数字副本使用说明

PB19020539:

您好！

您的借阅证是：PB19020539，所属单位是：中国科学技术大学图书馆。

线性代数
疑难问题选讲

蒲和平

Answers to Selected Perplexed Questions in Linear Algebra

高等教育出版社

线性代数疑难问题选讲

Xianxing Daishu Yinan Wenti Xuanjiang

蒲和平

高等教育出版社·北京

内容提要

本书对工科类线性代数课程中一些疑难问题作了较为深入的讨论。其内容主要涉及四个方面：一是对教材中一些重要的概念、理论和方法作剖析，揭示包含其中的数学原理与思想，归纳总结一些重要的数学方法；二是对于一些容易混淆的概念辨明它们之间的区别与联系；三是对教材中一些常用而又没给出证明的定理补充证明，并对部分内容作了适当的延伸；四是收集整理了大量的应用实例供读者参考。

本书可作为线性代数课程的教学参考书，对教师的教学与学生的学习都有很好的启迪与帮助。

图书在版编目（CIP）数据

线性代数疑难问题选讲 / 蒲和平主编. -- 北京：高等教育出版社，2014.8

ISBN 978-7-04-040392-3

Ⅰ. ①线… Ⅱ. ①蒲… Ⅲ. ①线性代数 - 高等学校 - 教材 Ⅳ. ① O151.2

中国版本图书馆 CIP 数据核字（2014）第 162162 号

策划编辑	蒋　青	责任编辑	蒋　青	特约编辑	师钦贤	封面设计	赵　阳
版式设计	范晓红	插图绘制	黄建英	责任校对	刘娟娟	责任印制	张泽业

出版发行	高等教育出版社	咨询电话	400-810-0598
社　　址	北京市西城区德外大街4号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	三河市华骏印务包装有限公司	网上订购	http://www.landraco.com
开　　本	787mm×960mm　1/16		http://www.landraco.com.cn
印　　张	11.5	版　　次	2014年8月第1版
字　　数	200千字	印　　次	2014年8月第1次印刷
购书热线	010-58581118	定　　价	21.80元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　40392-00

前言

线性代数是高等学校理工科各专业的一门公共基础课程，该课程涉及学生面宽，基础性强，应用范围广，具有一定的抽象性；该课程对培养学生能力、提高学生素质具有重要的奠基作用。因此，加强线性代数课程的建设，提高线性代数课程的教学质量是一件具有深远意义的工作。为了帮助广大线性代数课程教师(特别是青年教师) 深入钻研教材，准确把握教材内容，提高教学质量，我们编写了这本《线性代数疑难问题选讲》，旨在为任课教师提供一本有益的教学参考书，同时也为广大学生深入理解线性代数课程知识提供良好的帮助与指导。

本书不以知识总结与解题技巧为目的，而是从历史的、全局的、多角度的去分析线性代数课程知识的意义与内在联系，帮助教师站在一个较高的平台驾驭教材。本书内容主要涉及以下四个方面：一是针对教材中的一些疑难问题 (一些重要的概念、理论、方法) 加以剖析，揭示这些问题的实质以及包含其中的数学原理与思想，使我们明确学习与教学目标，对其中部分问题还介绍了具体的处理与讲授方法。二是对于一些容易混淆的概念辨明它们之间的区别与联系，归纳总结一些重要的数学方法。三是对教材中一些常用而又没给出证明的定理补充证明，对教材内容作适当的延伸与拓展，为教师开阔视野搭建必要的平台。最后针对现行教材普遍存在的应用实例较少的现象，本书收集整理了大量的应用实例 (第六章)，这些例子涉及物理、化学、生物、经济、交通、几何、生活等领域，是教学中很好的素材，供读者选用。

为了适应不同层次读者的需求，本书所写内容可分为两个层次：大部分问题是为深入理解线性代数课程的内容而写的，各问题之间相对独立，读者可根据自己的情况有选择地阅读；部分问题是教材内容的延伸与拓展，可供教师及学有余力的学生参考。

本书以问题与解答的形式展开，这些问题部分来自于教师的教学实践与教学研究，部分来自于教学中的学生提问，内容都是作者多年教学工作的积累与总结，编写时还参考了国内外大量的资料。为便于理解，对较为抽象的问题我们都给出了具体例子，部分例子是我们编创的，具有一定的新意，教学中可有针对性地选用。多年的教学实践及青年教师培训表明，书中涉及的问题，以及对这些问题的认识与处理方法都是十分重要的。期望本书能对教师的教学起到有力的推

动作用，对学生的学习起到启迪思维、开阔视野的作用。

本书由蒲和平编写，在编写初期，电子科技大学李厚彪、何军华、于泰彬三位老师参加了部分内容的编写工作。“高等学校大学数学教学研究与发展中心”为本书的编写提供了项目研究资助；在编写过程中得到了西安交通大学马知恩教授的悉心指导与热情帮助，还得到了电子科技大学黄廷祝教授、谢云荪教授的支持与帮助，他们对本书的编写提出了很多宝贵的意见与建议，在此一并表示衷心的感谢!

限于作者水平，不妥之处在所难免，恳请同行专家及读者批评指正。

作者

2014 年 3 月于成都

目录

第一章 矩阵

问题 1.1 矩阵乘法中的几个问题

(1) 矩阵的乘法为什么要像教材中那样定义

矩阵是线性代数的主要研究对象, 是解决线性问题的有力工具. 在矩阵的运算中, 从形式上看, 加法与数乘都与数或向量的对应运算相一致, 容易被初学者接受. 但矩阵的乘法却是一种不常见的规则, 会使初学者感到 "奇异", 甚至难以理解. 事实上, 矩阵的乘法规则并不是数学家的刻意所为, 从下面两个方面我们可以看到矩阵乘法定义的自然性与科学性.

1° 描述实际问题的需要

例 1.1.1 三家商店 S_i $(i=1,2,3)$ 销售四种食品 F_j $(j=1,2,3,4)$, 每单位售价的价格矩阵为 (以元计):

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{array}{c} \begin{array}{cccc} F_1 & F_2 & F_3 & F_4 \end{array} \\ \begin{pmatrix} 17 & 7 & 11 & 21 \\ 15 & 9 & 13 & 19 \\ 18 & 8 & 15 & 19 \end{pmatrix} \end{array}.$$

矩阵每一行的元表示每家商店 4 种食品的单位售价. 若欲在 3 家商店 S_i $(i=1,2,3)$ 分别购买 F_j 食品 x_j 个单位 $(j=1,2,3,4)$, 试问: 应分别付给 3 家商店多少金额? 并用矩阵表示.

解 记 $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, $\boldsymbol{x}$ 表示购买量矩阵. 根据给定的单位售价, 应分别付给

3 家商店的金额为

$$S_1 : 17x_1 + 7x_2 + 11x_3 + 21x_4,$$
$$S_2 : 15x_1 + 9x_2 + 13x_3 + 19x_4,$$
$$S_3 : 18x_1 + 8x_2 + 15x_3 + 19x_4.$$

用 3×1 矩阵表示为

$$\begin{pmatrix} 17x_1 + 7x_2 + 11x_3 + 21x_4 \\ 15x_1 + 9x_2 + 13x_3 + 19x_4 \\ 18x_1 + 8x_2 + 15x_3 + 19x_4 \end{pmatrix}.$$

由于所用金额应该是单位售价与购买量之积, 因此, 我们自然可把上面的金额矩阵看作是单位售价矩阵与购买量矩阵的乘积, 即有

$$\begin{pmatrix} 17 & 7 & 11 & 21 \\ 15 & 9 & 13 & 19 \\ 18 & 8 & 15 & 19 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 17x_1 + 7x_2 + 11x_3 + 21x_4 \\ 15x_1 + 9x_2 + 13x_3 + 19x_4 \\ 18x_1 + 8x_2 + 15x_3 + 19x_4 \end{pmatrix} \triangleq \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

这就给出了两个矩阵相乘应该遵循的规则.

2° 科学表示与计算的需要

我们将上面的例子作以下两个方面的延伸:

一方面的延伸是: 给定确定的金额 $\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, 反求各商品的购买量 $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, 这就是求解线性方程组 $\begin{cases} 17x_1 + 7x_2 + 11x_3 + 21x_4 = b_1, \\ 15x_1 + 9x_2 + 13x_3 + 19x_4 = b_2, \\ 18x_1 + 8x_2 + 15x_3 + 19x_4 = b_3 \end{cases}$ 的问题, 按上面"乘法"的约定, 该方程组就可简记为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

其中 $\boldsymbol{A} = \begin{pmatrix} 17 & 7 & 11 & 21 \\ 15 & 9 & 13 & 19 \\ 18 & 8 & 15 & 19 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. 这种表示很像我们熟知的一元一次方程, 这不仅简化了方程组的表示, 在以后的学习中我们会感受到这种简化带来的很多好处.

另一方面的延伸是: 如果我们计划在这 3 家商店分两批次购买对应的 4 种食品, 假设两次的购买量分别为 $\begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{pmatrix}$, $\begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \end{pmatrix}$, 若用 $\begin{pmatrix} b_{11} \\ b_{12} \\ b_{13} \end{pmatrix}$, $\begin{pmatrix} b_{21} \\ b_{22} \\ b_{23} \end{pmatrix}$ 分别表示两次购买需付出的对应金额, 按上面的 "乘法" 约定, 自然会有

$$\begin{pmatrix} 17 & 7 & 11 & 21 \\ 15 & 9 & 13 & 19 \\ 18 & 8 & 15 & 19 \end{pmatrix} \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ x_{13} & x_{23} \\ x_{14} & x_{24} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \\ b_{13} & b_{23} \end{pmatrix},$$

其中 b_{ij} 也就是等式左端前一个矩阵的第 i 行元与后一个矩阵的第 j 列对应元乘积之和.

通过上面的描述, 我们就会很自然地得到教材中矩阵乘法的 "定义".

有了矩阵乘法的概念, 我们再来看看线性变换的复合运算问题:

例 1.1.2　设有两个线性变换

$$(\mathrm{I}) \begin{cases} x_1 = a_{11}y_1 + a_{12}y_2 + a_{13}y_3, \\ x_2 = a_{21}y_1 + a_{22}y_2 + a_{23}y_3, \end{cases}$$

$$(\mathrm{II}) \begin{cases} y_1 = b_{11}z_1 + b_{12}z_2, \\ y_2 = b_{21}z_1 + b_{22}z_2, \\ y_3 = b_{31}z_1 + b_{32}z_2. \end{cases}$$

试求出从 z_1, z_2 到 x_1, x_2 的线性变换.

解　将两线性变换分别用矩阵表示为

$$(\mathrm{I})\ \boldsymbol{x} = \boldsymbol{A}\boldsymbol{y}, \quad (\mathrm{II})\ \boldsymbol{y} = \boldsymbol{B}\boldsymbol{z}.$$

将 (II) 代入 (I), 得

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{z}$$

可见 $\boldsymbol{AB}$ 就是从 z_1, z_2 到 x_1, x_2 的线性变换的矩阵.

如果将以上运算用分量形式具体写出, 应为

$$\begin{cases} x_1 = a_{11}(b_{11}z_1 + b_{12}z_2) + a_{12}(b_{21}z_1 + b_{22}z_2) + a_{13}(b_{31}z_1 + b_{32}z_2), \\ x_2 = a_{21}(b_{11}z_1 + b_{12}z_2) + a_{22}(b_{21}z_1 + b_{22}z_2) + a_{23}(b_{31}z_1 + b_{32}z_2), \end{cases}$$

即

$$\begin{cases} x_1 = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})z_1 + (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32})z_2, \\ x_2 = (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31})z_1 + (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32})z_2. \end{cases}$$

易见

$$\boldsymbol{AB} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}$$

就是复合变换 $z \to y \to x$ 的矩阵, 由此会进一步加深我们对矩阵乘法 "定义" 的认识.

最后谈谈矩阵乘法对矩阵意义的影响. 我们通常说, 矩阵是一张数表, 对这张数表我们定义了多种运算, 如加法、数乘、转置等; 当有了矩阵的乘法, 这张 "静" 态的数表就 "动" 了起来, 可用它来表示一个线性变换 $\mathcal{A}: \boldsymbol{x} \mapsto \boldsymbol{Ax}$, 同时任意一个线性变换也都可以用矩阵来表示 (问题 4.8 定理 4.8.3), 这就使得矩阵的意义更加丰富, 其应用的范围也更为广泛.

(2) 矩阵乘法还有其他定义形式吗

数学中的各种概念都不是凭空产生的, 它通常是具体问题的概括与抽象. 与已有概念作类比或延伸也是一种常见的方式. 比如, 能否像矩阵加法那样, 将矩阵的乘法规定为对应元之积 (这也是初学者常提出的问题)? 其实这样的乘积定义是有的, 同时也能找到它的一些应用背景, 历史上有很多人都曾研究过这类运算. 除此之外, 矩阵乘积还有一些其他的定义, 这些定义在理论与应用方面都很有价值. 下面介绍两个不同的乘法定义.

定义 1.1.1 设 $\boldsymbol{A} = (a_{ij})_{m\times n}, \boldsymbol{B} = (b_{ij})_{m\times n}$ 是两个同型矩阵, 定义两矩阵的乘法为

$$\boldsymbol{A} \circ \boldsymbol{B} = (a_{ij}b_{ij}) = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{pmatrix}.$$

矩阵的这种乘法称为 **Hadamard** (阿达马, 法国数学家, 1865—1963) 积, 或 Schur (舒尔, 1875—1941) 积.

这种运算简单, 显然满足交换律 ($\boldsymbol{A} \circ \boldsymbol{B} = \boldsymbol{B} \circ \boldsymbol{A}$), 并且具有很丰富的结构和各种有趣的结果, 例如:

定理 1.1.1 (Schur 积定理) 设 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 阶实对称矩阵. 则当 $\boldsymbol{A}, \boldsymbol{B}$ 都是半正定矩阵时, $\boldsymbol{A} \circ \boldsymbol{B}$ 也是半正定矩阵; 特别地, 若 $\boldsymbol{A}, \boldsymbol{B}$ 都是正定矩阵, 那么 $\boldsymbol{A} \circ \boldsymbol{B}$ 也是正定矩阵.

这表明, 矩阵的正定性 (或半正定性) 在 Hadamard 积运算下是封闭的, 但这对普通矩阵乘法运算未必成立, 因为两个正定矩阵的乘积未必是对称的, 从而也未必是正定的矩阵. 另外, 该运算还有与普通矩阵乘法运算不一样的其他性质, 如:

定理 1.1.2 (Oppenheim (奥本海姆) 不等式) 设 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 阶正定矩阵, 那么

$$|\boldsymbol{A} \circ \boldsymbol{B}| \geqslant a_{11}a_{22}\cdots a_{nn}|\boldsymbol{B}| \geqslant |\boldsymbol{A}||\boldsymbol{B}| = |\boldsymbol{AB}|,$$

且最左端等号成立的充要条件为 $\boldsymbol{B}$ 是对角矩阵, 中间等号成立的充要条件为 $\boldsymbol{A}$ 是对角矩阵, 因此左右两端等号成立的充要条件为 $\boldsymbol{A}, \boldsymbol{B}$ 均为对角矩阵.

由此可知, Hadamard 积矩阵的行列式不同于普通矩阵乘法的行列式性质:

$$|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|.$$

Hadamard 积还有许多有趣的问题值得研究, 例如 Hadamard 积的逆、Hadamard 积的特征值和特征向量、Hadamard 积的对角化等问题.

另外一种应用较为广泛的矩阵乘积叫 “Kronecker 积” (克罗内克, 1823—1891), 它的定义为

定义 1.1.2 设 $\boldsymbol{A} = (a_{ij})_{m\times n}, \boldsymbol{B} = (b_{ij})_{p\times q}$, 定义

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{pmatrix}_{mp\times nq}$$

为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 Kronecker 积 (或称直积, 张量积).

这种乘积具有很丰富的结构和有趣的结果, 它在张量代数、矩阵方程求解、统计的正交设计、信号传输预处理、自动控制、规划理论和图像处理等工程领域中都有着广泛的应用. 有兴趣的读者可以查询相关的书籍, 在此不作详述.

任何一种运算, 只有具备了坚实的应用背景和良好的运算律, 才能得以广泛的应用, 才具有其强大的生命力.

在教学中我们发现学生经常不知道为什么要如此定义一些数学概念, 而习惯于从老师那里得到问题和结果. 在课堂上, 很多学生常常有这样的想法: 所有与数学有关的课程都是系统的、逻辑严密的, 是 “完美” 的、“无懈可击” 的; 并产生一种这样一种错觉: 数学就是从定义到定理最后才应用到实际中去的. 所有这些都潜移默化地阻碍了学生创造力的发展, 抑制了他们在思维上的主动性和

积极性. 德国数学家 Cantor (康托尔, 1845—1918) 曾说 "数学的本质在于它的自由". 因此, 适当鼓励学生大胆尝试一些新的运算, 给以正确的引导, 也不失为体现数学的 "自由" 精神之策, 这对学生创新能力的培养很有好处.

(3) 如何理解双重求和符号的可交换性

在矩阵的乘法结合律的证明中, 要用到双重求和符号的次序交换. 对这个问题, 初学者经常报以疑问. 若我们根据双重求和的意义将其 "平铺" 开来, 就很容易从 "直观" 上得以理解, 事实上 $\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{n} a_{ij}$ 是下面 $m\times n$ 个数的求和:

$$a_{11}, a_{12}, a_{13}, \cdots, a_{1m},$$

$$a_{21}, a_{22}, a_{23}, \cdots, a_{2m},$$

$$a_{31}, a_{32}, a_{33}, \cdots, a_{3m},$$

$$\cdots,$$

$$a_{n1}, a_{n2}, a_{n3}, \cdots, a_{nm}.$$

如果将上面的 $m\times n$ 个数先按行求和, 有

$$\sum_{j=1}^{m} a_{1j} + \sum_{j=1}^{m} a_{2j} + \sum_{j=1}^{m} a_{3j} + \cdots + \sum_{j=1}^{m} a_{nj} = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij}\right);$$

先按列求和, 有

$$\sum_{i=1}^{n} a_{i1} + \sum_{i=1}^{n} a_{i2} + \sum_{i=1}^{n} a_{i3} + \cdots + \sum_{i=1}^{n} a_{im} = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij}\right),$$

所以

$$\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} a_{ij}\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij}\right).$$

(4) 矩阵的乘法运算要注意哪些问题

初学者在矩阵的运算方面经常会犯一些习惯性的错误, 究其原因是将实数乘法的性质套用到了矩阵乘法上, 现就矩阵乘法运算中要注意的问题作几点评述.

1° 不是任意两个矩阵都可以相乘

按矩阵乘法的定义, 只有乘积左侧矩阵的列数与右侧矩阵的行数相等, 两矩阵才能相乘. 因而只有方阵才有幂运算, 才有多项式运算; 在矩阵作分块乘积时, 左侧矩阵列数的划分必须与右侧矩阵行数的划分相同, 也就是说如果左侧矩阵

在某两列之间添加了一条竖线来分块, 那么乘积右侧的矩阵就必须在相应的两行之间添加一条横线来作相应的分块, 反之亦然.

2° 矩阵乘法不满足交换律

交换顺序后两矩阵不一定能相乘, 即使能乘也未必相等. 因此当 $\boldsymbol{AB} \neq \boldsymbol{BA}$ 时, 下列等式均是错误的:

$$\begin{aligned}
\boldsymbol{A}^2 - \boldsymbol{B}^2 &= (\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B}),\\
(\boldsymbol{A}+\boldsymbol{B})^2 &= \boldsymbol{A}^2 + \boldsymbol{B}^2 + 2\boldsymbol{AB},\\
\boldsymbol{A}^3 - \boldsymbol{B}^3 &= (\boldsymbol{A}-\boldsymbol{B})(\boldsymbol{A}^2+\boldsymbol{AB}+\boldsymbol{B}^2),\\
(\boldsymbol{AB})^2 &= \boldsymbol{A}^2\boldsymbol{B}^2.
\end{aligned}$$

那么上面各等式成立的条件是什么? 显然 $\boldsymbol{AB}$ 可交换是它们都成立的充分条件, 但非共同 (仅是前两式) 的必要条件!

一般来说, $\boldsymbol{AB}$ 可交换是两矩阵多项式可交换

$$f(\boldsymbol{A})g(\boldsymbol{B}) = g(\boldsymbol{B})f(\boldsymbol{A}) \quad (f(x), g(x)\text{ 是多项式函数})$$

的充分条件, 但两矩阵多项式可交换的必要条件却各有各的不同.

两矩阵乘积的可交换性是一个形式简单而传统的问题, 但鲜有简洁有效的判别方法, 下面是一些较简单的特殊情况:

例 1.1.3 (1) 设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶对 (反) 称矩阵, 则 $\boldsymbol{A}, \boldsymbol{B}$ 可交换的充要条件是 $\boldsymbol{AB}$ 为对称矩阵;

(2) 设 $\boldsymbol{A}$ 为 n 阶对称矩阵, $\boldsymbol{B}$ 为 n 阶反称矩阵, 则 $\boldsymbol{A}, \boldsymbol{B}$ 可交换的充要条件是 $\boldsymbol{AB}$ 为反称矩阵.

例 1.1.4 与任一 n 阶矩阵可交换的矩阵只能是数量矩阵.

证 设 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 与任一 n 阶矩阵可交换. 记 $\boldsymbol{E}_i(c)$ 为单位矩阵的第 i 行乘数 c 所成的初等矩阵, $\boldsymbol{E}_{ij}$ 是单位矩阵交换第 i, j 两行所成的初等矩阵, 则由

$$\boldsymbol{E}_i(2)\boldsymbol{A} = \boldsymbol{AE}_i(2) \quad (i = 1, 2, \cdots, n),$$

可得 $a_{ij} = 0\ (i \neq j)$, 即 $\boldsymbol{A}$ 是对角矩阵.

再由 $\boldsymbol{E}_{ij}\boldsymbol{A} = \boldsymbol{AE}_{ij}\ (i \neq j)$ 可得 $a_{ii} = a_{jj}$, 所以 $\boldsymbol{A}$ 是数量矩阵.

例 1.1.5 设 $\boldsymbol{A} = \mathrm{diag}(a_1, \cdots, a_n)$ (当 $i \neq j$ 时 $a_i \neq a_j$), 则与 $\boldsymbol{A}$ 可交换的矩阵只能是对角矩阵.

例 1.1.6 设 $\boldsymbol{AB} = \alpha\boldsymbol{A} + \beta\boldsymbol{B}$, 其中 α, β 为非零常数, 则 $\boldsymbol{A}, \boldsymbol{B}$ 可交换.

证 由 $\boldsymbol{AB}=\alpha\boldsymbol{A}+\beta\boldsymbol{B}$ 可得

$$(\boldsymbol{A}-\beta\boldsymbol{I})(\boldsymbol{B}-\alpha\boldsymbol{I})=\alpha\beta\boldsymbol{I},\quad \frac{1}{\alpha\beta}(\boldsymbol{A}-\beta\boldsymbol{I})(\boldsymbol{B}-\alpha\boldsymbol{I})=\boldsymbol{I}\quad (\boldsymbol{I}\text{ 是单位矩阵}),$$

于是

$$\frac{1}{\alpha\beta}(\boldsymbol{B}-\alpha\boldsymbol{I})(\boldsymbol{A}-\beta\boldsymbol{I})=\boldsymbol{I}.$$

由此可得

$$\boldsymbol{BA}=\alpha\boldsymbol{A}+\beta\boldsymbol{B}=\boldsymbol{AB}.$$

矩阵乘积的可交换性还与矩阵的特征向量有关, 我们将在问题 4.9 中作些讨论.

3° 矩阵乘法不满足消去律

矩阵乘法不满足消去律是指: 若 $\boldsymbol{AB}=\boldsymbol{O}\not\Rightarrow\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$ ($\boldsymbol{O}$ 为零矩阵).

由此派生出的结果有

$$\boldsymbol{AB}=\boldsymbol{O},\quad \boldsymbol{A}\neq\boldsymbol{O}\not\Rightarrow\boldsymbol{B}=\boldsymbol{O}.$$

$$\boldsymbol{AB}=\boldsymbol{AC},\quad \boldsymbol{A}\neq\boldsymbol{O}\not\Rightarrow\boldsymbol{B}=\boldsymbol{C}.$$

那么, 在什么条件下矩阵乘法消去律是正确的? 即由 $\boldsymbol{AB}=\boldsymbol{O}$ 可推出 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$?

容易证明, 以下结论是正确的:

例 1.1.7 $\boldsymbol{AB}=\boldsymbol{O}\Rightarrow\boldsymbol{B}=\boldsymbol{O}$ 的充要条件是 $\boldsymbol{A}$ 为列满秩矩阵, $\boldsymbol{AB}=\boldsymbol{O}\Rightarrow\boldsymbol{A}=\boldsymbol{O}$ 的充要条件是 $\boldsymbol{B}$ 为行满秩矩阵.

我们只需证明结论中的第一部分, 因为第二部分可由 $\boldsymbol{AB}=\boldsymbol{O}$ 两边取转置而得到. 下面用两种方法来予以证明.

证法 1 充分性. 设 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵, 若 $\boldsymbol{A}$ 是列满秩的, 则必有 $m\geqslant n$, 且经过一系列初等行变换可将 $\boldsymbol{A}$ 化为标准形 $\begin{pmatrix}\boldsymbol{I}_n\\ \boldsymbol{O}\end{pmatrix}$ (其中 $\boldsymbol{I}_n$ 是 n 阶单位矩阵, $\boldsymbol{O}$ 是零矩阵). 即存在 m 阶可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_n\\ \boldsymbol{O}\end{pmatrix}$, 则 $\boldsymbol{AB}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_n\\ \boldsymbol{O}\end{pmatrix}\boldsymbol{B}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{B}\\ \boldsymbol{O}\end{pmatrix}$. 所以由 $\boldsymbol{AB}=\boldsymbol{O}$, 及 $\boldsymbol{P}$ 可逆可得 $\boldsymbol{B}=\boldsymbol{O}$.

必要性. 若 $\boldsymbol{A}$ 不是列满秩的, 设 $\boldsymbol{A}$ 的秩为 r, 则存在可逆矩阵 $\boldsymbol{P},\boldsymbol{Q}$ 使 $\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}$. 记 $\boldsymbol{QB}=\begin{pmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2\end{pmatrix}$ (其中 $\boldsymbol{B}_1$ 是 r 行的矩阵), 则 $\boldsymbol{AB}=$

$\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}_2\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{B}_1\\ \boldsymbol{O}\end{pmatrix}=\boldsymbol{O}$, 由此可得 $\boldsymbol{B}_1=\boldsymbol{O},\boldsymbol{B}_2$ 可为任意非零矩阵, 从而 $\boldsymbol{B}\neq\boldsymbol{O}$.

证法 2　若 $\boldsymbol{AB}=\boldsymbol{O}$, 则 $\boldsymbol{B}$ 的列向量一定是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解向量, 而该方程组只有零解的充要条件是矩阵 $\boldsymbol{A}$ 的列向量组线性无关, 即 $\boldsymbol{A}$ 是列满秩矩阵.

显然, 对可逆矩阵而言, 乘法消去律总是成立的.

问题 1.2　计算一个矩阵的方幂有哪些常见方法

方阵的任意次方幂计算是一件困难的事, 没有一种通用的方法能直接找到矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^n$ 的关系. 但对一些特殊矩阵或满足一定条件的矩阵, 也有一些方法可循.

(1) 用归纳法计算

当一个矩阵的阶数不高, 并且通过低次幂计算能找到 $\boldsymbol{A}^n$ 与 n 关系的, 可用归纳法证明这种关系, 并用于计算.

例 1.2.1　设 $\boldsymbol{A}=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}$, 计算 $\boldsymbol{A}^n$.

解　经简单计算易得

$$\boldsymbol{A}^2=\begin{pmatrix}1&0&0\\2&1&0\\3&2&1\end{pmatrix},\quad \boldsymbol{A}^3=\begin{pmatrix}1&0&0\\3&1&0\\6&3&1\end{pmatrix},\quad \boldsymbol{A}^4=\begin{pmatrix}1&0&0\\4&1&0\\10&4&1\end{pmatrix}.$$

假设 $\boldsymbol{A}^k=\begin{pmatrix}1&0&0\\k&1&0\\\dfrac{k(k+1)}{2}&k&1\end{pmatrix}$ 成立, 则

$$\boldsymbol{A}^{k+1}=\begin{pmatrix}1&0&0\\k&1&0\\\dfrac{k(k+1)}{2}&k&1\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}=\begin{pmatrix}1&0&0\\k+1&1&0\\\dfrac{(k+1)(k+2)}{2}&k+1&1\end{pmatrix},$$

由数学归纳法得 $\boldsymbol{A}^n=\begin{pmatrix}1&0&0\\n&1&0\\\dfrac{n(n+1)}{2}&n&1\end{pmatrix}$.

(2) 用递推公式计算

如果能找到 $\boldsymbol{A}^n$ 与其低次幂的关系 (递推公式), 就可利用递推公式作降幂计算.

例 1.2.2 设 $\boldsymbol{A}=\begin{pmatrix}1&0&0\\1&0&1\\0&1&0\end{pmatrix}$, 计算 $\boldsymbol{A}^{100}$.

解 $\boldsymbol{A}^2=\begin{pmatrix}1&0&0\\1&1&0\\1&0&1\end{pmatrix}$, $\boldsymbol{A}^3=\begin{pmatrix}1&0&0\\2&0&1\\1&1&0\end{pmatrix}=\boldsymbol{A}+\boldsymbol{A}^2-\boldsymbol{I}$ ($\boldsymbol{I}$ 是单位矩阵).

$$\begin{aligned}\boldsymbol{A}^4&=\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{A}^2-\boldsymbol{I})=\boldsymbol{A}^2+\boldsymbol{A}^3-\boldsymbol{A}\\&=\boldsymbol{A}^2+(\boldsymbol{A}+\boldsymbol{A}^2-\boldsymbol{I})-\boldsymbol{A}=\boldsymbol{A}^2+\boldsymbol{A}^2-\boldsymbol{I},\end{aligned}$$

用归纳法易证 $\boldsymbol{A}^n=\boldsymbol{A}^{n-2}+\boldsymbol{A}^2-\boldsymbol{I}\ (n\geqslant 3)$. 所以

$$\boldsymbol{A}^{100}=\boldsymbol{A}^{98}+\boldsymbol{A}^2-\boldsymbol{I}=\boldsymbol{A}^{96}+2(\boldsymbol{A}^2-\boldsymbol{I})=\cdots=\boldsymbol{A}^2+49(\boldsymbol{A}^2-\boldsymbol{I})=\begin{pmatrix}1&0&0\\50&1&0\\50&0&1\end{pmatrix}.$$

(3) 将矩阵作拆分计算

常见的拆分有两种:

拆分为和 $\boldsymbol{A}=a\boldsymbol{I}+b\boldsymbol{B}$, 其中 a,b 是数, 且 $\boldsymbol{B}$ 的方幂较容易计算;

拆分为乘积 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}$, 其中 $\boldsymbol{B},\boldsymbol{C}$ 分别为列满秩与行满秩矩阵. 这种方法只适用于 $\boldsymbol{A}$ 的秩较低的情况 (如 $\boldsymbol{A}$ 的秩为 1 或 2).

例 1.2.3 设 $\boldsymbol{A}=\begin{pmatrix}a&b&c\\0&a&b\\0&0&a\end{pmatrix}$, 计算 $\boldsymbol{A}^n$.

解 将矩阵 $\boldsymbol{A}$ 作拆分, 有

$$\boldsymbol{A}=a\boldsymbol{I}+b\boldsymbol{B},$$

其中 $\boldsymbol{B}=\begin{pmatrix}0&1&c/b\\0&0&1\\0&0&0\end{pmatrix}$. 由于 $\boldsymbol{B}^2=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}$, $\boldsymbol{B}^3=\boldsymbol{O}$, 所以

$$\begin{aligned}\boldsymbol{A}^n&=(b\boldsymbol{B}+a\boldsymbol{I})^n=\sum_{k=0}^{n}\mathrm{C}_n^k(b\boldsymbol{B})^k(a\boldsymbol{I})^{n-k}.\\&=a^n\boldsymbol{I}+\mathrm{C}_n^1ba^{n-1}\boldsymbol{B}+\mathrm{C}_n^2b^2a^{n-2}\boldsymbol{B}^2\end{aligned}$$

$$= \begin{pmatrix} a^n & nba^{n-1} & nca^{n-1} + \dfrac{n(n-1)}{2}b^2a^{n-2} \\ 0 & a^n & nba^{n-1} \\ 0 & 0 & a^n \end{pmatrix}.$$

例 1.2.4　设 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 2 & 4 & \cdots & 2n \\ \vdots & \vdots & & \vdots \\ n & 2n & \cdots & n^2 \end{pmatrix}$. 计算 $\boldsymbol{A}^n$ $(n > 1)$.

解　矩阵 $\boldsymbol{A}$ 的秩为 1, 作满秩分解, 有

$$\boldsymbol{A} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} (1, 2, \cdots, n) \triangleq \boldsymbol{BC}.$$

所以

$$\boldsymbol{A}^n = \boldsymbol{B}(\boldsymbol{CB})^{n-1}\boldsymbol{C} = \left(\sum_{k=1}^{n} k^2\right)^{n-1} \boldsymbol{BC} = \left(\sum_{k=1}^{n} k^2\right)^{n-1} \boldsymbol{A}.$$

注　有时需要上面两种方式的综合运用, 读者可考虑 $\begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}^n$ 的计算.

例 1.2.5　求下列矩阵 $\boldsymbol{A}$ 的满秩分解, 并据此计算 $\boldsymbol{A}^n$.

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 & -1 & 2 \\ 2 & -3 & -2 & 3 \\ 2 & -2 & -2 & 2 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

解　首先, 将矩阵 $\boldsymbol{A}$ 经初等行变换化为行最简形矩阵.

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

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

由此可得 $\boldsymbol{A}$ 的满秩分解 (见下面**定理 1.2.1**)

$$\boldsymbol{A} = \begin{pmatrix} 1 & -2 \\ 2 & -3 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \triangleq \boldsymbol{BD}.$$

同前面秩 1 矩阵的幂计算一样, 当 $n \geqslant 1$ 时

$$\boldsymbol{A}^n = (\boldsymbol{BD})^n = \boldsymbol{B}(\boldsymbol{DB})^{n-1}\boldsymbol{D} = \begin{pmatrix} 1 & -2 \\ 2 & -3 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & -2 \end{pmatrix}^{n-1} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -2 \\ 2 & -3 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} (-1)^{n-1} & 0 \\ (-1)^{n-1} - (-2)^{n-1} & (-2)^{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} -(-2)^n + (-1)^n & (-2)^n & (-2)^n - (-1)^n & -(-2)^n \\ \frac{1}{2}[-3(-2)^n + 2(-1)^n] & \frac{3}{2}(-2)^n & \frac{1}{2}[3(-2)^n - 2(-1)^n] & -\frac{3}{2}(-2)^n \\ -(-2)^n & (-2)^n & (-2)^n & -(-2)^n \\ -\frac{1}{2}(-2)^n & \frac{1}{2}(-2)^n & \frac{1}{2}(-2)^n & -\frac{1}{2}(-2)^n \end{pmatrix}.$$

注 关于矩阵的满秩分解有下面结论:

定理 1.2.1 设 $\boldsymbol{A}$ 是秩为 r 的 $m \times n$ 矩阵, 则存在 $m \times r$ 列满秩矩阵 $\boldsymbol{B}$ 与 $r \times n$ 行满秩矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{A} = \boldsymbol{BC}$.

证 因为 $\boldsymbol{A}$ 的秩为 r, 则存在 m 阶可逆矩阵 $\boldsymbol{P}$ 与 n 阶可逆矩阵 $\boldsymbol{Q}$, 使得

$$\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \boldsymbol{Q} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{I}_r \\ \boldsymbol{O} \end{pmatrix} (\boldsymbol{I}_r \quad \boldsymbol{O})\boldsymbol{Q}.$$

记 $\boldsymbol{B} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{I}_r \\ \boldsymbol{O} \end{pmatrix}$, $\boldsymbol{C} = (\boldsymbol{I}_r \ \boldsymbol{O})\boldsymbol{Q}$, 则 $\boldsymbol{B}$ 恰为矩阵 $\boldsymbol{P}$ 的前 r 列, $\boldsymbol{C}$ 恰为矩阵 $\boldsymbol{Q}$ 的前 r 行, 且有 $\boldsymbol{A} = \boldsymbol{BC}$.

从上面的证明中看到, 要得到分解式 $\boldsymbol{A}=\boldsymbol{BC}$, 就要记录初等变换 $\boldsymbol{A}\to\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$ 中的可逆矩阵 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$. 但事实上没有这个必要. 因为对矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1\boldsymbol{\alpha}_2\cdots\boldsymbol{\alpha}_n)$ 作初等行变换不改变其列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的线性相关性, 因此若将矩阵 $\boldsymbol{A}$ 经初等行变换化为行最简形

$$\boldsymbol{H}_A=\begin{pmatrix}\boldsymbol{H}_r\\ \boldsymbol{O}\end{pmatrix},$$

则 $\boldsymbol{C}=\boldsymbol{H}_r$, 若 $\boldsymbol{H}_r$ 的各行非零首元所在的列为 $j_1,j_2,\cdots,j_r$, 则 $\boldsymbol{B}=(\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_r})$.

(4) 将矩阵相似对角化后作计算

若矩阵 $\boldsymbol{A}$ 可相似对角化, 即

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\boldsymbol{P}^{-1},\quad 则\quad \boldsymbol{A}^k=\boldsymbol{P}\begin{pmatrix}\lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k\end{pmatrix}\boldsymbol{P}^{-1}.$$

由于这种计算在一般的教材中都有详细讨论, 这里不再赘述. 当然这种计算矩阵方幂的方法也是最为普遍适用的.

(5) 用 Hamilton-Cayley 定理简化计算

定理 1.2.2 (Hamilton-Cayley (哈密顿 – 凯莱)) n 阶矩阵 $\boldsymbol{A}$ 是它特征多项式

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

的根, 即

$$f(\boldsymbol{A})=\boldsymbol{A}^n+a_1\boldsymbol{A}^{n-1}+\cdots+a_{n-1}\boldsymbol{A}+a_n\boldsymbol{I}=\boldsymbol{O}.$$

定理证明见例 1.4.1 (此处略).

利用该定理及多项式的带余除法, 可以将任一大于或等于 n 次的 $\boldsymbol{A}$ 的乘幂用小于 n 次的 $\boldsymbol{A}$ 的乘幂来表示, 这样就可简化矩阵多项式的计算.

例 1.2.6 设 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$, 计算 $g(\boldsymbol{A})=\boldsymbol{A}^7-\boldsymbol{A}^5-19\boldsymbol{A}^4+28\boldsymbol{A}^3+6\boldsymbol{A}-4\boldsymbol{I}$.

解 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{vmatrix}=\lambda^3-4\lambda^2+5\lambda-2.$$

用 $f(x)$ 除 $g(x)$ 得余式为

$$r(x)=-3x^2+22x-8,$$

于是

$$\begin{aligned}g(\boldsymbol{A})&=r(\boldsymbol{A})=-3\boldsymbol{A}^2+22\boldsymbol{A}-8\boldsymbol{I}\\&=-3\begin{pmatrix}-3&2&0\\-8&5&0\\1&1&4\end{pmatrix}+22\begin{pmatrix}-1&1&0\\-4&3&0\\1&0&2\end{pmatrix}-8\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}-21&16&0\\-64&43&0\\19&-3&24\end{pmatrix}.\end{aligned}$$

问题 1.3 如何理解逆矩阵的概念

在数的运算中, 由乘法导出了除法: 若两数 a,x 的乘积为 b, 即 $ax=b$, 则称由 $a,b\ (a\neq 0)$ 求因数 x 的运算为除法. 求解的方法是: 先求一个数 a^{-1} 满足 $a^{-1}a=1$, 称 a^{-1} 为 a 的 (乘法) 逆, 在方程 $ax=b$ 两边同乘 a^{-1} 得 $x=a^{-1}b$.

对矩阵的乘法运算自然会提出类似的问题, 即已知矩阵 $\boldsymbol{A},\boldsymbol{B}$, 求矩阵 $\boldsymbol{X}$ 使之满足矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 与 $\boldsymbol{XA}=\boldsymbol{B}$ (由于矩阵乘法不满足交换律, 所以这里有两个方程). 如果能找到一个矩阵 $\boldsymbol{A}^{-1}$ 使得 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$, 则在方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的两边左乘 $\boldsymbol{A}^{-1}$ 就可得到 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$, 在方程 $\boldsymbol{XA}=\boldsymbol{B}$ 的两边右乘 $\boldsymbol{A}^{-1}$ 就可得到 $\boldsymbol{X}=\boldsymbol{B}\boldsymbol{A}^{-1}$, 这样就可以实现矩阵乘法的逆运算问题. 因此人们将满足条件

$$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I} \tag{1.3.1}$$

的矩阵 $\boldsymbol{B}$ 称为矩阵 $\boldsymbol{A}$ 的逆矩阵 (也称 $\boldsymbol{A}$ 是可逆的), 并记为 $\boldsymbol{B}=\boldsymbol{A}^{-1}$. 这就是逆矩阵概念的缘由.

对一个数 a 来说, 只要 $a\neq 0$, a^{-1} 就一定存在. 对一个矩阵 $\boldsymbol{A}$ 来说, 要使 $\boldsymbol{A}^{-1}$ 存在就不是那么简单. 首先能使 (1.3.1) 成立的矩阵 $\boldsymbol{A}$ 必须是非零方阵, 但

仅有非零方阵的条件还不能保证 $\boldsymbol{A}$ 可逆. 比如 $\boldsymbol{A}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$ 就没有逆矩阵. 另一方面, 的确又存在大量的矩阵可逆, 从变换的角度容易想到, 单位矩阵 $\boldsymbol{I}$ 经一次初等变换得到的矩阵 (初等矩阵) $\boldsymbol{A}=\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}$ 一定可逆, 其逆矩阵就是其逆变换所对应的矩阵 $\boldsymbol{A}^{-1}=\begin{pmatrix}1 & -k\\ 0 & 1\end{pmatrix}$; 再如, 平面上绕原点的旋转变换对应的矩阵 $\boldsymbol{A}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$, 其逆矩阵也是其逆变换所对应的矩阵

$$\boldsymbol{A}^{-1}=\begin{pmatrix}\cos(-\theta) & \sin(-\theta)\\ -\sin(-\theta) & \cos(-\theta)\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

这些结果都很容易由逆矩阵的定义来加以验证. 因此, 寻求矩阵可逆的条件以及求逆的方法就成了矩阵运算中十分重要的内容.

(1.3.1) 式给出了矩阵 $\boldsymbol{A}$ 的逆矩阵的定义, 但在判断矩阵 $\boldsymbol{B}$ 是否为 $\boldsymbol{A}$ 的逆矩阵时, 却只需验证 $\boldsymbol{AB}=\boldsymbol{I}$ 或 $\boldsymbol{BA}=\boldsymbol{I}$ 其中之一就行了. 其原因为

若 $\boldsymbol{AB}=\boldsymbol{I}$, 则 $|\boldsymbol{A}|\cdot|\boldsymbol{B}|=1\Rightarrow|\boldsymbol{A}|\neq 0$, 所以 $\boldsymbol{A}$ 可逆, 则有

$$\boldsymbol{BA}=(\boldsymbol{A}^{-1}\boldsymbol{A})(\boldsymbol{BA})=\boldsymbol{A}^{-1}(\boldsymbol{AB})\boldsymbol{A}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}.$$

当然我们也可以不依赖行列式来完成上面的证明, 利用矩阵的等价标准形也可完成其证明, 具体方法如下:

例 1.3.1　若 n 阶方阵 $\boldsymbol{AB}=\boldsymbol{I}_n$, 则必有 $\boldsymbol{BA}=\boldsymbol{I}_n$.

证　设存在可逆矩阵 $\boldsymbol{P},\boldsymbol{Q}$ 使得

$$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}.$$

以 $\boldsymbol{Q}^{-1}\boldsymbol{B}$ 右乘上式两边, 并由 $\boldsymbol{AB}=\boldsymbol{I}_n$ 可得

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}^{-1}\boldsymbol{B}\Rightarrow\boldsymbol{I}_n=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}^{-1}\boldsymbol{BP}^{-1}.$$

由上式可知, 必有 $r=n$. 否则上式右边矩阵的最后一行必全为 0, 这是不可能的. 于是我们得到

$$\boldsymbol{PAQ}=\boldsymbol{I}_n,\quad \boldsymbol{Q}^{-1}\boldsymbol{BP}^{-1}=\boldsymbol{I}_n.$$

由上面两式可得

$$\boldsymbol{BA}=(\boldsymbol{QP})(\boldsymbol{P}^{-1}\boldsymbol{Q}^{-1})=\boldsymbol{I}_n.$$

问题 1.4 如何理解伴随矩阵的意义

设矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$, 其元 a_{ij} 的代数余子式为 A_{ij}, 则称矩阵 $\boldsymbol{A}^*=(A_{ji})_{n\times n}$ 为 $\boldsymbol{A}$ 的**伴随矩阵**. 伴随矩阵是为了解决求逆矩阵问题而提出来的, 它与矩阵 $\boldsymbol{A}$ 的最基本也是最重要的关系为

$$\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{I}. \tag{1.4.1}$$

(1.4.1) 式揭示了 $\boldsymbol{A}^*$ 与 $\boldsymbol{A}$ 的关系, 也建立了 $\boldsymbol{A}$ 与其行列式 $|\boldsymbol{A}|$ 的联系, 许多关于伴随矩阵的性质及其应用都是由该等式来建立的. 从另一个角度讲, (1.4.1) 式也是行列式展开性质

$$\sum_{k=1}^{n}a_{ki}A_{kj}=\begin{cases}|\boldsymbol{A}|, & i=j,\\ 0, & i\neq j\end{cases}$$

的矩阵表现形式. 因此, 伴随矩阵又是行列式在矩阵中的融合.

(1.4.1) 式告诉我们: 若 $|\boldsymbol{A}|\neq 0$, 则 $\boldsymbol{A}$ 可逆, 并且

$$\boldsymbol{A}^{-1}=\frac{1}{\boldsymbol{A}}\boldsymbol{A}^*. \tag{1.4.2}$$

(1.4.2) 给出了求矩阵逆的公式. 由于当矩阵的阶数稍大 ($n>3$) 时, $\boldsymbol{A}^*$ 的计算量就较大, 所以用它来计算 $\boldsymbol{A}^{-1}$ 的价值不高, 但在理论推导中却十分有用, 因为其他求逆矩阵的方法很难用较简单的数学公式来表示.

利用 (1.4.2) 式我们容易推出线性方程组求解的 Cramer (克拉默, 1704—1752) 法则 (见问题 2.4), 利用 (1.4.2) 式还能得到关于矩阵特征多项式的 **Hamilton-Cayley 定理**.

例 1.4.1 (Hamilton-Cayley 定理) n 阶矩阵 $\boldsymbol{A}$ 是它特征多项式

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

的根, 即 $f(\boldsymbol{A})=\boldsymbol{A}^n+a_1\boldsymbol{A}^{n-1}+\cdots+a_{n-1}\boldsymbol{A}+a_n\boldsymbol{I}=\boldsymbol{O}$.

证 设 $\boldsymbol{B}(\lambda)$ 是特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 的伴随矩阵, 由 (1.4.1) 式有

$$\boldsymbol{B}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})=|\lambda\boldsymbol{I}-\boldsymbol{A}|\boldsymbol{I}=f(\lambda)\boldsymbol{I}. \tag{1.4.3}$$

因为 $\boldsymbol{B}(\lambda)$ 的元是 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 中各元的代数余子式, 所以都是 $\boldsymbol{\lambda}$ 的多项式, 其次数不大于 $n-1$. 因此 $\boldsymbol{B}(\lambda)$ 可以写成矩阵多项式

$$\boldsymbol{B}(\lambda)=\boldsymbol{B}_0\lambda^{n-1}+\boldsymbol{B}_1\lambda^{n-2}+\cdots+\boldsymbol{B}_{n-2}\lambda+\boldsymbol{B}_{n-1},$$

其中 $\boldsymbol{B}_0, \boldsymbol{B}_1, \cdots, \boldsymbol{B}_{n-1}$ 是 n 阶常数矩阵, 于是

$$\begin{aligned}\boldsymbol{B}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A}) &= (\boldsymbol{B}_0\lambda^{n-1}+\boldsymbol{B}_1\lambda^{n-2}+\cdots+\boldsymbol{B}_{n-2}\lambda+\boldsymbol{B}_{n-1})(\lambda\boldsymbol{I}-\boldsymbol{A})\\ &= \boldsymbol{B}_0\lambda^n+(\boldsymbol{B}_1-\boldsymbol{B}_0\boldsymbol{A})\lambda^{n-1}+\cdots+(\boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}\boldsymbol{A})\lambda-\boldsymbol{B}_{n-1}\boldsymbol{A}.\end{aligned}$$

上式代入 (1.4.3) 比较两边系数, 得

$$\begin{aligned}\boldsymbol{B}_0 &= \boldsymbol{I},\\ \boldsymbol{B}_1-\boldsymbol{B}_0\boldsymbol{A} &= a_1\boldsymbol{I},\\ &\cdots\cdots\cdots\cdots\\ \boldsymbol{B}_{n-1}-\boldsymbol{B}_{n-2}\boldsymbol{A} &= a_{n-1}\boldsymbol{I},\\ -\boldsymbol{B}_{n-1}\boldsymbol{A} &= a_n\boldsymbol{I}.\end{aligned}$$

用 $\boldsymbol{A}^n, \boldsymbol{A}^{n-1}, \cdots, \boldsymbol{A}, \boldsymbol{I}$ 分别乘上式两边, 然后相加, 即得

$$\boldsymbol{A}^n+a_1\boldsymbol{A}^{n-1}+\cdots+a_{n-1}\boldsymbol{A}+a_n\boldsymbol{I}=\boldsymbol{O}.$$

伴随矩阵有很多良好的性质, 这里将一些较常见的性质列于下面, 并粗略地将它们归为三类 (由于这些性质的证明并不困难, 我们仅给出其中个别的证明):

1° 运算性质

(i) $(k\boldsymbol{A})^*=k^{n-1}\boldsymbol{A}^*$ (k 为常数, n 为 $\boldsymbol{A}$ 的阶数);

(ii) $(\boldsymbol{A}\boldsymbol{B})^*=\boldsymbol{B}^*\boldsymbol{A}^*$;

(iii) $(\boldsymbol{A}^{\mathrm{T}})^*=(\boldsymbol{A}^*)^{\mathrm{T}}$;

(iv) 若 $\boldsymbol{A}$ 可逆, 则 $(\boldsymbol{A}^{-1})^*=(\boldsymbol{A}^*)^{-1}=\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}$;

(v) $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$ (n 为 $\boldsymbol{A}$ 的阶数);

(vi) $\operatorname{rank}\boldsymbol{A}^*=\begin{cases}n, & \operatorname{rank}\boldsymbol{A}=n,\\ 1, & \operatorname{rank}\boldsymbol{A}=n-1,\\ 0, & \operatorname{rank}\boldsymbol{A}<n-1.\end{cases}$

证　(ii) 当 $\boldsymbol{A}, \boldsymbol{B}$ 可逆时, 由 $\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}$ 可得

$$(\boldsymbol{A}\boldsymbol{B})^*=|\boldsymbol{A}\boldsymbol{B}|(\boldsymbol{A}\boldsymbol{B})^{-1}=|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}=(|\boldsymbol{B}|\boldsymbol{B}^{-1})(|\boldsymbol{A}|\boldsymbol{A}^{-1})=\boldsymbol{B}^*\boldsymbol{A}^*;$$

当 $\boldsymbol{A}$ 不可逆时, 因为 $\boldsymbol{A}$ 的特征值只有有限多个, 故必存在实数 $\delta>0$, 使得对所有满足 $0<\lambda<\delta$ 的 λ, 都有 $|\boldsymbol{A}-\lambda\boldsymbol{I}|\neq 0$, 因而 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 可逆. 对 $\boldsymbol{B}$ 作同样处理. 则有

$$[(\boldsymbol{A}-\lambda\boldsymbol{I})(\boldsymbol{B}-\lambda\boldsymbol{I})]^*=(\boldsymbol{B}-\lambda\boldsymbol{I})^*(\boldsymbol{A}-\lambda\boldsymbol{I})^*.$$

上式两边矩阵中的元都是 λ 的多项式, 是 λ 的连续函数, 令 $\lambda \to 0$ 即得 $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$.

2° 关联性质

(i) 若 $\boldsymbol{A}$ 是上 (下) 三角形 (对角) 矩阵, 则 $\boldsymbol{A}^*$ 也是上 (下) 三角形 (对角) 矩阵;

(ii) 若 $\boldsymbol{A}$ 是对称矩阵, 则 $\boldsymbol{A}^*$ 也是对称矩阵;

(iii) 若 $\boldsymbol{A}$ 是可逆矩阵, 则 $\boldsymbol{A}^*$ 也是可逆矩阵;

(iv) 若 $\boldsymbol{A}$ 是正交矩阵, 则 $\boldsymbol{A}^*$ 也是正交矩阵;

(v) 若 $\boldsymbol{A}$ 是 (半) 正定矩阵, 则 $\boldsymbol{A}^*$ 也是 (半) 正定矩阵.

证 (iv) 若 $\boldsymbol{A}$ 是正交矩阵, 则

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}.$$

上式两边取伴随矩阵, 利用 1° 中的运算性质, 并注意 $\boldsymbol{I}^* = \boldsymbol{I}$, 可得

$$(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^* = \boldsymbol{I}^* \Rightarrow (\boldsymbol{A}^{\mathrm{T}})^*\boldsymbol{A}^* = \boldsymbol{I} \Rightarrow (\boldsymbol{A}^*)^{\mathrm{T}}\boldsymbol{A}^* = \boldsymbol{I},$$

所以 $\boldsymbol{A}^*$ 也是正交矩阵.

关联性质反映了 $\boldsymbol{A}^*$ 对原矩阵 $\boldsymbol{A}$ 某些性质的良好继承, 且有些还是充分必要的. 借助这性质也可以由 $\boldsymbol{A}^*$ 来研究 $\boldsymbol{A}$ 的相关问题.

3° 传导性质

(i) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 乘积可交换, 则 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 的乘积也可交换;

(ii) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 则 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 也等价;

(iii) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 也相似;

(iv) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同, 则 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 也合同.

证 (iii) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则存在可逆矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}.$$

上式两边取伴随矩阵, 并利用运算性质可得

$$(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^* = \boldsymbol{B}^* \Rightarrow \boldsymbol{P}^*\boldsymbol{A}^*(\boldsymbol{P}^{-1})^* = \boldsymbol{B}^* \Rightarrow \boldsymbol{P}^*\boldsymbol{A}^*(\boldsymbol{P}^*)^{-1} = \boldsymbol{B}^*,$$

所以 $\boldsymbol{A}^*$ 与 $\boldsymbol{B}^*$ 相似.

传导性质告诉我们, 对伴随矩阵而言, 已没有再单独研究其等价性、相似性、合同性的必要.

总之, 伴随矩阵是依附于原矩阵而存在的一个概念, 它将行列式的某些性质融入了矩阵之中, 从而使矩阵的一些特殊性态被揭示了出来; 同时伴随矩阵又保持了原矩阵的某些特性, 通过它可以帮助我们研究原矩阵的性质.

问题 1.5　矩阵的秩有何意义? 它有哪些等价的描述

秩是矩阵初等变换 (等价) 下的不变量, 反映了矩阵的固有的特质. 正是这种不变性, 才能够使我们通过初等变换及相关运算将许多问题中隐藏的 “本质” 给揭示出来.

例如, 矩阵是否可逆, 乘法消去律是否成立 (见问题 1.1(4)) 这都是由矩阵的秩所确定的; 向量组是否线性相 (无) 关, 最大无关组所含向量的个数, 两组向量是否等价 (见问题 3.6); 齐次线性方程组解空间的维数, 非齐次线性方程组是否有解, 解是否唯一; 线性变换 $\mathcal{A}:\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 的单、满性 (见问题 3.7); 矩阵是否有零特征值, 零特征值重数的估计 (见问题 4.4); 矩阵是否等价, 矩阵等价、相似、合同的标准形的构成, 等等. 众多问题都离不开矩阵秩的刻画, 可以说矩阵的秩贯穿了线性代数的始终. 正是因为秩的重要, 才使得线性代数课程中有许多关于矩阵秩的讨论.

下面再列举几个有关矩阵秩的简单应用.

例 1.5.1　证明向量组 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$ 与向量组 $\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 等价的充要条件是 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{B})=\mathrm{rank}(\boldsymbol{A},\boldsymbol{B})$.

证　由于向量 $\boldsymbol{\beta}_i$ $(1\leqslant i\leqslant n)$ 可由向量组 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$ 线性表出的充要条件是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}_i$ 有解, 即 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A},\boldsymbol{\beta}_i)$, 所以向量组 $\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 可由向量组 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$ 线性表出的充要条件是 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A},\boldsymbol{B})$.

同理可得向量组 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$ 可由向量组 $\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ 线性表出的充要条件是 $\mathrm{rank}(\boldsymbol{B})=\mathrm{rank}(\boldsymbol{B},\boldsymbol{A})$, 而 $\mathrm{rank}(\boldsymbol{A},\boldsymbol{B})=\mathrm{rank}(\boldsymbol{B},\boldsymbol{A})$. 结论得证.

上例说明两个向量组是否可以相互线性表示, 完全由它们对应矩阵的秩来确定. 这就是问题本质的揭示.

例 1.5.2　设 $\boldsymbol{A}$ 是三阶矩阵, 求满足 $\boldsymbol{A}^2=\boldsymbol{A}$ 的全体非零矩阵 $\boldsymbol{A}$.

解　由于 $\boldsymbol{A}^2=\boldsymbol{A}\Rightarrow\boldsymbol{A}(\boldsymbol{A}-\boldsymbol{I})=\boldsymbol{O}\Rightarrow\mathrm{rank}(\boldsymbol{A})+\mathrm{rank}(\boldsymbol{A}-\boldsymbol{I})\leqslant 3$. 又

$$\mathrm{rank}(\boldsymbol{A})+\mathrm{rank}(\boldsymbol{A}-\boldsymbol{I})=\mathrm{rank}(\boldsymbol{A})+\mathrm{rank}(\boldsymbol{I}-\boldsymbol{A})\geqslant\mathrm{rank}(\boldsymbol{A}+\boldsymbol{I}-\boldsymbol{A})=3,$$

所以

$$\mathrm{rank}(\boldsymbol{A})+\mathrm{rank}(\boldsymbol{A}-\boldsymbol{I})=3.$$

1° 若 $\mathrm{rank}(\boldsymbol{A}-\boldsymbol{I})=0$, 则 $\boldsymbol{A}=\boldsymbol{I}$;

2° 若 $\mathrm{rank}(\boldsymbol{A}-\boldsymbol{I})=1$, 则 $\boldsymbol{A}-\boldsymbol{I}=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}(b_1,b_2,b_3)$ (定理 1.2.1).

记 $\boldsymbol{B} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} (b_1, b_2, b_3)$, 则

$$\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{I}.$$

由于

$$\begin{aligned} \boldsymbol{A}^2 &= \boldsymbol{B}^2 + 2\boldsymbol{B} + \boldsymbol{I} = (a_1b_1 + a_2b_2 + a_3b_3)\boldsymbol{B} + 2\boldsymbol{B} + \boldsymbol{I} \\ &= (a_1b_1 + a_2b_2 + a_3b_3 + 1)\boldsymbol{B} + \boldsymbol{A}, \end{aligned}$$

所以

$$(a_1b_1 + a_2b_2 + a_3b_3 + 1)\boldsymbol{B} = \boldsymbol{O} \Rightarrow a_1b_1 + a_2b_2 + a_3b_3 + 1 = 0.$$

满足题设条件的全体矩阵为

$$\boldsymbol{A} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} (b_1, b_2, b_3) + \boldsymbol{I}, \quad a_1b_1 + a_2b_2 + a_3b_3 = -1;$$

3° 若 $\operatorname{rank}(\boldsymbol{A} - \boldsymbol{I}) = 2$, 则 $\operatorname{rank}(\boldsymbol{A}) = 1$, 满足题设条件的全体矩阵为

$$\boldsymbol{A} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} (b_1, b_2, b_3), \quad a_1b_1 + a_2b_2 + a_3b_3 = 1.$$

4° 若 $\operatorname{rank}(\boldsymbol{A} - \boldsymbol{I}) = 3$, 则 $\operatorname{rank}(\boldsymbol{A}) = 0$, $\boldsymbol{A} = \boldsymbol{O}$ 不合题意.

例 1.5.3 讨论空间三个平面

$$\pi_1 : a_1x + b_1y + c_1z = d_1, \quad \pi_2 : a_2x + b_2y + c_2z = d_2, \quad \pi_3 : a_3x + b_3y + c_3z = d_3$$

的位置关系.

解 记 $\boldsymbol{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \boldsymbol{\alpha}_3 \end{pmatrix}$, $\overline{\boldsymbol{A}} = \begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$, 则三个平面的位置关系完全由矩阵 $\boldsymbol{A}$ 与 $\overline{\boldsymbol{A}}$ 的秩来确定.

(i) 若 $\operatorname{rank}(\boldsymbol{A}) = 3$.

这时必有 $\operatorname{rank}(\overline{\boldsymbol{A}}) = 3$, 则三平面确定的线性方程组有唯一解, 所以三个平面相交于一点 (见图 1.5.1(1)).

(ii) 若 $\operatorname{rank}(\boldsymbol{A}) = 2, \operatorname{rank}(\overline{\boldsymbol{A}}) = 3$.

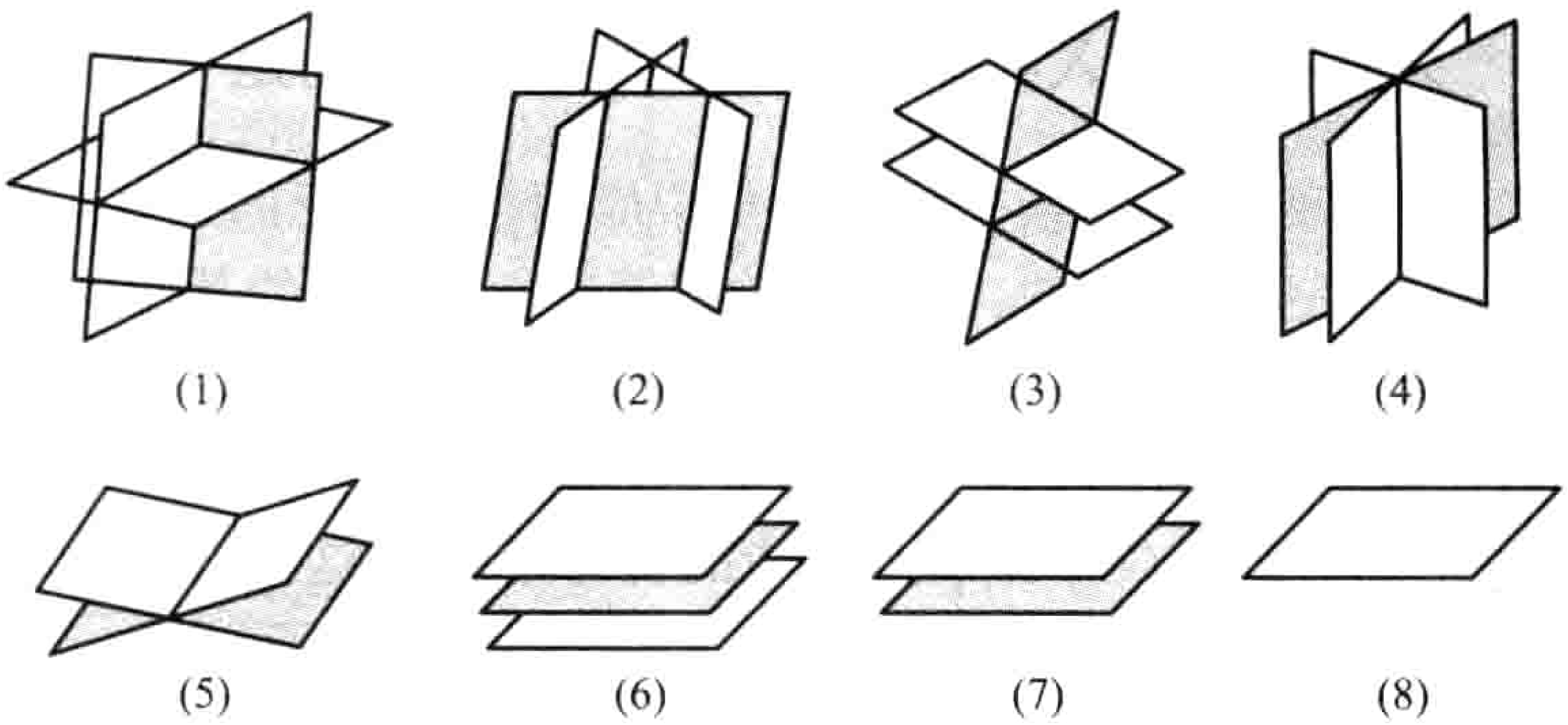


图 1.5.1

这时三平面所确定的线性方程组无解, 三平面无公共交点. 由于 $\boldsymbol{A}$ 中三个行向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ (分别是三个平面的法向量) 是线性相关的 (共面), 当它们中两两线性无关时, 三平面两两相交形成一个三棱柱 (见图 1.5.1(2)); 当三向量中有两个线性相关时, 则这两个都必与另一个线性无关, 此时三平面中有两平面平行, 另一平面与这两平面相交 (见图 1.5.1(3)).

(iii) 若 $\operatorname{rank}(\boldsymbol{A}) = 2, \operatorname{rank}(\overline{\boldsymbol{A}}) = 2$.

这时三平面所确定的线性方程组有解, 且解中含有一个参数, 故三平面相交于一条直线. 当 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 两两线性无关时, 三平面互异 (见图 1.5.1(4)); 当 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 中有两个线性相关时, 三平面中有两个重合 (见图 1.5.1(5)).

(iv) 若 $\operatorname{rank}(\boldsymbol{A}) = 1, \operatorname{rank}(\overline{\boldsymbol{A}}) = 2$.

这时三平面所确定的线性方程组无解, 三平面无公共交点. 由于 $\operatorname{rank}(\boldsymbol{A}) = 1$, 故三平面平行. 由于 $\operatorname{rank}(\overline{\boldsymbol{A}}) = 2$, 故三平面中至少有两平面互异 (见图 1.5.1(6), (7)).

(v) 若 $\operatorname{rank}(\boldsymbol{A}) = 1, \operatorname{rank}(\overline{\boldsymbol{A}}) = 1$.

这时三平面所确定的线性方程组有解, 且解中含有两个参数, 即解集合构成一张平面, 也就是说三平面重合 (见图 1.5.1(8))

三平面的位置关系有且仅有上述八种不同情况.

矩阵秩的等价描述主要源于下面结论:

定理 1.5.1　矩阵行向量组的秩等于其列向量组的秩, 等于矩阵的非零子式的最大阶数.

所以在现行教材中, 矩阵的秩有用其非零子式的最大阶数来定义的, 有用其等价行阶梯形矩阵中非零行的行数来定义的, 也有用等价行阶梯形矩阵中的主元列 (行非零首元素所对应的列) 数来定义的. 不同的定义方式蕴含着不同的处理方法, 也显现出不同教材的风格与特色.

问题 1.6 为什么要引入初等矩阵? 它的主要作用是什么

我们知道矩阵的初等变换是线性代数中应用十分广泛的运算, 许多问题需要通过矩阵的初等变换来解决, 但矩阵的初等变换与矩阵的其他运算具有怎样的联系? 这自然是人们很关心的问题. 初等矩阵刚好起到了这样一个桥梁作用, 它将矩阵的初等变换与矩阵的乘法运算有机地联系了起来 (矩阵的一次初等行(列) 变换等价于矩阵左 (右) 乘一个对应的初等矩阵). 这种联系为人们认识问题、解决问题开阔了视野.

一方面, 矩阵的初等变换可由矩阵的等式 (乘法) 来表示, 这使得由初等变换所得到的结论有了精准的刻画.

另一方面, 矩阵乘法中的某些问题又可通过矩阵的初等变换来得到较方便的解决. 比如矩阵的等价、秩的相关理论, 求逆矩阵的方法, 矩阵的乘积分解 (见问题 1.10), 行列式某些性质的证明, 等等. 这些都常借助初等矩阵来实现矩阵的乘积与初等变换的相互转换.

再则, 初等矩阵形式简单, 且有很多良好的性质, 比如: 任意可逆矩阵都可表示成一系列初等矩阵的乘积, 初等矩阵对逆运算、转置运算封闭 (初等矩阵的逆矩阵与转置矩阵还是初等矩阵), 这就使得许多与矩阵的逆和转置运算有关的问题都可以借助于初等矩阵来研究.

例 1.6.1 证明: 初等变换不改变矩阵的秩.

证 利用行列式的性质, 该结论是不难证明的 (一般教材都有证明). 如果将矩阵的初等变换用矩阵乘积来表示, 则证明就更为简单. 下面仅证初等行变换的情况.

设 $\boldsymbol{A} \xrightarrow{\text{行变换}} \boldsymbol{B}$, 则存在初等矩阵 $\boldsymbol{E}_1, \boldsymbol{E}_2, \cdots, \boldsymbol{E}_t$, 使

$$\boldsymbol{A} = \boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_t\boldsymbol{B}.$$

记 $\boldsymbol{P} = \boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_t$, 则 $\boldsymbol{P}$ 可逆, 且 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}$, 从而

$$\operatorname{rank}(\boldsymbol{A}) \leqslant \operatorname{rank}(\boldsymbol{B}).$$

由于 $\boldsymbol{P}$ 可逆, 又有 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}$, 从而

$$\operatorname{rank}(\boldsymbol{B}) \leqslant \operatorname{rank}(\boldsymbol{A}),$$

所以

$$\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{B}).$$

例 1.6.2 将矩阵 $\boldsymbol{A}=\begin{pmatrix}1&2&0\\0&2&2\\1&1&3\end{pmatrix}$ 表示为若干初等矩阵的乘积.

解 易知 $\boldsymbol{A}$ 可逆, 作初等变换将 $\boldsymbol{A}$ 化为单位矩阵, 并记录每次所作的变换, 再用初等矩阵的乘积表示出来.

$$\boldsymbol{A}=\begin{pmatrix}1&2&0\\0&2&2\\1&1&3\end{pmatrix}\xrightarrow{r_3-r_1}\begin{pmatrix}1&2&0\\0&2&2\\0&-1&3\end{pmatrix}\xrightarrow{\frac{1}{2}r_2}\begin{pmatrix}1&2&0\\0&1&1\\0&-1&3\end{pmatrix}\xrightarrow{r_3+r_2}\begin{pmatrix}1&2&0\\0&1&1\\0&0&4\end{pmatrix}$$
$$\xrightarrow{\frac{1}{4}r_3}\begin{pmatrix}1&2&0\\0&1&1\\0&0&1\end{pmatrix}\xrightarrow{r_2-r_3}\begin{pmatrix}1&2&0\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow{r_1-2r_2}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

以上过程用初等矩阵表示为

$$\boldsymbol{E}_{12}(-2)\boldsymbol{E}_{23}(-1)\boldsymbol{E}_3\left(\frac{1}{4}\right)\boldsymbol{E}_{32}(1)\boldsymbol{E}_2\left(\frac{1}{2}\right)\boldsymbol{E}_{31}(-1)\boldsymbol{A}=\boldsymbol{I},$$

其中 $\boldsymbol{E}_{ij}(c)$ 是单位矩阵的第 j 行乘数 c 加到第 i 行所成的初等矩阵, $\boldsymbol{E}_i(c)$ 是单位矩阵的第 i 行乘数 c 所成的初等矩阵. 则

$$\begin{aligned}\boldsymbol{A}&=[\boldsymbol{E}_{31}(-1)]^{-1}\left[\boldsymbol{E}_2\left(\frac{1}{2}\right)\right]^{-1}[\boldsymbol{E}_{32}(1)]^{-1}\left[\boldsymbol{E}_3\left(\frac{1}{4}\right)\right]^{-1}[\boldsymbol{E}_{23}(-1)]^{-1}[\boldsymbol{E}_{12}(-2)]^{-1}\\&=\boldsymbol{E}_{31}(1)\boldsymbol{E}_2(2)\boldsymbol{E}_{32}(-1)\boldsymbol{E}_3(4)\boldsymbol{E}_{23}(1)\boldsymbol{E}_{12}(2).\end{aligned}$$

即

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

以上分解显然不是唯一的. 如果不借助矩阵的初等变换, 我们很难找出一个正确的分解.

例 1.6.3 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵, 证明: $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$.

证 (1) 当 $\boldsymbol{A}$ 为初等矩阵时, 利用初等矩阵行列式的值及行列式的初等变换性质, 易验证等式成立.

(2) 当 $\boldsymbol{A}$ 不是初等矩阵时, 将 $\boldsymbol{A}$ 作初等行变换化为行最简形矩阵 $\boldsymbol{R}$, 即存在初等矩阵 $\boldsymbol{E}_1,\boldsymbol{E}_2,\cdots,\boldsymbol{E}_t$, 使

$$\boldsymbol{A}=\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_t\boldsymbol{R}.$$

利用 (1) 的结果

$$|\boldsymbol{AB}| = |\boldsymbol{E}_1 \cdots \boldsymbol{E}_t \boldsymbol{RB}| = |\boldsymbol{E}_1| \cdots |\boldsymbol{E}_t| \cdot |\boldsymbol{RB}|.$$

若 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{R} = \boldsymbol{I}$, 有

$$|\boldsymbol{AB}| = |\boldsymbol{E}_1| \cdots |\boldsymbol{E}_t| \cdot |\boldsymbol{B}| = |\boldsymbol{E}_1 \cdots \boldsymbol{E}_t| \cdot |\boldsymbol{B}| = |\boldsymbol{A}||\boldsymbol{B}|,$$

等式成立;

若 $\boldsymbol{A}$ 不可逆, 则 $\boldsymbol{R}$ 的最后一行全为零, $\boldsymbol{RB}$ 的最后一行也全为零. 则 $|\boldsymbol{AB}| = 0$, 且 $|\boldsymbol{A}| = 0$, 此时有

$$|\boldsymbol{AB}| = 0 = |\boldsymbol{A}||\boldsymbol{B}|,$$

等式也成立.

综上, 命题得证.

我们知道, 等式 $|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$ 的证明方法有多种, 上面的证明不失为一种较好的方法, 从中我们可得到许多借鉴. 比如, 我们可类似证明 $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}|$, 还可类似讨论有关矩阵秩的一些问题, 等等.

问题 1.7 如何理解矩阵等价中的三条基本性质? 矩阵的等价标准形有何意义

(1) 矩阵等价中的三条基本性质

线性代数中讨论了矩阵变换中的三种关系: 等价、相似、合同. 其中“等价”是最基本的, 也是最为普遍的关系. 三种关系都满足以下三条基本性质 (以等价为例):

1° 反身性 $\boldsymbol{A} \sim \boldsymbol{A}$;

2° 对称性 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{B} \sim \boldsymbol{A}$;

3° 传递性 若 $\boldsymbol{A} \sim \boldsymbol{B}, \boldsymbol{A} \sim \boldsymbol{C}$, 则 $\boldsymbol{A} \sim \boldsymbol{C}$.

事实上在数学或逻辑学领域, 通常将满足以上三条性质的关系统称为“等价”关系. 如数学运算中的等号“=”, 各种可逆映射, 命题中的充分必要条件, 几何学中的图形“相似”、“全等”、可逆变换, 等等. 所有这些都具有上面三条共性, 它们都属于“等价”的范畴, 所以等价关系是一种十分普遍的关系. 矩阵的等价 (初等变换前后的矩阵) 是“等价”关系中的一种具体形式.

三条性质中最基本的是 2°、3° 两条, 若这两条满足必有第 1° 条也满足, 原因如下:

由对称性: 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{B} \sim \boldsymbol{A}$; 再由传递性得: $\boldsymbol{A} \sim \boldsymbol{B} \sim \boldsymbol{A}$, 即 $\boldsymbol{A} \sim \boldsymbol{A}$.

对称性表明, 等价关系是可逆的.

传递性的作用在于, 通过一步步的等价连接, 使得所关心的两个问题产生联系, 揭示出等价关系下的共性; 另一方面, 通过等价传递使形式复杂的变得形式简单 (标准形), 以利于规律的发现与问题的解决 (见下面 (2)).

教学中, 这些研究问题的基本思想与方法需要教师明确地传递给学生.

(2) 矩阵等价标准形的意义

等价标准形是矩阵等价关系中最简单的形式, 讨论标准形通常有以下几方面的意义:

1° 解决分类问题. 即从等价的角度看, 不同的矩阵有多少类 (等价矩阵归为同一类)?

2° 找到同类矩阵的共性 (有相同的秩).

3° 由于标准形简单, 便于分析、处理, 利用它容易找到解决相关问题的一般性方法.

在前面的例 1.3.1 中我们已看到了矩阵等价标准形的应用, 下面再举两个例子:

例 1.7.1 (矩阵的秩 1 和分解) 若矩阵 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A}) = r$, 则 $\boldsymbol{A}$ 可分解为 r 个秩为 1 的矩阵之和.

分析 秩为 r 的 $m \times n$ 矩阵的标准形为 $\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$ ($\boldsymbol{I}_r$ 是 r 阶单位矩阵), 容易将其作秩 1 和分解

$$\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} = \boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots + \boldsymbol{E}_r,$$

其中 $\boldsymbol{E}_i\,(i = 1, 2, \cdots, r)$ 是第 i 个主对角元是 1, 其余元全为零的 $m \times n$ 矩阵. 这样任意秩为 r 的 $m \times n$ 矩阵 $\boldsymbol{A}$ 的秩 1 和分解也就解决了, 即有

$$\begin{aligned} \boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix} \boldsymbol{Q} &= \boldsymbol{P}(\boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots + \boldsymbol{E}_r)\boldsymbol{Q} \\ &= \boldsymbol{P}\boldsymbol{E}_1\boldsymbol{Q} + \boldsymbol{P}\boldsymbol{E}_2\boldsymbol{Q} + \cdots + \boldsymbol{P}\boldsymbol{E}_r\boldsymbol{Q}, \end{aligned}$$

其中 $\boldsymbol{P}, \boldsymbol{Q}$ 分别是 m 阶与 n 阶满秩矩阵.

类似我们会想到矩阵的满秩分解 (定理 1.2.1).

因为秩为 r 的 $m\times n$ 矩阵 $\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$ 容易作满秩分解:

$$\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}(\boldsymbol{I}_r\ \boldsymbol{O}).$$

由此得到一般秩为 r 的 $m\times n$ 矩阵的满秩分解:

$$\boldsymbol{A}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}(\boldsymbol{I}_r\ \boldsymbol{O})\boldsymbol{Q}\triangleq\boldsymbol{BC},$$

其中 $\boldsymbol{B}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix},\boldsymbol{C}=(\boldsymbol{I}_r\ \boldsymbol{O})\boldsymbol{Q}$ 分别是列满秩与行满秩矩阵 ($\boldsymbol{P},\boldsymbol{Q}$ 分别是 m 阶与 n 阶满秩矩阵).

利用等价标准形还容易完成对线性非齐次方程组有解的充要条件的证明.

例 1.7.2 设 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$, 试证: 线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解的充要条件是 $\mathrm{rank}(\boldsymbol{A},\boldsymbol{b})=\mathrm{rank}(\boldsymbol{A})$.

分析 设 $\mathrm{rank}(\boldsymbol{A})=r$, 若 $\boldsymbol{A}$ 就是标准形 $\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$, 则方程组为 $\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{x}=\begin{pmatrix}\boldsymbol{c}\\ \boldsymbol{d}\end{pmatrix}$, 结论是显然的. 若 $\boldsymbol{A}$ 不是标准形, 则可转化为标准形来处理.

证 设 $\mathrm{rank}(\boldsymbol{A})=r$, 则存在满秩矩阵 $\boldsymbol{P},\boldsymbol{Q}$ 使

$$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}.$$

令 $\boldsymbol{y}=\boldsymbol{Q}^{-1}\boldsymbol{x},\boldsymbol{Pb}=\begin{pmatrix}\boldsymbol{c}'\\ \boldsymbol{d}'\end{pmatrix}$, 于是

$$\boldsymbol{Ax}=\boldsymbol{b}\text{ 有解}\Leftrightarrow\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{y}=\begin{pmatrix}\boldsymbol{c}'\\ \boldsymbol{d}'\end{pmatrix}\text{ 有解}\Leftrightarrow\boldsymbol{d}'=\boldsymbol{0}.$$

又

$$\mathrm{rank}(\boldsymbol{A},\boldsymbol{b})=\mathrm{rank}\boldsymbol{P}(\boldsymbol{A},\boldsymbol{b})\begin{pmatrix}\boldsymbol{Q} & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}=\mathrm{rank}(\boldsymbol{PAQ},\boldsymbol{Pb})=\mathrm{rank}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O} & \boldsymbol{c}'\\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{d}'\end{pmatrix},$$

所以

$$\boldsymbol{Ax}=\boldsymbol{b}\text{ 有解}\Leftrightarrow\boldsymbol{d}'=\boldsymbol{0}\Leftrightarrow\mathrm{rank}(\boldsymbol{A},\boldsymbol{b})=r=\mathrm{rank}(\boldsymbol{A}).$$

问题 1.8 矩阵分块有何意义? 分块要注意哪些问题

矩阵分块是我们在相关计算或问题的处理中经常采用的方法, 它的意义主要体现在以下两个方面.

(1) 降阶, 使计算得到简化

我们知道矩阵的阶数越高, 相关计算越困难. 矩阵分块以后, 将每个子块当作一个元, 就可达到降阶的效果, 使运算得到简化. 这种为简化计算而分块的情况主要用在矩阵的乘法、逆、行列式运算等方面, 其中乘法是应用最为广泛也是最为基础的运算.

1° 作乘法运算.

分块要注意两点: 一是左侧矩阵列的划分要与右侧矩阵行的划分一致, 这样分块才能相乘; 二是要将一些特殊形式的子块 (如零矩阵、数量矩阵、对角矩阵、三角形矩阵等) 划分出来.

例如, 计算两矩阵相乘 $\boldsymbol{AB}$, 其中

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

按如下形式分块会使计算更为容易

$$\boldsymbol{A}=\left(\begin{array}{cc:ccc}1&0&0&0&0\\0&1&0&0&0\\\hdashline-1&2&1&0&0\\1&1&0&1&0\\-2&0&0&0&1\end{array}\right)=\begin{pmatrix}\boldsymbol{I}_2&\boldsymbol{O}\\\boldsymbol{A}_1&\boldsymbol{I}_3\end{pmatrix},$$

$$\boldsymbol{B}=\left(\begin{array}{ccc:cc}3&2&0&1&0\\1&3&0&0&1\\\hdashline-1&0&0&0&0\\0&-1&0&0&0\\0&0&-1&0&0\end{array}\right)=\begin{pmatrix}\boldsymbol{B}_1&\boldsymbol{I}_2\\-\boldsymbol{I}_3&\boldsymbol{O}\end{pmatrix}.$$

2° 求逆运算.

分块偏重于准主 (副) 对角分块或分块三角形. 常见的分块逆公式有

$$\begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}_1^{-1} & & & \\ & \boldsymbol{A}_2^{-1} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^{-1} \end{pmatrix},$$

$$\begin{pmatrix} & & & \boldsymbol{A}_1 \\ & & \boldsymbol{A}_2 & \\ & \iddots & & \\ \boldsymbol{A}_k & & & \end{pmatrix}^{-1} = \begin{pmatrix} & & & \boldsymbol{A}_k^{-1} \\ & & \boldsymbol{A}_{k-1}^{-1} & \\ & \iddots & & \\ \boldsymbol{A}_1^{-1} & & & \end{pmatrix} \quad (\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k \text{ 均可逆}),$$

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{C}^{-1} \\ \boldsymbol{O} & \boldsymbol{C}^{-1} \end{pmatrix} \quad (\boldsymbol{A}, \boldsymbol{C} \text{ 均可逆}).$$

3° 作行列式计算.

分块要便于利用行列式的性质及 Laplace (拉普拉斯, 1749—1827) 展开定理 (见问题 2.6, 2.7).

(2) 分割, 使问题得到转化

矩阵的分块相当于对矩阵的一种 "解剖", 这种 "解剖" 有助于看清整体与局部的关系, 从而使问题得到转化, 找到解决问题的有效方法.

例如, 利用分块的矩阵我们可以得到求逆矩阵的方法:

设矩阵 $\boldsymbol{A}$ 可逆, 按分块矩阵的乘法

$$\boldsymbol{A}^{-1}(\boldsymbol{A}, \boldsymbol{I}) = (\boldsymbol{I}, \boldsymbol{A}^{-1}).$$

由于 $\boldsymbol{A}^{-1}$ 可表示为一系列初等矩阵的乘积, 所以我们得到求逆矩阵的方法

$$(\boldsymbol{A}, \boldsymbol{I}) \xrightarrow{\text{初等行变换}} (\boldsymbol{I}, \boldsymbol{A}^{-1}).$$

又如, 利用矩阵的分块 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)$, 我们将线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 写成了向量的线性组合

$$x_1\boldsymbol{\alpha}_1 + \cdots + x_n\boldsymbol{\alpha}_n = \boldsymbol{b},$$

这就使得矩阵、向量、线性方程组构成了一个有机的整体, 使很多问题得以相互转换, 有效解决.

再如, 由于矩阵的分块才使矩阵的相似对角化与矩阵的特征值、特征向量建立了联系, 即

$$\begin{aligned} \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} &= \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \\ &\Leftrightarrow \boldsymbol{A}(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n) = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \end{aligned}$$

$$
\begin{aligned}
&(\text{记 } \boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n))\\
\Leftrightarrow\ &(\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \cdots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \cdots, \lambda_n\boldsymbol{p}_n)\\
\Leftrightarrow\ &\boldsymbol{A}\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i\ (i = 1, 2, \cdots, n).
\end{aligned}
$$

利用矩阵分块来处理问题的方法贯穿了线性代数的始终, 可以毫不夸张地说, 没有矩阵的分块, 线性代数的知识构架将很难建立!

问题 1.9 如何将矩阵的初等变换与初等矩阵推广到分块矩阵

我们知道矩阵初等变换与矩阵的分块在矩阵的运算与理论研究中有十分重要的作用, 一种很自然的想法, 是希望将矩阵的初等变换推广到分块矩阵上, 并且使推广后的"变换"能与通常矩阵的初等变换等价 (在初等变换下互化).

事实上, 这种想法是可行的, 仿照通常矩阵的初等变换可定义分块矩阵的初等变换如下:

定义 1.9.1 设 $\boldsymbol{M}$ 是一个分块矩阵, 称以下三种形式的变换叫分块矩阵的初等变换:

1° 交换 $\boldsymbol{M}$ 某两块行 (列) 的位置;

2° 用可逆矩阵左 (右) 乘 $\boldsymbol{M}$ 的某一块行 (列);

3° 用非零矩阵左 (右) 乘 $\boldsymbol{M}$ 的某一块行 (列) 加到另一块行 (列) 上.

容易验证上面定义与通常矩阵的初等变换是等价的.

为将分块矩阵的初等变换与分块矩阵的乘法联系起来, 也可引入分块初等矩阵的概念:

定义 1.9.2 将单位矩阵按准对角形式分块, 叫分块单位矩阵. 分块单位矩阵经过一次初等变换得到的矩阵称为**分块初等矩阵**.

值得注意的是, 分块初等矩阵可能不再是通常意义下的初等矩阵. 例如

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

是一个分块初等矩阵, 但它不是初等矩阵.

与普通矩阵一样, **分块矩阵的行 (列) 初等变换可通过左 (右) 乘对应的初等矩阵来实现**. 下面以 2×2 分块矩阵为例来具体说明:

设 $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}$, 其中 $\boldsymbol{A}_{11}$ 是 $m_1\times n_1$ 矩阵, $\boldsymbol{A}_{22}$ 是 $m_2\times n_2$ 矩阵, 则有

1° $\begin{pmatrix} O & I_{m_2} \\ I_{m_1} & O \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}=\begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{pmatrix}$;

2° $\begin{pmatrix} C & O \\ O & I_{m_2} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}=\begin{pmatrix} CA_{11} & CA_{12} \\ A_{21} & A_{22} \end{pmatrix}$ (C 是可逆矩阵);

3° $\begin{pmatrix} I_{m_1} & O \\ P & I_{m_2} \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}=\begin{pmatrix} A_{11} & A_{12} \\ PA_{11}+A_{21} & PA_{12}+A_{22} \end{pmatrix}$.

可类似讨论列初等变换的情况.

分块矩阵的初等变换能给我们的计算带来更多的方便, 在行列式的计算中会更显出它的优越性 (见问题 2.7). 下面举两个例子:

例 1.9.1 设 A, C 分别为 r 与 $n-r$ 阶可逆矩阵, 用初等变换求 $M=\begin{pmatrix} O & A \\ C & B \end{pmatrix}$ 的逆矩阵.

解

$$\begin{aligned}(M\,\vdots\, I)=&\left(\begin{array}{cc:cc} O & A & I_r & O \\ C & B & O & I_{n-r} \end{array}\right)\xrightarrow{r_1\leftrightarrow r_2}\left(\begin{array}{cc:cc} C & B & O & I_{n-r} \\ O & A & I_r & O \end{array}\right)\\ &\xrightarrow{r_1-BA^{-1}r_2}\left(\begin{array}{cc:cc} C & O & -BA^{-1} & I_{n-r} \\ O & A & I_r & O \end{array}\right)\\ &\xrightarrow{C^{-1}r_1,A^{-1}r_2}\left(\begin{array}{cc:cc} I_{n-r} & O & -C^{-1}BA^{-1} & C^{-1} \\ O & I_r & A^{-1} & O \end{array}\right),\end{aligned}$$

所以

$$M^{-1}=\begin{pmatrix} -C^{-1}BA^{-1} & C^{-1} \\ A^{-1} & O \end{pmatrix}.$$

以上步骤用分块初等矩阵的乘法来描述就是

$$\begin{pmatrix} I_{n-r} & O \\ O & A^{-1} \end{pmatrix}\begin{pmatrix} C^{-1} & O \\ O & I_r \end{pmatrix}\begin{pmatrix} I_{n-r} & -BA^{-1} \\ O & I_r \end{pmatrix}\begin{pmatrix} O & I_{n-r} \\ I_r & O \end{pmatrix}\begin{pmatrix} O & A \\ C & B \end{pmatrix}=\begin{pmatrix} I_{n-r} & O \\ O & I_r \end{pmatrix}.$$

由于

$$\begin{pmatrix} I_{n-r} & O \\ O & A^{-1} \end{pmatrix}\begin{pmatrix} C^{-1} & O \\ O & I_r \end{pmatrix}\begin{pmatrix} I_{n-r} & -BA^{-1} \\ O & I_r \end{pmatrix}\begin{pmatrix} O & I_{n-r} \\ I_r & O \end{pmatrix}=\begin{pmatrix} -C^{-1}BA^{-1} & C^{-1} \\ A^{-1} & O \end{pmatrix},$$

所以

$$\begin{pmatrix} O & A \\ C & B \end{pmatrix}^{-1}=\begin{pmatrix} -C^{-1}BA^{-1} & C^{-1} \\ A^{-1} & O \end{pmatrix}$$

例 1.9.2 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 皆是 n 阶方阵, 证明 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$.

证 在例 1.6.3 已经有过证明, 这里用分块矩阵的初等变换来做.

首先, 考虑下述分块矩阵

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}. \tag{1.9.1}$$

显然该矩阵的行列式等于 $|\boldsymbol{A}||\boldsymbol{B}|$, 为了便于与 $|\boldsymbol{AB}|$ 联系起来, 可将 (1.9.1) 式中的第二行的 $\boldsymbol{A}$ 倍加到第一行上, 即

$$\begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{AB} \\ -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}. \tag{1.9.2}$$

因 $\begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}$ 可由 $\begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}$ 经过 n^2 次行初等 (倍加) 变换而得到, 即

$$\begin{pmatrix} \boldsymbol{I} & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}=\boldsymbol{E}_{n^2}\boldsymbol{E}_{n^2-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1\begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}=\boldsymbol{E}_{n^2}\boldsymbol{E}_{n^2-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1.$$

其中 $\boldsymbol{E}_i\ (i=1,\cdots,n^2)$ 是初等倍加矩阵, 则 (1.9.2) 式为

$$\boldsymbol{E}_{n^2}\boldsymbol{E}_{n^2-1}\cdots\boldsymbol{E}_2\boldsymbol{E}_1\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{AB} \\ -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}. \tag{1.9.3}$$

上式两边取行列式, 并注意倍加变换不改变行列式的值, 有

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{B} \end{vmatrix}=\begin{vmatrix} \boldsymbol{O} & \boldsymbol{AB} \\ -\boldsymbol{I} & \boldsymbol{B} \end{vmatrix}.$$

用 Laplace 展开定理, 有

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{B} \end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{B}|,$$

$$\begin{vmatrix} \boldsymbol{O} & \boldsymbol{AB} \\ -\boldsymbol{I} & \boldsymbol{B} \end{vmatrix}=(-1)^{n^2}|\boldsymbol{AB}|\cdot|-\boldsymbol{I}|=(-1)^{n^2+n}|\boldsymbol{AB}|=|\boldsymbol{AB}|.$$

所以 $|\boldsymbol{AB}|=|\boldsymbol{A}||\boldsymbol{B}|$.

由于分块矩阵的初等变换同样不改变矩阵的秩, 所以我们也可用分块矩阵的初等变换来探求一些有关矩阵秩的问题.

例 1.9.3 设 $\boldsymbol{A}$ 是 r 阶可逆矩阵, 分块矩阵 $\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}$.

1° 试证明: $\operatorname{rank}(\boldsymbol{M})=r+\operatorname{rank}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})$;

2° 设 $\boldsymbol{M}=\begin{pmatrix}1 & 0 & 3 & 2 & 1\\ 4 & 2 & 6 & -4 & 18\\ -2 & 6 & -8 & -8 & -4\\ 1 & -6 & 5 & 6 & 3\\ 2 & -4 & 6 & 4 & 6\end{pmatrix}$, 求 $\operatorname{rank}(\boldsymbol{M})$.

证 1° 由于分块矩阵的初等变换不改变矩阵的秩, 所以

$$\begin{aligned}\operatorname{rank}(\boldsymbol{M})&=\operatorname{rank}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}=\operatorname{rank}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{pmatrix}\\ &=\operatorname{rank}(\boldsymbol{I}_r)+\operatorname{rank}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})=r+\operatorname{rank}(\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}).\end{aligned}$$

解 2° 将 $\boldsymbol{M}$ 按如下方式分块

$$\boldsymbol{M}=\left(\begin{array}{cc:ccc}1 & 0 & 3 & 2 & 1\\ 4 & 2 & 6 & -4 & 18\\ \hdashline -2 & 6 & -8 & -8 & -4\\ 1 & -6 & 5 & 6 & 3\\ 2 & -4 & 6 & 4 & 6\end{array}\right)=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}.$$

作行初等变换, 求 $\boldsymbol{A}^{-1}\boldsymbol{B}$:

$$(A\,\vdots\,B)=\left(\begin{array}{cc:ccc}1 & 0 & 3 & 2 & 1\\ 4 & 2 & 6 & -4 & 18\end{array}\right)\to\left(\begin{array}{cc:ccc}1 & 0 & 3 & 2 & 1\\ 0 & 1 & -3 & -6 & 7\end{array}\right),$$

得

$$\boldsymbol{A}^{-1}\boldsymbol{B}=\begin{pmatrix}3 & 2 & 1\\ -3 & -6 & 7\end{pmatrix},$$

$$\begin{aligned}\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}&=\begin{pmatrix}-8 & -8 & -4\\ 5 & 6 & 3\\ 6 & 4 & 6\end{pmatrix}-\begin{pmatrix}-2 & 6\\ 1 & -6\\ 2 & -4\end{pmatrix}\begin{pmatrix}3 & 2 & 1\\ -3 & -6 & 7\end{pmatrix}\\ &=\begin{pmatrix}16 & 32 & -44\\ -16 & -32 & 44\\ -12 & -24 & 32\end{pmatrix},\end{aligned}$$

所以

$$\operatorname{rank}(\boldsymbol{M}) = 2 + \operatorname{rank}(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}) = 2 + 2 = 4.$$

例 1.9.3 中 1° 的结论又称为矩阵**秩的第一降阶定理**. 它将高阶矩阵的秩转化成了较低阶矩阵的秩, 该结论的意义主要体现在理论推导上. 比如用它可以很容易得到矩阵秩的 **Sylvesterl (西尔维斯特, 1814—1897) 不等式**:

$$\operatorname{rank}(\boldsymbol{AB}) \geqslant \operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B}) - n, \tag{1.9.4}$$

其中 $\boldsymbol{A}, \boldsymbol{B}$ 分别是 $m \times n$ 与 $n \times l$ 矩阵.

其原因为

$$\begin{aligned}\operatorname{rank}(\boldsymbol{AB}) &= \operatorname{rank}[\boldsymbol{O} - (-\boldsymbol{A})\boldsymbol{I}_n\boldsymbol{B}] = \operatorname{rank}\begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{B} \\ -\boldsymbol{A} & \boldsymbol{O} \end{pmatrix} - n = \operatorname{rank}\begin{pmatrix} \boldsymbol{B} & \boldsymbol{I}_n \\ \boldsymbol{O} & -\boldsymbol{A} \end{pmatrix} - n \\ &\geqslant \operatorname{rank}(\boldsymbol{B}) + \operatorname{rank}(-\boldsymbol{A}) - n = \operatorname{rank}(\boldsymbol{B}) + \operatorname{rank}(\boldsymbol{A}) - n.\end{aligned}$$

问题 1.10 什么是矩阵的三角分解 ($\boldsymbol{LU}$ 分解), 有何应用

矩阵 $\boldsymbol{A}$ 的分解是指将 $\boldsymbol{A}$ 表示成两个或多个矩阵乘积的形式. 由于矩阵乘法涉及数据合成 (通过两个或多个线性变换的作用复合成一个矩阵), 所以矩阵分解是一种数据分解. 在计算机语言中, 将 $\boldsymbol{A}$ 表示成矩阵乘积的形式相当于对 $\boldsymbol{A}$ 中的数据进行预处理, 即把 $\boldsymbol{A}$ 中的数据分成两个或多个部分, 使其结构在某种意义上更有用, 而且更易于计算.

矩阵的三角 ($\boldsymbol{LU}$) 分解是指将一个 $m \times n$ 矩阵 $\boldsymbol{A}$ 分解成如下形式

$$\boldsymbol{A} = \boldsymbol{LU},$$

其中 $\boldsymbol{L}$ 是一个 $m \times m$ 下三角形矩阵, $\boldsymbol{U}$ 是一个行阶梯形矩阵. 当 $\boldsymbol{A}$ 为方阵时, $\boldsymbol{U}$ 就是一个上三角形矩阵.

将矩阵作三角分解的方法有多种, 下面介绍由矩阵的初等变换化阶梯形所得到的分解方法.

设 $\boldsymbol{A}$ 是一个 $m \times n$ 矩阵, 并且不使用行交换就可以化为阶梯形式 $\boldsymbol{U}$. 将所作的初等行变换表示为初等矩阵左乘 $\boldsymbol{A}$, 即

$$\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A} = \boldsymbol{U}.$$

记变换矩阵 $\boldsymbol{L} = (\boldsymbol{E}_k \cdots \boldsymbol{E}_2\boldsymbol{E}_1)^{-1}$, 则

$$\boldsymbol{A} = \boldsymbol{LU}.$$

由于 $\boldsymbol{E}_i\ (i=1,\cdots,k)$ 均是对角线元全为 1 的 $m\times m$ 下三角形矩阵, 所以 $\boldsymbol{L}$ 是一个对角线元全为 1 的 $m\times m$ 下三角形矩阵.

例 1.10.1 设 $\boldsymbol{A}=\begin{pmatrix}3&-7&-2&2\\-3&5&1&0\\6&-4&0&-5\\-9&5&-5&12\end{pmatrix}$, 将矩阵 $\boldsymbol{A}$ 作三角分解.

解 作初等行变换将 $\boldsymbol{A}$ 化为阶梯形, 并记录出每一步所作的变换:

$$\boldsymbol{A}\xrightarrow[r_4+3r_1]{\substack{r_2+r_1\\ r_3-2r_1}}\begin{pmatrix}3&-7&-2&2\\0&-2&-1&2\\0&10&4&-9\\0&-16&-11&18\end{pmatrix}\xrightarrow[r_4-8r_2]{r_3+5r_2}\begin{pmatrix}3&-7&-2&2\\0&-2&-1&2\\0&0&-1&1\\0&0&-3&2\end{pmatrix}$$

$$\xrightarrow{r_4-3r_3}\begin{pmatrix}3&-7&-2&2\\0&-2&-1&2\\0&0&-1&1\\0&0&0&-1\end{pmatrix}\triangleq\boldsymbol{U}, \tag{1.10.1}$$

即

$$\boldsymbol{E}_{43}(-3)\boldsymbol{E}_{42}(-8)\boldsymbol{E}_{32}(5)\boldsymbol{E}_{41}(3)\boldsymbol{E}_{31}(-2)\boldsymbol{E}_{21}(1)\boldsymbol{A}=\boldsymbol{U}.$$

记变换矩阵

$$\begin{aligned}\boldsymbol{L}&=(\boldsymbol{E}_{43}(-3)\boldsymbol{E}_{42}(-8)\boldsymbol{E}_{32}(5)\boldsymbol{E}_{41}(3)\boldsymbol{E}_{31}(-2)\boldsymbol{E}_{21}(1))^{-1}\\&=\boldsymbol{E}_{21}(-1)\boldsymbol{E}_{31}(2)\boldsymbol{E}_{41}(-3)\boldsymbol{E}_{32}(-5)\boldsymbol{E}_{42}(8)\boldsymbol{E}_{43}(3)\\&=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\2&0&1&0\\-3&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&-5&1&0\\0&8&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&3&1\end{pmatrix}\\&=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\2&-5&1&0\\-3&8&3&1\end{pmatrix},\end{aligned}\tag{1.10.2}$$

得到

$$\boldsymbol{A}=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\2&-5&1&0\\-3&8&3&1\end{pmatrix}\begin{pmatrix}3&-7&-2&2\\0&-2&-1&2\\0&0&-1&1\\0&0&0&-1\end{pmatrix}=\boldsymbol{LU}.$$

注意, 等式 (1.10.2) 中的 3 个矩阵完全可由初等变换式 (1.10.1) 中的 3 个步骤直接写出.

从上面化矩阵为阶梯形的过程中可以看到, 当途中某个矩阵的主对角元为零时, 要继续化阶梯形就需要对矩阵作行位置交换. 这相当于左乘一个对换矩阵 $\boldsymbol{E}(i,j)$ (i,j 两行对换), 由此得到的变换矩阵 $\boldsymbol{L}$ 就不再是一个下三角形矩阵了, 因此不是任何一个矩阵都能够作三角分解. 下面是矩阵能作三角分解的一个充分条件:

定理 1.10.1　若 n 阶矩阵 $\boldsymbol{A}=(a_{ij})$ 的所有 k 阶 $(1\leqslant k\leqslant n-1)$ 顺序主子式不为零, 则 $\boldsymbol{A}$ 能进行 $\boldsymbol{LU}$ 分解.

注　$\boldsymbol{A}$ 的 k 阶顺序主子式是指 k 阶行列式 $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$.

证　对 $\boldsymbol{A}$ 的阶用归纳法.

当 $n=1$ 时, $\boldsymbol{A}=(a_{11})=(1)(a_{11})\triangleq \boldsymbol{L}_1\boldsymbol{U}_1$, 定理成立.

设 $\boldsymbol{A}$ 为 $n-1$ 阶矩阵时, 定理成立; 当 $\boldsymbol{A}$ 为 n 阶矩阵时, 将 $\boldsymbol{A}$ 分块为

$$\begin{pmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{t} \\ \boldsymbol{u} & a_{nn} \end{pmatrix},$$

其中 $\boldsymbol{t}=(a_{1n},a_{2n},\cdots,a_{n-1,n})^{\mathrm{T}}, \boldsymbol{u}=(a_{n1},a_{n2},\cdots,a_{n,n-1})$. 由归纳假设

$$\boldsymbol{A}_{n-1}=\boldsymbol{L}_{n-1}\boldsymbol{U}_{n-1},$$

于是

$$\begin{pmatrix} \boldsymbol{A}_{n-1} & \boldsymbol{t} \\ \boldsymbol{u} & a_{nn} \end{pmatrix}=\begin{pmatrix} \boldsymbol{L}_{n-1} & \boldsymbol{0} \\ \boldsymbol{l} & 1 \end{pmatrix}\begin{pmatrix} \boldsymbol{U}_{n-1} & \boldsymbol{v} \\ \boldsymbol{0} & w_{nn} \end{pmatrix},$$

则

$$\boldsymbol{A}_{n-1}=\boldsymbol{L}_{n-1}\boldsymbol{U}_{n-1},\quad \boldsymbol{t}=\boldsymbol{L}_{n-1}\boldsymbol{v},\quad \boldsymbol{u}=\boldsymbol{l}\boldsymbol{U}_{n-1},\quad a_{nn}=\boldsymbol{l}\boldsymbol{v}+w_{nn}.$$

由于 $\boldsymbol{A}$ 的 $n-1$ 阶顺序主子式不为 0, 有 $\boldsymbol{A}_{n-1}$ 可逆, 则 $\boldsymbol{L}_{n-1}$ 与 $\boldsymbol{U}_{n-1}$ 均可逆. 因此, 从上面式子中可以确定唯一的 $\boldsymbol{l}$ 和 $\boldsymbol{v}$. 所以 $\boldsymbol{A}$ 的 $\boldsymbol{LU}$ 分解是存在的.

一般情况下, 在对矩阵 $\boldsymbol{A}$ 作初等行变换化为阶梯形 $\boldsymbol{U}$ 的过程中, 若要实施若干次行位置交换, 即左乘若干个类型 $\boldsymbol{E}(i,j)$ (对换矩阵) 的初等矩阵 $\boldsymbol{E}_1,\boldsymbol{E}_2,\cdots,\boldsymbol{E}_l$. 记 $\boldsymbol{P}=\boldsymbol{E}_l\boldsymbol{E}_{l-1}\cdots\boldsymbol{E}_1$ (称 $\boldsymbol{P}$ 为**置换矩阵**, 即单位矩阵经过若干次行交换得到的矩阵), 则有三角分解

$$\boldsymbol{PA}=\boldsymbol{LU}.$$

易验证 $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{I}$, 所以

$$\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{L}\boldsymbol{U}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{U},$$

上式又称为 $\boldsymbol{A}$ 的 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{U}$ 分解.

可以证明: **任何 n 阶可逆矩阵 $\boldsymbol{A}$ 总有一个 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{U}$ 分解, 当 $\boldsymbol{P}$ 确定后, $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 都是唯一的**.

矩阵的 $\boldsymbol{L}\boldsymbol{U}$ 分解或 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{U}$ 分解都可以由计算机编程来计算, 读者可查阅相关书籍, 这里不再讨论.

矩阵的 $\boldsymbol{L}\boldsymbol{U}$ 分解有很多应用, 特别是在科学计算方面. 比如, 若 $\boldsymbol{A}$ 是稀疏矩阵 (大多数元是零), 则 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 也多为稀疏的, 而 $\boldsymbol{A}^{-1}$ 可能是稠密的. 如果我们的问题需要用 $\boldsymbol{A}^{-1}$ 去解决, 这时对 $\boldsymbol{A}$ 作 $\boldsymbol{L}\boldsymbol{U}$ 分解是一种很好的方法, 并且三角形矩阵的逆矩阵也很容易计算.

再如, 解线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, 若 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$, 则方程组等价于以下两个方程组

$$\boldsymbol{L}\boldsymbol{y}=\boldsymbol{b},\quad \boldsymbol{U}\boldsymbol{x}=\boldsymbol{y},$$

而以上两个方程组都是三角状系数方程组, 求解将十分方便.

在计算行列式时, 因为 $\det(\boldsymbol{A})=\det(\boldsymbol{L})\det(\boldsymbol{U})$, 而三角形矩阵的行列式就是其对角线元的乘积, 显然也是很方便的.

第二章　行列式

问题 2.1　行列式的历史沿革

行列式的概念最早是由 17 世纪日本数学家关孝和 (约 1642—1708) 提出来的, 他在 1683 年写了一部叫做《解伏题之法》的著作, 意思是 “解行列式问题的方法”, 书里对行列式的概念和它的展开已经有了清楚的叙述. 同时代的 Leibnitz (莱布尼茨, 1646—1716) 是欧洲第一个提出行列式概念的人. 他在 1693 年 4 月写给 L'Hospital (洛必达, 1661—1704) 的一封信中使用并给出了行列式, 还给出方程组的系数行列式为零的条件.

1750 年, 瑞士数学家 G. Cramer (克拉默, 1704—1752) 在其著作《线性代数分析导言》中, 对行列式的定义和展开法则给出了比较完整、明确的阐述, 并给出了现在我们所称的解线性方程组的 Cramer 法则. 稍后, 法国数学家 É. Bézout (贝祖, 1739—1783) 将确定行列式每一项符号的方法进行了系统化, 利用系数行列式概念指出了如何判断一个 n 个未知量的 n 个齐次线性方程组有非零解的方法, 即系数行列式等于零是这方程组有非零解的条件.

在很长一段时间内, 行列式只是作为解线性方程组的一种工具使用, 并没有人意识到它可以独立于线性方程组之外单独形成一门理论加以研究.

在行列式的发展史上, 第一个对行列式理论作出连贯的逻辑的阐述, 即把行列式理论与线性方程组求解相分离的人, 是法国数学家 Vandermonde (范德蒙, 1735—1796), 他给出了用二阶子式和它们的余子式来展开行列式的法则. 就对行列式来说, 他是这门理论的奠基人. 1772 年, Laplace (拉普拉斯, 1749—1826) 在一篇论文中证明了 Vandermonde 提出的一些规则, 推广了他的展开行列式的方法. 用 r 行中所含的子式和它们的余子式的集合来展开行列式, 这个方法现在仍然以他的名字命名.

继 Vandermonde 之后, 在行列式的理论方面, 又一位做出突出贡献的就是法国大数学家 Cauchy (柯西, 1789—1857), 1812 年他发表论文, 用行列式给出了某些多面体体积的计算公式, 并且将这些公式与先前行列式的研究结果联系起来. Cauchy 所发现的行列式在解析几何中的应用激起了人们研究行列式应用的浓厚兴趣, 前后持续了近 100 年. 1815 年, Cauchy 在一篇论文中给出了行列式的第一个系统的、几乎是近代的处理. 其中主要结果之一是行列式的乘法定理. 另外, 他第一个把行列式的元排成方阵, 采用双足标记法; 引进了行列式特征方程的术语; 给出了相似行列式概念; 改进了 Laplace 的行列式展开定理并给出了一个证明.

继 Cauchy 之后, 在行列式理论方面最多产的人就是德国数学家 Jacobi (雅可比, 1804—1851), 他引进了函数行列式, 即 " Jacobi 行列式", 指出函数行列式在多重积分的变量替换中的作用, 给出了函数行列式的导数公式. Jacobi 的著名论文《论行列式的形成和性质》标志着行列式系统理论的建立. 由于行列式在数学分析、几何学、线性方程组理论、二次型理论等多方面的应用, 促使行列式理论自身在 19 世纪也得到了很大发展. 整个 19 世纪都有行列式的新结果. 除了一般行列式的大量定理之外, 还有许多有关特殊行列式的其他定理都相继得到.

问题 2.2 行列式有哪些不同的定义方式

虽然行列式的概念最初是由线性方程组的求解而提出来的, 但从其发展来看, 它的意义远不在求解线性方程组上. 为揭示行列式的基本规律, 人们所选择的视角可能就有所不同, 这主要体现在不同的定义方式上. 目前较常见的有以下三种:

(1) "逆序数法" 定义

"逆序数法" 定义又称为行列式的完全展开定义, 具体表述为

定义 2.2.1 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶方阵, 定义行列式

$$|\boldsymbol{A}|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n},$$

其中 $j_1j_2\cdots j_n$ 遍取 $1,2,\cdots,n$ 的全部 $n!$ 个排列, $\tau(j_1j_2\cdots j_n)$ 是排列 $j_1j_2\cdots j_n$ 的逆序数.

(2) "归纳法" 定义

"归纳法" 定义又称按行展开法定义, 该定义直接给出了行列式的计算式 (展开式).

定义 2.2.2　设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶方阵, 定义行列式 $|\boldsymbol{A}|$ 是按下列法则确定的一个数:

1° 当 $n=1$ 时 $|a_{11}|=a_{11}$;

2° 当 $n \geqslant 2$ 时, $|\boldsymbol{A}|=a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n}$, 其中 $A_{1j}\ (j=1,2,\cdots,n)$ 是 $\boldsymbol{A}$ 中元 a_{1j} 的代数余子式.

(3) "函数法" 定义

"函数法" 定义又称公理化定义, 该定义将行列式归结为满足 3 条性质的线性映射, 而不是直接给出行列式的具体形式.

定义 2.2.3　设 F 是数域, $\boldsymbol{\alpha}_j=(a_{1j},a_{2j}\ldots,a_{nj})^{\mathrm{T}}\in \boldsymbol{F}^n, 1\leqslant j\leqslant n$. 若有 $\boldsymbol{F}^n$ 上的 n 元函数, 即映射

$$f:\boldsymbol{F}^n\times\boldsymbol{F}^n\times\cdots\times\boldsymbol{F}^n\to\boldsymbol{F}$$

满足下列条件:

1° 反对称性　$\forall 1<i<j<n, f(\cdots,\boldsymbol{\alpha}_i,\cdots,\boldsymbol{\alpha}_j,\cdots)=-f(\cdots,\boldsymbol{\alpha}_j,\cdots,\boldsymbol{\alpha}_i,\cdots)$;

2° 线性性　$f(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=f(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)+f(\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$,

$$f(k\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=kf(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),k\in\boldsymbol{F};$$

3° 规范性　对 $\boldsymbol{F}^n$ 的自然基底 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n, f(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=1$.

则称 f 是 (数域 $\boldsymbol{F}$ 上的) n 阶行列式函数.

以上定义各不相同, 但可以证明它们都是彼此等价的. 当然其等价性的证明也并非一件易事, 读者可参见《线性代数与矩阵论》(许以超, 高等教育出版社, 1992) 与《高等代数与解析几何》(牛兴文, 化学工业出版社, 2005) 的相关章节. 下面仅对 (1) 与 (2) 之间的等价性给以简要的证明.

例 2.2.1　n 阶行列式的归纳法定义与逆序数法定义是等价的.

证　对矩阵阶数 n 用数学归纳法:

当 $n=1,2$ 时容易验证两种定义式的等价性;

假设 $n=k-1$ 时两种定义式仍然等价;

当 $n=k$ 时, 由定义 2.2.1

$$|\boldsymbol{A}|=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1k}\\a_{21}&a_{22}&\cdots&a_{2k}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kk}\end{vmatrix}=\sum_{j_1j_2\cdots j_k}(-1)^{\tau(j_1j_2\cdots j_k)}a_{1j_1}a_{2j_2}\cdots a_{kj_k}$$

$$
\begin{aligned}
&= a_{11}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=1}}(-1)^{\tau(j_1j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}+\\
&\quad a_{12}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=2}}(-1)^{\tau(j_1j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}+\cdots+\\
&\quad a_{1k}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=k}}(-1)^{\tau(j_1j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}\\
&= a_{11}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=1}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{nj_n}+\\
&\quad a_{12}(-1)^1\sum_{\substack{j_1j_2\cdots j_k\\ j_1=2}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}+\cdots+\\
&\quad a_{1k}(-1)^{k-1}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=k}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}\\
&= a_{11}(-1)^{1+1}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=1}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{nj_n}+\\
&\quad a_{12}(-1)^{1+2}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=2}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}+\cdots+\\
&\quad a_{1k}(-1)^{1+k}\sum_{\substack{j_1j_2\cdots j_k\\ j_1=k}}(-1)^{\tau(j_2\cdots j_k)}a_{2j_2}\cdots a_{kj_k}\\
&= a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1k}A_{1k},
\end{aligned}
$$

这便是定义 2.2.2 中的表示式. 这说明两种定义式对一切自然数均等价.

从教学的角度来看, 以上不同方式的定义各有利弊. 定义 2.2.1 容易使初学者理解行列式定义的来源, 掌握 n 阶行列式的特征 (共有 $n!$ 项, 每项取自矩阵不同行不同列的 n 个元的乘积), 容易利用定义作简单计算, 便于行列式理论的系统证明, 但排列的 "逆序数" 常常使初学者难以理解. 定义 2.2.2 避开了 "逆序数", 简洁易懂、易于计算, 但难以看出 n 阶行列式的特征及定义的来源, 在行列式个别性质的证明中会有一定的障碍. 定义 2.2.3 揭示了行列式的本质特征, 符合现代数学的观点, 便于行列式的性质推导及一般性理论研究, 适用面广, 但抽象难懂, 在具化到行列式的表达式时, 同样避免不了逆序数的概念与性质.

例 2.2.2 设函数行列式 $D(x)=\begin{vmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x)\\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x)\\ \vdots & \vdots & & \vdots\\ a_{n1}(x) & a_{n2}(x) & \cdots & a_{nn}(x)\end{vmatrix}$ 中各元均

是 x 的可微函数, 求其导函数 $D'(x)$.

解法 1　因为 $D(x)=\sum\limits_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}a_{1j_1}(x)a_{2j_2}(x)\cdots a_{nj_n}(x)$, 由导数的运算法则

$$\begin{aligned}
D'(x)&=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}[a_{1j_1}(x)a_{2j_2}(x)\cdots a_{nj_n}(x)]'\\
&=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}[a'_{1j_1}(x)a_{2j_2}(x)\cdots a_{nj_n}(x)+\\
&\quad a_{1j_1}(x)a'_{2j_2}(x)\cdots a_{nj_n}(x)+\cdots+a_{1j_1}(x)a_{2j_2}(x)\cdots a'_{nj_n}(x)]\\
&=\begin{vmatrix} a'_{11}(x) & a'_{12}(x) & \cdots & a'_{1n}(x)\\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x)\\ \vdots & \vdots & & \vdots\\ a_{n1}(x) & a_{n2}(x) & \cdots & a_{nn}(x)\end{vmatrix}+\begin{vmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x)\\ a'_{21}(x) & a'_{22}(x) & \cdots & a'_{2n}(x)\\ \vdots & \vdots & & \vdots\\ a_{n1}(x) & a_{n2}(x) & \cdots & a_{nn}(x)\end{vmatrix}+\cdots+\\
&\quad\begin{vmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x)\\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x)\\ \vdots & \vdots & & \vdots\\ a'_{n1}(x) & a'_{n2}(x) & \cdots & a'_{nn}(x)\end{vmatrix}
\end{aligned}\tag{2.2.1}$$

上面结果表明: n 阶函数行列式的导数等于逐行求导所得的 n 个行列式之和. 由于转置运算不改变行列式的值, 所以逐行求导也可以换为逐列求导.

解法 2　用数学归纳法证明 (2.2.1) 式成立.

当 $n=1$ 时, (2.2.1) 式显然成立. 假设 $n-1$ 时 (2.2.1) 式成立.

当 $D(x)$ 的阶数为 n 时, 将行列式按第 1 行展开为

$$D(x)=a_{11}(x)A_{11}(x)+a_{12}(x)A_{12}(x)+\cdots+a_{1n}(x)A_{1n}(x),\tag{2.2.2}$$

其中 $A_{1j}(x)$ $(j=1,2,\cdots,n)$ 是 $D(x)$ 中元 $a_{1j}(x)$ 的代数余子式.

对 (2.2.2) 式求导得

$$\begin{aligned}
D'(x)=&[a'_{11}(x)A_{11}(x)+a'_{12}(x)A_{12}(x)+\cdots+a'_{1n}(x)A_{1n}(x)]+\\
&[a_{11}(x)A'_{11}(x)+a_{12}(x)A'_{12}(x)+\cdots+a_{1n}(x)A'_{1n}(x)].
\end{aligned}\tag{2.2.3}$$

容易看出 (2.2.3) 中前一括号中的表达式就是 (2.2.1) 中的第 1 个行列式, 而 (2.2.3) 中后一括号中的表达式, 利用归纳假设, 则是 (2.2.1) 中的第 2 到第 n 个行列式的和.

解法 3 记

$$D(x)=f(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2(x),\cdots,\boldsymbol{\alpha}_n(x)),\tag{2.2.4}$$

其中 $\boldsymbol{\alpha}_j(x)$ 是 $D(x)$ 中第 j 个列向量.

先考察行列式 $f(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 的导数 (其中 $\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 与 x 无关). 由导数的定义及行列式的性质可得

$$\begin{aligned}f'(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)&=\lim_{\Delta x\to 0}\frac{1}{\Delta x}[f(\boldsymbol{\alpha}_1(x+\Delta x),\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)-f(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)]\\&=\lim_{\Delta x\to 0}f\left(\frac{\boldsymbol{\alpha}_1(x+\Delta x)-\boldsymbol{\alpha}_1(x)}{\Delta x},\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\right)\\&=f(\boldsymbol{\alpha}_1'(x),\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n).\end{aligned}$$

一般有

$$\begin{aligned}&f'(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{j-1},\boldsymbol{\alpha}_j(x),\boldsymbol{\alpha}_{j+1},\cdots,\boldsymbol{\alpha}_n)=f(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{j-1},\boldsymbol{\alpha}_j'(x),\boldsymbol{\alpha}_{j+1},\cdots,\boldsymbol{\alpha}_n)\\&\qquad(j=1,2,\cdots,n).\end{aligned}\tag{2.2.5}$$

将 (2.2.4) 写成多元复合函数

$$D(x)=f(\boldsymbol{\alpha}_1(x_1),\boldsymbol{\alpha}_2(x_2),\cdots,\boldsymbol{\alpha}_n(x_n)),\quad x_1=x_2=\cdots=x_n=x,$$

利用多元复合函数的全导数公式, 及 (2.2.5) 可得

$$\begin{aligned}D'(x)&=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\frac{\mathrm{d}x_j}{\mathrm{d}x}\\&=f(\boldsymbol{\alpha}_1'(x),\boldsymbol{\alpha}_2(x),\cdots,\boldsymbol{\alpha}_n(x))+f(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2'(x),\cdots,\boldsymbol{\alpha}_n(x))+\cdots+\\&\quad f(\boldsymbol{\alpha}_1(x),\boldsymbol{\alpha}_2(x),\cdots,\boldsymbol{\alpha}_n'(x)).\end{aligned}\tag{2.2.6}$$

以上三种解法可分别视为在定义 2.2.1、定义 2.2.2、定义 2.2.3 基础上演绎的结果, 解法 3 具有普适性, 因为无论哪种定义方式, 推导中所用的知识都是具备的. 当然, 解法 2 中的方法在不知道求导结果的情况下是难以想到的.

问题 2.3 行列式有何几何意义

行列式的几何意义有两种解释:

一种解释是行列式是其行或列向量所构成的超平行多面体的**有向面积**或**有向体积**; 另一种解释是矩阵 $\boldsymbol{A}$ 的行列式 $\det\boldsymbol{A}$ 就是在线性变换 $\mathcal{A}:\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 下的图形面积或体积的**伸缩因子**. 两种说法一个是静态的体积概念, 一个是动态的变换比例概念, 它们都有共同的几何本质. 下面我们来做具体的讨论.

(1) 超平行多面体的有向面积或体积

从静态的角度看, 行列式是以其行或列向量为这 (棱) 所构成的超平行多面体的有向面积或有向体积.

先考察二阶行列式的情况.

设 $D_2=\det(\boldsymbol{a},\boldsymbol{b})$, 它的两个列向量为 $\boldsymbol{a}=(a_1,a_2)^{\mathrm{T}},\boldsymbol{b}=(b_1,b_2)^{\mathrm{T}}$, 不妨设它们就是 $\mathbf{R}^2$ 中自然基底 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$ 下的坐标, 即有

$$\boldsymbol{a}=a_1\boldsymbol{\varepsilon}_1+a_2\boldsymbol{\varepsilon}_2,\quad \boldsymbol{b}=b_1\boldsymbol{\varepsilon}_1+b_2\boldsymbol{\varepsilon}_2,$$

作向量积, 有

$$\boldsymbol{a}\times\boldsymbol{b}=(a_1b_2-a_2b_1)\boldsymbol{\varepsilon}_3=D_2\boldsymbol{\varepsilon}_3 \tag{2.3.1}$$

其中 $\boldsymbol{\varepsilon}_3=\boldsymbol{\varepsilon}_1\times\boldsymbol{\varepsilon}_2$ 是垂直于 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$ 所在平面的单位向量, $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 构成右手系.

(2.3.1) 式两边取模可得 $|D_2|=\|\boldsymbol{a}\times\boldsymbol{b}\|$, 也就是说 D_2 的绝对值是以 $\boldsymbol{a},\boldsymbol{b}$ 为邻边的平行四边形面积.

下面我们引入有向面积的概念: 若平行四边形是由向量 $\boldsymbol{a}$ 沿逆 (顺) 时针方向转到 $\boldsymbol{b}$ 而得到的, 则面积取正 (负) 值, 所以平面上平行四边形有两种定向. 若 $\boldsymbol{a},\boldsymbol{b}$ 与 $\boldsymbol{a}',\boldsymbol{b}'$ 张成的平行四边形的有向面积符号相同, 称两个向量组有相同的定向.

由 (2.3.1) 式可看出, 二阶行列式 D_2 就是它的两个列 (或行) 向量 $\boldsymbol{a},\boldsymbol{b}$ 为邻边的平行四边形的有向面积. 当 $\boldsymbol{a},\boldsymbol{b}$ 与 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$ 定向相同时, D_2 的值为正; $\boldsymbol{a},\boldsymbol{b}$ 与 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2$ 定向相反时, D_2 为负; $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 共线时, D_2 的值为 0.

对三阶行列式, 有

$$\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}. \tag{2.3.2}$$

我们知道混合积 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$ 在几何上就是 $\mathbf{R}^3$ 中以 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 为棱的平行六面体的有向体积. 这里空间平行六面体也有两种定向: 当 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 构成右手系时 (此时 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 与 $\mathbf{R}^3$ 的自然基底 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3$ 定向相同), 体积取正值; 当 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 构成左手系时, 体积取负值. 显然, 平行六面体体积为零的充分必要条件是 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 共面.

有序向量组 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 的定向还有另一种理解方式: 右手四个指头从向量 $\boldsymbol{a}$ 的方向转过一个 $0\sim\pi$ 之间的角度到达向量 $\boldsymbol{b}$ 的方向, 拇指方向是向量 $\boldsymbol{a},\boldsymbol{b}$ 所确定平面的正向, 如果向量 $\boldsymbol{c}$ 在此平面正向的一侧, 有序向量组 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 就确定了一个正的定向; 如果向量 $\boldsymbol{c}$ 在平面正向的另一侧, $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 就确定了一个负的定向. 与向量的向量积联系起来讲, 第三个向量 $\boldsymbol{c}$ 与向量 $\boldsymbol{a}\times\boldsymbol{b}$ 在 $\boldsymbol{a},\boldsymbol{b}$ 所确定平面的同一侧时, 有序向量组 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 确定一个正的定向; 否则, 确定了一个负的定向. 当 $\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})>0$ 时, 意味着 $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}>0$. 也就是说向量 $\boldsymbol{a}\times\boldsymbol{b}$ 的方向与向量 $\boldsymbol{c}$

的方向成锐角, 此时 $\boldsymbol{c}$ 与 $\boldsymbol{a}\times\boldsymbol{b}$ 在 $\boldsymbol{a},\boldsymbol{b}$ 所确定平面的同一侧, 所以 $\det(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$ 的符号就确定了有序向量组 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 的定向.

当 $n>3$ 时, 我们可以类似地把 n 阶行列式视为由它的 n 个列 (或行) 向量为棱所构成的超平行多面体的有向体积. 行列式的符号同样反映了组成它的 n 个列 (或行) 向量有序组的定向关系.

(2) 线性变换下图形面积或体积的伸缩因子

从动态的角度看, 矩阵 $\boldsymbol{A}$ 的行列式是线性变换 $\mathcal{A}:\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 下图形面积或体积的伸缩因子 (由于 $\mathcal{A}$ 由矩阵 $\boldsymbol{A}$ 唯一确定, 在下面的叙述中将不再对 $\mathcal{A}$ 与 $\boldsymbol{A}$ 作严格区分).

设 n 阶行列式 D_n 所对应的矩阵为 $\boldsymbol{A}$, 即 $D_n=\det(\boldsymbol{A})=\det(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$. 由 $\boldsymbol{A}$ 可确定一个线性变换 $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$. 设 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 是 $\mathbf{R}^n$ 中的自然基底. 由矩阵的乘法知

$$\boldsymbol{\alpha}_i=\boldsymbol{A}\boldsymbol{\varepsilon}_i\quad(i=1,2,\cdots,n).\tag{2.3.3}$$

当 $n=2$ 时, (2.3.3) 说明 $\boldsymbol{A}$ 将 $\mathbf{R}^2$ 中的单位正方形变成了 $\mathbf{R}^2$ 中的以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 为邻边的平行四边形; 如果原图形为一个圆, 则线性变换 $\boldsymbol{A}$ 将之变成一个椭圆.

同样, 在 $n=3$ 时, $\boldsymbol{A}$ 将 $\mathbf{R}^3$ 中的一个单位立方体变成了 $\mathbf{R}^3$ 中由 $\boldsymbol{A}$ 的列向量确定的平行六面体; 如果原图形为一个球, 则线性变换 $\boldsymbol{A}$ 将之变成一个椭球.

一般地, 一个 n 阶矩阵 $\boldsymbol{A}$ 将 $\mathbf{R}^n$ 中的超单位立方体变成 $\mathbf{R}^n$ 中由 $\boldsymbol{A}$ 列向量确定的 n 维超平行多面体.

作为 (2.3.3) 式更一般的情形, 设 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 是 $\mathbf{R}^n$ 中的一个线性无关列向量组, 记

$$\boldsymbol{\gamma}_i=\boldsymbol{A}\boldsymbol{\beta}_i\quad(i=1,2,\cdots,n),\tag{2.3.4}$$

将 (2.3.4) 写成矩阵的形式, 有

$$(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n)=\boldsymbol{A}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n).\tag{2.3.5}$$

两边取行列式, 得

$$\det(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n)=\det(\boldsymbol{A})\cdot\det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n).\tag{2.3.6}$$

由此可见

$$\det(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n)/\det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=\det(\boldsymbol{A}).$$

即

$$\frac{\text{像域的体积}}{\text{原域的体积}}=\det(\boldsymbol{A}).$$

所以 $\det \boldsymbol{A}$ 是在线性变换 $\boldsymbol{A}$ 下的图形面积或体积的伸缩因子.

特别地, 当 $\boldsymbol{\beta}_i = \varepsilon_i\ (i=1,\cdots,n)$ 时, $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=\boldsymbol{I}, \det(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)=1$, 这个伸缩因子就是 $\boldsymbol{A}$ 的列向量所确定的 n 维超平行多面体的体积, 这与 (1) 中的结论是一致的.

例 2.3.1 求曲面 $x^2+5y^2+11z^2+2xy-2xz-6yz=1$ 所围立体的体积.

解 配方, 曲面方程为

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

令 $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, 曲面化为

$$x'^2+y'^2+z'^2=1,$$

这是单位球面, 它所围成的体积 $V'=\dfrac{4}{3}\pi$.

由于所作线性变换的行列式 $\det(\boldsymbol{A})=6$, 所以原曲面所围立体的体积为

$$V=\frac{V'}{\det(\boldsymbol{A})}=\frac{2}{9}\pi.$$

问题 2.4 Cramer 法则的多种证明与几何意义

Cramer 法则给出了线性方程组当变量个数和方程个数相等时, 在方程组的系数行列式不等于零的情况下, 解的公式表达式. 由于其结论简洁优美, 便于作符号运算, 所以被广为流传. 同时, 也由于 Cramer 法则的多种证明方法的产生, 使得该法则更显璀璨. 这里介绍几种不同的证明方法, 以体现数学的精彩与魅力.

(1) Cramer 法则的多种证明

定理 2.4.1 (Cramer 法则) 若方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases} \tag{2.4.1}$$

的系数行列式 $\det(\boldsymbol{A})\neq 0$, 则方程组有唯一解

$$x_i=\frac{\det[\boldsymbol{A}_i(\boldsymbol{b})]}{\det(\boldsymbol{A})},\quad i=1,2,\cdots,n,$$

其中 $\boldsymbol{A}_i(\boldsymbol{b})$ 是以向量 $\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$ 替代 $\boldsymbol{A}$ 中第 i 列所得的矩阵.

证法 1 将方程组 (2.4.1) 写成矩阵形式

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}. \tag{2.4.2}$$

由于 $\det(\boldsymbol{A})\neq 0$, 所以 $\boldsymbol{A}$ 可逆. 方程组 (2.4.2) 两边同时左乘 $\boldsymbol{A}^{-1}$, 得

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}. \tag{2.4.3}$$

易验证 (2.4.3) 的确是方程组 (2.4.2) 的解. 另一方面, 若 $\boldsymbol{x}_1$ 是方程组 (2.4.2) 的另一个解, 则有

$$\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{b}. \tag{2.4.4}$$

(2.4.2) 与 (2.4.4) 两式相减, 有 $\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}_1)=\boldsymbol{0}$, 左乘 $\boldsymbol{A}^{-1}$, 得

$$\boldsymbol{x}-\boldsymbol{x}_1=\boldsymbol{0}.$$

所以解是唯一的.

设 $\boldsymbol{A}$ 的伴随矩阵为

$$\boldsymbol{A}^*=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix},$$

由 (2.4.3) 式

$$\begin{aligned}\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}=\frac{1}{\det(\boldsymbol{A})}\boldsymbol{A}^*\boldsymbol{b}&=\frac{1}{\det(\boldsymbol{A})}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{pmatrix}\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}\\ &=\frac{1}{\det(\boldsymbol{A})}\begin{pmatrix} b_1A_{11}+b_2A_{21}+\cdots+b_nA_{n1}\\ b_1A_{12}+b_2A_{22}+\cdots+b_nA_{n2}\\ \vdots\\ b_1A_{1n}+b_2A_{2n}+\cdots+b_nA_{nn}\end{pmatrix}=\frac{1}{\det(\boldsymbol{A})}\begin{pmatrix}\det[\boldsymbol{A}_1(\boldsymbol{b})]\\ \det[\boldsymbol{A}_2(\boldsymbol{b})]\\ \vdots\\ \det[\boldsymbol{A}_n(\boldsymbol{b})]\end{pmatrix}.\end{aligned}$$

定理得证.

证法 2 先证明 $x_i=\dfrac{\det[\boldsymbol{A}_i(\boldsymbol{b})]}{\det(\boldsymbol{A})}$ $(i=1,2,\cdots,n)$ 是方程组的解, 也就是要验证

$$a_{i1}\frac{\det[\boldsymbol{A}_1(\boldsymbol{b})]}{\det(\boldsymbol{A})}+a_{i2}\frac{\det[\boldsymbol{A}_2(\boldsymbol{b})]}{\det(\boldsymbol{A})}+\cdots+a_{in}\frac{\det[\boldsymbol{A}_n(\boldsymbol{b})]}{\det(\boldsymbol{A})}=b_i\ (i=1,2,\cdots,n),$$

即

$$a_{i1}\det[\boldsymbol{A}_1(\boldsymbol{b})]+a_{i2}\det[\boldsymbol{A}_2(\boldsymbol{b})]+\cdots+a_{in}\det[\boldsymbol{A}_n(\boldsymbol{b})]=b_i\det(\boldsymbol{A})\ (i=1,2,\cdots,n). \tag{2.4.5}$$

为此, 我们作有两行相同的 $n+1$ 阶行列式

$$\begin{vmatrix} b_i & a_{i1} & \cdots & a_{in}\\ b_1 & a_{11} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ b_n & a_{n1} & \cdots & a_{nn}\end{vmatrix}\ (i=1,2,\cdots,n).$$

按第一行展开. 因为第一行第 $j+1$ 列的元 a_{ij} 的代数余子式是

$$\begin{aligned}&(-1)^{1+j+1}\begin{vmatrix} b_1 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots & \vdots & & \vdots\\ b_n & a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn}\end{vmatrix}\\ &=(-1)^{2+j}(-1)^{j-1}\det[\boldsymbol{A}_j(\boldsymbol{b})]=-\det[\boldsymbol{A}_j(\boldsymbol{b})].\end{aligned}$$

又因为这个 $n+1$ 阶行列式是零, 所以有

$$b_i\det(\boldsymbol{A})-a_{i1}\det[\boldsymbol{A}_1(\boldsymbol{b})]-a_{i2}\det[\boldsymbol{A}_2(\boldsymbol{b})]-\cdots-a_{in}\det[\boldsymbol{A}_n(\boldsymbol{b})]=0.$$

上式即为 (2.4.5).

再证解的唯一性.

设 $x_i=c_i$ $(i=1,2,\cdots,n)$ 是方程组的解, 则

$$a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n=b_i\ (i=1,2,\cdots,n).$$

由行列式的性质

$$\begin{aligned}c_1\det(\boldsymbol{A})&=\begin{vmatrix} a_{11}c_1 & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{n1}c_1 & a_{n2} & \cdots & a_{nn}\end{vmatrix}=\begin{vmatrix} a_{11}c_1+a_{12}c_2+\cdots+a_{1n}c_n & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{n1}c_1+a_{n2}c_2+\cdots+a_{nn}c_n & a_{n2} & \cdots & a_{nn}\end{vmatrix}\\ &=\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ b_n & a_{n2} & \cdots & a_{nn}\end{vmatrix}=\det[\boldsymbol{A}_1(\boldsymbol{b})].\end{aligned}$$

一般可得

$$c_i\det(\boldsymbol{A})=\det[\boldsymbol{A}_i(\boldsymbol{b})]\ (i=1,2,\cdots,n),$$

由于 $\det(\boldsymbol{A})\neq 0$, 所以 $c_i=\dfrac{\det[\boldsymbol{A}_i(\boldsymbol{b})]}{\det(\boldsymbol{A})}\ (i=1,2,\cdots,n)$.

证法 3 记 $\boldsymbol{A}$ 的 n 个列向量为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$, 将方程组改写为向量形式

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\ldots+x_n\boldsymbol{a}_n=\boldsymbol{b}. \tag{2.4.6}$$

作向量 $\boldsymbol{A}_i=(A_{1i},A_{2i},\cdots,A_{ni})$, A_{ki} 是 $\boldsymbol{A}$ 中元 a_{ki} 的代数余子式, 用 A_i 与 (2.4.6) 两边的向量作内积可得

$$x_i\boldsymbol{a}_i\cdot A_i=\boldsymbol{b}\cdot A_i,$$

上式即为

$$x_i\det(\boldsymbol{A})=\det[\boldsymbol{A}_i(\boldsymbol{b})]\ (i=1,2,\cdots,n).$$

证法 4 记 $\boldsymbol{A}$ 的 n 个列向量为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$, 单位矩阵 $\boldsymbol{I}$ 的 n 个列向量为 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$. 如果 $\boldsymbol{Ax}=\boldsymbol{b}$, 由分块矩阵的乘法可得

$$\begin{aligned}\boldsymbol{AI}_i(\boldsymbol{x})&=\boldsymbol{A}(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{x},\cdots,\boldsymbol{\varepsilon}_n)=(\boldsymbol{A\varepsilon}_1,\cdots,\boldsymbol{Ax},\cdots,\boldsymbol{A\varepsilon}_n)\\&=(\boldsymbol{a}_1,\cdots,\boldsymbol{b},\cdots,\boldsymbol{a}_n)=\boldsymbol{A}_i(\boldsymbol{b}).\end{aligned}$$

上式两边取行列式, 有

$$\det(\boldsymbol{A})\cdot x_i=\det[\boldsymbol{A}_i(\boldsymbol{b})]\quad(\text{注意 }\det[\boldsymbol{I}_i(\boldsymbol{x})]=x_i).$$

(2) Cramer 法则的几何意义

为便于几何直观, 我们仅对 $n=3$ 的情形作讨论, 一般情形可类比.

Cramer 法则中

$$x_1=\frac{\det(\boldsymbol{b},\boldsymbol{a}_2,\boldsymbol{a}_3)}{\det(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)}. \tag{2.4.7}$$

由问题 2.3 我们知道, (2.4.7) 的分子与分母是两个平行六面体的有向体积之比, 这两个平行六面体具有相同的底面 (以 $\boldsymbol{a}_2,\boldsymbol{a}_3$ 为邻边的平行四边形). 由此我们可进一步看出, x_1 也是向量 $\boldsymbol{b}$ 与 $\boldsymbol{a}_1$ 分别到 $\boldsymbol{a}_2,\boldsymbol{a}_3$ 所张成的平面的有向距离之比, 将 (2.4.7) 式用如下的向量投影形式写出更为清晰

$$x_1=\frac{\boldsymbol{b}\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3)}{\boldsymbol{a}_1\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3)}=\frac{\boldsymbol{b}\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3)^0}{\boldsymbol{a}_1\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3)^0}=\frac{\mathrm{Pr\ j}_{\boldsymbol{a}_2\times\boldsymbol{a}_3}\boldsymbol{b}}{\mathrm{Pr\ j}_{\boldsymbol{a}_2\times\boldsymbol{a}_3}\boldsymbol{a}_1}$$

($\boldsymbol{a}^0$ 表示 $\boldsymbol{a}$ 的单位向量) 当 $\boldsymbol{b}$ 与 $\boldsymbol{a}_1$ 都在 $\boldsymbol{a}_2,\boldsymbol{a}_3$ 所张成的平面的同侧时 (即有序组 $\boldsymbol{b},\boldsymbol{a}_2,\boldsymbol{a}_3$ 与 $\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3$ 有相同的定向) x_1 为正, 在异侧时为负.

x_2,x_3 有完全类似的几何意义.

问题 2.5 行列式的计算有哪些常用方法

行列式的计算是一个困难的问题, 方法很灵活, 几乎无法找出一个通用的方法来计算我们可能遇到的行列式. 计算中通常是利用行列式的性质, 根据行列式元的分布特点及行列式的结构特点进行化简处理.

常用的计算方法有:

1° 定义法 (主要适用于阶数较低的行列式或零元较多的行列式).

2° 降阶法 (利用行列式的初等变换性质使某行或列中的零元增多, 再按零元多的行或列展开, 化高阶为低阶, 有时也可按多行或多列展开, 即作 Laplace 展开).

3° 化为三角形行列式 (主要适用于元都是具体数字或字母元很有规律的行列式).

4° 加边法 (也叫升阶法, 给行列式增加一行与一列, 并保持行列式的值不变, 加边后要便于使用行列式的性质简化行列式).

5° 拆项法 (将行列式的某一行或列的元写成两个数的和, 利用行列式的性质将行列式拆成两个行列式之和, 而所得的两个行列式容易计算, 或得到行列式的递推公式).

6° 递推公式法 (按某行或列展开得到递推公式, 有时需要得到两个不同的递推公式, 再用解方程组的办法求解).

7° 数学归纳法 (从低阶到高阶, 找规律猜结果, 用归纳法证明).

8° 利用矩阵的乘法分解或已知的行列式 (如 Vandermonde 行列式) 作计算.

9° 利用函数行列式的微分性质 (见例 2.2.2) 作计算.

例 2.5.1 计算行列式 $\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{vmatrix}$.

解法 1

$$D = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & b-a & c-a & d-a \\ 0 & b^2-a^2 & c^2-a^2 & d^2-a^2 \\ 0 & b^4-a^4 & c^4-a^4 & d^4-a^4 \end{vmatrix}$$

$$= \begin{vmatrix} b-a & c-a & d-a \\ b^2-a^2 & c^2-a^2 & d^2-a^2 \\ (b^2+a^2)(b^2-a^2) & (c^2+a^2)(c^2-a^2) & (d^2+a^2)(d^2-a^2) \end{vmatrix}$$

$$
\begin{aligned}
&= (b-a)(c-a)(d-a)\begin{vmatrix} 1 & 1 & 1 \\ b+a & c+a & d+a \\ (b^2+a^2)(b+a) & (c^2+a^2)(c+a) & (d^2+a^2)(d+a) \end{vmatrix} \\
&= (b-a)(c-a)(d-a)\cdot \\
&\quad \begin{vmatrix} 1 & 1 & 1 \\ 0 & c-b & d-b \\ 0 & (c^2+a^2)(c+a)-(b^2+a^2)(b+a) & (d^2+a^2)(d+a)-(b^2+a^2)(b+a) \end{vmatrix} \\
&= (b-a)(c-a)(d-a)(c-b)(d-b)\cdot \\
&\quad \begin{vmatrix} 1 & 1 \\ (c^2+bc+b^2)+a(c+b)+a^2 & (d^2+bd+b^2)+a(d+b)+a^2 \end{vmatrix} \\
&= (a-b)(a-c)(a-d)(b-c)(b-d)(c-d)(a+b+c+d).
\end{aligned}
$$

评注 利用行列式的初等变换性质使某行 (列) 中的零元增多, 再按零元多的行 (列) 展开达到降阶的效果, 这是行列式计算中最基本的方法.

解法 2 令 $f(x)=\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ a & b & c & d & x \\ a^2 & b^2 & c^2 & d^2 & x^2 \\ a^3 & b^3 & c^3 & d^3 & x^3 \\ a^4 & b^4 & c^4 & d^4 & x^4 \end{vmatrix}$. 设按第五列展开为

$$
f(x)=A_{15}+A_{25}x+A_{35}x^2+A_{45}x^3+A_{55}x^4,
$$

其中 A_{ij} 是 $f(x)$ 中元 a_{ij} 的代数余子式.

因为 $A_{45}=-D$,

$$
A_{55}=\begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix} \xlongequal[\text{行列式}]{\text{Vandermonde}} (b-a)(c-a)(d-a)(c-b)(d-b)(d-c),
$$

易知 a,b,c,d 是 $f(x)$ 的根, 由 Vieta (韦达, 1540—1603) 定理

$$
a+b+c+d=-\frac{A_{45}}{A_{55}}=\frac{D}{A_{55}},
$$

所以 $D=(a+b+c+d)A_{55}=(a-b)(a-c)(a-d)(b-c)(b-d)(c-d)(a+b+c+d)$.

例 2.5.2 计算 n 阶行列式 $D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}$.

解法 1 化为三角形行列式计算.

将 $i\ (i=2,3,\cdots,n)$ 行全加到第 1 行, 然后提公因子得

$$D_n = \begin{vmatrix} x+(n-1)a & x+(n-1)a & \cdots & x+(n-1)a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}$$

$$= [x+(n-1)a]\begin{vmatrix} 1 & 1 & \cdots & 1 \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n]{r_i - a r_1}[x+(n-1)a]\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & x-a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x-a \end{vmatrix} = [x+(n-1)a](x-a)^{n-1}.$$

也可用以下方法化为三角形

$$D_n \xlongequal[i=2,\cdots,n]{r_i - r_1} \begin{vmatrix} x & a & \cdots & a \\ a-x & x-a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a-x & 0 & \cdots & x-a \end{vmatrix} \xlongequal[i=2,\cdots,n]{c_1 + c_i} \begin{vmatrix} x+(n-1)a & a & \cdots & a \\ 0 & x-a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x-a \end{vmatrix}$$

$$= [x+(n-1)a](x-a)^{n-1}.$$

解法 2 用加边法计算.

$$D_n = \begin{vmatrix} 1 & a & a & \cdots & a \\ 0 & x & a & \cdots & a \\ 0 & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a & a & \cdots & x \end{vmatrix} \xlongequal[i=2,\cdots,n+1]{r_i - r_1} \begin{vmatrix} 1 & a & a & \cdots & a \\ -1 & x-a & 0 & \cdots & 0 \\ -1 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & x-a \end{vmatrix}.$$

若 $x \neq a$, 则

$$D_n = \begin{vmatrix} 1 & a & a & \cdots & a \\ -1 & x-a & 0 & \cdots & 0 \\ -1 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & x-a \end{vmatrix}$$

$$\xlongequal[i=2,\cdots,n+1]{c_1 - \frac{1}{x-a} c_i} \begin{vmatrix} 1 + \dfrac{na}{x-a} & a & a & \cdots & a \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x-a \end{vmatrix}$$

$$= \left(1 + \frac{na}{x-a}\right)(x-a)^n = [x+(n-1)a](x-a)^{n-1}.$$

若 $x = a$, 则 $D_n = 0$.

总之, $D_n = [x+(n-1)a](x-a)^{n-1}$.

评注 只有当行列式对角线以下 (上) 的元与行列式中某行 (列) 的对应元成比例, 或行列式的元均为数字时, 才比较容易化为三角形.

解法 3 利用函数的 Taylor (泰勒, 1685—1731) 公式计算.

由函数行列式的求导公式 (例 2.2.2), 有

$$D_n'(x) = \begin{vmatrix} 1 & a & \cdots & a \\ 0 & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ 0 & a & \cdots & x \end{vmatrix} + \begin{vmatrix} x & 0 & \cdots & a \\ a & 1 & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & 0 & \cdots & x \end{vmatrix} + \cdots + \begin{vmatrix} x & a & \cdots & 0 \\ a & x & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & 1 \end{vmatrix}.$$

将上面 n 个行列式分别按第 $1, 2, \cdots, n$ 列展开得

$$D_n'(x) = nD_{n-1}(x).$$

类推可得

$$D_n''(x) = n(n-1)D_{n-2}(x), \quad \cdots, \quad D_n^{(n-1)}(x) = n!D_1(x) = n!x, \quad D_n^{(n)}(x) = n!.$$

由 Taylor 公式, 得

$$D_n(x) = D_n(a) + D_n'(a)(x-a) + \frac{D_n''(a)}{2!}(x-a)^2 + \cdots + \frac{D_n^{(n)}(a)}{n!}(x-a)^n.$$

注意到 $D_n(a)=D'_n(a)=D_n^{(n-2)}(a)=0$, 得

$$\begin{aligned}D_n(x)&=\frac{D_n^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+\frac{D_n^{(n)}(a)}{n!}(x-a)^n\\&=na(x-a)^{n-1}+(x-a)^n.\end{aligned}$$

评注 解法 3 虽然已不属于线性代数的范畴, 但作适当介绍, 对学生创新思维的培养是很有帮助的.

例 2.5.3 计算 n 阶行列式 $D_n=\begin{vmatrix}a&b&\cdots&b\\c&a&\cdots&b\\\vdots&\vdots&&\vdots\\c&c&\cdots&a\end{vmatrix}(b\neq c)$.

解法 1 拆项法计算.

$$\begin{aligned}D_n&=\begin{vmatrix}c&b&\cdots&b\\c&a&\cdots&b\\\vdots&\vdots&&\vdots\\c&c&\cdots&a\end{vmatrix}+\begin{vmatrix}a-c&b&\cdots&b\\0&a&\cdots&b\\\vdots&\vdots&&\vdots\\0&c&\cdots&a\end{vmatrix}\\&=\begin{vmatrix}0&b-a&0&\cdots&0\\0&a-c&b-a&\cdots&0\\0&0&a-c&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\c&c&c&\cdots&a\end{vmatrix}+(a-c)D_{n-1}\\&=c(a-b)^{n-1}+(a-c)D_{n-1}.\end{aligned}\tag{2.5.1}$$

将行列式转置, (2.5.1) 变为

$$D_n=b(a-c)^{n-1}+(a-b)D_{n-1}.\tag{2.5.2}$$

(2.5.1), (2.5.2) 联立求解, 得

$$D_n=\frac{c(a-b)^n-b(a-c)^n}{c-b}.$$

解法 2 利用函数的 Taylor 公式计算.

记 $D_n(x)=\begin{vmatrix}x&b&\cdots&b\\c&x&\cdots&b\\\vdots&\vdots&&\vdots\\c&c&\cdots&x\end{vmatrix}$, 则 $D_n=D_n(a)$. 将 $D_n(x)$ 在 $x=b$ 处作 Taylor

展开, 有

$$D_n(x)=D_n(b)+D_n'(b)(x-b)+\frac{D_n''(b)}{2!}(x-b)^2+\cdots+\frac{D_n^{(n)}(b)}{n!}(x-b)^n. \tag{2.5.3}$$

由于

$$\begin{aligned}D_n(b)&=\begin{vmatrix}b&b&b&\cdots&b\\c&b&b&\cdots&b\\c&c&b&\cdots&b\\\vdots&\vdots&\vdots&&\vdots\\c&c&c&\cdots&b\end{vmatrix}\xlongequal[\text{后列减前列}]{\text{从第 } n \text{ 列起}}\begin{vmatrix}b&0&0&\cdots&0\\c&b-c&0&\cdots&0\\c&0&b-c&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\c&0&0&\cdots&b-c\end{vmatrix}\\&=b(b-c)^{n-1},\end{aligned} \tag{2.5.4}$$

$$D_n'(x)=\begin{vmatrix}1&b&\cdots&b\\0&x&\cdots&b\\\vdots&\vdots&&\vdots\\0&c&\cdots&x\end{vmatrix}+\begin{vmatrix}x&0&\cdots&b\\c&1&\cdots&b\\\vdots&\vdots&&\vdots\\c&0&\cdots&x\end{vmatrix}+\cdots+\begin{vmatrix}x&b&\cdots&0\\c&x&\cdots&0\\\vdots&\vdots&&\vdots\\c&c&\cdots&1\end{vmatrix}=nD_{n-1}(x).$$

类推可得

$$D_n''(x)=n(n-1)D_{n-2}(x),\cdots,D_n^{(n-1)}(x)=n!D_1(x)=n!x,\ D_n^{(n)}(x)=n!.$$

利用 (2.5.4) 得

$$\begin{aligned}D_n'(b)&=nb(b-c)^{n-2},\\D_n''(b)&=n(n-1)b(b-c)^{n-3},\\&\cdots\\D_n^{(n-1)}(b)&=n!b,\\D_n^{(n)}(b)&=n!.\end{aligned} \tag{2.5.5}$$

将 (2.5.4), (2.5.5) 代入 (2.5.3) 得

$$\begin{aligned}D_n(x)&=b(b-c)^{n-1}+nb(b-c)^{n-2}(x-b)+\frac{n(n-1)b(b-c)^{n-3}}{2!}(x-b)^2+\cdots+\\&\quad nb(x-b)^{n-1}+(x-b)^n\\&=\frac{b}{b-c}[(b-c)^n+n(b-c)^{n-1}(x-b)+\frac{n(n-1)}{2!}(b-c)^{n-2}(x-b)^2+\cdots+\\&\quad n(b-c)(x-b)^{n-1}+(x-b)^n]-\frac{c}{b-c}(x-b)^n\\&=\frac{b}{b-c}[(b-c)+(x-b)]^n-\frac{c}{b-c}(x-b)^n\\&=\frac{b(x-c)^n-c(x-b)^n}{b-c}.\end{aligned}$$

从而 $D_n = D_n(a) = \dfrac{b(a-c)^n - c(a-b)^n}{b-c}$.

例 2.5.4 计算 n 阶行列式 $D_n = \begin{vmatrix} a+b & ab & & & & \\ 1 & a+b & ab & & & \\ & 1 & & \ddots & & \\ & & & \ddots & & \\ & & \ddots & & & ab \\ & & & & 1 & a+b \end{vmatrix}$

解法 1 对行列式的阶数作归纳.

当 $n=1$ 时, $D_1 = a+b$; 当 $n=2$ 时, $D_2 = \begin{vmatrix} a+b & ab \\ 1 & a+b \end{vmatrix} = a^2+ab+b^2$.

将上面两式作恒等变形得

$$D_1 = \begin{cases} 2a, & a=b, \\ \dfrac{a^2-b^2}{a-b}, & a\neq b. \end{cases} \quad D_2 = \begin{cases} 3a^2, & a=b, \\ \dfrac{a^3-b^3}{a-b}, & a\neq b. \end{cases}$$

假设当 $n<k$ 时, 有 $D_n = \begin{cases} (n+1)a^n, & a=b, \\ \dfrac{a^{n+1}-b^{n+1}}{a-b}, & a\neq b. \end{cases}$

当 $n=k$ 时, 将 D_k 按第一行展开有

$$D_k = (a+b)D_{k-1} - abD_{k-2} \xlongequal{\text{归纳假设}} \begin{cases} 2a\cdot ka^{k-1} - a^2\cdot(k-1)a^{k-2}, & a=b, \\ (a+b)\dfrac{a^k-b^k}{a-b} - ab\dfrac{a^{k-1}-b^{k-1}}{a-b}, & a\neq b. \end{cases}$$

$$= \begin{cases} (k+1)a^k, & a=b, \\ \dfrac{a^{k+1}-b^{k+1}}{a-b}, & a\neq b. \end{cases}$$

所以对一切正整数 n, 有 $D_n = \begin{cases} (n+1)a^n, & a=b, \\ \dfrac{a^{a+1}-b^{n+1}}{a-b}, & a\neq b. \end{cases}$

注意到, $\dfrac{a^{n+1}-b^{n+1}}{a-b} = a^n + a^{n-1}b + \cdots + ab^{n-1} + b^n \xlongequal{a=b} (n+1)a^n$, 所以总有

$$D_n = a^n + a^{n-1}b + \cdots + ab^{n-1} + b^n.$$

解法 2 将 D_n 按第一行展开有

$$D_n = (a+b)D_{n-1} - abD_{n-2}.$$

将上式变形并递推, 得

$$D_n - aD_{n-1} = b(D_{n-1} - aD_{n-2}) = \cdots = b^{n-2}(D_2 - aD_1) = b^n. \tag{2.5.6}$$

当 $a = b$ 时, 有 $D_n = aD_{n-1} + a^n$, 递推得

$$\begin{aligned} D_n &= a(aD_{n-2} + a^{n-1}) + a^n = a^2 D_{n-2} + 2a^n = \cdots \\ &= a^{n-1}D_1 + (n-1)a^n = (n+1)a^n. \end{aligned}$$

当 $a \neq b$ 时, 由于 (2.5.6) 式中交换 a, b 也成立, 有

$$D_n - bD_{n-1} = a^n. \tag{2.5.7}$$

由 (2.5.6), (2.5.7) 可联立解出

$$D_n = \frac{a^{n+1} - b^{n+1}}{a - b}$$

例 2.5.5 计算 n 阶行列式

$$D_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 + 1 & x_2 + 1 & x_3 + 1 & \cdots & x_n + 1 \\ x_1^2 + x_1 & x_2^2 + x_2 & x_3^2 + x_3 & \cdots & x_n^2 + x_n \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} + x_1^{n-2} & x_2^{n-1} + x_2^{n-2} & x_3^{n-1} + x_3^{n-2} & \cdots & x_n^{n-1} + x_n^{n-2} \end{vmatrix}.$$

解法 1 利用行列式的性质, 将该行列式化为 Vandermonde 行列式

$$D_n \xlongequal[i=1,\cdots,n-1]{r_{i+1}-r_i} \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j).$$

解法 2 利用矩阵的乘法分解, 该行列式对应的矩阵可分解为两个矩阵的乘积

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix},$$

故

$$D_n=\det(\boldsymbol{A})=\begin{vmatrix}1&0&0&\cdots&0&0\\1&1&0&\cdots&0&0\\0&1&1&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&1&0\\0&0&0&\cdots&1&1\end{vmatrix}\begin{vmatrix}1&1&1&\cdots&1\\x_1&x_2&x_3&\cdots&x_n\\x_1^2&x_2^2&x_3^2&\cdots&x_n^2\\\vdots&\vdots&\vdots&&\vdots\\x_1^{n-1}&x_2^{n-1}&x_3^{n-1}&\cdots&x_n^{n-1}\end{vmatrix}$$

$$=\prod_{1\leqslant j<i\leqslant n}(x_i-x_j).$$

例 2.5.6 设 n 为偶数, 计算行列式 $D_{n+1}=\begin{vmatrix}1&1&\cdots&1&1\\2&2^2&\cdots&2^n&2^{n+1}\\\vdots&\vdots&&\vdots&\vdots\\n&n^2&\cdots&n^n&n^{n+1}\\\dfrac{n}{2}&\dfrac{n^2}{3}&\cdots&\dfrac{n^n}{n+1}&\dfrac{n^{n+1}}{n+2}\end{vmatrix}$.

解 令 $F(x)=\dfrac{1}{n}\begin{vmatrix}1&1&\cdots&1&1\\2&2^2&\cdots&2^n&2^{n+1}\\\vdots&\vdots&&\vdots&\vdots\\n&n^2&\cdots&n^n&n^{n+1}\\\dfrac{x^2}{2}&\dfrac{x^3}{3}&\cdots&\dfrac{x^{n+1}}{n+1}&\dfrac{x^{n+2}}{n+2}\end{vmatrix}$, 则 $D_{n+1}=F(n)$.

由于

$$F'(x)=\frac{1}{n}\begin{vmatrix}1&1&\cdots&1&1\\2&2^2&\cdots&2^n&2^{n+1}\\\vdots&\vdots&&\vdots&\vdots\\n&n^2&\cdots&n^n&n^{n+1}\\x&x^2&\cdots&x^n&x^{n+1}\end{vmatrix}=(n-1)!x\begin{vmatrix}1&1&\cdots&1^{n-1}&1^n\\1&2&\cdots&2^{n-1}&2^n\\\vdots&\vdots&&\vdots&\vdots\\1&n&\cdots&n^{n-1}&n^n\\1&x&\cdots&x^{n-1}&x^n\end{vmatrix}$$

$$\xlongequal{\text{Vandermonde 行列式}}cx(x-1)(x-2)\cdots(x-n),$$

其中 $c=(n-1)!\prod\limits_{1\leqslant j<i\leqslant n}(i-j)$.

而 $F(0)=0$, 所以

$$\begin{aligned}
D_{n+1}=F(n)&=\int_0^n F'(x)\mathrm{d}x=c\int_0^n x(x-1)(x-2)\cdots(x-n)\mathrm{d}x\\
&\xlongequal{x=t+\frac{n}{2}} c\int_{-\frac{n}{2}}^{\frac{n}{2}}\left(t+\frac{n}{2}\right)\left(t+\frac{n}{2}-1\right)\cdots(t+1)t(t-1)\cdots\left(t-\frac{n}{2}+1\right)\left(t-\frac{n}{2}\right)\mathrm{d}t\\
&=c\int_{-\frac{n}{2}}^{\frac{n}{2}}\left[t^2-\left(\frac{n}{2}\right)^2\right]\left[t^2-\left(\frac{n}{2}-1\right)^2\right]\cdots\left(t^2-1^2\right)t\mathrm{d}t\\
&=0\quad(\text{被积函数是奇函数}).
\end{aligned}$$

行列式的计算方法还有很多, 这里仅仅是一些较常见的方法.

问题 2.6 行列式的 Laplace 展开定理如何证明

工科线性代数教材一般都会介绍行列式的 Laplace 展开定理, 但由于学时与篇幅所限通常都没给出证明. 这里对定理的证明作一介绍.

为便于定理的证明, 我们先介绍一个引理.

引理 行列式 D 中任一个子式 M 与它的代数余子式 A 的乘积中的每一项都是行列式 D 的展开式中的一项, 而且符号也一致.

证 首先讨论 M 位于行列式 D 的左上角的情形.

设 M 是位于 D 的左上角的一个 k 阶子式, M 的余子式为 M', 代数余子式为 A, 即

$$D=\left|\begin{array}{ccc:ccc}
a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n}\\
\vdots & & \vdots & \vdots & & \vdots\\
a_{k1} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn}\\
\hdashline
a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n}\\
\vdots & & \vdots & \vdots & & \vdots\\
a_{n1} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{nn}
\end{array}\right|,$$

则

$$A=(-1)^{(1+2+\cdots+k)+(1+2+\cdots+k)}M'=M'.$$

M 的每一项都可写作

$$a_{1s_1}a_{2s_2}\cdots a_{ks_k},$$

其中 $s_1,s_2,\cdots,s_k$ 是 $1,2,\cdots,k$ 的一个排列, 所以这一项前面所带的符号为 $\tau(s_1,s_2,\cdots,s_k)$.

M' 的每一项都可写作

$$a_{k+1,h_1}a_{k+2,h_2}\cdots a_{n,h_{n-k}}$$

其中 $h_1, h_2, \cdots, h_{n-k}$ 是 $k+1, k+2, \cdots, n$ 的一个排列, 这一项前面所带的符号为 $\tau(h_1, h_2, \cdots, h_{n-k})$.

注意到 $\tau(s_1, s_2, \cdots, s_k) + \tau(h_1, h_2, \cdots, h_{n-k}) = \tau(s_1, s_2, \cdots, s_k, h_1, h_2, \cdots, h_{n-k})$, 所以上面两项 (连同其符号) 的乘积为

$$(-1)^{\tau(s_1,s_2,\cdots,s_k,h_1,h_2,\cdots,h_{n-k})}a_{1s_1}a_{2s_2}\cdots a_{ks_k}a_{k+1,h_1}a_{k+2,h_2}\cdots a_{n,h_{n-k}}.$$

当 M 位于 D 的第 $i_1, i_2, \cdots, i_k$ $(i_1 < i_2 < \cdots < i_k)$ 行, 第 $j_1, j_2, \cdots, j_k$ $(j_1 < j_2 < \cdots < j_k)$ 列时, 变动 D 的行列次序使 M 位于 D 的左上角. 为此, 先把第 i_1 行依次与第 $i_1-1, i_1-2, \cdots, 2, 1$ 行对换, 这样经过了 i_1-1 次对换, 将第 i_1 行换到了第 1 行. 再将第 i_2 行依次与第 $i_2-1, i_2-2, \cdots, 2$ 行对换而换到第 2 行, 一共经过了 i_2-1 次对换. 如此继续进行, 一共经过了

$$(i_1-1)+(i_2-1)+\cdots+(i_k-1)=(i_1+i_2+\cdots+i_k)-(1+2+\cdots+k)$$

次对换而把第 $i_1, i_2, \cdots, i_k$ 行依次换到了第 $1, 2, \cdots, k$ 行.

利用类似的列变换, 可以将 M 的列换到第 $1, 2, \cdots, k$ 列, 一共作了

$$(j_1+j_2+\cdots+j_k)-(1+2+\cdots+k)$$

次列变换.

我们用 D' 表示这样变换后所得的新行列式, 那么

$$\begin{aligned} D' &= (-1)^{(i_1+i_2+\cdots+i_k)-(1+2+\cdots+k)+(j_1+j_2+\cdots+j_k)-(1+2+\cdots+k)}D \\ &= (-1)^{i_1+i_2+\cdots+i_k+j_1+j_2+\cdots+j_k}D. \end{aligned}$$

由此看出, D' 和 D 的展开式中出现的项是一样的, 只是每一项都差符号 $(-1)^{i_1+i_2+\cdots+i_k+j_1+j_2+\cdots+j_k}$.

现在 M 位于 D' 的左上角, 所以 $M \cdot M'$ 中每一项都是 D' 中的一项, 而且符号一致. 但是

$$M \cdot A = (-1)^{i_1+i_2+\cdots+i_k+j_1+j_2+\cdots+j_k}M \cdot M',$$

所以 $M \cdot A$ 中的每一项都与 D 中的一项相等而且符号一致.

定理 2.6.1 (Laplace 展开定理) 在 n 阶行列式 D 中取定 k 行, 这 k 行的所有 k 阶子式与其对应的代数余子式乘积之和等于行列式的值.

证 设 D 中取定 k 行后得到的全体子式为 $M_1, M_2, \cdots, M_t$, 对应的代数余子式为 $A_1, A_2, \cdots, A_t$, 显然 $t = \mathrm{C}_n^k = \dfrac{n!}{k!(n-k)!}$. 定理 2.6.1 即是要证明

$$D = M_1A_1 + M_2A_2 + \cdots + M_tA_t.$$

根据引理, M_iA_i 中每一项都是 D 中一项而且符号相同, 由此只需证明上式两边的项数相等就可以了. 我们知道等式左边共有 $n!$ 项, 而右边的 M_i 中共有 $k!$ 项, A_i 中共有 $(n-k)!$ 项, 所以右边共有 $t \cdot k!(n-k)! = n!$ 项. 定理得证.

问题 2.7 分块行列式也有初等变换性质吗

我们知道行列式的初等变换性质对行列式的化简的作用巨大, 那么分块矩阵的行列式 (下称分块行列式) 是否也有类似的性质? 具体方法怎样? 有何应用? 下面我们作一简略的讨论.

为便于叙述, 我们仅对形如 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix}$ 的分块行列式 ($\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ 均为矩阵) 作讨论, 更一般的分块情形, 可类似处理.

与通常行列式的初等变换性质相类似, 分块行列式有以下初等变换性质:

性质 1 分块行列式的某一块行 (列) 的元左 (右) 乘方阵 $\boldsymbol{K}$, 等于该分块行列式乘数 $|\boldsymbol{K}|$. 即

$$\begin{vmatrix} \boldsymbol{KA} & \boldsymbol{KB} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{K}| \cdot \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix}.$$

性质 2 对换分块行列式中两块行 (列) 的位置, 行列式增加一符号因子. 即

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = (-1)^{lm} \begin{vmatrix} \boldsymbol{C} & \boldsymbol{D} \\ \boldsymbol{A} & \boldsymbol{B} \end{vmatrix},$$

其中 l, m 分别为 $\boldsymbol{A}, \boldsymbol{C}$ 所对应的行数.

性质 3 分块行列式的某一块行 (列) 左 (右) 乘矩阵 $\boldsymbol{K}$ 加到另一块行 (列), 行列式值不变. 即

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C}+\boldsymbol{KA} & \boldsymbol{D}+\boldsymbol{KB} \end{vmatrix}.$$

注 以上性质中所涉及的矩阵运算, 都假定是有意义的.

证 利用分块矩阵的运算及行列式的 Laplace 展开定理, 有

$$\begin{vmatrix} \boldsymbol{KA} & \boldsymbol{KB} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \left| \begin{pmatrix} \boldsymbol{K} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \right| = \begin{vmatrix} \boldsymbol{K} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{vmatrix} \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{K}| \cdot \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix},$$

故性质 1 成立.

在性质 2 与性质 3 中, 设 $\boldsymbol{A},\boldsymbol{C}$ 的行数分别为 l,m, 则

$$\begin{aligned}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} &= \left|\begin{pmatrix}\boldsymbol{O} & \boldsymbol{I}_l\\ \boldsymbol{I}_m & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{C} & \boldsymbol{D}\\ \boldsymbol{A} & \boldsymbol{B}\end{pmatrix}\right| = \begin{vmatrix}\boldsymbol{O} & \boldsymbol{I}_l\\ \boldsymbol{I}_m & \boldsymbol{O}\end{vmatrix}\begin{vmatrix}\boldsymbol{C} & \boldsymbol{D}\\ \boldsymbol{A} & \boldsymbol{B}\end{vmatrix}\\ &= (-1)^{(1+2+\cdots+l)+[(m+1)+(m+2)+\cdots+(m+l)]}\begin{vmatrix}\boldsymbol{C} & \boldsymbol{D}\\ \boldsymbol{A} & \boldsymbol{B}\end{vmatrix}\\ &= (-1)^{2(1+2+\cdots+l)+lm}\begin{vmatrix}\boldsymbol{C} & \boldsymbol{D}\\ \boldsymbol{A} & \boldsymbol{B}\end{vmatrix} = (-1)^{lm}\begin{vmatrix}\boldsymbol{C} & \boldsymbol{D}\\ \boldsymbol{A} & \boldsymbol{B}\end{vmatrix},\end{aligned}$$

故性质 2 成立.

$$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C}+\boldsymbol{K}\boldsymbol{A} & \boldsymbol{D}+\boldsymbol{K}\boldsymbol{B}\end{vmatrix} = \left|\begin{pmatrix}\boldsymbol{I}_l & \boldsymbol{O}\\ \boldsymbol{K} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}\right| = \begin{vmatrix}\boldsymbol{I}_l & \boldsymbol{O}\\ \boldsymbol{K} & \boldsymbol{I}_m\end{vmatrix}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = \begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix},$$

故性质 3 成立.

以上性质给行列式的计算带来了更多的方便, 同时还可得到分块行列式的一些非常有用的结果, 使某些行列式的计算大大简化.

例 2.7.1　设 $\boldsymbol{A},\boldsymbol{B}$ 是同阶方阵, 求证: $\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{B} & \boldsymbol{A}\end{vmatrix} = |\boldsymbol{A}+2\boldsymbol{B}||\boldsymbol{A}-\boldsymbol{B}|^2$.

证

$$\begin{aligned}\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{B} & \boldsymbol{A}\end{vmatrix} &\xlongequal[i=2,3]{r_1+r_i} \begin{vmatrix}\boldsymbol{A}+2\boldsymbol{B} & \boldsymbol{A}+2\boldsymbol{B} & \boldsymbol{A}+2\boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B} & \boldsymbol{B} & \boldsymbol{A}\end{vmatrix}\\ &\xlongequal[i=2,3]{c_i-c_1} \begin{vmatrix}\boldsymbol{A}+2\boldsymbol{B} & \boldsymbol{O} & \boldsymbol{O}\\ \boldsymbol{B} & \boldsymbol{A}-\boldsymbol{B} & \boldsymbol{O}\\ \boldsymbol{B} & \boldsymbol{O} & \boldsymbol{A}-\boldsymbol{B}\end{vmatrix} = |\boldsymbol{A}+2\boldsymbol{B}||\boldsymbol{A}-\boldsymbol{B}|^2.\end{aligned}$$

例 2.7.2　设 $\boldsymbol{A}$ 是可逆矩阵, 证明:

$$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|. \tag{2.7.1}$$

证　因为 $\boldsymbol{A}$ 可逆, 则

$$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = |\boldsymbol{A}|\begin{vmatrix}\boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix} = |\boldsymbol{A}|\begin{vmatrix}\boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{vmatrix} = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|.$$

关于例 2.7.2 的几点说明:

1° 若例 2.7.2 增加条件 $\boldsymbol{AC}=\boldsymbol{CA}$, 则结论可改为

$$\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{AD}-\boldsymbol{CB}|. \tag{2.7.2}$$

这是因为

$$\begin{aligned}\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} &= |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{CA}^{-1}\boldsymbol{B}| = |\boldsymbol{AD}-\boldsymbol{ACA}^{-1}\boldsymbol{B}| \\ &= |\boldsymbol{AD}-\boldsymbol{CAA}^{-1}\boldsymbol{B}| = |\boldsymbol{AD}-\boldsymbol{CB}|.\end{aligned}$$

(2.7.2) 在形式上类似于以数为元的二阶行列 $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad-bc$. 要注意两者的区别.

2° (2.7.1) 式右边的行列式阶数低于左边行列式的阶数, 它又称为行列式的**第一降阶公式**, 用它可以简化某些行列式的计算.

3° 由于 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = \begin{vmatrix} \boldsymbol{D} & \boldsymbol{C} \\ \boldsymbol{B} & \boldsymbol{A} \end{vmatrix}$, 所以当 $\boldsymbol{A},\boldsymbol{D}$ 分别是 m 与 n 阶可逆矩阵时, 有

$$|\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{BD}^{-1}\boldsymbol{C}| = |\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{CA}^{-1}\boldsymbol{B}|. \tag{2.7.3}$$

若 $m\neq n$, (2.7.3) 左右两边行列式的阶数也不相同, 通常称它为行列式的**第二降阶公式**.

下面举例说明它们的应用.

例 2.7.3 计算行列式 $D=\begin{vmatrix} 2 & 1 & 0 & 1 & -1 \\ 3 & 2 & 1 & 0 & 2 \\ 1 & 1 & 2 & -1 & 5 \\ 0 & -1 & 6 & 1 & 4 \\ 4 & 5 & 8 & 7 & -1 \end{vmatrix}$.

解 将 1, 4 列互换, 2, 3 列互换, 并按下列方式分块, 再由 (2.7.1) 式得

$$D=\left|\begin{array}{cc:ccc} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & 3 & 2 \\ \hdashline -1 & 2 & 1 & 1 & 5 \\ 1 & 6 & -1 & 0 & 4 \\ 7 & 8 & 5 & 4 & -1 \end{array}\right| = \left|\begin{pmatrix} 1 & 1 & 5 \\ -1 & 0 & 4 \\ 5 & 4 & -1 \end{pmatrix} - \begin{pmatrix} -1 & 2 \\ 1 & 6 \\ 7 & 8 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 2 \end{pmatrix}\right|$$

$$= \left| \begin{pmatrix} 1 & 1 & 5 \\ -1 & 0 & 4 \\ 5 & 4 & -1 \end{pmatrix} - \begin{pmatrix} 3 & 4 & 5 \\ 13 & 20 & 11 \\ 23 & 38 & 9 \end{pmatrix} \right| = \begin{vmatrix} -2 & -3 & 0 \\ -14 & -20 & -7 \\ -18 & -34 & -10 \end{vmatrix} = 118.$$

例 2.7.4 计算 n 阶行列式 $D_n = \begin{vmatrix} 2 & 2 & 3 & \cdots & n \\ 1 & 3 & 3 & \cdots & n \\ 1 & 2 & 4 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & \cdots & n+1 \end{vmatrix}$.

解 先将 D_n 改写为下列两个方阵之和的行列式, 再凑成 $\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}$ 的形状:

$$D_n = \left| \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \\ \vdots & \vdots & & \vdots \\ 1 & 2 & \cdots & n \end{pmatrix} \right|$$

$$= \left| \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (1)^{-1} (1, 2, \cdots, n) \right|$$

利用 (2.7.3) 式得

$$D_n = \begin{vmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{vmatrix} \cdot \left| 1 + (1, 2, \cdots, n) \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right|$$

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

第三章　向量空间与线性方程组

问题 3.1　如何认识“n 维向量空间”所研究的问题

n 维向量空间是线性代数课程的一个重要部分, 这部分内容涉及的概念较多, 理论性强, 是学生学习中的难点. 不少学生学完以后也不能把握这部分知识的缘由及整体构架, 因此, 我们有必要对 n 维向量空间所研究的内容从宏观上作一粗略的勾画.

(1) n 维向量空间是 3 维几何空间的推广

我们知道, n 维向量空间是 3 维几何空间的推广, 要了解 n 维向量空间所研究的内容, 不妨先考察 3 维空间中关于向量部分的内容, 然后再作类比.

首先, 3 维空间中的向量有几何表示形式 (有向线段) 、坐标表示形式, 向量之间有线性运算 (加法 、数乘) 、乘法运算 (内积 、外积). 对 n $(n>3)$ 维空间, 向量已没有几何形式, 但有坐标形式, 只涉及线性运算的空间叫 n 维线性空间 (又叫 n 维向量空间), 既有线性运算又有内积运算的实向量空间称为欧氏空间.

向量的线性运算解决了向量间的表示问题. 3 维空间中的任一向量都可由坐标轴方向的单位向量 (基向量) 线性表示, 而且表示是唯一的. 这种表示十分重要, 它将无穷多个向量的问题转化成了 3 个基向量的问题, 基向量组就成了整个空间的 “代表”, 使整个空间得到了简化. 空间中, 能由一个向量线性表示的向量称为共线的, 能由两个向量线性表示的向量称为共面的. 在 n 维空间中, 若一个向量组中的某一个能由其余的线性表示, 则称这个向量组是线性相关的, 否则称线性无关的. 3 维空间中向量组的线性相 (无) 关性容易判定, 高维空间中少了几何直观, 问题也更为复杂, 就需要研究向量组线性相关性的一般性理论, 就会有极大线性无关组 (简称极大无关组) 、秩的概念, 对整个空间来说就是基与维数的概念.

向量的内积为空间增添了几何度量 (距离、夹角), 由于 3 维空间中我们关心的主要是几何问题, 所以内积是很有用的工具. 但是在 n 维线性空间中我们主要关心的不再是几何问题, 而是代数问题, 所以一般不讨论与内积有关的内容 (与内积有关的内容通常在 n 维欧氏空间中去讨论).

谈到 “空间” 就离不开 “运动”, 现实生活中的 “空间” 通常是指一个 “运动的范围”. n 维线性空间中向量的线性运算可以视为向量的 “运动”, 但更一般意义的运动是空间中的线性变换, 3 维仿射空间 (没有内积) 中有仿射变换, 直角坐标空间中有正交变换. 这些变换也可以推广到 n 维空间, 但由于工科线性代数的学时与内容所限, 有关线性变换的内容一般都很简略, 甚至没有介绍.

(2) n 维向量空间是线性空间的一个代表

将 n 元数组称为向量并不是因为它们有大小和方向, 而是它们保持了 3 维几何空间中向量的线性运算规律 (8 条运算律). 同样的理由, 可以将很多别的集合 V 中的元素称为向量, 只要该集合中定义了线性运算并满足 8 条运算律, 就称该集合 V 为一个向量空间或线性空间. 我们通常只讨论由 n 元数组所作成的向量空间, 是因为这个空间中的元素最具代表, 并且形式简单, 其他线性空间中元素间的运算都可以转化为 n 元数组来进行. 比如, 设 V 是数域 P 上的线性空间, $\boldsymbol{\alpha} \in V$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 下的表示为

$$\boldsymbol{\alpha} = a_1\boldsymbol{\alpha}_1 + a_2\boldsymbol{\alpha}_2 + \cdots + a_n\boldsymbol{\alpha}_n \quad (a_i \in P, i = 1, \cdots, n),$$

则 n 元数组 $\boldsymbol{a} \triangleq (a_1, a_2, \cdots, a_n) \in P^n$ 就是 $\boldsymbol{\alpha}$ 的代表 (也称为 $\boldsymbol{\alpha}$ 的坐标), 这样就可以建立 V 与 n 元数组作成的向量空间 P^n 之间的一一映射 $\varphi : \boldsymbol{\alpha} \to \boldsymbol{a}$ 而且这种映射还保持线性关系:

$$\varphi(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) = k_1\varphi(\boldsymbol{\alpha}_1) + k_2\varphi(\boldsymbol{\alpha}_2) \quad (k_1, k_2 \in P),$$

称 φ 为同构映射 (也称 V 与 P^n 是同构的). 对 n 维向量空间 P^n 作研究, 从而避免对每一个具体的空间做重复的工作.

(3) 线性方程组与向量组、向量空间

由于 n 元线性方程之间可以作加法运算, 可以作数乘运算, 所以可以将 n 元线性方程 $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ 用数组 $(a_1, a_2, \cdots, a_n, b)$ 代表, 或者说这个数组就是这个方程的坐标. 线性方程组的化简过程就是用向量组等价的方法找出极大无关组的过程. 从这层意义上讲以线性方程为元素的集合就是一个具体的线性空间,n 元线性方程组就是一个向量组.

另一方面, 一个 n 元线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\cdots\cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m, \end{cases} \tag{3.1.1}$$

又可以用向量的线性组合来表示

$$x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{b}, \tag{3.1.2}$$

其中 $\boldsymbol{\alpha}_i=(a_{1i},a_{2i},\cdots,a_{mi})^{\mathrm{T}}$ $(i=1,2,\cdots,n)$, $\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$. 方程组的求解就是求 $\boldsymbol{b}$ 与向量组 $\boldsymbol{\alpha}_i$ $(i=1,2,\cdots,n)$ 的线性关系. 当 $\boldsymbol{b}=\boldsymbol{0}$ 时, 它反映了向量组 $\boldsymbol{\alpha}_i$ $(i=1,2,\cdots,n)$ 的线性相关性. 从这层意义上讲, 线性方程组是线性空间中一组向量间的线性关系的反映.

若记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 则方程组 (3.1.2) 又可表示为

$$\boldsymbol{Ax}=\boldsymbol{b}. \tag{3.1.3}$$

线性方程组的求解又可看成是求向量 $\boldsymbol{b}$ 在线性变换 $\mathcal{A}:\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 下的原像问题. 当 $\boldsymbol{b}=\boldsymbol{0}$ 时, 这种原像的全体又构成 n 维向量空间的一个子空间.

因此, 无论是从哪个层面来看, 线性方程组与向量组、向量空间都有密切的联系, 它们之间是难以分割的.

问题 3.2　如何引入线性表出与线性相关等概念

向量组的线性相关、线性无关等概念是线性代数课程中的基本概念, 这些概念较为抽象, 学生难以理解. 如何引入这些概念, 使学生能有效地理解和掌握是教学中值得探讨的. 我们根据自己的教学经验, 介绍一种由方程组 (或向量组) 引入线性表出、线性相关等概念及其意义的方法.

(1) 线性表出概念的引入

考虑线性方程组:

$$\begin{cases} x_1+x_2-2x_3=2, & ① \\ 2x_1+x_2+3x_3=-3, & ② \\ 4x_1+3x_2-x_3=1. & ③ \end{cases} \tag{3.2.1}$$

不难发现:

1° 将第一个方程的两倍加上第二个方程就会得到第三个方程, 即

$$2 \times ① + ② = ③. \tag{3.2.2}$$

也就是说, 第三个方程可以由前两个方程经线性运算得到. 对此, 我们也称第三个方程可由前两个方程线性表出, 或称第三个方程是前两个方程的线性组合.

2° 方程组 (3.2.1) 是否有解等价于如下向量方程是否有解

$$x_1 \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}. \tag{3.2.3}$$

也就是说, 方程右边的向量如果能由方程左边的 3 个向量通过线性运算得到, 那么方程组就有解; 否则, 方程组无解. 在有解时, 也称方程右边的向量能由左边的 3 个向量线性表出, 或称方程右边向量是左边的 3 个向量的线性组合.

因此我们就可以很自然地给出线性表出或线性组合的概念:

定义 3.2.1 设向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$, 如果存在 $k_1, k_2, \cdots, k_m$ 使得

$$\boldsymbol{\beta} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m, \tag{3.2.4}$$

就称向量 $\boldsymbol{\beta}$ 可以由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ **线性表出**, 也称 $\boldsymbol{\beta}$ 是向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的一个**线性组合**.

同时我们看到向量 $\boldsymbol{\beta}$ 能否由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表出就等价于线性方程组 (3.2.4) 是否有解. 因此, 线性表出的问题就经常与线性方程组是否有解的问题联系在一起. 但是我们又不能将线性表出的定义式完全等同于一个线性方程组. 比如, 我们常说 "零向量可由任一向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表出", 这是因为

$$\mathbf{0} = 0\boldsymbol{\alpha}_1 + 0\boldsymbol{\alpha}_2 + \cdots + 0\boldsymbol{\alpha}_m.$$

或 "一个向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 中的任一个向量 $\boldsymbol{\alpha}_i$ 都可由该向量组自身线性表出", 这是因为

$$\boldsymbol{\alpha}_i = 0\boldsymbol{\alpha}_1 + \cdots + 0\boldsymbol{\alpha}_{i-1} + 1\boldsymbol{\alpha}_i + 0\boldsymbol{\alpha}_{i+1} \cdots + 0\boldsymbol{\alpha}_m,$$

这两句话更多的是表述一种客观事实, 而并非线性方程组的问题. 再如上面的 (3.2.2) 式也是一个线性表示式, 但它却不是通常意义下的方程组了.

(2) 线性相关概念的引入

我们再来考察线性方程组的情况. 前面我们谈到方程组 (3.2.1) 中的第三个方程能由前两个方程线性表出, 那么对方程组的求解来说第三个方程就是多余的. 在方程组的求解中我们有必要将那些"多余的"方程找出来并"清除掉". 那么如何来判断一个方程组中是否有"多余的"方程? 为方便起见, 我们将每个方程所对应的向量拿出来讨论, 比如方程组 (3.2.1) 对应的三个向量为 $\boldsymbol{\alpha}_1 = (1,1,-2,2), \boldsymbol{\alpha}_2 = (2,1,3,-3), \boldsymbol{\alpha}_3 = (4,3,-1,1)$, 这三个向量中若有某个能由其余的线性表出, 对应的方程组中就有"多余的"方程, 此时我们称这三个向量作成的向量组是**线性相关**的. 要考察一个向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 中有哪些向量能由其余的线性表出, 我们不可能将向量组中的每一个向量当作定义 3.2.1 中的 $\boldsymbol{\beta}$, 然后利用 (3.2.4) 逐一地解方程组, 这样的工作量太大, 一种更科学的办法是作线性方程组

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_m\boldsymbol{\alpha}_m = \mathbf{0}.$$

若该齐次线性方程组有非零解, 则非零系数 x_i 所对应的向量 $\boldsymbol{\alpha}_i$ 就能够由其余的向量线性表出. 于是得到向量组线性相关的概念如下:

定义 3.2.2 对向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$, 若存在一组不全为零的数 $k_1,k_2,\cdots,k_m$ 满足

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}, \tag{3.2.5}$$

则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ **线性相关**; 否则, 若当且仅当 $k_1,k_2,\cdots,k_m$ 全为零时才有 (3.2.5) 成立, 则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ **线性无关**.

通过上面的描述, 我们很容易得到结论:

向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关的充分必要条件是其中至少有一个向量能由其余的向量线性表出, 线性无关的充分必要条件是其中每一个向量都不能由其余的向量线性表出.

同时我们还可看到, 向量组的线性相关性总是与齐次线性方程组是否有非零解密切相连.

问题 3.3 判定向量组线性相关性的常见方法

我们已经知道向量组的线性相关性与齐次线性方程组是否有非零解密切相关, 而齐次线性方程组是否有非零解又取决于系数矩阵的秩, 具体说来即有下面的定理成立:

定理 3.3.1 设 $\boldsymbol{A} = (\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 是一个 $m \times n$ 矩阵, 则以下各结论等价:

1° $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性相关 (无关);

2° $\boldsymbol{Ax} = \mathbf{0}$ 有非零解 (只有零解);

3° $\mathrm{rank}(\boldsymbol{A}) < n$ $(\mathrm{rank}(\boldsymbol{A}) = n)$.

上面定理给出了判定向量组是否线性相关最常用的方法. 由于矩阵的秩又涉及许多其他问题 (如行列式、特征值、标准形等), 因而又会引申出许多别的判定方法, 同时这也从另一个侧面反映出讨论向量组线性相关性的意义.

例 3.3.1 设列向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性无关, $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{C}_{s\times s}$, 证明: $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性无关的充分必要条件是 $\det(\boldsymbol{C}) \neq 0$.

证法 1 令 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s), \boldsymbol{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)$, 则

$$\begin{aligned} \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s \text{ 线性无关} &\Leftrightarrow (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)\boldsymbol{x} = \mathbf{0} \text{ 只有零解} \\ &\Leftrightarrow \boldsymbol{ACx} = \mathbf{0} \text{ 只有零解} \\ &\Leftrightarrow \boldsymbol{A}(\boldsymbol{Cx}) = \mathbf{0} \text{ 只有零解}, \end{aligned}$$

再由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性无关, 知

$$\boldsymbol{A}(\boldsymbol{Cx}) = \mathbf{0} \Leftrightarrow \boldsymbol{Cx} = \mathbf{0} \Leftrightarrow \det(\boldsymbol{C}) \neq 0.$$

证法 2 **充分性** 因为

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{C}.$$

若 $\det(\boldsymbol{C}) \neq 0$, 则 $\boldsymbol{C}$ 可逆, 有

$$(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)\boldsymbol{C}^{-1}.$$

上面两式说明向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 与 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 等价. 所以

$$\mathrm{rank}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = \mathrm{rank}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s) = s,$$

故 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性无关.

必要性 若 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性无关, 则 $\mathrm{rank}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = s$. 又 $\mathrm{rank}(\boldsymbol{C}) \geqslant \mathrm{rank}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)$, 所以 $\mathrm{rank}(\boldsymbol{C}) = s, \det(\boldsymbol{C}) \neq 0$.

需要注意的是, 以下解法是不正确的:

因为 $(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\boldsymbol{C}_{s\times s}$, 两边取行列式得

$$\det(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = \det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s) \cdot \det(\boldsymbol{C}).$$

由于 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 线性无关, 所以 $\det(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s) \neq 0$, 故

$$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s \text{ 线性无关} \Leftrightarrow \det(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) \neq 0 \Leftrightarrow \det(\boldsymbol{C}) \neq 0.$$

以上不正确的原因是 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$ 与 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)$ 不一定是方阵, 所以不能取行列式.

例 3.3.2 已知向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ $(s\geqslant 2)$ 线性无关. 设

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\quad \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\cdots,\boldsymbol{\beta}_{s-1}=\boldsymbol{\alpha}_{s-1}+\boldsymbol{\alpha}_s,\quad \boldsymbol{\beta}_s=\boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1.$$

试讨论向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 的线性相关性.

解法 1 考察 $x_1\boldsymbol{\beta}_1+x_2\boldsymbol{\beta}_2+\cdots+x_s\boldsymbol{\beta}_s=\mathbf{0}$, 即

$$x_1(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+x_2(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+\cdots+x_s(\boldsymbol{\alpha}_s+\boldsymbol{\alpha}_1)=\mathbf{0},$$
$$(x_1+x_n)\boldsymbol{\alpha}_1+(x_1+x_2)\boldsymbol{\alpha}_2+\cdots+(x_{s-1}+x_s)\boldsymbol{\alpha}_s=\mathbf{0}.$$

因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关, 所以

$$\begin{cases} x_1+x_n=0, \\ x_1+x_2=0, \\ \cdots\cdots\cdots\cdots \\ x_{s-1}+x_s=0. \end{cases} \tag{3.3.1}$$

方程组的系数行列式为

$$|\boldsymbol{A}|=\begin{vmatrix} 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{vmatrix}=1+(-1)^{s+1}=\begin{cases} 2, & s\text{ 为奇数}, \\ 0, & s\text{ 为偶数}. \end{cases}$$

当 s 为奇数时, 方程组 (3.3.1) 只有零解, 向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性无关; 当 s 是偶数时, 方程组 (3.3.1) 有非零解, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 线性相关.

解法 2 因为 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{pmatrix}$, 而

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{vmatrix}=1+(-1)^{s+1}=\begin{cases} 2, & s\text{ 为奇数}, \\ 0, & s\text{ 为偶数}. \end{cases}$$

由例 3.3.1 知, 当 s 是奇数时, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性无关; 当 s 是偶数时, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 线性相关.

例 3.3.3　设 $\boldsymbol{A}$ 是 n 阶方阵, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 分别是 $\boldsymbol{A}$ 的 s 个互不相同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 所对应的特征向量, 则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$ 是线性无关的.

证　一般教材中对该结论都有证明, 基本上都是对向量个数 s 用数学归纳法. 这里用线性方程组与矩阵为工具来处理.

设

$$k_1\boldsymbol{\alpha}_1 + \cdots + k_s\boldsymbol{\alpha}_s = \boldsymbol{0}.$$

在该式两端分别用 $\boldsymbol{A}^0 = \boldsymbol{I}, \boldsymbol{A}^1, \boldsymbol{A}^2, \cdots, \boldsymbol{A}^{s-1}$ 左乘之, 利用 $\boldsymbol{A}\boldsymbol{\alpha}_i = \lambda_i\boldsymbol{\alpha}_i$, 就可以得到如下关系式:

$$\begin{cases} k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_s\boldsymbol{\alpha}_s = \boldsymbol{0}, \\ k_1\lambda_1\boldsymbol{\alpha}_1 + k_2\lambda_2\boldsymbol{\alpha}_2 + \cdots + k_s\lambda_s\boldsymbol{\alpha}_s = \boldsymbol{0}, \\ \cdots\cdots\cdots\cdots \\ k_1\lambda_1^{s-1}\boldsymbol{\alpha}_1 + k_2\lambda_2^{s-1}\boldsymbol{\alpha}_2 + \cdots + k_s\lambda_s^{s-1}\boldsymbol{\alpha}_s = \boldsymbol{0}. \end{cases}$$

写成矩阵形式, 就有

$$\begin{aligned} \boldsymbol{O} = (\boldsymbol{0}, \boldsymbol{0}, \cdots, \boldsymbol{0}) &= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\begin{pmatrix} k_1 & k_1\lambda_1 & \cdots & k_1\lambda_1^{s-1} \\ k_2 & k_2\lambda_2 & \cdots & k_2\lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ k_s & k_s\lambda_s & \cdots & k_s\lambda_s^{s-1} \end{pmatrix} \\ &= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s)\begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_s \end{pmatrix}\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_s & \cdots & \lambda_s^{s-1} \end{pmatrix}, \end{aligned}$$

注意到矩阵 $\boldsymbol{R} = \begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_s & \cdots & \lambda_s^{s-1} \end{pmatrix}$ 的行列式为 Vandermonde 行列式, 再由 $\lambda_1, \cdots, \lambda_s$ 互异知该行列式非零, 所以 $\boldsymbol{R}$ 可逆, 上式两端同时右乘 $\boldsymbol{R}^{-1}$, 于是就

有

$$O=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)\begin{pmatrix}k_1&&&\\&k_2&&\\&&\ddots&\\&&&k_s\end{pmatrix}=(k_1\boldsymbol{\alpha}_1,k_2\boldsymbol{\alpha}_2,\cdots,k_s\boldsymbol{\alpha}_s),$$

从而有 $k_1\boldsymbol{\alpha}_1=k_2\boldsymbol{\alpha}_2=\cdots=k_s\boldsymbol{\alpha}_s=\mathbf{0}$, 由 $\boldsymbol{\alpha}_i\neq\mathbf{0}$ 立得 $k_i=0$. 因此这些属于不同特征值的特征向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是线性无关的.

判定一个向量组是否线性相关还有很多其他方法, 我们这里不是为了总结或罗列这些方法, 而是要在学习中把握知识间的相互联系, 增强分析问题、解决问题的能力.

下面再列举一个有关矩阵的秩问题, 该问题需要借助线性方程组、向量组的线性相关性等知识才能得到比较好的解决.

例 3.3.4 设 $\boldsymbol{A}$ 是 n 阶非零矩阵, 证明: $\operatorname{rank}(\boldsymbol{A}^n)=\operatorname{rank}(\boldsymbol{A}^{n+1})$.

证 只需证明线性方程组 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 与 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 同解.

显然方程组 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 的解都是 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 的解.

下证 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 的解也是 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 的解.

运用反证法, 设 $\boldsymbol{\alpha}\neq\mathbf{0}$ 是 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 的解向量, 但不是 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 的解, 则 $\boldsymbol{A}^n\boldsymbol{\alpha}\neq\mathbf{0}$, 下证 $\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha},\cdots,\boldsymbol{A}^n\boldsymbol{\alpha}$ 线性无关. 作线性组合

$$k_0\boldsymbol{\alpha}+k_1\boldsymbol{A}\boldsymbol{\alpha}+\cdots+k_n\boldsymbol{A}^n\boldsymbol{\alpha}=\mathbf{0},\tag{3.3.2}$$

用 $\boldsymbol{A}^n$ 左乘以上等式两边可得 $k_0\boldsymbol{A}^n\boldsymbol{\alpha}=\mathbf{0}$, 因为 $\boldsymbol{A}^n\boldsymbol{\alpha}\neq\mathbf{0}$, 故 $k_0=0$.

将 $k_0=0$ 代入 (3.3.2) 式, 并用 $\boldsymbol{A}^{n-1}$ 左乘等式两边, 同理可得 $k_1=0$; 以此类推可知 $k_0=k_1=\cdots=k_n=0$, 所以 $\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha},\cdots,\boldsymbol{A}^n\boldsymbol{\alpha}$ 线性无关. 由此我们得到有 $n+1$ 个 n 维向量线性无关, 矛盾. 所以方程组 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 的解也都是 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 的解.

从而 $\boldsymbol{A}^n\boldsymbol{x}=\mathbf{0}$ 与 $\boldsymbol{A}^{n+1}\boldsymbol{x}=\mathbf{0}$ 同解, 所以 $\operatorname{rank}(\boldsymbol{A}^n)=\operatorname{rank}(\boldsymbol{A}^{n+1})$.

问题 3.4 向量组的极大线性无关组有何意义与等价形式

(1) 向量组的极大线性无关组的概念及其意义

不同的教材对向量组的极大线性无关组的定义在形式上可能有所不同, 但这些不同都没有本质上的差异, 换言之, 它们都是些等价的说法. 这里选取一般工科线性代数教材较常见形式给出定义如下:

定义 3.4.1 设有向量组 A, 如果在 A 中能选出 r 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$, 满足

1° 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关;

2° 向量组 A 中任意 $r+1$ 个向量 (如果 A 中有 $r+1$ 个向量的话) 都线性相关, 那么称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是向量组 A 的一个极大线性无关向量组, 简称**极大无关组**, 称 r 为向量组 A 的秩.

从形式上看, 定义着重于 “线性无关” 与 “极大” 两层含义. 极大无关组所关心的是向量组的一个部分组, 这个部分组首先是线性无关的, 同时又是线性无关部分组中向量个数最多的, 否则再添加一个向量进去就线性相关了.

从本质上看, 向量组的极大无关组是向量组自身的含向量个数最少的 “代表” (这个 “代表” 与整个向量组等价). 个数最少体现在两个方面: 一是线性无关 (极大无关组中没有任何一个向量能由其余的向量线性表出); 二是不能再少, 否则就能找到还能加入进去且保持线性无关的向量, 而这些新加入的向量又不能由原有的向量线性表出.

如果我们所关心的向量组是一个向量空间, 我们就称它的一个极大无关组为空间的一组**基**, 向量组的秩叫空间的**维数**. 由于一个非零的向量空间中有无穷多个元素, 而它的每个元素都能由它的一组基线性表示, 或者说向量空间是由它的一组基生成的, 所以对有限维空间来说, 基向量 (极大无关组) 就实现了用有限来表示无限的简化目标.

比如求解 n 元齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$, 我们的目的就是要找到方程组的一个基础解系 (解空间的一组基) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ $(r=\operatorname{rank}(\boldsymbol{A}))$, 这样方程组的全部解就可表示为

$$\boldsymbol{x}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}.$$

我们通常将它称为方程组的通解, 也就是整个解空间.

从更一般的意义来讲, 对任意一个有限维线性空间, 由于空间中的每一个向量都能由它的基向量唯一地线性表出, 利用线性表出的系数所作成的有序数组 (称为向量在这组基下的坐标), 就可以将抽象的向量问题转化为具体的向量 (有序数组) 来处理了. 或者说, 基向量能使我们实现同维空间的相互转换 (同构).

由于极大无关组所含向量个数较少, 且其中的向量又不能相互线性表出, 在很多时候用它去 “代表” 整个向量组, 就会使问题简化, 便于处理.

例 3.4.1 设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 矩阵. 证明: $\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})\leqslant\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B})$.

证 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),\boldsymbol{B}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$, 则

$$\boldsymbol{A}+\boldsymbol{B}=(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_n+\boldsymbol{\beta}_n).$$

因为 $\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_n+\boldsymbol{\beta}_n$ 可由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 线性表出, 所以

$$\begin{aligned}\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})&=\operatorname{rank}(\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1,\boldsymbol{\alpha}_2+\boldsymbol{\beta}_2,\cdots,\boldsymbol{\alpha}_n+\boldsymbol{\beta}_n)\\&\leqslant\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n).\end{aligned}$$

设 $\operatorname{rank}(\boldsymbol{A})=r_1$, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{r_1}$ 为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 的一个极大无关组; $\operatorname{rank}(\boldsymbol{B})=r_2$, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{r_2}$ 为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 的一个极大无关组.

由向量组与其极大无关组等价, 知

$$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{r_1},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{r_2}\text{ 与 }\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\text{ 等价.}$$

从而它们有相同的秩. 即有

$$\begin{aligned}\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)&=\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{r_1},\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_{r_2})\\&\leqslant r_1+r_2,\end{aligned}$$

所以

$$\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})\leqslant\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B}).$$

例 3.4.2 设向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 可由向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 线性表出, 试证: 若两向量组秩相等, 则两向量组等价.

证 设 $\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)=\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)=r$, 且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 分别为两向量组的极大无关组.

由于 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ 线性表出, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 可由 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 线性表出, 从而存在 $r\times r$ 矩阵 $\boldsymbol{C}$ 使得

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r)\boldsymbol{C}.$$

由于 $\operatorname{rank}(\boldsymbol{C})\geqslant\operatorname{rank}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)=r$, 则 $\boldsymbol{C}$ 可逆, 有

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)\boldsymbol{C}^{-1}.$$

上面两式说明向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$ 等价, 从而原来的两个向量组等价.

例 3.4.3 线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 求解中, 将其增广矩阵 $\bar{\boldsymbol{A}}=(\boldsymbol{A},\boldsymbol{b})$ 作初等行变换化为阶梯形的过程就是寻找方程组中极大无关组的过程, 阶梯形矩阵中非零行所对应的 $\bar{\boldsymbol{A}}$ 中的向量就是 $\bar{\boldsymbol{A}}$ 中行向量组的一个极大无关组, 对应的方程组就是原方程组的一个极大无关 (方程) 组.

(2) 极大无关组的等价描述

为了使学生掌握极大无关组的本质，学习中有必要使他们了解一些相关的等价描述，我们将一些等价命题不加证明地列在下面，供读者参考.

定理 3.4.1　设有向量组 A, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是 A 的一个部分向量组，则如下条件彼此等价:

1° $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 是 A 的极大无关组;

2° $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且对任一 $\boldsymbol{\beta}\in A,\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r,\boldsymbol{\beta}$ 线性相关;

3° $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且向量组 A 可以由向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性表出;

4° $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且向量组 A 与向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 等价;

5° 向量组 A 中任一向量都可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 唯一地线性表出;

6° $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$ 线性无关，且向量组 A 的秩为 r.

问题 3.5　何谓两个向量组各个向量之间有相同的线性关系

称 n 维向量组 $A:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 与 l 维向量组 $B:\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$ (n 与 l 不一定相等) 各个向量之间有相同的线性关系是指: 在不改变向量组各向量排列顺序的条件下，对任意的 $1\leqslant i_1<i_2<\cdots<i_k\leqslant m$，向量组 $\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k}$ 具有怎样的线性关系，则 $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_k}$ 也具有相同的线性关系，反之亦然.

特别地,

$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k}$ 线性相 (无) 关 $\Leftrightarrow$ $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_k}$ 线性相 (无) 关;

$\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_r}$ 是 A 的极大无关组 $\Leftrightarrow$ $\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_r}$ 是 B 的极大无关组;

$\boldsymbol{\alpha}_j=k_{i_1}\boldsymbol{\alpha}_{i_1}+k_{i_2}\boldsymbol{\alpha}_{i_2}+\cdots+k_{i_r}\boldsymbol{\alpha}_{i_r}\Leftrightarrow\boldsymbol{\beta}_j=k_{i_1}\boldsymbol{\beta}_{i_1}+k_{i_2}\boldsymbol{\beta}_{i_2}+\cdots+k_{i_r}\boldsymbol{\beta}_{i_r}$.

以上关系看起来很复杂，但实际上向量组 A 与 B 各个向量之间有相同的线性关系的充分必要条件是齐次线性方程组$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)\boldsymbol{x}=\boldsymbol{0}$ 与 $(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)\boldsymbol{x}=\boldsymbol{0}$ 同解. 这是因为

$$
\begin{aligned}
&(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)\boldsymbol{x}=\boldsymbol{0} \text{ 与 } (\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)\boldsymbol{x}=\boldsymbol{0} \text{ 同解} &(3.5.1)\\
\Leftrightarrow\ &(\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k})\boldsymbol{y}=\boldsymbol{0} \text{ 与 } (\boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_k})\boldsymbol{y}=\boldsymbol{0} \text{ 同解}\\
&(1\leqslant i_1<i_2<\cdots<i_k\leqslant m) &(3.5.2)\\
\Leftrightarrow\ &\boldsymbol{\alpha}_{i_1},\boldsymbol{\alpha}_{i_2},\cdots,\boldsymbol{\alpha}_{i_k} \text{ 与 } \boldsymbol{\beta}_{i_1},\boldsymbol{\beta}_{i_2},\cdots,\boldsymbol{\beta}_{i_k} \text{ 具有相同的线性关系.}
\end{aligned}
$$

为便于对 (3.5.1) 与 (3.5.2) 等价性的理解，我们不妨设 (3.5.1) 式中两方程组的系数矩阵均为行最简形 (否则可化为行最简形). 由于矩阵的行最简形是唯

一的, 所以两方程组同解的充要条件是它们的系数矩阵有完全相同的非零行. 故 (3.5.1) 与 (3.5.2) 等价性就显然了.

教材中的结论: “矩阵的初等行变换不改变其列向量组的线性相关性” 就是指行变换前后的列向量组具有相同的线性关系, 我们常利用这种关系来求向量组的极大无关组以及用极大无关组来表示其余向量.

例 3.5.1　设向量组

$$\begin{aligned}\boldsymbol{\alpha}_1 &= (1,1,1,3)^{\mathrm{T}}, & \boldsymbol{\alpha}_2 &= (-1,-3,5,1)^{\mathrm{T}},\\ \boldsymbol{\alpha}_3 &= (3,2,-1,p+2)^{\mathrm{T}}, & \boldsymbol{\alpha}_4 &= (4,1,6,10-p)^{\mathrm{T}},\end{aligned}$$

当 p 为何值时, 该向量组线性相关? 此时求出它的秩和一个极大无关组, 并将其余向量用该极大无关组线性表示出来.

解　对矩阵 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)$ 作初等行变换:

$$\begin{array}{c} \begin{array}{cccc} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \boldsymbol{\alpha}_3 & \boldsymbol{\alpha}_4 \end{array} \\ \begin{pmatrix} 1 & -1 & 3 & 4\\ 1 & -3 & 2 & 1\\ 1 & 5 & -1 & 6\\ 3 & 1 & p+2 & 10-p \end{pmatrix} \end{array} \to \begin{pmatrix} 1 & -1 & 3 & 4\\ 0 & -2 & -1 & -3\\ 0 & 6 & -4 & 2\\ 0 & 4 & p-7 & -2-p \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & -1 & 3 & 4\\ 0 & -2 & -1 & -3\\ 0 & 0 & -7 & -7\\ 0 & 0 & p-9 & -8-p \end{pmatrix} \to \begin{array}{c} \begin{array}{cccc} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \boldsymbol{\beta}_3 & \boldsymbol{\beta}_4 \end{array} \\ \begin{pmatrix} 1 & -1 & 3 & 4\\ 0 & -2 & -1 & -3\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1-2p \end{pmatrix} \end{array}.$$

当 $p=\dfrac{1}{2}$ 时, 向量组 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$ 线性相关, 此时向量组的秩为 3, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 为它的一个极大无关组; 因而 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 也线性相关, 向量组的秩为 3, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 为它的一个极大无关组. 为将 $\boldsymbol{\alpha}_4$ 用该极大无关组线性表示, 再对最后一个矩阵作初等行变换化为行最简形

$$\begin{pmatrix} 1 & -1 & 3 & 4\\ 0 & -2 & -1 & -3\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 & 1\\ 0 & -2 & 0 & -2\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 2\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 \end{pmatrix},$$

由此可知 $\boldsymbol{\alpha}_4 = 2\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$.

问题 3.6　矩阵的等价与向量组的等价有何区别与联系

由于矩阵可看成是由列向量组或行向量组构成的, 所以学习中难免会将矩阵的等价与向量组的等价相混淆. 这里来谈谈它们的区别与联系. 首先给出它们各自的定义:

定义 3.6.1　如果矩阵 $\boldsymbol{A}$ 经过有限次初等变换化为 $\boldsymbol{B}$, 那么称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **等价**.

定义 3.6.2　如果向量组 A 与向量组 B 可以相互线性表出, 就称它们是**等价的向量组**, 也称向量组 A 与 B **等价**.

从定义可以看出: 两个等价的矩阵必须是同型矩阵, 同型矩阵所含列 (行) 向量的个数是相同的 (有限个); 两个等价的向量组必须是同维数的向量组, 但它们所含向量的个数可以不同 (可以有无穷个). 如向量组 $A: \boldsymbol{\alpha} = (0,1)$ 与向量组 $B: \{\boldsymbol{\beta} = k\boldsymbol{\alpha} | k \in \mathbf{R}\}$ 是等价的, 前者只含一个向量, 后者含有无穷多个向量. 所以要讨论矩阵等价与向量组等价之间的联系, 只能在向量组所含向量个数有限的条件下才有意义, 下面就在这种条件下来作些讨论.

为便于叙述, 我们将向量组也作成矩阵 (称为矩阵的行向量组或列向量组), 其具体情况如下:

首先, 若矩阵 $\boldsymbol{A}$ 经初等行变换化为 $\boldsymbol{B}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 行等价, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行向量组也等价; 若 $\boldsymbol{A}$ 经初等列变换化为 $\boldsymbol{C}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 列等价, $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 的列向量组也等价; 若 $\boldsymbol{A}$ 化为 $\boldsymbol{D}$ 既有行变换又有列变换, 则 $\boldsymbol{A}$ 与 $\boldsymbol{D}$ 等价, 但 $\boldsymbol{A}$ 与 $\boldsymbol{D}$ 的行向量组或列向量组就未必等价了. 例如

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \quad 与 \quad \boldsymbol{D} = \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix}$$

是两个等价的矩阵, 但它们的行向量组与列向量组都不等价.

反过来, 设两向量组等价. 若它们所含向量个数不相同, 则分别以两组向量为列 (行) 向量的矩阵不同型, 因而不等价; 若两组向量个数相同, 则分别以两组向量为列 (行) 向量的矩阵列 (行) 等价, 从而一定等价, 但两矩阵不一定行 (列) 等价. 例如

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

的列向量组是等价的, 两个矩阵列等价, 从而等价, 但不行等价.

最后从秩的角度来看看两者的联系与区别. 我们以定理的形式叙述如下:

定理 3.6.1　两同型矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价的充要条件是 $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{B})$.

但两个向量组的秩相等却不是它们等价的充要条件. 例如, 向量组 $A:\boldsymbol{\alpha}_1=(1,0,0),\boldsymbol{\alpha}_2=(0,1,0)$ 与向量组 $B:\boldsymbol{\beta}_1=(0,1,0),\boldsymbol{\beta}_2=(0,0,1)$ 有相同的秩, 但它们不等价.

定理 3.6.2 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的列向量组等价的充要条件是 $\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{B})=\operatorname{rank}(\boldsymbol{A},\boldsymbol{B})$.

将式中的矩阵取转置就可得到行向量组等价的情况.

定理 3.6.1 是我们熟知的结果, 定理 3.6.2 的证明可参见第一章问题 1.5 中的例 1.5.1.

例 3.6.1 设有向量组 $A:\boldsymbol{\alpha}_1=(1,0,-1,-2)^{\mathrm{T}},\boldsymbol{\alpha}_2=(2,1,0,-1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,3,5,8)^{\mathrm{T}}$ 和向量组 $B:\boldsymbol{\beta}_1=(0,2,4,7)^{\mathrm{T}},\boldsymbol{\beta}_2=(1,1,1,1)^{\mathrm{T}},\boldsymbol{\beta}_3=(2,4,6,9)^{\mathrm{T}}$. 试判断这两个向量组是否等价?

解 对 $(\boldsymbol{A},\boldsymbol{B})=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3|\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$ 施行初等行变换:

$$
\begin{aligned}
(\boldsymbol{A},\boldsymbol{B})&=\left(\begin{array}{ccc|ccc}1&2&1&0&1&2\\0&1&3&2&1&4\\-1&0&5&4&1&6\\-2&-1&8&7&1&9\end{array}\right)\to\left(\begin{array}{ccc|ccc}1&2&1&0&1&2\\0&1&3&2&1&4\\0&2&6&4&2&8\\0&3&10&7&3&13\end{array}\right)\\
&\to\left(\begin{array}{ccc|ccc}1&2&1&0&1&2\\0&1&3&2&1&4\\0&0&0&0&0&0\\0&0&1&1&0&1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1&2&1&0&1&2\\0&1&3&2&1&4\\0&0&1&1&0&1\\0&0&0&0&0&0\end{array}\right)\\
&\to\left(\begin{array}{ccc|ccc}1&2&1&0&1&2\\0&1&1&0&1&2\\0&0&1&1&0&1\\0&0&0&0&0&0\end{array}\right).
\end{aligned}
$$

通过最后一个矩阵可以看出

$$\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A},\boldsymbol{B})=3>\operatorname{rank}(\boldsymbol{B})=2,$$

故这两个向量组不等价, 但向量组 $\boldsymbol{B}$ 可由向量组 $\boldsymbol{A}$ 线性表出, 反之不成立.

问题 3.7 线性方程组中的几个问题

(1) 如何理解 Gauss 消元法解线性方程组的正确性

线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 求解的 Gauss (高斯, 1777—1855) 消元法就是将方程组的增广矩阵 $\bar{\boldsymbol{A}}=(\boldsymbol{A},\boldsymbol{b})$ 作初等行变换化为阶梯形 $\bar{\boldsymbol{C}}=(\boldsymbol{C},\boldsymbol{d})$, 然后再求解方

程组 $\boldsymbol{Cx}=\boldsymbol{d}$, 从而得到原方程组解的方法. 在这个过程中我们既没验证方程组 $\boldsymbol{Cx}=\boldsymbol{d}$ 的解就是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解, 又没证明求解过程中方程组的解不会丢失. 换句话说, 为什么说方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 与 $\boldsymbol{Cx}=\boldsymbol{d}$ 一定同解? 这个问题用方程组的向量形式来描述就会看得很清楚.

首先将方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 写成向量形式, 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, 其中 $\boldsymbol{\alpha}_i=(a_{1i},a_{2i},\cdots,a_{mi})^{\mathrm{T}}$ $(i=1,2,\cdots,n)$, 则方程组可写成

$$x_1\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}+x_2\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{m2}\end{pmatrix}+\cdots+x_n\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{mn}\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_m\end{pmatrix} \tag{3.7.1}$$

容易看出, 对 $\bar{\boldsymbol{A}}=(\boldsymbol{A},\boldsymbol{b})$ 作初等行变换等价于对方程组 (3.7.1) 中的方程作相应的运算, 这些运算不会改变 (3.7.1) 式中各列向量的线性关系. 也就是各列向量的系数 $x_1,x_2,\cdots,x_n$ 与初等行变换无关. 另一方面, 初等变换是可逆的, 从而方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 与 $\boldsymbol{Cx}=\boldsymbol{d}$ 在初等行变换下可以互化, 所以它们是同解的.

(2) 为什么要用向量来表示线性方程组的解

线性方程组 (齐次或非齐次) 在有解且解不唯一的时候就会有无穷多解, 如何将无穷多解完全表示出来, 一种很好的方法就是用向量来表示.

首先, 用向量来表示齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解, 使我们知道了它的全部解构成一个向量空间, 一个向量空间最简单最科学的表示方式就是用它的一组基来生成. 如果 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ $(r=\mathrm{rank}(\boldsymbol{A}))$ 是解空间的一组基, 则方程组的全部解就是

$$\boldsymbol{x}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}. \tag{3.7.2}$$

(3.7.2) 式不仅给出了方程组的全部解, 还使我们对解的构成看得很清楚, 这就是常说的**齐次方程组解的结构**.

其次, 用向量来表示非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解, 很容易得到非齐次方程与对应齐次方程组解向量之间的关系, 从而非齐次线性方程组解的问题也就解决了, 其通解为

$$\boldsymbol{x}=\boldsymbol{x}_0+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}, \tag{3.7.3}$$

其中 $\boldsymbol{x}_0$ 是非齐次线性方程组的一个特解. 通常也称 (3.7.3) 为**非齐次方程组解的结构**.

需要注意的是, 非齐次方程组的全部解不构成一个向量空间.

另一方面, 也正是因为用向量来表示线性方程组的解, 我们才有如下十分重要的定理:

定理 3.7.1 若 n 元齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的系数矩阵的秩 $\operatorname{rank}(\boldsymbol{A})=r<n$, 则方程组有基础解系, 且基础解系含有 $n-r$ 个向量.

该定理将向量、方程组、矩阵的秩紧紧地联系在一起, 是线性代数中的一个基本定理. 该定理为许多问题的解决提供了方便.

例 3.7.1 设 $\boldsymbol{A}_{m\times n},\boldsymbol{B}_{n\times l}$ 是两个矩阵, 证明: 若 $\boldsymbol{AB}=\boldsymbol{O}$, 则 $\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B})\leqslant n$.

证 设 $\operatorname{rank}(\boldsymbol{A})=r, B=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l)$, 则

$$\boldsymbol{AB}=(\boldsymbol{A\beta}_1,\boldsymbol{A\beta}_2,\cdots,\boldsymbol{A\beta}_l)=\boldsymbol{O}.$$

这说明 $\boldsymbol{\beta}_i(i=1,2,\cdots,l)$ 是方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解向量, 从而 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l$ 可由方程组的一个基础解系 (含 $n-r$ 个向量) 线性表出, 所以

$$\operatorname{rank}(\boldsymbol{B})=\operatorname{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l)\leqslant n-r=n-\operatorname{rank}(\boldsymbol{A}),$$

即 $\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{B})\leqslant n$.

定理 3.7.1 不仅仅给出了齐次线性方程组基础解系所含向量的个数, 它所揭示的规律对研究线性空间与线性变换也是十分重要的.

比如, 设 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵, 我们将方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解集 $\{\boldsymbol{x}|\boldsymbol{Ax}=\boldsymbol{0},\boldsymbol{x}\in V\}$ 看成是 n 维线性空间 V 上线性变换 $\mathcal{A}:\boldsymbol{x}\mapsto\boldsymbol{Ax}$ 的**零空间** ($\boldsymbol{0}$ 的原像集), 又称为线性变换 $\mathcal{A}$ 的**核**, 记为 $\mathcal{A}^{-1}(\boldsymbol{0})$. 定理 3.7.1 告诉我们, 这个零空间的维数

$$\dim\mathcal{A}^{-1}(\boldsymbol{0})=n-\operatorname{rank}(\boldsymbol{A}).$$

记 $\mathcal{A}V=\{\boldsymbol{Ax}|\boldsymbol{x}\in V\}$, 容易证明 $\mathcal{A}V$ 也是一个线性空间 (称为线性变换 $\mathcal{A}$ 的像空间), 并且 $\operatorname{rank}(\boldsymbol{A})=\dim(\mathcal{A}V)$. 由此可见: $\dim\mathcal{A}^{-1}(\boldsymbol{0})$ 越大, $\dim(\mathcal{A}V)$ 就越小, 经线性变换 $\mathcal{A}$ 所 "丢失" 的信息也就越多. 人们形象地将**零空间**称为在变换 $\mathcal{A}$ 下 "消失" 的空间. 线性变换还有很多性质都与这个空间的维数有关.

例 3.7.2 设 $\boldsymbol{A}$ 是一个 $m\times n$ 实矩阵, 由 $\boldsymbol{A}$ 确定了空间 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的线性变换 $\mathcal{A}:\boldsymbol{x}\to\boldsymbol{Ax}$. 试证:

(1) $\mathcal{A}$ 是单射的充要条件是 $\operatorname{rank}(\boldsymbol{A})=n$ (此时 $n\leqslant m$);

(2) $\mathcal{A}$ 是满射的充要条件是 $\operatorname{rank}(\boldsymbol{A})=m$ (此时 $m\leqslant n$).

证 (1) $\mathcal{A}$ 是单射的充要条件是: $\forall\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in\mathbf{R}^n$, 若 $\boldsymbol{A\alpha}_1=\boldsymbol{A\alpha}_2$, 即 $\boldsymbol{A}(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)=\boldsymbol{0}$, 则 $\boldsymbol{\alpha}_1=\boldsymbol{\alpha}_2$. 这等价于齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解, 故其充要条件是 $\operatorname{rank}(\boldsymbol{A})=n$ (此时 $n\leqslant m$).

(2) $\mathcal{A}$ 是满射的充要条件是: $\forall \boldsymbol{\beta} \in \mathbf{R}^m$, 都 $\exists \boldsymbol{\alpha} \in \mathbf{R}^n$ 使 $\boldsymbol{A}\boldsymbol{\alpha} = \boldsymbol{\beta}$. 这等价于 $\forall \boldsymbol{\beta} \in \mathbf{R}^m$, 非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\beta}$ 都有解.

当 $\operatorname{rank}(\boldsymbol{A}) = m$ 时, 显然 $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}, \boldsymbol{\beta}) = m$, 非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\beta}$ 一定有解;

反过来, 取 $\boldsymbol{\varepsilon}_i$ $(i = 1, 2, \cdots, m)$ 为空间 $\mathbf{R}^m$ 的一组基, 若对每个 $\boldsymbol{\varepsilon}_i$, 方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\varepsilon}_i$ 都有解, 说明向量组 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_m$ 可由矩阵 $\boldsymbol{A}$ 的列向量组线性表出, 所以

$$\operatorname{rank}(\boldsymbol{A}) \geqslant \operatorname{rank}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_m) = m.$$

而 $\boldsymbol{A}$ 是 $m \times n$ 矩阵, 故只可能有 $\operatorname{rank}(\boldsymbol{A}) = m$ (此时 $m \leqslant n$).

(3) 非齐次线性方程组线性无关解向量的个数与通解

我们知道, 若 $\operatorname{rank}(\boldsymbol{A}) = r$ $(r < n)$, 则 n 元齐次方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 有 $n - r$ 个线性无关的解向量, 并且方程组的任一个解都可由这组解向量线性表出, 那么非齐次方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 在有解的情况下, 会有多少个线性无关的解向量? 或者说解向量集合 $S = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ 的秩为多少? S 是否也能由它的一个极大线性无关组线性表出? 下面我们就来讨论这两个问题.

先讨论第一个问题, 即 $S = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ 的秩为多少? 由非齐次方程组解的结构, 我们容易得到下面的结论:

定理 3.7.2　若 n 元非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 有解, 且 $\operatorname{rank}(\boldsymbol{A}) = r < n$, 则其线性无关解向量的最大个数为 $n - r + 1$, 即 $\operatorname{rank}(\{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}) = n - r + 1$.

证　由非齐次线性方程组解的结构知, $S = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ 中任一向量均可表示为

$$\boldsymbol{x} = \boldsymbol{x}_0 + k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{n-r}\boldsymbol{\alpha}_{n-r},$$

其中 $\boldsymbol{x}_0$ 是非齐次线性方程组的一个特解, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{n-r}$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的一个基础解系. 这说明 S 可由向量组 $\boldsymbol{x}_0, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{n-r}$ 线性表出, 所以

$$\operatorname{rank}(S) \leqslant n - r + 1. \tag{3.7.4}$$

另一方面, 易知 $\boldsymbol{x}_0, \boldsymbol{x}_0 + \boldsymbol{\alpha}_1, \boldsymbol{x}_0 + \boldsymbol{\alpha}_2, \cdots, \boldsymbol{x}_0 + \boldsymbol{\alpha}_{n-r}$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的解, 即它们都是 S 中的元素, 下面证明它们是线性无关的.

设有

$$k_0\boldsymbol{x}_0 + k_1(\boldsymbol{x}_0 + \boldsymbol{\alpha}_1) + k_2(\boldsymbol{x}_0 + \boldsymbol{\alpha}_2) + \cdots + k_{n-r}(\boldsymbol{x}_0 + \boldsymbol{\alpha}_{n-r}) = \boldsymbol{0},$$

整理后得

$$(k_0 + k_1 + k_2 + \cdots + k_{n-r})\boldsymbol{x}_0 + k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{n-r}\boldsymbol{\alpha}_{n-r} = \boldsymbol{0}. \tag{3.7.5}$$

等式两端左乘 $\boldsymbol{A}$, 得

$$(k_0+k_1+k_2+\cdots+k_{n-r})\boldsymbol{A}\boldsymbol{x}_0+k_1\boldsymbol{A}\boldsymbol{\alpha}_1+k_2\boldsymbol{A}\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{A}\boldsymbol{\alpha}_{n-r}=\boldsymbol{0}.$$

代入 $\boldsymbol{A}\boldsymbol{x}_0=\boldsymbol{b}$ 与 $\boldsymbol{A}\boldsymbol{\alpha}_i=\boldsymbol{0}\ (i=1,2,\cdots,n-r)$, 得到

$$(k_0+k_1+k_2+\cdots+k_{n-r})\boldsymbol{b}=\boldsymbol{0},$$

由 $\boldsymbol{b}\neq\boldsymbol{0}$ 知

$$k_0+k_1+k_2+\cdots+k_{n-r}=0. \tag{3.7.6}$$

代入 (3.7.5) 式, 得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}=\boldsymbol{0},$$

由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关即知 $k_1=k_2=\cdots=k_{n-r}=0$, 代入 (3.7.6) 式得到 $k_0=0$. 因此

$$k_0=k_1=k_2=\cdots=k_{n-r}=0,$$

所以 $\boldsymbol{x}_0,\boldsymbol{x}_0+\boldsymbol{\alpha}_1,\boldsymbol{x}_0+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{x}_0+\boldsymbol{\alpha}_{n-r}$ 线性无关. 这说明 S 中能找到 $n-r+1$ 个线性无关的向量, 故

$$\operatorname{rank}(S)\geqslant n-r+1.$$

结合 (3.7.4) 得到 $\operatorname{rank}(S)=n-r+1$.

下面再来讨论 S 是否能由它的一个极大线性无关组线性表出? 也就是非齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解能否由它自己的一组解向量线性表示的问题.

我们知道 $S=\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\}$ 不构成一个向量空间, 也就是它对线性运算不封闭, 所以非齐次线性方程组解的线性组合一般不再是它的解. 但这并不意味着我们上面所提问题的答案是否定的, 只要略加一点条件, 就有

定理 3.7.3 设 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l$ 是 n 元非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的 l 个线性无关的解向量, 若 $l=n-\operatorname{rank}(\boldsymbol{A})+1$, 则方程组的通解为

$$\boldsymbol{x}=k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_l\boldsymbol{\beta}_l\quad(k_1+k_2+\cdots+k_l=1). \tag{3.7.7}$$

证 首先, (3.7.7) 给出向量是方程组的解向量. 这是因为

$$\begin{aligned}\boldsymbol{A}(k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_l\boldsymbol{\beta}_l)&=k_1\boldsymbol{A}\boldsymbol{\beta}_1+k_2\boldsymbol{A}\boldsymbol{\beta}_2+\cdots+k_l\boldsymbol{A}\boldsymbol{\beta}_l\\&=(k_1+k_2+\cdots+k_l)\boldsymbol{b}=\boldsymbol{b}.\end{aligned}$$

其次, 证明 $\boldsymbol{\beta}_2-\boldsymbol{\beta}_1,\boldsymbol{\beta}_3-\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l-\boldsymbol{\beta}_1$ 是齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系.

它们是齐次方程组的 $n-\mathrm{rank}(\boldsymbol{A})$ 个解是显然的, 下面只需证明它们线性无关. 设

$$m_1(\boldsymbol{\beta}_2-\boldsymbol{\beta}_1)+m_2(\boldsymbol{\beta}_3-\boldsymbol{\beta}_1)+\cdots+m_{l-1}(\boldsymbol{\beta}_l-\boldsymbol{\beta}_1)=\boldsymbol{0},$$

整理得

$$-(m_1+m_2+\cdots+m_{l-1})\boldsymbol{\beta}_1+m_1\boldsymbol{\beta}_2+m_2\boldsymbol{\beta}_3+\cdots+m_{l-1}\boldsymbol{\beta}_l=\boldsymbol{0}.$$

由于 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_l$ 线性无关, 所以

$$m_1=m_2=\cdots=m_{l-1}=0,$$

所以 $\boldsymbol{\beta}_2-\boldsymbol{\beta}_1,\boldsymbol{\beta}_3-\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_l-\boldsymbol{\beta}_1$ 线性无关, 它们是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的一个基础解系.

方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解为

$$\begin{aligned}\boldsymbol{x}&=\boldsymbol{\beta}_1+t_2(\boldsymbol{\beta}_2-\boldsymbol{\beta}_1)+t_3(\boldsymbol{\beta}_3-\boldsymbol{\beta}_1)+\cdots+t_l(\boldsymbol{\beta}_l-\boldsymbol{\beta}_1)\\&=(1-t_2-t_3-\cdots-t_l)\boldsymbol{\beta}_1+t_2\boldsymbol{\beta}_2+t_3\boldsymbol{\beta}_3+\cdots+t_l\boldsymbol{\beta}_l.\end{aligned}$$

记 $k_1=1-t_2-t_3-\cdots-t_l, k_i=t_i\ (i=2,3,\cdots,l)$, 上式即为 (3.7.7) 式.

例 3.7.3 设 $\boldsymbol{A}$ 是 3 阶非零矩阵, 满足 $\boldsymbol{A}^2=\boldsymbol{O}$, 若非齐次方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解, 那么最多有多少个线性无关的解向量?

解 因为 $\boldsymbol{A}^2=\boldsymbol{O}$, 由例 3.7.1 知

$$2\mathrm{rank}(\boldsymbol{A})\leqslant 3\Rightarrow \mathrm{rank}(\boldsymbol{A})=1.$$

于是方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 最多有 $3-\mathrm{rank}(\boldsymbol{A})+1=3$ 个线性无关的解向量.

例 3.7.4 设 $\boldsymbol{A}$ 是 3 阶非零矩阵, $\boldsymbol{\beta}_1=(1,2,-2)^{\mathrm{T}},\boldsymbol{\beta}_2=(2,1,-1)^{\mathrm{T}},\boldsymbol{\beta}_3=(1,1,t)^{\mathrm{T}}$ 是非齐次方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解向量, 试确定 $\boldsymbol{A}$ 的秩.

解 先求所给解向量组的秩.

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=\begin{pmatrix}1&2&1\\2&1&1\\-2&-1&t\end{pmatrix}\to\begin{pmatrix}1&2&1\\0&-3&-1\\0&3&t+2\end{pmatrix}\to\begin{pmatrix}1&2&1\\0&-3&-1\\0&0&t+1\end{pmatrix},$$

则

$$\mathrm{rank}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=\begin{cases}3, & t\neq -1,\\ 2, & t=-1.\end{cases}$$

由于 $\boldsymbol{A}$ 是 3 阶非零矩阵, 所以 $\mathrm{rank}(\boldsymbol{A})\geqslant 1$, 方程组最多有 $3-\mathrm{rank}(\boldsymbol{A})+1\leqslant 3$ 个线性无关的解向量.

若 $t \neq -1$, 则 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 是方程组的 3 个线性无关的解向量, 此时有

$$3 - \operatorname{rank}(\boldsymbol{A}) + 1 = 3 \Rightarrow \operatorname{rank}(\boldsymbol{A}) = 1.$$

若 $t = -1$, 则题设给出了方程组两个线性无关的解向量, 则有

$$3 - \operatorname{rank}(\boldsymbol{A}) + 1 \geqslant 2 \Rightarrow \operatorname{rank}(\boldsymbol{A}) \leqslant 2,$$

所以 $\operatorname{rank}(\boldsymbol{A}) = 1$ 或 $\operatorname{rank}(\boldsymbol{A}) = 2$.

第四章　特征值与特征向量

问题 4.1　如何理解矩阵特征值与特征向量的几何意义

(1) 实特征值的情况

从映射的角度讲, 我们经常把矩阵看成一种线性映射 (又叫线性变换), 设 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$, 则 $\boldsymbol{A}$ 可以看成 $\mathbf{R}^n$ 到自身的一个映射. 设 $\mathcal{A}: \boldsymbol{\alpha} \mapsto \boldsymbol{\beta} = \boldsymbol{A\alpha}$, 从几何上看, 向量 $\boldsymbol{\beta}$ 是向量 $\boldsymbol{\alpha}$ 绕原点旋转一定角度后再沿径向作伸缩的结果 (如图 4.1.1, $n=2$).

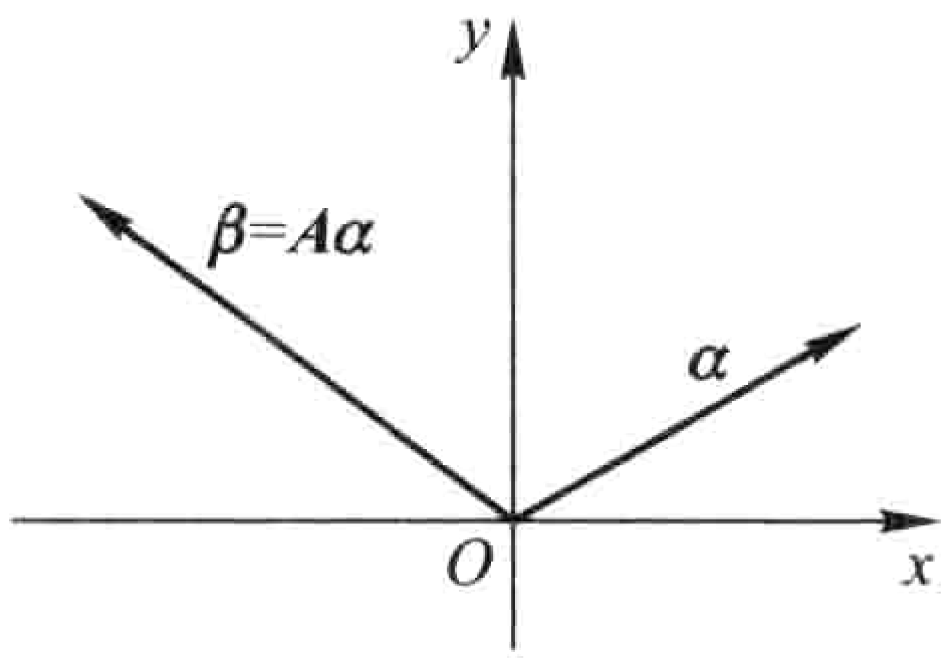


图 4.1.1

如果 $\boldsymbol{\alpha}$ 是 $\boldsymbol{A}$ 的一个特征向量, $\lambda \in \mathbf{R}$ 为对应的特征值, 即有

$$\boldsymbol{A\alpha} = \lambda\boldsymbol{\alpha},$$

则向量 $\boldsymbol{A\alpha}$ 就是 $\boldsymbol{\alpha}$ 沿自身 (或相反) 方向的一个伸缩. 也就是说矩阵 $\boldsymbol{A}$ 的特征向量的方向, 就是线性变换 $\mathcal{A}$ 不变的方向. 利用这种不变性, 可以使许多几何或图形学问题得到简化.

例 4.1.1　将曲面方程 $3x^2+2y^2+z^2-4xy-4yz=1$ 化为标准形, 并说明曲面的类型.

解 曲面方程用矩阵形式表示为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$, 其中 $\boldsymbol{A}=\begin{pmatrix}3&-2&0\\-2&2&-2\\0&-2&1\end{pmatrix}$, $\boldsymbol{x}=(x,y,z)^{\mathrm{T}}$.

可求得 $\boldsymbol{A}$ 的特征值为 $\lambda_1=5$, $\lambda_2=2$, $\lambda_3=-1$, 对应的三个线性无关的单位正交特征向量为 $\boldsymbol{\alpha}_1=\left(\frac{2}{3},-\frac{2}{3},\frac{1}{3}\right)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=\left(\frac{2}{3},\frac{1}{3},-\frac{2}{3}\right)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3=\left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right)^{\mathrm{T}}$.

取正交矩阵 $\boldsymbol{C}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=\begin{pmatrix}\frac{2}{3}&\frac{2}{3}&\frac{1}{3}\\-\frac{2}{3}&\frac{1}{3}&\frac{2}{3}\\\frac{1}{3}&-\frac{2}{3}&\frac{2}{3}\end{pmatrix}$, 作正交变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$, $\boldsymbol{y}=(x',y',z')^{\mathrm{T}}$, 在新坐标系中原曲面方程为

$$5x'^2+2y'^2-z'^2=1,$$

曲面是单叶双曲面.

由于正交变换也就是绕原点的旋转变换, 在新坐标系中三个坐标轴的方向向量 $\boldsymbol{y}_1=(1,0,0)^{\mathrm{T}}$, $\boldsymbol{y}_2=(0,1,0)^{\mathrm{T}}$, $\boldsymbol{y}_3=(0,0,1)^{\mathrm{T}}$ 所对应于原坐标系中的向量刚好就是矩阵 $\boldsymbol{A}$ 的三个特征向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$, 即有 $\boldsymbol{\alpha}_i=\boldsymbol{C}\boldsymbol{y}_i$ $(i=1,2,3)$. 该结论具有普适性, 也就是说, 对于一般二次曲面 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$, 若以 $\boldsymbol{A}$ 的特征方向为坐标轴的方向建立坐标系, 则曲面在该坐标系中的方程形式是最简单的, 也就是曲面的标准方程.

再来谈谈特征值的几何意义:

如果 $\boldsymbol{A}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}$, 且 $\|\boldsymbol{\alpha}\|=1$, 则

$$\|\boldsymbol{A}\boldsymbol{\alpha}\|=|\lambda|\cdot\|\boldsymbol{\alpha}\|=|\lambda|.$$

即特征值的绝对值反映了线性变换在特征方向上的伸缩程度 (伸缩率). 如果 λ_1 是绝对值最大的特征值, 则 $\boldsymbol{A}$ 在对应特征向量方向上的拉伸效果最明显.

利用这一原理, 我们可以求得一般线性变换矩阵 $\boldsymbol{A}$ ($m\times n$ 矩阵) 的最大 (小) 拉伸方向与伸缩率. 其原理为

$$\|\boldsymbol{A}\boldsymbol{x}\|^2=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}, \tag{4.1.1}$$

由于 (4.1.1) 确定的二次型是半正定的, 即 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是半正定矩阵, 则存在正交变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ ($\boldsymbol{C}$ 的列向量是矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的 n 个相互正交的单位特征向量) 将 (4.1.1) 化为标准形

$$\|\boldsymbol{A}\boldsymbol{x}\|^2\xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}}\lambda_1y_1^2+\cdots+\lambda_ny_n^2. \tag{4.1.2}$$

由于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的特征值是非负的, 设有

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0,$$

则

$$\|\boldsymbol{A}\boldsymbol{x}\|^2 \xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}} \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \leqslant \lambda_1 \|\boldsymbol{y}\|.$$

取 $\|\boldsymbol{y}\| = 1$, 有

$$\|\boldsymbol{A}\boldsymbol{x}\|^2 \leqslant \lambda_1.$$

另一方面, 取 $\boldsymbol{\alpha}_1$ 为 λ_1 所对应的单位特征向量, 即 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{\alpha}_1 = \lambda_1\boldsymbol{\alpha}_1, \|\boldsymbol{\alpha}_1\| = 1$, 则有

$$\|\boldsymbol{A}\boldsymbol{\alpha}_1\|^2 = \boldsymbol{\alpha}_1^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{\alpha}_1 = \lambda_1 \Rightarrow \|\boldsymbol{A}\boldsymbol{\alpha}_1\| = \sqrt{\lambda_1}.$$

所以 $\boldsymbol{\alpha}_1$ 的方向就是矩阵 $\boldsymbol{A}$ 所确定的线性变换的最大拉伸方向, $\sqrt{\lambda_1}$ 是最大拉伸率.

(2) 复特征值的情况

我们知道一个实矩阵的特征值有可能是复数, 复特征值所对应的特征向量也自然是复向量. 下面来讨论复特征值与其特征向量所蕴含的几何意义. 这里对 $n=2$ 的情况通过具体的例子来予以说明.

例 4.1.2　求矩阵 $\boldsymbol{A} = \begin{pmatrix} 5 & -6 \\ 7.5 & 11 \end{pmatrix}$ 的特征值与特征向量, 并讨论线性变换 $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 的几何特征.

解　$\boldsymbol{A}$ 的特征多项式为

$$|\lambda\boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda-5 & 6 \\ -7.5 & \lambda-11 \end{vmatrix} = \lambda^2 - 16\lambda + 100,$$

特征值为 $\lambda_1 = \mathbf{8} - 6\mathrm{i}, \lambda_2 = \mathbf{8} + 6\mathrm{i}$.

解齐次线性方程组 $(\lambda_1\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0}$, 可求得 λ_1 所对应的一个特征向量为 $\boldsymbol{\alpha}_1 = \begin{pmatrix} -2-4\mathrm{i} \\ 5 \end{pmatrix}$, 类似可求得 λ_2 所对应的一个特征向量为 $\boldsymbol{\alpha}_2 = \begin{pmatrix} -2+4\mathrm{i} \\ 5 \end{pmatrix}$.

用 $\boldsymbol{\alpha}_1$ 的实部与虚部构造矩阵

$$\boldsymbol{P} = (\operatorname{Re}\boldsymbol{\alpha}_1, \operatorname{Im}\boldsymbol{\alpha}_1) = \begin{pmatrix} -2 & -4 \\ 5 & 0 \end{pmatrix},$$

并令

$$B = P^{-1}AP = \frac{1}{20}\begin{pmatrix} 0 & 4 \\ -5 & -2 \end{pmatrix}\begin{pmatrix} 5 & -6 \\ 7.5 & 11 \end{pmatrix}\begin{pmatrix} -2 & -4 \\ 5 & 0 \end{pmatrix} = \begin{pmatrix} 8 & -6 \\ 6 & 8 \end{pmatrix}$$
$$= |\lambda_1|\begin{pmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{pmatrix}.$$

易见 $\begin{pmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{pmatrix}$ 是一个正交矩阵 (旋转变换), 从而 $\boldsymbol{B}$ 是一个绕原点旋转再以 $|\lambda_1|$ 为伸缩率的变换.

由于

$$\boldsymbol{A} = \boldsymbol{PBP}^{-1},$$

对 $\boldsymbol{x} \in \mathbf{R}^2$, 令 $\boldsymbol{x} = \boldsymbol{Py}$, 从几何角度看, $\boldsymbol{A}$ 的作用相当于先将变量 $\boldsymbol{x}$ 替换成 $\boldsymbol{y}$, 然后旋转、伸缩 ($|\lambda|$ 为其伸缩率), 最后回到 $\boldsymbol{Ax}$ (见图 4.1.2)

$$\begin{tikzcd}
\boldsymbol{x} \arrow[r, "\boldsymbol{A}"] \arrow[d, "\boldsymbol{P}^{-1}", "\text{变量替换}"'] & \boldsymbol{Ax} \\
\boldsymbol{y} \arrow[r, "\boldsymbol{B}", "\text{旋转、伸缩}"'] & \boldsymbol{By} \arrow[u, "\boldsymbol{P}", "\text{变量替换}"']
\end{tikzcd}$$

图 4.1.2

例 4.1.2 的一般情形为: 设 $\boldsymbol{A} \in \mathbf{R}^{2\times 2}$, $\lambda = a - bi$ $(b \neq 0)$ 为复特征值, $\boldsymbol{\alpha}$ 是对应的复特征向量, 则 $\boldsymbol{A} = \boldsymbol{PBP}^{-1}$, 其中 $\boldsymbol{P} = (\mathrm{Re}\,\boldsymbol{\alpha}, \mathrm{Im}\,\boldsymbol{\alpha})$, $\boldsymbol{B} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = |\lambda|\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 是一个旋转与伸缩的复合.

对矩阵 $\dfrac{1}{|\lambda|}\boldsymbol{A}$ 来说, 对应的 $\boldsymbol{B} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 就仅仅是一个旋转变换.

上面结论也可向高维空间作推广.

问题 4.2 若矩阵多项式 $f(\boldsymbol{A}) = \boldsymbol{O}$, 则方程 $f(\lambda) = 0$ 的根与 $\boldsymbol{A}$ 的特征值有怎样的关系

若矩阵多项式 $f(\boldsymbol{A}) = \boldsymbol{O}$, 则 $\boldsymbol{A}$ 的特征值一定是方程 $f(\lambda) = 0$ 的根.

设 $f(\boldsymbol{A})=a_n\boldsymbol{A}^n+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{I}$, $\boldsymbol{A\alpha}=\lambda\boldsymbol{\alpha}$ $(\boldsymbol{\alpha}\neq\boldsymbol{0})$, 则有

$$\begin{aligned}f(\boldsymbol{A})\boldsymbol{\alpha}&=a_n\boldsymbol{A}^n\boldsymbol{\alpha}+\cdots+a_1\boldsymbol{A\alpha}+a_0\boldsymbol{I\alpha}\\&=a_n\lambda^n\boldsymbol{\alpha}+\cdots+a_1\lambda\boldsymbol{\alpha}+a_0\boldsymbol{\alpha}\\&=(a_n\lambda^n+\cdots+a_1\lambda+a_0)\boldsymbol{\alpha}=f(\lambda)\boldsymbol{\alpha}.\end{aligned}\tag{4.2.1}$$

这说明 $f(\lambda)$ 是 $f(\boldsymbol{A})$ 的特征值.

若 $f(\boldsymbol{A})=\boldsymbol{O}$, 即 $f(\boldsymbol{A})$ 是零矩阵, 其特征值只能是零, 故 $f(\lambda)=0$. 所以 $\boldsymbol{A}$ 的特征值一定是方程 $f(\lambda)=0$ 的根.

反之, 若 $f(\boldsymbol{A})=\boldsymbol{O}$, 但方程 $f(\lambda)=0$ 的根却未必是 $\boldsymbol{A}$ 的特征值. 比如, 设 $\boldsymbol{A}=\boldsymbol{I}$, 显然有 $\boldsymbol{A}^2-\boldsymbol{A}=\boldsymbol{O}$, 且 $\lambda=0$ 是方程 $\lambda^2-\lambda=0$ 的根, 但 $\lambda=0$ 显然不是 $\boldsymbol{A}$ 的特征值.

利用以上性质, 我们可以确定矩阵的某些特征值.

例 4.2.1 设 3 阶实对称矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^3+\boldsymbol{A}^2+\boldsymbol{A}=3\boldsymbol{I}$, 求 $\boldsymbol{A}$ 的特征值.

解 设矩阵 $\boldsymbol{A}$ 的特征值为 λ, 则有 $\lambda^3+\lambda^2+\lambda=3$, 即 $(\lambda-1)(\lambda^2+2\lambda+3)=0$. 由于实对称矩阵的特征值是实数, 故 $\lambda^2+2\lambda+3=(\lambda+1)^2+2>0$, 由此可得 $\boldsymbol{A}$ 只有唯一的 3 重特征值 1.

事实上 $\boldsymbol{A}$ 只能是 3 阶单位矩阵, 因为实对称矩阵一定与对角矩阵相似, 即存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{I}$ (因为 $\boldsymbol{A}$ 的特征值全为 1), 于是有 $\boldsymbol{A}=\boldsymbol{PIP}^{-1}=\boldsymbol{I}$.

那么为什么会出现 $\boldsymbol{A}$ 满足方程 $f(\boldsymbol{A})=\boldsymbol{O}$, 但 $f(\lambda)=0$ 的根又不是 $\boldsymbol{A}$ 的特征值的情况?

我们来看下面两种情况:

1° 若 $\boldsymbol{A}$ 满足方程 $g(\boldsymbol{A})=\boldsymbol{O}$, 则对 $\boldsymbol{A}$ 的任意多项式 $h(\boldsymbol{A})$ 均有 $g(\boldsymbol{A})h(\boldsymbol{A})=\boldsymbol{O}$;

2° 由于两个非零矩阵的乘积也可能是零矩阵, 即 $g(\boldsymbol{A})\neq\boldsymbol{O}, h(\boldsymbol{A})\neq\boldsymbol{O}$, 但 $g(\boldsymbol{A})h(\boldsymbol{A})=\boldsymbol{O}$.

无论上面哪种情况, 令 $f=g\cdot h$, 都有 $f(\boldsymbol{A})=\boldsymbol{O}$. 显然 $f(\lambda)=0$ 的根中就会出现不是 $\boldsymbol{A}$ 的特征值的情况. 所以我们不能仅凭方程 $f(\lambda)=0$ 来确定 $\boldsymbol{A}$ 的特征值, 除非 $f(\lambda)$ 是 $\boldsymbol{A}$ 的特征多项式.

从 (4.2.1) 式, 我们还可以看出: 若 $f(\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的多项式, 则 $\boldsymbol{A}$ 的特征向量一定是 $f(\boldsymbol{A})$ 的特征向量, 但其逆命题也是不成立的.

例如: $\boldsymbol{A}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, $f(\boldsymbol{A})=\boldsymbol{A}^2=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, 显然任一 2 维非零向量都是 $f(\boldsymbol{A})$ 特征向量, 但却不一定是 $\boldsymbol{A}$ 的特征向量.

问题 4.3 如何确定一个数 λ 是矩阵 A 的特征值

特征值和特征向量是一对相互依存的概念. 如果 λ 是 $\boldsymbol{A}$ 的一个特征值, 那么至少存在一个非零列向量 $\boldsymbol{\alpha}$ 使得 $\boldsymbol{A\alpha}=\lambda\boldsymbol{\alpha}$, 此时 $\boldsymbol{\alpha}$ 称为属于特征值 λ 的特征向量; 不过属于特征值 λ 的特征向量不是唯一的; 如果 $\boldsymbol{\beta}$ 是 $\boldsymbol{A}$ 的一个特征向量, 那么存在由 $\boldsymbol{A}$ 和 $\boldsymbol{\beta}$ 唯一确定的数 μ 使得 $\boldsymbol{A\beta}=\mu\boldsymbol{\beta}$, 即属于每一个特征向量 $\boldsymbol{\beta}$ 的特征值 μ 是唯一的.

对于一个确定的矩阵 $\boldsymbol{A}$, 其特征值通常都是由特征方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$ 来计算. 可是当 $\boldsymbol{A}$ 的阶数较大, 比如 $n\geqslant 5$ 时, 由于一元 n 次方程没有一般的求根公式, 所以在实际应用中, 通常需要借助于数值计算方法求近似解. 但是作为理论上的判断或对一些特殊情况的处理, 了解以下等价命题是十分必要的. 在很多时候, 这些命题能帮助我们间接地找到矩阵的某些特征值.

λ 是 $\boldsymbol{A}$ 的特征值 $\Leftrightarrow$ 存在非零向量 $\boldsymbol{\alpha}$ 使得 $\boldsymbol{A\alpha}=\lambda\boldsymbol{\alpha}$

$$\begin{aligned}
&\Leftrightarrow \text{齐次线性方程组 } (\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0} \text{ 有非零解}\\
&\Leftrightarrow |\lambda\boldsymbol{I}-\boldsymbol{A}|=0\\
&\Leftrightarrow \lambda\boldsymbol{I}-\boldsymbol{A} \text{ 不可逆}\\
&\Leftrightarrow \operatorname{rank}(\lambda\boldsymbol{I}-\boldsymbol{A})<n \ (n \text{ 为 } \boldsymbol{A} \text{ 的阶数}).
\end{aligned}$$

另一方面, 前面给出的 (4.2.1) 式以及相似矩阵的性质 "相似矩阵必有相同的特征值." 也可以帮助我们确定某些特征值.

例 4.3.1 已知矩阵 $2\boldsymbol{A}+3\boldsymbol{I}$ 不可逆, 求矩阵 $\boldsymbol{B}=2\boldsymbol{A}^2+\boldsymbol{A}-\boldsymbol{I}$ 的一个特征值.

解 由于 $2\boldsymbol{A}+3\boldsymbol{I}$ 不可逆, 则

$$|2\boldsymbol{A}+3\boldsymbol{I}|=0\Rightarrow\left|-\frac{3}{2}\boldsymbol{I}-\boldsymbol{A}\right|=0,$$

所以 $\lambda=-\dfrac{3}{2}$ 是 $\boldsymbol{A}$ 的一个特征值, $\boldsymbol{B}$ 的一个特征值为 $2\left(-\dfrac{3}{2}\right)^2-\dfrac{3}{2}-1=2$.

例 4.3.2 设 3 阶矩阵 $\boldsymbol{A}$ 的各行元之和均为 3, 且线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有两个线性无关的解向量, 求 $\boldsymbol{A}$ 的特征值.

解 由 $\boldsymbol{A}$ 的各行元之和均为 3, 知

$$\boldsymbol{A}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}3\\3\\3\end{pmatrix}=3\begin{pmatrix}1\\1\\1\end{pmatrix},$$

故 $\lambda=3$ 是 $\boldsymbol{A}$ 的一个特征值.

又 $\boldsymbol{Ax}=\boldsymbol{0}$ 有两个线性无关的解向量, 故 $\operatorname{rank}(\boldsymbol{A})=1$, 从而 $\lambda=0$ 是 $\boldsymbol{A}$ 的一个二重特征值.

例 4.3.3 已知 3 阶矩阵 $\boldsymbol{A}$ 与 3 维列向量 $\boldsymbol{x}$, 使得向量组 $\boldsymbol{x}$, $\boldsymbol{Ax}$, $\boldsymbol{A}^2\boldsymbol{x}$ 线性无关, 且满足 $\boldsymbol{A}^3\boldsymbol{x}=3\boldsymbol{Ax}-2\boldsymbol{A}^2\boldsymbol{x}$, 求 $\boldsymbol{A}$ 的特征值.

解 作矩阵 $\boldsymbol{P}=(\boldsymbol{x},\boldsymbol{Ax},\boldsymbol{A}^2\boldsymbol{x})$, 因为向量组 $\boldsymbol{x},\boldsymbol{Ax},\boldsymbol{A}^2\boldsymbol{x}$ 线性无关, 所以 $\boldsymbol{P}$ 可逆, 且有

$$\begin{aligned}\boldsymbol{AP}&=(\boldsymbol{Ax},\boldsymbol{A}^2\boldsymbol{x},\boldsymbol{A}^3\boldsymbol{x})=(\boldsymbol{Ax},\boldsymbol{A}^2\boldsymbol{x},3\boldsymbol{Ax}-2\boldsymbol{A}^2\boldsymbol{x})\\&=(\boldsymbol{x},\boldsymbol{Ax},\boldsymbol{A}^2\boldsymbol{x})\begin{pmatrix}0&0&0\\1&0&3\\0&1&-2\end{pmatrix}=\boldsymbol{PB},\end{aligned}$$

其中 $\boldsymbol{B}=\begin{pmatrix}0&0&0\\1&0&3\\0&1&-2\end{pmatrix}$.

则 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$, 即 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 从而有相同的特征值. 易求得

$$|\lambda\boldsymbol{I}-\boldsymbol{B}|=\lambda(\lambda-1)(\lambda+3),$$

所以 $\boldsymbol{A}$ 的特征值为 $0,1,-3$.

问题 4.4 如何理解矩阵特征多项式的系数

设矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$,我们容易得到 $\boldsymbol{A}$ 的特征多项式为

$$\begin{aligned}f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|&=\begin{vmatrix}\lambda-a_{11}&-a_{12}&\cdots&-a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&-a_{2n}\\\vdots&\vdots&&\vdots\\-a_{n1}&-a_{n2}&\cdots&\lambda-a_{nn}\end{vmatrix}\\&=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n|\boldsymbol{A}|.\end{aligned}\tag{4.4.1}$$

但要问及 λ^k $(k=1,2,\cdots,n-2)$ 项的系数是多少, 则是一个较困难的事. 事实上关于 n 阶矩阵 $\boldsymbol{A}$ 的特征多项式有以下结论:

定理 4.4.1 设 $\boldsymbol{A}$ 是 n 阶矩阵, 则 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=\lambda^n-a_1\lambda^{n-1}+\cdots+(-1)^ka_k\lambda^{n-k}+\cdots+(-1)^na_n,\tag{4.4.2}$$

其中 a_k 是 $\boldsymbol{A}$ 的所有 k 阶主子式之和. 即

$$a_k=\sum_{1\leqslant j_1<\cdots<j_k\leqslant n}\begin{vmatrix} a_{j_1j_1} & \cdots & a_{j_1j_k}\\ \vdots & & \vdots\\ a_{j_kj_1} & \cdots & a_{j_kj_k}\end{vmatrix},$$

特别 $a_1=a_{11}+a_{22}+\cdots+a_{nn}$, $a_n=|\boldsymbol{A}|$.

证 下面用两种方法来证明.

方法 1 我们先观察 $n=3$ 的情形, 利用行列式的性质有

$$\begin{aligned}
f(\lambda)&=\begin{vmatrix}\lambda-a_{11} & 0-a_{12} & 0-a_{13}\\ 0-a_{21} & \lambda-a_{22} & 0-a_{23}\\ 0-a_{31} & 0-a_{32} & \lambda-a_{33}\end{vmatrix}\\
&=\begin{vmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{vmatrix}+\begin{vmatrix}-a_{11} & 0 & 0\\ -a_{21} & \lambda & 0\\ -a_{31} & 0 & \lambda\end{vmatrix}+\begin{vmatrix}\lambda & -a_{12} & 0\\ 0 & -a_{22} & 0\\ 0 & -a_{32} & \lambda\end{vmatrix}+\begin{vmatrix}\lambda & 0 & -a_{13}\\ 0 & \lambda & -a_{23}\\ 0 & 0 & -a_{33}\end{vmatrix}+\\
&\quad\begin{vmatrix}-a_{11} & -a_{12} & 0\\ -a_{21} & -a_{22} & 0\\ -a_{31} & -a_{32} & \lambda\end{vmatrix}+\begin{vmatrix}-a_{11} & 0 & -a_{13}\\ -a_{21} & \lambda & -a_{23}\\ -a_{31} & 0 & -a_{33}\end{vmatrix}+\begin{vmatrix}\lambda & -a_{12} & -a_{13}\\ 0 & -a_{22} & -a_{23}\\ 0 & -a_{32} & -a_{33}\end{vmatrix}+|-\boldsymbol{A}|\\
&=\lambda^3-(a_{11}+a_{22}+a_{33})\lambda^2+\left(\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix}a_{11} & a_{13}\\ a_{31} & a_{33}\end{vmatrix}+\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}\right)\lambda-|\boldsymbol{A}|.
\end{aligned}$$

定理成立.

对 n 阶矩阵的情况, 将 $\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}$ 按列分块, 记 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, $\lambda\boldsymbol{I}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$ ($\boldsymbol{\beta}_k$ 是第 k 个分量为 λ, 其他分量全为 0 的列向量), 由行列式的性质 (与 $n=3$ 的情形相仿)

$$\begin{aligned}
&f(\lambda)\\
&=|\lambda\boldsymbol{I}-\boldsymbol{A}|=|\boldsymbol{\beta}_1-\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2-\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\beta}_n-\boldsymbol{\alpha}_n|\\
&=|\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n|-\sum_{j_1=1}^{n}|\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{j_1-1},\boldsymbol{\alpha}_{j_1},\boldsymbol{\beta}_{j_1+1},\cdots,\boldsymbol{\beta}_n|+\\
&\quad(-1)^2\sum_{1\leqslant j_1<j_2\leqslant n}|\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{j_1-1},\alpha_{j_1},\boldsymbol{\beta}_{j_1+1},\cdots,\boldsymbol{\beta}_{j_2-1},\boldsymbol{\alpha}_{j_2},\boldsymbol{\beta}_{j_2+1},\cdots,\boldsymbol{\beta}_n|+\cdots+\\
&\quad(-1)^{n-1}\sum_{1\leqslant j_1<\cdots<j_{n-1}\leqslant n}|\boldsymbol{\alpha}_{j_1},\cdots,\boldsymbol{\beta}_{j_n},\cdots,\boldsymbol{\alpha}_{j_{n-1}}|+\\
&\quad(-1)^n|\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n|.
\end{aligned}\tag{4.4.3}$$

利用行列式的 Laplace 展开定理将 (4.4.3) 式中的各行列式按 $\boldsymbol{\alpha}_{j_1}, \cdots, \boldsymbol{\alpha}_{j_{n-1}}$ 所在列展开, 就得到定理的结论.

方法 2　(4.4.2) 式右边对 λ 求 $n-k$ 阶导数, 有

$$f^{(n-k)}(\lambda) = n(n-1)\cdots(k+1)\lambda^k + \cdots + (-1)^k(n-k)!a_k.$$

所以

$$f^{(n-k)}(0) = (-1)^k(n-k)!a_k. \tag{4.4.4}$$

(4.4.2) 式左边对 λ 求导数, 有

$$f'(\lambda) = \sum_{j_1=1}^{n} \begin{array}{c} \begin{vmatrix} \lambda - a_{11} & \cdots & 0 & \cdots & -a_{1n} \\ \vdots & & \vdots & & \vdots \\ -a_{j_1 1} & \cdots & 1 & \cdots & -a_{j_1 n} \\ \vdots & & \vdots & & \vdots \\ -a_{n1} & \cdots & 0 & \cdots & \lambda - a_{nn} \end{vmatrix} \\ \text{第 } j_1 \text{ 列} \end{array}. \tag{4.4.5}$$

右边各行列式的第 j_1 列中仅第 j_1 个元为 1 其余全为 0, 按第 j_1 列展开, 记为

$$\begin{array}{cl} \left|\begin{array}{c:c} \lambda - a_{11} & \\ \hdashline & \lambda - a_{nn} \end{array}\right| & \text{划掉第 } j_1 \text{ 行} \\ \text{划掉第 } j_1 \text{ 列} & \end{array}.$$

(4.4.5) 再求导一次, 有

$$\begin{aligned} f''(\lambda) &= \sum_{j_1 \neq j_2} \begin{array}{cl} \left|\begin{array}{c:c:c} \lambda - a_{11} & & \\ \hdashline \hdashline & & \lambda - a_{nn} \end{array}\right| & \begin{array}{l} \text{第 } j_1 \text{ 行} \\ \text{第 } j_2 \text{ 行} \end{array} \\ \text{第 } j_1 \text{ 列} \quad \text{第 } j_2 \text{ 列} & \end{array} \\ &= 2 \sum_{1 \leqslant j_1 < j_2 \leqslant n} \begin{array}{cl} \left|\begin{array}{c:c:c} \lambda - a_{11} & & \\ \hdashline \hdashline & & \lambda - a_{nn} \end{array}\right| & \begin{array}{l} \text{第 } j_1 \text{ 行} \\ \text{第 } j_2 \text{ 行} \end{array} \\ \text{第 } j_1 \text{ 列} \quad \text{第 } j_2 \text{ 列} & \end{array}. \end{aligned}$$

这样对 (4.4.2) 式左边求 $n-k$ 阶导数, 有

$$
\begin{aligned}
f^{(n-k)}(\lambda) &= (n-k)!\sum[\text{从 } |\lambda \boldsymbol{I}-\boldsymbol{A}| \text{ 中划掉 } (n-k) \text{ 行和 } (n-k) \text{ 列余下的子式}]\\
&= (n-k)! \sum_{1\leqslant j_1<\cdots<j_k\leqslant n} \begin{vmatrix} \lambda - a_{j_1j_1} & \cdots & -a_{j_1j_k} \\ \vdots & & \vdots \\ -a_{j_kj_1} & \cdots & \lambda - a_{j_kj_k} \end{vmatrix}.
\end{aligned}
$$

所以

$$
f^{(n-k)}(0) = (-1)^k (n-k)! \sum_{1\leqslant j_1<\cdots<j_k\leqslant n} \begin{vmatrix} a_{j_1j_1} & \cdots & a_{j_1j_k} \\ \vdots & & \vdots \\ a_{j_kj_1} & \cdots & a_{j_kj_k} \end{vmatrix}. \tag{4.4.6}
$$

比较 (4.4.4) 与 (4.4.6) 就得到 $a_k = \displaystyle\sum_{1\leqslant j_1<\cdots<j_k\leqslant n} \begin{vmatrix} a_{j_1j_1} & \cdots & a_{j_1j_k} \\ \vdots & & \vdots \\ a_{j_kj_1} & \cdots & a_{j_kj_k} \end{vmatrix}$, 定理得证.

利用定理 4.4.1, 我们可立即得到:

推论 若 n 阶矩阵 $\boldsymbol{A}$ 的秩 $\operatorname{rank}(\boldsymbol{A}) = r < n$, 则 $\lambda = 0$ 至少是 $\boldsymbol{A}$ 的 $n-r$ 重特征根.

证 因为 $\operatorname{rank}(\boldsymbol{A}) = r < n$, 则 (4.4.2) 式中 $a_{r+1} = \cdots = a_n = 0$, 从而特征多项式

$$
\begin{aligned}
f(\lambda) &= \lambda^n - a_1\lambda^{n-1} + a_2\lambda^{n-2} - \cdots + (-1)^r a_r \lambda^{n-r}\\
&= \lambda^{n-r}[\lambda^r - a_1\lambda^{r-1} + a_2\lambda^{r-2} - \cdots + (-1)^r a_r],
\end{aligned}
$$

所以 $\lambda = 0$ 至少是 $\boldsymbol{A}$ 的 $n-r$ 重特征根.

(4.4.2) 式对计算低秩 (秩为 1 或 2) 矩阵的特征多项式尤为方便. 比如

$$
\boldsymbol{A} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 2 & 4 & \cdots & 2n \\ \vdots & \vdots & & \vdots \\ n & 2n & \cdots & n^2 \end{pmatrix},
$$

有 $\operatorname{rank}(\boldsymbol{A}) = 1$, 则 $|\lambda \boldsymbol{I} - \boldsymbol{A}| = \lambda^n - \left(\displaystyle\sum_{k=1}^n k^2\right)\lambda^{n-1}$.

需要指出的是, 由于相似矩阵有相同的特征多项式, 因而矩阵的各阶主子式之和是相似条件下的不变量.

问题 4.5 设 A 是 $m\times n$ 矩阵, B 是 $n\times m$ 矩阵, 则 AB 与 BA 是否有相同的特征值

当 $m=n$, 且 A 可逆 (或 B 可逆) 时, $BA=A^{-1}(AB)A$ 与 AB 相似, 此时它们有相同的特征值; 当 A 与 B 均不可逆时, AB 与 BA 可能不再相似, 比如 $A=\begin{pmatrix}1&1\\0&0\end{pmatrix}$, $B=\begin{pmatrix}0&0\\0&1\end{pmatrix}$, 有 $AB=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $BA=\begin{pmatrix}0&0\\0&0\end{pmatrix}$, 显然 AB 与 BA 不相似, 但它们却有相同的特征值.

对更一般的情况, 其结果怎样? 即 AB 与 BA 是否有相同的特征值? 对此有以下结论:

定理 4.5.1 设 A 是 $m\times n$ 矩阵, B 是 $n\times m$ 矩阵, $m\geqslant n$, 则

$$|\lambda I-AB|=\lambda^{m-n}|\lambda I-BA|. \tag{4.5.1}$$

证 构造 $m+n$ 阶矩阵 $\begin{pmatrix}AB&O\\B&O\end{pmatrix}$ 与 $\begin{pmatrix}O&O\\B&BA\end{pmatrix}$, 下面证明它们是相似的. 由于

$$\begin{aligned}&\begin{pmatrix}I&-A\\O&I\end{pmatrix}\begin{pmatrix}AB&O\\B&O\end{pmatrix}\begin{pmatrix}I&-A\\O&I\end{pmatrix}^{-1}\\=&\begin{pmatrix}I&-A\\O&I\end{pmatrix}\begin{pmatrix}AB&O\\B&O\end{pmatrix}\begin{pmatrix}I&A\\O&I\end{pmatrix}=\begin{pmatrix}O&O\\B&BA\end{pmatrix},\end{aligned}$$

所以 $\begin{pmatrix}AB&O\\B&O\end{pmatrix}$ 与 $\begin{pmatrix}O&O\\B&BA\end{pmatrix}$ 相似, 它们有相同的特征多项式

$$\begin{aligned}&\begin{vmatrix}\lambda I-AB&O\\-B&\lambda I\end{vmatrix}=\begin{vmatrix}\lambda I&O\\-B&\lambda I-BA\end{vmatrix}\\\Rightarrow\ &|\lambda I-AB|\cdot\lambda^n=\lambda^m\cdot|\lambda I-BA|\\\Rightarrow\ &|\lambda I-AB|=\lambda^{m-n}\cdot|\lambda I-BA|.\end{aligned}$$

(4.5.1) 式表明 AB 与 BA 的特征多项式只相差一个 λ^{m-n} 因子, 因而有相同的非零特征值.

当 $m=n$ 时, (4.5.1) 式为

$$|\lambda I-AB|=|\lambda I-BA|. \tag{4.5.2}$$

(4.5.2) 式还可用以下方法证明:

当 $|\boldsymbol{A}|\,|\boldsymbol{B}| \neq 0$ 时, 有

$$
\begin{aligned}
&|\boldsymbol{AB}|\,|\lambda \boldsymbol{I} - \boldsymbol{AB}| = \left|\lambda \boldsymbol{AB} - (\boldsymbol{AB})^2\right| = |\boldsymbol{A}|\,|\lambda \boldsymbol{I} - \boldsymbol{BA}|\,|\boldsymbol{B}| \\
\Rightarrow\ & |\lambda \boldsymbol{I} - \boldsymbol{AB}| = |\lambda \boldsymbol{I} - \boldsymbol{BA}|\,.
\end{aligned}
$$

当 $|\boldsymbol{A}|\,|\boldsymbol{B}| = 0$ 时, 由于矩阵的特征值只有有限个, 故 $\exists \delta > 0$, 使得 $\forall t \in (0, \delta)$ 有 $|t\boldsymbol{I} - \boldsymbol{A}| \neq 0$ 且 $|t\boldsymbol{I} - \boldsymbol{B}| \neq 0$, 利用上面结论有

$$
|\lambda \boldsymbol{I} - (t\boldsymbol{I} - \boldsymbol{A})(t\boldsymbol{I} - \boldsymbol{B})| = |\lambda \boldsymbol{I} - (t\boldsymbol{I} - \boldsymbol{B})(t\boldsymbol{I} - \boldsymbol{A})|\,,
$$

取 $t \to 0$ 便得 (4.5.2) 式.

结合定理 4.4.1 我们还能得到结论: **$\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 具有相同的各阶主子式之和.**

例 4.5.1 求下列矩阵的特征值及各阶主子式之和.

$$
\boldsymbol{A} = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & -2 & -1 & 2 \\ 2 & -3 & -2 & 3 \\ 2 & -2 & -2 & 2 \\ 1 & -1 & -1 & 1 \end{pmatrix}.
$$

解 (1) 由于

$$
\boldsymbol{A} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} \triangleq \boldsymbol{A}_1\boldsymbol{A}_2,
$$

$$
\boldsymbol{A}_2\boldsymbol{A}_1 = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (a_1b_1 + a_2b_2 + a_3b_3)\,.
$$

而 $\boldsymbol{A}_2\boldsymbol{A}_1$ 的特征值及 1 阶主子式均为 $a_1b_1 + a_2b_2 + a_3b_3$ (无 2, 3 阶主子式, 视其值为 0), 所以 $\boldsymbol{A} = \boldsymbol{A}_1\boldsymbol{A}_2$ 的三个特征值分别为 $a_1b_1 + a_2b_2 + a_3b_3$, 0, 0; 它的 1, 2, 3 阶主子式之和也分别为 $a_1b_1 + a_2b_2 + a_3b_3, 0, 0$.

(2) 由第一章的例 1.2.5 我们知道

$$
\boldsymbol{B} = \begin{pmatrix} 1 & -2 \\ 2 & -3 \\ 2 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \triangleq \boldsymbol{B}_1\boldsymbol{B}_2.
$$

由于 $\boldsymbol{B}_2\boldsymbol{B}_1=\begin{pmatrix}-1 & 0\\ 1 & -2\end{pmatrix}$, 其特征值为 $-1,-2$, 所以 $\boldsymbol{B}=\boldsymbol{B}_1\boldsymbol{B}_2$ 的全部特征值为 $-1,-2,0,0$.

$\boldsymbol{B}$ 的 1 阶与 2 阶主子式之和分别为

$$\operatorname{tr}(\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}_2\boldsymbol{B}_1)=-1-2=-3,\det(\boldsymbol{B}_2\boldsymbol{B}_1)=\begin{vmatrix}-1 & 0\\ 1 & -2\end{vmatrix}=2;$$

$\boldsymbol{B}$ 的 3 阶与 4 阶主子式之和均为 0.

问题 4.6 如何理解矩阵的不同特征值所对应的特征向量的线性无关性

我们知道矩阵的任一特征值 λ 所对应的全体特征向量 $V_\lambda=\{\boldsymbol{\alpha}|\boldsymbol{A}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}\}$ 构成一个线性空间, 称为特征子空间. 通常的线性代数教材都会给出如下结论:

定理 4.6.1 设 $\boldsymbol{A}$ 是 n 阶方阵, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是分别属于互异特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 的特征向量, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关.

较定理 4.6.1 更一般的结论是

定理 4.6.2 设 $\boldsymbol{A}$ 是 n 阶方阵, $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $\boldsymbol{A}$ 的互异特征值, 若 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir_i}\ (i=1,2,\cdots,s)$ 是属于 λ_i 的线性无关的特征向量, 则 $\boldsymbol{\alpha}_{11},\cdots,\boldsymbol{\alpha}_{1r_1},\boldsymbol{\alpha}_{21},\cdots,\boldsymbol{\alpha}_{2r_2},\cdots,\boldsymbol{\alpha}_{s1},\cdots,\boldsymbol{\alpha}_{sr_s}$ 线性无关.

上面两定理用特征子空间的语言来描述便是

1° 从不同特征子空间中各取一个非零向量所构成的向量组是线性无关的;

2° 从不同特征子空间中各取一组线性无关的向量, 则全体向量构成的向量组是线性无关的.

结论 2° 的更一般形式是: 矩阵不同特征子空间的和是直和.

从特征子空间的角度来看特征向量的线性无关性可使问题更加清晰、本质, 使问题的证明思路更为简捷. 比如, 通常教材都会对定理 4.6.1 给出证明 (数学归纳法), 同时指出可类似证明定理 4.6.2 (证明通常略). 事实上用归纳法来证明定理 4.6.2 是很麻烦的, 若用特征子空间直和的观点来看则十分简单, 具体证法如下:

设

$$c_{11}\boldsymbol{\alpha}_{11}+\cdots+c_{1r_1}\boldsymbol{\alpha}_{1r_1}+c_{21}\boldsymbol{\alpha}_{21}+\cdots+c_{2r_2}\boldsymbol{\alpha}_{2r_2}+\cdots+c_{s1}\boldsymbol{\alpha}_{s1}+\cdots+c_{sr_s}\boldsymbol{\alpha}_{sr_s}=\mathbf{0} \tag{4.6.1}$$

记 $\boldsymbol{\beta}_i = c_{i1}\boldsymbol{\alpha}_{i1} + \cdots + c_{ir_i}\boldsymbol{\alpha}_{ir_i}$ $(i=1,2,\cdots,s)$, 则 $\boldsymbol{\beta}_i$ 是特征子空间 V_{λ_i} 中的元素 $(i=1,2,\cdots,s)$. 此时 (4.6.1) 式写为

$$\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 + \cdots + \boldsymbol{\beta}_s = \mathbf{0}. \tag{4.6.2}$$

由定理 4.6.1 知, (4.6.2) 成立的充分必要条件是 $\boldsymbol{\beta}_i = \mathbf{0}$ $(i=1,2,\cdots,s)$, 即

$$c_{i1}\boldsymbol{\alpha}_{i1} + \cdots + c_{ir_i}\boldsymbol{\alpha}_{ir_i} = \mathbf{0}$$

再由 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir_i}$ 线性无关知

$$c_{ij} = 0 \ (i=1,2,\cdots,s;\ j=1,\cdots,r_i),$$

所以 $\boldsymbol{\alpha}_{11},\cdots,\boldsymbol{\alpha}_{1r_1},\boldsymbol{\alpha}_{21},\cdots,\boldsymbol{\alpha}_{2r_2},\cdots,\boldsymbol{\alpha}_{s1},\cdots,\boldsymbol{\alpha}_{sr_s}$ 线性无关.

从直和的观点来看, 下面的结论也是显然的:

矩阵 $\boldsymbol{A}$ 不同特征值所对应的特征向量的和一定不是 $\boldsymbol{A}$ 的特征向量.

问题 4.7 什么是特征值的代数重数与几何重数? 二者有何关系

设 $\boldsymbol{A}$ 是 n 阶矩阵, 若 λ_0 是特征多项式 $f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|$ 的 k 重根, 即

$$f(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{A}|=(\lambda-\lambda_0)^k g(\lambda), \quad 且\ g(\lambda_0)\neq 0,$$

则称数 k 为特征值 λ_0 的**代数重数**.

特征值 λ_0 所对应的线性无关的特征向量的最大个数称为 λ_0 的**几何重数**. 换言之, λ_0 的几何重数就是其特征子空间 $V_{\lambda_0}=\{\boldsymbol{\alpha}|\boldsymbol{A\alpha}=\lambda_0\boldsymbol{\alpha}\}$ 的维数.

注意到 V_{λ_0} 即齐次方程组 $(\lambda_0\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\mathbf{0}$ 的解空间, 因此

$$\lambda_0\ 的几何重数 = n - \operatorname{rank}(\lambda_0\boldsymbol{I}-\boldsymbol{A}).$$

对任一特征值 λ_0 来说, 至少有一个属于 λ_0 的特征向量, 即 V_{λ_0} 中含有非零向量, 因此 λ_0 的几何重数至少为 1. 那么特征值的代数重数与几何重数具有怎样的关系? 我们先看几个具体的例子.

例 4.7.1 设对角矩阵 $\boldsymbol{A}=\begin{pmatrix}5&&&\\&2&&\\&&2&\\&&&2\end{pmatrix}$, 显然 $\boldsymbol{A}$ 的特征值为 5, 2, 它们的重数 (在不加说明的情况下, 特征值的重数总是指代数重数) 分别为 1, 3. 显然 $\operatorname{rank}(5\boldsymbol{I}-\boldsymbol{A})=3$, 因此特征值 5 的几何重数为 $4-3=1$; $\operatorname{rank}(2\boldsymbol{I}-\boldsymbol{A})=1$, 因此特征值 2 的几何重数为 $4-1=3$.

容易证明, 对角矩阵任一特征值的几何重数与代数重数都是相等的.

例 4.7.2 设

$$\boldsymbol{A}=\begin{pmatrix}7&1&0&0&0\\0&7&0&0&0\\0&0&9&1&0\\0&0&0&9&1\\0&0&0&0&9\end{pmatrix},$$

易见 $\boldsymbol{A}$ 的特征值为 7 (2 重), 9 (3 重).

由 $\operatorname{rank}(7\boldsymbol{I}-\boldsymbol{A})=4$, 知 2 重特征值 7 的几何重数为 1; 由 $\operatorname{rank}(9\boldsymbol{I}-\boldsymbol{A})=4$, 知 3 重特征值 9 的几何重数也为 1.

一般地, n 阶 Jordan 块矩阵

$$J_{a,\,n}=\begin{pmatrix}a&1&&\boldsymbol{O}\\&a&\ddots&\\&&\ddots&1\\\boldsymbol{O}&&&a\end{pmatrix}$$

具有一个 n 重特征值 a, 其几何重数为 1.

由上面的例子可见, 矩阵 $\boldsymbol{A}$ 的特征值 λ 的代数重数与几何重数一般来说是不相等的, 它们之间的关系可以叙述为如下定理:

定理 4.7.1 矩阵的任一特征值的几何重数都小于或等于其代数重数.

证 设 λ_0 是 n 阶矩阵 $\boldsymbol{A}$ 的 s 重特征值, 几何重数为 k, 下面来证明 $k\leqslant s$.

选取 V_{λ_0} 的一组基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k$, 并将其扩充为 V 的一组基

$$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k+1},\cdots,\boldsymbol{\beta}_n,$$

则有 $\boldsymbol{A\alpha}_1=\lambda_0\boldsymbol{\alpha}_1,\cdots,\boldsymbol{A\alpha}_k=\lambda_0\boldsymbol{\alpha}_k$, 而 $\boldsymbol{A\beta}_{k+1},\cdots,\boldsymbol{A\beta}_n$ 可以由如上基向量组线性表出, 不妨设

$$\boldsymbol{A\beta}_i=c_{1i}\boldsymbol{\alpha}_1+\cdots+c_{ki}\boldsymbol{\alpha}_k+c_{(k+1)i}\boldsymbol{\beta}_{k+1}+\cdots+c_{ni}\boldsymbol{\beta}_n\ (i=k+1,\cdots,n),$$

于是有

$$
\begin{aligned}
&\boldsymbol{A}(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k+1},\cdots,\boldsymbol{\beta}_n)\\
=&(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k+1},\cdots,\boldsymbol{\beta}_n)\begin{pmatrix}
\lambda_0 & & & c_{1,k+1} & \cdots & c_{1n}\\
& \ddots & & \vdots & & \vdots\\
& & \lambda_0 & c_{k,k+1} & \cdots & c_{kn}\\
& & & c_{k+1,k+1} & \cdots & c_{k+1,n}\\
& \boldsymbol{O} & & \vdots & & \vdots\\
& & & c_{n,k+1} & \cdots & c_{nn}
\end{pmatrix}
\end{aligned}
\tag{4.7.1}
$$

记

$$
\boldsymbol{P}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k+1},\cdots,\boldsymbol{\beta}_n),\quad \lambda_0\boldsymbol{I}=\begin{pmatrix}\lambda_0 & & \\ & \ddots & \\ & & \lambda_0\end{pmatrix},
$$

$$
\boldsymbol{C}_1=\begin{pmatrix}c_{1,k+1} & \cdots & c_{1n}\\ \vdots & & \vdots\\ c_{k,k+1} & \cdots & c_{kn}\end{pmatrix},\boldsymbol{C}_2=\begin{pmatrix}c_{k+1,k+1} & \cdots & c_{k+1,n}\\ \vdots & & \vdots\\ c_{n,k+1} & \cdots & c_{nn}\end{pmatrix},\boldsymbol{T}=\begin{pmatrix}\lambda_0\boldsymbol{I} & \boldsymbol{C}_1\\ \boldsymbol{O} & \boldsymbol{C}_2\end{pmatrix},
$$

则 (4.7.1) 式为 $\boldsymbol{AP}=\boldsymbol{PT}$. 由 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\beta}_{k+1},\cdots,\boldsymbol{\beta}_n$ 线性无关, 知 $\boldsymbol{P}$ 可逆, 且 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{T}$, 于是 $\boldsymbol{A}$ 与 $\boldsymbol{T}$ 具有相同的特征多项式

$$
|\lambda\boldsymbol{I}-\boldsymbol{A}|=|\lambda\boldsymbol{I}-\boldsymbol{T}|=\begin{vmatrix}(\lambda-\lambda_0)\boldsymbol{I} & -\boldsymbol{C}_1\\ \boldsymbol{O} & \lambda\boldsymbol{I}-\boldsymbol{C}_2\end{vmatrix}=|(\lambda-\lambda_0)\boldsymbol{I}|\cdot|\lambda\boldsymbol{I}-\boldsymbol{C}_2|=(\lambda-\lambda_0)^k g(\lambda),
$$

这里 $g(\lambda)=|\lambda\boldsymbol{I}-\boldsymbol{C}_2|$.

由上式可知, λ_0 的代数重数 $s\geqslant k=\lambda_0$ 的几何重数.

代数基本定理告诉我们, 任意复系数的 n 次多项式在复数域上恰好有 n 个根 (重根按重数计算). 因此, 若 n 阶矩阵的各特征值的代数重数都等于其几何重数, 则该矩阵一定存在 n 个线性无关的特征向量, 从而可相似对角化, 反之亦然.

问题 4.8 矩阵相似中的几个问题

(1) 矩阵的相似与等价具有怎样的关系

由矩阵相似的定义容易知道, 相似是一种特殊的等价关系, 但两矩阵等价却未必相似. 比如

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

是等价的, 但它们却不相似. 否则有

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

上式显然是矛盾的.

但另一方面, 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似却与其对应的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$, $\lambda\boldsymbol{I}-\boldsymbol{B}$ 等价是充分必要的关系. 即有

定理 4.8.1 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似的充分必要条件是它们的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}-\boldsymbol{B}$ 等价.

其理由为: 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 因此

$$\lambda\boldsymbol{I}-\boldsymbol{B}=\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}.$$

所以 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}-\boldsymbol{B}$ 等价.

反之若 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 与 $\lambda\boldsymbol{I}-\boldsymbol{B}$ 等价, 可证明一定存在常数矩阵 $\boldsymbol{R}$, $\boldsymbol{S}$ 使得

$$\lambda\boldsymbol{I}-\boldsymbol{B}=\boldsymbol{R}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{S},$$

于是 $\boldsymbol{R}\boldsymbol{S}=\boldsymbol{I}$, $\boldsymbol{B}=\boldsymbol{R}\boldsymbol{A}\boldsymbol{S}$, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似.

定理 4.8.1 不只是给出了矩阵相似的充要条件, 更重要的是把相似关系转化成了等价关系. 相似关系难以处理, 但等价关系更为简单, 它可以通过初等变换来互化.

(2) 相似矩阵的特征向量有何关系

我们知道相似矩阵有相同的特征值, 却未必有相同的特征向量. 但这并不意味着相似矩阵的特征向量之间没有任何关系, 相反, 相似矩阵的特征向量之间有十分紧密的联系, 具体表现为

定理 4.8.2 相似矩阵的特征向量只相差一可逆矩阵因子.

具体来讲就是: 若 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 且 $\boldsymbol{\alpha}$ 是 $\boldsymbol{B}$ 的特征向量, 则 $\boldsymbol{P}\boldsymbol{\alpha}$ 就是 $\boldsymbol{A}$ 的特征向量.

证 设有 $\boldsymbol{B}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}$, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则

$$\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{B},\ \boldsymbol{A}\boldsymbol{P}\boldsymbol{\alpha}=\boldsymbol{P}\boldsymbol{B}\boldsymbol{\alpha}=\lambda\boldsymbol{P}\boldsymbol{\alpha},$$

即 $\boldsymbol{P}\boldsymbol{\alpha}$ 是 $\boldsymbol{A}$ 的特征向量.

由此我们可以看出: **相似矩阵的特征子空间具有相同的结构, 它们能在某一可逆线性变换下互换.**

(3) 为什么要讨论矩阵的相似对角形

由于彼此相似的矩阵有许多共同的性质, 它们可以用来表示某些相同的研究对象. 譬如, 实对称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ ($\boldsymbol{Q}$ 是正交矩阵) 对应于可以用正交变换相互转化的二次型, 即有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \xlongequal{\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}} \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q})\boldsymbol{y}$. 再如, 从线性变换的角度看, $\boldsymbol{A}$ 与 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是同一线性变换在不同基下的矩阵. 即有

定理 4.8.3 设线性空间 V 中线性变换 $\mathcal{A}$ 在两组基

$$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n; \tag{4.8.1}$$

$$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n \tag{4.8.2}$$

下的矩阵分别为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 从基 (4.8.1) 到 (4.8.2) 的过渡矩阵是 $\boldsymbol{P}$, 则 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

证 由于 $\mathcal{A}\boldsymbol{\varepsilon}_i \in V$ $(i = 1, \cdots, n)$, 故它们能由两组基分别线性表出, 设

$$\begin{cases} \mathcal{A}\boldsymbol{\varepsilon}_1 = a_{11}\boldsymbol{\varepsilon}_1 + a_{21}\boldsymbol{\varepsilon}_2 + \cdots + a_{n1}\boldsymbol{\varepsilon}_n, \\ \mathcal{A}\boldsymbol{\varepsilon}_2 = a_{12}\boldsymbol{\varepsilon}_1 + a_{22}\boldsymbol{\varepsilon}_2 + \cdots + a_{n2}\boldsymbol{\varepsilon}_n, \\ \qquad\cdots\cdots\cdots\cdots \\ \mathcal{A}\boldsymbol{\varepsilon}_n = a_{1n}\boldsymbol{\varepsilon}_1 + a_{2n}\boldsymbol{\varepsilon}_2 + \cdots + a_{nn}\boldsymbol{\varepsilon}_n. \end{cases}$$

用矩阵表示为

$$(\mathcal{A}\boldsymbol{\varepsilon}_1, \mathcal{A}\boldsymbol{\varepsilon}_2, \cdots, \mathcal{A}\boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}.$$

记

$$\mathcal{A}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n) = (\mathcal{A}\boldsymbol{\varepsilon}_1, \mathcal{A}\boldsymbol{\varepsilon}_2, \cdots, \mathcal{A}\boldsymbol{\varepsilon}_n),$$

则有

$$\mathcal{A}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}.$$

同理

$$\mathcal{A}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)\boldsymbol{B}.$$

又

$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{B},$$

于是

$$\begin{aligned} \mathcal{A}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) &= \mathcal{A}[(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{P}] \\ &= [\mathcal{A}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)]\boldsymbol{P} \\ &= (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}\boldsymbol{P} \\ &= (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}, \end{aligned}$$

由于 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n$ 线性无关, 所以 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 定理得证.

设 $\boldsymbol{A}$ 是线性变换 $\mathcal{A}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 下的矩阵, 我们常提到的线性变换 $\mathcal{A}: \boldsymbol{x} \to \boldsymbol{A}\boldsymbol{x}$ (或 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$) 实际上是 $\mathcal{A}$ 的坐标表现形式. 具体描述如下: 设

$$\boldsymbol{\alpha} = x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + \cdots + x_n\boldsymbol{\varepsilon}_n = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{x}, \quad \mathcal{A}\boldsymbol{\alpha} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{y},$$

其中$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}$, 则有

$$\mathcal{A}\boldsymbol{\alpha} = \mathcal{A}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{x} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}\boldsymbol{x}.$$

由于 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 线性无关, 所以 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$.

容易证明定理 4.8.3 的逆命题也是正确的, 即两个相似矩阵是同一个线性变换在两组基下对应的矩阵.

为便于研究, 人们当然希望在彼此相似的一类矩阵中找一个最简单的 (如对角形矩阵) 作代表, 用它来表示所研究的对象.

此外, 矩阵相似对角形在简化计算方面也有广泛的应用, 比如矩阵的方幂运算、行列式计算、解齐次线性微分方程组, 等等.

例 4.8.1 计算 n 阶行列式

$$D_n = \begin{vmatrix} 2\cos\theta & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2\cos\theta & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2\cos\theta & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2\cos\theta \end{vmatrix} (\sin\theta \neq 0).$$

解 该行列式有多种计算方法. 这里将行列式的递推公式用矩阵形式来表示, 从而行列式的值就转化为矩阵方幂的计算.

将 D_n 按第一行展开, 有

$$D_n = 2\cos\theta \cdot D_{n-1} - D_{n-2},$$

由此可得到

$$\begin{pmatrix} D_n \\ D_{n-1} \end{pmatrix} = \begin{pmatrix} 2\cos\theta & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} D_{n-1} \\ D_{n-2} \end{pmatrix} = \cdots = \begin{pmatrix} 2\cos\theta & -1 \\ 1 & 0 \end{pmatrix}^{n-2} \begin{pmatrix} D_2 \\ D_1 \end{pmatrix}. \tag{4.8.3}$$

记 $\boldsymbol{A} = \begin{pmatrix} 2\cos\theta & -1 \\ 1 & 0 \end{pmatrix}$, 其特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \lambda^2 - 2\lambda\cos\theta + 1,$$

得特征根 $\lambda_1 = \mathrm{e}^{\mathrm{i}\theta}$, $\lambda_2 = \mathrm{e}^{-\mathrm{i}\theta}$ (这里用了 Euler (欧拉, 1707—1783) 公式: $\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$).

λ_1, λ_2 所对应的特征向量为 $\boldsymbol{\alpha}_1 = (\mathrm{e}^{\mathrm{i}\theta}, 1)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (\mathrm{e}^{-\mathrm{i}\theta}, 1)^{\mathrm{T}}$.

取 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$, 有 $\boldsymbol{P}^{-1} = \dfrac{1}{2\mathrm{i}\sin\theta}\begin{pmatrix} 1 & -\mathrm{e}^{-\mathrm{i}\theta} \\ -1 & \mathrm{e}^{\mathrm{i}\theta} \end{pmatrix}$, 且

$$\boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} \mathrm{e}^{\mathrm{i}\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\theta} \end{pmatrix}\boldsymbol{P}^{-1},$$

由 (4.8.3) 式可得

$$\begin{aligned}\begin{pmatrix} D_n \\ D_{n-1} \end{pmatrix} &= \boldsymbol{P}\begin{pmatrix} \mathrm{e}^{\mathrm{i}(n-2)\theta} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}(n-2)\theta} \end{pmatrix}\boldsymbol{P}^{-1}\begin{pmatrix} 4\cos^2\theta - 1 \\ 2\cos\theta \end{pmatrix} \\ &= \frac{1}{\sin\theta}\begin{pmatrix} \sin(n-1)\theta & -\sin(n-2)\theta \\ \sin(n-2)\theta & \sin(n-3)\theta \end{pmatrix}\begin{pmatrix} 2\cos 2\theta + 1 \\ 2\cos\theta \end{pmatrix}.\end{aligned} \tag{4.8.4}$$

由 (4.8.4) 式可求得 $D_n = \dfrac{\sin(n+1)\theta}{\sin\theta}$.

一般来说, 数列 $\{x_n\}$ 的 $k+1$ 阶常系数线性差分方程

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_k x_{n-k}$$

都可以表示成矩阵形式

$$\begin{aligned}\begin{pmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-k+1} \end{pmatrix} &= \begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}\begin{pmatrix} x_{n-1} \\ x_{n-2} \\ \vdots \\ x_{n-k} \end{pmatrix} = \cdots \\ &= \begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}^{n-k}\begin{pmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_1 \end{pmatrix},\end{aligned}$$

求通项 x_n 表达式的计算, 就转化成了矩阵方幂的计算. 矩阵的相似对角化就是很常用的方法.

例 4.8.2 求解线性微分方程组

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t} = x_1 - 3x_2 + 3x_3, \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t} = 3x_1 - 5x_2 + 3x_3, \\ \dfrac{\mathrm{d}x_3}{\mathrm{d}t} = 6x_1 - 6x_2 + 4x_3. \end{cases}$$

解 记 $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$, $\boldsymbol{A} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$, 方程组写为

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{x}. \tag{4.8.5}$$

求得 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = \lambda_2 = -2, \lambda_3 = 4$.
$\lambda_1 = \lambda_2 = -2$ 对应的特征向量为 $\boldsymbol{\alpha}_1 = (1,1,0)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2 = (-1,0,1)^{\mathrm{T}}$;
$\lambda_3 = 4$ 对应的特征向量为 $\boldsymbol{\alpha}_3 = (1,1,2)^{\mathrm{T}}$.

取 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$, 则有 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} -2 & & \\ & -2 & \\ & & 4 \end{pmatrix} \triangleq \boldsymbol{\Lambda}$.

作线性变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$, 其中 $\boldsymbol{y} = (y_1, y_2, y_3)^{\mathrm{T}}$, 则方程组 (4.8.5) 化为

$$\boldsymbol{P}\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{y}, \quad 即 \ \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{\Lambda}\boldsymbol{y}. \tag{4.8.6}$$

由 (4.8.6) 易求得 $\boldsymbol{y} = (c_1\mathrm{e}^{-2t}, c_2\mathrm{e}^{-2t}, c_3\mathrm{e}^{4t})^{\mathrm{T}}$, 所以原方程组的解为

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y} = \begin{pmatrix} c_1\mathrm{e}^{-2t} - c_2\mathrm{e}^{-2t} + c_3\mathrm{e}^{4t} \\ c_1\mathrm{e}^{-2t} + c_3\mathrm{e}^{4t} \\ c_2\mathrm{e}^{-2t} + 2c_3\mathrm{e}^{4t} \end{pmatrix}.$$

(4) 能否用初等变换化矩阵为相似对角形

由于可逆矩阵可以表示成若干初等矩阵的乘积, 所以两个相似矩阵总可以通过初等变换互化. 如果矩阵 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda}$ 相似, 当然也能通过初等变换将 $\boldsymbol{A}$ 化为对角矩阵 $\boldsymbol{\Lambda}$. 下面来讨论具体的方法.

设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$, 将 $\boldsymbol{P}$ 表示成若干初等矩阵的乘积, 设 $\boldsymbol{P} = \boldsymbol{E}_1 \cdots \boldsymbol{E}_m$, 则 $\boldsymbol{P}^{-1} = \boldsymbol{E}_m^{-1} \cdots \boldsymbol{E}_1^{-1}$, 从而

$$\boldsymbol{E}_m^{-1} \cdots \boldsymbol{E}_1^{-1}\boldsymbol{A}\boldsymbol{E}_1 \cdots \boldsymbol{E}_m = \boldsymbol{\Lambda}. \tag{4.8.7}$$

注意到

$$\boldsymbol{P}=\boldsymbol{I}\boldsymbol{E}_1\cdots\boldsymbol{E}_m, \tag{4.8.8}$$

(4.8.7)、(4.8.8) 式表明, 对 $\boldsymbol{A}$ 作一系列的初等列变换与对应的行逆变换 (称为列行互逆变换) 将 $\boldsymbol{A}$ 化为对角矩阵 $\boldsymbol{\Lambda}$, 则同样的列变换就将单位矩阵 $\boldsymbol{I}$ 化成了矩阵 $\boldsymbol{P}$. 即有

$$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{I}\end{pmatrix}\xrightarrow{\text{列行互逆变换}}\begin{pmatrix}\boldsymbol{\Lambda}\\ \boldsymbol{P}\end{pmatrix}$$

将其转置有

$$\left(\boldsymbol{A}^{\mathrm{T}},\boldsymbol{I}\right)\xrightarrow{\text{行列互逆变换}}\left(\boldsymbol{\Lambda},\boldsymbol{P}^{\mathrm{T}}\right). \tag{4.8.9}$$

该方法也同时求得了 $\boldsymbol{A}$ 的全部特征值 ($\boldsymbol{\Lambda}$ 的对角元) 与对应的线性无关的特征向量 ($\boldsymbol{P}^{\mathrm{T}}$ 的行向量).

上面谈到的矩阵的 "行列互逆变换", 其具体形式有以下三种:

(i) 互换 i,j 两行 ($r_i\leftrightarrow r_j$), 同时互换 i,j 两列 ($c_i\leftrightarrow c_j$);

(ii) 第 i 行乘非零数 k($k\cdot r_i$), 同时第 i 列乘数 $\dfrac{1}{k}\left(\dfrac{1}{k}\cdot c_i\right)$;

(iii) 第 i 行 k 倍加到第 j 行 ($r_j+k\cdot r_i$), 同时第 j 列 $-k$ 倍加到第 i 列 ($c_i-k\cdot c_j$).

例 4.8.3 用行列互逆变换将 $\boldsymbol{A}=\begin{pmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{pmatrix}$ 相似对角化.

解

$$\left(\boldsymbol{A}^{\mathrm{T}},\boldsymbol{I}\right)=\begin{pmatrix}1&-2&2&1&0&0\\-2&-2&4&0&1&0\\2&4&-2&0&0&1\end{pmatrix}\xrightarrow[(1)]{\substack{r_2-2r_1\\c_1+2c_2}}\begin{pmatrix}-3&-2&2&1&0&0\\0&2&0&-2&1&0\\10&4&-2&0&0&1\end{pmatrix}$$

$$\xrightarrow[(2)]{\substack{r_3+2r_1\\c_1-2c_3}}\begin{pmatrix}-7&-2&2&1&0&0\\0&2&0&-2&1&0\\0&0&2&2&0&1\end{pmatrix}\xrightarrow[(3)]{\substack{r_1-\frac{2}{9}r_3\\c_3+\frac{2}{9}c_1}}\begin{pmatrix}-7&-2&0&\frac{5}{9}&0&-\frac{2}{9}\\0&2&0&-2&1&0\\0&0&2&2&0&1\end{pmatrix}$$

$$\xrightarrow[(4)]{\substack{r_1+\frac{2}{9}r_2\\c_2-\frac{2}{9}c_1}}\begin{pmatrix}-7&0&0&\frac{1}{9}&\frac{2}{9}&-\frac{2}{9}\\0&2&0&-2&1&0\\0&0&2&2&0&1\end{pmatrix}\xrightarrow[(5)]{9r_1,\frac{1}{9}c_1}\begin{pmatrix}-7&0&0&1&2&-2\\0&2&0&-2&1&0\\0&0&2&2&0&1\end{pmatrix}.$$

按 (4.8.9), 有 $\boldsymbol{P}=\begin{pmatrix}1 & -2 & 2\\ 2 & 1 & 0\\ -2 & 0 & 1\end{pmatrix}$, $\boldsymbol{\Lambda}=\begin{pmatrix}-7 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}$, 且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$.

注 步骤 (1) 处的 $k=-2$ 是由方程 $-2+2k+(-2-2k)(-k)=0$ 确定的, 步骤 (3) 处的 $k=-\dfrac{2}{9}$ 是由方程 $2k+2+(-7)(-k)=0$ 确定的. 其他各处的 k 值类似可定.

例 4.8.4 求矩阵 $\boldsymbol{A}=\begin{pmatrix}2 & -1 & 1\\ 0 & 3 & -1\\ 2 & 1 & 3\end{pmatrix}$ 的特征值与特征向量, 并判定 $\boldsymbol{A}$ 是否能相似对角化.

解

$$(\boldsymbol{A}^{\mathrm{T}},\boldsymbol{I})=\begin{pmatrix}2 & 0 & 2 & 1 & 0 & 0\\ -1 & 3 & 1 & 0 & 1 & 0\\ 1 & -1 & 3 & 0 & 0 & 1\end{pmatrix}\xrightarrow[c_3+c_1]{r_1-r_3}\begin{pmatrix}1 & 1 & 0 & 1 & 0 & -1\\ -1 & 3 & 0 & 0 & 1 & 0\\ 1 & -1 & 4 & 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow[c_1+c_2]{r_2-r_1}\begin{pmatrix}2 & 1 & 0 & 1 & 0 & -1\\ 0 & 2 & 0 & -1 & 1 & 1\\ 0 & -1 & 4 & 0 & 0 & 1\end{pmatrix}\xrightarrow[c_2+\frac{1}{2}c_3]{r_3-\frac{1}{2}r_2}\begin{pmatrix}2 & 1 & 0 & 1 & 0 & -1\\ 0 & 2 & 0 & -1 & 1 & 1\\ 0 & 0 & 4 & \dfrac{1}{2} & -\dfrac{1}{2} & \dfrac{1}{2}\end{pmatrix}$$

$$\xrightarrow{2r_3,\frac{1}{2}c_3}\begin{pmatrix}2 & 1 & 0 & 1 & 0 & -1\\ 0 & 2 & 0 & -1 & 1 & 1\\ 0 & 0 & 4 & 1 & -1 & 1\end{pmatrix}.$$

最后一个矩阵中, 由于方程 $1+2k+(-2k)=0$ 无解, 所以无法将元 $a_{12}=1$ 化为 0, 故矩阵 $\boldsymbol{A}$ 不能与对角矩阵相似. 事实上, 矩阵 $\boldsymbol{J}=\begin{pmatrix}2 & 0 & 0\\ 1 & 2 & 0\\ 0 & 0 & 4\end{pmatrix}$ 是 $\boldsymbol{A}$ 的 Jordan 标准形.

$\boldsymbol{A}$ 的特征值为 $\lambda_1=\lambda_2=2,\lambda_3=4$.

$\lambda_1=\lambda_2=2$ 对应的全部特征向量为 $k_1(-1,1,1)^{\mathrm{T}}$ $(k_1\neq 0)$;

$\lambda_3=4$ 对应的全部特征向量为 $k_2(1,-1,1)^{\mathrm{T}}$ $(k_2\neq 0)$.

问题 4.9 乘积可交换矩阵一定有公共的特征向量吗

乘积可交换的矩阵存在许多内在的联系, 就特征向量而言, 有以下结论:

定理 4.9.1 若 n 阶方阵 $\boldsymbol{A}, \boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{BA}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 一定有公共的特征向量.

证 任取 $\boldsymbol{A}$ 的一个特征值 λ, 考虑 λ 的特征子空间

$$V_\lambda=\{\boldsymbol{\xi}|\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}\}.$$

设 $\dim V_\lambda=k$, $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_k$ 是 V_λ 的一组基, 则 $\boldsymbol{A\varepsilon}_i=\lambda\boldsymbol{\varepsilon}_i$ $(i=1,\cdots,k)$.

记

$$\boldsymbol{\eta}=c_1\boldsymbol{\varepsilon}_1+c_2\boldsymbol{\varepsilon}_2+\cdots+c_k\boldsymbol{\varepsilon}_k, \tag{4.9.1}$$

当 $c_1,c_2,\cdots,c_k$ 不全为零时, $\boldsymbol{\eta}$ 也是 $\boldsymbol{A}$ 的特征向量.

由于

$$\boldsymbol{A}(\boldsymbol{B\varepsilon}_i)=\boldsymbol{B}(\boldsymbol{A\varepsilon}_i)=\lambda(\boldsymbol{B\varepsilon}_i),$$

故 $\boldsymbol{B\varepsilon}_i\in V_\lambda$ $(i=1,\cdots,k)$, 从而存在数 l_{ij} $(i,j=1,\cdots,k)$ 使得

$$\boldsymbol{B\varepsilon}_i=l_{1i}\boldsymbol{\varepsilon}_1+l_{2i}\boldsymbol{\varepsilon}_2+\cdots+l_{ki}\boldsymbol{\varepsilon}_k\ (i=1,\cdots,k).$$

则

$$\begin{aligned}\boldsymbol{B\eta}&=c_1\boldsymbol{B\varepsilon}_1+c_2\boldsymbol{B\varepsilon}_2+\cdots+c_k\boldsymbol{B\varepsilon}_k\\&=c_1(l_{11}\boldsymbol{\varepsilon}_1+l_{21}\boldsymbol{\varepsilon}_2+\cdots+l_{k1}\boldsymbol{\varepsilon}_k)+\cdots+c_k(l_{1k}\boldsymbol{\varepsilon}_1+l_{2k}\boldsymbol{\varepsilon}_2+\cdots+l_{kk}\boldsymbol{\varepsilon}_k)\\&=(c_1l_{11}+c_2l_{12}+\cdots+c_kl_{1k})\boldsymbol{\varepsilon}_1+\cdots+(c_1l_{k1}+c_2l_{k2}+\cdots+c_kl_{kk})\boldsymbol{\varepsilon}_k.\end{aligned}$$

若 $\boldsymbol{\eta}$ 也是 $\boldsymbol{B}$ 的特征向量, 即有 $\boldsymbol{B\eta}=\mu\boldsymbol{\eta}$. 由 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_k$ 线性无关, 得

$$\begin{cases}c_1l_{11}+c_2l_{12}+\cdots+c_kl_{1k}=\mu c_1,\\c_1l_{21}+c_2l_{22}+\cdots+c_kl_{2k}=\mu c_2,\\\qquad\cdots\cdots\cdots\cdots\\c_1l_{k1}+c_2l_{k2}+\cdots+c_kl_{kk}=\mu c_k.\end{cases} \tag{4.9.2}$$

记 $\boldsymbol{L}=\begin{pmatrix}l_{11}&l_{12}&\cdots&l_{1k}\\l_{21}&l_{22}&\cdots&l_{2k}\\\vdots&\vdots&&\vdots\\l_{k1}&l_{k2}&\cdots&l_{kk}\end{pmatrix}$, (4.9.2) 写为

$$(\mu\boldsymbol{I}-\boldsymbol{L})\begin{pmatrix}c_1\\\vdots\\c_k\end{pmatrix}=\boldsymbol{0}. \tag{4.9.3}$$

这说明 μ 是矩阵 $\boldsymbol{L}$ 的特征值, $(c_1,\cdots,c_k)^{\mathrm{T}}$ 是对应于 μ 的特征向量. 由于矩阵的特征向量一定存在, 即方程组 (4.9.3) 一定有非零解 $(c_1,\cdots,c_k)^{\mathrm{T}}\neq\mathbf{0}$, 代入 (4.9.1) 所确定的向量 $\boldsymbol{\eta}$ 就是矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 公共的特征向量.

根据上面的证明, 我们可立即得到

推论 1　若 n 阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{BA}$, 且 $\boldsymbol{A}$ 有 $r\ (r\leqslant n)$ 个互不相同的特征值, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 至少有 r 个线性无关的公共特征向量.

推论 2　若 n 阶方阵 $\boldsymbol{A}$ 有 n 个互不相同的特征值, 则 $\boldsymbol{AB}=\boldsymbol{BA}$ 的充分必要条件是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有完全相同的特征向量.

证　只需证明推论 2 中的充分性.

$\boldsymbol{A}$ 有 n 个互不相同的特征值, 则 $\boldsymbol{A}$ 可相似对角化, 即存在可逆矩阵 $\boldsymbol{P}$, 使

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\triangleq\boldsymbol{\Lambda}_1.$$

由于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有完全相同的特征向量, 同样有

$$\boldsymbol{P}^{-1}\boldsymbol{BP}=\begin{pmatrix}\mu_1&&\\&\ddots&\\&&\mu_n\end{pmatrix}\triangleq\boldsymbol{\Lambda}_2.$$

则

$$\begin{aligned}\boldsymbol{AB}&=(\boldsymbol{P\Lambda}_1\boldsymbol{P}^{-1})(\boldsymbol{P\Lambda}_2\boldsymbol{P}^{-1})=\boldsymbol{P\Lambda}_1\boldsymbol{\Lambda}_2\boldsymbol{P}^{-1}=\boldsymbol{P\Lambda}_2\boldsymbol{\Lambda}_1\boldsymbol{P}^{-1}\\&=(\boldsymbol{P\Lambda}_2\boldsymbol{P}^{-1})(\boldsymbol{P\Lambda}_1\boldsymbol{P}^{-1})=\boldsymbol{BA}.\end{aligned}$$

注　推论 2 中的必要性也可以不依赖定理 4.9.1 中的证明. 现另证如下:

由于 $\boldsymbol{A}$ 有 n 个互不相同的特征值, 从而对应 n 个线性无关的特征向量, 以这 n 个线性无关的特征向量为列向量构成矩阵 $\boldsymbol{P}$, 则

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}.$$

由 $\boldsymbol{AB}=\boldsymbol{BA}\Rightarrow(\boldsymbol{P}^{-1}\boldsymbol{AP})(\boldsymbol{P}^{-1}\boldsymbol{BP})=(\boldsymbol{P}^{-1}\boldsymbol{BP})(\boldsymbol{P}^{-1}\boldsymbol{AP})$, 即

$$\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\begin{pmatrix}c_{11}&\cdots&c_{1n}\\\vdots&&\vdots\\c_{n1}&\cdots&c_{nn}\end{pmatrix}=\begin{pmatrix}c_{11}&\cdots&c_{1n}\\\vdots&&\vdots\\c_{n1}&\cdots&c_{nn}\end{pmatrix}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix},$$

由此可得到 $c_{ij}=0\ (i\neq j)$, 所以

$$
\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}=\begin{pmatrix} c_{11} & & \\ & \ddots & \\ & & c_{nn} \end{pmatrix}.
$$

这表明 $\boldsymbol{P}$ 的 n 个列向量也是 $\boldsymbol{B}$ 的特征向量.

例 4.9.1 求可交换矩阵 $\boldsymbol{A}=\begin{pmatrix} 3 & 0 & 0 \\ -1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix}$ 与 $\boldsymbol{B}=\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix}$ 的所有公共特征向量. 问是否存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 与 $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ 同为对角矩阵?

解 容易验证 $\boldsymbol{AB}=\boldsymbol{BA}$. 下求 $\boldsymbol{A}$ 的特征值与特征向量:

$$
|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix} \lambda-3 & 0 & 0 \\ 1 & \lambda-3 & -1 \\ -2 & 0 & \lambda-1 \end{vmatrix}=(\lambda-1)(\lambda-3)^2,
$$

特征值 $\lambda_1=1$, $\lambda_2=\lambda_3=3$.

对 $\lambda_1=1$, 方程组 $\begin{pmatrix} -2 & 0 & 0 \\ 1 & -2 & -1 \\ -2 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ 的基础解系为

$$
\boldsymbol{\varepsilon}=\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.
$$

而

$$
\boldsymbol{B\varepsilon}=\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}=-\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix},
$$

$\boldsymbol{\varepsilon}$ 也是 $\boldsymbol{B}$ 的特征向量, 对应特征值 $\mu=-1$.

此时 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的公共特征向量为 $c\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}\ (c\neq 0)$.

对 $\lambda_2=\lambda_3=3$, 方程组 $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ 的基础解系为

$\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\boldsymbol{\varepsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. 而

$$\boldsymbol{B}\boldsymbol{\varepsilon}_1 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = 2\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2,$$
$$\boldsymbol{B}\boldsymbol{\varepsilon}_2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = 2\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2.$$

定理 4.9.1 中的 $\boldsymbol{L} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}$.

由 $|\mu \boldsymbol{I} - \boldsymbol{L}| = \begin{vmatrix} \mu - 2 & -2 \\ -1 & \mu - 1 \end{vmatrix} = \mu(\mu - 3) = 0$, 得 $\mu_1 = 0, \mu_2 = 3$.

对 $\mu_1 = 0$, $\begin{pmatrix} -2 & -2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, 得 $c_2 = -c_1$, 于是公共特征向量为

$$\boldsymbol{\eta} = c_1(\boldsymbol{\varepsilon}_1 - \boldsymbol{\varepsilon}_2) = c_1 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \ c_1 \neq 0.$$

对 $\mu_2 = 3$, $\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, 得 $c_1 = 2c_2$, 于是公共特征向量为

$$\boldsymbol{\eta} = c_2(2\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2) = c_2 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \ c_2 \neq 0.$$

综上, 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的所有公共特征向量为

$$\boldsymbol{\eta}_1 = c_1 \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}, \boldsymbol{\eta}_2 = c_2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \boldsymbol{\eta}_3 = c_3 \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} (c_1, c_2, c_3 \text{ 是不为零的任意常数}).$$

由于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同有三个线性无关的特征向量, 所以它们能够同时相似对角

化. 取

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

则有

$$P^{-1}AP=\begin{pmatrix}1&&\\&3&\\&&3\end{pmatrix}, P^{-1}BP=\begin{pmatrix}-1&&\\&0&\\&&3\end{pmatrix}.$$

例 4.9.2　设 n 阶方阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}$, 证明:

(1) $\boldsymbol{A}$, $\boldsymbol{B}$ 乘法可交换;

(2) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有完全相同的特征向量.

证　(1) 因为 $\boldsymbol{AB}=\boldsymbol{A}+\boldsymbol{B}\Rightarrow(\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{I}$, 所以 $(\boldsymbol{B}-\boldsymbol{I})(\boldsymbol{A}-\boldsymbol{I})=\boldsymbol{I}$, 从而

$$(\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{B}-\boldsymbol{I})=(\boldsymbol{B}-\boldsymbol{I})(\boldsymbol{A}-\boldsymbol{I})\Rightarrow\boldsymbol{AB}-\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{I}=\boldsymbol{BA}-\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{I}$$

故

$$\boldsymbol{AB}=\boldsymbol{BA}.$$

(2) 由 $(\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{B}-\boldsymbol{I})=\boldsymbol{I}$ 知 $|\boldsymbol{I}-\boldsymbol{A}|\neq 0$, $|\boldsymbol{I}-\boldsymbol{B}|\neq 0$, 所以 1 不是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征值.

设 $\boldsymbol{\beta}$ 是 $\boldsymbol{B}$ 的任一特征向量, 且有 $\boldsymbol{B\beta}=\mu\boldsymbol{\beta}$, 则

$$\boldsymbol{AB\beta}=\boldsymbol{A\beta}+\boldsymbol{B\beta}\Rightarrow\mu\boldsymbol{A\beta}=\boldsymbol{A\beta}+\mu\boldsymbol{\beta}.$$

因为 $\mu\neq 1$, 上式整理得

$$\boldsymbol{A\beta}=\frac{\mu}{\mu-1}\boldsymbol{\beta},$$

这说明 $\boldsymbol{\beta}$ 也是 $\boldsymbol{A}$ 的特征向量.

由于 $\boldsymbol{A}$, $\boldsymbol{B}$ 乘法可交换, 又有 $\boldsymbol{BA}=\boldsymbol{A}+\boldsymbol{B}$, 故 $\boldsymbol{A}$ 的任一特征向量也是 $\boldsymbol{B}$ 的特征向量. 所以 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有完全相同的特征向量.

问题 4.10　为什么实对称矩阵一定能正交对角化

实对称矩阵能正交对角化的最基本原因是其特征值均为实数. 一般工科线性代数教材均给出了这一性质, 以及实对称矩阵正交对角化的方法, 但多数却没给出后者的一般性证明, 这里介绍两种较常见的证明, 它们都是容易被学生所理解的.

方法 1 对实对称矩阵 $\boldsymbol{A}$ 的阶数 n 作数学归纳法, 证明 $\boldsymbol{A}$ 正交相似于对角矩阵.

$n=1$ 时, 一阶矩阵已经是对角矩阵, 结论显然成立.

设 $n-1$ 阶实对称矩阵能正交相似于对角矩阵.

对 n 阶实对称矩阵 $\boldsymbol{A}$, 取 $\boldsymbol{A}$ 的一个特征值 λ_1, 由于 $\boldsymbol{A}$ 的特征值为实数, 所以对应的特征向量可以取为实向量, 设 $\boldsymbol{\alpha}_1$ 是 λ_1 所对应的一个单位实向量, 将 $\boldsymbol{\alpha}_1$ 扩充为空间 $\mathbf{R}^n$ 中的一组标准正交基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$. 以它们为列向量作成正交矩阵 $\boldsymbol{Q}_1=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, 则

$$\boldsymbol{A}\boldsymbol{Q}_1=\boldsymbol{Q}_1\begin{pmatrix}\lambda_0 & \boldsymbol{B}_1\\ \boldsymbol{O} & \boldsymbol{B}_2\end{pmatrix}\Rightarrow \boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\lambda_0 & \boldsymbol{B}_1\\ \boldsymbol{O} & \boldsymbol{B}_2\end{pmatrix}.$$

由于 $\boldsymbol{A}$ 是实对称矩阵, 则 $\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1$ 也是实对称矩阵 $(\boldsymbol{Q}_1^{-1}=\boldsymbol{Q}_1^{\mathrm{T}})$, 于是 $\boldsymbol{B}_1=\boldsymbol{O}$, 且 $\boldsymbol{B}_2$ 是 $n-1$ 阶实对称矩阵. 由归纳假设存在 $n-1$ 阶正交矩阵 $\boldsymbol{Q}_2$ 使 $\boldsymbol{B}_2$ 相似于对角矩阵

$$\boldsymbol{Q}_2^{-1}\boldsymbol{B}_2\boldsymbol{Q}_2=\begin{pmatrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

于是正交矩阵 $\boldsymbol{Q}=\boldsymbol{Q}_1\begin{pmatrix}1 & \\ & \boldsymbol{Q}_2\end{pmatrix}$ 就使 $\boldsymbol{A}$ 相似于对角矩阵

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

方法 2 证明实对称矩阵每个特征值的代数重数都等于其几何重数, 因而可正交对角化.

设 λ_1 是 n 阶实对称矩阵 $\boldsymbol{A}$ 的任一 s 重特征值, 其几何重数为 k. 由定理 4.7.1 我们知道 $k\leqslant s$, 下面证明 $k=s$.

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_k$ 是 λ_1 所对应的 k 个线性无关的实特征向量, 不妨设它们已经正交化、单位化. 将它们扩充为 $\mathbf{R}^n$ 的一组标准正交基 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k,\boldsymbol{\alpha}_{k+1},\cdots,\boldsymbol{\alpha}_n$, 以它们为列向量作成正交矩阵

$$\boldsymbol{Q}_1=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_l,\boldsymbol{\alpha}_{l+1},\cdots,\boldsymbol{\alpha}_n).$$

由 $\boldsymbol{A}$ 是实对称矩阵知 $\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1$ 也是实对称矩阵, 且有

$$\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\underbrace{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}_{k\text{ 个}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix},$$

其中 $\boldsymbol{B}$ 也是实对称矩阵.

若 $k<s$, 由于 $\boldsymbol{A}$ 与 $\boldsymbol{Q}_1^{-1}\boldsymbol{A}\boldsymbol{Q}_1$ 有相同的特征值, 所以 λ_1 还是 $\boldsymbol{B}$ 的特征值. 同上面一样, 我们有 $n-k$ 阶正交矩阵 $\boldsymbol{Q}_2$ 使

$$\boldsymbol{Q}_2^{-1}\boldsymbol{B}\boldsymbol{Q}_2=\begin{pmatrix}\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{C}\end{pmatrix}.$$

令 $\boldsymbol{Q}=\boldsymbol{Q}_1\begin{pmatrix}\underbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}_{k\text{ 个}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q}_2\end{pmatrix}$, 易知 $\boldsymbol{Q}$ 仍是正交矩阵, 且有

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}\underbrace{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}_{k\text{ 个}} & & \boldsymbol{O} \\ & \begin{matrix}\ddots & \\ & \lambda_1\end{matrix} & \\ \boldsymbol{O} & & \boldsymbol{C}\end{pmatrix}.$$

这样就有

$$\lambda_1 \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \begin{pmatrix} \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{k\ \text{个}} & & & \boldsymbol{O} \\ & \ddots & & \\ \boldsymbol{O} & & 0 & \\ & & & \lambda_1 \boldsymbol{I} - \boldsymbol{C} \end{pmatrix}.$$

于是 $\mathrm{rank}\,(\lambda_1 \boldsymbol{I} - \boldsymbol{A}) = \mathrm{rank}\,(\lambda_1 \boldsymbol{I} - \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}) < n - k$, 从而齐次方程组 $(\lambda_1 \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的基础解系所含解向量的个数 (也就是 λ_1 的几何重数) $n - \mathrm{rank}\,(\lambda_1 \boldsymbol{I} - \boldsymbol{A}) > k$. 这与所设矛盾. 所以只能有 $k = s$. 故实对称矩阵一定与对角矩阵相似, 并且可以选对应的变换矩阵为正交矩阵.

从几何上看, n 阶实对称矩阵所对应的线性变换在 $\mathbf{R}^n$ 空间中有 n 个相互正交的不变方向 (特征方向), 选这 n 个方向的单位向量为基底, 则该线性变换在这组基下的矩阵形式最简单, 它就是对角矩阵 (见定理 4.8.3 与问题 4.1 (1)).

第五章　二次型

问题 5.1　如何认识二次型所研究的问题

线性代数主要研究线性运算与线性变换中的问题，二次多项式按理不在该讨论范围内，但二次齐次式 (二次型) 是双线性式的特例，自然也就属于线性运算的范畴. 不仅如此，如果我们将二次型用矩阵来表示，那么二次型在满秩线性变换下所呈现的规律就完全可以通过其矩阵的运算 (合同) 揭示出来. 所以，二次型的问题可看成是矩阵的一个应用.

那么二次型需要讨论的一般性问题是什么?

从平面解析几何中我们知道，对于二次曲线

$$ax^2 + bxy + cy^2 = d,$$

总能通过适当的坐标旋转变换 $\begin{cases} x = x' \cos\theta - y' \sin\theta, \\ y = x' \sin\theta + y' \cos\theta \end{cases}$ 化为标准形

$$a'x'^2 + c'y'^2 = d.$$

从标准形，我们容易识别曲线的类型，便于研究曲线的性质. 由此自然会想到空间二次曲面的相应问题. 作为更一般的情况，就需要人们去研究含任意 n 个变量的二次型. 讨论在满秩线性变换下二次型的最简单形式 (标准形)，将二次型化为标准形的方法，以及满秩线性变换保持了二次型的哪些性质，有何应用，等等. 这些就构成了二次型最基本的内容. 同时，对一些特殊二次型 (如正定、负定，半正定、半负定) 的研究，以及将二次型问题转化为矩阵问题，使之得到更广泛的发展与应用，则进一步丰富了二次型的研究内容.

从历史发展来看，二次型的系统研究是从 18 世纪开始的，它起源于对二次曲线和二次曲面的分类问题的讨论，将二次曲线和二次曲面的方程变形，选择主

轴方向的轴作为坐标轴以简化方程的形状. Euler 在他的《引论》(1748 年) 中就讨论了用旋转变换化二次曲面方程为标准形的方法; Cauchy 在他的著作中曾给出结论: 当方程式为标准形时, 二次曲面用二次项的符号来进行分类. Cauchy 还讨论了化一般二次型为标准形的方法. 然而, 那时并不太清楚, 在化简成标准形时, 为何总是得到同样数目的正项和负项. 19 世纪中叶, 英国数学家 Sylvester 回答了这个问题, 1852 年他给出了 n 个变数的二次型的惯性定律, 并用反证法给出了严格的证明. 1801 年, Gauss 在《算术研究》中引进了二次型的正定、负定、半正定和半负定等术语. 1868 年, Weierstrass (魏尔斯特拉斯, 1815—1897) 较系统地完成了二次型的理论, 并将其推广到双线性型.

二次型的理论在最优化、物理学、几何学、概率论等学科中都有广泛的应用.

最后我们对双线性式的概念作一个简单的介绍:

设 $f(x_1,\cdots,x_n;y_1,\cdots,y_n)$ 是两组变量 $x_1,\cdots,x_n$ 及 $y_1,\cdots,y_n$ 的多项式, 如果对每一组变量, f 都是线性齐次式, 那么 f 叫**双线性式**. 一般双线性式 f 可写成

$$f(\boldsymbol{x},\boldsymbol{y})=\sum_{i,j=1}^{n}a_{ij}x_iy_j=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{y},$$

其中 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$, $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$, $\boldsymbol{A}=(a_{ij})$ 是 n 阶矩阵.

若 $f(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{y}$ 的矩阵 $\boldsymbol{A}$ 是对称矩阵, 则称 f 叫**对称双线性式**. 显然这时 $f(\boldsymbol{x},\boldsymbol{y})=f(\boldsymbol{y},\boldsymbol{x})$. 若 $\boldsymbol{A}$ 是反称矩阵, 则称 f 叫**反对称双线性式**, 这时 $f(\boldsymbol{x},\boldsymbol{y})=-f(\boldsymbol{y},\boldsymbol{x})$. 反过来也都成立. $\boldsymbol{x}=\boldsymbol{y}$ 的对称双线性式也就是**二次型**.

问题 5.2 为什么要确定二次型的矩阵为对称矩阵

二次型就是一个二次齐次多项式, 将二次型写成矩阵形式是便于利用矩阵这个强有力的工具对二次型作研究. 但将一个二次型写成矩阵形式的方法有很多, 比如

$$\begin{aligned}f(x_1,x_2)=x_1^2+4x_1x_2+4x_2^2&=(x_1,x_2)\begin{pmatrix}1&3\\1&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}\\&=(x_1,x_2)\begin{pmatrix}1&0\\4&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=(x_1,x_2)\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix},\end{aligned}\tag{5.2.1}$$

这种表达形式的多样性不仅不利于问题的研究, 而且还会出现一些混乱的现象. 比如 (5.2.1) 式中的前两个矩阵的秩为 2, 最后一个矩阵的秩为 1. 由于满秩线性变换不改变二次型的秩, 从而二次型化简后的形式 (标准形) 完全取决于对最初表达式的选择, 这样的研究结果显然没有任何意义.

当然使二次型的矩阵表达形式唯一的方法有很多 (比如可以规定为上三角形矩阵或下三角形矩阵等), 但为什么一定要规定为对称矩阵? 事实上, 除了对称矩阵, 没有其他任何一种形式可选. 其原因是, 在满秩线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 下, 二次型

$$f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\xlongequal{\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}}\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{y},$$

如果表示唯一, 必有 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$, 也就是说矩阵 $\boldsymbol{B}$ 可由 $\boldsymbol{A}$ 经过一系列初等行变换与相同的列变换 (称为合同变换) 而得到, 这种变换能保持矩阵 $\boldsymbol{A}$ 的对称性, 而其他形式 (如三角形矩阵) 是不能保持的. 也就是说, 选择二次型的矩阵为对称矩阵并不是数学家们的心血来潮, 而是客观规律的必然之选.

问题 5.3　对二次型的研究为什么要以满秩线性变换为手段

首先, 对一个对象用怎样的手段或方法去研究, 取决于我们的研究目标与所限定的范围. 比如, 如果将二次型视为一个多元函数, 如果我们要考察它的连续性、可微性或极值等问题, 我们所使用的手段或方法就可能不一样. 线性代数中研究二次型的主要目的是要弄清楚在满秩线性变换下二次型有什么变化, 有哪些不变量, 并利用这些变与不变的特性去解决我们需要解决的问题.

线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 的一个显著特点是它能保持二次型的 "二次齐次多项式" 形式不变, 即只有线性变换才能将二次型变为二次型. 但仅保持 "形式" 还不够, 我们还希望这种变换具有 "普适性", 即对任意的 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$, 都有唯一的 $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$ 与之对应. 要做到这一点, 就必须要求 $\boldsymbol{C}$ 是满秩的, 这是因为 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 相当于一个线性方程组

$$\begin{cases} x_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n, \\ x_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n, \\ \cdots\cdots\cdots\cdots \\ x_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n. \end{cases} \tag{5.3.1}$$

如果 $\boldsymbol{C}$ 是降秩的, 线性方程组 (5.3.1) 没有唯一解. 它可能无解, 也可能有无穷多解. 这样的变换就没有 "普适性", 也不便于讨论和应用. 事实上, 这种变换可能会 "丢失" 许多信息, 比如它会将点集 $\{\boldsymbol{y}|\boldsymbol{C}\boldsymbol{y}=\boldsymbol{0},\boldsymbol{y}\in\mathbf{R}^n\}$ 中的元全部变为零. 从另一个角度看, 要求 $\boldsymbol{C}$ 是满秩的, 也就是要求变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 是可逆的, 只有当所用的变换可逆, 我们才能从变换的结果中得到原问题的相关信息 (变换中的不变量).

例 5.3.1　设 $f(x_1,x_2,x_3)=x_1x_2+x_1x_3-3x_2x_3$, 作线性变换将 f 化为标准形, 并讨论方程 $f(x_1,x_2,x_3)=1$ 表示什么类型的曲面.

解法 1　作线性变换 $\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_3, \end{cases}$ 则

$$f = y_1^2 - y_2^2 - 2y_1y_3 + 4y_2y_3 = (y_1 - y_3)^2 - (y_2 - 2y_3)^2 + 3y_3^2.$$

再令 $\begin{cases} z_1 = y_1 - y_3, \\ z_2 = y_2 - 2y_3, \\ z_3 = y_3, \end{cases}$ 则 f 的标准形为

$$f = z_1^2 - z_2^2 + 3z_3^2.$$

$f(x_1, x_2, x_3) = 1$ 表示的曲面是单叶双曲面.

解法 2　作线性变换 $\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_1, \end{cases}$ 则

$$f = -y_1^2 - y_2^2 + 4y_1y_2 = -(y_1 - 2y_2)^2 + 3y_2^2.$$

再令 $\begin{cases} z_1 = y_1 - 2y_2, \\ z_2 = y_2, \\ z_3 = y_3, \end{cases}$ 则 f 的标准形为

$$f = -z_1^2 + 3z_2^2.$$

$f(x_1, x_2, x_3) = 1$ 表示的曲面是双曲柱面.

上面两种不同的做法得到完全不一样的结论, 问题出在什么地方? 显然解法 2 中的做法是错误的. 其原因是线性变换 $\begin{cases} x_1 = y_1 + y_2, \\ x_2 = y_1 - y_2, \\ x_3 = y_1 \end{cases}$ 不是满秩的, 由此得到的标准形不能通过可逆线性变换回到原二次型. 从几何上看, 该变换将空间 $\boldsymbol{O}$-$y_1y_2y_3$ 中的直线 $\begin{cases} y_1 = 0, \\ y_2 = 0, \quad t \in \mathbf{R}, \\ y_3 = t, \end{cases}$ 映射成了空间 $\boldsymbol{O}$-x_1, x_2, x_3 中的一个点, 即坐标原点.

在解析几何中, 点与坐标的一一对应是讨论全部问题的基础, 如果用了降秩的坐标变换, 那么可能有的点就没有新坐标对应, 而有的点却又有无穷多个新坐标对应, 这样的坐标变换显然是没有意义的. 所以在讨论变量之间的线性关系时需要用满秩变换.

问题 5.4 矩阵的等价、相似与合同有何区别与联系

矩阵的等价、相似与合同是同型矩阵的三种不同的关系, 它们既有区别又有一定的联系.

矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 是指存在可逆矩阵 $\boldsymbol{P},\boldsymbol{Q}$ 使得 $\boldsymbol{A}=\boldsymbol{PBQ}$, 等价矩阵一定可通过初等变换互化, 其本质特征是它们具有相同的秩 (充要条件).

矩阵的相似是方阵的一种特殊等价关系. 两个方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 如果存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{BP}$, 则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似. 两个矩阵是否相似很难通过初等变换来判别, 但我们可以通过对它们的特征矩阵 $\lambda\boldsymbol{I}-\boldsymbol{A}$, $\lambda\boldsymbol{I}-\boldsymbol{B}$ 作初等变换来判别, 因为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似的充分必要条件是它们的特征矩阵等价 (定理 4.8.1). 两个矩阵相似的共同特征是它们有相同的特征值, 但有相同的特征值仅仅是两矩阵相似的必要条件, 我们只能据此来否定两个矩阵的相似. 对于两个实对称矩阵来说, 有相同的特征值则是它们相似的充要条件.

矩阵的合同是也是方阵的一种特殊等价关系. 两个方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 如果存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{BP}$, 则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同. 两个合同的矩阵一定可以通过对称初等变换 (合同变换) 互化. 所谓对称初等变换是指对一个矩阵每作一次初等行变换, 同时作一次相应的初等列变换. 两个矩阵合同的充要条件是它们对应的二次型有相同的规范形. 矩阵合同与所考虑的数域有关, 在复数域中两个同型对称矩阵合同的充要条件是它们有相同的秩, 在实数域中两个同型对称矩阵合同的充要条件是它们有相同的正、负惯性指数 (即有相同的正特征值个数与负特征值个数).

从以上描述, 我们容易知道, 两矩阵的相似或合同关系一定是等价关系, 反之则未必; 相似与合同一般没有从属关系. 我们既能找到合同但不相似的矩阵, 又能找到相似但不合同的矩阵. 例如:

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

是合同的, 因为它们的正惯性指数都是 2, 负惯性指数都是 0. 也可由下式得到合同

$$\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}.$$

但 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 却不相似. 否则会有

$$\begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}=\boldsymbol{P}^{-1}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\boldsymbol{P}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

再如

$$C=\begin{pmatrix}2&1\\1&0\end{pmatrix},\quad D=\begin{pmatrix}1&1\\2&1\end{pmatrix}$$

是相似的, 因为

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

但 $\boldsymbol{C}$ 与 $\boldsymbol{D}$ 却不合同, 因为与 $\boldsymbol{C}$ 合同的矩阵一定是对称矩阵, 而 $\boldsymbol{D}$ 不是对称矩阵.

需要说明的是, 矩阵的相似与合同没有从属关系是就一般情况而言的, 假若变换矩阵 $\boldsymbol{P}$ 是正交矩阵, 即 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{\mathrm{T}}$, 则 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$, 此时矩阵的相似与合同就是一致的. 又假若两个实对称矩阵相似, 那么它们一定是合同的, 反之却未必.

例 5.4.1　若 n 阶矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 中有一个是正交矩阵, 则 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似且合同.

证　不妨设 $\boldsymbol{A}$ 是正交矩阵, 则 $\boldsymbol{A}$ 可逆, 且

$$\boldsymbol{A}^{-1}(\boldsymbol{AB})\boldsymbol{A}=\boldsymbol{BA},$$

故 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 相似. 又 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$, 所以 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 也合同.

最后, 谈谈矩阵三种关系中的标准形:

1° 等价: $\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$ (r 是 $\boldsymbol{A}$ 的秩);

2° 合同: $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\begin{pmatrix}\boldsymbol{I}_p&&\\&-\boldsymbol{I}_q&\\&&\boldsymbol{O}\end{pmatrix}$ (p,q 是 $\boldsymbol{A}$ 的正、负惯性指数);

3° 相似: $\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{pmatrix}$ ($\lambda_1,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的特征值).

归纳其特点可得到: 在矩阵的变换中, 限制越少则适应的范围就越宽, 化简后的形式就越简单, 但变换中丢掉的原矩阵的性质就越多. 例如 1° 的限制最少, 对所有矩阵都适用, 其标准形最简单, 但变换中只保留了秩不变; 2° 与 3° 都增加了条件, 就不是对任意矩阵都成立, 其标准形也略为复杂, 但变换中所保留的不变量也就更多.

问题 5.5　实二次型 $x^{\mathrm{T}}Ax$ 在条件 $\|x\|=1$ 下一定有最大与最小值吗

实二次型 $f(x)=x^{\mathrm{T}}Ax$ 是定义在 $\mathbf{R}^n$ 上的 n 元连续函数, 单位球面 $\|x\|=1$ 是 $\mathbf{R}^n$ 中的有界闭集, 因此 $f(x)$ 在 $\|x\|=1$ 上一定能取到最大值和最小值, 而且这个最大值与最小值分别就是实对称矩阵 A 的最大特征值和最小特征值. 事实上, 因为 A 是实对称矩阵, 故存在正交矩阵 Q 使得

$$A=Q^{\mathrm{T}}\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)Q,$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的 n 个特征值, 不妨设 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$.

对 $\forall x\in\mathbf{R}^n$, $\|x\|=1$, 记 $Qx=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}=y$, 则

$$y^{\mathrm{T}}y=(Qx)^{\mathrm{T}}(Qx)=x^{\mathrm{T}}Q^{\mathrm{T}}Qx=x^{\mathrm{T}}x=1$$

且

$$\begin{aligned}f(x)&=x^{\mathrm{T}}Ax=x^{\mathrm{T}}Q^{\mathrm{T}}\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)Qx=y^{\mathrm{T}}\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)y\\&=\sum_{i=1}^{n}\lambda_iy_i^2\leqslant\lambda_1\sum_{i=1}^{n}y_i^2=\lambda_1y^{\mathrm{T}}y=\lambda_1.\end{aligned}$$

另一方面设 α_1 是矩阵 A 的对应于特征值 λ_1 的单位特征向量, 则

$$f(\alpha_1)=\alpha_1^{\mathrm{T}}A\alpha_1=\boldsymbol{\lambda}_1\alpha_1^{\mathrm{T}}\alpha_1=\boldsymbol{\lambda}_1.$$

所以 $f(x)$ 在条件 $\|x\|=1$ 下的最大值是矩阵 A 的最大特征值.

类似地, 可证明 $f(x)$ 在条件 $\|x\|=1$ 下的最小值是矩阵 A 的最小特征值.

问题 5.6　为什么说二次型的规范形是唯一的, 它有何意义

设 $f(x)=x^{\mathrm{T}}Ax$ 是一个 n 元实二次型, 我们知道, 对实对称矩阵 A, 一定存在正交矩阵 Q 使得

$$A=Q^{\mathrm{T}}\mathrm{diag}\,(\lambda_1,\lambda_2,\cdots,\lambda_n)Q,\tag{5.6.1}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 A 的 n 个特征值. 该结论对二次型来说, 就是存在正交变换 $y=Qx$ 使得

$$f=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.\tag{5.6.2}$$

不妨设 rank $(\boldsymbol{A}) = r \leqslant n$, $\lambda_1, \cdots, \lambda_p > 0, \lambda_{p+1}, \cdots, \lambda_r < 0\ (0 \leqslant p \leqslant r)$. 再作可逆线性变换

$$\begin{cases} z_1 = \sqrt{\lambda_1} y_1, \\ \cdots\cdots\cdots\cdots \\ z_p = \sqrt{\lambda_p} y_p, \\ z_{p+1} = \sqrt{-\lambda_{p+1}} y_{p+1}, \\ \cdots\cdots\cdots\cdots \\ z_r = \sqrt{-\lambda_r} y_r, \\ z_{r+1} = y_{r+1}, \\ \cdots\cdots\cdots\cdots \\ z_n = y_n, \end{cases}$$

则 (5.6.2) 式可化为

$$f = z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2. \tag{5.6.3}$$

由于矩阵的特征值是由矩阵自身唯一确定的, 所以 (5.6.2) 式中正负号的项数也就是由二次型唯一确定的, 因而 (5.6.3) 式也就是唯一的. 我们通常将 (5.6.3) 称为二次型 f 的**规范形**, 它完全由二次型的秩 r 及正惯性指数 p 所确定, 并且 r 和 p 与所作的线性变换无关. 这就是二次型的**惯性定理**.

以上表述仅仅是对惯性定理的一种看似合理的解释, 但它并不是一种严格的证明. 下面来证明 (5.6.3) 式的唯一性, 即证明 r 和 p 与所作的线性变换无关.

证 由于可逆线性变换不改变二次型的秩, 不妨设实二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 经不同的可逆线性变换化为了规范形 (5.6.3) 与下面的形式

$$f = t_1^2 + t_2^2 + \cdots + t_q^2 - t_{q+1}^2 - \cdots - t_r^2, \tag{5.6.4}$$

即有

$$z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 = t_1^2 + t_2^2 + \cdots + t_q^2 - t_{q+1}^2 - \cdots - t_r^2. \tag{5.6.5}$$

下面证明一定有 $p = q$. 否则, 不妨设 $p > q$. 并设 (5.6.3) 与 (5.6.4) 所用的可逆线性变换为 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{z}$, $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{t}$, 则

$$\boldsymbol{t} = \boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{z}. \tag{5.6.6}$$

记 $\boldsymbol{C}^{-1}\boldsymbol{B}=\boldsymbol{D}=(d_{ij})_{n\times n}$, 将 (5.6.6) 具体写出来就是

$$\begin{cases} t_1=d_{11}z_1+d_{12}z_2+\cdots+d_{1n}z_n, \\ t_2=d_{21}z_1+d_{22}z_2+\cdots+d_{2n}z_n, \\ \cdots\cdots\cdots\cdots \\ t_n=d_{n1}z_1+d_{n2}z_2+\cdots+d_{nn}z_n. \end{cases} \tag{5.6.7}$$

考虑齐次线性方程组

$$\begin{cases} d_{11}z_1+d_{12}z_2+\cdots+d_{1n}z_n=0, \\ \cdots\cdots\cdots\cdots \\ d_{q1}z_1+d_{q2}z_2+\cdots+d_{qn}z_n=0, \\ z_{p+1}=0, \\ \cdots\cdots\cdots\cdots \\ z_n=0, \end{cases} \tag{5.6.8}$$

方程组含有 n 个未知量, 而方程的个数为 $q+(n-p)=n-(p-q)<n$, 故方程组有非零解. 设其非零解为

$$(z_1,\cdots,z_p,z_{p+1},\cdots,z_n)=(k_1,\cdots,k_p,0,\cdots,0),$$

将这组非零解代入 (5.6.5) 的左边, 得其值为

$$f=k_1^2+k_2^2+\cdots+k_p^2>0;$$

通过 (5.6.7) 代入 (5.6.5) 的右边, 得其值为

$$f=-t_{q+1}^2-\cdots-t_r^2\leqslant 0.$$

这就出现了矛盾. 所以只可能有 $p\leqslant q$; 同理可证 $q\leqslant p$. 故 $p=q$.

惯性定理也可用矩阵合同的语言来描述, 下面给出具体的描述与不同于上面方法的证明.

定理 5.6.1 (惯性定理) 若存在可逆矩阵 $\boldsymbol{Q}$ 使得

$$\boldsymbol{Q}^{\mathrm{T}}\begin{pmatrix} \boldsymbol{I}_p & & \\ & -\boldsymbol{I}_q & \\ & & \boldsymbol{O} \end{pmatrix}\boldsymbol{Q}=\begin{pmatrix} \boldsymbol{I}_{p'} & & \\ & -\boldsymbol{I}_{q'} & \\ & & \boldsymbol{O} \end{pmatrix} \tag{5.6.9}$$

则必有 $p=p'$, $q=q'$.

证 由于合同矩阵必有相同的秩, 故 $p+q=p'+q'$.

下证 $p=p'$. 若不然, 设有 $p'>p$, 将 $\boldsymbol{Q}$ 按列分块, 设 $\boldsymbol{Q}_1$ 是 $\boldsymbol{Q}$ 的前 p' 列构成的子块, 根据分块矩阵的乘法, 并由 (5.6.9) 式可得

$$\boldsymbol{I}_{p'}=\boldsymbol{Q}_1^{\mathrm{T}}\begin{pmatrix}\boldsymbol{I}_p & & \\ & -\boldsymbol{I}_q & \\ & & \boldsymbol{O}\end{pmatrix}\boldsymbol{Q}_1. \tag{5.6.10}$$

再将 $\boldsymbol{Q}_1$ 按行分块为 $\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{L}\\ \boldsymbol{M}\\ \boldsymbol{N}\end{pmatrix}$ (其中 $\boldsymbol{L}$ 是 $p\times p'$ 矩阵, $\boldsymbol{M}$ 是 $q\times p'$ 矩阵), (5.6.10) 可表示为

$$\boldsymbol{I}_{p'}=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}-\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}.$$

从而有

$$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}=\boldsymbol{I}_{p'}+\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}. \tag{5.6.11}$$

由于 $\operatorname{rank}(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})\leqslant\operatorname{rank}(\boldsymbol{L})\leqslant p<p'$; 而 $\boldsymbol{I}_{p'}+\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}$ 显然是正定矩阵 (或 $\boldsymbol{I}_{p'}+\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}=(\boldsymbol{I}_{p'},\boldsymbol{M}^{\mathrm{T}})\begin{pmatrix}\boldsymbol{I}_{p'}\\ \boldsymbol{M}\end{pmatrix}$ 是满秩矩阵), $\operatorname{rank}(\boldsymbol{I}_{p'}+\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})=p'$. 所以 (5.6.11) 式不可能成立, 矛盾. 所以 $p'\leqslant p$. 同理可证 $p\leqslant p'$, 从而必有 $p=p'$, 自然必有 $q=q'$.

惯性定理告诉我们, 两个实二次型可以通过可逆线性变换互化 (或两个实对称矩阵合同) 的充要条件是它们有相同的正、负惯性指数. 惯性指数是合同变换下的不变量, 从而可利用惯性指数来对二次型进行合同分类. 特别地, 几何空间中二次曲面的分类问题也就得到了相应的解决 (见问题 5.9, 第六章例 6.30).

问题 5.7　化二次型为标准形有哪些常见方法? 这些方法有何改进

我们知道, 化实二次型为标准形的常用方法有配方法与正交变换法. 利用矩阵合同的概念, 还不难得到合同变换法.

设实对称矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 合同, 则存在可逆矩阵 $\boldsymbol{C}$ 使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{\Lambda}.$$

将 $\boldsymbol{C}$ 表示为一系列初等矩阵的乘积, 设 $\boldsymbol{C}=\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k$, 则

$$\boldsymbol{E}_k^{\mathrm{T}}\boldsymbol{E}_{k-1}^{\mathrm{T}}\cdots\boldsymbol{E}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k=\boldsymbol{\Lambda},\quad \text{且 } \boldsymbol{C}=\boldsymbol{I}\boldsymbol{E}_1\boldsymbol{E}_2\cdots\boldsymbol{E}_k. \tag{5.7.1}$$

(5.7.1) 式表明, 将 $\boldsymbol{A}$ 作一系列初等行变换与相同的列变换 (称为合同变换) 化为对角矩阵 $\boldsymbol{\Lambda}$, 则所用列变换就将单位矩阵化成了 $\boldsymbol{C}$. 此过程可用图表示如下

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{pmatrix} \xrightarrow{\text{相同的行、列变换}} \begin{pmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{C} \end{pmatrix}, \text{ 则 } \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{\Lambda}. \tag{5.7.2}$$

例 5.7.1 用合同变换将矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & -4 \end{pmatrix}$ 化为标准形.

解 $$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{I} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & -4 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[c_2-c_1]{r_2-r_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & -4 \\ \hline 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[c_3+c_2]{r_3+r_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \\ \hline 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

由此得 $\boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$, $C = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, 且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{\Lambda}$.

下面就配方法与合同变换法的改进作一些讨论.

(1) 配方法的改进 —— 偏导数法

配方法是利用完全平方公式对二次型中的变量 $x_1, x_2 \cdots, x_n$ 从左到右逐一进行配方, 直到不含任何变量的交叉项为止. 但在配方的过程中, 一些初学者常常会有变量组合上的困难. 对此, 如果我们借助于函数偏导数的记号, 就可使这一问题得到较好的解决, 并且可使求解过程更为简便. 针对配方法的两种不同的情形, 其具体方法如下:

情形 1 如果 $f(x_1, \cdots, x_n)$ 中含有某变量的平方项, 即 $a_{ii}\ (i = 1, \cdots, n)$ 中至少有一个不为零, 不妨设 $a_{11} \neq 0$, 记 $f_1 = \dfrac{1}{2}\dfrac{\partial f}{\partial x_1}$, 令

$$f(x_1, \cdots, x_n) = \frac{1}{a_{11}}(f_1)^2 + g,$$

求得 g. 此时 g 中已不含 x_1, 再记 $g_1 = \dfrac{1}{2}\dfrac{\partial g}{\partial x_2}$, 并令

$$f(x_1, \cdots, x_n) = \frac{1}{a_{11}}(f_1)^2 + \frac{1}{a_{22}}(g_1)^2 + h,$$

此时 h 中已不含 x_1 与 x_2. 按这种程序继续运算, 可将二次型化为标准形.

情形 2 如果 $f(x_1,\cdots,x_n)$ 中不含有任一变量的平方项, 即 $a_{ii}=0$ $(i=1,\cdots,n)$, 但至少有一个 $a_{1j}\neq 0$ $(j>1)$ 不为零 (a_{ij} 是 x_ix_j 项的系数), 不妨设 $a_{12}\neq 0$, 记 $f_1=\dfrac{1}{2}\dfrac{\partial f}{\partial x_1}$, $f_2=\dfrac{1}{2}\dfrac{\partial f}{\partial x_2}$, 令

$$f(x_1,\cdots,x_n)=\frac{1}{a_{12}}[(f_1+f_2)^2-(f_1-f_2)^2]+\varphi,$$

求得 φ. 此时 φ 中已不含 x_1 与 x_2. 观察 φ 的结构, 如果 φ 中含有变量的平方项, 则按情形 1 中的方法进行, 否则按情形 2 中的方法进行, 直至二次型化为标准形.

例 5.7.2 用偏导数法化二次型 $f(x_1,x_2,x_3)=x_1^2+2x_2^2-4x_3^2+2x_1x_2-2x_2x_3$ 为标准形.

解 二次型中含有 x_1^2, 符合情形 1, 记 $f_1=\dfrac{1}{2}\dfrac{\partial f}{\partial x_1}=x_1+x_2$, 令

$$f(x_1,x_2,x_3)=\frac{1}{a_{11}}(f_1)^2+g=(x_1+x_2)^2+g.$$

整理并与原二次型对比可得到

$$g=x_2^2-4x_3^2-2x_2x_3.$$

记 $g_1=\dfrac{1}{2}\dfrac{\partial g}{\partial x_2}=x_2-x_3$, 有

$$f(x_1,x_2,x_3)=(x_1+x_2)^2+(x_2-x_3)^2-5x_3^2.$$

作可逆线性变换 $\begin{cases}y_1=x_1+x_2,\\ y_2=\qquad x_2-x_3,\\ y_3=\qquad\qquad x_3,\end{cases}$ 即 $\begin{cases}x_1=y_1-y_2-y_3,\\ x_2=\qquad y_2+y_3,\\ x_3=\qquad\qquad y_3,\end{cases}$ 则有

$$f=y_1^2+y_2^2-5y_3^2.$$

例 5.7.3 用偏导数法化二次型 $f(x_1,x_2,x_3)=-4x_1x_2+2x_1x_3+2x_2x_3$ 为标准形.

解 二次型中不含有任一变量的平方项, 符合情形 2, 记 $f_1=\dfrac{1}{2}\dfrac{\partial f}{\partial x_1}=-2x_2+x_3$, $f_2=\dfrac{1}{2}\dfrac{\partial f}{\partial x_2}=-2x_1+x_3$, 令

$$\begin{aligned}f(x_1,x_2,x_3)&=\frac{1}{a_{12}}[(f_1+f_2)^2-(f_1-f_2)^2]+\varphi\\&=-\frac{1}{4}[(-2x_1-2x_2+2x_3)^2-(2x_1-2x_2)^2]+\varphi,\end{aligned}$$

整理并与原二次型对比可得到 $\varphi = x_3^2$. 所以

$$f = -(x_1+x_2-x_3)^2 + (x_1-x_2)^2 + x_3^2.$$

作可逆线性变换 $\begin{cases} y_1 = x_1 + x_2 - x_3, \\ y_2 = x_1 - x_2, \\ y_3 = \qquad\qquad x_3, \end{cases}$ 即 $\begin{cases} x_1 = \dfrac{1}{2}(y_1+y_2+y_3), \\ x_2 = \dfrac{1}{2}(y_1-y_2+y_3), \\ x_3 = \qquad\qquad\quad y_3, \end{cases}$ 则有

$$f = -y_1^2 + y_2^2 + y_3^2.$$

(2) 合同变换法的改进 —— 行初等变换法

对上面的 (5.7.1) 式, 我们也可以这样看: 对 $\boldsymbol{A}$ 作一系列初等行变换化为上三角形矩阵, 相同的列变换就将它化成了对角矩阵 $\boldsymbol{\varLambda}$ (上三角形矩阵的对角元就是 $\boldsymbol{\varLambda}$ 的对角元), 所用行变换同时将单位矩阵化成了 $\boldsymbol{C}^{\mathrm{T}}$. 此过程可用图表示如下

$$(\boldsymbol{A},\boldsymbol{I}) \xrightarrow{\text{行初等变换}} \left(\begin{pmatrix} d_1 & & * \\ & \ddots & \\ & & d_n \end{pmatrix}, \boldsymbol{C}^{\mathrm{T}}\right), \text{ 则 } \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}. \tag{5.7.3}$$

(5.7.3) 显然较 (5.7.2) 节省了大量的计算.

例 5.7.4 用行初等变换法将例 5.7.1 中的矩阵化为标准形.

解 利用 (5.7.3) 有

$$(\boldsymbol{A},\boldsymbol{I}) = \left(\begin{array}{ccc:ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & -4 & 0 & 0 & 1 \end{array}\right) \xrightarrow{r_2-r_1} \left(\begin{array}{ccc:ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & -1 & -4 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{r_3+r_2} \left(\begin{array}{ccc:ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -5 & -1 & 1 & 1 \end{array}\right).$$

由此得 $\boldsymbol{\varLambda} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, 且 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{\varLambda}$. 这与例 5.7.1 中的结果是一致的.

问题 5.8 如何理解正定二次型与正定矩阵的重要性

正定二次型与正定矩阵是二次型部分的一个十分重要的内容, 这是因为正定二次型与正定矩阵具有良好的性质和广泛的应用. 下面我们就以矩阵形式为主, 作些举例说明.

例 5.8.1 正定矩阵具有我们熟知的如下充分必要条件:

实对称矩阵 $\boldsymbol{A}$ 为正定矩阵

$\Leftrightarrow$ $\boldsymbol{A}$ 的特征值全为正实数

$\Leftrightarrow$ $\boldsymbol{A}$ 与单位矩阵合同, 即存在可逆矩阵 $\boldsymbol{C}$, 使 $\boldsymbol{A}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$

$\Leftrightarrow$ $\boldsymbol{A}$ 的各阶顺序主子式全大于零.

以上结论在一般线性代数教材中都有较详细的讨论, 但最后一条很少有证明, 这里给出它的一个证明.

证 必要性. 设 $\boldsymbol{A}$ 是 n 阶正定矩阵, $\boldsymbol{A}_k$ $(1\leqslant k\leqslant n)$ 是 $\boldsymbol{A}$ 的 n 阶顺序主子矩阵. 对任一 k 维非零实向量 $\boldsymbol{x}_k=(x_1,x_2,\cdots,x_k)^{\mathrm{T}}$, 作 n 维非零向量 $\boldsymbol{x}=(x_1,\cdots,x_k,0,\cdots,0)^{\mathrm{T}}$, 则

$$\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{x}_k=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0,$$

所以二次型 $\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{x}_k$ 正定, $\boldsymbol{A}_k$ 是正定矩阵, 因而 $\det(\boldsymbol{A}_k)>0$.

充分性. 对实对称矩阵 $\boldsymbol{A}$ 的阶数 n 用数学归纳法.

当 $n=1$ 时显然. 设命题对 $n-1$ 阶实对称矩阵已成立.

对 n 阶实对称矩阵 $\boldsymbol{A}$, 将其分块为

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

由于 $|\boldsymbol{A}_{n-1}|>0$, $\boldsymbol{A}_{n-1}$ 可逆, 作合同变换

$$\begin{pmatrix}\boldsymbol{I}_{n-1} & \boldsymbol{0}\\ -\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1} & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\beta}\\ \boldsymbol{\beta}^{\mathrm{T}} & a_{nn}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_{n-1} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\\ \boldsymbol{0} & 1\end{pmatrix}$$
$$=\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{0}\\ \boldsymbol{0} & a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\end{pmatrix}.$$

记 $d_n=a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}$, 上式两边取行列式得 $|\boldsymbol{A}|=|\boldsymbol{A}_{n-1}|d_n$, 因此 $d_n=\dfrac{|\boldsymbol{A}|}{|\boldsymbol{A}_{n-1}|}>0$.

由归纳假设 $\boldsymbol{A}_{n-1}$ 是正定的, 则存在 $n-1$ 阶可逆矩阵 $\boldsymbol{C}_{n-1}$ 使 $\boldsymbol{A}_{n-1}=\boldsymbol{C}_{n-1}^{\mathrm{T}}\boldsymbol{C}_{n-1}$. 则

$$\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{0}\\ \boldsymbol{0} & d_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{C}_{n-1} & \boldsymbol{0}\\ \boldsymbol{0} & \sqrt{d_n}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\boldsymbol{C}_{n-1} & \boldsymbol{0}\\ \boldsymbol{0} & \sqrt{d_n}\end{pmatrix}$$

也是正定矩阵, 故与之合同的矩阵 $\boldsymbol{A}$ 也是正定矩阵.

例 5.8.2 设 $\boldsymbol{A}$ 是 n 阶正定矩阵, 则以下命题成立:

(1) $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$ 及伴随矩阵 $\boldsymbol{A}^{*}$ 也都是正定矩阵;

(2) 若 $\boldsymbol{B}$ 是 n 阶非零半正定矩阵, 则 $\boldsymbol{A}+\boldsymbol{B}$ 是正定矩阵, 且 $|\boldsymbol{A}+\boldsymbol{B}|>|\boldsymbol{A}|+|\boldsymbol{B}|$;

(3) 存在正定矩阵 $\boldsymbol{B}$, 使 $\boldsymbol{A}=\boldsymbol{B}^2$.

证 (1) 设 $\lambda_i\ (i=1,\cdots,n)$ 是 $\boldsymbol{A}$ 的 n 个特征值, 由于 $\boldsymbol{A}$ 是正定矩阵, 所以它们全大于零.

易验证 $\boldsymbol{A}^{-1}$ 与 $\boldsymbol{A}^{*}$ 都是实对称矩阵, 且 $\boldsymbol{A}^{-1}$ 与 $\boldsymbol{A}^{*}$ 的特征值分别为 $\dfrac{1}{\lambda_i}$ 与 $\dfrac{|\boldsymbol{A}|}{\lambda_i}\ (i=1,\cdots,n)$ 也全大于零, 所以 $\boldsymbol{A}^{-1}$ 及 $\boldsymbol{A}^{*}$ 也都是正定矩阵.

(2) 因为对任一非零实向量 $\boldsymbol{x}$, 有

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}>0,$$

所以 $\boldsymbol{A}+\boldsymbol{B}$ 是正定矩阵.

下证 $|\boldsymbol{A}+\boldsymbol{B}|>|\boldsymbol{A}|+|\boldsymbol{B}|$.

由 $\boldsymbol{A}$ 正定知, 存在可逆矩阵 $\boldsymbol{P}$, 使

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{I} \tag{5.8.1}$$

而 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$ 仍是非零半正定矩阵.

对 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$, 存在正交矩阵 $\boldsymbol{Q}$, 使

$$\boldsymbol{Q}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P})\boldsymbol{Q}=\operatorname{diag}(b_1,\cdots,b_n)\triangleq\boldsymbol{\Lambda}, \tag{5.8.2}$$

其中 $b_i\geqslant 0\ (i=1,\cdots,n)$ 是 $\boldsymbol{B}$ 的 n 个特征值 (不全为零).

记 $\boldsymbol{R}=\boldsymbol{P}\boldsymbol{Q}$, 利用 (5.8.1) 与 (5.8.2) 式, 有

$$|\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{R}|=|\boldsymbol{I}+\boldsymbol{\Lambda}|=\prod_{i=1}^{n}(1+b_i)>1+\prod_{i=1}^{n}b_i=|\boldsymbol{I}|+|\boldsymbol{\Lambda}|=|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{R}|+|\boldsymbol{R}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{R}|.$$

利用上式左右两端, 得

$$|\boldsymbol{R}|^2|\boldsymbol{A}+\boldsymbol{B}|>|\boldsymbol{R}|^2(|\boldsymbol{A}|+|\boldsymbol{B}|),$$

所以 $|\boldsymbol{A}+\boldsymbol{B}|>|\boldsymbol{A}|+|\boldsymbol{B}|$.

(3) 由 $\boldsymbol{A}$ 正定知, 存在正交矩阵 $\boldsymbol{Q}$使

$$\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{T}}\operatorname{diag}(\lambda_1,\cdots,\lambda_n)\boldsymbol{Q}, \tag{5.8.3}$$

其中 $\lambda_i>0\ (i=1,\cdots,n)$ 是 $\boldsymbol{A}$ 的 n 个特征值. 从而

$$\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})\boldsymbol{Q}.$$

记 $\boldsymbol{B}=\boldsymbol{Q}^{\mathrm{T}}\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})\boldsymbol{Q}$, 由于 $\operatorname{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})$ 正定, 所以 $\boldsymbol{B}$ 正定, 且有 $\boldsymbol{A}=\boldsymbol{B}^2$.

例 5.8.3　设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是 n 阶实对称矩阵, $\boldsymbol{A}$ 的特征值大于 a, $\boldsymbol{B}$ 的特征值大于 b, 证明 $\boldsymbol{A}+\boldsymbol{B}$ 的特征值大于 $a+b$.

证　设 $\boldsymbol{A}$ 的 n 个特征值为 $\lambda_i\ (i=1,\cdots,n)$, 则 $\boldsymbol{A}-a\boldsymbol{I}$ 的 n 个特征值为 $\lambda_i-a\ (i=1,\cdots,n)$. 已知 $\boldsymbol{A}$ 的特征值大于 a, 故 $\lambda_i-a>0\ (i=1,\cdots,n)$, 所以 $\boldsymbol{A}-a\boldsymbol{I}$ 是正定矩阵. 同理 $\boldsymbol{B}-b\boldsymbol{I}$ 是正定矩阵. 由此得到 $(\boldsymbol{A}-a\boldsymbol{I})+(\boldsymbol{B}-b\boldsymbol{I})=\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{I}$ 是正定矩阵.

设 $\boldsymbol{A}+\boldsymbol{B}$ 的特征值为 $l_i\ (i=1,\cdots,n)$, 则 $\boldsymbol{A}+\boldsymbol{B}-(a+b)\boldsymbol{I}$ 的特征值为

$$l_i-(a+b)>0,\ 即有\ l_i>(a+b)\ (i=1,\cdots,n).$$

例 5.8.4　利用多元函数 $f(x_1,\cdots,x_n)$ 在其驻点 $\boldsymbol{x}_0$ 处 Hesse (黑塞, 1811—1874) 矩阵 $\boldsymbol{H}_f(\boldsymbol{x}_0)=\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}\right)_{\boldsymbol{x}_0}$ 的正 (负) 定性, 可判定 f 在 $\boldsymbol{x}_0$ 点是否取得极值.

设 $f(x_1,\cdots,x_n)$ 在 $\boldsymbol{x}_0$ 点附近具有二阶连续偏导数, $\boldsymbol{x}_0$ 是其驻点, 即 $\nabla f(\boldsymbol{x}_0)=0$, 由 Taylor 公式

$$f(\boldsymbol{x}_0+\Delta\boldsymbol{x})=f(\boldsymbol{x}_0)+\frac{1}{2!}(\Delta\boldsymbol{x})^{\mathrm{T}}\boldsymbol{H}_f(\boldsymbol{x}_0)\Delta\boldsymbol{x}+o(\|\Delta\boldsymbol{x}\|^2),$$

即有

$$f(\boldsymbol{x}_0+\Delta\boldsymbol{x})-f(\boldsymbol{x}_0)=\frac{1}{2!}(\Delta\boldsymbol{x})^{\mathrm{T}}\boldsymbol{H}_f(\boldsymbol{x}_0)\Delta\boldsymbol{x}+o(\|\Delta\boldsymbol{x}\|^2).$$

当 $\|\Delta\boldsymbol{x}\|$ 充分小时, 上式右端的符号由二次型 $(\Delta\boldsymbol{x})^{\mathrm{T}}\boldsymbol{H}_f(\boldsymbol{x}_0)\Delta\boldsymbol{x}$ 确定. 由此可得:

若 $\boldsymbol{H}_f(\boldsymbol{x}_0)$ 正定, 则在 $\boldsymbol{x}_0$ 点附近有 $f(\boldsymbol{x}_0+\Delta\boldsymbol{x})-f(\boldsymbol{x}_0)\geqslant 0$, f 在点 $\boldsymbol{x}_0$ 取得极小值;

若 $\boldsymbol{H}_f(\boldsymbol{x}_0)$ 负定, 则在 $\boldsymbol{x}_0$ 点附近有 $f(\boldsymbol{x}_0+\Delta\boldsymbol{x})-f(\boldsymbol{x}_0)\leqslant 0$, f 在点 $\boldsymbol{x}_0$ 取得极大值;

若 $\boldsymbol{H}_f(\boldsymbol{x}_0)$ 不定, 则在 $\boldsymbol{x}_0$ 点附近 $f(\boldsymbol{x}_0+\Delta\boldsymbol{x})-f(\boldsymbol{x}_0)$ 的符号可正可负, f 在点 $\boldsymbol{x}_0$ 不取得极值.

例 5.8.5 正定矩阵与 $\mathbf{R}^n$ 空间中的内积.

$\mathbf{R}^n$ 中向量的内积给空间增添了"度量", 由此可诱导出一系列数学概念, 如长度、夹角、正交、极限等. 我们知道 $\mathbf{R}^n$ 中向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积定义为实数 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}$, 但这仅仅是内积的一种具体形式, 我们还可以有其他不同的方式来定义两向量的内积. 抽象来看, 实线性空间 V 中向量的内积可定义如下:

定义 5.8.1 设 V 是一实线性空间, 若 $\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, 都有一实数与之对应, 记成 $(\boldsymbol{\alpha},\boldsymbol{\beta})$, 并且满足以下三条, 则称实数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$ 叫 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积.

1° 对称性 $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$;

2° 线性性 $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$, $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$, k 是任意实数;

3° 非负性 $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, $(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$.

定义了内积的 $\mathbf{R}^n$ 空间称为欧氏空间.

事实上, 任意一个 n 阶正定矩阵 $\boldsymbol{A}$ 都可以确定 $\mathbf{R}^n$ 中的一个内积:

$$(\boldsymbol{\alpha},\boldsymbol{\beta})\triangleq\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\beta} \tag{5.8.4}$$

容易验证由 (5.8.4) 定义的运算完全满足定义 5.8.1 中的各项条件.

反之 $\mathbf{R}^n$ 中任何一个内积都可确定出无限多个正定矩阵. 具体方法如下:

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\mathbf{R}^n$ 的一组基, 记 $a_{ij}=(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)$ $(i,j=1,2,\cdots,n)$, 则 n 阶矩阵 $\boldsymbol{A}=(a_{ij})$ 就是一个正定矩阵. 其原因为

(i) 由内积的对称性知 $\boldsymbol{A}$ 是一个实对称矩阵;

(ii) 对 $\forall\boldsymbol{x}\in\mathbf{R}^n$, 设 $\boldsymbol{x}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n\neq\mathbf{0}$, 由内积的性质有

$$0<(\boldsymbol{x},\boldsymbol{x})=\sum_{i,j=1}^{n}(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)x_ix_j=\sum_{i,j=1}^{n}a_{ij}x_ix_j,$$

这说明二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定的.

也就是说, 欧氏空间 $\mathbf{R}^n$ 中任意一组基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 都可以确定一个正定矩阵 $\boldsymbol{A}=(a_{ij})$, 该矩阵又叫由基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 确定的**度量矩阵**, 它具体刻画了 $\mathbf{R}^n$ 中的内积. 当且仅当 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\mathbf{R}^n$ 的一组标准正交基时, 其度量矩阵是单位矩阵. 此时, 向量 $\boldsymbol{x}=x_1\boldsymbol{\alpha}_1+\cdots+x_n\boldsymbol{\alpha}_n$ 与 $\boldsymbol{y}=y_1\boldsymbol{\alpha}_1+\cdots+y_n\boldsymbol{\alpha}_n$ 的内积为 $(\boldsymbol{x},\boldsymbol{y})=\sum\limits_{i=1}^{n}x_iy_i$.

读者不难证明: **欧氏空间中两组不同基的度量矩阵是合同的**.

例 5.8.6 正定线性变换的几何特性.

我们知道, 若 $\boldsymbol{A}$ 是一个 n 阶正交矩阵, 则对应的线性变换 $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$ 是空间 $\mathbf{R}^n$ 中的保形变换, 它能保持 $\mathbf{R}^n$ 中的度量不变. 其原因是 $\forall \boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$, 有

$$(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{y})=\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{y}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}, \tag{5.8.5}$$

即正交变换保持向量的内积.

如果 $\boldsymbol{A}$ 是一个正定矩阵, 显然 (5.8.5) 就不一定成立了, 也就是说正定变换一般不能保持 $\mathbf{R}^n$ 中的度量不变, 但此时有

$$(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0\ (\boldsymbol{x}\neq\boldsymbol{0}). \tag{5.8.6}$$

从几何上看, $\mathbf{R}^n$ 中的任一非零向量 $\boldsymbol{x}$, 与它垂直的超平面把空间分成两部分, 一部分和 $\boldsymbol{x}$ 在该超平面的同一侧, 即满足与 $\boldsymbol{x}$ 的内积为正的那侧; 另一部分和 $\boldsymbol{x}$ 在该超平面的异侧, 它们与 $\boldsymbol{x}$ 的内积为负. (5.8.6) 式表明, 正定的线性变换把任意一个非零向量 $\boldsymbol{x}$ 都变到 $\boldsymbol{x}$ 的同侧, 这就是正定线性变换的几何特性.

由此我们也容易解释为什么正定矩阵的特征值都必定是正数, 否则的话, 它会把对应的特征向量变到与之垂直的超平面的另一侧而出现矛盾.

正定矩阵还有很多良好的性质, 其理论在科学计算、计算机图形学、几何学、微分方程、概率统计、物理学等领域中都有广泛的应用.

问题 5.9 如何作直角坐标变换化一般二次曲面方程为标准方程

直角坐标变换是指对变量的正交变换或平移变换, 在该变换下曲面的形状保持不变. 将一般二次曲面方程化为标准方程的关键是将方程中的二次齐次项作正交变换化为标准形, 然后再对具有二次与一次项的变量作配方, 并作坐标平移变换, 消去一次项. 具体方法如下:

设一般二次曲面的方程为

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz+2b_1x+2b_2y+2b_3z+c=0.$$

写成矩阵形式为

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}+c=0, \tag{5.9.1}$$

其中 $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}(a_{ij}=a_{ji}),\ \boldsymbol{x}=(x,y,z)^{\mathrm{T}},\ \boldsymbol{B}=(b_1,b_2,b_3)^{\mathrm{T}}.$

作正交矩阵 $\boldsymbol{Q}$, 使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}\,(\lambda_1,\lambda_2,\lambda_3)$, 其中 $\lambda_1,\lambda_2,\lambda_3$ 是 $\boldsymbol{A}$ 的特征值. 化 (5.9.1) 为标准方程可分两种情况来进行.

(1) 若方程组 $\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}=\boldsymbol{0}$ 有解, 则取其任意一个解 $\boldsymbol{\delta}$, 作直角坐标变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{x}'+\boldsymbol{\delta}$, 代入 (5.9.1) 得

$$\boldsymbol{x}'^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{x}'+2(\boldsymbol{A}\boldsymbol{\delta}+\boldsymbol{B})^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}'+\boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\delta}+2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\delta}+c=0.$$

由 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}\,(\lambda_1,\lambda_2,\lambda_3)$, $\boldsymbol{A}\boldsymbol{\delta}+\boldsymbol{B}=\boldsymbol{0}$, 可得

$$\lambda_1x'^2+\lambda_2y'^2+\lambda_3z'^2=d, \tag{5.9.2}$$

其中 $d=-(\boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\delta}+2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\delta}+c)$. (5.9.2) 就是曲面 (5.9.1) 的标准方程.

(2) 若方程组 $\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}=\boldsymbol{0}$ 无解, 则必有 $|\boldsymbol{A}|=0$. 由于 $|\boldsymbol{A}|=\lambda_1\lambda_2\lambda_3$, 则 $\lambda_1,\lambda_2,\lambda_3$ 中至少有一个为 0, 不妨设 $\lambda_1\neq 0$, $\lambda_3=0$. 作正交变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{x}'$, 代入 (5.9.1) 得

$$\lambda_1x'^2+\lambda_2y'^2+2b_1'x'+2b_2'y'+2b_3'z'+c=0. \tag{5.9.3}$$

1° 若 $\lambda_2\neq 0$, 将 (5.9.3) 配方得

$$\lambda_1(x'+b_1')^2+\lambda_2(y'+b_2')^2=d-2b_3'z', \tag{5.9.4}$$

其中 $d=\lambda_1b_1'^2+\lambda_2b_2'^2-c$.

当 $b_3'\neq 0$ 时, 作平移 $\begin{cases}x''=x'+b_1',\\ y''=y'+b_2',\\ z''=z'-\dfrac{d}{2b_3'},\end{cases}$ 则 (5.9.4) 化为标准方程

$$\lambda_1x''^2+\lambda_2y''^2=-2b_3'z''. \tag{5.9.5}$$

当 $b_3'=0$ 时, 作平移 $\begin{cases}x''=x'+b_1',\\ y''=y'+b_2',\\ z''=z',\end{cases}$ 则 (5.9.4) 化为标准方程

$$\lambda_1x''^2+\lambda_2y''^2=d. \tag{5.9.6}$$

2° 若 $\lambda_2=0$, 将 (5.9.3) 配方得

$$\lambda_1(x'+b_1')^2=d-2b_2'y'-2b_3'z', \tag{5.9.7}$$

其中 $d=\lambda_1b_1'^2-c$.

当 b_2', b_3' 不全为 0 时, 作直角坐标变换

$$\begin{cases} x'' = x' + b_1', \\ y'' = \dfrac{1}{\sqrt{b_2'^2 + b_3'^2}}(b_2'y' + b_3'z') - \dfrac{d}{2}, \\ z'' = \dfrac{1}{\sqrt{b_2'^2 + b_3'^2}}(-b_2'y' + b_3'z'), \end{cases}$$

(5.9.7) 化为标准方程

$$\lambda_1 x''^2 = -2\sqrt{b_2'^2 + b_3'^2}\,y''. \tag{5.9.8}$$

当 b_2', b_3' 全为 0 时, 令 $\begin{cases} x'' = x' + b_1', \\ y'' = y', \\ z'' = z', \end{cases}$ 则 (5.9.7) 化为标准方程

$$\lambda_1 x''^2 = d. \tag{5.9.9}$$

例 5.9.1 作直角坐标变换化二次曲面

$$x^2 - 2y^2 + 10z^2 + 28xy + 20xz - 8yz - 26x + 32y + 28z - 38 = 0$$

为标准方程, 并说明是什么曲面.

解 设 $\boldsymbol{A} = \begin{pmatrix} 1 & 14 & 10 \\ 14 & -2 & -4 \\ 10 & -4 & 10 \end{pmatrix}$, $\boldsymbol{x} = (x, y, z)^{\mathrm{T}}$, $\boldsymbol{B} = (-13, 16, 14)^{\mathrm{T}}, c = -38$, 曲面方程记为

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + 2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{x} + c = 0. \tag{5.9.10}$$

$\boldsymbol{A}$ 的特征多项式为

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & -14 & -10 \\ -14 & \lambda + 2 & 4 \\ -10 & 4 & \lambda - 10 \end{vmatrix} = (\lambda - 9)(\lambda - 18)(\lambda + 18).$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1 = 9$, $\lambda_2 = 18$, $\lambda_3 = -18$, 求得对应的单位特征向量为

$$\boldsymbol{\alpha}_1 = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \boldsymbol{\alpha}_2 = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \quad \boldsymbol{\alpha}_3 = \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}.$$

得正交矩阵 $\boldsymbol{Q} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$.

因为 $|\boldsymbol{A}| \neq 0$, 方程组 $\boldsymbol{Ax}+\boldsymbol{B}=\boldsymbol{0}$ 有唯一解 $\boldsymbol{\delta}=(-1,1,0)^{\mathrm{T}}$, 作直角坐标变换 $\boldsymbol{x}=\boldsymbol{Qx}'+\boldsymbol{\delta}$, 方程 (5.9.10) 化为

$$9x'^2+18y'^2-18z'^2=9,$$

得曲面的标准方程为

$$x'^2+2y'^2-2z'^2=1.$$

曲面为单叶双曲面.

例 5.9.2 作直角坐标变换化二次曲面

$$x^2+4y^2+4z^2+4xy+4xz+8yz-6x+24y+6z-27=0$$

为标准方程, 并说明是什么曲面.

解 设 $\boldsymbol{A}=\begin{pmatrix}1&2&2\\2&4&4\\2&4&4\end{pmatrix}$, $\boldsymbol{x}=(x,y,z)^{\mathrm{T}}$, $\boldsymbol{B}=(-3,12,3)^{\mathrm{T}}$, $c=-27$, 曲面方程记为

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}+2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{x}+c=0. \tag{5.9.11}$$

$\boldsymbol{A}$ 的特征多项式为

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-4&-4\\-2&-4&\lambda-4\end{vmatrix}=\lambda^2(\lambda-9),$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1=9$, $\lambda_2=0$ (二重).

求得对应于 $\lambda_1=9$ 的特征向量为 $\boldsymbol{\alpha}_1=\left(\dfrac{1}{2},1,1\right)^{\mathrm{T}}$, 单位化得 $\boldsymbol{\beta}_1=\dfrac{1}{3}(1,2,2)^{\mathrm{T}}$;

求得对应于 $\lambda_2=0$ 的线性无关的特征向量为 $\boldsymbol{\alpha}_2=(-2,1,0)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3=(-2,0,1)^{\mathrm{T}}$, 正交化、单位化得

$$\boldsymbol{\beta}_2=\frac{1}{\sqrt{5}}(-2,1,0)^{\mathrm{T}},\ \boldsymbol{\beta}_3=\frac{1}{3\sqrt{5}}(-2,-4,5)^{\mathrm{T}}.$$

得正交矩阵 $\boldsymbol{Q}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$.

易判定, 方程组 $\boldsymbol{Ax}+\boldsymbol{B}=\boldsymbol{0}$ 无解. 作正交变换 $\boldsymbol{x}=\boldsymbol{Qx}'$, 方程 (5.9.11) 化为

$$9x'^2+18x'+\frac{36}{\sqrt{5}}y'-\frac{18}{\sqrt{5}}z'-27=0,$$

配方化为

$$(x'+1)^2=2\left[\frac{1}{\sqrt{5}}(-2y'+z')+2\right]. \tag{5.9.12}$$

作直角坐标变换

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dfrac{-2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \\ 0 & \dfrac{1}{\sqrt{5}} & \dfrac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

(5.9.12) 化为标准方程 $x''^2 = 2y''$, 这是抛物柱面.

第六章　应用实例

例 6.1　在电路设计中, 经常要把复杂的电路分割为局部电路, 每一个电路都用一个网络 "黑盒子" 来表示. 黑盒子的输入为 u_1, i_1, 输出为 u_2, i_2, 其输入输出关系用一个矩阵 $\boldsymbol{A}$ 来表示 (如图 6.1.1 所示): $\begin{pmatrix} u_2 \\ i_2 \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} u_1 \\ i_1 \end{pmatrix}$, $\boldsymbol{A}$ 是 2×2 矩阵, 称为该局部电路的传输矩阵.

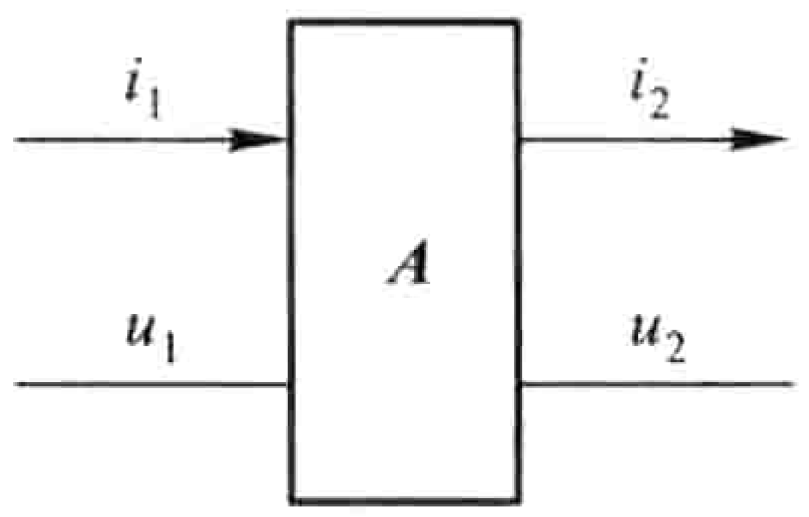


图 6.1.1　单个子网络模型

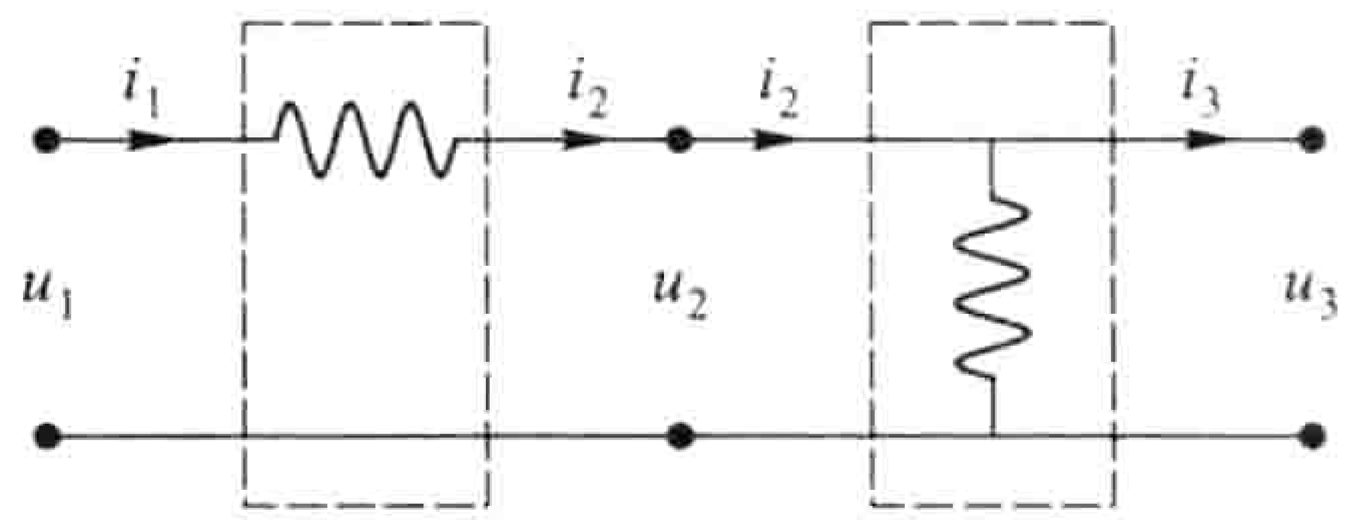


图 6.1.2　两个子网络模型

把复杂的电路分成许多串接局部电路, 分别求出它们的传输矩阵, 再相乘起来, 得到总的传输矩阵, 可以简化分析电路的工作. 以图 6.1.2 为例, 把两个电阻组成的分压电路分成两个串接的子网络. 第一个子网络只包含电阻 R_1, 第二个子网络只包含电阻 R_2, 请分别写出两个子网络的传输矩阵及整个电路的传输矩阵.

解　第一个子网络的电路方程为 $i_2 = i_1$, $u_2 = u_1 - i_1 R_1$. 写成矩阵形式为

$$\begin{pmatrix} u_2 \\ i_2 \end{pmatrix} = \begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ i_1 \end{pmatrix} = \boldsymbol{A}_1 \begin{pmatrix} u_1 \\ i_1 \end{pmatrix}.$$

同样可列出第二个子网络的电路方程, $i_3 = i_2 - u_2/R_2$, $u_3 = u_2$. 写成矩阵形式为

$$\begin{pmatrix} u_3 \\ i_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\dfrac{1}{R_2} & 1 \end{pmatrix} \begin{pmatrix} u_2 \\ i_2 \end{pmatrix} = \boldsymbol{A}_2 \begin{pmatrix} u_2 \\ i_2 \end{pmatrix}.$$

从上分别得到两个子网络的传输矩阵

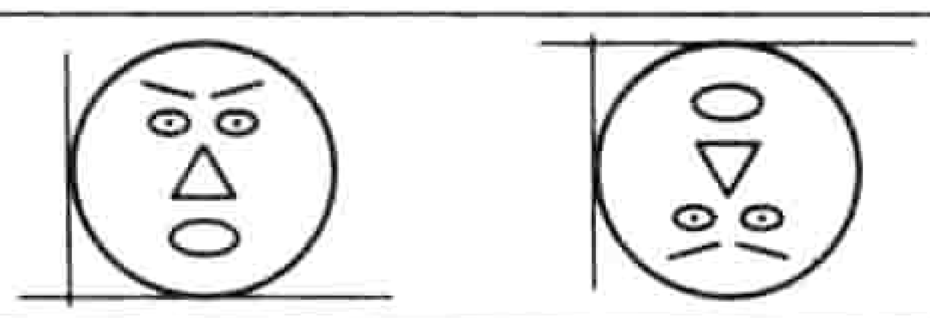
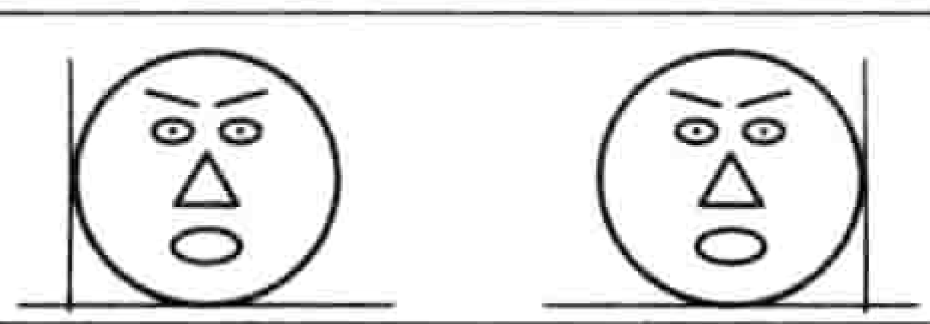
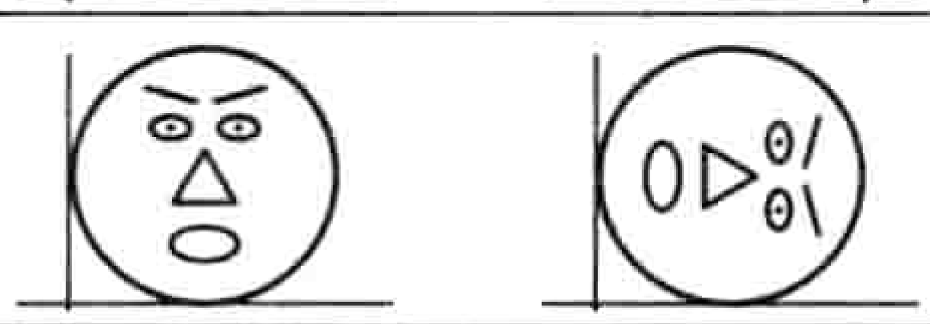
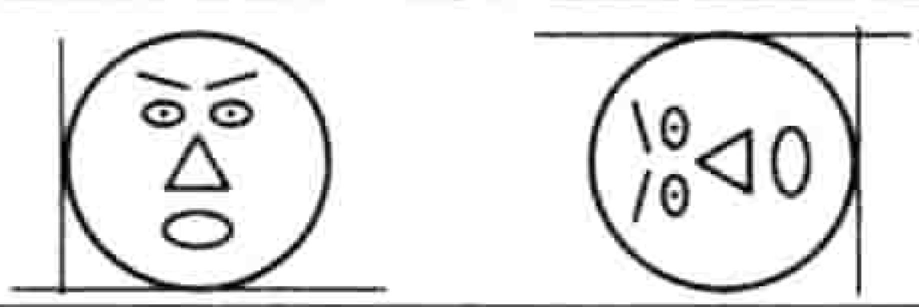
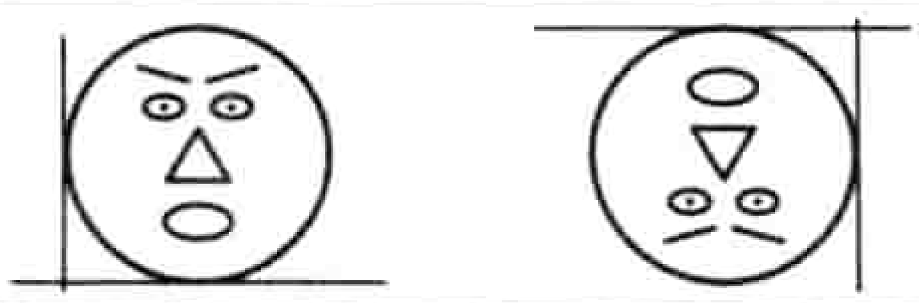
$$
\boldsymbol{A}_1 = \begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 \\ -\dfrac{1}{R_2} & 1 \end{pmatrix},
$$

整个电路的传输矩阵为两者的乘积

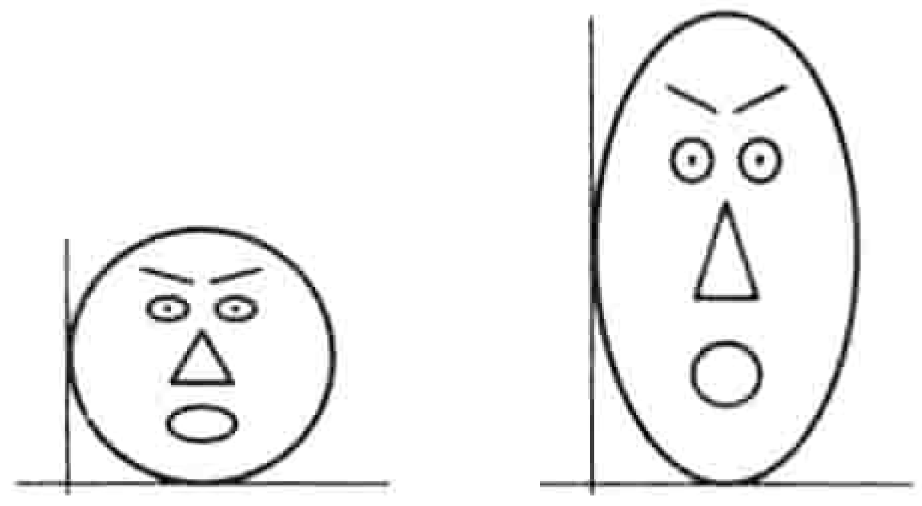
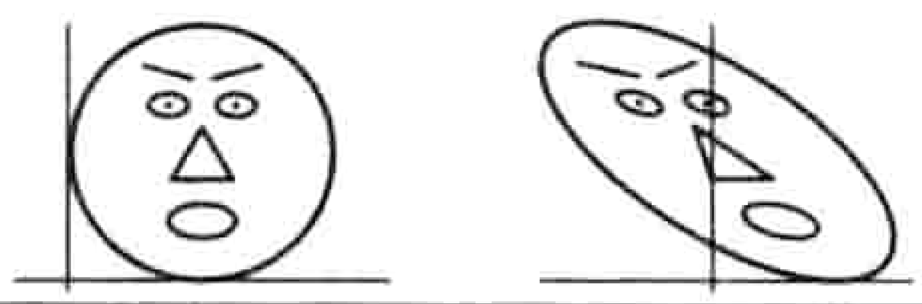
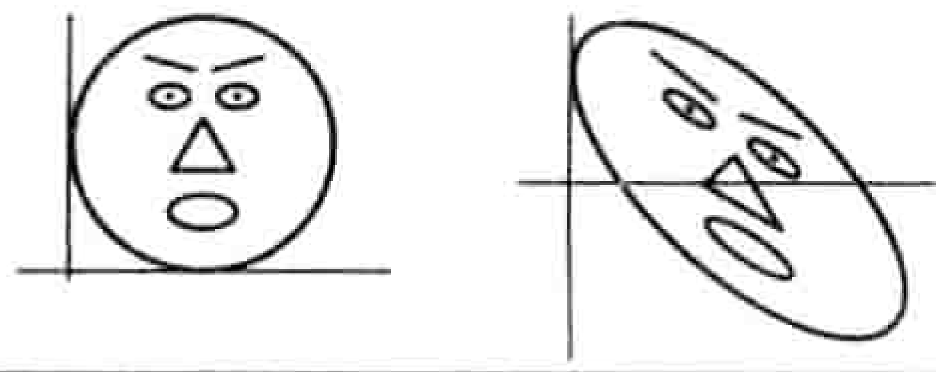
$$
\boldsymbol{A} = \boldsymbol{A}_1\boldsymbol{A}_2 = \begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\dfrac{1}{R_2} & 1 \end{pmatrix} = \begin{pmatrix} 1+\dfrac{R_1}{R_2} & -R_1 \\ -\dfrac{1}{R_2} & 1 \end{pmatrix}.
$$

例 6.2 (平面图形变换) 设 $\boldsymbol{x}, \boldsymbol{y}$ 为平面上的两个点 (或 $\mathbf{R}^2$ 中的两个列向量),$\boldsymbol{A}$ 为二阶矩阵, 则 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 为平面上的一个线性变换, 它把点 $\boldsymbol{x}$ 映射成点 $\boldsymbol{y}$. 若 $\boldsymbol{x}$ 为平面图形 G 上的任一点, $\boldsymbol{x}$ 的像 $\boldsymbol{y}$ 构成的平面图形记为 G_1, 则线性变换 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 的几何意义是把图形 G 变成图形 G_1, 称为平面图形变换. 平面图形变换有三种基本变换: 对称变换、伸缩变换、剪切 (错切) 变换, 如表 6.2.1 所示, 其他可逆变换均可由这三种变换复合而成.

表 6.2.1 三种基本变换

变换	变换前后的图像	变换矩阵
关于横轴的对称变换		$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
关于竖轴的对称变换		$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
关于 $y=x$ 的对称变换		$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
关于 $y=-x$ 的对称变换		$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
关于原点的对称变换		$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
水平伸缩变换		$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$

续表

变换	变换前后的图像	变换矩阵
垂直伸缩变换		$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$
水平剪切变换		$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$
垂直剪切变换		$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$

例 6.3 如图 6.3.1 所示, 水平弹性横梁的支撑点位于两端, 且在 #1、#2、#3 处被施加了作用力. 设 $\boldsymbol{f} \in \mathbf{R}^3$ 列出了施加在这三个点上的作用力, $\boldsymbol{y} \in \mathbf{R}^3$ 列出了横梁在这三个点的偏移 (移动). 利用物理学中的 Hooke (胡克, 1635—1703) 定律, 可以证明

$$\boldsymbol{y} = \boldsymbol{D}\boldsymbol{f},$$

其中 $\boldsymbol{D}$ 是柔度矩阵, 其逆矩阵称为刚度矩阵. 试描述矩阵 $\boldsymbol{D}$ 和 $\boldsymbol{D}^{-1}$ 中列向量的物理意义.

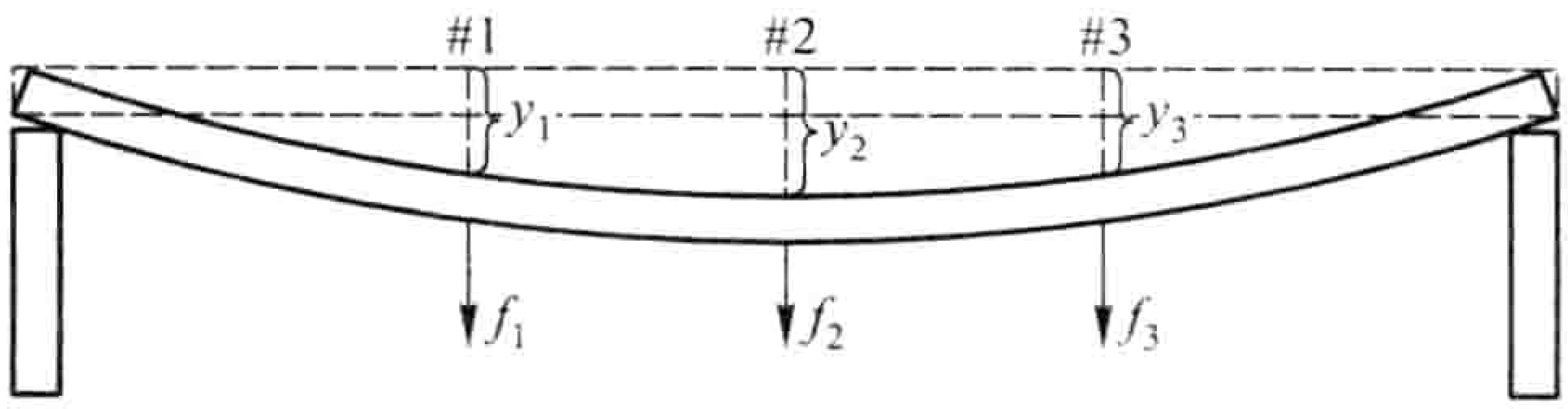


图 6.3.1 弹性横梁的偏移

解 记 $\boldsymbol{I}_3 = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$, 观察

$$\boldsymbol{D} = \boldsymbol{D}\boldsymbol{I}_3 = (\boldsymbol{D}\boldsymbol{e}_1, \boldsymbol{D}\boldsymbol{e}_2, \boldsymbol{D}\boldsymbol{e}_3),$$

将向量 $\boldsymbol{e}_1 = (1, 0, 0)^{\mathrm{T}}$ 解释为在横梁 #1 处施加向下的 1 单位作用力 (而其余两点处的作用力为 0). 则 $\boldsymbol{D}\boldsymbol{e}_1$, 即 $\boldsymbol{D}$ 的第一列, 列出了由点 #1 处的 1 单位作用力所引起的横梁偏移. $\boldsymbol{D}$ 的第二、三列也有类似的解释.

为了研究刚度矩阵 $\boldsymbol{D}^{-1}$, 注意到, 当偏移向量 $\boldsymbol{y}$ 给定时, 由方程 $\boldsymbol{f} = \boldsymbol{D}^{-1}\boldsymbol{y}$

可计算出作用力向量 $\boldsymbol{f}$. 记

$$\boldsymbol{D}^{-1}=\boldsymbol{D}^{-1}\boldsymbol{I}_3=(\boldsymbol{D}^{-1}\boldsymbol{e}_1,\boldsymbol{D}^{-1}\boldsymbol{e}_2,\boldsymbol{D}^{-1}\boldsymbol{e}_3),$$

将 $\boldsymbol{e}_1$ 解释成偏移向量, 则 $\boldsymbol{D}^{-1}\boldsymbol{e}_1$ 列出了产生这个偏移的作用力. 即 $\boldsymbol{D}^{-1}$ 的第一列列出了横梁三个点处的作用力, 它使点 #1 产生 1 单位偏移, 其余两点的偏移为零. 类似地, $\boldsymbol{D}^{-1}$ 的第二、三列分别列出了在点 #2 和点 #3 产生 1 单位偏移的作用力. 每一列中, 为使横梁在某点产生 1 单位偏移在另两点不偏移, 必须有一个或两个力是负 (向上) 的.

例 6.4　在行列式游戏 $Tic-Tac-Toc$ 中, 游戏者 1 在一个空的 3×3 矩阵中填入一个 1, 游戏者 0 在某一个空位置中填入一个 0, 游戏如此继续, 直到 3×3 矩阵填入了 5 个 1 和 4 个 0. 若此行列式值为 0, 则游戏者 0 获胜, 否则游戏者 1 获胜. 有没有一种策略保证某一个游戏者一定获胜?

解　游戏者 0 一定能获胜. 其原因如下:

如果游戏者 0 能够填好一行 0 或者一列 0, 或者一个 2×2 的元均为 0 的子式, 则行列式值为 0.

由于行的交换和列的交换仅仅影响行列式的符号, 因此不妨设游戏者 1 的第一次填入是在左上角. 此时游戏者 0 可以将 0 填入中心格. 利用对称性, 只需考虑游戏者 1 填入如下位置:

$$(2,\ 1)\,,\ (3,\ 1)\,,\ (3,\ 2)\,,\ (3,\ 3)\,,$$

其中 $(i,\ j)$ 表示 i 行 j 列交叉点位置.

设游戏者 1 随后的每次填入策略为: 避免产生一个全 0 行或全 0 列, 我们将说明此时游戏者 0 一定可以保证出现一个 2×2 的元均为 0 的子式即可.

下列图形列出了 1 和 0 的每次步骤, 游戏者 1 填充步骤依 $1,2,3,\cdots$ 的次序进行, 游戏者 0 的步骤则按加圆括号的号码进行:

1	3	
2	(1)	(4)
4	(2)	(3)

1	4	
3	(1)	(2)
2	(3)	(4)

1	(4)	(3)
3	(1)	(2)
	2	4

1	(3)	(4)
3	(1)	(2)
	4	2

例 6.5 已知直角坐标平面上三点 $A(a_1,a_2)$, $B(b_1,b_2)$, $O(0,0)$. 求以 OA,OB 为一组邻边的平行四边形 $OACB$ 的面积 S_{OACB}. 作为结论的推广, 设 $A(a_1,\cdots,a_n)$, $B(b_1,\cdots,b_n)$, $O(0,\cdots,0)$ 是 $\mathbf{R}^n$ 中的三点. 求以 OA,OB 为一组邻边的平行四边形 $OACB$ 的面积 S_{OACB}.

解 将 A,B,O 作为 $\mathbf{R}^3$ 中的点, 即 $A(a_1,a_2,0)$, $B(b_1,b_2,0)$, $O(0,0,0)$. 则

$$S_{OACB}=\left\|\overrightarrow{OA}\times\overrightarrow{OB}\right\|=\left.\begin{vmatrix}\vec{i}&\vec{j}&\vec{k}\\a_1&a_2&0\\b_1&b_2&0\end{vmatrix}\right|_{\text{取模}}=\left|\det\begin{pmatrix}a_1&a_2\\b_1&b_2\end{pmatrix}\right|.$$

在 $\mathbf{R}^n$ 中, 以平面 OAB 上建立以 O 为原点的平面直角坐标系. 设基向量为 $\boldsymbol{x}=(x_1,\cdots,x_n)^{\mathrm{T}}$ 和 $\boldsymbol{y}=(y_1,\cdots,y_n)^{\mathrm{T}}$, 在基 $\{\boldsymbol{x},\boldsymbol{y}\}$ 下

$$\boldsymbol{A}=(\boldsymbol{x},\boldsymbol{y})\begin{pmatrix}u_1\\u_2\end{pmatrix},\ \boldsymbol{B}=(\boldsymbol{x},\boldsymbol{y})\begin{pmatrix}v_1\\v_2\end{pmatrix}.$$

于是

$$\begin{aligned}S_{OACB}&=\left|\det\begin{pmatrix}u_1&v_1\\u_2&v_2\end{pmatrix}\right|\\&=\sqrt{\det\begin{pmatrix}u_1&u_2\\v_1&v_2\end{pmatrix}\begin{pmatrix}\boldsymbol{x}^{\mathrm{T}}\\\boldsymbol{y}^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}&\boldsymbol{y}\end{pmatrix}\begin{pmatrix}u_1&v_1\\u_2&v_2\end{pmatrix}}\\&=\sqrt{\det\begin{pmatrix}a_1&\cdots&a_n\\b_1&\cdots&b_n\end{pmatrix}\begin{pmatrix}a_1&b_1\\\vdots&\vdots\\a_n&b_n\end{pmatrix}}.\end{aligned}$$

例 6.6 已知空间中平面 π 过三点 $A(a_1,a_2,a_3)$, $B(b_1,b_2,b_3)$, $O(0,0,0)$, 求空间中任意 $P(x,y,z)$ 关于 π 的对称点 $P'(x',y',z')$ 的坐标.

解 设 P 在 π 上的投影点 Q, 令 $\overrightarrow{OQ}=\lambda\overrightarrow{OA}+\mu\overrightarrow{OB}$, 则有

$$\overrightarrow{PQ}\cdot\overrightarrow{OA}=\overrightarrow{PQ}\cdot\overrightarrow{OB}=0.$$

上式可化为线性方程组

$$\begin{pmatrix}\overrightarrow{OA}\cdot\overrightarrow{OA}&\overrightarrow{OB}\cdot\overrightarrow{OA}\\\overrightarrow{OA}\cdot\overrightarrow{OB}&\overrightarrow{OB}\cdot\overrightarrow{OB}\end{pmatrix}\begin{pmatrix}\lambda\\\mu\end{pmatrix}=\begin{pmatrix}\overrightarrow{OP}\cdot\overrightarrow{OA}\\\overrightarrow{OP}\cdot\overrightarrow{OB}\end{pmatrix},$$

解得

$$\overrightarrow{OQ}=(\overrightarrow{OA},\overrightarrow{OB})\begin{pmatrix}\overrightarrow{OA}\cdot\overrightarrow{OA}&\overrightarrow{OB}\cdot\overrightarrow{OA}\\\overrightarrow{OA}\cdot\overrightarrow{OB}&\overrightarrow{OB}\cdot\overrightarrow{OB}\end{pmatrix}^{-1}\begin{pmatrix}\overrightarrow{OP}\cdot\overrightarrow{OA}\\\overrightarrow{OP}\cdot\overrightarrow{OB}\end{pmatrix}.$$

于是 $\overrightarrow{OP'} = 2\overrightarrow{OQ} - \overrightarrow{OP}$, 或矩阵形式

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \left(2\boldsymbol{M}(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M})^{-1}\boldsymbol{M}^{\mathrm{T}} - \boldsymbol{I}_3\right)\begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

其中 $\boldsymbol{M} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}$, $\boldsymbol{I}_3$ 是 3 阶单位矩阵.

例 6.7 设三种食物每 100 g 中蛋白质、碳水化合物和脂肪的含量如下表. 如果这三种食物作为每天的主要食物, 那么它们的用量应各取多少才能实现表中给出的减肥营养要求?

20 世纪 80 年代美国流行的剑桥大学医学院的简捷营养处方

营养	每 100 g 食物所含营养/g			减肥所要求的每日营养量
	脱脂牛奶	大豆面粉	乳清	
蛋白质	36	51	13	33
碳水化合物	52	34	74	45
脂肪	0	7	1.1	3

解 设每天脱脂牛奶、大豆面粉、乳清的使用量分别为 x_1, x_2, x_3 个单位 (100 g), 由题意可得方程组

$$\begin{cases} 36x_1 + 51x_2 + 13x_3 = 33, \\ 52x_1 + 34x_2 + 74x_3 = 45, \\ \qquad\qquad 7x_2 + 1.1x_3 = 3. \end{cases}$$

解方程组可得 $x_1 = 0.277\,2$, $x_2 = 0.391\,9$, $x_3 = 0.233\,2$. 即脱脂牛奶用量约为 27.7 g, 大豆面粉用量约为 39.2 g, 乳清用量约为 23.3 g.

例 6.8 考察氨水氧化为二氧化碳的化学反应, 反应式为

$$4\mathrm{NH_3}+5\mathrm{O_2} = 4\mathrm{NO} + 6\mathrm{H_2O},$$

$$4\mathrm{NH_3}+3\mathrm{O_2} = 2\mathrm{N_2}+6\mathrm{H_2O},$$

$$4\mathrm{NH_3}+6\mathrm{NO} = 5\mathrm{N_2}+6\mathrm{H_2O},$$

$$2\mathrm{NO} + \mathrm{O_2} = 2\mathrm{NO_2},$$

$$2\mathrm{NO} = \mathrm{N_2}+\mathrm{O_2},$$

$$\mathrm{N_2}+2\mathrm{O_2} = 2\mathrm{NO_2},$$

求出表述此系统所需的最少独立化学反应式.

解 所给化学反应式给出了各元素与化合物之间的线性方程组, 写成矩阵形式为

$$\begin{pmatrix} 4 & -4 & 0 & -6 & 5 & 0 \\ 4 & 0 & 0 & -6 & 3 & -2 \\ 4 & 6 & 0 & -6 & 0 & -5 \\ 0 & 2 & -2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & -1 & -1 \\ 0 & 0 & -2 & 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{NH_3} \\ \mathrm{NO} \\ \mathrm{NO_2} \\ \mathrm{H_2O} \\ \mathrm{O_2} \\ \mathrm{N_2} \end{pmatrix} = 0.$$

将系数矩阵作初等行变换

$$\begin{pmatrix} 4 & -4 & 0 & -6 & 5 & 0 \\ 4 & 0 & 0 & -6 & 3 & -2 \\ 4 & 6 & 0 & -6 & 0 & -5 \\ 0 & 2 & -2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & -1 & -1 \\ 0 & 0 & -2 & 0 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 4 & -4 & 0 & -6 & 5 & 0 \\ 0 & 4 & 0 & 0 & -2 & -2 \\ 0 & 10 & 0 & 0 & -5 & -5 \\ 0 & 2 & -2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & -1 & -1 \\ 0 & 0 & -2 & 0 & 2 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 4 & -4 & 0 & -6 & 5 & 0 \\ 0 & 2 & 0 & 0 & -1 & -1 \\ 0 & 0 & -2 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

系数矩阵的秩为 3. 所以此系统所需的最少独立化学反应式仅有 3 个, 这 3 个可取原反应式中的第 1, 第 5, 第 6 个.

例 6.9 设 100 只昆虫分布在有四个格子的密闭盒子内, 格子间有如图 6.9.1 所示通道. 每个格子每分钟有 60% 的昆虫通过通道离开原来的格子, 均匀地进入和它相连的格子. 若 1 min 后四个格子昆虫数量分别为 12, 25, 26 和 37, 试问初始状态四个格子各有多少只昆虫?

解 设初始状态四个格子的昆虫数分别为 x_1, x_2, x_3 和 x_4, 根据题意可以

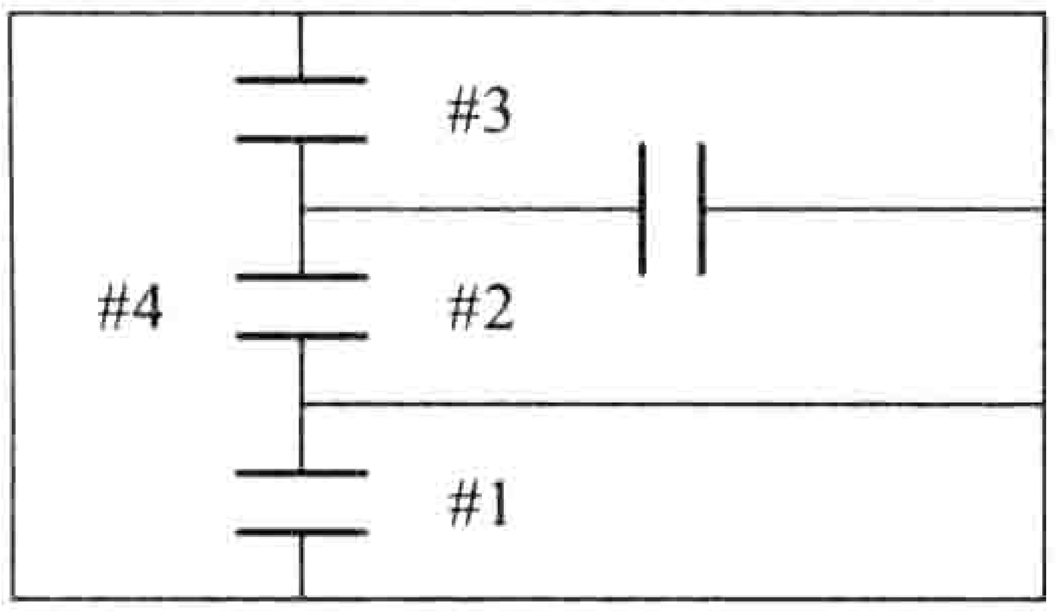


图 6.9.1

建立如下方程组

$$\begin{cases}\dfrac{0.6x_4}{3}+0.4x_1=12,\\[2mm]\dfrac{0.6x_4}{3}+\dfrac{0.6x_3}{2}+0.4x_2=25,\\[2mm]\dfrac{0.6x_4}{3}+\dfrac{0.6x_2}{2}+0.4x_3=26,\\[2mm]0.6x_1+\dfrac{0.6x_2}{2}+\dfrac{0.6x_3}{2}+0.4x_4=37.\end{cases}$$

易判定方程组的系数行列式不为零, 方程组有唯一解, 求得其解为

$$x_1=10,\ x_2=20,\ x_3=30,\ x_4=40.$$

例 6.10　一个家具厂生产桌子、椅子和沙发, 该厂一个月可用的原料包括 550 单位木材, 475 单位劳力及 222 单位纺织品. 家具厂要为每月用完这些原料制订生产计划表, 不同产品所需原料数量如下表:

	桌子	椅子	沙发
木材	4	2	5
劳力	3	2	5
纺织品	0	2	4

试确定:

(1) 每种产品应生产出多少个?

(2) 若纺织品的数量增加 10 个单位, 所生产沙发的数量改变多少?

解　(1) 设桌子、椅子和沙发的数量分别为 x_1,x_2,x_3, 由题意得

$$\begin{cases}4x_1+2x_2+5x_3=550,\\3x_1+2x_2+5x_3=475,\\\qquad\quad 2x_2+4x_3=222.\end{cases}$$

记 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$, $\boldsymbol{A}=\begin{pmatrix}4&2&5\\3&2&5\\0&2&4\end{pmatrix}$, $\boldsymbol{b}=\begin{pmatrix}550\\475\\222\end{pmatrix}$, 算出 $\boldsymbol{A}^{-1}=\begin{pmatrix}1&-1&0\\6&-8&\frac{5}{2}\\-3&4&-1\end{pmatrix}$,

所以

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}=\begin{pmatrix}75\\55\\28\end{pmatrix}.$$

(2) 纺织品的数量增加 10 个单位, 使得 $\boldsymbol{b}$ 改变 $\Delta\boldsymbol{b}=\begin{pmatrix}0\\0\\10\end{pmatrix}$, 研究 $\Delta\boldsymbol{b}$ 对解的影响, 我们建立新的关系式

$$\boldsymbol{A}\boldsymbol{x}^*=\boldsymbol{b}+\Delta\boldsymbol{b},$$

则

$$\boldsymbol{x}^*=\boldsymbol{A}^{-1}(\boldsymbol{b}+\Delta\boldsymbol{b})=\boldsymbol{A}^{-1}\boldsymbol{b}+\boldsymbol{A}^{-1}\Delta\boldsymbol{b}=\boldsymbol{x}+\Delta\boldsymbol{x},$$

$$\Delta\boldsymbol{x}=\boldsymbol{x}^*-\boldsymbol{x}=\boldsymbol{A}^{-1}\Delta\boldsymbol{b}=\begin{pmatrix}0\\25\\-10\end{pmatrix}.$$

故所生产的沙发数量减少 10 个.

例 6.11 假设一个经济系统由煤炭、电力、钢铁行业组成, 每个行业的产出在各个行业中的分配如表 6.11.1 所示. 每一列中的元素表示占该行业总产出的比例. 以表 6.11.1 的第 2 列为例, 电力行业的总产出分配如下: 40% 分配到煤炭行业, 50% 分配到钢铁行业, 余下的 10% 分配到电力行业 (电力行业把这 10% 当作部门营运所需的投入). 把煤炭、电力、钢铁行业每年总产出的价格 (即货币价值) 分别用 p_{C}, p_{E} 和 p_{S} 表示. 试求使得每个行业的投入与产出都相等的平衡价格.

解 表 6.11.1 中, 按列可以看出每个行业的产出分配到何处, 按行则可以看出这个行业所需的投入. 为了使煤炭行业的收入 p_{C} 等于它的支出, 则有

$$p_{\mathrm{C}}=0.4p_{\mathrm{E}}+0.6p_{\mathrm{S}}. \tag{6.11.1}$$

电力行业的收支平衡条件为

$$p_E=0.6p_{\mathrm{C}}+0.1p_{\mathrm{E}}+0.2p_{\mathrm{S}}. \tag{6.11.2}$$

表 6.11.1　一个简单的经济系统

产出分配			购买者
煤炭	电力	钢铁	
0.0	0.4	0.6	煤炭
0.6	0.1	0.2	电力
0.4	0.5	0.2	钢铁

钢铁行业的收支平衡条件为

$$p_{\mathrm{S}} = 0.4p_{\mathrm{C}} + 0.5p_{\mathrm{E}} + 0.2p_{\mathrm{S}}. \tag{6.11.3}$$

解由方程 (6.11.1)、(6.11.2) 和 (6.11.3) 构成的方程组, 得通解为

$$p_{\mathrm{C}} = 0.94p_{\mathrm{S}}, \quad p_{\mathrm{E}} = 0.85p_{\mathrm{S}},$$

其中 p_{S} 为自由变量. 这个经济系统的平衡价格向量有如下形式

$$\boldsymbol{p} = \begin{pmatrix} p_{\mathrm{C}} \\ p_{\mathrm{E}} \\ p_{\mathrm{S}} \end{pmatrix} = \begin{pmatrix} 0.94p_{\mathrm{S}} \\ 0.85p_{\mathrm{S}} \\ p_{\mathrm{S}} \end{pmatrix} = p_{\mathrm{S}} \begin{pmatrix} 0.94 \\ 0.85 \\ 1 \end{pmatrix}.$$

每个 p_{S} 的 (非负) 取值都确定一个平衡价格的取值. 例如, 取 $p_{\mathrm{S}} = 1$, 则 $p_{\mathrm{C}} = 0.94$, $p_{\mathrm{E}} = 0.85$. 即如果钢铁行业产出价格为 1 亿元, 则煤炭行业产出价格为 0.94 亿元, 电力行业产出价格为 0.85 亿元, 那么每个行业的收入和支出相等.

例 6.12　化学方程式表示化学反应中消耗和产生的物质的量. 例如, 燃烧丙烷时丙烷 (C_3H_8) 和氧 (O_2) 结合, 生成二氧化碳 (CO_2) 和水 (H_2O), 其方程式为

$$(x_1)\mathrm{C_3H_8} + (x_2)\mathrm{O_2} \to (x_3)\mathrm{CO_2} + (x_4)\mathrm{H_2O}. \tag{6.12.1}$$

为了 "配平" 这个方程式, 必须找出对应的系数 $x_1, \cdots, x_4$, 使得方程式左端的碳原子 (C)、氢原子 (H) 和氧原子 (O) 的总数与右端对应的原子总数相等. 试给出配平 (6.12.1) 式的对应系数.

解　方程 (6.12.1) 包含了 3 种不同的原子 (C、H、O), 用 3 维向量来表示每个分子所包含的不同原子数, 即

$$\mathrm{C_3H_8} : \begin{pmatrix} 3 \\ 8 \\ 0 \end{pmatrix}, \quad \mathrm{O_2} : \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}, \quad \mathrm{CO_2} : \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \mathrm{H_2O} : \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \begin{array}{l} \leftarrow \mathrm{C} \\ \leftarrow \mathrm{H} \\ \leftarrow \mathrm{O} \end{array}.$$

为了配平方程式 (6.12.1), 系数 $x_1, \cdots, x_4$ 必须满足

$$x_1\begin{pmatrix}3\\8\\0\end{pmatrix}+x_2\begin{pmatrix}0\\0\\2\end{pmatrix}=x_3\begin{pmatrix}1\\0\\2\end{pmatrix}+x_4\begin{pmatrix}0\\2\\1\end{pmatrix}.$$

解上面线性方程组得

$$x_1=\frac{1}{4}x_4,\ \ x_2=\frac{5}{4}x_4,\ \ x_3=\frac{3}{4}x_4,\quad x_4\text{ 是自由变量.}$$

由于化学方程式中的系数必须为整数, 取 $x_4=4$, 此时 $x_1=1$, $x_2=5$ 且 $x_3=3$. 配平后的方程式为

$$C_3H_8+5O_2\to 3CO_2+4H_2O.$$

例 6.13　图 6.13.1 中的网络给出了在下午一两点钟, 某市区一些单行道的交通流量 (以每小时的汽车数量来度量). 试确定网络的流量模式.

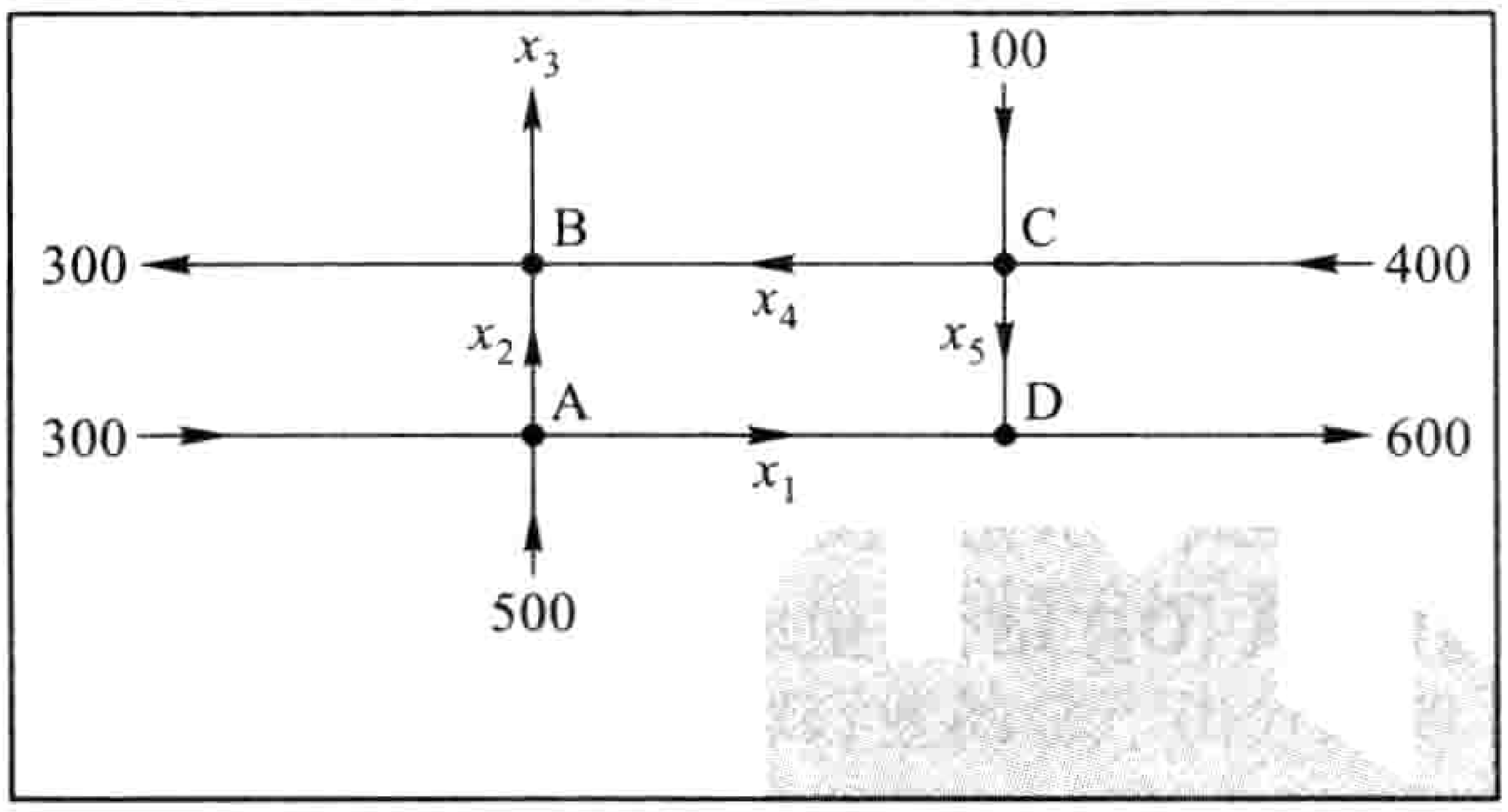


图 6.13.1　街道交通网络图

解　由所给交通网络图列表标出街道交叉点和分支中未知的流量如下表:

交叉点	注入	流出
A	$300+500$	x_1+x_2
B	x_2+x_4	$300+x_3$
C	$100+400$	x_4+x_5
D	x_1+x_5	600

由每个交叉点的流入量和流出量相等, 得方程组

$$\begin{cases} x_1 + x_2 & & & = 800, \\ & x_2 - x_3 + x_4 & & = 300, \\ & & x_4 + x_5 & = 500, \\ x_1 & & + x_5 & = 600. \end{cases}$$

此外, 该网络的总流入量 $(500+300+100+400)$ 等于网络的总流出量 $(300+x_3+600)$, 化简得 $x_3=400$. 代入方程组求解得

$$\begin{cases} x_1 = 600 - x_5, \\ x_2 = 200 + x_5, \\ x_3 = 400, \\ x_4 = 500 - x_5, \end{cases}$$

x_5 是自由变量. 由于街道是单行道, 因此变量不能取负值 (负流量表示与模型中指定的方向相反). 所以上面通解中的自由变量 $0 \leqslant x_5 \leqslant 500$.

例 6.14 确定图 6.14.1 电网中的回路电流 I_1, I_2, I_3.

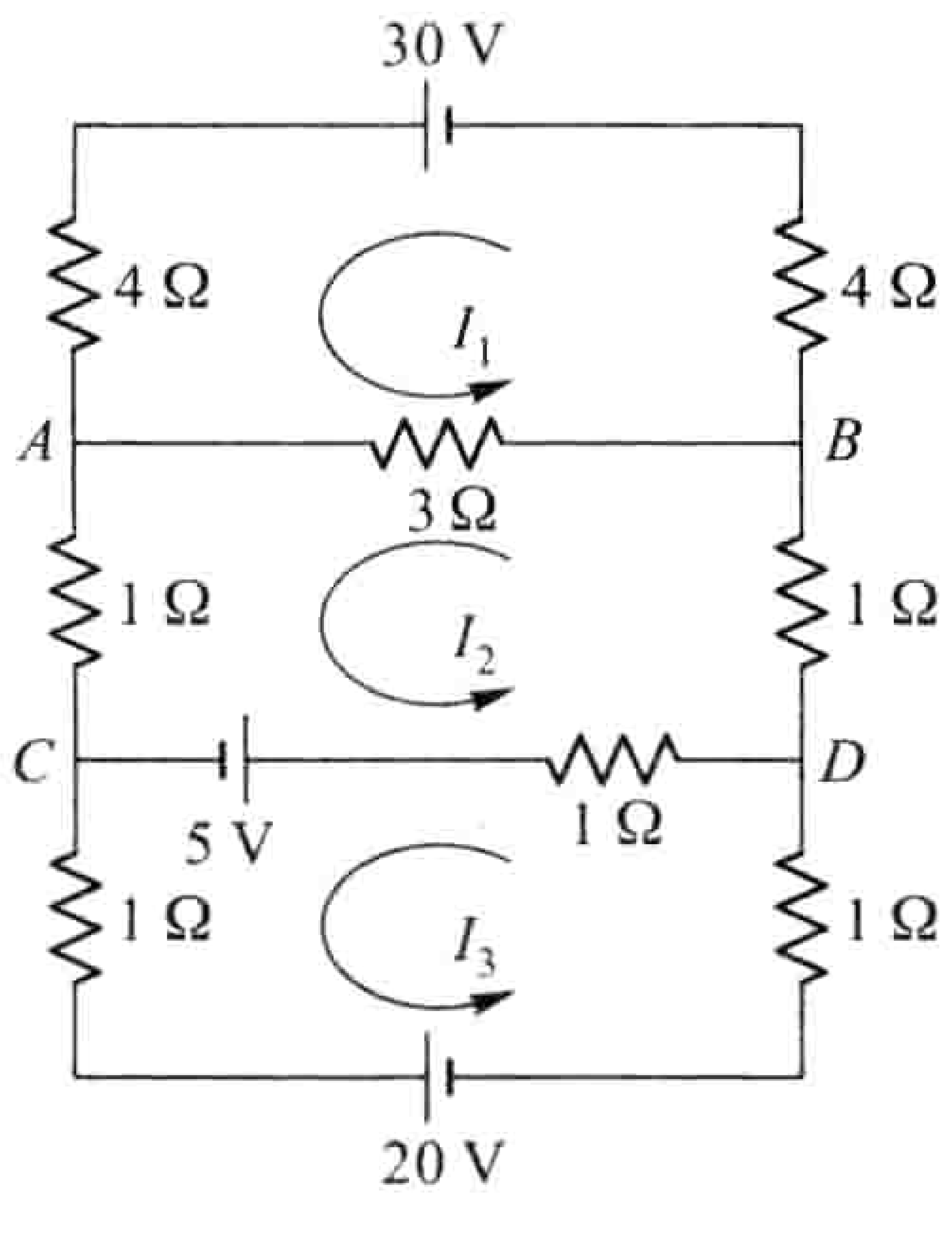


图 6.14.1 电网图

解 回路 1 中, 电流 I_1 流过三个电阻, 且电压降 RI 为

$$4I_1 + 4I_1 + 3I_1 = 11I_1.$$

回路 2 中的电流也流经回路 1 的一部分, 即从 A 到 B 的分支, 对应的电压降 RI 为 $3I_2$ V. 然而, 回路 1 中电流在 AB 段的方向与回路 2 中选定的方向相反, 因此

回路 1 中所有电压降 R_1 的代数和为 $11I_1-3I_2$. 由于回路 1 中的电压为 +30 V, Kirchhoff (基尔霍夫, 1824—1887) 电压定律表明

$$11I_1-3I_2=30.$$

类似可得回路 2 与回路 3 的方程分别为

$$-3I_1+6I_2-I_3=5,$$
$$-I_2+3I_3=-25.$$

解上面三个方程所确定的方程组得: $I_1=3$ A, $I_2=1$ A, $I_3=-8$ A.

I_3 取负值说明回路 3 中的实际电流与图 6.14.1 中显示的电流方向相反.

例 6.15 一个木工, 一个电工, 一个油漆工, 三人相互同意彼此装修他们自己的房子. 在装修之前, 他们达成了如下协议: (1) 每人总共工作十天 (包括给自己家干活在内); (2) 每人的日工资根据一般的市价在 120 元 ~ 150 元之间; (3) 每人的日工资数应使得每人的总收入与总支出相等. 下面的表是他们协商后制定出的工作天数的分配方案:

	木工	电工	油漆工
在木工家的工作天数	2	1	6
在电工家的工作天数	4	5	1
在油漆工家的工作天数	4	4	3

问三人的日工资各为多少?

解 根据协议中每人总支出与总收入相等的原则, 分别考虑木工、电工及油漆工的总收入和总支出. 设木工、电工、油漆工的日工资分别为 x_1,x_2,x_3, 而木工、电工及油漆工三人在木工家工作的天数分别为 2 天, 1 天, 6 天, 按日工资累计木工的总支出为 $2x_1+x_2+6x_3$, 于是木工的收支平衡可描述为等式

$$2x_1+x_2+6x_3=10x_1.$$

同理, 可建立描述电工、油漆工各自的收支平衡关系的另外两个等式, 将三个等式联立, 可得描述实际问题的方程组.

$$\begin{cases}2x_1+x_2+6x_3=10x_1,\\4x_1+5x_2+x_3=10x_2,\\4x_1+4x_2+3x_3=10x_3,\end{cases}$$

整理, 得

$$\begin{cases} -8x_1 + x_2 + 6x_3 = 0, \\ 4x_1 - 5x_2 + x_3 = 0, \\ 4x_1 + 4x_2 - 7x_3 = 0. \end{cases}$$

将其系数矩阵作初等行变换化为行最简形矩阵

$$\boldsymbol{A} = \begin{pmatrix} -8 & 1 & 6 \\ 4 & -5 & 1 \\ 4 & 4 & -7 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -\dfrac{31}{36} \\ 0 & 1 & -\dfrac{8}{9} \\ 0 & 0 & 0 \end{pmatrix}.$$

得方程组的通解

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = k \begin{pmatrix} \dfrac{31}{36} \\ \dfrac{8}{9} \\ 1 \end{pmatrix} \quad (k \text{ 为任意实数}).$$

由于 $120 \leqslant x_1, x_2, x_3 \leqslant 150$, 为了使日工资为整数值可取 $k = 144$, 满足题意的解为 $x_1 = 124$, $x_2 = 128$, $x_3 = 144$.

例 6.16 某中药厂用 9 种中草药 A—I, 根据不同的比例配制成了 7 种特效药, 各用量成分见表 6.16.1 (单位: g).

表 6.16.1

中药	1 号中药	2 号中药	3 号中药	4 号中药	5 号中药	6 号中药	7 号中药
A	10	2	14	12	20	38	100
B	12	0	12	25	35	60	55
C	5	3	11	0	5	14	0
D	7	9	25	5	15	47	35
E	0	1	2	25	5	33	6
F	25	5	35	5	35	55	50
G	9	4	17	25	2	39	25
H	6	5	16	10	10	35	10
I	8	2	12	0	0	6	20

试解答:

(1) 某医院要购买这 7 种特效药, 但药厂的第 3 号药和第 6 号药已经卖完, 请问能否用其他特效药配制出这两种脱销的药品;

(2) 现在该医院想用这 7 种草药配制三种新的特效药, 表 6.16.2 给出了三种新的特效药的成分, 请问能否配制? 如何配制?

表 6.16.2

中药	1 号新药	2 号新药	3 号新药
A	40	162	88
B	62	141	67
C	14	27	8
D	44	102	51
E	53	60	7
F	50	155	80
G	71	118	38
H	41	68	21
I	14	52	30

解 (1) 把每一种特效药看成一个 9 维列向量, 记第 i 号中药为 $\boldsymbol{\alpha}_i$ $(i=1,2,\cdots,7)$, 考察这 7 个列向量构成向量组的线性相关性. 若向量组线性无关, 则无法配制脱销的特效药; 若向量组线性相关, 且 $\boldsymbol{\alpha}_k$ 能由其余的向量线性表示, 则可以配制第 k 号药品.

由题设

$$
\begin{aligned}
\boldsymbol{\alpha}_1 &= (10,\ 12,\ 5,\ 7,\ 0,\ 25,\ 9,\ 6,\ 8)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_2 &= (2,\ 0,\ 3,\ 9,\ 1,\ 5,\ 4,\ 5,\ 2)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_3 &= (14,\ 12,\ 11,\ 25,\ 2,\ 35,\ 17,\ 16,\ 12)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_4 &= (12,\ 25,\ 0,\ 5,\ 25,\ 5,\ 25,\ 10,\ 0)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_5 &= (20,\ 35,\ 5,\ 15,\ 5,\ 35,\ 2,\ 10,\ 0)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_6 &= (38,\ 60,\ 14,\ 47,\ 33,\ 55,\ 39,\ 35,\ 6)^{\mathrm{T}}, \\
\boldsymbol{\alpha}_7 &= (100,\ 55,\ 0,\ 35,\ 6,\ 50,\ 25,\ 10,\ 20)^{\mathrm{T}}.
\end{aligned}
$$

以上面 7 个列向量作矩阵 $\boldsymbol{A}$, 并作初等行变换化为阶梯形, 有

$$\boldsymbol{A}=\begin{pmatrix}10&2&14&12&20&38&100\\12&0&12&25&35&60&55\\5&3&11&0&5&14&0\\7&9&25&5&15&47&35\\0&1&2&25&5&33&6\\25&5&35&5&35&55&50\\9&4&17&25&2&39&25\\6&5&16&10&10&35&10\\8&2&12&0&0&6&20\end{pmatrix}\rightarrow\begin{pmatrix}1&0&1&0&0&0&0\\0&1&2&0&0&3&0\\0&0&0&1&0&1&0\\0&0&0&0&1&1&0\\0&0&0&0&0&0&1\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\\0&0&0&0&0&0&0\end{pmatrix}$$

知 $\operatorname{rank}(\boldsymbol{A})=5$, 向量组的最大无关组为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7$, 余下的 $\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_6$ 可由最大无关组线性表示, 故可以配制 3 号和 6 号药.

(2) 三种新药用向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 表示, 问题化为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 能否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7$ 线性表示, 若能表示, 则可配制; 否则, 不能配制.

记 $\boldsymbol{B}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)$, 作初等行变换化为阶梯形, 有

$$\begin{array}{c} \begin{array}{cccccccc}\boldsymbol{\alpha}_1&\boldsymbol{\alpha}_2&\boldsymbol{\alpha}_4&\boldsymbol{\alpha}_5&\boldsymbol{\alpha}_7&\boldsymbol{\beta}_1&\boldsymbol{\beta}_2&\boldsymbol{\beta}_3\end{array}\\ \boldsymbol{B}\rightarrow\begin{pmatrix}1&0&0&0&0&1&3&0\\0&1&0&0&0&3&4&0\\0&0&1&0&0&2&2&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&1&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix}\end{array}$$

可以看出 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_4$, $\boldsymbol{\beta}_2=3\boldsymbol{\alpha}_1+4\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_4+\boldsymbol{\alpha}_7$, 这说明 1, 2 号新药可以配制, 其表达式给出了配置方法; 但 $\boldsymbol{\beta}_3$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5,\boldsymbol{\alpha}_7$ 线性表出, 即 3 号新药不能配制.

例 6.17 设有 $n+1$ 个人以及供他们读的 n 种小册子, 假定每人都读了一些小册子 (至少一本), 试证: 这 $n+1$ 个人中必可找到甲、乙两组人读过的小册子种类相同 (即, 将甲组人中每人读过的小册子合在一起, 其种类与乙组人每人读过的小册子合在一起的种类相同).

证 把 n 种小册子编号, 记

$$a_{ij}=\begin{cases}1, \text{第 } i \text{ 人读过第 } j \text{ 号小册子},\\ 0, \text{第 } i \text{ 人没读过第 } j \text{ 号小册子}\end{cases}(i=1,2,\cdots,n+1;\ j=1,2,\cdots,n),$$

则 $\boldsymbol{\alpha}_i=(a_{i1},a_{i2},\cdots,a_{in})$ 反映了第 i 人读过 n 种小册子的情况. 因为 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n+1}$ 一定线性相关, 所以存在不全为 0 的数 $k_i\ (i=1,2,\cdots,n+1)$ 使

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_{n+1}\boldsymbol{\alpha}_{n+1}=\mathbf{0}.$$

由于 $\boldsymbol{\alpha}_i$ 的分量非负, 所以 k_i 中必有正有负. 去掉那些等于 0 的 k_i, 并把所有负的 k_i 项移到等式右端, 记为

$$l_1\boldsymbol{\alpha}_{i_1}+l_2\boldsymbol{\alpha}_{i_2}+\cdots+l_t\boldsymbol{\alpha}_{i_t}=m_1\boldsymbol{\alpha}_{j_1}+m_2\boldsymbol{\alpha}_{j_2}+\cdots+m_s\boldsymbol{\alpha}_{j_s},$$

其中 l_i, m_j 均为正数, 且 $\{i_1,\cdots,i_t,j_1,\cdots,j_s\}\subseteq\{1,2,\cdots,n+1\}$.

将第 i_1 人, 第 i_2 人, $\cdots$, 第 i_t 人记为甲组; 第 j_1 人, $\cdots$, 第 j_s 人记为乙组. 则这两组人读过的小册子种类相同, 这是因为

$$l_1\boldsymbol{\alpha}_{i_1}+l_2\boldsymbol{\alpha}_{i_2}+\cdots+l_t\boldsymbol{\alpha}_{i_t}=m_1\boldsymbol{\alpha}_{j_1}+m_2\boldsymbol{\alpha}_{j_2}+\cdots+m_s\boldsymbol{\alpha}_{j_s}\triangleq(x_1,x_2,\cdots,x_n),$$

当 $x_i>0$ 时, 说明这两组人都读过第 i 种小册子, 当 $x_i=0$ 时, 这两组人都没读过这种小册子.

例 6.18 矩阵的相似对角化可帮助我们求矩阵的方幂运算. 矩阵的方幂运算在很多线性递归问题中都会涉及. 例如: Fibonacci (斐波那契, 约 1170—约 1250) 数列

$$F_0=1,\quad F_1=1,\quad F_{n+1}=F_n+F_{n-1}\ (n\geqslant 1).$$

(1) 试求二阶方阵 $\boldsymbol{A}$, 使得 $\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}F_n\\F_{n-1}\end{pmatrix}$;

(2) 根据 (1) 中结果得到 Fibonacci 数列通项公式;

(3) 根据以上思想, 试求解满足

$$\begin{cases}a_n=2a_{n-1}-5b_{n-1},\\ b_n=\dfrac{1}{3}a_{n-1}-\dfrac{2}{3}b_{n-1},\end{cases}$$

且 $a_0=b_0=1$ 的数列 $\{a_n\}$ 和 $\{b_n\}$ 的通项公式.

解 (1) 注意到

$$\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}F_n\\F_{n-1}\end{pmatrix}$$

由 $F_{n+1}=F_n+F_{n-1}\ (n\geqslant 1)$, 可得

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

(2) 由 (1) 中的结论作递推, 并注意到 $F_0=1$, $F_1=1$ 得

$$\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}F_n\\F_{n-1}\end{pmatrix}=\boldsymbol{A}^2\begin{pmatrix}F_{n-1}\\F_{n-2}\end{pmatrix}=\cdots=\boldsymbol{A}^n\begin{pmatrix}F_1\\F_0\end{pmatrix}=\boldsymbol{A}^n\begin{pmatrix}1\\1\end{pmatrix}.\quad(6.18.1)$$

为计算 $\boldsymbol{A}^n$, 将 $\boldsymbol{A}$ 相似对角化. $\boldsymbol{A}$ 的特征多项式为

$$|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-1\\-1&\lambda\end{vmatrix}=\lambda^2-\lambda-1,\ \text{特征值为}\ \lambda_1=\frac{1-\sqrt{5}}{2},\ \lambda_2=\frac{1+\sqrt{5}}{2}.$$

$(\lambda_1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_1=\left(1,-\dfrac{\sqrt{5}+1}{2}\right)^{\mathrm{T}}$,

$(\lambda_2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_2=\left(1,\dfrac{\sqrt{5}-1}{2}\right)^{\mathrm{T}}$.

令 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$, 则

$$\boldsymbol{P}^{-1}=\frac{1}{\sqrt{5}}\begin{pmatrix}\dfrac{\sqrt{5}-1}{2}&-1\\\dfrac{\sqrt{5}+1}{2}&1\end{pmatrix},\quad\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\dfrac{1-\sqrt{5}}{2}&0\\0&\dfrac{1+\sqrt{5}}{2}\end{pmatrix},$$

$$\boldsymbol{A}^n=\frac{1}{\sqrt{5}}\begin{pmatrix}1&1\\-\dfrac{\sqrt{5}+1}{2}&\dfrac{\sqrt{5}-1}{2}\end{pmatrix}\begin{pmatrix}\dfrac{1-\sqrt{5}}{2}&0\\0&\dfrac{1+\sqrt{5}}{2}\end{pmatrix}^n\begin{pmatrix}\dfrac{\sqrt{5}-1}{2}&-1\\\dfrac{\sqrt{5}+1}{2}&1\end{pmatrix}.$$

利用 (6.18.1) 式可得到

$$F_n=\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n-\left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(3) 首先用矩阵表示通项, 递推关系如下:

$$\begin{cases}a_n=2a_{n-1}-5b_{n-1},\\b_n=\dfrac{1}{3}a_{n-1}-\dfrac{2}{3}b_{n-1},\end{cases}$$

即

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix}^2 \begin{pmatrix} a_{n-2} \\ b_{n-2} \end{pmatrix} = \cdots$$
$$= \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix}^n \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}.$$

令

$$\boldsymbol{A} = \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix},$$

对角化可得

$$\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \dfrac{1}{3} & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{P}^{-1},$$

其中 $\boldsymbol{P} = \begin{pmatrix} 3 & 5 \\ 1 & 1 \end{pmatrix}$. 从而

$$\boldsymbol{A}^n = \boldsymbol{P} \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix}^n \boldsymbol{P}^{-1} = \frac{1}{2} \begin{pmatrix} 5 - 3 \times \left(\dfrac{1}{3}\right)^n & -15 + 15 \times \left(\dfrac{1}{3}\right)^n \\ 1 - \left(\dfrac{1}{3}\right)^n & -3 + 5 \times \left(\dfrac{1}{3}\right)^n \end{pmatrix}.$$

于是

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ \dfrac{1}{3} & -\dfrac{2}{3} \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 + 6 \times \left(\dfrac{1}{3}\right)^n \\ -1 + 2 \times \left(\dfrac{1}{3}\right)^n \end{pmatrix},$$

因此

$$a_n = -5 + 6 \times \left(\frac{1}{3}\right)^n, \quad b_n = -1 + 2 \times \left(\frac{1}{3}\right)^n.$$

例 6.19 某君举步上高楼, 每跨一次或上一个台阶或上两个台阶, 若要上 n 个台阶, 问有多少种不同的方式?

解 设登上 n 个台阶的不同方式数为 F_n.

显然有 $F_1 = 1$ (即登上一个台阶只有一种方式), $F_2 = 2$ (即登上两个台阶只有两种方式); 在登上 n 个台阶的方式中, 跨一步只有两种可能:

(1) 第一步跨一个台阶, 后面登 $n-1$ 个台阶的方式有 F_{n-1} 个;

(2) 第一步跨两个台阶, 后面登 $n-2$ 个台阶的方式有 F_{n-2} 个.

所以

$$F_n = F_{n-1} + F_{n-2}\ (n = 3, 4, 5, \cdots).$$

这里再定义 $F_0 = 1$, 记 $\boldsymbol{\alpha}_0 = \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\boldsymbol{\alpha}_n = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}$, 与上例相同的方法可得

$$\boldsymbol{\alpha}_n = \boldsymbol{A}^n \boldsymbol{\alpha}_0\ (n = 1, 2, 3, \cdots),$$

其中 $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. 由此可求得上台阶的方式数 F_n 为

$$1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34, \cdots.$$

利用上例 (2) 中的结论有

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right].$$

若某君住在二楼 (共有 18 个台阶), 则上楼的方式共有 $F_{18} = 4\,181$ 种. 因此, 如果某君每天以一种方式上楼, 那么他可以在 11.45 年内, 每天以不同的方式上楼.

例 6.20 出租汽车问题. 某一大型城市的出租汽车公司为了方便司机在城东和城西设了 A, B 两个营业部. 如果周一 A 营业部有 120 辆出租汽车, 而 B 营业部有 150 辆. 统计数据表明, 平均每天 A 营业部汽车的 10% 被开到 B 营业部, B 营业部汽车的 12% 被开到了 A 营业部. 假设所有汽车正常, 试寻找一种方案使每天汽车正常流动而 A 营业部和 B 营业部的汽车数量不增不减.

解 设第 n 天 A 营业部的汽车数为 $x_1^{(n)}$, B 营业部的汽车数为 $x_2^{(n)}$, 则有

$$\begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \end{pmatrix} = \begin{pmatrix} 0.9 & 0.12 \\ 0.1 & 0.88 \end{pmatrix} \begin{pmatrix} x_1^{(n-1)} \\ x_2^{(n-1)} \end{pmatrix}.$$

求得状态转换矩阵 $\boldsymbol{L} = \begin{pmatrix} 0.9 & 0.12 \\ 0.1 & 0.88 \end{pmatrix}$ 的两个特征值

$$\lambda_1 = 1, \quad \lambda_2 = 0.78.$$

求得 $\lambda_1 = 1$ 所对应的单位特征向量

$$\boldsymbol{\alpha} = (0.768\,2,\ 0.640\,2)^{\mathrm{T}}.$$

由于 $\boldsymbol{L\alpha}=\boldsymbol{\alpha}$, 所以当 A 营业部和 B 营业部的汽车数量按 $\boldsymbol{\alpha}$ 比例重新分配, 就能保持每天汽车正常流动而 A 营业部和 B 营业部的汽车数量保持一种"稳定状态".

即 A 营业部的汽车数为

$$(120+150)\frac{0.768\,2}{0.768\,2+0.640\,2}\approx 147\ (\text{辆}),$$

B 营业部的汽车数为

$$(120+150)\frac{0.640\,2}{0.768\,2+0.640\,2}\approx 123\ (\text{辆}).$$

例 6.21 在一个核反应堆中有 α 和 β 两种粒子, 在每秒里 1 个 α 粒子分裂为 3 个 β 粒子, 1 个 β 粒子分裂为 2 个 β 粒子和 1 个 α 粒子. 若在时刻 $t=0$, 反应堆中只有 1 个 α 粒子, 那么在 $t=100$ s 时反应堆中共有多少个粒子?

解 设在 $t=n$ s 时, 反应堆中共有 a_n 个 α 粒子和 b_n 个 β 粒子, 由题设可列出下面的 $\{a_n\}$, $\{b_n\}$ 定解问题

$$\begin{cases} a_n=b_{n-1}, \\ b_n=3a_{n-1}+2b_{n-1}, \\ a_0=1, b_0=0. \end{cases}$$

由上可得

$$a_n=b_{n-1}=3a_{n-2}+2b_{n-2}=2a_{n-1}+3a_{n-2},\ a_0=1,\ a_1=0.$$

可用矩阵表示为

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}=\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix}=\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}^{n-1}\begin{pmatrix} a_1 \\ a_0 \end{pmatrix}=\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}^{n-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{6.21.1}$$

$\boldsymbol{A}=\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$ 的特征多项式为

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

特征值为 $\lambda_1=3,\ \lambda_2=-1$.

$(3\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_1=(3,1)^{\mathrm{T}}$;

$(-\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_2=(-1,1)^{\mathrm{T}}$.

令 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$, 则 $\boldsymbol{P}^{-1}=\dfrac{1}{4}\begin{pmatrix}1&1\\-1&3\end{pmatrix}$, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}3&0\\0&-1\end{pmatrix}$.

$$\boldsymbol{A}^{n-1}=\boldsymbol{P}\begin{pmatrix}3&0\\0&-1\end{pmatrix}^{n-1}\boldsymbol{P}^{-1}=\frac{1}{4}\begin{pmatrix}3^n-(-1)^n&3^n+3(-1)^n\\3^{n-1}-(-1)^{n-1}&3^{n-1}+3(-1)^{n-1}\end{pmatrix}.$$

代入上面 (6.21.1) 式, 得

$$a_n=\frac{1}{4}\cdot 3^n+\frac{3}{4}\cdot(-1)^n,$$

从而

$$b_n=a_{n+1}=\frac{1}{4}\cdot 3^{n+1}+\frac{3}{4}\cdot(-1)^{n+1}.$$

所以在第 n s 时, 反应堆的总粒子数为

$$a_n+b_n=\frac{1}{4}\cdot 3^n+\frac{3}{4}\cdot(-1)^n+\frac{1}{4}\cdot 3^{n+1}+\frac{3}{4}\cdot(-1)^{n+1}=3^n.$$

在 $t=100$ s 时反应堆中共有粒子数 3^{100} 个.

例 6.22 伴性基因是一种位于 X 染色体上的基因. 例如, 红绿色盲基因 (简称色盲基因) 是一种隐性的伴性基因. 为给出一个描述给定的人群中色盲的数学模型, 需要将人群分为两类 —— 男性和女性. 令 $x_1^{(0)}$,$x_2^{(0)}$ 分别为男性与女性中有色盲基因的比例, 由于男性从母亲处获得一个 X 染色体, 且不从父亲处获得 X 染色体, 所以下一代的男性中色盲的比例 $x_1^{(1)}$ 将和上一代的女性中含有隐性色盲基因的比例相同. 由于女性从双亲处分别得到一个 X 染色体, 所以下一代女性中含有隐形基因的比例 $x_2^{(1)}$ 将为 $x_1^{(0)}$ 和 $x_2^{(0)}$ 的平均值, 写出第 $n+1$ 代男性和女性中色盲的比例, 并分析变化趋势.

解 第一代与第二代间的变化关系可写为如下形式:

$$\begin{pmatrix}x_1^{(1)}\\x_2^{(1)}\end{pmatrix}=\begin{pmatrix}0&1\\\dfrac{1}{2}&\dfrac{1}{2}\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}.$$

令 $\boldsymbol{x}^{(n)}=(x_1^{(n)},x_2^{(n)})^{\mathrm{T}}$ 表示第 $n+1$ 代男性和女性中色盲比例, $\boldsymbol{A}=\begin{pmatrix}0&1\\\dfrac{1}{2}&\dfrac{1}{2}\end{pmatrix}$,

则

$$\boldsymbol{x}^{(n)}=\boldsymbol{A}^n\boldsymbol{x}^{(0)}.$$

又矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-2\\1&1\end{pmatrix}\begin{pmatrix}1&0\\0&-\dfrac{1}{2}\end{pmatrix}\begin{pmatrix}\dfrac{1}{3}&\dfrac{2}{3}\\-\dfrac{1}{3}&\dfrac{1}{3}\end{pmatrix}$, 故

$$\begin{aligned}\boldsymbol{x}^{(n)}&=\begin{pmatrix}1&-2\\1&1\end{pmatrix}\begin{pmatrix}1&0\\0&-\dfrac{1}{2}\end{pmatrix}^n\begin{pmatrix}\dfrac{1}{3}&\dfrac{2}{3}\\-\dfrac{1}{3}&\dfrac{1}{3}\end{pmatrix}\boldsymbol{x}^{(0)}\\&=\frac{1}{3}\begin{pmatrix}1-\left(-\dfrac{1}{2}\right)^{n-1}&2+\left(-\dfrac{1}{2}\right)^{n-1}\\1-\left(-\dfrac{1}{2}\right)^{n}&2+\left(-\dfrac{1}{2}\right)^{n}\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}.\end{aligned}$$

于是

$$\lim_{n\to\infty}\boldsymbol{x}^{(n)}=\frac{1}{3}\begin{pmatrix}1&2\\1&2\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\end{pmatrix}=\begin{pmatrix}\dfrac{x_1^{(0)}+2x_2^{(0)}}{3}\\\dfrac{x_1^{(0)}+2x_2^{(0)}}{3}\end{pmatrix}.$$

当代数增加时, 男性和女性中含色盲基因的比例将趋于相同的数值.

注释 据医学研究, 男性只要有红绿色盲基因就表现为红绿色盲. 而女性只有隐性纯合子时才表现为红绿色盲, 而携带有红绿色盲基因时女性仍表现为正常. 因而在临床表现上, 往往男性色盲人数高于女性.

例 6.23 一种昆虫按年龄分为三个组, 第一组为幼虫 (不产卵), 第二组每个成虫在两周内平均产卵 100 个, 第三组每个成虫在两周内平均产卵 150 个. 假设现有三个组的昆虫各 100 只, 每个卵的成活率为 0.09, 第一组和第二组的昆虫能顺利进入下一个成虫组的存活率分别为 0.1 和 0.2.

(1) 以两周为一时间段, 分析这种昆虫各周龄组数目演变趋势, 昆虫数目是无限增长还是趋于灭亡?

(2) 一种除虫剂可以控制昆虫的数目, 使得各组昆虫的成活率减半, 问这种除虫剂是否有效?

解 开始时刻 $\boldsymbol{x}^{(0)}=(100,100,100)^{\mathrm{T}}$, 两周后昆虫的数目向量

$$\begin{pmatrix}x_1^{(1)}\\x_2^{(1)}\\x_3^{(1)}\end{pmatrix}=\begin{pmatrix}0&9&13.5\\0.1&0&0\\0&0.2&0\end{pmatrix}\begin{pmatrix}x_1^{(0)}\\x_2^{(0)}\\x_3^{(0)}\end{pmatrix}.$$

若记第 k 个阶段的数目向量为 $\boldsymbol{x}^{(k)}=\begin{pmatrix}x_1^{(k)}\\x_2^{(k)}\\x_3^{(k)}\end{pmatrix}$, $\boldsymbol{L}=\begin{pmatrix}0&9&13.5\\0.1&0&0\\0&0.2&0\end{pmatrix}$, 类似地

可以得到第 k 个阶段与第 $k+1$ 个阶段昆虫数目间的关系

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{L}\boldsymbol{x}^{(k)},$$

从而第 n 个阶段与初始阶段昆虫数目间的关系

$$\boldsymbol{x}^{(n)} = \boldsymbol{L}^n\boldsymbol{x}^{(0)}.$$

(1) 求得 $\boldsymbol{L}$ 的特征值为

$$\lambda_1 = \frac{587}{547}, \quad \lambda_2 = -\frac{1\,420}{1\,953}, \quad \lambda_3 = -\frac{1\,171}{3\,384},$$

且由 $|\lambda_1| > |\lambda_2| > |\lambda_3|$ 可知特征值互不相等, 因而它们对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性无关, 故 $\boldsymbol{L}$ 可相似对角化, 且 $\mathbf{R}^3$ 中任意向量都可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性表出. 设

$$\boldsymbol{x}^{(0)} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + c_3\boldsymbol{\alpha}_3,$$

则有

$$\begin{aligned}\boldsymbol{x}^{(n)} &= \boldsymbol{L}^n\boldsymbol{x}^{(0)} = \boldsymbol{L}^n(c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + c_3\boldsymbol{\alpha}_3)\\ &= c_1\lambda_1^n\boldsymbol{\alpha}_1 + c_2\lambda_2^n\boldsymbol{\alpha}_2 + c_3\lambda_3^n\boldsymbol{\alpha}_3\\ &= \lambda_1^n\left[c_1\boldsymbol{\alpha}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^n\boldsymbol{\alpha}_2 + c_3\left(\frac{\lambda_3}{\lambda_1}\right)^n\boldsymbol{\alpha}_3\right].\end{aligned}$$

那么

$$\lim_{n\to\infty}\boldsymbol{x}^{(n)} = \lim_{n\to\infty}\lambda_1^n c_1\boldsymbol{\alpha}_1.$$

由于 $|\lambda_1| > 1$, 故昆虫数目无限增长.

(2) 若使用除虫剂, 使得各组昆虫的成活率减半, 则此时矩阵

$$\boldsymbol{L} = \frac{1}{2}\begin{pmatrix} 0 & 9 & 13.5\\ 0.1 & 0 & 0\\ 0 & 0.2 & 0\end{pmatrix},$$

其特征值分别为

$$\lambda_1 = \frac{587}{1\,094}, \quad \lambda_2 = -\frac{710}{1\,953}, \quad \lambda_3 = -\frac{1\,171}{6\,768}.$$

由于 $|\lambda_1| < 1$, 故

$$\lim_{n\to\infty}\boldsymbol{x}^{(n)} = \lim_{n\to\infty}\lambda_1^n c_1\boldsymbol{\alpha}_1 = 0,$$

即昆虫数目趋于消亡, 这种除虫剂是有效的.

例 6.24 图 6.24.1 中的电路可以由下面的微分方程刻画

$$\begin{pmatrix} v_1'(t) \\ v_2'(t) \end{pmatrix} = \begin{pmatrix} -\dfrac{\dfrac{1}{R_1}+\dfrac{1}{R_2}}{C_1} & \dfrac{1}{R_2C_1} \\ \dfrac{1}{R_2C_2} & -\dfrac{1}{R_2C_2} \end{pmatrix} \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix},$$

其中 $v_1(t)$ 和 $v_2(t)$ 分别为两个电容器在时刻 t 的电压. 设电阻 R_1 等于 $1\,\Omega$, R_2 等于 $2\,\Omega$, 电容 C_1 等于 1 F, C_2 等于 0.5 F, 并且设电容器 C_1 和 C_2 的初始电压分别是 5 V 和 4 V. 试求 $v_1(t)$ 和 $v_2(t)$ 的公式, 并描述电压是如何随着时间变化的.

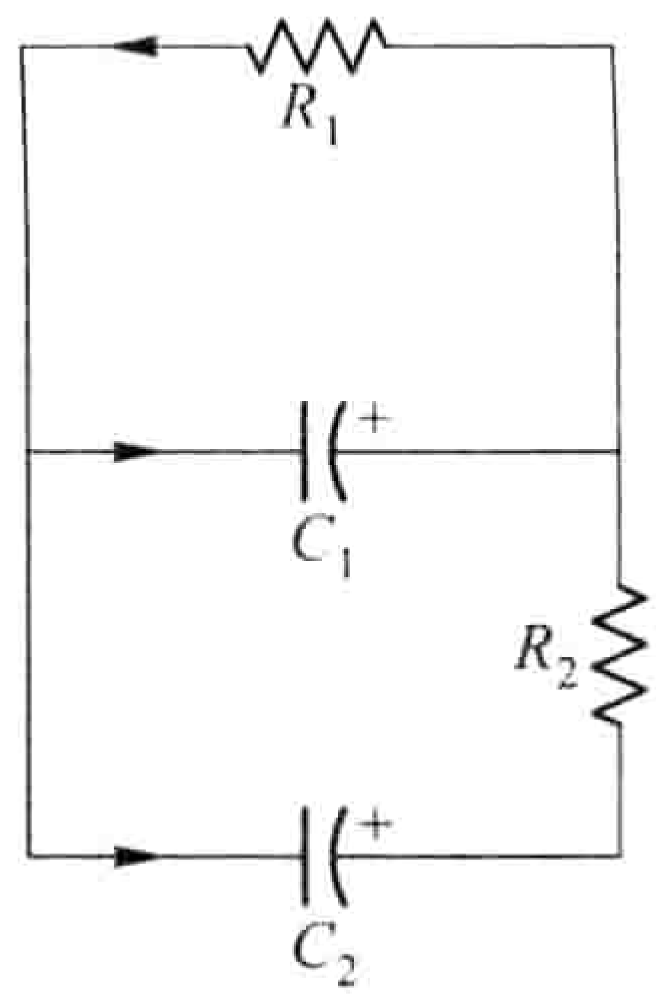


图 6.24.1 电路图

解 根据已知数据, 令 $\boldsymbol{A} = \begin{pmatrix} -1.5 & 0.5 \\ 1 & -1 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\boldsymbol{x}_0 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$. 向量 $\boldsymbol{x}_0$ 给出了 $\boldsymbol{x}$ 的初始值. 解特征方程

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = 0,$$

得到 $\boldsymbol{A}$ 的特征值 $\lambda_1 = -0.5$ 和 $\lambda_2 = -2$, 对应的特征向量为

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \text{ 和 } \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

特征函数 $\boldsymbol{x}_1(t) = \boldsymbol{v}_1 \mathrm{e}^{\lambda_1 t}$, $\boldsymbol{x}_2(t) = \boldsymbol{v}_2 \mathrm{e}^{\lambda_2 t}$, 得微分方程 $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ 的通解

$$\boldsymbol{x}(t) = c_1 \boldsymbol{v}_1 \mathrm{e}^{\lambda_1 t} + c_2 \boldsymbol{v}_2 \mathrm{e}^{\lambda_2 t} = c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mathrm{e}^{-0.5t} + c_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \mathrm{e}^{-2t}.$$

由初值条件 $\boldsymbol{x}(0)=\boldsymbol{x}_0=\begin{pmatrix}5\\4\end{pmatrix}$, 得 $c_1=3$, $c_2=-2$, 所以

$$\boldsymbol{x}(t)=3\begin{pmatrix}1\\2\end{pmatrix}\mathrm{e}^{-0.5t}-2\begin{pmatrix}-1\\1\end{pmatrix}\mathrm{e}^{-2t}$$

或

$$\begin{pmatrix}v_1(t)\\v_2(t)\end{pmatrix}=\begin{pmatrix}3\mathrm{e}^{-0.5t}+2\mathrm{e}^{-2t}\\6\mathrm{e}^{-0.5t}-2\mathrm{e}^{-2t}\end{pmatrix}.$$

当 $t\to\infty$ 时, 函数 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 都衰减至零, 但因为 $\boldsymbol{x}_2$ 的负指数较小, 所以它的衰退速度较快. 对应的特征向量 $\boldsymbol{v}_2$ 中的元说明, 如果初始电压绝对值相等但符号相反, 则电容器所带的电压将迅速地衰减至零.

例 6.25 一只虫子在平面直角坐标系内爬行 (图 6.25.1), 开始时位于点 $P_0(1,0)$ 处. 如果知道虫子在点 $P(x,\ y)$ 处沿 x 轴正向的速率为 $4x-5y$, 沿 y 轴正向的速率为 $2x-3y$. 如何确定虫子爬行的轨迹的参数方程?

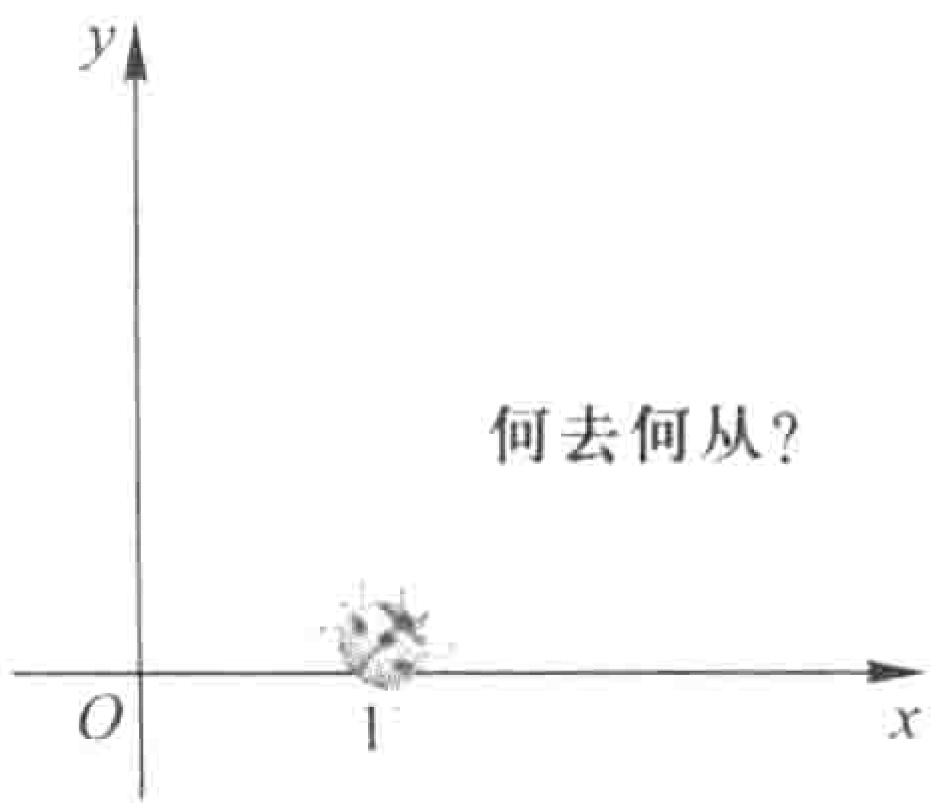


图 6.25.1 虫子爬行的轨迹图

解 设 t 时刻虫子所处位置的坐标为 $(x(t),y(t))$. 由已知条件和上述假设可知

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=4x-5y,\\[2mm]\dfrac{\mathrm{d}y}{\mathrm{d}t}=2x-3y,\end{cases}\quad 且\ (x(0),y(0))=(1,0).$$

现要由此得出虫子爬行的轨迹的参数方程.

令 $\boldsymbol{A}=\begin{pmatrix}4&-5\\2&-3\end{pmatrix}$, 则 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=\begin{vmatrix}\lambda-4&5\\-2&\lambda+3\end{vmatrix}=(\lambda+1)(\lambda-2)$.

$\boldsymbol{A}$ 的特征值为 $\lambda_1=-1$, $\lambda_2=2$.

$(-\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_1=(1,1)^{\mathrm{T}}$;

$(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的一个基础解系为 $\boldsymbol{\xi}_2=(5,2)^{\mathrm{T}}$.

令 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$, 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}-1&0\\0&2\end{pmatrix}$.

记 $\boldsymbol{X}=\begin{pmatrix}x\\y\end{pmatrix}$, $\boldsymbol{Y}=\begin{pmatrix}u\\v\end{pmatrix}$, 并且作线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$, 则 $\boldsymbol{Y}=\boldsymbol{P}^{-1}\boldsymbol{X}$.

$$\frac{\mathrm{d}\boldsymbol{Y}}{\mathrm{d}t}=\boldsymbol{P}^{-1}\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{X}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Y}=\begin{pmatrix}-1&0\\0&2\end{pmatrix}\boldsymbol{Y},$$

即

$$\begin{pmatrix}\dfrac{\mathrm{d}u}{\mathrm{d}t}\\\dfrac{\mathrm{d}v}{\mathrm{d}t}\end{pmatrix}=\begin{pmatrix}-1&0\\0&2\end{pmatrix}\begin{pmatrix}u\\v\end{pmatrix},$$

解得 $u=c_1\mathrm{e}^{-t}$, $v=c_2\mathrm{e}^{2t}$, 即 $\boldsymbol{Y}=\begin{pmatrix}c_1\mathrm{e}^{-t}\\c_2\mathrm{e}^{2t}\end{pmatrix}$. 因而

$$\begin{pmatrix}c_1\\c_2\end{pmatrix}=\boldsymbol{Y}\,|_{t=0}=\boldsymbol{P}^{-1}\boldsymbol{X}\,|_{t=0}=\begin{pmatrix}-\dfrac{2}{3}&\dfrac{5}{3}\\\dfrac{1}{3}&-\dfrac{1}{3}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}-\dfrac{2}{3}\\\dfrac{1}{3}\end{pmatrix}.$$

于是

$$\boldsymbol{Y}=\begin{pmatrix}-\dfrac{2}{3}\mathrm{e}^{-t}\\\dfrac{1}{3}\mathrm{e}^{2t}\end{pmatrix},\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}=\begin{pmatrix}1&5\\1&2\end{pmatrix}\begin{pmatrix}-\dfrac{2}{3}\mathrm{e}^{-t}\\\dfrac{1}{3}\mathrm{e}^{2t}\end{pmatrix}=\begin{pmatrix}-\dfrac{2}{3}\mathrm{e}^{-t}+\dfrac{5}{3}\mathrm{e}^{2t}\\-\dfrac{2}{3}\mathrm{e}^{-t}+\dfrac{2}{3}\mathrm{e}^{2t}\end{pmatrix}.$$

这就是说, 虫子爬行的轨迹的参数方程为 $\begin{cases}x=-\dfrac{2}{3}\mathrm{e}^{-t}+\dfrac{5}{3}\mathrm{e}^{2t},\\y=-\dfrac{2}{3}\mathrm{e}^{-t}+\dfrac{2}{3}\mathrm{e}^{2t}.\end{cases}$

例 6.26 设单位圆的坐标向量为 $\left\{\begin{pmatrix}x\\y\end{pmatrix}\middle|\,x^2+y^2=1\right\}$, $\boldsymbol{A}$ 为 2 阶正定矩阵, 且 $\begin{pmatrix}u\\v\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}x\\y\end{pmatrix}$, 则以 (u,v) 为坐标的点构成什么图形?

解 由于 $\boldsymbol{A}$ 为 2 阶正定矩阵, 则存在正交矩阵 $\boldsymbol{C}$ 使得

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix},$$

其中 $\lambda_i > 0\ (i = 1,2)$ 为 $\boldsymbol{A}$ 的特征值. 则

$$\boldsymbol{A}^{-1} = \boldsymbol{C}\begin{pmatrix} \dfrac{1}{\lambda_1} & 0 \\ 0 & \dfrac{1}{\lambda_2} \end{pmatrix}\boldsymbol{C}^{\mathrm{T}},$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{A}^{-1}\begin{pmatrix} u \\ v \end{pmatrix},$$

则方程 $x^2 + y^2 = 1$ 可化为

$$1 = (x,y)\begin{pmatrix} x \\ y \end{pmatrix} = (u,v)\left(\boldsymbol{A}^{-1}\right)^{\mathrm{T}}\boldsymbol{A}^{-1}\begin{pmatrix} u \\ v \end{pmatrix} = (u,v)\boldsymbol{C}\begin{pmatrix} \dfrac{1}{\lambda_1^2} & 0 \\ 0 & \dfrac{1}{\lambda_2^2} \end{pmatrix}\boldsymbol{C}^{\mathrm{T}}\begin{pmatrix} u \\ v \end{pmatrix}.$$

再作正交变换 (旋转) $\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \boldsymbol{C}^{\mathrm{T}}\begin{pmatrix} u \\ v \end{pmatrix}$, 方程化为

$$\frac{\tilde{u}^2}{\lambda_1^2} + \frac{\tilde{v}^2}{\lambda_2^2} = 1,$$

这是椭圆且椭圆的长半轴与短半轴分别为矩阵 $\boldsymbol{A}$ 的最大与最小特征值.

例 6.27 在下一年度, 某县政府计划用一笔资金修 x 百千米的公路, 修整 y 百平方千米的公园, 政府部门必须确定在两个项目上如何分配它的资金, 如果可能的话, 可以同时开始两个项目, 而不是仅开始一个项目. 假设 x 和 y 必须满足下面限制条件

$$16x^2 + 25y^2 \leqslant 400,$$

见图 6.27.1. 每个阴影可行集合的点 (x,y) 表示一个可能的年度工作计划. 为使资金效用达到最大, 经济学家常利用函数

$$q(x,y) = xy,$$

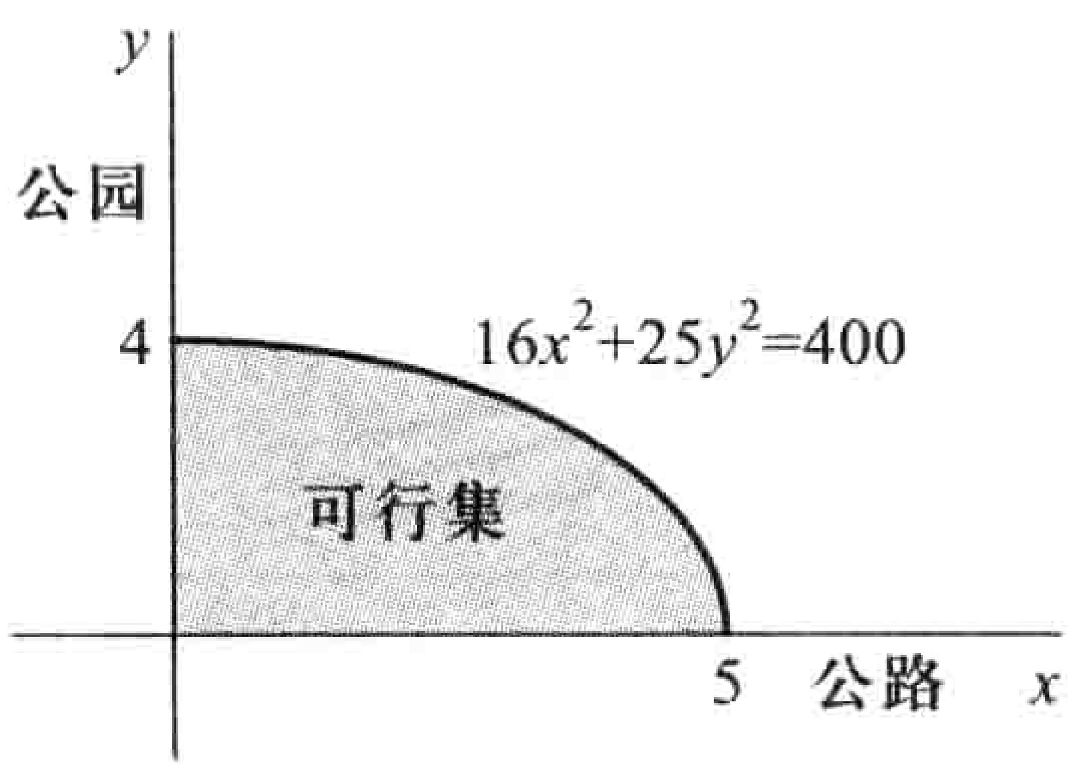


图 6.27.1

称之为效用函数, 曲线 $xy=c$ (c 为常数) 称之为无差异曲线, 因为在该曲线上的任意点的效用值相等. 现制定一个工作计划, 使得效用函数达到最大.

解 显然效用函数在可行集中的最大值只可能在边界曲线 $16x^2+25y^2=400$ 上取到, 作变量代换将该曲线化为单位圆. 因为

$$\left(\frac{x}{5}\right)^2+\left(\frac{y}{4}\right)^2=1.$$

令 $u=\frac{x}{5}$, $v=\frac{y}{4}$, 曲线化为 $u^2+v^2=1$. 效用函数变成

$$q(x,y)=q(5u,4v)=5u\cdot 4v=20uv.$$

令 $\boldsymbol{X}=\begin{pmatrix}u\\v\end{pmatrix}$, 则原问题变为在限制条件 $\|\boldsymbol{X}\|=1$ 下 $f(\boldsymbol{X})=20uv$ 的最大值. 我们知道这个最大值是二次型 $f(\boldsymbol{X})=20uv$ 的矩阵为 $\boldsymbol{A}=\begin{pmatrix}0&10\\10&0\end{pmatrix}$ 的最大特征值.

易求得 $\boldsymbol{A}$ 的特征值为 ± 10, 对应特征值 10 的单位特征向量为 $\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}$, 所以 $f(\boldsymbol{X})=20uv$ 的最大值为 10, 且在 $u=v=\frac{1}{\sqrt{2}}$ 处取得.

于是, 最优的工作计划是修建 $x=5u=\frac{5}{\sqrt{2}}\approx 3.5$ 百千米的公路, 修整 $y=4v=\frac{4}{\sqrt{2}}\approx 2.8$ 百平方千米的公园. 最优工作计划是限制曲线和无差异曲线的切点, 具有更大效用的点 (x,y) 位于和限制曲线不相交的无差异曲线上, 见图 6.27.2.

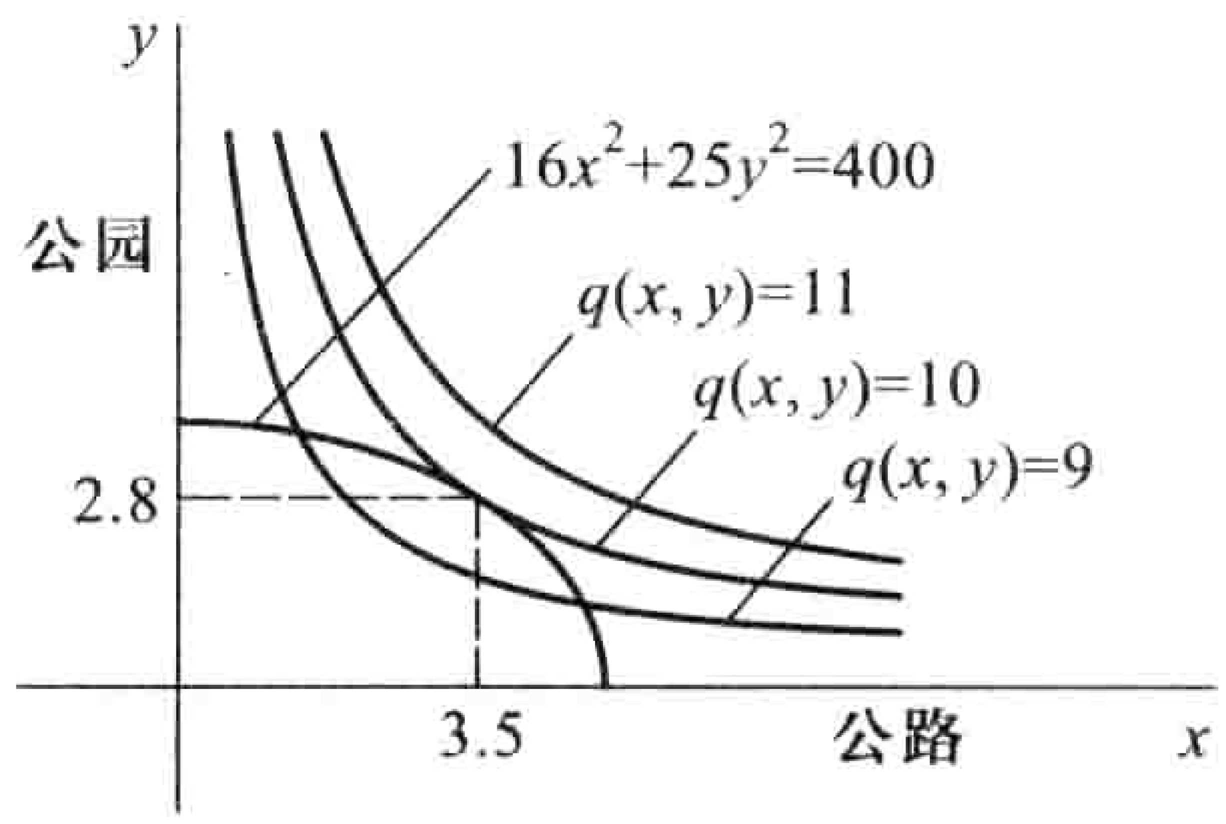


图 6.27.2

例 6.28 天文学家要确定一颗小行星绕太阳运行的轨道, 在轨道平面内建立以太阳为原点的直角坐标系, 在两坐标轴上取天文测量单位 (一天文单位为地球到太阳的平均距离约为 9300 万英里①). 在五个不同的时间对小行星作了观察, 测得轨道上五个点的坐标数据 (单位: 英里):

x	4.559 6	5.081 6	5.554 6	5.963 6	6.275 6
y	0.814 5	1.368 5	1.989 5	2.692 5	3.526 5

由 Kepler (开普勒, 1571—1630) 第一定律知, 小行星轨道为一椭圆, 设方程为

$$a_1x^2+2a_2xy+a_3y^2+2a_4x+2a_5y+1=0,$$

试确定椭圆的长短半轴.

解 记已知的 5 个数据点为 $(x_j,\ y_j)$ $(j=1,2,3,4,5)$, 将这 5 个数据点分别代入小行星轨道方程得关于 5 个系数 a_1, a_2, a_3, a_4, a_5 的线性方程组:

$$a_1x_j^2+2a_2x_jy_j+a_3y_j^2+2a_4x_j+2a_5y_j+1=0\ (j=1,2,3,4,5),$$

即

$$\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 \\ x_2^2 & 2x_2y_2 & y_2^2 & 2x_2 & 2y_2 \\ x_3^2 & 2x_3y_3 & y_3^2 & 2x_3 & 2y_3 \\ x_4^2 & 2x_4y_4 & y_4^2 & 2x_4 & 2y_4 \\ x_5^2 & 2x_5y_5 & y_5 & 2x_5 & 2y_5 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix}=\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

解该方程组得

$$a_1=-0.337\,8,\quad a_2=0.189\,2,\quad a_3=-0.381\,8,\quad a_4=0.460\,9,\quad a_5=0.410\,4,$$

即椭圆方程为

$$-0.337\,8x^2+0.378\,4xy-0.381\,8y^2+0.921\,8x+0.820\,8y+1=0.$$

为确定椭圆的长短半轴需将椭圆化为标准方程.

(1) 作坐标平移消去一次项

首先将二次曲线方程写成矩阵形式

$$(x\ ,y)\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}+2(x\ ,y)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}+1=0$$

① 1 英里 = 1609.344 千米.

或

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{c}+1=0,$$

其中

$$\boldsymbol{A}=\begin{pmatrix}-0.337\,8 & 0.189\,2\\ 0.189\,2 & -0.381\,8\end{pmatrix},\quad c=\begin{pmatrix}0.460\,9\\ 0.410\,4\end{pmatrix}.$$

作平移变换, $\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{z}$, 其中 $\boldsymbol{x}_0$ (椭圆中心) 待定, 代入方程得

$$(\boldsymbol{x}_0+\boldsymbol{z})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}_0+\boldsymbol{z})+2(\boldsymbol{x}_0+\boldsymbol{z})^{\mathrm{T}}\boldsymbol{c}+1=0,$$

即

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{z}+2\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{c})+(\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0+2\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{c}+1)=0.$$

记 $F=\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}_0+2\boldsymbol{x}_0^{\mathrm{T}}\boldsymbol{c}+1$, 为了消去方程中一次项, 令 $\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{c}=0$, 得

$$\boldsymbol{x}_0=(2.721\,3,2.423\,4)^{\mathrm{T}},\quad F=3.248\,8.$$

经坐标平移变换 $\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{z}$, 椭圆方程化简为

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{z}+F=0.$$

(2) 作正交 (旋转) 变换, 化椭圆为标准方程

令 $\boldsymbol{z}=\boldsymbol{U}\begin{pmatrix}u\\ v\end{pmatrix}$, 其中 $\boldsymbol{U}$ 是正交矩阵, 满足

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\ (\lambda_1,\lambda_2\text{ 是 }\boldsymbol{A}\text{ 的两个特征值}),$$

则椭圆化为标准方程

$$\lambda_1u^2+\lambda_2v^2+F=0.$$

解特征方程 $|\lambda\boldsymbol{I}-\boldsymbol{A}|=0$, 得特征值

$$\lambda_1=-0.550\,2,\quad \lambda_2=-0.169\,4,$$

由此可知椭圆的长短半轴分别为

$$a=\frac{1}{F\sqrt{-\lambda_1}}=4.379\,9,\quad b=\frac{1}{F\sqrt{-\lambda_2}}=2.429\,9.$$

例 6.29 若 $\boldsymbol{A}=\begin{pmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{pmatrix}$, 则线性变换 $\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 将 $\mathbf{R}^3$ 的单位球面 $\{\boldsymbol{x}:\|\boldsymbol{x}\|=1\}$ 映到 $\mathbf{R}^2$ 中的一个椭圆, 如图 6.29.1, 求最大化长度 $\|\boldsymbol{A}\boldsymbol{x}\|$ 的一个单位向量, 并且算出这个最大长度.

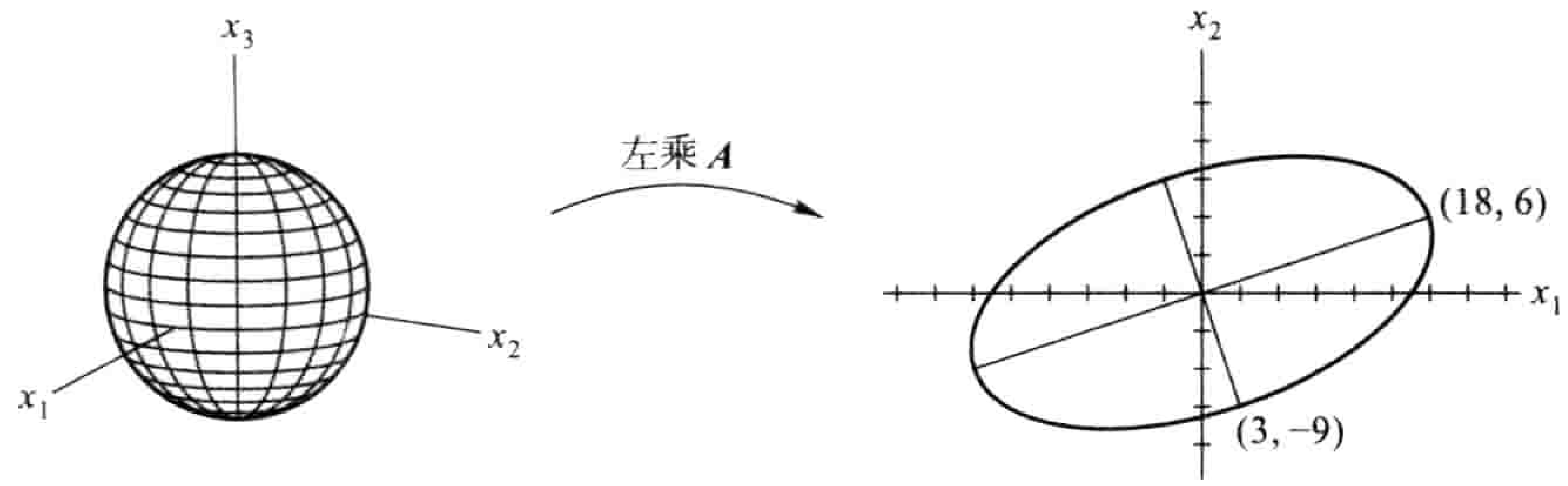


图 6.29.1 从 $\mathbf{R}^3$ 到 $\mathbf{R}^2$ 的变换

解 显然 $\|\boldsymbol{Ax}\|^2$ 与 $\|\boldsymbol{Ax}\|$ 在相同的 $\boldsymbol{x}$ 处取到最大值, 而

$$\|\boldsymbol{Ax}\|^2=(\boldsymbol{Ax})^{\mathrm{T}}(\boldsymbol{Ax})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x},$$

且由 $(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$, 知 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是对称矩阵.

记 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{B}$, 作正交变换 $\boldsymbol{x}=\boldsymbol{Py}$ 将二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Bx}$ 化为标准形

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Bx}=\lambda_1y_1^2+\lambda_2y_2^2+\lambda_3y_3^2\ (\lambda_1,\lambda_2,\lambda_3\ \text{是}\ \boldsymbol{B}\ \text{的特征值}).$$

设 $\lambda_1\geqslant\lambda_2\geqslant\lambda_3$, 则

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Bx}\leqslant\lambda_1(y_1^2+y_2^2+y_3^2)=\lambda_1\|\boldsymbol{y}\|^2.$$

由于

$$\|\boldsymbol{x}\|^2=(\boldsymbol{Py})^{\mathrm{T}}(\boldsymbol{Py})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{Py}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}=\|\boldsymbol{y}\|^2,$$

即 $\|\boldsymbol{x}\|=\|\boldsymbol{y}\|$, 所以当 $\|\boldsymbol{x}\|=1$ 时, 有 $\|\boldsymbol{y}\|=1$, 此时

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Bx}\leqslant\lambda_1,$$

所以 λ_1 就是二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Bx}$ 的最大值, 且这个值在 $\boldsymbol{x}$ 取 λ_1 对应的单位特征向量处被取到.

对本题中的矩阵 $\boldsymbol{A}$, 有

$$\boldsymbol{B}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}4&8\\11&7\\14&-2\end{pmatrix}\begin{pmatrix}4&11&14\\8&7&-2\end{pmatrix}=\begin{pmatrix}80&100&40\\100&170&140\\40&140&200\end{pmatrix}.$$

$\boldsymbol{B}$ 的特征值是 $\lambda_1 = 360$, $\lambda_2 = 90$ 及 $\lambda_3 = 0$, $\lambda_1 = 360$ 对应的单位特征向量为

$$\boldsymbol{v}_1 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix},$$

于是 $\|\boldsymbol{A}\boldsymbol{x}\|^2$ 的最大值为 360, 并且当 $\boldsymbol{x} = \boldsymbol{v}_1$ 时, 可以取到这个值. 向量 $\boldsymbol{A}\boldsymbol{v}_1$ 是图 6.29.1 中椭圆上与原点相距最远的点, 即

$$\boldsymbol{A}\boldsymbol{v}_1 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 18 \\ 6 \end{pmatrix}.$$

所以在约束条件 $\|\boldsymbol{x}\| = 1$ 下, $\|\boldsymbol{A}\boldsymbol{x}\|$ 的最大值是 $\|\boldsymbol{A}\boldsymbol{v}_1\| = \sqrt{360} = 6\sqrt{10}$.

例 6.30 n 维欧氏空间中二次曲面的一般形式是

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{x} + 2\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{x} + c = 0,$$

此处 $\boldsymbol{S}$ 是 n 阶实对称方阵, $\boldsymbol{\xi}$ 是常向量, c 是常实数. 求二次曲面在仿射变换

$$\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}, \quad \det(\boldsymbol{A}) \neq 0$$

下的分类.

解 首先, 将曲面方程写成矩阵形式

$$\begin{pmatrix} \boldsymbol{x}^{\mathrm{T}} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{S} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^{\mathrm{T}} & c \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix} = 0.$$

经过仿射变换 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\beta}$ 后曲面方程变为

$$\begin{pmatrix} \boldsymbol{y}^{\mathrm{T}} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{\beta} \\ \boldsymbol{0} & 1 \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \boldsymbol{S} & \boldsymbol{\xi} \\ \boldsymbol{\xi}^{\mathrm{T}} & c \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{\beta} \\ \boldsymbol{0} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix} = 0,$$

即

$$\begin{pmatrix} \boldsymbol{y}^{\mathrm{T}} & 1 \end{pmatrix} \begin{pmatrix} (\boldsymbol{A}^{-1})^{\mathrm{T}}\boldsymbol{S}\boldsymbol{A}^{-1} & (\boldsymbol{A}^{-1})^{\mathrm{T}}(\boldsymbol{\xi} - \boldsymbol{S}\boldsymbol{A}^{-1}\boldsymbol{\beta}) \\ (\boldsymbol{\xi} - \boldsymbol{S}\boldsymbol{A}^{-1}\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{A}^{-1} & c - \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{\beta} \end{pmatrix} \begin{pmatrix} \boldsymbol{y} \\ 1 \end{pmatrix} = 0.$$

由于 $\boldsymbol{G}=\begin{pmatrix}\boldsymbol{S} & \boldsymbol{\xi}\\ \boldsymbol{\xi}^{\mathrm{T}} & c\end{pmatrix}$ 和 $\widetilde{\boldsymbol{G}}=\begin{pmatrix}\widetilde{\boldsymbol{S}} & \widetilde{\boldsymbol{\xi}}\\ \widetilde{\boldsymbol{\xi}}^{\mathrm{T}} & \tilde{c}\end{pmatrix}$ 所代表的二次曲面仿射等价当且仅当 $\boldsymbol{G}$ 与 $\widetilde{\boldsymbol{G}}$ 合同且 $\boldsymbol{S}$ 与 $\widetilde{\boldsymbol{S}}$ 合同. 故所研究的二次曲面的仿射分类情况如下:

情形 1 $\operatorname{rank}(\boldsymbol{G})=\operatorname{rank}(\boldsymbol{S})$, 则存在仿射变换使得

$$\widetilde{\boldsymbol{G}}=\begin{pmatrix}\boldsymbol{I}_p & & \\ & -\boldsymbol{I}_q & \\ & & \boldsymbol{O}\end{pmatrix}.$$

此时, 曲面方程化为

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2=0.$$

(1) $pq\neq 0$, 曲面为锥面.

(2) $pq=0$ 且 $p+q<n$, 曲面为 $n-p-q$ 维平面.

(3) $pq=0$ 且 $p+q=n$, 曲面为点.

情形 2 $\operatorname{rank}(\boldsymbol{G})=1+\operatorname{rank}(\boldsymbol{S})$, 则存在仿射变换使得

$$\widetilde{\boldsymbol{G}}=\begin{pmatrix}\boldsymbol{I}_p & & & \\ & -\boldsymbol{I}_q & & \\ & & \boldsymbol{O} & \\ & & & \delta\end{pmatrix},\quad \delta=\pm 1.$$

此时, 曲面方程化为

$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2+\delta=0.$$

(4) $pq\neq 0$, 曲面为双曲面.

(5) $pq=0$ 且 $p+q<n$, 曲面为柱面或“虚柱面”.

(6) $pq=0$ 且 $p+q=n$, 曲面为球面或“虚球面”.

情形 3 $\operatorname{rank}(\boldsymbol{G})=2+\operatorname{rank}(\boldsymbol{S})$, 则存在仿射变换使得

$$\widetilde{\boldsymbol{G}}=\begin{pmatrix}\boldsymbol{I}_p & & & & \\ & -\boldsymbol{I}_q & & & \\ & & \boldsymbol{O} & & \\ & & & 0 & 1\\ & & & 1 & 0\end{pmatrix}.$$

此时, 曲线方程化为

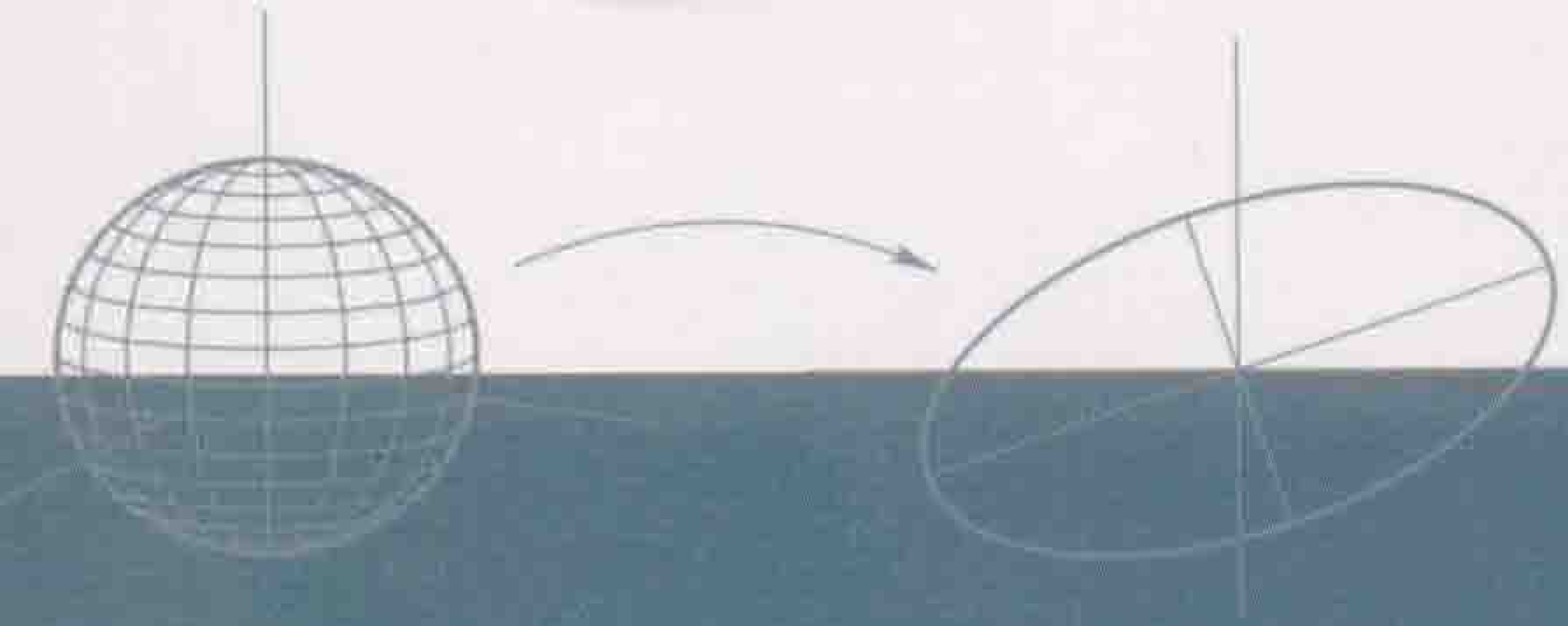
$$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2+2y_n=0.$$

(7) $pq \neq 0$ 且 $p+q<n-1$, 曲面为双曲抛物柱面.

(8) $pq \neq 0$ 且 $p+q=n-1$, 曲面为双曲抛物面.

(9) $pq=0$ 且 $p+q<n-1$, 曲面为椭圆抛物柱面.

(10) $pq=0$ 且 $p+q=n-1$, 曲面为椭圆抛物面.



ISBN 978-7-04-040392-3

9 787040 403923 >

定价 21.80 元