普通高等教育“十一五”国家级规划教材

线性代数

Linear Algebra

李尚志

高等教育出版社
HIGHER EDUCATION PRESS

普通高等教育"十一五"国家级规划教材

线性代数

Xianxing Daishu

李尚志

高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书是在作者主持的国家精品课程“线性代数(非数学专业)”的建设过程中形成的教材，是作者主持的国家级教学成果奖二等奖项目“数学建模思想融入基础课教学”的重要成果之一。

本书不是“奉天承运皇帝诏曰”从天而降的抽象定义和推理，而是一部由创造发明的系列故事组成的连续剧。每个故事从颇具悬念的问题开始，在解决问题的过程中将所要学习的知识一步一步“发明”出来。随着剧情的发展，知识的引入如“随风潜入夜”，知识的应用如“润物细无声”，都成为自然而然的了。

本书适合作为大学本科非数学类专业线性代数、工科高等代数课程的教材，也可作为需要或关心线性代数和矩阵论知识的科技工作者、工程技术人员、大专院校师生及其他读者的参考书。

图书在版编目(CIP)数据

线性代数／李尚志编著．—北京：高等教育出版社，2011.6

ISBN 978-7-04-031795-4

Ⅰ.①线… Ⅱ.①李… Ⅲ.①线性代数-高等学校-教材

Ⅳ.①O151.23

中国版本图书馆CIP数据核字(2011)第056928号

策划编辑 兰莹莹　责任编辑 张彦云　封面设计 张 楠　版式设计 王 莹

责任绘图 尹文军　责任校对 殷 然　责任印制 朱学忠

出版发行 高等教育出版社
社　　址 北京市西城区德外大街4号
邮政编码 100120
印　　刷 山东省高唐印刷有限责任公司
开　　本 787×960 1/16
印　　张 19
字　　数 350 000
购书热线 010—58581118
咨询电话 400—810—0598
网　　址 http://www.hep.edu.cn
　　　　 http://www.hep.com.cn
网上订购 http://www.landraco.com
　　　　 http://www.landraco.com.cn
版　　次 2011年6月第1版
印　　次 2011年6月第1次印刷
定　　价 26.00元

物 料 号 31795—00

前　言

有人说，文学的永恒主题是爱与死。

数学的永恒主题是什么？问题的回答很可能众说纷纭，并且随着科学的发展而不断变化。但无论如何，可以说函数与方程是数学的重要主题，至少是中学数学与大学数学的重要主题。

最简单的函数是一次函数，最简单的方程是一次方程。

中学数学中已经学习过一元的一次函数与一次方程。

大学数学最重要的两门课程是微积分和线性代数。

微积分是把非一次的函数与方程化成一次函数与一次方程来研究。

线性代数干什么？“线性”就是“一次”，线性代数的主要内容就是研究多元的一次方程组与一次函数组。一次方程组也称为线性方程组。常数项为 0 的多元一次函数 $y = a_1x_1 + \cdots + a_nx_n$ 称为线性函数，n 个 n 元线性函数组成的函数组 $y_i = a_{i1}x_1 + \cdots + a_{in}x_n$ $(1 \leqslant i \leqslant n)$ 称为线性变换。大学非数学类专业线性代数课程的两大主题就是多元线性方程组与线性变换。线性方程组与线性变换可以由方程组或函数组中的系数排成矩阵来表示，通过矩阵运算来求解方程及解决相关的应用问题，其中最重要最基本的矩阵运算是初等变换和乘法这两个算法。用这两个算法为两大主题服务，是线性代数课程教学的基本任务。算法只有两个，要解决的问题却是千千万万，这就需要练就一双慧眼，从不同问题中看出共同点，将众多的问题转化为两个算法可以解决的形式。算法有一定之规，不难学会，难的是练就这双慧眼，提高实现这种转化的能力，这需要在长期实践中努力，不是仅仅通过几十学时的课堂教学所能做到的。但是，课堂教学至少应当让学生有一个良好的开端，朝正确的方向前进，而不要背道而驰，南辕北辙。

本书是为大学本科非数学类专业线性代数、工科高等代数课程编写的教材，基本目标不是将学生培养成数学理论的研究人员，而是让他们熟练掌握矩阵的初等变换与乘法这两种算法，逐步学会用这两种算法来解决学习和工作中与线性方程组和线性函数组有关的问题。我们不是以“奉天承运皇帝诏曰”的方式从天而降概念、算法和定理，也不是在学生不知有何用处的情况下先学好算法再拿去应用，而是从问题出发，在尝试解决问题的过程中将所需的算法“发明”出来。书中所有推理和证明的目的都是为了训练学生应用算法解决问

题的能力, 使学生练就从纷纭复杂的问题中看出通向已有算法的出路的一双慧眼, 而不是为了“数学的严密性”。数学的严密性, 本身就是为了搞清楚相关结论和算法成立的理由和适用范围, 保证算法在一定范围内的正确性和可靠性, 既要防止超出适用范围发生错误, 又要在适用范围内让结论和算法大显身手。考虑到很多非数学类专业的线性代数课程课时比较少, 我们将推理和证明写得比较简略, 尽量通过具体例子来体现普遍规律。有些结论和算法的证明和推理难度较大, 我们就将它们写成附录, 仅供教师或一部分感兴趣的学生参考, 不作为课程学习内容, 学生知道结论、会算会用就行了, 暂时不必知其所以然。第 2 至 5 章有 6 节的标题用星号 * 标注出来, 供工科高等代数课程选用, 不作为线性代数课程的学习内容。线性代数课程只需在没有标 * 的章节中选择教学内容, 并且还可以根据不同层次再略去一些内容。简而言之, 本书为非数学类专业所有的教学和学生设置了共同的起点, 共同的前进方向, 但有几个不同的下车站。不同院校不同专业的教学可根据学生的不同层次选择不同的下车站, 有的可以一直到终点站才下车, 有的则可以选择适当的中途站, 提前下车。

既然线性代数研究的是最简单的方程和函数, 算法又比微积分少得多, 按道理应当容易学。但实际情况是: 很多学生学起来并不轻松。主要的困难是太抽象。比如, 微积分中的导数可以理解为切线的斜率、运动的速度, 定积分可以理解为求图形的面积、由速度求路程, 这都比较具体。但线性代数从一开始就是一个接一个从天而降的抽象定义, 使初学者难以理解。比如: 行列式为什么要这样定义? 矩阵为什么要这样相乘? 向量到底是有方向和大小的量, 还是数组, 还是定义了加法和数乘的任意非空集合中的元素? 线性相关、线性无关是什么意思, 有什么用处? 这些问题都让学生甚至很多讲授线性代数课程的教师迷惑不解。如果问学生为什么要学习行列式、矩阵和向量, 为什么要学矩阵乘法和初等变换, 很多学生的回答将是: 教材上写了这些内容, 考试要考这些内容, 考试不及格就不能毕业, 找不到工作, 我敢不学吗? 很多学生被线性代数折腾得很苦, 死记硬背定义和算法, 不知道这些定义和算法除了应付考试还有什么别的用处。即使是将定义和算法背熟了、考好了、毕业了的学生, 在工程应用和理论研究中遇到问题需要解决的时候也不会应用线性代数的知识和算法。例如, 需要解线性方程组时只会用中学的加减消去法而不会用矩阵的初等变换或求逆; 需要根据顶点坐标计算平行四边形和三角形面积的时候不知道用二阶行列式; 需要判断线性方程组是否有唯一解的时候不知道这就是判断系数矩阵各列是否线性无关; 需要计算平面上和空间中的旋转时不知道用矩阵作乘法来实现; 需要计算空间中的旋转轴时不知道这就是求特征值为 1 的特征向量, 等等。

针对这种情况, 本教材不从定义出发而从问题出发引入概念, 引导学生在

尝试解决这些问题的过程中将所要讲授的知识重新“发明”出来。选择的问题，尽量是在现实生活和学生今后的工作中有用的问题，并且希望是学生感兴趣、容易懂的，因此也不选择需要较多专业知识或者综合性太强的实际问题。本书第一个例子是求前 n 个正整数的 k 次方和 $S_n = 1^k + 2^k + \cdots + n^k$ 的公式。这个问题看起来很困难，通过将 S_n 看成满足条件 $S_n - S_{n-1} = n^k$ 的多项式函数 $f(n) = a_1 n + a_2 n^2 + \cdots + a_{k+1} n^{k+1}$，就归结为求解待定系数 $a_1, a_2, \cdots, a_{k+1}$ 满足的三角形线性方程组，变得容易了。很自然提出问题：如果线性方程组不是三角形方程组，怎么求解？解决问题的思路也很自然：变成三角形方程组来求解。本书第 1 章第二个例子是求二次函数满足三组对应值，归结为解三元一次方程组。用中学的加减消去法求解这个三元一次方程组，引入了线性组合、方程组的同解变形、初等变换等概念和方法，再进一步将字母省去，将方程组用矩阵表示，方程组的初等变换用矩阵的初等行变换表示，得到了矩阵消去法。我们不是先讲矩阵的初等变换再用其来解方程组，而是为了解方程组“发明”出矩阵的初等行变换，这样引入初等变换就比较自然，学生容易理解和接受。但这还不够，矩阵的初等变换一旦发明出来，其用途就不限于解线性方程组，以后还要用来解决许许多多其他的问题，例如：判断线性相关和无关、求极大线性无关组和秩、计算行列式、求二次型的标准型、……人们买东西的时候都希望花一笔钱买来的一样东西有多样的用途，在科学研究中，用从许多事物中总结出来的结论和方法来解决千千万万个不同的问题，这就是抽象的威力，这样的抽象是好事。可是，如果在教学中不举任何一个实例，不讲任何一个应用，只让学生像念咒语一样死记硬背来应付考试，当然就会导致学生害怕抽象，这不是真正的抽象而是冒牌的抽象。因而不是抽象的错，也不是学生的错，而是教学方式的错。

本书第 2 章仍然围绕线性方程组这个主题，开始第一个例子就是讨论多项式函数 $f(x)$ 各项系数 $a_0, a_1, \cdots, a_m$ 满足的线性方程组什么时候有唯一解，在第 2.4 节中得到了求唯一解的公式，在第 3 章中又通过计算这个方程的系数行列式 (也就是范德蒙德 (Vandermonde) 行列式) 再次讨论了唯一解条件。为了讨论线性方程组的唯一解条件，我们先将二元和三元线性方程组写成向量形式，转化为平面和空间向量共线、共面的几何问题来讨论，再将共线、共面条件用向量的代数运算来描述，推广到 n 维空间，引出了线性相关、线性无关、维数、基、坐标等重要概念，并且在讨论方程组的解不唯一的情况时引出了秩、子空间等重要概念。很多学校因为课时不够而不讲线性空间。我们只花了课时讨论线性方程组的解的唯一性条件、解集合的大小和构造，与此同时也就讲完了线性空间的主要内容，不需要另外花课时，岂不两全其美？

第 2 章通过对平面向量共线与共面的几何问题的代数描述引出了线性相

关、线性无关等代数概念,解决了线性方程组解的唯一性及解集合的结构问题。第 3 章 (行列式) 仍然围绕线性方程组这个主题,仍然在几何和代数之间左右逢源。对 2 维和 3 维的情形,向量是否共线或共面可以通过平行四边形面积或平行六面体体积是否为 0 来描述。我们先由平行四边形面积和平行六面体体积引入二阶和三阶行列式,由它们的几何定义得到代数性质,由代数性质得出代数算法。将代数性质和代数算法推广,就得到 n 阶行列式。我们所用的三条代数性质是:第一,既然矩形的面积是长宽相乘,长方体或正方体的体积是长宽高相乘,将平行四边形面积和平行六面体体积看成两边或三条棱所代表的向量的某种乘积,按分配律和对于数乘的结合律展开。第二,更加理所当然的是,如果平行四边形或平行六面体有两条边或两条棱重合,面积或体积就等于 0。第三,单位正方形的面积和单位正方体的体积等于 1。我们正是从这三条性质推出了二阶和三阶行列式的计算公式,并且将算法和计算公式推广到了 n 阶行列式。不过,在这一章中我们并没有强调行列式的定义,行列式性质的证明也尽量淡化,将其中比较困难的证明放到附录中仅供参考,我们只突出了行列式性质的应用,尤其是初等变换对行列式的影响,特别是第三类初等变换不改变行列式的性质,通过初等变换计算行列式以及证明相关的定理 (主要是用行列式判断线性相关和无关的定理,及关于线性方程组唯一解的判定和唯一解公式的克拉默 (Cramer) 法则)。

矩阵的分块运算,是以华罗庚为代表的中国代数学家们进行科学研究的一个重要武器。然而,在很多线性代数教材和课堂中,分块运算却被认为没有什么用处,这是因为在这些教材和课堂中讲分块运算的时候没有给它用武之地,没有举出有用的例子。本教材中淡化分块运算的定义和证明,但将分块运算大用特用,让它大显神通。第 2 章中,在引入矩阵乘法之前,我们将向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合 $x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m$ 写成由向量组成的“行向量” $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 与系数组成的列向量 $\boldsymbol{X}$ 的乘积的形式 $\boldsymbol{AX}$,就是按分块运算的方式定义了矩阵 $\boldsymbol{A}$ 与列向量 $\boldsymbol{X}$ 的乘积。其后我们又将 S 的若干个线性组合组成的向量组 $(\boldsymbol{AX}_1,\cdots,\boldsymbol{AX}_p)$ 写成 $\boldsymbol{A}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_p)$,这就是用分块运算一般地定义了两个矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{X}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_p)$ 的乘积。第 4 章正式定义矩阵乘法之后,又用分块运算讨论矩阵乘法的各种重要性质,包括对角矩阵、纯量矩阵、单位矩阵的乘法性质,矩阵乘法的分配律和结合律。矩阵乘法的结合律,本来是被认为最繁琐最无法直观理解的,通过分块运算得到了自然的解释。在第 4 章的以后各节中,矩阵的分块运算还被用来计算逆矩阵、将初等变换转化为矩阵乘法、证明关于矩阵秩的不等式、计算行列式。

当我们通过字母运算得到 $(a+b)^2=a^2+2ab+b^2$ 时,不必考虑当 a,b 为分数时是否需要通分、当 a,b 符号相反时是否需要将绝对值相减、当 a,b 是无理

数时是否需要极限运算，只要利用所有的复数共同满足的运算律：乘法对于加法的分配律、加法结合律、乘法交换律，就得到了所需的结论。这体现了运算律的威力，也体现了抽象的威力。如果将 a,b 换成同阶方阵 $\boldsymbol{A},\boldsymbol{B}$，也不必考虑矩阵乘法与数的乘法有多大的区别，只需考虑所用的三个运算律是否仍成立。矩阵乘法仍满足对加法的分配律，矩阵加法仍满足结合律，但矩阵乘法不满足交换律。不过，某些特殊的方阵 $\boldsymbol{A},\boldsymbol{B}$ 作乘法时可以交换：$\boldsymbol{AB}=\boldsymbol{BA}$，对这样的 $\boldsymbol{A},\boldsymbol{B}$，不但完全平方公式 $(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2$ 成立，而且可以用牛顿二项式定理展开 $(\boldsymbol{A}+\boldsymbol{B})^n$。在第 4 章中，以若尔当 (Jordan) 形矩阵

$$\boldsymbol{J}=\begin{pmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{pmatrix}=\lambda\boldsymbol{I}+\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$$

为例，用牛顿二项式定理计算 $\boldsymbol{J}^{10}$；当 $\lambda=1$ 时用泰勒 (Taylor) 展开式 $(1+x)^{-1}=1-x+x^2+\cdots$ 求逆 $\boldsymbol{J}^{-1}$；用 $(1+x)^{\frac{1}{n}}=1+\dfrac{1}{n}x+\dfrac{\dfrac{1}{n}\left(\dfrac{1}{n}-1\right)}{2}x^2+\cdots$ 求 $\boldsymbol{J}$ 的一个“n 次方根” $\boldsymbol{X}$ 使其满足 $\boldsymbol{X}^n=\boldsymbol{J}$；还可以用泰勒 (Taylor) 展开式 $\mathrm{e}^x=1+x+\dfrac{1}{2!}x^2+\cdots$ 计算微分方程组 $\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}=\boldsymbol{JX}$ 的解 $\boldsymbol{X}=\mathrm{e}^{\boldsymbol{J}t}C$。很自然引出一个问题：怎样求任意方阵 $\boldsymbol{A}$ 的 n 次幂、方根和指数函数 $\mathrm{e}^{\boldsymbol{A}}$？这就为将 $\boldsymbol{A}$ 相似于对角矩阵或若尔当 (Jordan) 标准形埋下了伏笔。

我参加面试 2010 年入学的研究生的时候，问了一个问题：求矩阵

$$\boldsymbol{A}=\begin{pmatrix}\cos\alpha&-\sin\alpha\\\sin\alpha&\cos\alpha\end{pmatrix}$$

的 2010 次幂。几乎所有的线性代数教材上都有对这个矩阵 $\boldsymbol{A}$ 求 $\boldsymbol{A}^n$ 的习题，按道理学生应当会做。让我震惊的是，将近十名在笔试中考高分的非数学类专业的学生没有一个能够给出答案。几乎每个考生都说，先算 $\boldsymbol{A}$ 的平方，再算 $\boldsymbol{A}$ 的 3 次方，观察规律，然后用数学归纳法证明。我问：你们的教科书上都有这道题，为什么以前没有算过 $\boldsymbol{A}$ 的平方与 3 次方，没有观察过规律？他们没有回答。我猜想，大概他们的老师也认为这个题只能先算平方、3 次方观察规律，再用数学归纳法证明，认为这样的题目没有什么意思，因此没有布置学生做，学生也就没有做过。然而，只要知道用矩阵 $\boldsymbol{A}$ 左乘列向量 $\boldsymbol{X}$ 的效果

$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}\cos\alpha&-\sin\alpha\\\sin\alpha&\cos\alpha\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x'\\y'\end{pmatrix}$$

是将点 $P(x,y)$ 旋转角 α 变到点 $P'(x',y')$, 就可立即知道: $\boldsymbol{A}^n$ 乘 $\boldsymbol{X}$ 的效果是将旋转 α 的动作重复进行 n 次, 总效果是旋转 $n\alpha$, 因而

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}^n = \begin{pmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{pmatrix}。$$

可是, 这些学生都不知道 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 是平面旋转, 因此都觉得要计算 $\boldsymbol{A}$ 的几千次幂是一件不可能完成的事情。为什么不知道, 是因为他们的老师没有告诉他们。为什么不告诉? 老师的解释可能是: 我们的课时不够, 不能讲线性变换。实际上, 不需要多少课时, 只要一句话就可以讲线性变换: 将列向量 $\boldsymbol{X}$ 乘方阵 $\boldsymbol{A}$ 得到列向量 $\boldsymbol{Y}$, 就是线性变换。将平面或空间每个点 P 通过旋转或轴对称变到点 P', 从 P 的坐标 $\boldsymbol{X}$ (写成列向量) 到 P' 的坐标 $\boldsymbol{Y}$ 的变换可以用某个方阵 $\boldsymbol{A}$ 左乘 $\boldsymbol{X}$ 来实现: $\boldsymbol{Y} = \boldsymbol{AX}$, 因此是线性变换。

本书没有花很多篇幅讨论线性变换, 只是在第 4.2 节中以平面上的旋转与轴对称两个例子来加以说明。线性变换 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 的各种性质 (保加法、保数乘、矩阵 $\boldsymbol{A}$ 的各列是自然基的像) 都由矩阵乘法的性质得出, 这一节对线性变换的讲解就变成了对矩阵乘法的训练, 训练学生理解矩阵乘法的代数性质的几何意义。第 4.3 节讲矩阵的逆, 第一个例子就是求平面旋转矩阵 $\boldsymbol{A}$ 的逆, 不是通过代数运算求逆, 而是通过几何观点求逆: 既然 $\boldsymbol{A}$ 表示旋转角 α, 旋转 $-\alpha$ 的矩阵就是 $\boldsymbol{A}^{-1}$。第 4.4 节的一个例子中, 将椭圆通过线性变换 $(x,y) \mapsto \left(x, \dfrac{a}{b}y\right)$ 拉伸成圆, 由圆的面积及内接四边形的最大面积来求椭圆面积及内接四边形的最大面积。通过这个简单易懂的例子解释了变换矩阵的行列式的几何意义, 引出了行列式的乘法性质。这个“拉伸”变换的矩阵是对角矩阵, 为第 5 章的特征向量和相似对角化埋下了伏笔。第 5 章第 5.5 节前两个例子与此一脉相承, 用非对角矩阵将斜着摆放的椭圆拉伸成圆。通过计算被拉伸的向量引出了特征向量, 并让学生一开始就知道了特征向量的几何意义。本书特征向量的例子, 一类是通过计算方阵 $\boldsymbol{A}$ 的特征向量研究对应的线性变换 $\boldsymbol{X} \mapsto \boldsymbol{AX}$ 的几何性质, 既有将向量拉长或缩短的变换的例子, 也有空间中旋转和对称变换的例子; 另一类是以计算 $\boldsymbol{A}$ 的函数如 $\boldsymbol{A}^n, \mathrm{e}^{\boldsymbol{A}t}$ 为目标的代数例子, 解释了将方阵 $\boldsymbol{A}$ 相似于对角矩阵或若尔当形矩阵的必要性和算法。

线性代数的算法只对行数和列数很少的矩阵才能用手算实现。而在实际工作中, 经常需要处理几十、几百甚至更多行和列的矩阵, 难以用手工实现算法, 必须求助于计算机及其软件。已经有很多软件能够满足这样的要求, 例如 MATLAB 与 Mathematica。西安电子科技大学陈怀琛教授主持的教育部教

改项目“用 MATLAB 和建模实践改造工科线性代数课程”在利用信息技术工具改革线性代数课程方面进行了有益的探索和实践。本书作者也组织了一些教师参加这个项目。本书以数学实验的方式附上了线性代数一些主要算法的 MATLAB 命令和 Mathematica 命令。其中 MATLAB 的例子是青年教师李娅为本书选择和整理的，谨在此感谢。读者可以“依样画葫芦”运行这些命令，再“照葫芦画瓢”替换其中的数据计算别的习题。关于 MATLAB 命令和 Mathematica 命令更详细的知识，请参考相关的书籍和资料，例如本书参考文献 [2],[3]。

李尚志

2010 年 12 月 31 日

目　录

第 1 章　线性方程组的解法

§1.1　线性方程组的初等变换

1.1.1　多元线性方程组

中学数学学习了二元一次方程组的解法, 但在科学研究和实际应用中经常需要解更多未知数的一次方程组.

例 1　已知正整数 n, 求

(1) $1^2+2^2+\cdots+n^2$;

(2) $1^4+2^4+\cdots+n^4$.

分析　中学数学给出了前 n 个正整数的平方和 $1^2+2^2+\cdots+n^2$ 的公式, 并且用数学归纳法证明了公式的正确性, 但没有教学生怎么"发明"出这个公式, 也不能类似地得到更高次数的幂的和 $1^k+2^k+\cdots+n^k$ 的公式.

本例是已知数列通项公式 $u_n=n^k$ 求前 n 项和 S_n, 这很困难. 反过来, 已知前 n 项和 S_n 求 u_n 却很容易: $u_n=S_n-S_{n-1}$ (当 $n\geqslant 2$), $u_1=S_1=1$.

如果将 S_n 看成 n 的函数 $S_n=f(n)$, 则上述等式就是函数 $f(n)$ 应当满足的必要条件:

$$f(n)-f(n-1)=u_n=n^k\ (\forall\ n\geqslant 2),\qquad f(1)=u_1=1.$$

不难想到, 如果 $f(n)$ 是 n 的多项式函数, 则 $f(n)-f(n-1)$ 也是多项式, 并且次数比 $f(n)$ 低一次. 如果能用待定系数法求 $k+1$ 次多项式 $f(n)=a_0+a_1n+\cdots+a_{k+1}n^{k+1}$ 的系数, 使 $f(n)$ 满足恒等式

$$f(n)-f(n-1)=a_1[n-(n-1)]+\cdots+a_{k+1}[n^{k+1}-(n-1)^{k+1}]=n^k,$$

则

$$\begin{aligned}S_n&=1^k+2^k+\cdots+n^k\\&=(f(1)-f(0))+(f(2)-f(1))+\cdots+(f(n)-f(n-1))\\&=f(n)-f(0).\end{aligned}$$

只要取 $f(0)=a_0=0$, 即可使 $S_n=f(n)$ 对所有的正整数 n 成立.

解　(1) 设 $f(n)=a_1n+a_2n^2+a_3n^3$, 满足条件

$$f(n)-f(n-1)=a_1[n-(n-1)]+a_2[n^2-(n-1)^2]+a_3[n^3-(n-1)^3]$$

$$= a_1 + a_2(2n-1) + a_3(3n^2 - 3n + 1)$$
$$= (a_1 - a_2 + a_3) + (2a_2 - 3a_3)n + 3a_3n^2 = n^2, \text{ 即}$$

$$\begin{cases} a_1 - a_2 + a_3 = 0, & ① \\ 2a_2 - 3a_3 = 0, & ② \\ 3a_3 = 1. & ③ \end{cases} \tag{1.1.1}$$

由方程 ③ 解出 $a_3 = \frac{1}{3}$, 代入 ② 解出 $a_2 = \frac{1}{2}$, 再代入 ① 解出 $a_1 = \frac{1}{6}$, 得到

$$S_n = \frac{1}{6}n + \frac{1}{2}n^2 + \frac{1}{3}n^3 = \frac{2n^3 + 3n^2 + n}{6} = \frac{n(n+1)(2n+1)}{6}.$$

(2) 将 $S_n = f(n) = a_1n + a_2n^2 + a_3n^3 + a_4n^4 + a_5n^5$ 代入 $f(n) - f(n-1) = n^4$, 整理并比较对应项系数得

$$\begin{cases} a_1 - a_2 + a_3 - a_4 + a_5 = 0, & ① \\ 2a_2 - 3a_3 + 4a_4 - 5a_5 = 0, & ② \\ 3a_3 - 6a_4 + 10a_5 = 0, & ③ \\ 4a_4 - 10a_5 = 0, & ④ \\ 5a_5 = 1. & ⑤ \end{cases} \tag{1.1.2}$$

由下而上依次从各方程中解出 $a_5 = \frac{1}{5}, a_4 = \frac{1}{2}, a_3 = \frac{1}{3}, a_2 = 0, a_1 = -\frac{1}{30}$, 得到

$$S_n = \frac{1}{5}n^5 + \frac{1}{2}n^4 + \frac{1}{3}n^3 - \frac{1}{30}n = \frac{n(6n^4 + 15n^3 + 10n^2 - 1)}{30}$$ □

例 1 中的 (1.1.1),(1.1.2) 分别是 3 元一次方程组与 5 元一次方程组.

一般地, n 个未知数 $x_1, \cdots, x_n$ 的如下形式的方程

$$a_1x_1 + \cdots + a_nx_n = b$$

称为 n 元一次方程, 也称为 n 元线性方程, 其中 $a_1, \cdots, a_n, b$ 是已知的常数, $a_1, \cdots, a_n$ 是一次项系数, b 是常数项. 具有同样 n 个未知数 $x_1, \cdots, x_n$ 的若干个一次方程组成的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.1.3}$$

称为 n 元一次方程组, 也称 n **元线性方程组** (linear equations in n variables).

注意, 在线性方程组中, 并不要求方程的个数 m 等于未知数个数 n, $m<n, m=n, m>n$ 等三种情况都允许. $m=1$ 的情形也允许. 也就是说: 一个线性方程也可以组成线性方程组.

如果将方程组 (1.1.3) 中的 n 个字母 $x_1,\cdots,x_n$ 分别替换成某 n 个已知数 $c_1,\cdots,c_n$, 得到的 m 个等式

$$a_{i1}c_1+\cdots+a_{in}c_n=b_i \quad (1\leqslant i\leqslant m)$$

全部成立, 就称 $c_1,\cdots,c_n$ 组成的有序数组 $(c_1,\cdots,c_n)$ 是方程组 (1.1.3) 的一个**解** (solution). 方程组的全体解组成的集合称为方程组的解集.

1.1.2 线性方程组的同解变形

例 1 的线性方程组 (1.1.1),(1.1.2) 具有特殊的形状: 从上到下每个方程比上一个方程少含一个未知数, 最后一个方程只含一个未知数. 将各方程的等号上下对齐, 同一未知数的项也上下对齐, 则等号左边左下角是空白 (其中各项系数都是 0), 其余各项组成一个摆放在右上角的三角形. 我们称这样的线性方程组为“上三角形”. 例 1 的解法可以推广到一般的上三角形方程组: 从最后一个方程解出所含的未知数的值, 代入上一个方程再解出一个未知数的值, 由下而上依次将已求出的未知数的值代入上一个方程, 可以依次求出所有的未知数的值, 从而求出方程组的解.

任给的线性方程组 U 不一定是三角形, 但如果能够变成例 1 那样的三角形方程组 T, 就可以求出方程组 T 的解. 但是, 必须首先保证 T 与 U 同解, 由 T 得到的才是 U 的全部解.

例 2 二次函数 $y=f(x)$ 的图像经过三个已知点 $(1,1),(2,2),(3,0)$. 求 $f(4)$.

分析 设所求二次函数为 $f(x)=ax^2+bx+c$, 其中 a,b,c 是待定常数. 则

$$\begin{cases} a+\ \ b+c=1, & ① \\ 4a+2b+c=2, & ② \\ 9a+3b+c=0. & ③ \end{cases} \tag{1.1.4}$$

尝试用中学的加减消去法解这个三元一次方程组.

方程② − 方程①, 得 $\qquad 3a+b=1.\qquad$ ④

方程③ − 方程②, 得 $\qquad 5a+b=-2.\qquad$ ⑤

方程⑤ − 方程④, 得 $\qquad 2a=-3.\qquad$ ⑥

由 ⑥ 可解出 a, 代入 ④ 可解出 b. 将 a,b 代入 ① 可解出 c. 这其实是从方程组

$$\begin{cases} a+\ \ b+c=1, & ① \\ 3a+b \qquad\ \ =1, & ④ \\ 2a \qquad\qquad\ =-3 & ⑥ \end{cases} \tag{1.1.5}$$

解出 a,b,c. 方程组 (1.1.5) 等号左边的右下角是空白, 左上角组成三角形, 可以由下而上依次求出各未知数的值. 但有一个问题: 由方程组 (1.1.5) 求出的是否是原方程组 (1.1.4) 的解集?

新方程 ④,⑤ 由原方程 ①,②,③ 相减得到, 新方程 ⑥ 又由 ④,⑤ 相减得到. 原方程组 (1.1.4) 的解使原方程 ①,②,③ 成为等式, 这些等式相减得到的 ④,⑤ 以及 ⑥ 仍然是等式. 这说明了原方程组 (1.1.4) 的解一定是新方程组 (1.1.5) 的解. 反过来, 如果 (1.1.5) 中的方程 ①,④,⑥ 乘常数再相加能够反过来得到原方程 ②,③, 则 (1.1.5) 的解也是原方程组 (1.1.4) 的解, 就可以保证原方程组 (1.1.4) 与新方程组 (1.1.5) 同解.

由原方程组 (1.1.4) 直接变成 (1.1.5), 变化太大, 不容易看出 (1.1.5) 怎样变回 (1.1.4). 我们将这个大变化分成若干个小步骤来实现, 使每一小步得到的方程组容易变回前一个方程组.

任意方程组 U 中各方程分别乘常数再相加得到的新方程称为方程组 U 的**线性组合** (linear combination), 由 U 的若干个线性组合组成的方程组 W 也称为 U 的线性组合. 容易验证, 方程组 U 的解一定是它的线性组合 W 的解. 反过来, 如果 W 可以通过线性组合变回 U, 则方程组 W 的解也是 U 的解, $U \to W$ 是同解变形.

定理 1.1.1 对线性方程组 U 进行以下三类变形, 得到的新方程组 W 与原方程组 U 互为线性组合, $U \longrightarrow W$ 是同解变形.

(1) 将第 i 个方程与第 j 个方程互相交换位置: $U \xrightarrow{(i,j)} W$.

(2) 将第 i 个方程两边同乘非零常数 λ: $U \xrightarrow{\lambda(i)} W$.

(3) 将第 i 个方程的 λ 倍加到第 j 个方程: $U \xrightarrow{\lambda(i)+(j)} W$.

证明 易见 W 的每个方程都是 U 的线性组合.

反过来, W 可以通过同样类型的变形变回 U :

(1) $U \xrightarrow{(i,j)} W \xrightarrow{(i,j)} U$.

(2) $U \xrightarrow{\lambda(i)} W \xrightarrow{\lambda^{-1}(i)} U$.

(3) $U \xrightarrow{\lambda(i)+(j)} W \xrightarrow{-\lambda(i)+(j)} U$.

可见 U 也是 W 的线性组合.

这就证明了方程组 U 与 W 互为线性组合. 由 U 到 W 的变形是同解变形. □

定理 1.1.1 中所说的三类变形称为方程组的**初等变换** (elementary transformations).

例 2 解法 1 所求二次函数 $f(x)=ax^2+bx+c$ 的系数满足方程组

$$\begin{cases} a+b+c=1, \\ 4a+2b+c=2, \\ 9a+3b+c=0. \end{cases} \qquad (U)$$

利用初等变换将原方程组 (U) 变成三角形方程组.

$$(U)\xrightarrow{-(2)+(3)}\begin{cases} a+b+c=1, \\ 4a+2b+c=2, \\ 5a+b=-2 \end{cases}\xrightarrow{-(1)+(2)}\begin{cases} a+b+c=1, \\ 3a+b=1, \\ 5a+b=-2 \end{cases}$$

$$\xrightarrow{-(2)+(3)}\begin{cases} a+b+c=1, \\ 3a+b=1, \\ 2a=-3. \end{cases} \qquad (T)$$

从方程组 (T) 由下而上解出 $a=-\dfrac{3}{2}, b=\dfrac{11}{2}, c=-3$.

所求函数 $f(x)=-\dfrac{3}{2}x^2+\dfrac{11}{2}x-3$, $f(4)=-5$. □

例 2 的方程组化成三角形方程组 (T) 之后, 还可以用初等变换将 (T) 进一步化简, 使每个方程都只包含一个未知数, 这些未知数系数都是 1:

$$(T)\xrightarrow{-\frac{1}{2}(3)+(1),-\frac{3}{2}(3)+(2)}\begin{cases} b+c=\dfrac{5}{2}, \\ b=\dfrac{11}{2}, \\ 2a=-3 \end{cases}$$

$$\xrightarrow{-(2)+(1)}\begin{cases} c=-3, \\ b=\dfrac{11}{2}, \\ 2a=-3 \end{cases}\xrightarrow{\frac{1}{2}(3)}\begin{cases} c=-3, \\ b=\dfrac{11}{2}, \\ a=-\dfrac{3}{2}, \end{cases} \qquad (\Lambda)$$

最后的方程组 (A) 的解可以直接写出, 为 $\left(-\dfrac{3}{2},\dfrac{11}{2},-3\right)$.

例 2 的以上的解法是从右到左先消去 c 再消去 b, 消去了方程组左边右下角的各项, 得到的三角形方程组 (T) 的三角形摆在左上角, 与例 1 的方程组的三角形的左右方向相反. 也可以按从左到右的顺序消元, 先消去 a, 再消去 b, 化简后的三角形方程组的三角形在右上角, 与例 1 的方程组的形状相同. 具体过程如下:

例 2 解法 2　所求二次函数 $f(x)=ax^2+bx+c$ 的系数满足方程组

$$\begin{cases} a+\ \ b+c\ =1,\\ 4a+2b+c\ =2,\\ 9a+3b+c\ =0.\end{cases}\qquad (U)$$

通过初等变换将 (U) 化简得

$$(U)\xrightarrow{-4(1)+(2),-9(1)+(3)}\begin{cases} a+\ b+\ c=1,\\ \quad -2b-3c=-2,\\ \quad -6b-8c=-9\end{cases}\xrightarrow{-3(2)+(3)}\begin{cases} a+\ b+\ c=1,\\ \quad -2b-3c=-2,\\ \qquad\qquad c=-3\end{cases}$$

$$\xrightarrow{3(3)+(2),-(3)+(1)}\begin{cases} a+b\qquad =4,\\ \ -2b\qquad =-11,\\ \qquad\quad c=-3\end{cases}\xrightarrow{\frac{1}{2}(2)+(1),-\frac{1}{2}(2)}\begin{cases} a\qquad\qquad =-\dfrac{3}{2},\\ \quad b\qquad =\dfrac{11}{2},\\ \qquad\ c\ =-3.\end{cases}$$

直接写出方程组解 $\left(-\dfrac{3}{2},\dfrac{11}{2},-3\right)$. 仍得 $f(x)=-\dfrac{3}{2}x^2+\dfrac{11}{2}x-3$, $f(4)=-5$.　□

例 3　在一次智力测验中, 老师写出某个数列的前两项 1, 2, 让学生按照前两项的规律写出第 3 项. 有的学生写 3, 有的学生写 4, 老师都判为正确. 有一个学生给的答案是 0, 老师判为错误.

试给出某个数列的通项公式使这个数列的前 3 项依次是 1, 2, 0, 来说明这个学生的答案也是正确的. 并按照这个通项公式写出第 4 项.

分析　与例 2 同样可求出二次函数 $f(x)=-\dfrac{3}{2}x^2+\dfrac{11}{2}x-3$ 满足 $f(1)=$

$1, f(2)=2, f(3)=0$. 以 $a_n=f(n)=-\dfrac{3}{2}n^2+\dfrac{11}{2}n-3$ 为通项公式的数列的前 3 项就是 $1,2,0$. 第 4 项为 $f(4)=-5$. □

习题 1.1

1. (1) 求 $1^3+2^3+\cdots+n^3$;　　(2) 求 $1^5+2^5+\cdots+n^5$.

2. 求二次函数 $y=f(x)$, 使其具有如下对应值

x	2	3	4
y	7	16	29

3. (1) 证明: 任给 3 个数 y_1,y_2,y_3, 存在函数 $f(n)=an^2+bn+c$, 使以 $a_n=f(n)$ 为通项公式的数列的前 3 项为 y_1,y_2,y_3.

(2) 在一次智力测验中, 老师给了一个数列的前 3 项 $1,2,3$, 让学生填写第 4 项. 试证明: 无论在第 4 项填上什么数 y, 都存在一个函数 $f(n)=an^3+bn^2+cn+d$, 使以 $a_n=f(n)$ 为通项公式的数列的前 4 项为 $1,2,3,y$.

4. 试用初等变换解线性方程组

(1) $$\begin{cases} x_1+2x_2+3x_3=1,\\ 2x_1+2x_2+5x_3=2,\\ 3x_1+5x_2+\ \ x_3=3;\end{cases}$$

(2) $$\begin{cases} \qquad\ x_2+x_3+x_4=1,\\ x_1\qquad\ +x_3+x_4=2,\\ x_1+x_2\qquad\ +x_4=3,\\ x_1+x_2+x_3\qquad\ =4.\end{cases}$$

5. 方程组 U 经过初等变换 $U\xrightarrow{-(1)+(2),-(1)+(3)}U_1\xrightarrow{-3(2)+(3)}U_2\xrightarrow{-\frac{1}{8}(3)}W$ 化成

$$\begin{cases} x+y+z=0,\\ \qquad y+3z=-1,\\ \qquad\qquad z=-1,\end{cases}\qquad (W)$$

写出方程组 U, 并求出它的解.

§1.2 矩阵消元法

1.2.1 矩阵的初等行变换

考察 §1.1 例 2 的方程组通过初等变换化简的过程可以发现, 实际上只有各项系数参加了加、减、乘、除运算, 表示未知数的字母并没有参加运算. 字母起的作用只是用来辨认哪些系数是同类项系数, 可以将它们合并起来.

既然如此, 将代表未知数的字母反复抄写是一种不必要的累赘. 为了书写的简便, 更为了突出解方程组中本质的东西——系数的运算, 我们采用分离系

数法, 将线性方程组中代表未知数的字母略去, 将等号也略去, 只剩下各项系数及常数项, 将 §1.1 例 2 的线性方程组

$$\begin{cases} a+\ \ b+c=1, \\ 4a+2b+c=2, \\ 9a+3b+c=0 \end{cases}$$

简写为一个矩形数表

$$\boldsymbol{M}=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 2 & 1 & 2 \\ 9 & 3 & 1 & 0 \end{pmatrix}$$

这样的矩形数表称为矩阵. 其中每一行由相应方程的各项系数组成, 代表这个方程; 同一个未知数的系数上下对齐组成一列, 最后一列由常数项组成.

一般地, 我们有:

定义 1.2.1(矩阵) 对任意的正整数 m,k, 由 mk 个数排成的 m 行 k 列的矩形数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix}$$

称为一个 $m\times k$ **矩阵** (matrix), 数表中每一个数 a_{ij} 称为这个矩阵的一个**元素** (element), 也称为矩阵的一个**分量** (entry). 矩阵的第 i 行第 j 列的元素 a_{ij} 称为这个矩阵的第 (i,j) 元, 也称为第 (i,j) 分量. 特别, $1\times k$ 矩阵 $(a_1,a_2,\cdots,a_k)$ 只有一行, 称为 k 维**行矩阵** (row matrix), 也称 k 维**行向量** (row vector). 类似地, $m\times 1$ 矩阵 $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$ 只有一列, 称为 m 维**列矩阵** (column matrix), 也称 m 维**列向量** (column vector). 当 $m=k$ 时, $m\times m$ 矩阵是一个正方形的数表, 称为 m 阶**方阵** (square matrix).

元素全为 0 的矩阵称为**零矩阵** (zero matrix), 通常记为 $\boldsymbol{O}$. 元素全为 0 的行向量与列向量都称为**零向量** (zero vector), 记为 $\mathbf{0}$. □

这样, 任意 n 元线性方程 $a_1x_1+\cdots+a_nx_n=b$ 就由 $n+1$ 维行向量

$(a_1,\cdots,a_n,b)$ 表示, 线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \tag{1.2.1}$$

由 $m\times(n+1)$ 矩阵

$$\boldsymbol{M}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

表示. 矩阵 $\boldsymbol{M}$ 的每一行 $(a_{i1},\cdots,a_{in},b_i)$ 表示一个方程 $a_{i1}x_1+\cdots+a_{in}x_n=b_i$. 方程组左边的全体未知数系数组成 $\boldsymbol{M}$ 的前 n 列, 所有这些系数组成一个 $m\times n$ 矩阵 $\boldsymbol{A}$, 称为这个线性方程组的**系数矩阵** (coefficient matrix), 其中每一列 $\boldsymbol{a}_j$ 由同一个未知数 x_j 在各方程中的系数 a_{ij} 组成. $\boldsymbol{M}$ 的最后一列 $\boldsymbol{b}$ 由各方程中的常数项 b_i 组成. 这样, $\boldsymbol{M}$ 由方程组的系数矩阵 $\boldsymbol{A}$ 增添由常数项组成的一列 $\boldsymbol{b}$ 得到, 因此也将 $\boldsymbol{M}$ 称为线性方程组的**增广矩阵** (augmented matrix).

用矩阵表示方程组, 不仅是为了更简单地表示方程组, 更重要的是为了简便地表示解方程组的过程. 将方程组用矩阵表示, 方程组的初等变换就变成对矩阵的各行进行相应的变换.

将两个方程 $a_1x_1+\cdots+a_nx_n=b$ 与 $a'_1x_1+\cdots+a'_nx_n=b'$ 相加, 就是将它们的同类项系数相加得到新方程 $(a_1+a'_1)x_1+\cdots+(a_n+a'_n)x_n=b+b'$. 将两个方程分别用行向量表示为 $(a_1,\cdots,a_n,b),(a'_1,\cdots,a'_n,b')$, 将方程相加也就是将这两个行向量的对应分量相加, 得到

$$(a_1,\cdots,a_n,b)+(a'_1,\cdots,a'_n,b')=(a_1+a'_1,\cdots,a_n+a'_n,b+b').$$

类似地, 将方程 $a_1x_1+\cdots+a_nx_n=b$ 乘常数 λ, 就是将 λ 乘各项系数得到新方程 $(\lambda a_1)x_1+\cdots+(\lambda a_n)x_n=\lambda b$, 也就是将表示这个方程的行向量的各分量同乘 λ:

$$\lambda(a_1,\cdots,a_n,b)=(\lambda a_1,\cdots,\lambda a_n,\lambda b).$$

行向量乘 0 得到零向量 $\mathbf{0}$, 它表示的方程就是恒等式 $0=0$.

方程组的线性组合, 相应的就是增广矩阵的各行的线性组合, 也就是各行的常数倍之和.

方程组的三类初等变换, 相应地可以用增广矩阵 $\boldsymbol{M}$ 的各行的如下三类初等变换来实现:

(1) 将第 i 行与第 j 行互相交换位置: $\boldsymbol{M} \xrightarrow{(i,j)} \boldsymbol{M}_1$.

(2) 将第 i 行乘非零常数 λ : $\boldsymbol{M} \xrightarrow{\lambda(i)} \boldsymbol{M}_1$.

(3) 将第 i 行的 λ 倍加到第 j 行: $\boldsymbol{M} \xrightarrow{\lambda(i)+(j)} \boldsymbol{M}_1$.

这三类变换称为矩阵的**初等行变换** (elementary transformations of rows) .

利用初等变换化简方程组的过程, 就变成对增广矩阵作初等行变换化简系数矩阵的过程. 例如, §1.1 例 2 解法 2 中的方程组 (U) 的同解变形过程用矩阵的初等行变换表示如下:

$$\boldsymbol{M}=\begin{pmatrix}1&1&1&1\\4&2&1&2\\9&3&1&0\end{pmatrix}\xrightarrow{-4(1)+(2),-9(1)+(3)}\begin{pmatrix}1&1&1&1\\0&-2&-3&-2\\0&-6&-8&-9\end{pmatrix}$$

$$\xrightarrow{-3(2)+(3)}\boldsymbol{T}=\begin{pmatrix}1&1&1&1\\0&-2&-3&-2\\0&0&1&-3\end{pmatrix}\xrightarrow{-(3)+(1),3(3)+(2)}\begin{pmatrix}1&1&0&4\\0&-2&0&-11\\0&0&1&-3\end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}(2)+(1)}\boldsymbol{D}=\begin{pmatrix}1&0&0&-\dfrac{3}{2}\\0&-2&0&-11\\0&0&1&-3\end{pmatrix}\xrightarrow{-\frac{1}{2}(2)}\boldsymbol{A}=\begin{pmatrix}1&0&0&-\dfrac{3}{2}\\0&1&0&\dfrac{11}{2}\\0&0&1&-3\end{pmatrix}.$$

§1.1 例 2 的方程组的未知数个数与方程个数都是 3, 增广矩阵 $\boldsymbol{M}$ 的前 3 列组成的系数矩阵 $\boldsymbol{A}$ 是方阵, 从左上角到右下角作一条对角线, 称为方阵 $\boldsymbol{A}$ 的主对角线. 对 $\boldsymbol{M}$ 作初等行变换先将系数方阵主对角线左下方元素全部变成 0, 增广矩阵化为 $\boldsymbol{T}$; 再将主对角线右上方的元素全部变成 0, 增广矩阵化为 $\boldsymbol{D}$; 再将系数方阵主对角线上元素全部变成 1, 增广矩阵化为 $\boldsymbol{A}$, 所代表的方程组

$$\begin{cases}a\qquad\quad=-\dfrac{3}{2},\\ \qquad b\quad=\dfrac{11}{2},\\ \qquad\quad c=-3\end{cases}$$

的解可直接写出为 $\left(-\dfrac{3}{2},\dfrac{11}{2},-3\right)$.

一般地, 每个 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 中的 n^2 个元素排成一个正方形数表. 从这个正方形左上角到右下角画一条对角线, 称为方阵 $\boldsymbol{A}$ 的**主对角线** (principal diagonal), 位于主对角线上的就是 $i=j$ 的 n 个元素 a_{ii} $(1\leqslant i\leqslant n)$, 称为**对角元** (diagonal elements). 如果方程 $\boldsymbol{A}$ 的主对角线左下方所有的元素 a_{ij} $(i>j)$ 全部为 0, 这样的方阵就称为**上三角形矩阵** (upper triangle matrix). 如果方阵的主对角线右上方的元素 a_{ij} $(i<j)$ 全为 0, 这样的方阵就称为**下三角形矩阵** (lower triangle matrix). 如果 $\boldsymbol{A}$ 既是上三角形矩阵也是下三角形矩阵, 只有对角元有可能不为 0, 就称 $\boldsymbol{A}$ 为**对角矩阵** (diagonal matrix). 对角矩阵 $\boldsymbol{A}$ 可以只写出对角元 $a_{11},\cdots,a_{nn}$ 来表示, 记为 $\mathrm{diag}(a_{11},\cdots,a_{nn})$. 特别地, 对角元全为 1 的对角矩阵

$$\mathrm{diag}(1,\cdots,1)=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&1\end{pmatrix}$$

称为**单位矩阵** (identity matrix 或 unit matrix) , 记为 $\boldsymbol{I}$ (有的书籍和文章也记为 $\boldsymbol{E}$).

以上对 §1.1 例 2 的方程组的增广矩阵 $\boldsymbol{M}$ 作初等行变换的过程, 就是将 $\boldsymbol{M}$ 的前 3 列组成的系数方阵 $\boldsymbol{A}$ 先变成上三角形矩阵、再变成对角矩阵、再变成单位矩阵的过程.

1.2.2 用矩阵的初等行变换解线性方程组

例 1 解方程组

$$\begin{cases}x_1+2x_2+3x_3+4x_4=-3,\\x_1+2x_2\qquad\quad-5x_4=1,\\3x_1-\ x_2-\ x_3\qquad\quad=1,\\x_1\qquad\ +\ x_3+2x_4=-1.\end{cases}$$

解 将方程组用增广矩阵表示为

$$\boldsymbol{M}=\begin{pmatrix}1&2&3&4&-3\\1&2&0&-5&1\\3&-1&-1&0&1\\1&0&1&2&-1\end{pmatrix}.$$

注意, 原方程组中等于 1 或 0 的未知数系数都被省略了, 但在矩阵中不能省略.

方程组的系数矩阵仍是方阵. 我们仍希望先将系数方阵化为上三角形矩阵. 为此, 先将第 1 行的适当常数倍加到以下各行, 将这些行的第 1 列元素全部变成 0:

$$\boldsymbol{M} \xrightarrow{-(1)+(2),-3(1)+(3),-(1)+(4)} \boldsymbol{M}_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 0 & -3 & -9 & 4 \\ 0 & -7 & -10 & -12 & 10 \\ 0 & -2 & -2 & -2 & 2 \end{pmatrix}$$

$\boldsymbol{M}_1$ 第 2 至 4 行第 2 列元素不全为 0, 但第 2 行第 2 列元素为 0. 将第 4 行与第 2 行互换, 使第 2 行第 2 列元素变成非零的数, 以便将以下各行的第 2 列元素全变成 0. 互换后得到的第 2 行各元素有公因子 2, 乘 $-\dfrac{1}{2}$ 约去这个公因子:

$$\boldsymbol{M}_1 \xrightarrow{(2,4),-\frac{1}{2}(2)} \boldsymbol{M}_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & -7 & -10 & -12 & 10 \\ 0 & 0 & -3 & -9 & 4 \end{pmatrix}$$

依次将 $\boldsymbol{M}_2$ 的第 2 行与第 3 行的适当常数倍加到它们以下各行, 将系数方阵变成上三角形矩阵:

$$\boldsymbol{M}_2 \xrightarrow{7(2)+(3)} \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & -3 & -5 & 3 \\ 0 & 0 & -3 & -9 & 4 \end{pmatrix} \xrightarrow{-(3)+(4)} \boldsymbol{T} = \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & -3 & -5 & 3 \\ 0 & 0 & 0 & -4 & 1 \end{pmatrix}.$$

由下而上依次将 $\boldsymbol{T}$ 的各行的适当常数倍加到以上各行, 将系数方阵变成对角矩阵, 再将各行乘适当常数将系数方阵变成单位矩阵:

$$\boldsymbol{T} \xrightarrow{(4)+(1),\frac{1}{4}(4)+(2),-\frac{5}{4}(4)+(3)} \begin{pmatrix} 1 & 2 & 3 & 0 & -2 \\ 0 & 1 & 1 & 0 & -\dfrac{3}{4} \\ 0 & 0 & -3 & 0 & \dfrac{7}{4} \\ 0 & 0 & 0 & -4 & 1 \end{pmatrix}$$

$$\xrightarrow{-\frac{1}{3}(3),-3(3)+(1),-(3)+(2),-\frac{1}{4}(4)}\begin{pmatrix}1&2&0&0&-\frac{1}{4}\\0&1&0&0&-\frac{1}{6}\\0&0&1&0&-\frac{7}{12}\\0&0&0&1&-\frac{1}{4}\end{pmatrix}\xrightarrow{-2(2)+(1)}\begin{pmatrix}1&0&0&0&\frac{1}{12}\\0&1&0&0&-\frac{1}{6}\\0&0&1&0&-\frac{7}{12}\\0&0&0&1&-\frac{1}{4}\end{pmatrix}$$

最后一个矩阵代表的方程为

$$\begin{cases}x_1 = \dfrac{1}{12},\\ x_2 = -\dfrac{1}{6},\\ x_3 = -\dfrac{7}{12},\\ x_4 = -\dfrac{1}{4}.\end{cases}$$

它的解显然是 $\left(\dfrac{1}{12},-\dfrac{1}{6},-\dfrac{7}{12},-\dfrac{1}{4}\right)$. □

前面的例题中的线性方程组的系数矩阵都是方阵, 都可以通过初等行变换化成单位矩阵, 方程组都有唯一解. 但是, 并非所有的线性方程组都是如此.

例 2 在空间直角坐标系中, 描述下列线性方程组的图像.

(1) $x+y+z=2$; (2) $\begin{cases}x+y+\ \ z=2,\\2x+y+5z=0.\end{cases}$

解 方程组的图像中的点的坐标 (x,y,z) 都是实数. 本题所求的图像就是以方程组的实数解为坐标的点的全体组成的集合.

(1) 原方程即 $x=2-y-z$. 将 y,z 分别取任意实数值 t_1,t_2 代入, 就得到一个解 $(x,y,z)=(2-t_1-t_2,t_1,t_2)$, 解集合 $\varPi_1=\{(2-t_1-t_2,t_1,t_2)\mid t_1,t_2\in\mathbf{R}\}$. 为了便于理解 $\varPi_1$ 中所有的点 $P(2-t_1-t_2,t_1,t_2)$ $(t_1,t_2\in\mathbf{R})$ 组成的图像的形状, 我们将 P 的坐标 (也就是向量 $\overrightarrow{OP}$ 的坐标) 写成线性组合的形式

$$\begin{pmatrix}2-t_1-t_2\\t_1\\t_2\end{pmatrix}=\begin{pmatrix}2\\0\\0\end{pmatrix}+t_1\begin{pmatrix}-1\\1\\0\end{pmatrix}+t_2\begin{pmatrix}-1\\0\\1\end{pmatrix},$$

以上表达式说明: 从图像上一个点 $P_0(2,0,0)$ 到任意点 P 的有向线段表示的向量 $\overrightarrow{P_0P}$ 取遍两个给定向量 $\boldsymbol{v}_1=(-1,1,0)$ 与 $\boldsymbol{v}_2=(-1,0,1)$ 的全体实系数线性组合. 这样的 P 组成的图像 Π_1 就是过 P_0 且与 $\boldsymbol{v}_1,\boldsymbol{v}_2$ 平行的平面. 取 $\overrightarrow{P_0P_1}=\boldsymbol{v}_1,\overrightarrow{P_0P_2}=\boldsymbol{v}_2$, 则 Π_1 就是过 3 点 P_0,P_1,P_2 的平面, 其中 $P_0=(2,0,0)$, $P_1(1,1,0)$, $P_2(1,0,1)$ 是分别取 $(t_1,t_2)=(0,0),(1,0)$ 和 $(0,1)$ 得到的图像上的三个点.

(2) 将方程组用矩阵表示, 再进行初等行变换, 得

$$\begin{pmatrix}1&1&1&2\\2&1&5&0\end{pmatrix}\xrightarrow{-2(1)+(2)}\begin{pmatrix}1&1&1&2\\0&-1&3&-4\end{pmatrix}\xrightarrow{(2)+(1),-1(2)}\begin{pmatrix}1&0&4&-2\\0&1&-3&4\end{pmatrix}.$$

最后一个矩阵所代表的方程组为

$$\begin{cases}x\quad+4z=-2,\\ \quad y-3z=4,\end{cases}\qquad 即 \qquad \begin{cases}x=-2-4z,\\ y=\quad 4+3z.\end{cases}$$

将 z 取任意一个实数值 t 代入, 得到一个解 $(-4t-2,3t+4,t)$. 解集合 $L=\{(-2-4t,4+3t,t)\mid t\in\mathbf{R}\}$. 图像中任意一点 P 的坐标

$$\begin{pmatrix}-2-4t\\4+3t\\t\end{pmatrix}=\begin{pmatrix}-2\\4\\0\end{pmatrix}+t\begin{pmatrix}-4\\3\\1\end{pmatrix}$$

以上表达式说明: 从图像上给定点 $P_0(-2,4,0)$ 到任一点 P 的有向线段表示的向量 $\overrightarrow{P_0P}$ 取遍给定向量 $\boldsymbol{v}_1=(-4,3,1)$ 的全体实数倍, 这样的 P 组成的图像是过 P_0 且与 $\boldsymbol{v}_1$ 平行的直线. 取 $\overrightarrow{P_0P_1}=\boldsymbol{v}_1$, 则 L 就是直线 P_0P_1, 其中 $P_0(-2,4,0),P_1(-6,7,1)$ 分别是取 $t=0$ 和 1 得到的图像上的两点. □

本节例 2 两个方程组的图像 (平面, 直线) 都不是一个点而有无穷多个点, 方程组都不止一个解而有无穷多个解. 方程组全部解的表达式称为这个方程组的**通解** (general solution), 其中的一个解称为方程组的**特解** (special solution). 本节例 2(1) 的 P_0 的坐标 $(2,0,0)$ 就是方程组的一个特解, 而 $(2-t_1-t_2,t_1,t_2)$ (其中 $t_1,t_2\in\mathbf{R}$) 是通解. 本节例 2 中的通解都含有可以任意取值的参数, 如 (1) 中的 t_1,t_2, (2) 中的 t. 这些参数都是点的坐标, 只能取实数值. 通解的每个分量都是这些参数的实系数一次多项式, 让参数取遍实数值, 就得到方程组的全部实数解.

注　我们将本节例 2 的图像上任一点 P 的坐标不写成行向量而写成列向量的形式, 是为了将通解中同一参数的系数上下对齐, 使分解式看起来更加清楚和自然.

例 3 解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+\ 4x_4=-3, \\ x_1+2x_2 \qquad\quad -\ 5x_4=1, \\ 2x_1+4x_2-3x_3-19x_4=6, \\ 3x_1+6x_2-3x_3-24x_4=7. \end{cases}$$

解

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 1 & 2 & 0 & -5 & 1 \\ 2 & 4 & -3 & -19 & 6 \\ 3 & 6 & -3 & -24 & 7 \end{pmatrix} \xrightarrow{-(1)+(2),-2(1)+(3),-3(1)+(4)} \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 0 & -3 & -9 & 4 \\ 0 & 0 & -9 & -27 & 12 \\ 0 & 0 & -12 & -36 & 16 \end{pmatrix}$$

$$\xrightarrow{-3(2)+(3),-4(2)+(4),(2)+(1),-\frac{1}{3}(2)} \begin{pmatrix} 1 & 2 & 0 & -5 & 1 \\ 0 & 0 & 1 & 3 & -\dfrac{4}{3} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

将最后这个矩阵写成方程组的形式. 注意, 这个矩阵的最后两行 $(0,0,0,0,0)$ 所代表的方程都是恒等式 $0=0$, 可以从方程组中删去而不会改变方程组的解. 方程组为

$$\begin{cases} x_1+2x_2 \quad -5x_4=1, \\ \qquad\qquad x_3+3x_4=-\dfrac{4}{3}. \end{cases} \tag{1.2.2}$$

未知数 x_1 只含于第 1 个方程, 未知数 x_3 只含于第 2 个方程. 将两个方程中含其余两个未知数 x_2,x_4 的项移到等号右边, 方程组化为

$$\begin{cases} x_1 \ = \ 1-2x_2+5x_4, \\ x_3 \ = \ -\dfrac{4}{3} \qquad -3x_4. \end{cases} \tag{1.2.3}$$

等号右边的未知数 x_2,x_4 可以任意取值. 让 x_2 取定任意值 t_1, x_4 取定任意值 t_2, 则由 (1.2.3) 的两式分别可算出 x_1,x_3 的值, 得到一组解. 让 x_2,x_4 取遍所有的可能的值 t_1,t_2, 就得到方程组的通解

$$\{(1-2t_1+5t_2,t_1,-\frac{4}{3}-3t_2,t_2)\mid t_1,t_2\text{在允许范围内任意取值}\}. \qquad \square$$

同样可以将通解写成线性组合的形式

$$\begin{pmatrix} 1-2t_1+5t_2 \\ t_1 \\ -\dfrac{4}{3} \qquad -3t_2 \\ t_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -\dfrac{4}{3} \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 5 \\ 0 \\ -3 \\ 1 \end{pmatrix}$$

解集合可以想象为在 4 维空间中过给定点 $P_0\left(1,0,-\frac{4}{3},0\right)$ 且平行于两个给定向量 $\boldsymbol{v}_1=(-2,1,0,0)$ 与 $\boldsymbol{v}_2=(5,0,-3,1)$ 的“平面”.

例 4 解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+4x_4=0, \\ x_1+2x_2 \qquad\quad -5x_4=0, \\ 2x_1+4x_2-3x_3-19x_4=0, \\ 3x_1+6x_2-3x_3-24x_4=0. \end{cases}$$

解 这个方程组与本节例 3 的方程组左边完全相同, 右边的常数项全改成了 0. 方程组的矩阵 $\boldsymbol{M}$ 的最后一列全是 0, 经过初等行变换之后也仍然全是 0. 可以将这一列略去不写, 直接用系数矩阵 $\boldsymbol{A}$ (由 $\boldsymbol{M}$ 的前 4 列组成) 代表方程组, 经过与例 3 完全相同的初等行变换:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 0 & -5 \\ 2 & 4 & -3 & -19 \\ 3 & 6 & -3 & -24 \end{pmatrix} \xrightarrow{-(1)+(2),-2(1)+(3),-3(1)+(4)} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & -3 & -9 \\ 0 & 0 & -9 & -27 \\ 0 & 0 & -12 & -36 \end{pmatrix}$$

$$\xrightarrow{-3(2)+(3),-4(2)+(4),(2)+(1),-\frac{1}{3}(2)} \begin{pmatrix} 1 & 2 & 0 & -5 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

得到化简后的方程组为

$$\begin{cases} x_1+2x_2 \quad -5x_4=0, \\ \qquad\qquad x_3+3x_4=0. \end{cases}$$

解之得

$$\begin{cases} x_1 = -2x_2 + 5x_4, \\ x_3 = \qquad\quad -3x_4. \end{cases}$$

通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2t_1 + 5t_2 \\ t_1 \\ -3t_2 \\ t_2 \end{pmatrix} = t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 5 \\ 0 \\ -3 \\ 1 \end{pmatrix} \qquad \square$$

与例 3 的通解相比较, 将例 3 的通解中的常数项改成 0 就得到例 4 的解. 例 4 的解同样可以想象成平行于两个给定向量 $\boldsymbol{v}_1 = (-2, 1, 0, 0)$ 与 $\boldsymbol{v}_2 = (5, 0, -3, 1)$ 的 "平面". 只不过例 3 的 "平面" 过点 $P_0\left(1, 0, -\frac{4}{3}, 0\right)$, 例 4 的 "平面" 过原点 $(0, 0, 0, 0)$. 可以认为: 将例 4 的 "图像" Π_0 平移, 使 "原点" $O(0, 0, 0, 0)$ 移到点 P_0, 则 Π_0 就移到例 3 的图像. 也就是说: 将例 4 的所有解都加上例 3 的一个解就得到例 3 的所有解.

像例 4 这样常数项全为 0 的线性方程组称为**齐次线性方程组** (homogeneous linear equations). 齐次线性方程组的增广矩阵 $\boldsymbol{M}$ 的最后一列全是 0, 经过初等行变换之后这一列也始终全是 0. 因此, 可以将 $\boldsymbol{M}$ 的最后一列略去不写, 直接用系数矩阵 $\boldsymbol{A}$ 代表齐次线性方程组. 不需解方程组就知道齐次线性方程组至少有一个解 $(0, \cdots, 0)$, 称为**零解** (zero solution), 也称**平凡解** (trivial solution). 是否还有其他的解, 总共有多少解, 都由它的系数矩阵决定. 通过初等行变换将系数矩阵 $\boldsymbol{A}$ 化成最简单的形式, 就可以得到齐次线性方程组的全部解.

如果线性方程组的常数项不全为 0, 就称为**非齐次线性方程组** (inhomogeneous linear equations). 除了本节例 4 之外, 本章前面所有例题中的方程组都是非齐次线性方程组. 非齐次线性方程组不一定有解. 例如, 方程组

$$\begin{cases} x + y = 0, \\ x + y = 1 \end{cases}$$

显然没有解. 因为, 对 x, y 的任何一组值, $x + y$ 不可能同时等于 0 与 1.

例 5 解线性方程组

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = -3, \\ x_1 + 2x_2 \qquad\quad - 5x_4 = 1, \\ 2x_1 + 4x_2 - 3x_3 - 19x_4 = 6, \\ 3x_1 + 6x_2 - 3x_3 - 24x_4 = 8. \end{cases}$$

解

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 1 & 2 & 0 & -5 & 1 \\ 2 & 4 & -3 & -19 & 6 \\ 3 & 6 & -3 & -24 & 8 \end{pmatrix} \xrightarrow{-(1)+(2),-2(1)+(3),-3(1)+(4)} \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 0 & -3 & -9 & 4 \\ 0 & 0 & -9 & -27 & 12 \\ 0 & 0 & -12 & -36 & 17 \end{pmatrix}$$

$$\xrightarrow{-3(2)+(3),-4(2)+(4)} \begin{pmatrix} 1 & 2 & 3 & 4 & -3 \\ 0 & 0 & -3 & -9 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

最后一个矩阵最后一行 $(0,0,0,0,1)$ 代表的方程 $0=1$ 无解. 因此原方程组无解. □

一般地, 如果将线性方程组化成最简形式之后, 出现了一个形如 $0=b$ 的方程, 其中 $b\neq 0$, 就说明原方程组无解.

1.2.3 线性方程组的系数范围, 数域

不难发现, 线性方程组的求解过程都是从原方程组各方程的系数出发作加、减、乘、除运算. 如果线性方程组有唯一解, 解的各分量都是由方程系数经过加、减、乘、除运算得到. 如果线性方程组有一个以上的解, 通解中含有可以任意取值的参数, 通解的各个分量都是这些参数的一次多项式, 这些一次多项式的系数由原方程系数经过加、减、乘、除运算得到.

实数经过加、减、乘、除得到的一定是实数. 因此, 实系数线性方程组如果有唯一解, 唯一解的各分量一定是实数; 如果有不止一个解, 通解的各分量一定是各参数的实系数一次多项式. 如果限制各参数在实数范围内任意取值, 得到的无穷多解的分量也一定都是实数.

有理数经过加、减、乘、除得到的一定是有理数. 有理系数线性方程组如果有唯一解, 唯一解的各分量一定是有理数; 如果有不止一个解, 通解的各分量一定是任意参数的有理系数多项式, 如果限制任意参数在有理数范围内取值, 得到的无穷多解的分量也都是有理数.

整数经过加、减、乘、除得到的不一定是整数. 因此, 整系数线性方程组的唯一解不一定是整数, 有一个以上的解 (因而有无穷多解) 时通解的系数也不一定是整数. 即使将任意参数取整数值, 得到的解的分量也不一定是整数, 只能保证是有理数. 例如, 本节例 1 与例 3 中的线性方程组的系数都是整数, 但方程组解的分量不全是整数.

注 线性方程组的系数范围与解的取值范围的这种关系与一元二次方程不同. 由一元二次方程的系数求方程的根, 不仅是进行加、减、乘、除运算, 还要开平方. 即使是有理系数一元二次方程, 经过开平方运算得到的根有可能不是有理数甚至不是实数.

定义 1.2.2(数域) 设 F 是复数集合 $\mathbf{C}$ 的子集 F, 至少包含 0 和 1, 并且 F 中任意两个数 a,b 经过加、减、乘、除运算得到的和 $a+b$、差 $a-b$、积 ab、商 $\dfrac{a}{b}$ $(b\neq 0)$ 仍属于 F, 就称 F 为**数域** (number field) . □

例如, 有理数集合 $\mathbf{Q}$、实数集合 $\mathbf{R}$、复数集合 $\mathbf{C}$ 都是数域.

如果某个集合 S 中任意两个元素 a,b 经过某种运算得到的结果仍含于 S, 就称 S 对这个运算封闭. 按照这个术语, 我们可以说数域 F 对加、减、乘运算封闭, F 中非零的数组成的集合 F^* 对除法封闭. 由于作除法时除数不能为 0, 凡是包含 0 的集合都不可能对除法封闭. 不过, 为了叙述简单起见, 如果集合 S 中任何两个元素 a,b 当 $b\neq 0$ 时作除法的商 $\dfrac{a}{b}$ 都含于 S, 也可以称 S 对除法封闭, 同时注明除数不为 0. 这样, 也可以说数域对加、减、乘、除封闭 (作除法时除数不为 0).

只要 $\mathbf{C}$ 的子集 F 包含一个非零的数 a, 并且对加、减、乘、除封闭 (作除法时除数不为 0), F 就必然包含 $a-a=0$ 及 $\dfrac{a}{a}=1$. 将 1 不断作加法可以得到所有的正整数, 作减法可以得到 0 与所有的负整数, 作除法可以得到所有的有理数. 这说明任何一个数域 F 至少包含全体有理数.

非零整数的集合对除法不封闭, 因此整数集合 $\mathbf{Z}$ 不是数域. 但是, 整数集合 $\mathbf{Z}$ 对加、减、乘运算封闭, 我们称它是数环. 一般地, 如果复数集合 $\mathbf{C}$ 的子集 D 对加、减、乘运算封闭, 就称 D 为**数环** (number ring) .

以后在讨论线性方程组的求解时, 都将系数范围限制在某个预先给定的数域 F 内, 并不一定求出所有的复数解. 在实际应用中, 很多时候都在实数域 $\mathbf{R}$

范围内解方程组.

既然线性方程组用矩阵表示. 很自然我们将矩阵的元素也限制在某个数域 F 内, 称为数域 F 上的矩阵. 数域 F 上的全体 $m\times n$ 矩阵组成的集合记为 $F^{m\times n}$. 特别, $F^{1\times n}$ 中的行向量 $(a_1,\cdots,a_n)$ 的分量都在 F 中, 这样的行向量称为数域 F 上的行向量. 类似地, $F^{m\times 1}$ 由 F 上的全体列向量组成. 用数域 F 上的行向量乘常数 λ 时, 我们要求 $\lambda\in F$. 对数域 F 上的矩阵作初等行变换 $\boldsymbol{M}\xrightarrow{\lambda(i)}\boldsymbol{M}_1$ 与 $\boldsymbol{M}\xrightarrow{\lambda(i)+(j)}\boldsymbol{M}_1$ 时, 要求所乘的常数 $\lambda\in F$.

习题 1.2

1. 用矩阵消元法解线性方程组, 并将其中 (1),(2) 的过程及结果与习题 1.1 第 4 题比较.

(1) $\begin{cases} x_1+2x_2+3x_3=1,\\ 2x_1+2x_2+5x_3=2,\\ 3x_1+5x_2+x_3=3;\end{cases}$　(2) $\begin{cases} x_2+x_3+x_4=1,\\ x_1+x_3+x_4=2,\\ x_1+x_2+x_4=3,\\ x_1+x_2+x_3=4;\end{cases}$

(3) $\begin{cases} 2x_1+x_2-5x_3+x_4=8,\\ x_1-3x_2-6x_4=9,\\ 2x_2-x_3+2x_4=-5,\\ x_1+4x_2-7x_3+6x_4=0;\end{cases}$　(4) $\begin{cases} x_1+3x_2-5x_3-5x_4=2,\\ x_1+2x_2+2x_3-2x_4+x_5=-2,\\ 2x_1+x_2+3x_3-3x_4=2,\\ x_1-4x_2+x_3+x_4-x_5=3,\\ x_1+3x_3-x_4+x_5=1.\end{cases}$

2. 在空间直角坐标系中, 求三个平面 $9x-3y+z=20$, $x+y+z=0$ 和 $-x+2y+z=-10$ 的公共点集合.

3. 已知两个变量 x,y 之间有某种函数关系 $y=f(x)$, 并且有如下对应值

x	1	2	3	4
y	2	7	16	29

问 y 是否可能是 x 的二次函数? 如果可能, 试求出满足要求的二次函数.

4. 二元一次方程组

$$\begin{cases} a_1x+b_1y=c_1,\\ a_2x+b_2y=c_2\end{cases}$$

的系数都是整数.

(1) 它的解是否一定是整数? 说明你的理由.

(2) 如果 $a_1b_2-a_2b_1=\pm1$, 证明以上整系数二元一次方程组的解一定是整数.

§1.3 线性方程组解集合的初步讨论

1.3.1 线性方程组求解过程总结

用 §1.2 介绍的矩阵消元法可以求解任何一个线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{1.3.1}$$

求解过程为

先将方程组用增广矩阵表示为

$$\boldsymbol{M} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} = (\boldsymbol{A}, \boldsymbol{b}),$$

其中 $\boldsymbol{A}$ 是方程组的系数矩阵, 由各方程的未知数系数组成, 组成 $\boldsymbol{M}$ 的前 n 列; $\boldsymbol{b}$ 是 $\boldsymbol{M}$ 的最后一列, 由各方程的常数项组成.

再对增广矩阵 $\boldsymbol{M} = (\boldsymbol{A}, \boldsymbol{b})$ 进行一系列初等行变换, 将系数矩阵 $\boldsymbol{A}$ 化为**阶梯形矩阵** (echelon matrix). 阶梯形矩阵 $\boldsymbol{T}$ 是指这样的矩阵: 如果 $\boldsymbol{T}$ 不为零, 每个非零行的上方没有零行, 且各非零行从左到右第一个非零元 $t_{1j_1}, t_{2j_2}, \cdots, t_{pj_p}$ 所在列的编号满足条件 $j_1 < j_2 < \cdots < j_p$.

为叙述方便, 我们将任何一个矩阵 $\boldsymbol{A}$ 中这样的元素 a_{pq} 称为阶梯元: $a_{pq} \neq 0$, 且 a_{pq} 的左方、下方、左下方所有元素 a_{ij} ($\forall\, i \geqslant p, j \leqslant q$ 且 $(i,j) \neq (p,q)$) 全都为 0 . 阶梯元 a_{pq} 所在位置 (p,q) 称为矩阵的一个阶梯. 按照这个说法, 非零矩阵 $\boldsymbol{A}$ 是阶梯形矩阵 $\Leftrightarrow$ $\boldsymbol{A}$ 的每个非零行的上方都没有零行, 且各非零行从左到右第一个非零元都是阶梯元.

如果阶梯形矩阵 $\boldsymbol{T}$ 的各阶梯元都等于 1, 各阶梯元所在列的其余元素都等于 0, 就称 $\boldsymbol{T}$ 为**最简阶梯形矩阵** (reduced echelon matrix).

我们约定, 元素全为 0 的矩阵也是最简阶梯形矩阵.

引理 1.3.1 设 $\boldsymbol{B} = (b_{ij})_{m\times n}$ 是数域 F 上任意 $m \times n$ 矩阵. 则

(1) $\boldsymbol{B}$ 能够通过有限次初等行变换化成最简阶梯形矩阵.

(2) $\boldsymbol{B}$ 能够通过有限次第三类初等变换化成阶梯形矩阵, 且各阶梯元所在列的其余元素都为 0.

证明 第一步 先将 $\boldsymbol{B}$ 化成阶梯形矩阵 $\boldsymbol{T}$.

如果 $\boldsymbol{B}$ 的所有元素都是 0, 已经是最简阶梯形矩阵.

设 $\boldsymbol{B}$ 的元素不全为 0, 我们对 $\boldsymbol{B}$ 进行一系列初等行变换, 从上到下依次在每一行构造一个阶梯, 直到得到一个阶梯形矩阵.

作为第一步, 我们先通过第三类初等行变换在 $\boldsymbol{B}$ 的第 1 行造出一个阶梯. 这个阶梯必然位于 $\boldsymbol{B}$ 中从左到右第 1 个非零列.

设 $\boldsymbol{B}$ 从左到右第一个不为零的列是第 j_1 列 $\boldsymbol{b}_{j_1}$, 即

$$\boldsymbol{B}=\begin{pmatrix} 0 & \cdots & 0 & b_{1j_1} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{mj_1} & \cdots & b_{mn} \end{pmatrix}$$

情况 1 $\boldsymbol{B}$ 的第 j_1 列第 1 行元素 $b_{1j_1}\neq 0$. 将 $\boldsymbol{B}$ 的第 1 行的适当常数倍加到其余各行:

$$\boldsymbol{B}\xrightarrow{-b_{ij_1}b_{1j_1}^{-1}(1)+(i),\forall 2\leqslant i\leqslant m}\boldsymbol{B}_1.$$

可将第 j_1 列第 2 至 m 行元素全部变成 0, 得到

$$\boldsymbol{B}_1=\begin{pmatrix} 0 & \cdots & 0 & b_{1j_1} & b_{1,j_1+1} & \cdots & b_{1n} \\ 0 & \cdots & 0 & 0 & c_{2,j_1+1} & \cdots & c_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & c_{m,j_1+1} & \cdots & c_{mn} \end{pmatrix}=\begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{C}_1 \end{pmatrix}. \tag{1.3.2}$$

它的第 1 行 $\boldsymbol{\beta}_1=(0,\cdots,0,b_{1j_1},\cdots,b_{1n})$ 的第 1 个非零元 $b_{1j_1}\neq 0$ 是 $\boldsymbol{B}_1$ 的阶梯元; $\boldsymbol{B}_1$ 的第 2 至 m 行组成的 $(m-1)\times n$ 矩阵 $\boldsymbol{C}_1$ 的前 j_1 列全是零.

情况 2 $b_{1j_1}=0$. 非零列 $\boldsymbol{b}_{j_1}$ 必有另一个元素 $b_{kj_1}\neq 0$. 将 $\boldsymbol{B}$ 的第 k 行加到第 1 行: $\boldsymbol{B}\xrightarrow{(k)+(1)}\tilde{\boldsymbol{B}}$, 得到的 $\tilde{\boldsymbol{B}}$ 的第 $(1,j_1)$ 元素不为 0. 用 $\tilde{\boldsymbol{B}}$ 代替 $\boldsymbol{B}$ 即化为情况 1, 仍可按情况 1 所说步骤得到 (1.3.2) 中的 $\boldsymbol{B}_1$.

如果不限定用第一类初等行变换, 可将 $\boldsymbol{B}$ 的第 k 行与第 1 行互换可化为 $b_{1j_1}\neq 0$ 的情形, 比用第三类初等行变换更简便 (解线性方程组时通常这样做).

这就证明了, 总可对非零矩阵 $\boldsymbol{B}$ 作一系列第三类初等变换将 $\boldsymbol{B}$ 变成 (1.3.2) 中 $\boldsymbol{B}_1$ 的形状, 使它的第 1 行有一个阶梯 $(1,j_1)$.

如果 $\boldsymbol{B}_1$ 只有 1 行, 或者第 2 至 m 行全为零, $\boldsymbol{B}_1$ 已经是阶梯形矩阵. 否则, 对 $\boldsymbol{B}_1$ 的第 2 至 m 行重复刚才对 $\boldsymbol{B}$ 的操作, 在第 2 行再造出一个阶梯. 重复这个过程, 从上到下依次在各行造出阶梯, 直到剩下的各行全是零, 或者在全部 m 行都造出阶梯, 就得到了阶梯形矩阵 $\boldsymbol{T}$.

第二步 由下至上将 $\boldsymbol{T}$ 的各非零行的适当常数倍加到以上各行, 可以将各阶梯元所在列的其余元素都变成 0, 得到矩阵 $\boldsymbol{T}_0$.

第三步 将 $\boldsymbol{T}_0$ 的各行乘适当非零常数可以将各阶梯元全部变成 1, 得到最简阶梯形矩阵. □

算法 1.3.1(解线性方程组) 将线性方程组用增广矩阵 $\boldsymbol{M}=(\boldsymbol{A},\boldsymbol{b})$ 表示.

按照引理 1.3.1 的步骤, 对 $\boldsymbol{M}$ 进行一系列初等行变换, 先将 $\boldsymbol{M}$ 的前 n 列化成阶梯形, 其中前 n 列只有前 r 行不全为 $0,t_{11},t_{2j_2},\cdots,t_{rj_r}$ 是阶梯元:

$$\boldsymbol{M}\to\cdots\to\boldsymbol{T}=\begin{pmatrix} t_{11} & \cdots & t_{1j_2} & \cdots & t_{1j_r} & \cdots & t_{1,n+1}\\ 0 & \cdots & t_{2j_2} & \cdots & t_{2j_r} & \cdots & t_{2,n+1}\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & t_{rj_r} & \cdots & t_{r,n+1}\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & 0 & \cdots & t_{m,n+1}\end{pmatrix}$$

注 $\boldsymbol{M}$ 的第一列的元素是各方程中未知数 x_1 的系数. 方程组中至少有一个方程含有未知数 x_1, 这个方程中 x_1 的系数不为 0, 这说明 $\boldsymbol{M}$ 的第一列不为零, $\boldsymbol{T}$ 的第一个阶梯在第 1 列.

情况 1 $\boldsymbol{T}$ 的最后 $m-r$ 行不全为 0, 至少有一个元素 $t_{k,n+1}\neq 0$, 其中 $k\geqslant r+1$. $\boldsymbol{T}$ 的第 k 行 $(0,\cdots,0,t_{k,n+1})$ 代表的方程 $0=t_{k,n+1}$ 无解. 原方程组无解.

情况 2 $\boldsymbol{T}$ 的最后 $m-r$ 行全为零, 此时 $\boldsymbol{T}$ 是 r 个阶梯的阶梯形矩阵. 再通过一系列初等行变换化为最简阶梯形矩阵

$$\boldsymbol{\Lambda}=(\lambda_{ij})_{m\times(n+1)}=\begin{pmatrix} \lambda_{11} & \cdots & 0 & \cdots & 0 & \cdots & \lambda_{1,n+1}\\ 0 & \cdots & \lambda_{2j_2} & \cdots & 0 & \cdots & \lambda_{2,n+1}\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & \lambda_{rj_r} & \cdots & \lambda_{r,n+1}\\ & & & & & & \boldsymbol{O}\end{pmatrix}$$

其中各阶梯元 $\lambda_{11}=\lambda_{2j_2}=\cdots=\lambda_{rj_r}=1$. $\boldsymbol{\Lambda}$ 所代表的方程组

$$\begin{cases} x_1+\cdots+\cdots+\cdots=\lambda_{1,n+1},\\ \qquad x_{j_2}+\cdots+\cdots=\lambda_{2,n+1},\\ \qquad\cdots\cdots\cdots\cdots\\ \qquad\qquad x_{j_r}+\cdots=\lambda_{r,n+1}\end{cases}\tag{1.3.3}$$

称为具有**最简形式**：其中每个方程各独占一个未知数 $x_1, x_{j_2}, \cdots, x_{j_r}$, 这些未知数只在所属的唯一的方程中出现而且系数为 1, 在其余方程中都不出现. 当 $r=n$ 时它们就是全部 n 个未知数 $x_1, x_2, \cdots, x_n$, 此时 $j_2, j_3, \cdots, j_n$ 分别等于 $2, 3, \cdots, n$. 当 $r<n$ 时, 除了这 r 个被独占的未知数之外还剩下另外 $n-r$ 个未知数, 称为独立未知数, 依次记为 $x_{j_{r+1}}, \cdots, x_{j_n}$.

最简形式的线性方程组 (1.3.3) 的解可以立即写出来:

情况 2.1 $r=n$. 此时方程组 (1.3.3) 具有形式

$$\begin{cases} x_1 = \lambda_{1,n+1}, \\ x_2 = \lambda_{2,n+1}, \\ \cdots\cdots\cdots\cdots \\ x_n = \lambda_{n,n+1}, \end{cases}$$

即有唯一解 $(\lambda_{1,n+1}, \lambda_{2,n+1}, \cdots, \lambda_{n,n+1})$.

情况 2.2 $r<n$. 此时方程组有 $n-r$ 个独立未知数 $x_{j_{r+1}}, \cdots, x_{j_n}$. 将各方程所独占的未知数的项留在左边, 其余各项 (含独立未知数的项) 移到右边, 得到

$$\begin{cases} x_{j_1} = \lambda_{1,n+1} - \lambda_{1,j_{r+1}} x_{j_{r+1}} - \cdots - \lambda_{1,j_n} x_{j_n}, \\ x_{j_2} = \lambda_{2,n+1} - \lambda_{2,j_{r+1}} x_{j_{r+1}} - \cdots - \lambda_{2,j_n} x_{j_n}, \\ \qquad\qquad \cdots\cdots\cdots\cdots \\ x_{j_r} = \lambda_{r,n+1} - \lambda_{r,j_{r+1}} x_{j_{r+1}} - \cdots - \lambda_{r,j_n} x_{j_n}. \end{cases}$$

右边的各独立未知数在允许范围 F 内任取一组值, 代入以上各等式得到方程组的一个解. 独立未知数取遍所有的允许值, 就得到方程组的通解, 每个分量都是独立未知数的一次多项式. □

1.3.2 对解集合的初步讨论

综上所述, 总结一般的线性方程组的求解过程:

当方程组有解且 $n=r$ 时, 方程组有唯一解.

当方程组有解且 $n>r$ 时, 方程组有无穷多解.

由此可见, n 与 r 是决定方程组的解集大小的两个关键性的数. 其中 n 是未知数的个数, 解方程组之前就知道. 而 r 却不是方程的个数 m, 而是当方程组有解而且系数矩阵 $\boldsymbol{A}$ 化成阶梯形矩阵之后的非零行的个数. 甚至 r 也不一定是方程组化成阶梯形方程组 (以阶梯形矩阵为系数矩阵的方程组) 之后的方

程个数: 当方程组无解时, 方程组化成阶梯形方程组后的方程个数大于 r; 当且仅当方程组有解时, 化成阶梯形方程组后的方程个数才是 r.

在解方程组之前, 不知道 r 等于多少, 也不知道方程组是否一定有解, 但可知两点:

(1) $r \leqslant m$.

(2) 齐次线性方程组 (常数项全为 0 的方程组) 至少有一个解 $(0,\cdots,0)$.

由此可得到

定理 1.3.1 如果方程个数 $m <$ 未知数个数 n, 则齐次线性方程组有无穷多解.

证明 此时有 $n > m \geqslant r$, 有 $n-r>0$ 个独立未知数可以任意取值, 因而有无穷多解. □

方程组是否有解, 既依赖于系数矩阵, 也依赖于常数项, 但有一种情况下不依赖于常数项.

定理 1.3.2 如果 $m=n=r$, 无论常数项 $b_1,\cdots,b_n$ 取什么值, 线性方程组都有唯一解.

证明 将增广矩阵化为阶梯形矩阵. 系数矩阵有 $r=n$ 个非零行, 因此全部 n 行都是非零行. 方程个数一开始就是 n, 化简后也不超过 n. 可见方程组一定有解. 且由 $n=r$ 知有唯一解. □

习题 1.3

1. a,b 取什么值时, 下面的方程组有解, 并求出其解:

$$\begin{cases} 3x_1+2x_2+ax_3+x_4-3x_5=4, \\ 5x_1+4x_2+3x_3+3x_4-x_5=3, \\ x_1+x_2+3x_3+2x_4+x_5=1, \\ x_2+2x_3+2x_4+6x_5=-3, \\ x_3+bx_4+x_5=1. \end{cases}$$

2. 讨论当 λ 取什么值时下面的方程组有解:

$$\begin{cases} \lambda x_1+x_2+x_3=1, \\ x_1+\lambda x_2+x_3=\lambda, \\ x_1+x_2+\lambda x_3=\lambda^2. \end{cases}$$

当方程组有解时求出解来, 并讨论 λ 取什么值时方程组有唯一解, 什么时候有无穷多组解.

3. 不解方程组, 判断下面的方程组是否有非零解:

(1) $\begin{cases} x+y+z=0, \\ 2x+y+5z=0; \end{cases}$ (2) $\begin{cases} x+y+z=0, \\ 2x+y+5z=0, \\ 3x+2y+6z=0. \end{cases}$

第2章 向量空间

§2.1 线性方程组的几何意义

2.1.1 线性方程组的唯一解问题

很多时候, 我们关注线性方程组是否有唯一解. 由第 1 章已经知道, 如果未知数个数 > 方程个数, 当方程组有解时一定有无穷多组解. 因此, 我们关注: 如果未知数个数 = 方程个数, 方程组是否一定有解? 有解时是否只有唯一解?

例 1 在平面上建立了直角坐标系, $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ 是任意 n 个横坐标不同的已知点. 是否存在唯一的实系数多项式 $f(x) = a_0 + a_1x + \cdots + a_mx^m$, 它的图像曲线经过这 n 个已知点?

分析 问题归结为: 是否存在唯一一组待定系数 $a_0, a_1, \cdots, a_m$ 满足条件

$$\begin{cases} a_0 + a_1x_1 + \cdots + a_mx_1^m = y_1, \\ a_0 + a_1x_2 + \cdots + a_mx_2^m = y_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_0 + a_1x_n + \cdots + a_mx_n^m = y_n. \end{cases}$$

这是一个以 $a_0, a_1, \cdots, a_m$ 为未知数的线性方程组. 未知数个数为 $m+1$, 方程个数为 n. 当 $m+1 > n$ 时, 即使有解也不唯一. 我们关心: 其余情况下何时有唯一解? 特别地, 当 $m+1 = n$ 即 $m = n-1$ 时, 上述方程组是否一定有唯一解? □

例 2 由 5 个电阻 R_1, R_2, R_3, R_4, R_5 组成如图 2–1 的电路. 设在两点 A, B 之间加电压 V. 求在 5 个电阻上流过的电流 I_1, I_2, I_3, I_4, I_5.

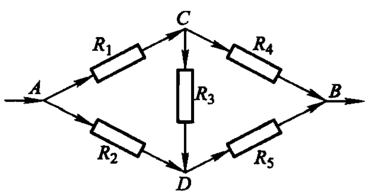

图 2–1

分析 在 C, D 两点流进与流出的电流的代数和为 0, 得到两个方程:

$$I_1 - I_3 - I_4 = 0,$$
$$I_2 + I_3 - I_5 = 0.$$

两条回路 $A \to C \to D \to A$ 与 $C \to B \to D \to C$ 的总电势差都为 0, 得到两个方程:

$$I_1R_1 - I_2R_2 + I_3R_3 = 0,$$
$$I_3R_3 - I_4R_4 + I_5R_5 = 0.$$

从 A 到 B 的总电压为 V, 得到一个方程

$$I_1R_1 + I_4R_4 = V.$$

以上 5 个方程组成线性方程组, 未知数 I_1, I_2, I_3, I_4, I_5 的个数为 5. 未知数个数=方程个数, 方程组是否一定有唯一解? □

例 1 与例 2 中的方程组, 如果已经知道方程系数的具体的数值, 可以对系数矩阵消元求解, 判断它是否有唯一解. 如果求解过程太复杂, 可以借助于计算机. 但是, 如果要对一般的系数讨论方程组是否有唯一解, 就需要对由字母组成的矩阵消元, 遇到除法就要对分母是否为 0 进行讨论. 即使可以用计算机中的符号运算进行消元, 例 1 中的方程个数 n 本身就是可以任意变动的值, 符号计算就永远无法结束. 本质上, 人工计算与计算机都只能进行有限次运算. 要得出适用于无限多种一般情况的普遍结论, 必须进行理论分析, 这是用手算和计算机运算都不能代替的.

2.1.2 二元一次方程组有唯一解的条件

例 3 试研究实系数二元一次方程组

$$\begin{cases} a_1x + b_1y = c_1, \\ a_2x + b_2y = c_2 \end{cases} \tag{2.1.1}$$

有唯一解的充分必要条件.

解 方程组 (2.1.1) 可以写成向量形式:

$$x\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + y\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \tag{2.1.2}$$

在平面直角坐标系中取点 $A(a_1, a_2), B(b_1, b_2), C(c_1, c_2)$, 如图 2–2. 则 (2.1.2) 就是

$$x\overrightarrow{OA} + y\overrightarrow{OB} = \overrightarrow{OC}. \tag{2.1.3}$$

其几何意义就是要将 $\overrightarrow{OC}$ 表示成 $\overrightarrow{OA},\overrightarrow{OB}$ 的线性组合, 求组合系数 x,y.

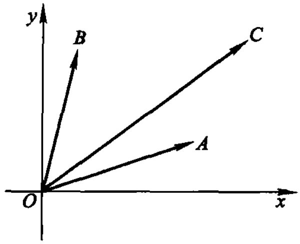

图 2–2

先设 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 不共线, 则它们组成平面上的一组基, 平面上每个向量 $\overrightarrow{OC}$ 在这组基下的坐标 (x,y) 就是方程组 (2.1.1) 的唯一解. 特别地, 零向量 $\mathbf{0}$ 的坐标 $(0,0)$ 是方程组

$$x\overrightarrow{OA}+y\overrightarrow{OB}=\mathbf{0} \tag{2.1.4}$$

的唯一解, 方程组 (2.1.4) 没有非零解.

再设 $\overrightarrow{OA},\overrightarrow{OB}$ 共线, 存在实数 λ 使 $\overrightarrow{OB}=\lambda\overrightarrow{OA}$ 或 $\overrightarrow{OA}=\lambda\overrightarrow{OB}$, 方程 (2.1.4) 有非零解 $(x_0,y_0)=(\lambda,-1)$ 或 $(x_0,y_0)=(-1,\lambda)$. 如果方程组 (2.1.1) 有解 (x_1,y_1), 将等式 $x_0\overrightarrow{OA}+y_0\overrightarrow{OB}=\mathbf{0}$ 乘任意实数 t 再与等式 $x_1\overrightarrow{OA}+y_1\overrightarrow{OB}=\overrightarrow{OC}$ 相加得

$$(x_1+tx_0)\overrightarrow{OA}+(y_1+ty_0)\overrightarrow{OB}=\overrightarrow{OC}.$$

这说明方程组 $x\overrightarrow{OA}+y\overrightarrow{OB}=\overrightarrow{OC}$ 有无穷多个解 (x_1+tx_0,y_1+ty_0) $(t\in\mathbf{R})$.

由此可知, 二元一次方程组 (2.1.1) 有唯一解的充分必要条件是: 以它的系数矩阵的两列为坐标的几何向量 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 不共线 $\Leftrightarrow$ (a_1,a_2) 与 (b_1,b_2) 不成比例 $\Leftrightarrow a_1b_2-a_2b_1\neq 0$. □

第 3 章中将看到,$\Delta=a_1b_2-a_2b_1$ 就是方程组 (2.1.1) 系数矩阵的行列式, 其绝对值 $|\Delta|$ 就是以 OA,OB 为邻边的平行四边形 $OADB$ 的面积, $\Delta\neq 0$ 当然是 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 不共线的充分必要条件.

2.1.3 三元一次方程组有唯一解的条件

例 4 试研究实系数三元一次方程组

$$\begin{cases} a_1x+b_1y+c_1z=d_1, \\ a_2x+b_2y+c_2z=d_2, \\ a_3x+b_3y+c_3z=d_3 \end{cases} \tag{2.1.5}$$

有唯一解的充分必要条件.

解 方程组 (2.1.5) 可以写成向量形式:

$$x\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}+y\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}+z\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix} \tag{2.1.6}$$

即

$$x\overrightarrow{OA}+y\overrightarrow{OB}+z\overrightarrow{OC}=\overrightarrow{OD}, \tag{2.1.7}$$

其中 $\overrightarrow{OA}=(a_1,a_2,a_3),\overrightarrow{OB}=(b_1,b_2,b_3),\overrightarrow{OC}=(c_1,c_2,c_3),\overrightarrow{OD}=(d_1,d_2,d_3)$ 是空间坐标系中的几何向量.

方程 (2.1.7) 的几何意义是: 将 $\overrightarrow{OD}$ 表示成 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 的线性组合, 求组合系数 x,y,z.

先设 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OB}$ 不共面, 组成空间的一组基, 空间每个向量 $\overrightarrow{OD}$ 在这组基下的坐标 (x,y,z) 就是方程组 (2.1.7) 的唯一解. 特别地, 零向量 $\mathbf{0}$ 的坐标 $(0,0,0)$ 是方程组

$$x\overrightarrow{OA}+y\overrightarrow{OB}+z\overrightarrow{OC}=\mathbf{0} \tag{2.1.8}$$

的唯一解, 方程组 (2.1.8) 没有非零解.

再设 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 共面, 其中一个向量是另外两个的实系数线性组合. 即: 存在实数 λ_1,λ_2 使 $\overrightarrow{OC}=\lambda_1\overrightarrow{OA}+\lambda_2\overrightarrow{OB}$ 或 $\overrightarrow{OB}=\lambda_1\overrightarrow{OA}+\lambda_2\overrightarrow{OC}$ 或 $\overrightarrow{OA}=\lambda_1\overrightarrow{OB}+\lambda_2\overrightarrow{OC}$, 方程组 (2.1.8) 有非零解 $(x_0,y_0,z_0)=(\lambda_1,\lambda_2,-1)$ 或 $(\lambda_1,-1,\lambda_2)$ 或 $(-1,\lambda_1,\lambda_2)$. 如果此时方程 $x\overrightarrow{OA}+y\overrightarrow{OB}+z\overrightarrow{OC}=\overrightarrow{OD}$ 有解 (x_1,y_1,z_1), 一定有无穷多解 $(x_1+tx_0,y_1+ty_0,z_1+tz_0)$ $(t\in\mathbf{R})$.

由此可知, 三元一次方程组 (2.1.5) 有唯一解的充分必要条件是: 以它的系数矩阵的 3 列为坐标的几何向量 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 不共面 $\Leftrightarrow$ 方程组 $x\overrightarrow{OA}+y\overrightarrow{OB}+z\overrightarrow{OC}=\mathbf{0}$ 即

$$\begin{cases}a_1x+b_1y+c_1z=0,\\a_2x+b_2y+c_2z=0,\\a_3x+b_3y+c_3z=0\end{cases}$$

只有唯一解 $(x,y,z)=(0,0,0)$ 而没有非零解. □

例 5 在平面上建立了直角坐标系,$(x_1,y_1),(x_2,y_2),(x_3,y_3)$ 是任意三个横坐标不同的已知点. 是否一定存在唯一的不超过二次的实系数多项式 $f(x)=a_0+a_1x+a_2x^2$, 它的图像曲线经过这三个已知点?

解 问题归结为: 待定系数 a_0,a_1,a_2 满足的线性方程组

$$\begin{cases} a_0+x_1a_1+x_1^2a_2=y_1, \\ a_0+x_2a_1+x_2^2a_2=y_2, \\ a_0+x_3a_1+x_3^2a_2=y_3 \end{cases} \tag{2.1.9}$$

是否一定有唯一解. 由例 4 的结论知: 方程组 (2.1.9) 有唯一解 $\Leftrightarrow$ 方程组

$$\begin{cases} a_0+x_1a_1+x_1^2a_2=0, \\ a_0+x_2a_1+x_2^2a_2=0, \\ a_0+x_3a_1+x_3^2a_2=0 \end{cases} \tag{2.1.10}$$

有唯一解 (0,0,0). 对 (2.1.10) 的系数矩阵 $\boldsymbol{M}$ 作初等行变换得

$$\boldsymbol{M}=\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix} \xrightarrow{-(1)+(2),-(1)+(3)} \begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & x_2-x_1 & x_2^2-x_1^2 \\ 0 & x_3-x_1 & x_3^2-x_1^2 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{x_2-x_1}(2),\frac{1}{x_3-x_1}(3)} \begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & 1 & x_2+x_1 \\ 0 & 1 & x_3+x_1 \end{pmatrix}$$

$$\xrightarrow{-(2)+(3),\frac{1}{x_3-x_2}(3)} \boldsymbol{T}=\begin{pmatrix} 1 & x_1 & x_1^2 \\ 0 & 1 & x_2+x_1 \\ 0 & 0 & 1 \end{pmatrix}.$$

以 $\boldsymbol{T}$ 为系数矩阵的齐次线性方程组只有唯一解 $(0,0,0)$,(2.1.10) 也只有唯一解 $(0,0,0)$. 因此方程组 (2.1.9) 总有唯一解. 满足要求的多项式 $f(x)=a_0+a_1x+a_2x^2$ 一定存在且唯一. □

本书第 3 章中将给出计算方程组 (2.1.5) 的系数矩阵的行列式 Δ 的公式, 其绝对值 $|\Delta|$ 就是以 OA,OB,OC 为棱的平行六面体的体积,$\Delta\neq 0$ 可以作为判定 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 不共面的充分必要条件.

习题 2.1

1. 已知平面直角坐标系中的三点 $A(2,3),B(3,4),C(10,5)$. 将几何向量 $\overrightarrow{OC}$ 表示成 $\overrightarrow{OA},\overrightarrow{OB}$ 的线性组合 $\overrightarrow{OC}=x\overrightarrow{OA}+y\overrightarrow{OB}$, 求系数 x,y.

2. 已知空间直角坐标系中的四点 $A(1,1,1)$, $B(1,2,3)$, $C(1,4,9)$, $D(1,8,27)$, 根据线性方程组

$$x\overrightarrow{AB}+y\overrightarrow{AC}+z\overrightarrow{AD}=\mathbf{0}$$

是否有非零解 (x,y,z) 判定这四点是否在同一个平面上.

3. (1) 将如下方程组写成向量形式并求方程组的通解.

$$\begin{cases} x+\ \ y+\ \ z=0, \\ x+2y+4z=0, \\ x+3y+9z=0. \end{cases}$$

(2) 已知空间直角坐标系中的三点 $A(1,1,1),B(1,2,4),C(1,3,9)$, 是否能将空间任意一个向量 $\overrightarrow{OD}$ 写成 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 的线性组合? 以 OA,OB,OC 为棱的平行六面体的体积是否等于 0 ?

§2.2 线性相关与线性无关

2.2.1 向量定义的推广

现在来研究任意数域 F 上的 n 元一次方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases} \tag{2.2.1}$$

有唯一解的条件.

与 $n=2,3$ 的情况类似, 将方程组 (2.2.1) 写成向量形式

$$\boldsymbol{x_1a_1+x_2a_2+\cdots+x_na_n=b}, \tag{2.2.2}$$

其中每个 $\boldsymbol{a}_j$ $(1\leqslant j\leqslant n)$ 是由各方程中同一个未知数 x_j 的系数 a_{ij} $(1\leqslant i\leqslant n)$ 组成的列向量, $\boldsymbol{b}$ 是由各方程的常数项组成的列向量:

$$\boldsymbol{a}_j=\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix},\ \forall\, 1\leqslant j\leqslant n;\qquad \boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

当 $n=2$ 或 3 且各系数 a_{ij} 及 b_i 是实数时, 我们将各 $\boldsymbol{a}_j$ 与 $\boldsymbol{b}$ 看作几何空间中的向量, 由几何知识知道, 当各 $\boldsymbol{a}_j$ 不共线 (当 $n=2$) 或不共面 (当 $n=3$) 时, 它们组成几何空间 (平面或空间) 的一组基, 空间中任何一个向量 $\boldsymbol{b}$ 可以唯

一地用这组基的线性组合表示出来, 线性组合式中的系数 $x_1,\cdots,x_n$ 组成 $\boldsymbol{b}$ 在这组基下的坐标, 也就是方程 (2.2.1) 的唯一解.

将任意数域 F 上的 n 维有序数组 $(a_1,\cdots,a_n)$ 也看成向量, 将这些数组的全体组成的集合 F^n 看成向量空间, 称为 **n 维数组空间** . 虽然 F^n 中的向量不能画出图来表示, 也不能利用几何知识来研究它们的性质, 但是, 它们有与几何向量类似的代数运算: 向量的加法, 向量与数域 F 中的数的乘法.

n 维数组空间 F^n 中的向量的加法:

F^n 中任意两个向量 $\boldsymbol{a}=(a_1,\cdots,a_n),\boldsymbol{b}=(b_1,\cdots,b_n)$ 可以按分量对应相加得到

$$\boldsymbol{a}+\boldsymbol{b}=(a_1+b_1,\cdots,a_n+b_n)\in F^n.$$

n 维数组空间 F^n 中的向量的数乘:

F^n 中任意向量 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 可以与任意 $\lambda\in F$ 相乘得到

$$\lambda\boldsymbol{a}=(\lambda a_1,\cdots,\lambda a_n)\in F^n.$$

这样定义的向量加法与数乘运算满足我们熟悉的运算律:

(1) **加法交换律**: $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$ 对任意 $\boldsymbol{a},\boldsymbol{b}\in F^n$ 成立.

(2) **加法结合律**: $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$ 对任意 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in F^n$ 成立.

(3) **存在零向量**: $\boldsymbol{0}=(0,\cdots,0)\in F^n$ 对任意 $\boldsymbol{a}\in F^n$ 满足条件 $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{0}+\boldsymbol{a}$.

(4) **存在负向量**: 对每个 $\boldsymbol{a}=(a_1,\cdots,a_n)\in F^n$, 存在负向量 $-\boldsymbol{a}=(-a_1,\cdots,-a_n)\in F^n$ 满足 $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$.

(5) **数乘对于数的加法的分配律**: $(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$ 对任意 $\lambda,\mu\in F$ 及 $\boldsymbol{a}\in F^n$ 成立.

(6) **数乘对于向量加法的分配律**: $\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$ 对任意 $\lambda\in F$ 及 $\boldsymbol{a},\boldsymbol{b}\in F^n$ 成立.

(7) **数乘结合律**: $\lambda(\mu\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$ 对任意 $\lambda,\mu\in F$ 及 $\boldsymbol{a}\in F^n$ 成立.

(8) **1 乘向量**: $1\boldsymbol{a}=\boldsymbol{a}$ 对任意 $\boldsymbol{a}\in F^n$ 成立.

对于数的运算和几何向量的运算, 这些运算律都是我们已经熟悉的. 但既然 F^n 中的向量是与几何向量不同的新概念, F^n 中的向量的运算也是与几何向量运算不同的新运算, 对这些运算律的叙述与验证就是必需的, 以便在运算时合法地应用. 我们知道, 欧几里得几何体系是由少数公理推出很多结论. 代数的体系则是由少数的运算律推出很多的结论, 运算律就是代数的公理.

n 维数组空间 F^n 中向量组的线性组合:

F^n 中任意一组向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 分别乘 F 中的常数 $\lambda_1,\cdots,\lambda_m$ 再相加, 得到的向量

$$\boldsymbol{b}=\lambda_1\boldsymbol{a}_1+\cdots+\lambda_m\boldsymbol{a}_m$$

称为这组向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 的**线性组合**, 也称为向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合. 由 S 的若干个线性组合 $\boldsymbol{b}_i=\lambda_{i1}\boldsymbol{a}_1+\cdots+\lambda_{im}\boldsymbol{a}_m$ 组成的向量组 T 也称为 S 的线性组合.

F^n 中的 n 维数组向量可以写成一行的形式 $(a_1,\cdots,a_n)$, 称为 n 维行向量. F 上全体 n 维行向量组成的空间记为 $F^{1\times n}$, 称为 F 上 n 维行空间.

也可以将 n 维数组写成一列的形式

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

称为 n 维列向量. F 上全体 n 维列向量组成的空间记为 $F^{n\times 1}$, 称为 F 上 n 维列空间.

当 $n=mk$ 时, 还可以将 n 维数组写成 m 行 k 列的矩阵的形式

$$(a_{ij})_{m\times k}=\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} \end{pmatrix}$$

将 F 上全体 $m\times k$ 矩阵组成的集合 $F^{m\times k}$ 看作 F 上的 mk 维数组空间, 将两个 $m\times k$ 矩阵 $\boldsymbol{A}=(a_{ij})_{m\times k}$ 与 $\boldsymbol{B}=(b_{ij})_{m\times k}$ 的对应分量相加得到这两个矩阵的和 $\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times k}\in F^{m\times k}$, 将 $\lambda\in F$ 与 $\boldsymbol{A}$ 的每个分量相乘得到 λ 与 $\boldsymbol{A}$ 的积 $\lambda\boldsymbol{A}=(\lambda a_{ij})_{m\times k}\in F^{m\times k}$.

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} \end{pmatrix}+\begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mk} \end{pmatrix}$$

$$=\begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1k}+b_{1k} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mk}+b_{mk} \end{pmatrix}$$

$$\lambda\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1k} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mk} \end{pmatrix}$$

行空间 $F^{1\times n}$ 与列空间 $F^{n\times 1}$ 分别是矩阵空间 $F^{m\times k}$ 当 $m=1$ 或 $k=1$ 时的特殊情形.

不论将 F^n 写成行空间、列空间还是矩阵空间, 如果只关心其中的加法及数乘这两种运算, 这些空间实质上都是相同的, 由这两种运算所导出的定义和性质对它们都同样成立.

将 $F^{1\times n}$ 中的每个行向量 $\boldsymbol{\alpha}=(a_1,\cdots,a_n)$ 中的元素按原来的顺序从上到下依次排列, 得到的列向量 $\boldsymbol{a}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ 称为 $\boldsymbol{\alpha}$ 的转置, 记为 $\boldsymbol{\alpha}^{\mathrm{T}}$. 反过来, 将列向量 $\boldsymbol{a}\in F^{n\times 1}$ 中各元素按原来顺序从左到右得到的行向量 $\boldsymbol{\alpha}$ 也称为 $\boldsymbol{a}$ 的转置, 记为 $\boldsymbol{a}^{\mathrm{T}}$. 更一般地, 对 $F^{m\times k}$ 中的任意矩阵 $\boldsymbol{A}=(a_{ij})_{m\times k}$, 依次以 $\boldsymbol{A}$ 的各行 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 的转置 $\boldsymbol{\alpha}_1^{\mathrm{T}},\cdots,\boldsymbol{\alpha}_m^{\mathrm{T}}$ 为各列排成的 $k\times m$ 矩阵记作 $\boldsymbol{A}^{\mathrm{T}}$, 称为 $\boldsymbol{A}$ **的转置** (transpose) :

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{pmatrix}, \quad \boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} a_{11} & \cdots & a_{m1} \\ a_{12} & \cdots & a_{m2} \\ \vdots & & \vdots \\ a_{1k} & \cdots & a_{mk} \end{pmatrix},$$

则 $\boldsymbol{A}$ 的第 (i,j) 元素 a_{ij} 成为 $\boldsymbol{A}^{\mathrm{T}}$ 的第 (j,i) 元素, $\boldsymbol{A}$ 的第 i 行 $\boldsymbol{\alpha}_i$ 的转置 $\boldsymbol{\alpha}_i^{\mathrm{T}}$ 成为 $\boldsymbol{A}^{\mathrm{T}}$ 的第 i 列, $\boldsymbol{A}$ 的第 i 列 $\boldsymbol{a}_i$ 的转置 $\boldsymbol{a}_i^{\mathrm{T}}$ 成为 $\boldsymbol{A}^{\mathrm{T}}$ 的第 i 行.

由同样的元素 $a_1,\cdots,a_n$ 排成的列向量 $\boldsymbol{a}$ 比行向量 $\boldsymbol{\alpha}$ 在书写和排版时占的篇幅更大. 因此, 我们常将列向量 $\boldsymbol{a}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ 写成 $(a_1,\cdots,a_n)^{\mathrm{T}}$ 的形式, 以节省排版空间.

2.2.2 线性相关与线性无关

几何向量的共线与共面难以用几何语言推广到 n 维数组向量, 但可以通过代数运算来推广.

平面向量 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 共线, 可以用向量运算描述为: $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=\boldsymbol{0}$ 有非零解 (x_1,x_2).

空间向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 共面, 可以用向量运算描述为: $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+x_3\boldsymbol{a}_3=\boldsymbol{0}$ 有非零解 (x_1,x_2,x_3).

一般地, 有:

定义 2.2.1(线性相关与线性无关) 对 F^n 中 k 个向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k\in F^n$, 如果存在不全为 0 的数 $x_1,\cdots,x_k$ 满足条件 $x_1\boldsymbol{a}_1+\cdots+x_k\boldsymbol{a}_k=\boldsymbol{0}$, 就称 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ **线性相关** (linearly dependent) .

反过来, 如果 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 不是线性相关, 就称它们**线性无关** (linearly independent) , 也就是说: $x_1\boldsymbol{a}_1+\cdots+x_k\boldsymbol{a}_k=\boldsymbol{0}$ 仅当 $x_1=\cdots=x_k=0$ 时才成立. □

几何向量的共线与共面还可以描述为: 其中某个向量可以写成其余向量的线性组合.

类似地, 一般的向量的线性相关也可以这样描述:

定理 2.2.1 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关 $\Leftrightarrow$ 其中某个向量 $\boldsymbol{a}_i$ 可以写成其余向量 $\boldsymbol{a}_j\ (j\neq i)$ 的线性组合 $\Leftrightarrow$ 其中某个向量 $\boldsymbol{a}_i$ 可以写成它前面的向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1}$ 的线性组合.

证明 先设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关, 即存在不全为 0 的 $x_1,\cdots,x_k\in F$ 使

$$x_1\boldsymbol{a}_1+\cdots+x_k\boldsymbol{a}_k=\boldsymbol{0}.$$

设 $x_i\neq 0$, 将其余各项 $x_j\boldsymbol{a}_j\ (j\neq i)$ 移到等式右边, 两边同除以 x_i 得

$$\boldsymbol{a}_i=-\frac{x_1}{x_i}\boldsymbol{a}_1-\cdots-\frac{x_{i-1}}{x_i}\boldsymbol{a}_{i-1}-\frac{x_{i+1}}{x_i}\boldsymbol{a}_{i+1}-\cdots-\frac{x_k}{x_i}\boldsymbol{a}_k,$$

$\boldsymbol{a}_i$ 就被表示成了其余 $\boldsymbol{a}_j\ (j\neq i)$ 的线性组合.

当 $x_1,\cdots,x_k$ 不全为 0 时, 可以选择 x_i 是其中最后一个不为 0 的数, 于是

$$x_1\boldsymbol{a}_1+\cdots+x_i\boldsymbol{a}_i=\boldsymbol{0}\quad\Rightarrow\quad \boldsymbol{a}_i=-\frac{x_1}{x_i}\boldsymbol{a}_1-\cdots-\frac{x_{i-1}}{x_i}\boldsymbol{a}_{i-1},$$

$\boldsymbol{a}_i$ 是它前面的 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1}$ 的线性组合.

反过来, 一开始就设某个 $\boldsymbol{a}_i$ 是其余 $\boldsymbol{a}_j$ 的线性组合:

$$\boldsymbol{a}_i=\lambda_1\boldsymbol{a}_1+\cdots+\lambda_{i-1}\boldsymbol{a}_{i-1}+\lambda_{i+1}\boldsymbol{a}_{i+1}+\cdots+\lambda_k\boldsymbol{a}_k,$$

则

$$\lambda_1\boldsymbol{a}_1+\cdots+\lambda_{i-1}\boldsymbol{a}_{i-1}+(-1)\boldsymbol{a}_i+\lambda_{i+1}\boldsymbol{a}_{i+1}+\cdots+\lambda_n\boldsymbol{a}_k=\boldsymbol{0},$$

其中 $-1\neq 0$. 因而 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关. □

推论 2.2.1 如果向量组 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 中 $\boldsymbol{a}_1\neq\mathbf{0}$, 且每一个向量 $\boldsymbol{a}_i$ $(2\leqslant i\leqslant k)$ 都不是它前面的向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1}$ 的线性组合, 则 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性无关. □

例如, 在三维几何向量的空间中, 取非零向量 $\boldsymbol{a}_1$, 取 $\boldsymbol{a}_2$ 不在 $\boldsymbol{a}_1$ 所决定的直线上, 再取 $\boldsymbol{a}_3$ 不在 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 所决定的平面上, 则 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关, 且组成三维空间的一组基.

例 1 一个向量组成的向量组 $S=\{\boldsymbol{a}\}$ 何时线性相关, 何时线性无关?

解 按照定义, S 线性相关 $\Leftrightarrow$ 存在 $\lambda\neq 0$ 使 $\lambda\boldsymbol{a}=\mathbf{0}\Leftrightarrow\boldsymbol{a}=\mathbf{0}$.

因此, 零向量 $\mathbf{0}$ 线性相关, 非零向量 $\boldsymbol{a}$ 线性无关. □

引理 2.2.1 向量集合 S 如果包含一部分线性相关的向量, 则 S 线性相关.

证明 设 S 中一部分向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关, S 中其余向量为 $\boldsymbol{a}_{k+1},\cdots,\boldsymbol{a}_m$. 则 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k\}$ 线性相关 $\Rightarrow$ 存在不全为 0 的 $\lambda_1,\cdots,\lambda_k$ 使 $\lambda_1\boldsymbol{a}_1+\cdots+\lambda_k\boldsymbol{a}_k=\mathbf{0}\Rightarrow\lambda_1\boldsymbol{a}_1+\cdots+\lambda_k\boldsymbol{a}_k+0\boldsymbol{a}_{k+1}+\cdots+0\boldsymbol{a}_m=\mathbf{0}$, 其中 $\lambda_1,\cdots,\lambda_k,0,\cdots,0$ 不全为 $0\Rightarrow S$ 线性相关. □

2.2.3 通过解方程组判定线性相关

设

$$\boldsymbol{a}_j=\begin{pmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{nj}\end{pmatrix}\in F^n,\quad 1\leqslant j\leqslant k.$$

则 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关的充分必要条件是: 方程

$$x_1\boldsymbol{a}_1+\cdots+x_k\boldsymbol{a}_k=\mathbf{0}\tag{2.2.3}$$

有非零解 $(x_1,\cdots,x_k)$. 也就是: 齐次线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1k}x_k=0,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2k}x_k=0,\\\cdots\cdots\cdots\cdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nk}x_k=0\end{cases}\tag{2.2.4}$$

有非零解 $(x_1,\cdots,x_k)$.

反过来, $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性无关的充分必要条件则是: 齐次线性方程组 (2.2.4) 只有唯一解 $(0,\cdots,0)$.

给定了各个向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 的坐标 a_{ij} 的具体数值, 就可以通过解齐次线性方程组 (2.2.4) 来判定这些向量是线性相关还是线性无关.

例 2 在空间直角坐标系中, 判断如下 4 点 A,B,C,D 是否在同一平面上:

(1) $A(1,2,3)$, $B(2,3,4)$, $C(3,3,8)$, $D(2,-1,16)$;

(2) $A(1,2,3)$, $B(2,3,4)$, $C(3,3,8)$, $D(2,-1,7)$.

解 (1) $\overrightarrow{AB}=(2,3,4)-(1,2,3)=(1,1,1)$, $\overrightarrow{AC}=(3,3,8)-(1,2,3)=(2,1,5)$, $\overrightarrow{AD}=(2,-1,16)-(1,2,3)=(1,-3,13)$.

A,B,C,D 共面 $\Leftrightarrow$ $\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}$ 共面 $\Leftrightarrow$ $\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}$ 线性相关

$\Leftrightarrow$ 方程组 $x_1\overrightarrow{AB}+x_2\overrightarrow{AC}+x_3\overrightarrow{AD}=\mathbf{0}$ 有非零解 (x_1,x_2,x_3).

此方程组即

$$\begin{cases}x_1+2x_2+x_3=0,\\ x_1+x_2-3x_3=0,\\ x_1+5x_2+13x_3=0,\end{cases}$$

经过计算求得通解 (计算过程略) $(x_1,x_2,x_3)=t(7,-4,1)$, 当 $t\neq 0$ 时就得到非零解. 例如, 取 $t=1$ 得到非零解 $(7,-4,1)$, 说明

$$7\overrightarrow{AB}-4\overrightarrow{AC}+\overrightarrow{AD}=\mathbf{0}.$$

可见 $\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}$ 线性相关. 从而 A,B,C,D 共面.

(2) $\overrightarrow{AB}=(1,1,1),\overrightarrow{AC}=(2,1,5),\overrightarrow{AD}=(1,-3,4)$. 解方程组

$$x_1\overrightarrow{AB}+x_2\overrightarrow{AC}+x_3\overrightarrow{AD}=\mathbf{0},$$

即

$$\begin{cases}x_1+2x_2+x_3=0,\\ x_1+x_2-3x_3=0,\\ x_1+5x_2+4x_3=0,\end{cases}$$

得唯一解 $(0,0,0)$, 这说明 $\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}$ 线性无关. 从而 A,B,C,D 不共面. □

例 3 当 b_1,b_2,b_3 取什么实数值时, 如下方程组有唯一解:

$$(1)\ \begin{cases}x+2y+z=b_1,\\ x+y-3z=b_2,\\ x+5y+13z=b_3;\end{cases}\qquad (2)\ \begin{cases}x+2y+z=b_1,\\ x+y-3z=b_2,\\ x+5y+4z=b_3.\end{cases}$$

解 (1) 原方程组可以写成向量形式

$$x\boldsymbol{a}_1+y\boldsymbol{a}_2+z\boldsymbol{a}_3=\boldsymbol{b},$$

其中

$$\boldsymbol{a}_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \boldsymbol{a}_2=\begin{pmatrix}2\\1\\5\end{pmatrix},\quad \boldsymbol{a}_3=\begin{pmatrix}1\\-3\\13\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}$$

解齐次线性方程组 (见本节例 2(1))

$$\begin{cases}x+2y+\ \ z=0,\\x+\ \ y-\ \ 3z=0,\\x+5y+13z=0,\end{cases}$$

知它有非零解. 也就是

$$x\boldsymbol{a}_1+y\boldsymbol{a}_2+z\boldsymbol{a}_3=\boldsymbol{0}$$

有非零解 (从而有无穷多解), 向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性相关, 它们所代表的几何向量在同一平面上. 当 $\boldsymbol{b}$ 不在这个平面上时, 原方程组无解; 当 $\boldsymbol{b}$ 在这个平面上时, 原方程组有无穷多解. 无论 b_1,b_2,b_3 取什么实数值, 方程组都没有唯一解.

(2) 原方程组即

$$x\boldsymbol{a}_1+y\boldsymbol{a}_2+z\boldsymbol{a}_3=\boldsymbol{b},$$

其中

$$\boldsymbol{a}_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \boldsymbol{a}_2=\begin{pmatrix}2\\1\\5\end{pmatrix},\quad \boldsymbol{a}_3=\begin{pmatrix}1\\-3\\4\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix},$$

解齐次线性方程组 (见本节例 2(2))

$$\begin{cases}x+2y+\ z=0,\\x+\ \ y-3z=0,\\x+5y+4z=0,\end{cases}$$

知它只有唯一解 $(0,0,0)$. 也就是说

$$x\boldsymbol{a}_1+y\boldsymbol{a}_2+z\boldsymbol{a}_3=\boldsymbol{0}$$

只有唯一解 $(0,0,0)$, 向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关, 它们组成三维几何空间的一组基, 空间每一个向量 $\boldsymbol{b}$ 都能唯一地写成这组基的线性组合. 无论 b_1,b_2,b_3 取什么实数值, 方程组都有唯一解. □

例 4 设 $\boldsymbol{A} = (a_{ij})_{m\times n} \in F^{m\times n}$. $\boldsymbol{A}_0$ 是 $\boldsymbol{A}$ 的子矩阵, 即由 $\boldsymbol{A}$ 的第 $i_1,\cdots,i_p$ 行与第 $j_1,\cdots,j_q$ 列交叉位置的元素组成的 $p\times q$ 矩阵. 求证: 如果 $\boldsymbol{A}_0$ 的各行线性无关, 则将它的各行扩充得到的 $\boldsymbol{A}$ 的第 $i_1,\cdots,i_p$ 行线性无关. 如果 $\boldsymbol{A}_0$ 的各列线性无关, 则将它的各列扩充得到的 $\boldsymbol{A}$ 的第 $j_1,\cdots,j_q$ 列线性无关.

证明 先证明: $\boldsymbol{A}_0$ 的各列线性无关 $\Rightarrow$ $\boldsymbol{A}$ 的第 $j_1,\cdots,j_q$ 列线性无关.

记 $\boldsymbol{A}$ 的第 $j_1,\cdots,j_q$ 列分别为 $\boldsymbol{a}_{j_1},\cdots,\boldsymbol{a}_{j_q}$. 考虑以 $x_1,\cdots,x_q$ 为未知数的方程组

$$x_1\boldsymbol{a}_{j_1}+\cdots+x_q\boldsymbol{a}_{j_q}=\boldsymbol{0}, \tag{2.2.5}$$

即

$$\begin{cases} a_{1j_1}x_1+\cdots+a_{1j_q}x_q=0,\\ \qquad\cdots\cdots\cdots\cdots\\ a_{mj_1}x_1+\cdots+a_{mj_q}x_q=0. \end{cases} \tag{2.2.6}$$

已知 A_0 的各列 $\tilde{\boldsymbol{a}}_{j_1},\cdots,\tilde{\boldsymbol{a}}_{j_q}$ 线性无关, 也就是说: 方程组

$$x_1\tilde{\boldsymbol{a}}_{j_1}+\cdots+x_q\tilde{\boldsymbol{a}}_{j_q}=\boldsymbol{0}, \tag{2.2.7}$$

即

$$\begin{cases} a_{i_1j_1}x_1+\cdots+a_{i_1j_q}x_q=0,\\ \qquad\cdots\cdots\cdots\cdots\\ a_{i_pj_1}x_1+\cdots+a_{i_pj_q}x_q=0. \end{cases} \tag{2.2.8}$$

只有唯一解 $(0,\cdots,0)$. 方程组 (2.2.8) 由 (2.2.6) 的一部分方程组成, (2.2.6) 的每个解都必须是 (2.2.8) 的解, 只能是 $(0,\cdots,0)$. 这就证明了 $\boldsymbol{A}$ 的第 $j_1,\cdots,j_q$ 列线性无关.

现在设 $\boldsymbol{A}_0$ 的各行线性无关, 则 $\boldsymbol{A}_0^{\mathrm{T}}$ 的各列线性无关. $\boldsymbol{A}_0$ 的转置矩阵 $\boldsymbol{A}_0^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的子矩阵, 按照刚才所证, $\boldsymbol{A}_0^{\mathrm{T}}$ 的线性无关列向量扩充得到的 $\boldsymbol{A}^{\mathrm{T}}$ 的第 $i_1,\cdots,i_p$ 列线性无关. 也就是: $\boldsymbol{A}$ 的第 $i_1,\cdots,i_p$ 行线性无关 □

例 4 的结论可以作为一个引理供以后应用:

引理 2.2.2 如果 $\boldsymbol{A}$ 的子矩阵 $\boldsymbol{A}_0$ 的各行 (列) 线性无关, 则由 $\boldsymbol{A}_0$ 的这些行 (列) 扩充得到的 $\boldsymbol{A}$ 的行 (列) 线性无关. □

例 5 已知数域 F 上的向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关. 试判断 $\boldsymbol{a}_1+\boldsymbol{a}_2,\boldsymbol{a}_2+\boldsymbol{a}_3,\boldsymbol{a}_3+\boldsymbol{a}_1$ 是否线性无关?

解　设 $\lambda_1,\lambda_2,\lambda_3\in F$ 满足 $\lambda_1(\boldsymbol{a}_1+\boldsymbol{a}_2)+\lambda_2(\boldsymbol{a}_2+\boldsymbol{a}_3)+\lambda_3(\boldsymbol{a}_3+\boldsymbol{a}_1)=\mathbf{0}$. 整理得

$$(\lambda_1+\lambda_3)\boldsymbol{a}_1+(\lambda_1+\lambda_2)\boldsymbol{a}_2+(\lambda_2+\lambda_3)\boldsymbol{a}_3=\mathbf{0}. \tag{2.2.9}$$

由于 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 线性无关, (2.2.9) 成立仅当

$$\begin{cases}\lambda_1+\lambda_3=0,\\ \lambda_1+\lambda_2=0,\\ \lambda_2+\lambda_3=0.\end{cases} \tag{2.2.10}$$

解方程组 (2.2.10) 得唯一解 $(\lambda_1,\lambda_2,\lambda_3)=(0,0,0)$. 因此 $\boldsymbol{a}_1+\boldsymbol{a}_2,\boldsymbol{a}_2+\boldsymbol{a}_3,\boldsymbol{a}_3+\boldsymbol{a}_1$ 线性无关. □

习题　2.2

1. 判定 $\mathbf{R}^3$ 中的下述向量是线性相关还是线性无关:

(1) $\boldsymbol{\alpha}_1=(1,1,1)$, $\boldsymbol{\alpha}_2=(1,2,3)$, $\boldsymbol{\alpha}_3=(1,4,9)$;

(2) $\boldsymbol{\alpha}_1=(1,1,1)$, $\boldsymbol{\alpha}_2=(1,2,3)$, $\boldsymbol{\alpha}_3=(1,4,9)$, $\boldsymbol{\alpha}_4=(1,8,27)$.

2. 判定 $\mathbf{R}^4$ 中的下述向量是线性相关还是线性无关:

(1) $\boldsymbol{\alpha}_1=(2,0,-1,2)$, $\boldsymbol{\alpha}_2=(0,-2,1,-3)$, $\boldsymbol{\alpha}_3=(3,-1,2,1)$, $\boldsymbol{\alpha}_4=(-2,4,-7,5)$;

(2) $\boldsymbol{\alpha}_1=(1,-1,0,0)$, $\boldsymbol{\alpha}_2=(0,1,-1,0)$, $\boldsymbol{\alpha}_3=(0,0,1,-1)$, $\boldsymbol{\alpha}_4=(-1,0,0,1)$.

3. 设在 3 维几何空间中建立了直角坐标系. 判定如下 4 点是否共面:

(1) $A(1,1,1)$, $B(1,2,3)$, $C(1,4,9)$, $D(1,8,27)$;

(2) $A(1,1,1)$, $B(1,2,3)$, $C(2,5,8)$, $D(3,7,15)$.

4. 设复数域上的向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 线性无关. λ 取什么复数值时, 向量 $\boldsymbol{\alpha}_1-\lambda\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2-\lambda\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\alpha}_{n-1}-\lambda\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n-\lambda\boldsymbol{\alpha}_1$ 线性无关?

附录 1　关于向量定义与线性相关的进一步说明

1. 向量定义的进一步推广

将 F^n 中的数组称为向量并不是因为它们有方向和大小, 而是因为它们可以作加法和数乘运算. 同样的理由, 可以将很多别的集合 V 中的元素称为向量. 只要集合 V 中定义了某种加法与数乘, 使得 V 中任何两个元素 $\boldsymbol{u},\boldsymbol{v}$ 可以按照所定义的加法法则相加得到一个和 $\boldsymbol{u}+\boldsymbol{v}\in V$, V 中任何一个元素 $\boldsymbol{u}$ 可以按照所定义的数乘法则与数域 F 中每个数 λ 相乘得到一个乘积 $\lambda\boldsymbol{u}\in V$, 就可以将 V 中的元素称为向量, V 称为数域 F 上的向量空间. 我们不限制 V 由什么样的元素组成, 也不限制 V 中的加法与数乘怎样进行, 但是, 既然称为加法与数乘, 就必须满足前面所说的 8 条运算律, 否则就不是合法的加法与数乘.

由于数组向量的运算最简单, 我们经常将别的向量空间 V 中的向量 $\boldsymbol{u}$ 也都用数组向量 $\boldsymbol{a}$ 来代表, 将 V 中向量的运算用数组向量的运算来代表, 这样的数组 $\boldsymbol{a}$ 称为向量 $\boldsymbol{u}$ 的坐标.

例如, 方程可以相加得到方程, 可以乘常数得到方程, 因此方程是向量. 将每个 n 元线性方程 $a_1x_1+\cdots+a_nx_n=b$ 用数组 $(a_1,\cdots,a_n,b)$ 代表, 这个数组就是这个线性方程的坐标.

又如, 全体实系数一元多项式 $f(x)$ 组成的集合 $\mathbf{R}[x]$, 其中任何两个多项式可以相加, V 中每个多项式可以乘实数仍得到实系数多项式, 这个集合 $\mathbf{R}[x]$ 是实数域 $\mathbf{R}$ 上的向量空间. 考虑 $\mathbf{R}[x]$ 中次数不超过 $n-1$ 的全体多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 组成的子集 $\mathbf{R}_n[x]$, 将 $\mathbf{R}_n[x]$ 中任意两个多项式相加、任意一个多项式乘实数, 得到的多项式的次数仍不超过 $n-1$, 仍在 $\mathbf{R}_n[x]$ 中, 因此 $\mathbf{R}_n[x]$ 仍是 $\mathbf{R}$ 上的向量空间. $\mathbf{R}_n[x]$ 中每个多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 可以由它的系数组成的 n 元数组 $(a_0,a_1,\cdots,a_{n-1})$ 代表, 作为 $f(x)$ 的坐标. 如果将 $\mathbf{R}[x]$ 中每个多项式 $f(x)=a_0+a_1x+\cdots+a_mx^m$ 也用数组 $(a_0,a_1,\cdots,a_m)$ 作为坐标, 由于次数 m 没有限制, 数组的长度可以无限增大, 只有用无限序列 $(a_0,a_1,\cdots,a_m,0,\cdots)$ 为坐标才能表示出 $\mathbf{R}[x]$ 中所有的多项式. 这样的空间 $\mathbf{R}[x]$ 称为无穷维空间.

将许多不同的东西都称为向量, 不是为了故弄玄虚, 也不是为了在考试时增加一个考题: "解释如下词语: 向量". 而是为了将向量的 8 条运算律推导出来的结论直接应用于千千万万的不同对象, 避免对它们一次又一次作重新推导的重复劳动.

2. 空集合的线性组合

定理 2.2.1 给出了 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 线性相关的两个充分必要条件:

(1) 其中某个向量 $\boldsymbol{a}_i$ 可以写成其余向量 $\boldsymbol{a}_j$ $(j\neq i)$ 的线性组合;

(2) 其中某个向量 $\boldsymbol{a}_i$ 可以写成它前面的向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1}$ 的线性组合.

这两个条件各有一个特殊情形值得注意:

情况 1. 一个向量的集合 $\{\boldsymbol{a}_1\}$.

按定理应有: $\boldsymbol{a}_1$ 线性相关 $\Leftrightarrow$ $\boldsymbol{a}_1$ 是集合 $\{\boldsymbol{a}_1\}$ 中其余向量的线性组合.

但集合 $\{\boldsymbol{a}_1\}$ 中除了 $\boldsymbol{a}_1$ 没有其余向量, 其余向量组成的集合是空集 $\varnothing$.

情况 2. $\boldsymbol{a}_1=\mathbf{0}$ 且 $\boldsymbol{a}_2,\cdots,\boldsymbol{a}_k$ 线性无关. 此时 $\mathbf{0},\boldsymbol{a}_2,\cdots,\boldsymbol{a}_k$ 线性相关.

按定理的证明, 如果满足 $x_1\boldsymbol{a}_1+\cdots+x_k\boldsymbol{a}_k=\mathbf{0}$ 的 $x_1,\cdots,x_k$ 中最后一个不为 0 的系数是 x_i, 则 $\boldsymbol{a}_i$ 是它前面的 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1}$ 的线性组合.

在此情况下 $(x_1,x_2,\cdots,x_k)=(1,0,\cdots,0)$, 最后一个 $x_i\neq 0$ 是 $x_1=1$. 但 $\boldsymbol{a}_1=\mathbf{0}$ 前面没有向量, $\boldsymbol{a}_1$ 无法等于"前面的向量"的线性组合.

我们规定: **空集合 $\varnothing$ 的线性组合是 $\mathbf{0}$.**

则以上两个特殊情况下定理 2.2.1 仍成立:

情况 1. $\{\boldsymbol{a}_1\}$ 线性相关 $\Leftrightarrow \boldsymbol{a}_1=\boldsymbol{0} \Leftrightarrow \boldsymbol{a}_1$ 是“其余向量” $\varnothing$ 的线性组合.

情况 2. $\boldsymbol{a}_1=\boldsymbol{0}$ 是它“前面的向量” $\varnothing$ 的线性组合.

3. 无穷集合的线性组合与线性相关

线性空间 F^n 是无穷多个向量组成的集合. 但线性代数中没有定义无穷多个向量的加法, 因此也不能进行无穷多个向量的线性组合.

引理 2.2.1 说: 向量集合 S 如果包含一部分线性相关的向量, 则 S 线性相关.

我们将引理 2.2.1 推广到无穷集合, 对 F^n 的任意子集 (包括无限子集) S 定义线性相关与线性无关:

如果 S 的某个有限子集线性相关, 则称 S 线性相关.

如果 S 的每个有限子集都线性无关, 则称 S 线性无关.

集合 S 中任意有限子集 $S_0=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合 $x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m$ 称为 S 的线性组合. 可以认为 S 中所有的向量都参加了这个线性组合, 只不过 S_0 之外的那些向量 $\boldsymbol{a}\in S\setminus S_0$ 的系数是 0:

$$\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m+\sum_{\boldsymbol{a}\in S\setminus S_0}0\boldsymbol{a}.$$

§2.3 基

平面上最多有两个线性无关的向量, 因此平面称为 2 维空间. 平面上任意两个线性无关 (不共线) 向量 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 组成一组基, 平面上每个向量 $\boldsymbol{b}$ 可以唯一地写成这组基的线性组合 $x\boldsymbol{a}_1+y\boldsymbol{a}_2$, 组合系数 x,y 组成的有序数组 (x,y) 称为 $\boldsymbol{b}$ 在这组基下的坐标, 用坐标 (x,y) 代表向量 $\boldsymbol{b}$ 参加运算, 通过坐标的运算来研究向量的各种性质.

几何空间中最多有 3 个线性无关的向量, 因此几何空间称为 3 维空间. 空间中任意 3 个线性无关 (不共面) 向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 组成一组基, 空间中每个向量 $\boldsymbol{b}$ 可以唯一地写成这组基的线性组合 $x\boldsymbol{a}_1+y\boldsymbol{a}_2+z\boldsymbol{a}_3$, 组合系数组成的有序数组 (x,y,z) 称为 $\boldsymbol{b}$ 在这组基下的坐标, 坐标 (x,y,z) 代表向量 $\boldsymbol{b}$ 参加运算.

很自然要问: F^n 中最多有多少个线性无关的向量? F^n 中任何一组线性无关向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 是否组成一组基 S, 使每个向量 $\boldsymbol{b}$ 可以唯一地写成 S 的线性组合 $\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m$?

2.3.1 维数

例 1 F^n 中最多有多少个线性无关的向量?

解 对 F^n 中 m 个向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$, 其中

$$\boldsymbol{a}_j=(a_{1j},a_{2j},\cdots,a_{nj})^{\mathrm{T}}\in F^n,\ \forall\ 1\leqslant j\leqslant m,$$

考虑方程组

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_m\boldsymbol{a}_m=\mathbf{0},\tag{2.3.1}$$

即

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=0,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=0,\\ \qquad\cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m=0.\end{cases}\tag{2.3.2}$$

齐次线性方程组 (2.3.2) 的未知数个数为 m, 方程个数为 n. 当 $m>n$ 时方程组 (2.3.2) (即 (2.3.1)) 必有非零解, $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性相关. 可见 F^n 中线性无关向量个数 $m\leqslant n$.

设 $m=n$. 我们设法找出 n 个向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 使方程 (2.3.1) 即方程组 (2.3.2) 只有唯一解 $(0,\cdots,0)$. 显然, 方程组

$$\begin{cases}x_1=b_1,\\ x_2=b_2,\\ \cdots\cdots\cdots\cdots\\ x_n=b_n\end{cases}\tag{2.3.3}$$

对任意 $b_1,\cdots,b_n$ 都有唯一解 $(b_1,\cdots,b_n)$. 特别当 $b_1=\cdots=b_n=0$ 时, (2.3.3) 是齐次线性方程组, 有唯一解 $(0,\cdots,0)$. 方程组 (2.3.3) 的系数矩阵是单位矩阵

$$\boldsymbol{I}=\begin{pmatrix}1&0&\cdots&0\\ 0&1&\cdots&0\\ \vdots&\vdots&&\vdots\\ 0&0&\cdots&1\end{pmatrix},$$

它的各列为

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

其中 $\boldsymbol{e}_j$ $(1\leqslant j\leqslant n)$ 是第 j 分量为 1、其余分量为 0 的列向量. 此时方程组 (2.3.3) 即为

$$x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n=\mathbf{0}.$$

这个方程组只有零解 $(0,\cdots,0)$, 说明 n 个向量 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n$ 线性无关.

F^n 中存在 n 个线性无关的向量, 不存在多于 n 个的线性无关向量. 可见, F^n 中最多有 n 个线性无关的向量. □

n 数组空间 F^n 中最多只存在 n 个线性无关的向量. 因此 F^n 称为 n **维空间** (n dimensional space) .

2.3.2 基的定义

例 1 中举出了由 n 个线性无关向量组成的向量组 $E = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$, 其中每个 $\boldsymbol{e}_i$ 的第 i 分量等于 1、其余分量都是 0. 例 1 中的方程组 (2.3.3) 即

$$x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = \boldsymbol{b} \tag{2.3.4}$$

对每个 $\boldsymbol{b} = (b_1, \cdots, b_n)^{\mathrm{T}}$ 有唯一解 $(x_1, \cdots, x_n) = (b_1, \cdots, b_n)$, 这说明 F^n 中每个向量 $\boldsymbol{b} = (b_1, \cdots, b_n)^{\mathrm{T}}$能够唯一地写成向量组$E$=$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 的线性组合:

$$\boldsymbol{b} = b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n.$$

很自然, 我们将这个向量组 E 称为 n 维空间 F^n 的一组基.

一般地, 我们有:

定义 2.3.1(基与坐标) 如果 $T = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m\}$ 是 F^n 中的一个向量组, 使 F^n 中每个向量 $\boldsymbol{b}$ 都能唯一地写成 T 的线性组合 $\boldsymbol{b} = x_1\boldsymbol{a}_1 + \cdots + x_m\boldsymbol{a}_m$, 则 T 称为 F^n 的一组**基** (basis), 由线性组合系数组成的有序数组 $(x_1, \cdots, x_m)$ 称为向量 $\boldsymbol{a}$ 在这组基 T 下的**坐标** (coordinates). □

例 1 中举出的线性无关向量组 $E = \{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 是 F^n 的最简单也最重要的基, 每个数组向量 $\boldsymbol{b} = (b_1, \cdots, b_n)$ 在这组基下的坐标就是 $(b_1, \cdots, b_n)$ 本身. 这组基 E 称为 F^n 的**自然基** (natural basis) .

F^n 可以写成行向量空间 $F^{1\times n}$, 此时它的自然基向量 $\boldsymbol{e}_i$ 就是行向量. F^n 也可以写成列向量空间 $F^{n\times 1}$, 此时它的自然基向量 $\boldsymbol{e}_i$ 就是列向量.

当 $n = mk$ 是两个大于 1 的正整数 m 与 k 的乘积时, F^n 还可以写成由 F 上全体 $m\times k$ 矩阵组成的空间 $F^{m\times k}$, 此时 F^n 的自然基变成由只有一个分量为 1、其余分量全为 0 的 $m \times k$ 矩阵组成的基 $E = \{\boldsymbol{E}_{ij} \mid 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant k\}$, 其中 $\boldsymbol{E}_{ij}$ 是第 i 行第 j 列元素为 1、其余元素都为 0 的矩阵. $F^{m\times k}$ 中每个矩阵 $\boldsymbol{A} = (a_{ij})_{m\times k}$ 唯一地写成这组基的线性组合

$$\boldsymbol{A} = \sum_{1\leqslant i\leqslant m,\, 1\leqslant j\leqslant k} a_{ij}\boldsymbol{E}_{ij},$$

其中 $\boldsymbol{E}_{ij}$ 的系数 a_{ij} 就是矩阵 $\boldsymbol{A}$ 的第 i 行第 j 列的元素. 不过, 将线性组合表达式中的这些系数组成 $\boldsymbol{A}$ 的坐标 $\boldsymbol{X}$ 的时候, 并不将它们排成 $m \times k$ 的形式,

而仍然排成行向量或列向量的形式, 例如, 排成行向量时 $\boldsymbol{A}$ 的坐标为

$$\boldsymbol{X} = (a_{11}, \cdots, a_{1k}, a_{21}, \cdots, a_{2k}, \cdots, a_{m1}, \cdots, a_{mk}).$$

例 2 在复数域 $\mathbf{C}$ 上的 3 维空间 $\mathbf{C}^3$ 中, 向量 $\boldsymbol{a}_1 = (2,0,0), \boldsymbol{a}_2 = (a_2,3,0)$, $\boldsymbol{a}_3 = (a_3,a_4,5)$ 是否组成一组基?

解 只需看方程组 $x\boldsymbol{a}_1 + y\boldsymbol{a}_2 + z\boldsymbol{a}_3 = \boldsymbol{b}$ 即

$$\begin{cases} 2x + a_2y + a_3z = b_1, \\ \qquad 3y + a_4z = b_2, \\ \qquad\qquad 5z = b_3 \end{cases} \tag{2.3.5}$$

是否对任意 $\boldsymbol{b} = (b_1,b_2,b_3)$ 有唯一解.

将方程组 (2.3.5) 用矩阵 $\boldsymbol{M}$ 表示, 经过一系列初等行变换化成最简阶梯形矩阵, 得:

$$\boldsymbol{M} = \begin{pmatrix} 2 & a_2 & a_3 & b_1 \\ 0 & 3 & a_4 & b_2 \\ 0 & 0 & 5 & b_3 \end{pmatrix} \to \cdots \to \boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix}.$$

不论其中表示常数项的 $*$ 取什么值, 最简阶梯形 $\boldsymbol{A}$ 所表示的方程组都有唯一解. 这说明: 不论原方程组 (2.3.5) 中的常数项 b_1,b_2,b_3 取什么复数值, 方程组 (2.3.5) 总有唯一解. 原题中的 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 组成的向量组都是 $\mathbf{C}^3$ 的基. □

例 2 的方程组的增广矩阵 $\boldsymbol{M}$ 已经是阶梯形矩阵, 前 3 列组成的系数矩阵

$$\boldsymbol{A} = \begin{pmatrix} 2 & a_2 & a_3 \\ 0 & 3 & a_4 \\ 0 & 0 & 5 \end{pmatrix}$$

是上三角形矩阵, 且对角元 2,3,5 都不为 0. 将 $\boldsymbol{M}$ 化成最简阶梯形矩阵 $\boldsymbol{A}$ 时, 系数矩阵 $\boldsymbol{A}$ 化成了单位矩阵 $\boldsymbol{I}$, 无论 $\boldsymbol{A}$ 的第 4 列的元素取什么值, 所代表的方程组都有唯一解.

n 阶方阵 $\boldsymbol{A}$ 如果是阶梯形矩阵, $\boldsymbol{A}$ 一定是上三角形矩阵. 如果 $\boldsymbol{A}$ 有 n 个阶梯, 则每行各有一个阶梯, 所有的阶梯都在主对角线上, $\boldsymbol{A}$ 是对角元全不为 0 的上三角形矩阵. 如果每个阶梯元 (也就是对角元) 所在列的其余元素都是 0, 它就是对角矩阵. 如果 $\boldsymbol{A}$ 还是最简阶梯形矩阵, 它就是单位矩阵.

与例 2 同理可知: 对角元全不为 0 的 n 阶上三角形矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

($a_{11},a_{22},\cdots,a_{nn}$ 都不为 0) 可以通过一系列初等行变换化成单位矩阵 $\boldsymbol{I}$, 因而以 $\boldsymbol{A}$ 为系数矩阵的方程组有唯一解, $\boldsymbol{A}$ 的各列组成 F^n 的一组基.

类似地可验证: 以对角元全不为 0 的 n 阶下三角形矩阵 $\boldsymbol{A}$ 为系数矩阵的线性方程组也都有唯一解, 这样的方阵 $\boldsymbol{A}$ 的各列也组成 F^n 的基. 还可知道: 对角元全不为 0 的上三角形矩阵或下三角形矩阵的各行也组成 F^n 的基.

2.3.3 向量运算与坐标运算的对应关系

向量本来是指有方向与大小的量, 也就是几何向量. 几何向量的运算通过几何方式定义, 比数的运算更困难一些. 例如, 几何向量的加法由三角形法则或平行四边形法则定义, 要由两个几何向量的长度和方向计算出它们的和的长度和方向就不大容易. 在引入坐标后, 将每个几何向量用坐标表示, 将几何向量的运算转化为坐标的运算来进行, 运算就变得简单了. 几何向量的运算与坐标运算的这种对应关系很容易推广到 F^n 中的向量.

设向量 $\boldsymbol{b},\boldsymbol{c}$ 在基 $T=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 下的坐标分别是 $(x_1,\cdots,x_m)$ 与 $(y_1,\cdots,y_m)$, λ 是 F 中任意数. 则

$$\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m,\qquad \boldsymbol{c}=y_1\boldsymbol{a}_1+\cdots+y_m\boldsymbol{a}_m, \tag{2.3.6}$$

$$\boldsymbol{b}+\boldsymbol{c}=(x_1+y_1)\boldsymbol{a}_1+\cdots+(x_m+y_m)\boldsymbol{a}_m, \tag{2.3.7}$$

$$\lambda\boldsymbol{b}=(\lambda x_1)\boldsymbol{a}_1+\cdots+(\lambda x_m)\boldsymbol{a}_m. \tag{2.3.8}$$

这说明, 向量 $\boldsymbol{b}$ 与 $\boldsymbol{c}$ 之和 $\boldsymbol{b}+\boldsymbol{c}$ 的坐标为 $(x_1+y_1,\cdots,x_n+y_n)$, 等于 $\boldsymbol{b},\boldsymbol{c}$ 的坐标之和; 向量 $\boldsymbol{b}$ 的 λ 倍 $\lambda\boldsymbol{b}$ 的坐标为 $(\lambda x_1,\cdots,\lambda x_n)$, 等于 $\boldsymbol{b}$ 的坐标的 λ 倍. 向量的加法与数乘运算, 可以转化为坐标的加法与数乘来进行.

为了将向量与坐标的这种对应关系表达得更清楚, 我们将任意一个向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 中的各向量排成 "行向量" 的形式 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 来表示这个向量组, 将线性组合式 $\boldsymbol{b}=x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m$ 中的系数 $x_1,\cdots,x_m$ 排成列向量 $\boldsymbol{X}$, 将这个线性组合式写成 "行向量" $\boldsymbol{A}$ 与列向量 $\boldsymbol{X}$ 的乘积的形式:

$$\boldsymbol{b}=\boldsymbol{A}\boldsymbol{X}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}=\boldsymbol{a}_1x_1+\cdots+\boldsymbol{a}_mx_m.$$

乘法法则是: 将 $\boldsymbol{A}$ 的各"分量" $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 与 $\boldsymbol{X}$ 的各分量 $x_1,\cdots,x_n$ 对应相乘再相加.

按照这种表达方式, 等式 (2.3.6) 就成为 $\boldsymbol{b}=\boldsymbol{AX}$, $\boldsymbol{c}=\boldsymbol{AY}$, 其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$, $\boldsymbol{X}=(x_1,\cdots,x_m)^{\mathrm{T}},\boldsymbol{Y}=(y_1,\cdots,y_m)^{\mathrm{T}}$. 而 (2.3.7) 与 (2.3.8) 分别成为:

$$\boldsymbol{AX}+\boldsymbol{AY}=\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{Y}),\qquad (\boldsymbol{AX})\lambda=\boldsymbol{A}(\boldsymbol{X}\lambda).\tag{2.3.9}$$

将向量组 $\boldsymbol{A}$ 看成一个"向量", 系数组 $\boldsymbol{X},\boldsymbol{Y}$ 各看成一个"数", 则 (2.3.9) 中第一个等式看起来好像是按照乘法对于加法的分配律将 $\boldsymbol{A}$ 的 $\boldsymbol{X}$ 倍与 $\boldsymbol{Y}$ 倍合并"同类项"得到 $\boldsymbol{X}+\boldsymbol{Y}$ 倍, 第二个等式就好像是按照乘法结合律将 $\boldsymbol{A}$ 的 $\boldsymbol{X}$ 倍乘 λ 得到 $\boldsymbol{X}\lambda$ 倍.

特别地, 如果 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 表示的向量组是 F^n 的一组基, $\boldsymbol{X},\boldsymbol{Y}$ 就分别是向量 $\boldsymbol{b}=\boldsymbol{AX}$ 与 $\boldsymbol{c}=\boldsymbol{AY}$ 的坐标, (2.3.9) 的结论就是: 向量 $\boldsymbol{b},\boldsymbol{c}$ 的和与 $\boldsymbol{b}$ 的常数倍的坐标, 分别等于 $\boldsymbol{b},\boldsymbol{c}$ 的坐标 $\boldsymbol{X},\boldsymbol{Y}$ 之和与 $\boldsymbol{b}$ 的坐标 X 的常数倍.

由 (2.3.9) 可以进一步算出: $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 表示的向量组 S 的若干个线性组合 $\boldsymbol{b}_i=\boldsymbol{AX}_i$ $(1\leqslant i\leqslant k)$ 的线性组合

$$\begin{aligned}\boldsymbol{b}_1\lambda_1+\cdots+\boldsymbol{b}_k\lambda_k&=(\boldsymbol{AX}_1)\lambda_1+\cdots+(\boldsymbol{AX}_k)\lambda_k\\&=\boldsymbol{A}(\boldsymbol{X}_1\lambda_1+\cdots+\boldsymbol{X}_k\lambda_k)\end{aligned}\tag{2.3.10}$$

仍是 S 的线性组合, 其系数组 $\boldsymbol{X}_1\lambda_1+\cdots+\boldsymbol{X}_k\lambda_k$ 等于各系数组 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_k$ 的相应线性组合.

为了便于以后的应用, 我们将以上结论叙述为一个引理:

引理 2.3.1(线性组合的传递性) 向量组 S 的线性组合的线性组合仍是 S 的线性组合. □

特别地, 如果 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 表示的向量组是 F^n 的基, 则 $\boldsymbol{X}_i$ 是 $\boldsymbol{b}_i$ 的坐标 $(1\leqslant i\leqslant k)$, 等式 (2.3.10) 说的就是:

向量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k$ 的线性组合的坐标, 等于它们的坐标 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_k$ 的相应的线性组合.

由此立即得到:

引理 2.3.2 向量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k$ 线性相关 (无关) $\Leftrightarrow$ 它们在同一组基下的坐标 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_k$ 线性相关 (无关).

证明 向量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k$ 线性相关 $\Leftrightarrow$ 存在不全为 0 的 $\lambda_1,\cdots,\lambda_k\in F$ 使 $\boldsymbol{b}_1\lambda_1+\cdots+\boldsymbol{b}_k\lambda_k=\boldsymbol{0}$ 即 $\boldsymbol{X}_1\lambda_1+\cdots+\boldsymbol{X}_k\lambda_k=\boldsymbol{0}\Leftrightarrow\boldsymbol{X}_1,\cdots,\boldsymbol{X}_k$ 线性相关. □

可见，不但向量的加法、数乘、线性组合等运算可以由坐标“全权代表”，对向量的线性相关与线性无关等性质的判定和研究也可以由坐标“全权代表”.

2.3.4 基的判定定理

平面上任意两个线性无关 (不共线) 的几何向量组成一组基，空间中任意 3 个线性无关 (不共面) 的几何向量组成一组基. 对任意 n 维数组空间 F^n，很自然猜测:

F^n 中任意 n 个线性无关的向量 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 组成一组基.

F^n 的基 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 由两条性质刻画:

(1) (坐标的存在性) 每个向量 $\boldsymbol{b}\in F^n$ 都能写成 S 的线性组合 $\boldsymbol{b}=\boldsymbol{AX}$，其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)$.

(2) (坐标的唯一性) 每个 $\boldsymbol{b}=\boldsymbol{AX}$ 中的 $\boldsymbol{X}$ 由 $\boldsymbol{b}$ 唯一决定. 即: $\boldsymbol{AX}=\boldsymbol{AY} \Leftrightarrow \boldsymbol{X}=\boldsymbol{Y}$.

我们来看怎样的向量组满足这两个条件.

坐标的唯一性

F^n 的零向量 $\boldsymbol{0}$ 显然是任意 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合: $\boldsymbol{0}=0\boldsymbol{a}_1+\cdots+0\boldsymbol{a}_m=\boldsymbol{A0}_m$，其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$，$\boldsymbol{0}_m$ 是全部由 0 组成的 m 维列向量.

要使零向量 $\boldsymbol{0}$ 的坐标唯一，必须要求 $\boldsymbol{AX}=\boldsymbol{0}\Rightarrow\boldsymbol{X}=\boldsymbol{0}$，即: $x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m=\boldsymbol{0}\Rightarrow x_1=\cdots=x_m=0$. 这就是要求向量组 S 线性无关.

设 $\boldsymbol{b}=\boldsymbol{AX}$ 是 S 的线性组合. 则 $\boldsymbol{AY}=\boldsymbol{b}\Leftrightarrow\boldsymbol{AX}-\boldsymbol{AY}=\boldsymbol{0}\Leftrightarrow\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{Y})=\boldsymbol{0}$. 因此，$\boldsymbol{AY}=\boldsymbol{b}$ 有唯一解 $\boldsymbol{Y}=\boldsymbol{X}\Leftrightarrow\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{Y})=\boldsymbol{0}$ 有唯一解 $\boldsymbol{X}-\boldsymbol{Y}=\boldsymbol{0}\Leftrightarrow$ 向量组 S 线性无关.

可见，坐标是否唯一由 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 是否线性无关决定. 某个向量的坐标唯一 $\Leftrightarrow$ 零向量 $\boldsymbol{0}$ 坐标唯一 $\Leftrightarrow$ S 线性无关 $\Leftrightarrow$ S 的每个线性组合的坐标唯一.

因此，基 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 一定线性无关. 由于 F^n 中线性无关的向量不超过 n 个，每组基中所含向量个数 $m\leqslant n$. 进一步，我们有:

引理 2.3.3 F^n 的任何两组基所含向量的个数相等. 每组基都由 n 个线性无关向量组成.

证明 设 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 与 $T=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k\}$ 都是 F^n 的基. 基 T 中 k 个向量 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k$ 线性无关，它们在基 S 下的坐标 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_k$ 是 k 个线性无关的 m 维列向量. 线性无关的 m 维列向量个数不超过 m，因此 $k\leqslant m$. 反过来，S 的 m 个基向量在基 T 下的坐标 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_m$ 是 m 个线性无关的 k 维

列向量, 这又使得 $m \leqslant k$. 这就证明了 $m = k$. F^n 的任何两组基所含向量个数相等.

已经知道 F^n 的自然基 $E = \{e_1, \cdots, e_n\}$ 由 n 个向量组成. 因此 F^n 的每一组基都由 n 个向量组成. □

坐标的存在性

反过来, 我们证明:

引理 2.3.4 F^n 中任意 n 个线性无关向量 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ 组成一组基.

证明 由于向量组 $S = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$ 线性无关, 其中每个向量 $\boldsymbol{a}_i$ 都不是其余 $\boldsymbol{a}_j$ $(j \neq i)$ 的线性组合. 如果 F^n 中有某个向量 $\boldsymbol{b}$ 不是 S 的线性组合, 则向量组 $M = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{b}\}$ 中每个向量都不是它前面的向量的线性组合, 由推论 2.2.1 知 M 线性无关. 但 M 包含 $n+1$ 个向量, 而 n 维空间 F^n 中 $n+1$ 个向量一定线性相关, 矛盾.

这证明了 F^n 中每个向量 $\boldsymbol{b}$ 都必然是 S 的线性组合, 即 $\boldsymbol{b} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n$, 由 S 线性无关知系数组 $(x_1, \cdots, x_n)$ 由 $\boldsymbol{b}$ 唯一决定. 这证明了 S 是 F^n 的基. □

这就得到了:

定理 2.3.1 F^n 中的向量组 S 是基 $\Leftrightarrow$ $S = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$ 由 n 个线性无关向量组成.

2.3.5 判定线性方程组的唯一解

例 3 在复数范围内求常数 b_1, b_2, b_3, 使线性方程组

$$\begin{cases} x + y + z = b_1, \\ x + 2y + 4z = b_2, \\ x + 3y + 9z = b_3 \end{cases} \tag{2.3.11}$$

有唯一解.

解 将原方程组 (2.3.11) 写成向量形式

$$x\boldsymbol{a}_1 + y\boldsymbol{a}_2 + z\boldsymbol{a}_3 = \boldsymbol{b}. \tag{2.3.12}$$

只需判定 $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ 代表的向量组 S 是否线性无关, 就知道 S 是否是 $\mathbf{C}^3$ 的基, 方程组 (2.3.11) 是否有唯一解. S 线性无关的充分必要条件是: 方程组

$$x\boldsymbol{a}_1 + y\boldsymbol{a}_2 + z\boldsymbol{a}_3 = \boldsymbol{0} \tag{2.3.13}$$

即

$$\begin{cases} x+\ \ y+\ \ z=0, \\ x+2y+4z=0, \\ x+3y+9z=0 \end{cases} \tag{2.3.14}$$

有唯一解 $(0,0,0)$. $\boldsymbol{A}$ 就是方程组 (2.3.11) 与 (2.3.14) 的系数矩阵, 经过初等行变换化简:

$$\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \xrightarrow{-(2)+(3),-(1)+(2),-(2)+(3)} \boldsymbol{B}=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

$\boldsymbol{B}$ 是对角元全不为 0 的上三角形矩阵, 以 $\boldsymbol{B}$ 为系数矩阵的齐次线性方程组有唯一解 $(0,0,0)$, 因而 (2.3.14) 有唯一解 $(0,0,0)$. 可见 (2.3.14) 与 (2.3.11) 共同的系数矩阵 $\boldsymbol{A}$ 的各列线性无关, 组成 $\mathbf{C}^3$ 的基. 原方程组 (2.3.11) 对每一组 b_1,b_2,b_3 都有唯一解. □

例 3 的推理和结论可以推广到任意的线性方程组, 得到

定理 2.3.2 线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=b_1, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1+\cdots+a_{mn}x_n=b_m. \end{cases} \tag{2.3.15}$$

对任意一组 $b_1,\cdots,b_m$ 都有唯一解的充分必要条件是: $m=n$, 且齐次线性方程组

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n=0, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1+\cdots+a_{nn}x_n=0 \end{cases} \tag{2.3.16}$$

有唯一解 $(0,\cdots,0)$.

证明 将原方程组 (2.3.15) 写成向量形式

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}, \tag{2.3.17}$$

即可知道: 方程组 (2.3.17) 对任意的 $\boldsymbol{b}$ 有唯一解 $\Leftrightarrow$ 系数矩阵 $\boldsymbol{A}$ 的列向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$ 是 $F^{m\times 1}$ 的一组基 $\Leftrightarrow$ $m=n$ 且 S 线性无关 $\Leftrightarrow$ $m=n$ 且

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{0} \tag{2.3.18}$$

有唯一解. 而 (2.3.18) 就是 (2.3.16) 的向量形式. □

将定理 2.3.2 中的方程组 (2.3.17) 等号左边的线性组合 $x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n$ 写成 $\boldsymbol{AX}$ 的形式, (2.3.17) 就写成了

$$\boldsymbol{AX}=\boldsymbol{b} \tag{2.3.19}$$

的形式, 其中

$$\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn}\end{pmatrix},\quad \boldsymbol{X}=\begin{pmatrix} x_1 \\ \vdots \\ x_n\end{pmatrix},$$

$\boldsymbol{A}$ 是依次以 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 为“元素”排成的“行向量”, 也可看成依次以 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 为各列排成的 $m\times n$ 矩阵, 就是定理 2.3.2 中的方程组 (2.3.15) 的系数矩阵. $\boldsymbol{AX}$ 可以看成矩阵 $\boldsymbol{A}$ 与列向量 $\boldsymbol{X}$ 的乘积, 乘法法则是: 将 $\boldsymbol{A}$ 的各列 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 与 $\boldsymbol{X}$ 的各分量 $x_1,\cdots,x_n$ 对应相乘再相加.

$\boldsymbol{AX}=\boldsymbol{b}$ 称为方程组 (2.3.15) 的矩阵形式. 齐次线性方程组 (2.3.16) 同样地可以写成矩阵乘法的形式 $\boldsymbol{AX}=\boldsymbol{0}$.

如果 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 都是 $\boldsymbol{AX}=\boldsymbol{b}$ 的解, 则 $\boldsymbol{AX}_1=\boldsymbol{AX}_2 \Leftrightarrow \boldsymbol{AX}_1-\boldsymbol{AX}_2=\boldsymbol{A}(\boldsymbol{X}_1-\boldsymbol{X}_2)=\boldsymbol{0},\boldsymbol{X}_1-\boldsymbol{X}_2$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解. 这再次证明了: 方程 $\boldsymbol{AX}=\boldsymbol{b}$ 有唯一解 $\Leftrightarrow$ $\boldsymbol{AX}=\boldsymbol{0}$ 有唯一解.

引理 2.3.5 设 F 上 n 阶方阵 $\boldsymbol{A}$ 经过一系列初等行变换变成 $\boldsymbol{B}$, 并经过一系列初等行变换变成阶梯形方阵 $\boldsymbol{T}$. 则

$\boldsymbol{A}$ 的各列组成 F^n 的基 $\Leftrightarrow$ $\boldsymbol{B}$ 的各列组成 F^n 的基 $\Leftrightarrow$ $\boldsymbol{T}$ 的对角元全不为0.

证明 方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{BX}=\boldsymbol{0}$ 同解. $\boldsymbol{AX}=\boldsymbol{0}$ 没有非零解 $\Leftrightarrow$ $\boldsymbol{BX}=\boldsymbol{0}$ 没有非零解. 这就是说: $\boldsymbol{A}$ 的各列组成 $F^{n\times1}$ 的基 $\Leftrightarrow$ $\boldsymbol{B}$ 的各列组成 $F^{n\times1}$ 的基.

阶梯形方阵 $\boldsymbol{T}$ 一定是上三角形矩阵. $\boldsymbol{A}$ 的各列组成 F^n 的基 $\Leftrightarrow$ 上三角形矩阵 $\boldsymbol{T}$ 的各列组成 F^n 的基 $\Leftrightarrow$ T 的对角元全不为 0. □

例 4 当 b_1,b_2,b_3 取什么复数值时, 如下方程组有唯一解:

$$(1)\ \begin{cases} x+2y+\ \ z=b_1, \\ x+\ \ y-\ 3z=b_2, \\ x+5y+13z=b_3; \end{cases} \qquad (2)\ \begin{cases} x+2y+\ z=b_1, \\ x+\ \ y-3z=b_2, \\ x+5y+4z=b_3. \end{cases}$$

解 两个方程组的系数方阵经过初等行变换化成的阶梯形矩阵分别为

$$(1) \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -4 \\ 0 & 0 & 0 \end{pmatrix}; \qquad (2) \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -4 \\ 0 & 0 & -9 \end{pmatrix}.$$

不论 b_1, b_2, b_3 取什么值, (1) 的方程组都没有唯一解, (2) 的方程组都有唯一解. □

习题 2.3

1. 判断如下哪些向量组是 $\mathbf{R}^3$ 的基:

(1) $\boldsymbol{\alpha}_1 = (1,2,1), \boldsymbol{\alpha}_2 = (1,1,-3), \boldsymbol{\alpha}_3 = (1,5,13)$;

(2) $\boldsymbol{\alpha}_1 = (1,2,4), \boldsymbol{\alpha}_2 = (1,-1,1), \boldsymbol{\alpha}_3 = (1,3,9)$;

(3) $\boldsymbol{\alpha}_1 = (1,2,4), \boldsymbol{\alpha}_2 = (1,0,5), \boldsymbol{\alpha}_3 = (6,0,0)$.

2. 已知 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是 $\mathbf{R}^3$ 的基. 判断如下哪些向量组是 $\mathbf{R}^3$ 的基:

(1) $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 - 3\boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_1 + 5\boldsymbol{\alpha}_3 + 13\boldsymbol{\alpha}_3$;

(2) $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_1 + 9\boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_1 + 4\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3$;

(3) $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_2 = 2\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3$, $\boldsymbol{\beta}_3 = 4\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + 9\boldsymbol{\alpha}_3$.

3. b_1, b_2, b_3 取什么实常数时, 如下方程组有唯一解?

$$(1)\begin{cases} x + y + z = b_1, \\ 2x + y + 5z = b_2, \\ x - 3y + 13z = b_3; \end{cases} \qquad (2)\begin{cases} b_1 x + y + z = 1, \\ b_2 x + 5z = 2, \\ -3x + 13z = b_3. \end{cases}$$

4. 将方程组

$$x\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix} + z\begin{pmatrix} 1 \\ 5 \\ 13 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

写成矩阵形式 $\boldsymbol{AX} = \boldsymbol{b}$. 并写出它的系数矩阵 $\boldsymbol{A}$.

5. 已知

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

(1) 求 $\boldsymbol{b}_1 = \boldsymbol{Ab}$, $\boldsymbol{b}_2 = \boldsymbol{Ab}_1$, $\boldsymbol{b}_3 = \boldsymbol{Ab}_2$;

(2) 求方程组 $\boldsymbol{AX} = \boldsymbol{b}$ 的通解 $\boldsymbol{X}$.

§2.4 坐 标 变 换

2.4.1 求向量的坐标

例 1 求向量 $\boldsymbol{\varepsilon}_1 = (1,0,0)$ 在 F^3 的基 $T = \{\boldsymbol{\alpha}_1 = (1,1,1), \boldsymbol{\alpha}_2 =$

$(1,2,3)$, $\boldsymbol{\alpha}_3=(1,4,9)\}$ 下的坐标.

解 $\boldsymbol{\varepsilon}_1$ 在基 T 下的坐标 (x_1,x_2,x_3) 满足条件

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3=\boldsymbol{\varepsilon}_1. \tag{2.4.1}$$

将 $\boldsymbol{\varepsilon}_1,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 分别转置写成列向量 $\boldsymbol{e}_1,\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$, 则条件 (2.4.1) 成为

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+x_3\boldsymbol{a}_3=\boldsymbol{e}_1,$$

即方程组

$$\boldsymbol{AX}=\boldsymbol{e}_1, \tag{2.4.2}$$

其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$ 是方程组 (2.4.2) 的系数矩阵, $\boldsymbol{X}=(x_1,x_2,x_3)^{\mathrm{T}}$.

将方程组 $\boldsymbol{AX}=\boldsymbol{e}_1$ 的增广矩阵 $(\boldsymbol{A},\boldsymbol{e}_1)$ 经过初等行变换化简得

$$(\boldsymbol{A},\boldsymbol{e}_1)=\begin{pmatrix}1&1&1&1\\1&2&4&0\\1&3&9&0\end{pmatrix}\xrightarrow{-(2)+(3),-(1)+(2)}\begin{pmatrix}1&1&1&1\\0&1&3&-1\\0&1&5&0\end{pmatrix}\xrightarrow{-(2)+(3),-(2)+(1)}$$

$$\begin{pmatrix}1&0&-2&2\\0&1&3&-1\\0&0&2&1\end{pmatrix}\xrightarrow{(3)+(1),\frac{1}{2}(3),-3(3)+(2)}\begin{pmatrix}1&0&0&3\\0&1&0&-\dfrac{5}{2}\\0&0&1&\dfrac{1}{2}\end{pmatrix},$$

得到方程组 (2.4.2) 的解 $\boldsymbol{X}=\left(3,-\dfrac{5}{2},\dfrac{1}{2}\right)^{\mathrm{T}}$, 就是 $\boldsymbol{\varepsilon}_1$ 在 T 下的坐标. □

一般地, 设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 组成列向量空间 $F^{n\times 1}$ 的一组基, 则列向量 $\boldsymbol{b}\in F^{n\times 1}$ 在这组基下的坐标 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 是方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的唯一解, 其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)$ 是依次以 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 为各列排成的矩阵.

例 2 设 $\mathbf{R}_3[x]$ 是由不超过 2 次的全体实系数多项式 $f(x)$ 组成的向量空间. 求多项式 $(x-5)^2$ 在基 $S=\{1,x-2,(x-2)^2\}$ 下的坐标.

解法 1 $\mathbf{R}_3[x]$ 中每个多项式 $f(x)=a_0+a_1x+a_2x^2$ 可以由系数组 $\boldsymbol{a}=(a_0,a_1,a_2)$ 代表, 参加加法与数乘运算.

将 S 中的多项式 $1,-2+x,4-4x+x^2$ 分别用系数组 $\boldsymbol{a}_1=(1,0,0)^{\mathrm{T}},\boldsymbol{a}_2=(-2,1,0)^{\mathrm{T}},\boldsymbol{a}_3=(4,-4,1)^{\mathrm{T}}$ 代表. 多项式 $(x-5)^2=25-10x+x^2$ 由系数组 $\boldsymbol{b}=(25,-10,1)^{\mathrm{T}}$ 代表. 解方程组 $\boldsymbol{AX}=\boldsymbol{b}$(其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$), 求 $\boldsymbol{b}$ 在基 $S_1=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$ 下的坐标 $\boldsymbol{X}$, 则 $\boldsymbol{X}$ 就是 $(x-5)^2$ 在 S 下的坐标.

对方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的增广矩阵 $(\boldsymbol{A},\boldsymbol{b})$ 作初等行变换, 得:

$$(\boldsymbol{A},\boldsymbol{b})=\begin{pmatrix}1&-2&4&25\\0&1&-4&-10\\0&0&1&1\end{pmatrix}\to\begin{pmatrix}1&0&0&9\\0&1&0&-6\\0&0&1&1\end{pmatrix},$$

$\boldsymbol{b}$ 在基 S_1 下的坐标 $\boldsymbol{X}=(9,-6,1)^{\mathrm{T}}$, 这就是 $(x-5)^2$ 在 S 下的坐标.

解法 2 令 $y=x-2$, 即 $x=y+2$, 代入 $(x-5)^2$ 并整理得:

$$(x-5)^2=(y+2-5)^2=(y-3)^2=y^2-6y+9=9-6(x-2)+(x-2)^2,$$

可见 $(x-5)^2$ 在 S 下的坐标为 $(9,-6,1)$. □

设 $F_n[x]$ 是系数在数域 F 中、次数不超过 $n-1$ 的全体一元多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$ 组成的向量空间. 则 $T=\{1,x,\cdots,x^{n-1}\}$ 是 $F_n[x]$ 的一组基, $f(x)$ 的系数组 $\boldsymbol{\alpha}=(a_0,a_1,\cdots,a_{n-1})$ 就是 $f(x)$ 在基 T 下的坐标, 可以用 $\boldsymbol{\alpha}$ 代表 $f(x)$ 参加加法与数乘运算, 包括求 $f(x)$ 在另一组基 S 下的坐标. 例 2 解法 1 就是这样做的. 这样做虽然可能比较繁琐, 但按部就班, 比较容易掌握.

例 2 解法 2 看起来比解法 1 简便, 但适用范围较窄, 可用来求 $f(x)\in F_n[x]$ 在基 $S=\{1,(x-c),(x-c)^2,\cdots,(x-c)^{n-1}\}$ 下的坐标 $(b_0,b_1,\cdots,b_{n-1})$:

$$f(x)=b_0+b_1(x-c)+b_2(x-c)^2+\cdots+b_{n-1}(x-c)^{n-1},$$

这实际上是在将 $f(x)$ 在 $x=c$ 作泰勒展开:

$$f(x)=\lambda_0+\lambda_1(x-c)+\lambda_2\frac{(x-c)^2}{2!}+\cdots+\lambda_{n-1}\frac{(x-c)^{n-1}}{(n-1)!},$$

其中的系数 $\lambda_k=(k!)b_k=f^{(k)}(c)$ $(0\leqslant k\leqslant n-1)$ 就是 $f(x)$ 在 c 对 x 的 k 阶导数, $(\lambda_0,\lambda_1,\cdots,\lambda_{n-1})$ 就是 $f(x)$ 在基 $\left\{1,x-c,\cdots,\dfrac{(x-c)^k}{k!},\cdots,\dfrac{(x-c)^{n-1}}{(n-1)!}\right\}$ 下的坐标.

例 3 求自然基向量 $\boldsymbol{\varepsilon}_1=(1,0,0),\boldsymbol{\varepsilon}_2=(0,1,0),\boldsymbol{\varepsilon}_3=(0,0,1)$ 在 F^3 的基 $T=\{\boldsymbol{\alpha}_1=(1,1,1),\ \boldsymbol{\alpha}_2=(1,2,3),\ \boldsymbol{\alpha}_3=(1,4,9)\}$ 下的坐标.

分析 将各向量 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 都写成列向量的形式 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$. 基向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 排成矩阵 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$, 则 $\boldsymbol{\varepsilon}_i$ $(i=1,2,3)$ 在基 T 下的坐标 $\boldsymbol{X}_i$ 就是方程组 $\boldsymbol{AX}=\boldsymbol{e}_i$ 的解. 例 1 中已经通过对 $(\boldsymbol{A},\boldsymbol{e}_1)$ 作初等行变换求出了 $\boldsymbol{\varepsilon}_1$ 的坐标 $\boldsymbol{X}_1$. 同样地, 对 $(\boldsymbol{A},\boldsymbol{e}_2),(\boldsymbol{A},\boldsymbol{e}_3)$ 作初等行变换可以得到另外两个 $\boldsymbol{\varepsilon}_i$ 的坐标.

将增广矩阵 $(\boldsymbol{A},\boldsymbol{e}_i)$ 经过初等行变换化成最简阶梯形矩阵, 所用的初等行变换取决于系数矩阵 $\boldsymbol{A}$, 而与常数项 $\boldsymbol{e}_i$ 无关. 将3个增广矩阵 $(\boldsymbol{A},\boldsymbol{e}_1),(\boldsymbol{A},\boldsymbol{e}_2),(\boldsymbol{A},\boldsymbol{e}_3)$ 分别通过初等行变换化成最简阶梯形矩阵 $(\boldsymbol{I},\boldsymbol{X}_1),(\boldsymbol{I},\boldsymbol{X}_2),(\boldsymbol{I},\boldsymbol{X}_3)$, 所用的初等行变换完全相同, 3 个增广矩阵共同的系数矩阵 $\boldsymbol{A}$ 被重复进行了三遍同样的初等行变换. 为了避免这种重复劳动, 我们将 3 个增广矩阵共同的系数矩阵 $\boldsymbol{A}$ 放在前面, 各不相同的第四列 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 依次放在后面, 组成 6 列的矩阵 $\boldsymbol{M}=(\boldsymbol{A},\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$. 对 $\boldsymbol{M}$ 进行初等行变换将前 3 列的矩阵 $\boldsymbol{A}$ 化为 $\boldsymbol{I}$:

$$(\boldsymbol{A},\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)\to\cdots\to(\boldsymbol{I},\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3),$$

则由后 3 列 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 化成的 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 就分别是 3 个线性方程组的解.

解 将 6 个列向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 排成的 3×6 矩阵 $\boldsymbol{M}=(\boldsymbol{A},\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)$ 经过与例 1 中完全相同的初等行变换化为最简阶梯形矩阵, 得

$$(\boldsymbol{A},\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)=\begin{pmatrix}1&1&1&1&0&0\\1&2&4&0&1&0\\1&3&9&0&0&1\end{pmatrix}\to\begin{pmatrix}1&0&0&3&-3&1\\0&1&0&-\dfrac{5}{2}&4&-\dfrac{3}{2}\\0&0&1&\dfrac{1}{2}&-1&\dfrac{1}{2}\end{pmatrix}$$

由最后一个矩阵的后3列得到 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 在基 T 下的坐标分别为 $\left(3,-\dfrac{5}{2},\dfrac{1}{2}\right)$, $(-3,4,-1)$, $\left(1,-\dfrac{3}{2},\dfrac{1}{2}\right)$. □

例4 求向量 $\boldsymbol{\beta}=(y_1,y_2,y_3)$ 在基 $T=\{\boldsymbol{\alpha}_1=(1,1,1),\ \boldsymbol{\alpha}_2=(1,2,3),\ \boldsymbol{\alpha}_3=(1,4,9)\}$ 下的坐标 (x_1,x_2,x_3).

解法 1 将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\beta}$ 分别写成列向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\boldsymbol{b}$. 将 $\boldsymbol{b}=(y_1,y_2,y_3)^{\mathrm{T}}$ 的分量 y_1,y_2,y_3 看成已知数, 解方程组 $\boldsymbol{AX}=\boldsymbol{b}$, 其中 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$. 对增广矩阵 $(\boldsymbol{A},\boldsymbol{b})$ 经过与本节例 1 完全相同的初等行变换化为最简阶梯形矩阵:

$$(\boldsymbol{A},\boldsymbol{b})=\begin{pmatrix}1&1&1&y_1\\1&2&4&y_2\\1&3&9&y_3\end{pmatrix}\to\begin{pmatrix}1&0&0&3y_1-3y_2+y_3\\0&1&0&-\dfrac{5}{2}y_1+4y_2-\dfrac{3}{2}y_3\\0&0&1&\dfrac{1}{2}y_1-y_2+\dfrac{1}{2}y_3\end{pmatrix},$$

得到 $\boldsymbol{\beta}$ 在基 T 下的坐标为 $\left(3y_1-3y_2+y_3,-\dfrac{5}{2}y_1+4y_2-\dfrac{3}{2}y_3,\dfrac{1}{2}y_1-y_2+\dfrac{1}{2}y_3\right)$.

解法 2 本节例 3 已经求出了自然基向量 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 在基 $T = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 下的坐标分别为 $\boldsymbol{P}_1 = \left(3, -\frac{5}{2}, \frac{1}{2}\right), \boldsymbol{P}_2 = (-3, 4, -1), \boldsymbol{P}_3 = \left(1, -\frac{3}{2}, \frac{1}{2}\right)$. $\boldsymbol{\beta}$ 是自然基向量 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 的线性组合

$$\boldsymbol{\beta} = (y_1, y_2, y_3) = y_1\varepsilon_1 + y_2\varepsilon_2 + y_3\varepsilon_3.$$

$\boldsymbol{\beta}$ 在 T 下的坐标也就是 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 在 T 下的坐标的相应的线性组合

$$y_1\boldsymbol{P}_1 + y_2\boldsymbol{P}_2 + y_3\boldsymbol{P}_3 = \left(3y_1 - 3y_2 + y_3, -\frac{5}{2}y_1 + 4y_2 - \frac{3}{2}y_3, \frac{1}{2}y_1 - y_2 + \frac{1}{2}y_3\right). \quad \Box$$

2.4.2 坐标变换公式

本节例 1—例 4 都是求向量在基下的坐标. 例 2 解法 1 是根据向量 $\boldsymbol{u}$ 在一组基下的坐标求 $\boldsymbol{u}$ 在另一组基下的坐标. 其余 3 个例题, 也是根据数组向量 $\boldsymbol{u}$ 在自然基下的坐标求它在另一组基下的坐标.

一般地, 我们来研究向量空间 V 中同一个向量 γ 在两组不同的基 S, T 下的坐标 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 之间的关系.

按照 γ 在基 T 下的坐标 $\boldsymbol{Y} = (y_1, \cdots, y_n)$ 将 γ 写成 $T = \{\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n\}$ 的线性组合

$$\gamma = y_1\boldsymbol{\beta}_1 + \cdots + y_n\boldsymbol{\beta}_n, \tag{2.4.3}$$

将等式中出现的向量 $\gamma, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n$ 分别用它们在基 S 下的坐标 $\boldsymbol{X}, \boldsymbol{P}_1, \cdots, \boldsymbol{P}_n$ (都写成列向量形式) 代替, 得到

$$\boldsymbol{X} = y_1\boldsymbol{P}_1 + \cdots + y_n\boldsymbol{P}_n = (\boldsymbol{P}_1, \cdots, \boldsymbol{P}_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{P}\boldsymbol{Y}, \tag{2.4.4}$$

其中 $\boldsymbol{P} = (\boldsymbol{P}_1, \cdots, \boldsymbol{P}_n)$ 是依次以 T 的各个基向量在基 S 下的坐标 $\boldsymbol{P}_1, \cdots, \boldsymbol{P}_n$ 为各列的矩阵, 称为由基 S 到 T 的**过渡矩阵** (transition matrix). 等式 (2.4.4) 给出了同一个向量在两组基 S 与 T 下的坐标 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的关系, 称为**坐标变换公式** (formula for coordinate transformation). 如果已知 $\boldsymbol{Y}$, 可以直接由矩阵乘法算出 $\boldsymbol{X} = \boldsymbol{P}\boldsymbol{Y}$. 如果已知 $\boldsymbol{X}$, 可通过解方程组 $\boldsymbol{P}\boldsymbol{Y} = \boldsymbol{X}$ 求出 $\boldsymbol{Y}$, 也可以反过来先求出基 T 到 S 的过渡矩阵 $\boldsymbol{K}$, 再由坐标变换公式 $\boldsymbol{Y} = \boldsymbol{K}\boldsymbol{X}$ 通过矩阵乘法 $\boldsymbol{K}\boldsymbol{X}$ 算出 $\boldsymbol{Y}$ (如例 4 解法 2).

过渡矩阵 $\boldsymbol{P}$ 的每一列 $\boldsymbol{P}_j$ 是 $\boldsymbol{B}$ 的一个基向量 $\boldsymbol{\beta}_j$ 在基 $S = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$ 下的坐标, 将 $\boldsymbol{A} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$ 与 $\boldsymbol{P}_j$ 相乘就得到基向量 $\boldsymbol{\beta}_j$. 矩阵 $\boldsymbol{P} =$

$(\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n)$ 可以看成是 n 个坐标 $\boldsymbol{P}_j$ 排成的"行向量"，定义 $\boldsymbol{A}$ 与这个"行向量"的乘积为

$$\boldsymbol{AP}=\boldsymbol{A}(\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n)=(\boldsymbol{AP}_1,\cdots,\boldsymbol{AP}_n).$$

乘法的法则是：将 $\boldsymbol{A}$ 分别与 $\boldsymbol{P}$ 的各分量相乘，得到的各乘积 $\boldsymbol{AP}_j$ 排成一行 $(\boldsymbol{AP}_1,\cdots,\boldsymbol{AP}_n)$ 作为 $\boldsymbol{AP}$，就好像是一个"数"与"行向量" $\boldsymbol{P}=(\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n)$ 相乘一样. 由于 $\boldsymbol{AP}$ 的每个分量 $\boldsymbol{AP}_j=\boldsymbol{\beta}_j$, $\boldsymbol{AP}=(\boldsymbol{AP}_1,\cdots,\boldsymbol{AP}_n)=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)$ 就是由各个基向量 $\boldsymbol{\beta}_j$ 排成的"行向量" $\boldsymbol{B}$.

为了便于应用，将以上结论总结为：

定义 2.4.1(过渡矩阵) 设 $S=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$ 与 $T=\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}$ 是数域 F 上 n 维向量空间 V 的两组基，T 的基向量 $\boldsymbol{\beta}_j$ $(1\leqslant j\leqslant n)$ 在基 S 下的坐标为 $\boldsymbol{P}_j=(p_{1j},\cdots,p_{nj})^{\mathrm{T}}$. 依次以 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n$ 为各列组成的方阵

$$\boldsymbol{P}=(\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n)=(p_{ij})_{n\times n}$$

称为由基 S 到 T 的**过渡矩阵** (transition matrix). □

定理 2.4.1 设 $S=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$ 与 $T=\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n\}$ 是数域 F 上 n 维向量空间 V 的两组基，由基 S 到 T 的过渡矩阵为 $\boldsymbol{P}$，即

$$(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n)\boldsymbol{P}. \qquad \text{(基变换公式)}$$

则 V 中同一个向量在两组基 S 与 T 下的坐标 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 与 $\boldsymbol{Y}=(y_1,\cdots,y_n)^{\mathrm{T}}$ 之间有关系式

$$\boldsymbol{X}=\boldsymbol{PY}=y_1\boldsymbol{P}_1+\cdots+y_n\boldsymbol{P}_n, \qquad \text{(坐标变换公式)}$$

其中 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_n$ 依次为 $\boldsymbol{P}$ 的各列. □

例 5 描述平面直角坐标系中方程 $2x^2+4xy+5y^2=6$ 的图像曲线的形状.

解 将平面直角坐标系绕原点旋转适当的角 α，将方程化为我们熟悉的曲线的标准形.

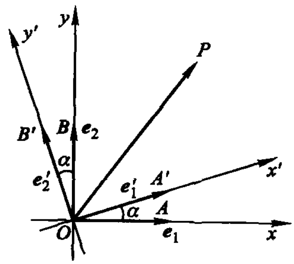

图 2–3

每一点 P 在原坐标系下的坐标 (x,y) 就是向量 $\overrightarrow{OP}$ 在自然基 $E=\{\mathbf{e}_1,\mathbf{e}_2\}$ 下的坐标，其中 $\mathbf{e}_1,\mathbf{e}_2$ 分别是 x 轴和 y 轴正方向上的单位向量 $\overrightarrow{OA},\overrightarrow{OB}$，坐标分别为 $(1,0),(0,1)$. 坐标轴 Ox,Oy 绕原点 O 沿逆时针方向旋转 α 变成 Ox',Oy'，原来的基向量 $\mathbf{e}_1,\mathbf{e}_2$ 分别变成 $\mathbf{e}_1'=\overrightarrow{OA'}$ 与 $\mathbf{e}_2'=\overrightarrow{OB'}$，如图 2–3. 点 P 的新坐标 (x',y') 就是向量 $\overrightarrow{OP}$ 在新的基 $E'=\{\mathbf{e}_1',\mathbf{e}_2'\}$ 下的坐标. 为了得到每一点 P 的原坐标 (x,y) 与新坐标 (x',y') 之间的变换公式，先求出新的基向量 $\mathbf{e}_1'=\overrightarrow{OA'}$ 与 $\mathbf{e}_2'=\overrightarrow{OB'}$ 在原来的基 E 下的坐标，也就是求点 A',B' 在原来的坐标系下的坐标 $(a_1,a_2),(b_1,b_2)$. 我们有 $\angle XOA'=\alpha$，由三角函数定义知 $\cos\alpha=\dfrac{a_1}{|OA'|}=a_1,\sin\alpha=\dfrac{a_2}{|OA'|}=a_2$，因此 A' 的原坐标为 $(\cos\alpha,\sin\alpha)$. 类似地，由 $|OB'|=1$ 及 $\angle XOB'=\dfrac{\pi}{2}+\alpha$ 知 B' 的坐标为 $\left(\cos\left(\dfrac{\pi}{2}+\alpha\right),\sin\left(\dfrac{\pi}{2}+\alpha\right)\right)=(-\sin\alpha,\cos\alpha)$. 将等式

$$\overrightarrow{OP}=x'\overrightarrow{OA'}+y'\overrightarrow{OB'}.$$

两边的向量 $\overrightarrow{OP},\overrightarrow{OA'},\overrightarrow{OB'}$ 用它们在基 $E=\{\mathbf{e}_1,\mathbf{e}_2\}$ 下的坐标代替，得到坐标变换公式

$$\begin{pmatrix} x \\ y \end{pmatrix}=x'\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}+y'\begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix},\quad \text{即}\begin{cases} x=x'\cos\alpha-y'\sin\alpha, \\ y=x'\sin\alpha+y'\cos\alpha. \end{cases}$$

代入曲线方程 $2x^2+4xy+5y^2=6$ 得

$$\begin{aligned} 2(x'\cos\alpha-y'\sin\alpha)^2+4(x'\cos\alpha-y'\sin\alpha)(x'\sin\alpha+y'\cos\alpha)& \\ +5(x'\sin\alpha+y'\cos\alpha)^2=6&, \end{aligned}$$

整理得

$$\begin{aligned} x'^2\left(\frac{7}{2}+2\sin 2\alpha-\frac{3}{2}\cos 2\alpha\right)+x'y'(3\sin 2\alpha+4\cos 2\alpha)& \\ +y'^2\left(\frac{7}{2}+\frac{3}{2}\cos 2\alpha-2\sin 2\alpha\right)=6&. \end{aligned}$$

选 α 使 $3\sin 2\alpha+4\cos 2\alpha=0$，即 $\tan 2\alpha=-\dfrac{4}{3}$. 选 2α 在第 2 象限，$\cos 2\alpha=-\dfrac{1}{\sqrt{1+\tan^2 2\alpha}}=-\dfrac{3}{5}$，$\sin 2\alpha=\dfrac{4}{5}$. 代入曲线方程，整理得

$$6x'^2+y'^2=6,\quad \text{即}\quad x'^2+\frac{y'^2}{6}=1.$$

图像曲线是椭圆. 半长轴长度为 $\sqrt{6}$，半短轴长度为 1. □

习题 2.4

1. $\mathbf{R}^3$ 中的向量 $\boldsymbol{\alpha}_1=(3,1,0),\boldsymbol{\alpha}_2=(6,3,2),\boldsymbol{\alpha}_3=(1,3,5)$ 组成向量组 S.

(1) 证明 S 是 $\mathbf{R}^3$ 的基.

(2) 求向量 $\boldsymbol{\beta}=(2,-1,2)$ 在基 S 下的坐标.

(3) 求自然基向量 $\boldsymbol{\varepsilon}_1=(1,0,0),\boldsymbol{\varepsilon}_2=(0,1,0),\boldsymbol{\varepsilon}_3=(0,0,1)$ 在基 S 下的坐标.

2. $\boldsymbol{\alpha}_1=(3,1,0),\boldsymbol{\alpha}_2=(6,3,2),\boldsymbol{\alpha}_3=(1,3,5)$ 组成 $\mathbf{R}^3$ 的一组基 S.

(1) 求 S 到 $\mathbf{R}^3$ 的自然基的过渡矩阵.

(2) 写出 $\mathbf{R}^3$ 中任意向量 (x,y,z) 在基 S 下的坐标.

3. 向量 $\boldsymbol{\alpha}_1=(1,1,1,1),\boldsymbol{\alpha}_2=(0,1,-1,-1),\boldsymbol{\alpha}_3=(0,0,1,-1),\boldsymbol{\alpha}_4=(0,0,0,1)$ 组成 $\mathbf{R}^4$ 的一组基. 求这组基到自然基的过渡矩阵 $\boldsymbol{P}$.

4. 描述平面直角坐标系中方程 $x^2+6xy+5y^2=1$ 的图像的形状.

§2.5 向量组的秩

2.5.1 极大线性无关组与秩

在本章前面几节, 我们重点研究了线性方程组有唯一解的情况. 但是, 线性方程组的解不唯一的情况同样很重要. 例如, 在空间坐标系中, 一个三元一次方程的图像通常是一个平面, 两个三元一次方程组成的方程组的图像通常是一条直线, 三个三元一次方程组成的方程组的图像在大多数情况下是一个点. 平面和直线上的点有无穷多个, 代表了方程组的无穷多个解, 图像是一个点的方程组的解才是唯一的. 然而, 图像是平面和直线的情形的重要性并不亚于图像是一个点的情形.

不过, 三个三元一次方程构成的方程组的图像并不一定都是一个点.

例 1 如下方程组在空间直角坐标系中的图像各是什么形状?

$$(1)\begin{cases} x+\ y+\ z=0,\\ 2x+\ y+5z=0,\\ 3x+2y+6z=0;\end{cases}\qquad (2)\begin{cases} x+\ y+\ z=0,\\ 2x+2y+2z=0,\\ 3x+3y+3z=0.\end{cases}$$

解 (1) 易见第 3 个方程是前两个方程之和, 前两个方程的公共解一定是第 3 个方程的解. 可以将第 3 个方程删去而不改变方程组的解. 可见, 方程组 "实质上" 只有两个方程, 方程组的图像是过原点且不重合的两个平面 $x+y+z=0$ 与 $2x+y+5z=0$ 的交集, 是一条直线.

(2) 后两个方程都是第一个方程的常数倍, 可以将后两个方程都删去而不改变方程组的解. 方程组 "实质上" 只有一个方程 $x+y+z=0$, 图像是一个平面. □

例 1 中的两个方程组看起来都有三个方程. 但这些方程线性相关, 其中有的方程是其余方程的线性组合, 可以从方程组中删去而不改变方程组的解, 这可以认为是对方程的个数进行“打假”. 将“打假”进行到底, 删去尽可能多的方程, 直到剩下的方程线性无关, 而被删去的方程都是剩下的方程的线性组合, 剩下的方程的个数才是原方程组中方程的“真正个数”. 方程组解集的大小不是由原来的方程个数决定, 而是由经过“打假”之后得到的“真正个数”决定的. 例 1(1) 的方程组剩下两个方程, 所以图像是直线. 例 1(2) 的方程组剩下一个方程, 所以图像是平面.

定义 2.5.1(极大线性无关组) 设 S 是向量组, $T=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r\}$ 是 S 的线性无关子集, 并且将 S 中任何一个 $\boldsymbol{b}\notin T$ 添加到 T 中得到的集合 $T_1=T\cup\{\boldsymbol{b}\}$ 线性相关, 就称 T 是 S 的**极大线性无关组** (maximal linearly independent system). □

引理 2.5.1 S 是它的任何一个极大线性无关组 T 的线性组合.

证明 如果某个 $\boldsymbol{b}$ 不是 $T=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r\}$ 的线性组合, 则由 T 添加 $\boldsymbol{b}$ 得到的集合 $T_1=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{b}\}$ 中的每个向量都不是它前面的向量的线性组合, T_1 线性无关. 这与 T 是极大线性无关组相矛盾. 因此, S 中所有的向量都是 T 的线性组合. □

将向量组 S 删去尽可能多的向量, 使剩下的向量组成的子集 S_0 线性无关, 并且被删去的向量都是 S_0 的线性组合, 则 S_0 就是 S 的极大线性无关组. 如果 S 是由方程构成的方程组, 由 S 删去一些方程得到极大线性无关组 S_0 的过程就是前面所说的对方程个数“打假”, S_0 所含方程个数 r 就是方程组中的方程的“真正个数”.

有一个问题值得注意: 同一个方程组 S 可能有不同的极大线性无关组. 如果某两个极大线性无关组 T_1,T_2 所含方程的个数 $|T_1|$ 与 $|T_2|$ 不同, 将其中哪一个作为方程的“真正个数”呢?

我们证明: 这样的矛盾不会发生, 同一个向量集合 S 可能有很多不同的极大线性无关组, 但所有这些极大线性无关组所含向量个数一定都相等.

根据引理 2.5.1, S 中每个向量都是它的任何一个极大线性无关组的线性组合, 因此 S 的两个极大线性无关组 T_1,T_2 互为线性组合. 我们证明: 如果 T_1 与 T_2 包含的向量个数 $|T_1|$ 与 $|T_2|$ 不相等, 含有更多向量的那个 T_i 一定线性相关.

引理 2.5.2 设向量集合 $T=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_k\}$ 是 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 的线性组合.

(1) 如果 $k>m$, 则 T 线性相关.

(2) T 中线性无关向量的个数 $r\leqslant m$.

证明 将向量组 S 用 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 表示. T 中每个向量 $\boldsymbol{b}_j$ 是

S 的线性组合, 可以写成 $\boldsymbol{b}_j = b_{1j}\boldsymbol{a}_1 + \cdots + b_{mj}\boldsymbol{a}_m = \boldsymbol{A}\boldsymbol{X}_j$ 的形式, 其中 $\boldsymbol{X}_j = (b_{1j}, \cdots, b_{mj})^{\mathrm{T}}$ 是 m 维列向量.

(1) 设 $k > m$, 则 k 个 m 维列向量 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_k$ 线性相关, 存在不全为 0 的 $\lambda_1, \cdots, \lambda_k$ 使 $\lambda_1\boldsymbol{X}_1 + \cdots + \lambda_k\boldsymbol{X}_k = \boldsymbol{0}$. 从而

$$\lambda_1\boldsymbol{b}_1 + \cdots + \lambda_k\boldsymbol{b}_k = \lambda_1(\boldsymbol{A}\boldsymbol{X}_1) + \cdots + \lambda_k(\boldsymbol{A}\boldsymbol{X}_k) = \boldsymbol{A}(\lambda_1\boldsymbol{X}_1 + \cdots + \lambda_k\boldsymbol{X}_k) = \boldsymbol{0}.$$

这说明 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_k$ 线性相关. 这就证明了结论 (1).

(2) 设 T_0 是 T 的线性无关子集, 由 r 个向量组成. T_0 当然也是 S 的线性组合. 由结论 (1) 知不可能 $r > m$ (否则 T_0 线性相关), 只能 $r \leqslant m$. □

推论 2.5.1 如果两个线性无关向量组互为线性组合, 它们包含的向量个数一定相等. 特别地, 同一个向量组的任何两个极大线性无关组包含的向量个数相等. □

定义 2.5.2(秩) 向量组 S 的任何一个极大线性无关组所含向量个数称为 S 的**秩** (rank), 记为 $\operatorname{rank} S$. 矩阵 $\boldsymbol{A}$ 的行向量组 $R(\boldsymbol{A})$ 的秩称为 $\boldsymbol{A}$ 的**行秩** (row rank), $\boldsymbol{A}$ 的列向量组 $C(\boldsymbol{A})$ 的秩称为 $\boldsymbol{A}$ 的**列秩** (column rank). □

将线性方程组的方程看成向量, 方程组看成这些向量组成的向量组 S, 则 $\operatorname{rank} S$ 就是方程组中线性无关方程的最大个数, 可以看成是方程的“真正个数”. 将每个方程用数组向量表示, 方程组用数组向量组成的向量组 S 表示, S 也就是方程组的增广矩阵 $\boldsymbol{M}$ 的行向量组. $\boldsymbol{M}$ 的行秩就是方程的“真正个数”.

由引理还可得出:

推论 2.5.2 (1) 如果向量集合 T 是 S 的线性组合, 则 $\operatorname{rank} T \leqslant \operatorname{rank} S$.

(2) 如果向量集合 S 与 T 互为线性组合, 则 $\operatorname{rank} S = \operatorname{rank} T$.

证明 设 S_0 是 S 的极大线性无关组, 则 S_0 所含向量个数 $s = \operatorname{rank} S$.

由引理 2.5.1, S 是 S_0 的线性组合. 而 T 是 S 的线性组合. 由向量的线性组合的传递性 (引理 2.3.1) 知 T 是 S_0 的线性组合. 由引理 2.5.2(2) 立即可知 T 的极大线性无关子集 T_0 所含向量个数 $t \leqslant s$. 也就是 $\operatorname{rank} T \leqslant \operatorname{rank} S$. 结论 (1) 成立.

由结论 (1) 立即得出结论 (2). □

推论 2.5.2(2) 对于计算向量组 $S = \{\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m\}$ 的秩很有用处. 将 S 中各向量 $\boldsymbol{\alpha}_i$ 写成行向量, 排成矩阵 $\boldsymbol{M}$, 则 $\boldsymbol{M}$ 的行向量组 $R(\boldsymbol{M})$ 的秩 $\operatorname{rank} R(\boldsymbol{M}) = \operatorname{rank} S$. 将 $\boldsymbol{M}$ 经过一系列初等行变换化成阶梯形矩阵 $\boldsymbol{T}$. 由于每次初等行变换前后的矩阵的行向量组互为线性组合, 秩相等. 因此 $\operatorname{rank} \boldsymbol{M} = \operatorname{rank} \boldsymbol{T}$. 阶梯形矩阵 $\boldsymbol{T}$ 的行秩很容易看出来, 由此容易得到 $\operatorname{rank} S$.

刚才这段话中证明了如下结论, 可以作为推论 2.5.2(2) 的推论:

推论 2.5.3 初等行变换不改变矩阵的行秩. □

互为线性组合的向量集合的用处很广泛, 专门给出如下定义, 以便应用:

定义 2.5.3 如果向量集合 S 与 T 互为线性组合, 就称 S 与 T **等价** (equivalent). 如果矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行向量组 $R(\boldsymbol{A})$ 与 $R(\boldsymbol{B})$ 等价, 就称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **行等价** (row equivalent). 如果矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的列向量组 $C(\boldsymbol{A})$ 与 $C(\boldsymbol{B})$ 等价, 就称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **列等价** (column equivalent). □

由线性组合的传递性 (引理 2.3.1) 立即得出向量组等价的传递性:

推论 2.5.4 如果向量组 S_1 与 S_2 等价, S_2 与 S_3 等价, 则 S_1 与 S_3 等价.

按照定义 2.5.3, 推论 2.5.2 (2) 就成为:

□

如果向量集合 S 与 T 等价, 则它们的秩相等: $\operatorname{rank} S = \operatorname{rank} T$.

注 "等价"就是"价值相等". 向量组的"价值"是用它所包含的线性无关向量的个数来衡量的.

向量组也可以是无限集合, 甚至是整个空间 V. V 的任意一组基 S 就是 V 的极大线性无关组, 基所含向量个数就是 V 的秩 $\operatorname{rank} V$, 也就是 V 的维数.

向量的个数经过线性组合可以大幅度增加甚至增加到无穷, 但线性无关向量的个数不会增加. 例如, 在几何空间中, 一个非零的几何向量经过线性组合能得到一条直线上的全体向量, 直线上的向量有无穷多, 但线性无关向量至多只有一个; 两个不共线的向量能够线性组合出一个平面上的全体向量, 其中的线性无关向量至多只有两个. 这使我们联想到一条著名的科学定律——能量守恒定律: 能量可以从一种形式转化为另一种形式, 但不能创生也不会消亡. 向量组的秩可以通过线性组合转移到另一个向量组, 但绝不会增加.

2.5.2 用初等行变换计算秩

例 2 试求线性方程组

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1, \\ x_1 + 2x_2 + 3x_3 + 4x_4 = 5, \\ x_1 + 3x_2 + 5x_3 + 7x_4 = 9, \\ x_1 + 5x_2 + 8x_3 + 10x_4 = 15, \\ x_1 + 4x_2 + 9x_3 + 16x_4 = 17 \end{cases} \tag{2.5.1}$$

的秩.

解 将线性方程组用它的增广矩阵 $\boldsymbol{M}$ 表示. 方程组的秩就是矩阵 $\boldsymbol{M}$ 的行向量组 $R(\boldsymbol{M})$ 的秩 $\operatorname{rank} R(\boldsymbol{M})$. 通过一系列初等行变换将矩阵 $\boldsymbol{M}$ 化成阶

梯形矩阵:

$$M=\begin{pmatrix}1&1&1&1&1\\1&2&3&4&5\\1&3&5&7&9\\1&5&8&10&15\\1&4&9&16&17\end{pmatrix}\to\cdots\to T=\begin{pmatrix}1&1&1&1&1\\0&1&2&3&4\\0&0&-1&-3&-2\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}$$

$\boldsymbol{T}$ 中只有 3 行不为零, $\operatorname{rank}\boldsymbol{T}\leqslant 3$. $\boldsymbol{T}$ 的前 3 行的前 3 列组成对角元全不为 0 的上三角阵 $\boldsymbol{T}_0$, $\boldsymbol{T}_0$ 的 3 行线性无关, 并组成 F^3 的基. 由 $\boldsymbol{T}_0$ 的 3 行扩充得到的 $\boldsymbol{T}$ 的前 3 行也线性无关 (引理 2.2.2). 于是 $\operatorname{rank}R(\boldsymbol{M})=\operatorname{rank}R(\boldsymbol{T})=3$. 方程组的秩为 3. □

一般地, 我们有:

算法 2.5.1(求向量组的秩) 求向量组 $S=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 的秩.

将 S 中各向量 $\boldsymbol{\alpha}_i$ 写成行向量, 以它们为各行排成矩阵 $\boldsymbol{A}$. 将矩阵 $\boldsymbol{A}$ 经过一系列初等行变换化成阶梯形矩阵 $\boldsymbol{T}$. 则 $\boldsymbol{T}$ 中非零行的个数 $r=\operatorname{rank}S$. □

例 2 的方程组有 5 个方程, 通过初等变换化简后只剩下 3 个方程. 另外的两个方程到哪里去了? 由计算的结果知道, 原方程组的秩为 3, 其中线性无关的方程只有 3 个, 另外两个方程是这 3 个方程的线性组合, 可以删去. 初等行变换并没有减少方程的 "真正个数", 只不过让原来虚假的个数 5 暴露原形, 真正的个数 3 "水落石出". 可见, 初等行变换的过程也是在对方程个数进行 "打假". 不过, 这个打假过程只是另外组合出了一个阶梯形方程组 W 与 U 等价, 由 W 的秩得到了 U 的秩. 能不能直接在原方程组 U 中找出一个极大线性无关组 U_0? 我们有下面的算法.

2.5.3 用初等行变换求极大线性无关组

例 3 求例 2 中的方程组 U 的极大线性无关组 U_0. 并将其余方程表示成 U_0 的线性组合.

分析 将方程组 U 中的 5 个方程分别用行向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_5$ 表示. 要判断其中某 3 个方程 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\boldsymbol{\alpha}_{i_3}$ 是否组成极大线性无关组, 需要做如下两件事:

首先, 解方程组

$$x_1\boldsymbol{\alpha}_{i_1}+x_2\boldsymbol{\alpha}_{i_2}+x_3\boldsymbol{\alpha}_{i_3}=\boldsymbol{0},\tag{2.5.2}$$

看它是否只有零解.

然后, 对 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\boldsymbol{\alpha}_{i_3}$ 之外的每个 $\boldsymbol{\alpha}_i$ 解方程组

$$x_1\boldsymbol{\alpha}_{i_1}+x_2\boldsymbol{\alpha}_{i_2}+x_3\boldsymbol{\alpha}_{i_3}=\boldsymbol{\alpha}_i,\tag{2.5.3}$$

求出唯一解.

我们不能预先断定哪 3 个行向量可以充当 $\alpha_{i_1},\alpha_{i_2},\alpha_{i_3}$. 不过, 不论找了哪 3 个行向量, 都必须将方程组 (2.5.2) 与 (2.5.3) 中出现的行向量 $\alpha_{i_1},\alpha_{i_2},\alpha_{i_3},\alpha_i$ 写成列向量, 排成矩阵, 对这个矩阵进行初等行变换, 将矩阵化成阶梯形矩阵. 因此, 我们先将 $\alpha_1,\cdots,\alpha_5$ 都写成列向量, 排成矩阵 $\boldsymbol{A}$, 经过初等行变换化成最简阶梯形矩阵 $\boldsymbol{\Lambda}$, 然后根据 $\boldsymbol{\Lambda}$ 的形状从 $\boldsymbol{A}$ 中选取恰当的 3 列来充当 $\alpha_{i_1},\alpha_{i_2},\alpha_{i_3}$.

解法 1 将各方程用数组向量表示为 $\alpha_1=(1,1,1,1,1)$, $\alpha_2=(1,2,3,4,5)$, $\alpha_3=(1,3,5,7,9)$, $\alpha_4=(1,5,8,10,15)$, $\alpha_5=(1,4,9,16,17)$. 将这些数组向量转置成列向量, 排成矩阵 $\boldsymbol{A}=(\alpha_1^{\mathrm{T}},\alpha_2^{\mathrm{T}},\alpha_3^{\mathrm{T}},\alpha_4^{\mathrm{T}},\alpha_5^{\mathrm{T}})$, 通过初等行变换化成最简阶梯形矩阵 $\boldsymbol{\Lambda}$, 得

$$\boldsymbol{A}=\begin{pmatrix}1&1&1&1&1\\1&2&3&5&4\\1&3&5&8&9\\1&4&7&10&16\\1&5&9&15&17\end{pmatrix}\to\boldsymbol{\Lambda}=\begin{pmatrix}1&0&-1&0&-8\\0&1&2&0&11\\0&0&0&1&-2\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

$\boldsymbol{\Lambda}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_5)$ 的 3 个 "阶梯" 分别位于第 1,2,4 列. 如果取 $\boldsymbol{A}$ 的第 1,2,4 列排成矩阵 $\boldsymbol{A}_1=(\alpha_1^{\mathrm{T}},\alpha_2^{\mathrm{T}},\alpha_4^{\mathrm{T}})$, 则 $\boldsymbol{A}_1$ 经过与 $\boldsymbol{A}$ 同样的初等行变换变成 $\boldsymbol{\Lambda}_1=(\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_4)$:

$$\boldsymbol{A}_1=(\alpha_1^{\mathrm{T}},\alpha_2^{\mathrm{T}},\alpha_4^{\mathrm{T}})\to\boldsymbol{\Lambda}_1=(\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_4)=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\\0&0&0\end{pmatrix}.$$

齐次线性方程组 $\boldsymbol{\Lambda}_1\boldsymbol{X}=\boldsymbol{0}$ 只有零解. 因此 $\boldsymbol{A}_1\boldsymbol{X}=\boldsymbol{0}$ 也只有零解, 这说明 $\boldsymbol{A}_1$ 的各列 $\alpha_1^{\mathrm{T}},\alpha_2^{\mathrm{T}},\alpha_4^{\mathrm{T}}$ 线性无关.

以 $\boldsymbol{A}_1$ 为系数矩阵的非齐次线性方程组 $\boldsymbol{A}_1\boldsymbol{X}=\alpha_3^{\mathrm{T}}$ 与 $\boldsymbol{A}_1\boldsymbol{X}=\alpha_5^{\mathrm{T}}$ 的增广矩阵 $(\boldsymbol{A}_1,\alpha_3^{\mathrm{T}})$ 与 $(\boldsymbol{A}_1,\alpha_5^{\mathrm{T}})$ 经过与 $\boldsymbol{A}$ 同样的初等行变换, 分别得到:

$$(\boldsymbol{A}_1,\alpha_3^{\mathrm{T}})\to(\boldsymbol{\Lambda}_1,\boldsymbol{b}_3)=\begin{pmatrix}1&0&0&-1\\0&1&0&2\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{pmatrix},$$

$$(\boldsymbol{A}_1, \boldsymbol{\alpha}_5^{\mathrm{T}}) \to (\boldsymbol{\varLambda}_1, \boldsymbol{b}_5) = \begin{pmatrix} 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & 11 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

因而分别得到方程组 $\boldsymbol{A}_1\boldsymbol{X} = \boldsymbol{\alpha}_3^{\mathrm{T}}$ 与 $\boldsymbol{A}_1\boldsymbol{X} = \boldsymbol{\alpha}_5^{\mathrm{T}}$ 的唯一解 $(-1,2,0)$ 与 $(-8,11,-2)$. 这说明

$$(-1)\boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 + 0\boldsymbol{\alpha}_4 = \boldsymbol{\alpha}_3, \quad (-8)\boldsymbol{\alpha}_1 + 11\boldsymbol{\alpha}_2 + (-2)\boldsymbol{\alpha}_4 = \boldsymbol{\alpha}_5.$$

由第 1,2,4 个方程组成的 $U_0 = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_4\}$ 线性无关, 并且将其余方程 $\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_5$ 写成了 U_0 的上述线性组合, 因此 U_0 是原方程组 U 的极大线性无关组. □

一般地, 将矩阵 $\boldsymbol{A} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m)$ 通过初等行变换化成 $\boldsymbol{B} = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m)$, 不但将方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 通过同解变形化成了 $\boldsymbol{\varLambda X} = \boldsymbol{0}$, 而且将 $\boldsymbol{A}$ 的任何一部分列向量决定的方程组

$$x_1\boldsymbol{a}_{i_1} + \cdots + x_k\boldsymbol{a}_{i_k} = \boldsymbol{0} \quad 与 \quad y_1\boldsymbol{a}_{i_1} + \cdots + y_k\boldsymbol{a}_{i_k} = \boldsymbol{a}_i$$

化成了由 $\boldsymbol{\varLambda}$ 的对应列决定的方程组

$$x_1\boldsymbol{b}_{i_1} + \cdots + x_k\boldsymbol{b}_{i_k} = \boldsymbol{0} \quad 与 \quad y_1\boldsymbol{b}_{i_1} + \cdots + y_k\boldsymbol{b}_{i_k} = \boldsymbol{b}_i,$$

由此可知:

$\boldsymbol{b}_{i_1}, \cdots, \boldsymbol{b}_{i_k}$ 线性相关 (或无关) $\Leftrightarrow$ $\boldsymbol{a}_{i_1}, \cdots, \boldsymbol{a}_{i_k}$ 线性相关 (或无关).

$\{\boldsymbol{b}_{i_1}, \cdots, \boldsymbol{b}_{i_k}\}$ 是 $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m\}$ 的极大线性无关组 $\Leftrightarrow$ $\{\boldsymbol{a}_{i_1}, \cdots, \boldsymbol{a}_{i_k}\}$ 是 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m\}$ 的极大线性无关组.

$$y_1\boldsymbol{b}_{i_1} + \cdots + y_k\boldsymbol{b}_{i_k} = \boldsymbol{b}_i \;\Leftrightarrow\; y_1\boldsymbol{a}_{i_1} + \cdots + y_k\boldsymbol{a}_{i_k} = \boldsymbol{a}_i.$$

如果 $\boldsymbol{B}$ 是最简阶梯形矩阵, 则很容易找出它的列向量组 $C(\boldsymbol{B})$ 的一个极大线性无关组 $\boldsymbol{T}_0$ 并将其余各列写成 $\boldsymbol{T}_0$ 的线性组合, 对 $\boldsymbol{A}$ 的列向量组可立即得出同样的结论.

按照这一思路, 例 3 的解法可以进一步简化为:

例 3 解法 2 将 $\boldsymbol{A} = (\boldsymbol{\alpha}_1^{\mathrm{T}}, \cdots, \boldsymbol{\alpha}_5^{\mathrm{T}})$ 通过初等行变换化成最简阶梯形矩阵 $\boldsymbol{\varLambda}$ (如解法 1).

$\boldsymbol{\varLambda} = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_5)$ 的各列最后两行都是 0, 其中 $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4$ 依次是前 3 个自然基向量 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, 显然它们组成 $\boldsymbol{\varLambda}$ 的列向量组 $C(\boldsymbol{\varLambda})$ 的一个极大线性无关组 T_0, 并且有

$$\boldsymbol{b}_3 = (-1)\boldsymbol{b}_1 + 2\boldsymbol{b}_2 + 0\boldsymbol{b}_4, \quad \boldsymbol{b}_5 = (-8)\boldsymbol{b}_1 + 11\boldsymbol{b}_2 + (-2)\boldsymbol{b}_4.$$

对应地, $\boldsymbol{A}$ 的第 1,2,4 列组成 $\boldsymbol{A}$ 的列向量组的极大线性无关组, 原方程组 $U=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_5\}$ 的第 1,2,4 个方程 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 组成 U 的极大线性无关组.且

$$\boldsymbol{\alpha}_3=(-1)\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2,\quad \boldsymbol{\alpha}_5=(-8)\boldsymbol{\alpha}_1+11\boldsymbol{\alpha}_2+(-2)\boldsymbol{\alpha}_4. \qquad \square$$

一般地, 我们有

算法 2.5.2 (求极大线性无关组与秩) 已知 F^n 中的向量组 $S=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$.

(1) 求 $\operatorname{rank} S$, 并求出 S 的一个极大线性无关组 S_0;

(2) 将 S 中各向量写成极大线性无关组 S_0 的线性组合.

将 S 中各向量写成列向量, 以它们为列组成矩阵 $\boldsymbol{A}$. 将 $\boldsymbol{A}$ 通过一系列初等行变换化成阶梯形矩阵 $\boldsymbol{B}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_m)$. 设其中有 r 个非零行, 各行的阶梯元依次在第 $j_1,\cdots,j_r$ 列. 则:

(1) $\operatorname{rank} S=r$, $S_0=\{\boldsymbol{\alpha}_{j_1},\cdots,\boldsymbol{\alpha}_{j_r}\}$ 是 S 的一个极大线性无关组.

(2) 如果需要将 S 中各向量写成 S_0 的线性组合, 将 $\boldsymbol{B}$ 进一步化成最简阶梯形矩阵 $\boldsymbol{A}=(\lambda_{ij})_{n\times m}$ 使它的第 $j_1,\cdots,j_r$ 列依次为自然基向量 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r$,则

$$\boldsymbol{\alpha}_j=\lambda_{1j}\boldsymbol{\alpha}_{j_1}+\cdots+\lambda_{rj}\boldsymbol{\alpha}_{j_r}. \qquad \square$$

引理 2.5.3 初等行变换不改变矩阵的行秩与列秩, 矩阵的行秩与列秩相等.

证明 初等行变换前后的矩阵的行向量组等价, 行秩相等.

初等行变换将矩阵列向量组的极大线性无关组化为极大线性无关组, 当然保持列秩.

通过初等行变换将矩阵 $\boldsymbol{A}$ 变成阶梯形矩阵 $\boldsymbol{T}$. 阶梯形矩阵 $\boldsymbol{T}$ 的行秩与列秩都等于 $\boldsymbol{T}$ 中非零行的个数 r, 也就是 $\boldsymbol{T}$ 中阶梯元的个数. 因而 $\boldsymbol{A}$ 的行秩与列秩也都等于 r. $\square$

既然矩阵 $\boldsymbol{A}$ 的行秩与列秩相等, 我们将行秩与列秩统称为**矩阵 $\boldsymbol{A}$ 的秩** (rank of matrix $\boldsymbol{A}$), 记为 $\operatorname{rank}\boldsymbol{A}$.

例 4 试将 F^4 中的线性无关向量组 $\boldsymbol{\alpha}_1=(1,1,1,1),\boldsymbol{\alpha}_2=(1,2,3,3)$ 添加一些向量扩充为 F^4 的一组基.

解 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 添加 F^4 的自然基 $E=\{\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4\}$ 得到 $S=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4\}$, 则 E 是 S 的极大线性无关组. 按照算法 2.5.2 求出 S 的极大线性无关组 S_0 包含 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$. 则 S_0 与 E 等价, 就是由 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 扩充成的 F^4 的基.

将 S 中各向量写成列向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$, 依次排成 4×6 矩阵 $\boldsymbol{M}$, 经过初等行变换化成阶梯形矩阵 $\boldsymbol{T}$：

$$\boldsymbol{M}=\begin{pmatrix}1&1&1&0&0&0\\1&2&0&1&0&0\\1&3&0&0&1&0\\1&3&0&0&0&1\end{pmatrix}\to\boldsymbol{T}=\begin{pmatrix}1&1&1&0&0&0\\0&1&-1&1&0&0\\0&0&1&-2&1&0\\0&0&0&0&-1&1\end{pmatrix}.$$

$\boldsymbol{T}$ 的第 1,2,3,5 列组成 $\boldsymbol{T}$ 的列向量组 $C(\boldsymbol{T})$ 的极大线性无关组.

由此可知: $\boldsymbol{M}$ 的第 1,2,3,5 列 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{e}_1, \boldsymbol{e}_3$ 组成 $\boldsymbol{M}$ 的列向量组 $C(\boldsymbol{M})$ 的极大线性无关组, 从而组成 F^4 的基.

因此, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 添加自然基向量 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_3$ 扩充为 F^4 的基. □

例 4 的算法和结论可以推广到任意 n 维向量空间 V 中的任意线性无关向量组:

引理 2.5.4 n 维向量空间 V 中任意线性无关子集 S 可以扩充为 V 的基.

证明 设 B 是 V 的一组基. 按算法 2.5.2 可以求出 $S\cup B$ 的极大线性无关组 B_1 包含 S. B_1 就是由 S 扩充得到的 V 的基. □

习题 2.5

1. 求由以下每个小题中的向量组成的向量组的秩, 并求出一个极大线性无关组.

(1) $\boldsymbol{\alpha}_1=(6,4,1,-1,2)$, $\boldsymbol{\alpha}_2=(1,0,2,3,4)$, $\boldsymbol{\alpha}_3=(1,4,-9,-16,22)$, $\boldsymbol{\alpha}_4=(7,1,0,-1,3)$;

(2) $\boldsymbol{\alpha}_1=(1,\ -1,\ 2,\ 4)$, $\boldsymbol{\alpha}_2=(0,\ 3,\ 1,\ 2)$, $\boldsymbol{\alpha}_3=(3,\ 0,\ 7,\ 14)$, $\boldsymbol{\alpha}_4=(1,\ -1,\ 2,\ 0)$, $\boldsymbol{\alpha}_5=(2,\ 1,\ 5,\ 6)$.

2. 设 $\boldsymbol{\alpha}_1=(0,1,2,3), \boldsymbol{\alpha}_2=(1,2,3,4), \boldsymbol{\alpha}_3=(3,4,5,6), \boldsymbol{\alpha}_4=(4,3,2,1), \boldsymbol{\alpha}_5=(6,5,4,3)$. 将 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 扩充成 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_5\}$ 的一个极大线性无关组.

3. 求下列矩阵的秩. 并求出它们的行向量组和列向量组的一个极大线性无关组.

(1) $\begin{pmatrix}2&-1&-1\\-1&2&-1\\-1&-1&2\end{pmatrix}$; (2) $\begin{pmatrix}1&1&1&1&1\\1&2&3&4&5\\5&4&3&1&2\end{pmatrix}$.

4. 设复数域上线性空间 V 中的向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 线性无关. 对复数 λ 的不同值, 判断向量组 $S=\{\boldsymbol{\alpha}_1+\lambda\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-1}+\lambda\boldsymbol{\alpha}_n,\boldsymbol{\alpha}_n+\lambda\boldsymbol{\alpha}_1\}$ 是否线性无关, 并求 S 的秩.

§2.6 子 空 间

2.6.1 子集生成的子空间

例 1 设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 是线性无关的向量. 对不同的实数值 λ, 求向量组 $S = \{\lambda\boldsymbol{a}_1 + \boldsymbol{a}_2,\ \lambda\boldsymbol{a}_2 + \boldsymbol{a}_3,\ \lambda\boldsymbol{a}_3 + \boldsymbol{a}_1\}$ 的秩.

分析 如果 $T = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ 是向量空间的基, 则 S 所含的三个向量可以分别由它们的坐标 $\boldsymbol{X}_1 = (\lambda, 1, 0), \boldsymbol{X}_2 = (0, \lambda, 1), \boldsymbol{X}_3 = (1, 0, \lambda)$ "全权代表". 很容易用 §2.5 的算法求出 $S_1 = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3\}$ 的秩, 就得到了 $\operatorname{rank} S = \operatorname{rank} S_1$.

已经知道 T 线性无关, 但 T 不一定能将整个空间 V 中所有的向量都线性组合出来, 因此不一定是空间 V 的基. 考虑由 T 的全体实系数线性组合构成的集合

$$W = \{x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 \mid x_1, x_2, x_3 \in \mathbf{R}\} = \{\boldsymbol{A}\boldsymbol{X} \mid \boldsymbol{X} \in \mathbf{R}^{3\times 1}\},$$

其中 $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$. 则 W 中任何两个向量 $\boldsymbol{A}\boldsymbol{X}_1, \boldsymbol{A}\boldsymbol{X}_2$ 之和 $\boldsymbol{A}\boldsymbol{X}_1 + \boldsymbol{A}\boldsymbol{X}_2 = \boldsymbol{A}(\boldsymbol{X}_1 + \boldsymbol{X}_2)$ 仍在 W 中, 任何一个 $\boldsymbol{A}\boldsymbol{X} \in W$ 与实数 λ 之积 $(\boldsymbol{A}\boldsymbol{X})\lambda = \boldsymbol{A}(\boldsymbol{X}\lambda)$ 仍在 W 中, 这说明 W 对加法与数乘运算封闭, 并且这两个运算在 W 的范围内仍满足 8 条运算律. 因此 W 本身就是实数域 $\mathbf{R}$ 上的向量空间, T 是它的基. S 中的向量都是 T 的线性组合, 都在 W 中. 只要在 W 中求出 S 的秩 r, 在更大的范围 V 内 S 的秩仍是 r.

解 记 $W = \{x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 \mid x_1, x_2, x_3 \in \mathbf{R}\}$ 是 T 在实数域 $\mathbf{R}$ 上的全体线性组合组成的集合. 则 W 是 $\mathbf{R}$ 上的一个向量空间. $T = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ 是 W 的一组基.

S 中 3 个向量可以由它们在 T 下的坐标 $\boldsymbol{X}_1 = (\lambda, 1, 0)^{\mathrm{T}}$, $\boldsymbol{X}_2 = (0, \lambda, 1)^{\mathrm{T}}$, $\boldsymbol{X}_3 = (1, 0, \lambda)^{\mathrm{T}}$ 代表. 以 $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3$ 为各列排成矩阵 $\boldsymbol{M}$, 通过初等行变换化成阶梯形矩阵:

$$\boldsymbol{M} = \begin{pmatrix} \lambda & 0 & 1 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} \xrightarrow{(2)+(1),(3)+(1)} \boldsymbol{M}_1 = \begin{pmatrix} \lambda+1 & \lambda+1 & \lambda+1 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}.$$

当 $\lambda = -1$,

$$\boldsymbol{M}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

的第一行为零, 后两行线性无关, 行秩为 2, 列秩也为 2, 因此 $\operatorname{rank} S = 2$.

当 $\lambda \neq -1$,

$$M_1 \xrightarrow{\frac{1}{\lambda+1}(1)} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} \xrightarrow{-(1)+(2)} \begin{pmatrix} 1 & 1 & 1 \\ 0 & \lambda-1 & -1 \\ 0 & 1 & \lambda \end{pmatrix}$$

$$\xrightarrow{(1-\lambda)(3)+(2),(2,3)} B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \lambda \\ 0 & 0 & -\lambda^2+\lambda-1 \end{pmatrix}.$$

当 λ 为实数时, $-\lambda^2+\lambda-1 \neq 0$, B 是对角元全不为 0 的上三角形矩阵, 各列线性无关, 列秩为 3. 因此 S 的秩为 3. □

如果例 1 中的 λ 允许在复数范围内任意取值, 则当 $-\lambda^2+\lambda-1=0$ 时 $\operatorname{rank} S$ 也是 2.

定义 2.6.1 设 V 是数域 F 上的向量空间. V 的非空子集 W 如果满足以下两个条件:

(1) $\boldsymbol{u}, \boldsymbol{v} \in W \quad \Rightarrow \quad \boldsymbol{u}+\boldsymbol{v} \in W$,

(2) $\boldsymbol{u} \in W, \lambda \in F \quad \Rightarrow \quad \lambda\boldsymbol{u} \in W$,

就称 W 是 F^n 的**子空间** (subspace). □

V 中的加法与数乘满足 8 条运算律. W 中的加法与数乘运算法则是从 V 中照搬过来的, 显然满足 8 条运算律中的 (1),(2) 以及 (5)— (8). 非空子集 W 至少包含一个 $\boldsymbol{u} \in V$, 于是 W 包含 $0\boldsymbol{u} = \boldsymbol{0}$, 这说明运算律 (3) 成立. 对每个 $\boldsymbol{u} \in W$, 有 $(-1)\boldsymbol{u} = -\boldsymbol{u} \in W$, 这说明运算律 (4) 成立. 可见, V 的子空间 W 也可以按 V 中的加法与数乘成为一个独立的向量空间. 关于向量空间的定义与定理也都适用于子空间, 例如 W 也有维数、基、坐标等等.

定义 2.6.2 子空间 W 包含的线性无关向量的最大个数称为 W 的维数, 记作 $\dim W$. □

注 向量空间 V 有两个显然的子空间: V 与 $\{\boldsymbol{0}\}$. 由零向量 $\boldsymbol{0}$ 单独组成的空间 $\{\boldsymbol{0}\}$, 称为零空间, 通常记作 O. 零空间仅有的一个向量是零向量且线性相关, 求极大线性无关组时应当删去. 因此零空间 O 的极大线性无关组是空集合 $\varnothing$, 基也是空集合 $\varnothing$, 包含元素个数是 0, 因此零空间的维数 $\dim O = 0$. 按照规定 (参见附录 1), 空集合 $\varnothing$ 的线性组合不是空集合而是零向量, 因此它确实是零空间的基.

例 2 将 3 维几何空间 $V = \mathbf{R}^3$ 中每个向量 $\boldsymbol{\alpha} = \overrightarrow{OA}$ 用点 A 表示. 则 V 的每个一维子空间 W 是过原点的一条直线, 二维子空间 W 是过原点的一个平面, 零维子空间 W 仅由原点组成, 三维子空间 W 是整个空间 V. □

引理 2.6.1 设 W 是 F^n 的 r 维子空间. T 是 W 的子集. 则

T 是 W 的基 $\Leftrightarrow T$ 是 W 的极大线性无关组 $\Leftrightarrow$ T 由 r 个线性无关向量组成. □

如果 V 的两个子空间 W_1 与 W_2 有包含关系 $W_1\subseteq W_2$, 显然 W_1 也是 W_2 的子空间.

引理 2.6.2 设 W_1 是 W_2 的子空间. 则

(1) $\dim W_1\leqslant \dim W_2$.

(2) $\dim W_1=\dim W_2 \ \Leftrightarrow\ W_1=W_2$.

证明 设 $T=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r\}$ 是 W_1 的一组基. 则 $r=\dim W_1$.

(1) T 也是 W_2 的线性无关向量组, 因此 $\dim W_1=r\leqslant \dim W_2$.

(2) 显然 $W_1=W_2\ \Rightarrow\ \dim W_1=\dim W_2$. 只需证 $\dim W_1=\dim W_2\Rightarrow W_1=W_2$.

设 $\dim W_1=\dim W_2=r$, 则 W_1 的基 $T=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r\}$ 也是 W_2 的基. W_1 与 W_2 都由 T 的全体线性组合组成, 因此 $W_1=W_2$. □

注 引理 2.6.2 (2) 中的 $\dim W_1=\dim W_2\Rightarrow W_1=W_2$ 是在 $W_1\subseteq W_2$ 的前提下成立的. 如果没有这个前提, 则可能有不同的 W_1,W_2 具有相同的维数. 例如, 例 2 中所举的空间 V 中过原点的不同直线是不同的子空间, 但维数都等于 1.

引理 2.6.3 数域 F 上的向量空间 V 的任何一个子集 S 的全体线性组合组成的集合

$$W=\{x_1\boldsymbol{a}_1+\cdots+x_m\boldsymbol{a}_m\mid \boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\in S,\ x_1,\cdots,x_m\in F,\ m\text{为正整数}\}$$

是 V 的子空间, 并且是 V 中包含 S 的最小的子空间.

证明 S 的任意两个线性组合 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 之和 $\boldsymbol{b}_1+\boldsymbol{b}_2$ 及任意常数倍 $\lambda\boldsymbol{b}_1$ $(\lambda\in F)$ 仍是 S 的线性组合, 这说明 W 对加法与数乘封闭, 是 V 的子空间.

如果 V 的子空间 W_1 包含 S, 则 W_1 包含 S 中任何两个向量 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 之和 $\boldsymbol{a}_1+\boldsymbol{a}_2$ 及任意常数倍 $\lambda\boldsymbol{a}_1$, 从而包含 S 中任意有限个向量的线性组合, 也就是包含 W. 这说明 W 是 V 中包含 S 的最小子空间. □

注 我们允许 S 是 F^n 的无限集合. 此时 S 的全体线性组合就是 S 中全体有限子集的全部线性组合.

定义 2.6.3(子集生成的子空间) 由数域 F 上向量空间 V 的子集 S 的全体线性组合组成的子空间称为 S **生成的子空间** (subspace generated by S), 记作 $L(S)$.

例 3 设 V 是数域 F 上 n 维向量空间. 证明下列命题:

(1) V 的子集 S 的每个极大线性无关组是 $L(S)$ 的一组基, 且 $\dim L(S) = \operatorname{rank} S$.

(2) V 的两个子集 S 与 T 等价 $\Leftrightarrow L(S) = L(T)$.

证明 (1) 设 $S_0 = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m\}$ 是 S 的极大线性无关组. 则 S_0 的线性组合 S 的线性组合 $L(S)$ 是 S_0 的线性组合. S_0 是 $L(S)$ 的极大线性无关组, 因而是 $L(S)$ 的基. S_0 所含元素个数 $r = \operatorname{rank} S = \dim L(S)$.

(2) S 与 T 等价 $\Leftrightarrow \{T \subseteq L(S)$ 且 $S \subseteq L(T)\}$

$$\Leftrightarrow \{L(T) \subseteq L(S) \text{且} L(S) \subseteq L(T)\} \Leftrightarrow L(S) = L(T). \qquad \square$$

2.6.2 齐次线性方程组的解空间

例 4 求证: 数域 F 上 n 元齐次线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 的全体解组成的集合 $V_{\boldsymbol{A}}$ 是 F^n 的子空间.

证明 $\boldsymbol{X}_1, \boldsymbol{X}_2 \in V_{\boldsymbol{A}} \Rightarrow \left\{\begin{array}{l} \boldsymbol{AX}_1 = \boldsymbol{0} \\ \boldsymbol{AX}_2 = \boldsymbol{0} \end{array}\right\} \Rightarrow$

$$\left\{\begin{array}{l} \boldsymbol{A}(\boldsymbol{X}_1 + \boldsymbol{X}_2) = \boldsymbol{AX}_1 + \boldsymbol{AX}_2 = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0} \\ \boldsymbol{A}(\lambda \boldsymbol{X}_1) = \lambda(\boldsymbol{AX}_1) = \lambda \boldsymbol{0} = \boldsymbol{0} \ (\forall \lambda \in F) \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \boldsymbol{X}_1 + \boldsymbol{X}_2 \in V_{\boldsymbol{A}} \\ \lambda \boldsymbol{X}_1 \in V_{\boldsymbol{A}} \ (\forall \lambda \in F) \end{array}\right\},$$

这证明了 $V_{\boldsymbol{A}}$ 是子空间. $\square$

齐次线性方程组的解集称为解空间 (subspace of solutions).

例 5 求下列齐次线性方程组在实数域 $\mathbf{R}$ 上的解空间 W 的维数及一组基.

(1) 三元一次方程 $x + y + z = 0$;

(2) 三元一次方程组 $\left\{\begin{array}{l} x + y + z = 0, \\ 2x + y + 5z = 0; \end{array}\right.$

(3) 线性方程组

$$\left\{\begin{array}{l} x_1 + 2x_2 - 5x_4 + x_5 = 0, \\ x_1 + 2x_2 + 3x_3 + 4x_4 - 5x_5 = 0, \\ x_1 + 2x_2 + 2x_3 + x_4 - 3x_5 = 0, \\ 2x_1 + 4x_2 - 3x_3 - 19x_4 + 8x_5 = 0, \\ 3x_1 + 6x_2 - 3x_3 - 24x_4 + 9x_5 = 0. \end{array}\right.$$

解 (1) 方程 $x+y+z=0$ 的通解是

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -t_1-t_2 \\ t_1 \\ t_2 \end{pmatrix} = t_1\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t_2\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix},$$

分别取 $(t_1,t_2)=(1,0)$, $(t_1,t_2)=(0,1)$, 得到两个解

$$\boldsymbol{X}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{X}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

解空间 $W=\{t_1\boldsymbol{X}_1+t_2\boldsymbol{X}_2 \mid t_1,t_2\in\mathbf{R}\}=L(\boldsymbol{X}_1,\boldsymbol{X}_2)$. 易见 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关, 组成 W 的一组基. $\dim W=2$.

(2) 解方程组得通解为

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -4t \\ 3t \\ t \end{pmatrix} = t\begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix},$$

取 $t=1$ 得一个解

$$\boldsymbol{X}_1 = \begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix}.$$

解空间 $W=\{t\boldsymbol{X}_1 \mid t_1\in\mathbf{R}\}=L(\boldsymbol{X}_1)$. 非零向量 $\boldsymbol{X}_1$ 组成 W 的一组基. $\dim W=1$.

(3) 方程组的系数矩阵 $\boldsymbol{A}$ 经过一系列初等行变换化为最简阶梯形矩阵 $\boldsymbol{\Lambda}$:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 & -5 & 1 \\ 1 & 2 & 3 & 4 & -5 \\ 1 & 2 & 2 & 1 & -3 \\ 2 & 4 & -3 & -19 & 8 \\ 3 & 6 & -3 & -24 & 9 \end{pmatrix} \to \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 2 & 0 & -5 & 1 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

原方程组化简为以 $\boldsymbol{\Lambda}$ 为系数矩阵的齐次线性方程组

$$\begin{cases} x_1+2x_2-5x_4+x_5=0, \\ x_3+3x_4-2x_5=0, \end{cases} \quad \text{即} \quad \begin{cases} x_1=-2x_2+5x_4-x_5, \\ x_3=-3x_4+2x_5, \end{cases}$$

通解为

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=\begin{pmatrix}-2t_1+5t_2-t_3\\t_1\\-3t_2+2t_3\\t_2\\t_3\end{pmatrix}=t_1\begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix}+t_2\begin{pmatrix}5\\0\\-3\\1\\0\end{pmatrix}+t_3\begin{pmatrix}-1\\0\\2\\0\\1\end{pmatrix}.$$

解空间 $W=\{t_1\boldsymbol{X}_1+t_2\boldsymbol{X}_2+t_3\boldsymbol{X}_3\mid t_1,t_2,t_3\in\mathbf{R}\}=L(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)$ 由

$$\boldsymbol{X}_1=\begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix},\quad \boldsymbol{X}_2=\begin{pmatrix}5\\0\\-3\\1\\0\end{pmatrix},\quad \boldsymbol{X}_3=\begin{pmatrix}-1\\0\\2\\0\\1\end{pmatrix}$$

的全体线性组合组成. 在每个线性组合 $\boldsymbol{X}=t_1\boldsymbol{X}_1+t_2\boldsymbol{X}_2+t_3\boldsymbol{X}_3$ 中, 线性组合系数 t_1,t_2,t_3 是 $\boldsymbol{X}$ 的某几个分量 (第 2,4,5 分量), 因此: $\boldsymbol{X}=\mathbf{0}\Leftrightarrow t_1=t_2=t_3=0$. 这说明 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 线性无关, 组成 W 的一组基. $\dim W=3$. □

例 5(1) 的 2 维解空间的图像是 3 维几何空间 $\mathbf{R}^3$ 中过原点的一个平面, 例 5(2) 的 1 维解空间的图像是过原点的一条直线. 例 5(3) 的解空间则是 5 维空间 $\mathbf{R}^5$ 的一个 3 维子空间.

例 5(2) 的方程组包括了例 5(1) 的方程, 因此例 5(2) 的解空间是例 5(1) 的解空间的子空间, 例 5(2) 的图像直线被例 5(1) 的图像平面所包含.

在以上各题中可以看到, 解空间的维数等于通解中可以自由取值的未知数的个数 $n-r$, 其中 r 是系数矩阵化成阶梯形矩阵之后的非零行的个数, 也就是系数矩阵 $\boldsymbol{A}$ 的秩.

这一结论及其推理过程可以推广到一般的齐次线性方程组.

定理 2.6.1(齐次线性方程组解空间的维数) 齐次线性方程组 $\boldsymbol{AX}=\mathbf{0}$ 的解空间 $V_{\boldsymbol{A}}$ 的维数

$$\dim V_{\boldsymbol{A}}=n-\operatorname{rank}\boldsymbol{A}.$$

证明 略去, 参见附录 2. □

齐次线性方程组的解空间的一组基称为这个方程组的一个**基础解系** (system of fundamental solutions).

比如, 例 5 的三个小题求出的向量组 $\{\boldsymbol{X}_1,\boldsymbol{X}_2\}$, $\{\boldsymbol{X}_1\}$, $\{\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3\}$ 分别是各方程组的基础解系.

注 方程组 $\boldsymbol{AX}=\mathbf{0}$ 的解集的大小是未知数个数与方程个数"对抗"的结果. 解集 $V_{\boldsymbol{A}}$ 是 F^n 的子集, 通常是无穷集合, 不宜用元素个数来衡量大小, 而用其中的线性无关的向量个数——维数 $\dim V_{\boldsymbol{A}}$ 来衡量. 方程个数可能有假, 也用其中线性无关方程的个数——秩 $\operatorname{rank}\boldsymbol{A}$ 来衡量. 每个方程好比一次"考试", F^n 好比"考生". 参加考试的"考生" F^n 有 n 维, 每次"真正"的考试淘汰一批不合格的考生, 维数减少 1. 经过 r 次"真正"的考试, 最后通过全部考试的 $V_{\boldsymbol{A}}$ 还剩 $n-r$ 维. 这样的公式 $\dim V_{\boldsymbol{A}}=n-\operatorname{rank}\boldsymbol{A}$ 合乎情理.

例 6 已知 F^5 中的向量

$$\boldsymbol{X}_1=(1,2,3,4,5),\quad \boldsymbol{X}_2=(1,3,2,1,2).$$

求一个齐次线性方程组, 使 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 组成这个方程组的基础解系.

解 将 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 代入方程组中的每个方程

$$a_1x_1+a_2x_2+a_3x_3+a_4x_4+a_5x_5=0,$$

得到的等式

$$\begin{cases}a_1+2a_2+3a_3+4a_4+5a_5=0,\\ a_1+3a_2+2a_3+\ \ a_4+2a_5=0\end{cases}\tag{2.6.1}$$

都成立, 可见, 所求方程组中每个方程的系数组 $(a_1,\cdots,a_5)$ 都是方程组 (2.6.1) 的解.

方程组 (2.6.1) 的系数矩阵

$$\boldsymbol{B}=\begin{pmatrix}1&2&3&4&5\\1&3&2&1&2\end{pmatrix}=\begin{pmatrix}\boldsymbol{X}_1\\ \boldsymbol{X}_2\end{pmatrix}$$

就是以 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 为两行组成的矩阵. 对 $\boldsymbol{B}$ 作初等行变换化成最简阶梯形矩阵,得

$$\boldsymbol{B}\to\begin{pmatrix}1&2&3&4&5\\0&1&-1&-3&-3\end{pmatrix}\to\boldsymbol{A}=\begin{pmatrix}1&0&5&10&11\\0&1&-1&-3&-3\end{pmatrix}.$$

方程组 (2.6.1) 化为 $\boldsymbol{Aa}=\mathbf{0}$, 即

$$\begin{cases}a_1=-5a_3-10a_4-11a_5,\\ a_2=a_3+3a_4+3a_5,\end{cases}$$

得通解

$$\begin{aligned}(a_1,a_2,a_3,a_4,a_5)&=(-5a_3-10a_4-11a_5,a_3+3a_4+3a_5,a_3,a_4,a_5)\\&=a_3(-5,1,1,0,0)+a_4(-10,3,0,1,0)+a_5(-11,3,0,0,1).\end{aligned}$$

方程组 (2.6.1) 的一个基础解系是

$$(-5,1,1,0,0),\quad (-10,3,0,1,0),\quad (-11,3,0,0,1).$$

以各基础解为各行组成矩阵

$$\boldsymbol{A}=\begin{pmatrix} -5 & 1 & 1 & 0 & 0 \\ -10 & 3 & 0 & 1 & 0 \\ -11 & 3 & 0 & 0 & 1 \end{pmatrix},$$

则 $\operatorname{rank}\boldsymbol{A}=3$. 以 $\boldsymbol{A}$ 为系数矩阵的齐次线性方程组

$$\begin{cases} -\ 5x_1+\ \ x_2+x_3=0, \\ -10x_1+3x_2+\qquad x_4=0, \\ -11x_1+3x_2+\qquad\qquad x_5=0 \end{cases} \tag{2.6.2}$$

的解空间的维数为 $5-\operatorname{rank}\boldsymbol{A}=5-3=2$. 而 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是方程组 (2.6.2) 的两个线性无关解, 因此组成 (2.6.2) 的基础解系.

因此, 方程组 (2.6.2) 符合要求. □

2.6.3 非齐次线性方程组

有解条件

定理 2.6.2 线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 有解 $\Leftrightarrow \operatorname{rank}\boldsymbol{A}=\operatorname{rank}(\boldsymbol{A},\boldsymbol{b})$.

证明 将方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 写成向量形式

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}, \tag{2.6.3}$$

其中 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 依次是系数矩阵 $\boldsymbol{A}$ 的各列. 记 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$, $S_1=S\cup\{\boldsymbol{b}\}$. 则 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}S$, $\operatorname{rank}(\boldsymbol{A},\boldsymbol{b})=\operatorname{rank}S_1$.

显然 S 是 S_1 的线性组合, 因而 $\operatorname{rank}S\leqslant\operatorname{rank}S_1$.

(2.6.3) 有解 $\Rightarrow$ $\boldsymbol{b}$ 是 S 的线性组合 $\Rightarrow$ S_1 是 S 的线性组合 $\Rightarrow$ S 与 S_1 等价 $\Rightarrow \operatorname{rank}S=\operatorname{rank}S_1\Rightarrow\operatorname{rank}\boldsymbol{A}=\operatorname{rank}(\boldsymbol{A},\boldsymbol{b})$.

设 $S_0=\{\boldsymbol{a}_{i_1},\cdots,\boldsymbol{a}_{i_r}\}$ 是 S 的极大线性无关组. 如果 $\operatorname{rank}S_1=\operatorname{rank}S=r$, 则线性无关子集 S_0 也是 $S_1=S\cup\{\boldsymbol{b}\}$ 的极大线性无关组, 因而 $\boldsymbol{b}$ 是 S_0 的线性组合, 当然也是 S 的线性组合. 这说明方程组 (2.6.3) 有解. □

解集的结构

设 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 是非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的任意两个解 $\boldsymbol{X}_1,\boldsymbol{X}_2$. 则将 $\boldsymbol{AX}_1=\boldsymbol{b}$ 与 $\boldsymbol{AX}_2=\boldsymbol{b}$ 两式相减得 $\boldsymbol{A}(\boldsymbol{X}_2-\boldsymbol{X}_1)=\boldsymbol{0}$, 可见 $\boldsymbol{X}_0=\boldsymbol{X}_2-\boldsymbol{X}_1$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解, $\boldsymbol{X}_2=\boldsymbol{X}_1+\boldsymbol{X}_0$.

反过来, 设 $\boldsymbol{X}_1$ 是 $\boldsymbol{AX}=\boldsymbol{b}$ 的解, $\boldsymbol{X}_0$ 是 $\boldsymbol{AX}=\boldsymbol{0}$ 的解, 则将 $\boldsymbol{AX}_1=\boldsymbol{b}$ 与 $\boldsymbol{AX}_0=\boldsymbol{0}$ 相加得 $\boldsymbol{A}(\boldsymbol{X}_1+\boldsymbol{X}_0)=\boldsymbol{b}$, 可见 $\boldsymbol{X}_2=\boldsymbol{X}_1+\boldsymbol{X}_0$ 是 $\boldsymbol{AX}=\boldsymbol{b}$ 的解.

定理 2.6.3 如果非齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 有一个解 $\boldsymbol{X}_1$, 则:

(1) $\boldsymbol{AX}=\boldsymbol{b}$ 的解集为 $\{\boldsymbol{X}_1+\boldsymbol{X}_0 \mid \boldsymbol{AX}_0=\boldsymbol{0}\}$.

(2) 设 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_{n-r}$ 是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的一个基础解系, 则 $\boldsymbol{AX}=\boldsymbol{b}$ 的解集为

$$\{\boldsymbol{X}_1+t_1\boldsymbol{Y}_1+\cdots+t_{n-r}\boldsymbol{Y}_{n-r} \mid t_1,\cdots,t_{n-r}\in F\}. \qquad \square$$

只要得到 $\boldsymbol{AX}=\boldsymbol{b}$ 的一个解 $\boldsymbol{X}_1$, 将它与 $\boldsymbol{AX}=\boldsymbol{0}$ 的全部解相加就得到 $\boldsymbol{AX}=\boldsymbol{b}$ 的全部解. $\boldsymbol{X}\mapsto\boldsymbol{X}+\boldsymbol{X}_1$ 建立了 $\boldsymbol{AX}=\boldsymbol{0}$ 的解集与 $\boldsymbol{AX}=\boldsymbol{b}$ 的解集的 1–1 对应, 说明两个方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{AX}=\boldsymbol{b}$ 的解集的大小是相同的.

例如, 在空间直角坐标系中, 三元一次方程 $Ax+By+Cz=0$ 的图像是过原点的平面 π_0. 在三元一次非齐次方程 $Ax+By+Cz=D$ 的图像上任取一点 $\boldsymbol{P}_1(x_1,y_1,z_1)$. 将平面 π_0 上每一点 $\boldsymbol{P}(x,y,z)$ 的坐标加上 (x_1,y_1,z_1), 也就是将平面 π_0 平移使圆点移动到 $\boldsymbol{P}_1$, 则 π_0 平移得到的平面 π_1 就是 $Ax+By+Cz=D$ 的图像.

例 7 设 4 元线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}$ 的秩等于 3. $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 是它的 3 个解,

$$\boldsymbol{X}_1=(1,-2,-3,4), \quad 5\boldsymbol{X}_2-2\boldsymbol{X}_3=(2,0,1,0).$$

求这个线性方程组的通解.

解 齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间 $V_{\boldsymbol{A}}$ 的维数 $\dim V_{\boldsymbol{A}}=4-\operatorname{rank}\boldsymbol{A}=4-3=1$.

方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的任意两个解的差是齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解. 因此 $V_{\boldsymbol{A}}$ 包含 $\boldsymbol{X}_2-\boldsymbol{X}_1,\boldsymbol{X}_3-\boldsymbol{X}_1$, 从而包含它们的线性组合

$$\begin{aligned}\boldsymbol{Y}&=5(\boldsymbol{X}_2-\boldsymbol{X}_1)-2(\boldsymbol{X}_3-\boldsymbol{X}_1)=(5\boldsymbol{X}_2-2\boldsymbol{X}_3)-3\boldsymbol{X}_1\\&=(2,0,1,0)-3(1,-2,-3,4)=(-1,6,10,-12),\end{aligned}$$

$V_{\boldsymbol{A}}=\{t\boldsymbol{Y}=t(-1,6,10,-12)\mid t\in F\}$. 显然 $\boldsymbol{X}_1\notin V_{\boldsymbol{A}}$, 原方程组非齐次, 通解为

$$\boldsymbol{X}_1+t\boldsymbol{Y}=(1,-2,-3,4)+t(-1,6,10,-12). \qquad \square$$

习题 2.6

1. 求由以下每个小题中的向量生成的子空间的维数, 并求出一组基.

(1) $\boldsymbol{\alpha}_1=(6,4,1,-1,2)$, $\boldsymbol{\alpha}_2=(1,0,2,3,4)$, $\boldsymbol{\alpha}_3=(1,4,-9,-16,22)$, $\boldsymbol{\alpha}_4=(7,1,0,-1,3)$;

(2) $\boldsymbol{\alpha}_1=(1,\ -1,\ 2,\ 4)$, $\boldsymbol{\alpha}_2=(0,\ 3,\ 1,\ 2)$, $\boldsymbol{\alpha}_3=(3,\ 0,\ 7,\ 14)$, $\boldsymbol{\alpha}_4=(1,\ -1,\ 2,\ 0)$, $\boldsymbol{\alpha}_5=(2,\ 1,\ 5,\ 6)$.

2. 下列方程的解集合 W 是否是 $\mathbf{R}^4$ 的子空间:

(1) $x_1+2x_2=3x_3+4x_4$;

(2) $x_1+2x_2=3x_3+4-x_4$;

(3) $(x_1+2x_2)^2=(3x_3+4x_4)^2$;

(4) $(x_1+2x_2)^2+(3x_3+4x_4)^2=0$.

3. 求下列每个齐次线性方程组的一个基础解系, 并用它表出全部解.

(1) $\begin{cases}x_1+x_2+x_3+x_4+x_5=0,\\x_1+2x_2+3x_3+4x_4+5x_5=0,\\x_1-x_3-2x_4-3x_5=0;\end{cases}$

(2) $\begin{cases}x_1+x_2+x_3+x_4-4x_5=0,\\x_1-2x_2+3x_3-4x_4+2x_5=0,\\-x_1+3x_2-5x_3+7x_4-4x_5=0,\\x_1+2x_2-x_3+4x_4-6x_5=0.\end{cases}$

4. 已知 F^5 中的向量

$$\boldsymbol{X}_1=(1,2,3,4,5),\quad \boldsymbol{X}_2=(1,-1,1,-1,1),\quad \boldsymbol{X}_3=(1,2,4,8,16),$$

求一个齐次线性方程组, 使 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 组成这个方程组的基础解系.

5. 已知 5 元线性方程组的系数矩阵的秩为 3, 且以下向量是它的解:

$$\boldsymbol{X}_1=(1,1,1,1,1),\quad \boldsymbol{X}_2=(1,2,3,4,5),\quad \boldsymbol{X}_3=(1,0,-3,-2,-3).$$

(1) 求方程组的通解.

(2) $\boldsymbol{X}_1+\boldsymbol{X}_2+\boldsymbol{X}_3$ 是否是方程组的解?

(3) $\dfrac{1}{3}(\boldsymbol{X}_1+\boldsymbol{X}_2+\boldsymbol{X}_3)$ 是否是方程组的解?

6. 在平面上建立了直角坐标系, A,B 是两圆 $x^2+y^2-x+2y-10=0$ 及 $x^2+y^2+3x-4y-1=0$ 的交点. 求过 A,B 及 $C(2,0)$ 的圆的方程.

附录 2 齐次线性方程组解空间的维数公式

定理 2.6.1 (齐次线性方程组解空间的维数) 齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的解空间 $V_{\boldsymbol{A}}$ 的维数

$$\dim V_{\boldsymbol{A}}=n-\operatorname{rank}\boldsymbol{A}.$$

证明　将系数矩阵 $\boldsymbol{A}$ 经过一系列初等行变换化为最简阶梯形矩阵

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_{11} & \cdots & 0 & \cdots & 0 & \cdots & \lambda_{1n}\\ 0 & \cdots & \lambda_{2j_2} & \cdots & 0 & \cdots & \lambda_{2n}\\ \vdots & & \vdots & & \vdots & & \vdots\\ 0 & \cdots & 0 & \cdots & \lambda_{rj_r} & \cdots & \lambda_{rn}\\ & & & & & & \boldsymbol{O}\end{pmatrix},$$

其中 $r=\operatorname{rank}\boldsymbol{\Lambda}=\operatorname{rank}\boldsymbol{A}$. $\boldsymbol{\Lambda}$ 只有前 r 行不为零, 各行第一个非零元 $\lambda_{11},\lambda_{2j_2},\cdots,\lambda_{rj_r}$ 都等于 1, 分别位于第 $1,j_2,\cdots,j_r$ 列, 这 r 列的其余元素全部为 0. 原方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 化为 $\boldsymbol{\Lambda X}=\boldsymbol{0}$ 即

$$\begin{cases}x_{j_1}+\lambda_{1j_{r+1}}x_{j_{r+1}}+\cdots+\lambda_{1j_n}x_{j_n}=0,\\ x_{j_2}+\lambda_{2j_{r+1}}x_{j_{r+1}}+\cdots+\lambda_{2j_n}x_{j_n}=0,\\ \qquad\cdots\cdots\cdots\cdots\\ x_{j_r}+\lambda_{rj_{r+1}}x_{j_{r+1}}+\cdots+\lambda_{rj_n}x_{j_n}=0.\end{cases}\tag{2.6.4}$$

未知数 $x_{j_1},x_{j_2},\cdots,x_{j_r}$ 仅在其中一个方程中出现, 且系数都是 1. 将这 r 个未知数留在等号左边, 其余所有的项移到右边, 方程组变形为:

$$\begin{cases}x_{j_1}=-\lambda_{1j_{r+1}}x_{j_{r+1}}-\cdots-\lambda_{1j_n}x_{j_n},\\ x_{j_2}=-\lambda_{2j_{r+1}}x_{j_{r+1}}-\cdots-\lambda_{2j_n}x_{j_n},\\ \qquad\cdots\cdots\cdots\cdots\\ x_{j_r}=-\lambda_{rj_{r+1}}x_{j_{r+1}}-\cdots-\lambda_{rj_n}x_{j_n}.\end{cases}\tag{2.6.5}$$

等式右边的 $n-r$ 个未知数 $x_{j_{r+1}},\cdots,x_{j_n}$ 都可以在规定的数域 F 中任意取值 $t_1,\cdots,t_{n-r}$, 代入方程组 (2.6.5) 右边, 将等号左边的 $x_{j_1},x_{j_2},\cdots,x_{j_r}$ 表示成 $t_1,\cdots,t_{n-r}$ 的线性组合, 就得到原方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 的通解

$$\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}=(f_1(t_1,\cdots,t_{n-r}),\cdots,f_n(t_1,\cdots,t_{n-r}))^{\mathrm{T}},$$

其中每个分量 x_i $(1\leqslant i\leqslant n)$ 都是 $t_1,\cdots,t_{n-r}$ 的一个常系数线性组合, 通解

$$\boldsymbol{X}=t_1\boldsymbol{X}_1+\cdots+t_{n-r}\boldsymbol{X}_{n-r}$$

由某 $n-r$ 个 n 维列向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{n-r}$ 在数域 F 上的全体线性组合组成, 其中每个 $\boldsymbol{X}_i$ 也是 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{n-r}$ 的线性组合, 也是一个解. $V_{\boldsymbol{A}}$ 就是由这 $n-r$ 个解生成的子空间 $L(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{n-r})$.

线性组合式 $\boldsymbol{X}=t_1\boldsymbol{X}_1+\cdots+t_{n-r}\boldsymbol{X}_{n-r}$ 的系数 $t_1,\cdots,t_{n-r}$ 本身就是 $\boldsymbol{X}$ 的分量 $x_1,x_{j_2},\cdots,x_{j_r}$ 所取的值, 因此 $\boldsymbol{X}=t_1\boldsymbol{X}_1+\cdots+t_{n-r}\boldsymbol{X}_{n-r}=\boldsymbol{0}$ 当且仅当 $t_1=\cdots=t_{n-r}=0$, 这说明 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{n-r}$ 线性无关, 组成 $V_{\boldsymbol{A}}$ 的一组基. 由此得到

$$\dim V_{\boldsymbol{A}}=n-r=n-\operatorname{rank}\boldsymbol{A}. \qquad \square$$

§2.7* 子空间的交与和

2.7.1 子空间的交

例 1 在空间直角坐标系中, 已知点 $A(1,1,1),B(1,2,3),C(2,4,5),D(1,-2,-2)$. 求平面 OAB 与 OCD 的交线 l 的方程.

解 将向量 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC},\overrightarrow{OD}$ 的坐标分别记为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1,\boldsymbol{b}_2$. 则

$P\in$ 平面 $OAB\Leftrightarrow\overrightarrow{OP}=x_1\overrightarrow{OA}+x_2\overrightarrow{OB}$; $P\in$ 平面 $OCD\Leftrightarrow\overrightarrow{OP}=y_1\overrightarrow{OC}+y_2\overrightarrow{OD}$.

$P\in$ 交线 $l\Leftrightarrow\overrightarrow{OP}=x_1\overrightarrow{OA}+x_2\overrightarrow{OB}=y_1\overrightarrow{OC}+y_2\overrightarrow{OD}\Leftrightarrow P$ 的坐标 $\boldsymbol{X}$ 满足方程

$$\boldsymbol{X}=x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2=y_1\boldsymbol{b}_1+y_2\boldsymbol{b}_2.$$

解如下齐次线性方程组求 (x_1,x_2) :

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2-y_1\boldsymbol{b}_1-y_2\boldsymbol{b}_2=\boldsymbol{0},$$

$$\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,-\boldsymbol{b}_1,-\boldsymbol{b}_2)=\begin{pmatrix}1&1&-2&-1\\1&2&-4&2\\1&3&-5&2\end{pmatrix}\to\begin{pmatrix}1&0&0&-4\\0&1&0&-3\\0&0&1&-3\end{pmatrix}.$$

$$\begin{cases}x_1-4y_2=0,\\x_2-3y_2=0,\\y_1-3y_2=0.\end{cases}\quad\Rightarrow\quad(x_1,x_2)=(4,3)t\ (\forall\ t\in\mathbf{R}).$$

因此, l 上全体点的坐标的集合为 $\{4t\boldsymbol{a}_1+3t\boldsymbol{a}_2\mid t\in\mathbf{R}\}=\{(7,10,13)t\mid t\in\mathbf{R}\}$. l 的参数方程为 $(x,y,z)=(7t,10t,13t)\ (t\in\mathbf{R})$. $\square$

例 2 已知平面 $\pi_1:x+y+z=0$, $\pi_2:2x+y+5z=0$ 相交于一条直线 l. 过 l 及点 $A(2,5,1)$ 作平面 π, 求 π 的方程.

解 易验证平面 π_1 不过点 A. 对任意实数 λ,

$$\lambda(x+y+z)+(2x+y+5z)=0 \tag{2.7.1}$$

是 3 元一次方程, 它的图像是平面. 显然, 方程 $x+y+z=0$ 与 $2x+y+5z=0$ 的公共解是方程 (2.7.1) 的解. 因此, 方程 (2.7.1) 所表示的平面 π 一定经过平面 π_1,π_2 的交线. 只需适当选择 λ 的值使 π 经过点 $A(2,5,1)$. 为此, 将 $x=2,y=5,z=1$ 代入方程 (2.7.1) 得

$$8\lambda+14=0 \quad\Leftrightarrow\quad \lambda=-\frac{7}{4}.$$

将 $\lambda=-\frac{7}{4}$ 代入方程 (2.7.1) 得 $\frac{1}{4}x-\frac{3}{4}y+\frac{13}{4}z=0$, 即

$$x-3y+13z=0.$$

此方程的图像是过 l 及点 A 的平面, 而满足此条件的平面是唯一的. 因此这就是所求的方程. □

例 3 (1) 设 W_1,W_2 分别是数域 F 上的线性方程组

$$\begin{cases}x_1+x_2+\ \ x_3-x_4-x_5=0,\\ \qquad x_2+2x_3+x_4\qquad=0\end{cases} \quad 与 \quad \begin{cases}x_1+2x_2+7x_3+5x_4-4x_5=0,\\ \qquad x_2+4x_3+3x_4-x_5=0\end{cases}$$

的解空间, 求 $W_1\cap W_2$.

(2) 设 π_1 是建立了空间直角坐标系的 3 维几何空间 $\mathbf{R}^3$ 中过点 $(0,0,0),(1,-1,0),(1,2,-3)$ 的平面, π_2 是过点 $(0,0,0),(1,-1,-1),(2,3,1)$ 的平面, 求这两个平面的交集 $\pi_1\cap\pi_2$.

解 (1) 将两个方程组的 4 个方程共同组成一个方程组, 求得的通解为

$$\left(-\frac{1}{2}t_1+3t_2,\ 3t_1-3t_2,\ -\frac{3}{2}t_1+t_2,\ t_1,\ t_2\right),$$

即

$$t_1\left(-\frac{1}{2},\ 3,\ -\frac{3}{2},\ 1,\ 0\right)+t_2\,(3,\ -3,\ 1,\ 0,\ 1)$$

组成的集合就是 $W_1\cap W_2$. 容易看出, $W_1\cap W_2$ 是由两个线性无关向量生成的 2 维子空间.

(2) 用几何向量 $\overrightarrow{OP}$ 表示 3 维几何空间中的点 $P(x,y,z)$, 则 π_1 是向量 $\boldsymbol{\alpha}_1=(1,-1,0)$, $\boldsymbol{\alpha}_2=(1,2,-3)$ 生成的子空间, π_2 是 $\boldsymbol{\beta}_1=(1,-1,-1)$ 与 $\boldsymbol{\beta}_2=(2,3,1)$ 生成的子空间.

$$\boldsymbol{\alpha}\in\pi_1\cap\pi_2 \quad\Leftrightarrow\quad \boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2\ (x_1,x_2,y_1,y_2\in\mathbf{R}), \tag{2.7.2}$$

条件 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2$, 即

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2-y_1\boldsymbol{\beta}_1-y_2\boldsymbol{\beta}_2=0,$$

将 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 的坐标代入得

$$x_1\begin{pmatrix}1\\-1\\0\end{pmatrix}+x_2\begin{pmatrix}1\\2\\-3\end{pmatrix}-y_1\begin{pmatrix}1\\-1\\-1\end{pmatrix}-y_2\begin{pmatrix}2\\3\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}. \qquad (2.7.3)$$

这是以 x_1,x_2,y_1,y_2 为未知数的线性方程组, 求得通解为

$$(x_1,\ x_2,\ y_1,\ y_2)=t\left(\frac{19}{3},\ \frac{5}{3},\ 6,\ 1\right).$$

将 $x_1=\dfrac{19}{3}t,\ x_2=\dfrac{5}{3}t$ 代入 (2.7.2) 得

$$\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2=\frac{19}{3}t(1,-1,0)+\frac{5}{3}t(1,2,-3)=t(8,-3,-5).$$

因此 $\pi_1\cap\pi_2=\{t\,(8,-3,-5)\mid t\in\mathbf{R}\}$ 是 $(8,-3,-5)$ 生成的 1 维子空间, 图像是过原点和点 $(8,-3,-5)$ 的直线. □

例 1, 例 3 求得的子空间的交都是子空间. 这不是偶然的. 一般地, 线性空间 V 中任意多个子空间的交仍是子空间.

定理 2.7.1 F 上线性空间 V 的任意一组子空间 $W_1,\cdots,W_k$ 的交集

$$W=W_1\cap\cdots\cap W_k$$

是 V 的子空间.

证明 交集 W 中任意两个向量 $\boldsymbol{u},\boldsymbol{v}$ 同时含于每个子空间 W_i, 因而它们的和 $\boldsymbol{u}+\boldsymbol{v}$ 含于每个 W_i, 且任意常数倍 $\lambda\boldsymbol{u}$ $(\forall\lambda\in F)$ 含于每个 W_i. 这又导致 $\boldsymbol{u}+\boldsymbol{v}$ 与 $\lambda\boldsymbol{u}$ 含于所有 W_i 的交 W. 这证明了 W 是子空间. □

定理 2.7.1 的证明其实也适用于无穷多个子空间. 任意多个 (包括无穷多个) 子空间的交仍是子空间.

2.7.2 子空间的和

虽然任意多个子空间 W_i 的交集仍是子空间, 但两个子空间 W_1,W_2 的并集 $W_1\cup W_2$ 不一定是子空间. 例如, 平面上过原点的两条不同的直线都是子空间, 但它们的并集对加法不封闭, 不是子空间.

两个子空间的并集 $W_1\cup W_2$ 对数乘封闭, 但对加法不封闭. $W_1\cup W_2$ 生成的子空间为

$$W_1+W_2=\{\boldsymbol{w}_1+\boldsymbol{w}_2\mid \boldsymbol{w}_1\in W_1,\boldsymbol{w}_2\in W_2\},$$

称为 W_1,W_2 的和 (sum).

例 4 给定 F^4 的子空间 W_1 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 和子空间 W_2 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$, 其中

$$\begin{cases}\boldsymbol{\alpha}_1=(1,1,0,0),\\ \boldsymbol{\alpha}_2=(0,1,1,0),\end{cases}\qquad \begin{cases}\boldsymbol{\beta}_1=(1,2,3,4),\\ \boldsymbol{\beta}_2=(0,1,2,2).\end{cases}$$

(1) 求 W_1+W_2 的维数并求出一组基.

(2) 求 $W_1\cap W_2$ 的维数并求出一组基, 并将它扩充为 W_1+W_2 的一组基.

解 (1) W_1+W_2 就是由 W_1,W_2 共同生成的子空间, 也就是由 $S=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$ 生成的子空间. S 的极大线性无关组就是 W_1+W_2 的基. 将 S 中各向量写成列向量的形式, 以它们为各列排成矩阵 $\boldsymbol{A}$, 通过初等行变换化成阶梯形矩阵 $\boldsymbol{T}$, 得

$$\boldsymbol{A}=\begin{pmatrix}1&0&1&0\\1&1&2&1\\0&1&3&2\\0&0&4&2\end{pmatrix}\to\cdots\to\boldsymbol{T}=\begin{pmatrix}1&0&1&0\\0&1&1&1\\0&0&2&1\\0&0&0&0\end{pmatrix}.$$

T 的前 3 列组成列向量组的极大线性无关组. 可见 S 的前 3 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1$ 组成 W_1+W_2 的基, $\dim(W_1+W_2)=3$.

(2) 仿照例 3(2) 求 $\pi_1\cap\pi_2$ 的方法求

$$W_1\cap W_2=\{x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2=y_1\boldsymbol{\beta}_1+y_2\boldsymbol{\beta}_2\mid x_1,x_2,y_1,y_2\in F\}.$$

解线性方程组 $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\beta}_1+x_4\boldsymbol{\beta}_2=0$. 将系数矩阵 $\boldsymbol{A}$ 通过初等行变换化简, 得:

$$\boldsymbol{A}=\begin{pmatrix}1&0&1&0\\1&1&2&1\\0&1&3&2\\0&0&4&2\end{pmatrix}\to\boldsymbol{T}=\begin{pmatrix}1&0&1&0\\0&1&1&1\\0&0&2&1\\0&0&0&0\end{pmatrix}\to\boldsymbol{\Lambda}=\begin{pmatrix}1&0&1&0\\0&1&-1&0\\0&0&2&1\\0&0&0&0\end{pmatrix}.$$

$\boldsymbol{\Lambda}$ 对应的方程组

$$\begin{cases}x_1+x_3=0,\\ x_2-x_3=0,\\ x_4+2x_3=0\end{cases}$$

的通解为 $(x_1,x_2,x_3,x_4)=t(-1,1,1,-2)$. 因此

$$W_1\cap W_2=\{t(-\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)=t(-\boldsymbol{\beta}_1+2\boldsymbol{\beta}_2)\mid t\in F\},$$

$\boldsymbol{\alpha}_0=\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2=\boldsymbol{\beta}_1-2\boldsymbol{\beta}_2=(1,0,-1,0)$ 组成 $W_1\cap W_2$ 的基, $\dim(W_1\cap W_2)=1$.

$\boldsymbol{\alpha}_0$ 既属于 W_1 又属于 W_2. $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1$ 线性无关, 组成 W_1 的基; $\boldsymbol{\alpha}_0,\boldsymbol{\beta}_1$ 线性无关, 组成 W_2 的基. W_1+W_2 中每个向量 $\boldsymbol{w}$ 可写成 $\boldsymbol{w}=\boldsymbol{w}_1+\boldsymbol{w}_2$ 的形式, 其中 $\boldsymbol{w}_1\in W_1$ 是 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1$ 的线性组合, $\boldsymbol{w}_2\in W_2$ 是 $\boldsymbol{\alpha}_0,\boldsymbol{\beta}_1$ 的线性组合. 因此 $\boldsymbol{w}$ 是 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1$ 的线性组合.

我们证明 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1$ 线性无关. 设 $\lambda_0\boldsymbol{\alpha}_0+\lambda_1\boldsymbol{\alpha}_1+\lambda_2\boldsymbol{\beta}_1=\mathbf{0}$, 则

$$\lambda_0\boldsymbol{\alpha}_0+\lambda_1\boldsymbol{\alpha}_1=-\lambda_2\boldsymbol{\beta}_1\in W_1\cap W_2,$$

因而 $\lambda_0\boldsymbol{\alpha}_0+\lambda_1\boldsymbol{\alpha}_1$ 即 $-\lambda_2\boldsymbol{\beta}_1$ 是 $\boldsymbol{\alpha}_0$ 的某个线性组合 $\lambda_3\boldsymbol{\alpha}_0$.

我们有 $-\lambda_2\boldsymbol{\beta}_1=\lambda_3\boldsymbol{\alpha}_0$ 即 $\lambda_3\boldsymbol{\alpha}_0+\lambda_2\boldsymbol{\beta}_1=\mathbf{0}$. 由 $\boldsymbol{\alpha}_0,\boldsymbol{\beta}_1$ 线性无关知 $\lambda_3=\lambda_2=0$. 于是又有 $\lambda_0\boldsymbol{\alpha}_0+\lambda_1\boldsymbol{\alpha}_1=-\lambda_2\boldsymbol{\beta}_1=\mathbf{0}$, 再由 $\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1$ 线性无关知 $\lambda_0=\lambda_1=0$. 这证明了 $\{\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1\}$ 线性无关, 是由 $W_1\cap W_2$ 的基 $\{\boldsymbol{\alpha}_0\}$ 扩充成的 W_1+W_2 的基. □

例 4(2) 对 $\{\boldsymbol{\alpha}_0,\boldsymbol{\alpha}_1,\boldsymbol{\beta}_1\}$ 是 W_1+W_2 的基的证明可以推广到一般情形:

定理 2.7.2 设 W_1,W_2 是 F^n 的两个子空间. $S_0=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\}$ 是 $W_1\cap W_2$ 的一组基, S_0 添加 $T_1=\{\boldsymbol{\alpha}_{s+1},\cdots,\boldsymbol{\alpha}_r\}$ 扩充为 W_1 的一组基, S_0 添加 $T_2=\{\boldsymbol{\beta}_{s+1},\cdots,\boldsymbol{\beta}_t\}$ 扩充为 W_2 的一组基. 则

(1) $S_0\cup T_1\cup T_2$ 是 W_1+W_2 的一组基.

(2) $\dim(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)$.

证明 (1) 每个 $\boldsymbol{w}\in W_1+W_2$ 可以写成 $\boldsymbol{w}=\boldsymbol{w}_1+\boldsymbol{w}_2$ 的形式, 其中 $\boldsymbol{w}_1\in W_1$, $\boldsymbol{w}_2\in W_2$. $\boldsymbol{w}_1$ 是 $S_0\cup T_1$ 的线性组合, $\boldsymbol{w}_2$ 是 $S_0\cup T_2$ 的线性组合, 因此 $\boldsymbol{w}$ 是 $S_0\cup T_1\cup T_2$ 的线性组合.

设

$$\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_s\boldsymbol{\alpha}_s+\lambda_{s+1}\boldsymbol{\alpha}_{s+1}+\cdots+\lambda_r\boldsymbol{\alpha}_r+\mu_{s+1}\boldsymbol{\beta}_{s+1}+\cdots+\mu_t\boldsymbol{\beta}_t=\mathbf{0} \quad (2.7.4)$$

对 $\lambda_1,\cdots,\lambda_r,\mu_{s+1},\cdots,\mu_t\in F$ 成立. 则

$$\begin{aligned}\boldsymbol{w}_0 &= \lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_s\boldsymbol{\alpha}_s+\lambda_{s+1}\boldsymbol{\alpha}_{s+1}+\cdots+\lambda_r\boldsymbol{\alpha}_r\\ &= -\mu_{s+1}\boldsymbol{\beta}_{s+1}-\cdots-\mu_t\boldsymbol{\beta}_t\in W_1\cap W_2,\end{aligned}$$

$\boldsymbol{w}_0$ 是 S_0 的线性组合, 可以写成

$$\boldsymbol{w}_0=-\mu_{s+1}\boldsymbol{\beta}_{s+1}-\cdots-\mu_t\boldsymbol{\beta}_t=\mu_1\boldsymbol{\alpha}_1+\cdots+\mu_s\boldsymbol{\alpha}_s$$

的形式, 其中 $\mu_1,\cdots,\mu_t\in F$. 移项得

$$\mu_1\boldsymbol{\alpha}_1+\cdots+\mu_s\boldsymbol{\alpha}_s+\mu_{s+1}\boldsymbol{\beta}_{s+1}+\cdots+\mu_t\boldsymbol{\beta}_t=\mathbf{0},$$

$\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_{s+1},\cdots,\boldsymbol{\beta}_t\}=S_0\cup T_2$ 是 W_2 的基, 线性无关, 因此 $\mu_1=\cdots=\mu_s=\mu_{s+1}=\cdots=\mu_t=0$. 代入 (2.7.4) 得

$$\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_r\boldsymbol{\alpha}_r=\mathbf{0}.$$

再由 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r\}=S_0\cup T_1$ 是 W_1 的基知 $\lambda_1=\cdots=\lambda_r=0$.

这证明了 $S_0\cup T_1\cup T_2$ 是 W_1+W_2 的基.

(2) 显然 $S_0\cup T_1\cup T_2$ 这组基包含的向量个数为 $r+(t-s)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)$, 这就是 $\dim(W_1+W_2)$. □

推论 2.7.1 如果 F^n 的子空间 W_1,W_2 的维数之和 $\dim W_1+\dim W_2>n$, 则 $W_1\cap W_2\neq\mathbf{0}$.

证明 $W_1+W_2\subseteq F^n$, 因而 $\dim(W_1+W_2)\leqslant n$, $\dim(W_1\cap W_2)=\dim W_1+\dim W_2-\dim(W_1+W_2)\geqslant\dim W_1+\dim W_2-n>0$. □

定理 2.7.3 设 W_1,W_2 是 F^n 的两个子空间. 则如下命题等价:

(1) $W_1\cap W_2=\{\mathbf{0}\}$.

(2) $\dim(W_1+W_2)=\dim W_1+\dim W_2$.

(3) 每个 $\boldsymbol{w}\in W_1+W_2$ 的分解式 $\boldsymbol{w}=\boldsymbol{w}_1+\boldsymbol{w}_2$ $(\boldsymbol{w}_1\in W_1,\boldsymbol{w}_2\in W_2)$ 由 $\boldsymbol{w}$ 唯一决定.

(4) 设 $\boldsymbol{w}_1\in W_1,\boldsymbol{w}_2\in W_2$, 则 $\boldsymbol{w}_1+\boldsymbol{w}_2=\mathbf{0}\Leftrightarrow\boldsymbol{w}_1=\boldsymbol{w}_2=\mathbf{0}$.

定义 2.7.1 如果 F^n 的子空间 W_1,W_2 满足定理 2.7.3 的四个等价命题的任何一个命题, 则 W_1+W_2 称为 W_1,W_2 的**直和** (direct sum), 记作 $W_1\oplus W_2$. □

习题 2.7

1. 给定 F^4 的子空间 W_1 的基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\}$ 和子空间 W_2 的基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\}$, 其中

$$\begin{cases}\boldsymbol{\alpha}_1=(1,1,0,0),\\ \boldsymbol{\alpha}_2=(0,1,1,0),\end{cases}\qquad\begin{cases}\boldsymbol{\beta}_1=(1,2,3,4),\\ \boldsymbol{\beta}_2=(0,1,2,2).\end{cases}$$

(1) 求 W_1+W_2 的维数并求出一组基.

(2) 求 $W_1\cap W_2$ 的维数并求出一组基.

2. 设 W_1,W_2 分别是数域 F 上的线性方程组

$$\begin{cases}x_1+x_2+x_3\qquad\qquad=0,\\ \qquad x_2+2x_3+x_4=0\end{cases}\quad\text{与}\quad\begin{cases}x_1+2x_2+4x_3+2x_4=0,\\ \qquad x_2+4x_3+3x_4=0\end{cases}$$

的解空间. 分别求 W_1+W_2 及 $W_1\cap W_2$ 的维数并各求出一组基.

3. 设 W_1,W_2 分别是数域 F 上齐次线性方程组 $x_1+x_2+\cdots+x_n=0$ 与 $x_1=x_2=\cdots=x_n$ 的解空间. 求证: $F^n=W_1\oplus W_2$.

§2.8* 更多的例子

2.8.1 线性相关应用例

例 1 求方程 $\sqrt{x^2+x+1}+\sqrt{2x^2+x+5}=\sqrt{x^2-3x+13}$ 的实数解.

解 令 $u=\sqrt{x^2+x+1}, v=\sqrt{2x^2+x+5}, w=\sqrt{x^2-3x+13}$. 则原方程成为

$$u+v=w. \tag{2.8.1}$$

设法求常数 λ_1,λ_2 使 $\lambda_1(x^2+x+1)+\lambda_2(2x^2+x+5)=x^2-3x+13$,

即 $$\lambda_1(1,1,1)+\lambda_2(2,1,5)=(1,-3,13).$$

解方程组 $\begin{cases}\lambda_1+2\lambda_2=1,\\ \lambda_1+\lambda_2=-3,\\ \lambda_1+5\lambda_2=13,\end{cases}$ 得 $\begin{cases}\lambda_1=-7,\\ \lambda_2=4.\end{cases}$ 因而

$$-7(x^2+x+1)+4(2x^2+x+5)=x^2-3x+13,$$

即 $$-7u^2+4v^2=w^2. \tag{2.8.2}$$

将 (2.8.1) 代入 (2.8.2) 得 $-7u^2+4v^2=(u+v)^2$,

整理得 $$3v^2-2uv-8u^2=0.$$

左边因式分解得

$$(v-2u)(3v+4u)=0, \tag{2.8.3}$$

易见当 x 为实数时, $x^2+x+1, 2x^2+x+5$ 都是正实数, 而 u,v 分别是它们的算术平方根, 恒为正, 因此 $3v+4u>0$, 只能 $v-2u=0$, 即 $v=2u$, 即

$$\sqrt{2x^2+x+5}=2\sqrt{x^2+x+1},$$

两边平方, 整理得

$$2x^2+3x-1=0 \quad\Leftrightarrow\quad x=\frac{-3\pm\sqrt{3^2+8}}{4}=\frac{-3\pm\sqrt{17}}{4}.$$

经检验, $x=\dfrac{-3\pm\sqrt{17}}{4}$ 确实是原方程的解, 因此就是原方程的全部实数解. □

例 1 解答的关键是多项式 $x^2+x+1, 2x^2+x+5, x^2-3x+13$ 线性相关, 利用它们之间的线性关系式可求得方程的解.

2.8.2 基与坐标应用例

例 2(拉格朗日 (Lagrange) 插值公式) 平面上建立了直角坐标系，(x_1,y_1), $(x_2,y_2),\cdots,(x_n,y_n)$ 是任意 n 个横坐标不同的已知点. 求不超过 $n-1$ 次的实系数多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$, 使它的图像曲线经过这 n 个已知点.

解 所求多项式 $f(x)$ 满足的条件为: $f(x_i)=y_i\ (1\leqslant i\leqslant n)$. 记

$$\sigma(f)=\begin{pmatrix} f(x_1)\\ \vdots \\ f(x_n)\end{pmatrix},\quad \boldsymbol{Y}=\begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix},$$

则 $f(x)$ 满足的条件成为:

$$\sigma(f)=\boldsymbol{Y}. \tag{2.8.4}$$

对任意两个多项式 f,g 及常数 λ, 易验证 $\sigma(f+g)=\sigma(f)+\sigma(g),\sigma(\lambda f)=\lambda\sigma(f)$.

将 $\boldsymbol{Y}$ 写成自然基 $E=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 的线性组合 $\boldsymbol{Y}=y_1\boldsymbol{e}_1+\cdots+y_n\boldsymbol{e}_n$. 只要对每个自然基向量 $\boldsymbol{e}_i\ (1\leqslant i\leqslant n)$ 找到满足条件 $\sigma(f_i)=\boldsymbol{e}_i$ 的多项式 f_i, 则所有这些 f_i 的线性组合 $f=y_1f_1+\cdots+y_nf_n$ 满足条件 (2.8.4):

$$\sigma(f)=y_1\sigma(f_1)+\cdots+y_n\sigma(f_n)=y_1\boldsymbol{e}_1+\cdots+y_n\boldsymbol{e}_n=\boldsymbol{Y}.$$

先求 f_1 满足条件 $\sigma(f_1)=\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}}$, 即 $f_1(x_1)=1$ 且 $f_1(x_2)=\cdots=f_1(x_n)=0$. 满足条件 $f_1(x_2)=\cdots=f_1(x_n)=0$ 且次数不超过 $n-1$ 的多项式具有形式 $f_1(x)=\lambda_1(x-x_2)\cdots(x-x_n)$, 其中 λ_1 是待定常数 λ_1. 又因为

$$f_1(x_1)=\lambda_1(x_1-x_2)\cdots(x_1-x_n)=1 \quad\Leftrightarrow\quad \lambda_1=\frac{1}{(x_1-x_2)\cdots(x_1-x_n)}.$$

于是

$$f_1(x)=\frac{(x-x_2)\cdots(x-x_n)}{(x_1-x_2)\cdots(x_1-x_n)}.$$

类似地, 对每个 $1\leqslant i\leqslant n$, 可求得满足条件 $\sigma(f_i)=\boldsymbol{e}_i$ 的

$$f_i(x)=\prod_{1\leqslant j\leqslant n,\, j\neq i}\frac{x-x_j}{x_i-x_j},$$

于是得

$$f(x)=y_1f_1(x)+\cdots+y_nf_n(x)=\sum_{i=1}^{n}\left(y_i\prod_{1\leqslant j\leqslant n,\, j\neq i}\frac{x-x_j}{x_i-x_j}\right)$$

满足条件 $\sigma(f)=\boldsymbol{Y}$.

如果多项式 $\tilde{f}(x)$ 也满足 $\sigma(\tilde{f})=\boldsymbol{Y}$. 则 $\sigma(\tilde{f}-f)=\sigma(\tilde{f})-\sigma(f)=\boldsymbol{Y}-\boldsymbol{Y}=\mathbf{0}$. 这说明 $x_1,\cdots,x_n$ 都是方程 $\tilde{f}(x)-f(x)=0$ 的根, $\tilde{f}(x)-f(x)$ 被 $(x-x_1)\cdots(x-x_n)$ 整除, $f(x)$ 是 $\tilde{f}(x)$ 被 $(x-x_1)\cdots(x-x_n)$ 除得到的余式, 即

$$\tilde{f}(x)=f(x)+q(x)(x-x_1)\cdots(x-x_n),$$

$q(x)$ 是某个多项式. 特别地, 当 $\tilde{f}(x)$ 的次数不超过 $n-1$ 时, 只能 $\tilde{f}(x)=f(x)$.□

例 2 求出的 f 的表达式称为拉格朗日插值公式.

例如, 在本书第 1 章 §1.1 例 2 中求抛物线 $y=ax^2+bx+c$ 经过三个已知点 $(1,1),(2,2),(3,0)$. 直接利用拉格朗日插值公式得

$$f(x)=\frac{(x-2)(x-3)}{(1-2)(1-3)}+2\frac{(x-1)(x-3)}{(2-1)(2-3)}=-\frac{3}{2}x^2+\frac{11}{2}x-3,$$

所求抛物线方程为 $y=-\frac{3}{2}x^2+\frac{11}{2}x-3$.

例 2 中将每个不超过 $n-1$ 次的多项式 f 对应于唯一的 n 维列向量 $\sigma(f)$, 满足条件

$$\sigma(f+g)=\sigma(f)+\sigma(g),\sigma(\lambda f)=\lambda\sigma(f).$$

$\sigma(f)$ 其实是 f 在某组基 $S=\{f_1,\cdots,f_n\}$ 下的坐标. 例 2 实际上是反过来由坐标 $\boldsymbol{Y}$ 求向量 f. 基向量 $f_1,\cdots,f_n$ 的坐标分别等于列向量空间的自然基向量 $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$. 由 $\sigma(f_i)=\boldsymbol{e}_i$ 求出各个基向量 $f_1,\cdots,f_n$, 则坐标为 $\boldsymbol{Y}=(y_1,\cdots,y_n)^{\mathrm{T}}$ 的向量 $f=y_1f_1+\cdots+y_nf_n$.

例 3(中国剩余定理) 对任意非负整数 $r_1<3,r_2<5,r_3<7$, 求最小的非负整数使它除以 3,5,7 的余数分别是 r_1,r_2,r_3.

分析 在例 2 中, 将每个多项式 f 对应于一个数组 $\sigma(f)=(f(x_1),\cdots,f(x_n))$, 这个数组 $\sigma(f)$ 可以认为是 f 的坐标. 反过来, 要对任意给定的坐标 $\boldsymbol{Y}=(y_1,\cdots,y_n)$ 求 f, 我们先将坐标 $\boldsymbol{Y}$ 分解为自然基的线性组合

$$\boldsymbol{Y}=y_1\boldsymbol{e}_1+\cdots+y_n\boldsymbol{e}_n,$$

求出每个自然基向量 $\boldsymbol{e}_i$ 对应的多项式 f_i, 则 $\boldsymbol{Y}$ 对应的 $f=y_1f_1+\cdots+y_nf_n$.

本题也可以用类似的思路来解决: 将每个整数 x 对应于 3 维坐标 $\sigma(x)=(r_1,r_2,r_3)$, 其中 r_1,r_2,r_3 分别是 x 除以 3,5,7 得到的余数. 为了对任意给定的坐标 $\boldsymbol{Y}=(r_1,r_2,r_3)$ 反过来求 x 使 $\sigma(x)=\boldsymbol{Y}$, 先将 $\boldsymbol{Y}$ 分解为“自然基向量” $(1,0,0),(0,1,0),(0,0,1)$ 的线性组合

$$\boldsymbol{Y}=r_1(1,0,0)+r_2(0,1,0)+r_3(0,0,1),$$

分别对三个“自然基向量”求整数 x_1,x_2,x_3 使 $\sigma(x_1)=(1,0,0),\sigma(x_2)=(0,1,0),\sigma(x_3)=(0,0,1)$. 则 $x=r_1x_1+r_2x_2+r_3x_3$ 满足所需条件 $\sigma(x)=r_1\sigma(x_1)+r_2\sigma(x_2)+r_3\sigma(x_3)=(r_1,r_2,r_3)$.

解 对每个整数 x, 设 x 除以 3,5,7 的余数分别是 r_1,r_2,r_3, 记 $\sigma(x)=(r_1,r_2,r_3)$. 先设法求出整数 x_1,x_2,x_3 使

$$\sigma(x_1)=(1,0,0),\quad \sigma(x_2)=(0,1,0),\quad \sigma(x_3)=(0,0,1). \tag{2.8.5}$$

由 $\sigma(x_1)=(1,0,0)$ 知 x_1 除以 5,7 的余数都是 0, x_1 是 5,7 的公倍数, $x_1=35k$, k 为整数. 依次试验 $k=1,2,\cdots$, 使 $35k$ 除以 3 余 1. 发现 $k=2$ 时的 $x_1=70$ 符合要求.

类似地, $\sigma(x_2)=(0,1,0)$ 意味着 x_2 是 3,7 的公倍数, $x_2=21k$, $k=1$ 时的 $x_2=21$ 除以 5 余 1. $\sigma(x_3)=(0,0,1)$ 意味着 x_3 是 3,5 的公倍数, $x_3=15k$, $k=1$ 时的 $x_3=15$ 除以 7 余 1.

于是, $\tilde{x}=70r_1+21r_2+15r_3$ 被 3,5,7 除的余数分别等于 r_1,r_2,r_3. 将 $\tilde{x}$ 加上或减去 3,5,7 的任何一个公倍数 $3\times5\times7q=105q$, 被 3,5,7 除的余数仍符合要求. 因而将 $\tilde{x}$ 除以 105 得到的余数 x 就是满足条件的最小非负整数. □

例如, 求最小的正整数 x 使它被 3,5,7 除的余数分别是 2,3,4. 取 $\tilde{x}=2\times70+3\times21+4\times15=263$ 除以 105 的余数 53 即符合要求.

例 4 的关键是求出了三个整数 $x_1=70,x_2=21,x_3=15$ 分别满足条件 $\sigma(x_1)=(1,0,0),\sigma(x_2)=(0,1,0),\sigma(x_3)=(0,0,1)$. 中国古代还有一首诗帮助记忆这个问题的算法:

三人同行七十稀, 五树梅花廿一枝, 七子团圆正半月, 除百零五便得知.

前三句是帮助记忆 70(七十稀),21(廿一枝), 15(正半月, 就是 15 天). 将 r_1,r_2,r_3 分别乘 70,21,15 再相加. 最后一句说: 如果求出的数大于 105, 就减去 105, 还大于 105 就再减 105, 直到小于 105, 就是所求的最小非负整数.

2.8.3 子空间应用例

例 4(斐波那契 (Fibonacci) 数列的通项公式) 满足条件

$$F_1=F_2=1,\quad F_n=F_{n-1}+F_{n-2}\quad (\text{对所有的正整数 } n\geqslant 3)$$

的数列 $F_1,F_2,\cdots,F_n,\cdots$ 称为斐波那契数列. 求这个数列的通项公式.

解 对任意正整数 $m\geqslant 3$, 由复数组成的全体 m 项数列 $\boldsymbol{\alpha}=(a_1,\cdots,a_m)$ 组成的集合 $\mathbf{C}^m$ 是复数域 $\mathbf{C}$ 上的 m 维向量空间.

$\mathbf{C}^m$ 中满足条件 $a_n=a_{n-1}+a_{n-2}$ $(\forall\ n\geqslant 3)$ 的全体数列 $\boldsymbol{\alpha}=(a_1,\cdots,a_m)$ 组成的集合 W 对加法与数乘封闭, 是 $\mathbf{C}^m$ 的子空间. 每个 $\boldsymbol{\alpha}=(a_1,\cdots,a_m)\in$

W 的前两项 a_1,a_2 可以取任意复数值, 以后的各项 a_n $(n\geqslant 3)$ 由前两项 a_1,a_2 完全决定, (a_1,a_2) 可以看成 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_m)\in W$ 的坐标, W 是二维子空间. 设法选取 W 中两个线性无关等比数列 $\boldsymbol{\alpha}_i=(1,q_i,q_i^2,\cdots,q_i^{m-1})$ $(i=1,2)$ 组成 W 的一组基, 则斐波那契数列 $\boldsymbol{\varphi}=(F_1,F_2,\cdots,F_m)\in W$ 可以分解成这组基的线性组合 $\boldsymbol{\varphi}=x\boldsymbol{\alpha}_1+y\boldsymbol{\alpha}_2$, 由等比数列 $\boldsymbol{\alpha}_i$ 的通项公式即可得出 $\boldsymbol{\varphi}=(F_1,F_2,\cdots,F_m)$ 的通项公式.

任一等比数列 $\boldsymbol{\alpha}=(a_1,\cdots,a_m)$ 有通项公式 $a_n=a_1q^{n-1}$ $(\forall\ 1\leqslant n\leqslant m)$, 这个等比数列 $\boldsymbol{\alpha}\in W$ 的充分必要条件为

$$a_n=a_{n-1}+a_{n-2}\quad 即\quad a_1q^{n-1}=a_1q^{n-2}+a_1q^{n-3}\ (\forall\ n\geqslant 3),$$

也就是

$$q^2=q+1\ \Leftrightarrow\ q^2-q-1=0\ \Leftrightarrow\ q=q_1=\frac{1-\sqrt{5}}{2}\ 或\ q=q_2=\frac{1+\sqrt{5}}{2}.$$

由于 $q_1\neq q_2$, 等比数列 $\boldsymbol{\alpha}_1=(1,q_1,\cdots,q_1^{m-1})$ 与 $\boldsymbol{\alpha}_2=(1,q_2,\cdots,q_2^{m-1})$ 线性无关, 组成 W 的一组基. 斐波那契数列 $\boldsymbol{\varphi}=(F_1,F_2,\cdots)$ 可以分解为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的线性组合 $\boldsymbol{\varphi}=x\boldsymbol{\alpha}_1+y\boldsymbol{\alpha}_2$, 即

$$(F_1,F_2,\cdots)=x(1,q_1,\cdots)+y(1,q_2,\cdots)=(x+y,xq_1+yq_2,\cdots),$$

其中 x,y 满足条件

$$\begin{cases}x+y=F_1=1,\\ q_1x+q_2y=F_2=1,\end{cases}\qquad 解之得\qquad (x,y)=\left(\frac{q_2-1}{q_2-q_1},\frac{1-q_1}{q_2-q_1}\right),$$

从而

$$F_n=xq_1^{n-1}+yq_2^{n-1}=\frac{q_1^{n-1}(q_2-1)+q_2^{n-1}(1-q_1)}{q_2-q_1},$$

由于 $q_1+q_2=1,q_2-1=-q_1,1-q_1=q_2$, 又 $q_2-q_1=\sqrt{5}$. 因此

$$F_n=\frac{q_2^n-q_1^n}{q_2-q_1}=\frac{\left(\dfrac{1+\sqrt{5}}{2}\right)^n-\left(\dfrac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}=\frac{(1+\sqrt{5})^n-(1-\sqrt{5})^n}{2^n\sqrt{5}}.$$

□

例 5(幻方) 试将正整数 $1,2,\cdots,9$ 填入 3×3 的方格中, 使每行、每列、每条对角线上的 3 个数之和都取同一个值.

分析 看起来只需用到小学算术, 但要将 $1,2,\cdots,9$ 填入表中同时满足这些条件并不容易.

先将 $0,1,\cdots,8$ 这 9 个非负整数填入 3×3 方格中, 得到一个 3 阶方阵 $\boldsymbol{B}=(b_{ij})_{3\times 3}$, 希望它的每行、每列、每条对角线上的 3 个数之和相等. $\boldsymbol{B}$ 中每个整数 b_{ij} 可以分解为 $b_{ij}=3q_{ij}+r_{ij}$, 其中 q_{ij},r_{ij} 分别是 a_{ij} 被 3 除得到的商和余数, 取值都在 $\{0,1,2\}$ 中. $\boldsymbol{B}$ 被相应地分解为 $\boldsymbol{A}=3\boldsymbol{Q}+\boldsymbol{R}$, 其中

$\boldsymbol{Q}=(q_{ij})_{3\times 3},\boldsymbol{R}=(r_{ij})_{3\times 3}$ 分别是由商与余数组成的 3 阶方阵. 如果能够用 $0,1,2$ 构造出 $\boldsymbol{Q},\boldsymbol{R}$ 使它们的每行、每列、每条对角线的 3 个数之和相等, 则在 $\boldsymbol{Q},\boldsymbol{R}$ 的线性组合 $\boldsymbol{B}=3\boldsymbol{Q}+\boldsymbol{R}$ 中这些和也都相等. 适当设计 $\boldsymbol{Q},\boldsymbol{R}$ 使 $3\boldsymbol{Q}+\boldsymbol{R}$ 中的 9 个数各不重复, 取遍 $0,1,\cdots,8$. 将 $\boldsymbol{B}$ 的各数同加 1, 就得到 $1,2,\cdots,9$ 组成的幻方 $\boldsymbol{A}$.

解 在 3×3 方格表的每行以适当顺序填入 $0,1,2$ 这三个数, 使每行、每列、每条对角线上三个数之和相等, 得到 $\boldsymbol{Q}$. 将 $\boldsymbol{Q}$ 关于第 2 列作轴对称得到 $\boldsymbol{R}$. 取全由 1 组成的方格表 $\boldsymbol{H}$, 则 $\boldsymbol{A}=3\boldsymbol{Q}+\boldsymbol{R}+\boldsymbol{H}$ 是 $1,2,\cdots,9$ 组成的幻方.

$$\boldsymbol{Q}=\begin{array}{|c|c|c|}\hline 1&2&0\\\hline 0&1&2\\\hline 2&0&1\\\hline\end{array},\quad \boldsymbol{R}=\begin{array}{|c|c|c|}\hline 0&2&1\\\hline 2&1&0\\\hline 1&0&2\\\hline\end{array},\quad \boldsymbol{A}=3\boldsymbol{Q}+\boldsymbol{R}+\boldsymbol{H}=\begin{array}{|c|c|c|}\hline 4&9&2\\\hline 3&5&7\\\hline 8&1&6\\\hline\end{array}.\ \square$$

例 5 的关键是: 如果两个方格表 (n 阶方阵) $\boldsymbol{B}_1,\boldsymbol{B}_2$ 的每行、每列、每条对角线上的 n 个数之和相等, 则在 $\boldsymbol{B}_1$ 与 $\boldsymbol{B}_2$ 之和 $\boldsymbol{B}_1+\boldsymbol{B}_2$ 以及它们的常数倍 $\lambda\boldsymbol{B}_1,\lambda\boldsymbol{B}_2$ 中这些和仍然相等. 换句话说, 这些和相等的方阵所组成的集合是子空间. 既然是子空间, 就可以分解, 将复杂繁多的向量分解为很少很简单的向量的线性组合, 正如将几何平面和空间中无穷多个向量分解为两个或三个向量 (基) 的线性组合, 又如在本节例 4 中将不会求通项公式的斐波那契数列分解为会求通项公式的等比数列的线性组合, 以及本例中将 $1,2,\cdots,9$ 组成的复杂幻方分解为 $0,1,2$ 组成的简单幻方的线性组合.

例 5 的思想方法可以加以推广, 用来对更大的正整数 n 设计 n 阶幻方 (其中 $\boldsymbol{H}$ 都表示全由 1 组成的方格表).

5 阶幻方

$$5\times\begin{array}{|c|c|c|c|c|}\hline 0&1&2&3&4\\\hline 3&4&0&1&2\\\hline 1&2&3&4&0\\\hline 4&0&1&2&3\\\hline 2&3&4&0&1\\\hline\end{array}+\begin{array}{|c|c|c|c|c|}\hline 0&1&2&3&4\\\hline 2&3&4&0&1\\\hline 4&0&1&2&3\\\hline 1&2&3&4&0\\\hline 3&4&0&1&2\\\hline\end{array}+\boldsymbol{H}=\begin{array}{|c|c|c|c|c|}\hline 1&7&13&19&25\\\hline 18&24&5&6&12\\\hline 10&11&17&23&4\\\hline 22&3&9&15&16\\\hline 14&20&21&2&8\\\hline\end{array}.$$

4 阶幻方

$$4\times\begin{array}{|c|c|c|c|}\hline 0&1&2&3\\\hline 3&2&1&0\\\hline 3&2&1&0\\\hline 0&1&2&3\\\hline\end{array}+\begin{array}{|c|c|c|c|}\hline 0&3&3&0\\\hline 1&2&2&1\\\hline 2&1&1&2\\\hline 3&0&0&3\\\hline\end{array}+\boldsymbol{H}=\begin{array}{|c|c|c|c|}\hline 1&8&12&13\\\hline 14&11&7&2\\\hline 15&10&6&3\\\hline 4&5&9&16\\\hline\end{array}.$$

由 4 阶幻方 A_4 构造 8 阶幻方 A_8

$$
\boldsymbol{A}_4=\begin{array}{|c|c|c|c|}\hline 1&8&12&13\\\hline 14&11&7&2\\\hline 15&10&6&3\\\hline 4&5&9&16\\\hline\end{array}\ \rightarrow\ \boldsymbol{X}_8=\begin{array}{|cc|cc|cc|cc|}\hline 1&1&8&8&12&12&13&13\\ 1&1&8&8&12&12&13&13\\\hline 14&14&11&11&7&7&2&2\\ 14&14&11&11&7&7&2&2\\\hline 15&15&10&10&6&6&3&3\\ 15&15&10&10&6&6&3&3\\\hline 4&4&5&5&9&9&16&16\\ 4&4&5&5&9&9&16&16\\\hline\end{array}.
$$

$$
\boldsymbol{Y}_4=\begin{array}{|cc|cc|}\hline 0&1&2&3\\ 3&2&1&0\\\hline 3&2&1&0\\ 0&1&2&3\\\hline\end{array}\ \rightarrow\ \boldsymbol{Y}_8=\begin{array}{|c|c|}\hline \boldsymbol{Y}_4&\boldsymbol{Y}_4\\\hline \boldsymbol{Y}_4&\boldsymbol{Y}_4\\\hline\end{array}=\begin{array}{|cc|cc|cc|cc|}\hline 0&1&2&3&0&1&2&3\\ 3&2&1&0&3&2&1&0\\\hline 3&2&1&0&3&2&1&0\\ 0&1&2&3&0&1&2&3\\\hline 0&1&2&3&0&1&2&3\\ 3&2&1&0&3&2&1&0\\\hline 3&2&1&0&3&2&1&0\\ 0&1&2&3&0&1&2&3\\\hline\end{array}.
$$

$$\boldsymbol{A}_8=\boldsymbol{X}_8+16\boldsymbol{Y}_8.$$

由 3 阶幻方 A_3 构造 6 阶幻方 A_6

$$
\boldsymbol{A}_3=\begin{array}{|c|c|c|}\hline 4&9&2\\\hline 3&5&7\\\hline 8&1&6\\\hline\end{array}\ \rightarrow\ \boldsymbol{X}_6=\begin{array}{|cc|cc|cc|}\hline 4&4&9&9&2&2\\ 4&4&9&9&2&2\\\hline 3&3&5&5&7&7\\ 3&3&5&5&7&7\\\hline 8&8&1&1&6&6\\ 8&8&1&1&6&6\\\hline\end{array}.
$$

设计一个由 2 阶块组成的 6 阶方阵, 使它的各行、各列、两条对角线上 6 个数之和都等于同一个值 9, 且每个 2 阶块由 0, 1, 2, 3 组成:

$$
\boldsymbol{Y}_6=\begin{array}{|cc|cc|cc|}\hline 0&1&2&3&3&0\\ 3&2&1&0&2&1\\\hline 0&3&0&1&2&3\\ 1&2&2&3&1&0\\\hline 2&1&3&0&1&2\\ 3&0&1&2&0&3\\\hline\end{array}.
$$

$$A_6 = X_6 + 9Y_6.$$

一般地, 仿照 3 阶幻方的构造法可以得到任意奇数 n 阶幻方. 仿照 6 阶和 8 阶幻方的构造法可以由任意奇数或偶数阶幻方 A_n 构造出 $2n$ 阶幻方. 这样就可以对每个正整数 $n \geqslant 3$ 构造出一批幻方.

将例 5 得到的 3 阶幻方旋转 90° 的整数倍以及作轴对称可以得到 8 个不同的 3 阶幻方. 可以证明这就是全部的 3 阶幻方. 当 $n > 3$ 时的幻方个数都非常大, 不但难以全部设计出来, 即使设计出来了也难以全部写出来.

习题 2.8

1. 求多项式 $f(x)$ 使 $f(1) = 1$, $f(2) = 2$, $f(3) = 4$.

2. 设 $f_1 = (x-1)(x-2)(x-3), f_2 = x(x-2)(x-3), f_3 = x(x-1)(x-3), f_4 = x(x-1)(x-2)$. 试将 $1, x, x^2, x^3$ 分别表示为 f_1, f_2, f_3, f_4 的线性组合.

3. 求最小的正整数 x 使它被 4,7,9 除的余数分别是 1,2,3.

4. 数列 $\{a_n\}$ 满足条件 $a_1 = a_2 = 1$, 且 $a_n = 3a_{n-1} - 2a_{n-2}$ 对 $n \geqslant 3$ 成立. 求通项公式.

5. 数列 $\{a_n\}$ 满足条件 $a_1 = a_2 = 1$, 且 $a_n = 6a_{n-1} - 9a_{n-2}$ 对 $n \geqslant 3$ 成立. 求通项公式. (提示: 将数列 $\{a_n\}$ 分解为等比数列 $\{b_n\} = \{1, q, q^2, \cdots\}$ 及数列 $\{c_n\} = \{1, 2q, 3q^2, \cdots, nq^{n-1}, \cdots\}$ 的线性组合, 使 $b_n = 6b_{n-1} - 9b_{n-2}, c_n = 6c_{n-1} - 9c_{n-2}$ 对 $n \geqslant 3$ 成立.)

6. 试设计一个 7 阶幻方.

第3章 行 列 式

几何向量不共线与不共面, 可以刻画为所生成的平行四边形面积与平行六面体体积不为 0.

几何向量不共线与不共面, 推广为 n 维空间中的向量线性无关. 能不能将平行四边形面积和平行六面体体积推广为"n 维平行体体积", 用来刻画和判定 n 维向量线性无关?

n 阶行列式就是"n 维平行体体积". 我们先根据平行四边形面积与平行六面体体积的代数运算性质得出它们的算法, 再将代数算法推广到 n 维空间, 定义 n 阶行列式.

n 阶行列式可以用来判定线性相关和线性无关, 还可以得出求 n 元一次方程组唯一解的"求根公式".

§3.1 二阶与三阶行列式

3.1.1 二阶行列式的几何定义

例 1 当实系数二元一次方程组

$$\begin{cases} a_1x+b_1y=c_1, \\ a_2x+b_2y=c_2 \end{cases} \tag{3.1.1}$$

有唯一解时, 用几何方法求出它的唯一解.

解 方程组可以写成向量形式:

$$x\overrightarrow{OA}+y\overrightarrow{OB}=\overrightarrow{OC}, \tag{3.1.2}$$

其中 $A(a_1,a_2),B(b_1,b_2),C(c_1,c_2)$ 是平面直角坐标系中的点, 如图 3–1.

OB 沿顺时针方向旋转直角得 $\overrightarrow{OB'}=(b_2,-b_1)$, 与 (3.1.2) 两边作内积消去 y, 得:

$$x\overrightarrow{OA}\cdot\overrightarrow{OB'}=\overrightarrow{OC}\cdot\overrightarrow{OB'}.$$

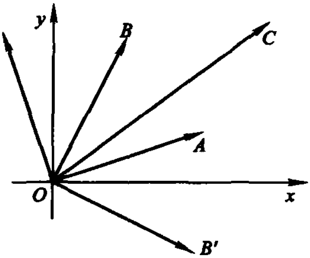

图 3–1

(3.1.2) 有唯一解 $\Leftrightarrow OA$ 与 OB 不共线 $\Leftrightarrow OA$ 与 OB' 不垂直 $\Leftrightarrow \overrightarrow{OA}\cdot\overrightarrow{OB'}\neq 0$. 此时

$$x=\frac{\overrightarrow{OC}\cdot\overrightarrow{OB'}}{\overrightarrow{OA}\cdot\overrightarrow{OB'}}=\frac{(c_1,c_2)\cdot(b_2,-b_1)}{(a_1,a_2)\cdot(b_2,-b_1)}=\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1}.$$

OA 沿逆时针方向旋转直角得 $\overrightarrow{OA'}=(-a_2,a_1)$, 与 (3.1.1) 两边作内积消去 x, 得:

$$y\overrightarrow{OB}\cdot\overrightarrow{OA'}=\overrightarrow{OC}\cdot\overrightarrow{OA'},\quad y=\frac{\overrightarrow{OC}\cdot\overrightarrow{OA'}}{\overrightarrow{OB}\cdot\overrightarrow{OA'}}=\frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}.$$

这就得到原方程组 (3.1.1) 的唯一解

$$\left(\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1},\frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}\right). \tag{3.1.3}$$

□

我们来看唯一解表达式 (3.1.3) 分母中 $a_1b_2-a_2b_1=\overrightarrow{OA}\cdot\overrightarrow{OB'}$ 的几何意义. 如图 3–2,

$$\begin{aligned}\Delta=\overrightarrow{OA}\cdot\overrightarrow{OB'}&=|OA||OB'|\cos\angle AOB'\\&=|OA||OB|\cos\left(\angle AOB-\frac{\pi}{2}\right)=|OA||OB|\sin\angle AOB.\end{aligned}$$

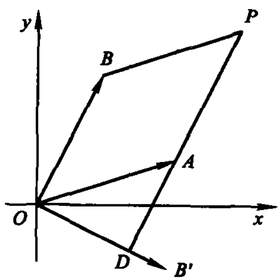

图 3–2

当 $\Delta \neq 0$ 时, $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 不共线. 此时 Δ 的绝对值就是以 OA, OB 为一组邻边的平行四边形 $OAPB$ 的面积 S_{OAPB}. Δ 的符号就是 $\sin\angle AOB$ 的符号: 当 $\angle AOB \in (0, \pi)$ 时为正, 当 $\angle AOB \in (-\pi, 0)$ 时为负. Δ 可看成平行四边形 $OAPB$ 的有向面积.

$\Delta = 0 \ \Leftrightarrow$ 平行四边形 $OAPB$ 退化为一条线段 $\Leftrightarrow$ $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 共线.

$\Delta = |OA||OB|\sin\angle AOB$ 是向量 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 的函数, 记为 $\det(\overrightarrow{OA}, \overrightarrow{OB})$. Δ 也可看作以 $\overrightarrow{OA}, \overrightarrow{OB}$ 的坐标 $\boldsymbol{a} = (a_1, a_2)^{\mathrm{T}}, \boldsymbol{b} = (b_1, b_2)^{\mathrm{T}}$ 为两列组成的方阵

$$\boldsymbol{M} = (\boldsymbol{a}, \boldsymbol{b}) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

的函数, 还可看作 $\boldsymbol{M}$ 的 4 个元素 a_1, b_1, a_2, b_2 的函数, 记作

$$\det \boldsymbol{M} \quad \text{或} \quad |\boldsymbol{M}| \quad \text{或} \quad \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$

称为 **2 阶行列式** (determinant of order 2).

3.1.2 利用代数性质计算二阶行列式

虽然例 1 中利用内积 $\overrightarrow{OA} \cdot \overrightarrow{OB'} = (a_1, a_2) \cdot (b_2, -b_1) = a_1 b_2 - a_2 b_1$ 得出了二阶行列式 $\Delta = \det(\overrightarrow{OA}, \overrightarrow{OB})$ 的计算公式, 但是, 为了推广到 n 阶行列式, 我们先研究二阶行列式的代数性质, 再由代数性质重新得出行列式的计算公式.

引理 3.1.1 二阶行列式

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \det(\boldsymbol{a}, \boldsymbol{b}) = \det(\overrightarrow{OA}, \overrightarrow{OB})$$

$$= |OA||OB|\sin\angle AOB = \overrightarrow{OA} \cdot \overrightarrow{OB'}$$

有如下性质:

性质 1 可以看作两列的乘积, 按分配律展开:
$\det(\boldsymbol{a}+\boldsymbol{h}, \boldsymbol{b}) = \det(\boldsymbol{a}, \boldsymbol{b}) + \det(\boldsymbol{h}, \boldsymbol{b})$, $\det(\boldsymbol{a}, \boldsymbol{b}+\boldsymbol{h}) = \det(\boldsymbol{a}, \boldsymbol{b}) + \det(\boldsymbol{a}, \boldsymbol{h})$.
即

$$\begin{vmatrix} a_1 + h_1 & b_1 \\ a_2 + h_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} h_1 & b_1 \\ h_2 & b_2 \end{vmatrix},$$

$$\begin{vmatrix} a_1 & b_1 + h_1 \\ a_2 & b_2 + h_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} a_1 & h_1 \\ a_2 & h_2 \end{vmatrix}.$$

性质 2 每列的公因子可以提出来: $\det(\lambda\boldsymbol{a},\boldsymbol{b})=\lambda\det(\boldsymbol{a},\boldsymbol{b})=\det(\boldsymbol{a},\lambda\boldsymbol{b})$. 即

$$\begin{vmatrix}\lambda a_1 & b_1\\ \lambda a_2 & b_2\end{vmatrix}=\lambda\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}=\begin{vmatrix}a_1 & \lambda b_1\\ a_2 & \lambda b_2\end{vmatrix}.$$

性质 3 两列互换, 行列式 Δ 变号: $\det(\boldsymbol{a},\boldsymbol{b})=-\det(\boldsymbol{b},\boldsymbol{a})$. 即

$$\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}=-\begin{vmatrix}b_1 & a_1\\ b_2 & a_2\end{vmatrix}.$$

性质 4 两列成比例, 行列式为 0: $\det(\lambda\boldsymbol{a},\mu\boldsymbol{a})=0$. 即

$$\begin{vmatrix}\lambda a_1 & \mu a_1\\ \lambda a_2 & \mu a_2\end{vmatrix}=0=\begin{vmatrix}a_1 & a_1\\ a_2 & a_2\end{vmatrix}.$$

证明 见附录 3. □

利用以上代数性质可以得出二阶行列式的计算公式.

行列式

$$\Delta=\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}$$

的两列 $\boldsymbol{a},\boldsymbol{b}$ 都可以分解为自然基向量 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 的线性组合

$$\boldsymbol{a}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2=\begin{pmatrix}a_1\\0\end{pmatrix}+\begin{pmatrix}0\\a_2\end{pmatrix},\quad \boldsymbol{b}=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2=\begin{pmatrix}b_1\\0\end{pmatrix}+\begin{pmatrix}0\\b_2\end{pmatrix},$$

按照性质 1 与性质 2, 行列式 $\det(\boldsymbol{a},\boldsymbol{b})$ 看成这两个线性组合的乘积 $\boldsymbol{a}*\boldsymbol{b}$, 就像是以 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 为字母的两个一次多项式相乘, 按照乘法分配律展开成 4 项之和:

$$\begin{aligned}\Delta&=(a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2)*(b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2)\\&=a_1b_1(\boldsymbol{e}_1*\boldsymbol{e}_1)+a_1b_2(\boldsymbol{e}_1*\boldsymbol{e}_2)+a_2b_1(\boldsymbol{e}_2*\boldsymbol{e}_1)+a_2b_2(\boldsymbol{e}_2*\boldsymbol{e}_2).\end{aligned}\tag{3.1.4}$$

注意, 这个乘法不满足交换律, 而是按照性质 3 有 $\boldsymbol{e}_2*\boldsymbol{e}_1=-\boldsymbol{e}_1*\boldsymbol{e}_2$. 再由性质 4 有 $\boldsymbol{e}_1*\boldsymbol{e}_1=0$, $\boldsymbol{e}_2*\boldsymbol{e}_2=0$. 代入 (3.1.4) 得:

$$\Delta=0+a_1b_2(\boldsymbol{e}_1*\boldsymbol{e}_2)-a_2b_1(\boldsymbol{e}_1*\boldsymbol{e}_2)+0=(a_1b_2-a_2b_1)\det(\boldsymbol{e}_1,\boldsymbol{e}_2).\tag{3.1.5}$$

将 $\det(\boldsymbol{e}_1,\boldsymbol{e}_2)=|\mathrm{e}_1||\mathrm{e}_2|\sin 90^\circ=1$ 代入 (3.1.5) 得

$$\Delta=\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}=a_1b_2-a_2b_1.\tag{3.1.6}$$

注意, $\det(\boldsymbol{e}_1, \boldsymbol{e}_2)$ 是由坐标轴上的单位向量 $\overrightarrow{OE_1} = \boldsymbol{e}_1$ 与 $\overrightarrow{OE_2} = \boldsymbol{e}_2$ 决定的单位正方形 OE_1DE_2 的面积, 面积当然为 1, 如图 3–3. 行列式 $\Delta = \det(\overrightarrow{OA}, \overrightarrow{OB})$ 就是用这个单位正方形 OE_1DE_2 的面积去度量平行四边形 $OAPB$ 的有向面积得到的量数.

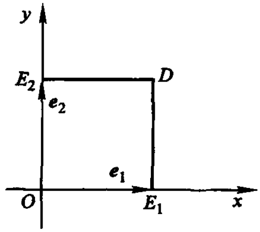

图 3–3

这个结论既然在推导二阶行列式计算公式中用到, 也作为一个性质列出来:

性质 0

$$\det(\boldsymbol{e}_1, \boldsymbol{e}_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

将由 (3.1.4) 到 (3.1.6) 的计算过程用行列式的记号写出来, 就是

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} a_1 & 0 \\ 0 & b_2 \end{vmatrix} + \begin{vmatrix} 0 & b_1 \\ a_2 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ a_2 & b_2 \end{vmatrix} \qquad (3.1.4')$$

$$= a_1b_2 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_2b_1 \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = a_1b_2 - a_2b_1. \qquad (3.1.5'), (3.1.6')$$

例 2 求行列式:

(1) $\begin{vmatrix} a_1 & b_1 \\ 0 & b_2 \end{vmatrix}$; (2) $\begin{vmatrix} a_1 & 0 \\ a_2 & b_2 \end{vmatrix}$.

解 由计算公式 (3.1.5) 得到这两个行列式的值都是 a_1b_2. □

例 2 的结论是: 二阶上三角阵或下三角阵的行列式都等于它的对角元的乘积.

例 3 以二阶实方阵 $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ 的两行为坐标作几何向量 $\overrightarrow{OA_1} = (a_1, b_1)$ 与 $\overrightarrow{OA_2} = (a_2, b_2)$, 求 OA_1, OA_2 决定的平行四边形的有向面积 $D = |OA_1||OA_2| \sin \angle A_1OA_2$.

解

$$D = \det(\overrightarrow{OA_1}, \overrightarrow{OA_2}) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1. \qquad \square$$

例 3 所求的就是 $\boldsymbol{A}$ 的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的行列式. 例 3 的结论就是: $\det(\boldsymbol{A}^{\mathrm{T}})=\det\boldsymbol{A}$.

$\boldsymbol{A}^{\mathrm{T}}$ 的各列就是 $\boldsymbol{A}$ 的各行. 关于行列式 $\det(\boldsymbol{A}^{\mathrm{T}})$ 的各列的性质对 $\det\boldsymbol{A}$ 的各行也都成立:

性质 1′ $|\boldsymbol{A}|$ 可以看成两行的乘积, 按分配律展开.

性质 2′ 每一行的公因子都可以提出来.

性质 3′ 两行互换位置, 行列式值 Δ 变为 $-\Delta$.

性质 4′ 两行成比例, 行列式值为 0.

3.1.3 三阶行列式的几何定义和代数性质

实系数 3 阶方阵 $\boldsymbol{M}=(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ 的行列式

$$\Delta=|\boldsymbol{M}|=\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=\begin{vmatrix} a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix}$$

定义为以它的 3 列 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为坐标的几何向量 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 的混合积

$$\Delta=\overrightarrow{OA}\cdot(\overrightarrow{OB}\times\overrightarrow{OC}),$$

当 $\Delta\neq 0$ 时, 它的绝对值 $|\Delta|$ 是以 OA,OB,OC 为棱的平行六面体的体积; Δ 的符号表示 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 成右手系 (当 $\Delta>0$) 还是左手系 (当 $\Delta<0$). $\Delta=0\Leftrightarrow OA,OB,OC$ 共面.

三阶行列式具有与二阶行列式类似的代数性质:

性质 0 自然基决定的单位正方体的体积

$$\det(\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)=\begin{vmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{vmatrix}=1.$$

性质 1 可以看成 3 列的乘积, 按乘法对于加法的分配律展开.

性质 2 任何一列的公因子可以提到行列式符号外面.

性质 3 两列互换, 行列式 Δ 变号, 值变为 $-\Delta$.

性质 4 如果某两列成比例, 则行列式为 0.

任何一个三阶行列式

$$\Delta=\begin{vmatrix} a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix}$$

的各列 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 都可以分解为自然基向量 $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ 的线性组合

$$\boldsymbol{a}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2+a_3\boldsymbol{e}_3,\quad \boldsymbol{b}=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2+b_3\boldsymbol{e}_3,\quad \boldsymbol{c}=c_1\boldsymbol{e}_1+c_2\boldsymbol{e}_2+c_3\boldsymbol{e}_3.$$

将 Δ 看成这三个线性组合的某种乘积, 类似于以 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 为字母的三个一次多项式相乘, 展开成"单项式"

$$\Delta_{(ikt)}=\det(a_i\boldsymbol{e}_i,b_k\boldsymbol{e}_k,c_t\boldsymbol{e}_t)=a_ib_kc_t\,\delta(ikt)\ (1\leqslant i,k,t\leqslant 3)$$

之和, 其中每个 $\delta(ikt)=\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)$ 是依次以自然基向量 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 为各列组成的行列式. 每个 $\Delta_{(ikt)}$ 的每列各有一个元素 a_i,b_k,c_t 与 Δ 相同 (分别位于第 i,k,t 行), 其余元素都为 0. 如果某项 $\Delta_{(ikt)}$ 对应的 3 个数字 i,k,t 中某两个相等, 则行列式 $\Delta_{(ikt)}$ 有两列成比例, 值为 0, 可以从 Δ 的展开式中删去. 因此 Δ 等于 i,k,t 两两不同的 6 个 $\Delta_{(ikt)}$ 之和:

$$\begin{aligned}\Delta&=\Delta_{(123)}+\Delta_{(132)}+\Delta_{(213)}+\Delta_{(231)}+\Delta_{(312)}+\Delta_{(321)}\\&=a_1b_2c_3\delta(123)+a_1b_3c_2\delta(132)+a_2b_1c_3\delta(213)\\&\quad+a_2b_3c_1\delta(231)+a_3b_1c_2\delta(312)+a_3b_2c_1\delta(321).\end{aligned}\tag{3.1.7}$$

以下只需再算出各个 $\delta(ikt)=\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)$.

3.1.4 排列的奇偶性

每个 (ikt) 是 3 个数字 1,2,3 的一个排列. $\delta(ikt)$ 是两两垂直的单位向量 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 决定的单位正方体的有向体积, 体积大小等于 1, 正负号由 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 的排列顺序是右手系还是左手系决定. 将 $\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)$ 经过 s 次两列互换变成 $\det(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=1$, 则 $(-1)^s\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)=1$, $\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)=(-1)^s$. 行列式 $\det(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t)$ 的两列互换对应于排列 (ikt) 中相应的两个数字的互换. 将数字排列 (ikt) 中两个数字的位置互换称为**对换** (transposition) . 如果 (ikt) 经过 s 次对换变成 (123), 则

$$\delta(ikt)=(-1)^s=\begin{cases}1 & (\text{当 } s \text{ 为偶数}),\\-1 & (\text{当 } s \text{ 为奇数}).\end{cases}$$

当 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 成右手系时, $\delta(ikt)=(-1)^s=1$ 为正, s 一定是偶数, 这样的排列 (ikt) 称为**偶排列** (even permutation). 当 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 成左手系时, $\delta(ikt)=(-1)^s=-1$ 为负, s 一定是奇数, 这样的排列 (ikt) 称为**奇排列** (odd permutation) . $\delta(ikt)=(-1)^s$ 是排列 (ikt) 的函数, 函数值取 1 或 -1 代表了排列的奇偶性, 我们称 $\delta(ikt)$ 为排列的奇偶性符号函数.

具体计算结果为:

$\delta(123)=1$.

$(132)\to(123)$, $s=1$, $\delta(132)=-1$.

$(213)\to(123)$, $s=1$, $\delta(213)=-1$.

$(231)\to(213)\to(123)$, $s=2$, $\delta(231)=1$.

$(312)\to(132)\to(123)$, $s=2$, $\delta(312)=1$.

$(321)\to(312)\to(132)\to(123)$, $s=3$, $\delta(321)=-1$.

以上采用的是如下标准程序: 先将 (ikt) 中的 1 与排在它前面的 s_1 个数字依次对换, 将 1 排到第一位. 再将 2 与排在它前面且比 2 大的 s_2 个数字依次对换, 将 2 排到第二位. 这就将 (ikt) 变成了 (123), $s=s_1+s_2$.

虽然将 (ikt) 变成 (123) 的过程并不唯一, 得到的 s 也不唯一. 但 $\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_t$ 是右手系还是左手系是唯一的, 因此 $\delta(ikt)=(-1)^s$ 的正负号也是确定的, 与 (ikt) 变成 (123) 的过程无关, 即 s 的奇偶性是唯一的. 例如, (321) 可以直接将 3 与 1 对换得到 (123), 从而 $s=1$, $\delta(321)=-1$. 比以上的 $s=3$ 更简便.

以上标准程序不一定最简便, 但最大的优点是: 不需要实施对换过程, 根据排列 (ikt) 可直接知道排在 1 前面的数字个数 s_1, 以及排在 2 前面且比 2 大的数字个数 s_2. 从而直接得到 $\delta(ikt)=(-1)^{s_1+s_2}$. 例如, 在 (321) 中, 排在 1 前面的数字为 3,2, 共有两个, $s_1=2$; 排在 2 前面且比 2 大的数字为 3, 共有 1 个, $s_2=1$. 不需进行对换就知道了 $s=2+1=3$ 是奇数, $\delta(321)=-1$.

3.1.5 三阶行列式计算公式

将各 $\delta(ikt)$ 的值代入 (3.1.7), 就得到三阶行列式计算公式

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \sum_{(ikt)} \delta(ikt)a_ib_kc_t \tag{3.1.8}$$

$$= a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1.$$

为了帮助记忆, 将三阶行列式中添正号的项与添负号的项分别集中写到一起:

$$\varDelta = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1,$$

观察添正号和添负号的项中的三个因子 a_i,b_k,c_t 的位置走向. 将每项的 3 个因子在矩阵中的位置用线连接起来, 如图 3–4. 可以发现, 走向与主对角线 (从左上角到右下角) 相同或平行的项添正号 (用实线表示), 走向与副对角线 (从右上角到左下角) 相同或平行的项添负号 (用虚线表示).

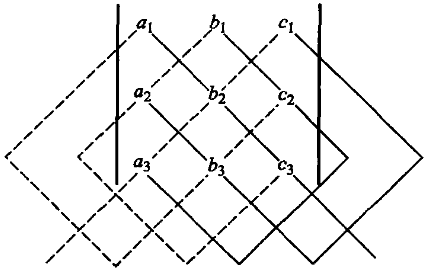

图 3-4

例 4 求三阶行列式:

$$(1)\begin{vmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{vmatrix}; \qquad (2)\begin{vmatrix} a_1 & 0 & 0 \\ a_2 & b_2 & 0 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

解 根据三阶行列式计算公式 (3.1.8), 两个小题 (1),(2) 中的行列式都等于 $a_1b_2c_3$. □

例 4 的结论是: 上三角与下三角 3 阶方阵 (包括对角阵) 的行列式等于对角元的乘积.

3.1.6 初等变换与行列式

直接用公式 (3.1.8) 计算三阶行列式, 共有 6 项, 计算起来太繁琐. 如果先将方阵用初等变换化成三角阵, 展开式只有一项, 容易计算.

要利用初等变换化简行列式, 必须先知道初等变换对行列式的影响. 在三阶行列式的性质中, 有两条性质讲了初等变换对行列式的影响:

性质 3 两列互换, 行列式 Δ 变号, 值变为 $-\Delta$.

性质 2 每列的公因子可以提出来. 也就是: 某列乘常数 λ, 行列式 Δ 的值变为 $\lambda\Delta$.

两列互换, 某列乘非零常数 λ, 是两类初等列变换. 性质 3 与性质 2 分别说明了这两类初等列变换对行列式 Δ 值的影响.

还有一类更重要的初等列变换是: 某列的常数倍加到另一列. 我们来看它对行列式值的影响.

例如, 将 Δ 的第 1 列的 λ 倍加到第 2 列, 得到

$$\det(\boldsymbol{a}, \boldsymbol{b}+\lambda\boldsymbol{a}, \boldsymbol{c}) \xlongequal{\text{性质 }1,2} \det(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) + \lambda\det(\boldsymbol{a}, \boldsymbol{a}, \boldsymbol{b}) \xlongequal{\text{性质 }4} \Delta + 0 = \Delta.$$

同理, 可以证明更一般的结论:

性质 5 三阶行列式的任何一列的常数倍加到另外任何一列, 行列式值不变.

例 5 求三阶行列式

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 9 \end{vmatrix}.$$

解 利用初等列变换将 $\Delta = |\boldsymbol{A}|$ 的方阵 $\boldsymbol{A}$ 化成三角阵, 再计算行列式.

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 9 \end{vmatrix} \underset{-(1)+(2),-(1)+(3)}{=\!=\!=\!=\!=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 8 \end{vmatrix} \underset{-2(2)+(3)}{=\!=\!=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 4 \end{vmatrix}$$

$$= 1 \times 1 \times 4 = 4.$$ □

在第 1 章中, 我们用箭头 $\xrightarrow{\lambda(i)+(j)}$ 的 $\lambda(i)+(j)$ 表示行变换: 将第 i 行的 λ 倍加到第 j 行. 我们约定: 将同样的 $\lambda(i)+(j)$ 写在箭头或等号下方表示列变换. 例如, 例 5 中等号下方的 $-2(2)+(3)$ 表示将第 2 列的 -2 倍加到第 3 列.

例 6 三元一次方程组

$$\begin{cases} a_1x + b_1y + c_1z = d_1, \\ a_2x + b_2y + c_2z = d_2, \\ a_3x + b_3y + c_3z = d_3 \end{cases} \tag{3.1.9}$$

的系数矩阵 $\boldsymbol{A}$ 的行列式 $\Delta = |\boldsymbol{A}| \neq 0$. 试求由方程组各项系数计算唯一解 (x, y, z) 的公式.

解 设 (x, y, z) 是方程组的唯一解. 分别用列向量等式

$$\begin{pmatrix} a_1x + b_1y + c_1z \\ a_2x + b_2y + c_2z \\ a_3x + b_3y + c_3z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \tag{3.1.10}$$

的左边与右边取代系数行列式 Δ 的第 1 列, 得到的两个行列式相等:

$$\begin{vmatrix} a_1x + b_1y + c_1z & b_1 & c_1 \\ a_2x + b_2y + c_2z & b_2 & c_2 \\ a_3x + b_3y + c_3z & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}. \tag{3.1.11}$$

将 (3.1.11) 左边的行列式的第 2 列的 $-y$ 倍加到第 1 列, 第 3 列的 $-z$ 倍加到

第 1 列, 再将第 1 列的公因子 x 提出来, 得:

$$x\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, \tag{3.1.12}$$

将 (3.1.12) 右边的行列式记为 Δ_1. 则 (3.1.12) 成为 $x\Delta = \Delta_1$. 由于 $\Delta \neq 0$, 得 $x = \dfrac{\Delta_1}{\Delta}$.

类似地, 分别用 (3.1.10) 等号两边的列向量代替 Δ 的第 2 列, 可得 $y\Delta = \Delta_2$, $y = \dfrac{\Delta_2}{\Delta}$; 分别用 (3.1.10) 等号两边的列向量代替 Δ 的第 3 列, 可得 $z\Delta = \Delta_3$, $z = \dfrac{\Delta_3}{\Delta}$. 其中 Δ_2, Δ_3 分别是用常数项组成的列向量代替 Δ 的第 2 列、第 3 列得到的行列式.

利用三阶行列式计算公式可以将 $\Delta, \Delta_1, \Delta_2, \Delta_3$ 表示为方程组 (3.1.9) 各系数的多项式, 从而将方程组的唯一解 $(x, y, z) = \left(\dfrac{\Delta_1}{\Delta}, \dfrac{\Delta_2}{\Delta}, \dfrac{\Delta_3}{\Delta}\right)$ 表示成方程组系数的公式. □

3.1.7 转置矩阵的行列式

根据三阶行列式的计算公式可以验证, $\boldsymbol{A}$ 的转置方阵 $\boldsymbol{A}^{\mathrm{T}}$ 的行列式 $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}|$.

对二阶或三阶行列式 $\Delta = |\boldsymbol{A}|$, 定义 $\Delta^{\mathrm{T}} = |\boldsymbol{A}^{\mathrm{T}}|$, 称为 $\Delta = |\boldsymbol{A}|$ 的转置. 则有:

性质 6 行列式转置, 值不变: $\Delta^{\mathrm{T}} = \Delta$.

由此可以推出: 对三阶行列式的列成立的各个性质也对行成立.

特别地, 将任何一行的常数倍加到另一行, 行列式不变.

例 5 解法 2 利用初等行变换将 $\Delta = |\boldsymbol{A}|$ 的方阵 $\boldsymbol{A}$ 化成三角阵, 再计算行列式.

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 9 \end{vmatrix} \xlongequal{-(1)+(2),-(1)+(3)} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{vmatrix} \xlongequal{-2(2)+(3)} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{vmatrix} = 4. \qquad □$$

习题　3.1

1. 已知平面直角坐标系中的三点 $A(1,-2), B(2,5), C(-4,1)$. 求 $\triangle ABC$ 的面积.

2. 已知空间直角坐标系中的四点 $A(1,-1,-3), B(2,0,-2), C(0,-5,-12), D(2,1,0)$.

(1) 求以 AB, AC, AD 为棱的平行六面体的体积;

(2) 判断 AB, AC, AD 成右手系还是左手系;

(3) 求平面 ABC 上任意一点 $P(x,y,z)$ 的坐标满足的充分必要条件.

3. 计算行列式

(1) $\begin{vmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & 0 & c_3 \end{vmatrix}$;　(2) $\begin{vmatrix} a_1 & b_1 & c_1 \\ 0 & b_2 & c_2 \\ 0 & b_3 & c_3 \end{vmatrix}$;　(3) $\begin{vmatrix} a & b & b \\ b & a & b \\ b & b & a \end{vmatrix}$;

(4) $\begin{vmatrix} x & y & x+y \\ y & x+y & x \\ x+y & x & y \end{vmatrix}$;　(5) $\begin{vmatrix} a^2 & (a+1)^2 & (a+2)^2 \\ b^2 & (b+1)^2 & (b+2)^2 \\ c^2 & (c+1)^2 & (c+2)^2 \end{vmatrix}$.

4. 已知 x, y, z 两两不相等. 求证:

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} \neq 0.$$

附录 3　二阶与三阶行列式的性质

1. 由几何定义推导二阶行列式性质

作 $\overrightarrow{OA} = \boldsymbol{a}$, $\overrightarrow{OB} = \boldsymbol{b}$. 则有:

性质 3　$\det(\boldsymbol{a},\boldsymbol{b}) = |OA||OB|\sin\angle AOB = -|OB||OA|\sin\angle BOA = -\det(\boldsymbol{b},\boldsymbol{a})$.

性质 4　当 $\boldsymbol{a},\boldsymbol{b}$ 成比例时, OA, OB 共线, $\angle AOB = 0$ 或 π, $\sin\angle AOB = 0$. 于是 $\det(\boldsymbol{a},\boldsymbol{b}) = |OA||OB|\sin\angle AOB = 0$.

性质 1 与性质 2　将 OB 绕原点 O 沿顺时针方向旋转直角得到 $\boldsymbol{b}' = \overrightarrow{OB'}$. 则

$$\det(\lambda\boldsymbol{a}+\boldsymbol{h},\boldsymbol{b}) = (\lambda\boldsymbol{a}+\boldsymbol{h})\cdot\boldsymbol{b}' = \lambda(\boldsymbol{a}\cdot\boldsymbol{b}') + \boldsymbol{h}\cdot\boldsymbol{b}' = \lambda\det(\boldsymbol{a},\boldsymbol{b}) + \det(\boldsymbol{h},\boldsymbol{b}).$$

再由性质 3 得

$$\begin{aligned}\det(\boldsymbol{a},\lambda\boldsymbol{b}+\boldsymbol{h}) &= -\det(\lambda\boldsymbol{b}+\boldsymbol{h},\boldsymbol{a}) = -\lambda\det(\boldsymbol{b},\boldsymbol{a}) - \det(\boldsymbol{h},\boldsymbol{a}). \\ &= \lambda\det(\boldsymbol{a},\boldsymbol{b}) + \det(\boldsymbol{a},\boldsymbol{h}).\end{aligned}$$

取 $\lambda = 1$ 得性质 1:

$$\det(\boldsymbol{a}+\boldsymbol{h},\boldsymbol{b}) = \det(\boldsymbol{a},\boldsymbol{b}) + \det(\boldsymbol{h},\boldsymbol{b}),\ \det(\boldsymbol{a},\boldsymbol{b}+\boldsymbol{h}) = \det(\boldsymbol{a},\boldsymbol{b}) + \det(\boldsymbol{a},\boldsymbol{h}).$$

取 $\boldsymbol{h}=\mathbf{0}$ 得性质 2: $\det(\lambda\boldsymbol{a},\boldsymbol{b})=\lambda\det(\boldsymbol{a},\boldsymbol{b})=\det(\boldsymbol{a},\lambda\boldsymbol{b})$.

2. 图解二阶行列式性质

由图 3–5 看出 $S_{OHQB}=S_{OAPB}+S_{AHQP}$, 即 $\det(\boldsymbol{a}+\boldsymbol{h},\boldsymbol{b})=\det(\boldsymbol{a},\boldsymbol{b})+\det(\boldsymbol{h},\boldsymbol{b})$, 如性质 1 所说.

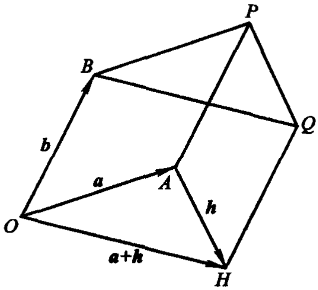

图 3–5

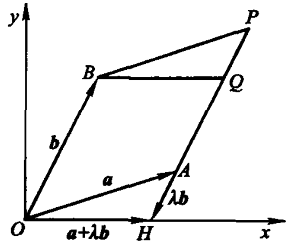

图 3–6

如果取 $\overrightarrow{AH}=\boldsymbol{h}=\lambda\boldsymbol{b}$ 与 $\overrightarrow{OB}$ 平行, AH 与平行四边形 $OAPB$ 的边 AP 共线, 如图 3–6. 则平行四边形 $AHQP$ 退化为一条线段, 面积 $S_{AHQP}=\det(\lambda\boldsymbol{b},\boldsymbol{b})=0$, 如性质 4 所说. 此时平行四边形 $OHQB$ 与 $OAPB$ 有公共底边 OB, 且高相等, 因而面积相等: $S_{OHQB}=S_{OAPB}+S_{AHQP}=S_{OAPB}$. 即

$$\det(\boldsymbol{a}+\lambda\boldsymbol{b},\boldsymbol{b})=\det(\boldsymbol{a},\boldsymbol{b})+\det(\lambda\boldsymbol{b},\boldsymbol{b})=\det(\boldsymbol{a},\boldsymbol{b}).$$

这说明了二阶行列式的又一个重要性质:

性质 5 将行列式某列 (某行) 的常数倍加到另一列 (另一行), 行列式值不变.

特别地, 如果取 H 为直线 AP 与 x 轴的交点 $(x_1,0)$, 也就是选 λ 使 $\overrightarrow{OH}=\boldsymbol{a}+\lambda\boldsymbol{b}$ 的纵坐标为 0, 平行四边形 $OHQB$ 的底边 OH 在 x 轴上, 如图 3–6. 则 $B(b_1,b_2)$ 的纵坐标 b_2 就是平行四边形 $OHQB$ 在底边 OH 上的高. 则 $S_{OHQB}=x_1b_2$. 即: $\det(\overrightarrow{OH},\overrightarrow{OB})=\begin{vmatrix}x_1 & b_1\\ 0 & b_2\end{vmatrix}=x_1b_2$. 这正是例 2 所说的: 上三角阵的行列式等于对角元的乘积.

具体写出 $S_{OAPB}=\det(\boldsymbol{a},\boldsymbol{b})=\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}$ 可知, 当 $b_2\neq 0$ 时可以将第二列的 $-\dfrac{a_2}{b_2}$ 倍加到第一列, 将行列式化成上三角形:

$$\begin{vmatrix}a_1 & b_1\\ a_2 & b_2\end{vmatrix}=\begin{vmatrix}a_1-\dfrac{a_2}{b_2}b_1 & b_1\\ 0 & b_2\end{vmatrix}=\left(a_1-\frac{a_2}{b_2}b_1\right)b_2=a_1b_2-a_2b_1\,.$$

其中的 $a_1-\dfrac{a_2}{b_2}b_1$ 就是 H 的横坐标 x_1 .

3. 三阶行列式转置公式的证明

利用三阶行列式计算公式可以直接验证转置公式 $\Delta^{\mathrm{T}} = \Delta$. 但这样的验证缺乏启发性, 使这个等式的成立显得偶然, 而且不容易推广到 n 阶行列式. 如下的推理可以让我们对等式的必然性有所体会.

在计算三阶行列式 $|\boldsymbol{A}|$ 的时候, 我们先将它分解为 6 个简单行列式 $\Delta_{(ikt)} = \det(a_i\boldsymbol{e}_i, b_k\boldsymbol{e}_k, c_t\boldsymbol{e}_t)$ 之和, 每个 $\Delta_{(ikt)}$ 的每行每列有一个位置 (依次在各列的第 i,k,t 行) 的元素 a_i, b_k, c_t 与 Δ 相同, 其余元素都是 0. 每个 $\Delta_{(ikt)}$ 又经过各列提取公因子化成更简单的行列式 $\delta(ikt) = \det(\boldsymbol{e}_i, \boldsymbol{e}_k, \boldsymbol{e}_t)$ 的 $a_ib_kc_t$ 倍, $\delta(ikt)$ 由 $\Delta_{(ikt)}$ 将 a_i, b_k, c_t 变成 1 得到. 将 $|\boldsymbol{A}|$ 变成它的转置 $|\boldsymbol{A}^{\mathrm{T}}|$, $|\boldsymbol{A}|$ 分解得到的每个 $\Delta_{(ikt)}$ 及 $\delta(ikt)$ 相应地被转置为 $\Delta_{(ikt)}^{\mathrm{T}}$ 与 $\delta(ikt)^{\mathrm{T}}$, $|\boldsymbol{A}^{\mathrm{T}}|$ 等于各 $\Delta_{(ikt)}^{\mathrm{T}} = a_ib_kc_t\delta(ikt)^{\mathrm{T}}$ 之和. 只要证明每个 $\delta(ikt)^{\mathrm{T}} = \delta(ikt)$, 则 $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}|$ 成立.

以

$$\Delta_{(231)} = \begin{vmatrix} 0 & 0 & c_1 \\ a_2 & 0 & 0 \\ 0 & b_3 & 0 \end{vmatrix} = a_2b_3c_1\delta(231) = a_2b_3c_1\begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix},$$

$$\Delta_{(231)}^{\mathrm{T}} = \begin{vmatrix} 0 & a_2 & 0 \\ 0 & 0 & b_3 \\ c_1 & 0 & 0 \end{vmatrix} = a_2b_3c_1\delta(231)^{\mathrm{T}} = a_2b_3c_1\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}$$

为例. 当 $\delta(231)$ 经过 s 次两列互换变成 $\delta(123)$, $\delta(231)^{\mathrm{T}}$ 经过 s 次相应的两行互换变成 $\delta(123)^{\mathrm{T}} = \delta(123)$:

$$\delta(231) = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} \xrightarrow[(2,3)]{} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} \xrightarrow[(1,2)]{} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \delta(123),$$

$$\delta(231)^{\mathrm{T}} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} \xrightarrow{(2,3)} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} \xrightarrow{(1,2)} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \delta(123).$$

易见 $\delta(231)^{\mathrm{T}}$ 到 $\delta(123)$ 的两次两行互换 $\xrightarrow{(2,3)}$, $\xrightarrow{(1,2)}$ 可以通过两次两列互换 $\xrightarrow[(2,3)]{}$, $\xrightarrow[(1,2)]{}$ 实现. 因此 $\delta(231)^{\mathrm{T}} = (-1)^2\delta(123) = \delta(231)$.

每个 $\delta(ikt)$ 可经过 s 次两列互换变成 $\delta(123)$, 则 $\delta(ikt)^{\mathrm{T}}$ 经过相应的 s 次两行互换变成 $\delta(123)$. 由 $\delta(ikt)^{\mathrm{T}}$ 经过两行互换得到的每个 δ 中, 每行各有唯一的非零元 1. 将 δ 的某两行互换, 就是将这两行所含的两个 1 互换. 可以通过将这两个 1 所在的两列互换来实现, 行列式变号. 经过 s 次两行互换得到 $\delta(ikt)^{\mathrm{T}} = (-1)^s\delta(123) = \delta(ikt)$.

由此可得 $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}|$.

§3.2 n 阶行列式的定义与性质

3.2.1 n 阶行列式的定义

仿照三阶行列式计算公式 (3.1.8), 对任意 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 定义行列式

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}=\sum_{(i_1\cdots i_n)}\delta(i_1i_2\cdots i_n)\,a_{i_11}a_{i_22}\cdots a_{i_nn}, \tag{3.2.1}$$

其中 $(i_1i_2\cdots i_n)$ 取遍 $1,2,\cdots,n$ 的全部 $n!$ 个排列; $a_{i_11}a_{i_22}\cdots a_{i_nn}$ 是在 $\boldsymbol{A}$ 的各列的不同行中取出的一组元素的乘积; $\delta(i_1i_2\cdots i_n)=\pm1$ 的正负号由排列 $(i_1i_2\cdots i_n)$ 决定, 定义如下:

对每个正整数 $k\leqslant n-1$, 设 s_k 是 $(i_1i_2\cdots i_n)$ 中排在 k 的前面且 $>k$ 的数字个数, $s=s_1+s_2+\cdots+s_{n-1}$, 则 $\delta(i_1i_2\cdots i_n)=(-1)^s$.

行列式 $|\boldsymbol{A}|$ 也记作 $\det\boldsymbol{A}$ 或 $\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)$, 其中 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 依次是 $\boldsymbol{A}$ 的各列.

用几何面积或体积定义二阶与三阶方阵的行列式时, 要求方阵的元素都是实数. 但在 n 阶方阵的行列式的上述定义 (3.2.1) 中, 矩阵元素可以是任意复数.

例 1 求 n 阶行列式:

$$(1)\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix};\qquad (2)\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

解 (1) 所求行列式 $\varDelta=|\boldsymbol{A}|$ 的矩阵 $\boldsymbol{A}=\mathrm{diag}(a_{11},\cdots,a_{nn})$ 是对角矩阵, 每列只有一个可能不为 0 的元素. 因此 $|\boldsymbol{A}|$ 只有一项

$$|\boldsymbol{A}|=\delta(12\cdots n)\,a_{11}a_{22}\cdots a_{nn}.$$

排列 $(12\cdots n)$ 中的数按从小到大的顺序排列, 其中每个 k 前面都没有比它更大的数, 因此 $s_k=0$. $\delta(12\cdots n)=(-1)^{0+\cdots+0}=1$, $|\boldsymbol{A}|=a_{11}a_{22}\cdots a_{nn}$.

(2) 所求行列式 $|\boldsymbol{A}|$ 的矩阵是上三角形矩阵. 从各列中取出两两不同行的元素的乘积 $a_{i_11}\cdots a_{i_nn}$, 第 1 列只有 a_{11} 可能不为 0, 第 2 列与 a_{11} 不同行的只有 a_{22} 可能不为 0. 依此类推, 可知 $|\boldsymbol{A}|$ 展开式中只有一项可能不为 0:

$$|\boldsymbol{A}|=\delta(12\cdots n)\,a_{11}a_{22}\cdots a_{nn}=a_{11}a_{22}\cdots a_{nn}.$$ □

由例 1 可知, 上三角形矩阵 (包括对角矩阵) 的行列式等于对角元的乘积. 类似地, 下三角形矩阵的行列式也等于对角元的乘积. 特别地, 单位矩阵 $\boldsymbol{I}=\operatorname{diag}(1,1,\cdots,1)$ 的行列式

$$|\boldsymbol{I}|=\begin{vmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{vmatrix}=1.$$

这是二阶与三阶行列式性质 0 的推广.

3.2.2 逆序数

例 2 求 4 阶行列式

$$\begin{vmatrix}0 & 0 & a_1 & 0\\ 0 & 0 & 0 & a_2\\ 0 & a_3 & 0 & 0\\ a_4 & 0 & 0 & 0\end{vmatrix}.$$

解 各列只有一个元素 a_4,a_3,a_1,a_2 有可能不为 0, 分别位于第 4, 3, 1, 2 行. 所求行列式

$$\Delta=\delta(4312)\,a_4a_3a_1a_2\ .$$

在排列 (4312) 中, 有两个数字 4,3 排在 1 前面, 因此 $s_1=2$; 有两个大于 2 的数字 4, 3 排在 2 前面, 因此 $s_2=2$; 有 1 个大于 3 的数字 4 排在 3 前面, 因此 $s_3=1$. $s=s_1+s_2+s_3=5$ 是奇数, 因此 $\delta(4312)=(-1)^s=-1$.

所求行列式 $\Delta=-a_4a_3a_1a_2$. □

在标准排列 $(12\cdots n)$ 中, n 个数字 $1,2,\cdots,n$ 全部按从小到大的顺序排列, 从排列中任取两个数字比较大小, 都是较小的数在前, 较大的数在后. 如果 $(i_1i_2\cdots i_n)$ 不是标准排列, 其中必有某两个数字 i_p,i_q 的排列顺序是颠倒的: 排在前面的 i_p 反而比排在后面的 i_q 更大, 也就是有 $p<q$ 且 $i_p>i_q$. 排列 $(i_1i_2\cdots i_n)$ 中每一对这样的数字 (i_p,i_q) 称为一个**逆序** (inverted sequence). 逆序的总数称为排列 $(i_1i_2\cdots i_n)$ 的**逆序数**, 记作 $\tau(i_1i_2\cdots i_n)$. 为了计算逆序 (i_p,i_q) 的总数, 可以按两个数中较小的数 i_q 的不同取值分类来统计个数. $i_q=1$ 的逆序 $(i_p,1)$ 的个数就是排在 1 前面的数字 i_p 的个数 s_1 . 一般地, 对每个正整数 $k\leqslant n-1$, 形如 (i_p,k) 的逆序总数就是 $(i_1i_2\cdots i_n)$ 中排在 k 前面且大于 k 的数字的个数 s_k. 于是逆序数 $\tau(i_1i_2\cdots i_n)=s=s_1+s_2+\cdots+s_{n-1}$. 行列式定义中的 $\delta(i_1\cdots i_n)=(-1)^{\tau(i_1i_2\cdots i_n)}$.

n 个元素 $1,2,\cdots,n$ 的 $n!$ 个排列 $(i_1i_2\cdots i_n)$ 按照逆序数 $\tau(i_1i_2\cdots i_n)$ 的奇偶性即 $\delta(i_1i_2\cdots i_n)=(-1)^{\tau(i_1i_2\cdots i_n)}$ 的正负号分成两类: 逆序数为偶数的排列称为**偶排列** (even permutation) , 逆序数为奇数的排列称为**奇排列** (odd permutation) . 行列式展开式中的 $n!$ 项也就按照排列的奇偶性分成两类, 偶排列 $(i_1i_2\cdots i_n)$ 对应的项 $a_{i_11}a_{i_22}\cdots a_{i_nn}$ 等于各列所取出的元素的乘积; 奇排列 $(i_1i_2\cdots i_n)$ 对应的项 $-a_{i_11}a_{i_22}\cdots a_{i_nn}$ 等于这种乘积的相反数.

在行列式 Δ 的各列的不同的行各选定一个元素 $a_{i_11},\cdots,a_{i_nn}$, 保持这 n 个元素不变, 将 Δ 中其余元素全部变成 0, 得到一个行列式 $\det(a_{i_11}e_{i_1},\cdots,a_{i_nn}e_{i_n})$, 记作 $\Delta_{(i_1\cdots i_n)}$. 这个行列式等于 $a_{i_11}\cdots a_{i_nn}\delta(i_1\cdots i_n)$, 就是 Δ 的展开式中由所选定的 $a_{i_11},\cdots,a_{i_nn}$ 的乘积决定的项. Δ 等于所有这些行列式 $\Delta_{(i_1\cdots i_n)}$ 之和. 我们将每个这样的行列式 $\Delta_{(i_1\cdots i_n)}$ 看成 Δ 的一项.

在 Δ 的每项 $\Delta_{(i_1\cdots i_n)}$ 中将 $a_{i_11},\cdots,a_{i_nn}$ 都换成 1, 得到更简单的行列式 $\det(e_{i_1},\cdots,e_{i_n})$, 它的各列是自然基向量 $e_1,\cdots,e_n$ 的一个排列, 它的值等于 $\delta(i_1\cdots i_n)$. 因此, 我们可以将刻画排列 $(i_1\cdots i_n)$ 奇偶性的函数 $\delta(i_1\cdots i_n)=(-1)^s$ 看成行列式 $\det(e_{i_1},\cdots,e_{i_n})$, Δ 看成所有的这样的行列式的线性组合.

3.2.3 行列式的性质

二阶与三阶行列式是由几何定义得出代数性质, 再得出代数算法. n 阶行列式是由代数算法定义的, 需要由算法得出性质.

定理 3.2.1 n 阶行列式

$$\Delta=\det(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{vmatrix}$$

有如下性质:

性质 1 将行列式 Δ 的任何一列 (或行) $\boldsymbol{a}$ 分解成两个列 (或行) 向量 $\boldsymbol{b},\boldsymbol{c}$ 之和, 在 Δ 中分别用 $\boldsymbol{b},\boldsymbol{c}$ 取代 $\boldsymbol{a}$ 得到行列式 $\Delta(\boldsymbol{b}),\Delta(\boldsymbol{c})$, 则 $\Delta=\Delta(\boldsymbol{b})+\Delta(\boldsymbol{c})$. 即:

$$\begin{aligned}&\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{j-1},\boldsymbol{b}+\boldsymbol{c},\boldsymbol{a}_{j+1},\cdots,\boldsymbol{a}_n)\\ =\ &\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{j-1},\boldsymbol{b},\boldsymbol{a}_{j+1},\cdots,\boldsymbol{a}_n)+\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{j-1},\boldsymbol{c},\boldsymbol{a}_{j+1},\cdots,\boldsymbol{a}_n),\end{aligned}$$

也就是

$$\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1+c_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n+c_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$
$$= \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$
$$+ \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & c_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & c_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

性质 2 任何一列 (行) 的公因子可以提到行列式符号外面:

$$\begin{aligned} &\det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \lambda\boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n) \\ =& \lambda\det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \cdots, \boldsymbol{a}_n), \end{aligned}$$

即

$$\begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & \lambda b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & \lambda b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$
$$= \lambda \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

性质 3 任意两列 (行) 互换位置, 行列式 Δ 变成 $-\Delta$.

性质 4 如果某两列 (行) 成比例, 则行列式值为 0.

性质 5 任意一列 (行) 的常数倍加到另一列 (行), 行列式值不变.

性质 6 方阵转置, 行列式值不变, 即 $\Delta^{\mathrm{T}} = \Delta$.

证明 性质 3 与性质 6 的证明参见附录 4.

对行列式 $\Delta = |\boldsymbol{A}|$ 记 $\Delta^{\mathrm{T}} = |\boldsymbol{A}^{\mathrm{T}}|$. 根据性质 6 有 $\Delta = \Delta^{\mathrm{T}}$. 只要证明了性质 1 − 5 对任意行列式 Δ 的列成立, 则由性质 1 − 5 对 Δ^{T} 的列成立知同样的性质对 Δ 的行成立.

性质 1 与性质 2 将行列式 Δ 看成以第 k 列 $\boldsymbol{a}_k = (a_{1k}, \cdots, a_{nk})^{\mathrm{T}}$ 的各分量 $a_{1k}, a_{2k}, \cdots, a_{nk}$ 为自变量的函数 $\Delta = f(\boldsymbol{a}_k) = f(a_{1k}, \cdots, a_{nk})$, 其余各列

的元素都看成常数. 则 Δ 的展开式

$$\Delta = \sum_{(i_1 i_2 \cdots i_n)} \delta(i_1 i_2 \cdots i_n)\, a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

中每一项都是某个自变量 $a_{i_k k}$ 的常数倍, Δ 是若干个这种常数倍之和, 经过合并同类项整理为各自变量 $a_{1k}, a_{2k}, \cdots, a_{nk}$ 的线性组合

$$\Delta = f(\boldsymbol{a}_k) = \lambda_1 a_{1k} + \lambda_2 a_{2k} + \cdots + \lambda_n a_{nk} = \boldsymbol{\Lambda} \boldsymbol{a}_k,$$

其中 $\boldsymbol{\Lambda} = (\lambda_1, \cdots, \lambda_n)$. 于是对任意 $\boldsymbol{b}, \boldsymbol{c} \in F^{n\times 1}$ 及 $\lambda \in F$ 有

$$\Delta(\lambda \boldsymbol{b} + \boldsymbol{c}) = f(\lambda \boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{\Lambda}(\lambda \boldsymbol{b} + \boldsymbol{c}) = \lambda \boldsymbol{\Lambda} \boldsymbol{b} + \boldsymbol{\Lambda} \boldsymbol{c} = \lambda \Delta(\boldsymbol{b}) + \Delta(\boldsymbol{c}).$$

取 $\lambda = 1$ 得性质 1: $\Delta(\boldsymbol{b} + \boldsymbol{c}) = \Delta(\boldsymbol{b}) + \Delta(\boldsymbol{c})$.

取 $\boldsymbol{c} = \boldsymbol{0}$ 得性质 2: $\Delta(\lambda \boldsymbol{b}) = \lambda \Delta(\boldsymbol{b})$.

性质 4 设 Δ 的第 p, q 两列成比例. 将这两列各提取适当的公因子 λ, μ 之后得到的行列式 Δ_0 的第 p, q 两列相等, $\Delta = \lambda\mu\Delta_0$. 将 Δ_0 的第 p, q 两列互换, 仍得到 Δ_0. 但由性质 3 知行列式两列互换将 Δ_0 变成 $-\Delta_0$, 因此 $-\Delta_0 = \Delta_0$, $\Delta_0 = 0$, $\Delta = \lambda\mu\Delta_0 = 0$.

性质 5 将 $\det(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots)$ 的第 i 列的 λ 倍加到第 j 列得到

$$\begin{aligned}
\det\ & (\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j + \lambda \boldsymbol{a}_i, \cdots) \\
&= \det(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots) + \lambda \det(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_i, \cdots) && (\text{性质 } 1,2) \\
&= \det(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots) + \lambda \cdot 0 && (\text{性质 } 4) \\
&= \det(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots). && \square
\end{aligned}$$

行列式的两列 (或两行) 成比例, 是说其中一列 (或一行) 是另一列 (或另一行) 的常数倍. 当常数 $\lambda = 1$ 或 $\lambda = 0$ 时就得到性质 4 的两种特殊情形:

性质 4′ 如果某两列 (或两行) 相等, 则行列式为 0.

性质 4″ 如果某列 (或某行) 元素全为 0, 则行列式为 0.

由于解线性方程组的需要, 我们将行列式定义为方阵的各列的函数. 根据行列式的性质 6, 行列式的行与列处于平等的地位, 因此行列式也可以按行定义如下:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(j_1 j_2 \cdots j_n)} (-1)^{\tau(j_1 j_2 \cdots j_n)} a_{1j_1} a_{2j_2} \cdots a_{nj_n}. \tag{3.2.2}$$

3.2.4 利用初等变换计算行列式

在第 1 章中我们学会了通过矩阵的初等行变换来进行线性方程组的同解变形, 将方程组化成阶梯形, 就可以直接写出它的解.

定理 3.2.1 中行列式性质 2,3,5 讲了初等行变换和列变换对行列式的影响:

性质 3 任意两列 (行) 互换位置, 行列式 Δ 变成 $-\Delta$.

性质 2 任意一列 (行) 同乘非零常数 λ, 行列式 Δ 变成 $\lambda\Delta$.

性质 5 任意一列 (行) 的常数倍加到另一列 (行), 行列式值不变.

通过矩阵的初等变换可以将行列式化成三角形, 再将对角元相乘求出行列式.

例 3 求行列式

$$\Delta=\begin{vmatrix}1&2&3&4\\1&2&0&-5\\3&-1&-1&0\\1&0&1&2\end{vmatrix}.$$

解

$$\Delta\xlongequal{-(1)+(2),-3(1)+(3),-(1)+(4)}\begin{vmatrix}1&2&3&4\\0&0&-3&-9\\0&-7&-10&-12\\0&-2&-2&-2\end{vmatrix}$$

$$\xlongequal{-\frac{1}{2}(4),(2,4),7(2)+(3)}-(-2)\begin{vmatrix}1&2&3&4\\0&1&1&1\\0&0&-3&-5\\0&0&-3&-9\end{vmatrix}\xlongequal{-(3)+(4)}2\begin{vmatrix}1&2&3&4\\0&1&1&1\\0&0&-3&-5\\0&0&0&-4\end{vmatrix}$$

$$=2\times1\times1\times(-3)(-4)=24.$$ □

在等号上方标注出了所用的初等行变换:

$\lambda(i)+(j)$ 表示将第 i 行的 λ 倍加到第 j 行, 行列式不变.

(i,j) 表示将第 i,j 两行互换, 此时行列式前面**添一个负号** (注意不要忘记添这个负号).

$\lambda^{-1}(i)$ 表示将第 i 行的公因子提到行列式外面, 即将第 i 行元素同除以 λ, 行列式前面乘 λ (注意不要忘记乘这个 λ).

如果将这些记号写在等号的下方, 就表示进行相应的初等列变换.

有时候我们也直接用文字叙述所用的初等行变换或列变换, 从上到下写在等号两旁.

例 4 计算 n 阶行列式

$$\Delta=\begin{vmatrix} x & a & a & \cdots & a \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & x \end{vmatrix}.$$

解

$$\Delta \xlongequal{\text{其余各行加到第 1 行}}$$

$$\begin{vmatrix} x+(n-1)a & x+(n-1)a & x+(n-1)a & \cdots & x+(n-1)a \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & x \end{vmatrix}$$

$$\xlongequal{\text{第 1 行提公因子 } x+(n-1)a} [x+(n-1)a]\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & x \end{vmatrix}$$

$$\xlongequal{\text{其余各列减第 1 列}} [x+(n-1)a]\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ a & x-a & 0 & \cdots & 0 \\ a & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a & 0 & 0 & \cdots & x-a \end{vmatrix}$$

$$=[x+(n-1)a](x-a)^{n-1}.$$ □

习题 3.2

1. 计算行列式

(1) $\begin{vmatrix} 0 & 0 & a_1 & b_1 \\ 0 & a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 & d_3 \\ 0 & 0 & 0 & a_4 \end{vmatrix}$; (2) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 3 \\ 3 & 0 & 0 & 2 \\ 4 & 0 & 0 & 1 \end{vmatrix}$;

(3) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 9 \end{vmatrix}$;　(4) $\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1-a & 1 & 1 \\ 1 & 1 & 1+b & 1 \\ 1 & 1 & 1 & 1-b \end{vmatrix}$;

(5) $\begin{vmatrix} a & b & c & d \\ a & a+b & a+b+c & a+b+c+d \\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d \\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d \end{vmatrix}$.

2. 行列式

$$\begin{vmatrix} \lambda-1 & 2 & -4 \\ -1 & \lambda-3 & -9 \\ -1 & 4 & \lambda-16 \end{vmatrix}$$

是 λ 的多项式. 求出其中的 3 次项、 2 次项和常数项.

3. 写出行列式 $\begin{vmatrix} x & 1 & 2 & 3 \\ x & x & 1 & 2 \\ 2 & 3 & x & 1 \\ x & 2 & 3 & x \end{vmatrix}$ 中含 x^4 和 x^3 的项.

4. (1) 求 $\tau(n(n-1)\cdots 21)$, 并讨论排列 $n(n-1)\cdots 21$ 的奇偶性.

(2) 求 $\tau(678354921)$,$\tau(87654321)$ 的奇偶性.

5. 证明:

(1) $\begin{vmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$;　(2) $\begin{vmatrix} 1 & a & a^2-bc \\ 1 & b & b^2-ac \\ 1 & c & c^2-ab \end{vmatrix} = 0$.

6. 计算 n 阶行列式

(1) $\begin{vmatrix} a_1-b_1 & a_1-b_2 & \cdots & a_1-b_n \\ a_2-b_1 & a_2-b_2 & \cdots & a_2-b_n \\ \vdots & \vdots & & \vdots \\ a_n-b_1 & a_n-b_2 & \cdots & a_n-b_n \end{vmatrix}$;　(2) $\begin{vmatrix} 1 & 2 & \cdots & n \\ -2 & 2 & & \\ \vdots & & \ddots & \\ -n & & & n \end{vmatrix}$;

(3) $\begin{vmatrix} \lambda & & & & -a_n \\ -1 & \lambda & & & -a_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & -1 & \lambda & -a_2 \\ & & & -1 & \lambda-a_1 \end{vmatrix}$;　(4) $\begin{vmatrix} x & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & x & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & a_2 & x & \cdots & a_{n-2} & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_{n-1} & x \end{vmatrix}$.

7. 满足条件 $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$ 的方阵 $\boldsymbol{A}$ 称为反对称方阵, 也称斜对称方阵. 证明: 奇数阶反对称方阵的行列式等于 0.

附录 4 排列的奇偶性与行列式性质

1. 行列式性质 3 任意两列互换位置, 行列式 Δ 变成 $-\Delta$.

证明 对每个 n 阶行列式 D, 用 $\tau(D)$ 表示将 D 的第 p 列与第 q 列互换得到的行列式. 当 Δ 经过第 p,q 两列互换变成 $\tau(\Delta)$ 时, Δ 的各项 $\Delta_{(i_1\cdots i_n)}$ 经过同样的两列互换得到 $\tau(\Delta)$ 的各项 $\tau(\Delta_{(i_1\cdots i_n)})$:

$$\tau(\Delta)=\sum_{(i_1\cdots i_n)}\tau(\Delta_{(i_1\cdots i_n)})=\sum_{(i_1\cdots i_n)}a_{i_11}\cdots a_{i_nn}\,\tau(\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n})).$$

只要证明每个 $\delta(i_1\cdots i_n)=\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n})$ 经过两列互换变成自己的相反数:

$$\tau(\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n}))=-\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n}),$$

则所需的结论 $\tau(\Delta)=-\Delta$ 得证. □

将行列式 $\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n})$ 的第 p,q 两列 $\boldsymbol{e}_{i_p},\boldsymbol{e}_{i_q}$ 互换, 也就是将 $\delta(i_1\cdots i_n)$ 对应的排列 $(i_1\cdots i_n)$ 中第 p,q 位置的两个数 i_p,i_q 对换. 我们证明:

引理 3.2.1 将排列 $(i_1\cdots i_n)$ 中任何两个数 i_p,i_q 作对换, 排列奇偶性改变, $\delta(i_1\cdots i_n)$ 变号.

证明 先设 $q=p+1$, 考虑将排列 $(i_1\cdots i_n)$ 中相邻两数 i_p,i_{p+1} 对换所引起的逆序数变化.

将相邻两数 i_p,i_{p+1} 位置互换, 这两个数的先后顺序由 (i_p,i_{p+1}) 变成 (i_{p+1},i_p). 如果 $i_p<i_{p+1}$, 对换前的 (i_p,i_{p+1}) 不是逆序, 对换后的 (i_{p+1},i_p) 是逆序, 逆序增加了一个; 如果 $i_p>i_{p+1}$, 对换前是逆序, 对换后不是逆序, 逆序减少一个. 除此之外其余数对的先后顺序 (包括其余任意两数的顺序, 以及这两数之一与其余任意一个数的顺序) 都没有改变, 这些先后顺序中的逆序数也没有改变. 因此, 经过相邻两数的对换, 逆序数增加一个或减少一个, 奇偶性都改变, $\delta(i_1\cdots i_n)$ 的正负号改变.

再考虑不相邻两个数 i_p,i_q 的对换, 其中 $p<q-1$. 设在排列中 i_p 与 i_q 之间隔着 d 个数, 如下所示:

$$(i_1,\cdots,i_p,\boxed{d\text{ 个数}},i_q,\cdots,i_n),$$

先将 i_q 往前依次与这 d 个数作对换, 经过 d 次相邻两数的对换排到这 d 个数与 i_p 之间, 再将 i_p 往后先与 i_q 作对换, 再依次与 d 个数作对换, 到达 i_q 原来的位置. 这就经过了 $d+1+d=2d+1$ 次相邻两数的对换实现了不相邻的 i_p 与 i_q 的位置互换. 每次相邻两数对换改变排列的奇偶性, $2d+1$ 是奇数, 排列的奇偶性经过 $2d+1$ 次改变, 总的效果是改变了.

这就证明了: 将排列 $(i_1\cdots i_n)$ 中的任何两个数字 i_p, i_q 互换位置, $\delta(i_1\cdots i_n)$ 改变符号. □

引理 3.2.1 的结论说明: 将行列式 $\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n})$ 的任何两列 $\boldsymbol{e}_{i_p}, \boldsymbol{e}_{i_q}$ 互换, 行列式变号. 进而证明了: 行列式 Δ 的任意两列互换位置, 行列式变号, 值变为 $-\Delta$.

对任意行列式 $\Delta=|\boldsymbol{A}|$, 将 $\boldsymbol{A}$ 的转置矩阵的行列式 $|\boldsymbol{A}^{\mathrm{T}}|$ 定义为行列式的转置 Δ^{T}.

2. 行列式性质 6 行列式转置, 值不变. 即 $\Delta^{\mathrm{T}}=\Delta$.

证明 当 Δ 经过转置得到 Δ^{T} 时, Δ 的各项 $\Delta_{(i_1\cdots i_n)}$ 经过转置得到 Δ^{T} 的各项 $(\Delta_{(i_1\cdots i_n)})^{\mathrm{T}}$:

$$\Delta^{\mathrm{T}}=\sum_{(i_1\cdots i_n)}(\Delta_{(i_1\cdots i_n)})^{\mathrm{T}}=\sum_{(i_1\cdots i_n)}a_{i_1 1}\cdots a_{i_n n}\,\delta(i_1\cdots i_n)^{\mathrm{T}},$$

只要证明每个 $\delta(i_1\cdots i_n)^{\mathrm{T}}=\delta(i_1\cdots i_n)$, 则所需结论 $\Delta^{\mathrm{T}}=\Delta$ 得证.

$\delta=\det(\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n})=\det\boldsymbol{N}$, 其中 $\boldsymbol{N}=\boldsymbol{E}_{i_1 1}+\cdots+\boldsymbol{E}_{i_n n}$. 则 $\delta^{\mathrm{T}}=\det\boldsymbol{N}^{\mathrm{T}}=\det(\boldsymbol{E}_{1i_1}+\cdots+\boldsymbol{E}_{ni_n})$.

$\boldsymbol{N}$ 可以经过 s 次两列互换变成单位阵: $\boldsymbol{N}\to\boldsymbol{N}_1\to\cdots\to\boldsymbol{N}_s=\boldsymbol{I}$, 这说明 $\det\boldsymbol{N}=(-1)^s$. $\boldsymbol{N}^{\mathrm{T}}$ 经过 s 次相应的两行互换变成单位阵: $\boldsymbol{N}^{\mathrm{T}}\to\boldsymbol{N}_1^{\mathrm{T}}\to\cdots\to\boldsymbol{N}_s^{\mathrm{T}}=\boldsymbol{I}$. 我们证明其中每次两行互换 $\boldsymbol{N}_k^{\mathrm{T}}\to\boldsymbol{N}_{k+1}^{\mathrm{T}}$ 可以通过两列互换来实现, 这将说明每个 $\det(\boldsymbol{N}_{k+1}^{\mathrm{T}})=-\det(\boldsymbol{N}_k^{\mathrm{T}})$, 从而 $\delta^{\mathrm{T}}=\det(\boldsymbol{N}^{\mathrm{T}})=(-1)^s=\delta$.

$\boldsymbol{N}$ 的各行是自然基向量的一个排列, 经过两行互换得到的每个 $\boldsymbol{N}_k$ 也是如此, 因此 $\boldsymbol{N}_k=\boldsymbol{E}_{1k_1}+\cdots+\boldsymbol{E}_{nk_n}$, 它的每行各有唯一的非零元等于 1, 分别位于不同的列 (第 $k_1,\cdots,k_n$ 列), 其余元素都是 0. 将 $\boldsymbol{N}_k$ 的第 p,q 两行互换, 就是将这两行位于第 (p,k_p) 与 (q,k_q) 两个位置的 1 分别移到第 (q,k_p) 与 (p,k_q) 两个位置, 其余的 1 位置不变. 将 $\boldsymbol{N}_k$ 的第 k_p 列与第 k_q 列互换, 产生同样的效果. 如图 3 – 7 所示:

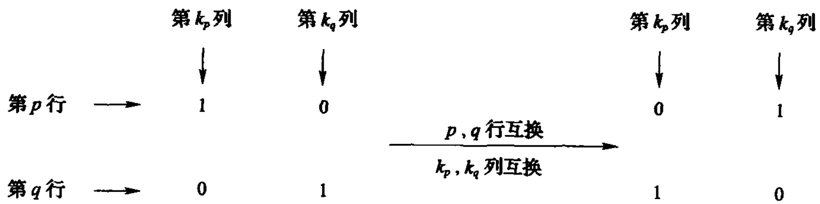

图 3–7

这证明从 $\boldsymbol{N}^{\mathrm{T}}$ 变成 $\boldsymbol{I}$ 的每次两行互换都可以通过两列互换来实现. 恰如所需. □

§3.3 线性方程组唯一解公式

3.3.1 线性方程组唯一解条件

在第 2 章中已经知道, n 元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n, \end{cases}$$

即

$$\boldsymbol{AX} = \boldsymbol{b}$$

有唯一解的充分必要条件是: 系数矩阵 $\boldsymbol{A} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$ 的各列 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性无关.

既然我们将 $\boldsymbol{A}$ 的行列式 $\Delta = \det \boldsymbol{A}$ 想象成以各列 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 为棱的 n 维 "平行体" 的体积. 很自然猜想: 各条 "棱" $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性无关 $\Leftrightarrow$ "n 维体积" $\Delta \neq 0$.

几何观念启发想象. 再用代数运算验证这个猜想是否正确.

定理 3.3.1 设 $\boldsymbol{A}$ 是由 n 个方程组成的 n 元一次方程组 $\boldsymbol{AX} = \boldsymbol{b}$ 的系数矩阵. 则:

$$\det \boldsymbol{A} \neq 0 \Leftrightarrow \text{方程组 } \boldsymbol{AX} = \boldsymbol{b} \text{ 有唯一解}.$$

证明 只需证明: $\det A \neq 0 \Leftrightarrow A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$ 的各列线性无关.

先设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性相关, 则其中某个 $\boldsymbol{a}_k$ 是其余 $\boldsymbol{a}_i$ 的线性组合:

$$\boldsymbol{a}_k = \lambda_1 \boldsymbol{a}_1 + \cdots + \lambda_{k-1}\boldsymbol{a}_{k-1} + \lambda_{k+1}\boldsymbol{a}_{k+1} + \cdots + \lambda_n \boldsymbol{a}_n.$$

对每个 $i \neq k$, 在行列式 $\Delta = \det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k, \cdots, \boldsymbol{a}_n)$ 中将第 i 列的 $-\lambda_i$ 倍 $-\lambda_i \boldsymbol{a}_i$ 加到第 k 列, 行列式值不变, 最后将第 k 列变成 $\boldsymbol{0}$. 这就得到

$$\begin{aligned} \Delta &= \det\Big(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \sum_{i \neq k} \lambda_i \boldsymbol{a}_i, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n\Big) \\ &\xlongequal[-\lambda_i(i)+(k), \forall i \neq k]{} \det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \boldsymbol{0}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n) = 0. \end{aligned}$$

这证明了: $\boldsymbol{A}$ 的各列线性相关 $\Rightarrow \det \boldsymbol{A} = 0$. 反过来有: $\det \boldsymbol{A} \neq 0 \Rightarrow \boldsymbol{A}$ 的各列线性无关.

再设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 线性无关, 则以 $\boldsymbol{A}$ 为系数矩阵的齐次线性方程组 $\boldsymbol{AX} = \boldsymbol{0}$ 有唯一解. 根据引理 1.3.1, $\boldsymbol{A}$ 可以通过有限次第三类初等行变

换变成阶梯形矩阵 $\boldsymbol{T}$. 由于第三类初等行变换不改变行列式, 所以 $\det \boldsymbol{A} = \det \boldsymbol{T}$. 由于 $\boldsymbol{AX} = \boldsymbol{0}$ 有唯一解, $\boldsymbol{T}$ 的非零行的个数 $r = n$, 每行的阶梯元 $\lambda_{1j_1}, \lambda_{2j_2}, \cdots, \lambda_{nj_n}$ 恰是 $\boldsymbol{T}$ 的全部对角元, $\boldsymbol{T}$ 是对角元 $\lambda_{11}, \cdots, \lambda_{nn}$ 全不为零的上三角形矩阵, $\det \boldsymbol{T} = \lambda_{11}\lambda_{22}\cdots\lambda_{nn} \neq 0$. 因此 $\det \boldsymbol{A} = \det \boldsymbol{T} \neq 0$.

这就证明了:

$$\det \boldsymbol{A} \neq 0 \Leftrightarrow \boldsymbol{A} \text{ 的各列线性无关 } \Leftrightarrow \text{ 方程组 } \boldsymbol{AX} = \boldsymbol{b} \text{ 有唯一解}.$$

不论 $\boldsymbol{b}$ 取什么值, 这个结论都成立. □

例 1 λ 取什么值时, 方程组

$$\begin{cases} x_2 + x_3 = \lambda x_1, \\ x_1 + x_3 = \lambda x_2, \\ x_1 + x_2 = \lambda x_3 \end{cases}$$

有非零解. 并在有非零解时求出通解.

解 移项, 整理得

$$\begin{cases} \lambda\, x_1 - \quad x_2 - \quad x_3 = 0, \\ -x_1 + \lambda\, x_2 - \quad x_3 = 0, \\ -x_1 - \quad x_2 + \lambda\, x_3 = 0. \end{cases} \tag{3.3.1}$$

方程组 (3.3.1) 有非零解的条件是系数行列式 $\Delta = 0$. 而

$$\begin{aligned}\Delta &= \begin{vmatrix} \lambda & -1 & -1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} \xlongequal{(2)+(1),(3)+(1)} \begin{vmatrix} \lambda-2 & \lambda-2 & \lambda-2 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} \\ &= (\lambda-2)\begin{vmatrix} 1 & 1 & 1 \\ -1 & \lambda & -1 \\ -1 & -1 & \lambda \end{vmatrix} \xlongequal{(1)+(2),(1)+(3)} (\lambda-2)\begin{vmatrix} 1 & 1 & 1 \\ 0 & \lambda+1 & 0 \\ 0 & 0 & \lambda+1 \end{vmatrix} \\ &= (\lambda-2)(\lambda+1)^2.\end{aligned}$$

由 $(\lambda-2)(\lambda+1)^2 = 0$ 解出 $\lambda = 2$ 或 $\lambda = -1$.

当 $\lambda = 2$, 对方程组的系数矩阵进行初等行变换得

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \xrightarrow{(2)+(1),(3)+(1),(1,3)} \begin{pmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{-(1),(1)+(2)}$$

$$\begin{pmatrix} 1 & 1 & -2 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\frac{1}{3}(2),-(2)+(1)} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

对应的方程组

$$\begin{cases} x_1 - x_3 = 0, \\ x_2 - x_3 = 0 \end{cases}$$

的通解为 $t(1,1,1)$.

当 $\lambda = -1$, 对方程组的系数矩阵进行初等行变换得

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

对应的方程组为 $x_1 + x_2 + x_3 = 0$, 通解为 $t_1(1,-1,0) + t_2(1,0,-1)$. □

例 2 设实数 a,b,c 不全为 0, α,β,γ 为任意实数. 且

$$\begin{cases} a = b\cos\gamma + c\cos\beta, \\ b = c\cos\alpha + a\cos\gamma, \\ c = a\cos\beta + b\cos\alpha. \end{cases}$$

求证:

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

证明 将 a,b,c 看成未知数, α,β,γ 看成已知数, 则以上等式为齐次线性方程组, 整理得

$$\begin{cases} -a + b\cos\gamma + c\cos\beta = 0, \\ a\cos\gamma - b + c\cos\alpha = 0, \\ a\cos\beta + b\cos\alpha - c = 0. \end{cases}$$

方程组有非零解 (a,b,c) 的条件是系数行列式 $\Delta = 0$, 即

$$\begin{vmatrix} -1 & \cos\gamma & \cos\beta \\ \cos\gamma & -1 & \cos\alpha \\ \cos\beta & \cos\alpha & -1 \end{vmatrix} = 0.$$

将左边的行列式展开并整理, 即得

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

3.3.2 行列式秩

矩阵 $\boldsymbol{A}$ 的线性无关的行的最大个数称为行秩, 线性无关的列的最大个数称为列秩. 同一个矩阵 $\boldsymbol{A}$ 的行秩与列秩相等, 称为 $\boldsymbol{A}$ 的秩, 记为 $\operatorname{rank}\boldsymbol{A}$.

如果方阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\neq 0$, 就称 $\boldsymbol{A}$ 是**非退化的** (non-degenerate). 行列式等于零的方阵则称为**退化的** (degenerate). 定理 3.3.1 证明了: 非退化方阵 $\boldsymbol{A}$ 的各列 (行) 线性无关, $\operatorname{rank}\boldsymbol{A}=n$, 这样的方阵也称为**满秩的** (full rank).

对任意的 $m\times n$ 矩阵 $\boldsymbol{A}$, 将它的 k 个不同的行与 k 个不同的列的交叉位置的元素组成的 k 阶行列式 Δ_k 称为 $\boldsymbol{A}$ 的**子式** (minor) . 如果 $\boldsymbol{A}$ 的某个 k 阶子式 $\Delta_k\neq 0$, 则 Δ_k 的 k 行线性无关, 由 Δ_k 的 k 行扩充成的 $\boldsymbol{A}$ 的 k 行也线性无关 (见引理 2.2.2), 这说明 $k\leqslant\operatorname{rank}\boldsymbol{A}$. 设 $\boldsymbol{A}$ 中不为零的子式的最大阶为 r, 也就是说: $\boldsymbol{A}$ 存在 r 阶非零子式, 不存在更大阶的非零子式, r 就称为 $\boldsymbol{A}$ 的**行列式秩** (determinant rank) . 当然也有 $r\leqslant\operatorname{rank}A$.

反过来, 设 $s=\operatorname{rank}\boldsymbol{A}$, 则 $\boldsymbol{A}$ 有某 s 列 $\boldsymbol{A}_{j_1},\cdots,\boldsymbol{A}_{j_s}$ 线性无关, 这 s 列排成的矩阵 $\boldsymbol{A}_s=(\boldsymbol{A}_{j_1},\cdots,\boldsymbol{A}_{j_s})$ 的列秩为 s. 因而 $\boldsymbol{A}_s$ 的行秩也为 s, 可见 $\boldsymbol{A}_s$ 中有某 s 行 $\boldsymbol{\alpha}_{i_1},\cdots,\boldsymbol{\alpha}_{i_s}$ 线性无关, 这 s 行排成的 s 阶方阵 $\boldsymbol{B}_s$ 的各行线性无关, 其行列式 $|\boldsymbol{B}_s|\neq 0$. 但 $|\boldsymbol{B}_s|$ 就是由 $\boldsymbol{A}$ 的第 $i_1,\cdots,i_s$ 行和第 $j_1,\cdots,j_s$ 列交叉位置的元素组成的 s 阶子式. 这说明 $\boldsymbol{A}$ 存在 s 阶非零子式. 既然 r 是非零子式的最大阶数, 就有 $r\geqslant s=\operatorname{rank}\boldsymbol{A}$. 这就证明了 $r=\operatorname{rank}\boldsymbol{A}$. 也就是:

定理 3.3.2 矩阵 $\boldsymbol{A}$ 的行列式秩 $=\boldsymbol{A}$ 的行秩 $=\boldsymbol{A}$ 的列秩. □

既然矩阵 $\boldsymbol{A}$ 的行列式秩、行秩、列秩相等, 其中任何一个都可以定义为矩阵的秩.

定义 3.3.1 矩阵 $\boldsymbol{A}$ 的非零子式的最大阶数 r 称为 $\boldsymbol{A}$ 的**秩** (rank), 记作 $\operatorname{rank}\boldsymbol{A}$. □

3.3.3 线性方程组唯一解公式——克拉默 (Cramer) 法则

例 3 设由 n 个方程组成的 n 元线性方程组

$$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b} \tag{3.3.2}$$

的系数行列式 $\Delta=\det\boldsymbol{A}\neq 0$. 试求由方程组各系数计算唯一解 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 的公式.

解 由定理 3.3.1 知, 当且仅当 $\Delta\neq 0$ 时方程组 (3.3.2) 有唯一解. 设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 依次是系数矩阵 $\boldsymbol{A}$ 的各列, 则方程组 (3.3.2) 的唯一解 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 满足等式

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}. \tag{3.3.3}$$

对每个 $1 \leqslant k \leqslant n$, 分别用等式 (3.3.3) 两边的列向量替代 Δ 的第 k 列, 得到行列式等式

$$\det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n) = \Delta_k, \tag{3.3.4}$$

等式 (3.3.4) 右边的 Δ_k 是用 $\boldsymbol{b}$ 替代 Δ 的第 k 列得到的行列式. 将等式 (3.3.4) 左边的行列式除了第 k 列之外的每一列 $\boldsymbol{a}_j$ $(j \neq k)$ 的 $-x_j$ 倍加到第 k 列, 消去第 k 列的项 $x_j\boldsymbol{a}_j$, 行列式不变. 第 k 列只剩下一项 $x_k\boldsymbol{a}_k$. 等式 (3.3.4) 变成

$$\det(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, x_k\boldsymbol{a}_k, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n) = \Delta_k,$$

将这个等式左边的行列式第 k 列的公因子 x_k 提出来, 得:

$$x_k\Delta = \Delta_k, \qquad \text{从而} \qquad x_k = \frac{\Delta_k}{\Delta}.$$

这就得到原方程组 (3.3.2) 的唯一解

$$\boldsymbol{X} = (x_1, \cdots, x_k, \cdots, x_n)^{\mathrm{T}} = \left(\frac{\Delta_1}{\Delta}, \cdots, \frac{\Delta_k}{\Delta}, \cdots, \frac{\Delta_n}{\Delta}\right)^{\mathrm{T}}. \qquad \square$$

例 3 证明了:

定理 3.3.3 (克拉默法则) 设方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

的系数行列式 $\Delta = \det \boldsymbol{A} \neq 0$, 则方程组有唯一解

$$(x_1, \cdots, x_k, \cdots, x_n) = \left(\frac{\Delta_1}{\Delta}, \cdots, \frac{\Delta_k}{\Delta}, \cdots, \frac{\Delta_n}{\Delta}\right),$$

其中每个 Δ_k $(1 \leqslant k \leqslant n)$ 是依次用各方程的常数项 $b_1, \cdots, b_n$ 替换 Δ 的第 k 列各行元素得到的行列式. $\square$

习题 3.3

1. 用克拉默 (Cramer) 法则解线性方程组:

$$\begin{cases} 2x_1 + x_2 - 5x_3 + x_4 = 4, \\ x_1 - 3x_2 - 6x_4 = 3, \\ 2x_2 - x_3 + 2x_4 = -5, \\ x_1 + 4x_2 - 7x_3 + 6x_4 = 0. \end{cases}$$

2. 分别求复数 λ 使其满足下面的条件:

(1) 向量 $(1+\lambda, 1-\lambda)$, $(1-\lambda, 1+\lambda)$ 线性相关.

(2) 向量 $(\lambda, 1, 0), (1, \lambda, 1), (0, 1, \lambda)$ 组成 $\mathbf{C}^3$ 的基.

(3) 方程组

$$\begin{cases} x_2 + x_3 = \lambda x_1, \\ x_1 + x_3 = \lambda x_2, \\ x_1 + x_2 = \lambda x_3 \end{cases}$$

有非零解.

3. 对不同的复数值 λ, 求如下方阵 $\boldsymbol{A}$ 的秩 r, 并找出 $\boldsymbol{A}$ 的一个 r 阶非零子式.

$$(1)\ \boldsymbol{A} = \begin{pmatrix} \lambda & 1 & 0 \\ 1 & \lambda & 1 \\ 0 & 1 & \lambda \end{pmatrix}; \qquad (2)\ \boldsymbol{A} = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 1 & 0 & 0 & \lambda \end{pmatrix}.$$

§3.4 展开定理

3.4.1 行列式降阶法

例 1 求行列式

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ 0 & 0 & c_3 & d_3 \\ 0 & 0 & c_4 & d_4 \end{vmatrix}.$$

解 根据引理 1.3.1, 方阵 $\boldsymbol{A}$ 可以通过有限次第三类初等行变换变成上三角形矩阵 $\boldsymbol{T}$.

对 Δ 的前两行和后两行各进行一系列第三类初等行变换, 可以将左上角和右下角的 2 阶方阵都变成上三角形矩阵, 保持这两个 2 阶方阵的行列式值不变, 也保持整个 4 阶行列式 Δ 不变. 因此

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} \lambda_1 & * \\ 0 & \lambda_2 \end{vmatrix} = \lambda_1\lambda_2; \quad \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} = \begin{vmatrix} \lambda_3 & * \\ 0 & \lambda_4 \end{vmatrix} = \lambda_3\lambda_4;$$

$$\begin{aligned} \Delta &= \begin{vmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & \lambda_4 \end{vmatrix} = \lambda_1\lambda_2\lambda_3\lambda_4 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_3 & d_3 \\ c_4 & d_4 \end{vmatrix} \\ &= (a_1b_2 - a_2b_1)(c_3d_4 - c_4d_3). \end{aligned}$$

□

例 1 的方法适用于计算具有如下形式

$$\boldsymbol{A} \doteq \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{O} & \boldsymbol{A}_4 \end{pmatrix}$$

的 n 阶方阵 $\boldsymbol{A}$ 的行列式, 其中 $\boldsymbol{A}_1, \boldsymbol{A}_4$ 分别是 r 阶方阵和 $n-r$ 阶方阵, $\boldsymbol{O}$ 是零矩阵. 如果将组成方阵 $\boldsymbol{A}$ 的 4 个小矩阵 $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{O}, \boldsymbol{A}_4$ 各看成一个整体, $\boldsymbol{A}$ 看起来就好像一个上三角形矩阵. 这样的方阵 $\boldsymbol{A}$ 称为**准上三角形矩阵**, 组成 $\boldsymbol{A}$ 的每个小矩阵称为一个**块**, 处于对角线上的 $\boldsymbol{A}_1, \boldsymbol{A}_4$ 称为**对角块**. 对 $\boldsymbol{A}$ 的前 r 行和后 $n-r$ 行各作一系列第三类初等变换, 可以将两个对角块 $\boldsymbol{A}_1, \boldsymbol{A}_4$ 分别化成上三角形矩阵 $\boldsymbol{T}_1, \boldsymbol{T}_4$, 从而将 $\boldsymbol{A}$ 化成上三角形矩阵

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{T}_1 & \boldsymbol{T}_2 \\ \boldsymbol{O} & \boldsymbol{T}_4 \end{pmatrix}.$$

我们有

$$|\boldsymbol{A}_1| = |\boldsymbol{T}_1| = \begin{vmatrix} \lambda_{11} & \cdots & * \\ & \ddots & \vdots \\ \boldsymbol{O} & & \lambda_{rr} \end{vmatrix} = \lambda_{11} \cdots \lambda_{rr},$$

$$|\boldsymbol{A}_4| = |\boldsymbol{T}_4| = \begin{vmatrix} \lambda_{r+1,r+1} & \cdots & * \\ & \ddots & \vdots \\ \boldsymbol{O} & & \lambda_{nn} \end{vmatrix} = \lambda_{r+1,r+1} \cdots \lambda_{nn}.$$

从而

$$|\boldsymbol{A}| = |\boldsymbol{T}| = \lambda_{11} \cdots \lambda_{rr} \lambda_{r+1,r+1} \cdots \lambda_{nn} = |\boldsymbol{A}_1| \cdot |\boldsymbol{A}_4|.$$

类似地, 可定义准下三角形矩阵

$$\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix},$$

其行列式

$$\begin{vmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{vmatrix} = \left| \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{O} \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix}^{\mathrm{T}} \right| = \begin{vmatrix} \boldsymbol{A}_1^{\mathrm{T}} & \boldsymbol{A}_3^{\mathrm{T}} \\ \boldsymbol{O} & \boldsymbol{A}_4^{\mathrm{T}} \end{vmatrix} = |\boldsymbol{A}_1^{\mathrm{T}}| \cdot |\boldsymbol{A}_4^{\mathrm{T}}| = |\boldsymbol{A}_1| \cdot |\boldsymbol{A}_4|.$$

这就证明了:

引理 3.4.1 准上三角形矩阵与准下三角形矩阵的行列式等于它们的对角块的行列式的乘积:

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \cdot \begin{vmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots \\ a_{n,r+1} & \cdots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} & 0 & \cdots & 0 \\ a_{r+1,1} & \cdots & a_{r+1,r} & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nr} & a_{n,r+1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix} \cdot \begin{vmatrix} a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \vdots & & \vdots \\ a_{n,r+1} & \cdots & a_{nn} \end{vmatrix}.$$

特别地

$$\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}. \quad \square$$

例 2 计算行列式

$$\varDelta = \begin{vmatrix} -1 & 0 & 0 & 0 & 0 \\ a & 2 & 3 & 0 & 0 \\ b & 4 & 5 & 0 & 0 \\ c_1 & c_2 & c_3 & 1 & 3 \\ d_1 & d_2 & d_3 & 2 & 5 \end{vmatrix}.$$

解

$$\varDelta = \begin{vmatrix} -1 & 0 & 0 \\ a & 2 & 3 \\ b & 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix} = (-1) \cdot \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix}$$
$$= (-1) \cdot (2 \times 5 - 3 \times 4) \cdot (1 \times 5 - 3 \times 2) = -2. \quad \square$$

例 3 平面上建立了直角坐标系, 求证: 对任意 n 个横坐标不同的已知点 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$ 都存在唯一的不超过 $n-1$ 次的多项式 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$, 使它的图像曲线经过这 n 个已知点.

解 问题归结为: $f(x)$ 的系数 $a_0,a_1,\cdots,a_{n-1}$ 满足的线性方程组

$$\begin{cases} a_0+a_1x_1+\cdots+a_{n-1}x_1^{n-1}=y_1, \\ a_0+a_1x_2+\cdots+a_{n-1}x_2^{n-1}=y_2, \\ \cdots\cdots\cdots\cdots \\ a_0+a_1x_n+\cdots+a_{n-1}x_n^{n-1}=y_n \end{cases} \tag{3.4.1}$$

有唯一解.

进一步归结为: 方程组 (3.4.1) 的系数行列式 $\Delta\neq 0$. 计算得

$$\Delta=V(x_1,x_2,\cdots,x_n)=\begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix}$$

$$\xlongequal[-x_1(i)+(i+1),\,\forall i=n-1,\cdots,2,1]{}\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & x_2-x_1 & x_2(x_2-x_1) & \cdots & x_2^{n-2}(x_2-x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n-x_1 & x_n(x_n-x_1) & \cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix}$$

$$=\begin{vmatrix} x_2-x_1 & x_2(x_2-x_1) & \cdots & x_2^{n-2}(x_2-x_1) \\ \vdots & \vdots & & \vdots \\ x_n-x_1 & x_n(x_n-x_1) & \cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix}$$

$$\xlongequal[]{(x_i-x_1)^{-1}(i-1),\,\forall 2\leqslant i\leqslant n}(x_2-x_1)\cdots(x_n-x_1)\begin{vmatrix} 1 & x_2 & \cdots & x_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} \end{vmatrix}$$

$$=(x_2-x_1)\cdots(x_n-x_1)\,V(x_2,\cdots,x_n).$$

可见, 从 $V(x_1,x_2,\cdots,x_n)$ 中提取因子 $x_j-x_1\ (\forall j>1)$ 得到 $V(x_2,\cdots,x_n)$. 同理, 再提取因子 $x_j-x_2\ (\forall j>2)$ 得到 $V(x_3,\cdots,x_n)$. 依此类推, 依次对 $i=1,2,\cdots,n-2$ 提取所有的 $x_j-x_i\ (j>i)$ 的乘积之后得到

$$V(x_{n-1},x_n)=\begin{vmatrix} 1 & x_{n-1} \\ 1 & x_n \end{vmatrix}=x_n-x_{n-1}.$$

可知

$$V(x_1,x_2,\cdots,x_n)=\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

这一证明过程可用数学归纳法的语言给出严格叙述 (略).

n 个已知点的横坐标 $x_1,x_2,\cdots,x_n$ 两两不同, 所以 $V(x_1,x_2,\cdots,x_n)\neq 0$, 所以满足条件的多项式 $f(x)$ 存在且唯一. □

例 3 中的行列式 $V(x_1,\cdots,x_n)$ 称为范德蒙德 (Vandermonde) 行列式.

3.4.2 行列式按一行或一列展开

n 阶行列式的展开式中有 $n!$ 项. 随着 n 的增加, $n!$ 增加得很快, 计算的复杂程度也急剧增加. 计算 3 阶行列式的 6 项已经不容易, 计算 4 阶行列式的 24 项就更艰难了.

但是, 如果行列式含有足够多的零元素, 需要计算的非零项就要少得多. 例如, 若某一行或某一列如果只有一个非零元, 就可以降阶为 $n-1$ 阶行列式来计算:

如果第 1 行只有一个非零元, 例如

$$\Delta=\begin{vmatrix}0&\cdots&0&a_{1j}&0&\cdots&0\\a_{21}&\cdots&a_{2,j-1}&a_{2j}&a_{2,j+1}&\cdots&a_{2n}\\\vdots&&\vdots&\vdots&\vdots&&\vdots\\a_{n1}&\cdots&a_{n,j-1}&a_{nj}&a_{n,j+1}&\cdots&a_{nn}\end{vmatrix},$$

将非零元所在的第 j 列依次与左边的 $j-1$ 列互换位置, 经过 $j-1$ 次变号, 化为准下三角形矩阵, 从而化为 $n-1$ 阶行列式来计算:

$$\begin{aligned}\Delta&=(-1)^{j-1}\begin{vmatrix}a_{1j}&0&\cdots&0&0&\cdots&0\\a_{2j}&a_{21}&\cdots&a_{2,j-1}&a_{2,j+1}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots&\vdots&&\vdots\\a_{nj}&a_{n1}&\cdots&a_{n,j-1}&a_{n,j+1}&\cdots&a_{nn}\end{vmatrix}\\&=(-1)^{j-1}\cdot a_{1j}\begin{vmatrix}a_{21}&\cdots&a_{2,j-1}&a_{2,j+1}&\cdots&a_{2n}\\\vdots&&\vdots&\vdots&&\vdots\\a_{n1}&\cdots&a_{n,j-1}&a_{n,j+1}&\cdots&a_{nn}\end{vmatrix}\\&=a_{1j}\cdot(-1)^{j-1}M_{1j},\end{aligned}$$

其中 M_{1j} 是在 Δ 中将 a_{1j} 所在的第 1 行和第 j 列删去之后剩下的元素组成的行列式, 称为 a_{1j} 在 Δ 中的**余子式**. 由于 $(-1)^{j-1}=(-1)^{1+j}$, 我们将 $(-1)^{j-1}M_{1j}$

改写为 $(-1)^{1+j}M_{1j}$, 其中 -1 的指数 $1+j$ 等于 a_{1j} 的行指标 1 与列指标 j 之和. $(-1)^{1+j}M_{1j}$ 称为 a_{1j} 在 Δ 中的**代数余子式** (algebraic cofactor), 记作 A_{1j}. 这样, 等式 $\Delta = a_{1j}\cdot(-1)^{j-1}M_{1j}$ 就成为

$$\Delta = a_{1j}A_{1j}.$$

对任意 n 阶行列式 Δ, 可以将它的第 1 行 $(a_{11},\cdots,a_{1n})$ 分解成自然基向量的倍向量之和:

$$(a_{11},\cdots,a_{1n}) = (a_{11},0,\cdots,0)+(0,a_{12},0,\cdots,0)+\cdots+(0,\cdots,0,a_{1n}),$$

将 Δ 按行列式性质 1 展开成 n 个行列式之和:

$$\Delta = \Delta_1+\cdots+\Delta_j+\cdots+\Delta_n,$$

其中每个 Δ_j 是用行向量 $(0,\cdots,0,a_{1j},0,\cdots,0)$ 替换 Δ 的第 1 行得到的行列式

$$\Delta_j = \begin{vmatrix} 0 & \cdots & 0 & a_{1j} & 0 & \cdots & 0 \\ a_{21} & \cdots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{nj} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = a_{1j}A_{1j},$$

于是 Δ 被分解为所有这些 $a_{1j}A_{1j}$ $(1\leqslant j\leqslant n)$ 之和:

$$\Delta = a_{11}A_{11}+\cdots+a_{1j}A_{1j}+\cdots+a_{1n}A_{1n} = \sum_{j=1}^{n} a_{1j}A_{1j},$$

这称为行列式按第 1 行展开.

行列式 Δ 也可以按第 i 行展开. 为此, 只需将第 i 行依次与前面各行互换, 将它换到第 1 行, 再将得到的 $\widetilde{\Delta}$ 按第 1 行展开:

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{i+1,1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = (-1)^{i-1} \begin{vmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

$$=(-1)^{i-1}\sum_{j=1}^{n} a_{ij}\cdot(-1)^{1+j}M_{ij} = \sum_{j=1}^{n} a_{ij}\cdot(-1)^{i+j}M_{ij} = \sum_{j=1}^{n} a_{ij}A_{ij},$$

其中 M_{ij} 是从 $\widetilde{\Delta}$ 中删去 a_{ij} 所在的第 1 行和第 j 列之后剩下的元素组成的 $n-1$ 阶行列式, 也就是从 Δ 中删去 a_{ij} 所在的第 i 行和第 j 列之后剩下的元素组成的 $n-1$ 阶行列式, 称为 a_{ij} 在 Δ 中的**余子式** (cofactor), $A_{ij}=(-1)^{i+j}M_{ij}$ 称为 a_{ij} 的 **代数余子式** (algebraic cofactor). 这样就得到:

引理 3.4.2 行列式 Δ 等于它的任意一行的各元素与各自的代数余子式的乘积之和:

$$\Delta=\sum_{j=1}^{n}a_{ij}A_{ij}=a_{i1}A_{i1}+a_{i2}A_{i2}+\cdots+a_{in}A_{in}. \qquad \square$$

引理 3.4.2 中的等式称为行列式按第 i 行的展开式.

由于行与列平等, 类似地可以得到行列式按列的展开式:

引理 3.4.3 行列式 Δ 等于它的任意一列的各元素与各自的代数余子式的乘积之和:

$$\Delta=\sum_{i=1}^{n}a_{ij}A_{ij}=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}. \qquad \square$$

引理 3.4.3 中的等式称为行列式按第 j 列的展开式.

例 4 求 4 阶行列式

$$\Delta=\begin{vmatrix}2&1&0&0\\1&2&1&0\\0&1&2&1\\0&0&1&2\end{vmatrix}.$$

解 按第 1 行展开得

$$\Delta=2\begin{vmatrix}2&1&0\\1&2&1\\0&1&2\end{vmatrix}-1\begin{vmatrix}1&1&0\\0&2&1\\0&1&2\end{vmatrix},$$

将其中第 1 个行列式按第 1 行展开, 第 2 个行列式按第 1 列展开, 得

$$\Delta=2\left(2\begin{vmatrix}2&1\\1&2\end{vmatrix}-1\begin{vmatrix}1&1\\0&2\end{vmatrix}\right)-\begin{vmatrix}2&1\\1&2\end{vmatrix}=2\times(2\times3-2)-3=5. \qquad \square$$

3.4.3 展开定理

将实系数三阶方阵 $\boldsymbol{A}=(a_{ij})_{3\times3}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$ 的行列式 $\Delta=|\boldsymbol{A}|$ 按第 1 列展开得

$$\Delta = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31},$$

展开式右边可以写成 $\boldsymbol{a}_1 = (a_{11}, a_{21}, a_{31})^{\mathrm{T}}$ 与 $\boldsymbol{A}_1 = (A_{11}, A_{21}, A_{31})^{\mathrm{T}}$ 的内积:

$$\boldsymbol{a}_1 \cdot \boldsymbol{A}_1 = \Delta = \det(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3), \tag{3.4.2}$$

其中 $\boldsymbol{a}_1$ 是行列式 Δ 的第 1 列. $\boldsymbol{A}_1$ 的各分量 A_{i1} 是 Δ 的第 1 列各元素的代数余子式, 由 Δ 的第 2,3 两列的元素计算得出, 与第 1 列的元素没有关系. 因此, 在等式 (3.4.2) 中将 $\boldsymbol{a}_1$ 替换成别的 3 维列向量 $\boldsymbol{b} = (b_1, b_2, b_3)^{\mathrm{T}}$, $\boldsymbol{A}_1$ 不变, 得到的等式

$$b_1A_{11} + b_2A_{21} + b_3A_{31} = \boldsymbol{b} \cdot \boldsymbol{A}_1 = \det(\boldsymbol{b}, \boldsymbol{a}_2, \boldsymbol{a}_3)$$

仍成立. 特别地, 取 $\boldsymbol{b} = \boldsymbol{a}_2$ 或 $\boldsymbol{a}_3$ 得

$$a_{1j}A_{11} + a_{2j}A_{21} + a_{3j}A_{31} = \boldsymbol{a}_j \cdot \boldsymbol{A}_1 = \det(\boldsymbol{a}_j, \boldsymbol{a}_2, \boldsymbol{a}_3) = 0 \quad (j = 2, 3).$$

$\boldsymbol{A}_j$ 与 $\boldsymbol{a}_2, \boldsymbol{a}_3$ 内积都为 0 (垂直), 而与 $\boldsymbol{a}_1$ 的内积 $\boldsymbol{a}_1 \cdot \boldsymbol{A}_1 = \Delta$. 容易联想到 $\boldsymbol{a}_2$ 与 $\boldsymbol{a}_3$ 的外积 $\boldsymbol{a}_2 \times \boldsymbol{a}_3$ 也有同样的性质:

$$\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \Delta, \quad \boldsymbol{a}_2 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = \boldsymbol{a}_3 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3) = 0.$$

事实上, 具有同样性质

$$\boldsymbol{a}_1 \cdot \boldsymbol{X} = \Delta, \quad \boldsymbol{a}_2 \cdot \boldsymbol{X} = \boldsymbol{a}_3 \cdot \boldsymbol{X} = 0$$

的向量 $\boldsymbol{X} = (x_1, x_2, x_3)^{\mathrm{T}}$ 是线性方程组

$$\begin{cases} a_{11}x_1 + a_{21}x_2 + a_{31}x_3 = \Delta, \\ a_{12}x_1 + a_{22}x_2 + a_{32}x_3 = 0, \\ a_{13}x_1 + a_{23}x_2 + a_{33}x_3 = 0 \end{cases}$$

的解. 这个线性方程组的系数行列式为 $\Delta^{\mathrm{T}}=\Delta$, 当 $\Delta \neq 0$ 时方程组有唯一解

$$\begin{aligned} \boldsymbol{a}_2 \times \boldsymbol{a}_3 = (A_{11}, A_{21}, A_{31})^{\mathrm{T}} &= \left(\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}, \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \right)^{\mathrm{T}} \\ &= (a_{22}a_{33} - a_{32}a_{23}, -(a_{12}a_{33} - a_{13}a_{32}), a_{12}a_{23} - a_{13}a_{22})^{\mathrm{T}}. \end{aligned} \tag{3.4.3}$$

这就是根据向量 $\boldsymbol{a}_2, \boldsymbol{a}_3$ 的坐标计算外积的公式.

注 虽然公式 (3.4.3) 是在 $\Delta \neq 0$ 的情况下得出的, 但不难看出 $\boldsymbol{a}_2 \times \boldsymbol{a}_3$ 与各 A_{1j} 都是坐标的连续函数, 让坐标连续变化, 使 Δ 由不为 0 趋于 0, 等式两边取极限可知 (3.4.3) 在 $\Delta = 0$ 时也成立.

以上推理可以推广到任意域 F (不限定实数域) 上任意 n 阶 (不限定 $n=3$) 行列式 Δ. 将 Δ 按第 j 列展开 (不限定 $j=1$), 得

$$\Delta=\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{j-1},\boldsymbol{a}_j,\boldsymbol{a}_{j+1},\cdots,\boldsymbol{a}_n)=a_{1j}A_{1j}+\cdots+a_{nj}A_{nj}. \tag{3.4.4}$$

在等式 (3.4.4) 的行列式 Δ 中将第 j 列 $\boldsymbol{a}_j$ 替换成另外一列 $\boldsymbol{a}_k$ $(k\neq j)$, 则等式右边的各 a_{ij} 变成 a_{ik}, 但 A_{ij} 都不变, 得到新的等式:

$$\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{j-1},\boldsymbol{a}_k,\boldsymbol{a}_{j+1},\cdots,\boldsymbol{a}_n)=a_{1k}A_{1j}+\cdots+a_{nk}A_{nj}, \tag{3.4.5}$$

等式左边的行列式的第 k 列与第 j 列都等于 $\boldsymbol{a}_k$. 两列相等, 行列式为 0. 等式 (3.4.5) 成为

$$a_{1k}A_{1j}+\cdots+a_{nk}A_{nj}=0. \tag{3.4.6}$$

类似地, Δ 的任意一行的代数余子式 $A_{i1},\cdots,A_{in}$ 与另外一行的元素 $a_{i1},\cdots,a_{in}$ (其中 $i\neq k$) 的对应乘积之和也满足

$$a_{k1}A_{i1}+\cdots+a_{kn}A_{in}=0.$$

由此得到:

定理 3.4.1(展开定理) 设 A_{ij} 是 a_{ij} 在 $\Delta=\det\boldsymbol{A}$ 中的代数余子式, 则:

$$\sum_{i=1}^{n}a_{ik}A_{ij}=a_{1k}A_{1j}+\cdots+a_{nk}A_{nj}=\begin{cases}\Delta, & 当\ k=j,\\ 0, & 当\ k\neq j;\end{cases}$$

$$\sum_{j=1}^{n}a_{kj}A_{ij}=a_{k1}A_{i1}+\cdots+a_{kn}A_{in}=\begin{cases}\Delta, & 当\ k=i,\\ 0, & 当\ k\neq i.\end{cases}$$ □

习题 3.4

1. 计算行列式

(1) $\begin{vmatrix} a_{11} & a_{12} & 0 & 0 & 0\\ a_{21} & a_{22} & 0 & 0 & 0\\ 0 & 0 & b_{11} & 0 & 0\\ 0 & 0 & b_{21} & b_{22} & 0\\ 0 & 0 & b_{31} & b_{32} & b_{33}\end{vmatrix}$; (2) $\begin{vmatrix} 2 & 1 & 0 & \cdots & 0\\ 1 & 2 & 1 & \ddots & \vdots\\ 0 & 1 & 2 & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & 1\\ 0 & \cdots & 0 & 1 & 2\end{vmatrix}_{n\times n}$.

2. 求行列式

$$\Delta=\begin{vmatrix} 1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\\ 3 & 4 & 1 & 2\\ 4 & 1 & 2 & 3\end{vmatrix}$$

的值. 并求 Δ 的第 1 列各元素的代数余子式之和 $A_{11}+A_{21}+A_{31}+A_{41}$.

3. 求如下行列式 Δ 展开式中的二次项系数:

$$\Delta=\begin{vmatrix}1&1&1&1\\1&2&3&x\\1&4&9&x^2\\1&8&27&x^3\end{vmatrix}.$$

4. 求行列式

$$\Delta=\begin{vmatrix}1&1&1\\x_1&x_2&x_3\\x_1^3&x_2^3&x_3^3\end{vmatrix}.$$

(提示: 参照第 3 题.)

§3.5* 更多的例子

3.5.1 方程组唯一解条件应用例

例 1(勾股数) 求方程 $a^2+b^2=c^2$ 的全体正整数解 (a,b,c).

说明 根据勾股定理, 直角三角形的三边长 a,b,c 满足方程 $a^2+b^2=c^2$, 求这个方程的正整数解 (a,b,c) 也就是求边长为整数的直角三角形. 每一组正整数解 (a,b,c) 称为一组**勾股数**, 国外称为毕达哥拉斯数 (Pythagorean number). 勾股数 $(3,4,5)$ 就是大家都熟悉的"勾三股四弦五". 容易验证, 对任意正整数 $n<m$, $(2mn,m^2-n^2,m^2+n^2)$ 是勾股数, 它的正整数倍 $(2kmn,k(m^2-n^2),k(m^2+n^2))$ 也是勾股数. 我们证明, 这就是全部勾股数.

解 勾股数 (a,b,c) 满足的条件 $a^2+b^2=c^2$ 可以写成

$$0=c^2-b^2-a^2=(c+b)(c-b)-a^2=\begin{vmatrix}c+b&-a\\-a&c-b\end{vmatrix}.$$

齐次线性方程组

$$\begin{cases}(c+b)x-ay=0,\\-ax+(c-b)y=0\end{cases}\tag{3.5.1}$$

的系数行列式等于 0, 因此有非零解 (x,y). 易见 x,y 中只要有一个不为 0, 则另一个也不为 0, 满足条件

$$\frac{c+b}{a}=\frac{a}{c-b}=\frac{y}{x}=\frac{m}{n},$$

其中 $\dfrac{m}{n}$ 是由大于 1 的正有理数 $\dfrac{y}{x}=\dfrac{c+b}{a}$ 约分得到的最简分数, 分子分母 m,n 是互素的正整数. 于是

$$\begin{cases}c+b=\dfrac{m}{n}a,\\ c-b=\dfrac{n}{m}a\end{cases}\quad\Rightarrow\quad (c,b)=\left(\frac{m^2+n^2}{2mn}a,\frac{m^2-n^2}{2mn}a\right).$$

由于 m,n 互素, m,n 的每个素因子都不整除 $m^2\pm n^2$, 因此 mn 与 $m^2\pm n^2$ 互素. 由 $\dfrac{(m^2+n^2)a}{2mn}=c_1$ 是整数知 mn 整除 $(m^2+n^2)a$ 从而整除 a, $k_1=\dfrac{a}{mn}$ 是正整数. 我们有

$$a=k_1mn,\quad b=\frac{k_1(m^2-n^2)}{2},\quad c=\frac{k_1(m^2+n^2)}{2},$$

其中 k_1 是正整数.

如果 m,n 的奇偶性相反, 则 $m^2\pm n^2$ 都是奇数, 由 b,c 是整数知 k_1 必须是偶数, 设 $k_1=2k$, 则

$$a=2kmn,\quad b=k(m^2-n^2),\quad c=k(m^2+n^2).$$

若 m,n 的奇偶性相同, 则 $m\pm n$ 是偶数, $m_1=\dfrac{m+n}{2}$ 与 $n_1=\dfrac{m-n}{2}$ 都是正整数. 由 $m=m_1+n_1$ 及 $n=m_1-n_1$ 知 m_1,n_1 的公因子必是 m 与 n 的公因子, 只能为 1. 可见 m_1,n_1 互素. 如果 m_1,n_1 的奇偶性相同, 将导致 m 与 n 都是偶数, 不可能互素. 因此 m_1,n_1 的奇偶性必然相反. 我们有

$$\begin{aligned}a&=k_1(m_1+n_1)(m_1-n_1)=k_1(m_1^2-n_1^2),\\ b&=2k_1\left(\frac{m+n}{2}\right)\left(\frac{m-n}{2}\right)=2k_1m_1n_1,\\ c&=k_1\frac{(m_1+n_1)^2+(m_1-n_1)^2}{2}=k_1(m_1^2+n_1^2).\end{aligned}$$

这证明了方程 $a^2+b^2=c^2$ 的全体正整数解 (a,b,c) 为

$$(2kmn,k(m^2-n^2),k(m^2+n^2))\ \text{ 或 }\ (k(m^2-n^2),2kmn,k(m^2+n^2)),$$

其中 $m>n$ 是互素的正整数且奇偶性相反, k 是任意正整数. □

3.5.2 展开定理应用例

例 2 n 阶行列式

$$\Delta=\begin{vmatrix}\lambda_1 & 1 & \cdots & 1\\ 1 & \lambda_2 & \cdots & 1\\ \vdots & \vdots & & \vdots\\ 1 & 1 & \cdots & \lambda_n\end{vmatrix}$$

的对角元依次为 $\lambda_1,\cdots,\lambda_n$, 非对角元全为 1. 求 Δ 的值.

解 如果某个对角元 $\lambda_i=1$, 将第 i 行的 -1 倍加到其余各行, 将其余各行的非对角元全部变成 0, 对角元分别变成 $\lambda_1-1,\cdots,\lambda_{i-1}-1,\lambda_{i+1}-1,\cdots,\lambda_n-1$, 行列式化为

$$\Delta=\begin{vmatrix}\lambda_1-1 & & & & \\ & \ddots & & & \\ 1 & \cdots & \lambda_i & \cdots & 1\\ & & & \ddots & \\ & & & & \lambda_n-1\end{vmatrix}$$

$$=\begin{vmatrix}\lambda_1-1 & & \\ & \ddots & \\ & & \lambda_{i-1}-1\end{vmatrix}\cdot\begin{vmatrix}1 & \cdots & \cdots & 1\\ & \lambda_{i+1}-1 & & \\ & & \ddots & \\ & & & \lambda_n-1\end{vmatrix}$$

$$=(\lambda_1-1)\cdots(\lambda_{i-1}-1)(\lambda_{i+1}-1)\cdots(\lambda_n-1).$$

设所有的 λ_i 都不等于 1. 在 Δ 左边添加 n 个 0 组成的列, 顶上添上 $n+1$ 个 1 组成的行, 将 Δ 变成与 Δ 的值相等的 $n+1$ 阶行列式, 再将新的第一行的 -1 倍加到其余各行:

$$\Delta=\begin{vmatrix}1 & 1 & 1 & \cdots & 1\\ 0 & \lambda_1 & 1 & \cdots & 1\\ 0 & 1 & \lambda_2 & \cdots & 1\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 1 & 1 & \cdots & \lambda_n\end{vmatrix}\xlongequal{-(1)+(i),\forall i\geqslant 2}\begin{vmatrix}1 & 1 & \cdots & 1\\ -1 & \lambda_1-1 & & \\ \vdots & & \ddots & \\ -1 & & & \lambda_n-1\end{vmatrix},$$

再依次将第 2 至 $n+1$ 列的 $\dfrac{1}{\lambda_i-1}$ 倍 $(i=1,\cdots,n)$ 加到第一列, 将第一列第 2 至 $n+1$ 行元素全部变成 0, 从而将 Δ 变成上三角行列式:

$$\Delta=\begin{vmatrix}1+\sum\limits_{i=1}^{n}\dfrac{1}{\lambda_i-1} & 1 & \cdots & 1\\ & \lambda_1-1 & & \\ & & \ddots & \\ & & & \lambda_n-1\end{vmatrix}$$

$$=\left(1+\sum_{i=1}^{n}\frac{1}{\lambda_i-1}\right)(\lambda_1-1)\cdots(\lambda_n-1),$$

最后得到

$$\Delta=(\lambda_1-1)\cdots(\lambda_n-1)+\sum_{i=1}^{n}\left(\prod_{1\leqslant j\leqslant n,j\neq i}(\lambda_j-1)\right). \tag{3.5.2}$$

最后的答案 (3.5.2) 当某个 $\lambda_i=1$ 时也正确, 因此适合于任意的 $\lambda_1,\cdots,\lambda_n$. □

例 3(广义范德蒙德 (Vandermonde) 行列式) 求 n 阶行列式

$$\Delta=\begin{vmatrix}1 & x_1 & \cdots & x_1^{k-1} & x_1^{k+1} & \cdots & x_1^n\\ 1 & x_2 & \cdots & x_2^{k-1} & x_2^{k+1} & \cdots & x_2^n\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ 1 & x_n & \cdots & x_n^{k-1} & x_n^{k+1} & \cdots & x_n^n\end{vmatrix}.$$

解 考虑由 $n+1$ 个字母组成的范德蒙德 (Vandermonde) 行列式

$$\begin{aligned}V_{n+1}=V(x_1,\cdots,x_n,x)&=\begin{vmatrix}1 & x_1 & \cdots & x_1^k & \cdots & x_1^n\\ \vdots & \vdots & & \vdots & & \vdots\\ 1 & x_n & \cdots & x_n^k & \cdots & x_n^n\\ 1 & x & \cdots & x^k & \cdots & x^n\end{vmatrix}\\ &=(x-x_1)\cdots(x-x_n)\cdot V(x_1,\cdots,x_n).\end{aligned} \tag{3.5.3}$$

将 $x_1,\cdots,x_n$ 看成常数, x 看成自变量, 将行列式 V_{n+1} 按第 $n+1$ 行展开成 x 的多项式 $f(x)$, 则等式 (3.5.3) 成为

$$\begin{aligned}f(x)&=A_{n+1,1}+A_{n+1,2}x+\cdots+A_{n+1,k+1}x^k+\cdots+A_{n+1,n+1}x^n\\ &=(x-x_1)\cdots(x-x_n)\cdot V(x_1,\cdots,x_n),\end{aligned} \tag{3.5.4}$$

$f(x)$ 中 x^k 的系数 $A_{n+1,k+1}=(-1)^{n+k}\Delta$ 是 V_{n+1} 第 $n+1$ 行第 $k+1$ 列元素 x^k 的代数余子式. Δ 是 x^k 的余子式, 就是本题中需要计算的行列式. 另一方面, 等式 (3.5.4) 右边的 $(x-x_1)\cdots(x-x_n)V(x_1,\cdots,x_n)$ 展开式中的 x^k 项系数等于 $(-1)^{n-k}\sigma_{n-k}(x_1,\cdots,x_n)V(x_1,\cdots,x_n)$, 其中

$$\sigma_{n-k}(x_1,\cdots,x_n)=\sum_{1\leqslant i_1<\cdots<i_{n-k}\leqslant n}x_{i_1}\cdots x_{i_{n-k}}$$

是 n 个数 $x_1,\cdots,x_n$ 中每次取 $n-k$ 个相乘得到的所有乘积之和. 比较等式 (3.5.4) 两边 x^k 项系数就得到

$$\begin{aligned}\Delta&=\sigma_{n-k}(x_1,\cdots,x_n)V(x_1,\cdots,x_n)\\ &=\left(\sum_{1\leqslant i_1<\cdots<i_{n-k}\leqslant n}x_{i_1}\cdots x_{i_{n-k}}\right)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).\end{aligned}$$

□

例如, 在例 3 中取 $k=n-1$ 就得到

$$\begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-2} & x_1^n \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-2} & x_n^n \end{vmatrix} = (x_1+\cdots+x_n)\prod_{1\leqslant i<j\leqslant n}(x_j-x_i).$$

例 4 由 5 个电阻 R_1,R_2,R_3,R_4,R_5 组成如图 3–8 的电路. 设在两点 A,B 之间加电压 V, 根据系数行列式分析 5 个电阻上流过的电流 I_1,I_2,I_3,I_4,I_5 所满足的线性方程组的解的唯一性.

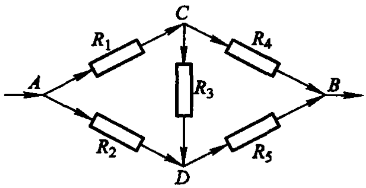

图 3–8

解 按图中所标的箭头方向分别流过 5 个电阻的电流 I_1,I_2,I_3,I_4,I_5 满足方程组:

$$\begin{cases} I_1 - I_3 - I_4 = 0, \\ I_2 + I_3 - I_5 = 0, \\ I_1R_1 - I_2R_2 + I_3R_3 = 0, \\ I_3R_3 - I_4R_4 + I_5R_5 = 0, \\ I_1R_1 + I_4R_4 = V. \end{cases} \tag{3.5.5}$$

方程组的系数行列式

$$\Delta = \begin{vmatrix} 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ R_1 & -R_2 & R_3 & 0 & 0 \\ 0 & 0 & R_3 & -R_4 & R_5 \\ R_1 & 0 & 0 & R_4 & 0 \end{vmatrix} \xlongequal{-R_1(1)+(3),-R_1(1)+(5)}$$

$$\begin{vmatrix} 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & -R_2 & R_1+R_3 & R_1 & 0 \\ 0 & 0 & R_3 & -R_4 & R_5 \\ 0 & 0 & R_1 & R_1+R_4 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 & -1 \\ -R_2 & R_1+R_3 & R_1 & 0 \\ 0 & R_3 & -R_4 & R_5 \\ 0 & R_1 & R_1+R_4 & 0 \end{vmatrix}$$

$$\xlongequal{R_2(1)+(2)} \begin{vmatrix} 1 & 1 & 0 & -1 \\ 0 & R_1+R_2+R_3 & R_1 & -R_2 \\ 0 & R_3 & -R_4 & R_5 \\ 0 & R_1 & R_1+R_4 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} R_1+R_2+R_3 & R_1 & -R_2 \\ R_3 & -R_4 & R_5 \\ R_1 & R_1+R_4 & 0 \end{vmatrix}$$

$$\xlongequal{\text{按第 3 行展开}} R_1(R_1R_5 - R_2R_4) - (R_1+R_4)[(R_1+R_2+R_3)R_5 + R_2R_3]$$

$$= -R_1R_2(R_4+R_5) - R_4R_5(R_1+R_2) - R_3(R_1+R_4)(R_2+R_5).$$

当然应当假定所有的电阻 $R_i \geqslant 0 (1 \leqslant i \leqslant 5)$. 除了如下情况 (1) – (6) 之外, 其余的情况下都有 $\Delta < 0$, 方程组 (3.5.5) 有唯一解.

(1) $R_1 = R_4 = 0$;

(2) $R_2 = R_5 = 0$;

(3) $R_1 = R_3 = R_5 = 0$;

(4) $R_2 = R_3 = R_4 = 0$;

(5) $R_1 = R_2 = R_3 = 0$ 且 $R_4 > 0$ 且 $R_5 > 0$;

(6) $R_1 > 0$ 且 $R_2 > 0$ 且 $R_3 = R_4 = R_5 = 0$.

在以上 6 种使 $\Delta = 0$ 的情况中, 在情况 (1) – (4) 下都存在从 A 到 B 的总电阻为 0 的通路, 电路“短路”, 沿着这条电路的总电压只能为零, 方程组无解. 理论上, 当 $V > 0$ 时, 短路电路的电流只能是无穷大. 实际情况下, 短路电阻不会真正等于零, 只能是很小, 短路电路上的电流虽然不是无穷大, 但也会太大, 烧毁电路.

情况 (5) 相当于 R_4, R_5 并联后直接接在 A, B 之间, 经过 R_4, R_5 的电流 I_4, I_5 是唯一决定的; 但经过 R_3 的电流 I_3 可以取任意实数值, 再由 $I_1 = I_3 + I_4$ 及 $I_2 = I_5 - I_3$ 算出 I_1, I_2, 得到的都是方程组的解, 可见此时方程组有无穷多解. 类似地, 情况 (6) 相当于将 R_1, R_2 并联后直接接在 A, B 之间, I_1, I_2 由方程组唯一决定, I_3 仍可任意取值. 这说明, 在情况 (5),(6) 下方程组都有无穷多

解. 在实际情况下, 电阻都不会真正为 0, 说它为 0 只是说电阻很小, 可以忽略不计. 但在情况 (5),(6) 下, 无论题中这些约等于 0 的电阻多么小, 它们都能唯一决定所通过的电流的值, 而且这些值在正常范围内, 不会烧毁电路. □

3.5.3 行列式的归纳定义

利用展开定理, 可以将 n 阶行列式转化为 $n-1$ 阶行列式的线性组合来计算.

也可以一开始就直接利用这样的计算公式来定义行列式: 先将 1 阶方阵 (a) 的行列式定义为数 a. 再由 1 阶行列式定义 2 阶行列式, 2 阶行列式定义 3 阶行列式, ……. 一般地, 由 $n-1$ 阶行列式定义 n 阶行列式.

定义 3.5.1(行列式的归纳定义) 定义 1 阶方阵 (a) 的行列式等于数 a.

设已经对正整数 k 定义了 k 阶行列式

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix},$$

则定义 $k+1$ 阶行列式

$$\begin{aligned} \Delta_{k+1} &= \begin{vmatrix} a_{11} & \cdots & a_{1,k+1} \\ \vdots & & \vdots \\ a_{k+1,1} & \cdots & a_{k+1,k+1} \end{vmatrix} \\ &= a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{1+j}a_{1j}M_{1j} + \cdots + (-1)^{1+(k+1)}a_{1,k+1}M_{1,k+1}, \end{aligned}$$

其中 M_{1j} 是在 Δ_{k+1} 中将第 1 行及第 j 列删去之后剩下的元素组成的 k 阶行列式. □

本书直接由行列式计算公式得出行列式的各种性质, 包括按行和按列的展开式. 有些教材采用上述归纳定义来定义行列式, 由这个定义得出行列式的各种性质和计算公式.

3.5.4 利用展开定理求线性方程组唯一解

n 个方程组成的 n 元线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 当系数行列式 $\Delta=|\boldsymbol{A}|\neq 0$ 时有唯一解. 将方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 写成向量形式

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=\boldsymbol{b}. \tag{3.5.6}$$

则当 $n=3$ 且方程组各系数都是实数时, 可以将等式 (3.5.6) 两边与 $\boldsymbol{a}_2\times\boldsymbol{a}_3$ 作内积消去 x_2,x_3, 得到

$$x_1\boldsymbol{a}_1\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3)=\boldsymbol{b}\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3),\quad \text{即}\quad x_1\Delta=\Delta_1,\quad \text{从而}\quad x_1=\frac{\Delta_1}{\Delta}.$$

类似地, 将 (3.5.6) 两边与 $\boldsymbol{a}_1\times\boldsymbol{a}_3,\boldsymbol{a}_1\times\boldsymbol{a}_2$ 作内积可解出 x_2,x_3.

在任意域 F 上任意维空间 F^n 中, 我们还没有定义内积与外积. 但可以仿照 $\mathbf{R}^3$ 中的内积定义 "点乘"

$$(x_1,\cdots,x_n)\cdot(y_1,\cdots,y_n)=x_1y_1+\cdots+x_ny_n.$$

并将 Δ 的第 1 列元素的代数余子式组成的向量 $\boldsymbol{A}_1=(A_{11},\cdots,A_{n1})^{\mathrm{T}}$ 看成 $\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 的 "外积", 与等式 (3.5.6) 两边作 "点乘", 消去 $x_2,\cdots,x_n$, 得到

$$x_1\Delta=\Delta_1,\text{从而}x_1=\frac{\Delta_1}{\Delta.}$$

类似地可以对 $j=2,\cdots,n$ 得到 $x_j=\dfrac{\Delta_j}{\Delta}$.

习题 3.5

1. 求 n 阶行列式

$$(1)\begin{vmatrix}1&-1&\cdots&-1\\-1&2&\cdots&-1\\\vdots&\vdots&&\vdots\\-1&-1&\cdots&n\end{vmatrix};\quad(2)\begin{vmatrix}\lambda_1+a_1b_1&a_1b_2&\cdots&a_1b_n\\a_2b_1&\lambda_2+a_2b_2&\cdots&a_2b_n\\\vdots&\vdots&&\vdots\\a_nb_1&a_nb_2&\cdots&\lambda_n+a_nb_n\end{vmatrix}.$$

2. 计算偶数阶方阵的行列式

$$\Delta_n=\begin{vmatrix}0&1&\cdots&1\\-1&0&\cdots&1\\\vdots&\vdots&&\vdots\\-1&-1&\cdots&0\end{vmatrix},$$

它的对角元全为 0, 主对角线右上方的元素全为 1, 左下方的元素全为 -1.

3. $\boldsymbol{A}$ 是偶数阶反对称方阵, 满足条件 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$. 求证: 将 $\Delta=\det\boldsymbol{A}$ 的所有的元素加上同一个数 λ, 行列式值不变.

4. 设 n 元线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\\cdots\cdots\cdots\cdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n\end{cases}$$

的系数矩阵 $\boldsymbol{A}$ 的行列式 $\Delta \neq 0$. 对每个 $1 \leqslant j \leqslant n$, 令 $\Delta_j = b_1A_{1j} + \cdots + b_nA_{nj}$ 是在 Δ 中将第 j 列元素分别换成 $b_1, b_2, \cdots, b_n$ 得到的行列式. 将

$$(x_1, \cdots, x_n) = \left(\frac{\Delta_1}{\Delta}, \cdots, \frac{\Delta_j}{\Delta}, \cdots, \frac{\Delta_n}{\Delta}\right)$$

代入原方程组检验, 证明它确实是原方程组的解.

第 4 章　矩阵的代数运算

§4.1　矩阵运算的定义与运算律

4.1.1　矩阵运算的定义

1. 矩阵的线性运算——加法与数乘

定义 4.1.1(矩阵加法)　$F^{m\times n}$ 中任意两个矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 与 $\boldsymbol{B}=(b_{ij})_{m\times n}$ 可以相加, 得到的和 $\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}$ 仍是 $F^{m\times n}$ 的矩阵, 它的第 (i,j) 分量等于 $\boldsymbol{A},\boldsymbol{B}$ 的第 (i,j) 分量 a_{ij},b_{ij} 之和.

定义 4.1.2(矩阵与数的乘法)　$F^{m\times n}$ 中任意矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 可以与 F 中任意数 λ 相乘, 得到的乘积 $\lambda\boldsymbol{A}=(\lambda a_{ij})_{m\times n}$ 仍是 $m\times n$ 矩阵, 称为 $\boldsymbol{A}$ 的 λ 倍, 它的第 (i,j) 分量等于 $\boldsymbol{A}$ 的第 (i,j) 分量 a_{ij} 的 λ 倍.

矩阵与数的乘法也称为矩阵的数乘.

$F^{m\times n}$ 中矩阵的加法与数乘, 与 F^{mn} 中数组向量的加法与数乘, 本质上是相同的. 因此, $F^{m\times n}$ 中的加法与数乘同样满足 8 条运算律. 对 $F^{m\times n}$ 任意两个矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{B}=(b_{ij})_{m\times n}$ 定义减法, 将对应位置元素相减得到差 $\boldsymbol{A}-\boldsymbol{B}=(a_{ij}-b_{ij})_{m\times n}$, 满足条件 $(\boldsymbol{A}-\boldsymbol{B})+\boldsymbol{B}=\boldsymbol{A}$. $F^{m\times n}$ 在这样的加法与数乘下成为 mn 维向量空间, $E=\{\boldsymbol{E}_{ij}\mid 1\leqslant i\leqslant m,1\leqslant j\leqslant n\}$ 是这个空间的自然基, 每个 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 在这组基下的坐标是列向量 $(a_{11},\cdots,a_{1n},a_{21},\cdots,a_{2n},\cdots,a_{m1},\cdots,a_{mn})^{\mathrm{T}}$ (也可写成行向量).

值得提醒的是: 同一数域 F 上两个矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 与 $\boldsymbol{B}=(b_{ij})_{k\times t}$, 只有当它们的行数相等 $(m=k)$ 且列数相等 $(n=t)$ 时才能相加, 如果 $m\neq k$ 或 $n\neq t$, 即使 $mn=kt$ 也不能将 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相加.

2. 矩阵的乘法

如果矩阵只有加法与数乘运算, 那就与向量完全相同, 没有专门研究的必要. 矩阵最精彩最有用的运算是乘法.

第 2 章在讨论向量组的线性组合的时候已经引入并利用了矩阵的乘法:

同一数域 F 上的 $m\times n$ 矩阵 $\boldsymbol{A}$ 与 $n\times 1$ 矩阵 (n 维列向量) $\boldsymbol{X}$ 可以相乘,

乘积

$$\boldsymbol{A}\boldsymbol{X} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{a}_1 x_1 + \cdots + \boldsymbol{a}_n x_n$$

等于 $\boldsymbol{A}$ 的各列与 $\boldsymbol{X}$ 的各元素对应乘积之和.

同一数域 F 上的 $m \times n$ 矩阵 $\boldsymbol{A}$ 与 $n \times p$ 矩阵 $\boldsymbol{X}$ 可以相乘, 乘积

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_p) = (\boldsymbol{A}\boldsymbol{X}_1, \cdots, \boldsymbol{A}\boldsymbol{X}_p)$$

的各列由 $\boldsymbol{A}$ 分别乘 $\boldsymbol{X}$ 的各列 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_p$ 得到, $\boldsymbol{A}$ 与每列 $\boldsymbol{X}_j$ 的乘积 $\boldsymbol{A}\boldsymbol{X}_j$ 按前述法则得到.

这样得到的矩阵乘积 $\boldsymbol{A}\boldsymbol{X}$ 的第 (i,j) 元素就是第 j 列 $\boldsymbol{A}\boldsymbol{X}_j$ 的第 i 分量, 它等于 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{X}$ 的第 j 列对应元素乘积之和.

由此得到矩阵乘法的定义如下:

定义 4.1.3(矩阵乘法)

数域 F 上任意行向量 $\boldsymbol{\alpha} = (a_1, \cdots, a_n)$ 与同维数的列向量 $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 可以相乘, 乘积

$$\boldsymbol{\alpha}\boldsymbol{b} = (a_1, \cdots, a_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n$$

是一个数, 等于 $\boldsymbol{\alpha}$ 与 $\boldsymbol{b}$ 对应分量的乘积之和.

数域 F 上任意 $m \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$ 与 $n \times p$ 矩阵 $\boldsymbol{B} = (b_{ij})_{n\times p}$ 可以相乘, 乘积 $\boldsymbol{C} = (c_{ij})_{m\times p}$ 是 $m \times p$ 矩阵, 它的第 (i,j) 元等于 $\boldsymbol{A}$ 的第 i 行 $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{B}$ 的第 j 列 $\boldsymbol{b}_j$ 之积:

$$c_{ij} = \boldsymbol{\alpha}_i \boldsymbol{b}_j = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + \cdots + a_{in} b_{nj}\,. \qquad \square$$

可乘条件: 同一个数域 F 上的矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可以相乘的充分必要条件是: $\boldsymbol{A}$ 的列数与 $\boldsymbol{B}$ 的行数相等. 也就是说: $\boldsymbol{A}$ 的每一行 $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{B}$ 的每一列 $\boldsymbol{b}_j$ 的维数相同, 这才能够将 $\boldsymbol{\alpha}_i$ 与 $\boldsymbol{b}_j$ 的对应元素相乘再相加得到 $\boldsymbol{A}\boldsymbol{B}$ 的第 (i,j) 元素.

例 1 建立了直角坐标系的三维几何空间中任何两个向量 $\boldsymbol{a}_1, \boldsymbol{a}_2$ 的坐标如果分别用行向量表示为 $\boldsymbol{X}_1 = (x_1, y_1, z_1), \boldsymbol{X}_2 = (x_2, y_2, z_2)$. 试将它们的内积 $\boldsymbol{a}_1 \cdot \boldsymbol{a}_2$ 用矩阵乘法表示.

解

$$\boldsymbol{a}_1\cdot\boldsymbol{a}_2=x_1x_2+y_1y_2+z_1z_2=(x_1,y_1,z_1)\begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix}=\boldsymbol{X}_1\boldsymbol{X}_2^{\mathrm{T}}.$$ □

如果 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 的坐标分别写成列向量 $\boldsymbol{X}_1,\boldsymbol{X}_2$, 则易见 $\boldsymbol{a}_1\cdot\boldsymbol{a}_2=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2$.

例 2 设 $\boldsymbol{A}\in F^{m\times n}$, $\boldsymbol{B}\in F^{p\times q}$. 试给出发生下列情况的充分必要条件, 并举例.

(1) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可以相乘, 但 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 不能相乘.

(2) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可以相乘, $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 也可以相乘, 但乘积 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 不能相加.

解 (1) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可乘 $\Leftrightarrow n=p$. $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 不能相乘 $\Leftrightarrow m\neq q$.

因此, 所求条件为 $n=p$ 且 $m\neq q$. 例如

$$\boldsymbol{A}=\begin{pmatrix}a_1\\a_2\end{pmatrix},\boldsymbol{B}=(b_1,b_2,b_3),$$

$\boldsymbol{A},\boldsymbol{B}$ 可相乘得

$$\boldsymbol{AB}=\begin{pmatrix}a_1\\a_2\end{pmatrix}(b_1,b_2,b_3)=\begin{pmatrix}a_1b_1&a_1b_2&a_1b_3\\a_2b_1&a_2b_2&a_2b_3\end{pmatrix},$$

但 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 不能相乘.

(2) 所求条件为 $n=p\neq m=q$. 例如

$$\boldsymbol{A}=\begin{pmatrix}a_1\\a_2\end{pmatrix},\boldsymbol{B}=(b_1,b_2),\boldsymbol{AB}=\begin{pmatrix}a_1b_1&a_1b_2\\a_2b_1&a_2b_2\end{pmatrix},\boldsymbol{BA}=(b_1a_1+b_2a_2).$$

但 2 阶方阵 $\boldsymbol{AB}$ 与 1 阶方阵 $\boldsymbol{BA}$ 不能相加. □

注 任意 m 维列向量 $\boldsymbol{A}$ 与 n 维行向量 $\boldsymbol{B}$ 都可以相乘得到 $m\times n$ 矩阵 $\boldsymbol{AB}$, 这是因为 $\boldsymbol{A}$ 的列数与 $\boldsymbol{B}$ 的行数都等于 1. 反过来, 当 $m\neq n$ 时, $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 不能相乘.

数 λ 可以与任意矩阵 $\boldsymbol{A}$ 相乘, 乘积 $\lambda\boldsymbol{A}$ 也可以写成 $\boldsymbol{A}\lambda$. 但是, 如果将 λ 看成一阶方阵, 仅当 $\boldsymbol{A}$ 是行向量时 $\lambda\boldsymbol{A}$ 才是矩阵乘法, 仅当 $\boldsymbol{A}$ 是列向量时 $\boldsymbol{A}\lambda$ 才是矩阵乘法. $m\times n$ 矩阵 $\boldsymbol{A}$ 与 n 维列向量 $\boldsymbol{X}$ 的乘积 $\boldsymbol{AX}=\boldsymbol{a}_1x_1+\cdots+\boldsymbol{a}_nx_n$ 中的每个 $\boldsymbol{a}_jx_j$ 不仅是列向量 $\boldsymbol{a}_j$ 与数 x_j 的乘积, 也是列向量 $\boldsymbol{a}_j$ 与 1 阶方阵 x_j 按矩阵乘法的乘积.

例 3 设 $\boldsymbol{A}=\begin{pmatrix}1&1\\1&1\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}1&1\\-1&-1\end{pmatrix}$. 求 $\boldsymbol{AB},\boldsymbol{BA}$.

解
$$AB=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad BA=\begin{pmatrix}2&2\\-2&-2\end{pmatrix}.$$
□

矩阵乘法不满足交换律: 矩阵加法满足交换律. 但从例 2 与例 3 中看到, 矩阵乘法不满足交换律. $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可乘, $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 不一定可乘. 即使两者都可乘, 乘积 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 也有可能不相等. 这与数的乘法是不同的.

零矩阵: 所有的元素都为 0 的矩阵称为**零矩阵** (zero matrix), 记为 $\boldsymbol{O}$, 必要时写成 $\boldsymbol{O}_{(m\times n)}$ 来指明是 m 行 n 列的零矩阵.

$m\times n$ 零矩阵 $\boldsymbol{O}$ 就是向量空间 $F^{m\times n}$ 中的零向量. $\boldsymbol{O}_{(m\times n)}$ 与任何一个 $m\times n$ 矩阵 $\boldsymbol{A}$ 相加, 得到的和仍为 $\boldsymbol{A}$: $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$.

零矩阵如果能与 $m\times n$ 矩阵 $\boldsymbol{A}$ 相乘, 乘积仍为零矩阵: $\boldsymbol{O}_{(p\times m)}\boldsymbol{A}=\boldsymbol{O}_{(p\times n)}$, $\boldsymbol{AO}_{(n\times q)}=\boldsymbol{O}_{(m\times q)}$.

可见, 零矩阵在矩阵加法与乘法中的作用, 相当于数 0 在数的加法与乘法中的作用.

矩阵乘法不满足消去律：数的运算中, 如果 $a\neq 0$, 且 $ab=ac$, 则可从等式两边消去 a 得到 $b=c$. 这称为**消去律** (cancellation law). 特别地, 两个非零数 a,b 的乘积 ab 一定不为 0.

例 3 中的两个矩阵都不为零, 它们的乘积却为零, 这样的事情在数的乘法中是不可能发生的. 可见, 消去律对矩阵乘法不成立, 不能由 $\boldsymbol{AB}=\boldsymbol{AC}$ 及 $\boldsymbol{A}\neq\boldsymbol{O}$ 断定 $\boldsymbol{B}=\boldsymbol{C}$.

4.1.2　矩阵的分块运算

如果将 $m\times n$ 矩阵 $\boldsymbol{A}$ 看成由它的各行 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 从上到下排成的“列向量”
$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\vdots\\\boldsymbol{\alpha}_m\end{pmatrix},$$
$n\times p$ 矩阵 $\boldsymbol{B}$ 看成由它的各列从左到右排成的“行向量” $\boldsymbol{B}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p)$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘法法则可写为:
$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{\alpha}_1\\\vdots\\\boldsymbol{\alpha}_m\end{pmatrix}(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p)=\begin{pmatrix}\boldsymbol{\alpha}_1\boldsymbol{b}_1&\cdots&\boldsymbol{\alpha}_1\boldsymbol{b}_p\\\vdots&&\vdots\\\boldsymbol{\alpha}_m\boldsymbol{b}_1&\cdots&\boldsymbol{\alpha}_m\boldsymbol{b}_p\end{pmatrix}.\tag{4.1.1}$$

特别地, 按照这个法则, $\boldsymbol{A}$ 与 $n\times 1$ 矩阵 $\boldsymbol{b}_j$ ($\boldsymbol{B}$ 的第 j 列) 的乘积

$$\boldsymbol{A}\boldsymbol{b}_j=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}\boldsymbol{b}_j=\begin{pmatrix}\boldsymbol{\alpha}_1\boldsymbol{b}_j\\ \vdots\\ \boldsymbol{\alpha}_m\boldsymbol{b}_j\end{pmatrix}, \tag{4.1.2}$$

$\boldsymbol{A}$ 与每个 $\boldsymbol{b}_j$ $(1\leqslant j\leqslant p)$ 的乘积 $\boldsymbol{A}\boldsymbol{b}_j$ 就是 $\boldsymbol{A}\boldsymbol{B}$ 的第 j 列. 因此

$$\boldsymbol{A}\boldsymbol{B}=\boldsymbol{A}(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p)=(\boldsymbol{A}\boldsymbol{b}_1,\cdots,\boldsymbol{A}\boldsymbol{b}_p). \tag{4.1.3}$$

我们还有

$$\boldsymbol{A}\boldsymbol{b}_j=\begin{pmatrix}a_{11}b_{1j}+\cdots+a_{1n}b_{nj}\\ \vdots\\ a_{m1}b_{1j}+\cdots+a_{mn}b_{nj}\end{pmatrix}=\begin{pmatrix}a_{11}\\ \vdots\\ a_{m1}\end{pmatrix}b_{1j}+\cdots+\begin{pmatrix}a_{1n}\\ \vdots\\ a_{mn}\end{pmatrix}b_{nj},$$

即

$$(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)\begin{pmatrix}b_{1j}\\ \vdots\\ b_{nj}\end{pmatrix}=\boldsymbol{a}_1b_{1j}+\cdots+\boldsymbol{a}_nb_{nj}. \tag{4.1.4}$$

这就是我们为表示向量组 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n\}$ 的线性组合 $\boldsymbol{a}_1b_{1j}+\cdots+\boldsymbol{a}_nb_{nj}$ 规定的乘法法则.

如果暂时忘记等式 (4.1.1),(4.1.2) 中的 $\boldsymbol{\alpha}_i,\boldsymbol{b}_j$ 代表一行和一列, 而将它们看成数, 将等式 (4.1.1) 当作一列乘一行, 等式 (4.1.2) 当作一列乘一个元素, 得到的结果仍与 (4.1.1),(4.1.2) 一致. 在等式 (4.1.3),(4.1.4) 中将 $\boldsymbol{A},\boldsymbol{b}_j,\boldsymbol{a}_j$ 都看成数, 将等式 (4.1.3) 当作一个元素乘一行, 等式 (4.1.4) 当作一行乘一列, 仍得到同样的结论.

由此看来, 矩阵乘法的法则不仅适用于以数为元素组成的矩阵, 还适用于以“小矩阵”为元素组成的“大矩阵”. 我们可以将任意矩阵 $\boldsymbol{A}$ 的某些相邻的行和相邻的列归并为一组, 从而将矩阵分成若干小的矩形区域, 每个矩形区域由行与列都在同一组的元素组成, 成为一个小矩阵, 称为 $\boldsymbol{A}$ 的一个**块**. 将 $\boldsymbol{A}$ 看成以块为元素的矩阵, 称为**分块矩阵** (partitioned matrix) .

将两个可以作乘法的矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 划分成分块矩阵. 如果分块恰当, 将 $\boldsymbol{A},\boldsymbol{B}$ 看成以这些块为元素组成的矩阵仍然能够相乘:

$$\boldsymbol{A}\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}&\cdots&\boldsymbol{A}_{1r}\\ \vdots&&\vdots\\ \boldsymbol{A}_{q1}&\cdots&\boldsymbol{A}_{qr}\end{pmatrix}\begin{pmatrix}\boldsymbol{B}_{11}&\cdots&\boldsymbol{B}_{1s}\\ \vdots&&\vdots\\ \boldsymbol{B}_{r1}&\cdots&\boldsymbol{B}_{rs}\end{pmatrix}=\begin{pmatrix}\boldsymbol{C}_{11}&\cdots&\boldsymbol{C}_{1s}\\ \vdots&&\vdots\\ \boldsymbol{C}_{q1}&\cdots&\boldsymbol{C}_{qs}\end{pmatrix},$$

其中 $\boldsymbol{C}_{ij}=\boldsymbol{A}_{i1}\boldsymbol{B}_{1j}+\cdots+\boldsymbol{A}_{ir}\boldsymbol{B}_{rj}$, 这样的乘法就称为分块矩阵的乘法. 具体地说, 要使分块乘法能够进行, 分块方式应满足如下条件: 设 $\boldsymbol{A}$ 的 m 行分成的 q 组依次包含 $h_1,\cdots,h_q$ 行, $\boldsymbol{A}$ 的 n 列分成的 r 组依次包含 $t_1,\cdots,t_r$ 列, 则 $\boldsymbol{A}$ 的第 (i,j) 块 $\boldsymbol{A}_{ij}$ 是 $h_i\times t_j$ 矩阵. $\boldsymbol{B}$ 的 n 行与 $\boldsymbol{A}$ 的列应按同样的方式分组, 各组依次包含 $t_1,\cdots,t_r$ 行. 设 $\boldsymbol{B}$ 的 p 列分成的 s 组依次包含 $l_1,\cdots,l_s$ 列, 则 $\boldsymbol{B}$ 的块 $\boldsymbol{B}_{ij}$ 为 $t_i\times l_j$ 矩阵. 每个 $h_i\times t_k$ 矩阵 $\boldsymbol{A}_{ik}$ 能够与 $t_k\times l_j$ 矩阵 $\boldsymbol{B}_{kj}$ 相乘得到 $h_i\times l_j$ 矩阵 $\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}$. r 个 $h_i\times l_j$ 矩阵 $\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}$ $(1\leqslant k\leqslant r)$ 能够相加得到 $\boldsymbol{C}_{ij}$. 以所有这些 $\boldsymbol{C}_{ij}$ 为块能够组成矩阵 $\boldsymbol{C}$. 可以验证, 通过分块矩阵乘法得到的这个矩阵 $\boldsymbol{C}$ 与直接按元素相乘得到的乘积 $\boldsymbol{AB}$ 一定相等. (验证过程略去.)

两个 $m\times n$ 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 以完全相同的方式分块之后, 可以看成以块为元素的矩阵, 将对应的块相加减:

$$\boldsymbol{A}\pm\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}_{11}&\cdots&\boldsymbol{A}_{1s}\\\vdots&&\vdots\\\boldsymbol{A}_{r1}&\cdots&\boldsymbol{A}_{rs}\end{pmatrix}\pm\begin{pmatrix}\boldsymbol{B}_{11}&\cdots&\boldsymbol{B}_{1s}\\\vdots&&\vdots\\\boldsymbol{B}_{r1}&\cdots&\boldsymbol{B}_{rs}\end{pmatrix}$$
$$=\begin{pmatrix}\boldsymbol{A}_{11}\pm\boldsymbol{B}_{11}&\cdots&\boldsymbol{A}_{1s}\pm\boldsymbol{B}_{1s}\\\vdots&&\vdots\\\boldsymbol{A}_{r1}\pm\boldsymbol{B}_{r1}&\cdots&\boldsymbol{A}_{rs}\pm\boldsymbol{B}_{rs}\end{pmatrix}.$$

矩阵的分块运算是一个有用的工具. 本节我们先利用这个工具来研究矩阵乘法的运算律.

4.1.3 矩阵乘法运算律

1. 单位矩阵、纯量矩阵与对角矩阵的乘法性质

例 4 试求 n 阶方阵 $\boldsymbol{X}$, 使 $\boldsymbol{AX}=\boldsymbol{A}$ 及 $\boldsymbol{XB}=\boldsymbol{B}$ 对所有能与 $\boldsymbol{X}$ 相乘的矩阵 $\boldsymbol{A},\boldsymbol{B}$ 成立.

解 n 维行向量空间 $\mathbf{R}^{1\times n}$ 中每个向量 $\boldsymbol{\alpha}=(a_1,\cdots,a_n)$ 在自然基 $E\coloneqq\{\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_n\}$

下的坐标等于 $\boldsymbol{\alpha}$ 本身:

$$\boldsymbol{\alpha}=a_1\boldsymbol{\varepsilon}_1+\cdots+a_n\boldsymbol{\varepsilon}_n=(a_1,\cdots,a_n)\begin{pmatrix}\boldsymbol{\varepsilon}_1\\\vdots\\\boldsymbol{\varepsilon}_n\end{pmatrix}=\boldsymbol{\alpha I},\qquad(4.1.5)$$

其中每个基向量 $\boldsymbol{\varepsilon}_i$ 是第 i 分量为 1、其余分量为 0 的行向量,

$$I = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

是以自然基向量为各行排成的方阵, 就是单位矩阵. 在 (4.1.5) 中将 α 的各分量 $a_1, \cdots, a_n$ 换成任意 n 个 m 维列向量, 从而将 α 换成任意 $m \times n$ 矩阵 A, 得到的等式 $A = AI$ 仍成立.

行向量空间 $\mathbf{R}^{1\times n}$ 的自然基 E 中各向量转置成列向量 $\varepsilon_1^{\mathrm{T}}, \cdots, \varepsilon_n^{\mathrm{T}}$ 组成列向量空间 $\mathbf{R}^{n\times 1}$ 的自然基 E', 每个列向量 $b = (b_1, \cdots, b_n)^{\mathrm{T}}$ 在这组基下的坐标等于 b 本身:

$$b = \varepsilon_1^{\mathrm{T}} b_1 + \cdots + \varepsilon_n^{\mathrm{T}} b_n = (\varepsilon_1^{\mathrm{T}}, \cdots, \varepsilon_n^{\mathrm{T}}) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = Ib, \qquad (4.1.6)$$

将列向量 b 的各分量换成任意 p 维行向量, 从而将 b 换成任意 $n \times p$ 矩阵 B, 得到等式 $B = IB$.

可见 $AI = A$, $IB = B$ 对能够与单位矩阵 I 相乘的所有矩阵 A, B 成立.

如果方阵 X 也具有同样的性质, 则 $X = XI = I$. 可见单位矩阵 I 是唯一满足条件的方阵. □

单位矩阵 I 的乘法性质 $AI = A, IB = B$ 与 1 在数的乘法中的性质 $1a = a1 = a$ 类似.

引理 4.1.1(单位矩阵的乘法性质) 单位矩阵 I 与任何一个可与它相乘的矩阵的乘积等于这个矩阵本身:

$$IA = A, \qquad BI = B.$$ □

例 5 已知

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

求 DA, AD.

解 将 $\boldsymbol{A}$ 的第 i 行记为 $\boldsymbol{\alpha}_i$, 第 j 列记为 $\boldsymbol{a}_j$. 则

$$\boldsymbol{DA}=\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_1\\\boldsymbol{\alpha}_2\\\boldsymbol{\alpha}_3\end{pmatrix}=\begin{pmatrix}\lambda_1\boldsymbol{\alpha}_1\\\lambda_2\boldsymbol{\alpha}_2\\\lambda_3\boldsymbol{\alpha}_3\end{pmatrix}=\begin{pmatrix}\lambda_1a_1&\lambda_1b_1&\lambda_1c_1\\\lambda_2a_2&\lambda_2b_2&\lambda_2c_2\\\lambda_3a_3&\lambda_3b_3&\lambda_3c_3\end{pmatrix},$$

$$\boldsymbol{AD}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix}=(\boldsymbol{a}_1\lambda_1,\boldsymbol{a}_2\lambda_2,\boldsymbol{a}_3\lambda_3)$$

$$=\begin{pmatrix}a_1\lambda_1&b_1\lambda_2&c_1\lambda_3\\a_2\lambda_1&b_2\lambda_2&c_2\lambda_3\\a_3\lambda_1&b_3\lambda_2&c_3\lambda_3\end{pmatrix}.$$ □

例 5 的算法可以推广到任意的 n 阶对角矩阵与矩阵的乘积.

引理 4.1.2(对角矩阵的乘法性质) 用对角矩阵 $\boldsymbol{D}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 左乘 $n\times p$ 矩阵 $\boldsymbol{B}$, 各对角元 $\lambda_1,\cdots,\lambda_n$ 分别乘 $\boldsymbol{B}$ 的各行得到 $\boldsymbol{DB}$; 用 $\boldsymbol{D}$ 右乘 $m\times n$ 矩阵 $\boldsymbol{A}$, 各对角元分别乘 $\boldsymbol{A}$ 的各列得到 $\boldsymbol{AD}$. □

注 我们说"用 $\boldsymbol{D}$ 左乘 $\boldsymbol{B}$", 是 $\boldsymbol{D}$ 在左边, 得到乘积 $\boldsymbol{DB}$. 说"用 $\boldsymbol{D}$ 右乘 $\boldsymbol{A}$", 是 $\boldsymbol{D}$ 在右边, 得到乘积 $\boldsymbol{AD}$. 矩阵乘法不满足交换律, 左乘与右乘在很多情况下是不一样的.

特别地, 当各对角元 $\lambda_1,\cdots,\lambda_n$ 等于同一个值 λ 时, 对角矩阵 $\mathrm{diag}(\lambda,\cdots,\lambda)$ 等于单位矩阵 $\boldsymbol{I}$ 的 λ 倍 $\lambda\boldsymbol{I}$, 称为**纯量矩阵** (scalar matrix).

引理 4.1.3(纯量矩阵的乘法性质) 纯量矩阵 $\lambda\boldsymbol{I}$ 与矩阵相乘, 相当于用数 λ 乘这个矩阵. 当 $\boldsymbol{A}\in F^{m\times n}$ 时, 有

$$\lambda\boldsymbol{A}=(\lambda\boldsymbol{I}_{(m)})\boldsymbol{A}=\boldsymbol{A}(\lambda\boldsymbol{I}_{(n)}).$$

□

$\lambda=1$ 时的纯量矩阵就是单位矩阵 $\boldsymbol{I}$, 它与矩阵 $\boldsymbol{A}$ 相乘相当于用 1 乘 $\boldsymbol{A}$, 等于 $\boldsymbol{A}$ 本身.

2. 分配律与结合律

引理 4.1.4(对加法的分配律) 对任意 $\boldsymbol{A}\in F^{m\times n}$, $\boldsymbol{B},\boldsymbol{C}\in F^{n\times p}$, $\boldsymbol{M},\boldsymbol{N}\in F^{q\times m}$, 有

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC},\qquad(\boldsymbol{M}+\boldsymbol{N})\boldsymbol{A}=\boldsymbol{MA}+\boldsymbol{NA}.$$

证明 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)$ 与 n 维列向量 $\boldsymbol{b}$ 的乘积 $\boldsymbol{Ab}=b_1\boldsymbol{a}_1+\cdots+b_n\boldsymbol{a}_n$ 是 $\boldsymbol{A}$ 的列向量组 S 的线性组合. S 的两个线性组合之和

$$\boldsymbol{Ab}+\boldsymbol{Ac}=(b_1\boldsymbol{a}_1+\cdots+b_n\boldsymbol{a}_n)+(c_1\boldsymbol{a}_1+\cdots+c_n\boldsymbol{a}_n)$$

$$=(b_1+c_1)\boldsymbol{a}_1+\cdots+(b_n+c_n)\boldsymbol{a}_n=\boldsymbol{A}(\boldsymbol{b}+\boldsymbol{c}).$$

在得到的等式 $\boldsymbol{A}(\boldsymbol{b}+\boldsymbol{c})=\boldsymbol{Ab}+\boldsymbol{Ac}$ 中将列向量 $\boldsymbol{b},\boldsymbol{c}$ 的各分量 b_i,c_i 都换成任意 p 维行向量 $\boldsymbol{\beta}_i,\boldsymbol{\gamma}_i$, 则 $\boldsymbol{b},\boldsymbol{c}$ 换成任意两个 $n\times p$ 矩阵 $\boldsymbol{B},\boldsymbol{C}$, 得到等式

$$A(B+C)=AB+AC.$$

类似地, A 的行向量组的两个线性组合之和 $\beta A+\gamma A=(\beta+\gamma)A$, 将行向量 β,γ 的各分量换成任意 q 维列向量, 从而 β,γ 换成任意两个 $q\times m$ 矩阵 M,N, 得到等式

$$(M+N)A=MA+NA.$$ □

引理 4.1.5(与数乘的结合律) 对任意 $A\in F^{m\times n},B\in F^{n\times p}$ 及 $\lambda\in F$, 有

$$\lambda(AB)=(\lambda A)B=A(\lambda B).$$

证明 设 A 的第 i 行为 α_i, B 的第 j 列为 B_j. 则易见三个矩阵 $\lambda(AB),(\lambda A)B,A(\lambda B)$ 的第 (i,j) 元 $\lambda(\alpha_iB_j),(\lambda\alpha_i)B_j,\alpha_i(\lambda B_j)$ 相等, 因此这三个矩阵相等. □

引理 4.1.6(乘法结合律) 对任意 $A\in F^{m\times n},B\in F^{n\times p},C\in F^{p\times q}$, 有

$$(AB)C=A(BC).$$

证明 设 $B=(b_1,\cdots,b_p)$, 则 $AB=(Ab_1,\cdots,Ab_p)$ 的各列的线性组合

$$(Ab_1)c_1+\cdots+(Ab_p)c_p=A(b_1c_1+\cdots+b_pc_p), \tag{4.1.7}$$

其中 $(Ab_1)c_1+\cdots+(Ab_p)c_p=(AB)c$, $b_1c_1+\cdots+b_pc_p=Bc$, c 是系数 $c_1,\cdots,c_p$ 排成的列向量. 等式 (4.1.7) 成为

$$(AB)c=A(Bc).$$

在这个等式中将 c 的各分量 c_i 换成任意 q 维行向量 γ_i, 则 c 变成任意 $p\times q$ 矩阵 C, 得到

$$(AB)C=A(BC).$$ □

例 6 $A=\begin{pmatrix}1&2&1&2\\-1&-2&-1&-2\\1&2&1&2\\-1&-2&-1&-2\end{pmatrix}$, 求 A^{10}.

解 观察到 A 的各行都可以写成同一个行向量 $\beta=(1,2,1,2)$ 的常数倍.

$$A=\begin{pmatrix}\beta\\-\beta\\\beta\\-\beta\end{pmatrix}=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}\beta=a\beta,\quad 其中a=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}.$$

$$A^{10}=\underbrace{(a\beta)(a\beta)\cdots(a\beta)}_{10个\ (a\beta)}=a\underbrace{(\beta a)(\beta a)\cdots(\beta a)}_{9个\ (\beta a)}\beta=a(\beta a)^9\beta.$$

由

$$\boldsymbol{\beta}\boldsymbol{a} = (1,2,1,2)\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = 1-2+1-2=-2,$$

得

$$\boldsymbol{A}^{10} = \boldsymbol{a}(-2)^9\boldsymbol{\beta} = -512\boldsymbol{a}\boldsymbol{\beta} = -512\boldsymbol{A} = \begin{pmatrix} -512 & -1024 & -512 & -1024 \\ 512 & 1024 & 512 & 1024 \\ -512 & -1024 & -512 & -1024 \\ 512 & 1024 & 512 & 1024 \end{pmatrix}.$$

□

在例 6 中, 按照分块运算的观点, 我们将 $\boldsymbol{\beta}$ 暂时看成"一阶方阵", 作为各行的公因子, 从右边提出来, 从而将 $\boldsymbol{A}$ 写成了列向量 $\boldsymbol{a}$ 与行向量 $\boldsymbol{\beta}$ 的乘积. 10 个 $\boldsymbol{A}$ 的乘积, 变成了 20 个矩阵的乘积, 似乎更繁琐了. 然而, 利用结合律将中间的 18 个矩阵重新分组, 每组都是一行乘一列得到一个数, 运算就变简单了.

例 6 的方法和结论可以推广到秩为 1 的任意方阵 $\boldsymbol{A}$.

例 7 设方阵 $\boldsymbol{A}$ 的秩为 1, 对角元之和为 λ, 求证: $\boldsymbol{A}^n = \lambda^{n-1}\boldsymbol{A}$.

证明 $\boldsymbol{A}$ 的秩为 1, 行向量组的极大线性无关组由一个非零行向量 $\boldsymbol{\beta} = (b_1,\cdots,b_n)$ 组成, $\boldsymbol{A}$ 的每一行 $\boldsymbol{\alpha}_i$ 都是 $\boldsymbol{\beta}$ 的常数倍: $\boldsymbol{\alpha}_i = a_i\boldsymbol{\beta}$. 于是

$$\boldsymbol{A} = \begin{pmatrix} a_1\boldsymbol{\beta} \\ \vdots \\ a_n\boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\boldsymbol{\beta} = \boldsymbol{a}\boldsymbol{\beta} = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & & \vdots \\ a_nb_1 & \cdots & a_nb_n \end{pmatrix},$$

$$\begin{aligned}\boldsymbol{A}^n &= \underbrace{(\boldsymbol{a}\boldsymbol{\beta})(\boldsymbol{a}\boldsymbol{\beta})\cdots(\boldsymbol{a}\boldsymbol{\beta})}_{n\text{个}(\boldsymbol{a}\boldsymbol{\beta})} = \boldsymbol{a}\underbrace{(\boldsymbol{\beta}\boldsymbol{a})(\boldsymbol{\beta}\boldsymbol{a})\cdots(\boldsymbol{\beta}\boldsymbol{a})}_{n-1\text{个}(\boldsymbol{\beta}\boldsymbol{a})}\boldsymbol{\beta} \\ &= \boldsymbol{a}\lambda^{n-1}\boldsymbol{\beta} = \lambda^{n-1}(\boldsymbol{a}\boldsymbol{\beta}) = \lambda^{n-1}\boldsymbol{A}.\end{aligned}$$

其中 $\lambda = \boldsymbol{\beta}\boldsymbol{a} = b_1a_1 + \cdots + b_na_n$ 等于 $\boldsymbol{A}$ 的各对角元 $a_1b_1,\cdots,a_nb_n$ 之和. □

方阵 $\boldsymbol{A}$ 的对角元之和称为 $\boldsymbol{A}$ 的**迹** (trace), 记作 $\operatorname{tr}\boldsymbol{A}$. 例 7 的结论是: 如果方阵 $\boldsymbol{A}$ 的秩 $\operatorname{rank}\boldsymbol{A} = 1$, 则 $\boldsymbol{A}^n = (\operatorname{tr}\boldsymbol{A})^{n-1}\boldsymbol{A}$.

4.1.4 方阵的多项式

由矩阵的乘法可以定义任一方阵 $\boldsymbol{A}$ 的正整数次幂:

$$\boldsymbol{A}^1 = \boldsymbol{A};\quad \boldsymbol{A}^2 = \boldsymbol{A}\boldsymbol{A},\quad \boldsymbol{A}^3 = (\boldsymbol{A}^2)\boldsymbol{A},\quad \boldsymbol{A}^4 = (\boldsymbol{A}^3)\boldsymbol{A},\quad \cdots$$

一般地, 设已对正整数 k 定义了 $\boldsymbol{A}^k$, 则定义 $\boldsymbol{A}^{k+1}=(\boldsymbol{A}^k)\boldsymbol{A}$. 这样就可定义 $\boldsymbol{A}$ 的任意正整数次幂.

由于矩阵乘法的结合律, 对任意正整数 m,s, 有

$$\boldsymbol{A}^m\boldsymbol{A}^s=\boldsymbol{A}^{m+s}.$$

有了方阵的各次幂, 可以将方阵代入多项式求值. 设 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_mx^m\in F[x]$ 是以 x 为字母、系数 $a_0,a_1,\cdots,a_m$ 在 F 中的多项式, $\boldsymbol{A}$ 是任一 n 阶方阵, 则

$$f(\boldsymbol{A})=a_0\boldsymbol{I}_n+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+\cdots+a_m\boldsymbol{A}^m$$

是 n 阶方阵. 注意多项式的常数项 a_0 应当换成 a_0 对应的纯量矩阵 $a_0\boldsymbol{I}$, 才能与 $\boldsymbol{A}$ 的各次幂的线性组合相加.

对于任意两个多项式 $f(x),g(x)\in F[x]$, 设 $s(x)=f(x)+g(x)$, $p(x)=f(x)g(x)$, 则对任意方阵 $\boldsymbol{A}\in F^{n\times n}$, 有

$$\begin{aligned}f(\boldsymbol{A})+g(\boldsymbol{A})&=s(\boldsymbol{A}),\\ f(\boldsymbol{A})g(\boldsymbol{A})&=p(\boldsymbol{A}).\end{aligned}$$

这是因为, 由多项式 $f(x),g(x)$ 计算它们的和 $s(x)$ 与积 $p(x)$ 所用到的加法交换律、加法结合律、乘法结合律、乘法对于加法的分配律, 方阵的运算都满足; 而在 $f(\boldsymbol{A}),g(\boldsymbol{A})$ 中出现的方阵都是一个方阵 $\boldsymbol{A}$ 的各次幂的线性组合, 它们在乘法中相互可交换; 因此由多项式的加法与乘法得出的结果, 将 $\boldsymbol{A}$ 代入后仍然成立.

例 8 求方阵 $\boldsymbol{A}$ 的幂 $\boldsymbol{A}^{10}$,

$$\boldsymbol{A}=\begin{pmatrix}\lambda&1&0\\0&\lambda&1\\0&0&\lambda\end{pmatrix}.$$

解 $\boldsymbol{A}=\lambda\boldsymbol{I}+\boldsymbol{N}$, 其中

$$\boldsymbol{N}=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix},\quad \boldsymbol{N}^2=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},\quad \boldsymbol{N}^3=\boldsymbol{O}.$$

由于纯量矩阵 $\lambda\boldsymbol{I}$ 与 $\boldsymbol{N}$ 做乘法时可交换: $(\lambda\boldsymbol{I})\boldsymbol{N}=\boldsymbol{N}(\lambda\boldsymbol{I})$, 可以用牛顿

二项式定理展开 $\boldsymbol{A}^{10}=(\lambda\boldsymbol{I}+\boldsymbol{N})^{10}$ 得:

$$\begin{aligned}\boldsymbol{A}^{10}=(\lambda\boldsymbol{I}+\boldsymbol{N})^{10}&=\lambda^{10}\boldsymbol{I}+10\lambda^9\boldsymbol{N}+\frac{10\times 9}{2}\lambda^8\boldsymbol{N}^2\\&=\begin{pmatrix}\lambda^{10}&10\lambda^9&45\lambda^8\\0&\lambda^{10}&10\lambda^9\\0&0&\lambda^{10}\end{pmatrix}.\end{aligned}$$ □

对任意数 a,b 证明关于 $(a+b)^n$ 展开式的牛顿二项式定理时, 用到了乘法对于加法的分配律、乘法交换律与结合律、加法交换律与结合律. 矩阵运算不满足乘法交换律, 但满足上述其余所有的运算律. 因此, 不能对任意同阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 用牛顿二项式定理, 也不能用依赖于乘法交换律的其他乘法公式 (例如 $(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2$). 但是, 某些方阵 $\boldsymbol{A},\boldsymbol{B}$ 作乘法时可交换: $\boldsymbol{AB}=\boldsymbol{BA}$, 就能对它们应用包括牛顿二项式定理在内的乘法公式. 例 8 中的纯量矩阵 $\lambda\boldsymbol{I}$ 与同阶方阵 $\boldsymbol{A}$ 作乘法时可交换, 因此可以用二项式定理. 在牛顿二项式定理展开式中本来还应当有 $\boldsymbol{N}^3,\boldsymbol{N}^4,\cdots$ 等项, 但这些项都等于零, 就不需要再考虑了.

4.1.5 转置与共轭

$m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 的转置矩阵 $\boldsymbol{A}^{\mathrm{T}}=(a_{ji})_{n\times m}$ 是一个 $n\times m$ 矩阵, 它的第 (j,i) 元等于 $\boldsymbol{A}$ 的第 (i,j) 元.

引理 4.1.7 矩阵的转置满足如下的运算律:

(1) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}$.

(2) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$.

(3) $(\lambda\boldsymbol{A})^{\mathrm{T}}=\lambda\boldsymbol{A}^{\mathrm{T}}$, λ 是任意数.

(4) 对 n 阶方阵 $\boldsymbol{A}$, 有 $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$.

(5) $\boldsymbol{A}=(\boldsymbol{A}_{ij})_{p\times q}$, 则 $\boldsymbol{A}^{\mathrm{T}}=(\boldsymbol{A}_{ji}^{\mathrm{T}})_{q\times p}$, 即

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}&\cdots&\boldsymbol{A}_{1q}\\\boldsymbol{A}_{21}&\boldsymbol{A}_{22}&\cdots&\boldsymbol{A}_{2q}\\\vdots&\vdots&&\vdots\\\boldsymbol{A}_{p1}&\boldsymbol{A}_{p2}&\cdots&\boldsymbol{A}_{pq}\end{pmatrix},\quad \boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{A}_{11}^{\mathrm{T}}&\boldsymbol{A}_{21}^{\mathrm{T}}&\cdots&\boldsymbol{A}_{p1}^{\mathrm{T}}\\\boldsymbol{A}_{12}^{\mathrm{T}}&\boldsymbol{A}_{22}^{\mathrm{T}}&\cdots&\boldsymbol{A}_{p2}^{\mathrm{T}}\\\vdots&\vdots&&\vdots\\\boldsymbol{A}_{1q}^{\mathrm{T}}&\boldsymbol{A}_{2q}^{\mathrm{T}}&\cdots&\boldsymbol{A}_{pq}^{\mathrm{T}}\end{pmatrix}.$$

(6) 对 $\boldsymbol{A}\in F^{m\times n},\boldsymbol{B}\in F^{n\times p}$ 有 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

证明 (1),(2),(3),(5) 容易直接验证. (4) 是 3.2.3 中的行列式的性质 6. 以下证明 (6).

将 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{B}=(b_{ij})_{n\times p}$ 按列分块得 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n),\boldsymbol{B}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p)$.

则 $\boldsymbol{AB}=(\boldsymbol{Ab}_1,\cdots,\boldsymbol{Ab}_p)$ 的各列依次是 $\boldsymbol{Ab}_1,\cdots,\boldsymbol{Ab}_p$, 将它们转置得到的行向量 $(\boldsymbol{Ab}_1)^{\mathrm{T}},\cdots,(\boldsymbol{Ab}_p)^{\mathrm{T}}$ 依次是 $(\boldsymbol{AB})^{\mathrm{T}}$ 的各行.

$\boldsymbol{AB}$ 的第 i 列 $\boldsymbol{Ab}_i$ 是 $\boldsymbol{A}$ 的各列的线性组合 $\boldsymbol{Ab}_i=b_{1i}\boldsymbol{a}_1+\cdots+b_{ni}\boldsymbol{a}_n$, 将这个等式中各个列向量转置成行向量得到 $(\boldsymbol{AB})^{\mathrm{T}}$ 的第 i 行

$$(\boldsymbol{Ab}_i)^{\mathrm{T}}=b_{1i}\boldsymbol{a}_1^{\mathrm{T}}+\cdots+b_{ni}\boldsymbol{a}_n^{\mathrm{T}}=(b_{1i},\cdots,b_{ni})\begin{pmatrix}\boldsymbol{a}_1^{\mathrm{T}}\\\vdots\\\boldsymbol{a}_n^{\mathrm{T}}\end{pmatrix}=\boldsymbol{b}_i^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}},$$

由此得到

$$(\boldsymbol{AB})^{\mathrm{T}}=\begin{pmatrix}(\boldsymbol{Ab}_1)^{\mathrm{T}}\\\vdots\\(\boldsymbol{Ab}_n)^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\\\vdots\\\boldsymbol{b}_n^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{b}_1^{\mathrm{T}}\\\vdots\\\boldsymbol{b}_n^{\mathrm{T}}\end{pmatrix}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}.\quad\square$$

设 $\boldsymbol{A}$ 是方阵. 如果 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$, 则称 $\boldsymbol{A}$ 为**对称方阵** (symmetric matrix). 如果 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$, 就称 $\boldsymbol{A}$ 是**反对称方阵** (anti-symmetric matrix), 也称**斜对称方阵** (skew symmetric matrix).

如果 $\boldsymbol{A}$ 既是对称方阵又是反对称方阵: $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$, 则 $\boldsymbol{A}=\boldsymbol{O}$.

例 9 求证:

(1) $\boldsymbol{A}$ 是任意矩阵, 则 $\boldsymbol{AA}^{\mathrm{T}}$ 是对称方阵.

(2) 任意方阵 $\boldsymbol{A}$ 可以唯一地写成对称方阵 $\boldsymbol{S}$ 与反对称方阵 $\boldsymbol{K}$ 之和, 即 $\boldsymbol{A}=\boldsymbol{S}+\boldsymbol{K}$.

证明 (1) $(\boldsymbol{AA}^{\mathrm{T}})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{AA}^{\mathrm{T}}$, 可见 $\boldsymbol{AA}^{\mathrm{T}}$ 是对称方阵.

(2) 取 $\boldsymbol{S}=\dfrac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}),\boldsymbol{K}=\dfrac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}})$. 则 $\boldsymbol{A}=\boldsymbol{S}+\boldsymbol{K}$.

由 $\boldsymbol{S}^{\mathrm{T}}=\dfrac{1}{2}(\boldsymbol{A}^{\mathrm{T}}+(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}})=\dfrac{1}{2}(\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A})=\boldsymbol{S}$ 及 $\boldsymbol{K}^{\mathrm{T}}=\dfrac{1}{2}(\boldsymbol{A}^{\mathrm{T}}-(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}})=-\dfrac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}})=-\boldsymbol{K}$ 知道 $\boldsymbol{S}$ 是对称方阵, $\boldsymbol{K}$ 是反对称方阵.

设 $\boldsymbol{A}=\boldsymbol{S}+\boldsymbol{K}=\boldsymbol{S}_1+\boldsymbol{K}_1$, $\boldsymbol{S},\boldsymbol{S}_1$ 都是对称方阵, $\boldsymbol{K},\boldsymbol{K}_1$ 都是反对称方阵. 则 $\boldsymbol{S}-\boldsymbol{S}_1=\boldsymbol{K}_1-\boldsymbol{K}$. 等式左边的 $\boldsymbol{S}_0=\boldsymbol{S}-\boldsymbol{S}_1$ 对称, 右边的 $\boldsymbol{K}_0=\boldsymbol{K}_1-\boldsymbol{K}$ 反对称. $\boldsymbol{S}_0=\boldsymbol{K}_0$ 既对称又反对称, 只能为零. 从而 $\boldsymbol{S}=\boldsymbol{S}_1$ 且 $\boldsymbol{K}_1=\boldsymbol{K}$. $\square$

对复数域 $\mathbf{C}$ 上任意矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 记 $\overline{\boldsymbol{A}}=(\overline{a_{ij}})_{m\times n}$. 也就是说: 将矩阵 $\boldsymbol{A}$ 中每个元素 a_{ij} 换成与它共轭的复数 $\overline{a_{ij}}$ 得到 $\overline{\boldsymbol{A}}$, 称为 $\boldsymbol{A}$ 的**共轭** (conjugation) . 容易验证, 关于矩阵的共轭有以下性质成立:

(1) $\forall\, \boldsymbol{A}, \boldsymbol{B} \in \mathbf{C}^{m\times n}$, $\overline{\boldsymbol{A}+\boldsymbol{B}} = \overline{\boldsymbol{A}} + \overline{\boldsymbol{B}}$;

(2) $\forall\, \lambda \in \mathbf{C}$, $\boldsymbol{A} \in \mathbf{C}^{m\times n}$, $\overline{\lambda \boldsymbol{A}} = \overline{\lambda}\,\overline{\boldsymbol{A}}$;

(3) $\forall\, \boldsymbol{A} \in \mathbf{C}^{m\times n}$, $\boldsymbol{B} \in \mathbf{C}^{n\times p}$, $\overline{\boldsymbol{A}\boldsymbol{B}} = \overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}$;

(4) $\forall\, \boldsymbol{A} \in \mathbf{C}^{m\times n}$, $\overline{\boldsymbol{A}}^{T} = \overline{\boldsymbol{A}^{\mathrm{T}}}$.

我们将复矩阵 $\boldsymbol{A}$ 的共轭矩阵 $\overline{\boldsymbol{A}}$ 的转置矩阵 $\overline{\boldsymbol{A}}^{\mathrm{T}}$ 简记为 $\boldsymbol{A}^*$. 则由共轭与转置的性质知如下性质成立:

(1) $\boldsymbol{A}^{**} = \boldsymbol{A}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})^* = \boldsymbol{A}^* + \boldsymbol{B}^*$;

(3) $(\lambda \boldsymbol{A})^* = \overline{\lambda} \boldsymbol{A}^*$;

(4) $(\boldsymbol{A}\boldsymbol{B})^* = \boldsymbol{B}^*\boldsymbol{A}^*$;

(5) $|\boldsymbol{A}^*| = \overline{|\boldsymbol{A}|}$.

如果复方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^* = \boldsymbol{A}$, 就称 $\boldsymbol{A}$ 为**埃尔米特方阵** (Hermitian matrix). 如果 $\boldsymbol{A}^* = -\boldsymbol{A}$, 就称 $\boldsymbol{A}$ 为 **斜埃尔米特方阵** (skew Hermitian matrix). 对实矩阵 $\boldsymbol{A}$ 有 $\boldsymbol{A}^* = \boldsymbol{A}^{\mathrm{T}}$, 因此实埃尔米特方阵就是对称方阵, 实斜埃尔米特方阵就是斜对称方阵.

习题 4.1

1. 设

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 2 & -1 & 4 \\ 1 & 0 & -2 \\ 0 & 3 & 1 \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} 0 & 2 \\ -1 & 0 \\ 3 & 1 \end{pmatrix},$$

计算矩阵 $\boldsymbol{AB}, \boldsymbol{B}^2, \boldsymbol{AC}, \boldsymbol{CA}, \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$. 根据计算结果, $\boldsymbol{AC}$ 与 $\boldsymbol{CA}$ 是否相等? $(\boldsymbol{AB})^{\mathrm{T}}$ 与 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$ 是否相等?

2. 设

(1) $\boldsymbol{A} = \begin{pmatrix} -1 & -2 & -4 \\ -1 & -2 & -4 \\ 1 & 2 & 4 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$;

(2) $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$.

计算矩阵 $\boldsymbol{AB}, \boldsymbol{BA}$. 得到的 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 是否相等?

3. 计算:

(1) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2$; (2) $\begin{pmatrix} 0 & 1 \\ -a & 0 \end{pmatrix}^2$; (3) $\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}^9$;

(4) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$; (5) $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^9$; (6) $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}^9$;

(7) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{10}$; (8) $\begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -2 \\ 4 & 4 & -4 \end{pmatrix}^{11}$

4. 已知方阵 $\boldsymbol{A}=\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$. 分别对如下的多项式 $f(x)$ 求 $f(\boldsymbol{A})$:

(1) $f(x)=1+x+x^2$; (2) $f(x)=x^3$; (3) $f(x)=x^{100}$.

5. 设 $n\geqslant 2$. 是否存在一个方阵 $\boldsymbol{A}\in F^{n\times n}$, 使 $F^{n\times n}$ 中所有的方阵都可以写成 $\boldsymbol{A}$ 的多项式的形式 $a_0\boldsymbol{I}+a_1\boldsymbol{A}+\cdots+a_m\boldsymbol{A}^m$ (m 为正整数, $a_0,a_1,\cdots,a_m\in F$)? 并说明理由.

6. 举出分别满足下列条件的整系数 2 阶方阵 $\boldsymbol{A}$:

(1) $\boldsymbol{A}\neq\pm\boldsymbol{I}$ 但 $\boldsymbol{A}^2=\boldsymbol{I}$; (2) $\boldsymbol{A}^2=-\boldsymbol{I}$;

(3) $\boldsymbol{A}\neq\boldsymbol{I}$ 且 $\boldsymbol{A}^3=\boldsymbol{I}$; (4) $\boldsymbol{A}^2=-2\boldsymbol{I}$.

7. 计算:

(1) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}^5$; (2) $\begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}^5$.

§4.2 矩阵乘法与线性变换

4.2.1 平面上的旋转变换

例 1 在平面上建立直角坐标系. 将平面上每个点 P 绕原点旋转 α 角到点 P'. 试写出由点 P 的坐标 (x,y) 计算点 P' 的坐标 (x',y') 的函数关系式.

注 通常规定沿逆时针方向旋转得到的角为正角, 顺时针方向旋转得到的角为负角. 因此, 本书所说“旋转 α 角”都是“沿逆时针方向旋转 α 角”; 如果沿顺时针方向旋转 α 角, 就说“旋转 $-\alpha$ 角”.

解 将 OP 旋转 $90°$ 得到 OQ, 如图 4–1 . 则

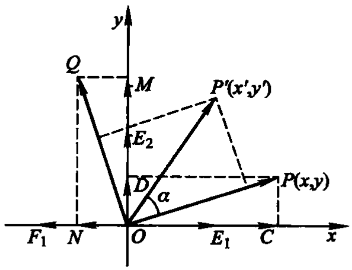

图 4–1

$$\overrightarrow{OP'}=(\cos\alpha)\overrightarrow{OP}+(\sin\alpha)\overrightarrow{OQ}. \tag{4.2.1}$$

点 P 的坐标 (x,y) 就是向量 $\overrightarrow{OP}$ 在自然基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2\}$ 下的坐标. 再求出 $\overrightarrow{OQ}$ 的坐标. 代入 (4.2.1) 即可求出 $\overrightarrow{OP'}$ 的坐标 (x',y').

$\overrightarrow{OP}=x\boldsymbol{e}_1+y\boldsymbol{e}_2=x\overrightarrow{OE_1}+y\overrightarrow{OE_2}=\overrightarrow{OC}+\overrightarrow{OD}$, 其中 $\overrightarrow{OE_1}=\boldsymbol{e}_1,\overrightarrow{OE_2}=\boldsymbol{e}_2$ 分别是 x 轴与 y 轴正方向上的单位向量. 将平面上所有的图形绕原点 O 旋转直角, 则矩形 $OCPD$ 变成矩形 $OMQN$, 相互垂直的单位线段 OE_1,OE_2 分别变成相互垂直的单位线段 OE_2,OF_1, 因而

$$\overrightarrow{OQ}=\overrightarrow{OM}+\overrightarrow{ON}=x\overrightarrow{OE_2}+y\overrightarrow{OF_1}=x\boldsymbol{e}_2+y(-\boldsymbol{e}_1)=(-y)\boldsymbol{e}_1+x\boldsymbol{e}_2. \quad (4.2.2)$$

这就求出了 $\overrightarrow{OQ}$ 的坐标为 $(-y,x)$. 将 $\overrightarrow{OP'},\overrightarrow{OP},\overrightarrow{OQ}$ 的坐标都写成列向量, 代入 (4.2.1) 得

$$\begin{pmatrix} x' \\ y' \end{pmatrix}=(\cos\alpha)\begin{pmatrix} x \\ y \end{pmatrix}+(\sin\alpha)\begin{pmatrix} -y \\ x \end{pmatrix}=\begin{pmatrix} x\cos\alpha-y\sin\alpha \\ x\sin\alpha+y\cos\alpha \end{pmatrix}, \quad (4.2.3)$$

即

$$\begin{cases} x'=x\cos\alpha-y\sin\alpha, \\ y'=x\sin\alpha+y\cos\alpha. \end{cases}$$

□

例 1 得到的旋转后的每个坐标 x',y' 都是旋转前的坐标 x,y 的线性组合, 因此, 旋转前后坐标 $\boldsymbol{X}=\begin{pmatrix} x \\ y \end{pmatrix}$ 与 $\boldsymbol{Y}=\begin{pmatrix} x' \\ y' \end{pmatrix}$ 之间的关系能够写成矩阵乘积的形式:

$$\boldsymbol{Y}=\begin{pmatrix} x' \\ y' \end{pmatrix}=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\boldsymbol{AX}, \quad (4.2.4)$$

一般地, 如果 n 元函数 $y=f(x_1,\cdots,x_n)$ 的因变量是自变量的常系数线性组合 $y=a_1x_1+\cdots+a_nx_n$, 这样的函数就称为**线性函数** (linear function) . 将 n 个自变量看成一个整体, 写成列向量的形式 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$, f 就成为以 $F^{n\times 1}$ 中的列向量 $\boldsymbol{X}$ 为自变量的一元函数, 也就是 $F^{n\times 1}$ 到 F 的一个映射. 因变量 y 可以写成常数组成的行向量 $\boldsymbol{\alpha}=(a_1,\cdots,a_n)$ 与自变量 $\boldsymbol{X}$ 的乘积的形式: $y=f(\boldsymbol{X})=\boldsymbol{\alpha X}$, f 的作用 $\boldsymbol{X}\mapsto\boldsymbol{\alpha X}$ 看起来好像是用 "常数" $\boldsymbol{\alpha}$ 乘自变量 $\boldsymbol{X}$. 这个行向量 $\boldsymbol{\alpha}$ 称为线性函数 f 的矩阵.

设 $\sigma: F^{n\times 1}\to F^{m\times 1},\boldsymbol{X}\mapsto\boldsymbol{Y}$ 是数域 F 上两个列向量空间之间的映射. 如果 $\boldsymbol{Y}=(y_1,\cdots,y_m)^{\mathrm{T}}$ 的每个分量 y_i 都是 $\boldsymbol{X}$ 的线性函数 $y_i=\boldsymbol{\alpha}_i\boldsymbol{X}=a_{i1}x_1+\cdots+a_{in}x_n$, 则

$$Y=\begin{pmatrix} y_1\\ \vdots\\ y_m\end{pmatrix}=\begin{pmatrix} \alpha_1 X\\ \vdots\\ \alpha_m X\end{pmatrix}=AX,\quad 其中\quad A=\begin{pmatrix} \alpha_1\\ \vdots\\ \alpha_m\end{pmatrix}.$$

映射 $\sigma: X\mapsto AX$ 可以通过用矩阵 A 左乘自变量 X 实现, 这样的映射 σ 称为**线性映射**. 当 $m=1$ 时 σ 就是线性函数. 如果 $m=n$, σ 是 $F^{n\times 1}$ 到自身的映射, 也就是 $F^{n\times 1}$ 上的一个变换, 称为**线性变换**.

两个线性映射 $\sigma: F^{n\times 1}\to F^{m\times 1}, X\mapsto AX$ 与 $\tau: F^{m\times 1}\to F^{p\times 1}, X\mapsto BX$ 的复合映射 $\tau\sigma: F^{n\times 1}\to F^{p\times 1}, X\mapsto B(AX)=(BA)X$ 是由矩阵 BA 决定的线性映射, $\tau\sigma$ 的矩阵 BA 是 τ 与 σ 的矩阵 B, A 的乘积.

例 2 已知

$$A=\begin{pmatrix} \cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix},\qquad B=\begin{pmatrix} \cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}.$$

(1) 求 AB.

(2) n 是任意正整数, 求 A^n.

解 将平面上每个点 $P(x,y)$ 的坐标写成列向量 $X=\begin{pmatrix} x\\ y\end{pmatrix}$, 用 X 代表点 P. 则 $\sigma: X\mapsto AX$ 是将每个点绕原点旋转 α 角的变换, $\tau: X\mapsto BX$ 是绕原点旋转 β 角的变换.

(1) 方阵 AB 决定的变换 $X\mapsto (AB)X=A(BX)$ 是 σ 与 τ 的复合变换, 将每个点 P 先旋转 β 角再旋转 α 角, 总的效果是旋转了 $\alpha+\beta$ 角. 这个变换的矩阵

$$AB=\begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{pmatrix}.$$

(2) $X\mapsto A^nX$ 将旋转变换 $\sigma: X\mapsto AX$ 重复进行 n 次, 总效果 $\sigma^n: X\mapsto A^nX$ 是将每个点 P 旋转了角 $n\alpha$, 这个变换的矩阵

$$A^n=\begin{pmatrix} \cos n\alpha & -\sin n\alpha\\ \sin n\alpha & \cos n\alpha\end{pmatrix}.$$ □

例 3 的结果也可以直接用矩阵运算和数学归纳法得到. 但由矩阵乘法所代表的几何变换得出结论更自然, 更简明. 如果认为这样的解法的叙述在数学上不够严密, 可以先由几何解法想出结果, 再用矩阵乘法法则和数学归纳法写出严格的验证过程.

4.2.2 线性变换的矩阵

例 3 是否存在 $\mathbf{R}^{2\times 1}$ 上的线性变换 σ 将 $\boldsymbol{e}_1=\begin{pmatrix}1\\0\end{pmatrix},\boldsymbol{e}_2=\begin{pmatrix}0\\1\end{pmatrix}$ 分别映射到 $\begin{pmatrix}2\\3\end{pmatrix},\begin{pmatrix}4\\5\end{pmatrix}$? 如果存在, 是否唯一?

解 设 $\mathbf{R}^{2\times 1}$ 上的线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 的矩阵为 $\boldsymbol{A}$. 则

$$\sigma(\boldsymbol{e}_1)=\boldsymbol{A}\boldsymbol{e}_1=\boldsymbol{A}_1 \text{ 与 } \sigma(\boldsymbol{e}_2)=\boldsymbol{A}\boldsymbol{e}_2=\boldsymbol{A}_2$$

分别是 $\boldsymbol{A}$ 的第 1 列和第 2 列. 因此, σ 的矩阵 $\boldsymbol{A}$ 所满足的充分必要条件为

$$\boldsymbol{A}_1=\sigma(\boldsymbol{e}_1)=\begin{pmatrix}2\\3\end{pmatrix} \quad \text{且} \quad \boldsymbol{A}_2=\sigma(\boldsymbol{e}_2)=\begin{pmatrix}4\\5\end{pmatrix},$$

满足此条件的矩阵 $\boldsymbol{A}$ 显然存在且唯一, 就是以 $\begin{pmatrix}2\\3\end{pmatrix},\begin{pmatrix}4\\5\end{pmatrix}$ 为两列排成的矩阵

$$\boldsymbol{A}=\begin{pmatrix}2&4\\3&5\end{pmatrix}.$$

因此线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 存在且唯一. □

一般地, 有:

引理 4.2.1 从 $F^{n\times 1}$ 到 $F^{m\times 1}$ 的线性映射 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 的矩阵 $\boldsymbol{A}=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$ 的各列 $\boldsymbol{A}_j=\sigma(\boldsymbol{e}_j)$, 分别等于各个自然基向量 $\boldsymbol{e}_j$ $(1\leqslant j\leqslant n)$ 在映射 σ 下的像.

例 4 设平面直角坐标系中的直线 l 由 x 轴绕原点 O 旋转角 α 得到, 如图 4–2. 将平面上每一点 $P(x,y)$ 的坐标写成列向量 $\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}$ 的形式, 将 P 关于 l 的对称点 $P'(x',y')$ 的坐标 $\boldsymbol{Y}=\begin{pmatrix}x'\\y'\end{pmatrix}$ 记为 $\sigma(\boldsymbol{X})$. 如果已经知道 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{Y}$ 是 $\mathbf{R}^{2\times 1}$ 上的线性变换, 求 (x',y') 与 (x,y) 之间的函数关系式.

解 线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{Y}=\boldsymbol{AX}$ 的矩阵 $\boldsymbol{A}$ 的两列分别等于 $\sigma(\boldsymbol{e}_1),\sigma(\boldsymbol{e}_2)$. 如图 4–2, E_1,E_2 分别是自然基向量 $\boldsymbol{e}_1=\overrightarrow{OE_1},\boldsymbol{e}_2=\overrightarrow{OE_2}$ 所代表的点, E_1',E_2' 分别是它们关于 l 的对称点. 以表示 E_1',E_2' 的向量 $\overrightarrow{OE_1'},\overrightarrow{OE_2'}$ 为两列组成的矩阵 $\boldsymbol{A}$ 就是线性变换 σ 的矩阵.

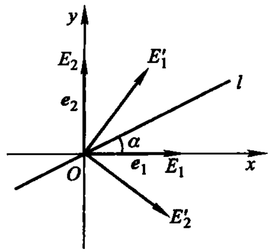

图 4-2

E_1 与 E_1' 关于直线 l 对称 $\Rightarrow$ 直线 l 平分 $\angle E_1OE_1' \Rightarrow \angle xOE_1' = 2\alpha$.

l 平分 $\angle E_2OE_2' \Rightarrow \dfrac{1}{2}(\angle xOE_2+\angle xOE_2') = \alpha \Rightarrow \angle xOE_2' = 2\alpha - \angle xOE_2 = 2\alpha - \dfrac{\pi}{2}$. 再由 $|OE_1'| = |OE_2'| = 1$, 即得

$$\sigma(\boldsymbol{e}_1) = \overrightarrow{OE_1'} = \begin{pmatrix} \cos 2\alpha \\ \sin 2\alpha \end{pmatrix},$$

$$\sigma(\boldsymbol{e}_2) = \overrightarrow{OE_2'} = \begin{pmatrix} \cos\left(2\alpha - \dfrac{\pi}{2}\right) \\ \sin\left(2\alpha - \dfrac{\pi}{2}\right) \end{pmatrix} = \begin{pmatrix} \sin 2\alpha \\ -\cos 2\alpha \end{pmatrix},$$

$$\boldsymbol{A} = (\sigma(\boldsymbol{e}_1), \sigma(\boldsymbol{e}_2)) = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}.$$

$$x' = x\cos 2\alpha + y\sin 2\alpha, \qquad y' = x\sin 2\alpha - y\cos 2\alpha. \qquad \square$$

在例 1 中, 如果预先知道旋转变换 σ 可以通过矩阵乘法 $\boldsymbol{Y} = \boldsymbol{AX}$ 实现, 也可以直接由 $\sigma(\boldsymbol{e}_1), \sigma(\boldsymbol{e}_2)$ 的坐标得到 $\boldsymbol{A}$. 很自然要问: 怎样知道变换 $\sigma : \boldsymbol{X} \mapsto \boldsymbol{Y}$ 可以由矩阵乘法 $\boldsymbol{Y} = \boldsymbol{AX}$ 实现?

由矩阵乘法的性质 $\boldsymbol{A}(\boldsymbol{X}_1 + \boldsymbol{X}_2) = \boldsymbol{AX}_1 + \boldsymbol{AX}_2$ 及 $\boldsymbol{A}(\lambda\boldsymbol{X}_1) = \lambda\boldsymbol{A}(\boldsymbol{X}_1)$ 知: 由矩阵 $\boldsymbol{A}$ 左乘引起的线性映射 $\sigma : \boldsymbol{X} \mapsto \boldsymbol{Y} = \boldsymbol{AX}$ 满足必要条件

$$\sigma(\boldsymbol{X}_1 + \boldsymbol{X}_2) = \sigma(\boldsymbol{X}_1) + \sigma(\boldsymbol{X}_2), \quad \sigma(\lambda\boldsymbol{X}_1) = \lambda\sigma(\boldsymbol{X}_1).$$

反过来, 容易证明这个条件也是充分的.

定理 4.2.1 如果映射 $\sigma : F^{n\times 1} \to F^{m\times 1}$ 对任意 $\boldsymbol{X}_1, \boldsymbol{X}_2 \in F^{n\times 1}$ 与 $\lambda \in F$ 满足条件

$$\sigma(\boldsymbol{X}_1 + \boldsymbol{X}_2) = \sigma(\boldsymbol{X}_1) + \sigma(\boldsymbol{X}_2), \qquad \sigma(\lambda\boldsymbol{X}_1) = \lambda\sigma(\boldsymbol{X}_1), \tag{4.2.5}$$

则存在矩阵 $\boldsymbol{A} \in F^{m\times n}$ 使 $\sigma(\boldsymbol{X}) = \boldsymbol{AX}$ 对所有的 $\boldsymbol{X} \in F^{n\times 1}$ 成立.

解 由条件 4.2.5 可以推出:

$$\sigma(\lambda_1 \boldsymbol{X}_1 + \cdots + \lambda_k \boldsymbol{X}_k) = \lambda_1 \sigma(\boldsymbol{X}_1) + \cdots + \lambda_k \sigma(\boldsymbol{X}_k) \tag{4.2.6}$$

对任意 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_k \in F^{n\times 1}$ 和 $\lambda_1, \cdots, \lambda_k \in F$ 成立. 特别地, 对任意

$$\boldsymbol{X} = (x_1, \cdots, x_n)^{\mathrm{T}} = x_1 \boldsymbol{e}_1 + \cdots + x_n \boldsymbol{e}_n \in F^{n\times 1}$$

有

$$\sigma(\boldsymbol{X}) = x_1 \sigma(\boldsymbol{e}_1) + \cdots + x_n \sigma(\boldsymbol{e}_n) = (\sigma(\boldsymbol{e}_1), \cdots, \sigma(\boldsymbol{e}_n)) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{A}\boldsymbol{X},$$

其中 $\boldsymbol{A} = (\sigma(\boldsymbol{e}_1), \cdots, \sigma(\boldsymbol{e}_n))$ 是以各自然基向量 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 在 σ 作用下的像 $\sigma(\boldsymbol{e}_1), \cdots, \sigma(\boldsymbol{e}_n)$ 为各列排成的矩阵. □

因此, 线性映射可以定义如下:

定义 4.2.1 设 V, U 是数域 F 上的向量空间, σ 是 U 到 V 的映射. 如果 (1) (保加法) $\sigma(\boldsymbol{a} + \boldsymbol{b}) = \sigma(\boldsymbol{a}) + \sigma(\boldsymbol{b})$;(2) (保数乘) $\sigma(\lambda \boldsymbol{a}) = \lambda \sigma(\boldsymbol{a})$ 对任意向量 $\boldsymbol{a}, \boldsymbol{b} \in V$ 和纯量 $\lambda \in F$ 成立, 就称 σ 是**线性映射** (linear mapping). □

各取定 V, U 中的一组基 S, T, 将 V, U 中的向量在各自的基 S, T 下用坐标来表示, 则线性映射 $\sigma : \boldsymbol{X} \mapsto \sigma(\boldsymbol{X})$ 可以通过用某个矩阵 $\boldsymbol{A}$ 左乘 $\boldsymbol{X}$ 来实现: $\sigma : \boldsymbol{X} \mapsto \boldsymbol{A}\boldsymbol{X}$. 这个矩阵 $\boldsymbol{A}$ 称为 σ 在基 S, T 下的矩阵.

当 $U = F$ 时称 σ 是 V 上的**线性函数** (linear function), 当 $U = V$ 时称 σ 是 V 上的**线性变换** (linear transformation) .

例 5 l 是平面上过原点的直线. 试证明: 将平面上每个点 P 的坐标 $\boldsymbol{X}$ 对应到它关于直线 l 的对称点 P' 的坐标的变换 $\sigma : \boldsymbol{X} \mapsto \boldsymbol{Y}$ 是线性变换.

证明 只需证明 $\sigma(\boldsymbol{a} + \boldsymbol{b}) = \sigma(\boldsymbol{a}) + \sigma(\boldsymbol{b})$ 及 $\sigma(\lambda \boldsymbol{a}) = \lambda \sigma(\boldsymbol{a})$ 对任意向量 $\boldsymbol{a}, \boldsymbol{b}$ 及实数 λ 成立. 作 $\boldsymbol{a} = \overrightarrow{OA}$, $\boldsymbol{b} = \overrightarrow{OB}$, $\overrightarrow{OC} = \boldsymbol{a} + \boldsymbol{b}$, $\overrightarrow{OM} = \lambda \boldsymbol{a}$ $(\lambda > 0)$; $\sigma(\boldsymbol{a}) = \overrightarrow{OA'}$, $\sigma(\boldsymbol{b}) = \overrightarrow{OB'}$, $\sigma(\boldsymbol{a} + \boldsymbol{b}) = \overrightarrow{OC'}$, $\sigma(\lambda \boldsymbol{a}) = \overrightarrow{OM'}$. 如图 4–3.

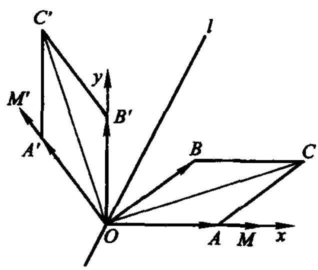

图 4–3

当 $\boldsymbol{a}=\boldsymbol{0}$ 或 $\boldsymbol{b}=\boldsymbol{0}$ 或 $\lambda=0$ 时命题显然成立. 以下设 $\boldsymbol{a}\neq\boldsymbol{0},\boldsymbol{b}\neq\boldsymbol{0},\lambda\neq 0$.

关于轴对称的变换 σ 保持图形的形状大小不变, 变换前后的图形相互全等.

由于 $\lambda>0$, $\overrightarrow{OM}=\lambda\overrightarrow{OA}\Rightarrow OA,OM$ 方向相同且 $|OM|=\lambda|OA|$. 经过 σ 变换得到的 OA',OM' 方向仍相同且 $|OM'|=\lambda|OA'|$, 因此 $\overrightarrow{OM'}=\lambda\overrightarrow{OA'}$, 即 $\sigma(\lambda\boldsymbol{a})=\lambda\sigma(\boldsymbol{a})$. 易见 $\sigma(-\overrightarrow{OM})=-\sigma(\overrightarrow{OM})$, 因此 $\sigma((-\lambda)\boldsymbol{a})=-\lambda\sigma(\boldsymbol{a})$, 说明 $\sigma(\lambda\boldsymbol{a})=\lambda\sigma(\boldsymbol{a})$ 当 $\lambda<0$ 也成立.

当 OA,OB 不共线时, $\overrightarrow{OA}+\overrightarrow{OB}=\overrightarrow{OC}\Leftrightarrow OACB$ 是平行四边形, 经过 σ 变成的 $OA'C'B'$ 仍是平行四边形, 因此 $\overrightarrow{OC'}=\overrightarrow{OA'}+\overrightarrow{OB'}$, 即 $\sigma(\boldsymbol{a}+\boldsymbol{b})=\sigma(\boldsymbol{a})+\sigma(\boldsymbol{b})$.

如果 OA,OB 共线, 则 $\boldsymbol{b}=\lambda\boldsymbol{a}$ 对某个实数 λ 成立, $\sigma(\boldsymbol{a}+\boldsymbol{b})=\sigma((1+\lambda)\boldsymbol{a})=(1+\lambda)\sigma(\boldsymbol{a})=\sigma(\boldsymbol{a})+\lambda\sigma(\boldsymbol{a})=\sigma(\boldsymbol{a})+\sigma(\lambda\boldsymbol{a})=\sigma(\boldsymbol{a})+\sigma(\boldsymbol{b})$ 仍成立. □

例 5 中的证明不但适用于对轴对称变换 σ, 也适用于几何平面或三维空间中保持图形的形状大小不变并且原点不变的所有变换 σ, 例如: 平面上绕原点的旋转, 空间中绕过原点的直线的旋转, 空间中关于原点的对称或关于过原点的平面的对称. 所有这些变换 σ 都是线性变换, 都可以通过左乘某个矩阵 $\boldsymbol{A}$ 来实现, $\boldsymbol{A}$ 的各列依次是各自然基向量 $\boldsymbol{e}_i$ 的像 $\sigma(\boldsymbol{e}_i)$ 的坐标.

例 6 在三维几何空间中建立了直角坐标系. 变换 σ 将每个点 P 绕 Oz 轴旋转角 α 到 P'. 求 $P'(x',y',z')$ 与 $P(x,y,z)$ 的坐标之间的函数关系式.

注 在三维空间中, 将旋转轴看成一条射线, 面对它的正方向看, 以沿逆时针方向旋转角为正角, 顺时针方向为负角. 绕 Oz 旋转角 α, 就是面对 Oz 的正方向看起来是沿逆时针方向旋转角 α .

解 三维几何空间中保持原点不动的旋转变换 σ 是线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$. 变换矩阵 $\boldsymbol{A}$ 的各列 $\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$ 依次是自然基向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3$ 在 σ 作用下的像 $\sigma(\boldsymbol{e}_1),\sigma(\boldsymbol{e}_2),\sigma(\boldsymbol{e}_3)$. 转轴 Oz 上的 $\boldsymbol{e}_3$ 保持不变: $\sigma(\boldsymbol{e}_3)=\boldsymbol{e}_3=(0,0,1)^{\mathrm{T}}$. $\boldsymbol{e}_1,\boldsymbol{e}_2$ 在 Oxy 平面内旋转角 α: $\sigma(\boldsymbol{e}_1)=(\cos\alpha,\sin\alpha,0)^{\mathrm{T}}$, $\sigma(\boldsymbol{e}_2)=(-\sin\alpha,\cos\alpha,0)^{\mathrm{T}}$. 因此

$$\boldsymbol{A}=\begin{pmatrix}\cos\alpha & -\sin\alpha & 0\\ \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix},$$

$$\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}x\cos\alpha-y\sin\alpha\\ x\sin\alpha+y\cos\alpha\\ z\end{pmatrix},$$

$$\begin{cases} x' = x\cos\alpha - y\sin\alpha, \\ y' = x\sin\alpha + y\cos\alpha, \\ z' = z. \end{cases}$$ □

习题 4.2

1. 计算矩阵乘积. 并指出这些乘积代表平面上什么变换.

(1) $\begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}^{2010}$; (2) $\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}\begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix}$;

(3) $\begin{pmatrix} \cos\beta & \sin\beta \\ \sin\beta & -\cos\beta \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$.

2. 在平面直角坐标系中, 将 x 轴绕原点沿逆时针旋转 α 角和 β 角得到直线 l_1 与 l_2. 设平面上每个点 P 关于直线 l_1 的对称点为 P_1, P_1 关于 l_2 的对称点为 P_2. 试求 P_2 的坐标 (x', y') 与 P 的坐标 (x, y) 之间的关系式. 映射 $\sigma: P \mapsto P_2$ 是什么变换?

3. 将空间直角坐标系中每一点 $P(x, y, z)$ 绕 Ox 轴旋转角 α 到 $P'(x', y', z')$, 求旋转前后的坐标 (x, y, z) 与 (x', y', z') 之间的关系式.

4. 以下矩阵分别代表三维几何空间中的什么变换? 并分别求出它们的 19 次幂.

(1) $\begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$; (3) $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

附录 5 复数乘法的几何意义

中学数学硬性规定符号 i 满足 $\mathrm{i}^2 = -1$, 不讲任何道理. 虽然也用平面向量 (a, b) 来表示复数 $a + b\mathrm{i}$, 用向量加法 $(a, b) + (c, d) = (a + c, b + d)$ 表示复数加法 $(a + b\mathrm{i}) + (c + d\mathrm{i}) = (a + c) + (b + d)\mathrm{i}$, 但不讲复数乘法的几何意义, 学生仍然不知 $\mathrm{i}^2 = -1$ 有什么意义, 还容易误认为复数乘法代表向量内积.

其实, $\mathrm{i}^2 = -1$ 有一个简单易懂的几何模型. -1 乘向量 $\boldsymbol{u}$ 将向量“向后转” (旋转 180°): $\boldsymbol{u} \mapsto (-1)\boldsymbol{u} = -\boldsymbol{u}$. $(-1)^2$ 就是将“向后转”进行两次: $\boldsymbol{u} \mapsto -\boldsymbol{u} \mapsto -(-\boldsymbol{u}) = \boldsymbol{u}$, 旋转 360° 回到原来的 $\boldsymbol{u}$, 这解释了 $(-1)^2 = 1$. 如果将“向后转”进行“半次”, 向左 (逆时针方向) 旋转 90°, 并且将这个动作记为 i, 则 i^2 就是旋转 180°, 相当于乘 -1, 这就是 $\mathrm{i}^2 = -1$. 向右转进行两次也是向后转, 这是说 $(-\mathrm{i})^2 = -1$.

将向量 $\boldsymbol{u}$ 与 $\mathrm{i}\boldsymbol{u}$ 的实系数线性组合 $a\boldsymbol{u} + b(\mathrm{i}\boldsymbol{u})$ 写成 $(a + b\mathrm{i})\boldsymbol{u}$, 看成 $\boldsymbol{u}$ 的 $a + b\mathrm{i}$ 倍. 特别地, 坐标为 (x, y) 的平面向量 $\boldsymbol{u}$ 可以写成自然基向量 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 的

实系数线性组合 $\boldsymbol{u}=x\boldsymbol{e}_1+y\boldsymbol{e}_2$. 由于 $\boldsymbol{e}_2=\mathrm{i}\boldsymbol{e}_1$, $\boldsymbol{u}=x\boldsymbol{e}_1+y\mathrm{i}\boldsymbol{e}_1=(x+y\mathrm{i})\boldsymbol{e}_1$ 就是 $\boldsymbol{e}_1$ 的 $x+y\mathrm{i}$ 倍, 可以用复数 $x+y\mathrm{i}$ 表示. 两个复数 $a+b\mathrm{i}$ 与 $x+y\mathrm{i}$ 的乘积 $(a+b\mathrm{i})(x+y\mathrm{i})$ 代表的就是 $\boldsymbol{u}=(x+y\mathrm{i})\boldsymbol{e}_1$ 的 $a+b\mathrm{i}$ 倍, 也就是 $a\boldsymbol{u}+b(\mathrm{i}\boldsymbol{u})$.

用 $\rho_\alpha(\boldsymbol{u})$ 表示将 $\boldsymbol{u}=\overrightarrow{OP}$ 旋转角 α 得到的向量 $\overrightarrow{OP'}$. 如 §4.2 例 1 中的图 4–1 及等式 (4.2.1) 所示,

$$\overrightarrow{OP'}=(\cos\alpha)\overrightarrow{OP}+(\sin\alpha)\overrightarrow{OQ}=(\cos\alpha)\boldsymbol{u}+(\sin\alpha)(\mathrm{i}\boldsymbol{u})=(\cos\alpha+\mathrm{i}\sin\alpha)\boldsymbol{u}$$

是 $\boldsymbol{u}$ 的 $\cos\alpha+\mathrm{i}\sin\alpha$ 倍. 可见, 用 $\cos\alpha+\mathrm{i}\sin\alpha$ 乘 $x+y\mathrm{i}$ 表示将平面向量 $\boldsymbol{u}=(x+y\mathrm{i})\boldsymbol{e}_1$ 旋转角 α. 任意复数 $a+b\mathrm{i}$ 可以写成 $r(\cos\alpha+\mathrm{i}\sin\alpha)$ 的形式, 其中 $r=\sqrt{a^2+b^2}$, 用 $a+b\mathrm{i}$ 乘 $x+y\mathrm{i}$, 就是将 $(x+y\mathrm{i})\boldsymbol{e}_1$ 先旋转角 α, 再将长度乘 r.

用 $\cos\alpha+\mathrm{i}\sin\alpha$ 乘 $x+y\mathrm{i}$ 表示将向量 $\boldsymbol{u}=(x+y\mathrm{i})\boldsymbol{e}_1$ 旋转角 α, 用 $(\cos\alpha+\mathrm{i}\sin\alpha)^n$ 乘 $x+y\mathrm{i}$ 就表示将这个旋转动作重复 n 次, 一共旋转 $n\alpha$, 相当于用 $\cos n\alpha+\mathrm{i}\sin n\alpha$ 乘 $x+y\mathrm{i}$. 这就得到

$$(\cos\alpha+\mathrm{i}\sin\alpha)^n=\cos n\alpha+\mathrm{i}\sin n\alpha,$$

这就是著名的**棣莫弗公式** (De Moivre formula).

将平面向量 $\boldsymbol{u}=(x+y\mathrm{i})\boldsymbol{e}_1$ 用坐标 $\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}$ 表示, 则左转 90° 的变换 $\boldsymbol{u}\mapsto\mathrm{i}\boldsymbol{u}$ 是线性变换 $\boldsymbol{X}\mapsto\boldsymbol{J}\boldsymbol{X}$, 它的矩阵 $\boldsymbol{J}$ 的两列分别等于 $\mathrm{i}\boldsymbol{e}_1=\boldsymbol{e}_2$ 与 $\mathrm{i}\boldsymbol{e}_2=-\boldsymbol{e}_1$ 的坐标 $\begin{pmatrix}0\\1\end{pmatrix}$, $\begin{pmatrix}-1\\0\end{pmatrix}$, 因此

$$\boldsymbol{J}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}.$$

易验证 $\boldsymbol{J}^2=-\boldsymbol{I}$. 而 $-\boldsymbol{I}$ 是 -1 引起的变换 $\boldsymbol{u}\mapsto(-1)\boldsymbol{u}$ 的矩阵. $\boldsymbol{J}^2=-\boldsymbol{I}$ 是 $\mathrm{i}^2=-1$ 的“矩阵版本”.

用复数 $a+b\mathrm{i}$ 作乘法引起的变换 $\sigma:x+y\mathrm{i}\mapsto(a+b\mathrm{i})(x+y\mathrm{i})$ 则代表了线性变换

$$\sigma:\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}\mapsto a\boldsymbol{X}+b\boldsymbol{J}\boldsymbol{X}=(a\boldsymbol{I}+b\boldsymbol{J})\boldsymbol{X},$$

它的矩阵

$$a\boldsymbol{I}+b\boldsymbol{J}=a\begin{pmatrix}1&0\\0&1\end{pmatrix}+b\begin{pmatrix}0&-1\\1&0\end{pmatrix}=\begin{pmatrix}a&-b\\b&a\end{pmatrix}.$$

特别地, 旋转角 α 的变换 $x+y\mathrm{i}\mapsto(\cos\alpha+\mathrm{i}\sin\alpha)(x+y\mathrm{i})$ 的矩阵是

$$(\cos\alpha)\boldsymbol{I}+(\sin\alpha)\boldsymbol{J}=\begin{pmatrix}\cos\alpha&-\sin\alpha\\\sin\alpha&\cos\alpha\end{pmatrix}.$$

§4.3 逆 矩 阵

4.3.1 逆矩阵的定义

在数的运算中, 由乘法导出了除法: 已知两数的乘积 b 和其中一个因数 a, 求另一个因数 x 满足方程 $ax=b$, 这样的运算称为除法. 求解办法是: 先求一个数 a^{-1} 满足 $a^{-1}a=1$, 称为 a 的倒数, 也称为 a 的逆. 方程 $ax=b$ 两边同乘 a^{-1} 消去 a, 得到 $x=a^{-1}b$. 代入原方程 $ax=b$ 检验: 左边 $=a(a^{-1}b)=b=$ 右边.$x=a^{-1}b$ 确实是 $ax=b$ 的解.

定义了矩阵乘法之后, 也可提出同样的问题: 已知矩阵 $\boldsymbol{A},\boldsymbol{B}$, 求矩阵 $\boldsymbol{X}$ 满足方程 $\boldsymbol{AX}=\boldsymbol{B}$. 由于矩阵乘法不满足交换律, 还要考虑求另一个矩阵方程 $\boldsymbol{YA}=\boldsymbol{B}$ 的解. 如果能够求出一个矩阵 $\boldsymbol{A}^{-1}$ 满足条件 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$, 同样能够在矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 两边左乘 $\boldsymbol{A}^{-1}$ 得到 $\boldsymbol{A}^{-1}\boldsymbol{AX}=\boldsymbol{A}^{-1}\boldsymbol{B}$, 将 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$ 代入得到 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$, 再由 $\boldsymbol{A}(\boldsymbol{A}^{-1}\boldsymbol{B})=\boldsymbol{B}$(这里利用了 $\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$) 知 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ 确实是 $\boldsymbol{AX}=\boldsymbol{B}$ 的唯一解. 类似地可以得到 $\boldsymbol{YA}=\boldsymbol{B}$ 的唯一解 $\boldsymbol{Y}=\boldsymbol{B}\boldsymbol{A}^{-1}$.

我们从第一章开始就研究线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$. 如果能够找到 $\boldsymbol{A}^{-1}$, 立即可以得到方程组的唯一解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{b}$.

定义 4.3.1 对于矩阵 $\boldsymbol{A}\in F^{n\times n}$, 如果存在矩阵 $\boldsymbol{B}\in F^{n\times n}$ 满足条件 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$, 就称 $\boldsymbol{A}$ **可逆** (invertible), 并且称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ **的逆** (inverse).

例 1 求方阵 $\boldsymbol{A}$ 的逆,

$$\boldsymbol{A}=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}.$$

解 将平面上每个点 P 用它的坐标 $\boldsymbol{X}$(写成列向量) 代表, 则 $\boldsymbol{X}\mapsto\boldsymbol{AX}$ 代表平面上的变换 ρ_α, 将每个点 P 绕原点旋转 α 角变到 P', 再作相反的变换 $\rho_{-\alpha}:\boldsymbol{X}\mapsto\boldsymbol{BX}$ 将每个点 $P'=\rho_\alpha$ 旋转 $-\alpha$ 角回到原来的 P, 则复合变换 $\rho_{-\alpha}\rho_\alpha:\boldsymbol{X}\mapsto\boldsymbol{BAX}$ 使得每个 $\boldsymbol{X}$ 固定不变, 矩阵 $\boldsymbol{BA}=\boldsymbol{I}$. 同理, 先做 $\rho_{-\alpha}$ 后做 ρ_α 的复合变换 $\rho_\alpha\rho_{-\alpha}:\boldsymbol{X}\mapsto\boldsymbol{ABX}$ 也使得每个 $\boldsymbol{X}$ 固定不变, 即 $\boldsymbol{AB}=\boldsymbol{I}$. $\rho_{-\alpha}$ 的矩阵

$$\boldsymbol{B}=\begin{pmatrix}\cos(-\alpha) & -\sin(-\alpha)\\ \sin(-\alpha) & \cos(-\alpha)\end{pmatrix}=\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}$$

就是 $\boldsymbol{A}$ 的逆. □

例 2 求 n 阶方阵 $\boldsymbol{P}$ 的逆,

$$\boldsymbol{P}=\begin{pmatrix}1 & \lambda & \\ 0 & 1 & \\ & & \boldsymbol{I}_{(n-2)}\end{pmatrix}.$$

解 对任意 $n\times m$ 矩阵 $\boldsymbol{A}$, 设它的各行依次为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 则

$$\sigma:\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}\mapsto \boldsymbol{PA}=\begin{pmatrix}1 & \lambda & \\ 0 & 1 & \\ & & \boldsymbol{I}_{(n-2)}\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1+\lambda\boldsymbol{\alpha}_2\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}.$$

可见 $\sigma:\boldsymbol{A}\to\boldsymbol{PA}$ 的效果是将 $\boldsymbol{A}$ 的第 2 行的 λ 倍加到第 1 行, 这是一个初等行变换 $\sigma:\boldsymbol{A}\xrightarrow{\lambda(2)+(1)}\boldsymbol{PA}$. 如果再做初等行变换 $\tau:\boldsymbol{A}\xrightarrow{-\lambda(2)+(1)}\boldsymbol{B}$ 将 σ 的效果抵消, 则每个 $\boldsymbol{PA}$ 被变回 $\boldsymbol{A}$. 在 $\boldsymbol{P}$ 中将 λ 换成 $-\lambda$ 得到

$$\boldsymbol{Q}=\begin{pmatrix}1 & -\lambda & \\ 0 & 1 & \\ & & \boldsymbol{I}_{(n-2)}\end{pmatrix},$$

则 $\tau:\boldsymbol{A}\xrightarrow{-\lambda(2)+(1)}\boldsymbol{QA}$,$\tau\sigma:\boldsymbol{A}\mapsto\boldsymbol{QPA}=\boldsymbol{A}$ 及 $\sigma\tau:\boldsymbol{A}\mapsto\boldsymbol{PQA}=\boldsymbol{A}$ 对每个 $\boldsymbol{A}$ 成立. 特别可以取 $\boldsymbol{A}=\boldsymbol{I}$, 得到 $\boldsymbol{QP}=\boldsymbol{PQ}=\boldsymbol{I}$. 可见 $\boldsymbol{Q}$ 就是 $\boldsymbol{P}$ 的逆. □

4.3.2 逆矩阵的性质

定义 4.3.1 给出了 $\boldsymbol{A}$ 的逆 $\boldsymbol{B}$ 满足的条件 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$, 但还没有说将 $\boldsymbol{A}$ 的逆记为 $\boldsymbol{A}^{-1}$. 这是因为, 还不知道满足条件的矩阵 $\boldsymbol{B}$ 是否唯一. 如果有两个不同的 $\boldsymbol{B}$ 都满足条件, 都是 $\boldsymbol{A}$ 的逆, 我们用 $\boldsymbol{A}^{-1}$ 来代表哪一个呢?

另外还有一个问题:$\boldsymbol{A}$ 的逆 $\boldsymbol{B}$ 要求同时满足两个条件 $\boldsymbol{AB}=\boldsymbol{I}$ 与 $\boldsymbol{BA}=\boldsymbol{I}$. 由于矩阵乘法不满足交换律, 一般来说不能断定 $\boldsymbol{AB}=\boldsymbol{BA}$. 但在 $\boldsymbol{AB}=\boldsymbol{I}$ 这个特殊情形下, 能否断定 $\boldsymbol{BA}=\boldsymbol{I}$ 自动成立?

引理 4.3.1 设 $\boldsymbol{A}$ 是 n 阶方阵, 如果有 n 阶方阵 $\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{I}$, 则 $\boldsymbol{B}$ 是方程 $\boldsymbol{AX}=\boldsymbol{I}$ 的唯一解, 且满足条件 $\boldsymbol{BA}=\boldsymbol{I}$, 因而是 $\boldsymbol{A}$ 的唯一的逆.

证明 将 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{I}$ 按列分块为 $\boldsymbol{A}=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$,$\boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$,$\boldsymbol{I}=(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n)$. 则

$$\boldsymbol{AB}=(\boldsymbol{AB}_1,\cdots,\boldsymbol{AB}_n)=\boldsymbol{I}=(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n)\ \Leftrightarrow\ \boldsymbol{AB}_j=\boldsymbol{e}_j\ (\forall\, 1\leqslant j\leqslant n).$$

$\boldsymbol{AB}$ 的各列 $\boldsymbol{AB}_j$ 都是 $\boldsymbol{A}$ 的各列的线性组合,所组成的向量组 $C(\boldsymbol{AB})$ 的秩不超过 $\boldsymbol{A}$ 的列向量组 $C(\boldsymbol{A})$ 的秩, 即 $\operatorname{rank}\boldsymbol{A} \geqslant \operatorname{rank}(\boldsymbol{AB}) = \operatorname{rank}\boldsymbol{I} = n$. 这迫使 $\operatorname{rank}\boldsymbol{A} = n$. 说明 $\boldsymbol{A}$ 的 n 列线性无关, 组成 $F^{n\times 1}$ 的一组基. 行列式 $|\boldsymbol{A}| \neq 0$. 每个方程组 $\boldsymbol{AX} = \boldsymbol{e}_j (1 \leqslant j \leqslant n)$ 的解都有唯一解 $\boldsymbol{X}_j$, 矩阵方程 $\boldsymbol{AX} = \boldsymbol{I}$ 也有唯一解 $\boldsymbol{X} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n) = \boldsymbol{B}$.

现在证明矩阵方程 $\boldsymbol{YA} = \boldsymbol{I}$ 也有唯一解 $\boldsymbol{Y}_0$, 再证明 $\boldsymbol{Y}_0 = \boldsymbol{B}$.

$\boldsymbol{YA} = \boldsymbol{I} \Leftrightarrow (\boldsymbol{YA})^{\mathrm{T}} = \boldsymbol{I}$ 即 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}^{\mathrm{T}} = \boldsymbol{I}$. 由于 $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}| \neq 0$, 矩阵方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{I}$ 也有唯一解 $\boldsymbol{X}_0$, 因而 $\boldsymbol{YA} = \boldsymbol{I}$ 有唯一解 $\boldsymbol{Y}_0 = \boldsymbol{X}_0^{\mathrm{T}}$. 等式 $\boldsymbol{Y}_0\boldsymbol{A} = \boldsymbol{I}$ 两边右乘 $\boldsymbol{B}$ 得 $\boldsymbol{Y}_0\boldsymbol{AB} = \boldsymbol{B}$, 而 $\boldsymbol{AB} = \boldsymbol{I}$, 因此 $\boldsymbol{Y}_0 = \boldsymbol{B}$. □

引理 4.3.1 证明了: 可逆方阵 $\boldsymbol{A}$ 的逆 $\boldsymbol{B}$ 一定是唯一的, 因此可以记 $\boldsymbol{B}$ 为 $\boldsymbol{A}^{-1}$. 反过来也有 $\boldsymbol{B}^{-1} = \boldsymbol{A}$, 也就是 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.

引理 4.3.1 还给出了 $\boldsymbol{A}$ 可逆的必要条件 $|\boldsymbol{A}| \neq 0$. 这个条件也是充分的:

推论 4.3.1(矩阵可逆的条件) 方阵 $\boldsymbol{A}$ 可逆 $\Leftrightarrow |\boldsymbol{A}| \neq 0$.

证明 引理 4.3.1 已证明 $\boldsymbol{A}$ 可逆 $\Rightarrow |\boldsymbol{A}| \neq 0$.

反过来, $|\boldsymbol{A}| \neq 0 \Rightarrow$ 对每个 $j = 1, 2, \cdots, n$, 方程组 $\boldsymbol{AX} = \boldsymbol{e}_i (1 \leqslant i \leqslant n)$ 有唯一解 $\boldsymbol{X} = \boldsymbol{B}_j$. 以这 n 个方程组 $\boldsymbol{AX} = \boldsymbol{e}_j$ 的解 $\boldsymbol{B}_j$ 为各列排成矩阵 $\boldsymbol{B} = (\boldsymbol{B}_1, \cdots, \boldsymbol{B}_n)$, 则 $\boldsymbol{AB} = \boldsymbol{I}, \boldsymbol{B} = \boldsymbol{A}^{-1}$. □

引理 4.3.2 (逆矩阵的性质) 矩阵的逆具有如下性质:

1. $\boldsymbol{A}$ 可逆 $\Rightarrow$ 它的逆 $\boldsymbol{A}^{-1}$ 可逆, 且 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.
2. 同阶方阵 $\boldsymbol{A}, \boldsymbol{B}$ 可逆 $\Rightarrow$ 它们的乘积 $\boldsymbol{AB}$ 可逆, 且 $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
3. $\boldsymbol{A}$ 可逆 $\Rightarrow$ 它的非零常数倍 $\lambda\boldsymbol{A}$ 可逆, 且 $(\lambda\boldsymbol{A})^{-1} = \lambda^{-1}\boldsymbol{A}^{-1}$.
4. $\boldsymbol{A}$ 可逆 $\Rightarrow$ 它的转置 $\boldsymbol{A}^{\mathrm{T}}$ 可逆, 且 $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$.
5. 方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可逆

$$\Rightarrow \text{准对角阵} \begin{pmatrix} \boldsymbol{A} & \\ & \boldsymbol{B} \end{pmatrix} \text{可逆, 且} \begin{pmatrix} \boldsymbol{A} & \\ & \boldsymbol{B} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & \\ & \boldsymbol{B}^{-1} \end{pmatrix}.$$

证明 所有性质都容易根据逆矩阵的定义直接验证. 我们仅验证性质 2.

$\boldsymbol{A}, \boldsymbol{B}$ 可逆 $\Rightarrow (\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{A}(\boldsymbol{BB}^{-1})\boldsymbol{A}^{-1} = \boldsymbol{AA}^{-1} = \boldsymbol{I}$, 类似地有 $(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})(\boldsymbol{AB}) = \boldsymbol{I}$. 这说明 $\boldsymbol{AB}$ 可逆, 且 $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. □

性质 2 可以推广到任意有限个同阶可逆方阵的乘积:

$$\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k \text{ 可逆 } \Rightarrow (\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)^{-1} = \boldsymbol{A}_k^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}.$$

注 $\boldsymbol{AB}$ 的逆不是 $\boldsymbol{A}^{-1}, \boldsymbol{B}^{-1}$ 按原来顺序相乘, 而是顺序颠倒之后的乘积, 这才能在乘积 $(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})$ 中让 $\boldsymbol{AB}$ 右端的 $\boldsymbol{B}$ 与 $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ 的左端的 $\boldsymbol{B}^{-1}$ 相互抵消, 剩下的 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{-1}$ 再相乘得 $\boldsymbol{I}$. 如果按原来的顺序相乘为 $\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}$, 则在乘积 $(\boldsymbol{AB})(\boldsymbol{A}^{-1}\boldsymbol{B}^{-1})$ 中 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{-1}$ 不能相互抵消, $\boldsymbol{B}$ 与 $\boldsymbol{B}^{-1}$ 也不能相互抵消, 此时只有当 $\boldsymbol{AB} = \boldsymbol{BA}$ 时才有 $(\boldsymbol{AB})(\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}) = (\boldsymbol{BA})(\boldsymbol{A}^{-1}\boldsymbol{B}^{-1}) = \boldsymbol{I}$.

生活中也有类似的例子: 用 $\boldsymbol{A}$ 表示穿袜,$\boldsymbol{B}$ 表示穿鞋, 则 $\boldsymbol{A}^{-1}$ 表示脱袜,$\boldsymbol{B}^{-1}$ 表示脱鞋. 在穿的时候应当先穿袜后穿鞋, 顺序是 $\boldsymbol{AB}$. 脱的顺序就应当反过来: 先脱鞋后脱袜, 就是 $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$, 这也说明 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. 因此, 公式 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ 也称为**穿脱原理**.

4.3.3 逆矩阵的算法

推论 4.3.1 给出了当 $|\boldsymbol{A}|\neq 0$ 时求 $\boldsymbol{A}^{-1}$ 的算法: 解 n 个方程组 $\boldsymbol{AX}=\boldsymbol{e}_j(1\leqslant j\leqslant n)$, 以得到的 n 个解 $\boldsymbol{X}_j$ 为各列排成矩阵 $\boldsymbol{X}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$, 则 $\boldsymbol{AX}=\boldsymbol{I}$.

一般地, 当 $\boldsymbol{A}$ 是可逆方阵时, 解矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 都是同时解若干个线性方程组 $\boldsymbol{AX}=\boldsymbol{b}_j$, 其中 $\boldsymbol{b}_j$ 依次取遍 $\boldsymbol{B}$ 的各列, 以各方程组的唯一解 $\boldsymbol{X}_j$ 为各列排成矩阵 $\boldsymbol{X}$, 则 $\boldsymbol{AX}=\boldsymbol{B}$. 解各方程组 $\boldsymbol{AX}=\boldsymbol{b}_j$ 时需要将各增广矩阵 $(\boldsymbol{A},\boldsymbol{b}_j)$ 经过一系列初等行变换化为 $(\boldsymbol{I},\boldsymbol{X}_j)$ 来得到解 $\boldsymbol{X}_j$. 由于各方程组 $\boldsymbol{AX}=\boldsymbol{b}_j$ 的系数矩阵相同, 对这些不同的增广矩阵所作的初等行变换完全相同, 可以直接将 $\boldsymbol{A},\boldsymbol{B}$ 排在一起成为 $\boldsymbol{M}=(\boldsymbol{A},\boldsymbol{B})$, 作一系列初等行变换将 $\boldsymbol{A}$ 变成 $\boldsymbol{I}$, 则 $\boldsymbol{B}$ 的各列 $\boldsymbol{b}_j$ 经过同样的一系列初等行变换分别变成各方程组 $\boldsymbol{AX}=\boldsymbol{b}_j$ 的唯一解 $\boldsymbol{X}_j$, 从而 $\boldsymbol{B}$ 变成矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$. 特别地, 当 $\boldsymbol{B}=\boldsymbol{I}$ 时得到的就是 $\boldsymbol{A}^{-1}$.

算法 4.3.1(求矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的解) 设方阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|\neq 0$, 将 n 阶方阵 $\boldsymbol{A}$ 与 $n\times m$ 矩阵 $\boldsymbol{B}$ 排成 $n\times(n+m)$ 矩阵 $\boldsymbol{M}=(\boldsymbol{A},\boldsymbol{B})$. 作有限次初等行变换将 $\boldsymbol{M}$ 的前 n 列化成单位矩阵:$(\boldsymbol{A},\boldsymbol{B})\to(\boldsymbol{I},\boldsymbol{X})$, 则 $\boldsymbol{M}$ 的后 m 列变成的 $\boldsymbol{X}$ 就是矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的解 $\boldsymbol{A}^{-1}\boldsymbol{B}$.

特别地, 由 $(\boldsymbol{A},\boldsymbol{I})$ 经过一系列初等行变换得到的 $(\boldsymbol{I},\boldsymbol{X})$ 的后 n 列组成的 $\boldsymbol{X}=\boldsymbol{A}^{-1}$. □

如果要求矩阵方程 $\boldsymbol{XA}=\boldsymbol{B}$ 的解, 先将方程两边转置成为 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}$. 求出 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{B}^{\mathrm{T}}$ 的唯一解 $\boldsymbol{Y}$, 则 $\boldsymbol{X}=\boldsymbol{Y}^{\mathrm{T}}$ 是 $\boldsymbol{XA}=\boldsymbol{B}$ 的唯一解.

例 3 求方阵 $\boldsymbol{A}$ 的逆,

$$\boldsymbol{A}=\begin{pmatrix}1&1&1\\1&2&3\\1&4&9\end{pmatrix}.$$

解

$$(\boldsymbol{A},\boldsymbol{I})=\begin{pmatrix}1&1&1&1&0&0\\1&2&3&0&1&0\\1&4&9&0&0&1\end{pmatrix}\xrightarrow{-(1)+(2),-(1)+(3)}\begin{pmatrix}1&1&1&1&0&0\\0&1&2&-1&1&0\\0&3&8&-1&0&1\end{pmatrix}$$

$$\xrightarrow{-(2)+(1),-3(2)+(3)}\begin{pmatrix}1&0&-1&2&-1&0\\0&1&2&-1&1&0\\0&0&2&2&-3&1\end{pmatrix}\to\begin{pmatrix}1&0&0&3&-\frac{5}{2}&\frac{1}{2}\\0&1&0&-3&4&-1\\0&0&1&1&-\frac{3}{2}&\frac{1}{2}\end{pmatrix},$$

因此

$$A^{-1}=\begin{pmatrix}3&-\frac{5}{2}&\frac{1}{2}\\-3&4&-1\\1&-\frac{3}{2}&\frac{1}{2}\end{pmatrix}.$$

□

例 4 求 n 阶方阵 A 的逆,

$$(1)\ A=\begin{pmatrix}0&1&\cdots&1\\1&0&\cdots&1\\\vdots&\vdots&&\vdots\\1&1&\cdots&0\end{pmatrix};\qquad (2)\ A=\begin{pmatrix}1&1&&\\&1&\ddots&\\&&\ddots&1\\&&&1\end{pmatrix}.$$

解法 1 (1) 将 (A,I) 的第 2 至 n 行加到第 1 行得

$$\begin{pmatrix}n-1&n-1&n-1&\cdots&n-1&1&1&1&\cdots&1\\1&0&1&\cdots&1&0&1&0&\cdots&0\\1&1&0&\cdots&1&0&0&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots\\1&1&1&\cdots&0&0&0&0&\cdots&1\end{pmatrix}.$$

将第 1 行乘 $\dfrac{1}{n-1}$. 得到的矩阵的第 1 行的 -1 倍加到其余各行, 得

$$\begin{pmatrix}1&1&1&\cdots&1&\frac{1}{n-1}&\frac{1}{n-1}&\frac{1}{n-1}&\cdots&\frac{1}{n-1}\\0&-1&0&\cdots&0&-\frac{1}{n-1}&\frac{n-2}{n-1}&-\frac{1}{n-1}&\cdots&-\frac{1}{n-1}\\0&0&-1&\cdots&0&-\frac{1}{n-1}&-\frac{1}{n-1}&\frac{n-2}{n-1}&\cdots&-\frac{1}{n-1}\\\vdots&\vdots&\vdots&&\vdots&\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&-1&-\frac{1}{n-1}&-\frac{1}{n-1}&-\frac{1}{n-1}&\cdots&\frac{n-2}{n-1}\end{pmatrix}.$$

第 2 至 n 行加到第 1 行, 再将第 2 至 n 行都乘 -1, 矩阵的前 n 列就变为单位矩阵, 后面 n 列组成 A^{-1}.(略, 请自己完成答案.)

(2) 对 (A,I) 作一系列初等行变换将前面 n 列的 A 变成单位矩阵, 则后面 n 列组成的方阵即为 A^{-1}.(具体过程略, 请自己完成答案.)

解法 2 (1) 令

$$B = A + I = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

则 $B^2 = nB$(参见 §4.1 例 7), 即 $(A+I)^2 = n(A+I)$. 整理得

$$A^2 + (2-n)A = (n-1)I, \text{即} \quad A \cdot \left(\frac{1}{n-1}A + \frac{2-n}{n-1}I \right) = I,$$

这说明

$$A^{-1} = \frac{1}{n-1}A + \frac{2-n}{n-1}I = \begin{pmatrix} \frac{2-n}{n-1} & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \frac{1}{n-1} & \frac{2-n}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n-1} & \frac{1}{n-1} & \cdots & \frac{2-n}{n-1} \end{pmatrix}.$$

(2) 取

$$N = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$$

则

$$N^k = \begin{pmatrix} & I_{(n-k)} \\ O_{(k)} & \end{pmatrix} \ (\forall\ 1 \leqslant k \leqslant n-1), \ N^n = O.$$

我们有 $A = I + N$. 由

$$(I+N)(I - N + N^2 - \cdots + (-1)^{n-1}N^{n-1}) = I - (-N)^n = I$$

知

$$A^{-1} = I - N + N^2 - \cdots + (-1)^{n-1}N^{n-1}$$

$$= \begin{pmatrix} 1 & -1 & 1 & \cdots & (-1)^{n-1} \\ & 1 & -1 & \cdots & (-1)^{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix}. \qquad \square$$

注 例 4(1) 解法 2 的推广: 一般地, 如果已经知道 $\boldsymbol{A}$ 满足一个多项式方程

$$a_m\boldsymbol{A}^m+a_{m-1}\boldsymbol{A}^{m-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}=\boldsymbol{O},$$

且常数项 $a_0\neq 0$. 则

$$\begin{aligned}\boldsymbol{A}\cdot a_0^{-1}(-a_m\boldsymbol{A}^{m-1}-a_{m-1}\boldsymbol{A}^{m-2}-\cdots-a_1\boldsymbol{I})&=\boldsymbol{I},\\ \boldsymbol{A}^{-1}&=a_0^{-1}(-a_m\boldsymbol{A}^{m-1}-a_{m-1}\boldsymbol{A}^{m-2}-\cdots-a_1\boldsymbol{I}).\end{aligned}$$

例 4(2) 的答案是根据 $(1+x)^{-1}$ 的泰勒 (Taylor) 展开式

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

得出来的. 这个泰勒展开式既可以看成牛顿二项式定理的展开式

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

当 $n=-1$ 时的特殊情形. 也可以看成无穷递缩等比数列求和公式

$$1+(-x)+(-x)^2+\cdots=\frac{1}{1+x}.$$

例 5 求矩阵方程的解 $\boldsymbol{X}$:

(1) $\begin{pmatrix}2&5\\1&3\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}4&-1\\2&1\end{pmatrix}$; (2) $\boldsymbol{X}\begin{pmatrix}1&-2&0\\2&1&3\\0&2&1\end{pmatrix}=\begin{pmatrix}1&-1&1\\2&-3&1\end{pmatrix}$.

解 (1) 通过一系列初等行变换将

$$\begin{pmatrix}2&5&4&-1\\1&3&2&1\end{pmatrix}\to\begin{pmatrix}1&0&2&-8\\0&1&0&3\end{pmatrix}.\quad 得\quad \boldsymbol{X}=\begin{pmatrix}2&-8\\0&3\end{pmatrix}.$$

(2) 将矩阵方程 $\boldsymbol{XA}=\boldsymbol{B}$ 两边同时转置, 变成 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}$, 先求 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{B}^{\mathrm{T}}$ 的解.

通过一系列初等行变换得

$$(\boldsymbol{A}^{\mathrm{T}},\boldsymbol{B}^{\mathrm{T}})=\begin{pmatrix}1&2&0&1&2\\-2&1&2&-1&-3\\0&3&1&1&1\end{pmatrix}\to\begin{pmatrix}1&0&0&-1&0\\0&1&0&1&1\\0&0&1&-2&-2\end{pmatrix}.$$

于是得

$$\boldsymbol{Y}=\begin{pmatrix}-1&0\\1&1\\-2&-2\end{pmatrix},\quad \boldsymbol{X}=\boldsymbol{Y}^{\mathrm{T}}=\begin{pmatrix}-1&1&-2\\0&1&-2\end{pmatrix}.$$ □

4.3.4 矩阵求逆公式

例 6 设方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的行列式 $|\boldsymbol{A}|\neq 0$. 求 $\boldsymbol{A}^{-1}$.

分析 $\boldsymbol{A}^{-1}$ 是矩阵方程 $\boldsymbol{AX}=\boldsymbol{I}$ 的唯一解 $\boldsymbol{X}$. 设 $\boldsymbol{A}$ 的第 i 行为 $\boldsymbol{\alpha}_i$,$\boldsymbol{X}$ 的第 j 列为 $\boldsymbol{X}_j$. 要使

$$\boldsymbol{AX}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)=\begin{pmatrix}\boldsymbol{\alpha}_1\boldsymbol{X}_1 & \cdots & \boldsymbol{\alpha}_1\boldsymbol{X}_n\\ \vdots & & \vdots\\ \boldsymbol{\alpha}_n\boldsymbol{X}_1 & \cdots & \boldsymbol{\alpha}_n\boldsymbol{X}_n\end{pmatrix}=\boldsymbol{I}=\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix},$$

应选择 $\boldsymbol{X}$ 的各列 $\boldsymbol{X}_j$ 使所有的非对角元 $\boldsymbol{\alpha}_i\boldsymbol{X}_j=0$(当 $i\neq j$), 对角元 $\boldsymbol{\alpha}_i\boldsymbol{X}_i=1(1\leqslant i\leqslant n)$.

先设法寻找 $\boldsymbol{X}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$ 使 $\boldsymbol{\alpha}_i\boldsymbol{X}_j=0(\forall\, i\neq j)$, 对角元 $\boldsymbol{\alpha}_i\boldsymbol{X}_i=\lambda_i\neq 0$. 则 $\boldsymbol{AX}=\boldsymbol{D}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 是可逆对角矩阵. 再由 $\boldsymbol{A}^{-1}=\boldsymbol{XD}^{-1}$ 得到 $\boldsymbol{A}^{-1}$.

已知 n 个行向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$, 要寻找列向量 $\boldsymbol{X}_i$ 与其中的 $n-1$ 个 $\boldsymbol{\alpha}_k(k\neq i)$ 的乘积 $\boldsymbol{\alpha}_k\boldsymbol{X}_i$ 都是 0, 只有一个乘积 $\boldsymbol{\alpha}_i\boldsymbol{X}_i\neq 0$.(可以看成: 与 $n-1$ 个 $\boldsymbol{\alpha}_k(k\neq i)$ 都垂直, 只与 $\boldsymbol{\alpha}_i$ 不垂直.) 我们想到了由 $\boldsymbol{\alpha}_i=(a_{i1},\cdots,a_{in})$ 各元素的代数余子式 A_{ij} 组成的行向量 $\boldsymbol{A}_i=(A_{i1},\cdots,A_{in})$, 它与每个 $\boldsymbol{\alpha}_k$ 的"内积"

$$\boldsymbol{\alpha}_k\boldsymbol{A}_i^{\mathrm{T}}=a_{k1}A_{i1}+\cdots+a_{kn}A_{in}=\begin{cases}0, & \forall\, k\neq i,\\ |\boldsymbol{A}|, & k=i.\end{cases}$$

取 $\boldsymbol{X}=(\boldsymbol{A}_1^{\mathrm{T}},\cdots,\boldsymbol{A}_n^{\mathrm{T}})$, 则 $\boldsymbol{AX}=|\boldsymbol{A}|\boldsymbol{I}$ 是纯量矩阵. 当行列式 $|\boldsymbol{A}|\neq 0$ 时就得到 $\boldsymbol{A}^{-1}=|\boldsymbol{A}|^{-1}\boldsymbol{X}$.

解 取 $\boldsymbol{A}$ 的每一行 $\boldsymbol{\alpha}_i=(a_{i1},\cdots,a_{in})$ 中各元素在行列式 $|\boldsymbol{A}|$ 中的代数余子式 A_{ij} 组成列向量 $\boldsymbol{A}_i'=(A_{i1},\cdots,A_{in})^{\mathrm{T}}$, 依次以这些列向量为各列组成方阵 $\boldsymbol{A}^*=(\boldsymbol{A}_1',\cdots,\boldsymbol{A}_n')$. 则

$$\boldsymbol{AA}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{I},$$

当 $|\boldsymbol{A}|\neq 0$ 时 $\boldsymbol{A}^{-1}=|\boldsymbol{A}|^{-1}\boldsymbol{A}^*$. □

例 6 中得到的矩阵

$$\boldsymbol{A}^*=\begin{pmatrix}A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix}$$

称为 A 的**伴随方阵** (adjoint matrix), 它的第 (i,j) 元等于 A 的第 (j,i) 元的代数余子式 A_{ji}, 它的第 i 列由 A 的第 i 行各元素的代数余子式组成.

我们将例 6 的结论作为一个引理, 以便应用:

引理 4.3.3(矩阵求逆公式)　$\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A} = |\boldsymbol{A}|\boldsymbol{I}$. 当 $|\boldsymbol{A}| \neq 0$ 时,$\boldsymbol{A}^{-1} = |\boldsymbol{A}|^{-1}\boldsymbol{A}^*$.

例 7　求 F^2 中的线性变换 σ 将 $\boldsymbol{\alpha}_1 = (2,1), \boldsymbol{\alpha}_2 = (5,3)$ 分别映到 $\boldsymbol{\beta}_1 = (4,2), \boldsymbol{\beta}_2 = (-1,1)$.

解　将 F^2 中的向量都写成列向量的形式, 则线性变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{A}\boldsymbol{X}$ 由某个方阵 $\boldsymbol{A}$ 左乘 $\boldsymbol{X}$ 实现. 本题也就是求方阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}\boldsymbol{\alpha}_1^{\mathrm{T}} = \boldsymbol{\beta}_1, \quad \boldsymbol{A}\boldsymbol{\alpha}_2^{\mathrm{T}} = \boldsymbol{\beta}_2^{\mathrm{T}}, \qquad \text{即} \qquad \boldsymbol{A}(\boldsymbol{\alpha}_1^{\mathrm{T}}, \boldsymbol{\alpha}_2^{\mathrm{T}}) = (\boldsymbol{\beta}_1^{\mathrm{T}}, \boldsymbol{\beta}_2^{\mathrm{T}}),$$

$$\boldsymbol{A}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}. \tag{4.3.1}$$

等式两边右乘

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} = \frac{1}{2\times 3 - 1\times 5}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix},$$

得

$$\boldsymbol{A} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 13 & -22 \\ 5 & -8 \end{pmatrix}.$$

所求线性变换为

$$\sigma: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 13 & -22 \\ 5 & -8 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \qquad \square$$

例 7 中的 (4.3.1) 仍是矩阵方程, 可以采用例 5(2) 的方法, 将等式两边转置之后用初等行变换来解. 例 7 中直接将矩阵求逆公式代入, 得到了答案. 例 5(1) 中也可以直接将等式两边左乘 $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}^{-1}$ 得到答案. 但是, 在实际计算中, 矩阵求逆公式也只有对 2 阶方阵求逆比较简便, 对更高阶的方阵, 还是用初等变换求逆更简便些. 矩阵求逆公式更多的是用于理论分析.

例 8　设 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{b}$ 是 n 个方程组成的 n 元线性方程组, 系数行列式 $|\boldsymbol{A}| \neq 0$. 试由矩阵求逆公式得出由各方程系数求唯一解 $\boldsymbol{X}$ 的公式.

解 $\Delta=|\boldsymbol{A}|\neq 0\Rightarrow\boldsymbol{A}$ 可逆 $\Rightarrow\quad \boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{b}$. 将矩阵求逆公式 $\boldsymbol{A}^{-1}=\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$ 代入得

$$\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=|\boldsymbol{A}|^{-1}\boldsymbol{A}^*\boldsymbol{b}=|\boldsymbol{A}|^{-1}\begin{pmatrix}b_1A_{11}+\cdots+b_nA_{n1}\\ \vdots\\ b_1A_{1n}+\cdots+b_nA_{nn}\end{pmatrix},$$

$\boldsymbol{X}$ 的第 i 分量

$$x_i=\frac{1}{|\boldsymbol{A}|}(b_1A_{1i}+\cdots+b_nA_{ni}),$$

其中

$$b_1A_{1i}+\cdots+b_nA_{ni}=\det(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{i-1},\boldsymbol{b},\boldsymbol{a}_{i+1},\cdots,\boldsymbol{a}_n)=\Delta_i,$$

就是将行列式 $\Delta=|\boldsymbol{A}|$ 的第 i 列换成 $\boldsymbol{b}$ 得到的 Δ_i 按第 i 列的展开式. 因此 $x_i=\dfrac{\Delta_i}{\Delta}$. 方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的解为

$$\boldsymbol{X}=\left(\frac{\Delta_1}{\Delta},\cdots,\frac{\Delta_n}{\Delta}\right)^{\mathrm{T}},$$

这就是 §3.3 中讲过的克拉默法则. □

习题 4.3

1. 求下列矩阵的逆矩阵.

(1) $\boldsymbol{A}=\begin{pmatrix}0&0&-4\\-1&0&0\\0&2&0\end{pmatrix}$; (2) $\boldsymbol{A}=\begin{pmatrix}1&5&3&0\\0&4&6&2\\0&0&9&1\\0&0&0&1\end{pmatrix}$;

(3) $\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$; (4) $\boldsymbol{A}=\begin{pmatrix}1&2&4&8\\0&1&2&4\\0&0&1&2\\0&0&0&1\end{pmatrix}$.

2. 设 $\boldsymbol{A}$ 是方阵,$\boldsymbol{A}^k=\boldsymbol{O}$ 对某个正整数 k 成立. 求证下列方阵可逆, 并分别求它们的逆.

(1) $\boldsymbol{I}-\boldsymbol{A}$; (2) $\boldsymbol{I}+\boldsymbol{A}$; (3) $\boldsymbol{I}+\boldsymbol{A}+\dfrac{1}{2!}\boldsymbol{A}^2+\cdots+\dfrac{1}{(k-1)!}\boldsymbol{A}^{k-1}$.

3. 求下式中的矩阵 $\boldsymbol{X}$:

(1) $\begin{pmatrix}7&3\\2&1\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}4&5\\3&1\end{pmatrix}$; (2) $\boldsymbol{X}\begin{pmatrix}1&-2&0\\2&1&3\\0&2&1\end{pmatrix}=\begin{pmatrix}1&-1&1\\2&-3&1\\3&-4&1\end{pmatrix}$.

4. 设 $\boldsymbol{A}$ 是 n 阶方阵, 证明: 若 $\boldsymbol{A}^2=\boldsymbol{I}$, 且 $\boldsymbol{A}\neq\boldsymbol{I}$, 则 $\boldsymbol{A}+\boldsymbol{I}$ 不是可逆矩阵.

5. 在以下每个小题中, 是否存在 $\mathbf{R}^2$ 上的线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 满足所说的要求? 如果存在, 求出变换矩阵 $\boldsymbol{A}$.

(1) $\sigma(1,2)=(2,3),\ \sigma(2,5)=(5,8)$;

(2) $\sigma(1,1)=(2,3),\ \sigma(2,3)=(3,5),\ \sigma(3,5)=(2,3)$;

(3) $\sigma(1,1)=(2,3),\ \sigma(2,3)=(7,7),\ \sigma(1,-1)=(0,1)$.

6. 设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$ 组成 F^n 的一组基,$\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n$ 是 F^n 中任意 n 个向量. 求证: 存在唯一的线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 满足 $\sigma(\boldsymbol{a}_i)=\boldsymbol{b}_i,\forall\ 1\leqslant i\leqslant n$.

§4.4 初等方阵及应用

4.4.1 初等变换与初等方阵

矩阵的初等行变换 $\sigma:\boldsymbol{A}\mapsto\boldsymbol{B}$ 将 $\boldsymbol{A}$ 的行向量组 $R(\boldsymbol{A})=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 变到 $\boldsymbol{B}$ 的行向量组 $R(\boldsymbol{B})=\{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_m\}$,$\boldsymbol{B}$ 的每一行 $\boldsymbol{\beta}_i$ 都是 $\boldsymbol{A}$ 的行的线性组合:

$$\boldsymbol{\beta}_i=p_{i1}\boldsymbol{\alpha}_1+\cdots+p_{im}\boldsymbol{\alpha}_m=(p_{i1},\cdots,p_{im})\begin{pmatrix}\boldsymbol{\alpha}_1\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}=\boldsymbol{\pi}_i\boldsymbol{A},$$

因此

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{\pi}_1\\ \vdots\\ \boldsymbol{\pi}_m\end{pmatrix}\boldsymbol{A}=\boldsymbol{PA}$$

可由方阵 $\boldsymbol{P}$ 左乘 $\boldsymbol{A}$ 得到,$\boldsymbol{P}$ 以各个 $\boldsymbol{\pi}_i$ 为各行排列而成.

例如, 设 $\sigma:\boldsymbol{A}\xrightarrow{\lambda(2)+(1)}\boldsymbol{B}$ 是将矩阵的第 2 行的 λ 倍加到第 1 行的初等行变换, 则

$$\sigma:\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}\xrightarrow{\lambda(2)+(1)}\begin{pmatrix}\boldsymbol{\alpha}_1+\lambda\boldsymbol{\alpha}_2\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}=\begin{pmatrix}1&\lambda&\\ 0&1&\\ &&\boldsymbol{I}_{(n-2)}\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_n\end{pmatrix}.$$

一般地, 每个初等行变换 $\sigma:\boldsymbol{A}\to\boldsymbol{B}$ 可以看成向量空间 $F^{n\times m}$ 上的一个线性变换, 通过左乘某个 n 阶方阵 $\boldsymbol{P}$ 实现:$\sigma(\boldsymbol{A})=\boldsymbol{PA}$. 为了求出这个方阵 $\boldsymbol{P}$, 只要将 σ 看成 $F^{n\times n}$ 上的变换, 将 σ 作用于单位矩阵就得到 $\sigma(\boldsymbol{I})=\boldsymbol{PI}=\boldsymbol{P}$. 也

就是说: 将 n 阶单位矩阵 $\boldsymbol{I}$ 经过初等行变换 σ 得到方阵 $\boldsymbol{P}$, 则 $\sigma(\boldsymbol{A})=\boldsymbol{PA}$ 对所有的 $n\times m$ 矩阵成立. 例如:

$$\sigma:\boldsymbol{I}=\begin{pmatrix}1 & 0 & \\ 0 & 1 & \\ & & \boldsymbol{I}\end{pmatrix}\xrightarrow{\lambda(2)+(1)}=\begin{pmatrix}1 & \lambda & \\ 0 & 1 & \\ & & \boldsymbol{I}\end{pmatrix}=\boldsymbol{P},$$

得到的 $\boldsymbol{P}$ 就是初等行变换 $\sigma:\boldsymbol{A}\xrightarrow{\lambda(2)+(1)}\boldsymbol{PA}$ 的矩阵 $\boldsymbol{P}$.

如果要作相应的初等列变换 $\sigma':\boldsymbol{A}\xrightarrow[\lambda(2)+(1)]{}\boldsymbol{B}$, 将 $\boldsymbol{A}$ 的第 2 列的 λ 倍加到第 1 列, 只要先对 A 的转置 $\boldsymbol{A}^{\mathrm{T}}$ 作初等行变换 $\sigma:\boldsymbol{A}^{\mathrm{T}}\xrightarrow{\lambda(2)+(1)}\boldsymbol{PA}^{\mathrm{T}}$, 再将 $\boldsymbol{PA}^{\mathrm{T}}$ 转置回来得到的 $(\boldsymbol{PA}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{AP}^{\mathrm{T}}$ 就是 $\sigma'(\boldsymbol{A})$. 可见 $\sigma':\boldsymbol{A}\xrightarrow[\lambda(2)+(1)]{}\boldsymbol{AP}^{\mathrm{T}}$ 可以用矩阵

$$\boldsymbol{P}^{\mathrm{T}}=\begin{pmatrix}1 & 0 & \\ \lambda & 1 & \\ & & \boldsymbol{I}\end{pmatrix}$$

右乘每个 $\boldsymbol{A}$ 实现.

以上结论总结为

定理 4.4.1 对 $n\times m$ 矩阵 $\boldsymbol{A}$ 进行的任何一个初等行变换 $\sigma:\boldsymbol{A}\to\boldsymbol{B}$ 都可以通过左乘某个 n 阶方阵 $\boldsymbol{P}$ 实现:$\sigma:\boldsymbol{A}\to\boldsymbol{PA}$, 其中 $\boldsymbol{P}=\sigma(\boldsymbol{I})$ 可以由 n 阶单位矩阵 $\boldsymbol{I}$ 经过同样的初等行变换 σ 得到.

对 $m\times n$ 矩阵 $\boldsymbol{A}$ 进行的任何一个初等列变换 $\sigma':\boldsymbol{A}\to\boldsymbol{B}$ 可以通过右乘某个 n 阶方阵 $\boldsymbol{P}'$ 实现: $\sigma:\boldsymbol{A}\to\boldsymbol{AP}'$, 其中 $\boldsymbol{P}'=\sigma'(\boldsymbol{I})$ 可以由 n 阶单位矩阵 $\boldsymbol{I}$ 经过同样的初等列变换 σ' 得到, 也可以由相应的初等行变换 σ 的方阵 $\boldsymbol{P}$ 转置得到:$\boldsymbol{P}'=\boldsymbol{P}^{\mathrm{T}}$. □

例 1 求实现如下初等行变换 $\sigma:\boldsymbol{A}\to\boldsymbol{PA}$ 的方阵 $\boldsymbol{P}$.

(1) $\sigma:\boldsymbol{A}\xrightarrow{(1,2)}\boldsymbol{B}$; (2) $\sigma:\boldsymbol{A}\xrightarrow{\lambda(1)}\boldsymbol{B}$.

解 (1)

$$\boldsymbol{I}=\begin{pmatrix}1 & 0 & \\ 0 & 1 & \\ & & \boldsymbol{I}\end{pmatrix}\xrightarrow{(1,2)}\begin{pmatrix}0 & 1 & \\ 1 & 0 & \\ & & \boldsymbol{I}\end{pmatrix}=\boldsymbol{P};$$

(2)

$$\boldsymbol{I}=\begin{pmatrix}1 & 0 & \\ 0 & 1 & \\ & & \boldsymbol{I}\end{pmatrix}\xrightarrow{\lambda(1)}\begin{pmatrix}\lambda & 0 & \\ 0 & 1 & \\ & & \boldsymbol{I}\end{pmatrix}=\boldsymbol{P}.$$

□

定义 4.4.1(初等方阵) 由单位矩阵经过一次初等行变换得到的方阵称为**初等方阵** (elementary matrix). □

引理 4.4.1 初等方阵包括如下三种类型:

(1) 对 $1 \leqslant i < j \leqslant n$, 将 n 阶单位矩阵 $\boldsymbol{I}_{(n)}$ 的第 i,j 两行互换得到的方阵

$$\begin{aligned} \boldsymbol{P}_{ij} &= \boldsymbol{I} - \boldsymbol{E}_{ii} - \boldsymbol{E}_{jj} + \boldsymbol{E}_{ij} + \boldsymbol{E}_{ji} \\ &= \begin{array}{c} \begin{pmatrix} \boldsymbol{I}_{(i-1)} & & & & \\ & 0 & & 1 & \\ & & \boldsymbol{I}_{(j-i-1)} & & \\ & 1 & & 0 & \\ & & & & \boldsymbol{I}_{(n-j)} \end{pmatrix} \end{array} \begin{array}{l} \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行}. \\ \\ \end{array} \end{aligned}$$

(2) 对 $1 \leqslant i \leqslant n$,$\lambda \neq 0$ 且 $\lambda \neq 1$, 将 n 阶单位矩阵 $\boldsymbol{I}_{(n)}$ 的第 i 行乘 λ 得到的方阵

$$\begin{aligned} \boldsymbol{D}_i(\lambda) &= \boldsymbol{I} + (\lambda - 1)\boldsymbol{E}_{ii} \\ &= \begin{pmatrix} \boldsymbol{I}_{(i-1)} & & \\ & \lambda & \\ & & \boldsymbol{I}_{(n-i)} \end{pmatrix} \begin{array}{l} \\ \leftarrow \text{第 } i \text{ 行}. \\ \\ \end{array} \end{aligned}$$

(3) 对 $1 \leqslant i,j \leqslant n$,$i \neq j$,$\lambda \neq 0$, 将 n 阶单位矩阵 $\boldsymbol{I}_{(n)}$ 的第 j 行的 λ 倍加到第 i 行得到的方阵

$$\begin{aligned} \boldsymbol{T}_{ij}(\lambda) &= \boldsymbol{I} + \lambda\boldsymbol{E}_{ij} \\ &= \begin{pmatrix} \boldsymbol{I}_{(i-1)} & & & & \\ & 1 & & \lambda & \\ & & \boldsymbol{I}_{(j-i-1)} & & \\ & & & 1 & \\ & & & & \boldsymbol{I}_{(n-j)} \end{pmatrix} \begin{array}{l} \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行}. \\ \\ \end{array} \end{aligned}$$

□

推论 4.4.1 用各初等方阵左乘矩阵实现如下初等行变换:

(1) $\boldsymbol{A} \xrightarrow{(i,j)} \boldsymbol{P}_{ij}\boldsymbol{A}$. (2) $\boldsymbol{A} \xrightarrow{\lambda(i)} \boldsymbol{D}_i(\lambda)\boldsymbol{A}$. (3) $\boldsymbol{A} \xrightarrow{\lambda(j)+(i)} \boldsymbol{T}_{ij}(\lambda)\boldsymbol{A}$.

用各初等方阵右乘矩阵实现如下初等列变换:

(1) $\boldsymbol{A} \xrightarrow[(i,j)]{} \boldsymbol{A}\boldsymbol{P}_{ij}$. (2) $\boldsymbol{A} \xrightarrow[\lambda(i)]{} \boldsymbol{A}\boldsymbol{D}_i(\lambda)$. (3) $\boldsymbol{A} \xrightarrow[\lambda(j)+(i)]{} \boldsymbol{A}\boldsymbol{T}_{ji}(\lambda)$. □

注 实现初等列变换 $\sigma' : \boldsymbol{A} \mapsto \boldsymbol{A}\boldsymbol{P}$ 的初等方阵 $\boldsymbol{P}^{\mathrm{T}}$ 由相应的初等行变换 $\sigma : \boldsymbol{A} \to \boldsymbol{P}\boldsymbol{A}$ 的方阵 $\boldsymbol{P}$ 转置得到. 前两类初等方阵 $\boldsymbol{P}_{ij}$ 与 $\boldsymbol{D}_i(\lambda)$ 是对称方阵, 转置之后等于自己. 第 3 类初等方阵的转置 $\boldsymbol{T}_{ij}(\lambda)^{\mathrm{T}} = \boldsymbol{T}_{ji}(\lambda)$ 不等于自己.

4.4.2 行列式的乘法性质

例 2 求椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$(其中 $a>b>0$) 的面积. 并求出它的内接四边形的最大面积.

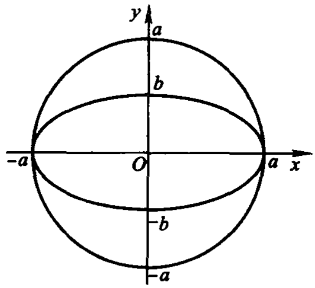

图 4–4

分析 将椭圆在 y 轴方向"拉长" $\dfrac{a}{b}$ 倍得到圆 $x^2+y^2=a^2$. 反过来, 将圆 $x^2+y^2=a^2$ 在 y 轴方向按比例 $\dfrac{b}{a}$ "压缩"得到所说的椭圆.

在 y 轴方向上的"压缩"可以通过平面上的线性变换 $(x,y)\mapsto\left(x,\dfrac{b}{a}y\right)$ 来实现. 经过这个压缩变换, 所有图形的面积都被压缩了同样的比例, 变成原来的 $\dfrac{b}{a}$ 倍.

圆面积 πa^2 "压缩"成椭圆面积 $\dfrac{b}{a}\cdot\pi a^2=\pi ab$.

圆内接四边形最大面积 (内接正方形面积)$2a^2$ 压缩成椭圆内接四边形最大面积 $\dfrac{b}{a}\cdot 2a^2=2ab$.

(解答叙述过程略.) □

一般地, 平面上的线性变换 $\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 将所有的图形的面积放大或缩小同一个倍数 $\lambda=|\det\boldsymbol{A}|$, 等于变换矩阵 $\boldsymbol{A}$ 的行列式 $\det\boldsymbol{A}$ 的绝对值. 行列式 $\det\boldsymbol{A}$ 的正负号表示是保持还是改变图形旋转方向 (逆时针方向还是顺时针方向). 例如, 本节例 2 中将圆压缩成椭圆的变换 $(x,y)\mapsto\left(x,\dfrac{b}{a}\right)$ 的矩阵 $\boldsymbol{A}=\operatorname{diag}\left(1,\dfrac{b}{a}\right)$, 行列式 $\det\boldsymbol{A}=\dfrac{b}{a}$, 图形面积都被压缩成原来的 $\dfrac{b}{a}$, 旋转方向不变.§4.2 例 2 的旋转变换的矩阵 $\boldsymbol{A}$ 的行列式 $\det\boldsymbol{A}=1$, 该变换保持图形的

形状大小不变, 面积不变, 旋转方向也不变.§4.2 例 4 的轴对称变换的矩阵 $\boldsymbol{A}$ 的行列式 $\det \boldsymbol{A}=-1$, 该变换也保持面积不变, 但旋转方向改变, 从 $\boldsymbol{e}_1$ 到 $\boldsymbol{e}_2$ 是逆时针方向旋转 $90°$, 变换之后的 $\boldsymbol{e}_1'$ 到 $\boldsymbol{e}_2'$ 则是顺时针方向旋转 $90°$.

类似地, 三维几何空间的线性变换 $\boldsymbol{X}\mapsto \boldsymbol{AX}$ 将图形的体积放大或缩小同一个倍数 $\det \boldsymbol{A}, \det \boldsymbol{A}$ 的符号表示是保持还是改变图形的旋转方向 (右手系还是左手系).

如果先经过变换 $\sigma:\boldsymbol{X}\mapsto \boldsymbol{AX}$ 将面积或体积放大或缩小到 $\lambda=\det \boldsymbol{A}$ 倍, 再经过变换 $\tau:\boldsymbol{X}\mapsto \boldsymbol{BA}$ 放大或缩小到 $\mu=\det \boldsymbol{B}$ 倍, 复合变换 $\tau\sigma:\boldsymbol{X}\mapsto (\boldsymbol{BA})\boldsymbol{X}$ 的放缩倍数就应当是

$$\det(\boldsymbol{BA})=\mu\lambda=(\det \boldsymbol{A})(\det \boldsymbol{B}).$$

一般地, 有:

定理 4.4.2 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶方阵, 则 $\det(\boldsymbol{AB})=(\det \boldsymbol{A})(\det \boldsymbol{B})$.

证明 按引理 1.3.1 所说, 方阵 $\boldsymbol{A}$ 可以经过有限次第三类初等行变换变成阶梯形矩阵 $\boldsymbol{D}$, 其中阶梯元所在列的其余元素都为 0. 第三类初等行变换不改变行列式, 因此 $\det \boldsymbol{A}=\det \boldsymbol{D}$.

每个第三类初等行变换都可以通过左乘第三类初等方阵 $\boldsymbol{P}_i$ 实现. 因此,$\boldsymbol{D}$ 可由 $\boldsymbol{A}$ 左乘有限个第三类初等方阵 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_s$ 得到:$\boldsymbol{D}=\boldsymbol{P}_s\cdots\boldsymbol{P}_1\boldsymbol{A}$. 于是 $\boldsymbol{DB}=(\boldsymbol{P}_s\cdots\boldsymbol{P}_1\boldsymbol{A})\boldsymbol{B}=\boldsymbol{P}_s\cdots\boldsymbol{P}_1(\boldsymbol{AB})$ 由 $\boldsymbol{AB}$ 左乘有限个第三类初等方阵 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_s$ 得到, 也就是经过有限次第三类初等行变换得到, 因此 $\det(\boldsymbol{AB})=\det(\boldsymbol{DB})$.

阶梯形方阵 $\boldsymbol{D}$ 的阶梯元的个数等于非零行的个数, 也就是 $\operatorname{rank}\boldsymbol{A}$. 如果 $\det \boldsymbol{A}\neq 0$, 则 $\boldsymbol{D}$ 的非零行个数 $\operatorname{rank}\boldsymbol{D}=\operatorname{rank}\boldsymbol{A}=n$. $\boldsymbol{D}$ 是对角元全不为 0 的对角阵 $\boldsymbol{D}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n), \det \boldsymbol{A}=\det \boldsymbol{D}=\lambda_1\cdots\lambda_n$. $\boldsymbol{DB}$ 由 $\boldsymbol{B}$ 的各行分别乘 $\lambda_1,\cdots,\lambda_n$ 得到, 将 $\boldsymbol{DB}$ 的各行分别提出公因子 $\lambda_1,\cdots,\lambda_n$ 仍得到 $\boldsymbol{B}$, 由此得到

$$\det(\boldsymbol{AB})=\det(\boldsymbol{DB})=\lambda_1\cdots\lambda_n\det \boldsymbol{B}=(\det \boldsymbol{A})(\det \boldsymbol{B}).$$

当 $\det \boldsymbol{A}=0$ 时,$\boldsymbol{D}$ 的非零行的个数 $\operatorname{rank}\boldsymbol{D}=\operatorname{rank}\boldsymbol{A}<n$,$\boldsymbol{D}$ 的最后一行一定是零,$\boldsymbol{DB}$ 的最后一行也是零. 此时 $\det \boldsymbol{D}=\det(\boldsymbol{DB})=0$.

$$\det(\boldsymbol{AB})=\det(\boldsymbol{DB})=0=0\det \boldsymbol{B}=(\det \boldsymbol{A})(\det \boldsymbol{B}). \qquad \square$$

例 3 求证: (1) 设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶方阵, 则 $|\boldsymbol{AB}|=|\boldsymbol{BA}|$.

(2) 设 $\boldsymbol{A}\in F^{m\times n},\boldsymbol{B}\in F^{n\times m}$, 如果 $m>n$, 则 $|\boldsymbol{AB}|=0$, 但 $|\boldsymbol{BA}|$ 不一定为 0.

(3) 设 $\boldsymbol{A}$ 是可逆方阵, 则 $|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$.

(4) 设方阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{-1}$ 的元素都是整数, 则 $|\boldsymbol{A}|=\pm 1$.

证明 (1) $|\boldsymbol{AB}|=|\boldsymbol{A}|\,|\boldsymbol{B}|=|\boldsymbol{B}|\,|\boldsymbol{A}|=|\boldsymbol{BA}|$.

(交换律对矩阵乘法不成立,$\boldsymbol{AB}$ 不一定等于 $\boldsymbol{BA}$. 但行列式 $|\boldsymbol{A}|,|\boldsymbol{B}|$ 是数, 数的乘法一定可交换,$|\boldsymbol{A}||\boldsymbol{B}|=|\boldsymbol{B}||\boldsymbol{A}|$.)

(2) 将 $\boldsymbol{A}$ 补充 $m-n$ 个等于零的列向量成为 m 阶方阵 $\boldsymbol{A}_1=(\boldsymbol{A},\boldsymbol{O})$,$\boldsymbol{B}$ 补充 $m-n$ 个等于零的行向量成为 m 阶方阵 $\boldsymbol{B}_1=\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{O}\end{pmatrix}$. 则 $|\boldsymbol{A}_1|=|\boldsymbol{B}_1|=0$.

$$\boldsymbol{A}_1\boldsymbol{B}_1=(\boldsymbol{A},\boldsymbol{O})\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{O}\end{pmatrix}=\boldsymbol{AB}+\boldsymbol{OO}=\boldsymbol{AB}.$$

因而 $|\boldsymbol{AB}|=|\boldsymbol{A}_1\boldsymbol{B}_1|=|\boldsymbol{A}_1||\boldsymbol{B}_1|=0\times 0=0$.

($|\boldsymbol{AB}|=0$ 的另一种证法:$\boldsymbol{AB}$ 的 m 行是 $\boldsymbol{B}$ 的 n 行的线性组合, 当 $m>n$ 时这 m 个线性组合一定线性相关, 因此 $|\boldsymbol{AB}|=0$.)

$|\boldsymbol{BA}|\neq 0$ 的例子:$\boldsymbol{A}=(1,0)^{\mathrm{T}},\boldsymbol{B}=(1,0)$, 则 $|\boldsymbol{BA}|=1\neq 0$.

(3) $|\boldsymbol{A}||\boldsymbol{A}^{-1}|=|\boldsymbol{AA}^{-1}|=|\boldsymbol{I}|=1$, 因此 $|\boldsymbol{A}^{-1}|=|\boldsymbol{A}|^{-1}$.

(4) $|\boldsymbol{A}|,|\boldsymbol{A}^{-1}|$ 都是整数且乘积为 1, 仅当 $|\boldsymbol{A}|=|\boldsymbol{A}^{-1}|=\pm 1$ 时可能成立.□

4.4.3 相抵标准形

每个初等行变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{PX}$ 都是可逆变换, 存在初等行变换 $\tau:\boldsymbol{X}\mapsto\boldsymbol{QX}$ 将每个 $\sigma(A)$ 送回 $\boldsymbol{A}$. 反过来,σ 也将每个 $\tau(\boldsymbol{A})$ 送回 $\boldsymbol{A}$.τ 称为 σ 的逆变换, 记为 $\tau=\sigma^{-1}$. 反过来也有 $\sigma=\tau^{-1}=(\sigma^{-1})^{-1}$. σ,σ^{-1} 的初等方阵 $\boldsymbol{P},\boldsymbol{Q}$ 也互为逆方阵:$\boldsymbol{Q}=\boldsymbol{P}^{-1},\boldsymbol{P}=\boldsymbol{Q}^{-1}$. 可见, 初等方阵都是可逆方阵.

具体地, 有

1. $\sigma:\boldsymbol{A}\xrightarrow{(i,j)}\boldsymbol{B}=\boldsymbol{P}_{ij}\boldsymbol{A}$, 则 $\sigma^{-1}=\sigma,\boldsymbol{P}_{ij}^{-1}=\boldsymbol{P}_{ij}$.

2. $\sigma:\boldsymbol{A}\xrightarrow{\lambda(i)}\boldsymbol{B}=\boldsymbol{D}_i(\lambda)\boldsymbol{A}$, 则 $\sigma^{-1}:\boldsymbol{B}\xrightarrow{\lambda^{-1}(i)}\boldsymbol{A}=\boldsymbol{D}_i(\lambda^{-1})\boldsymbol{B}$,$\boldsymbol{D}_i(\lambda)^{-1}=\boldsymbol{D}_i(\lambda^{-1})$.

3. $\sigma:\boldsymbol{A}\xrightarrow{\lambda(j)+(i)}\boldsymbol{B}=\boldsymbol{T}_{ij}(\lambda)\boldsymbol{A}$, 则 $\sigma^{-1}:\boldsymbol{B}\xrightarrow{-\lambda(j)+(i)}\boldsymbol{A}=\boldsymbol{T}_{ij}(-\lambda)\boldsymbol{B}$,$\boldsymbol{T}_{ij}(\lambda)^{-1}=\boldsymbol{T}_{ij}(-\lambda)$.

当然, 每个初等列变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{XP}$ 也有逆变换 $\sigma^{-1}:\boldsymbol{X}\mapsto\boldsymbol{XP}^{-1}$.

在解线性方程组的时候, 我们需要对代表方程组的矩阵 $\boldsymbol{M}$ 作一系列 (有限次) 初等行变换将它化为最简阶梯形矩阵 $\boldsymbol{\Lambda}$, 也就是将 $\boldsymbol{M}$ 左乘一系列初等方阵 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_s$ 变成 $\boldsymbol{\Lambda}$:$\boldsymbol{P}_s\cdots\boldsymbol{P}_1\boldsymbol{M}=\boldsymbol{\Lambda}$. 这也就是用这些初等方阵的乘积 $\boldsymbol{P}=\boldsymbol{P}_s\cdots\boldsymbol{P}_1$ 将 $\boldsymbol{M}$ 变成 $\boldsymbol{\Lambda}$. 若干个初等方阵的乘积 $\boldsymbol{P}$ 不一定是初等方阵. 但

每个初等方阵都是可逆方阵, 有限个可逆方阵的乘积仍是可逆方阵. 因此, 有限个初等方阵的乘积 $\boldsymbol{P}$ 仍是可逆方阵. 反过来, 有:

引理 4.4.2 每个可逆方阵可以分解为有限个初等方阵乘积.

证明 由引理 1.3.1, 每个 n 阶方阵 $\boldsymbol{P}$ 可以通过有限次初等行变换化成最简阶梯形矩阵 $\boldsymbol{\Lambda}$. 当 $\boldsymbol{P}$ 可逆时 $\boldsymbol{\Lambda}$ 也可逆, 每行都不为零, 都有一个阶梯元, 共有 n 个阶梯元, 每列各有一个. 最简阶梯形矩阵 $\boldsymbol{\Lambda}$ 每列除了阶梯元为 1 外其余元素都是 0, 这迫使 $\boldsymbol{\Lambda}=\boldsymbol{I}$ 是单位矩阵.$\boldsymbol{P}$ 可以通过有限次初等行变换变成单位矩阵 $\boldsymbol{I}$, 也就是通过左乘一系列初等方阵 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_s$ 变成 $\boldsymbol{I}$. 于是

$$\boldsymbol{P}_s\cdots\boldsymbol{P}_1\boldsymbol{P}=\boldsymbol{I} \quad\Rightarrow\quad \boldsymbol{P}=\boldsymbol{P}_1^{-1}\cdots\boldsymbol{P}_s^{-1}.$$

初等方阵的逆方阵 $\boldsymbol{P}_1^{-1},\cdots,\boldsymbol{P}_s^{-1}$ 都是初等方阵,$\boldsymbol{P}$ 是这些初等方阵的乘积. □

每个矩阵可以经过一系列初等行变换化成最简阶梯形矩阵. 如果允许作初等列变换, 还可以化成更简单的形状.

定理 4.4.3 每个 $m\times n$ 矩阵 $\boldsymbol{A}$ 可以经过有限次初等变换变成唯一的最简形状

$$\boldsymbol{D}=\begin{pmatrix}\boldsymbol{I}_{(r)} & \\ & \boldsymbol{O}\end{pmatrix}, \tag{4.4.1}$$

其中 $r=\operatorname{rank}\boldsymbol{A}$.

证明 $\boldsymbol{A}$ 可以通过有限次初等行变换化成最简阶梯形矩阵 $\boldsymbol{\Lambda}=(\lambda_{ij})_{m\times n}$. 其中非零行集中在前 r 行, 每行第 1 个非零元 (阶梯元)$\lambda_{kj_k}=1$, 它所在的第 j_k 列的其余元素全为 0. 阶梯元所在各列 $\boldsymbol{\Lambda}_{j_1},\cdots,\boldsymbol{\Lambda}_{j_r}$ 依次等于前 r 个自然基向量 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r$, 将它们的适当常数倍加到不含阶梯元的各列 $\boldsymbol{\Lambda}_j$ ($\boldsymbol{\Lambda}_{j_k}$ 的 $-\lambda_{kj}$ 倍加到 $\boldsymbol{\Lambda}_j$) 可以将所有这些 $\boldsymbol{\Lambda}_j$ 全变成零. 再将阶梯元所在各列 $\boldsymbol{\Lambda}_{j_1},\cdots,\boldsymbol{\Lambda}_{j_r}$ 通过两列互换依次换到第 $1,2,\cdots,r$ 列, 就得到 (4.4.1) 中的最简形状 $\boldsymbol{D}=(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r,\boldsymbol{0},\cdots,\boldsymbol{0})$.

由于矩阵经过初等变换秩不变, 因此 $r=\operatorname{rank}\boldsymbol{D}=\operatorname{rank}\boldsymbol{A}$.$\boldsymbol{D}$ 由 r 唯一决定. □

定义 4.4.2(矩阵的相抵) 如果矩阵 $\boldsymbol{A}\in F^{m\times n}$ 可以经过有限次初等变换变成 $\boldsymbol{B}$, 就称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相抵** (equivalent), 也称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **等价**. 定理 4.4.3 中的 $\boldsymbol{D}$ 称为 $\boldsymbol{A}$ 的**相抵标准形** (canonical form of equivalent matrices). □

显然有:

引理 4.4.3 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相抵 $\Leftrightarrow$ 存在可逆方阵 $\boldsymbol{P},\boldsymbol{Q}$ 使 $\boldsymbol{PAQ}=\boldsymbol{B}$. □

引理 4.4.4 矩阵的相抵关系具有如下性质:

(1) $\boldsymbol{A}$ 与 $\boldsymbol{A}$ 自己相抵.

(2) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相抵, 则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相抵.

(3) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相抵, 且 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 相抵, 则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 相抵. □

由定理 4.4.3 得出

推论 4.4.2 对每个矩阵 $\boldsymbol{A}$, 存在可逆方阵 $\boldsymbol{P},\boldsymbol{Q}$ 使 $\boldsymbol{PAQ}$ 等于 $\boldsymbol{A}$ 的相抵标准形 $\boldsymbol{D}$. □

推论 4.4.3 设 $\boldsymbol{A},\boldsymbol{B}\in F^{m\times n}$. 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相抵 $\Leftrightarrow$rank $\boldsymbol{A}=$ rank $\boldsymbol{B}$. □

4.4.4 矩阵相抵应用例

例 4 试利用矩阵的相抵标准形研究齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 解空间 $V_{\boldsymbol{A}}$ 的维数.

解 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, 秩 rank $\boldsymbol{A}=r$. 则存在 m 阶可逆方阵 $\boldsymbol{P}$ 和 n 阶可逆方阵 $\boldsymbol{Q}$ 使 $\boldsymbol{A}=\boldsymbol{PDQ}$ 对 $\boldsymbol{A}$ 的相抵标准形

$$\boldsymbol{D}=\begin{pmatrix}\boldsymbol{I}_{(r)} & \\ & \boldsymbol{O}\end{pmatrix}$$

成立. 方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 成为 $\boldsymbol{PDQX}=\boldsymbol{0}$. 这个方程两边同时左乘 $\boldsymbol{P}^{-1}$, 并令 $\boldsymbol{Y}=\boldsymbol{QX}=(y_1,\cdots,y_n)^{\mathrm{T}}$, 得 $\boldsymbol{DY}=\boldsymbol{0}$, 即

$$\begin{pmatrix}\boldsymbol{I}_{(r)} & \\ & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}=\boldsymbol{0},\quad \text{即}\quad \begin{pmatrix}y_1\\ \vdots\\ y_r\\ \boldsymbol{O}\end{pmatrix}=\boldsymbol{0}.$$

其解为 $y_1=\cdots=y_r=0$, 即 $\boldsymbol{Y}=(0,\cdots,0,y_{r+1},\cdots,y_n)^{\mathrm{T}}$. 原方程组通解为

$$\boldsymbol{X}=\boldsymbol{Q}^{-1}\boldsymbol{Y}=\boldsymbol{Q}^{-1}\begin{pmatrix}\boldsymbol{0}\\ y_{r+1}\\ \vdots\\ y_n\end{pmatrix}=y_{r+1}\boldsymbol{b}_{r+1}+\cdots+y_n\boldsymbol{b}_n.$$

其中 $\boldsymbol{b}_{r+1},\cdots,\boldsymbol{b}_n$ 是可逆方阵 $\boldsymbol{Q}^{-1}$ 的后 $n-r$ 列,$y_{r+1},\cdots,y_n$ 可以在 F 中相互独立地任意取值. 因此, 原方程组的解集 $V_{\boldsymbol{A}}$ 由 $n-r$ 个线性无关的向量 $\boldsymbol{b}_{r+1},\cdots,\boldsymbol{b}_n$ 的全体线性组合组成, 就是由这 $n-r$ 个线性无关向量生成的子空间, 维数 $\dim V_{\boldsymbol{A}}=n-r=n-\operatorname{rank}\boldsymbol{A}$. □

例 5 求证: 两个矩阵 $\boldsymbol{A},\boldsymbol{B}$ 的乘积的秩 rank $(\boldsymbol{AB})\leqslant$ rank $\boldsymbol{A}$ 且 rank $(\boldsymbol{AB})\leqslant$ rank $\boldsymbol{B}$.

证法 1 设 $\operatorname{rank}\boldsymbol{A}=r$,$\operatorname{rank}\boldsymbol{B}=s$. 则存在可逆方阵 $\boldsymbol{P}_1,\boldsymbol{Q}_1,\boldsymbol{P}_2,\boldsymbol{Q}_2$ 使

$$\boldsymbol{A}=\boldsymbol{P}_1\begin{pmatrix}\boldsymbol{I}_{(r)} & \\ & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}_1,\quad \boldsymbol{B}=\boldsymbol{P}_2\begin{pmatrix}\boldsymbol{I}_{(s)} & \\ & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}_2,$$

于是

$$\boldsymbol{AB}=\boldsymbol{P}_1\boldsymbol{M}\boldsymbol{Q}_2,$$

其中

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{I}_{(r)} & \\ & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}_1\boldsymbol{P}_2\begin{pmatrix}\boldsymbol{I}_{(s)} & \\ & \boldsymbol{O}\end{pmatrix}=\begin{pmatrix}\boldsymbol{D}_{(r\times s)} & \\ & \boldsymbol{O}\end{pmatrix},$$

$\boldsymbol{D}_{r\times s}$ 是由 $\boldsymbol{Q}_1\boldsymbol{P}_2$ 的前 r 行与前 s 列的交叉位置的元素组成的矩阵.$\boldsymbol{AB}$ 与 $\boldsymbol{M}$ 相抵, 因此

$$\operatorname{rank}(\boldsymbol{AB})=\operatorname{rank}\boldsymbol{M}=\operatorname{rank}\boldsymbol{D}\leqslant r\text{且}\leqslant s.$$

证法 2 $\boldsymbol{AB}$ 的列向量组 $C(\boldsymbol{AB})$ 是 $\boldsymbol{A}$ 的列向量组 $C(\boldsymbol{A})$ 的线性组合, 由推论 2.5.2 知 $\operatorname{rank}C(\boldsymbol{AB})\leqslant\operatorname{rank}C(\boldsymbol{A})$. 即 $\operatorname{rank}(\boldsymbol{AB})\leqslant\operatorname{rank}\boldsymbol{A}$.

类似地,$\boldsymbol{AB}$ 的行向量组 $R(\boldsymbol{AB})$ 是 $\boldsymbol{A}$ 的行向量组 $R(\boldsymbol{B})$ 的线性组合, 所以 $\operatorname{rank}(\boldsymbol{AB})=\operatorname{rank}R(\boldsymbol{AB})\leqslant\operatorname{rank}R(\boldsymbol{B})=\operatorname{rank}\boldsymbol{B}$.

证法 3 设 $\boldsymbol{A}\in F^{m\times n}$,$\boldsymbol{B}\in F^{n\times p}$, 则 $(\boldsymbol{A},\boldsymbol{O}_{(m\times p)})$ 相抵于

$$(\boldsymbol{A},\boldsymbol{O}_{(m\times p)})\begin{pmatrix}\boldsymbol{I}_{(n)} & \boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_{(p)}\end{pmatrix}=(\boldsymbol{A},\boldsymbol{AB}),\tag{4.4.2}$$

因此

$$\operatorname{rank}\boldsymbol{A}=\operatorname{rank}(\boldsymbol{A},\boldsymbol{O})=\operatorname{rank}(\boldsymbol{A},\boldsymbol{AB}).$$

$\operatorname{rank}(\boldsymbol{AB})$ 等于 $\boldsymbol{AB}$ 的最大非零子式的阶, 但 $\boldsymbol{AB}$ 的最大非零子式也是 $(\boldsymbol{A},\boldsymbol{AB})$ 的非零子式, 因此 $\operatorname{rank}(\boldsymbol{A},\boldsymbol{AB})\geqslant\operatorname{rank}(\boldsymbol{AB})$. 由此得到 $\operatorname{rank}\boldsymbol{A}\geqslant\operatorname{rank}(\boldsymbol{AB})$.

类似地, 由

$$\begin{pmatrix}\boldsymbol{I}_{(n)} & \boldsymbol{O}\\ \boldsymbol{A} & \boldsymbol{I}_{(m)}\end{pmatrix}\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{O}_{(m\times p)}\end{pmatrix}=\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{AB}\end{pmatrix}\tag{4.4.3}$$

得

$$\operatorname{rank}\boldsymbol{B}=\operatorname{rank}\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{O}\end{pmatrix}=\operatorname{rank}\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{AB}\end{pmatrix}\geqslant\operatorname{rank}(\boldsymbol{AB}).\qquad\square$$

注 例 5 的三种证明方法, 证法 1 是按部就班利用相抵标准形, 最容易模仿. 证法 2 最简便, 但要求对矩阵乘法的几何意义非常熟悉. 证法 3 依赖于两个巧妙的等式 (4.4.2),(4.4.3),

这两个等式不难通过分块运算验证，问题是：这两个等式是怎么设计来的？假如 a,b 是两个数，不难想到将行向量 $(a,0)$ 的第 1 列乘 b 加到第 2 列得到 (a,ab). 这个列变换可以通过右乘初等方阵实现：

$$(a,0)\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}=(a,ab).$$

将 $a,b,1,0$ 分别替换成矩阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{I},\boldsymbol{O}$，就得到 (4.4.2). 类似地可得 (4.4.3). 更多例子见 §4.5*.

例 6 秩为 1 的矩阵 $\boldsymbol{A}$ 可以写成某个非零列向量 $\boldsymbol{\beta}$ 和非零行向量 $\boldsymbol{\alpha}$ 的乘积:$\boldsymbol{A}=\boldsymbol{\beta\alpha}$.

解 存在可逆方阵 $\boldsymbol{P},\boldsymbol{Q}$ 使

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}=\boldsymbol{P}\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}(1,0,\cdots,0)\boldsymbol{Q}=\boldsymbol{\beta\alpha},$$

其中 $\boldsymbol{\beta}=\boldsymbol{P}\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}$ 是 $\boldsymbol{P}$ 的第一列,$\boldsymbol{\alpha}=(1,0,\cdots,0)\boldsymbol{Q}$ 是 $\boldsymbol{Q}$ 的第一行. □

例 7 设 $\boldsymbol{A}\in F^{n\times n}$,$\operatorname{rank}\boldsymbol{A}=r$. 则

(1) 存在 $\boldsymbol{B}\in F^{n\times n}$ 满足条件 $\operatorname{rank}\boldsymbol{B}=n-r$ 且 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{O}$.

(2) 存在 $\boldsymbol{C}\in F^{n\times n}$ 满足条件 $\boldsymbol{AC}=\boldsymbol{O}$ 且 $\boldsymbol{A}+\boldsymbol{C}$ 可逆.

证明 存在可逆方阵 $\boldsymbol{P},\boldsymbol{Q}\in F^{n\times n}$ 使

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_{(r)} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q},$$

(1) 取 $\boldsymbol{B}=\boldsymbol{Q}^{-1}\begin{pmatrix}\boldsymbol{O}_{(r)} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_{(n-r)}\end{pmatrix}\boldsymbol{P}^{-1}$ 即满足要求.

(2) 取 $\boldsymbol{C}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{O}_{(r)} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{I}_{(n-r)}\end{pmatrix}\boldsymbol{Q}$ 即满足要求. □

习题 4.4

1. 已知方阵 $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$.

(1) 通过初等行变换与初等列变换化成单位矩阵 $\boldsymbol{I}$;

(2) 将 $\boldsymbol{A}$ 左乘或右乘初等方阵化成单位矩阵 $\boldsymbol{I}$;

(3) 将 $\boldsymbol{A}$ 写成初等方阵的乘积.

2. 已知

$$\boldsymbol{A}=\begin{pmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 1-\lambda\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \lambda^{-1} & 1\end{pmatrix}\begin{pmatrix}\lambda & 0\\ 0 & \lambda^{-1}\end{pmatrix},$$

(1) 将 $\boldsymbol{A}$ 经过一系列初等行变换得到 $\boldsymbol{B}$;

(2) 将 $\boldsymbol{A}$ 写成第三类初等方阵的乘积.

3. 试将 $\boldsymbol{A}=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$ 写成第三类初等方阵的乘积.

4. 举出矩阵 $\boldsymbol{A},\boldsymbol{B}$ 使 $\det(\boldsymbol{AB})\neq\det(\boldsymbol{BA})$. 其中的 $\boldsymbol{A},\boldsymbol{B}$ 是否可能是方阵? 为什么?

5. 已知

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & \cdots & n\\ 2 & 3 & \cdots & n+1\end{pmatrix},\quad n\geqslant 2,$$

求 $\det(\boldsymbol{AA}^{\mathrm{T}})$ 与 $\det(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$.

6. 证明: 任意一个秩为 r 的矩阵都可以表示为 r 个秩为 1 的矩阵之和.

7. 已知方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 的秩等于 1, $\lambda=a_{11}+\cdots+a_{nn}$.

(1) 求证:$\boldsymbol{A}^2=\lambda\boldsymbol{A}$; (2) 求 $\det(\boldsymbol{I}+\boldsymbol{A})$;

(3) 当 $\boldsymbol{I}+\boldsymbol{A}$ 可逆时求 $(\boldsymbol{I}+\boldsymbol{A})^{-1}$.

§4.5* 更多的例子

4.5.1 运算律的应用

例 1 已知

$$\boldsymbol{A}=\begin{pmatrix}1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix},$$

求一个 3 阶方阵 $\boldsymbol{B}$ 使其满足条件 $\boldsymbol{B}^{10}=\boldsymbol{A}$.

解 $\boldsymbol{A}=\boldsymbol{I}+\boldsymbol{N}$, 其中

$$\boldsymbol{N}=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix},\quad \boldsymbol{N}^2=\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},\quad \boldsymbol{N}^3=\boldsymbol{O}.$$

当实数 x 满足条件 $|x|<1$ 时,$f(x)=(1+x)^{\frac{1}{10}}$ 的泰勒 (Taylor) 展开式

$$f(x)=(1+x)^{\frac{1}{10}}=1+\frac{1}{10}x+\frac{\frac{1}{10}\left(\frac{1}{10}-1\right)}{2}x^2+\cdots \tag{4.5.1}$$

满足条件 $[f(x)]^{10}=1+x$. 将 $x=\boldsymbol{N}$ 代入 (4.5.1) 得

$$\boldsymbol{B}=f(\boldsymbol{N})=\boldsymbol{I}+\frac{1}{10}\boldsymbol{N}-\frac{9}{200}\boldsymbol{N}^2=\begin{pmatrix}1 & \frac{1}{10} & -\frac{9}{200}\\ 0 & 1 & \frac{1}{10}\\ 0 & 0 & 1\end{pmatrix}.$$

经验证

$$\begin{aligned}\boldsymbol{B}^{10}&=\left(\boldsymbol{I}+\frac{1}{10}\boldsymbol{N}-\frac{9}{200}\boldsymbol{N}^2\right)^{10}\\ &=\boldsymbol{I}+10\left(\frac{1}{10}\boldsymbol{N}-\frac{9}{200}\boldsymbol{N}^2\right)+\frac{10\times 9}{2}\left(\frac{1}{10}\boldsymbol{N}-\frac{9}{200}\boldsymbol{N}^2\right)^2\\ &=\boldsymbol{I}+\boldsymbol{N}-\frac{9}{20}\boldsymbol{N}^2+\frac{10\times 9}{2}\frac{1}{10^2}\boldsymbol{N}^2=\boldsymbol{I}+\boldsymbol{N}=\boldsymbol{A}.\end{aligned}$$

可见 $\boldsymbol{B}$ 符合要求. □

函数 $f(x)=(1+x)^{\frac{1}{10}}$ 的泰勒 (Taylor) 展开式 (4.5.1) 是无穷级数, 要求 $|x|<1$ 是为了保证级数的收敛. 将 x 替换成 $\boldsymbol{N}$ 之后, 由于 $\boldsymbol{N}^3=\boldsymbol{O}$, 收敛性没有任何问题.(4.5.1) 右边的级数既然是 $(1+x)^{\frac{1}{10}}$ 的泰勒 (Taylor) 展开式, 它的 10 次幂必然应当等于 $1+x$. 特别地,(4.5.1) 右边的前 3 项之和的 10 次幂的前 3 项等于 $1+x$:

$$\left(1+\frac{1}{10}x+\frac{\frac{1}{10}\left(\frac{1}{10}-1\right)}{2}x^2\right)^{10}=1+x+0x^2+\cdots \tag{4.5.2}$$

将 (4.5.2) 左边展开的过程只用到数的加法与乘法的交换律、结合律、分配律等运算律. 将 x 替换成方阵 $\boldsymbol{N}$, 不但分配律与结合律仍成立, 而且由于其中涉及的矩阵乘法都是 $\boldsymbol{N}$ 的各次幂之间的乘法, 交换律也没有违背. 因此, 将 $x=\boldsymbol{N}$ 代入 (4.5.2) 之后等号仍成立. 而且, 由于 $\boldsymbol{N}^3=\boldsymbol{O}$,(4.5.2) 右边的 2 次以上的项都不需考虑, 必然得到 $[f(\boldsymbol{N})]^{10}=\boldsymbol{I}+\boldsymbol{N}=\boldsymbol{A}$. 不过, 在实际解题时, 可以不说明刚才这些道理, 只要对 $\boldsymbol{B}=f(\boldsymbol{N})$ 实际计算验证 $\boldsymbol{B}^{10}=\boldsymbol{A}$, 理由就已经充分了.

如果有些读者还没有学过泰勒 (Taylor) 级数, 只要知道牛顿二项式定理

$$(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\cdots$$

中的指数 α 不仅可以取正整数而且可以取任意实数, 取 $\alpha=\frac{1}{10}$ 就得到展开式 (4.5.1).

4.5.2 分块初等变换的应用

例 2 求证:

(1) 对 $\boldsymbol{A}\in F^{m\times n},\boldsymbol{B}\in F^{p\times q},\boldsymbol{C}\in F^{p\times n}$, 有

$$\operatorname{rank}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}=\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}\leqslant\operatorname{rank}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{C}&\boldsymbol{B}\end{pmatrix}.$$

(2) 对 $\boldsymbol{A},\boldsymbol{B}\in F^{m\times n}$, 有 $\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})\leqslant\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}$.

(3) 对 $\boldsymbol{A}\in F^{m\times n},\boldsymbol{B}\in F^{n\times p}$, 有

$$\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}-n\leqslant\operatorname{rank}\boldsymbol{AB}\leqslant\min\{\operatorname{rank}\boldsymbol{A},\operatorname{rank}\boldsymbol{B}\}.$$

证明 (1) 设 $r=\operatorname{rank}\boldsymbol{A},s=\operatorname{rank}\boldsymbol{B}$. 则存在可逆方阵 $\boldsymbol{P}_1,\boldsymbol{Q}_1,\boldsymbol{P}_2,\boldsymbol{Q}_2$ 使

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_{(r)}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix},\quad \boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_{(s)}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

取可逆方阵

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{P}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{P}_2\end{pmatrix},\quad \boldsymbol{Q}=\begin{pmatrix}\boldsymbol{Q}_1&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}_2\end{pmatrix},$$

则

$$\boldsymbol{S}_1=\boldsymbol{P}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{C}&\boldsymbol{B}\end{pmatrix}\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1&\boldsymbol{O}\\\boldsymbol{P}_2\boldsymbol{C}\boldsymbol{Q}_1&\boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{(r)}&&&\\&\boldsymbol{O}&&\\&\boldsymbol{C}_1&\boldsymbol{I}_{(s)}&\\&&&\boldsymbol{O}\end{pmatrix}.$$

$\boldsymbol{S}_1$ 中的块 $\boldsymbol{I}_{(r)}$ 与 $\boldsymbol{I}_{(s)}$ 所在的第 $1,2,\cdots,r,m+1,\cdots,m+s$ 行及第 $1,2,\cdots,r,n+1,\cdots,n+s$ 列交叉位置的元素组成 $r+s$ 阶非零子式 $\begin{vmatrix}\boldsymbol{I}_{(r)}&\boldsymbol{O}\\ *&\boldsymbol{I}_{(s)}\end{vmatrix}=1$, 因此

$$\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}=r+s\leqslant\operatorname{rank}\boldsymbol{S}_1=\operatorname{rank}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{C}&\boldsymbol{B}\end{pmatrix}.$$

当 $\boldsymbol{C}=\boldsymbol{O}$ 时 $\boldsymbol{C}_1=\boldsymbol{O}.\boldsymbol{S}_1$ 中除了块 $\boldsymbol{I}_{(r)}$ 与 $\boldsymbol{I}_{(s)}$ 所在的 $r+s$ 个非零行以外其余的行全为零, $\operatorname{rank}\boldsymbol{S}_1\leqslant r+s$, 因而 $\operatorname{rank}\boldsymbol{S}_1=r+s$. 于是

$$\operatorname{rank}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}=\operatorname{rank}\boldsymbol{S}_1=r+s=\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}.$$

(2) 易验证等式

$$\begin{pmatrix} I_{(m)} & O \\ I & I_{(m)} \end{pmatrix}\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} I_{(n)} & O \\ I & I_{(n)} \end{pmatrix}=\begin{pmatrix} A & O \\ A+B & B \end{pmatrix}. \qquad (4.5.3)$$

将等号左边三个矩阵的乘积记为 S_1, 等号右边的矩阵记为 S_2. 则 S_1 与 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 相抵, 秩相等, 因而 $\operatorname{rank} S_1=\operatorname{rank} A+\operatorname{rank} B$.

$A+B$ 是 S_2 的子矩阵. 设 $\operatorname{rank}(A+B)=d$, 则 $A+B$ 的最大非零子式为 d 阶. 这个 d 阶非零子式也是 S_2 的非零子式, 因此 $\operatorname{rank} S_2 \geqslant d=\operatorname{rank}(A+B)$.

因此

$$\operatorname{rank} A+\operatorname{rank} B=\operatorname{rank} S_1=\operatorname{rank} S_2 \geqslant \operatorname{rank}(A+B).$$

(3) §4.4 例 5 中已经证明 $\operatorname{rank}(AB) \leqslant \operatorname{rank} A$ 且 $\leqslant \operatorname{rank} B$, 因此

$$\operatorname{rank} AB \leqslant \min\{\operatorname{rank} A, \operatorname{rank} B\}.$$

另一方面, 由

$$\begin{pmatrix} I & -A \\ O & I \end{pmatrix}\begin{pmatrix} A & O \\ I_{(n)} & B \end{pmatrix}\begin{pmatrix} I & -B \\ O & I \end{pmatrix}\begin{pmatrix} O & I \\ -I & O \end{pmatrix}=\begin{pmatrix} AB & O \\ O & I_{(n)} \end{pmatrix}, \qquad (4.5.4)$$

知

$$\operatorname{rank}\begin{pmatrix} A & O \\ I_{(n)} & B \end{pmatrix}=\operatorname{rank}\begin{pmatrix} AB & O \\ O & I_{(n)} \end{pmatrix}=\operatorname{rank}(AB)+\operatorname{rank} I_{(n)},$$

于是

$$\operatorname{rank} A+\operatorname{rank} B \leqslant \operatorname{rank}\begin{pmatrix} A & O \\ I_{(n)} & B \end{pmatrix}=\operatorname{rank}(AB)+n,$$

$$\operatorname{rank} A+\operatorname{rank} B-n \leqslant \operatorname{rank}(AB).$$

□

例 3 设 $A \in F^{m\times n}, B \in F^{n\times m}$. 求证: 如果 $I_{(m)}-AB$ 可逆, 则 $I_{(n)}-BA$ 可逆. 并求出 $(I_{(n)}-BA)^{-1}$.

证明 令

$$S=\begin{pmatrix} I_{(n)} & B \\ O & I_{(m)}-AB \end{pmatrix}=\begin{pmatrix} I_{(n)} & O \\ O & I_{(m)}-AB \end{pmatrix}\begin{pmatrix} I_{(n)} & B \\ O & I_{(m)} \end{pmatrix},$$

则

$$\begin{pmatrix} I_{(n)} & O \\ A & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & B \\ O & I_{(m)}-AB \end{pmatrix} = \begin{pmatrix} I_{(n)} & B \\ A & I_{(m)} \end{pmatrix},$$

$$\begin{pmatrix} I_{(n)} & -B \\ O & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & B \\ A & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & O \\ -A & I_{(m)} \end{pmatrix} = \begin{pmatrix} I_{(n)}-BA & O \\ O & I_{(m)} \end{pmatrix}.$$

于是

$$\begin{aligned}\begin{pmatrix} I_{(n)}-BA & O \\ O & I_{(m)} \end{pmatrix} = &\begin{pmatrix} I_{(n)} & -B \\ O & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & O \\ A & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & O \\ O & I_{(m)}-AB \end{pmatrix} \\ &\begin{pmatrix} I_{(n)} & B \\ O & I_{(m)} \end{pmatrix} \begin{pmatrix} I_{(n)} & O \\ -A & I_{(m)} \end{pmatrix}.\end{aligned} \tag{4.5.5}$$

如果 $I_{(m)}-AB$ 可逆, 则等式 (4.5.5) 右边的每个方阵都可逆, 左边的矩阵等于可逆方阵的乘积, 仍然可逆, 因此 $I_{(n)}-BA$ 可逆.

等式 (4.5.5) 两边同时求逆得

$$\begin{aligned}&\begin{pmatrix} (I_{(n)}-BA)^{-1} & O \\ O & I_{(m)} \end{pmatrix} \\ =&\begin{pmatrix} I_{(n)} & O \\ A & I_{(m)} \end{pmatrix}\begin{pmatrix} I_{(n)} & -B \\ O & I_{(m)} \end{pmatrix}\begin{pmatrix} I_{(n)} & O \\ O & (I_{(m)}-AB)^{-1} \end{pmatrix}\begin{pmatrix} I_{(n)} & O \\ -A & I_{(m)} \end{pmatrix}\begin{pmatrix} I_{(n)} & B \\ O & I_{(m)} \end{pmatrix} \\ =&\begin{pmatrix} I_{(n)} & O \\ A & I_{(m)} \end{pmatrix}\begin{pmatrix} I_{(n)}-B(I_{(m)}-AB)^{-1}A & * \\ * & * \end{pmatrix}\begin{pmatrix} I_{(n)} & B \\ O & I_{(m)} \end{pmatrix} \\ =&\begin{pmatrix} I_{(n)}-B(I_{(m)}-AB)^{-1}A & * \\ * & * \end{pmatrix}.\end{aligned}$$

因此

$$(I_{(n)}-BA)^{-1} = I_{(n)}-B(I_{(m)}-AB)^{-1}A. \qquad \square$$

4.5.3 利用矩阵乘法计算行列式

例 4 设 $A \in F^{n\times m}, B \in F^{m\times n}$. 求证:

$$|I_{(n)}-AB| = |I_{(m)}-BA|. \tag{4.5.6}$$

证明 在例 3 的等式 (4.5.5) 两边取行列式即可得到行列式等式 (4.5.6). 不过,(4.5.5) 是为了由 $(I-AB)^{-1}$ 求 $(I-BA)^{-1}$ 设计的. 如果只是为了证明

(4.5.6), 可以设计更简单的分块乘法等式

$$\begin{pmatrix} I_{(m)} & O \\ A & I_{(n)} \end{pmatrix}\begin{pmatrix} I_{(m)} & B \\ O & I_{(n)}-AB \end{pmatrix}\begin{pmatrix} I_{(m)} & O \\ -A & I_{(n)} \end{pmatrix}=\begin{pmatrix} I_{(m)}-BA & B \\ O & I_{(n)} \end{pmatrix}, \quad (4.5.7)$$

等式两边取行列式即得 (4.5.6). □

例 4 中的等式 (4.5.6) 可以作为公式来计算行列式. 特别是当 $m=1$ 时, 将 n 阶行列式 $|I_{(n)}-AB|$ 化为 1 阶行列式 $|I_{(m)}-BA|$ 即 $1-BA$ 可以立即算出结果.

例 5 计算行列式

$$\Delta=\begin{vmatrix} 1+a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & 1+a_nb_n \end{vmatrix}$$

解

$$\Delta=\left|I-\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}(-b_1,\cdots,-b_n)\right|$$

$$=1-(-b_1,\cdots,-b_n)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}=1+a_1b_1+\cdots+a_nb_n. \qquad □$$

例 6 设 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 及 A,D 都是数域 F 上的方阵. 求证:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=\begin{cases} |A||D-CA^{-1}B|, & \text{当 } A \text{ 可逆 ;} \\ |D||A-BD^{-1}C|, & \text{当 } D \text{ 可逆 .} \end{cases} \quad (4.5.8)$$

证明 当 A 可逆时,

$$\begin{pmatrix} I & O \\ -CA^{-1} & I \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A & B \\ O & D-CA^{-1}B \end{pmatrix}.$$

两边同时取行列式, 得

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}=\begin{vmatrix} A & B \\ O & D-CA^{-1}B \end{vmatrix}=|A|\cdot|D-CA^{-1}B|. \quad (4.5.9)$$

当 $\boldsymbol{D}$ 可逆时,

$$\begin{pmatrix} \boldsymbol{I} & -\boldsymbol{B}\boldsymbol{D}^{-1} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}.$$

两边同时取行列式, 得

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \begin{vmatrix} \boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{D}||\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|.$$ □

例 6 中的公式可以将行列式 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix}$ 化为较低阶行列式来计算, 也称为行列式的降阶公式.

例 7 已知 n 阶方阵

$$\boldsymbol{A} = \begin{pmatrix} 0 & & & -a_n \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_2 \\ & & 1 & -a_1 \end{pmatrix},$$

求 n 阶行列式 $|\lambda \boldsymbol{I}-\boldsymbol{A}|$.

解 将 $|\lambda \boldsymbol{I}-\boldsymbol{A}|$ 的第 2 至 n 行分别乘 -1, 再将第 1 行依次与第 2 至 n 行互换, 得到

$$|\lambda \boldsymbol{I}-\boldsymbol{A}| = \begin{vmatrix} \lambda & & & a_n \\ -1 & \ddots & & \vdots \\ & \ddots & \lambda & a_2 \\ & & -1 & \lambda+a_1 \end{vmatrix} = (-1)^{n-1}(-1)^{n-1} \begin{vmatrix} \boldsymbol{\Lambda} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix},$$

其中

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & -\lambda & & \\ & 1 & \ddots & \\ & & \ddots & -\lambda \\ & & & 1 \end{pmatrix} \in F^{(n-1)\times(n-1)}, \quad \boldsymbol{B} = \begin{pmatrix} -a_{n-1} \\ \vdots \\ -\lambda-a_1 \end{pmatrix} \in F^{(n-1)\times 1},$$

$$\boldsymbol{C} = (\lambda, 0, \cdots, 0) \in F^{1\times(n-1)}, \boldsymbol{D} = a_n \in F^{1\times 1}.$$

记

$$N=\begin{pmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}\in F^{(n-1)\times(n-1)},$$

则

$$\begin{aligned}\Lambda^{-1}&=(I-\lambda N)^{-1}=I+\lambda N+\lambda^2N^2+\cdots+\lambda^{n-2}N^{n-2}\\ &=\begin{pmatrix}1 & \lambda & \cdots & \lambda^{n-2}\\ 0 & 1 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \lambda\\ 0 & \cdots & 0 & 1\end{pmatrix}.\end{aligned}$$

由例 6 的行列式降阶公式得

$$\begin{aligned}|\lambda I-A|&=|\Lambda||D-C\Lambda^{-1}B|\\ &=1\left[a_n-(\lambda,\lambda^2,\cdots,\lambda^{n-1})\begin{pmatrix}-a_{n-1}\\ \vdots\\ -\lambda-a_1\end{pmatrix}\right]\\ &=1(a_n+a_{n-1}\lambda+a_{n-2}\lambda^2+\cdots+a_1\lambda^{n-1}+\lambda^n)\\ &=\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}+\cdots+a_{n-1}\lambda+a_n.\end{aligned}$$

□

习题 4.5

1. 已知上三角形方阵

$$A=\begin{pmatrix}1 & 1 & \cdots & 1\\ 0 & 1 & \cdots & 1\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{pmatrix},$$

(1) 求 A^{10}.

(2) 求 B 使 $B^{10}=A$.

2. 设 A,B 是行数相同的矩阵,(A,B) 是由 A,B 并排组成的矩阵. 证明:

$$\operatorname{rank}(A,B)\leqslant\operatorname{rank}A+\operatorname{rank}B.$$

3. 对角阵 $A=\operatorname{diag}(a_1,a_2,\cdots,a_n)$ 可逆,$\beta=(a_1,a_2,\cdots,a_n)$. 求行列式 $\det(A-\beta^{\mathrm T}\beta)$.

4. 设 $A\in F^{n\times m}$, $B\in F^{m\times n}$. 求证: 多项式 $\lambda^m|\lambda I_{(n)}-AB|=\lambda^n|\lambda I_{(m)}-BA|$.

5. 设 n 阶方阵 $\boldsymbol{A}$ 满足条件 $\boldsymbol{A}^2=\boldsymbol{I}$, 求证:$\operatorname{rank}(\boldsymbol{A}-\boldsymbol{I})+\operatorname{rank}(\boldsymbol{A}+\boldsymbol{I})=n$.

6. 设 $\boldsymbol{A}$ 是 n 阶方阵. 证明: 如果 $\operatorname{rank}\boldsymbol{A}^m=\operatorname{rank}\boldsymbol{A}^{m+1}$ 对某个正整数 m 成立, 则 $\operatorname{rank}\boldsymbol{A}^m=\operatorname{rank}\boldsymbol{A}^{m+k}$ 对所有的正整数 k 成立.

7. 设 $\boldsymbol{A}^*$ 表示 n 阶方阵 $\boldsymbol{A}$ 的伴随方阵. 证明:

(1) $\operatorname{rank}\boldsymbol{A}^*=n \quad \Leftrightarrow \quad \operatorname{rank}\boldsymbol{A}=n$;

(2) $\operatorname{rank}\boldsymbol{A}^*=1 \quad \Leftrightarrow \quad \operatorname{rank}\boldsymbol{A}=n-1$;

(3) $\operatorname{rank}\boldsymbol{A}^*=0 \quad \Leftrightarrow \quad \operatorname{rank}\boldsymbol{A}<n-1$.

第 5 章　矩阵的相合与相似

§5.1　欧 氏 空 间

5.1.1　最小二乘法

例 1　研究表明, 某种材料在生产过程中的废品率 y 与某种化学成分的含量百分比 x 近似地成一次函数关系 $y = kx + b$. 试根据 x, y 的如下三组实测数据求待定常数 k, b 的值.

$x(\%)$	3.7	4.0	4.2
$y(\%)$	0.9	0.6	0.35

分析　将 x, y 的三组对应值代入函数关系式 $kx + b = y$, 得到 k, b 满足的方程组

$$\begin{cases} 3.7k + b = 0.9, \\ 4.0k + b = 0.6, \\ 4.2k + b = 0.35. \end{cases} \tag{5.1.1}$$

如果方程组 (5.1.1) 有解 (k, b), 就得到了函数式 $y = kx + b$. 然而, 由于 x, y 只是近似地而不是完全精确地成一次函数关系, 实测到的 (x, y) 数据也是近似的, 方程组 (5.1.1) 没有精确解. 我们求使总体误差尽可能小的近似解 (k, b).

方程组 (5.1.1) 可以写成向量形式和矩阵形式

$$k\boldsymbol{a}_1 + b\boldsymbol{a}_2 = \boldsymbol{c}, \tag{5.1.2}$$

$$\boldsymbol{AX} = \boldsymbol{c}, \tag{5.1.3}$$

其中

$$\boldsymbol{a}_1 = \begin{pmatrix} 3.7 \\ 4.0 \\ 4.2 \end{pmatrix}, \boldsymbol{a}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} 0.9 \\ 0.6 \\ 0.35 \end{pmatrix};$$

$$A = (a_1, a_2) = \begin{pmatrix} 3.7 & 1 \\ 4.0 & 1 \\ 4.2 & 1 \end{pmatrix}, \ X = \begin{pmatrix} k \\ b \end{pmatrix}.$$

(5.1.2) 的几何意义是: 将三维几何向量 $\boldsymbol{c} = \overrightarrow{OC}$ 表示成 $\boldsymbol{a}_1 = \overrightarrow{OA_1}$ 与 $\boldsymbol{a}_2 = \overrightarrow{OA_2}$ 的线性组合 $k\boldsymbol{a}_1 + b\boldsymbol{a}_2$, 求线性组合系数 k, b. $\overrightarrow{OA_1}, \overrightarrow{OA_2}$ 的全体实系数线性组合 $\overrightarrow{OP}$ 对应的点 P 组成过 O, A_1, A_2 的平面 π (如图 5–1). 方程组 (5.1.2) 没有精确解, 就是说 $\overrightarrow{OC} = \boldsymbol{c}$ 对应的点 C 不在平面 π 内. 我们在平面 π 内求点 D 与 C 的距离 $|DC|$ 最近, 也就是求线性组合 $k\boldsymbol{a}_1 + b\boldsymbol{a}_2 = \overrightarrow{OD}$ 与 $\boldsymbol{c} = \overrightarrow{OC}$ 之差 $\overrightarrow{CD} = \overrightarrow{OD} - \overrightarrow{OC} = k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{c}$ 的长度 $d = |CD|$ 最短, 也就是使

$$\begin{aligned} d^2 &= |CD|^2 = |k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{b}|^2 \\ &= (3.7k + b - 0.9)^2 + (4.0k + b - 0.6)^2 + (4.2k + b - 0.35)^2 \end{aligned} \tag{5.1.4}$$

取最小值, 得到的 (k, b) 就可以认为是 (5.1.1) 的最优近似解.

根据几何知识, 平面 π 内与 C 距离最近的点 D 是从点 C 到平面 π 的垂线的垂足, 也就是说: $\overrightarrow{CD} = k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{c}$ 与 $\boldsymbol{a}_1, \boldsymbol{a}_2$ 都垂直:

$$\boldsymbol{a}_1 \cdot (k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{c}) = \boldsymbol{a}_2 \cdot (k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{c}) = 0. \tag{5.1.5}$$

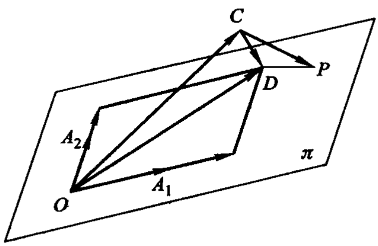

图 5–1

一般地, $\mathbf{R}^{3\times 1}$ 中两个列向量 $\boldsymbol{a} = (a_1, a_2, a_3)^{\mathrm{T}}$ 与 $\boldsymbol{b} = (b_1, b_2, b_3)^{\mathrm{T}}$ 的内积

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = (a_1, a_2, a_3) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \boldsymbol{a}^{\mathrm{T}} \boldsymbol{b}$$

可以写成行向量 $\boldsymbol{a}^{\mathrm{T}}$ 与列向量 $\boldsymbol{b}$ 的矩阵乘积. 因此, (k, b) 满足的条件 (5.1.5) 可以写成

$$\begin{pmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \end{pmatrix} (k\boldsymbol{a}_1 + b\boldsymbol{a}_2 - \boldsymbol{c}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{即} \quad \begin{pmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \end{pmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_2) \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1^{\mathrm{T}} \\ \boldsymbol{a}_2^{\mathrm{T}} \end{pmatrix} \boldsymbol{c},$$

也就是

$$A^{\mathrm{T}}AX = A^{\mathrm{T}}c. \tag{5.1.6}$$

(5.1.6) 可以由 (5.1.3) 两边同时左乘 A^{T} 得到. 如果 (5.1.3) 有解, 一定是 (5.1.6) 的解. 虽然本例题中的 (5.1.3) 无解, 但 (5.1.6) 的唯一解是 (5.1.3) 的最优近似解, 最优的标准就是使 (5.1.4) 中的 d^2 取最小值.

解 将 x, y 的三组对应值代入函数式 $kx + b = y$, 得到 k, b 满足的方程组 $AX = c$, 即

$$\begin{pmatrix} 3.7 & 1 \\ 4.0 & 1 \\ 4.2 & 1 \end{pmatrix} \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.6 \\ 0.35 \end{pmatrix}.$$

两边同时左乘 A^{T} 得方程组 $A^{\mathrm{T}}AX = A^{\mathrm{T}}c$, 即

$$\begin{pmatrix} 47.33 & 11.9 \\ 11.9 & 3 \end{pmatrix} \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} 7.2 \\ 1.85 \end{pmatrix},$$

解之得 $(k, b) = (-1.09, 4.95)$. 所求函数为

$$y = -1.09x + 4.95. \qquad \square$$

例 2 已知某种材料在生产过程中的废品率 y 与某种化学成分的含量百分比 x 有关. 以下是某工厂在生产过程中实测到的 x, y 的一些对应数据:

$x(\%)$	3.6	3.7	3.8	3.9	4.0	4.1	4.2
$y(\%)$	1.00	0.9	0.9	0.81	0.60	0.56	0.35

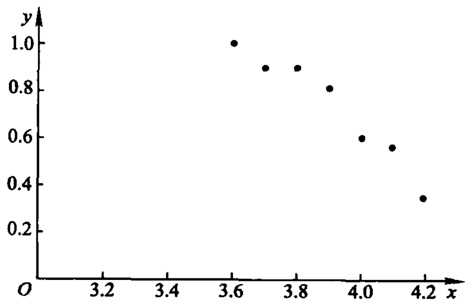

图 5–2

试给出 x, y 之间函数关系的一个近似公式.

解 以表中对应数据为坐标 (x_i, y_i) $(1 \leqslant i \leqslant 7)$ 画出 7 个点 (图 5–2).

观察发现, 这些点近似地在一条直线上. 可将 x, y 的函数关系用一次函数 $y = kx + b$ 近似地表示. 待定系数 k, b 满足方程组

$$\begin{cases} kx_1 + b = y_1, \\ \cdots\cdots\cdots\cdots \\ kx_7 + b = y_7, \end{cases} \tag{5.1.1$'$}$$

写成向量形式和矩阵形式, 分别是

$$k\boldsymbol{a}_1 + b\boldsymbol{a}_2 = \boldsymbol{c}, \tag{5.1.2$'$}$$

$$\boldsymbol{AX} = \boldsymbol{c}, \tag{5.1.3$'$}$$

其中

$$\boldsymbol{a}_1 = (x_1, \cdots, x_7)^{\mathrm{T}},\ \boldsymbol{a}_2 = (1, \cdots, 1)^{\mathrm{T}},\ \boldsymbol{c} = (y_1, \cdots, y_7)^{\mathrm{T}},$$

$$\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2),\ \boldsymbol{X} = (k, b)^{\mathrm{T}}.$$

由图 5–2 看出, 不可能有一条直线同时通过 7 个点 (x_i, y_i) $(1 \leqslant i \leqslant 7)$, 方程组 (5.1.1$'$) 没有精确解. 与例 1 类似, 仍求 (5.1.1$'$) 的近似解 (k, b) 使

$$d^2 = (kx_1 + b - y_1)^2 + \cdots + (kx_7 + b - y_7)^2$$

取最小值. 仍沿用例 1 的方法, 将 $\boldsymbol{AX} = \boldsymbol{c}$ 两边同时左乘 $\boldsymbol{A}^{\mathrm{T}}$ 得到 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$ 即

$$\begin{pmatrix} 106.75 & 27.3 \\ 27.3 & 7 \end{pmatrix}\begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} 19.675 \\ 5.12 \end{pmatrix}, \quad \text{解之得} \quad \begin{pmatrix} k \\ b \end{pmatrix} = \begin{pmatrix} -1.04643 \\ 4.8125 \end{pmatrix}.$$

因此, 与 7 个点 (x_i, y_i) 总体上最接近的直线是 $y = -1.04643x + 4.8125$. 将这条直线与 7 个点画在同一坐标系中, 如图 5–3 所示.

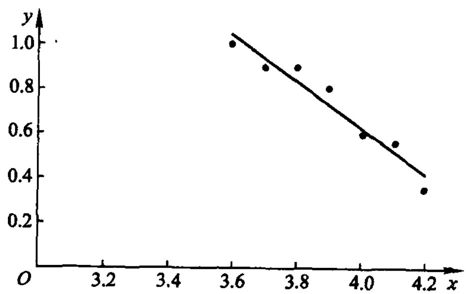

图 5–3

通过观察, 可以认为这条直线在总体上确实接近这 7 个点. □

例 1 与例 2 中求方程组 $\boldsymbol{AX}=\boldsymbol{c}$ 的最优近似解的方法称为**最小二乘法**. 它的基本目标是: 在方程组 $\boldsymbol{AX}=\boldsymbol{c}$ 没有精确解的情况下, 求它的近似解 $\boldsymbol{X}$ 使各个方程左右两边的差的平方和最小. 对于三个方程组成的方程组 $\boldsymbol{AX}=\boldsymbol{c}$, 例 1 中利用立体几何知识得到了它的最优近似解满足的方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$. 例 2 中将这个方法照搬到 $\boldsymbol{A}$ 是 7×2 矩阵的情形, 将方程组 $\boldsymbol{AX}=\boldsymbol{c}$ 仍归结为 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$ 来求解, 这在计算上并没有本质的困难. 但是例 1 中对于三维几何向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{c}$ 的论证并不能直接适用于例 2 中的 7 维空间中的向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{c}$. 例 2 中求出的 (k,b) 是否是最优近似解, 仍需要论证. 为此, 我们首先将三维几何空间中的内积推广到任意 n 维实向量空间 $\mathbf{R}^n$, 并证明它具有与三维空间中类似的性质. 特别是: 与子空间 π 外任一点 C 距离最近的点 $D\in\pi$ 仍是 π 的垂线段 CD 的垂足.

5.1.2 内积的推广

定义 5.1.1(内积) 设 $\mathbf{R}^n$ 是实数域上 n 数组空间. 对 $\mathbf{R}^n$ 中任意两个向量 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 与 $\boldsymbol{b}=(b_1,\cdots,b_n)$, 定义它们的**内积** (inner product)

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

定义了内积的实向量空间称为**欧氏空间** (Euclidean space).

如果将 $\mathbf{R}^n$ 中的数组向量写成列向量的形式, 则 $\boldsymbol{a}=(a_1,\cdots,a_n)^{\mathrm{T}}$ 与 $\boldsymbol{b}=(b_1,\cdots,b_n)^{\mathrm{T}}$ 的内积等于行向量 $\boldsymbol{a}^{\mathrm{T}}$ 与列向量 $\boldsymbol{b}$ 的矩阵乘积 $(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$.

实数组空间 $\mathbf{R}^n$ 按上述方式定义了内积, 成为欧氏空间. 对 $\mathbf{R}^n$ 的任何一个子空间 W 中的任意两个向量 $\boldsymbol{a},\boldsymbol{b}$ 也可以按 $\mathbf{R}^n$ 中的算法计算内积 $(\boldsymbol{a},\boldsymbol{b})$, 这就在 W 中定义了内积, W 也是欧式空间.

定理 5.1.1 $\mathbf{R}^n$ 中的内积具有性质:

对任意 $\boldsymbol{a},\boldsymbol{a}_1,\boldsymbol{b},\boldsymbol{b}_1\in\mathbf{R}^n$, $\lambda\in\mathbf{R}$, 有

(1) (双线性) $(\boldsymbol{a}+\boldsymbol{a}_1,\boldsymbol{b})=(\boldsymbol{a},\boldsymbol{b})+(\boldsymbol{a}_1,\boldsymbol{b})$, $(\boldsymbol{a},\boldsymbol{b}+\boldsymbol{b}_1)=(\boldsymbol{a},\boldsymbol{b})+(\boldsymbol{a},\boldsymbol{b}_1)$,
$(\lambda\boldsymbol{a},\boldsymbol{b})=\lambda(\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{a},\lambda\boldsymbol{b})$.

(2) (对称性) $(\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{b},\boldsymbol{a})$.

(3) (正定性) 当 $\boldsymbol{a}\neq\mathbf{0}$ 时, 有 $(\boldsymbol{a},\boldsymbol{a})>0$.

证明 (1) 由矩阵乘法法则容易直接验证 (略).

(2) 设 $\boldsymbol{a}=(a_1,\cdots,a_n)$, $\boldsymbol{b}=(b_1,\cdots,b_n)$. 则

$$(\boldsymbol{a},\boldsymbol{b})=a_1b_1+\cdots+a_nb_n=b_1a_1+\cdots b_na_n=(\boldsymbol{b},\boldsymbol{a}).$$

也可由矩阵乘法验证: $(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ 是一阶方阵, 等于自身的转置, 因此

$$(\boldsymbol{a},\boldsymbol{b})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b})^{\mathrm{T}}=\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}=(\boldsymbol{b},\boldsymbol{a}).$$

(3) 设 $\boldsymbol{a}=(a_1,\cdots,a_n)$. 则 $(\boldsymbol{a},\boldsymbol{a})=a_1^2+\cdots+a_n^2\geqslant 0$. 当 $\boldsymbol{a}\neq\boldsymbol{0}$ 时 $(\boldsymbol{a},\boldsymbol{a})>0$. □

对任意 $\boldsymbol{a}\in\mathbf{R}^n$, 记 $\boldsymbol{a}^2=(\boldsymbol{a},\boldsymbol{a})$. 定义 $|\boldsymbol{a}|=\sqrt{\boldsymbol{a}^2}$ 为 $\boldsymbol{a}$ 的**模** (module). $|\boldsymbol{a}|$ 是几何向量的长度的推广. 模 $|\boldsymbol{a}|=1$ 的向量 $\boldsymbol{a}$ 称为**单位向量** (unit vector). 任何一个非零向量 $\boldsymbol{a}$ 都可以乘常数$\dfrac{1}{|\boldsymbol{a}|}$ 得到同方向的单位向量 $\dfrac{1}{|\boldsymbol{a}|}\boldsymbol{a}$.

如果 $\boldsymbol{a},\boldsymbol{b}\in\mathbf{R}^n$ 的内积 $(\boldsymbol{a},\boldsymbol{b})=0$, 则称 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ **垂直** (perpendicular), 也称 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ **正交** (orthogonal), 记为 $\boldsymbol{a}\perp\boldsymbol{b}$.

例 3(勾股定理) 证明: 设 $\boldsymbol{a},\boldsymbol{b}\in\mathbf{R}^n$, 则 $\boldsymbol{a}\perp\boldsymbol{b}\ \Leftrightarrow\ |\boldsymbol{a}\pm\boldsymbol{b}|^2=|\boldsymbol{a}|^2+|\boldsymbol{b}|^2$.

证明 由定理 5.1.1 得

$$|\boldsymbol{a}\pm\boldsymbol{b}|^2=(\boldsymbol{a}\pm\boldsymbol{b})^2=\boldsymbol{a}^2+2(\boldsymbol{a},\boldsymbol{b})+\boldsymbol{b}^2=|\boldsymbol{a}|^2+|\boldsymbol{b}|^2+2(\boldsymbol{a},\boldsymbol{b}),$$

因此, $|\boldsymbol{a}\pm\boldsymbol{b}|^2=|\boldsymbol{a}|^2+|\boldsymbol{b}|^2\ \Leftrightarrow\ (\boldsymbol{a},\boldsymbol{b})=0\ \Leftrightarrow\ \boldsymbol{a}\perp\boldsymbol{b}$. □

现在可以将例 1 的解法推广到一般的实系数线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$. 同时也证明例 2 的算法的正确性.

例 4 对任意 $\boldsymbol{A}\in\mathbf{R}^{m\times n},\boldsymbol{b}\in\mathbf{R}^{m\times 1}$, 求证:

(1) $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的解 $\boldsymbol{X}_0$ 使 $|\boldsymbol{AX}-\boldsymbol{b}|^2$ 取到最小值 $|\boldsymbol{AX}_0-\boldsymbol{b}|^2$.

(2) 齐次线性方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}$ 同解. $\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\operatorname{rank}\boldsymbol{A}$.

(3) 方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 总是有解, 且当 $\operatorname{rank}\boldsymbol{A}=n$ 时有唯一解 $\boldsymbol{X}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.

证明 (1) 在图 5–1 中, $CD\perp$ 平面 π, P 是平面 π 内另外任意一点. $|CD|^2<|CP|^2$ 的理由是: 由于 $CD\perp DP$, 由勾股定理有 $|CP|^2=|CD|^2+|DP|^2>|CD|^2$.

这个推理可以推广到一般的实系数线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$.

设 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}_0=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. 对任意 $\boldsymbol{X}\in\mathbf{R}^{n\times 1}$, $\boldsymbol{AX}-\boldsymbol{b}=(\boldsymbol{AX}_0-\boldsymbol{b})+\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{X}_0)$, 且

$$\begin{aligned}(\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{X}_0))^{\mathrm{T}}(\boldsymbol{AX}_0-\boldsymbol{b})&=(\boldsymbol{X}-\boldsymbol{X}_0)^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{AX}_0-\boldsymbol{b})\\&=(\boldsymbol{X}-\boldsymbol{X}_0)^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}_0-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})=0.\end{aligned}$$

可见 $\boldsymbol{AX}_0-\boldsymbol{b}$ 与 $\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{X}_0)$ 相互垂直 (分别相当于图 5–1 中的 $\overrightarrow{CD}$ 与

$\overrightarrow{DP}$). 由勾股定理得

$$(\boldsymbol{AX}-\boldsymbol{b})^2=(\boldsymbol{AX}_0-\boldsymbol{b})^2+(\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{X}_0))^2\geqslant(\boldsymbol{AX}_0-\boldsymbol{b})^2.$$

(2) 分别记方程组 $\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}$ 解空间为 $V_{\boldsymbol{A}},V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}$.

首先, $\boldsymbol{X}\in V_{\boldsymbol{A}}\Leftrightarrow \boldsymbol{AX}=\boldsymbol{0}$, 两边左乘 $\boldsymbol{A}^{\mathrm{T}}$ 得 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}$, 这说明 $V_{\boldsymbol{A}}\subseteq V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}$.

反过来, $\boldsymbol{X}\in V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}\Rightarrow \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=0\Rightarrow(\boldsymbol{AX})^{\mathrm{T}}(\boldsymbol{AX})=0$. 记 $\boldsymbol{Y}=\boldsymbol{AX}=(y_1,\cdots,y_m)^{\mathrm{T}}$. 则 $\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}=y_1^2+\cdots+y_m^2=0\Rightarrow\boldsymbol{Y}=\boldsymbol{0}$. 可见 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}\Rightarrow\boldsymbol{AX}=\boldsymbol{0}$. $V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}\subseteq V_{\boldsymbol{A}}$.

这就证明了 $V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}=V_{\boldsymbol{A}}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{0}$ 与 $\boldsymbol{AX}=\boldsymbol{0}$ 同解, 它们的解空间的维数 $\dim V_{\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}}=n-\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$ 与 $\dim V_{\boldsymbol{A}}=n-\operatorname{rank}\boldsymbol{A}$ 相等, 因而 $\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\operatorname{rank}\boldsymbol{A}$.

(3) §4.4 例 5 中已经证明: 任意两个矩阵的乘积 $\boldsymbol{AB}$ 的秩 $\operatorname{rank}(\boldsymbol{AB})\leqslant\operatorname{rank}\boldsymbol{A}$. 由此可知方程组 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的增广矩阵 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A},\boldsymbol{b})$ 的秩

$$\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})\leqslant\operatorname{rank}\boldsymbol{A}^{\mathrm{T}}=\operatorname{rank}\boldsymbol{A}=\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}).$$

另一方面, 系数矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是增广矩阵 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})$ 的子矩阵, 当然有

$$\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})\geqslant\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}).$$

因此, 方程组 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的系数矩阵与增广矩阵的秩相等, 一定有解.

当 $\operatorname{rank}\boldsymbol{A}=n$ (即 $\boldsymbol{A}$ 的各列线性无关) 时, $\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\operatorname{rank}\boldsymbol{A}=n$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶可逆方阵, 方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AX}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 有唯一解 $\boldsymbol{X}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. □

算法 5.1.1(最小二乘法) 设 n 元实系数线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 的系数矩阵的秩为 n. 如果方程组无解, 则 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{X}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$ 的唯一解 $\boldsymbol{X}_0$ 是使 $|\boldsymbol{AX}-\boldsymbol{b}|^2$ 最小的近似解.

5.1.3 角度的计算公式

$\mathbf{R}^3$ 中的数组向量 $\boldsymbol{a}=(a_1,a_2,a_3),\boldsymbol{b}=(b_1,b_2,b_3)$ 可以看成几何向量, 当 $\boldsymbol{a},\boldsymbol{b}$ 都不为零时, 它们的夹角 θ 可以通过公式 $\cos\theta=\dfrac{(\boldsymbol{a},\boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|}$ 来计算. 且由 $|\cos\theta|\leqslant 1$ 知

$$(\boldsymbol{a},\boldsymbol{b})^2\leqslant|\boldsymbol{a}|^2|\boldsymbol{b}|^2,\text{即}(a_1b_1+a_2b_2+a_3b_3)^2\leqslant(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2).$$

在任意 n 维实向量空间 $\mathbf{R}^n$ 中, 自然希望用同样的公式 $\cos\theta=\dfrac{(\boldsymbol{a},\boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|}$ 来定义非零向量 $\boldsymbol{a},\boldsymbol{b}$ 的夹角 θ. 为此, 必须先证明这样计算出来的 $|\cos\theta|\leqslant 1$,

才能确定一个角 θ 来作为 $\boldsymbol{a},\boldsymbol{b}$ 的夹角. 也就是说, 必须证明 $\mathbf{R}^n$ 中任意 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 与 $\boldsymbol{b}=(b_1,\cdots,b_n)$ 满足不等式

$$(\boldsymbol{a},\boldsymbol{b})^2\leqslant|\boldsymbol{a}|^2|\boldsymbol{b}|^2,\text{即}(a_1b_1+\cdots+a_nb_n)^2\leqslant(a_1^2+\cdots+a_n^2)(b_1^2+\cdots+b_n^2).$$

这个不等式称为柯西 – 施瓦茨 (Cauchy–Schwarz) 不等式.

为了对 $\mathbf{R}^n$ 中的向量证明这个不等式, 先来看两个几何向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的夹角 θ 为什么满足 $|\cos\theta|\leqslant 1$. 作 $\overrightarrow{OA}=\boldsymbol{a},\overrightarrow{OB}=\boldsymbol{b}$, 如图 5–4.

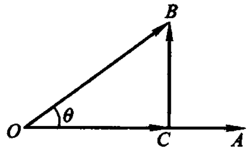

图 5–4

从 B 作 OA 的垂线 BC 交 OA 于 C. 则 $|\cos\theta|=\dfrac{|OC|}{|OB|}\leqslant 1$ 由 $|OC|\leqslant|OB|$ 决定. 而 $|OC|\leqslant|OB|$ 可以由勾股定理得到: $|OB|^2-|OC|^2=|CB|^2\geqslant 0$.

$\overrightarrow{OC}$ 是 $\boldsymbol{a}$ 的实数倍 $\lambda\boldsymbol{a}$, 满足条件 $\overrightarrow{OA}\perp(\overrightarrow{OB}-\lambda\overrightarrow{OA})$, 即 $(\boldsymbol{a},\boldsymbol{b}-\lambda\boldsymbol{a})=0$, 从而 $\lambda=\dfrac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}$, $\overrightarrow{CB}=\boldsymbol{b}-\dfrac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}$. 不等式 $|OB|^2-|OC|^2=|CB|^2\geqslant 0$ 就成为

$$\boldsymbol{b}^2-\left(\frac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}\right)^2=\left(\boldsymbol{b}-\frac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}\right)^2\geqslant 0.$$

这可以推广到任意 $\mathbf{R}^n$ 中的非零向量 $\boldsymbol{a},\boldsymbol{b}$.

定理 5.1.2(柯西 – 施瓦茨 (Cauchy–Schwarz) 不等式) 对任意实数 $a_1,\cdots,a_n,b_1,\cdots,b_n$, 不等式

$$(a_1^2+\cdots+a_n^2)(b_1^2+\cdots+b_n^2)\geqslant(a_1b_1+\cdots+a_nb_n)^2$$

成立, 其中等号成立当且仅当 $(a_1,\cdots,a_n)$ 与 $(b_1,\cdots,b_n)$ 成比例.

证明 记 $\boldsymbol{a}=(a_1,\cdots,a_n),\boldsymbol{b}=(b_1,\cdots,b_n)$. 当 $\boldsymbol{a}=\boldsymbol{0}$ 时定理的等号成立. 设 $\boldsymbol{a}\neq\boldsymbol{0}$, 则 $\boldsymbol{a}^2>0$. 由内积的正定性知不等式

$$\left(\boldsymbol{b}-\frac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}\right)^2\geqslant 0$$

成立. 将左边展开得

$$\boldsymbol{b}^2-2(\boldsymbol{b},\boldsymbol{a})\frac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}+\frac{(\boldsymbol{a},\boldsymbol{b})^2}{(\boldsymbol{a}^2)^2}\boldsymbol{a}^2=\boldsymbol{b}^2-\frac{(\boldsymbol{a},\boldsymbol{b})^2}{\boldsymbol{a}^2}\geqslant 0\ \Leftrightarrow\ \boldsymbol{a}^2\boldsymbol{b}^2\geqslant(\boldsymbol{a},\boldsymbol{b})^2,$$

即

$$(a_1^2+\cdots+a_n^2)(b_1^2+\cdots+b_n^2)\geqslant(a_1b_1+\cdots+a_nb_n)^2,$$

其中的等号成立 $\Leftrightarrow$ $\boldsymbol{b}-\dfrac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}=\mathbf{0}$ $\Leftrightarrow$ $\boldsymbol{b}=\dfrac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}\boldsymbol{a}$ $\Rightarrow$ $\boldsymbol{b}=(b_1,\cdots,b_n)$ 与 $\boldsymbol{a}=(a_1,\cdots,a_n)$ 成比例. 反过来, 当 $\boldsymbol{a},\boldsymbol{b}$ 成比例时, $\boldsymbol{b}=\lambda\boldsymbol{a}$, $\dfrac{(\boldsymbol{a},\boldsymbol{b})}{\boldsymbol{a}^2}=\lambda$, 定理的等号成立. □

现在可以对 $\mathbf{R}^n$ 中任意两个非零向量 $\boldsymbol{a},\boldsymbol{b}$ 定义夹角 $\langle\boldsymbol{a},\boldsymbol{b}\rangle=\arccos\dfrac{(\boldsymbol{a},\boldsymbol{b})}{|\boldsymbol{a}||\boldsymbol{b}|}\in[0,\pi]$. 特别地, 当 $(\boldsymbol{a},\boldsymbol{b})=0$ 时, 夹角$\langle\boldsymbol{a},\boldsymbol{b}\rangle=\dfrac{\pi}{2}$, $\boldsymbol{a}\perp\boldsymbol{b}$.

按照这个定义, 任意两个线性无关向量 $\boldsymbol{a},\boldsymbol{b}$ 之差的完全平方展开式 $(\boldsymbol{a}-\boldsymbol{b})^2=\boldsymbol{a}^2+\boldsymbol{b}^2-2(\boldsymbol{a},\boldsymbol{b})$ 就成为

$$|\boldsymbol{a}-\boldsymbol{b}|^2=|\boldsymbol{a}|^2+|\boldsymbol{b}|^2-2|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

如图 5–5. 这就是余弦定理.

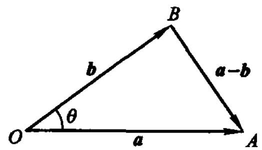

图 5–5

例 5(三角形不等式) 对任意 $\boldsymbol{a},\boldsymbol{b}\in\mathbf{R}^n$, 求证 $|\boldsymbol{a}|+|\boldsymbol{b}|\geqslant|\boldsymbol{a}+\boldsymbol{b}|$. 等号仅当 $\boldsymbol{a},\boldsymbol{b}$ 方向相同时成立.

证明 $|\boldsymbol{a}+\boldsymbol{b}|^2=(\boldsymbol{a}+\boldsymbol{b})^2=\boldsymbol{a}^2+2(\boldsymbol{a},\boldsymbol{b})+\boldsymbol{b}^2$.

由柯西 – 施瓦茨不等式, 有 $(\boldsymbol{a},\boldsymbol{b})\leqslant|\boldsymbol{a}||\boldsymbol{b}|$. 于是

$$|\boldsymbol{a}+\boldsymbol{b}|^2\leqslant|\boldsymbol{a}|^2+2|\boldsymbol{a}||\boldsymbol{b}|+|\boldsymbol{b}|^2=(|\boldsymbol{a}|+|\boldsymbol{b}|)^2\ \Rightarrow\ |\boldsymbol{a}+\boldsymbol{b}|\leqslant|\boldsymbol{a}|+|\boldsymbol{b}|.$$

等号成立 $\Leftrightarrow$ $|\boldsymbol{a}||\boldsymbol{b}|=(\boldsymbol{a},\boldsymbol{b})=|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle$ $\Leftrightarrow$ $\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle=1\Leftrightarrow$ $\langle\boldsymbol{a},\boldsymbol{b}\rangle=0$ $\Leftrightarrow$ $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 方向相同. □

习题 5.1

1. 在欧氏空间 $\mathbf{R}^4$ 中求向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的长度和夹角:

(1) $\boldsymbol{\alpha}=(1,3,2,-1),\boldsymbol{\beta}=(-4,2,-3,1)$;

(2) $\boldsymbol{\alpha}=(1,2,0,2),\boldsymbol{\beta}=(3,5,-1,1)$.

2. 求方程组

$$\begin{cases} 0.39x - 1.89y = 1, \\ 0.61x - 1.80y = 1, \\ 0.93x - 1.68y = 1, \\ 1.35x - 1.50y = 1 \end{cases}$$

的最小二乘解.

3. 在平面直角坐标系中, 作抛物线 $y = ax^2 + bx + c$, 使总体上尽可能接近如下 8 个点

$$(-4, 8.915), (-3, 2.587), (-2, -1.481), (-1, -3.384),$$

$$(0, -2.655), (1, -1.052), (2, 3.102), (3, 9.224).$$

4. 设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_n$ 是欧氏空间 $\mathbf{R}^n$ 的一组基, $\boldsymbol{a}, \boldsymbol{\beta} \in \mathbf{R}^n$. 求证:

(1) $\boldsymbol{\alpha} = \mathbf{0} \quad \Leftrightarrow \quad (\boldsymbol{\alpha}, \boldsymbol{\alpha}_i) = 0$ 对所有 $1 \leqslant i \leqslant n$ 成立;

(2) $\boldsymbol{\alpha} = \boldsymbol{\beta} \quad \Leftrightarrow \quad (\boldsymbol{\alpha}, \boldsymbol{\alpha}_i) = (\boldsymbol{\beta}, \boldsymbol{\alpha}_i)$ 对所有 $1 \leqslant i \leqslant n$ 成立.

5. 三维几何空间中两两成钝角的向量最多有几条? n 维欧氏空间 $\mathbf{R}^n$ 中两两成钝角的向量最多有几条? 试证明你的结论.

§5.2 正 交 化

5.2.1 标准正交基

例 1 实数域 $\mathbf{R}$ 上方程 $x_1 + x_2 + x_3 + x_4 = 0$ 的解空间 W 是欧氏空间 $\mathbf{R}^4$ 的子空间, $\boldsymbol{a}_1 = (1, -1, 0, 0)^{\mathrm{T}}, \boldsymbol{a}_2 = (0, 1, -1, 0)^{\mathrm{T}}, \boldsymbol{a}_3 = (0, 0, 1, -1)^{\mathrm{T}}$ 组成 W 的一组基 S. 将 W 中每个向量用它在 S 下的坐标表示. 试根据 $\boldsymbol{a}, \boldsymbol{b} \in W$ 在 S 下的坐标 $\boldsymbol{X} = (x_1, x_2, x_3)^{\mathrm{T}}$ 与 $\boldsymbol{Y} = (y_1, y_2, y_3)^{\mathrm{T}}$ 计算内积 $(\boldsymbol{a}, \boldsymbol{b})$.

解

$$\boldsymbol{a} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \boldsymbol{AX},$$

$$\boldsymbol{b} = y_1\boldsymbol{a}_1 + y_2\boldsymbol{a}_2 + y_3\boldsymbol{a}_3 = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \boldsymbol{AY},$$

其中 $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ 是依次以 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 为列排成的矩阵. 于是

$$(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = (\boldsymbol{AX})^{\mathrm{T}}(\boldsymbol{AY}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AY} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{GY},$$

其中

$$G=A^{\mathrm{T}}A=\begin{pmatrix}a_1^{\mathrm{T}}\\a_2^{\mathrm{T}}\\a_3^{\mathrm{T}}\end{pmatrix}(a_1,a_2,a_3)=\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix}$$

的 (i,j) 元 $a_i^{\mathrm{T}}a_j=(a_i,a_j)$ 是 W 的第 i 个基向量 a_i 与第 j 个向量 a_j 的内积. 于是

$$\begin{aligned}(a,b)=X^{\mathrm{T}}GY&=(x_1,x_2,x_3)\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}\\&=2x_1y_1+2x_2y_2+2x_3y_3-x_1y_2-x_2y_1-x_2y_3-x_3y_2.\end{aligned}$$

□

例 1 中由 a,b 的坐标 X,Y 计算内积的公式比较复杂, 是因为由基向量 a_1,a_2,a_3 的内积 (a_i,a_j) 组成的方阵 G 比较复杂. 而在 3 维几何空间中, 坐标为 $X=(x_1,x_2,x_3)^{\mathrm{T}},Y=(y_1,y_2,y_3)^{\mathrm{T}}$ 的两个向量 $a=x_1e_1+x_2e_2+x_3e_3$ 与 $b=y_1e_1+y_2e_2+y_3e_3$ 的内积

$$\begin{aligned}(a,b)&=((x_1e_1+x_2e_2+x_3e_3),(y_1e_1+y_2e_2+y_3e_3))\\&=\sum_{1\leqslant i,j\leqslant 3}x_i(e_i,e_j)y_j\\&=(x_1,x_2,x_3)\begin{pmatrix}(e_1,e_1)&(e_1,e_2)&(e_1,e_3)\\(e_2,e_1)&(e_2,e_2)&(e_2,e_3)\\(e_3,e_1)&(e_3,e_2)&(e_3,e_3)\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=x_1y_1+x_2y_2+x_3y_3.\end{aligned}$$

特别简单, 是因为所选取的基向量 e_1,e_2,e_3 是两两垂直的单位向量, 基向量的内积 $(e_i,e_j)=0$ (当 $i\neq j$) 或 1 (当 $i=j$), 以 (e_i,e_j) 为第 (i,j) 元组成的矩阵为单位矩阵 I, 因此由坐标 X,Y 计算内积的公式具有最简单的形式 $X^{\mathrm{T}}Y=x_1y_1+x_2y_2+x_3y_3$. 如果在例 1 的子空间 W 中也选取两两垂直的单位向量 b_1,b_2,b_3 组成基, 则以 (b_i,b_j) 为第 (i,j) 元的矩阵也是单位矩阵 I, 在这组基下的坐标为 X,Y 的两个向量 a,b 的内积也具有最简单的形式 $(a,b)=X^{\mathrm{T}}Y$.

定义 5.2.1 欧氏空间中两两垂直的非零向量组成的向量组称为**正交向量组** (orthogonal vectors). 如果某组基是正交向量组, 就称它为**正交基** (orthogonal basis). 由单位向量组成的正交向量组称为**标准正交向量组** (orthonormal vectors). 如果某组基是标准正交向量组, 就称它为**标准正交基** (orthonormal basis). □

将正交向量组中各向量分别除以它们各自的模, 就得到标准正交向量组.

以向量组 $T=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 中两两向量的内积 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$ 为第 (i,j) 元组成的 m 阶方阵 $\boldsymbol{G}$ 称为向量组 T 的**度量方阵** (metric matrix), 也称为格拉姆 (Gram) 方阵. 特别地, 将向量组中每个向量 $\boldsymbol{\alpha}_i$ 写成列向量 $\boldsymbol{a}_i$ 的形式, 以这些列向量为各列排成矩阵 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$, 则 $\boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=(s_{ij})_{m\times m}$ 的第 (i,j) 元 $\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{a}_j=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$ 就是向量组中第 i 个向量与第 j 个向量的内积, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 就是这个向量组的格拉姆方阵. 对任意实矩阵 $\boldsymbol{A}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 $\boldsymbol{A}$ 的列向量组的格拉姆方阵, 而 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的行向量组的格拉姆方阵.

引理 5.2.1 设 $\boldsymbol{G}=(s_{ij})_{m\times m}$ 是向量组 $T=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 的格拉姆方阵. 则

(1) $\boldsymbol{G}$ 是对称方阵: $\boldsymbol{G}^{\mathrm{T}}=\boldsymbol{G}$.

(2) 对 T 的任意两个线性组合 $\boldsymbol{\alpha}=x_1\boldsymbol{\alpha}_1+\cdots+x_m\boldsymbol{\alpha}_m$ 及 $\boldsymbol{\beta}=y_1\boldsymbol{\alpha}_1+\cdots+y_m\boldsymbol{\alpha}_m$, 有

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y},$$

其中 $\boldsymbol{X}=(x_1,\cdots,x_m)^{\mathrm{T}},\boldsymbol{Y}=(y_1,\cdots,y_m)^{\mathrm{T}}$. 特别地, 如果 T 是欧氏空间的一组基, 则 $\boldsymbol{X},\boldsymbol{Y}$ 就是 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 在这组基下的坐标.

(3) 对任意 $\boldsymbol{X}\in\mathbf{R}^{m\times 1}$, 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{X}\geqslant 0$. 特别地, 如果 T 线性无关, 则当 $\boldsymbol{X}\neq\mathbf{0}$ 时有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{X}>0$.

证明 (1) 对任意 $1\leqslant i<j\leqslant m$, 由内积的对称性有 $s_{ij}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_i)=s_{ji}$. 可见 $\boldsymbol{G}^{\mathrm{T}}=\boldsymbol{G}$, $\boldsymbol{G}$ 是对称方阵.

(2) 将 T 中每个向量 $\boldsymbol{\alpha}_i$ 写成列向量 $\boldsymbol{a}_i$, 排成矩阵 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$. 则 $\boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$, $\boldsymbol{\alpha}=\boldsymbol{A}\boldsymbol{X},\boldsymbol{\beta}=\boldsymbol{A}\boldsymbol{Y}$.

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{Y})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{Y}.$$

(3) 令 $\boldsymbol{\alpha}=\boldsymbol{A}\boldsymbol{X}$, 则由内积的正定性知 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{X}=(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$. 特别地, 当 T 线性无关且 $\boldsymbol{X}\neq\mathbf{0}$ 时 $\boldsymbol{\alpha}=\boldsymbol{A}\boldsymbol{X}\neq\mathbf{0}$, $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{X}=(\boldsymbol{\alpha},\boldsymbol{\alpha})>0$. □

定义 5.2.2 设 $\boldsymbol{S}$ 是 n 阶实对称方阵. 如果对任意 $\mathbf{0}\neq\boldsymbol{X}\in\mathbf{R}^{n\times 1}$ 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}>0$, 就称 $\boldsymbol{S}$ **是正定的** (positive definite), 记为 $\boldsymbol{S}>0$. 如果对任意 $\mathbf{0}\neq\boldsymbol{X}\in\mathbf{R}^{n\times 1}$ 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}\geqslant 0$, 就称 $\boldsymbol{S}$ **是半正定的** (positive semidefinite), 记为 $\boldsymbol{S}\geqslant 0$. 如果对任意 $\mathbf{0}\neq\boldsymbol{X}\in\mathbf{R}^{n\times 1}$ 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}<0$, 就称 $\boldsymbol{S}$ **是负定的** (negative definite), 记为 $\boldsymbol{S}<0$. 如果对任意 $\mathbf{0}\neq\boldsymbol{X}\in\mathbf{R}^{n\times 1}$ 有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}\leqslant 0$, 就称 $\boldsymbol{S}$ **是半负定的** (negative semidefinite), 记为 $\boldsymbol{S}\leqslant 0$. □

例 2 欧氏空间中任意向量组 $T=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 的格拉姆方阵 $\boldsymbol{G}$ 是半正定的. $\boldsymbol{G}$ 正定 $\Leftrightarrow T$ 线性无关.

对任意实矩阵 $\boldsymbol{A}$, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是半正定的. $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 正定 $\Leftrightarrow\boldsymbol{A}$ 的列向量组线性无关 $\Leftrightarrow\boldsymbol{A}^{\mathrm{T}}$ 的行向量组线性无关.

引理 5.2.2 (1) 实矩阵 $\boldsymbol{A}$ 的列向量组是正交向量组 $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是可逆对角矩阵.

(2) 实矩阵 $\boldsymbol{A}$ 的列向量组是标准正交向量组 $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$.

(3) n 阶实方阵 $\boldsymbol{A}$ 的列向量组是 $\mathbf{R}^{n\times 1}$ 的标准正交基 $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$ $\Leftrightarrow$ $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$ $\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{I}$ $\Leftrightarrow$ $\boldsymbol{A}$ 的行向量组是 $\mathbf{R}^{1\times n}$ 的标准正交基. □

定义 5.2.3 满足条件 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$ 即 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$ 的方阵 $\boldsymbol{A}$ 称为**正交方阵** (orthogonal matrix). □

n 阶正交方阵的列向量组与行向量组都是 $\mathbf{R}^n$ 的标准正交基.

例 3 求证: 正交向量组线性无关.

证法 1(几何证法) 设非零向量 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$ 两两垂直, 实数 $\lambda_1,\cdots,\lambda_m$ 满足

$$\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m=\mathbf{0},$$

等式两边与每个 $\boldsymbol{\alpha}_i$ $(1\leqslant i\leqslant m)$ 作内积得

$$\lambda_1(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_i)+\cdots+\lambda_m(\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_i)=0. \tag{5.2.1}$$

将 $(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_i)=0$ $(\forall\, j\neq i)$ 代入等式 (5.2.1) 得 $\lambda_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0$. 再由 $\boldsymbol{\alpha}_i\neq\mathbf{0}$ 知 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)>0$, 因此 $\lambda_i=0$. 对 $1\leqslant i\leqslant m$ 成立. 这证明了正交向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 线性无关.

证法 2(矩阵证法) 将正交向量组 $T=\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m\}$ 中各向量 $\boldsymbol{\alpha}_i$ 写成列向量 $\boldsymbol{X}_i$, 排成矩阵 $\boldsymbol{A}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)$. 则 $\boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 $\boldsymbol{T}$ 的格拉姆方阵.

$$\begin{aligned}T \text{ 是正交向量组} &\Leftrightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\operatorname{diag}(s_{11},\cdots,s_{mm})\ (s_{ii}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)>0)\\ &\Rightarrow\ m=\operatorname{rank}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\leqslant\operatorname{rank}\boldsymbol{A}\ \Rightarrow\ \boldsymbol{A}\text{ 的各列线性无关}\\ &\Rightarrow\ T\text{ 线性无关.}\end{aligned}$$ □

5.2.2 格拉姆 – 施密特 (Gram–Schmidt) 正交化方法

欧氏空间的任何一组基都可以改造成正交基, 再改造成标准正交基.

例 4 求 $x_1+x_2+x_3+x_4=0$ 的解空间 W 的一组标准正交基.

分析 §5.1 图 5–4 中, 将向量 $\boldsymbol{b}=\overrightarrow{OB}$ 减去适当的 $\overrightarrow{OC}=\lambda\boldsymbol{a}$ 就得到与 $\boldsymbol{a}$ 垂直的 $\boldsymbol{b}_1=\overrightarrow{CB}=\boldsymbol{b}-\lambda\boldsymbol{a}$. 图 5–1 中, 将 $\boldsymbol{c}=\overrightarrow{OC}$ 减去适当的 $k\boldsymbol{a}_1+b\boldsymbol{a}_2$ 可以得到与 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 都垂直的 $\overrightarrow{DC}=\boldsymbol{c}-k\boldsymbol{a}_1-b\boldsymbol{a}_2$. 一般地, 对于线性无关的 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k,\boldsymbol{a}_{k+1}$, 将 $\boldsymbol{a}_{k+1}$ 减去 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 的适当的线性组合可以得到与 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_k$ 都正交的 $\boldsymbol{b}_{k+1}=\boldsymbol{a}_{k+1}-\lambda_1\boldsymbol{a}_1-\cdots-\lambda_k\boldsymbol{a}_k$.

解 列向量 $\boldsymbol{a}_1=(1,-1,0,0)^{\mathrm{T}},\boldsymbol{a}_2=(0,1,-1,0)^{\mathrm{T}},\boldsymbol{a}_3=(0,0,1,-1)^{\mathrm{T}}$ 组成 W 的基 T. 先将它改造成一组正交基, 再改造成标准正交基.

选择适当的常数 λ 使 $\boldsymbol{\xi}_2=\boldsymbol{a}_2-\lambda\boldsymbol{a}_1$ 与 $\boldsymbol{a}_1$ 正交:

$$(\boldsymbol{a}_1,\boldsymbol{\xi}_2)=(\boldsymbol{a}_1,\boldsymbol{a}_2)-\lambda(\boldsymbol{a}_1,\boldsymbol{a}_1)=0\ \Rightarrow\ \lambda=\frac{(\boldsymbol{a}_1,\boldsymbol{a}_2)}{(\boldsymbol{a}_1,\boldsymbol{a}_1)}=-\frac{1}{2},$$

$$\boldsymbol{\xi}_2=\boldsymbol{a}_2+\frac{1}{2}\boldsymbol{a}_1=\left(\frac{1}{2},\frac{1}{2},-1,0\right)^{\mathrm{T}}.$$

用 $\boldsymbol{\xi}_2$ 代替 $\boldsymbol{a}_2$. 选择适当的常数 λ_1,λ_2 使 $\boldsymbol{\xi}_3=\boldsymbol{a}_3-\lambda_1\boldsymbol{a}_1-\lambda_2\boldsymbol{\xi}_2$ 与 $\boldsymbol{a}_1,\boldsymbol{\xi}_2$ 都正交:

$$\begin{cases}(\boldsymbol{a}_1,\boldsymbol{\xi}_3)=(\boldsymbol{a}_1,\boldsymbol{a}_3)-\lambda_1(\boldsymbol{a}_1,\boldsymbol{a}_1)=0,\\(\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=(\boldsymbol{\xi}_2,\boldsymbol{a}_3)-\lambda_2(\boldsymbol{\xi}_2,\boldsymbol{\xi}_2)=0\end{cases}$$

$$\Rightarrow\lambda_1=\frac{(\boldsymbol{a}_1,\boldsymbol{a}_3)}{(\boldsymbol{a}_1,\boldsymbol{a}_1)}=0,\qquad \lambda_2=\frac{(\boldsymbol{\xi}_2,\boldsymbol{a}_3)}{(\boldsymbol{\xi}_2,\boldsymbol{\xi}_2)}=-\frac{2}{3}.$$

$$\boldsymbol{\xi}_3=\boldsymbol{a}_3+\frac{2}{3}\boldsymbol{\xi}_2=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},-1\right)^{\mathrm{T}}.$$

$\boldsymbol{a}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 组成 W 的正交基, 分别除以各自的长度得到单位向量

$$\boldsymbol{b}_1=\frac{1}{\sqrt{2}}\boldsymbol{a}_1=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0,0\right)^{\mathrm{T}},\quad \boldsymbol{b}_2=\sqrt{\frac{2}{3}}\xi_2=\left(\frac{1}{\sqrt{6}},\frac{1}{\sqrt{6}},-\frac{2}{\sqrt{6}},0\right)^{\mathrm{T}},$$

$$\boldsymbol{b}_3=\frac{\sqrt{3}}{2}\xi_3=\left(\frac{\sqrt{3}}{6},\frac{\sqrt{3}}{6},\frac{\sqrt{3}}{6},-\frac{\sqrt{3}}{2}\right)^{\mathrm{T}}.$$

$\{\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3\}$ 就是 W 的标准正交基. □

算法 5.2.1(格拉姆 – 施密特 (Gram–Schmidt) 正交化) 设 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 是 $\mathbf{R}^{n\times 1}$ 中任意一组线性无关向量. 取 $\boldsymbol{b}_1=\boldsymbol{a}_1$, 并依次对 $2\leqslant k\leqslant m$ 取

$$\boldsymbol{b}_k=\boldsymbol{a}_k-\frac{(\boldsymbol{a}_k,\boldsymbol{b}_1)}{\boldsymbol{b}_1^2}\boldsymbol{b}_1-\cdots-\frac{(\boldsymbol{a}_k,\boldsymbol{b}_{k-1})}{\boldsymbol{b}_{k-1}^2}\boldsymbol{b}_{k-1},$$

得到正交向量组 $\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_m\}$.

再将其中各向量 $\boldsymbol{b}_k$ 分别乘常数 $\dfrac{1}{|\boldsymbol{b}_k|}$ 得到标准正交向量组 $\{\dfrac{1}{|\boldsymbol{b}_1|}\boldsymbol{b}_1,\cdots,\dfrac{1}{|\boldsymbol{b}_m|}\boldsymbol{b}_m\}$. □

注 将 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ 写成列向量, 则可用 §5.1 例 2 的算法求列向量 $\boldsymbol{\Lambda}_{k-1}=(\lambda_1,\cdots,\lambda_{k-1})^{\mathrm{T}}$ 使

$$\boldsymbol{b}_k=\boldsymbol{a}_k-\lambda_1\boldsymbol{a}_1-\cdots-\lambda_{k-1}\boldsymbol{a}_{k-1}=\boldsymbol{a}_k-\boldsymbol{A}_{k-1}\boldsymbol{\Lambda}_{k-1}$$

与 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{k-1}$ 都正交, 其中 $\boldsymbol{A}_{k-1}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{k-1})$. $\boldsymbol{\Lambda}_{k-1}$ 满足条件

$$\boldsymbol{A}_{k-1}^{\mathrm{T}}(\boldsymbol{a}_k-\boldsymbol{A}_{k-1}\boldsymbol{\Lambda}_{k-1})=\mathbf{0}\ \Rightarrow\ \boldsymbol{\Lambda}_{k-1}=(\boldsymbol{A}_{k-1}^{\mathrm{T}}\boldsymbol{A}_{k-1})^{-1}(\boldsymbol{A}_{k-1}^{\mathrm{T}}\boldsymbol{a}_k).$$

如果用已求出的两两正交的 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_{k-1}$ 代替 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{k-1}$, 也就是用 $\boldsymbol{B}_{k-1}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_{k-1})$ 代替 $\boldsymbol{A}_{k-1}$, 则 $\boldsymbol{B}_{k-1}^{\mathrm{T}}\boldsymbol{B}_{k-1}=\mathrm{diag}(\boldsymbol{b}_1^2,\cdots,\boldsymbol{b}_{k-1}^2)$,

$$\boldsymbol{\Lambda}_{k-1}=(\boldsymbol{B}_{k-1}^{\mathrm{T}}\boldsymbol{B}_{k-1})^{-1}(\boldsymbol{B}_{k-1}^{\mathrm{T}}\boldsymbol{a}_k)=\left(\frac{\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{a}_k}{\boldsymbol{b}_1^2},\cdots,\frac{\boldsymbol{b}_{k-1}^{\mathrm{T}}\boldsymbol{a}_k}{\boldsymbol{b}_{k-1}^2}\right)^{\mathrm{T}},$$

得到的

$$\boldsymbol{b}_k=\boldsymbol{a}_k-\boldsymbol{B}_{k-1}\boldsymbol{\Lambda}_{k-1}=\boldsymbol{a}_k-\frac{\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{a}_k}{\boldsymbol{b}_1^2}\boldsymbol{b}_1-\cdots-\frac{\boldsymbol{b}_{k-1}^{\mathrm{T}}\boldsymbol{a}_k}{\boldsymbol{b}_{k-1}^2}\boldsymbol{b}_{k-1},$$

与算法 5.2.1 的结果相同.

推论 5.2.1 欧氏空间 W 中任一个标准正交向量组都能扩充为 W 的标准正交基. 特别地, W 中任意单位向量能够扩充为 W 的标准正交基.

证明 先将标准向量组扩充为 W 的一组基, 再进行正交化和单位化得到标准正交基. □

5.2.3 矩阵的相合

将线性无关向量组 $S=\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m\}$ 改造成正交向量组 $T=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_m\}$ 的正交化过程可以通过矩阵的乘法和初等变换来实现. 具体做法如下:

设 S 中各向量 $\boldsymbol{a}_k$ 已写成列向量形式, 排成矩阵 $\boldsymbol{A}=(\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m)$ 来表示, T 中的每个向量 $\boldsymbol{b}_k$ 都是 $\boldsymbol{A}$ 的线性组合 $\boldsymbol{b}_k=\boldsymbol{A}\boldsymbol{X}_k$, 其中 $\boldsymbol{X}_k$ 是 m 维列向量, T 也可以由各 $\boldsymbol{b}_k=\boldsymbol{A}\boldsymbol{X}_k$ 排成矩阵 $\boldsymbol{B}=(\boldsymbol{b}_1,\cdots,\boldsymbol{b}_m)=\boldsymbol{A}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)=\boldsymbol{A}\boldsymbol{P}$ 来表示, 其中 $\boldsymbol{P}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)$ 是 m 阶方阵. 且当 T 线性无关时 $\boldsymbol{P}$ 的各列 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m$ 也线性无关, $\boldsymbol{P}$ 是可逆方阵.

设 S 的格拉姆方阵 $\boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. 则 T 的格拉姆方阵

$$\boldsymbol{H}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=(\boldsymbol{A}\boldsymbol{P})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}$$

只要能找到可逆方阵 $\boldsymbol{P}$ 将 S 的格拉姆方阵 $\boldsymbol{G}$ 变到 $\boldsymbol{H}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}=\boldsymbol{I}$, 则 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{P}$ 是标准正交向量组.

定义 5.2.4 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶方阵. 如果存在可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$, 则称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相合 (congruent). □

只要求出可逆方阵 $\boldsymbol{P}$ 将 $\boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 相合到 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}=\boldsymbol{I}$, 则 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{P}$ 的列向量组就是由 $\boldsymbol{A}$ 的列向量组改造成的标准正交向量组.

如果 $\boldsymbol{P}_i$ 是初等方阵, 则 $\boldsymbol{G}\to\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}_i$ 就是先对 $\boldsymbol{G}$ 作某个初等行变换 $\boldsymbol{G}\to\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{G}$, 再作相应的初等列变换 $\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{G}\to\boldsymbol{P}_i^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}_i$. 每个可逆方阵 $\boldsymbol{P}$ 都可以分解为有限个初等方阵的乘积, 相合变换 $\boldsymbol{G}\to\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}$ 可以分解为有限次初等行变换和初等列变换来完成.

例 4 解法 2　仍将 W 的一组基 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$ 改造成标准正交基, 其中

$$\boldsymbol{a}_1=(1,-1,0,0)^{\mathrm{T}},\boldsymbol{a}_2=(0,1,-1,0)^{\mathrm{T}},\boldsymbol{a}_3=(0,0,1,-1)^{\mathrm{T}},$$

从而

$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\-1&1&0\\0&-1&1\\0&0&-1\end{pmatrix},\quad \boldsymbol{G}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}2&-1&0\\-1&2&-1\\0&-1&2\end{pmatrix}.$$

将 $\boldsymbol{G}$ 与单位矩阵排成 3×6 阶矩阵 $(\boldsymbol{G},\boldsymbol{I})$, 经过一系列初等行变换并对前三列进行相应的列变换, 将矩阵的前 3 列变成单位矩阵 $\boldsymbol{I}$, 则后 3 列变成 $\boldsymbol{P}^{\mathrm{T}}$:

$$(\boldsymbol{G},\boldsymbol{I})=\begin{pmatrix}2&-1&0&1&0&0\\-1&2&-1&0&1&0\\0&-1&2&0&0&1\end{pmatrix}\xrightarrow[\frac{1}{2}(1)+(2)]{\frac{1}{2}(1)+(2)}\begin{pmatrix}2&0&0&1&0&0\\0&\frac{3}{2}&-1&\frac{1}{2}&1&0\\0&-1&2&0&0&1\end{pmatrix}$$

$$\xrightarrow[\frac{2}{3}(2)+(3)]{\frac{2}{3}(2)+(3)}\begin{pmatrix}2&0&0&1&0&0\\0&\frac{3}{2}&0&\frac{1}{2}&1&0\\0&0&\frac{4}{3}&\frac{1}{3}&\frac{2}{3}&1\end{pmatrix}$$

$$\xrightarrow[\frac{1}{\sqrt{2}}(1),\sqrt{\frac{2}{3}}(2),\frac{\sqrt{3}}{2}(3)]{\frac{1}{\sqrt{2}}(1),\sqrt{\frac{2}{3}}(2),\frac{\sqrt{3}}{2}(3)}\begin{pmatrix}1&0&0&\frac{\sqrt{2}}{2}&0&0\\0&1&0&\frac{\sqrt{6}}{6}&\frac{\sqrt{6}}{3}&0\\0&0&1&\frac{\sqrt{3}}{6}&\frac{\sqrt{3}}{3}&\frac{\sqrt{3}}{2}\end{pmatrix}.$$

可知 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}=\boldsymbol{I}$ 对

$$\boldsymbol{P}=\begin{pmatrix}\frac{\sqrt{2}}{2}&\frac{\sqrt{6}}{6}&\frac{\sqrt{3}}{6}\\0&\frac{\sqrt{6}}{3}&\frac{\sqrt{3}}{3}\\0&0&\frac{\sqrt{3}}{2}\end{pmatrix}$$

成立.

$$
B = AP = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{3} \\ 0 & 0 & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} \\ 0 & -\frac{\sqrt{6}}{3} & \frac{\sqrt{3}}{6} \\ 0 & 0 & -\frac{\sqrt{3}}{2} \end{pmatrix}
$$

的 3 列组成 W 的标准正交基. □

在例 4 解法 2 中, 从 $(\boldsymbol{G}, \boldsymbol{I})$ 出发所作的每次行变换必然紧跟着同样的列变换, 在每个箭头上下方标注的内容完全相同, 分别表示行变换和相应的列变换. 为了简单起见, 以后可以只在箭头上方标出行变换, 略去箭头下方的标注, 记住按照上方的标注作相应的列变换就行了.

引理 5.2.3 矩阵的相合具有如下性质:

(1) 自反性: $\boldsymbol{A}$ 与自身相合.

(2) 对称性: $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相合 $\Rightarrow$ $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相合.

(3) 传递性: $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相合, $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 相合 $\Rightarrow$ $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 相合.

证明 (1) $\boldsymbol{A} = \boldsymbol{I}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{I}$.

(2) $\boldsymbol{B} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} \Rightarrow \boldsymbol{A} = (\boldsymbol{P}^{-1})^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}^{-1}$.

(3) $\boldsymbol{B} = \boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}_1, \boldsymbol{C} = \boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}_2 \Rightarrow \boldsymbol{C} = (\boldsymbol{P}_1\boldsymbol{P}_2)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P}_1\boldsymbol{P}_2)$. □

引理 5.2.4 对称方阵相合之后仍是对称方阵. 反对称方阵相合之后仍是反对称方阵.

证明 设 $\boldsymbol{S}, \boldsymbol{A} \in \mathbf{R}^{n\times n}$, $\boldsymbol{X} \in \mathbf{R}^{n\times m}$, 且 $\boldsymbol{S}^{\mathrm{T}} = \boldsymbol{S}$, $\boldsymbol{A}^{\mathrm{T}} = -\boldsymbol{A}$. 则

$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X})^{\mathrm{T}} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$. $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X})^{\mathrm{T}} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X} = -\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$.

这说明 $\boldsymbol{S}_1 = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 是 m 阶对称方阵, $\boldsymbol{A}_1 = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是 m 阶反对称方阵.

特别地, 当 $\boldsymbol{X}$ 是 n 阶可逆方阵时, 对称方阵 $\boldsymbol{S}$ 相合到对称方阵 $\boldsymbol{S}_1$, 反对称方阵 $\boldsymbol{A}$ 相合到反对称方阵 $\boldsymbol{A}_1$. □

注 如果引理 5.2.4 证明中的 $\boldsymbol{X}$ 是 n 维列向量, 则 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 是一阶反对称方阵, 只能是数 0. 这说明: 反对称方阵 $\boldsymbol{A}$ 定义的 $Q(\boldsymbol{X}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 恒等于 0.

习题 5.2

1. 设 W 是 $\mathbf{R}^3$ 中过点 $(0,0,0), (1,2,2), (3,4,0)$ 的平面, 求点 $A(5,0,0)$ 到平面 W 的最短距离.

2. 求齐次方程 $x_1-x_2+x_3-x_4=0$ 在实数域上的解空间的一组标准正交基.

3. (1) 求齐次线性方程组 $\begin{cases}x_1+x_2+x_3+x_4+x_5=0,\\ 2x_1+3x_2+5x_3+8x_4=0\end{cases}$ 的解空间的一组标准正交基;

(2) 将 $(1,1,1,1,1),(2,3,5,8,0)$ 扩充为 $\mathbf{R}^5$ 的一组标准正交基.

4. 求 $\mathbf{R}^4$ 中向量

$$\boldsymbol{a}_1=(1,1,1,-3),\ \boldsymbol{a}_2=(0,1,2,-3),\ \boldsymbol{a}_3=(0,0,1,-1),\ \boldsymbol{a}_4=(0,0,0,1)$$

的格拉姆方阵 $\boldsymbol{G}$. 求可逆方阵 $\boldsymbol{P}$ 使其将 $\boldsymbol{G}$ 相合到单位矩阵, 即 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{P}=\boldsymbol{I}$.

§5.3 二 次 型

5.3.1 二次型的配方

例 1 求以下实函数的最大值或最小值.

(1) $f(x,y,z)=x^2+5y^2+5z^2+4xy-2xz+4yz$;

(2) $f(x,y,z)=x^2+5y^2+5z^2+4xy-2xz-2yz$.

解 (1) 先将 y,z 看成常数, 将 $f(x,y,z)$ 看成一个自变量 x 的二次多项式, 配方得

$$\begin{aligned}f(x,y,z)&=x^2+2x(2y-z)+5y^2+4yz+5z^2\\&=(x+2y-z)^2-(2y-z)^2+5y^2+4yz+5z^2\\&=(x+2y-z)^2+y^2+8yz+4z^2.\end{aligned}$$

再将其中不含 x 的项看成以 y 为自变量 (z 看成常数) 的二次多项式, 配方得

$$f(x,y,z)=(x+2y-z)^2+(y+4z)^2-12z^2.$$

令

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\begin{pmatrix}x+2y-z\\y+4z\\z\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}x\\y\\z\end{pmatrix},\quad \text{其中}\quad \boldsymbol{P}=\begin{pmatrix}1&2&-1\\0&1&4\\0&0&1\end{pmatrix}\ \text{可逆},$$

则

$$g(x',y',z')=f(x,y,z)=x'^2+y'^2-12z'^2$$

是 x',y',z' 的函数, 且 (x',y',z') 取遍 $\mathbf{R}^3 \Leftrightarrow (x,y,z)$ 取遍 $\mathbf{R}^3$.

当 x' 取遍 $\mathbf{R}$ 时 $g(x',0,0)=x'^2$ 取遍 $[0,+\infty)$, 当 z' 取遍 $\mathbf{R}$ 时 $g(0,0,z')=-12z'^2$ 取遍 $(-\infty,0]$. 可见 $f(x,y,z)$ 的值域是 $(-\infty,+\infty)$, 既没有最大值也没有最小值.

(2) 与第 (1) 小题类似地配方得

$$f(x,y,z)=(x+2y-z)^2+(y+z)^2+3z^2=g(x',y',z')=x'^2+y'^2+3z'^2,$$

其中

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}x\\y\\z\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}1&2&-1\\0&1&1\\0&0&1\end{pmatrix}\ \text{可逆}.$$

同样地有: (x',y',z') 取遍 $\mathbf{R}^3\Leftrightarrow(x,y,z)$ 取遍 $\mathbf{R}^3$.

易见当(x',y',z')取遍$\mathbf{R}^3$ 时有$g(x',y',z')\geqslant 0$ 且当x' 取遍$\mathbf{R}$ 时$g(x',0,0)=x'^2$ 取遍 $[0,+\infty)$. 因此 $g(x',y',z')=f(x,y,z)$ 的值域是 $[0,+\infty)$, 没有最大值. 当且仅当 $(x',y',z')=(0,0,0)$ 即 $(x,y,z)=(0,0,0)$ 时 $g(x',y',z')=f(x,y,z)$ 取最小值 0. □

例 1 中的函数 $f(x,y,z)$ 都是自变量的二次多项式, 并且只有二次项而不含一次项与常数项. 这样的函数称为二次齐次函数, 也称为二次型. 一般地, 有

定义 5.3.1 n 元二次函数 $Q(x_1,\cdots,x_n)$ 如果只含二次项而不含一次项与常数项, 具有形式

$$Q(x_1,\cdots,x_n)=\sum_{1\leqslant i\leqslant j\leqslant n}a_{ij}x_ix_j,$$

就称为**二次型** (quadratic form). □

二次型 $Q(x_1,\cdots,x_n)$ 可以看成以 n 维数组向量 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 为自变量的一元函数 $Q(\boldsymbol{X})$, 也就是看成 n 维向量空间 $\mathbf{R}^n$ 上的函数.

微积分的重要思想是将函数 $f(x_1,\cdots,x_n)$ 在任一点 $P_0(c_1,\cdots,c_n)$ 附近用自变量增量 $\Delta x_i=x_i-c_i$ $(1\leqslant i\leqslant n)$ 的一次函数近似代替, 通过这个一次函数的变化速度来研究 f 在 P_0 附近的变化速度. 如果要进一步研究函数 f 在这一点 P_0 是否有极大值或极小值, 用一次函数就不够了, 需要用二次函数来近似代替 f, 通过二次函数的二次项部分组成的二次型是否有最大值或最小值来研究 f 的极大值极小值.

对一般的二次型 $Q(X)$, 不容易判定它是否有最大值或最小值. 但如果能够像例 1 那样, 通过可逆线性代换 $\boldsymbol{Y}=(y_1,\cdots,y_n)^{\mathrm{T}}=\boldsymbol{PX}$ ($\boldsymbol{P}$ 是可逆方阵) 将 $Q(\boldsymbol{X})$ 化为只含平方项 $d_iy_i^2$, 不含交叉项 $a_{ij}y_iy_j$ $(i\neq j)$ 的二次型

$$Q_1(\boldsymbol{Y})=Q_1(y_1,\cdots,y_n)=d_1y_1^2+\cdots+d_ny_n^2,$$

就容易根据各平方项系数 d_i 的正负号判定 $Q_1(\boldsymbol{Y})$ 是否有最大值或最小值. 容易看出:

引理 5.3.1 设二次型 $Q(\boldsymbol{Y})=d_1y_1^2+\cdots+d_ny_n^2$. 则

(1) $d_i\geqslant 0\ (\forall\ 1\leqslant i\leqslant n)\quad\Leftrightarrow\quad Q(\boldsymbol{Y})$ 有最小值 0.

(2) $d_i>0\ (\forall\ 1\leqslant i\leqslant n)\quad\Leftrightarrow$ 当且仅当 $\boldsymbol{Y}=\boldsymbol{0}$ 时 $Q(\boldsymbol{Y})$ 取最小值 0.

(3) $d_i\leqslant 0\ (\forall\ 1\leqslant i\leqslant n)\quad\Leftrightarrow\quad Q(\boldsymbol{Y})$ 有最大值 0.

(4) $d_i<0\ (\forall\ 1\leqslant i\leqslant n)\quad\Leftrightarrow$ 当且仅当 $\boldsymbol{Y}=\boldsymbol{0}$ 时 $Q(\boldsymbol{Y})$ 取最大值 0. □

定义 5.3.2 设存在 n 阶可逆方阵 $\boldsymbol{P}$ 使

$$Q(\boldsymbol{X})=Q_1(\boldsymbol{Y})=d_1y_1^2+\cdots+d_ny_n^2$$

对 $\boldsymbol{Y}=(y_1,\cdots,y_n)^{\mathrm{T}}=\boldsymbol{PX}$ 成立, 则称 $Q_1(\boldsymbol{Y})$ 是 $Q(\boldsymbol{X})$ 的**标准形** (canonical form). □

定义 5.3.3 设 $Q(\boldsymbol{X})$ 是 $\mathbf{R}^n$ 上的二次型. 如果 $Q(\boldsymbol{X})\geqslant 0$ 对所有的 $\boldsymbol{X}\in\mathbf{R}^n$ 成立, 就称 Q 是**半正定的** (positive semidefinite). 如果 $Q(\boldsymbol{X})>0$ 对所有的 $\boldsymbol{0}\neq\boldsymbol{X}\in\mathbf{R}^n$ 成立, 就称 Q 是**正定的** (positive definite). 如果 $Q(\boldsymbol{X})\leqslant 0$ 对所有的 $\boldsymbol{X}\in\mathbf{R}^n$ 成立, 就称 Q 是**半负定的** (negative semidefinite). 如果 $Q(\boldsymbol{X})<0$ 对所有的 $\boldsymbol{0}\neq\boldsymbol{X}\in\mathbf{R}^n$ 成立, 就称 Q 是**负定的** (negative definite). □

由于 $\boldsymbol{Y}=\boldsymbol{PX}$ 是可逆线性变换, $\boldsymbol{X}$ 取遍 $\mathbf{R}^n\Leftrightarrow\boldsymbol{Y}$ 取遍 $\mathbf{R}^n$. $Q(\boldsymbol{X})$ 与它的标准形 $Q_1(\boldsymbol{Y})$ 有同样的值域, 同为正定、半正定、负定或半负定. 由引理 5.3.1 及定义 5.3.2 得

引理 5.3.2 设 $Q(\boldsymbol{X})$ 的标准形为 $Q_1(\boldsymbol{Y})=d_1y_1^2+\cdots+d_ny_n^2$. 则:

$Q(\boldsymbol{X})$ 正定 $\Leftrightarrow d_i>0$ 对所有 $1\leqslant i\leqslant n$ 成立.

$Q(\boldsymbol{X})$ 半正定 $\Leftrightarrow d_i\geqslant 0$ 对所有 $1\leqslant i\leqslant n$ 成立.

$Q(\boldsymbol{X})$ 负定 $\Leftrightarrow d_i<0$ 对所有 $1\leqslant i\leqslant n$ 成立.

$Q(\boldsymbol{X})$ 半负定 $\Leftrightarrow d_i\leqslant 0$ 对所有 $1\leqslant i\leqslant n$ 成立. □

例 2 求二次实函数的最大值或最小值.

(1) $f(x,y)=x^2+5y^2+4xy-2x+4y+5$;

(2) $f(x,y)=x^2+4y^2+4xy-2x+4y+5$.

解 (1) $f(x,y)=(x+2y-1)^2+(y+4)^2-12\geqslant -12$.

当 $y=-4$ 时 $f(x,-4)=(x-9)^2-12$ 可取遍 $[-12,+\infty)$.

可见, 函数 f 没有最大值, 当 $(x,y)=(9,-4)$ 时取最小值 -12.

(2) $f(x,y)=(x+2y-1)^2+8y+4$.

取 $x=-2y+1$ 得 $f(-2y+1,y)=8y+4$, 当 y 取遍 $\mathbf{R}=(-\infty,+\infty)$ 时 $f(-2y+1,y)$ 也取遍 $(-\infty,+\infty)$. 因此 f 没有最大值, 也没有最小值. □

例 2(1) 的二次项部分 $x^2+4xy+5y^2=(x+2y)^2+y^2$ 正定，可以保证 $f(x,y)$ 有最小值. 例 2(2) 的二次项部分 $x^2+4xy+4y^2=(x+2y)^2$ 半正定，不能保证 $f(x,y)$ 有最小值.

注 一般地，n 元函数 $f(x_1,\cdots,x_n)$ 如果能够在某一点 $P_0(c_1,\cdots,c_n)$ 附近用自变量增量 $\Delta x_i=x_i-c_i$ 的二次函数

$$g(\Delta x_1,\cdots,\Delta x_n)=f(c_1,\cdots,c_n)+\sum_{i=1}^{n}a_i\Delta x_i+\sum_{1\leqslant i\leqslant j\leqslant n}a_{ij}\Delta x_i\Delta x_j$$

近似代替，使误差是 $\rho=\sqrt{(\Delta x_1)^2+\cdots+(\Delta x_n)^2}$ 的高阶无穷小，则 f 在 P_0 有极值的必要条件是 g 中所有的一次项系数 $a_i=\dfrac{\partial f}{\partial x_i}(P_0)=0\ (1\leqslant i\leqslant n)$，且当 g 的二次项部分 Q 正定时 f 在 P_0 取极小值，当 Q 负定时 f 在 P_0 取极大值.

例 3 将二次型 $Q(x_1,x_2,x_3)=x_1x_2+x_1x_3+x_2x_3$ 化成标准形. 并求出值域.

解 这个二次型没有平方项，不能像例 1 那样直接配方. 先作可逆线性变换 $x_1=u_1+u_2$, $x_2=u_1-u_2$, $x_3=u_3$, 也就是

$$\begin{pmatrix}u_1\\u_2\\u_3\end{pmatrix}=\begin{pmatrix}\frac{1}{2}x_1+\frac{1}{2}x_2\\ \frac{1}{2}x_1-\frac{1}{2}x_2\\ x_3\end{pmatrix}=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&0\\ \frac{1}{2}&-\frac{1}{2}&0\\ 0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

得

$$\begin{aligned}Q(x_1,x_2,x_3)&=(u_1+u_2)(u_1-u_2)+(u_1+u_2)u_3+(u_1-u_2)u_3\\&=u_1^2-u_2^2+2u_1u_3=(u_1+u_3)^2-u_2^2-u_3^2\\&=Q_1(y_1,y_2,y_3)=y_1^2-y_2^2-y_3^2,\end{aligned}$$

其中

$$\begin{cases}y_1=u_1+u_3=\dfrac{1}{2}x_1+\dfrac{1}{2}x_2+x_3,\\ y_2=u_2=\dfrac{1}{2}x_1-\dfrac{1}{2}x_2,\\ y_3=u_3=x_3.\end{cases}$$

所求值域为 $(-\infty,+\infty)$. □

5.3.2 用矩阵相合化简二次型

二次型可以用矩阵乘积来表示, 通过矩阵相合来化成标准形.

例 1(1) 解法 2

$$\begin{aligned} f(x,y,z) &= x^2+5y^2+5z^2+4xy-2xz+4yz \\ &= x(x+2y-z)+y(2x+5y+2z)+z(-x+2y+5z) \\ &= (x,y,z)\begin{pmatrix} x+2y-z \\ 2x+5y+2z \\ -x+2y+5z \end{pmatrix} = (x,y,z)\begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}, \end{aligned}$$

其中

$$\boldsymbol{X} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \boldsymbol{S} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & 5 \end{pmatrix}.$$

设 $\boldsymbol{Y} = (x',y',z')^{\mathrm{T}} = \boldsymbol{PX}$, 其中 $\boldsymbol{P}$ 是可逆方阵. 则 $\boldsymbol{X} = \boldsymbol{P}^{-1}\boldsymbol{Y}$. 代入 $f(\boldsymbol{X}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{SX}$ 得

$$f(\boldsymbol{X}) = (\boldsymbol{P}^{-1}\boldsymbol{Y})^{\mathrm{T}}\boldsymbol{S}(\boldsymbol{P}^{-1}\boldsymbol{Y}) = \boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{P}^{-1})^{\mathrm{T}}\boldsymbol{SP}^{-1}\boldsymbol{Y} = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{S}_1\boldsymbol{Y},$$

其中 $\boldsymbol{S}_1 = (\boldsymbol{P}^{-1})^{\mathrm{T}}\boldsymbol{SP}^{-1}$ 与 $\boldsymbol{S}$ 相合. 设法求可逆方阵 $\boldsymbol{K}$ 将 $\boldsymbol{S}$ 相合到对角矩阵 $\boldsymbol{K}^{\mathrm{T}}\boldsymbol{SK} = \boldsymbol{D} = \mathrm{diag}(d_1,d_2,d_3)$, 则

$$f(\boldsymbol{X}) = \boldsymbol{Y}^{\mathrm{T}}\boldsymbol{DY} = d_1y_1^2 + d_2y_2^2 + d_3y_3^2$$

就是二次型 $f(\boldsymbol{X})$ 经过可逆线性变换 $\boldsymbol{Y} = \boldsymbol{K}^{-1}\boldsymbol{X}$ 化成的标准形. 具体操作过程如下:

$$\boldsymbol{S} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & 5 \end{pmatrix} \xrightarrow{-2(1)+(2),\,(1)+(3)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 4 & 4 \end{pmatrix} \xrightarrow{-4(2)+(3)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -12 \end{pmatrix} = \boldsymbol{D}.$$

(注意, 我们只在箭头上方标出了所作的初等行变换, 但在每次初等行变换之后都作了相应的初等列变换.) 每次初等行变换和列变换可以通过左乘和右乘初等方阵实现:

$$\boldsymbol{S} \to \boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{SP}_1 \to \boldsymbol{P}_2^{\mathrm{T}}(\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{SP}_1)\boldsymbol{P}_2 = (\boldsymbol{P}_1\boldsymbol{P}_2)^{\mathrm{T}}\boldsymbol{S}(\boldsymbol{P}_1\boldsymbol{P}_2) = \boldsymbol{D},$$

其中

$$\boldsymbol{P}_1=\begin{pmatrix}1&-2&1\\0&1&0\\0&0&1\end{pmatrix},\quad \boldsymbol{P}_2=\begin{pmatrix}1&0&0\\0&1&-4\\0&0&1\end{pmatrix}.$$

取

$$\begin{aligned}\boldsymbol{P}&=(\boldsymbol{P}_1\boldsymbol{P}_2)^{-1}=(\boldsymbol{P}_2^{-1})(\boldsymbol{P}_1^{-1})\\&=\begin{pmatrix}1&0&0\\0&1&4\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&-1\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&2&-1\\0&1&4\\0&0&1\end{pmatrix},\\ \boldsymbol{Y}&=\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\boldsymbol{PX}=\begin{pmatrix}1&2&-1\\0&1&4\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+2y-z\\y+4z\\z\end{pmatrix}.\end{aligned}$$

则 $Q(\boldsymbol{X})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{DY}=x'^2+y'^2-12z'^2$ 的值域为 $(-\infty,+\infty)$, 没有最大值或最小值. □

注 由于对角矩阵是对称方阵, 能够相合于对角矩阵的方阵 $\boldsymbol{S}$ 只能是对称方阵. 如果一开始将二次型 $f(x,y,z)=x^2+5y^2+5z^2+4xy-2xz+4yz$ 写成

$$f(x,y,z)=x(x+4y-2z)+y(5y+4z)+5z^2=(x,y,z)\begin{pmatrix}1&4&-2\\0&5&4\\0&0&5\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix},$$

由于其中的方阵

$$\boldsymbol{A}=\begin{pmatrix}1&4&-2\\0&5&4\\0&0&5\end{pmatrix}$$

不是对称方阵, 就不能将它相合到对角矩阵 $\boldsymbol{D}$.

一般地, 对任意 n 阶实对称方阵 $\boldsymbol{S}=(s_{ij})_{n\times n}$ 及 n 维列向量 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$,

$$Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{SX}=\sum_{i=1}^{n}s_{ii}x_i^2+\sum_{1\leqslant i<j\leqslant n}2s_{ij}x_ix_j$$

是 $\boldsymbol{X}$ 的二次型.

反过来, 每个 n 元二次型

$$Q(\boldsymbol{X})=Q(x_1,\cdots,x_n)=\sum_{i=1}^{n}a_{ii}x_i^2+\sum_{1\leqslant i<j\leqslant n}a_{ij}x_ix_j$$

可以写成矩阵乘积 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$, 其中 $\boldsymbol{S}=(s_{ij})_{n\times n}$ 是由 Q 唯一决定的实对称方阵:

$$s_{ii}=a_{ii}\ (\forall\ 1\leqslant i\leqslant n),\quad s_{ij}=\frac{1}{2}a_{ij}\ (\forall\ 1\leqslant i<j\leqslant n).$$

求 n 阶可逆方阵 $\boldsymbol{K}$ 将 $\boldsymbol{S}$ 相合到对角矩阵 $\boldsymbol{K}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{K}=\boldsymbol{D}=\mathrm{diag}(d_1,\cdots,d_n)$. 则 $\boldsymbol{S}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{P}$ 对 $\boldsymbol{P}=\boldsymbol{K}^{-1}$ 成立. 取 $\boldsymbol{Y}=\boldsymbol{P}\boldsymbol{X}=(y_1,\cdots,y_n)^{\mathrm{T}}$. 则 $Q(\boldsymbol{X})$ 被化成标准形

$$Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{P}\boldsymbol{X}=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{Y}=d_1y_1^2+\cdots+d_ny_n^2.$$

例 3 解法 2 记 $\boldsymbol{X}=(x_1,x_2,x_3)^{\mathrm{T}}$. 则

$$Q(x_1,x_2,x_3)=x_1x_2+x_1x_3+x_2x_3=\boldsymbol{X}^{\mathrm{T}}\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix}\boldsymbol{X}.$$

$$\boldsymbol{S}=\begin{pmatrix}0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix}\xrightarrow{(2)+(1)}\begin{pmatrix}1&\frac{1}{2}&1\\\frac{1}{2}&0&\frac{1}{2}\\1&\frac{1}{2}&0\end{pmatrix}$$

$$\xrightarrow{-\frac{1}{2}(1)+(2),-(1)+(3)}\begin{pmatrix}1&0&0\\0&-\frac{1}{4}&0\\0&0&-1\end{pmatrix}\xrightarrow{2(2)}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}=\boldsymbol{D},$$

其中 $\boldsymbol{D}=\boldsymbol{P}_3^{\mathrm{T}}\boldsymbol{P}_2^{\mathrm{T}}\boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$ 对初等方阵

$$\boldsymbol{P}_1=\begin{pmatrix}1&0&0\\1&1&0\\0&0&1\end{pmatrix},\ \boldsymbol{P}_2=\begin{pmatrix}1&-\frac{1}{2}&-1\\0&1&0\\0&0&1\end{pmatrix},\ \boldsymbol{P}_3=\begin{pmatrix}1&0&0\\0&2&0\\0&0&1\end{pmatrix}$$

成立. 取

$$P=(P_1P_2P_3)^{-1}=P_3^{-1}P_2^{-1}P_1^{-1}$$

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

则

$$Q(X)=X^{\mathrm{T}}SX=Y^{\mathrm{T}}P_3^{\mathrm{T}}P_2^{\mathrm{T}}P_1^{\mathrm{T}}SP_1P_2P_3Y=Y^{\mathrm{T}}DY=y_1^2-y_2^2-y_3^3$$

对

$$Y=\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=PX=\begin{pmatrix}\frac{1}{2}x_1+\frac{1}{2}x_2+x_3\\-\frac{1}{2}x_1+\frac{1}{2}x_2\\x_3\end{pmatrix}$$

成立. □

习题 5.3

1. 用配方法化下列二次型为标准形:

(1) $Q(x_1,x_2,x_3)=x_1^2+x_2^2-2x_3^2+2x_1x_2-4x_2x_3$;

(2) $Q(x_1,x_2,x_3)=-4x_1x_2+2x_1x_3+2x_2x_3$;

(3) $Q(x_1,x_2,x_3)=x_1^2+5x_2^2-4x_3^2+2x_1x_3-4x_2x_3$;

(4) $Q(x_1,x_2,x_3,x_4)=x_1x_2+x_2x_3+x_3x_4+x_4x_1$.

2. 写出第 1 题中各二次型的实对称方阵. 利用方阵相合求它们的标准形.

3. 设 x,y,z 是任意实数, A,B,C 是任意三角形的三个内角. 证明不等式

$$x^2+y^2+z^2\geqslant 2xy\cos A+2xz\cos B+2yz\cos C.$$

4. 求实函数 f 的最大值或最小值:

(1) $f(x,y)=x^2+4xy+5y^2-6x+8y+9$;

(2) $f(x,y,z)=x^2+5y^2+6z^2+4xy+2xz+6yz-4x+6y+8z+9$.

§5.4 实对称方阵相合标准形

5.4.1 实对称方阵相合标准形

算法 5.4.1(求实对称方阵相合标准形) 任意实对称方阵 $\boldsymbol{S}=(s_{ij})_{n\times n}$ 可以通过如下步骤相合到对角矩阵.

基本步骤 将 $\boldsymbol{S}$ 相合到准对角矩阵 $\boldsymbol{S}_1=\begin{pmatrix} d_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}_{n-1}\end{pmatrix}$, 其中 $\boldsymbol{B}_{n-1}$ 是 $n-1$ 阶实对称方阵.

情况 1 $s_{11}\neq 0$, 将 $\boldsymbol{S}$ 的第 1 列的适当常数倍加到其余各列, 并作相应的初等行变换, 完成基本步骤:

$$\boldsymbol{S}\xrightarrow[-s_{1j}s_{11}^{-1}(1)+(j),\ \forall\ 2\leqslant j\leqslant n]{-s_{1j}s_{11}^{-1}(1)+(j),\ \forall\ 2\leqslant j\leqslant n}\begin{pmatrix} s_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B}_{n-1}\end{pmatrix}.$$

情况 2 $s_{11}=0$. 如果 $\boldsymbol{S}$ 的第 1 行元素全为 0, 则第 1 列元素也全为 0, 基本步骤已经完成. 否则, 第 1 行与第 1 列某一对元素 $s_{1j}=s_{j1}\neq 0$. 将 $\boldsymbol{S}$ 的第 j 列加到第 1 列, 第 j 行加到第 1 行, 将 S 的第 $(1,1)$ 元变成 $2s_{1j}\neq 0$, 化为情况 1.

这样, 在任何情况下都能将 $\boldsymbol{S}$ 相合到基本步骤所说的准对角矩阵 $\boldsymbol{S}_1$. 再对 $\boldsymbol{B}_{n-1}$ 重复基本步骤. 不断重复下去, 经过有限步之后将 $\boldsymbol{S}$ 相合到对角矩阵 $\boldsymbol{D}=\mathrm{diag}(d_1,\cdots,d_n)$.

通过两列互换和两行互换 $\xrightarrow[(i,j)]{(i,j)}$ 可以将 $\boldsymbol{D}$ 的任何两个对角元 d_i,d_j 的位置互换. 这样就可以通过相合变换将 $\boldsymbol{D}$ 的对角元任意重新排列顺序, 使最前面 p 个 $d_1,\cdots,d_p$ 是正实数, 其后的 q 个是负实数, 最后 $n-p-q$ 个是 0. 并且可以用对角矩阵

$$\boldsymbol{D}_1=\mathrm{diag}\left(\frac{1}{\sqrt{|d_1|}},\cdots,\frac{1}{\sqrt{|d_{p+q}|}},1,\cdots,1\right)$$

将 $\boldsymbol{D}$ 进一步相合到

$$\boldsymbol{D}_1^{\mathrm{T}}\boldsymbol{D}\boldsymbol{D}_1=\mathrm{diag}(\boldsymbol{I}_{(p)},-\boldsymbol{I}_{(q)},\boldsymbol{O}_{(n-p-q)}).$$

其中的正实数都化为 1, 负实数都化为 -1. □

定理 5.4.1 任何一个实对称方阵 $\boldsymbol{S}$ 都能相合到唯一的标准形

$$\boldsymbol{\varLambda}=\begin{pmatrix}\boldsymbol{I}_{(p)} & & \\ & -\boldsymbol{I}_{(q)} & \\ & & \boldsymbol{O}_{(n-p-q)}\end{pmatrix}.$$ □

定义 5.4.1 实对称方阵 $\boldsymbol{S}$ 的相合标准形中 1 的个数 p 称为 $\boldsymbol{S}$ 的**正惯性指数** (positive index of inertia), -1 的个数 q 称为 $\boldsymbol{S}$ 的**负惯性指数** (negative index of inertia), 它们的和 $p+q$ 就是方阵 $\boldsymbol{S}$ 的秩 $\operatorname{rank}\boldsymbol{S}$, 差 $p-q$ 称为 $\boldsymbol{S}$ 的**符号差** (signature). □

正惯性指数与负惯性指数由 $\boldsymbol{S}$ 唯一决定, 定理 5.4.1 称为**西尔维斯特惯性定律** (Sylvester law of inertia). 这个定律的证明在本教材中不作为教学要求. 感兴趣的读者可参见附录 6.

推论 5.4.1 实二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 可以通过可逆线性代换化为**规范形** (normal form):

$$Q_1(\boldsymbol{Y})=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2,$$

其中 p,q 分别是 Q 的方阵 $\boldsymbol{S}$ 的正惯性指数和负惯性指数. □

实对称方阵 $\boldsymbol{S}$ 的正惯性指数 p、负惯性指数 q、秩 $\operatorname{rank}\boldsymbol{S}=p+q$、符号差 $p-q$ 也分别称为二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 的正惯性指数、负惯性指数、秩、符号差.

实二次型 $Q(\boldsymbol{X})$ 的标准形 $Q_1(\boldsymbol{Y})=d_1y_1^2+\cdots+d_ny_n^2$ 并不唯一, 但其中的正系数 d_i 的个数 p、负系数 d_i 的个数 q 以及等于 0 的系数 d_i 的个数 $n-p-q$ 都是唯一的, p,q 分别是 Q 的正惯性指数和负惯性指数. 将各个非零的 d_i 对应的 y_i 替换成 $z_i=\sqrt{|d_i|}y_i$, 并将各项顺序重新排列, 就得到 $Q(\boldsymbol{X})$ 的规范形 $z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_{p+q}^2$.

由引理 5.3.2 得

推论 5.4.2 设 $Q(\boldsymbol{X})$ 的规范形为

$$Q_1(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{Y}=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2,$$

则

$$Q(\boldsymbol{X})\text{ 正定 }\Leftrightarrow p=n,\ Q_1(\boldsymbol{Y})=y_1^2+\cdots+y_n^2\quad\Leftrightarrow\boldsymbol{\Lambda}=\boldsymbol{I}..$$

$$Q(\boldsymbol{X})\text{ 半正定}\Leftrightarrow q=0,\ Q_1(\boldsymbol{Y})=y_1^2+\cdots+y_p^2\quad\Leftrightarrow\boldsymbol{\Lambda}=\begin{pmatrix}\boldsymbol{I}_{(p)} & \\ & \boldsymbol{O}_{(n-p)}\end{pmatrix}.$$

$$Q(\boldsymbol{X})\text{ 负定 }\Leftrightarrow q=n,\ Q_1(\boldsymbol{Y})=-y_1^2+\cdots-y_n^2\quad\Leftrightarrow\boldsymbol{\Lambda}=-\boldsymbol{I}.$$

$$Q(\boldsymbol{X})\text{ 半负定}\Leftrightarrow p=0,\ Q_1(\boldsymbol{Y})=-y_1^2-\cdots-y_q^2\quad\Leftrightarrow\boldsymbol{\Lambda}=\begin{pmatrix}-\boldsymbol{I}_{(q)} & \\ & \boldsymbol{O}_{(n-q)}\end{pmatrix}.$$

□

例 1 求二次型 $Q(x_1,\cdots,x_n)=(x_1-x_2)^2+\cdots+(x_i-x_{i+1})^2+\cdots+(x_{n-1}-x_n)^2+(x_n-x_1)^2$ 的正惯性指数和负惯性指数.

解　对任意 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 有 $Q(\boldsymbol{X})\geqslant 0$, 可见 Q 半正定, 负惯性指数 $q=0$. 通过可逆线性变换 $\boldsymbol{Y}=\boldsymbol{P}\boldsymbol{X}$ 将 $Q(\boldsymbol{X})$ 化成的规范形 $Q_1(\boldsymbol{Y})$ 也半正定, 具有形状 $Q_1(\boldsymbol{Y})=y_1^2+\cdots+y_p^2$.

$Q_1(\boldsymbol{Y})=0\Leftrightarrow y_1=\cdots=y_p=0\ \Leftrightarrow\ \boldsymbol{Y}=(0,\cdots,0,y_{p+1},\cdots,y_n)^{\mathrm{T}}$.

$Q(\boldsymbol{X})=Q_1(\boldsymbol{Y})=0\ \Leftrightarrow\ \boldsymbol{X}=\boldsymbol{P}^{-1}\boldsymbol{Y}=y_{p+1}\boldsymbol{K}_{p+1}+\cdots+y_n\boldsymbol{K}_n$, 其中 $\boldsymbol{K}_{p+1},\cdots,\boldsymbol{K}_n$ 是可逆方阵 $\boldsymbol{P}^{-1}$ 的后 $n-p$ 列, 线性无关, $Q(\boldsymbol{X})=0$ 的解集合是 $\boldsymbol{K}_{p+1},\cdots,\boldsymbol{K}_n$ 的全体实系数线性组合组成的 $n-p$ 维子空间.

实际上, $Q(\boldsymbol{X})=0$ 的解集合为 $\{\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}\mid x_1=x_2=\cdots=x_n\}$, 是 1 维子空间, 因此 $n-p=1$, $p=n-1$.

结论是: Q 的正惯性指数为 $n-1$, 负惯性指数为 0. □

5.4.2　正定方阵的判定

由正定二次型和正定方阵的定义知道: 实对称方阵 $\boldsymbol{S}$ 正定 $\Leftrightarrow$ 二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 正定. 设 $\boldsymbol{S}_1$ 与 $\boldsymbol{S}$ 相合, 存在可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}=\boldsymbol{S}_1$. 通过可逆线性变换 $\boldsymbol{Y}=\boldsymbol{P}\boldsymbol{X}$ 将正定二次型 $Q(\boldsymbol{X})$ 化为 $Q_1(\boldsymbol{Y})$, 则 $Q_1(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{S}_1\boldsymbol{Y}$ 仍正定, 因此 $\boldsymbol{S}_1$ 仍正定. 可见, 方阵的正定性经过相合变换仍保持. 类似地, 负定 (半正定、半负定) 的实对称方阵经过相合之后仍然负定 (半正定、半负定).

根据推论 5.4.2, $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 正定 $\Leftrightarrow Q(\boldsymbol{X})$ 的标准形 $Q_1(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{Y}$ 的矩阵 $\boldsymbol{\varLambda}=\boldsymbol{I}$. 而 $Q(\boldsymbol{X})$ 与 $Q_1(\boldsymbol{Y})$ 对应的实对称方阵 $\boldsymbol{S}$ 与 $\boldsymbol{\varLambda}$ 相合. 由此得到:

引理 5.4.1　实对称方阵 $\boldsymbol{S}>0\Leftrightarrow\boldsymbol{S}$ 与单位矩阵 $\boldsymbol{I}$ 相合 $\Leftrightarrow$ 存在可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{S}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$. □

推论 5.4.3　实对称方阵 $\boldsymbol{S}>0\Rightarrow|\boldsymbol{S}|>0$.

证明　$\boldsymbol{S}>0\Rightarrow$ 存在可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{S}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$, $|\boldsymbol{S}|=|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}|=|\boldsymbol{P}^{\mathrm{T}}||\boldsymbol{P}|=|\boldsymbol{P}|^2>0$. □

推论 5.4.4　设正整数 $k\leqslant n$, $\boldsymbol{S}_k$ 是 $\boldsymbol{S}$ 的前 k 行与前 k 列交叉位置组成的方阵. 则 $\boldsymbol{S}>0\Rightarrow|\boldsymbol{S}_k|>0$.

证明　正定实对称方阵 $\boldsymbol{S}=(s_{ij})_{n\times n}$ 对应的二次型

$$Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}=\sum_{i=1}^{n}s_{ii}x_i^2+\sum_{1\leqslant i<j\leqslant n}2s_{ij}x_ix_j$$

正定. 对实数域 $\mathbf{R}$ 上任意 k 维列向量 $\boldsymbol{\xi}=(x_1,\cdots,x_k)^{\mathrm{T}}$, 添加 $n-k$ 个 0 分量扩充成 $\boldsymbol{X}=(x_1,\cdots,x_k,0,\cdots,0)^{\mathrm{T}}\neq\mathbf{0}$. 则

$$f(\boldsymbol{\xi})=Q(\boldsymbol{X})=\sum_{i=1}^{k}s_{ii}x_i^2+\sum_{1\leqslant i<j\leqslant k}2s_{ij}x_ix_j=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{S}_k\boldsymbol{\xi}>0,$$

这说明 k 阶实对称方阵 $\boldsymbol{S}_k$ 决定的二次型 $f(\boldsymbol{\xi})=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{S}_k\boldsymbol{\xi}$ 正定, 由推论 5.4.3 知行列式 $|\boldsymbol{S}_k|>0$. □

推论 5.4.4 中 $\boldsymbol{S}$ 的子方阵 $\boldsymbol{S}_k$ 的行列式 $|\boldsymbol{S}_k|$ 称为 $\boldsymbol{S}$ 的顺序主子式. 推论 5.4.4 指出: $\boldsymbol{S}$ 正定 $\Rightarrow$ 所有的顺序主子式 $|\boldsymbol{S}_k|>0$ 是 $\boldsymbol{S}$ 正定的必要条件. 我们指出, 这个条件也是充分的.

定理 5.4.2　n 阶实对称方阵 $\boldsymbol{S}$ 正定 $\Leftrightarrow$ $\boldsymbol{S}$ 的全体顺序主子式 $|\boldsymbol{S}_k|>0$ $(\forall\ 1\leqslant k\leqslant n)$. □

定理的证明在附录 6 中给出. 先利用它来解决问题.

例 2　给出正定实二次型 $Q(x,y)=ax^2+bxy+cy^2$ 的系数 a,b,c 满足的充分必要条件.

解

$$Q(x,y)=(x,y)\boldsymbol{S}\begin{pmatrix}x\\y\end{pmatrix},\quad \boldsymbol{S}=\begin{pmatrix}a&\dfrac{b}{2}\\\dfrac{b}{2}&c\end{pmatrix},$$

$$Q>0\ \Leftrightarrow\ \boldsymbol{S}>0\ \Leftrightarrow\ \begin{cases}|\boldsymbol{S}_1|=a>0,\\|\boldsymbol{S}_2|=ac-\dfrac{b^2}{4}>0.\end{cases}$$

Q 正定的充分必要条件为 $a>0$ 且 $b^2-4ac<0$. □

例 3　已知实二次型 $Q(x,y,z)=\lambda(x^2+y^2+z^2)-2xy-2xz+2yz$.

(1) λ 取什么值时 $Q(x,y,z)$ 正定?

(2) λ 取什么值时 $Q(x,y,z)$ 负定?

证明　$Q(x,y,z)$ 的矩阵为

$$\boldsymbol{S}=\begin{pmatrix}\lambda&-1&-1\\-1&\lambda&1\\-1&1&\lambda\end{pmatrix}.$$

(1) Q 正定 $\Leftrightarrow\boldsymbol{S}>0\Leftrightarrow\boldsymbol{S}$ 的顺序主子式 $|\boldsymbol{S}_k|>0,\ \forall\ k=1,2,3$

$$\Leftrightarrow\begin{cases}|\boldsymbol{S}_1|=\lambda>0,\\|\boldsymbol{S}_2|=\lambda^2-1>0,\\|\boldsymbol{S}_3|=(\lambda-1)^2(\lambda+2)>0\end{cases}\quad\Leftrightarrow\quad\lambda>1.$$

结论: Q 正定 $\Leftrightarrow\lambda>1$.

(2) Q 负定 $\Leftrightarrow \boldsymbol{S}<0 \Leftrightarrow -\boldsymbol{S}>0 \Leftrightarrow -\boldsymbol{S}$ 的所有顺序主子式 $|(-\boldsymbol{S})_k|>0$

$$\Leftrightarrow \begin{cases} |(-\boldsymbol{S})_1| = -\lambda > 0, \\ |(-\boldsymbol{S})_2| = \lambda^2 - 1 > 0, \\ |(-\boldsymbol{S})_3| = -(\lambda-1)^2(\lambda+2) > 0 \end{cases} \qquad \Leftrightarrow \lambda < -2.$$

结论: Q 负定 $\Leftrightarrow$ $\lambda < -2$. □

习题　5.4

1. 求实对称方阵在实相合下的标准形:

(1) $\begin{pmatrix} 2 & 4 & -2 \\ 4 & 5 & -1 \\ -2 & -1 & 0 \end{pmatrix}$;　(2) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$;　(3) $\begin{pmatrix} \boldsymbol{O}_{(n)} & \boldsymbol{I}_{(n)} \\ \boldsymbol{I}_{(n)} & \boldsymbol{I}_{(n)} \end{pmatrix}$.

2. 设 $Q = ax_1^2 + bx_2^2 + ax_3^2 + 2cx_1x_3$, 问 a, b, c 满足什么条件时, Q 正定.

3. 给出负定二次型 $Q(x_1, \cdots, x_n) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的方阵 $\boldsymbol{A}$ 的顺序主子式满足的充要条件.

4. 在二次型 $Q(x,y,z) = \lambda(x^2+y^2+z^2) + 3y^2 - 4xy - 2xz + 4yz$ 中, 问:

(1) λ 取什么值时 Q 正定;

(2) λ 取什么值时 Q 半负定;

(3) λ 取什么值使 Q 为实一次多项式的完全平方.

5. 设 $\boldsymbol{A}$ 是 n 阶实对称正定矩阵, 证明 $\boldsymbol{A}^{-1}$ 正定.

6. 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶实对称正定矩阵, 证明: 如果 $\boldsymbol{A}-\boldsymbol{B}$ 正定, 则 $\boldsymbol{B}^{-1}-\boldsymbol{A}^{-1}$ 亦正定.

7. 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶实对称阵, 其中 $\boldsymbol{A}$ 正定. 证明: 当实数 t 充分大后, $t\boldsymbol{A}+\boldsymbol{B}$ 亦正定.

8. 求下列二次型的秩、正惯性指数和符号差.

(1) $Q(x,y,z) = x^2+y^2+z^2-xy-xz-yz$;

(2) $Q(x_1,\cdots,x_n) = (x_1-\overline{x})^2+\cdots+(x_n-\overline{x})^2$, 其中 $\overline{x} = \dfrac{1}{n}(x_1+\cdots+x_n)$.

附录 6　惯性定律与正定性判定

1. 西尔维斯特 (Sylvester) 惯性定律的证明

定理 5.4.1(西尔维斯特惯性定律)　设 n 阶实对称方阵 $\boldsymbol{S}$ 通过两个不同的可逆方阵 $\boldsymbol{P}, \boldsymbol{P}_1$ 相合到标准形

$$\boldsymbol{A} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P} = \mathrm{diag}(\boldsymbol{I}_{(p)}, -\boldsymbol{I}_{(q)}, \boldsymbol{O}), \quad \boldsymbol{A}_1 = \boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{S}\boldsymbol{P}_1 = \mathrm{diag}(\boldsymbol{I}_{(p_1)}, -\boldsymbol{I}_{(q_1)}, \boldsymbol{O}),$$

则 $\boldsymbol{A} = \boldsymbol{A}_1$.

证明　由于 $\boldsymbol{P},\boldsymbol{P}_1$ 可逆, $\boldsymbol{S}$ 与 $\boldsymbol{\Lambda},\boldsymbol{\Lambda}_1$ 都相抵, 因此 $p+q=\operatorname{rank}\boldsymbol{\Lambda}=\operatorname{rank}\boldsymbol{S}=\operatorname{rank}\boldsymbol{\Lambda}_1=p_1+q_1$. 只要能够证明 $p=p_1$, 则一定有 $q=q_1$, 从而 $\boldsymbol{\Lambda}=\boldsymbol{\Lambda}_1$.

如果 $p=n$, 则 $\boldsymbol{\Lambda}$ 正定 $\Rightarrow \boldsymbol{S}$ 正定 $\Rightarrow \boldsymbol{\Lambda}_1$ 正定 $\Rightarrow \boldsymbol{\Lambda}=\boldsymbol{I}=\boldsymbol{\Lambda}_1$.

如果 $p=0$, 则 $\boldsymbol{\Lambda}$ 半负定 $\Rightarrow \boldsymbol{S}$ 与 $\boldsymbol{\Lambda}_1$ 半负定 $\Rightarrow p_1=p=0\Rightarrow q=q_1\Rightarrow \boldsymbol{\Lambda}=\boldsymbol{\Lambda}_1$.

当 $0<p<n$ 时, $\boldsymbol{S}$ 决定的二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 既不正定也不半负定. 对任意 $\boldsymbol{Y}=(y_1,\cdots,y_n)^{\mathrm{T}}$, 有

$$\begin{aligned}Q(\boldsymbol{PY})&=(\boldsymbol{PY})^{\mathrm{T}}\boldsymbol{S}(\boldsymbol{PY})=\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{SP})\boldsymbol{Y}\\&=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{\Lambda Y}=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2.\end{aligned}$$

让 $\boldsymbol{Y}$ 取遍所有的 $(y_1,\cdots,y_p,0,\cdots,0)^{\mathrm{T}}$, 让 $\boldsymbol{X}=\boldsymbol{PY}=y_1\boldsymbol{P}_1+\cdots+y_p\boldsymbol{P}_p$ 取遍可逆方阵 $\boldsymbol{P}$ 的前 p 列 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_p$ 的所有的线性组合, 也就是取遍这 p 列在 $\mathbf{R}^{n\times1}$ 中生成的 p 维子空间 W_+, 则当 $\mathbf{0}\neq\boldsymbol{PY}\in W_+$ 时

$$Q(\boldsymbol{PY})=y_1^2+\cdots+y_p^2>0,$$

我们称二次型在子空间 W_+ 上正定. 类似地, 如果让 $\boldsymbol{Y}$ 取遍所有的 $(0,\cdots,0,y_{p+1},\cdots,y_n)^{\mathrm{T}}$, $\boldsymbol{X}=\boldsymbol{PY}$ 取遍 $\boldsymbol{P}$ 的后 $n-p$ 列在 $\mathbf{R}^{n\times1}$ 中生成的 $n-p$ 维子空间 $W_{-,0}$, 则 $Q(\boldsymbol{PY})=-y_{p+1}^2-\cdots-y_{p+q}^2\leqslant 0$, 我们称二次型在子空间 $W_{-,0}$ 上半负定.

同理, 设可逆方阵 $\boldsymbol{P}_1$ 的前 p_1 列和后 $n-p_1$ 列在 $\mathbf{R}^{n\times1}$ 中分别生成 p_1 维子空间 U_+ 与 $n-p_1$ 维子空间 $U_{-,0}$, 则 Q 在 U_+ 上正定, 在 $U_{-,0}$ 上半负定.

如果 $p\neq p_1$, 比如 $p>p_1$, 则正定子空间 W_+ 与半负定子空间 $U_{-,0}$ 的维数之和 $p+(n-p_1)>n$. 根据 §2.7 的推论 2.7.1, 这将导致 $W_+\cap U_{-,0}\neq\mathbf{0}$, 正定子空间 W_+ 与半负定子空间 $U_{-,0}$ 将包含有公共的非零向量 $\boldsymbol{X}$. 正定子空间中的 $\boldsymbol{X}\neq\mathbf{0}$ 应满足 $Q(\boldsymbol{X})>0$, 半负定子空间中的 $\boldsymbol{X}$ 又必须满足 $Q(\boldsymbol{X})\leqslant 0$. 同一个 $Q(\boldsymbol{X})$ 不能既 >0 又 $\leqslant 0$, 这个矛盾说明不可能 $p>p_1$. 同理, $p_1>p$ 也将导致正定子空间 U_+ 与 $W_{-,0}$ 的维数之和 $p_1+(n-p)>n$, 这两个子空间同样包含公共的非零向量, 仍矛盾. 可见只能 $p=p_1$. □

$p+(n-p_1)>n$ 时 $W_+\cap U_{-,0}\neq\mathbf{0}$ 也可直接证明如下: 取 W_+ 的基 $\{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_p\}$ 与 $U_{-,0}$ 的基 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_{n-p_1}\}$. 两组基的并集包含的向量个数超过 n, 必然线性相关, 存在不全为 0 的系数使

$$\lambda_1\boldsymbol{w}_1+\cdots+\lambda_p\boldsymbol{w}_p+\mu_1\boldsymbol{u}_1+\cdots+\mu_{n-p_1}\boldsymbol{u}_{n-p_1}=\mathbf{0},$$

移项得

$$\boldsymbol{X}_1=\lambda_1\boldsymbol{w}_1+\cdots+\lambda_p\boldsymbol{w}_p=-\mu_1\boldsymbol{u}_1-\cdots-\mu_{n-p_1}\boldsymbol{u}_{n-p_1},$$

等号两边至少有一边的系数不全为零, 是线性无关向量的非零线性组合. 可见 $\boldsymbol{X}_1 \neq \mathbf{0}$. $\boldsymbol{X}_1$ 既是 W_+ 的基的线性组合也是 $U_{-,0}$ 的基的线性组合, 同时含于 W_+ 与 $U_{-,0}$, 是它们包含的公共非零向量.

2. 利用顺序主子式判定正定性

定理 5.4.2 n 阶实对称方阵 $\boldsymbol{S}$ 正定 $\Leftrightarrow$ $\boldsymbol{S}$ 的所有的顺序主子式 $|\boldsymbol{S}_k| > 0$ $(\forall 1 \leqslant k \leqslant n)$.

证明 只需证明: $|\boldsymbol{S}_k| > 0$ $(1 \leqslant k \leqslant n) \Rightarrow \boldsymbol{S}$ 正定.

对 $\boldsymbol{S}$ 的阶数 n 作数学归纳法. 当 $n = 1$ 时显然. 设命题对 $n-1$ 阶实对称方阵已成立.

将 $\boldsymbol{S}$ 分块为

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{S}_{n-1} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & s_{nn} \end{pmatrix},$$

由于 $|\boldsymbol{S}_{n-1}| > 0$, $\boldsymbol{S}_{n-1}$ 可逆. 因此可以作相合变换

$$\begin{pmatrix} \boldsymbol{I}_{(n-1)} & \mathbf{0} \\ -\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{S}_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{S}_{n-1} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & s_{nn} \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_{(n-1)} & -\boldsymbol{S}_{n-1}^{-1}\boldsymbol{\beta} \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}_{n-1} & \\ & d_n \end{pmatrix},$$

其中 $d_n = s_{nn} - \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{S}_{n-1}\boldsymbol{\beta}$. 在上述矩阵等式两边取行列式得 $|\boldsymbol{S}_n| = |\boldsymbol{S}_{n-1}|d_n$, 因此 $d_n = \dfrac{|\boldsymbol{S}_n|}{|\boldsymbol{S}_{n-1}|} > 0$.

$n-1$ 阶实对称方阵 $\boldsymbol{S}_{n-1}$ 的所有的顺序主子式也就是 $\boldsymbol{S}$ 的前 $n-1$ 个顺序主子式, 都 > 0. 由归纳假设知 $\boldsymbol{S}_{n-1}$ 正定, 可写成 $\boldsymbol{S}_{n-1} = \boldsymbol{P}_1^{\mathrm{T}}\boldsymbol{P}_1$, 其中 $\boldsymbol{P}_1$ 是 $n-1$ 阶可逆方阵. 于是

$$\begin{pmatrix} \boldsymbol{S}_{n-1} & \mathbf{0} \\ \mathbf{0} & d_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_1 & \\ & \sqrt{d_n} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \boldsymbol{P}_1 & \\ & \sqrt{d_n} \end{pmatrix}$$

正定, 与之相合的 $\boldsymbol{S}$ 也正定. □

§5.5 特征向量与相似矩阵

5.5.1 特征向量

例 1 平面直角坐标系中, 方程 $x^2 + 2xy + 5y^2 = 4$ 的图像是否为封闭曲线? 如果是, 试求它围成的图形的面积.

解 将方程左边的二次型 $Q(x,y)=x^2+2xy+5y^2$ 经过可逆线性变换化成规范形: $Q(x,y)=(x+y)^2+(2y)^2=x'^2+y'^2$, 其中

$$\boldsymbol{Y}=\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}x+y\\2y\end{pmatrix}=\boldsymbol{PX},\quad \boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix},\quad \boldsymbol{P}=\begin{pmatrix}1&1\\0&2\end{pmatrix}$$

原方程经过可逆线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{Y}=\boldsymbol{PX}$ 得到的新方程 $x'^2+y'^2=4$ 的图像 C_1 是圆心在原点、半径为 2 的圆, 其面积 $S_1=4\pi$. 圆是封闭曲线, 经过可逆线性变换 $\sigma^{-1}:\boldsymbol{Y}\mapsto\boldsymbol{P}^{-1}\boldsymbol{X}$ 得到的原方程 $x^2+2xy+5y^2=4$ 的图像 C 仍应是封闭曲线. 线性变换 σ^{-1} 将所有图形的面积变到原来面积的 $\det(\boldsymbol{P}^{-1})$ 倍. 因此

$$\boldsymbol{S}=\det(\boldsymbol{P}^{-1})\boldsymbol{S}_1=\begin{vmatrix}1&1\\0&2\end{vmatrix}^{-1}\cdot\boldsymbol{S}_1=\frac{1}{2}\cdot 4\pi=2\pi. \qquad \square$$

在 §4.4 例 2 中, 我们将椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 通过可逆线性变换 $(x,y)\mapsto(x,\dfrac{a}{b}y)$ 在与 y 轴平行的方向上“拉长”成为圆, 求出了椭圆的面积. 本例中通过可逆线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{PX}$ 将方程 $x^2+2xy+5y^2=4$ 的图像 C 变成了圆 C_1, 求出了 C 所围的面积. §4.4 例 2 中的“拉伸”变换 $(x,y)\mapsto(x,\dfrac{a}{b}y)$ 不改变 x 轴方向上的向量, 而将 y 轴方向上的非零向量“拉伸”到原来的 $\dfrac{a}{b}$ 倍. 很自然要问: 本例中的曲线 C 是否也是椭圆, 变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{PX}$ 是否也将某个方向上的非零向量“拉伸”或“压缩”到自己的实数倍?

例 2 平面直角坐标系中的线性变换

$$\sigma:\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x'\\y'\end{pmatrix}=\boldsymbol{PX}=\begin{pmatrix}1&1\\0&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

是否将某些非零向量 $\boldsymbol{X}$ 变到自己的实数倍?

解 设非零向量 $\boldsymbol{X}$ 被 σ 变到自己的实数倍 $\boldsymbol{PX}=\lambda\boldsymbol{X}$. 即

$$\lambda\boldsymbol{X}-\boldsymbol{PX}=(\lambda\boldsymbol{I}-\boldsymbol{P})\boldsymbol{X}=\boldsymbol{0}. \tag{5.5.1}$$

齐次线性方程组 (5.5.1) 有非零解 $\boldsymbol{X}$ $\Leftrightarrow$ 系数行列式 $|\lambda\boldsymbol{I}-\boldsymbol{P}|=0$. 解方程

$$|\lambda\boldsymbol{I}-\boldsymbol{P}|=\begin{vmatrix}\lambda-1&-1\\0&\lambda-2\end{vmatrix}=(\lambda-1)(\lambda-2)=0,$$

得 $\lambda=1$ 或 2. 分别代入 (5.5.1).

当 $\lambda=1$ 时, 方程组 (5.5.1) 即 $\begin{pmatrix}0&-1\\0&-1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$, 求得基础解 $\boldsymbol{X}_1=\begin{pmatrix}1\\0\end{pmatrix}$.

当 $\lambda=2$ 时, 方程组 (5.5.1) 即 $\begin{pmatrix}1&-1\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$, 求得基础解 $\boldsymbol{X}_2=\begin{pmatrix}1\\1\end{pmatrix}$.

图 5–6 左边的图形被 σ 变成右边的图形. 水平方向的向量 ($\boldsymbol{X}_1$ 的实数倍) 都保持不变 (被 σ 变到自己的 1 倍). 与 $\boldsymbol{X}_2=(1,1)^{\mathrm{T}}$ 平行的向量 ($\boldsymbol{X}_2$ 的实数倍) 都被 σ 拉长到自己的 2 倍. 左图中 $x^2+2xy+5y^2=4$ 的图像曲线被 σ 斜着拉伸成为圆.

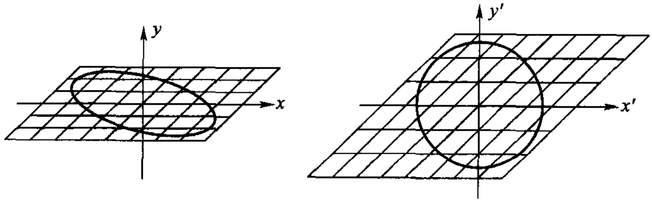

图 5–6

□

定义 5.5.1(特征值, 特征向量)　设 $\boldsymbol{A}\in F^{n\times n}$ 是 n 阶方阵. 如果存在非零向量 $\boldsymbol{X}\in F^{n\times 1}$ 使 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 对某个常数 $\lambda\in F$ 成立, 则称 λ 是 $\boldsymbol{A}$ 的**特征值** (eigenvalue), $\boldsymbol{X}$ 是属于特征值 λ 的**特征向量** (eigenvector).

设 σ 是数域 F 上向量空间 V 上的线性变换, 如果某个非零向量 $\boldsymbol{u}\in V$ 被 σ 映到自己的常数倍 $\sigma(\boldsymbol{u})=\lambda\boldsymbol{u}$, 则称常数 $\lambda\in F$ 是 σ 的**特征值**, 向量 $\boldsymbol{u}$ 是属于特征值 λ 的**特征向量**. □

例 2 中求出的 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 就是方阵 $\boldsymbol{P}$ 的特征向量, 也是线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{PX}$ 的特征向量, 所属的特征值分别是 1,2 .

要同时找到满足同一个等式 $\boldsymbol{AX}=\lambda\boldsymbol{X}$ 的非零向量 $\boldsymbol{X}\neq\boldsymbol{0}$ 及常数 λ 通常是难以做到的. 例 2 中是先由方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 求特征值 λ, 再由 $(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 求特征向量 $\boldsymbol{X}$. 这个算法可以推广到一般情形.

引理 5.5.1　(1) λ_i 是 n 阶方阵 $\boldsymbol{A}$ 的特征值 $\Leftrightarrow$ λ_i 是一元 n 次方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 的根.

(2) $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的属于特征值 λ_i 的特征向量 $\Leftrightarrow$ $\boldsymbol{X}$ 是齐次线性方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ (也就是 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$) 的解空间 V_{λ_i} 中的非零向量.

证明 (1) λ_i 是 $\boldsymbol{A}$ 的特征值 $\Leftrightarrow$ 存在非零向量 $\boldsymbol{X}$ 满足方程 $\boldsymbol{AX}=\lambda_i\boldsymbol{X}$ 即 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ $\Leftrightarrow$ 齐次线性方程组 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 有非零解 $\Leftrightarrow$ 系数行列式 $|\lambda_i\boldsymbol{I}-\boldsymbol{A}|=0$ $\Leftrightarrow$ λ_i 是方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 的根.

设 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 则行列式

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}=\lambda^n+\cdots$$

是 λ 的 n 次多项式 $\varphi_{\boldsymbol{A}}(\lambda)$, 特征值 λ_i 是一元 n 次方程 $\varphi_{\boldsymbol{A}}(\lambda)=0$ 的根.

(2) 设 $\boldsymbol{X}\neq\boldsymbol{0}$. 则: $\boldsymbol{X}$ 是属于 λ_i 的特征向量 $\Leftrightarrow$ $\boldsymbol{AX}=\lambda_i\boldsymbol{X}$

$$\Leftrightarrow(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}\Leftrightarrow\boldsymbol{X}\in V_{\lambda_i}.$$ □

定义 5.5.2(特征多项式, 特征根, 特征子空间) 设 $\boldsymbol{A}\in F^{n\times n}$ 是 n 阶方阵. 则方阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 称为 $\boldsymbol{A}$ 的**特征方阵** (eigenmatrix); 特征方阵的行列式 $\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 称为方阵 $\boldsymbol{A}$ 的**特征多项式** (eigenpolynomial); 特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的根称为 $\boldsymbol{A}$ 的**特征根** (eigenroot). 对 $\boldsymbol{A}$ 的每个特征值 λ_i, 方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的解空间 V_{λ_i} (也就是 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的解空间) 称为 $\boldsymbol{A}$ 的属于特征值 λ_i 的**特征子空间** (eigensubspace). □

由引理 5.5.1 知道: 方阵 $\boldsymbol{A}$ 的全体特征根也就是 $\boldsymbol{A}$ 的全体特征值. 特征子空间 V_{λ_i} 中的全体非零向量, 也就是方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的任何一个基础解系的全体非零线性组合, 就是 $\boldsymbol{A}$ 的属于特征值 λ_i 的全体特征向量.

算法 5.5.1(求方阵的特征向量) 设 $\boldsymbol{A}$ 是 n 阶方阵.

1. 求特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$. 解方程 $\varphi_{\boldsymbol{A}}(\lambda)=0$ 求出所有的根 $\lambda_1,\cdots,\lambda_t$, 即是 $\boldsymbol{A}$ 的全部不同的特征值.

2. 对每个特征值 λ_i, 解齐次线性方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ (也就是 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$) 求出解空间 (特征子空间) V_{λ_i} 的一组基, 这组基的所有的非零线性组合就是 $\boldsymbol{A}$ 的属于 λ_i 的全部特征向量. □

例 3 求如下方阵 $\boldsymbol{A},\boldsymbol{B}$ 的特征值与特征向量:

$$\boldsymbol{A}=\begin{pmatrix}3 & 2 & 1\\ 0 & -1 & -2\\ 0 & 0 & 3\end{pmatrix};\qquad \boldsymbol{B}=\begin{pmatrix}3 & 2 & 1\\ 0 & -1 & -2\\ 0 & 0 & -1\end{pmatrix}.$$

解　$\boldsymbol{A}$ 的特征多项式

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-3 & -2 & -1\\ 0 & \lambda+1 & 2\\ 0 & 0 & \lambda-3\end{vmatrix}=(\lambda-3)^2(\lambda+1).$$

特征值为 $3,-1$.

将特征值 $\lambda=3$ 代入方程组 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得 $(\boldsymbol{A}-3\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 即

$$\begin{pmatrix}0 & 2 & 1\\ 0 & -4 & -2\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}.$$

求得基础解 $\boldsymbol{X}_1=(1,0,0)^{\mathrm{T}},\boldsymbol{X}_2=(0,1,-2)^{\mathrm{T}}$ 组成解空间 V_3 的基. $\boldsymbol{X}_1,\boldsymbol{X}_2$ 的全体非零线性组合 $c_1\boldsymbol{X}_1+c_2\boldsymbol{X}_2\neq\boldsymbol{0}$ $(c_1,c_2\in F)$ 就是属于特征值 3 的全部特征向量.

将特征值 $\lambda=-1$ 代入方程组 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得 $(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 即

$$\begin{pmatrix}4 & 2 & 1\\ 0 & 0 & -2\\ 0 & 0 & 4\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}.$$

求得基础解 $\boldsymbol{X}_3=(1,-2,0)^{\mathrm{T}}$. $\boldsymbol{X}_3$ 的全体非零常数倍 $c\boldsymbol{X}_3\neq\boldsymbol{0}$ 就是属于特征值 -1 的全体特征向量.

$\boldsymbol{B}$ 的特征多项式 $|\lambda\boldsymbol{I}-\boldsymbol{B}|=(\lambda-3)(\lambda+1)^2$, 特征值仍为 $3,-1$.

当 $\lambda=3$, 解方程组 $(\boldsymbol{B}-3\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得基础解 $\boldsymbol{X}_1=(1,0,0)^{\mathrm{T}}$, $c\boldsymbol{X}_1=(c,0,0)^{\mathrm{T}}$ $(0\neq c\in F)$ 就是属于特征值 3 的全体特征向量.

当 $\lambda=-1$, 解方程组 $(\boldsymbol{B}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得基础解 $\boldsymbol{X}_2=(1,-2,0)^{\mathrm{T}}$, $c\boldsymbol{X}_2=(c,-2c,0)^{\mathrm{T}}$ $(0\neq c\in F)$ 就是属于特征值 -1 的全体特征向量. □

例 3 中的两个方阵 $\boldsymbol{A},\boldsymbol{B}$ 都是上三角形矩阵, 对角元就是 $\boldsymbol{A}$ 的全体特征值. 一般地, 我们有:

引理 5.5.2　三角形矩阵 $\boldsymbol{A}$ 的对角元就是 $\boldsymbol{A}$ 的全体特征值. □

对角矩阵既是上三角形矩阵也是下三角形矩阵, 因此对角矩阵 $\boldsymbol{D}$ 的对角元就是 $\boldsymbol{D}$ 的全体特征值.

例 3 中的方阵 $\boldsymbol{A}$ 只有两个不同的特征值 $3,-1$, 但特征值 3 在 $\boldsymbol{A}$ 的对角元 $3,-1,3$ 中重复了两次, 特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=(\lambda-3)^2(\lambda+1)$ 中以 3 为根的一次因式 $\lambda-3$ 也就重复了两次, $(\lambda-3)^2$ 是 $\varphi_{\boldsymbol{A}}(\lambda)$ 的因子. 我们称 3 是多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的 2 重根. 以 $\varphi_{\boldsymbol{A}}(\lambda)$ 的另一根 -1 为根的一次因式 $\lambda+1$ 在多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的分解式中的次数为 1, -1 是 $\varphi_{\boldsymbol{A}}(\lambda)$ 的一重根.

一般地, 如果 c 是多项式 $f(x)$ 的根, 则存在正整数 $m \geqslant 1$ 使 $(x-c)^m$ 整除 $f(x)$ 而 $(x-c)^{m+1}$ 不整除 $f(x)$, 我们称 m 是 $f(x)$ 的根 c 的**重数** (multiplicity), 称 c 是 $f(x)$ 的 m **重根**. 当 $m=1$ 时称 c 是 $f(x)$ 的**单根** (simple root), 当 $m \geqslant 2$ 时称 c 是 $f(x)$ 的**重根** (multiple root). 方阵 $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的根 c 的重数也称为特征值 c 的重数.

每个复系数 n 次多项式 $f(x)$ 在复数范围内都有标准分解式

$$f(x)=a(x-\beta_1)^{n_1}\cdots(x-\beta_t)^{n_t},$$

其中 a 是 n 次项系数, $\beta_1,\cdots,\beta_t$ 是 $f(x)$ 中全部不同的根, n_i 就是根 β_i 的重数. $f(x)$ 的次数 n 等于各因式 $(x-\beta_1)^{n_1},\cdots,(x-\beta_t)^{n_t}$ 的次数之和 $n_1+\cdots+n_t$, 也就是各根 $\beta_1,\cdots,\beta_t$ 的重数之和.

三角形矩阵 (包括对角矩阵) 的全体特征值就是它的全体对角元, 每个特征值 λ_i 的重数就是它在对角元中出现的次数. 例如, 例 3 的方阵 $\boldsymbol{A},\boldsymbol{B}$ 都只有两个不同的特征值 $3,-1$, 但在 $\boldsymbol{A}$ 的 3 个对角元 $3,-1,3$ 中, 3 出现两次, -1 出现一次, 特征值 3 的重数就是 2, 特征值 -1 的重数是 1, 特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=(\lambda-3)^2(\lambda+1)$. 而在 $\boldsymbol{B}$ 的对角元 $3,-1,-1$ 中 3 只出现一次, -1 出现两次, 特征多项式 $\varphi_{\boldsymbol{B}}(\lambda)=(\lambda-3)(\lambda+1)^2$ 与 $\varphi_{\boldsymbol{A}}(\lambda)$ 不同.

将方阵 $\boldsymbol{A}$ 的特征多项式 $|\lambda \boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_0$ 分解为一次因式的乘积

$$|\lambda \boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_0=(\lambda-\lambda_1)\cdots(\lambda-\lambda_n). \tag{5.5.2}$$

在等式 (5.5.2) 中取 $\lambda=0$ 得:

引理 5.5.3 n 阶方阵 $\boldsymbol{A}$ 的 n 个特征值的乘积 $\lambda_1\cdots\lambda_n$ 等于 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}|$. 特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的常数项等于 $(-1)^n|\boldsymbol{A}|$. □

引理 5.5.3 所说的 $\boldsymbol{A}$ 的 n 个特征值 $\lambda_1,\cdots,\lambda_n$, 每个特征值 λ_i 按照它的重数 n_i 重复出现了 n_i 次. 如果 $\beta_1,\cdots,\beta_t$ 是 $\boldsymbol{A}$ 的所有的不同的特征值, 重数分别为 $n_1,\cdots,n_t$, 则 $\beta_1^{n_1}\cdots\beta_t^{n_t}=|\boldsymbol{A}|$.

5.5.2 相似矩阵

例 4 方程 $x^2+2xy+5y^2=4$ 的图像是什么曲线?

分析 这就是例 1 中的方程, 它的图像曲线经过例 2 的线性变换之后被拉伸成圆, 从图 5-6 看出图像曲线是椭圆. 但是, 要确切地证明曲线是椭圆, 需要将原直角坐标系 $O-xy$ 旋转适当角度变成新的直角坐标系 $O-x'y'$, 使图像在新坐标系下的方程是椭圆的标准方程.

将原坐标轴 Ox,Oy 正方向上的单位向量 $\boldsymbol{e}_1,\boldsymbol{e}_2$ 旋转同一角度 α, 得到新坐标轴 Ox',Oy' 正方向上的单位向量 $\boldsymbol{e}_1',\boldsymbol{e}_2'$, 仍组成标准正交基, 以它们为两列

排成的方阵 $\boldsymbol{U}=(\boldsymbol{e}_1',\boldsymbol{e}_2')$ 是正交方阵, 满足条件 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}$ 即 $\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{-1}$, 且 $\boldsymbol{e}_1'$ 到 $\boldsymbol{e}_2'$ 是沿逆时针方向旋转直角, 因此 $\det\boldsymbol{U}=1$. 每一点 $P(x,y)$ 的原坐标 $\boldsymbol{X}=(x,y)^{\mathrm{T}}$ 与新坐标 $\boldsymbol{\xi}=(x',y')^{\mathrm{T}}$ 之间有坐标变换式 $\boldsymbol{X}=x'\boldsymbol{e}_1'+y'\boldsymbol{e}_2'=\boldsymbol{U}\boldsymbol{\xi}$.

原方程左边是 $\boldsymbol{X}$ 的二次型 $Q(\boldsymbol{X})=x^2+2xy+5y^2=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{A}=\begin{pmatrix}1&1\\1&5\end{pmatrix}$$

经过坐标变换 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$ 变成 $\boldsymbol{Y}=(x',y')$ 的二次型:

$$Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=(\boldsymbol{U}\boldsymbol{Y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{U}\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Y}=Q_1(\boldsymbol{\xi}),$$

$Q_1(\boldsymbol{\xi})$ 的方阵 $\boldsymbol{B}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ 与 $\boldsymbol{A}$ 相合. 如果能找到正交方阵 $\boldsymbol{U}$ 将 $\boldsymbol{A}$ 相合到对角矩阵

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}=\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix},$$

则 $Q_1(\boldsymbol{\xi})=\lambda_1x'^2+\lambda_2y'^2$, 曲线的新方程为 $\lambda_1x'^2+\lambda_2y'^2=4$, 容易化成椭圆的标准方程.

按照 §5.3 中讲过的方法, 可以通过二次型的配方或矩阵的初等变换法得到可逆方阵 $\boldsymbol{P}$ 将实对称方阵 $\boldsymbol{A}$ 相合到对角矩阵 $\boldsymbol{D}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$, 但这样求出的 $\boldsymbol{P}$ 很可能不是正交方阵, 它的两列组成的基不是标准正交基, 不能用来建立直角坐标系 $O-x'y'$.

由于正交方阵 $\boldsymbol{U}$ 满足条件 $\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{-1}$, 相合等式 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ 可以写成 $\boldsymbol{D}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$, 从而 $\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}\boldsymbol{D}$. 当 $\boldsymbol{D}=\mathrm{diag}(\lambda_1,\lambda_2)$ 时就有

$$\boldsymbol{A}\boldsymbol{U}=\boldsymbol{A}(\boldsymbol{e}_1',\boldsymbol{e}_2')=\boldsymbol{U}\boldsymbol{D}=(\boldsymbol{e}_1',\boldsymbol{e}_2')\mathrm{diag}(\lambda_1,\lambda_2)=(\lambda_1\boldsymbol{e}_1',\lambda_2\boldsymbol{e}_2'),$$

这就是说 $\boldsymbol{A}\boldsymbol{e}_1'=\lambda_1\boldsymbol{e}_1'$, $\boldsymbol{A}\boldsymbol{e}_2'=\lambda_2\boldsymbol{e}_2'$. 因此 $\boldsymbol{e}_1',\boldsymbol{e}_2'$ 是 $\boldsymbol{A}$ 的特征向量.

由此得到本题的解法: 求 $\boldsymbol{A}$ 的两个特征向量 $\boldsymbol{e}_1',\boldsymbol{e}_2'$ 组成标准正交基, 且 $\boldsymbol{U}=(\boldsymbol{e}_1',\boldsymbol{e}_2')$ 的行列式 $\det\boldsymbol{U}=1$.

解 先求出方程左边的二次型 $x^2+2xy+5y^2=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 的矩阵 $\boldsymbol{A}$ 的特征值与特征向量.

$$\boldsymbol{A}=\begin{pmatrix}1&1\\1&5\end{pmatrix},\quad \varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-1\\-1&\lambda-5\end{vmatrix}=\lambda^2-6\lambda+4.$$

$\varphi_{\boldsymbol{A}}(\lambda)$ 的两根为 $3\pm\sqrt{5}$, 就是 $\boldsymbol{A}$ 的全部特征值.

对特征值 $\lambda_1=3+\sqrt{5}$, 解方程组 $(\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$, 即

$$\begin{pmatrix}-2-\sqrt{5}&1\\1&2-\sqrt{5}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},$$

得基础解 $\boldsymbol{X}_1=(1,2+\sqrt{5})^{\mathrm{T}}$. 类似地, 对特征值 $\lambda_2=3-\sqrt{5}$ 解方程组 $(\boldsymbol{A}-\lambda_2\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得基础解 $\boldsymbol{X}_2=(1,2-\sqrt{5})^{\mathrm{T}}$.

$\boldsymbol{X}_1,\boldsymbol{X}_2$ 的内积 $\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=1+(2+\sqrt{5})(2-\sqrt{5})=0$, 因此 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 垂直. 取单位向量

$$\boldsymbol{P}_1=\frac{1}{|\boldsymbol{X}_1|}\boldsymbol{X}_1=\begin{pmatrix}\dfrac{1}{\sqrt{10+4\sqrt{5}}}\\ \dfrac{2+\sqrt{5}}{\sqrt{10+4\sqrt{5}}}\end{pmatrix},\quad \boldsymbol{P}_2=\frac{1}{|\boldsymbol{X}_2|}\boldsymbol{X}_2=\begin{pmatrix}\dfrac{1}{\sqrt{10-4\sqrt{5}}}\\ \dfrac{2-\sqrt{5}}{\sqrt{10-4\sqrt{5}}}\end{pmatrix}$$

组成标准正交基, 排成的方阵 $\boldsymbol{P}=(\boldsymbol{P}_1,\boldsymbol{P}_2)$ 是正交方阵. 计算得 $\det\boldsymbol{P}=-1$. 用 $-\boldsymbol{P}_2$ 取代 $\boldsymbol{P}_2$ 得到正交方阵 $\boldsymbol{U}=(\boldsymbol{e}_1',\boldsymbol{e}_2')$, 其中 $\boldsymbol{e}_1'=\boldsymbol{P}_1,\boldsymbol{e}_2'=-\boldsymbol{P}_2$, 行列式 $\det\boldsymbol{U}=1$.

$$\boldsymbol{AU}=(\boldsymbol{Ae}_1',\boldsymbol{Ae}_2')=(\lambda_1\boldsymbol{e}_1',\lambda_2\boldsymbol{e}_2')=\boldsymbol{UD},\qquad \boldsymbol{D}=\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix},$$

即 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{AU}=\boldsymbol{U}^{-1}\boldsymbol{AU}=\boldsymbol{D}$. 分别以 $\boldsymbol{e}_1',\boldsymbol{e}_2'$ 为 Ox',Oy' 轴正方向上的单位向量建立直角坐标系 $O-x'y'$, 则原坐标系中坐标为 $\boldsymbol{X}=(x,y)^{\mathrm{T}}$ 的点在坐标系 $O-x'y'$ 中的新坐标 $\boldsymbol{\xi}=(x',y')^{\mathrm{T}}$ 满足条件 $\boldsymbol{X}=\boldsymbol{U\xi}$. 于是

$$\begin{aligned}x^2+2xy+5y^2&=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}=(\boldsymbol{U\xi})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{U\xi})=\boldsymbol{\xi}^{\mathrm{T}}(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{AU})\boldsymbol{\xi}\\&=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{D\xi}=(3+\sqrt{5})x'^2+(3-\sqrt{5})y'^2,\end{aligned}$$

原方程的图像在新坐标系下的方程为

$$(3+\sqrt{5})x'^2+(3-\sqrt{5})y'^2=4,\quad 即\ \frac{x'^2}{b^2}+\frac{y'^2}{a^2}=1,$$

其中

$$a=\sqrt{\frac{4}{3-\sqrt{5}}}=\sqrt{3+\sqrt{5}},\quad b=\sqrt{\frac{4}{3+\sqrt{5}}}=\sqrt{3-\sqrt{5}}.$$

这是椭圆的标准方程.

可见原方程的图像是椭圆, 半长轴和半短轴分别为 $a=\sqrt{3+\sqrt{5}},b=\sqrt{3-\sqrt{5}}$. □

例 3 中将相合关系 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{AU}$ 写成了 $\boldsymbol{D}=\boldsymbol{U}^{-1}\boldsymbol{AU}$, 由此得到 $\boldsymbol{AU}=\boldsymbol{UD}$, 从而 $\boldsymbol{U}$ 的各列是 $\boldsymbol{A}$ 的特征向量.

一般地, 只要 $\boldsymbol{P}$ 是可逆方阵 (不一定是正交方阵), 都可由 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$ 得到 $\boldsymbol{AP}=\boldsymbol{PB}$, 当 $\boldsymbol{B}$ 是对角矩阵时都能得到 $\boldsymbol{P}$ 的各列是 $\boldsymbol{A}$ 的特征向量.

定义 5.5.3(相似矩阵) 设 $\boldsymbol{A},\boldsymbol{B}$ 是 n 阶方阵. 如果存在 n 阶可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似** (similar). □

相似方阵具有如下一些简单而有用的性质.

引理 5.5.4 (1) $\boldsymbol{A}$ 与 $\boldsymbol{A}$ 本身相似.

(2) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似.

(3) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 且 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 相似, 则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 相似.

证明 (1) $\boldsymbol{A}=\boldsymbol{I}^{-1}\boldsymbol{A}\boldsymbol{I}$.

(2) $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} \Rightarrow \boldsymbol{A}=\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}=(\boldsymbol{P}^{-1})^{-1}\boldsymbol{B}(\boldsymbol{P}^{-1})$.

(3) $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}, \boldsymbol{C}=\boldsymbol{K}^{-1}\boldsymbol{B}\boldsymbol{K} \Rightarrow \boldsymbol{C}=(\boldsymbol{P}\boldsymbol{K})^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{K})$. □

引理 5.5.5 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似. 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 特征值以及每个特征值的重数相同, 行列式 $|\boldsymbol{A}|=|\boldsymbol{B}|$.

证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似 $\Leftrightarrow$ 存在可逆方阵 $\boldsymbol{P}$ 使 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 从而

$$\begin{aligned}|\lambda\boldsymbol{I}-\boldsymbol{B}| &= |\lambda\boldsymbol{I}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}|=|\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}|\\ &=|\boldsymbol{P}|^{-1}|\lambda\boldsymbol{I}-\boldsymbol{A}||\boldsymbol{P}|=|\lambda\boldsymbol{I}-\boldsymbol{A}|.\end{aligned}$$

这证明了 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 与 $\varphi_{\boldsymbol{B}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{B}|$ 相同.

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值是同一个特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=\varphi_{\boldsymbol{B}}(\lambda)$ 的根, 这些根以及重数当然相同.

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相等, 它们的常数项 $a_n=(-1)^n|\boldsymbol{A}|$ 与 $b_n=(-1)^n|\boldsymbol{B}|$ 相等, 因此行列式 $|\boldsymbol{A}|$ 与 $|\boldsymbol{B}|$ 相等. □

相似方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_0$ 与 $|\lambda\boldsymbol{I}-\boldsymbol{B}|=\lambda^n+b_{n-1}\lambda^{n-1}+\cdots+b_0$ 相等, 对应项系数 a_i 与 b_i 也相等. 已经知道 $|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的常数项 a_0 是 $\boldsymbol{A}$ 的行列式的 $(-1)^n$ 倍, 我们来看 $n-1$ 次项系数 a_{n-1} 与 $\boldsymbol{A}$ 的元素之间的关系.

行列式

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_0$$

是 $n!$ 个乘积的代数和, 每个乘积由不同行不同列的 n 个元素相乘得到, 其中含有 $n-1$ 次项的乘积的 n 个因子中必须有 $n-1$ 个含 λ, 也就是 $n-1$ 个对角元, 剩下一个因子与这 $n-1$ 个对角元都不同行且不同列, 只能是剩下的一个对角元. 可见, $|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的 $n-1$ 次项只能含于 n 个对角元的乘积

$$(\lambda-a_{11})\cdots(\lambda-a_{nn})=\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}+\cdots$$

中, 等于这个乘积展开式的 $n-1$ 次项 $-(a_{11}+\cdots+a_{nn})\lambda^{n-1}$, 其系数 $a_{n-1}=-(a_{11}+\cdots+a_{nn})$, 即等于 $\boldsymbol{A}$ 的各对角元之和的相反数.

定义 5.5.4 方阵 $\boldsymbol{A}$ 的全体对角元之和称为 $\boldsymbol{A}$ 的**迹** (trace), 记为 tr $\boldsymbol{A}$. □

推论 5.5.1 相似方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的迹相等: $\mathrm{tr}\boldsymbol{A}=\mathrm{tr}\boldsymbol{B}$. □

引理 5.5.6 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{P}$ 是同阶方阵且 $\boldsymbol{P}$ 可逆. 则

(1) $(\boldsymbol{PAP}^{-1})(\boldsymbol{PBP}^{-1})=\boldsymbol{P}(\boldsymbol{AB})\boldsymbol{P}^{-1}$.

(2) $(\boldsymbol{PAP}^{-1})^n=\boldsymbol{PA}^n\boldsymbol{P}^{-1}$.

(3) 对任意多项式 $f(x)$, 有 $f(\boldsymbol{PAP}^{-1})=\boldsymbol{P}f(\boldsymbol{A})\boldsymbol{P}^{-1}$.

证明 (1) 显然.

(2) 对 k 作数学归纳法. $k=1$ 时显然. 归纳假设 $\boldsymbol{B}^{k-1}=\boldsymbol{P}^{-1}\boldsymbol{A}^{k-1}\boldsymbol{P}$, 则 $\boldsymbol{B}^k=\boldsymbol{B}^{k-1}\boldsymbol{B}=(\boldsymbol{P}^{-1}\boldsymbol{A}^{k-1}\boldsymbol{P})(\boldsymbol{P}^{-1}\boldsymbol{AP})=\boldsymbol{P}^{-1}(\boldsymbol{A}^{k-1}\boldsymbol{A})\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P}$.

(3) 设 $f(x)=a_0+a_1x+\cdots+a_mx^m$. 则

$$\begin{aligned}f(\boldsymbol{PAP}^{-1})&=a_0\boldsymbol{I}+a_1(\boldsymbol{PAP}^{-1})+\cdots+a_m(\boldsymbol{PAP}^{-1})^m\\&=\boldsymbol{P}(a_0\boldsymbol{I}+a_1\boldsymbol{A}+\cdots+a_m\boldsymbol{A}^m)\boldsymbol{P}^{-1}=\boldsymbol{P}f(\boldsymbol{A})\boldsymbol{P}^{-1}.\end{aligned}$$ □

例 5 对任意正整数 n, 求例 3 中的方阵

$$\boldsymbol{A}=\begin{pmatrix}3&2&1\\0&-1&-2\\0&0&3\end{pmatrix}$$

的 n 次幂 $\boldsymbol{A}^n$.

解 例 3 中对 $\boldsymbol{A}$ 求出了属于特征值 3 的特征向量 $\boldsymbol{X}_1=(1,0,0)^{\mathrm{T}},\boldsymbol{X}_2=(0,1,-2)^{\mathrm{T}}$, 它们组成特征子空间 V_3 的一组基, 属于特征值 -1 的特征向量 $\boldsymbol{X}_3=(1,-2,0)^{\mathrm{T}}$ 组成特征子空间 V_{-1} 的一组基. 以 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 为各列排成的方阵

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\begin{pmatrix}1&0&1\\0&1&-2\\0&-2&0\end{pmatrix}$$

可逆. 由 $\boldsymbol{AX}_1=3\boldsymbol{X}_1,\boldsymbol{AX}_2=3\boldsymbol{X}_2,\boldsymbol{AX}_3=-\boldsymbol{X}_3$ 得

$$\begin{aligned}\boldsymbol{AP}&=(\boldsymbol{AX}_1,\boldsymbol{AX}_2,\boldsymbol{AX}_3)=(3\boldsymbol{X}_1,3\boldsymbol{X}_2,-\boldsymbol{X}_3)\\&=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)\begin{pmatrix}3&0&0\\0&3&0\\0&0&-1\end{pmatrix}=\boldsymbol{PD},\ \text{其中 }\boldsymbol{D}=\begin{pmatrix}3&0&0\\0&3&0\\0&0&-1\end{pmatrix}.\end{aligned}$$

因而

$$\begin{aligned}
\boldsymbol{A} &= \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}, \\
\boldsymbol{A}^n &= \boldsymbol{P}\boldsymbol{D}^n\boldsymbol{P}^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix} \begin{pmatrix} 3^n & 0 & 0 \\ 0 & 3^n & 0 \\ 0 & 0 & (-1)^n \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{4} \end{pmatrix} \\
&= \begin{pmatrix} 3^n & \frac{1}{2}\times 3^n - \frac{1}{2}\times(-1)^n & \frac{1}{4}\times 3^n - \frac{1}{4}\times(-1)^n \\ 0 & (-1)^n & -\frac{1}{2}\times 3^n + \frac{1}{2}\times(-1)^n \\ 0 & 0 & 3^n \end{pmatrix}.
\end{aligned}$$

□

例 6(斐波那契数列) 已知数列 $\{F_n\}$ 的前两项 $F_1 = F_2 = 1$, 且从第 3 项起每项 $F_n = F_{n-1} + F_{n-2}$ $(\forall\ n \geqslant 3)$. 求这个数列的通项公式.

解 将这个数列的每相邻两项组成一个 2 维数组, 这些数组组成一个序列:

$$(F_1, F_2), (F_2, F_3), \cdots, (F_{n-2}, F_{n-1}), (F_{n-1}, F_n), \cdots, \tag{5.5.3}$$

序列中相邻两项之间的关系可以用矩阵乘法写为

$$(F_{n-1}, F_n) = (F_{n-2}, F_{n-1}) \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{5.5.4}$$

序列 (5.5.3) 每一项右乘矩阵 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ 得到下一项. 第 1 项 (F_1, F_2) 连续右乘 $n-2$ 个 $\boldsymbol{A}$ 得到第 $n-1$ 项 (F_{n-1}, F_n):

$$(F_{n-1}, F_n) = (F_1, F_2)\boldsymbol{A}^{n-2} = (1, 1) \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{n-2} \tag{5.5.5}$$

为了计算 $\boldsymbol{A}^{n-2}$, 求可逆方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于对角矩阵.

$\boldsymbol{A}$ 的特征多项式

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{vmatrix} = \lambda^2 - \lambda - 1$$

的两根为 $\lambda_1 = \dfrac{1-\sqrt{5}}{2}$, $\lambda_2 = \dfrac{1+\sqrt{5}}{2}$. 将 $\lambda = \lambda_1, \lambda_2$ 分别代入方程组 $(\lambda \boldsymbol{I} -$

$\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$, 各求出一个基础解

$$\boldsymbol{X}_1=\begin{pmatrix}1\\\lambda_1\end{pmatrix},\qquad \boldsymbol{X}_2=\begin{pmatrix}1\\\lambda_2\end{pmatrix}$$

满足条件 $\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1$, $\boldsymbol{A}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_2$. 以 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 为两列组成可逆方阵 $\boldsymbol{P}$. 则 $\boldsymbol{AP}=\boldsymbol{PD}$ 对 $\boldsymbol{D}=\mathrm{diag}(\lambda_1,\lambda_2)$ 成立, 从而 $\boldsymbol{A}=\boldsymbol{PDP}^{-1}$,

$$\begin{aligned}\boldsymbol{A}^{n-2}&=\boldsymbol{PD}^{n-2}\boldsymbol{P}^{-1}=\begin{pmatrix}1&1\\\lambda_1&\lambda_2\end{pmatrix}\begin{pmatrix}\lambda_1^{n-2}&0\\0&\lambda_2^{n-2}\end{pmatrix}\cdot\frac{1}{\lambda_2-\lambda_1}\begin{pmatrix}\lambda_2&-1\\-\lambda_1&1\end{pmatrix}\\&=\frac{1}{\sqrt5}\begin{pmatrix}*&\lambda_2^{n-2}-\lambda_1^{n-2}\\ *&\lambda_2^{n-1}-\lambda_1^{n-1}\end{pmatrix},\end{aligned}$$

代入 (5.5.5) 得

$$\begin{aligned}(F_{n-1},F_n)&=(1,1)\boldsymbol{A}^{n-2}=\frac{1}{\sqrt5}(*,\lambda_2^{n-2}-\lambda_1^{n-2}+\lambda_2^{n-1}-\lambda_1^{n-1}),\\F_n&=\frac{1}{\sqrt5}(\lambda_2^{n-2}(1+\lambda_2)-\lambda_1^{n-2}(1+\lambda_1))\\&=\frac{1}{\sqrt5}(\lambda_2^{n-2}\lambda_2^2-\lambda_1^{n-2}\lambda_1^2)\\&=\frac{1}{\sqrt5}\left(\left(\frac{1+\sqrt5}{2}\right)^n-\left(\frac{1-\sqrt5}{2}\right)^n\right).\end{aligned}$$

□

5.5.3　相似于对角矩阵的条件

例 7　对任意正整数 n, 求例 3 中的方阵

$$\boldsymbol{B}=\begin{pmatrix}3&2&1\\0&-1&-2\\0&0&-1\end{pmatrix}$$

的 n 次幂 $\boldsymbol{B}^n$.

分析　仍希望仿照例 5 和例 6 的方法, 求有可逆方阵 $\boldsymbol{P}$ 将 $\boldsymbol{B}$ 相似到对角矩阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{BP}$, 再由 $\boldsymbol{B}^n=(\boldsymbol{PDP}^{-1})^n=\boldsymbol{PD}^n\boldsymbol{P}^{-1}$ 求出 $\boldsymbol{B}^n$.

如果存在上述可逆方阵 $\boldsymbol{P}$ 和对角矩阵 $\boldsymbol{D}$, 则由引理 5.5.5 知 $\boldsymbol{D}$ 的特征值及其重数与 $\boldsymbol{B}$ 相同, 为 3(1 重),−1(2 重), 因而 $\boldsymbol{D}=\mathrm{diag}(3,-1,-1)$. 由

$$\boldsymbol{BP}=(\boldsymbol{BX}_1,\boldsymbol{BX}_2,\boldsymbol{BX}_3)=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)\,\mathrm{diag}(3,-1,-1)=(3\boldsymbol{X}_1,-\boldsymbol{X}_2,-\boldsymbol{X}_3)$$

知 $\boldsymbol{BX}_1=3\boldsymbol{X}_1,\boldsymbol{BX}_2=-\boldsymbol{X}_2,\boldsymbol{BX}_3=-\boldsymbol{X}_3$, 即 $\boldsymbol{X}_1$ 是属于 $\boldsymbol{B}$ 的特征值 3 的特征向量, $\boldsymbol{X}_2,\boldsymbol{X}_3$ 是属于特征值 −1 的线性无关特征向量.

解 例 3 中求出了方程组 $(\boldsymbol{B}-3\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的基础解 $\boldsymbol{X}_1=(1,0,0)^{\mathrm{T}}$, 方程组 $(\boldsymbol{B}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的基础解 $\boldsymbol{X}_2=(1,-2,0)^{\mathrm{T}}$, $\boldsymbol{X}_1,\boldsymbol{X}_2$ 分别单独组成特征子空间 V_3 与 V_{-1} 的基. 属于 B 的特征值 -1 的特征子空间 V_{-1} 只有 1 维, 不可能包含两个线性无关的特征向量 $\boldsymbol{X}_2,\boldsymbol{X}_3$. 可见 $\boldsymbol{B}$ 不能相似于对角矩阵.

方程组 $(\boldsymbol{B}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的解空间 V_{-1} 维数为 1 是因为它的系数矩阵 $\boldsymbol{B}+\boldsymbol{I}$ 的秩为 2. 由

$$(\boldsymbol{B}+\boldsymbol{I})^2=\begin{pmatrix}4&2&1\\0&0&-2\\0&0&0\end{pmatrix}^2=\begin{pmatrix}16&8&0\\0&0&0\\0&0&0\end{pmatrix}$$

的秩为 1 知方程组 $(\boldsymbol{B}+\boldsymbol{I})^2\boldsymbol{X}=\boldsymbol{0}$ 的解空间的维数为 $3-\operatorname{rank}(\boldsymbol{B}+\boldsymbol{I})^2=2$, 因此存在与 $\boldsymbol{X}_2$ 线性无关的解 $\boldsymbol{X}_3$. 易见 $\boldsymbol{X}_3=(0,0,1)^{\mathrm{T}}$ 是 $(\boldsymbol{B}+\boldsymbol{I})^2\boldsymbol{X}=\boldsymbol{0}$ 的解, 且 $(\boldsymbol{B}+\boldsymbol{I})\boldsymbol{X}_3=(1,-2,0)^{\mathrm{T}}=\boldsymbol{X}_2$, 因而 $\boldsymbol{B}\boldsymbol{X}_3=\boldsymbol{X}_2-\boldsymbol{X}_3$. 以 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 为 3 列排成的方阵

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)=\begin{pmatrix}1&1&0\\0&-2&0\\0&0&1\end{pmatrix}$$

可逆,

$$\begin{aligned}\boldsymbol{BP}&=(\boldsymbol{BX}_1,\boldsymbol{BX}_2,\boldsymbol{BX}_3)=(3\boldsymbol{X}_1,-\boldsymbol{X}_2,\boldsymbol{X}_2-\boldsymbol{X}_3)\\&=(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3)\begin{pmatrix}3&0&0\\0&-1&1\\0&0&-1\end{pmatrix}=\boldsymbol{PJ},\quad \text{其中}\quad \boldsymbol{J}=\begin{pmatrix}3&0&0\\0&-1&1\\0&0&-1\end{pmatrix},\end{aligned}$$

于是得 $\boldsymbol{B}=\boldsymbol{PJP}^{-1}$. 只要计算出 $\boldsymbol{J}^n$, 就得到 $\boldsymbol{B}^n=\boldsymbol{PJ}^n\boldsymbol{P}^{-1}$.

将 $\boldsymbol{J}$ 分解为 $\boldsymbol{J}=\boldsymbol{D}+\boldsymbol{N}$, 其中

$$\boldsymbol{D}=\begin{pmatrix}3&0&0\\0&-1&0\\0&0&-1\end{pmatrix},\quad \boldsymbol{N}=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix},$$

满足条件 $\boldsymbol{DN}=\boldsymbol{ND}$ 及 $\boldsymbol{N}^2=\boldsymbol{O}$. 可用牛顿二项式定理算出

$$\boldsymbol{J}^n=(\boldsymbol{D}+\boldsymbol{N})^n=\boldsymbol{D}^n+n\boldsymbol{D}^{n-1}\boldsymbol{N}=\begin{pmatrix}3^n&0&0\\0&(-1)^n&n(-1)^{n-1}\\0&0&(-1)^n\end{pmatrix},$$

进而得到

$$B^n = PJ^nP^{-1} = \begin{pmatrix} 3^n & \frac{1}{2}\times 3^n - \frac{1}{2}(-1)^n & n(-1)^{n-1} \\ 0 & (-1)^n & -2n(-1)^{n-1} \\ 0 & 0 & (-1)^n \end{pmatrix}. \qquad \square$$

例 3 的方阵 $\boldsymbol{A}$ 相似于对角矩阵以及 $\boldsymbol{B}$ 不相似于对角矩阵的原因可以推广到一般的方阵:

定理 5.5.1(相似于对角矩阵的条件) n 阶复方阵 $\boldsymbol{A}$ 相似于对角矩阵 $\Leftrightarrow$ $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$.

证明 设 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D} = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, 其中 $\boldsymbol{P}$ 为可逆方阵. 可逆方阵 $\boldsymbol{P} = (\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$ 的各列 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$ 线性无关, 且由 $\boldsymbol{AP} = \boldsymbol{PD}$ 即

$$\boldsymbol{A}(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n) = (\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = (\lambda_1\boldsymbol{X}_1,\cdots,\lambda_n\boldsymbol{X}_n)$$

知 $\boldsymbol{A}\boldsymbol{X}_j = \lambda_j\boldsymbol{X}_j$ 对 $1 \leqslant j \leqslant n$ 成立, $\boldsymbol{P}$ 的 n 列 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$ 是 n 个线性无关的特征向量, 分别属于特征值 $\lambda_1,\cdots,\lambda_n$.

反过来, 设 $\boldsymbol{A}$ 有 n 个线性无关特征向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$, 分别属于特征值 $\lambda_1,\cdots,\lambda_n$, 则 $\boldsymbol{A}\boldsymbol{X}_j = \lambda_j\boldsymbol{X}_j$ 对 $1 \leqslant j \leqslant n$ 成立. 以这些特征向量为各列排成的矩阵 $\boldsymbol{P} = (\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$ 是可逆方阵. 且

$$\boldsymbol{AP} = (\boldsymbol{A}\boldsymbol{X}_1,\cdots,\boldsymbol{A}\boldsymbol{X}_n) = (\lambda_1\boldsymbol{X}_1,\cdots,\lambda_n\boldsymbol{X}_n) = \boldsymbol{PD}$$

对 $\boldsymbol{D} = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 成立. $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{D}$ 是对角矩阵. $\square$

例 3 中的方阵 $\boldsymbol{B}$ 的两个特征子空间 V_3, V_{-1} 都只有一维, 各求出一个基向量 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 之后, 所有的特征向量都是 $\boldsymbol{X}_1$ 或 $\boldsymbol{X}_2$ 的常数倍, 不存在 3 个线性无关特征向量, 因此 $\boldsymbol{B}$ 不相似于对角矩阵. 导致这个结果的原因是: 特征值 -1 的重数 2 大于特征子空间 V_{-1} 的维数 1. 如果 $\boldsymbol{B}$ 相似于对角矩阵 $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{BP}$, 则 -1 在 $\boldsymbol{D} = \mathrm{diag}(3,-1,-1)$ 的对角元中出现两次, 由 $\boldsymbol{BP} = \boldsymbol{B}(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3) = \boldsymbol{PD} = (3\boldsymbol{X}_1,-\boldsymbol{X}_2,-\boldsymbol{X}_3)$ 知 $\boldsymbol{P}$ 有两列 $\boldsymbol{X}_2,\boldsymbol{X}_3$ 是属于特征值 -1 的线性无关特征向量. 而特征子空间 V_{-1} 只有 1 维, 不可能有两个线性无关的特征向量.

这个推理也适用于任意方阵 $\boldsymbol{A}$, 得到:

定理 5.5.2 复方阵 $\boldsymbol{A}$ 相似于对角矩阵 $\Leftrightarrow$ $\boldsymbol{A}$ 的每个特征值 λ_i 的重数等于特征子空间 V_{λ_i} 的维数.

定理 5.5.2 的详细证明参见附录 7.

例 8 实方阵

$$\varphi_{\boldsymbol{A}}(\lambda)=\boldsymbol{A}=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix},$$

其中 α 不是 π 的整数倍. 是否存在实方阵或复方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$?

解 如果 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{D}$ 是对角矩阵, 则 $\boldsymbol{D}$ 的对角元都是 $\boldsymbol{A}$ 的特征值.

$\boldsymbol{A}$ 的特征多项式

$$\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-\cos\alpha & \sin\alpha\\ -\sin\alpha & \lambda-\cos\alpha\end{vmatrix}=\lambda^2-2\lambda\cos\alpha+1.$$

解方程 $\varphi_{\boldsymbol{A}}(\lambda)=0$ 得特征值 $\lambda=\cos\alpha\pm\mathrm{i}\sin\alpha$. 当 α 不是 π 的整数倍时, $\sin\alpha\neq0$, $\boldsymbol{A}$ 的特征值都是虚数, 如果 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角矩阵, 对角元只能是虚数. 而当 $\boldsymbol{P}$ 是实方阵时 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 只能是实方阵, 不可能是对角矩阵.

当 $\lambda=\cos\alpha+\mathrm{i}\sin\alpha$ 时, 解方程组 $(\boldsymbol{A}-(\cos\alpha+\mathrm{i}\sin\alpha)\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得基础解 $X_1=\begin{pmatrix}\mathrm{i}\\1\end{pmatrix}$.

当 $\lambda=\cos\alpha-\mathrm{i}\sin\alpha$ 时, 解方程组 $(\boldsymbol{A}-(\cos\alpha-\mathrm{i}\sin\alpha)\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的基础解 $X_2=\begin{pmatrix}-\mathrm{i}\\1\end{pmatrix}$.

取

$$\boldsymbol{P}=(\boldsymbol{X}_1,\boldsymbol{X}_2)=\begin{pmatrix}\mathrm{i} & -\mathrm{i}\\ 1 & 1\end{pmatrix},$$

则 $\boldsymbol{A}\boldsymbol{P}=(\boldsymbol{A}\boldsymbol{X}_1,\boldsymbol{A}\boldsymbol{X}_2)=(\lambda_1\boldsymbol{X}_1,\lambda_2\boldsymbol{X}_2)=\boldsymbol{P}\boldsymbol{D}$ 对

$$\boldsymbol{D}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}=\begin{pmatrix}\cos\alpha+\mathrm{i}\sin\alpha & 0\\ 0 & \cos\alpha-\mathrm{i}\sin\alpha\end{pmatrix}$$

成立. 可逆复方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似到对角阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. □

例 8 中的 $\boldsymbol{A}$ 引起的线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 将每个非零向量 $\boldsymbol{X}=\overrightarrow{OP}$ 绕 O 旋转角 α 到 $\boldsymbol{A}\boldsymbol{X}=\overrightarrow{OP'}$, 当 α 不是 π 的整数倍时, OP' 与 OP 不共线. $\boldsymbol{A}\boldsymbol{X}$ 不是 $\boldsymbol{X}$ 的实数倍, 可见 $\boldsymbol{A}$ 没有实特征向量.

习题 5.5

1. 已知矩阵 $\boldsymbol{A}=\begin{pmatrix}\frac{3}{5}&\frac{4}{5}\\ \frac{4}{5}&-\frac{3}{5}\end{pmatrix}$. 平面直角坐标系中的变换 $\sigma:\begin{pmatrix}x\\y\end{pmatrix}\mapsto \boldsymbol{A}\begin{pmatrix}x\\y\end{pmatrix}$ 将哪些向量变到自己的实数倍? σ 是什么变换?

2. 求下列矩阵 $\boldsymbol{A}$ 的全部特征值和特征向量:

(1) $\boldsymbol{A}=\begin{pmatrix}0&0&1\\0&2&0\\4&0&0\end{pmatrix}$; (2) $\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&2&1\\0&0&3\end{pmatrix}$;

(3) $\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$; (4) $\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&2&1\\0&0&0&2\end{pmatrix}$.

3. (1) 构造一个二阶整系数方阵 $\boldsymbol{A}$, 使它的特征根是 $2+\sqrt{3}$ 与 $2-\sqrt{3}$;

(2) 构造一个三阶对角矩阵 $\boldsymbol{A}$, 使它的特征多项式是 $\lambda^3+2\lambda^2-3\lambda$.

4. (1) 求方阵

$$\boldsymbol{A}=\begin{pmatrix}0&0&-a_3\\1&0&-a_2\\0&1&-a_1\end{pmatrix}$$

的特征多项式;

(2) 构造一个三阶整系数方阵 $\boldsymbol{A}$, 使它的特征多项式为 $\lambda^3-3\lambda^2+2$.

5. 是否存在可逆方阵 $\boldsymbol{P}$ 将下列矩阵 $\boldsymbol{A}$ 相似于对角阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$? 如果存在, 求出 $\boldsymbol{P},\boldsymbol{D}$.

(1) $\boldsymbol{A}=\begin{pmatrix}0&0&1\\0&2&0\\4&0&0\end{pmatrix}$; (2) $\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&2&1\\0&0&3\end{pmatrix}$;

(3) $\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&1&-1&1\end{pmatrix}$; (4) $\boldsymbol{A}=\begin{pmatrix}1&1&1&1\\0&1&1&1\\0&0&2&1\\0&0&0&2\end{pmatrix}$.

6. 已知 $\boldsymbol{A}=\begin{pmatrix}0&2\\1&-1\end{pmatrix}$, 求 $\boldsymbol{A}^n$.

7. 已知方阵 $\boldsymbol{A}$ 满足条件 $\boldsymbol{A}^2=\boldsymbol{I}$. 求 $\boldsymbol{A}$ 的特征值.

8. 求证: 方阵 $\boldsymbol{A}$ 可逆 $\Leftrightarrow$ 0 不是 $\boldsymbol{A}$ 的特征值.

9. 求证: 如果方阵 $\boldsymbol{A}$ 的某个正整数次幂 $\boldsymbol{A}^m=\boldsymbol{O}$, 则 $\boldsymbol{A}$ 只有一个特征值 0. 反过来, 如果 $\boldsymbol{A}$ 只有唯一的特征值 0, 一定存在正整数 m 使 $\boldsymbol{A}^m=\boldsymbol{O}$.

10. 设 $n \geqslant 2$, $V = F^{n\times n}$ 是数域 F 上全体 n 阶方阵组成的向量空间, V 的线性变换 $\tau : \boldsymbol{X} \mapsto \boldsymbol{X}^{\mathrm{T}}$ 将 V 中每个方阵 $\boldsymbol{X}$ 映到它的转置 $\boldsymbol{X}^{\mathrm{T}}$. 求 τ 的特征值和特征向量.

附录 7 复方阵的对角化与三角化

将复方阵 $\boldsymbol{A}$ 相似到最简单的方阵——对角矩阵, 称为复方阵的对角化. 并非所有的复方阵都能够对角化, 如例 3 中的复方阵 $\boldsymbol{B}$ 就不能对角化. 但所有的复方阵都能相似到三角形矩阵, 称为方阵的三角化.

1. 代数重数与几何重数

定理 5.5.2 给出了复方阵 $\boldsymbol{A}$ 可对角化的充分必要条件: 每个特征值 λ_i 的重数等于特征子空间 V_{λ_i} 的维数.

定义 5.5.5 $\boldsymbol{A}$ 的特征值 λ_i 在特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 中的重数称为特征值 λ_i 的**代数重数** (algebraic multiplicity), 特征子空间 V_{λ_i} (即方程组的解空间 $(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$) 的维数称为 λ_i 的**几何重数** (geometric multiplicity).

按照这个定义, 定理 5.5.2 就叙述为:

复方阵 $\boldsymbol{A}$ 相似于对角阵 $\Leftrightarrow$ $\boldsymbol{A}$ 的每个特征值 λ_i 的代数重数等于几何重数.

定理 5.5.2 的证明

先设存在可逆复方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于对角阵 $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. 设特征值 λ_i 的代数重数为 n_i, 则 $\boldsymbol{D}$ 恰有 n_i 个对角元等于 λ_i, $\boldsymbol{D} - \lambda_i\boldsymbol{I}$ 恰有 n_i 个对角元为 0, $\mathrm{rank}\,(\boldsymbol{D} - \lambda_i\boldsymbol{I}) = n - n_i$. 从而 $\mathrm{rank}\,(\boldsymbol{A} - \lambda_i\boldsymbol{I}) = \mathrm{rank}\,(\boldsymbol{D} - \lambda_i\boldsymbol{I}) = n - n_i$, 方程组 $(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 的解空间 V_{λ_i} 的维数 (即 λ_i 的几何重数) $\dim V_{\lambda_i} = n - \mathrm{rank}\,(\boldsymbol{A} - \lambda_i\boldsymbol{I}) = n - (n - n_i) = n_i$. 几何重数 $\dim V_{\lambda_i}$ 等于代数重数 n_i.

再设 $\boldsymbol{A}$ 的各个不同特征值 $\lambda_1, \cdots, \lambda_t$ 的代数重数 $n_1, \cdots, n_t$ 分别等于它们的几何重数 $\dim V_{\lambda_1}, \cdots, \dim V_{\lambda_t}$. 对每个特征值 λ_i, 取特征子空间 V_{λ_i} 的一组基 $S_i = \{\boldsymbol{X}_{i1}, \cdots, \boldsymbol{X}_{in_i}\}$, 由 n_i 个特征向量组成. 各个 S_i 的并集 $S = S_1 \cup \cdots \cup S_t$ 由 $n_1 + \cdots + n_t = n$ 个特征向量组成. 只要证明 S 线性无关, 以 S 中 n 个特征向量 $\boldsymbol{X}_{ij}$ 为各列排成的方阵 $\boldsymbol{P}$ 可逆, 则由定理 5.5.1 知 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于对角阵.

设 n 个复数 μ_{ij} $(1 \leqslant i \leqslant t, 1 \leqslant j \leqslant n_i)$ 使

$$\sum_{j=1}^{n_1} \mu_{1j}\boldsymbol{X}_{1j} + \cdots + \sum_{j=1}^{n_t} \mu_{tj}\boldsymbol{X}_{tj} = \boldsymbol{0}, \tag{5.5.6}$$

我们证明所有的 μ_{ij} 全部等于 0, 则 S 线性无关. 对每个 $1 \leqslant i \leqslant t$, 记

$$\boldsymbol{a}_i = \sum_{j=1}^{n_i} \mu_{ij} \boldsymbol{X}_{ij},$$

代入 (5.5.6) 得

$$\boldsymbol{a}_1 + \cdots + \boldsymbol{a}_t = \boldsymbol{0}, \tag{5.5.7}$$

每个 $\boldsymbol{a}_i$ 是同一个特征子空间 V_{λ_i} 的基 $S_i = \{\boldsymbol{X}_{i1}, \cdots, \boldsymbol{X}_{in_i}\}$ 的线性组合, 仍在 V_{λ_i} 中, 满足条件 $(\boldsymbol{A} - \lambda_i \boldsymbol{I})\boldsymbol{a}_i = \boldsymbol{0}$ 即 $\boldsymbol{A}\boldsymbol{a}_i = \lambda_i \boldsymbol{a}_i$. 因而 $\boldsymbol{A}^k \boldsymbol{a}_i = \lambda_i^k \boldsymbol{a}_i$ 对任意正整数 k 成立. 进而 $f(\boldsymbol{A})\boldsymbol{a}_i = f(\lambda_i)\boldsymbol{a}_i$ 对任意多项式 $f(x)$ 成立. 对每个 $1 \leqslant i \leqslant t$, 取

$$f_i(x) = \prod_{1 \leqslant p \leqslant t, p \neq i} (x - \lambda_p) = (x - \lambda_1) \cdots (x - \lambda_{i-1})(x - \lambda_{i+1}) \cdots (x - \lambda_t),$$

这是除了 $x - \lambda_i$ 之外的所有的 $x - \lambda_p$ 的乘积, 则 $f_i(\lambda_p) = 0$ (当 $p \neq i$) 且 $f_i(\lambda_i) \neq 0$. 用 $f_i(\boldsymbol{A})$ 左乘 (5.5.7) 等式两边, 得到

$$f_i(\lambda_1)\boldsymbol{a}_1 + \cdots + f_i(\lambda_t)\boldsymbol{a}_t = \boldsymbol{0}, \tag{5.5.8}$$

即 $f_i(\lambda_i)\boldsymbol{a}_i = \boldsymbol{0}$, 进而 $\boldsymbol{a}_i = \boldsymbol{0}$. 也就是

$$\mu_{i1}\boldsymbol{X}_{i1} + \cdots + \mu_{in_i} X_{in_i} = \boldsymbol{0},$$

但 $\boldsymbol{X}_{i1}, \cdots, \boldsymbol{X}_{in_i}$ 是 V_{λ_i} 的基向量, 线性无关, 因此所有的系数 $\mu_{i1} = \cdots = \mu_{in_i} = 0$. 这证明了 (5.5.6) 中所有的系数 $\mu_{ij} = 0$, 特征向量集合 $S = S_1 \cup \cdots \cup S_t$ 线性无关, 组成 $\mathbf{C}^{n \times 1}$ 的基, 以它们为各列排成的矩阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于对角阵. □

定理 5.5.2 的上述证明过程中得到: 属于同一个方阵 $\boldsymbol{A}$ 的不同特征值的特征子空间 V_{λ_i} 的基的并集线性无关, 这个并集的子集当然也线性无关. 特别地, 有:

推论 5.5.2　属于同一个方阵 $\boldsymbol{A}$ 的不同特征值的特征向量线性无关. □

方阵的每个特征值 λ_i 都存在特征向量 $\boldsymbol{X}_i \in V_{\lambda_i}$, 可见几何重数 $\dim V_{\lambda_i} \geqslant 1$. 如果 n 阶方阵 $\boldsymbol{A}$ 有 n 个不同的特征值, 每个特征值 λ_i 的代数重数与几何重数都只能为 1, 定理 5.5.2 的条件必然满足. 由此得到:

推论 5.5.3　如果 n 阶方阵 $\boldsymbol{A}$ 有 n 个不同的特征值, 则 $\boldsymbol{A}$ 相似于对角矩阵. □

引理 5.5.7　n 阶方阵 $\boldsymbol{A}$ 的每个特征值的几何重数 $\leqslant$ 代数重数.

证明　属于 $\boldsymbol{A}$ 的各个不同特征值 λ_i 的特征子空间 V_{λ_i} 的基 S_i 的并集 $S=S_1\cup\cdots\cup S_t$ 由线性无关特征向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_d$ 组成, 其中的向量个数 d 等于各特征值 λ_i 的几何重数 d_i 之和. 将 S 扩充为 $\mathbf{C}^{n\times 1}$ 的一组基 $T=\{\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n\}$, 以 T 中各向量为各列排成可逆方阵 $\boldsymbol{P}$. 每个 $\boldsymbol{A}\boldsymbol{X}_j$ 可以写成 T 的线性组合

$$\boldsymbol{A}\boldsymbol{X}_j=b_{1j}\boldsymbol{X}_1+\cdots+b_{nj}\boldsymbol{X}_n=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)(b_{1j},\cdots,b_{nj})^{\mathrm{T}}=\boldsymbol{P}\boldsymbol{B}_j\ ,$$

其中的列向量 $\boldsymbol{B}_j$ 是 $\boldsymbol{A}\boldsymbol{X}_j$ 在基 T 下的坐标. 于是

$$\boldsymbol{A}\boldsymbol{P}=(\boldsymbol{A}\boldsymbol{X}_1,\cdots,\boldsymbol{A}\boldsymbol{X}_n)=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=\boldsymbol{P}\boldsymbol{B}.$$

$\boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$ 是以 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n$ 为各列排成的矩阵. $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 与 $\varphi_{\boldsymbol{B}}(\lambda)$ 相等, 每个特征值 λ_i 在 $\boldsymbol{B}$ 中的代数重数也就是它在 $\boldsymbol{A}$ 中的代数重数 n_i.

$\boldsymbol{B}$ 的前 d 列 $\boldsymbol{X}_j$ $(1\leqslant j\leqslant d)$ 是 $\boldsymbol{A}$ 的特征向量, $\boldsymbol{A}\boldsymbol{X}_j=\lambda_j\boldsymbol{X}_j$ 在 T 下的坐标 $\boldsymbol{B}_j=\lambda_j\boldsymbol{e}_j$ 是自然基向量 $\boldsymbol{e}_j$ 的 λ_j 倍. 如果 $d=n$, 则 $\boldsymbol{B}=(\lambda_1,\cdots,\lambda_n)$ 是对角阵, 每个特征值 λ_i 在 $\boldsymbol{B}$ 的对角元中出现的次数就是 λ_i 在 $\boldsymbol{A}$ 中的代数重数, 等于它的几何重数 d_i. 当 $d<n$ 时, $\boldsymbol{B}$ 具有分块形式

$$\boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=\begin{pmatrix}\boldsymbol{B}_{11}&\boldsymbol{B}_{12}\\ \boldsymbol{O}&\boldsymbol{B}_{22}\end{pmatrix},\tag{5.5.9}$$

其中 $\boldsymbol{B}_{11}=\mathrm{diag}(\lambda_1,\cdots,\lambda_d)$ 是 d 阶对角矩阵, $\boldsymbol{B}_{22}$ 为 $n-d$ 阶方阵. $\boldsymbol{B}$ 的特征多项式

$$\varphi_{\boldsymbol{B}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{B}|=|\lambda\boldsymbol{I}-\boldsymbol{B}_{11}||\lambda\boldsymbol{I}-\boldsymbol{B}_{22}|=\varphi_{\boldsymbol{B}_{11}}(\lambda)\varphi_{\boldsymbol{B}_{22}}(\lambda),$$

其中 $\varphi_{\boldsymbol{B}_{11}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{B}_{11}|=(\lambda-\lambda_1)\cdots(\lambda-\lambda_d)$ 是 $\varphi_{\boldsymbol{B}}(\lambda)$ 的因式, λ_i 在其中的重数等于 d_i, $(\lambda-\lambda_i)^{d_i}$ 整除 $\varphi_{\boldsymbol{B}_{11}}(\lambda)$ 从而整除 $\varphi_{\boldsymbol{A}}(\lambda)=\varphi_{\boldsymbol{B}}(\lambda)$, 这证明了 λ_i 在 $\varphi_{\boldsymbol{A}}(\lambda)$ 中的重数 $n_i\geqslant d_i$, 也就是: $\boldsymbol{A}$ 的特征值 λ_i 的代数重数 $n_i\geqslant$ 几何重数 d_i. □

引理 5.5.7 的证明过程中将复方阵 $\boldsymbol{A}$ 相似到具有形式 (5.5.9) 的准上三角形矩阵 $\boldsymbol{B}$, 其中右下角的块 $\boldsymbol{B}_{22}$ 又可以进一步相似到同样的形状. 照此下去可以将 $\boldsymbol{A}$ 相似到上三角形矩阵.

2. 三角化

引理 5.5.8　每个复方阵 $\boldsymbol{A}$ 可以相似于上三角形矩阵.

证明　对 $\boldsymbol{A}$ 的阶数 n 作数学归纳法. 当 $n=1$ 时显然. 假定 $n-1$ 阶复方阵可以相似于上三角形矩阵.

任取 $\boldsymbol{A}$ 的特征值 λ_1 和特征向量 $\boldsymbol{X}_1$, 将 $\boldsymbol{X}_1$ 扩充为 $\mathbf{C}^{n\times 1}$ 的基 $S=\{\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n\}$, 与引理 5.5.7 类似, 以 T 中各向量为各列排成可逆方阵 $\boldsymbol{P}_1=$

$(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$, 将每个 $\boldsymbol{AX}_j$ 写成 T 的线性组合 $b_{1j}\boldsymbol{X}_1+\cdots+b_{nj}\boldsymbol{X}_n=\boldsymbol{PB}_j$, 其中 $\boldsymbol{B}_j$ 是 $\boldsymbol{AX}_j$ 在基 T 下的坐标, 则 $\boldsymbol{AP}_1=\boldsymbol{P}_1\boldsymbol{B}$ 从而 $\boldsymbol{B}=\boldsymbol{P}_1^{-1}\boldsymbol{AP}_1$ 对 $\boldsymbol{B}=(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$ 成立. 由 $\boldsymbol{AX}_1=\lambda_1\boldsymbol{X}_1$ 知 $\boldsymbol{B}_1=(\lambda_1,0,\cdots,0)^{\mathrm{T}}$, 从而

$$\boldsymbol{B}=\begin{pmatrix}\lambda_1 & \boldsymbol{B}_{12}\\ \boldsymbol{0} & \boldsymbol{B}_{22}\end{pmatrix}.$$

由归纳假设, 存在 $n-1$ 阶可逆方阵 $\boldsymbol{P}_2$ 将 $n-1$ 阶复方阵 $\boldsymbol{B}_{22}$ 相似到上三角形矩阵 $\boldsymbol{\Lambda}_{22}=\boldsymbol{P}_2^{-1}\boldsymbol{B}_{22}\boldsymbol{P}_2$. 于是

$$\boldsymbol{\Lambda}=\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_2\end{pmatrix}^{-1}\boldsymbol{B}\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{0} & \boldsymbol{\Lambda}_{22}\end{pmatrix}=\boldsymbol{P}^{-1}\boldsymbol{AP}$$

是与 $\boldsymbol{A}$ 相似的上三角形矩阵, 其中

$$\boldsymbol{P}=\boldsymbol{P}_1\begin{pmatrix}1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_2\end{pmatrix}.$$

□

3. 化零多项式

定理 5.5.3(凯莱 – 哈密顿 (Cayley-Hamilton) 定理) 设 $\varphi_{\boldsymbol{A}}(\lambda)$ 是方阵 $\boldsymbol{A}$ 的特征多项式, 则 $\varphi_{\boldsymbol{A}}(\boldsymbol{A})=\boldsymbol{O}$.

证明 由引理 5.5.8, 存在复可逆方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似到上三角形矩阵 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}=(\lambda_{ij})_{n\times n}$, 其对角元 $\lambda_{11},\cdots,\lambda_{nn}$ 就是 $\boldsymbol{A}$ 的全部特征值. 特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=\varphi_{\boldsymbol{\Lambda}}(\lambda)=(\lambda-\lambda_{11})\cdots(\lambda-\lambda_{nn})$. 只要证明 $\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})=0$, 则 $\varphi_{\boldsymbol{A}}(\boldsymbol{A})=\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{P\Lambda P}^{-1})=\boldsymbol{P}\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})\boldsymbol{P}^{-1}=\boldsymbol{O}$ 成立.

对 $\boldsymbol{\Lambda}$ 的阶 n 作数学归纳法证明 $\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})=0$. $n=1$ 时显然. 假定 $\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})$ 对 $n-1$ 阶上三角形矩阵 $\boldsymbol{\Lambda}$ 成立. 将 $\boldsymbol{\Lambda}$ 分块为

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_{11} & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{0} & \boldsymbol{\Lambda}_{22}\end{pmatrix},$$

其中 $\boldsymbol{\Lambda}_{22}$ 是 $n-1$ 阶上三角形矩阵, 特征多项式 $\varphi_{\boldsymbol{\Lambda}_{22}}(\lambda)=(\lambda-\lambda_{22})\cdots(\lambda-\lambda_{nn})$. 根据归纳假设, 有 $\varphi_{\boldsymbol{\Lambda}_{22}}(\boldsymbol{\Lambda}_{22})=\boldsymbol{O}$. 于是

$$\begin{aligned}\varphi_{\boldsymbol{\Lambda}}(\boldsymbol{\Lambda})&=(\boldsymbol{\Lambda}-\lambda_{11}\boldsymbol{I})\cdots(\boldsymbol{\Lambda}-\lambda_{nn}\boldsymbol{I})\\&=\begin{pmatrix}0 & \boldsymbol{\Lambda}_{12}\\ 0 & \boldsymbol{\Lambda}_{22}-\lambda_{11}\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\lambda_{11}-\lambda_{22} & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{0} & \boldsymbol{\Lambda}_{22}-\lambda_{22}\boldsymbol{I}\end{pmatrix}\end{aligned}$$

$$\cdots\begin{pmatrix}\lambda_{11}-\lambda_{nn} & \Lambda_{12}\\ 0 & \Lambda_{22}-\lambda_{nn}\boldsymbol{I}\end{pmatrix}$$

$$=\begin{pmatrix}0 & \Lambda_{12}\\ \mathbf{0} & \Lambda_{22}-\lambda_{11}\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\varphi_{\Lambda_{22}}(\lambda_1) & *\\ \mathbf{0} & \varphi_{\Lambda_{22}}(\Lambda_{22})\end{pmatrix}$$

$$=\begin{pmatrix}0 & \Lambda_{12}\\ \mathbf{0} & \Lambda_{22}-\lambda_{11}\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\varphi_{\Lambda_{22}}(\lambda_1) & *\\ \mathbf{0} & \boldsymbol{O}_{(n-1)}\end{pmatrix}=\boldsymbol{O}.$$

从而$\varphi_{\boldsymbol{A}}(\boldsymbol{A})=O$. □

如果将数 c 代入多项式 $f(x)$ 得到 $f(c)=0$, 就称 c 是多项式 $f(x)$ 的根. 因此, 定理 5.5.3 可以理解为, 方阵 $\boldsymbol{A}$ 是它自己的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 的" 根". 反过来, $\varphi_{\boldsymbol{A}}(\lambda)$ 称为 $\boldsymbol{A}$ 的**化零多项式** (annihilator). 显然, $\varphi_{\boldsymbol{A}}(\lambda)$ 的任何一个倍式 $f(\lambda)=q(\lambda)\varphi_{\boldsymbol{A}}(\lambda)$ 也是 $\boldsymbol{A}$ 的化零多项式. 反过来, 可能有更低次数的多项式 $f(\lambda)$ 满足 $g(\boldsymbol{A})=\boldsymbol{O}$, 因而 $g(\lambda)$ 也是 $\boldsymbol{A}$ 的化零多项式. 对每个方阵 $\boldsymbol{A}$, 存在最低次数的首项系数为 1 的多项式 $d(\lambda)$ 使 $d(\boldsymbol{A})=\boldsymbol{O}$, 这个 $d(\lambda)$ 是唯一的, 称为 $\boldsymbol{A}$ 的**最小多项式** (minimal polynomial). $\boldsymbol{A}$ 的所有的化零多项式都是最小多项式的倍式. 例如, 例 3 的方阵 $\boldsymbol{A}$ 相似于对角阵 $\boldsymbol{D}=\mathrm{diag}(3,-1,3)=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, $\boldsymbol{A}$ 与 $\boldsymbol{D}$ 的特征多项式都是 $(\lambda-3)^2(\lambda+1)$, 最小多项式则都是 $(\lambda-3)(\lambda+1)$.

§5.6 正 交 相 似

5.6.1 二次曲线与二次曲面方程的化简

§5.5 例 4 将平面直角坐标系中的方程 $x^2+2xy+5y^2=4$ 通过坐标轴的旋转化成了椭圆的标准方程. 类似地可以将一般的二元或三元二次方程化为二次曲线或二次曲面的标准方程.

例 1 空间直角坐标系中, 方程 $x^2+y^2+z^2+2xy+2xz-2yz=2$ 的图像是什么曲面?

解 方程 $x^2+y^2+z^2+2xy+2xz-2yz=2$ 左边的二次型的方阵为

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

$\boldsymbol{A}$ 的特征多项式 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda+1)(\lambda-2)^2$, 特征值为 $-1,2(2$ 重).

对特征值 -1, 解方程组 $(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得 $\boldsymbol{X}_1=(-1,1,1)^{\mathrm{T}}$ 组成解空间 V_{-1} 的基.

对特征值 2, 解方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得 $\boldsymbol{X}_2=(1,1,0)^{\mathrm{T}},\boldsymbol{X}_3=(1,0,1)^{\mathrm{T}}$ 组成解空间 V_2 的基.

$\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2,\boldsymbol{X}_3$ 的内积 $\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=0=\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_3$, 可见 $\boldsymbol{X}_1$ 与 $\boldsymbol{X}_2,\boldsymbol{X}_3$ 都垂直, 因而直线 V_{-1} 与平面 V_2 垂直. V_{-1} 中的单位向量 $\boldsymbol{\xi}_1$ 与 V_2 中的标准正交基 $\{\boldsymbol{\xi}_2,\boldsymbol{\xi}_3\}$ 共同组成 $\mathbf{R}^{3\times1}$ 的标准正交基 S. 如果 S 是左手系, 用 $-\boldsymbol{\xi}_1$ 代替 $\boldsymbol{\xi}_1$ 可得到右手系标准正交基. 总之, 能够求出 $\mathbf{R}^{3\times1}$ 的右手系标准正交基 S, 以其中各向量为 Ox',Oy',Oz' 轴正方向上的单位向量, 建立新的直角坐标系. 则每一点 P 的原坐标 $\boldsymbol{X}=(x,y,z)^{\mathrm{T}}$ 与新坐标 $\boldsymbol{Y}=(x',y',z')^{\mathrm{T}}$ 之间有变换公式 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$, 其中 $\boldsymbol{U}$ 是由 S 的各列排成的正交方阵且 $\det\boldsymbol{U}=1$.

曲面方程左边的二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$ 经过坐标变换 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$ 变成 $Q(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{Y}$, 其中 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}(-1,2,2)$, 因而 $Q(\boldsymbol{Y})=-x'^2+2y'^2+2y'^2$. 曲面在新坐标系下的方程为

$$-x'^2+2y'^2+2z'^2=2, \text{即} -\frac{x'^2}{2}+y'^2+z'^2=1,$$

图像是单叶旋转双曲面, 可由 $x'Oy'$ 平面内的双曲线 $-\dfrac{x'^2}{2}+y'^2=1$ 绕 x' 轴旋转得到. □

例 1 中只要求判断曲面的形状, 因此不必具体求出标准正交基 S 中的各向量. 只要由特征子空间 V_{-1} 与 V_2 相互垂直证明存在由特征向量组成的标准正交基 S 就行了.

例 2 求行列式为 1 的正交方阵 $\boldsymbol{U}$, 使其将实对称方阵

$$\boldsymbol{A}=\begin{pmatrix}1&1&1\\1&1&-1\\1&-1&1\end{pmatrix}$$

相合到对角阵 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$.

解 与例 1 同样先求出 $\boldsymbol{A}$ 的特征值 $-1,2$, 解方程组 $(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得 $\boldsymbol{X}_1=(-1,1,1)^{\mathrm{T}}$ 组成基础解系, 解方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得 $\boldsymbol{X}_2=(1,1,0)^{\mathrm{T}},\boldsymbol{X}_3=(1,0,1)^{\mathrm{T}}$ 组成基础解系.

对 $\boldsymbol{X}_2,\boldsymbol{X}_3$ 作正交化, 得到

$$\boldsymbol{Y}_3=\boldsymbol{X}_3-\frac{\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_3}{\boldsymbol{X}_2^{\mathrm{T}}\boldsymbol{X}_2}\boldsymbol{X}_2=\boldsymbol{X}_3-\frac{1}{2}\boldsymbol{X}_2=\left(\frac{1}{2},-\frac{1}{2},1\right)^{\mathrm{T}}$$

取代 $\boldsymbol{X}_3$, 则 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y}_3$ 是由 $\boldsymbol{A}$ 的特征向量组成的 $\mathbf{R}^{3\times1}$ 的正交基. 将它们分

别除以各自的模, 得到的三个单位向量

$$\boldsymbol{\xi}_1=\frac{1}{\sqrt{3}}(-1,1,1)^{\mathrm{T}},\boldsymbol{\xi}_2=\frac{1}{\sqrt{2}}(1,1,0)^{\mathrm{T}},\boldsymbol{\xi}_3=\frac{1}{\sqrt{6}}(1,-1,2)^{\mathrm{T}}$$

组成标准正交基. 以它们为各列排成的方阵 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)$ 是正交方阵. 但 $\det\boldsymbol{P}=-1$. 用 $-\boldsymbol{\xi}_1$ 代替 $\boldsymbol{\xi}_1$, 得到行列式为 1 的正交方阵

$$\boldsymbol{U}=(-\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=\begin{pmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{3}}&0&\frac{2}{\sqrt{6}}\end{pmatrix},$$

将 $\boldsymbol{A}$ 相合到对角矩阵

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\operatorname{diag}(-1,2,2).$$ □

例 3　求平面直角坐标系中方程 $x^2+4xy+4y^2-2x+6y-5=0$ 的图像曲线的形状.

解　方程左边的二次项部分 $Q(x,y)=x^2+4xy+4y^2=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix},\boldsymbol{A}=\begin{pmatrix}1&2\\2&4\end{pmatrix}.$$

$\boldsymbol{A}$ 的特征多项式 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^2-5\lambda$, 特征值为 $5,0$.

对特征值 5, 解方程组 $(\boldsymbol{A}-5\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得基础解 $\boldsymbol{X}_1=(1,2)^{\mathrm{T}}$.

对特征值 0, 解方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{0}$ 得基础解 $X_2=(-2,1)^{\mathrm{T}}$.

$\boldsymbol{X}_1,\boldsymbol{X}_2$ 的内积 $\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{X}_2=(1,2)\begin{pmatrix}-2\\1\end{pmatrix}=0$, 相互垂直. 分别除以各自的长度 $|\boldsymbol{X}_1|=|\boldsymbol{X}_2|=\sqrt{5}$ 得相互垂直的单位向量 $\boldsymbol{\xi}_1=\frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix},\boldsymbol{\xi}_2=\frac{1}{\sqrt{5}}\begin{pmatrix}-2\\1\end{pmatrix}$, 且 $\boldsymbol{U}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ 的行列式 $|\boldsymbol{U}|=1$. 分别以 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 为 x' 轴、y' 轴正方向上的单位向量建立直角坐标系, 则 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}=\operatorname{diag}(5,0)$. 每个点 $P(x,y)$ 在这个新坐标系下的坐标 $\boldsymbol{\xi}=(x',y')^{\mathrm{T}}$ 与原坐标 $\boldsymbol{X}=(x,y)^{\mathrm{T}}$ 之间有坐标变换公式 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{\xi}$, 即

$$x=\frac{x'-2y'}{\sqrt{5}},\quad y=\frac{2x'+y'}{\sqrt{5}}.$$

于是 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}=(\boldsymbol{U}\boldsymbol{\xi})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{U}\boldsymbol{\xi})=\boldsymbol{\xi}^{\mathrm{T}}(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U})\boldsymbol{\xi}=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{\xi}=5x'^2$. 曲线在新坐标系下的方程为

$$5x'^2-2\frac{x'-2y'}{\sqrt{5}}+6\frac{2x'+y'}{\sqrt{5}}-5=0,$$

即

$$5x'^2+2\sqrt{5}\,x'+2\sqrt{5}\,y'-5=0,$$

方程两边同乘 $\frac{1}{5}$, 并对左边配方得

$$\left(x'+\frac{1}{\sqrt{5}}\right)^2+\frac{2}{\sqrt{5}}\left(y'-\frac{3}{\sqrt{5}}\right)=0$$

将新直角坐标系 $Ox'y'$ 平移, 使原点移到点 $O'\left(-\frac{1}{\sqrt{5}},\frac{3}{\sqrt{5}}\right)$. 则在坐标系 $Ox'y'$ 中坐标为 (x',y') 的点在平移之后的坐标系中的坐标 $(x'',y'')=\left(x'+\frac{1}{\sqrt{5}},y'-\frac{3}{\sqrt{5}}\right)$. 曲线方程变成

$$x''^2=-\frac{2\sqrt{5}}{5}y''.$$

这是抛物线的标准方程. 图像曲线是抛物线, 焦点与准线的距离为 $\frac{\sqrt{5}}{5}$. □

5.6.2 实对称方阵的正交相似

§5.5 例 4 以及本节例 1、例 3 中的方程都具有形式 $f(\boldsymbol{X})=0$ (其中 $\boldsymbol{X}=(x,y)^{\mathrm{T}}$ 或 $(x,y,z)^{\mathrm{T}}$), f 是二次多项式. 这样的方程 $f(\boldsymbol{X})=0$ 的图像称为二次曲线或二次曲面. 在解析几何中学习过二次曲线的标准方程. 这几个例题中的方程都不是标准方程, 我们通过坐标变换将方程化成了标准方程, 从而知道了曲线或曲面的形状. 将二次方程化成标准方程的关键是: 将二次多项式 $f(\boldsymbol{X})$ 中的二次项组成的二次型 $Q(\boldsymbol{X})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{X}$ 的对称方阵 $\boldsymbol{S}$ 用适当的正交方阵 $\boldsymbol{U}$ 相合到对角矩阵 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{U}$, 则二次型 $Q(\boldsymbol{X})$ 经过变换 $\boldsymbol{X}=\boldsymbol{U}\boldsymbol{Y}$ 化成了标准形 $Q_1(\boldsymbol{Y})=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{Y}$.

由于 $\boldsymbol{U}$ 是正交方阵, 满足条件 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}$ 即 $\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}^{-1}$. $\boldsymbol{S}$ 与 $\boldsymbol{D}$ 的相合关系式 $\boldsymbol{D}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{U}$ 可以写成 $\boldsymbol{D}=\boldsymbol{U}^{-1}\boldsymbol{S}\boldsymbol{U}$.

当 $\boldsymbol{U}$ 是正交方阵时, $\boldsymbol{U}^{-1}=\boldsymbol{U}^{\mathrm{T}}$, 通过 $\boldsymbol{U}$ 与 $\boldsymbol{A}$ 相似的方阵 $\boldsymbol{B}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ 也与 $\boldsymbol{A}$ 相合, 在此情形下称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交相似. 前述几个例题中实对称方阵 $\boldsymbol{A}$ 正交相似到对角矩阵 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}$ 的算法适用于所有的实对称方阵.

算法 5.6.1(实对称方阵正交相似到对角矩阵) 设 $\boldsymbol{A}$ 是 n 阶实对称方阵.

1. 求出特征多项式 $\varphi_A(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的全部复数根 λ_i, 一定都是实数.

2. 对每个 λ_i, 求出齐次线性方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 实数域上的解空间(特征子空间) V_{λ_i} 的一组基 S_i.

3. 将每个特征子空间 V_{λ_i} 的基 S_i 正交化和单位化得到 V_{λ_i} 的一组标准正交基 T_i. 各 T_i 的并集一定是 $\mathbf{R}^n$ 的一组标准正交基. 以其中的基向量为各列排成正交方阵 $\boldsymbol{U}$. 则 $\boldsymbol{D}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ 是对角矩阵, 它的对角元就是 $\boldsymbol{A}$ 的全部特征根. □

如果 $\det\boldsymbol{U}=-1$, 将它的任何一列乘 -1 或任何两列互换, 可以化成 $\det\boldsymbol{U}=1$ 的情形.

很自然对算法 5.6.1 提出以下疑问:

1. 实系数 n 次多项式 $f(x)$ 的复数根可能是虚数. 实对称方阵 $\boldsymbol{A}$ 的特征多项式 $|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的复数根为什么一定都是实数?

2. 属于每个实特征值 λ_i 的特征子空间 V_{λ_i} 一定存在标准正交基 T_i. 为什么所有这些 T_i 的并集一定线性无关, 一定是 $\mathbf{R}^{n\times 1}$ 的基, 一定两两正交?

如下的定理回答了这些疑问.

定理 5.6.1 任意 n 阶实对称方阵 $\boldsymbol{S}$ 可通过正交方阵相似到对角矩阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{S}\boldsymbol{P}$. □

为了证明定理 5.6.1, 先证明:

引理 5.6.1 实对称方阵 $\boldsymbol{S}$ 的特征值一定都是实数.

证明 对任意复矩阵 $\boldsymbol{A}$, 记 $\boldsymbol{A}^*=\overline{\boldsymbol{A}}^{\mathrm{T}}$ 为 $\boldsymbol{A}$ 的转置共轭矩阵. 则实对称方阵 $\boldsymbol{S}$ 满足 $\boldsymbol{S}^*=\boldsymbol{S}$. 且易验证 $(\boldsymbol{A}^*\boldsymbol{B}\boldsymbol{A})^*=\boldsymbol{A}^*\boldsymbol{B}^*\boldsymbol{A}$.

设 λ 是 $\boldsymbol{S}$ 的任意复特征值, $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 是属于 λ 的特征向量, 即 $\boldsymbol{S}\boldsymbol{X}=\lambda\boldsymbol{X}$. 在等式 $\boldsymbol{S}\boldsymbol{X}=\lambda\boldsymbol{X}$ 两边同时左乘 $\boldsymbol{X}^*=(\overline{x_1},\cdots,\overline{x_n})$ 得 $\boldsymbol{X}^*\boldsymbol{S}\boldsymbol{X}=\lambda\boldsymbol{X}^*\boldsymbol{X}$. 由 $(\boldsymbol{X}^*\boldsymbol{S}\boldsymbol{X})^*=\boldsymbol{X}^*\boldsymbol{S}^*\boldsymbol{X}=\boldsymbol{X}^*\boldsymbol{S}\boldsymbol{X}$ 知一阶方阵 $\boldsymbol{X}^*\boldsymbol{S}\boldsymbol{X}$ 共轭转置之后等于自己, 是实数. 且 $\boldsymbol{X}^*\boldsymbol{X}=\overline{x_1}x_1+\cdots+\overline{x_n}x_n=|x_1|^2+\cdots+|x_n|^2$ 是正实数. 因此 $\lambda=\dfrac{\boldsymbol{X}^*\boldsymbol{S}\boldsymbol{X}}{\boldsymbol{X}^*\boldsymbol{X}}$ 是实数. □

容易看出, 引理 5.6.1 的证明过程其实并不需要 $\boldsymbol{S}$ 是实方阵. 只要 $\boldsymbol{S}$ 是满足条件 $\boldsymbol{S}^*=\boldsymbol{S}$ 的复方阵, 证明过程就仍然正确, $\boldsymbol{S}$ 的特征值就都是实数.

满足条件 $\boldsymbol{H}^*=\boldsymbol{H}$ 的复方阵 $\boldsymbol{H}$ 称为埃尔米特 (Hermite) 方阵. 元素全是实数的埃尔米特方阵就是实对称方阵. 实际上我们证明了:

引理 5.6.2 埃尔米特方阵的全部特征值都是实数.

引理 5.6.1 是引理 5.6.2 的特殊情形.

定理 5.6.1 的证明

对 n 阶实对称方阵 $\boldsymbol{A}$ 的阶 n 作数学归纳法, 证明 $\boldsymbol{A}$ 正交相似于对角矩阵.

$n=1$ 时, 一阶实对称方阵已经是对角矩阵, 命题显然成立.

设 $n-1$ 阶实对称方阵正交相似于对角矩阵. 证明 n 阶实对称方阵 $\boldsymbol{S}$ 正交相似于对角矩阵.

任取 $\boldsymbol{A}$ 的一个特征值 λ_1. 由引理 5.6.1 知 λ_1 是实数, 方程组 $(\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 在实数域上有非零解 $\boldsymbol{X}_1$, 也就是属于特征值 λ_1 的特征向量. 与 $\boldsymbol{X}_1$ 同方向上的单位向量 $\boldsymbol{\xi}_1=\dfrac{1}{|\boldsymbol{X}_1|}\boldsymbol{X}_1$ 也是属于特征值 λ_1 的特征向量, 满足 $\boldsymbol{A}\boldsymbol{\xi}_1=\lambda_1\boldsymbol{\xi}_1$. $\boldsymbol{\xi}_1$ 可以扩充为欧氏空间 $\mathbf{R}^{n\times 1}$ 的标准正交基 $\boldsymbol{T}=\{\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n\}$. 以这些基向量为各列排成的方阵 $\boldsymbol{U}_1=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n)$ 是正交方阵. 且

$$\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{A}(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n)=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_n)\boldsymbol{B}=\boldsymbol{U}_1\boldsymbol{B},$$

其中 $\boldsymbol{B}$ 的各列 $\boldsymbol{b}_i$ 分别是各 $\boldsymbol{A}\boldsymbol{\xi}_i$ 在基 T 下的坐标. 特别地, $\boldsymbol{A}\boldsymbol{\xi}_1=\lambda_1\boldsymbol{\xi}_1$ 在基 T 下的坐标 $\boldsymbol{b}_1=(\lambda_1,0,\cdots,0)^{\mathrm{T}}$. 因此 $\boldsymbol{B}$ 具有形式

$$\boldsymbol{B}=\begin{pmatrix}\lambda_1 & \boldsymbol{B}_{12}\\ \boldsymbol{O} & \boldsymbol{B}_{22}\end{pmatrix},$$

且由 $\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{U}_1\boldsymbol{B}$ 知 $\boldsymbol{B}=\boldsymbol{U}_1^{-1}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{U}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}_1$ 与 $\boldsymbol{A}$ 相似, 也与 $\boldsymbol{A}$ 相合.

$\boldsymbol{A}$ 是对称方阵, 与 $\boldsymbol{A}$ 相合的 $\boldsymbol{B}=\boldsymbol{U}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}_1$ 仍应是对称方阵. 这迫使 $\boldsymbol{B}_{12}=\boldsymbol{O}$, 且 $\boldsymbol{B}_{22}$ 是 $n-1$ 阶实对称方阵. 由归纳假设知: 存在 $n-1$ 阶正交方阵 $\boldsymbol{U}_2$ 将 $\boldsymbol{B}_{22}$ 相似到对角矩阵

$$\boldsymbol{U}_2^{-1}\boldsymbol{B}_{22}\boldsymbol{U}_2=\boldsymbol{D}_2=\begin{pmatrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

于是 $\boldsymbol{U}=\boldsymbol{U}_1\begin{pmatrix}1 & \\ & \boldsymbol{U}_2\end{pmatrix}$ 将 $\boldsymbol{S}$ 正交相似到对角矩阵

$$\boldsymbol{U}^{-1}\boldsymbol{S}\boldsymbol{U}=\boldsymbol{D}=\begin{pmatrix}\lambda_1 & \\ & \boldsymbol{D}_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$ □

推论 5.6.1 实对称方阵 $\boldsymbol{A}$ 的不同特征值 λ_i,λ_j 的特征向量 $\boldsymbol{X}_i,\boldsymbol{X}_j$ 相互垂直.

证法 1 存在正交方阵 $\boldsymbol{U}$ 将 $\boldsymbol{A}$ 相似到对角矩阵 $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{D}$. $\boldsymbol{U}$ 的各列都是 $\boldsymbol{A}$ 的特征向量, 且两两相互正交, 其中属于特征值 λ_i 的特征向量组成特

征子空间 V_{λ_i} 的标准正交基 T_i, 属于特征值 λ_j 的特征向量组成 V_{λ_j} 的标准正交基 T_j. 属于特征值 λ_i 的特征向量 $\boldsymbol{X}_i$ 都是 T_i 的线性组合 $\boldsymbol{X}_i=\sum_{\boldsymbol{\xi}\in T_i}\lambda_{\xi}\boldsymbol{\xi}$, $\boldsymbol{X}_j$ 则是 T_j 的线性组合 $\boldsymbol{X}_j=\sum_{\boldsymbol{\eta}\in T_j}\lambda_{\eta}\boldsymbol{\eta}$. T_i 中每个基向量 $\boldsymbol{\xi}$ 与 T_j 中每个基向量 η 是 $\boldsymbol{U}$ 的不同列, 相互内积 $(\boldsymbol{\xi},\boldsymbol{\eta})=0$. 因此

$$(\boldsymbol{X}_i,\boldsymbol{X}_j)=\left(\sum_{\boldsymbol{\xi}}\lambda_{\xi}\boldsymbol{\xi},\sum_{\boldsymbol{\eta}}\lambda_{\eta}\boldsymbol{\eta}\right)=\sum_{\boldsymbol{\xi},\boldsymbol{\eta}}\lambda_{\xi}\lambda_{\eta}(\boldsymbol{\xi},\boldsymbol{\eta})=0.$$

证法 2 我们有 $\boldsymbol{AX}_i=\lambda_i\boldsymbol{X}_i$, $\boldsymbol{AX}_j=\lambda_j\boldsymbol{X}_j$. 将

$$\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{AX}_j=\boldsymbol{X}_i^{\mathrm{T}}(\lambda_j\boldsymbol{X}_j) \text{ 与 } \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{AX}_j=(\boldsymbol{AX}_i)^{\mathrm{T}}\boldsymbol{X}_j=\lambda_i\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_j$$

相减得 $(\lambda_j-\lambda_i)\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_j=0$. 由 $\lambda_j-\lambda_i\neq 0$ 即得 $\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_j=0$, $\boldsymbol{X}_i\perp\boldsymbol{X}_j$. □

5.6.3 三维几何空间中的旋转与对称

例 4 求实方阵

$$\boldsymbol{A}=\begin{pmatrix}\frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\end{pmatrix}$$

的实特征值与实特征向量. 并分析 $\boldsymbol{A}$ 在三维几何空间 $V=\mathbf{R}^{3\times 1}$ 中引起的变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 的几何性质.

解法 1 $\boldsymbol{A}$ 的特征多项式

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

$$\xlongequal[-(1)+(2),-(1)+(3)]{}\begin{vmatrix}\lambda-1 & 0 & 0\\ \frac{1}{3} & \lambda-1 & -1\\ -\frac{2}{3} & 1 & \lambda\end{vmatrix}=(\lambda-1)(\lambda^2-\lambda+1).$$

方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-1)(\lambda^2-\lambda+1)=0$ 的唯一实根为 1, 也就是 $\boldsymbol{A}$ 的唯一的实特征值.

对特征值 1 解方程组 $(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 求得特征向量 $\boldsymbol{X}_1=(1,1,1)^{\mathrm{T}}$ 组成解空间 V_1 的基. V_1 是过原点和点 $(1,1,1)$ 的直线, V_1 中所有的向量在 σ 作用下保持不动, 其中所有的非零向量就是属于特征值 1 的全体特征向量.

为了分析变换 $\sigma: X\mapsto \boldsymbol{AX}$ 的几何性质, 将 V_1 中的单位向量 $\boldsymbol{e}_1'=\dfrac{1}{|\boldsymbol{X}_1|}\boldsymbol{X}_1$ 扩充为 $\mathbf{R}^{3\times 1}$ 的一组右手系标准正交基 $E'=\{\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'\}$, 依次以 E' 中各向量作为坐标轴 Ox',Oy',Oz' 正方向上的单位向量建立新的空间直角坐标系 $Ox'y'z'$, 求出在这个新坐标系下 $\sigma:\boldsymbol{\xi}\mapsto\boldsymbol{B\xi}$ 的矩阵 $\boldsymbol{B}$, 使其中 $\boldsymbol{\xi},\boldsymbol{B\xi}$ 分别是向量 $\boldsymbol{X},\sigma(\boldsymbol{X})$ 在新坐标系下的坐标.

先将 $\boldsymbol{X}_1$ 扩充为一组正交基 $\{\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3\}$. 易验证 $\boldsymbol{X}_2=(1,-1,0)^{\mathrm{T}}$ 与 $\boldsymbol{X}_3=(1,1,-2)^{\mathrm{T}}$ 符合要求. 分别取与 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 同方向的单位向量

$$\boldsymbol{e}_1'=\frac{1}{|\boldsymbol{X}_1|}\boldsymbol{X}_1=\begin{pmatrix}\dfrac{1}{\sqrt{3}}\\ \dfrac{1}{\sqrt{3}}\\ \dfrac{1}{\sqrt{3}}\end{pmatrix},\boldsymbol{e}_2'=\frac{1}{|\boldsymbol{X}_2|}\boldsymbol{X}_2=\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ -\dfrac{1}{\sqrt{2}}\\ 0\end{pmatrix},$$

$$\boldsymbol{e}_3'=\frac{1}{|\boldsymbol{X}_3|}\boldsymbol{X}_3=\begin{pmatrix}\dfrac{1}{\sqrt{6}}\\ \dfrac{1}{\sqrt{6}}\\ -\dfrac{2}{\sqrt{6}}\end{pmatrix},$$

则 $\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'$ 组成 V 的标准正交基. 依次以 $\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'$ 为各列组成的方阵

$$\boldsymbol{U}=\begin{pmatrix}\dfrac{1}{\sqrt{3}} & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{6}}\\ \dfrac{1}{\sqrt{3}} & -\dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{6}}\\ \dfrac{1}{\sqrt{3}} & 0 & -\dfrac{2}{\sqrt{6}}\end{pmatrix}$$

是正交方阵, 且 $\det\boldsymbol{U}=1$, 因而 $\boldsymbol{U}$ 的各列组成右手系标准正交基. 分别以它们为 Ox',Oy',Oz' 轴正方向上的单位向量建立新的坐标系. 则每一点 P 的原坐标 $\boldsymbol{X}=(x,y,z)^{\mathrm{T}}$ 与新坐标 $\boldsymbol{\xi}=(x',y',z')^{\mathrm{T}}$ 之间有坐标变换公式 $\boldsymbol{X}=\boldsymbol{U\xi}$. $\sigma(P)$ 的原坐标 $\boldsymbol{AX}$ 与新坐标 $\boldsymbol{\eta}$ 之间也有同样的关系式 $\boldsymbol{AX}=\boldsymbol{U\eta}$, 将 $\boldsymbol{X}=\boldsymbol{U\xi}$ 代入得到 $\boldsymbol{AU\xi}=\boldsymbol{U\eta}$, 于是 $\boldsymbol{\eta}=(\boldsymbol{U}^{-1}\boldsymbol{AU})\boldsymbol{\xi}$. 变换 $\sigma:P\mapsto\sigma(P)$

在新坐标系下成为 $\sigma:\boldsymbol{\xi}\mapsto\boldsymbol{\eta}=(\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U})\boldsymbol{\xi}$, 由方阵 $\boldsymbol{B}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ 左乘 P 的新坐标 $\boldsymbol{\xi}$ 得到 $\sigma(P)$ 的新坐标.

$$\begin{aligned}\boldsymbol{B}=&\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}\\ =&\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}}\end{pmatrix}\begin{pmatrix}\frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}}\end{pmatrix}\\ =&\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2}\\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(-60^\circ) & -\sin(-60^\circ)\\ 0 & \sin(-60^\circ) & \cos(-60^\circ)\end{pmatrix}\end{aligned}$$

是变换 σ 在新坐标系下的矩阵. $\boldsymbol{B}$ 表示的变换是将所有的点绕 Ox' 轴旋转 -60°. 如果面对 Ox' 的正方向看, 就是沿顺时针方向旋转 60°. 在原坐标系中, 是沿 $X_1=(1,1,1)^{\mathrm{T}}$ 所在的直线旋转, 如果面对 $\boldsymbol{X}_1$ 的正方向看, 就是沿顺时针方向旋转 60°. □

由例 4 中的方阵 $\boldsymbol{A}$ 很难直接看出线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 的几何性质. 选取了一组右手系标准正交基 $E'=\{\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'\}$, 以 E' 中各个基向量作为坐标轴正方向上的单位向量建立新的直角坐标系 $Ox'y'z'$ 之后, σ 在新坐标系下具有形式 $\boldsymbol{\xi}\mapsto\boldsymbol{B}\boldsymbol{\xi}$, 具有新的矩阵 $\boldsymbol{B}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$. 容易看出 $\boldsymbol{B}$ 是绕 Ox' 轴的旋转变换的矩阵, 从而 σ 是旋转变换.

一般地, 如果 3 阶实正交方阵 $\boldsymbol{A}$ 的行列式 $\det\boldsymbol{A}=1$, 它在空间直角坐标系中引起的变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 就是旋转变换 (这一结论的证明见 §5.7.2 例 2), 特征值 1 的特征向量所在的直线就是旋转轴, 由 $\boldsymbol{A}$ 的虚特征值 $\cos\alpha\pm\mathrm{i}\sin\alpha$ 可以决定旋转角 α. 由此得到例 4 更简捷的解法:

例 4 解法 2 易验证 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$ 且 $\det\boldsymbol{A}=1$. 可知 σ 是绕轴旋转. 由 $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=(\lambda-1)(\lambda^2-\lambda+1)$ 解出实特征值 1, 虚特征值 $\frac{1}{2}\pm\frac{\sqrt{3}}{2}\mathrm{i}=\cos 60^\circ\pm\mathrm{i}\sin 60^\circ$, 可知旋转角大小为 60°. 求出特征值 1 的特征向量 $\boldsymbol{X}_1=(1,1,1)^{\mathrm{T}}$, 则 $\boldsymbol{X}_1$ 所在的过原点的直线 l 就是旋转轴. σ 将 Ox 正方向上的单位向量 $\boldsymbol{e}_1=(1,0,0)^{\mathrm{T}}$ 旋转到 $\boldsymbol{A}\boldsymbol{e}_1=(\frac{2}{3},-\frac{1}{3},\frac{2}{3})^{\mathrm{T}}$. 由 $\det(\boldsymbol{X}_1,\boldsymbol{e}_1,\boldsymbol{A}\boldsymbol{e}_1)=-1$ 知 $\boldsymbol{X}_1,\boldsymbol{e}_1,\sigma(\boldsymbol{e}_1)$ 成左手系, 面对 $\boldsymbol{X}_1$ 正方向看到的旋

转是沿顺时针方向旋转 60°. □

例 4 通过坐标变换将线性变换 σ 的方阵 $\boldsymbol{A}$ 变成了 $\boldsymbol{B}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$. 一般地, 有:

引理 5.6.3 设 $\sigma: \boldsymbol{X}\mapsto \boldsymbol{A}\boldsymbol{X}$ 是 $V=F^{n\times 1}$ 上的线性变换, $S=\{\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n\}$ 是 V 的任意一组基. 将每个向量 $\boldsymbol{X}$ 用它在 S 下的坐标 $\boldsymbol{\xi}$ 表示. 则 $\sigma(\boldsymbol{X})$ 在 S 下的坐标 $\boldsymbol{\eta}=\boldsymbol{B}\boldsymbol{\xi}$, 其中 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 与 $\boldsymbol{A}$ 相似, $\boldsymbol{P}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n)$ 是以 S 中各个基向量为各列排成的可逆方阵.

证明 $\boldsymbol{X}$ 与它在 S 下的坐标 $\boldsymbol{\xi}=(\xi_1,\cdots,\xi_n)^{\mathrm{T}}$ 之间有坐标变换关系式 $\boldsymbol{X}=\xi_1\boldsymbol{X}_1+\cdots+\xi_n\boldsymbol{X}_n=\boldsymbol{P}\boldsymbol{\xi}$. $\sigma(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}$ 与它的新坐标 $\boldsymbol{\eta}$ 有同样的关系 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{P}\boldsymbol{\eta}$. 将 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{\xi}$ 代入得 $\boldsymbol{A}\boldsymbol{P}\boldsymbol{\xi}=\boldsymbol{P}\boldsymbol{\eta}$, 从而 $\boldsymbol{\eta}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{\xi}=\boldsymbol{B}\boldsymbol{\xi}$ 对 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 成立. □

将引理 5.6.3 中每个向量 $\boldsymbol{X}\in F^{n\times 1}$ 用它在基 S 下的新坐标 $\boldsymbol{\xi}$ 表示, 则 σ 将新坐标为 $\boldsymbol{\xi}$ 的向量映到新坐标为 $\boldsymbol{B}\boldsymbol{\xi}$ 的向量. σ 在基 S 下具有形式 $\sigma:\boldsymbol{\xi}\mapsto\boldsymbol{B}\boldsymbol{\xi}$, $\boldsymbol{B}$ 称为 σ 在基 S 下的矩阵. 由于 $\boldsymbol{X}$ 本身是它在 $F^{n\times 1}$ 的自然基 $E=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 下的坐标, $\boldsymbol{A}$ 就是 σ 在自然基 E 下的矩阵. 同一个变换 σ 在两组不同基 E 与 S 下的矩阵 $\boldsymbol{A},\boldsymbol{B}$ 相似: $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 适当选择 S, 可以使 σ 的矩阵 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 具有简单的形式, 将 σ 的性质刻画得更清楚. 特别地, 如果 $\boldsymbol{P}$ 的各列都是 $\boldsymbol{A}$ 的特征向量, 则 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 是对角矩阵, σ 将各个基向量映到各自的常数倍.

如果 $V=\mathbf{R}^{2\times 1}$ 或 $\mathbf{R}^{3\times 1}$ 是几何平面, $\boldsymbol{U}$ 是行列式为 1 的正交方阵, 则 $\boldsymbol{U}$ 的各列组成 V 的标准正交基 S, 依次以 S 中的各向量为各坐标轴正方向上的单位向量可以建立新的直角坐标系, 每个点 M 在新坐标系下的坐标就等于 $\overrightarrow{OM}$ 在新基 S 下的坐标.

例 5 试分析实方阵

$$\boldsymbol{A}=\begin{pmatrix}\frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\end{pmatrix}$$

在三维几何空间 $V=\mathbf{R}^{3\times 1}$ 中引起的变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 的几何性质.

解 求得 $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-1)^2(\lambda+1)$. 特征值为 1(2 重),$-1$.

$\boldsymbol{A}$ 是实对称方阵, 存在行列式等于 1 的 3 阶正交方阵 $\boldsymbol{U}$ 将 $\boldsymbol{A}$ 相似于对角矩阵 $\boldsymbol{D}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\mathrm{diag}(1,1,-1)$. $\boldsymbol{U}$ 的各列 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3$ 组成右手

系标准正交基 S. 依次以 $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3$ 为坐标轴 Ox', Oy', Oz' 正方向上的单位向量建立新的直角坐标系 $Ox'y'z'$, 则 σ 在新坐标系下的矩阵为对角矩阵 $\boldsymbol{D} = \mathrm{diag}(1,1,-1)$. σ 将 $x'Oy'$ 平面内的向量都保持不变, 将 Oz' 轴上的向量 $\boldsymbol{X}$ 都变到 $-\boldsymbol{X}$. σ 是关于 $Ox'y'$ 平面的对称变换.

$x'Oy'$ 平面也就是原坐标系中 $\boldsymbol{A}$ 的属于特征值 1 的特征子空间 V_1. 解方程组 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 可求得 $\boldsymbol{X}_1 = (1,-1,0)^{\mathrm{T}}, \boldsymbol{X}_2 = (0,2,1)^{\mathrm{T}}$ 组成解空间 V_1 的基. 过原坐标系中原点 O 及坐标为 $(1,-1,0),(0,2,1)$ 的点作平面 π, σ 就是关于平面 π 的对称变换. □

例 5 中的方阵 $\boldsymbol{A}$ 是实对称方阵且特征值为 $1,1,-1$, 根据实对称方阵正交相似于对角矩阵的定理, 可以断定存在行列式为 1 的正交方阵 $\boldsymbol{U}$ 将 $\boldsymbol{A}$ 相似到 $\mathrm{diag}(1,1,-1)$, 从而断定 σ 是关于平面的对称变换. 一般地, 并不需要 $\boldsymbol{A}$ 是实对称方阵, 只要它的特征值是 $1,1,-1$, 通过解方程组得到了属于特征值 1 与 -1 的特征子空间 V_1 与 V_{-1} 的维数分别是 2 与 1, 并且 V_1 与 V_{-1} 垂直, 仍能断定 σ 是关于平面的对称变换.

5.6.4 正交变换

例 4 与例 5 中的线性变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 分别是三维几何空间中绕轴的旋转变换和关于平面的对称变换, 保持图形的形状和大小不变.

其实, 不必知道 σ 是旋转或对称, 只要验证它们的方阵 $\boldsymbol{A}$ 是正交方阵, 就可以断定 σ 保持图形的形状和大小不变. 一般地, 只要变换 σ 保持任何两点的距离 AB 和任何一个角 $\angle BAC$ 的大小不变, 就可以断定 σ 保持图形的大小和形状不变. 而距离和角度可以通过向量内积算出来. 因此, 只需证明 σ 保持任何两个向量 $\boldsymbol{X}, \boldsymbol{Y}$ 的内积不变, 就知道 σ 保持距离和角度不变, 从而保持图形的形状和大小不变.

定义 5.6.1 设 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 是欧氏空间 V 上的线性变换. 如果 σ 保持向量内积不变, 即对任意 $\boldsymbol{X}, \boldsymbol{Y} \in V$ 有 $(\sigma(\boldsymbol{X}), \sigma(\boldsymbol{Y})) = (\boldsymbol{X}, \boldsymbol{Y})$, 就称 σ 是**正交变换** (orthogonal transformation).

引理 5.6.4 欧氏空间 $V = \mathbf{R}^{n\times 1}$ 上的线性变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 是正交变换的充分必要条件是: 变换方阵 $\boldsymbol{A}$ 是正交方阵.

证明 先证充分性, 设 $\boldsymbol{A}$ 是正交方阵, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$. 由 $(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$ 知

$$(\sigma(\boldsymbol{X}), \sigma(\boldsymbol{Y})) = (\boldsymbol{AX}, \boldsymbol{AY}) = (\boldsymbol{AX})^{\mathrm{T}}(\boldsymbol{AY}) = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{AY} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = (\boldsymbol{X}, \boldsymbol{Y}).$$

可见 σ 保内积, 是正交变换.

再证必要性, 设 σ 保内积. $\boldsymbol{A}$ 的第 j 列 $\boldsymbol{A}_j = \boldsymbol{A}\boldsymbol{e}_j = \sigma(\boldsymbol{e}_j)$ 是自然基向量 $\boldsymbol{e}_j$ 的像. 自然基 $E = \{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 是标准正交基, 它的像 $\sigma(E) =$

$\{\sigma(\boldsymbol{e}_1),\cdots,\sigma(\boldsymbol{e}_n)\}=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n\}$ 也是标准正交基, 排成的方阵 $\boldsymbol{A}=(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$ 的格拉姆方阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$, $\boldsymbol{A}$ 是正交方阵. □

推论 5.6.2 同阶正交方阵的乘积仍是正交方阵. 正交方阵的逆仍是正交方阵.

证法 1 设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶正交方阵, 则线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 与 $\tau:\boldsymbol{X}\mapsto\boldsymbol{B}\boldsymbol{X}$ 保内积. 因而复合变换 $\sigma\tau:\boldsymbol{X}\mapsto(\boldsymbol{A}\boldsymbol{B})\boldsymbol{X}$ 保内积, 逆变换 $\sigma^{-1}:\boldsymbol{X}\mapsto\boldsymbol{A}^{-1}\boldsymbol{X}$ 保内积, $\sigma\tau$ 与 σ^{-1} 的矩阵 $\boldsymbol{A}\boldsymbol{B}$ 与 $\boldsymbol{A}^{-1}$ 是正交方阵.

证法 2 $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{B})=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{I}\ \Rightarrow\ \boldsymbol{A}\boldsymbol{B}$ 是正交方阵.

$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$ 两边求逆得 $\boldsymbol{A}^{-1}(\boldsymbol{A}^{-1})^{\mathrm{T}}=\boldsymbol{I}$, 这说明 $\boldsymbol{A}^{-1}$ 是正交方阵. □

正交变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 保持向量的内积不变, 当然也保持每个向量 $\boldsymbol{X}$ 与自己的内积 $(\boldsymbol{X},\boldsymbol{X})=|\boldsymbol{X}|^2$ 不变, 也就是保持每个向量 $\boldsymbol{X}$ 的模 (即长度) $|\boldsymbol{X}|=\sqrt{(\boldsymbol{X},\boldsymbol{X})}$ 不变. 如果 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的实特征向量, 要使 $\sigma(\boldsymbol{X})=\lambda\boldsymbol{X}$ 与 $\boldsymbol{X}$ 的长度 $|\lambda\boldsymbol{X}|=|\lambda||\boldsymbol{X}|$ 与 $|\boldsymbol{X}|$ 相等, 当且仅当 $|\lambda|=1$ 即 $\lambda=\pm1$. 更进一步, 正交方阵的每个复特征值 (不一定是实数) λ 的模也等于 1.

引理 5.6.5 正交方阵 $\boldsymbol{A}$ 的特征值 λ 的模 $|\lambda|=1$.

证明 设 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 即

$$\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}. \tag{5.6.1}$$

等式两边取共轭转置得 $\boldsymbol{X}^*\boldsymbol{A}^*=\boldsymbol{X}^*\overline{\lambda}$. 左乘等式 (5.6.1) 两边得

$$\boldsymbol{X}^*(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{X}=\boldsymbol{X}^*(\overline{\lambda}\lambda)\boldsymbol{X}. \tag{5.6.2}$$

$\boldsymbol{A}$ 是实方阵, $\boldsymbol{A}^*=\boldsymbol{A}^{\mathrm{T}}$,从而 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$. 又 $\overline{\lambda}\lambda=|\lambda|^2$, $\boldsymbol{X}^*\boldsymbol{X}=\overline{x_1}x_1+\cdots+\overline{x_n}x_n=|x_1|^2+\cdots+|x_n|^2>0$. 代入 (5.6.2) 得 $\boldsymbol{X}^*\boldsymbol{X}=|\lambda|^2\boldsymbol{X}^*\boldsymbol{X}$, $|\lambda|^2=1$, $|\lambda|=1$. □

引理 5.6.5 实际上证明了: 满足条件 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{I}$ 的复方阵 $\boldsymbol{A}$ 的特征值 λ 的模 $|\lambda|=1$. 满足条件 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{I}$ 的复方阵 $\boldsymbol{A}$ 称为**酉方阵** (unitary matrix). 因此我们有:

引理 5.6.6 酉方阵的特征值的模等于 1.

元素全为实数的酉方阵就是正交方阵. 因此引理 5.6.5 是引理 5.6.6 的特殊情形.

习题 5.6

1. 设

$$S=\begin{pmatrix}1&-2&0\\-2&2&-2\\0&-2&3\end{pmatrix},$$

求正交矩阵 $\boldsymbol{U}$ 使 $\boldsymbol{U}^{-1}\boldsymbol{S}\boldsymbol{U}$ 是对角矩阵.

2. 用正交方阵化下列二次型为标准形:

(1) $Q(x_1,x_2,x_3)=2x_1^2+x_2^2-4x_1x_2-4x_2x_3$;

(2) $Q(x_1,x_2,x_3)=3x_1^2+4x_1x_2+8x_1x_3+4x_2x_3+3x_3^2$;

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

3. 在建立了直角坐标系的 3 维几何空间中, 如下的方程的图像是什么?

(1) $x^2+2y^2+3z^2-4xy-4yz=10$;

(2) $x^2+y^2+z^2-xy-xz-yz=6$;

(3) $xy+yz+zx=5$;

(4) $xy+yz+zx=0$.

4. 在建立了直角坐标系的 3 维几何空间中, 证明方程

$$2x^2+4y^2+5z^2-2xy+4xz+6yz-20=0$$

的图像是椭球面, 并求出这个椭球面所围成的立体的体积.

5. 如下方阵在几何空间 $\mathbf{R}^{3\times 1}$ 中引起的变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 是否是旋转变换或对称变换? 如果是旋转变换, 求旋转轴及旋转角. 如果是关于某直线 L 或平面 π 的对称变换, 求直线 L 或平面 π.

(1) $\boldsymbol{A}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$; (2) $\boldsymbol{A}=\begin{pmatrix}-1&0&0\\0&0&-1\\0&-1&0\end{pmatrix}$.

§5.7* 更多的例子

5.7.1 变换的迭代——利用计算机求特征向量

例 1 平面上建立了直角坐标系. 用 $\mathbf{R}^{2\times 1}$ 中每个向量 $\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}$ 代表平面上坐标为 (x,y) 的点. 取 2 阶实方阵

$$\boldsymbol{A}=\begin{pmatrix}1.1&0.4\\0.2&0.9\end{pmatrix},$$

则 $\mathbf{R}^{2\times 1}$ 的线性变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{AX}$ 代表了平面上的点的变换.

在正方形区域 $\{(x,y)\mid -1\leqslant x\leqslant 1,-1\leqslant y\leqslant 1\}$ 内随机地产生 200 个点 P_i $(1\leqslant i\leqslant 200)$. 用 σ 作用产生 200 个新的点 $\sigma(P_i)$, 再作用再产生 200 个新的点 $\sigma^2(P_i)$. 重复作用 n 次, 产生 $200(n+1)$ 个点 $\sigma^k(P_i)$ $(1\leqslant i\leqslant 200, 0\leqslant k\leqslant n)$. 将它们全部画出来观察变化趋势, 图 5–7 是 $n=10$ 时的所有点.

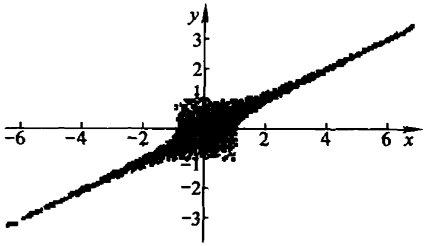

图 5–7

观察发现, 大量的点趋向于一条直线 l . 试求这条直线的斜率 k_l .

分析 设 $\boldsymbol{X}_0=(x_0,y_0)^{\mathrm{T}}\neq\boldsymbol{0}$ 是随机产生的非零向量. 通过观察图 5–7 可以认为: 当 $n\to\infty$ 时向量 $\boldsymbol{X}_n=\sigma^n(\boldsymbol{X}_0)=\boldsymbol{A}^n\boldsymbol{X}_0=\begin{pmatrix}x_n\\y_n\end{pmatrix}$ 趋于与 l 平行的方向,

$$k_l=\lim_{n\to\infty}\frac{y_n}{x_n}.$$

$\boldsymbol{A}$ 的特征向量 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 组成 $\mathbf{R}^{2\times 1}$ 的一组基. 利用 $\boldsymbol{A}$ 对 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 的左乘作用可以知道 $\boldsymbol{A}^n$ 对 $\mathbf{R}^{2\times 1}$ 中任意向量 $a_1\boldsymbol{X}_1+a_2\boldsymbol{X}_2$ 的作用.

解 先求 $\boldsymbol{A}$ 的特征值与特征向量.

$\boldsymbol{A}$ 的特征多项式

$$\varphi_{\boldsymbol{A}}(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1.1 & -0.4\\ -0.2 & \lambda-0.9\end{vmatrix}=\lambda^2-2\lambda+0.91$$

的两根为 $\lambda_1=0.7$ 与 $\lambda_2=1.3$.

当 $\lambda=0.7$, 求得方程组 $(\boldsymbol{A}-0.7\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的基础解 $\boldsymbol{X}_1=(1,-1)^{\mathrm{T}}$. 当 $\lambda=1.3$ 时, 求得方程组 $(\boldsymbol{A}-1.3\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的基础解 $\boldsymbol{X}_2=(2,1)^{\mathrm{T}}$. $\boldsymbol{X}_1,\boldsymbol{X}_2$ 线性无关, 组成 $\mathbf{R}^2$ 的一组基, $\mathbf{R}^2$ 中的任何向量 $\boldsymbol{X}_0$ 都能写成 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 的线性组合 $\boldsymbol{X}_0=a_1\boldsymbol{X}_1+a_2\boldsymbol{X}_2$. 于是

$$\begin{cases}\boldsymbol{A}\boldsymbol{X}_1=\lambda_1\boldsymbol{X}_1,\\ \boldsymbol{A}\boldsymbol{X}_2=\lambda_2\boldsymbol{X}_2\end{cases}\Rightarrow\begin{cases}\boldsymbol{A}^n\boldsymbol{X}_1=\lambda_1^n\boldsymbol{X}_1,\\ \boldsymbol{A}^n\boldsymbol{X}_2=\lambda_2^n\boldsymbol{X}_2\end{cases}\Rightarrow\quad \boldsymbol{A}^n\boldsymbol{X}_0=a_1\lambda_1^n\boldsymbol{X}_1+a_2\lambda_2^n\boldsymbol{X}_2.$$

$\boldsymbol{X}_n=\boldsymbol{A}^n\boldsymbol{X}_0$ 与 $\boldsymbol{\xi}_n=(\lambda_2^n)^{-1}\boldsymbol{A}^n\boldsymbol{X}_0$ 的方向相同. 当 $a_2\neq 0$ 时,

$$\lim_{n\to\infty}\boldsymbol{\xi}_n=\lim_{n\to\infty}\left(a_1\left(\frac{0.7}{1.3}\right)^n\boldsymbol{X}_1+a_2\boldsymbol{X}_2\right)=a_2\boldsymbol{X}_2.$$

当 $a_2=0\neq a_1$ 时, $\boldsymbol{X}_0=a_1\boldsymbol{X}_1$ 是 $\boldsymbol{A}$ 的属于特征值 $\lambda_1=0.7$ 的特征向量, 无论被 $\boldsymbol{A}$ 作用多少次, 仍保持在原来的 $\boldsymbol{X}_1$ 的方向上. 然而, 随机产生的向量 $\boldsymbol{X}_0$ 正好与 $\boldsymbol{X}_1$ 共线的概率很小, 绝大多数情形下 $\boldsymbol{X}_0$ 都满足条件 $a_2\neq 0$, 当 n 无限增大时 $\boldsymbol{A}^n\boldsymbol{X}_0$ 的方向趋于与 $\boldsymbol{X}_2$ 平行, 无限接近于过原点且与 $\boldsymbol{X}_2$ 平行的直线 l. l 的斜率 k_l 等于 $\boldsymbol{X}_2=(2,1)$ 的斜率 $\dfrac{1}{2}$. □

求方阵 $\boldsymbol{A}$ 的特征向量要先解一元 n 次方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 以求出特征根. 解一元 n 次方程一般是很困难的, 这就给求特征向量带来困难.

例 1 的结果为求 $\boldsymbol{A}$ 的特征向量提供了一种方法: 从任何一个非零向量 $\boldsymbol{X}_0$ 出发 (例如取 $\boldsymbol{X}_0=(1,\cdots,1)^{\mathrm{T}}$), 不断地用 $\boldsymbol{A}$ 左乘, 得到 $\boldsymbol{X}_n=\boldsymbol{A}^n\boldsymbol{X}_0$ $(n=1,2,\cdots)$, 并且用与 $\boldsymbol{X}_n$ 方向相同的某个 $\boldsymbol{Y}_n$ 代替 (例如, 将 $\boldsymbol{X}_n$ 除以它的各分量绝对值之和得到 $\boldsymbol{Y}_n$, 则 $\boldsymbol{Y}_n$ 的各分量绝对值之和为 1.) 当 n 无限增大时, 如果 $\boldsymbol{Y}_n$ 趋于一个极限位置 $\tilde{\boldsymbol{Y}}$, 则 $\tilde{\boldsymbol{Y}}$ 就是 $\boldsymbol{A}$ 的特征向量 (所属特征值通常是 $\boldsymbol{A}$ 的所有特征值中绝对值最大的.)

例如:

$$\boldsymbol{A}=\begin{pmatrix}1&0.5&0.4&0.3\\2&1&0.7&0.6\\2.5&1.4&2&1.3\\0.3&0.4&0.5&1\end{pmatrix},$$

从 $\boldsymbol{X}_0=(1,1,1,1)^{\mathrm{T}}$ 开始, 将 $\boldsymbol{X}_1=\boldsymbol{A}\boldsymbol{X}_0$ 除以各分量之和得到 $\boldsymbol{Y}_1$, $\boldsymbol{X}_2=\boldsymbol{A}\boldsymbol{Y}_1$ 除以各分量之和得到 $\boldsymbol{Y}_2$. 一般地, 如果已经得到 $\boldsymbol{Y}_k$, 将 $\boldsymbol{A}\boldsymbol{Y}_k$ 除以各分量之和得到 $\boldsymbol{Y}_{k+1}$. 这样就得到一系列的列向量 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n,\cdots$, 直到某个 $|\boldsymbol{Y}_{m+1}-\boldsymbol{Y}_m|$ 小到可以忽略不计的程度, 可以近似地认为 $\boldsymbol{Y}_{m+1}=\boldsymbol{Y}_m$, $\boldsymbol{A}\boldsymbol{Y}_m=\lambda\boldsymbol{Y}_m$, 则 $\boldsymbol{Y}_m$ 是特征向量, 所属的特征值 λ 等于 $\boldsymbol{A}\boldsymbol{Y}_m$ 各分量之和. 这个过程很容易用计算机实现. 具体计算出的各 $\boldsymbol{Y}_k$ 如下 (为节省篇幅写成行向量形式):

(0.138365,0.27044,0.45283,0.138365), (0.131698,0.251377,0.480387,0.136538),
(0.130463,0.248149,0.483933,0.137455),(0.130229,0.247579,0.484386,0.137806),
(0.130184,0.247475,0.484443,0.137899),(0.130175,0.247456,0.484449,0.13792),
(0.130173,0.247452,0.48445,0.137924), (0.130173,0.247451,0.48445,0.137925),
(0.130173,0.247451,0.48445,0.137926), (0.130173,0.247451,0.48445,0.137926),

其中 $\boldsymbol{Y}_9$ 与 $\boldsymbol{Y}_{10}$ 已经看不出差别, 可以认为是特征向量. $\boldsymbol{A}\boldsymbol{Y}_9=\lambda\boldsymbol{Y}_9$, 其中 $\lambda=3.75697$ 是 $\boldsymbol{A}\boldsymbol{Y}_9$ 的各分量和, 就是 $\boldsymbol{Y}_9$ 所属于的特征值.

这个方法求出的是绝对值最大的特征值的特征向量.

在科学计算中，这个方法叫做幂方法. 什么样的方阵可以用这个方法来求特征向量，是值得探讨的理论问题. 不过，只要计算机做出来的结果收敛，$\boldsymbol{Y}_n$ 与 $\boldsymbol{Y}_{n+1}$ 的差别趋于零，就求出了一个特征向量.

5.7.2 几何空间中的旋转

例 2 (1) 行列式为 1 的二阶正交方阵 $\boldsymbol{A}$ 在几何平面 $\mathbf{R}^{2\times 1}$ 上引起的变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 是绕原点的旋转变换.

(2) 行列式为 1 的三阶正交方阵 $\boldsymbol{A}$ 在几何空间 $\mathbf{R}^{3\times 1}$ 中引起的变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 是绕某条过原点的直线的旋转变换.

解 (1) $\boldsymbol{A}$ 的两列 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 组成标准正交基，且由 $\det \boldsymbol{A} = 1$ 知 $\boldsymbol{A}_1$ 沿逆时针方向旋转直角得到 $\boldsymbol{A}_2$. 设从 Ox 轴旋转到 $\boldsymbol{A}_1$ 的方向所成的角为 α，则 Ox 旋转到 $\boldsymbol{A}_2$ 的方向所成的角为 $\alpha + \dfrac{\pi}{2}$. 由于 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 都是单位向量，得:

$$\boldsymbol{A}_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \boldsymbol{A}_2 = \begin{pmatrix} \cos\left(\alpha + \dfrac{\pi}{2}\right) \\ \sin\left(\alpha + \dfrac{\pi}{2}\right) \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}, \boldsymbol{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

$\boldsymbol{A}$ 左乘引起的线性变换 $\sigma: \boldsymbol{X} \mapsto \boldsymbol{AX}$ 是绕原点旋转角 α 的旋转变换.

(2) $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda) = |\lambda \boldsymbol{I} - \boldsymbol{A}|$ 的三个复数根的模都等于 1，其中的实根为 ± 1. 且三个根的乘积等于 $\boldsymbol{A}$ 的行列式 1.

如果三个根都是实数，不可能三个都等于 -1 (否则三个根的乘积为 $(-1)^3 = -1 \neq 1$)，至少有一个实根 $\lambda_1 = 1$.

三次实系数多项式 $\varphi_{\boldsymbol{A}}(\lambda)$ 如果有虚根，必然是两个相互共轭的虚根 $\lambda_2, \overline{\lambda_2}$，另一个根 λ_1 是实根. 两个虚根的乘积 $\lambda_2\overline{\lambda_2} = |\lambda_2|^2 = 1$. 代入 $\lambda_1\lambda_2\overline{\lambda_2} = 1$ 得 $\lambda_1 = 1$.

因此，所有情况下 $\boldsymbol{A}$ 都有一个特征值为 1，存在实特征向量 $\boldsymbol{X}_1$ 满足条件 $\boldsymbol{AX}_1 = \boldsymbol{X}_1$，$\boldsymbol{X}_1$ 的全体实数倍组成的直线 L 上所有的点在 σ 作用下都固定不动. 我们证明 σ 是绕轴 L 的旋转.

取与 $\boldsymbol{X}_1$ 同方向上的单位向量 $\boldsymbol{e}_1' = \dfrac{1}{|\boldsymbol{X}_1|}\boldsymbol{X}_1$ 扩充为 $\mathbf{R}^{3\times 1}$ 的右手系标准正交基 $E' = \{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\}$. 以 E' 中的基向量为各列排成的方阵 $\boldsymbol{U} = (\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3')$ 是正交方阵且 $\det \boldsymbol{U} = 1$. σ 在基 E' 下的矩阵 $\boldsymbol{B} = \boldsymbol{U}^{-1}\boldsymbol{AU}$ 仍是正交方阵，且 $\boldsymbol{B}$ 的每列 $\boldsymbol{B}_j$ $(1 \leqslant j \leqslant 3)$ 是 $\sigma(\boldsymbol{e}_j')$ 在基 E' 下的坐标. 特别地，$\sigma(\boldsymbol{e}_1') = \boldsymbol{e}_1'$ 在基 E' 下的坐标 $\boldsymbol{B}_1 = (1,0,0)^{\mathrm{T}}$. 由 $\boldsymbol{B}$ 是正交方阵知 $\boldsymbol{B}$ 的各列两两正交，$\boldsymbol{B}_1 = (1,0,0)^{\mathrm{T}}$ 与另外两列 $\boldsymbol{B}_j = (b_{1j}, b_{2j}, b_{3j})^{\mathrm{T}}$ $(j = 2,3)$ 的内积

$(\boldsymbol{B}_1,\boldsymbol{B}_j)=b_{1j}=0$. 于是 $\boldsymbol{B}_j=(0,b_{2j},b_{3j})^{\mathrm{T}}$ 对 $j=2,3$ 成立.

$$\boldsymbol{B}=\begin{pmatrix}1&0&0\\0&b_{22}&b_{23}\\0&b_{32}&b_{33}\end{pmatrix}=\begin{pmatrix}1&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{B}_{22}\end{pmatrix},$$

其中 $\boldsymbol{B}_{22}$ 是二阶实方阵, 行列式 $|\boldsymbol{B}_{22}|=|\boldsymbol{B}|=1$. 且由

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\begin{pmatrix}1&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{B}_{22}^{\mathrm{T}}\boldsymbol{B}_{22}\end{pmatrix}=\boldsymbol{I}\Rightarrow\ \boldsymbol{B}_{22}^{\mathrm{T}}\boldsymbol{B}_{22}=\boldsymbol{I}$$

知 $\boldsymbol{B}_{22}$ 是二阶正交方阵. 由 (1) 的结果知 $\boldsymbol{B}_{22}$ 具有形状

$$\boldsymbol{B}_{22}=\begin{pmatrix}\cos\alpha&-\sin\alpha\\\sin\alpha&\cos\alpha\end{pmatrix}.$$

依次以 $\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'$ 为坐标轴正方向上的单位向量建立新的直角坐标系 $Ox'y'z'$, 则 $\boldsymbol{B}$ 是绕 Ox' 轴旋转变换的矩阵. σ 是绕 Ox' 轴 (即直线 L) 旋转角 α 的变换. □

例 3 写出空间中绕轴 L 旋转角 α 的变换 σ 的矩阵 $\boldsymbol{A}$, 旋转轴 L 由 z 轴在 yOz 平面内绕原点往 y 轴正方向旋转角 ω 得到.

解 将原来的坐标轴绕 Ox 轴旋转角 $-\omega$, 得到新的坐标轴 Ox',Oy',Oz', 建立新的坐标系 $Ox'y'z'$. 则新的 Oz' 轴就是旋转变换 σ 的旋转轴 L. 因此, σ 就是绕 Oz' 轴旋转角 α 的变换, σ 在新坐标系下的矩阵

$$\boldsymbol{B}=\begin{pmatrix}\cos\alpha&-\sin\alpha&0\\\sin\alpha&\cos\alpha&0\\0&0&1\end{pmatrix}.$$

$\boldsymbol{B}$ 的 3 列 $\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_3$ 分别是 $\sigma(\boldsymbol{e}_1'),\sigma(\boldsymbol{e}_2'),\sigma(\boldsymbol{e}_3')$ 在基 $\{\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'\}$ 下的坐标, 其中 $\boldsymbol{e}_1',\boldsymbol{e}_2',\boldsymbol{e}_3'$ 依次是新坐标系的 3 个坐标轴正方向上的单位向量, 也就是绕 x 轴旋转角 $-\omega$ 的变换矩阵

$$\boldsymbol{U}=\begin{pmatrix}1&0&0\\0&\cos\omega&\sin\omega\\0&-\sin\omega&\cos\omega\end{pmatrix}$$

的三列. 我们有 $\boldsymbol{AU}=\boldsymbol{UB}$. 从而

$$
\begin{aligned}
\boldsymbol{A} &= \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{-1} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{U}^{\mathrm{T}}\\
&= \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\omega & \sin\omega\\ 0 & -\sin\omega & \cos\omega \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha & 0\\ \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\omega & -\sin\omega\\ 0 & \sin\omega & \cos\omega \end{pmatrix}\\
&= \begin{pmatrix} \cos\alpha & -\cos\omega\sin\alpha & \sin\omega\sin\alpha\\ \cos\omega\sin\alpha & \cos^2\omega\cos\alpha+\sin^2\omega & \sin\omega\cos\omega(1-\cos\alpha)\\ -\sin\omega\sin\alpha & \sin\omega\cos\omega(1-\cos\alpha) & \sin^2\omega\cos\alpha+\cos^2\omega \end{pmatrix}. \qquad \square
\end{aligned}
$$

例 4 在三维空间中, 先绕 x 轴旋转角 α, 再绕 z 轴旋转角 β. 这两个旋转的复合变换是否仍是绕某条轴的旋转? 如果是, 找出旋转轴.

解 两个旋转变换 $\sigma:\boldsymbol{X}\mapsto\boldsymbol{A}\boldsymbol{X}$ 与 $\tau:\boldsymbol{X}\mapsto\boldsymbol{B}\boldsymbol{X}$ 的复合变换 $\tau\sigma:\boldsymbol{X}\mapsto(\boldsymbol{B}\boldsymbol{A})\boldsymbol{X}$ 的矩阵

$$
\begin{aligned}
\boldsymbol{K}=\boldsymbol{B}\boldsymbol{A} &= \begin{pmatrix} \cos\beta & -\sin\beta & 0\\ \sin\beta & \cos\beta & 0\\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha\\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}\\
&= \begin{pmatrix} \cos\beta & -\sin\beta\cos\alpha & \sin\beta\sin\alpha\\ \sin\beta & \cos\beta\cos\alpha & -\cos\beta\sin\alpha\\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.
\end{aligned}
$$

由 $\boldsymbol{A},\boldsymbol{B}$ 是正交方阵知 $\boldsymbol{K}=\boldsymbol{B}\boldsymbol{A}$ 是正交方阵. 由行列式 $|\boldsymbol{A}|=|\boldsymbol{B}|=1$ 知 $|\boldsymbol{K}|=|\boldsymbol{B}||\boldsymbol{A}|=1$. 因此 $\tau\sigma:\boldsymbol{X}\mapsto\boldsymbol{K}\boldsymbol{X}$ 仍是绕某条过原点 O 的直线 L 的旋转. 转轴 L 上所有的点 P 在旋转变换 $\tau\sigma$ 下都不动, L 上所有的 $\overrightarrow{OP}\neq\mathbf{0}$ 都是属于 $\boldsymbol{K}$ 的特征值 1 的特征向量. 解方程组 $(\boldsymbol{K}-\boldsymbol{I})\boldsymbol{X}=\mathbf{0}$ 求出非零解 $\boldsymbol{X}$, 就得到旋转轴 $L=\{a\boldsymbol{X}\mid a\in\mathbf{R}\}$.

如果 $\cos\alpha=1$, 则 $\boldsymbol{A}=\boldsymbol{I}$, $\tau\sigma=\tau$ 的旋转轴是 z 轴. 如果 $\cos\beta=1$, 则 $\boldsymbol{B}=\boldsymbol{I}$, $\tau\sigma=\sigma$ 的旋转轴是 x 轴. 以下设 $\cos\alpha\neq1$ 且 $\cos\beta\neq1$. 对 $\boldsymbol{K}-\boldsymbol{I}$ 作初等行变换得

$$
\begin{aligned}
\boldsymbol{K}-\boldsymbol{I} &= \begin{pmatrix} \cos\beta-1 & -\sin\beta\cos\alpha & \sin\beta\sin\alpha\\ \sin\beta & \cos\beta\cos\alpha-1 & -\cos\beta\sin\alpha\\ 0 & \sin\alpha & \cos\alpha-1 \end{pmatrix}\\
&\xrightarrow{\frac{\sin\beta}{1-\cos\beta}(1)+(2)} \begin{pmatrix} \cos\beta-1 & -\sin\beta\cos\alpha & \sin\beta\sin\alpha\\ 0 & -1-\cos\alpha & \sin\alpha\\ 0 & \sin\alpha & \cos\alpha-1 \end{pmatrix}
\end{aligned}
$$

$$\xrightarrow{\frac{\sin\alpha}{1-\cos\alpha}(3)+(2),\frac{\sin\beta\sin\alpha}{1-\cos\alpha}(3)+(1)}\begin{pmatrix}\cos\beta-1 & \sin\beta & 0\\ 0 & 0 & 0\\ 0 & \sin\alpha & \cos\alpha-1\end{pmatrix},$$

求得齐次线性方程组 $(\boldsymbol{K}-\boldsymbol{I})\boldsymbol{X}=\mathbf{0}$ 的解空间为

$$\begin{aligned}L&=\left\{\lambda\left(\frac{\sin\beta}{1-\cos\beta},1,\frac{\sin\alpha}{1-\cos\alpha}\right)^{\mathrm{T}}\middle|\lambda\in\mathbf{R}\right\}\\&=\left\{\lambda\left(\cot\frac{\beta}{2},1,\cot\frac{\alpha}{2}\right)^{\mathrm{T}}\middle|\lambda\in\mathbf{R}\right\}.\end{aligned}$$

直线 L 就是所求的旋转轴. □

5.7.3 一阶常系数微分方程组的矩阵解法

微分方程在科学研究和工程技术中的应用非常广泛. 例如, 放射性物质的质量 x 的衰变速度 $\dfrac{\mathrm{d}x}{\mathrm{d}t}$ 与它的现存量 x 成正比, 在资源充沛的情况下生物数量 x 的增长速度 $\dfrac{\mathrm{d}x}{\mathrm{d}t}$ 在不长的时期内与现有数量 x 成正比, 函数 $x=f(t)$ 都满足条件

$$\frac{\mathrm{d}x}{\mathrm{d}t}=ax, \tag{5.7.1}$$

其中 $a\neq 0$ 是常数. 条件 (5.7.1) 是含有微分运算的方程, 称为微分方程. 微分方程是数学研究中一个专门的研究方向, 也是一门单独的数学课程. 但只要学过基本的求导公式, 就知道指数函数 $x=\mathrm{e}^{at}$ 满足条件

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{at})=a\mathrm{e}^{at}, \tag{5.7.2}$$

$x=\mathrm{e}^{at}$ 是微分方程 (5.7.1) 的解.

要求出微分方程 (5.7.1) 的全部解, 考虑这个微分方程的任一解 $x=x(t)$ 与已求得的特解 e^{at} 之比 $c(t)=x(t)\mathrm{e}^{-at}$. 对 $c(t)$ 求导得

$$\frac{\mathrm{d}}{\mathrm{d}t}c(t)=x'(t)\mathrm{e}^{-at}+x(t)(\mathrm{e}^{-at})'=ax(t)\mathrm{e}^{-at}-ax(t)\mathrm{e}^{-at}=0. \tag{5.7.3}$$

可见 $c(t)$ 是常数 c, 因此 $x(t)=c\,\mathrm{e}^{at}$ 是微分方程 (5.7.1) 的全部解, 也就是它的通解.

如果同一个自变量 t 的若干个函数 $x_i = f_i(t)$ 满足条件

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = a_{11}x_1 + \cdots + a_{1n}x_n, \\ \qquad\cdots\cdots\cdots\cdots \\ \dfrac{\mathrm{d}x_n}{\mathrm{d}t} = a_{n1}x_1 + \cdots + a_{nn}x_n, \end{cases} \tag{5.7.4}$$

条件 (5.7.4) 就是一个微分方程组, 可以写成矩阵形式

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X}, \tag{5.7.1$'$}$$

其中

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

微分方程组 (5.7.1$'$) 与微分方程 (5.7.1) 在形式上看起来类似, 是否也有类似形式的通解 $\boldsymbol{X} = \mathrm{e}^{\boldsymbol{A}t}C$? 其中 $\mathrm{e}^{\boldsymbol{A}t}$ 可以在 e^{at} 的泰勒展开式

$$\mathrm{e}^{at} = 1 + at + \frac{1}{2!}(at)^2 + \cdots + \frac{1}{k!}(at)^k + \cdots \tag{5.7.5}$$

中将实数 a 替换成方阵 $\boldsymbol{A}$ 得到:

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{I} + \boldsymbol{A}t + \frac{1}{2!}(\boldsymbol{A}t)^2 + \cdots + \frac{1}{k!}(\boldsymbol{A}t)^k + \cdots \tag{5.7.5$'$}$$

虽然指数函数的求导公式 (5.7.2) 得来并不容易, 难以推广到指数为矩阵的情况. 但既然 (5.7.5) 的右边是指数函数的泰勒展开式, 对它求导可以得到与 (5.7.2) 同样的结论 (读者试自己验证). 对 (5.7.5) 等号右边求导只涉及 t 的非负整数次幂 t^k 的导数公式, 将 (5.7.5) 右边的 a 换成方阵 $\boldsymbol{A}$ 得到 (5.7.5$'$) 之后, 同样可以对右边求导得到与 (5.7.2) 类似的结论

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{\boldsymbol{A}t}) = \boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}. \tag{5.7.2$'$}$$

还可以对 (5.7.1$'$) 的通解 $\boldsymbol{X}$ 求 $\boldsymbol{C}(t) = \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{X}$ 对 t 的导数, 与 (5.7.3) 类似地得到 $\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{C}(t) = \boldsymbol{0}$, 从而 $\boldsymbol{C}(t) = \boldsymbol{C} = \mathrm{e}^{-\boldsymbol{A}t}\boldsymbol{X}$ 是由实常数组成的 n 维列向量, $\boldsymbol{X} = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{C}$ 是 (5.7.1$'$) 的通解.

展开式 (5.7.5$'$) 右边是无穷级数, 实际计算 $\mathrm{e}^{\boldsymbol{A}t}$ 时还需要研究它的收敛问题. 如果 $\boldsymbol{A}$ 的某个正整数次幂等于零: $\boldsymbol{A}^m = \boldsymbol{O}$, 这样的方阵 $\boldsymbol{A}$ 称为**幂零的** (nilpotent). 将幂零方阵 $\boldsymbol{A}$ 代入 (5.7.5$'$), 右边的无穷级数就变成有限的多项

式, 不存在收敛问题. 另外一种容易处理的情形是: $\boldsymbol{A}$ 是对角矩阵. 将对角矩阵 $\boldsymbol{A}=\mathrm{diag}(a_1,\cdots,a_n)$ 代入 (5.7.5′) 得到

$$\mathrm{e}^{\boldsymbol{A}t}=\begin{pmatrix} 1+\cdots+\frac{1}{k!}(a_1t)^k+\cdots & & \\ & \ddots & \\ & & 1+\cdots+\frac{1}{k!}(a_nt)^k+\cdots \end{pmatrix}$$

$$=\begin{pmatrix} \mathrm{e}^{a_1t} & & \\ & \ddots & \\ & & \mathrm{e}^{a_nt} \end{pmatrix}.$$

由泰勒展开式还可得出指数函数的性质 $\mathrm{e}^{at}\mathrm{e}^{bt}=\mathrm{e}^{at+bt}$, 推导过程用到乘法交换律 $ab=ba$. 虽然矩阵乘法不满足交换律, 但如果某两个同阶实方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{BA}$, 同样能得到等式 $\mathrm{e}^{\boldsymbol{A}t}\mathrm{e}^{\boldsymbol{B}t}=\mathrm{e}^{\boldsymbol{A}t+\boldsymbol{B}t}$ (如果 $\boldsymbol{AB}\neq\boldsymbol{BA}$, 这个等式就可能不成立) . 如果能将 $\boldsymbol{A}$ 分解为 $\boldsymbol{A}=\boldsymbol{A}_1+\boldsymbol{A}_2$ 使 $\boldsymbol{A}_1\boldsymbol{A}_2=\boldsymbol{A}_2\boldsymbol{A}_1$, $\boldsymbol{A}_1$ 是对角矩阵且 $\boldsymbol{A}_2$ 是幂零矩阵, 则 $\mathrm{e}^{\boldsymbol{A}t}=\mathrm{e}^{\boldsymbol{A}_1t}\mathrm{e}^{\boldsymbol{A}_2t}$ 容易计算. 即使 $\boldsymbol{A}_1$ 不是对角矩阵, 而只是相似于对角矩阵 $\boldsymbol{D}=\boldsymbol{P}^{-1}\boldsymbol{A}_1\boldsymbol{P}$, 仍可以通过 $f(\boldsymbol{A}_1t)=\boldsymbol{P}f(\boldsymbol{D}t)\boldsymbol{P}^{-1}$ 将 $f(\boldsymbol{A}_1t)=\mathrm{e}^{\boldsymbol{A}_1t}$ 转化为 $f(\boldsymbol{D}t)$ 来计算.

例 5 求函数 $x=f(t)$, 使其满足微分方程 $x''=-x$.

解 令 $y=x'=f'(t)$, 则微分方程变成方程组

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=y,\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=-x,\end{cases}\quad 即\quad \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}=\boldsymbol{AX},\quad 其中\quad \boldsymbol{A}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

通解为 $\boldsymbol{X}=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{C}$. 注意到 $\boldsymbol{A}^2=-\boldsymbol{I}$, 我们有

$$\begin{aligned}\mathrm{e}^{\boldsymbol{A}t}&=\boldsymbol{I}+\boldsymbol{A}t+\frac{1}{2!}\boldsymbol{A}^2t^2+\cdots+\frac{1}{k!}\boldsymbol{A}^kt^k+\cdots\\&=(1-\frac{t^2}{2!}+\frac{t^4}{4!}-\cdots)\boldsymbol{I}+(t-\frac{t^3}{3!}+\frac{t^5}{5!}-\cdots)\boldsymbol{A}\\&=(\cos t)\boldsymbol{I}+(\sin t)\boldsymbol{A}=\begin{pmatrix}\cos t&\sin t\\-\sin t&\cos t\end{pmatrix}.\end{aligned}$$

微分方程组 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}=\boldsymbol{AX}$ 的通解为

$$\boldsymbol{X}=\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{C}=\begin{pmatrix}\cos t&\sin t\\-\sin t&\cos t\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix}=\begin{pmatrix}c_1\cos t+c_2\sin t\\-c_1\sin t+c_2\cos t\end{pmatrix}.$$

原微分方程的通解为 $x = c_1\cos t + c_2\sin t$. □

例 5 中的方阵 $\boldsymbol{A}$ 满足条件 $\boldsymbol{A}^2 = -\boldsymbol{I}$, 指数函数 $\mathrm{e}^{\boldsymbol{A}t}$ 正好是平面上表示旋转角 $-t$ 的矩阵. 我们已经熟悉的等式

$$\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}^n = \begin{pmatrix} \cos nt & \sin nt \\ -\sin nt & \cos nt \end{pmatrix},$$

就是

$$(\mathrm{e}^{\boldsymbol{A}t})^n = \mathrm{e}^{n\boldsymbol{A}t},$$

这是指数函数理所当然的性质.

如果在指数函数 e^{at} 的泰勒展开式 (5.7.5) 中将 a 换成具有性质 $\mathrm{i}^2 = -1$ 的虚数单位 i, 类似地可以得到

$$\mathrm{e}^{t\mathrm{i}} = \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \cdots\right) + \mathrm{i}\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots\right) = \cos t + \mathrm{i}\sin t,$$

这恰是表示旋转角 t 的复数. 我们已经熟悉的棣莫弗公式

$$(\cos t + \mathrm{i}\sin t)^n = \cos nt + \mathrm{i}\sin nt$$

就是指数函数的理所当然的性质 $(\mathrm{e}^{t\mathrm{i}})^n = \mathrm{e}^{nt\mathrm{i}}$.

例 6 求同一自变量 t 的函数 $x = f(t)$ 与 $y = g(t)$, 使其满足条件

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = 1.1x + 0.4y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = 0.2x + 0.9y. \end{cases} \tag{5.7.6}$$

解 微分方程组 (5.7.6) 可以写成矩阵形式 $\dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X} = \boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 1.1 & 0.4 \\ 0.2 & 0.9 \end{pmatrix}.$$

通解为 $\boldsymbol{X} = \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{C}$, 其中 $\boldsymbol{C} = (c_1, c_2)^{\mathrm{T}}$ 是任意实常数组成的 2 维列向量. 只需求出 $\mathrm{e}^{\boldsymbol{A}t}$.

求得特征多项式 $\varphi_{\boldsymbol{A}}(\lambda) = |\lambda\boldsymbol{I} - \boldsymbol{A}| = \lambda^2 - 2\lambda + 0.91$ 的两个根为 $\lambda_1 = 0.7, \lambda_2 = 1.3$. 解方程组 $(\boldsymbol{A} - 0.7\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 与 $(\boldsymbol{A} - 1.3\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 分别求得基础解 $\boldsymbol{X}_1 = (1, -1)^{\mathrm{T}}$, $\boldsymbol{X}_2 = (2, 1)^{\mathrm{T}}$, 也就是分别属于特征值 0.7, 1.3 的特征向量. 以 $\boldsymbol{X}_1, \boldsymbol{X}_2$ 为两列排成可逆方阵 $\boldsymbol{P} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$, 则

$$\boldsymbol{AP} = \boldsymbol{A}(\boldsymbol{X}_1, \boldsymbol{X}_2) = (\boldsymbol{X}_1, \boldsymbol{X}_2)\boldsymbol{D} = \boldsymbol{PD}, \qquad \text{其中} \quad \boldsymbol{D} = \begin{pmatrix} 0.7 & 0 \\ 0 & 1.3 \end{pmatrix}.$$

由 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$ 得

$$\mathrm{e}^{\boldsymbol{A}t}=\boldsymbol{P}\mathrm{e}^{\boldsymbol{D}t}\boldsymbol{P}^{-1}=\begin{pmatrix}1 & 2\\ -1 & 1\end{pmatrix}\begin{pmatrix}\mathrm{e}^{0.7t} & 0\\ 0 & \mathrm{e}^{1.3t}\end{pmatrix}\begin{pmatrix}\frac{1}{3} & -\frac{2}{3}\\ \frac{1}{3} & \frac{1}{3}\end{pmatrix}$$

$$=\begin{pmatrix}\frac{1}{3}\mathrm{e}^{0.7t}+\frac{2}{3}\mathrm{e}^{1.3t} & -\frac{2}{3}\mathrm{e}^{0.7t}+\frac{2}{3}\mathrm{e}^{1.3t}\\ -\frac{1}{3}\mathrm{e}^{0.7t}+\frac{1}{3}\mathrm{e}^{1.3t} & \frac{2}{3}\mathrm{e}^{0.7t}+\frac{1}{3}\mathrm{e}^{1.3t}\end{pmatrix}.$$

(5.7.6) 的通解为 $\boldsymbol{X}=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{C}$, 即

$$\begin{cases}x=c_1\left(\frac{1}{3}\mathrm{e}^{0.7t}+\frac{2}{3}\mathrm{e}^{1.3t}\right)+c_2\left(-\frac{2}{3}\mathrm{e}^{0.7t}+\frac{2}{3}\mathrm{e}^{1.3t}\right),\\ y=c_1\left(-\frac{1}{3}\mathrm{e}^{0.7t}+\frac{1}{3}\mathrm{e}^{1.3t}\right)+c_2\left(\frac{2}{3}\mathrm{e}^{0.7t}+\frac{1}{3}\mathrm{e}^{1.3t}\right),\end{cases}$$

也就是

$$\begin{cases}x=(b_1-2b_2)\mathrm{e}^{0.7t}+2(b_1+b_2)\mathrm{e}^{1.3t},\\ y=(-b_1+2b_2)\mathrm{e}^{0.7t}+(b_1+b_2)\mathrm{e}^{1.3t},\end{cases}$$

其中 $b_1=\frac{1}{3}c_1$ 与 $b_2=\frac{1}{3}c_2$ 为任意实常数. □

例 7 求同一自变量 t 的两个函数 $x=f(t)$ 与 $y=g(t)$, 使其满足微分方程组

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=3x-y,\\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=x+y.\end{cases}$$

解 将原微分方程组写成矩阵形式 $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}=\boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{X}=\begin{pmatrix}x\\ y\end{pmatrix},\quad \boldsymbol{A}=\begin{pmatrix}3 & -1\\ 1 & 1\end{pmatrix}.$$

通解为 $\boldsymbol{X}=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{C}$, 其中 $\boldsymbol{C}=\begin{pmatrix}c_1\\ c_2\end{pmatrix}\in\mathbf{R}^{2\times1}$ 是由任意实常数组成的 2 维列向量.

求得 $\boldsymbol{A}$ 的特征多项式 $\varphi_{\boldsymbol{A}}(\lambda)=(\lambda-2)^2$. 令

$$\boldsymbol{N}=\boldsymbol{A}-2\boldsymbol{I}=\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix},$$

则易验证 $N^2=O$. $A=2I+N$ 且 $(2I)N=N(2I)$, 因此

$$e^{At}=e^{2tI+tN}=e^{2tI}e^{tN}=e^{2t}(I+Nt)=e^{2t}\begin{pmatrix}1+t & -t\\ t & 1-t\end{pmatrix}.$$

原微分方程组的通解为

$$X=\begin{pmatrix}x\\ y\end{pmatrix}=e^{tA}\begin{pmatrix}c_1\\ c_2\end{pmatrix}=\begin{pmatrix}c_1e^{2t}(1+t)-c_2e^{2t}t\\ c_1e^{2t}t+c_2e^{2t}(1-t)\end{pmatrix},$$

即

$$\begin{cases}x=c_1e^{2t}(1+t)-c_2e^{2t}t,\\ y=c_1e^{2t}t+c_2e^{2t}(1-t),\end{cases}$$

c_1,c_2 为任意常数. □

例 7 中 $N^2=O$ 对 $N=A-2I$ 成立并非偶然. 而是附录 7 定理 5.5.3(凯莱 -哈密顿定理) 的一个具体例子: 二阶方阵 A 如果不相似于对角矩阵, 必然只有一个 2 重特征值 λ_1, 特征多项式 $\varphi_A(\lambda)=(\lambda-\lambda_1)^2$, 由凯莱 -哈密顿定理就有 $(A-\lambda_1I)^2=O$, 就可以与例 7 同样得到 $A=\lambda_1I+N$ 使 $N^2=O$, 从而算出 $e^{tA}=e^{\lambda_1t}(I+tN)$.

更一般地, 任意 n 阶方阵 A 总可以分解为 $A=D+N$ 使 $DN=ND$, D 相似于对角矩阵, N 幂零, 仍可由 $e^{At}=e^{Dt}e^{Nt}$ 求出 e^{At}. 关于 A 的分解方法, 参见 §5.8.

例 8 求 e^{Bt}, 其中

$$B=\begin{pmatrix}3 & 2 & 1\\ 0 & -1 & -2\\ 0 & 0 & -1\end{pmatrix}.$$

解 §5.5 例 7 中已经求出可逆方阵 P 将 B 相似到

$$P^{-1}BP=J=\begin{pmatrix}3 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}=D+N,$$

其中

$$P=\begin{pmatrix}1 & 1 & 0\\ 0 & -2 & 0\\ 0 & 0 & 1\end{pmatrix},\quad D=\begin{pmatrix}3 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{pmatrix},\ N=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}.$$

于是

$$\begin{aligned}e^{Bt} &= Pe^{Jt}P^{-1} = P(e^{Dt}e^{Nt})P^{-1} = Pe^{Dt}(I+tN)P^{-1}\\ &= P\begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^{-t} & te^{-t} \\ 0 & 0 & e^{-t}\end{pmatrix}P^{-1} = \begin{pmatrix} e^{3t} & \frac{1}{2}(e^{3t}-e^{-t}) & te^{-t} \\ 0 & e^{-t} & -2te^{-t} \\ 0 & 0 & e^{-t}\end{pmatrix}. \quad \square\end{aligned}$$

习题 5.7

1. 平面上两条直线 l_1, l_2 的夹角为 α, σ_1, σ_2 分别是关于直线 l_1, l_2 的轴对称变换. 以 l_1, l_2 的交点 O 为原点建立适当的直角坐标系, 写出 σ_1, σ_2 及它们的复合变换 $\sigma_2\sigma_1$ 与 $\sigma_1\sigma_2$ 的矩阵. 复合变换 $\sigma_2\sigma_1, \sigma_1\sigma_2$ 是否相同, 是什么样的变换?

2. 已知方阵

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

(1) 求 $\boldsymbol{\Sigma}$ 的特征值与特征向量. 并求可逆方阵 P 使其将 $\boldsymbol{\Sigma}$ 相似到对角矩阵;

(2) 求 $\boldsymbol{\Sigma}^2, \boldsymbol{\Sigma}^3, \boldsymbol{\Sigma}^4$. 求证: 方阵

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ a_4 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_1 \end{pmatrix}$$

可以写成 $\boldsymbol{\Sigma}$ 的多项式;

(3) 求可逆方阵使其将 (2) 中的方阵 A 相似于对角矩阵. 并求出 A 的行列式.

3. 空间中建立了直角坐标系, 旋转变换 σ 将 Ox 轴变到 Oy 轴, Oy 轴变到 Oz 轴. 求 σ 的旋转轴和旋转角度.

4. 空间中建立了直角坐标系. σ 是关于 xOy 平面的对称变换. τ 是关于平面 $x=y$ 的对称变换. 求证: $\sigma\tau$ 与 $\tau\sigma$ 都是旋转变换, 并分别求出它们的旋转轴与旋转角度.

5. 求下列微分方程组的通解:

(1) $\begin{cases} \dfrac{dx}{dt} = 3x + y, \\ \dfrac{dy}{dt} = x + 3y; \end{cases}$ (2) $\begin{cases} \dfrac{dx}{dt} = 3x - 2y, \\ \dfrac{dy}{dt} = 5x - 3y. \end{cases}$

§5.8 * 若尔当标准形

在 §5.5 — §5.7 的例题中我们初步看到了将方阵 $\boldsymbol{A}$ 相似于尽可能简单的形状 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 带来的好处. 例如, 对幂函数 $f(x)=x^k$ 和指数函数 $f(x)=\mathrm{e}^x$ 计算 $f(\boldsymbol{A})$ 比较困难, 通过 $f(\boldsymbol{A})=\boldsymbol{P}f(\boldsymbol{B})\boldsymbol{P}^{-1}$ 转化为 $f(\boldsymbol{B})$ 来计算就更容易些. 特别地, 如果能将 $\boldsymbol{A}$ 相似到对角矩阵 $\boldsymbol{B}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, 则

$$f(\boldsymbol{B})=\mathrm{diag}(f(\lambda_1),\cdots,f(\lambda_n))$$

容易计算.

并非所有的方阵都能相似于对角矩阵. §5.5 例 7 中的方阵 $\boldsymbol{B}$ 不能相似于对角矩阵, 将它相似到另一种简单的方阵

$$\boldsymbol{J}=\begin{pmatrix}3&0&0\\0&-1&1\\0&0&-1\end{pmatrix}, \tag{5.8.1}$$

通过 $\boldsymbol{J}^n$ 算出了 $\boldsymbol{B}^n$, §5.7 中再次对同样的 $\boldsymbol{J}$ 算出了 $\mathrm{e}^{\boldsymbol{J}t}$, 从而算出了 $\mathrm{e}^{\boldsymbol{B}t}$.

(5.8.1) 中的方阵 $\boldsymbol{J}$ 是由对角块 $\boldsymbol{J}_1=3$ 与 $\boldsymbol{J}_2=\begin{pmatrix}-1&1\\0&-1\end{pmatrix}$ 组成的准对角矩阵 $\mathrm{diag}(\boldsymbol{J}_1,\boldsymbol{J}_2)$. 一般地, 我们有:

定义 5.8.1 具有如下形式的方阵

$$\boldsymbol{J}_m(a)=\begin{pmatrix}a&1&&\\&a&\ddots&\\&&\ddots&1\\&&&a\end{pmatrix}_{m\times m}=a\boldsymbol{I}_{(m)}+\boldsymbol{J}_m(0)$$

称为**若尔当块** (Jordan block), 其中 $\boldsymbol{J}_m(0)=(c_{ij})_{m\times m}$ 中所有的 $c_{i,i+1}=1(1\leqslant i\leqslant m-1)$, 其余元素都为 0. 由有限个若尔当块组成的准对角矩阵

$$\boldsymbol{J}=\mathrm{diag}(\boldsymbol{J}_{r_1}(\lambda_1),\cdots,\boldsymbol{J}_{r_p}(\lambda_p))$$

称为**若尔当矩阵** (Jordan matrix).

定理 5.8.1 每个复方阵 $\boldsymbol{A}$ 都能被复可逆方阵 $\boldsymbol{P}$ 相似于若尔当形矩阵 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 称为 $\boldsymbol{A}$ 的**若尔当标准形** (Jordan canonical form).

定理 5.8.1 的证明超出非数学类专业线性代数及工科高等代数课程的教学要求. 以下仅通过几个例子介绍求任意复方阵 $\boldsymbol{A}$ 的若尔当标准形 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$

以及可逆方阵 $\boldsymbol{P}$ 的算法, 为在实际应用和继续学习中需要这样的算法的读者提供参考和帮助.

例 1 设 m 阶可逆方阵 $\boldsymbol{P}$ 将方阵 $\boldsymbol{A}$ 相似于若尔当块 $\boldsymbol{J}_m(a)=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 试分析 $\boldsymbol{A}$ 和 $\boldsymbol{P}$ 应满足的条件.

解 $\boldsymbol{A}$ 与 $\boldsymbol{J}_m(a)$ 具有同样的特征多项式 $(\lambda-a)^m$, 同样地只有一个 m 重特征值 a.

$\boldsymbol{J}_m(0)=\boldsymbol{J}_m(a)-a\boldsymbol{I}=\boldsymbol{P}^{-1}(\boldsymbol{A}-a\boldsymbol{I})\boldsymbol{P}$ 与 $\boldsymbol{A}-a\boldsymbol{I}$ 相似. 由 $\boldsymbol{J}_m(0)^m=\boldsymbol{O}\neq\boldsymbol{J}_m(0)^{m-1}$ 知 $(\boldsymbol{A}-a\boldsymbol{I})^m=\boldsymbol{O}\neq(\boldsymbol{A}-a\boldsymbol{I})^{m-1}$.

设 $\boldsymbol{P}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)$ 的各列依次为 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m$. 则由 $(\boldsymbol{A}-a\boldsymbol{I})\boldsymbol{P}=\boldsymbol{P}\boldsymbol{J}_m(0)$ 知

$$(\boldsymbol{A}-a\boldsymbol{I})(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_m)=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_m)\boldsymbol{J}_m(0)=(\boldsymbol{0},\boldsymbol{X}_1,\cdots,\boldsymbol{X}_{m-1}).$$

$\boldsymbol{A}-a\boldsymbol{I}$ 对 $\boldsymbol{P}$ 的各列 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m$ 的左乘作用可用箭头图表示为

$$\boldsymbol{0}\leftarrow\boldsymbol{X}_1\leftarrow\boldsymbol{X}_2\leftarrow\cdots\leftarrow\boldsymbol{X}_{m-1}\leftarrow\boldsymbol{X}_m \tag{5.8.2}$$

反过来, 只要求出了满足箭头图 (5.8.2) 的 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m$ 且 $\boldsymbol{X}_1\neq\boldsymbol{0}$, 则 $\boldsymbol{P}=(\boldsymbol{X}_1,\cdots,\boldsymbol{X}_m)$ 的各列线性无关 (证明见例 2), $\boldsymbol{P}$ 是可逆方阵, 且将 $\boldsymbol{A}$ 相似到若尔当块 $\boldsymbol{J}_m(a)$.

由此得到: m 阶方阵 $\boldsymbol{A}$ 相似于若尔当块 $\boldsymbol{J}_m(a)$ $\Leftrightarrow$ $(\boldsymbol{A}-a\boldsymbol{I})^m=\boldsymbol{O}\neq(\boldsymbol{A}-a\boldsymbol{I})^{m-1}$.

如果此条件满足, 必可求出 $\boldsymbol{X}_m$ 使 $\boldsymbol{X}_1=(\boldsymbol{A}-a\boldsymbol{I})^{m-1}\boldsymbol{X}_m\neq\boldsymbol{0}$. 对 $1\leqslant i\leqslant m-1$ 取 $\boldsymbol{X}_i=(\boldsymbol{A}-a\boldsymbol{I})^{m-i}\boldsymbol{X}_m$, 则各 $\boldsymbol{X}_i$ 满足箭头图 (5.8.2), 排成可逆方阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 相似于若尔当块 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}_m(a)$. $\boldsymbol{J}_m(a)$ 就是 $\boldsymbol{A}$ 的若尔当标准形. □

例 2 求证: 如果向量组 $S=\{\boldsymbol{X}_{ij}\mid 1\leqslant i\leqslant d,1\leqslant j\leqslant m_i\}$ 在某个方阵 $\boldsymbol{B}$ 的左乘作用下满足箭头图

$$\begin{array}{c}\boldsymbol{0}\leftarrow\boldsymbol{X}_{11}\leftarrow\boldsymbol{X}_{12}\leftarrow\cdots\cdots\leftarrow\boldsymbol{X}_{1m_1}\\ \cdots\cdots\cdots\cdots\\ \boldsymbol{0}\leftarrow\boldsymbol{X}_{d1}\leftarrow\boldsymbol{X}_{d2}\leftarrow\cdots\leftarrow\boldsymbol{X}_{dm_d}\end{array} \tag{5.8.3}$$

且图中除了 $\boldsymbol{0}$ 之外的第一列向量 $\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{d1}$ 线性无关. 则 S 线性无关.

证明 设 S 的线性组合

$$\sum_{i=1}^{d}\sum_{j=1}^{m_i}\mu_{ij}\boldsymbol{X}_{ij}=\boldsymbol{0}, \tag{5.8.4}$$

我们证明其中的系数 μ_{ij} 必须全部为 0.

若不然, 设系数 μ_{ij} 不全为 0, 必有最大的正整数 k 使某个 $\mu_{tk} \neq 0$, 而当 $j > k$ 时所有的 $\mu_{ij} = 0$. 用 $\boldsymbol{B}^{k-1}$ 左乘等式 (5.8.4) 两边, 将所有的 $\mu_{ij}\boldsymbol{X}_{ij}$ $(j < k)$ 全部变成 $\mathbf{0}$, 每个 $\mu_{ik}\boldsymbol{X}_{ik}$ 变成 $\mu_{ik}\boldsymbol{X}_{i1}$, 得到

$$\mu_{1k}\boldsymbol{X}_{11} + \cdots + \mu_{tk}\boldsymbol{X}_{t1} + \cdots = \mathbf{0},$$

其中 $\mu_{tk} \neq 0$, 与 $\boldsymbol{X}_{11}, \cdots, \boldsymbol{X}_{d1}$ 线性无关矛盾.

这证明了 (5.8.3) 中所有的 $\mu_{ij} = 0$, S 线性无关. □

特别地, 当例 1 的箭头图 (5.8.2) 中的 $\boldsymbol{X}_1 \neq \mathbf{0}$ 时, $S = \{\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m\}$ 就是例 2 中当 $d = 1$ 的情形, 由例 2 证明的结论知道 $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_m$ 线性无关.

例 3 求可逆方阵 $\boldsymbol{P}$,使其将

$$\boldsymbol{A} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 \\ 2 & 3 & 5 & 0 \\ 1 & 2 & 3 & 5 \end{pmatrix}$$

相似到若尔当标准形 $\boldsymbol{J} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

解 $\boldsymbol{A}$ 只有一个 4 重特征值 5. 计算

$$\boldsymbol{A} - 5\boldsymbol{I} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 1 & 2 & 3 & 0 \end{pmatrix}$$

的各次幂得

$$(\boldsymbol{A} - 5\boldsymbol{I})^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 \\ 12 & 9 & 0 & 0 \end{pmatrix}, \quad (\boldsymbol{A} - 5\boldsymbol{I})^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 27 & 0 & 0 & 0 \end{pmatrix}.$$

由 $(\boldsymbol{A} - 5\boldsymbol{I})^4 = \boldsymbol{O} \neq (\boldsymbol{A} - 5\boldsymbol{I})^3$ 知 $\boldsymbol{A}$ 相似于若尔当块 $J_4(5)$.

取 $\boldsymbol{X}_4 = (1, 0, 0, 0)^{\mathrm{T}}$ 满足条件 $\boldsymbol{X}_1 = (\boldsymbol{A} - 5\boldsymbol{I})^3\boldsymbol{X}_4 = (0, 0, 0, 27)^{\mathrm{T}} \neq \mathbf{0}$. 算出

$$\boldsymbol{X}_3 = (\boldsymbol{A} - 5\boldsymbol{I})\boldsymbol{X}_4 = (0, 3, 2, 1)^{\mathrm{T}}, \quad \boldsymbol{X}_2 = (\boldsymbol{A} - 5\boldsymbol{I})^2\boldsymbol{X}_4 = (0, 0, 9, 12)^{\mathrm{T}}.$$

则各 $\boldsymbol{X}_i$ $(1 \leqslant i \leqslant 4)$ 在 $\boldsymbol{A} - 5\boldsymbol{I}$ 作用下满足箭头图

$$\mathbf{0} \leftarrow \boldsymbol{X}_1 \leftarrow \boldsymbol{X}_2 \leftarrow \boldsymbol{X}_3 \leftarrow \boldsymbol{X}_4$$

依次以 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3,\boldsymbol{X}_4$ 为各列排成可逆方阵

$$\boldsymbol{P}=\begin{pmatrix}0&0&0&1\\0&0&3&0\\0&9&2&0\\27&12&1&0\end{pmatrix},\quad 则\quad \boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}=\begin{pmatrix}5&1&0&0\\0&5&1&0\\0&0&5&1\\0&0&0&5\end{pmatrix}.\qquad\square$$

一般地, 设复方阵 $\boldsymbol{A}$ 被可逆方阵 $\boldsymbol{P}$ 相似到若尔当标准形

$$\boldsymbol{J}=\mathrm{diag}(\boldsymbol{J}_{m_1}(\lambda_1),\cdots,\boldsymbol{J}_{m_d}(\lambda_1),\boldsymbol{J}_{m_{d+1}}(\lambda_2),\cdots),$$

且不妨假定其中属于同一个特征值 λ_i 的若尔当块排在一起, 前 d 个若尔当块 $\boldsymbol{J}_{m_i}(\lambda_1)$ 属于特征值 λ_1, 以后的若尔当块属于其他的特征值 λ_j. 则 $\boldsymbol{P}$ 的前 $m_1+\cdots+m_d$ 列在 $\boldsymbol{A}-\lambda_1\boldsymbol{I}$ 的左乘作用下满足例 2 中的箭头图 (5.8.3), 其中每行对应于一个若尔当块, 各行所含非零向量的个数 $m_1,\cdots,m_d$ 就是属于特征值 λ_1 的各若尔当块 $\boldsymbol{J}_{m_i}(\lambda_1)$ 的阶数. 由箭头图 (5.8.3) 可以看出 (有兴趣的读者可以自己作出严格的论证): 其中第 1 个非零列的 d 个向量 $\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{d1}$ 被 $\boldsymbol{A}-\lambda_1\boldsymbol{I}$ 左乘为 $\boldsymbol{0}$, 组成 $(\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_1 的基, $d=\dim W_1$. 一般地, 箭头图 (5.8.3) 的前 k 个非零列组成 $(\boldsymbol{A}-\lambda_1\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_k 的基, 解空间的维数 $d_k=\dim W_k$ 就是箭头图 (5.8.3) 的前 k 个非零列所含向量的个数.

反过来, 只要对 $\boldsymbol{A}$ 的每个特征值 λ_i 求出了各方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ $(k=1,2,\cdots)$ 的解空间 W_k 的维数 d_k, 就可画出箭头图 (5.8.3) 使图中非零向量组成的第 1 列、前 2 列、 $\cdots$、前 k 列、 $\cdots$ 分别含 $d_1,d_2,\cdots,d_k,\cdots,d_p$ 个向量, 直到 $d_p=m=d_{p+1}$ (即第 $p+1$ 不含向量) 为止, 其中 m 是特征值 λ_i 的代数重数. $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_1 就是属于特征值 λ_i 的特征子空间 V_{λ_i}, 它的维数 d_1 就是对角元为 λ_i 的若尔当块的个数. 根据箭头图中各行所含非零向量个数 m_j 得到相应的若尔当块 $\boldsymbol{J}_{m_j}(\lambda_i)$ 的阶数. 对各个特征值 λ_i 都按这样的方法求出若尔当块, 就得到 $\boldsymbol{A}$ 的若尔当标准形 $\boldsymbol{J}$.

不需解方程组 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$, 只要算出系数矩阵 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})^k$ 的秩, 就能得到解空间 W_k 的维数 $d_k=n-\mathrm{rank}(\boldsymbol{A}-\lambda_i\boldsymbol{I})^k$. 根据这些维数就可画出箭头图, 得到属于特征值 λ_i 的各若尔当块.

要求出满足条件 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{AP}$ 的过渡矩阵 $\boldsymbol{P}$, 需要对每个特征值 λ_i 先求出箭头图 (5.8.3) 各行最后一个向量 $\boldsymbol{X}_{1m_1},\boldsymbol{X}_{2m_2},\cdots,\boldsymbol{X}_{dm_d}$, 使它们分别左乘 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})^{m_1-1},\cdots,(\boldsymbol{A}-\lambda_i\boldsymbol{I})^{m_d-1}$ 得到的 $\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{d1}$ 线性无关, 可以充当箭头图 (5.8.3) 中除了 $\boldsymbol{0}$ 之外的第一列各向量. 用 $\boldsymbol{A}-\lambda_i\boldsymbol{I}$ 的各次幂左乘各行最后的 $\boldsymbol{X}_{im_i}$ 就得到箭头图中所有的 $\boldsymbol{X}_{ij}$. 将各个特征值 λ_i 的箭头图中各行的向量从左到右依次排列, 就得到符合要求的可逆方阵 $\boldsymbol{P}$.

例 4 求方阵

$$A=\begin{pmatrix}3&0&0&0&1&0\\-2&3&1&1&2&-1\\-3&-1&5&5&-1&-1\\0&0&0&3&1&0\\0&0&0&-1&5&0\\0&0&0&2&-2&4\end{pmatrix}$$

的若尔当标准形 J. 并求可逆方阵 P 使 $P^{-1}AP=J$.

解 求得 A 的特征多项式 $|\lambda I-A|=(\lambda-3)(\lambda-4)^5$, 特征值为 3(1 重) 与 4(5 重).

属于特征值 3 的若尔当块阶数为 1, 只能等于 3. 解方程组 $(A-3I)X=0$ 求出非零解 $P_1=(1,1,2,0,0,0)^{\mathrm{T}}$ 作为 P 的第 1 列.

属于特征值 4 的若尔当块对应于 P 的其余 5 列. 计算出 $A-4I$ 与 $(A-4I)^2$ 分别等于

$$\begin{pmatrix}-1&0&0&0&1&0\\-2&-1&1&1&2&-1\\-3&-1&1&5&-1&-1\\0&0&0&-1&1&0\\0&0&0&-1&1&0\\0&0&0&2&-2&0\end{pmatrix},\quad\begin{pmatrix}1&0&0&-1&0&0\\1&0&0&-1&0&0\\2&0&0&-2&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}.$$

对 $A-4I$ 作初等行变换化成最简阶梯形矩阵

$$A-4I\rightarrow\begin{pmatrix}1&0&0&0&-1&0\\0&1&-1&0&-1&1\\0&0&0&1&-1&0\\&&&&&O\end{pmatrix},$$

求得 $\operatorname{rank}(A-4I)=3$, $(A-4I)X=0$ 的解空间 W_1 维数 $d_1=6-3=3$.

又由 $\operatorname{rank}(A-4I)^2=1$ 知 $(A-4I)^2X=0$ 的解空间 W_2 的维数 $d_2=5$, 已达到特征值 4 的代数重数.

由 $d_1=3,d_2=5$ 可以画出在 $A-4I$ 的左乘作用下的箭头图, 其中第一个非零列由 3 个向量 X_1,X_2,X_3 组成, 前两列由 5 个向量 $X_1,\cdots,X_5$ 组成.

$$\begin{array}{l}\mathbf{0}\leftarrow X_1\leftarrow X_4\\\mathbf{0}\leftarrow X_2\leftarrow X_5\\\mathbf{0}\leftarrow X_3\end{array}\tag{5.8.5}$$

箭头图共三行, 每行含非零向量个数分别为 2,2,1. 说明属于特征值 4 的若尔当块有三块, 阶数分别为 2,2,1, 分别为 $\boldsymbol{J}_2(4),\boldsymbol{J}_2(4),\boldsymbol{J}_1(4)$. $\boldsymbol{A}$ 的若尔当标准形为

$$\boldsymbol{J}=\begin{pmatrix}3&&&\\&\boldsymbol{J}_2(4)&&\\&&\boldsymbol{J}_2(4)&\\&&&4\end{pmatrix}=\begin{pmatrix}3&&&&&\\&4&1&&&\\&&4&&&\\&&&4&1&\\&&&&4&\\&&&&&4\end{pmatrix}.$$

为了求出满足条件 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$ 的可逆方阵 $\boldsymbol{P}$, 需要求出满足箭头图 (5.8.5) 的各向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_5$. 先求 $(\boldsymbol{A}-4\boldsymbol{I})^2\boldsymbol{X}=\boldsymbol{0}$ 的解 $\boldsymbol{X}_4,\boldsymbol{X}_5$ 使 $\boldsymbol{X}_1=(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}_4$ 与 $\boldsymbol{X}_2=(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}_5$ 线性无关. 取 $\boldsymbol{X}_4=(0,0,0,0,1,0)^{\mathrm{T}}$ 与 $\boldsymbol{X}_5=(0,0,0,0,0,1)^{\mathrm{T}}$, 则 $\boldsymbol{X}_1=(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}_4=(1,2,-1,1,1,-2)^{\mathrm{T}}$ 与 $\boldsymbol{X}_2=(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}_5=(0,-1,-1,0,0,0)^{\mathrm{T}}$ 线性无关, 符合要求.

$\boldsymbol{X}_1,\boldsymbol{X}_2$ 是 $(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的线性无关解, 需要再添加一个解 $\boldsymbol{X}_3$ 扩充为 $(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_1 的一组基. 为此, 解方程组求出 W_1 的一个基础解系 $\boldsymbol{Y}_1,\boldsymbol{Y}_2,\boldsymbol{Y}_3$, 再求 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y}_1,\boldsymbol{Y}_2,\boldsymbol{Y}_3$ 的极大线性无关组 $\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y}_i$, 取 $\boldsymbol{X}_3=\boldsymbol{Y}_i$ 即可. 计算得到 $\boldsymbol{X}_3=(1,1,0,1,1,0)^{\mathrm{T}}$.

依次以 $\boldsymbol{P}_1,\boldsymbol{X}_1,\boldsymbol{X}_4,\boldsymbol{X}_2,\boldsymbol{X}_5,\boldsymbol{X}_3$ 为各列排成可逆方阵

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

则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$. 恰如所需. □

如果只求例 5 的方阵 $\boldsymbol{A}$ 的若尔当标准形 $\boldsymbol{J}$ 而不求可逆方阵 $\boldsymbol{P}$, 则在画箭头图 (5.8.5) 时可以将向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_5$ 直接用 5 个数字 $1,2,\cdots,5$ 代替: 根据 $d_1=3$ 在第一列从上到下写 $1,2,3$, 再根据 $d_2=5$ 在第二列继续写 4,5 使前两列共有 5 个数字, 得到

$$\begin{bmatrix}1&4\\2&5\\3&\end{bmatrix},$$

由各行所含数字的个数 2,2,1 可知属于特征值 4 的若尔当块阶数分别为 2,2,1.

一般地, 我们有:

算法 5.8.1(求若尔当标准形及过渡矩阵) 已知 n 阶复方阵 $\boldsymbol{A}$ 及其全部不同的特征值, 求可逆方阵 $\boldsymbol{P}$ 使其将 $\boldsymbol{A}$ 相似到若尔当标准形 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

第 1 步 求若尔当标准形 $\boldsymbol{J}$.

对 $\boldsymbol{A}$ 的每个特征值 a 和每个正整数 $k=1,2,\cdots$, 求出各方程组 $(\boldsymbol{A}-a\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_k 的维数 $d_k=n-\operatorname{rank}(\boldsymbol{A}-a\boldsymbol{I})^k$, 得到 $d_1<d_2<\cdots<d_{p-1}<d_p=m$, 其中 m 是 a 的代数重数. 如果不知道 m, 则当 $d_{p-1}<d_p=d_{p+1}$ 时 $d_p=m$.

将前 m 个正整数 $1,2,\cdots,m$ 从左到右依次排成 p 列: 第 1 列从上到下依次填入 $1,2,\cdots,d_1$, 第 2 列继续依次从上到下填入后面的整数 $d_1+1,\cdots,d_2$. 对 $k=1,2,\cdots,p$, 第 k 列从上到下依次填入 $d_{k-1}+1,\cdots,d_k$. 得到维数图:

$$\begin{array}{cccc}
1 & d_1+1 & \cdots & d_{p-1}+1 \\
\vdots & \vdots & \cdots & \vdots \\
\vdots & \vdots & \cdots & d_p \\
\vdots & \vdots & \cdots & \\
\vdots & d_2 & & \\
d_1 & & &
\end{array} \tag{5.8.6}$$

维数图 (5.8.6) 共有 $d=d_1$ 行, 就是 $\boldsymbol{J}$ 中属于特征值 a 的若尔当块的个数. 各行所含整数的个数 $m_1,\cdots,m_d$ 就是属于特征值 a 的各若尔当块的阶数, 其中 $m_1=p$. 根据这些阶数 m_i 可以写出属于特征值 a 的各若尔当块 $\boldsymbol{J}_{m_1}(a),\cdots,\boldsymbol{J}_{m_d}(a)$.

将各个不同特征值的各若尔当块排成准对角矩阵, 就得到 $\boldsymbol{J}$.

第 2 步 求 $\boldsymbol{P}$ 的各列.

对 $\boldsymbol{A}$ 的每个特征值 a, 根据维数图 (5.8.6) 中各行所含数字的个数 (即各若尔当块的阶数) $m_1,\cdots,m_d$, 求满足相应的箭头图

$$\begin{array}{c}
\boldsymbol{0}\leftarrow\boldsymbol{X}_{11}\leftarrow\boldsymbol{X}_{12}\leftarrow\cdots\cdots\leftarrow\boldsymbol{X}_{1m_1} \\
\cdots\cdots\cdots\cdots \\
\boldsymbol{0}\leftarrow\boldsymbol{X}_{d1}\leftarrow\boldsymbol{X}_{d2}\leftarrow\cdots\leftarrow\boldsymbol{X}_{dm_d}
\end{array} \tag{5.8.7}$$

的各向量 $\boldsymbol{X}_{ij}$, 使其中第一个非零列的向量 $\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{d1}$ 线性无关. 依次将各特征值的箭头图中的各行的向量排成矩阵 $\boldsymbol{P}$, 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$ 符合要求.

要求出满足箭头图 (5.8.7) 的全部向量 $\boldsymbol{X}_{ij}$, 只要先求出各行最右边的向量 $\boldsymbol{X}_{1m_1},\cdots,\boldsymbol{X}_{dm_d}$, 使按箭头要求将它们分别左乘 $(\boldsymbol{A}-a\boldsymbol{I})^{m_1-1},\cdots,(\boldsymbol{A}-a\boldsymbol{I})^{m_d-1}$ 后得到的第 1 列的 $\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{d1}$ 线性无关, 则由 $\boldsymbol{X}_{1m_1},\cdots,\boldsymbol{X}_{dm_d}$ 左乘 $\boldsymbol{A}-a\boldsymbol{I}$ 的各次幂得到的各 $\boldsymbol{X}_{ij}$ 符合箭头图的要求.

要得到箭头图 (5.8.7) 各行最后右边的向量 $\boldsymbol{X}_{1m_1},\cdots,\boldsymbol{X}_{dm_d}$, 可以采用如下算法:

步骤 2.1 对每个正整数 k, 记 W_k 为 $(\boldsymbol{A}-a\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ 的解空间. 先解方程组 $(\boldsymbol{A}-a\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得解空间 W_1 的一组基 S_1, 添加 T_2 扩充为 W_2 的一组基 $S_2=S_1\cup T_2$. 一般地, 已求得解空间 W_k 的基 S_k 之后, 添加 T_{k+1} 扩充为解空间 W_{k+1} 的基 S_{k+1}. 最后得到 W_p 的基 $S_p=S_1\cup T_2\cup\cdots\cup T_p$ 使 $S_k=S_1\cup T_2\cup\cdots\cup T_k$ 是 W_k 的基 $(\forall 2\leqslant k\leqslant p)$. 我们设法从 S_p 中选出箭头图 (5.8.7) 中每行最后一个向量 $\boldsymbol{X}_{1m_1},\cdots,\boldsymbol{X}_{dm_d}$.

步骤 2.2 $T_k\subset W_k \Rightarrow (\boldsymbol{A}-a\boldsymbol{I})^kT_k=\boldsymbol{0}$, 因而 $(\boldsymbol{A}-a\boldsymbol{I})^{k-1}T_k\subset W_1$, 从而

$$M=(\boldsymbol{A}-a\boldsymbol{I})^{p-1}T_p\cup\cdots\cup(\boldsymbol{A}-a\boldsymbol{I})^{k-1}T_k\cup\cdots\cup(\boldsymbol{A}-a\boldsymbol{I})T_2\cup S_1\subset W_1.$$

求 M 的极大线性无关组

$$H=Q_p\cup Q_{p-1}\cup\cdots\cup Q_2\cup Q_1,$$

使每个 $H_k=Q_p\cup Q_{p-1}\cup\cdots\cup Q_k$ 是 $M_k=(\boldsymbol{A}-a\boldsymbol{I})^{p-1}T_p\cup\cdots\cup(\boldsymbol{A}-a\boldsymbol{I})^{k-1}T_k$ 的极大线性无关组. 当 $2\leqslant k\leqslant p$ 时, 每个 $Q_k\subseteq(\boldsymbol{A}-a\boldsymbol{I})^{k-1}T_k$, 因而 $Q_k=(\boldsymbol{A}-a\boldsymbol{I})^{k-1}T_k^*$ 由 T_k 的某个子集 T_k^* 左乘 $(\boldsymbol{A}-a\boldsymbol{I})^{k-1}$ 得到. 记 $T_1^*=Q_1\subseteq S_1$. 取

$$M^*=T_p^*\cup\cdots\cup T_k^*\cup\cdots\cup T_1^*$$

中各个向量作为箭头图 (5.8.7) 各行最后一个向量 $\boldsymbol{X}_{1m_1},\cdots,\boldsymbol{X}_{dm_d}$, 则 $H=(\boldsymbol{A}-a\boldsymbol{I})^{p-1}T_k^*\cup\cdots(\boldsymbol{A}-a\boldsymbol{I})T_2^*\cup T_1^*$ 中各向量就是各个 $\boldsymbol{X}_{im_i}$ 被 $(\boldsymbol{A}-a\boldsymbol{I})^{m_i-1}$ 的左乘得到的箭头图 (5.8.7) 中各行第一个非零向量 $\boldsymbol{X}_{i1}$. H 是 M 的极大线性无关组, 当然线性相关. 符合箭头图要求. □

注 以上算法第 2 步需要将各个解空间 W_{k-1} 的基 S_{k-1} 添加 T_k 得到 W_k 的基 S_k. 实际操作时, 可以解方程组 $(\boldsymbol{A}-a\boldsymbol{I})^{k-1}\boldsymbol{X}=\boldsymbol{0}$ 与 $(\boldsymbol{A}-a\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ 分别求出解空间的基 S_{k-1},S_k, 但不一定能使 $S_{k-1}\subseteq S_k$. 此时不一定要从 S_k 中选出 T_k 使 $S_{k-1}\cup T_k$ 正好是 W_k 的基, 只要取 S_k 的子集 T_k' 使 $S_{k-1}\cup T_k'$ 包含 W_k 的一组基就行了. 甚至可以直接取 S_k 作为 T_k'. 用 $M'=(\boldsymbol{A}-a\boldsymbol{I})^{p-1}T_p'\cup\cdots\cup(\boldsymbol{A}-a\boldsymbol{I})T_2'\cup S_1$ 的极大线性无关组 $H'=Q_p'\cup\cdots\cup Q_2'\cup Q_1'$ 代替 H, 自然就将各 T_k' 中多余的向量清除出去了.

例 5 求可逆方阵 $\boldsymbol{P}$ 将方阵

$$\boldsymbol{A}=\begin{pmatrix}1&1&1&1&1&1\\0&2&0&1&2&3\\0&0&2&0&1&-1\\0&0&0&2&1&2\\0&0&0&0&2&0\\0&0&0&0&0&2\end{pmatrix}$$

相似到若尔当标准形 $\boldsymbol{J}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

解 $\boldsymbol{A}$ 有两个不同的特征值 1,2, 代数重数分别为 1,5.

属于特征值 1 的若尔当块只有 1 个 1 阶块 $\boldsymbol{J}_1(1)=1$.

属于特征值 2 的若尔当块的阶数的和为 5. 计算可得:

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

$$(\boldsymbol{A}-2\boldsymbol{I})^3=\begin{pmatrix}-1&1&1&0&-2&-1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}.$$

由 $\boldsymbol{A}-2\boldsymbol{I}$ 的各次幂的秩分别求出方程组 $(\boldsymbol{A}-2\boldsymbol{I})^k\boldsymbol{X}=\boldsymbol{0}$ 的解空间 W_k 的维数 d_k, 得:

$d_1=6-\text{rank}(\boldsymbol{A}-2\boldsymbol{I})=2$, $d_2=6-\text{rank}(\boldsymbol{A}-2\boldsymbol{I})^2=4$, $d_3=6-\text{rank}(\boldsymbol{A}-2\boldsymbol{I})^3=5$.

根据这 3 个维数 2,4,5 对特征值 2 画出箭头图

$$\begin{matrix}\boldsymbol{0}\leftarrow\boldsymbol{X}_1\leftarrow\boldsymbol{X}_3\leftarrow\boldsymbol{X}_5\\ \boldsymbol{0}\leftarrow\boldsymbol{X}_2\leftarrow\boldsymbol{X}_4\end{matrix}\tag{5.8.8}$$

箭头图有两行, 所含非零向量个数分别为 3,2, 说明属于特征值 2 的若尔当块有两个, 分别为 $\boldsymbol{J}_3(2),\boldsymbol{J}_2(2)$. $\boldsymbol{A}$ 的若尔当标准形为

$$\boldsymbol{J}=\text{diag}(1,\boldsymbol{J}_3(2),\boldsymbol{J}_2(2)).$$

现在来求 $\boldsymbol{P}$ 的各列 $\boldsymbol{P}_1,\cdots,\boldsymbol{P}_6$.

对特征值 1, 解方程组 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 求得基础解 $(1,0,0,0,0,0)^{\mathrm{T}}$ 作为 $\boldsymbol{P}_1$.

为了对特征值 2 求出满足箭头图 (5.8.8) 的 5 个向量 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_5$, 只要求出其中两行的最后一个向量 $\boldsymbol{X}_5,\boldsymbol{X}_4$ 即可.

$\boldsymbol{X}_5$ 应是 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{X}=\boldsymbol{0}$ 的解且 $\boldsymbol{X}_1=(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X}_5\neq\boldsymbol{0}$. 求得 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{X}=\boldsymbol{0}$ 的解 $\boldsymbol{X}_5=(0,0,0,0,1,-2)^{\mathrm{T}}$ 满足 $\boldsymbol{X}_1=(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X}_5=$

$(-3,-3,0,0,0,0)^{\mathrm{T}} \neq \mathbf{0}$, 符合要求. 再按箭头图 (5.8.8) 的要求取 $\boldsymbol{X}_3 = (\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}_5 = (-1,-4,3,-3,0,0)^{\mathrm{T}}$.

$\boldsymbol{X}_4$ 应是 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X} = \mathbf{0}$ 的解, 且 $\boldsymbol{X}_2 = (\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}_4$ 与 $\boldsymbol{X}_1$ 线性无关. 求得 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X} = \mathbf{0}$ 的解 $\boldsymbol{X}_4 = (3,0,0,0,-2,1)^{\mathrm{T}}$, 且 $\boldsymbol{X}_2 = (\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}_4 = (-4,-1,-3,0,0,0)^{\mathrm{T}}$ 与 $\boldsymbol{X}_1$ 线性无关, 符合要求.

$$\boldsymbol{P} = (\boldsymbol{P}_1,\boldsymbol{X}_1,\boldsymbol{X}_3,\boldsymbol{X}_5,\boldsymbol{X}_2,\boldsymbol{X}_4) = \begin{pmatrix} 1 & -3 & -1 & 0 & -4 & 3 \\ 0 & -3 & -4 & 0 & -1 & 0 \\ 0 & 0 & 3 & 0 & -3 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & -2 & 0 & 1 \end{pmatrix}$$

满足条件 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J} = \mathrm{diag}(1,\boldsymbol{J}_3(2),\boldsymbol{J}_2(2))$. 恰如所需. □

注意例 5 中并没有按算法 5.8.1 所说步骤依次求出各方程组 $(\boldsymbol{A}-2\boldsymbol{I})^k\boldsymbol{X} = 0$ 的解空间 W_k 的基, 只对方程组 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{X} = \mathbf{0}$ 与 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X} = \mathbf{0}$ 各求出了一个解 $\boldsymbol{X}_5,\boldsymbol{X}_4$, 根本没有解方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X} = \mathbf{0}$. 这是因为, 箭头图 (5.8.8) 只有两行, 每行最后一个向量 $\boldsymbol{X}_5,\boldsymbol{X}_4$ 分别是 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{X} = \mathbf{0}$ 与 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{X} = \mathbf{0}$ 的解, 只要解这两个方程组求出满足条件的 $\boldsymbol{X}_5,\boldsymbol{X}_4$ 就行了. 如果按照算法 5.8.1 求出 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X} = \mathbf{0}$ 的基 S_1, 在求极大线性无关组之后 S_1 中的向量将被全部删去, 没有一个入选, 可见根本不需要解方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X} = \mathbf{0}$.

不妨将算法 5.8.1 的步骤与例 3, 例 4 的求解过程相比较. 例 3 的方阵 $\boldsymbol{A}$ 只有一个特征值, 特征子空间只有 1 维, 箭头图只有一行. 由于 $(\boldsymbol{A}-5\boldsymbol{I})^4 = \boldsymbol{O}$, 所有的向量都是 $(\boldsymbol{A}-5\boldsymbol{I})^4\boldsymbol{X} = \mathbf{0}$ 的解, 只要求一个 $\boldsymbol{X}_4$ 不是 $(\boldsymbol{A}-5\boldsymbol{I})^3\boldsymbol{X} = \mathbf{0}$ 的解, 就符合要求. 例 4 的方阵 $\boldsymbol{A}$ 的特征值 4 的箭头图有三行, 各行最后一个向量 $\boldsymbol{X}_4,\boldsymbol{X}_5,\boldsymbol{X}_3$ 需要解两个方程组 $(\boldsymbol{A}-4\boldsymbol{I})^2\boldsymbol{X} = \mathbf{0}$ 与 $(\boldsymbol{A}-4\boldsymbol{I})\boldsymbol{X} = \mathbf{0}$ 才能得到, $\{\boldsymbol{X}_4,\boldsymbol{X}_5\}$ 相当于算法 5.8.1 中的 T_2. 例 4 中确实按照算法 5.8.1 的步骤求得了 $M = (\boldsymbol{A}-4\boldsymbol{I})T_2 \cup S_1$ 的极大线性无关组 $H = (\boldsymbol{A}-4\boldsymbol{I})T_2 \cup \{\boldsymbol{Y}_i\}$, 从而得到 $M^* = T_2 \cup \{\boldsymbol{Y}_i\}$ 中的 $\boldsymbol{X}_4,\boldsymbol{X}_5,\boldsymbol{X}_3$ 作为箭头图中各行的最后一个向量.

习题 5.8

1. 求下列复方阵的若尔当标准形.

(1) $\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 4 & 0 \\ 2 & 3 & 0 & 4 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & 2 & 4 & 7 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 3 \end{pmatrix}$; (3) $\begin{pmatrix} 4 & 3 & 0 & 0 \\ -3 & -2 & 0 & 0 \\ 1 & 2 & -3 & 2 \\ 4 & 3 & -8 & 5 \end{pmatrix}$;

(4) $\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix}_{n\times n}$; (5) $\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n}^2$;

(6) $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \end{pmatrix}$; (7) $\begin{pmatrix} 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{pmatrix}_{n\times n}^n$

2. 求可逆复方阵 $\boldsymbol{P}$ 将第 1 题各小题中的复方阵相似到若尔当标准形.

3. 已知方阵

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

求: (1) $\boldsymbol{A}^{10}$; (2) $e^{\boldsymbol{A}t}$.

数 学 实 验

I 线性代数中常用的 MATLAB 命令

1. 求解线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$

求解

$$\begin{cases} x+\ y+\ z=1, \\ x-\ y-\ z=2, \\ 5x-2y+2z=3, \end{cases}$$

将此方程组写为

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 5 & -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

方法 1　通过矩阵求逆

MATLAB 命令

输入：A=[1, 1, 1; 1, −1, −1; 5, −2, 2]

输出:

```
A =
     1     1     1
     1    -1    -1
     5    -2     2
```

输入：b = [1; 2; 3]

输出：

```
b =
     1
     2
     3
```

输入：det(A)　　　　　　　　　　　% 求 $\boldsymbol{A}$ 的行列式

输出：−8

输入：inv(A)*b　　　　　　　　　　%inv(A) 为对矩阵 $\boldsymbol{A}$ 求逆

（此处也可写成 A\b）

输出：

ans =

1.5000

0.8750

−1.3750

故此方程组的解为 $x = 1.5, y = 0.875, z = -1.375$.

方法 2 化简为三角矩阵求解

MATLAB 命令

输入：rref([A,b])

输出：ans =

1.0000	0	0	1.5000
0	1.0000	0	0.8750
0	0	1.0000	−1.3750

即原方程组简化为

$$\begin{cases} 1x + 0y + 0z = 1.5, \\ 0x + 1y + 0z = 0.875, \\ 0x + 0y + 1z = -1.375. \end{cases}$$

2. 求矩阵的特征值和特征向量

求：

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 5 & -2 & 2 \end{pmatrix}$$

的特征值与特征向量.

MATLAB 命令

输入：A=[1, 1, 1; 1, −1, −1; 5, −2, 2]

输出：

A =

1	1	1
1	−1	−1
5	−2	2

输入：eig(A) % 求$\boldsymbol{A}$的特征值

输出：

```
ans =
      3.8108
      0.8031
     -2.6139
```

输入： [lamda,e]=eig(A)　% 求A的特征值和特征向量，其中特征向量为 lamda 的列向量

输出：

```
lamda =
     -0.2884  -0.4060   0.3643
      0.1370  -0.6050  -0.6431
     -0.9477   0.6849  -0.6736
e =
     3.8108        0         0
          0   0.8031         0
          0        0   -2.6139
```

3. 实对称方阵正交相似于对角阵

MATLAB 命令

输入： A=[1, 2, 3; 2, 4, 5; 3, 5, 6]　% 输入实对称方阵A

输出: A =

```
     1  2  3
     2  4  5
     3  5  6
```

输入: [P,D] = eig(A)　% 求正交方阵 P 使$D=P^{-1}AP$ 是对角矩阵

输出: P=

```
      0.7370   0.5910  0.3280
      0.3280  -0.7370  0.5910
     -0.5910   0.3280  0.7370
D =
     -0.5157        0        0
           0   0.1709        0
           0        0  11.3448
```

4. 矩阵运算其他命令

MATLAB 命令

输入： A=[1, 1, 1; 1, −1, −1; 3, −1, −1]

输出： A =

```
      1    1    1
      1   -1   -1
      3   -1   -1
```

输入：rank(A) % 求矩阵的秩

输出：ans =

2

输入：trace(A) % 求 trace

输出：ans =

−1

输入：orth(A) % 求 A 的列向量空间的一组标准正交基

输出：ans =

```
 -0.0455    0.9117
 -0.4285   -0.3870
 -0.9024    0.1378
```

输入：A′ % 矩阵转置

输出：ans =

```
  1    1    3
  1   -1   -1
  1   -1   -1
```

输入：poly(A) % 求特征多项式，输出为 x^n 从高阶到低阶系数

输出：ans =

```
  1.0000  1.0000  - 6.0000  - 0.0000
```

输入：A=[1, -3, -2; -1, 1,-1; 2, 4, 5]

输出：A =

```
  1  -3  -2
 -1   1  -1
  2   4   5
```

输入：[V,J]=jordan(A) % 求 A 的若尔当标准形 $J = V^{-1}AV$

输出：V =

```
 -1  -1  1
  0  -1  0
  1   2  0
```

J =

```
     3  0  0
     0  2  1
     0  0  2
```

Ⅱ 线性代数中常用的 Mathematica 命令

1. 求解线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$

例 1 解线性方程组

$$\begin{cases} x+y+z=1, \\ x-y-z=2, \\ 5x-2y+2z=3. \end{cases}$$

In[1]:= A={{1,1,1},{1,−1,−1},{5,−2,2}}; b={1,2,3};
MatrixForm[A] (∗ 输入矩阵 $\boldsymbol{A}, \boldsymbol{b}$. 按矩阵形式输出 $\boldsymbol{A}$.)

Out[1]//MatrixForm=

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 5 & -2 & 2 \end{pmatrix}$$

方法 1 通过矩阵求逆求 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$.

In[2]:= Det[A] (∗ 求行列式 $\det \boldsymbol{A}$.)

Out[2]= −8 (∗ 输出答案 $\det \boldsymbol{A}=-8$, $\boldsymbol{A}$ 可逆.)

In[3]:= Inverse[A].b (∗ 由 Inverse[A] 求 $\boldsymbol{A}^{-1}$, 再乘 $\boldsymbol{b}$.)

Out[3]= $\left\{\frac{3}{2}, \frac{7}{8}, -\frac{11}{8}\right\}$ (∗ 输出原方程组解.)

In[4]:= N[%] (∗ 将求得的答案化成小数近似值.)

Out[4]= {1.5, 0.875, −1.375}

例 2 解线性方程组

$$\begin{cases} x_1+2x_2+3x_3+4x_4=-3, \\ x_1+2x_2-5x_4=1, \\ 2x_1+4x_2-3x_3-19x_4=6, \\ 3x_1+6x_2-3x_3-24x_4=7. \end{cases}$$

In[4]:= A={{1,2,3,4},{1,2,0,−5},{2,4,−3,−19},{3,6,−3,−24}};

b={−3,1,6,7}

方法 2 求 $\boldsymbol{Ax}=\boldsymbol{b}$ 的特解与 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解.

In[5]:= LinearSolve[A] (* 求 $\boldsymbol{Ax}=\boldsymbol{b}$ 的特解.)

Out[5]= $\left\{1,0,-\frac{4}{3},0\right\}$

In[6]:= NullSpace[A] (* 求 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解.)

Out[6]= {{5,0,−3,1}, {−2,1,0,0}}

(*$\boldsymbol{Ax}=\boldsymbol{b}$ 的通解 $\left(1,0,-\frac{4}{3},0\right)+c_1(5,0,-3,1)+c_2(-2,1,0,0)$.)

方法 3 化简增广矩阵.

In[7]:= M={{1,2,3,4,−3},{1,2,0,−5,1},{2,4,−3,−19,6},
{3,6,−3,−24,7}}; RowReduce[M]//MatrixForm

Out[7]//MatrixForm=

$$\begin{pmatrix} 1 & 2 & 0 & -5 & 1 \\ 0 & 0 & 1 & 3 & \frac{4}{3} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(* 由 $\boldsymbol{M}$ 化成的最简阶梯形矩阵可知 $\operatorname{rank}\boldsymbol{M}=2$. 也可用以下命令求 $\operatorname{rank}\boldsymbol{M}$, 其中 Length[NullSpace[M]] 是方程组 $\boldsymbol{Mx}=\boldsymbol{0}$ 的解空间维数.)

In[8]:= 5−Length[NullSpace[M]]
Out[8]= 2

2. 求矩阵的特征值和特征向量

例 3 求如下方阵 $\boldsymbol{A}$ 的特征值和特征向量.

$$\boldsymbol{A}=\begin{pmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

In[9]:= A={{2/3,2/3,−1/3},{−1/3,2/3,2/3},{2/3,−1/3,2/3}};

Eigensystem[A]

Out[9]= $\left\{\left\{\frac{1}{2}(1+\mathrm{i}\sqrt{3}),\frac{1}{2}(1-\mathrm{i}\sqrt{3}),1\right\},\right.$

$\left\{\left\{-1+\frac{1}{2}(1+\mathrm{i}\sqrt{3}),\frac{1}{2}(-1-\mathrm{i}\sqrt{3}),1\right\},\right.$

$\left.\left.\left\{-1-\frac{1}{2}(1+\mathrm{i}\sqrt{3}),\frac{1}{2}(-1+\mathrm{i}\sqrt{3}),1\right\},\{1,1,1\}\right\}\right\}$

(* 输入语句 "Print[Det[A]]; Transpose[A].A//MatrixForm" 可验证 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$ 且 $\det\boldsymbol{A}=1$, $\boldsymbol{A}$ 是三维几何空间中旋转变换的矩阵. 旋转轴就是属于特征值 1 的特征向量 $(1,1,1)$ 决定的直线. 由虚特征值 $\frac{1}{2}(1\pm\mathrm{i}\sqrt{3})=\cos 60^\circ+\mathrm{i}\sin 60^\circ$ 知旋转角为 60°.)

In[11]:= P=Det[x IdentityMatrix[3]−A]; Print[P]; Factor[P]

(* 求 $\boldsymbol{A}$ 的特征多项式 $\det(x\boldsymbol{I}-\boldsymbol{A})$ 并进行因式分解.)

$-1+2x-2x^2+x^3$

Out[11]= $(-1+x)(1-x+x^2)$

例 4 求可逆方阵 $\boldsymbol{P}$ 将如下实对称方阵 $\boldsymbol{S}$ 相似到对角阵.

$$\boldsymbol{S}=\begin{pmatrix}1&2&3\\2&4&5\\3&5&6\end{pmatrix}.$$

In[12]:= S={{1.0,2,3},{2,4,5},{3,5,6}};P=Eigenvectors[S];
MatrixForm[P]

(* 将 $\boldsymbol{S}$ 中的 1 写成 1.0 为了做数值计算. 矩阵 $\boldsymbol{P}$ 的各行是特征向量.)

Out[12]//MatrixForm=

$$\begin{pmatrix}0.327985&0.591009&0.736976\\-0.736976&-0.327985&0.591009\\0.591009&-0.736976&0.327985\end{pmatrix}$$

In[13]:= P.Transpose[P]//MatrixForm (* 计算 $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}$.)

Out[13]//MatrixForm=

$$\begin{pmatrix} 1. & 1.99114\times10^{-16} & 1.6282\times10^{-16} \\ 1.99114\times10^{-16} & 1. & 8.32803\times10^{-17} \\ 1.6282\times10^{-16} & 8.32803\times10^{-17} & 1. \end{pmatrix}$$

(*$\boldsymbol{PP}^{\mathrm{T}}$ 的非对角元都 $< 10^{15}$, 可以认为 $\boldsymbol{PP}^{\mathrm{T}}=\boldsymbol{I}$, $\boldsymbol{P}$ 是正交方阵.)

In[14]:= P.S.Transpose[P]//MatrixForm

Out[14]//MatrixForm=

$$\begin{pmatrix} 11.3448 & 3.62991\times10^{-16} & -2.09902\times10^{-16} \\ 3.09499\times10^{-16} & -0.515729 & 6.04714\times10^{-16} \\ -1.13096\times10^{-17} & 6.047\times10^{-16} & 0.170915 \end{pmatrix}$$

(*对称方阵 $\boldsymbol{S}$ 被正交方阵 $\boldsymbol{P}$ 相似到对角矩阵 diag(11.3448,−0.515729,0.170915).)

3. 方阵的幂与指数函数

例 5 求 $\boldsymbol{A}^n$ 与 $\mathrm{e}^{t\boldsymbol{A}}$, 其中 t 是常数,

$$\boldsymbol{A}=\begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{pmatrix}.$$

In[15]:= A={{3,2,1},{0,−1,−2},{0,0,−1}};P=MatrixPower[A,n];
Q=MatrixExp[t A]; Print[{MatrixForm[P],MatrixForm[Q]}]

$$\left\{\begin{pmatrix} 3^n & \frac{1}{2}(-(-1)^n+3^n) & -(-1)^n n \\ 0 & (-1)^n & 2(-1)^n n \\ 0 & 0 & (-1)^n \end{pmatrix}, \begin{pmatrix} e^{3t} & \frac{1}{2}e^{-t}(-1+e^{4t}) & e^{-t}t \\ 0 & e^{-t} & -2e^{-t}t \\ 0 & 0 & e^{-t} \end{pmatrix}\right\}$$

参 考 文 献

[1] 李尚志. 线性代数 (数学专业用). 北京: 高等教育出版社,2006.

[2] 陈怀琛, 高淑萍, 杨威. 工程线性代数 (MATLAB 版). 北京: 电子工业出版社, 2011.

[3] 张韵华. 符号计算系统 Mathematica 教程. 北京: 科学出版社,2001.

[4] 李尚志. 让抽象变得自然——建设国家精品课程的体会. 中国大学教学,2006(7):11-13.

[5] 李尚志. 线性代数精彩应用案例 (之一). 大学数学,2006,22(3):1-8.

[6] 李尚志. 线性代数精彩应用案例 (之二). 大学数学,2006,22(4):1-3.

[7] 李尚志. 若当标准形的计算. 大学数学.2006,22(5):1-10.

[8] 李尚志. 从问题出发引入线性代数概念. 高等数学研究,2006(5):6-8.

[9] 李尚志. 从问题出发引入线性代数概念 (续). 高等数学研究,2006(6): 12-15.

郑重声明

线性代数

Linear Algebra

ISBN 978-7-04-031795-4

9 787040 317954 >

定价 26.00 元