.

5. 已知 $\boldsymbol{b} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$, $c = (c_1, c_2, \cdots, c_n)^{\mathrm{T}}$, 定义矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 为 $a_{ij} = b_i b_j + c_i c_j$. 证明:

(1) $R(\boldsymbol{A}) \leq 2$, 且 $R(\boldsymbol{A}) = 2$ 的充要条件为 $\boldsymbol{b}, \boldsymbol{c}$ 线性无关;

(2) 求 $\boldsymbol{A}$ 的特征多项式.

第七章　二　次　型

这一章首先研究如何通过满秩线性变换将 n 元变量的二次齐次多项式——二次型化为标准形(即化为平方和的形式) .从第六章例 1 看到,通过正交变换把二次曲线化为标准方程这个几何问题就可归结为本章所要研究的问题 .这章还要研究所谓正定二次型的问题,即二次型满足什么条件,在 n 个变量任意变化时,二次型的值恒为正或恒为负 .多元函数极值问题、数理统计、数学规划、物理、力学、控制理论的某些问题可归结为这一问题 .最后,略述二次曲面分类问题 .

第一节　二　次　型

1.1　二次型与矩阵

首先,给出 n 元二次型的定义 .

定义 7.1　n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次多项式

$$f(x_1, x_2, \cdots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \tag{7.1}$$

其中 a_{ij} 为实数, $a_{ij} = a_{ji}(i, j = 1, 2, \cdots, n)$, 称 f 为实数域 $\mathbf{R}$ 上的 n 元二次型 .简称为二次型 .自然可以研究复数域上的二次型问题,但本书仅局限于实数域上的二次型的讨论 .

令 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, $\boldsymbol{A} = (a_{ij})_{n\times n}$, 其中 $a_{ij} = a_{ji}(i, j = 1, 2, \cdots, n)$, 则 (7.1)可写为

$$f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x}, \tag{7.2}$$

矩阵 $\boldsymbol{A}$ 称为二次型 f 对应的矩阵 . 显然,二次型与其对应的矩阵一一对应 .

二次型研究的主要问题之一是如何使用满秩线性变换 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, 其中 $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}$, 把 n 元二次型

$$f = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{C}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$$

化为完全平方和的形式:

$$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 . \tag{7.3}$$

把只含平方项不含混合项的二次型称为标准二次型,且把二次型(7.3)称为二次型(7.2)的标准形.

上述问题的几何起源就是通过坐标变换将二次曲线和二次曲面的方程化为标准形,以便判断曲线或曲面的类型,本节 1.2 将讨论这个问题.用矩阵理论可将二次型化为标准形问题可表述如下:对一个 n 阶实对称矩阵 $\boldsymbol{A}$,找一个 n 阶可逆矩阵 $\boldsymbol{C}$,使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 是对角形矩阵 $\Lambda=\operatorname{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)$.因此,有如下定义.

定义 7.2 对于 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$,若存在可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\boldsymbol{B}$,则称 $\boldsymbol{A}$ 合同于 $\boldsymbol{B}$,记为 $\boldsymbol{A}\simeq\boldsymbol{B}$.

显然,矩阵的合同关系是同阶矩阵集合上的等价关系.引用合同矩阵定义,可把二次型主要问题叙述为,寻找合同于 n 阶实对称矩阵 $\boldsymbol{A}$ 的对角形矩阵.下面讨论二次型化为标准形的两种主要方法——正交变换法及配方法.

1.2 正交变换法

1. 坐标变换

首先讨论一个几何问题,它是二次型问题的几何方面的来源.在 3 维欧氏几何空间中建立了直角坐标系$\{O,\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$,一个二次曲面的方程为

$$f=x_1^2-2x_2^2-2x_3^2-4x_1x_2+4x_1x_3+8x_2x_3=1. \tag{7.4}$$

试问这个方程的曲面是什么类型的曲面?

第二章研究了二次曲面,只有将方程(7.4)化为只含变量的平方项不含混合项,才容易判断其类型.建立了坐标系使曲面得以用代数方程来表示,但因坐标系不同曲面的方程也不同,因而希望找到一个合适的坐标系来定位这个曲面,使其方程最简单.为此,下面首先讨论一般仿射坐标系的坐标变换公式.

设在 3 维几何空间中建立了仿射坐标系$\{O;\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$,坐标向量(基)为 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$,如果坐标原点不动,将坐标向量(基)变为 $\varepsilon_1,\varepsilon_2,\varepsilon_3$,由旧的基到新的基的过渡矩阵为 $\boldsymbol{P}$,即

$$(\varepsilon_1,\varepsilon_2,\varepsilon_3)=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\boldsymbol{P}, \tag{7.5}$$

其中 $\boldsymbol{P}=(p_{ij})_{3\times3}$, $\varepsilon_i=\sum\limits_{1\leq k\leq3}p_{ki}\boldsymbol{e}_k$,由于一个点 M 对应一个向径 α,点 M 在旧的基下坐标为 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$,在新的基下坐标为 $\boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}}$,则

$$x_1\boldsymbol{e}_1+x_2\boldsymbol{e}_2+x_3\boldsymbol{e}_3=\alpha=y_1\varepsilon_1+y_2\varepsilon_2+y_3\varepsilon_3,$$

即

$$(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\boldsymbol{x}=\alpha=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\boldsymbol{y}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\boldsymbol{Py},$$

于是

$$\boldsymbol{x}=\boldsymbol{Py}. \tag{7.6}$$

这是一个点 M 在旧坐标系中的坐标与新坐标系中的坐标之间的变换公式,其

中 $\boldsymbol{P}$ 是可逆矩阵.如果两个基都是标准正交基,则由(7.5),有

$$\boldsymbol{I}=\begin{pmatrix}\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_1 & \boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2 & \boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_3\\ \boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_1 & \boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_2 & \boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_3\\ \boldsymbol{\varepsilon}_3\cdot\boldsymbol{\varepsilon}_1 & \boldsymbol{\varepsilon}_3\cdot\boldsymbol{\varepsilon}_2 & \boldsymbol{\varepsilon}_3\cdot\boldsymbol{\varepsilon}_3\end{pmatrix}$$

$$=\begin{pmatrix}\sum\limits_{1\le i,j\le 3}p_{i1}p_{j1}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i1}p_{j2}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i1}p_{j3}\boldsymbol{e}_i\cdot\boldsymbol{e}_j\\ \sum\limits_{1\le i,j\le 3}p_{i2}p_{j1}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i2}p_{j2}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i2}p_{j3}\boldsymbol{e}_i\cdot\boldsymbol{e}_j\\ \sum\limits_{1\le i,j\le 3}p_{i3}p_{j1}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i3}p_{j2}\boldsymbol{e}_i\cdot\boldsymbol{e}_j & \sum\limits_{1\le i,j\le 3}p_{i3}p_{j3}\boldsymbol{e}_i\cdot\boldsymbol{e}_j\end{pmatrix}$$

$$=\begin{pmatrix}\sum\limits_{1\le i\le 3}p_{i1}p_{i1} & \sum\limits_{1\le i\le 3}p_{i1}p_{i2} & \sum\limits_{1\le i\le 3}p_{i1}p_{i3}\\ \sum\limits_{1\le i\le 3}p_{i2}p_{i1} & \sum\limits_{1\le i\le 3}p_{i2}p_{i2} & \sum\limits_{1\le i\le 3}p_{i2}p_{i3}\\ \sum\limits_{1\le i\le 3}p_{i3}p_{i1} & \sum\limits_{1\le i\le 3}p_{i3}p_{i2} & \sum\limits_{1\le i\le 3}p_{i3}p_{i3}\end{pmatrix}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P},$$

于是 $\boldsymbol{P}$ 为正交矩阵.

从代数角度看,公式(7.6)与公式(5.8)都是线性变换,但两个公式的几何含义不同.坐标变换公式(7.6)表示同一点关于两个不同坐标系的两个坐标向量之间的关系;公式(5.8)表示在一个线性变换下,一点与其象在同一个坐标系下的坐标向量之间的关系.

下面,讨论坐标系平移.在 3 维几何空间的仿射坐标系$\{O;\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$中,点 O'的坐标 $\boldsymbol{c}=(c_1,c_2,c_3)^{\mathrm{T}}$,将坐标原点移至 O',建立新的仿射坐标系$\{O';\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$.如果点 M 在旧坐标系中坐标 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$,而在新坐标系中坐标 $\boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}}$,则得到坐标平移公式:

$$\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{c}. \tag{7.7}$$

注意,(7.7)式不是线性变换.

下面通过例子说明,如何使用正交坐标变换及平移坐标变换将表示二次曲面的三元二次方程化为标准形式.

例 1　在3维欧氏几何空间 $\mathbf{R}^3$ 中,将二次曲面方程

$$F=2x_1^2+3x_2^2+23x_3^2+72x_1x_3+6x_2-147=0 \tag{7.8}$$

化为标准形.

解　(1) 使用正交变换,将 F 中二次项部分化为完全平方和的形式.F 的二次项部分相应的二次型为

$$f=2x_1^2+3x_2^2+23x_3^2+72x_1x_3.$$

二次型 f 相应的矩阵

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 36 \\ 0 & 3 & 0 \\ 36 & 0 & 23 \end{pmatrix},$$

其特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 2 & 0 & -36 \\ 0 & \lambda - 3 & 0 \\ -36 & 0 & \lambda - 23 \end{vmatrix} = (\lambda - 3)(\lambda - 50)(\lambda + 25).$$

矩阵 $\boldsymbol{A}$ 的特征值 $\lambda_1 = 50, \lambda_2 = 3, \lambda_3 = -25$,相应的单位化且彼此正交的特征向量为

$$\boldsymbol{p}_1 = \begin{pmatrix} 3\ 5 \\ 0 \\ 4\ 5 \end{pmatrix}, \boldsymbol{p}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{p}_3 = \begin{pmatrix} -4\ 5 \\ 0 \\ 3\ 5 \end{pmatrix}.$$

使用正交变换

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}, \text{其中 } \boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3), \tag{7.9}$$

将(7.9)代入方程(7.8)得

$$\begin{aligned} F &= 50y_1^2 + 3y_2^2 - 25y_3^2 + 6y_2 - 147 \\ &= 50y_1^2 + 3(y_2^2 + 2y_2 + 1) - 25y_3^2 - 150 \\ &= 50y_1^2 + 3(y_2 + 1)^2 - 25y_3^2 - 150 \end{aligned}$$

(2) 平移变换

令

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{7.10}$$

则 $F = 50z_1^2 + 3z_2^2 - 25z_3^2 - 150$.于是曲面方程为

$$\frac{z_1^2}{(\ 3)^2} + \frac{z_2^2}{(5\ \ 2)^2} - \frac{z_3^2}{(\ 6)^2} = 1. \tag{7.11}$$

与(7.5)和(7.7)相对照,在旧坐标系中,原点在(0,0,0),坐标向量为 $\boldsymbol{e}_1 = (1,0,0)^{\mathrm{T}}$, $\boldsymbol{e}_2 = (0,1,0)^{\mathrm{T}}$, $\boldsymbol{e}_3 = (0,0,1)^{\mathrm{T}}$.在这个坐标系中,原点不动,以向量 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$ 作为坐标向量构成第二个坐标系 .在第二坐标系中找一个坐标为(0,1,0)的点,以这点为新原点,将第二个坐标系由原点平移至新原点而构成第三个坐标系 .在第三个坐标系中,方程化为标准形式(7.11) .

2. 主轴定理

现在研究一般二次型问题,由定理 6.7,对一般 n 元二次型有下面的定理(主轴定理) .

定理 7.1(主轴定理)**对于任何** n **元二次型** $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,**其中** $\boldsymbol{x}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)^{\mathrm{T}}$,$\boldsymbol{A}$ **为** n **阶实对称矩阵,存在一个正交变换** $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$,**其中** $\boldsymbol{P}$ **为** n **阶正交矩阵,** $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$,**使得** $f=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}=\sum_{i=1}^{n}\lambda_i y_i^2$,**其中** $\lambda_1,\lambda_2,\cdots,\lambda_n$ **为矩阵** $\boldsymbol{A}$ **的** n **个特征值,** $\boldsymbol{P}$ **的** n **个列向量依次为矩阵** A **对应于特征值** $\lambda_1,\lambda_2,\cdots,\lambda_n$ **的特征向量**.

例 2　将二次型

$$f=x_1^2-2x_2^2-2x_3^2-4x_1x_2+4x_1x_3+8x_2x_3.$$

化为标准形.

解　f **相应的矩阵为**

$$\boldsymbol{A}=\begin{pmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{pmatrix}.$$

矩阵 $\boldsymbol{A}$ **的特征多项式** $|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-2)^2(\lambda+7)$,**故特征值** $\lambda_1=\lambda_2=2$,$\lambda_3=-7$.**相应于** $\lambda_1=\lambda_2=2$ **的特征向量满足方程组**

$$\begin{pmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

也即

$$x_1+2x_2-2x_3=0.$$

因而特征向量 $\boldsymbol{x}_1=(-2,1,0)^{\mathrm{T}}$,$\boldsymbol{x}_2=(2,0,1)^{\mathrm{T}}$.**正交单位化后,得**

$$\boldsymbol{p}_1=\left(-\frac{2}{\sqrt{5}},\frac{1}{\sqrt{5}},0\right)^{\mathrm{T}},\ \boldsymbol{p}_2=\left(\frac{2}{3\sqrt{5}},\frac{4}{3\sqrt{5}},\frac{5}{3\sqrt{5}}\right)^{\mathrm{T}}.$$

对于特征值 $\lambda_3=-7$,**特征向量满足方程组**

$$\begin{pmatrix}-8&2&-2\\2&-5&-4\\-2&-4&-5\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

取特征向量 $\boldsymbol{x}_3=(1,2,-2)^{\mathrm{T}}$.**单位化后,得** $\boldsymbol{p}_3=(1/3,2/3,-2/3)^{\mathrm{T}}$.**用正交变换** $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$,**其中**

$$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix}-\frac{2}{\sqrt{5}}&\frac{2}{3\sqrt{5}}&\frac{1}{3}\\\frac{1}{\sqrt{5}}&\frac{4}{3\sqrt{5}}&\frac{2}{3}\\0&\frac{5}{3\sqrt{5}}&-\frac{2}{3}\end{pmatrix},$$

将二次型化为

$$f = 2y_1^2 + 2y_2^2 - 7y_3^2 .$$

1.3 配方法

从 1.2 小节知道,如何通过正交变换把二次型化为标准形,一般也可通过非正交的满秩线性变换将二次型化为标准形,下面介绍一种方法,即配方法.配方法的想法很简单,下面用两个例子说明这个方法.事实上,这两个例子也说明用配方法将二次型化为标准形的步骤.

例 3 将二次型

$$f(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$$

化为标准型,且写出所用满秩线性变换.

解 首先注意,该例中含有变量 x_1 的平方项.先对所有含 x_1 的项配方,如有 x_2 的平方项,再对所有含 x_2 的项配方等等:

$$\begin{aligned} f &= x_1^2 + 2(x_2 + x_3)x_1 + 2x_2^2 + 5x_3^2 + 4x_2x_3 \\ &= x_1^2 + 2(x_2 + x_3)x_1 + (x_2 + x_3)^2 - (x_2 + x_3)^2 + 2x_2^2 + 5x_3^2 + 4x_2x_3 \\ &= (x_1 + x_2 + x_3)^2 + x_2^2 + 4x_3^2 + 2x_2x_3 \\ &= (x_1 + x_2 + x_3)^2 + (x_2^2 + 2x_2x_3 + x_3^2) + 3x_3^2 \\ &= (x_1 + x_2 + x_3)^2 + (x_2 + x_3)^2 + 3x_3^2 . \end{aligned}$$

令

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3, & x_1 &= y_1 - y_2, \\ y_2 &= x_2 + x_3, & x_2 &= y_2 - y_3, \\ y_3 &= x_3, & x_3 &= y_3, \end{aligned} \tag{7.12}$$

则

$$f(x_1, x_2, x_3) = y_1^2 + y_2^2 + 3y_3^2 .$$

令 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$, $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}}$,二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,其中

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}.$$

由(7.3)知 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$ 是所用的满秩线性变换,其中

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

并且

$$C^{\mathrm{T}} AC = \mathrm{diag}(1,1,3).$$

例 4　将二次型

$$f(x_1, x_2, x_3) = x_1 x_2 + 2x_1 x_3 - 2x_2 x_3$$

化为标准形,且写出所用满秩线性变换.

解　注意,本例中 $f(x_1, x_2, x_3)$不含平方项,无法立即进行配方.先作一个满秩线性变换,使之出现平方项,然后再按例 1 方法处理.令

$$\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3. \end{cases} \tag{7.13}$$

于是

$$\begin{aligned} f &= (y_1 + y_2)(y_1 - y_2) + 2(y_1 + y_2)y_3 - 2(y_1 - y_2)y_3 \\ &= y_1^2 - y_2^2 + 4y_2 y_3 \\ &= y_1^2 - y_2^2 + 4y_2 y_3 - 4y_3^2 + 4y_3^2 \\ &= y_1^2 - (y_2 - 2y_3)^2 + 4y_3^2. \end{aligned}$$

令

$$\begin{cases} z_1 = y_1, \\ z_2 = y_2 - 2y_3, \\ z_3 = y_3, \end{cases} \qquad \begin{cases} y_1 = z_1, \\ y_2 = z_2 + 2z_3, \\ y_3 = z_3, \end{cases} \tag{7.14}$$

则

$$f(x_1, x_2, x_3) = z_1^2 - z_2^2 + 4z_3^2.$$

令 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$, $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}}$, $\boldsymbol{z} = (z_1, z_2, z_3)^{\mathrm{T}}$,二次型 f 对应的矩阵

$$\boldsymbol{A} = \begin{pmatrix} 0 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix},$$

(7.13)及(7.14)式可分别记为 $\boldsymbol{x} = \boldsymbol{D}\boldsymbol{y}$, $\boldsymbol{y} = \boldsymbol{F}\boldsymbol{z}$,则 $\boldsymbol{x} = \boldsymbol{D}\boldsymbol{F}\boldsymbol{z} = \boldsymbol{C}\boldsymbol{z}$ 是所用的满秩线性变换,其中

$$\boldsymbol{D} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{F} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \boldsymbol{C} = \boldsymbol{D}\boldsymbol{F} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

并且

$$\boldsymbol{C}^{\mathrm{T}} \boldsymbol{A}\boldsymbol{C} = \mathrm{diag}(1, -1, 4).$$

一般来说,对 n 个变量的二次型都可通过配方法将其化为标准型:(1) 如果二次型没有平方项,使用如(7.13)式那样的满秩线性变换,将它化为含有平方项

的二次型;(2) 如果二次型有平方项,首先找出一个有平方项的变量,再集中含有这个变量的所有项,并进行配方.如同例 2 那样.反复使用以上两个步骤,有限步后,一定能将含有 n 个变量的二次型化为标准形.因为以上所使用的两种变换均为满秩线性变换,因而它们的乘积也是满秩线性变换.这样,有如下定理.

定理 7.2　可以通过满秩线性变换,将二次型化为标准形.

注意,用配方法将二次型化为标准形所使用的线性变换是满秩的,不一定是正交变换,因而其标准形中平方项的系数也不一定是矩阵 A 的特征值.还应注意,如果例 2 中分别使用如下两个不同的线性变换

$$\begin{aligned} y_1 &= x_1 + x_2 + x_3, \\ y_2 &= \phantom{x_1 + {}} x_2 + x_3, \\ y_3 &= \phantom{x_1 + x_2 + {}} x_3. \end{aligned} \qquad \begin{aligned} y_1 &= x_1 + x_2 + x_3, \\ y_2 &= \phantom{x_1 + {}} x_2 + x_3, \\ y_3 &= \phantom{x_1 + x_2 + {}} 3x_3. \end{aligned}$$

可将 f 分别化为两个不同的标准形

$$f(x_1, x_2, x_3) = y_1^2 + y_2^2 + 3y_3^2 \quad \text{及} \quad f(x_1, x_2, x_3) = y_1^2 + y_2^2 + y_3^2.$$

事实上,存在其他的满秩线性变换,也可将 f 化为标准形.对一个二次型而言,它的标准形中平方项的系数可能不一样,但其正平方项与负平方项的个数是否唯一确定的呢?下面的小节将回答这个问题.

1.4　惯性定理

下面定理回答了第 1.3 节提出的问题.

定理 7.3(惯性定理)　若 n 元二次型 $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 中的矩阵 $\boldsymbol{A}$ 的秩为 r,用任意满秩线性变换 $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$,使

$$f = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q})\boldsymbol{y} = b_1 y_1^2 + \cdots + b_p y_p^2 - b_{p+1} y_{p+1}^2 - \cdots - b_r y_r^2,$$

则正项个数 p 与负项个数 $r - p$ 都是唯一确定的.

证明见附录 B.3.

由上面定理,引出如下定义.

定义 7.3　在二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的标准形中,正平方项个数称为二次型正惯性指标,负平方项个数称为二次型负惯性指标,正、负惯性指标的差称为符号差.

定理 7.3 表明,二次型的正惯性指标、负惯性指标及符号差在满秩线性变换下是不变的.

当二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 通过满秩线性变换 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$ 化为标准形 $f = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y}$,其中

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \operatorname{diag}(d_1, \cdots, d_p, -d_{p+1}, \cdots, -d_r, 0, \cdots, 0).$$

令

$$D = \operatorname{diag}\left(d_1, \cdots, d_p, d_{p+1}, \cdots, d_r, 1, \cdots, 1\right)^{-1},$$

则 $\boldsymbol{F}=\boldsymbol{CD}$ 使

$$\boldsymbol{F}^{\mathrm{T}}\boldsymbol{AF}=(\boldsymbol{CD})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{CD})=\operatorname{diag}(1,\cdots 1,-1,\cdots-1,0,\cdots,0).$$

这表明如果使用满秩线性变换 $\boldsymbol{x}=\boldsymbol{CDz}$,其中 $\boldsymbol{z}=(z_1,z_2,\cdots,z_n)^{\mathrm{T}}$,则

$$f=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2$$

称为二次型 f 的规范形,其中 $+1,-1$ 个数分别为二次型的正、负惯性指标 .

例 5　将二次型

$$6x_1x_3+8x_2x_3$$

先用正交变换化为标准形再用满秩线性变换将其化为规范形 .

解　1°　令 $f=6x_1x_3+8x_2x_3$,则二次型 f 对应的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}0&0&3\\0&0&4\\3&4&0\end{pmatrix}.$$

$\boldsymbol{A}$ 的特征方程为

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda&0&-3\\0&\lambda&-4\\-3&-4&\lambda\end{vmatrix}=\lambda(\lambda^2-25).$$

其特征值 $\lambda_1=5,\lambda_2=-5,\lambda_3=0$.

2°　特征值 $\lambda_1=5,\lambda_2=-5,\lambda_3=0$,其相应的特征向量分别满足如下三个方程:

$$\begin{pmatrix}5&0&-3\\0&5&-4\\-3&-4&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},\quad\begin{pmatrix}-5&0&-3\\0&-5&-4\\-3&-4&-5\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\text{ 及}$$

$$\begin{pmatrix}0&0&-3\\0&0&-4\\-3&-4&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

解方程组,分别得到对应于 $\lambda_1,\lambda_2,\lambda_3$ 的特征向量:

$$\boldsymbol{x}_1=(3,4,5)^{\mathrm{T}},\boldsymbol{x}_2=(3,4,-5)^{\mathrm{T}},\boldsymbol{x}_3=(-4,3,0)^{\mathrm{T}}.$$

单位化得

$$\boldsymbol{p}_1=\left(\frac{3}{5\sqrt{2}},\frac{4}{5\sqrt{2}},\frac{1}{\sqrt{2}}\right)^{\mathrm{T}},\boldsymbol{p}_2=\left(\frac{3}{5\sqrt{2}},\frac{4}{5\sqrt{2}},\frac{-1}{\sqrt{2}}\right)^{\mathrm{T}},\boldsymbol{p}_3=\left(-\frac{4}{5},\frac{3}{5},0\right)^{\mathrm{T}}.$$

3°　用正交变换 $\boldsymbol{x}=\boldsymbol{Py}$,其中 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)$,则 $f=5y_1^2-5y_2^2$,再令 $\boldsymbol{y}=\boldsymbol{Qz}$,其中

$$\boldsymbol{z}=(z_1,z_2,z_3)^{\mathrm{T}},Q=\operatorname{diag}\left(\frac{1}{\sqrt{5}},\frac{1}{\sqrt{5}},1\right),$$

则满秩线性变换 $\boldsymbol{x}=\boldsymbol{PQz}$,使

$$f=z_1^2-z_2^2.$$

该二次型已化为规范形,其中的正惯性指标为 1,负惯性指标为 1,符号差为 0,相应矩阵的秩为 2 .

例 6 设 $\boldsymbol{A}$ 是 n 阶实对称矩阵,$\boldsymbol{A}$ 的秩为 r .证明存在秩为 $n-r$ 的实对称矩阵 $\boldsymbol{B}$,使 $\boldsymbol{AB}=\boldsymbol{O}$.

证 因为 $\boldsymbol{A}$ 是秩为 r 的 n 阶实对称矩阵,所以存在可逆矩阵 $\boldsymbol{P}$,使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\boldsymbol{D}=\operatorname{diag}(d_1,d_2,\cdots,d_r,0,\cdots,0),$$

其中 $d_i\neq 0(i=1,2,\cdots,r)$.令

$$\boldsymbol{C}=\operatorname{diag}(0,\cdots,0,c_{r+1},c_{r+2},\cdots,c_n),$$

其中 $c_i\neq 0(i=r+1,r+2,\cdots,n)$.显然 $\boldsymbol{DC}=\boldsymbol{O}$,也即 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APC}=\boldsymbol{O}$,因为 $\boldsymbol{P}$ 可逆,所以

$$\boldsymbol{APCP}^{\mathrm{T}}=(\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APCP}^{\mathrm{T}}=(\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{OP}^{\mathrm{T}}=\boldsymbol{O}.$$

令 $\boldsymbol{PCP}^{\mathrm{T}}=\boldsymbol{B}$,因为 $\boldsymbol{P}$ 可逆,所以矩阵 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 的秩都等于 $n-r$,其中 $\boldsymbol{B}$ 是对称矩阵且使 $\boldsymbol{AB}=\boldsymbol{O}$.

第二节 正定二次型

2.1 正定二次型

首先讨论一个例子 .令

$$f=2x_1^2+2x_2^2+3x_3^2-2x_1x_2+4x_1x_3,$$

问函数 f 是否有最小值 . f 可写成二次型形式,即 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$.其中 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$,且

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

如果 $\boldsymbol{x}\neq\boldsymbol{0}$ 时,总有 $f=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}>0$,则 f 在 $\boldsymbol{x}=\boldsymbol{0}$ 处达到最小值 .

对一般二次型,有下面定义 .

定义 7.4 若对任意非零向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,恒有二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}>0$,其中 $\boldsymbol{A}$ 为实对称矩阵,则称 f 为正定二次型,$\boldsymbol{A}$ 为正定矩阵 .

从上面讨论看到,正定二次型(及后面讨论的负定二次型)在判断函数极值中有重要作用 .因而,在数学自身及实际应用中,正定二次型或正定矩阵有重要的意义 .在这一节后半部分再讨论它的应用 .首先,给出 $\boldsymbol{A}$ 是正定矩阵的充要

条件．

定理 7.4　设 $\boldsymbol{A}$ 为 n 阶实对称矩阵,则下列命题等价:

(1) $\boldsymbol{A}$ 是正定矩阵;

(2) $\boldsymbol{A}$ 的正惯性指标为 n 即 $\boldsymbol{A} \simeq \boldsymbol{I}_n$;

(3) $\boldsymbol{A}$ 的 n 个特征值都大于零．

证　(1)　(2)　设 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型,据定理 7.1,存在满秩线性变换 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$ 使

$$f = b_1 y_1^2 + \cdots + b_n y_n^2,$$

其中 $\boldsymbol{y} = (y_1, \cdots, y_n)^{\mathrm{T}}$．如果某个 $b_k \leq 0 (1 \leq k \leq n)$,令 $\boldsymbol{y}_k$ 是第 k 个分量为 1 且其余分量为 0 的向量,则 $\boldsymbol{x}_k = \boldsymbol{C}\boldsymbol{y}_k$ 也为非零向量,这时

$$f = \boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_k = b_k \leq 0,$$

这与 f 是正定二次型矛盾．于是, $\boldsymbol{A}$ 的正惯性指标为 n．从而 $\boldsymbol{A}$ 与 $\boldsymbol{I}_n$ 合同,即 $\boldsymbol{A} \simeq \boldsymbol{I}_n$．

(2)　(3) 设 $\boldsymbol{A}$ 的正惯指标为 n,使用正交变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$,使二次型化为标准形,则

$$f = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{A}$ 的特征值．由惯性定理知,特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 都大于零．

(3)　(1) 设 $\boldsymbol{A}$ 的 n 个特征值都大于零,使用正交变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$,使二次型化为标准形

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2,$$

其中特征值 $\lambda_i > 0 (i = 1, \cdots, n)$．对任意 $\boldsymbol{x} \neq \boldsymbol{0}$,则 $\boldsymbol{y} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{x} \neq \boldsymbol{0}$,因而

$$f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 > 0.$$

这样, $\boldsymbol{A}$ 是正定矩阵．□

下面,给出如何利用二次型对应的矩阵的顺序主子式(左上角主子式)判断二次型正定的方法．

定理 7.5　n 元实二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的充分必要条件是 $\boldsymbol{A}$ 的 n 个顺序主子式(左上角主子式)全大于零,即

$$|\boldsymbol{A}_k| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} > 0 (k = 1, 2, \cdots, n).$$

证明见附录 B.4.

例 1　试问 t 取什么值,二次型

$$f = tx_1^2 + 2x_2^2 + 3x_3^2 - 2x_1x_2 + 4x_1x_3$$

是正定的.

解 二次型对应的矩阵为

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

各阶主子式为

$$|\boldsymbol{A}_1| = t,\ |\boldsymbol{A}_2| = \begin{vmatrix} t & -1 \\ -1 & 2 \end{vmatrix} = 2t - 1,\ |\boldsymbol{A}_3| = |\boldsymbol{A}| = 6t - 11.$$

当 $t > \dfrac{11}{6}$时,$\boldsymbol{A}$ 为正定的,所以 f 也是正定的,因而 f 在 $\boldsymbol{x} = (0,0,0)^{\mathrm{T}}$ 点取得最小值 0.本节开始举的例子,正是这个例子中 $t = 2$ 的情形.

例 2 证明:若 $\boldsymbol{A}$ 是正定矩阵,则存在正定矩阵 $\boldsymbol{B}$ 使 $\boldsymbol{A} = \boldsymbol{B}^2$.

证 因为 $\boldsymbol{A}$ 是实对称矩阵,而且特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n > 0$,所以存在正交矩阵 $\boldsymbol{P}$(即 $\boldsymbol{P}^{-1} = \boldsymbol{P}^{\mathrm{T}}$),使得

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

于是

$$\begin{aligned} \boldsymbol{A} &= P\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\boldsymbol{P}^{\mathrm{T}} \\ &= \boldsymbol{P}\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\boldsymbol{P}^{\mathrm{T}}. \end{aligned}$$

令

$$\boldsymbol{B} = \boldsymbol{P}\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\boldsymbol{P}^{\mathrm{T}},$$

则 $\boldsymbol{A} = \boldsymbol{B}^2$.因为 $\boldsymbol{B}$ 的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_n > 0$,所以 $\boldsymbol{B}$ 也是正定的.

2.2 负定二次型

定义 7.5 若对任意非零向量 $\boldsymbol{x} = (x_1, \cdots, x_n)^{\mathrm{T}}$,恒有二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax} < 0$,称 f 为负定二次型,且 $\boldsymbol{A}$ 为负定矩阵.

根据定义,如果 $\boldsymbol{A} = (a_{ij})_{n\times n}$为正定的,那么 $-\boldsymbol{A} = (-a_{ij})_{n\times n}$为负定的,相应于定理 7.4 及定理 7.5,有如下定理,

定理 7.6 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵,则下列命题等价:

(1) $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 是负定的;

(2) $\boldsymbol{A}$ 的负惯性指标为 n;

(3) $\boldsymbol{A}$ 的 n 个特征值都小于零;

(4) $\boldsymbol{A}$ 的奇数阶顺序主子式小于 0,偶数阶顺序主子式大于 0.

例 3 试证矩阵

$$A=\begin{pmatrix}-5&2&2\\2&-6&0\\2&0&-4\end{pmatrix}$$

是负定的.

证　因为 A 的各阶顺序主子式为

$$|A_1|=-5<0,\ |A_2|=\begin{vmatrix}-5&2\\2&-6\end{vmatrix}=26>0,\ |A_3|=|A|=-80<0.$$

由定理 7.6 知,A 是负定的.

*2.3　多元函数极值存在的充分条件

从二元函数微分学知道

$$\begin{aligned}f(x+h,y+k)&=f(x,y)+f'_x(x,y)h+f'_y(x,y)k+\\&\quad f''_{xx}(x,y)h^2+2f''_{xy}(x,y)hk+f''_{yy}(x,y)k^2+o(h^2+k^2),\end{aligned}$$

其中 $o(h^2+k^2)$是关于 h^2+k^2 的高阶无穷小,函数在(x,y)处取极值的必要条件为 $f'_x(x,y)=f'_y(x,y)=0$.当 $f'_y(x,y)=f'_x(x,y)=0$ 时,

$$\begin{aligned}\Delta f&=f(x+h,y+k)-f(x,y)\\&=(h,k)\begin{pmatrix}f''_{xx}(x,y)&f''_{xy}(x,y)\\f''_{xy}(x,y)&f''_{yy}(x,y)\end{pmatrix}\begin{pmatrix}h\\k\end{pmatrix}+o(h^2+k^2)\\&=(h,k)A\begin{pmatrix}h\\k\end{pmatrix}+o(h^2+k^2),\end{aligned}$$

在点(x,y)的某一个邻域内,当按上式定义的矩阵 A 为正定(或负定),则 Δf 恒正(或恒负),因而 f 在点(x,y)处取极小值(或极大值).

例 4　求函数 $f=x^2-xy+y^2-2x+y$ 的极小值.

解　由

$$\frac{\partial f}{\partial x}=2x-y-2=0,$$

$$\frac{\partial f}{\partial y}=-x+2y+1=0,$$

知$(x,y)=(1,0)$是函数 f 的唯一驻点,且

$$A=\begin{pmatrix}f''_{xx}(x,y)&f''_{xy}(x,y)\\f''_{xy}(x,y)&f''_{yy}(x,y)\end{pmatrix}_{(x=1,y=0)}=\begin{pmatrix}2&-1\\-1&2\end{pmatrix}.$$

因为 A 是正定的,所以 f 在$(1,0)$处有极小值,且 f 的极小值为 $f(1,0)=-1$.

对 n 元函数 $f(x_1,x_2,\cdots,x_n)$有类似的结论,不再详细论述.

例 5　设 A 为正定矩阵,证明:x^* 是 $Ax=b$ 的解的充分必要条件是 x^* 使

二次函数

$$f(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x})$$

达到极小值,即对任意 $\boldsymbol{x} \in \mathbf{R}^n$,有 $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$.

证　设 $\boldsymbol{x}^*$ 为 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解,即 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$.于是

$$\begin{aligned} f(\boldsymbol{x}^*) &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*) - (\boldsymbol{b}, \boldsymbol{x}^*) \\ &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*) - (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*) = -\frac{1}{2}(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*), \end{aligned}$$

并且

$$\begin{aligned} F(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x}^*) &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) + \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*) \\ &= \frac{1}{2}[(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) + (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*)] \\ &= \frac{1}{2}[(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}) - (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}) + (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*)] \\ &= \frac{1}{2}[(\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*), \boldsymbol{x}) - (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^*)]. \end{aligned}$$

因为 $\boldsymbol{A}$ 是实对称矩阵,$\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$,所以

$$\begin{aligned} (\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^*) &= (\boldsymbol{A}\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) = \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{x}^*) \\ &= \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*) = (\boldsymbol{x}^*, \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*)) = (\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*), \boldsymbol{x}^*). \end{aligned}$$

从而

$$\begin{aligned} F(\boldsymbol{x}) &= \frac{1}{2}(\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*), (\boldsymbol{x} - \boldsymbol{x}^*)) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^*, \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*)) \\ &= \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^*). \end{aligned}$$

于是 $F(\boldsymbol{x})$是二次型,由题设,其对应的矩阵 $\boldsymbol{A}$ 是正定的.$f(\boldsymbol{x})$取极小值等价于 $F(\boldsymbol{x})$取极小值.因为 $\boldsymbol{A}$ 是正定的,当 $\boldsymbol{x} - \boldsymbol{x}^* \neq \boldsymbol{0}$,恒有 $F(\boldsymbol{x}) > 0$.因此,$F(\boldsymbol{x})$达到极小值的充分必要条件是 $\boldsymbol{x} - \boldsymbol{x}^* = \boldsymbol{0}$,其中 $\boldsymbol{x}^*$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解.

这个定理是解线性方程组两种迭代法——最速下降法、共轭斜量法的基础,它把求解系数矩阵为正定矩阵的线性方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 问题,转化为求二次函数的极小值问题.可以用迭代法求函数的极小值点 $\boldsymbol{x}^*$,从而得到方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解 $\boldsymbol{x}^*$.

*第三节　二次曲面的度量分类

一个三元二次方程所表示的几何空间中的曲面称为二次曲面,椭球面、双曲面、抛物面等都是二次曲面.已知一个二次曲面的方程,常常不能立刻看出它所

表示的曲面类型,二次型的理论有助于这个问题的解决.如果从传统解析几何观点研究空间二次曲面,应从曲面几何特征出发,给出直径面、中心及主方向等概念,进而使用二次曲面不变量理论,给出二次曲面的完整分类.由于篇幅所限,本节仅略述如何通过坐标变换将二次曲面化为标准方程,并在此基础上,简述二次曲面分类.

在 3 维欧氏几何空间 $\mathbf{R}^3$ 中,也即假设已经建立了直角坐标系$\{O;\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$,若点 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}$ 满足方程 $F(x_1,x_2,x_3)=0$,其中

$$\begin{aligned}F(\boldsymbol{x})={}&a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+\\&2a_{13}x_1x_3+2a_{23}x_2x_3+b_1x_1+b_2x_2+b_3x_3+c,\end{aligned}\tag{7.15}$$

是三元二次多项式,则称方程确定的曲面为二次曲面.令

$$\boldsymbol{b}=(b_1,b_2,b_3)^{\mathrm{T}},\ \boldsymbol{A}=(a_{ij})_{3\times3},$$

其中 $a_{ij}=a_{ji}$, $i,j=1,2,3$,即 $\boldsymbol{A}$ 是实对称矩阵,则(7.15)可写为

$$F(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c.\tag{7.16}$$

由定理 6.12 可知,存在正交矩阵 $\boldsymbol{P}$,使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3)$,且 $\lambda_1,\lambda_2,\lambda_3$ 为矩阵 $\boldsymbol{A}$ 的特征值,$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)$,其中 $\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3$ 分别为对应于 $\lambda_1,\lambda_2,\lambda_3$ 的正交单位化的特征向量.使用正交的坐标变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$,有

$$\begin{aligned}F(\boldsymbol{x})&=(\boldsymbol{P}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{y})+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}+c\\&=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}+c\\&=\lambda_1y_1^2+\lambda_2y_2^2+\lambda_3y_3^2+2\mu_1y_1+2\mu_2y_2+2\mu_3y_3+c,\end{aligned}\tag{7.17}$$

其中

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{P}=(b_1,b_2,b_3)\boldsymbol{P}=2(\mu_1,\mu_2,\mu_3)$$

于是,当 $\lambda_1\lambda_2\lambda_3\neq0$ 时,有

$$\begin{aligned}F={}&\lambda_1\left(y_1+\frac{\mu_1}{\lambda_1}\right)^2+\lambda_2\left(y_2+\frac{\mu_2}{\lambda_2}\right)^2+\lambda_3\left(y_3+\frac{\mu_3}{\lambda_3}\right)^2+\\&c-\left(\frac{\mu_1^2}{\lambda_1}+\frac{\mu_2^2}{\lambda_2}+\frac{\mu_3^2}{\lambda_3}\right)=0.\end{aligned}\tag{7.18}$$

对方程(7.18)使用平移坐标变换 $\boldsymbol{z}=\boldsymbol{y}+\boldsymbol{y}_0$,其中

$$\boldsymbol{z}=(z_1,z_2,z_3)^{\mathrm{T}},\ \boldsymbol{y}=(y_1,y_2,y_3)^{\mathrm{T}},\ \boldsymbol{y}_0=(\mu_1/\lambda_1,\mu_2/\lambda_2,\mu_3/\lambda_3)^{\mathrm{T}},$$

则方程 $F(\boldsymbol{x})=0$ 变为

$$\lambda_1z_1^2+\lambda_2z_2^2+\lambda_3z_3^2+d=0,\tag{7.19}$$

其中 $d=c-(\mu_1^2/\lambda_1+\mu_2^2/\lambda_2+\mu_3^2/\lambda_3)$.

下面讨论二次曲面 $F(\boldsymbol{x})=0$ 的分类.用 $R(\boldsymbol{A})$记矩阵 $\boldsymbol{A}$ 的秩

情形 1 $R(\boldsymbol{A})=3$,即 $\lambda_1\lambda_2\lambda_3\neq0$.考察(7.19)式.

情形 1.1　$\lambda_1,\lambda_2,\lambda_3$ **同号(符号相同)**.

(a) $\lambda_1,\lambda_2,\lambda_3,d$ **同号,(方程表示的)曲面是虚椭球面(在** $\mathbf{R}^3$ **中不存在实在曲面);**

(b) $\lambda_1,\lambda_2,\lambda_3,d$ **异号,曲面是椭球面;**

(c) $d=0$,**曲面退化为一点**.

情形 1.2　$\lambda_1,\lambda_2,\lambda_3$ **异号**.

(a) $\lambda_1,\lambda_2,\lambda_3$ **中两个为正,一个为负**, $d<0$;**或** $\lambda_1,\lambda_2,\lambda_3$ **中两个为负,一个为正**, $d>0$ **时,曲面都是单叶双曲面;**

(b) $\lambda_1,\lambda_2,\lambda_3$ **中两个为正,一个为负**, $d>0$;**或** $\lambda_1,\lambda_2,\lambda_3$ **中两个为负,一个为正**, $d<0$,**曲面都是双叶双曲面;**

(c) $d=0$,**曲面是二次锥面**.

情形 2　$R(\boldsymbol{A})=2$,**即** $\lambda_1,\lambda_2,\lambda_3$ **有一个为零,其余两个不为零.不失一般性,设** $\lambda_3=0$.

情形 2.1　**在**(7.17)**中,若** $\mu_3=0$,**于是**

$$F=\lambda_1\left(y_1+\frac{\mu_1}{\lambda_1}\right)^2+\lambda_2\left(y_2+\frac{\mu_2}{\lambda_2}\right)^2+\left(c-\left(\frac{\mu_1^2}{\lambda_1}+\frac{\mu_2^2}{\lambda_2}\right)\right)=0. \tag{7.20}$$

对方程(7.20)**使用平移变换** $\boldsymbol{z}=\boldsymbol{y}+\boldsymbol{y}_0$,**其中** $\boldsymbol{y}_0=(\mu_1/\lambda_1,\mu_2/\lambda_2,0)^{\mathrm{T}}$,**则方程** $F=0$ **化为**

$$\lambda_1 z_1^2+\lambda_2 z_2^2+d=0, \tag{7.21}$$

其中 $d=c-(\mu_1^2/\lambda_1+\mu_2^2/\lambda_2)$.

情形 2.1.1　λ_1,λ_2 **同号**

(a) λ_1,λ_2,d **同号,曲面是虚椭圆柱面;**

(b) λ_1,λ_2 **同号且与** d **异号,曲面是椭圆柱面;**

(c) λ_1,λ_2 **同号且** $d=0$,**一条直线**.

情形 2.1.2　λ_1,λ_2 **异号**.

(a) λ_1,λ_2 **异号**, $d\neq0$,**曲面是双曲柱面;**

(b) λ_1,λ_2 **异号**, $d=0$,**曲面是相交的两个平面**.

情形 2.2　**在**(7.17)**中,若** $\mu_3\neq0$,**于是曲面方程**(7.17)**可由平移变换化为(略去具体变换过程)**

$$\lambda_1 z_1^2+\lambda_2 z_2^2+2\mu_3 z_3=0.$$

(a) λ_1,λ_2 **同号,曲面是椭圆抛物面;**

(b) λ_1,λ_2 **异号,曲面是双曲抛物面**.

情形 3　$R(\boldsymbol{A})=1$,**即** $\lambda_1,\lambda_2,\lambda_3$ **中有两个为零,一个不为零.不失一般性,**

设 $\lambda_1 \neq 0$.

表形 3 .1 曲面方程(7.17)可化为如下形式(略去具体变换过程)

$$\lambda_1 z_1^2 + r z_2 + s z_3 = 0, \tag{7.22}$$

其中 $r \neq 0, s \neq 0$.进行正交坐标变换 $\boldsymbol{z} = \boldsymbol{Q}\boldsymbol{w}$,其中 $\boldsymbol{z} = (z_1, z_2, z_3)^{\mathrm{T}}$, $\boldsymbol{w} = (w_1, w_2, w_3)^{\mathrm{T}}$,且

$$\boldsymbol{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

其中 $\cos\theta = r / \sqrt{r^2 + s^2}$, $\sin\theta = s / \sqrt{r^2 + s^2}$,则方程(7.17)呈

$$\lambda_1 w_1^2 + \sqrt{r^2 + s^2}\, w_2 = 0,$$

曲面是抛物柱面 .

情形 3.2 曲面方程(7.17)可化为如下形式

$$\lambda_1 z_1^2 + r z_2 = 0 \text{(或 } \lambda_1 z_1^2 + s z_3 = 0\text{)},$$

其中 $r \neq 0$(或 $s \neq 0$) .曲面为抛物柱面 .

情形 3.3 曲面方程(7.17)可化为如下形式

$$\lambda_1 z_1^2 + d = 0 .$$

(a) λ_1 与 d 同号,曲面是一对虚平行平面;

(b) λ_1 与 d 异号,曲面是一对平行平面;

(c) $d = 0$,曲面是一对重合平面 .

例 1 在空间直角坐标系中,二次曲面方程为

$$F = 2x_1x_2 + 2x_1x_3 + 2x_2x_3 - 6x_1 - 6x_2 - 4x_3 + 9 = 0,$$

判断曲面类型 .

解 F 中二次项部分构成的二次型的相应矩阵及特征多项式分别为

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

于是,矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = -1$, $\lambda_2 = -1$, $\lambda_3 = 2$,按(7.17)～(7.19)的步骤计算得 $d = 25$.与上面讨论情形 1.2(a)对照,曲面是单叶双曲面 .

习 题 七

习题 A

1 . 用配方法,将下列二次型化为标准形,并写出所用的满秩线性变换 .

(1) $f=x_1^2+x_2^2+x_3^2+2x_1x_2+2x_1x_3+8x_2x_3$;

(2) $f=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$;

(3) $f=2x_1x_2+2x_1x_3-6x_2x_3$;

(4) $f=2x_1x_2+4x_1x_3$;

(5) $f=x_1x_{2n}+x_2x_{2n-1}+\cdots+x_nx_{n+1}$.

2. **用正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$,将下列二次型化为标准形,并写出正交矩阵 $\boldsymbol{P}$.**

(1) $f=2x_1^2+2x_2^2-2x_1x_2$;

(2) $f=5x_1^2+2x_2^2+4x_1x_2$;

(3) $f=2x_1^2+5x_2^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3$;

(4) $f=-5x_1^2+x_2^2-x_3^2+4x_1x_2+6x_1x_3$.

3. **用正交坐标变换和平移变换把如下二次曲面化为标准形,并写出坐标变换公式:**

(1) $2x^2-4xy-y^2+8=0$;

(2) $5x^2+4xy+5y^2-9=0$;

(3) $9x^2-4xy+6y^2-10x-20y-5=0$;

(4) $21x^2+6xy+13y^2-114x+34y+73=0$.

4. **证明:实数域上两个二次型合同的充要条件是两个二次型对应的矩阵秩和符号差相同.**

5. **已知 $\boldsymbol{A}$ 是 $m\times n$ 实矩阵,试证: $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是具有非负特征值的实对称矩阵.**

6. **判断如下矩阵是否正定,负定:**

(1) $\begin{pmatrix}1&1&1\\1&2&3\\1&3&6\end{pmatrix}$; (2) $\begin{pmatrix}99&-6&24\\-6&130&-30\\24&-30&71\end{pmatrix}$;

(3) $\begin{pmatrix}-5&2&2\\2&-6&0\\2&0&-4\end{pmatrix}$; (4) $\begin{pmatrix}10&4&12\\4&2&-14\\12&-14&1\end{pmatrix}$.

7. **试问 t 取何值,下列二次型为正定的:**

(1) $f=2x_1^2+x_2^2+x_3^2+2x_1x_2+tx_2x_3$;

(2) $f=t(x_1^2+x_2^2+x_3^2)+2x_1x_2+2x_1x_3-2x_2x_3+x_4^2$;

8. **设实对称矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 是正定的, $b_1,b_2,\cdots,b_n$ 是任意 n 个非零实数.证明: $\boldsymbol{B}=(a_{ij}b_ib_j)_{n\times n}$ 也是正定的.**

9. **设 $\boldsymbol{A}$ 是正定矩阵,证明: $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^{\mathrm{T}}$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 也是正定矩阵.**

10. **设 $f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是一个实二次型, $\boldsymbol{A}$ 为实对称矩阵. λ_1,**

$\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的特征值,且 $\lambda_1\leq\lambda_2\cdots\leq\lambda_n$.证明:对 $\mathbf{R}^n$ 任意向量 $\boldsymbol{x}$,都有

$$\lambda_1\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\leq\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leq\lambda_n\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}.$$

11. 设 A 与 B 为实正定矩阵 .证明: AB 正定的充要条件 $\boldsymbol{AB}=\boldsymbol{BA}$.

12 . 设 $\boldsymbol{A}$ 为实对称矩阵 .证明:对任意奇数 m,必有实对称矩阵 $\boldsymbol{B}$,使 $\boldsymbol{B}^m=\boldsymbol{A}$.

13 . 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为 n 维向量空间 $\mathbf{R}^n$ 的一个基 .试证:

(1) 任给 $\boldsymbol{a},\boldsymbol{b}\in\mathbf{R}^n$. $\boldsymbol{a}=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n$, $\boldsymbol{b}=b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n$.若定义运算

$$(\boldsymbol{a},\boldsymbol{b})=(a_1,a_2,\cdots,a_n)\boldsymbol{B}\begin{pmatrix}b_1\\b_2\\\cdots\\b_n\end{pmatrix},$$

其中 $\boldsymbol{B}$ 为 n 阶实正定矩阵 .则 $(\boldsymbol{a},\boldsymbol{b})$ 为 $\mathbf{R}^n$ 的内积运算且在此内积运算下 $\mathbf{R}^n$ 为一欧氏空间 .

(2) 若 $\mathbf{R}^n$ 在内积运算 $(\ .,\ .)$ 下为欧氏空间,令 $b_{ij}=(\alpha_i,\alpha_j)$,则矩阵 $\boldsymbol{B}=(b_{ij})$,为正定矩阵 .

14 . 设二次型 $f(x_1,x_2,x_3)=5x_1^2+5x_2^2+cx_3^2-2x_1x_2+6x_1x_3-6x_2x_3$ 的秩为 2 .

(1) 求参数 c 及此二次型对应矩阵的特征值;

(2) 指出方程 $f(x_1,x_2,x_3)=1$ 表示何种二次曲面 .

15 . 设 $\boldsymbol{A}$ 是 n 阶正定阵,求证: $|\boldsymbol{A}+\boldsymbol{I}|>1$.

16 . 二次型 $f(x_1,x_2,x_3)=2x_1^2+3x_2^2+3x_3^2+2ax_2x_3$,其中 $a>0$,通过正交变换可化为标准形式: $f=y_1^2+2y_2^2+5y_3^2$,求参数 a 及所用的正交变换矩阵 .

17 . 设二次曲面 $x^2+ay^2+z^2+2bxy+2xz+2yz=4$ 可经过正交变换

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}\xi\\\eta\\\zeta\end{pmatrix}$$

化为椭圆柱面方程 $\eta^2+4\zeta^2=4$,求 a,b 及所用的正交变换矩阵 $\boldsymbol{P}$.

18 . $\boldsymbol{A}$ 为 n 阶实对称矩阵,证明:对于任意实列向量 $\boldsymbol{x}\in\mathbf{R}^n$ 都有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=0$ 的充要条件为 $\boldsymbol{A}=\boldsymbol{O}$.

19 . 证明:秩为 r 的对称矩阵可以表为 r 个秩为 1 的对称矩阵之和 .

20 . $\boldsymbol{A}$ 为 n 阶实对称矩阵,证明:存在一正实数 γ,使得对于任意实列向量 $\boldsymbol{x}\in\mathbf{R}^n$ 都有 $|\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}|\leq\gamma\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

习题 B

1．设 $\boldsymbol{A}$ 是 n 阶可逆矩阵，证明：存在正交矩阵 $\boldsymbol{S}$ 和正定矩阵 $\boldsymbol{U}$，使 $\boldsymbol{A}=\boldsymbol{US}$．

2．设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为 n 阶实对称矩阵，且 $\boldsymbol{B}$ 为正定矩阵．证明：存在 n 阶实可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 与 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}$ 同时为对角矩阵．

3．判定下列二次曲面类型．

(1) $5x^2-3y^2-2xy+8yz+8zx+2y+4z=0$;

(2) $25x^2+9y^2+5z^2+20xy+10yz+10x+4y+2\quad 6z+12=0$;

(3) $2x^2+3y^2+4z^2+4xy+4yz+4x-2y-12z+10=0$;

(4) $-2x^2+7y^2+4z^2+4xy+20yz+16zx+4x+2y-4z+8=0$．

4．设 $\boldsymbol{A}$ 为 m 阶实正定对称矩阵，$\boldsymbol{B}$ 为 $m\times n$ 实矩阵 $m\geq n$，证明：$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{AB}$ 为正定矩阵的充要条件是 $\boldsymbol{B}$ 的秩为 n．

5．证明：二次型

$$f=\sum_{i=1}^{s}(a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n)^2$$

系数矩阵的秩等于下列矩阵的秩

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ a_{s1} & a_{s2} & \cdots & a_{sn} \end{matrix}.$$

6．若分块矩阵

$$\boldsymbol{M}=\begin{matrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{D} \end{matrix}$$

为正定矩阵，证明：

(1) $\boldsymbol{D}-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}$ 为正定矩阵；

(2) $|\boldsymbol{M}|\leq|\boldsymbol{A}||\boldsymbol{D}|$，且等号成立的充要条件为 $\boldsymbol{B}=\boldsymbol{O}$．

7．设 $\boldsymbol{A}$ 为 n 阶正定阵，

$$\boldsymbol{A}=\begin{matrix} \boldsymbol{A}_{n-1} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^{\mathrm{T}} & a_{nn} \end{matrix},$$

证明：$a_{nn}|\boldsymbol{A}_{n-1}|>\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{*}\boldsymbol{\alpha}$．

8．$\boldsymbol{A}$ 正定，证明：$\boldsymbol{A}$ 的绝对值最大的元素一定位于主对角线上．

9．实列向量 $\boldsymbol{x}\in\mathbf{R}^n$，$a\in\mathbf{R}$，$\boldsymbol{A}$ 为 n 阶正定矩阵．

(1) 若 $1+a\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}>0$,证明:矩阵 $\boldsymbol{I}+a\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}$ 正定;

(2) 若 $1+a\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}>0$,证明:矩阵 $\boldsymbol{A}+a\boldsymbol{x}\boldsymbol{x}^{\mathrm{T}}$ 正定.

附录 A.1　群　环　域

代数结构是基本的数学结构,而群、环和域又是基本的代数结构.不仅许多数学问题的研究以群、环和域为基本工具,而且这些结构的观点已经渗透到数学的所有分支之中,也渗透到物理,化学、通讯、计算机科学及工程的许多理论和方法之中.群是历史上最早提出的最简单的一种代数结构.在 18 世纪 30 年代,19 岁的法国数学家 E.Galois(1811—1832)用群论的方法建立了代数方程的可解性理论.他从而证明了五次和五次以上的多项式方程的根不能用其系数的有限次四则运算与开方根运算组成的公式来表示.这不仅结束了众多数学家寻求五次以上的多项式方程求根公式的徒劳努力,而且使数学研究进入了从局部性研究转向整体结构研究的新时代.群论自身研究也在其间不断延伸.在有限群中,有些群称为单群.从 1830 年开始,数学家找到一些有限单群无穷族,到 20 世纪 80 年代末,数学家才列出所有单群并证明列举是完全的,其证明长达 15 000 页,有些人认为这是数学史最伟大成就之一.下面通过例子简要地说明群、环和域的定义.

1. 群

首先看两个例子.考察两个自然数的加法,两个加数与其和的奇偶性有下面的规律(表 1):

表 1

+	偶	奇
偶	偶	奇
奇	奇	偶

表 2

$\times$	1	-1
1	1	-1
-1	-1	1

表 1 表示集合{偶,奇}上关于数的加法运算规律:偶数加奇数为奇数,奇数加奇数为偶数.表 2 表示集合$\{1, -1\}$上关于数 1 和 -1 的乘法运算规律.下面的定义是这类运算属性的数学表述.

定义 1　对于集合 S 中任意两个元素 a, b,有一个确定的关系,使它们对应于 S 中唯一的元素 c,记作 $c = \phi(a, b)$,或者简单地用符号 $c = ab$ 记之,这时 ϕ 就称为代数运算(即 $\phi: S\times S\to S$).注意,任意 $a, b\in S$, $\phi(a, b)$还属于 S.

定义 2　给定集合 G,如果在 G 中定义了一个代数运算"·"且满足下面四条性质,则称 G 为群:

(1) 任意 $a, b\in G$, $a\cdot b$(一般记为 ab)为 G 中一个确定的元素(封闭性);

(2) 任意 $a, b, c\in G$,有 $a(bc) = (ab)c$(结合性);

(3) 存在 $e\in G$,对于任意 $a\in G$,有

$$ae = ea = a,$$

称 e 为单位元;

(4) 任给 $a\in G$,存在 $b\in G$,使得 $ab = ba = e$;将 b 记为 a^{-1},称为 a 的逆元.

在群的定义中,只要求所定义的代数运算满足四条性质,至于这个运算用什么符号来表示,可以自由选取.

定理 1　G 为群.则 G 的单位元唯一.

证　设 $e, e'\in G$ 都为 G 的单位元.由单位元定义可知 $e = ee' = e'$,即 G 单位元唯一.□

定理 2　G 为群.对于给定的 $a\in G$,则 a 的逆元 a^{-1} 唯一.

证　设 $b, b'\in G$ 都为 a 的逆元,$e\in G$ 为单位元.由逆元定义可知

$$b = be = bab' = (ba)b' = eb' = b'.\square$$

定义 3　G 为群,A　G,若 A 在 G 的群运算下也为一个群,则称 A 为 G 的子群.

定义 4　G 为群,且对任意 $a, b\in G$,有 $ab = ba$.则 G 称为交换群或 **Abel** 群.

一般,当 G 为 Abel 群时,将运算"·"记为"+",这时 G 的单位元称为零元,而 a 的逆元称为 a 的负元,记为 $-a$.并记 $a + (-b) = a - b$.

例 1　验证表 1 为群.用 e 及 o 分别表示偶数与奇数,在集合 $G = \{o, e\}$ 上定义运算"+"如表 1.验证 G 为一个 Abel 群.

(1) 由于每个格内只有一个元素,故第一条满足.

(2) 任给 $a, b, c\in\{e, o\}$,验证 $a + (b + c) = (a + b) + c$.共有 $2^3 = 8$ 个等式需要验证:

$o + (o + o) = o = (o + o) + o,$

$e + (e + e) = e = (e + e) + e,$

$o + (o + e) = e = (o + o) + e,$

…………

(3) e 为零元,因为

$o + e = o,$

$e + e = e,$

(4) $o + o = e$　$-o = o,$

$e + e = e$　$-e = e.$

故 G 为一群.再由于任给 $a, b\in G$,有 $a + b = b + a$,所以 G 为 Abel 群.

例 2　实数集 $\mathbf{R}$ 在实数加法运算下为 Abel 群.整数集 $\mathbf{N}$ 在整数加法运算下为 $\mathbf{R}$ 的子群.类似地 $(\mathbf{Q}^+, \times)$, $(\mathbf{R}^+, \times)$, $(\mathbf{Q}^*, \times)$, $(\mathbf{R}^*, \times)$ 都为 Abel 群,其中 $\mathbf{Q}^+$ 为所有大于零的有理数构成的集合, $\mathbf{R}^+$ 为所有大于零的实数构成的集合, $\mathbf{Q}^* = \mathbf{Q} \setminus \{0\}$; $\mathbf{R}^* = \mathbf{R} \setminus \{0\}$.

例 3　$\mathbf{Z}_n$ 为 $\mathbf{Z}$ 关于模 n 同余关系 R 的商集,即

$$\mathbf{Z}_n = \mathbf{Z} \quad R = \{\overline{0}, \overline{1}, \cdots, \overline{n-1}\},$$

对任给 瑏,瑏∈ $\mathbf{Z}_n$,在 $\mathbf{Z}_n$ 上定义二元运算　为:

瑏　瑏 $= \overline{a+b}$.

若 瑏$= \overline{a_1}$,瑏$= \overline{b_1}$,则 $a - a_1$ 和 $b - b_1$ 都被 n 整除.因此 $a - a_1 + b - b_1 = a + b - (a_1 + b_1)$ 被 n 整除,即

瑏　瑏 $= \overline{a+b} = \overline{a_1 + b_1} = \overline{a_1} \quad \overline{b_1}$,

这表明　为 $\mathbf{Z}_n$ 上的二元运算.显然它是封闭的,而且满足结合律和交换律,因而 $\mathbf{Z}_n$ 是 Abel 群.$\mathbf{Z}_n$ 关于　的零元 $e = \overline{0}$.因为

瑏　$\overline{n-a} = \overline{n-a}$　瑏 $= \overline{a+n-a} = \overline{n} = \overline{0}$,

所以 $\mathbf{Z}_n$ 中任意元素瑏的负元为 $\overline{n-a}$.$(\mathbf{Z}_n,\ )$ 叫做模 n 剩余类加法群.

2. 环

定义 5　如果在集合 S 中引进了两种代数运算“　”和“·”,通常 $a \quad b$, $a \cdot b$ 简记为 $a + b$, ab,分别称为“加法”和“乘法”.其中 $(S, +)$ 为一个 Abel 群,即“+”满足:

(1) 任给 $a, b, c \in S$, $a + (b + c) = (a + b) + c$;

(2) 任给 $a, b \in S$, $a + b = b + a$;

(3) 存在零元 $o \in S$,对任给 $a \in S$, o 具有性质 $a + o = o + a = a$;

(4) 任给 $a \in S$,存在 $b \in S$,使得 $a + b = b + a = o$; b 称为元素 a 的负元,记作 $-a$.

运算“·”满足:

(5) 任给 $a, b, c \in S$, $a(bc) = (ab)c$;

运算“·”与运算“+”满足:

(6) 任给 $a, b, c \in S$, $a(b + c) = ab + ac$, $(b + c)a = ba + ca$.

这时 $(S, +, \cdot)$ 称为环.

注意:在环的定义中并未要求 $(S, \cdot)$ 为一群.并且运算“·”也不一定满足交换律.

3. 域

定义 6　设 $(S, +, \cdot)$ 为一环,若它还满足:

(1) 任给 $a, b \in S$, $ab = ba$;

(2) 存在 $e\in S$，对任给 $a\in S$，e 具有性质 $ae=ea=a$；

(3) 任给 $a\in S\setminus\{o\}$，存在 $b\in S$，使得 $ab=ba=e$.

即 $(S\setminus\{o\},\cdot)$ 也为一 Abel 群，则 $(S,+,\cdot)$ 称为域.一般地 $(S,+,\cdot)$ 的运算“·”的单位元称为域的单位元，记为 1；运算“+”的单位元称为零元，记为 0.

注意，域是一种特殊的环.可验证全体复数 **C**，全体实数 **R** 和全体有理数 **Q** 关于数的加法和乘法满足定义 6.因此它们都为域，分别称为复数域 **C**，实数域 **R** 和有理数域 **Q**.

定义 7　复数域的子集 K 称为一个数域，如果它满足：

(1) $0,1\in K$；

(2) 任给 $a,b\in K$，都有 $a\pm b\in K$，$ab\in K$，并且当 $b\neq 0$ 时，有 $\frac{a}{b}\in K$.

下面再给出一些环及域的例子.

例 4　设 S 为所有偶数的集合，则 S 关于整数的加法和乘法为一个环.但 S 若为所有奇数的集合，则 S 关于整数加法和乘法不构成环.

例 5　令 $\boldsymbol{Z}(\sqrt{2})=\{m+n\sqrt{2}\mid m,n\in\mathbf{Z}\}$，其中 **Z** 为整数集，则关于实数的加法与乘法构成一个环，但不是域.

证　$\boldsymbol{Z}(\sqrt{2})$ 关于实数的加法，乘法封闭，并且满足定义 5.故 $\boldsymbol{Z}(\sqrt{2})$ 是环.当 $m\in\boldsymbol{Z}(\sqrt{2})\setminus\{0\}$，$m$ 为整数且 $m\geq 2$ 时，$m^{-1}\notin\boldsymbol{Z}(\sqrt{2})$，所以 $\boldsymbol{Z}(\sqrt{2})$ 不是域.□

例 6　令 $\boldsymbol{Q}(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbf{Q}\}$，其中 **Q** 为所有有理数的集合，则关于实数的加法和乘法构成一个域.

证　$\boldsymbol{Q}(\sqrt{2})$ 为环是显然的.任给 $0\neq a+b\sqrt{2}\in\boldsymbol{Q}(\sqrt{2})\setminus\{0\}$，若 $a^2-2b^2=0$，则 $a=\pm b\sqrt{2}$，于是 $\sqrt{2}=\pm\frac{a}{b}$，这与 a,b 为有理数且 $\sqrt{2}$ 为无理数矛盾.所以 $a^2-2b^2\neq 0$，

$$(a+b\sqrt{2})^{-1}=\frac{1}{a+b\sqrt{2}}=\frac{a-b\sqrt{2}}{a^2-2b^2}\in\boldsymbol{Q}(\sqrt{2}).$$

于是 $\boldsymbol{Q}(\sqrt{2})$ 为域.

下面是熟知的代数结构之间的关系：

环 $(\mathbf{Z},+,\times)$　域 $(\mathbf{Q},+,\times)$　域 $(\boldsymbol{Q}(\sqrt{2}),+,\times)$　域 $(\mathbf{R},+,\times)$　域 $(\mathbf{C},+,\times)$.

环 $(\mathbf{Z},+,\times)$　环 $(\boldsymbol{Z}(\sqrt{2}),+,\times)$　域 $(\boldsymbol{Q}(\sqrt{2}),+,\times)$　域 $(\mathbf{R},+,\times)$　域 $(\mathbf{C},+,\times)$.

下面给出关于整数的两个基本的定理.

定理 3　a 是任一整数，b 是一非零整数，则存在唯一一对整数 q 及 r 且 0

$\leq r < |b|$,使得 $a = bq + r$.

证明略

定理 4 对任意两非零整数 a 和 b,则存在整数 r 和 s,使得 $ra + sb = (a, b)$((a, b)为 a 和 b 的最大公因子).a 和 b 互素的充分必要条件为存在整数 r 和 s 使得 $ra + sb = 1$.

证明略.

下面,给出最简单、最常用的有限环和有限域的例子:

例 7 设 n 为任意正整数,$\boldsymbol{Z}_n$ 为 $\boldsymbol{Z}$ 关于模 n 同余关系 R 的商集,即

$$\boldsymbol{Z}_n = \boldsymbol{Z}\ R = \{\bar{0}, \bar{1}, \cdots, \overline{n-1}\},$$

对于 $\bar{a}, \bar{b} \in \boldsymbol{Z}_n$,在 $\boldsymbol{Z}_n$ 上定义“ ”,“·”分别为:

$$\bar{a} \quad \bar{b} = \overline{a+b}, \quad \bar{a} \cdot \bar{b} = \overline{ab}.$$

由例 3 知,$(\boldsymbol{Z}_n, \ )$是一个交换群.为了简便把 记做+.下面证明“·”为一个二元运算.若 $\bar{a} = \bar{a}_1$,$\bar{b} = \bar{b}_1$,则 $a - a_1$ 和 $b - b_1$ 都被 n 整除,即 $a = kn + a_1$,$b = qn + b_1$.于是

$$\bar{a} \cdot \bar{b} = \overline{(kn + a_1)(qn + b_1)} = \overline{kqn^2 + knb_1 + qna_1 + a_1 b_1} = \bar{a}_1 \cdot \bar{b}_1,$$

这表明“·”为 $\boldsymbol{Z}_n$ 上的二元运算.它显然是封闭的,而且满足结合律和交换律,且 $\boldsymbol{Z}_n$ 关于“·”的单位元 $e = \bar{1}$.因为

$$\begin{aligned} \bar{a} \cdot (\bar{b} + \bar{c}) &= \bar{a} \cdot \overline{(b + c)} = \overline{a(b + c)} = \overline{ab + ac} \\ &= \overline{ab} + \overline{ac} = \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}, \end{aligned}$$

所以运算“·”及“+”满足分配律.也易验证“·”满足结合律.因此$(\boldsymbol{Z}_n, +, \cdot)$为环,称此环为整数模 $\boldsymbol{n}$ 的环,记为 $\boldsymbol{Z}_n$.

例 8 当 p 为素数时,$\boldsymbol{Z}_p$ 为域.

证 在例 7 中已验证 $\boldsymbol{Z}_p$ 为环,下面证明:任给 $\bar{i} \in \boldsymbol{Z}_p$ 且 $\bar{i} \neq \bar{0}$,存在 $\bar{j} \in \boldsymbol{Z}_p$,使得 $\bar{i}\bar{j} = \bar{1}$.由 p 为素数及 $\bar{i} \neq \bar{0}$ 知,$(i, p) = 1$.存在整数 u, v,使得 $iu + pv = 1$,于是

$$\bar{1} = \overline{iu} + \overline{pv} = \bar{i} \cdot \bar{u} + \bar{p} \cdot \bar{v} = \bar{i} \cdot \bar{u} + \bar{0}\ \bar{v} = \bar{i} \cdot \bar{u},$$

即 $\bar{i}^{-1} = \bar{u}$,故 $\boldsymbol{Z}_p$ 为域.□

附录 A.2　应用实例——投入产出综合平衡的数学模型

1　价值型投入产出模型

投入产出综合平衡模型,简称投入产出模型,是研究一个经济系统(例如国民经济)在社会再生产过程中各部门之间‘投入”与‘产出”的平衡关系.该模型是经济学家 W.W.Leontief 在 1936 年提出的,他将一个经济系统中各个部门的投入产出,通过表格联系起来,综合分析各部门间的关系,以线性代数为工具,建立了投入产出方法.这一方法得到了广泛应用,Leontief 也因此获得 1973 年 Nobel 经济学奖.

如果一个经济系统各部门投入和产出都用价值(货币数量)来衡量时,投入产出模型称为价值型模型,它与实物型模型构成了投入产出的两个基本模型.我们只讨论价值型投入产出模型.

现在考虑一个经济系统内部各部门在一个生产周期内的生产与分配之间的平衡关系.设该系统有 n 个部门,每个部门既是生产者又是消耗者.设第 i 个部门总的生产价值为 x_i,用 b_i 表示系统外部对第 i 个部门产值的需求量(例如用于消费、积累等等),它们是不参加本周期生产过程的最终产值.用 x_{ij} 表示第 j 个部门在生产过程中所消耗的第 i 个部门产值数.第 i 个部门分配给系统内部各部门(包括自己)生产性消耗总产值(可称中间产品价值)为 $\sum\limits_{j=1}^{n} x_{ij}$,它与供给系统外的产值 b_i 之和应等于总产值 x_i,即

$$\sum_{j=1}^{n} x_{ij} + b_i = x_i\ (i = 1,2,\cdots,n), \tag{1}$$

这个方程组称为分配平衡方程组.

下面考虑系统内各部门的总产值与新创造的价值(净产值)和内部的生产性消耗之间的平衡关系.第 j 个部门在生产过程中消耗本部门及其它部门的产值数为 $\sum\limits_{i=1}^{n} x_{ij}$,第 j 个部门固定资产折旧为 d_j,以上两项总和为第 j 个部门在生产过程中生产性消耗价值(生产资料补偿价值).第 j 个部门的劳动报酬 v_j 及纯收入 m_j 是新创造价值,第 j 个部门生产资料消耗价值与新创造价值之和为其总价值 x_j,即

$$\sum_{i=1}^{n} x_{ij} + d_j + v_j + m_j = x_j\ (j = 1,2,\cdots,n). \tag{2}$$

该方程组称为消耗平衡方程组,方程组(1)及(2)构成了价值型投入产出模型的两个基本方程组.上面讨论的参数关系可列成一个价值型投入产出表 1.

表 1

<table>
<tr><td colspan="3" rowspan="2">产出
投入</td><td>中间产品</td><td rowspan="2">最终产品
(外部需求)</td><td rowspan="2">总产值</td></tr>
<tr><td>消耗部门
1 2 … n</td></tr>
<tr><td rowspan="2">生产资料补偿价值</td><td>生产部门</td><td>1
2
…
n</td><td>x_{11} x_{12} … x_{1n}
x_{21} x_{22} … x_{2n}
… … …
x_{n1} x_{n2} … x_{nn}</td><td>b_1
b_2
…
b_n</td><td>x_1
x_2
…
x_n</td></tr>
<tr><td colspan="2">固定资产折旧</td><td>d_1 d_2 … d_n</td><td></td><td></td></tr>
<tr><td>新创造价值</td><td colspan="2">劳动报酬
社会纯收入</td><td>v_1 v_2 … v_n
m_1 m_2 … m_n</td><td></td><td></td></tr>
<tr><td colspan="3">总产值</td><td>x_1 x_2 … x_n</td><td></td><td></td></tr>
</table>

2 直接消耗系数、分配与消耗平衡方程的解

(1) 用 a_{ij} 表示第 j 个部门生产单位产值需要消耗第 i 个部门的产值数,称为第 j 个部门对第 i 个部门直接消耗系数,于是

$$a_{ij} = \frac{x_{ij}}{x_j}, (i, j = 1, 2, \cdots, n) \tag{3}$$

矩阵 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 称为直接消耗系数矩阵.$\boldsymbol{A}$ 不仅反映各部门间技术经济联系,也反映各部门的技术水平,有时称为技术矩阵.利用投入产出法进行经济分析与预测时,常假设 $\boldsymbol{A}$ 不变或略作调整.

由(3)知

$$x_{ij} = a_{ij} x_j .$$

代入(1),有

$$\sum_{j=1}^{n} a_{ij} x_j + b_i = x_i (i = 1, 2, \cdots, n) .$$

令

$$\boldsymbol{A} = (a_{ij})_{n\times n}, \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}},$$
$$\boldsymbol{b} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}},$$

则分配平衡方程可写为

$$\boldsymbol{Ax} + \boldsymbol{b} = \boldsymbol{x},$$

或

$$(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{b}. \tag{4}$$

(2) 将 $x_{ij} = a_{ij}x_j$ 代入消耗平衡方程组(2),得到

$$\sum_{i=1}^{n} a_{ij}x_j + d_j + v_j + m_j = x_j (j = 1, 2, \cdots, n),$$

令

$$\boldsymbol{D} = \operatorname{diag}\left(\sum_{i=1}^{n} a_{i1}, \sum_{i=1}^{n} a_{i2}, \cdots, \sum_{i=1}^{n} a_{in}\right),$$

$$\boldsymbol{d} = (d_1, d_2, \cdots, d_n)^{\mathrm{T}},$$

$$\boldsymbol{v} = (v_1, v_2, \cdots, v_n)^{\mathrm{T}},$$

$$\boldsymbol{m} = (m_1, m_2, \cdots, m_n)^{\mathrm{T}},$$

$$\boldsymbol{z} = \boldsymbol{d} + \boldsymbol{v} + \boldsymbol{m},$$

则消耗平衡方程可写成

$$\boldsymbol{Dx} + \boldsymbol{z} = \boldsymbol{x},$$

或

$$(\boldsymbol{I} - \boldsymbol{D})\boldsymbol{x} = \boldsymbol{z}. \tag{5}$$

由直接消耗系数 a_{ij} 的定义,可知如下两个性质成立:

$$0 \leq a_{ij} < 1;$$

$$\sum_{i=1}^{n} a_{ij} < 1.$$

根据直接消耗系数 a_{ij} 的这两个性质,可以证明 $\boldsymbol{I} - \boldsymbol{A}$ 及 $\boldsymbol{I} - \boldsymbol{D}$ 可逆.因篇幅限制,证明略去.从(4)及(5),有

$$\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}, \qquad \boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{D})^{-1}\boldsymbol{z}.$$

最后,我们给出第 j 个部门折旧系数、劳动报酬系数及社会收入系数定义依次为

$$\frac{d_j}{x_j}, \frac{v_j}{x_j}, \frac{m_j}{x_j} \quad (j = 1, 2, \cdots, n). \tag{6}$$

3　投入产出方法在编制计划中的应用

投入产出表建立在所研究的各生产部门生产与消耗的实际统计资料基础上,比较客观地反映了各生产部门间生产与需求的综合平衡关系,这样建立的投入产出模型,揭示了各生产部门的经济活动规律,反映了多种经济现象的内在联系.它广泛应用于编制计划,进行经济分析与宏观的经济管理之中.

下面以一个简单例子,说明投入产出模型在编制计划工作中的应用.根据一

个时期(报告期)统计资料得到投入产出表,计算直接消耗系数 .它反映了过去或目前各部门间投入产出关系,但由于一般直接消耗系数在二、三年变化不大,因而报告期直接消耗系数可作为或略作修改后权作计划期的直接消耗系数 .如果先确定了各部门计划期最终产品向量 $\boldsymbol{b}$,就可由最终产品出发安排生产计划 .下面是一个简单的数值例子 .

例 1　设某经济系统包括三个部门:农场、工厂及其它产业,某一年期间的简化投入产出表如下,产值单位:亿元 .

投入 \ 产出		中间产品			最终产品	总产值
		农场(1)	工厂(2)	其它(3)		
生产资料补偿价值	农场(1)	0.20	0.20	0.10	0.70	1.20
	工厂(2)	0.24	0.60	0.36	0.80	2.00
	其它(3)	0.05	0.25	0.40	0.90	1.60
	固定资产折旧	0.08	0.15	0.10		
新创造价值	劳动报酬	0.40	0.40	0.32		
	社会纯收入	0.23	0.40	0.32		
总产值		1.20	2.00	1.60		

如果直接消耗系数矩阵 $\boldsymbol{A}$ 不变,固定资产折旧系数、劳动报酬系数及收入系数不变情况下,若要求农场、工厂及其它产业最终产品产值分别为 1.00,1.00,2.00 亿元,问各部门总产值、中间产品分配情况,固定资产折旧、劳动报酬及收入各多少亿元?

解　a.由投入产出表,利用公式(3),可求出直接消耗系数矩阵

$$\boldsymbol{A} = \begin{pmatrix} 0.1667 & 0.1000 & 0.0625 \\ 0.2000 & 0.3000 & 0.2250 \\ 0.0417 & 0.1250 & 0.2500 \end{pmatrix}.$$

b.由分配平衡方程组(4)及 $\boldsymbol{b} = (1.00, 1.00, 2.00)^{\mathrm{T}}$ 知

$$\boldsymbol{x} = (I - A)^{-1}\boldsymbol{b} = \begin{pmatrix} 1.2588 & 0.2098 & 0.1678 \\ 0.4038 & 1.5767 & 0.5067 \\ 0.1373 & 0.2745 & 1.4271 \end{pmatrix} \begin{pmatrix} 1.0000 \\ 1.0000 \\ 2.0000 \end{pmatrix} = \begin{pmatrix} 1.8043 \\ 2.9939 \\ 3.2660 \end{pmatrix},$$

于是,农场、工厂及其它产业总产值分别为 $x_1 = 1.8043$(亿元),$x_2 = 2.9939$(亿元),$x_3 = 3.2660$(亿元).

c.由公式(3),有

$$x_{ij} = a_{ij}x_j,$$

中间产品分配可写成矩阵形式:

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}$$

$$= \boldsymbol{A} \begin{pmatrix} 1.8043 & 0 & 0 \\ 0 & 2.9939 & 0 \\ 0 & 0 & 3.2660 \end{pmatrix} = \begin{pmatrix} 0.3008 & 0.2994 & 0.2041 \\ 0.3609 & 0.8982 & 0.7348 \\ 0.0752 & 0.3742 & 0.8165 \end{pmatrix}.$$

d. **由(6)知**

	农场(1)	工厂(2)	其它部门(3)	
折旧系数	$\frac{0.08}{1.20}$	$\frac{0.15}{2.00}$	$\frac{0.10}{1.60}$	$= B$
劳动报酬系数	$\frac{0.40}{1.20}$	$\frac{0.40}{2.00}$	$\frac{0.32}{1.60}$	
收入系数	$\frac{0.23}{1.20}$	$\frac{0.40}{2.00}$	$\frac{0.32}{1.60}$	
总产值	1.804 7	2.994 0	3.266 1	

将上表前三行的数构成的矩阵记为 B,**则各部门固定资产折旧、劳动报酬及收入可写成如下形式:**

$$\begin{pmatrix} d_1 & d_2 & d_3 \\ v_1 & v_2 & v_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \boldsymbol{B} \begin{pmatrix} 1.8047 & 0 & 0 \\ 0 & 2.9940 & 0 \\ 0 & 0 & 3.2661 \end{pmatrix} = \begin{pmatrix} 0.1203 & 0.2245 & 0.2041 \\ 0.6014 & 0.5988 & 0.6532 \\ 0.3459 & 0.5988 & 0.6532 \end{pmatrix}.$$

附录 A .3　Jordan 标准形

Jordan 标准形在微分方程、力学及控制论等方面有重要应用 .因篇幅所限,本书不拟详细讨论此问题,仅给出将矩阵化为 Jordan 标准形的一种算法 .

定理 6 .7 已证明, n 阶方阵 $\boldsymbol{A}$ 与对角形相似的充要条件:对应 $\boldsymbol{A}$ 的每个 k_i 重特征值 λ_i,特征矩阵 $\lambda_i \boldsymbol{I} - \boldsymbol{A}$ 的秩为 $n - k_i$,也即对于 k_i 重特征值 λ_i 有 k_i 个线性无关的特征向量 .如果矩阵 $\boldsymbol{A}$ 不满足这个条件,就不能与对角形相似,也即不能对角化 .但 $\boldsymbol{A}$ 可与 Jordan 标准形相似,即存在可逆矩阵 $\boldsymbol{P}$,使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_l \end{pmatrix}, \text{其中 } \boldsymbol{J}_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \tag{1}$$

λ_i, $i = 1, 2, \cdots, l$,为 $\boldsymbol{A}$ 的特征值 .第一等式右端矩阵称为 $\boldsymbol{A}$ 的 **Jordan** 标准形,其中每一个 $\boldsymbol{J}_i$ 称为对应于 λ_i 的 **Jordan** 块 .当每个 Jordan 块都为 1 阶时, Jordan 标准形为对角形 .(熊全淹,叶明川 线性代数(第 3 版) .北京:高等教育出版社, 1985) .

下面,用一个例子说明将一个矩阵化为与其相似的 Jordan 标准形的算法的思路 .

首先,假设已将 6 阶矩阵 $\boldsymbol{A}$ 化为 Jordan 标准形 .然后分析矩阵 $\boldsymbol{A}$ 与它的特征值,特征向量及矩阵 $\boldsymbol{P}$ 的列向量之间的关系,以便理出算法思路(注意下述思路逻辑) .

令 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_6)$,其中 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_6$ 为 $\boldsymbol{P}$ 的列向量 .设可逆矩阵 $\boldsymbol{P}$ 将矩阵 $\boldsymbol{A}$ 化为 Jordan 标准形,且设 $\boldsymbol{A}$ 的 Jordan 标准形是式(2)右端的第二个矩阵,由式(1)可知

$$\boldsymbol{A}(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_6) = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_6)\left(\begin{array}{ccc|cc|c} \lambda_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{array}\right) \tag{2}$$

其中 λ_1, λ_2 为 A 的特征值, λ_1 为 A 的 5 重特征值.由式(2)知, p_1, p_2, …, p_6 满足如下等式

$$\begin{aligned} &(A-\lambda_1 I)p_1 = \mathbf{0}, (A-\lambda_1 I)p_2 = P_1, (A-\lambda_1 I)p_3 = p_2; \\ &(A-\lambda_1 I)p_4 = \mathbf{0}, (A-\lambda_1 I)p_5 = p_4; \\ &(A-\lambda_2 I)p_6 = \mathbf{0}. \end{aligned} \tag{3}$$

因为 p_1, p_2, …, p_6 为可逆矩阵 P 的列向量, 它们都是非零向量且线性无关.由式(3)知

$$\begin{aligned} &(A-\lambda_1 I)^3 p_3 = \mathbf{0}, (A-\lambda_1 I)^2 p_3 \neq \mathbf{0}, \\ &(A-\lambda_1 I)^2 p_2 = \mathbf{0}, (A-\lambda_1 I)p_2 \neq \mathbf{0}, p_2 = (A-\lambda_1 I)p_3, \\ &(A-\lambda_1 I)p_1 = \mathbf{0}, p_1 = (A-\lambda_1 I)p_2, \\ &(A-\lambda_1 I)^2 p_5 = \mathbf{0}, (A-\lambda_1 I)p_5 \neq \mathbf{0}, \\ &(A-\lambda_1 I)p_4 = \mathbf{0}, p_4 = (A-\lambda_1 I)p_5. \end{aligned} \tag{4}$$

如果能找到满足上面这些等式的 p_1, p_2, …, p_6, 且它们线性无关, 则它们一定满足式(2), 于是, $P=(p_1, p_2, \cdots, p_6)$就可把 A 化为 Jordan 标准形.为此引入下面的定义.

定义 1　一个向量 v 称为矩阵 A 的特征值 λ 的 k 阶广义特征向量, 当且仅当

$$(A-\lambda I)^k v = \mathbf{0} \text{ 且 } (A-\lambda I)^{k-1} v \neq \mathbf{0},$$

其中 k 为某一正整数.

注意, 当 $k=1$ 时, 广义特征向量即为特征向量.

设 v 是矩阵 A 的特征值 λ 的 k 阶广义特征向量.定义 v_k, v_{k-1}, …, v_1 如下:

$$\begin{aligned} v_k &= v, \\ v_{k-1} &= (A-\lambda I)v = (A-\lambda I)v_k, \\ v_{k-2} &= (A-\lambda I)^2 v = (A-\lambda I)v_{k-1}, \\ &\cdots\cdots \\ v_1 &= (A-\lambda I)^{k-1} v = (A-\lambda I)v_2. \end{aligned} \tag{5}$$

向量组 v_1, v_2, …, v_k 称为矩阵 A 的特征值 λ 的长度为 k 的一个广义特征向量链, 记为$\{v_1, v_2, \cdots, v_k\}$.

关于广义特征向量, 有如下两个定理(证明略):

定理 1　对 A 的每个 k_i 重特征值 λ_i, 存在 k_i 个线性无关的对应特征值 λ_i 的广义特征向量(可能是不同阶).

定理 2　不同特征值的广义特征向量线性无关.

下面用式(2)表示的例子来说明算法.

将$(A-\lambda_1 I)^i v=0$的解空间记为 N_i,其维数为 $d_i(i=1,2,3)$,由式(3)可知

(1) $N_3=[p_3,p_2,p_1,p_5,p_4]$, $d_3=6-1=5$;

$N_2=[\quad p_2,p_1,p_5,p_4]$, $d_2=6-2=4$;

$N_1=[\quad\quad p_1,\quad p_4]$, $d_1=6-4=2$.

(2) $\{p_1,p_2,p_3\}$是 A对应λ_1 的长度为 3 的广义特征向量链.$\{p_4,p_5\}$是 A 对应λ_1 的长度为 2 的广义特征向量链.

从广义特征向量链$\{p_1,p_2,p_3\}$中三个向量 p_1,p_2,p_3 之间关系,以及它们与另一个广义特征向量链之间关系,可得到相应于 λ_1 的广义特征向量链及 Jordan 矩阵的算法:

(Ⅰ) 计算$(A-\lambda_1 I)^i$ 及其秩 R_i,使 $i=1,2,\cdots$,直至

$$R_k = 6(A\text{ 的阶数}) - 5(\lambda_1\text{ 的重数}) = 1.$$

上例中 $k=3$,且可知 $d_3=5$, $d_2=4$, $d_1=2$.

(Ⅱ) 在 N_3 中找到 1 个线性无关解向量,要求它不在 N_2 中,将它记 p_3.令 $p_2=(A-\lambda_1 I)p_3$, $p_1=(A-\lambda_1 I)p_2$,则 $p_2\in N_2$, $p_1\in N_1$, $\{p_1,p_2,p_3\}$为 A 的对应λ_1 的一个长度为 3 的广义特征向量链.

(Ⅲ) 因为 $d_3-d_2=1$,在 N_3 中再也找不到与 p_3 线性无关的且不在 N_2 中的向量,转到下一步.如果 $d_3-d_2>1$,仍重复(2)的步骤.

(Ⅳ) 在 N_2 中找到一个与 p_2 线性无关且不在 N_1 中的解向量.应找的解向量个数$=d_2-d_1-l_2$(在已形成的广义特征向量链内的 N_2 的解向量的个数)$=4-2-1=1$.令其为 p_5,再令 $p_4=(A-\lambda_1 I)p_5$, $p_5\in N_1$, $\{p_4,p_5\}$形成 A 的对应λ_1 的第二个长度为 2 的广义特征向量链.

在 N_1 中找到线性无关的解向量,使解向量个数$=d_1-l_1$(在已形成的广义特征向量链内的 N_1 的解向量的个数),但现在解向量个数$=2-2=0$,停止.

(Ⅴ) 这时,p_1,p_2,p_3,p_4,p_5 在矩阵 P 中位置及相应的 Jordan 块如式(2).

下面是一个具体例子.

例 1 把下列矩阵 A 化为 Jordan 标准形:

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

解 第一步:计算 A 的特征值.

$$|A-\lambda I| = [(3-\lambda)(1-\lambda)+1](\lambda-2)^2[(1-\lambda)^2-1] = (\lambda-2)^5\lambda.$$

所以 $\boldsymbol{A}$ 有 5 重特征值 2, 1 重特征值 0 .这时 $\boldsymbol{A}$ 的阶 $n=6$, 特征值 2 的重数 $m=5$.

第二步:计算 $(\boldsymbol{A}-2\boldsymbol{I})^i, (i=1,2,\cdots)$ 的秩直到 $R((\boldsymbol{A}-\lambda\boldsymbol{I})^k)=n-m=6-5=1$.

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

$R((\boldsymbol{A}-2\boldsymbol{I}))=4$, 解空间 N_1 的维数 $d_1=6-4=2$.因此, $\boldsymbol{A}$ 的特征值 2 的特征向量只能有 $6-R((\boldsymbol{A}-2\boldsymbol{I}))=2$ 个线性无关的特征向量, 而 $\boldsymbol{A}$ 有 5 重特征值 2, 因此 $\boldsymbol{A}$ 不能对角化 .继续计算 $(\boldsymbol{A}-2\boldsymbol{I})^i (i=1,2,\cdots)$ 的秩 .

$$(\boldsymbol{A}-2\boldsymbol{I})^2=\begin{pmatrix}0&0&2&2&0&0\\0&0&2&2&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&2&-2\\0&0&0&0&-2&2\end{pmatrix},$$

$R((\boldsymbol{A}-2\boldsymbol{I})^2)=2$, 解空间 N_2 的维数 $d_2=6-2=4$.

$$(\boldsymbol{A}-2\boldsymbol{I})^3=\begin{pmatrix}0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&-4&4\\0&0&0&0&4&-4\end{pmatrix},$$

$R((\boldsymbol{A}-2\boldsymbol{I})^3)=1$, 解空间 N_3 的维数 $d_3=6-1=5$.因为 $R((\boldsymbol{A}-2\boldsymbol{I})^3)=1=6-5=n-m$, 所以在此停止 .由于 $d_3-d_2=1$, 能找到一个 3 阶的广义特征向量 $\boldsymbol{u}$ 满足 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{u}=\boldsymbol{0}$ 且 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{u}\neq\boldsymbol{0}$.易验证 $\boldsymbol{u}=(0,0,1,0,0,0)^{\mathrm{T}}$ 就是这样的向量 .定义

$$\boldsymbol{u}_1=(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{u}=(2,2,0,0,0,0)^{\mathrm{T}},$$

$$\boldsymbol{u}_2=(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{u}=(1,-1,0,0,0,0)^{\mathrm{T}}, \boldsymbol{u}_3=\boldsymbol{u} .$$

这是一个长度为 3 的广义特征向量链 .因为 $d_2-d_1=2$, 所以存在两个线性无关的向量在 N_2 中而不在 N_1 中, $\boldsymbol{u}_2$ 是其中一个 .因此能找到一个 $\boldsymbol{v}$ 与 $\boldsymbol{u}_2$ 线性无

并且满足$(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{v}=\boldsymbol{0}$及$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{v}\neq\boldsymbol{0}$.易验证 $\boldsymbol{v}=(0,0,1,-1,1,1)^{\mathrm{T}}$就是这样的向量.定义

$$\boldsymbol{v}_1=(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{v}=(0,0,2,-2,0,0)^{\mathrm{T}},\ \boldsymbol{v}_2=\boldsymbol{v}.$$

现在已找到 $\boldsymbol{A}$ 的关于特征值 2 的 5 个广义特征向量.很容易验证以上 5 个广义特征向量是线性无关的.

第三步:计算 $\boldsymbol{A}$ 的关于特征值 0 的特征向量.易验证 $\boldsymbol{w}=(0,0,0,0,1,-1)^{\mathrm{T}}$是 $\boldsymbol{A}$ 的关于特征值 0 的一个特征向量.

第四步:令

$$\boldsymbol{P}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{u}_3,\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{w})=\begin{pmatrix}2&1&0&0&0&0\\2&-1&0&0&0&0\\0&0&1&2&1&0\\0&0&0&-2&-1&0\\0&0&0&0&1&1\\0&0&0&0&1&-1\end{pmatrix}.$$

由前面的算法分析知

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

例 2 将下列矩阵 $\boldsymbol{A}$ 化为 Jordan 标准形:

$$\boldsymbol{A}=\begin{pmatrix}1&1&2\\0&1&3\\0&0&2\end{pmatrix}.$$

解 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-1)^2(\lambda-2),\lambda_1=\lambda_2=1,\lambda_3=2$.因为

$$\lambda_1\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}0&-1&-2\\0&0&-3\\0&0&-2\end{pmatrix},\ \lambda_3\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}1&-1&-2\\0&1&-3\\0&0&0\end{pmatrix}$$

的秩都为 2,所以$(\lambda_1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$的基础解系由一个向量构成.因此不存在三个线性无关的特征向量.这样 $\boldsymbol{A}$ 就不能对角化.下面用上面给出的方法求一个可逆矩阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 化为 Jordan 标准形.首先计算$(\boldsymbol{A}-\lambda_1\boldsymbol{I})^i$ 的秩,$i=1,2$.矩阵 $\boldsymbol{A}$ 的阶 $n=3$,特征值 $\lambda_1=1$ 的重数 $m=2$.

$$A - \lambda_1 I = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix}, R(A - \lambda_1 I) = 2 \geq n - m = 1 .$$

$$(A - \lambda_1 I)^2 = \begin{pmatrix} 0 & 0 & 7 \\ 0 & 0 & 6 \\ 0 & 0 & 4 \end{pmatrix}, R((A - \lambda_1 I)^2) = 1 = n - m .$$

寻找一个向量 $\boldsymbol{v}$ 满足 $(A - \lambda_1 I)^2 \boldsymbol{v} = O$ 而 $(A - \lambda_1 I)\boldsymbol{v} \neq O$.显然 $\boldsymbol{v} = (0,1,0)^{\mathrm{T}}$ 就是这样的向量 .$\boldsymbol{v}$ 是 A 的 2 阶广义特征向量 .令

$$\boldsymbol{v}_2 = \boldsymbol{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_1 = (A - \lambda_1 I)\boldsymbol{v} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} .$$

对应于特征值 $\lambda_3 = 2$ 的特征向量为 $\boldsymbol{v}_3 = (5,3,1)^{\mathrm{T}}$.令

$$P = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

可得 A 的 Jordan 标准形如下

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

通过上面例子,不难把例子中使用的算法一般化 .

设 A 是一个 n 阶矩阵并且 λ 为 A 的 m 重的特征值 .上面定理 1、定理 2 保证能找到 m 个线性无关的 λ 的广义特征向量 .这将通过寻找不同长度 λ 的广义特征向量链来实现 .注意, $(A - \lambda I)^i$ 的秩随着指数 i 的增加而递减 .计算 $(A - \lambda I)^i (i = 1,2,\cdots)$ 的秩直到 $R((A - \lambda I)^k) = n - m$.设 d_i 为解空间 N_i 的维数 .由齐次方程组理论知 $d_i = n - R((A - \lambda I)^i)$.令 N_i 为线性方程 $(A - \lambda I)^i \boldsymbol{x} = \boldsymbol{0}$ 的解空间 $i = 1,2,\cdots$.如果 $\alpha \in N_i$,显然有 $\alpha \in N_{i+1}$.即 N_i 为 N_{i+1} 的子空间 .这样 $N_0 \subseteq N_1 \subseteq N_2 \subseteq \cdots \subseteq N_k$,所以有 $0 = d_0 \leq d_1 \leq d_2 \cdots \leq d_k = m$.$d_i - d_{i-1}$ 为在 N_i 中而不在 N_{i-1} 中的线性无关的广义特征向量个数 .设 $\boldsymbol{v}$ 为定义 1 中的矩阵 A 的特征值 λ 的 k 阶广义特征向量,于是 $\boldsymbol{v} \in N_k$,且 $\boldsymbol{v} \notin N_{k-1}$,因而

$$(A - \lambda I)^i \boldsymbol{v}_i = (A - \lambda I)^i (A - \lambda I)^{k-i} \boldsymbol{v} = (A - \lambda I)^k \boldsymbol{v} = \boldsymbol{0},$$

$$(A - \lambda I)^{i-1} \boldsymbol{v}_i = (A - \lambda I)^{i-1} (A - \lambda I)^{k-i} \boldsymbol{v} = (A - \lambda I)^{k-1} \boldsymbol{v} \neq \boldsymbol{0} .$$

因此 $\boldsymbol{v}_i \in N_i$, $\boldsymbol{v}_i \notin N_{i-1}$.

如何寻找矩阵 A 的相应 m 重特征值 λ 的所有不同长度的广义特征向量链呢 ?

(1) 在集合 $N_k \setminus N_{k-1}$ 中,寻找 $d_k - d_{k-1}$ 个彼此线性无关的 k 阶广义特征向量.例如 $\boldsymbol{v}_k$ 为其中一个广义特征向量,令 $\boldsymbol{v}_k = \boldsymbol{v}$,用公式(5)可产生长度为 k 的一个广义特征向量链.顺序地对每个 k 阶的广义特征向量都这样做,都产生一个长度为 k 的广义特征向量链,要求这些广义特征向量与所有在它之前产生的广义特征向量链中在同一个空间 $N_i(i = k, k-1, \cdots, 1)$ 的所有广义特征向量都线性无关.

(2) 令 l_{k-1} 是已产生的广义特征向量链中在 N_{k-1} 中的广义特征向量的个数.在 $N_{k-1} \setminus N_{k-2}$ 中寻找 $d_{k-1} - d_{k-2} - l_{k-1}$ 个彼此线性无关的 $k-1$ 阶广义特征向量,并要求它们与已产生的广义特征向量链中在 N_{k-1} 中的广义特征向量也线性无关.例如 $\boldsymbol{u}_{k-1}$ 是其中一个 $k-1$ 阶广义特征向量,令 $\boldsymbol{u}_{k-1} = \boldsymbol{v}$,利用公式(5)(公式(5)中的 k 换为 $k-1$)产生长度为 $k-1$ 的一个广义特征向量链.顺序地对每个 $k-1$ 阶广义特征向量都这样做,都产生一个长度为 $k-1$ 阶的广义特征向量链,要求这些广义特征向量与所有在它之前产生的广义特征向量链中在同一个空间 $N_i(i = k-1, k-2, \cdots, 1)$ 的所有广义特征向量都线性无关.

这样继续下去,直到产生出 $\boldsymbol{A}$ 的相应于 m 重特征值 λ 的所有 m 个广义特征向量,如公式(2)一样,每个 k 长度的广义特征向量链都对应公式(2)中矩阵 $\boldsymbol{P}$ 的 k 个列向量,也对应 Jordan 标准形中一个 k 阶 Jordan 块.

附录 B.1　定理 3.3 的证明

定理 3.3　n 阶行列式

$$\Delta(a_{ij}) = \sum_{i=1}^{n}\sum_{k=1}^{i-1}(a_{1i}a_{2k} - a_{1k}a_{2i})(-1)^{i+k}(M_{1i})_{1k}$$
$$= \sum_{i=1}^{n}\sum_{k=1}^{i-1}\begin{vmatrix} a_{1k} & a_{1i} \\ a_{2k} & a_{2i} \end{vmatrix}(-1)^{i+k+1}(M_{1i})_{1k},$$

其中$(M_{1k})_{1j}$是从行列式 M_{1k}划去第 1 行与第 j 列之后得到的 $n-2$ 阶行列式，$i=1,2,\cdots,n$, $k=1,2,\cdots,n-1$, $j=1,2,\cdots,n-1$.

证　n 阶行列式 $\Delta(a_{ij})$

$$\Delta(a_{ij}) = a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{1+n}a_{1n}M_{1n} = \sum_{i=1}^{n}(-1)^{1+i}a_{1i}M_{1i}. \tag{1}$$

对于(1)中的 $n-1$ 阶行列式 M_{1i}，再用递推公式有：

$$M_{1i} = \sum_{k=1}^{i-1}(-1)^{1+k}a_{2k}(M_{1i})_{1k} + \sum_{k=i}^{n-1}(-1)^{1+k}a_{2k+1}(M_{1i})_{1k},\ i = 1,2,\cdots,n.$$

代入(1)有

$$\Delta(a_{ij}) = \sum_{i=1}^{n}\sum_{k=1}^{i-1}(-1)^{i+k}a_{1i}a_{2k}(M_{1i})_{1k} + \sum_{i=1}^{n}\sum_{k=i}^{n-1}(-1)^{i+k}a_{1i}a_{2k+1}(M_{1i})_{1k}. \tag{2}$$

对任意 u_k，规定 $\sum_{k=1}^{0}u_k = 0$. 由 $(M_{1i})_{ij}$ 的定义知，当 $f \leq g$ 时，$(M_{1f})_{1g} = (M_{1g+1})_{1f}$. 因此有

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1}(-1)^{i+k}a_{1i}a_{2k+1}(M_{1i})_{1k} = \sum_{i=1}^{n}\sum_{k=i}^{n-1}(-1)^{i+k}a_{1i}a_{2k+1}(M_{1k+1})_{1i}$$
$$= \sum_{i=1}^{n}\sum_{k=i+1}^{n}(-1)^{i+k-1}a_{1i}a_{2k}(M_{1k})_{1i}.$$

由于标号 i, k 仅仅是记号，所以对换 i, k 后，原式的值不变.

$$\sum_{i=1}^{n}\sum_{k=i}^{n-1}(-1)^{i+k}a_{1i}a_{2k+1}(M_{1i})_{1k} = \sum_{k=1}^{n}\sum_{i=k+1}^{n}(-1)^{i+k-1}a_{1k}a_{2i}(M_{1i})_{1k}.$$

再考察二元数组(k, i). 右式的双重和等于加遍下列的二元数组(k, i)，与加的先后顺序无关.

$$\begin{array}{cccccc} (1,n) & (2,n) & (3,n) & \cdots & (n-1,n) & (n,n) \\ (1,n-1) & (2,n-1) & (3,n-1) & \cdots & (n-2,n-2) & \\ \cdots & \cdots & \cdots & \cdots & & \\ (1,3) & (2,3) & & & & \\ (1,2) & & & & & \end{array}$$

因此,若一行一行地加则有

$$\sum_{k=1}^{n}\sum_{i=k+1}^{n}(-1)^{i+k-1}a_{1k}a_{2i}(M_{1i})_{1k} = \sum_{i=2}^{n}\sum_{k=1}^{i-1}(-1)^{i+k-1}a_{1k}a_{2i}(M_{1i})_{1k}$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{i-1}(-1)^{i+k-1}a_{1k}a_{2i}(M_{1i})_{1k}.$$

因此

$$\begin{aligned}\Delta(a_{ij}) &= \sum_{i=1}^{n}\sum_{k=1}^{i-1}(-1)^{i+k}a_{1i}a_{2k}(M_{1i})_{1k} + \sum_{i=1}^{n}\sum_{k=i}^{n-1}(-1)^{i+k}a_{1i}a_{2k+1}(M_{1i})_{1k} \\ &= \sum_{i=1}^{n}\sum_{k=1}^{i-1}(-1)^{i+k}a_{1i}a_{2k}(M_{1i})_{1k} + \sum_{i=1}^{n}\sum_{k=1}^{i-1}(-1)^{i+k-1}a_{1k}a_{2i}(M_{1i})_{1k} \\ &= \sum_{i=1}^{n}\sum_{k=1}^{i-1}[(-1)^{i+k}a_{1i}a_{2k}(M_{1i})_{1k} + (-1)^{k+(i-1)}a_{1k}a_{2i}(M_{1i})_{1k}] \\ &= \sum_{i=1}^{n}\sum_{k=1}^{i-1}(a_{1i}a_{2k} - a_{1k}a_{2i})(-1)^{i+k}(M_{1i})_{1k} \\ &= \sum_{i=1}^{n}\sum_{k=1}^{i-1}\begin{vmatrix} a_{1k} & a_{1i} \\ a_{2k} & a_{2i}\end{vmatrix}(-1)^{i+k+1}(M_{1i})_{1k}. \square\end{aligned}$$

附录 B.2　定理 6.8 的证明

定理 6.8　设 λ_0 是 n 阶方阵 $\boldsymbol{A}$ 的一个 k 重特征值,则对应于 λ_0 的线性无关的特征向量最大个数 l 小于等于 k.

证　用反证法,设 $l>k$,即 $\boldsymbol{A}$ 对 λ_0 存在 l 个线性无关的特征向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l$,把它扩充为数域 R 上 n 维线性空间 V_n 的一个基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n$.令

$$\boldsymbol{P}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n),$$

则 $\boldsymbol{P}$ 可逆.于是 $\boldsymbol{AP}=\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n)$.因为 $\boldsymbol{Ax}_i=\lambda_0\boldsymbol{x}_i\ (i=1,2,\cdots,l)$,所以

$$\boldsymbol{Ax}_i=\lambda_0\boldsymbol{x}_i=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{pmatrix}0\\ \vdots\\ \lambda_0\\ \vdots\\ 0\end{pmatrix}\begin{array}{l}\\ \\ \cdots i\text{ 列}\\ \\ \\ \end{array}\qquad(i=1,2,\cdots,l).$$

又因为 $\boldsymbol{Ax}_j\in V_n\ (j=l+1,\cdots,n)$,所以 $\boldsymbol{Ax}_j$ 可写为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 的线性组合,即

$$\boldsymbol{Ax}_j=b_{1j}\boldsymbol{x}_1+b_{2j}\boldsymbol{x}_2+\cdots+b_{nj}\boldsymbol{x}_n=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{pmatrix}b_{1j}\\ b_{2j}\\ \vdots\\ b_{nj}\end{pmatrix}.$$

这样,

$$\begin{aligned}&\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n)\\ &=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n)\left(\begin{array}{ccc|ccc}\lambda_0&&&b_{1,l+1}&\cdots&b_{1n}\\ &\ddots&&&&\\ &&\lambda_0&b_{l,l+1}&\cdots&b_{ln}\\ \hline &&&b_{l+1,l+1}&\cdots&b_{l+1,n}\\ &\boldsymbol{O}&&&&\\ &&&b_{n,l+1}&\cdots&b_{nn}\end{array}\right)\\ &=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n)\left(\begin{array}{ccc|c}\lambda_0&&&\\ &\ddots&&\boldsymbol{B}_1\\ &&\lambda_0&\\ \hline &\boldsymbol{O}&&\boldsymbol{B}_2\end{array}\right).\end{aligned}$$

其中 $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ 表示相应的矩阵,上式可简记为

$$\boldsymbol{AP} = \boldsymbol{P}\begin{pmatrix} \lambda_0 \boldsymbol{I}_l & \boldsymbol{B}_1 \\ \boldsymbol{O} & \boldsymbol{B}_2 \end{pmatrix}.$$

由于 $\boldsymbol{P}$ 可逆,有

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} \lambda_0 \boldsymbol{I}_l & \boldsymbol{B}_1 \\ \boldsymbol{O} & \boldsymbol{B}_2 \end{pmatrix}. \tag{1}$$

因为相似矩阵的特征多项式相同,故

$$|\lambda \boldsymbol{I}_n - \boldsymbol{A}| = |\lambda \boldsymbol{I}_n - \boldsymbol{P}^{-1}\boldsymbol{AP}| = \begin{vmatrix} \lambda \boldsymbol{I}_l - \lambda_0 \boldsymbol{I}_l & -\boldsymbol{B}_1 \\ \boldsymbol{O} & \lambda \boldsymbol{I}_{n-l} - \boldsymbol{B}_2 \end{vmatrix}$$

$$= |(\lambda - \lambda_0)\boldsymbol{I}_l|\,|\lambda \boldsymbol{I}_{n-l} - \boldsymbol{B}_2| = (\lambda - \lambda_0)^l f_1(\lambda),$$

其中 $f_1(\lambda) = |\lambda \boldsymbol{I}_{n-l} - \boldsymbol{B}_2|$ 是 λ 的 $n-l$ 次多项式.这样 λ_0 是 $|\lambda \boldsymbol{I}_n - \boldsymbol{A}| = 0$ 至少 l 重根.也即是 $\boldsymbol{A}$ 的 l 重特征值,但 $l > k$,这与 λ_0 是 $\boldsymbol{A}$ 的一个 k 重特征值矛盾.所以 $l \leq k$.□

附录 B.3　定理 7.3(惯性定理)的证明

定理 7.3(惯性定理)　若 n 元二次型 $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 中的矩阵 $\boldsymbol{A}$ 的秩为 r,对任意满秩线性变换 $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$,使

$$f = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{y} = b_1 y_1^2 + \cdots + b_q y_q^2 + \cdots + b_p y_p^2 - b_{p+1} y_{p+1}^2 - \cdots - b_r y_r^2,$$

则正项个数 p 与负项个数 $r - p$ 都是唯一确定的.

证　因为 $\boldsymbol{Q}$ 为满秩矩阵,所以 $\boldsymbol{A}$ 与

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}(b_1, \cdots, b_q, \cdots, b_p, -b_{p+1}, \cdots, -b_r, 0, \cdots, 0)$$

的秩相同,它们的秩都是 r,这是由矩阵 $\boldsymbol{A}$ 唯一确定的.下面证明 p 是由 $\boldsymbol{A}$ 唯一确定的.无损一般性,设 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经过两个不同的满秩线性变换 $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$, $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{z}$,分别变为

$$f = b_1 y_1^2 + \cdots + b_q y_q^2 \boxed{+ b_{q+1} y_{q+1}^2 + \cdots + b_p y_p^2} - b_{p+1} y_{p+1}^2 - \cdots - b_r y_r^2, \quad (1)$$

$$f = c_1 z_1^2 + \cdots + c_q z_q^2 \boxed{- c_{q+1} z_{q+1}^2 - \cdots - c_p z_p^2} - c_{p+1} z_{p+1}^2 - \cdots - c_r z_r^2. \quad (2)$$

其中 $b_i, c_i > 0\ (i = 1, 2, \cdots, r)$,且 $p > q$.由于 $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ 及 $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{z}$,可知 $\boldsymbol{z} = \boldsymbol{S}^{-1}\boldsymbol{Q}\boldsymbol{y}$,令 $\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{Q}$,则 $\boldsymbol{z} = \boldsymbol{D}\boldsymbol{y}$,其中 $\boldsymbol{D}$ 也是满秩矩阵.令 $\boldsymbol{D} = (d_{ij})_{n\times n}$.适当选取 $\boldsymbol{y}$ 与 $\boldsymbol{z}$ 的值,能使式(1)中 $f > 0$,而式(2)中 $f \leqslant 0$,导致矛盾,从而证明 $P = q$.令 $z_1 = z_2 = \cdots = z_q = 0$, $y_{p+1} = \cdots = y_n = 0$,即

$$\left.\begin{array}{ll}
\left.\begin{array}{l}
z_1 = d_{11} y_1 + \cdots + d_{1p} y_p + d_{1,p+1} y_{p+1} + \cdots + d_{1n} y_n = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
z_q = d_{q1} y_1 + \cdots + d_{qp} y_p + d_{q,p+1} y_{p+1} + \cdots + d_{qn} y_n = 0
\end{array}\right\} & q\text{ 个方程} \\
\left.\begin{array}{r}
y_{p+1} = 0 \\
\cdots\cdots\cdots\cdots \\
y_n = 0
\end{array}\right\} & n - p\text{ 个方程}
\end{array}\right\}, \quad (3)$$

把(3)看成关于 n 个未知数 $y_1, y_2, \cdots, y_n$ 的线性方程组,其中方程个数为 $n - p + q = n - (p - q) < n$,因而有非零解 $\boldsymbol{y} = (y_1, \cdots, y_p, y_{p+1}, \cdots, y_n)$.但已令 $y_{p+1} = \cdots = y_n = 0$,所以 $\boldsymbol{y} = (y_1, \cdots, y_p, 0, 0, \cdots, 0)^{\mathrm{T}}$,其中 $y_1, \cdots, y_p$ 不全为零.因为 $\boldsymbol{z} = \boldsymbol{D}\boldsymbol{y}$ 且 $\boldsymbol{D}$ 为满秩矩阵,所以 $\boldsymbol{z}$ 也是非零向量.因为已令 $z_1 = z_2 = \cdots = z_q = 0$,所以 $\boldsymbol{z} = (0, \cdots, 0, z_{q+1}, \cdots, z_r, z_{r+1}, \cdots, z_n)^{\mathrm{T}}$,其中 $z_{q+1}, \cdots, z_n$ 不全为零.将 $\boldsymbol{y}$ 及 $\boldsymbol{z}$ 分别代入(1)及(2),则有

$$f = b_1 y_1^2 + \cdots + b_q y_q^2 + \cdots + b_p y_p^2 > 0, \text{且 } f = -c_{q+1} z_{q+1}^2 - \cdots - c_r z_r^2 \leqslant 0,$$

矛盾.因此,$p = q$.□

附录B.4　定理7.5的证明

定理7.5　n元实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 正定的充分必要条件是 $\boldsymbol{A}$ 的 n 个顺序主子式(左上角主子式)全大于零,即 $|\boldsymbol{A}_k|=|(a_{ij})_{k\times k}|>0(k=1,2,\cdots,n)$.

证　必要性.首先证明,对任意 l 元的正定二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}>0$,其矩阵的行列式 $|\boldsymbol{B}|>0$.因为 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}$ 是正定 l 元二次型,所以存在满秩(l元的)线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{C}=\boldsymbol{I}_l$.则 $|\boldsymbol{B}|=|(\boldsymbol{C}^{\mathrm{T}})^{-1}\boldsymbol{I}_l\boldsymbol{C}^{-1}|=|\boldsymbol{C}^{-1}|^2>0$.

设 f 为 n 元正定二次型,对任何 n 元向量 $\boldsymbol{x}$, $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$.取一个特殊的 n 元向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_k,0,\cdots,0)^{\mathrm{T}}$,及相应的 k 元向量 $\boldsymbol{x}_k=(x_1,x_2,\cdots,x_k)^{\mathrm{T}}$, $1\leq k\leq n$,则 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{x}_k>0$,其中 $\boldsymbol{A}_k=(a_{ij})_{k\times k}$ 为 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的第 k 个顺序主子式相应的矩阵.这时 f 变为正定 k 元二次型了,根据上段的论证知,它的行列式 $|\boldsymbol{A}_k|>0(k=1,2,\cdots,n)$.

充分性.对 n 作数学归纳法.当 $n=1$, $f=a_{11}x_1^2$,因为 $a_{11}>0$,所以 f_n 为正定的.假设充分性对 $k\leq n-1$ 元二次型成立.下面证明,充分性对 n 元二次型也成立.把二次型 $f_n=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j$ 改写为

$$f_n=\frac{1}{a_{11}}(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)^2+\sum_{i,j=2}^{n}b_{ij}x_ix_j,\tag{1}$$

其中

$$b_{ij}=a_{ij}-\frac{a_{1i}a_{1j}}{a_{11}}.$$

因为 $a_{ij}=a_{ji}$,所以 $b_{ij}=b_{ji}(i,j=2,\cdots,n)$.$f_{n-1}=\sum\limits_{i,j=2}^{n}b_{ij}x_ix_j$ 是 $n-1$ 元的二次型,如果 f_{n-1} 是正定的, f_n 就自然是正定的.下面证明 f_{n-1} 是正定的.由行列式性质可知

$$|\boldsymbol{A}_k|=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1k}\\a_{21}&a_{22}&\cdots&a_{2k}\\\cdots&\cdots&&\cdots\\a_{k1}&a_{k2}&\cdots&a_{kk}\end{vmatrix}=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1k}\\0&b_{22}&\cdots&b_{2k}\\\cdots&\cdots&&\cdots\\0&b_{k2}&\cdots&b_{kk}\end{vmatrix}=a_{11}|\boldsymbol{B}_{k-1}|,\tag{2}$$

其中

$$|\boldsymbol{B}_{k-1}|=\begin{vmatrix}b_{22}&\cdots&b_{2k}\\\cdots&&\cdots\\b_{k2}&\cdots&b_{kk}\end{vmatrix}(k=2,3,\cdots,n).$$

由假设知 $a_{11}>0$，$|\boldsymbol{A}_k|>0(k=2,3,\cdots,n)$，从(2)可见 $|\boldsymbol{B}_{k-1}|>0(k=2,3,\cdots,n)$.而 $\boldsymbol{B}_{k-1}$正是 f_{n-1}的顺序主子式，由归纳假设知，f_{n-1}是正定的.于是，f_n 是正定的.□

习题参考答案

习　题　一

习题 A

3．A:Ⅳ, B: Ⅴ, C:Ⅷ, D:Ⅲ．

4．xOy 面:$(x_0, y_0, 0)$, yOz 面:$(0, y_0, z_0)$, xOz 面:$(x_0, 0, z_0)$; x 轴:$(x_0, 0, 0)$, y 轴:$(0, y_0, 0)$, z 轴:$(0, 0, z_0)$．

5．x 轴:　34, y 轴:　41, z 轴:5．

8．$5a - 11b + 7c$．

10．13, $7\boldsymbol{j}$．

11．$\{6\ 11, 7\ 11, -6\ 11\}$或$\{-6\ 11, -7\ 11, 6\ 11\}$．

12．(1)3, $5\boldsymbol{i} + \boldsymbol{j} + 7\boldsymbol{k}$; (2) -18, $10\boldsymbol{i} + 2\boldsymbol{j} + 14\boldsymbol{k}$; (3)$\cos(\boldsymbol{a}, \boldsymbol{b}) = \dfrac{3}{2\quad 21}$．

13．$-3\ 2$．

14．$\lambda = 2\mu$．

15．4．

16．(1)否;(2)否;(3)否;(4)是;(5)是．

17．(1)是;(2)否;(3)否;(4)是．

18．(1) $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$必线性无关;(2)不能判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性无关或线性相关性;(3)不能判断 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_{n-1}$线性无关或线性相关性;(4) $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 必线性相关．

27．(1) $\left(\dfrac{1}{3}, \dfrac{1}{3}, \dfrac{1}{3}\right)$, $\left(-\dfrac{2}{6}, \dfrac{1}{6}, \dfrac{1}{6}\right)$, $\left(0, -\dfrac{1}{2}, \dfrac{1}{2}\right)$;

(2) $\dfrac{1}{10}(1, 2, 2, -1)$, $\dfrac{1}{26}(2, 3, -3, 2)$, $\dfrac{1}{10}(2, -1, -1, -2)$;

(3) $\dfrac{1}{7}(1, 1, -1, 2)$, $\dfrac{1}{4431}(26, 47, -5, -39)$, $\dfrac{1}{422}(-4, 9, 17, 6)$;

(4) $\dfrac{1}{15}(2, 1, 3, -1)$, $\dfrac{1}{23}(3, 2, -3, -1)$, $\dfrac{1}{127}(1, 5, 1, 10)$．

习题 B

8．$\dfrac{1}{2}(\boldsymbol{e}_1 + \boldsymbol{e}_2)$, $\dfrac{1}{10}(\boldsymbol{e}_1 - 2\boldsymbol{e}_2 + 2\boldsymbol{e}_4 - \boldsymbol{e}_5)$, $\dfrac{1}{2}(\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 - \boldsymbol{e}_5)$．

习　题　二

1．$x - 3y - 2z = 0$．

2．$x + y - 3z = 4$．

3. (1) $y+5=0$; (2) $x+3y=0$; (3) $9y-z=2$.

4. $2x+9y-6z-121=0$.

5. (1) 即 yOz 面; (2) 平行于 xOz 面的平面; (3) 平行于 z 轴的平面; (4) 通过 z 轴的平面; (5) 平行于 x 轴的平面; (6) 通过 y 轴的平面; (7) 通过原点的平面.

6. $(1,-1,3)$.

7. (1) $\begin{cases} x=\dfrac{3}{2}\cos t, \\ y=\dfrac{3}{2}\cos t, \quad (0\leq t\leq 2\pi) \\ z=3\sin t. \end{cases}$

(2) $\begin{cases} x=1+3\cos t, \\ y=3\sin t, \quad (0\leq t\leq 2\pi) \\ z=0. \end{cases}$

8. 1.

9. $\dfrac{x-4}{2}=\dfrac{y+1}{1}=\dfrac{z-3}{5}$.

10. $\dfrac{x-3}{-4}=\dfrac{y+2}{2}=\dfrac{z-1}{1}$.

11. $\phi=0$.

12. $x-y+z=0$.

13. $\dfrac{1}{3},\dfrac{2}{3},\dfrac{2}{3}$.

14. $\begin{cases} 17x+31y-37z=117, \\ 4x-y+z=1. \end{cases}$

15. $(1,1,2)$.

16. $x-3y+z+2=0$.

17. $\begin{cases} x-y+2z-1=0, \\ x-3y-2z+1=0. \end{cases}$

18. $(-5\ 3, 2\ 3, 2\ 3)$.

19. $\dfrac{3\ 2}{2}$.

20. 以 $(1,-2,-1)$ 点为球心,半径等于 6 的球面.

21. $y^2+z^2=5x$.

22. 绕 x 轴: $4x^2-9(y^2+z^2)=36$; 绕 y 轴: $4(x^2+z^2)-9y^2=36$.

23. 母线平行于 x 轴的柱面方程: $3y^2-z^2=16$; 母线平行于 y 轴的柱面方程: $3x^2+2z^2=16$.

24. $\begin{cases} x^2+y^2+(1-x)^2=9, \\ z=0. \end{cases}$

25. $x^2+y^2\leq ax$; $x^2+z^2\leq a^2$, $x\geq 0$, $z\geq 0$.

习 题 三

习题 A

1. (1) $\begin{pmatrix} 35 \\ 6 \\ 49 \end{pmatrix}$. (2) 10. (3) $\begin{pmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{pmatrix}$. (4) $\begin{pmatrix} a_1 b_1 & \cdots & a_1 b_n \\ \vdots & & \vdots \\ a_n b_1 & \cdots & a_n b_n \end{pmatrix}$. (5) $\sum_{i=1}^{n} a_i b_i$.

(6) $\sum_{1 \leq i, j \leq 3} a_{ij} x_i x_j$. (7) $\begin{pmatrix} 0 & a_1 & 0 \\ 0 & 0 & a_2 \\ a_3 & 0 & 0 \end{pmatrix}$.

2. (1)否. (2)否. (3)否.

3. (1) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. (2) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. (3) $\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $\boldsymbol{X} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\boldsymbol{Y} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.

12. $\boldsymbol{A}^k = \begin{pmatrix} \lambda^k & C_k^1 \lambda^{k-1} & C_k^2 \lambda^{k-2} \\ 0 & \lambda^k & C_k^1 \lambda^{k-1} \\ 0 & 0 & \lambda^k \end{pmatrix}$.

13. (1) 0. (2) 1. (3) $-2(x^3 + y^3)$. (4) 0.

14. (1) $[x + (n-1)a](x-a)^{n-1}$. (2) a^4. (3) $\prod_{i=1}^{n} (a_i d_i - b_i c_i)$. (4) $(-1)^{n-1}(n-1)2^{n-2}$.

(5) $a_1 a_2 \cdots a_n \left(1 + \sum_{i=1}^{n} \frac{1}{a_i}\right)$.

16. $x_1 = \frac{1\,507}{665}$, $x_2 = -\frac{1\,145}{665}$, $x_3 = \frac{703}{665}$, $x_4 = -\frac{395}{665}$, $x_5 = \frac{212}{665}$.

17. $\lambda = 1$, $\mu = 0$.

26. (1) $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. (2) $\begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$. (3) $\begin{pmatrix} a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1} \end{pmatrix}$.

27. (1) $\boldsymbol{B} = \begin{pmatrix} 5 & -2 & -2 \\ 4 & -3 & -2 \\ -2 & 2 & 3 \end{pmatrix}$, (2) $\boldsymbol{B} = \begin{pmatrix} 3 & -8 & -6 \\ 2 & -9 & -6 \\ -2 & 12 & 99 \end{pmatrix}$.

28. (1) $\begin{pmatrix} \boldsymbol{A}^{-1} & \\ & \boldsymbol{B}^{-1} \end{pmatrix}$, $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{B}^{-1} \\ \boldsymbol{A}^{-1} & \boldsymbol{O} \end{pmatrix}$, $\begin{pmatrix} \boldsymbol{I} & -\boldsymbol{X} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}$, $\begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ -\boldsymbol{Y} & \boldsymbol{I} \end{pmatrix}$.

(2) $\frac{1}{24}\begin{pmatrix} 24 & 0 & 0 & 0 \\ -12 & 12 & 0 & 0 \\ -12 & -4 & 8 & 0 \\ 3 & -5 & -2 & 6 \end{pmatrix}$, $\begin{pmatrix} 1 & -2 & 0 & 0 \\ -2 & 5 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 0 & -5 & 8 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 2 & -3 \\ 0 & 0 & -5 & 8 \\ 1 & -2 & 0 & 0 \\ -2 & 5 & 0 & 0 \end{pmatrix}$.

29. (1) $\boldsymbol{X} = \begin{pmatrix} 2 & -23 \\ 0 & 8 \end{pmatrix}$; (2) $\boldsymbol{X} = \begin{pmatrix} -2 & 2 & 1 \\ -\frac{8}{3} & 5 & \frac{2}{3} \end{pmatrix}$; $\boldsymbol{X} = \begin{pmatrix} 1 & 1 \\ \frac{1}{4} & 0 \end{pmatrix}$; $\boldsymbol{X} = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 3 & -4 \\ 1 & 0 & -2 \end{pmatrix}$.

30. (1) $\begin{cases} x_1 = 1, \\ x_2 = 0, \\ x_3 = 0. \end{cases}$ (2) $\begin{cases} x_1 = 5, \\ x_2 = 0, \\ x_3 = 3. \end{cases}$

31. $\frac{1}{3}\begin{pmatrix} 1+2^{13} & 4+2^{13} \\ -1-2^{11} & -4-2^{11} \end{pmatrix}$.

36. 3, 4.

习题 B

2. $\boldsymbol{A}^{-1} = \boldsymbol{I} - \frac{1}{3} X Y^{\mathrm{T}}$.

3. $\begin{cases} a_1 - b_1 & n = 1 \\ (a_1 - a_2)(b_1 - b_2) & n = 2. \\ 0 & n \geq 3 \end{cases}$

6. $a = 6$, $\boldsymbol{A} = \frac{1}{5}\begin{pmatrix} 3 & -3 \\ -2 & 2 \end{pmatrix}$, $\boldsymbol{B} = \frac{1}{5}\begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}$.

10. $\boldsymbol{A}^{-1} = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{I})$, $(\boldsymbol{A} + 2\boldsymbol{I})^{-1} = \frac{1}{4}(3\boldsymbol{I} - \boldsymbol{A})$.

16. $\begin{pmatrix} 2t^2 e^{2t} + 2te^{2t} + 3t\cos t - t^3 \sin t & 12t \\ e^{2t} + 2te^{2t} + \cos^2 t - \sin^2 t & 4\cos t \end{pmatrix}$.

18. $\begin{pmatrix} \frac{t^3}{3} & 1 - e^{-t} \\ 1 - \cos t & \sin t \end{pmatrix}$, $\begin{pmatrix} 1 & \frac{1}{2} - \frac{1}{2}\cos 2 \\ e - 1 & \sin 1 \end{pmatrix}$.

习 题 四

习题 A

1. (1) 牛 1 200;羊 500;猪 300. (2) $(1, -11, 16)$. (3) $(1, 5, -1)$. (4) $(1, 2, 3)$. (5) $(-6.5, 7.5, 2)$.

2. (1) $(\frac{3}{2}, \frac{3}{2}, 1, 0)^{\mathrm{T}}, (-\frac{3}{4}, \frac{7}{4}, 0, 1)^{\mathrm{T}}, (\frac{5}{4}, -\frac{1}{4}, 0, 0)^{\mathrm{T}}$.

(2) $(-1, -2, 3, 0)^{\mathrm{T}}$.

(3) $(1, -2, 1, 0, 0)^{\mathrm{T}}, (1, -2, 0, 1, 0)^{\mathrm{T}}, (5, -6, 0, 0, 1)^{\mathrm{T}}$.

(4) $(-1, -1, 1, 2, 0)^{\mathrm{T}}, (1, 0, 0, 5, 4)^{\mathrm{T}}$.

3. (1) $k(-1, -1, 1)^{\mathrm{T}} + (-2, 7, 0)^{\mathrm{T}}$.

(2) $k_1(-9, 1, 0, 11)^{\mathrm{T}} + k_2(1, -5, 11, 0)^{\mathrm{T}} + (0, 0, 2, 0)^{\mathrm{T}}$.

(3) $k(1, 1, 0, 1, -2)^{\mathrm{T}} + (1, 0, 0, 0, -2)^{\mathrm{T}}$.

(4) $k_1(3, 3, 2, 0)^{\mathrm{T}} + k_2(-3, 7, 0, 4)^{\mathrm{T}} + (\frac{5}{4}, -\frac{1}{4}, 0, 0)^{\mathrm{T}}$.

4. (1) $\lambda = 5$, $k_1(-1, 3, 5, 0)^{\mathrm{T}} + k_2(-6, -7, 0, 5)^{\mathrm{T}} + (\frac{4}{5}, \frac{3}{5}, 0, 0)^{\mathrm{T}}$.

(2) $\lambda = 1$, $k_1(-1, 1, 0)^{\mathrm{T}} + k_2(-1, 0, 1)^{\mathrm{T}} + (1, 0, 0)^{\mathrm{T}}$;

$\lambda \neq 1$ 且 $\lambda \neq -2$ 有唯一解, $\left(-\frac{\lambda+1}{\lambda+2}, \frac{1}{\lambda+2}, \frac{(\lambda+1)^2}{\lambda+2}\right)^{\mathrm{T}}$.

9. $(a_1 + a_2 + a_3 + a_4, a_2 + a_3 + a_4, a_3 + a_4, a_4, 0)^{\mathrm{T}} + k(1, 1, 1, 1, 1)^{\mathrm{T}}$.

11. $\begin{cases} 8x_1 - 3x_2 + 2x_3 = 0, \\ -2x_2 + x_4 = 0. \end{cases}$

13. 当 $a=1$ 且 $b\neq -1$ 时,方程组无解;当 $a\neq 1$ 时,方程组有唯一解;当 $a=1$ 且 $b=-1$ 时,方程组有无穷多解且其通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + k_1 \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}, k_1, k_2 \in \mathbf{R}.$$

14. $(x_1, x_2, \cdots, x_n)^{\mathrm{T}} = k(1, 1, \cdots, 1)^{\mathrm{T}}, k\in\mathbf{R}$.

15. $(x_1, x_2, \cdots, x_{2n})^{\mathrm{T}} = k_1(b_{11}, b_{12}, \cdots, b_{1,2n})^{\mathrm{T}} + k_2(b_{21}, b_{22}, \cdots, b_{2,2n})^{\mathrm{T}} + \cdots + k_n(b_{n1}, b_{n2}, \cdots, b_{n,2n})^{\mathrm{T}}$.

16. $C(126°31'12'', 45°36'54'')$.

习题 B

1. (1) $(0.342\,5, 0.231\,2, 0.146\,1)$. (2) $(0.999\,0, 0.243\,1, 0.123\,4)$.

4. 各部门总产值 $(4.00, 3.00, 3.00)$,中间产品分配

$$\begin{pmatrix} 1.00 & 0.30 & 0.35 \\ 0.80 & 0.60 & 0.35 \\ 0.40 & 0.30 & 0.70 \end{pmatrix}.$$

5. $(y_1, y_2, y_3) = (2\,450, 900, 1\,750)$, $(u_1, u_2, u_3) = (1\,700, 1\,450, 1\,700)$.

6. (1) $y = -\frac{1}{2} + \frac{7}{2}x$. (2) $y = \frac{2}{3} + \frac{1}{6}x$

7. $y = 2 + 5x - 3x^2$.

习　题　五

习题 A

1. A:非线性变换;B:线性变换但非满秩变换;C:满秩线性变换.

(1) A;(2) C;(3) C;(4) B;(5) A.

4. (1) $\begin{pmatrix} 3 & -1 \\ 2 & 4 \end{pmatrix}$;(2) $\begin{pmatrix} 2 & 3 & -8 \\ 1 & 1 & 1 \\ 4 & 0 & -5 \end{pmatrix}$.

5. (1) $\begin{pmatrix} 2 & 3 & 5 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$;(2) $\begin{pmatrix} -\frac{5}{7} & \frac{20}{7} & -\frac{20}{7} \\ -\frac{4}{7} & -\frac{5}{7} & -\frac{2}{7} \\ \frac{27}{7} & \frac{18}{7} & \frac{24}{7} \end{pmatrix}$.

6. (1) $\boldsymbol{A} = \frac{1}{2}\begin{pmatrix} -4 & -3 & 3 \\ 2 & 3 & 3 \\ 2 & 1 & -5 \end{pmatrix}$;(2) 仍然为 $\boldsymbol{A}$.

习题 B

5. (1) $\mathrm{Im}(T)$的基：$(1,5,7)^T$, $(0,1,1)^T$, $\dim(\mathrm{Im}(T))=2$;

$\mathrm{Ker}(T)$的基：$(-14/11, 19/11, 1)^T$, $\dim(\mathrm{Ker}(T))=1$.

(2) $\mathrm{Im}(T)$的基：$(1,2,0)^T$, $\dim(\mathrm{Im}(T))=1$;

$\mathrm{Ker}(T)$的基：$(1/2,0,1)^T$, $(0,1,0)^T$, $\dim(\mathrm{Ker}(T))=2$.

6. $\mathrm{Im}(T)$: $14x-8y-5z=0$; $\mathrm{Ker}(T)$: $x=-t, y=-t, z=t$. $-\infty<t<\infty$.

7. $\mathrm{Im}(T)$的基：$(1,1,1,1)^T$, $(1,2,3,4)^T$, $(1,3,6,10)^T$, $\dim(\mathrm{Im}(T))=3$;

$\mathrm{Ker}(T)$的基：$(-1,3,-3,1)^T$, $\dim(\mathrm{Ker}(T))=1$

$\boldsymbol{b}$的原象：$k(-1,3,-3,1)^T+(3,-8,6,0)^T$

习　题　六

习题 A

1. (1) $\lambda=1,2,3$; $(0,1,0)^T$, $(-1/2,1,1)^T$, $(-1,1,1)^T$.

(2) $\lambda=0,\sqrt{2},-\sqrt{2}$; $\left(\frac{5}{3},\frac{1}{3},1\right)^T$, $\left(\frac{1}{7}(15+5\sqrt{2}),\frac{1}{7}(2\sqrt{2}-1),1\right)^T$, $\left(\frac{1}{7}(15-5\sqrt{2}),\frac{1}{7}(-1-2\sqrt{2}),1\right)^T$.

(3) $\lambda=-2,1,1,-1$; $(-1,0,1,0)^T$, $(0,0,0,1)^T$, $(2,3,1,0)^T$, $(-2,1,1,0)^T$.

(4) $\lambda=2,2,2$; $(-1/3,-1/3,1)^T$.

5. (1) 不能对角化.

(2) $\boldsymbol{P}=\begin{pmatrix}1&\frac{2}{3}&1\\1&1&3\\1&1&4\end{pmatrix}$, $\boldsymbol{P}^{-1}=\begin{pmatrix}3&-5&3\\-3&7&-6\\0&-1&1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}1&&\\&2&\\&&3\end{pmatrix}$.

(3) $\boldsymbol{P}=\begin{pmatrix}1&0&0&0\\0&1&1&0\\0&0&1&1\\0&0&0&1\end{pmatrix}$, $\boldsymbol{P}^{-1}=\begin{pmatrix}1&0&0&0\\0&1&-1&1\\0&0&1&-1\\0&0&0&1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}-2&&&\\&-2&&\\&&3&\\&&&3\end{pmatrix}$.

(4) 不能对角化.

6. (1) $\begin{pmatrix}-2^{100}+2&-2^{101}+2&0\\2^{100}-1&2^{101}-1&0\\2^{100}-1&2^{101}-2&1\end{pmatrix}$. (2) $\begin{pmatrix}-1&10237&-2047\\0&1&0\\0&10245&-2048\end{pmatrix}$.

7. (1) $\boldsymbol{P}=\frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\\1&1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}4&\\&2\end{pmatrix}$. (2) $\boldsymbol{P}=\frac{1}{\sqrt{5}}\begin{pmatrix}1&2\\2&-1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}2&\\&7\end{pmatrix}$.

(3) $\boldsymbol{P}=\begin{pmatrix}0&1&0\\\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}2&&\\&4&\\&&4\end{pmatrix}$.

(4) $\boldsymbol{P}=\begin{pmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{3}}&-\frac{2}{\sqrt{6}}&0\\\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}0&&\\&3&\\&&3\end{pmatrix}$.

(5) $\boldsymbol{P}=\frac{1}{5}\begin{pmatrix}3 & 4 & 0 & 0\\ 4 & -3 & 0 & 0\\ 0 & 0 & 3 & 4\\ 0 & 0 & 4 & -3\end{pmatrix}$，$\boldsymbol{\Lambda}=\begin{pmatrix}25 & & & \\ & -25 & & \\ & & 25 & \\ & & & 25\end{pmatrix}$.

(6) $\boldsymbol{P}=\begin{pmatrix}1\ 2 & 1\ 2 & 1\ 6 & -\ 3\ 6\\ -1\ 2 & 1\ 2 & -1\ 6 & 3\ 6\\ -1\ 2 & 0 & 2\ 6 & 3\ 6\\ 1\ 2 & 0 & 0 & 3\ 2\end{pmatrix}$，$\boldsymbol{\Lambda}=\begin{pmatrix}-3 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1\end{pmatrix}$.

20．$\boldsymbol{A}^n\boldsymbol{b}=\begin{pmatrix}2-2^{n+1}+3^n\\ 2-2^{n+2}+3^{n+1}\\ 2-2^{n+3}+3^{n+2}\end{pmatrix}$.

21．$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & -1\\ 0 & -1 & 0\end{pmatrix}$.

22．(1) $\lambda=-1$, $a=-3$, $b=0$.(2) **因** 3 **阶方阵** $\boldsymbol{A}$ **的与** 3 **重特征值** $\lambda=-1$ **对应的特征向量仅有一个**,**所以** $\boldsymbol{A}$ **不与对角阵相似**．

23．$\left(\frac{|\boldsymbol{A}|}{\lambda_i}\right)^2+1$, $i=1,2,\cdots,n$.

24．$a=2$, $b=-3$, $c=2$, $\lambda_0=1$.

习题 B

4．$\boldsymbol{\Lambda}=\begin{pmatrix}n & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \cdots & \cdots & & \cdots\\ 0 & 0 & 0 & 0\end{pmatrix}$

5．$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n+\sum_{i=1}^{n}(b_i^2+c_i^2)\lambda^{n-1}+\sum_{i<j}\begin{vmatrix}b_i & c_i\\ b_j & c_j\end{vmatrix}^2\lambda^{n-2}$.

习　题　七

习题 A

1．(1) $f=z_1^2+6z_2^2-6z_3^2$, $\boldsymbol{x}=\boldsymbol{Cz}$, $\boldsymbol{C}=\begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 1\\ 0 & 1 & -1\end{pmatrix}$.

(2) $f=y_1^2+y_2^2$, $\boldsymbol{x}=\boldsymbol{Cy}$, $\boldsymbol{C}=\begin{pmatrix}1 & -1 & 1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{pmatrix}$.

(3) $f=2z_1^2-2z_2^2+6z_3^2$, $\boldsymbol{x}=\boldsymbol{Cz}$, $\boldsymbol{C}=\begin{pmatrix}1 & -1 & 1\\ 0 & 1 & 2\\ 0 & 0 & 1\end{pmatrix}$.

(4) $f = 2z_1^2 - 2z_2^2$, $\boldsymbol{x} = \boldsymbol{Cz}$, $\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -2 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.

(5) $f = z_1^2 - z_2^2 + z_3^2 - z_4^2 + \cdots + z_{2n-1}^2 - z_{2n}^2$, $\boldsymbol{x} = \boldsymbol{Cz}$, $\boldsymbol{C} = \begin{pmatrix} \boldsymbol{P} & & \\ & \ddots & \\ & & \boldsymbol{P} \end{pmatrix}$, 其中 $\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

2. (1) $f = y_1^2 + 3y_2^3$, $\boldsymbol{x} = \boldsymbol{Cy}$, $\boldsymbol{C} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

(2) $f = y_1^2 + 6y_2^3$, $\boldsymbol{x} = \boldsymbol{Cy}$, $\boldsymbol{C} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$.

(3) $f = 10y_1^2 + y_2^2 + y_3^2$, $\boldsymbol{x} = \boldsymbol{Cy}$, $\boldsymbol{C} = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & -2 & 1 \\ -2 & 1 & 2 \end{pmatrix}$.

(4) $f = 2y_2^2 - 7y_3^2$, $\boldsymbol{x} = \boldsymbol{Cy}$, $\boldsymbol{C} = \begin{pmatrix} 1\ \ 14 & 1\ \ 6 & -4\ \ 21 \\ -2\ \ 14 & 2\ \ 6 & 1\ \ 21 \\ 3\ \ 14 & 1\ \ 6 & 2\ \ 21 \end{pmatrix}$.

3. (1) $\frac{x'^2}{4} - \frac{y'^2}{\frac{8}{3}} = 1$, $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}$.

(2) $\frac{x'^2}{\frac{9}{7}} + \frac{y'^2}{\frac{9}{3}} = 1$, $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}$.

(3) $\frac{x'^2}{6} + \frac{y'^2}{3} = 1$, $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 5 \\ 0 \end{pmatrix}$.

(4) $\frac{x'^2}{11} + \frac{y'^2}{6} = 1$, $\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 1 & 3 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 9 \\ 7 \end{pmatrix}$.

6. (1),(2) 正定.(3)负定.(4) 非正定非负定.

7. (1) $- 2 < t < 2$, f 为正定.(2) $t > 2$, f 为正定.

14. (1) $c = 3$, $\lambda_1 = 0$, $\lambda_2 = 4$, $\lambda_3 = 9$;(2) 椭圆柱面.

16. $a = 2$,所用正交变换矩阵为

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

17. $a = 3$, $b = 1$,所用正交变换矩阵为

$$\begin{pmatrix} 1\ \ 2 & 1\ \ 3 & 1\ \ 6 \\ 0 & -1\ \ 3 & 2\ \ 6 \\ -1\ \ 2 & 1\ \ 3 & 1\ \ 6 \end{pmatrix}$$

习题 B

3. (1) 双曲抛物面: $4x'^2 - 3y'^2 + 3z' = 0$.

(2) **椭圆抛物面**:$90x'^2+27y'^2+10z'=0$.

(3) **椭圆柱面**:$6x'^2+3y'^2-1=0$.

(4) **双曲柱面**:$2x'^2-y'^2+1=0$.

主要参考文献

1 同济大学数学教研室 .线性代数 .第 3 版 .北京:高等教育出版社,1999
2 同济大学数学教研室 .高等数学 .第 4 版 .北京:高等教育出版社,1996
3 上海交通大学线性代数编写组 .线性代数 .第 3 版 .北京:高等教育出版社,1991
4 萧树铁,居余马 .高等数学:第 I 卷　基础与代数 .北京:清华大学出版社,1995
5 俞正光,李永乐,詹汉生 .线性代数与解析几何 .北京:清华大学出版社,1998
6 许以超 .代数学引论 .上海:上海科学技术出版社,1966
7 谢邦杰 .线性代数 .北京:人民教育出版社,1978
8 熊全淹,叶明训 .线性代数 .第 3 版 .北京:高等教育出版社,1985
9 方德植 .解析几何 .北京:高等教育出版社,1986
10 丘维声 .解析几何 .北京:北京大学出版社,1988
11 Strang G .线性代数及其应用 .侯自新,等译 .天津:南开大学出版社,1990
12 Kalman B .应用线性代数 .方世荣译 .台湾:晓园出版社,1982
13 李有法 .数值计算方法 .北京:高等教育出版社,1996
14 COMAP .数学的原理与实践 .申大为,等译 .北京:高等教育出版社,施普林格出版社,1999
15 梁宗巨 .世界数学史简编 .沈阳:辽宁人民出版社,1980
16 Polya G .数学与猜想 .第 1 卷 .李心灿,等译 .北京:科学出版社,1984
17 Dieudonne J .当代数学——为了人类心智的荣耀 .沈永欢译 .上海:上海教育出版社,1999
18 黎光耀 .智慧的星光——诺贝尔自然科学奖获奖者文萃 .北京:经济日报出版社,2000