

序 言

长期以来,我国高等学校各类非数学专业的数学基础课都限于以微积分为主要内容的“高等数学”。面临 21 世纪各门知识的相互渗透和自身加速更新的形势以及全面提高人才素质的需要,数学的作用将显得日益重要。而作为高等学校数学基础课的作用,除了作为各门学科的重要工具以外,它在提高人才全面素质中起着重要作用的培育理性思维和审美功能方面也应得到充分的重视。这就需要一部与之相适应的教材。

这套“大学数学”教材是在前国家教委“面向 21 世纪教学内容和课程体系改革”研究课题的支持下完成的。共有五本《一元微积分》、《多元微积分及其应用》、《代数与几何》、《随机数学》与《数学实验》。我们认为它们是 21 世纪高级人才应该普遍具备的数学基础。希望学生通过对它们的学习,能使他们在掌握数学工具、提高理性思维和审美素质以及获取新知识的能力诸方面打下一个良好的基础。这种要求应该对任何专业都一样,只是在深度上及侧重的方面可能会有些区别。

在现行的《高等数学》中,微积分和数学分析之间的关系一直是一个难以处理的问题。19 世纪以前的微积分,以它的直观性和不断扩展的应用显示了数学的威力,但同时也暴露出其缺乏严格逻辑基础的缺点。诞生于 19 世纪的数学分析则以其逻辑的完美显示了数学的理性精神。这两个方面在教材中如果结合得好,可以激发初学者对数学的兴趣;但如果结合得不好,则很可能失去两者的活力而形成一堆枯燥的形式推理和繁琐的计算。在本书中我们力图按其本来的面目来编写,把一元微积分分为两部分:前一部分注重直观,着重训练应用和运算,后一部分则着重培育理性思维。

《多元微积分及其应用》的应用内容包括复变函数、微分几何及常微分方程。

《代数与几何》的代数部分基本上是线性代数,其内容也可分为两部分:一部分是以算法为主的求解一般线性方程组的内容;另一部分则主要研究线性空间及其上的线性映射。由于后者是前者的理论框架,而且它已成为近代数学普遍使用的基本语言,因此本书在集合、关系、运算、代数结构之后,较快地进入后者的讨论,并且通过数值表示把两者结合起来。

至于几何,尽管它在古希腊及 19 世纪有其辉煌的历史,在本世纪后半叶也进入了数学研究的主流行列,但近 50 年来,在我国高校的数学基础课中,却一直被压

缩到只剩下一点空间解析几何.这对培养学生的形象思维及理性思维的习惯极为不利.本书除了在多元微积分应用中加上古典微分几何基础(曲线和曲面)以外,在几何部分则增加了“仿射及射影几何”及非欧几何的两个初等模型.

本世纪后半叶以来,人们对事物认识演化的表现之一是从单纯的确定性思维模式进入确定—随机性模式.这一趋势还在发展,在高校数学教学中已受到广泛的关注.我们提出把《随机数学》正式列入基础课.本书内容的重点是通过几个典型范例的讨论,使学者学会描述与表达随机性及随机变化的过程,即集中于对随机模式认识的训练.

在这套系列教材中《数学实验》有其独特性.它的知识内容包含数值方法、统计计算和优化计算的基本概念和初等方法,其目的是为学生自己动手解决问题提供必要的数学知识和软件平台.这是一门以学生独立动手,教师起辅导作用的课程.这类课程的教材如何编写,本书只是一种尝试.

以上是这套教材的一个简要介绍.这套教材既是一个统一的整体,各部分之间又有相对的独立性,可以独立讲授.在内容方面,它包含了现行的高等数学、线性代数、复变函数、微分方程、微分几何、数值分析、概率统计、优化计算等课程最基本的内容,而总学时则大为减少.我们在清华大学几个班的试验表明,全部讲完上述内容所需的学时大约为340左右.除数学实验外,如果再减掉一些内容,280学时左右也是可以的,可由教师灵活掌握.

这套教材在有些大段落后面,附有一段“评注”,主要讲述这一段的重要思想和可能的发展,为有兴趣的学生进一步学习数学开一点小小的窗口.

大凡一本可用的教材,往往有两种写法:尽量多写一点,以便于教师选择;或尽量写少一点,以便于教师发挥.这套教材似乎偏于前者.原因是这是一个尝试,对习惯讲授传统“高等数学”的教师来说,对这套教材可能不太适应,也许需要多一些说明.

这套教材原有的基础是清华大学出版社1995年出版,萧树铁、居余马、葛严林等主编的三卷本《高等数学》.参与现在这套教材编写的有朱学贤、郑建华、章纪民、居余马、李海中、钱敏平、叶俊、姜启源、高立、何青等人.谭泽光、白峰杉、韩云瑞等同志为本书的编写作了大量的工作.高教出版社对本书的编写和出版始终给予热情的支持.

前面已说过,这套教材的编写是一个尝试.目的在于根据“百家争鸣”的精神,参与探索大学数学基础课在培养下一世纪高素质人才中所应起的作用,以及与之相适应的教材建设.我们衷心欢迎各方人士对这套教材评头论足,指出缺点和错误.如果这套教材能起到抛砖引玉的作用,我们就很满足了.

萧树铁

1999年6月

再版序言

提高大学数学教学质量的关键在于教师,但一套较好的教材也是重要的.随着我国大学数学教学内容改革的逐步深入,当前不少高等学校在基础数学教学内容的改革方面有了一些进展,例如单纯“面向专业”的观念有所淡化,代数课程的内容和学时有所增加,开设了一些新的课程,如“数学实验”和“随机数学”等,相应地有一批新教材出版.本套教材也在试用了两年多以后,进行了部分修订.这就是《大学数学》的第二版.

在保持原有的指导思想和风格的前提下,这一套教材由原来的五本《一元微积分》、《多元微积分及其应用》、《代数与几何》、《随机数学》及《数学实验》改编、扩充为七本,即《微积分(一)》、《微积分(二)》、《多元微积分及其应用》、《流形上的微积分》、《代数与几何》、《随机数学》及《数学实验》,其中《流形上的微积分》是新编入的,《随机数学》正在在修订中,《数学实验》这次还来不及修订.除了这三本以外,另外四本改编的大致情况如下:

《微积分(一)》以原来的《一元微积分》中的第一篇,即“直观基础上的微积分”为其主要内容,力求做到“返璞归真”.除了进一步强调了计算和应用之外,还增加了一些对“极限”的朴素描述.

《微积分(二)》是把原来《一元微积分》中的第二篇,即“理性微积分”的内容作一些修改而成.其中为了使读者能更好体会数学分析中的一些基本手法,对用阶梯函数逼近的办法来处理定积分(即函数集扩张的思想)又作了一些改进.

《多元微积分及其应用》是把原书加以适当精简而成.原书中“复变函数”部分重新改写以求突出重点和更加精练,原书的“微分几何”部分移到《代数与几何》.

以上三本教材的习题也都作了调整.

《代数与几何》内容的变动是适当精简了代数的内容,增加了“行列式的几何意义”,几何部分则增加了“微分几何”的基本内容.

《流形上的微积分》与前面三本微积分教材合在一起,就显示了微积分从古典一直到现代的基本面貌,而且也是一个理解当代数学和物理一个不可缺少的台阶.虽然目前它并不属于数学基础课的范围,但可供对此有兴趣的学生选修.此外,对从事微积分教学而在这方面有所欠缺的教师来讲,不妨顺便补

上这一课.

《代数与几何》中的几何部分包括了仿射、射影和微分几何,还有两个非欧几何的模型.它所需的学时不多(不超过 30 学时).这些内容的选取和写法是否合适,能在多大程度上体现数学理性思维和“数学美”,还有待进一步讨论.人们对大学数学课程中几何被严重削弱的缺陷已有共识,但又往往以“课内学时不够”或“没有用”等理由保留了这个缺陷.精简课内学时是必要的,内容的选取更可以讨论.希望有志于此的教师能先试开一些这方面的选修课,供大家来讨论.

这次内容的调整主要是为了增加这套教材的灵活性,不同的学校或专业在内容上可以有不同的选择:可以选择其中的某几本,或删去某些用小字写的部分.例如在清华,这套教材就初步适应了一个较为稳定的教学计划.即除了部分文科和艺术类专业以外,数学基础课的内容确定为:“微积分(一)”(3 学分);“微积分(二)”(3 学分);“多元微积分及其应用”(4 学分);“代数”(4 学分);“几何”(2 学分);“随机数学”(3 学分);“数学实验”(3 学分),其中 1 学分表示一个学期(实上课 15 周)上 15 节课(每节 45 分钟),另外适当安排少数课外习题课.这样数学基础课的总学时就是 330 学时,而其中被列为必修基础课的只有“微积分(一)”和“代数”两门.但实际上多数专业的学生几乎都选了大部分甚至全部数学基础课.

参加这一版改写工作的有朱学贤、郑建华、章纪民、华苏、居余马、萧树铁、李津、陈维恒等同志,谭泽光、白峰杉同志参加了讨论并提出很多好的意见.

编者于清华园

2002 年 10 月

第二版前言

本书第二版主要是考虑到目前国内非数学专业的大学数学课程中,各校有关“代数与几何”的课时相差较大,为了便于使用本教材的教师合理地取舍教材内容进行教学,我们把教材内容线性代数与空间解析几何部分分为三个档次:第一档次是不打*的节,这是教学的基本要求,其中有些用小字排印的内容可以不作为基本要求,例如,分块矩阵的初等变换,正交矩阵中的 $Q-R$ 分解, Hadamard 不等式等;第二档次是用大字排印的打*的节,例如, 2.6 子空间的交与和,直和, 2.9 正交子空间,正交补;第三档次是用小字排印的打*的节,例如, 7.2 二次曲线一般方程化为标准方程及其分类, 8.4 二次曲面的分类, 7.7 中的双线性函数, 1.4 中的序关系,偏序集,全序集,但其中的第二数学归纳法原理应作为教学基本要求,让学生有所了解并会用.这三个档次中的第二、三档次对课时少的学校都可以不作为教学的基本要求,仅供学有余力且有兴趣的学生课外自己阅读.

与第一版相比,我们对 1.10 基本代数结构——群、环、域的基本概念作了较大的修改,去掉了较多内容,只保留了最基本的概念,其中有些例子也用了小字.另外,增加了 n 阶行列式的几何意义(用小字排印,可不作基本要求).

与第一版相比,还有一个较大的变化是把多元微积分中的微分几何的基础知识——“空间曲线与空间曲面”安排为本书的第 9 章,主要的考虑是把大学数学中有关几何几门课(空间解析几何,微分几何,射影几何,非欧几何)的内容都集中在一起,以便于更好地安排教学.

本书第 1~8 章由居余马修订,第 9 章由萧树铁编写并修订,第 10~11 章由萧树铁、李海中修订.书稿最后由主编萧树铁教授审定.

编者于清华园
二〇〇二年十二月

前 言

多年来,我国高校非数学专业的数学基础课只有微积分以及为之服务的一点解析几何.80年代开始,由于计算机解题的需要,一些学校陆续增加了“线性代数”课.它的内容主要是行列式、线性方程组和矩阵运算.而对于线性代数的核心内容——线性空间及其上的线性映射则涉及很少.面对21世纪培养高素质人才的需要,这种数学基础课的结构是难以适应的.建立一个以微积分、代数、几何和随机数学为基本内容的数学基础课新体系的任务随着新世纪的来临就显得急迫了.教育部为此从1995年起就开始立项研究.本书就是在此背景下进行的一个尝试.

这本教材包含了代数和几何两部分.关于其内容,有以下几点考虑:

一、关于代数与几何内容的整体安排

把代数与几何合写成一本教材.这种做法可能会削弱几何的训练,但也可使代数与几何更好地结合,互相渗透,互相促进.代数为研究几何问题提供有效的方法;几何为抽象的代数结构和方法提供形象的几何模型和背景.基于这样的考虑,本书大致分为四个部分.

第一部分是第一章中阐述的基础知识——集合、关系、运算与映射及代数结构的基本概念.其重点是关系和运算.特别要让学生了解运算的对象是多种多样的,而不是局限于数的运算.这里讨论了:映射(是一种一元运算)的一般概念;集合运算;命题运算;几何向量的运算; n 元向量的线性运算及高斯消元法中的矩阵初等行运算.它们不仅拓宽了学生对“运算”概念的认识,而且也启发学生以后在探索新问题时要善于把实际问题变为数学问题;同时也是学习本书必备的基础.关于群、环、域,重点是群的概念,这里主要是让学生初步了解从数学结构上区分各种数学问题的异同,知道近世代数主要是讨论各种代数结构的性质.

第二部分是由第二章至第五章所阐述的线性代数的基本内容——线性空间与内积空间、线性映射、矩阵和行列式.

第三部分是由第六章至第八章所组成,它们是线性代数与几何相结合的内容.线性方程组的解的理论与线性图形(平面与空间直线)的位置关系和度量关系;特征值与特征向量、正交变换、二次型与二次曲线、二次曲面的不变量及其分类.

第四部分是由第九章与第十章所介绍的“仿射与射影几何”及“非欧几何”的两个初等模型.

二、关于线性代数的内容和体系

线性代数的内容大致可分为两部分:一部分是以算法为主的求解线性方程组和矩阵运算(包括特征值与特征向量、矩阵的三种标准形及二次型).另一部分则主要是研究线性空间和内积空间的结构,以及有限维线性空间上的线性映射.由于后者是前者的理论框架,是线性代数的核心内容,而且它也是近代数学普遍使用的基本语言.为使学习较全面深入地掌握和理解线性代数,提高学习现代数学新知识的能力,我们认为必须加强线性空间和线性映射的教学.除了增加必要的课时,还要突出它们的核心地位.为此,我们没有采用国内现行线性代数教材的传统模式,而是直接从讨论线性空间的结构和研究线性映射入手,展开线性代数的内容.通过有限维线性空间的线性映射的数值表示,建立了线性映射与矩阵的对应关系,从而矩阵的基本运算的定义也由相应的线性映射的运算所确定;至于线性方程组的求解问题也对应于已知线性映射的像求完全原像(或核)的问题,矩阵的相似标准形问题也就是从线性映射在不同基下对应的矩阵构成的等价类中找一个最简单的代表元的问题.如此,线性空间和线性映射的概念就贯穿始终,起到了统领全局的核心作用.当然这样的课程体系,初学者在开始学习时会有些困难,感到抽象和不容易理解,但是只要教学得法,按照“从特殊到一般,从具体到抽象”的原则,从具体的模型、背景和实例抽象为一般的概念,学生还是不难接受的.学完整个课程,学生最终会较好和较深入地掌握线性代数的内容,能熟练使用线性空间和线性映射的语言,特别是一开始就接触公理化的定义和方法,对学生以后自学现代数学新知识是大有裨益的.

此外,我们在第六至第八章中还尽量把代数与几何有机地结合起来,例如,特征值的概念是由二次曲线在正交变换下的不变量引出来的,然后再利用特征值与特征向量的概念研究一般的二次型,把二次曲面的一般方程化为标准方程,并对二次曲面作正交分类.

关于行列式,我们也采用了公理化的定义,也就是把 n 阶行列式定义为 n 重反对称线性函数,这样定义不仅方便快捷,而且与全书的风格协调,同时证明行列式的一些重要性质(如 $\det A = \det A^T$ 、行列式按一行(列)的展开定理、 $|AB| = |A||B|$ 等)时,充分利用了矩阵的工具,使证明更为简明.

三、关于几何,除上述内容外我们还选择了“仿射与射影几何”、“非欧几何简介”.目的是加深读者对“形”的理解和认识.

在仿射几何中,研究了仿射变换下的不变性质(如共线性、平行性)和不变量(如简单比),对二次曲线作了更简单的仿射分类.

在射影几何中,我们把欧氏(仿射)平面扩大为射影平面,引进齐次坐标来研究射影平面上的几何问题,研究了对偶原理和射影变换下的不变量(交比),并对二次曲线作了射影分类.

在非欧几何简介中,我们介绍了非欧几何中两个典型的模型.椭圆几何的球面模型和双曲几何的庞加莱模型.它们除了有重要的应用价值外,还希望能通过对这种人类重要文化遗产的认识,培养学生的理性和审美意识.

使用本教材所需课时大约在 80 左右,如果课时少,可根据实际情况适当取舍内容.

本教材的编写工作是在萧树铁教授领导下进行的,第 1~8 章由居余马、第 10~11 章由李海中编写,书稿最后经萧教授审定,高教出版社胡乃啟同志对本书的出版作了认真仔细的编辑工作,对此我们深表感谢.

更新非数学专业的数学基础课的课程体系和教学内容的工作是十分艰巨的,在代数部分我们虽然已经历了九年的教学实践,但探索仍是初步的,缺陷和不妥之处在所难免,恳请同行专家和读者赐教和指正.

编者于清华园
一九九九年九月

目 录

第 1 章 集合 关系 运算 结构.....	(1)
1.1 集合 子集 幂集 直积	(1)
1.2 二元关系及其性质	(3)
1.3 等价关系 等价类 商集	(4)
1.4 * 序关系 偏序集 全序集 数学归纳法原理	(6)
1.5 运算与映射	(7)
1.6 命题运算 量词	(13)
1.7 几何向量的运算 空间直角坐标系	(17)
1.8 n 元向量的线性运算 高斯消元法	(28)
1.9 平面方程与空间直线方程.....	(36)
1.10 基本代数结构——群、环、域的基本概念	(41)
习题	(46)
第 2 章 线性空间 内积空间	(58)
2.1 线性空间的定义及其简单性质.....	(58)
2.2 线性子空间.....	(62)
2.3 线性相关性.....	(65)
2.4 有限维线性空间的基和维数 向量组的秩	(69)
2.5 向量的坐标.....	(72)
2.6 * 子空间的交与和 直和	(74)
2.7 内积空间.....	(77)
2.8 欧氏空间的单位正交基.....	(80)
2.9 * 正交子空间 正交补	(82)
附录 双重连加号 $\sum \sum$ 连乘号 \prod	(84)
习题	(85)
第 3 章 线性映射	(95)
3.1 线性映射的定义及例.....	(95)
3.2 线性映射的像和核	(101)
3.3 线性映射的运算 空间 $L(V_1, V_2)$	(103)
3.4 有限维线性空间的线性映射 线性映射的秩.....	(104)
3.5 线性空间的同构	(109)
习题.....	(113)
第 4 章 矩阵.....	(119)
4.1 矩阵的定义	(119)

4.2	线性映射的矩阵表示	(120)
4.3	矩阵的加法与数量乘法	(122)
4.4	矩阵的乘法	(124)
4.5	可逆矩阵	(129)
4.6	矩阵的转置	(133)
4.7	矩阵的初等变换和初等矩阵	(135)
4.8	矩阵的秩 相抵标准形	(139)
4.9	分块矩阵	(144)
4.10	基的变换矩阵与坐标变换	(150)
	习题	(153)
第 5 章	行列式	(167)
5.1	n 阶行列式的定义	(168)
5.2	行列式按一列(行)的展开式	(174)
5.3	方阵乘积的行列式	(181)
5.4	Cramer 法则	(184)
	习题	(188)
第 6 章	线性方程组与线性几何	(193)
6.1	齐次线性方程组	(193)
6.2	非齐次线性方程组	(196)
6.3	线性图形的几何问题	(201)
	习题	(209)
第 7 章	特征值与特征向量 矩阵的标准形	(216)
7.1	正交变换与正交矩阵	(216)
7.2 *	二次曲线一般方程化为标准方程及其分类	(222)
7.3	线性变换在不同基下的矩阵表示 相似矩阵	(231)
7.4	特征值与特征向量	(234)
7.5	可对角化的条件 相似标准形	(241)
7.6	实对称矩阵的对角化	(247)
7.7	双线性函数 * 二次型	(252)
7.8	实二次型的标准形 实对称矩阵的相合标准形	(255)
7.9	正定二次型与正定矩阵 其它有定二次型	(263)
	习题	(269)
第 8 章	常见曲面及二次曲面的分类	(283)
8.1	球面 柱面 锥面 旋转面	(283)
8.2	空间曲线的方程	(291)

8.3	二次曲面	(292)
8.4 *	二次曲面的分类	(298)
	习题.....	(305)
第 9 章	空间曲线与空间曲面.....	(309)
9.1	向量函数及其微积分	(309)
9.2	曲线的弧长和弗雷耐标架	(315)
9.3	曲线的曲率 挠率 弗雷耐公式	(323)
9.4	特殊的空间曲线	(331)
9.5	曲面的表示 切平面 参数变换	(332)
9.6	曲面的第一基本形式	(338)
9.7	曲面上曲线的法曲率 曲面的第二基本形式.....	(341)
	习题.....	(352)
第 10 章	平面正交变换 仿射变换 射影变换	(360)
10.1	平面正交变换.....	(360)
10.2	平面的仿射变换.....	(362)
10.3	射影平面与齐次坐标.....	(365)
10.4	射影映射和射影变换.....	(379)
	习题.....	(383)
第 11 章	非欧几何学简介	(387)
11.1	球面几何.....	(387)
11.2	双曲几何的庞加莱模型.....	(393)
索引	(398)

第1章 集合 关系 运算 结构

中学数学主要是讨论数的运算、几何图形的性质以及相应的各种数量关系和图形关系,而现代数学的领域非常广阔,客观世界众多的实际问题可以抽象为各种数学问题,并采用不同的数学方法加以解决.所以就整个数学而言,它研究的是各种各样可以表述为数学概念的事物所组成的各类集合,以及同一集合或不同集合中元素间的各种数学关系(运算或变换)和相应的数学结构.因此,我们将在本章介绍大学数学的一些基础知识,首先简要地叙述“集合”的一些基本概念,然后给出“二元关系”、“运算”和“映射”的概念,其中重点是“关系”(主要是等价关系)、“运算”和“映射”,特别是要列举很多例子让读者了解运算的对象非常广泛,而不只局限于“数”.这里读者不仅要掌握“命题运算”、“几何向量的运算”及“高斯消元法中的 n 元向量或矩阵运算”的一些具体内容,而且要从中得到启发,十分重视如何根据实际问题的需要抽象出数学问题,把某些“操作”定义为运算,并在以后的学习中自觉培养这方面的能力.本章最后介绍基本的代数结构——群、环、域的基本概念,让读者初步了解从数学结构上区分各种数学问题的异同.

1.1 集合 子集 幂集 直积

集合是数学的最基本的一个概念,它不能用更简单的概念来定义,而只能对它作些解释.所谓集合(简称集)是指由一些确定的对象(或事物)汇集成的整体,其中每个对象叫集合的元素(简称元).通常用大写字母 A, B, X, Y 等表示集合,用小写字母 a, b, x, y 等表示集的元素.如果元 a 在集 A 中,就说“ a 属于 A ”,记作“ $a \in A$ ”;否则就说“ a 不属于 A ”,记作“ $a \notin A$ ”.

全体自然数、整数、有理数、实数、复数分别组成的集合,依次记作 $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$;全体正整数 $1, 2, \dots, m, \dots$ 组成的集合记作 \mathbf{N}^* .

表述集合的方式:一种是穷举法,即列出集中全部元素;另一种是描述法,即用集中全部元素所具有的特性来表述集合.例如:

$$A = \{\sqrt{2}, -\sqrt{2}\} = \{x | x^2 - 2 = 0, x \in \mathbf{R}\},$$

其中前者是穷举法,后者是描述法.

$$M = \{(x, y) | x - 2y = 1; x, y \in \mathbf{R}\}$$

是用描述法表述直线 $x - 2y = 1$ 上的所有点组成的集合.

含有限个元素的集合叫有限集, 含无穷多个元素的集合叫无限集, 不含任何元素的集合叫空集, 记作 \emptyset , 例如,

$$\{x | x^2 + 1 = 0; x \in \mathbf{R}\}$$

是一个空集. 把空集也看作集合是必要的, 正如把 0 也看作数一样.

定义 1.1 设 A, B 是两个集合, 如果 A, B 含的元素全相同, 就说 A, B 相等, 记作 $A = B$; 如果对任意的 $a \in A$ 均有 $a \in B$, 则称 A 是 B 的子集, 或者说 A 含于 B , B 包含 A , 记作 $A \subset B$.

对任意的集合 A 均有 $\emptyset \subset A, A \subset A$.

显然, 集合 A, B 相等, 当且仅当 $A \subset B$ 与 $B \subset A$ 同时成立.

定义 1.2 非空集合 A 的所有子集组成的集合称为 A 的幂集, 记作: $P(A)$ 或 2^A .

若集 A 含有 n 个不同元素, 则其幂集 $P(A)$ 含有 2^n 个不同元素. 例如集合 $A = \{a, b, c\}$ 的幂集

$$P(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, A\}.$$

我们把非空集 A, B 的任一对有次序的元素 $x \in A, y \in B$ 叫有序二元组(或称序偶), 记作 (x, y) , 并称 x 为第一元素, y 为第二元素. 序偶中的两个元素也可属于同一集合(即 $A = B$). 两个序偶 $(x_1, y_1), (x_2, y_2)$ 相等, 当且仅当 $x_1 = x_2, y_1 = y_2$. 由集合 A, B 中所有元素构成的序偶组成的集合, 叫做集合 A 和 B 的笛卡儿乘积(或称直积), 即

定义 1.3 设 A, B 是两个非空集, 我们把集合

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

称为 A 和 B 的笛卡儿乘积(或称直积).

特别, 当 $B = A$ 时, $A \times A$ (也常记作 A^2) 是 A 中一切元素所作成的序偶集合. 例如 $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ 可表示平面直角坐标系中全部点的坐标的集合.

定义 1.4 设 A, B 是两个集合, 我们把集合

$$A \cap B = \{x | x \in A \text{ 且 } x \in B\},$$

$$A \cup B = \{x | x \in A \text{ 或 } x \in B\},$$

$$A \setminus B = \{x | x \in A \text{ 但 } x \notin B\}$$

分别叫做 A 和 B 的交集, A 和 B 的并集, B 在 A 中的余集. 余集也可记作 $C_A B$ 或 $A - B$, 称 $A - B$ 为 A 与 B 的差集.

如果 $B \in P(X)$, 即 $B \subset X$, 则 B 在 X 中的余集 $X \setminus B$ 常记作 \bar{B} .

1.2 二元关系及其性质

对于集中的元素,我们要研究它们相互间具有的各种关系.例如:在数集中比较数的大小,有 \leq 、 $=$ 、 $<$ 等关系;在直线集中,直线间有平行、垂直、相交等关系;在三角形集中,相似、全等是两种基本关系;在人口集中,人们之间有性别、国籍、肤色相同或不不同的关系;在电路的节点集和支路集中,节点与支路间有关联或不关联的关系.由这些例子可见,一个集合或不同集合中,两个元素间的“关系”是由某种特性或规则来描述的.

定义 1.5 设 X 是一个集合, R 是涉及两个元素的一个规则,如果对于 X 中的任两个元素 a, b 均可确定它们是适合 R (记作 aRb) 或不适合 R (记作 $a\bar{R}b$) 就称 R 是集 X 中的一个二元关系.

如果把 aRb 用序偶 (a, b) 表示,那末集 X 中所有适合关系 R 的元素组组成的集合是 $X \times X$ 的一个子集.因此,我们也可把 $X \times X$ 的一个子集 R 定义为集 X 中的一个二元关系.更一般,我们把 $X \times Y$ 的一个子集 R 定义为集 X 与 Y 间的一个二元关系.

下面讨论二元关系的一些基本性质.

定义 1.6 设 R 是集 X 中的一个二元关系,如果

- (1) 对任意的 $a \in X$ 均有 aRa , 则称 R 是自反的(或称反身的);
- (2) 对任意的 $a, b \in X$ 若有 aRb 就有 bRa , 则称 R 是对称的;
- (3) 对任意的 $a, b \in X$, 由 aRb 和 bRa , 可推出 $a = b$, 则称 R 是反对称的;

- (4) 对任意的 $a, b, c \in X$ 若有 aRb 和 bRc , 就有 aRc , 则称 R 是传递的.

关系 R 是自反的、对称的、反对称的和传递的,也常说成 R 具有自反性、对称性、反对称性和传递性.

例 在实数集中,小于($<$)关系具有传递性;小于或等于(\leq)关系具有自反性、反对称性和传递性.

在正整数集中,如果 aRb 表示 $a \mid b$ (即 a 整除 $b, b/a \in \mathbb{N}^*$) 则 R 具有自反性、反对称性和传递性.

在三角形集中,三角形的相似关系 R 具有自反性、对称性和传递性.在平面直线集中,直线的平行关系 R 也具有自反性、对称性和传递性;而直线的垂直关系 R 只有对称性.在平面的点集中,点在某固定直线上投影相同的关系 R 具有自反性、对称性和传递性.

在集 X 的幂集 $P(X)$ 中, 集合的包含关系“ \subset ”是自反的、反对称的和传递的.

在由世界各国组成的集合中, 国家间有共同陆地边界的关系是自反和对称的.

在人口集合中, 父子关系 R (即 aRb 表示 a 是 b 的父亲) 显然没有自反性、对称性和传递性.

1.3 等价关系 等价类 商集

集合中的等价关系是把集合中元素分类的一种二元关系. 要把集合中元素分类, 自然会提出一个问题——什么叫做两个元素“相同”(或“相等”、“一样”)呢? 事实上天下没有完全不能区别的两个事物, 因此, 所谓两个元素“相同”通常只是指它们具有某种共性, 或者说在某种意义下是“相同”的. 例如: 在动物世界里从动物的性质来看, “人类”中的每个人都是“相同”的; 如果在人类中只考虑人的肤色, 则所有黄种人都“相同”; 如果只考虑人的性别, 则男人都“相同”. 如果对正整数只考虑能否被 2 整除, 则偶数都“相同”; 如果对平面上由直线围成的封闭图形只考虑图形的顶点数, 则所有三角形都“相同”, 如此等等. 这样, 我们就可把两个元素“相同”也看成一种“关系”, 这种“关系”当然有别于其它二元关系. 下面的定义刻画了表示两个元素“相同”这个关系的特点.

定义 1.7 (等价关系) 集合 X 中的一个二元关系 R 称为等价关系, 如果 R 是自反的、对称的和传递的.

例如: 数的相等关系、直线的平行关系、三角形的相似关系、多边形的顶点数相等的关系、人口集合中肤色相同或性别相同的关系、平面点集的点在固定直线上投影相同的关系都是等价关系. 再看一个例子:

设 $a, b \in \mathbb{Z}$, 规定 aRb 为“ a, b 模 n 同余”(记作: $a \equiv b \pmod{n}$), 即 a, b 分别被固定的正整数 n 相除后, 其余数为相同的非负整数 r ($0 \leq r < n$). 整数集 \mathbb{Z} 的这个二元关系 R 是 \mathbb{Z} 的一个等价关系. 因为: (1) 对任意的 $a \in \mathbb{Z}$, 均有 $a \equiv a \pmod{n}$; (2) 若 $a \equiv b \pmod{n}$, 则 $b \equiv a \pmod{n}$; (3) 若 $a \equiv b \pmod{n}, b \equiv c \pmod{n}$, 则必有 $a \equiv c \pmod{n}$.

此例如取 $n = 3$, 容易看出 \mathbb{Z} 的模 3 同余关系 R 把 \mathbb{Z} 划分为三个互不相交的子集:

$$\bar{0} = \{3k \mid k \in \mathbb{Z}\},$$

$$\bar{1} = \{3k + 1 \mid k \in \mathbb{Z}\},$$

$$\bar{2} = \{3k + 2 \mid k \in \mathbb{Z}\}.$$

它们分别是 \mathbb{Z} 中模 3 余 0, 模 3 余 1, 模 3 余 2 的所有整数组成的子集. 显然任一整数都属于且仅属于上述某一子集. 这表明模 3 同余关系把全体整数分成上述三类.

要把集合中的元素按某种意义分类, 自然要求不漏分任一元素, 也不会把任一元素分在两个不同类之中. 下面两个定理表明, 集合的等价关系可把集合的元素分类.

定义 1.8 (元素的等价和等价类) 设 R 是集 X 中的一个等价关系, $a, b \in X$, 如果 aRb , 则称 a, b 关于 R 是等价的, 并把所有与 a 等价的元素集合

$$\bar{a} = \{x \mid xRa, x \in X\}$$

称为 a 关于 R 的等价类 (简称 a 的等价类).

定理 1.1 设 R 是集 X 中的一个等价关系, $a, b \in X$, 则 $\bar{a} = \bar{b}$ 当且仅当 aRb .

证 若 $\bar{a} = \bar{b}$, 则由 $a \in \bar{a} = \bar{b}$ 得 aRb ; 反之, 若 aRb , 则对任意的 $x \in \bar{a}$, 即 xRa , 由传递性得 xRb , 即 $x \in \bar{b}$, 故 $\bar{a} \subset \bar{b}$; 同理可证 $\bar{b} \subset \bar{a}$, 因此 $\bar{a} = \bar{b}$.

通常我们把 a 称为 \bar{a} 的代表元. 由定理 1.1 可见, \bar{a} 中任一元素 b 均可作为 \bar{a} 的代表元.

推论 若 R 是集 X 中的等价关系, 则对任意的 $a, b \in X$, 只能是 $\bar{a} = \bar{b}$ 或 $\bar{a} \cap \bar{b} = \emptyset$.

证 设 $\bar{a} \cap \bar{b} \neq \emptyset$, 则存在 $x \in \bar{a} \cap \bar{b}$, 于是有 xRa 和 xRb , 再由 R 的对称性得 aRx , 由 R 的传递性得 aRb , 从而 $\bar{a} = \bar{b}$.

定理 1.2 若 R 是集 X 中的一个等价关系, 则 X 中存在关于 R 的一族互不相交的等价类

$$\{\bar{a}_i \mid a_i \in X, i \in I\}$$

(其中 I 是所有等价类的代表元 a_i 的下标 i 组成的指标集), 使得 $X = \bigcup_{i \in I} \bar{a}_i$.

证 由于对任意的 a 均有 $a \in \bar{a}$, 所以显然有 $X = \bigcup_{a \in X} \bar{a}$. 根据定理 1.1 的推论, 这个式子右边的任意两个等价类不是相等就是互不相交, 因此, 取 $\bar{a}_i (a_i \in X)$ 中所有互不相交的 $\bar{a}_i (i \in I)$, 其并集就等于 X .

定理 1.2 表明, 集 X 可按等价关系 R 划分为一族等价类, 使 X 中任一元素必属于唯一的一个等价类.

例如,正整数集按模 2 同余关系分为奇数和偶数两个等价类;整数集按模 n 同余关系分为 n 个等价类 $\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}$;人口集合按性别关系分为男女两个等价类.

定义 1.9* 以集 X 的等价关系 R 来划分的所有等价类作为元素所组成的集合,称为 X 关于 R 的商集,记作 X/R .

例 整数集 \mathbb{Z} 关于模 n 同余关系 R 的商集

$$\mathbb{Z}/R = \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}\}.$$

平面直线集 L 关于直线平行关系 R 的商集

$$L/R = \{\bar{l} \mid l \text{ 为过某个定点的直线}\}.$$

设 M 为清华大学学生集,规定 aRb 为 a 与 b 是同班学生,则

$$M/R = \{x_1, x_2, \dots, x_n\},$$

其中 x_i 是第 i 班全体学生的集合, n 为清华大学学生班的总数.

1.4* 序关系 偏序集 全序集 数学归纳法原理

序关系是对集合元素排序的一种二元关系.集合的元素可按大小、先后、重要性等原则排序,它们的共同特性反映在下面两个定义中.

定义 1.10 集 X 中的一个二元关系 R 称为偏序关系,如果 R 具有自反性、反对称性和传递性.

偏序关系 R 常记作 \leq , $a \leq b$ 读作“ a 小于或等于 b ”.具有偏序关系 \leq 的集 X 称为偏序集,记作 (X, \leq) .

例 实数集中的 \leq 关系,幂集 $P(X)$ 中的 \subset 关系,正整数集 \mathbb{N}^* 中的整除关系“ \mid ”,都是偏序关系.

在一个集合中,可以定义不同的偏序关系.对于数集

$$A = \{1, 2, 3, 5, 6, 8, 10, 30\},$$

虽然 (A, \leq) 和 $(A, \text{整除})$ 都是偏序集,但二者有所不同.对于 \leq 而言, A 中任两个元素 a, b 都有 $a \leq b$ 或 $b \leq a$. 因此 A 中全部元素可按 \leq 排序,即

$$1 \leq 2 \leq 3 \leq 5 \leq 6 \leq 8 \leq 10 \leq 30.$$

而对整除关系来说, A 中任两个元素未必有 $a \mid b$ 或者 $b \mid a$ 的关系,因此, A 只有某些子集的元素可按整除关系排序.例如 A 的子集

$$A_1 = \{1, 2, 6, 30\}.$$

对整除关系有

$$1 \mid 2, 2 \mid 6, 6 \mid 30.$$

此外,由自反性和传递性还可得其它一些整除的情况.

定义 1.11 设 $(X, <)$ 是一个偏序集,

(1) 如果对任意的 $a, b \in X$ 均有 $a < b$ 或 $b < a$, 则称 $(X, <)$ 为全序集, $<$ 为全序关系;

(2) 如果 X 的任意非空子集都有最小元 a (即对任意的 $x \in A$ 均有 $a \in A, a < x$), 则称 $(X, <)$ 为良序集.

良序集必是全序集, 因为良序集中任两个元素 a, b 组成的子集必有 $a < b$ 或 $b < a$.

例 (\mathbf{R}, \leq) (\mathbf{Z}, \leq) 都是全序集, 但不是良序集 (\mathbf{N}^*, \leq) 既是全序集也是良序集.

数学归纳法原理的基础是正整数集为良序集, 它依据下面的定理.

定理 1.3 设 $M \subset \mathbf{N}^*$, 如果 $1 \in M$ 且当 $n-1 \in M$ 时可推出 $n \in M$, 则 $M = \mathbf{N}^*$.

证 设 $M' = \mathbf{N}^* \setminus M \neq \emptyset$, 则 $1 \notin M'$, 由于 $M' \subset \mathbf{N}^*$, 所以 M' 必有最小数 $a \in M'$, 于是 $a > 1, a-1 \in M'$, 即 $a-1 \in M$, 如此由定理假设又得 $a \in M$, 这与 $a \in M'$ 矛盾. 故 $M' = \emptyset$, 即 $M = \mathbf{N}^*$.

依据这个定理, 要证明一个命题对所有正整数成立, 只需证明 (1) 命题对 $n=1$ 成立; (2) 若命题对 $k-1$ 成立, 则命题对 k 也成立. 这就是通常的数学归纳法. 此外, 数学归纳法还有另一种形式, 称为第二数学归纳法.

第二数学归纳法原理 设 $P(n)$ 是与正整数有关的一个命题, 如果:

(1) $P(n)$ 对 $n=1$ 成立;

(2) 假设 $P(n)$ 对任意的 $n < k$ 成立, 则 $P(n)$ 对 $n=k$ 也成立.

那末命题 $P(n)$ 对一切正整数 n 都成立.

1.5 运算与映射

中学数学对于数和函数作过各种运算, 例如, 数的四则运算、开方和求对数等运算是由两个数或一个数按一定规则得到另一个数; 又如, 求两个正整数 a, b 的最大公因数 (a, b) , 最小公倍数 $[a, b]$, 算术平均值 $\frac{(a+b)}{2}$, 最大值 $\max(a, b)$, 最小值 $\min(a, b)$ 等, 都是由两个数按一定规则确定第三个数; 再如求集 X 的两子集 A, B 的交或并, 是在幂集 $P(X)$ 中由两个元素依一定规则得到第三个元素. 可见, “运算”是十分广泛的, 它反映了数学应用的广泛性. 下面对“运算”这个概念给出一般的定义.

定义 1.12 设 X, Y, Z 都是非空集, 所谓定义在 X, Y 上取值于 Z 的二元运算“ \circ ”是一个法则, 它使任意的 $x \in X$ 和 $y \in Y$ 都有唯一的 $z \in Z$ 与之对应, 记作 $x \circ y = z$. 如果 $X = Y = Z$, 则称这个二元运算为 X 上的一个代数运算, 或说是 X 上的封闭的二元运算.

按照这个定义, 读者不难定义 n 个集合上的 n 元运算和一个集合上的 n

元运算.

定义 1.13 设 X, Y 是非空集, 定义在 X 上取值于 Y 的一元运算 σ 是一个法则, 它使 X 中每个元素 α 都有 Y 中唯一确定的元素 β 与之对应. 这个一元运算 σ 也称为集合 X 到 Y 的一个映射, 记作

$$\sigma: X \rightarrow Y,$$

并称 β 为 α 在 σ 下的像, α 为 β 在 σ 下的一个原像, 记作

$$\sigma: \alpha \mapsto \beta, \text{ 或 } \sigma(\alpha) = \beta.$$

称 X 为 σ 的定义域, 并把 X 中全体元素在 σ 下的像的集合称为 σ 的值域, 或称 σ 的像, 记作 $\sigma(X)$ 或 $\text{Im}\sigma$, 即

$$\sigma(X) = \{\sigma(\alpha) \mid \alpha \in X\}.$$

集 X 到它自身的映射, 有时也称为 X 的变换.

需要注意, α 的像是唯一的, 但 β 的原像不一定是唯一的.

例 1 设 \mathbf{R} 是实数集, 对任意的 $a, b \in \mathbf{R}$, 定义

$$a \circ b = (a, b),$$

则“ \circ ”是定义在 \mathbf{R} 上取值于 $\mathbf{R} \times \mathbf{R}$ 的一个二元运算, 也称它为“笛卡儿乘积”, 一般 $a \circ b \neq b \circ a$.

定义在两个有限集合上的运算往往用一个表格来表示(最简单的例子是九九表).

例 2 自动售货机售货问题

自动售货机能接受币值为 a, b 的硬币, 投入任两枚这种硬币, 它按下表出售商品.

$*$	a	b
a	桔子水一瓶	糖一包
b	糖一包	冰淇淋一杯

即 $a * a =$ 桔子水一瓶, $b * b =$ 冰淇淋一杯, $a * b = b * a =$ 糖一包. 所以这架自动售货机售货是定义在硬币集 $\{a, b\}$ 上取值于食品集 $\{\text{桔子水}, \text{糖}, \text{冰淇淋}\}$ 的一个二元运算.

例 3 开关电路——逻辑运算

在图 1-1 的电路中, 开关 K_1, K_2 有“接通”和“断开(不通)”两种状态, 电路 S 也有“导通”和“不通”两种状态. 两个开关处于任一状态都确定了 S 的一

种状态. 这就定义了 K_1, K_2 状态集上的一个二元运算. 如果用“0”和“1”分别表示开关与电路“不通”和“通”, 用“ \wedge ”和“ \vee ”分别表示开关串联和并联的运算, 那末定义在开关 K_1, K_2 状态集 $\{0, 1\}$ 上的两种运算 \wedge, \vee 的结果如下:

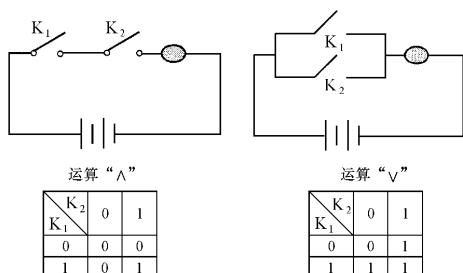


图 1-1

“ \wedge ”叫逻辑乘, 也叫“与”运算; “ \vee ”叫逻辑加, 也叫“或”运算. “与”、“或”和“非”(即 $\bar{1} = 0, \bar{0} = 1$) 运算一起统称逻辑运算, 它们是二值布尔代数中的基本运算.

例 4 设 $P(X)$ 是集 X 的幂集, $A, B \in P(X)$ 则

$$A \cup B, A \cap B, A \setminus B$$

都是 $P(X)$ 上的二元运算; \bar{A} 是 $P(X)$ 上的一元运算. 集合的交、并、余(或差)三种运算适合以下运算律:

(1) $A \cup B = B \cup A, A \cap B = B \cap A$ (交换律)

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

$(A \cap B) \cap C = A \cap (B \cap C)$ (结合律)

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

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (分配律)

(4) $A \cup \emptyset = A; A \cap X = A$ (恒等律)

(5) $A \cap \emptyset = \emptyset, A \cup X = X$ (零律)

(6) $A \cup \bar{A} = X, A \cap \bar{A} = \emptyset$ (互补律)

(7) $\bar{\bar{A}} = A$ (否定律)

(8) $\overline{A \cup B} = \bar{A} \cap \bar{B}, \overline{A \cap B} = \bar{A} \cup \bar{B}$ (DeMorgan 律)

(9) $A \setminus B = A \cap \bar{B}.$

这些运算律可用运算的定义作推理证明, 也可画图证明. 图 1-2 是运算

律(3)与(8)的示意图.

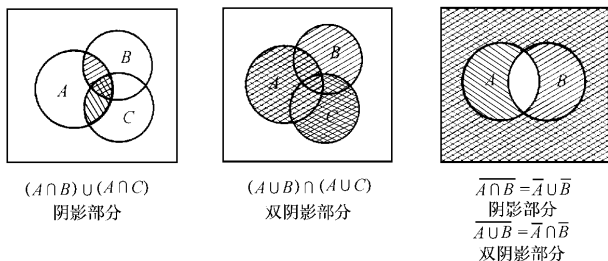


图 1-2

我们用推理证明 $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

设 $x \in \overline{A \cap B}$ 则 $x \notin A \cap B$, 即 x 不同属于 A, B , 所以 $x \notin A$ 或 $x \notin B$, 从而 $x \in \overline{A}$ 或 $x \in \overline{B}$, 即 $x \in \overline{A} \cup \overline{B}$. 故 $\overline{A \cap B} \subset \overline{A} \cup \overline{B}$. 反之, 若 $x \in \overline{A} \cup \overline{B}$, 则 $x \notin A$ 或 $x \notin B$, 所以 $x \notin A \cap B$, 即 $x \in \overline{A \cap B}$. 故 $\overline{A} \cup \overline{B} \subset \overline{A \cap B}$; 于是 $\overline{A} \cup \overline{B} = \overline{A \cap B}$.

下面我们举一些映射的例子, 介绍单射(内射)、满射和双射的概念, 定义映射相等和乘法(或称合成)运算, 并讨论映射可逆的充要条件.

例 5 定义在 $(-\infty, +\infty)$ 上的一元实值函数 $y = f(x)$ 是实数集 \mathbf{R} 到自身的映射.

例 6 对任意的 $n \in \mathbf{Z}$, 定义 $\sigma(n) = 2|n|$, 则 σ 是整数集 \mathbf{Z} 到非负偶数(包括数零)集 E 的映射.

例 7 设 $A = \{a, b, c\}, B = \{1, 2, 3, 4\}$, 定义

$$f: a \mapsto 1, \quad b \mapsto 4, \quad c \mapsto 1,$$

$$g: a \mapsto 3, \quad b \mapsto 1,$$

则 f 是 A 到 B 的映射, 但 g 不是 A 到 B 的映射, 只是 A 的子集 $X = \{a, b\}$ 到 B 的映射.

例 8 设 X 是清华大学历届学生组成的集合, \mathbf{N}^* 是正整数集, 每个学生对应到他的学号是 X 到 \mathbf{N}^* 的一个映射.

例 9 设 X 是某班全体学生的集合, M 是由 12 个月份组成的集合, 每个人对应于他的生日月份是 X 到 M 的一个映射. 如果集 X 的人数大于 12, 那么

至少有两个人的生日对应于同一个月份. 这就是所谓的狄利克雷(Dirichlet)抽屉原理(或称鸽舍原理).

例 10 幂集 $P(X)$ 上的“交、并”运算都是集合 $P(X) \times P(X)$ 到 $P(X)$ 的映射, 但不是 $P(X)$ 到 $P(X)$ 的映射.

例 11 在集 X 上, 把 X 中的每个元素都映射为自身的映射 I (即 $\alpha \in X$, $I(\alpha) = \alpha$) 称为 X 上的恒等映射(或单位映射).

定义 1.14 设 σ 是 X 到 Y 的一个映射, 如果:

(1) 对任意的 $\alpha_1, \alpha_2 \in X$, 当 $\alpha_1 \neq \alpha_2$ 时, $\sigma(\alpha_1) \neq \sigma(\alpha_2)$, 则称 σ 为单射(或称内射, injection);

(2) $\sigma(X) = Y$, 即对于任意的 $\beta \in Y$, 存在 $\alpha \in X$, 使 $\sigma(\alpha) = \beta$, 则称 σ 为满射(或称映上的, surjection);

(3) σ 既是单射又是满射, 则称 σ 为双射(或称一一对应, bijection).

例如, 在前面的例子中, 例 6 是满射而不是单射, 例 8 是单射而不是满射; 例 9 (人数大于 12) 不是单射, 也可能不是满射; 例 7 的 f 既不是单射, 也非满射; 例 11 是双射.

有限集到有限集的映射的三种情况, 可用图 1-3 示意.

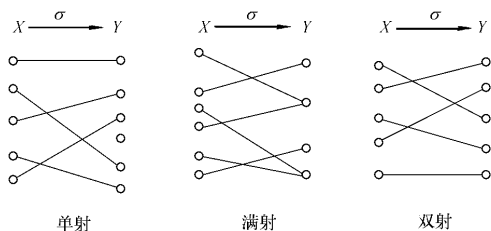


图 1-3

定义 1.15 如果 σ, τ 都是 $X \rightarrow Y$ 的映射, 且对任意的 $\alpha \in X$, $\sigma(\alpha) = \tau(\alpha)$, 则称 $\sigma = \tau$.

定义 1.16 设 σ 和 τ 分别是集合 A 到 B 和 B 到 C 的两个映射, 我们规定其乘积

$$(\tau\sigma)(\alpha) = \tau(\sigma(\alpha)), \quad \alpha \in A.$$

显然 $\tau\sigma$ 是集合 A 到 C 的一个映射.

对于集 X 到 Y 的任一个映射 σ , 显然有

$$I_Y \sigma = \sigma I_X = \sigma,$$

其中 I_X 和 I_Y 分别是 X 和 Y 的恒等映射(即单位映射).

映射的乘法满足结合律,即

如果 σ, τ, φ 分别是 $A \rightarrow B, B \rightarrow C, C \rightarrow D$ 的映射,则

$$\varphi(\tau\sigma) = (\varphi\tau)\sigma.$$

事实上,上式两边都是 $A \rightarrow D$ 的映射,而且对任意的 $\alpha \in A$ 均有

$$\begin{aligned} [\varphi(\tau\sigma)](\alpha) &= \varphi[(\tau\sigma)(\alpha)] = \varphi[\tau(\sigma(\alpha))] \\ &= (\varphi\tau)(\sigma(\alpha)) = [(\varphi\tau)\sigma](\alpha), \end{aligned}$$

因此, $\varphi(\tau\sigma) = (\varphi\tau)\sigma$ 成立.

但映射的乘法不满足交换律,例如, $f(x) = \sin x, g(x) = 1 + x$, 则

$$g(f(x)) = 1 + \sin x; f(g(x)) = \sin(1 + x),$$

故 $gf \neq fg$.

定义 1.17 设 σ 是集 X 到 Y 的一个映射,如果存在集 Y 到 X 的一个映射 τ ,使

$$\tau\sigma = I_X \quad \text{和} \quad \sigma\tau = I_Y$$

同时成立,则称 σ 是可逆映射(简称 σ 可逆),并称 τ 为 σ 的逆映射,记作 $\sigma^{-1} = \tau$.

定义中 σ 与 τ 的地位是相同的,此时也说 τ 可逆,且 $\tau^{-1} = \sigma$.

如果映射是可逆的,则其逆映射是唯一的.事实上,如果 τ_1 和 τ_2 是 σ 的两个逆映射,即

$$\tau_1\sigma = \tau_2\sigma = I_X \quad \text{且} \quad \sigma\tau_1 = \sigma\tau_2 = I_Y,$$

则有

$$\tau_1 = \tau_1 I_Y = \tau_1(\sigma\tau_2) = (\tau_1\sigma)\tau_2 = I_X\tau_2 = \tau_2,$$

故 σ 的逆映射是唯一的.

定理 1.4 集 X 到 Y 的映射 σ 可逆的充要条件是 σ 为双射.

证 条件是必要的:设 σ 可逆,即有唯一的 $\tau: Y \rightarrow X$,使

$$\tau\sigma = I_X \quad \text{且} \quad \sigma\tau = I_Y.$$

于是,对任意的 $\beta \in Y, \beta = I_Y(\beta) = (\sigma\tau)(\beta) = \sigma(\tau(\beta))$, 由于 $\tau(\beta) \in X$, 故 σ 是满射;再则,若 $\sigma(\alpha_1) = \sigma(\alpha_2)$, 就必有 $\alpha_1 = I_X(\alpha_1) = \tau(\sigma(\alpha_1)) = \tau(\sigma(\alpha_2)) = I_X(\alpha_2) = \alpha_2$, 故 σ 是单射;从而 σ 是双射.

条件是充分的:事实上,当 σ 为双射时,对任意的 $\beta \in Y$,存在唯一的 $\alpha \in X$,使 $\sigma(\alpha) = \beta$,于是可定义 $\tau: Y \rightarrow X$,使得 $\tau(\beta) = \alpha$,其中 α 是 X 中与

β 一一对应的元素, 这样, 对任意的 $\alpha \in X$, 都有

$$(\tau\sigma)(\alpha) = \tau(\sigma(\alpha)) = \tau(\beta) = \alpha,$$

所以 $\tau\sigma = I_X$; 同样, 对任意的 $\beta \in Y$, 都有

$$(\sigma\tau)(\beta) = \sigma(\tau(\beta)) = \sigma(\alpha) = \beta,$$

所以 $\sigma\tau = I_Y$. 因此 σ 是可逆的.

例 12 设 f 和 g 都是 \mathbb{Z} 到自身的映射, 且

$$f: n \mapsto n+5, g: n \mapsto n-5,$$

则 $(fg)(n) = (gf)(n) = n$, 即 fg 和 gf 都是 \mathbb{Z} 的恒等映射, 所以 f, g 都是可逆的, 它们互为逆映射.

定义 1.18 设 σ 是集 X 到 Y 的一个映射, B 是 Y 的一个子集, 我们把 X 的子集

$$A = \{\alpha \mid \alpha \in X \text{ 且 } \sigma(\alpha) \in B\}$$

称为 B 在 σ 下的完全原像(或逆像), 记作 $\sigma^{-1}(B)$.

特别, 当 $B = \{\beta\}$ 时, $\sigma^{-1}(\beta)$ 表示 β 在 X 中的所有原像的集合. 这里用记号 $\sigma^{-1}(\beta)$ 并不意味着映射 σ 是可逆的.

显然, 如果 σ 不是满射, 则存在 $\beta \in Y$, 使得 $\sigma^{-1}(\beta)$ 为空集; 如果 σ 不是单射, 则存在 $\beta \in Y$, 使得集合 $\sigma^{-1}(\beta)$ 至少含两个元素. 如果 σ 是双射, 则对任意的 $\beta \in Y$, 集合 $\sigma^{-1}(\beta)$ 都含且只含 X 中的一个元素.

下面, 我们介绍命题运算, 几何向量的运算, n 元向量与矩阵的某些运算及高斯消元法. 它们都是大学数学的基础知识.

1.6 命题运算 量词

所谓命题, 就是一个陈述句. 严格地说, 命题不是陈述句本身, 而是陈述句所表达的含义, 因为“语句”是语言学的概念, 而“命题”是逻辑学的概念. 然而“命题”是由“语句”表达的, 因此在不致引起误解的情况下, 我们就把陈述句叫做命题.

作为命题的陈述句, 它所使用的词和词组的含义都是明确的. 命题所陈述的事理总是真的或假的, 不能模棱两可. 如“1 000 是很大的数”; “南京靠近上海”; “保定是一个大城市”等语句的含义都有某种不确定性, 它们都不是传统逻辑学中的命题.

感叹句、疑问句、祈使句都不能作为命题. 如“天气真好啊!” “今天上课吗?” “请坐下!” 都不是命题, 这些句子的含义无所谓真假.

下面的语句 (1) $3 < 5$ (2) $3 \geq 5$ (3) $3 < 5$ 且 $3 \mid 5$ (4) 他学英语或者他学法语 (5) 如果天不下雨, 我就出去散步 (6) 三角形三边相等当且仅当其三内角相等都是命题. 其中(1)(2)是简单命题 (3)~(6)都是由两个命题与逻辑联结词组成的复合命题, 实际上就是命题的二元运算. 一个命题的否定也是一个命题, 命题(1)是命题(2)的否定, 这是命题的一元运算.

定义 1.19 设 P 是命题集合, $p, q \in P$, 我们规定:

$p \wedge q$ 表示命题“ p 且 q ”;

$p \vee q$ 表示命题“ p 或 q ”;

$p \rightarrow q$ 表示命题“若 p 则 q ”(或说 p 蕴涵 q);

$p \Leftrightarrow q$ 表示命题“ p 当且仅当 q ”(或说 p 与 q 等价);

$\neg p$ 表示命题“非 p ”, 即 p 的否定命题.

$\wedge, \vee, \rightarrow, \Leftrightarrow$ 是命题的二元运算, \neg 是命题的一元运算. 在逻辑学中它们称为命题的联结词, \wedge 叫合取词, \vee 叫析取词, \rightarrow 叫蕴涵词, \Leftrightarrow 叫双蕴涵词或等值词, \neg 叫否定词.

由于每个命题的含义或是真的, 或是假的, 所以可说命题有一个值, 称为真值. 真值只取“真、假”(或“1、0”)两个值, 它与开关电路中开关状态只有“通和不通”是类似的. 等价命题 p, q (即 $p \Leftrightarrow q$) 其真值是相同的.

一般把恒取“1”的命题叫“恒真命题”, 恒取“0”的命题叫“恒假命题”.

下面用表格分别表示二元运算“ \wedge ”, “ \vee ”, “ \rightarrow ”的运算规则:

$p \backslash q$	0	1
0	0	0
1	0	1

“ $p \vee q$ ”

$p \backslash q$	0	1
0	0	0
1	0	1

“ $p \wedge q$ ”

$p \backslash q$	0	1
0	1	1
1	0	1

“ $p \rightarrow q$ ”

$p \backslash q$	0	1
0	1	1
1	0	1

“ $\neg p \vee q$ ”

注意 (1) 对于命题 $p \rightarrow q$, 当前题 p 为假时, 不管 q 是真还是假, 都不好说 $p \rightarrow q$ 是假的, 因而把它规定为真的, 这称为“善意的推定”. 例如“若太阳从西边出, 则人永不死”; “若太阳从西边出, 则人都要死”, 它们都不好说是假的, 按“善意推定”的原则, 它们都是真的. $p \rightarrow q$ 也可说 p 是 q 成立的充分条件, q 是 p 成立的必要条件.

为了省略命题运算中的一些括号, 我们规定运算的先后次序为 $\neg, \wedge, \vee, \rightarrow, \Leftrightarrow$.

$\vee, \rightarrow, \Leftrightarrow$.

(2) 由运算表可以看出: “ $p \rightarrow q$ ” 和 “ $\neg q \vee q$ ” 两个命题取值完全相同, 因此它们是一样的.

(3) 设 t 是一个恒真命题, f 是一个恒假命题, 则对一切 p :

$$t \vee p = t, \quad t \wedge p = p;$$

$$f \vee p = p, \quad f \wedge p = f.$$

(4) 由上述运算表, 容易得到下列结论: $p \wedge \neg p = f, p \vee \neg p = t$ (这两个等式分别被称为“矛盾律”和“排中律”).

由五种命题运算的法则, 容易证明命题运算具有以下运算律:

(1) $p \vee q \Leftrightarrow q \vee p, p \wedge q \Leftrightarrow q \wedge p$ (交换律)

(2) $(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r),$

$(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$ (结合律)

(3) $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r),$

$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ (分配律)

(4) $\neg(\neg p) \Leftrightarrow p$ (否定律)

(5) $p \rightarrow q \Leftrightarrow \neg q \rightarrow \neg p, p \rightarrow q \Leftrightarrow \neg p \vee q;$

(6) $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q, \neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q.$

证 我们只证(5).

(5) 命题 $p \rightarrow q$ 为假, 当且仅当命题 p 真 q 假; $\neg q \rightarrow \neg p$ 为假, 当且仅当 $\neg q$ 真 $\neg p$ 假, 即 p 真 q 假; $\neg p \vee q$ 为假, 当且仅当 $\neg p$ 与 q 都假, 即 p 真 q 假. 这三个命题真值相同, 所以它们是等价的.

例如: “如果下雨, 我就带伞” 等价于 “如果我不带伞, 就不下雨”, 也等价于 “要么不下雨, 要么带伞”. 但不等价于 “如果不下雨, 我就不带伞”. 一般而论, 命题 $p \rightarrow q$ 与命题 $\neg p \rightarrow \neg q$ 是不等价的.

我们用反证法证明一个数学定理 “若 p 则 q ”, 就是证明它的等价命题 (即逆否命题) “若非 q , 则非 p ”.

再如: 若 p 则 q , 且若 r 则 p , 可推出若 r 则 q . 这样的命题, 符号形式为

$$(p \rightarrow q) \wedge (r \rightarrow p) \rightarrow (r \rightarrow q),$$

这是一个真值为恒真的命题. (它就是有名的 “三段论”). 因为 $(p \rightarrow q) \wedge (r \rightarrow p)$ 为真, 当且仅当 $(p \rightarrow q)$ 和 $(r \rightarrow p)$ 都真, 即 $\neg p \vee q$ 和 $\neg r \vee p$ 都真. 此时, 若 r 为真 (即 $\neg r$ 为假), 则 p 为真 ($\neg p$ 为假), 从而必有 q 为真, 即 $(r \rightarrow q)$ 为真. 若 r 为假, q 可真可假, 但 $(r \rightarrow q)$ 为真, 故是恒真命题. 读者不妨利用运算律推演这个式子.

对比一下集合运算与命题运算所适合的运算律, 读者不难发现, 如果把命题的五种运算 $\neg, \wedge, \vee, \rightarrow, \Leftrightarrow$ 依次对应于集合的余, 交, 并, 包含和相等, 则它们的运算律是 “相同的”.

最后再介绍一下量词 \forall, \exists .

有些命题,常用两种断言:“集 X 中每个元素具有性质 p ”;“集 X 中有一个元素 x 具有性质 p ”.为表述简便起见,我们用逻辑符号:“ $\forall x \in X, p$ ”(或“($\forall x \in X$) p ”);“ $\exists x \in X, p$ ”(或“($\exists x \in X$) p ”)来分别表示.这里,“ $\forall x$ ”表示“对于任意的元素 x ”;“ \forall ”叫做全称量词,它是 Any 字头 A 的倒写;“ \exists ”叫做存在量词,表示“存在”,它是 Exist 字头 E 的反写.例如对于集合 A 与 B , $A \subset B$ 的含义是“若 $a \in A$ 则 $a \in B$ ”,即“ $\forall a(a \in A \rightarrow a \in B)$ ”也可表述为“ $\forall a \in A, a \in B$ ”.

$A \subset B$ 的否定为 $A \not\subset B$,其含义是“ $\exists a \in A, a \notin B$ ”.

“每一个(\forall)人都要死”的否定命题为“有(\exists)人不会死”.

“甲班有(\exists)学生得奖学金”的否定命题为“甲班所有学生(即任何一个学生)都未得奖学金”.

一般,含有量词的命题的否定命题,满足下面两个基本的等价规则:

$$\neg(\forall x \in X) p \Leftrightarrow (\exists x \in X) \neg p, \quad (1-1)$$

$$\neg(\exists x \in X) p \Leftrightarrow (\forall x \in X) \neg p. \quad (1-2)$$

数学中,常遇到含有一个或多个量词的命题,借助于上述规则便能对其否命题作机械运算.例如:

设 X 是甲班学生的集合, Y 是乙班学生的集合, $p(x, y)$ 表示 x 和 y 同姓(其中 $x \in X, y \in Y$) 则命题 $(\forall x \in X)(\exists y \in Y) p(x, y)$ 表示“所有甲班的学生都能在乙班找到与之同姓的学生”.按上述(1-1)(1-2)规则,由

$$\begin{aligned} \neg(\forall x \in X)(\exists y \in Y) p(x, y) &\Leftrightarrow (\exists x \in X) \neg(\exists y \in Y) p(x, y) \\ &\Leftrightarrow (\exists x \in X)(\forall y \in Y) \neg p(x, y) \end{aligned}$$

可知,上述命题的否命题的含义为“甲班有一学生与乙班所有学生都不同姓”.

例如,数列 $\{u_n\}$ 以 a 为极限,用 $\epsilon - N$ 语言可定义为:

“对任意的 $\epsilon > 0$, 存在 $N > 0$, 使对任何 $n > N$, 恒有 $|u_n - a| < \epsilon$ ”,这可用多个量词组合为命题 I :

$$(\forall \epsilon > 0)(\exists N > 0)[(\forall n > N)|u_n - a| < \epsilon],$$

于是数列 $\{u_n\}$ 不以 a 为极限的命题为 $\neg(I)$, 即

$$\begin{aligned} \neg(I) &\Leftrightarrow (\exists \epsilon > 0) \neg(\exists N > 0)[(\forall n > N)|u_n - a| < \epsilon] \\ &\Leftrightarrow (\exists \epsilon > 0)(\forall N > 0) \neg[(\forall n > N)|u_n - a| < \epsilon] \\ &\Leftrightarrow (\exists \epsilon > 0)(\forall N > 0)(\exists n > N) \neg|u_n - a| < \epsilon \\ &\Leftrightarrow (\exists \epsilon > 0)(\forall N > 0)(\exists n > N)|u_n - a| \geq \epsilon, \end{aligned}$$

即“存在 $\varepsilon > 0$,对任何 $N > 0$,存在 $n > N$,使 $|u_n - a| \geq \varepsilon$.”

1.7 几何向量的运算 空间直角坐标系

描述质点的位移、速度、加速度和作用于物体的力、力矩等类型的量 ,既要指出大小 ,也要指明方向 .这种既有大小又有方向的量称为向量(或矢量) .

向量有两个特征——大小和方向 .方向是一个几何性质 ,它反映从一点 O 到另一点 P 的顺序关系 ,而两点之间又有一个距离 ,因此 ,常用有向线段 \overrightarrow{OP} 表示向量 ,用其长度 $|\overrightarrow{OP}|$ 表示向量的大小(或称为向量的模) ,用箭头 \rightarrow 表示向量的方向 ,即端点 $O \rightarrow P$ 所指的方向 ,端点 O, P 分别叫向量的起点和终点 .今后也常用黑体字 a, b, x 等表示向量 .

用有向线段表示的向量常称为“几何向量” .我们在空间所有几何向量组成的集合中定义一个二元关系 R .

$$aRb \Leftrightarrow |a| = |b|, a // b \text{ 且方向相同.}$$

容易验证 R 是一个等价关系 .这里不考虑 a, b 向量的起点(通常把不考虑起点的几何向量称为自由向量) ,只要它们的大小相等、方向相同就称其为等价向量(或相等的向量 ,即 $a = b$) .等价关系 R 把空间几何向量分为等价类 ,每个等价类中可任取一个代表元素(向量) ,为方便起见 ,我们选以定点 O 为起点的向量为代表元(如图 1-4 中 \overrightarrow{OP} 和 \overrightarrow{OQ} 分别是两个等价类的代表元向量) .以后如不指明空间几何向量的起点都认为起点在原点 .

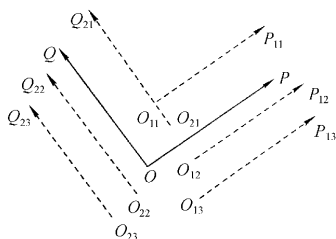


图 1-4

与 a 大小相等、方向相反的向量称为 a 的反向量(或负向量) ,记作 $-a$ (如图 1-5 所示) .显然 ,

$$-(-a) = a.$$

大小为零的向量叫零向量,记作 $\mathbf{0}$. 零向量没有确定的方向,或说方向任意.

向量的基本运算有加法、数量乘法(即数乘向量)、内积、外积和混合积,它们主要源于力学问题.

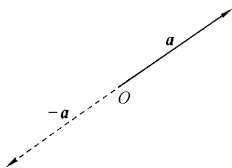


图 1-5

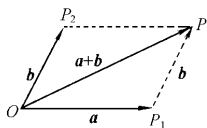


图 1-6

1.7.1 向量的加法

根据质点由 O 点移至 P_1 点,再由 P_1 点移至 P 点所得总位移 \overrightarrow{OP} ,以及由两个力求合力的方法,我们定义向量的加法如下.

定义 1.20 设 $a = \overrightarrow{OP_1}$, $b = \overrightarrow{OP_2}$, 我们规定 $a + b$ 是一个向量,它是以 a , b 为邻边的平行四边形中由点 O 与其相对的顶点 P 联成的向量 \overrightarrow{OP} ,即 $a + b = \overrightarrow{OP}$ (见图 1-6).

在图 1-6 中 $\overrightarrow{P_1P} = \overrightarrow{OP_2} = b$, 所以,将 a 的终点与 b 的起点重合在一起,则由 a 的起点与 b 的终点相连的向量 $\overrightarrow{OP} = a + b$,这叫求两个向量之和的三角形法(见图 1-6 中右边的三角形部分).定义中求 $a + b$ 的方法叫平行四边形法.

按定义作图,容易证明向量加法满足下列运算规则:

- (1) $a + b = b + a$ (交换律)
- (2) $(a + b) + c = a + (b + c)$ (结合律)
- (3) $a + \mathbf{0} = a$;
- (4) $a + (-a) = \mathbf{0}$.

由于向量加法满足结合律,所以多个向量相加,不必加括号指明相加的次序.求多个向量 a_1, a_2, \dots, a_n 之和,可利用三角形法,将 a_i 的终点与 a_{i+1} 的起点重合在一起($i = 1, 2, \dots, n-1$)则 a_1 的起点与 a_n 的终点相连并指向后者的向量等于 $a_1 + a_2 + \dots + a_n$.当 $n = 4$ 时,如图 1-7 所示.

向量运算常要用到加法的逆运算,即向量的减法,我们定义

$$a - b = a + (-b).$$

求 $a - b$ 的方法,见图 1-8.

根据加法的运算规则和减法的定义,容易得到向量方程 $a + x = b$ 的解为 $x = b - a$.

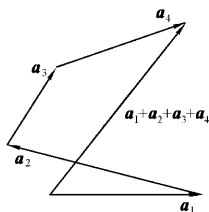


图 1-7

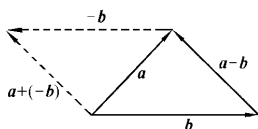


图 1-8

注意,不要把向量与数混淆.例如实数是有序的,即可比大小,而向量分子 $a < b$ 无意义.当然,向量的长度可以比大小,根据三角形两边长度之和不小于第三边长度, $a, b, a + b$ 的长度满足三角不等式

$$|a + b| \leq |a| + |b|.$$

1.7.2 向量与数量的乘法(简称数乘向量)

如果 $a \parallel b$ 且 $|b| = 2|a|$,那么当 a, b 同向时, b 可表示为 $2a$;当 a, b 反向时, b 可表示为 $-2a$.一般为了描述向量的“伸缩”,我们定义实数与向量的乘法(也称数量乘法)如下:

定义 1.21 设 $\lambda \in \mathbf{R}$,定义 λ 与 a 的乘积 λa 是一个向量,其大小 $|\lambda a| = |\lambda| |a|$;其方向为: $\lambda > 0$ 时, λa 与 a 同向, $\lambda < 0$ 时, λa 与 a 反向.

由数乘向量的定义,容易证明数乘向量的运算满足以下规则:

- (1) $1a = a$ ($-1)a = -a$;
- (2) $\mu(\lambda a) = (\mu\lambda)a$;
- (3) $(\mu + \lambda)a = \mu a + \lambda a$;
- (4) $\lambda(a + b) = \lambda a + \lambda b$.

长度为 1 的向量叫做单位向量.由于单位向量的长度已确定,所以,单位向量的特征就是方向.因此一个单位向量确定了一个方向.反之,给定一个方向,在这个方向上也唯一确定一个单位向量.也就是说,方向与单位向量是一一对应的.

当 $a \neq 0$ 时, $\frac{a}{|a|}$ 是 a 的方向的单位向量,记作 a^0 ,于是 $a = |a|a^0$,这就

明确表达了一个向量的长度和方向. 但要注意, 式子 $\frac{1}{a} \cdot \frac{b}{a}$ (其中 a, b 是向量) 都没有意义.

1.7.3 向量的内积

大家知道力 F 作用在质点上使之产生位移 S , 力所作的功为 $|F| |S| \cos \theta$, 其中 θ 为 F 与 S 的夹角(如图 1-9).

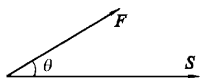


图 1-9

两个向量按这样的算式对应于一个实数, 这在很多问题中都有用. 于是我们把它抽象为两个向量的内积(或点积, 或数量积).

定义 1.22 向量 a 与 b 的内积 $a \cdot b$ 是一个实数, 且

$$a \cdot b = |a| |b| \cos \theta,$$

其中 θ 为 a 与 b 的夹角(以后 θ 也常记作 a, b), 并规定 $0 \leq \theta \leq \pi$. 若 a, b 有一个是零向量, 则规定 $a \cdot b = 0$.

由定义可见, 非零向量 a, b 的内积 $a \cdot b > 0$ (< 0) 当且仅当 a, b 为锐(钝)角; $a \cdot b = 0$ 当且仅当 $a \perp b$ (也说 a 与 b 正交). 由于零向量与任何向量的内积为零, 所以也说零向量与任何向量正交.

由内积的定义可得:

$$|a| = \sqrt{a \cdot a};$$

$$\cos a, b = \frac{a \cdot b}{|a| |b|}. \quad (1-3)$$

$a \cdot a$ 也常记作 a^2 , 所以, $|a|^2 = a^2$.

向量的内积满足以下运算规则:

- (1) $a \cdot b = b \cdot a$; (交换律)
- (2) $(\lambda a) \cdot b = \lambda(a \cdot b)$;
- (3) $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$; (分配律)
- (4) $a \cdot a \geq 0$, 等号成立当且仅当 $a = 0$.

由定义容易证明规则(1)(2)(4). 下面证明(3), 我们仅就夹角 a, c , b, c , $a + b, c$ 均小于 $\frac{\pi}{2}$ 的情形给以证明. 如图 1-10.

记 $a = \overrightarrow{OA}$, $b = \overrightarrow{AB}$, $c = \overrightarrow{OC}$, 则 $a + b = \overrightarrow{OB}$.

$$\begin{aligned}
 (a+b) \cdot c &= |a+b| |c| \cos \theta = |\overrightarrow{OQ}| |c| = (|\overrightarrow{OP}| + |\overrightarrow{PQ}|) |c| \\
 &= |\overrightarrow{OP}| |c| + |\overrightarrow{PQ}| |c| = a \cdot c + b \cdot c
 \end{aligned}$$

对于一般情形的证明, 需要利用向量的投影定理(略去).

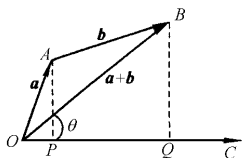


图 1-10

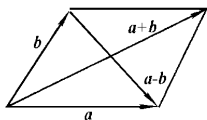


图 1-11

要注意由 $a \cdot b = 0$ 不能得到 $a = 0$ 或 $b = 0$, 也不能由 $a \cdot b = a \cdot c$ 推出 $b = c$.

例 1 利用向量内积证明恒等式:

$$|a+b|^2 + |a-b|^2 = 2|a|^2 + 2|b|^2,$$

即平行四边形对角线的平方和等于四边的平方和(如图 1-11).

$$\begin{aligned}
 \text{证} \quad & |a+b|^2 + |a-b|^2 \\
 &= (a+b) \cdot (a+b) + (a-b) \cdot (a-b) \\
 &= (a \cdot a + 2a \cdot b + b \cdot b) + (a \cdot a - 2a \cdot b + b \cdot b) \\
 &= 2a \cdot a + 2b \cdot b = 2|a|^2 + 2|b|^2.
 \end{aligned}$$

其中第二个等号成立是利用了内积运算规则(1)(2)(3)及向量减法定义和 $-b = (-1)b$.

1.7.4 向量的外积

除了内积, 向量还有另一种乘法, 称为向量的外积(也称叉积或向量积). 它的定义如下:

定义 1.23 向量 a 与 b 的外积 $a \times b$ 是一个向量, 其长度为

$$|a \times b| = |a| |b| \sin \angle a, b.$$

$a \times b$ 的方向为 (1) $a \times b$ 与 a, b 都垂直 (2) $a, b, a \times b$ 按“右手法则”确定 $a \times b$ 的指向, 即把 $a, b, a \times b$ 的起点放在一起, 将右手的四指(不含拇指)伸开由 a 转到 b 转过的角度为 $\angle a, b$, 此时张开的拇指(与四指垂直)的指向就是 $a \times b$ 的方向(这种由 a, b 确定 $a \times b$ 的指向的方法称为“右手法则”). 若 a, b 有一个是零向量, 规定 $a \times b = 0$.

由定义可见, $|a \times b|$ 的几何意义是以 a, b 为邻边的平行四边形的面积.

向量的外积满足以下运算规则：

$$(1) a \times b = -(b \times a);$$

$$(2) (\lambda a) \times b = \lambda(a \times b);$$

$$(3) (a + b) \times c = a \times c + b \times c.$$

其中(1)(2)容易证明(3)表示叉积对向量加法满足右分配律,我们略去它的证明.

读者不难利用上述三条运算规则,证明下列等式：

$$a \times (\lambda b) = \lambda(a \times b);$$

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

下面我们再介绍向量外积的一个物理背景.

设有一个圆锥形刚体(所谓刚体就是在运动过程中不变形的物体)以等角速度 ω 按逆时针方向绕中心轴 L 转动, ω 的方向按“右手法则”规定为伸开的拇指的指向(如图 1-12 所示).此时,点 P 的线速度 v 的大小为

$$|\omega| |\overrightarrow{O_1 P}| = |\omega| |\overrightarrow{OP}| \sin \theta = |\omega \times \overrightarrow{OP}|.$$

(其中 $\theta = \angle \omega, \overrightarrow{OP}$), v 的方向恰为 $\omega \times \overrightarrow{OP}$ 的方向,所以

$$v = \omega \times \overrightarrow{OP}.$$

这里的参考点 O 可在轴 L 上任选,如以 O_1 为参考点,则

$$v = \omega \times \overrightarrow{O_1 P}.$$

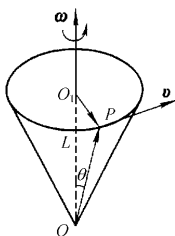


图 1-12

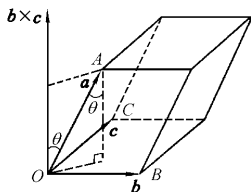


图 1-13

1.7.5 向量的混合积

定义 1.24 对向量 a, b, c , 先将 b, c 作外积, 再将其与 a 作内积, 即

$$a \cdot (b \times c),$$

称为向量的混合积.

向量 a, b, c 的混合积是一个数量, 它的绝对值 $|a \cdot (b \times c)|$ 的几何意义

是以 a, b, c 为邻边的平行六面体的体积(图 1-13). 事实上,

$$V = \text{高} \times \text{底面积}$$

$$= h |(b \times c)| = |a| |\cos \theta| |(b \times c)| = |a \cdot (b \times c)|,$$

其中 $\theta \in (0, \pi)$, 当 $\theta > \frac{\pi}{2}$ 时, $\cos \theta < 0$, $h = |a| |\cos \theta|$.

向量内积和外积所满足的运算规则, 对混合积中相应的运算也都适用, 例如

$$a \cdot (c \times b) = a \cdot (-b \times c) = -a \cdot (b \times c).$$

向量的混合积还有下面重要的性质(证明见本节 1.7.7 段)

$$a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b)$$

根据向量的混合积的定义, 读者不难证明, 向量 a, b, c 共面的充要条件是,

$$a \cdot (b \times c) = 0.$$

向量共面的另一个充要条件是, 存在不全为零的常数 k_1, k_2, k_3 , 使

$$k_1 a + k_2 b + k_3 c = 0.$$

几何向量的运算实际上是把有向线段作为运算对象. 向量的加法和外积是定义在向量集 V 上取值于 V 的二元运算; 内积和混合积分别是定义在 V 上取值于实数集 \mathbf{R} 的二元运算和三元运算; 数量乘法是定义在 $\mathbf{R} \times V$ 上取值于 V 的一元运算. 虽然我们由五种运算的定义, 得到了相应的运算规则, 但以有向线段作运算却很不方便.

下面我们进一步讨论如何把几何向量的运算化为数的代数运算. 为此要引入空间直角坐标系中点的坐标概念, 并给出几何向量的坐标表示式.

1.7.6 空间直角坐标系与几何向量的坐标表示式

大家知道, 平面上点的位置可与两个实数组成的有序数组 (x, y) 一一对应. 同样空间中点的位置可与三个实数组成的有序数组 (x, y, z) 一一对应, 其对应规则如下:

在空间中取交于 O 点的三条互相垂直的有向直线(取定它们的正方向并给定长度单位), 排定次序, 依次记作 Ox, Oy 和 Oz , 它们就构成一个空间直角坐标系, 记作 $Oxyz$. 点 O 叫做坐标原点, 三条互相垂直的有向直线叫做坐标轴, 依次称为 x 轴, y 轴和 z 轴. 每两条轴所决定的平面叫做坐标平面, 三个坐标平面分别垂直于 z 轴, x 轴和 y 轴. 常用的空间直角坐标系是“右手系”, 即 Ox, Oy 和 Oz 的方向符合“右手规则”.

确定一个点 P 在空间直角坐标系 $Oxyz$ 中坐标的方法为: 过点 P 作平行

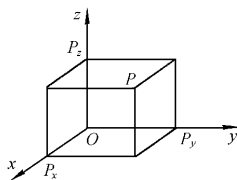


图 1-14

于坐标平面的三个平面(图 1-14), 它们分别与 x 、 y 、 z 轴交于点 P_x 、 P_y 、 P_z (即点 P 在三个坐标轴上的投影点), 它们在 x 、 y 、 z 轴上的坐标依次记作 x 、 y 、 z . 这样, 点 P 就唯一地确定了一个三元有序数组 (x, y, z) ; 反之, 任给一个三元有序数组, 按同样的含义也唯一确定了 $Oxyz$ 中的一个点. 因此在空间直角坐标系中, 点与三元有序数组之间建立了一一对应的关系. 于是我们把与点 P 对应的 (x, y, z) 叫做点 P 在 $Oxyz$ 中的坐标, 记作 $P(x, y, z)$.

例如, 在 $Oxyz$ 中点 $P(3, 3, 4)$ 与点 $Q(-2, 3, -3)$ 的位置如图 1-15 所示.

下面, 我们介绍空间直角坐标系中几何向量的坐标表示式. 为此先把 x 、 y 、 z 轴的正方向依次对应于三个单位向量 e_1 、 e_2 、 e_3 (或表示为 i 、 j 、 k), 并称之为基本向量, 它们具有性质: $e_i \cdot e_i = 1$, $e_i \cdot e_j = 0$ ($i \neq j$), $i, j = 1, 2, 3$. 任一空间几何向量都可用基本向量的数乘和加法运算来表示.

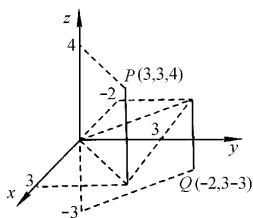


图 1-15

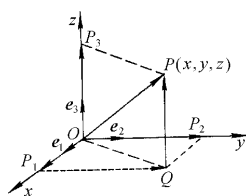


图 1-16

设 $r = \overrightarrow{OP}$, 点 $P(x, y, z)$ 在 x 、 y 、 z 轴上的投影点为 P_1 、 P_2 、 P_3 , 点 P 在坐标平面 xOy 上的投影点为 Q (图 1-16), 于是, 按几何向量的加法和数乘, 有

$$\begin{aligned}
 \vec{r} &= \overrightarrow{OP} = \overrightarrow{OP_1} + \overrightarrow{P_1Q} + \overrightarrow{QP} \\
 &= \overrightarrow{OP_1} + \overrightarrow{OP_2} + \overrightarrow{OP_3} = xe_1 + ye_2 + ze_3.
 \end{aligned} \quad ①$$

上式称为基本向量组 $\{e_1, e_2, e_3\}$ 的线性组合, 其中系数 x, y, z 恰是点 P 的坐标. 我们把这个三元有序数组 (x, y, z) 也称为向量关于基本向量组 $\{e_1, e_2, e_3\}$ 的坐标(简称向量的坐标). 所以向量(起点在原点)的坐标与向量终点 P 的坐标是一样的.

零向量关于基本向量组 $\{e_1, e_2, e_3\}$ 的坐标为 $(0, 0, 0)$, 即

$$xe_1 + ye_2 + ze_3 = \mathbf{0} \Leftrightarrow x = y = z = 0. \quad ②$$

事实上, 当后者成立时, 前者显然成立; 反之, 当前者成立时, 必有

$$(xe_1 + ye_2 + ze_3) \cdot e_1 = \mathbf{0} \cdot e_1 = 0,$$

于是

$$xe_1 \cdot e_1 + ye_2 \cdot e_1 + ze_3 \cdot e_1 = x = 0.$$

同理可证 $y = 0, z = 0$.

利用命题②易知, 两个相等向量的坐标也是相同的, 即

若 $a = a_1e_1 + a_2e_2 + a_3e_3, b = b_1e_1 + b_2e_2 + b_3e_3$, 则

$$a = b \Leftrightarrow a_i = b_i \quad (i = 1, 2, 3). \quad ③$$

事实上, $a = b$ 即 $a - b = \mathbf{0}$, 而

$$a - b = (a_1 - b_1)e_1 + (a_2 - b_2)e_2 + (a_3 - b_3)e_3,$$

于是由命题②立即可得命题③.

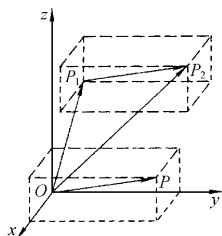


图 1-17

对于起点不在原点的向量 $\overrightarrow{P_1P_2}$, 如果点 P_1, P_2 的坐标分别为 $(x_1, y_1, z_1), (x_2, y_2, z_2)$ (如图 1-17 所示), 则

$$\begin{aligned}
 \overrightarrow{P_1P_2} &= \overrightarrow{OP_2} - \overrightarrow{OP_1} = (x_2e_1 + y_2e_2 + z_2e_3) - (x_1e_1 + y_1e_2 + z_1e_3) \\
 &= (x_2 - x_1)e_1 + (y_2 - y_1)e_2 + (z_2 - z_1)e_3.
 \end{aligned} \quad ④$$

这个向量与起点在原点 O , 方向及大小都相同的向量 \overrightarrow{OP} 是被看作相等的, 即

$$\overrightarrow{P_1P_2} = \overrightarrow{OP} = xe_1 + ye_2 + ze_3.$$

根据命题 ③ 就有 $x = x_2 - x_1$, $y = y_2 - y_1$, $z = z_2 - z_1$, 因此, 我们把 ④ 式中的三元有序数组 $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$ 叫做向量 $\overrightarrow{P_1P_2}$ 关于基本向量组 $\{e_1, e_2, e_3\}$ 的坐标.

建立了几何向量关于基本向量组 $\{e_1, e_2, e_3\}$ 的坐标概念, 可把

$$a = a_1e_1 + a_2e_2 + a_3e_3$$

记作 $a = (a_1, a_2, a_3)$ 并称其为向量 a 的坐标表示式. 例如:

$$e_1 = 1e_1 + 0e_2 + 0e_3 = (1, 0, 0);$$

$$e_2 = 0e_1 + 1e_2 + 0e_3 = (0, 1, 0);$$

$$e_3 = 0e_1 + 0e_2 + 1e_3 = (0, 0, 1);$$

$$0 = 0e_1 + 0e_2 + 0e_3 = (0, 0, 0).$$

1.7.7 向量五种运算在坐标表示式下的算式

引入了向量的坐标表示式, 就可把向量的五种运算化为向量坐标的代数运算, 它们的计算公式如下:

设 $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$, $c = (c_1, c_2, c_3)$ 则

$$(1) a + b = (a_1 + b_1, a_2 + b_2, a_3 + b_3);$$

$$(2) \lambda a = (\lambda a_1, \lambda a_2, \lambda a_3);$$

$$(3) a \cdot b = a_1b_1 + a_2b_2 + a_3b_3;$$

$$(4) a \times b = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1);$$

$$(5) a \cdot (b \times c) = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1).$$

其中(4)(5)式常用三阶行列式表示为:

$$a \times b = \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} e_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} e_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} e_3;$$

$$a \cdot (b \times c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$$

其中二阶行列式的定义为:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

(关于 n 阶行列式的定义、性质及其计算,将在第 5 章详述.)

(1)(2)式的证明留给读者练习,下面证明(3)(4)(5)式:

$$\begin{aligned} \boldsymbol{a} \cdot \boldsymbol{b} &= (a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3) \cdot (b_1 \boldsymbol{e}_1 + b_2 \boldsymbol{e}_2 + b_3 \boldsymbol{e}_3) \\ &= (a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3) \cdot (b_1 \boldsymbol{e}_1 + b_2 \boldsymbol{e}_2 + b_3 \boldsymbol{e}_3) \\ &= a_1 b_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 + a_1 b_2 \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 + a_1 b_3 \boldsymbol{e}_1 \cdot \boldsymbol{e}_3 + a_2 b_1 \boldsymbol{e}_2 \cdot \boldsymbol{e}_1 \\ &\quad + a_2 b_2 \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 + a_2 b_3 \boldsymbol{e}_2 \cdot \boldsymbol{e}_3 + a_3 b_1 \boldsymbol{e}_3 \cdot \boldsymbol{e}_1 + a_3 b_2 \boldsymbol{e}_3 \cdot \boldsymbol{e}_2 + a_3 b_3 \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3; \end{aligned}$$

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= (a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3) \times (b_1 \boldsymbol{e}_1 + b_2 \boldsymbol{e}_2 + b_3 \boldsymbol{e}_3) \\ &= a_1 b_1 \boldsymbol{e}_1 \times \boldsymbol{e}_1 + a_2 b_2 \boldsymbol{e}_2 \times \boldsymbol{e}_2 + a_3 b_3 \boldsymbol{e}_3 \times \boldsymbol{e}_3 + a_1 b_2 \boldsymbol{e}_1 \times \boldsymbol{e}_2 + a_2 b_1 \boldsymbol{e}_2 \times \boldsymbol{e}_1 \\ &\quad + a_2 b_3 \boldsymbol{e}_2 \times \boldsymbol{e}_3 + a_3 b_2 \boldsymbol{e}_3 \times \boldsymbol{e}_2 + a_3 b_1 \boldsymbol{e}_3 \times \boldsymbol{e}_1 + a_1 b_3 \boldsymbol{e}_1 \times \boldsymbol{e}_3. \end{aligned}$$

由于 $|\boldsymbol{e}_1 \times \boldsymbol{e}_1| = 0$, 所以 $\boldsymbol{e}_1 \times \boldsymbol{e}_1 = \mathbf{0}$; 同理 $\boldsymbol{e}_2 \times \boldsymbol{e}_2 = \mathbf{0}$, $\boldsymbol{e}_3 \times \boldsymbol{e}_3 = \mathbf{0}$; $|\boldsymbol{e}_1 \times \boldsymbol{e}_2| = 1$, $\boldsymbol{e}_1 \times \boldsymbol{e}_2$ 与 \boldsymbol{e}_3 同向, 所以 $\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3$, $\boldsymbol{e}_2 \times \boldsymbol{e}_1 = -\boldsymbol{e}_3$; 同理 $\boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1$, $\boldsymbol{e}_3 \times \boldsymbol{e}_2 = -\boldsymbol{e}_1$, $\boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2$, $\boldsymbol{e}_1 \times \boldsymbol{e}_3 = -\boldsymbol{e}_2$. 于是

$$\boldsymbol{a} \times \boldsymbol{b} = (a_2 b_3 - a_3 b_2) \boldsymbol{e}_1 + (a_3 b_1 - a_1 b_3) \boldsymbol{e}_2 + (a_1 b_2 - a_2 b_1) \boldsymbol{e}_3$$

利用(3)(4)的结论, 即得(5)式.

利用(5)的结论, 读者容易证明混合积的下列性质:

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

根据内积定义和上述式(3), 对于向量 $\boldsymbol{a} = (a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3)$, 易得:

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

$$\cos \angle(\boldsymbol{a}, \boldsymbol{e}_1) = \frac{\boldsymbol{a} \cdot \boldsymbol{e}_1}{|\boldsymbol{a}| |\boldsymbol{e}_1|} = \frac{a_1}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

同理

$$\cos \angle(\boldsymbol{a}, \boldsymbol{e}_2) = \frac{a_2}{\sqrt{a_1^2 + a_2^2 + a_3^2}};$$

$$\cos \angle(\boldsymbol{a}, \boldsymbol{e}_3) = \frac{a_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

因此, 给定了向量 \boldsymbol{a} 的坐标表示式, 既确定了 \boldsymbol{a} 的长度, 也确定了 \boldsymbol{a} 与三个坐标轴之间的夹角, 依次记作 α, β, γ , 从而确定了 \boldsymbol{a} 的方向.

向量 $\boldsymbol{a} = (a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3)$ 与三坐标轴之间的夹角 α, β, γ 称为 \boldsymbol{a} 的方向角,

$\cos \alpha, \cos \beta, \cos \gamma$ 称为 \boldsymbol{a} 的方向余弦, 其坐标表示式为:

$$\cos \alpha = \frac{a_1}{\sqrt{a_1^2 + a_2^2 + a_3^2}}, \quad \cos \beta = \frac{a_2}{\sqrt{a_1^2 + a_2^2 + a_3^2}},$$

$$\cos \gamma = \frac{a_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

显然

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1.$$

例2 已知:点 $P_1(1, 1, 1), P_2(2, 3, -1), P_3(1, 3, 5)$. 试求

(1) $\triangle P_1 P_2 P_3$ 的面积;

(2) 与点 P_1, P_2, P_3 所确定的平面相垂直的单位向量及其方向余弦.

解 取 $\mathbf{a} = \overrightarrow{P_1 P_2} = (2-1)\mathbf{e}_1 + (3-1)\mathbf{e}_2 + (-1-1)\mathbf{e}_3 = (1, 2, -2),$

$$\mathbf{b} = \overrightarrow{P_1 P_3} = (1-1)\mathbf{e}_1 + (3-1)\mathbf{e}_2 + (5-1)\mathbf{e}_3 = (0, 2, 4).$$

(1) $\triangle P_1 P_2 P_3$ 的面积

$$S = \frac{|\mathbf{a} \times \mathbf{b}|}{2} = \frac{|(1, 2, -2) \times (0, 2, 4)|}{2} = \frac{1}{2} \sqrt{(12)^2 + (-4)^2 + 2^2} = \sqrt{41}.$$

(2) $\mathbf{a} \times \mathbf{b}$ 与 $\mathbf{b} \times \mathbf{a}$ (即 $-\mathbf{a} \times \mathbf{b}$) 垂直于点 P_1, P_2, P_3 所确定的平面, 记 $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, 则

$$\begin{aligned} \mathbf{c}^0 &= \frac{(\mathbf{a} \times \mathbf{b})}{|\mathbf{a} \times \mathbf{b}|} = \left(\frac{6}{\sqrt{41}}, \frac{-2}{\sqrt{41}}, \frac{1}{\sqrt{41}} \right), \\ -\mathbf{c}^0 &= -\left(\frac{6}{\sqrt{41}}, \frac{-2}{\sqrt{41}}, \frac{1}{\sqrt{41}} \right) \end{aligned}$$

是所求的两个单位向量, 它们的方向余弦分别为:

$$\begin{aligned} \cos \alpha &= \frac{6}{\sqrt{41}}, \cos \beta = \frac{-2}{\sqrt{41}}, \cos \gamma = \frac{1}{\sqrt{41}}, \\ \cos \alpha' &= -\frac{6}{\sqrt{41}}, \cos \beta' = \frac{2}{\sqrt{41}}, \cos \gamma' = -\frac{1}{\sqrt{41}}. \end{aligned}$$

1.8 n 元向量的线性运算 高斯消元法

在很多实际问题中, 有些研究对象也可用多个数组成的有序数组来描述. 例如:

由 n 个未知数 x_1, x_2, \dots, x_n 组成的 n 元一次方程式

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b$$

可用其系数及常数 b 排成的有序数组 $(a_1, a_2, \dots, a_n, b)$ 来表示.

按升幂排成的 n 次多项式

$$P(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

也可用其系数组成的 $n+1$ 元有序数组 $(a_0, a_1, a_2, \dots, a_n)$ 来表示.

在由 m 个节点和 n 条支路组成的电路中, m 个节点的电位和 n 条支路的电流也分别可由 m 元和 n 元有序数组来描述.

定义 1.25 由 n 个数 a_1, a_2, \dots, a_n 组成的有序数组称为 n 元向量, 记作 (a_1, a_2, \dots, a_n) , 其中 a_i 称为 n 元向量的第 i 个分量. 如果 $a_i (i = 1, 2, \dots, n)$ 是实(复)数, 叫做实(复)向量. 如果 n 个分量全为零, 叫做零向量. 全体 n 元实向量组成的集合记作 \mathbf{R}^n .

n 元向量可以看成是空间几何向量(三元向量)的推广. 因此, 可把几何向量在坐标表示式下相等及加法和数乘运算推广到 n 元向量. 以后常用 α, β, γ 等表示 n 元向量, 用 $\mathbf{0}$ 表示零向量.

定义 1.26 设 $\alpha = (a_1, a_2, \dots, a_n)$, $\beta = (b_1, b_2, \dots, b_n)$, λ 是一个数, 我们定义:

$$\alpha = \beta \text{ 当且仅当 } a_i = b_i, \quad i = 1, 2, \dots, n;$$

$$\alpha + \beta = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n);$$

$$\lambda\alpha = (\lambda a_1, \lambda a_2, \dots, \lambda a_n).$$

其中 $\alpha + \beta$, $\lambda\alpha$ 分别称为 α 与 β 之和及数 λ 与 α 之乘积.

显然, n 元向量的加法与数量乘法也满足几何向量相应的运算规则. 向量的加法与数量乘法统称为向量的线性运算. 在下一章还将定义 n 元向量的内积运算.

下面介绍高斯消元法.

用高斯消元法求解 n 元一次方程的联立方程组(以后常称 n 元线性方程组), 是用规范化的加减消元法将方程组化为容易求解的同解方程组, 从而求得原方程组的解.

例 1 求解三元一次方程组

$$\begin{cases} x_1 + 2x_2 - 5x_3 = 2 & \text{①} \\ 2x_1 - 3x_2 + 4x_3 = 11 & \text{②} \\ -4x_1 - 7x_2 + 17x_3 = -7. & \text{③} \end{cases}$$

解 这里用高斯消元法, 其消元步骤如下:

先将方程 ① 分别乘 (-2) 和 4 加到方程 ②, ③ 上去, 消去方程 ②, ③ 中的 x_1 , 得

$$-7x_2 + 14x_3 = 7, \quad \text{④}$$

$$x_2 - 3x_3 = 1 \quad \text{⑤}$$

(这里容易证明 满足方程 ①, ② 的解必满足方程 ①, ④, 反之亦然; 同理 ①, ③ 与 ①, ⑤ 也同解; 从而 ①, ④, ⑤ 与 ①, ②, ③ 是同解方程组). 然后, 将方程 ④ 乘 $1/7$ 加到方程 ⑤ 上去, 消去方程 ⑤ 中的 x_2 , 得

$$-x_3 = 2 \quad (6)$$

同样可证, 方程组 ①, ④, ⑥ 与原方程组也是同解方程组. 于是, 由 ⑥ 得 $x_3 = -2$, 将其代入 ④ 得 $x_2 = -5$, 再将它们代入 ① 得 $x_1 = 2$. 所以原方程组的解为 $x_1 = 2, x_2 = -5, x_3 = -2$.

对于一般的 n 元线性方程组(其方程个数 m 不一定等于未知元个数 n)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (1-4)$$

(其中 $a_{ij}, b_i (i = 1 \dots m, j = 1 \dots n)$ 是已知数, $x_1 \dots x_n$ 是未知元), 用高斯消元法求解, 就是把方程组变为易于求解的同解方程组来求解, 还可以判定 (1-4) 是有唯一解, 无穷多解, 还是无解. 所谓消元就是把某些未知元的系数化为零. 因此, 为书写简便, 可省略未知元, 把方程组 (1-4) 中每个方程的 n 个系数和等号右边的常数看成一个 $n+1$ 元向量. 这样, 方程组 (1-4) 就对应于下列 m 个 $n+1$ 元向量(称为行向量)

$$\alpha_i = (a_{i1} \ a_{i2} \ \dots \ a_{in} \ b_i), \quad i = 1 \ 2 \ \dots \ m$$

所排成的一个矩形数表

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \quad \text{即} \quad \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix}. \quad (1-5)$$

这种由 $m \times (n+1)$ 个数排成 m 个(横)行、 $n+1$ 个(竖)列的矩形数表称为 $m \times (n+1)$ 矩阵. 上述矩阵 (1-5) 称为方程组 (1-4) 的增广矩阵, 它常记作 (A, b) , 其中

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix},$$

并称 A 为方程组 (1-4) 的系数矩阵. 再把 n 个未知元 x_1, x_2, \dots, x_n 记为 n 元

列向量

$$\mathbf{X} = (x_1 \ x_2 \ \cdots \ x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

如此, 我们可把方程组 (1-4) 简记为

$$\mathbf{A}\mathbf{X} = \mathbf{b}. \quad (1-4)^*$$

用高斯消元法求解方程组 (1-4), 其消元步骤可以简化到在增广矩阵 (1-5) 上进行. 一般来说, 其消元过程就是对矩阵 (1-5) 的行向量 $\alpha_1, \alpha_2, \dots, \alpha_m$ 作以下三种运算:

- (1) 以非零常数乘 (1-5) 中的某一行 (即 $n+1$ 元向量);
- (2) 将 (1-5) 中某行乘以非零常数并加到另一行上去;
- (3) 将 (1-5) 中某两行对换位置.

如果把所有的 $m \times (n+1)$ 矩阵组成的集合记作 $M_{m \times (n+1)}$, 即

$$M_{m \times (n+1)} = \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix} \mid \alpha_i \in \mathbf{R}^{n+1}, i = 1, 2, \dots, m \right\},$$

则上述三种运算都是定义在集 $M_{m \times (n+1)}$ 上取值于 $M_{m \times (n+1)}$ 的一元运算, 并称为矩阵的初等行运算 (通常称为矩阵的初等行变换).

在集 $M_{m \times (n+1)}$ 中定义二元关系 R : 设 $M_1, M_2 \in M_{m \times (n+1)}$, 规定 $M_1 R M_2$ 当且仅当 M_1 和 M_2 所表示的线性方程组有相同的解. 容易看出 R 是一个等价关系, 而且如果 M_2 是由 M_1 经过初等行运算而得到, 则 $M_1 R M_2$. 高斯消元法就是在集 $M_{m \times (n+1)}$ 中关于这个二元关系 R 的一个等价类中, 找到一个容易求解的线性方程组所对应的矩阵 M_0 的一种方法.

下面举例说明在线性方程组的增广矩阵上作初等行运算 (变换) 进行消元和求解的方法步骤.

例 2 求解线性方程组

$$\begin{cases} x_1 - x_2 - x_3 + 3x_5 = -1 \\ 2x_1 - 2x_2 - x_3 + 2x_4 + 4x_5 = -2 \\ 3x_1 - 3x_2 - x_3 + 4x_4 + 5x_5 = -3 \\ x_1 - x_2 + x_3 + x_4 + 8x_5 = 2. \end{cases} \quad (1-6)$$

解 方程组(1-6)的增广矩阵为

$$(A, b) = \left(\begin{array}{ccccc|c} 1 & -1 & -1 & 0 & 3 & -1 \\ 2 & -2 & -1 & 2 & 4 & -2 \\ 3 & -3 & -1 & 4 & 5 & -3 \\ 1 & -1 & 1 & 1 & 8 & 2 \end{array} \right). \quad (1-7)$$

以下把矩阵的初等行运算:第①行乘常数 c ,第②行乘 c 加到第③行上,第③行与第⑤行对换位置,依次记作 $(c) \times \textcircled{1}$, $\textcircled{3} + (c) \times \textcircled{2}$, $\textcircled{3} \leftrightarrow \textcircled{5}$.于是,

对矩阵(1-7)作初等行运算: $\textcircled{2} + (-2) \times \textcircled{1}$, $\textcircled{3} + (-3) \times \textcircled{1}$, $\textcircled{4} + (-1) \times \textcircled{1}$, 即得

$$(A, b) \rightarrow \left(\begin{array}{ccccc|c} 1 & -1 & -1 & 0 & 3 & -1 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 2 & 4 & -4 & 0 \\ 0 & 0 & 2 & 1 & 5 & 3 \end{array} \right). \quad (1-8)$$

对矩阵(1-8)作行运算: $\textcircled{3} + (-2) \times \textcircled{2}$, $\textcircled{4} + (-2) \times \textcircled{2}$, 即得

$$(1-8) \text{右} \rightarrow \left(\begin{array}{ccccc|c} 1 & -1 & -1 & 0 & 3 & -1 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 9 & 3 \end{array} \right). \quad (1-9)$$

对矩阵(1-9)作行运算 $(-\frac{1}{3}) \times \textcircled{4}$, $\textcircled{3} \leftrightarrow \textcircled{4}$, 即得

$$(1-9) \text{右} \rightarrow \left(\begin{array}{ccccc|c} 1 & -1 & -1 & 0 & 3 & -1 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right). \quad (1-10)$$

对矩阵(1-10)作行运算: $\textcircled{2} + (-2) \times \textcircled{3}$, 再将变换后的第②行加到第①行上, 即得

$$(1-10) \text{右} \rightarrow \left(\begin{array}{ccccc|c} 1 & -1 & 0 & 0 & 7 & 1 \\ 0 & 0 & 1 & 0 & 4 & 2 \\ 0 & 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right). \quad (1-11)$$

增广矩阵(1-11)所对应的线性方程组

$$\begin{cases} x_1 - x_2 & + 7x_5 = 1 \\ & x_3 & + 4x_5 = 2 \\ & & x_4 - 3x_5 = -1 \end{cases} \quad (1-12)$$

与原方程组同解. 现在是三个方程, 五个未知数, 任取 $x_2 = k_1$, $x_5 = k_2$ 代入方程组(1-12)可解出 x_1, x_3, x_4 , 从而得方程组的全部解: $x_1 = 1 + k_1 - 7k_2$, $x_2 = k_1$, $x_3 = 2 - 4k_2$, $x_4 = -1 + 3k_2$, $x_5 = k_2$ (其中 k_1, k_2 为任意常数). 以后把方程组的解也写成向量形式(称解向量)

$$\begin{aligned} X &= (x_1, x_2, x_3, x_4, x_5)^T \\ &= (1 + k_1 - 7k_2, k_1, 2 - 4k_2, -1 + 3k_2, k_2)^T \end{aligned}$$

(其中右上角的 T 表示 X 为列向量), 用向量的加法和数乘, 可把上述解向量表示为:

$$X = (1, 0, 2, -1, 0)^T + k_1(1, 1, 0, 0, 0)^T + k_2(-7, 0, -4, 3, 1)^T.$$

当方程组(1-4)中常数项 $b_1 = b_2 = \dots = b_m = 0$ 时, 称为齐次线性方程组, 否则叫非齐次线性方程组.

如果例2中四个方程的右端常数项全为零, 用同样方法得该齐次线性方程组的解向量为:

$$X = k_1(1, 1, 0, 0, 0)^T + k_2(-7, 0, -4, 3, 1)^T,$$

其中 k_1, k_2 为任意常数.

具有上面(1-10)(1-11)形式的矩阵分别称为阶梯形矩阵和行简化阶梯形矩阵(即阶梯形矩阵中每个非零行第一个非零元所在列的其余元素全为零). 化为后者求解时可避免繁琐的回代步骤. 求解方程组(1-12)时, x_2, x_5 可取任意常数, 它们称为自由未知量, 相应地 x_1, x_3, x_4 称为基本未知量.

例3 判断下列线性方程组是否有解?

$$\begin{cases} x_1 + x_2 + x_3 = 1 & \text{①} \\ x_1 + 2x_2 - 5x_3 = 2 & \text{②} \\ 2x_1 + 3x_2 - 4x_3 = 5. & \text{③} \end{cases}$$

对它的增广矩阵作初等行运算: ② + (-1) × ①, ③ + (-2) × ①, 然后再把变换后的矩阵作行运算: ③ + (-1) × ②, 即依次得

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & -5 & 2 \\ 2 & 3 & -4 & 5 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -6 & 1 \\ 0 & 1 & -6 & 3 \end{array} \right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -6 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right).$$

其中第三行所表示的方程 $0x_1 + 0x_2 + 0x_3 = 2$ 显然无解, 故原方程组无解. 这

是由于第三方程的左端等于前两个方程左端之和,而右端不等于前两方程右端之和.因此,满足前两个方程之解不满足第三方程,所以后者与前两个方程矛盾.含矛盾方程而无解的方程组称为不相容方程组;有解的方程组称为相容方程组.例2是相容的,它在消元时于矩阵中出现全零行,这是因为:方程③ = 方程① \times (-1) + 方程② \times (2).因此,满足前两个方程的解都满足第三方程,所以后者对求解是多余的,称之为多余方程.高斯消元法在消元过程中,会在增广矩阵上揭示出多余方程和矛盾方程.

对一般的线性方程组的增广矩阵(1-5)作三种初等行运算,都可把 (A, b) 化为如下形式的矩阵

$$(A, b) \rightarrow \left(\begin{array}{cccccccc|c} c_{11} & c_{12} & \cdots & c_{1j_2} & c_{1j_2+1} & \cdots & c_{1j_r} & c_{1j_r+1} & \cdots & c_{1n} & d_1 \\ 0 & 0 & \cdots & c_{2j_2} & c_{2j_2+1} & \cdots & c_{2j_r} & c_{2j_r+1} & \cdots & c_{2n} & d_2 \\ \cdots & & & & & & & & & & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & c_{rj_r} & c_{rj_r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1} \\ \cdots & & & & & & & & & & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array} \right). \quad (1-13)$$

其中,0行全在矩阵的下方,而第*i*个非零行的第一个不为零的元素 c_{rj_i} 满足

$$1 < j_2 < \cdots < j_r \leq n.$$

这种矩阵叫阶梯矩阵,每行第一个不为零的元素叫主元素.

我们还可通过行变换继续把主元素都变成1,而每一主元素所在列的其余元素都是零,这种矩阵叫行简化阶梯矩阵.

由增广矩阵(1-13)(假设它已是行简化阶梯矩阵)易见,方程组有解的充要条件是 $d_{r+1} = 0$,因为 $d_{r+1} \neq 0$ 时,其中第 $r+1$ 行对应的方程显然无解.

在有解的情况下:

(1)当 $r = n$ 且 $c_{11} = c_{22} = \cdots = c_{nn} = 1$ 时,有唯一解 $x_1 = d_1, x_2 = d_2, \cdots, x_n = d_n$;

(2)当 $r < n$ 时,有无穷多个解,求解时把(1-13)中每行的主元素所在列对应的未知量取作基本未知量,其余的取作自由未知量,并把自由未知量依次取任意常数 $k_1, k_2, \cdots, k_{n-r}$,将它们代入(1-13)所对应的方程组即可解得基本未知量,从而得到方程组的全部解.

齐次线性方程组总是有解的,因为 $b_1 = \cdots = b_m = 0$,从而 $d_1 = \cdots = d_r = d_{r+1} = 0$.当 $r = n$ 时,只有零解,即 $x_1 = \cdots = x_n = 0$;当 $r < n$ 时,有无穷多解.

无穷多解 求解的方法同上. 如果齐次线性方程组中方程个数 m 小于未知量个数 n , 则必有无穷多个非零解.

线性代数在它的发展历史上, 最早研究的问题就是一般线性方程组的求解问题. 线性方程组的解的基本问题是: 有解的条件(对于齐次方程组则是有非零解的条件)以及解的结构.

(1-4)式所示的线性方程组 $AX = b$, 实际上是一个映射, 它是 $m \times n$ 系数矩阵 A 把 n 元列向量 $X = (x_1 \ x_2 \ \dots \ x_n)^T$ 映射(或说变换)为 m 元列向量 $b = (b_1 \ b_2 \ \dots \ b_m)^T$, 即

$$A: X \mapsto b.$$

求一个线性方程组 $AX = b$ 的解, 也就是给定了映射 A (即系数矩阵)和这个映射的象 b , 求 b 在这个映射下的完全原象或逆象(即线性方程组的解集合 S)

$$S = A^{-1}(b) = \{X | X \in \mathbb{R}^n \text{ 且 } AX = b\}.$$

求 b 的原象是映射(或变换)的反问题, 是反变换, 用高斯消元法我们可以实现这个反变换. 对于给定的 A 和 b , b 的原象可能不存在, 即 $A^{-1}(b) = \emptyset$; 也可能存在且唯一(即线性方程组有唯一解); 也可能存在, 但不唯一, 即线性方程组有无穷多解. 但是高斯消元法只能在线性方程组的增广矩阵经初等行运算化为阶梯形矩阵后, 才能知道方程组的解的情况. 因此我们还需要把消元法求解的具体操作上升到理性的高度, 揭示线性方程组的解的情况与增广矩阵结构(即增广矩阵的 m 个行向量及 n 个列向量在线性运算下的相互关系)间的内在联系. 事实上, 如果把线性方程组(1-4)的系数矩阵 A 的 n 个列向量记作

$$\beta_j = (a_{1j} \ a_{2j} \ \dots \ a_{mj})^T = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \quad j = 1 \ 2 \ \dots \ n,$$

则线性方程组(1-4)等价于一个向量方程

$$x_1 \beta_1 + x_2 \beta_2 + \dots + x_n \beta_n = b. \quad (1-4)^{**}$$

这里使 $(1-4)^{**}$ 成立的 $x_1 \ x_2 \ \dots \ x_n$ 就是方程组(1-4)的解. 所以方程组解的情况取决于向量组 $\beta_1 \ \beta_2 \ \dots \ \beta_n$ 与 b 在向量线性运算下的相互关系. 因此要搞清一般线性方程组的解的理论, 就要研究全体 n 元(或 m 元)向量(记作 \mathbb{R}^n 或 \mathbb{R}^m)在线性运算下的相互关系, 也就是要搞清楚下一章要讨论的有限维

向量空间的结构,这正是线性代数的核心内容之一.此外,系数矩阵为 $m \times n$ 矩阵 A 的线性方程组作为 $\mathbf{R}^n \rightarrow \mathbf{R}^m$ 的映射,它是一种线性映射(以后将论述),所以在第三、四章讨论了一般线性映射的理论及其矩阵表示以后,我们就明白线性方程组的解的问题,只是线性映射的一个具体模型.因此线性映射的理论是线性代数的又一个核心内容.

1.9 平面方程与空间直线方程

在平面直角坐标系中的一条直线对应于一个二元一次方程 $ax + by = c$. 这一节讨论在空间直角坐标系中一个平面和一条直线对应于怎样的三元方程,进而对三元一次方程组解的情况作几何解释,并为第二章中子空间的概念提供几何背景.

1.9.1 平面方程

在空间直角坐标系中,欲确定一个平面 π ,只要给定平面上一个定点 $P_0(x_0, y_0, z_0)$ 以及与平面垂直的一个向量

$$\mathbf{n} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} = (a, b, c),$$

并称 \mathbf{n} 为平面的法向量(如图 1-18). 所谓平面 π 的方程,就是平面 π 上任一点 $P(x, y, z)$ 的坐标 x, y, z 所满足的一个关系式.

点 $P(x, y, z)$ 在平面 π 上当且仅当 $\overrightarrow{P_0P} \perp \mathbf{n}$, 即 $\overrightarrow{P_0P} \cdot \mathbf{n} = 0$, 其中

$$\overrightarrow{P_0P} = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k},$$

于是得

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \quad (1-14)$$

方程(1-14)称为平面 π 的点法式方程,它可化为

$$ax + by + cz + d = 0. \quad (1-15)$$

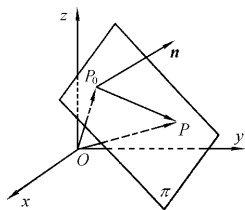


图 1-18

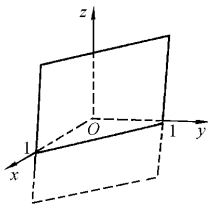


图 1-19

其中 $d = -ax_0 - by_0 - cz_0$. 方程 (1-15) 称为平面 π 的一般方程, 它是一个三元一次方程. 事实上, 任何三元一次方程都可化为 (1-14) 的形式, 因而都是平面的方程. 在平面一般方程 (1-15) 中, 系数 a, b, c 是平面法向量 \mathbf{n} 的三个坐标, 即 $\mathbf{n} = (a, b, c)$.

(1) 当 a, b, c 有一个为 0 时, 平面 π 与某个坐标轴平行, 例如 $x + y = 1$, 其中 $c = 0$, $\mathbf{n} \perp z$ 轴, 平面平行于 z 轴 (如图 1-19).

(2) 当 a, b, c 有两个为 0 时, 平面 π 平行于某坐标平面, 例如 $x = 1$ 平行于坐标平面 yOz (如图 1-20).

(3) 当 $d = 0$ 时, 平面 π 过坐标原点 $O(0, 0, 0)$, 因为原点坐标满足方程

$$ax + by + cz = 0.$$

(4) 当 $abcd \neq 0$ 时, 方程 (1-15) 可化为

$$\frac{x}{a_1} + \frac{y}{b_1} + \frac{z}{c_1} = 1. \quad (1-16)$$

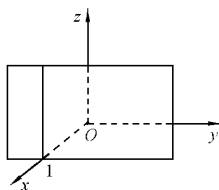


图 1-20

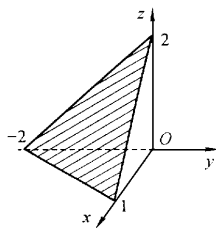


图 1-21

(1-16) 称为平面的截矩式方程, 其中 a_1, b_1, c_1 分别称为平面在 x, y, z 轴上的截矩. 例如 $\frac{x}{1} + \frac{y}{-2} + \frac{z}{2} = 1$ 的图形如图 1-21 所示.

例 1 已知平面过三点 $A(1, 2, -1), B(2, 1, -3), C(5, 2, -4)$. 试求平面方程.

解 该平面的法向量 \mathbf{n} 与 $\overrightarrow{AB}, \overrightarrow{AC}$ 都垂直, 所以

$$\mathbf{n} = \overrightarrow{AB} \times \overrightarrow{AC} = (1, -1, -2) \times (4, 0, -3) = (3, -5, 4).$$

又已知平面过点 $A(1, 2, -1)$, 所以平面方程为

$$3(x-1) - 5(y-2) + 4(z+1) = 0,$$

即

$$3x - 5y + 4z + 11 = 0.$$

例2 已知平面 π 过点 $P_0(1, -2, 1)$ 且平行于 $\beta = i + j - k$ 和 z 轴, 试求平面 π 的一般方程.

解法1 该平面的法向量

$n = \beta \times k = (i + j - k) \times k = i \times k + j \times k = -j + i = i - j$,
所以平面 π 的点法式方程为

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

由此得平面的一般方程 $x - y - 3 = 0$.

解法2 (待定系数法) 由于平面平行于 z 轴, 即法向量 n 垂直于 z 轴, $n = (a, b, 0)$, 所以其一般方程为

$$ax + by + d = 0. \quad (1)$$

由点 P_0 的坐标满足方程得

$$a - 2b + d = 0, \quad (2)$$

再由法向量 $n = ai + bj$ 垂直于 β , 又得

$$n \cdot \beta = a + b = 0. \quad (3)$$

从 (2), (3) 解得 $b = -a$, $d = -3a$, 代入 (1) 式得

$$ax - ay - 3a = 0, \quad (4)$$

其中 $a \neq 0$, 消去 a , 得平面的一般方程 $x - y - 3 = 0$.

1.9.2 空间直线方程

在空间直角坐标系中确定一直线 L , 只要给定 L 上一个定点 $P_0(x_0, y_0, z_0)$ 以及与直线平行的一个向量 $S = li + mj + nk = (l, m, n)$, S 称为直线的方向向量 (如图 1-22).

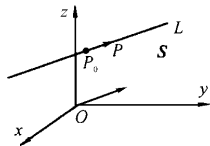


图 1-22

点 $P(x, y, z)$ 在直线上当且仅当 $\overrightarrow{P_0P} \parallel S$, 即当且仅当 $\overrightarrow{P_0P} = tS$, 于是

$$\begin{cases} x - x_0 = lt \\ y - y_0 = mt \\ z - z_0 = nt \end{cases} \quad \text{即} \quad \begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases} \quad (1-17)$$

方程(1-17)称为直线的参数方程,其中 t 称为参数.在方程中消去参数 t ,得到

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}. \quad (1-18)$$

方程(1-18)称为直线的标准方程,式中的两个连等号实际上是给了由两个独立方程所组成的方程组,即

$$\begin{cases} \frac{x-x_0}{l} = \frac{y-y_0}{m} \\ \frac{x-x_0}{l} = \frac{z-z_0}{n} \end{cases} \quad \text{或} \quad \begin{cases} \frac{x-x_0}{l} = \frac{y-y_0}{m} \\ \frac{y-y_0}{m} = \frac{z-z_0}{n} \end{cases} \quad \text{或} \quad \begin{cases} \frac{x-x_0}{l} = \frac{z-z_0}{n} \\ \frac{y-y_0}{m} = \frac{z-z_0}{n} \end{cases}.$$

直线的参数方程给出了直线上的点 P 与参数 t 的一一对应关系.

例如,质点 m 从点 $P_0(x_0, y_0, z_0)$ 起,以等速度 $v = (v_x, v_y, v_z)$ 运动,则其轨迹(为直线)方程(或称运动方程)为

$$\begin{cases} x = x_0 + v_x t \\ y = y_0 + v_y t \\ z = z_0 + v_z t \end{cases} \quad t \geq 0.$$

直线的标准方程显示了直线过定点 $P_0(x_0, y_0, z_0)$ 和其方向向量 $S = (l, m, n)$.

如果 l, m, n 中有一个为0,不妨设 $l = 0$,此时 $S \perp x$ 轴,即直线 $L \perp x$ 轴,于是方程(1-18)应理解为

$$\begin{cases} x - x_0 = 0 \\ \frac{y - y_0}{m} = \frac{z - z_0}{n} \end{cases}.$$

如果 $l = 0$ 且 $m = 0$,此时 $S \perp x$ 轴且 $S \perp y$ 轴,所以 $S \parallel z$ 轴,即直线 $L \parallel z$ 轴,相应的方程应理解为

$$\begin{cases} x - x_0 = 0 \\ y - y_0 = 0 \end{cases}.$$

此时直线 L 上点 P 的 z 坐标可任取,这是过点 $P_0(x_0, y_0, 0)$ 且平行于 z 轴的一条直线.

一般也可把直线视为两个平面的交线.因此,如果两个平面 π_1 与 π_2 相交,则它们的两个三元一次方程所组成的方程组

$$\begin{cases} a_1x + b_1y + c_1z + d_1 = 0 & (\pi_1 \text{ 的方程}) \\ a_2x + b_2y + c_2z + d_2 = 0 & (\pi_2 \text{ 的方程}) \end{cases} \quad (1-19)$$

就是 π_1 与 π_2 的交线的方程式, 并称为直线的一般方程. 当然, 任意两个三元一次方程所组成的方程组并不一定是直线的方程.

例3 求过点 $P_0(1, 0, -4)$ 且与直线

$$\begin{cases} x - y + 3z + 1 = 0 \\ 2x - 2y + z - 3 = 0 \end{cases}$$

平行的直线方程.

解 由于已知直线(两平面的交线)与两个平面的法向量 n_1, n_2 都垂直, 所以 $n_1 \times n_2$ 是其方向向量 S , 即

$$S = n_1 \times n_2 = (1, -1, 3) \times (2, -2, 1) = (5, 5, 0) // (1, 1, 0),$$

故所求直线的标准方程为

$$\frac{x-1}{1} = \frac{y}{1} = \frac{z+4}{0}.$$

其一般方程为

$$\begin{cases} x - y - 1 = 0 \\ z + 4 = 0. \end{cases}$$

其参数方程为

$$\begin{cases} x = 1 + t \\ y = 0 + t \\ z = -4. \end{cases}$$

例4 将例3中已知直线的一般方程化为标准方程.

解 上例已得直线的方向向量 $S = (1, 1, 0)$, 因此只要再求直线上一个点的坐标. 此时令 $x = 0$, 将其代入直线一般方程, 可解得 $y = -2, z = -1$, 故 $P_0(0, -2, -1)$ 是直线上一个点. 因此, 直线的标准方程为

$$\frac{x}{1} = \frac{y+2}{1} = \frac{z+1}{0}.$$

1.9.3 三元一次方程组的解的几何解释

我们把三元一次方程组

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \\ a_3x + b_3y + c_3z = d_3 \end{cases} \quad (1-20)$$

中的三个方程对应的平面分别叫做 π_1, π_2, π_3 .

(1) 当三个平面交于一点(即三个法向量不共面)时, 方程组有唯一解.

(2) 当三个平面共线时, 方程组有无穷多解, 其解向量可表示为

$$X = (x_0, y_0, z_0)^T + k(l, m, n)^T,$$

其中 (x_0, y_0, z_0) 和 (l, m, n) 为两个常向量, k 为任意常数, 这实际上是直线参数方程的向量形式.

(3) 当三个平面重合(即共面)时, 方程组有无穷多解, 它就是其中任一个方程的解. 分别取 y, z 为任意常数 k_1 和 k_2 , 代入方程解得 x , 此时解向量可表示为

$$X = (x, y, z)^T = (x_0, y_0, z_0)^T + k_1(x_1, y_1, z_1)^T + k_2(x_2, y_2, z_2)^T$$

其中 $(x_0, y_0, z_0)^T, (x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T$ 为三个常向量.

(4) 方程组无解时, 三个平面间的相互关系可能有以下两种类型: 一是有两个平面平行而不重合; 二是有两个平面(如 π_1, π_2)相交, 交线为 L , 第三个平面 π_3 与 L 平行, 但 L 不在 π_3 上, 此时满足方程 ①、② 的解都不满足方程 ③, 故无解.

对于三元齐次线性方程组, 即方程组(1-20)中 $d_1 = d_2 = d_3 = 0$, 由于每个平面都过原点, 所以方程组必有零解 $(0, 0, 0)$; 至于何时只有零解, 何时有无穷多解, 读者不难给出几何解释.

1.10 基本代数结构——群、环、域的基本概念

代数结构是基本的数学结构, 群、环、域则是基本的代数结构. 这里列举一些例子, 让读者了解如何从数学结构上区分各种数学问题的异同, 了解群、环、域(主要是群)三种结构最基本的概念.

如果集合中不定义运算, 集合不过是一堆元素而已. 如果集 X 中定义了代数运算(即 X 上封闭的二元运算), 则集 X 中的元素由运算而确定了相互的关系, 也就是有了一定的“结构”. 通常把一个非空集合 X 和在 X 上定义的若干个代数运算 f_1, \dots, f_k 组成的系统称为代数系统(简称代数系), 记作 $X; f_1, \dots, f_k$. 一个代数系统的性质, 或说代数系统的数学结构是由其中运算的性质所决定的. 在不同的集合上定义不同的代数运算, 虽然其运算方法会不同, 但运算的性质却有可能是相同的. 例如, 在整数集 \mathbf{Z} 和实数集 $\mathbf{R} \setminus \{0\}$ 中, 分别定义加法和乘法, 则代数系统 $\mathbf{Z}; +$ 与 $\mathbf{R} \setminus \{0\}; \cdot$ 有以下相同的性质:

(1) 加法和乘法都满足结合律和交换律;

(2) $\exists 0 \in \mathbf{Z}$, 使 $\forall k \in \mathbf{Z}$ 均有 $0 + k = k$; $\exists 1 \in \mathbf{R} \setminus \{0\}$, 使 $\forall a \in \mathbf{R}$ 均有 $1 \cdot a = a$. 这里 0 和 1 分别称为加法和乘法的单位元;

(3) $\forall k \in \mathbf{Z}, \exists -k \in \mathbf{Z}$, 使得 $k + (-k) = 0$ (加法单位元); $\forall a \in \mathbf{R} \setminus \{0\}, \exists a^{-1} = \frac{1}{a} \in \mathbf{R} \setminus \{0\}$, 使得 $a \cdot a^{-1} = 1$ (乘法单位元). 以后, 我们把

$-k$ 叫做 k 关于加法的逆元, a^{-1} 叫做 a 关于乘法的逆元.

这些性质的共同点概括起来为 (1) 运算满足结合律和交换律 (2) 集合关于运算存在单位元 (3) 集合中每个元素关于运算都有逆元. 以后, 我们把具有上述性质的代数系统称为“交换群”.

下面介绍三种基本代数结构——群、环、域的基本概念.

1.10.1 半群和群

群是历史上最早提出, 也是最简单的一种代数结构. 在 19 世纪 30 年代, 年青的法国数学家伽罗瓦 (Galois) 开创性地用群论方法建立了方程的可解性理论, 证明了 5 次和 5 次以上的代数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. 群论的发现使代数的研究进入了新时代, 从局部性研究转向整体结构的分析研究. 群结构的观点已经渗透到数学所有的分支. 这里我们只简要地叙述群的基本概念.

群是具有一个代数运算的代数系统, 它的定义如下:

定义 1.27 代数系统 $G : \circ$ 称为群, 如果:

(1) 运算“ \circ ”满足结合律, 即 $\forall a, b, c \in G, a \circ (b \circ c) = (a \circ b) \circ c$;

(2) G 关于运算“ \circ ”存在单位元, 即 $\exists e \in G$, 使 $\forall a \in G$, 有 $a \circ e = e \circ a = a$;

(3) G 中每个元素关于“ \circ ”都可逆, 即 $\forall a \in G, \exists b \in G$ 使得 $a \circ b = b \circ a = e$ (单位元), 并称 a 为可逆元, b 为 a 的逆元, 记作 $b = a^{-1}$.

$G : \circ$ 是一个群, 也说 G 关于“ \circ ”构成群. 如果运算还满足交换律, 即 $\forall a, b \in G$, 有 $a \circ b = b \circ a$, 则称 $G : \circ$ 为交换群, 也称 Abel 群.

当 $G : \circ$ 适合条件 (1) 时, 称之为半群; 当 $G : \circ$ 适合条件 (1) 和 (2) 时, 称之为含么半群.

如果群 G 的子集 H 关于 G 的运算也构成群, 则称 H 为 G 的子群, 记作 $H \leq G$.

例 1 设 $+$, \cdot 分别为数的加法和乘法, 则 $\mathbf{N}^* : +$, $\mathbf{N}^* : \cdot$, $\mathbf{Z} : \cdot$ 都是交换半群; $\mathbf{Q}^+ : \cdot$, $\mathbf{R}^* : \cdot$ 都是交换群 (其中 \mathbf{Q}^+ 是正有理数集, $\mathbf{R}^* = \mathbf{R} \setminus \{0\}$), 且 \mathbf{Q}^+ 是 \mathbf{R}^* 关于乘法运算的子群, 而 $\{1, -1\} : \cdot$ 是仅有两个元素的交换群. 一般, 含有限个元素的群称为有限群, 否则叫无限群.

例 2 设 \mathbf{R}^3 是全体空间几何向量组成的集合, 则 \mathbf{R}^3 关于向量的加法构成一个交换群. 其中单位元 e 为零向量 $\mathbf{0}$, 任一向量 α 的逆元为 $-\alpha$. 同理 \mathbf{R}^n 关于向量的加法也构成一个交换群.

例 3 设 $\mathcal{R}(X)_3 = \{a_0 + a_1x + a_2x^2 \mid a_0, a_1, a_2 \in \mathbf{R}\}$; $\mathcal{R}(X)$ 是由所有实系数多项

式组成的集合,则它们关于多项式的加法都构成交换群,单位元 e 都是零多项式(即系数全为零的多项式);对于多项式乘法, $R(X)$ 是含么半群(其单位元 $e=1$);而 $R(X)$ 不是半群,因为它关于乘法不封闭.

例4 设 $G=P(X)$ (集 X 的幂集),则 $G:\cap, G:\cup$ 都是含么半群,其中运算 \cup 的单位元是 \emptyset ,运算 \cap 的单位元是 X .

例5 开关电路(见节1.5例3)的状态集 $G=\{0,1\}$ 关于运算 \vee, \wedge 都构成含么半群.因为集 G 关于这两种运算都封闭,容易验证两种运算都满足结合律;关于运算 \vee, \wedge 的单位元分别为0和1,非单位元不可逆.

例6 设 G 为某班学生集,运算“ \circ ”为比高矮,规定

$$a \circ b = b \circ a = a \Leftrightarrow a \text{ 比 } b \text{ 高}; a \circ a = a, \forall a \in G.$$

由于任两个人的个子不可能绝对相等,都可比高矮,所以该运算是 G 的代数运算,且满足结合律,运算的单位元为最矮的学生,因此 G 关于该运算构成含么半群.

例7 设 $G=\{1, 2, \dots, 12\}$ 在 G 上定义时钟加法 \oplus 为

$$a \oplus b = \begin{cases} a+b, & a+b \leq 12, \\ a+b-12, & a+b \geq 13, \end{cases}$$

则 $G:\oplus$ 是交换群,因为 G 关于运算 \oplus 封闭,且运算满足结合律和交换律, G 关于 \oplus 的单位元 $e=12$,小于12的任何元素 a 的逆元为 $12-a$, $e^{-1}=e$.

例8* 设 Z_n 是 \mathbf{Z} 关于模 n 同余关系 R 的商集,即

$$Z_n = \mathbf{Z}/R = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}.$$

在 Z_n 上定义加法 \oplus 为

$$\overline{a} \oplus \overline{b} = \overline{a+b}.$$

由于等价类 \overline{a} 的代表元不是唯一的,即 $\overline{a} = \overline{kn+a}$ ($k \in \mathbf{Z}$),因此首先要证明 $\overline{a} \oplus \overline{b}$ 不管取怎样的代表元,其运算的结果是唯一的.

设 $\overline{a} = \overline{a_1}, \overline{b} = \overline{b_1}$,则 $a_1 = pn+a, b_1 = qn+b$,其中 $p, q \in \mathbf{Z}$,于是

$$a_1 + b_1 = (p+q)n + (a+b),$$

因此 $(a_1 + b_1) \equiv (a+b) \pmod{n}$,即 $(a_1 + b_1)R(a+b)$,从而

$$\overline{a_1 + b_1} = \overline{a+b},$$

这就证明了加法 \oplus 是 Z_n 上的二元运算. Z_n 对加法 \oplus 显然封闭,而且满足结合律和交换律, Z_n 关于 \oplus 的单位元 $e = \overline{0}$, Z_n 中任一元素 \overline{a} 的逆元为 $\overline{n-a}$,因为

$$\overline{a} \oplus \overline{n-a} = \overline{n-a} \oplus \overline{a} = \overline{a+n-a} = \overline{n} = \overline{0} = e,$$

所以 $Z_n:\oplus$ 是个交换群,叫做模 n 剩余类加群.

容易验证: Z_6 的子集 $\{\overline{0}, \overline{3}\}, \{\overline{0}, \overline{2}, \overline{4}\}$ 关于 Z_6 的加法 \oplus 也都构成群,所以它们都是 Z_6 的子群.

例9 从集 A 到自身的所有映射(变换) A^A 关于映射的乘法(或合成)运算构成一个含么半群 A^A, \circ ,其单位元为恒等映射.集 A 到自身的所有双射(一一变换) $E(A)$ 关于映

射的乘法构成一个群, $E(A)$ 称为 A 的一个变换群. $E(A)$ 的子群称为变换群. 例如, $E_f(I) = \{f(x) = x^\alpha \mid 0 < \alpha \in \mathbf{R}, x \in I = (0, 1)\}$ 是区间 I 上的一个变换群, 这个变换群是可交换的, 但一般变换群是不可交换的.

例如, 集 $X = \{1, 2, 3\}$ 到自身的一个双射 (一一变换) 是

$$\sigma = \begin{pmatrix} i_1 & i_2 & i_3 \\ j_1 & j_2 & j_3 \end{pmatrix},$$

即 $\sigma(i_k) = j_k$, 其中 $i_k, j_k \in X$ ($k = 1, 2, 3$), 且 $k \neq m$ 时, $i_k \neq i_m, j_k \neq j_m$.

此时也把 σ 称为 X 的一个置换, 也称三元置换. 集 X 上的三元置换共有 $3!$ 个, 即

$$\begin{aligned} \sigma_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \\ \sigma_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}. \end{aligned}$$

我们把全体三元置换组成的集合记作 S_3 , 容易验证 S_3 关于变换 (即映射) 的乘法构成一个群, 称为置换群, 它的单位元为恒等置换 σ_1 ; $\sigma_2^{-1} = \sigma_2, \sigma_3^{-1} = \sigma_3, \sigma_4^{-1} = \sigma_4, \sigma_5^{-1} = \sigma_6, \sigma_6^{-1} = \sigma_5$. 但这个置换群是不可交换的, 例如,

$$\sigma_3 \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

因为

$$\sigma_3 \sigma_4(1) = \sigma_3(1) = 3, \quad \sigma_3 \sigma_4(2) = \sigma_3(3) = 1, \quad \sigma_3 \sigma_4(3) = \sigma_3(2) = 2,$$

而

$$\sigma_4 \sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \neq \sigma_3 \sigma_4.$$

关于群的定义 1.27 还需指出以下两点:

(1) 群的单位元是唯一的. 如果 e_1, e_2 都是群 G 的单位元, 则

$$e_1 = e_1 \circ e_2 = e_2;$$

(2) 群的每个元的逆元也是唯一的. 如果 b, c 都是 a 的逆元, 则由

$$a \circ b = b \circ a = e \text{ 和 } a \circ c = c \circ a = e,$$

即得

$$b = b \circ e = b \circ (a \circ c) = (b \circ a) \circ c = e \circ c = c.$$

1.10.2 环与域

环与域都是有二个代数运算“ \oplus, \circ ” (通常称为“加法, 乘法”) 的代数系, $a \oplus b$ 和 $a \circ b$ 简记为 $a + b$ 和 ab .

定义 1.28 代数系 $R : +, \circ$ 称为环, 如果

(1) $R : +$ 是交换群(加法群), 其单位元记作 0 (也称为环 R 的乘法零元);

(2) $R : \circ$ 是半群;

(3) 运算“ \circ ”对“ $+$ ”满足左、右分配律, 即 $\forall a, b, c \in R$,

$$a(b + c) = ab + ac;$$

$$(b + c)a = ba + ca.$$

定义中的(3)是重要的, 没有它, R 只是对“ $+$ ”和“ \circ ”分别构成交换群和半群, 而不能成为区别于群结构的另一种代数结构.

如果环 $R : +, \circ$ 中的乘法满足交换律, 则称其为交换环; 如环关于乘法存在单位元(乘法单位元 e 也常记作 1) 则称之为含么环.

例 1 对于数的加法和乘法, 代数系 $\mathbf{Z} : +, \circ$, $\mathbf{Q} : +, \circ$, $\mathbf{R} : +, \circ$, $\mathbf{C} : +, \circ$ 都是含么交换环.

$\mathbf{Z}_e : +, \circ$ (\mathbf{Z}_e 为偶数集) 是交换环, 但不是含么环.

$\mathbf{N}^* : +, \circ$ 不是环, 因为 $\mathbf{N}^* : +$ 不是群.

例 2 整系数多项式集合

$$\mathcal{A}[x] = \{p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \mid a_i \in \mathbf{Z}, n \in \mathbf{N}^*\}$$

对多项式加法和乘法构成一个含么交换环, 乘法单位元为 $p(x) = 1$, 但是对于正整数 n ,

$$\mathcal{A}[x]_{n+1} = \{p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \mid a_i \in \mathbf{Z}\}$$

对同样的加法和乘法不构成环, 因为 $\mathcal{A}[x]_{n+1}$ 对乘法不封闭, 不是半群.

例 3* 在模 n 剩余类加法群 $Z_n : \oplus$ 中定义乘法运算“ \circ ”为

$$\overline{a} \circ \overline{b} = \overline{ab}, \quad (\forall \overline{a}, \overline{b} \in Z_n),$$

则 $Z_n : \oplus, \circ$ 是一个含么交换环. 这是因为

(1) 这里定义的乘法“ \circ ”是 Z_n 上的一个代数运算, 因为 $\overline{a} \circ \overline{b}$ 的运算结果是唯一的, 它与等价类 $\overline{a}, \overline{b}$ 的代表元的选择无关, 其证明如下:

设 $\overline{a}_1 = \overline{a}, \overline{b}_1 = \overline{b}$, 即 $a_1 = pn + a, b_1 = qn + b$ 其中 $p, q \in \mathbf{Z}$, 则

$$a_1 b_1 = (pqn + pb + aq)n + ab,$$

因此 $a_1 b_1 \equiv ab \pmod{n}$, 从而 $\overline{a_1 b_1} = \overline{ab}$.

(2) 乘法运算“ \circ ”显然满足结合律与交换律; 单位元为 $\overline{1}$.

(3) 乘法对加法满足分配律, 因为

$$\begin{aligned} \overline{a} \circ (\overline{b} \oplus \overline{c}) &= \overline{a} \circ \overline{b + c} = \overline{a(b + c)} = \overline{ab + ac} \\ &= \overline{ab} \oplus \overline{ac} = \overline{a} \circ \overline{b} \oplus \overline{a} \circ \overline{c}. \end{aligned}$$

定义 1.29 代数系 $F : +, \cdot$ 称为一个域, 如果它是至少含有两个元的交换环, 且 $F \setminus \{0\}$ 关于乘法运算是交换群.

由定义可见, F 至少含加法单位元 (即环的零元 0) 和乘法单位元 e .

任一个数集对于数的加法和乘法要构成一个域都必须含 0 和 1 . 有理数集 \mathbf{Q} 、实数集 \mathbf{R} 和复数集 \mathbf{C} 对数的加法和乘法都构成域, 分别称为有理数域、实数域、复数域.

根据域的公理化定义, 数集 F 对数的加法和乘法构成数域的条件也可表述为: 数集 F 含 $0, 1$, 且对数的加、减、乘、除 (除数不为 0) 运算封闭. 这是因为: 对减法封闭 (即 $\forall a, b \in F, a - b \in F$) 保证了 F 中任何非零数 a 对加法可逆 (即 $(-a) = 0 - a \in F$); 对除法封闭 (即 $\forall a, b \in F$ 且 $a \neq 0$ 均有 $\frac{b}{a} \in F$) 保证了 F 中任何非零数 a 对乘法可逆 (即 $a^{-1} = \frac{1}{a} \in F$).

例 4 设 $Q(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbf{Q}\}$, 证明 $Q(\sqrt{2})$ 是一个数域.

证 在集 $Q(\sqrt{2})$ 中含 $0, 1$, 且数的加、减、乘运算显然封闭 (证明留给读者). 如果 $c + d\sqrt{2} \neq 0 (c, d \in \mathbf{Q})$ 则 $c - d\sqrt{2} \neq 0$, 且

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}} = \frac{(a + b\sqrt{2})(c - d\sqrt{2})}{c^2 - 2d^2} = \frac{ac - 2bd}{c^2 - 2d^2} + \frac{bc - ad}{c^2 - 2d^2}\sqrt{2} \in Q(\sqrt{2}),$$

故在集 $Q(\sqrt{2})$ 上除法运算也封闭, 所以 $Q(\sqrt{2})$ 是一个数域.

任何数域 F 都包含有理数域 \mathbf{Q} , 即 \mathbf{Q} 是最小的数域. 事实上, 由 $0, 1 \in F$, 得 $n = 1 + 1 + \dots + 1 \in F, 0 - n = -n \in F$, 从而 $\mathbf{Z} \subset F$; 又 $\forall p, q \in \mathbf{Z} \subset F, p \neq 0$ 均有 $\frac{q}{p} \in F$, 而 $q/p \in \mathbf{Q}$, 所以 $\mathbf{Q} \subset F$. 这就证明了 \mathbf{Q} 是最小的数域.

任何一个数域显然都含有无穷多个元素. 但是一般的域并不都像数域那样含有无穷多个元素. 例如, $\mathbf{Z}_2, \oplus, \cdot$ 是仅含两个元素 $\bar{0}, \bar{1}$ 的有限域. 因为由例 3 已知它是一个交换环, 又非零元 $\bar{1}$ 是乘法单位元, 它是可逆的, 其逆元为自身.

习 题

1. 设 $A = \{1, 2, 7, 8\}, B = \{n \mid n^2 < 50, n \in \mathbf{N}^*\}, C = \{n \mid 0 < n \leq 30 \text{ 且 } \frac{n}{3} \in \mathbf{N}^*\}$, 求:

$$B \setminus (A \cup C), \quad (B \cap C) \setminus A.$$

2. 已知 $A, B, C \subseteq \mathbb{Z}^*$ 都是整数 \mathbb{Z} 集的子集, $A = \{n \mid n = 3m, m \geq 4\}$; $B = \{n \mid n = 2m, |m| \geq 1\}$; $C = \{n \mid |n| \leq 10\}$, $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$. 试用 $A, B, C \subseteq \mathbb{Z}^*$ 的运算, 表达下列集合:

(1) $\{n \mid n = 6m, m \geq 2\}$;

(2) 奇数组成的集合;

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

3. 设 \mathbb{N}^* 是正整数集, $M_n = \left\{m \mid \frac{m}{n} \in \mathbb{N}^*, m \text{ 是某个正整数}\right\}$. 求:

(1) $\bigcup_{n \in \mathbb{N}} M_n$; (2) $M_n \cap M_k$; (3) $\bigcap_{n \in \mathbb{N}} M_n$.

4. 设 $A = \{x \mid x^2 - 1 = 0\}$, 求 A 的幂集 $P(A)$; $A \times A$.

5. 京、津、沪三支足球队的联赛规定, 每队都在主、客场与另一队打两场比赛, 以(京、津)表示北京队在主场与天津队比赛. 试问: 用怎样的集合及其运算来表示全部六场比赛所组成的集合?

6. 设 $A = \{(x, y) \mid x^2 + y^2 \leq 4\}$, $B = \{x \mid -\infty < x < +\infty\}$, $C = \{y \mid -1 \leq y \leq 1\}$. 在平面直角坐标系内画出下列集合:

(1) $A \cup (B \times C)$; (2) $A \cap (B \times C)$;

(3) $A \setminus (B \times C)$; (4) $(B \times C) \setminus A$.

7. 设 A 是含有 n 个元素的有限集, 证明 A 的幂集 $P(A)$ 所含元素为 2^n 个.

8. 设 A, B 为有限集, 以 $|A|, |B|$ 表示 A, B 中元素的个数.

(1) 确定 $|A \cup B| = |A| + |B|$ 成立的条件;

(2) 若 $|A \cap B| \neq 0$, 求 $|A \cup B|$;

(3) 画图说明: $|A \cup B| + |A \cap B| = |A| + |B|$;

(4) 已知某班 50 个学生第一次考试有 26 人不及格, 第二次考试有 21 人不及格, 又已知两次考试都及格的有 17 人. 问: 两次考试都不及格的有多少人?

(5)* 利用(3)的结论, 证明:

$$\begin{aligned} & |A \cup B \cup C| \\ &= |A| + |B| + |C| - (|A \cap B| + |B \cap C| + |A \cap C|) \\ & \quad + |A \cap B \cap C|. \end{aligned}$$

(6)* 设某校足球队有队员 22 人, 篮球队有队员 10 人, 排球队有队员 12 人, 但参加三支球队的学生只有 35 人, 其中有 2 人同时参加三个队, 问: 同时只参加两个队的有几人?

(7)* 在 1 到 40 之间的 40 个整数中, 能被 2, 3, 5 任何一个整除的数有几

个?

9. 已知 $A \cup B = A \cup C$, 是否有 $B = C$?

已知 $A \cap B = A \cap C$, 是否有 $B = C$?

* 如果上述两个条件都成立 , 是否有 $B = C$?

10. 画图说明 $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$.

问 $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$ 是否成立?

11. 利用图形说明 $(A \cup B) \cap C \neq A \cup (B \cap C)$, 并问 : 在什么条件下此式等号成立?

12. 证明 : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

13. 下列集合中的关系 R 是否是等价关系? 如果是 , 说明等价类的意义 , 并写出集合关于 R 的商集 .

(1) 实数集上定义关系 : $xRy \Leftrightarrow x - y$ 为无理数或 0 ;

(2) 集合 $S = \{ (x, y) \mid |x| \leq 1, |y| \leq 1 \}$ 上定义关系 :

$$(x_1, y_1)R(x_2, y_2) \Leftrightarrow x_1 = x_2 ;$$

(3) 集合 $S = \{ (x, y) \mid |x| < +\infty, |y| < +\infty \}$ 上定义关系 :

$$(x_1, y_1)R(x_2, y_2) \Leftrightarrow x_1^2 + y_1^2 = x_2^2 + y_2^2 ;$$

(4) 平面上所有直线组成的集合 L 上分别定义关系 R_1 和 R_2 为 :

$$l_1 R_1 l_2 \Leftrightarrow l_1 \perp l_2 ; \quad l_1 R_2 l_2 \Leftrightarrow l_1 \text{ 与 } l_2 \text{ 相交} ;$$

(5) 集合 $S = \{ (x, y) \mid 0 < |x| < +\infty, 0 < |y| < +\infty \}$ 上定义关系 :

$$(x_1, y_1)R(x_2, y_2) \Leftrightarrow x_1 x_2 > 0 \text{ 且 } y_1 y_2 > 0 ;$$

(6) 集合 $C^* = \{ a + bi \mid a, b \in \mathbf{R}, a \neq 0 \}$ 上定义关系 :

$$(a_1 + b_1 i)R(a_2 + b_2 i) \Leftrightarrow a_1 a_2 > 0 ;$$

(7) 在集合 $A = (a_1, a_2, a_3, a_4)$ 的幂集 $P(A)$ 上定义关系 :

$$A_1 R A_2 \Leftrightarrow |A_1| = |A_2| \text{ (即子集 } A_1, A_2 \text{ 所含元素个数相等)} ;$$

(8) 在集合 $\mathbf{N}^* \times \mathbf{N}^*$ 上定义关系 :

$$(m_1, n_1)R(m_2, n_2) \Leftrightarrow m_1 n_2 = m_2 n_1 .$$

14*. 证明 : 若集合 A 中的二元关系 R 满足 (1) $\forall a \in A, aRa$,

(2) $\forall a, b, c \in A$, 如果有 aRb 和 aRc , 就必有 bRc , 则 R 是等价关系 .

15*. 有人说 , 二元关系 R 如果具有对称性和传递性 , 就必有自反性 , 因为有 aRb , 就有 bRa , 从而有 aRa , 这种说法对吗?

16*. 判断下列集合 A 中的二元关系 R 是否为偏序关系、全序关系 .

(1) $A = \mathbf{Z}^* = \{ a \mid a \in \mathbf{Z}, a \neq 0 \}$, $aRb \Leftrightarrow ab$ 是正整数 , 而且 a 整除 b ;

(2) $A = \mathbf{Q} \times \mathbf{Q} (x_1, y_1) R (x_2, y_2) \Leftrightarrow x_1 < x_2$ 或 $x_1 = x_2$ 而 $y_1 \leq y_2$;

(3) A 是由全体英文单词组成的集合 (单词 1) R (单词 2) 是指: 单词 1 与单词 2 相同或按词典排列法单词 2 排单词 1 之后;

(4) A 是由实轴上所有开区间组成的集合 $I_1 R I_2 \Leftrightarrow I_1 \subset I_2$ (其中 $I_1, I_2 \in A$).

17. 设 f, g, h 均是 \mathbf{Z} 到 \mathbf{Z} 的映射:

$$f: x \mapsto 3x; \quad g: x \mapsto 3x + 1; \quad h: x \mapsto 3x + 2.$$

计算 fg, gf, hg, fgh .

18. 设 $S = \{(x, y) | x, y \in \mathbf{R}\}, X = \{(x, 0) | x \in \mathbf{R}\}$. 定义 S 到 X 的映射 p 为:

$$p(x, y) = (x, 0),$$

即 p 为平面直角坐标系 xOy 中点到 X 轴上的投影. 问: p 是否单射, 满射, 并求 $p^{-1}(p(2, 1)), p^{-1}(p(2, 0)), p^{-1}(1, 0) \cap p^{-1}(2, 0)$.

19. 集合 S 如上题, 定义 S 到 S 的下列映射:

$$(1) r(x, y) = (-y, x); \quad (2) p_1(x, y) = (x, 0);$$

$$(3) p_2(x, y) = (0, y); \quad (4) \varphi_1(x, y) = (x, -y);$$

$$(5) \varphi_2(x, y) = (-y, -x); \quad (6) l(x, y) = (x, y).$$

问: 哪些映射是双射? 如是双射, 求其逆映射.

计算: $rp_1, p_1r, p_1\varphi_2, \varphi_2p_1, p_1p_2, \varphi_1\varphi_2, l p_2$, 说明它们的几何意义, 并分析映射的乘法是否满足交换律和消去律.

20. 设 A, B 是两个集合, $f: A \rightarrow B, g: B \rightarrow A$. 证明: 若 gf 是 A 到 A 的恒等映射, 则 f 是单射, g 是满射.

21*. 设 A, B, C 是三个集合, $f: A \rightarrow B, g: B \rightarrow C$. 证明: 若 gf 是 A 到 C 的双射, 则 f 是单射, g 是满射.

22*. 设 A, B, C, D 是四个集合, $f: A \rightarrow B, g: B \rightarrow C, h: C \rightarrow D$. 证明: 若 gf 和 hg 都是双射, 则 f, g, h 也都是双射.

23*. 设 f 是 A 到 B 的映射, g 是 B 到 A 的满射. 证明: 如果 $fg = I_B$, 则 $gf = I_A$ (即 g 是可逆映射, f 是 g 的逆映射).

24. 先用符号表示简单命题, 然后用符号运算表示下列语句 (复合命题):

例: “今天与明天都下雨”. 如果用 p, q 分别表示“今天下雨”和“明天下雨”, 则此命题可表示为 $p \wedge q$.

(1) 如果他和她都不去, 你就去;

- (2) 我不能既爬山又划船 ;
 (3) 如果你来了 ,那么他在会上发不发言就看你请不请他发言 ;
 (4) 占据空间的、有质量的而且不断变化的叫物质 ;
 (5) 占据空间的、有质量的叫物质 ,而物质是不断变化的 ;
 (6) 除非你努力 ,否则你将失败 ;
 (7) 某次列车不是 8 点开就是 20 点开 .

25. 用量词 \forall, \exists 表示下列语句内容 :

- (1) 数集 S 中任一数 x 都大于 4 而且被 4 整除 ;
 (2) 数集 S 中有的数不大于 4 或不被 4 整除 ;
 (3) 甲班学生都很聪明且很文明或很努力且很文明 (以 $p(x), q(x), r(x)$ 分别表示学生很聪明、很努力、很文明 ,以 X 表示甲班学生集) ;
 (4) 上述命题 (3) 的否命题 ;
 (5) 有些实数是有理数 ;
 (6) 甲班有一个学生叫王强 .

26. 设 $x \in X$ (甲班学生集), $y \in Y$ (乙班学生集), $p(x, y)$ 表示 x 与 y 同姓, $q(x, y)$ 表示 x 与 y 同年龄.

(1) 说明下列命题的含义 :

- (a) $(\forall x \wedge \forall y \wedge p(x, y) \vee q(x, y))$;
 (b) $(\exists x \wedge \exists y \wedge p(x, y) \wedge q(x, y))$;
 (c) $(\forall x \wedge \exists y \wedge p(x, y) \wedge q(x, y))$;
 (d) $(\exists x \wedge \forall y \wedge p(x, y) \vee q(x, y))$.

(2) 写出 (1) 中 4 个命题的否命题 .

27. 设有 m 个 n 元有序数组

$$(a_{i1} \ a_{i2} \ \dots \ a_{in}), \quad i = 1 \ 2 \ \dots \ m.$$

已知: $\forall i \in \{1 \ 2 \ \dots \ m\}, |a_{ii}| > |a_{i1}| + \dots + |a_{i, i-1}| + |a_{i, i+1}| + \dots + |a_{in}|$, 试写出这个已知命题的否命题 .

28. 设平行四边形 $ABCD$ 的对角线向量 $\overrightarrow{AC} = a, \overrightarrow{BD} = b$, 试用 a, b 表示 $\overrightarrow{AB}, \overrightarrow{BC}, \overrightarrow{CD}, \overrightarrow{DA}$.

29. 设 P 是线段 AB 的中点 , 证明 : 对任意一点 $O, \overrightarrow{OP} = \frac{1}{2}(\overrightarrow{OA} + \overrightarrow{OB})$.

30. 设 A, B, C 是任意三点 , 求 $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA}$.

31. 设 M 是三角形 ABC 的重心 , 证明 : 对任意一点 O ,

$$\overrightarrow{OM} = \frac{1}{3}(\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}).$$

如果点 O 是直角坐标系的原点, A, B, C 的坐标分别是 $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$, 试求点 M 的坐标.

32. 设四面体的四个顶点为 O, A, B, C , 边 OA, BC 的中点分别为 M, P , 证明:

$$\overrightarrow{MP} = \frac{1}{2}(\overrightarrow{OC} + \overrightarrow{AB}).$$

33*. 设 O 是点 A 和点 B 连线外一点, 证明: 三点 A, B, C 共线的充要条件是 $\overrightarrow{OC} = \lambda \overrightarrow{OA} + \mu \overrightarrow{OB}$, 其中 $\lambda + \mu = 1$.

34. 已知点 $A(1, 2, 4)$ 和 $B(2, 0, -1)$,

(1) 求 \overrightarrow{AB} 的坐标表示式及 $|\overrightarrow{AB}|$;

(2) 求点 C 的坐标, 使 $\overrightarrow{AC} = (0, 4, 1)$;

(3) 求点 C , 使 $\overrightarrow{CA} + \overrightarrow{CB} = \frac{1}{3} \overrightarrow{AB}$.

35. 已知 $a = (1, -2, 3), b = (2, 1, 0), c = (6, -2, 6)$,

(1) $a + b$ 和 c 是否平行?

(2) 求 $a \cdot b, a \cdot c, a \cdot b, a \cdot c$;

(3) 求 $a \times b, (a \times b) \cdot c$;

(4) 设 $x = 3a + 4b - c, y = 2b + c$, 求 x, y .

36. 已知 $|a| = 3, |b| = 2, a \cdot b = \frac{\pi}{3}$, 试求 $3a + 2b$ 与 $2a - 5b$ 的内积及其夹角.

37. 证明下列各对向量互相垂直.

(1) $(4, -1, 1)$ 与 $(-2, -5, 3)$;

(2) $(c \cdot a)b - (b \cdot a)c$ 与 a ;

(3) $a \times b$ 与 $\lambda a + \mu b$ (λ, μ 为任意常数).

38. 证明: 三角形三条中线的长度平方和等于三条边长之平方和的 $\frac{3}{4}$.

39. 证明: $\overrightarrow{AB} \cdot \overrightarrow{CD} + \overrightarrow{BC} \cdot \overrightarrow{AD} + \overrightarrow{CA} \cdot \overrightarrow{BD} = 0$.

40. 已知 $\overrightarrow{OA} = (1, 0, -1), \overrightarrow{OB} = (1, -2, 0), \overrightarrow{OC} = (3, -4, -1)$,

(1) 求 $\overrightarrow{OA} \times \overrightarrow{OB}$, 并证明 $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ 共面;

(2) 求 λ , 使 $(\lambda \overrightarrow{OA} + \overrightarrow{OB}) \perp \overrightarrow{OC}$;

(3) 求 λ, μ , 使 $\lambda \overrightarrow{OA} + \mu \overrightarrow{OB} \parallel \overrightarrow{OC}$.

41. 已知 $a = (1, 0, -1), b = (1, -2, 0), c = (1, 1, 1)$,

(1) 求 $a \times b$ 和以 a, b 为邻边的平行四边形的面积;

(2) 求以 a, b, c 为邻边的平行六面体的体积;

(3) 求 $(a \times b) \times c, a \times (b \times c)$.

42. 设 $a = (a_1, a_2, a_3), b = (b_1, b_2, b_3), c = (c_1, c_2, c_3)$,

(1) 证明 $a \cdot (b \times c) = b \cdot (c \times a)$;

(2) 证明 如果 $a \times b + b \times c + c \times a = 0$ 则 a, b, c 共面.

43. 下列命题是否成立?

(1) 如果 $a \cdot b = a \cdot c$ 且 $a \neq 0$ 则 $b = c$;

(2) 如果 $a \times b = a \times c$ 且 $a \neq 0$ 则 $b = c$.

44. 已知 a, b 满足下列条件, 讨论 a, b 之间的关系:

(1) $(a \cdot b)^2 = a^2 b^2$ 其中 $a^2 = a \cdot a, b^2 = b \cdot b$;

(2) a 与 $a \times b$ 共线;

(3) $a, b, a \times b$ 共面.

45. 求下列平面的方程:

(1) 过点 $P_0(3, 0, -1)$ 且平行于平面 $x - 2y + 5z - 1 = 0$;

(2) 过点 $P_0(1, 2, -3)$ 且垂直于点 $A(0, 0, 1)$ 和 $B(1, -3, 4)$ 点的连线;

(3) 过 x 轴且平行于向量 $a = (1, -2, -1)$;

(4) 过三点 $A(1, 1, 1), B(-1, 0, 2), C(4, -3, 1)$;

(5) 过 $2x + y - 4 = 0$ 与 $y + 2z = 0$ 的交线, 且过点 $P_0(2, -1, -1)$;

(6) 在 x, z 轴上的截距依次为 (-3) 和 2 , 且过点 $P_0(6, -10, 1)$.

46. 求下列直线的标准方程和参数方程:

(1) 过点 $A(2, 5, -3)$ 和 $B(0, 4, -3)$;

(2) 过点 $A(2, -1, 0)$ 且垂直于平面 $x - 2y + 5z = 1$;

(3) 过点 $A(2, -1, 0)$ 且与 z 轴垂直相交;

(4) 过点 $A(2, -1, 0)$ 且垂直于 $A, B(1, 1, 1), C(0, 1, -2)$ 所确定的平面;

(5) 平面 $x - y = 0$ 与 $x + y + z = 1$ 的交线;

(6) 过点 $(-1, -4, 3)$ 且与直线 L_1, L_2 都垂直,

$$L_1: \begin{cases} 2x - 4y + z = 1 \\ x + 3y = -5 \end{cases}, \quad L_2: \begin{cases} x = 2 + 4t \\ y = -1 - t \\ z = -3 + 2t \end{cases}.$$

47. 已知: 平面 $\pi: 4x + 4y - 5z + 12 = 0$;

$$\text{直线 } L_1: \frac{x}{2} = \frac{y+3}{3} = \frac{z}{4}; \quad L_2: \frac{x-1}{1} = \frac{y+2}{1} = \frac{z-2}{2}.$$

(1) 直线 L_1 是否在平面 π 上?

(2) L_1 与 L_2 共面吗? 相交吗? 如相交, 求交点.

48. 用高斯消元法解下列方程组:

$$(1) \begin{cases} x_1 - x_2 + x_3 = 1 \\ 2x_1 - x_2 + 2x_3 = 3; \end{cases} \quad (2) \begin{cases} 2x_1 - x_2 + x_3 - x_4 = 0 \\ x_1 - x_2 + x_3 - \frac{1}{2}x_4 = 0; \end{cases}$$

$$(3) \begin{cases} 4x_1 - x_2 - x_3 = 0 \\ -x_1 + 4x_2 - x_4 = 3 \\ x_1 - 4x_3 + x_4 = -3 \\ x_2 + x_3 - 4x_4 = 0; \end{cases} \quad (4) \begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 = 4 \\ x_2 - x_3 + x_4 = -3 \\ x_1 + 3x_2 - 3x_4 = 1 \\ 7x_2 - 3x_3 - x_4 = 3; \end{cases}$$

$$(5) \begin{cases} x_1 + 2x_2 - x_3 + 2x_4 = 3 \\ x_1 + x_2 - x_3 + x_4 = 2 \\ x_1 - x_2 - x_3 - x_4 = 1 \\ 3x_1 + 2x_2 - 3x_3 + 2x_4 = 5; \end{cases} \quad (6) \begin{cases} x_1 - x_2 + x_3 - x_4 = 0 \\ 2x_1 + 4x_2 - 5x_3 + 7x_4 = 0 \\ 3x_1 + 3x_2 - 4x_3 + 6x_4 = 0 \\ 3x_1 - 3x_2 + 3x_3 - 2x_4 = 0. \end{cases}$$

49. 将军点兵, 三三数之剩二, 五五数之剩三, 七七数之剩二, 问兵几何(求在 500 至 1 000 范围内的解)?

50. 百鸡术: 母鸡每只 5 元, 公鸡每只 3 元, 小鸡三只 1 元, 百元买百鸡, 各买几何?

51. 设“ $*$ ”是正整数集 \mathbf{N}^* 上的一个二元运算, 判断下列运算是否满足结合律:

$$(1) a * b = \max(a, b); \quad (2) a * b = \min(a, b);$$

$$(3) a * b = a + b + 1; \quad (4) a * b = a + 2b;$$

$$(5) a * b = \begin{cases} \min(a, b), & \text{当 } \min(a, b) < 10, \\ \max(a, b), & \text{当 } \min(a, b) \geq 10. \end{cases}$$

52. 集 X 上的一个二元关系“ \oplus ”定义为: $a \oplus b = a$. 讨论运算 \oplus 是否可结合和可交换.

53. 孩子头发的颜色是由其父母头发的颜色所决定, 其对应规则如下表所示:

孩子		母	
		浅色	深色
父	浅色	浅色	深色
	深色	深色	深色

这是集合 $S = \{\text{浅色}, \text{深色}\}$ 上的一个二元运算“ $*$ ”的表达式. 试问: S 关于“ $*$ ”的单位元是什么?

54. 在正实数集 \mathbf{R}^+ 上定义两个二元运算“ \circ ”和“ $*$ ”为：

$$a \circ b = ab, \quad a * b = a^b.$$

证明：“ $*$ ”对“ \circ ”是不可左分配的，“ $*$ ”是不可结合的。

55. 设 $G = \{e, a\}$, 试列出二元素群 G 的运算表。

56. 设 $G = \{e, a, b\}$, 试列出三元素群 G 的运算表。

57. 设 $R_\theta = \left\{0, \frac{\pi}{3}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{5\pi}{3}\right\}$ 表示平面几何图形绕形心逆时针旋转角度的 6 种可能情况。在 R_θ 上定义二元运算“ $*$ ”为： $\forall \alpha, \beta \in R_\theta, \alpha * \beta = \alpha + \beta$ (表示平面图形连续旋转 α 和 β 所得到的总旋转角度), 并规定旋转 2π 等于原来的状态, 即看作没有旋转。验证： $R_\theta : *$ 是一个交换群。

58*. 写出模 8 剩余类加群 $Z_8 : \oplus$ 的所有子群。

59*. 在 Z_{15} 中, 找出方程 $x^2 = \bar{1}$ 的全部根。

补 充 题

1. 设 $A_i (i = 1, 2, \dots, n)$ 是 X 的子集, 证明：

$$(1) \overline{\bigcup_{i=1}^n A_i} = \bigcap_{i=1}^n \overline{A_i}; \quad (2) \overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \overline{A_i}.$$

2. 在复数集 \mathbf{C} 中, 能否定义二元关系 R 使 \mathbf{C} 成为全序集?

3. 在正有理数集 $\mathbf{Q}^+ = \left\{ \frac{p}{q} \mid p, q \in \mathbf{N}^* \right\}$ 中, 如何定义一种二元关系 R , 使 (\mathbf{Q}^+, R) 为全序集。

4. 在集合 $A = \{a, b, c\}$ 上定义一个既非对称, 又非反对称的二元关系。

5. 我们把 $A \times A$ 的一个子集 R 定义为 A 上的一个二元关系, 即 $\forall a, b \in A$, 如果 $(a, b) \in R$ 就说 a, b 具有关系 R , 即 aRb 。

设 $A = \{a, b, c\}$, 试判断 A 上的下列二元关系是否具有自反性、对称性、反对称性和传递性：

$$(1) R_1 = \{(a, a), (a, b), (a, c), (c, c)\};$$

$$(2) R_2 = \{(a, a), (b, b), (c, c), (a, b), (b, a)\};$$

$$(3) R_3 = \{(a, a), (a, b), (b, b), (b, c)\};$$

$$(4) R_4 = \{(a, a), (b, b), (c, c)\};$$

$$(5) R_5 = \{(a, b), (a, c), (c, a)\}.$$

6. 设 $A = \{1, 2, 3, 4\}$, R_1, R_2 均是 $A \times A$ 的子集, 且

$$R_1 = \{(x, y) : 2 \mid (x - y)\}; \quad R_2 = \{(x, y) : 3 \mid (x - y)\}.$$

(1) 求 $R_1, R_2, R_1 \cup R_2, R_1 \cap R_2, R_1 \setminus R_2, \overline{R_1 \cup R_2}$;

(2) 指出集合 A 上的上述二元关系的意义.

7. 举出一个没有自反性、对称性和传递性的关系的例子.

8*. 证明:

(1) $p \Leftrightarrow q$ 与 $(p \rightarrow q) \wedge (q \rightarrow p)$ 是等价命题;

(2) $p \Leftrightarrow q$ 与 $(p \wedge q) \vee (\neg p \wedge \neg q)$ 是等价命题;

(3) $p \vee (p \wedge q) \Leftrightarrow p \wedge (p \vee q)$ (吸收律);

(4) $p \rightarrow q$ 与 $\neg p \rightarrow \neg q$ 不是等价命题.

9*. 利用 $(p \rightarrow q) \Leftrightarrow \neg p \vee q$ 证明:

(1) $p \rightarrow (q \rightarrow r) \Leftrightarrow q \rightarrow (p \rightarrow r)$;

(2) $p \rightarrow (q \rightarrow r) \Leftrightarrow \neg r \rightarrow (q \rightarrow \neg p)$;

(3) $\neg(p \rightarrow q) \Leftrightarrow p \wedge \neg q$.

10*. 试从逻辑的角度分析下面命题所存在的问题:

(1) “苛政猛于虎”

某地有虎,一老妇人的亲人为虎所害,但她不愿搬走,因为“无苛政”.由此推出:“苛政猛于虎”.

(2) 资本主义社会有商品经济,社会主义社会不是资本主义社会,由此推出:社会主义社会没有商品经济.

11. 在有理数集 \mathbb{Q} 上定义二元运算“ $*$ ”为 $a * b = a + b - ab$. 讨论运算“ $*$ ”是否满足结合律和交换律. \mathbb{Q} 关于运算“ $*$ ”是否存在单位元. \mathbb{Q} 中每个元素是否可逆. 可逆元的逆元是什么?

12. 设 $S = \mathbb{Q} \setminus \{1\}$, S 上的二元运算“ $*$ ”如第 11 题所定义.

(1) 证明 $S; *$ 是一个群;

(2) 求方程 $2 * x * 3 = 7$ 在 S 中的解.

13. 求有理数加群 $\mathbb{Q}; +$ 的一个子群 H , 使 $H \neq \mathbb{Z}$ 且 $H \neq \mathbb{Q}$.

部分习题和补充题答案

习题

1. $\{4, 5\}, \{3, 6\}$. 2. $A \cap B, \mathbb{Z}^* \setminus B, (C \cap \mathbb{Z}^*) \setminus B$.

3. N, M (l 为 n 与 k 的最小公倍数), \emptyset .

4. $\{\emptyset, \{1\}, \{-1\}, \{-1, 1\}\}, \{(-1, -1), (-1, 1), (1, -1), (1, 1)\}$.

5. $A = \{\text{京, 津, 沪}\}, (A \times A) \setminus B, B = \{(\text{京, 京}), (\text{津, 津}), (\text{沪, 沪})\}$.

8. (1) $A \cap B = \emptyset$; (2) $|A| + |B| = |A \cup B|$; (4) 14; (6) 5;

(7) 30.9. 未必有, 未必有, 必有.

13. (1) 不是; (2) 是, 矩形内平行于 y 轴的直线段, $S/R = \{\text{线段 } x = c \mid |y| \leq 1\} : \{c \mid |c| \leq 1\}$; (3) 是, 以原点为心的同心圆, $S/R = \{x^2 + y^2 = r^2 \mid r \geq 0\}$; (4) 均不是; (5) 是, $S/R = \{\text{不含坐标轴的四个象限}\}$; (6) 是, $S/R = \{\text{复平面的左半平面和右半平面}\}$; (7) 是, $S/R = \{S_i \mid i = 0, 1, 2, 3, 4\}$, S_i 是含 i 个元素的所有子集组成的集合, $S_0 = \{\emptyset\}$; (8) 是, $S/R = \{(m, n) \mid \frac{m}{n} = c\} \mid c \in \mathbf{Q}\}$.

15. 不对. 16. (1)(4) 均是偏序, 不是全序; (2)(3) 均是全序.

17. $x \mapsto 9x + 3$; $x \mapsto 9x + 1$; $x \mapsto 9x + 5$; $x \mapsto 27x + 21$.

18. 不单, 满, $\{(2, y) \mid y \in R\}, \{(2, 0)\}, \emptyset$.

19. (1) 双射, $r^{-1}(x, y) = (y, -x)$; (2)(3) 非双射; (4) $\varphi_1^{-1} = \varphi_1$; (5) $\varphi_2^{-1} = \varphi_2$; (6) $I^{-1} = I$.

28. $\frac{a-b}{2}, \frac{a+b}{2}, \frac{b-a}{2}, \frac{-a-b}{2}$. 30. 0.

34. (1) $(1, -2, -5), \sqrt{30}$; (2) $(1, 6, 5)$; (3) $(\frac{4}{3}, \frac{4}{3}, \frac{7}{3})$.

35. (1) 平行; (2) 0, 28, $\frac{\pi}{2}$, $\arccos \frac{14}{\sqrt{266}}$; (3) $(-3, 6, 5), 0$; (4) $\arccos 1$.

36. $-19, \pi - \arccos \frac{1}{\sqrt{28}}$.

40. (1) $(-2, -1, -2)$; (2) $-\frac{11}{4}$; (3) $\mu = 2\lambda$.

41. (1) $(-2, -1, -2), 3$; (2) 5; (3) $(1, 0, -1), (-1, -1, -1)$.

43. 均不成立. 44. (1) 共线; (2)(3) $a \times b = 0$.

45. (1) $x - 2y + 5z + 2 = 0$; (2) $x - 3y + 3z + 14 = 0$;

(3) $y - 2z = 0$; (4) $4x + 3y + 11z - 18 = 0$;

(5) $3x + y - z - 6 = 0$; (6) $-\frac{x}{3} - \frac{y}{4} + \frac{z}{2} = 1$.

46. (1) $\frac{x}{2} = \frac{y-4}{1} = \frac{z+3}{0}$; $x = 2t$, $y = 4+t$, $z = -3$;

(2) $\frac{x-2}{1} = \frac{y+1}{-2} = \frac{z}{5}$; $x = 2+t$, $y = -1-2t$, $z = 5t$;

(3) $\frac{x}{2} = \frac{y}{-1} = \frac{z}{0}$; $x = 2t$, $y = -t$, $z = 0$;

(4) $\frac{x-2}{3} = \frac{y+1}{2} = \frac{z}{-1}$; (5) $-\frac{x}{-1} = \frac{y}{-1} = \frac{z-1}{2}$;

$$(6) \frac{x+1}{12} = \frac{y+4}{46} = \frac{z-3}{-1}.$$

47. (1) 在; (2) 共面 相交 $(0, -3, 0)$.

48. $(1) \lambda(2, 1, 0) + \mu(-1, 0, 1)$; $(2) \lambda_1(1, 0, 0, 2) + \lambda_2(0, 1, 1, 0)$;

$(3) \left(\frac{1}{2}, 1, 1, \frac{1}{2} \right)$; $(4) \lambda(-8, 3, 6, 0) + \mu(0, 1, 2, 1)$; (5) 无解;

(6) $\mu(1, 7, 6, 0)$.

49. 548, 653, 758, 863, 968.

50. $(8, 11, 81)$ 或 $(12, 4, 84)$ 或 $(4, 18, 78)$.

51. (1)(2)(3)(5) 满足 (4) 不满足.

52. 可结合, 不可交换. 53. 浅色. 55.

\circ	e	a
e	e	a
a	a	e

56. 提示 运算表中每行(列)的三个元素互不相同.

58. $\{\bar{0}\}, \{\bar{0}, \bar{A}\}, \{\bar{0}, \bar{2}, \bar{A}, \bar{6}\}, \mathbb{Z}_8$.

59. $\bar{1}, \bar{A}, \bar{11}, \bar{14}$.

补充题

2. 能, 关系如习题 16(2) 所给, 将其中 $\mathbf{Q} \times \mathbf{Q}$ 改为 $\mathbf{R} \times \mathbf{R}$ 即可.

4. 定义 $(a, b) \in R$ 为 aRb , $R = \{(a, a), (b, b), (c, c), (a, b), (a, c), (c, a)\}$.

5. (1) 具有反对称性和传递性, (2) 具有自反性、对称性和传递性, (3) 具有反对称性, (4) 全具有, (5) 全不具有.

6. (1) $R_1 = \{(1, 1), (2, 2), (3, 3), (4, 4), (1, 3), (3, 1), (2, 4), (4, 2)\}$,
 $R_2 = \{(1, 1), (2, 2), (3, 3), (4, 4), (1, 4), (4, 1)\}$, $\overline{R_1 \cup R_2} = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3)\}$ (2) R_1 表示 2 整除 x 与 y 之差, R_2 表示 x 与 y 之差为 3 的倍数. $R_1 \cap R_2$ 表示 x 与 y 之差既是 2 又是 3 的倍数.

7. 父子关系或 5 题 (3) 所给关系.

11. 满足结合律和交换律, 单位元是 0, $a \neq 1$ 时可逆, $a^{-1} = a/(a-1)$.

12. $(2)x = 4$. 13. $H = \{q/p \mid q \in \mathbf{Z}, p \text{ 为固定的非 1 正整数}\}$.

第 2 章 线性空间 内积空间

线性空间是一个重要的代数结构,它广泛应用于现代数学的很多领域.内积空间是具有几何度量性的线性空间.

线性代数主要是研究有限维线性空间和线性映射的基本性质.线性代数在其发展史上,最早是研究一般线性方程组的求解问题.随着研究的深入,人们发现线性方程组的解的理论,仅仅是有限维线性空间和线性映射理论的一个具体应用(或说一个具体模型).为了更好地培养学生的数学抽象思维能力以及应用理论解决具体问题的能力,我们遵循“在宏观上从一般到具体,在微观上从具体到一般”的原则,将从阐明线性空间和线性映射的基本理论入手,来揭示线性代数所研究的主要问题.学好第 2、3 章是学好线性代数的关键.

本章主要讨论有限维线性空间和内积空间的结构.首先介绍线性空间的公理化定义及其简单性质,进而利用线性相关性的概念,讨论有限维线性空间的基和维数以及向量在基下的坐标;为搞清线性空间的结构,还要了解子空间的概念以及子空间的交、和与直和.内积空间是定义了内积运算从而具有几何度量性的线性空间,重点是内积的公理化定义,向量长度和夹角的度量方法,有限维实内积空间即欧氏空间的单位正交基以及正交子空间与正交补的概念.

下面先阐述线性空间的定义及其简单性质,然后详细讨论其结构.

2.1 线性空间的定义及其简单性质

在第 1 章中讲过,全体空间几何向量组成的集合 \mathbf{R}^3 关于向量的加法构成交换群,再定义向量的数量乘法,就使 \mathbf{R}^3 中任一向量 $\alpha = (a_1, a_2, a_3)$ 可用基本向量组 $\{e_1, e_2, e_3\}$ 作数乘和加法运算来表示,即

$$\alpha = a_1 e_1 + a_2 e_2 + a_3 e_3.$$

这表明 \mathbf{R}^3 中的向量通过两种运算确立了相互间的关系.这是区别于群、环、域的一种代数结构.

与 \mathbf{R}^3 的情况类似, $R[x]_n$ 关于多项式加法也构成交换群,如果再定义数与多项式相乘的运算,则 $R[x]_n$ 中任一多项式显然可用其中 n 个基本多项式

$\{1, x, \dots, x^{n-1}\}$ 作数乘和加法运算来表示. 在数学中类似的情况还有很多.

线性空间(或称向量空间)正是以上述数学问题为背景而抽象出来的一种代数结构, 它的下述公理化定义就是上述不同数学问题的共有特征的概括.

定义 2.1 设 V 是一个非空集合, F 是一个域, 在 V 和 $F \times V$ 上定义两种运算: 一是 V 上的代数运算, 称为加法, 记作“+”; 二是定义在集 $F \times V$ 上取值于 V 的一元运算, 称为数量乘法(简称数乘), 即 $F \times V$ 中每个元素 (λ, α) $\mapsto \lambda\alpha \in V$, 如果:

(1) $V : +$ 是一个交换群(加法群);

(2) 数量乘法满足 4 条性质, 即 $\forall \alpha, \beta \in V, \forall \lambda, \mu \in F$ 以及域 F 的乘法单位元 1, 有:

$$\begin{aligned} 1\alpha &= \alpha, \\ \lambda(\mu\alpha) &= (\lambda\mu)\alpha, \\ (\lambda + \mu)\alpha &= \lambda\alpha + \mu\alpha, \\ \lambda(\alpha + \beta) &= \lambda\alpha + \lambda\beta, \end{aligned}$$

则称 V 对于上述两种运算在域 F 上构成一个线性空间, 简称 V 为域 F 上的线性空间, 记作 $V(F)$. 如果 F 是实(复)数域, 则称 V 为实(复)数域上的线性空间, 简称实(复)空间.

线性空间也称为向量空间, 其元素也称为向量. 线性空间中的运算称为线性运算. 加法群 $V : +$ 的单位元叫做零元, 记作 0 .

显然, 全体空间几何向量和全体 n 元实向量分别组成的集合 \mathbf{R}^3 和

$$\mathbf{R}^n = \{ (a_1, a_2, \dots, a_n) \mid a_1, a_2, \dots, a_n \in \mathbf{R} \}$$

对向量的加法和数乘运算都构成实数域上的线性空间. 它们是抽象的线性空间的两个基本的具体模型. 再举一些例如下:

例 1 系数属于数域 F 的全体多项式组成的集合, 以及次数小于 n 的全体多项式同零多项式一起组成的集合, 它们对多项式加法与数乘多项式运算在数域 F 上都构成线性空间. 因为: 它们关于两种运算封闭; 它们对加法构成交换群(见节 1.10.1 例 3); 数乘多项式运算显然满足定义中的 4 条性质. 但是对固定的 n , 数域 F 上的多项式集合

$$\{p(x) \mid p(x) = a_0 + a_1x + \dots + a_nx^n, a_n \neq 0\}$$

对同样的运算不构成线性空间, 因为集合对加法不封闭.

例 2 第 1 章节 1.8 中例 2 的线性方程组, 当其右端常数项全为零时的齐次线性方程组, 其解集合

$$S = \{X = k_1\alpha + k_2\beta \mid \alpha = (1 \ 1 \ 0 \ 0 \ 0), \\ \beta = (-7 \ 0 \ -4 \ 3 \ 1), k_1, k_2 \in \mathbf{R}\}$$

关于向量的加法和数乘运算在 \mathbf{R} 上构成一个线性空间(称为齐次线性方程组的解空间). 因为: 容易验证集 S 关于两种运算封闭, 即 $\forall X_1, X_2 \in S, \forall k \in \mathbf{R}$, 均有

$$X_1 + X_2 \in S, \quad kX_1 \in S,$$

S 关于 $+$ 是一个交换群, 零元 $\mathbf{0} = (0 \ 0 \ 0 \ 0 \ 0)$ 是齐次线性方程组的零解, 每个元 X 的负元是 $(-1)X$; 向量的数乘运算满足定义中的 4 条性质.

例 3 定义在区间 $[a, b]$ 上的全体实值函数, 对通常的函数加法和实数与函数的乘法在实数域 \mathbf{R} 上构成一个线性空间. 其零元是零函数 $\theta(x)$, 即 $\forall x \in [a, b], \theta(x) = 0$; 每个函数 $f(x)$ 的负元是 $-f(x)$.

例 4 实数集 \mathbf{R} 在实数域 \mathbf{R} 上对实数的加法和实数与实数的乘法, 构成实数域 \mathbf{R} 上的一个线性空间. 但是, 实数集 \mathbf{R} 在复数域 \mathbf{C} 上, 对实数的加法和复数与实数的乘法不构成复数域 \mathbf{C} 上的线性空间, 因为复数与实数的乘积一般不属于实数集 \mathbf{R} , 即 \mathbf{R} 对数乘不封闭. 而复数集 \mathbf{C} 在实数域和复数域上, 对相应的数的加法和数乘运算分别构成实数域 \mathbf{R} 和复数域 \mathbf{C} 上的线性空间.

总之, 对给定的非空集 V 和域 F , 如果 V 对定义的加法和数乘运算不封闭, 或者 V 对加法不构成交换群, 或者数乘不满足 4 条性质, 那么 V 对定义的两运算不构成线性空间.

还要强调指出, 线性空间作为一种代数结构, 定义中非空集是抽象的, 域是泛指, 两种运算不是具体的. 研究一个具体问题要将其视为线性空间, 其中的集合、域和运算需要符合定义的规定. 由运算性质约定运算正是代数结构中运算的本质的特征. 正是由于代数结构的这种高度抽象性, 就使它有更广泛的应用性.

在线性空间 $V(F)$ 中, 我们定义减法为: $\alpha - \beta = \alpha + (-\beta)$, 其中 $(-\beta)$ 为 β 的负元(即加法的逆元).

由线性空间的定义可推出下面的简单性质:

1. 线性空间作为加法群, 它的零元(加法群的单位元)是唯一的, 每一个元的负元(即加法的逆元)是唯一的;
2. 数乘运算的分配律对元素的减法和数量的减法也都成立, 即 $\forall \alpha, \beta \in V, \lambda, \mu \in F$,

$$\lambda(\alpha - \beta) = \lambda\alpha - \lambda\beta, \quad (2-1)$$

$$(\lambda - \mu)\alpha = \lambda\alpha - \mu\alpha. \quad (2-2)$$

事实上,由

$$\begin{aligned}\lambda(\alpha - \beta) + \lambda\beta &= \lambda[(\alpha - \beta) + \beta] = \lambda[\alpha + ((-\beta) + \beta)] \\ &= \lambda(\alpha + 0) = \lambda\alpha,\end{aligned}\quad ①$$

$$(\lambda - \mu)\alpha + \mu\alpha = [(\lambda - \mu) + \mu]\alpha = \lambda\alpha,\quad ②$$

并在①、②式两边分别加 $-(\lambda\beta)$, $-(\mu\alpha)$ 即得上述分配律.

3. 在(2-1)式中分别令 $\alpha = \beta$, $\alpha = 0$ 在(2-2)式中分别令 $\lambda = \mu$ 和 $\lambda = 0$, 即得数乘运算的下列性质:

$$\lambda 0 = 0, \quad \lambda(-\beta) = -(\lambda\beta), \quad (2-3)$$

$$0\alpha = 0, \quad (-\mu)\alpha = -(\mu\alpha). \quad (2-4)$$

特别当 $\mu = 1$ 时,有 $(-1)\alpha = -\alpha$.

4. 若 $\lambda\alpha = 0$ 则 $\lambda = 0$ 或 $\alpha = 0$.

事实上,如果 $\lambda \neq 0$ 则 $\alpha = 1\alpha = (\lambda^{-1}\lambda)\alpha = \lambda^{-1}(\lambda\alpha) = \lambda^{-1}0 = 0$.

由上述性质可知 线性空间 $V(F)$ 中元素作线性运算所得的方程

$$\lambda\beta + \lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_r\alpha_r = 0,$$

当 $\lambda \neq 0$ 时,其解为

$$\beta = -\lambda^{-1}\lambda_1\alpha_1 - \lambda^{-1}\lambda_2\alpha_2 - \dots - \lambda^{-1}\lambda_r\alpha_r.$$

定义2.2 设 $V(F)$ 是一个线性空间, $\alpha_i \in V$, $\lambda_i \in F$ ($i = 1, 2, \dots, m$), 则向量

$$\alpha = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_m\alpha_m$$

称为向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ 在域 F 上的线性组合,或说 α 在域 F 上可用向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ 线性表示.

例如,在 \mathbb{R}^3 中,任一向量 $\alpha = (a_1, a_2, a_3)$ 可由基本向量组 $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ 线性表示,即

$$\alpha = a_1e_1 + a_2e_2 + a_3e_3.$$

例5 证明 在 $\mathbb{R}[x]_3 = \{a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$ 中任一个多项式 $p(x)$ 都可唯一地表示为 $p_1(x) = 1 + x$, $p_2(x) = 1 - x$, $p_3(x) = x + x^2$ 的线性组合.

证 设

$$p(x) = a + bx + cx^2.$$

令

$$p(x) = \lambda_1 p_1(x) + \lambda_2 p_2(x) + \lambda_3 p_3(x),$$

于是得

$$\begin{cases} \lambda_1 + \lambda_2 = a \\ \lambda_1 - \lambda_2 + \lambda_3 = b \\ \lambda_3 = c. \end{cases}$$

这个 $\lambda_1, \lambda_2, \lambda_3$ 的三元线性方程组有唯一解

$$\lambda_1 = \frac{a+b-c}{2}, \quad \lambda_2 = \frac{a-b+c}{2}, \quad \lambda_3 = c,$$

故 $\mu(x)$ 可唯一地表示为 $p_1(x), p_2(x), p_3(x)$ 的线性组合, 即

$$\mu(x) = \frac{a+b-c}{2} p_1(x) + \frac{a-b+c}{2} p_2(x) + c p_3(x).$$

2.2 线性子空间

在线性空间 $V(F)$ 中, V 的子集 W 关于 $V(F)$ 中的线性运算可能封闭, 也可能不封闭. 例如 \mathbf{R}^3 的下列子集:

$$W_1 = \{ (x_1, x_2, x_3) \mid x_1 - x_2 + 5x_3 = 0 \},$$

$$W_2 = \{ (x_1, x_2, x_3) \mid x_1 - x_2 + 5x_3 = 1 \},$$

它们分别是由起点在原点, 终点在平面 $x_1 - x_2 + 5x_3 = 0$ 和平面 $x_1 - x_2 + 5x_3 = 1$ 上的全体向量所组成的两个子集. 以后, 简称 W_1 是过原点的平面 $x_1 - x_2 + 5x_3 = 0$ 上的全体向量; W_2 是不过原点的平面 $x_1 - x_2 + 5x_3 = 1$ 上的全体向量. 容易验证, W_1 关于向量的加法和数乘是封闭的, 而 W_2 不封闭.

由于 W_1 关于 \mathbf{R}^3 的线性运算是封闭的, 所以 W_1 关于 \mathbf{R}^3 的线性运算也构成一个线性空间(下面将一般地证明), 并称之为 \mathbf{R}^3 的子空间. 而 W_2 关于 \mathbf{R}^3 的线性运算不封闭, 自然对此运算不能构成线性空间, 因而不是 \mathbf{R}^3 的子空间. 下面给出子空间的一般定义及判别定理.

定义 2.3 设 W 是线性空间 $V(F)$ 的非空子集, 如果 W 对 V 中的运算也构成域 F 上的线性空间, 则称 W 为 V 的线性子空间(简称子空间).

定理 2.1 线性空间 $V(F)$ 的非空子集 W 为 V 的子空间的充分必要条件是 W 对于 $V(F)$ 的线性运算封闭.

证 必要性是显然的, 否则 W 关于 V 的运算不构成线性空间. 下面证充分性.

由于 W 是 V 的子集, 所以 $V(F)$ 中数乘满足的 4 条性质及加法的交换律

与结合律对 W 都成立. 因此只要再证 V 的零元 $\mathbf{0} \in W$, W 中每个元 α 的负元 $(-\alpha) \in W$. 由于 W 对数乘封闭, 所以取 $\lambda = 0$ 和 -1 , 即得

$$0\alpha = \mathbf{0} \in W, \quad (-1)\alpha = -\alpha \in W,$$

故 W 是 $V(F)$ 的线性子空间.

例 1 在线性空间 V 中, 仅含零元的子集 $\{\mathbf{0}\}$ 是 V 的一个子空间, 叫做零子空间; V 本身也是 V 的一个子空间. 它们称为 V 的平凡子空间, V 的其它子空间称为非平凡子空间.

例 2 节 2.1 例 2 的解空间 S 是 \mathbf{R}^5 的一个子空间.

例 3 $R[x]_3$ 是 $R[x]$ 的一个子空间,

$R[x]_2 = \{a_0 + a_1x \mid a_0, a_1 \in \mathbf{R}\}$ 是 $R[x]_3$ 的一个子空间.

例 4 \mathbf{R}^3 的下列子集

$$W_1 = \left\{ (x, y, z) \mid \frac{x}{3} = \frac{y}{2} = z \right\},$$

$$W_2 = \{ (x, y, z) \mid x + y + z = 1 \text{ 且 } x - y + z = 1 \}$$

分别是过原点的直线 $\frac{x}{3} = \frac{y}{2} = z$ 和不过原点的直线 $x + y + z = 1, x - y + z = 1$ 上的全体向量组成的 \mathbf{R}^3 的子集. 读者不难验证 W_1 是 \mathbf{R}^3 的一个子空间, 而 W_2 不是子空间.

定义 2.4 设 S 是线性空间 $V(F)$ 的非空子集, 我们把 S 中所有的有限子集 (即 S 中任意 k 个向量 ($k = 1, 2, 3, \dots$) 组成的子集) 在域 F 上的一切线性组合所组成的 $V(F)$ 的子集合, 称为 S 的线性扩张, 记作 $L(S)$, 即

$$L(S) = \{ \lambda_1 \alpha_1 + \dots + \lambda_k \alpha_k \mid \lambda_1, \dots, \lambda_k \in F, \alpha_1, \dots, \alpha_k \in S, k \in \mathbf{N}^* \}.$$

定理 2.2 线性空间 $V(F)$ 的非空子集 S 的线性扩张 $L(S)$ 是 V 中包含 S 的最小子空间.

证 $L(S)$ 显然包含 S . 下面先证 $L(S)$ 是 V 的一个子空间.

设 $\alpha, \beta \in L(S)$ 则存在 $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n \in S$ 使得

$$\alpha = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m, \quad \beta = \mu_1 \beta_1 + \dots + \mu_n \beta_n,$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_m, \mu_1, \dots, \mu_n \in F$, 于是

$$\alpha + \beta = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m + \mu_1 \beta_1 + \dots + \mu_n \beta_n \in L(S).$$

$\forall \lambda \in F$, 也有

$$\lambda \alpha = \lambda \lambda_1 \alpha_1 + \lambda \lambda_2 \alpha_2 + \dots + \lambda \lambda_m \alpha_m \in L(S),$$

所以 $L(S)$ 是 V 的一个子空间. 再证 $L(S)$ 是包含 S 的最小子空间.

设 W 是 V 中包含 S 的任一子空间, 则对任意的

$$\alpha = \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m \in L(S).$$

由于 $\alpha_1, \alpha_2, \dots, \alpha_m \in S \subset W$, 所以 $\alpha \in W$, 从而 $L(S) \subset W$. 因此 $L(S)$ 是 V 中包含 S 的最小子空间.

由于 V 的非空子集 S 的线性扩张 $L(S)$ 是 V 的一个子空间, 所以也称 $L(S)$ 是由 S “张成”的子空间(或生成的子空间). 如果 S 是有限子集 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$, 就称 $L(\alpha_1, \alpha_2, \dots, \alpha_m)$ 是由向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ “张成”的子空间.

以上讨论了由一部分元素经过一定的运算“扩张”成对此运算封闭的一个集合(代数结构). 但更重要的是相反的问题: 即对一个代数结构是否能找到在所给运算下“扩张”成这个结构的一组元素, 当然这种元素越少越好. 例如中国古人长期以来就认为万物都由“金、木、水、火、土”所构成(当然古人不会有“线性”的概念). 又如取各种颜料为元素组成的集合 A , 其运算为将颜料(元素)按一定比例混合, 由于用红、黄、蓝三种颜料可以配成各种颜色的颜料, 所以红、黄、蓝三种颜料可以“张成” A . 下面对这个问题给出一般的数学描述, 先给出有限维线性空间的定义.

定义 2.5 $V(F)$ 称为有限维线性空间, 如果 V 中存在一个有限子集 S , 使得 $L(S) = V$; 否则, 称为无穷维线性空间.

例 5 \mathbf{R}^3 是一个实数域上的有限维线性空间, 因为 \mathbf{R}^3 中存在有限子集 $S = \{e_1, e_2, e_3\}$, 其中 $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, 使得 $L(S) = \mathbf{R}^3$. 事实上, 显然有 $L(S) \subset \mathbf{R}^3$; 又 \mathbf{R}^3 中任一向量

$$\alpha = (a_1, a_2, a_3) = a_1 e_1 + a_2 e_2 + a_3 e_3 \in L(S),$$

所以 $\mathbf{R}^3 \subset L(S)$, 因此 $L(S) = \mathbf{R}^3$.

例 6 $\mathbf{R}[x]_3$ 也是一个实数域上的有限维线性空间. 在节 2.1 例 5 中已给出它的一个有限子集 $S_1 = \{p_1(x), p_2(x), p_3(x)\}$, 使得 $L(S_1) = \mathbf{R}[x]_3$. 此外, $\mathbf{R}[x]_3$ 有更简单的有限子集 $S = \{1, x, x^2\}$, 也使得 $L(S) = \mathbf{R}[x]_3$.

同理, $\mathbf{R}[x]_n$ 也是实数域上的一个有限维线性空间. 而全体实系数多项式构成的线性空间 $\mathbf{R}[x]$ 则是实数域上的一个无穷维线性空间, 因为它不存在一个有限子集 S , 使得 $L(S) = \mathbf{R}[x]$.

我们说一个线性空间是有限维的, 是指存在有限个元素能张成这个空间, 但更重要的是这有限个元素“至少”要多少个. 具体地说, 应该如何区别 \mathbf{R}^3 和 \mathbf{R}^n 这两个有限维线性空间? 这就要回答下面的问题: 如果 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 使得 $L(B) = W$, 那么在什么条件下可在 B 中去掉一部分元素, 使余下的

元素仍可张成 W 呢? 为了解决这个问题, 必须先引入线性空间中向量的线性相关性的概念.

2.3 线性相关性

线性空间中向量的线性相关性是向量在线性运算下的一种性质, 它是阐明线性空间理论的一个重要的基本概念. 对它下定义之前, 先看一下它在 \mathbb{R}^3 中的几何背景.

若三个非零向量 $\alpha_1, \alpha_2, \alpha_3$ 共面(起点皆在原点, 如图 2-1, 2-2)则至少有一个向量可由另两个向量线性表示(例如 α_3 可由 α_1, α_2 线性表示)这等价于: 存在不全为零的数 $\lambda_1, \lambda_2, \lambda_3$ 使 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3 = \mathbf{0}$ 成立. 这时如果 $W = L(\alpha_1, \alpha_2, \alpha_3)$ 则 W 中任一向量也可由 α_1, α_2 线性表示, 从而 α_1, α_2 也可张成 W , 即 $W = L(\alpha_1, \alpha_2)$. 若 $\alpha_1, \alpha_2, \alpha_3$ 不共面(如 \mathbb{R}^3 中基本向量 e_1, e_2, e_3)则任一向量都不能由另两个向量线性表示, 即只有 $\lambda_1 = \lambda_2 = \lambda_3 = 0$ 才使 $\lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3 = \mathbf{0}$. 这表明 $\alpha_1, \alpha_2, \alpha_3$ 中缺少任一个都不能张成 $W = L(\alpha_1, \alpha_2, \alpha_3)$.

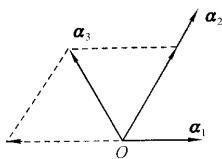


图 2-1

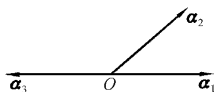


图 2-2

上述 \mathbb{R}^3 中向量在线性运算下的性质, 就是向量的线性相关性, 它的一般定义如下:

定义 2.6 设 $V(F)$ 是一个线性空间, $\alpha_1, \alpha_2, \dots, \alpha_m \in V$, 如果存在不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_m \in F$ 使

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_m\alpha_m = \mathbf{0} \quad (2-5)$$

成立, 则称 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关, 否则称为线性无关.

定义中的“否则”是指: 不线性相关就线性无关, 即“没有不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_m \in F$ 使得 (2-5) 式成立”, 也就是“对任何不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_m \in F$ (2-5) 式都不成立”, 这等价于“如果 (2-5) 式成立, 则 $\lambda_1, \lambda_2, \dots,$

λ_m 必须全为零”即“仅当 $\lambda_1, \lambda_2, \dots, \lambda_m$ 全为零时,才使(2-5)式成立”.

所谓向量的线性相关性是指向量线性相关和线性无关的性质.

向量 $\alpha_1, \alpha_2, \dots, \alpha_m$ ($m \geq 2$) 线性相关的等价定义是:若 $\alpha_1, \alpha_2, \dots, \alpha_m$ 中有一个向量可由其余向量在域 F 上线性表示,则称 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关. 所谓两个定义等价,就是可以互相推出.现在把等价定义写成定理.

定理 2.3 $V(F)$ 中的向量组 $\alpha_1, \alpha_2, \dots, \alpha_m$ ($m \geq 2$) 线性相关的充分必要条件是 $\alpha_1, \alpha_2, \dots, \alpha_m$ 中有一个向量可由其余向量在域 F 上线性表示.

证 设 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关,则存在不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_m \in F$, 使得

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m = 0.$$

为表述简便,不妨设 $\lambda_1 \neq 0$, 于是

$$\alpha_1 = -\lambda_1^{-1} \lambda_2 \alpha_2 - \dots - \lambda_1^{-1} \lambda_m \alpha_m,$$

必要性得证.反之,若 $\alpha_1, \alpha_2, \dots, \alpha_m$ 中的一个向量

$$\alpha_j = \mu_1 \alpha_1 + \dots + \mu_{j-1} \alpha_{j-1} + \mu_{j+1} \alpha_{j+1} + \dots + \mu_m \alpha_m,$$

则

$$\mu_1 \alpha_1 + \dots + \mu_{j-1} \alpha_{j-1} - \alpha_j + \mu_{j+1} \alpha_{j+1} + \dots + \mu_m \alpha_m = 0,$$

其中 $\mu_1, \dots, \mu_{j-1}, -1, \mu_{j+1}, \dots, \mu_m$ 不全为零,所以 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关,充分性得证.

定理 2.3 的等价命题 $\alpha_1, \alpha_2, \dots, \alpha_m$ ($m \geq 2$) 线性无关的充分必要条件是其中任一个向量都不能由其余向量线性表示.

例 1 \mathbf{R}^n 中的 e_1, e_2, \dots, e_n 是线性无关的,其中 $e_i = (0 \dots 0 \underset{i}{1} 0 \dots 0)$ 是第 i 个分量为 1,其余分量全为零的向量.因为,由

$$\lambda_1 e_1 + \lambda_2 e_2 + \dots + \lambda_n e_n = 0$$

即

$$(\lambda_1, \lambda_2, \dots, \lambda_n) = (0, 0, \dots, 0),$$

必有 $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

例 2 线性空间中单个向量 α 线性相关的充分必要条件是 α 为零向量. 因为 $\lambda \neq 0$ 时,等式 $\lambda \alpha = 0$ 成立的充要条件是 $\alpha = 0$.

例 3 含零向量的任何向量组 $\{0, \alpha_1, \alpha_2, \dots, \alpha_m\}$ 都线性相关. 因为 $\exists \lambda \neq 0$ 使

$$\lambda 0 + 0 \alpha_1 + 0 \alpha_2 + \dots + 0 \alpha_m = 0.$$

例 4 判别 \mathbf{R}^3 中向量组 $\{\alpha_1, \alpha_2, \alpha_3\}$ 和 $\{\beta_1, \beta_2, \beta_3\}$ 的线性相关性,其中

$$\alpha_1 = (1, 1, 0), \quad \alpha_2 = (0, 1, 1), \quad \alpha_3 = (1, 0, -1),$$

$$\beta_1 = (1, -3, 1), \quad \beta_2 = (-1, 2, -2), \quad \beta_3 = (1, 1, 3).$$

解 容易看出 $\alpha_3 = \alpha_1 - \alpha_2$, 所以 $\alpha_1, \alpha_2, \alpha_3$ 线性相关. 但对 $\beta_1, \beta_2, \beta_3$ 不易看出是否有线性关系, 要按定义 2.6 来判别. 设

$$x_1\beta_1 + x_2\beta_2 + x_3\beta_3 = \mathbf{0}, \quad (1)$$

即

$$x_1(1, -3, 1) + x_2(-1, 2, -2) + x_3(1, 1, 3) = (0, 0, 0).$$

这个向量方程等价于下面的三元线性齐次方程组:

$$\begin{cases} x_1 - x_2 + x_3 = 0 \\ -3x_1 + 2x_2 + x_3 = 0 \\ x_1 - 2x_2 + 3x_3 = 0, \end{cases}$$

容易解得这个方程组只有零解 $x_1 = x_2 = x_3 = 0$. 即只有全为零的 x_1, x_2, x_3 才使 (1) 成立, 故 $\beta_1, \beta_2, \beta_3$ 线性无关.

一般若 $\beta_1 = (a_1, b_1, c_1), \beta_2 = (a_2, b_2, c_2), \beta_3 = (a_3, b_3, c_3)$ 则 $\beta_1, \beta_2, \beta_3$ 线性相关(线性无关)的充要条件是三元线性齐次方程组

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 = 0 \\ c_1x_1 + c_2x_2 + c_3x_3 = 0 \end{cases}$$

有非零解(只有零解). 用此法可得 \mathbf{R}^3 中任何 4 个向量, \mathbf{R}^n 中任何 $n+1$ 个向量都线性相关.

例 5 节 2.1 例 5 中的 $p_1(x) = 1 + x, p_2(x) = 1 - x, p_3(x) = x + x^2$ 是线性无关的. 因为设

$$\lambda_1(1 + x) + \lambda_2(1 - x) + \lambda_3(x + x^2) = 0 \quad (\text{零多项式}), \quad (2)$$

即

$$(\lambda_1 + \lambda_2) + (\lambda_1 - \lambda_2 + \lambda_3)x + \lambda_3x^2 = 0.$$

于是

$$\lambda_1 + \lambda_2 = 0; \quad \lambda_1 - \lambda_2 + \lambda_3 = 0; \quad \lambda_3 = 0,$$

从而得 $\lambda_1 = \lambda_2 = \lambda_3 = 0$ 故 $p_1(x), p_2(x), p_3(x)$ 线性无关.

例 6 如果向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性无关, 则其任一子集也线性无关; 如果向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关, 则任何包含它的向量组也线性相关.

我们只证明前者, 不失一般性, 设子集为 $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$, 如果

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \dots + \lambda_k\alpha_k = \mathbf{0},$$

则必有 $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ 否则有不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_k$ 使

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_k \alpha_k + 0 \alpha_{k+1} + 0 \alpha_{k+2} + \dots + 0 \alpha_n = 0$$

成立,这与假设矛盾),所以 $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 线性无关.

定理 2.4 若向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性无关,而向量组 $\{\beta, \alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关,则 β 可由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表示,且表示法唯一.

证 由于向量组 $\{\beta, \alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关,所以存在不全为零的数量 $\lambda, \lambda_1, \lambda_2, \dots, \lambda_n$ 使得

$$\lambda \beta + \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_n \alpha_n = 0,$$

其中 λ 必不等于零(如果 $\lambda = 0$,则由 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性无关又得 $\lambda_1, \lambda_2, \dots, \lambda_n$ 必全为零,与题设矛盾),于是

$$\beta = -\lambda^{-1} \lambda_1 \alpha_1 - \lambda^{-1} \lambda_2 \alpha_2 - \dots - \lambda^{-1} \lambda_n \alpha_n.$$

再证表示法唯一,设有两种表示法:

$$\beta = b_1 \alpha_1 + b_2 \alpha_2 + \dots + b_n \alpha_n,$$

$$\beta = c_1 \alpha_1 + c_2 \alpha_2 + \dots + c_n \alpha_n.$$

于是

$$(b_1 - c_1) \alpha_1 + (b_2 - c_2) \alpha_2 + \dots + (b_n - c_n) \alpha_n = 0$$

而 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性无关,所以 $b_i - c_i = 0$, 即 $b_i = c_i$ ($i = 1, 2, \dots, n$), 故 β 由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 的表示法唯一.

因为 \mathbf{R}^n 中任何 $n+1$ 个向量都线性相关,故有

推论 如果 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 \mathbf{R}^n 中线性无关的 n 个向量,则 \mathbf{R}^n 中任一向量 α 可由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表示,且表示法唯一.

定理 2.5 设 $V(F)$ 中向量组 $\{\beta_1, \beta_2, \dots, \beta_s\}$ 的每个向量可由另一向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_r\}$ 线性表示.如果 $s > r$, 则 $\{\beta_1, \beta_2, \dots, \beta_s\}$ 线性相关.

这个定理在 \mathbf{R}^3 中的几何背景是:如果 α_1, α_2 线性无关, $\beta_1, \beta_2, \beta_3$ 可由 α_1, α_2 线性表示,则 $\beta_1, \beta_2, \beta_3$ 都位于 α_1, α_2 所确定的平面上,故 $\beta_1, \beta_2, \beta_3$ 线性相关.下面给出一般的证明:

证 设 $\beta_j = \sum_{i=1}^r \lambda_{ij} \alpha_i$ (其中 $\lambda_{ij} \in F$, $j = 1, 2, \dots, s$), 又设

$$x_1 \beta_1 + x_2 \beta_2 + \dots + x_s \beta_s = 0,$$

①

即

$$\sum_{j=1}^s x_j \beta_j = \sum_{j=1}^s x_j \left(\sum_{i=1}^r \lambda_{ij} \alpha_i \right) = \sum_{i=1}^r \left(\sum_{j=1}^s \lambda_{ij} x_j \right) \alpha_i = 0$$

(第二个等号成立见本章附录),令上式中 α_i ($i = 1, 2, \dots, r$) 的系数全为零, 即

$$\sum_{j=1}^s \lambda_{ij} x_j = 0, \quad i = 1, \dots, r, \quad (2)$$

② 式是关于 x_1, x_2, \dots, x_s 的齐次线性方程组, 由于其方程个数 r 小于未知量个数 s , 因此方程组 ② 必有非零解, 从而有不全为零的 x_1, x_2, \dots, x_s 使 ① 式成立, 故 $\{\beta_1, \beta_2, \dots, \beta_s\}$ 线性相关.

定理 2.5 的等价命题 如定理 2.5 所设, 若 $\{\beta_1, \beta_2, \dots, \beta_s\}$ 线性无关, 则 $s \leq r$.

定理 2.4 与定理 2.5 是了解清楚线性空间结构的重要定理, 以后常要用到它们.

2.4 有限维线性空间的基和维数 向量组的秩

如果线性空间 $V(F)$ 的线性无关的两个子集 $B_1 = \{\alpha_1, \dots, \alpha_n\}$ 和 $B_2 = \{\beta_1, \dots, \beta_m\}$ 都能扩张成 $V(F)$ 则由

$$\beta_j \in V(F) = L(\alpha_1, \dots, \alpha_n), \quad j = 1, \dots, m$$

以及定理 2.5 的等价命题, 即得 $m \leq n$; 再由

$$\alpha_i \in V(F) = L(\beta_1, \dots, \beta_m), \quad i = 1, \dots, n$$

又得 $n \leq m$, 从而就有 $m = n$. 这就是说, 线性空间 $V(F)$ 如能被其两组线性无关的有限子集所张成, 则它们所含向量个数相同. 如此, 我们可以给出下面的定义.

定义 2.7 如果线性空间 $V(F)$ 的有限子集 $B = \{\alpha_1, \dots, \alpha_n\}$ 线性无关, 且 $L(B) = V$, 则称 B 为 V 的一组基, 并称 n 为 V 的维数(或说 V 是 n 维线性空间), 记作 $\dim V = n$.

如果 $L(S) = V$, 且 S 是 V 的有限子集, 则 S 中最大的线性无关向量的个数就是能张成 V 的最少的向量个数, 也就是 V 的维数.

例 $\mathbb{R}[x]$ (\mathbb{R}^n) 的维数是 $3(n)$, 所以称它为 3 维(n 维)向量空间, 其中的向量也称为 3 维(n 维)向量. 它的基 $\{e_1, e_2, e_3\}$ ($\{e_1, e_2, \dots, e_n\}$) 叫做自然基.

$\mathbb{R}[x]$ 是 3 维线性空间, $B_1 = \{1, x, x^2\}$ 和 $B_2 = \{1+x, 1-x, x+x^2\}$ 都是 $\mathbb{R}[x]$ 的基(见节 2.2 例 6). $\mathbb{R}[x]_n$ 是 n 维线性空间, $B = \{1, x, x^2, \dots, x^{n-1}\}$ 是它的一组基, 也叫自然基.

节 2.1 例 2 的解空间 S 是二维的, $\{\alpha, \beta\}$ 是它的一组基, 它是 \mathbb{R}^5 的一个二

维子空间.

线性空间 V 的零子空间 $\{0\}$ 的维数为零, 因为其中没有线性无关的向量.

在 n 维线性空间 $V = L(B)$ 中, 任一向量都可唯一地表示为基 $B = \{\alpha_1, \dots, \alpha_n\}$ 的线性组合. 由定理 2.5 可知, 在 n 维线性空间 V 中, 任何 $n+1$ 个向量 $\eta_1, \eta_2, \dots, \eta_{n+1}$ 都是线性相关的(因为它们都可由基 B 线性表示). 于是由定理 2.4 又可知, n 维线性空间 V 中任何 n 个线性无关的向量组成的子集 $B^* = \{\beta_1, \dots, \beta_n\}$ 都是 V 的基(因为 $\forall \alpha \in V(F), \{\alpha, \beta_1, \dots, \beta_n\}$ 是线性相关的, 因此 α 可由 B^* 线性表示)故 $V = L(B^*)$. 这表明, 有限维线性空间的基并不唯一, 但任一组基所含向量的个数是唯一确定的.

下面讨论 V 的子空间的基与 V 的基的关系.

定理 2.6 如果 W 是 n 维线性空间 V 的一个子空间, 则 W 的基可以扩充为 V 的基(即 W 的基可添加 V 中若干向量成为 V 的基).

证 设 W 的基 $B_1 = \{\alpha_1, \dots, \alpha_m\}$, 如果 $m = n$, B_1 就是 V 的基. 如果 $m < n$, 则必存在 $\alpha_{m+1} \in V$ 使 $\{\alpha_1, \dots, \alpha_m, \alpha_{m+1}\}$ 线性无关, 不然的话, $\dim V < n$, 与假设矛盾. 如果 $m+1 = n$, 定理已得证. 如果 $m+1 < n$, 继续上述步骤, 必存在 $\alpha_{m+2}, \dots, \alpha_n \in V$, 使 $\{\alpha_1, \dots, \alpha_m, \alpha_{m+1}, \dots, \alpha_n\}$ 线性无关, 这就是 V 的基.

如果 n 维线性空间 $V(F)$ 的有限子集 $S = \{\alpha_1, \dots, \alpha_n\}$ 中的 n 个向量线性相关, 那么 V 的子空间 $L(S) \neq V$, 为了确定它的维数和基, 我们先给出下面的定义.

定义 2.8 设 S 是线性空间 $V(F)$ 的一个子集, 如果 S 中存在线性无关的向量组 $B = \{\alpha_1, \dots, \alpha_r\}$, 且 S 中每个向量可由 B 线性表示, 则 B 中向量的个数 r 叫做 S 的秩, 记作 $\text{秩}(S) = r$.

如果 S 是有限维线性空间 $V(F)$ 的子空间, 那么 S 的秩就是 S 的维数.

由 $\text{秩}(S)$ 与 $L(S)$ 的定义以及定理 2.5 可得以下结论:

(1) $\text{秩}(S) = r$, 则 S 中任何 $r+1$ 个向量都线性相关. 因此 S 中任何线性无关的向量组至多含 r 个向量, 并把含 r 个线性无关向量的向量组称为 S 的极大线性无关组.

(2) 如果 $\text{秩}(S) = r$, $B = \{\alpha_1, \dots, \alpha_r\}$ 是 S 的极大线性无关组, 则 $L(S) = L(B)$, 即 $\dim L(S) = \text{秩}(S)$.

(3) 如果 S, T 是 $V(F)$ 的两个有限子集, 且 S 中每个向量可由 T 线性表示, 则 $\text{秩}(S) \leq \text{秩}(T)$.

(4) 如果 S, T 是 $V(F)$ 的两个有限子集, 则 $L(S) = L(T)$ 的充要条件

是 S 中每个元素可由 T 线性表示, 且 T 中每个元素也可由 S 线性表示(此时也称 S 与 T 是等价向量组).

下面证明结论(3)(其它结论的证明留给读者去完成).

证 设 $S = \{\alpha_1, \dots, \alpha_s\}, T = \{\beta_1, \dots, \beta_t\}$ 秩 $(S) = p$ 秩 $(T) = r$; 为简便起见, 不妨设 S 和 T 的极大线性无关组分别为 $B_S = \{\alpha_1, \dots, \alpha_p\}$ 和 $B_T = \{\beta_1, \dots, \beta_r\}$. 由于 S 中每个向量可由 T 线性表示, 所以 S 中每个向量也可由 B_T 线性表示, 从而 B_S 中每个向量可由 B_T 线性表示. 因此, 根据定理 2.5 的等价命题即得 $p \leq r$.

例 已知 \mathbb{R}^4 的一个子集 $S = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ 其中

$$\alpha_1 = (1, 1, 0, 1), \quad \alpha_2 = (0, 1, 2, 4),$$

$$\alpha_3 = (2, 1, -2, -2), \quad \alpha_4 = (0, 1, 1, 1).$$

试求 $L(S)$ 的维数及其一组基 B .

解 S 的任一个极大线性无关组 B 都是 $L(S)$ 的基. 求 S 的一个极大线性无关组可用如下的方法: 先考察 α_1, α_2 , 由于 $\alpha_1 \neq 0, \alpha_2 \neq \lambda\alpha_1$, 所以 α_1, α_2 线性无关; 再考察 $\alpha_1, \alpha_2, \alpha_3$, 如果方程 $x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = 0$ 只有零解, 则 $\alpha_1, \alpha_2, \alpha_3$ 线性无关(然后再考察 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$), 如果方程有非零解, 则 $\alpha_1, \alpha_2, \alpha_3$ 线性相关, 去掉 α_3 , 再考察 $\alpha_1, \alpha_2, \alpha_4$, 如此继续下去, 即可找到 S 的一个极大线性无关组. 但当 S 所含的向量较多时, 用此法求其极大线性无关组比较繁.

下面再介绍一种通用而简便的方法. 设

$$x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 + x_4\alpha_4 = 0, \quad (1)$$

得方程组

$$\begin{cases} x_1 + 0x_2 + 2x_3 + 0x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0 \\ 0x_1 + 2x_2 - 2x_3 + x_4 = 0 \\ x_1 + 4x_2 - 2x_3 + x_4 = 0, \end{cases} \quad (2)$$

其增广矩阵为

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -2 & 1 & 0 \\ 1 & 4 & -2 & 1 & 0 \end{array} \right). \quad (3)$$

对矩阵(3)作初等行变换, 所得同解方程组的增广矩阵为

$$\left(\begin{array}{cccc|c} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right). \quad (4)$$

(4)所对应的齐次线性方程组显然有非零解,所以 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 线性相关.由于(3)中的系数矩阵的4个列向量分别是 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$,因此,如果去掉 α_3 ,考察 $\alpha_1, \alpha_2, \alpha_4$ 的线性相关性,设

$$x_1\alpha_1 + x_2\alpha_2 + x_4\alpha_4 = \mathbf{0}, \quad (5)$$

则由(5)所得的方程组的增广矩阵就是把(3)中第3列去掉的矩阵.用高斯消元法解方程组(5)得到的同解方程组对应的增广矩阵就是在矩阵(4)中去掉第3列,即

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right). \quad (6)$$

(6)所对应的方程组只有零解 $x_1 = x_2 = x_4 = 0$,故 $\alpha_1, \alpha_2, \alpha_4$ 线性无关,它就是 S 的一个极大线性无关组.于是 $\dim L(S) = 3$, $L(S)$ 的一组基为 $\{\alpha_1, \alpha_2, \alpha_4\}$.

上述求 S 的一个极大线性无关组的方法是:把增广矩阵(3)用高斯消元法(即对矩阵作初等行变换)化为增广矩阵(4)时,(4)中每一行第一个非零元1所在的列(第1,2,4列)所对应的列向量(即 $\alpha_1, \alpha_2, \alpha_4$)就是 S 的一个极大线性无关组.

2.5 向量的坐标

第一章中,已介绍过空间几何向量关于基本向量组 $\{e_1, e_2, e_3\}$ 的坐标,这里对 n 维线性空间引进向量的坐标的概念.

定义 2.9 设 $B = \{\beta_1, \beta_2, \dots, \beta_n\}$ 是 n 维线性空间 $V(F)$ 的一组基,如果 V 中元素 α 表示为

$$\alpha = a_1\beta_1 + a_2\beta_2 + \dots + a_n\beta_n,$$

则其系数组 a_1, a_2, \dots, a_n 叫做 α 在基 B 下的坐标,记作 $\alpha_B = (a_1, a_2, \dots, a_n)$.

由定义可见, V 中元素的坐标是由所选的基决定的,同一元素在不同的基下一般有不同的坐标;又基 B 的向量组是有序的,因此坐标 α_B 是有序数组,

它是 F^n 中的一个向量,也称为 α 关于基 B 的坐标向量.

由定理 2.4 可知,有限维线性空间 $V(F)$ 中每一个元素在给定的基 B 下的坐标是唯一确定的,因此,给定了 $V(F)$ 的一个基 B , V 中元素与 F^n 中的向量一一对应,而且保持元素间的线性运算关系不变,即如果

$$\alpha_B = (a_1 \ a_2 \ \dots \ a_n), \quad \beta_B = (b_1 \ b_2 \ \dots \ b_n),$$

则 $(\alpha + \beta)_B = (a_1 + b_1 \ a_2 + b_2 \ \dots \ a_n + b_n) = \alpha_B + \beta_B$,

$$(\lambda\alpha)_B = (\lambda a_1 \ \lambda a_2 \ \dots \ \lambda a_n) = \lambda\alpha_B.$$

由此可见,研究任何 n 维线性空间 $V(F)$, 都可以通过基和坐标,归结为研究 n 维向量空间 F^n . 这样我们对各种各样的 n 维线性空间就有了统一的研究方法,统一到研究 F^n . 如果把 F^n 的元素 $(a_1 \ a_2 \ \dots \ a_n)$ 看成是 n 维几何向量(即以 $(e_1 \ e_2 \ \dots \ e_n)$ 为基的向量的坐标),那么 F^n 又可以看成通常的几何向量空间 \mathbf{R}^n . 正因为如此,我们有时把线性空间和向量空间,线性空间中的元素和向量不加区别也就是合理的了.

例 1 \mathbf{R}^n 中的向量 $\alpha = (a_1 \ a_2 \ \dots \ a_n)$ 在自然基 $B = (e_1 \ e_2 \ \dots \ e_n)$ 下的坐标也是 $(a_1 \ a_2 \ \dots \ a_n)$; $\mathbf{R}[x]_3$ 中的 $p(x) = a_0 + a_1x + a_2x^2$ 在自然基 $B = (1 \ x \ x^2)$ 下的坐标 $(p(x))_B = (a_0 \ a_1 \ a_2)$.

例 2 \mathbf{R}^n 有一组基 $B = \{\beta_1 \ \beta_2 \ \dots \ \beta_n\}$ 其中

$$\beta_1 = (1 \ 1 \ \dots \ 1) \ \beta_2 = (0 \ 1 \ \dots \ 1) \ \dots \ \beta_n = (0 \ 0 \ \dots \ 1),$$

试求 $\alpha = (a_1 \ a_2 \ \dots \ a_n) = a_1e_1 + a_2e_2 + \dots + a_ne_n$ 在基 B 下的坐标 α_B .

解 设 $\alpha_B = (x_1 \ x_2 \ \dots \ x_n)$, 即

$$\alpha = x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n,$$

于是

$$\begin{cases} x_1 = a_1 \\ x_1 + x_2 = a_2 \\ \dots\dots\dots \\ x_1 + x_2 + \dots + x_n = a_n. \end{cases}$$

解这个方程组,得 $x_1 = a_1 \ x_2 = a_2 - a_1 \ \dots \ x_n = a_n - a_{n-1}$.

例 3 若 $\mathbf{R}[x]_3$ 的基 $B = (1 \ x - 1 \ (x - 1)^2)$ 则容易证明: $p(x) = a_0 + a_1x + a_2x^2$ 在基 B 下的坐标为 $(p(x))_B = (a_0 + a_1 + a_2 \ a_1 + 2a_2 \ a_2)$.

在 n 维线性空间 $V(F)$ 中,同一个向量在不同基下的坐标之间的关系,我们将在第 4 章用矩阵的工具给出简明的结论.

2.6* 子空间的交与和 直和

定义 2.10 设 W_1, W_2 是线性空间 $V(F)$ 的两个子空间, 则

$$W_1 \cap W_2 = \{\alpha \mid \alpha \in W_1 \text{ 且 } \alpha \in W_2\},$$

$$W_1 + W_2 = \{\alpha \mid \alpha = \alpha_1 + \alpha_2, \alpha_1 \in W_1, \alpha_2 \in W_2\}$$

分别称为 W_1, W_2 的交与和.

需要注意, W_1, W_2 的和与 W_1, W_2 的并是不同的概念. W_1, W_2 的交与和仍是 V 的子空间. 下面证明后者.

设 $\lambda \in F, \alpha, \beta \in W_1 + W_2$, 即存在 $\alpha_1, \beta_1 \in W_1, \alpha_2, \beta_2 \in W_2$, 使得

$$\alpha = \alpha_1 + \alpha_2, \quad \beta = \beta_1 + \beta_2.$$

于是

$$\alpha + \beta = (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) = (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) \in W_1 + W_2,$$

$$\lambda \alpha = \lambda(\alpha_1 + \alpha_2) = \lambda \alpha_1 + \lambda \alpha_2 \in W_1 + W_2,$$

故 $W_1 + W_2$ 是 $V(F)$ 的一个子空间.

两个子空间的交与和也可推广到多个子空间, 而且 V 的子空间 W_1, W_2, \dots, W_m 的交

$$\bigcap_{i=1}^m W_i = \{\alpha \mid \alpha \in W_i, i = 1, 2, \dots, m\}$$

与和

$$\sum_{i=1}^m W_i = \{\alpha_1 + \alpha_2 + \dots + \alpha_m \mid \alpha_i \in W_i, i = 1, 2, \dots, m\}$$

也都是 V 的子空间.

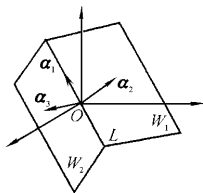


图 2-3

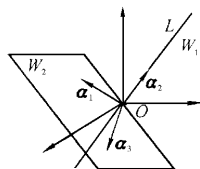


图 2-4

例1 设 $W_1 = L(\alpha_1, \alpha_2), W_2 = L(\alpha_1, \alpha_3)$ 是 \mathbb{R}^3 中两个不同的子空间 (如图 2-3 所示), 则 $W_1 \cap W_2 = L(\alpha_1)$ 是两个平面交线 L 上全体向量构成

的一维子空间; $W_1 + W_2 = L(\alpha_1, \alpha_2, \alpha_3) = \mathbf{R}^3$, 因为 $\{\alpha_1, \alpha_2, \alpha_3\}$ 是线性无关的(如果 $\{\alpha_1, \alpha_2, \alpha_3\}$ 线性相关, 则 α_3 可由 α_1, α_2 线性表示, 从而 $W_1 = W_2$ 与题设矛盾).

例2 设 W_1 与 W_2 分别是 \mathbf{R}^3 中过原点的直线和平面(直线不在平面上)上的全体向量构成的子空间(如图 2-4), 则

$$W_1 \cap W_2 = \{\mathbf{0}\},$$

$$W_1 + W_2 = L(\alpha_1, \alpha_2, \alpha_3) = \mathbf{R}^3.$$

从以上两例我们可猜测: 子空间 $W_1, W_2, W_1 \cap W_2, W_1 + W_2$ 的维数满足以下关系式:

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2). \quad (2-6)$$

(2-6)式叫做子空间的维数公式. 下面给出它的证明, 设:

$$\dim W_1 = s, \dim W_2 = t, \dim(W_1 \cap W_2) = r,$$

$$W_1 \cap W_2 = L(\alpha_1, \alpha_2, \dots, \alpha_r),$$

$$W_1 = L(\alpha_1, \alpha_2, \dots, \alpha_r, \beta_1, \dots, \beta_{s-r}),$$

$$W_2 = L(\alpha_1, \alpha_2, \dots, \alpha_r, \gamma_1, \dots, \gamma_{t-r}),$$

于是

$$W_1 + W_2 = L(\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{s-r}, \gamma_1, \dots, \gamma_{t-r}).$$

如此, 只要证明 $\dim(W_1 + W_2) = s + t - r$, 即 $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{s-r}, \gamma_1, \dots, \gamma_{t-r}$ (共 $s + t - r$ 个向量)是线性无关的. 为此, 设

$$a_1 \alpha_1 + \dots + a_r \alpha_r + b_1 \beta_1 + \dots + b_{s-r} \beta_{s-r} + c_1 \gamma_1 + \dots + c_{t-r} \gamma_{t-r} = \mathbf{0}, \quad (1)$$

即

$$a_1 \alpha_1 + \dots + a_r \alpha_r + b_1 \beta_1 + \dots + b_{s-r} \beta_{s-r} = -c_1 \gamma_1 - \dots - c_{t-r} \gamma_{t-r}. \quad (2)$$

易见, (2)式两端的向量分别属于 W_1, W_2 , 所以它们都属于 $W_1 \cap W_2$, 于是

$$-c_1 \gamma_1 - \dots - c_{t-r} \gamma_{t-r} = d_1 \alpha_1 + \dots + d_r \alpha_r,$$

即

$$c_1 \gamma_1 + \dots + c_{t-r} \gamma_{t-r} + d_1 \alpha_1 + \dots + d_r \alpha_r = \mathbf{0}. \quad (3)$$

(3)式左端是 W_2 的基的线性组合, 所以 $c_1 = \dots = c_{t-r} = 0$, 代入(2), 再由 $\{\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_{s-r}\}$ 是 W_1 的基, 又得 $a_1 = \dots = a_r = b_1 = \dots = b_{s-r} = 0$, 所以(1)式中的向量组线性无关. (2-6)式得证.

定义 2.11 设 W_1, W_2 是 $V(F)$ 两个子空间, 如果 $W_1 \cap W_2 = \{\mathbf{0}\}$ 则 $W_1 + W_2$ 叫做 W_1 与 W_2 的直和, 记作 $W_1 \oplus W_2$.

例2中的 $W_1 + W_2$ 是直和, 由此例易见有下面的定理.

定理 2.7 对子空间 W_1, W_2 , 下列命题等价:

(1) $W_1 + W_2$ 是直和, 即 $W_1 \cap W_2 = \{0\}$;

(2) $W_1 + W_2$ 中的每个向量 α 的分解式 $\alpha = \alpha_1 + \alpha_2$ ($\alpha_1 \in W_1, \alpha_2 \in W_2$) 是唯一的;

(3) 零向量 0 的分解式 $0 = \alpha_1 + \alpha_2$ ($\alpha_1 \in W_1, \alpha_2 \in W_2$), 仅当 $\alpha_1 = \alpha_2 = 0$ 才成立;

(4) $\dim(W_1 + W_2) = \dim W_1 + \dim W_2$.

证 上述命题等价, 只需证明 (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1).

先证 (1) \Rightarrow (2). 设 $W_1 + W_2$ 中的 α 有两个分解式

$$\alpha = \alpha_1 + \alpha_2 = \beta_1 + \beta_2, \quad \alpha_1, \beta_1 \in W_1, \quad \alpha_2, \beta_2 \in W_2,$$

则 $\alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in W_1 \cap W_2 = \{0\}$, 于是得 $\beta_1 = \alpha_1, \beta_2 = \alpha_2$, 故 α 的分解式是唯一的.

其次证 (2) \Rightarrow (3): 由 $W_1 + W_2$ 中零向量 0 的分解式唯一性以及 $0 = 0 + 0$, 立即得命题 (3) 成立.

再证 (3) \Rightarrow (4): 由命题 (3) 可推出 $W_1 \cap W_2 = \{0\}$, 因为若 $W_1 \cap W_2 \neq \{0\}$ 则存在 $0 \neq \alpha \in W_1 \cap W_2$, 使得 $0 = \alpha + (-\alpha)$ (其中 $\alpha \in W_1, -\alpha \in W_2$), 这与命题 (3) 相矛盾. 再根据维数公式 (2-6) 就得命题 (4).

最后由命题 (4) 及维数公式 (2-6) 立即得命题 (1) 成立.

子空间的直和也可推广到多个子空间 W_1, W_2, \dots, W_m , 此时应定义为:

如果 W_1, W_2, \dots, W_m 中每个子空间与其余子空间的和的交为 $\{0\}$, 即

$$W_i \cap \sum_{j \neq i} W_j = \{0\}, \quad i = 1, 2, \dots, m,$$

则它们的和称为直和, 记作 $W_1 \oplus W_2 \oplus \dots \oplus W_m$.

如果 $W = W_1 \oplus W_2 \oplus \dots \oplus W_m$, 则 W 叫做 W_1, W_2, \dots, W_m 的直和, 此时 W 中任一向量 α 将唯一地分解为

$$\alpha = \alpha_1 + \alpha_2 + \dots + \alpha_m, \quad \alpha_i \in W_i, \quad i = 1, 2, \dots, m.$$

如果 $\alpha = 0$, 仅当 $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ 时, 上式才成立, 而且

$$\dim(W_1 \oplus W_2 \oplus \dots \oplus W_m) = \sum_{j=1}^m \dim W_j.$$

如果线性空间 $V = W_1 \oplus W_2$, 则 W_2 叫做 W_1 的补空间, W_1 叫做 W_2 的补空间, 或者说, W_1 与 W_2 是互补子空间.

有限维线性空间的一个子空间的补空间不是唯一的. 例如 \mathbb{R}^3 中, $W_1 = L(\alpha_1), W_2 = L(\alpha_2, \alpha_3)$, 只要 $\{\alpha_1, \alpha_2, \alpha_3\}$ 线性无关, 那么 W_1, W_2 就都是互补子空间.

例如

$$\mathbb{R}^3 = L(e_1, e_2) \oplus L(\alpha) \text{ 其中 } \alpha = (1, 1, 1),$$

$$\mathbb{R}^3 = L(e_1, e_2) + L(e_3, \alpha) \text{ 同上.}$$

例3 设 \mathbf{R}^4 的两个子空间 S_1 和 S_2 为:

$$S_1 = \{ (x_1, x_2, x_3, x_4) \mid x_1 + x_2 + x_3 + x_4 = 0 \},$$

$$S_2 = \{ (x_1, x_2, x_3, x_4) \mid x_1 - x_2 - x_3 + x_4 = 0, x_1 + x_2 + x_3 - x_4 = 0 \}.$$

试求 $S_1 + S_2$, $S_1 \cap S_2$ 和 S_2 的补空间的基.

解 先求 S_1 和 S_2 的基. S_1 是齐次方程 $x_1 + x_2 + x_3 + x_4 = 0$ 的解空间, 任取 $x_2 = k_1$, $x_3 = k_2$, $x_4 = k_3$, 解得 $x_1 = -k_1 - k_2 - k_3$, 于是

$$S_1 = \{ k_1 \alpha_1 + k_2 \alpha_2 + k_3 \alpha_3 \mid k_1, k_2, k_3 \in \mathbf{R} \} = L(\alpha_1, \alpha_2, \alpha_3),$$

其中 $\alpha_1 = \{-1, 1, 0, 0\}$, $\alpha_2 = \{-1, 0, 1, 0\}$, $\alpha_3 = \{-1, 0, 0, 1\}$, 易证 $\{\alpha_1, \alpha_2, \alpha_3\}$ 是 S_1 的一组基. S_2 是齐次线性方程组

$$\begin{cases} x_1 - x_2 - x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 0 \end{cases}$$

的解空间, 用高斯消元法容易求得

$$S_2 = \{ k_1 \beta_1 + k_2 \beta_2 \mid k_1, k_2 \in \mathbf{R} \} = L(\beta_1, \beta_2),$$

其中 $\beta_1 = (0, -1, 1, 0)$, $\beta_2 = (0, 1, 0, 1)$ 线性无关, 因此 $\{\beta_1, \beta_2\}$ 是 S_2 的基. 如此即得

$$S_1 + S_2 = L(\alpha_1, \alpha_2, \alpha_3) + L(\beta_1, \beta_2) = L(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2).$$

用节 2.4 例题的方法, 求得 $S_1 + S_2$ 的一组基为 $\{\alpha_1, \alpha_2, \alpha_3, \beta_2\}$. 由维数公式可知

$$\dim(S_1 \cap S_2) = \dim S_1 + \dim S_2 - \dim(S_1 + S_2) = 3 + 2 - 4 = 1.$$

再由 S_1 与 S_2 的假设可见, $S_1 \cap S_2$ 是齐次线性方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 - x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 - x_4 = 0 \end{cases}$$

的解空间. 于是用高斯消元法求得 $S_1 \cap S_2 = L(\alpha)$, 其中 $\alpha = (0, -1, 1, 0)$ 是 $S_1 \cap S_2$ 的基.

欲求 S_2 的补空间的基, 只需把 S_2 的基扩充为 \mathbf{R}^4 的基. 此时可利用习题 14 的结论. 由于 $\xi_1 = (0, 0, 0, 1)$ 不能由 β_1, β_2 线性表示, 所以 ξ_1, β_1, β_2 线性无关. 再根据

$$\lambda_1 \xi_1 + \lambda_2 \beta_1 + \lambda_3 \beta_2 = (0, \lambda_3 - \lambda_2, \lambda_2, \lambda_1 + \lambda_3),$$

则 $\xi_2 = (1, 0, 0, 0)$ 不能由 ξ_1, β_1, β_2 线性表示, 所以 $\{\beta_1, \beta_2, \xi_1, \xi_2\}$ 是 \mathbf{R}^4 的基, 其中 $\{\xi_1, \xi_2\}$ 便是 S_2 的一个补空间的一组基. 以后学了行列式, 利用定理 5.4 和节 5.2 中例 4 的结论, 也很容易确定 ξ_1, ξ_2 .

2.7 内积空间

线性空间中的线性运算不能描述向量的度量性质, 如向量的长度、夹角等. 而向量的度量性质在分析、几何问题中是不可缺少的, 所以我们必须引入

度量的概念.

在节 1.7 空间几何向量的运算中,讲过向量的长度、夹角都可由向量的内积表示,而且向量的内积满足 4 条运算规则.现在,我们在抽象的线性空间中先定义内积运算(要求它满足几何向量的内积的 4 条性质),再用内积定义线性空间中向量的长度和向量之间的夹角.定义了内积的线性空间就称为内积空间.这里我们仅讨论实数域上的内积空间.

定义 2.12 在实空间 $V(\mathbf{R})$ 上定义一个二元运算,使 V 中元素 α, β 与一个实数相对应,记作 (α, β) , 如果 $\forall \alpha, \beta \in V, \lambda \in \mathbf{R}$ 满足:

$$(1) (\alpha, \beta) = (\beta, \alpha);$$

$$(2) (\alpha + \beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma);$$

$$(3) (\lambda \alpha, \beta) = \lambda (\alpha, \beta);$$

$$(4) (\alpha, \alpha) \geq 0, \text{ 等号成立当且仅当 } \alpha = 0,$$

则称实数 (α, β) 为向量 α, β 的内积,定义了内积的 $V(\mathbf{R})$ 称为实内积空间,有限维实内积空间叫做欧氏空间(Euclid 空间).

这里附带说一下,在复空间 $V(\mathbf{C})$ 上的内积定义,只是把(1)改为

$$(\alpha, \beta) = \overline{(\beta, \alpha)},$$

即 (α, β) 与 (β, α) 互为共轭复数.复内积空间通常称为酉空间(Unitary space).

由定义可得,零向量与任何向量的内积等于零,即

$$(0, \beta) = (0\alpha, \beta) = 0(\alpha, \beta) = 0.$$

又

$$\left(\sum_{i=1}^m \lambda_i \alpha_i, \sum_{j=1}^n \mu_j \beta_j \right) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j (\alpha_i, \beta_j).$$

根据内积的第(4)条性质,我们可用内积定义向量的长度.

定义 2.13 实内积空间 $V(\mathbf{R})$ 中向量 α 的长度定义为

$$|\alpha| = \sqrt{(\alpha, \alpha)}. \quad (2-7)$$

例1 $\forall \alpha = (a_1, a_2, \dots, a_n), \beta = (b_1, b_2, \dots, b_n) \in \mathbf{R}^n$, 定义

$$(\alpha, \beta) = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

容易验证它是 \mathbf{R}^n 的一个内积,并称之为 \mathbf{R}^n 的标准内积.此时向量 α 的长度

$$|\alpha| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}.$$

在 \mathbf{R}^n 中定义内积的方法不是唯一的.

例2 $\forall \alpha = (a_1, a_2), \beta = (b_1, b_2) \in \mathbf{R}^2$, 定义

$$(\alpha, \beta) = a_1 b_1 - a_2 b_1 - a_1 b_2 + 3a_2 b_2.$$

它也是 \mathbf{R}^2 的一个内积, 我们验证它满足内积的第(4)条性质:

$$(\alpha, \alpha) = a_1 a_1 - 2a_1 a_2 + 3a_2 a_2 = (a_1 - a_2)^2 + 2a_2^2 \geq 0.$$

其等号成立当且仅当 $a_1 = a_2 = 0$, 即 $\alpha = 0$.

$$\text{此时向量 } \alpha \text{ 的长度 } |\alpha| = \sqrt{(a_1 - a_2)^2 + 2a_2^2}.$$

关于该内积, 对 $\alpha = (1, 0)$ 和 $\beta = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, 有 $|\alpha| = |\beta|$ 且 $(\alpha, \beta) = 0$.

例 3 在 $R[x]_3$ 中, 对任意的 $f(x) = a_0 + a_1 x + a_2 x^2$, $g(x) = b_0 + b_1 x + b_2 x^2$, 定义

$$(f(x), g(x)) = a_0 b_0 + a_1 b_1 + a_2 b_2.$$

容易验证它是 $R[x]_3$ 中的一个内积. 但在 $R[x]_n$ 空间中, 常用积分定义

$$(f(x), g(x)) = \int_a^b f(x)g(x)dx.$$

定理 2.8 设 $V(\mathbf{R})$ 是一个内积空间, 则 $\forall \alpha, \beta \in V$ 和 $\lambda \in \mathbf{R}$, 有

$$(1) |\lambda\alpha| = |\lambda| |\alpha|;$$

$$(2) |(\alpha, \beta)| \leq |\alpha| |\beta|; \quad (2-8)$$

$$(3) |\alpha + \beta| \leq |\alpha| + |\beta|,$$

其中(2)称为柯西-施瓦茨(Cauchy-Schwarz)不等式, (3)称为三角不等式.

证 (1) $|\lambda\alpha| = \sqrt{(\lambda\alpha, \lambda\alpha)} = \sqrt{\lambda^2(\alpha, \alpha)} = |\lambda| |\alpha|.$

(2) 当 $\beta = 0$ 时, 结论成立; 当 $\beta \neq 0$ 时, 令 $\gamma = \alpha + \lambda\beta$, 则

$$(\gamma, \gamma) = (\alpha + \lambda\beta, \alpha + \lambda\beta)$$

$$= (\alpha, \alpha) + 2(\alpha, \beta)\lambda + (\beta, \beta)\lambda^2 \geq 0.$$

由这个关于 λ 的二次三项式的非负性, 即得其判别式

$$4(\alpha, \beta)^2 - 4(\alpha, \alpha)(\beta, \beta) \leq 0,$$

即 $(\alpha, \beta)^2 \leq |\alpha|^2 |\beta|^2$, 从而 $|(\alpha, \beta)| \leq |\alpha| |\beta|.$

上式等号成立当且仅当 α, β 线性相关(证明留给读者完成).

(3) 由

$$\begin{aligned} |\alpha + \beta|^2 &= (\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) + 2(\alpha, \beta) + (\beta, \beta) \\ &\leq |\alpha|^2 + 2|\alpha| |\beta| + |\beta|^2, \end{aligned}$$

即得三角不等式(其中“ \leq ”成立是利用了(2)的结论).

有了 Cauchy-Schwarz 不等式, 就可用内积定义向量的夹角.

定义 2.14 对任意非零向量 $\alpha, \beta \in V(\mathbf{R})$, 定义其夹角

$$\langle \alpha, \beta \rangle = \arccos \frac{(\alpha, \beta)}{|\alpha| |\beta|}. \quad (2-9)$$

非零向量 α, β 正交 (即 $\alpha \perp \beta$), 当且仅当 $(\alpha, \beta) = 0$. 由于零向量与任何向量的内积等于零, 所以也说零向量与任何向量正交. 于是一般说, α, β 正交当且仅当 $(\alpha, \beta) = 0$.

如果 $\alpha \perp \beta$, 由定理 2.8 中三角不等式 (3) 的证明, 即得勾股定理

$$|\alpha + \beta|^2 = |\alpha|^2 + |\beta|^2. \quad (2-10)$$

例 在 \mathbf{R}^n 中, 对标准内积, 有 $e_i \perp e_j$, $i \neq j$, $i, j = 1, 2, \dots, n$.

\mathbf{R}^2 中, 对例 2 定义的内积, 与 $e_1 = (1, 0)$ 正交的向量为 (b, b) , 其中 b 为任意实数.

定理 2.9 在内积空间 $V(R)$ 中两两正交的非零向量组 $S = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ 是线性无关的.

证 该定理证明可用直接证法, 现用反证法: 设 S 线性相关, 于是 S 中有一个向量可由其余向量线性表示, 不妨设

$$\alpha_1 = \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m,$$

则当 $j \neq 1$, 即 $j = 2, 3, \dots, m$ 时, 均有

$$\begin{aligned} (\alpha_1, \alpha_j) &= (\lambda_2 \alpha_2 + \dots + \lambda_j \alpha_j + \dots + \lambda_m \alpha_m, \alpha_j) \\ &= \lambda_2 (\alpha_2, \alpha_j) + \dots + \lambda_j (\alpha_j, \alpha_j) + \dots + \lambda_m (\alpha_m, \alpha_j) \\ &= \lambda_j (\alpha_j, \alpha_j) = 0, \end{aligned}$$

而 $(\alpha_j, \alpha_j) > 0$, 故 $\lambda_j = 0$ ($j = 2, 3, \dots, m$), 从而 $\alpha_1 = 0$, 与定理的假设矛盾.

2.8 欧氏空间的单位正交基

由于欧氏空间中的向量具有几何度量性, 因此, 在欧氏空间中常用有某种度量性质的基.

定义 2.15 设 $B = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 是 n 维欧氏空间 $V(R)$ 的一个子集, 如果

$$(\varepsilon_i, \varepsilon_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i, j = 1, 2, \dots, n,$$

则称 B 为 V 的单位正交基 (或称标准正交基).

由定义易见, 所谓单位正交基就是每个基向量都是单位长, 而且基向量两两正交.

例如, 在 \mathbf{R}^n 中, 自然基 $B = \{e_1, e_2, \dots, e_n\}$ 关于标准内积是单位正交基.

但在抽象的欧氏空间中,有必要证明其单位正交基存在.

定理 2.10 欧氏空间 $V(\mathbf{R})$ 恒有单位正交基.

证 设 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 n 维欧氏空间 $V(\mathbf{R})$ 的一组基,我们采用下面的作法(通常称为 Schmidt 正交化过程),由 B 构造出 $V(\mathbf{R})$ 的一组单位正交基.令

$$\begin{aligned}\beta_1 &= \alpha_1, \\ \beta_2 &= \alpha_2 + \lambda_{12}\beta_1.\end{aligned}$$

由于 β_1, α_2 线性无关,所以 $\beta_2 \neq 0$,为使 β_2, β_1 正交,即

$$(\beta_2, \beta_1) = (\alpha_2 + \lambda_{12}\beta_1, \beta_1) = (\alpha_2, \beta_1) + \lambda_{12}(\beta_1, \beta_1) = 0,$$

可取

$$\lambda_{12} = -\frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}.$$

如果已求出两两正交的非零向量 $\beta_1, \beta_2, \dots, \beta_{m-1}$,再令

$$\beta_m = \alpha_m + \lambda_{m-1,m}\beta_{m-1} + \dots + \lambda_{2,m}\beta_2 + \lambda_{1,m}\beta_1. \quad (1)$$

为使 β_m 与 β_k ($k = 1, 2, \dots, m-1$) 正交,即

$$(\beta_m, \beta_k) = (\alpha_m, \beta_k) + \lambda_{km}(\beta_k, \beta_k) = 0,$$

可取

$$\lambda_{km} = -\frac{(\alpha_m, \beta_k)}{(\beta_k, \beta_k)}, \quad k = 1, 2, \dots, m-1. \quad (2)$$

由于 $\beta_k \in L(\alpha_1, \alpha_2, \dots, \alpha_k)$, $k = 1, 2, \dots, m-1$,所以 (1) 式中的 $\beta_m \neq 0$,否则 $\alpha_m \in L(\alpha_1, \alpha_2, \dots, \alpha_{m-1})$ 这与 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性无关矛盾.

按数学归纳法原理,用上述方法就可由基 B 构造出如 (1)、(2) 式所示的两两正交的非零向量组 $\beta_1, \beta_2, \dots, \beta_n$,即

$$\begin{aligned}\beta_1 &= \alpha_1, \\ \beta_2 &= \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1, \\ &\dots\dots\dots \\ \beta_n &= \alpha_n - \frac{(\alpha_n, \beta_{n-1})}{(\beta_{n-1}, \beta_{n-1})}\beta_{n-1} - \dots - \frac{(\alpha_n, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \frac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)}\beta_1.\end{aligned}$$

再将它们单位化,即

$$\varepsilon_m = \frac{\beta_m}{|\beta_m|}, \quad m = 1, 2, \dots, n.$$

如此即得到 $V(\mathbf{R})$ 的单位正交基 $B^* = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$.

例1 已知 $B = \{\alpha_1, \alpha_2, \alpha_3\}$ 是 \mathbf{R}^3 的一组基, 其中 $\alpha_1 = (1, -1, 0)$, $\alpha_2 = (1, 0, 1)$, $\alpha_3 = (1, -1, 1)$, 试用 Schmidt 正交化方法, 由 B 构造 \mathbf{R}^3 的一组单位正交基.

解 取

$$\beta_1 = \alpha_1 = (1, -1, 0),$$

$$\beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1 = (1, 0, 1) - \frac{(1, -1, 0)}{2} = \left(\frac{1}{2}, \frac{1}{2}, 1\right),$$

$$\begin{aligned}\beta_3 &= \alpha_3 - \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)}\beta_1 \\ &= (1, -1, 1) - \frac{2}{3}\left(\frac{1}{2}, \frac{1}{2}, 1\right) - \frac{2(1, -1, 0)}{2} \\ &= \left(-\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}\right).\end{aligned}$$

再将 $\beta_1, \beta_2, \beta_3$ 单位化, 得 \mathbf{R}^3 的一组单位正交基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$, 即

$$\varepsilon_1 = \frac{\beta_1}{|\beta_1|} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right),$$

$$\varepsilon_2 = \frac{\beta_2}{|\beta_2|} = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right),$$

$$\varepsilon_3 = \frac{\beta_3}{|\beta_3|} = \left(-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

例2 如果 $B = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 是 n 维欧氏空间 $V(\mathbf{R})$ 的一组单位正交基, α, β 关于基 B 的坐标分别为 (a_1, a_2, \dots, a_n) 和 (b_1, b_2, \dots, b_n) , 则

$$(1) a_i = (\alpha, \varepsilon_i), i = 1, 2, \dots, n;$$

$$(2) (\alpha, \beta) = a_1 b_1 + a_2 b_2 + \dots + a_n b_n.$$

解 由 $\alpha = a_1 \varepsilon_1 + \dots + a_i \varepsilon_i + \dots + a_n \varepsilon_n$, 即得

$$\begin{aligned}(1) (\alpha, \varepsilon_i) &= a_1 (\varepsilon_1, \varepsilon_i) + \dots + a_i (\varepsilon_i, \varepsilon_i) + \dots + a_n (\varepsilon_n, \varepsilon_i) \\ &= a_i (\varepsilon_i, \varepsilon_i) = a_i, \quad i = 1, 2, \dots, n.\end{aligned}$$

$$\begin{aligned}(2) (\alpha, \beta) &= ((a_1 \varepsilon_1 + a_2 \varepsilon_2 + \dots + a_n \varepsilon_n)(b_1 \varepsilon_1 + b_2 \varepsilon_2 + \dots + b_n \varepsilon_n)) \\ &= a_1 b_1 + a_2 b_2 + \dots + a_n b_n.\end{aligned}$$

2.9* 正交子空间 正交补

本节讨论欧氏空间中子空间的正交关系.

定义 2.16 设 $\alpha \in V(\mathbf{R})$, W 是 $V(\mathbf{R})$ 的一个子空间, 如果 $\forall \gamma \in W$ 均

有 $(\alpha, \gamma) = 0$ 则称 α 与 W 正交, 记作 $\alpha \perp W$.

定义 2.17 设 W_1, W_2 是 $V(\mathbf{R})$ 的两个子空间. 如果 $\forall \alpha \in W_1, \beta \in W_2$ 均有 $(\alpha, \beta) = 0$ 则称 W_1 与 W_2 互相正交, 记作 $W_1 \perp W_2$.

如果子空间 W_1, W_2 正交, 则 $W_1 + W_2$ 是直和. 这是因为 $\forall \alpha \in W_1 \cap W_2$, 均有 $(\alpha, \alpha) = 0$ 故 $\alpha = 0$, 从而 $W_1 \cap W_2 = \{0\}$.

同理, 若 W_1, W_2, \dots, W_m 两两正交, 则它们之和也是直和.

定义 2.18 设 W_1, W_2 是 $V(\mathbf{R})$ 的两个子空间, 如果 $W_1 \perp W_2$, 且 $W_1 + W_2 = V$, 则称 W_2 是 W_1 的正交补, 记作 W_1^\perp .

显然, V 的子空间 W 与其正交补 W^\perp 的和是直和, 而且

$$W \oplus W^\perp = V.$$

例如, 两个子空间 $W_1 = L(\alpha_1)$ 和 $W_2 = L(\alpha_2, \alpha_3)$ (如图 2-5), 当 $\alpha_1 \perp W_2$ 时, W_1 与 W_2 是正交子空间, 且互为正交补, 而 \mathbf{R}^3 中两个一维子空间 $W_1 = L(\alpha_1)$ 和 $W_2 = L(\alpha_2)$ (如图 2-6), 当 $\alpha_1 \perp \alpha_2$ 时, $W_1 \perp W_2$, 但 W_1 与 W_2 并不互为正交补, 因为 $W_1 + W_2 = L(\alpha_1, \alpha_2) \neq \mathbf{R}^3$.

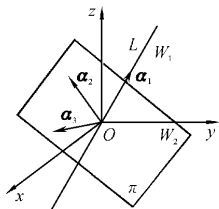


图 2-5

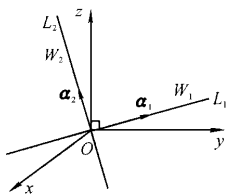


图 2-6

需要注意, 在 \mathbf{R}^3 中过原点的两个垂直平面 π_1 与 π_2 上所有向量分别形成的子空间 W_1 与 W_2 并不是正交子空间, 因为此时 $W_1 \cap W_2 \neq \{0\}$, 而是一个一维子空间.

下面的定理告诉我们如何求一个子空间的正交补.

定理 2.11 如果 W_1 是 n 维欧氏空间 $V(\mathbf{R})$ 的一个子空间, 则

$$W_2 = \{\alpha \mid \alpha \in V \text{ 且 } \alpha \perp W_1\}$$

是 W_1 的正交补.

证 设 $B_1 = \{\varepsilon_1, \dots, \varepsilon_m\}$ 是 W_1 的单位正交基, 将 B_1 扩充为 V 的单位正交基 $B = \{\varepsilon_1, \dots, \varepsilon_m, \varepsilon_{m+1}, \dots, \varepsilon_n\}$. 于是对任意的

$$\alpha = a_1 \varepsilon_1 + \dots + a_m \varepsilon_m + a_{m+1} \varepsilon_{m+1} + \dots + a_n \varepsilon_n \in W_2$$

均有

$$(\alpha, \varepsilon_i) = a_i = 0, \quad i = 1, 2, \dots, m,$$

所以

$$\alpha = a_{m+1} \varepsilon_{m+1} + \dots + a_n \varepsilon_n \in L(\varepsilon_{m+1}, \dots, \varepsilon_n),$$

故 $W_2 \subset L(\varepsilon_{m+1}, \dots, \varepsilon_n)$; 显然, 后者中任一向量 $\alpha \perp W_1$, 所以 $W_2 = L(\varepsilon_{m+1}, \dots, \varepsilon_n)$ 且 $W_2 \perp W_1$; 又 $W_2 + W_1 = V$ 故 W_2 是 W_1 的正交补.

由定理的证明可见

$$\dim W + \dim W^\perp = \dim V.$$

此结论由节 2.6 中的维数公式(2-6)及 $W \cap W^\perp = \{0\}$ 也可得到.

例 设 $W = L(\alpha_1, \alpha_2)$ 是 \mathbb{R}^4 的子空间, 其中 $\alpha_1 = (1, 0, -1, 0)$, $\alpha_2 = (1, 1, 1, 1)$. 求 W^\perp .

解 设 $x = (x_1, x_2, x_3, x_4) \in W^\perp$, 于是 $x \perp W$. 容易证明 $x \perp W$ 的充要条件是

$$(x, \alpha_1) = 0 \quad \text{且} \quad (x, \alpha_2) = 0.$$

因此 x 满足

$$\begin{cases} x_1 - x_3 = 0 \\ x_1 + x_2 + x_3 + x_4 = 0. \end{cases}$$

容易解得 $x = \lambda_1 \beta_1 + \lambda_2 \beta_2$ 其中 $\beta_1 = (1, -2, 1, 0)$, $\beta_2 = (0, -1, 0, 1)$, λ_1, λ_2 为任意实数. 所以 $W^\perp = L(\beta_1, \beta_2)$.

附录 双重连加号 $\sum \sum$ 连乘号 \prod

n 个数 a_1, a_2, \dots, a_n 求和时, 可将 $a_1 + a_2 + \dots + a_n$ 表示为 $\sum_{k=1}^n a_k$, \sum 叫做求和的连加号. 例如

$$\sum_{j=1}^n a_{1j} x_j = a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n.$$

用两个角标编号的 $m \times n$ 个数 a_{ij} ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$)

$$a_{11}, a_{12}, \dots, a_{1j}, \dots, a_{1n}$$

$$\dots, \dots, \dots, \dots, \dots, \dots$$

$$a_{i1}, a_{i2}, \dots, a_{ij}, \dots, a_{in}$$

$$\dots, \dots, \dots, \dots, \dots, \dots$$

$$a_{m1}, a_{m2}, \dots, a_{mj}, \dots, a_{mn}$$

①

求它们的和 S 时, 可以先把第一个角标为 i 的 n 个数相加, 记作

$$S_i = a_{i1} + a_{i2} + \dots + a_{ij} + \dots + a_{in} = \sum_{j=1}^n a_{ij}, \quad i = 1, 2, \dots, m,$$

然后再把 S_1, S_2, \dots, S_m 相加, 得

$$S = \sum_{i=1}^m S_i = \sum_{i=1}^m \sum_{j=1}^n a_{ij}. \quad (2)$$

同样也可以先固定第二个角标, 而对第一个角标 i 求和, 得

$$S'_j = a_{1j} + a_{2j} + \dots + a_{ij} + \dots + a_{mj} = \sum_{i=1}^m a_{ij}, \quad j = 1, \dots, n,$$

然后把 S'_1, S'_2, \dots, S'_n 相加得

$$S = \sum_{j=1}^n S'_j = \sum_{j=1}^n \sum_{i=1}^m a_{ij}. \quad (3)$$

显然, (2), (3) 式相等. 这表明对角标 i 与 j 的求和次序可颠倒.

欲对 (1) 式中部分数求和, 可注明角标应满足的条件, 例如

$$\begin{aligned} \sum_{1 \leq j < i \leq n} a_{ij} &= (a_{21} + a_{31} + \dots + a_{n1}) + (a_{32} + a_{42} + \dots + a_{n2}) \\ &\quad + \dots + (a_{n-1, n-2} + a_{n, n-2}) + a_{n, n-1}. \end{aligned}$$

再如 $a_1 + a_2 + \dots + a_{i-1} + a_{i+1} + \dots + a_n$ 可表示为 $\sum_{j \neq i} a_j$.

以后, 我们有时也用连乘积记号 \prod , 例如 $\prod_{j=1}^n a_j = a_1 a_2 \dots a_n$.

习 题

1. 检验下列集合对指定的加法和数量乘法, 是否构成实数域上的线性空间.

(1) $\mathbf{R}^2 = \{(x, y) \mid x, y \in \mathbf{R}\}$ 对通常的向量加法和如下定义的数量乘法:

$$\lambda \circ (x, y) = (\lambda x, y);$$

(2) 集合 \mathbf{R}^2 其加法同(1), 数量乘法为:

$$\lambda \circ (x, y) = (x, y), \quad \forall \lambda \in \mathbf{R};$$

(3) 集合 \mathbf{R}^2 其加法同(1), 数量乘法为:

$$\lambda \circ (x, y) = \begin{cases} (0, 0), & \lambda = 0, \\ \left(\lambda x, \frac{y}{\lambda}\right), & \lambda \neq 0; \end{cases}$$

(4) 平面上不平行于某一向量 α_0 的所有起点在原点的向量, 对通常的向量加法和数量乘法;

(5) 有理数集 \mathbf{Q} 对普通的数的加法和乘法;

(6) 复数集 \mathbf{C} 对普通的数的加法和乘法;

(7)* 正实数集 \mathbf{R}^+ 对如下定义的加法和数量乘法:

$$a \oplus b = ab, \quad \lambda \circ a = a^\lambda;$$

(8)* $\mathbf{R}_+^n = \{ (a_1, \dots, a_n) \mid a_i \in \mathbf{R}^+, \text{即 } a_i > 0 \}$ 对如下定义的向量加法和数量乘法:

$$(a_1, \dots, a_n) \oplus (b_1, \dots, b_n) = (a_1 b_1, \dots, a_n b_n),$$

$$\lambda \circ (a_1, \dots, a_n) = (a_1^\lambda, \dots, a_n^\lambda);$$

(9) 集合 V 为区间 $[a, b]$ 上所有函数值 ≥ 0 的实变量函数, 即

$$V = \{ f \mid f(x) \geq 0, \forall x \in [a, b] \}.$$

对通常的函数加法和数与函数的乘法, 即:

$$(f \oplus g)(x) = f(x) + g(x),$$

$$(\lambda \circ f)(x) = \lambda f(x);$$

(10) $V_1 = \{ f \mid x \in \mathbf{R}, f(x) \in \mathbf{R} \text{ 且 } f(-x) = -f(x) \},$

$$V_2 = \{ f \mid x \in \mathbf{R}, f(x) \in \mathbf{R}, f(0) = 1 \text{ 且 } f(-x) = f(x) \}.$$

对题(9)所定义的加法和数量乘法;

(11)* $V = \{ f \mid x \in \mathbf{R}, f(x) \in \mathbf{C} \text{ (即 } f \text{ 为实变量复值函数), 且 } f(-x) = \overline{f(x)} \text{ 后者为 } f(x) \text{ 的共轭复数)} \}$ 对题(9)所定义的加法和数量乘法.

2. 在向量空间 $V(F)$ 中:

(1) $\forall \lambda, \mu \in F, \alpha, \beta \in V$, 计算 $(\lambda - \mu)(\alpha - \beta)$.

(2) 设 $x \in V$, 对 $\forall \alpha_i \in V$ 及 $\forall \lambda_i \in F$ (但 $\lambda_1 + \dots + \lambda_n \neq 0$) 求解向量方程 (即求解 x): $\lambda_1(\alpha_1 + x) + \dots + \lambda_n(\alpha_n + x) = 0$.

3. 判断下列子集是否为给定线性空间的子空间 (对 \mathbf{R}^3 中的子集并说明其几何意义):

(1) $W = \{ (x_1, \dots, x_n) \in F^n \mid a_1 x_1 + \dots + a_n x_n = 0 \}$ 其中 a_1, \dots, a_n 为域 F 中的固定数量};

(2) $W_1 = \{ (x, y, 0) \in \mathbf{R}^3 \}, \quad W_2 = \{ (x, y, 0) \in \mathbf{R}^3 \};$

(3) $W_1 = \{ (x, y, z) \in \mathbf{R}^3 \mid x - 3y + z = 0 \},$

$$W_2 = \{ (x, y, z) \in \mathbf{R}^3 \mid x - 3y + z = 1 \};$$

$$(4) W_1 = \left\{ (x, y, z) \in \mathbf{R}^3 \mid \frac{x}{2} = \frac{y-4}{1} = \frac{z-1}{-3} \right\},$$

$$W_2 = \{ (x, y, z) \in \mathbf{R}^3 \mid x - y = 0 \text{ 且 } x + y + z = 0 \};$$

$$(5) W_1 = \{ f(x) \in R[x] \mid f(1) = 0 \},$$

$$W_2 = \{ f(x) \in R[x] \mid f(1) = f(0) \};$$

(6) $W = \{ f \in F(-\infty, +\infty) \mid f(-x) = f(x), \forall x \in \mathbf{R} \}$, 其中 $F(-\infty, +\infty)$ 是所有定义在 $(-\infty, +\infty)$ 上的实值函数对通常的函数加法及数与函数的乘法在实数域上构成的线性空间.

4. 设 $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{R}^n$, $c_1, c_2, c_3 \in \mathbf{R}$, 如果 $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \mathbf{0}$, 且 $c_1c_3 \neq 0$, 证明: $L(\alpha_1, \alpha_2) = L(\alpha_2, \alpha_3)$.

5. 设 W 是线性空间 V 的一个非平凡子空间, 给定非零向量 $\beta \in V \setminus W$, 证明: $W_1 = \{ \beta + \alpha \mid \alpha \in W \}$ 不是 V 的一个子空间.

6. 设 $\alpha_1 = (1, 0, 1)$, $\alpha_2 = (1, 1, 0)$, $\alpha_3 = (1, -1, 2)$, 问下列 β_1, β_2 属于 $L(\alpha_1, \alpha_2, \alpha_3)$ 吗? 如属于, 它们由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示唯一吗? 为什么?

$$(1) \beta_1 = (1, -1, -1); \quad (2) \beta_2 = (1, 2, -1).$$

7. 判别下列向量组的线性相关性:

$$(1) \alpha_1 = (1, 1, 1), \quad \alpha_2 = (0, 2, 5), \quad \alpha_3 = (1, 3, 6);$$

$$(2) \beta_1 = (1, -1, 2, 4), \quad \beta_2 = (0, 3, 1, 2), \quad \beta_3 = (3, 0, 7, 14);$$

(3) $R[x]$ 中: $p_1(x) = 1$, $p_2(x) = (x - x_0)$, $p_3(x) = (x - x_0)^2$, 其中常数 $x_0 \in \mathbf{R}$.

8. 证明: 如果向量组有一个部分组线性相关, 则整个向量组也线性相关.

9. 设向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_s\} \subset \mathbf{R}^n$, $\beta_1, \beta_2, \dots, \beta_s$ 是分别在 $\alpha_1, \alpha_2, \dots, \alpha_s$ 的第 n 个分量后任意加一个分量 b_1, b_2, \dots, b_s 而形成的 $n+1$ 元向量组, 证明: 若 $\alpha_1, \alpha_2, \dots, \alpha_s$ 线性无关, 则 $\beta_1, \beta_2, \dots, \beta_s$ 也线性无关. 并叙述这个命题的等价命题.

10. 下列命题是否正确? 如正确, 证明之; 如不正确, 举反例.

(1) 若 $\alpha_1, \dots, \alpha_m$ ($m > 2$) 线性相关, 则其中每一向量都是其余向量的线性组合;

(2) 若 $\alpha_1, \dots, \alpha_m$ 线性无关, 则其中每一向量都不是其余向量的线性组合, 这个命题的等价命题应如何叙述?

(3) $\alpha_1, \dots, \alpha_m$ ($m > 2$) 线性无关的充要条件是任意两个向量都线性无关;

(4) 若 α_1, α_2 线性相关, β_1, β_2 线性相关, 则 $\alpha_1 + \beta_1, \alpha_2 + \beta_2$ 也线性相关;

(5) 若 $\alpha_1, \dots, \alpha_n$ 线性无关, 则 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \dots, \alpha_{n-1} + \alpha_n, \alpha_n + \alpha_1$ 也线性无关;

(6) 若 $\alpha_1, \alpha_2, \alpha_3$ 线性相关, 则 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_1$ 也线性相关;

(7)* 设 $B = \{\alpha_1, \alpha_2, \alpha_3\}$ 是 \mathbb{R}^3 的一组基, 非零向量 $\alpha_0 \in \mathbb{R}^3$, 则 $\{\alpha_0 + \alpha_1, \alpha_0 + \alpha_2, \alpha_0 + \alpha_3\}$ (其中三个向量均是非零向量) 也是 \mathbb{R}^3 的一组基;

(8) 设 $B = \{\alpha_1, \alpha_2\}$ 是 \mathbb{R}^2 的一组基, 则 $\{\alpha_1 + \alpha_2, \alpha_1 - \alpha_2\}$ 也是 \mathbb{R}^2 的一组基;

(9) 一个有限维线性空间只含有有限个子空间;

(10)* 如果 W_1, W_2 是 \mathbb{R}^n 的两个子空间, B_1, B_2 分别是 W_1, W_2 的基, 则存在 \mathbb{R}^n 的一组基 B , 使得 $B \supset \{B_1 \cup B_2\}$.

11. 在函数空间中, 判断下列向量组的线性相关性, 并证明之:

(1) $1, \sin^2 x, \cos^2 x$; (2)* $1, 2^x, 2^{-x}$.

12. 设在线性空间 $V(F)$ 中, 向量 β 是 $\{\alpha_1, \dots, \alpha_r\}$ 的线性组合, 但不是 $\{\alpha_1, \dots, \alpha_{r-1}\}$ 的线性组合, 证明: $L(\alpha_1, \dots, \alpha_{r-1}, \alpha_r) = L(\alpha_1, \dots, \alpha_{r-1}, \beta)$.

13. 若 $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ 线性相关, 但其中任意三个向量线性无关, 则存在一组全不为零的数 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, 使得 $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \lambda_3 \alpha_3 + \lambda_4 \alpha_4 = 0$.

14. 设向量组 $\alpha_1, \dots, \alpha_r$ 线性无关, 证明: $\beta, \alpha_1, \dots, \alpha_r$ 线性无关的充要条件是 β 不能由 $\alpha_1, \dots, \alpha_r$ 线性表示.

15*. 设 $\alpha_1, \dots, \alpha_n \in F^n$, 证明: $\alpha_1, \dots, \alpha_n$ 线性无关的充要条件是 F^n 中任一向量都可由它们线性表示.

16*. 证明: 若向量 α 可经向量组 $\{\alpha_1, \dots, \alpha_r\}$ 线性表示, 则表示法唯一的充要条件是 $\{\alpha_1, \dots, \alpha_r\}$ 线性无关.

17. 在线性空间 $V(F)$ 中, 对于给定的一个向量组 $\{\alpha_1, \dots, \alpha_n\}$, 如何判断它是否是 $V(F)$ 的一组基向量? 如果已知 $\dim V = n$, 又如何判断 $\{\alpha_1, \dots, \alpha_n\}$ 是否是 $V(F)$ 的一组基向量? 什么是有限维线性空间的维数?

18. 求下列线性空间的维数及其一组基(向量):

(1) 第1题中的(6)(7)(8);

(2) 全体二元复数向量在复数域上构成的线性空间;

(3)* 全体二元复数向量在实数域上构成的线性空间.

19. 求下列子空间的一组基及其维数:

(1) $W_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0, x_1 - x_2 + x_3 = 0\}$;

(2) $W_2 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 + x_2 - x_3 - x_4 = 0\}$;

$$(3) W_3 = \{ (x_1, x_2, x_3, x_4) \in \mathbf{R}^4 \mid x_1 - x_2 - x_3 + x_4 = 0 \};$$

$$(4) W_4 = L(\alpha_1, \alpha_2, \alpha_3) \text{ 其中}$$

$$\alpha_1 = (1, 2, 1), \alpha_2 = (1, 1, -1), \alpha_3 = (1, 3, 3);$$

$$(5) W_5 = L(\beta_1, \beta_2, \beta_3) \text{ 其中}$$

$$\beta_1 = (2, 3, -1), \beta_2 = (1, 2, 2), \beta_3 = (1, 1, -3).$$

20. 设

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

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

证明 $(1) \{e_1 + e_4, e_2 + e_3\}$ 是 W_1 的基, $\{e_1 - e_4, e_2 - e_3\}$ 是 W_2 的基, 其中 $\{e_1, e_2, e_3, e_4\}$ 是 \mathbf{R}^4 的自然基.

(2)* $\mathbf{R}^4 = W_1 \oplus W_2$, 即 \mathbf{R}^4 是 W_1 与 W_2 的直和.

21. 已知 $\{\alpha_1, \dots, \alpha_n\}$ 是线性空间 $V(F)$ 的一组基向量, 如何求 $V(F)$ 中任一向量关于这一组基的坐标?

22. 证明 $\alpha_1 = (1, 0, 1), \alpha_2 = (1, 1, 0), \alpha_3 = (0, 1, 1)$ 是 \mathbf{R}^3 的一组基, 并求 $\beta = (1, -2, 4)$ 关于这组基的坐标.

23. 已知 \mathbf{R}^2 的两组基为 $\{\alpha_1, \alpha_2\}$ 和 $\{e_1, e_2\}$, 其中 $\alpha_1 = (2, -1), \alpha_2 = (5, -4), e_1 = (1, 0), e_2 = (0, 1)$. 试求一个非零向量 $\beta \in \mathbf{R}^2$, 使 β 关于这两组基有相同的坐标, 并求这个 β 关于基 $\{\xi_1, \xi_2\}$ 的坐标, 其中 $\xi_1 = (-1, 1), \xi_2 = (1, 1)$.

24. 证明 $\{1, x - 2, (x - 2)^2\}$ 是 $\mathbf{R}[x]_3$ 的一组基, 并求 $f(x) = a + bx + cx^2$ 关于这组基的坐标.

25. 已知 $B = \{1, x - a, (x - a)^2 \mid a \neq 0\}$ 是 $\mathbf{R}[x]_3$ 的一组基, 求 $\mathbf{R}[x]_3$ 的自然基 $\{1, x, x^2\}$ 中每个向量关于基 B 的坐标.

26*. 求下列子空间的交与和的维数及其一组基:

(1)* 第 19 题中的 W_4 与 W_5 ;

(2) 第 19 题中的 W_2 与 W_3 .

27. (1) 把 19 题中 W_4 的基扩充为 \mathbf{R}^3 的基;

(2) 把 19 题中的 W_2 基扩充为 \mathbf{R}^4 的基.

28*. 设 W_1, W_2 是线性空间 V 的两个子空间, $\dim W_1 = m, \dim W_2 = n, m \leq n$, 证明:

$$(1) \dim(W_1 \cap W_2) \leq m; \quad (2) \dim(W_1 + W_2) \leq m + n.$$

29*. (1) 对 19 题中的 W_1 , 求 V_1 , 使 $\mathbf{R}^3 = W_1 \oplus V_1$, 并问这样的 V_1 是

否唯一?

(2) 对 19 题中的 W_5 求其补空间 V_5 ;

(3) 设 W 是 \mathbf{R}^n 的 k 维子空间 ($0 < k < n$) 如何求 W 的补空间 V .

30*. 证明: 每个 n 维线性空间都可以表示成 n 个一维子空间的直和.

31. 已知 $\alpha = (1, 2, -1, 1)$, $\beta = (2, 3, 1, -1)$, $\gamma = (-1, -1, -2, 2)$:

(1) 求 α, β, γ 的长度及夹角 α, β 与 α, γ ;

(2) 求与 α, β, γ 都正交的所有向量.

32. 求与向量 $(1, 1, -1, 1)$ $(1, -1, -1, 1)$ $(2, 1, 1, 3)$ 都正交的单位向量.

33*. $W = L(\alpha_1, \dots, \alpha_m)$ 是 \mathbf{R}^n 的一个子空间, 已知 β 与向量组 $\{\alpha_1, \dots, \alpha_m\}$ 中每个向量都正交, 证明: $\beta \perp W$.

34*. $W = L(\alpha_1, \dots, \alpha_m)$ 是内积空间 V 的一个子空间, 设 $\beta_1, \dots, \beta_k \in V$, 证明: 如果 $(\beta_i, \alpha_j) = 0$ $i = 1, \dots, k$ $j = 1, \dots, m$ 则 $L(\beta_1, \dots, \beta_k) \perp W$.

35. $\forall \alpha = (x_1, x_2), \beta = (y_1, y_2) \in \mathbf{R}^2$ 定义:

$$(\alpha, \beta) = ax_1y_1 + bx_1y_2 + cx_2y_1 + dx_2y_2$$

(其中 $a, b, c, d \in \mathbf{R}$) 问: a, b, c, d 满足什么条件时 (α, β) 是 \mathbf{R}^2 上的一个内积?

36. 用 Schmidt 正交化过程, 由下列向量组关于 \mathbf{R}^n 的标准内积分别构造一组单位正交向量组:

(1) $\alpha_1 = (1, 1, 1)$, $\alpha_2 = (0, 1, 1)$, $\alpha_3 = (0, 0, 1)$;

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

37. 设 $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5\}$ 是欧氏空间 V 的一组单位正交基, $W = L(\alpha_1, \alpha_2, \alpha_3)$ 其中 $\alpha_1 = \varepsilon_1 + \varepsilon_5$, $\alpha_2 = \varepsilon_1 - \varepsilon_2 + \varepsilon_4$, $\alpha_3 = 2\varepsilon_1 + \varepsilon_2 + \varepsilon_3$. 试求 W 的一组单位正交基.

38*. 求齐次线性方程组

$$\begin{cases} 2x_1 + x_2 + 3x_3 - x_4 = 0 \\ 3x_1 + 2x_2 - 2x_4 = 0 \\ 3x_1 + x_2 + 9x_3 - x_4 = 0 \end{cases}$$

的解空间 S 的正交补 S^\perp .

39*. 设 W 是 \mathbf{R}^n 的非平凡子空间, $\alpha \in W$, 证明: $\exists \beta \in \mathbf{R}^n$ 使 $\beta \in W^\perp$,

且 $(\alpha, \beta) \neq 0$.

40. 设 $\{\varepsilon_1, \dots, \varepsilon_n\}$ 是 n 维欧氏空间 V 的一组单位正交基, 证明:

(1) 如果 $\beta \in V$, 且 $(\beta, \varepsilon_i) = 0 (i = 1, \dots, n)$ 则 $\beta = 0$;

(2) 如果 $\beta_1, \beta_2 \in V$, 且 $\forall \alpha \in V$ 均有 $(\beta_1, \alpha) = (\beta_2, \alpha)$ 则 $\beta_1 = \beta_2$.

41. 设 ξ 是 n 维欧氏空间中的一个固定的非零向量, 证明:

(1) $W = \{\alpha | (\alpha, \xi) = 0, \alpha \in V\}$ 是 V 的一个子空间;

(2) $\dim W = n - 1$.

补充题

1*. 设 E 是域 F 的一个子域.

(1) 证明: F 关于自身的加法和乘法, 构成一个 E 上的向量空间, 并举一例;

(2) 举例说明, $E (E \neq F)$ 不是 F 上的向量空间;

(3) 证明: 若 V 是 F 上的一个向量空间, 则 V 也是 E 上的一个向量空间.

2. 设 V 是一个线性空间, W 是 V 的子集. 证明: W 是 V 的子空间 $\Leftrightarrow L(W) = W$.

3. 设 S_1, S_2 是线性空间 V 的两个子集, 证明:

$$L(S_1 \cap S_2) \subseteq L(S_1) \cap L(S_2),$$

并在 \mathbb{R}^3 中分别举出两个例子, 使得上式中等号成立和等号不成立.

4. 证明: 若 \mathbb{R}^2 中的向量 $\alpha = (a_1, a_2), \beta = (b_1, b_2)$ 满足

$$a_1 b_1 + a_2 b_2 = 0 \quad a_1^2 + a_2^2 = b_1^2 + b_2^2 = 1,$$

则 α, β 是 \mathbb{R}^2 的一组基.

5. 设 $\{\alpha, \beta\}$ 是 \mathbb{R}^2 的一组基, 又

$$\alpha' = c_{11}\alpha + c_{12}\beta, \quad \beta' = c_{21}\alpha + c_{22}\beta$$

证明: $\{\alpha', \beta'\}$ 是 \mathbb{R}^2 的基 $\Leftrightarrow c_{11}c_{22} - c_{12}c_{21} \neq 0$.

6*. 设向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_r\}$ 线性无关, 证明: 在向量组 $\{\beta, \alpha_1, \alpha_2, \dots, \alpha_r\}$ 中至多有一个向量 $\alpha_i (1 \leq i \leq r)$ 可被其前面的 i 个向量 $\{\beta, \alpha_1, \dots, \alpha_{i-1}\}$ 线性表示.

7*. 证明: $\alpha_1, \alpha_2, \dots, \alpha_r$ (其中 $\alpha_1 \neq 0$) 线性相关 \Leftrightarrow 存在一个 $\alpha_i (1 < i \leq r)$, 使得 α_i 可以由 $\alpha_1, \alpha_2, \dots, \alpha_{i-1}$ 线性表示, 且表示法唯一.

8*. 设 V_1, V_2 是线性空间 V 的两个非平凡子空间, 证明: $\exists \alpha \in V$, 使

$\alpha \in V_1$ 和 $\alpha \in V_2$ 同时成立,并在 \mathbf{R}^3 中举一例.

9*. 设 V_1, \dots, V_m ($m > 2$) 是线性空间 V 的 m 个非平凡子空间,证明: V 中存在同时不属于 V_1, \dots, V_m 的向量. 并在 \mathbf{R}^3 中举一例.

10*. 设 $S_1 = \{\alpha_1, \dots, \alpha_s\}$ 和 $S_2 = \{\beta_1, \dots, \beta_t\}$ 是向量空间 V 的两个线性无关的子集,证明:

$$(\alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_t) \text{ 线性无关} \Leftrightarrow L(S_1) \cap L(S_2) = \{0\}.$$

11*. 设 W_1, W_2, W_3, W_4 都是 \mathbf{R}^3 的子空间,如果 $W_1 \subseteq W_2 = W_3 + W_4$, 是否必有

$$W_1 = W_1 \cap W_2 = (W_1 \cap W_3) + (W_1 \cap W_4).$$

12. 在三维复向量空间 \mathbf{C}^3 中,对任意向量 $\alpha = (x_1, x_2, x_3)$ 和 $\beta = (y_1, y_2, y_3)$ 定义 $(\alpha, \beta) = \overline{x_1 y_1} + \overline{x_2 y_2} + \overline{x_3 y_3}$.

(1) 验证 (α, β) 是 \mathbf{C}^3 上的一个内积(称为 \mathbf{C}^3 上的标准内积).

(2)* 已知 $W = L(\xi)$ 其中 $\xi = (i, 0, 1)$, 求 W 及 W^\perp 关于 \mathbf{C}^3 的标准内积的单位正交基.

13. 设 V 是实数域 \mathbf{R} 或复数域 \mathbf{C} 上的内积空间,证明:

$$(1) ||\alpha| - |\beta|| \leq |\alpha - \beta|, \quad \alpha, \beta \in V;$$

$$(2) |\alpha \pm \beta|^2 = |\alpha|^2 \pm 2\operatorname{Re}(\alpha, \beta) + |\beta|^2, \quad \alpha, \beta \in V,$$

其中 $\operatorname{Re}(\alpha, \beta)$ 表示复数 (α, β) 的实部;

$$(3) \forall \alpha, \beta \in V(\mathbf{R}) (\alpha, \beta) = \frac{1}{4} |\alpha + \beta|^2 - \frac{1}{4} |\alpha - \beta|^2.$$

部分习题和补充题答案

习题

1. (6)(7)(8)(11) 均是 (1)(2)(3)(4)(5)(9) 均不是 (10) V_1 是, V_2 不是.

$$2. (1) \lambda\alpha - \mu\alpha - \lambda\beta + \mu\beta; (2) -(\lambda_1 + \dots + \lambda_n)^{-1}(\lambda_1\alpha_1 + \dots + \lambda_n\alpha_n).$$

3. (1) 是; (2) W_1 不是, W_2 是 (xOy 平面上的全体向量); (3) W_1 是, W_2 不是 (不过原点的平面上的全体向量); (4) W_1 不是 (不过原点的直线上的全体向量), W_2 是 (过原点的直线上的全体向量); (5)(6) 均是.

5. (提示: 证明零向量 $0 \notin W_1$.)

6. (1) 不属于; (2) 属于, 不唯一, $\alpha_1, \alpha_2, \alpha_3$ 线性相关.

7. (1)(2) 均线性相关; (3) 线性无关.

10.(1)不正确;(3)(4)均不正确;(5) n 为奇数时线性无关, n 为偶数时线性相关;(6)正确;(7)不正确(提示:取 α_0 使 $\alpha_0 + \alpha_3 = (\alpha_0 + \alpha_1) + (\alpha_0 + \alpha_2)$).(8)正确;(9)不正确;(10)不正确.

11.(1)线性相关;(2)线性无关.

12.提示:利用节 2.4 定义 2.8 后的结论(4)及 $\beta = c_1\alpha_1 + \dots + c_{r-1}\alpha_{r-1} + c_r\alpha_r$, 其中 $c_r \neq 0$.

14.提示:证充分性时,用反证法,设 $\beta, \alpha_1, \dots, \alpha_r$ 线性相关.

15.提示:证充分性时,利用 F^n 的自然基可被 $\alpha_1, \dots, \alpha_n$ 线性表示.

16.提示:证必要性时,用反证法,设 $\alpha_1, \dots, \alpha_r$ 线性相关.

18.(1)⑥2 $\{1, i\}$, ⑦1 $\{c\}$ $c > 0, c \neq 1$,

⑧ n $\{e_1, e_2, \dots, e_n\}$ 其中 $e_i = (1, \dots, 1, c_i, 1, \dots, 1)$ $c_i > 0, c_i \neq 1, i = 1, \dots, n$;

(2)2 $\{e_1, e_2\}$, (3)4 $\{(1, 0), (i, 0), (0, 1), (0, i)\}$.

19.(1)1 $\{(-1, 0, 1)\}$; (2)3 $\{(-1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)\}$;
(3)3 (4)2 $\{\alpha_1, \alpha_2\}$; (5)2 $\{\beta_1, \beta_2\}$.

22. $\left(\frac{7}{2}, -\frac{5}{2}, \frac{1}{2}\right)$. 23. $\beta = (-5, 1)$ 在 $\{\xi_1, \xi_2\}$ 下坐标为 $(3, -2)$.

24. $(a + 2b + 4c, b + 4c, c)$. 25. $(1, 0, 0), (a, 1, 0), (a^2, 2a, 1)$.

26.(1)交: 1 $\{\beta_1 + \beta_2\}$ 和 3 $\{\alpha_1, \alpha_2, \beta_1\}$; (2)交: 2 $\{(0, 1, 0, 1), (1, 0, 1, 0)\}$ 和 4 \mathbb{R}^4 的基.

27.(提示:利用第 14 题的结论.) 1 $\{\alpha_1, \alpha_2, (1, 0, 0)\}$; (2) $\{(-1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1), (1, 0, 0, 0)\}$.

28.利用维数公式. 29.(1) $V_1 = L\{(0, 1, 0), (0, 0, 1)\}$, 不唯一;
(2) $V_5 = L\{(0, 1, 0)\}$, 不唯一.

31.(1) $\sqrt{7}, \sqrt{15}, \sqrt{10}, \arccos \frac{6}{\sqrt{105}}, \arccos \frac{1}{\sqrt{70}}$.

32. $\pm \frac{(-4, 0, -1, 3)}{\sqrt{26}}$. 35. $c = b, a > 0, ad - b^2 > 0$.

36.(1) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right), \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$;

(2) $\frac{1}{\sqrt{10}}(1, 2, 2, -1), \frac{1}{\sqrt{26}}(2, 3, -3, 2), \frac{1}{\sqrt{10}}(2, -1, -1, -2)$;

(3) $\frac{1}{\sqrt{7}}(1, 1, -1, 2), \frac{1}{\sqrt{4431}}(26, 47, -5, -39), \frac{1}{\sqrt{422}}(-4, 9, 17, 6)$;

$$(4) \frac{1}{\sqrt{15}}(2, 1, 3, -1), \frac{1}{\sqrt{23}}(3, 2, -3, -1), \frac{1}{\sqrt{127}}(1, 5, 1, 10).$$

$$37. \frac{1}{\sqrt{2}}(\varepsilon_1 + \varepsilon_5), \frac{1}{\sqrt{10}}(\varepsilon_1 - 2\varepsilon_2 + 2\varepsilon_4 - \varepsilon_5), \frac{1}{2}(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 - \varepsilon_5).$$

$$38. L\{(2, 1, 3, -1), (3, 2, 0, -2)\}.$$

39. 用反证法, 或用构造性证法(设 $W = L(\varepsilon_1, \dots, \varepsilon_k), \{\varepsilon_1, \dots, \varepsilon_k, \varepsilon_{k+1}, \dots, \varepsilon_n\}$ 为单位正交基).

40. (1) 提示: 设 $\beta = a_1\varepsilon_1 + \dots + a_n\varepsilon_n$, 证明 a_1, \dots, a_n 全等于零;

(2) 提示: 利用内积的性质及本题(1)的结论.

补 充 题

1. 如 E 是实数域, F 是复数域.

3. 例 1: $\alpha_1, \alpha_2, \alpha_3$ 线性无关, $S_1 = \{\alpha_1, \alpha_1\}, S_2 = \{\alpha_1, \alpha_2\}$ 等号成立; 例 2. $\alpha_1, \alpha_2, \alpha_3$ 共面, S_1, S_2 同例 1, 等号不成立.

6. 用反证法.

$$8. V_1 = L\{(0, 0, 1)\}, V_2 = L\{(1, 0, 0), (0, 1, 0)\}, \alpha = (1, 1, 1).$$

9. 利用上题结论, 用数学归纳法证明.

10. 用维数公式.

11. 第一个等号成立, 第二个等号不一定成立.

$$12. (2) \left\{ \frac{1}{\sqrt{2}}(i, 0, 1) \right\}; \left\{ \frac{1}{\sqrt{2}}(1, 0, i), (0, 1, 0) \right\}.$$

第3章 线性映射

线性空间上的线性映射概念是一元线性函数概念的推广. 我们主要讨论有限维线性空间的线性映射. 本章的重点是线性映射的定义和简单性质, 线性映射的像和核的概念, 有限维线性空间的线性映射与其基像的关系, 线性映射的秩, 以及两个有限维线性空间同构的充分必要条件是它们的维数相同.

3.1 线性映射的定义及例

一元线性函数 $f(x) = ax$ (a 为常数) 是最简单的一类数值函数, 它是 \mathbf{R} 到 \mathbf{R} 的一种映射, 一元线性函数的基本性质可用

$$f(a_1x_1 + a_2x_2) = a_1f(x_1) + a_2f(x_2)$$

来表示. 如果把这类函数(即映射)的定义域和值域推广到一般的线性空间 $V_1(F)$ 和 $V_2(F)$, 并保留上述性质, 就得到线性空间上的线性映射的概念.

定义 3.1 从线性空间 $V_1(F)$ 到 $V_2(F)$ 的一个映射 σ 是线性的, 如果 $\forall \alpha, \beta \in V_1$ 和 $\forall \lambda, \mu \in F$ 都有

$$\sigma(\lambda\alpha + \mu\beta) = \lambda\sigma(\alpha) + \mu\sigma(\beta). \quad (3-1)$$

从线性空间 V 到自身的线性映射 σ 也叫做 V 上的线性变换.

从线性空间 $V(F)$ 到域 F 的线性映射 f 叫做 V 上的线性函数(或称线性形式).

设 $f: \mathbf{R} \rightarrow \mathbf{R}$ 为满足(3-1)的线性函数, 则 $\forall \lambda \in \mathbf{R} (\lambda \neq 0)$,

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

于是

$$f(y) = \frac{y}{\lambda} f(\lambda) = \frac{f(\lambda)}{\lambda} y,$$

$\forall \lambda \in \mathbf{R}$ 成立, 所以 $\frac{f(\lambda)}{\lambda}$ 必为常数 a , 否则同一个 y 将映射到不同的像.

注意 $f(x) = ax + b$ ($b \neq 0$) 不是 \mathbf{R} 上的线性函数.

式(3-1)也可等价地表述为: $\forall \alpha, \beta \in V_1$ 和 $\forall \lambda \in F$ 都有

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta),$$

$$\sigma(\lambda\alpha) = \lambda\sigma(\alpha). \quad (3-2)$$

式(3-1)与(3-2)的等价性,读者不难自行证明.

在式(3-1)中取 $\lambda = \mu = 0$ 和 $\lambda = -1, \mu = 0$, 立即可得

$$\alpha(\mathbf{0}_1) = \mathbf{0}_2, \alpha(-\alpha) = -\alpha(\alpha). \quad (3-3)$$

例1 线性空间 $V(F)$ 上的恒等变换 I , 数乘变换(位似变换) Λ 和零变换 θ , 即

$$I(\alpha) = \alpha, \quad \forall \alpha \in V,$$

$$\Lambda(\alpha) = \lambda\alpha \quad \forall \alpha \in V (\lambda \text{ 是 } F \text{ 中固定数量}),$$

$$\theta(\alpha) = \mathbf{0}, \quad \forall \alpha \in V (\mathbf{0} \text{ 是 } V \text{ 的零元})$$

都是 V 上的线性变换(读者容易按定义 3.1 加以验证).

下面的例 2—例 5 是 \mathbf{R}^2 中常用的线性变换.

例2 旋转变换— \mathbf{R}^2 中每个向量绕原点按逆时针方向旋转 θ 角的变换 r_θ , 即 $\forall \alpha = (x, y) \in \mathbf{R}^2$,

$$r_\theta(\alpha) = \alpha' \quad \text{即} \quad r_\theta(x, y) = (x', y') \quad (1)$$

(如图 3-1) 是 \mathbf{R}^2 上的一个线性变换. 因为: ① 式中的 x', y' 分别为

$$x' = r \cos(\theta + \beta) = x \cos \theta - y \sin \theta,$$

$$y' = r \sin(\theta + \beta) = x \sin \theta + y \cos \theta,$$

其中 $\beta = \angle \alpha, \overrightarrow{Ox}$, 即

$$r_\theta(\alpha) = r_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta), \quad (2)$$

于是 $\forall \alpha_1 = (x_1, y_1), \alpha_2 = (x_2, y_2) \in \mathbf{R}^2$ 和 $\forall \lambda, \mu \in \mathbf{R}$,

$$\begin{aligned} r_\theta(\lambda\alpha_1 + \mu\alpha_2) &= r_\theta(\lambda x_1 + \mu x_2, \lambda y_1 + \mu y_2) \\ &= ((\lambda x_1 + \mu x_2) \cos \theta - (\lambda y_1 + \mu y_2) \sin \theta, \\ &\quad (\lambda x_1 + \mu x_2) \sin \theta + (\lambda y_1 + \mu y_2) \cos \theta) \\ &= \lambda(x_1 \cos \theta - y_1 \sin \theta, x_1 \sin \theta + y_1 \cos \theta) \\ &\quad + \mu(x_2 \cos \theta - y_2 \sin \theta, x_2 \sin \theta + y_2 \cos \theta) \\ &= \lambda r_\theta(x_1, y_1) + \mu r_\theta(x_2, y_2) = \lambda r_\theta(\alpha_1) + \mu r_\theta(\alpha_2), \end{aligned}$$

故 r_θ 是 \mathbf{R}^2 上的一个线性变换.

例3 镜像变换(或镜面反射)— \mathbf{R}^2 中每个向量关于过原点的直线 L (看作镜面)相对称的变换 φ , 即 $\forall \alpha = \overrightarrow{OA} \in \mathbf{R}^2$,

$$\varphi(\alpha) = \alpha' = \overrightarrow{OB}$$

(如图 3-2, 其中 A, B 对称于直线 L) 也是 \mathbf{R}^2 上的一个线性变换.

事实上, 设直线 L 的某个方向的单位向量为 ω (如图 3-2), 则 $\overrightarrow{OC} = (\alpha, \omega)\omega$, 其中 (α, ω) 为向量 α 与 ω 的内积, 于是

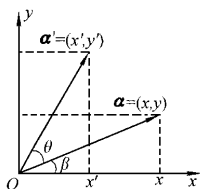


图 3-1

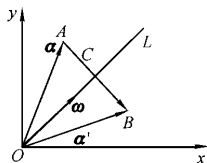


图 3-2

所以 $\alpha' = \alpha + \overrightarrow{AB} = \alpha + 2\overrightarrow{AC} = \alpha + 2(\overrightarrow{OC} - \alpha) = -\alpha + 2(\alpha, \omega)\omega$

$$\varphi(\alpha) = \alpha' = -\alpha + 2(\alpha, \omega)\omega$$

下面验证 φ 是线性变换. $\forall \alpha, \beta \in \mathbb{R}^2$ 和 $\forall \lambda, \mu \in \mathbb{R}$ 都有

$$\begin{aligned}\varphi(\lambda\alpha + \mu\beta) &= -(\lambda\alpha + \mu\beta) + 2(\lambda\alpha + \mu\beta, \omega)\omega \\ &= \lambda[-\alpha + 2(\alpha, \omega)\omega] + \mu[-\beta + 2(\beta, \omega)\omega] \\ &= \lambda\varphi(\alpha) + \mu\varphi(\beta),\end{aligned}$$

故 φ 是线性变换.

如果设 $\alpha = (x, y)$, $\omega = (\cos\theta, \sin\theta)$ 则

$$\begin{aligned}\varphi(\alpha) &= -\alpha + 2(\alpha, \omega)\omega \\ &= -(x, y) + 2(x\cos\theta + y\sin\theta)(\cos\theta, \sin\theta) \\ &= (-x + 2x\cos^2\theta + 2y\sin\theta\cos\theta, -y + 2x\cos\theta\sin\theta + 2y\sin^2\theta) \\ &= (x\cos 2\theta + y\sin 2\theta, x\sin 2\theta - y\cos 2\theta)\end{aligned}$$

例 4 比例变换— \mathbb{R}^2 中每个向量 $\alpha = (x, y)$ 的坐标各按一定比例放大或缩小的变换 ρ :

$$\rho(x, y) = (cx, dy) = \alpha'$$

也是 \mathbb{R}^2 上的一个线性变换(其中 $c, d \in \mathbb{R}$)(验证留给读者).

当 $c = 1, d = 2$ 时, 圆 $x^2 + y^2 = 1$ 变换为椭圆 $x^2 + \frac{y^2}{4} = 1$.

当 $c = d = 2$ 时, 圆 $x^2 + y^2 = 1$ 变换为圆 $x^2 + y^2 = 4$.

例 5 在 \mathbb{R}^2 上作怎样的变换, 会使矩形 $OAPB$ 所围的平面区域变换为平行四边形 $OAP'B'$ 所围的平面区域?

任作平行于 x 轴的直线 $B_1P'_1$, 如果此直线上任一向量 $\alpha = (x, y)$ 变换为同一直线上的向量 $\alpha' = (x', y')$ (如图 3-3) 其中

$$y' = y, \quad x' - x = |B_1B'_1| = y \tan \theta = cy,$$

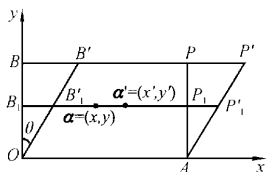


图 3-3

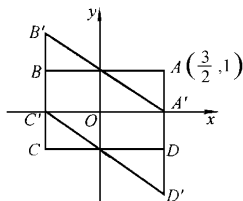


图 3-4

那么这个变换 q_x :

$$q_x(x, y) = (x + cy, y)$$

就会实现上述要求.

我们称变换 q_x 为沿 x 轴方向关于 y 的错切变换. 类似地

$$q_y(x, y) = (x, y + bx).$$

称为沿 y 轴方向关于 x 的错切变换. 变换 q_y (取 $b = -\frac{3}{2}$) 把图 3-4 中的矩形 $ABCD$ 变换为平行四边形 $A'B'C'D'$.

不难验证, 错切变换 q_x, q_y 都是线性变换

以上例子中的图象变换与数的变换虽然是两种不同的现象, 但它们都是映射. 在中学解析几何里讨论了平面图形和方程的简单关系, 而在高等数学中更重于讨论图形之间的映射, 乃至于一般现象的转换(也可视为映射)和数的映射之间的关系.

读者可能从上述例子中已察觉到, \mathbb{R}^2 上的线性变换总是把直线变换为直线(在例 6 之后将给以一般的证明). 这是线性变换的特性, 也是它的一种局限性. 非线性映射(变换)则不一样, 例如

$$\sigma(x, y) = (x + y^2, x^2 + y)$$

是非线性的映射, 记 $x' = x + y^2, y' = x^2 + y$, 容易证明该映射把图 3-5 中正方形 $OABCO$ 变换为图 3-6 中斜 8 字形 $OA'B'C'O$.

例如 σ 把直线 OA ($y = 0, 0 \leq x \leq 1$) 变换为抛物线 OA' . 因为此时

$$\begin{cases} x' = x + y^2 = x + 0^2 = x \\ y' = x^2 + y = x^2 + 0 = x^2, \end{cases} \quad 0 \leq x \leq 1,$$

消去 x 即得 $y' = (x')^2, 0 \leq x' \leq 1$. 同理可证 σ 分别把直线 AB, BC, CO 变换为曲线 $A'B', B'C', C'O$.

在 \mathbb{R}^3 中也有类似上述 \mathbb{R}^2 的线性变换, 但旋转变换是向量绕过原点的直线(轴)旋转, 镜像变换的镜面是过原点的平面.

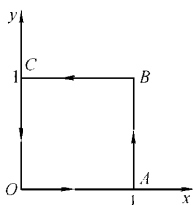


图 3-5

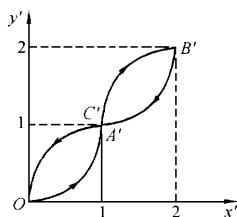


图 3-6

例 6 三维几何向量空间 \mathbf{R}^3 中的投影(内映射)变换. 例如(下面的 e_1, e_2, e_3 是 \mathbf{R}^3 的自然基)

(1) \mathbf{R}^3 到子空间 $W_1 = L(e_1)$ 的投影变换 p_1 (图 3-7)为

$$p_1(\alpha) = p_1(x_1, x_2, x_3) = (x_1, 0, 0), \quad \forall (x_1, x_2, x_3) \in \mathbf{R}^3,$$

(2) \mathbf{R}^3 到子空间 $W_2 = L(e_1, e_2)$ 的投影变换 p_2 (图 3-8)为

$$p_2(\alpha) = p_2(x_1, x_2, x_3) = (x_1, x_2, 0), \quad \forall (x_1, x_2, x_3) \in \mathbf{R}^3.$$

读者不难验证 p_1, p_2 都是 \mathbf{R}^3 上的线性变换.

这里 \mathbf{R}^3 关于镜面 W_2 的镜像变换 φ (如图 3-8)为

$$\begin{aligned} \varphi(\alpha) &= \alpha' = \overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OA} + 2\overrightarrow{AC} \\ &= \overrightarrow{OA} - 2\overrightarrow{OD} = \alpha - 2(\alpha, e_3)e_3, \end{aligned}$$

其中 e_3 是镜面 W_2 的单位法向量. 一般 \mathbf{R}^3 关于镜面 W (过原点的某个平面, 其单位法向量为 ω) 的镜像变换 φ 为

$$\varphi(\alpha) = \alpha - 2(\alpha, \omega)\omega, \quad \forall \alpha \in \mathbf{R}^3.$$

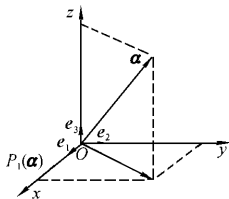


图 3-7

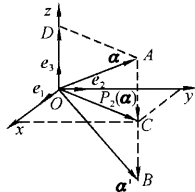


图 3-8

下面证明 \mathbf{R}^3 (或 \mathbf{R}^2) 中任何线性变换 σ 都把直线变为直线.

在第1章习题中已知, A, B, C 三点共线的充要条件是

$$\overrightarrow{OC} = \lambda \overrightarrow{OA} + (1 - \lambda) \overrightarrow{OB}.$$

分别记 $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ 为 α, β, γ , 它们在 σ 作用下的像为 $\sigma(\alpha) = \overrightarrow{OA'}$, $\sigma(\beta) = \overrightarrow{OB'}$, $\sigma(\gamma) = \overrightarrow{OC'}$, 又 σ 作用于上式的两边得

$$\sigma(\gamma) = \sigma(\lambda\alpha + (1 - \lambda)\beta) = \lambda\sigma(\alpha) + (1 - \lambda)\sigma(\beta),$$

即

$$\overrightarrow{OC'} = \lambda \overrightarrow{OA'} + (1 - \lambda) \overrightarrow{OB'},$$

所以 A', B', C' 三点也共线(见图3-9), 这就证明了线性变换 σ 把 A, B, C 三点所在的直线变成了 A', B', C' 三点所在的直线. 这也正是满足(3-1)式的映射 σ 称为线性映射的理由.

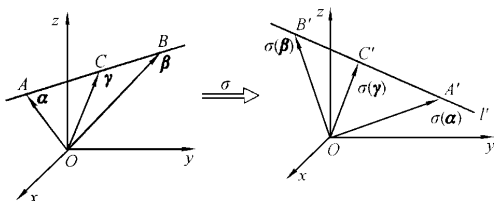


图 3-9

反过来, \mathbf{R}^3 中把直线变为直线的变换并不都是线性的, 例如 \mathbf{R}^3 中向量的平移变换 $\tau(\alpha) = \alpha + \alpha_0$ (其中 α_0 是 \mathbf{R}^3 非零常向量), 便不是线性变换(验证留给读者).

例7 定义 \mathbf{R}^3 到 \mathbf{R}^2 的两个映射 σ 和 τ 为:

$$\sigma(x_1, x_2, x_3) = (x_1 + x_2, x_2 - x_3),$$

$$\tau(x_1, x_2, x_3) = (x_1 x_2, x_2 - x_3).$$

容易验证: σ 是线性的(一般, 当 σ 的象的每个分量为 x_1, x_2, x_3 的线性组合, 即 x_1, x_2, x_3 的一次式时, σ 是线性的); τ 不是线性的. 因为 $\forall \alpha = (a_1, a_2, a_3)$, $\beta = (b_1, b_2, b_3) \in \mathbf{R}^3$, 但

$$\begin{aligned} \tau(\alpha + \beta) &= \tau(a_1 + b_1, a_2 + b_2, a_3 + b_3) \\ &= ((a_1 + b_1)(a_2 + b_2), (a_2 + b_2) - (a_3 + b_3)) \\ &= (a_1 a_2, a_2 - a_3) + (b_1 b_2, b_2 - b_3) + (a_1 b_2 + b_1 a_2, 0) \\ &= \tau(\alpha) + \tau(\beta) + (a_1 b_2 + b_1 a_2, 0) \\ &\neq \tau(\alpha) + \tau(\beta). \end{aligned}$$

例8 设 $a_1, \dots, a_n \in \mathbf{R}$, 如果 $\forall X = (x_1, \dots, x_n) \in \mathbf{R}^n$, 定义

$$f(\mathbf{X}) = a_1x_1 + a_2x_2 + \dots + a_nx_n,$$

则 f 是 \mathbf{R}^n 到 \mathbf{R} 的线性映射(也叫做 f 是 \mathbf{R}^n 上的线性函数).

例 9 设 V 是复数集 \mathbf{C} 在实数域 \mathbf{R} 上的线性空间, 定义

$$\sigma(z) = \bar{z}, \quad \forall z \in \mathbf{C},$$

则 σ 是 V 上的线性变换, 但如果 V 是复数集 \mathbf{C} 在复数域 \mathbf{C} 上的线性空间, 则上述 σ 不是线性的, 因为当 λ 的虚部不为零时,

$$\sigma(\lambda z) = \overline{\lambda z} = \bar{\lambda} \bar{z} = \bar{\lambda} \sigma(z) \neq \lambda \sigma(z).$$

线性映射 $\sigma: V_1(F) \rightarrow V_2(F)$ 有以下简单性质:

$$\begin{aligned} (i) \quad & \sigma(\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m) \\ &= \lambda_1 \sigma(\alpha_1) + \lambda_2 \sigma(\alpha_2) + \dots + \lambda_m \sigma(\alpha_m) \end{aligned} \quad (3-4)$$

(其中 $\alpha_i \in V_1, \lambda_i \in F, i = 1, 2, \dots, m$), 用数学归纳法易证上式成立. 这个性质表明, 线性映射保持线性运算关系不变, 也就是说线性映射把 $\alpha_1, \alpha_2, \dots, \alpha_m$ 的线性组合映射为 $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_m)$ 的同样的线性组合.

(ii) 若 $V_1(F)$ 中的向量组 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性相关, 则它们的像 $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_m)$ 也线性相关.

事实上, 由不全为零的数 $\lambda_1, \lambda_2, \dots, \lambda_m \in F$ 使得

$$\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \dots + \lambda_m \alpha_m = \mathbf{0}_1 \quad (1)$$

以及(1)式两边在 σ 作用下的像的等式

$$\lambda_1 \sigma(\alpha_1) + \lambda_2 \sigma(\alpha_2) + \dots + \lambda_m \sigma(\alpha_m) = \sigma(\mathbf{0}_1) = \mathbf{0}_2 \quad (2)$$

即得性质(ii). 但(ii)的逆命题不成立, 即 $\alpha_1, \alpha_2, \dots, \alpha_m$ 线性无关, 其像 $\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_m)$ 可能线性相关. 如例 6 的投影变换 p_2 把 $\alpha_1 = (x_1, y_1, z_1)$ 和 $\alpha_2 = (x_1, y_1, z_2)$ 变换为

$$p_2(\alpha_1) = p_2(\alpha_2) = (x_1, y_1, 0).$$

显然, 当 $z_1 \neq z_2$ 时, α_1, α_2 线性无关, 而其像 $p_2(\alpha_1), p_2(\alpha_2)$ 线性相关.

3.2 线性映射的像和核

定义 3.2 设 σ 是线性空间 $V_1(F)$ 到 $V_2(F)$ 的线性映射, V_1 的所有元素在 σ 下的像所组成的集合

$$\sigma(V_1) = \{\beta \mid \beta = \sigma(\alpha), \alpha \in V_1\}$$

称为 σ 的像(或称 σ 的值域), V_2 的零元 $\mathbf{0}_2$ 在 σ 下的完全原像

$$\sigma^{-1}(\mathbf{0}_2) = \{\alpha \mid \sigma(\alpha) = \mathbf{0}_2, \alpha \in V_1\}$$

称为 σ 的核 $\mathcal{A}(V_1)$ 和 $\sigma^{-1}(\mathbf{0}_2)$ 也常记作 $\text{Im}\sigma$ 和 $\text{Ker}\sigma$.

例1 \mathbf{R}^2 上的旋转变换 r_θ 和镜像变换 φ 的像(值域)都是 \mathbf{R}^2 自身, 它们的核都只含一个零向量 $\{\mathbf{0}\}$.

例2 节3.1例6的 \mathbf{R}^3 上的投影变换 p_1, p_2 的像和核为

$$p_1(\mathbf{R}^3) = \text{Im} p_1 = L(e_1), \quad p_1^{-1}(\mathbf{0}) = \text{Ker} p_1 = L(e_2, e_3),$$

$$p_2(\mathbf{R}^3) = \text{Im} p_2 = L(e_1, e_2), \quad p_2^{-1}(\mathbf{0}) = \text{Ker} p_2 = L(e_3).$$

关于 σ 的像和核, 有以下结论.

(1) $\mathcal{A}(V_1)$ 是 V_2 的一个子空间. 因为: 由 $\mathcal{A}(\mathbf{0}_1) = \mathbf{0}_2$, 可知 $\mathcal{A}(V_1)$ 是非空集, 而且 $\forall \beta_1, \beta_2 \in \mathcal{A}(V_1), \exists \alpha_1, \alpha_2 \in V_1$ 使得 $\mathcal{A}(\alpha_1) = \beta_1$ 和 $\mathcal{A}(\alpha_2) = \beta_2$, 于是 $\forall \lambda_1, \lambda_2 \in F$ 有

$$\lambda_1 \beta_1 + \lambda_2 \beta_2 = \lambda_1 \mathcal{A}(\alpha_1) + \lambda_2 \mathcal{A}(\alpha_2) = \mathcal{A}(\lambda_1 \alpha_1 + \lambda_2 \alpha_2) \in \mathcal{A}(V_1).$$

(2) $\text{Ker}\sigma$ 是 V_1 的一个子空间. 因为: $\sigma^{-1}(\mathbf{0}_2)$ 不是空集(事实上 $\mathbf{0}_1 \in \sigma^{-1}(\mathbf{0}_2)$), 而且 $\forall \alpha_1, \alpha_2 \in \sigma^{-1}(\mathbf{0}_2)$ 和 $\forall \lambda_1, \lambda_2 \in F$ 有

$$\mathcal{A}(\lambda_1 \alpha_1 + \lambda_2 \alpha_2) = \lambda_1 \mathcal{A}(\alpha_1) + \lambda_2 \mathcal{A}(\alpha_2) = \lambda_1 \mathbf{0}_2 + \lambda_2 \mathbf{0}_2 = \mathbf{0}_2,$$

所以 $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 \in \sigma^{-1}(\mathbf{0}_2)$.

同样可证: 如果 W_1, W_2 分别是 V_1, V_2 的子空间, 则 $\mathcal{A}(W_1), \sigma^{-1}(W_2)$ 也分别是 V_2 和 V_1 的子空间.

定理3.1 线性映射 $\sigma: V_1 \rightarrow V_2$ 是单射 $\Leftrightarrow \sigma^{-1}(\mathbf{0}_2) = \{\mathbf{0}_1\}$.

证 如果 σ 是单射, 则由 $\mathcal{A}(\alpha) = \mathbf{0}_2 = \mathcal{A}(\mathbf{0}_1)$, 可推出 $\alpha = \mathbf{0}_1$, 故 $\sigma^{-1}(\mathbf{0}_2) = \{\mathbf{0}_1\}$; 反之, 如果 $\sigma^{-1}(\mathbf{0}_2) = \{\mathbf{0}_1\}$ 则由 $\mathcal{A}(\alpha_1) = \mathcal{A}(\alpha_2)$ 即 $\mathcal{A}(\alpha_1) - \mathcal{A}(\alpha_2) = \mathcal{A}(\alpha_1 - \alpha_2) = \mathbf{0}_2$, 可得 $\alpha_1 - \alpha_2 = \mathbf{0}_1$, 即 $\alpha_1 = \alpha_2$, 故 σ 是单射.

例3 已知 \mathbf{R}^3 到 \mathbf{R}^2 的映射 σ 为: $\mathcal{A}(x_1, x_2, x_3) = (x_1 + x_2, x_2 - x_3)$, 试求 σ 的像和核.

解

$$\begin{aligned} \mathcal{A}(x_1, x_2, x_3) &= (x_1 + x_2, x_2 - x_3) \\ &= x_1(1, 0) + x_2(1, 1) + x_3(0, -1). \end{aligned} \quad \textcircled{1}$$

记 $\alpha_1 = (1, 0), \alpha_2 = (1, 1), \alpha_3 = (0, -1)$, 由①式可知 σ 的像是 $\{\alpha_1, \alpha_2, \alpha_3\}$ 的线性扩张, $\{\alpha_1, \alpha_3\}$ 显然是 $\{\alpha_1, \alpha_2, \alpha_3\}$ 的极大线性无关组, 所以

$$\mathcal{A}(\mathbf{R}^3) = L(\alpha_1, \alpha_3) = \mathbf{R}^2,$$

即 σ 是 \mathbf{R}^3 到 \mathbf{R}^2 的满射. 再由 $\text{Ker}\sigma = \{\mathcal{A}(x_1, x_2, x_3) | \mathcal{A}(x_1, x_2, x_3) = \mathbf{0}\}$, 可知 σ 的核是齐次线性方程组 $x_1 + x_2 = 0, x_2 - x_3 = 0$ 的解空间 S , 易见 $S = L(\alpha)$, 其中 $\alpha = (-1, 1, 1)$, 故 $\text{Ker}\sigma = L(\alpha)$.

3.3 线性映射的运算 空间 $L(V_1, V_2)$

线性映射乘法运算的定义与一般映射的乘法一样. 需要指出的是线性映射的乘积和可逆线性映射的逆映射仍是线性映射, 因为:

如果 σ_1 和 σ_2 分别是线性空间 $V_1(F)$ 到 $V_2(F)$ 和 $V_2(F)$ 到 $V_3(F)$ 的线性映射, 那么 $\sigma_2\sigma_1$ 是 $V_1(F)$ 到 $V_3(F)$ 的映射, 而且 $\forall \alpha, \beta \in V_1$ 和 $\forall \lambda, \mu \in F$, 有

$$\begin{aligned}(\sigma_2\sigma_1)(\lambda\alpha + \mu\beta) &= \sigma_2(\sigma_1(\lambda\alpha + \mu\beta)) = \sigma_2(\lambda\sigma_1(\alpha) + \mu\sigma_1(\beta)) \\&= \lambda(\sigma_2(\sigma_1(\alpha))) + \mu(\sigma_2(\sigma_1(\beta))) \\&= \lambda(\sigma_2\sigma_1)(\alpha) + \mu(\sigma_2\sigma_1)(\beta),\end{aligned}$$

所以 $\sigma_2\sigma_1$ 是 V_1 到 V_3 的线性映射.

如果可逆线性映射 $\sigma: V_1 \rightarrow V_2$ 的逆映射为 $\sigma^{-1}: V_2 \rightarrow V_1$, 那么 $\sigma^{-1}\sigma = I_{V_1}$ 且 $\sigma\sigma^{-1} = I_{V_2}$, 于是 $\forall \beta_1, \beta_2 \in V_2$ 和 $\forall \lambda_1, \lambda_2 \in F$, 有

$$\begin{aligned}\sigma^{-1}(\lambda_1\beta_1 + \lambda_2\beta_2) &= \sigma^{-1}[\lambda_1(\sigma\sigma^{-1})(\beta_1) + \lambda_2(\sigma\sigma^{-1})(\beta_2)] \\&= \sigma^{-1}[\lambda_1\sigma^{-1}(\beta_1) + \lambda_2\sigma^{-1}(\beta_2)] \\&= \lambda_1\sigma^{-1}(\beta_1) + \lambda_2\sigma^{-1}(\beta_2),\end{aligned}$$

所以 σ^{-1} 是线性的.

我们把线性空间 $V_1(F)$ 到 $V_2(F)$ 的所有线性映射组成的集合记作 $L(V_1, V_2)$. 下面说明这个集合也是一个线性空间. 为此要定义 $L(V_1, V_2)$ 上的加法运算和 $F \times L(V_1, V_2)$ 上的数乘运算.

定义 3.3 设 $\sigma, \tau \in L(V_1, V_2)$, 规定 σ 与 τ 之和 $\sigma + \tau$ 及 λ 与 σ 之数量乘积 $\lambda\sigma$ 分别为

$$\begin{aligned}(\sigma + \tau)(\alpha) &= \sigma(\alpha) + \tau(\alpha), \quad \alpha \in V_1, \\(\lambda\sigma)(\alpha) &= \lambda(\sigma(\alpha)), \quad \alpha \in V_1.\end{aligned}$$

上述定义的 $(\sigma + \tau)(\lambda\sigma) \in L(V_1, V_2)$. 事实上, $\forall \alpha_1, \alpha_2 \in V_1$, 和 $\forall \lambda_1, \lambda_2 \in F$, 有

$$\begin{aligned}(\sigma + \tau)(\lambda_1\alpha_1 + \lambda_2\alpha_2) &= \sigma(\lambda_1\alpha_1 + \lambda_2\alpha_2) + \tau(\lambda_1\alpha_1 + \lambda_2\alpha_2) \\&= \lambda_1\sigma(\alpha_1) + \lambda_2\sigma(\alpha_2) + \lambda_1\tau(\alpha_1) + \lambda_2\tau(\alpha_2) \\&= \lambda_1(\sigma + \tau)(\alpha_1) + \lambda_2(\sigma + \tau)(\alpha_2), \\(\lambda\sigma)(\lambda_1\alpha_1 + \lambda_2\alpha_2) &= \lambda(\sigma(\lambda_1\alpha_1 + \lambda_2\alpha_2)) = \lambda(\lambda_1\sigma(\alpha_1) + \lambda_2\sigma(\alpha_2))\end{aligned}$$

$$= \lambda_1(\lambda\sigma(\alpha_1)) + \lambda_2(\lambda\sigma(\alpha_2))$$

$L(V_1, V_2)$ 对上述定义的加法和数乘运算构成域 F 上的一个线性空间, 因为: $L(V_1, V_2)$ 中的加法显然满足交换律和结合律, 加法的单位元(即零元)是 V_1 到 V_2 的零映射 θ , 每个元 σ 的逆元为 $(-\sigma)$, 即

$$(-\sigma)(\alpha) = -\sigma(\alpha), \quad \forall \alpha \in V_1.$$

容易验证 $(-\sigma) \in L(V_1, V_2)$. 因此 $L(V_1, V_2)$ 关于线性映射的加法构成一个交换群(加法群).

至于数乘运算, 读者不难验证它满足以下 4 条性质:

$$1\sigma = \sigma, \quad \lambda(\mu\sigma) = (\lambda\mu)\sigma,$$

$$(\lambda + \mu)\sigma = \lambda\sigma + \mu\sigma, \quad \lambda(\sigma + \tau) = \lambda\sigma + \lambda\tau.$$

其中 1 是域 F 的乘法单位元, $\lambda, \mu \in F$; $\sigma, \tau \in L(V_1, V_2)$.

此外, 乘法对加法也满足左、右分配律(验证留给读者)

$$\varphi(\sigma + \tau) = \varphi\sigma + \varphi\tau,$$

$$(\sigma + \tau)\Psi = \sigma\Psi + \tau\Psi.$$

其中 $\sigma, \tau \in L(V_1, V_2)$, $\varphi \in L(V_2, V_3)$, $\Psi \in L(V_4, V_1)$. 于是由线性空间 V 到自身的所有线性映射组成的集合 $L(V, V)$ 关于加法和乘法构成的代数系统 $L(V, V); +, \circ$ 是一个含么环, 其乘法单位元是恒等映射. 在 $L(V, V)$ 中全体可逆线性映射组成的集合, 对映射的乘法构成一个群, 当 $\dim V = n$ 时, 叫做 n 阶线性群.

3.4 有限维线性空间的线性映射 线性映射的秩

设 $\sigma \in L(V_1, V_2)$, $V_1(F)$ 和 $V_2(F)$ 的维数分别为 n 和 m , $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 V_1 的基, 那么 V_1 中任一向量 ξ 可表示为

$$\xi = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n,$$

于是 ξ 在 σ 下的像为

$$\sigma(\xi) = x_1\sigma(\alpha_1) + x_2\sigma(\alpha_2) + \dots + x_n\sigma(\alpha_n).$$

因此, 知道了 σ 关于基 B 的像, 就可求得 σ 对任一向量 ξ 的像. 所以, 有限维线性空间的线性映射 σ 被它在基上的像所确定. 从以上两式可见, σ 的像(值域) $\sigma(V_1)$ 是基 B 在 σ 下的像 $\sigma(B) = \{\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)\}$ 的线性扩张, 即

$$\sigma(V_1) = L(\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)). \quad (3-5)$$

因此, $\sigma(V_1)$ 的维数等于基像组 $\sigma(B)$ 的秩, 并称之为 σ 的秩.

定义 3.4 (线性映射的秩) 设 $\sigma \in L(V_1, V_2)$, 如果 $\sigma(V_1)$ 是 V_2 的有限维子空间, 则 $\sigma(V_1)$ 的维数称为 σ 的秩, 记作 $r(\sigma)$, 即

$$r(\sigma) = \dim \sigma(V_1). \quad (3-6)$$

显然,有限维线性空间 V_1 到 V_2 的线性映射的秩都 $\leq \dim V_1$.

有限维线性空间的线性映射的像和核有如下的维数公式:

定理 3.2 设 $\sigma \in L(V_1, V_2)$, 如果 $\dim(V_1) = n$ 则

$$r(\sigma) + \dim(\text{Ker} \sigma) = n. \quad (3-7)$$

证 设 $\dim(\text{Ker} \sigma) = k$, $B_1 = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 是 $\text{Ker} \sigma$ 的一个基, 把 B_1 扩充为 V_1 的基 $B = \{\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$. 由于 $\sigma(\alpha_i) = 0$ ($i = 1, \dots, k$), 所以 σ 的值域 (像)

$$\sigma(V_1) = L(\sigma(B)) = L(\sigma(\alpha_{k+1}), \dots, \sigma(\alpha_n)).$$

因此, 只需证明 $r(\sigma) = n - k$, 即 $\{\sigma(\alpha_{k+1}), \dots, \sigma(\alpha_n)\}$ 线性无关. 设

$$c_{k+1}\sigma(\alpha_{k+1}) + \dots + c_n\sigma(\alpha_n) = 0,$$

即

$$\sigma(c_{k+1}\alpha_{k+1} + \dots + c_n\alpha_n) = 0.$$

所以 $c_{k+1}\alpha_{k+1} + \dots + c_n\alpha_n (\in \text{Ker} \sigma)$ 可被 B_1 线性表示, 于是有

$$c_1\alpha_1 + \dots + c_k\alpha_k - c_{k+1}\alpha_{k+1} - \dots - c_n\alpha_n = 0,$$

从而 $c_1 = \dots = c_k = c_{k+1} = \dots = c_n = 0$, 故 $\sigma(\alpha_{k+1}), \dots, \sigma(\alpha_n)$ 线性无关.

由定理 3.2 立即可得 $r(\sigma) = \dim V_1 = n$ 的充要条件是 $\text{Ker} \sigma = \{0\}$, 即 σ 为单射. 如果 V_2 是有限维线性空间, 根据定义 3.4 即得 $r(\sigma) = \dim V_2$ 的充要条件是 $\sigma(V_1) = V_2$, 即 σ 是满射.

综上所述, 我们可得下面的定理.

定理 3.3 设 $\sigma \in L(V_1, V_2)$, 如果 V_1 和 V_2 都是 n 维线性空间, 则下列命题等价:

(1) $r(\sigma) = n$ (或说 σ 满秩) (2) σ 是单射 (3) σ 是满射 (4) σ 是可逆线性映射.

定理 3.3 自然适用于有限维线性空间到自身的线性变换.

例 1 节 3.1 例 2 和例 3 的 \mathbb{R}^2 上的旋转变换 r_θ 和镜面反射 φ 的秩都为 2, 核都是零子空间 $\{0\}$, 它们都是可逆线性映射.

节 3.1 例 6 的投影变换 p_1 和 p_2 的秩和核为 秩(p_1) = 1, $\text{Ker} p_1 = W_1^\perp$, $\dim(\text{Ker} p_1) = 2$, 秩(p_2) = 2, $\text{Ker} p_2 = W_2^\perp$, $\dim(\text{Ker} p_2) = 1$. p_1 和 p_2 都是不可逆的线性映射.

例 2 设 $\sigma \in L(\mathbb{R}^3, \mathbb{R}^3)$, 已知 $\sigma(x_1, x_2, x_3) = (x_1 + x_2, x_2 - x_3, x_1 + x_3)$, 求秩 σ 及 $\text{Ker} \sigma$ 的维数.

解 由 (3-5) 式知, σ 的象 (值域) 为 $\sigma(\mathbb{R}^3) = L(\sigma(\beta_1), \sigma(\beta_2), \sigma(\beta_3))$,

其中 $\{\beta_1, \beta_2, \beta_3\}$ 是 \mathbf{R}^3 的任一组基, 取自然基 $\{e_1, e_2, e_3\}$ 最简单, 得

$$\alpha(\mathbf{R}^3) = L(\alpha(e_1), \alpha(e_2), \alpha(e_3)),$$

其中 $\alpha(e_1) = \alpha(1, 0, 0) = (1, 0, 1)$, $\alpha(e_2) = \alpha(0, 1, 0) = (1, 1, 0)$, $\alpha(e_3) = \alpha(0, 0, 1) = (0, -1, 1)$. 易见 $\alpha(e_3) = \alpha(e_1) - \alpha(e_2)$, 所以 $\alpha(\mathbf{R}^3) = L(\alpha(e_1), \alpha(e_2))$, 因此 $\text{秩}(\sigma) = \dim \alpha(\mathbf{R}^3) = 2$.

$\text{Ker} \sigma$ 是齐次线性方程组 $x_1 + x_2 = 0$, $x_2 - x_3 = 0$, $x_1 + x_3 = 0$ 的解空间 S , 用高斯消元法易得 $\text{Ker} \sigma = S = L(\alpha)$, 其中 $\alpha = (-1, 1, 1)$, 故 $\dim(\text{Ker} \sigma) = 1$.

这里附带说一下, 用节 3.2 例题提供的方法求得的 $\alpha(\mathbf{R}^3)$ 恰好就是 $L(\alpha(e_1), \alpha(e_2), \alpha(e_3))$.

下面讨论两个线性映射之和与乘积的秩.

定理 3.4 设 V_1, V_2, V_3 分别是 m, n, s 维线性空间, $\sigma \in L(V_1, V_2)$, $\tau \in L(V_2, V_3)$ 则

$$\text{秩}(\sigma) + \text{秩}(\tau) - n \leq \text{秩}(\tau\sigma) \leq \min(\text{秩}(\sigma), \text{秩}(\tau)). \quad (3-8)$$

证 先证右边. 由于 $\sigma(V_1) \subset V_2$, 所以 $(\tau\sigma)(V_1) \subset \tau(V_2)$, 如图 3-10 所示. 因此

$$\dim(\tau\sigma)(V_1) \leq \dim \tau(V_2), \text{ 即 } \text{秩}(\tau\sigma) \leq \text{秩}(\tau)$$

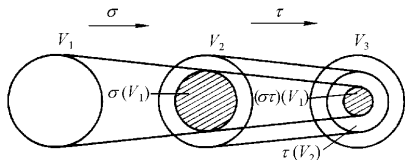


图 3-10

又因为 $(\tau\sigma)(V_1) = \tau(\sigma(V_1))$, 所以又有

$$\dim(\tau\sigma)(V_1) \leq \dim \sigma(V_1), \text{ 即 } \text{秩}(\tau\sigma) \leq \text{秩}(\sigma).$$

再证左边: 由 (3-7),

$$\text{秩}(\tau) + \dim(\text{Ker} \tau) = n, \quad \text{秩}(\tau\sigma) + \dim(\text{Ker}(\tau\sigma)) = m.$$

又

$$\dim(\text{Ker}(\tau\sigma)) \leq \dim(\text{Ker} \tau),$$

所以

$$m - \text{秩}(\tau\sigma) = \dim(\text{Ker}(\tau\sigma)) \leq \dim(\text{Ker} \tau).$$

代入前面的式子, 得到

$$\dim(\text{Ker} \tau) = n - \text{秩}(\tau) \geq m - \text{秩}(\tau\sigma),$$

即

$$\begin{aligned} r(\tau\sigma) &\geq m + r(\tau) - n \\ &\geq r(\sigma) + r(\tau) - n. \end{aligned}$$

推论 在定理 3.4 中, 如果 $\dim(V_1) = \dim(V_2) = n$, 秩 $(\sigma) = n$ (即 σ 是可逆线性映射), 则

$$\text{秩}(\tau\sigma) = \text{秩}(\tau). \quad (3-9)$$

事实上, 由 $\sigma(V_1) = V_2$, 即得 $(\tau\sigma)(V_1) = \tau(V_2)$, 故 (3-9) 式成立.

如果 $\dim(V) = n$, $\sigma, \tau \in L(V, V)$, 那么定理 3.4 及其推论也成立.

定理 3.5 设 V_1 是有限维线性空间, $\sigma, \tau \in L(V_1, V_2)$, 则

$$\text{秩}(\sigma + \tau) \leq \text{秩}(\sigma) + \text{秩}(\tau). \quad (3-10)$$

证 由于 $\forall \beta \in (\sigma + \tau)(V_1), \exists \alpha \in V_1$, 使 $\beta = (\sigma + \tau)(\alpha) = \sigma(\alpha) + \tau(\alpha) \in \sigma(V_1) + \tau(V_1)$, 所以

$$(\sigma + \tau)(V_1) \subset \sigma(V_1) + \tau(V_1),$$

因此

$$\begin{aligned} \dim(\sigma + \tau)(V_1) &\leq \dim[\sigma(V_1) + \tau(V_1)] \\ &\leq \dim\sigma(V_1) + \dim\tau(V_1), \end{aligned}$$

故 (3-10) 式成立.

例如 \mathbf{R}^3 分别到子空间 $W_1 = L(e_1), W_2 = L(e_1, e_2)$ 的投影变换 p_1, p_2 , 其值域: $p_1(\mathbf{R}^3) = W_1, p_2(\mathbf{R}^3) = W_2$, 而

$$(p_1 + p_2)(\mathbf{R}^3) = W_2, \quad (p_1 p_2)(\mathbf{R}^3) = W_1.$$

相应地

$$\begin{aligned} \text{秩}(p_1 + p_2) &= 2 < \text{秩}(p_1) + \text{秩}(p_2) = 3, \\ \text{秩}(p_1 p_2) &= 1 \leq \min(\text{秩}(p_1), \text{秩}(p_2)) = 1. \end{aligned}$$

关于线性映射与其基像的关系, 我们还要指出以下事实:

有限维空间的线性映射不仅被它在基上的像所确定, 而且被基像唯一确定, 也就是说两个不同的线性映射作用于基上的像不可能完全相同, 即如果 $\sigma, \tau \in L(V_1, V_2)$ 作用于 V_1 的基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 上的像完全相同, 则必有 $\sigma = \tau$. 事实上, 如果

$$\sigma(\alpha_i) = \tau(\alpha_i), \quad i = 1, 2, \dots, n$$

则 $\forall \xi = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n \in V_1$ 都有

$$\begin{aligned} \sigma(\xi) &= x_1\sigma(\alpha_1) + x_2\sigma(\alpha_2) + \dots + x_n\sigma(\alpha_n) \\ &= x_1\tau(\alpha_1) + x_2\tau(\alpha_2) + \dots + x_n\tau(\alpha_n) \\ &= \tau(x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n) = \tau(\xi). \end{aligned}$$

故 $\tau = \sigma$, 这就证明了线性映射被它在基上的像所唯一确定.

进而又有: 如果 $\dim V_1 = n$, 则 V_2 中任意 n 个向量 $\beta_1, \beta_2, \dots, \beta_n$ 都可以是某个 $\sigma \in L(V_1, V_2)$ 作用于 V_1 的一组基的像, 即有

定理 3.6 设 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 $V_1(F)$ 的基, $S = \{\beta_1, \beta_2, \dots, \beta_n\}$ 是 $V_2(F)$ 中

任意 n 个向量, 则存在唯一的 $\sigma \in L(V_1, V_2)$, 使得

$$\sigma(\alpha_i) = \beta_i, \quad i = 1, 2, \dots, m. \quad (3-11)$$

证 先定义满足 (3-11) 式的从 V_1 到 V_2 的一个映射 σ , 即

$$\forall \xi = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n \in V_1,$$

规定

$$\sigma(\xi) = x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n.$$

事实上, 这个映射也是线性的. 因为 $\forall \xi_1 = b_1\alpha_1 + b_2\alpha_2 + \dots + b_n\alpha_n, \xi_2 = c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n$ 和 $\forall \lambda_1, \lambda_2 \in F$ 有

$$\begin{aligned} \sigma(\lambda_1\xi_1 + \lambda_2\xi_2) &= \sigma((\lambda_1b_1 + \lambda_2c_1)\alpha_1 + (\lambda_1b_2 + \lambda_2c_2)\alpha_2 + \dots + (\lambda_1b_n + \lambda_2c_n)\alpha_n) \\ &= (\lambda_1b_1 + \lambda_2c_1)\beta_1 + (\lambda_1b_2 + \lambda_2c_2)\beta_2 + \dots + (\lambda_1b_n + \lambda_2c_n)\beta_n \\ &= \lambda_1(b_1\beta_1 + b_2\beta_2 + \dots + b_n\beta_n) + \lambda_2(c_1\beta_1 + c_2\beta_2 + \dots + c_n\beta_n) \\ &= \lambda_1\sigma(\xi_1) + \lambda_2\sigma(\xi_2), \end{aligned}$$

这就证明了 σ 的线性性. σ 的唯一性前已论述.

* * 最后, 我们来看线性空间 $L(V_1, V_2)$ 的维数是多少.

定理 3.7 若 $V_1(F)$ 和 $V_2(F)$ 分别是 n 和 m 维线性空间, 则空间 $L(V_1, V_2)$ 的维数为 nm .

证 我们对 $n = m = 2$ 的情形证明. 设 V_1 的基为 $B_1 = \{\alpha_1, \alpha_2\}$, V_2 的基为 $B_2 = \{\beta_1, \beta_2\}$. 根据定理 3.6 定义映射 $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22} \in L(V_1, V_2)$ 为

$$\begin{aligned} \sigma_{11}(\alpha_1) &= \beta_1, \quad \sigma_{11}(\alpha_2) = 0_2; \quad \sigma_{12}(\alpha_1) = 0_2, \quad \sigma_{12}(\alpha_2) = \beta_1, \\ \sigma_{21}(\alpha_1) &= \beta_2, \quad \sigma_{21}(\alpha_2) = 0_2; \quad \sigma_{22}(\alpha_1) = 0_2, \quad \sigma_{22}(\alpha_2) = \beta_2. \end{aligned}$$

我们要证 $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}$ 是 $L(V_1, V_2)$ 的基. 先证它们线性无关. 设

$$c_{11}\sigma_{11} + c_{12}\sigma_{12} + c_{21}\sigma_{21} + c_{22}\sigma_{22} = \theta,$$

其中 θ 是空间 $L(V_1, V_2)$ 的零元 (即零映射), 于是

$$\begin{aligned} \theta(\alpha_1) &= c_{11}\sigma_{11}(\alpha_1) + c_{12}\sigma_{12}(\alpha_1) + c_{21}\sigma_{21}(\alpha_1) + c_{22}\sigma_{22}(\alpha_1) \\ &= c_{11}\beta_1 + c_{21}\beta_2 = 0_2. \end{aligned}$$

因为 $B_2 = \{\beta_1, \beta_2\}$ 线性无关, 所以 $c_{11} = c_{21} = 0$. 同理, 由 $\theta(\alpha_2) = 0_2$ 又得 $c_{12} = c_{22} = 0$. 因此 $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}$ 线性无关.

再证明任一 $\sigma \in L(V_1, V_2)$ 可被 $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}$ 线性表示. 由于 $\forall \alpha \in V_1, \sigma(\alpha) \in V_2$, 所以

$$\begin{aligned} \sigma(\alpha_1) &= b_{11}\beta_1 + b_{21}\beta_2, \\ \sigma(\alpha_2) &= b_{12}\beta_1 + b_{22}\beta_2. \end{aligned}$$

令

$$\tau = b_{11}\sigma_{11} + b_{12}\sigma_{12} + b_{21}\sigma_{21} + b_{22}\sigma_{22},$$

显然 $\tau \in L(V_1, V_2)$, 且

$$\begin{aligned}\tau(\alpha_1) &= b_{11}\sigma_{11}(\alpha_1) + b_{12}\sigma_{12}(\alpha_1) + b_{21}\sigma_{21}(\alpha_1) + b_{22}\sigma_{22}(\alpha_1) \\ &= b_{11}\beta_1 + b_{21}\beta_2 = \sigma(\alpha_1).\end{aligned}$$

同理, $\tau(\alpha_2) = \sigma(\alpha_2)$. 所以 τ 与 σ 关于基 $B_1 = \{\alpha_1, \alpha_2\}$ 的像相同. 因此 $\sigma = \tau$, 即 σ 可由 $\{\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}\}$ 线性表示, 从而 $\{\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}\}$ 是 $L(V_1, V_2)$ 的一组基, 故 $\dim L(V_1, V_2) = 2 \times 2 = 4$.

特别地, 当 $\dim V = n$ 时, $\dim L(V, V) = n^2$.

\mathbf{R}^n 上的线性函数是 \mathbf{R}^n 到 \mathbf{R}^1 的线性映射, 所以空间 $L(\mathbf{R}^n, \mathbf{R})$ 与 \mathbf{R}^n 一样, 也是 n 维的. 同样, 如果线性空间 $V(\mathbf{R})$ 是 n 维的, 则 $V^* = L(V, \mathbf{R})$ 也是 n 维的. 我们把 V^* 叫做 V 的对偶空间, 由定理 3.7 的证明可知:

如果 $B = \{e_1, e_2, \dots, e_n\}$ 是 V 的一个基, 则 $B^* = \{\overline{e_1}, \overline{e_2}, \dots, \overline{e_n}\}$ 是 V^* 的一个基, 其中

$$\begin{aligned}\overline{e_i}(e_j) &= \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases} \quad j = 1, 2, \dots, n, \\ &= \delta_{ij}, \quad i = 1, 2, \dots, n.\end{aligned}$$

这里 V^* 的基 B^* 叫做 V 的基 B 的对偶基. * *

3.5 线性空间的同构

在刻画集合中二元素“相同”时, 我们引进了“等价关系”的概念. 对于各种各样的线性空间, 它们的元素不相同, 具体的运算也不一样, 那么何谓两个线性空间的结构是“相同”的呢? 为了弄清这一点, 也要引进一种“等价关系”, 然后就可把对各种 n 维线性空间的研究归结为对 F^n 的研究.

我们在第二章 2.5 中讲过, 研究任何 n 维线性空间 $V(F)$ 都可通过基和坐标归结为研究 n 维向量空间 F^n . 因为, 给定了 $V(F)$ 的基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, 在 $V(F)$ 与 F^n 之间存在一个一一对应的映射 (即双射) σ , 即 $\forall \xi = x_1\alpha_1 + x_2\alpha_2 + \dots + x_n\alpha_n \in V(F)$,

$$\sigma(\xi) = X = (x_1, x_2, \dots, x_n) \in F^n, \quad (3-12)$$

使得 $V(F)$ 中的任一向量 ξ 一一对应地映射为 ξ 在基 B 下的坐标向量 X , 而且这个映射 σ 也是线性的, 因为: $\forall \xi_1, \xi_2 \in V(F)$,

$$\xi_1 = a_1\alpha_1 + a_2\alpha_2 + \dots + a_n\alpha_n, \quad \xi_2 = b_1\alpha_1 + b_2\alpha_2 + \dots + b_n\alpha_n,$$

和 $\forall \lambda, \mu \in F$, 都有

$$\begin{aligned}\sigma(\lambda\xi_1 + \mu\xi_2) &= \sigma((\lambda a_1 + \mu b_1)\alpha_1 + \dots + (\lambda a_n + \mu b_n)\alpha_n) \\ &= (\lambda a_1 + \mu b_1, \dots, \lambda a_n + \mu b_n)\end{aligned}$$

$$\begin{aligned}
 &= \lambda(a_1 \dots a_n) + \mu(b_1 \dots b_n) \\
 &= \lambda(\xi_1) + \mu(\xi_2),
 \end{aligned}$$

所以在 $V(F)$ 与 F^n 之间, 不仅元素一一对应, 而且任意一组对应元素之间也保持线性运算关系不变, 即 $V(F)$ 中 $\xi_1, \xi_2, \dots, \xi_m$ 的某个线性组合 $c_1\xi_1 + c_2\xi_2 + \dots + c_m\xi_m$ 对应于 F^n 中相应的 $\sigma(\xi_1), \sigma(\xi_2), \dots, \sigma(\xi_m)$ 的同样的线性组合 $c_1\sigma(\xi_1) + c_2\sigma(\xi_2) + \dots + c_m\sigma(\xi_m)$. 这表明 $V(F)$ 与 F^n 的“结构”是相同的. 对 $V(F)$ 的研究可以化为对 F^n 的研究.

由此可见, 两个线性空间的结构“相同”(即同构)的本质是在它们之间存在一个线性的双射, 以保证元素一一对应, 且保持线性运算关系不变. 下面给出正式的定义.

定义 3.5 如果由线性空间 $V_1(F)$ 到 $V_2(F)$ 存在一个线性的双射 σ , 就说 $V_1(F)$ 和 $V_2(F)$ 是同构的, 记作 $V_1(F) \cong V_2(F)$. 这个 σ 叫做从 $V_1(F)$ 到 $V_2(F)$ 的一个同构映射(或同构).

由于同构映射 $\sigma: V_1 \rightarrow V_2$ 是线性的双射, 所以 σ 使 V_1 与 V_2 之间的元素一一对应, 而且 V_1 中的 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ 与 V_2 中对应的 $\{\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_m)\}$ 有相同的线性相关性. 事实上, 等式

$$c_1\alpha_1 + c_2\alpha_2 + \dots + c_m\alpha_m = \mathbf{0}_1 \quad (3-13)$$

成立, 当且仅当

$$c_1\sigma(\alpha_1) + c_2\sigma(\alpha_2) + \dots + c_m\sigma(\alpha_m) = \mathbf{0}_2 \quad (3-14)$$

成立. 因为 $(3-13) \Rightarrow (3-14)$ 是显然的; 反之, 当 $(3-14)$ 式成立时, 由该式得

$$\sigma(c_1\alpha_1 + c_2\alpha_2 + \dots + c_m\alpha_m) = \mathbf{0}_2.$$

由于 σ 是双射, 所以只有零向量才映射为零向量, 故 $(3-13)$ 式成立.

综上所述, 两个同构的线性空间, 其元素一一对应而且对应元素在各自的线性运算下的性质也一样, 因此, 本质上是相同的.

如果把线性空间看作元素, 把两个同构的线性空间看作“相同”, 那么线性空间之间的同构关系“ \cong ”自然应该是等价关系. 这是显然的, 因为

(1) $V(F) \cong V(F)$, 其同构映射为 $V(F)$ 到自身的恒等映射;

(2) 若 $V_1(F) \cong V_2(F)$, V_1 到 V_2 的同构映射为 σ , 则 $V_2(F) \cong V_1(F)$, 且 V_2 到 V_1 的同构映射为 σ^{-1} ;

(3) 若 $V_1(F) \cong V_2(F)$, $V_2(F) \cong V_3(F)$, V_1 到 V_2 和 V_2 到 V_3 的同构映射分别为 σ_1 和 σ_2 , 则 $V_1(F) \cong V_3(F)$, 且 V_1 到 V_3 的同构映射为 $\sigma_2\sigma_1$.

定理 3.8 两个有限维线性空间 $V_1(F)$ 和 $V_2(F)$ 同构的充要条件是它们的维数相等.

证 设 $\dim V_1(F) = n$, 如果 $V_1(F) \cong V_2(F)$ 则存在 V_1 到 V_2 的一个线性双射 σ , 由于 σ 为单射, 所以 $\text{Ker} \sigma = \{0_1\}$, 因此由定理 3.2 得 $\alpha(\sigma) = n$, 即 $\dim \alpha(V_1) = n$; 由于 σ 为满射, 所以 $\alpha(V_1) = V_2$, 因此 $\dim V_2(F) = n = \dim V_1(F)$, 必要性得证.

反之, 如果 $\dim V_1(F) = \dim V_2(F) = n$, 则由本节开始所述可见, 域 F 上的任何 n 维线性空间都与 F^n 同构, 因此, $V_1(F) \cong F^n$, $V_2(F) \cong F^n$, 于是由同构关系“ \cong ”的对称性和传递性即得 $V_1(F) \cong V_2(F)$, 充分性得证.

例如 $\mathbb{R}[x]_3$ 与 \mathbb{R}^3 是同构的, 把 $\mathbb{R}[x]_3$ 的基 $\{1, x, x^2\}$ 映射为 \mathbb{R}^3 的基 $\{e_1, e_2, e_3\}$ 的映射 σ , 即

$$\sigma(1) = e_1, \quad \sigma(x) = e_2, \quad \sigma(x^2) = e_3$$

是 $\mathbb{R}[x]_3$ 到 \mathbb{R}^3 的一个同构映射, 同样 $\mathbb{R}[x]_n$ 与 \mathbb{R}^n 也是同构的.

再如 m 维线性空间 $V(\mathbb{R})$ 与 $V^* = L(V, \mathbb{R})$ 也是同构的 (因为 $\dim V^* = n$).

评注:

学完了前三章, 有人可能会问: 大学“代数”的基本内容为什么叫“线性代数”? 它为什么从抽象的线性空间结构和线性映射的性质入手来讲? 它和中学代数怎么如此不同?

代数在它的发展史上分为两个阶段, 一为古典代数, 或说初等代数 (中学代数的主要内容), 它以数的运算律为基础, 以方程论为中心, 研究代数方程的求解问题. 它的研究方法是高度计算性的, 求解代数方程的基本问题是解的存在问题和求解方法. 高斯 (Gauss) 在 1799 年证明了复系数代数方程至少有一个复数根 (称为代数基本定理), 从而可知, n 次代数方程在复数域中有 n 个根, 解决了根的存在问题. 至于求解方法, 5 次以下的代数方程都有求根公式, 而 5 次和 5 次以上的代数方程能否用它的系数通过四则运算与根式运算的公式来求根, 经过很多数学家 (如拉格朗日 (Lagrange), 鲁菲尼 (Ruffini), 阿贝尔 (Abel) 等) 特别是法国天才的伽罗瓦 (Galois) 研究了拉格朗日, 阿贝尔和高斯的著作后, 创造性地引进置换群的正规子群和数域的扩域以及群的同构等概念, 证明了它们没有求根公式. 至此, 古典代数的问题已经终结. 代数进入了抽象代数 (也称近世代数) 的新阶段. 抽象代数研究的对象是非特定的任意集合和定义在集合上满足某些条件或公理的代数运算构成的代数系统, 也就是它以研究各种代数结构的性质为中心. 它的研究方法主要是公理化的.

自古以来“解方程”就是数学(尤其是代数)中最常被研究的对象之一.从最简单的一元线性(一次)方程和方程组(例如鸡兔同笼问题),线性同余方程组(例如第一章习题中的“将军点兵”问题),二次及高次方程和方程组,超越方程(例如 Kepler 方程 $a - \varepsilon \sin x = x$)乃至微分方程,积分方程,泛函方程等等,都是历史上或当代被研究的对象.本世纪电子计算机的出现,更使这方面的研究如虎添翼.尤其是“线性化”的思想,使很多解方程的问题都近似地转化为求解大量未知数(成千上万乃至上亿个未知数)的线性方程组的问题.这里面除了要研究“快而准”的算法外,也要求从理论上对“解方程”这一古老的问题进行各种探索.

所谓“解方程”,就是指由已知某些“量”(可以是各种数,向量,矩阵,函数,乃至某一类集合)之间应该满足的某种“关系”(即列方程)出发,求出其中的未知量.这里首先需要根据经验或理论分析认定这样的量是存在的.上面所说的“关系”(通常表现为方程)往往可以看作是一种映射.例如“鸡兔同笼”问题,设 x_1, x_2 分别为鸡,兔的数量, y_1, y_2 分别为鸡,兔的总头数和总腿数,于是得到方程组

$$\begin{cases} x_1 + x_2 = y_1 \\ 2x_1 + 4x_2 = y_2. \end{cases}$$

设 $X = (x_1, x_2)^T, Y = (y_1, y_2)^T$, 则 $X, Y \in \mathbb{N}^{*2}$ (表示元素为正整数的二维向量集合).这样,这个问题就可以看成是 \mathbb{N}^{*2} 上的一个映射 A (即方程组的系数矩阵): $X \mapsto Y$ (记作 $AX = Y$).因此,解这个方程就是求 A 的逆映射 A^{-1} , 即 $X = A^{-1}(Y)$.

这只是一种形式的写法,还需要进一步分析.容易证明这个映射 A 是线性的,但 \mathbb{N}^{*2} 不是线性空间(加法没有逆元素),所以 A 的值域(即 A^{-1} 的定义域)就不清楚.事实上,当 y_2 为奇数或 $y_2 < 2x_1$ 时,这个方程都无解.因此方程 $AX = Y$ 对任意的 $Y \in \mathbb{N}^{*2}$ 的可解性除了要分析 A 本身的性质外,还要分析 A 的定义域和值域的构造.读者可以自己分析一下 $\mathbb{R}^3 \rightarrow W$ (\mathbb{R}^3 中过原点的某一个平面)的投影变换 p , 即 $p(X) = Y$ (其中 $X \in \mathbb{R}^3, Y \in W$) 这个方程的解存在,但其解不唯一.

由此可见,一个解方程的问题如果能化成两个集合的映射问题,那么解方程就是“求映射的逆象”的问题.对方程可解性的分析,也就变成对相应的映射的分析,具体说就是对两个集合的构造及映射本身特性的分析.而 n 个元 m 个线性方程的方程组 $AX = Y$ (其中 $X \in \mathbb{R}^n, Y \in \mathbb{R}^m$) 的可解性问题就是其

中的典型问题. 这里两个集合 $\mathbf{R}^n, \mathbf{R}^m$ 都是线性空间, 映射是线性映射. 对于较简单的齐次线性方程组, 它总是可解的, 零向量就是解(注意这一点可由映射的线性性推出来), 问题是是否有非零解, 也就是映射的核空间的维数是多少? 而对于一般的非齐次方程组可能无解(因为方程组中可能有矛盾方程), 它是否有解, 取决于非齐次项所组成的向量是否在映射的值域(值空间)内, 求它的解也就是求向量 Y 在映射 A 下的完全逆象 $A^{-1}(Y)$. 因此, 对于线性方程组 $AX = Y$ (其中 $Y = 0$ 或 $Y \neq 0$) 的求解问题, 需要从理论上搞清有解(对于齐次方程组是有非零解)的条件以及解有多少和解的结构. 通过第二、三章的学习, 我们感到这个问题与线性映射 A 的值域 ($\text{Im}A$) 和核 ($\text{Ker}A$) 有关. 由于 $\dim(\text{Im}A) + \dim(\text{Ker}A) = n$, 而 $\dim(\text{Im}A) = \text{秩}(A)$, 所以, 这里的关键是线性映射的秩.

对有限维空间 \mathbf{R}^n 到 \mathbf{R}^m 的线性映射 A , 它的秩总是有限的. 其核空间的维数反映了齐次方程 $AX = 0$ 的线性无关的解的个数, 其值域是判断非齐次方程可解性的根据(关于它们的解的结构和求解方法将在第 6 章详述). 但可惜的是这一点只适用于有限维线性空间的线性映射. 把这一重要概念推广至无穷维空间的线性映射(即函数空间的线性微分方程和积分方程等), 已有较为满意的结果. 但要进一步推广到非线性映射(超越方程和非线性代数方程, 非线性微分方程等), 虽然多年来已有很多工作, 但还远没有像线性情况解决得那么清楚.

从上面分析可见, 一般的 n 元线性方程组 $AX = Y$ 的求解问题就是在线性空间中求线性映射 A 关于 Y 的完全逆象. 所以它只是 \mathbf{R}^n 到 \mathbf{R}^m 的线性映射的一个具体问题(或具体模型), 而一般的线性空间结构和线性映射的理论在现代数学乃至很多科学技术领域里都有着广泛的应用. 因此线性空间的结构和线性映射的基本理论是线性代数的核心内容. 从这里入手, 就能较好地展开线性代数所涉及的其他问题. 同时我们主要用公理化的方法讲述线性空间(一种重要的代数结构)和线性映射的理论, 也将有助于读者以后进一步学习现代数学的其他内容.

习 题

1. 判别下列映射哪些是线性的(并指出它们是从哪个空间到哪个空间的线性映射), 哪些不是线性的:

$$(1) \sigma(x_1, x_2, x_3) = (x_1 + x_2, x_1 - x_3, x_2), \quad \forall (x_1, x_2, x_3) \in \mathbf{R}^3;$$

$$(2) \alpha(x_1, x_2) = (x_1 - x_2, x_1 + x_2), \quad \forall (x_1, x_2) \in \mathbf{R}^2;$$

$$(3) \alpha(x_1, x_2) = (x_1^2, x_1 + x_2), \quad \forall (x_1, x_2) \in \mathbf{R}^2;$$

(4) $\alpha(\xi) = 2\xi + \xi_0$, $\forall \xi \in V$ (线性空间), ξ_0 是 V 中的一个固定向量;

$$(5) \alpha(f(x)) = f(x+1) - f(x), \quad \forall f(x) \in \mathbf{R}[x]_n;$$

$$(6) \alpha(f(x)) = f(a), \quad \forall f(x) \in \mathbf{R}[x], a \text{ 是一个常数.}$$

2. 求 \mathbf{R}^2 的线性变换 σ , 使得正方形 $ABCD$ (其中 $A(1, 0), B(0, 1), C(-1, 0), D(0, -1)$) 变换为下列四边形 $A'B'C'D'$ (如图 3-11 所示).

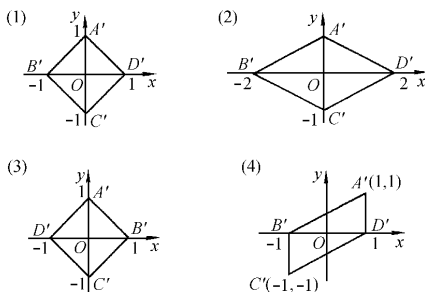


图 3-11

3. 设 \mathbf{R}^3 的一个二维子空间 W (W 是过原点的平面上的全体向量) 为

$$W = \{\xi \mid (\xi, \omega) = 0, \xi \in \mathbf{R}^3\},$$

其中 ω 是一个固定单位向量 (ξ, ω) 表示 ξ 与 ω 的内积.

(1) 求 \mathbf{R}^3 到 W 上的投影变换 p , 并证明 $p^2 = p$;

(2) 求 \mathbf{R}^3 关于镜面 W 的镜象变换 φ .

4. 设 $\{\alpha_1, \alpha_2\}$ 是线性空间 $V(F)$ 的一组基, $x_1\alpha_1 + x_2\alpha_2 \in V$, 定义

$$T(x_1\alpha_1 + x_2\alpha_2) = r_1x_1\alpha_1 + r_2x_2\alpha_2,$$

其中 r_1, r_2 是域 F 中的两个常数量. 证明: T 是 V 上的一个线性变换, 当 $V(F) = \mathbf{R}^2$ 时, 说明线性变换 T 的几何意义.

5. 求下列线性变换 σ 的象(值域)和核以及 σ 的秩:

$$(1) \alpha(x_1, x_2, x_3) = (x_1 + x_2 + x_3, -x_1 - 2x_3, x_2 - x_3);$$

$$(2) \alpha(x_1, x_2) = (x_2, x_1, x_2 - x_1);$$

$$(3) \alpha(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_2 + x_3, x_3);$$

(4) σ 是 n 维线性空间 V 的零变换;

(5) σ 是 n 维线性空间 V 的恒等变换;

(6) $\sigma: \mathbf{R}^n \rightarrow \mathbf{R}^n$, 且 $\sigma(x_1, x_2, \dots, x_n) = (x_1, 0, \dots, 0)$.

6. 求 \mathbf{R}^3 的一个线性变换 σ , 使得 σ 的象(值域)为 $\sigma(\mathbf{R}^3) = L(\alpha_1, \alpha_2)$, 其中 $\alpha_1 = (1, 0, -1)$, $\alpha_2 = (1, 2, 2)$.

7. 已知 \mathbf{R}^2 的线性变换 $\sigma(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$.

(1) 求 $\sigma^2(x_1, x_2) = ?$

(2) σ 是否可逆? 如可逆, 求 $\sigma^{-1}(x_1, x_2) = ?$

8. 已知 \mathbf{R}^2 的线性变换 $\sigma(x_1, x_2) = (x_1 - x_2, x_2 - x_1)$.

(1) 求 σ 的秩和 $\text{Ker}\sigma$;

(2) 求 $\tau \in L(\mathbf{R}^2, \mathbf{R}^2)$ 使 $\tau\sigma = \theta$ (零变换).

9. 已知 \mathbf{R}^3 的两个线性变换 σ, τ 为:

$$\sigma(x_1, x_2, x_3) = (x_3, 0, 0),$$

$$\tau(x_1, x_2, x_3) = (x_1 + x_2 + x_3, x_1 - x_2, 0).$$

(1) 求秩(σ) 秩(τ), $\text{Im}\sigma, \text{Ker}\sigma$; (2) 求秩($\sigma\tau$) 秩($\tau\sigma$);

(3) 求秩($\sigma + \tau$);

(4)* 求 $\text{Im}\tau + \text{Ker}\tau$.

10. 已知

$$\alpha_1 = (1, -1), \alpha_2 = (2, -1), \alpha_3 = (-3, 2),$$

$$\beta_1 = (1, 0), \beta_2 = (0, 1), \beta_3 = (1, 1).$$

问: 是否存在 $\sigma \in L(\mathbf{R}^2, \mathbf{R}^2)$, 使得 $\sigma(\alpha_i) = \beta_i, i = 1, 2, 3$.

11. 设 $\sigma \in L(\mathbf{R}^n, \mathbf{R}^m)$, $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 \mathbf{R}^n 的基, 当 $m < n$ 时, $\{\sigma(\alpha_1), \sigma(\alpha_2), \dots, \sigma(\alpha_n)\}$ 是否线性无关? 为什么?

12. 已知 σ 是 n 维线性空间 V 的线性变换, 且 σ 的像(值域)等于 σ 的核, 证明 n 必为偶数, 并在 \mathbf{R}^2 中举出一个这种线性变换的例子.

13*. 已知 \mathbf{R}^3 的一个子空间 $W_1 = \{x, 0, 0) | x \in \mathbf{R}\}$.

(1) 求 \mathbf{R}^3 的另一子空间 W_2 , 使 $\mathbf{R}^3 = W_1 \oplus W_2$, 这样的 W_2 是否唯一? 如果 W_2 是 W_1 的正交补, W_2 是否唯一?

(2) 求 \mathbf{R}^3 上的一个投影变换 p , 使 $\text{Im}p = W_1$, 并问: $\text{Im}p + \text{Ker}p = \mathbf{R}^3$ 是否成立?

14*. 已知 $\sigma \in L(V, V)$, $\dim V = n$, 有人认为“由秩(σ) + $\dim(\text{Ker}\sigma) = n$ 可得 $\text{Im}\sigma + \text{Ker}\sigma = V$ ”, 这个说法正确吗?

15*. 已知 V 是一个线性空间, $\sigma \in L(V, V)$, 证明:

(1) $\text{Im}\sigma \subset \text{Ker}\sigma \Leftrightarrow \sigma^2 = \mathcal{O}$ (零变换);

(2) $\text{Ker}\sigma^k \subset \text{Ker}\sigma^{k+1}$, $k = 1, 2, \dots$;

(3) $\text{Im}\sigma^k \supset \text{Im}\sigma^{k+1}$, $k = 1, 2, \dots$.

16. 设 $\sigma \in L(V, V)$, $\xi \in V$. 证明: 若 $\sigma^{k-1}(\xi) \neq 0$, $\sigma^k(\xi) = 0$, 则 $\{\xi, \sigma(\xi), \dots, \sigma^{k-1}(\xi)\}$ 是线性无关的 ($k > 0$).

17. 下列线性空间 V_1 与 V_2 是否同构? 如果同构, 给出 V_1 到 V_2 的一个同构映射.

(1) $V_1 = \mathcal{C}(\mathbf{R})$ (全体复数 \mathbf{C} 在实数域 \mathbf{R} 上构成的线性空间), $V_2 = \mathbf{R}^2$;

(2) $V_1 = \mathbf{C}^2(\mathbf{R})$ (全体二元复向量 \mathbf{C}^2 在实数域 \mathbf{R} 上构成的线性空间), $V_2 = \mathbf{R}[x]_4$;

(3) $V_1 = \mathbf{C}^2(\mathbf{R})$, $V_2 = \mathbf{R}[x]_4$;

(4)* $V_1 = \mathbf{R}^2$, $V_2 = L(\mathbf{R}^2, \mathbf{R}^1)$.

18. 设 $V(\mathbf{R})$ 是线性空间, σ 是 $V(\mathbf{R})$ 到 \mathbf{R}^3 的同构映射, 且 $\sigma(\alpha_1) = (1, 0, 1)$, $\sigma(\alpha_2) = (-2, 1, 0)$, $\sigma(\alpha_3) = (-3, 2, 1)$, $\sigma(\alpha_4) = (1, 1, 2)$.

(1) α_1 在 V 的子空间 $L(\alpha_2, \alpha_3)$ 中吗?

(2)* 设 $W_1 = L(\alpha_1, \alpha_2)$, $W_2 = L(\alpha_3, \alpha_4)$, 求 $W_1 \cap W_2$.

补充题

1. 是否存在 \mathbf{R}^3 到 \mathbf{R}^2 的一个线性映射 σ , 使得 $\sigma(1, -1, 1) = (1, 0)$, $\sigma(1, 1, 1) = (0, 1)$?

2*. 设 V 是有限维线性空间, $T \in L(V, V)$ 且 T 不是恒等变换也不是零变换. 问: 下列情况是否可能发生 (在 \mathbf{R}^2 或 \mathbf{R}^3 中举例).

(1) $\text{Im}T \cap \text{Ker}T = \{0\}$; (2) $\text{Im}T \subset \text{Ker}T$;

(3) $\text{Im}T = \text{Ker}T$; (4) $\text{Ker}T \subset \text{Im}T$.

又问: 当 T 为投影变换时, 上述哪些情况发生?

3*. 已知 V 是一个线性空间, $\sigma \in L(V, V)$. 证明:

(1) $\text{Ker}\sigma = \text{Ker}\sigma^2 \Leftrightarrow (\text{Ker}\sigma) \cap (\text{Im}\sigma) = \{0\}$ (零子空间);

(2) $\text{Im}\sigma^2 = \text{Im}\sigma \Leftrightarrow V = (\text{Ker}\sigma) + (\text{Im}\sigma)$.

4*. 设 V 是一个 n 维线性空间, $V = W_1 \oplus W_2$, $\sigma \in L(V, V)$, 证明:

$$\sigma \text{ 可逆} \Leftrightarrow V = \sigma(W_1) + \sigma(W_2).$$

5. 设 $\sigma, \tau \in L(V, V)$, $\sigma^2 = \sigma$, $\tau^2 = \tau$, 证明:

(1) $\sigma^k = \sigma$ ($k \geq 2$ 即 σ 是幂等变换);

(2) 若 $(\sigma + \tau)^2 = \sigma + \tau$ 则 $\sigma\tau = \theta$ (零变换);

(3) 若 $\sigma\tau = \tau\sigma$ 则 $(\sigma + \tau - \sigma\tau)^2 = \sigma + \tau - \sigma\tau$.

6*. 设 $V(F)$ 是一个 n 维线性空间, $\sigma \in L(V, V)$, 证明:

(1) 在 $F[x]$ 中有一个次数 $\leq n^2$ 的非零多项式 $f(x)$ 使 $f(\sigma) = \theta$ (零变换);

(2) σ 可逆 \Leftrightarrow 有一常数项不为零的多项式 $f(x)$ 使 $f(\sigma) = \theta$.

7*. 设 c_1, c_2, \dots, c_n 是 n 个互异的实常数, 证明: $\mathbb{R}[x]_n$ 到 \mathbb{R}^n 的一个映射 σ :

$$\sigma(f(x)) = (f(c_1), f(c_2), \dots, f(c_n))$$

是 $\mathbb{R}[x]_n$ 到 \mathbb{R}^n 的一个同构映射.

这里要求读者对 $n = 3$ 的情形给以证明.

对一般的 n , 其证明要用到第 5 章中的 Vandermonde 行列式的性质.

部分习题和补充题答案

习题

1. (1)(2)(5)(6) 是 (3)(4) ($\xi_0 \neq 0$) 不是.

2. (1) $\sigma(x, y) = (-y, x)$; (2) $\sigma(x, y) = (-2y, x)$;

(3) $\sigma(x, y) = (y, x)$; (4) $\sigma(x, y) = (x - y, x)$;

3. (1) $f(\alpha) = \alpha - (\alpha, \omega)\omega$; (2) $f(\alpha) = \alpha - 2(\alpha, \omega)\omega$.

5. (1) $L\{(1, -1, 0), (1, 0, 1)\}, L\{(-2, 1, 1)\} \cong \mathbb{R}^2$;

(2) $L\{(1, 0, 1), (0, 1, -1)\}, \{0\} \cong \mathbb{R}^2$; (3) $\mathbb{R}^3, \{0\} \cong \mathbb{R}^3$; (4) $\{0\}, V \cong \{0\}$;

(5) $V, \{0\} \cong \mathbb{R}^m$; (6) $L\{e_1\}, L\{e_2, \dots, e_n\} \cong \mathbb{R}^1$.

6. 提示: 参考 3.3 例 3 的解法. $\sigma(x_1, x_2, x_3) = (x_1 + x_2, 2x_2, 2x_2 - x_1)$.

7. (1) $\sigma^2(x_1, x_2) = (-2x_2, 2x_1)$;

$$(2) \sigma^{-1}(x_1, x_2) = \left(\frac{x_1 + x_2}{2}, \frac{x_2 - x_1}{2} \right).$$

8. (1) $L\{(1, 1)\}$; (2) $\tau(x_1, x_2) = (x_1 + x_2, x_1 + x_2)$ 或 $\tau = \theta$.

9. (1) $L\{(1, 0, 0)\}, L\{(1, 0, 0), (0, 1, 0)\}$;

(2) \mathbb{R}^1 ; (3) \mathbb{R}^2 ; (4) \mathbb{R}^3 .

10. 否. 11. 否. 12. 例 $\sigma(x, y) = (x + y, -x - y)$.

13. (1) $L\{(0, 1, 0), (0, 0, 1)\}$ 否; (2) $f(x_1, x_2, x_3) = (x_1, 0, 0)$ 是.

16. 提示: 用线性相关性的定义证明, 并利用 $\lambda\alpha = 0, \alpha \neq 0 \Rightarrow \lambda = 0$.

17. (1) $a + ib \mapsto (a, b)$; (2) $(1, 0) \mapsto 1, (i, 0) \mapsto x, (0, 1) \mapsto x^2, (0, i) \mapsto$

$\mapsto x^3$; (3) 否; (4) $(a, b) \mapsto ax + by$ 或 $\overline{e_1}(e_1) = 1, \overline{e_1}(e_2) = 0, \overline{e_2}(e_1) = 0, \overline{e_2}(e_2) = 1, e_i \mapsto \overline{e_i} (i = 1, 2)$.

18. (1) 在, (提示: 考虑 $\alpha(\alpha_1)$ 与 $\alpha(\alpha_2), \alpha(\alpha_3)$ 之间的关系);
(2) $L(\alpha_3)$.

补充题

1. 存在. 2. (1)(2)(3)(4) 可能 (1) 发生.

6. (1) 提示: 利用 $\dim L(V, V) = n^2$, $L(V, V)$ 中任何 $n^2 + 1$ 个元素 (即线性映射) 都是线性相关的. 若 $p(x) = a_0 + a_1x + \dots + a_mx^m$, 则 $p(\sigma) = a_0I + a_1\sigma + \dots + a_m\sigma^m$;

(2) 提示: 利用 (1) 的结果以及“若 $\sigma\tau = I$ 则 σ 可逆”.

7. 提示: 只要证明 σ 把 $\mathbb{R}[x]$ 的基 $\{1, x, x^2\}$ 变为 \mathbb{R}^3 的基.

第 4 章 矩 阵

本章是前一章有限维线性空间线性映射的继续. 这里我们引入线性映射的矩阵表示, 相应地定义矩阵的运算, 从而把对线性映射的演算化为矩阵的运算. 本章主要内容有: 线性映射与矩阵的对应关系, 矩阵的基本运算——加法, 数乘, 乘法, 求逆, 转置以及矩阵的分块运算, 矩阵的初等变换与矩阵的秩; 有限维线性空间的基的变换矩阵与向量的坐标变换. 由于矩阵在科学技术领域里的应用非常广泛, 因此读者要特别注重培养熟练的矩阵计算能力.

4.1 矩阵的定义

定义 4.1 域 F 中的 $m \times n$ 个元素 a_{ij} ($i = 1, \dots, m$; $j = 1, \dots, n$) 排成 m 行 n 列的矩形数表, 称为域 F 上的一个 $m \times n$ 矩阵, 记作

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (4-1)$$

或简记为 $(a_{ij})_{m \times n}$, 其中 a_{ij} 叫做矩阵第 i 行, 第 j 列的元素. 元素 a_{ij} 都是零 (域 F 的零元) 的矩阵称为零矩阵.

矩阵常用大写字母 A, B, C 表示. (4-1) 式也可记作 $A_{m \times n}$. 如果域 F 为实 (复) 数域 $\mathbf{R}(\mathbf{C})$, (4-1) 式就称为实 (复) 矩阵. 域 F 上全体 $m \times n$ 矩阵组成的集合, 记作 $M_{m \times n}(F)$ 或 $F^{m \times n}$.

当 $m = n$ 时, 矩阵 (4-1) 称为方阵, 或称 n 阶矩阵. 元素 $a_{11}, a_{22}, \dots, a_{nn}$ 叫做方阵的主对角元, 它们所在位置叫做方阵的主对角线, 方阵的另一条对角线叫做副对角线. 域 F 上全体 n 阶矩阵组成的集合, 记作 $M_n(F)$ 或 $F^{n \times n}$.

如果方阵中非主对角线上的所有元素都是零 (即当 $i \neq j$ 时 $a_{ij} = 0$) 称之为对角矩阵 (简称对角阵); 如果在主对角线之下 (上) 的所有元素都是零, 即当 $i > j$ 时 $a_{ij} = 0$ ($i < j$ 时 $a_{ij} = 0$) 称之为上 (下) 三角矩阵.

如果 $A = (a_{ij})$ 和 $B = (b_{ij})$ 都是 $m \times n$ 矩阵, 我们规定 $A = B$ 当且仅当 $a_{ij} = b_{ij}$ ($i = 1, \dots, m$; $j = 1, \dots, n$). 所以一个 $m \times n$ 矩阵的等式, 等价于

$m \times n$ 个数量(域 F 的元)等式.

4.2 线性映射的矩阵表示

在第3章中我们看到一些线性映射的例子,其中每个向量及其像都有明确的数值表示.这一章我们探讨如何用矩阵表示一般的有限维线性空间的线性映射 $\sigma \in L(V_1, V_2)$. 由于 V_1, V_2 中的向量都可以用基来线性表示,而 σ 又是线性的,所以我们就只需考察 V_1 的基映射到 V_2 后如何用 V_2 的基表示.

设 $B_1 = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 是 $V_1(F)$ 的基, $B_2 = \{e_1, e_2, \dots, e_m\}$ 是 $V_2(F)$ 的基,则线性映射 $\sigma \in L(V_1, V_2)$ 被它作用于基 B_1 的象

$$\sigma(B_1) = \{\sigma(\varepsilon_1), \sigma(\varepsilon_2), \dots, \sigma(\varepsilon_n)\} \quad (4-2)$$

所唯一确定,而 $\sigma(B_1)$ 是 V_2 的一个子集,于是

$$\begin{cases} \sigma(\varepsilon_1) = a_{11}e_1 + a_{21}e_2 + \dots + a_{m1}e_m \\ \sigma(\varepsilon_2) = a_{12}e_1 + a_{22}e_2 + \dots + a_{m2}e_m \\ \dots\dots\dots \\ \sigma(\varepsilon_n) = a_{1n}e_1 + a_{2n}e_2 + \dots + a_{mn}e_m. \end{cases} \quad (4-3)$$

因此,在取定了 V_1 和 V_2 的基 B_1 和 B_2 时, σ 由(4-3)式中的 $m \times n$ 个系数 a_{ij} 所确定,而且由第3章定理3.6还可知,任意给定(4-3)式中的 $m \times n$ 个系数 a_{ij} ,即任意给定 V_2 中 n 个向量,使之分别为 V_1 的 n 个基向量的像,也唯一地确定了 V_1 到 V_2 的一个线性映射,于是,我们给出下面的定义.

定义4.2 σ, V_1, V_2, B_1 和 B_2 如上所设,我们把 $\sigma(\varepsilon_1), \sigma(\varepsilon_2), \dots, \sigma(\varepsilon_n)$ 关于基 B_2 的坐标按列排成的矩阵 $M(\sigma)$,即

$$M(\sigma) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad (4-4)$$

称为 σ 关于基 B_1 和 B_2 的矩阵.

从上面的分析可知,选定了 V_1 和 V_2 的基之后, $L(V_1, V_2)$ 中的元素(线性映射) σ 与域 F 上 $m \times n$ 型矩阵 A 是一一对应的.因此, σ 可由对应的矩阵 $A = M(\sigma)$ 来表示.要注意的是,对 V_1, V_2 中不同的基,矩阵 $M(\sigma)$ 一般是不同的,这个问题将在第7章中讨论.

当 $\sigma \in L(V_1, V_2)$ 且 $V_2 = V_1$ 时,通常取 $B_2 = B_1$,此时,相应的 $n \times n$

方阵 $M(\sigma)$ 称为 σ 关于基 B_1 的矩阵.

例 1 n 维线性空间 $V(F)$ 上的恒等变换 I , 数乘变换 Δ , 零变换 θ 关于 V 的任何基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 的矩阵分别为 n 阶方阵:

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & & & \dots \\ 0 & 0 & \dots & \lambda \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & & & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

它们分别称为单位矩阵, 数量矩阵和零矩阵.

例 2 \mathbf{R}^2 上的旋转变换 r_θ (向量绕原点按逆时针方向旋转 θ 角) 关于单位正交基 $B = \{e_1, e_2\}$ (图 4-1) 的矩阵为

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

这是因为 r_θ 关于单位向量 e_1, e_2 的象为

$$\begin{cases} r_\theta(e_1) = (\cos \theta)e_1 + (\sin \theta)e_2 \\ r_\theta(e_2) = (-\sin \theta)e_1 + (\cos \theta)e_2. \end{cases}$$

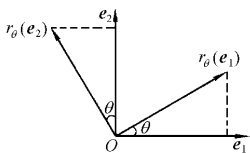


图 4-1

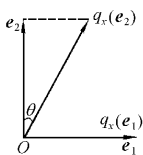


图 4-2

例 3 第 3 章 3.2 例 5 中的错切变换 (图 4-2)

$$q_x(x, y) = (x + cy, y)$$

(其中 $c = \tan \theta$), 关于自然基 $e_1 = (1, 0), e_2 = (0, 1)$ 的象为

$$q_x(e_1) = q_x(1, 0) = (1, 0) = e_1,$$

$$q_x(e_2) = q_x(0, 1) = (c, 1) = ce_1 + e_2.$$

因此 q_x 关于基 $B = \{e_1, e_2\}$ 的矩阵为

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

例 4 \mathbf{R}^3 上以坐标平面 xOy 为镜面的镜象变换 φ 和 \mathbf{R}^3 到 xOy 坐标平面

的投影变换 p 关于自然基 $B = \{e_1, e_2, e_3\}$ 的矩阵分别为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

例 5 已知 $\sigma \in L(\mathbf{R}^3, \mathbf{R}^2)$ 且 $\sigma(x_1, x_2, x_3) = (x_1 + x_2, 2x_1 - x_2 - 4x_3)$, 求 σ 关于 \mathbf{R}^3 的自然基 $B_1 = \{e_1, e_2, e_3\}$ 和 \mathbf{R}^2 的自然基 $B_2 = \{e_1, e_2\}$ 的矩阵 $M(\sigma)$.

解 由

$$\sigma(e_1) = \sigma(1, 0, 0) = (1, 2) = e_1 + 2e_2,$$

$$\sigma(e_2) = \sigma(0, 1, 0) = (1, -1) = e_1 - e_2,$$

$$\sigma(e_3) = \sigma(0, 0, 1) = (0, -4) = -4e_2,$$

立即可得

$$M(\sigma) = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & -4 \end{pmatrix}.$$

如果取基 $B_1 = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$, 读者不难求得

$$M(\sigma) = \begin{pmatrix} 2 & 2 & 1 \\ -3 & 1 & 2 \end{pmatrix}.$$

4.3 矩阵的加法与数量乘法

设 $\sigma, \tau \in L(V_1, V_2)$ 线性空间 $V_1(F), V_2(F)$ 的基分别为 $B_1 = \{e_1, e_2, \dots, e_n\}, B_2 = \{e_1, e_2, \dots, e_m\}$, σ, τ 关于基 B_1 和 B_2 的矩阵为 $M(\sigma) = (a_{ij})_{m \times n}, M(\tau) = (b_{ij})_{m \times n}; \lambda \in F$, 则根据定义 4.2 由

$$(\sigma + \tau)(e_j) = \sigma(e_j) + \tau(e_j)$$

$$= (a_{1j}e_1 + \dots + a_{mj}e_m) + (b_{1j}e_1 + \dots + b_{mj}e_m)$$

$$= (a_{1j} + b_{1j})e_1 + \dots + (a_{mj} + b_{mj})e_m, \quad j = 1, 2, \dots, n,$$

$$(\lambda\sigma)(e_j) = \lambda\sigma(e_j) = \lambda(a_{1j}e_1 + \dots + a_{mj}e_m)$$

$$= (\lambda a_{1j})e_1 + \dots + (\lambda a_{mj})e_m, \quad j = 1, 2, \dots, n,$$

即得

$$M(\sigma + \tau) = (a_{ij} + b_{ij})_{m \times n}, \quad (4-5)$$

$$M(\lambda\sigma) = (\lambda a_{ij})_{m \times n}. \quad (4-6)$$

为了使线性映射的加法和数乘运算也对应于矩阵的加法和数乘运算,根据(4-5)(4-6)式,我们对矩阵的加法和数量乘法作如下的定义.

定义 4.3 设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{m \times n}$, $A, B \in M_{m \times n}(F)$, $\lambda \in F$, 我们规定

$$A + B = (a_{ij} + b_{ij})_{m \times n}, \quad (4-7)$$

$$\lambda A = (\lambda a_{ij})_{m \times n}. \quad (4-8)$$

即 $A + B$ 是由 A 与 B 的所有对应元素分别相加得到的 $m \times n$ 型矩阵; λA 是由数量 λ 与 A 的每个元素相乘而得到的 $m \times n$ 型矩阵.

此外我们还定义 $A - B = A + (-B)$, 其中 $(-B) = (-b_{ij})_{m \times n}$.

域 F 上全体矩阵组成的集合 $M_{m \times n}(F)$ 对上面定义的矩阵的加法和数乘运算在域 F 上构成一个线性空间, 因为由定义 4.3 容易证明, 在集合 $M_{m \times n}(F)$ 上:

(1) 加法和数乘运算封闭;

(2) 加法满足结合律, 交换律, 加法的单位元为零矩阵, 任一矩阵 $A = (a_{ij})_{m \times n}$ 关于加法的逆元为 $-A = (-a_{ij})_{m \times n}$, 即 $A + (-A) = O$, 所以 $M_{m \times n}(F): +$ 是一个交换群;

(3) 数乘运算满足以下四条规则:

$$1A = A; \quad \lambda(\mu A) = (\lambda\mu)A;$$

$$(\lambda + \mu)A = \lambda A + \mu A; \quad \lambda(A + B) = \lambda A + \lambda B.$$

下面证明线性空间 $M_{m \times n}(F)$ 的维数等于 $m \times n$. 为了简明(也不失一般性), 我们证明 $M_{2 \times 2}(F)$ 的维数等于 4.

事实上, $M_{2 \times 2}(F)$ 中存在 4 个线性无关的元素(二阶矩阵):

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

因为由

$$\lambda_{11}e_{11} + \lambda_{12}e_{12} + \lambda_{21}e_{21} + \lambda_{22}e_{22} = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

得 $\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}$ 全为零, 故 $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ 线性无关. 又任意的

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}e_{11} + a_{12}e_{12} + a_{21}e_{21} + a_{22}e_{22},$$

因此 $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ 是 $M_{2 \times 2}(F)$ 的一组基, 所以 $\dim M_{2 \times 2}(F) = 4$.

在第 3 章中讲过, 若 $\dim V_1(F) = n$, $\dim V_2(F) = m$, 则 $\dim L(V_1, V_2) = mn$, 但一

般的证明比较难. 现在我们可以先证明 $L(V_1, V_2) \cong M_{m \times n}(F)$. 然后根据定理 3.8 得

$$\dim L(V_1, V_2) = \dim M_{m \times n}(F) = mn$$

$L(V_1, V_2) \cong M_{m \times n}(F)$ 是容易证明的. 事实上, 给定了 $V_1(F)$ 与 $V_2(F)$ 的基 B_1 和 B_2 后, 由定义 4.2 所确定的从 $L(V_1, V_2)$ 到 $M_{m \times n}(F)$ 的映射 φ :

$$\varphi: \sigma \mapsto M(\sigma) = (a_{ij})_{m \times n} \quad \text{即} \quad \varphi(\sigma) = M(\sigma)$$

是双射 (即 σ 与矩阵 $(a_{ij})_{m \times n}$ 是一一对应的), 而且是线性的, 因为我们在定义矩阵加法和数量乘法时, 就有

$$\varphi(\sigma + \tau) = M(\sigma + \tau) = M(\sigma) + M(\tau) = \varphi(\sigma) + \varphi(\tau),$$

$$\varphi(\lambda\sigma) = M(\lambda\sigma) = \lambda M(\sigma) = \lambda\varphi(\sigma).$$

因此 φ 是从 $L(V_1, V_2)$ 到 $M_{m \times n}(F)$ 的一个同构映射. 故 $L(V_1, V_2) \cong M_{m \times n}(F)$.

当 $\dim V = n$ 时, $L(V, V) \cong M_n(F)$ (域 F 上的全体 n 阶方阵的集合), $\dim L(V, V) = n^2$.

4.4 矩阵的乘法

设线性空间 $V_1(F), V_2(F), V_3(F)$ 的基分别为 $B_1 = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m\}$, $B_2 = \{e_1, e_2, \dots, e_m\}$, $B_3 = \{\zeta_1, \zeta_2, \dots, \zeta_p\}$; $\sigma \in L(V_1, V_2)$, $\tau \in L(V_2, V_3)$, 且 σ, τ 分别关于基 B_1 和 B_2 及基 B_2 和 B_3 的矩阵为 $B = (b_{ij})_{m \times n}$ 和 $A = (a_{ij})_{p \times m}$, 即

$$M(\tau) = A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & & & \cdots \\ a_{p1} & a_{p2} & \cdots & a_{pm} \end{pmatrix},$$

$$M(\sigma) = B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & & & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix},$$

则 $\tau\sigma \in L(V_1, V_3)$ 关于基 B_1 和 B_3 的矩阵 $C = (c_{ij})_{p \times n}$ 中第 j 列元素 $(c_{1j}, c_{2j}, \dots, c_{pj})$ 是 $\tau(\sigma(\varepsilon_j))$ 在基 B_3 下的坐标. 于是由

$$\begin{aligned} \tau(\sigma(\varepsilon_j)) &= \tau(\sigma(\varepsilon_j)) = \tau\left(\sum_{k=1}^m b_{kj} e_k\right) = \sum_{k=1}^m b_{kj} \tau(e_k) \\ &= \sum_{k=1}^m b_{kj} \left(\sum_{i=1}^p a_{ik} \zeta_i\right) = \sum_{k=1}^m \sum_{i=1}^p a_{ik} b_{kj} \zeta_i = \sum_{i=1}^p \left(\sum_{k=1}^m a_{ik} b_{kj}\right) \zeta_i, \end{aligned}$$

即得

$$c_{ij} = \sum_{k=1}^m a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{im}b_{mj}, \quad (4-9')$$

它是 A 中第 i 行 m 个元素组成的 m 元(行)向量 $\alpha_i = (a_{i1} \ a_{i2} \ \dots \ a_{im})$ 与 B 中第 j 列的 m 个元素组成的 m 元(列)向量 $\beta_j = (b_{1j} \ b_{2j} \ \dots \ b_{mj})$ 的内积,即

$$c_{ij} = (\alpha_i, \beta_j), \quad i = 1 \dots, p; \quad j = 1 \dots, n. \quad (4-9)$$

为了使矩阵的乘法与映射的乘法相对应,我们把按(4-9)式所得到的 $\tau\sigma$ 对应的矩阵 $C = M(\tau\sigma)$ 定义为 $A = M(\tau)$ 与 $B = M(\sigma)$ 之乘积.因此两个矩阵的乘积的一般定义如下.

定义 4.4 设 $A = (a_{ij})_{p \times m}$, $B = (b_{ij})_{m \times n}$, 我们规定 A 与 B 之乘积 $AB = C = (c_{ij})$ 是一个 $p \times n$ 型矩阵, 它的第 i 行, 第 j 列元素

$$c_{ij} = \sum_{k=1}^m a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{im}b_{mj}, \\ i = 1 \dots, p; \quad j = 1 \dots, n.$$

必须注意, 乘积 AB 当且仅当 A 的列数等于 B 的行数时才有意义, 否则 A 不能左乘 B . 与映射的乘法一样, 矩阵的乘法也不满足交换律, 即一般来说, $AB \neq BA$. 当然, 这并不是说任何情况下 AB 都不等于 BA . 如果 $AB = BA$ (此时 A 与 B 必为同阶方阵), 则称 A 与 B 相乘可交换, 简称 A 与 B 可交换.

根据矩阵的乘法定义, 容易证明矩阵的乘法满足以下运算律:

- (1) $(AB)C = A(BC)$ (结合律);
- (2) $\lambda(AB) = (\lambda A)B = A(\lambda B)$ (其中 λ 是数量);
- (3) $A(B + C) = AB + AC$ (左分配律),
 $(B + C)P = BP + CP$ (右分配律).

下面证明结合律(其余的证明留给读者完成).

设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$, $C = (c_{ij})_{p \times r}$. 显然 $(AB)C$ 与 $A(BC)$ 都是 $m \times r$ 矩阵; 又 $\forall i = 1 \dots, m, \forall j = 1 \dots, r$, 有

$$\begin{aligned} [(AB)C]_{ij} &= \sum_{k=1}^p (AB)_{ik}c_{kj} = \sum_{k=1}^p \left(\sum_{l=1}^n a_{il}b_{lk} \right) c_{kj} \\ &= \sum_{l=1}^n a_{il} \left(\sum_{k=1}^p b_{lk}c_{kj} \right) = \sum_{l=1}^n a_{il} (BC)_{lj} = [A(BC)]_{ij}, \end{aligned}$$

所以 $(AB)C = A(BC)$.

由矩阵的加法和乘法满足的运算律可知, 全体 n 阶矩阵组成的集合

$M_n(F)$ 关于矩阵的乘法构成半群 ; 关于矩阵的加法和乘法构成环 .

例 1 设

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

则

$$AB = \begin{pmatrix} 0-3 & 3-1 \\ 1-9 & 3+2-3 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ -8 & 2 \end{pmatrix},$$

$$BA = \begin{pmatrix} 0+3 & 0+3 & 0-9 \\ 1+2 & 0+2 & -1-6 \\ 3+1 & 0+1 & -3-3 \end{pmatrix} = \begin{pmatrix} 3 & 3 & -9 \\ 3 & 2 & -7 \\ 4 & 1 & -6 \end{pmatrix}.$$

例 2 设

$$A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad B = (b_1 \ b_2),$$

则

$$AB = \begin{pmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{pmatrix}, \quad BA = b_1 a_1 + b_2 a_2.$$

BA 是一阶矩阵 , 此时不必加矩阵符号 .

例 3 设

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

则

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}.$$

以上各例都说明矩阵的乘法不满足交换律 . 从例 3 还可见 , A, B 都是非零矩阵 , 但 $AB = O$, 因此 , 对于矩阵乘法 , 由 $AB = O$, 不能推出 $A = O$ 或 $B = O$; 进而又得矩阵的乘法不满足消去律 , 也就是 $A \neq O$, 由 $AB = AC$ 不能推出 $B = C$, 这是因为 , 由 $AB - AC = A(B - C) = O$, 不能推出 $B - C = O$ (即 $B = C$) .

一般来说 , 在矩阵环中存在非零矩阵的零因子 , 即 $A \neq O$, 存在 $B \neq O$ 使 $AB = O$, 此时称 A 为左零因子 , B 为右零因子 .

例 4 n 阶单位矩阵

$$E_n = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}_{n \times n} \quad (\text{常简记作 } E \text{ 或 } I)$$

对任一 $A \in M_n(F)$, 都有 $EA = AE = A$. 所以 E 是 $M_n(F)$ 中乘法运算的单位元.

n 阶数量矩阵 λE (λ 是一个数量, E 是单位矩阵) 为

$$\lambda E = \begin{pmatrix} \lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda \end{pmatrix},$$

它与任意 n 阶矩阵 A 相乘可交换, 即 $(\lambda E)A = A(\lambda E) = \lambda A$.

例 5 对角矩阵

$$A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{pmatrix},$$

常记作 $A = \text{diag}(a_1, a_2, \dots, a_n)$, $B = \text{diag}(b_1, b_2, \dots, b_n)$, 其乘积仍为对角阵, 且

$$AB = BA = \text{diag}(a_1 b_1, a_2 b_2, \dots, a_n b_n).$$

例 6 两个上(下)三角矩阵 A 与 B 的乘积 AB 仍是上(下)三角矩阵, 且其主对角元 $(AB)_{ii} = a_{ii}b_{ii}$ (证明留给读者作为练习).

利用矩阵乘法, 可以把线性方程组简洁地表示为一个矩阵的等式. 例如对于线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (4-10)$$

我们令

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix},$$

则(4-10)式就可以用矩阵等式

$$AX = b$$

来表示,其中 A 叫做线性方程组的系数矩阵,在 A 的第 n 列后面添上一列 b , 叫做方程组的增广矩阵,记为 (A, b) .

利用矩阵的乘法,也可以把节 4.2 中的(4-3)式形式地表示为

$$(\sigma(\varepsilon_1), \sigma(\varepsilon_2), \dots, \sigma(\varepsilon_n)) = (e_1, e_2, \dots, e_m) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

这种表示完全是形式的,因为 $e_1, e_2, \dots, e_m, \sigma(\varepsilon_1), \sigma(\varepsilon_2), \dots, \sigma(\varepsilon_n)$ 并不是域 F 中的元素.通常我们把这个式子简记作

$$\sigma(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (e_1, e_2, \dots, e_m)A, \quad (4-11)$$

其中 A 是线性映射 $\sigma \in L(V_1, V_2)$ 关于基 $B_1 = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 和基 $B_2 = \{e_1, e_2, \dots, e_m\}$ 的矩阵.

同样,利用矩阵乘法,也可将

$$\alpha = x_1 \varepsilon_1 + x_2 \varepsilon_2 + \dots + x_n \varepsilon_n \in V_1,$$

$$\sigma(\alpha) = y_1 e_1 + y_2 e_2 + \dots + y_m e_m \in V_2$$

形式地表示为

$$\alpha = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)X, \quad \sigma(\alpha) = (e_1, e_2, \dots, e_m)Y, \quad (4-12)$$

其中

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{pmatrix}$$

分别是 α 关于基 B_1 的坐标和 $\sigma(\alpha)$ 关于基 B_2 的坐标.

定理 4.1 设 $\sigma \in L(V_1, V_2)$ 关于 V_1 和 V_2 的基 B_1 和基 B_2 的矩阵 $A = (a_{ij})_{m \times n}$ 如(4-11)式, α 与 $\sigma(\alpha)$ 如(4-12)式, 则

$$Y = AX. \quad (4-13)$$

证

$$\begin{aligned}\sigma(\alpha) &= x_1\sigma(\varepsilon_1) + x_2\sigma(\varepsilon_2) + \dots + x_n\sigma(\varepsilon_n) \\ &= (\sigma(\varepsilon_1) \sigma(\varepsilon_2) \dots \sigma(\varepsilon_n))X \\ &= ((e_1 \ e_2 \ \dots \ e_m)A)X = (e_1 \ e_2 \ \dots \ e_m)(AX).\end{aligned}$$

由于 $\sigma(\alpha)$ 关于基 $B_2 = \{e_1 \ e_2 \ \dots \ e_m\}$ 的坐标 Y 是唯一的, 所以 $Y = AX$.

我们在定义矩阵的加法、数乘和乘法时, 使得它们与线性映射同样的运算相对应. 因此, 研究线性映射的问题就可转化为研究矩阵的问题; 反之亦然.

如果 $\sigma \in L(V, V)$ 关于 $V(F)$ 的基 $B = \{\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n\}$ 对应的矩阵为 A , 即

$$\sigma(\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n) = (\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n)A,$$

则 σ 在域 F 上的多项式

$$p(\sigma) = a_m\sigma^m + a_{m-1}\sigma^{m-1} + \dots + a_1\sigma + a_0I$$

对应于矩阵 A 在域 F 上有相同系数的多项式

$$p(A) = a_mA^m + a_{m-1}A^{m-1} + \dots + a_1A + a_0E,$$

其中 $a_k \in F$, $\sigma^k = \sigma\sigma\dots\sigma$ (k 个 σ 之乘积) $\in L(V, V)$, $A^k = AA\dots A$ (k 个 A 之乘积) $\in M_n(F)$, $k = 0, 1, 2, \dots, m$, 规定 $\sigma^0 = I$ (恒等变换), $A^0 = E$.

显然, 若 A 为方阵, k, m, n 为非负整数, 则

$$A^kA^m = A^{k+m}, \quad (A^k)^n = A^{kn},$$

$$(A + \lambda E)^n = A^n + C_n^1\lambda A^{n-1} + \dots + C_n^k\lambda^k A^{n-k} + \dots + C_n^{n-1}\lambda^{n-1}A + \lambda^n E, \quad (4-14)$$

其中 $\lambda \in F$, C_n^k 为组合数. (4-14) 式成立是因为数量矩阵 λE 与 A 相乘可交换.

当 A, B 为同阶方阵时, 若 $AB \neq BA$, 则

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B) = A^2 + AB + BA + B^2 \\ &\neq A^2 + 2AB + B^2.\end{aligned}$$

4.5 可逆矩阵

如果 V_1, V_2 的维数都是 n , $\sigma \in L(V_1, V_2)$ 是可逆的, 则存在唯一的 $\tau \in L(V_2, V_1)$, 使得 $\tau\sigma = I_{V_1}$, $\sigma\tau = I_{V_2}$. 设 V_1 的基 $B_1 = \{\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n\}$, V_2 的基 $B_2 = \{e_1 \ e_2 \ \dots \ e_n\}$, 且

$$\sigma(\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n) = (e_1 \ e_2 \ \dots \ e_n)A,$$

$$\tau(e_1, e_2, \dots, e_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)B,$$

则 $\tau\sigma$ 关于 V_1 的基 B_1 和 $\sigma\tau$ 关于 V_2 的基 B_2 对应的矩阵分别为 BA 和 AB . 由于恒等变换 I 关于任何基对应于单位矩阵 E 因此

$$BA = AB = E$$

我们称可逆线性映射对应的矩阵为可逆矩阵(也称非奇异矩阵). 于是有下面的定义.

定义 4.5 设 $A \in M_n(F)$ 如果存在 $B \in M_n(F)$ 使得

$$BA = AB = E, \quad (4-15)$$

则称矩阵 A 是可逆的, 并把 B 叫做 A 的逆矩阵.

可逆矩阵 A 的逆矩阵是唯一的, 事实上, 如果 B, C 都是 A 的逆矩阵, 即

$$BA = AB = CA = AC = E,$$

则必有

$$B = BE = B(AC) = (BA)C = EC = C.$$

由于逆矩阵的唯一性, 我们可以把 A 的逆矩阵 B 记作 $B = A^{-1}$.

在 (4-15) 式中, A, B 的地位是相同的, 所以此时也说 B 是可逆的, 且 $A = B^{-1}$. 由此即得

$$(A^{-1})^{-1} = A. \quad (4-16)$$

定理 4.2 设 $B, A \in M_n(F)$ 若 $AB = E$ 则必有 $BA = E$ 即 A, B 互为逆矩阵.

证 设 A, B 分别是 $\sigma, \tau \in L(V, V)$ 关于 V 的基 $B = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 所对应的矩阵, 则

$$AB = E \Leftrightarrow \sigma\tau = I_V.$$

因此要证明 $BA = E$, 只要证明 $\tau\sigma = I_V$, 由 $\text{秩}(\sigma\tau) = \text{秩}(I_V) = n \leq \min(\text{秩}(\sigma), \text{秩}(\tau))$ 得

$$\text{秩}(\sigma) = \text{秩}(\tau) = n,$$

因此 $\forall \alpha \in V$ 均存在唯一的 $\beta \in V$ 使得 $\tau(\beta) = \alpha$, 从而

$$(\tau\sigma)(\alpha) = (\tau\sigma)(\tau(\beta)) = \tau(\sigma\tau(\beta)) = \tau(I_V(\beta)) = \tau(\beta) = \alpha,$$

故 $\tau\sigma = I_V$, 定理得证.

例 1 单位矩阵 E 是可逆的, 且 $E^{-1} = E$ 因为 $EE = E$.

例 2 主对角元都是非零数的对角阵是可逆的, 且

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1} \end{pmatrix}.$$

有很多矩阵是不可逆的,例如有一行全为零的方阵

$$A = \begin{pmatrix} 0 & 0 \\ a_{21} & a_{22} \end{pmatrix}$$

就是不可逆的. 因为 $\forall B \in M_2(F)$ 都有

$$AB = \begin{pmatrix} 0 & 0 \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ * & * \end{pmatrix} \neq E.$$

关于逆矩阵的运算性质, 现在先讲以下几点.

若 $A, B \in M_n(F)$ 都可逆, $\lambda \in F (\lambda \neq 0)$ 则 λA 和 AB 也可逆, 且

$$(\lambda A)^{-1} = \lambda^{-1} A^{-1}, \quad (4-17)$$

$$(AB)^{-1} = B^{-1} A^{-1}. \quad (4-18)$$

事实上, 由

$$(\lambda A)(\lambda^{-1} A^{-1}) = (\lambda \lambda^{-1})(AA^{-1}) = 1E = E$$

及

$$(AB)(B^{-1} A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$$

即得 (4-17)(4-18) 式.

对于 k 个可逆矩阵 A_1, A_2, \dots, A_k 及可逆矩阵 A , 用数学归纳法易证:

$$(A_1 A_2 \dots A_k)^{-1} = A_k^{-1} \dots A_2^{-1} A_1^{-1},$$

$$(A^k)^{-1} = (A^{-1})^k. \quad (4-19)$$

通常把 $(A^{-1})^k$ 记作 A^{-k} , 这样, $\forall k, m \in \mathbb{Z}$, 都有

$$A^k A^m = A^{k+m}, \quad (A^k)^m = A^{mk}. \quad (4-20)$$

当 A 可逆时, 由 $AB = AC$ 可以推出 $B = C$, 即乘法的消去律成立, 这是因为在等式 $AB = AC$ 的两边都左乘 A^{-1} , 即得 $B = C$.

关于逆矩阵运算要注意: A, B 都可逆, 而 $A + B$ 不一定可逆, 即使 $A + B$ 可逆, $(A + B)^{-1} \neq A^{-1} + B^{-1}$, 读者不难举出这样的例子(可用对角阵举例).

下面, 我们介绍求逆矩阵的一种方法(在本章节 4.7 和第 5 章中还要介绍两种常用的求逆矩阵的方法).

由定理 4.1 可知, $\sigma \in L(V, V)$ 关于基 $B = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 对应的矩阵为 A , 则

$$\sigma(\alpha) = \beta \quad \text{对应于} \quad AX = b,$$

其中 X, b 分别是 α, β 关于基 B 的坐标. 如果方程组 $AX = b$ 对任意的 b 都有唯一解, 也就是对任意的 β , 方程 $\sigma(\alpha) = \beta$ 有唯一解, 则 σ 是双射, 即 σ 可逆, 从而 A 是可逆矩阵, $AX = b$ 的唯一解为

$$X = A^{-1}b.$$

如果存在 $b \in F^n$, 使方程组 $AX = b$ 无解, 或有解而不唯一, 则 A 所对应的映射 σ 不是满射, 或 σ 不是单射, 这都表明 A 不是可逆矩阵.

现在我们举例说明, 如何由方程组 $AX = b$ 的解, 来判断 A 是否可逆, 可逆时如何求逆.

例3 判断下面矩阵 A, B 是否可逆. 如果可逆, 求其逆矩阵.

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix}.$$

解 以 A 为系数矩阵的非齐次线性方程组 $AX = b$ 对任意的 b (其分量为 b_1, b_2, b_3), 用高斯消元法可将其增广矩阵

$$(A, b) = \left(\begin{array}{ccc|c} 1 & -1 & 1 & b_1 \\ 0 & 1 & 2 & b_2 \\ 1 & 0 & 4 & b_3 \end{array} \right)$$

化为

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 4b_1 + 4b_2 - 3b_3 \\ 0 & 1 & 0 & 2b_1 + 3b_2 - 2b_3 \\ 0 & 0 & 1 & -b_1 - b_2 + b_3 \end{array} \right).$$

因此, 对任意的 b , 方程组 $AX = b$ 有唯一解

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4b_1 + 4b_2 - 3b_3 \\ 2b_1 + 3b_2 - 2b_3 \\ -b_1 - b_2 + b_3 \end{pmatrix} = \begin{pmatrix} 4 & 4 & -3 \\ 2 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

所以 A 是可逆矩阵, 其逆矩阵就是上式右边的三阶矩阵(因为 A 的逆矩阵是唯一的), 即

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

同理, 将

$$(B, \mathbf{b}) = \left(\begin{array}{ccc|c} 1 & -1 & 1 & b_1 \\ 0 & 1 & 2 & b_2 \\ 1 & 0 & 3 & b_3 \end{array} \right)$$

化为

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & b_1 + b_2 \\ 0 & 1 & 2 & b_2 \\ 0 & 0 & 0 & -b_1 - b_2 + b_3 \end{array} \right).$$

当 $-b_1 - b_2 + b_3 \neq 0$ 时, $B\mathbf{X} = \mathbf{b}$ 无解; 当 $-b_1 - b_2 + b_3 = 0$ 时, $B\mathbf{X} = \mathbf{b}$ 有无穷多解, 故 B 是不可逆矩阵.

用同样的方法, 读者不难证明主对角元都非零的上(下)三角矩阵 $A = (a_{ij})_{n \times n}$ 是可逆的, 其逆矩阵仍是上(下)三角矩阵, 且 A^{-1} 的主对角元为 $\frac{1}{a_{ii}}, i = 1, 2, \dots, n$.

下面再从矩阵 A 满足的关系式来判断其可逆性和求逆.

例 4 设方阵 B 为幂等矩阵(即 $B^2 = B$, 从而 $\forall k \in \mathbf{N}^*, B^k = B$), $A = E + B$, 证明 A 是可逆阵, 且

$$A^{-1} = \frac{3E - A}{2}.$$

证 由 $B = A - E$ 及 $B^2 = B$ 得

$$B^2 = (A - E)^2 = A^2 - 2A + E = A - E,$$

所以

$$A^2 - 3A = A(A - 3E) = -2E, \text{ 即 } A\left(\frac{3E - A}{2}\right) = E.$$

因此, 根据定理 4.2 可知, A 是可逆矩阵, 且

$$A^{-1} = \frac{3E - A}{2}.$$

域 F 上的全体 n 阶矩阵 $M_n(F)$ 关于矩阵的乘法只构成含么半群, 但全体 n 阶可逆矩阵(记作 $GL_n(F)$)关于矩阵的乘法构成群(这是因为 $GL_n(F)$ 关于矩阵的乘法是封闭的; 其单位元是单位矩阵; 可逆矩阵的逆矩阵仍可逆), 这个群叫作 n 阶线性群.

4.6 矩阵的转置

定义 4.6 把矩阵 $A = (a_{ij})_{m \times n}$ 的行列依次互换得到的一个 $n \times m$ 矩

阵称为 A 的转置矩阵, 记作 $A^T = (a'_{ji})_{n \times m}$, 其中 $a'_{ji} = a_{ij}$ ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$), 即

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & & & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

矩阵的转置运算满足以下运算律:

- (1) $(A^T)^T = A$; (2) $(A + B)^T = A^T + B^T$;
 (3) $(\lambda A)^T = \lambda A^T$ (λ 是数量); (4) $(AB)^T = B^T A^T$;
 (5) $(A^T)^{-1} = (A^{-1})^T$.

前三条是显然的, 下面证明(4)(5). 设:

$A = (a_{ij})_{m \times n}$, $A^T = (a'_{ji})_{n \times m}$, $B = (b_{ij})_{n \times s}$, $B^T = (b'_{ji})_{s \times n}$
 (其中 $a'_{ji} = a_{ij}$, $b'_{ji} = b_{ij}$) 则 $(AB)^T$ 与 $B^T A^T$ 都是 $s \times m$ 矩阵, 且

$$(B^T A^T)_{ji} = \sum_{k=1}^n b'_{jk} a'_{ki} = \sum_{k=1}^n a_{ik} b_{kj} = (AB)_{ij} = (AB)^T_{ji},$$

$$j = 1, \dots, s; \quad i = 1, \dots, m,$$

故 $(AB)^T = B^T A^T$.

由 $AA^{-1} = E$ (AA^{-1})^T = $(A^{-1})^T A^T = E$, 及定理 4.2 即得(5).

对于多个矩阵的乘积的转置, 用数学归纳法容易证明:

$$(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T.$$

定义 4.7 设 $A = (a_{ij})_{n \times n}$, 如果 $\forall i, j = 1, \dots, n$ 均有 $a_{ji} = a_{ij}$, 则 A 称为对称矩阵; 如均有 $a_{ji} = -a_{ij}$, 则 A 称为反对称矩阵.

n 阶为反对称矩阵 A 的主对角元都为零, 因为由 $a_{ii} = -a_{ii}$ 即得 $a_{ii} = 0$ ($i = 1, 2, \dots, n$).

根据定义 4.6 和 4.7, 容易证明:

A 为对称矩阵的充要条件是 $A^T = A$;

A 为反对称矩阵的充要条件是 $A^T = -A$.

例 1 设 A 是 $m \times n$ 矩阵, 则 $A^T A$ 和 AA^T 都是对称矩阵. 因为 $A^T A$ 是 n 阶矩阵, 且 $(A^T A)^T = A^T (A^T)^T = A^T A$; 同理 AA^T 是 m 阶对称矩阵.

例 2 设 A, B 分别是 n 阶对称和反对称矩阵, 则 $AB + BA$ 是反对称矩阵. 这是因为

$$(AB + BA)^T = B^T A^T + A^T B^T = (-B)A + A(-B) = -(AB + BA)$$

必须注意, 两个对称矩阵 A 和 B 的乘积不一定是对称矩阵. 因为,

$(AB)^T = B^T A^T = BA$, 而 BA 不一定等于 AB .

利用运算律(5), 容易证明: 可逆的对称(反对称)矩阵的逆矩阵也是对称(反对称)矩阵.

4.7 矩阵的初等变换和初等矩阵

我们在节 1.8 中讲过, 用高斯消元法解线性方程组, 其消元步骤是对增广矩阵作下面 3 种行初等运算:

- (1) 以非零常数 c 乘矩阵的某一行;
- (2) 将矩阵的某一行乘以非零常数 c 后加到另一行;
- (3) 将矩阵的某两行对换位置.

矩阵的这三种行初等运算统称为矩阵的初等行变换 (1) 称为倍乘行变换 (2) 称为倍加行变换 (3) 称为对换行变换.

研究矩阵的其他一些问题, 有时也要对矩阵作上述三种类型的初等列变换. 矩阵的初等行、列变换统称为矩阵的初等变换, 我们用:

$e_i^c(A)$, $e_{ij}^c(A)$, $e_{ij}(A)$ 依次表示以非零常数 c 乘 A 的第 i 行, 将 A 的第 i 行乘 c 加到第 j 行, 将 A 的第 i, j 行对换;

$\bar{e}_i^c(B)$, $\bar{e}_{ij}^c(B)$, $\bar{e}_{ij}(B)$ 依次表示以非零常数 c 乘 B 的第 i 列, 将 B 的第 j 列乘 c 加到第 i 列, 将 B 的第 j, i 列对换.

对 $m \times n$ 矩阵作的上述 6 种初等变换, 都是集 $M_{m \times n}(F)$ 上的线性变换, 它们也可以用矩阵表示, 但按节 4.1 中的方法来做, 那将复杂而不方便. 人们通过对矩阵乘法运算特点的分析研究, 发现引入下面定义的三种“初等矩阵”, 就可把矩阵的初等行、列变换用“初等矩阵”与该矩阵作乘法运算来实现.

定义 4.8 将单位矩阵 E 作一次初等变换所得到的矩阵称为初等矩阵, 与三种初等行、列变换对应的三类初等矩阵为:

- (1) 将单位矩阵第 i 行(或列)乘 c , 得到初等倍乘矩阵 $E_i(c)$;
- (2) 将单位矩阵第 i 行乘 c 加到第 j 行, 或将第 j 列乘 c 加到第 i 列, 得到初等倍加矩阵 $E_{ij}(c)$;
- (3) 将单位矩阵的第 i, j 行(或列)对换, 得到初等对换矩阵 E_{ij} .

即:

$$E_i(c) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & c & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \quad i \text{ 行},$$

$$E_{ij}(c) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & c & & 1 \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix} \quad \begin{matrix} i \text{ 行} \\ j \text{ 行} \end{matrix},$$

$$E_{ij} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & & 1 \\ & & & \ddots & \\ & & 1 & & 0 \\ & & & & \ddots \\ & & & & & 1 \end{pmatrix} \quad \begin{matrix} i \text{ 行} \\ j \text{ 行} \end{matrix}.$$

读者不难验证,三种初等矩阵左(右)乘矩阵 A (B) 是对 A (B) 作相应的初等行(列)变换,也就是由单位矩阵变为初等矩阵的行(列)变换,即

$$e_i^c(A) = E_i(c)A, \quad e_{ij}^c(A) = E_{ij}(c)A, \quad e_{ij}(A) = E_{ij}A, \\ \bar{e}_i^c(B) = BE_i(c), \quad \bar{e}_{ij}^c(B) = BE_{ij}(c), \quad \bar{e}_{ij}(B) = BE_{ij}.$$

例如:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \begin{pmatrix} a_{11} + ca_{31} & a_{12} + ca_{32} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix},$$

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{13} & b_{12} \\ b_{21} & b_{23} & b_{22} \\ b_{31} & b_{33} & b_{32} \end{pmatrix}.$$

初等矩阵都是可逆矩阵. 这是因为对初等矩阵再作一次同类型的初等变换都可化为单位矩阵, 即

$$E_i\left(\frac{1}{c}\right)E_i(c) = E, \quad E_{ij}(-c)E_{ij}(c) = E, \quad E_{ij}E_{ij} = E.$$

由此可见, 初等矩阵的逆矩阵是同类初等矩阵, 且

$$E_i^{-1}(c) = E_i\left(\frac{1}{c}\right), \quad E_{ij}^{-1}(c) = E_{ij}(-c), \quad E_{ij}^{-1} = E_{ij}.$$

对于任一个 $m \times n$ 型矩阵 $A = (a_{ij})_{m \times n}$, 如果它的第一列元素不全为零, 则可以对它作若干次初等行变换, 将其化为行简化阶梯形矩阵

$$U = \begin{pmatrix} c_{11} & \cdots & 0 & \cdots & 0 & \cdots & c_{1n} \\ 0 & \cdots & c_{2j_2} & \cdots & 0 & \cdots & c_{2n} \\ \cdots & & & & & & \cdots \\ 0 & \cdots & 0 & \cdots & c_{rj_r} & \cdots & c_{rn} \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \cdots & & & & & & \cdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

(其中 $c_{11} = c_{2j_2} = \cdots = c_{rj_r} = 1$ 它们分别是第 1, 2, ..., r 行的第一个非零元素, 而且它们所在的第 1, j_2 , ..., j_r 列的其它元素全为零, 第 $r+1$, ..., m 行是全零行) 把所作的初等行变换所对应的初等矩阵依次记为 P_1, P_2, \dots, P_k , 则上述结论可以表示为

$$P_k \cdots P_2 P_1 A = U.$$

如果 A 是 n 阶可逆矩阵, 则上式左端的 $k+1$ 个矩阵的乘积也是可逆矩阵, 即 U 是可逆矩阵, 从而 U 不能有全零行, 于是 U 必是 n 阶单位矩阵. 这样我们就有下面的定理.

定理 4.3 对任一个可逆矩阵 A , 都可以作若干次初等行变换将其化为单位矩阵 E , 即存在初等矩阵 P_1, P_2, \dots, P_k , 使得

$$P_k \cdots P_2 P_1 A = E. \quad (4-21)$$

推论 1 可逆矩阵 A 可以表示为若干个初等矩阵的乘积.

证 由 (4-21) 式得

$$A = (P_k \dots P_2 P_1)^{-1} = P_1^{-1} P_2^{-1} \dots P_k^{-1},$$

其中 $P_1^{-1}, P_2^{-1}, \dots, P_k^{-1}$ 仍是初等矩阵, 推论 1 得证.

由 (4-21) 式又可知

$$A^{-1} = (P_k \dots P_2 P_1) = P_k \dots P_2 P_1 E. \quad (4-22)$$

于是由 (4-21) (4-22) 式, 又可得以下推论.

推论 2 如果对可逆矩阵 A 和同阶单位矩阵 E 作同样的初等行变换, 那么当 A 变为 E 时, E 就变为 A^{-1} , 即

$$(A, E) \xrightarrow{\text{初等行变换}} (E, A^{-1}).$$

推论 2 提供了用初等行变换求逆矩阵的方法. 由 (4-22) 式又可得

$$AP_k \dots P_2 P_1 = E, \quad EP_k \dots P_2 P_1 = A^{-1}.$$

这两式子表明, 也可用初等列变换求逆矩阵, 即

$$\begin{pmatrix} A \\ E \end{pmatrix} \xrightarrow{\text{初等列变换}} \begin{pmatrix} E \\ A^{-1} \end{pmatrix}.$$

用初等变换的方法求逆矩阵是求逆矩阵的常用方法. 但必须注意, 用初等行(列)变换求逆矩阵时, 要始终作行(列)变换.

例 用初等行变换求矩阵 A 的逆矩阵.

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

解

$$\begin{aligned} (A, E) &= \left(\begin{array}{ccc|ccc} 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ -1 & -1 & -1 & 0 & 0 & 1 \end{array} \right) \\ &\xrightarrow{[1] \leftrightarrow [2]} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ -1 & -1 & -1 & 0 & 0 & 1 \end{array} \right) \\ &\xrightarrow{[3] + [1]} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right) \\ &\xrightarrow{\begin{matrix} [1] + [3] \times (-2) \\ [2] + [3] \end{matrix}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & -1 & -2 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right) \end{aligned}$$

$$\xrightarrow{\begin{matrix} [2] \times (1/2) \\ [1] + [2] \times (-1) \end{matrix}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{5}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right),$$

所以
$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & -\frac{3}{2} & -\frac{5}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 1 \end{pmatrix}.$$

注意,求逆矩阵容易算错,求得 A^{-1} 后,应验算 $AA^{-1} = E$.

4.8 矩阵的秩 相抵标准形

线性映射 σ 的秩是一个重要的概念,它刻画了 σ 的值域(子空间)的维数. 由于线性映射可以用矩阵表示,所以矩阵也有秩的概念. 我们先从不同角度来引进矩阵的秩的概念.

定义 4.9 设 $A = (a_{ij})_{m \times n}$ 是线性映射 σ 对应的矩阵,我们把秩(σ)也称为矩阵 A 的秩,记作秩(A)或 $r(A)$.

矩阵 A 的 n 个列向量的秩称为 A 的列秩;

矩阵 A 的 m 个行向量的秩称为 A 的行秩.

矩阵 A 的秩与 A 的列秩及 A 的行秩之间有着密切的关系. 下面我们将证明这三者是相等的.

定理 4.4 设矩阵 $A = (a_{ij})_{m \times n}$ 是 $\sigma \in L(V_1, V_2)$ 关于 V_1 和 V_2 的基 $B_1 = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 和 $B_2 = \{e_1, e_2, \dots, e_m\}$ 对应的矩阵,则秩(A) = A 的列秩.

证 分别记 A 的秩、行秩、列秩为 r, r_r, r_c , 其中 r 为 $\text{Im} \sigma$ 的维数,即 $\{\alpha(\varepsilon_1), \dots, \alpha(\varepsilon_n)\}$ 的极大无关组所含向量所含向量的个数. 由

$$\alpha(\varepsilon_k) = \sum_{i=1}^m a_{ik} e_i,$$

$\alpha(\varepsilon_k)$ 对应于 (a_{1k}, \dots, a_{mk}) , 后者就是 A 的第 k 个列向量, 所以 $\{\alpha(\varepsilon_1), \dots, \alpha(\varepsilon_n)\}$ 的极大无关组所含向量的个数就是 A 的列向量的极大无关组所含向量的个数, 也就是 A 的列秩 r_c , 即 $r = r_c$.

定理 4.5 对于任一矩阵 $A = (a_{ij})_{m \times n}$, 都有

A 的行秩 = A 的列秩.

证 设 A 的行秩为 r_r , 即 A 有 r_r 个行向量线性无关. 不妨就是前 r_r 行 $\alpha_1, \dots, \alpha_{r_r}$. 于是

$$\alpha_i = \sum_{k=1}^{r_r} c_{ik} \alpha_k, \quad i = 1, 2, \dots, m,$$

即

$$(\alpha_{i1} \dots \alpha_{in}) = \sum_{k=1}^{r_r} (c_{ik} \alpha_{k1} \dots \alpha_{kn}),$$

或

$$a_{ij} = \sum_{k=1}^{r_r} c_{ik} a_{kj}, \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

也就是

$$(\alpha_{1j} \dots \alpha_{m_j})^T = \sum_{k=1}^{r_r} (c_{1k} \dots c_{mk})^T \alpha_{kj}.$$

这个式子说明: A 所有的列向量都可以被 r_r 个向量 $(c_{1k} \dots c_{mk})^T (k = 1, \dots, r_r)$ 线性表示, 所以 A 的列秩 $r_c \leq r_r$.

同理 A^T 的列秩 $\leq A^T$ 的行秩, 也就是 A 的行秩 $\leq A$ 的列秩, 于是 $r_c = r_r$. 即 A 的行秩 = A 的列秩.

以上我们证明了 秩 $(A) = A$ 的列秩 = A 的行秩. 由此可见:

对任意的 $A \in M_{m \times n}(F)$, 都有秩 $(A^T) =$ 秩 (A) ; 又当 $m \leq n$ 时 秩 $(A) = A$ 的行秩 $\leq m$; 当 $n \leq m$ 时 秩 $(A) = A$ 的列秩 $\leq n$.

阶梯形矩阵的秩等于它的非零行的行数. 例如

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} \\ 0 & 0 & 0 & a_{34} & a_{35} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

其中 $a_{11}a_{22}a_{34} \neq 0$, 读者不难证明: A 的三个非零行的行向量是线性无关的. A 的列向量组 $\{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5\}$ 的极大线性无关组为 $\{\beta_1, \beta_2, \beta_4\}$, 故 秩 $(A) = A$ 的行秩 = A 的列秩 = 3.

将矩阵 A 作初等行变换化为阶梯形矩阵 U , 则秩 A 等于 U 的非零行的行数, 因为矩阵的秩有下列重要的“不变”性质.

定理 4.6 初等行变换和初等列变换都不改变矩阵的秩.

证 仅就初等行变换的情形给以证明. 设 $m \times n$ 矩阵 A 的 m 个行向量为

$\{\alpha_1, \alpha_2, \dots, \alpha_m\}$,

(1) 将 A 的第 i, j 行对换得到 B , 则 B 与 A 的行向量组相同, 故 A, B 的行秩相等.

(2) 将 A 的第 i 行乘非零常数 c 得到 B , 则 B 的行向量组: $\beta_i = c\alpha_i, \beta_k = \alpha_k (k \neq i)$. 因此 B 与 A 的行向量组可以互相线性表示, 所以 A 与 B 的行秩相等.

(3) 将 A 的第 i 行乘常数 c 加到第 j 行得到 B , 则 B 的行向量组: $\beta_j = c\alpha_i + \alpha_j, \beta_k = \alpha_k (k \neq j)$. 相应地也有 $\alpha_j = \beta_j - c\beta_i, \alpha_k = \beta_k (k \neq j)$. 因此 A 与 B 的行向量组可以互相线性表示, 所以 A 与 B 的行秩相等.

这就证明了初等行变换不改变矩阵的秩. 同理, 初等列变换也不改变矩阵的秩. 这个定理给出了求矩阵(线性映射)的秩的一个简单算法.

例 1 求下列矩阵 A 的秩

$$A = \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ -1 & 2 & 1 & 3 & 6 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & 1 & 2 \end{pmatrix}.$$

解 对 A 作初等行变换, 容易将其化为阶梯形矩阵 U , 即

$$A \longrightarrow \begin{pmatrix} 1 & -1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 2 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = U,$$

于是 $\text{秩}(A) = \text{秩}(U) = 3$.

如果对上例中的阶梯形矩阵 U 再作列变换: 将第一列乘某些常数加到第 2, 3, 4, 5 列, 使第一行其余元素全为零, 再将第 2 列乘某些常数加到第 3, 4, 5 列, 使第二行其余元素全为零, 然后将第 4 列乘 -2 加到第 5 列, 并将第 3, 4 列对换, 即得

$$U \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = U_3.$$

任何秩为 3 的 4×5 矩阵通过初等行、列变换都可化为上面最简单的阶梯形矩阵 U_3 . 一般情况的结论如下.

定理 4.7 若 $\text{秩}(A_{m \times n}) = r$, 则存在可逆矩阵 P 和 Q , 使得

$$PAQ = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & & & & & & \cdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & & & & & & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} = U_r, \quad (4-24)$$

其中 r 个非零行向量为 n 维单位向量 e_1, e_2, \dots, e_r .

证 先将 A 作初等行变换化为有 r 个非零行的阶梯形矩阵 U (非零行的第一个非零元都是 1), 再对 U 作初等列变换, 即得 U_r .

集合 $M_{m \times n}(F)$ 中的矩阵按秩分类, 当 $m \leq n$ 时, 有 $m+1$ 类; 当 $m > n$ 时, 有 $n+1$ 类. 矩阵按秩分类确定了一种等价关系.

定义 4.10 设 $A, B \in M_{m \times n}(F)$ 如果 A 经过初等变换可以化为 B , 就称 A 相抵于 B (或说 A 等价于 B), 记作 $A \cong B$.

根据定义, 所谓 $A \cong B$, 即存在 m 和 n 阶初等矩阵 P_1, P_2, \dots, P_s 和 Q_1, Q_2, \dots, Q_t , 使得

$$P_s \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_t = B.$$

由于可逆矩阵 P, Q 可以表示为若干初等矩阵的乘积, 所以 $A \cong B$ 的充要条件是存在可逆矩阵 P 和 Q , 使得 $PAQ = B$.

集合 $M_{m \times n}(F)$ 中的相抵关系是一个等价关系. 因为

(1) $E_m A E_n = A$, 即 $A \cong A$;

(2) 若 $A \cong B$, 则存在可逆阵 P 和 Q , 使得 $PAQ = B$, 于是 $P^{-1} B Q^{-1} = A$, 从而 $B \cong A$;

(3) 若 $A \cong B, B \cong C$, 则存在可逆矩阵 P, S 和 Q, T , 使得 $PAQ = B, SBT = C$, 于是就有 $(SP)A(QT) = C$, 从而 $A \cong C$.

(4-24) 式的右端的矩阵 U_r 称为 A 的相抵 (或等价) 标准形. 所有秩为 r 的 $m \times n$ 矩阵都相抵于 U_r . 所以 $M_{m \times n}(F)$ 中所有秩为 r 的矩阵, 对于相抵关系构成的等价类最简单的代表元为 U_r , 于是集合 $M_{m \times n}(F)$ 关于 \cong 的商集为 $\{\overline{O}, \overline{U}_1, \overline{U}_2, \dots, \overline{U}_k\}$, 其中 $k = \min\{m, n\}$, O 为零矩阵.

根据相抵的定义及定理 4.6, 4.7 可见, $\forall A, B \in M_{m \times n}(F)$,

$A \cong B$ 的充要条件是 $\text{秩}(A) = \text{秩}(B)$.

定理 4.8 设 $A \in M_n(F)$ 则下列命题等价:

(1) A 可逆; (2) $\text{秩}(A) = n$; (3) A 的 n 个列 (行) 向量线性无关;

(4) 齐次线性方程组 $AX = 0$ 只有零解.

证 (1) \Rightarrow (2) 因为可逆矩阵 A 经初等行变换可化为 E 故 $r(A) = n$.

(2) \Rightarrow (3) 因为 $r(A) = A$ 的列秩 (行秩), 故 (3) 成立.

(3) \Rightarrow (4) 设 A 的 n 个列向量为 $\beta_1, \beta_2, \dots, \beta_n$, 则齐次线性方程组 $AX = 0$ 等价于

$$x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n = 0. \quad (*)$$

因为 $\beta_1, \beta_2, \dots, \beta_n$ 线性无关, 由 (*) 式即得 x_1, x_2, \dots, x_n 全为零, 故 (4) 成立.

(4) \Rightarrow (1) 反证法, 设 A 不可逆, 则对 A 作初等行变换必出现全为零元的行, 从而 $AX = 0$ 有非零解, 矛盾.

由于 n 阶可逆矩阵的秩等于 n , 所以可逆矩阵也称为满秩矩阵.

定理 4.9

$$r(A+B) \leq r(A) + r(B), \quad (4-25)$$

$$r(AB) \leq \min(r(A), r(B)). \quad (4-26)$$

这个定理可由定义 4.9 及 A, B 所对应的线性映射 σ, τ 的和 $\sigma + \tau$ 与积 $\sigma\tau$ 的秩的结论 (3-10) (3-8) 而直接得到. 它也可以利用定理 4.4, 4.5 以及向量组的秩的有关性质加以证明. 例如证明 (4-26): 由

$$\begin{aligned} AB &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{ns} \end{bmatrix} \\ &= \begin{bmatrix} | & | & & | \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1s} \\ b_{21} & b_{22} & \dots & b_{2s} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{ns} \end{bmatrix} \end{aligned}$$

得到 AB 的 s 个列向量

$$\gamma_j = \sum_{i=1}^n b_{ij}\alpha_i, \quad j = 1, 2, \dots, s$$

均可由 A 的列向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性表示, 所以,

$$r(AB) = AB \text{ 的列秩} \leq A \text{ 的列秩} = r(A). \quad (1)$$

同理, 记 B 的 n 个行向量为 $\beta_1, \beta_2, \dots, \beta_n$, 则 AB 的 m 个行向量

$$\xi_i = \sum_{j=1}^n a_{ij}\beta_j, \quad i = 1, 2, \dots, m,$$

均可由 B 的行向量组线性表示, 所以

$$\text{秩}(AB) = AB \text{ 的行秩} \leq B \text{ 的行秩} = \text{秩}(B). \quad (2)$$

综合(1)(2), 即得(4-26)式成立.

如果 P, Q 分别是 m, n 阶可逆矩阵, A 为 $m \times n$ 矩阵, 则

$$\text{秩}(PA) = \text{秩}(AQ) = \text{秩}(PAQ) = \text{秩}(A). \quad (4-27)$$

这是因为: $PA \cong A$, 所以 $\text{秩}(PA) = \text{秩}(A)$; 或利用(4-26)式来证明, 事实上, 由 $A = P^{-1}(PA)$, 即得

$$\text{秩}(A) \leq \text{秩}(PA) \leq \text{秩}(A),$$

所以 $\text{秩}(PA) = \text{秩}(A)$. 同理可证(4-27)式中的其它等式.

例2 若 A 为 3×4 矩阵, 则 $A^T A$ 是不可逆矩阵.

证 由于 $\text{秩}(A^T A) \leq \text{秩}(A) \leq 3$, 而 $A^T A$ 是4阶矩阵, 故 $A^T A$ 是不可逆矩阵.(在节6.2例3中将证明, 当 A 为实矩阵时, $\text{秩}(A^T A) = \text{秩}(A)$.)

4.9 分块矩阵

把一个大型矩阵分成若干小块, 构成一个分块矩阵, 从而把大型矩阵的运算化为若干小型矩阵的运算, 使运算更为简明, 这是矩阵运算中的一个重要技巧. 下面通过例子说明矩阵如何分块及分块矩阵的运算方法.

例如, 一个5阶矩阵可用纵横垂直的两条线将其分成4块, 构成一个分块矩阵, 即

$$\left(\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & -1 & 0 \\ \hline 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{array} \right) = \begin{pmatrix} E_3 & A_1 \\ O & A_2 \end{pmatrix}, \quad (4-28)$$

其中 E_3 为三阶单位矩阵, O 为 2×3 零矩阵,

$$A_1 = \begin{pmatrix} 0 & 2 \\ 1 & -3 \\ -1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}.$$

一般, 对于 $m \times n$ 矩阵 A , 如果在行的方向分成 s 块, 在列的方向分成 t 块, 就得到 A 的一个 $s \times t$ 分块矩阵, 记作

$$A = (A_{kl})_{s \times t},$$

其中 A_{kl} ($k = 1, \dots, s; l = 1, \dots, t$) 称为 A 的子块.

常用的分块矩阵 除了上面(4-28)式的 2×2 分块矩阵 还有以下几种:
将 $A = (A_{ij})_{m \times n}$ 按行分块或按列分块为:

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix}, \quad A = (\beta_1 \beta_2 \dots \beta_n). \quad (4-29)$$

当 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 中非零元素都集中在主对角线附近时,有时可将 A 分块成下面的对角块矩阵(又称准对角矩阵)

$$A = \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{mm} \end{pmatrix}, \quad (4-30)$$

或记作 $A = \text{diag}(A_{11}, A_{22}, \dots, A_{mm})$ 其中 $A_{ii} (i = 1, 2, \dots, m)$ 是 r_i 阶方阵;
 $A_{ij} = O (i \neq j)$.

下面讨论分块矩阵的运算.

1 分块矩阵的加法

设分块矩阵 $A = (A_{kl})_{s \times t}$, $B = (B_{kl})_{s \times t}$ 如果 A 与 B 的对应子块 A_{kl} 和 B_{kl} 都是同型矩阵,则

$$A + B = (A_{kl} + B_{kl})_{s \times t}. \quad (4-31)$$

2 分块矩阵的数量乘法

设分块矩阵 $A = (A_{kl})_{s \times t}$, λ 是一个数,则

$$\lambda A = (\lambda A_{kl})_{s \times t}. \quad (4-32)$$

3 分块矩阵的乘法

设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{n \times p}$ 如果把 A, B 分别分块为 $r \times s$ 和 $s \times t$ 分块矩阵,且 A 的列的分块法与 B 的行的分块法相同,则

$$AB = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & & & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1t} \\ B_{21} & B_{22} & \dots & B_{2t} \\ \dots & & & \dots \\ B_{s1} & B_{s2} & \dots & B_{st} \end{pmatrix} = C = (C_{kl})_{r \times t},$$

其中 C 是 $r \times t$ 分块矩阵,且

$$C_{kl} = A_{k1}B_{1l} + A_{k2}B_{2l} + \dots + A_{ks}B_{sl}, \quad k = 1, 2, \dots, r; l = 1, \dots, t. \quad (4-33)$$

可以证明(略),用分块乘法和不分块乘法求得的 AB 相同.

前面定理4.9证明中的 γ_j 也可能为将 A 列分块为 $[\alpha_1, \alpha_2, \dots, \alpha_n]$ 与矩阵 B 相乘的结果;同理, ξ_i 是将 B 按行分块为 $\begin{bmatrix} \beta_1 \\ \dots \\ \beta_n \end{bmatrix}$ 后与不分块的 A 相乘的结果.

例1 对(4-28)式所给的矩阵 A ,用分块乘法求 A^2 .

解

$$A^2 = \begin{bmatrix} E_3 & A_1 \\ O & A_2 \end{bmatrix}^2 = \begin{bmatrix} E_3 & A_1 + A_1 A_2 \\ O & A_2^2 \end{bmatrix}.$$

读者容易计算出 $A_1 + A_1 A_2$ 和 A_2^2 得到

$$A^2 = \begin{bmatrix} 1 & 0 & 0 & 2 & 10 \\ 0 & 1 & 0 & 2 & -14 \\ 0 & 0 & 1 & -5 & -1 \\ 0 & 0 & 0 & 17 & 8 \\ 0 & 0 & 0 & 8 & 17 \end{bmatrix}.$$

例2 设 A 是 $m \times n$ 矩阵, B 是 $n \times s$ 矩阵.将 B 按列分块为 $1 \times s$ 分块矩阵,将 A 视为 1×1 分块矩阵,则

$$AB = A(B_1, B_2, \dots, B_s) = (AB_1, AB_2, \dots, AB_s).$$

例3 若 n 阶矩阵 C 和 D 分块成同型对角块矩阵,即

$$C = \text{diag}(C_1, C_2, \dots, C_s), \quad D = \text{diag}(D_1, D_2, \dots, D_s),$$

其中 C_i 和 D_i 是同阶方阵($i = 1, 2, \dots, s$)则

$$CD = \text{diag}(C_1 D_1, C_2 D_2, \dots, C_s D_s).$$

用矩阵分块乘法,可以证明矩阵的一些重要命题.

例4 证明: n 阶可逆上三角矩阵 A 的逆矩阵也是上三角矩阵.

证 对 n 作数学归纳法: $n = 1$ 时 $(a)^{-1} = \left(\frac{1}{a}\right)$ 结论成立(一阶矩阵可视为特殊的三角矩阵).假设命题对 $n - 1$ 阶可逆的上三角矩阵成立.对 n 阶矩阵 A ,设 B 为 A 的逆矩阵,并把 A, B 分块为同样的 2×2 分块矩阵,即

$$A = \begin{bmatrix} a_{11} & \alpha \\ O & A_1 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \beta \\ \gamma & B_1 \end{bmatrix},$$

其中 A_1 是 $n - 1$ 阶可逆的上三角矩阵, B_1 是 $n - 1$ 矩阵, α, β 为 $1 \times (n - 1)$ 矩阵, γ 为 $(n - 1) \times 1$ 矩阵,于是

$$AB = \begin{bmatrix} a_{11}b_{11} + \alpha\gamma & a_{11}\beta + \alpha B_1 \\ A_1\gamma & A_1B_1 \end{bmatrix} = \begin{bmatrix} 1 & O \\ O & E_{n-1} \end{bmatrix}.$$

这样,由 $A_1\gamma = O$ 得 $\gamma = A_1^{-1}O = O$,由 $A_1B_1 = E_{n-1}$ 得 $B_1 = A_1^{-1}$. 根据归纳假设, B_1 是上三角矩阵,所以 $A^{-1} = B$ 是上三角矩阵.

4 分块矩阵的转置

$s \times t$ 分块矩阵 $A = (A_{kl})_{s \times t}$ 的转置 A^T 为 $t \times s$ 分块矩阵. 如果记 $A^T = (B_{lk})_{t \times s}$, 则 $B_{lk} = A_{kl}^T$, $l = 1, 2, \dots, t$; $k = 1, 2, \dots, s$. 要注意的是, 此时大的分块形状要转置, 其中每一小分块也要转置.

例如

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A^T = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}.$$

5 可逆分块矩阵的逆矩阵

对角块矩阵(准对角矩阵) $A = \text{diag}(A_1, A_2, \dots, A_m)$ 可逆的充要条件是子块 A_i ($i = 1, 2, \dots, m$) 都可逆, 且 A^{-1} 也是对角块矩阵, 即

$$A^{-1} = \text{diag}(A_1^{-1}, A_2^{-1}, \dots, A_m^{-1}).$$

例5 设 n 阶矩阵 A 分块为

$$A = \begin{bmatrix} B & O \\ C & D \end{bmatrix},$$

其中 B, D 分别为 k 阶和 m 阶矩阵. 证明: A 可逆的充要条件为 B, D 都是可逆矩阵, 并求 A^{-1} .

证 充分性: 由习题 36 知, $r(A) \geq r(B) + r(D) = k + m$ (因为 D 中线性相关的非零行向量添加了 C 中的分量后, 可能成为线性无关的向量), 又 $r(A) \leq k + m$ 所以 $r(A) = k + m$, 即 A 可逆.

必要性: 如果 B 的行秩小于 k 或 D 的列秩小于 m , 则 $r(A) < k + m$, 与 A 可逆矛盾, 故 B, D 满秩, 皆可逆. 设

$$A^{-1} = \begin{bmatrix} X & Y \\ Z & T \end{bmatrix},$$

其中 X, T 分别为 k 阶和 m 阶矩阵. 于是由

$$\begin{bmatrix} B & O \\ C & D \end{bmatrix} \begin{bmatrix} X & Y \\ Z & T \end{bmatrix} = \begin{bmatrix} BX & BY \\ CX + DZ & CY + DT \end{bmatrix} = \begin{bmatrix} E_k & O \\ O & E_m \end{bmatrix},$$

得

$$BX = E_k, \quad \text{故} \quad X = B^{-1};$$

$$BY = O, \quad \text{故} \quad Y = B^{-1}O = O;$$

$$CX + DZ = O, \quad \text{故} \quad Z = -D^{-1}CB^{-1};$$

$$CY + DT = E_m, \quad \text{故} \quad DT = E_m, T = D^{-1}.$$

所以

$$A^{-1} = \begin{pmatrix} B^{-1} & O \\ -D^{-1}CB^{-1} & D^{-1} \end{pmatrix}.$$

最后,我们讨论一下分块矩阵的初等变换,这里仅以 2×2 分块矩阵为例来讨论.对于分块矩阵

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

可以同样地定义它的三种初等(列)变换,并相应地定义三类分块初等矩阵:

$$(1) \text{分块倍乘矩阵} \quad \begin{pmatrix} C_1 & O \\ O & E_m \end{pmatrix} \quad \text{或} \quad \begin{pmatrix} E_k & O \\ O & C_2 \end{pmatrix};$$

$$(2) \text{分块倍加矩阵} \quad \begin{pmatrix} E_k & O \\ C_3 & E_m \end{pmatrix} \quad \text{或} \quad \begin{pmatrix} E_k & C_4 \\ O & E_m \end{pmatrix};$$

$$(3) \text{分块对换矩阵} \quad \begin{pmatrix} O & E_m \\ E_k & O \end{pmatrix},$$

其中 E_k, E_m 表示 k, m 阶单位矩阵, C_1, C_2 是可逆矩阵.

分块初等矩阵左乘或右乘分块矩阵 A , 在保证可乘的情况下,其作用于节 4.7 中讲的首等矩阵左乘或右乘矩阵 A 的作用是一样的.分块矩阵的初等变换也是矩阵的一个重要技巧.下面举两个例子.

例 6 设 n 阶矩阵 A 分块表示为

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

其中 A_{11}, A_{22} 为方阵.证明:如果 A 和 A_{11} 可逆,则 $A_{22} - A_{21}A_{11}^{-1}A_{12}$ 可逆,并求 A^{-1} .

解 对 A 作倍加变换把 A 的第一行左乘 $(-A_{21}A_{11}^{-1})$ 并加到第二行(相应的分块倍加初等阵为 P_1),使 A 化为上三角块矩阵 B ,即

$$P_1 A = \begin{pmatrix} E_1 & O \\ -A_{21}A_{11}^{-1} & E_2 \end{pmatrix} A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} = B,$$

其中 E_1, E_2 分别是与 A_{11}, A_{22} 同阶的单位阵.由 $P_1 A = B$ 可逆及例 5 的转置(其可逆性不变)得 $A_{22} - A_{21}A_{11}^{-1}A_{12}$ 也可逆.记

$$Q = A_{22} - A_{21}A_{11}^{-1}A_{12},$$

为了求 A^{-1} , 先对 B 作行变换将其化为对角块矩阵 C , 即

$$P_2 B = \begin{pmatrix} E_1 & -A_{12}Q^{-1} \\ O & E_2 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ O & Q \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ O & Q \end{pmatrix} = C,$$

于是 $P_2 P_1 A = C$ 故 $A^{-1} = [(P_2 P_1)^{-1} C]^{-1} = C^{-1} P_2 P_1$, 即

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & O \\ O & Q^{-1} \end{pmatrix} \begin{pmatrix} E_1 & -A_{12} Q^{-1} \\ O & E_2 \end{pmatrix} \begin{pmatrix} E_1 & O \\ -A_{21} A_{11}^{-1} & E_2 \end{pmatrix} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}.$$

其中

$$D_{11} = A_{11}^{-1} + A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1},$$

$$D_{12} = -A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1},$$

$$D_{21} = -(A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1},$$

$$D_{22} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}.$$

例7 证明 若 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 的各阶左上角子块矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}, \quad k = 1, 2, \dots, n$$

都可逆, 则存在主对角元全为 1 的下三角矩阵 L 和上三角矩阵 U , 使得 $A = LU$ (通常称为矩阵 A 的 $L-U$ 分解).

证 由于主对角元全为 1 的下三角矩阵可逆, 其逆矩阵也是主对角元全为 1 的下三角矩阵, 因此只要证明存在主对角元全为 1 的下三角矩阵 S , 使得 $SA = U$. 下面用数学归纳法证明之.

$n = 1$ 时, 命题显然成立; 假设命题对满足条件的 $n-1$ 阶矩阵成立. 对于 n 阶矩阵 A , 把 A 分块为

$$A = \begin{pmatrix} A_1 & \alpha_1 \\ \alpha_2 & a_{nn} \end{pmatrix},$$

其中 A_1 为满足命题条件的 $n-1$ 阶矩阵, 根据归纳假设, 存在 $n-1$ 阶主对角元全为 1 的下三角矩阵 S_1 和上三角矩阵 U_1 , 使得

$$S_1 A_1 = U_1.$$

于是先对 A 作倍加初等变换将其化为上三角块矩阵, 即

$$PA = \begin{pmatrix} E_{n-1} & O \\ -\alpha_2 A_1^{-1} & 1 \end{pmatrix} A = \begin{pmatrix} A_1 & \alpha_1 \\ O & a_{nn} - \alpha_2 A_1^{-1} \alpha_1 \end{pmatrix},$$

再取

$$Q = \begin{pmatrix} S_1 & O \\ O & 1 \end{pmatrix},$$

即得

$$QPA = \begin{pmatrix} U_1 & S_1 \alpha_1 \\ O & a_{nn} - \alpha_2 A_1^{-1} \alpha_1 \end{pmatrix} = U.$$

其中 U 为上三角阵, $S = QP$ 是主对角元全为 1 的下三角阵, 故存在 $L = S^{-1}$ 和 U , 使得

$$A = LU.$$

4.10 基的变换矩阵与坐标变换

我们在节 2.5 中讲过,有限维线性空间中向量在不同基下的坐标一般是不同的.下面讨论基变换与坐标变换的问题.

设 $B_1 = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 与 $B_2 = \{\beta_1, \beta_2, \dots, \beta_n\}$ 是线性空间 $V(F)$ 的任意两组基, B_2 中每个基向量被基 B_1 表示为

$$\begin{cases} \beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \dots + a_{n1}\alpha_n \\ \beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \dots + a_{n2}\alpha_n \\ \dots\dots\dots \\ \beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \dots + a_{nn}\alpha_n. \end{cases} \quad (4-34)$$

将上式用矩阵形式地表示为

$$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \quad (4-35)$$

我们把(4-34)式右端的矩阵 $A = (a_{ij})_{n \times n}$ 叫做基 B_1 变为基 B_2 的变换矩阵(或称过渡矩阵).

由(4-34)(4-35)式可见,变换矩阵 A 是(4-34)中基向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 的系数矩阵的转置.它的第 j 列是 β_j 关于基 B_1 的坐标.由于 $\beta_1, \beta_2, \dots, \beta_n$ 是线性无关的,所以 A 中对应的 n 个列向量也线性无关.因此基的变换矩阵 A 是可逆矩阵.

下面论述同一个向量在两组不同基下的坐标间的关系.

定理 4.10 设上述基 B_1 变为基 B_2 的变换矩阵为 A , 如果 $\xi \in V(F)$, 且在 B_1, B_2 下的坐标分别为

$$\xi_{B_1} = X = (x_1, \dots, x_n)^T, \quad \xi_{B_2} = Y = (y_1, \dots, y_n)^T,$$

则

$$Y = A^{-1}X. \quad (4-36)$$

证 由已知条件, ξ 可以形式地表示为

$$\xi = (\alpha_1, \dots, \alpha_n)X = (\beta_1, \dots, \beta_n)Y.$$

将已知条件 $(\beta_1, \dots, \beta_n) = (\alpha_1, \dots, \alpha_n)A$ 代入上式右端,得

$$\xi = ((\alpha_1, \dots, \alpha_n)A)Y = (\alpha_1, \dots, \alpha_n)(AY).$$

由于 ξ 在基 $B_1 = (\alpha_1, \dots, \alpha_n)$ 下的坐标是唯一的, 所以 $AY = X$, 从而

$$Y = A^{-1}X.$$

例1 已知 $\beta_1 = (1, 1, 1), \beta_2 = (0, 2, 1), \beta_3 = (0, 0, 4)$ 是 \mathbf{R}^3 一组基, 求 \mathbf{R}^3 的自然基 $\{e_1, e_2, e_3\}$ 变为基 $\{\beta_1, \beta_2, \beta_3\}$ 的变换矩阵 A .

解 由

$$\begin{cases} \beta_1 = e_1 + e_2 + e_3 \\ \beta_2 = 0 + 2e_2 + e_3 \\ \beta_3 = 4e_3 \end{cases}$$

即得自然基 $\{e_1, e_2, e_3\}$ 变为基 $\{\beta_1, \beta_2, \beta_3\}$ 的变换矩阵

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 4 \end{pmatrix} = \begin{pmatrix} | & | & | \\ \beta_1 & \beta_2 & \beta_3 \\ | & | & | \end{pmatrix}.$$

例2 已知 \mathbf{R}^3 的两组基 $B_1 = \{\alpha_1, \alpha_2, \alpha_3\}, B_2 = \{\beta_1, \beta_2, \beta_3\}$ 为

$$\alpha_1 = (1, 1, 0), \alpha_2 = (0, 1, 1), \alpha_3 = (1, 0, 1),$$

$$\beta_1 = (1, 0, 0), \beta_2 = (1, 1, 0), \beta_3 = (1, 1, 1).$$

ξ 在基 B_1 下的坐标为 $X = (1, 0, 2)^T$, 试求 ξ 在基 B_2 下的坐标 Y .

解 先求基 B_1 变为基 B_2 的变换矩阵 A . 根据变换矩阵的定义, 需要分别将 $\beta_1, \beta_2, \beta_3$ 用基 B_1 线性表示, 这样需要解系数矩阵为

$$P = \begin{pmatrix} | & | & | \\ \alpha_1 & \alpha_2 & \alpha_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

的三个非齐次线性方程组. 此时可按下述方法求变换矩阵 $A = (a_{ij})_{3 \times 3}$, 即由

$$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)A,$$

将基 B_1 和 B_2 中的基向量以列向量的形式代入上式, 得矩阵方程

$$Q = PA.$$

于是变换矩阵

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

根据(4-36)式, 得 ξ 在基 B_2 下的坐标 Y 为

$$Y = A^{-1}X = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}.$$

此题的另一种解法:先由 $\xi_{B_1} = X = (1 \ 0 \ 2)^T$ 得

$$\xi = \alpha_1 + 2\alpha_3 = (3 \ 1 \ 2).$$

再令 $\xi = y_1\beta_1 + y_2\beta_2 + y_3\beta_3$, 可解得 $\xi_{B_2} = Y = (y_1 \ y_2 \ y_3)^T = (2 \ -1 \ 2)^T$.

例3 已知 $R[x]_4$ 的两组基 $B_1 = \{g_1 \ g_2 \ g_3 \ g_4\}$ 和 $B_2 = \{h_1 \ h_2 \ h_3 \ h_4\}$ 为

$$g_1 = 1 + x + x^2 + x^3, \quad h_1 = 1 - x + x^3,$$

$$g_2 = x + x^2 + x^3, \quad h_2 = 1 + x,$$

$$g_3 = x^2 + x^3, \quad h_3 = 1 + x - x^2,$$

$$g_4 = x^3, \quad h_4 = 1 - x + x^2.$$

又知 $\mu(x)$ 在基 B_1 下的坐标 $X = (1 \ -2 \ 0 \ 1)^T$, 试求:

(1) 基 B_1 变为基 B_2 的变换矩阵 C ;

(2) $\mu(x)$ 在基 B_2 下的坐标 $Y = (y_1 \ y_2 \ y_3 \ y_4)^T$.

解 (1) 对于 $R[X]_4$ 的自然基 $B_0 = \{1 \ x \ x^2 \ x^3\}$ 按等式

$$(g_1 \ g_2 \ g_3 \ g_4) = (1 \ x \ x^2 \ x^3)A_1, \quad (1)$$

$$(h_1 \ h_2 \ h_3 \ h_4) = (1 \ x \ x^2 \ x^3)A_2, \quad (2)$$

易得基 B_0 变为基 B_1 和基 B_2 的变换矩阵 A_1 和 A_2 为

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

由 ① 式又得

$$(1 \ x \ x^2 \ x^3) = (g_1 \ g_2 \ g_3 \ g_4)A_1^{-1}. \quad (3)$$

将 ③ 式代入 ② 式 则得

$$(h_1 \ h_2 \ h_3 \ h_4) = (g_1 \ g_2 \ g_3 \ g_4)(A_1^{-1}A_2).$$

所以基 B_1 变为基 B_2 的变换矩阵为

$$C = A_1^{-1}A_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -2 & 0 & 0 & -2 \\ 1 & -1 & -2 & 2 \\ 1 & 0 & 1 & -1 \end{pmatrix}.$$

(2) 根据 (4-36) 式 $\mu(x)$ 在基 B_2 下的坐标 Y 为

$$Y = C^{-1}X = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 3 & 3 & 2 & 1 \\ -2 & -\frac{5}{2} & -2 & -1 \\ -1 & -\frac{3}{2} & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 2 \\ 1 \end{pmatrix}.$$

习 题

1. 求下列线性映射关于自然基所对应的矩阵.

(1) $\alpha(x_1, x_2, x_3) = (x_1 + x_2, x_1 - x_3, x_2)$;

(2) $\alpha(x_1, x_2) = (x_1 - x_2, x_1, x_1 + x_2)$;

(3) $\alpha(p(x)) = p(x+1) - p(x)$, $p(x) \in R[x]_n$;

(4) 第3章习题2(2)(4);

(5) σ 是 n 维线性空间 $V(F)$ 的恒等变换(对 $V(F)$ 的任一组基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$).

2*. 设 C 为全体复数在实数域上构成的线性空间, $z_0 = a_0 + b_0i \in C$, 定义 C 到 C 的映射为: σ_{Z_0} 为 $\sigma_{Z_0}(Z) = Z_0Z$, 证明: σ_{Z_0} 是线性的, 并求:

(1) σ_{Z_0} 关于基 $B_1 = \{1, i\}$ 所对应的矩阵;

(2) σ_{Z_0} 关于基 $B_2 = \{1, 1+i\}$ 所对应的矩阵.

3. 设 $D: R[x]_4 \rightarrow R[x]_3$.

$$D(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2.$$

(1) 证明 D 是线性的;

(2) 求 D 关于 $R[x]_4$ 的基 $B_1 = \{1, x, x^2, x^3\}$ 和 $R[x]_3$ 的基 $B_2 = \{1, x, x^2\}$ 所对应的矩阵;

(3) 求 $p(x) = 4 + 3x - x^2 + 2x^3$ 的象 $D(p(x))$ 关于基 B_2 的坐标;

(4) 如果把 D 看作 $R[x]_4$ 到自身的线性变换, 那么 D 关于基 B_1 所对应的矩阵是什么?

4. 对第3章习题3中的投影变换 p 和镜象变换 φ :

(1) 分别求 p 和 φ 关于 R^3 的自然基 $\{e_1, e_2, e_3\}$ 所对应的矩阵(此时令 $\omega = (c_1, c_2, c_3)^T$);

(2) 求 R^3 的一组基, 使 p 和 φ 关于这组基所分别对应的矩阵都是对角阵.

5. 计算 $2A + 3B, AB - BA$, 其中 A, B 为

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

6. 已知：

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

证明： $A^2 = A$, $B^k = B$ ($k \geqslant 2$), $AB = BA = O$.

7. 计算下列的矩阵乘积：

$$(1) \begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \\ 4 & 2 & 1 \end{pmatrix}; \quad (2) \begin{pmatrix} 2 & -1 & 0 \\ 1 & 1 & 3 \\ 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix};$$

$$(3) \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; \quad (4) \begin{pmatrix} 2 \\ 1 \end{pmatrix} (a \ b);$$

$$(5) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ c_1 & c_2 & \cdots & c_n \end{pmatrix}; \quad (6) \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \cdots & \cdots & \cdots \\ a_n & b_n & c_n \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix};$$

$$(7) \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}; \quad (8) \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

8. 计算 $QP(PAQ)^k$ (k 为正整数), 其中 P, A, Q 为

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

9. 证明 若 $AB = BA$, $AC = CA$ 则 A, B, C 为同阶方阵, 且

$$A(BC) = (BC)A, \quad A(B+C) = (B+C)A.$$

10. A, B 都是 n 阶矩阵, 问下列等式成立的充分条件是什么?

$$(1) (A+B)^3 = A^3 + 3A^2B + 3AB^2 + B^3;$$

$$(2) (A+B)(A-B) = A^2 - B^2.$$

11. 计算下列矩阵的幂：

$$(1) \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}^n; \quad (2) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^n;$$

$$(3) \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}^n; \quad (4) \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{pmatrix}^n.$$

12. 求平方等于零矩阵的所有二阶矩阵.

13. 求所有与 A 可交换的矩阵, 其中

$$(1) A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}; \quad (2) A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 1 & -2 \end{pmatrix}.$$

14. 设对角矩阵 $A = \text{diag}(a_1, a_2, \dots, a_n)$ 其中 $a_i \neq a_j (i \neq j)$ 证明: 与 A 可交换的矩阵必是对角阵.

15. 设 $f(x), g(x) \in F[x], A, B \in M_n(F)$ 证明:

$$f(A)g(A) = g(A)f(A);$$

(1) 如果 $AB = BA$ 则 $f(A)g(B) = g(B)f(A)$;

(2) 设 $f(x) = 1 + x + \dots + x^{m-1}, g(x) = 1 - x, A = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$, 计算 $f(A)g(A) = ?$

16. 证明: 若 A, B 皆为 n 阶下三角矩阵, 则 $AB = C$ 也是下三角矩阵, 且 $c_{ii} = a_{ii}b_{ii}$, 其中 a_{ii}, b_{ii}, c_{ii} 为 A, B, C 的主对角元, $i = 1, 2, \dots, n$.

17. 证明: 若 A 是主对角元全为 1 的上三角矩阵, 则 A^2 也是主对角元全为 1 的上三角矩阵.

$$18. \text{ 设 } S = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid a, b, c \in \mathbf{Z} \right\}.$$

(1) 验证 S 对矩阵的加法和乘法构成一个环;

(2) 求 S 的幂零元素 (幂零矩阵, 即 $\exists k \in \mathbf{N}^*$ 使 $A^k = O$) 与幂等元素 (幂等矩阵, 即 $A^2 = A$), 并证明 S 的幂零元素组成的集合是 S 的一个子环.

19. 设 $B = \{\alpha_1, \alpha_2, \alpha_3\}$ 是三维向量空间 V 的一组基, $\sigma \in L(V, V)$ 关于基 B 所对应的矩阵为 A , 求 $\text{Im}\sigma$ 和 $\text{Ker}\sigma$:

$$(1) A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 1 \\ 1 & 2 & 5 \end{pmatrix}; \quad (2) A = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ -2 & 4 & -2 \end{pmatrix}.$$

20. “若 A 是 n 阶可逆矩阵, 则方程组 $AX = b$ 对任意的 b 有唯一解, 且其解 $X = A^{-1}b$ ”. 利用这个结论, 判断下列矩阵是否可逆? 如可逆, 并求其逆矩阵.

$$(1) A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}; \quad (2) B = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix};$$

$$(3) C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 3 & 2 & 3 \end{pmatrix}; \quad (4) D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix};$$

(5) 证明 : 上三角矩阵 $A = (a_{ij})_{n \times n}$ ($i > j$ 时 $a_{ij} = 0$) 可逆的充要条件是 $a_{ii} \neq 0, i = 1, 2, \dots, n$.

21. 设方阵 A 满足 $A^2 - A - 2E = O$, 证明 :

(1) A 和 $E - A$ 都是可逆矩阵 , 并求它们的逆矩阵 ;

(2) $A + E$ 和 $A - 2E$ 不可能同时都是可逆的.

22. 设 A 为幂零矩阵(即 $\exists k$, 使 $A^k = O$), 证明 $E - A$ 是可逆矩阵 , 并求其逆矩阵.

23. 设 A, B 是同阶非奇异矩阵(即可逆矩阵), 证明下列等式彼此等价 :

$$AB = BA; \quad AB^{-1} = B^{-1}A;$$

$$B^{-1}A^{-1} = A^{-1}B^{-1}; \quad A^{-1}B = BA^{-1}.$$

24. 计算 $AB^T, A^TB, A^TA + B^TB$, 其中 A, B 为 :

$$A = \begin{bmatrix} 5 & -2 & 1 \\ 3 & -4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & 2 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

25. 证明 $(A_1 A_2 \dots A_k)^T = A_k^T \dots A_2^T A_1^T$.

26. 对于任意的 n 阶矩阵 A , 证明 :

(1) $A + A^T$ 是对称矩阵, $A - A^T$ 是反对称矩阵 ;

(2) A 可表示为对称矩阵与反对称矩阵之和.

27. A, B 都是 n 阶对称矩阵, 证明 :

(1) $A + B, A - 2B$ 也都是对称矩阵 ;

(2) AB 是对称矩阵 $\Leftrightarrow A$ 与 B 可交换.

28. 证明 : 若 A 是实对称矩阵, 且 $A^2 = O$ 则 $A = O$.

29. 已知 A 是 n 阶对称矩阵, B 是 n 阶反对称矩阵.

(1) 问 : A^k, B^k 是对称矩阵或反对称矩阵吗 ?

(2) 证明 : $AB^3 + B^3A$ 是反对称矩阵.

30. 证明 : 可逆的对称(反对称)矩阵的逆矩阵仍是对称(反对称)矩阵.

31. 用初等变换法求下列矩阵的逆矩阵 :

$$(1) \begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & -2 \\ 2 & -2 & 1 \end{bmatrix};$$

$$(2) \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix};$$

$$(3) \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \end{bmatrix};$$

$$(4) \begin{bmatrix} 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \\ 1 & -1 & 1 & -1 \end{bmatrix};$$

$$(5) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}; \quad (6) \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix};$$

$$(7) \begin{pmatrix} 1 & a & a^2 & a^3 \\ 0 & 1 & a & a^2 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad (8) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 4 & 0 & 0 & 0 \end{pmatrix}.$$

32. 已知

$$A = \begin{pmatrix} a & a & 1 \\ a & 1 & a \\ 1 & a & a \end{pmatrix}, \quad 0 < a < 1.$$

证明 :A 是可逆阵 ,并求 A^{-1} .

33. 利用逆矩阵(或用初等变换法) ,解下列矩阵方程 :

$$(1) \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}; \quad (2) \begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 0 \\ -1 & 2 & -2 \end{pmatrix} X = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 3 & 1 \end{pmatrix};$$

$$(3) A \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

34. 利用逆矩阵 ,解线性方程组

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_2 + 3x_3 = 1 \\ x_1 - x_2 = 2. \end{cases}$$

35. 用初等变换法 ,求下列矩阵的秩 :

$$(1) \begin{pmatrix} 1 & 1 & 2 \\ 4 & 5 & 5 \\ 5 & 8 & 1 \\ -1 & -2 & 2 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & 3 & -2 & 5 & 4 \\ 1 & 4 & 1 & 3 & 5 \\ 1 & 4 & 2 & 4 & 3 \\ 2 & 7 & -3 & 6 & 13 \end{pmatrix};$$

$$(3) \begin{pmatrix} 3 & 2 & -1 & -3 \\ 2 & -1 & 3 & 1 \\ 4 & 5 & -5 & -6 \end{pmatrix}; \quad (4) \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

36. 设 $A, B, C \in M_n(F)$, 证明:

$$(1) \text{秩} \begin{pmatrix} A & O \\ O & B \end{pmatrix} = \text{秩}(A) + \text{秩}(B);$$

$$(2) \text{秩}(A) + \text{秩}(B) \leq \text{秩} \begin{pmatrix} A & C \\ O & B \end{pmatrix} \leq \text{秩}(A) + \text{秩}(B) + \text{秩}(C).$$

37. 证明: 矩阵添加一列(或一行), 其秩或不变, 或增加 1.

38. 设 A 是 $s \times n$ 矩阵, B 是 A 的前 m 行构成的 $m \times n$ 矩阵, 证明:

$$\text{秩}(B) \geq \text{秩}(A) + m - s.$$

39*. 设 $A \in M_n(F)$, $\text{秩}(A) = 1$, 证明:

$$(1) A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} [b_1 \ b_2 \ \dots \ b_n];$$

(2) $A^2 = kA$, 其中 k 是 A 的主对角元之和.

40. 设 A 是 $m \times n$ 矩阵 ($m \leq n$), $\text{秩}(A) = m$, 证明: 存在 $n \times m$ 矩阵 B , 使 $AB = E$.

41. 设 $A, B \in M_n(F)$, $\text{秩}(A) + \text{秩}(B) \leq n$, 证明: 存在可逆阵 M , 使 $AMB = O$.

42. 设 $A, B \in M_{m \times n}(F)$, 证明: $A \cong B \Leftrightarrow \text{秩}(A) = \text{秩}(B)$.

43. 设 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 n 维空间 $V(F)$ 的一组基, $A \in M_{n \times k}(F)$, 且

$$(\beta_1, \beta_2, \dots, \beta_k) = (\alpha_1, \alpha_2, \dots, \alpha_n)A.$$

证明: $L(\beta_1, \beta_2, \dots, \beta_k)$ 的维数等于 $\text{秩}(A)$.

44. 设

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 2 & 3 & 0 \\ 1 & 2 & 0 & 4 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 4 \end{pmatrix}.$$

用矩阵分块的方法 (1) 计算 A^2, AB (2) 求 A^{-1} .

45. 设 $A, B, C \in M_n(F)$, $P = \begin{pmatrix} A & B \\ O & C \end{pmatrix}$ 给出 P 可逆的条件, 并求其逆矩阵.

46. 设 B, C 分别是 m 阶和 n 阶矩阵, $A = \begin{pmatrix} O & B \\ C & O \end{pmatrix}$,

(1) 证明: A 可逆 $\Leftrightarrow B, C$ 都可逆;

(2) 当 A 可逆时, 求 A^{-1} ;

(3) 求 $\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1}$.

47. 设 B 是 n 阶可逆矩阵,

$$C = \begin{pmatrix} 1 & 2 & \cdots & n \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

求一个矩阵 A , 使 $A \begin{pmatrix} B \\ C \end{pmatrix} = E_n$.

48. 将 n 阶矩阵 A 分块为

$$A = \begin{pmatrix} A_1 & B \\ C & a_{nn} \end{pmatrix},$$

其中 A_1 是 $n-1$ 阶矩阵, 设 A 与 A_1 都是可逆矩阵, 且已知 A_1^{-1} , 试求 A^{-1} (这种利用 A_1^{-1} 来求 A^{-1} 的方法称为加边法). 利用这个结果, 求

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

49*. 对下列矩阵作 $L-U$ 分解:

$$(1) \begin{pmatrix} 1 & -2 & 3 \\ 2 & 4 & 2 \\ 0 & 1 & 1 \end{pmatrix}; \quad (2) \begin{pmatrix} 3 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 2 & 3 \end{pmatrix}.$$

50. 已知 \mathbf{R}^3 的基 $B_1 = \{\alpha_1, \alpha_2, \alpha_3\}$ 变为基 $B_2 = \{\xi_1, \xi_2, \xi_3\}$ 的变换矩阵 $A = (a_{ij})_{3 \times 3}$, 求:

- (1) 基 $B_3 = \{\alpha_2, \alpha_1, \alpha_3\}$ 变为基 B_2 的变换矩阵;
 (2) 基 $B_4 = \{-\alpha_1, \alpha_2, \alpha_3\}$ 变为基 B_2 的变换矩阵;
 (3) 基 B_4 变为基 $B_5 = \{\xi_3, \xi_2, -\xi_1\}$ 的变换矩阵;
 (4) 基 B_4 变为基 $B_6 = \{\xi_1 + \xi_2, \xi_2 + \xi_3, \xi_3 + \xi_1\}$ 的变换矩阵.

51. 已知 \mathbf{R}^4 的一组基为: $\alpha_1 = (2, 1, -1, 1), \alpha_2 = (0, 3, 1, 0), \alpha_3 = (5, 3, 2, 1), \alpha_4 = (6, 6, 1, 3)$, 求一个非零向量 β , 使它关于这组基和关于自然基 $\{e_1, e_2, e_3, e_4\}$ 有相同的坐标.

52. 已知 \mathbf{R}^3 的基 $B_1 = \{\alpha_1, \alpha_2, \alpha_3\}$ 和 $B_2 = \{\beta_1, \beta_2, \beta_3\}$ 为

$$\alpha_1 = (1, 1, 1), \quad \alpha_2 = (0, 1, 1), \quad \alpha_3 = (0, 0, 1),$$

$$\beta_1 = (1, 0, 1), \quad \beta_2 = (1, 1, 0), \quad \beta_3 = (0, 1, 1).$$

- (1) 求基 B_1 变为基 B_2 的变换矩阵, 如果 ξ 关于基 B_1 的坐标为 $(1, -1, -1)^T$, 求 ξ 关于基 B_2 的坐标;
 (2) 求基 B_2 变为基 B_1 的变换矩阵.

53. 已知 \mathbf{R}^3 的基 $B_1 = \{\alpha_1, \alpha_2, \alpha_3\}$ 和 $B_2 = \{\beta_1, \beta_2, \beta_3\}$ 为

$$\alpha_1 = (1, 2, 1), \quad \alpha_2 = (2, 3, 3), \quad \alpha_3 = (3, 7, 1),$$

$$\beta_1 = (3, 1, 4), \quad \beta_2 = (5, 2, 1), \quad \beta_3 = (1, 1, -6).$$

- (1) 求 $\gamma = (3, 6, 2)$ 在基 B_1 下的坐标;
 (2) 求基 B_1 变为基 B_2 的变换矩阵;
 (3) 利用变换矩阵, 求 γ 在基 B_2 下的坐标.

补 充 题

1. 证明: 与任意 n 阶矩阵都可交换的矩阵 A 必是 n 阶数量矩阵.
 2. 矩阵 $A = (a_{ij})_{n \times n}$ 的主对角元之和称为矩阵 A 的迹, 记作

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

证明: 若 $A \in M_{n \times m}(F), B \in M_{m \times n}(F)$ 则 $\text{tr}(AB) = \text{tr}(BA)$.

3. 若 $A, B \in M_n(F)$ 则 $AB - BA \neq E$.

4*. 设

$$H = \left\{ \begin{pmatrix} a & b \\ -\frac{1}{b} & \frac{1}{a} \end{pmatrix} \mid a, b \in \mathbb{C} \text{ (复数域)} \right\}.$$

证明: H 是一个除环(含乘法单位元的无零因子环), 但不是域.

5. 设 $X = [x_1 \cdots x_n]^T, Y = [y_1 \cdots y_n]^T$ 且 $X^T Y = 2$ 证明:

(1) $(XY^T)^k = 2^{k-1}(XY^T), k \geq 2$;

(2) 如果 $A = E + XY^T$ 则 A 可逆 并求其逆矩阵.

6. 证明 对换变换可以通过若干次倍加和倍乘变换来实现.

7. 把二阶矩阵 $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ 表示为若干个倍加初等阵的乘积.

8. 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, ad - bc = 1$. 证明 矩阵 A 可以表示为若干个倍加初等矩阵的乘积.

9. 设 $A, B \in M_n(F)$, 证明 $r(AB) \geq r(A) + r(B) - n$. 并问: 如果 $A \in M_{m \times n}(F), B \in M_{n \times s}(F)$, 上述结论是否成立?

10. 设

$$B_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

$$B_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}.$$

(1) 证明 B_2 也是线性空间 $M_2(R)$ 的基;

(2) 求基 B_2 变为基 B_1 的变换矩阵;

(3) 求 $M_2(R)$ 的一组基 $B_3 = \{A_1, A_2, A_3, A_4\}$ 使得 $A_i^2 = A_i (i = 1, \dots, 4)$;

(4) 已知矩阵 A 关于基 B_2 的坐标为 $(1, 1, 1, 1)^T$ 求 A 关于基 B_3 的坐标.

11. 设

$$W_1 = \left\{ \begin{pmatrix} x & -x \\ y & z \end{pmatrix} \mid x, y, z \in F(\text{域}) \right\},$$

$$W_2 = \left\{ \begin{pmatrix} a & b \\ -a & c \end{pmatrix} \mid a, b, c \in F \right\}.$$

(1) 证明 W_1, W_2 是 $M_2(F)$ 的子空间 并求 $\dim W_1, \dim W_2, \dim(W_1 + W_2), \dim(W_1 \cap W_2)$;

(2) 求 $W_1 \cap W_2$ 的一组基 并求 $A = \begin{pmatrix} 3 & -3 \\ -3 & 1 \end{pmatrix}$ 关于这组基的坐标.

12. 设 $V_1 = L(f_1, f_2, f_3), V_2 = L(g_1, g_2)$ 其中

$$f_1 = 1 + 2x - x^2 - 2x^3, \quad f_2 = 3 + x + x^2 + x^3,$$

$$f_3 = -1 + x^2 - x^3,$$

$$g_1 = 2 + 5x - 6x^2 - 5x^3, \quad g_2 = -1 + 2x - 7x^2 + 3x^3.$$

求 $\dim(V_1 \cap V_2)$ 和 $(V_1 \cap V_2)$ 的基; $\dim(V_1 + V_2)$ 和 $(V_1 + V_2)$ 的基.

部分习题和补充题答案

习题

$$1.(1) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix};$$

$$(3) \begin{pmatrix} 0 & 1 & 1 & 1 & \cdots & 1 \\ & 0 & 2 & 3 & \cdots & C_{n-1}^{n-2} \\ & & 0 & 3 & \cdots & C_{n-1}^{n-3} \\ & & & \ddots & \ddots & \cdots \\ & & & & \ddots & C_{n-1}^1 \\ & & & & & 0 \end{pmatrix}; \quad (4) \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}; \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix};$$

$$(5) E_n.$$

$$2.(1) \begin{pmatrix} a_0 & -b_0 \\ b_0 & a_0 \end{pmatrix}; \quad (2) \begin{pmatrix} a_0 - b_0 & -2b_0 \\ b_0 & a_0 + b_0 \end{pmatrix}.$$

$$3.(2) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}; \quad (3) (3, -2, 6)^T; \quad (4) \begin{pmatrix} 0 & 1 & 0 & 0 \\ & 0 & 2 & 0 \\ & & 0 & 3 \\ & & & 0 \end{pmatrix}.$$

$$4.(1) E - A, E - 2A, A = \begin{pmatrix} c_1^2 & c_1 c_2 & c_1 c_3 \\ c_1 c_2 & c_2^2 & c_2 c_3 \\ c_1 c_3 & c_2 c_3 & c_3^2 \end{pmatrix};$$

(2) 基 $\{\alpha_1, \alpha_2, \alpha_3\}$: $\alpha_1 = \omega, \alpha_2 \perp \omega, \alpha_3 \perp \omega, \alpha_3$ 与 α_2 不共线.

$$7.(1) (9, 2, -1); \quad (2) (3, 6, 4)^T.$$

$$8. \begin{pmatrix} 2^{k+2} + (-1)^{k+1} 3 & -3 \cdot 2^{k+1} + 6(-1)^k \\ 2^{k+1} + (-1)^{k+1} 2 & -3 \cdot 2^k + 4(-1)^k \end{pmatrix}.$$

$$10. AB = BA.$$

$$11.(1) \begin{pmatrix} \cos n\varphi & \sin n\varphi \\ -\sin n\varphi & \cos n\varphi \end{pmatrix}; \quad (2) \begin{cases} (-1)^k E, & m = 2k \\ (-1)^k A, & m = 2k + 1; \end{cases}$$

提示 (3)(4)题利用 $(\lambda E + A)^k$ 的二项式展开式 (4-14).

$$(3) \begin{pmatrix} 1 & 2n & 2n^2 + n \\ & 1 & 2n \\ & & 1 \end{pmatrix}; \quad (4) \begin{pmatrix} a^n & a^{n-1}c_n^1 & a^{n-2}c_n^2 & a^{n-3}c_n^3 \\ & a^n & a^{n-1}c_n^1 & a^{n-2}c_n^2 \\ & & a^n & a^{n-1}c_n^1 \\ & & & a^n \end{pmatrix}.$$

$$12. \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, bc = -a^2.$$

$$13. (1) \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}; \quad (2) \begin{pmatrix} a & 0 & 0 \\ 0 & b+3c & 2c \\ 0 & c & b \end{pmatrix}.$$

$$15. (3) E - A^m. \quad 18. \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}; \begin{pmatrix} 0 & b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

$$19. (1) L\{\alpha_1 - \alpha_2 + \alpha_3, \alpha_2 + \alpha_3\}; L\{-2\alpha_1 - \frac{3}{2}\alpha_2 + \alpha_3\}; \\ (2) L\{\alpha_1 - \alpha_2 - 2\alpha_3\}; L\{-\alpha_1 + \alpha_3, 2\alpha_1 + \alpha_2\}.$$

$$20. (2)(4) \text{不可逆} (1) \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}; (3) \begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ 3 & 0 & -1 \\ -1 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

$$21. (1) \frac{A-E}{2}, \frac{A}{2}. \quad 22. E + A + \dots + A^{k-1}. \text{提示 找矩阵 } A \text{ 的一个多项式 } p(A) \text{ 使 } (E-A)p(A) = E.$$

$$24. \begin{pmatrix} -19 & -9 \\ -17 & -5 \end{pmatrix}, \begin{pmatrix} -21 & 10 & 3 \\ 14 & -4 & -4 \\ -5 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 47 & -28 & 6 \\ -28 & 24 & -6 \\ 6 & -6 & 3 \end{pmatrix}.$$

28. 提示 考虑 A^2 的主对角元即可.

30. 提示 利用 $(A^{-1})^T = (A^T)^{-1}$.

$$31. (1) \frac{1}{21} \begin{pmatrix} 5 & 6 & 2 \\ 6 & 3 & -6 \\ 2 & -6 & 5 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix};$$

$$(3) \begin{pmatrix} 1 & 0 & 0 & 0 \\ -a & 1 & 0 & 0 \\ a^2 & -a & 1 & 0 \\ -a^3 & a^2 & -a & 1 \end{pmatrix}; \quad (4) \frac{1}{20} \begin{pmatrix} -4 & 5 & 1 & 9 \\ 1 & 5 & -4 & -1 \\ 6 & -5 & 1 & -1 \\ 1 & -5 & 6 & -11 \end{pmatrix};$$

$$(5) \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}; \quad (6) \frac{1}{40} \begin{pmatrix} -9 & 11 & 1 & 1 \\ 1 & -9 & 11 & 1 \\ 1 & 1 & -9 & 11 \\ 11 & 1 & 1 & -9 \end{pmatrix};$$

$$(7) \begin{pmatrix} 1 & -a & 0 & 0 \\ 0 & 1 & -a & 0 \\ 0 & 0 & 1 & -a \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad (8) \begin{pmatrix} 0 & 0 & 0 & \frac{1}{4} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}.$$

$$32. \frac{1}{(2a+1)(1-a)} \begin{pmatrix} -a & -a & a+1 \\ -a & a+1 & -a \\ a+1 & -a & -a \end{pmatrix}.$$

$$33. (1) \begin{pmatrix} 8 & 3 \\ -3 & -1 \end{pmatrix}; (2) \begin{pmatrix} 1 & \frac{1}{3} \\ -1 & -\frac{1}{6} \\ -3 & -\frac{5}{6} \end{pmatrix}; (3) \begin{pmatrix} 3 & -3 & 1 \\ -1 & 2 & -1 \end{pmatrix}.$$

$$34. (1, -1, 1)^T. \quad 35. (1)(2)(3)3(4)4.$$

39. (1) 提示: A 的每个行向量成比例. (2) 利用 (1) 的结论.

40. 提示: 利用矩阵的相抵标准形.

$$44. (1) A^2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \text{ 其中 } A_1 = \begin{pmatrix} 5 & 12 \\ 12 & 29 \end{pmatrix}, A_2 = \begin{pmatrix} 4 & -4 & 1 \\ 0 & 4 & -4 \\ 0 & 0 & 4 \end{pmatrix};$$

$$(2) \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \text{ 其中 } B_1 = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}, B_2 = -\frac{1}{8} \begin{pmatrix} 4 & 2 & 1 \\ 0 & 4 & 2 \\ 0 & 0 & 4 \end{pmatrix}.$$

$$45. \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ O & C^{-1} \end{pmatrix}.$$

$$46. (2) \begin{pmatrix} O & C^{-1} \\ B^{-1} & O \end{pmatrix}; (3) \begin{pmatrix} 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \end{pmatrix}.$$

$$47. [B^{-1} \ P] P = [O_{n \times 1} \ D_{n \times 2}], \forall D.$$

$$49. (1) \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & \frac{1}{8} & 1 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 3 \\ 0 & 8 & -4 \\ 0 & 0 & \frac{3}{2} \end{pmatrix};$$

$$(2) \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 1 & 0 & 0 \\ 0 & \frac{6}{7} & 1 & 0 \\ 0 & 0 & \frac{14}{15} & 1 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & \frac{7}{3} & 1 & 0 \\ 0 & 0 & \frac{15}{7} & 1 \\ 0 & 0 & 0 & \frac{31}{15} \end{pmatrix}.$$

$$50. (1) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} A;$$

$$(2) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A;$$

$$(3) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}; (4) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

$$51. -e_1 - e_2 - e_3 + e_4.$$

$$52. (1) \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} [0 \ 1 \ -1]^T; (2) \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & 2 & 1 \end{pmatrix}.$$

$$53. (1) \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}; (2) \begin{pmatrix} -27 & -71 & -41 \\ 9 & 20 & 9 \\ 4 & 12 & 8 \end{pmatrix}; (3) \frac{1}{4} [153 \ -106 \ 83]^T.$$

补充题

1. 提示: 挑选一些简单的矩阵, 如 $\text{diag}(1 \dots, 1, c, 1 \dots, 1)$ 等与 A 可交换, 即可证得所要的结果.

3. 提示: 利用上题结果.

4. 提示 :证明 H 中非零矩阵可逆 ,乘法不可交换. 5.(2)提示 :考虑 A^2 与的 A 关系. 6. 提示 :只要证明初等对换阵可表示为初等倍加和初等倍乘阵的乘积 ,以二阶为例即可. 9. 提示 :利用矩阵的相抵标准形和(4-24)式

及习题 38 的结论 ,或者利用习题 36 的结论对分块矩阵 $\begin{bmatrix} E & O \\ O & AB \end{bmatrix}$ 作初等变换使之变为 $\begin{bmatrix} * & B \\ A & O \end{bmatrix}$ 型矩阵.

$$10.(2) \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

$$(3) \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \right\}; \quad (4) [1 \ 3 \ -1 \ 2]^T.$$

$$11.(1) B \ 3 \ A \ 2; \quad (2) \left\{ \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

$$12. 1: \{g_1\} \nsubseteq \{f_1, f_2, f_3, g_2\}.$$

第 5 章 行 列 式

在线性代数中,行列式是一个基本工具,用行列式可以判断矩阵是否可逆,向量组是否线性相关;以后研究线性映射的特征值以及欧氏空间中的一些几何问题也都离不开行列式.

本章主要内容有: n 阶行列式的定义及其性质;行列式按一行(行)的展开式;Cramer 法则;矩阵的行列式秩;利用方阵的行列式及其子式求方阵的逆矩阵;方阵 A 的行列式不等于零是 A 可逆和齐次线性方程组 $AX = 0$ 只有零解的充要条件等.

行列式的概念起源于线性方程组的求解.例如二元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad (5-1)$$

当 $a_{11}a_{22} - a_{12}a_{21} \neq 0$ 时,用消元法求得解为

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{b_2a_{11} - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}. \quad (5-2)$$

这个解的分子和分母都是 4 个元素的二次齐次多项式,它们都有两项(正负各一项).人们从(5-2)的表达式中受到启发,定义

$$D = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix} = d_{11}d_{22} - d_{12}d_{21} \quad (5-3)$$

(其中每项都是不同行、列的二元素之积),并称它为二阶行列式.如此(5-2)式可表示为

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad (5-2')$$

其中:

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad D_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

对于由 9 个元素 a_{ij} ($i, j = 1, 2, 3$) 排成的三行三列的式子定义为

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

$$-a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}, \quad (5-4)$$

并称它为三阶行列式. 它是 9 个元素的三次齐次多项式, 其中有 $3! = 6$ 项, 正负项各半, 每项都是不同行、列的三个元素之积. 这 6 项可用下面(5-5)式所示的方法得到.



$$(5-5)$$

如果把三元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

中未知元 x_1, x_2, x_3 的系数排成三阶行列式(称系数行列式)

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

那么, 当 $D \neq 0$ 时, 读者可以用消元法验证这个方程组的解为

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad x_3 = \frac{D_3}{D}. \quad (5-6)$$

其中 D_j 是由 b_1, b_2, b_3 替换 D 中第 j 列所得的三阶行列式, 即

$$D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.$$

对于 4 元乃至一般的 n 元 n 个一次方程构成的方程组是否也有类似于(5-2)及(5-6)式那样的结论呢? 答案是肯定的(见定理 5.6). 但是, 对于 $n > 3$ 的 n 阶行列式, 不能如(5-5)式那样定义. n 阶行列式可以用几种不同的方法来定义, 我们采用公理化方法定义.

5.1 n 阶行列式的定义

我们把数域 F 上的 n^2 个数 a_{ij} ($i, j = 1, 2, \dots, n$) 排成的 n 行 n 列并在两

边作两条直线的式子

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \quad \text{或} \quad |a_{ij}|_{n \times n} \quad (5-7)$$

称为数域 F 上的 n 阶行列式, 它表示从集合 $F^n \times F^n \times \cdots \times F^n$ (n 个) 到数域 F 的一个映射, 记作 $D(\alpha_1, \alpha_2, \dots, \alpha_n)$, 其中 $\alpha_j \in F^n$ 是 (5-7) 式中的第 j 个列向量 ($j = 1, 2, \dots, n$).

定义 5.1 数域 F 上的一个 n 阶行列式是取值于 F 的 n 个 n 维向量 $\alpha_1, \alpha_2, \dots, \alpha_n \in F^n$ 的一个函数, 而且 $\forall \alpha_i, \beta_i \in F^n$ 和 $\forall \lambda \in F$, 满足下列规则:

$$(1) D(\alpha_1, \dots, \lambda \alpha_i, \dots, \alpha_n) = \lambda D(\alpha_1, \dots, \alpha_i, \dots, \alpha_n);$$

$$(2) D(\alpha_1, \dots, \alpha_i + \beta_i, \dots, \alpha_n)$$

$$= D(\alpha_1, \dots, \alpha_i, \dots, \alpha_n) + D(\alpha_1, \dots, \beta_i, \dots, \alpha_n);$$

$$(3) D(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_n) = -D(\alpha_1, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_n);$$

$$(4) D(e_1, e_2, \dots, e_n) = 1, \text{ 即}$$

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & & \cdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

当 $n = 1$ 时, 一阶行列式 $D = |a_{11}| = a_{11}$.

定义中的 (1)(2) 表示 D 是线性的, (3) 表示第 i 列与第 j 列对换, 行列式反号, (4) 表示 D 关于 F^n 的自然基的函数值为 1, 即 n 阶单位矩阵的行列式等于 1.

关于满足上述 (1)~(4) 的映射 $D: F^n \rightarrow F$ 的唯一性问题这里不讨论. 先看下面的例子;

例 1 用定义计算下列行列式:

$$(1) D_n = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & & & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}, \quad (2) D_n = \begin{vmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & a_2 & 0 \\ \cdots & & & & \cdots \\ 0 & a_{n-1} & \cdots & 0 & 0 \\ a_n & 0 & \cdots & 0 & 0 \end{vmatrix},$$

其中 (1) 称为对角行列式 (即, 对角矩阵的行列式).

$$\begin{aligned} \text{解 } (1) D_n &= D(\alpha_{11}e_1, \alpha_{22}e_2, \dots, \alpha_{nn}e_n) \\ &= a_{11}a_{22}\dots a_{nn}D(e_1, e_2, \dots, e_n) = a_{11}a_{22}\dots a_{nn}. \end{aligned}$$

(2) 对 $D_n = D(\alpha_n e_n, \alpha_{n-1} e_{n-1}, \dots, \alpha_2 e_2, \alpha_1 e_1)$ 将 $\alpha_1 e_1$ 与前 $n-1$ 列依次对换 $n-1$ 次, 换到第 1 列; 再将 $\alpha_2 e_2$ 与前 $n-2$ 列依次对换 $n-2$ 次, 换到第 2 列; 如此, 依次地将 $\alpha_3 e_3, \dots, \alpha_{n-1} e_{n-1}$ 分别与它们前面的 $n-3, \dots, 1$ 列对换 $n-3, \dots, 1$ 次, 换到第 3, $\dots, n-1$ 列, 得

$$\begin{aligned} D_n &= (-1)^{(n-1)+(n-2)+\dots+1} D(\alpha_1 e_1, \alpha_2 e_2, \dots, \alpha_{n-1} e_{n-1}, \alpha_n e_n) \\ &= (-1)^{\binom{n-1}{2}} a_1 a_2 \dots a_{n-1} a_n D(e_1, e_2, \dots, e_{n-1}, e_n) \\ &= (-1)^{\binom{n-1}{2}} a_1 a_2 \dots a_{n-1} a_n. \end{aligned}$$

对一般的 n 阶行列式 D_n 按定义计算其值是相当繁的, 为使计算简便, 我们要根据定义推出行列式的一些性质. 然后按行列式定义及其性质来计算(或称展开)行列式.

性质 1 若行列式有一列为零向量, 则行列式的值等于零, 即

$$D(\alpha_1, \dots, 0, \dots, \alpha_n) = 0.$$

证 因为零向量 $0 = 0\alpha_i$ (右端的 0 为数 0), 所以

$$\begin{aligned} D(\alpha_1, \dots, 0, \dots, \alpha_n) &= D(\alpha_1, \dots, 0\alpha_i, \dots, \alpha_n) \\ &= 0D(\alpha_1, \dots, \alpha_i, \dots, \alpha_n) = 0. \end{aligned}$$

性质 2 若行列式有两列元素相同, 则行列式的值等于零.

证 设 $\alpha_i = \alpha_j$ ($i \neq j$), 根据定义 5.1 中(3)则有

$$D(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_n) = -D(\alpha_1, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_n),$$

而等式两边行列式相同, 故 $D(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_n) = 0$.

性质 3 若行列式有两列元素成比例, 即 $\alpha_i = \lambda\alpha_j$ ($i \neq j$) ($\lambda \in F$), 则行列式的值等于零. 由定义 5.1 中(1)及性质 2, 立即得此性质.

性质 4 将行列式的某一列乘以常数加到另一列(即对行列式作倍加列变换)则行列式的值不变. 即

$$D(\alpha_1, \dots, \alpha_i, \dots, \lambda\alpha_i + \alpha_j, \dots, \alpha_n) = D(\alpha_1, \dots, \alpha_i, \dots, \alpha_j, \dots, \alpha_n).$$

由定义 5.1 中(2)及性质 3, 立即得此性质.

性质 5 若 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关, 则 $D(\alpha_1, \alpha_2, \dots, \alpha_n) = 0$.

证 由于 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性相关, 所以有一个向量可被其余向量线性表示, 不妨设

$$\alpha_1 = \lambda_2 \alpha_2 + \dots + \lambda_n \alpha_n.$$

于是, 利用性质 4, 将 D 中第 2, \dots, n 列依次乘 $-\lambda_2, \dots, -\lambda_n$ 加到第一列, 即得

$$D(\alpha_1, \alpha_2, \dots, \alpha_n) = D(0, \alpha_2, \dots, \alpha_n) = 0.$$

例2 用定义和性质计算二阶行列式 $D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

解

$$\begin{aligned} D &= D(\alpha_1, \alpha_2) = D(a_{11}e_1 + a_{21}e_2, a_{12}e_1 + a_{22}e_2) \\ &= D(a_{11}e_1, a_{12}e_1) + D(a_{11}e_1, a_{22}e_2) + D(a_{21}e_2, a_{12}e_1) \\ &\quad + D(a_{21}e_2, a_{22}e_2) \\ &= 0 + a_{11}a_{22}D(e_1, e_2) + a_{21}a_{12}D(e_2, e_1) + 0 \\ &= a_{11}a_{22} - a_{21}a_{12}D(e_1, e_2) = a_{11}a_{22} - a_{21}a_{12} \end{aligned}$$

这个结果与本章开始定义的二阶行列式是相同的. 用同样的方法计算三阶行列式, 其结果与前面的定义也相同.

我们把 n 阶行列式 $D = |a_{ij}|_{n \times n}$ 也叫作 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 的行列式, 并记作 $D = \det A$ 或 $|A|$.

根据定义 5.1 及性质 4 可知三种初等矩阵的行列式为:

$$|E_i(c)| = D(e_1, \dots, ce_i, \dots, e_n) = cD(e_1, \dots, e_i, \dots, e_n) = c,$$

$$\begin{aligned} |E_{ij}| &= D(e_1, \dots, e_j, \dots, e_i, \dots, e_n) \\ &= -D(e_1, \dots, e_i, \dots, e_j, \dots, e_n) = -1 \quad (i < j), \end{aligned}$$

$$\begin{aligned} |E_{ij}(c)| &= D(e_1, \dots, e_i + ce_j, \dots, e_j, \dots, e_n) \\ &= D(e_1, \dots, e_i, \dots, e_j, \dots, e_n) = 1. \end{aligned}$$

于是对 n 阶矩阵 A 作三种初等列变换所得矩阵的行列式与 A 及初等矩阵的行列式有如下关系:

$$|AE_i(c)| = c|A| = |A||E_i(c)|, \quad (5-8)$$

$$|AE_{ij}| = -|A| = |A||E_{ij}|, \quad (5-9)$$

$$|AE_{ij}(c)| = |A| = |A||E_{ij}(c)|. \quad (5-10)$$

因此, 如果 P_1, P_2, \dots, P_k 均为初等矩阵, 则有

$$|AP_1P_2\dots P_k| = |A||P_1||P_2|\dots|P_k|. \quad (5-11)$$

利用这个结果, 可得行列式的另一个重要性质.

性质 6 设 A 是数域 F 上的 n 阶矩阵, 则

$$|A^T| = |A|, \quad (5-12)$$

即 n 阶行列式的行与列依次互换, 则行列式的值不变.

证 若秩 $(A) < n$, 则秩 $(A^T) = \text{秩}(A) < n$, 由性质 5 得,

$$|A^T| = |A| = 0.$$

若秩 $(A) = n$, 即 A 可逆, 则存在初等矩阵 P_1, P_2, \dots, P_k , 使得

$$A = P_1 P_2 \dots P_k, \quad A^T = P_k^T \dots P_2^T P_1^T.$$

于是, $|A| = |P_1| |P_2| \dots |P_k|$, $|A^T| = |P_k^T| \dots |P_2^T| |P_1^T|$, 而初等矩阵 P_i 与 P_i^T 是同类初等矩阵, 且 $|P_i^T| = |P_i|$, 所以 $|A^T| = |A|$.

性质6表明如果把 n 阶行列式 D 看作 n 个行向量的函数, 则定义中4条规则及行列式的5条性质对行向量都成立.

例3 证明: n 阶上三角行列式

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = \prod_{i=1}^n a_{ii}.$$

证 如果 $|A|$ 中 $\exists a_{ii} = 0$, 则秩 $(A) < n$, 从而 $|A| = 0$, 故结论成立.

如果 $a_{ii} \neq 0 (i = 1, 2, \dots, n)$, 则对 A 作倍加列变换可使 A 化为对角阵 $\text{diag}(a_{11}, a_{22}, \dots, a_{nn})$, 而倍加列变换不改变行列式的值, 故

$$|A| = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) = a_{11} a_{22} \dots a_{nn}.$$

同理或利用性质6可得, 下三角行列式也等于主对角元之积.

计算行列式常利用行列式的性质(包括定义中的基本性质), 用初等变换把它化为一些简单的形式(如上, 下三角行列式), 然后再来求它的值. 例如

例4 计算四阶行列式

$$D = \begin{vmatrix} 1 & 2 & -2 & 3 \\ -1 & -2 & 4 & -2 \\ 0 & 1 & 2 & -1 \\ 2 & 3 & -3 & 10 \end{vmatrix}.$$

解 对 D 作行变换将其化为上三角行列式

$$\begin{aligned} D &\xrightarrow[\substack{[4]+[1] \times (-2)}]{\substack{[2]+[1] \\ [4]+[1] \times (-2)}} \begin{vmatrix} 1 & 2 & -2 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & -1 & 1 & 4 \end{vmatrix} \xrightarrow[\substack{[2] \leftrightarrow [3]}]{\substack{[4]+[3] \\ [2] \leftrightarrow [3]}} \begin{vmatrix} 1 & 2 & -2 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & 3 \end{vmatrix} \\ &\xrightarrow{\substack{[4]+[3] \times (-3/2)}} \begin{vmatrix} 1 & 2 & -2 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & \frac{3}{2} \end{vmatrix} = -2 \times \frac{3}{2} = -3. \end{aligned}$$

例5 计算 n 阶行列式 $D_n = \begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \cdots & & & \cdots \\ b & b & \cdots & a \end{vmatrix}$.

解 D_n 中主对角元均为 a , 其余元素都是 b , 所以 D_n 的每行元素之和均为 $a + (n-1)b$, 此时, 可把各列加到第一列, 然后将第一行乘 (-1) 加到其余各行, 如此就化为上三角行列式, 即

$$\begin{aligned} D_n &= \begin{vmatrix} a + (n-1)b & b & \cdots & b \\ a + (n-1)b & a & \cdots & b \\ \cdots & & & \cdots \\ a + (n-1)b & b & \cdots & a \end{vmatrix} \\ &= [a + (n-1)b] \begin{vmatrix} 1 & b & \cdots & b \\ 1 & a & \cdots & b \\ \cdots & & & \cdots \\ 1 & b & \cdots & a \end{vmatrix} \\ &= [a + (n-1)b] \begin{vmatrix} 1 & b & \cdots & b \\ 0 & a-b & & 0 \\ \cdots & & \ddots & \cdots \\ 0 & 0 & \cdots & a-b \end{vmatrix} \\ &= [a + (n-1)b] (a-b)^{n-1}. \end{aligned}$$

例6 证明

$$\begin{vmatrix} a_1 + b_1 & b_1 + c_1 & c_1 + a_1 \\ a_2 + b_2 & b_2 + c_2 & c_2 + a_2 \\ a_3 + b_3 & b_3 + c_3 & c_3 + a_3 \end{vmatrix} = 2 \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

证法一 把左端行列式的第 2、3 列加到第一列, 提出公因子 2, 再把第一列乘 (-1) 加到第 2、3 列, 得

$$\text{左式} = 2 \begin{vmatrix} a_1 + b_1 + c_1 & -a_1 & -b_1 \\ a_2 + b_2 + c_2 & -a_2 & -b_2 \\ a_3 + b_3 + c_3 & -a_3 & -b_3 \end{vmatrix},$$

再将第 2、3 列加到第一列, 然后提出第 2、3 列的公因数 (-1) , 再作两次列对换, 等式就得证.

证法二 对左式的各列依次用定义 5.1 中的 (2), 即

$$\begin{aligned}
\text{左式} &= \begin{vmatrix} a_1 & b_1 + c_1 & c_1 + a_1 \\ a_2 & b_2 + c_2 & c_2 + a_2 \\ a_3 & b_3 + c_3 & c_3 + a_3 \end{vmatrix} + \begin{vmatrix} b_1 & b_1 + c_1 & c_1 + a_1 \\ b_2 & b_2 + c_2 & c_2 + a_2 \\ b_3 & b_3 + c_3 & c_3 + a_3 \end{vmatrix} \\
&= \begin{vmatrix} a_1 & b_1 & c_1 + a_1 \\ a_2 & b_2 & c_2 + a_2 \\ a_3 & b_3 & c_3 + a_3 \end{vmatrix} + \begin{vmatrix} a_1 & c_1 & c_1 + a_1 \\ a_2 & c_2 & c_2 + a_2 \\ a_3 & c_3 & c_3 + a_3 \end{vmatrix} + 0 + \begin{vmatrix} b_1 & c_1 & c_1 + a_1 \\ b_2 & c_2 & c_2 + a_2 \\ b_3 & c_3 & c_3 + a_3 \end{vmatrix} \\
&= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + 0 + 0 + 0 + 0 + 0 + \begin{vmatrix} b_1 & c_1 & a_1 \\ b_2 & c_2 & a_2 \\ b_3 & c_3 & a_3 \end{vmatrix} = \text{右式}.
\end{aligned}$$

5.2 行列式按一行(行)的展开式

行列式按一行(或行)的展开式是计算行列式的一种常用方法,它把计算一个 n 阶行列式化为计算 n 个 $n-1$ 阶行列式.

定义 5.2 在 n 阶行列式 $D = |a_{ij}|_{n \times n}$ 中,去掉元素 a_{ij} 所在的第 i 行和第 j 列的所有元素而得到的 $n-1$ 阶行列式,称为元素 a_{ij} 的余子式,记作 M_{ij} ,并把数

$$A_{ij} = (-1)^{i+j} M_{ij} \quad (5-13)$$

称为元素 a_{ij} 的代数余子式.

例如三阶行列式 $D = |a_{ij}|_{3 \times 3}$ 中 a_{11} a_{12} a_{13} 的代数余子式

$$A_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad A_{12} = - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \quad A_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

为了证明行列式按一行(或行)展开的定理,先证一个引理.

引理 设 $A = (a_{ij})_{n \times n}$ 中 $a_{in} = 0$ ($i = 1, 2, \dots, n-1$) 则

$$D = |A| = a_{nn} M_{nn} = a_{nn} A_{nn}. \quad (5-14)$$

证 将矩阵 A 分块表示为

$$A = \begin{bmatrix} A_1 & \mathbf{0} \\ \boldsymbol{\alpha} & a_{nn} \end{bmatrix}, \quad \text{即} \quad D = \begin{vmatrix} A_1 & \mathbf{0} \\ \boldsymbol{\alpha} & a_{nn} \end{vmatrix}.$$

其中 $|A_1| = M_{nn} = (-1)^{n+n} M_{nn} = A_{nn}$. 若 $a_{nn} = 0$ 则 $D = 0$ (5-14) 式成立,若 $a_{nn} \neq 0$ 对 D 作倍加列变换,可得

$$D = |A| = \begin{vmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & a_{nn} \end{vmatrix}.$$

若 $r(A_1) < n-1$ 则 $|A_1| = M_{nn} = 0$ 且 $r(A) < n, D = |A| = 0$ 故 (5-14) 式成立. 若 $r(A_1) = n-1$ 则对矩阵 A 作倍加行变换可将 A_1 化为上三角矩阵 B_1 , 而对方阵作倍加变换, 其行列式的值不变, 于是

$$\begin{aligned} D &= \begin{vmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & a_{nn} \end{vmatrix} = \begin{vmatrix} B_1 & \mathbf{0} \\ \mathbf{0} & a_{nn} \end{vmatrix} = a_{nn} |B_1| \\ &= a_{nn} |A_1| = a_{nn} M_{nn} = a_{nn} A_{nn}. \end{aligned}$$

定理5.1 设 $D = |a_{ij}|_{n \times n}$, 则

$$D = \sum_{k=1}^n a_{kj} A_{kj} = a_{1j} A_{1j} + a_{2j} A_{2j} + \dots + a_{nj} A_{nj}, \quad j = 1, \dots, n, \quad (5-15)$$

$$D = \sum_{k=1}^n a_{ik} A_{ik} = a_{i1} A_{i1} + a_{i2} A_{i2} + \dots + a_{in} A_{in}, \quad i = 1, \dots, n. \quad (5-16)$$

(5-15) 式称为 D 对第 j 列的展开式, (5-16) 式称为 D 对第 i 行的展开式.

下面证 (5-15) 式:

证 先证明第 j 列元素 $a_{ij} \neq 0$ ($a_{kj} = 0$ ($k \neq i$)) 的情形 (即第 j 列除 $a_{ij} \neq 0$ 其余元素均为零). 此时行列式记作 D_{ij} , 把 D_{ij} 中第 j 列依次与第 $j+1, j+2, \dots, n$ 列对换 (共对换 $n-j$ 次), 使第 j 列换到第 n 列, 再把第 i 行依次与第 $i+1, i+2, \dots, n$ 行对换 (共对换 $n-i$ 次), 使第 i 行换到第 n 行, 于是

$$D_{ij} = (-1)^{n-j+i} \begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} & 0 \\ \dots & & & & & & \dots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} & 0 \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} & 0 \\ \dots & & & & & & \dots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} & 0 \\ a_{i1} & \dots & a_{i,j-1} & a_{i,j+1} & \dots & a_{in} & a_{ij} \end{vmatrix},$$

其中左上角的 $n-1$ 阶行列式就是 D_{ij} 中元素 a_{ij} 的余子式 M_{ij} . 根据引理, 就有

$$D_{ij} = (-1)^{n-i-j} a_{ij} M_{ij} = a_{ij} (-1)^{n-i+j} M_{ij} = a_{ij} A_{ij}. \quad (5-17)$$

对于一般的 n 阶行列式, 将其第 j 列元素 $a_{1j}, a_{2j}, \dots, a_{nj}$ 写成:

$$\begin{array}{ccccccc} a_{1j} & + & 0 & + & \dots & + & 0 \\ 0 & + & a_{2j} & + & \dots & + & 0 \\ & & & & \dots & & \\ 0 & + & 0 & + & \dots & + & a_{nj} \end{array}$$

由定义 5.1 中的 (2) 和 (5-17) 式的结果, 即得

$$D = D_{1j} + D_{2j} + \dots + D_{nj} = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj}.$$

(5-15) 式得证. 利用 $|A^T| = |A|$ 和 (5-15) 式可证 (5-16) 式也成立.

例如, 三阶行列式 $D = |a_{ij}|_{3 \times 3}$ 对第 2 行的展开式为:

$$\begin{aligned} D &= a_{21}A_{21} + a_{22}A_{22} + a_{23}A_{23} \\ &= -a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned}$$

根据定理 5.1, 用数学归纳法可以证明: n ($n > 1$) 阶行列式 $|a_{ij}|_{n \times n}$ 是 n^2 个元 a_{ij} ($i, j = 1, 2, \dots, n$) 的 n 次齐次多项式, 其中共有 n 项, 每项都是不同行不同列的 n 个元素之积, 且正负项各占一半. 证明时, 只要注意到结论对二阶行列式成立.

定理 5.2 n 阶行列式 $D = |a_{ij}|_{n \times n}$ 的某一行 (或列) 元素与另一行 (或列) 相应元素的代数余子式的乘积之和等于零, 即

$$\begin{aligned} \sum_{k=1}^n a_{kj}A_{ki} &= a_{1j}A_{1i} + a_{2j}A_{2i} + \dots + a_{nj}A_{ni} = 0, \quad j \neq i, \\ \sum_{k=1}^n a_{jk}A_{ik} &= a_{j1}A_{i1} + a_{j2}A_{i2} + \dots + a_{jn}A_{in} = 0, \quad j \neq i. \end{aligned} \quad (5-18)$$

证 对列的情况证明. 将 D 对第 i 列展开, 得

$$D = \sum_{k=1}^n a_{ki}A_{ki}.$$

如果将 D 中第 i 列元素 $a_{1i}, a_{2i}, \dots, a_{ni}$ 换为第 j 列元素 $a_{1j}, a_{2j}, \dots, a_{nj}$ 所得的行列式为 D_i , 则 $D_i = 0$ (因为 D_i 中第 i 列与第 j 列相同), 即此时 D_i 对第 i 列的展开式为

$$D_i = \sum_{k=1}^n a_{kj}A_{ki} = 0 \quad (j \neq i).$$

引入克罗内克 (Kronecker) 符号

$$\delta_{ij} = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases}$$

定理 5.1 和定理 5.2 的结论可以统一写成

$$\sum_{k=1}^n a_{kj}A_{ki} = \delta_{ij}D, \quad \sum_{k=1}^n a_{jk}A_{ik} = \delta_{ij}D. \quad (5-19)$$

定理 5.1 表明, 计算一个 n 阶行列式可以化为计算 n 个 $n-1$ 阶行列式. 通常我们对行列式作倍加初等变换把某行 (或列) 化为只有一个非零元, 然后再对该行 (或列) 展开, 就可简化计算.

例1 计算四阶行列式

$$D = \begin{vmatrix} 2 & -4 & 2 & 0 \\ 2 & 1 & 4 & 1 \\ 1 & 7 & 4 & 2 \\ -3 & -2 & 1 & 1 \end{vmatrix}.$$

解 将第3列乘-1加到第1列,再乘2加到第2列,然后对第1行展开,得

$$D = \begin{vmatrix} 0 & 0 & 2 & 0 \\ -2 & 9 & 4 & 1 \\ -3 & 15 & 4 & 2 \\ -4 & 0 & 1 & 1 \end{vmatrix} = 2(-1)^{+3} \begin{vmatrix} -2 & 9 & 1 \\ -3 & 15 & 2 \\ -4 & 0 & 1 \end{vmatrix}.$$

再将上面的三阶行列式的第3列乘4加到第1列,然后对第3行展开,得

$$D = 2 \begin{vmatrix} 2 & 9 & 1 \\ 5 & 15 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 2(-1)^{+3} \begin{vmatrix} 2 & 9 \\ 5 & 15 \end{vmatrix} = 2(30 - 45) = -30.$$

例2 证明: n 阶范德蒙(Vandermonde)行列式

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (x_j - x_i),$$

其连乘积中因式 $(x_j - x_i)$ 共有 $(n-1) + \cdots + 2 + 1 = \frac{n(n-1)}{2}$ 个,即

$$V_n = [(x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1)][(x_3 - x_2) \cdots (x_n - x_2)] \\ \cdots [(x_{n-1} - x_{n-2})(x_n - x_{n-2})] \cdots (x_n - x_{n-1}).$$

证 对 n 作归纳法,当 $n=2$ 时,结论显然成立,假设结论对 $n-1$ 阶范德蒙行列式成立,现对于 V_n ,从第 n 行起,依次将前一行乘 $(-x_1)$ 加到后一行,得

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \cdots & & & \cdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}.$$

上式右端是 x_2, \dots, x_n 的 $n-1$ 阶范德蒙行列式, 由归纳假设得

$$\begin{aligned} V_n &= [(x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1)] \prod_{2 \leq i < j \leq n} (x_j - x_i) \\ &= \prod_{1 \leq i < j \leq n} (x_j - x_i). \end{aligned}$$

例3 计算 n 阶三对角行列式

$$D_n = \begin{vmatrix} a & b & & & \\ c & a & b & & \\ & c & a & \ddots & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & a & b \\ & & & & c & a \end{vmatrix}_{n \times n} \quad (a \neq 0, b \neq 0, c \neq 0).$$

解 解题的方法是找一个递推公式, 然后算出结果. 先对第1列展开, 得

$$\begin{aligned} D_n &= aD_{n-1} - c \begin{vmatrix} b & 0 \\ c & a & b \\ & c & a & \ddots \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & a & b \\ & & & & c & a \end{vmatrix}_{(n-1) \times (n-1)} \\ &= aD_{n-1} - bcD_{n-2}. \end{aligned} \quad ①$$

将递推公式 ① 化为

$$(D_n - kD_{n-1}) = l(D_{n-1} - kD_{n-2}), \quad ②$$

其中

$$k + l = a, \quad kl = bc. \quad ③$$

反复利用递推公式 ② 递推, 即得

$$(D_n - kD_{n-1}) = l^{n-2}(D_2 - kD_1),$$

其中

$$D_2 - kD_1 = \begin{vmatrix} a & b \\ c & a \end{vmatrix} - ka = a^2 - bc - k(k + l) = l^2,$$

所以

$$D_n = l^n + kD_{n-1}. \quad (4)$$

重复利用递推公式 (4), 得

$$\begin{aligned} D_n &= l^n + kD_{n-1} = l^n + k(l^{n-1} + kD_{n-2}) \\ &= l^n + kl_{n-1} + k^2(l^{n-2} + kD_{n-3}) \\ &= \dots \\ &= l^n + kl^{n-1} + k^2l^{n-2} + \dots + k^{n-2}l^2 + k^{n-1}D_1 \\ &= l^n + kl^{n-1} + k^2l^{n-2} + \dots + k^{n-2}l^2 + k^{n-1}l + k^n, \end{aligned}$$

即

$$D_n = \begin{cases} \frac{l^{n+1} - k^{n+1}}{l - k}, & l \neq k, \\ (n+1)k^n, & l = k. \end{cases}$$

例4 证明

$$D = \begin{vmatrix} A & O \\ C & B \end{vmatrix} = |A| |B|, \quad (5-20)$$

其中 $A = (a_{ij})_{k \times k}$, $B = (b_{ij})_{m \times m}$, C 是任意 $m \times k$ 矩阵.

证 对 A 的阶数 k 作数学归纳法. 当 $k=1$ 时, $D = |a_{11}| |B|$ (5-20) 式成立 (这里 $|a_{11}|$ 是一阶行列式). 假设 A 为 $k-1$ 阶时 (5-20) 式成立, 下面考虑 A 为 k 阶的情形:

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2k} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} & 0 & \dots & 0 \\ c_{11} & c_{12} & \dots & c_{1k} & b_{11} & \dots & b_{1m} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mk} & b_{m1} & \dots & b_{mm} \end{vmatrix}.$$

此时 将 D 对第一行展开, 得

$$D = a_{11}(-1)^{1+1}M_{11}^D + a_{12}(-1)^{1+2}M_{12}^D + \dots + a_{1k}(-1)^{1+k}M_{1k}^D, \quad (1)$$

其中 M_{1j}^D 是 a_{1j} 在 D 中的余子式 ($j=1, 2, \dots, k$). 显然 M_{1j}^D 也是 (5-20) 类型的行列式, 而且它的左上角是 $k-1$ 阶的. 根据归纳假设

$$M_{1j}^D = M_{1j}^A |B|, \quad j=1, 2, \dots, k, \quad (2)$$

其中 M_{1j}^A 是 a_{1j} 在 $|A|$ 中的余子式. 将 (2) 代入 (1), 即得

$$D = [a_{11}(-1)^{1+1}M_{11}^A + a_{12}(-1)^{1+2}M_{12}^A + \dots + a_{1k}(-1)^{1+k}M_{1k}^A] |B| \\ = |A| |B|.$$

A, B 如上所设, 同样也有

$$D = \begin{vmatrix} A & C_1 \\ O & B \end{vmatrix} = |A| |B|,$$

但要注意

$$D = \begin{vmatrix} O & A \\ B & C \end{vmatrix} \neq |A| |B|.$$

此时, 可将 A 和 C 中的每一列依次与其前面的 m 列逐列对换(共对换 $k \times m$ 次), 即得

$$\begin{vmatrix} O & A \\ B & C \end{vmatrix} = (-1)^{k \times m} \begin{vmatrix} A & O \\ C & B \end{vmatrix} = (-1)^{k \times m} |A| |B|.$$

例5 设几何向量 $\mathbf{a} = (a_1, a_2, a_3,)$, $\mathbf{b} = (b_1, b_2, b_3,)$, 即 $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, 节 1.7 中已得到的 $\mathbf{a} \times \mathbf{b}$ 的结果, 可用行列式表示, 即

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k}.$$

根据三阶行列式按一行的展开式, 上式又可形式地表示为

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

于是三个向量的混合积 $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ (其中 $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$), 可用如下的行列式计算, 即

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1) \\ = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

在节 1.7 中已证明过, $|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$ 表示以 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 为邻边的平行六面体的体积, 所以上述三阶行列式的绝对值是以三个行向量为邻边的平行六面体的体积.

根据行列式两行互换行列式反号的性质, 立即可得混合积的性质:

$$a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b). \quad (5-21)$$

5.3 方阵乘积的行列式

本节中将证明: 方阵乘积的行列式等于方阵行列式的乘积; n 阶矩阵 A 的 $\det A \neq 0$ 是 A 可逆的充要条件, 并给出求 A^{-1} 的另一种方法. 此外还将阐明矩阵 A 的秩与 A 的子行列式之间的关系.

定理 5.3 若 $A, B \in M_n(F)$ 则

$$|AB| = |A| |B|. \quad (5-22)$$

证 如果 $\text{秩}(B) < n$ 则 $\text{秩}(AB) \leq \text{秩}(B) < n$, 于是 $|AB| = |B| = 0$, 故 (5-22) 式成立.

如果 $\text{秩}(B) = n$ 即 B 可逆, 则存在初等矩阵 P_1, P_2, \dots, P_k 使

$$B = P_1 P_2 \dots P_k, \quad AB = AP_1 P_2 \dots P_k.$$

于是根据 (5-11) 式, 即得

$$|AB| = |A| |P_1| |P_2| \dots |P_k| = |A| |B|.$$

定理 5.4 n 阶矩阵 A 可逆的充要条件是 $|A| \neq 0$.

证 若 A 可逆, 则存在 B 使得 $AB = E$, 于是

$$|AB| = |A| |B| = |E| = 1,$$

从而 $|A| \neq 0$. 必要性得证.

若 $|A| \neq 0$, 则由行列式的性质 5 可知, A 的列向量线性无关, 即, $\text{秩}(A) = n$, 故 A 可逆.

充分性的另一种证法: 我们先作一个矩阵

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & & & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \quad (5-23)$$

A^* 称为 A 的伴随矩阵, 它是 $A = (a_{ij})_{n \times n}$ 的元素 a_{ij} 的代数余子式 A_{ij} 矩阵 $(A_{ij})_{n \times n}$ 的转置, 其中第 j 列元素 $A_{j1}, A_{j2}, \dots, A_{jn}$ 是 $|A|$ 中第 j 行元素 $a_{j1}, a_{j2}, \dots, a_{jn}$ 的代数余子式 ($j = 1, 2, \dots, n$). 于是, 由 (5-19) 式即得

$$AA^* = A^*A = |A|E.$$

如此, 则有

$$A \left(\frac{1}{|A|} A^* \right) = \left(\frac{1}{|A|} A^* \right) A = E,$$

故矩阵 A 可逆,且

$$A^{-1} = \frac{1}{|A|} A^*. \quad (5-24)$$

这种证明方法叫做构造性证法,它构造出了 A 的逆矩阵,这也是求逆矩阵的常用方法.

有些方阵用其行列式来判断是否可逆是很方便的,例如:

上(下)三角矩阵 $A = (a_{ij})_{n \times n}$ 可逆的充要条件为主对角元都非零,因为 $|A| = a_{11}a_{22}\cdots a_{nn}$.

上(下)三角块方阵

$$A = \begin{bmatrix} B & * \\ O & C \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} B^{-1} & O \\ * & C^{-1} \end{bmatrix}$$

(其中 B, C 皆为方阵)可逆的充要条件为, $|A| = |B||C| \neq 0$, 即 B, C 皆为可逆矩阵.

再如,奇数阶反对称矩阵 A 不可逆,因为 $A^T = -A$, 所以

$$|A| = |A^T| = |-A| = (-1)^n |A|.$$

当 n 为奇数时,便有 $|A| = 0$.

例1 下列矩阵 A, B, C 是否可逆?若可逆,求其逆矩阵.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 3 & -1 \\ -1 & -3 & 5 \\ 1 & 0 & 4 \end{bmatrix}.$$

解 $|A| = 4$, 故 A 可逆. $|A|$ 中各元素的代数余子式为:

$$A_{11} = \begin{vmatrix} 1 & 2 \\ 3 & 3 \end{vmatrix} = -3, \quad A_{12} = -\begin{vmatrix} 2 & 2 \\ 1 & 3 \end{vmatrix} = -4, \quad A_{13} = \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} = 5.$$

类似地有,

$$\begin{aligned} A_{21} &= 3, & A_{22} &= 0, & A_{23} &= -1, \\ A_{31} &= 1, & A_{32} &= 4, & A_{33} &= -3, \end{aligned}$$

所以

$$A^{-1} = \frac{1}{|A|} A^* = \frac{1}{4} \begin{bmatrix} -3 & 3 & 1 \\ -4 & 0 & 4 \\ 5 & -1 & -3 \end{bmatrix}.$$

$|B| = -2$, 故 B 可逆,其逆矩阵为

$$B^{-1} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

$|C| = 0$, 故 C 不可逆.

以前讨论矩阵的秩时, 都是从行(或列)向量的线性相关性出发. 现在从行列式的角度来讨论矩阵的秩. 为此先给一个定义.

定义 5.3 矩阵 $A = (a_{ij})_{m \times n}$ 的任意 k 行 ($i_1 < i_2 < \dots < i_k$ 行) 和任意 k 列 ($j_1 < j_2 < \dots < j_k$ 列) 的交点上的 k^2 个元素排成的行列式

$$\begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \cdots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \cdots & a_{i_2 j_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_k j_1} & a_{i_k j_2} & \cdots & a_{i_k j_k} \end{vmatrix} \quad (5-25)$$

称为矩阵 A 的一个 k 阶子行列式, 简称 A 的 k 阶子式. 当 (5-25) 式等于零时, 称为 k 阶零子式, 否则叫非零子式. 当 A 为方阵且 (5-25) 式中 $j_t = i_t$ ($t = 1, 2, \dots, k$) 时, 称为 A 的 k 阶主子式.

如果矩阵 A 存在 r 阶非零子式, 而所有 $r+1$ 阶子式(如果有的话)都等于零, 则矩阵 A 的非零子式的最高阶数为 r , 因为由所有 $r+1$ 阶子式都等于零, 可根据按一列(行)的展开式推出所有更高阶的子式也都等于零.

定义 5.4 矩阵 A 的非零子式的最高阶数 r 称为 A 的行列式秩.

定理 5.5 秩 $(A) = r$ 的充要条件是 A 的行列式的秩为 r .

证 先证必要性. 秩 $(A) = r$, 即 A 的行秩为 r , 不妨设 A 的前 r 行构成的矩阵 A_1 的行秩为 r , 其列秩也为 r ; 不妨再设 A_1 的前 r 列向量线性无关. 如此, A 的左上角的 r 阶子块是可逆矩阵, 其行列式是 A 的一个 r 阶非零子式; 又因为 A 的任意 $r+1$ 个行向量线性相关, 因此任意 $r+1$ 阶子式都是零子式. 故 A 的行列式的秩为 r .

再证充分性. 不妨设 A 的左上角 r 阶子式 $|A_0|$ 为非零子式. 于是 A_0 可逆, 其 r 个行向量线性无关, 将它们添分量成为 A 的前 r 个行向量, 它们也线性无关, 而且此时 A 中任何 $r+1$ 行线性相关(否则由必要性的证明可知 A 中存在 $r+1$ 阶非零子式), 故矩阵 A 的行秩为 r , 即秩 (A) 为 r .

由定理 5.4 和定理 5.5 都可得: n 阶矩阵 A 的秩小于 n (即 A 的列向量组, 行向量组都线性相关) 的充要条件是 $|A| = 0$.

例 2 证明: 若 A 是一个 $m \times n$ 矩阵 ($m < n$), 则 $|A^T A| = 0$.

证 $A^T A$ 是一个 $n \times n$ 矩阵, 而秩 $(A^T A) \leq (\text{秩 } A) \leq m < n$, 故

$$|A^T A| = 0.$$

例 3 设 $W = L(\alpha_1, \alpha_2)$ 是 \mathbf{R}^4 的一个子空间, 其中 $\alpha_1 = (1, 2, 1, -1)^T$,

$\alpha_2 = (1 \ 4 \ -1 \ 1)^T$. 试将 α_1, α_2 扩充为 \mathbb{R}^4 的基.

解 第2章讲过求解的一种方法. 现在取 $\alpha_3, \alpha_4 \in \mathbb{R}^4$, 使 $D(\alpha_1, \alpha_2, \alpha_3, \alpha_4) \neq 0$, 由(5-20)式可知

$$\begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 2 \neq 0.$$

其中前两列为 α_1, α_2 . 于是取 $\alpha_3 = (0 \ 0 \ 1 \ 0)^T, \alpha_4 = (0 \ 0 \ 0 \ 1)^T$ 就使 $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ 为 \mathbb{R}^4 的一组基.

5.4 Cramer 法则

这一节又回到线性方程组求解的讨论. 当 n 阶矩阵 A 的行列式 $|A| \neq 0$ 时, 线性方程组 $AX = b$ 的解, 可用行列式表示, 其解法称为 Cramer 法则.

定理 5.6 (Cramer 法则) 设 $A = (a_{ij})_{n \times n}$, 若线性方程组 $AX = b$ 即

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, 2, \dots, n \quad (5-26)$$

的系数行列式 $D = |A| \neq 0$ 则方程组有唯一解, 且

$$x_j = \frac{D_j}{D}, \quad j = 1, 2, \dots, n, \quad (5-27)$$

其中

$$D_j = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \cdots & & & & & & \cdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}. \quad (5-28)$$

证 由于 $|A| \neq 0$, 即 A 可逆, 所以方程组(5-26)有唯一解, 且

$$X = A^{-1}b = \frac{1}{|A|}A^*b = \frac{1}{D}A^*b,$$

即

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \frac{1}{D} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & & & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix},$$

于是

$$x_j = \frac{1}{D}(b_1 A_{1j} + b_2 A_{2j} + \dots + b_n A_{nj}) = \frac{D_j}{D}, \quad j = 1, 2, \dots, n.$$

由定理 5.6 可知, 齐次线性方程组 $AX = 0$, 当系数行列式 $|A| \neq 0$ 时, 只有零解. 因此 $AX = 0$ 有非零解的必要条件是 $|A| = 0$. 事实上, $|A| = 0$ 也是有非零解的充分条件, 因为当 $|A| = 0$ 时, $\text{秩}(A) < n$, 于是对 A 做初等行变换将其化为阶梯形矩阵 U 时, 其非零行的行数小于 n , 因此 $AX = 0$ 的同解方程组 $UX = 0$ 必有非零解(或, 当 $|A| = 0$ 时, A 的列向量组 $\beta_1, \beta_2, \dots, \beta_n$ 线性相关, 于是存在 $X = (x_1, x_2, \dots, x_n)^T \neq 0$, 使 $x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n = AX = 0$), 如此即得

定理 5.7 若 A 为 n 阶矩阵, 则齐次线性方程组 $AX = 0$ 有非零解的充要条件为 $|A| = 0$, 即 $\text{秩}(A) < n$.

其等价命题为: $AX = 0$ 只有零解的充要条件为 $|A| \neq 0$.

用 Cramer 法则解 n 元非齐次线性方程组, 需要算 $n+1$ 个 n 阶行列式, 计算量很大. 实际上解线性方程组(包括系数行列式等于零及方程个数与未知数个数不相等的方程组), 一般都采用高斯消元法或其他一些近似解法. 然而 Cramer 法则在理论上是重要的, 特别是它揭示了方程组的解与系数之间的关系.

例 已知三次曲线 $y = a_0 + a_1x + a_2x^2 + a_3x^3$ 过 4 个点 $P_i(x_i, y_i)$, ($i = 1, 2, 3, 4$, 其中 x_1, x_2, x_3, x_4 互异), 试求方程的系数 a_0, a_1, a_2, a_3 .

解 方法一 将 4 个点的坐标分别代入三次曲线的方程, 得到非齐次线性方程组

$$\sum_{j=0}^3 a_j x_i^j = y_i, \quad i = 1, 2, 3, 4.$$

这个关于 a_0, a_1, a_2, a_3 的方程组的系数行列式 D 是 Vandermonde 行列式, 即

$$D = \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{vmatrix} = \prod_{1 \leq i < j \leq 4} (x_j - x_i) \neq 0.$$

根据 Cramer 法则, 它有唯一解 $a_j = \frac{D_{j+1}}{D}$ ($j = 0, 1, 2, 3$), 其中 D_{j+1} 是以 y_1, y_2, y_3, y_4 替代 D 中第 j 列元素所得的行列式.

方法二 将三次曲线上的任意点 (x, y) 所满足的方程写成

$$a_0 + a_1x + a_2x^2 + a_3x^3 + a_4y = 0 ,$$

它与

$$\sum_{j=0}^3 a_j x_i^j + a_4 y_i = 0 , \quad i = 1 \ 2 \ 3 \ 4$$

一起构成的关于 $a_0 \ a_1 \ a_2 \ a_3 \ a_4$ 的齐次线性方程组有非零解(因为 $a_4 = -1$), 于是由定理 5.7 得

$$\left| \begin{array}{ccccc} 1 & x & x^2 & x^3 & y \\ 1 & x_1 & x_1^2 & x_1^3 & y_1 \\ 1 & x_2 & x_2^2 & x_2^3 & y_2 \\ 1 & x_3 & x_3^2 & x_3^3 & y_3 \\ 1 & x_4 & x_4^2 & x_4^3 & y_4 \end{array} \right| \xrightarrow{\text{列变换}} \left| \begin{array}{ccccc} 1 & * & * & * & y - p(x) \\ 1 & 0 & 0 & 0 & 0 \\ 1 & b & 0 & 0 & 0 \\ 1 & * & c & 0 & 0 \\ 1 & * & * & d & 0 \end{array} \right| = 0 , bcd \neq 0 .$$

于是 $y = p(x)$ 就是过 4 个已知点的三次曲线方程.

行列式的几何意义 \mathbf{R}^n 中 m 个线性无关向量所张成的体积.

定义 \mathbf{R}^n 中 $\alpha_1 \ \alpha_2 \ \dots \ \alpha_m$ 为 m 个线性无关的列向量. 它们所张的体积 $V_m(\alpha_1 \ \dots \ \alpha_m)$ 定义为(用归纳法):

$$V_1^2(\alpha_1) = |\alpha_1|^2 ,$$

$V_k^2(\alpha_1 \ \dots \ \alpha_k) = |\beta_k|^2 \cdot V_{k-1}^2(\alpha_1 \ \dots \ \alpha_{k-1})$ 其中 $\beta_k = \alpha_k - \sum_{j=1}^{k-1} \lambda_j \alpha_j$ ($\lambda_j \in R$) 满足

$$\beta_k \perp L(\alpha_1 \ \dots \ \alpha_{k-1}) .$$

先说明满足这个条件的 λ_j 是存在的(有可能全为零). 事实上由 $\alpha_i^T \beta_k = 0$ ($i = 1 \ \dots \ k-1$) 得到 λ_j 所满足的 $(k-1)$ 个线性方程:

$$\left\{ \begin{array}{l} \sum_{j=1}^{k-1} \lambda_j \alpha_1^T \alpha_j = \alpha_1^T \alpha_k , \\ \dots\dots\dots \\ \sum_{j=1}^{k-1} \lambda_j \alpha_{k-1}^T \alpha_j = \alpha_{k-1}^T \alpha_k \end{array} \right.$$

这个线性方程组有唯一解($\lambda_1 \ \dots \ \lambda_{k-1}$) 的充要条件是系数行列式不为零, 即

$$|\alpha_i^T \alpha_j| = |A_{k-1}^T A_{k-1}| \neq 0 , \quad i \ j = 1 \ 2 \ \dots \ k-1 ,$$

其中

$$A_{k-1} = \begin{pmatrix} a_{11} & \cdots & a_{k-1,1} \\ \cdots & \cdots & \cdots \\ a_{1n} & \cdots & a_{k-1,n} \end{pmatrix} = (\alpha_1 \cdots \alpha_{k-1}).$$

这个矩阵的行列式不为零, 因为秩 $(A_{k-1}) = \text{秩}(A_{k-1}^T) = k-1$ (利用 α_j 的线性无关性). 面由(3-8), 我们有

$$(k-1) + (k-1) - (k-1) \leq \text{秩}(A_{k-1}^T A_{k-1}) \leq k-1.$$

因此秩 $(A_{k-1}^T A_{k-1}) = k-1$, 也就是满秩, 于是其行列式不为零.

定理 5.8 设 \mathbf{R}^n 中 m 个列向量 $\alpha_1 \cdots \alpha_m$ 线性无关, 而矩阵

$$A = (\alpha_1 \cdots \alpha_m)_{n \times m},$$

则

$$V_m^2(\alpha_1 \cdots \alpha_m) = |A^T A|.$$

证 用归纳法, $m=1$ 时按定义成立. 设 $|A_{m-1}^T A_{m-1}| = V_{m-1}^2(\alpha_1 \cdots \alpha_{m-1})$ 则

$$\begin{aligned} |A^T A| &= \begin{vmatrix} \alpha_1^T \alpha_1 & \cdots & \alpha_1^T \alpha_m \\ \cdots & \cdots & \cdots \\ \alpha_m^T \alpha_1 & \cdots & \alpha_m^T \alpha_m \end{vmatrix} \\ &= \begin{vmatrix} \alpha_1^T \alpha_1 & \cdots & \alpha_1^T \left(\beta_m + \sum_{i=1}^{m-1} \lambda_i \alpha_i \right) \\ \cdots & \cdots & \cdots \\ \left(\beta_m^T + \sum_{i=1}^{m-1} \lambda_i \alpha_i^T \right) \alpha_1 & \cdots & \left(\beta_m^T + \sum_{i=1}^{m-1} \lambda_i \alpha_i^T \right) \left(\beta_m + \sum_{i=1}^{m-1} \lambda_i \alpha_i \right) \end{vmatrix} \\ &= \begin{vmatrix} \alpha_1^T \alpha_1 & \cdots & \alpha_1^T \alpha_{m-1} & \sum_{i=1}^{m-1} \lambda_i \alpha_1^T \alpha_i \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{m-1} \lambda_i \alpha_i^T \alpha_1 & \cdots & \sum_{i=1}^{m-1} \lambda_i \alpha_i^T \alpha_{m-1} & |\beta_m|^2 + \sum_{j=1}^{m-1} \lambda_j \sum_{i=1}^{m-1} \lambda_i \alpha_j^T \alpha_i \end{vmatrix}. \end{aligned}$$

把第一行乘以 $-\lambda_1 \cdots$, 第 $(m-1)$ 行乘以 $-\lambda_{m-1}$ 后加到最后一行, 然后按第 m 列展开, 由归纳法假设就得到

$$\begin{aligned} |A^T A| &= |\beta_m|^2 \begin{vmatrix} \alpha_1^T \alpha_1 & \cdots & \alpha_1^T \alpha_{m-1} \\ \cdots & \cdots & \cdots \\ \alpha_{m-1}^T \alpha_1 & \cdots & \alpha_{m-1}^T \alpha_{m-1} \end{vmatrix} \\ &= |\beta_m|^2 V_{m-1}^2(\alpha_1 \cdots \alpha_{m-1}) = V_m^2(\alpha_1 \cdots \alpha_m) \end{aligned}$$

习 题

1. 计算下列行列式：

$$(1) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 0 \\ 3 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{vmatrix};$$

$$(2) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{vmatrix};$$

$$(3) \begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{vmatrix};$$

$$(4) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{vmatrix};$$

$$(5) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix};$$

$$(6) \begin{vmatrix} \lambda + 2 & -1 & -1 & -1 \\ -1 & \lambda + 2 & -1 & -1 \\ -1 & -1 & \lambda + 2 & -1 \\ -1 & -1 & -1 & \lambda + 2 \end{vmatrix};$$

$$(7) \begin{vmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{vmatrix};$$

$$(8) \begin{vmatrix} 3 & 2 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{vmatrix}.$$

2. 按对一行(或一列)展开第一题中(1)–(6)的行列式.

3. 展开下列行列式：

$$(1) \begin{vmatrix} a & x & b & x \\ x & a & x & b \\ b & x & a & x \\ x & b & x & a \end{vmatrix};$$

$$(2) \begin{vmatrix} 1 & x & x^2 & x^3 \\ x^3 & 1 & x & x^2 \\ x^2 & x^3 & 1 & x \\ x & x^2 & x^3 & 1 \end{vmatrix}.$$

4. 证明下列恒等式：

$$(1) \begin{vmatrix} a_1 + b_1x & a_1x + b_1 & c_1 \\ a_2 + b_2x & a_2x + b_2 & c_2 \\ a_3 + b_3x & a_3x + b_3 & c_3 \end{vmatrix} = (1 - x^2) \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix};$$

$$(2) \begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 & (a+3)^2 \\ b^2 & (b+1)^2 & (b+2)^2 & (b+3)^2 \\ c^2 & (c+1)^2 & (c+2)^2 & (c+3)^2 \\ d^2 & (d+1)^2 & (d+2)^2 & (d+3)^2 \end{vmatrix} = 0;$$

$$(3) \begin{vmatrix} x & -1 & & & \\ & x & -1 & & \\ & & \ddots & \ddots & \\ & & & x & -1 \\ a_n & a_{n-1} & \cdots & a_2 & x + a_1 \end{vmatrix} = x^n + \sum_{k=1}^n a_k x^{n-k};$$

$$(4) \begin{vmatrix} \cos \theta & 1 & & & \\ 1 & 2\cos \theta & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \\ & & & 1 & 2\cos \theta & 1 \\ & & & & 1 & 2\cos \theta \end{vmatrix} = \cos n\theta \quad (n \text{ 阶行列式}).$$

5. 利用 Vandermonde 行列式, 展开下列行列式:

$$(1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a & a-1 & \cdots & a-n \\ a^2 & (a-1)^2 & \cdots & (a-n)^2 \\ \cdots & \cdots & \cdots & \cdots \\ a^n & (a-1)^n & \cdots & (a-n)^n \end{vmatrix};$$

$$(2) \begin{vmatrix} a_1 + a_2 & a_2 + a_3 & \cdots & a_{n-1} + a_n & a_n + a_1 \\ a_1^2 + a_2^2 & a_2^2 + a_3^2 & \cdots & a_{n-1}^2 + a_n^2 & a_n^2 + a_1^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_1^n + a_2^n & a_2^n + a_3^n & \cdots & a_{n-1}^n + a_n^n & a_n^n + a_1^n \end{vmatrix}.$$

6. 设

$$D = \begin{vmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1+a_n \end{vmatrix}.$$

(1) 用递推公式计算行列式 D ;

(2) 利用定义 5.1 中的 (2) 将 D 表示为 2^n 个行列式之和, 并算出结果.

7*. Fibonacci 数列 $\{F_n\}$: 1, 2, 3, 5, 8, 13, 21, ... 满足递推关系:

$$\begin{cases} F_1 = 1, & F_2 = 2, \\ F_n = F_{n-1} + F_{n-2}, & n \geq 3. \end{cases}$$

证明:

$$F_n = \begin{vmatrix} 1 & 1 & & \\ -1 & 1 & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & 1 \end{vmatrix} \quad (n \text{ 阶三对角行列式}).$$

8. 利用 5-2 例 4 的结论, 计算下列行列式:

$$(1) \begin{vmatrix} 1 & 0 & 2 & 0 \\ -1 & 0 & 3 & 0 \\ 0 & 2 & 0 & -1 \\ 0 & 1 & 0 & 3 \end{vmatrix}; \quad (2) \begin{vmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{vmatrix};$$

(3) $A \in M_3(R), B \in M_4(R)$. 已知: $|A| = -6, |B| = 4$,

计算 $\begin{vmatrix} O & \frac{1}{2}A \\ -B & C \end{vmatrix}$, 其中 C 是任意的 4×3 矩阵.

9. 计算下列行列式:

$$(1) \begin{vmatrix} 1 & 2 & \cdots & 2 & 2 \\ 2 & 2 & \cdots & 2 & 2 \\ \cdots & & & \cdots & \\ 2 & 2 & \cdots & n-1 & 2 \\ 2 & 2 & \cdots & 2 & n \end{vmatrix}; \quad (2)^* \begin{vmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \\ 3 & 4 & \cdots & 1 & 2 \\ \cdots & & & \cdots & \\ n & 1 & \cdots & n-2 & n-1 \end{vmatrix}.$$

10. 证明下列等式:

$$(1)^* \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \cdots & & & \cdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_n^{n-2} \\ x_1^n & x_2^n & \cdots & x_n^n \end{vmatrix} = \left(\sum_{i=1}^n x_i \right) \prod_{1 \leq i < j \leq n} (x_j - x_i);$$

$$(2)^* \begin{vmatrix} a & c & c & \cdots & c \\ b & a & c & \cdots & c \\ b & b & a & \cdots & c \\ \cdots & & & \cdots & \\ b & b & b & \cdots & a \end{vmatrix} = \frac{b(a-c)^n - c(a-b)^n}{b-c},$$

其中 $b \neq c$, 等式左端是 n 阶行列式.

11. 设 A 是 n 阶可逆对称矩阵, B 是 n 阶反对称矩阵. 证明: 当 n 为奇数时, 齐次线性方程组 $(AB)X = O$ 有非零解.

12. 用行列式判断下列矩阵是否可逆, 如可逆, 用伴随矩阵求其逆矩阵:

$$(1) \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix};$$

$$(2) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (ad - bc \neq 0);$$

$$(3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix};$$

$$(4) \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix};$$

$$(5) \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & -1 & -2 \\ -1 & 1 & 1 & 2 \\ -1 & -1 & 1 & 0 \end{pmatrix};$$

$$(6) \begin{pmatrix} 1 & -2 & 0 & 0 \\ -3 & 4 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

13. 设 $A, B \in M_n(R)$, 判断下列命题是否正确, 如正确, 证明之, 如不正确, 举反例:

(1) 若 A, B 皆可逆, 则 $A + B$ 也可逆;

(2) 若 $A + B$ 可逆, 则 A, B 皆可逆;

(3) 若 AB 不可逆, 则 A, B 皆不可逆;

(4) 若 $A^T A = E$, 则 $|A|$ 等于 1 或 -1 ;

(5) 若 A 可逆, 则 $|A^{-1}| = |A|^{-1}$;

(6) 若 A 是幂等矩阵 (即 $A^2 = A$), 且 A 可逆, 则 $|A| = 1$;

(7) 若 A 是幂零矩阵 (即 $k \in \mathbb{N}^*$, 使 $A^k = O$) 则 A 不可逆.

14*. 设 A, B, C, D 都是 n 阶矩阵, $|A| \neq 0$ 且 $AC = CA$, 证明:

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

15*. 证明: $\begin{vmatrix} E_m & B \\ A & E_n \end{vmatrix} = |E_n - AB| = |E_m - BA|.$

16*. 设 $A, B \in M_n(R), \lambda \in F$, 证明:

$$|\lambda E - AB| = |\lambda E - BA|.$$

17. 用 Cramer 法则, 解线性方程组 $AX = b$:

$$(1) A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix};$$

$$(2) A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & -1 & -1 \\ 2 & 1 & 1 & 0 \\ -1 & 2 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}.$$

18. 已知对称轴平行于 Y 轴的抛物线过三点 $(1, -1)$ $(2, 1)$ $(-1, 7)$, 试求该抛物线的方程.

19. 圆的一般方程为 $(x^2 + y^2) + ax + by + c = 0$, 已知圆过三点 $(6, -1)$ $(4, 3)$ $(-3, 2)$, 求该圆的一般方程.

部分习题答案

1. (1) 4^4 ; (2) -16 ; (3) 1 ; (4) $-8i$; (5) 160 ; (6) $\lambda - 1$; $\lambda + 3$; (7) 0 ; (8) 31 ;

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

5. (1) $\prod_{0 \leq i < j \leq n} (i-j)$; (2) $(1 + (-1)^{n-1} \prod_{i=1}^n a_i) \prod_{1 \leq i < j \leq n} (a_j - a_i)$.

6. $\prod_{k=1}^n a_k \left(1 + \sum_{i=1}^n \frac{1}{a_i} \right)$.

8. (1) -35 ; (2) -4 ; (3) -3 ;

9. (1) $-2(n-2)!$; (2) $(-1)^{n(n-1)/2} (n+1) \frac{n^{n-1}}{2}$.

12. (1)(2)(3)(4)(6)可逆 (5)不可逆.

(1) $\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$; (2) $\frac{1}{ad-bc} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$;

(3) $\frac{1}{2} \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$; (4) $\begin{bmatrix} 1 & -a & \\ & 1 & -a \\ & & 1 \end{bmatrix}$;

(6) $-\frac{1}{4} \begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}$, $A_1 = \begin{bmatrix} 8 & 4 \\ 6 & 2 \end{bmatrix}$, $A_2 = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$.

13. (1)(2)(3)否 (4)(5)(6)(7)是.

14. 提示: 对分块矩阵作倍加初等变换, 使之化为上三角块矩阵.

17. $(1)(2, 0, -1)^T$; $(2)(1, 0, -1, 2)^T$.

18. $y = 1 - 4x + 2x^2$. 19. $x^2 + y^2 - 2x + 2y - 23 = 0$.

第 6 章 线性方程组与线性几何

对于一般的线性方程组 $AX = b$ (系数矩阵 A 不一定是方阵,即使 A 为方阵,其行列式也可能等于零),我们虽已会用高斯消元法求解,但总体上对线性方程组的求解问题还缺乏深刻的理性认识.例如,系数矩阵与增广矩阵满足什么条件时,方程组必有解;有解时,自由未知量的个数是否唯一确定;取不同的自由未知量求得的解集合是否相同;解集合具有怎样的结构.这一章将讨论这些问题.先讨论齐次线性方程组 $AX = 0$ 有非零解的条件及其解的结构,然后讨论非齐次线性方程组 $AX = b$ 有解的条件及其解的结构.

我们在第 4 章中讲过,如果 $\dim V_1(F) = n$, $\dim V_2(F) = m$,并给定它们的基 B_1 和 B_2 ,则线性映射 $\sigma \in L(V_1, V_2)$ 与矩阵 $A \in M_{m \times n}(F)$ 一一对应;相应地, α 和 $\sigma(\alpha) = \beta$ 分别关于基 B_1 和基 B_2 的坐标 $\alpha_{B_1} = X \in F^n$ 和 $\beta_{B_2} = b \in F^m$ 满足 $AX = b$.因此,求解非齐次线性方程组 $AX = b$,也就是求矩阵 A 所对应的线性映射 σ 关于 β 的完全原像 α 在基 B_1 的坐标,即

$$\sigma^{-1}(\beta) = \{\alpha \mid \alpha_{B_1} = X, X \text{ 是 } AX = b \text{ 的解}\}.$$

求解齐次线性方程组 $AX = 0$,也就是求矩阵 A 所对应的线性映射 σ 的核 $\text{Ker} \sigma$ 中向量 α 关于基 B_1 的坐标,即

$$\sigma^{-1}(0_2) = \{\alpha \mid \alpha_{B_1} = X, X \text{ 是 } AX = 0 \text{ 的解}\},$$

其中 0_2 是线性空间 V_2 的零元.

根据方程组 $AX = b$ 与方程 $\sigma(\alpha) = \beta$ 之间的对应关系,利用矩阵和线性映射的有关理论,就能解决本章所要讨论的问题.

由于空间平面和直线的一般方程都是线性方程,因此我们把平面、直线的图形称为线性图形,它们的几何问题(主要是它们的位置关系和度量关系)称为线性几何,本章也将用代数方法来讨论线性几何问题.

6.1 齐次线性方程组

定理 6.1 设矩阵 $A \in M_{m \times n}(F)$,若 $r(A) = r$,则齐次线性方程组 $AX = 0$ 的解空间 $N(A)$ 是 F^n 的一个 $n - r$ 维子空间.

证 设 A 是线性映射 $\sigma \in L(V_1, V_2)$ 关于 V_1 和 V_2 的基 B_1 和基 B_2 所对应的矩阵, 由

$$\mathfrak{r}(\sigma) = \mathfrak{r}(A) = r, \quad \mathfrak{r}(\sigma) + \dim(\text{Ker}\sigma) = \dim V_1 = n,$$

即得

$$\dim(\text{Ker}\sigma) = n - r.$$

再根据方程 $\sigma(\alpha) = 0_2$ 的解与方程组 $AX = 0$ 的解的对应关系以及节 4.8 中的引理可知 $\text{Ker}\sigma$ 的基 $\alpha_1, \alpha_2, \dots, \alpha_{n-r}$ 关于基 B_1 的坐标 X_1, X_2, \dots, X_{n-r} 就是 $AX = 0$ 的解空间 $N(A)$ 的基, 所以

$$\dim N(A) = n - r.$$

由定理 6.1 立即可得下面的推论.

推论 以 $m \times n$ 矩阵 A 为系数矩阵的齐次线性方程组 $AX = 0$ 有非零解的充要条件是

$$\mathfrak{r}(A) < n.$$

我们常把齐次线性方程组 $AX = 0$ 的解空间 $N(A)$ 的基 X_1, X_2, \dots, X_p 叫做 $AX = 0$ 的基础解系, 由于 $N(A) = L(X_1, X_2, \dots, X_p)$, 所以 $AX = 0$ 的任一个解都可表示为它的基础解系的线性组合, 而基础解系的任何线性组合也都是它的解, 因此, 我们把

$$X = k_1 X_1 + k_2 X_2 + \dots + k_p X_p$$

(其中 k_1, k_2, \dots, k_p 为任意常数) 叫做 $AX = 0$ 的一般解.

由此可见, 求解齐次线性方程组 $AX = 0$ 的关键是求它的基础解系(解空间的基), 从而得到一般解(解空间中任意解向量的表示式). 第 1 章中用高斯消元法所求得解, 例如, 节 1.8 中例 2 (方程右端系数为 0) 所得的解 $X = k_1(1, 1, 0, 0, 0) + k_2(-7, 0, -4, 3, 1)$, 其中两个解向量就是解空间的基.

现在我们再介绍求 $AX = 0$ 的基础解系的常用方法.

如果 $m \times n$ 矩阵 A 的秩为 r , 对 A 作初等行变换可将其化为有 r 个非零行的行简化阶梯形矩阵, 再作一些列对换(此时相应地改变未知元的次序)可使阶梯形矩阵的左上角为 r 阶单位矩阵, 为了表述简便, 不妨假设

$$A \xrightarrow{\text{初等行变换}} \begin{pmatrix} 1 & 0 & \dots & 0 & c_{1,r+1} & \dots & c_{1n} \\ 0 & 1 & \dots & 0 & c_{2,r+1} & \dots & c_{2n} \\ \dots & & & & & & \\ 0 & 0 & \dots & 1 & c_{r,r+1} & \dots & c_{rn} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} = U,$$

于是 $UX = \mathbf{0}$ 与 $AX = \mathbf{0}$ 是同解方程. $UX = \mathbf{0}$ 只含 r 个方程, 把 r 个方程中带 $x_{r+1} \ x_{r+2} \ \dots \ x_n$ 的项移到右端, 并选 $x_{r+1} \ x_{r+2} \ \dots \ x_n$ 为自由未知量, 对它们取下列 $n - r$ 组值

$$(1 \ 0 \ \dots \ 0) (0 \ 1 \ \dots \ 0) \ \dots \ (0 \ 0 \ \dots \ 1),$$

再分别代入 $UX = \mathbf{0}$, 即可得到 $AX = \mathbf{0}$ 的 $n - r$ 个解:

$$X_1 = (-c_{1,r+1} \ -c_{2,r+1} \ \dots \ -c_{r,r+1} \ 1 \ 0 \ \dots \ 0)^T,$$

$$X_2 = (-c_{1,r+2} \ -c_{2,r+2} \ \dots \ -c_{r,r+2} \ 0 \ 1 \ \dots \ 0)^T,$$

.....

$$X_{n-r} = (-c_{1n} \ -c_{2n} \ \dots \ -c_{rn} \ 0 \ 0 \ \dots \ 1)^T.$$

这 $n - r$ 个解显然是线性无关的, 所以它们是齐次线性方程组 $AX = \mathbf{0}$ 的基础解系.

这里矩阵 A 的秩 r 确定了 $AX = \mathbf{0}$ 的解空间的维数为 $n - r$, 自由未知量有 $n - r$ 个, 但自由未知量的取法不是唯一的, 然而取不同的自由未知量所得到的不同的基础解系, 都是同一个解空间的基, 因此, 由这些不同的基础解系表示的一般解所含全部解向量的集合是相同的.

我们把矩阵 $A \in M_{m \times n}(F)$ 的 n 个列向量 $\beta_1 \ \dots \ \beta_n$ 张成的子空间 $L(\beta_1, \dots, \beta_n)$ 叫做 A 的列空间, 记作 $R(A)$; A 的 m 个行向量 $\alpha_1 \ \dots \ \alpha_m$ 张成的子空间 $L(\alpha_1 \ \dots \ \alpha_m)$ 叫做 A 的行空间, 记作 $R(A^T)$. 显然, $R(A)$ 是 F^n 的子空间, $R(A^T)$ 是 F^m 的子空间.

由于 A 的任一个行向量 α_i 与 $AX = \mathbf{0}$ 的解向量 X 是正交的, 即 $(\alpha_i, X) = 0$, 所以 $R(A^T)$ 与 $N(A)$ 是正交子空间, 由定理 6.1 的结论

$$r(A) + \dim N(A) = n \quad (A \text{ 的列数}) \quad (6-1)$$

和 $r(A) = \dim R(A^T)$, 又得

$$N(A) = (R(A^T))^\perp. \quad (6-2)$$

(6-1)(6-2) 式是齐次线性方程组 $AX = \mathbf{0}$ 的解空间的两个基本结论.

例 1 设

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 & -1 \\ 2 & 6 & 9 & 5 & 4 \\ -1 & -3 & 3 & 1 & 13 \\ 0 & 0 & -3 & 1 & -6 \end{bmatrix}.$$

试求齐次线性方程组 $AX = \mathbf{0}$ 的一般解.

解 对矩阵 A 作初等行变换, 将 A 化为行简化阶梯形矩阵

$$\begin{pmatrix} 1 & 3 & 0 & 0 & -7 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

选 x_2, x_5 为自由未知量 取 $x_2 = 1, x_5 = 0$ 和 $x_2 = 0, x_5 = 1$ 得基础解系:

$$\mathbf{X}_1 = (-3, 1, 0, 0, 0)^T, \quad \mathbf{X}_2 = (7, 0, -2, 0, 1)^T,$$

所以 $AX = \mathbf{0}$ 的一般解为 $X = k_1 \mathbf{X}_1 + k_2 \mathbf{X}_2$ 即

$$(x_1, x_2, x_3, x_4, x_5)^T = k_1(-3, 1, 0, 0, 0)^T + k_2(7, 0, -2, 0, 1)^T,$$

其中 k_1, k_2 为任意常数.

例2 设 A, B 分别是 $m \times n$ 和 $n \times s$ 矩阵, 且 $AB = O$. 证明:

$$\mathfrak{r}(A) + \mathfrak{r}(B) \leqslant n.$$

证 将 B 按列分块为 $B = (B_1, B_2, \dots, B_s)$, 由 $AB = O$ 得

$$AB_j = \mathbf{0}, \quad j = 1, 2, \dots, s,$$

即 B 的每一列 B_j 都是 $AX = \mathbf{0}$ 的解, 所以

$$L(B_1, B_2, \dots, B_s) \subseteq N(A).$$

又 $\mathfrak{r}(B) = B$ 的列秩 $= \dim L(B_1, B_2, \dots, B_s)$, 所以 $\mathfrak{r}(B) \leqslant \dim N(A)$, 于是便有

$$\mathfrak{r}(A) + \mathfrak{r}(B) \leqslant \mathfrak{r}(A) + \dim N(A) = n.$$

例3 设 A 是 $m \times n$ 实矩阵, 证明: $\mathfrak{r}(A^T A) = \mathfrak{r}(A)$.

证 由定理 4.9 可知 $\mathfrak{r}(A^T A) \leqslant \mathfrak{r}(A)$, 因此, 只需证明 $\mathfrak{r}(A^T A) \geqslant \mathfrak{r}(A)$. 根据定理 6.1, 为此只要证明 $N(A^T A) \subseteq N(A)$.

设 $(A^T A)X = \mathbf{0} (X \in \mathbf{R}^n)$, 则 $X^T(A^T A)X = 0$, 即

$$(AX)^T AX = |AX|^2 = 0.$$

又 $AX \in \mathbf{R}^m$, 故必有 $AX = \mathbf{0}$. 因此 $(A^T A)X = \mathbf{0}$ 的解都是 $AX = \mathbf{0}$ 的解, 即 $N(A^T A) \subseteq N(A)$, 所以 $\dim N(A^T A) \leqslant \dim N(A)$, 即

$$n - \mathfrak{r}(A^T A) \leqslant n - \mathfrak{r}(A),$$

由此便得

$$\mathfrak{r}(A^T A) \geqslant \mathfrak{r}(A).$$

6.2 非齐次线性方程组

对于非齐次线性方程组 $AX = b$ (其中 $A \in M_{m \times n}, b \in F^m$ 且 $b \neq 0$), 将

矩阵 A 按列分块为

$$A = (\beta_1, \beta_2, \dots, \beta_n),$$

则非齐次线性方程组

$$AX = \begin{pmatrix} | & | & & | \\ \beta_1 & \beta_2 & \dots & \beta_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = b$$

等价于向量方程

$$x_1\beta_1 + x_2\beta_2 + \dots + x_n\beta_n = b. \quad (6-3)$$

因此 $AX = b$ 有解 (即存在 (x_1, x_2, \dots, x_n) 使 (6-3) 式成立) 的充分必要条件是 b 可被 A 的列向量组 $\beta_1, \beta_2, \dots, \beta_n$ 线性表示, 即 $b \in R(A)$. 这又等价于

$$\text{秩}\{\beta_1, \beta_2, \dots, \beta_n, b\} = \text{秩}\{\beta_1, \beta_2, \dots, \beta_n\},$$

即 $\text{r}(A, b) = \text{r}(A)$. 如此, 就有下面的定理.

定理 6.2 对于非齐次线性方程组 $AX = b$, 下列命题等价:

- (1) $AX = b$ 有解;
- (2) $b \in R(A)$, 即 b 可被 A 的列向量组线性表示;
- (3) $\text{r}(A, b) = \text{r}(A)$, 即增广矩阵的秩等于系数矩阵的秩.

由定理 6.2 可知, $AX = b$ 有唯一解的充要条件是: $b \in R(A)$, 且 A 的列向量组 $\beta_1, \beta_2, \dots, \beta_n$ 线性无关. 于是又有下面的推论.

推论 非齐次线性方程组 $AX = b$ 有唯一解的充要条件是

$$\text{r}(A, b) = \text{r}(A) = A \text{ 的列数}.$$

非齐次线性方程组 $AX = b$ (其中 $A \in M_{m \times n}$, $b \in F^m$) 的解集合

$$S = \{X | AX = b, X \in F^n\}$$

不是 F^n 的子空间, 因为 S 对向量的加法和数乘运算不封闭. 但是如果 $X_1, X_2 \in S$, 则 $X_1 - X_2$ 是对应 (或相伴的) 齐次线性方程组 $AX = 0$ 的解, 因为

$$A(X_1 - X_2) = AX_1 - AX_2 = b - b = 0.$$

利用这个性质, 我们可得到非齐次线性方程组的解的结构 (如定理 6.3 所述).

定理 6.3 若非齐次线性方程组 $AX = b$ 有解, 则其一般解为

$$X = X_0 + \bar{X}, \quad (6-4)$$

其中 X_0 是 $AX = b$ 的一个特解, \bar{X} 是 $AX = 0$ 的一般解.

证 显然, $X = X_0 + \bar{X}$ 是方程组 $AX = b$ 的解. 再由于 X 和 X_0 都是 AX

$= b$ 的解, 所以 $X - X_0$ 是 $AX = 0$ 的解, 即 $X - X_0 \in N(A)$, 因此 $X - X_0$ 的一般形式为

$$X - X_0 = k_1 X_1 + k_2 X_2 + \dots + k_p X_p = \overline{X}, \quad (6-5)$$

其中 X_1, X_2, \dots, X_p 为 $AX = 0$ 的基础解系, \overline{X} 为其一般解, 由此即得 $AX = b$ 的一般解如(6-4)式所述.

需要指出(6-4)式中的 X_0 可以是 $AX = b$ 的任一个特解, 因为如果 X_0^* 也是 $AX = b$ 的一个特解, 则 $X_0^* - X_0$ 必是 $AX = 0$ 的一个解, 所以, $AX = b$ 的一般解也可表示为

$$X = X_0^* + \overline{X}. \quad (6-6)$$

读者不难验证, 由(6-4)和(6-6)式表示的 $AX = b$ 的解集合 S 和 S^* 是相等的.

例1 设非齐次线性方程组 $AX = b$ 的增广矩阵为

$$(A, b) = \left(\begin{array}{ccccc|c} 1 & 1 & -1 & -1 & 1 & 0 \\ 2 & 2 & 1 & 0 & 1 & 1 \\ 3 & 3 & 0 & -1 & 2 & 1 \\ 1 & 1 & 2 & 1 & 0 & 1 \end{array} \right).$$

试求 $AX = b$ 的一般解.

解

$$(A, b) \xrightarrow{\text{初等行变换}} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & -\frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) = (U, d).$$

取 $x_2 = x_4 = x_5 = 0$ 代入 $UX = d$ 求得 $AX = b$ 的一个特解

$$X_0 = \left(\frac{1}{3}, 0, \frac{1}{3}, 0, 0 \right)^T.$$

取自由未知量 x_2, x_4, x_5 的下列三组数

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

并依次代入 $UX = 0$ 得 $AX = 0$ 的基础解系:

$$X_1 = (-1, 1, 0, 0, 0)^T,$$

$$X_2 = \left(\frac{1}{3}, 0, -\frac{2}{3}, 1, 0 \right)^T,$$

$$\mathbf{X}_3 = \left(-\frac{2}{3} \ 0 \ \frac{1}{3} \ 0 \ 1 \right)^T.$$

显然, 基础解系也可取为:

$$\mathbf{X}_1^* = (-1 \ 1 \ 0 \ 0 \ 0)^T,$$

$$\mathbf{X}_2^* = (1 \ 0 \ -2 \ 3 \ 0)^T,$$

$$\mathbf{X}_3^* = (-2 \ 0 \ 1 \ 0 \ 3)^T.$$

于是, $AX = b$ 的一般解为

$$\begin{aligned} X &= X_0 + k_1 X_1^* + k_2 X_2^* + k_3 X_3^* \\ &= \left(\frac{1}{3} \ 0 \ \frac{1}{3} \ 0 \ 0 \right)^T + k_1 (-1 \ 1 \ 0 \ 0 \ 0)^T \\ &\quad + k_2 (1 \ 0 \ -2 \ 3 \ 0)^T + k_3 (-2 \ 0 \ 1 \ 0 \ 3)^T \end{aligned}$$

其中 k_1, k_2, k_3 为任意常数.

例 2 设线性方程组

$$\begin{cases} ax_1 + x_2 + x_3 = 4 \\ x_1 + bx_2 + x_3 = 3 \\ x_1 + 2bx_2 + x_3 = 4. \end{cases}$$

讨论方程组的解的情况与参数 a, b 的关系, 有解时, 求其解.

解 此题可以就系数行列式 $D = b(1-a)$ 的取值情况: $D \neq 0$; $D = 0$ ($(1) a = 1, b \neq 0$ ($2) b = 0, a \neq 1$ ($3) a = 1, b = 0$), 分别讨论方程组的解的情况. 也可用初等行变换将增广矩阵化为阶梯阵, 来讨论其解与参数 a, b 的关系. 这里用后面的方法更简便, 即

$$\begin{aligned} &\left(\begin{array}{ccc|c} a & 1 & 1 & 4 \\ 1 & b & 1 & 3 \\ 1 & 2b & 1 & 4 \end{array} \right) \xrightarrow{\text{行对换}} \left(\begin{array}{ccc|c} 1 & b & 1 & 3 \\ 1 & 2b & 1 & 4 \\ a & 1 & 1 & 4 \end{array} \right) \\ &\xrightarrow{\substack{[2]+[1] \times (-1) \\ [3]+[1] \times (-a)}} \left(\begin{array}{ccc|c} 1 & b & 1 & 3 \\ 0 & b & 0 & 1 \\ 0 & 1-ab & 1-a & 4-3a \end{array} \right) \\ &\xrightarrow{[3]+[2] \times a} \left(\begin{array}{ccc|c} 1 & b & 1 & 3 \\ 0 & b & 0 & 1 \\ 0 & 1 & 1-a & 4-2a \end{array} \right) \\ &\xrightarrow{\substack{[2]+[3] \times (-b) \\ [2] \leftrightarrow [3]}} \left(\begin{array}{ccc|c} 1 & b & 1 & 3 \\ 0 & 1 & 1-a & 4-2a \\ 0 & 0 & (a-1)b & 1-4b+2ab \end{array} \right). \end{aligned}$$

(1) 当 $(a-1)b \neq 0$ 时, 有唯一解

$$x_1 = \frac{2b-1}{(a-1)b}, \quad x_2 = \frac{1}{b}, \quad x_3 = \frac{1-4b+2ab}{(a-1)b}.$$

(2) 当 $a=1$ 且 $1-4b+2ab=1-2b=0$ 即 $b=\frac{1}{2}$ 时, 有无穷多解, 此时增广矩阵化为

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & 1 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right),$$

于是方程组的一般解为

$$\mathbf{X} = (2 \ 2 \ 0)^T + k(-1 \ 0 \ 1)^T \quad (k \text{ 为任意常数}).$$

(3) 当 $a=1, b \neq \frac{1}{2}$ 时, 方程组无解.

(4) 当 $b=0$ 时, $1-4b+2ab=1 \neq 0$, 故方程组无解. 此时, 原方程组中后两个方程是矛盾方程.

例3 证明 若 \mathbf{X}_0 是 $A\mathbf{X} = \mathbf{b}$ 的一个特解, $\mathbf{X}_1, \dots, \mathbf{X}_p$ 是 $A\mathbf{X} = \mathbf{0}$ 的基础解系, 则 $\mathbf{X}_0, \mathbf{X}_0 + \mathbf{X}_1, \mathbf{X}_0 + \mathbf{X}_2, \dots, \mathbf{X}_0 + \mathbf{X}_p$ 线性无关, 且 $A\mathbf{X} = \mathbf{b}$ 的任一个解 \mathbf{X} 可表示为

$$\mathbf{X} = k_0\mathbf{X}_0 + k_1(\mathbf{X}_0 + \mathbf{X}_1) + k_2(\mathbf{X}_0 + \mathbf{X}_2) + \dots + k_p(\mathbf{X}_0 + \mathbf{X}_p),$$

其中 $k_0 + k_1 + k_2 + \dots + k_p = 1$.

证 设

$$c_0\mathbf{X}_0 + c_1(\mathbf{X}_0 + \mathbf{X}_1) + c_2(\mathbf{X}_0 + \mathbf{X}_2) + \dots + c_p(\mathbf{X}_0 + \mathbf{X}_p) = \mathbf{0},$$

即

$$(c_0 + c_1 + c_2 + \dots + c_p)\mathbf{X}_0 + c_1\mathbf{X}_1 + c_2\mathbf{X}_2 + \dots + c_p\mathbf{X}_p = \mathbf{0},$$

则必有 $c_0 + c_1 + c_2 + \dots + c_p = 0$ (否则 $\mathbf{X}_0 = d_1\mathbf{X}_1 + d_2\mathbf{X}_2 + \dots + d_p\mathbf{X}_p$ 是 $A\mathbf{X} = \mathbf{0}$ 的解, 与题设矛盾). 再由 $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$ 线性无关得 $c_1 = c_2 = \dots = c_p = 0$, 从而 $c_0 = 0$, 故 $\mathbf{X}_0, \mathbf{X}_0 + \mathbf{X}_1, \dots, \mathbf{X}_0 + \mathbf{X}_p$ 线性无关.

根据定理 6.3, $A\mathbf{X} = \mathbf{b}$ 的任一个解, 可表示为

$$\begin{aligned} \mathbf{X} &= \mathbf{X}_0 + k_1\mathbf{X}_1 + k_2\mathbf{X}_2 + \dots + k_p\mathbf{X}_p \\ &= (1 - k_1 - \dots - k_p)\mathbf{X}_0 + k_1(\mathbf{X}_0 + \mathbf{X}_1) + \dots + k_p(\mathbf{X}_0 + \mathbf{X}_p). \end{aligned}$$

令 $(1 - k_1 - \dots - k_p) = k_0$, 则 $k_0 + k_1 + k_2 + \dots + k_p = 1$, 于是命题得证.

6.3 线性图形的几何问题

6.3.1 线性图形的位置关系

线性图形是指空间的平面和直线,它们之间的位置关系,就是平面与平面、平面与直线以及直线与直线之间的位置关系.由于空间的平面和直线在直角坐标系下的方程,是三元线性方程 $a_1x_1 + a_2x_2 + a_3x_3 = b$ 和两个三元线性方程组成的方程组,因此讨论它们之间的位置关系(如平行、重合、相交等),可用线性方程组的解的理论阐明.

1 两个平面间的位置关系有平行、重合和相交的三种情况.从两个平面方程构成的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases}$$

可见 (i) 当系数矩阵 A 与增广矩阵 (A, b) 的秩均为 2 时,方程组有无穷多组解,其一般解中含一个任意常数,即两平面相交为一直线; (ii) 当 $r(A) = r(A, b) = 1$ 时,增广矩阵的两个行向量成比例,故两个平面是重合的; (iii) 当 $r(A) = 1$ (此时两个平面的法向量平行), $r(A, b) = 2$ 时,方程组无解,即两个平面平行,且不重合.

2 平面与直线间的位置关系有直线平行于平面,直线在平面上和直线与平面相交于一点的三种情况.从三元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (1) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 & (2) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & (3) \end{cases}$$

(其中 (1) 是平面 π 的方程, (2) (3) 联立是直线 L 的方程,其系数矩阵与增广矩阵的秩均为 2),可见方程组的增广矩阵 (A, b) 的秩 ≥ 2 .

(i) 当 $r(A) = r(A, b) = 3$ 时,方程组有唯一解,直线 L 与平面 π 相交于一个点,此时 $\det A \neq 0$,三个平面法向量 n_1, n_2, n_3 线性无关,即不共面;

(ii) 当 $r(A) = r(A, b) = 2$ 时,方程组有无穷多解(求解时有一个自由未知量),解集合的图形是一条直线,即直线 L 位于平面 π 上,此时,平面 π 的法向量 n_1 垂直于直线 L 的方向向量 $S = n_2 \times n_3$,从而 $n_1 \cdot (n_2 \times n_3) = \det A = 0$;

(iii) 当 $r(A) = 2, r(A, b) = 3$ 时,方程组无解,直线 L 平行于平面 π (但

不位于平面 π 上), 此时 n_1 也垂直于 $S = (n_2 \times n_3)$.

平面与直线间的位置关系, 也可视为三个平面之间的关系. 如此, 以上前两种情况分别是三个平面交于一点和交于一直线; 第三种(方程无解)情况除了上面所说的两个平面的交线平行于另一个平面以外, 还可能三个平面平行且至少有两个不重合的情况. 此时 $r(A) = 1$, $r(A, b) = 2$. 对于三个平面间的关系, 还需指出三个平面重合(即共面)的情况, 即 $r(A) = r(A, b) = 1$.

至于 m ($m > 3$) 个平面间的关系, 从方程组的系数矩阵和增广矩阵的秩的关系看, 不外乎上述各种情况. 方程组有解时, 有三种情况: m 个平面交于一点, m 个平面交于一直线, m 个平面都重合. 方程组无解时, m 个平面间的关系有三种类型: 一是有两个平面平行而不重合; 二是有两个平面相交, 但交线平行于另一个平面; 三是有三个平面交于一点 P , 但点 P 不在另一个平面上. 后两种类型可以统一为: 平面两两的交线中, 有两条交线不相交.

3 两条直线间的位置关系有相交、平行和不共面(即交叉不相交)的三种情况. 从线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & (1) \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 & (2) \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & (3) \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 = b_4 & (4) \end{cases}$$

(其中前两个和后两个方程的联立分别表示直线 L_1 和 L_2) 可见:

(i) 当 $r(A) = r(A, b) = 3$ 时, 方程组有唯一解, 此时, 两条直线交于一点;

(ii) 当 $r(A) = r(A, b) = 2$ 时, 方程组有无穷多解, 此时两条直线重合;

(iii) 当 $r(A) = 2$, $r(A, b) = 3$ 时, 方程组无解, 此时四个平面的法向量共面, 从而两条直线的方向向量平行, 即 $S_1 = n_1 \times n_2 // n_3 \times n_4 = S_2$, 但直线 L_1 上任一点都不在直线 L_2 上, 所以 L_1 与 L_2 平行不相交;

(iv) 当 $r(A) = 3$, $r(A, b) = 4$ 时, 方程组也无解, 此时两条直线不平行, 也不相交(称为相错直线), 从而不共面.

如果从直线的标准方程

$$\begin{aligned} L_1: \quad & \frac{x-x_1}{l_1} = \frac{y-y_1}{m_1} = \frac{z-z_1}{n_1}, \\ L_2: \quad & \frac{x-x_2}{l_2} = \frac{y-y_2}{m_2} = \frac{z-z_3}{n_3} \end{aligned}$$

来考察两条直线间的关系, 则直线 L_1 与 L_2 重合, 当且仅当它们方向向量

$S_1 = (l_1, m_1, n_1)$, $S_2 = (l_2, m_2, n_2)$, 以及 L_1 和 L_2 上任意的点 $P_1(x_1, y_1, z_1)$ 和 $P_2(x_2, y_2, z_2)$ 连成的向量 $\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ 互相平行, 即秩 $\{S_1, S_2, \overrightarrow{P_1P_2}\} = 1$;

L_1 与 L_2 平行, 当且仅当 $S_1 \parallel S_2$, 但不平行于 $\overrightarrow{P_1P_2}$, 即秩 $\{S_1, S_2\} = 1$ 秩 $\{S_1, S_2, \overrightarrow{P_1P_2}\} = 2$;

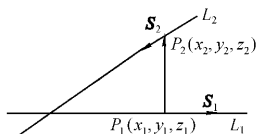


图 6-1

L_1 与 L_2 相交, 当且仅当 S_1, S_2 不平行, 且它们与 $\overrightarrow{P_1P_2}$ 共面, 即秩 $\{S_1, S_2\} = 2$ 秩 $\{S_1, S_2, \overrightarrow{P_1P_2}\} = 2$ 或

$$\overrightarrow{P_1P_2} \cdot (\vec{S}_1 \times \vec{S}_2) = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{vmatrix} = 0;$$

L_1 与 L_2 异面(交叉不相交), 当且仅当 S_1, S_2 与 $\overrightarrow{P_1P_2}$ 不共面, 即秩 $\{S_1, S_2, \overrightarrow{P_1P_2}\} = 3$ 或 $\overrightarrow{P_1P_2} \cdot (S_1 \times S_2) \neq 0$.

6.3.2 线性图形的度量关系

线性图形的度量关系是指线性图形间的距离和角度. 这里要讨论: 点到平面和直线的距离; 平面与平面, 直线与平面以及直线与直线间的距离和夹角.

1 点到平面和点到直线的距离

点 $P(x_1, y_1, z_1)$ 到平面 $\pi: a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ 的距离 d 通常是指点到平面的最短距离, 即点到平面的垂直距离 d . 因此求 d 的一种方法是: 过点 P 作垂直于平面 π 的直线 L :

$$\frac{x - x_1}{a} = \frac{y - y_1}{b} = \frac{z - z_1}{c}.$$

然后, 解方程组求出 L 与 π 的交点 $Q(x_2, y_2, z_2)$, 得

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

但是, 更简便的方法是利用几何向量的内积运算, 由图 6-2 易得

$$d = | \overrightarrow{P_0P} \cdot \mathbf{n}^0 | = | | \overrightarrow{P_0P} | \cos \theta |, \quad (6-7)$$

其中 $P_0(x_0, y_0, z_0)$ 为平面 π 上的任一个点, \mathbf{n}^0 为平面 π 的单位长的法向量, 即

$$\mathbf{n}^0 = \frac{1}{\sqrt{a^2 + b^2 + c^2}}(a, b, c),$$

于是

$$d = \frac{1}{\sqrt{a^2 + b^2 + c^2}} |a(x_1 - x_0) + b(y_1 - y_0) + c(z_1 - z_0)|.$$

如果将平面 π 的点法式方程 $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ 写为一般方程

$$ax + by + cz + f = 0,$$

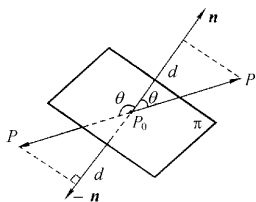


图 6-2

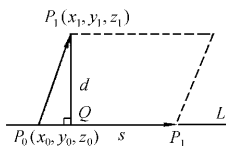


图 6-3

其中 $f = -ax_0 - by_0 - cz_0$, 则点 $P(x_1, y_1, z_1)$ 到平面 π 的距离为

$$d = \frac{1}{\sqrt{a^2 + b^2 + c^2}} |ax_1 + by_1 + cz_1 + f|. \quad (6-8)$$

点 $P(x_1, y_1, z_1)$ 到直线 $L: \frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n}$ 的距离 (即点 P 到 L 的垂直距离) d , 由图 6-3 易得

$$d = \frac{|\mathbf{S} \times \overrightarrow{P_0P}|}{|\mathbf{S}|}, \quad (6-9)$$

其中 $\mathbf{S} = \overrightarrow{P_0P_1} = (l, m, n)$; $\overrightarrow{P_0P} = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$.

$|\mathbf{S} \times \overrightarrow{P_0P}|$ 等于图 6-3 中平行四边形的面积.

求点 P 到 L 的垂直距离 d , 也可过点 P 作垂直于 L 的平面 $\pi: l(x - x_1) + m(y - y_1) + n(z - z_1) = 0$, 再求 L 与 π 的交点 $Q(x_2, y_2, z_2)$, 得 $d = |\overrightarrow{PQ}|$.

2 平面与平面间的距离和夹角

两个平面 π_1 与 π_2 之间的距离 d ,是指两个平面上任意两点 $P_1 \in \pi_1$ 与 $P_2 \in \pi_2$ 之间距离的最小值. 因此,两个相交平面的距离为零;两个平行平面(但不重合)

$$\pi_1: ax + by + cz + f_1 = 0, \quad \pi_2: ax + by + cz + f_2 = 0$$

($f_1 \neq f_2$) 间的距离,由图 6-4 易得

$$d = |\overrightarrow{P_0P_1} \cdot \mathbf{n}^0|. \quad (6-10)$$

其中:点 $P_1(x_1, y_1, z_1)$ 与点 $P_2(x_2, y_2, z_2)$ 分别是 π_1 与 π_2 上任意定点; \mathbf{n}^0 是平面法向量 $\mathbf{n} = (a, b, c)$ 方向的单位向量.

两个平面 π_1 与 π_2 之间的夹角 θ 就是两个法向量 \mathbf{n}_1 与 \mathbf{n}_2 的夹角,它们有互补的两个夹角,我们规定 $\leq 90^\circ$ 那个角为其夹角(如图 6-5),因此:

$$\cos \theta = |\mathbf{n}_1^0 \cdot \mathbf{n}_2^0|, \quad \text{或} \quad \theta = \arccos |\mathbf{n}_1^0 \cdot \mathbf{n}_2^0|. \quad (6-11)$$

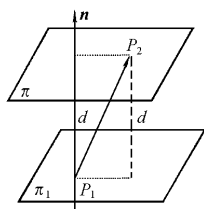


图 6-4

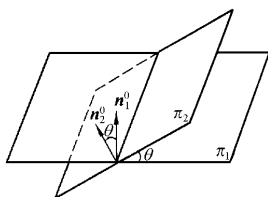


图 6-5

3 直线与直线间的距离和夹角

两条直线 L_1 与 L_2 间的距离 d ,也是指两条直线上任意两点 $P_1 \in L_1$ 与 $P_2 \in L_2$ 之间距离的最小值. 因此两条直线

$$L_1: \frac{x - x_1}{l_1} = \frac{y - y_1}{m_1} = \frac{z - z_1}{n_1},$$

$$L_2: \frac{x - x_2}{l_2} = \frac{y - y_2}{m_2} = \frac{z - z_2}{n_2}$$

之间的距离 d 为:

当 L_1 与 L_2 相交时 $d = 0$.

当 $\mathbf{S}_1 = (l_1, m_1, n_1) \nparallel \mathbf{S}_2 = (l_2, m_2, n_2)$ 时, L_1 与 L_2 之间的距离 d 是 L_2 上任一点(取 $P_2(x_2, y_2, z_2)$)到 L_1 的距离,于是由(6-9)式,得

$$d = \frac{|\mathbf{s}_1 \times \overrightarrow{P_1 P_2}|}{|\mathbf{s}_1|}, \quad (6-12)$$

其中 $P_1(x_1, y_1, z_1)$ 是直线 L_1 上的点.

当 L_1 与 L_2 为既不平行又不相交的相错直线时, 它们间的最短距离 d 是公垂线 L 与 L_1, L_2 的交点 M_1 和 M_2 之间的距离, 即 $d = |\overrightarrow{M_1 M_2}|$ (如图 6-6).

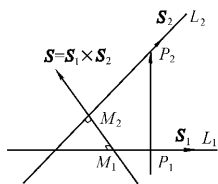


图 6-6

公垂线 L 的方向向量 $\mathbf{s} = \mathbf{s}_1 \times \mathbf{s}_2$, 公垂线 L 是平面 π_1 (过 L_1 且平行于 \mathbf{s} 的平面) 与 π_2 (过 L_2 且平行于 \mathbf{s} 的平面) 的交线. 公垂线 L 与 L_1, L_2 的交点 M_1, M_2 分别是 L_1 与 π_2 和 L_2 与 π_1 的交点. 由图 6-6 易见

$$d = |\overrightarrow{M_1 M_2}| = \left| |\overrightarrow{P_1 P_2}| \cos \angle \overrightarrow{P_1 P_2}, \mathbf{s} \right|,$$

故

$$d = \frac{|\overrightarrow{P_1 P_2} \cdot \mathbf{s}|}{|\mathbf{s}|} = |\overrightarrow{P_1 P_2} \cdot \mathbf{s}^0| \quad (6-13)$$

是 L_1 与 L_2 间的最短距离 (其中 $P_1 \in L_1, P_2 \in L_2$) 因为 $\forall P_1 \in L_1, P_2 \in L_2$, 有

$$|\overrightarrow{P_1 P_2}| = |\overrightarrow{P_1 P_2}| |\mathbf{s}^0| \geq |\overrightarrow{P_1 P_2} \cdot \mathbf{s}^0|.$$

两条直线 L_1 与 L_2 间的夹角 θ , 不论它们是相交、平行, 还是相错, 都规定 $\theta = \angle \mathbf{s}_1, \mathbf{s}_2$, 即

$$\theta = \arccos(\mathbf{s}_1^0 \cdot \mathbf{s}_2^0), \quad (6-14)$$

其中 $\mathbf{s}_1, \mathbf{s}_2$ 是 L_1, L_2 的方向向量. 由于直线的方向向量可取 \mathbf{s}_1 , 也可取 $-\mathbf{s}_1$, 所以 θ 和 $\pi - \theta$ 都是 L_1 与 L_2 的夹角.

4 直线与平面间的距离和夹角

直线 L 与平面 π 间的距离 d , 也是指 L 与 π 上任意两点 $P_1(x_1, y_1, z_1) \in L$ 与 $P_2(x_2, y_2, z_2) \in \pi$ 间距离的最小值. 当 L 与 π 相交时, $d = 0$; 当

$$L: \frac{x-x_1}{l} = \frac{y-y_1}{m} = \frac{z-z_1}{n},$$

$$\pi: a(x-x_2)+b(y-y_2)+c(z-z_2)=0$$

平行时, 即 $S \cdot n = (l, m, n) \cdot (a, b, c) = al + bm + cn = 0$ 时,

$$\begin{aligned} d &= |\overrightarrow{P_1P_2} \cdot n^0| \\ &= \frac{a(x_2-x_1)+b(y_2-y_1)+c(z_2-z_1)}{\sqrt{a^2+b^2+c^2}}. \end{aligned} \quad (6-14)$$

直线 L 与平面 π 的夹角: 当 $n \parallel S$ (即 $n = \lambda S$) 时, L 与 π 垂直 (或正交); 当 $n \perp S$ (即 $n \cdot S = 0$) 时, L 与 π 平行; 当 L 与 π 相交 (但不正交) 时, L 与 π 的夹角 θ 规定为, L 在 π 上的投影直线 L_1 (即过 L 且平行于平面 π 的法向量 n 的平面 π_1 与平面 π 的交线) 与 L 所夹之锐角 (如图 6-7), 于是

$$\theta = \arcsin \frac{|n \cdot S|}{|n| |S|}. \quad (6-15)$$

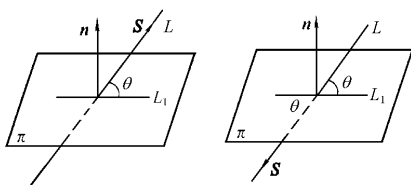


图 6-7

例1 已知直线 L_1, L_2, L_3 和平面 π 的方程为:

$$L_1: \frac{x-1}{3} = \frac{y-7}{-1} = \frac{z+4}{2}, \quad L_2: \frac{x-1}{1} = \frac{y+6}{4} = \frac{z-1}{-1},$$

$$L_3: \frac{x-1}{1} = \frac{y+2}{-2} = \frac{z}{2} \quad \pi: x - y - 3z - 1 = 0.$$

- (1) 验证 L_1 与 L_2 相交, 并求交点和夹角;
- (2) 求 L_1 与 L_3 之间的最短距离;
- (3) 求 L_2 与 π 间的距离; (4) 求 L_3 与 π 的交点及夹角 θ .

解 $P_1(1, 7, -4), P_2(1, -6, 1), P_3(1, -2, 0), P_0(1, 0, 0)$ 分别是 L_1, L_2, L_3, π 上的点.

(1) 由于

$$\overrightarrow{P_1P_2} \cdot (\mathbf{S}_1 \times \mathbf{S}_2) = \begin{vmatrix} 0 & -13 & 5 \\ 3 & -1 & 2 \\ 1 & 4 & -1 \end{vmatrix} = 0$$

及 $\mathbf{S}_1, \mathbf{S}_2$ 不平行, 所以 L_1 与 L_2 共面且相交. 求交点时可将 L_1, L_2 的方程改写成参数方程:

$$L_1: x = 1 + 3t_1, \quad y = 7 - t_1, \quad z = -4 + 2t_1,$$

$$L_2: x = 1 + t_2, \quad y = -6 + 4t_2, \quad z = 1 - t_2.$$

由 L_1 与 L_2 的交点 $P(x, y, z)$ 同时满足 L_1 与 L_2 的方程, 即得

$$\begin{cases} 1 + 3t_1 = 1 + t_2 \\ 7 - t_1 = -6 + 4t_2 \\ -4 + 2t_1 = 1 - t_2 \end{cases} \quad \text{即} \quad \begin{cases} 3t_1 - t_2 = 0 \\ t_1 + 4t_2 = 13 \\ 2t_1 + t_2 = 5 \end{cases} \quad (1)$$

解方程组(1)得: $t_1 = 1, t_2 = 3$, 故 L_1 与 L_2 的交点为 $P(4, 6, -2)$.

L_1 与 L_2 的夹角 θ 为

$$\begin{aligned} \theta &= \arccos(\mathbf{S}_1^0 \cdot \mathbf{S}_2^0) = \arccos \frac{1}{\sqrt{14}\sqrt{18}}(3, -1, 2) \cdot (1, 4, -1) \\ &= \arccos \frac{-1}{2\sqrt{7}} = 100.9^\circ. \end{aligned}$$

(2) 容易验证 L_1 与 L_3 不共面(即 $\overrightarrow{P_1P_3} \cdot (\mathbf{S}_1 \times \mathbf{S}_3) \neq 0$), 是两条相错直线, 它们的公垂线的方向向量为

$$\mathbf{S} = \mathbf{S}_1 \times \mathbf{S}_3 = (3, -1, 2) \times (1, -2, 2) = (2, -4, -5).$$

L_1 与 L_2 之间的最短距离为

$$d = |\overrightarrow{P_1P_3} \cdot \mathbf{S}^0| = \left| \frac{1}{\sqrt{45}}(2, -4, -5) \cdot (0, -9, 4) \right| = \frac{16}{3\sqrt{5}}.$$

(3) 由于 π 的法向量 \mathbf{n} 与 L_2 的方向向量 \mathbf{S}_2 的内积 $\mathbf{n} \cdot \mathbf{S}_2 = 1 - 4 + 3 = 0$, 所以 $L_2 \parallel \pi$, 它们间的距离

$$d = |\overrightarrow{P_0P_2} \cdot \mathbf{n}^0| = \left| (0, -6, 1) \cdot \frac{(1, -1, -3)}{\sqrt{11}} \right| = \frac{3}{\sqrt{11}}.$$

(4) 将 L_3 的参数方程 $x = 1 + t, y = -2 - 2t, z = 2t$ 代入平面 π 的方程, 得

$$2 - 3t = 0, \quad t = \frac{2}{3},$$

故 L_3 与 π 的交点 $P\left(\frac{5}{3}, -\frac{10}{3}, \frac{4}{3}\right)$. 它们间的夹角为

$$\theta = \arcsin \frac{|(1, -1, -3) \cdot (1, -2, 2)|}{\sqrt{11} \sqrt{9}} = \arcsin \frac{1}{\sqrt{11}} \approx 17.55^\circ.$$

例 2 点 P 把有向线段 $\overrightarrow{P_1P_2}$ 分成定比 λ , 即 $\overrightarrow{P_1P} = \lambda \overrightarrow{PP_2}$ (如图 6-8), 已知 $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2)$, 求 $P(x, y, z)$.

解 由

$$\overrightarrow{P_1P} = \mathbf{r} - \mathbf{r}_1, \quad \overrightarrow{PP_2} = \mathbf{r}_2 - \mathbf{r}, \quad \text{得 } \mathbf{r} - \mathbf{r}_1 = \lambda(\mathbf{r}_2 - \mathbf{r}),$$

于是

$$(1 + \lambda)\mathbf{r} = \mathbf{r}_1 + \lambda\mathbf{r}_2.$$

由于 $P_1 \neq P_2, \lambda \neq -1$, 所以 $\mathbf{r} = \frac{1}{1 + \lambda}(\mathbf{r}_1 + \lambda\mathbf{r}_2)$, 即

$$(x, y, z) = \left(\frac{x_1 + \lambda x_2}{1 + \lambda}, \frac{y_1 + \lambda y_2}{1 + \lambda}, \frac{z_1 + \lambda z_2}{1 + \lambda} \right).$$

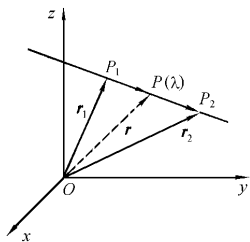


图 6-8

给定了 P_1 和 P_2 , 点 P 由 λ 唯一确定. 当 $\lambda = 1$ 时, 点 P 是线段 $\overrightarrow{P_1P_2}$ 的中点, 其坐标为

$$\left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}, \frac{z_1 + z_2}{2} \right).$$

当 $\lambda > 0$ 时, 点 P 在 P_1 与 P_2 之间, 称为内分点; 当 $-1 < \lambda < 0$ 时, 点 P 在 P_1 之“外侧”(即 $\overrightarrow{P_1P}$ 与 $\overrightarrow{PP_2}$ 反向)称为 P_1 侧的外分点; 当 $\lambda < -1$ 时, 点 P 在 P_2 之“外侧”, 称为 P_2 侧的外分点.

习 题

1. 求下列齐次线性方程组的一个基础解系及一般解:

$$(1) \begin{cases} x_1 - x_2 + 5x_3 - x_4 = 0 \\ x_1 + x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 - x_2 + 8x_3 + x_4 = 0 \\ x_1 + 3x_2 - 9x_3 + 7x_4 = 0; \end{cases}$$

$$(2) \begin{cases} 3x_1 + x_2 - 8x_3 + 2x_4 + x_5 = 0 \\ 2x_1 - 2x_2 - 3x_3 - 7x_4 + 2x_5 = 0 \\ x_1 + 11x_2 - 12x_3 + 3x_4 - 5x_5 = 0 \\ x_1 - 5x_2 + 2x_3 - 16x_4 + 3x_5 = 0. \end{cases}$$

2. 设

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

试求数 λ , 使齐次线性方程组 $AX = \lambda X$ 有非零解, 并求其基础解系.

3. 求一个齐次线性方程组 $AX = 0$, 使其解空间解 $N(A) = L(X_1, X_2)$, 其中:

$$X_1 = (1 \ 2 \ -1 \ 4)^T, \quad X_2 = (1 \ 0 \ -4 \ 0)^T.$$

4. 设 $A \in M_{m \times n}(F)$, 证明: 若 F^n 中任一个向量都是 $AX = 0$ 的解, 则 $A = O$. 并用线性映射的语言来叙述这个命题.

5. 设 $A \in M_{m \times n}(F)$, 且 $\text{r}(A) = r$, 证明: 齐次线性方程组 $AX = 0$ 的任何 $n - r$ 个线性无关的解都是它的基础解系.

6. 设 A, B 分别是 $m \times n$ 和 $n \times s$ 矩阵, 证明:

若方程组 $(AB)X = 0$ 与 $BX = 0$ 同解, 则 $\text{r}(AB) = \text{r}(B)$.

7*. 设 A^* 是 n 阶矩阵 A 的伴随矩阵, 证明:

$$(1) \text{r}(A^*) = \begin{cases} n, & \text{当 } \text{r}(A) = n \\ 1, & \text{当 } \text{r}(A) = n - 1 \\ 0, & \text{当 } \text{r}(A) < n - 1; \end{cases} \quad (2) |A^*| = |A|^{n-1}.$$

8*. 设 $A \in M_n(F)$, 证明: 若 $|A| = 0$, 则 A 中任意两行(或列)对应元素的代数余子式成比例.

9*. 证明: 若 A 是 n 阶幂等矩阵(即 $A^2 = A$), 则 $\text{r}(A) + \text{r}(E - A) = n$.

10*. 证明: 若 $A \in M_n(F)$, 且 $A^2 = E$, 则 $\text{r}(A + E) + \text{r}(A - E) = n$.

11. 求下列非齐次线性方程组 $AX = b$ 的一般解:

$$(1) A = \begin{pmatrix} 2 & 7 & 3 & 1 \\ 3 & 5 & 2 & 2 \\ 9 & 4 & 1 & 7 \end{pmatrix}, \quad (2) A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 1 & -3 \\ 0 & 1 & 2 & 2 & 6 \\ 5 & 4 & 3 & 3 & -1 \end{pmatrix},$$

$$b = (6, 4, 2)^T;$$

$$b = (7, -2, 23, 12)^T.$$

12. 求参数 t, p 的值, 使方程组有唯一解, 有无穷多数, 无解, 有解时, 求其解:

$$(1) \begin{cases} tx_1 + x_2 + x_3 = 1 \\ x_1 + tx_2 + x_3 = t \\ x_1 + x_2 + tx_3 = t^2; \end{cases}$$

$$(2) AX = b, A \text{ 如 11 题 (2)}, b = (1, t, 3, p)^T.$$

13*. 设 $A \in M_{m \times n}(F)$, 证明: 非齐次线性方程组 $AX = b$ 对任何 $b \in F^n$ 都有解的充分必要条件是 $r(A) = m$.

14. 设

$$\begin{cases} x_1 - x_2 = a_1 \\ x_2 - x_3 = a_2 \\ x_3 - x_4 = a_3 \\ x_4 - x_5 = a_4 \\ x_5 - x_1 = a_5. \end{cases}$$

证明: 该方程组有解的充分必要条件是 $a_1 + a_2 + a_3 + a_4 + a_5 = 0$, 有解时, 求其一般解.

15. 设 A 是 $m \times n$ 矩阵, $m = n + 1$, 证明:

(1) 线性方程组 $AX = b$ 有解的必要条件是 $\det(A, b) = 0$;

(2) 如果 $r(A) = n$ (1) 中条件也是充分的.

16. 已给平面 $\pi_1: x - 2y + 2z + d = 0$; $\pi_2: -2x + 4y + cz + 1 = 0$.

(1) 求 c, d , 使 $\pi_1 \parallel \pi_2$, 并问答案是否唯一?

(2) 求 c, d , 使 π_1 与 π_2 重合;

(3) 求 d , 使原点到 π_1 的距离为 1;

(4) 求 d , 使平面 π_1 到点 $M(1, 1, 1)$ 的距离为 1;

(5) 求 c, d , 使 $\pi_1 \parallel \pi_2$, 且它们之间的距离为 1.

17. 已给平面:

$$\pi_1: x - 2y + 2z + 1 = 0; \quad \pi_2: 2x + 3y - 6z - 6 = 0.$$

- (1) 求平面 π_1 与 π_2 之间的夹角;
 (2) 在 x 轴上确定一点,使它到平面 π_1 与 π_2 的距离相等.

18. 判断下列两条直线

$$L_1: \begin{cases} x = 2t \\ y = -3 + 3t \\ z = 4t \end{cases}, \quad L_2: \frac{x-1}{1} = \frac{y+2}{1} = \frac{z-2}{2}$$

是否共面?是否相交?如果相交求其交点.

19. 求点 $A(2, 4, 3)$ 在直线 $x = y = z$ 上的投影点的坐标及点 A 到该直线的垂直距离.

20. 设有两条直线

$$L_1: \frac{x-1}{2} = \frac{y+1}{-2} = \frac{z}{n}; \quad L_2: \begin{cases} x = -2 - 4t \\ y = 2 + mt \\ z = 3 + 2t \end{cases}$$

- (1) 求 m, n 使 $L_1 \parallel L_2$;
 (2) 当 $m = n = 1$ 时,求 L_1, L_2 之间的最短距离;
 (3) 当 $m = n = 1$ 时,求 L_1 与 L_2 的公垂线 L 的方程 (L 与 L_1, L_2 垂直且相交);
 (4) 求 m, n 使 $L_1 \perp L_2$ 并问 m, n 是否唯一?
 (5) 求 m, n 使 L_1 与 L_2 共面,这样的 m, n 是否唯一?
 (6) 当 $m = -4, n = -1$ 时,求 L_1 与 L_2 的夹角.

21. 已知:

$$\text{平面 } \pi: x - 2y - 2z + 4 = 0, \quad \text{直线 } L: \frac{x-1}{-1} = \frac{y}{2} = \frac{z+2}{n}.$$

- (1) 求 n 使 $L \perp \pi$; (2) 求 n 使 $L \parallel \pi$;
 (3) 当 $n = -2$ 时,求 L 与 π 之间的夹角;
 (4) 当 $n = -2$ 时,求 L 与 π 的交点,并求 L 在 π 上的投影线方程;
 (5) 求 L 在各坐标平面上的投影线方程;
 (6) 当 $n = -2$ 时,求直线 L_1 使 L_1 与 L 关于平面 π 对称;
 (7) 求原点关于平面 π 的对称点的坐标.

22. 设有两个相交的平面:

$$\pi_1: a_1x + b_1y + c_1z + d_1 = 0, \quad \pi_2: a_2x + b_2y + c_2z + d_2 = 0.$$

问当 λ, μ 是不全为零的任意常数时,方程

$$\lambda(a_1x + b_1y + c_1z + d_1) + \mu(a_2x + b_2y + c_2z + d_2) = 0$$

的图形有何共同特点?

23. 已知两个平面:

$$\pi_1: x - y - 2z = 2, \quad \pi_2: x + 2y + z = 8.$$

(1) 求过 π_1 与 π_2 的交线, 且与 $\pi_3: x + y + z = 0$ 垂直的平面的方程;

(2) 求 π_1 与 π_2 的平分面的方程.

补充题

1. 设 $A = (a_{ij})_{m \times n}$, $m < n$, $r(A) = m$; $B = (b_{ij})_{n \times (n-m)}$, $m < n$, $r(B) = n - m$. 已知齐次线性方程组 $AX = 0$ 的解空间 $N(A) = R(B)$ (矩阵 B 的列空间), 试求齐次线性方程组

$$\sum_{i=1}^n b_{ij} y_i = 0, \quad j = 1, 2, \dots, m - m$$

的一个基础解系.

2. 设 $A \in M_{m \times n}(F)$, $r(A) = m$; B 是 m 阶非奇异矩阵 (即可逆阵). 已知 A 的行空间 $R(A^T)$ 是方程组 $CX = 0$ 的解空间. 证明: BA 的行向量组也是 $CX = 0$ 的一个基础解系.

3. 设 A 是 n 阶矩阵 ($n > 2$), 证明 $(A^*)^* = |A|^{n-2} A$.

4. 设 $A = (a_{ij}) \in M_n(R)$, 证明: 如果 A 是对角绝对优势矩阵, 即

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, 2, \dots, n,$$

则 $\det A \neq 0$ (即 $r(A) = n$).

5. 证明: 欧氏空间 $V(R)$ 中向量组 $\{\alpha_1, \alpha_2, \dots, \alpha_r\}$ 线性相关的充分必要条件是 Gram 矩阵的行列式

$$\det \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \dots & (\alpha_1, \alpha_r) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \dots & (\alpha_2, \alpha_r) \\ \dots & \dots & \dots & \dots \\ (\alpha_r, \alpha_1) & (\alpha_r, \alpha_2) & \dots & (\alpha_r, \alpha_r) \end{pmatrix} = 0.$$

6. 设 $A = (a_{ij})_{m \times n}$, $b = (b_1, b_2, \dots, b_m)^T$, $Y = (y_1, y_2, \dots, y_n)^T$, $X = (x_1, x_2, \dots, x_m)^T$, 证明:

(1) 若 $AY = b$ 有解, 则 $A^T X = 0$ 的任一组解都满足方程 $b^T X = 0$;

(2) 方程组 $AY = b$ 有解的充要条件是, 方程组

$$\begin{pmatrix} A^T \\ b^T \end{pmatrix} X = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

无解(其中 $\mathbf{0}$ 是 $n \times 1$ 零矩阵).

7*. 相容的线性方程组 $AX = b$ 在怎样的条件下, 其解中第 k 个未知量 x_k 都是同一个值? 你给的条件是否是充分必要的?

8. 设 $A \in M_{m \times n}(R)$, $b = (b_1, b_2, \dots, b_m)^T \in R^m$, 证明: 正规方程

$$(A^T A)X = A^T b$$

有解; 当 $r(A) = n$ 时, 求其解, 并证明 $A(A^T A)^{-1}A^T$ 是幂等的对称矩阵.

9*. 设 $A \in M_{m \times n}(R)$ (其中 $m \gg n$), $AX = b$ 是不相容方程组(即无解). 证明:

$\exists X^* \in R^n$, 使 $\forall X \in R^n$ 有

$$\|b - AX^*\| \leq \|b - AX\|.$$

这个 X^* 称为不相容方程组 $AX = b$ 的最小二乘解.

部分习题和补充题答案

习题

1. (1) $k_1(-3, 7, 2, 0)^T + k_2(-1, -2, 0, 1)^T$;

(2) $k_1(19, 7, 8, 0, 0)^T + k_2(-1, 1, 0, 0, 2)^T + k_3(3, 25, 0, 8, 0)^T$.

2. $\lambda = -1, \{-1, 1, 1, 1\}^T$; $\lambda = 3, \{-1, 1, 0, 0\}^T, \{-1, 0, 1, 0\}^T, \{-1, 0, 0, 1\}^T$.

3. 提示: 利用 A 的行空间 $R(A^T)$ 是 $N(A)$ 的正交补, $8x_1 - 3x_2 + 2x_3 = 0, -2x_2 + x_4 = 0$.

7. 提示: 利用 $AA^* = A^*A = |A|E$ 和节 6.1 例 2 的结论以及 A^* 的定义.

8. 提示: 利用上题的结论.

9. 提示: 利用 $A(E - A) = O$ 和节 6.1 例 2 的结论以及 $r(A + B) \leq r(A) + r(B)$.

11. (1) $k_1(8, 0, 0, -10)^T + k_2(-4, 0, 1, 5)^T + k_3(-9, 1, 0, 11)^T$.

(2) $k_1(-16, 23, 0, 0, 0)^T + k_2(1, -2, 1, 0, 0)^T + k_3(1, -2, 0, 1, 0)^T + k_4(5, -6, 0, 0, 1)^T$.

12. (1) $t \neq 1, -2 \left(\frac{-(t+1)}{t+2}, \frac{1}{t+2}, \frac{(1+t)^2}{t+2} \right)^T$;

$t = 1, (1, 0, 0)^T + k_1(-1, 1, 0)^T + k_2(-1, 0, 1)^T$; $t = -2$, 无解;

(2) $t = 0, p = 2, k_1(-2, 3, 0, 0, 0)^T + k_2(1, -2, 1, 0, 0)^T + k_3(1, -2, 0, 1, 0)^T + k_4(5, -6, 0, 0, 1)^T$.

14. $a_1 + a_2 + a_3 + a_4 + a_5 = 0$;

$$(a_1 + a_2 + a_3 + a_4, a_2 + a_3 + a_4, a_3 + a_4, a_4, 0)^T + k(1, 1, 1, 1, 1)^T.$$

16. (1) $c = -4$, d 任意; (2) $-4, -\frac{1}{2}$; (3) ± 3 ;

(4) 2; (5) $-4, \frac{5}{2}$ 或 $-\frac{7}{2}$.

17. (1) $\pi - \arccos \frac{16}{21}$; (2) $(-25, 0, 0)$ 或 $(\frac{11}{13}, 0, 0)$.

18. 是, 是 $(0, -3, 0)$;

19. $(3, 3, 3); \sqrt{2}$. 20. (1) $4, -1$; (2) $\frac{27\sqrt{5}}{25}$;

(3) $20x + 7y - 26z - 13 = 0, 10x - 34y + 37z - 23 = 0$; (4) $m - n = 4$, 不唯一; (5) $mn + 2m - 4n - 8 = 0$, 不唯一; (6) $\pi - \arccos \frac{1}{9}$.

21. (1) 2; (2) $-\frac{5}{2}$; (3) $\arcsin \frac{1}{9}$; (4) $(-8, 18, -20)$; $\frac{x}{2} = \frac{y-2}{-4} = \frac{z}{5}$; (5) $\frac{x-1}{-1} = \frac{y}{2} = \frac{z}{0}$; $\frac{x-1}{-1} = \frac{y}{0} = \frac{z+2}{n}$; $\frac{x}{0} = \frac{y}{2} = \frac{z+2}{n}$; (6) $\frac{x+1}{7} = \frac{y-4}{-14} = \frac{z-2}{22}$; (7) $(-\frac{8}{9}, \frac{16}{9}, \frac{16}{9})$.

22. 过 π_1, π_2 交线的所有平面(即平面束).

23. (1) $x - z = 4$; (2) $2x + y - z = 10$ 或 $y + z = 2$.

补充题

1. (提示: 将方程组写成矩阵等式, 再利用 $AB = O$ 表示 B 的每个列向量都是方程组 $AX = 0$ 的解.) A 的行向量组. 3. 提示: 利用习题 7 的结论.

4. 提示: 用反证法, 即设 $AX = 0$ 有非零解. 7. A 的第 k 列不能用其余列线性表示. 8. 提示: 证明有解, 只要证增广矩阵与系数矩阵的秩相等, 此时要用到节 6.1 例 3 的结论.

9. (提示: 设 $W = R(A)$ 为 A 的列空间, $\alpha \in W$, 如果 $(b - \alpha) \perp W$, 则称 α 为 b 在 W 上的投影向量.) X^* 是正规方程 $(A^T A)X = A^T b$ 的解.

第7章 特征值与特征向量 矩阵的标准形

本章将进一步讨论 n 维线性空间的线性变换与 n 阶矩阵之间的对应关系. 我们已经讲过, 给定了线性空间的一组基, 线性变换与矩阵之间是一一对应的, 也就是说, 用矩阵 A 来表示线性变换 $\sigma \in L(V, V)$ 和 V 的基的选择是有关的. 那么, 对于一个线性变换 σ , 它有没有与基无关的不变量呢? 这就是本章要讨论的线性变换(矩阵)的特征值与特征向量的问题. 解决了这个问题, 我们就能选择线性空间 V 的适当的基, 使 $\sigma \in L(V, V)$ 在这个基下对应的矩阵最简单(称为 σ 对应的矩阵的标准形). 为了阐明这个问题, 我们将从二次曲线一般方程(含 xy 混合项)通过坐标变换化为标准方程入手.

本章主要内容有: 正交变换与正交矩阵; 二次曲线一般方程的化简及分类; 线性变换和矩阵的特征值与特征向量; 矩阵与对角阵相似的条件及相似标准形; 实对称矩阵的对角化; 实二次型的标准形; 矩阵的相合(或合同)标准形以及正定二次型与正定矩阵.

7.1 正交变换与正交矩阵

正交变换是欧氏空间中的一种重要的线性变换. 我们在第3章中讲过的旋转变换和镜象变换, 都使向量的长度及向量之间的夹角保持不变, 也就是任意两个向量之间的内积等于其象向量之间的内积. 具有这种性质的线性变换称为正交变换, 其定义如下:

定义 7.1 欧氏空间 $V(\mathbf{R})$ 的一个线性变换 σ 称为正交变换, 如果 $\forall \alpha, \beta \in V$, 都有

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta). \quad (7-1)$$

与(7-1)等价的条件是:

$$|\sigma(\alpha)| = |\alpha|, \quad \forall \alpha \in V. \quad (7-2)$$

这是因为: 在(7-1)式中, 取 $\beta = \alpha$, 即得

$$(\sigma(\alpha), \sigma(\alpha)) = (\alpha, \alpha), \quad \text{即} \quad |\sigma(\alpha)|^2 = |\alpha|^2,$$

故 $|\sigma(\alpha)| = |\alpha|$; 反之, 如果(7-2)式成立, 则任取 $\alpha, \beta \in V$, 都有

$$|\sigma(\alpha + \beta)| = |\alpha + \beta|.$$

而

$$\begin{aligned}
 |\sigma(\alpha) + \sigma(\beta)|^2 &= (\sigma(\alpha) + \sigma(\beta), \sigma(\alpha) + \sigma(\beta)) \\
 &= (\sigma(\alpha), \sigma(\alpha)) + (\sigma(\beta), \sigma(\beta)) + 2(\sigma(\alpha), \sigma(\beta)) \\
 &= |\sigma(\alpha)|^2 + |\sigma(\beta)|^2 + 2(\sigma(\alpha), \sigma(\beta)), \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 |\alpha + \beta|^2 &= (\alpha + \beta, \alpha + \beta) = (\alpha, \alpha) + (\beta, \beta) + 2(\alpha, \beta) \\
 &= |\alpha|^2 + |\beta|^2 + 2(\alpha, \beta). \quad (2)
 \end{aligned}$$

因此,由(1)(2)式,即得(7-1)式成立.

由(7-1)(7-2)式又可得,正交变换 σ 使任意两个向量 α, β 的夹角与其像 $\sigma(\alpha), \sigma(\beta)$ 的夹角是相等的,即

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta). \quad (3)$$

但是,这个命题的逆命题不成立.例如对于 $V(\mathbf{R})$ 的线性变换

$$\sigma(\alpha) = k\alpha \quad (k \in \mathbf{R}, k \neq 1, 0)$$

显然(3)式成立,而(7-2)式不成立,故这个 σ 不是正交变换.

定义 7.2 欧氏空间 $V(\mathbf{R})$ 的正交变换 σ 关于 V 的单位正交基所对应的矩阵 A 称为正交矩阵.

设 $A = (a_{ij})_{n \times n} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 是正交变换 σ 在 V 的单位正交基 $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 下所对应的矩阵,即

$$\sigma(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)A, \quad (1)$$

于是 $a_{ij} \in \mathbf{R}$,而且

$$\begin{aligned}
 (\varepsilon_i, \varepsilon_j) &= (\sigma(\varepsilon_i), \sigma(\varepsilon_j)) = \left(\sum_{k=1}^n a_{ki} \varepsilon_k, \sum_{l=1}^n a_{lj} \varepsilon_l \right) \\
 &= \sum_{k=1}^n a_{ki} a_{kj} = \alpha_j^T \alpha_i \\
 &= (\alpha_i, \alpha_j) = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases} \quad i, j = 1, 2, \dots, n, \quad (2)
 \end{aligned}$$

所以 A 的列向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是关于 \mathbf{R}^n 的标准内积的一组单位正交基.

反之,如果 A 的列向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 \mathbf{R}^n 的一组单位正交基,则 A 在 V 的单位正交基 $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 下所对应的线性变换 σ 也必是正交变换.因为,由(1)(2)式可见,此时 $\{\sigma(\varepsilon_1), \sigma(\varepsilon_2), \dots, \sigma(\varepsilon_n)\}$ 也是 V 的一组单位正交基.于是,对于 V 中任意的

$$\alpha = \sum_{i=1}^n a_i \varepsilon_i, \quad \beta = \sum_{j=1}^n b_j \varepsilon_j$$

都有

$$\begin{aligned} (\sigma(\alpha), \sigma(\beta)) &= \left(\sum_{i=1}^n a_i \sigma(\varepsilon_i), \sum_{j=1}^n b_j \sigma(\varepsilon_j) \right) = \sum_{i=1}^n a_i b_i \\ &= \left(\sum_{i=1}^n a_i \varepsilon_i, \sum_{j=1}^n b_j \varepsilon_j \right) = (\alpha, \beta), \end{aligned}$$

故 σ 是 $V(R)$ 的一个正交变换.

再由

$$A^T A = \begin{pmatrix} \alpha_1^T \\ \alpha_2^T \\ \vdots \\ \alpha_n^T \end{pmatrix} (\alpha_1, \alpha_2, \dots, \alpha_n) = \begin{pmatrix} \alpha_1^T \alpha_1 & \alpha_1^T \alpha_2 & \cdots & \alpha_1^T \alpha_n \\ \alpha_2^T \alpha_1 & \alpha_2^T \alpha_2 & \cdots & \alpha_2^T \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n^T \alpha_1 & \alpha_n^T \alpha_2 & \cdots & \alpha_n^T \alpha_n \end{pmatrix}$$

(其中 $\alpha_i^T \alpha_j = (\alpha_i, \alpha_j)$) 又可见 n 阶实矩阵 A 的列向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 \mathbf{R}^n 的一组单位正交基的充分必要条件是 $A^T A = E$.

综上所述, 正交矩阵也可等价地定义如下.

定义 7.3 n 阶实矩阵 A 称为正交矩阵, 如果 $A^T A = E$ (或: 如果 A 的列向量组是 \mathbf{R}^n 的一组单位正交基).

正交矩阵还有以下性质:

- (1) 若 A 为正交矩阵, 则 $A^{-1} = A^T$, 且 A^T 也是正交矩阵;
- (2) 若 A 是正交矩阵, 则 $|A| = 1$ 或 -1 ;
- (3) 若 A, B 都是正交矩阵, 则 AB 也是正交矩阵.

证 (1) 由 $A^T A = E$ 即得 $A^{-1} = A^T$. 又因为

$$(A^T)^T A^T = A A^T = A A^{-1} = E,$$

故 A^T (即 A^{-1}) 也是正交矩阵, 从而正交矩阵 A 的行向量组也是 \mathbf{R}^n 的一组单位正交基.

(2) 由 $A^T A = E$ 即得 $|A^T A| = |A^T| |A| = |A|^2 = 1$, 故 $|A| = \pm 1$.

(3) 由于 $(AB)^T (AB) = (B^T A^T) (AB) = B^T E B = B^T B = E$, 所以 AB 也是正交矩阵.

行列式为 $+1$ 的正交矩阵所对应的正交变换称为第一类正交变换, 行列式为 -1 的正交矩阵所对应的正交变换称为第二类正交变换. 对这个定义要作一点说明: 正交变换在不同的单位正交基下对应的矩阵一般是不同的, 但它的是相似矩阵 (见 7.3 节) 而相似矩阵的行列式的值是相等的.

例 1 \mathbf{R}^2 上的旋转变换 r_θ 是最基本的第一类正交变换. 它在单位正交基 $e_1 = (1 \ 0)$, $e_2 = (0 \ 1)$ 下对应的正交矩阵为

$$T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

(见第 4 章节 4-2 例 2), 它的行列式为 $+1$.

例 2 \mathbf{R}^3 关于镜面——与单位向量 $\eta_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T$ 垂直并过原点的平面(图 7-1)的镜象变换

$$\varphi(\alpha) = \alpha - 2(\eta_1, \alpha)\eta_1, \quad \forall \alpha \in \mathbf{R}^3$$

关于自然基 $\{e_1, e_2, e_3\}$ 也是 \mathbf{R}^3 的一组单位正交基)所对应的正交矩阵为

$$A = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

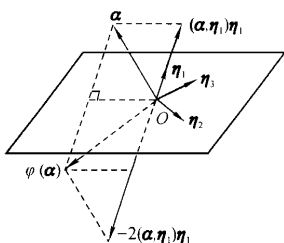


图 7-1

但如果在镜面上任取两个正交的单位向量 η_2, η_3 , 则

$$\varphi(\eta_2) = \eta_2, \quad \varphi(\eta_3) = \eta_3.$$

于是镜象变换 φ 关于 \mathbf{R}^3 的单位正交基 $\{\eta_1, \eta_2, \eta_3\}$ 所对应的对角阵 $B = \text{diag}(-1, 1, 1)$. 因为 $\varphi(\eta_1) = -\eta_1$, $\varphi(\eta_2) = \eta_2$, $\varphi(\eta_3) = \eta_3$, 所以

$$\varphi(\eta_1, \eta_2, \eta_3) = (\eta_1, \eta_2, \eta_3) \begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

这里 $|A| = |B| = -1$. 镜象变换是最基本的第二类正交变换.

下面,我们介绍用 Schmidt 正交化的方法对可逆实矩阵 A 作 $Q-R$ 分解.

定理 7.1 ($Q-R$ 分解) 若 A 为可逆实矩阵,则存在正交矩阵 Q 和主对角元为正数的上三角矩阵 R ,使得

$$A = QR.$$

证 n 阶可逆实矩阵 A 的列向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 \mathbb{R}^n 的一组基,用 Schmidt 正交化方法可由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 构造出 \mathbb{R}^n 的一组单位正交基.先正交化,得到一组两两正交的非零向量组 $\beta_1, \beta_2, \dots, \beta_n$,即:

$$\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = k_{12}\beta_1 + \alpha_2 \\ \dots\dots\dots \\ \beta_n = k_{1n}\beta_1 + k_{2n}\beta_2 + \dots + k_{n-1,n}\beta_{n-1} + \alpha_n, \end{cases} \quad (1)$$

其中

$$k_{ij} = -\frac{(\alpha_j, \beta_i)}{(\beta_i, \beta_i)} \quad (1 \leq i < j, \quad j = 2, \dots, n).$$

由(1)式可得

$$\begin{cases} \alpha_1 = \beta_1 \\ \alpha_2 = -k_{12}\beta_1 + \beta_2 \\ \dots\dots\dots \\ \alpha_n = -k_{1n}\beta_1 - k_{2n}\beta_2 - \dots - k_{n-1,n}\beta_{n-1} + \beta_n. \end{cases} \quad (2)$$

将(2)式写成一个矩阵等式

$$A = (\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_n) \begin{pmatrix} 1 & -k_{12} & \dots & -k_{1n} \\ 0 & 1 & \dots & -k_{2n} \\ \dots & & \ddots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \quad (3)$$

$$= (\beta_1, \beta_2, \dots, \beta_n) \begin{pmatrix} \frac{1}{|\beta_1|} & & & \\ & \frac{1}{|\beta_2|} & & \\ & & \ddots & \\ & & & \frac{1}{|\beta_n|} \end{pmatrix}$$

$$\begin{pmatrix} |\beta_1| & & & \\ & |\beta_2| & & \\ & & \ddots & \\ & & & |\beta_n| \end{pmatrix} \begin{pmatrix} 1 & -k_{12} & \dots & -k_{1n} \\ 0 & 1 & \dots & -k_{2n} \\ \dots & & \ddots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

$$= \left(\frac{\beta_1}{|\beta_1|}, \frac{\beta_2}{|\beta_2|}, \dots, \frac{\beta_n}{|\beta_n|} \right) \begin{pmatrix} |\beta_1| & -t_{12} & \dots & -t_{1n} \\ 0 & |\beta_2| & \dots & -t_{2n} \\ \dots & & \ddots & \dots \\ 0 & 0 & \dots & |\beta_n| \end{pmatrix} = QR.$$

上式的上三角矩阵 R 的主对角元 $|\beta_i| > 0 (i = 1, 2, \dots, n)$ $\left(\frac{\beta_1}{|\beta_1|}, \frac{\beta_2}{|\beta_2|}, \dots, \frac{\beta_n}{|\beta_n|} \right) = Q$ 是一个正交矩阵, 因为它的 n 个列向量是 \mathbf{R}^n 的一组标准正交基.

定理 7.2 (哈达马(Hadamard)不等式) n 阶实矩阵 A 的行列式的绝对值小于或等于 A 的 n 个列(行)向量长度的乘积, 即

$$|\det A| \leq \prod_{i=1}^n |\alpha_i|, \quad (7-3)$$

其中 α_i 为 A 的列(或行)向量. 当 $\det A \neq 0$ 时, 等号成立当且仅当 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为正交向量组.

证 当秩 $(A) < n$ 时, $\det A = 0$ (7-3) 式显然成立.

当秩 $(A) = n$ 时, A 的列向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 \mathbf{R}^n 的一组基. 用定理 7.1 的同样证法得其中的 (3) 式, 并将它记作

$$A = BR,$$

于是

$$\det A = (\det B)(\det R) = \det B. \quad (4)$$

由于 B 的列向量 $\beta_1, \beta_2, \dots, \beta_n$ 两两正交, 所以

$$\begin{pmatrix} \beta_1^T \\ \beta_2^T \\ \dots \\ \beta_n^T \end{pmatrix} (\beta_1, \beta_2, \dots, \beta_n) = \begin{pmatrix} |\beta_1|^2 & & & \\ & |\beta_2|^2 & & \\ & & \ddots & \\ & & & |\beta_n|^2 \end{pmatrix},$$

从而

$$(\det B)^2 = \det(B^T B) = \prod_{i=1}^n |\beta_i|^2. \quad (5)$$

根据勾股定理, 由定理 7.1 证明过程中的 (2) 式得

$$|\alpha_i|^2 \geq |\beta_i|^2, \quad i = 1, 2, \dots, n. \quad (6)$$

于是由 (4)(5)(6) 式, 即得哈达马不等式

$$|\det A| \leq \prod_{i=1}^n |\alpha_i|.$$

这里等号成立当且仅当 (6) 式中 等号成立, 即 (2) 式中 $k_{ij} = 0 (1 \leq i < j \leq n)$. 从而 $\alpha_i = \beta_i (i = 1, 2, \dots, n)$ 亦即 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为正交向量组.

若 A 的行向量组(即 A^T 的列向量组)为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 则

$$|\det A| = |\det A^T| \leq \prod_{i=1}^n |\alpha_i|.$$

哈达马不等式在三维空间中的几何意义: 三阶矩阵 A 的行列式等于 A 的三个行向量的混合积, 它的绝对值等于以三个行向量为邻边的平行六面体的体积, 它显然小于或等于三邻边边长的乘积, 等号成立当且仅当三邻边互相垂直.

7.2* 二次曲线一般方程化为标准方程及其分类

x_1, x_2 的一般二次方程

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_1x_1 + 2a_2x_2 + a_0 = 0 \quad (7-4)$$

(其中二次项系数不全为零), 在平面直角坐标系中一般是否都是圆锥截线(椭圆(包括圆) \ 双曲线和抛物线)呢? 如果我们通过坐标变换 $x_1 = g_1(y_1, y_2)$, $x_2 = g_2(y_1, y_2)$ 能将 (7-4) 式的二次项部分

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \quad (7-5)$$

中的混合项 $2a_{12}x_1x_2$ 消去(下面将证明这是可以办到的), 变为

$$\phi(y_1, y_2) = b_{11}y_1^2 + b_{22}y_2^2, \quad (7-6)$$

相应地 (7-4) 式变为

$$b_{11}y_1^2 + b_{22}y_2^2 + 2b_1y_1 + 2b_2y_2 + b_0 = 0, \quad (7-7)$$

再将 (7-7) 式配方(即作坐标平移)就可断言, 一般二次方程 (7-4) 所表示的图形都是圆锥截线(包括某些退化的情形, 如直线、点、虚椭圆和虚平行直线等), 因此, 它们也称为二次曲线.

7.2.1 作坐标轴旋转的坐标变换消去混合项, 再作坐标轴平移的坐标变换化为标准方程

欲消去 (7-5) 式中的混合项, 作坐标轴平移是不行的, 必须作坐标轴旋转的坐标变换.

设 (7-5) 式中 (x_1, x_2) 是点 P 在直角坐标系中的坐标(即向量 \vec{OP} 的坐标)是关于自

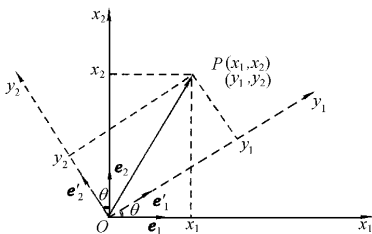


图 7-2

然基 $B_1 = \{e_1, e_2\} = \{(1, 0), (0, 1)\}$ 的坐标). 现将坐标轴绕原点 O 按逆时针方向旋转 θ 角, 相应地自然基变换为 $B_2 = \{e_1', e_2'\}$ (如图 7-2 所示), 于是

$$(e_1', e_2') = (e_1, e_2) \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (7-8)$$

其中右端的矩阵 Q (为正交矩阵) 是基 B_1 变为基 B_2 的变换矩阵. 于是根据定理 4.10 可知, 点 P 在基 B_2 下的坐标 (y_1, y_2) 满足

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Q^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (7-9)$$

所以作坐标轴旋转的坐标变换是一种正交变换.

将 (7-5) 式 $f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ 用矩阵表示为

$$f(x_1, x_2) = (x_1, x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{X}^T \mathbf{A} \mathbf{X}, \quad (7-10)$$

其中 $\mathbf{X} = (x_1, x_2)^T$, $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $a_{21} = a_{12}$.

将 (7-9) 式 $\mathbf{X} = \mathbf{Q}\mathbf{Y}$ (其中 $\mathbf{Y} = (y_1, y_2)^T$), 代入 (7-10) 式, 得

$$f(y_1, y_2) = (\mathbf{Q}\mathbf{Y})^T \mathbf{A} (\mathbf{Q}\mathbf{Y}) = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y},$$

其中

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (7-11)$$

记 $B = (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) = (b_{ij})_{2 \times 2}$, 则

$$\begin{cases} b_{11} = a_{11} \cos^2 \theta + a_{12} \sin 2\theta + a_{22} \sin^2 \theta \\ b_{12} = b_{21} = \frac{1}{2} (a_{22} - a_{11}) \sin 2\theta + a_{12} \cos 2\theta \\ b_{22} = a_{11} \sin^2 \theta - a_{12} \sin 2\theta + a_{22} \cos^2 \theta, \end{cases}$$

于是

$$f(y_1, y_2) = b_{11}y_1^2 + 2b_{12}y_1y_2 + b_{22}y_2^2. \quad (7-12)$$

为了消去混合项 y_1y_2 , 只须选择适当的 θ 角, 使 $b_{12} = 0$, 即由

$$\frac{1}{2} (a_{22} - a_{11}) \sin 2\theta + a_{12} \cos 2\theta = 0,$$

可得

$$\cot 2\theta = \frac{a_{11} - a_{22}}{2a_{12}}, \quad \theta = \frac{1}{2} \operatorname{arccot} \frac{a_{11} - a_{22}}{2a_{12}}. \quad (7-13)$$

此时, 由 (7-11) 式则得

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

比较上述等式中乘积矩阵的主对角元, 并依据 $a_{12} \neq 0$, 可知 $\cos \theta \neq 0$, 进而得到

$$b_{11} = a_{11} + a_{12}\tan\theta, \quad b_{22} = a_{22} - a_{12}\tan\theta. \quad (7-14)$$

例 1 把二次曲线的一般方程

$$5x_1^2 + 4x_1x_2 + 2x_2^2 - 24x_1 - 12x_2 + 18 = 0 \quad (1)$$

化为标准方程,并作其图形.

解 先作坐标轴旋转的坐标变换,消去其混合项.根据(7-13)式,转角 θ 满足

$$\cot 2\theta = \frac{a_{11} - a_{22}}{2a_{12}} = \frac{5-2}{4} = \frac{3}{4},$$

于是

$$2\tan^2\theta + 3\tan\theta - 2 = (2\tan\theta - 1)(\tan\theta + 2) = 0,$$

得 $\tan\theta = \frac{1}{2}$ 或 -2 . 取 $\tan\theta = \frac{1}{2}$ (且 θ 为锐角) 则

$$\sin\theta = \frac{1}{\sqrt{5}}, \quad \cos\theta = \frac{2}{\sqrt{5}}.$$

如此,方程(1)中的二次项部分 $5x_1^2 + 4x_1x_2 + 2x_2^2$ 化为 $b_{11}y_1^2 + b_{22}y_2^2$, 其中

$$b_{11} = a_{11} + a_{12}\tan\theta = 5 + 2 \times \left(\frac{1}{2}\right) = 6,$$

$$b_{22} = a_{22} - a_{12}\tan\theta = 2 - 2 \times \left(\frac{1}{2}\right) = 1.$$

再根据

$$X = QY \text{ 即 } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad (2)$$

方程(1)中的一次项与常数项 $-24x_1 - 12x_2 + 18$ 化为

$$\begin{aligned} (-24, -12) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 18 &= (-24, -12) \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + 18 \\ &= -12\sqrt{5}y_1 + 18. \end{aligned}$$

于是,原方程(1)化为

$$6y_1^2 + y_2^2 - 12\sqrt{5}y_1 + 18 = 0$$

即

$$6(y_1 - \sqrt{5})^2 + y_2^2 - 12 = 0. \quad (3)$$

再将坐标轴平移,令

$$z_1 = y_1 - \sqrt{5}, \quad z_2 = y_2, \quad (4)$$

则由方程(3)得到方程(1)所化成的标准方程

$$\frac{z_1^2}{2} + \frac{z_2^2}{12} = 1. \quad (5)$$

所以方程(1)的图形为椭圆,长轴在 z_2 轴上,长半轴为 $2\sqrt{3}$,短半轴为 $\sqrt{2}$ (如图7-3所示).这里通过坐标轴旋转和平移将(1)化为(5)的总的坐标变换公式由(2)(4)可得

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} z_1 + \sqrt{5} \\ z_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}}z_1 - \frac{1}{\sqrt{5}}z_2 + 2 \\ \frac{1}{\sqrt{5}}z_1 + \frac{2}{\sqrt{5}}z_2 + 1 \end{pmatrix}.$$

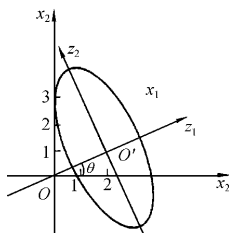


图 7-3

7.2.2 二次曲线的不变量

从上面的讨论和例题可见,任何一条二次曲线的一般二次方程(如(7-4)式)都可以通过坐标轴旋转和平移的坐标变换,将其化为标准方程.由此可见,同一条二次曲线在不同的直角坐标系中,其方程一般是不同的二次方程,也就是说,二次曲线的方程一般是随着坐标系的改变而改变的.这样,我们自然会提出一个问题:既然不同的方程代表着同一条曲线,那么这些不同的方程(表现在系数不同)中有没有某种共性,而这种共性是不随坐标系的改变而变化的.结论是肯定的,刻画这种共性的量称为二次曲线的不变量.

定理 7.3 二次曲线的方程(7-4)式,在直角坐标变换下,有以下三个不变量.

$$I_1 = a_{11} + a_{22}, \quad I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \quad I_3 = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{vmatrix}.$$

证 先证在坐标轴旋转的坐标变换下, I_1, I_2, I_3 不变.记 $\mathbf{X} = (x_1, x_2)^T, \mathbf{Y} = (y_1, y_2)^T$,转轴的坐标变换公式为

$$\mathbf{X} = \mathbf{QY} \quad (\mathbf{Q} \text{ 为正交矩阵}).$$

(7-4)式中的二次项部分为

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = \mathbf{X}^T \mathbf{A} \mathbf{X},$$

其中 $|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = I_2$.转轴变换后, $f(x_1, x_2)$ 变为

$$f(y_1, y_2) = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y}.$$

相应地, I_2 变为

$$I_2' = |Q^T A Q| = |Q|^2 |A| = |A| = I_2.$$

故 I_2 不变.

$I_1 = a_{11} + a_{22} = \text{tr}(A)$ 称为矩阵 A 的迹, 它是 A 的主对角元之和. 转轴变换后变为 $I_1' = \text{tr}(Q^T A Q)$, 记

$$Q = (q_{ij})_{2 \times 2}, \quad Q^T = (q'_{ji})_{2 \times 2} \quad q'_{ji} = q_{ij},$$

于是

$$\begin{aligned} I_1' &= \sum_{i=1}^2 (Q^T A Q)_{ii} = \sum_{i=1}^2 \sum_{j=1}^2 q'_{ij} (A Q)_{ji} \\ &= \sum_{i=1}^2 \sum_{j=1}^2 q_{ji} \left(\sum_{k=1}^2 a_{jk} q_{ki} \right) = \sum_{j=1}^2 \sum_{k=1}^2 a_{jk} \left(\sum_{i=1}^2 q_{ji} q_{ki} \right). \end{aligned}$$

由于 $Q = (q_{ij})_{2 \times 2}$ 是正交矩阵, 所以

$$\sum_{i=1}^2 q_{ji} q_{ki} = \delta_{jk} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases}$$

于是

$$I_1' = \sum_{j=1}^2 \sum_{k=1}^2 a_{jk} \delta_{jk} = \sum_{j=1}^2 (a_{j1} \delta_{j1} + a_{j2} \delta_{j2}) = a_{11} + a_{22} = I_1,$$

故 I_1 不变.

方程 (7-4) 的左端 (记为 $F(x_1, x_2)$) 可用矩阵表示为

$$\begin{aligned} F(x_1, x_2) &= (x_1 \ x_2 \ 1) \begin{pmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} \\ &= (X^T \ 1) \begin{pmatrix} A & \alpha \\ \alpha^T & a_0 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix}. \end{aligned} \quad (7-15)$$

其中,

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad P = \begin{pmatrix} A & \alpha \\ \alpha^T & a_0 \end{pmatrix}, \quad \alpha = (a_1 \ a_2)^T, \quad I_3 = |P|, \\ (X^T \ 1) &= (x_1 \ x_2 \ 1). \end{aligned}$$

转轴后 (7-15) 式变为

$$\begin{aligned} \varphi(y_1, y_2) &= (Y^T Q^T \ 1) \begin{pmatrix} A & \alpha \\ \alpha^T & a_0 \end{pmatrix} \begin{pmatrix} QY \\ 1 \end{pmatrix} \\ &= (Y^T \ 1) \begin{pmatrix} Q^T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & \alpha \\ \alpha^T & a_0 \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Y \\ 1 \end{pmatrix}, \end{aligned}$$

于是 I_3 变为

$$I'_3 = \begin{vmatrix} Q^T & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} A & \alpha \\ \alpha^T & a_0 \end{vmatrix} \begin{vmatrix} Q & 0 \\ 0 & 1 \end{vmatrix} \\ = |Q^T| |P| |Q| = |Q|^2 |P| = |P| = I_3,$$

故 I_3 不变.

再证, 在坐标轴平移的坐标变换下 I_1, I_2, I_3 不变.

将坐标轴平移的坐标变换公式

$$\begin{cases} x_1 = y_1 + c \\ x_2 = y_2 + d \end{cases}$$

代入原二次方程(7-4)式, 整理得

$$a_{11}y_1^2 + 2a_{12}y_1y_2 + a_{22}y_2^2 \\ + 2(a_{11}c + a_{12}d + a_1)y_1 + 2(a_{12}c + a_{22}d + a_2)y_2 \\ + a_{11}c^2 + 2a_{12}cd + a_{22}d^2 + 2a_1c + 2a_2d + a_0 = 0.$$

如果将此新方程记作

$$b_{11}y_1^2 + 2b_{12}y_1y_2 + b_{22}y_2^2 + 2b_1y_1 + 2b_2y_2 + b_0 = 0 \quad (7-16)$$

则

$$\begin{cases} b_{11} = a_{11}, & b_{12} = a_{12}, & b_{22} = a_{22} \\ b_1 = a_{11}c + a_{12}d + a_1 \\ b_2 = a_{12}c + a_{22}d + a_2 \\ b_0 = F(c, d) = a_{11}c^2 + 2a_{12}cd + a_{22}d^2 + 2a_1c + 2a_2d + a_0. \end{cases} \quad (7-17)$$

由(7-17)式, 立即可得

$$I'_1 = b_{11} + b_{22} = a_{11} + a_{22} = I_1,$$

$$I'_2 = \begin{vmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = I_2.$$

利用行列式的性质, 读者也不难得到

$$I'_3 = \begin{vmatrix} b_{11} & b_{12} & b_1 \\ b_{12} & b_{22} & b_2 \\ b_1 & b_2 & b_0 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_1 \\ a_{12} & a_{22} & a_2 \\ a_1 & a_2 & a_0 \end{vmatrix} = I_3.$$

综上所述, 在坐标轴旋转和平移的坐标变换下, 二次曲线方程中 I_1, I_2, I_3 是不变量.

7.2.3 利用不变量判定二次曲线的类型

二次曲线的一般二次方程(7-4)式, 经过直角坐标变换(坐标轴的旋转和平移), 都可化为二次曲线的最简方程(进而可化为标准方程), 利用最简方程中不变量满足的条件, 即可判定一般二次方程(7-4)所表示的二次曲线的类型.

1 椭圆型和双曲型曲线

这两种类型曲线的最简方程为

$$b_{11}y_1^2 + b_{22}y_2^2 + b_0 = 0. \quad (7-18)$$

当 b_{11} 与 b_{22} 同(异)号时,为椭圆(双曲)型.由于其不变量

$$I_1 = b_{11} + b_{22}, \quad I_2 = \begin{vmatrix} b_{11} & 0 \\ 0 & b_{22} \end{vmatrix}, \quad (7-19)$$

于是,曲线为椭圆型(b_{11}, b_{22} 同号)的充分必要条件是 $I_2 > 0$; 曲线为双曲型(b_{11} 与 b_{22} 异号)的充分必要条件是 $I_2 < 0$.

利用一元二次方程根与系数的关系,由(7-19)式易见 b_{11} 与 b_{22} 是一元二次方程

$$\lambda^2 - I_1\lambda + I_2 = 0 \quad (7-20)$$

的两个实根.将(7-20)式中的 I_1, I_2 用二次曲线一般方程(7-4)式中的系数表示,则有

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}^2) = 0,$$

即

$$(\lambda^2 - a_{11})(\lambda - a_{22}) - a_{12}^2 = 0.$$

此式可以用二阶行列式表示为

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{12} & \lambda - a_{22} \end{vmatrix} = 0 \quad (7-21)$$

(其中 E 为二阶单位矩阵).我们称(7-21)即(7-20)式是二次曲线(7-4)的特征方程,其中 $|\lambda E - A|$ 称为矩阵 A 的特征多项式,它的两个实根称为二次曲线(7-4)的特征根,记作 λ_1, λ_2 (即 $\lambda_1 = b_{11}, \lambda_2 = b_{22}$).再一般地说一下,由于方程(7-20)即(7-21)的判别式

$$\begin{aligned} I_1^2 - 4I_2 &= (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}^2) \\ &= (a_{11} - a_{22})^2 + 4a_{12}^2 \geq 0, \end{aligned}$$

所以此特征方程必有两个不等或相等的实根.

对于不变量 I_3 ,有

$$I_3 = \begin{vmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_0 \end{vmatrix} = b_{11}b_{22}b_0 = I_2b_0,$$

于是

$$b_0 = \frac{I_3}{I_2}. \quad (7-22)$$

综上所述,椭圆型和双曲型曲线的最简方程可写为.

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \frac{I_3}{I_2} = 0. \quad (7-23)$$

由(7-23)可以进而化成标准方程,从而确定椭圆的长、短半轴,双曲线的实、虚半轴.

由(7-23)根据不变量判别二次曲线(7-4)所属类型的方法如下:

(1)当 $I_2 = b_{11}b_{22} = \lambda_1\lambda_2 > 0$ 时(曲线为椭圆型),若 I_3 与 $I_1 = b_{11} + b_{22} = \lambda_1 + \lambda_2$

异号(即 $\frac{I_3}{I_2}$ 与 λ_1, λ_2 异号),则曲线是椭圆;若 I_3 与 I_1 同号,则曲线是虚椭圆;若 $I_3 = 0$,则

曲线是一个点.

(2) 当 $I_2 = b_{11}b_{22} = \lambda_1\lambda_2 < 0$ 时(曲线为双曲型) 若 $I_3 \neq 0$ 则曲线是双曲线 若 $I_3 = 0$ 则曲线是一对相交直线.

2 抛物型曲线

抛物型曲线的最简方程为

$$b_{22}y_2^2 + 2b_1y_1 = 0, \quad (7-24)$$

其中 $b_{22} \neq 0, b_1 \neq 0$. 由于其不变量

$$I_1 = b_{22}, \quad I_2 = \begin{vmatrix} 0 & 0 \\ 0 & b_{22} \end{vmatrix} = 0,$$

$$I_3 = \begin{vmatrix} 0 & 0 & b_1 \\ 0 & b_{22} & 0 \\ b_1 & 0 & 0 \end{vmatrix} = -b_{22}b_1^2 = -I_1b_1^2.$$

于是, 其最简方程(7-24)可写为

$$I_1y_2^2 \pm 2\sqrt{\frac{-I_3}{I_1}}y_1 = 0. \quad (7-25)$$

因此, 若 $I_2 = 0, I_3 \neq 0$ (此时 $\frac{a_{11}}{a_{12}} = \frac{a_{12}}{a_{22}} \neq \frac{a_1}{a_2}$) 则曲线为抛物线, 这是无心二次曲线.

若 $I_2 = 0$ 且 $I_3 = 0$ (此时 $\frac{a_{11}}{a_{12}} = \frac{a_{12}}{a_{22}} = \frac{a_1}{a_2}$) 则一般二次方程(7-4)可化为(证明略)

$$I_1y_2^2 + b_0 = 0.$$

当 b_0 与 I_1 异号时, 为一对平行直线; 当 b_0 与 I_1 同号时, 为一对虚平行直线; 当 $b_0 = 0$ 时, 为一对重合直线.

用不变量将二次曲线分类的情况可列表如下:

不变量			曲线形状	化简后的方程
$I_2 \neq 0$ 有心 二次 曲线	$I_2 > 0$ 椭圆型	$I_3 \neq 0$	$I_3I_1 < 0$ 为椭圆 $I_3I_1 > 0$ 为虚椭圆	$\lambda_1y_1^2 + \lambda_2y_2^2 + \frac{I_3}{I_2} = 0$
		$I_3 = 0$	一个点	
	$I_2 < 0$ 双曲线型	$I_3 \neq 0$	双曲线	
		$I_3 = 0$	一对相交直线	
$I_2 = 0$ 无心 二次 曲线	抛物型	$I_3 \neq 0$	抛物线	$I_1y_2^2 \pm 2\sqrt{\frac{-I_3}{I_1}}y_1 = 0$
		$I_3 = 0$	$K_1 < 0$, 一对平行直线 $K_1 = 0$, 一对重合直线 $K_1 > 0$, 一对虚平行线	$I_1y_2^2 + \frac{K_1}{I_1} = 0$

表中 $K_1 = \begin{vmatrix} a_{11} & a_1 \\ a_1 & a_0 \end{vmatrix} + \begin{vmatrix} a_{22} & a_2 \\ a_2 & a_0 \end{vmatrix}$ 称为二次曲线的半不变量, 它在坐标轴旋转的坐

标变换下不变, 对于 $I_2 = I_3 = 0$ 的二次曲线, 在坐标轴平移的坐标变换下也不变(证明略).

例 2 判断下列二次曲线的类型, 把方程化为最简形式, 并确定曲线的形状

$$(1) x_1^2 + 4x_1x_2 + 4x_2^2 + 12x_1 - x_2 + 1 = 0;$$

$$(2) 4x_1^2 - 4x_1x_2 + x_2^2 + 2x_1 - x_2 - 2 = 0;$$

$$(3) x_1^2 - 3x_1x_2 + x_2^2 + 10x_1 - 10x_2 + 21 = 0.$$

解 (1) $I_2 = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$ 这是抛物型曲线. 此时

$$I_1 = 1 + 4 = 5, \quad I_3 = \begin{vmatrix} 1 & 2 & 6 \\ 2 & 4 & -\frac{1}{2} \\ 6 & -\frac{1}{2} & 1 \end{vmatrix} = -\frac{625}{4},$$

所以这是一条抛物线, 化简后的方程为

$$5y_2^2 + 2y_1\sqrt{\frac{625}{4}} = 0$$

或

$$5y_2^2 - 2y_1\sqrt{\frac{625}{4}} = 0,$$

即

$$y_2^2 + \sqrt{5}y_1 = 0 \quad \text{或} \quad y_2^2 - \sqrt{5}y_1 = 0.$$

因此这条抛物线的焦参数 $p = \frac{\sqrt{5}}{2}$.

(2) $I_2 = \begin{vmatrix} 4 & -2 \\ -2 & 1 \end{vmatrix} = 0$ 这是抛物型曲线. 此时

$$I_3 = \begin{vmatrix} 4 & -2 & 1 \\ -2 & 1 & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -2 \end{vmatrix} = 0, \quad I_1 = 4 + 1 = 5,$$

$$K_1 = \begin{vmatrix} 4 & 1 \\ 1 & -2 \end{vmatrix} + \begin{vmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & -2 \end{vmatrix} = -\frac{45}{4} < 0,$$

所以这是两条平行直线, 化简后的方程为

$$5y_2^2 - \frac{9}{4} = 0, \quad \text{即} \quad y_2 = \pm \frac{3\sqrt{5}}{10}.$$

(3) $I_2 = \begin{vmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & 1 \end{vmatrix} = -\frac{5}{4} < 0$ 这是双曲型曲线. 此时

$$I_3 = \begin{vmatrix} 1 & -\frac{3}{2} & 5 \\ -\frac{3}{2} & 1 & -5 \\ 5 & -5 & 21 \end{vmatrix} = -\frac{5}{4}, \quad I_1 = 1 + 1 = 2.$$

因为 $I_3 \neq 0$, 所以这是双曲线. 解特征方程

$$\begin{aligned} |\lambda E - A| &= \begin{vmatrix} \lambda - 1 & \frac{3}{2} \\ \frac{3}{2} & \lambda - 1 \end{vmatrix} \\ &= (\lambda - 1)^2 - \frac{9}{4} = \left(\lambda + \frac{1}{2}\right)\left(\lambda - \frac{5}{2}\right) = 0, \end{aligned}$$

得特征根 $\lambda_1 = -\frac{1}{2}$, $\lambda_2 = \frac{5}{2}$.

又 $\frac{I_3}{I_2} = 1$, 于是化简后的方程为

$$-\frac{1}{2}y_1^2 + \frac{5}{2}y_2^2 + 1 = 0, \quad \text{即} \quad \frac{y_1^2}{2} - \frac{y_2^2}{5} = 1,$$

所以这条双曲线的实半轴 $a = \sqrt{2}$, 虚半轴 $b = \sqrt{\frac{2}{5}} = \frac{\sqrt{10}}{5}$.

7.3 线性变换在不同基下的矩阵表示 相似矩阵

在前一节中讲过, 下面的二次曲线方程

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 1$$

可以用矩阵表示为

$$f(x_1, x_2) = \mathbf{X}^T \mathbf{A} \mathbf{X} = 1, \quad (7-26)$$

其中

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

这就是说, 一个二次曲线在给定的直角坐标系(或说 \mathbf{R}^2 的一组单位正交基)下, 对应于一个实对称矩阵 \mathbf{A} . 如果我们把对方程(7-26)所作坐标轴旋转的坐标变换

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

记作 $\mathbf{X} = \mathbf{Q}\mathbf{Y}$ (\mathbf{Q} 为正交矩阵) 则二次曲线的方程(7-26), 在新的直角坐标系(或说 \mathbf{R}^2 的新的单位正交基)下, 变为

$$\varphi(y_1, y_2) = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y} = 1,$$

即二次曲线在新的直角坐标系下对应于实对称矩阵 $(Q^T A Q)$. 因此, 对坐标系作不同的旋转变换, 将使二次曲线对应于不同的实对称矩阵. 容易证明: 这些实对称矩阵

$$A_L = \{Q^T A Q \text{ 或 } Q^{-1} A Q \mid Q \text{ 为正交矩阵} \}$$

构成一个等价类, 它的最简单的代表元记为

$$A_0 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

其中 λ_1, λ_2 是 A 的特征方程 $|\lambda E - A|$ 的根 (称为二次曲线的特征根), 它们之和是二次曲线的不变量 I_1 , 即 $I_1 = \lambda_1 + \lambda_2 = a_{11} + a_{22}$. 如果 (7-26) 是椭圆方程, 它将化为标准方程

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = 1, \quad \text{即} \quad \frac{y_1^2}{\left(\sqrt{\frac{1}{\lambda_1}}\right)^2} + \frac{y_2^2}{\left(\sqrt{\frac{1}{\lambda_2}}\right)^2} = 1.$$

此时, 特征根的几何意义是: 它们的倒数的平方根是椭圆的长、短半轴, 这显然是二次曲线不随坐标系改变的不变量.

线性变换 $\sigma \in L(V, V)$ 也有类似的问题. 给定了线性空间 V 的一组基, σ 对应于一个确定的矩阵 A , 对于 V 的不同的基, σ 一般也将对应于不同的矩阵. 既然这些不同的矩阵对应于同一个线性变换 σ , 那么它们也应该是一个等价类. 在这个等价类中也应该有不变的量, 我们也希望找到这个等价类中最简单的代表元. 这就是本节和以后几节要讨论的主要问题.

定理 7.4 设线性变换 $\sigma \in L(V, V)$, $B_1 = \{\alpha_1, \dots, \alpha_n\}$ 和 $B_2 = \{\beta_1, \dots, \beta_n\}$ 是线性空间 $V(F)$ 的两组基. 基 B_1 变为基 B_2 的变换矩阵为 C . 如果 σ 在基 B_1 下的矩阵为 A , 则 σ 关于基 B_2 所对应的矩阵为 $C^{-1}AC$.

证 由假设

$$(\beta_1 \dots \beta_n) = (\alpha_1 \dots \alpha_n) C, \quad (1)$$

得

$$(\alpha_1 \dots \alpha_n) = (\beta_1 \dots \beta_n) C^{-1}. \quad (2)$$

将 (2) 式代入已知条件

$$\sigma(\alpha_1 \dots \alpha_n) = (\alpha_1 \dots \alpha_n) A,$$

得

$$\sigma((\beta_1 \dots \beta_n) C^{-1}) = ((\beta_1 \dots \beta_n) C^{-1}) A. \quad (3)$$

容易验证 (3) 式左端等于 $(\sigma(\beta_1 \dots \beta_n)) C^{-1}$, 所以

$$(\sigma(\beta_1 \dots \beta_n)) C^{-1} = (\beta_1 \dots \beta_n) (C^{-1} A),$$

从而得

$$\sigma(\beta_1 \dots \beta_n) = (\beta_1 \dots \beta_n) C^{-1} A C,$$

故 σ 关于基 B_2 所对应的矩阵为 $C^{-1} A C$.

反之, 如果 σ 在基 $B_1 = \{\alpha_1 \dots \alpha_n\}$ 下所对应的矩阵为 A , 则 $C^{-1} A C$ 必是 σ 关于另一个基 $\{\beta_1 \dots \beta_n\}$ 所对应的矩阵, 且基 $\{\beta_1 \dots \beta_n\}$ 是由基 B_1 经变换矩阵 C 而得到的, 即

$$(\beta_1 \dots \beta_n) = (\alpha_1 \dots \alpha_n) C.$$

定义 7.4 如果对于 $A, B \in M_n(F)$ 存在可逆矩阵 $C \in M_n(F)$ 使得

$$C^{-1} A C = B, \quad (7-27)$$

则称 A 相似于 B , 记作 $A \sim B$.

容易证明, 矩阵的相似关系是集合 $M_n(F)$ 上的一种等价关系, 即相似关系具有:

(1) 自反性, 即 $\forall A \in M_n(F), A \sim A$;

(2) 对称性, 即 $\forall A, B \in M_n(F)$ 若 $A \sim B$ 则 $B \sim A$;

(3) 传递性, 即 $\forall A_1, A_2, A_3 \in M_n(F)$, 若 $A_1 \sim A_2, A_2 \sim A_3$, 则 $A_1 \sim A_3$.

因此, 彼此相似的矩阵构成一个等价类, 等价的相似矩阵都表示同一个线性变换, 只不过取的基不同而已.

相似矩阵还有以下性质:

(1) $C^{-1}(kA + tB)C = kC^{-1}AC + tC^{-1}BC \quad (k, t \in F)$;

(2) $C^{-1}(AB)C = (C^{-1}AC)(C^{-1}BC)$;

(3) 若 $A \sim B$ 则 $A^m \sim B^m$ (m 为正整数);

(4) 若 $A \sim B$ 则 $f(A) \sim f(B)$,

其中 $f(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0$,

$$f(A) = a_m A^m + a_{m-1} A^{m-1} + \dots + a_1 A + a_0 E,$$

$$a_i \in F, i = 0, 1, \dots, m.$$

例 1 已知 \mathbb{R}^2 的线性变换 σ 关于自然基 e_1, e_2 所对应的矩阵为

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix},$$

求 σ 关于基 $\{\beta_1, \beta_2\}$ 所对应的矩阵 B , 其中 $\beta_1 = e_1 - e_2, \beta_2 = 3e_1 + 4e_2$.

解 由已知条件

$$\sigma(e_1, e_2) = (e_1, e_2) \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix},$$

及

$$(\beta_1, \beta_2) = (e_1, e_2) \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix},$$

即得

$$\begin{aligned} B &= \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix} \\ &= \frac{1}{7} \begin{pmatrix} 4 & -3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 6 \end{pmatrix}. \end{aligned}$$

有限维线性空间 V 上的线性变换关于不同基所对应的矩阵构成一个相似等价类, 这个等价类中的不变量(也就是线性变换与基无关的不变量)是什么呢? 它的最简单的代表元又是什么呢? 在例 1 中已经见到 \mathbf{R}^2 的线性变换

$$\sigma(x_1, x_2) = (2x_1 + 3x_2, 4x_1 + 3x_2)$$

关于自然基 $\{e_1, e_2\}$ 对应的矩阵为 A , 而关于基 $\{\beta_1, \beta_2\}$ 所对应的矩阵 B 为对角阵. 这是与矩阵 A 相似的最简单的矩阵. 在 7-1 的例 2 中, 对于 \mathbf{R}^3 中的镜像变换(也存在 \mathbf{R}^3 的单位正交基使 σ 对应的矩阵为对角阵 $\text{diag}(-1, 1, 1)$). 然而并不是所有的 n 阶矩阵都能与对角阵相似. 这一章我们将着重讨论线性空间 V 的线性变换 σ 的不变量以及 σ 在什么条件下对应于对角阵. 也就是 n 阶矩阵 A 在相似意义下的不变量以及 A 在什么条件下与对角阵相似. 此外还要简单地介绍一下, 任何矩阵 A 在复数域上都相似于一类较简单的对角块矩阵——约当形矩阵(对角阵是约当形矩阵的特殊情形).

7.4 特征值与特征向量

在 7-1 的例 2 中我们看到镜像变换 φ 对镜面的法向量 η_1 及镜面上的任一向量 η_2 的像 $\varphi(\eta_1), \varphi(\eta_2)$ 与原像 η_1, η_2 成比例, 即 $\varphi(\eta_1) = -\eta_1$, $\varphi(\eta_2) = \eta_2$. 这个比例系数 $-1, 1$ 与原像 η_1, η_2 称为镜像变换 φ 的特征值与特征向量. 显然 φ 的特征值与 \mathbf{R}^3 的基的选择无关. 下面对线性变换的特征值和特征向量给以一般的定义:

定义 7.5 设 σ 是线性空间 $V(F)$ 的一个线性变换, 如果存在数 $\lambda_0 \in F$ 和非零向量 $\xi \in V$, 使得

$$\sigma(\xi) = \lambda_0 \xi, \quad (7-28)$$

则称数 λ_0 为 σ 的一个特征值, 称非零向量 ξ 为 σ 的属于其特征值 λ_0 的特征向量.

这里需要注意, 特征值 λ_0 是域 F 中的数量, 特征向量 ξ 是非零向量. 显然, 零向量对任意的 λ_0 都满足 (7-28) 式, 因此它不具有“特征”意义.

在几何向量空间 \mathbf{R}^2 和 \mathbf{R}^3 中, 线性变换 σ 的特征值与特征向量的几何意义是: 特征向量 ξ (起点在坐标原点) 与其在 σ 作用下的像 $\sigma(\xi)$ 同向(或反向),

同向时, 特征值 $\lambda_0 > 0$, 反向时, $\lambda_0 < 0$, 且 λ_0 的绝对值等于 $|\alpha(\xi)|$ 与 $|\xi|$ 之比值. 如果特征值 $\lambda_0 = 0$, 则 $\alpha(\xi)$ 为零向量.

例如: 在 \mathbf{R}^2 中, 向量绕原点按逆时针方向旋转 θ 角的旋转变换 r_θ , 当 $0 < \theta < \pi$ 时, 对任意非零向量 $\alpha \in \mathbf{R}^2$, $r_\theta(\alpha)$ 与 α 都不共线 (图 7-4), 此时 r_θ 没有实特征值; 当 $\theta = \pi$ 时, \mathbf{R}^2 中任何非零向量 ξ 都与 $r_\pi(\xi)$ 共线, 且 $r_\pi(\xi) = -\xi$ (图 7-5), 所以 -1 是 r_π 的特征值, 而且任何非零向量 ξ 都是其特征向量.

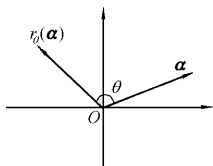


图 7-4

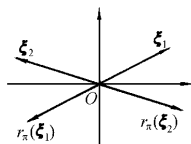


图 7-5

在 \mathbf{R}^3 中, 向量到二维子空间 W 上的投影变换 p : 当 $\xi \in W$ 时, $p(\xi) = \xi$; 当 $\xi \in W^\perp$ 时, $p(\xi) = \mathbf{0} = 0\xi$. 因此 1 与 0 是 p 的特征值, W 与 W^\perp 中的所有非零向量分别是属于特征值 1 和 0 的特征向量. 对于 \mathbf{R}^3 的基 $\{\xi_1, \xi_2, \xi_3\}$ 其中 $\xi_1, \xi_2 \in W, \xi_3 \in W^\perp$, 如图 7-6), p 所对应的矩阵是对角阵 $\text{diag}(1, 1, 0)$, 即

$$p(\xi_1, \xi_2, \xi_3) = (\xi_1, \xi_2, \xi_3) \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}.$$

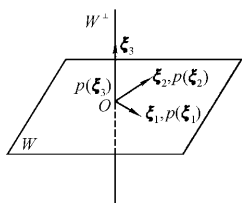


图 7-6

下面进一步讨论特征值与特征向量的一些性质及如何求特征值和特征向

量.

设 V 是域 F 上的线性空间, λ_0 是 $\sigma \in L(V, V)$ 的一个特征值, 根据定义 7.5, 使 (7-28) 式成立的非零向量 $\xi \in V$ 不止一个, 所有满足 (7-28) 式的向量组成的集合

$$V_{\lambda_0} = \{ \xi \mid \sigma(\xi) = \lambda_0 \xi, \xi \in V \} \quad (7-29)$$

是 V 的一个子空间. 这是因为 $\forall \xi_1, \xi_2 \in V_{\lambda_0}$ 和 $\forall k_1, k_2 \in F$ 均有

$$\sigma(k_1 \xi_1 + k_2 \xi_2) = k_1 \sigma(\xi_1) + k_2 \sigma(\xi_2) = \lambda_0(k_1 \xi_1 + k_2 \xi_2),$$

于是 $k_1 \xi_1 + k_2 \xi_2 \in V_{\lambda_0}$, 所以 V_{λ_0} 是 V 的一个子空间, 并称为 σ 关于其特征值 λ_0 的特征子空间. 在 V_{λ_0} 中的所有非零向量都是 σ 的特征向量.

由于 (7-28) 式等价于

$$(\lambda_0 I - \sigma)(\xi) = 0, \quad (7-30)$$

其中 I 是 V 上的恒等变换, 0 是 V 的零元. 所以特征子空间 V_{λ_0} 就是线性变换 $(\lambda_0 I - \sigma)$ 的核, 而该核中必有非零向量 ξ , 因此

$$\dim(\text{Ker}(\lambda_0 I - \sigma)) \geq 1.$$

如果 $\dim V = n$, 则由第 3 章定理 3.2 即得

$$\dim(\lambda_0 I - \sigma) \leq n - 1.$$

若 σ 关于基 $B = \{e_1, \dots, e_n\}$ 所对应的矩阵为 A , 则 $(\lambda_0 I - \sigma)$ 关于基 B 所对应的矩阵为 $(\lambda_0 E - A)$, 且其秩也小于等于 $n - 1$, 从而

$$|\lambda_0 E - A| = 0. \quad (7-31)$$

相应的齐次线性方程组

$$(\lambda_0 E - A)X = 0 \quad (7-32)$$

的非零解 $X = (x_1, \dots, x_n)^T$ 所对应的非零向量

$$\xi = x_1 e_1 + \dots + x_n e_n \in \text{Ker}(\lambda_0 I - \sigma). \quad (7-33)$$

所以 ξ 是 σ 的属于 λ_0 的特征向量, 而且 σ 的特征子空间 V_{λ_0} 中的 ξ 与 (7-32) 的解空间中的解向量 X 如 (7-33) 式那样一一对应.

从 (7-31) 式又可见, σ 的特征值 λ_0 是 λ 的 n 次代数方程

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & & & \dots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix} = 0 \quad (7-34)$$

在域 F 上的一个根. 事实上, 方程 (7-34) 在域 F 上的任一个根 λ_j 都是 σ 的特

征值,这是因为:与齐次线性方程组

$$(\lambda_j E - A)X = 0 \quad (7-35)$$

的非零解 X 按(7-33)式对应的非零向量 ξ 满足

$$\sigma(\xi) = \lambda_j \xi.$$

综上所述,对于线性空间 $V(F)$ 的一个线性变换 σ ,可以通过它在 V 的一个基 $B = \{e_1, \dots, e_n\}$ 下所对应的矩阵 A 按(7-34)(7-35)式求它的特征值及相应的特征向量关于基 B 的坐标.正因为如此,我们把方程(7-34)在域 F 上的根 λ_j 也称为矩阵 $A \in M_n(F)$ 的特征值,相应地把(7-35)式的非零解 X 称为矩阵 A 属于其特征值 λ_j 的特征向量.关于矩阵 A 的特征值与特征向量,也可以等价地定义如下.

定义 7.6 设矩阵 $A \in M_n(F)$, 如果存在数 $\lambda_0 \in F$ 和非零向量 $X \in F^n$, 使得

$$AX = \lambda_0 X, \quad (7-36)$$

则称数 λ_0 为 A 的一个特征值,称非零向量 X 为 A 的属于其特征值 λ_0 的特征向量.

域 F 上的 n 阶矩阵 A 的特征值是方程(7-34)在域 F 上的根,因此,通常把方程(7-34)叫做矩阵 A 的特征方程,把 λ 的 n 次多项式

$$f(\lambda) = |\lambda E - A| \quad (7-37)$$

叫做矩阵 A 的特征多项式.

如果域 F 是复数域 C , 则 n 阶矩阵 A 的特征多项式在复数域上的 n 个根都是矩阵 A 的特征值,其 k 重根叫做 k 重特征值.以后求矩阵 A 的特征值,如不加说明,都是求它在复数域上的特征值.

例 1 n 阶对角矩阵、上(下)三角形矩阵 A 的特征值是它们的 n 个主对角元 $a_{11}, a_{22}, \dots, a_{nn}$. 因为它们的特征多项式

$$|\lambda E - A| = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).$$

例 2 已知 R^3 的线性变换 σ 关于 R^3 的某一组基 $\{\alpha_1, \alpha_2, \alpha_3\}$ 所对应的矩阵为

$$A = \begin{pmatrix} 0 & -2 & -2 \\ 2 & -4 & -2 \\ -2 & 2 & 0 \end{pmatrix},$$

求 σ 的特征值及相应的特征子空间.

解 R^3 是实数域上的线性空间, σ 的特征值是矩阵 A 的特征多项式的实数根.

$$\begin{aligned}
 |\lambda E - A| &= \begin{vmatrix} \lambda & 2 & 2 \\ -2 & \lambda + 4 & 2 \\ 2 & -2 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 2 & 2 \\ 0 & \lambda + 2 & \lambda + 2 \\ 2 & -2 & \lambda \end{vmatrix} \\
 &= \begin{vmatrix} \lambda & 2 & 0 \\ 0 & \lambda + 2 & 0 \\ 2 & -2 & \lambda + 2 \end{vmatrix} = \lambda(\lambda + 2)^2,
 \end{aligned}$$

所以 A 的特征值为 $\lambda_1 = 0$ ($\lambda_2 = -2$ 二重特征值), 它们也都是 σ 的特征值.

对于 $\lambda_1 = 0$ 求解 $(\lambda_1 E - A)X = 0$, 即

$$\begin{pmatrix} 0 & 2 & 2 \\ -2 & 4 & 2 \\ 2 & -2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

得基础解系: $X_1 = (-1, -1, 1)^T$ (即 A 属于 λ_1 的一个特征向量). 于是

$$\xi_1 = -\alpha_1 - \alpha_2 + \alpha_3$$

是 σ 关于其特征值 $\lambda_1 = 0$ 的特征子空间 V_{λ_1} 的基.

对于 $\lambda_2 = -2$ 求解 $(\lambda_2 E - A)X = 0$, 即

$$\begin{pmatrix} -2 & 2 & 2 \\ -2 & 2 & 2 \\ 2 & -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

得基础解系: $X_2 = (1, 1, 0)^T$, $X_3 = (1, 0, 1)^T$ (它们是 A 关于 λ_2 的两个线性无关的特征向量), 于是

$$\xi_2 = \alpha_1 + \alpha_2, \quad \xi_3 = \alpha_1 + \alpha_3$$

是 σ 关于其特征值 $\lambda_2 = -2$ 的特征子空间 V_{λ_2} 的基.

关于矩阵和线性变换的特征值有以下几个基本结论.

定理 7.5 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 的特征多项式为

$$f(\lambda) = \lambda^n + b_1 \lambda^{n-1} + \dots + b_k \lambda^{n-k} + \dots + b_{n-1} \lambda + b_n, \quad (7-38)$$

其中系数 $b_k = (-1)^k S_k$, S_k 为 A 的全体 k 阶主子式之和, 即

$$S_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \dots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \dots & a_{i_2 i_k} \\ \dots & \dots & \dots & \dots \\ a_{i_k i_1} & a_{i_k i_2} & \dots & a_{i_k i_k} \end{vmatrix}.$$

定理 7.5 的证明留给有兴趣的读者自己去完成. (提示: 将

$f(x) = |\lambda E - A|$ 中, $i \neq j$ 的 $-a_{ij}$ 都表示为 $0 - a_{ij}$, 并将这个行列式表示为 2^n 个行列式之和, 再利用第 5 章节 5.2 例 4 的结论.)

推论 若 n 阶矩阵 $A = (a_{ij})_{n \times n}$ 的 n 个特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则

$$(1) \sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}, \quad (2) \prod_{i=1}^n \lambda_i = |A|. \quad (7-39)$$

证 (7-38) 式中的 b_1, b_n 为

$$b_1 = (-1)^1 S_1 = - \sum_{i=1}^n a_{ii}, \quad b_n = (-1)^n S_n = (-1)^n |A|,$$

所以

$$f(\lambda) = \lambda^n - \left(\sum_{i=1}^n a_{ii} \right) \lambda^{n-1} + \dots + (-1)^n |A|. \quad (1)$$

又已知

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n), \quad (2)$$

根据 (1)(2) 式中 λ^{n-1} 的系数及常数项相等, 即得要证的结论.

定理 7.6 若矩阵 A 与 B 相似, 则它们的特征多项式相等, 即

$$|\lambda E - A| = |\lambda E - B|.$$

证 $A \sim B$, 即存在可逆矩阵 P , 使得

$$P^{-1}AP = B,$$

于是

$$\begin{aligned} |\lambda E - B| &= |\lambda E - P^{-1}AP| = |P^{-1}(\lambda E - A)P| \\ &= |P^{-1}| |\lambda E - A| |P| = |\lambda E - A|. \end{aligned}$$

定理 7.6 的逆命题不成立. 例如:

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

$|\lambda E - A| = |\lambda E - B| = (\lambda - 2)^2$, 但 A 与 B 不相似, 因为对任何可逆矩阵 P ,

$$P^{-1}AP = P^{-1}(2E)P = 2E = A \neq B.$$

我们在 7.3 中讲过, 线性变换 σ 在不同基下对应的矩阵是相似矩阵, 而相似矩阵的特征多项式相同 (即特征值相同). 所以, 线性变换 σ 通过它在 V 的基下所对应的矩阵求其特征值, 不会由于基的不同而得到不同的结果, 也就是说 σ 的特征值与 V 的基是无关的, 所以特征值是线性变换 σ 的不变量. 正因为如此, σ 在任何基下所对应的矩阵的特征多项式 (它们都是相同的) 也称为 σ 的特征多项式.

例 3 设 $P^{-1}AP = B$, 证明: A, B 分别属于其特征值 λ_0 的特征向量 X 和 Y 满足 $Y = P^{-1}X$.

证 由 $AX = \lambda_0 X$ 及 $A = PBP^{-1}$ 得 $PBP^{-1}X = \lambda_0 X$, 从而就有

$$B(P^{-1}X) = \lambda_0(P^{-1}X),$$

故 $(P^{-1}X)$ 是 B 属于 λ_0 的特征向量, 即 $Y = (P^{-1}X)$.

由此例可见, 如果 A, B 是线性变换 σ 在基 B_1 和基 B_2 下对应的矩阵, 则例 3 中的 P 是基 B_1 变为 B_2 的变换矩阵, X 和 Y 是 σ 属于其特征值 λ_0 的特征向量 ξ 分别在基 B_1 和基 B_2 下的坐标.

例 4 设 λ_0 是线性空间 $V(F)$ 的线性变换 σ 的特征值, ξ 是 σ 属于 λ_0 的特征向量, 证明:

(1) $k\lambda_0, \lambda_0^m$ 分别是 $k\sigma$ 和 σ^m 的特征值, 且 ξ 仍是相应的特征向量 (其中 $k \in F, m \in \mathbb{N}^*$);

(2) 如果 $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$ 是 F 上的多项式, 并定义 σ 的多项式

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

(其中 I 是 $V(F)$ 的恒等变换) 则 $f(\sigma)(\xi) = f(\lambda_0)\xi$.

证 (1) 由假设 $\sigma(\xi) = \lambda_0 \xi (\xi \neq 0)$, 即得

$$(k\sigma)(\xi) = k(\sigma(\xi)) = (k\lambda_0)\xi,$$

$$\sigma^2(\xi) = \sigma(\sigma(\xi)) = \sigma(\lambda_0 \xi) = \lambda_0 \sigma(\xi) = \lambda_0^2 \xi.$$

用数学归纳法易证, $\forall m \in \mathbb{N}^*$, 有 $\sigma^m(\xi) = \lambda_0^m \xi$. 故 (1) 的结论成立.

(2) 利用 (1) 的结论, 立即可得. 如把本例中的 σ 换为矩阵 A , 结论也都成立.

由例 2 可见, 线性变换 σ 的两个特征子空间 V_{λ_1} 与 V_{λ_2} 的交为零子空间, 即 V_{λ_1} 的基 ξ_1 与 V_{λ_2} 的基 ξ_2, ξ_3 合在一起是线性无关的向量组. 也就是矩阵 A 属于 λ_1, λ_2 的三个特征向量 X_1, X_2, X_3 是线性无关的. 这个结论有普遍意义, 一般的结论如下面的定理所述.

定理 7.7 设 λ_j 和 $V_{\lambda_j} (j = 1, \dots, m)$ 是 n 维线性空间 $V(F)$ 的线性变换 σ 的 m 个互不相同的特征值及相应的特征子空间, 则 m 个特征子空间的和是直和, 即

$$\dim(V_{\lambda_1} + V_{\lambda_2} + \dots + V_{\lambda_m}) = \sum_{j=1}^m \dim V_{\lambda_j}. \quad (7-40)$$

证 欲证 $W = V_{\lambda_1} + \dots + V_{\lambda_m}$ 是直和, 只要证明 W 中零向量的分解式是唯一的, 即等式

$$\xi_1 + \xi_2 + \dots + \xi_m = 0 \quad (1)$$

(其中 $\xi_j \in V_{\lambda_j}$, 即 $\alpha(\xi_j) = \lambda_j \xi_j$ $j = 1, 2, \dots, m$) 成立, 仅当 $\xi_1 = \xi_2 = \dots = \xi_m = 0$.

对 m 作归纳法. $m = 2$ 时, 设 $\xi_1 + \xi_2 = 0$, 即 $\xi_1 = -\xi_2$ (其中 $\xi_j \in V_{\lambda_j}$ $j = 1, 2$) 则

$$\alpha(\xi_1 + \xi_2) = \lambda_1 \xi_1 + \lambda_2 \xi_2 = (\lambda_1 - \lambda_2) \xi_1 = 0.$$

由于 $\lambda_1 \neq \lambda_2$, 得 $\xi_1 = 0$, 从而 $\xi_2 = 0$, 故结论对 $m = 2$ 成立.

设结论对 $m - 1$ 成立, 下面考虑为 m 的情形. 设

$$\xi_1 + \xi_2 + \dots + \xi_m = 0, \quad \text{即} \quad \xi_m = -(\xi_1 + \xi_2 + \dots + \xi_{m-1}). \quad (2)$$

于是

$$\alpha(\xi_1 + \xi_2 + \dots + \xi_m) = \lambda_1 \xi_1 + \dots + \lambda_{m-1} \xi_{m-1} + \lambda_m \xi_m = 0, \quad (3)$$

将(2)式代入(3)式, 得

$$(\lambda_1 - \lambda_m) \xi_1 + \dots + (\lambda_{m-1} - \lambda_m) \xi_{m-1} = 0, \quad (4)$$

其中 $(\lambda_j - \lambda_m) \xi_j \in V_{\lambda_j}$ ($j = 1, 2, \dots, m-1$). 根据归纳假设, 由(4)式得 $(\lambda_j - \lambda_m) \xi_j = 0$ (其中 $\lambda_j \neq \lambda_m$) 从而 $\xi_j = 0$ ($j = 1, 2, \dots, m-1$). 再由(2)式又得 $\xi_m = 0$, 故结论对 m 成立.

由直和的等价条件及定理 7.7, 立即得以下推论:

推论 1 若 $\lambda_1, \dots, \lambda_m$ 是 σ 的互不相同的特征值, 则

$$V_{\lambda_i} \cap \sum_{j \neq i} V_{\lambda_j} = \{0\}, \quad i = 1, \dots, m. \quad (7-41)$$

由此又可得 $\lambda_i \neq \lambda_j$ 时, $V_{\lambda_i} \cap V_{\lambda_j} = \{0\}$, 这表明一个特征向量不能属于两个不同的特征值.

推论 2 σ 的不同特征值 $\lambda_1, \dots, \lambda_m$ 对应的特征向量 ξ_1, \dots, ξ_m 是线性无关的. 这是因为: 如果 ξ_1, \dots, ξ_m 线性相关, 则存在 ξ_j 可由其它特征向量线性表出, 不妨设

$$\xi_1 = c_2 \xi_2 + \dots + c_m \xi_m,$$

于是 $\xi_1 \in V_{\lambda_1} \cap \sum_{j=2}^m V_{\lambda_j}$, 这与推论 1 矛盾.

推论 3 σ 的不同特征值 $\lambda_1, \dots, \lambda_m$ 的特征子空间 $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_m}$ 的基向量合在一起构成的向量组是线性无关的. 这个线性无关的向量组是 $V_{\lambda_1} + V_{\lambda_2} + \dots + V_{\lambda_m}$ 的基.

定理 7.7 及其推论所阐述的线性变换的特征向量的性质同样适用于矩阵的特征向量.

7.5 可对角化的条件 相似标准形

如果有限维线性空间 $V(F)$ 的线性变换 σ 在 V 的某个基下对应的矩阵为

对角阵,我们就称 σ 为可对角化的线性变换. 同样,与对角阵相似的矩阵 A 称为可对角化矩阵.

在节 7.4 中,我们讲过 \mathbf{R}^3 中向量到二维子空间 W 上的投影变换 p 有三个线性无关的特征向量 ξ_1, ξ_2, ξ_3 , 因此投影变换 p 关于 \mathbf{R}^3 的这组基对应的矩阵为 $\text{diag}(1, 1, 0)$, 所以投影变换 p 可对角化.

在节 7.4 例 2 中, \mathbf{R}^3 的线性变换 σ 对应的三阶矩阵 A 有三个线性无关的特征向量 X_1, X_2, X_3 , 它们相应的特征值为 $0, -2, -2$, 如果将 X_1, X_2, X_3 按列排成一个三阶矩阵 P , 则有

$$P^{-1}AP = \text{diag}(0, -2, -2), \quad \text{即} \quad A \sim \text{diag}(0, -2, -2).$$

这里, 三维线性空间 \mathbf{R}^3 中的线性变换和三阶矩阵 A 均有三个线性无关的特征向量, 这是它们可对角化的充分必要条件, 更一般的结论为:

定理 7.8 n 维线性空间 $V(F)$ 的线性变换 α (或 $A \in M_n(F)$) 可对角化的充分必要条件为 α (或 A) 有 n 个线性无关的特征向量.

证 我们仅对 $A \in M_n(F)$ 的情形作证明. 如果 $A \sim \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 则存在可逆矩阵 P , 使得

$$P^{-1}AP = \Lambda, \quad \text{即} \quad AP = P\Lambda. \quad (1)$$

将矩阵 P 按列分块为 $P = (X_1, X_2, \dots, X_n)$, 代入 (1) 式得

$$A(X_1, X_2, \dots, X_n) = (X_1, X_2, \dots, X_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}. \quad (2)$$

于是

$$AX_j = \lambda_j X_j \quad (X_j \neq 0, j = 1, 2, \dots, n), \quad (3)$$

即 X_1, X_2, \dots, X_n 是 A 的 n 个线性无关的特征向量, 必要性得证.

反之, 如果 A 有 n 个线性无关的特征向量 X_1, X_2, \dots, X_n , 则 (3) 式成立, 由 (3) 式可得 (2) 式, 从而 (1) 式成立. 充分性也得证.

由定理 7.7 的推论 2 及本定理即得:

推论 若 n 维线性空间的线性变换 α (或 n 阶矩阵 A) 有 n 个互不相同的特征值, 则 α (或 A) 可对角化.

但是推论的条件对于 α (或 A) 可对角化不是必要的, 如节 7.4 例 2 中的 σ 和 A 可对角化, 而它们有二重特征值.

由定理 7.7 可见, 如果 n 维线性空间 $V(F)$ 的线性变换 σ 的 m 个互不相

同的特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ 的重数和等于 n , 且每个特征值的重数等于其特征子空间的维数, 则 σ 有 n 个线性无关的特征向量 (即 m 个特征子空间的基向量), 从而 σ 可对角化. 这个条件也是 σ 可对角化的必要条件, 为要证明其必要性, 需要先证明下面的定理.

定理 7.9 n 维线性空间 $V(F)$ 的线性变换 σ 的每个特征值 λ_i 的重数大于或等于其特征子空间 V_{λ_i} 的维数.

证 设 σ 在 V 的某个基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 下对应的矩阵为 A , 且 λ_i 是其 k 重特征值, 并设 $\{X_1, X_2, \dots, X_t\}$ 是矩阵 A 的特征子空间 $N(\lambda_i E - A)$ (简记作 N_{λ_i}) 的基, 即

$$\dim V_{\lambda_i} = \dim N_{\lambda_i} = t.$$

将 N_{λ_i} 的基扩充为 F^n 的基 $\{X_1, \dots, X_t, X_{t+1}, \dots, X_n\}$, 于是, 由

$$AX_j = \lambda_i X_j, \quad j = 1, \dots, t$$

及

$$AX_j = b_{1j}X_1 + \dots + b_{tj}X_t + b_{t+1,j}X_{t+1} + \dots + b_{nj}X_n, \quad j = t+1, \dots, n,$$

得

$$A(X_1, \dots, X_t, X_{t+1}, \dots, X_n) = (X_1, \dots, X_t, X_{t+1}, \dots, X_n) \begin{pmatrix} \lambda_i & \dots & 0 & b_{1,t+1} & \dots & b_{1n} \\ \dots & & & & & \dots \\ 0 & \dots & \lambda_i & b_{t,t+1} & \dots & b_{tn} \\ 0 & \dots & 0 & b_{t+1,t+1} & \dots & b_{t+1,n} \\ \dots & & & & & \dots \\ 0 & \dots & 0 & b_{n,t+1} & \dots & b_{nn} \end{pmatrix}. \quad (1)$$

将 (1) 式右端的矩阵分块表示为

$$\begin{pmatrix} \lambda_i E_t & B_1 \\ O & B_2 \end{pmatrix} \text{ 记作 } B,$$

并记 $(X_1, \dots, X_t, X_{t+1}, \dots, X_n) = P$ ($P \in M_n(F)$ 是可逆矩阵), 则由 (1) 式得

$$AP = PB, \quad \text{即} \quad P^{-1}AP = B.$$

于是

$$\begin{aligned} |\lambda E - A| &= |\lambda E - B| = \begin{vmatrix} (\lambda - \lambda_i)E_t & -B_1 \\ O & \lambda E_{n-t} - B_2 \end{vmatrix} \\ &= (\lambda - \lambda_i)^t |\lambda E_{n-t} - B_2| \end{aligned}$$

因此 λ_i 是 A 的大于或等于 t 重的特征值, 所以 $k \geq t$.

定理 7.10 n 维线性空间 $V(F)$ 的线性变换 σ 可对角化的充分必要条件是: σ 的每个特征值的重数等于其特征子空间的维数, 而且 σ 的不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ 的重数 r_1, r_2, \dots, r_m 之和等于 n .

证 充分性在定理 7.9 前已述. 下面证明必要性(用反证法).

如果 σ 在数域 F 上的 m 个不同特征值的重数和小于 n , 则相应的 m 个特征子空间的维数和也小于 n , 从而 σ 的线性无关的特征向量的个数小于 n , 故 σ 不可对角化.

如果 σ 的 m 个不同特征值的重数和等于 n , 但有一个特征值的重数大于其特征子空间的维数, 那么由定理 7.9 可知, σ 的 m 个特征子空间的维数和必小于 n , 故 σ 也不可对角化. 必要性得证.

定理 7.10 自然也适用于 n 阶矩阵 A . 如果在复数域上考虑 A 可否对角化, 那么 A 可对角化的充分必要条件只需说: A 的每个特征值的重数等于其特征子空间的维数. 因为 n 阶矩阵 A 在复数域上有 n 个特征值(r 重特征值算 r 个特征值).

对于一个 n 阶可对角化矩阵 A , 求变换矩阵 P , 使 $P^{-1}AP = \Lambda$ (对角阵), 其解题步骤如下:

(1) 求出 A 的所有不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$, 它们的重数 r_1, r_2, \dots, r_m 之和等于 n ;

(2) 求特征子空间 N_{λ_i} 的基 $\{X_{i1}, X_{i2}, \dots, X_{ir_i}\}$;

(3) 将 m 个特征子空间的基向量依次按列排成的 n 阶矩阵, 取作变换矩阵 P , 即

$$P = (X_{11} \dots X_{1r_1} \dots X_{m1} \dots X_{mr_m}), \quad (7-42)$$

则有

$$P^{-1}AP = \text{diag}(\lambda_1 \dots \lambda_1 \dots \lambda_m \dots \lambda_m).$$

对于可对角化矩阵 A , 如果不考虑其相似对角阵 Λ 中的对角元(即 A 的特征值)的排列次序, 则对角阵 Λ 是唯一确定的. 在这个意义下, 与 A 相似的对角阵 Λ 就称为 A 的相似标准形.

但必须注意, 由(7-42)式所示的变换矩阵 P 不是唯一的, 因为每个特征子空间 N_{λ_i} 的基不是唯一的. 因此, 可对角化的线性变换 σ 也不是只在唯一的一组基下对应于对角阵.

例 1 证明 n 阶幂等矩阵 A (即 $A^2 = A$) 相似于对角阵 $\Lambda = \text{diag}(1, 1, \dots, 1, 0, \dots, 0)$ 其中 1 的个数等于秩 (A) .

证 我们从线性变换的角度来证明这个矩阵的命题.

设 A 是 $\sigma \in L(V, V)$ 在 V 的某个基 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ 下所对应的矩阵, 于是由 $A^2 = A$ 可推出 $\sigma^2 = \sigma$ (即 σ 是幂等线性变换).

对于 n 维线性空间 V 的幂等线性变换 σ , 容易证明它的值域和核中的非零向量都是其特征向量, 而且值域与核之和为直和, 并等于 V , 故 σ 可对角化. 具体证明如下:

设 $\xi \in \mathcal{A}(V)$, 即存在 $\exists \beta \in V$, 使 $\mathcal{A}(\beta) = \xi$, 于是

$$\mathcal{A}(\xi) = \mathcal{A}(\mathcal{A}(\beta)) = \sigma^2(\beta) = \mathcal{A}(\beta) = \xi.$$

因此 σ 的值域为

$$\mathcal{A}(V) = \{\xi \mid \mathcal{A}(\xi) = \xi, \xi \in V\}, \quad (1)$$

所以 1 是 σ 的特征值, 值域中非零向量都是其特征向量. 又 σ 的核为

$$\sigma^{-1}(0) = \{\xi \mid \mathcal{A}(\xi) = 0, \xi \in V\}. \quad (2)$$

由(2)式可见 0 是 σ 的特征值, 核中所有非零向量都是其特征向量. 再由(1),

(2)式可得 $\mathcal{A}(V) \cap \sigma^{-1}(0) = \{0\}$, 而

$$\dim \mathcal{A}(V) + \dim \sigma^{-1}(0) = n,$$

于是

$$\mathcal{A}(V) \oplus \sigma^{-1}(0) = V.$$

又因为 $\dim \mathcal{A}(V) = r(\sigma) = r(A)$ (设 $r(A) = r$), 所以 $\dim \sigma^{-1}(0) = n - r$, 取值域与核的基分别为

$$\xi_1, \xi_2, \dots, \xi_r \quad \text{和} \quad \xi_{r+1}, \dots, \xi_n,$$

则由(1)(2)式得

$$\mathcal{A}(\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_n) = (\xi_1, \dots, \xi_r, \xi_{r+1}, \dots, \xi_n) \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

设 P 是基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 变为基 $\{\xi_1, \xi_2, \dots, \xi_n\}$ 的变换矩阵, 便有

$$P^{-1}AP = \text{diag}\{1, \dots, 1, 0, \dots, 0\}.$$

此例也可用幂等矩阵的特征值和特征向量的性质给以证明(证明留给读者作为练习, 证明时要用到第6章习题9的结论).

例2 \mathbf{R}^2 的旋转变换 r_θ ($0 < \theta < \pi$) 不可对角化, 因为它在自然基 $\{e_1, e_2\}$ 下对应的矩阵

$$T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

的特征方程 $|\lambda E - T| = \lambda^2 - 2\lambda \cos \theta + 1 = 0$ 没有实根, 所以旋转变换 r_θ 在

实数域上不可对角化.

例3 证明 r 阶上三角矩阵($r > 1$)

$$J_0 = \begin{pmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{pmatrix}$$

不与对角阵相似.

证 由于 $|\lambda E - J_0| = (\lambda - \lambda_0)^r$, 即 λ_0 是矩阵 J_0 的 r 重特征值, 而秩 $(\lambda_0 E - J_0) = r - 1$, 所以 λ_0 的特征子空间 N_{λ_0} 是一维的, 特征值的重数大于其特征子空间的维数, 故 J_0 不可对角化.

例3中的 J_0 称为约当块矩阵. 由约当块矩阵构成的对角块矩阵

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_m \end{pmatrix} \quad (7-43)$$

称为约当(Jordan)形矩阵, 其中

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \quad i = 1, \dots, m. \quad (7-44)$$

对于约当形矩阵 J , 需要指出 (1) 当 $i \neq j$ 时, 可能有 $\lambda_i = \lambda_j$; (2) 对角阵是最简单的约当形矩阵(此时每个约当块都是一阶的); (3) 约当块不全是一阶的约当形矩阵是不可对角化的.

可以证明: 任何一个 n 阶矩阵 A , 在复数域上都相似于唯一的约当形矩阵 J (唯一性是在不考虑约当块排列次序的意义下而言的), 并称 J 为 A 的相似标准形(或 A 的约当标准形), 于是由

$$P^{-1}AP = J, \quad \text{即} \quad AP = PJ$$

可见, 如果 J 如(7-43)式所示, 则 P 中有 m 个列向量是 A 分属于特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ 的特征向量. 反之, 如果 A 有 m 个线性无关的特征向量, 则 A 的约当标准形由 m 个约当块构成. 这里的复杂性在于: 如果 λ_1 是4重特征值, 其特征子空间 N_{λ_1} 是2维的, 那么相应的约当块可能有两种情况:

$$\left(\begin{array}{cc|cc} \lambda_1 & 1 & & \\ 0 & \lambda_1 & & \\ \hline & & \lambda_1 & 1 \\ & & 0 & \lambda_1 \end{array} \right), \quad \left(\begin{array}{c|ccc} \lambda_1 & & & \\ \hline & \lambda_1 & 1 & 0 \\ & 0 & \lambda_1 & 1 \\ & 0 & 0 & \lambda_1 \end{array} \right).$$

当然对于给定的矩阵 A , 只能是其中的某一个.

关于约当标准形, 我们不再深入讨论, 有兴趣的读者可参阅其它高等代数教材.

7.6 实对称矩阵的对角化

实对称矩阵是一类很重要的可对角化矩阵, 而且它的特征值都是实数. 为了证明这个结论, 我们先介绍复矩阵的共轭矩阵的概念及其基本运算.

定义 7.7 设 $A = (a_{ij}) \in M_{m \times n}(\mathbb{C})$ (\mathbb{C} 为复数域), 我们把 $\bar{A} = (\bar{a}_{ij})_{m \times n}$ 叫做 A 的共轭矩阵, 其中 \bar{a}_{ij} 是 a_{ij} 的共轭复数.

由定义可知: $\overline{\bar{A}} = A$; $(\bar{A}^T)^T = (\bar{A})^T$; A 为实矩阵时, $\bar{A} = A$; A 为实对称矩阵时, $(\bar{A})^T = A$.

根据定义及共轭复数的运算性质, 容易证明共轭矩阵有以下性质:

- (1) $\overline{kA} = \bar{k} \bar{A}$ ($k \in \mathbb{C}$);
- (2) $\overline{A+B} = \bar{A} + \bar{B}$;
- (3) $\overline{AB} = \bar{A} \bar{B}$;
- (4) $\overline{(AB)^T} = (\bar{B})^T (\bar{A})^T$;
- (5) 若 A 可逆, 则 $\overline{A^{-1}} = (\bar{A})^{-1}$;
- (6) 若 A 为方阵, 则 $\det \bar{A} = \overline{\det A}$;

(7) 若 $X = (x_1, x_2, \dots, x_n)^T$, 则 $(\bar{X})^T X = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n |x_i|^2 \geq 0$, 等号成立当且仅当 $X = 0$.

实矩阵的特征多项式是实系数多项式, 而实系数多项式的根并不一定都是实数. 但是, 实对称矩阵的特征多项式的根都是实数 (即特征值都是实数), 而且属于不同特征值的特征向量都是正交的. 下面给以证明.

定理 7.11 实对称矩阵 A 的特征值都是实数.

证 设 λ 是 A 的任一个特征值, 即存在非零向量 X , 使 $AX = \lambda X$. 欲证 λ 是实数, 只需证明 $\bar{\lambda} = \lambda$. 根据 $AX = \lambda X$ 和 $(\bar{A})^T = A$ 就有

$$\begin{aligned} \lambda (\bar{X})^T X &= (\bar{X})^T (\lambda X) = (\bar{X})^T AX = (\bar{X})^T (\bar{A})^T X \\ &= (\bar{AX})^T X = (\bar{\lambda X})^T X = \bar{\lambda} (\bar{X})^T X \end{aligned} \quad (1)$$

其中 $X \neq 0$, 即 $(\bar{X})^T X > 0$, 故由 (1) 式得 $\bar{\lambda} = \lambda$.

定理 7.12 实对称矩阵 A 的属于不同特征值的特征向量是正交的.

证 设 λ_1, λ_2 是 A 的两个不同的特征值(它们都是实数), 相应的特征向量 X_1, X_2 显然是实向量. 由 $AX_1 = \lambda_1 X_1, AX_2 = \lambda_2 X_2$, 即得

$$\begin{aligned}\lambda_1(X_1, X_2) &= (\lambda_1 X_1, X_2) = (AX_1, X_2) = (AX_1)^T X_2 = X_1^T A^T X_2 \\ &= X_1^T A X_2 = (X_1, \lambda_2 X_2) = (\lambda_2 X_2, X_1) = \lambda_2(X_2, X_1). \quad (2)\end{aligned}$$

而 $\lambda_1 \neq \lambda_2$, 所以由(2)式得 $(X_1, X_2) = 0$, 即 X_1 与 X_2 是正交的.

定理 7.13 若 A 是一个 n 阶实对称矩阵, 则存在 n 阶正交矩阵 Q , 使得

$$Q^{-1}AQ = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \quad (7-45)$$

证 对 n 作数学归纳法. $n=1$ 结论显然成立; 假设定理对 $n-1$ 阶实对称矩阵成立. 下面证明对 n 阶也成立.

设 λ_1 是 A 的一个特征值, $AX_1 = \lambda_1 X_1$, 其中 X_1 是长度为 1 的特征向量. 现将 X_1 扩充为 \mathbb{R}^n 的一组单位正交基 $\{X_1, X_2, \dots, X_n\}$, 于是 $AX_j \in \mathbb{R}^n$ 可由这组基线性表示($j=2, \dots, n$), 从而

$$A(X_1, X_2, \dots, X_n) = (X_1, X_2, \dots, X_n) \begin{pmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \dots & & & \dots \\ 0 & b_{n2} & \dots & b_{nn} \end{pmatrix}. \quad (3)$$

记 $P = (X_1, X_2, \dots, X_n)$ (P 为正交矩阵), 并将(3)式右端的矩阵用分块矩阵表示, 则(3)式可写为

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & \mathbf{b} \\ \mathbf{0} & B \end{pmatrix}. \quad (4)$$

由于 $P^{-1} = P^T$, $(P^{-1}AP)^T = P^T A^T (P^{-1})^T = P^{-1}AP$, 所以 $P^{-1}AP$ 为实对称矩阵, 因此(4)式中 $\mathbf{b} = \mathbf{0}$, B 为 $n-1$ 阶实对称矩阵. 根据归纳假设, 存在 $n-1$ 阶正交阵 Q_1 , 使得

$$Q_1^{-1}BQ_1 = \text{diag}(\lambda_2, \dots, \lambda_n).$$

取

$$S = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_1 \end{pmatrix} \quad (S \text{ 也是正交矩阵}),$$

便有

$$S^{-1}(P^{-1}AP)S = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_1^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q_1 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & Q_1^{-1} B Q_1 \end{pmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \quad (5)$$

于是,令 $Q = P$ (Q 也是正交矩阵),代入(5)式即得

$$Q^{-1} A Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的 n 个特征值.

定理 7.13 是一个重要的定理,它表明,任一个实对称矩阵 A 的相似标准形是由其特征值排成的一个对角矩阵,而且存在正交矩阵 Q 作为变换矩阵,使 $Q^{-1} A Q$ 为对角阵.但是定理的证明并没有提供求正交矩阵 Q 的方法.那么,对于给定的一个 n 阶实对称矩阵 A ,如何求正交矩阵 Q ,使 $Q^{-1} A Q$ 为对角阵呢?

由于 n 阶实对称矩阵 A 的不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_m$ 的特征子空间 $N_{\lambda_1}, \dots, N_{\lambda_m}$ 是两两正交的,而且它们的维数和等于 n (这是因为 A 可对角化).因此,把不同特征子空间的单位正交基合在一起,即 $X_{11}, \dots, X_{1r_1}, \dots, X_{m1}, \dots, X_{mr_m}$ 就是 \mathbf{R}^n 的一组单位正交基,将它们按列排成的 n 阶矩阵 Q (是一个正交矩阵)就使

$$Q^{-1} A Q = \text{diag}(\lambda_1, \dots, \lambda_1, \dots, \lambda_m, \dots, \lambda_m). \quad (6)$$

对每个特征子空间 N_{λ_i} ($i = 1, \dots, m$),求得任一组基后,再用 Schmidt 正交化方法可构造出它的单位正交基.显然,使(6)式成立的正交矩阵 Q 不是唯一的.

例 1 设

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

求正交矩阵 Q ,使 $Q^{-1} A Q$ 为对角矩阵,并求 A^k (k 为正整数).

解

$$\begin{aligned} |\lambda E - A| &= \begin{vmatrix} \lambda - 2 & -2 & 2 \\ -2 & \lambda - 5 & 4 \\ 2 & 4 & \lambda - 5 \end{vmatrix} = \begin{vmatrix} \lambda - 2 & -2 & 2 \\ 0 & \lambda - 1 & \lambda - 1 \\ 2 & 4 & \lambda - 5 \end{vmatrix} \\ &= \begin{vmatrix} \lambda - 2 & -2 & 4 \\ 0 & \lambda - 1 & 0 \\ 2 & 4 & \lambda - 9 \end{vmatrix} = (\lambda - 1) \begin{vmatrix} \lambda - 2 & 4 \\ 2 & \lambda - 9 \end{vmatrix} \\ &= (\lambda - 1)(\lambda - 10). \end{aligned}$$

对于特征值 $\lambda_1 = 1$ (二重), 由 $(\lambda_1 E - A)X = 0$ 即

$$\begin{pmatrix} -1 & -2 & 2 \\ -2 & -4 & 4 \\ 2 & 4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

解得特征子空间 N_{λ_1} 的一组基 (即基础解系)

$$X_1 = (-2 \ 1 \ 0)^T, \quad X_2 = (2 \ 0 \ 1)^T.$$

用 Schmidt 正交化方法求 N_{λ_1} 的单位正交基. 先正交化. 取

$$\beta_1 = X_1 = (-2 \ 1 \ 0)^T,$$

$$\beta_2 = X_2 - \frac{(X_2, \beta_1)}{(\beta_1, \beta_1)} \beta_1 = (2 \ 0 \ 1)^T - \frac{-4}{5}(-2 \ 1 \ 0)^T = \left(\frac{2}{5} \ \frac{4}{5} \ 1\right)^T.$$

再将 β_1, β_2 单位化, 得

$$X_1^* = \left(\frac{-2\sqrt{5}}{5} \ \frac{\sqrt{5}}{5} \ 0\right)^T, \quad X_2^* = \left(\frac{2\sqrt{5}}{15} \ \frac{4\sqrt{5}}{15} \ \frac{\sqrt{5}}{3}\right)^T.$$

对于特征值 $\lambda_2 = 10$, 由 $(\lambda_2 E - A)X = 0$, 即

$$\begin{pmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

解得特征子空间 N_{λ_2} 的一组基: $X_3 = (1 \ 2 \ -2)^T$, 将其单位化, 得 N_{λ_2} 的单位正交基

$$X_3^* = \left(\frac{1}{3} \ \frac{2}{3} \ \frac{-2}{3}\right)^T.$$

取正交矩阵

$$Q = \begin{pmatrix} | & | & | \\ X_1^* & X_2^* & X_3^* \\ | & | & | \end{pmatrix} = \begin{pmatrix} \frac{-2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3} \\ \frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3} \\ 0 & \frac{\sqrt{5}}{3} & \frac{-2}{3} \end{pmatrix},$$

则有

$$Q^{-1}AQ = \text{diag}(\lambda_1 \ \lambda_2 \ \lambda_3) = \text{diag}(1 \ 1 \ 10).$$

将上式的 $\text{diag}(1 \ 1 \ 10)$ 记作 Λ , 则

$$A = Q\Lambda Q^{-1},$$

从而

$$A^k = (Q\Lambda Q^{-1})^k = Q\Lambda^k Q^{-1}.$$

这里 $Q^{-1} = Q^T (\Lambda)^k = \text{diag}(1, 1, 10^k)$, 于是容易求得

$$A^k = \frac{1}{9} \begin{pmatrix} 8 + 10^k & 2(-1 + 10^k) & 2(1 - 10^k) \\ 2(-1 + 10^k) & 5 + 4 \cdot 10^k & 4(1 - 10^k) \\ 2(1 - 10^k) & 4(1 - 10^k) & 5 + 4 \cdot 10^k \end{pmatrix}.$$

在一些实用的问题中,常要求矩阵 A 的 k 次方. 如果 $A \sim \Lambda$ (对角阵) 则按上例的方法容易求得 A^k ; 一般的矩阵 A 在复数域上相似于约当形矩阵 J , 而 J^k 是易求的, 因为将其中的每个约当块 J_i 表示为

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix} = \lambda_i E + \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

按第4章节4.4中的(4-14)式, 容易求得 J_i^k . 因此求 A^k 的问题在理论上完满地解决了. 但实际上求一般的三阶实矩阵的特征值也不是轻而易举的, 所以, 有关特征值的计算问题, 会在《计算方法》课程中专门讨论.

例2 证明: 若 n 阶实对称矩阵 A 和 B 的特征值相同, 则 $A \sim B$, 且存在正交矩阵 Q 使 $Q^{-1}AQ = B$.

证 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 和 B 的特征值, 则存在正交阵 Q_1 和 Q_2 , 使得

$$Q_1^{-1}AQ_1 = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = Q_2^{-1}BQ_2,$$

于是

$$Q_2Q_1^{-1}AQ_1Q_2^{-1} = B.$$

令 $Q = Q_1Q_2^{-1}$ (Q 是正交矩阵, 且 $Q^{-1} = Q_2Q_1^{-1}$) 就有

$$Q^{-1}AQ = B.$$

事实上, 对任意两个可对角化的同阶矩阵 A 和 B , 它们相似的充分条件是, 它们的特征值相同. 当然, 此时对非实对称矩阵 A, B , 一般不存在正交矩阵 Q , 使 $Q^{-1}AQ = B$.

最后, 我们再附带地说一下: n 阶实对称矩阵是 n 维欧氏空间 $V(\mathbf{R})$ 的对称变换, 在单位正交基 $\{\varepsilon_1, \dots, \varepsilon_n\}$ 下所对应的矩阵, 所谓欧氏空间 $V(\mathbf{R})$ 的对称变换 σ , 即 σ 是 $V(\mathbf{R})$ 的一个线性变换, 而且 $\forall \alpha, \beta \in V$, 都有

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)).$$

投影变换和镜象变换都是对称变换. 事实上对于镜象变换

$$\varphi(\alpha) = \alpha - \chi(\omega, \alpha)\omega$$

(其中 ω 是与镜面正交的单位向量), $\forall \alpha, \beta \in V(\mathbf{R})$ 有

$$\begin{aligned}(\varphi(\alpha), \beta) &= (\alpha - \chi(\omega, \alpha)\omega, \beta) = (\alpha, \beta) - \chi(\omega, \alpha)\chi(\omega, \beta) \\ &= (\alpha, \beta - \chi(\omega, \beta)\omega) = (\alpha, \varphi(\beta)).\end{aligned}$$

7.7 双线性函数* 二次型

二次型是线性代数中的一个专题,它在几何、力学、物理和最优化等问题中有很多重要用途.一个 n 元二次型就是一个 n 元二次齐次函数.下面,我们先介绍双线性函数的概念及其矩阵表示.

定义 7.8 我们称 f 为线性空间 $V(F)$ 上的一个双线性函数,如果 f 是 $V \times V$ 到 F 的映射,而且 $\forall \alpha, \beta, \alpha_i, \beta_i \in V, k_i \in F (i = 1, 2)$ 均有

$$(1) f(\alpha, k_1\beta_1 + k_2\beta_2) = k_1f(\alpha, \beta_1) + k_2f(\alpha, \beta_2),$$

$$(2) f(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1f(\alpha_1, \beta) + k_2f(\alpha_2, \beta).$$

由定义可见, V 上的双线性函数 $f(\alpha, \beta)$ 对两个变元 α, β 都是线性函数,即将其中一个变元固定时是另一个变元的线性函数.

同样可以定义线性空间 $V(F)$ 上的多重线性函数 $f(\alpha_1, \alpha_2, \dots, \alpha_n)$. 数域 F 上的 n 阶行列式 $D(\alpha_1, \alpha_2, \dots, \alpha_n)$ 实际上就是 F^n 上的一个 n 重线性函数.

例 1 欧氏空间 $V(\mathbf{R})$ 上的内积 (α, β) 是 $V(\mathbf{R})$ 上的一个双线性函数.

例 2 如果 $f_1(\alpha), f_2(\alpha)$ 是线性空间 $V(F)$ 上的两个线性函数, 则

$$f(\alpha, \beta) = f_1(\alpha)f_2(\beta) \quad (\forall \alpha, \beta \in V)$$

是 V 上的一个双线性函数.

以上二例的验证留给读者去完成.

例 3 设 $A \in M_n(F)$ 是一个固定的矩阵, 则

$$f(X, Y) = X^T A Y = (X, AY)$$

(其中 X, Y 是 F^n 中任意两个列向量) 是 F^n 上的一个双线性函数. 事实上, $\forall X_1, X_2 \in F^n$ 和 $\forall k_1, k_2 \in F$,

$$\begin{aligned}f(k_1X_1 + k_2X_2, Y) &= (k_1X_1 + k_2X_2)^T A Y = k_1X_1^T A Y + k_2X_2^T A Y \\ &= k_1f(X_1, Y) + k_2f(X_2, Y),\end{aligned}$$

所以 f 对 X 是线性函数, 同理可证 f 对 Y 也是线性函数, 即

$$f(X, k_1Y_1 + k_2Y_2) = k_1f(X, Y_1) + k_2f(X, Y_2).$$

现在, 我们来讨论 n 维线性空间 $V(F)$ 上的双线性函数 $f(\alpha, \beta)$ 的矩阵表示.

设 $B = \{e_1, e_2, \dots, e_n\}$ 是 $V(F)$ 的一组基, V 中向量 α, β 关于基 B 的坐标分别为:

$$X = (x_1, \dots, x_n)^T, \quad Y = (y_1, \dots, y_n)^T,$$

即

$$\alpha = \sum_{i=1}^n x_i e_i, \quad \beta = \sum_{j=1}^n y_j e_j, \quad (1)$$

于是

$$f(\alpha, \beta) = f\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j f(e_i, e_j). \quad (2)$$

令 $a_{ij} = f(e_i, e_j)$ ($i, j = 1, \dots, n$), 则

$$\begin{aligned} f(\alpha, \beta) &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i y_j = \sum_{i=1}^n x_i (a_{i1} y_1 + a_{i2} y_2 + \dots + a_{in} y_n) \\ &= (x_1 \ x_2 \ \dots \ x_n) \begin{pmatrix} a_{11} y_1 + a_{12} y_2 + \dots + a_{1n} y_n \\ a_{21} y_1 + a_{22} y_2 + \dots + a_{2n} y_n \\ \dots \dots \dots \\ a_{n1} y_1 + a_{n2} y_2 + \dots + a_{nn} y_n \end{pmatrix} \\ &= (x_1 \ x_2 \ \dots \ x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \mathbf{X}^T \mathbf{A} \mathbf{Y}. \quad (3) \end{aligned}$$

其中 $A = (a_{ij})_{n \times n} = (f(e_i, e_j))_{n \times n}$ 称为双线性函数 $f(\alpha, \beta)$ 在基 B 下的度量矩阵. 显然, 如果已知一个双线性函数在一组基下的度量矩阵, 就可算出任意向量 α, β 所对应的函数值.

从上面的讨论可见: 取定了 V 的一组基, 每个双线性函数 f 都对应于唯一的一个矩阵 A (即在基下的度量矩阵). 而且不同的双线性函数 f_1, f_2 必对应不同的矩阵 (因为如果 f_1, f_2 对应同一个矩阵, 由 (3) 式可得 $f_1 = f_2$; 反之, 任给 $M_n(F)$ 中的一个矩阵 A , 如果对 V 中任意向量 α, β 定义如 (3) 式所示的一个函数 $f(\alpha, \beta)$, 则容易验证 $f(\alpha, \beta)$ 是一个双线性函数, 而且由 (3) 式容易算得 $f(e_i, e_j) = a_{ij}$ ($i, j = 1, \dots, n$), 即 A 就是这个双线性函数所对应的度量矩阵. 因此, 给定了 n 维线性空间 $V(F)$ 的一组基后, V 上的双线性函数与 $M_n(F)$ 中的矩阵一一对应, 所以双线性函数可用它的度量矩阵表示.

当然, 一个双线性函数在 V 的不同基下的度量矩阵一般是不同的, 其结论如下:

定理 7.14 设线性空间 $V(F)$ 的双线性函数 $f(\alpha, \beta)$ 在基 $B_1 = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$ 和 $B_2 = \{\eta_1, \eta_2, \dots, \eta_n\}$ 下的度量矩阵分别为 A 和 B , 如果

$$(\eta_1, \eta_2, \dots, \eta_n) = (\epsilon_1, \epsilon_2, \dots, \epsilon_n) C,$$

则

$$B = C^T A C. \quad (7-46)$$

证 设 $A = (a_{ij})_{n \times n}$, $a_{ij} = f(\epsilon_i, \epsilon_j)$; $B = (b_{ij})_{n \times n}$, $b_{ij} = f(\eta_i, \eta_j)$;
 $C = (c_{ij})_{n \times n}$, $C^T = (c'_{ij})_{n \times n}$, 其中 $c'_{ij} = c_{ji}$ ($i, j = 1, 2, \dots, n$), 于是

$$b_{ij} = f(\eta_i, \eta_j) = f\left(\sum_{k=1}^n c_{ki} \epsilon_k, \sum_{m=1}^n c_{mj} \epsilon_m\right) = \sum_{k=1}^n \sum_{m=1}^n c_{ki} c_{mj} f(\epsilon_k, \epsilon_m)$$

$$\begin{aligned}
 &= \sum_{k=1}^n \sum_{m=1}^n c_{ki} c_{mj} a_{km} = \sum_{m=1}^n \left(\sum_{k=1}^n c'_{ik} a_{km} \right)_{mj} \\
 &= \sum_{m=1}^n (C^T A)_{im} c_{mj} = (C^T A C)_{ij}, \quad i, j = 1, \dots, n,
 \end{aligned}$$

所以 $B = C^T A C$.

定义 7.9 我们称 n 阶矩阵 A 相合于 B (记作 $A \simeq B$) , 如果存在可逆矩阵 C , 使得

$$B = C^T A C.$$

容易验证 , 矩阵的相合 (或说 : 合同) 关系也是集合 $M_n(F)$ 中的一种等价关系 , 即相合关系具有 :

- (1) 自反性 , 即 $\forall A \in M_n(F), A \simeq A$;
- (2) 对称性 , 即若 $A \simeq B$, 则 $B \simeq A$;
- (3) 传递性 , 即若 $A \simeq B, B \simeq C$, 则 $A \simeq C$.

这个关系说明 , $V(F)$ 上的双线性函数 $f(\alpha, \beta)$ 对不同的基的度量矩阵是一个等价类 , 它们可以看成是“一样的”.

定义 7.10 线性空间 $V(F)$ 上的一个双线性函数 $f(\alpha, \beta)$, 如果 $\forall \alpha, \beta \in V$, 都有 :

- (1) $f(\alpha, \beta) = f(\beta, \alpha)$, 则 $f(\alpha, \beta)$ 叫做对称双线性函数 ;
- (2) $f(\alpha, \beta) = -f(\beta, \alpha)$, 则 $f(\alpha, \beta)$ 叫做反对称双线性函数 .

例如 : 欧氏空间上的内积 (α, β) 是对称的双线性函数 , 数域 F 上的 n 阶行列式 $D(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是反对称的 n 重线性函数 .

显然 , 对称双线性函数在 $V(F)$ 的一组基 $B = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$ 下对应的度量矩阵 $A = (a_{ji})_{n \times n}$ 是数域 F 上的对称矩阵 , 因为

$$a_{ij} = f(\epsilon_i, \epsilon_j) = f(\epsilon_j, \epsilon_i) = a_{ji} \quad (i, j = 1, \dots, n).$$

反对称双线性函数在基 B 下对应的度量矩阵 A 是数域 F 上的反对称矩阵 .

定义 7.11 n 个元 x_1, x_2, \dots, x_n 的二次齐次多项式

$$\begin{aligned}
 f(x_1, x_2, \dots, x_n) &= \sum_{i=1}^n a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j \quad (7-47) \\
 &= a_{11} x_1^2 + 2a_{12} x_1 x_2 + 2a_{13} x_1 x_3 + \dots + 2a_{1n} x_1 x_n \\
 &\quad + a_{22} x_2^2 + 2a_{23} x_2 x_3 + \dots + 2a_{2n} x_2 x_n \\
 &\quad + \dots + a_{n-1, n-1} x_{n-1}^2 + 2a_{n-1, n} x_{n-1} x_n + a_{nn} x_n^2
 \end{aligned}$$

(其中系数 a_{ij} 是数域 F 中的数) , 叫做数域 F 上的 n 元二次型 (简称二次型) .

我们主要研究实数域上的二次型 (简称实二次型) , 以后说到二次型 , 如不加说明 , 都是指实二次型 .

如果令 $a_{ji} = a_{ij} (1 \leq i < j \leq n)$, 则 (7-47) 式如前面的 (2) 式那样可以

表示为

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{X}^T \mathbf{A} \mathbf{X}, \quad (7-48)$$

其中 $\mathbf{X} = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$, $\mathbf{A} = (a_{ij})_{n \times n}$ 是实对称矩阵, 并称对称矩阵 \mathbf{A} 为 (7-47) 式所示的二次型的矩阵.

(7-48) 式恰是 n 维线性空间 $V(F)$ 上的对称双线性函数 $f(\alpha, \beta)$, 当 $\beta = \alpha$ 时, $f(\alpha, \beta)$ 在 V 的基 $B_1 = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 下的坐标表示式, 即

$$f(\alpha, \alpha) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \mathbf{X}^T \mathbf{A} \mathbf{X}, \quad (7-49)$$

其中 $(\alpha)_{B_1} = \mathbf{X} = (x_1, x_2, \dots, x_n)^T$, \mathbf{A} 是 $f(\alpha, \beta)$ 在基 B_1 下的度量矩阵.

因此 (7-49) 式所示的 n 元二次型 (令 $a_{ij} = a_{ji}$) 是与 n 维线性空间 $V(F)$ 上的对称双线性函数 $f(\alpha, \alpha)$ 一一对应的, 也就是与对称矩阵是一一对应的.

双线性函数在不同基下的度量矩阵是相合的. 根据定理 7.14 的结论, 如果对 α 作坐标变换, 令 α 在基 B_1 和基 B_2 下的坐标分别为 \mathbf{X} 和 \mathbf{Y} ($\mathbf{X} = \mathbf{C}\mathbf{Y}$, \mathbf{C} 为基 B_1 变为基 B_2 的变换矩阵), 则 (7-49) 式的 $f(\alpha, \alpha)$ 在基 B_2 下的坐标表示式为

$$f(\alpha, \alpha) = \mathbf{Y}^T (\mathbf{C}^T \mathbf{A} \mathbf{C}) \mathbf{Y}. \quad (7-50)$$

这与 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ 直接代入 (7-49) 式所得的结果是一致的.

因此, 对于一个对称双线性函数, 我们自然希望找到一组基, 使得 $f(\alpha, \alpha)$ 在这组基下的坐标表示式最简单. 显然, 如果 $(\mathbf{C}^T \mathbf{A} \mathbf{C}) = \text{diag}(d_1, d_2, \dots, d_n)$, 则

$$\mathbf{Y}^T (\mathbf{C}^T \mathbf{A} \mathbf{C}) \mathbf{Y} = d_1 y_1^2 + d_2 y_2^2 + \dots + d_n y_n^2 \quad (7-51)$$

是 $f(\alpha, \alpha)$ 的一个最简单的坐标表示式, 并把它叫做 $f(\alpha, \alpha) = \mathbf{X}^T \mathbf{A} \mathbf{X}$ 的标准形.

下一节, 我们将论述, 对于任一个实对称矩阵, 都存在可逆矩阵 \mathbf{C} , 使得 $\mathbf{C}^T \mathbf{A} \mathbf{C} = \text{diag}(d_1, d_2, \dots, d_n)$, 并将这个对角阵叫做 \mathbf{A} 的相合标准形.

7.8 实二次型的标准形 实对称矩阵的相合标准形

在节 7.6 中讲过, 任一个 n 阶实对称矩阵 \mathbf{A} , 都存在正交矩阵 \mathbf{Q} , 使得

$$\mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} = \mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

所以, 实对称矩阵的相似标准形也是它的相合标准形. 相应地, 对 n 元二次型

$$f(x_1, x_2, \dots, x_n) = \mathbf{X}^T \mathbf{A} \mathbf{X}$$

作坐标变换 $\mathbf{X} = \mathbf{Q}\mathbf{Y}$ (这是正交变换) 就得到它的一个标准形, 即

$$f(x_1, x_2, \dots, x_n) = \mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2,$$

于是有下面的重要定理.

定理 7.15 (主轴定理) 对于任一个 n 元二次型 $f(x_1, x_2, \dots, x_n) =$

$\mathbf{X}^T \mathbf{A} \mathbf{X}$ 都存在正交变换 $\mathbf{X} = \mathbf{Q} \mathbf{Y}$ 使得

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2. \quad (7-52)$$

其中 $\lambda_1, \dots, \lambda_n$ 是实对称矩阵 \mathbf{A} 的 n 个特征值, \mathbf{Q} 的 n 个列向量是 \mathbf{A} 属于 $\lambda_1, \dots, \lambda_n$ 的 n 个单位正交的特征向量.

例 1 对二次型 $f(x_1, x_2) = x_1^2 - 3x_2^2 + 2\sqrt{5}x_1x_2$ 作正交变换, 将其化为标准形, 并给出变换矩阵.

解 二次型所对应的矩阵为

$$\mathbf{A} = \begin{pmatrix} 1 & \sqrt{5} \\ \sqrt{5} & -3 \end{pmatrix}$$

\mathbf{A} 的特征多项式 $|\lambda \mathbf{E} - \mathbf{A}| = (\lambda - 2)(\lambda + 4)$, 对特征值 $\lambda_1 = 2, \lambda_2 = -4$, 容易求得特征子空间 N_{λ_1} 和 N_{λ_2} 的单位正交基为:

$$\xi_1 = \left(\sqrt{\frac{5}{6}}, \sqrt{\frac{1}{6}} \right)^T \quad \text{和} \quad \xi_2 = \left(-\sqrt{\frac{1}{6}}, \sqrt{\frac{5}{6}} \right)^T.$$

取正交矩阵

$$\mathbf{Q} = (\xi_1, \xi_2) = \begin{pmatrix} \sqrt{\frac{5}{6}} & -\sqrt{\frac{1}{6}} \\ \sqrt{\frac{1}{6}} & \sqrt{\frac{5}{6}} \end{pmatrix}, \quad \text{则} \quad \mathbf{Q}^T \mathbf{A} \mathbf{Q} = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}.$$

于是, 对二次型 $f(x_1, x_2) = \mathbf{X}^T \mathbf{A} \mathbf{X}$ (作坐标变换 $\mathbf{X} = \mathbf{Q} \mathbf{Y}$ (即将坐标轴绕原点按逆时针方向旋转 θ 角 ($\theta = \arctan(\frac{1}{\sqrt{5}})$)), 可将其化为标准形:

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{Y}^T (\mathbf{Q}^T \mathbf{A} \mathbf{Q}) \mathbf{Y} = 2y_1^2 - 4y_2^2.$$

对二次曲线方程 $f(x_1, x_2) = x_1^2 - 3x_2^2 + 2\sqrt{5}x_1x_2 = 1$, 按例 1 的方法作坐标变换 (坐标轴旋转) $\mathbf{X} = \mathbf{Q} \mathbf{Y}$, 即可将其化为标准方程

$$2y_1^2 - 4y_2^2 = \frac{y_1^2}{\left(\sqrt{\frac{1}{2}}\right)^2} - \frac{y_2^2}{\left(\sqrt{\frac{1}{4}}\right)^2} = 1.$$

这是一条双曲线, 它的实半轴和虚半轴分别是两个特征值的倒数的平方根. 这显然是二次曲线 (或说二次型矩阵) 的不变量. 如果双曲线上的点 P 关于自然基 $\{e_1, e_2\}$ 的坐标为 (x_1, x_2) , 则点 P 关于基 $\{\xi_1, \xi_2\}$ 的坐标为 (y_1, y_2) , 这里的 $\{\xi_1, \xi_2\}$ 是两个特征值的单位正交的特征向量, 它们的方向是双曲线实轴和虚轴的方向 (如图 7-7), 因此定理 7.15 也称为主轴定理.

例 2 求三元二次齐次函数

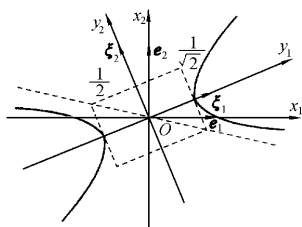


图 7-7

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

在 $\mathbf{X} = (x_1, x_2, x_3)^T$ 满足条件 $\mathbf{X}^T \mathbf{X} = x_1^2 + x_2^2 + x_3^2 = 1$ 时的最小值.

解 先对二次型 $f(\mathbf{X}) = \mathbf{X}^T \mathbf{A} \mathbf{X}$ 作正交变换 $\mathbf{X} = \mathbf{Q} \mathbf{Y}$ 将其化为标准形 $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2$ 然后在条件 $\mathbf{X}^T \mathbf{X} = \mathbf{Y}^T \mathbf{Q}^T \mathbf{Q} \mathbf{Y} = \mathbf{Y}^T \mathbf{Y} = y_1^2 + y_2^2 + y_3^2 = 1$ 下讨论函数的最小值. 该二次型对应的实对称矩阵为

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

它的特征多项式 $|\lambda \mathbf{E} - \mathbf{A}| = (\lambda + 2)(\lambda - 1)^2$.

对于特征值 $\lambda_1 = -2$, 求得一个特征向量 $(1, -1, 1)^T$ 将它单位化, 得

$$\xi_1 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)^T.$$

对于特征值 $\lambda_2 = 1$, 求得两个线性无关的特征向量 $(1, 1, 0)^T$ 及 $(-1, 0, 1)^T$, 再用 Schmidt 正交化方法, 得两个单位正交的特征向量

$$\xi_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right)^T, \quad \xi_3 = \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)^T.$$

取正交矩阵

$$\mathbf{Q} = (\xi_1, \xi_2, \xi_3) = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix},$$

则有

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

对二次型 $f(\mathbf{X}) = \mathbf{X}^T A \mathbf{X}$ 作正交变换 $\mathbf{X} = Q\mathbf{Y}$ 得

$$f(\mathbf{X}) = \mathbf{Y}^T (Q^T A Q) \mathbf{Y} = -2y_1^2 + y_2^2 + y_3^2. \quad (1)$$

相应地, 条件 $\mathbf{X}^T \mathbf{X} = x_1^2 + x_2^2 + x_3^2 = 1$ 化为

$$\mathbf{Y}^T (Q^T Q) \mathbf{Y} = \mathbf{Y}^T \mathbf{Y} = y_1^2 + y_2^2 + y_3^2 = 1. \quad (2)$$

于是原题化为对(1)式的三元二次齐次函数在满足条件(2)时求其最小值. 此时, 显然有

$$f(\mathbf{X}) = -2y_1^2 + y_2^2 + y_3^2 \geq -2(y_1^2 + y_2^2 + y_3^2) = -2$$

又当 $\mathbf{Y} = (y_1, y_2, y_3)^T = (\pm 1, 0, 0)^T$ 时, $f = -2$, 所以 f 满足条件(2)的最小值 $f_{\min} = -2$, 而且它仅在 $\mathbf{Y}_1 = (1, 0, 0)^T$ 和 $\mathbf{Y}_1 = (-1, 0, 0)^T$ 处取得最小值.

回到变元 $\mathbf{X} = (x_1, x_2, x_3)^T$, 则 $f(x_1, x_2, x_3)$ 在 $\mathbf{X}_1 = Q\mathbf{Y}_1 = \xi_1 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T$ 和 $\mathbf{X}_2 = Q\mathbf{Y}_2 = -\xi_1 = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)^T$ 处取得最小值.

n 元二次型 $f(\mathbf{X}) = \mathbf{X}^T A \mathbf{X}$ 也可以通过一般的坐标变换 $\mathbf{X} = C\mathbf{Y}$ (称为仿射变换, 矩阵 C 可逆) 化为标准形, 即

$$\mathbf{Y}^T (C^T A C) \mathbf{Y} = d_1 y_1^2 + \dots + d_n y_n^2.$$

即对任一个实对称矩阵 A , 都存在变换矩阵 C , 使得

$$C^T A C = \text{diag}(d_1, \dots, d_n).$$

但要注意, 这里的 d_1, \dots, d_n 一般不是 A 的特征值, 矩阵 C 的列向量一般也不是 A 的特征向量.

这里常用的方法是“配方法”和“初等变换法”. 下面, 我们通过例题介绍这两种方法.

例3 用配方法把三元二次型

$$f(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$$

化为标准形, 并求所用的坐标变换 $\mathbf{X} = C\mathbf{Y}$ 及变换矩阵 C .

解 先按 x_1^2 及含有 x_1 的混合项配成完全平方, 即

$$\begin{aligned} f(x_1, x_2, x_3) &= 2(x_1^2 + 2x_1(x_2 - x_3) + (x_2 - x_3)^2) \\ &\quad - 2(x_2 - x_3)^2 + 3x_2^2 + x_3^2 - 8x_2x_3 \\ &= 2(x_1 + x_2 - x_3)^2 + x_2^2 - x_3^2 - 4x_2x_3. \end{aligned}$$

在上式中, 再按 $x_2^2 - 4x_2x_3$ 配成完全平方, 于是

$$f(x_1, x_2, x_3) = 2(x_1 + x_2 - x_3)^2 + (x_2 - 2x_3)^2 - 5x_3^2. \quad (1)$$

令

$$\begin{cases} y_1 = x_1 + x_2 - x_3 \\ y_2 = x_2 - 2x_3 \\ y_3 = x_3 \end{cases} \quad (2)$$

将(2)式代入(1)式,得二次型的标准形

$$f(x_1, x_2, x_3) = 2y_1^2 + y_2^2 - 5y_3^2.$$

从(2)式中解出 x_1, x_2, x_3 得

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \quad (3)$$

(3)式是化二次型为标准形所作的坐标变换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ (变换矩阵 \mathbf{C} 是(3)式中的三阶矩阵).

例4 对三元二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 - 2x_1x_3 + 2x_2x_3 \quad (1)$$

作坐标变换 $\mathbf{X} = \mathbf{C}\mathbf{Z}$ 将其化为标准形.

解 此时欲用配方法,先要作如下的坐标变换:

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases}, \quad \text{即} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \quad (2)$$

将(2)式记为 $\mathbf{X} = \mathbf{C}_1\mathbf{Y}$, 并代入(1)式,得

$$f(x_1, x_2, x_3) = 2y_1^2 - 2y_2^2 - 4y_2y_3 = 2y_1^2 - 2(y_2 + y_3)^2 + 2y_3^2. \quad (3)$$

令

$$\begin{cases} z_1 = y_1 \\ z_2 = y_2 + y_3 \\ z_3 = y_3 \end{cases},$$

即

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}. \quad (4)$$

将(4)式记为 $\mathbf{Y} = \mathbf{C}_2\mathbf{Z}$, 并代入(3)式,得二次型的标准形

$$f(x_1, x_2, x_3) = 2z_1^2 - 2z_2^2 + 2z_3^2.$$

此时所作的坐标变换为 $\mathbf{X} = \mathbf{C}_1\mathbf{Y} = \mathbf{C}_1(\mathbf{C}_2\mathbf{Z}) = (\mathbf{C}_1\mathbf{C}_2)\mathbf{Z}$, 其变换矩阵 $\mathbf{C} =$

$(C_1 C_2)$, 即

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

按例 3、例 4 提供的方法, 任何 n 元二次型都可用配方法化为标准形. 相应的变换矩阵为主对角元为 1 的上三角矩阵和例 4 中的 (2) 式类型的对角块矩阵, 或者是这两类矩阵的乘积.

任一个 n 阶实对称矩阵 A , 也都可以通过一系列相同类型的初等行、列变换化成其相合标准形 (对角形矩阵). 所谓相同类型的初等行、列变换指的是:

(1) 如果用倍加初等阵 $E_{jk}(c)$ 右乘 A (即 A 的第 j 列乘 c 加到第 j 列), 那么再相应地用 $E_{jk}^T(c) = E_{ij}(c)$ 左乘 $AE_{jk}(c)$ (即列变换后的 A 的第 i 行乘 c 加到第 j 行). 变换后的矩阵 $E_{jk}^T(c)AE_{jk}(c)$ 仍是对称阵.

(2) 如果用 $E_i(c)$ 右乘 A , 则再用 $E_i^T(c)$ 左乘 $AE_i(c)$, 即 A 的第 i 列和第 i 行都乘非零常数 c , 显然 $E_i^T(c)AE_i(c)$ 仍是对称阵.

(3) 如果用 E_{ij} 右乘 A , 则再用 E_{ij}^T 左乘 AE_{ij} , 即 A 的第 i 列与第 j 列对换后, 再将第 i 行与第 j 行对换位置, 如此所得的 $E_{ij}^T AE_{ij}$ 也是对称阵.

对于一个 n 阶实对称矩阵

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

(1) 如果 $a_{11} \neq 0$, 由于 $a_{1j} = a_{j1}$ ($j = 1, 2, \dots, n$), 因此对 A 作相同的倍加行、列变换, 可将第 1 行与第 1 列的其它元素全化为零, 得

$$\begin{pmatrix} a_{11} & 0 \\ 0 & A_1 \end{pmatrix},$$

其中 A_1 为 $n-1$ 阶实对称矩阵.

(2) 如果 $a_{11} = 0$, 但存在 $a_{ii} \neq 0$, 此时, 先将第 1 列与第 i 列对换, 再将第 1 行与第 i 行对换, 把 a_{ii} 换到 1 行, 1 列的位置, 就化为 (1) 的情况.

(3) 如果 $a_{ii} = 0$ ($i = 1, 2, \dots, n$), $a_{ij} \neq 0$, 此时, 可将第 j 列加到第 i 列, 将第 j 行加到第 i 行, 如此 i 行 i 列元素化为 $2a_{ij} \neq 0$, 这就化为 (2) 的情况.

这样, 我们就可用数学归纳法证明: 对任一个 n 阶实对称矩阵 A , 都存在初等矩阵 P_1, P_2, \dots, P_k , 使得

$$P_k^T \cdots P_2^T P_1^T A P_1 P_2 \cdots P_k = \text{diag}(d_1, d_2, \dots, d_m).$$

这时使 A 相合于对角阵的变换矩阵为

$$C = P_1 P_2 \cdots P_k = E P_1 P_2 \cdots P_k.$$

因此,对 A 作的列变换同样施加于单位矩阵 E ,即得变换矩阵 C .

例 5 用初等变换法将例 4 的二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X} = f(x_1, x_2, x_3) = 2x_1x_2 - 2x_1x_3 + 2x_2x_3$ 化为标准形,并求所作的坐标变换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$ 的变换矩阵 C .

解 将二次型矩阵 A 与单位矩阵 E 排成 $\begin{pmatrix} A \\ E \end{pmatrix}$,对 A 作相同类型的初等行、列变换使之化为对角阵,相应地,对 E 作同样的初等列变换, E 就化为变换矩阵 C .(以下 $(i)[j]$ 分别表示 i 列 j 行).

$$\begin{pmatrix} A \\ E \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} (1)+(2) \end{smallmatrix}]{\begin{smallmatrix} (1)+(2) \\ [1]+[2] \end{smallmatrix}} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \xrightarrow[\begin{smallmatrix} (2)+(1) \times (-1/2) \\ [2]+[1] \times (-1/2) \end{smallmatrix}]{\begin{smallmatrix} (3)+(2) \times 2 \\ [3]+[2] \times 2 \end{smallmatrix}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 1 & 0 \\ 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[\begin{smallmatrix} (3)+(2) \times 2 \end{smallmatrix}]{\begin{smallmatrix} (3)+(2) \times 2 \\ [3]+[2] \times 2 \end{smallmatrix}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 2 \\ 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

于是,取

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

就有 $C^T A C = \text{diag}\left(2, -\frac{1}{2}, 2\right)$ 相应地作坐标变换 $\mathbf{X} = \mathbf{C}\mathbf{Y}$,则二次型化为标准形,即

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = 2y_1^2 - \frac{y_2^2}{2} + 2y_3^2.$$

从例 2、例 4 和例 5 可见,用三种不同的方法把同一个二次型化成了三个不同的标准形,即同一个三阶实对称矩阵 A 相合于三个不同的对角阵:

$$\text{diag}(-2, 1, 1); \quad \text{diag}(2, -2, 2); \quad \text{diag}\left(2, -\frac{1}{2}, 2\right),$$

但是它们的对角元中都有两个正数,一个负数.事实上,容易证明(留给读者证明)

上面三个对角阵都相合于对角阵 $\text{diag}(1, 1, \dots, 1, -1, -1, \dots, -1, 0, \dots, 0)$. 从而 A 也相合于这个对角阵.

一般, 我们把实对称矩阵 A 的相合标准形

$$\text{diag}(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$$

叫做 A 的相合规范形, 其中 $+1$ 的个数和 -1 的个数分别叫做 A 的正惯性指数和负惯性指数. 由

$$C^T A C = \Lambda = \text{diag}(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$$

及 C 为可逆矩阵, 立即可得

$$\text{秩}(A) = \text{秩}(\Lambda) = A \text{ 的正、负惯性指数之和.}$$

显然, 当 A 为满秩矩阵时, 其相合规范形的主对角元中没有数零.

定理 7.16 (惯性定理) 实对称矩阵 A 的正、负惯性指数是由 A 唯一确定的.

证 如果对实二次型 $f = \mathbf{X}^T A \mathbf{X}$ 作坐标变换 $\mathbf{X} = B\mathbf{Y}$ 和 $\mathbf{X} = C\mathbf{Z}$ (B, C 是可逆矩阵), 分别使之化为相合规范形

$$f = y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 \quad (1)$$

和

$$f = z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_r^2, \quad (2)$$

其中 r 为 A 的秩, 则必有 $p = k$. 我们用反证法:

假设 $p > k$, 此时 (1)(2) 式可写为

$$\begin{aligned} f &= y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 \\ &= z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_r^2. \end{aligned} \quad (3)$$

由 $\mathbf{Z} = C^{-1}\mathbf{X} = (C^{-1}B)\mathbf{Y} = D\mathbf{Y}$ (其中 $D = C^{-1}B$), 即

$$\begin{cases} z_1 = d_{11}y_1 + d_{12}y_2 + \dots + d_{1n}y_n \\ z_2 = d_{21}y_1 + d_{22}y_2 + \dots + d_{2n}y_n \\ \dots\dots\dots \\ z_k = d_{k1}y_1 + d_{k2}y_2 + \dots + d_{kn}y_n \\ \dots\dots\dots \\ z_n = d_{n1}y_1 + d_{n2}y_2 + \dots + d_{nn}y_n. \end{cases} \quad (4)$$

为了从 (3) 式中找到矛盾, 我们令 $z_1 = z_2 = \dots = z_k = 0, y_{p+1} = \dots = y_n = 0$, 再利用 (4) 式, 得到 y_1, y_2, \dots, y_n 的方程组

$$\begin{cases} d_{11}y_1 + d_{12}y_2 + \dots + d_{1n}y_n = 0 \\ d_{21}y_1 + d_{22}y_2 + \dots + d_{2n}y_n = 0 \\ \dots\dots\dots \\ d_{k1}y_1 + d_{k2}y_2 + \dots + d_{kn}y_n = 0 \\ y_{p+1} = 0 \\ \dots\dots\dots \\ y_n = 0. \end{cases} \quad (5)$$

齐次线性方程组(5)有 n 个未知量,但方程个数 $= k + (n - p) = n - (p - k) < n$,故必有非零解.由于 $y_{p+1} = \dots = y_n = 0$,所以(5)的非零解中必是 y_1, y_2, \dots, y_p 不全为零,将其代入(3)式,得

$$f = y_1^2 + \dots + y_k^2 + y_{k+1}^2 + \dots + y_p^2 > 0. \quad (6)$$

将(5)的非零解代入(4)式得到 $z_1, z_2, \dots, z_k, \dots, z_n$ 的一组值(这时 $z_1 = z_2 = \dots = z_k = 0$)将它们再代入(3)式,又得

$$f = -z_{k+1}^2 - \dots - z_p^2 - z_{p+1}^2 - \dots - z_n^2 \leq 0, \quad (7)$$

(6)(7)二式显然是矛盾的,故假设的 $p > k$ 不能成立,即必有 $p \leq k$.

同理可证 $k \leq p$,因此必有 $p = k$.

从定理 7.16 可知,一个实对称矩阵 A 的相合规范形是唯一的,两个 n 阶实对称矩阵 A 和 B 相合的充分必要条件是它们的正、负惯性指数分别相等;

全体 n 阶实对称矩阵按相合关系可以划分为 $\frac{(n+1)(n+2)}{2}$ 个等价类.

7.9 正定二次型与正定矩阵 其它有定二次型

正定二次型与正定矩阵在工程技术和最优化等问题中有着广泛的应用,讨论多元函数极值的充分条件也要用到它.

定义 7.12 n 元实二次型 $f(x_1, x_2, \dots, x_n) = \mathbf{X}^T \mathbf{A} \mathbf{X}$ 称为正定二次型,如果 $\forall \mathbf{X} \neq \mathbf{0} (\mathbf{X} \in \mathbf{R}^n)$ 恒有 $\mathbf{X}^T \mathbf{A} \mathbf{X} > 0$; 正定二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 所对应的矩阵 A 叫做正定矩阵.

由定义可得以下两个基本结论:

(1) n 元实二次型(标准形)

$$f = (x_1, x_2, \dots, x_n) = d_1 x_1^2 + d_2 x_2^2 + \dots + d_n x_n^2$$

正定的充分必要条件是 $d_i > 0 (i = 1, 2, \dots, n)$.

充分性是显然的,其必要性可用反证法证明,设存在 $d_i \leq 0$,于是取

$x_i = 1, x_j = 0 (j \neq i)$, 便有

$$f(0 \dots 0, 1, 0 \dots 0) = d_i \leq 0,$$

这与二次型正定相矛盾.

(2) 对正定二次型 $f = X^T A X$ 作坐标变换 $X = CY$ (C 为可逆矩阵), 所得的二次型

$$f = Y^T (C^T A C) Y$$

也是正定的. 这是因为: $\forall Y_0 \neq 0$, 相应的 $X_0 = CY_0 \neq 0$ (如果 $X_0 = 0$, 则 $Y_0 = C^{-1}X_0 = 0$), 于是由 $f = X^T A X$ 的正定性, 即得

$$f = Y_0^T (C^T A C) Y_0 = X_0^T A X_0 > 0.$$

这表明, 对二次型作坐标变换, 其正定性不变.

由以上两个基本结论可见, 一个二次型的正定性可由其相合标准形(或规范形)来判定.

定理 7.17 对于 n 阶实对称矩阵 A , 下列命题等价:

(1) $X^T A X$ 是正定二次型(或 A 是正定矩阵);

(2) A 的正惯性指数为 n , 即 $A \simeq E$;

(3) 存在可逆矩阵 P , 使得 $A = P^T P$;

(4) A 的 n 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都大于零.

证 (1) \Rightarrow (2) 由上述两个基本结论即得正定矩阵 $A \simeq E$, 即对正定二次型 $X^T A X$ 作坐标变换所化成的相合规范形必为

$$X^T A X = y_1^2 + y_2^2 + \dots + y_n^2.$$

(2) \Rightarrow (3) 设 $A \simeq E$, 即存在可逆阵 C , 使得 $C^T A C = E$, 由此即得 $A = (C^T)^{-1} C^{-1}$, 令 $P = C^{-1}$, 则 $P^T = (C^T)^{-1}$, 于是 $A = P^T P$.

(3) \Rightarrow (4) 设 λ 是 A 的任一个特征值, 于是存在实向量 $X \neq 0$, 使得 $A X = \lambda X$. 由已知条件 $A = P^T P$, 得

$$(P^T P) X = \lambda X,$$

从而有

$$(X^T P^T) (P X) = \lambda (X^T X),$$

即

$$(P X, P X) = \lambda (X, X).$$

由 P 是可逆矩阵和 $X \neq 0$, 可得 $P X \neq 0$, 因此上式中的两个内积都大于零, 故 $\lambda > 0$.

(4) \Rightarrow (1) 对于 n 元实二次型 $X^T A X$, 存在正交变换 $X = Q Y$, 使得

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2.$$

于是,由 $\lambda_1, \dots, \lambda_n$ 都大于零即得 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 是正定二次型.

例 1 证明:若 A 是正定矩阵,则 A^{-1} 也是正定矩阵.

证 由定义 7.12 可知,正定矩阵是实对称矩阵,由定理 7.17 又知,正定矩阵 A 是可逆的,且 $(A^{-1})^T = (A^T)^{-1} = A^{-1}$,因此 A^{-1} 也是实对称矩阵.证明其正定性的方法很多.

方法 1 对二次型 $\mathbf{X}^T A^{-1} \mathbf{X}$ 作坐标变换 $\mathbf{X} = \mathbf{A} \mathbf{Y}$,得

$$\mathbf{X}^T A^{-1} \mathbf{X} = \mathbf{Y}^T A^T A^{-1} \mathbf{A} \mathbf{Y} = \mathbf{Y}^T \mathbf{A} \mathbf{Y}.$$

由 $\mathbf{Y}^T \mathbf{A} \mathbf{Y}$ 正定,可知 $\mathbf{X}^T A^{-1} \mathbf{X}$ 也正定,故 A^{-1} 是正定矩阵.

方法 2 由于 $A \simeq E$,即存在可逆阵 C 使得 $C^T A C = E$,将两边求逆,得 $(C^{-1}) A^{-1} (C^{-1})^T = E$,故 $A^{-1} \simeq E$,因此 A^{-1} 是正定的.

方法 3 由于 A 正定,所以存在可逆阵 P ,使得 $A = P^T P$,于是 $A^{-1} = P^{-1} (P^{-1})^T$,因此 A^{-1} 也正定.

方法 4 由于 A 的 n 个特征值都大于零,所以存在正交阵 Q ,使得

$$Q^{-1} A Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

从而

$$Q^{-1} A^{-1} Q = \text{diag}\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}\right).$$

因此 A^{-1} 的 n 个特征值 $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$ 也都大于零,故 A^{-1} 正定.

例 2 判断三元二次型

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2 + x_2 x_3$$

是否是正定二次型.

解法 1 二次型的对应矩阵

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

的特征多项式 $|\lambda E - A| = (\lambda - 1) \left((\lambda - 1)^2 - \frac{1}{2} \right)$,特征值 $\lambda_1 = 1, \lambda_2 = 1 + \frac{\sqrt{2}}{2}, \lambda_3 = 1 - \frac{\sqrt{2}}{2}$ 都大于零,所以二次型正定.

解法 2 用配方法得

$$f(x_1, x_2, x_3) = \left(x_1 - \frac{x_2}{2}\right)^2 + \frac{3}{4}\left(x_2 + \frac{2}{3}x_3\right)^2 + \frac{2}{3}x_3^2. \quad (1)$$

显然 $f(x_1, x_2, x_3) \geq 0$, 等号成立当且仅当

$$\left(x_1 - \frac{x_2}{2}\right) = 0, \quad \left(x_2 + \frac{2}{3}x_3\right) = 0, \quad x_3 = 0,$$

即 $x_1 = x_2 = x_3 = 0$, 故二次型正定. 此时也可由 (1) 式得二次型标准形

$$f = y_1^2 + \frac{3}{4}y_2^2 + \frac{2}{3}y_3^2,$$

从而判定 $f(x_1, x_2, x_3)$ 是正定的.

下面, 我们从二次型矩阵 A 的子式来判别二次型 $\mathbf{X}^T A \mathbf{X}$ 的正定性. 先给出 A 正定的两个必要条件, 再给一个充分必要条件.

定理 7.18 若 n 元二次型 $\mathbf{X}^T A \mathbf{X}$ 正定, 则

(1) A 的主对角元 $a_{ii} > 0 (i = 1, 2, \dots, n)$;

(2) A 的行列式 $\det A > 0$.

证 (1) 因为

$$\mathbf{X}^T A \mathbf{X} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

正定, 所以取 $\mathbf{X}_i = (0, \dots, 0, 1, 0, \dots, 0)$ (其中第 i 个分量 $x_i = 1$), 便必有 $\mathbf{X}_i^T A \mathbf{X}_i = a_{ii} x_i^2 = a_{ii} > 0 (i = 1, \dots, n)$.

(2) 因为 A 正定, 所以存在可逆矩阵 P , 使得 $A = P^T P$, 从而

$$|A| = |P^T| |P| = |P|^2 > 0.$$

或根据正定矩阵 A 的特征值都大于零, 得 $|A| = \lambda_1 \lambda_2 \dots \lambda_n > 0$.

根据定理 7.18, 容易判别

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

都不是正定矩阵, 因为 $|A| = -1 < 0$; B 有一个主对角元为 0.

定理 7.19 n 元二次型 $\mathbf{X}^T A \mathbf{X}$ 正定的充分必要条件为 A 的 n 个顺序主子式 (左上角主子式) 都大于零, 即

$$\det A_k = \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & & & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix} > 0, \quad k = 1, 2, \dots, n.$$

证 必要性. 取 $\mathbf{X}_k = (x_1 \cdots x_k)^T \neq \mathbf{0}$, $\mathbf{X} = (x_1 \cdots x_k \ 0 \cdots 0)^T = (\mathbf{X}_k^T \ \mathbf{0}_{n-k}^T)^T \neq \mathbf{0}$ 则有

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = (\mathbf{X}_k^T \ \mathbf{0}_{n-k}^T) \begin{pmatrix} A_k & * \\ * & * \end{pmatrix} \begin{pmatrix} \mathbf{X}_k \\ \mathbf{0}_{n-k} \end{pmatrix} = \mathbf{X}_k^T A_k \mathbf{X}_k > 0,$$

即对一切 $\mathbf{X}_k \neq \mathbf{0}$, 恒有 $\mathbf{X}_k^T A_k \mathbf{X}_k > 0$, 所以 $x_1 \cdots x_k$ 的 k 元二次型 $\mathbf{X}_k^T A_k \mathbf{X}_k$ 是正定的, 因此, 由定理 7.18 得 $\det A_k > 0$ ($k = 1 \cdots n$).

充分性. 对 n 用数学归纳法. 当 $n = 1$ 时, $a_{11} > 0$, $\mathbf{X}^T \mathbf{A} \mathbf{X} = a_{11} x_1^2 > 0$ ($\forall x_1 \neq 0$), 命题成立. 假设命题对 $n - 1$ 元二次型成立; 下面证明对 n 元二次型也成立. 将 A 分块表示为

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

其中 $\boldsymbol{\alpha} = (a_{1n} \ a_{2n} \ \cdots \ a_{n-1,n})^T$.

根据定理 7.17, 只需证明 $A \simeq E$. 取

$$C_1 = \begin{pmatrix} E_{n-1} & -A_{n-1}^{-1} \boldsymbol{\alpha} \\ \mathbf{0} & 1 \end{pmatrix}, \quad \text{则} \quad C_1^T = \begin{pmatrix} E_{n-1} & \mathbf{0} \\ -\boldsymbol{\alpha}^T A_{n-1}^{-1} & 1 \end{pmatrix}.$$

于是

$$C_1^T A C_1 = \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0} & a_{nn} - \boldsymbol{\alpha}^T A_{n-1}^{-1} \boldsymbol{\alpha} \end{pmatrix}.$$

记 $b = a_{nn} - \boldsymbol{\alpha}^T A_{n-1}^{-1} \boldsymbol{\alpha}$, 由于 $|A| > 0$, $|A_{n-1}| > 0$, 从上式即得 $b > 0$. 根据归纳假设, A_{n-1} 正定, 故存在 $n - 1$ 阶可逆矩阵 P , 使得 $P^T A_{n-1} P = E_{n-1}$, 因此再取

$$C_2 = \begin{pmatrix} P & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{b}} \end{pmatrix}, \quad C_2^T = \begin{pmatrix} P^T & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{b}} \end{pmatrix},$$

就有

$$C_2^T (C_1^T A C_1) C_2 = E, \quad \text{故} \quad A \simeq E.$$

用定理 7.19 容易判断例 2 的二次型是正定的, 因为二次型矩阵 A 的三个顺序主子式:

$$|A_1| = 1 > 0, \quad |A_2| = \frac{3}{4} > 0, \quad |A_3| = |A| = \frac{1}{2} > 0.$$

例 3 证明 若 A 是 n 阶正定矩阵, 则存在正定矩阵 B , 使得 $A = B^2$.

证 因为正定矩阵 A 是实对称矩阵, 所以存在正交阵 Q , 使得

$$A = Q(\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n))Q^T$$

其中 $\lambda_i > 0 (i = 1, 2, \dots, n)$, 于是存在正定矩阵

$$B = Q(\text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}))Q^T$$

使得 $A = B^2$. 这里的 B 通常记作 $A^{\frac{1}{2}}$.

例4 设 A 为 n 阶正定矩阵, $\mathbf{X} = (x_1, \dots, x_n)^T \in \mathbf{R}^n$, \mathbf{b} 是一固定的实 n 维列向量. 证明:

$$p(\mathbf{X}) = \frac{\mathbf{X}^T A \mathbf{X}}{2} - \mathbf{X}^T \mathbf{b}$$

在 $\mathbf{X}_0 = A^{-1}\mathbf{b}$ 处取得最小值, 且 $p_{\min} = -\frac{1}{2}\mathbf{b}^T A^{-1}\mathbf{b}$.

证 当 A 为一阶正定矩阵时, $A = (a)$, $a > 0$, $\mathbf{X} = (x)^T \in \mathbf{R}^1$, $p(x) = \frac{1}{2}ax^2 - bx$ 是一条抛物线, 它在 $x = \frac{b}{a}$ 处取得最小值 $p_{\min} = -\frac{b^2}{2a}$. 本例是把一元二次函数的最小值问题推广到 n 元二次函数(其二次项部分是正定二次型).

这里欲证 $p(\mathbf{X}_0)$ 是 $p(\mathbf{X})$ 的最小值, 只要证恒有 $p(\mathbf{X}) - p(\mathbf{X}_0) \geq 0$. 由于 $\mathbf{b} = A\mathbf{X}_0$, 所以

$$\begin{aligned} p(\mathbf{X}) - p(\mathbf{X}_0) &= \frac{1}{2}\mathbf{X}^T A \mathbf{X} - \mathbf{X}^T \mathbf{b} - \frac{1}{2}\mathbf{X}_0^T A \mathbf{X}_0 + \mathbf{X}_0^T \mathbf{b} \\ &= \frac{1}{2}\mathbf{X}^T A \mathbf{X} - \mathbf{X}^T A \mathbf{X}_0 + \frac{1}{2}\mathbf{X}_0^T A \mathbf{X}_0. \end{aligned} \quad (1)$$

又因为 $\mathbf{X}^T A \mathbf{X}_0$ 是一阶矩阵, 所以

$$\mathbf{X}^T A \mathbf{X}_0 = (\mathbf{X}^T A \mathbf{X}_0)^T = \mathbf{X}_0^T A \mathbf{X} = \frac{1}{2}\mathbf{X}_0^T A \mathbf{X} + \frac{1}{2}\mathbf{X}^T A \mathbf{X}_0. \quad (2)$$

将(2)式代入(1)式, 即得

$$p(\mathbf{X}) - p(\mathbf{X}_0) = \frac{1}{2}(\mathbf{X} - \mathbf{X}_0)^T A (\mathbf{X} - \mathbf{X}_0).$$

因此, 由 A 的正定性, 即得 $\forall (\mathbf{X} - \mathbf{X}_0) \neq \mathbf{0}$, 即 $\forall \mathbf{X} \neq \mathbf{X}_0$, 恒有 $p(\mathbf{X}) - p(\mathbf{X}_0) > 0$. 故 $p(\mathbf{X}_0)$ 是 $p(\mathbf{X})$ 的最小值, 且

$$\begin{aligned} p_{\min} = p(\mathbf{X}_0) &= \frac{1}{2}\mathbf{X}_0^T A \mathbf{X}_0 - \mathbf{X}_0^T \mathbf{b} = -\frac{1}{2}\mathbf{X}_0^T \mathbf{b} \\ &= -\frac{1}{2}(\mathbf{A}^{-1}\mathbf{b})^T \mathbf{b} = -\frac{1}{2}\mathbf{b}^T \mathbf{A}^{-1}\mathbf{b}. \end{aligned}$$

最后, 我们再简要地介绍一下负定二次型, 半正定、半负定二次型以及不定二次型.

定义 7.12 如果 n 元实二次型 X^TAX 满足 $\forall X \neq 0$, 恒有:

- (1) $X^TAX < 0$ 则称之为负定二次型 相应地称 A 为负定矩阵;
 (2) $X^TAX \geq 0$ 则称之为半正定二次型 相应地称 A 为半正定矩阵;
 (3) $X^TAX \leq 0$ 则称之为半负定二次型 相应地称 A 为半负定矩阵.
 不是正定、半正定、负定、半负定的二次型叫做不定二次型.

根据定义, 同样可以证明:

对二次型作坐标变换, 其负定性、半正定性、半负定性及不定性都不变.

n 元二次型 X^TAX : 负定的充要条件是 A 的负惯性指数等于 n ; 半正定的充要条件是 A 的正惯性指数等于 $r(A)$ (当 $r(A) < n$ 时, 存在 $X \neq 0$, 使 $X^TAX = 0$); 半负定的充要条件是 A 的负惯性指数等于 $r(A)$ (当 $r(A) < n$ 时, 存在 $X \neq 0$, 使 $X^TAX = 0$). 如此, 读者也不难写出如定理 7.17 那样的有关负定、半正定、半负定的其它等价命题.

A 负定与 $(-A)$ 正定是等价的. 所以实对称矩阵 A 负定的充要条件是 A 的奇数阶顺序主子式都小于零, A 的偶数阶顺序主子式都大于零.

必须注意, 实对称矩阵 A 的各阶顺序主子式 ≥ 0 不是 A 半正定的充分必要条件, 它的充分必要条件应是 A 的各阶主子式 ≥ 0 (证明留给有兴趣的读者自己去完成).

习 题

1. 求 a, b, c, d, e , 使 Q 为正交矩阵

$$Q = \begin{pmatrix} a & -\frac{3}{7} & \frac{2}{7} \\ b & c & d \\ -\frac{3}{7} & \frac{2}{7} & e \end{pmatrix}.$$

2. 证明: 任一个方阵如果有三个性质(对称阵、正交阵、对合阵即 $A^2 = E$) 中的任两个性质, 则必有第三个性质.

3. 验证下列矩阵是对称阵、正交阵和对合阵:

$$(1) \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}; \quad (2) \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

4. 证明 若 A 是正交矩阵, 则其伴随矩阵 A^* 也是正交矩阵.

5. 证明 (1) 若 $\det A = 1$, 则 A 为正交矩阵 $\Leftrightarrow A$ 的每个元素等于自己的代数余子式;

(2) 若 $\det A = -1$, 则 A 为正交矩阵 $\Leftrightarrow A$ 的每个元素等于其代数余子式乘 (-1) .

6. 证明 如果三角矩阵 A 是正交矩阵, 则 A 必是主对角元为 1 或 (-1) 的对角矩阵.

7*. 设 A 为正交矩阵, 已知 $E + A$ 可逆, 证明:

(1) $(E - A)(E + A)^{-1}$ 可交换;

(2) $(E - A)(E + A)^{-1}$ 为反对称矩阵.

8*. 证明 欧氏空间的一组单位正交基变为另一组单位正交基的变换矩阵是正交矩阵.

9. 利用转轴与移轴, 化简下列二次曲线的方程, 并画出它们的图形.

(1) $5x^2 + 4xy + 2y^2 - 24x - 12y + 18 = 0$;

(2) $x^2 + 2xy + y^2 - 4x + y - 1 = 0$;

(3) $5x^2 + 12xy - 22x - 12y - 19 = 0$;

(4) $x^2 + 2xy + y^2 + 2x + 2y = 0$.

10. 利用不变量与半不变量, 判断下列二次曲线的类型, 并求化简后的方程和标准方程.

(1) $x^2 + 6xy + y^2 + 6x + 2y - 1 = 0$;

(2) $3x^2 - 2xy + 3y^2 + 4x + 4y - 4 = 0$;

(3) $x^2 - 4xy + 3y^2 + 2x - 2y = 0$;

(4) $x^2 - 4xy + 4y^2 + 2x - 2y - 1 = 0$;

(5) $x^2 - 2xy + 2y^2 - 4x - 6y + 29 = 0$.

11. 已知 \mathbf{R}^3 的一个线性变换

$$\sigma(x_1, x_2, x_3) = (2x_1 - 2x_2, -2x_1 + x_2 - 2x_3, -2x_2).$$

(1) 求 σ 关于自然基 $\{e_1, e_2, e_3\}$ 所对应的矩阵 A ;

(2) 求 σ 关于基 $\{(1, 1, 1), (0, 1, 1), (0, 0, 1)\}$ 所对应的矩阵 B ;

(3) 求矩阵 C_1 , 使 $C_1^{-1}BC_1 = A$.

12. 已知三维线性空间 V 的线性变换 σ 关于基 $\{\alpha_1, \alpha_2, \alpha_3\}$ 所对应的矩阵为

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

(1) 求 σ 在基 $\{\beta_1, \beta_2, \beta_3\}$ 下对应的矩阵 B , 其中:

$$\beta_1 = 2\alpha_1 + \alpha_2 + 3\alpha_3,$$

$$\beta_2 = \alpha_1 + \alpha_2 + 2\alpha_3,$$

$$\beta_3 = -\alpha_1 + \alpha_2 + \alpha_3;$$

(2) 求 σ 的值域 $\sigma(V)$ 和核 $\text{Ker}\sigma$;

(3) 把 $\sigma(V)$ 的基扩充为 V 的基, 并求 σ 在这个基下对应的矩阵;

(4) 把 $\text{Ker}\sigma$ 的基扩充为 V 的基, 并求 σ 在这个基下对应的矩阵.

13. 设 $A, B \in M_n(F)$, 证明: 若 A 可逆, 则 $AB \sim BA$.

14. 设 $A \sim B, C \sim D$, 证明:

$$\begin{pmatrix} A & O \\ O & C \end{pmatrix} \sim \begin{pmatrix} B & O \\ O & D \end{pmatrix}.$$

15. 设 $A \sim B, f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$. 证明: $f(A) \sim f(B)$.

16. 设 $V(C)$ 是一个线性空间, $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 V 的基, $\sigma \in L(V, V)$, 已知 σ 在基 B 下对应的矩阵 A 如下, 试求 σ 的特征值与特征向量 (只要求特征子空间的基):

$$(1) \begin{pmatrix} 2 & -3 \\ -3 & 1 \end{pmatrix};$$

$$(2) \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix};$$

$$(3) \begin{pmatrix} 4 & 5 & -2 \\ -2 & -2 & 1 \\ -1 & -1 & 1 \end{pmatrix};$$

$$(4) \begin{pmatrix} 0 & 4 & 0 \\ -4 & 0 & 3 \\ 0 & -3 & 0 \end{pmatrix};$$

$$(5) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix};$$

$$(6) \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

17. 设

$$A = \begin{pmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & x \end{pmatrix}.$$

试求 x 的值, 使 $\lambda_1 = 3$ 是 A 的二重特征值, 并求另一个特征值 λ_2 及两个特征子空间的基.

18. 对下列矩阵 A 的特征值, 能做出怎样的断言?

(1) A 可逆阵;

(2) A 不可逆;

- (3) $E + A$ 可逆; (4) $4E + A$ 不可逆;
 (5) $\det(E - A^2) = 0$; (6) $A^2 = E$;
 (7) $A^2 = A$ (幂等矩阵); (8) $A^k = O$ (幂零矩阵);
 (9) $A = \lambda_0 E + B$ (λ_0 为常数, 且已知 B 的 n 个特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$);
 (10) A 对角块矩阵, 即

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}.$$

19. 设 $A, B \in M_n(F)$; $AB = BA$. 证明: 若 X 是矩阵 A 属于特征值 λ_0 的特征向量, 则 $BX \in V_{\lambda_0}(A)$ 的特征子空间, 即当 $BX \neq 0$ 时, BX 也是 A 属于 λ_0 的特征向量).

20*. 设 $A, B \in M_n(F)$; A 有 n 个不同的特征值. 证明: $AB = BA \Leftrightarrow A$ 的特征向量也是 B 的特征向量.

21. 下列线性变换 $\sigma \in L(V, V)$ 可否对角化? 如果可以, 试求 V 的一组基, 使 σ 在这组基下对应的矩阵为对角阵, 并给出这个对角阵.

- (1) 第 11 题的 σ ; (2) 第 12 题的 σ ;
 (3) 第 1(4) 题的 σ ; (4) 第 1(5) 题的 σ ;
 (5) 第 1(6) 题的 σ .

22. 设 $A = (a_{ij})_{n \times n}$ 是上三角矩阵, 证明:

- (1) 如果主对角元互不相等, 则 A 与对角阵相似;
 (2) 如果 n 个主对角元相等, 且至少有一个元素 $a_{ij} \neq 0 (i < j)$, 则 A 不与对角阵相似;

(3) $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ 不可对角化.

23. 已知三阶矩阵 A 的特征值 $\lambda_1 = 1$ (二重), $\lambda_2 = -2$, 属于 λ_1 的特征向量有 $X_1 = (1, -1, 0)^T$, $X_2 = (1, 0, -1)^T$, 属于 λ_2 的特征向量有 $X_3 = (1, 1, 1)^T$. 问 A 可否对角化? 如能对角化, 求出 A 及 A^k (k 为正整数).

24. 设

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

试求 A^n (n 为正整数).

25*. 设 $\alpha = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, 其中 $a_i \neq 0$ ($i = 1, \dots, n$). 证明: 矩阵 $A = \alpha^T \alpha$ 可对角化, 并求可逆阵 P 及对角阵 Λ , 使 $P^{-1}AP = \Lambda$.

26. 对下列实对称矩阵 A , 求正交矩阵 Q 和对角阵 Λ , 使 $Q^{-1}AQ = \Lambda$.

$$(1) \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & -1 \end{pmatrix};$$

$$(3) \begin{pmatrix} 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & 4 \\ 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \end{pmatrix}; \quad (4) \begin{pmatrix} 1 & 3 & -3 & 3 \\ 3 & 1 & 3 & -3 \\ -3 & 3 & 1 & 3 \\ 3 & -3 & 3 & 1 \end{pmatrix}.$$

27. 设 A 是 n 阶实对称矩阵, 且 $A^2 = A$. 证明: 存在正交阵 Q , 使得

$$Q^{-1}AQ = \text{diag}(1, \dots, 1, 0, \dots, 0),$$

其中 1 的个数等于 $r(A)$.

28. 设 A 为 n 阶实对称幂等矩阵, $r(A) = r$, 求 $\det(A - 2E)$.

29. 设 n 阶实对称矩阵 A 的特征值 $\lambda_i \geq 0$ ($i = 1, \dots, n$). 证明: 存在特征值都是非负数的实对称矩阵 B (即半正定矩阵), 使得 $A = B^2$.

30. 证明: 反对称实矩阵的特征值 λ 必是零或纯虚数.

31*. 已知 A 是反对称实矩阵, 证明 $E - A^2$ 是可逆矩阵.

32*. 已知 $\alpha = (a_1, \dots, a_n), \beta = (b_1, \dots, b_n)$ 是 \mathbb{R}^n 中两个非零的正交向量, 证明: 矩阵 $A = \alpha^T \beta$ 的特征值全为零, 且 A 不可对角化.

33. 证明: 若二次型 $f(x_1, \dots, x_n) = X^T A X$ 对一切 $X = (x_1, \dots, x_n)^T$ 恒有 $f(x_1, \dots, x_n) = 0$, 则 $A = 0$.

34. 设二次型 $f(x_1, x_2, x_3) = X^T A X$, $g(x_1, x_2, x_3) = X^T B X$. 证明: 若 $f(x_1, x_2, x_3) = g(x_1, x_2, x_3)$, 则 $A = B$.

35. 设 $A \simeq B, C \simeq D$, 且它们都是 n 阶实对称矩阵, 下列结论成立吗?

$$(1) (A + C) \simeq (B + D);$$

$$(2) \begin{pmatrix} A & O \\ O & C \end{pmatrix} \simeq \begin{pmatrix} B & O \\ O & D \end{pmatrix}.$$

36. 用正交变换 $X = QY$ 将下列二次型化为标准形, 并给出 Q .

$$(1) f = 2x_1^2 + 3x_2^2 + 3x_3^2 + 4x_2x_3;$$

$$(2) f = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 - 2x_1x_4 - 2x_2x_3 - 2x_3x_4;$$

$$(3) f = x_1^2 - x_2^2 + x_3^2 - 2x_1x_3;$$

$$(4) f = 2x_1x_4 + 2x_2x_3.$$

37. 用配方法将下列二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 化为标准形, 并给出坐标变换 $\mathbf{X} = \mathbf{C} \mathbf{Y}$ 的变换矩阵 \mathbf{C} .

$$(1) A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 4 \end{pmatrix}; \quad (2) A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix};$$

$$(3) A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}; \quad (4) A = \begin{pmatrix} 1 & 3 & -2 & 0 \\ 3 & 7 & 0 & -2 \\ -2 & 0 & -4 & -4 \\ 0 & -2 & -4 & -1 \end{pmatrix}.$$

38. 用初等变换法将题 37 的二次型化为标准形, 并给出坐标变换 $\mathbf{X} = \mathbf{C} \mathbf{Y}$ 的变换矩阵 \mathbf{C} .

39. 证明 秩为 r 的 n 阶实对称矩阵 \mathbf{A} 可以表示为秩为 1 的 r 个实对称矩阵之和.

40. 已知 n 阶实对称幂等矩阵 \mathbf{A} 的秩为 r , 试求:

(1) 二次型 $\mathbf{X}^T \mathbf{A} \mathbf{X}$ 的一个标准形;

(2) $\text{de}(E + A + A^2 + \dots + A^k)$.

41. 判断题 36 中的二次型哪些是正定的.

42. 判断题 37 中的实对称矩阵哪些是正定的.

43. 求下列二次型中的参数 t , 使二次型正定:

(1) $5x_1^2 + x_2^2 + tx_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$;

(2) $2x_1^2 + x_2^2 + 3x_3^2 + 2tx_1x_2 + 2x_1x_3$.

44. 设二次型

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j.$$

(1) 用正交变换法将其化为标准形, 并判断它是否正定;

(2) 当 $n = 3$ 时, 求正定矩阵 \mathbf{B} , 使得 $\mathbf{A} = \mathbf{B}^2$.

45. 用定义证明: 正定矩阵的特征值大于零.

46. 用定义证明: 若 \mathbf{P} 是 n 阶实矩阵, 则 $\mathbf{P}^T \mathbf{P}$ 是半正定矩阵.

47. 设 \mathbf{A} 是半正定矩阵, \mathbf{C} 是可逆矩阵, 证明: $\mathbf{C}^T \mathbf{A} \mathbf{C}$ 也是半正定矩阵.

48. 设 \mathbf{A}, \mathbf{B} 皆是 n 阶正定矩阵, k, l 都是正数, 用定义证明 $k\mathbf{A} + l\mathbf{B}$ 也是正定矩阵.

49*. 设 \mathbf{A} 是 n 阶实对称矩阵, $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 \mathbf{A} 的 n 个互异的特征值,

X_1 (列向量) 是对应于 λ_1 的单位特征向量. 证明: $A - \lambda_1 X_1 X_1^T$ 的特征值为 $0, \lambda_2, \dots, \lambda_n$.

50. 设 n 阶实对称矩阵 A 的 n 个特征值为 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. 问 t 满足什么条件时,

(1) $A + tE$ 为正定矩阵; (2) $A - tE$ 为正定矩阵.

51*. 设 A 是实对称矩阵, B 是正定矩阵. 证明: 存在可逆阵 C , 使得 $C^T A C$ 和 $C^T B C$ 都成对角形.

52*. 证明: 若对 n 元实二次型 $X^T A X$ 有 X_1, X_2 使得 $X_1^T A X_1 > 0$, $X_2^T A X_2 < 0$, 则存在 $X_0 \neq 0$, 使得 $X_0^T A X_0 = 0$.

53*. 设 A 是 n 阶正定矩阵, $X = (x_1, x_2, \dots, x_n)^T$, 证明:

$$f(x) = \det \begin{pmatrix} 0 & X^T \\ X & A \end{pmatrix}$$

是一个负定二次型.

54*. 设 $A = (a_{ij})_{n \times n}$ 是正定矩阵, 证明:

(1) $|A| \leq a_{nn} |A_{n-1}|$, 其中 $|A_{n-1}|$ 是 A 的左上角 $n-1$ 阶主子式;

(2) $|A| \leq a_{11} a_{22} \dots a_{nn}$ (利用 (1) 的结果).

补 充 题

1. 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是矩阵 $A = (a_{ij})_{n \times n}$ 的 n 个特征值. 证明 $\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2$ 是 A^2 的 n 个特征值, 且

$$\sum_{i=1}^n \lambda_i^2 = \sum_{i=1}^n \sum_{k=1}^n a_{ik} a_{ki}.$$

2. 设 $A, B \in M_n(C)$, B 的特征多项式 $f(\lambda) = |\lambda E - B|$. 证明: $f(A)$ 可逆的充要条件是 B 的任一特征值都不是 A 的特征值.

3. 设 $V(F)$ 是 n 维线性空间, $\sigma \in L(V, V)$, 证明:

(1) 若 α, β 是 σ 的属于不同特征值的特征向量, 则当 $c_1 c_2 \neq 0$ 时, $c_1 \alpha + c_2 \beta$ 不是 σ 的特征向量;

(2) V 中每一非零向量都是 σ 的特征向量 $\Leftrightarrow \sigma = c_0 I_V$, 其中 $c_0 \in F$ 是一个常数, I_V 是恒等变换.

4. 证明: 对任一个 n 阶矩阵 $A \in M_n(C)$, 存在可逆阵 P , 使 $P^{-1} A P$ 为上三角阵.

5*. 设 A 相似于对角阵, λ_0 是 A 的特征值, X_0 是 A 对应于 λ_0 的特征向量. 证明:

$$(1) \text{秩}(A - \lambda_0 E) = \text{秩}(A - \lambda_0 E);$$

$$(2) \text{不存在 } Y, \text{使 } (A - \lambda_0 E)Y = X_0.$$

6. 证明 (1) 正交矩阵的实特征值为 ± 1 ; (2) 正交矩阵的复特征值的模为 1.

7. 证明: 在奇数维欧氏空间中的第一类正交变换必有一个特征值为 1.

8. 证明: 第二类正交变换必有 (-1) 为其特征值.

9. 设 σ 是欧氏空间 V 的一个变换. 证明: 如果 $\forall \alpha, \beta \in V (\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$, 则 σ 是线性变换, 即 $\forall \lambda, \mu \in \mathbf{R}$, 必有 $\sigma(\lambda\alpha + \mu\beta) = \lambda\sigma(\alpha) + \mu\sigma(\beta)$, 从而 σ 为正交变换.

10. 设 $A = (a_{ij})_{n \times m}$ 是实矩阵. 证明:

$$\det(A^T A) \leq \prod_{k=1}^m \sum_{i=1}^n a_{ik}^2.$$

11. 证明: σ 是 n 维欧氏空间 V 的对称变换的充要条件是 σ 关于 V 的单位正交基 $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 所对应的矩阵为 n 阶实对称矩阵.

12. 设 $A \in M_{m \times n}(C), B \in M_{n \times m}(C)$. 证明:

$$\begin{pmatrix} AB & O_1 \\ B & O_2 \end{pmatrix} \sim \begin{pmatrix} O_3 & O_1 \\ B & BA \end{pmatrix},$$

其中 O_1, O_2, O_3 分别是 $m \times n, m \times n, m \times m$ 的零矩阵. 并由此推出 AB 与 BA 的非零特征值相同. 如果 $m = n$, 则 $|\lambda E - AB| = |\lambda E - BA|$.

13*. 证明: 若 $AB = BA$, 则 A, B 至少有一个共同的特征向量.

14. 设 A 为 n 阶实对称矩阵, A 的 n 个特征值 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. 证明: $\forall X \in \mathbf{R}^n$,

$$\lambda_1(X, X) \leq (AX, X) \leq \lambda_n(X, X)$$

并指出分别取怎样的非零向量 X 使两个等号成立.

15. 设 A, B 为 n 阶正定矩阵. 证明: $A + B$ 的最大特征值大于 A 的最大特征值.

16. 设 λ_1 和 μ_1 分别是 n 阶实对称矩阵 A 和 B 的最小特征值. 证明: $A + B$ 的最小特征值大于或等于 $\lambda_1 + \mu_1$.

17*. 设 A, B 皆是正定矩阵. 证明: 若 $AB = BA$, 则 AB 也是正定矩阵.

18. 生物外部的某种特征由其内部的两个基因(A, a)组成的基因对 AA, Aa, aa 所确定. 例如, 某种花的三种颜色被三种基因对所确定. 常染色体

的遗传规律是,母体双方各自的两个基因等可能地遗传给后代一个.因此,母体(双方)基因型与后代基因型的关系如下表:

后代基因型		母体(双方)基因型					
		AA	AA - Aa	AA - aa	Aa - Aa	Aa - aa	aa - aa
AA	AA	1	1/2	0	1/4	0	0
Aa	Aa	0	1/2	1	1/2	1/2	0
aa	aa	0	0	0	1/4	1/2	1

例如,第4列表示母体皆为 Aa 型时,其后代为 AA, Aa, aa 型的可能性分别为 $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$.

(1) 设某种植物有三种基因型 AA, Aa, aa, 它们各占总数的百分数为 a_0, b_0, c_0 ($a_0 + b_0 + c_0 = 1$). 如果它们总是都只与 Aa 型结合而进行繁殖,问繁殖到第 n 代时,三种基因型植物占总数的百分数 a_n, b_n, c_n 各为多少? 并求其极限值;

(2) 在(1)中的植物,如果初始时都与 AA 型结合,其第一代都与 Aa 型结合,第二代又都与 AA 型结合,如此交替繁殖下去,求 a_n, b_n, c_n 及其极限值.

部分习题和补充题答案

习题

$$1. -\frac{6}{7}, \frac{2}{7}, -\frac{6}{7}, -\frac{3}{7}, -\frac{6}{7}.$$

$$5. \text{提示: 利用 } A^{-1} = A^T \text{ 并用伴随矩阵表示 } A^{-1}.$$

$$7. \text{提示: 利用 } E = A^T A \text{ 及(1)的结论.}$$

$$9. (1) 6x_2^2 + y_2^2 - 12 = 0; (2) 2\sqrt{2}x_2^2 + 5y_2 - \frac{25}{16}\sqrt{2} = 0;$$

$$(3) 9x_2^2 - 4y_2^2 - 36 = 0; (4) 2x_2^2 - 1 = 0.$$

$$10. (1) \text{双曲线 } 4x^2 - 2y^2 - 2 = 0, 2x^2 - y^2 = 1;$$

$$(2) \text{椭圆 } 2x^2 + 4y^2 - 8 = 0, \frac{x^2}{4} + \frac{y^2}{2} = 1;$$

$$(3) \text{两相交直线 } (2 + \sqrt{5})x_2^2 + (2 - \sqrt{5})y_2^2 = 0;$$

$$(4) \text{抛物线 } 5y^2 - \frac{2\sqrt{5}}{5}x = 0; y^2 = \frac{2\sqrt{5}}{25}x;$$

$$(5) \text{点(或称相交于一实点的两条虚直线)} \frac{(3 + \sqrt{5})x^2}{2} + \frac{(3 - \sqrt{5})y^2}{2} = 0.$$

$$11. (1) \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}; (2) \begin{pmatrix} 0 & -2 & 0 \\ -3 & 1 & -2 \\ 1 & -1 & 2 \end{pmatrix};$$

$$(3) \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

$$12. (1) \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ 3 & 1 & -1 \end{pmatrix};$$

$$(2) L(2\alpha_1 + \alpha_2, -\alpha_1 + \alpha_3), L(\alpha_1 - 2\alpha_2 - 3\alpha_3);$$

$$(3) \{\alpha_1, 2\alpha_1 + \alpha_2, -\alpha_1 + \alpha_3\}; \begin{pmatrix} 0 & 0 & 0 \\ 2 & 5 & -2 \\ 3 & 6 & -2 \end{pmatrix};$$

$$(4) \{\alpha_1, \alpha_2, \alpha_1 - 2\alpha_2 - 3\alpha_3\}; \begin{pmatrix} 2 & 2 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

$$16. (1) \frac{3 \pm \sqrt{37}}{2} \alpha_1 + \frac{1 \mp \sqrt{37}}{6} \alpha_2; (2) \pm a i, \pm i \alpha_1 + \alpha_2;$$

$$(3) 1; -\alpha_1 + \alpha_2 + \alpha_3; (4) 0, \frac{3}{4} \alpha_1 + \alpha_3; \pm 5i, -\frac{4}{3} \alpha_1 + \frac{5}{3} i \alpha_2 + \alpha_3;$$

$$(5) -2, -\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4; 2, \alpha_1 + \alpha_2, \alpha_1 + \alpha_3, \alpha_1 + \alpha_4;$$

$$(6) 1, -\alpha_2 + \alpha_3, -\alpha_1 + \alpha_4; -1, \alpha_2 + \alpha_3, \alpha_1 + \alpha_4.$$

$$17. 4, 12; \{(-1, 1, 0)^T, (1, 0, 4)^T\}; \{(1, 1, -1)^T\}.$$

$$18. (1) \lambda \neq 0; (2) \exists \lambda_1 = 0; (3) \lambda \neq -1; (4) \exists \lambda = -4;$$

$$(5) \exists \lambda = 1 \text{ 或 } -1; (6) \lambda^2 = 1; (7) \lambda^2 = \lambda; (8) \lambda = 0;$$

$$(9) \lambda_A = \lambda_0 + \lambda_i (i = 1, \dots, m); (10) \lambda_A = \{\lambda_{A_1}, \dots, \lambda_{A_m}\}.$$

20. 提示 利用上题结果以及 A 的每个特征子空间是一维的, A 可对角化.

$$21. (1) \{(-2, -1, 2)^T, (2, -2, 1)^T, (1, 2, 2)^T\}, \text{diag}(1, 4, -2);$$

$$(2) \{(-1, 2, 3)^T, (0, 1, 2)^T, (1, 2, 3)^T\}, \text{diag}(0, 1, 2);$$

(3) 可对角化;

$$(4) \{(1, 1, 0, 0)^T, (1, 0, 1, 0)^T, (1, 0, 0, 1)^T, (-1, 1, 1, 1)^T\}, \\ \text{diag}(2, 2, 2, -2);$$

$$(5) \{(0, -1, 1, 0)^T, (-1, 0, 0, 1)^T, (0, 1, 1, 0)^T, (1, 0, 0, 1)^T\}, \\ \text{diag}(1, 1, -1, -1).$$

本题(2)(4)(5)的基向量,都是它们在 B 下的坐标向量.

$$23. \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 2+(-2)^k & -1+(-2)^k & -1+(-2)^k \\ -1+(-2)^k & 2+(-2)^k & -1+(-2)^k \\ -1+(-2)^k & -1+(-2)^k & 2+(-2)^k \end{pmatrix}.$$

$$24. \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix}^n A_1^n = \begin{pmatrix} 4 \cdot 5^{n-1} - (-5)^{n-1} & 2 \cdot 5^{n-1} + 2(-5)^{n-1} \\ 2 \cdot 5^{n-1} + 2(-5)^{n-1} & 5^{n-1} - 4(-5)^{n-1} \end{pmatrix};$$

$$A_2^n = \begin{pmatrix} 2^n & n2^{n+1} \\ 0 & 2^n \end{pmatrix}. \quad \text{提示 按对角块矩阵求 } A^n.$$

25. 提示 利用 $A^2 = kA$ ($k = \sum a_i^2$) 求 A 的特征值.

$$26. (1) \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{5}} & \frac{2}{3} \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & 0 & \frac{2}{3} \end{pmatrix}, \begin{pmatrix} -1 & & \\ & -1 & \\ & & 8 \end{pmatrix};$$

$$(2) \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix}, \begin{pmatrix} 1 & & \\ & 3 & \\ & & -3 \end{pmatrix};$$

$$(3) \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 5 & & & \\ & -5 & & \\ & & 3 & \\ & & & -3 \end{pmatrix};$$

$$(4) \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{2\sqrt{3}} & \frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2} \\ 0 & 0 & \frac{3}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix} \text{diag}(4, 4, 4, -8).$$

28. 提示 利用上题结果及 $E = Q^{-1}Q$.

30. 提示 只要证明 $\bar{\lambda} = -\lambda$.

31. 提示 利用上题结果, 证明 $|E - A^2| \neq 0$.

32. 提示 由 $A^2 = 0$ 考虑 A 的特征值 利用 $\chi(A) = 1$ 证明 A 不可对角化.

34. 提示 利用上题结果.

35. (1) 不成立; (2) 成立.

$$36. (1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad 2y_1^2 + y_2^2 + 5y_3^2;$$

$$(2) \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \end{pmatrix},$$

$$(1 + \sqrt{2})y_1^2 + y_2^2 + (1 - \sqrt{2})y_3^2 + y_4^2.$$

$$(3) \begin{pmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}; \quad (4) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

$$37. (1) \begin{pmatrix} 1 & -1 & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & 1 & 3 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix};$$

$$2y_1^2 + 3y_2^2 + \frac{2}{3}y_3^2$$

$$(3) \begin{pmatrix} 1 & 1 & \frac{-1}{2} & \frac{-1}{2} \\ 1 & -1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{-1}{2} \end{pmatrix}; \quad (4) \begin{pmatrix} 1 & -3 & -7 & -4 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$y_1^2 - 2y_2^2 + 10y_3^2 - 9y_4^2$$

$$2y_1^2 - 2y_2^2 + \frac{1}{2}y_3^2 - \frac{1}{2}y_4^2$$

$$38. (1) \begin{pmatrix} 1 & -1 & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix};$$

$$2y_1^2 + 3y_2^2 + \frac{2}{3}y_3^2 \quad y_1^2 - \frac{1}{4}y_2^2 + 3y_3^2$$

$$(3) \begin{pmatrix} 1 & -\frac{1}{2} & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}; \quad (4) \text{同} 37(4).$$

$$2y_1^2 - \frac{1}{2}y_2^2 + 2y_3^2 - \frac{1}{2}y_4^2$$

39. 提示 若 $C^T A C = B$ 且 C 为可逆矩阵 则 $r(A) = r(B)$.

$$40. (1) \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{n \times n}; \quad (2) (1+k)^r.$$

41. (1) 正定; 42. (1) 正定; 43. (1) $t > 2$; (2) $-\frac{5}{3} < t^2 < \frac{5}{3}$.

44. (1) $\frac{1}{2}(y_1^2 + \dots + y_{n-1}^2) + \frac{n+1}{2}y_n^2$ 是;

$$(2) B = \frac{1}{3\sqrt{2}} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}.$$

46. 提示 证明 $\forall X \neq 0, X^T(P^T P)X \geq 0$.

49. 提示 利用定理 7.12 的结论.

50. (1) $t > -\lambda_1$; (2) $t < \lambda_1$.

51. 提示 对于 B 存在 C_1 使 $C_1^T B C_1 = E$ 再对实对称矩阵 $C_1^T A C_1$ 作正交变换使之相合于对角阵.

52. 提示 这是一个不定二次型 利用其标准形 即 $C^T A C = \text{diag}(1, -1, \dots, *)$ 其中 $*$ 为 ± 1 或 0 .

53. 提示 对分块矩阵作初等行变换 将其化为上(下)三角块矩阵 即可展开这个行列式.

54. 提示 : 将 $|A|$ 分块表示为 $\begin{vmatrix} A_{n-1} & \alpha \\ \alpha^T & a_{nn} \end{vmatrix}$, 再展开这个行列式, 并利用 A_{n-1} 的正定性.

补充题

2. 提示 : 设 B 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 即 $f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$, 再利用 $f(A)$ 可逆 $\Leftrightarrow |f(A)| \neq 0$.

4. 提示 : 用数学归纳法.

7. 提示 : 实系数多项式 $f(x)$ 若有 $f(z) = 0$, 则必有 $f(\bar{z}) = 0$.

14. 提示 : 对二次型 (AX, X) 作正交变换, 将其化为标准形.

15. 提示 : 利用上题结果.

第 8 章 常见曲面及二次曲面的分类

本章主要讨论三维几何空间中常见的曲面,包括球面、柱面、锥面、旋转面、椭球面、双曲面和抛物面以及它们在空间直角坐标系下的方程表示(其中也涉及空间曲线,如螺旋线的方程表示),并对二次曲面方程进行分类.

8.1 球面 柱面 锥面 旋转面

我们知道,空间直角坐标系中,一个平面上的所有点的坐标 (x, y, z) 满足一个三元一次方程;反之,坐标满足一个三元一次方程的那些点形成一个平面.而一般的空间曲面在直角坐标系中,对应于非线性的三元方程 $F(x, y, z) = 0$,因为一个曲面 S 上的点的要满足一定的几何条件,这个条件一般可以写成点的坐标 (x, y, z) 所满足的一个方程 $F(x, y, z) = 0$,曲面 S 上的点的坐标一定满足这个方程,而坐标满足这个方程的点也一定在曲面 S 上.所以 $F(x, y, z) = 0$ 是曲面 S 的方程,曲面 S 是这个方程的图形.

直线作为两个平面的交线可以用两个平面的方程的联立方程组来表示.同样,两个相交的曲面一般相交于一条空间曲线 L .它可以用两个曲面方程的联立方程组

$$\begin{cases} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0 \end{cases} \quad (8-1)$$

来表示,这个方程组叫做曲线 L 的一般方程.

知道了方程的图形,就可用图形的几何性质讨论方程的一些问题;知道了图形的方程,就能够从方程得到图形的几何性质.

8.1.1 球面

设球面的球心为 $P_0(x_0, y_0, z_0)$,半径为 r .因此,点 $P(x, y, z)$ 在球面上,当且仅当 $|P_0P| = r$,即 $|P_0P|^2 = r^2$,这个条件用点的坐标可表示为

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2, \quad (8-2)$$

这就是该球面的方程.

把这个方程展开,得

$$x^2 + y^2 + z^2 - 2x_0x - 2y_0y - 2z_0z + (x_0^2 + y_0^2 + z_0^2 - r^2) = 0.$$

由此可见,球面方程是一个特殊的三元二次方程,它的一般形式为

$$x^2 + y^2 + z^2 + 2a_0x + 2b_0y + 2c_0z + d = 0,$$

用配方法可将其化为

$$(x + a_0)^2 + (y + b_0)^2 + (z + c_0)^2 = a_0^2 + b_0^2 + c_0^2 - d.$$

令 $a_0^2 + b_0^2 + c_0^2 - d = r^2$, 当 $r^2 > 0$ 时, 它的图形是以点 $P_0(-a_0, -b_0, -c_0)$ 为球心, 以 r 为半径的球面; 当 $r^2 = 0$ 时, 它的图形只有一个点 $P_0(-a_0, -b_0, -c_0)$; 当 $r^2 < 0$ 时, 它的图形是虚球面.

空间中的一个圆可以看作一个球面与一个平面的交线, 因此可用一个球面方程和一个平面方程联立起来表示.

例 1 求过点 $P_1(a, 0, 0)$, $P_2(0, b, 0)$, $P_3(0, 0, c)$ 的圆的方程.

解 过点 P_1, P_2, P_3 的平面方程为

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

设过这三个点的球面方程为

$$x^2 + y^2 + z^2 + 2x_0x + 2y_0y + 2z_0z + d = 0,$$

将点 P_1, P_2, P_3 的坐标代入, 得

$$\begin{cases} a^2 + 2x_0a + d = 0 \\ b^2 + 2y_0b + d = 0 \\ c^2 + 2z_0c + d = 0. \end{cases}$$

这是关于 x_0, y_0, z_0, d 的四元方程组, 所以 d 可以任意取, 我们取 $d = 0$, 这样就得

$$2x_0 = -a, \quad 2y_0 = -b, \quad 2z_0 = -c,$$

从而球面方程为

$$x^2 + y^2 + z^2 - ax - by - cz = 0.$$

这是过原点 O 和 P_1, P_2, P_3 (4 个点不共面) 的球面方程, 故所求圆的方程为

$$\begin{cases} \frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1 \\ x^2 + y^2 + z^2 = ax + by + cz. \end{cases}$$

由于过三点的球面有无穷多个, 所以过三点的圆(这是唯一确定的)用球面方程与上述平面方程联立的表示式不是唯一的.

8.1.2 柱面

由一族平行直线形成的曲面叫做柱面, 这些平行直线叫做柱面的母线, 在

柱面上与每条母线相交的一条曲线叫做柱面的一条准线(图 8-1). 显然, 一个柱面的准线不是唯一的. 例如用任一个与母线相交的平面去截柱面, 其交线都是柱面的准线.

给定一个柱面, 我们选取一个空间直角坐标系, 使 z 轴平行于母线. 此时, 位于柱面同一条母线上的所有点的 x, y 坐标都相同, 而其 z 坐标的集合是实数集 \mathbf{R} , 这表明柱面在这个坐标系中的方程对于 z 是不加限制的. 因此柱面的方程不包含 z , 它的方程就是所有母线与坐标平面 xOy 的交点坐标 $(x, y, 0)$ 中的 x, y 应满足的方程

$$F(x, y) = 0. \quad (8-3) \quad \text{图 8-1}$$

如果在平面直角坐标系 xOy 内看这个方程, 它的图形是一条曲线, 也就是柱面在 xOy 坐标平面上的准线. 但是必须注意, 我们现在是在空间中考虑问题, 用的是空间直角坐标系, 因此, 这个柱面在 xOy 坐标平面上的准线的方程必须写作

$$\begin{cases} F(x, y) = 0 \\ z = 0, \end{cases} \quad (8-4)$$

所以方程 (8-3) 是以 (8-4) 式所表示的曲线为准线, 以平行于 z 轴的直线为母线的柱面方程.

一般地说, 母线平行于一个坐标轴的柱面方程不包含对应的坐标; 反之, 一个不包含某个坐标的方程的图形是母线平行于该坐标轴的柱面.

但要注意, 如果母线不平行于坐标轴, 其柱面方程就要包含所有的坐标. 例如所有的平面都可视为柱面, 它的一般方程是三元一次方程.

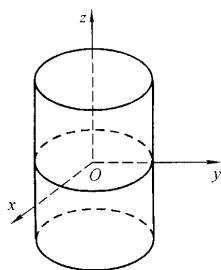


图 8-2

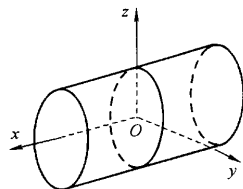


图 8-3

例2 下列二次方程：

$$(1) x^2 + y^2 = 1; (2) \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1; (3) \frac{y^2}{b^2} - \frac{x^2}{a^2} = 1; (4) y = x^2$$

各表示一个柱面,统称为二次柱面.(1)为圆柱面(在 xOy 平面上的准线是单位圆,母线平行于 z 轴,图8-2)(2)为椭圆柱面(在 yOz 平面上的准线为椭圆,母线平行于 x 轴,图8-3)(3)为双曲柱面(在 xOy 平面上的准线为双曲线,母线平行于 z 轴,图8-4)(4)为抛物柱面(在 xOy 平面上的准线为抛物线,母线平行于 z 轴,图8-5).

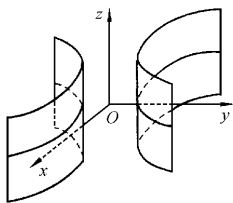


图8-4

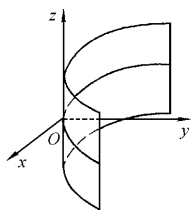


图8-5

8.1.3 锥面

过一个定点的直线族形成的曲面叫做锥面,这些直线叫做锥面的母线,定点叫锥面的顶点,在锥面上与每条母线都相交但不过顶点的一条曲线叫做锥面的准线.

给定了锥面的一条准线和锥面的顶点,把准线上的每个点与顶点用直线相连就得到一个锥面.如果把坐标系的原点 $O(0, 0, 0)$ 置于锥面的顶点,则锥面上任一点 $P(x, y, z)$ 与顶点 O 的连线上的点的坐标就是 (tx, ty, tz) , t 可为任意实数.因此,顶点在原点的锥面,其方程的特征是:如果 (x, y, z) 满足方程,那么对于任意实数 t (tx, ty, tz) 也满足方程.所以,如果 $F(x, y, z)$ 是一个齐次多项式,则方程 $F(x, y, z) = 0$ 就表示一个顶点在原点的锥面.

例3 求顶点在原点,准线为圆

$$\begin{cases} x^2 + y^2 = r^2 & (r^2 > 0) \\ z = c & (c \neq 0) \end{cases}$$

的圆锥面(图8-6)的方程.

解 设点 $P(x, y, z)$ 是锥面上非顶点的任意点,显然点 P 的 $z \neq 0$. 于是锥面上连线 OP 的方程为:

$$x' = 0 + tx, \quad y' = 0 + ty, \quad z' = 0 + tz,$$

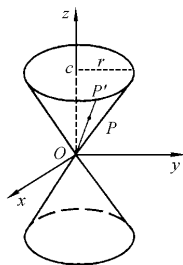


图 8-6

其中 (x', y', z') 是直线 OP 上的任意点的坐标. 容易求得直线 OP 与平面 $z = c$ 的交点 P' 的坐标为 $(\frac{cx}{z}, \frac{cy}{z}, c)$, 显然点 P' 在锥面的准线上, 把点 P' 的坐标代入准线方程, 得

$$\left(\frac{cx}{z}\right)^2 + \left(\frac{cy}{z}\right)^2 = r^2,$$

即

$$x^2 + y^2 = \left(\frac{r}{c}\right)^2 z^2,$$

这就是所求的圆锥面方程. 当 $c = r$ 时, 这个圆锥面的方程为

$$z^2 = x^2 + y^2.$$

它的每条母线与 z 轴的夹角均为 45° , $z = \sqrt{x^2 + y^2}$ 表示 xOy 坐标平面以上部分的圆锥面.

再如 二次齐次方程

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad (abc \neq 0)$$

也是表示顶点在原点的一个锥面, 要知道它是怎样的锥面, 只要知道它的一条准线就行了. 这个锥面与平面 $x = a$ 的交线

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0, \\ x = a, \end{cases} \quad \text{即} \quad \begin{cases} \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \\ x = a \end{cases}$$

是其一条准线, 它是一个椭圆, 所以此锥面称为椭圆锥面.

例 4 二次齐次方程 $xy + yz + zx = 0$ 表示顶点在原点的一个怎样的锥面?

解 这个锥面在平面 $x + y + z = 1$ 上的准线是

$$\begin{cases} xy + yz + zx = 0 \\ x + y + z = 1. \end{cases}$$

由于这条准线上的点 (x, y, z) 的坐标满足

$$x^2 + y^2 + z^2 = (x + y + z)^2 - 2(xy + yz + zx) = 1,$$

所以这条准线的方程也可表示为

$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ x + y + z = 1. \end{cases}$$

这表明准线是一个圆. 因为上式中的球面的球心正是锥面的顶点, 所以这个锥面是一个圆锥面, 它过坐标轴上的三个点 $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$.

此例也可用正交变换将二次型 $f(x, y, z) = xy + yz + zx$ 化为标准形, 来判断 $f(x, y, z) = 0$ 是一个圆锥面, 其解法如下:

$$f(x, y, z) = (x, y, z) \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

容易求得: 二次型矩阵 A 的特征值为 $\lambda_1 = 1$, $\lambda_2 = -\frac{1}{2}$ (二重); 特征子空间

V_{λ_1} 的单位正交基为 $\xi_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T$, V_{λ_2} 的单位正基为

$\xi_2 = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right)^T$, $\xi_3 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right)^T$. 作正交变换

$$\mathbf{X} = \mathbf{QY},$$

其中 $\mathbf{X} = (x, y, z)^T$, $\mathbf{Y} = (x_1, y_1, z_1)^T$,

$$\mathbf{Q} = (\xi_1, \xi_2, \xi_3) = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix},$$

得

$$f(x, y, z) = x_1^2 - \frac{y_1^2}{2} - \frac{z_1^2}{2} = 0.$$

显然,这是在直角坐标系 (O, ξ_1, ξ_2, ξ_3) 中的一个圆锥面的方程.

8.1.4 旋转面

由一条平面曲线绕该平面上一条直线旋转而形成的曲面称为旋转面,这条直线叫做它的轴,曲线叫做它的一条准线.圆柱面和圆锥面都是以直线为准线的旋转面.

为便于求旋转面的方程,我们把曲线所在的平面取作坐标平面,把旋转面的轴取作坐标轴.

设在 xOy 平面上给定了一条曲线 L :

$$\begin{cases} f(x, y) = 0, & y > 0 \\ z = 0. \end{cases}$$

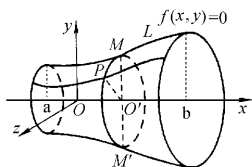


图 8-7

现在推导曲线 L 绕 x 轴旋转而形成的旋转面 S (图 8-7) 的方程.

在 S 上任取点 $P(x, y, z)$, 点 P 到 x 轴的垂直距离为

$$|O'P| = \sqrt{y^2 + z^2}.$$

既然 P 在 S 上, 那末 P 绕 x 轴旋转到 xOy 平面上去, 就得到两个点 M 和 M' , 它们的坐标分别为 $(x, \sqrt{y^2 + z^2}, 0)$ 和 $(x, -\sqrt{y^2 + z^2}, 0)$. 显然 M 和 M' 必有一点在旋转面的准线 L 上, 这里点 M 在图示的准线 L 上, 所以将点 M 的坐标代入 $f(x, y) = 0$ 所得的方程

$$f(x, \sqrt{y^2 + z^2}) = 0$$

就是旋转面 S 的方程.

同样, 把曲线 L 绕 y 轴旋转所形成的旋转面的方程是

$$f(\pm \sqrt{x^2 + z^2}, y) = 0,$$

这里要注意: 取正(负)号所得的方程是曲线 L 在 $x \geq 0$ ($x \leq 0$) 的部分绕 y 轴旋转而成的旋转面方程.

例 5 把双曲线 L :

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ z = 0 \end{cases}$$

分别绕虚轴(y 轴)和实轴(x 轴)旋转 ,求所得旋转面 S 的方程.

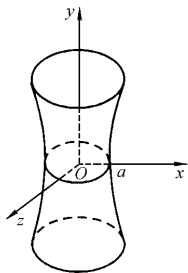


图 8-8

解 绕 y 轴旋转 , S 上的点 $P(x, y, z)$ 可旋转为准线 L 上的点 $(\pm\sqrt{x^2+z^2}, y, 0)$,所以旋转面 S 的方程为

$$\frac{x^2+z^2}{a^2} - \frac{y^2}{b^2} = 1,$$

绕 x 轴旋转 , S 上的点 $P(x, y, z)$ 可旋转为准线 L 上的点 $(x, \pm\sqrt{y^2+z^2}, 0)$,故 S 的方程为

$$\frac{x^2}{a^2} - \frac{y^2+z^2}{b^2} = 1.$$

这两种旋转面分别叫做单叶旋转双曲面(图 8-8)和双叶旋转双曲面(图 8-9).

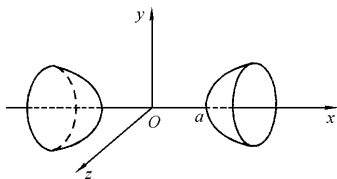


图 8-9

8.2 空间曲线的方程

前面讲过,空间曲线作为两个曲面的交线,其一般方程是两个曲面方程的联立方程组.

一条空间曲线的方程,也常用参数方程

$$\begin{cases} x = f(t), \\ y = g(t), \\ z = h(t), \end{cases} \quad a \leq t \leq b$$

来表示.这个方程如果是质点在空间直角坐标系中的运动方程,它的图象就是质点的运动轨迹,它确定了从 R 的子集 $[a, b]$ 到 \mathbf{R}^3 的一个映射,即把 $t \in [a, b]$ 映射为 $(x(t), y(t), z(t)) \in \mathbf{R}^3$, 此时曲线上的点与参数 t 相对应.

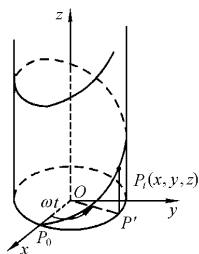


图 8-10

例 1 圆柱螺线 (或称螺旋线, 图 8-10). 设一动点 P 到某一个轴的垂直距离为 r , 如果点 P 以匀角速度 ω 绕轴转动, 同时又以匀速度 v 沿轴的某个方向移动, 则动点 P 的轨迹叫做圆柱螺线 (因为动点 P 始终在半径为 r 的圆柱面上). 圆柱形螺钉的螺纹线就是这种螺旋线. 选取空间直角坐标系, 规定点 P 按反时针方向绕 z 轴转动, 同时沿 z 正向移动, 其起点 ($t = 0$) 位置在点 $P_0(r, 0, 0)$ 处, 在时刻 t 时, 点 P_t 的坐标为 (x, y, z) (图 8-10). 由于点 P_t 在 xOy 平面上的投影点 P' 的坐标为 $(x, y, 0)$, 在 0 到 t 时间内, 点 P 绕 z 轴转过的角度为 ωt , 点 P 沿 z 轴正向移动的距离为 $|P'P_t| = vt$, 因此

$$\begin{cases} x = r \cos \omega t \\ y = r \sin \omega t \\ z = vt \end{cases} \quad t \geq 0,$$

这就是圆柱螺线(螺旋线)的参数方程, t 为参数.

下面讨论空间曲线在坐标平面上的投影曲线(简称投影).

设 L 是一条空间曲线, π 是一个平面, 过 L 上任一点 M , 作 π 的垂线, 此垂线与 π 的交点 P 叫做 M 在 π 上的投影点. 过 L 上所有点的这些垂线构成一个柱面, 叫做从 L 到 π 的投影柱面. 显然, L 在 π 上的投影就是上述投影柱面与 π 的交线.

例2 求曲线 L :

$$\begin{cases} x^2 + y^2 + z^2 = 1 \\ x^2 + y^2 - x = 0 \end{cases} \quad (1)$$

在 xOy , xOz 坐标面上的投影.

解 由于圆柱面 $x^2 + y^2 - x = 0$ 的母线平行于 z 轴(即垂直于 xOy 平面), 所以曲线 L 在该圆柱面上, 而这个圆柱面也是从曲线 L 到 xOy 平面的投影柱面. 因此, L 在 xOy 平面上的投影曲线的方程为

$$\begin{cases} x^2 + y^2 - x = 0 \\ z = 0. \end{cases}$$

从曲线 L 的方程组(1)中消去 y 得到方程

$$z^2 = -x + 1. \quad (2)$$

方程(2)是母线平行于 y 轴(即垂直于 xOz 平面)的柱面方程. 由于这个方程是从 L 的方程得到的, 所以 L 上的一切点的坐标都满足方程(2), 即 L 在柱面(2)上. 而柱面(2)也是从 L 到 xOz 平面的投影柱面, 因此 L 在 xOz 平面上的投影曲线的方程为

$$\begin{cases} z^2 = -x + 1 \\ y = 0, \end{cases}$$

这是 xOz 平面上的一条抛物线.

8.3 二次曲面

一个三元二次方程所表示的曲面叫做二次曲面. 球面、二次柱面和二次锥面都是二次曲面. 现在再介绍5种典型的二次曲面. 用几何特征来刻画这些二次曲面是比较复杂的, 但是, 适当地选取坐标系, 它们的方程可以化成形式简单的标准方程. 因此, 我们先从它们的标准方程来讨论它们的图形. 最后再介绍用坐标变换(正交变换和平移变换)把一般二次曲面方程化为标准方程.

8.3.1 椭球面

在 \mathbf{R}^3 中, 对向量 (x, y, z) 作比例变换

$$\rho(x, y, z) = (ax, by, cz)$$

其中 a, b, c 是不完全相同的正数, 则球心在原点的球面

$$x^2 + y^2 + z^2 = 1$$

在坐标轴方向按不完全相同的比例放大或缩小后, 变成一个椭球面, 其方程变为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (8-5)$$

当 $a = b = c$ 时, 它是一个半径为 a 的球面(特殊的椭球面).

从椭球面方程(8-5)可以看出它的图形有以下一些性质:

(1) 椭球面上点的坐标是有界的, 即

$$|x| \leq a, \quad |y| \leq b, \quad |z| \leq c.$$

这个性质是椭球面在二次曲面中最突出的特点.

(2) 椭球面的图形具有对称性. 它对于每个坐标平面对称的, 因为如果点 $P(x, y, z)$ 满足方程, 则点 P 分别关于 yOz , xOz , xOy 坐标平面的对称点 $P_1(-x, y, z)$, $P_2(x, -y, z)$, $P_3(x, y, -z)$ 也都满足方程. 由此可以推知椭球面关于每个坐标轴和原点也是对称的.

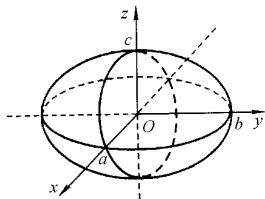


图 8-11

椭球面的对称中心、对称轴和对称平面分别叫做它的中心、主轴和主平面. 椭球面与三个对称轴的三对交点, 称为顶点. 如果 $a > b > c$, 那末 a, b, c 分别叫做半长轴、半中轴和半短轴. 椭球面与三个主平面(即坐标平面)的交线是三个椭圆(图 8-11).

如果 $a = b \neq c$, 则椭球面(8-5)是 yOz 平面上的椭圆

$$\begin{cases} \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \\ x = 0 \end{cases}$$

绕 z 轴旋转所成的旋转椭球面.

8.3.2 单叶双曲面

在 \mathbf{R}^3 中对向量 (x, y, z) 作比例变换

$$\rho(x, y, z) = \left(\frac{ax}{b}, y, z \right),$$

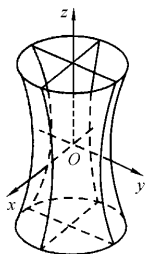


图 8-12

则单叶旋转双曲面

$$\frac{x^2}{b^2} + \frac{y^2}{c^2} - \frac{z^2}{a^2} = 1$$

上点的 x 坐标按比例 a/b 来改变 (即图形向 yOz 平面压缩或拉伸) 就得到一个曲面, 称为单叶双曲面 (图 8-12), 其方程

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (8-6)$$

称为单叶双曲面的标准方程.

显然, 方程 (8-6) 的图形对称于三个坐标平面和三个坐标轴及原点, 它的形状大体上是压扁了的单叶旋转双曲面.

用平行于 xOy 坐标平面的平面 $z = h$ 去截割它, 截线都是椭圆

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2} \\ z = h, \end{cases}$$

其 4 个顶点分别在另两个坐标平面上. 但另两个坐标面截割曲面的截线是下

列双曲线

$$\begin{cases} \frac{x^2}{a^2} - \frac{z^2}{c^2} = 1, \\ y = 0, \end{cases} \quad \text{和} \quad \begin{cases} \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \\ x = 0, \end{cases}$$

它们的虚轴为 z 轴, 虚轴长均为 $2c$. 单叶双曲面可以看作是由一个长、短轴可变的椭圆(所在平面垂直虚轴 z)沿这两条双曲线运动的轨迹, 这个椭圆的两对顶点分别在这两条双曲线上.

8.3.3 双叶双曲面

把双叶旋转双曲面 $\frac{y^2}{b^2} - \frac{x^2 + z^2}{c^2} = 1$ 向 yOz 坐标面压缩或拉伸, 使坐标 x 按比例 a/c 来改变而形成的曲面称为双叶双曲面(图 8-13), 其方程

$$-\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (8-7)$$

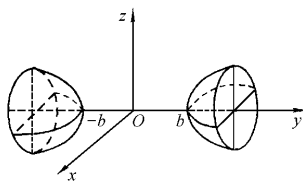


图 8-13

称为双叶双曲面的标准方程. 它的图形也对称于各坐标平面、坐标轴和原点, 大体形状是压扁了的双叶旋转双曲面. 该曲面上点的 y 坐标恒有

$$y^2 \geq b^2, \text{ 即 } |y| \geq b,$$

因而曲面分成两叶.

用平行于 xOz 坐标面的平面 $y = k$ ($|k| \geq b$) 去截割它, 截线都是椭圆

$$\begin{cases} \frac{x^2}{a^2} + \frac{z^2}{c^2} = \frac{k^2}{b^2} - 1 \\ y = k. \end{cases}$$

它的 4 个顶点分别在另两个坐标平面 $z = 0$ 和 $x = 0$ 上, 但曲面在这两个对称平面上的截线都是双曲线

$$\begin{cases} -\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \\ z = 0; \end{cases} \quad \text{和} \quad \begin{cases} \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \\ x = 0. \end{cases}$$

它们的实轴为 y 轴, 实轴长均为 $2b$. 双叶双曲面可以看作是由一个长、短轴可变的椭圆(所在平面垂直于实轴 y) 沿这两条双曲线运动的轨迹, 这个椭圆的两对顶点分别在这两条双曲线上.

以上三种二次曲面是有心二次曲面, 它们都有一个对称中心.

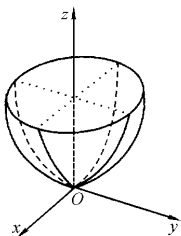


图 8-14

8.3.4 椭圆抛物面

把旋转抛物面 $z = \frac{x^2}{a^2} + \frac{y^2}{a^2}$ 向 xOz 坐标面压缩或拉伸, 使坐标 y 按比例 $\frac{b}{a}$ 来改变而形成的曲面, 称为椭圆抛物面(图 8-14), 其方程

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2} \quad (8-8)$$

称为椭圆抛物面的标准方程. 它的图形对称于 xOz 和 yOz 坐标面. 因而也对称于 z 轴, 但它没有对称中心. 它与对称轴的交点叫做图形的顶点. 它的大体形状是压扁了的旋转抛物面, 该曲面上点的 z 坐标恒 ≥ 0 , 图形在 xOy 坐标面的上方(顶点在原点).

用平行于坐标面 xOy 的平面 $z = h$ ($h \geq 0$) 去截割它, 截线都是椭圆

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = h \\ z = h. \end{cases}$$

它的 4 个顶点分别在另两个坐标平面 $y = 0$ 和 $x = 0$ 上, 但曲面在这两个对称平面上的截线是抛物线

$$\begin{cases} x^2 = a^2 z \\ y = 0, \end{cases} \quad \text{和} \quad \begin{cases} y^2 = b^2 z \\ x = 0. \end{cases}$$

它们的对称轴都是 z 轴, 顶点和开口方向都是相同的. 椭圆抛物面可以看作是

由一个长、短轴可变的椭圆(所在平面垂直于抛物线的对称轴)沿这两条抛物线运动的轨迹,这个椭圆的两对顶点分别在这两条抛物线上.

8.3.5 双曲抛物面(马鞍面)

双曲抛物面的标准方程为

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 2pz \quad (p \neq 0). \quad (8-9)$$

现假定 $p > 0$, 它的图形显然对称于 xOz 和 yOz 坐标面, 因而对称于 z 轴, 但它没有对称中心. 它也不能由压缩旋转面而得到, 为了了解它的图形结构, 我们用平行于坐标面的平面去截割它, 通过截线了解其图形的情况.

用平面 $z = h$ 去截割它, 截线为

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 2ph \\ z = h. \end{cases}$$

当 $h = 0$ 时, 它是 xOy 平面上的两条直线 $y = \pm bx/a$; 当 $h \neq 0$ 时, 它是双曲线, 如果 $h > 0$, 其实轴平行于 x 轴, 如果 $h < 0$, 其实轴平行于 y 轴. 曲面在 xOy 坐标面之上, 沿 x 轴的两个方向上升; 在 xOy 坐标面之下, 沿 y 轴的两个方向下降. 这些截线在 xOy 坐标面上的投影如图 8-15 所示.

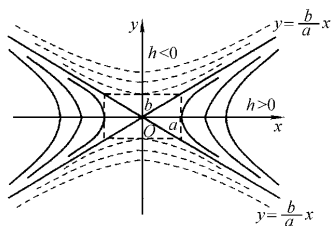


图 8-15

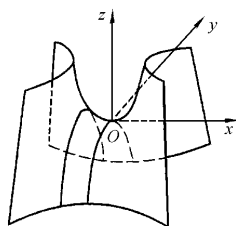


图 8-16

用平行于 yOz 坐标面的平面 $x = k$ 去截它, 截线都是抛物线

$$\begin{cases} y^2 = -2pb^2\left(z - \frac{k^2}{2pa^2}\right) \\ x = k, \end{cases} \quad (p > 0) \quad (1)$$

其对称轴平行于 z 轴, 开口为 z 轴的负方向.

在 xOz 坐标面($y = 0$)上的一条抛物线为

$$\begin{cases} x^2 = 2pa^2z \\ y = 0 \end{cases} \quad (p > 0). \quad (2)$$

这条抛物线的顶点为原点 O , 开口为 z 轴的正方向.

从上述的截线情况, 可以把双曲抛物面(图 8-16)看作是由抛物线(1)平行移动而形成的, 抛物线(1)平行移动时其顶点始终在另一条固定的抛物线(2)上. 我们可以把两条抛物线设想为两根抛物形铁丝, 一根挂在固定的另一根上滑动, 其轨迹就是双曲抛物面, 它的形状恰似马鞍形, 所以也叫马鞍面.

方程 $z = axy$ ($a \neq 0$) 也表示一个双曲抛物面, 它可以通过坐标变换(正交变换)化为标准方程. 理想气体的状态方程 $PV = RT$ (R 为常数)就是这种方程.

椭圆抛物面和双曲抛物面都是无心二次曲面, 因为它们都没有对称中心.

8.4* 二次曲面的分类

二次曲面在直角坐标系中的一般方程为

$$\begin{aligned} F(x_1, x_2, x_3) &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 \\ &\quad + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \\ &\quad + 2a_{14}x_1 + 2a_{24}x_2 + 2a_{34}x_3 + a_{44} = 0. \end{aligned} \quad (8-10)$$

记

$$\mathbf{X} = (x_1, x_2, x_3)^T, \quad \boldsymbol{\alpha} = (a_{14}, a_{24}, a_{34})^T,$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad (a_{ij} = a_{ji}),$$

则(8-10)式可用矩阵表示为

$$\begin{pmatrix} \mathbf{X}^T & 1 \end{pmatrix} \begin{pmatrix} A & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^T & a_{44} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix} = 0, \quad (8-11)$$

其中 $\mathbf{X}^T A \mathbf{X}$ 是(8-10)式中的二次型部分, 即

$$\begin{aligned} f(x_1, x_2, x_3) &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 \\ &\quad + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 = \mathbf{X}^T A \mathbf{X}. \end{aligned} \quad (8-12)$$

这里的 x_1, x_2, x_3 是曲面上点 P 的直角坐标, 即 $P(x_1, x_2, x_3)$, 或者说向量 \overrightarrow{OP} 在单位正交基 $\{e_1, e_2, e_3\}$ 下的坐标, 即 $\overrightarrow{OP} = (x_1, x_2, x_3)$.

我们在第7章中讲过, 对于三元二次型 $\mathbf{X}^T A \mathbf{X}$ 的矩阵 A , 存在正交阵 Q , 使得

$$Q^T A Q = \text{diag}(\lambda_1, \lambda_2, \lambda_3),$$

因此作转轴坐标变换 $\mathbf{X} = Q\mathbf{Y}$ (正交矩阵 Q 是把 \mathbf{R} 的自然基 $\{e_1, e_2, e_3\}$ 变为另一组单位正

交基 $\{\xi_1, \xi_2, \xi_3\}$ 的变换矩阵, 点 P 在基 $\{\xi_1, \xi_2, \xi_3\}$ 下的坐标为 (y_1, y_2, y_3) , 可将二次型化为标准形, 即

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{Y}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \mathbf{Y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2. \quad (8-13)$$

再将 $\mathbf{X} = \mathbf{Q} \mathbf{Y}$ 代入(8-11)或(8-10)式, 可得

$$F(x_1, x_2, x_3) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 + 2b_{14}y_1 + 2b_{24}y_2 + 2b_{34}y_3 + b_{44} = 0. \quad (8-14)$$

再对由基 $\{\xi_1, \xi_2, \xi_3\}$ 确定的空间直角坐标系作坐标的平移变换

$$z_1 = y_1 + \frac{b_{14}}{\lambda_1}, \quad z_2 = y_2 + \frac{b_{24}}{\lambda_2}, \quad z_3 = y_3 + \frac{b_{34}}{\lambda_3}$$

(这里假设 $\lambda_1 \lambda_2 \lambda_3 \neq 0$)进而可将(8-14)式化为只含平方项和常数项的二次曲面的方程.

$$\lambda_1 z_1^2 + \lambda_2 z_2^2 + \lambda_3 z_3^2 + d = 0. \quad (8-15)$$

我们知道, 对二次型(8-12)不论作怎样的正交变换 $\mathbf{X} = \mathbf{Q} \mathbf{Y}$, 将其化为标准形 $\lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2$, 其系数 $\lambda_1, \lambda_2, \lambda_3$ 都是唯一确定的, 因为它们是二次型矩阵 \mathbf{A} 的特征值, 即

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{Q}^{-1} \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \sim \mathbf{A},$$

因此, 由矩阵 $\mathbf{A} = (a_{ij})_{3 \times 3}$ 的特征多项式(见定理7.5)

$$\begin{aligned} |\lambda \mathbf{E} - \mathbf{A}| &= \lambda^3 - I_1 \lambda^2 + I_2 \lambda - I_3 \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \end{aligned} \quad (8-16)$$

及根 $\lambda_1, \lambda_2, \lambda_3$ 与系数的关系得:

$$\begin{aligned} I_1 &= \lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33}, \\ I_2 &= \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1 \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \\ I_3 &= \lambda_1 \lambda_2 \lambda_3 = |\mathbf{A}|. \end{aligned}$$

由于二次曲面的一般方程(8-10)经过坐标变换(转轴的正交变换和坐标平移)化简后的方程, 其二次型部分对应的矩阵

$$\mathbf{A} \sim \mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \lambda_3),$$

所以它们的特征多项式相同, 因此它们的 I_1 (主对角元之和), I_2 (三个二阶主子式之和), I_3 (二次型矩阵的行列式)是不变的, 称为二次曲面的不变量. 此外(8-11)式中4阶对称阵的行列式

$$I_4 = \det \begin{pmatrix} \mathbf{A} & \boldsymbol{\alpha} \\ \boldsymbol{\alpha}^T & a_{44} \end{pmatrix}$$

也是二次曲面(8-10)的不变量(证明略). 其中4阶矩阵称为二次曲面(8-10)对应的矩阵.

根据二次曲面的不变量, 可把二次曲面分为两大类五小类, 共十七种形状(包括虚曲

面等):

I 第一大类, $I_3 \neq 0$, 有心二次曲面:

由于 $I_3 = \lambda_1 \lambda_2 \lambda_3 \neq 0$, 即 $\lambda_1, \lambda_2, \lambda_3$ 皆为非零常数, 所以方程(8-10)化简后的方程(8-15)为(将 z_1, z_2, z_3 改用 x_1, x_2, x_3)

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + d = 0.$$

利用不变量

$$I_4 = |\text{diag}(\lambda_1, \lambda_2, \lambda_3, d)| = \lambda_1 \lambda_2 \lambda_3 d = I_3 d$$

得

$$d = \frac{I_4}{I_3},$$

所以, 这类曲面化简后的方程为

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + \frac{I_4}{I_3} = 0. \quad (8-17)$$

于是这类曲面就 I_4 又可分为三小类:

1° $I_4 > 0$

(1) $\lambda_1, \lambda_2, \lambda_3$ 同号时($\frac{I_4}{I_3}$ 也同号), 为虚椭球面.

(2) $\lambda_1, \lambda_2, \lambda_3$ 异号时, 不论有一个为负或两个为负, 均为单叶双曲面.

2° $I_4 < 0$

(3) $\lambda_1, \lambda_2, \lambda_3$ 同号时($\frac{I_4}{I_3}$ 则异号), 为椭球面.

(4) $\lambda_1, \lambda_2, \lambda_3$ 异号时, 不论有一个为负或两个为负, 均为双叶双曲面.

3° $I_4 = 0$

(5) $\lambda_1, \lambda_2, \lambda_3$ 同号时, 为虚二次锥面(或一个点)

(6) $\lambda_1, \lambda_2, \lambda_3$ 异号时, 为实二次锥面.

II 第二大类, $I_3 = 0$, 无心二次曲面:

4° $I_2 \neq 0$

此时, 由 $I_3 = \lambda_1 \lambda_2 \lambda_3 = 0$, $I_2 = \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1 \neq 0$, 可知特征值只有一个为 0, 不妨设 $\lambda_3 = 0$, 于是对方程(8-14)作坐标平移, 则化简后的方程为

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + 2px_3 = 0 \quad (8-18)$$

或

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + d = 0. \quad (8-19)$$

4°-1 $I_4 \neq 0$

此时方程化简为(8-18)的形式. 由

$$I_2 = \lambda_1 \lambda_2, \quad I_4 = \begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & p \\ 0 & 0 & p & 0 \end{vmatrix} = -\lambda_1 \lambda_2 p^2 = -I_2 p^2 \quad (8-20)$$

得 $p = \pm \sqrt{-I_4/I_2}$, 于是方程 (8-18) 即为

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 \pm \sqrt{\frac{-I_4}{I_2}} x_3 = 0, \quad (8-21)$$

因此, 当 $I_3 = 0, I_2 \neq 0, I_4 \neq 0$ 时, 有两种曲面:

(7) $I_4 < 0$, 由 (8-20) 得 $I_2 = \lambda_1 \lambda_2 > 0$, 即 λ_1, λ_2 同号, 此时, 方程 (8-21) 为椭圆抛物面.

(8) $I_4 > 0$ 则 $I_2 = \lambda_1 \lambda_2 < 0$, 即 λ_1, λ_2 异号, 此时方程 (8-21) 为双曲抛物面.

$$4^\circ - I_4 = 0$$

此时, 方程化为 (8-19) 的形式. 为了确定 (8-19) 式中的 d , 需要引入二次曲面的不变量 K_2 :

$$K_2 = \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{41} & a_{42} & a_{44} \end{vmatrix},$$

即 K_2 是 a_{11}, a_{22}, a_{33} 在 I_4 中的代数余子式之和.

可以证明 (略去), 当 $I_3 = I_4 = 0$ 时, K_2 是不变量 (其它情况 K_2 则是变量), 于是由方程 (8-19) 的不变量 $I_4 = |\text{diag}(\lambda_1, \lambda_2, \lambda_3, d)|$ 得

$$K_2 = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & d \end{vmatrix} = \lambda_1 \lambda_2 d = I_2 d,$$

从而有

$$d = \frac{K_2}{I_2}.$$

因此, 方程 (8-19) 即为

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \frac{K_2}{I_2} = 0. \quad (8-22)$$

于是这类曲面就 $K_2 \neq 0, K_2 = 0$ 又分为以下几种形状.

- * (9) $K_2 \neq 0$ 时, $\lambda_1, \lambda_2, K_2$ 同号, 为虚椭圆柱面.
- * (10) $K_2 \neq 0$ 时, λ_1, λ_2 同号, 但与 K_2 异号, 为椭圆柱面.
- * (11) $K_2 \neq 0$ 时, λ_1, λ_2 异号, 为双曲柱面.
- * (12) $K_2 = 0$ 时, λ_1, λ_2 异号, 为一对相交平面.
- * (13) $K_2 = 0$ 时, λ_1, λ_2 同号, 为虚相交平面.
- * 5° $I_2 = I_4 = 0$

此时, 由 $I_3 = \lambda_1 \lambda_2 \lambda_3 = 0$ 及 $I_2 = \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_1 = 0$, 可知必有 (也仅有) 两个特征值为 0. 不妨设 $\lambda_2 = \lambda_3 = 0, \lambda_1 \neq 0$, 于是方程 (8-14) 经坐标平移可化简为:

$$\lambda_1 z_1^2 + pz_2 + qz_3 = 0. \quad (8-23)$$

(i) 如果 p, q 均不为 0, 在新坐标系中作绕 z_1 轴的旋转变换, 令

$$u_1 = z_1, \quad u_2 = \frac{pz_2 + qz_3}{\sqrt{p^2 + q^2}}, \quad u_3 = \frac{-qz_2 + pz_3}{\sqrt{p^2 + q^2}},$$

则 (8-23) 式变换为

$$\lambda_1 u_1^2 + \sqrt{p^2 + q^2} u_2 = 0.$$

我们把上式改写为

$$\lambda_1 x_1^2 + 2p_0 x_2 = 0. \quad (8-24)$$

如果 p, q 有一个为 0, 方程 (8-23) 就是 (8-24) 的形式.

(ii) 如果 $b_{24} = b_{34} = 0$, 则方程 (8-14) 经坐标平移后, 化简为

$$\lambda_1 x_1^2 + d = 0. \quad (8-25)$$

当 $K_2 \neq 0$ 时, 方程取 (8-24) 的形式. 由方程 (8-24) 的不变量

$$I_4 = \begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_0 \\ 0 & 0 & 0 & 0 \\ 0 & p_0 & 0 & 0 \end{vmatrix}$$

可得

$$K_2 = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & p_0 \\ 0 & p_0 & 0 \end{vmatrix} = -\lambda_1 p_0^2 = -I_1 p_0^2,$$

所以 I_1 与 K_2 异号, 方程 (8-24) 可化为:

$$I_1 x_1^2 \pm 2\sqrt{\frac{-K_2}{I_1}} x_2 = 0. \quad (8-26)$$

当 $K_2 = 0$ 时, 方程取 (8-25) 的形式. 这时还需引入另一个半不变量 K_1 , 即

$$K_1 = \begin{vmatrix} a_{11} & a_{14} \\ a_{41} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{24} \\ a_{42} & a_{44} \end{vmatrix} + \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix}.$$

当 $I_3 = I_4 = I_2 = K_2 = 0$ 时, 可以证明 (略去), K_1 也是二次曲面的一个不变量. 此时从方程 (8-25) 可得

$$K_1 = \begin{vmatrix} \lambda_1 & 0 \\ 0 & d \end{vmatrix} = \lambda_1 d = I_1 d,$$

所以 $d = \frac{K_1}{I_1}$, 于是方程 (8-25) 可化为

$$I_1 x_1^2 + \frac{K_1}{I_1} = 0. \quad (8-27)$$

从方程 (8-26) (8-27), 又可得曲面的 4 种形状:

* (14) 当 $I_3 = I_4 = I_2 = 0, K_2 \neq 0$ 时, 化简后方程为 (8-26), 其图形为抛物柱面.

*(15) 当 $I_3 = I_4 = I_2 = K_2 = 0, K_1 < 0$ 时, 由 (8-27) 知, 图形为一对平行平面.

*(16) 当 $I_3 = I_4 = I_2 = K_2 = 0, K_1 > 0$ 时, 图形为虚平行平面.

*(17) 当 $I_3 = I_4 = I_2 = K_2 = K_1 = 0$ 时, 图形为一对重合平面.

例 1 利用不变量判断下列二次曲面的类型, 并把方程化为最简形式和确定曲面的形状.

$$(1) x_1^2 + x_2^2 + 5x_3^2 - 6x_1x_2 - 2x_1x_3 + 2x_2x_3 - 6x_1 + 6x_2 - 6x_3 + 10 = 0;$$

$$(2) 2x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3 - 4x_1 + 6x_2 - 2x_3 + 3 = 0.$$

解 (1) 二次曲面所对应的矩阵为

$$P = \begin{pmatrix} 1 & -3 & -1 & -3 \\ -3 & 1 & 1 & 3 \\ -1 & 1 & 5 & -3 \\ -3 & 3 & -3 & 10 \end{pmatrix} = \begin{pmatrix} A & \alpha \\ \alpha^T & a_{44} \end{pmatrix},$$

所以 $I_3 = \det A = -36, I_4 = \det P = -252$, 再由

$$|\lambda E - A| = \begin{vmatrix} \lambda - 1 & 3 & 1 \\ 3 & \lambda - 1 & -1 \\ 1 & -1 & \lambda - 5 \end{vmatrix} = (\lambda + 2)(\lambda - 3)(\lambda - 6)$$

得 $\lambda_1 = -2, \lambda_2 = 3, \lambda_3 = 6$ (三个特征值不同号), 于是化简后的方程为:

$$-2z_1^2 + 3z_2^2 + 6z_3^2 + \frac{-252}{-36} = 0,$$

即

$$2z_1^2 - 3z_2^2 - 6z_3^2 = 7,$$

故方程的图形为双叶双曲面.

(2) 二次曲面所对应的矩阵

$$P = \begin{pmatrix} 2 & 2 & 1 & -2 \\ 2 & 2 & 1 & 3 \\ 1 & 1 & 3 & -1 \\ -2 & 3 & -1 & 3 \end{pmatrix} = \begin{pmatrix} A & \alpha \\ \alpha^T & a_{44} \end{pmatrix},$$

所以 $I_3 = \det A = 0, I_4 = \det P = -125, I_2 = 10$, 再由

$$|\lambda E - A| = \begin{vmatrix} \lambda - 2 & -2 & -1 \\ -2 & \lambda - 2 & -1 \\ -1 & -1 & \lambda - 3 \end{vmatrix} = \lambda(\lambda - 2)(\lambda - 5)$$

得 $\lambda_1 = 5, \lambda_2 = 2, \lambda_3 = 0$, 于是化简后的方程为

$$5z_1^2 + 2z_2^2 + 2\sqrt{\frac{125}{10}}z_3 = 0, \quad \text{即} \quad 5z_1^2 + 2z_2^2 + 5\sqrt{2}z_3 = 0,$$

故方程的图形是椭圆抛物面.

例 2 对二次曲面方程

$$F(x_1, x_2, x_3) = 4x_1^2 + x_2^2 + 4x_3^2 - 4x_1x_2 + 8x_1x_3 - 4x_2x_3 - 12x_1 - 3x_2 + 6x_3 = 0$$

作坐标变换, 将其化为标准方程.

$$\text{解} \quad F(\mathbf{X}) = \mathbf{X}^T \mathbf{A} \mathbf{X} + \mathbf{b}^T \mathbf{X},$$

其中 $\mathbf{X} = (x_1, x_2, x_3)^T$, $\mathbf{b} = (-12, -3, 6)^T$; 二次型部分的矩阵为

$$A = \begin{vmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{vmatrix},$$

$|\lambda E - A| = \lambda^2(\lambda - 9)$ 属于特征值 $\lambda_1 = 9$, $\lambda_2 = \lambda_3 = 0$ 的单位正交的特征向量分别为:

$$\xi_1 = \left(\frac{2}{3}, \frac{-1}{3}, \frac{2}{3} \right)^T, \quad \xi_2 = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0 \right)^T, \quad \xi_3 = \left(\frac{-4}{3\sqrt{5}}, \frac{2}{3\sqrt{5}}, \frac{5}{3\sqrt{5}} \right)^T.$$

取正交矩阵 $Q = (\xi_1, \xi_2, \xi_3)$, 作坐标变换

$$\mathbf{X} = \mathbf{Q} \mathbf{Y}, \quad (2)$$

则

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{Y}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \mathbf{Y} = \mathbf{Y}^T (\text{diag}(9, 0, 0)) \mathbf{Y} = 9y_1^2,$$

$$\begin{aligned} \mathbf{b}^T \mathbf{X} &= (\mathbf{b}^T \mathbf{Q}) \mathbf{Y} = \left(-3, \frac{-18}{\sqrt{5}}, \frac{24}{\sqrt{5}} \right) (y_1, y_2, y_3)^T \\ &= -3y_1 - \frac{18}{\sqrt{5}}y_2 + \frac{24}{\sqrt{5}}y_3. \end{aligned}$$

于是方程 $F(\mathbf{X}) = 0$, 化为

$$\begin{aligned} G(\mathbf{Y}) &= 9y_1^2 - 3y_1 - \frac{18}{\sqrt{5}}y_2 + \frac{24}{\sqrt{5}}y_3 \\ &= 9\left(y_1 - \frac{1}{6}\right)^2 - \frac{18}{\sqrt{5}}\left(y_2 + \frac{\sqrt{5}}{72}\right) + \frac{24}{\sqrt{5}}y_3 = 0. \end{aligned} \quad (3)$$

再作坐标平移, 令 $z_1 = y_1 - \frac{1}{6}$, $z_2 = y_2 + \frac{\sqrt{5}}{72}$, $z_3 = y_3$, 即

$$\mathbf{Z} = \mathbf{Y} + \mathbf{Y}_0, \quad (4)$$

其中 $\mathbf{Y}_0 = \left(\frac{-1}{6}, \frac{\sqrt{5}}{72}, 0 \right)^T$, 于是方程 $G(\mathbf{Y}) = 0$ 化为

$$9z_1^2 - \frac{18}{\sqrt{5}}z_2 + \frac{24}{\sqrt{5}}z_3 = 0. \quad (5)$$

对于方程 (5), 再把平移后的坐标系 z_1 轴作旋转变换 (记 $p = -\frac{18}{\sqrt{5}}$, $q = \frac{24}{\sqrt{5}}$), 令

$$u_1 = z_1, \quad u_2 = \frac{pz_2 + qz_3}{\sqrt{p^2 + q^2}} = \frac{-3z_2}{5} + \frac{4z_3}{5},$$

$$u_3 = \frac{-qz_2 + pz_3}{\sqrt{p^2 + q^2}} = \frac{-4z_2}{5} + \frac{-3z_3}{5},$$

即

$$U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{-3}{5} & \frac{4}{5} \\ 0 & \frac{-4}{5} & \frac{-3}{5} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = TZ, \quad (6)$$

则方程(5)化为标准方程

$$u_1^2 + \frac{\sqrt{p^2 + q^2}}{9} u_2 = u_1^2 + \frac{2\sqrt{5}}{3} u_2 = 0. \quad (7)$$

这是一个抛物柱面的方程,所以方程(7)的图形是抛物柱面,其准线是新的直角坐标系 $O' - u_1 u_2 u_3$ 中坐标平面 $u_1 O' u_2$ 上的一条抛物线.方程(1)化为标准方程(7)所作的坐标变换(由(2)(4)(6)式可得)为

$$X = QY = Q(Z - Y_0) = Q(T^{-1}U) - QY_0,$$

其中坐标系 $O_{y_1 y_2 y_3}$ 在三个坐标轴方向的单位向量为 A 的特征向量(即正交矩阵 Q 的列向量) ξ_1, ξ_2, ξ_3 , 由于 $\det Q = 1$, 所以这个坐标系仍是右手系;坐标系 $O' z_1 z_2 z_3$ 是将坐标系 $O_{y_1 y_2 y_3}$ 在 y_1, y_2 轴方向平移所得;坐标系 $O' u_1 u_2 u_3$ 是将坐标系 $O' z_1 z_2 z_3$ 绕 z_1 轴顺时针方向旋转 $\theta = \pi - \arctan \frac{4}{3}$ 所得.

附带说一下,如果 $\det Q = |\xi_1, \xi_2, \xi_3| = -1$, 则坐标系 $O\xi_1 \xi_2 \xi_3$ 是左手系,只要将基向量 $\{\xi_1, \xi_2, \xi_3\}$ 中任意对换两个向量就构成右手系.

习 题

1. 求下列球面的球心与半径.

(1) $x^2 + y^2 + z^2 - 2x + 4y - 4z - 7 = 0$;

(2) $2x^2 + 2y^2 + 2z^2 - x + 3z = 0$;

2. 求过 4 点 $A(6 \ 0 \ 0), B(4 \ 4 \ 0), C(-3, -3 \ 0), D(1 \ 0 \ 5)$ 的球面方程,并求球心和半径.

3. 动点 P 到点 $A(2 \ 0 \ 0)$ 的距离为到点 $B(-4 \ 0 \ 0)$ 的距离的一半,求动点轨迹的方程.

4. 动点 P 到点 $F_1(-a \ 0 \ 0)$ 与到点 $F_2(a \ 0 \ 0)$ 的距离的平方和等于 $4a^2$, 求动点轨迹的方程.

5. 下列方程在空间直角坐标系中各表示什么曲面?并作其略图.

(1) $x^2 + y^2 - 2y = 0$; (2) $4x^2 + y^2 = 1$; (3) $x^2 = 2z$;

(4) $z^2 = 4$; (5) $x^2 + y^2 + z^2 = 0$; (6) $y^2 + z^2 = 0$;

(7) $x^2 - y^2 = 0$; (8) $x^2 - y^2 = 1$.

6. 下列方程各表示什么曲线:

$$(1) \begin{cases} x^2 + \frac{y^2}{4} = 1 \\ z = 1; \end{cases}$$

$$(2) \begin{cases} x^2 + y^2 + z^2 - 25 = 0 \\ x^2 + y^2 = 16; \end{cases}$$

$$(3) \begin{cases} (x-1)^2 + (y+2)^2 + (z-1)^2 = 25 \\ x^2 + y^2 + z^2 = 36; \end{cases}$$

$$(4) \begin{cases} x = \cos \varphi \\ y = \sin \varphi \quad (\varphi \text{ 为参数}) \\ z = -1. \end{cases}$$

7. 求在 yOz 坐标平面上以原点为圆心的单位圆的方程(写出两种以上不同形式的方程).

8. 求下列顶点在原点的锥面方程:

$$(1) \text{ 准线为 } \begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \\ z = k; \end{cases} \quad (2) \text{ 准线为 6 题(2).}$$

9. 求下列曲线绕指定轴旋转而成的旋转面方程:

$$(1) \begin{cases} x^2 + 4y^2 = 1 \\ z = 0; \end{cases} \text{ 分别绕 } x \text{ 轴、} y \text{ 轴旋转;} \quad (2) \begin{cases} z = \sqrt{y} \\ x = 0; \end{cases} \text{ 分别绕 } y \text{ 轴、} z \text{ 轴旋转;} \quad (3) \begin{cases} x - 4z^2 = 1 \\ y = 0; \end{cases} \text{ 分别绕 } x \text{ 轴、} z \text{ 轴旋转.}$$

10. 求下列曲线在指定平面上的投影:

$$(1) \begin{cases} x^2 + y^2 + z^2 = 4 \\ y = x; \end{cases} \text{ 在各坐标平面上的投影;} \quad (2) \begin{cases} x^2 + (y-1)^2 + z^2 = 1 \\ y + z = 2, \end{cases} \text{ 在平面 } z = 2 \text{ 上的投影.}$$

11. 下列二次方程各表示什么曲面 并作其略图:

$$(1) 16x^2 + 9y^2 + 16z^2 = 144; \quad (2) z = x^2 + (y-1)^2;$$

$$(3) 4x^2 - 4y^2 + 36z^2 = 144; \quad (4) x = (y+1)^2 + \frac{z^2}{4};$$

$$(5) x^2 + 4y^2 - z^2 + 9 = 0; \quad (6) z = \sqrt{x^2 + y^2};$$

$$(7) x^2 + 2y^2 - z^2 = 0; \quad (8) z = xy.$$

12. 画出下列各组曲面所围成的空间体的略图.

$$(1) 2x + 3y + 6z = 6, x = 0, y = 0, z = 0; \quad (2) z = \sqrt{1 - x^2 - y^2}, z = 0;$$

$$(3) x^2 + y^2 - z + 1 = 0, z = 3; \quad (4) y = x^2, y = 2, z = 0, z = 2.$$

13. 用坐标变换(转轴和平移),将下列一般二次曲面方程化为标准方程

(并写出坐标变换式,指出你所选取的新坐标系是右手系还是左手系).

$$(1) x_1^2 + x_2^2 - 2x_3^2 + 8x_1x_2 - 4x_1x_3 - 4x_2x_3 + 6x_1 + 6x_2 + 6x_3 - \frac{11}{2} = 0;$$

$$(2) 13x_1^2 + 13x_2^2 + 10x_3^2 + 8x_1x_2 - 4x_1x_3 - 4x_2x_3 + 36(x_1 + x_2 + x_3) + 81 = 0;$$

$$(3) x_1^2 - 2x_2^2 + 10x_3^2 + 28x_1x_2 + 20x_1x_3 - 8x_2x_3 - 26x_1 + 32x_2 + 28x_3 - 38 = 0;$$

$$(4) x_1^2 + 4x_2^2 + x_3^2 + 2x_1x_2 + 4x_1x_3 + 2x_2x_3 + \sqrt{3}x_1 - 6\sqrt{3}x_2 + \sqrt{3}x_3 + \frac{1}{2} = 0.$$

14. 利用不变量判断下列二次曲面的类型,并把方程化为最简形式和确定曲面的形状.

$$(1) x_1^2 - 2x_2^2 + x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3 - 14x_1 - 4x_2 + 14x_3 + 16 = 0;$$

$$(2) 4x_1^2 + 5x_2^2 + 6x_3^2 - 4x_1x_2 + 4x_2x_3 + 4x_1 + 6x_2 + 4x_3 - 27 = 0;$$

$$(3) 2x_1^2 + 5x_2^2 + 2x_3^2 - 2x_1x_2 - 4x_1x_3 + 2x_2x_3 + 2x_1 - 10x_2 - 2x_3 - 1 = 0;$$

$$(4) 7x_2^2 - 7x_3^2 - 8x_1x_2 + 8x_1x_3 = 0.$$

部分习题答案

$$1. (1) (1, -2, 2); (2) \left(\frac{1}{4}, 0, -\frac{3}{4} \right); \frac{\sqrt{10}}{4}.$$

$$2. (x-1)^2 + y^2 + z^2 = 25. \quad 3. (x-4)^2 + y^2 + z^2 = 16.$$

$$4. x^2 + y^2 + z^2 = a^2.$$

5. (1) 圆柱面; (2) 椭圆柱面; (3) 抛物柱面; (4) 两平行平面: $z = 2$ 和 $z = -2$; (5) 原点 $(0, 0, 0)$; (6) x 轴; (7) 两个过 z 轴的互相垂直的平面 $x = y$ 和 $x = -y$; (8) 双曲柱面.

6. (1) 椭圆(在平面 $z = 1$ 上); (2) 两个圆在平面 $z = \pm 3$ 上;

(3) 圆; (4) 圆(在平面 $z = -1$ 上);

$$7. (1) \begin{cases} y^2 + z^2 = 1 \\ x = 0; \end{cases} \quad (2) \begin{cases} x = 0 \\ y = \cos \theta \\ z = \sin \theta; \end{cases}$$

$$8. (1) z^2 = k^2 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} \right); \quad (2) x^2 + y^2 = \frac{16}{9} z^2.$$

$$9. (1) x^2 + 4y^2 + 4z^2 = 1; \quad x^2 + 4y^2 + z^2 = 1;$$

$$(2) y = z^2 + x^2; z^2 = \sqrt{y^2 + x^2};$$

$$(3) x^2 - 4y^2 - 4z^2 = 1; x^2 + y^2 - 4z^2 = 1.$$

$$10. (1) \begin{cases} y = x \\ z = 0, \left(\frac{-1}{\sqrt{5}} \leq x \leq \frac{1}{\sqrt{5}} \right); \end{cases} \begin{cases} 2x^2 + z^2 = 4 \\ y = 0; \end{cases} \begin{cases} 2y^2 + z^2 = 4 \\ x = 0; \end{cases}$$

$$(2) \begin{cases} x^2 + 2y^2 - 6y + 4 = 0 \\ z = 2. \end{cases}$$

11. (1) 椭球面; (2) 圆抛物面(或绕轴 $x = 0, y = 1$ 的旋转抛物面);
 (3) 单叶双曲面; (4) 椭圆抛物面; (5) 双叶双曲面; (6) 圆锥面;
 (7) 椭圆锥面; (8) 马鞍面.

$$13. (4) -\frac{z_1^2}{10} + \frac{z_2^2}{5} + \frac{z_3^2}{2} = 1 \text{ 左手系.}$$

$$14. (1) I_3 = 54 \neq 0, I_2 = -27 < 0, I_4 = 0, 2x_1^2 - x_2^2 - x_3^2 = 0;$$

$$(2) I_2 = 66 > 0, I_1 I_3 = 1200 > 0, I_4 = -2560 < 0, 2x_1^2 + 5x_2^2 + 8x_3^2 - 32 = 0;$$

$$(3) I_3 = I_4 = 0, I_2 = 18 > 0, I_1 K_2 = -972 < 0, \frac{x_1^2}{2} + x_2^2 = 1;$$

$$(4) I_3 = I_4 = K_2 = 0, I_2 = -81 < 0, 9x_1^2 - 9x_2^2 = 0.$$

第 9 章 空间曲线与空间曲面

在这一章,我们将介绍几何中的一个重要部分——微分几何的最基本内容,即 \mathbf{R}^3 中曲线、曲面在一点附近的局部性质和一些特征量.例如 \mathbf{R}^3 中的曲线在一点处的切线、法平面、密切平面、副法线、从切平面、主法线、曲率和挠率,以及 \mathbf{R}^3 中曲面的第一、第二基本形式和曲面上的曲率——法曲率、主曲率、高斯曲率和平均曲率.研究这些几何量的主要工具是微积分.

9.1 向量函数及其微积分

9.1.1 向量函数

本章我们只考虑 \mathbf{R}^3 中的曲线和曲面.由于这种限制,利用 \mathbf{R}^3 中的向量运算就比较方便,即记

$$\mathbf{r}(t) = (x(t), y(t), z(t)), \quad t \in [a, b]. \quad (9-1)$$

当 $x(t), y(t), z(t)$ 分别为常数时, $\mathbf{r}(t)$ 为常向量,它表示 \mathbf{R}^3 中的一个点.

如果令 $\mathbf{r}(t) = \overrightarrow{OP}$ (通常称它为点 P 的向径),则当 t 从 a 变到 b 时, $\mathbf{r}(t)$ 的图形就是点 P 的轨迹.当 $x(t), y(t), z(t) \in C[a, b]$ 时,点 P 的轨迹是 \mathbf{R}^3 中一条连续的空间曲线 Γ (如图 9-1).

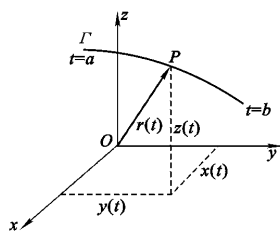


图 9-1

例如,在空间直角坐标系 $Oxyz$ 中,向量函数

$$\mathbf{r}(t) = (x_0 + lt, y_0 + mt, z_0 + nt), \quad t \in (-\infty, +\infty)$$

是通过点 $M_0(x_0, y_0, z_0)$ 方向为 (l, m, n) 的一条直线.而

$$\mathbf{r}(t) = (r \cos \omega t, r \sin \omega t, vt), \quad t \in [0, +\infty)$$

是一条圆柱螺线,它是点 P (起始位置为 $(r, 0, 0)$,距 z 轴始终为 r) 以等角速度 ω 绕 z 轴旋转,以等速度 v 沿 z 轴方向移动的轨迹.

在平面直角坐标系 Oxy 中,向量函数

$$\mathbf{r}(t) = (x(t), y(t)), \quad t \in [a, b]$$

表示一条平面曲线,例如

$$\mathbf{r}(t) = (r \cos t, r \sin t), \quad t \in [0, 2\pi]$$

表示以原点为中心,以 r 为半径的圆.

$$\mathbf{r}(\theta) = (a\theta \cos \theta, a\theta \sin \theta), \quad \theta \in [0, +\infty), a > 0$$

表示的曲线叫阿基米德螺线,其极坐标方程为

$$r = |\mathbf{r}| = \sqrt{(a\theta \cos \theta)^2 + (a\theta \sin \theta)^2} = a\theta.$$

同一条曲线的参数方程一般不是唯一的.例如 $\mathbf{r}(t) = (a \cos t, b \sin t), t \in [0, 2\pi]$, 以及 $\mathbf{r}(u) = (a \cos ku, b \sin ku)$ (其中 k 为任意非零常数, $0 \leq ku \leq 2\pi$), 都表示一个椭圆.

9.1.2 向量函数的分析性质

设向量函数 $\mathbf{r}(t) = (x(t), y(t), z(t))$ 在 t_0 的 η 邻域 $B_\eta(t_0, \eta)$ 有定义, $\mathbf{r}_0 = (x_0, y_0, z_0)$ 则

$$\lim_{t \rightarrow t_0} \mathbf{r}(t) = \mathbf{r}_0 \quad (9-2)$$

指的是: $\forall \varepsilon > 0, \exists \delta > 0 (\delta < \eta), \forall t \in B_\delta(t_0, \delta)$ 即 $t_0 - \delta < t < t_0 + \delta$, 恒有 $\|\mathbf{r}(t) - \mathbf{r}_0\| < \varepsilon$.

(9-2)式等价于

$$\begin{aligned} & \lim_{t \rightarrow t_0} \|\mathbf{r}(t) - \mathbf{r}_0\| \\ &= \lim_{t \rightarrow t_0} \sqrt{(x(t) - x_0)^2 + (y(t) - y_0)^2 + (z(t) - z_0)^2} = 0, \end{aligned}$$

因此(9-2)式成立的充分必要条件是

$$\lim_{t \rightarrow t_0} x(t) = x_0, \quad \lim_{t \rightarrow t_0} y(t) = y_0, \quad \lim_{t \rightarrow t_0} z(t) = z_0. \quad (9-3)$$

若向量函数 $\mathbf{r}(t)$ 在 $B(t_0, \eta)$ 有定义, 而且 $\lim_{t \rightarrow t_0} \mathbf{r}(t) = \mathbf{r}(t_0)$, 则称 $\mathbf{r}(t)$ 在 t_0 处连续. 由(9-3)式, 向量函数 $\mathbf{r}(t)$ 在 t_0 连续的充要条件是它的分量 $x(t), y(t), z(t)$ 在 t_0 连续. 同样若数值函数 $\lambda(t)$ 和向量函数 $\mathbf{r}(t), \mathbf{r}_1(t), \mathbf{r}_2(t)$ 在 t_0 连续, 则 $\lambda(t)\mathbf{r}(t), \mathbf{r}_1(t) \pm \mathbf{r}_2(t), \mathbf{r}_1(t) \cdot \mathbf{r}_2(t), \mathbf{r}_1(t) \times \mathbf{r}_2(t)$ 都在 t_0 连续.

在区间的端点, 向量函数的极限和连续的概念与数值函数一样, 也是单侧的. 若向量函数 $\mathbf{r}(t)$ 在某个区间上每点都连续, 则称 $\mathbf{r}(t)$ 在该区间上连续.

向量函数(9-1)在点 t_0 处的微分为

$$d\mathbf{r}(t_0) = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right)_{t_0}^T dt, \quad (9-4)$$

其中 $\mathbf{r}(t)$ 在 t_0 处的雅可比矩阵的转置 $\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right)_{t_0}$ 就是 $\mathbf{r}(t)$ 在 t_0 处的导数,

记作 $\left. \frac{d\mathbf{r}}{dt} \right|_{t_0}$ 或 $\dot{\mathbf{r}}(t_0)$. 相应地 $\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right)$ 也常记作 $(\dot{x}(t), \dot{y}(t), \dot{z}(t))$ 即

$$\dot{\mathbf{r}}(t_0) = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{r}(t_0 + \Delta t) - \mathbf{r}(t_0)}{\Delta t} = (\dot{x}(t_0), \dot{y}(t_0), \dot{z}(t_0)). \quad (9-5)$$

向量函数 $\mathbf{r}(t)$ 在可微区间 $[a, b]$ 上的导函数也是 $\mathbf{R}^1 \rightarrow \mathbf{R}^3$ 的映射, 它不是数量, 而是一个三维向量, 所以严格地说成 $\mathbf{r}(t)$ 的导向量函数(简称导向量).

(9-5) 式中的导向量习惯上不用雅可比矩阵的形式(即不写作列向量形式).

下面看导向量的几何意义. 设(9-1)式向量函数 $\mathbf{r}(t)$ 在 $[a, b]$ 上可微, 它对应的连续曲线为 C (如图 9-2). $t_0, t_0 + \Delta t \in [a, b], \mathbf{r}(t_0), \mathbf{r}(t_0 + \Delta t)$ 为对应曲线 C 上两个点, 它们决定一个向量 $\Delta \mathbf{r}$, 即

$$\Delta \mathbf{r} = \mathbf{r}(t_0 + \Delta t) - \mathbf{r}(t_0).$$

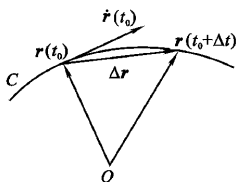


图 9-2

当 $\Delta t > 0$ 时, $\frac{\Delta \mathbf{r}}{\Delta t}$ 与 $\Delta \mathbf{r}$ 同向; 当 $\Delta t < 0$ 时, $\frac{\Delta \mathbf{r}}{\Delta t}$ 与 $\Delta \mathbf{r}$ 反向. 因此, 当 $\Delta t \rightarrow 0$ 时, 如果 $\frac{\Delta \mathbf{r}}{\Delta t}$ 的极限 $\dot{\mathbf{r}}(t_0)$ 为非零向量, 它就是曲线 C 在点 t_0 处的切向量, 其正方向指向 t 增加的方向(如图 9-2 所示, 图中 $\Delta t > 0$).

如果 $\dot{\mathbf{r}}(t_0) \neq 0$, 则称 t_0 点是曲线 C 的正则点, 否则称为奇点. 当曲线 C

上的点都是正则点时,就称 C 为正则曲线,当 $r(t)$ 在 $[a, b]$ 上处处有不等于零的连续导函数时,曲线 C 称为光滑曲线.

$\dot{r}(t_0) \neq 0$ 是曲线 C 在点 t_0 处切线存在的充分条件,但非必要条件,例如

$$r = (t^2, t^3, 0)$$

是半三次抛物线 $\begin{cases} y^2 = x^3 \\ z = 0 \end{cases}$ 的参数方程.在 $t = 0$ 处(即在原点 O 处), $\dot{r}(0) = 0$,故原点是曲线的奇点,但曲线在原点有切线 $y = 0, z = 0$ (见图 9-3).

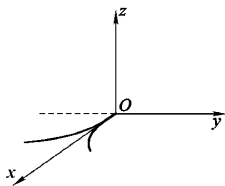


图 9-3

向量函数具有下列求导法则:

$$(1) \frac{d}{dt} \lambda r = \frac{d\lambda}{dt} r + \lambda \frac{dr}{dt};$$

$$(2) \frac{d}{dt} (r_1 \pm r_2) = \frac{dr_1}{dt} \pm \frac{dr_2}{dt};$$

$$(3) \frac{d}{dt} (r_1 \cdot r_2) = \frac{dr_1}{dt} \cdot r_2 + r_1 \cdot \frac{dr_2}{dt}, \quad \text{特别} \frac{d}{dt} r^2 = 2r \cdot \frac{dr}{dt};$$

$$(4) \frac{d}{dt} (r_1 \times r_2) = \frac{dr_1}{dt} \times r_2 + r_1 \times \frac{dr_2}{dt};$$

$$(5) \frac{d}{dt} (r_1, r_2, r_3) = \left(\frac{dr_1}{dt}, r_2, r_3 \right) + \left(r_1, \frac{dr_2}{dt}, r_3 \right) + \left(r_1, r_2, \frac{dr_3}{dt} \right);$$

$$(6) \text{对复合向量函数 } r = r(t), t = t(u), \text{ 有 } \frac{dr}{du} = \frac{dr}{dt} \frac{dt}{du}.$$

其中 $\lambda = \lambda(t)$ 是数值函数, $(r_1, r_2, r_3) = r_1 \cdot (r_2 \times r_3)$ 是三个向量函数的混合积.注意在向量积求导时,由于向量积和向量的次序有关,因此(4)(5)的次序不能随意交换.

向量函数 $r(t)$ 的导数称为 $r(t)$ 的二阶导数,记作 $\ddot{r}(t)$,更高阶的导数依此类推,我们有

$$r^{(n)}(t) = (x^{(n)}(t), y^{(n)}(t), z^{(n)}(t)).$$

若 $r^{(n)}(t)$ 连续, 则称 $r(t)$ 是 $C^{(n)}$ 类的向量函数, 记为 $r(t) \in C^{(n)}$. $r(t) \in C^{(n)}$ 的充要条件是它的每一个分量都属于 $C^{(n)}$, $n=0$ 表示 $r(t)$ 是连续的向量函数.

以后在讨论向量函数时, 我们总是假定 $r(t)$ 在所需要的阶数内是可微的.

我们知道向量函数也有泰勒公式: 设向量函数 $r(t) \in C^{(n)}[a, b]$, 则它的三个分量(数值函数)在点 $t \in (a, b)$ 处可展成 $(n-1)$ 阶泰勒公式, 即

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{\ddot{x}(t)}{2!}(\Delta t)^2 + \dots + \frac{x^{(n)}(\xi_1)}{n!}(\Delta t)^n,$$

$$y(t + \Delta t) = y(t) + \dot{y}(t)\Delta t + \frac{\ddot{y}(t)}{2!}(\Delta t)^2 + \dots + \frac{y^{(n)}(\xi_2)}{n!}(\Delta t)^n,$$

$$z(t + \Delta t) = z(t) + \dot{z}(t)\Delta t + \frac{\ddot{z}(t)}{2!}(\Delta t)^2 + \dots + \frac{z^{(n)}(\xi_3)}{n!}(\Delta t)^n,$$

其中 ξ_1, ξ_2, ξ_3 是三个在 t 与 $t + \Delta t$ 之间的值. 以上三式等价于下面向量函数的式子:

$$\begin{aligned} r(t + \Delta t) = & r(t) + \dot{r}(t)\Delta t + \frac{1}{2!}\ddot{r}(t)(\Delta t)^2 + \dots \\ & + \frac{r^{(n-1)}(t)}{(n-1)!}(\Delta t)^{n-1} + R_{n-1}(\Delta t)^n, \end{aligned} \quad (9-6)$$

其中

$$R_{n-1} = \frac{1}{n!} [x^{(n)}(\xi_1), y^{(n)}(\xi_2), z^{(n)}(\xi_3)].$$

(9-6)式称为向量函数 $r(t)$ 在点 $t \in (a, b)$ 处的 $n-1$ 阶泰勒公式. 注意其中 R_{n-1} 一般不能写成 $\frac{1}{n!}r^{(n)}(\xi)$ (ξ 在 $t, t + \Delta t$ 之间)的形式, 因为 R_{n-1} 中三个分量分别取值于 ξ_1, ξ_2, ξ_3 , 而它们一般是不相等的, 这和数量函数的泰勒公式不同.

由于 $r(t) \in C^{(n)}$, $x^{(n)}(t), y^{(n)}(t), z^{(n)}(t)$ 都是 t 的连续函数, 而 ξ_1, ξ_2, ξ_3 又都在 t 和 $t + \Delta t$ 之间, 所以

$$\lim_{\Delta t \rightarrow 0} R_{n-1} = \frac{1}{n!} \lim_{\Delta t \rightarrow 0} [x^{(n)}(\xi_1), y^{(n)}(\xi_2), z^{(n)}(\xi_3)] = \frac{1}{n!} r^{(n)}(t)$$

因而(9-6)式还可以写成

$$r(t + \Delta t) = r(t) + \dot{r}(t)\Delta t + \dots + \frac{1}{n!}r^{(n)}(t)(\Delta t)^n + \varepsilon(\Delta t)^n \quad (9-7)$$

其中 ε 是无穷小向量, 即 $\lim_{\Delta t \rightarrow 0} \varepsilon = 0$.

9.1.3 向量函数的积分

对于给定的向量函数 $\mathbf{r}(t)$, 若存在向量函数 $\mathbf{R}(t)$, 使对区间 $[a, b]$ 上的每个 t , 都有

$$\dot{\mathbf{R}}(t) = \mathbf{r}(t),$$

则称 $\mathbf{R}(t)$ 是区间 $[a, b]$ 上 $\mathbf{r}(t)$ 的一个原向量函数, 简称原函数. 若 $\mathbf{R}(t)$ 是 $\mathbf{r}(t)$ 的一个原函数, 则 $\mathbf{r}(t)$ 的任何原函数与 $\mathbf{R}(t)$ 只差一常向量 \mathbf{C} . $\mathbf{r}(t)$ 的全体原函数称为 $\mathbf{r}(t)$ 的不定积分, 记成

$$\int \mathbf{r}(t) dt = \mathbf{R}(t) + \mathbf{C}.$$

下列不定积分的运算法则, 读者可以自行证明.

$$(1) \int \mathbf{r}(t) dt = \left(\int x(t) dt, \int y(t) dt, \int z(t) dt \right);$$

$$(2) \int (\mathbf{r}_1(t) \pm \mathbf{r}_2(t)) dt = \int \mathbf{r}_1(t) dt \pm \int \mathbf{r}_2(t) dt;$$

$$(3) \int a \varphi(t) dt = a \int \varphi(t) dt;$$

$$(4) \int \mathbf{a} \cdot \mathbf{r}(t) dt = \mathbf{a} \cdot \int \mathbf{r}(t) dt;$$

$$(5) \int \mathbf{a} \times \mathbf{r}(t) dt = \mathbf{a} \times \int \mathbf{r}(t) dt.$$

其中 $\varphi(t)$ 是数量函数, \mathbf{a} 是常向量.

同样我们可以定义向量函数在区间 $[a, b]$ 上的定积分:

$$\int_a^b \mathbf{r}(t) dt = \left(\int_a^b x(t) dt, \int_a^b y(t) dt, \int_a^b z(t) dt \right).$$

设 $\mathbf{R}(t)$ 是 $\mathbf{r}(t)$ 的一个原函数, 同样有牛顿—莱布尼兹公式

$$\int_a^b \mathbf{r}(t) dt = \mathbf{R}(t) \Big|_a^b = \mathbf{R}(b) - \mathbf{R}(a)$$

9.1.4 三个特殊的向量函数

下面介绍三个特殊的向量函数(或特殊的空间曲线). 一方面为后面的内容作必要的准备, 另一方面也给读者指出一种方法, 就是利用向量函数及其导数所满足的代数关系, 来判别、研究向量函数 $\mathbf{r}(t)$ 的几何性质.

(1) 定长向量函数

若 $\|\mathbf{r}(t)\| = \text{常数}$, 则称 $\mathbf{r}(t)$ 为定长向量函数. 由于 $r^2(t) = \|\mathbf{r}(t)\|^2 = \text{常数}$, 对 t 求导, 即得 $\mathbf{r}(t)$ 为定长向量函数的充要条件

$$\mathbf{r}(t) \cdot \dot{\mathbf{r}}(t) = 0. \quad (9-8)$$

定长向量函数 $\mathbf{r} = \mathbf{r}(t)$ 的图形是一条位于以原点为中心的球面上的曲线. (9-8) 式

说明, 曲线在球面上的充要条件是, 它在每点处的切向量与该点的向径向量垂直.

(2) 定向向量函数

与一固定方向平行的非零向量函数 $\mathbf{r}(t)$ 称为定向向量函数, 即

$$\mathbf{r}(t) = \lambda(t) \mathbf{e}, \quad (9-9)$$

其中 \mathbf{e} 是一固定方向的单位向量, $\lambda(t) \neq 0$ 是一个数值函数, 显然, $|\lambda(t)| = \|\mathbf{r}(t)\|$.

容易证明, $\mathbf{r}(t)$ 为定向向量函数的充要条件是

$$\mathbf{r}(t) \times \dot{\mathbf{r}}(t) = \mathbf{0}, \quad (9-10)$$

即 $\dot{\mathbf{r}}$ 与 \mathbf{r} 平行.

(3) 与固定向量垂直的向量函数

与一非零定向向量 \mathbf{a} 垂直的非零向量函数 $\mathbf{r}(t)$ 称为与一定向垂直(或平行于固定平面)的向量函数. 向量函数 $\mathbf{r}(t)$ 与一定向垂直的充要条件是混合积为零

$$(\mathbf{r}(t), \dot{\mathbf{r}}(t), \ddot{\mathbf{r}}(t)) = 0. \quad (9-11)$$

结论的必要性是明显的, 只需证充分性. 如果 $(\mathbf{r}(t), \dot{\mathbf{r}}(t), \ddot{\mathbf{r}}(t)) = 0$, 不妨设 $\mathbf{r}(t)$ 与 $\dot{\mathbf{r}}(t)$ 线性无关, 即 $\mathbf{r} \times \dot{\mathbf{r}} \neq \mathbf{0}$ (否则 $\mathbf{r}(t)$ 为定向向量函数, 它与一定向垂直), 于是存在 $\lambda, \mu \in \mathbf{R}$ 使得

$$\ddot{\mathbf{r}} = \lambda \mathbf{r} + \mu \dot{\mathbf{r}}.$$

记 $\mathbf{a} = \mathbf{r} \times \dot{\mathbf{r}}$, 它与 \mathbf{r} 垂直, 而且

$$\mathbf{a} \times \dot{\mathbf{a}} = (\mathbf{r} \times \dot{\mathbf{r}}) \times (\dot{\mathbf{r}} \times \dot{\mathbf{r}} + \mathbf{r} \times \ddot{\mathbf{r}}) = \mathbf{0},$$

所以 \mathbf{a} 为定向向量函数, 而 $\mathbf{r}(t)$ 是与定向 \mathbf{a} 垂直的向量函数.

9.2 曲线的弧长和弗雷耐标架

9.2.1 曲线的弧长

设有光滑曲线 $C: \mathbf{r} = \mathbf{r}(t), a \leq t \leq b$, 我们再一次定义和计算曲线 C 的弧长.

将区间 $[a, b]$ 分成 n 份: $a = t_0, t_0 < t_1 < t_2 < \dots < t_n = b$, 则曲线 C 上的对应分点是 $\mathbf{r}(t_i)$, 相邻分点之间的线段长为 $\|\mathbf{r}(t_{i-1}) - \mathbf{r}(t_i)\|$, 记 $\Delta t_i = t_i - t_{i-1}$, $s = \max_{1 \leq i \leq n} |\Delta t_i|$. 如果极限

$$\lim_{s \rightarrow 0} \sum_{i=1}^n \|\mathbf{r}(t_{i-1}) - \mathbf{r}(t_i)\|$$

存在, 且不依赖于分法, 则称这个极限为曲线 C 在 $a \leq t \leq b$ 上所对应的曲线段的弧长 s .

我们知道, 这个极限等于 $\|\dot{\mathbf{r}}(t)\|$ 在 $[a, b]$ 上的定积分, 即

$$s = \lim_{s \rightarrow 0} \sum_{i=1}^n \|\mathbf{r}(t_{i-1}) - \mathbf{r}(t_i)\| = \int_a^b \|\dot{\mathbf{r}}(t)\| dt.$$

设 $t_0, t \in [a, b]$, 记 $s(t)$ 为 $r(t_0)$ 到 $r(t)$ 的弧长, 并规定 $t > t_0$ 时, $s(t) > 0$, $t < t_0$ 时 $s(t) < 0$. 则 $s(t) = \int_{t_0}^t \|\dot{r}(t)\| dt$ 是 t 的增函数, 而且

$$s(t) = \int_{t_0}^t \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} dt. \quad (9-12)$$

由 (9-12) 容易推出平面曲线在直角坐标系和极坐标系中的弧长公式:

$$s(x) = \int_{x_0}^x \sqrt{1 + [y'(x)]^2} dx,$$

$$s(\theta) = \int_{\theta_0}^{\theta} \sqrt{\rho^2(\theta) + [\rho'(\theta)]^2} d\theta.$$

我们称

$$ds = \|\dot{r}(t)\| dt = \sqrt{\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t)} dt \quad (9-13)$$

为弧微分, 若 $\dot{r}(t) \neq 0$, 则

$$\frac{ds}{dt} = \|\dot{r}(t)\| = \left\| \frac{dr}{dt} \right\| > 0. \quad (9-13')$$

这表示 $s(t)$ 的反函数 $t = t(s)$ 存在, 代入 (9-1), 得到以弧长 s 为参数的曲线 C 的方程

$$r = r(s) = (x(s), y(s), z(s)). \quad (9-1')$$

弧长参数 s 也称为曲线的自然参数.

例 1 R^3 中的直线方程为

$$r = at + b \quad (-\infty < t < +\infty),$$

其中 $a \neq 0$, b 是常向量, 求直线上由点 $t = 0$ 到点 t 之间的弧长 s 及直线的弧长参数表达式.

解 直线的方向向量为

$$\frac{dr}{dt} = a,$$

所求弧长是

$$s = \int_0^t \left\| \frac{dr}{dt} \right\| dt = \int_0^t \|a\| dt = \|a\| t \quad (-\infty < t < +\infty),$$

由此解得反函数

$$t = \frac{1}{\|a\|} s.$$

代入原直线方程, 即得直线由其弧长作为参数的方程

$$\mathbf{r} = s \frac{\mathbf{a}}{\|\mathbf{a}\|} + \mathbf{b} = se + \mathbf{b}$$

其中 e 表示直线方向的单位向量.

例 2 求圆柱螺线 $\mathbf{r} = (a \cos t, a \sin t, bt)$ 从 $t = 0$ 到 t 的弧长 s , 并求圆柱螺线的弧长参数表达式.

解 圆柱螺线在点 t 处的切向量是

$$\frac{d\mathbf{r}}{dt} = (-a \sin t, a \cos t, b),$$

所求弧长为

$$s = \int_0^t \left\| \frac{d\mathbf{r}}{dt} \right\| dt = \int_0^t \sqrt{a^2 + b^2} dt = \sqrt{a^2 + b^2} t,$$

由此我们得到圆柱螺线由自然参数 s 表示的方程

$$\mathbf{r} = \left(a \cos \frac{s}{\sqrt{a^2 + b^2}}, a \sin \frac{s}{\sqrt{a^2 + b^2}}, \frac{bs}{\sqrt{a^2 + b^2}} \right).$$

注意 $\dot{\mathbf{r}}$ 的第三个分量是常数 b , 与参数 t 无关, 故圆柱螺线的切向量和 z 轴的夹角 θ 的余弦为

$$\cos \theta = \dot{\mathbf{r}} \cdot \mathbf{k} = b$$

这就是说: 圆柱螺线上任一点的切线与 z 轴交成定角.

用具有明确几何意义的弧长 s 代替原来的任意参数 t , 会给问题的讨论带来很大的方便, 特别是理论问题. 用弧长作参数, 将使一些公式大为简化, 但在具体问题上, 要把一条原来由参数 t 表示的曲线规范化为由其弧长参数来表达, 可不是一件容易的事, 这是因为弧长的积分往往不好计算(如椭圆弧长的积分不能表示为初等函数). 由 $s = s(t)$ 求反函数 $t = t(s)$ 也不总是容易的. 在以下推导的公式中我们都用弧长作为参数, 至于一般参数 t 的相应公式可见后面的表格(9-1). 以下用“ \cdot ”表示对一般参数 t 求导, 以撇“ $'$ ”表示对弧长参数 s 求导, 即

$$\dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt}, \quad \mathbf{r}' = \frac{d\mathbf{r}}{ds}.$$

9.2.2 曲线的切线和法平面

设 \mathbf{R}^3 中曲线 C 用自然参数表示为

$$\mathbf{r} = \mathbf{r}(s) = (x(s), y(s), z(s)), \quad a \leq s \leq b, \quad (9-1)$$

$s_0 \in [a, b]$ 为 C 的正则点. 由图 9-2 及式(9-5)得 C 在点 s_0 的切线向量

$$\mathbf{T} = \mathbf{r}'(s_0) \quad (\|\mathbf{T}\| = 1),$$

切线方程为

$$\boldsymbol{\rho} = \boldsymbol{r}(s_0) + \lambda \boldsymbol{T} = \boldsymbol{r}(s_0) + \lambda \boldsymbol{r}'(s_0), \quad (9-14)$$

其中 $\boldsymbol{\rho}(x, y, z)$ 是切线上点的向径, λ 为参数.

过点 $\boldsymbol{r}(s_0)$ 与切线垂直的直线称为 C 在 $\boldsymbol{r}(s_0)$ 的法线, 与切线垂直的平面称为 C 在 $\boldsymbol{r}(s_0)$ 法平面, 法平面方程为

$$(\boldsymbol{\rho} - \boldsymbol{r}(s_0)) \cdot \boldsymbol{r}'(s_0) = 0, \quad (9-15)$$

其中 $\boldsymbol{\rho}(x, y, z)$ 是法平面上点 (x, y, z) 的向径.

若曲线 C 以一般参数 t 表示为 $\boldsymbol{r} = \boldsymbol{r}(t)$, 记 $t = t(s)$ (s 为弧长参数) 则

$$\boldsymbol{T} = \frac{d\boldsymbol{r}}{ds} = \boldsymbol{r}'(s) = \frac{\frac{d\boldsymbol{r}}{dt}}{\frac{ds}{dt}} = \frac{\dot{\boldsymbol{r}}(t)}{\|\dot{\boldsymbol{r}}(t)\|}.$$

此时曲线在 $\boldsymbol{r}(t_0)$ 处的切线方程和法平面方程 (与 (9-14) 和 (9-15) 形式相似) 分别为

$$\boldsymbol{\rho} = \boldsymbol{r}(t_0) + \lambda \dot{\boldsymbol{r}}(t_0), \quad (9-14')$$

$$\dot{\boldsymbol{r}}(t_0) \cdot (\boldsymbol{\rho} - \boldsymbol{r}(t_0)) = 0. \quad (9-15')$$

例3 求螺旋线 $\boldsymbol{r} = (a \cos \varphi, a \sin \varphi, b\varphi)$ 在点 $\varphi = 0$ 处的切线和法平面.

解 $\varphi = 0$ 时, $\boldsymbol{r}(0) = (a, 0, 0)$,

$$\dot{\boldsymbol{r}}(\varphi) = (-a \sin \varphi, a \cos \varphi, b),$$

$$\dot{\boldsymbol{r}}(0) = (0, a, b).$$

因此在 $\varphi = 0$ 处螺旋线的切线方程为

$$\boldsymbol{\rho} = \boldsymbol{r}(0) + \lambda \dot{\boldsymbol{r}}(0) = (a, a\lambda, b\lambda), \quad \lambda \text{ 是参数,}$$

或消去 λ ,

$$\frac{x-a}{0} = \frac{y-0}{a} = \frac{z-0}{b}.$$

法平面方程为

$$\dot{\boldsymbol{r}}(0) \cdot (\boldsymbol{\rho} - \boldsymbol{r}(0)) = 0,$$

即

$$ay + bz = 0.$$

9.2.3 密切平面和副法线

我们知道曲线上正则点有唯一的切线和法平面, 但却有很多切平面和法线. 经过曲线 C 上点 $\boldsymbol{r}(s_0)$ 的切线的平面称为切平面, C 在 $\boldsymbol{r}(s_0)$ 处有一族切

平面, 其中有一个最贴近 C 的平面可以如下得到: 过 C 上点 $\mathbf{r}(s_0)$ 的切线及其邻近点 $\mathbf{r}(s_0 + \Delta s)$ 作一平面 Π' . 当 $\Delta s \rightarrow 0$ 时, 若 Π' 有极限位置 Π , 则称 Π 为 C 在点 $\mathbf{r}(s_0)$ 的密切平面. 如果 $\mathbf{r} = \mathbf{r}(s)$ 是平面曲线, 则其上任一点的密切平面就是曲线所在平面. 密切平面的法线称为 C 在 $\mathbf{r}(s_0)$ 处的副法线, 单位副法线向量以 \mathbf{B} 表示, 则

$$\mathbf{B} = \frac{\mathbf{r}'(s_0) \times \mathbf{r}''(s_0)}{\|\mathbf{r}''(s_0)\|}, \quad (9-16)$$

这是由于 Π' 的法向量平行于向量

$$\begin{aligned} & \mathbf{r}'(s_0) \times (\mathbf{r}(s_0 + \Delta s) - \mathbf{r}(s_0)) \\ &= \mathbf{r}'(s_0) \times \left(\mathbf{r}'(s_0) \Delta s + \frac{1}{2} (\mathbf{r}''(s_0) + \boldsymbol{\varepsilon}) \Delta s^2 \right), \end{aligned}$$

即向量

$$\mathbf{r}'(s_0) \times (\mathbf{r}''(s_0) + \boldsymbol{\varepsilon}),$$

其中 $\lim_{\Delta s \rightarrow 0} \boldsymbol{\varepsilon} = 0$. 因此 $\mathbf{B} \parallel \mathbf{r}'(s_0) \times \mathbf{r}''(s_0)$. 再利用 $\|\mathbf{r}'(s_0)\| = \|\mathbf{T}\| = 1$ 和 $\mathbf{r}'(s_0) \perp \mathbf{r}''(s_0)$ (见(9-8)), 即得(9-16).

副法线方程为

$$\boldsymbol{\rho} = \mathbf{r}(s_0) + \lambda \mathbf{B} \quad (9-17)$$

或

$$\boldsymbol{\rho} = \mathbf{r}(s_0) + \lambda (\mathbf{r}'(s_0) \times \mathbf{r}''(s_0)). \quad (9-17')$$

密切平面方程为

$$(\boldsymbol{\rho} - \mathbf{r}(s_0)) \cdot \mathbf{B} = 0 \quad (9-18)$$

或

$$(\mathbf{r}'(s_0), \mathbf{r}''(s_0), \boldsymbol{\rho} - \mathbf{r}(s_0)) = 0. \quad (9-18')$$

对曲线 C 的一般参数方程 $\mathbf{r} = \mathbf{r}(t)$, 它在 $\mathbf{r}(t_0)$ 处的单位副法线向量为

$$\mathbf{B} = \frac{\dot{\mathbf{r}}(t_0) \times \ddot{\mathbf{r}}(t_0)}{\|\dot{\mathbf{r}}(t_0) \times \ddot{\mathbf{r}}(t_0)\|}. \quad (9-16')$$

副法线方程和密切平面方程形同(9-17)(9-18).

例 4 求圆锥螺线 $\mathbf{r} = (t \cos t, -t \sin t, at)$ 在坐标原点处的密切平面方程和副法线方程.

解 坐标原点对应于 $t = 0$,

$$\dot{\mathbf{r}} = (\cos t - t \sin t, -\sin t - t \cos t, a),$$

$$\dot{\mathbf{r}}(0) = (1, 0, a),$$

$$\ddot{\mathbf{r}} = (-2 \sin t - t \cos t, -2 \cos t + t \sin t, 0),$$

$$\ddot{\mathbf{r}}(0) = (0, -2, 0),$$

所以,螺旋线在点处的密切平面方程为

$$(\rho - \mathbf{r}(0), \dot{\mathbf{r}}(0), \ddot{\mathbf{r}}(0)) = 0,$$

即

$$\begin{vmatrix} x & y & z \\ 1 & 0 & a \\ 0 & -2 & 0 \end{vmatrix} = 0 \text{ 或 } -ax + z = 0.$$

螺旋线在点处的副法线方程为

$$\rho = \mathbf{r}(0) + \lambda(\dot{\mathbf{r}}(0) \times \ddot{\mathbf{r}}(0)),$$

即

$$\rho = (a\lambda, 0, -\lambda),$$

其中 λ 是参数.(其实,密切平面的法向量 $(-a, 0, 1)$ 就是副法线的方向向量,

可以直接得出副法线方程 $\frac{x}{-a} = \frac{y}{0} = \frac{z}{1}$.)

9.2.4 主法线和从切平面

在曲线 C 的任一点 $\mathbf{r}(s_0)$ 处,已经有切线和法平面、密切平面和副法线的概念,进而我们把密切平面和法平面的交线称为主法线,过 $\mathbf{r}(s_0)$ 点且以主法线为法线的平面称为从切平面(见图 9-4).

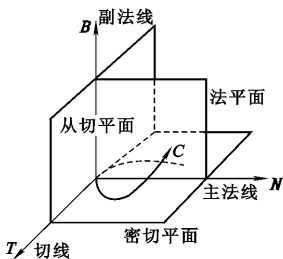


图 9-4

以 \mathbf{N} 表示过 $\mathbf{r}(s_0)$ 的单位主法线向量,取

$$\mathbf{N} = \mathbf{B} \times \mathbf{T}. \quad (9-19)$$

由于 $\mathbf{r}'(s_0) = \mathbf{T}$ 是定长向量,所以 $\mathbf{r}''(s_0) \perp \mathbf{r}'(s_0) = \mathbf{T}$,再由

$$\mathbf{B} = \frac{\mathbf{r}'(s_0) \times \mathbf{r}''(s_0)}{\|\mathbf{r}''(s_0)\|} \perp \mathbf{r}''(s_0),$$

即得 $\mathbf{N} \parallel \mathbf{r}''(s_0)$, 它的单位向量为

$$\mathbf{N} = \frac{\mathbf{r}''(s_0)}{\|\mathbf{r}''(s_0)\|}. \quad (9-19)$$

过 $\mathbf{r}(s_0)$ 的主法线和从切平面的方程分别为

$$\boldsymbol{\rho} = \mathbf{r}(s_0) + \lambda \mathbf{N}, \quad (9-20)$$

$$(\boldsymbol{\rho} - \mathbf{r}(s_0)) \cdot \mathbf{N} = 0. \quad (9-21)$$

若曲线 C 的方程为 $\mathbf{r} = \mathbf{r}(t)$ 则

$$\mathbf{N} = \frac{(\dot{\mathbf{r}}(t_0) \times \ddot{\mathbf{r}}(t_0)) \times \dot{\mathbf{r}}(t_0)}{\|\dot{\mathbf{r}}(t_0) \times \ddot{\mathbf{r}}(t_0)\| \|\dot{\mathbf{r}}(t_0)\|}. \quad (9-19')$$

例 5 求圆柱螺线 $\mathbf{r} = (a \cos t, a \sin t, bt)$ 在任一点 t 处的切线和法平面、密切平面和副法线、主法线和从切平面。

解

$$\mathbf{r} = (a \cos t, a \sin t, bt),$$

$$\dot{\mathbf{r}} = (-a \sin t, a \cos t, b),$$

$$\ddot{\mathbf{r}} = (-a \cos t, -a \sin t, 0).$$

得向量

$$\mathbf{T} // \dot{\mathbf{r}} = (-a \sin t, a \cos t, b),$$

副法线向量

$$\mathbf{B} // \dot{\mathbf{r}} \times \ddot{\mathbf{r}} = (ab \sin t, -ab \cos t, a^2) // (b \sin t, -b \cos t, a),$$

主法线向量

$$\mathbf{N} // (\dot{\mathbf{r}} \times \ddot{\mathbf{r}}) \times \dot{\mathbf{r}}$$

$$= (- (a^2 + b^2) \cos t, - (a^2 + b^2) \sin t, 0) // (\cos t, \sin t, 0).$$

于是得切线方程

$$\frac{x - a \cos t}{-a \sin t} = \frac{y - a \sin t}{a \cos t} = \frac{z - bt}{b},$$

副法线方程

$$\frac{x - a \cos t}{b \sin t} = \frac{y - a \sin t}{-b \cos t} = \frac{z - bt}{a},$$

主法线方程

$$\frac{x - a \cos t}{\cos t} = \frac{y - a \sin t}{\sin t} = \frac{z - bt}{0},$$

法平面方程

$$-a(\sin t)x + a(\cos t)y + bz = b^2t,$$

密切平面方程

$$b(\sin t)x - b(\cos t)y + az = abt,$$

从切平面方程

$$(\cos t)x + (\sin t)y = a.$$

9.2.5 弗雷耐标架

上面我们得到曲线 $C: \mathbf{r} = \mathbf{r}(s)$ 在 $\mathbf{r}(s_0)$ 处的单位切向量 \mathbf{T} , 单位副法线向量 \mathbf{B} 和单位主法线向量 \mathbf{N} , 它们以自然参数 s 表示和以一般参数 t 表示的公式如表(9-1)所示.

在曲线 C 上任一点 $\mathbf{r}(s)$ 处, 都有互相正交的三个单位向量 $\mathbf{T}(s)$, $\mathbf{N}(s)$ 和 $\mathbf{B}(s)$, 且依上述次序构成一个右旋坐标系, 我们把它看成是粘附在 C 上 $\mathbf{r}(s)$ 点处以 $\mathbf{r}(s)$ 为原点的右旋单位正交坐标系, 当 s 在 C 上移动时, 这个坐标系也随着运动, 是个活动的坐标系.

表 9-1

	空间曲线方程 $\mathbf{r} = \mathbf{r}(s)$	空间曲线方程 $\mathbf{r} = \mathbf{r}(t)$
弧长	s	$s = \int_0^t \ \dot{\mathbf{r}}\ dt$
弧微分	ds	$ds = \ \dot{\mathbf{r}}\ dt$
\mathbf{T}	\mathbf{r}'	$\frac{\dot{\mathbf{r}}}{\ \dot{\mathbf{r}}\ }$
\mathbf{N}	$\frac{\mathbf{r}''}{\ \mathbf{r}''\ }$	$\mathbf{B} \times \mathbf{T}$
\mathbf{B}	$\mathbf{T} \times \mathbf{N}$	$\frac{\dot{\mathbf{r}} \times \ddot{\mathbf{r}}}{\ \dot{\mathbf{r}} \times \ddot{\mathbf{r}}\ }$

定义 9.1 曲线 $\mathbf{r} = \mathbf{r}(s)$ 在点 s 处的单位正交向量 \mathbf{T} , \mathbf{N} , \mathbf{B} 按右手法则组成的坐标系, 称为曲线在点 s 处的弗雷耐(Frenet)标架, 记作

$$\{\mathbf{r}(s); \mathbf{T}(s), \mathbf{N}(s), \mathbf{B}(s)\},$$

$$\begin{cases} \mathbf{T} = \mathbf{N} \times \mathbf{B}, \mathbf{N} = \mathbf{B} \times \mathbf{T}, \mathbf{B} = \mathbf{T} \times \mathbf{N}, \\ (\mathbf{T}, \mathbf{N}, \mathbf{B}) = 1. \end{cases}$$

例 6 设圆柱螺线的方程为

$$\mathbf{r} = (a \cos \omega s, a \sin \omega s, b \omega s),$$

式中 $\omega = \frac{1}{\sqrt{a^2 + b^2}}$, $a > 0$, 求曲线在 s 处的弗雷耐标架.

解 $\frac{d\mathbf{r}}{ds} = (-a\omega \sin \omega s, a\omega \cos \omega s, b\omega)$, $\left\| \frac{d\mathbf{r}}{ds} \right\| = 1$ 知 s 为弧长参数, 所以

$$\mathbf{T} = \mathbf{r}' = \omega(-a \sin \omega s, a \cos \omega s, b),$$

$$\mathbf{N} = \frac{\mathbf{T}'}{\|\mathbf{T}'\|} = (-\cos \omega s, -\sin \omega s, 0),$$

$$\mathbf{B} = \mathbf{T} \times \mathbf{N} = \omega(b \sin \omega s, -b \cos \omega s, a).$$

图 9-5(a) 为 $b > 0$ 时的圆柱螺线, 它是右旋的, 因为 $b > 0$, \mathbf{T} 的第三分量 $b\omega$ 是向上倾斜的. \mathbf{N} 的第三分量为零, 它总是平行于水平平面且指向 z 轴. \mathbf{B} 的第三个分量 $\omega a > 0$, 它是向上倾斜的. 图 9-5(b) 为 $b < 0$ 时的圆柱螺线, 它是左旋的.

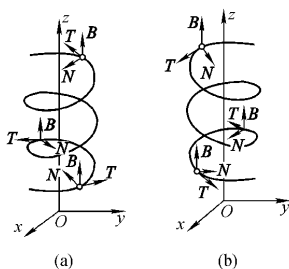


图 9-5

人们在实际生活中,也经常利用活动标架的方法.例如一条船航行在大海上,如果取其前进的方向为 \mathbf{T} ,垂直向上为 \mathbf{B} ,则 $\mathbf{N} = \mathbf{B} \times \mathbf{T}$ 是沿水平指向左舷的向量.这样 $\{r; \mathbf{T}, \mathbf{N}, \mathbf{B}\}$ 就组成了一个标架,由于船本身在前进,因此它们都是时间 t 的函数,是一个活动标架.当水手观察到海面上空一架飞机时,它报出了下列数据:正前方 x 米,左侧 y 米,高度 z 米,这三个数字就是飞机在活动标架 $\{r; \mathbf{T}, \mathbf{N}, \mathbf{B}\}$ 上的三个坐标分量,它们也都是 t 的函数,即使飞机不动,飞机的坐标也是要改变的.当然地面指挥所也能测得飞机的方位(这是固定坐标架下的坐标),通知舰船.但打仗时,对舰船有用的数据正是敌机在活动标架中的坐标 x, y, z .

9.3 曲线的曲率 挠率 弗雷耐公式

9.3.1 曲线的曲率

在很多实际问题中,例如,设计拐弯的铁道、弯曲的梁、曲线形的刀具,都必须考虑曲线的弯曲程度,数学上定量描述空间曲线弯曲程度的概念叫曲线

的曲率. 它应该如何定义呢?

设光滑曲线 C 的方程为 $r = r(s)$, s 为自然参数. 当 s 由 s_0 变到 $s_0 + \Delta s$ 时, 切向量 $T(s_0)$ 与 $T(s_0 + \Delta s)$ 的正向夹角为 $\Delta\theta$ (如图 9-6 所示). 显然, $\left| \frac{\Delta\theta}{\Delta s} \right|$ 越大, 则曲线 C 在 AB 弧段上平均弯曲程度越大. 于是, C 在点 A (或说点 s_0) 处的弯曲程度可以用 $\lim_{\Delta s \rightarrow 0} \left| \frac{\Delta\theta}{\Delta s} \right|$ 来描述. $\left| \frac{\Delta\theta}{\Delta s} \right|$ 越大, 等价于 $\left\| \frac{\Delta T}{\Delta s} \right\| = \left\| \frac{T(s_0 + \Delta s) - T(s_0)}{\Delta s} \right\|$ 越大. 因此, C 在点 s_0 的曲率可以用 $\lim_{\Delta s \rightarrow 0} \left\| \frac{\Delta T}{\Delta s} \right\| = \|T'(s_0)\| = \|r''(s_0)\|$ 来描述. 即有下面的定义.

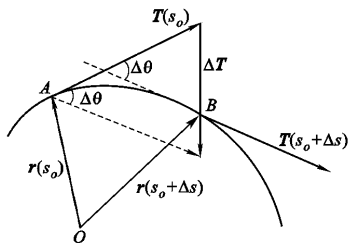


图 9-6

定义 9.2 光滑曲线 $r = r(s)$ 对自然参数 s 的二阶导向量的模 $\|r''(s)\|$ 称为曲线 $r(s)$ 在点 s 处的曲率, 记作 $k(s)$, 即

$$k(s) = \|r''(s)\|. \quad (9-22)$$

当 $k(s) \neq 0$ 时, 其倒数称为曲线 $r(s)$ 在点 s 处的曲率半径, 记作 $R(s)$, 即

$$R(s) = \frac{1}{k(s)}. \quad (9-23)$$

在弗雷耐标架中, 由于 $T = r'$, $T' = r''$ 而 $N = \frac{r''}{\|r''\|}$, 故

$$T'(s) = \|r''(s)\| N(s) = k(s) N(s).$$

我们称向量 $k(s)N(s)$ 为曲线 $r(s)$ 在点 s 处的曲率向量. 由向量 $r(s) + R(s)N(s)$ 表示的点 Q 称为曲线在点 s 处的曲率中心. 在点 s 的密切平面上, 以曲率中心 Q 为圆心, 曲率半径 $R(s)$ 为半径作圆 M , 称 M 为曲线在点 s 处的密切圆 (见图 9-7).

显然, 密切圆与曲线在点 s 相切, 而且在点 s 它们有相同的曲率. 工程中

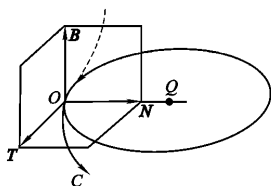


图 9-7

常常用密切圆的一段弧来近似代替点 s 邻近的一段曲线.

总结以上的讨论,可以归结为下面的引理:

引理 设 $e(t)$ 是单位向量函数, $\Delta e = e(t + \Delta t) - e(t)$, 而 $\Delta\theta$ 为 $e(t + \Delta t)$ 和 $e(t)$ 的夹角(见图(9-8)), 则

$$\left\| \frac{de(t)}{dt} \right\| = \lim_{\Delta t \rightarrow 0} \left| \frac{\Delta\theta}{\Delta t} \right|.$$

证 按照向量函数导数的定义, 有

$$\left\| \frac{de(t)}{dt} \right\| = \lim_{\Delta t \rightarrow 0} \left| \frac{\Delta e}{\Delta t} \right| = \lim_{\Delta t \rightarrow 0} \frac{\left| 2 \sin \frac{\Delta\theta}{2} \right|}{|\Delta t|} = \lim_{\Delta t \rightarrow 0} \left| \frac{\Delta\theta}{\Delta t} \right|.$$

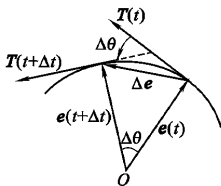


图 9-8

由曲率的定义及上述引理, 得

$$k(s) = \|r''(s)\| = \left\| \frac{dT(s)}{ds} \right\| = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta\theta}{\Delta s} \right|, \quad (9-24)$$

式中 $\Delta\theta$ 是 $T(s + \Delta s)$ 和 $T(s)$ 的夹角. 所以, 曲线在一点处的曲率就是曲线在该点的切向量倾角对其弧长的变化率, 或称切线对弧长的转动率.

例 7 求圆柱螺线 $r = (a \cos t, a \sin t, bt)$ 的曲率.

解

$$\begin{aligned}
 \mathbf{T} &= \frac{d\mathbf{r}}{ds} = \frac{\dot{\mathbf{r}}(t)}{\|\dot{\mathbf{r}}(t)\|} = \frac{1}{\sqrt{a^2 + b^2}}(-a \sin t, a \cos t, b), \\
 \frac{d\mathbf{T}}{ds} &= \frac{\frac{d\mathbf{T}}{dt}}{\frac{ds}{dt}} = \frac{1}{\sqrt{a^2 + b^2}} \frac{1}{\sqrt{a^2 + b^2}}(-a \cos t, -a \sin t, 0) \\
 &= \frac{a}{a^2 + b^2}(-\cos t, -\sin t, 0),
 \end{aligned}$$

故

$$k = \left\| \frac{d\mathbf{T}}{ds} \right\| = \frac{a}{a^2 + b^2}.$$

例 8 证明:空间曲线为直线的充要条件是每点的曲率为 0.

证 只证必要性. 设直线方程为 $\mathbf{r} = s\mathbf{e} + \mathbf{b}$, 有 $\mathbf{r}'' = 0$, 即 $k(s) = \|\mathbf{r}''\| \equiv 0$. 反之设 $k = \|\mathbf{r}''(s)\| \equiv 0$, 由 $\mathbf{r}''(s) = 0$ 解得 $\mathbf{r}(s) = a s + \mathbf{b}$ (其中 a, b 为常向量).

再看以一般参数 t 表示的曲线 $\mathbf{r} = \mathbf{r}(t)$ 的曲率表达式. 由

$$\begin{aligned}
 \mathbf{r}'(s) &= \dot{\mathbf{r}}(t)t'(s) = \frac{\dot{\mathbf{r}}(t)}{\|\dot{\mathbf{r}}(t)\|}, \\
 \mathbf{r}''(s) &= \ddot{\mathbf{r}}(t)t'(s)^2 + \dot{\mathbf{r}}(t)t''(s),
 \end{aligned}$$

及 $\mathbf{r}'(s) \cdot \mathbf{r}''(s) = 0$ 得

$$\begin{aligned}
 k(s) &= \|\mathbf{r}''(s)\| = \|\mathbf{r}'(s) \times \mathbf{r}''(s)\| \\
 &= \|\dot{\mathbf{r}}(t)t'(s) \times (\ddot{\mathbf{r}}(t)t'(s)^2 + \dot{\mathbf{r}}(t)t''(s))\| \\
 &= \|(t'(s))^2(\dot{\mathbf{r}}(t) \times \ddot{\mathbf{r}}(t))\| \\
 &= \frac{\|\dot{\mathbf{r}}(t) \times \ddot{\mathbf{r}}(t)\|}{\|\dot{\mathbf{r}}(t)\|^3}.
 \end{aligned}$$

9.3.2 挠率

空间曲线除了有“弯曲”的概念外,还有一个“挠扭”的概念;“弯曲”是相对于“直”而言的,而“挠扭”是相对于“平”而言的. 位于一个平面上的曲线称为平面曲线,它是不“挠扭”的,所以,我们称非平面曲线为挠曲线. 挠曲线上点 $\mathbf{r}(s + \Delta s)$ 将偏离 s 处的密切平面,或者说曲线上点 $\mathbf{r}(s)$ 和 $\mathbf{r}(s + \Delta s)$ 处的密切平面是不同的. 显然这种偏离也有程度上的差异. 例如,一条弹簧,在压得很紧的时候,它的每一圈都接近于(或者就是)平面曲线,曲线上点 $\mathbf{r}(s + \Delta s)$ 偏离 $\mathbf{r}(s)$ 处的密切平面较小;放开时,曲线上点 $\mathbf{r}(s + \Delta s)$ 偏离 $\mathbf{r}(s)$ 处的密切平面程度就会很大.

我们将用密切平面的变化率,即密切平面的法向量(也就是曲线的副法向量) $B(s)$ 的变化率来描述这种偏离程度,从而导出曲线论中的另一个重要的几何概念:挠率.

在弗雷耐标架 $\{r(s); T(s), N(s), B(s)\}$ 中,

$$B(s) \cdot T(s) = 0.$$

对上式两边求导,得

$$B'(s) \cdot T(s) + B(s) \cdot T'(s) = 0.$$

将 $T'(s) = k(s)N(s)$ 和 $B(s) \cdot N(s) = 0$ 代入得

$$B'(s) \cdot T(s) = 0.$$

又由 $B'(s) \cdot B(s) = 0$ 可知 $B'(s)$ 同时垂直于 $T(s)$ 和 $B(s)$,从而 $B'(s)$ 平行于 $N(s)$,即

$$B'(s) = -\tau(s)N(s),$$

其中 $\tau(s)$ 是数值函数,前面的负号是为了方便添加的.于是,

$$\|B'(s)\| = |\tau(s)|, \quad \tau(s) = -B'(s) \cdot N(s).$$

我们对曲线的挠率给出如下的定义.

定义 9.3 我们称函数

$$\tau(s) = -B'(s) \cdot N(s) \quad (9-25)$$

为曲线 $r(s)$ 在点 s 处的挠率.

由于

$$|\tau(s)| = \|B'(s)\| = \lim_{\Delta s \rightarrow 0} \left| \frac{\Delta \varphi}{\Delta s} \right|$$

其中 $\Delta \varphi$ 表示 $B(s)$ 和 $B(s + \Delta s)$ 间的夹角,即两相邻点的密切平面的夹角.所以,挠率的绝对值为曲线在该点密切平面法向量的倾角对其弧长的变化率,或曲线在该点“偏离”相应的密切平面的程度,这就是挠率的几何意义.

例 9 证明:曲线 $r(s)$ 是平面曲线的充要条件是其挠率恒为零.

证 必要性.因为平面曲线上任一点处的密切平面都是 $r(s)$ 所处的平面,即 $B(s)$ 是一个常向量,故有

$$|\tau(s)| = \|B'(s)\| = 0.$$

充分性.已知对任意 s 有 $\tau(s) = 0$,即 $B'(s) = 0$,即副法线向量 B 为常量.

又由

$$(r(s) \cdot B_0)' = r'(s) \cdot B_0 = T(s) \cdot B_0 = 0,$$

从而

$$(\mathbf{r}(s) \cdot \mathbf{B}_0) = \text{常数, 或 } (\mathbf{r}(s) - \mathbf{r}(0)) \cdot \mathbf{B}_0 = 0,$$

即曲线上任一点 $\mathbf{r}(s)$ 均位于过 $\mathbf{r}(0)$ 以 \mathbf{B}_0 为法向量的平面上.

例 10 求圆柱螺线 $\mathbf{r}(s) = (a \cos \omega s, a \sin \omega s, b\omega s)$ 的曲率和挠率, 其中 a, b 是常数, 且 $a > b, \omega = \frac{1}{\sqrt{a^2 + b^2}}$.

解 由 $\left\| \frac{d\mathbf{r}}{ds} \right\| = 1$ 知 s 为自然参数, 所以

$$\mathbf{N}(s) = \frac{\mathbf{T}'(s)}{|\mathbf{T}'(s)|} = (-\cos \omega s, -\sin \omega s, 0),$$

$$\mathbf{B}(s) = \mathbf{T}(s) \times \mathbf{N}(s) = \omega(b \sin \omega s, -b \cos \omega s, a),$$

$$\mathbf{B}'(s) = b\omega^2(\cos \omega s, \sin \omega s, 0) = -b\omega^2 \mathbf{N}(s).$$

曲率

$$k(s) = \|\mathbf{T}'(s)\| = a\omega^2 = \frac{a}{a^2 + b^2}.$$

挠率

$$\tau = \omega^2 b = \frac{b}{a^2 + b^2}.$$

一条圆柱螺线的曲率和挠率都是常数. 按定义, 曲率 $k(s)$ 总是正数 ($a > 0$), 而挠率 $\tau(s)$ 却可正可负: 当 $b > 0$ 时, $\tau > 0$; $b < 0$ 时, $\tau < 0$, 也就是说, 对于右手螺旋, 挠率为正; 对左手螺旋, 挠率取负值.

从挠率的角度来观察一条挠曲线时, 最简单的挠曲线是圆柱螺线, 如同从曲率角度圆是最简单的平面曲线一样.

下面我们分别给出用自然参数和一般参数来表达挠率 τ 的计算公式

$$\tau = \frac{1}{k}(\mathbf{r}''' \cdot \mathbf{B}) = \frac{(\dot{\mathbf{r}}' \ddot{\mathbf{r}}'' \ddot{\mathbf{r}}''')}{\|\ddot{\mathbf{r}}''\|^2}, \quad (9-26)$$

$$\tau = \frac{(\dot{\mathbf{r}} \ddot{\mathbf{r}} \ddot{\mathbf{r}}'')}{(\dot{\mathbf{r}} \times \ddot{\mathbf{r}})^2}. \quad (9-26')$$

先证 (9-26): 由 $\mathbf{r}'' = k\mathbf{N}$,

$$\begin{aligned} \frac{1}{k}(\mathbf{r}''' \cdot \mathbf{B}) &= \frac{1}{k}((k\mathbf{N})' \cdot \mathbf{B}) = \frac{1}{k}((k\mathbf{N} \cdot \mathbf{B})' - (k\mathbf{N} \cdot \mathbf{B}')) \\ &= -(\mathbf{N} \cdot \mathbf{B}') = \tau. \end{aligned}$$

将 $\mathbf{B} = \frac{\dot{\mathbf{r}}' \times \ddot{\mathbf{r}}''}{\|\ddot{\mathbf{r}}''\|} = \frac{1}{k}(\dot{\mathbf{r}}' \times \ddot{\mathbf{r}}'')$ 代入 $\mathbf{r}''' \cdot \mathbf{B}$, 即得式 (9-26).

再证第二个式子: 它可从参数变换 $s = s(t)$ 得到. 我们只须把下面的等式代入 (9-26) 就得到对参数 t 的挠率计算公式.

$$\mathbf{r}' = \dot{\mathbf{r}} \frac{dt}{ds},$$

$$\mathbf{r}'' = \ddot{\mathbf{r}} \left(\frac{dt}{ds} \right)^2 + \dot{\mathbf{r}} \frac{d^2t}{ds^2},$$

$$\mathbf{r}''' = \ddot{\mathbf{r}}' \left(\frac{dt}{ds} \right)^3 + 3\ddot{\mathbf{r}} \frac{dt}{ds} \frac{d^2t}{ds^2} + \dot{\mathbf{r}} \frac{d^3t}{ds^3},$$

$$\|\mathbf{r}''\| = k = \frac{\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\|}{\|\dot{\mathbf{r}}\|^3}, \|\dot{\mathbf{r}}\| = \frac{ds}{dt}.$$

由公式(9-26) (9-26')可知, $\tau = 0$ ($\mathbf{r}', \mathbf{r}'', \mathbf{r}'''$) = 0 ($\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \ddot{\mathbf{r}}'$) = 0 彼此是等价的, 它们都是平面曲线的充要条件, 也正是这个原因, 非平面曲线被称为挠曲线.

例 11 求三次参数曲线

$$\mathbf{r} = (3t - t^3, 3t^2, 3t + t^3)$$

的弗雷耐标架 $\mathbf{T}, \mathbf{N}, \mathbf{B}$ 及 k 和 τ .

$$\text{解 } \dot{\mathbf{r}} = 3(1 - t^2, 2t, 1 + t^2), \ddot{\mathbf{r}} = 6(-t, 1, t), \ddot{\mathbf{r}}' = 6(-1, 0, 1),$$

$$\dot{\mathbf{r}} \times \ddot{\mathbf{r}} = 18(-1 + t^2, -2t, 1 + t^2),$$

$$\|\dot{\mathbf{r}}\| = 3\sqrt{2}(1 + t^2),$$

$$\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\| = 18\sqrt{2}(1 + t^2).$$

按照标架的计算公式, 有

$$\mathbf{T} = \frac{\dot{\mathbf{r}}}{\|\dot{\mathbf{r}}\|} = \frac{1}{\sqrt{2}} \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}, 1 \right),$$

$$\mathbf{B} = \frac{\dot{\mathbf{r}} \times \ddot{\mathbf{r}}}{\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\|} = \frac{1}{\sqrt{2}} \left(\frac{-1+t^2}{1+t^2}, \frac{-2t}{1+t^2}, 1 \right),$$

$$\mathbf{N} = \mathbf{B} \times \mathbf{T} = \left(\frac{-2t}{1-t^2}, \frac{1-t^2}{1+t^2}, 0 \right)$$

$$k = \frac{\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\|}{\|\dot{\mathbf{r}}\|^3} = \frac{1}{3(1+t^2)^2}$$

$$\tau = \frac{\|(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \ddot{\mathbf{r}}')\|}{\|\dot{\mathbf{r}} \times \ddot{\mathbf{r}}\|^2} = \frac{1}{3(1+t^2)^2}.$$

注意, 对于参数 t , 我们先求 \mathbf{B} , 后求 \mathbf{N} , 这与弧长参数的情形正好相反.

9.3.3 弗雷耐公式

在 9.1 中我们介绍了向量函数 $\mathbf{r}(s)$ 及其各阶导数在某个固定标架 i, j, k 上的分解式:

$$\mathbf{r}(s) = (x(s), y(s), z(s)).$$

在曲线论中,我们常常把曲线 $\mathbf{r} = \mathbf{r}(s)$ 在点 s 处的弗雷耐标架 $(\mathbf{r}(s); \mathbf{T}, \mathbf{N}, \mathbf{B})$ 作为标架,把 $\mathbf{r}(s)$ 表成 $\mathbf{T}, \mathbf{N}, \mathbf{B}$ 的线性组合,即

$$\mathbf{r}(s) = \alpha(s)\mathbf{T} + \beta(s)\mathbf{N} + \gamma(s)\mathbf{B}.$$

由于这个标架是活动的,这时,一阶导数为

$$\frac{d\mathbf{r}(s)}{ds} = \alpha'\mathbf{T} + \beta'\mathbf{N} + \gamma'\mathbf{B} + \alpha\mathbf{T}' + \beta\mathbf{N}' + \gamma\mathbf{B}'.$$

这里 $\mathbf{T}', \mathbf{N}', \mathbf{B}'$ 还是以 s 为参数的向量函数,我们仍希望把它们表示成标架向量 $\mathbf{T}, \mathbf{N}, \mathbf{B}$ 的线性组合. 对于更高阶的导数,将有更多的项,但是最终它们都可表成标架向量 $\mathbf{T}, \mathbf{N}, \mathbf{B}$ 的线性组合. 为此,我们需要求出 $\mathbf{T}', \mathbf{N}', \mathbf{B}'$ 在 $\mathbf{T}, \mathbf{N}, \mathbf{B}$ 上的分解式.

设 s 是 $\mathbf{r}(s)$ 的自然参数,由(9-23)和(9-25),有

$$\mathbf{T}' = k\mathbf{N}, \quad \mathbf{B}' = -\tau\mathbf{N}.$$

对等式 $\mathbf{N} = \mathbf{B} \times \mathbf{T}$ 的两边求导数,得

$$\mathbf{N}' = \mathbf{B}' \times \mathbf{T} + \mathbf{B} \times \mathbf{T}' = -\tau\mathbf{N} \times \mathbf{T} + \mathbf{B} \times k\mathbf{N} = \tau\mathbf{B} - k\mathbf{T}.$$

我们得到公式:

$$\mathbf{T}' = k\mathbf{N}, \quad \mathbf{N}' = -k\mathbf{T} + \tau\mathbf{B}, \quad \mathbf{B}' = -\tau\mathbf{N}. \quad (9-27)$$

公式(9-27)称为弗雷耐公式,写成矩阵形式为

$$\begin{pmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{pmatrix} = \begin{pmatrix} 0 & k & 0 \\ -k & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{pmatrix}. \quad (9-27')$$

矩阵形式(9-27')有助于记忆和应用. 注意这个系数矩阵是反对称的,而且只有两个自由度:曲率 k 和挠率 τ . 弗雷耐公式是曲线论的基本公式.

弧长参数 s , 曲率 $k(s)$ 和挠率 $\tau(s)$ 是描述一条曲线的基本量. 它们的“基本”之处在于:只要知道了这三个量,曲线的形式就完全确定了;因此,这三个量是确定一条曲线的特征量,即有下面两个定理.

定理 9.1 设 $\mathbf{r} = \mathbf{r}_1(s)$, $\mathbf{r} = \mathbf{r}_2(s)$ 是 \mathbf{R}^3 中两条以弧长 s 为参数的正则曲线,如果它们的曲率处处不为零,而且 $k_1(s) = k_2(s)$, $\tau_1(s) = \tau_2(s)$, 则 $\mathbf{r}_1(s)$ 必可通过平移和旋转变换而与 $\mathbf{r}_2(s)$ 重合.

定理 9.2 设 $k(s), \tau(s)$ 是定义在 $[a, b]$ 上的连续可微函数,而且 $k(s) > 0$, 则在 \mathbf{R}^3 中存在正则曲线 $\mathbf{r} = \mathbf{r}(s)$, $s \in [a, b]$, 它以 s 为弧长参数,以 $k(s), \tau(s)$ 为其曲率和挠率.

9.4 特殊的空间曲线

9.4.1 一般螺线

圆柱螺线的曲率和挠率都是非零常数,从挠率的角度观察一条空间挠曲线时,圆柱螺线是最简单的挠曲线.圆柱螺线可以用一个锐角为 α 的直角三角形纸片(其底边 AB 垂直于圆柱轴)卷贴在圆柱上,它的斜边形成的曲线就是圆柱螺线(见图 9-9(a)),把圆柱改为一般柱面,所得到的斜边曲线就是一般的螺线(见图 9-9(b)).

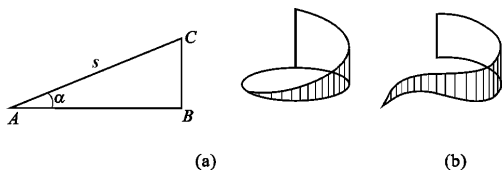


图 9-9

定义 9.4 当一条空间曲线上任一点的切向量和某一固定方向成定角时,该曲线称为一般螺线.

定理 9.3 下列命题(其中 $k, \tau \neq 0$)等价:

- (1) 曲线的切向量 $T(s)$ 和某固定方向成定角;
- (2) 曲线的主法向量 $N(s)$ 和某固定方向垂直;
- (3) 曲线的副法向量 $B(s)$ 和某固定方向成定角;
- (4) 曲线的挠率 τ 和曲率 k 之比等于常数.

9.4.2 球面曲线

位于球面上的曲线称为球面曲线,对球面曲线 $r(s)$ 有以下等价命题.

定理 9.4 当 $\tau \neq \frac{dk}{ds} \neq 0$ 时,下述命题等价:

- (1) $r(s)$ 是位于以 $A(\vec{OA} = a$ 为常向量)为球心的球面上的曲线(s 为自然参数);
- (2) $r(s)$ 的所有法平面过球心 A , 即 $(r(s) - a) \cdot T(s) = 0$; (9-28)
- (3) $r(s)$ 在弗雷耐标架上的表示式为

$$r(s) = a - RN - \frac{R'}{\tau} B \quad (R \text{ 为曲率半径}). \quad (9-29)$$

9.5 曲面的表示 切平面 参数变换

9.5.1 曲面的表示

在多元函数微分学中已讲过 \mathbf{R}^3 中的曲面 S 可以通过空间直角坐标系以显式表示为 $z = f(x, y)$, 或以隐式表示为 $F(x, y, z) = 0$, 还可用参数方程表示为

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \\ z = z(u, v), \end{cases} \quad (u, v) \in D \subset \mathbf{R}^2. \quad (9-30)$$

如果引入向量的记号, 则曲面方程 (9-30) 可以表示为如下的向量形式

$$\mathbf{r} = \mathbf{r}(u, v), \quad (u, v) \in D. \quad (9-30')$$

式中的 \mathbf{r} 表示曲面 S 上点的位置向量, 即

$$\mathbf{r} = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in D, \quad (9-30'')$$

其中 D 表示向量函数 $\mathbf{r}(u, v)$ 的定义域.

向量函数 (9-30') 实际上是一个映射:

$$\mathbf{r}: D \rightarrow \mathbf{R}^3$$

或

$$\mathbf{r}(u, v) \mapsto (x(u, v), y(u, v), z(u, v)).$$

在这个映射下, D 的象集构成 \mathbf{R}^3 的一个曲面 S (见图 9-10).

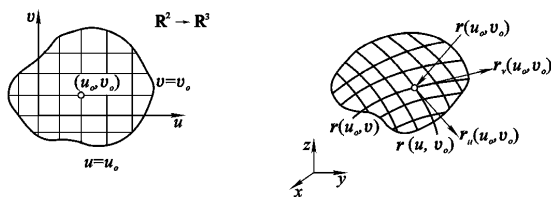


图 9-10

满足 $\mathbf{r}'_u \times \mathbf{r}'_v = \mathbf{0}$ 的点 (u, v) 称为曲面的奇点. 满足 $\mathbf{r}'_u \times \mathbf{r}'_v \neq \mathbf{0}$ 的点 (u, v) 称为曲面的正则点, u, v 称为正则参数. 若映射 $\mathbf{r}: D \rightarrow S$ 是 1-1 对应的, 且 $x, y, z \in C^1(D)$ 秩 $\begin{vmatrix} x'_u & y'_u & z'_u \\ x'_v & y'_v & z'_v \end{vmatrix} = 2, \forall (u, v) \in D$ 则称曲面 $\mathbf{r}(u, v)$

v) 是正则曲面. 为了方便起见, 以后总假定曲面是光滑的(即 $\mathbf{r}(u, v)$ 具有任意阶偏导数), 正则的(即无奇点). 以下将偏导数 $\mathbf{r}'_u, \mathbf{r}'_v$ 简记为 $\mathbf{r}_u, \mathbf{r}_v$.

设 (u_0, v_0) 是 D 中任一点, 固定参数 $v = v_0$, 而让 u 变动, 则直线 $v = v_0$ 被映射到 S 上的一条曲线:

$$\mathbf{r} = \mathbf{r}(u, v_0) = (x(u, v_0), y(u, v_0), z(u, v_0)), \quad (9-31)$$

称为曲面上过点 $\mathbf{r}(u_0, v_0)$ 的 u 坐标曲线, 简称 u 曲线. 同样固定 $u = u_0$, 而让 v 变动, 可得曲面 S 上过点 $\mathbf{r}(u_0, v_0)$ 的一条 v 坐标曲线:

$$\mathbf{r} = \mathbf{r}(u_0, v) = (x(u_0, v), y(u_0, v), z(u_0, v)). \quad (9-31')$$

这样, 在曲面 S 上每一点, 一般都有一条 u 曲线, 一条 v 曲线, 统称参数曲线, 他们的全体构成了曲面上的参数曲线网, 简称坐标网, u, v 也称为曲面上的曲线坐标(见图 9-10).

例 球面 $x^2 + y^2 + z^2 = a^2$ (见图 9-11) 常用的参数方程是

$$\mathbf{r} = \mathbf{r}(u, v) = (a \sin u \cos v, a \sin u \sin v, a \cos u)$$

($0 \leq u \leq \pi, 0 \leq v \leq 2\pi$), 其中常数 $a > 0$.

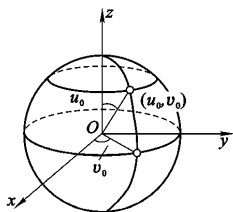


图 9-11

这样, v 曲线($u = u_0$)是平行环线(纬线), 而 u 曲线($v = v_0$)是子午线(经线).

显然, 曲面 S 上的 u 曲线、 v 曲线在点 $\mathbf{r}(u_0, v_0)$ 处的切向量分别是

$$\mathbf{r}_u(u_0, v_0) = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right) \Big|_{(u_0, v_0)},$$

$$\mathbf{r}_v(u_0, v_0) = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right) \Big|_{(u_0, v_0)}.$$

9.5.2 参数变换

同一个曲面, 可以用不同的参数表示, 设参数变换式是

$$\begin{cases} u = u(\bar{u}, \bar{v}) \\ v = v(\bar{u}, \bar{v}), \end{cases} \quad (9-32)$$

且映射 $(\bar{u}, \bar{v}) \mapsto (u, v)$ 满足:

- (1) $u = u(\bar{u}, \bar{v}), v = v(\bar{u}, \bar{v})$ 都是连续可微函数;
 (2) 变换的雅可比行列式不为零, 即

$$\frac{D(u, v)}{D(\bar{u}, \bar{v})} = \begin{vmatrix} \frac{\partial u}{\partial \bar{u}} & \frac{\partial u}{\partial \bar{v}} \\ \frac{\partial v}{\partial \bar{u}} & \frac{\partial v}{\partial \bar{v}} \end{vmatrix} \neq 0.$$

曲面 $\mathbf{r} = \mathbf{r}(u, v)$ 在新参数下的方程为

$$\mathbf{r} = \mathbf{r}(u(\bar{u}, \bar{v}), v(\bar{u}, \bar{v})).$$

将上式分别对 \bar{u}, \bar{v} 求偏导数, 得到曲面上新参数曲线的切向量:

$$\mathbf{r}_{\bar{u}} = \mathbf{r}_u \frac{\partial u}{\partial \bar{u}} + \mathbf{r}_v \frac{\partial v}{\partial \bar{u}}, \quad \mathbf{r}_{\bar{v}} = \mathbf{r}_u \frac{\partial u}{\partial \bar{v}} + \mathbf{r}_v \frac{\partial v}{\partial \bar{v}},$$

而且有

$$\mathbf{r}_{\bar{u}} \times \mathbf{r}_{\bar{v}} = \frac{D(u, v)}{D(\bar{u}, \bar{v})} (\mathbf{r}_u \times \mathbf{r}_v). \quad (9-33)$$

由(9-33)可知, 曲面的法线方向与参数的选择无关, 当雅可比行列式

$\frac{D(u, v)}{D(\bar{u}, \bar{v})} > 0$ 时, 法向量还保持方向一致, 同样, 曲面的正则性也和参数变换无关.

9.5.3 曲面的切平面与法向量

下面研究曲面上一条曲线 C 的切向量, 设曲线 C 在曲面 $S: \mathbf{r} = \mathbf{r}(u, v)$ 上, 可经参数 t 表示为

$$C: \mathbf{r} = \mathbf{r}(u(t), v(t)), \quad a \leq t \leq b. \quad (9-34)$$

根据复合函数的求导法则, 曲线 C 在点 $P(t)$, 即在点 $\mathbf{r}(t) = \overrightarrow{OP}$ 处的切向量是

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \mathbf{r}}{\partial u} \frac{du}{dt} + \frac{\partial \mathbf{r}}{\partial v} \frac{dv}{dt}, \quad (9-35)$$

其中 $\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u}, \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$ 分别是 u 曲线、 v 曲线在同一点处的切向量.

(9-35)式表明, 当 $\mathbf{r}_u \times \mathbf{r}_v \neq 0$, 即 $\mathbf{r}_u, \mathbf{r}_v$ 线性无关时, 过曲面上点 $P(t)$ 的任意曲线 C 的切向量均可由 $\mathbf{r}_u, \mathbf{r}_v$ 线性表出, 也就是它总是落在由 $\mathbf{r}_u, \mathbf{r}_v$ 张成的平面 T_p 内. 我们称由 $\mathbf{r}_u, \mathbf{r}_v$ 张成的平面 T_p 为曲面 S 在点 $P(t)$ 处的切平面, 也称为二维切空间, 并称 $\mathbf{r}_u, \mathbf{r}_v$ 是切空间的基. 任意过 P 点的切向量 $\frac{d\mathbf{r}}{dt}$ 可

用(9-35)表示, $\frac{du}{dt}, \frac{dv}{dt}$ 就是这个切向量关于基 r_u, r_v 的坐标, 我们称 $\left(\frac{du}{dt}, \frac{dv}{dt}\right)$ 或 (du, dv) 为过 P 点的切方向. 切平面的法向量是 $r_u \times r_v$, 也称为曲面 S 在点 $P(t)$ 处的法向量 n , 其单位法向量为

$$n^\circ = \frac{r_u \times r_v}{\|r_u \times r_v\|}. \quad (9-36)$$

曲面 S 在点 $P(t)$ 处的切平面方程为

$$n \cdot (\rho - r) = 0 \text{ 或 } (\rho - r, r_u, r_v) = 0, \quad (9-37)$$

其中 $\rho = (x, y, z)$ 是切平面上动点的位置向量, r 是点 P 的位置向量.

(9-8)式也可用参数方程表示为

$$\rho(\lambda, \mu) = r + \lambda r_u + \mu r_v, \quad (9-37')$$

其中 λ, μ 为参数, $\lambda, \mu \in (-\infty, +\infty)$.

曲面 S 在点 $P(t)$ 处的法线方程是

$$\rho(t) = r(u, v) + tn^\circ(u, v), \quad (9-38)$$

其中 t 是参数, $-\infty < t < +\infty$.

例1 旋转面: 将 xz 平面上的一条曲线

$$C: \begin{cases} x = f(v) \\ z = g(v) \end{cases}, \quad a \leq v \leq \beta$$

绕 z 轴旋转一周(假定 $f(v) > 0$), 所得的曲面称为旋转面, 它的参数方程是

$$r = (f(v)\cos u, f(v)\sin u, g(v)), \quad \begin{cases} 0 \leq u \leq 2\pi, \\ \alpha \leq v \leq \beta \end{cases}$$

曲线 C 称为旋转面的母线, u 曲线称为纬线, 是平行于 xy 平面的圆, v 曲线是经线, 每条经线都落在过 z 轴的一个平面上, 显然和曲线 C 合同(见图9-12), 它的坐标曲线的切向量是

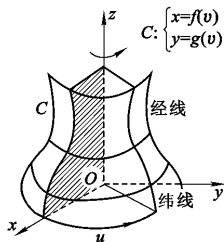


图 9-12

$$\mathbf{r}_u = (-f \sin u, f \cos u, 0), \mathbf{r}_v = (f' \cos u, f' \sin u, g').$$

法向量是

$$\mathbf{r}_u \times \mathbf{r}_v = (fg' \cos u, fg' \sin u, -ff').$$

单位法向量是

$$\mathbf{n}^\circ = \frac{1}{\sqrt{(f')^2 + (g')^2}} (g' \cos u, g' \sin u, -f').$$

例2 圆柱面：圆柱面是旋转面的特例. 图9-13所示圆柱面是例1中当 $f(v) = a$ (常数), $g(v) = v$ 时的旋转面, 因此其方程为

$$\mathbf{r} = (a \cos u, a \sin u, v), \quad 0 \leq u \leq 2\pi, -\infty < v < +\infty.$$

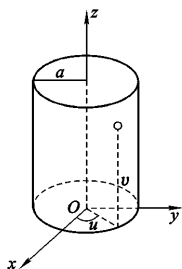


图9-13

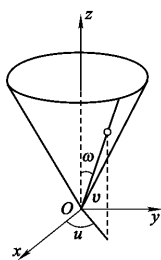


图9-14

例3 圆锥面：以原点为顶点, z 轴为对称轴, 半顶角为 ω ($\cot \omega = b$) (图9-14). 它也是一种旋转面. 在例1中令 $f(v) = v$, $g(v) = bv$ 所得旋转面就是这里的圆锥面, 其方程为

$$\mathbf{r} = (v \cos u, v \sin u, bv), \quad 0 \leq u \leq 2\pi, -\infty < v < +\infty.$$

此时

$$\mathbf{r}_u \times \mathbf{r}_v = (bv \cos u, bv \sin u, -v).$$

在顶点处 ($v = 0$), $\mathbf{r}_u \times \mathbf{r}_v = 0$, 因此顶点是奇点; 为了使曲面处处正则, 应该除去顶点, 即假定 $v \neq 0$, 此时单位法向量是,

$$\mathbf{n}^\circ = (\cos \omega \cos u, \cos \omega \sin u, -\sin \omega),$$

它与参数 v 无关.

例4 螺旋面：如果例1中的曲线 C 在绕 z 轴旋转 u 角的同时沿着 z 轴上升距离 bu ($b > 0$ 是常数), 此时, 所得曲面称为螺旋面, 它的参数方程是

$$\mathbf{r} = (f(v)\cos u, f(v)\sin u, g(v) + bu), \begin{cases} 0 \leq u \leq 2\pi \\ \alpha \leq v \leq \beta \end{cases}.$$

曲线 C 称为螺旋面的母线, u 曲线是圆柱螺线, 螺距是 $2b\pi$, v 曲线是过 z 轴的一个平面上的平面曲线, 合同于曲线 α (见图 9-15).

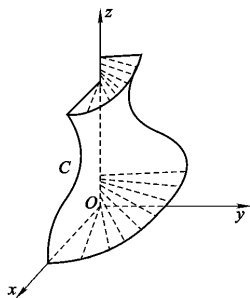


图 9-15

同样,

$$\mathbf{r}_u = (-f\sin u, f\cos u, b), \quad \mathbf{r}_v = (-f'\cos u, f'\sin u, g'),$$

$$\mathbf{r}_u \times \mathbf{r}_v = (fg'\cos u - bf'\sin u, bf'\cos u + fg'\sin u, -ff'),$$

$$\mathbf{n}^\circ = \frac{(fg'\cos u - bf'\sin u, bf'\cos u + fg'\sin u, -ff')}{\sqrt{(fg'\cos u - bf'\sin u)^2 + (bf'\cos u + fg'\sin u)^2 + (ff')^2}}.$$

当 $f(v) = v$, $g(v) = 0$ 时, 曲线 C 是和 x 轴重合的直线段, 这时所生成的螺旋面称为正螺面 (见图 9-16). 机械中的螺杆是正螺面, 有些大楼、场馆、立交桥中的旋转楼梯也是正螺面, 这时

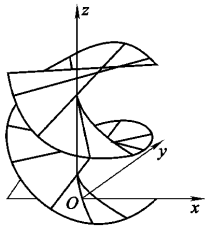


图 9-16

$$\begin{aligned} \mathbf{r} &= (v \cos u, v \sin u, bu), \\ \mathbf{n}^\circ &= \frac{1}{\sqrt{b^2 + v^2}}(-b \sin u, b \cos u, -v). \end{aligned}$$

9.6 曲面的第一基本形式

设 S 为正则曲面, 其方程为

$$S: \mathbf{r} = \mathbf{r}(u, v), \quad (u, v) \in D. \quad (9-39)$$

设曲线 $C \subset S$, 其参数方程为

$$C: \begin{cases} u = u(t) \\ v = v(t) \end{cases} \quad (\alpha < t < \beta).$$

C 的向量方程为

$$\mathbf{r} = \mathbf{r}(u(t), v(t)) \quad (\alpha < t < \beta). \quad (9-40)$$

C 在点 P (其位置向量为 \mathbf{r}) 处的切向量为

$$\frac{d\mathbf{r}}{dt} = \mathbf{r}_u \frac{du}{dt} + \mathbf{r}_v \frac{dv}{dt} = (\mathbf{r}_u, \mathbf{r}_v) \begin{pmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{pmatrix},$$

其微分形式是

$$d\mathbf{r} = \mathbf{r}_u du + \mathbf{r}_v dv = (\mathbf{r}_u, \mathbf{r}_v) \begin{pmatrix} du \\ dv \end{pmatrix}.$$

设 s 是曲线 C 的弧长, 由 $ds = \|d\mathbf{r}\|$ 得

$$\begin{aligned} ds^2 &= d\mathbf{r} \cdot d\mathbf{r} = (\mathbf{r}_u du + \mathbf{r}_v dv)^2 \\ &= \mathbf{r}_u^2 du^2 + 2\mathbf{r}_u \cdot \mathbf{r}_v du dv + \mathbf{r}_v^2 dv^2. \end{aligned}$$

$$\text{记} \quad E = \mathbf{r}_u^2, \quad F = \mathbf{r}_u \cdot \mathbf{r}_v, \quad G = \mathbf{r}_v^2, \quad (9-41)$$

将 ds^2 表示成矩阵形式:

$$ds^2 = d\mathbf{r}^2 = (du, dv) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \quad (9-42)$$

(9-42)式右端是关于微分 du, dv 的二次型, 称为曲面 S 的第一基本形式, 或称为曲面 S 的线素, 它表示 S 上以 (du, dv) 为切方向的一条曲线弧长元素的平方, 一般把它记作 $I(du, dv)$ 即

$$I(du, dv) = (du, dv) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \quad (9-42')$$

其中 E, F, G 称为曲面 S 的第一类基本量, 它们都是 u, v 的函数, 矩阵

$\begin{pmatrix} E & F \\ F & G \end{pmatrix}$ 称为第一基本形式的对应矩阵, 它是一个正定矩阵.

当曲面引入新参数 \bar{u}, \bar{v} , 即作参数变换

$$\begin{cases} u = u(\bar{u}, \bar{v}), \\ v = v(\bar{u}, \bar{v}), \end{cases} \quad \text{且} \quad \frac{D(u, v)}{D(\bar{u}, \bar{v})} > 0. \quad (9-32)$$

对于新参数, 曲面方程为

$$\mathbf{r} = \mathbf{r}(u, v) = \mathbf{r}(u(\bar{u}, \bar{v}), v(\bar{u}, \bar{v})).$$

我们来计算用新参数表示的第一基本形式, 新参数曲线的切向量是

$$\mathbf{r}_{\bar{u}} = \mathbf{r}_u \frac{\partial u}{\partial \bar{u}} + \mathbf{r}_v \frac{\partial v}{\partial \bar{u}},$$

$$\mathbf{r}_{\bar{v}} = \mathbf{r}_u \frac{\partial u}{\partial \bar{v}} + \mathbf{r}_v \frac{\partial v}{\partial \bar{v}}.$$

在点 $P(u, v)$ 的切空间中, 新基 $\mathbf{r}_{\bar{u}}, \mathbf{r}_{\bar{v}}$ 和旧基 $\mathbf{r}_u, \mathbf{r}_v$ 由过渡矩阵(雅可比矩阵 J)联系, 即

$$\begin{pmatrix} \mathbf{r}_{\bar{u}} \\ \mathbf{r}_{\bar{v}} \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial \bar{u}} & \frac{\partial v}{\partial \bar{u}} \\ \frac{\partial u}{\partial \bar{v}} & \frac{\partial v}{\partial \bar{v}} \end{pmatrix} \begin{pmatrix} \mathbf{r}_u \\ \mathbf{r}_v \end{pmatrix} = \left(\frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} \right)^T \begin{pmatrix} \mathbf{r}_u \\ \mathbf{r}_v \end{pmatrix}. \quad (9-43)$$

记在新参数下的第一类基本量为 $\bar{E}, \bar{F}, \bar{G}$, 其第一基本形式的对应矩阵为

$$\begin{pmatrix} \bar{E} & \bar{F} \\ \bar{F} & \bar{G} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_{\bar{u}} \\ \mathbf{r}_{\bar{v}} \end{pmatrix} (\mathbf{r}_{\bar{u}}, \mathbf{r}_{\bar{v}}) \\ = J^T \begin{pmatrix} \mathbf{r}_u \\ \mathbf{r}_v \end{pmatrix} (\mathbf{r}_u, \mathbf{r}_v) J = J^T \begin{pmatrix} E & F \\ F & G \end{pmatrix} J.$$

由此可见, 在不同参数下, 第一基本形式的对应矩阵彼此互为合同矩阵.

对(9-32)式两边微分并以矩阵形式表示, 则

$$\begin{pmatrix} du \\ dv \end{pmatrix} = \frac{\partial(u, v)}{\partial(\bar{u}, \bar{v})} \begin{pmatrix} d\bar{u} \\ d\bar{v} \end{pmatrix} = J \begin{pmatrix} d\bar{u} \\ d\bar{v} \end{pmatrix}.$$

对第一基本形式作参数变换, 即有

$$\begin{aligned} ds^2 &= I(du, dv) = (du, dv) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \\ &= (d\bar{u}, d\bar{v}) J^T \begin{pmatrix} E & F \\ F & G \end{pmatrix} J \begin{pmatrix} d\bar{u} \\ d\bar{v} \end{pmatrix} \end{aligned}$$

$$= (\mathrm{d}\bar{u} \ \mathrm{d}\bar{v}) \begin{pmatrix} \bar{E} & \bar{F} \\ \bar{F} & \bar{G} \end{pmatrix} \begin{pmatrix} \mathrm{d}\bar{u} \\ \mathrm{d}\bar{v} \end{pmatrix} = I(\mathrm{d}\bar{u} \ \mathrm{d}\bar{v}), \quad (9-44)$$

这说明第一基本形式与参数变换无关.

由于 $\mathrm{d}\mathbf{r}^2 = \mathrm{d}s^2 = I(\mathrm{d}u \ \mathrm{d}v)$, 故曲面上曲线段的弧长为

$$L = \int_a^b \sqrt{E\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 + 2F\frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}v}{\mathrm{d}t} + G\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^2} \mathrm{d}t. \quad (9-45)$$

利用第一基本形式还可以求曲面 S 上两条曲线的交角. 设 S 上的两条曲线 C_1 、 C_2 在交点 P 的切向量分别是

$$\mathrm{d}\mathbf{r} = \mathbf{r}_u \mathrm{d}u + \mathbf{r}_v \mathrm{d}v,$$

$$\delta\mathbf{r} = \mathbf{r}_u \delta u + \mathbf{r}_v \delta v,$$

则

$$\begin{aligned} \mathrm{d}\mathbf{r} \cdot \delta\mathbf{r} &= (\mathrm{d}u \ \mathrm{d}v) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} \\ &= E \mathrm{d}u \delta u + F(\mathrm{d}u \delta v + \mathrm{d}v \delta u) + G \mathrm{d}v \delta v. \end{aligned}$$

故两曲线在 P 点处的夹角的余弦为

$$\begin{aligned} \cos \angle \mathrm{d}\mathbf{r} \ \delta\mathbf{r} &= \frac{\mathrm{d}\mathbf{r} \cdot \delta\mathbf{r}}{\|\mathrm{d}\mathbf{r}\| \ \|\delta\mathbf{r}\|} \\ &= \frac{E \mathrm{d}u \delta u + F(\mathrm{d}u \delta v + \mathrm{d}v \delta u) + G \mathrm{d}v \delta v}{\sqrt{E \mathrm{d}u^2 + 2F \mathrm{d}u \mathrm{d}v + G \mathrm{d}v^2} \sqrt{E \delta u^2 + 2F \delta u \delta v + G \delta v^2}}. \end{aligned} \quad (9-46)$$

由(9-46)可知, 曲面上两条曲线相互垂直的充要条件是 $\cos \angle \mathrm{d}\mathbf{r} \ \delta\mathbf{r} = 0$, 即

$$E \mathrm{d}u \delta u + F(\mathrm{d}u \delta v + \mathrm{d}v \delta u) + G \mathrm{d}v \delta v = 0. \quad (9-47)$$

如果两条曲线是坐标参数曲线, 即 u 曲线($\mathrm{d}v = 0$)和 v 曲线($\delta u = 0$), 则坐标参数曲线的夹角余弦为

$$\cos \angle \mathrm{d}\mathbf{r} \ \delta\mathbf{r} = \frac{F}{\sqrt{E} \sqrt{G}}.$$

由此可知, 曲面的参数曲线网是正交坐标网的充分必要条件是 $F = 0$.

例 求正螺面

$$\mathbf{r} = (u \cos v, u \sin v, bv), \quad -\infty < u, v < +\infty$$

的第一基本形式, 并验证其参数曲线网是正交曲线网.

解

$$\mathbf{r}_u = (\cos v, \sin v, 0),$$

$$\mathbf{r}_v = (-u \sin v, u \cos v, b),$$

$$E = \mathbf{r}_u^2 = \cos^2 v + \sin^2 v = 1,$$

$$F = \mathbf{r}_u \cdot \mathbf{r}_v = -u \cos v \sin v + u \sin v \cos v = 0,$$

$$G = r_u^2 = u^2 \sin^2 v + u^2 \cos^2 v + b^2 = u^2 + b^2.$$

正螺面的第一基本形式为

$$I = du^2 + (u^2 + b^2)dv^2.$$

又因为 $F = 0$, 故其参数曲线网是正交曲线网.

9.7 曲面上曲线的法曲率 曲面的第二基本形式

在 9.6 中, 我们通过曲面的第一基本形式研究了曲面上曲线的度量(弧长、交角)性质. 现在要讨论曲面在一点的弯曲状况. 容易想到, 如果过曲面上一点所有曲面曲线的曲率都很大, 那么曲面在这一点将比较“弯曲”. 下面我们从讨论曲面上过一点 $P(u, v)$ 具有不同方向的曲线的曲率入手, 从而引入与曲面在一点处弯曲状况有关的第二基本形式.

9.7.1 曲面上曲线的法曲率

我们已经定义了空间曲线上一点的曲率, 现在来定义曲面在一点的曲率(曲面弯曲的程度). 自然的办法是通过曲面上过这一点所有曲线的曲率入手.

设 $S: \mathbf{r} = \mathbf{r}(u, v)$ 是一正则曲面, $P(u, v)$ 是 S 上一固定点, S 在 P 点的切平面为 Π_P , 其单位法向量为 \mathbf{n}° . S 上一曲线 C 的方程为 $\mathbf{r} = \mathbf{r}(u(s))$, (s 为弧长参数), C 在 P 点的曲率为 k , 标准正交标架为 $\{P; \mathbf{T}, \mathbf{n}^\circ, \mathbf{T} \times \mathbf{n}^\circ\}$, 其中 $\mathbf{T} = \frac{d\mathbf{r}}{ds}$, $\mathbf{T}' = k\mathbf{N}$, 由于 $k\mathbf{N}$ 与 \mathbf{T} 垂直, 故在 \mathbf{n}° 和 $\mathbf{T} \times \mathbf{n}^\circ$ 张成的平面内, $k\mathbf{N}$ 可以分解成

$$k\mathbf{N} = k_n \mathbf{n}^\circ + k_g (\mathbf{n}^\circ \times \mathbf{T}). \quad (9-48)$$

其中 k_n 和 k_g 分别称为曲线 C 在 P 点处的法曲率和测地曲率, $k_n \mathbf{n}^\circ$ 和 $k_g (\mathbf{n}^\circ \times \mathbf{T})$ 分别称为法曲率向量和测地曲率向量.(见图 9-17)

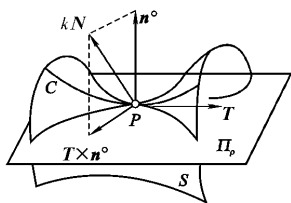


图 9-17

上式两边点乘 \mathbf{n}° 得

$$k_n = k\mathbf{N} \cdot \mathbf{n}^\circ, \quad (9-49)$$

即 P 点处 C 的法曲率是曲率向量 $k\mathbf{N}$ 在 \mathbf{n}° 上的投影, 可见 S 上任意一条曲线在 P 点的法曲率 k_n 不会超过 C 在该点的曲率 k .

9.7.2 曲面的第二基本形式

由 (9-49),

$$k_n = k\mathbf{N} \cdot \mathbf{n}^\circ = \mathbf{T}' \cdot \mathbf{n}^\circ = \mathbf{r}'' \cdot \mathbf{n}^\circ.$$

又

$$\begin{aligned} \mathbf{r}'' &= (\mathbf{r}_u u' + \mathbf{r}_v v')' \\ &= \mathbf{r}_{uu} u'' + \mathbf{r}_{uv} v'' + \mathbf{r}_{uu} (u')^2 + 2\mathbf{r}_{uv} u'v' + \mathbf{r}_{vv} (v')^2, \end{aligned}$$

由于 $\mathbf{r}_u, \mathbf{r}_v$ 垂直于 \mathbf{n}° , 所以有

$$\begin{aligned} k_n &= (\mathbf{r}_{uu} (u')^2 + 2\mathbf{r}_{uv} u'v' + \mathbf{r}_{vv} (v')^2) \cdot \mathbf{n}^\circ \\ &= (\mathbf{r}_{uu} \cdot \mathbf{n}^\circ) (u')^2 + (2\mathbf{r}_{uv} \cdot \mathbf{n}^\circ) u'v' + (\mathbf{r}_{vv} \cdot \mathbf{n}^\circ) (v')^2. \end{aligned}$$

记

$$\mathbf{r}_{uu} \cdot \mathbf{n}^\circ = L, \quad \mathbf{r}_{uv} \cdot \mathbf{n}^\circ = M, \quad \mathbf{r}_{vv} \cdot \mathbf{n}^\circ = N, \quad (9-50)$$

得

$$\begin{aligned} k_n &= L(u')^2 + 2Mu'v' + N(v')^2 \\ &= \frac{Ldu^2 + 2Mdudv + Ndv^2}{ds^2}. \end{aligned}$$

我们称 $Ldu^2 + 2Mdudv + Ndv^2$ 为曲面的第二基本形式, 记成 $II(du, dv)$ 即

$$II(du, dv) = Ldu^2 + 2Mdudv + Ndv^2. \quad (9-51)$$

L, M, N 称为曲面的第二类基本量, 由上一节知 $ds^2 = Edu^2 + 2Fdudv + Gdv^2$ 是曲面的第一基本形式, 所以

$$k_n = \frac{II(du, dv)}{I(du, dv)} = \frac{(du, dv) \begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}}{(du, dv) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}}. \quad (9-52)$$

曲面上曲线 C 在点 $P(u, v)$ 的法曲率 k_n 除与曲面在 P 点处的第一、二类基本量有关外, 仅仅依赖于曲线 C 在点 P 的切向量的方向 du, dv . 过正则曲面上一点的一条曲线对应一个切方向 (du, dv) , 然而给了过这一点的一个切方向, 却有远不止一条曲面上过此点的曲线与之对应, 它们都在这一点彼此相切, 所以曲面上相切于点 P 的两条曲线 C_1 和 C_2 , 有相同的法曲率, 今后我们称 k_n 为曲面的在点 P 沿方向 $d\mathbf{r}$ 的法曲率.

由于 $\mathbf{n}^\circ = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}$ 利用拉格朗日恒等式 (证明略)

$$(\mathbf{r}_1 \times \mathbf{r}_2) \cdot (\mathbf{r}_3 \times \mathbf{r}_4) = \begin{vmatrix} \mathbf{r}_1 \cdot \mathbf{r}_3 & \mathbf{r}_1 \cdot \mathbf{r}_4 \\ \mathbf{r}_2 \cdot \mathbf{r}_3 & \mathbf{r}_2 \cdot \mathbf{r}_4 \end{vmatrix},$$

易得

$$\|\mathbf{r}_u \times \mathbf{r}_v\| = \sqrt{\mathbf{r}_u^2 \mathbf{r}_v^2 - (\mathbf{r}_u \cdot \mathbf{r}_v)^2} = \sqrt{EG - F^2}. \quad (9-53)$$

于是,第二类基本量还可表示为

$$L = \frac{(\mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_{uu})}{\sqrt{EG - F^2}}, M = \frac{(\mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_{uv})}{\sqrt{EG - F^2}}, N = \frac{(\mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_{vv})}{\sqrt{EG - F^2}}. \quad (9-50')$$

再由等式 $\mathbf{r}_u \cdot \mathbf{n}^\circ = 0$, $\mathbf{r}_v \cdot \mathbf{n}^\circ = 0$ 两边分别对 u, v 求导,得

$$\begin{aligned} \mathbf{r}_{uu} \cdot \mathbf{n}^\circ &= -\mathbf{r}_u \cdot \mathbf{n}^\circ_u, & \mathbf{r}_{vv} \cdot \mathbf{n}^\circ &= -\mathbf{r}_v \cdot \mathbf{n}^\circ_v, \\ \mathbf{r}_{uv} \cdot \mathbf{n}^\circ &= -\mathbf{r}_u \cdot \mathbf{n}^\circ_v = -\mathbf{r}_v \cdot \mathbf{n}^\circ_u, \end{aligned}$$

于是又可将第二类基本量表示为

$$\begin{aligned} L &= -\mathbf{r}_u \cdot \mathbf{n}^\circ_u, M = -\mathbf{r}_u \cdot \mathbf{n}^\circ_v = -\mathbf{r}_v \cdot \mathbf{n}^\circ_u, \\ N &= -\mathbf{r}_v \cdot \mathbf{n}^\circ_v. \end{aligned} \quad (9-50'')$$

将(9-50'')代入(9-51),得

$$\begin{aligned} II(du, dv) &= -[\mathbf{r}_u \cdot \mathbf{n}^\circ_u du^2 + (\mathbf{r}_u \cdot \mathbf{n}^\circ_v + \mathbf{r}_v \cdot \mathbf{n}^\circ_u) du dv + (\mathbf{r}_v \cdot \mathbf{n}^\circ_v) dv^2] \\ &= -d\mathbf{r} \cdot d\mathbf{n}^\circ, \end{aligned}$$

其中

$$d\mathbf{r} = \mathbf{r}_u du + \mathbf{r}_v dv; \quad d\mathbf{n}^\circ = \mathbf{n}^\circ_u du + \mathbf{n}^\circ_v dv.$$

例1 求平面的第一、二基本形式.

解 设 S_1 是 \mathbf{R}^3 中的 xy 平面,它的参数方程的向量表示式为

$$\mathbf{r} = (u, v, 0).$$

它的单位法向量是常向量

$$\mathbf{n}^\circ = (0, 0, 1).$$

所以

$$I = d\mathbf{r} \cdot d\mathbf{r} = du^2 + dv^2,$$

$$II = -d\mathbf{r} \cdot d\mathbf{n}^\circ = 0$$

例2 求圆柱面的第一、二基本形式.

解 设圆柱面的方程是

$$\mathbf{r} = (a \cos u, a \sin u, v).$$

因

$$\mathbf{r}_u = (-a \sin u, a \cos u, 0),$$

$$\mathbf{r}_v = (0, 0, 1),$$

$$\begin{aligned}\text{故 } \mathbf{n}^\circ &= \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|} = \frac{(a \cos u, a \sin u, 0)}{a} = (\cos u, \sin u, 0), \\ \mathbf{r}_{uu} &= (-a \cos u, -a \sin u, 0), \\ \mathbf{r}_{vv} &= \mathbf{r}_{uv} = 0.\end{aligned}$$

因此

$$\begin{aligned}E &= \mathbf{r}_u \cdot \mathbf{r}_u = a^2, F = \mathbf{r}_u \cdot \mathbf{r}_v = 0, G = \mathbf{r}_v \cdot \mathbf{r}_v = 1, \\ L &= \mathbf{r}_{uu} \cdot \mathbf{n}^\circ = -a, M = \mathbf{r}_{uv} \cdot \mathbf{n}^\circ = 0, N = \mathbf{r}_{vv} \cdot \mathbf{n}^\circ = 0.\end{aligned}$$

所以

$$\begin{aligned}I &= a^2 du^2 + dv^2, \\ II &= -a du^2.\end{aligned}$$

例3 求旋转面

$$\mathbf{r} = (f(v) \cos u, f(v) \sin u, g(v))$$

的第二基本形式.

解 因

$$\begin{aligned}\mathbf{r}_{uu} &= (-f \cos u, -f \sin u, 0), \\ \mathbf{r}_{uv} &= (-f' \sin u, f' \cos u, 0), \\ \mathbf{r}_{vv} &= (f'' \cos u, f'' \sin u, g''(v)),\end{aligned}$$

单位向量

$$\mathbf{n}^\circ = \frac{1}{\sqrt{(f')^2 + (g')^2}} (g' \cos u, g' \sin u, -f'),$$

就得到

$$\begin{aligned}L &= \mathbf{r}_{uu} \cdot \mathbf{n}^\circ = \frac{-fg'}{\sqrt{(f')^2 + (g')^2}}, \quad M = \mathbf{r}_{uv} \cdot \mathbf{n}^\circ = 0, \\ N &= \mathbf{r}_{vv} \cdot \mathbf{n}^\circ = \frac{f''g' - f'g''}{(f')^2 + (g')^2}.\end{aligned}$$

故

$$II = \frac{-fg'}{\sqrt{(f')^2 + (g')^2}} du^2 + \frac{f''g' - f'g''}{(f')^2 + (g')^2} dv^2.$$

例4 求圆柱面 $\mathbf{r} = (a \cos u, a \sin u, v)$ 上 u 曲线的法曲率.

解 由例2, 圆柱面的第一、第二基本形式为

$$\begin{aligned}I &= a^2 du^2 + dv^2, \\ II &= -a du^2,\end{aligned}$$

$$k_n = - \frac{a du^2}{a^2 du^2 + dv^2}.$$

u 曲线 ($dv = 0$) 的法曲率为

$$k_n = - \frac{1}{a}.$$

若用 φ 表示切方向 $(a du, dv)$ 与 u 曲线的夹角, 则

$$\cos \varphi = \frac{a du}{\sqrt{(a du)^2 + dv^2}},$$

$$k_n = - \frac{1}{a} \frac{a^2 du^2}{(a du)^2 + dv^2} = - \frac{1}{a} \cos^2 \varphi.$$

所以 $-\frac{1}{a} \leq k_n \leq 0$, k_n 是 v 曲线处的法曲率. 法曲率是负的, 说明 n° 是曲面的外法线向量, 曲面是朝着 $-n^\circ$ 方向弯曲的.

9.7.3 梅尼埃定理

曲面 S 上过 P 点且与某个固定方向相切的曲线有很多, 而且其中还有很多是平面曲线, 它们的曲率可能不同, 但都有相同的法曲率. 我们希望把求 S 上曲线 C 在一点过方向 dr 的法曲率问题归结为求某一平面曲线 C_n 在这一点的曲率的问题. 过曲面在 P 点处的法线的平面, 称为曲面的法平面, 法平面与曲面的交线称为法截线, 由于过 P 的法平面很多, 相应的法截线也很多, 但过给定切方向 (du, dv) 的法平面只有一个, 所以相应的法截线也只有一条. 法截线在 P 点的曲率向量 kN 和 n° 平行, 法截线的曲率就是法曲率 k_n , 于是我们得到有关法曲率的下述定理.

定理 曲面 S 在点 $P(u, v)$ 处沿切方向 dr 的法曲率 k_n 等于沿同一方向的法截线在该点的曲率.

这个定理说明曲面在一点的法曲率就是法截线的曲率, 那么其他 S 上过 P 点而以 dr 为切方向的曲线 C_1 的曲率与法曲率又是什么关系呢?

设曲面 S 在点 P 的单位法向量 n° 和 kN 的夹角为 φ , 法平面为 Π_n , 法截线是 C_n , 法截线的曲率为 k_n , 曲率半径为 $R_n = k_n^{-1}$. 过点 P 和一固定切向量 dr 作一平面 Π_1 , 记 Π_1 与曲面的交线为 C_1 (见图 9-18), C_1 的曲率向量为 kN , 其曲率半径为 $R = k^{-1}$. 根据 (9-49) 式,

$$k_n = kN \cdot n^\circ = k \cos \varphi,$$

从而得

$$R = R_n \cos \varphi. \quad (9-54)$$

(9-54) 式的几何意义可表述成下面的定理.

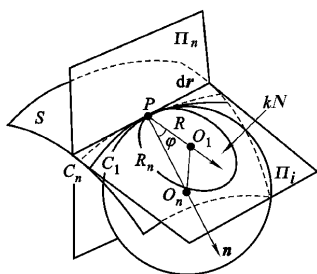


图 9-18

定理 (梅尼埃(Meusnier)定理) 曲面上任何曲线在某点的曲率半径 R 等于对应的法截线在这点的曲率半径 R_n 乘以曲面的法线与曲线的主法线之间夹角的余弦(或 R 等于在曲面的法线上截取的对应的法截线的曲率半径 R_n 在曲线的主法线上的投影).

换句话说,曲面上在 P 处相切(方向为 dr)的全体曲线的曲率中心的轨迹是在与该切方向垂直的平面上一个圆,此圆的直径就是法截线的曲率半径 R_n .

以球面为例,法截线是大圆,设 C 是球面上任意圆,公式(9-54)就变成两个圆半径的关系(如图 9-19),这显然是正确的.

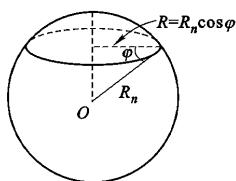


图 9-19

9.7.4 主方向和主曲率

(9-52)式表明,曲面 S 上点 P 的法曲率 k_n 是随切线方向

$$dr = r_u du + r_v dv = (r_u, r_v) \begin{pmatrix} du \\ dv \end{pmatrix}$$

而改变的.以后我们把切线方向 dr 用 $(du, dv)^T$ 表示(即用它在点 P 处的切空

间的基 $\{r_u, r_v\}$ 下的坐标表示), 把在 dr 方向的法曲率记作 $k_n(du, dv)$. 下面要研究的主方向和主曲率是在特殊方向上的法曲率.

定义 使法曲率 $k_n(du, dv)$ 取极值的方向 (du, dv) 称为曲面在该点的主方向, 该极值称为曲面在该点的主曲率.

由于第一、二基本形式与参数选择无关, 因此我们可以选取合适的新参数 (即作参数变换), 使第一、第二基本形式同时化成标准形 (把矩阵化成对角阵).

第一、二基本形式为

$$I(du, dv) = (du, dv) \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}, \quad (1)$$

$$II(du, dv) = (du, dv) \begin{pmatrix} L & M \\ M & N \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \quad (2)$$

其中 $A = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$ 是正定矩阵, $B = \begin{pmatrix} L & M \\ M & N \end{pmatrix}$ 是实对称矩阵, 所以存在可逆矩阵 W , 使得

$$W^T A W = E_2 \text{ (二阶单位矩阵)}, \quad (3)$$

$$W^T B W = \text{diag}(\lambda_1, \lambda_2). \quad (4)$$

这是由于存在可逆阵 R , 使得 $R^T A R = E_2$, 且 $R^T B R$ 仍为实对称矩阵, 于是存在正交矩阵 Q , 使得

$$Q^T (R^T B R) Q = \text{diag}(\lambda_1, \lambda_2), \quad (5)$$

$$Q^T (R^T A R) Q = Q^T E Q = E, \quad (6)$$

其中 λ_1, λ_2 是 $R^T B R$ 的特征值, 在 (5)、(6) 式中令 $W = RQ$ 即得 (3)、(4) 式. 作参数变换

$$\begin{pmatrix} du \\ dv \end{pmatrix} = W \begin{pmatrix} d\bar{u} \\ d\bar{v} \end{pmatrix}, \quad (7)$$

把 (7) 式代入第一、二基本形式 (即 du, dv 的两个二次型), 可将它们化为 $d\bar{u}, d\bar{v}$ 的平方和, 即

$$I(du, dv) = d\bar{u}^2 + d\bar{v}^2,$$

$$II(du, dv) = \lambda_1 d\bar{u}^2 + \lambda_2 d\bar{v}^2,$$

因此便有

$$k_n = \frac{II(du, dv)}{I(du, dv)} = \frac{II(d\bar{u}, d\bar{v})}{I(d\bar{u}, d\bar{v})} = \frac{\lambda_1 d\bar{u}^2 + \lambda_2 d\bar{v}^2}{d\bar{u}^2 + d\bar{v}^2}. \quad (9-55)$$

当 $\lambda_1 = \lambda_2 = \lambda$ 时, 由 (9-55) 可知任何方向的法曲率都是常数 λ , 主方

向变为不确定. 这样的点称为曲面上的脐点. 脐点就是 II 与 I (或 B 与 A) 成比例的点, 即使

$$\frac{L}{E} = \frac{M}{F} = \frac{N}{G}$$

成立的点. 若它们的比为零, 则称该脐点为平点. 若比值不是零, 则称该脐点为圆点. 曲面上每一点都是脐点的曲面称为全脐点曲面, 全脐点曲面必是平面或球面,

当 $\lambda_1 \neq \lambda_2$ 时, 设 $\lambda_1 < \lambda_2$, 由于

$$\frac{\lambda_1(d\bar{u}^2 + d\bar{v}^2)}{d\bar{u}^2 + d\bar{v}^2} \leq \frac{\lambda_1 d\bar{u}^2 + \lambda_2 d\bar{v}^2}{d\bar{u}^2 + d\bar{v}^2} \leq \frac{\lambda_2(d\bar{u}^2 + d\bar{v}^2)}{d\bar{u}^2 + d\bar{v}^2},$$

即知 $\lambda_1 \leq k_n \leq \lambda_2$, 并且 λ_1, λ_2 是 $k_n(du, dv)$ 可以取到的函数值, 故 λ_1, λ_2 是 k_n 的最小值和最大值. 由定义 9.5 知 λ_1, λ_2 是曲面 S 在 $P(u, v)$ 处的主曲率.

如果曲面 $S: \mathbf{r} = \mathbf{r}(u, v)$ 在点 P 处的主曲率 λ_1, λ_2 不相等, 则 λ 是主曲率的充要条件是它满足方程

$$|\lambda A - B| = \left| \lambda \begin{pmatrix} E & F \\ F & G \end{pmatrix} - \begin{pmatrix} L & M \\ M & N \end{pmatrix} \right| = 0, \quad (9-56)$$

即

$$(EG - F^2)\lambda^2 - (LG - 2MF + NE)\lambda + (LN - M^2) = 0, \quad (9-57)$$

而且两个主曲率对应的主方向是正交的.

事实上, 由 ⑤ 式可知, 主曲率 λ_1, λ_2 是 $R^T B R$ 的特征值, 所以 λ 为主曲率的充要条件是 λ 满足 $R^T B R$ 的特征方程, 即

$$|\lambda E - R^T B R| = 0.$$

由 $R^T A R = E$ 且 $|R| \neq 0$, 即得

$$|\lambda E_2 - R^T B R| = |\lambda R^T A R - R^T B R| = |R|^2 |\lambda A - B| = 0.$$

因此, 满足方程 $|\lambda A - B| = 0$ 是 λ 为主曲率的充要条件, 并称该方程为主曲率方程.

再由 (9-56) 式可知, 当 $d\bar{v} = 0$ 时, $k_n = \lambda_1$; 当 $d\bar{u} = 0$ 时, $k_n = \lambda_2$. 所以 $(d\bar{u}, 0)$ 和 $(0, d\bar{v})$ 是主曲率 λ_1, λ_2 对应的两个主方向. 再根据参数变换式 ⑦, 把这两个主方向变换为原参数 u, v 的表示式, 此时记变换矩阵 W 的两个二维列向量为 a_1 和 a_2 , 即

$$W = (\omega_{ij})_{2 \times 2} = (a_1, a_2).$$

于是在参数 u, v 下的两个主方向为

$$dr_1 = \begin{pmatrix} du \\ dv \end{pmatrix} = W \begin{pmatrix} d\bar{u} \\ 0 \end{pmatrix} = (a_1, a_2) \begin{pmatrix} d\bar{u} \\ 0 \end{pmatrix} = d\bar{u}a_1 // a_1, \quad (9-58)$$

$$dr_2 = \begin{pmatrix} du \\ dv \end{pmatrix} = W \begin{pmatrix} 0 \\ d\bar{v} \end{pmatrix} = (a_1, a_2) \begin{pmatrix} 0 \\ d\bar{v} \end{pmatrix} = d\bar{v}a_2 // a_2. \quad (9-59)$$

需要指出,这里的主方向 $(du, dv)^T$ 是主方向 dr_1, dr_2 在切空间的基 $\{r_u, r_v\}$ 下的坐标向量,为证明两个主方向正交,不妨设 a_1 与 a_2 是两个主方向向量 e_1, e_2 在基 $\{r_u, r_v\}$ 下的坐标向量,即

$$e_j = (r_u, r_v) \begin{pmatrix} w_{1j} \\ w_{2j} \end{pmatrix} = (r_u, r_v) a_j, \quad j = 1, 2. \quad (9-60)$$

如此即得 e_1 与 e_2 的内积:

$$\begin{aligned} (e_1, e_2) &= (w_{11}r_u + w_{21}r_v) \cdot (w_{12}r_u + w_{22}r_v) \\ &= w_{11}w_{12}r_u^2 + (w_{11}w_{22} + w_{21}w_{12})r_u \cdot r_v + w_{21}w_{22}r_v^2 \\ &= (w_{11}, w_{21}) \begin{pmatrix} r_u^2 & r_u \cdot r_v \\ r_u \cdot r_v & r_v^2 \end{pmatrix} \begin{pmatrix} w_{12} \\ w_{22} \end{pmatrix} = a_1^T A a_2. \end{aligned}$$

再根据⑥式

$$W^T A W = \begin{pmatrix} a_1^T \\ a_2^T \end{pmatrix} A (a_1, a_2) = \begin{pmatrix} a_1^T A a_1 & a_1^T A a_2 \\ a_2^T A a_1 & a_2^T A a_2 \end{pmatrix} = E,$$

即得 $(e_1, e_2) = a_1^T A a_2 = 0$, 故两个方向正交.

如果曲面 S 上的点 P 不是脐点,在点 P 的切空间中必存在两个单位正交的主方向向量,因为在(9-58)(9-59)中适当选取 $d\bar{u}, d\bar{v}$,即可使 dr_1 和 dr_2 都是单位向量.

定理 (欧拉(Euler)公式) 设 $\{e_1, e_2\}$ 是曲面 S 在点 P 的两个单位正交的主方向向量,它们对应的主曲率为 k_1, k_2 ,则在点 P 沿切向量

$$T = (\cos \theta)e_1 + (\sin \theta)e_2 \quad (0 \leq \theta \leq \pi)$$

的法曲率为

$$k_n(T) = k_1 \cos^2 \theta + k_2 \sin^2 \theta. \quad (9-61)$$

证 不妨设 e_1, e_2 的表达式如(9-60)式,于是

$$T = (e_1, e_2) \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} = (r_u, r_v) W \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix},$$

所以切向量 T 在切空间的基 (r_u, r_v) 下的坐标为

$$\begin{pmatrix} du \\ dv \end{pmatrix} = W \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

因此, 根据参数变换式 ⑦

$$\begin{pmatrix} du \\ dv \end{pmatrix} = W \begin{pmatrix} d\bar{u} \\ d\bar{v} \end{pmatrix},$$

即得 $(d\bar{u} \ d\bar{v})^T = (\cos \theta \ \sin \theta)^T$, 按 (9-55) 式就有

$$k_n(\mathbf{T}) = \frac{k_1 d\bar{u}^2 + k_2 d\bar{v}^2}{d\bar{u}^2 + d\bar{v}^2} = k_1 \cos^2 \theta + k_2 \sin^2 \theta.$$

9.7.5 高斯曲率(全曲率)和平均曲率

定义 曲面 S 在点 P 的高斯曲率 K 和平均曲率 H 定义为

$$K = k_1 k_2, \quad H = \frac{1}{2}(k_1 + k_2),$$

其中 k_1, k_2 为点 P 处的主曲率.

由主曲率方程 (9-57)

$$(EG - F^2)\lambda^2 - (LG - 2MF + NE)\lambda + (LN - M^2) = 0$$

和根与系数的关系, 即得高斯曲率和平均曲率的计算公式:

$$K = \frac{LN - M^2}{EG - F^2} = \frac{\begin{vmatrix} L & M \\ M & N \end{vmatrix}}{\begin{vmatrix} E & F \\ F & G \end{vmatrix}}, \quad (9-62)$$

$$H = \frac{1}{2} \frac{LG - 2MF + NE}{EG - F^2} = \frac{1}{2} \frac{\begin{vmatrix} E & M \\ F & N \end{vmatrix} + \begin{vmatrix} L & F \\ M & G \end{vmatrix}}{\begin{vmatrix} E & F \\ F & G \end{vmatrix}}. \quad (9-63)$$

这样, 主曲率方程 (9-57) 就可以表示成便于记忆的形式:

$$\lambda^2 - 2H\lambda + K = 0.$$

如果在整个曲面 S 上高斯曲率 $K = \text{常数}$, 则称 S 为常曲率曲面.

例 环面的方程是

$$\mathbf{r} = ((a + r \cos v) \cos u, (a + r \cos v) \sin u, r \sin v),$$

其中 a, r 是常数, 且 $a > r > 0$, 参数 u, v 满足 $0 \leq u, v < 2\pi$, 求其上任意一点的高斯曲率.

解 经计算:

$$E = (a + r \cos v)^2, F = 0, G = r^2,$$

$$L = (a + r \cos v) \cos v, M = 0, N = r,$$

因此高斯曲率为

$$K = \frac{LN - M^2}{EG - F^2} = \frac{\cos v}{r(a + r \cos v)},$$

可见 K 的符号由 $\cos v$ 的符号确定.

定理 若 $f(x, y) \in C^2(D)$ 则曲面 $z = f(x, y)$ 的高斯曲率和平均曲率分别为

$$K = \frac{z''_{xx}z''_{yy} - (z''_{xy})^2}{[1 + (z'_x)^2 + (z'_y)^2]^2}, \quad (9-64)$$

$$H = \frac{[1 + (z'_x)^2]z''_{yy} - 2z'_xz'_yz''_{xy} + [1 + (z'_y)^2]z''_{xx}}{[1 + (z'_x)^2 + (z'_y)^2]^{\frac{3}{2}}}. \quad (9-65)$$

证 把曲面方程写成向量形式

$$\mathbf{r} = (x, y, f(x, y)).$$

将 x, y 看作参数 u, v , 于是得

$$\mathbf{r}_x = (1, 0, z'_x), \quad \mathbf{r}_y = (0, 1, z'_y),$$

$$\mathbf{n}^0 = \frac{(-z'_x, -z'_y, 1)}{\sqrt{1 + (z'_x)^2 + (z'_y)^2}}.$$

又

$$\mathbf{r}_{xx} = (0, 0, z''_{xx}), \quad \mathbf{r}_{xy} = (0, 0, z''_{xy}), \quad \mathbf{r}_{yy} = (0, 0, z''_{yy}),$$

所以

$$E = 1 + (z'_x)^2, \quad F = z'_xz'_y, \quad G = 1 + (z'_y)^2,$$

$$L = \frac{z''_{xx}}{\mu}, \quad M = \frac{z''_{xy}}{\mu}, \quad N = \frac{z''_{yy}}{\mu},$$

其中 $\mu = \sqrt{1 + (z'_x)^2 + (z'_y)^2}$ 将上面算得的基本量代入(9-62)及(9-63)即得.

例 证明: 猴鞍面(图9-20) $z = x^3 - 3xy^2$ 上, 原点(0, 0, 0)是一个平点. 并求其上 $P(1, 1, 2)$ 处的高斯曲率和平均曲率.

解

$$z'_x = 3x^2 - 3y^2, \quad z'_y = -6xy,$$

$$z''_{xx} = 6x, \quad z''_{xy} = -6y, \quad z''_{yy} = -6x,$$

$$\begin{aligned} K &= \frac{-36(x^2 + y^2)}{[1 + (3x^2 - 3y^2)^2 + (-6xy)^2]^2} \\ &= \frac{-36(x^2 + y^2)}{[1 + (3x^2 + 3y^2)^2]} \leq 0. \end{aligned}$$

在原点(0, 0, 0)处, $z''_{xx} = z''_{yy} = z''_{xy} = 0$ 所以 $L = M = N = 0$ 第二基本形

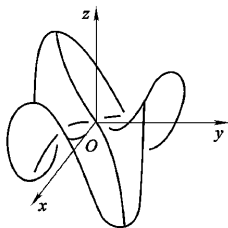


图 9-20

式对应的矩阵 $B = 0$, 由主曲率方程(9-57)得 $\lambda_2 = 0$, 即主曲率 $k_1 = k_2 = 0$, 从而 $k_n \equiv 0$, 因此原点为平点.

在 $(1, 1, -2)$ 处: $z'_x = 0, z'_y = -6, z'_{xx} = 6, z'_{xy} = z'_{yx} = -6$, 代入 (9-64)(9-65), 即得

$$K_p = \frac{-72}{(1+36)^2} \approx -0.0526,$$

$$H_p = \frac{36 \times 6}{(1+36)^{\frac{3}{2}}} \approx 0.96.$$

我们知道, 空间曲线的形状完全由弧长参数 s , 曲率 $k(s)$ 和挠率 $\tau(s)$ 三个特征量所决定. 自然就要问, 对曲面来说, 这样的“特征量”是什么呢? 下面的定理给出了部分的答案:

定理 设 S_1, S_2 是定义在 $D \subset \mathbf{R}^2$ 上的两个正则曲面. 如果在每一点 $(u, v) \in D$, S_1, S_2 都有相同的第一、第二基本形式, 则 S_1 必定能通过平移和旋转变换而与 S_2 重合.

但反过来, 给定了第一、第二基本形式, 还不足以唯一确定一个曲面 S .

习题与补充题

习题

1. 证明 $a(t)$ 是常向量的充要条件是 $\dot{a}(t) = 0$.

2. 设 λ_0 是常数, a_0 是常向量, 证明

$$(1) \frac{d}{dt}(\lambda_0 r(t)) = \lambda_0 \dot{r}; \quad (2) \frac{d}{dt}(\lambda(t) a_0) = \lambda'(t) a_0;$$

$$(3) \frac{d}{dt}(a_0 \cdot r(t)) = a_0 \cdot \dot{r}(t); \quad (4) \frac{d}{dt}(a_0 \times r(t)) = a_0 \times \dot{r}(t).$$

3. 下列等式成立吗?为什么?

$$(1) \mathbf{r}^2 = \|\mathbf{r}\|^2;$$

$$(2) \left\| \frac{d\mathbf{r}}{dt} \right\| = \frac{d}{dt} \|\mathbf{r}\|;$$

$$(3) \mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = \|\mathbf{r}\| \frac{d\|\mathbf{r}\|}{dt}.$$

4. 设向量函数 $\mathbf{a}(t)$ 满足 $\mathbf{a} \cdot \dot{\mathbf{a}} = 0$, $\mathbf{a} \times \dot{\mathbf{a}} = 0$, 证明 $\mathbf{a}(t)$ 是常向量.

5. 证明 $\mathbf{r}(t) = (2t-1, t^2-2, -t^2+4t)$ 为共面向量函数.

6. 证明 $\mathbf{r}(t) = at^3 + bt^2 + ct$ 为共面向量函数的充要条件是

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = 0.$$

7. 试证明

$$\mathbf{r}_1 = (t \sin t, e^t) \quad -\infty < t < +\infty$$

与

$$\mathbf{r}_2 = (\ln \theta, \sin \ln \theta, \theta) \quad 0 < \theta < +\infty$$

是同一条曲线的两种不同的表示式.

8. 求曲线 $\mathbf{r} = (e^t \cos t, e^t \sin t, e^t)$ 在 $t = 0$ 处的切线方程.

9. 求曲线 $x^2 + y^2 = 1, y^2 + z^2 = 1$ 在任意点处的法平面方程.

10. 求下列曲线的切线和法平面方程.

$$(1) \mathbf{r} = (a \cos t, a \sin t, bt), t = 0;$$

$$(2) \mathbf{r} = (t, t^2, t^3), t = 1;$$

$$(3) \begin{cases} f(x, y, z) = 0 \\ g(x, y, z) = 0, \end{cases} \quad P(x_0, y_0, z_0).$$

11. 求下列曲线的副法线和密切平面方程.

$$(1) \mathbf{r} = (a \cos t, b \sin t, e^t), t = 0;$$

$$(2) \mathbf{r} = (a \cos t + b \sin t, a \sin t + b \cos t, c \sin 2t), t = \frac{\pi}{2}.$$

12. 求曲线 $\mathbf{r} = (t, t^2, t^3)$ 在 $t = 1$ 处的主法线和从切平面方程.

13. 证明球面曲线的法平面通过球心.

14. 计算圆锥螺线 $\mathbf{r} = (e^t \cos t, e^t \sin t, e^t)$ 的弧长公式(从 0 到 t).

15. 求下列平面曲线的弧长公式及弧长.

$$(1) \text{曲线由直角坐标中显式表示 } y = f(x), y = \ln(1-x^2) \quad 0 \leq x \leq \frac{1}{2};$$

$$(2) \text{曲线由极坐标方程表示 } \rho = \rho(\varphi), \text{对数螺线 } \rho = e^{a\varphi} \quad 0 \leq \varphi \leq \varphi_0.$$

16. 将方程 $\mathbf{r} = (a \cos t, a \sin t, bt)$ (圆柱螺线)化成以弧长为参数的方程.

17. 求曲线 $\mathbf{r} = (t \sin t, t \cos t, te^t)$ 在 $t = 0$ 处的弗雷耐标架.

18. 求下列曲线的曲率 k 和挠率 τ .

(1) $\mathbf{r} = (a \cosh t, a \sinh t, at)$;

(2) $\mathbf{r} = (t - \sin t, 1 - \cos t, t)$;

(3) $\mathbf{r} = (t \cos t, t \sin t, at)$ (圆锥曲线);

(4) $\mathbf{r} = (t, t^2, t^3)$.

19. 证明曲线 $\mathbf{r} = \left(t, 1 + \frac{1}{t}, \frac{1}{t} - t \right)$ 是平面曲线.

20. 证明曲线 $\mathbf{r} = (1 + 3t + 2t^2, 2 - 2t + 5t^2, 1 - t^2)$ 是平面曲线.

21. 证明

(1) $\mathbf{T}' \cdot \mathbf{N}' = 0$; (2) $\mathbf{B}' \cdot \mathbf{N}' = 0$.

22. 已知曲线 $\mathbf{r} = \mathbf{r}(s)$, 证明:

(1) $\mathbf{r}' \cdot \mathbf{r}'' = 0$; (2) $\mathbf{r}''' = -k^2 \mathbf{T} + k' \mathbf{N} + k\tau \mathbf{B}$;

(3) $\mathbf{r}' \cdot \mathbf{r}''' = -k^2$; (4) $(\mathbf{r}, \mathbf{r}'', \mathbf{r}''') = k^2 \tau$;

(5) $\mathbf{r}'' \cdot \mathbf{r}''' = kk'$.

23. 证明

(1) $(\mathbf{T}, \mathbf{B}, \mathbf{B}') = \tau$; (2) $(\mathbf{B}', \mathbf{B}'', \mathbf{B}''') = \tau^2 \left(\frac{k}{\tau} \right)'$;

(3) $(\mathbf{T}', \mathbf{T}'', \mathbf{T}''') = k^2 \left(\frac{\tau}{k} \right)'$.

24. 试证明曲线 $\mathbf{r} = \mathbf{r}(t)$ 在一般参数下的弗雷耐公式为

$$\dot{\mathbf{T}} = kv\mathbf{N}, \dot{\mathbf{N}} = -kv\mathbf{T} + \tau v\mathbf{B}, \dot{\mathbf{B}} = -\tau v\mathbf{N}, \text{ 其中 } v = \|\dot{\mathbf{r}}\|.$$

25. 已知曲线 $C: \mathbf{r} = \mathbf{r}(s)$, 证明: 若曲线 C 的

(1) 所有切线通过定点, 则 C 是直线;

(2) 所有切线相互平行, 则 C 是直线;

(3) 所有主法线通过定点, 则 C 是圆;

(4) 所有切线平行于同一平面, 则 C 是平面曲线.

26. 设 $\tau = ck$, c 为常数, 写出此曲线的参数方程.

27. 已给 $\mathbf{r} = \mathbf{r}(t)$ 的单位副法线向量为 $\mathbf{B} = \frac{1}{\sqrt{2}}(-\sin t, \cos t, 1)$, 求它

的单位切向量 \mathbf{T} 和单位主法线向量 \mathbf{N} . 证明此曲线是一般螺线, 并求它的曲率和挠率的比值.

28. 证明曲线 $\mathbf{r}(t) = (3t, 3t^2, 2t^3)$ 是一般螺线.

29. 证明下列曲线是球面曲线.

(1) $\mathbf{r} = (a \sin^2 t, a \sin t \cos t, a \cos t)$;

$$(2) \mathbf{r} = (-\cos 2\theta, -2\cos \theta \sin 2\theta).$$

30. 证明: 当且仅当球面曲线是圆周时, 其曲率不变.

31. 证明曲面 $\mathbf{r} = (a \cos \varphi \cos \theta, b \sin \varphi \cos \theta, c \sin \theta)$ 是椭球面, 并求其法向量, 切平面及坐标曲线.

32. 求圆锥的参数方程和它的切平面.

33. 证明曲面

$$(1) \mathbf{r} = \left(u, v, \frac{1}{2} \left(\frac{u^2}{a^2} + \frac{v^2}{b^2} \right) \right) \text{ 是椭圆抛物面;}$$

$$(2) \mathbf{r} = (a(u+v), b(u-v), 2uv) \text{ 是双曲抛物面.}$$

34. 求题 3 中各曲面的法向量和切平面.

35. 求旋转曲面 $\mathbf{r} = (u \cos v, u \sin v, f(u))$ ($0 < v < 2\pi$) 的单位法向量.

36. 求劈锥面 $\mathbf{r} = (u \cos v, u \sin v, f(u))$ 的切平面和法线方程.

37. 证明一曲面是球面的充要条件是它的所有法线通过一定点.

38. 设曲面的表示式为 $z = f(x, y)$ 求它的法向量.

39. 求双曲抛物面 $\mathbf{r} = (u+v, u-v, uv)$ 在 $u=1, v=-1$ 点处的单位法向量和切平面方程.

40. 证明: 旋转面 $\mathbf{r} = (f(v) \cos u, f(v) \sin u, g(v))$ ($g'(v) \neq 0$) 上任一点所作的法线一定和 z 轴相交.

41. 已知曲面的第一基本形式为 $ds^2 = du^2 + f(u, v)dv^2$, 证明:

(1) u 曲线和 v 曲线正交;

(2) 任意两条 u 曲线被 v 曲线截成等长的弧, 即 (u_1, v_0) 到 (u_2, v_0) 的弧长与 v_0 无关.

42. 已知曲面 S 的第一基本形式为 $ds^2 = du^2 + \sinh^2 u dv^2$, 求 S 上的曲线 $C: u=t, v=t$ ($0 \leq t \leq 1$) 的弧长.

43. 已知曲面 S 的第一基本形式为 $ds^2 = du^2 + (u^2 + a^2)dv^2$, 求 S 上两条曲线 $C_1: u+v=0$ 和 $C_2: u-v=0$ 在交点 $(0, 0)$ 处的交角.

44. 求圆柱面 $\mathbf{r} = \left(a \cos \frac{u}{a}, a \sin \frac{u}{a}, v \right)$ 的法曲率.

45. 试证明: 在表面上的一点, 任何两个正交方向的法曲率之和相等.

46. 求下列曲面的第二基本形式.

$$(1) \text{ 正螺面 } \mathbf{r} = (v \cos u, v \sin u, bu);$$

$$(2) \text{ 环面 } \mathbf{r} = ((a + r \cos v) \cos u, (a + r \cos v) \sin u, r \sin v);$$

$$(3) \text{ 双曲抛物面 } \mathbf{r} = (a(u+v), b(u-v), 2uv).$$

47. 证明 : 曲面是平面的充要条件是 $L = M = N = 0$.

48. 试证明 : 在半径为 a 的球面上 , 高斯曲率和平均曲率都是常数.

49. 求下列曲面的高斯曲率和平均曲率.

(1) 旋转曲面 $\mathbf{r} = (g(t)\cos\theta, g(t)\sin\theta, f(t))$;

(2) 正螺面 $\mathbf{r} = (v\cos u, v\sin u, bu)$;

(3) 螺面 $\mathbf{r} = (u\cos v, u\sin v, u + v)$.

50. 试证明 :

(1) 平面和球面都是全脐点曲面 ;

(2) 平面上任何一点都是平点 , 球面上任何一点的法曲率与方向无关.

51. 试证明 : 直纹面的高斯曲率 $K \leq 0$.

52. 求椭球旋转面 $\mathbf{r} = (a\cos u\sin v, a\sin u\sin v, c\cos v)$ 的高斯曲率.

补充题

1. 证明 : 曲线 $\mathbf{r} = (x(t), y(t), z(t))$ 为平面曲线的充要条件是

$$(\dot{\mathbf{r}}, \ddot{\mathbf{r}}, \dddot{\mathbf{r}}) = 0.$$

2. 设 D 是半径为 a 的球面 , 而 C 是一条空间曲线 , 方程为 $\mathbf{r} = \mathbf{r}(s)$ (s 为弧长参数) , 证明.

(1) 若曲线的所有的法平面与 D 相切 , 则

$$r^2 = \pm 2as + r^2(0);$$

(2) 若曲线的所有密切平面与 D 相切 , 则 $\mathbf{r}(s)$ 或为平面曲线或满足方程

$$\mathbf{r} = \pm a \left(\frac{\tau}{k} \mathbf{T} + \mathbf{B} \right).$$

3. $C: \begin{cases} F(x, y, z) = 0 \\ \varphi(x, y, z) = 0 \end{cases}$ 秩 $\begin{vmatrix} F'_x & F'_y & F'_z \\ \varphi'_x & \varphi'_y & \varphi'_z \end{vmatrix} = 2$ 求 C 的切线与法平面.

4. 已知曲线 $\mathbf{r} = \mathbf{r}(s)$ 的曲率 k 和挠率 τ , 试求向量函数 $\mathbf{W}(s)$ 使下式成立 :

$$\begin{cases} \mathbf{T}' = \mathbf{W} \times \mathbf{T} \\ \mathbf{N}' = \mathbf{W} \times \mathbf{N} \\ \mathbf{B}' = \mathbf{W} \times \mathbf{B} \end{cases}$$

5. 证明曲线

$$\mathbf{r} = \left(t, \frac{k_0}{2}t^2, \frac{k_0\tau_0}{6}t^3 \right)$$

在 $t = 0$ 处的曲率和挠率分别为 k_0 和 τ_0 , 而且 $\mathbf{T}(0), \mathbf{N}(0), \mathbf{B}(0)$ 分别重合于

x, y, z 轴.

6. 证明曲线

$$\mathbf{r} = \left(a \int_0^t \sin a(t) dt, a \int_0^t \cos a(t) dt, bt \right)$$

是一般螺线.

7. 设曲线 $\mathbf{r} = (at, bt^2, t^3)$ 是一般螺线, 求出 a, b 之间的关系, 并求出对应的固定向量 \mathbf{u} .

8. 证明下列条件之一是曲线 $\mathbf{r} = \mathbf{r}(s)$ 为一般螺线的充要条件.

$$(1) (\mathbf{N}, \mathbf{N}', \mathbf{N}'') = 0; \quad (2) (\mathbf{N}'', \mathbf{N}''', \mathbf{r}^{(4)}) = 0.$$

9. 已知曲线 $C: \mathbf{r} = (a \cos \omega t, a \sin \omega t, b\omega t)$, 对每个 C 上的点 P , 沿 P 的主法向量 \mathbf{N} 的方向取单位长的点 Q , 求 Q 点形成的曲线的方程.

10. 求曲面上参数曲线 (u 曲线和 v 曲线) 的二等分角轨线的微分方程.

11. 在曲面 $\mathbf{r} = \mathbf{r}(u, v)$ 上任一点处, 求两个切向量 $\mathbf{e}_1, \mathbf{e}_2$, 使之构成曲面的单位正交参数(坐标)系.

12. 证明: 坐标曲线与任意曲线的交角公式为:

$$u \text{ 线: } \cos \theta = \frac{E du + F dv}{\sqrt{E} ds},$$

$$v \text{ 线: } \cos \theta = \frac{F du + G dv}{\sqrt{G} ds}.$$

13. 证明: 曲面上的点是脐点的充要条件为 $H^2 = K$.

部分习题和补充题解答

习题

$$3. (1) \text{ 成立 } (2) \text{ 不成立. } \quad 5. (\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}}) = 0.$$

$$8. x - 1 = y - z - 1. \quad 9. y_0 x - x_0 y + y_0 z = x_0 y_0.$$

$$10. (1) \begin{cases} x = a \\ az + by = 0 \end{cases}, \quad ay + bz = 0;$$

$$(2) \frac{x-1}{1} = \frac{y-1}{2} = \frac{z-1}{3}, \quad x + 2y + 3z = 6;$$

$$(3) \frac{x-x_0}{A} = \frac{y-y_0}{B} = \frac{z-z_0}{C},$$

$$\text{其中 } A = \begin{vmatrix} f'_y & f'_z \\ g'_y & g'_z \end{vmatrix}, B = \begin{vmatrix} f'_z & f'_x \\ g'_z & g'_x \end{vmatrix}, C = \begin{vmatrix} f'_x & f'_y \\ g'_x & g'_y \end{vmatrix},$$

$$A(x-x_0) + B(y-y_0) + C(z-z_0) = 0.$$

$$11. (1) \frac{x-b}{0} = \frac{y-a}{ab} = \frac{z-1}{a^2}, \quad az - by = a;$$

$$(2) \frac{x-b}{-2ac} = \frac{y-a}{2ac} = \frac{z}{a^2-b^2}, \quad -2acx + 2bcy + (a^2-b^2)z = 0.$$

$$12. \frac{x-1}{11} = \frac{y-1}{8} = \frac{z-1}{-9}, \quad 11x + 8y - 9z = 10. \quad 14. \sqrt{3}(e^t - 1).$$

$$15. (1) \int_0^x \sqrt{1+y_x^2} dx, \quad \ln 3 - \frac{1}{2};$$

$$(2) \int_0^{\varphi_0} \sqrt{\rho^2 + \rho_1^2} d\varphi, \quad \frac{\sqrt{1+a^2}(e^{a\varphi_0} - 1)}{a}.$$

$$16. \mathbf{r} = (a \cos \omega s, a \sin \omega s, bs), \quad \omega = \frac{1}{\sqrt{a^2 + b^2}}.$$

$$17. \mathbf{T} = \frac{1}{\sqrt{2}}(0, 1, 1), \mathbf{B} = \frac{1}{\sqrt{3}}(1, 1, -1), \mathbf{N} = \frac{1}{\sqrt{6}}(2, -1, 1).$$

$$18. (1) k = \tau = \frac{1}{sa \operatorname{ch}^2 t};$$

$$(2) \frac{\sqrt{1+(\cos t - 1)^2}}{(3 - 2\cos t)^{\frac{3}{2}}}, \quad -\frac{1}{1+(\cos t - 1)^2}; \quad (3) \frac{2}{1+a^2}, \frac{3}{2} \frac{1}{1+a^2};$$

$$(4) \frac{2\sqrt{9t^4 + 9t^2 + 1}}{(9t^4 + 4t^2 + 1)^{\frac{3}{2}}} \frac{3}{(9t^4 + 9t^2 + 1)}.$$

$$19. \ddot{\mathbf{r}} // \ddot{\mathbf{r}}'. \quad 20. \ddot{\mathbf{r}}' = \mathbf{0}. \quad 27. (1) k = \tau = \frac{1}{\sqrt{2}}.$$

$$31. \mathbf{n} = (bc \cos \varphi \cos \theta, ac \sin \varphi \cos \theta, abs \sin \theta).$$

$$34. \mathbf{n} = \left(-\frac{u}{a^2}, -\frac{v}{b^2}, 1 \right); \mathbf{n} = (b(u+v), a(v-u), -ab).$$

$$35. \mathbf{n}^\circ = (-f' \cos v, -f' \sin v) \sqrt{1+f'^2}.$$

$$36. \mathbf{n} = (f' \sin v, -f' \cos v, u); \mathbf{n} \cdot (\boldsymbol{\rho} - \mathbf{r}) = 0; \boldsymbol{\rho} - \mathbf{r} = \lambda \mathbf{n}.$$

$$38. (f'_x, f'_y, -1).$$

$$42. \operatorname{sh} 1.$$

$$43. \cos \beta = \sqrt{\frac{|a^2 - 1|}{a^2 + 1}}.$$

$$44. -du^2/a(du^2 + dv^2).$$

$$45. \text{利用欧拉公式}.$$

$$46. (1) -\frac{2bdu dv}{\sqrt{b^2 + v^2}}; \quad (2) -\cos v(a + r \cos v) du^2 - r dv^2;$$

$$(3) -4ab/\sqrt{a^2 b^2 + b^2(u+v)^2 + a^2(u-v)^2}.$$

$$48. K = 1/R = H.$$

$$49. (1) K = f'(g'f'' - g''f') \mathcal{Y} g(g'^2 + f'^2)^2,$$

$$H = [g(g''f' - f''g') \mathcal{Y} f'(g'^2 + f'^2)] \mathcal{Y} 2g(g'^2 + f'^2)^{\frac{3}{2}};$$

$$(2) K = -b^2 \mathcal{K}(b^2 + v^2)^2, H = 0;$$

$$(3) K = -1 \mathcal{K}(1 + 2u^2)^2, H = (u^2 + 1) \mathcal{K}(2u^2 + 1)^{\frac{3}{2}}.$$

$$51. N = 0. II = -M^2, I > 0. \quad 52. K = c^2 \mathcal{K}(a^2 \cos^2 v + c^2 \sin^2 v)^2.$$

补充题

$$3. \frac{\frac{X-x}{D(F, \varphi)}}{\frac{D(y, z)}{D(z, x)}} = \frac{\frac{Y-y}{D(F, \varphi)}}{\frac{D(z, x)}{D(x, y)}} = \frac{\frac{Z-z}{D(F, \varphi)}}{\frac{D(x, y)}{D(y, z)}}, \quad \begin{vmatrix} X-x & Y-y & Z-z \\ F'_x & F'_y & F'_z \\ \varphi'_x & \varphi'_y & \varphi'_z \end{vmatrix} = 0.$$

$$6. b = 0; a \text{ 任意}.$$

$$9. r_Q = ((a-1)\cos \omega t, (a-1)\sin \omega t, b\omega t).$$

$$10. \sqrt{E} du \pm \sqrt{G} dv = 0.$$

$$11. \text{将}\{r_u, r_v\}\text{正交化 } e_1 = r_u / \sqrt{E};$$

$$e_2 = (-Fr_u / \sqrt{E}, \sqrt{E}r_v) \sqrt{EG - F^2}.$$

$$12. u \text{ 线 } \delta v = 0.$$

$$13. H^2 = (\lambda_1 + \lambda_2)^2 / 4 = K = \lambda_1 \lambda_2 \Rightarrow \lambda_1 = \lambda_2.$$

第 10 章 平面正交变换 仿射变换 射影变换

在第 7 章我们从代数角度讨论了一般欧氏空间的正交变换,本章将从几何角度讨论平面上两种特殊的点变换,即正交变换与仿射变换,给出平面二次曲线的正交分类与仿射分类.在本章中我们还将对平面射影几何作一简单介绍,包括齐次坐标,对偶原理,射影坐标系,交比,射影坐标变换,平面二次曲线的射影分类等.

10.1 平面正交变换

首先回顾正交变换的概念.

定义 10.1 平面的一个点变换,如果保持点之间的距离不变,则称它为**正交变换**(或**保距变换**).

从定义可知正交变换具有下述性质:

性质 1 正交变换的乘积仍是正交变换.

性质 2 恒等变换是正交变换.

性质 3 (1)正交变换把共线的三点变成共线的三点,且保持它们的顺序不变.

(2)正交变换把不共线的三点变成不共线的三点.

证 (1)设 A, B, C 是共线的三点,且 B 位于 A, C 之间,正交变换 σ 把 A, B, C 分别变成 A', B', C' , 则 $|A'B'| = |AB|, |A'C'| = |AC|, |B'C'| = |BC|$, 由假设, $|AB| + |BC| = |AC|$, 则 $|A'B'| + |B'C'| = |A'C'|$. 这表明 A', B', C' 三点共线, 且 B' 在 $A'C'$ 之间.

(2)设 D, E, F 是不共线的三点, 则 $|DE| + |EF| > |DF|$. 设正交变换 σ 把 D, E, F 分别变成 D', E', F' , 则 $|D'E'| + |E'F'| > |D'F'|$. 这表明 D', E', F' 三点不共线.

由性质 3 即得

性质 4 正交变换把直线变成直线,把线段变成线段.

性质 5 正交变换是可逆的,且逆变换仍是正交变换.

证 设 σ 是正交变换, A, B 是平面上不同的两个点, 则 $|\sigma(A)\sigma(B)|$

$= |AB| > 0$, 即 σ 把不同的点 A, B 变成不同的点 $\sigma(A), \sigma(B)$. 这证明 σ 是单射. 还可以证明 σ 是满射, 因而 σ 是双射. 把 σ 的逆变换记为 σ^{-1} . 易见 σ^{-1} 也是正交变换.

性质 6 正交变换把平行直线变成平行直线.

证 设 σ 是正交变换, 设 $l_1 // l_2$, 由性质 4 σ 把直线 l_i 变成直线 $l'_i, i = 1, 2$. 如果 l'_1 与 l'_2 交于 M' , 则 $\sigma^{-1}(M') \in l_1 \cap l_2$. 这与 $l_1 // l_2$ 矛盾, 所以 $l'_1 // l'_2$.

性质 7 正交变换保持向量长度与夹角.

证 设向量 $a = \overrightarrow{P_1 P_2}$, 定义 $\sigma(a) = \sigma(\overrightarrow{P_1 P_2}) = \overrightarrow{\sigma(P_1)\sigma(P_2)}$, 易见性质 7 成立.

性质 8 正交变换把平面直角坐标系变成平面直角坐标系.

证 设 $I = \{O; e_1, e_2\}$ 为平面直角坐标系, σ 为正交变换, 记 $O' = \sigma(O), e'_i = \sigma(e_i), i = 1, 2$. 则 $I' = \{O'; e'_1, e'_2\}$ 为平面直角坐标系 (由性质 7).

定理 10.1 设正交变换 σ 把直角标架 $I = \{O; e_1, e_2\}$ 变成直角标架 $I' = \{O'; e'_1, e'_2\}$. 其中 $\overrightarrow{OO'} = x_0 e_1 + y_0 e_2, e'_1 = a_{11} e_1 + a_{21} e_2, e'_2 = a_{12} e_1 + a_{22} e_2$, 即 $(e'_1, e'_2) = (e_1, e_2)A$. 则 σ 在直角坐标系 I 中的表示是

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad (10-1)$$

其中 $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ 是正交矩阵.

证 设 $\overrightarrow{OP} = xe_1 + ye_2 = (e_1, e_2) \begin{pmatrix} x \\ y \end{pmatrix}, P' = \sigma(P)$ 则 $\overrightarrow{OP'} = \sigma(\overrightarrow{OP}) = xe'_1 + ye'_2$, 设

$$\overrightarrow{OP'} = (e_1, e_2) \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (10-2)$$

由

$$\begin{aligned} \overrightarrow{OP'} &= \overrightarrow{OO'} + \overrightarrow{O'P'} = (e_1, e_2) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + (e'_1, e'_2) \begin{pmatrix} x \\ y \end{pmatrix} \\ &= (e_1, e_2) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + (e_1, e_2)A \begin{pmatrix} x \\ y \end{pmatrix}, \end{aligned} \quad (10-3)$$

结合 (10-2) 与 (10-3) 就得到 (10-1).

注 (10-1) 表示平面上的平移、旋转、反射或它们的复合.

10.2 平面的仿射变换

定义 10.2 (代数定义) 平面的一个点变换, 如果它在一个不一定正交的坐标系(通常称之为仿射坐标系或“仿射标架”)中的表示为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \quad (10-4)$$

其中系数矩阵 $A = (a_{ij})_{2 \times 2}$ 是可逆的(即 $|A| \neq 0$), 则称之为平面的仿射变换.

本书第 3 章中提到的镜面反射、比例变换、错切变换等都是平面上的仿射变换.

注 上述定义与仿射坐标系的选择无关. 从定义可见正交变换是仿射变换, 反之不一定对.

仿射变换有以下性质:

性质 1 仿射变换的乘积仍是仿射变换.

性质 2 仿射变换是可逆的, 且逆变换仍是仿射变换.

上述性质从仿射变换定义与可逆矩阵的性质容易得到.

性质 3 仿射变换把直线变成直线.

证 设 τ 是仿射变换, 它在一个仿射坐标系中的表示为(10-4). 由(10-4)有

$$\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} x' - x_0 \\ y' - y_0 \end{pmatrix}. \quad (10-5)$$

在平面上任取一条直线 $l: ax + by + c = 0$, 将(10-5)代入得 $l': a'x' + b'y' + c' = 0$, 它仍是一条直线.

性质 4 仿射变换把平行直线变成平行直线.

证 与上节性质 6 的证明类似.

注 通常在仿射变换下, 两点之间的距离是会改变的, 即距离不是一个仿射变换的不变量. 由于仿射变换把直线变成直线, 因此它把共线的三点变成共线的三点, 把不共线的三点变成不共线的三点.

可以证明: 平面仿射变换的另一等价定义为

定义 10.2 (几何定义) 平面到自身的 1-1 对应的变换称为仿射变换, 如果它把共线的三点变成共线的三点.

仿射变换有一个重要的不变量,即共线三点的简单比.

定义 10.3 设 A, B, C 是共线的三点, 设

$$\overrightarrow{AB} = \lambda \overrightarrow{BC},$$

λ 叫共线三点 A, B, C 的简单比, 记为 (A, C, B) .

定理 10.2 仿射变换保持共线三点的简单比不变.

证 设 τ 是仿射变换, 任取平面上共线三点 A, B, C , 取它们所在的直线为 x 轴, 建立一个直角坐标系. 设它们的横坐标分别为 x_1, x_2, x_3 . 如果把 A, B, C 分别变成 A', B', C' , 则 A', B', C' 仍共线. 作一个正交变换 σ , 使得 A', B', C' 所在的直线 l' 变成 x 轴, 其中 A', B', C' 分别变成 A'', B'', C'' . A'', B'', C'' 的横坐标分别为 x''_1, x''_2, x''_3 . 设 $\sigma\tau$ 的表示为

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

因为 $\sigma\tau$ 把 A, B, C 分别变成 A'', B'', C'' , 因此

$$\begin{pmatrix} x''_i \\ y''_i \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} x_i \\ 0 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad i = 1, 2, 3$$

即

$$x''_i = b_{11}x_i + b_1, \quad i = 1, 2, 3,$$

从而

$$\frac{\overrightarrow{A''B''}}{\overrightarrow{B''C''}} = \frac{x''_2 - x''_1}{x''_3 - x''_2} = \frac{x_2 - x_1}{x_3 - x_2} = \frac{\overrightarrow{AB}}{\overrightarrow{BC}}.$$

又因 σ 是正交变换, 则 $\frac{\overrightarrow{A''B''}}{\overrightarrow{B''C''}} = \frac{\overrightarrow{A'B'}}{\overrightarrow{B'C'}}$, 所以有

$$\frac{\overrightarrow{A'B'}}{\overrightarrow{B'C'}} = \frac{\overrightarrow{AB}}{\overrightarrow{BC}}, \quad \text{即} \quad (A', C', B') = (A, C, B).$$

定理 10.3 平面上任给两组不共线的三点 A_1, A_2, A_3 和 B_1, B_2, B_3 , 则存在唯一的仿射变换把 A_i 变成 B_i , $i = 1, 2, 3$.

证 因为 A_1, A_2, A_3 不共线, 所以 $I = \{A_1, \overrightarrow{A_1A_2}, \overrightarrow{A_1A_3}\}$ 是仿射标架. 同理 $II = \{B_1, \overrightarrow{B_1B_2}, \overrightarrow{B_1B_3}\}$ 也是仿射标架. 设 I 到 II 的过渡矩阵是 A, B_1 的坐标是 (x_0, y_0) . 构造仿射变换 τ :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

则 $\tau(A_i) = B_i$, $i = 1, 2, 3$. 唯一性请读者自证.

推论 10.1 平面上任给两个仿射标架 I 和 II , 则存在唯一的仿射变换把

I 变成 II .

正交变换保持两点之间的距离不变,保持向量之间的夹角不变,从而保持图形的面积不变.而一般的仿射变换会改变两点之间的距离,也会改变向量之间的夹角,因而也会改变图形的面积.现在我们讨论仿射变换改变图形面积的规律.

定理 10.4 设仿射变换 τ 在仿射标架 $I = \{O; e_1, e_2\}$ 中的表示为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

则对任意不平行的向量 a, b 有

$$\frac{|\tau(a) \times \tau(b)|}{|a \times b|} = \left| \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right|.$$

证 由于 τ 是仿射变换,因此它把仿射标架 $I = \{O; e_1, e_2\}$ 变成仿射标架 $II = \{O'; e'_1, e'_2\}$. 其中

$$(e'_1, e'_2) = (e_1, e_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

设 $a = u_1 e_1 + v_1 e_2$, $b = u_2 e_1 + v_2 e_2$,

则 $\tau(a) = u_1 e'_1 + v_1 e'_2$, $\tau(b) = u_2 e'_1 + v_2 e'_2$.

直接计算可验证: $|\tau(a) \times \tau(b)| = |a \times b| |\det(A)|$.

利用极限理论可证下列推论:

推论 10.2 若平面上任一有面积的区域 D 经过仿射变换变成区域 D' , 则有

$$\frac{S_{D'}}{S_D} = |\det(A)|,$$

其中 $S_D, S_{D'}$ 分别表示 D, D' 的面积, A 为仿射变换公式(10.4)中的系数矩阵.

这一推论说明,仿射变换 τ 按一个比值 $|\det(A)|$ 改变平面上所有(有面积的)图形的面积,因此把 $|\det(A)|$ 称为仿射变换的变积系数.

在第 8 章中我们讨论了平面二次曲线的正交分类.现在来看它们的仿射分类.

设二次曲线的方程为

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c = 0. \quad (10-6)$$

我们知道,二次曲线(10-6)都可经过正交变换化为下列标准形之一:

$$\frac{(x')^2}{a^2} + \frac{(y')^2}{b^2} = \pm 1; \quad \frac{(x')^2}{a^2} - \frac{(y')^2}{b^2} = 1; \quad (y')^2 = 2px', p > 0;$$

$$\frac{(x')^2}{a^2} \pm \frac{(y')^2}{b^2} = 0; \quad (y')^2 = a.$$

现在设仿射变换为

$$\begin{cases} x = g_{11}x' + g_{12}y' + h_1 \\ y = g_{21}x' + g_{22}y' + h_2 \end{cases}, \quad \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \neq 0. \quad (10-7)$$

下面来看二次曲线(10-6)在仿射变换(10-7)下的标准形.

定理 10.5 二次曲线(10-6)可通过仿射变换化为下列标准形之一:

$$(x'')^2 + (y'')^2 = \pm 1, \quad (x'')^2 - (y'')^2 = 1, \quad (y'')^2 = x'',$$

$$(x'')^2 \pm (y'')^2 = 0, \quad (y'')^2 = \pm 1, \quad (y'')^2 = 0.$$

证 因为正交变换与仿射变换的复合仍是仿射变换,我们先通过正交变换把(10-6)变成正交标准型.然后再作仿射变换

$$\begin{cases} x' = ax'' \\ y' = by'' \end{cases}.$$

方程 $\frac{(x')^2}{a^2} + \frac{(y')^2}{b^2} = \pm 1$, $\frac{(x')^2}{a^2} - \frac{(y')^2}{b^2} = 1$, $\frac{(x')^2}{a^2} \pm \frac{(y')^2}{b^2} = 0$ 就分别变成

$$(x'')^2 + (y'')^2 = \pm 1, \quad (x'')^2 - (y'')^2 = 1, \quad (x'')^2 \pm (y'')^2 = 0.$$

如果作仿射变换 $\begin{cases} x' = \frac{1}{2p}x'' \\ y' = y'' \end{cases}$, 方程 $(y')^2 = 2px'$ 就变成 $(y'')^2 = x''$. 当 $a > 0$,

作仿射变换 $\begin{cases} x' = x'' \\ y' = \sqrt{a}y'' \end{cases}$, 方程 $(y')^2 = a$ 就变成 $(y'')^2 = 1$; 当 $a < 0$, 作仿射变

换 $\begin{cases} x' = x'' \\ y' = \sqrt{-a}y'' \end{cases}$, 方程 $(y')^2 = a$ 就变成 $(y'')^2 = -1$. 至此完成了定理的证明.

10.3 射影平面与齐次坐标

扩大的笛卡儿平面 在实际生活中我们常遇到把一个平面图形映射到另一个平面图形,例如,航空摄影把地面(假定是平的)上的景物摄到底片上,这可以看成是地面 π_0 到底片 π_1 的一种映射.飞机在飞行中,从不同的角度拍摄同一景物,其中各物体的相对距离或角度可能都不相同,但人们仍能从照片上进行识别.这说明从地面景物到底片之间的映射中,必有某种几何性质或几

何量是保持不变的. 在欧氏几何中, 我们知道在正交变换下, 长度和角度, 以及直线的平行, 曲线的斜率和曲率都是不变的. 而在现在所说的映射中, 这些都是可变的, 然而它却保持另外一些不变的性质, 例如直线变为直线, 两条直线的相交性, 三线交于一点或三点共线, 等等. 这一节我们就来研究这些所谓“射影变换”的一些性质.

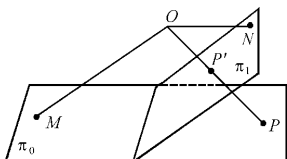


图 10 - 1

设 π_0, π_1 是空间中两个相交但不重合平面(见图 10 - 1), O 点是不在这两个平面上的点. 对于 π_0 上任一点 P , 连接 OP 的直线与 π_1 交于 P' , 这样建立的 π_0 和 π_1 上点之间的对应, 称为中心投影(透视), 点 O 称为投影中心(试与摄影比较一下).

射影变换就是由有限个中心透视投影所组合而成的变换(或称对应). 一个平面几何的性质或事物, 如果在任何中心透视下保持不变(当然在它们的组合之下也保持不变), 就称之为一种“射影性质”. 例如, 直线, 或共线的点列都是射影性质, 也是射影几何讨论的对象. 又例如, 两线相交, 三点共线, 三线共点, 两条曲线相切等等都是射影性质. 而长度, 角度, 面积等几何量则不是射影不变的, 所以也就不是射影性质. 从此可见, 射影几何学所研究的是一些比长度, 角度等欧氏几何性质更加基本的一些几何性质.

但是, 上面所描述的中心透视(投影)还有不足之处, 需要引进适当的概念来加以补充. 例如, 如果 π_0 上的点 M 满足条件 $OM \parallel \pi_1$, 则 M 在 π_1 上没有像. 同样地, π_1 上的点 N 若满足 $ON \parallel \pi_0$, 则 N 在 π_0 上没有原像. 于是为了保证在中心投影下, π_0 和 π_1 上点的 1 - 1 对应, 我们应该在平面上增加一些点, 叫做无穷远点, 使得

(1) 每一直线在无穷远处有一无穷远点(直线两端的无穷远点是同一个点);

(2) 平行直线上的无穷远点是同一个点(即: 平行直线交于无穷远点);

(3) 平面上所有无穷远点都在一条直线 l_∞ 上, 称 l_∞ 为无穷远直线.

例如:在图 10-1 中, π_0 上 M 这样的点形成一条直线 l , l 的像就是 π_1 上的无穷远直线.

注 这种给平面补充一条无穷远直线的思想归功于十七世纪天文学家 Kepler.

定义 10.4 添加了无穷远直线 l_∞ 的笛卡儿平面称为扩大的笛卡儿平面, 记为 E .

扩大的平面 E 有两个重要特点:

(1) E 上任意两点(包含无穷远点)必在一直线上;

(2) E 上任意两条不同直线交于一点.

引入扩大的笛卡儿平面 E 后, 保证了中心投影下的 1-1 对应. 在中心投影下, 扩大平面上的“普通点”和“无穷远点”并无本质差别.

定义 10.5 将“普通点”和“无穷远点”一视同仁, 不加区别的笛卡儿扩大平面叫射影平面. 在射影平面上, 任何两点在一条直线上, 任意两条直线必相交于一点.

齐次坐标 有了扩大平面, 就需引进它上面点的坐标. 由于在扩大平面上, 任意一点都是两条直线的交点, 设 l 与 l' 是扩大平面上的两条普通直线, 它们的方程分别为 $l: Ax + By + C = 0$, $l': A'x + B'y + C' = 0$, 如果 l 与 l' 不平行, 交点 (x, y) 就是

$$x = \frac{\begin{vmatrix} -C & B \\ -C' & B' \end{vmatrix}}{\begin{vmatrix} A & B \\ A' & B' \end{vmatrix}} = \frac{\begin{vmatrix} B & C \\ B' & C' \end{vmatrix}}{\begin{vmatrix} A & B \\ A' & B' \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} A & -C \\ A' & -C' \end{vmatrix}}{\begin{vmatrix} A & B \\ A' & B' \end{vmatrix}} = \frac{\begin{vmatrix} C & A \\ C' & A' \end{vmatrix}}{\begin{vmatrix} A & B \\ A' & B' \end{vmatrix}}.$$

如果 $l \parallel l'$ (但不重合), 交点是一个无穷远点. 此时, $\begin{vmatrix} A & B \\ A' & B' \end{vmatrix} = 0$, 但是

$\begin{vmatrix} B & C \\ B' & C' \end{vmatrix}$ 与 $\begin{vmatrix} C & A \\ C' & A' \end{vmatrix}$ 不同时为零, 这时无法用普通坐标 (x, y) 表示交点. 如

果改用三数组 $(\begin{vmatrix} B & C \\ B' & C' \end{vmatrix}, \begin{vmatrix} C & A \\ C' & A' \end{vmatrix}, \begin{vmatrix} A & B \\ A' & B' \end{vmatrix})$ 作为交点坐标, 且认为与之成比例的 $(\lambda \begin{vmatrix} B & C \\ B' & C' \end{vmatrix}, \lambda \begin{vmatrix} C & A \\ C' & A' \end{vmatrix}, \lambda \begin{vmatrix} A & B \\ A' & B' \end{vmatrix})$ 表示同一点, 则直线 l 与 l' 的

交点可统一地用它们的系数 A, B, C 与 A', B', C' 表示为 $(\begin{vmatrix} B & C \\ B' & C' \end{vmatrix}, \begin{vmatrix} C & A \\ C' & A' \end{vmatrix}, \begin{vmatrix} A & B \\ A' & B' \end{vmatrix})$. 当 l, l' 交于普通点时, 交点就是 $(x, y, 1)$ (这就是普通

点的坐标 (x, y)). 当 $l \parallel l'$ (但不重合) 交点就是 $\left(\begin{vmatrix} B & C \\ B' & C' \end{vmatrix}, \begin{vmatrix} C & A \\ C' & A' \end{vmatrix}, 0 \right) = \rho_1(-B, A, 0) = \rho_2(-B', A', 0)$ (ρ_1, ρ_2 均为非零实数). 最后一数为零, 这正好表示它是 l (或 l') 上的无穷远点. 一个明显的事实是: 三数组中的三个数不全为零. 因此在扩大的平面上, 可以用非零三元有序数组 (x_1, x_2, x_3) 来表示点的坐标. 这里要求

(1) 成比例的三元有序数组表示同一点;

(2) $(x_1, x_2, 0)$ (x_1, x_2 不全为0) 表示无穷远点;

(3) (x_1, x_2, x_3) ($x_3 \neq 0$) 表示普通点 (x, y) , 这时 $x = \frac{x_1}{x_3}, y = \frac{x_2}{x_3}$.

这样的坐标称为射影平面上的齐次坐标.

射影平面上的直线有两类: 无穷远直线和普通直线添加上无穷远点.

无穷远直线的方程为 $x_3 = 0$, 普通直线的方程为 $A \frac{x_1}{x_3} + B \frac{x_2}{x_3} + C = 0$, 添加上无穷远点 $(-B, A, 0)$ 后, 方程为三元一次齐次方程: $Ax_1 + Bx_2 + Cx_3 = 0$, 其中 A, B 不同时为零. 总之射影平面上的方程 $Ax_1 + Bx_2 + Cx_3 = 0$, 其中 A, B, C 不同时为零, 被称为射影直线. $A = B = 0$ 对应无穷远直线, A, B 不同时为零对应普通直线.

注 上面我们是从普通的笛卡儿平面出发, 添加一条无穷远直线后得到的射影平面. 但具有上述特点的还有其它模型, 下面介绍三个:

(1) 直线把: 过一个定点的所有直线和所有平面所成的集合称为直线把. 将把中的每一直线看成一个“点”, 每一平面看成一条“直线”, 则明显它有射影平面的特点: 任意两“点”在一“直线”上 (对应直线把中任意两直线在一平面上); 任意两“直线”交于一“点” (对应直线把中任意两平面有唯一交线). 这样的直线把模型将把中的直线视为“点”, 将把中的平面视为“直线”, 也可把它看成是一个射影平面.

(2) 球面模型: 在直线把模型中, 取一个球心在直线把中心的球面, 直线把中的直线交球面于两点 (即球面上的对径点), 直线把中的平面交球面于大圆. 于是得到射影平面的球面模型. 将球面上的对径点粘在一起作为一“点”, 将球面的大圆作为“直线”. 这个模型也可以看成是一个射影平面.

(3) 代数模型: 在全体非零三数组的集合 $\{(x_1, x_2, x_3)\}$ 中 (此处非零是指 x_1, x_2, x_3 不全为0), 引进等价关系“ \sim ” ($(x_1, x_2, x_3) \sim (y_1, y_2, y_3)$), 当且仅当存在非0的 λ 使得 $(x_1, x_2, x_3) = \lambda(y_1, y_2, y_3)$; “ \sim ”是一个等价关系. 记

由此决定的等价类为 $[(x_1, x_2, x_3)]$, 把每个等价类看作一点, 其全体构成一个射影平面 $P^2, P^2 = \{(x_1, x_2, x_3)\} \sim$.

对偶原理 在欧氏平面上建立了坐标系以后, 平面上每一点都有一个坐标, 而每一条直线都有一个方程. 点没有方程, 而直线也没有坐标. 但在射影平面上就不同了. 在射影平面上, 点的齐次坐标为 (x_1, x_2, x_3) , 与之成比例的点 $(\lambda x_1, \lambda x_2, \lambda x_3) (\lambda \neq 0)$ 表示同一点. 射影平面上直线的方程为 $u_1 x_1 + u_2 x_2 + u_3 x_3 = 0$ (不全为零), 自然 $\lambda u_1 x_1 + \lambda u_2 x_2 + \lambda u_3 x_3 = 0$ 也表示同一直线. 直线与非零三元有序数组 (u_1, u_2, u_3) 对应, 而且与之成比例的 $(\lambda u_1, \lambda u_2, \lambda u_3)$ 也表示同一直线. 于是称 (u_1, u_2, u_3) 为直线 $u_1 x_1 + u_2 x_2 + u_3 x_3 = 0$ 的齐次坐标. 同样, 固定 (x_1, x_2, x_3) , 上面的方程又可看成是所有过这点直线的轨迹, 因此就可以把它看成是这个点的“方程”.

用 X, U 分别表示点 (x_1, x_2, x_3) 和直线 (u_1, u_2, u_3) 的坐标, 下面讨论点与直线的结合关系: 点 X 在直线 U 上的充分必要条件是 $U \cdot X = u_1 x_1 + u_2 x_2 + u_3 x_3 = 0$ (这里 $U \cdot X$ 是借用向量内积的符号). 直线 U 通过点 X 的充分必要条件也是 $U \cdot X = u_1 x_1 + u_2 x_2 + u_3 x_3 = 0$. 两点 X_1, X_2 决定直线 U , 它们的坐标应满足方程: $U \cdot X_1 = 0, U \cdot X_2 = 0$ 或 $U = X_1 \times X_2$ (这里是借用向量外积的符号). 两条直线 U_1, U_2 决定交点 $X: U_1 \cdot X = 0, U_2 \cdot X = 0$ 则有 $X = U_1 \times U_2$. 三点 X_1, X_2, X_3 在一条直线上的充要条件是 $(X_1, X_2, X_3) = 0$ (这里 (X_1, X_2, X_3) 是借用向量混合积的符号). 三直线 U_1, U_2, U_3 共点的充要条件是 $(U_1, U_2, U_3) = 0$. 点 X 在 X_1, X_2 决定的直线 U 上的充要条件是 $X \cdot U = 0$, 即 $(X, X_1, X_2) = 0$, 也即 X 是 X_1, X_2 的线性组合: $X = \lambda X_1 + \mu X_2$, 直线 U 通过 U_1, U_2 决定的交点 X 的充要条件是 $U \cdot X = 0$, 即 $(U, U_1, U_2) = 0$, 也即 U 是 U_1, U_2 的线性组合: $U = \lambda U_1 + \mu U_2$.

在射影平面上, 点和直线的齐次坐标都是非零有序三元数组 (可差一常数倍). 点和直线间的结合关系包含: 点在直线上和直线经过点, 它们都用同一个三元一次方程来表示. 若仅研究射影平面上点和直线的结合关系, 则上面的讨论说明, 射影平面上点和直线的地位是对称的, 即用点和直线的齐次坐标, 在公式和命题的证明过程中看不出点和直线的区别. 为此引进下面定义:

定义 10.6 设 φ (点, 线)是关于射影平面上一些点和一些直线的结合关系的命题, 那么, 把此命题中的点换成直线, 把直线换成点, 把“点在直线上”换成“直线经过点”, 同时把“直线经过点”换成“点在直线上”, 则得到一个新命题 $\tilde{\varphi}$ (线, 点)称为原命题 φ (点, 线)的对偶命题.

注 一个几何图形在这种改换下,就成为它的对偶图形.例如:点的对偶图形是直线;两点形(两点及其连线)的对偶图形是两线形(二直线及其交点);三点形(不共线三点及两两连线)的对偶图形是三角形(不共点的三直线及其两两的交点).

下面我们列举一些命题和它们的对偶命题

原命题

对偶命题

(1)射影平面上三点共线的充要条件是它们的齐次坐标组成的三阶行列式等于零.

(1)射影平面上三线共点的充要条件是它们的齐次坐标组成的三阶行列式等于零.

(2)射影平面上若三点 P_1, P_2, P_3 不共线,则三线 P_1P_2, P_2P_3, P_3P_1 不共点.

(2)射影平面上若三线 l_1, l_2, l_3 不共点,则三点 l_1l_2, l_2l_3, l_3l_1 不共线.

(3)德沙格(Desargues)定理:射影平面上,如果两个三角形的对应顶点的连线共点,那么它们的对应边的交点共线.(见图 10-2)

(3)德沙格(Desargues)定理的逆定理:射影平面上,如果两个三角形的对应边的交点共线,那么它们对应顶点的连线共点.

(4)直线(看成点列)的点方程是三元一次齐次方程.

(4)点(看成线束)的线方程是三元一次齐次方程.

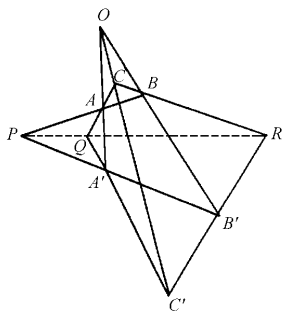


图 10-2

下面我们来证一下德沙格定理 设三角形 ABC 与三角形 $A'B'C'$ 的对应顶点的连线 AA', BB', CC' 相交于一点 O . 以下证对应边的交点 P (AB 与 $A'B'$ 的交点), Q (AC 与 $A'C'$ 的交点), R (BC 与 $B'C'$ 的交点)三点共线. 因为

O 点在直线 AA' , BB' 和 CC' 上, 适当选取 A, A', B, B', C, C' 的齐次坐标, 可使 $d = a + a' = b + b' = c + c'$, 这里 d, a, a', b, b', c, c' 分别是点 O 及 A, A', B, B', C, C' 的齐次坐标. 由上式令 $p = a - b = b' - a', q = c - a = a' - c', r = b - c = c' - b'$. 可看出, p, q, r 分别为 P, Q, R 的齐次坐标. 因为 $p + q + r = 0$, 所以 P, Q, R 在一条直线上. (见图 10-2).

射影坐标系 射影坐标变换 在欧氏平面上, 确定原点及两个线性无关向量就形成一个仿射坐标系. 在射影平面上, 一个点 M 的齐次坐标 (x_1, x_2, x_3) 有三个分量, 从形式上自然想到可以取三个基 $(1, 0, 0), (0, 1, 0), (0, 0, 1)$; 于是 M 就可分解为

$$(x_1, x_2, x_3) = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1), \quad (10-8)$$

其系数就是关于这个基的坐标. 把 $O_1(0, 0, 1), O_2(1, 0, 0), O_3(0, 1, 0)$ 作为三点形 OO_1O_2 的三顶点的特定齐次坐标 (在扩大平面上它们分别表示坐标原点, x 轴上的无穷远点和 y 轴上的无穷远点). 称 OO_1O_2 为坐标三点形. 然而我们发现光这三个点还不足以决定其它点的坐标. 因为 O_1, O_2, O 的齐次坐标还可以是 $(0, 0, c), (a, 0, 0), (0, b, 0)$, 而点 $M(x_1, x_2, x_3)$ 关于这组坐标的分解系数就不再是 x_1, x_2, x_3 了. 所以还需添加一个“单位点” $E(1, 1, 1)$, 它关于 O_1, O_2, O 的分解系数 (坐标) 为 $1, 1, 1$. 这就规范了 $a = b = c$. 从而唯一确定平面上点的齐次坐标 (在差一个常数倍意义下的唯一!). 称 E 为单位点, 而称 $\{O_1, O_2, O; E\}$ 为射影平面上的原始射影坐标系或原始标架.

注 坐标系 $\{O_1, O_2, O; E\}$ 中, 没有三点在一直线上.

定义 10.7 射影平面上任意四点 A, B, C 和 D , 其中没有三点在一直线上, 其齐次坐标分别为 $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$ 和 (d_1, d_2, d_3) , 称 $\{A, B, C; D\}$ 为射影坐标系. 其中 C, A, B 为坐标三点形, D 为单位点.

射影平面上, 任意点 $M(x_1, x_2, x_3)$ (其中 (x_1, x_2, x_3) 是 M 在原始射影坐标系下的坐标) 在射影坐标系 $\{A, B, C; D\}$ 中的坐标如何决定呢? 请看下面两个例子:

例 1 设点 M 在原始坐标系中的坐标为 $(-1, -1, 0)$, 求它在新射影坐标系 $\{A(1, -1, 2), B(0, -1, 2), C(3, 1, -4); D(1, 0, -2)\}$ 中的新坐标 (x'_1, x'_2, x'_3)

解 首先用单位点 D 决定 A, B, C 的特定坐标. 即单位点应满足如下方程:

$$(1, 0, -2) = \lambda(1, -1, 2) + \mu(0, -1, 2) + \nu(3, 1, -4)$$

或

$$\begin{cases} 1 = \lambda + 3v \\ 0 = -\lambda - \mu + v \\ -2 = 2\lambda + 2\mu - 4v. \end{cases}$$

解得 $\lambda = -2, \mu = 3, v = 1$, 即 A, B, C 应取的特定(规范化)坐标分别为 $(-2, 2, -4), (0, -3, 6), (3, 1, -4)$. 取定后 M 的新坐标应满足以下方程:

$$(-1, -1, 0) = x'_1(-2, 2, -4) + x'_2(0, -3, 6) + x'_3(3, 1, -4).$$

解得 $(x'_1, x'_2, x'_3) = (2, 2, 1)$. 因而 M 的新坐标为 $\rho(2, 2, 1)$, 其中 ρ 为任意非零实数.

通过这个例子, 我们可以建立原始射影坐标系 $\{O_1, O_2, O; E\}$ 和射影坐标系 $\{A, B, C; D\}$ 间的射影坐标变换公式如下: 设 $\{A, B, C; D\}$ 中各点在原始射影坐标系 $\{O_1, O_2, O; E\}$ 中的坐标为 $A(a_1, a_2, a_3), B(b_1, b_2, b_3), C(c_1, c_2, c_3)$ 和 $D(d_1, d_2, d_3)$, 不妨假定 $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3)$ 是 D 取为单位点的特定坐标, 即

$$(d_1, d_2, d_3) = (a_1, a_2, a_3) + (b_1, b_2, b_3) + (c_1, c_2, c_3), \quad (10-9)$$

则射影平面上任意一点 $M(x_1, x_2, x_3)$ 在射影坐标系 $\{A, B, C; D\}$ 中的齐次坐标 (x'_1, x'_2, x'_3) 由下列公式确定:

$$(x_1, x_2, x_3) = \rho[x'_1(a_1, a_2, a_3) + x'_2(b_1, b_2, b_3) + x'_3(c_1, c_2, c_3)],$$

即

$$\begin{cases} x_1 = \rho(a_1 x'_1 + b_1 x'_2 + c_1 x'_3) \\ x_2 = \rho(a_2 x'_1 + b_2 x'_2 + c_2 x'_3) \\ x_3 = \rho(a_3 x'_1 + b_3 x'_2 + c_3 x'_3), \end{cases}$$

或

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \rho \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}, \quad (10-10)$$

其中 ρ 是任意非零常数, 这是因为齐次坐标允许相差一个非零常数. (10-10) 中右边的系数矩阵是满秩的, 称(10-10)式为射影坐标变换公式.

反过来, 给定射影平面上满秩线性变换(10-10), 它一定表示某个射影坐标变换.

注 在射影坐标变换公式(10-10)中, 待定系数并不是九个 $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3$, 更不是十个(加上 ρ), 而是八个, 因为与 $(a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3)$ 九个数成比例的任意九个数 $(\tilde{a}_1, \tilde{a}_2, \tilde{a}_3, \tilde{b}_1, \tilde{b}_2, \tilde{b}_3, \tilde{c}_1, \tilde{c}_2, \tilde{c}_3)$

\sim_{c_3}) 决定同一个射影坐标变换. 所以独立的参数仅有八个. 从另一个角度也可说明这点: 因 $(10-10)$ 由 A, B, C, D 的坐标决定, 每个点的齐次坐标是三个参数, 但齐次坐标允许差一常数倍数, 故独立的参数只有两个, 四个点共计八个. 因此由射影平面上已知四点 (无三点共线) 的新, 旧坐标, 就可以唯一确定一个射影坐标变换.

例 2 已知在射影坐标变换下, 四个点 $A(1, 1, -2), B(-1, 2, -1), C(5, -7, -2), D(0, -3, -1)$ 的新坐标分别是 $(1, -1, 2), (0, -1, 2), (3, 1, -4), (1, 0, -2)$, 求射影坐标变换公式.

解 我们把 A, B, C, D 及其象看成两组坐标系, 为此分别把它们规范化, 得到 $(-2, -2, 4), (-3, 6, -3), (5, -7, -2)$ 及 $(-2, 2, -4), (0, -3, 6), (3, 1, -4)$. 设所求变换的矩阵为 P , 即

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \rho P \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix},$$

则有

$$Q = \begin{pmatrix} -2 & -3 & 5 \\ -2 & 6 & -7 \\ 4 & -3 & -2 \end{pmatrix} = P \begin{pmatrix} -2 & 0 & 3 \\ 2 & -3 & 1 \\ -4 & 6 & -4 \end{pmatrix} = PR.$$

从而得到

$$P = QR^{-1} = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix},$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \rho \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

交比 前面已经说过, 为什么同一景象从不同的角度投影到平面上后, 我们仍能识别. 这就提示我们, 虽然射影变换改变了物体之间的长度, 角度等几何量, 甚至改变了简单比 (它是仿射不变量), 但一定还有这些量的某种比例关系却是不变的. 我们来找这类比例关系.

平面上两个景物 A, B 称为是“等价的”, 如果可以用有限个射影变换把 A 变为 B . 这是一个等价关系. 我们的问题是找一个依赖于景象的数, 它对同一等价类的元素是一样的, 这就是所谓“射影不变量”. 对直线上的 4 个点, 人们发现它们之间距离的“比值之比” (称为“交比”) 就是一个射影不变量.

在普通平面上,共线四点 P_1, P_2, P_3, P_4 的交比定义为

$$(P_1, P_2; P_3, P_4) = \frac{\frac{\overrightarrow{P_1P_3}}{\overrightarrow{P_3P_2}}}{\frac{\overrightarrow{P_1P_4}}{\overrightarrow{P_4P_2}}} = \frac{\overrightarrow{P_1P_3} \cdot \overrightarrow{P_2P_4}}{\overrightarrow{P_1P_4} \cdot \overrightarrow{P_2P_3}},$$

由第一个等号,它是两个简单比的比值,故称为二重比;由第二个等号,它是线段交叉乘积之比,故称为交比.

在扩大平面上,若四个普通点 P_1, P_2, P_3, P_4 共线,则 $P_3 = \lambda P_1 + \mu P_2$, $P_4 = \lambda' P_1 + \mu' P_2$, 其中 P_1, P_2, P_3, P_4 同时也表示点 P_1, P_2, P_3, P_4 的齐次坐标向量. 适当选择坐标系,可设 $P_i = (x_1^{(i)} \ 0 \ x_3^{(i)})$, P_i 的非齐次坐标为 $(x^{(i)} \ 0 \ 1 \ i = 1, 2, 3, 4)$ 则

$$\begin{aligned} (P_1, P_2; P_3, P_4) &= \frac{\frac{x^{(3)} - x^{(1)}}{x^{(3)} - x^{(2)}}}{\frac{x^{(4)} - x^{(1)}}{x^{(4)} - x^{(2)}}} = \frac{\frac{\frac{x_1^{(3)}}{x_3^{(3)}} - \frac{x_1^{(1)}}{x_3^{(1)}}}{\frac{x_1^{(3)}}{x_3^{(3)}} - \frac{x_1^{(2)}}{x_3^{(2)}}}}{\frac{\frac{x_1^{(4)}}{x_3^{(4)}} - \frac{x_1^{(1)}}{x_3^{(1)}}}{\frac{x_1^{(4)}}{x_3^{(4)}} - \frac{x_1^{(2)}}{x_3^{(2)}}}} = \frac{\frac{\frac{\lambda x_1^{(1)} + \mu x_1^{(2)}}{\lambda x_3^{(1)} + \mu x_3^{(2)}} - \frac{x_1^{(1)}}{x_3^{(1)}}}{\frac{\lambda x_1^{(1)} + \mu x_1^{(2)}}{\lambda x_3^{(1)} + \mu x_3^{(2)}} - \frac{x_1^{(2)}}{x_3^{(2)}}}}{\frac{\frac{\lambda' x_1^{(1)} + \mu' x_1^{(2)}}{\lambda' x_3^{(1)} + \mu' x_3^{(2)}} - \frac{x_1^{(1)}}{x_3^{(1)}}}{\frac{\lambda' x_1^{(1)} + \mu' x_1^{(2)}}{\lambda' x_3^{(1)} + \mu' x_3^{(2)}} - \frac{x_1^{(2)}}{x_3^{(2)}}}} \\ &= \frac{\frac{\mu(x_1^{(2)}x_3^{(1)} - x_3^{(2)}x_1^{(1)})}{x_3^{(1)}}}{\frac{\lambda(x_1^{(2)}x_3^{(1)} - x_3^{(2)}x_1^{(1)})}{x_3^{(2)}}} = \frac{\mu}{\lambda} = \frac{\mu\lambda'}{\lambda\mu'}. \end{aligned}$$

由此引出下面的定义:

定义 10.8 射影平面上共线四点 P_1, P_2, P_3, P_4 的齐次坐标向量 P_1, P_2, P_3, P_4 如果满足 $P_3 = \lambda P_1 + \mu P_2$, $P_4 = \lambda' P_1 + \mu' P_2$ 则其交比定义为

$$(P_1, P_2; P_3, P_4) = \frac{\frac{\mu}{\lambda}}{\frac{\mu'}{\lambda'}} = \frac{\mu\lambda'}{\lambda\mu'}.$$

注 交比定义与齐次坐标的选取无关. 因若取 P_1, P_2, P_3, P_4 的齐次坐标分别为 $\rho_1 P_1, \rho_2 P_2, \rho_3 P_3, \rho_4 P_4$ 则

$$\rho_3 P_3 = \lambda(\rho_1 P_1) + \mu(\rho_2 P_2), \quad \rho_4 P_4 = \lambda'(\rho_1 P_1) + \mu'(\rho_2 P_2),$$

即

$$P_3 = \frac{\lambda \rho_1}{\rho_3} P_1 + \frac{\mu \rho_2}{\rho_3} P_2 = \tilde{\lambda} P_1 + \tilde{\mu} P_2,$$

$$P_4 = \frac{\lambda' \rho_1}{\rho_4} P_1 + \frac{\mu' \rho_2}{\rho_4} P_2 = \tilde{\lambda}' P_1 + \tilde{\mu}' P_2.$$

同样有

$$(P_1, P_2; P_3, P_4) = \frac{\frac{\tilde{\mu}}{\tilde{\lambda}}}{\frac{\tilde{\mu}'}{\tilde{\lambda}'}} = \frac{\frac{\tilde{\mu} \tilde{\lambda}'}{\tilde{\lambda} \tilde{\mu}'}}{\frac{\tilde{\mu} \tilde{\lambda}'}{\tilde{\lambda} \tilde{\mu}'}} = \frac{\mu \lambda'}{\lambda \mu'}.$$

对偶地,也可以定义共点四线 l_1, l_2, l_3, l_4 的交比. 如果 $l_3 = \lambda l_1 + \mu l_2, l_4 = \lambda' l_1 + \mu' l_2$, 此处 l_1, l_2, l_3, l_4 同时也表示 l_1, l_2, l_3, l_4 的齐次坐标向量, 则交比定义为

$$(l_1, l_2; l_3, l_4) = \frac{\frac{\mu}{\lambda}}{\frac{\mu'}{\lambda'}} = \frac{\mu \lambda'}{\lambda \mu'}.$$

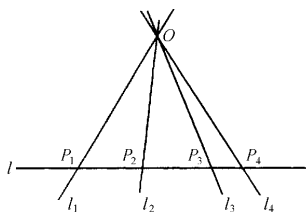


图 10-3

定理 10.6 设四直线 l_1, l_2, l_3, l_4 共点, 直线 l 与它们分别交于四点 P_1, P_2, P_3, P_4 , 则有(见图 10-3)

$$(l_1, l_2; l_3, l_4) = (P_1, P_2; P_3, P_4).$$

证 设 l 的齐次坐标为 U , l_i 的齐次坐标为 $U^{(i)}$, 则交点 P_i 的齐次坐标为 $U \times U^{(i)}$, 因 l_1, l_2, l_3 共点, l_1, l_2, l_4 共点, 可设

$$U^{(3)} = \lambda U^{(1)} + \mu U^{(2)}, \quad U^{(4)} = \lambda' U^{(1)} + \mu' U^{(2)}.$$

由定义知

$$(l_1, l_2; l_3, l_4) = \frac{\frac{\mu}{\lambda'}}{\frac{\mu'}{\lambda}} = \frac{\mu\lambda'}{\lambda\mu'},$$

又因

$$\begin{aligned} U \times U^{(3)} &= \lambda U \times U^{(1)} + \mu U \times U^{(2)}, \\ U \times U^{(4)} &= \lambda' U \times U^{(1)} + \mu' U \times U^{(2)}, \end{aligned}$$

即

$$P_3 = \lambda P_1 + \mu P_2, \quad P_4 = \lambda' P_1 + \mu' P_2,$$

这就得到

$$(l_1, l_2; l_3, l_4) = (P_1, P_2; P_3, P_4) = \frac{\frac{\mu}{\lambda'}}{\frac{\mu'}{\lambda}} = \frac{\mu\lambda'}{\lambda\mu'}.$$

共点四线的交比的几何意义如下：

定理 10.7 设四直线 l_1, l_2, l_3, l_4 共点 O (O 为普通点), 用 $\angle(l_i, l_j)$ 表示 l_i 绕 O 转到 l_j 的有向角(逆时针方向为正)则有(见图 10-3)

$$(l_1, l_2; l_3, l_4) = \frac{\frac{\sin \angle(l_1, l_3)}{\sin \angle(l_3, l_2)}}{\frac{\sin \angle(l_1, l_4)}{\sin \angle(l_4, l_2)}},$$

即: 共点四线的交比是四直线间有向角的正弦的复比.

这个证明留作习题.

共线四点 P_1, P_2, P_3, P_4 的交比, 因 P_1, P_2, P_3, P_4 的排列不同有不同的值, 但它们之间有如下定理所述的关系.

定理 10.8 记 $\sigma = (P_1, P_2; P_3, P_4)$ 则

$$(1) \sigma = (P_1, P_2; P_3, P_4) = (P_2, P_1; P_4, P_3) = (P_3, P_4; P_1, P_2) = (P_4, P_3; P_2, P_1).$$

$$(2) \frac{1}{\sigma} = (P_2, P_1; P_3, P_4) = (P_1, P_2; P_4, P_3) = (P_3, P_4; P_2, P_1) = (P_4, P_3; P_1, P_2).$$

$$(3) 1 - \sigma = (P_1, P_3; P_2, P_4) = (P_3, P_1; P_4, P_2) = (P_2, P_4; P_1, P_3) = (P_4, P_2; P_3, P_1).$$

由以上三条性质还可推出：

$$(4) 1 - \frac{1}{\sigma} = (P_2, P_3; P_1, P_4) = (P_3, P_2; P_4, P_1) = (P_1, P_4; P_2, P_3) =$$

$$(P_4, P_1; P_3, P_2).$$

$$(5) \frac{1}{1-\sigma} = (P_3, P_1; P_2, P_4) = (P_1, P_3; P_4, P_2) = (P_2, P_4; P_3, P_1) = (P_4, P_2; P_1, P_3).$$

$$(6) \frac{\sigma}{\sigma-1} = (P_3, P_2; P_1, P_4) = (P_2, P_3; P_4, P_1) = (P_1, P_4; P_3, P_2) = (P_4, P_1; P_2, P_3).$$

注 如果这六个比值中有值相等, 则交比的值一定是 $-1, -1/2, 2, 1/2$.

证 下面仅证(3)中第一个等号, 即证

$$1 - (P_1, P_2; P_3, P_4) = (P_1, P_3; P_2, P_4).$$

因 P_1, P_2, P_3, P_4 共线, 设 $P_3 = \lambda P_1 + \mu P_2, P_4 = \lambda' P_1 + \mu' P_2$, 则

$$(P_1, P_2; P_3, P_4) = \frac{\frac{\mu}{\lambda}}{\frac{\mu'}{\lambda'}} = \frac{\mu\lambda'}{\lambda\mu'}.$$

又有 $P_2 = -\frac{\lambda}{\mu}P_1 + \frac{1}{\mu}P_3, P_4 = \left(\lambda' - \frac{\lambda\mu'}{\mu}\right)P_1 + \frac{\mu'}{\mu}P_3$, 则

$$(P_1, P_2; P_3, P_4) = \frac{\frac{1}{\mu}\left(\lambda' - \frac{\lambda\mu'}{\mu}\right)}{\left(-\frac{\lambda}{\mu}\right)\frac{\mu'}{\mu}} = 1 - \frac{\mu\lambda'}{\lambda\mu'} = 1 - (P_1, P_2; P_3, P_4).$$

其余性质可类似证得.

交比为 -1 的情形是一种重要的特殊情形, 我们有下列定义

定义 10.9 射影平面上四点 P_1, P_2, P_3, P_4 共线, 若满足 $(P_1, P_2; P_3, P_4) = -1$, 则称它们是调和点列. 以下几种说法都表示 P_1, P_2, P_3, P_4 是调和点列: P_1, P_2 与 P_3, P_4 调和共轭; P_3, P_4 调和分割 P_1, P_2 ; P_4 为 P_1, P_2, P_3 的第四调和点. 类似地, 共点四线 l_1, l_2, l_3, l_4 , 如果满足 $(l_1, l_2; l_3, l_4) = -1$, 则称 l_1, l_2, l_3, l_4 是调和线束.

例 $\triangle ABC$ 中角 A 的内外角平分线分别交对边于 D, E , 则 $(B, C; D, E) = -1$.

解 如图 10-4 所示, 由定理 10.6 和定理 10.7, 有

$$(B, C; D, E) = (l_1, l_3; l_2, l_4) = \frac{\frac{\sin \angle(l_1, l_2)}{\sin \angle(l_2, l_3)}}{\frac{\sin \angle(l_1, l_4)}{\sin \angle(l_4, l_3)}} = -1.$$

定理 10.9 (交比与射影坐标系选取无关) 设 A, B, C, D 是射影平面 $\bar{\pi}_0$

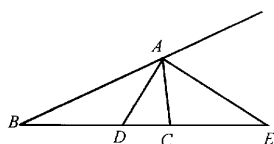


图 10-4

上共线四点, 设 A, B, C, D 在任一基底 $II[A_1, A_2, A_3; E]$ 中的射影坐标为: $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3), (d_1, d_2, d_3)$, 并且 $c_i = \lambda_1 a_i + \mu_1 b_i$, $d_i = \lambda_2 a_i + \mu_2 b_i, i = 1, 2, 3$, 则

$$(A, B; C, D) = \frac{\frac{\lambda_2}{\mu_2}}{\frac{\lambda_1}{\mu_1}}.$$

证 在 π_0 上取基底 $II[O_1, O_2, O; E]$, 则 A, B, C, D 在 II 中射影坐标 $(a'_1, a'_2, a'_3), (b'_1, b'_2, b'_3), (c'_1, c'_2, c'_3), (d'_1, d'_2, d'_3)$ 也就是它们的齐次坐标. 设

$$c'_i = \lambda'_1 a'_i + \mu'_1 b'_i, \quad d'_i = \lambda'_2 a'_i + \mu'_2 b'_i, \quad i = 1, 2, 3,$$

由交比定义有

$$(A, B; C, D) = \frac{\lambda'_2 \mu'_1}{\lambda'_1 \mu'_2}.$$

设基底 II 到 I 的射影坐标变换公式为

$$\rho \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = H \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix},$$

其中 H 为非奇异矩阵, 于是由 $c_i = \lambda_1 a_i + \mu_1 b_i$ 得

$$\rho_3^{-1} H \begin{pmatrix} c'_1 \\ c'_2 \\ c'_3 \end{pmatrix} = \lambda_1 \rho_1^{-1} H \begin{pmatrix} a'_1 \\ a'_2 \\ a'_3 \end{pmatrix} + \mu_1 \rho_2^{-1} H \begin{pmatrix} b'_1 \\ b'_2 \\ b'_3 \end{pmatrix},$$

所以 $\lambda'_1 = \lambda_1 \rho_3 \rho_1^{-1}, \mu'_1 = \mu_1 \rho_3 \rho_2^{-1}$. 同理可得

$$\lambda'_2 = \lambda_2 \rho_4 \rho_1^{-1}, \quad \mu'_2 = \mu_2 \rho_4 \rho_2^{-1}.$$

于是可以算出

$$\frac{\lambda'_2 \mu'_1}{\lambda'_1 \mu'_2} = \frac{\lambda_2 \mu_1}{\mu_2 \lambda_1}.$$

10.4 射影映射和射影变换

定义 10.10 (射影映射的代数定义) 设 $\overline{\pi}_0$ 是一射影平面, $\overline{\pi}_1$ 为另一射影平面, 在 $\overline{\pi}_0, \overline{\pi}_1$ 上分别取射影标架 $I = [A_1, A_2, A_3; E], II = [B_1, B_2, B_3; F]$. 对 $\overline{\pi}_0$ 中任意一点 P , P 在 I 中的射影坐标为 (x_1, x_2, x_3) , 定义 $\tau(P)$ 在 II 中的射影坐标 (x'_1, x'_2, x'_3) 为

$$\rho \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

其中 ρ 是非零实数, $A = (a_{ij})$ 为非奇异矩阵. $\tau: P \rightarrow \tau(P)$ 称为 $\overline{\pi}_0$ 到 $\overline{\pi}_1$ 的射影映射. 射影平面到自身的射影映射称为射影变换.

注 与平面仿射映射的定义类似, 我们有下面射影映射的几何定义, 可以证明它与定义 10.10 等价 (这里不证).

定义 10.10 (射影映射的几何定义) 设 $\overline{\pi}_0$ 是一射影平面, $\overline{\pi}_1$ 为另一射影平面, 设有双射 $\tau: \overline{\pi}_0 \rightarrow \overline{\pi}_1$. 如果 τ 把共线的三点变成共线的三点, 则称 τ 是一个射影映射.

由定义可见, 中心投影及其有限次乘积都是射影映射.

由定义可得射影映射的下述性质:

性质 1 射影映射把共线的三点变成共线的三点, 把不共线的三点变成不共线的三点.

性质 2 射影映射把直线变成直线.

性质 3 射影映射把一般位置的四个点 (即四个点中没有三点共线) 变成一般位置的四个点.

性质 4 射影映射保持共线四点的交比不变.

定理 10.10 (射影映射基本定理) 设 A_1, A_2, A_3, E 是射影平面 π_0 上一般位置的四个点, A'_1, A'_2, A'_3, E' 是射影平面 π_1 上一般位置的四个点, 则存在 π_0 到 π_1 的唯一的射影映射 τ 把 A_1, A_2, A_3, E 分别映射成 A'_1, A'_2, A'_3, E' .

证 存在性: 在 π_0 上取坐标系 $I\{A_1, A_2, A_3; E\}$, 在 π_1 上取坐标系 $I'\{A'_1, A'_2, A'_3; E'\}$. 规定 π_0 到 π_1 的映射 τ 如下: π_0 上任取一点 M , 它在 I 中的射影坐标为 (x_1, x_2, x_3) , M 在映射 τ 下的象 M' 定义为 I' 中的射影坐标

为 (x_1, x_2, x_3) 的点, τ 显然是双射. 设 P, Q, R 是 π_0 上共线三点, 则它们的射影坐标组成的三阶行列式等于零. 据 τ 的定义知, $\tau(P), \tau(Q), \tau(R)$ 在 I' 中的射影坐标组成的三阶行列式也等于零. 因此 $\tau(P), \tau(Q), \tau(R)$ 共线. 这就证明了 τ 是 π_0 到 π_1 的一个射影映射. 唯一性证明略.

射影变换 射影平面到自身的射影映射称为射影变换. 射影变换把点变到点, 直线变到直线.

容易验证, 射影变换有下列性质:

(1) 两个射影变换的乘积是一个射影变换.

(2) 恒等变换是射影变换.

(3) 射影变换的逆变换也是一个射影变换.

从而射影平面上的所有射影变换形成一个群, 称它为射影变换群.

我们有下面的重要论断:

(1) (基本定理) 设 P_1, P_2, P_3, P_4 是平面上的任意四个点, 其中任意三个点都不共线, P'_1, P'_2, P'_3, P'_4 也是平面上的任意四个点, 其中任何三个点都不共线, 则存在唯一一个射影变换它把 P_1, P_2, P_3, P_4 分别变为 P'_1, P'_2, P'_3, P'_4 .

(2) 直线上四点的交比在射影变换下是不变的.

(3) 射影平面上对一变换, 若把直线变成直线, 则它一定是射影变换.

二阶曲线的射影分类 普通平面上二阶曲线 C 的方程为

$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + 2a_{13}x + 2a_{23}y + a_{33} = 0.$$

作为扩大平面 (射影平面) 上的曲线, 它的方程应由齐次坐标表达, 即

$$a_{11}\left(\frac{x_1}{x_3}\right)^2 + a_{22}\left(\frac{x_2}{x_3}\right)^2 + 2a_{12}\frac{x_1x_2}{x_3^2} + 2a_{13}\frac{x_1}{x_3} + 2a_{23}\frac{x_2}{x_3} + a_{33} = 0.$$

如果允许 $x_3 = 0$, 即添加曲线 C 上的无穷远点, C 的方程可写为

$$\begin{aligned} a_{11}x_1^2 + a_{22}x_2^2 + 2a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \\ = \sum_{i,j=1}^3 a_{ij}x_ix_j = 0. \end{aligned} \quad (10-11)$$

记 $X = (x_1, x_2, x_3)$, $A = (a_{ij})$, 则 (10-11) 的矩阵形式为 $XAX^T = 0$.

命题 给定二阶曲线 $C: XAX^T = 0$ 以及不在 C 上的点 $P(Y)$, 过 $P(Y)$ 的动直线 l 交二阶曲线 C 于两点 $G(X_1), H(X_2)$, 则动直线 l 上 $P(Y)$ 关于 $G(X_1), H(X_2)$ 的第四调和点 $Q(Z)$ 的轨迹为一条直线 L_P , 其方程为 $YAZ^T = 0$. (见图 10-5)

证 设过点 $P(Y)$ 的直线 l 与二阶曲线 C 的交点坐标为 X , 则有 $X = Y + \lambda Z$. 因它在 C 上, 代入 C 的方程得 $XAX^T = 0$ 得

$$YAY^T + 2\lambda YAZ^T + \lambda^2 ZAZ^T = 0. \quad (10-12)$$

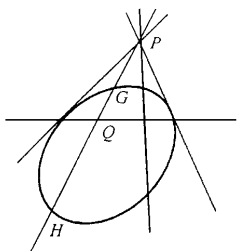


图 10-5

上述二次方程的两个解 λ_1, λ_2 确定 l 与 C 的两个交点 $G(X_1), H(X_2)$, 易见

$$X_1 = Y + \lambda_1 Z, \quad X_2 = Y + \lambda_2 Z.$$

若 P, Q 与 G, H 调和共轭, 则 $(P, Q; G, H) = -1$, 即 $\frac{\lambda_2}{\lambda_1} = -1$ 或 $\lambda_1 + \lambda_2 = 0$. 由根与系数的关系知方程 (10-12) 的一次项系数等于零, 即 $YAZ^T = 0$. 这表明 $Q(Z)$ 的轨迹是一条直线.

定义 10.11 设二阶曲线 C 的方程为 $XAX^T = 0$, 称上述命题中的直线 $L_P: YAX^T = 0$ 为点 $P(Y)$ 关于二阶曲线 C 的极线, 而称点 $P(Y)$ 为直线 L_P 关于二阶曲线 C 的极点.

注 若直线 L 与二阶曲线 C 有两个重合的交点, 或者整个在 C 上, 则称 L 是 C 的切线, 它们的交点成为切点. 若过 $P(Y)$ 可作曲线 C 的两条切线, 则切点必在极线上. 即 $P(Y)$ 关于 C 的极线是二切点的连线. 若 $P(Y)$ 在曲线 C 上, 则 $P(Y)$ 关于 C 的极线就是 C 在 $P(Y)$ 处的切线.

极线有下列重要性质:

定理 10.11 给定二阶曲线 $C: XAX^T = 0$, 设点 P 的极线过点 Q , 则点 Q 的极线也过点 P ; 点 P 的极线通过点 P 的充分必要条件为 P 在二阶曲线 C 上.

证 设 P, Q 的齐次坐标分别为 Y_1, Y_2 , 则 P 的极线为 $Y_1AX^T = 0$, Q 的极线为 $Y_2AX^T = 0$. 由假设 $Y_1AY_2^T = 0$, 因 A 为对称矩阵, 所以 $Y_2AY_1^T = (Y_1AY_2^T)^T = Y_1AY_2^T = 0$.

定义 10.12 设 S 是射影平面 π_0 上的一条二阶曲线, π_0 上的一个三角形 ABC 称为关于 S 的自配极三角形, 如果它的每一条边是其对顶点的极线.

定理 10.12 对于射影平面 π_0 上的任意一条二阶曲线 S , 都存在关于 S 的自配极三角形.

证 在二阶曲线 S 外任取一点 A , 设 A 的极线是 L_A , 则 A 不在 L_A 上.

如果 L_A 不是整个在 S 上, 此时在 L_A 上任取一点 B , 使 B 不在 S 上. 设 B 的极线是 L_B . 因为 B 在 L_A 上, 所以 A 在 L_B 上. 因为 $B \notin S$, B 不在 L_B 上, 从而 B 不是 L_A 与 L_B 的交点.

设 L_A 与 L_B 的交点为 C , 于是 A, B, C 可以构成一个三角形. 因为 C 在 L_A 上, 所以 A 在 C 的极线上. 同理, B 在 C 的极线上. 因此, A, B 的连线 L 是 C 的极线. 于是三角形 $\triangle ABC$ 是自配极三角形.

如果 L_A 整个在 S 上. 此时 S 为一对重合直线. 在 L_A 上任取两点 B, C , 则三角形 $\triangle ABC$ 是自配极三角形.

二阶曲线的射影分类 如图 10-6, 对于二阶曲线 C , 选取一自配极三角形为坐标三角形, 设二阶曲线 C 在坐标三角形 $O_1O_2O_3$ 中的矩阵为 (a_{ij}) . 则 $O_3(0, 0, 1)$ 的极线为 O_1O_2 $(0 \ 0 \ 1) \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} X^T = 0$, 即 $a_{13}x_1 + a_{23}x_2 + a_{33}x_3 = 0$, 但 O_1O_2 的方程是 $x_3 = 0$, 所以 $a_{13} = a_{23} = 0$. 用同样的推导可得 $a_{12} = 0$. 故二阶曲线 C 的方程可简化为

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 = 0. \quad (10-13)$$

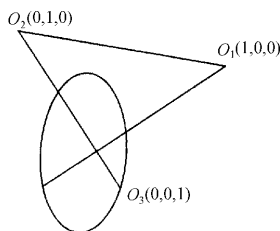


图 10-6

适当选取自配极三角形顶点的次序 (10-13) 可分为以下几种情形:

情形 1 $\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 = 0, \lambda_1 \lambda_2 \lambda_3 \neq 0$.

情形 2 $\lambda_1 x_1^2 + \lambda_2 x_2^2 = 0, \lambda_1 \lambda_2 \neq 0$.

情形 3 $\lambda_1 x_1^2 = 0, \lambda_1 \neq 0$.

对情形 1 再取坐标变换

$$\rho x_i = \frac{1}{\sqrt{|\lambda_i|}} \tilde{x}_i, \quad i = 1, 2, 3,$$

得到情形 1 的标准方程为

1° $\tilde{x}_1^2 + \tilde{x}_2^2 + \tilde{x}_3^2 = 0$ (无轨迹).

2° $\tilde{x}_1^2 + \tilde{x}_2^2 - \tilde{x}_3^2 = 0$ (椭圆, 双曲线, 抛物线).

同理可知, 经过适当的射影坐标变换, 情形 2 和情形 3 的标准方程为:

3° $\tilde{x}_1^2 + \tilde{x}_2^2 = 0$ (一点).

4° $\tilde{x}_1^2 - \tilde{x}_2^2 = 0$ (相交的两直线).

5° $\tilde{x}_1^2 = 0$ (重合的两直线).

注 在射影平面上,椭圆、双曲线、抛物线在同一个射影类里,这是不奇怪的.例如,在中心投影下,椭圆、双曲线、抛物线可以互变,如图 10-7 所示.

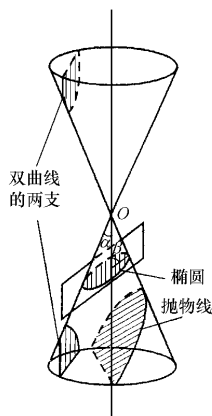


图 10-7

习 题

1. 证明 如果平面的一个仿射变换 τ 有两个不动点 A, B , 则直线 AB 上每个点在这个仿射变换下都不变.

2. 求把三条直线 $x = 0, x - y = 0, y = 1$ 依次变成 $3x - 2y - 3 = 0, x - 1 = 0, 4x - y - 9 = 0$ 的仿射变换.

3. 证明仿射变换把任意一个仿射标架 I 变成一个仿射标架 II , 并且任意一点 P 的 I 坐标等于它的像 $P' = \tau(P)$ 的 II 坐标.

4. 如果一条直线与它在仿射变换 τ 下的像重合, 则称这条直线为 τ 的不变直线. 求下述仿射变换的不变直线:

$$\begin{cases} x' = 7x - y + 1 \\ y' = 4x + 2y + 4. \end{cases}$$

5. 设仿射变换 τ 在仿射坐标系 I 中的公式为

$$\begin{cases} x' = x + 2y \\ y' = 4x + 3y. \end{cases}$$

(1) 求 τ 的不变直线;

(2) 以 τ 的两条不变直线为新坐标轴, 求 τ 在新的仿射坐标系中的公式.

6. 设 $\triangle ABC$ 的 AB, BC 边长分别为 10 cm, 6 cm, $\angle ABC = 30^\circ$. 设仿射变换 $\tau: \begin{cases} x' = 2x + y + 1 \\ y' = x - 4y + 2 \end{cases}$ 把 A, B, C 分别变成 A', B', C' , 求 $\triangle A'B'C'$ 的面积.

7*. 证明仿射变换的(代数)定义 10.2 与(几何)定义 10.2' 等价.

8. 证明平面上任意两个平行四边形都仿射等价.

9. 任给一个三角形 $\triangle ABC$, 设 D, E, F 分别是 AB, BC, CA 的中点. 证明: 存在一个椭圆, 它与 $\triangle ABC$ 的三边分别相切于 D, E, F ; 并且求出这个椭圆的面积与 $\triangle ABC$ 的面积之比.

10. 在扩大的欧氏平面 π_0 上, 两点 A, B 的齐次坐标分别为 $(3, -1, 2), (2, 0, 1)$, 求

(1) 直线 AB 在齐次坐标中的普通方程和参数方程;

(2) 直线 AB 上的无穷远点的齐次坐标和它所对应的参数值.

11. 在扩大的欧氏平面 π_0 上, 给了下列欧氏直线在仿射坐标中的方程, 求由它们所确定的射影直线在齐次坐标中的方程, 并且求出它上面的无穷远点.

(1) $x + 2y - 1 = 0$; (2) $3x - 2y = 0$.

12. 证明: 射影平面上若三点 A, B, C 不共线, 则三线 AB, BC, CA 不共点.

13. 设 A, B, C, D 为射影平面 π_0 上的共线四点, 其中 A, B, C 各不相同, 并且 D 与 A 不同. 在此直线上取两点 P, Q , 它们的齐次坐标分别为 $(p_1, p_2, p_3), (q_1, q_2, q_3)$, A, B, C, D 的齐次坐标分别为 $(a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3), (d_1, d_2, d_3)$, 并设

$$a_i = \lambda_1 p_i + \mu_1 q_i, \quad b_i = \lambda_2 p_i + \mu_2 q_i,$$

$$c_i = \lambda_3 p_i + \mu_3 q_i, \quad d_i = \lambda_3 p_i + \mu_3 q_i \quad (i = 1, 2, 3).$$

证明

$$(A, B; C, D) = \frac{\begin{vmatrix} \lambda_1 & \lambda_3 \\ \mu_1 & \mu_3 \end{vmatrix}}{\begin{vmatrix} \lambda_1 & \lambda_4 \\ \mu_1 & \mu_4 \end{vmatrix}} \cdot \frac{\begin{vmatrix} \lambda_2 & \lambda_4 \\ \mu_2 & \mu_4 \end{vmatrix}}{\begin{vmatrix} \lambda_2 & \lambda_3 \\ \mu_2 & \mu_3 \end{vmatrix}}.$$

14. 在射影平面 π_0 上, 设共线三点 A, B, C 的齐次坐标分别为 $(1, 2, 5), (1, 0, 3), (-1, 2, -1)$, 在直线 AB 上求一点 D 使得交比 $(A, B; C, D) = 5$.

15. 证明定理 10.7.

16. 设 A, B, C, D, E 是共线五点, 并且两两不同, 证明 $(A, B; C, D) \cdot$

$$(A, B; D, E) = (A, B; C, E).$$

17. 用交比 $(A, B; C, D)$ 表达 A, B, C, D 这些点按任何其他顺序所取的交比.

18. 在射影平面上, 求把点 $A(1, 0, 1), B(2, 1, 1), C(3, -1, 0), D(3, 5, 2)$ 分别变成点 $A'(-1, 0, 3), B'(1, 1, 3), C'(2, 3, 8), D'(2, 1, -2)$ 的射影坐标变换公式.

19. 在扩大的欧氏平面上, 求把直线 $x_1 = 0, x_2 = 0, x_3 = 0$ 分别变为直线 $a_i x_1 + b_i x_2 + c_i x_3 = 0 (i = 1, 2, 3)$ 的射影变换公式的一般形式.

20. 给定二阶曲线 C , 如图 10-8, 选取三角形 $\triangle O_1 O_2 O_3$ (其中 O_1, O_3 是切点) 为坐标三角形. 证明在上述坐标三角形中二阶曲线 C 的方程 (10.11) 简化为

$$a_{22}x_2^2 + 2a_{13}x_1x_3 = 0.$$

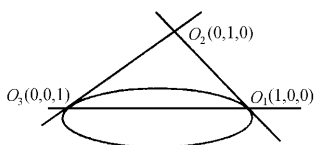


图 10-8

21. 正常成人的头长, 体长, 腿长与全身长之比分别为 14%, 35%, 51%. 某甲天生腿短, 腿长只占全身长的 45%. 在照相中他想弥补这一缺陷, 通过取一个合适的角度, 使照片上的这一比例变为 50%. 利用交比不变性, 问这时他的头长所占比例变成多少?

22. 设 A, B, C 和 A', B', C' 分别是直线 l_1, l_2 上任意三点, 试通过直线中心透视的乘积构造一射影对应 $\tau: l_1 \rightarrow l_2$, 使

$$\tau(A) = A', \quad \tau(B) = B', \quad \tau(C) = C'.$$

23. 设射影平面 π_0 上取了两个基底 $I = [A_1, A_2, A_3; E]$ 和 $II = [B_1, B_2, B_3; F]$. 已知点 $B_1, B_2, B_3; F$ 在基底 I 中的射影坐标分别为 $(1, -1, 2), (2, 0, 1), (-1, 2, 4)$ 和 $(1, -1, 0)$. 求基底 I 到基底 II 的点的射影坐标变换公式.

24. 在射影平面上已知射影变换(点变换) σ 的公式为

$$\rho \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

求 σ 把直线 $U = (u_1 \ u_2 \ u_3)$ 变为直线 $U' = (u'_1 \ u'_2 \ u'_3)$ 的公式.

25. 在射影平面上, 射影变换 σ 的公式为

$$\lambda \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

求 σ 的不动点和不变直线.

部分习题答案

2. $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ -8 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

4. $x - y - \frac{3}{2} = 0$ 或 $4x - y = 0.$

5. (1) $x + y = 0$, $y - 2x = 0$; (2) $\begin{cases} x' = -x \\ y' = 5y \end{cases}.$

6. 135. 9. $\frac{\sqrt{3}\pi}{9}.$

10. (1) $x - y - z = 0$, $X = (3, -1, 2) + \lambda(2, 0, 1), -2.$

11. (1) $x_1 + 2x_2 - x_3 = 0$ 不 $(-2, 1, 0)$;

(2) $3x_1 - 2x_2 = 0$ 不 $(2, 3, 0).$

14. $D = k(3, 10, 19).$

19. $\rho \begin{bmatrix} u'_1 \\ u'_2 \\ u'_3 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}.$

21. 11.75%.

23. $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{\rho}{15} \begin{bmatrix} 7 & 4 & 4 \\ -7 & 0 & -8 \\ 14 & 2 & -16 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix}.$

24. $U' = \rho(A^T)^{-1}U.$

25. 不动点: $\rho(1, 1, 0)$, $\rho(1, 0, 1)$, $\rho(1, 1, 2),$

不变直线: $\rho(-1, 1, 0)$, $\rho(-2, 0, 1)$, $\rho(1, -1, -1)$, $\rho(-1, -1, 1).$

第 11 章 非欧几何学简介

以欧几里德的“几何原本”(共 13 卷,出现于公元前 3 世纪)为代表的欧氏几何是人类理性思维的一大成就.它第一次系统地体现了“公理化”思维的模式.即一种从公认的定义和无须证明的公理和公设出发,经过合乎逻辑的推理,最后得出结论的模式.“原本”中用了 23 个定义,5 条公理(即对一切学科都通用的无须证明的命题,例如“整体大于部分”)和 5 条公设(对几何学无须证明的命题,例如“所有直角都相等”),由此推出了一系列的平面和空间图形的性质.出于理性和对简洁美的追求,后人对其中第 5 公设(平行公设)的公设地位有所怀疑,认为它可能是其它 4 条公设的推论.经过两千多年的探索,到了 19 世纪才得到解决.除了证明它确实独立于其它 4 条公设之外,还创立了两门新的几何学——椭圆几何(满足公设 1-4,但过直线外一点没有与之平行的直线)和双曲几何(满足公设 1-4,但过直线外一点有无穷多与之平行的直线),并为物理学的发展提供了有力的工具.在这一章中,我们将介绍这两种几何的简单模型.

11.1 球面几何

空间中和点 O 的距离等于定值 R 的所有点构成一个曲面,叫做以 O 点为球心,以 R 为半径的球面.例如:我们所居住的大地,其局部地貌虽然是丘陵起伏,山川纵横,但是其全局的形状却十分接近于一个球面,这也是为什么我们现在把它叫做“地球”的缘故.自工业革命以来,远洋航海日益发达,球面几何就成为航海、天文的基本工具.

将球面几何和大家熟悉的平面几何相比较,我们可看出以下几点:(1)在相对于半径来说,很小的一小片球面,看起来就几乎是一个平面,这也是早期历史上各民族将“大地”的整体形状误认为是“平面”的原因.(2)过球心的平面和球面的交线叫该球面的一个“大圆”.它们在球面几何里所扮演的角色相当于平面几何中的直线.例如一个小于半圆的大圆弧就是其两端点之间在球面上的所有道路之中唯一的最短道路.这表明,它们具有和平面上直线段相同的特征.(3)平面三角形的研究是平面几何学的核心问题,同样,球面三角

形亦即由连接三个顶点的大圆圆弧所构成的图形)的研究也是球面几何学的核心问题.

测地线 大家知道,连接平面上(任意)两点的最短曲线是直线.对一般空间,连接空间中任意两点的最短曲线是什么?这就是下面将要介绍的测地线的概念.

定义 11.1 连接空间 X 中任意两点的最短曲线叫 X 中极小测地线.

例 1 给定球面上两点 p 与 q ,紧贴球面上拉一根连接 p 与 q 的细绳,这条拉紧的细绳就是球面上连接两点 p 与 q 的测地线.

例 2 把一张纸卷成圆柱面,用一根细绳连接柱面上两点 p, q ,拉紧细绳就得到一条连接 p, q 的测地线.视细绳绕柱面的次数不同,可得到无穷多条连接 p, q 的测地线.但仅仅一条(或两条)是连接 p, q 的极小测地线.

定义 11.2 空间 X 上测地三角形定义为用测地线连接 X 上三点 A, B, C 构成的图形.

定义 11.3 球面 S 上的大圆是指一个过球心的平面在球面上的截线.球面上不是大圆的圆叫小圆(见图 11-1).

设 A, B 为球面上的两个点, C 为球面的中心,则弧 AB 的长度 $= R(\angle ACB)$,其中 R 为球面的半径, $\angle ACB$ 是弧度.

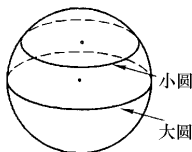


图 11-1

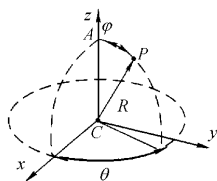


图 11-2

用球面坐标计算球面上曲线的长度比较方便.以 C 为原点建立三维欧氏空间中直角坐标系,正向 z 轴与球面交于 A 点,则球面上点 P 的球坐标为 (R, θ, φ) ,其中 R 为球面半径, $\varphi = \angle PCA$, θ 为 x 轴正向与 \vec{CP} 在 xOy 平面的投影向量之夹角(见图 11-2).用初等几何学可验证下面球面坐标与直角坐标之间的关系:

$$\begin{cases} x = R \sin \varphi \cos \theta \\ y = R \sin \varphi \sin \theta \\ z = R \cos \varphi \end{cases}$$

$$(0 \leq \varphi \leq \pi, 0 \leq \theta < 2\pi)$$

定理 11.1 球面上两点之间的最短道路是大圆弧. 即球面上的测地线是大圆弧.

证 设 $\sigma[a, b] \rightarrow S$ 为球面上任意一条连 A, B 的参数曲线, 即满足 $\sigma(a) = A, \sigma(b) = B$. 在直角坐标下记为 $\sigma(t) = (x(t), y(t), z(t))$, σ 的长度为

$$l(\sigma) = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} dt. \quad (11-1)$$

用球坐标

$$\begin{cases} x(t) = R \sin \varphi(t) \cos \theta(t) \\ y(t) = R \sin \varphi(t) \sin \theta(t) \\ z(t) = R \cos \varphi(t) \end{cases} \quad (11-2)$$

将下面求导后的式子

$$\begin{cases} x'(t) = R(\cos \varphi(t) \cos \theta(t) \varphi'(t) - \sin \varphi(t) \sin \theta(t) \theta'(t)) \\ y'(t) = R(\cos \varphi(t) \sin \theta(t) \varphi'(t) + \sin \varphi(t) \cos \theta(t) \theta'(t)) \\ z'(t) = -R \sin \varphi(t) \varphi'(t) \end{cases}$$

代入方程 (11-1) 得到

$$\begin{aligned} l(\sigma) &= \int_a^b R \sqrt{\varphi'(t)^2 + \sin^2 \varphi(t) \theta'(t)^2} dt \\ &\geq \left| \int_a^b R \varphi'(t) dt \right| = R(\varphi(b) - \varphi(a)) \\ &= R(\angle BCA) \\ &= \text{弧 } AB \text{ 的长度}, \end{aligned} \quad (11-3)$$

等号成立当且仅当 $\theta'(t) = 0$ 或 $\sin^2 \varphi(t) = 0, t \in [a, b]$, 即 σ 为弧 AB . 定理 11.1 证毕.

如果把球面上的大圆弧定义为球面上的“直线”, 则有下列的命题:

命题 11.1 球面上任意两个不在同一直径上的点决定唯一一条“直线”, 任意两条“直线”必交于在同一直径上的两点. 因而球面上没有“平行线”. (见图 11-3)

定理 11.2 在球面三角形 $\triangle ABC$ 中, $\angle A + \angle B + \angle C = \pi + \frac{V}{R^2}$, 其中 R 为球面半径, V 为 $\triangle ABC$ 的面积.

证 将 AB, AC 这两个大圆圆弧延长, 它们相交于 A 的对顶点 A' , 这两条半圆圆弧之间的区域叫做一个“梭形”, 它的面积是全球面面积的 $\frac{\angle A}{2\pi}$ 倍,

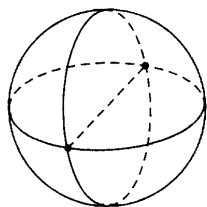


图 11-3

即 $\frac{\angle A}{2\pi} \cdot 4\pi R^2 = 2R^2 \angle A$ 其中 $\angle A$ 表示 $\angle A$ 的弧度. 记 A', B', C' 分别为 A, B, C 的对径点, 则有(见图 11-4)

$$\triangle ABC \text{ 的面积} + \triangle A'BC \text{ 的面积} = 2R^2 \angle A, \quad (11-4)$$

$$\triangle ABC \text{ 的面积} + \triangle AB'C \text{ 的面积} = 2R^2 \angle B, \quad (11-5)$$

$$\triangle ABC \text{ 的面积} + \triangle ABC' \text{ 的面积} = 2R^2 \angle C, \quad (11-6)$$

$$\triangle ABC' \text{ 的面积} = \triangle A'B'C \text{ 的面积 (因为两者对称)}. \quad (11-7)$$

又由 $\triangle ABC \text{ 的面积} + \triangle A'BC \text{ 的面积} + \triangle AB'C \text{ 的面积}$
 $+ \triangle A'B'C \text{ 的面积} = \text{半球的面积} = 2\pi R^2, \quad (11-8)$

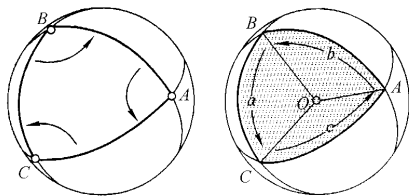


图 11-4

(11-4)+(11-5)+(11-6)后, 即得

$$3\triangle ABC \text{ 的面积} + \triangle A'BC \text{ 的面积} + \triangle AB'C \text{ 的面积} \\ + \triangle A'B'C \text{ 的面积} = 2R^2(\angle A + \angle B + \angle C). \quad (11-9)$$

综合(11-7)(11-8)和(11-9)就有

$$2\triangle ABC \text{ 的面积} + 2\pi R^2 = 2R^2(\angle A + \angle B + \angle C),$$

即 $\angle A + \angle B + \angle C = \pi + \frac{V}{R^2}$.

由定理立即得推论 11.1:

推论 11.1 在任意球面三角形 $\triangle ABC$ 中,

$$\angle A + \angle B + \angle C > \pi$$

如果我们在平面上讨论问题, 可以通过所谓的球极投影把球面映成平面.

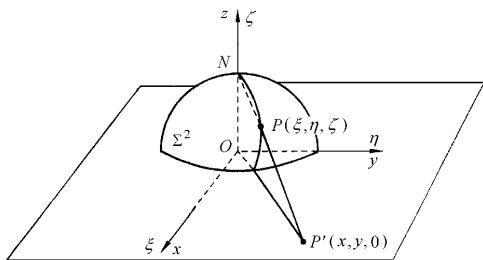


图 11-5

如图 11-5, Σ^2 是以原点 O 为球心的单位半径球面, π 为 xOy 坐标平面, 我们以 (ξ, η, ζ) 表示 Σ^2 上的动点 P 的坐标. N 为坐标为 $(0, 0, 1)$ 的点. 令 P' 为直线 NP 和 π 的交点, 以 $(x, y, 0)$ 表示 P' 的坐标. 由于三点 $N(0, 0, 1)$, $P(\xi, \eta, \zeta)$, $P'(x, y, 0)$ 共线, 而且满足 $\xi^2 + \eta^2 + \zeta^2 = 1$, 不难解出 (ξ, η, ζ) 和 $(x, y, 0)$ 之间的关系

$$x = \frac{\xi}{1 - \zeta}, y = \frac{\eta}{1 - \zeta}, \quad (11-10)$$

$$\xi = \frac{2x}{1 + x^2 + y^2}, \eta = \frac{2y}{1 + x^2 + y^2}, \zeta = \frac{-1 + x^2 + y^2}{1 + x^2 + y^2}. \quad (11-11)$$

定义 11.4 球极投影就是上述把球面上任一点 $P(\xi, \eta, \zeta)$ 映到平面上 $P'(x, y, 0)$ 点的映射, 并定义 N 点的像点为 ∞ , 即

$$\tau: \Sigma^2 \rightarrow \pi \cup \{\infty\}, \tau(N) = \infty$$

命题 11.2 球极投影 $\tau: \Sigma^2 \rightarrow \pi^* = \pi \cup \{\infty\}$ 具有下列重要性质

- (1) Γ 是 Σ^2 上的圆, 且 $N \in \Gamma$ 当且仅当 $\tau(\Gamma)$ 为 π^* 中直线.
- (2) Γ 是 Σ^2 上的圆, 且 $N \notin \Gamma$ 当且仅当 $\tau(\Gamma)$ 为 π^* 中圆.
- (3) τ 保持角度不变.

证 (3) 的证明较长, 这里省略. 只证 (1) (2).

设 Γ 是 Σ^2 上任意一圆, 则 Γ 中的动点 $P(\xi, \eta, \zeta)$ 除了满足球面方程 $\xi^2 + \eta^2 + \zeta^2 = 1$ 外, 还满足线性方程

$$A\xi + B\eta + C\zeta + D = 0, \quad (11-12)$$

用变换 (11-11), 即得 $\tau(\Gamma)$ 的方程为

$$A \frac{2x}{1+x^2+y^2} + B \frac{2y}{1+x^2+y^2} + C \frac{-1+x^2+y^2}{1+x^2+y^2} + D = 0, \quad (11-13)$$

亦即

$$(C+D)(x^2+y^2) + 2Ax + 2By + (D-C) = 0. \quad (11-14)$$

当 $N \notin \Gamma$, 即 $C+D \neq 0$ 时 (11-14) 说明 $\tau(\Gamma)$ 为 π^* 中的一个圆; 当 $N \in \Gamma$, 即 $C+D = 0$ 时 (11-14) 说明 $\tau(\Gamma)$ 为 π^* 中的直线. 反之亦然.

平面极投影的应用 万能变换

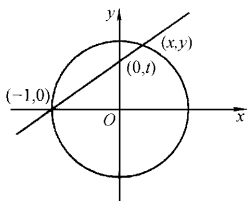


图 11-6

图 11-6 给出 \mathbf{R}^2 上以 $(-1, 0)$ 为基点, 把单位圆 $x^2 + y^2 = 1$ 映到 y 轴的投影, 叫做极投影. 设 (x, y) 是单位圆上的一点, $(0, t)$ 为它与基点连线在 y 轴的交点, 则三点 $(-1, 0)$, $(0, t)$, (x, y) 共线, 可得 $(0, t)$ 与 (x, y) 的关系

$$t = \frac{y}{x+1}. \quad (11-15)$$

注意到点 (x, y) 满足 $x^2 + y^2 = 1$, 结合 (11-15), 当 $(x, y) \neq (-1, 0)$ 时, 解出

$$\begin{cases} x = \frac{1-t^2}{1+t^2} \\ y = \frac{2t}{1+t^2} \end{cases} \quad (11-16)$$

考虑积分

$$\int R(\cos \theta, \sin \theta) d\theta,$$

其中 $R(x, y)$ 为 x, y 的有理函数 (两个多项式的商). 由 (11-16) 可设

$$x = \cos \theta = \frac{1-t^2}{1+t^2}, \quad y = \sin \theta = \frac{2t}{1+t^2},$$

这等价于

$$t = \tan \frac{\theta}{2}.$$

对上式微分得 $d\theta = \frac{2}{1+t^2}dt$, 从而得到

$$\int R(\cos \theta, \sin \theta) d\theta = \int R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \frac{2}{1+t^2} dt,$$

被积函数变为 t 的有理函数, 故可积. 上面推证可以看成是不定积分中“万能变换”的来源.

11.2 双曲几何的庞加莱模型

设 C 是欧氏平面上的单位圆, S 是它的内部. S 叫做绝对空间, 它是以下我们讨论的范围. 定义 S 中的“点”为欧氏平面的点; “直线”(双曲线)为过圆心的直径或与 C 正交的圆弧. C 上的点称为无穷远点. 而两条直线称为互相平行, 如果它们交于无穷远点.

我们有以下的命题:

命题 11.4 S 中任意两点必位于唯一的直线上.

这个命题的证明要用到以下一个简单的事实(我们不加证明): “欧氏平面中, 与单位圆正交的圆方程必具以下形式: $x^2 + y^2 + ax + by + 1 = 0$.” 用它来证明命题: 设两点的坐标分别为 (x_1, y_1) (x_2, y_2), 它们能否共线, 就看下列方程组是否有解:

$$\begin{cases} x_1 a + y_1 b = -x_1^2 - y_1^2 - 1 \\ x_2 a + y_2 b = -x_2^2 - y_2^2 - 1. \end{cases}$$

它唯一可解的条件是 $x_1 y_2 - x_2 y_1 \neq 0$, 而这正是这两点与 $(0, 0)$ 不共线的条件. 它说明: 这两点或者位于同一直径上, 或者位于某一直线. 不论哪种情况, 过这两点的直线都是唯一的.

我们定义 S 中任意两点 A, B 的距离(“双曲距离”) $d(A, B)$ 如下:

$$d(A, B) = \begin{cases} 0, & A = B, \\ \ln \frac{|AM| \cdot |BN|}{|AN| \cdot |BM|}, & A \neq B, \end{cases}$$

其中 $|AB|$ 表示欧氏线段 AB 的长度; M, N 表示 A, B 所在直线与 C 的两个交点. 由定义, M, N 与 A, B 的次序无关. 这个距离函数满足对称性及三角不等

式(这里不证). 如果 A, B 中有一个在 C 上, 易见其距离为 ∞ .

设两条双曲线交于 B 点, 定义它们的夹角就是两条曲线在这一点切线的夹角.

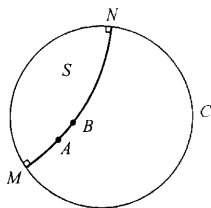


图 11-7

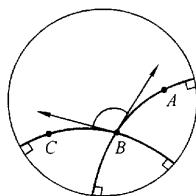


图 11-8

以上就构成了 S 中的一种几何. 在这个模型(称为庞加莱平面双曲几何模型)中, 可以证明欧氏几何的公设 1-4 都成立, 而第 5 公设明显不对(过线外一点有无穷多条直线与此线平行). 下面我们来证明 S 中任意三角形的内角和小于 π . 为了方便, 取三角形的两个边都是过 A (即原点 O) 的直线.

命题 11.5 O 为单位圆, O' 是任意与它正交的圆. 过 O 的直线交圆 O' 于两点 P, P' , 则

$$OP \cdot OP' = 1.$$

证 用初等几何.

命题 11.6 设在 S 中, 点 A 与圆心的欧氏距离为 r , 双曲距离为 r' , 则有以下关系:

$$\operatorname{sh} r' = \frac{2r}{1-r^2}, \quad \operatorname{ch} r' = \frac{1+r^2}{1-r^2}, \quad \operatorname{th} r' = \frac{2r}{1+r^2}.$$

证 三个等式都由下面式子推出:

$$r' = \ln \frac{1+r}{1-r}.$$

命题 11.7 图 11-9 中二圆正交, 直角双曲三角形 ABC 的两个双曲边长满足以下关系:

$$BB'' = \frac{2}{\operatorname{sh} c'}, \quad CC'' = \frac{2}{\operatorname{sh} b'}.$$

证 利用命题 11.5, $BB'' + AB = \frac{1}{AB}$,

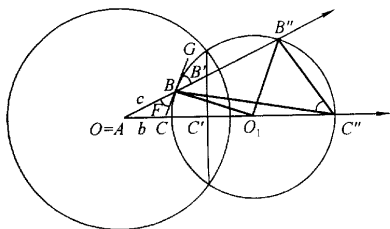


图 11-9

$$BB'' = \frac{1}{c} - c = \frac{2}{\operatorname{sh} c'}.$$

另一个同样证明.

定理 11.3 如图 11-9 对于直角双曲三角形 ABC , 下面 (1)(2)(3) 式成立:

$$(1) \sin A = \frac{\operatorname{sh} a'}{\operatorname{sh} c'} \quad \sin B = \frac{\operatorname{sh} b'}{\operatorname{sh} c'}; \quad (2) \cos A = \frac{\operatorname{th} b'}{\operatorname{th} c'};$$

$$(3) \operatorname{ch} c' = \operatorname{ch} a' \operatorname{ch} b' \quad (\text{勾股定理}).$$

(4) 对于一般的三角形 ABC , 有

$$\operatorname{ch} a' = \operatorname{ch} b' \operatorname{ch} c' - \operatorname{sh} b' \operatorname{sh} c' \cos A. \quad (\text{余弦定理}).$$

证 (1) 只证第二个等式. 利用命题 11.7, 可得

$$\sin B = \sin \frac{1}{2} \angle BA'B'' = \frac{BB''}{CC''} = \frac{\operatorname{sh} b'}{\operatorname{sh} c'}.$$

$$(2) \cos A = \frac{AC'}{AB'}, \quad 1 - AC'^2 = CC' \cdot C'C'' = (AC' - b) \left(\frac{1}{b} - AC' \right). \quad \text{解出}$$

$$AC' = \frac{2b}{1 + b^2} = \operatorname{th} b', \quad \text{同样可解出 } AB'.$$

$$(3) 1 = \cos^2 A + \sin^2 A = \frac{\operatorname{th}^2 b'}{\operatorname{th}^2 c'} + \frac{\operatorname{sh}^2 a'}{\operatorname{sh}^2 c'} \quad \operatorname{sh}^2 c' = \operatorname{sh}^2 a' + \frac{\operatorname{sh}^2 b' \operatorname{ch}^2 c'}{\operatorname{ch}^2 b'},$$

$$\operatorname{ch}^2 c' = \operatorname{ch}^2 a' + \frac{\operatorname{sh}^2 b' \operatorname{ch}^2 c'}{\operatorname{ch}^2 b'}, \quad \text{把它乘出即得.}$$

(4) 在任意双曲三角形 ABC 中, 过 B 作对边的垂线 h 并与它交于 P . 分别记双曲线段 AP, PC 为 b_1, b_2 . 在直角三角形 BPC 中, 利用 (3) 得到

$$\operatorname{ch} a' = \operatorname{ch} h' \operatorname{ch} (b' - b'_1) = \operatorname{ch} c' \operatorname{ch} b' - \operatorname{ch} h' \operatorname{sh} b' \operatorname{sh} b'_1;$$

其中用到直角三角形的勾股定理. 再用 (2) 把 $\operatorname{sh} b'_1 = \frac{\operatorname{ch} b'_1 \operatorname{sh} c' \sin A}{\operatorname{ch} c'}$ 代入

即得.

定理 11.4 在任意双曲三角形 ABC 中, 记 $s' = \frac{1}{2}(a' + b' + c')$, 则有半角公式

$$\sin \frac{A}{2} = \sqrt{\frac{\operatorname{sh}(s' - c') \operatorname{sh}(s' - b')}{\operatorname{sh} c' \operatorname{sh} b'}}.$$

证 利用余弦定理:

$$\begin{aligned} \sin \frac{A}{2} &= \sqrt{\frac{1}{2}(1 - \cos A)} = \frac{1}{\sqrt{2}} \sqrt{1 - \frac{\operatorname{ch} b' \operatorname{ch} c' - \operatorname{ch} a'}{\operatorname{sh} b' \operatorname{sh} c'}} \\ &= \frac{1}{\sqrt{2}} \sqrt{\frac{\operatorname{ch} a' - \operatorname{ch}(b' - c')}{\operatorname{sh} b' \operatorname{sh} c'}} \\ &= \sqrt{\frac{\operatorname{sh}\left(\frac{a' - b' + c'}{2}\right) \operatorname{sh}\left(\frac{a' + b' - c'}{2}\right)}{\operatorname{sh} b' \operatorname{sh} c'}}. \end{aligned}$$

定理 11.5 S 内欧氏三角形 ABC 和双曲三角形 DEF 的对应边都相等 (在各自长度的意义下), 则 DEF 的内角和小于 ABC 的内角和 π .

证 把 ABC 的三个边连同其长度都记为 a, b, c , DEF 与之相应的边及其长度记为 a', b', c' . 由假设 $a = a', s = s'$ 等等.

根据欧氏及双曲三角形的半角公式, 我们只须证明下式:

$$\frac{\operatorname{sh}(s' - c') \operatorname{sh}(s' - b')}{\operatorname{sh} c' \operatorname{sh} b'} < \frac{(s - c)(s - b)}{cb}$$

或
$$\frac{(\operatorname{sh}(s - c) \operatorname{sh}(s - b))}{(s - c)(s - b)} < \frac{\operatorname{sh} c \operatorname{sh} b}{cb}.$$

因为只要三角形不退化为一条线, 总有 $s - b < c$, 而函数 $\frac{\operatorname{sh} x}{x}$ 当 $x \geq 0$ 时上升, 所以上面不等式成立, 即三角形 ABC 的 A 角大于三角形 DEF 的 D 角, 同理, 另外两个角也有同样的不等式, 这就证明了定理.

评述 非欧几何的产生, 源于对欧氏几何平行公设地位的怀疑, 并不牵涉到欧氏几何本身的正确性. 虽然希腊人已认识到数学空间不同于感性认识的空间, 但长期以来人们都承认欧氏几何是物质空间正确的理想化. 非欧几何的出现, 自然引起了它们在物质空间的适用性问题. 这和当时流行的射影几何大不相同. 这个问题一直到爱因斯坦在相对论中用上了非欧几何才得到基本解决.

19 世纪各种几何的相继出现,自然引出一个问题,这些几何之间有什么内在的联系呢?这个问题到了 19 世纪末才由克莱因在他的“埃尔兰根纲领”中加以解决.其基本思路如下:

给了一个几何对象的集合,并在其元素间定义了一类变换(乘法),它们组成一个群.如果有一个变换,把集合中一个元素变为其中另一个元素,则称这两元素相等(或等价、合同).例如在欧氏几何中,任意两个半径相等的圆都相等.

设某类变换把元素 A 变为 B , A 的一些性质, B 未必也有.但既然把 A, B 视为相等,必然有一些它们所共有的性质,这种性质(依赖于所给的变换)叫做这类变换的不变性质.如果这种性质表现为某种量,就叫做不变量.把它们找出来之后,就可以直接从 A, B 本身的性质来判断它们是否相等.研究不同变换下的不变性质,就成为不同几何的研究内容.例如欧氏几何研究的是欧氏平面上的正交变换,不变量是直线段的长度等等.

现在说一下上述几种几何的关系.因为其细节已超出本书的范围,所以只能给一个轮廓.

在射影平面上,射影几何的变换群由全体射影变换组成.仿射几何的变换群由所有保持无穷远直线不变的射影变换组成,因而它是射影群的子群.欧氏几何的变换群由所有保持无穷远直线上两点 $(1, i, 0)$ $(1, -i, 0)$ 不变的仿射变换所组成,是仿射群的子群.另外,椭圆几何的变换群由所有保持二次曲线 $x_1^2 + x_2^2 + x_3^2 = 0$ 不变的射影变换组成,也是射影群的子群.而双曲几何的变换群由所有保持二次曲线 $x_1^2 + x_2^2 - x_3^2 = 0$ 不变的射影变换组成.

索引

一划

一元运算	1.5
一般解(通解)	6.1

二划

二元关系	1.2
反对称性	1.2
对称性	1.2
自反性	1.2
传递性	1.2
二元运算	1.5
二次曲线	7.2
二次曲线的不变量	7.2.2
双曲型曲线	7.2.3
抛物型曲线	7.2.3
椭圆型曲线	7.2.3
二次曲面	8.3
二次曲面的不变量	8.4
无心二次曲面	8.4
双叶双曲面	8.3.3
双曲抛物面(马鞍面)	8.3.5
有心二次曲面	8.4
单叶双曲面	8.3.2
球面	8.1
椭球面	8.3.1
椭圆抛物面	8.3.4
二次型	7.7
二次型的标准形	7.8
二次形的规范形	7.8

不定二次型	7.9
半正定二次型	7.9
半负定二次型	7.9
正定二次型	7.9
负定二次型	7.9
几何向量的坐标表示式	1.7

三划

子式	5.3
主子式	5.3
代数余子式	5.2
余子式	5.2
顺序主子式	7.9
子空间	2.2
平凡子空间	2.2
正交子空间	2.9
向量组张成的子空间	2.2
非平凡子空间	2.2
零子空间	2.2
子空间的交	2.6
子空间的和	2.6
子空间的直和	2.6
子空间的维数公式	2.6

四划

中心投影	10.3
无穷远点	10.3
无穷远直线	10.3
双曲几何的庞加莱模型	11.2
双线性函数	7.7
不相容方程组	1.9

内积	2.7		
内积空间	2.7		六划
实内积空间	2.7	仿射变换	10.2
复内积空间	2.7	扩大的笛卡尔平面	10.3
切平面	9.2	交比	10.3
从切平面	9.2	行列式	5.
密切平面	9.2	n 阶行列式	5.1
切线	9.2	二阶行列式	5.
切向量	9.2	三阶行列式	5.
		三对角行列式	5.2
		上(下)三角行列式	5.1
		范德蒙(Vandermonde)行列式	
			5.2
可对角化	7.5	齐次坐标	10.3
可交换	4.4	全序集	1.4
正交	2.7	全序关系	1.4
正交补	2.9	自配极三角形	10.4
正交变换	7.1, 10.1	向量	1.7 1.8
可逆元	1.10.1	几何向量	1.7
主轴定理	7.8	负向量	1.7
平面方程	1.9.1	单位向量	1.7.2
平面的一般方程	1.9.1	零向量	1.7
平面的点法式方程	1.9.1	n 元(维)向量	1.8 2.4
平面的截距式方程	1.9.1	向量的内积(点积, 数量积)	1.7.3
矛盾律	1.6	向量的外积(叉积, 向量积)	1.7.4
对偶原理	10.3	向量的加法	1.7.1 1.8
代数系统	1.10	向量的坐标	2.5
代数运算	1.5	向量的混合积	1.7.5
代数结构	1.10	向量的数量乘法	1.7.2 1.8
主方向	9.7	向量的模	1.7
主法线	9.2	向量空间	2.1
主法线向量	9.2	向量组的秩	2.4
弗雷耐(Frenet)标架	9.2	约当标准形	7.5
弗雷耐(Frenet)公式	9.3		

曲线	9.1	无零因子环	1.10.2
正则曲线	9.1	交换环	1.10.2
光滑曲线	9.7	含么环	1.10.2
曲面	9.5	整环	1.10.2
曲面第一类基本量	9.6	极线	10.4
曲面第一基本形式	9.6	极点	10.4
曲面第二类基本量	9.7	极大线性无关组	2.4
曲面第二基本形式	9.7	欧氏(Euclid)空间	2.7
曲率	9.3	单位元	1.10.1
主曲率	9.7	单位点	10.3
平均曲率	9.7	单位正交基(标准正交基)	2.8
曲率中心	9.3	空间直角坐标系	1.7.6
曲率本径	9.3	空间直线方程	1.9.2
曲率本径	9.3	直线的一般方程	1.9.2
曲率向量	9.3	直线的参数方程	1.9.2
法曲率	9.7	直线的标准方程	1.9.2
法曲率向量	9.7	空间 $L(V_1, V_2)$	3.3
测地曲率	9.7	线性几何	
测地曲率向量	9.7	线性方程组	1.8
高斯曲率	9.7	齐次线性方程组	1.8
向量函数	9.1	非齐次线性方程组	1.8
定长向量函数	9.1	线性无关	2.3
定向向量函数	9.1	线性扩张	2.2
原向量函数	9.1	线性组合(线性表示)	2.1
		线性空间	2.1
		无限维线性空间	2.2
		有限维线性空间	2.2
坐标三点形	10.3	实空间	2.1
序偶(有序二元组)	1.1	复空间	2.1
良序集	1.4	线性空间的同构	3.5
酉空间	4.10	线性空间的基	2.4
		线性空间的维数	2.4
		线性变换	3.1
环	1.10.2		

恒等变换	3.1	命题运算	1.6
数乘变换	3.1	双蕴涵词	1.6
零变换	3.1	否定词	1.6
旋转变换	3.1	合取词	1.6
镜象变换(镜面反射)	3.1	析取词	1.6
比例变换	3.1	蕴涵词	1.6
错切变换	3.1		
线性相关	2.3	九划	
线性相关性	2.3	逆元	1.10.1
线性映射	3.1	矩阵	1.8 4.1
线性映射的加法	3.3	上(下)三角矩阵	4.1
线性映射的秩	3.4	反对称矩阵	4.6
线性映射的逆映射	3.3	方阵	4.1
线性映射的秩	3.4	分块矩阵	4.9
线性映射的核	3.2	正交矩阵	7.1
线性映射的乘法	3.3	正定矩阵	7.9
线性映射的乘积的秩	3.4	半正定矩阵	7.9
线性映射的象(值域)	3.2	半负定矩阵	7.9
线性映射的数乘	3.3	对角矩阵	4.1
线性映射象和核的维数公式		可逆矩阵	4.5
	3.4	对称矩阵	4.6
直线把	10.3	负定矩阵	7.9
直积(笛卡儿乘积)	1.1	约当型矩阵	7.5
变积系数	10.2	初等矩阵	4.7
命题	1.6	系数矩阵	1.8
永真命题	1.6	伴随矩阵	5.3
永假命题(矛盾命题)	1.6	实对称矩阵	7.6
条件命题	1.6	单位矩阵	4.2
否定命题	1.6	线性映射的矩阵(表示)	4.2
复合命题	1.6	逆矩阵(非奇异矩阵)	4.5
原子命题	1.6	度量矩阵	7.7
等价(等值)命题	1.6	相似矩阵	7.3
命题的真值	1.6	相合(合同矩阵)	7.7

相抵矩阵	4.8	恒等(单位)映射	1.5
基的变换矩阵(过渡矩阵)	4.10	满射	1.5
满秩矩阵	4.8	标准内积	2.7
零矩阵	4.2	弧长	9.2
数量矩阵	4.2	弧长参数	9.2
增广矩阵	1.8	十划	
幂等矩阵	4.5 例 4		
幂零矩阵	第 4 章习题 18	调和线束	10.3
矩阵的列秩	4.8	调和点列	10.3
矩阵的行秩	4.7	高斯消元法	1.8
矩阵的行列式秩	5.3	特解	6.2
矩阵的列空间	6.1	特征值	7.4
矩阵的行空间	6.1	特征向量	7.4
矩阵的初等变换	4.7	特征子空间	7.4
矩阵的初等行变换	4.7	特征多项式	7.4
矩阵的初等列变换	4.7	原始射影坐标系	10.3
矩阵的转置	4.6	射影平面	10.3
矩阵的迹	第 4 章补充题 2	射影坐标变换	10.3
矩阵的秩	4.8	射影直线	10.3
相似	7.3	射影映射	10.4
相似标准形	7.5	射影变换	10.4
相合(合同)	7.7	十一划	
相合(合同)标准形	7.8		
相抵	4.8	基	2.4
相抵标准形	4.8	域	1.10.2
测地线	11.1	有理数域	1.10.2
测地三角形	11.1	实数域	1.10.2
相容方程组	1.8	复数域	1.10.2
柱面	8.1	像	1.5
映射	1.5	完全原像	1.5
双射	1.5	原像	1.5
可逆映射	1.5	维数	2.4
单射	1.5	排中律	1.6

旋转面	8.1	子群	1.10.1
偏序集	1.4	无限群	1.10.1
偏序关系	1.4	半群	1.10.1
基础解系	6.1	交换群(Abel 群)	1.10.1
惯性定理	7.8	有限群	1.10.1
正惯性指数	7.8	含幺半群	1.10.1
负惯性指数	7.8	克莱因(Klein)群	1.10.1
球极投影	11.1	变换群	1.10.1
副法线	9.2	置换群	1.10.1
副法线向量	9.2	模 n 剩余类加群	1.10.1
十二划		n 阶线性群	3.3
		简单比	10.2
		锥面	8.1
集合	1.1	置换	1.10.1
子集	1.1	零因子	1.10.2
交集	1.1	左零因子	1.10.2
并集	1.1	右零因子	1.10.2
余集	1.1	解向量	1.8
空集	1.1	解空间	2.1
商集	1.1	数学归纳法原理	1.4
幂集	1.1	第二数学归纳法	1.4
量词	1.6	微分几何	9.0
存在量词 \exists	1.6	Cramer 法则	5.4
全称量词 \forall	1.6	Cauchy – Schwarz 不等式	2.7
等价类	1.3	Schmidt 正交化过程	2.8
等价关系	1.3		
幂等线性变换	7.5		

十三划

群	1.10.1
---	--------