面向 21 世纪全国高职高专数学规划教材
* 江苏省省级精品课程配套教材

线性代数与概率统计

朱文辉　陈　刚　编著

北京大学出版社
PEKING UNIVERSITY PRESS

内 容 简 介

本书是江苏省一类优秀课程“线性代数与概率统计”的配套教材，该成果获2004年江苏省高等教育教学成果一等奖.本书贯彻“淡化严密性，强调思维性”的编写思路，使得必需够用为度和应用能力培养落到实处.

全书包括线性代数和概率统计方面的教学基本内容，并配有建模应用方面的资料，内容翔实，语言通俗，可读性强.每章后附有数量充足、难易适中的习题，书后附有答案.

本书可作为高职高专院校工程类与经管类专业相关课程的教材,也可供工程技术人员和高校学生业务性学习或扩充性学习的参考.

图书在版编目（CIP）数据

线性代数与概率统计/朱文辉，陈刚编著．—北京：北京大学出版社，2005.8
（面向21**世纪全国高职高专数学规划教材）**
ISBN 7-301-09220-2

I．线… II．①朱… ②陈… III．①线性代数—高等学校：技术学校—教材 ②概率论—高等学校：技术学校—教材 ③数理统计—高等学校：技术学校—教材 IV．①0151.2 ②021

中国版本图书馆CIP**数据核字（**2005**）第**069414**号**

书　　名：线性代数与概率统计
著作责任者：朱文辉　陈刚　编著
责 任 编 辑：胡伟晔　刘建龙
标 准 书 号：ISBN 7-301-09220-2/O·0656
出　版　者：北京大学出版社
地　　址：北京市海淀区成府路205**号**　100871
电　　话：邮购部62752015　**发行部**62750672　**编辑部**62765126
网　　址：http://cbs.pku.edu.cn
电 子 信 箱：xxjs@pup.pku.edu.cn
印　刷　者：
发　行　者：北京大学出版社
经　销　者：新华书店
787**毫米×**1092**毫米**　16**开本**　16.5**印张**　360**千字**
2005**年**8**月第**1**版**　2005**年**8**月第**1**次印刷**
定　　价：26.00**元**

前　言

本书是高职高专院校工程类与经管类专业的数学基础课教材之一。针对高职高专学生的特点，我们确立了适合于数学教育的“淡化严密性，强调思维性”的指导思想，能从观念上、方法上解决数学抽象性与人才培养对象之间的矛盾。“淡化严密性”就是以人为本，以高职学生的基础状况和实际需要为依据，调整和控制数学知识传授中的严密程度，降解抽象性，营造适度的数学环境，不追求每个细节的严格表述；“强调思维性”就是以能力为本，关注学生在掌握和应用数学知识时所需要的思维性，包括对专业技术的数学语言表达与交流、对实际问题的数学模型建立与运行。

本书从内容到形式，都力求贯彻“淡化严密性，强调思维性”的指导思想。在介绍各种概念时，均不以严格的“定义”形式出现，而是结合自然的叙述，辅以背景材料顺势引入，并注重从正反两方面对概念内涵加以充分的解读，首次出现的专用名词用黑体字明显标出，这种自然界面既便于查找，同时也减少了数学形式的抽象感；在介绍数学理论时，不拘泥于“定理—证明”的单一模式，也不是简单地删去证明了事，而是尽可能地在通俗易懂的叙述中渐入主题，既交代了来龙去脉，又冲淡了抽象成分，让读者有一种“水到渠成”之感，同时注重对抽象内容的形象化处理，大量地设计一些文字语言解读数学知识。这样的叙述方式，可使读者对数学教材不再望而生畏。

本书特别尊重读者的习惯性与情理性思维，介绍比较复杂的内容都从案例引入，并且注重典型例题的多层次开发，这样一方面可以改变数学教材严肃有余活泼不足的局面，另一方面能踏准读者的思维节奏，便于他们学习数学知识、理解数学原理。

作者运用教学研究成果，对传统的材料组织结构进行了较大幅度的改革，如线性代数中将线性方程组与向量的线性关系分离倒置，增加了初等变换的标准程序和基础解系的读取规则等操作性内容；弱化伴随矩阵在求逆运算中的作用；强化求特征向量的矩阵化技术，增加基础特征向量的新概念等。又如在概率统计中，凸现古典概率与统计定义各自的优势功能；概率的极限定理不再单独成章，而是结合数字特征，一改其高不可攀的理论形象，成了应用性内容；假设检验突破一题一招式的思路，以显著性原理为引导，在适当分解步骤的基础上，重点提示每个步骤的操作思路和执行要领，便于读者面对各种情况融会贯通；在方差分析和回归分析中废止列表计算模式，利用计算器中能直接显示的均值与方差来构造统计量，既避免了复杂的数学推导，能在情理之中掌握计算公式，又大大简化了计算，能在意料之外获得数值结果。本书增加了有关的数学建模内容，以便将数学建模思想有机地融入到基础性教学之中。

全书共分三篇。第一篇线性代数，包括行列式、矩阵、线性方程组、向量的线性关系、矩阵的特征值与特征向量、二次型等 6 章；第二篇概率统计，包括随机事件与概率、随机变量及其概率分布、随机变量的数字特征、统计量与参数估计、假设检验、方差分析与回归分析等 6 章；第三篇建模应用，包括线性代数和概率统计的应用共 2 章。各章虽有联系，但又能各自独立成篇，便于实施模块式教学构建。

随着教育大众化形势的发展，高职教育的专业设置多种多样，新专业不断涌现，对数学基础课提出了多元化、小型化、分散化的要求。本书的第 5 章、第 6 章、第 11 章、第 12 章等部分内容，就是为适应这些需求而特意安排的。

讲授本书的全部内容约需 60～100 学时。通常各专业要根据不同需求进行模块化搭建，因为本书已对各章节进行了独立化模型化处理，所以本教材实际上可适应 20～60 学时的工程数学课程教学需要。

本书文笔简洁明了，可读性强。“淡化严密性”带来了篇幅的减少以及信息量的增加；“强调思维性”使得理论高度不仅没有降低，相反在通俗易懂的铺叙中生动地展示了思维的全过程，突现了线性代数与概率统计的主流方法和操作技术，使这门基础课对于读者来说更易掌握，更为实用，同时也能使读者的思维在适度的数学环境中得到潜移默化的熏陶，成为培养思维能力的重要载体。

本书的编写特色得到了江苏省教育厅专家组的认可，“线性代数与概率统计”被评为江苏省一类优秀课程，同时，本书还是数学基础课教学改革成果的组成部分，该成果获 2004 年江苏省高等教育教学成果一等奖。

编　者

2005 年 6 月

目　　录

第一篇　线性代数 1

第 1 章　行列式 1

1.1　**行列式的概念与性质** 1

1.1.1　**二阶、三阶行列式** 1

1.1.2　***n* 阶行列式的全面展开** 2

1.1.3　**行列式的性质** 2

1.2　**行列式的降阶算法** 5

1.2.1　**代数余子式** 5

1.2.2　**特殊行列式的计算公式** 5

1.2.3　**行列式的降阶算法** 6

1.3　**克莱姆法则** 8

1.3.1　**行列式的按行（列）展开** 8

1.3.2　**代数余子式组合定理** 9

1.3.3　**克莱姆法则** 9

习题一 12

第 2 章　矩阵 15

2.1　**矩阵的概念及其线性运算** 15

2.1.1　**矩阵的概念** 15

2.1.2　**矩阵的加、减运算** 16

2.1.3　**矩阵的数乘** 16

2.2　**矩阵的乘法与转置** 17

2.2.1　**矩阵的乘法** 17

2.2.2　**矩阵乘法的性质** 18

2.2.3　**矩阵的转置** 20

2.2.4　**方阵行列式的乘积定理** 21

2.3　**逆矩阵** 21

2.3.1　**逆矩阵的概念** 21

2.3.2　**矩阵可逆的条件** 22

2.3.3 逆矩阵的性质 23
2.4 矩阵的初等变换 23
2.4.1 矩阵的初等行变换 23
2.4.2 初等变换的标准程序 24
2.4.3 用初等变换法求逆矩阵 25
2.4.4 初等矩阵 26
2.5 分块矩阵 27
2.5.1 分块矩阵的概念 27
2.5.2 分块矩阵的运算 28
习题二 29
第 3 章 线性方程组 32
3.1 线性方程组的矩阵消元解法 32
3.2 矩阵的秩 34
3.2.1 秩的概念 34
3.2.2 秩的求法 35
3.2.3 矩阵的秩与线性方程组的解 36
3.3 线性方程组解的结构 37
3.3.1 齐次线性方程组解的结构 37
3.3.2 非齐次线性方程组解的结构 39
3.4 矩阵方程的矩阵消元解法 42
习题三 44
第 4 章 向量的线性关系 48
4.1 向量的线性相关性 48
4.1.1 两种线性关系 48
4.1.2 线性关系和线性方程组 50
4.2 极大线性无关组与向量组的秩 52
4.2.1 极大线性无关组的概念 52
4.2.2 极大线性无关组的求法 52
4.2.3 向量组的秩 53
习题四 55
第 5 章 矩阵的特征值与特征向量 57
5.1 特征值与特征向量 57
5.2 矩阵的相似与矩阵的对角化 59
5.3 实对称矩阵的对角化 62
5.3.1 向量的内积与正交矩阵 62

5.3.2 实对称矩阵的特征值与特征向量 63
习题五 66
第 6 章 二次型 69
6.1 二次型及其标准形 69
6.1.1 二次方程与几何图形 69
6.1.2 二次型的矩阵表示 69
6.1.3 二次型的标准形 71
6.2 二次型的线性变换与惯性定理 72
6.2.1 二次型中的线性变换 72
6.2.2 化二次型为标准形的矩阵变换法 74
6.2.3 二次型的惯性定理 76
6.3 二次型的正交变换与有定性 78
6.3.1 用正交变换化二次型为标准形 78
6.3.2 二次型的有定性 79
习题六 81
第二篇 概率统计 84
第 7 章 随机事件及其概率 84
7.1 随机事件 84
7.1.1 随机试验与样本空间 84
7.1.2 随机事件与集合 85
7.1.3 事件的关系与运算 86
7.2 事件的概率 88
7.2.1 古典概率 88
7.2.2 概率的性质 88
7.2.3 古典概率的计算 89
7.2.4 概率的统计定义 91
7.3 事件的独立性 93
7.3.1 条件概率 93
7.3.2 乘法公式 93
7.3.3 事件的独立性 94
7.3.4 全概率公式 96
习题七 98
第 8 章 随机变量及其概率分布 101
8.1 离散型随机变量及其分布律 101

8.1.1 随机变量 101
8.1.2 离散型随机变量 101
8.1.3 两点分布 103
8.1.4 二项分布 103
8.1.5 泊松（Poisson）分布 105
8.2 连续型随机变量及其概率密度 106
8.2.1 连续型随机变量 106
8.2.2 均匀分布 109
8.2.3 指数分布 110
8.3 分布函数与函数的分布 110
8.3.1 分布函数 110
8.3.2 函数的分布 111
8.4 正态分布 112
8.4.1 正态分布的定义与性质 112
8.4.2 正态分布的概率计算 113
习题八 115
第 9 章 随机变量的数字特征 118
9.1 数学期望 118
9.1.1 数学期望的概念与计算公式 118
9.1.2 常用分布的数学期望 120
9.1.3 数学期望的运算规则 121
9.1.4 随机变量函数的数学期望 123
9.2 方差与标准差 124
9.2.1 方差的概念与计算公式 124
9.2.2 方差的运算规则 125
9.2.3 常用分布的方差 126
9.2.4 协方差与相关系数 128
9.3 大数定律与中心极限定理 129
习题九 132
第 10 章 统计量与参数估计 135
10.1 样本与统计量 135
10.1.1 总体与样本 135
10.1.2 统计量及其分布 136
10.1.3 临界值的概念及其概率意义 138
10.2 点估计 140

10.2.1 点估计的概念 140
10.2.2 点估计的方法 141
10.2.3 估计量的评价标准 142
10.3 区间估计 143
10.3.1 置信度与置信区间 143
10.3.2 正态总体的区间估计 143
10.3.3 置信度的选择 146
习题十 146
第 11 章 假设检验 150
11.1 单个正态总体的参数检验 150
11.1.1 假设检验的一般步骤 150
11.1.2 正态总体均值与方差的假设检验 152
11.1.3 显著性原理 154
11.2 两个正态总体的参数检验 156
11.2.1 两个样本的统计量及其分布 156
11.2.2 两个正态总体的均值与方差的假设检验 158
11.3 非参数检验 160
11.3.1 直方图法 160
11.3.2 皮尔逊检验 161
11.3.3 秩和检验 163
习题十一 164
第 12 章 方差分析与回归分析 168
12.1 方差分析 168
12.1.1 单因素方差分析 168
12.1.2 双因素方差分析 171
12.2 一元回归分析 176
12.2.1 最小二乘法 177
12.2.2 线性化方法 178
12.2.3 相关性检验 180
12.2.4 预测与控制 181
12.3 正交试验设计 183
12.3.1 多因素试验与正交表 183
12.3.2 正交表的应用 184
12.3.3 考虑交互作用的正交试验设计 187
12.3.4 正交试验的方差分析 189

习题十二 191

第三篇 建模应用 196

第 13 章 线性代数的应用 196
13.1 矩阵的简化作用 196
13.2 线性运算技术 201
第 14 章 概率统计的应用 212
14.1 现实中的概率 212
14.2 统计推断 219
部分习题答案 230
附表 241
附表 1 标准正态分布表 241
附表 2 泊松分布数值表 242
附表 3 t 分布临界值表 243
附表 4 χ^2 分布临界值表 244
附表 5 F 分布临界值表 245
附表 6 秩和检验临界值表 250
附表 7 相关系数临界值表 251
附表 8 正交表 252
参考文献 254

第一篇　线 性 代 数

第 1 章　行 列 式

1.1　行列式的概念与性质

1.1.1　二阶、三阶行列式

当一个代数式很复杂时，需要引入适当的记号，以便简化它的表达形式、突现它的构成规则．行列式就是一种代数式的简要记号，如

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}\text{，} \tag{1.1}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}\text{，} \tag{1.2}$$

分别是二阶、三阶行列式，两式的左端表示行列式的记号，右端是行列式的全面展开式．行列式的元素有两个下标，分别称为行标和列标，如 a_{32} 表示该元素位于第 3 行、第 2 列．

二阶、三阶行列式的全面展开可以用对角线法：在图 1-1 中，实对角线上元素的乘积，前置符号取正；虚对角线上元素的乘积，前置符号取负．它们的代数和即行列式的值．

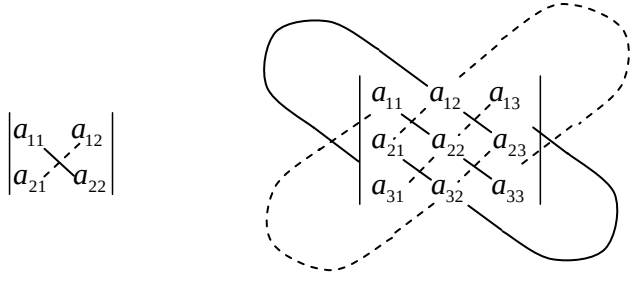

图 1-1　行列式的对角线

例如

$$\begin{vmatrix} 5 & -1 \\ 3 & 2 \end{vmatrix} = 5\times 2-(-1)\times 3 = 13\ ;\quad \begin{vmatrix} a & b \\ -b & a \end{vmatrix} = a^2-(-b^2) = a^2+b^2\ ;$$

$$\begin{vmatrix} 2 & -5 & 0 \\ 1 & 3 & -3 \\ 4 & -1 & 6 \end{vmatrix} = 2\times 3\times 6+1\times(-1)\times 0+(-5)\times(-3)\times 4-0\times 3\times 4$$

$$-(-1)\times(-3)\times 2-1\times(-5)\times 6 = 36+0+60-0-6-(-30) = 120\ .$$

1.1.2 n 阶行列式的全面展开

用 n^2 个元素可以构成 n 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & \mathrm{L} & a_{nn} \end{vmatrix}.$$

行列式有时简记为 $|a_{ij}|$．一阶行列式 $|a|$ 就是 a，为了避免与绝对值记号相混，一阶行列式通常不加行列式号．

高于 4 阶的行列式不能用对角线法展开．参照二阶、三阶行列式的展开式 (1.1)、(1.2)，规定 n 阶行列式的全面展开按如下方式进行：

(1) 展开式的每一项都是不同行、不同列的 n 个元素的乘积．

(2) 取自不同行、不同列的 n 个元素要出现所有不同的搭配．若将行标顺序安排，则每一项对应列标的一个排列．如 $a_{12}a_{21}a_{33}$ 对应的排列是 2 1 3．所有不同的搭配，对应所有不同的列标排列，n 个自然数共有 $n!$ 种排列，因而全面展开式共有 $n!$ 项．

(3) 各项的前置符号，偶排列取正，奇排列取负．所谓偶（奇）排列是指该排列的逆序数为偶（奇）数．比如排列 4 3 1 2 中，4 后面有比它小的 3，1，2（算作 3 个逆序），3 后面有 1，2，合计共有 5 个逆序，是奇排列．全面展开式的 $n!$ 项中有半数的前置符号为正，另一半为负．

通过全面展开来计算行列式显然是很复杂的，应该考虑简便的方法．

1.1.3 行列式的性质

将行列式 D 的行与列互换后得到的行列式，称为 D 的转置行列式，记为 D^{T}．即

$$\boldsymbol{D}=\begin{vmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & \mathrm{L} & a_{nn} \end{vmatrix},\quad \boldsymbol{D}^{\mathrm{T}}=\begin{vmatrix} a_{11} & a_{21} & \mathrm{L} & a_{n1} \\ a_{12} & a_{22} & \mathrm{L} & a_{n2} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{1n} & a_{2n} & \mathrm{L} & a_{nn} \end{vmatrix}.$$

实际书写时，“横着看，竖着写”，便可得到转置行列式．

性质 1　行列式转置后，其值不变，即 $\boldsymbol{D}^{\mathrm{T}}=\boldsymbol{D}$．

例如 $\begin{vmatrix} 1 & 2 & 3 \\ 8 & 9 & 4 \\ 7 & 6 & 5 \end{vmatrix}=\begin{vmatrix} 1 & 8 & 7 \\ 2 & 9 & 6 \\ 3 & 4 & 5 \end{vmatrix}$．

证　在行列式 $\boldsymbol{D}$ 中，每一行取一个元素，这 n 个元素位于不同的列，它们的乘积添上前置符号构成了 $\boldsymbol{D}$ 的展开式中的一项．该项中的元素也可以理解为取自不同的列，并位于不同的行，而这正是 $\boldsymbol{D}^{\mathrm{T}}$ 的展开式中的一项．可见 $\boldsymbol{D}$ 和 $\boldsymbol{D}^{\mathrm{T}}$ 的展开式中各项都对应相同，因此它们相等．

这条性质告诉我们，行列式的行具有某一性质，它们的列也具有相同的性质．

性质 2　交换行列式的两行（列），行列式的值变号．

例如 $\begin{vmatrix} 1 & 2 & 3 \\ 8 & 9 & 4 \\ 7 & 6 & 5 \end{vmatrix}=-\begin{vmatrix} 1 & 2 & 3 \\ 7 & 6 & 5 \\ 8 & 9 & 4 \end{vmatrix}$．

证　交换行列式的两行，相当于在展开式每一项所对应的列标排列中，交换了两个数字的位置．这两个数字之间的逆序发生了变化，而这两个数字和其他数字之间的逆序变化是成对发生的，因此整个排列的逆序数变化量为奇数，从而排列的奇偶性发生改变．即行列式展开式中的每一项都改变了符号，于是行列式的值变号．

性质 3　行列式的某一行（列）元素有公因子，可以提到行列式的外面．

例如 $\begin{vmatrix} 1 & 2 & 3 \\ 8k & 9k & 4k \\ 7 & 6 & 5 \end{vmatrix}=k\begin{vmatrix} 1 & 2 & 3 \\ 8 & 9 & 4 \\ 7 & 6 & 5 \end{vmatrix}$．

证　行列式的某一行有公因子 k 时，因为行列式展开式的每一项中都出现了该行的一个元素，所以每一项都有了公因子 k，当然可以提取出来．

这条性质也可以反向运用：行列式乘以数 k，等于把 k 乘到行列式的某一行（列）上去．

推论　以下 3 种行列式的值为零．

（1）行列式有某一行（列）的元素全为零；

（2）行列式有两行（列）完全相同；

（3）行列式有两行（列）的元素对应成比例．

证 其中第 1 种行列式有公因子 0 ；第 2 种行列式交换两行（列）后，其值不变，同时又改变符号，即 $D=-D$ ，故 $D=0$ ；第 3 种行列式提取公因子后，即第 2 种行列式.

性质 4 一个行列式可以拆分成两个行列式的和，这两个行列式的某对应行（列）上相同位置的元素之和，正好等于原行列式的对应位置的元素，而其他行（列）的元素都与原行列式相同.

例如
$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ 1 & 2 & 3 & 4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ 2 & 0 & -1 & 3 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ -1 & 2 & 4 & 1 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix}.$$

证 因为在行列式展开式的各项中，可以把来自于某行（列）的元素拆分成两数之和，再利用分配律将每一项都拆成两项之和，由此组合成两个行列式，而且行列式中除被拆分的元素外，其他元素都未变.

这条性质给出了行列式的**拆分规则**. 若反向运用，则成了行列式的**合并规则**. 拆分与合并规则特别强调：除某一对应行（列）外，其余元素都相同.

性质 5 把行列式的某一行（列）的各元素乘以同一数后加到另一行（列）的对应元素上去，行列式的值不变.

证 做了这种变换后的行列式可以拆分成两个行列式，一个是原行列式，另一个是推论中的第 3 种行列式（其值为零）.

比如在一个三阶行列式中将第 1 列乘以数 k 后加到第 3 列，所得行列式可拆分为
$$\begin{vmatrix} a_1 & b_1 & c_1+ka_1 \\ a_2 & b_2 & c_2+ka_2 \\ a_3 & b_3 & c_3+ka_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a_1 & b_1 & ka_1 \\ a_2 & b_2 & ka_2 \\ a_3 & b_3 & ka_3 \end{vmatrix}.$$

性质 5 将是计算行列式的主要工具，务必正确理解：“ 把某一行（列）乘以同一数后加到另一行（列）上去 ”，运算的对象是整行（列）中的每一个元素，不可偏漏；运算后，其中的 “ 某一行（列）” 应保持原样，而 “ 另一行（列）” 须发生变化；行列式的元素发生了变化，但行列式的值不变. 例如在下面的行列式中，把第 3 行乘以 3 加到第 1 行上去，可以得到等式
$$\begin{vmatrix} 7 & 2 & 0 & 3 \\ -5 & 6 & 2 & 0 \\ 2 & 4 & 5 & -1 \\ -3 & 1 & -1 & 5 \end{vmatrix} = \begin{vmatrix} 13 & 14 & 15 & 0 \\ -5 & 6 & 2 & 0 \\ 2 & 4 & 5 & -1 \\ -3 & 1 & -1 & 5 \end{vmatrix}.$$

1.2　行列式的降阶算法

1.2.1　代数余子式

在 n 阶行列式中，把元素 a_{ij} 所在的第 i 行和第 j 列划去后，余下的 $(n-1)$ 阶行列式叫做元素 a_{ij} 的**余子式**，记为 M_{ij}．再记

$$A_{ij}=(-1)^{i+j}M_{ij}\text{，}\qquad (1.3)$$

A_{ij} 叫做 a_{ij} 的**代数余子式**.

1.2.2　特殊行列式的计算公式

若行列式第 1 行元素除 a_{11} 外均为零，则有如下计算式：

$$\boldsymbol{D}=\begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=a_{11}\begin{vmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

证　等式右端的 $(n-1)$ 阶行列式是 a_{11} 的余子式 M_{11}．在行列式 $\boldsymbol{D}$ 中，每行选取一个元素（不同列）做乘积，构成展开式的各项．可以只考虑不等于零的项，显然第 1 行只能取 a_{11}，于是各项有了公因子 a_{11}．选取第 2 至第 n 行的元素，和全面展开 M_{11} 的选取方式完全相同；另一方面 a_{11} 与其他元素不构成逆序，因此 $\boldsymbol{D}$ 的展开项与 M_{11} 的展开项前置符号也相同．提取公因子 a_{11} 后，即得 $\boldsymbol{D}=a_{11}M_{11}$．

若行列式的第 1 列除 a_{11} 外均为零，则有同样的计算式．

现在考虑一般的降阶条件．如果在行列式 $\boldsymbol{D}$ 中，非零元素 a_{ij} 所在的行（第 i 行）或列（第 j 列）的其他元素均为零，那么可以将第 i 行依次与第 $(i-1)$，…，2，1 行交换，然后再将第 j 列依次与第 $(j-1)$，…，2，1 列交换，共经过 $(i+j-2)$ 次换行与换列，把 a_{ij} 换到左上角的位置，其所在的行与列也换到了第 1 行和第 1 列，而其他行、列的排列顺序没有改变．根据前面的公式可知，最后的行列式的值变成了 $a_{ij}M_{ij}$．由于行列式的值变了 $(i+j-2)$ 次符号，所以

$$\boldsymbol{D}=(-1)^{i+j-2}a_{ij}M_{ij}=a_{ij}A_{ij}\text{．}\qquad (1.4)$$

定理 1.1　如果行列式的某行（或某列）中，仅有一个元素非零，那么行列式的值等于该非零元素与它的代数余子式的乘积．

定理 1.1 是行列式**降阶算法**的基础．比如下面的行列式经逐次降阶，很容易得到它的

值：

$$\begin{vmatrix} a_{11} & 0 & 0 & \mathrm{L} & 0 \\ a_{21} & a_{22} & 0 & \mathrm{L} & 0 \\ a_{31} & a_{32} & a_{33} & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & a_{n3} & \mathrm{L} & a_{nn} \end{vmatrix} = a_{11}a_{22}a_{33}\mathrm{L}\ a_{nn}.$$

我们称这种行列式为下三角行列式．类似地，上三角行列式也有同样的计算公式：

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \mathrm{L} & a_{1n} \\ 0 & a_{22} & a_{23} & \mathrm{L} & a_{2n} \\ 0 & 0 & a_{33} & \mathrm{L} & a_{3n} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & 0 & \mathrm{L} & a_{nn} \end{vmatrix} = a_{11}a_{22}a_{33}\mathrm{L}\ a_{nn}.$$

更特殊的对角行列式的值为

$$\begin{vmatrix} a_{11} & 0 & 0 & \mathrm{L} & 0 \\ 0 & a_{22} & 0 & \mathrm{L} & 0 \\ 0 & 0 & a_{33} & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & 0 & \mathrm{L} & a_{nn} \end{vmatrix} = a_{11}a_{22}a_{33}\mathrm{L}\ a_{nn}.$$

我们把行列式中从左上角到右下角的对角线称为主对角线．三角行列式中，主对角线一侧的元素全为零；对角行列式中，主对角线两侧的元素全为零．

推论　三角行列式及对角行列式的值等于其主对角线上所有元素的乘积．

1.2.3　行列式的降阶算法

对于一般的行列式，可以利用行列式的性质 5 把某些元素变为零，来达到降阶条件．具体做法是：在行列式中选定一个非零元素，称之为主元，将主元所在的行乘以适当的数加到其他行上，由此把主元所在列的其他元素都变为零．这个过程叫做消元．

例 1.1　计算行列式 $\boldsymbol{D} = \begin{vmatrix} -10 & 8 & -7 & 2 \\ 7 & -9 & 13 & 1 \\ -5 & -1 & 5 & 3 \\ 3 & 5 & -1 & -2 \end{vmatrix}$.

解　选 $a_{24} = 1$ **为主元，为了将第 4 列的元素** $2, 3, -2$ **变成零，进行行变换：第 2 行乘** -2

加到第 1 行；第 2 行乘 –3 加到第 3 行；第 2 行乘 2 加到第 4 行，得到

$$\boldsymbol{D}=\begin{vmatrix} -24 & 26 & -33 & 0 \\ 7 & 9 & 13 & 1 \\ -26 & 26 & -34 & 0 \\ 17 & -13 & 25 & 0 \end{vmatrix}=1\times(-1)^{2+4}\begin{vmatrix} -24 & 26 & -33 \\ -26 & 26 & -34 \\ 17 & -13 & 25 \end{vmatrix}.$$

降阶后的行列式可继续降阶：取第 2 行的 26 为主元，将第 2 行分别乘 –26/26 和 13/26 加到第 1、第 3 行，得到

$$\boldsymbol{D}=\begin{vmatrix} 2 & 0 & 1 \\ -26 & 26 & -34 \\ 4 & 0 & 8 \end{vmatrix}=26\times(-1)^{2+2}\begin{vmatrix} 2 & 1 \\ 4 & 8 \end{vmatrix}=312\ .$$

从本例看出，要注重主元的选取，适当地选择主元可以简化计算．

行列式的消元运算也可以做列变换，即通过列变换把主元所在行的其他元素都变为零．

阶数较高的行列式计算常常借助于计算机．计算机不怕烦琐的计算，但无法“灵活”地选择主元，所以计算机一般是先把行列式变为三角行列式，即从第 1 列开始，通过行变换依次将各列主对角线下方的元素变为零．

例 1.2　将例 1.1 的行列式 $\boldsymbol{D}$ 化为上三角行列式，并计算其值．

解　以下计算过程中，所选主元都用“〈 〉”号做了标记，这样可使消元的对象清晰，便于突出变换的有序性：

$$\boldsymbol{D}=\begin{vmatrix} \langle -10\rangle & 8 & -7 & 2 \\ 7 & -9 & 13 & 1 \\ -5 & -1 & 5 & 3 \\ 3 & 5 & -1 & -2 \end{vmatrix}=\begin{vmatrix} -10 & 8 & -7 & 2 \\ 0 & \langle -3.4\rangle & 8.1 & 2.4 \\ 0 & -5 & 8.5 & 2 \\ 0 & 7.4 & -3.1 & -1.4 \end{vmatrix}$$

$$=\begin{vmatrix} -10 & 8 & -7 & 2 \\ 0 & -3.4 & 8.1 & 2.4 \\ 0 & 0 & \langle -3.412\rangle & -1.529 \\ 0 & 0 & 14.529 & 3.824 \end{vmatrix}=\begin{vmatrix} -10 & 8 & -7 & 2 \\ 0 & -3.4 & 8.1 & 2.4 \\ 0 & 0 & -3.412 & -1.529 \\ 0 & 0 & 0 & -2.687 \end{vmatrix}=311.7\ .$$

数值计算有误差是难免的．为了减少误差，可以选取绝对值较大的数作为主元，同时通过换行把主元换到主对角线上的位置．这种做法还有一个好处是能适应主对角线上元素出现零的情况．

1.3 克莱姆法则

1.3.1 行列式的按行（列）展开

按照行列式的拆分规则，一个三阶行列式可作如下拆分：

$$\boldsymbol{D}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

再根据定理 1.1**，得到**

$$\boldsymbol{D}=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13}.$$

这叫做行列式按第 1 **行展开．因为拆分可针对不同的行、列进行，所以行列式也可按其他行展开，还可以按各列展开，而且这种展开能直接推广到更高阶的行列式．一般地，对于** n **阶行列式** $\boldsymbol{D}$**，有如下展开式：**

$$\boldsymbol{D}=a_{i1}A_{i1}+a_{i2}A_{i2}+\mathrm{L}+a_{in}A_{in} \quad (i=1,2,\cdots,n),$$

$$\boldsymbol{D}=a_{1j}A_{1j}+a_{2j}A_{2j}+\mathrm{L}+a_{nj}A_{nj} \quad (j=1,2,\cdots,n).$$

这两个公式分别称为行列式的按行展开式与按列展开式．展开式右端是某行（列）元素与它们的代数余子式乘积之和．以后，我们把这种两两相乘再相加的运算叫做“组合”，组合运算是线性代数中常见的运算．

行列式的按行或按列展开，起到了降阶的效果，可以用来计算行列式的值．比如例 1.1 **的行列式，按第** 3 **行展开为**

$$\boldsymbol{D}=\begin{vmatrix} -10 & 8 & -7 & 2 \\ 7 & -9 & 13 & 1 \\ -5 & -1 & 5 & 3 \\ 3 & 5 & -1 & -2 \end{vmatrix}=(-5)\times(-1)^{3+1}\begin{vmatrix} 8 & -7 & 2 \\ -9 & 13 & 1 \\ 5 & -1 & -2 \end{vmatrix}+(-1)\times(-1)^{3+2}\begin{vmatrix} -10 & -7 & 2 \\ 7 & 13 & 1 \\ 3 & -1 & -2 \end{vmatrix}$$

$$+5\times(-1)^{3+3}\begin{vmatrix} -10 & 8 & 2 \\ 7 & -9 & 1 \\ 3 & 5 & -2 \end{vmatrix}+3\times(-1)^{3+4}\begin{vmatrix} -10 & 8 & -7 \\ 7 & -9 & 13 \\ 3 & 5 & -1 \end{vmatrix}.$$

不过用这种方法要计算 n **个** $(n-1)$ **阶行列式，如果连环地不断展开，计算量与全面展开不相上下，是非常庞大的．所以计算行列式还是以上节介绍的降阶算法为主．**

1.3.2　代数余子式组合定理

在三阶行列式

$$\boldsymbol{D}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

中，改写第 1 行，使之与第 2 行相同，将这个值为零的行列式按第 1 行展开，注意到代数余子式未变，有

$$0=\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=a_{21}A_{11}+a_{22}A_{12}+a_{23}A_{13}\text{．}$$

一般地，在 n 阶行列式 $\boldsymbol{D}$ 中，把第 i 行改写为第 k 行 $(k\neq i)$，再按第 i 行展开，可得

$$a_{k1}A_{i1}+a_{k2}A_{i2}+\mathrm{L}\ +a_{kn}A_{in}=0\text{；}$$

把第 j 列改写成第 k 列 $(k\neq j)$，再按第 j 列展开，可得

$$a_{1k}A_{1j}+a_{2k}A_{2j}+\mathrm{L}\ +a_{nk}A_{nj}=0\text{．}$$

将这两个公式与行列式的按行（列）展开式结合起来，即

$$\sum_{j=1}^{n}a_{kj}A_{ij}=\begin{cases}\boldsymbol{D}, & k=i\\ 0, & k\neq i\end{cases}\text{，}\tag{1.5}$$

$$\sum_{i=1}^{n}a_{ik}A_{ij}=\begin{cases}\boldsymbol{D}, & k=j\\ 0, & k\neq j\end{cases}\text{．}\tag{1.6}$$

我们把这些公式归结为**代数余子式组合定理**：

定理 1.2　行列式的某行（列）元素与本行（列）元素的代数余子式相组合，结果等于该行列式的值；行列式的某行（列）元素与另一行（列）元素的代数余子式相组合，结果等于零．

1.3.3　克莱姆法则

含有 n 个未知量、n 个方程的线性方程组为

$$\begin{cases} a_{11}x_1+a_{12}x_2+\mathrm{L}\ +a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\mathrm{L}\ +a_{2n}x_n=b_2\\ \qquad\qquad \mathrm{L}\ \ \mathrm{L}\\ a_{n1}x_1+a_{n2}x_2+\mathrm{L}\ +a_{nn}x_n=b_n\end{cases}\text{．}\tag{1.7}$$

未知量的所有系数构成的行列式

$$\boldsymbol{D}=\begin{vmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & \mathrm{L} & a_{nn} \end{vmatrix},$$

称为方程组（1.7）的**系数行列式**. 用 $\boldsymbol{D}$ 的代数余子式 A_{i1}（$i=1,2,\cdots,n$）依次乘方程组（1.7）的各个方程，再把所有方程相加，得到

$$\sum_{i=1}^{n}A_{i1}(a_{i1}x_1+a_{i2}x_2+\mathrm{L}\ +a_{in}x_n)=\sum_{i=1}^{n}b_iA_{i1}\ ,$$

将左端的乘积展开，再合并含同一未知量的项，对照公式（1.6）发现，除 x_1 的系数是 $\boldsymbol{D}$ 外，其余未知量的系数都是零；右端的和式正是以下行列式按第 1 列展开的展开式：

$$\boldsymbol{D}_1=\begin{vmatrix} b_1 & a_{12} & \mathrm{L} & a_{1n} \\ b_2 & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ b_n & a_{n2} & \mathrm{L} & a_{nn} \end{vmatrix}.$$

于是得到 $\boldsymbol{D}x_1=\boldsymbol{D}_1$. 如果 $\boldsymbol{D}\neq 0$，可立即算出 x_1 的值. 一般地，用代数余子式 A_{ij}（$i=1,2,\cdots,n$）乘方程组（1.7）的各个方程后再相加，能够保留 x_j 而消去其他所有未知量. 综合以上推导，可形成解线性方程组的**克莱姆法则**：

定理 1.3 线性方程组（1.7）当其系数行列式 $\boldsymbol{D}\neq 0$ 时，有且仅有惟一解

$$x_j=\frac{\boldsymbol{D}_j}{\boldsymbol{D}}\ (j=1,2,\cdots,n), \tag{1.8}$$

其中 $\boldsymbol{D}_j$（$j=1,2,\cdots,n$）是以方程组的常数项 b_1，b_2，…，b_n 替换 $\boldsymbol{D}$ 的第 j 列元素后形成的行列式.

例 1.3 求解线性方程组 $\begin{cases} x_1-x_2+x_3-2x_4=2 \\ 2x_1-x_3+4x_4=4 \\ 3x_1+2x_2+x_3=-1 \\ -x_1+2x_2-x_3+2x_4=-4 \end{cases}$.

解 用降阶法计算行列式，所选主元做了标记：

$$\boldsymbol{D}=\begin{vmatrix} 1 & -1 & \langle 1\rangle & -2 \\ 2 & 0 & -1 & 4 \\ 3 & 2 & 1 & 0 \\ -1 & 2 & -1 & 2 \end{vmatrix}=\begin{vmatrix} 1 & -1 & 1 & -2 \\ 3 & -1 & 0 & 2 \\ 2 & 3 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{vmatrix}=\begin{vmatrix} 3 & -1 & 2 \\ 2 & 3 & 2 \\ 0 & 1 & 0 \end{vmatrix}=-2\neq 0\ ,$$

$$\boldsymbol{D}_1=\begin{vmatrix}2&-1&\langle 1\rangle&-2\\4&0&-1&4\\-1&2&1&0\\-4&2&-1&2\end{vmatrix}=\begin{vmatrix}2&-1&1&-2\\6&-1&0&2\\-3&3&0&2\\-2&1&0&0\end{vmatrix}=\begin{vmatrix}6&-1&2\\-3&3&2\\-2&1&0\end{vmatrix}=-2,$$

$$\boldsymbol{D}_2=\begin{vmatrix}1&2&\langle 1\rangle&-2\\2&4&-1&4\\3&-1&1&0\\-1&-4&-1&2\end{vmatrix}=\begin{vmatrix}1&2&1&-2\\3&6&0&2\\2&-3&0&2\\0&-2&0&0\end{vmatrix}=\begin{vmatrix}3&6&2\\2&-3&2\\0&-2&0\end{vmatrix}=4,$$

类似地可算得 $\boldsymbol{D}_3=\begin{vmatrix}1&\langle -1\rangle&2&-2\\2&0&4&4\\3&2&-1&0\\-1&2&-4&2\end{vmatrix}=0$，$\boldsymbol{D}_4=\begin{vmatrix}1&-1&\langle 1\rangle&2\\2&0&-1&4\\3&2&1&-1\\-1&2&-1&-4\end{vmatrix}=-1$，

所以 $x_1=\dfrac{\boldsymbol{D}_1}{\boldsymbol{D}}=1$，$x_2=\dfrac{\boldsymbol{D}_2}{\boldsymbol{D}}=-2$，$x_3=\dfrac{\boldsymbol{D}_3}{\boldsymbol{D}}=0$，$x_4=\dfrac{\boldsymbol{D}_4}{\boldsymbol{D}}=\dfrac{1}{2}$.

用克莱姆法则解线性方程组要计算 $(n+1)$ 个 n 阶行列式，计算量很大而且包含了大量的重复计算．我们将在第 3 章介绍更好的方法．

如果线性方程组（1.7）的常数项全为零，即

$$\begin{cases}a_{11}x_1+a_{12}x_2+\mathrm{L}+a_{1n}x_n=0\\a_{21}x_1+a_{22}x_2+\mathrm{L}+a_{2n}x_n=0\\\qquad\qquad\mathrm{L}\ \mathrm{L}\\a_{n1}x_1+a_{n2}x_2+\mathrm{L}+a_{nn}x_n=0\end{cases},\tag{1.9}$$

则称之为**齐次线性方程组**，而方程组（1.7）当常数项不全为零时，称为**非齐次线性方程组**．

显然，方程组（1.9）一定有**零解** $x_j=0$（$j=1,\ 2,\ \cdots,\ n$）．如果系数行列式 $\boldsymbol{D}\neq 0$，则根据克莱姆法则，它只有惟一的零解，没有非零解．以后还可以证明：如果 $\boldsymbol{D}=0$，则方程组（1.9）一定有**非零解**（未知量的取值不全为零）．我们提前归纳为以下定理：

定理 1.4　齐次线性方程组（1.9）有非零解的充分必要条件是它的系数行列式等于零．

定理 1.4 将在第 5 章有重要的应用．

例 1.4　已知齐次线性方程组 $\begin{cases}x+\ y+\ z=0\\2x-\ y+3z=0\\Ax+By+Cz=0\end{cases}$ 有非零解，问 A，B，C 应满足什么条件？

解　根据定理 1.4，齐次线性方程组有非零解的充分必要条件是它的系数行列式等于

零，即 $\begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ A & B & C \end{vmatrix} = 0$．计算行列式的值，得到

$$4A - B - 3C = 0,$$

这就是 A，B，C 应满足的条件．

本例的几何意义是过坐标原点的 3 个平面共线．

习 题 一

1．计算下列行列式．

(1) $\begin{vmatrix} n+1 & n \\ n & n-1 \end{vmatrix}$；　(2) $\begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix}$；　(3) $\begin{vmatrix} 1 & \log_b a \\ \log_a b & 1 \end{vmatrix}$；

(4) $\begin{vmatrix} 3 & 2 & 1 \\ 2 & 5 & 3 \\ 3 & 4 & 2 \end{vmatrix}$；　(5) $\begin{vmatrix} 4 & -3 & 5 \\ 3 & -2 & 8 \\ 1 & -7 & -5 \end{vmatrix}$；　(6) $\begin{vmatrix} 3 & 4 & -5 \\ 8 & 7 & -2 \\ 2 & -1 & 8 \end{vmatrix}$；

(7) $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$；　(8) $\begin{vmatrix} \cos\theta\cos\phi & \sin\theta\cos\phi & \sin\phi \\ -r\sin\theta\cos\phi & r\cos\theta\cos\phi & 0 \\ -r\cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\phi \end{vmatrix}$．

2．求解方程 $\begin{vmatrix} x-5 & 2 & 0 \\ 2 & x-6 & -2 \\ 0 & -2 & x-7 \end{vmatrix} = 0$．

3．$\begin{vmatrix} 1 & a & -1 \\ a & 1 & 2 \\ -1 & 2 & 5 \end{vmatrix} > 0$ 的充分必要条件是什么？

4．利用行列式的性质证明下列各式．

(1) $\begin{vmatrix} b+c & c+a & a+b \\ a+b & b+c & c+a \\ c+a & a+b & b+c \end{vmatrix} = 2\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$；　(2) $\begin{vmatrix} \lambda-3 & -2 & -4 \\ -2 & \lambda & -2 \\ -4 & -2 & \lambda-3 \end{vmatrix} = (\lambda+1)^2(\lambda-8)$；

(3) $\begin{vmatrix} x & a & a & a \\ a & x & a & a \\ a & a & x & a \\ a & a & a & x \end{vmatrix} = (3a+x)(x-a)^3$．

5．已知行列式 $\boldsymbol{D} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ a & b & c & d \\ 1 & 4 & 10 & 20 \end{vmatrix}$，求元素 a，b 的代数余子式的值．

6．运用行列式的降阶算法计算下列行列式．

(1) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix}$；　(2) $\begin{vmatrix} 2 & -5 & 1 & 2 \\ -3 & 7 & -1 & 4 \\ 5 & -9 & 2 & 7 \\ 4 & -6 & 1 & 2 \end{vmatrix}$；　(3) $\begin{vmatrix} 3 & 1 & 4 & 2 \\ 15 & 2 & 14 & 29 \\ 16 & 3 & 17 & 19 \\ 33 & 8 & 38 & 39 \end{vmatrix}$；

(4) $\begin{vmatrix} 3 & -5 & -2 & 2 \\ -4 & 7 & 4 & 4 \\ 4 & -9 & -3 & 7 \\ 2 & -6 & -3 & 2 \end{vmatrix}$；　(5) $\begin{vmatrix} 3 & -5 & 2 & -4 \\ -3 & 4 & -5 & 3 \\ -5 & 7 & -7 & 5 \\ 8 & -8 & 5 & -6 \end{vmatrix}$；　(6) $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 7 & 10 & 13 \\ 3 & 5 & 11 & 16 & 21 \\ 2 & -7 & 7 & 7 & 2 \\ 1 & 4 & 5 & 3 & 10 \end{vmatrix}$．

7．设 $\boldsymbol{D}=\begin{vmatrix} 6 & -5 & 8 & 4 \\ 9 & 7 & 5 & 2 \\ a & b & c & 1 \\ -4 & 8 & -8 & -3 \end{vmatrix}=1$，行列式 $\boldsymbol{D}$ 的代数余子式记为 A_{ij}（$i,j=1,2,3,4$），求下列各式的值：

(1) $9A_{21}+7A_{22}+5A_{23}+2A_{24}$；　(2) $9A_{31}+7A_{32}+5A_{33}+2A_{34}$；

(3) $4A_{12}+2A_{22}+A_{32}-3A_{42}$；　(4) $4A_{13}+2A_{23}+A_{33}-3A_{43}$．

8．用克莱姆法则解方程组 $\begin{cases} x_1+x_2+x_3+x_4=5 \\ x_1+2x_2-x_3+4x_4=-2 \\ 2x_1-3x_2-x_3-5x_4=-2 \\ 3x_1+x_2+2x_3+11x_4=0 \end{cases}$．

9．已知齐次线性方程组 $\begin{cases} (15-2a)x_1+11x_2+10x_3=0 \\ (11-3a)x_1+17x_2+16x_3=0 \\ (7-a)x_1+14x_2+13x_3=0 \end{cases}$ 有非零解，问 a 应取什么值？

10．单项选择题

(1) 如果 $\boldsymbol{D}=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=5$，那么 $\begin{vmatrix} 2a_{11} & 2a_{12} & 2a_{13} \\ 2a_{21} & 2a_{22} & 2a_{23} \\ 2a_{31} & 2a_{32} & 2a_{33} \end{vmatrix}=$　（　）

A．10；　B．−10；　C．40；　D．−40．

(2) 已知 $\boldsymbol{D}=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}=1$，则 $\begin{vmatrix} 2a_1 & 3a_1-4b_1 & c_1 \\ 2a_2 & 3a_2-4b_2 & c_2 \\ 2a_3 & 3a_3-4b_3 & c_3 \end{vmatrix}=$　（　）

A．6；　B．−2；　C．−8；　D．−24．

(3) 三阶行列式 $\begin{vmatrix} -2 & 3 & 1 \\ 503 & 201 & 298 \\ 5 & 2 & 3 \end{vmatrix}=$　（　）

A．63；　B．70；　C．−70；　D．82．

(4) 行列式 $\begin{vmatrix} 0 & a & 0 & 0 \\ b & c & 0 & 0 \\ 0 & 0 & d & e \\ 0 & 0 & 0 & f \end{vmatrix} =$ ( )

A．$abcdef$；　　B．cdf；　　C．$abdf$；　　D．$-abdf$．

(5) 下列 n（$n>2$）阶行列式的值必为零的是 ( )

A．行列式主对角线上的元素全为零；　　B．行列式中有一半的元素等于零；

C．行列式中非零元素的个数小于 n；　　D．行列式的元素中每个数都重复出现 n 次．

(6) 设 $\boldsymbol{D}=|a_i|$ 是 n 阶行列式，且 $\boldsymbol{D}\neq 0$，A_i 是元素 a_i 的代数余子式，则 $\sum_{i=1}^{n} a_{i2}A_{i3} =$ ( )

A．$\boldsymbol{D}$；　　B．0；　　C．$\frac{1}{\boldsymbol{D}}$；　　D．难以确定其值．

(7) 已知 $a_1b_2-a_2b_1=\frac{1}{m}$，则方程组 $\begin{cases} a_1x+b_1y=c_1 \\ a_2x+b_2y=c_2 \end{cases}$ 的解是 ( )

A．$x=\frac{1}{m}\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$，$y=\frac{1}{m}\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$；　　B．$x=\frac{1}{m}\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$，$y=\frac{1}{m}\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$；

C．$x=m\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$，$y=m\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$；　　D．$x=m\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}$，$y=m\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$．

(8) 设 $\boldsymbol{D}$ 是含有 n 个变量和 n 个方程组的线性方程组的系数行列式，下列说法中正确的是 ( )

A．若 $\boldsymbol{D}\neq 0$，则线性方程组有解；　　B．若 $\boldsymbol{D}=0$，则线性方程组无解；

C．若线性方程组有解，则必有 $\boldsymbol{D}\neq 0$；　　D．若线性方程组无非零解，则必有 $\boldsymbol{D}=0$．

(9) 已知方程组 $\begin{cases} kx+3y+z=0 \\ 2x+ky+z=0 \\ kx-2y+z=0 \end{cases}$ 有非零解，则 $k=$ ( )

A．0；　　B．1；　　C．2；　　D．3．

(10) 方程组 $\begin{cases} 3x+ky+z=0 \\ 4y+z=0 \\ kx-5y-z=0 \end{cases}$ 只有零解的充分必要条件是 ( )

A．$k\neq 1$；　　B．$k\neq 3$；　　C．$k\neq 1$ 或 $k\neq 3$；　　D．$k\neq 1$ 且 $k\neq 3$．

第 2 章　矩　　阵

2.1　矩阵的概念及其线性运算

学习本节内容，要特别注意与行列式的有关概念、运算相区别．

2.1.1　矩阵的概念

矩阵是一张简化了的表格，一般地，

$$\begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{m1} & a_{m2} & \mathrm{L} & a_{mn} \end{pmatrix}$$

称为 $\boldsymbol{m\times n}$ **矩阵**，它有 m 行、n 列，共 $m\times n$ 个元素，其中第 i 行、第 j 列的元素用 a_{ij} 表示．通常我们用大写黑体字母 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$，…表示矩阵．为了标明矩阵的行数 m 和列数 n，可用 $\boldsymbol{A}_{m\times n}$ 或 $\left(a_{ij}\right)_{m\times n}$ 表示．矩阵既然是一张表，就不能像行列式那样算出一个数来．

所有元素均为 0 的矩阵，称为**零矩阵**，记作 $\boldsymbol{O}$．

两个矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 相等，意味着不仅它们的行、列数相同，而且所有对应元素都相同，记作 $\boldsymbol{A}=\boldsymbol{B}$．

如果矩阵 $\boldsymbol{A}$ 的行、列数都是 n，则称 $\boldsymbol{A}$ 为 $\boldsymbol{n}$ **阶矩阵**，或称为 $\boldsymbol{n}$ **阶方阵**．n 阶矩阵有一条从左上角到右下角的**主对角线**．n 阶矩阵 $\boldsymbol{A}$ 的元素按原次序构成的 n 阶行列式，称为**矩阵 $\boldsymbol{A}$ 的行列式**，记作 $|\boldsymbol{A}|$．

在 n 阶矩阵中，若主对角线左下侧的元素全为零，则称之为**上三角矩阵**；若主对角线右上侧的元素全为零，则称之为**下三角矩阵**；若主对角线两侧的元素全为零，则称之为**对角矩阵**．主对角线上元素全为 1 的对角矩阵，叫做**单位矩阵**，记为 $\boldsymbol{E}$，即

$$\boldsymbol{E}=\begin{pmatrix} 1 & 0 & \mathrm{L} & 0 \\ 0 & 1 & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & \mathrm{L} & 1 \end{pmatrix}.$$

$1\times n$ 矩阵（只有一行）又称为 $\boldsymbol{n}$ **维行向量**；$n\times 1$ 矩阵（只有一列）又称为 $\boldsymbol{n}$ **维列向量**．行

向量、列向量统称为**向量**．向量通常用小写黑体字母 $\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{x}$，$\boldsymbol{y}$，…表示．向量中的元素又称为向量的**分量**．1×1 矩阵因只有一个元素，故视之为数量，即 $(a)=a$．

2.1.2 矩阵的加、减运算

如果矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 的行数和列数都相同，那么它们可以相加、相减，记为 $\boldsymbol{A}+\boldsymbol{B}$，$\boldsymbol{A}-\boldsymbol{B}$，分别称为矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 的**和与差**．$\boldsymbol{A}\pm\boldsymbol{B}$ 表示将 $\boldsymbol{A}$，$\boldsymbol{B}$ 中所有对应位置的元素相加、减得到的矩阵．例如

$$\boldsymbol{A}=\begin{pmatrix}-1 & 2 & 3\\ 0 & 3 & -2\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}4 & 3 & 2\\ 5 & -3 & 0\end{pmatrix},$$

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

$$\boldsymbol{A}-\boldsymbol{B}=\begin{pmatrix}-1-4 & 2-3 & 3-2\\ 0-5 & 3-(-3) & -2-0\end{pmatrix}=\begin{pmatrix}-5 & -1 & 1\\ -5 & 6 & -2\end{pmatrix}.$$

2.1.3 矩阵的数乘

矩阵 $\boldsymbol{A}$ 与数 k 相乘记为 $k\boldsymbol{A}$ 或 $\boldsymbol{A}k$．$k\boldsymbol{A}$ 表示将 k 乘 $\boldsymbol{A}$ 中的所有元素得到的矩阵．例如

$$\boldsymbol{A}=\begin{pmatrix}2 & 4\\ 3 & 0\\ 5 & 1\end{pmatrix},\quad 3\boldsymbol{A}=\begin{pmatrix}3\times2 & 3\times4\\ 3\times3 & 3\times0\\ 3\times5 & 3\times1\end{pmatrix}=\begin{pmatrix}6 & 12\\ 9 & 0\\ 15 & 3\end{pmatrix}.$$

当 $k=-1$ 时，我们简记 $(-1)\boldsymbol{A}=-\boldsymbol{A}$，称为 $\boldsymbol{A}$ 的**负矩阵**．

矩阵的加减与**数乘**统称为**线性运算**．不难验证线性运算满足交换律、结合律与分配律，这与数量的运算规律相同，所以在数量运算中形成的诸如提取公因子、合并同类项、移项变号、正负抵消等运算习惯，在矩阵的线性运算中都可以保留、沿用．

例 2.1 设

$$\boldsymbol{A}=\begin{pmatrix}3 & -1 & 2 & 0\\ 1 & 5 & 7 & 9\\ 2 & 4 & 6 & 8\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}7 & 5 & -2 & 4\\ 5 & 1 & 9 & 7\\ 3 & 2 & -1 & 6\end{pmatrix},$$

已知 $\boldsymbol{A}+2\boldsymbol{X}=\boldsymbol{B}$，求 $\boldsymbol{X}$．

解 在等式中移项得 $2\boldsymbol{X}=\boldsymbol{B}-\boldsymbol{A}$，再除以 2 得 $\boldsymbol{X}=\dfrac{1}{2}(\boldsymbol{B}-\boldsymbol{A})$．通过心算立得

$$\boldsymbol{X}=\begin{pmatrix}2 & 3 & -2 & 2\\ 2 & -2 & 1 & -1\\ 0.5 & -1 & -3.5 & -1\end{pmatrix}.$$

例 2.2　设 $\boldsymbol{A}$ 为三阶矩阵．已知 $|\boldsymbol{A}|=-2$，求行列式 $|3\boldsymbol{A}|$ 的值．

解　设 $\boldsymbol{A}=\begin{pmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{pmatrix}$，则 $3\boldsymbol{A}=\begin{pmatrix}3a_1 & 3a_2 & 3a_3\\ 3b_1 & 3b_2 & 3b_3\\ 3c_1 & 3c_2 & 3c_3\end{pmatrix}$．显然行列式 $|3\boldsymbol{A}|$ 中每行都有公因子 3，因此 $|3\boldsymbol{A}|=3^3\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}=27|\boldsymbol{A}|=-54$．

2.2　矩阵的乘法与转置

2.2.1　矩阵的乘法

如果矩阵 $\boldsymbol{A}$ 的列数与矩阵 $\boldsymbol{B}$ 的行数相同，即 $\boldsymbol{A}$ 是 $m\times s$ 矩阵，$\boldsymbol{B}$ 是 $s\times n$ 矩阵，那么 $\boldsymbol{A}$，$\boldsymbol{B}$ 可以相乘，记为 $\boldsymbol{AB}$ 或 $\boldsymbol{A}\cdot\boldsymbol{B}$，称为矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 的**乘积**．$\boldsymbol{AB}=\boldsymbol{C}$ 表示一个 $m\times n$ 矩阵，矩阵 $\boldsymbol{C}$ 的构成规则如下：

$\boldsymbol{B}$ 的第 1 列元素依次与 $\boldsymbol{A}$ 的各行元素相组合，形成 $\boldsymbol{C}$ 的第 1 列元素；$\boldsymbol{B}$ 的第 2 列元素依次与 $\boldsymbol{A}$ 的各行元素相组合，形成 $\boldsymbol{C}$ 的第 2 列元素……以此类推，最后 $\boldsymbol{B}$ 的第 n 列元素依次与 $\boldsymbol{A}$ 的各行元素相组合，形成 $\boldsymbol{C}$ 的第 n 列元素．这里的“组合”表示两两相乘再相加．

若记

$$\boldsymbol{A}=(a_{ij})_{m\times s},\quad \boldsymbol{B}=(b_{ij})_{s\times n},\quad \boldsymbol{C}=(c_{ij})_{m\times n},$$

且 $\boldsymbol{C}=\boldsymbol{AB}$，则乘积矩阵 $\boldsymbol{C}$ 的元素可用公式表示为

$$c_i=\sum_{k=1}^{s}a_{ik}b_k\quad (i=1,2,\cdots,m;\ j=1,2,\cdots,n).\tag{2.1}$$

例如

$$\begin{pmatrix}3 & -1\\ 0 & 3\\ 1 & 4\\ 2 & 1\end{pmatrix}\begin{pmatrix}1 & -2 & 3\\ 2 & 1 & 0\end{pmatrix}=\begin{pmatrix}3\times1+(-1)\times2 & 3\times(-2)+(-1)\times1 & 3\times3+(-1)\times0\\ 0\times1+3\times2 & 0\times(-2)+3\times1 & 0\times3+3\times0\\ 1\times1+4\times2 & 1\times(-2)+4\times1 & 1\times3+4\times0\\ 2\times1+1\times2 & 2\times(-2)+1\times1 & 2\times3+1\times0\end{pmatrix}=\begin{pmatrix}1 & -7 & 9\\ 6 & 3 & 0\\ 9 & 2 & 3\\ 4 & -3 & 6\end{pmatrix}.$$

利用矩阵的乘法可以简化线性方程组的表示形式．设

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \mathrm{L} + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \mathrm{L} + a_{2n}x_n = b_2 \\ \qquad\qquad \mathrm{L}\ \mathrm{L} \\ a_{m1}x_1 + a_{m2}x_2 + \mathrm{L} + a_{mn}x_n = b_m \end{cases} \tag{2.2}$$

是含有 m 个方程、 n 个变量的线性方程组，若记

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{m1} & a_{m2} & \mathrm{L} & a_{mn} \end{pmatrix},\quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \mathrm{M} \\ x_n \end{pmatrix},\quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \mathrm{M} \\ b_m \end{pmatrix},$$

则方程组可表示为矩阵方程

$$\boldsymbol{Ax} = \boldsymbol{b} \tag{2.3}$$

这个矩阵方程两端都是 $m\times 1$ 矩阵，对应元素都相同，相当于 m 个等式，恰好是（2.2）式的 m 个方程．（2.3）式称为线性方程组（2.2）的**矩阵形式**．以后，矩阵形式（2.3）将成为我们表示线性方程组的主要形式．其中 $\boldsymbol{A}$ 称为线性方程组的**系数矩阵**，$\boldsymbol{x}$ 称为**变量列**，$\boldsymbol{b}$ 称为**常数列**．

2.2.2 矩阵乘法的性质

两个矩阵相乘，要求行、列数相匹配，即在乘积 $\boldsymbol{AB}$ 中，矩阵 $\boldsymbol{A}$ 的列数必须等于矩阵 $\boldsymbol{B}$ 的行数，因此当 $\boldsymbol{AB}$ 有意义时，$\boldsymbol{BA}$ 未必有意义．即使 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 都有意义，它们也可能表示不同阶数的矩阵．比如 $\boldsymbol{A}$ 是 $1\times n$ 矩阵（行向量）、$\boldsymbol{B}$ 是 $n\times 1$ 矩阵（列向量）时，$\boldsymbol{AB}$ 是 1×1 矩阵而 $\boldsymbol{BA}$ 为 $n\times n$ 矩阵．当 $\boldsymbol{A}$ ， $\boldsymbol{B}$ 都是 n 阶方阵时，情况又怎样呢？

例 2.3 设 $\boldsymbol{A} = \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}$，$\boldsymbol{B} = \begin{pmatrix} 2 & 4 \\ -3 & -6 \end{pmatrix}$，$\boldsymbol{C} = \begin{pmatrix} 8 & 8 \\ 0 & -4 \end{pmatrix}$，求 $\boldsymbol{AB}$ ， $\boldsymbol{BA}$ ， $\boldsymbol{AC}$ ．

解 利用乘积的构成规则容易得到

$$\boldsymbol{AB} = \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 2 & 4 \\ -3 & -6 \end{pmatrix} = \begin{pmatrix} -16 & -32 \\ 8 & 16 \end{pmatrix},$$

$$\boldsymbol{BA} = \begin{pmatrix} 2 & 4 \\ -3 & -6 \end{pmatrix}\begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\boldsymbol{AC} = \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 8 & 8 \\ 0 & -4 \end{pmatrix} = \begin{pmatrix} -16 & -32 \\ 8 & 16 \end{pmatrix}.$$

从例 2.3 可以看到矩阵乘法的两个重要特点：

（1）矩阵乘法不满足交换律，即一般情况下 $\boldsymbol{AB} \neq \boldsymbol{BA}$ ．

（2）矩阵乘法不满足消去律，即从 $\boldsymbol{A}\neq\boldsymbol{O}$ 和 $\boldsymbol{AB}=\boldsymbol{AC}$ 不能推得 $\boldsymbol{B}=\boldsymbol{C}$．特别地，当 $\boldsymbol{BA}=\boldsymbol{O}$ 时，不能断定 $\boldsymbol{A}=\boldsymbol{O}$ 或者 $\boldsymbol{B}=\boldsymbol{O}$．

这两个特点与数量乘法的规律不同，所以在数量运算中形成的交换与消去习惯必须改变．矩阵相乘时要注意顺序，有左乘、右乘之分，“ $\boldsymbol{A}$ 左乘 $\boldsymbol{B}$ ”和“ $\boldsymbol{A}$ 右乘 $\boldsymbol{B}$ ”是不同的运算，“ $\boldsymbol{A}$ 乘以 $\boldsymbol{B}$ ”是不明确的、无法执行的运算．

不过，矩阵的自乘无须区别左乘右乘，因此，可以引入**矩阵乘幂**的记号，比如

$$\boldsymbol{AAA}=\boldsymbol{A}^3,$$

这里 $\boldsymbol{A}$ 是 n 阶方阵．方阵的乘幂显然有下列性质

$$\boldsymbol{A}^k\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\quad (\boldsymbol{A}^k)^l=\boldsymbol{A}^{kl},$$

其中 k，l 是自然数．但是因为 $\boldsymbol{A}$，$\boldsymbol{B}$ 的乘积不能交换顺序，所以

$$(\boldsymbol{AB})^2=(\boldsymbol{AB})(\boldsymbol{AB})\neq(\boldsymbol{AA})(\boldsymbol{BB})=\boldsymbol{A}^2\boldsymbol{B}^2.$$

一般情况下，当 $k\geqslant 2$ 时，$(\boldsymbol{AB})^k\neq\boldsymbol{A}^k\boldsymbol{B}^k$．这与数量的乘幂运算规则大不相同，相应的运算习惯要改变，切不可混淆．

例 2.4　设 $\boldsymbol{A}=\begin{pmatrix}-3&2&-1\\0&3&0\\1&4&-2\end{pmatrix}$，求 $P(\boldsymbol{A})=2\boldsymbol{A}^2-3\boldsymbol{A}+4\boldsymbol{E}$．

解
$$P(\boldsymbol{A})=2\begin{pmatrix}-3&2&-1\\0&3&0\\1&4&-2\end{pmatrix}\begin{pmatrix}-3&2&-1\\0&3&0\\1&4&-2\end{pmatrix}-3\boldsymbol{A}+4\boldsymbol{E}$$
$$=2\begin{pmatrix}8&-4&5\\0&9&0\\-5&6&3\end{pmatrix}-3\begin{pmatrix}-3&2&-1\\0&3&0\\1&4&-2\end{pmatrix}+4\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}29&-14&13\\0&13&0\\-13&0&16\end{pmatrix}.$$

本例中，$P(\boldsymbol{A})$ 与多项式 $P(x)=2x^2-3x+4$ 有类似的形式，因此称它为矩阵多项式．一般地，如果一个矩阵式的每一项都是带系数的同一方阵 $\boldsymbol{A}$ 的非负整数幂，“常数项”（零次幂项）是带系数的单位矩阵，那么称这个矩阵式为关于 $\boldsymbol{A}$ 的**矩阵多项式**．

如果矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 满足 $\boldsymbol{AB}=\boldsymbol{BA}$，那么称 $\boldsymbol{A}$，$\boldsymbol{B}$ 是**可交换**的．可交换是个很强的条件，下面介绍 2 种特殊情况．

一种是对角矩阵．容易验证

$$\begin{pmatrix}a_1&0&\cdots&0\\0&a_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&a_n\end{pmatrix}\begin{pmatrix}b_1&0&\cdots&0\\0&b_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&b_n\end{pmatrix}=\begin{pmatrix}a_1b_1&0&\cdots&0\\0&a_2b_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&a_nb_n\end{pmatrix},\tag{2.4}$$

交换乘积的顺序，结果显然相同．由此可知：2 个同阶对角矩阵是可交换的，它们的乘积矩阵由对应位置元素的乘积构成．

另一种是单位矩阵．设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{E}_m$，$\boldsymbol{E}_n$ 分别为 m 阶、n 阶单位矩阵，不难验证 $\boldsymbol{E}_m\boldsymbol{A}=\boldsymbol{A}$，$\boldsymbol{A}\boldsymbol{E}_n=\boldsymbol{A}$．特别地，当 $m=n$ 时

$$\boldsymbol{E}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{E}=\boldsymbol{A}\text{．}\tag{2.5}$$

可见单位矩阵 $\boldsymbol{E}$ 在矩阵乘法中，与数 1 在数量乘法中有类似的作用．单位矩阵与任何同阶矩阵可交换．

矩阵的乘法虽然不满足交换律，但仍满足下列运算规律（假设运算都是可行的）：

（1）乘法结合律　$(\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$；

（2）左、右分配律　$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{A}\boldsymbol{C}+\boldsymbol{B}\boldsymbol{C}$，$\boldsymbol{C}(\boldsymbol{A}+\boldsymbol{B})=\boldsymbol{C}\boldsymbol{A}+\boldsymbol{C}\boldsymbol{B}$；

（3）数乘结合律　$k(\boldsymbol{A}\boldsymbol{B})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$．

这些运算律的证明，都可以利用乘法公式（2.1）以及通过和式的乘积展开与重组来完成，此处从略．这些运算律与数量的运算规律相同，所以在数量运算中形成的诸如多项乘积展开、系数归并化简、因式分解、连乘重组等运算习惯，在矩阵的运算中，仍可保留沿用．当然应该特别注意不可随意交换乘法顺序，不可随意约简非零因子．

2.2.3 矩阵的转置

把矩阵 $\boldsymbol{A}$ 的行与列互换所得到的矩阵称为矩阵 $\boldsymbol{A}$ 的**转置矩阵**，记为 $\boldsymbol{A}^{\mathrm{T}}$，即

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{m1} & a_{m2} & \mathrm{L} & a_{mn} \end{pmatrix},\quad \boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} a_{11} & a_{21} & \mathrm{L} & a_{m1} \\ a_{12} & a_{22} & \mathrm{L} & a_{m2} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{1n} & a_{2n} & \mathrm{L} & a_{mn} \end{pmatrix}\text{．}$$

矩阵的转置方法与行列式相类似，可以“横着看，竖着写”．但是矩阵转置后，行、列数都变了，各元素的位置也变了，所以通常 $\boldsymbol{A}\neq\boldsymbol{A}^{\mathrm{T}}$．

转置矩阵有如下性质（其中 $\boldsymbol{A}$，$\boldsymbol{B}$ 是矩阵，k 是数，涉及的运算都可行）：

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

（2）$(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$；

（3）$(k\boldsymbol{A})^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}$；

（4）$(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$．

这里性质（1）～（3）是显然的，性质（4）可利用乘法公式（2.1）证明．

例 2.5　设 $\boldsymbol{A}=\begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 2 \end{pmatrix}$，计算 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$．

解　$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 3 \\ -1 & 2 \end{pmatrix}=\begin{pmatrix} 5 & 0 \\ 0 & 14 \end{pmatrix}$，

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}2&1\\0&3\\-1&2\end{pmatrix}\begin{pmatrix}2&0&-1\\1&3&2\end{pmatrix}=\begin{pmatrix}5&3&0\\3&9&6\\0&6&5\end{pmatrix}.$$

若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$，则称 $\boldsymbol{A}$ 为**对称矩阵**．比如例 2.5 所求的两个矩阵都是对称矩阵．

2.2.4　方阵行列式的乘积定理

设 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是 n 阶方阵．一般地 $\boldsymbol{AB}\neq\boldsymbol{BA}$，但它们的行列式相等，并且

$$|\boldsymbol{AB}|=|\boldsymbol{BA}|=|\boldsymbol{A}|\cdot|\boldsymbol{B}|\ .\tag{2.6}$$

定理 2.1　方阵乘积的行列式等于各因子行列式的乘积．

这个定理的结论简明、自然，但它的证明很复杂，而且需要用到特殊的构造性技巧，此处从略．

2.3　逆 矩 阵

2.3.1　逆矩阵的概念

设 $\boldsymbol{A}$ 是 n 阶矩阵（方阵），如果存在 n 阶矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E}$，则称矩阵 $\boldsymbol{A}$ 是**可逆的**，并称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的**逆矩阵**．

矩阵 $\boldsymbol{A}$ 可逆时，逆矩阵 $\boldsymbol{B}$ 必惟一．事实上，若另有一逆矩阵 $\boldsymbol{B}_1$，则由 $\boldsymbol{AB}=\boldsymbol{E}$ 和 $\boldsymbol{B}_1\boldsymbol{A}=\boldsymbol{E}$ 得到 $\boldsymbol{B}_1=\boldsymbol{B}_1\boldsymbol{E}=\boldsymbol{B}_1(\boldsymbol{AB})=(\boldsymbol{B}_1\boldsymbol{A})\boldsymbol{B}=\boldsymbol{EB}=\boldsymbol{B}$．这样，逆矩阵可以有惟一的记号．记 $\boldsymbol{A}$ 的逆矩阵为 $\boldsymbol{A}^{-1}$，即

$$\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{E}\ .\tag{2.7}$$

比如不难验证

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

逆矩阵相当于矩阵的“倒数”，但是因为矩阵的乘法有左乘、右乘之分，所以不允许以分数线表示逆矩阵．

如果 3 个矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 满足 $\boldsymbol{AB}=\boldsymbol{AC}$，且 $\boldsymbol{A}$ 可逆，那么在等式两边左乘逆矩阵 $\boldsymbol{A}^{-1}$，可得 $\boldsymbol{A}^{-1}\boldsymbol{AB}=\boldsymbol{A}^{-1}\boldsymbol{AC}$，即 $\boldsymbol{EB}=\boldsymbol{EC}$，从而 $\boldsymbol{B}=\boldsymbol{C}$．这说明利用逆矩阵可以实现“约简”，换言之，矩阵的乘法并非没有消去规则，但消去规则必须通过逆矩阵的乘法来实现，可逆才有消去律．当然，在等式两边乘逆矩阵时应当注意分清左乘还是右乘．

逆矩阵为求解矩阵方程带来了方便．比如线性方程组 $Ax=b$ 中，若 A 可逆，则左乘 A^{-1} 后，即得 $x=A^{-1}b$，事先求出逆矩阵 A^{-1}，只要做一次乘法，便可求得所有变量（x 的分量）的值．又如矩阵方程 $AXB=C$ 中，若 A，B 均可逆，则未知矩阵直接可求，即 $X=A^{-1}CB^{-1}$．

2.3.2 矩阵可逆的条件

设有 n 阶方阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & \mathrm{L} & a_{nn} \end{pmatrix}.$$

它的行列式 $|A|$ 有 n^2 个代数余子式 A_{ij}（$i,j=1,2,\cdots,n$），将它们按转置排列，得到矩阵

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

称 A^{*} 为矩阵 A 的**伴随矩阵**．利用定理 1.2（代数余子式组合定理）容易验证

$$\boldsymbol{A}\boldsymbol{A}^{*}=\boldsymbol{A}^{*}\boldsymbol{A}=\begin{pmatrix} |\boldsymbol{A}| & 0 & \mathrm{L} & 0 \\ 0 & |\boldsymbol{A}| & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & \mathrm{L} & |\boldsymbol{A}| \end{pmatrix}=|\boldsymbol{A}|\boldsymbol{E}.$$

如果 $|A|\neq 0$，则上式两端除以非零数 $|A|$，可得

$$\boldsymbol{A}\left(\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}\right)=\left(\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}\right)\boldsymbol{A}=\boldsymbol{E}.$$

这说明矩阵 A 可逆，并且

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}. \tag{2.8}$$

定理 2.2 方阵 A 可逆的充分必要条件是它的行列式不等于零，即 $|A|\neq 0$．

证 （2.8）式已给出充分性证明，现证必要性．如果矩阵 A 可逆，则由 $AA^{-1}=E$ 取行列式，根据定理 2.1 的（2.6）式得 $|\boldsymbol{A}||\boldsymbol{A}^{-1}|=|\boldsymbol{A}\boldsymbol{A}^{-1}|=|\boldsymbol{E}|=1$，因而必有 $|A|\neq 0$．

行列式非零的方阵又叫做**非奇异矩阵**．显然，非奇异矩阵和可逆矩阵是等价的概念．行列式等于零的矩阵自然叫做**奇异矩阵**．奇异矩阵即不可逆矩阵有无数多个，这与数量中唯有数 0 没有倒数大不相同．

例 2.6　设 $\boldsymbol{A}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$，求 $\boldsymbol{A}^{-1}$．

解　显然 $|\boldsymbol{A}|=-2\neq 0$，$|\boldsymbol{A}|$ 的代数余子式都是一阶行列式，不需要计算，只要附上适当的符号，并注意转置排列即可：

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^{*}=\frac{1}{-2}\begin{pmatrix}4 & -2\\ -3 & 1\end{pmatrix}=\begin{pmatrix}-2 & 1\\ 1.5 & -0.5\end{pmatrix}.$$

公式（2.8）给出了求逆矩阵的方法，但是求伴随矩阵 $\boldsymbol{A}^{*}$ 要计算 n^2 个 $(n-1)$ 阶行列式，当 n 较大时，计算量非常大．我们将在下一节介绍更好的方法．

定理 2.3　设 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是 n 阶矩阵，则 $\boldsymbol{B}=\boldsymbol{A}^{-1}$ 的充分必要条件是 $\boldsymbol{AB}=\boldsymbol{E}$ 或者 $\boldsymbol{BA}=\boldsymbol{E}$．

证　必要性显然，只证充分性．若 $\boldsymbol{AB}=\boldsymbol{E}$，取行列式得 $|\boldsymbol{A}||\boldsymbol{B}|=1$，故 $|\boldsymbol{A}|\neq 0$，则根据定理 2.2，$\boldsymbol{A}^{-1}$ 存在．等式两端左乘 $\boldsymbol{A}^{-1}$，立得 $\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{AB}=\boldsymbol{A}^{-1}\boldsymbol{E}=\boldsymbol{A}^{-1}$．$\boldsymbol{BA}=\boldsymbol{E}$ 的情况相同，证毕．

定理 2.3 表明，检验或者证明 $\boldsymbol{B}$ 是否 $\boldsymbol{A}$ 的逆矩阵，只要做一个乘法即可．比如从公式（2.4）很容易求得对角矩阵的逆矩阵：

$$\begin{pmatrix}a_1 & 0 & \mathrm{L} & 0\\ 0 & a_2 & \mathrm{L} & 0\\ \mathrm{M} & \mathrm{M} & & \mathrm{M}\\ 0 & 0 & \mathrm{L} & a_n\end{pmatrix}^{-1}=\begin{pmatrix}1/a_1 & 0 & \mathrm{L} & 0\\ 0 & 1/a_2 & \mathrm{L} & 0\\ \mathrm{M} & \mathrm{M} & & \mathrm{M}\\ 0 & 0 & \mathrm{L} & 1/a_n\end{pmatrix},\tag{2.9}$$

其中 $a_1a_2\mathrm{L}\ a_n\neq 0$．

2.3.3　逆矩阵的性质

（1）若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{-1}$ 也可逆，且 $(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$．

证　根据定理（2.3），只需做一个乘法：因为 $\boldsymbol{AA}^{-1}=\boldsymbol{E}$，故得证．

（2）若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{\mathrm{T}}$ 也可逆，且 $(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$．

证　因为 $\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{-1}\boldsymbol{A})^{\mathrm{T}}=\boldsymbol{E}^{\mathrm{T}}=\boldsymbol{E}$，故得证．

（3）若 $\boldsymbol{A}$，$\boldsymbol{B}$ 是同阶矩阵且都可逆，则 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$．

证　因为 $(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1})=\boldsymbol{A}(\boldsymbol{BB}^{-1})\boldsymbol{A}^{-1}=\boldsymbol{AEA}^{-1}=\boldsymbol{AA}^{-1}=\boldsymbol{E}$，故得证．

2.4　矩阵的初等变换

2.4.1　矩阵的初等行变换

在第 1 章中，我们已经看到了行（列）变换在行列式计算中的重要作用．对矩阵也有

类似的变换.

对矩阵施行下列 3 种变换，统称为矩阵的**初等行变换**：

(1) **换行变换** 将矩阵的两行互换位置；

(2) **倍缩变换** 以非零数 k 乘矩阵某一行的所有元素；

(3) **消去变换** 把矩阵某一行所有元素乘同一数 k 加到另一行对应的元素上去.

例如对下列矩阵作初等行变换：先将第 3 行乘 -2 加到第 1 行，再将第 1，3 行互换，得到

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 5 \end{pmatrix} \to \begin{pmatrix} 0 & -1 & -9 \\ 0 & 1 & 3 \\ 1 & 2 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 5 \\ 0 & 1 & 3 \\ 0 & -1 & -9 \end{pmatrix}.$$

由于矩阵的初等变换改变了矩阵的元素，因此初等变换前后的矩阵是不相等的，应该用“ $\to$ ”连接而不可用“ = ”连接. 矩阵的初等变换可以链锁式地反复进行，以便达到简化矩阵的目的.

类似地，可引入**初等列变换**的概念.

2.4.2 初等变换的标准程序

例 2.7 已知 $\boldsymbol{A}=\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 3 \\ 1 & 2 & 4 \end{pmatrix}$，求 $\boldsymbol{A}^{-1}$.

解 将矩阵 $\boldsymbol{A}$ 和单位矩阵 $\boldsymbol{E}$ 拼成一个 3×6 矩阵 $(\boldsymbol{A}\quad \boldsymbol{E})$. 类似于行列式的消元运算（参看 1.2 节），对 $(\boldsymbol{A}\quad \boldsymbol{E})$ 施行一系列初等行变换：

$$(\boldsymbol{A}\quad \boldsymbol{E})=\begin{pmatrix} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ \langle 1\rangle & 2 & 4 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & -1 & -7 & 1 & 0 & -2 \\ 0 & \langle 1\rangle & 3 & 0 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{*} \begin{pmatrix} 0 & 0 & \langle -4\rangle & 1 & 1 & -2 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 1 & 0 & -2 & 0 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & -0.25 & -0.25 & 0.5 \\ 0 & 1 & 0 & 0.75 & 1.75 & -1.5 \\ 1 & 0 & 0 & -0.5 & -2.5 & 2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 0 & -0.5 & -2.5 & 2 \\ 0 & 1 & 0 & 0.75 & 1.75 & -1.5 \\ 0 & 0 & 1 & -0.25 & -0.25 & 0.5 \end{pmatrix}.$$

可以验证，最后的矩阵中，右半部分就是逆矩阵，即

$$\boldsymbol{A}^{-1}=\begin{pmatrix}-0.5 & -2.5 & 2\\ 0.75 & 1.75 & -1.5\\ -0.25 & -0.25 & 0.5\end{pmatrix}=\frac{1}{4}\begin{pmatrix}-2 & -10 & 8\\ 3 & 7 & -6\\ -1 & -1 & 2\end{pmatrix}.$$

本例的结果不是偶然的．在论证这一方法的原理之前，我们先结合例 2.7 介绍**矩阵初等变换的标准程序**：

（1）变换分步进行，每步选一个非零元素，称为**主元**．利用行倍缩变换把主元变为 1，并且通过行消去变换把主元所在列的其他元素全都变为 0．

（2）各步所选的主元，必须位于不同的行．逐步重复上述变换，直至选不出新的主元为止．

（3）穿插换行变换，使主元呈左上到右下排列．

简单地说，标准程序就是通过初等行变换（不允许做列变换），变出一个一个不同的基本单位列，直至变不出新的基本单位列为止．**基本单位列**是指一个元素为 1，其余元素全都为 0 的列向量．

比如在例 2.7 的运算中，带“ * ”号的第 2 步是以元素 1 为主元，将第 2 行乘 1 和 –2 分别加到第 1，3 行上去；最后一步（第 4 步）并未选主元，而是作了一个互换第 1，3 行的换行变换．在所有的行消去变换中，主元都用“〈 〉”号做了标记，这样可使消元的目标清晰，变换的过程简洁．

标准程序体现了初等变换的目的性和条理性．矩阵的初等变换将贯穿本书的始终，初等变换的标准程序也将反复多次得到应用．

2.4.3　用初等变换法求逆矩阵

设 $\boldsymbol{A}$ 是 n 阶矩阵，$\boldsymbol{E}$ 是 n 阶单位矩阵，对 $n\times 2n$ 矩阵 $(\boldsymbol{A}\quad \boldsymbol{E})$ 按标准程序作初等行变换，主元在左半部分（即前 n 列）的范围内选取．当把子块 $\boldsymbol{A}$ 变成单位矩阵 $\boldsymbol{E}$ 的同时，右半部分必然变成了 $\boldsymbol{A}^{-1}$（参看例 2.7）．

例 2.8　设 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}$，问 $\boldsymbol{A}^{-1}$ 是否存在？

解　运用初等变换法：

$$(\boldsymbol{A}\,\vdots\,\boldsymbol{E})=\left(\begin{array}{ccc:ccc}\langle 1\rangle & 2 & 3 & 1 & 0 & 0\\ 4 & 5 & 6 & 0 & 1 & 0\\ 7 & 8 & 9 & 0 & 0 & 1\end{array}\right)\to\left(\begin{array}{ccc:ccc}1 & 2 & 3 & 1 & 0 & 0\\ 0 & \langle -3\rangle & -6 & -4 & 1 & 0\\ 0 & -6 & -12 & -7 & 0 & 1\end{array}\right)$$

$$\rightarrow \begin{pmatrix} 1 & 0 & -1 & -\frac{5}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 2 & \frac{4}{3} & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

标准程序已执行完毕，但子块 A 未变成单位矩阵，即在（前 3 列）未选主元的行中没有非零元素，无法选出新的主元，此时可以断定 A 不可逆．其理由如下：

设想对行列式 $|A|$ 施行初等变换．如果将换行变换、倍缩变换或消去变换施加于行列式，则行列式的值仅仅是改变符号、非零倍缩或保持不变，总之，初等变换不改变行列式的非零性，因此能通过初等变换检验矩阵的可逆性（参看定理 2.2）．

例 2.8 说明用初等变换法求逆矩阵，不必事先知道矩阵是否可逆．

2.4.4 初等矩阵

对单位矩阵 E 施行一次初等变换得到的矩阵称为**初等矩阵**．例如下面 3 个矩阵：

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

都是初等矩阵．与它们相对应的初等行变换分别是“互换第 1、第 2 行”、“以 3 乘第 2 行”、“第 1 行乘 2 加到第 3 行”；相对应的初等列变换分别是“互换第 1、第 2 列”、“以 3 乘第 2 列”、“第 3 列乘 2 加到第 1 列”．以下定理不难验证：

定理 2.4 对矩阵 A 施行初等行（列）变换，相当于左（右）乘一个相应的初等矩阵．

现在对初等变换法求逆矩阵的有效性进行证明．

对 $n\times 2n$ 矩阵 $(A \quad E)$ 作一系列初等行变换，相当于左乘一系列 n 阶初等矩阵 G_1，G_2，…，G_k（定理 2.4）．最后得到的矩阵是 $(E \mid B)$，即

$$G_k \cdots G_2 G_1 (A \mid E) = (E \mid B).$$

需要证明 $B = A^{-1}$．令 $G = G_k \cdots G_2 G_1$，G 仍是 n 阶矩阵．按矩阵乘法规则，$G(A \mid E)$ 的前 n 列由 GA 给出（A 的各列与 G 的各行相组合），而后 n 列由 $GE = G$ 给出，因此

$$G(A \mid E) = (GA \mid G) = (E \mid B).$$

由于矩阵相等意味着对应元素相等，所以应有

$$GA = E, \quad G = B.$$

根据定理 2.3 知 $G = A^{-1}$，从而 $B = A^{-1}$．证毕．

对矩阵$(A\,|\,E)$作初等变换，实际上初等变换的“对象”是左半部分，而右半部分则起到了“记录”这些变换过程的作用，A^{-1}是所有这些初等变换的总结果．由此可知，一系列初等变换（初等矩阵的乘积）构成可逆矩阵；反之，可逆矩阵能分解为一系列初等变换（初等矩阵的乘积）.

2.5　分 块 矩 阵

2.5.1　分块矩阵的概念

上节中曾把矩阵A和E拼成一个大矩阵$(A\quad E)$，反过来可以看作把矩阵$(A\quad E)$划分成两部分，这就是分块矩阵的概念．对于一个矩阵，可以根据需要用贯穿整个矩阵的横线或竖线把它划分成若干子块（或称子矩阵），便形成了**分块矩阵**．分块的方式有许多，例如对于矩阵

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

可例举出以下 3 种不同的分块方式：

$$\left(\begin{array}{cc:cc}1&0&0&3\\0&1&0&-1\\\hdashline 0&0&1&0\\0&0&0&1\end{array}\right),\quad \left(\begin{array}{ccc:c}1&0&0&3\\0&1&0&-1\\0&0&1&0\\\hdashline 0&0&0&1\end{array}\right),\quad \left(\begin{array}{c:c:c}1&0&0&3\\\hdashline 0&1&0&-1\\\hdashline 0&0&1&0\\\hdashline 0&0&0&1\end{array}\right).$$

在第 1 个分块矩阵中，左上角和右下角子块都是二阶单位矩阵$\boldsymbol{E}_2$，左下角子块是零矩阵，若记$\boldsymbol{A}_1=\begin{pmatrix}0&3\\0&-1\end{pmatrix}$，则$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{E}_2&\boldsymbol{A}_1\\\boldsymbol{O}&\boldsymbol{E}_2\end{pmatrix}$.

在第 2 个分块矩阵中，左上角子块和右下角子块分别是三阶单位矩阵$\boldsymbol{E}_3$和一阶单位矩阵$\boldsymbol{E}_1$，左下角子块是零矩阵，若记$\boldsymbol{A}_2=\begin{pmatrix}3\\-1\\0\end{pmatrix}$，则$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{E}_3&\boldsymbol{A}_2\\\boldsymbol{O}&\boldsymbol{E}_1\end{pmatrix}$.

在第 3 个分块矩阵中，若记

$$\boldsymbol{e}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\quad \boldsymbol{e}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\quad \boldsymbol{e}_3=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\quad \boldsymbol{u}=\begin{pmatrix}3\\-1\\0\\1\end{pmatrix},$$

则 $\boldsymbol{A}=(\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \boldsymbol{e}_3,\ \boldsymbol{u})$，这里 $\boldsymbol{e}_1$，$\boldsymbol{e}_2$，$\boldsymbol{e}_3$ 都是一个分量为 1、其余分量为 0 的列向量，这种向量称为**基本单位向量**.

对矩阵 $\boldsymbol{A}$ 的分块，当然不止这 3 种方式．分块矩阵也可以理解为是以若干子块为元素组成的矩阵.

2.5.2 分块矩阵的运算

上节得到过简单分块矩阵的乘法规则 $\boldsymbol{G}(\boldsymbol{A}\,|\,\boldsymbol{E})=(\boldsymbol{GA}\,|\,\boldsymbol{GE})$，相当于把 $\boldsymbol{G}$，$\boldsymbol{A}$，$\boldsymbol{E}$ 都当作数量看待．对于一般的分块矩阵有同样的结论：分块矩阵运算时，可以把子块当作数量元素处理．例如

$$\begin{pmatrix}\boldsymbol{X}_{11} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22}\end{pmatrix}+\begin{pmatrix}\boldsymbol{Y}_{11} & \boldsymbol{Y}_{12}\\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22}\end{pmatrix}=\begin{pmatrix}\boldsymbol{X}_{11}+\boldsymbol{Y}_{11} & \boldsymbol{X}_{12}+\boldsymbol{Y}_{12}\\ \boldsymbol{X}_{21}+\boldsymbol{Y}_{21} & \boldsymbol{X}_{22}+\boldsymbol{Y}_{22}\end{pmatrix},$$

$$\begin{pmatrix}\boldsymbol{X}_{11} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{Y}_1\\ \boldsymbol{Y}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{X}_{11}\boldsymbol{Y}_1+\boldsymbol{X}_{12}\boldsymbol{Y}_2\\ \boldsymbol{X}_{21}\boldsymbol{Y}_1+\boldsymbol{X}_{22}\boldsymbol{Y}_2\end{pmatrix}.$$

不过在把子块当作元素对待时，还须遵循以下规则.

（1）矩阵的分块方式要与运算相配套.

具体而言：两个矩阵相加时，它们的行、列划分方式应完全相同，以保证相加的子块有同样的行、列数；两个矩阵相乘时，左矩阵列的划分与右矩阵行的划分方式应一致，以保证相乘子块的行、列数相匹配.

（2）对子块的乘积要分清左、右顺序，不能随意交换.

比如上面运算中的 $\boldsymbol{X}_{11}\boldsymbol{Y}_1$ 不能写作 $\boldsymbol{Y}_1\boldsymbol{X}_{11}$.

（3）分块矩阵转置时，除子块的位置转置外，子块本身也要转置．比如

$$(\boldsymbol{A}_1\,|\,\boldsymbol{A}_2)^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{A}_1^{\mathrm{T}}\\ \boldsymbol{A}_2^{\mathrm{T}}\end{pmatrix}.$$

例 2.9 设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\\ \boldsymbol{O} & \boldsymbol{A}_3\end{pmatrix}$，其中 $\boldsymbol{A}_1$，$\boldsymbol{A}_3$ 是可逆矩阵．试用 $\boldsymbol{A}_1$，$\boldsymbol{A}_2$，$\boldsymbol{A}_3$ 及其运算式构造出逆矩阵 $\boldsymbol{A}^{-1}$.

解 将 $\boldsymbol{A}^{-1}$ 分块，令 $\boldsymbol{A}^{-1}=\begin{pmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2\\ \boldsymbol{X}_3 & \boldsymbol{X}_4\end{pmatrix}$，则应有 $\begin{pmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_2\\ \boldsymbol{O} & \boldsymbol{A}_3\end{pmatrix}\begin{pmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2\\ \boldsymbol{X}_3 & \boldsymbol{X}_4\end{pmatrix}=\begin{pmatrix}\boldsymbol{E} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{E}\end{pmatrix}$，等式右

端是分块的单位矩阵．矩阵相等意味着所有对应子块都相同，即

$$\boldsymbol{A}_1\boldsymbol{X}_1+\boldsymbol{A}_2\boldsymbol{X}_3=\boldsymbol{E}\ ,\ \boldsymbol{O}+\boldsymbol{A}_3\boldsymbol{X}_3=\boldsymbol{O}\ ,$$

$$\boldsymbol{A}_1\boldsymbol{X}_2+\boldsymbol{A}_2\boldsymbol{X}_4=\boldsymbol{O}\ ,\ \boldsymbol{O}+\boldsymbol{A}_3\boldsymbol{X}_4=\boldsymbol{E}\ .$$

因为 $\boldsymbol{A}_1^{-1}$，$\boldsymbol{A}_3^{-1}$ 存在，所以可逐个解得 $\boldsymbol{X}_3=\boldsymbol{O}$，$\boldsymbol{X}_4=\boldsymbol{A}_3^{-1}$，$\boldsymbol{X}_1=\boldsymbol{A}_1^{-1}$，$\boldsymbol{X}_2=-\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{A}_3^{-1}$，于是

$$\boldsymbol{A}^{-1}=\begin{pmatrix}\boldsymbol{A}_1^{-1} & -\boldsymbol{A}_1^{-1}\boldsymbol{A}_2\boldsymbol{A}_3^{-1}\\ \boldsymbol{O} & \boldsymbol{A}_3^{-1}\end{pmatrix}.$$

习 题 二

1．设 $\boldsymbol{A}=\begin{pmatrix}1&2&1&2\\2&1&2&1\\1&2&3&4\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}4&3&2&1\\-2&1&-2&1\\0&-1&0&-1\end{pmatrix}$，

（1）求 $3\boldsymbol{A}+2\boldsymbol{B}$；　（2）已知 $3(\boldsymbol{A}+\boldsymbol{X})-2(\boldsymbol{B}-\boldsymbol{X})=\boldsymbol{O}$，求 $\boldsymbol{X}$．

2．已知 $\boldsymbol{A}$ 是 5 阶矩阵，k 是常数，问下列哪个等式是正确的？

（1）$|k\boldsymbol{A}|=k|\boldsymbol{A}|$；　（2）$|k\boldsymbol{A}|=|k||\boldsymbol{A}|$；

（3）$|k\boldsymbol{A}|=k^5|\boldsymbol{A}|$；　（4）$|k\boldsymbol{A}|=|k|^5|\boldsymbol{A}|$．

3．计算下列矩阵的乘积．

（1）$(2,\ 3,\ -1)\begin{pmatrix}1\\-1\\-1\end{pmatrix}$；　（2）$\begin{pmatrix}1\\-1\\-1\end{pmatrix}(2,\ 3,\ -1)$；

（3）$\begin{pmatrix}4&3&2\\1&-2&3\\5&7&-1\end{pmatrix}\begin{pmatrix}2\\3\\-1\end{pmatrix}$；　（4）$\begin{pmatrix}1&-3&2\\3&-4&1\\2&-5&3\end{pmatrix}\begin{pmatrix}2&5&6\\1&2&5\\1&3&2\end{pmatrix}$；

（5）$\begin{pmatrix}1&0&3&1\\0&1&2&-1\\0&0&-2&3\\0&0&0&-3\end{pmatrix}\begin{pmatrix}2&1&-3&2\\0&2&1&-1\\0&0&3&1\\0&0&0&3\end{pmatrix}$；　（6）$\begin{pmatrix}2&1&4&0\\1&-1&3&4\end{pmatrix}\begin{pmatrix}1&3&1\\0&-1&2\\1&-3&1\\4&0&-2\end{pmatrix}$．

4．计算矩阵多项式 $f(\boldsymbol{A})=3\boldsymbol{A}^2-2\boldsymbol{A}+5\boldsymbol{E}$，其中 $\boldsymbol{A}=\begin{pmatrix}1&-2&3\\2&-4&1\\3&-5&2\end{pmatrix}$．

5．设 $\boldsymbol{A}=\begin{pmatrix}1&0&0&0\\0&-2&0&0\\0&0&3&0\\0&0&0&4\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}3&0&0&0\\0&-5&0&0\\0&0&6&0\\0&0&0&-2\end{pmatrix}$，计算 $\boldsymbol{A}^3\boldsymbol{B}$，$\boldsymbol{B}\boldsymbol{A}^3$，$\boldsymbol{A}^2\boldsymbol{B}\boldsymbol{A}$，$\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^2$．

6．试将两个线性变换方程组 $\begin{cases} y_1 = 2x_1 + x_2 \\ y_2 = 3x_1 - x_2 + 2x_3 \\ y_3 = 4x_1 + 2x_2 - x_3 \end{cases}$ 和 $\begin{cases} z_1 = y_1 + 2y_2 - y_3 \\ z_2 = 2y_1 - 2y_2 + 2y_3 \\ z_3 = y_1 - 3y_2 - 4y_3 \end{cases}$ 写成矩阵形式，并由此求出 z_1, z_2, z_3 与 x_1, x_2, x_3 之间的线性变换关系式．

7．设 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 阶矩阵．举例说明以下问题．

（1） $\boldsymbol{A} \neq \boldsymbol{O}, \boldsymbol{B} \neq \boldsymbol{O}$，但可能有 $\boldsymbol{AB} = \boldsymbol{O}$；

（2）如果 $\boldsymbol{A} \neq \boldsymbol{O}$ 且 $\boldsymbol{A}^2 = \boldsymbol{A}$，则 $\boldsymbol{A}(\boldsymbol{A} - \boldsymbol{E}) = \boldsymbol{O}$，但未必有 $\boldsymbol{A} = \boldsymbol{E}$；

（3） $\boldsymbol{AX} = \boldsymbol{AY}$ 且 $\boldsymbol{A} \neq \boldsymbol{O}$ 时，不一定有 $\boldsymbol{X} = \boldsymbol{Y}$；

（4） $(\boldsymbol{A} + \boldsymbol{B})^2 = \boldsymbol{A}^2 + 2\boldsymbol{AB} + \boldsymbol{B}^2$ 和 $(\boldsymbol{A} + \boldsymbol{B})(\boldsymbol{A} - \boldsymbol{B}) = \boldsymbol{A}^2 - \boldsymbol{B}^2$ 是否一定成立?

8．设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶方阵，且 $2\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{E}$（$\boldsymbol{E}$ 是单位矩阵）．证明：$\boldsymbol{A}^2 = \boldsymbol{A}$ 的充分必要条件是 $\boldsymbol{B}^2 = \boldsymbol{E}$．

9．设 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，证明 $\boldsymbol{AA}^{\mathrm{T}}$ 和 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 都是对称矩阵．

10．设 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 1 & 2 \end{pmatrix}$，$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$．试通过计算验证 $\boldsymbol{AB}^{\mathrm{T}} = (\boldsymbol{BA}^{\mathrm{T}})^{\mathrm{T}}$．

11．设 $\boldsymbol{A} = \begin{pmatrix} 3 & 0 \\ 1 & 3 \\ 2 & 1 \\ -1 & 0 \end{pmatrix}$，$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \\ 0 & 3 & 0 \end{pmatrix}$，$\boldsymbol{C} = \begin{pmatrix} -1 & 0 \\ 1 & 5 \\ 0 & 2 \end{pmatrix}$，问乘积 $\boldsymbol{ABC}$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{BC}$，$\boldsymbol{AB}^{\mathrm{T}}\boldsymbol{C}$，$\boldsymbol{ABC}^{\mathrm{T}}$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{BC}^{\mathrm{T}}$，$\boldsymbol{AB}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}$，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}$ 哪些是可行的？求出这些可行的乘积结果．

12．利用伴随矩阵，求下列矩阵的逆矩阵：

（1） $\begin{pmatrix} 3 & 5 \\ 5 & 9 \end{pmatrix}$；　（2） $\begin{pmatrix} 1 & -2 \\ -4 & 3 \end{pmatrix}$；　（3） $\begin{pmatrix} 2a+b & 3a-b \\ b-2a & a+2b \end{pmatrix}$．

13．设 $\boldsymbol{A}$ 是 n 阶矩阵，且 $\boldsymbol{A}^2 - 2\boldsymbol{A} - 4\boldsymbol{E} = \boldsymbol{O}$（$\boldsymbol{E}$ 是单位矩阵）．证明：矩阵 $\boldsymbol{A} + \boldsymbol{E}$ 可逆，且 $(\boldsymbol{A} + \boldsymbol{E})^{-1} = \boldsymbol{A} - 3\boldsymbol{E}$．

14．已知 $\boldsymbol{A}$ 是可逆的对称矩阵，证明：

（1） $\boldsymbol{A}^{-1}$ 也是对称矩阵；　（2） $\boldsymbol{X}$ 与 $\boldsymbol{A}$ 是同阶矩阵时，$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{AX}$ 是对称矩阵．

15．已知 $\boldsymbol{A}^2 = \boldsymbol{A}$，试用反证法证明：若 $\boldsymbol{A}$ 不是单位矩阵，则 $\boldsymbol{A}$ 必为奇异矩阵．

16．已知 $\boldsymbol{A}, \boldsymbol{B}$ 都是 n 阶非奇异矩阵，证明：$\boldsymbol{A}, \boldsymbol{B}$ 可交换的充分必要条件是 $(\boldsymbol{AB})^2 = \boldsymbol{A}^2\boldsymbol{B}^2$．

17．求下列矩阵的逆矩阵：

（1） $\begin{pmatrix} 2 & 2 & -1 \\ 1 & -2 & 4 \\ 5 & 8 & 2 \end{pmatrix}$；　（2） $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 4 & 3 \end{pmatrix}$；　（3） $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$；

（4） $\begin{pmatrix} 1 & 3 & -5 & 7 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$；　（5） $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$；　（6） $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 7 & 2 & 3 \\ 2 & 5 & 1 & 2 \end{pmatrix}$．

18．设 $\boldsymbol{A}=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&3&0&1\end{pmatrix}$，$\boldsymbol{C}=\begin{pmatrix}2&-3&1&1\\1&4&-1&2\\2&-1&3&2\\-3&2&1&5\end{pmatrix}$，利用矩阵的初等变换计算乘积 $\boldsymbol{AC}$，$\boldsymbol{CA}$，$\boldsymbol{BC}$，$\boldsymbol{CB}$．

19．设 $\boldsymbol{X}=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{A}\\\boldsymbol{B}&\boldsymbol{O}\end{pmatrix}$，其中 $\boldsymbol{A}=\begin{pmatrix}2&1\\5&3\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}4&-5\\-2&3\end{pmatrix}$，$\boldsymbol{O}$ 是零矩阵．利用分块矩阵的乘法规则求 $\boldsymbol{X}^{-1}$．

20．单项选择题

（1）已知 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是 n 阶方阵，则必有　（　）

A．$|\boldsymbol{AB}|=|\boldsymbol{BA}|$；　B．$\boldsymbol{AB}=\boldsymbol{BA}$；　C．$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}=(\boldsymbol{AB})^{\mathrm{T}}$；　D．$(\boldsymbol{AB})^2=\boldsymbol{A}^2\boldsymbol{B}^2$．

（2）已知 $(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2$，则矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 必定满足　（　）

A．$\boldsymbol{A}=\boldsymbol{B}$；　B．$\boldsymbol{AB}=\boldsymbol{BA}$；　C．$\boldsymbol{AB}$ 是对称矩阵；　D．$\boldsymbol{A}$，$\boldsymbol{B}$ 都是对角矩阵．

（3）设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是同阶的非零矩阵，则 $\boldsymbol{AB}=\boldsymbol{AC}$ 是 $\boldsymbol{B}=\boldsymbol{C}$ 的　（　）

A．充分非必要条件；　B．必要非充分条件；

C．充分必要条件；　D．非充分非必要条件．

（4）矩阵 $\boldsymbol{A}=\begin{pmatrix}5&6\\7&8\end{pmatrix}$ 的伴随矩阵 $\boldsymbol{A}^*=$　（　）

A．$\begin{pmatrix}-8&7\\6&-5\end{pmatrix}$；　B．$\begin{pmatrix}5&-6\\-7&8\end{pmatrix}$；　C．$\begin{pmatrix}8&-6\\-7&5\end{pmatrix}$；　D．$\begin{pmatrix}-5&7\\6&-8\end{pmatrix}$．

（5）设 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是可逆矩阵，且 $\boldsymbol{AC}=\boldsymbol{B}$，则 $\boldsymbol{C}=$　（　）

A．$\boldsymbol{A}^{-1}\boldsymbol{B}$；　B．$\boldsymbol{AB}^{-1}$；　C．$\boldsymbol{BA}^{-1}$；　D．$\boldsymbol{B}^{-1}\boldsymbol{A}$．

（6）若 $\boldsymbol{A}^2=\boldsymbol{O}$，$\boldsymbol{E}$ 是单位矩阵，则 $(\boldsymbol{E}-\boldsymbol{A})^{-1}=$　（　）

A．$\boldsymbol{E}-\boldsymbol{A}^{-1}$；　B．$\boldsymbol{E}-\boldsymbol{A}$；　C．$\boldsymbol{E}+\boldsymbol{A}^{-1}$；　D．$\boldsymbol{E}+\boldsymbol{A}$．

（7）设 $\boldsymbol{A}$ 是 n 阶方阵，则 $|\boldsymbol{A}|=0$ 是 $\boldsymbol{A}$ 不可逆的　（　）

A．充分非必要条件；　B．必要非充分条件；

C．充分必要条件；　D．非充分非必要条件．

（8）设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 都是 n 阶方阵，下面 4 个等式中，必定成立的有几个？　（　）

$(\boldsymbol{A}+\boldsymbol{B})-\boldsymbol{C}=\boldsymbol{B}-(\boldsymbol{C}-\boldsymbol{A})$；　$\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{BC}$；

$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{B}(\boldsymbol{AC})$；　$[(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}]^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}$．

A．1 个；　B．2 个；　C．3 个；　D．4 个．

（9）设 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是可逆的对称矩阵，则不一定对称的矩阵是　（　）

A．$\boldsymbol{A}+\boldsymbol{B}$；　B．$\boldsymbol{AB}+\boldsymbol{BA}$；　C．$(\boldsymbol{AB})^{-1}$；　D．$\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$．

（10）设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 为同阶方阵，$\boldsymbol{E}$ 是同阶单位矩阵，若 $\boldsymbol{ABC}=\boldsymbol{E}$，则下列各式中，必定成立的是　（　）

A．$\boldsymbol{ACB}=\boldsymbol{E}$；　B．$\boldsymbol{BCA}=\boldsymbol{E}$；　C．$\boldsymbol{CBA}=\boldsymbol{E}$；　D．$\boldsymbol{BAC}=\boldsymbol{E}$．

第 3 章　线性方程组

3.1　线性方程组的矩阵消元解法

例 3.1　求解线性方程组 $\begin{cases} 2x_1+2x_2-x_3=6 \\ x_1-2x_2+4x_3=3 \\ 5x_1-x_2+x_3=4 \end{cases}$ **．**

解方程组通常采用消元法，比如将第 2 **个方程乘** -2 **加到第** 1 **个方程，可消去** x_1 **得到**

$$6x_2-9x_3=0\text{，}$$

将此方程两边除以 3**，约简可得**

$$2x_2-3x_3=0\text{．}$$

除了消元和约简，有时还要交换两个方程的位置．这些变形运算实际上仅在变量的系数之间进行，所以只需将所有的系数和常数项列成一个矩阵，作初等行变换即可．显然消元、约简和交换方程位置分别相当于矩阵的消去变换、倍缩变换和换行变换．比如上面对本例的两个具体变形相当于以下矩阵初等行变换：

$$\begin{pmatrix} 2 & 2 & -1 & 6 \\ 1 & -2 & 4 & 3 \\ 5 & -1 & 1 & 4 \end{pmatrix} \to \begin{pmatrix} 0 & 6 & -9 & 0 \\ 1 & -2 & 4 & 3 \\ 5 & -1 & 1 & 4 \end{pmatrix} \to \begin{pmatrix} 0 & 2 & -3 & 0 \\ 1 & -2 & 4 & 3 \\ 5 & -1 & 1 & 4 \end{pmatrix}\text{，}$$

其中第 1 **个变换是第** 2 **行乘** -2 **加到第** 1 **行，第** 2 **个变换是以** 1/3 **乘第** 1 **行．矩阵的初等变换可以使解方程组的过程显得紧凑、快捷、简洁．**

下面我们运用初等变换的标准程序（参看 2.4 **节）来解例** 3.1 **的线性方程组：**

$$\begin{pmatrix} 2 & 2 & -1 & 6 \\ \langle 1 \rangle & -2 & 4 & 3 \\ 5 & -1 & 1 & 4 \end{pmatrix} \to \begin{pmatrix} 0 & \langle 6 \rangle & -9 & 0 \\ 1 & -2 & 4 & 3 \\ 0 & 9 & -19 & -11 \end{pmatrix}$$

$$\xrightarrow{*} \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1.5 & 0 \\ 0 & 0 & \langle -5.5 \rangle & -11 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}\text{，}$$

其中，主元都用“ $\langle\ \rangle$ **”号做了标记．消元与换行可同步进行（如带“** $*$ **”号的第** 2 **步），**

换行的目的是为了使主元呈左上到右下排列．最后一个矩阵对应方程组

$$\begin{cases} x_1 + 0 + 0 = 1 \\ 0 + x_2 + 0 = 3 \\ 0 + 0 + x_3 = 2 \end{cases},$$

实际上已得到方程组的解是 $x_1 = 1$，$x_2 = 3$，$x_3 = 2$．写成列向量为 $\boldsymbol{x} = (1,\ 3,\ 2)^{\mathrm{T}}$，$\boldsymbol{x}$ 叫做解向量．显然解向量可以从最后一个矩阵右侧的常数列直接读出，无须写出对应的方程组．

第 2 章中曾经把一般的线性方程组（2.2）写成矩阵形式 $\boldsymbol{Ax} = \boldsymbol{b}$，比如例 3.1 的线性方程组，写成矩阵形式是

$$\begin{pmatrix} 2 & 2 & -1 \\ 1 & -2 & 4 \\ 5 & -1 & 1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix}.$$

我们把系数矩阵和常数列拼成的 $m \times (n+1)$ 矩阵 $(\boldsymbol{A} \mid \boldsymbol{b})$ 称为线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 的**增广矩阵**，满足方程的列向量 $\boldsymbol{x}$ 称为方程 $\boldsymbol{Ax} = \boldsymbol{b}$ 的**解向量**．线性方程组的**矩阵消元解法**，就是通过对增广矩阵施行一系列的初等行变换来求得方程组的解向量，矩阵消元解法中不允许作初等列变换．矩阵消元解法通常执行初等变换的标准程序，主元在系数矩阵范围内选取．初等变换的标准程序可以保证消元的有序性和完全性．

例 3.2　求解以下两个线性方程组：

（1）$\begin{pmatrix} 1 & -2 & 3 \\ 3 & -1 & 5 \\ 2 & 1 & 2 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix}$；　　（2）$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 1 \\ 1 & 5 & 2 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 \\ 3 \\ 7 \end{pmatrix}$.

解　（1）对增广矩阵作初等行变换：

$$\begin{pmatrix} \langle 1 \rangle & -2 & 3 & 1 \\ 3 & -1 & 5 & 6 \\ 2 & 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & \langle 5 \rangle & -4 & 3 \\ 0 & 5 & -4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 5 & -4 & 3 \\ 0 & 0 & 0 & -2 \end{pmatrix},$$

其中第 2 步仅仅将第 2 行乘 -1 加到第 3 行，即发现第 3 行对应的方程为 $0 + 0 + 0 = -2$，这是一个**矛盾方程**，无论 x 取何值，都不能使这个方程成立，所以原方程组无解．

（2）按标准程序作初等行变换：

$$\begin{pmatrix} \langle 1 \rangle & -1 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 1 & 5 & 2 & 7 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 3 & \langle 1 \rangle & 3 \\ 0 & 6 & 2 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

最后矩阵的第 3 行对应方程 $0 + 0 + 0 = 0$，这是一个恒等式，说明原方程组有多余的方程．比如在原方程组中，将第 1 个方程与第 2 个方程的 2 倍相加，恰好得到第 3 个方程，即第 3 个

方程是多余的．多余方程在对增广矩阵的初等行变换中会自动地显现出来，无须刻意寻找．

本例最后矩阵的前两行对应方程组

$$\begin{cases} x_1 - x_2 + 0 = 1 \\ 0 + 3x_2 + x_3 = 3 \end{cases},$$

这 2 个方程不能完全约束 3 个变量，其中必有一个变量可以取任意值，称之为**自由变量**，这里把 x_2 看作自由变量比较方便，任意指定 x_2 的一个值，就能够立即得到方程组的一个解向量．比如令 $x_2 = 0$，得解 $\boldsymbol{x} = (1,\ 0,\ 3)^{\mathrm{T}}$；令 $x_2 = 1$，得解 $\boldsymbol{x} = (2,\ 1,\ 0)^{\mathrm{T}}$；令 $x_2 = 2$，得解 $\boldsymbol{x} = (3,\ 2,\ -3)^{\mathrm{T}}$，等等．

从例 3.1 和例 3.2 可以看出，实施线性方程组的矩阵消元解法，可能出现下列情况：

（1）若某一行的元素全为零，则该行对应**多余方程**；

（2）若某一行除最右侧一个元素非零外，其余元素都为零，则该行对应矛盾方程，说明原方程组无解；

（3）如果标准程序执行完毕未出现矛盾方程，那么当所选主元个数等于变量个数时，方程组有惟一解；当所选主元个数少于变量个数时，方程组含有自由变量，从而有无穷多解．

3.2 矩 阵 的 秩

3.2.1 秩的概念

一个线性方程组虽含有多个方程，但其中可能有多余方程，因此线性方程组解的情况取决于去除多余方程后留下的独立方程．这说明方程组存在某种秩序，决定了方程组的解．这种秩序可以反映在系数矩阵和增广矩阵之中，它是矩阵的固有特性．尽管多余方程会在初等变换中显现出来，但我们有必要对这种秩序做出明确的定义．

在矩阵 $\boldsymbol{A}$ 中，任选 k 行、k 列，这些行、列交叉处的元素可构成一个 k 阶行列式，称为矩阵 $\boldsymbol{A}$ 的 $\boldsymbol{k}$ **阶子式**．$\boldsymbol{A}$ 的子式有许多，包括 1 阶子式、2 阶子式……子式的最高阶数不超过 $\boldsymbol{A}$ 的行数和列数．

如果矩阵 $\boldsymbol{A}$ 的所有子式中，不等于零的子式的最高阶数为 r，则 r 称为矩阵 $\boldsymbol{A}$ 的**秩**，记为 $R(\boldsymbol{A}) = r$ 或秩 $(\boldsymbol{A}) = r$．换一种说法，即矩阵 $\boldsymbol{A}$ 有一个 r 阶子式非零，而所有 $(r+1)$ 阶子式皆为零．对于零矩阵 $\boldsymbol{O}$，规定 $R(\boldsymbol{O}) = 0$．

由于行列式转置后，其值不变，所以 $R(\boldsymbol{A}) = R(\boldsymbol{A}^{\mathrm{T}})$．

设 $\boldsymbol{A}$ 是 $m \times n$ 矩阵，显然 $0 \leq R(\boldsymbol{A}) \leq \min(m,n)$．如果 $R(\boldsymbol{A}) = 0$，则 $\boldsymbol{A}$ 是零矩阵．如果 $R(\boldsymbol{A}) = \min(m,n)$，则称 $\boldsymbol{A}$ 为**满秩矩阵**．当 $m = n$ 时，因 $\boldsymbol{A}$ 的 n 阶子式只有一个，即 $|\boldsymbol{A}|$，故

对于方阵来说，满秩矩阵与非奇异矩阵、可逆矩阵是同一个概念，相互等价．

3.2.2　秩的求法

计算矩阵所有的子式后再确定矩阵的秩，显然太麻烦了．一般是通过矩阵的初等变换来求秩，由于初等变换不改变行列式的非零性，所以有以下定理．

定理 3.1　矩阵经初等变换后，其秩不变．

例 3.3　求矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-2&-1&0&2\\-2&4&2&6&-6\\2&-1&0&2&3\\3&3&3&3&4\end{pmatrix}$ 的秩．

解　按标准程序对 $\boldsymbol{A}$ 施行初等行变换：

$$\boldsymbol{A}=\begin{pmatrix}\langle 1\rangle&-2&-1&0&2\\-2&4&2&6&-6\\2&-1&0&2&3\\3&3&3&3&4\end{pmatrix}\to\begin{pmatrix}1&-2&-1&0&2\\0&0&0&6&\langle -2\rangle\\0&3&2&2&-1\\0&9&6&3&-2\end{pmatrix}$$

$$\to\begin{pmatrix}1&-2&-1&6&0\\0&0&0&-3&1\\0&3&\langle 2\rangle&-1&0\\0&9&6&-3&0\end{pmatrix}\to\begin{pmatrix}1&-0.5&0&5.5&0\\0&1.5&1&-0.5&0\\0&0&0&-3&1\\0&0&0&0&0\end{pmatrix},$$

其中第 3 步消元和换行同时进行．在最后的矩阵中，显然所有 4 阶子式均为零（有一行元素全为零），有一个 3 阶子式非零（选前 3 行及第 1，3，5 列）：

$$\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix}=1\neq 0,$$

所以 $R(\boldsymbol{A})=3$．

实际上矩阵的秩可以直接从最后的矩阵中读出：

$$\text{秩}=\text{所选主元个数}=\text{非零行数}=\text{基本单位列数}．\tag{3.1}$$

应该注意的是：在施行初等变换时，主元应在整个矩阵范围内选取，读取秩数必须在标准程序执行完毕后进行，这时最后的矩阵称为**最简矩阵**．

如果仅仅是求矩阵的秩，那么初等变换的程序可以简化：主元不必变为 1；已选过主元的行可不再参加变换．我们称之为**初等行变换的简化程序**．比如对本例运用简化程序的过程为

$$\boldsymbol{A}=\begin{pmatrix}\langle 1\rangle & -2 & -1 & 0 & 2\\ -2 & 4 & 2 & 6 & -6\\ 2 & -1 & 0 & 2 & 3\\ 3 & 3 & 3 & 3 & 4\end{pmatrix}\to\begin{pmatrix}1 & -2 & -1 & 0 & 2\\ 0 & 0 & 0 & 6 & -2\\ 0 & \langle 3\rangle & 2 & 2 & -1\\ 0 & 9 & 6 & 3 & -2\end{pmatrix}$$

$$\to\begin{pmatrix}1 & -2 & -1 & 0 & 2\\ 0 & 3 & 2 & 2 & -1\\ 0 & 0 & 0 & \langle 6\rangle & -2\\ 0 & 0 & 0 & -3 & 1\end{pmatrix}\to\begin{pmatrix}1 & -2 & -1 & 0 & 2\\ 0 & 3 & 2 & 2 & -1\\ 0 & 0 & 0 & 6 & -2\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

简化程序中，选取主元通常是从左至右按第 1 列，第 2 列，第 3 列……的顺序进行，最后的矩阵是一个阶梯形矩阵．秩数也可直接读出：

$$\text{秩}=\text{所选主元个数}=\text{非零行数}. \tag{3.2}$$

注意秩数的读取必须在简化程序执行完毕后进行，即只有当矩阵中选不出新的主元时，读出的秩数才是正确的．简化程序还允许对矩阵作初等列变换．

3.2.3 矩阵的秩与线性方程组的解

例 3.2（1）的线性方程组无解，其增广矩阵经初等行变换化为矩阵

$$\begin{pmatrix}1 & -2 & 3 & 1\\ 0 & 5 & -4 & 3\\ 0 & 0 & 0 & -2\end{pmatrix},$$

可以读出增广矩阵的秩等于 3．这个矩阵的前 3 列从系数矩阵变化而来，故由此可读出系数矩阵的秩等于 2，这两个秩数不相等．反之从例 3.1 和例 3.2（2）可读出两个秩数相等，这时方程组有解（惟一解或无穷多解）．

对于齐次线性方程组

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0},$$

常数列为零向量 $\boldsymbol{0}$（常数项全为 0）．其增广矩阵 $(\boldsymbol{A}\,|\,\boldsymbol{0})$ 中右侧的列全为零元素，因此，$R(\boldsymbol{A}\,|\,\boldsymbol{0})=R(\boldsymbol{A})$．显然齐次线性方程组至少有一个零解 $\boldsymbol{x}=\boldsymbol{0}$．如果 $R(\boldsymbol{A})$ 小于变量的个数，那么像例 3.2（2）那样会含有自由变量，从而方程组有无穷多解，即除了零解外，还有非零解（至少有一个变量取非零值）．

我们把以上分析归纳为下面的定理．

定理 3.2　设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{b}$ 是 $m\times 1$ 矩阵，$R(\boldsymbol{A})=r$，则方程组的解与矩阵的秩有如下关系：

（1）非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有解的充分必要条件是 $R(\boldsymbol{A}\,|\,\boldsymbol{b})=r$，即系数矩阵的秩等于增广矩阵的秩．

（2）若 $R(\boldsymbol{A}\vert\boldsymbol{b})=r$，则当 $r=n$ 时，方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有惟一解；当 $r<n$ 时，方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解．

（3）齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充分必要条件是 $r<n$．

当 $m=n$，即 $\boldsymbol{A}$ 是 n 阶方阵时，条件 $R(\boldsymbol{A})<n$ 等价于 $|\boldsymbol{A}|=0$．定理 3.2 的结论（3）在这个特殊情况下正是定理 1.4．当时我们并未证明充分性，现在得到了证明．

在对线性方程组实施矩阵消元解法时，$R(\boldsymbol{A})$ 和 $R(\boldsymbol{A}\vert\boldsymbol{b})$ 可从最后的矩阵中直接读出．$R(\boldsymbol{A}\vert\boldsymbol{b})$ 反映了线性方程组独立方程的个数，因此线性方程组解的状况与方程的个数 m 无关，而取决于增广矩阵的秩．

3.3　线性方程组解的结构

在 3.1 节中提出要用解向量来表示线性方程组的解．现在考虑当线性方程组有无穷多解时，如何用解向量表示全部解．

3.3.1　齐次线性方程组解的结构

例 3.4　求解齐次线性方程组 $\begin{pmatrix}1&-1&4&-1\\1&1&-2&3\\3&-1&6&1\\1&3&-8&7\end{pmatrix}\boldsymbol{x}=\boldsymbol{0}$．

解　运用矩阵消元解法．由于齐次线性方程组的增广矩阵中，最右侧的零元素在初等行变换中不起作用，因此可以仅对系数矩阵按标准程序作初等行变换：

$$\begin{pmatrix}\langle 1\rangle&-1&4&-1\\1&1&-2&3\\3&-1&6&1\\1&3&-8&7\end{pmatrix}\to\begin{pmatrix}1&-1&4&-1\\0&\langle 2\rangle&-6&4\\0&2&-6&4\\0&4&-12&8\end{pmatrix}\to\begin{pmatrix}1&0&1&1\\0&1&-3&2\\0&0&0&0\\0&0&0&0\end{pmatrix},$$

最后的矩阵显示只有 2 个独立方程，现有 4 个变量（和矩阵的第 1～4 列相对应），所以应含有 2 个自由变量．取非主元列对应的 x_3,x_4 为自由变量，它们可以是任意常数，即令 $x_3=c_1$，$x_4=c_2$，则 2 个独立方程变为

$$\begin{cases}x_1+0+c_1+c_2=0\\0+x_2-3c_1+2c_2=0\end{cases}.$$

移项后得到方程组的解及其向量形式（含任意常数 c_1,c_2）为

$$\begin{cases} x_1 = -c_1 - c_2 \\ x_2 = 3c_1 - 2c_2 \\ x_3 = c_1 \\ x_4 = c_2 \end{cases} \text{和} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = c_1 \begin{pmatrix} -1 \\ 3 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

向量形式可简写为：$\boldsymbol{x} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2$，其中向量

$$\boldsymbol{\alpha}_1 = (-1,\ 3,\ 1,\ 0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2 = (-1,\ -2,\ 0,\ 1)^{\mathrm{T}}$$

显然都是解向量（取 $c_1 = 1$，$c_2 = 0$ 和 $c_1 = 0$，$c_2 = 1$ 得到）．这样，我们仅用两个解向量就表示出了所有无穷多个解向量．

一般地，设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}\ , \boldsymbol{\alpha}_s$ 是齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的 s 个解向量，即有

$$\boldsymbol{A\alpha}_1 = \boldsymbol{A\alpha}_2 = \mathrm{L} = \boldsymbol{A\alpha}_s = \boldsymbol{0}\ . \tag{3.3}$$

如果 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}\ , \boldsymbol{\alpha}_s$ 满足条件：

（1）它们的线性表示式

$$\boldsymbol{x} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2 + \mathrm{L}\ + c_s\boldsymbol{\alpha}_s \tag{3.4}$$

给出 $\boldsymbol{Ax} = \boldsymbol{0}$ 的所有的解向量，其中 $c_1, c_2, \mathrm{L}\ , c_s$ 都是任意常数；

（2）它们之间不能相互线性表示．

则称这组解向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}\ , \boldsymbol{\alpha}_s$ 为齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的一个**基础解系**．由基础解系表示的解（3.4）称为方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的**通解**．

上面的条件（2）可以保证通解（3.4）是最简单的，没有多余的项．例 3.4 中的 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 就是一个基础解系．这是因为解向量可表示为 $\boldsymbol{x} = c_1\boldsymbol{\alpha}_1 + c_2\boldsymbol{\alpha}_2$，同时容易验证 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 不可能相互线性表示，即按向量相等规则，$\boldsymbol{\alpha}_1 = a\boldsymbol{\alpha}_2$ 或 $\boldsymbol{\alpha}_2 = b\boldsymbol{\alpha}_1$ 都不可能成立．

从例 3.4 还能够发现，基础解系可以从矩阵消元解法的最后矩阵中直接读出．**读取规则**是：

（1）找出自由变量．在最后的矩阵中，**非主元列**（不含主元的列）对应自由变量．

（2）令自由变量依次取 1（其余自由变量取 0），读取最后矩阵中相应列的元素，改变符号后，即为**主元变量**（与含主元的列相对应）的取值．

比如例 3.4 中，最后矩阵显示 x_3, x_4 是自由变量．令 $x_3 = 1$（$x_4 = 0$），则矩阵第 3 列元素（多余方程对应的零元素除外）的相反数 -1，3 给出主元变量 x_1, x_2 的取值，由此得到解向量 $\boldsymbol{\alpha}_1$；令 $x_4 = 1$（$x_3 = 0$），则从矩阵第 4 列读出 $x_1 = -1$，$x_2 = -2$，由此得到解向量 $\boldsymbol{\alpha}_2$．从矩阵中读出数值后需要变号，那是因为含有移项的过程．

应该指出的是，在读取基础解系之前，必须先将初等变换的标准程序执行完毕，亦即基础解系必须在最简矩阵中读出，否则会发生读取错误．这里再强调一下最简矩阵的三条标准：

（1）已经选不出新的主元；

（2）所有的主元都已经变为 1；

（3）主元的位置已经理顺（呈左上到右下排列）.

例 3.5　求解齐次线性方程组 $\begin{pmatrix}1&1&-1&1&1\\3&1&1&2&-3\\0&2&-4&1&6\\5&3&-1&4&-1\end{pmatrix}\boldsymbol{x}=\boldsymbol{0}$.

解　按标准程序对系数矩阵作初等行变换：

$$\begin{pmatrix}\langle 1\rangle&1&-1&1&1\\3&1&1&2&-3\\0&2&-4&1&6\\5&3&-1&4&-1\end{pmatrix}\to\begin{pmatrix}1&1&-1&1&1\\0&-2&4&-1&-6\\0&2&-4&\langle 1\rangle&6\\0&-2&4&-1&-6\end{pmatrix}\to\begin{pmatrix}\langle 1\rangle&-1&3&0&-5\\0&0&0&0&0\\0&2&-4&\langle 1\rangle&6\\0&0&0&0&0\end{pmatrix}.$$

为了读取方便，在最后的矩阵中，将选择过的主元都做了标记．最后的矩阵中，第 2，3，5 列是非主元列，显示 x_2,x_3,x_5 是自由变量，取 $(x_2,\ x_3,\ x_5)$ 依次为 $(1,\ 0,\ 0)$、$(0,\ 1,\ 0)$、$(0,\ 0,\ 1)$，依次读取第 2，3，5 列非零行元素，改变符号后即 x_1,x_4 的取值，从而得到基础解系

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

原方程组的通解为

$$\boldsymbol{x}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+c_3\boldsymbol{\alpha}_3,$$

这里 c_1,c_2,c_3 是任意常数.

本例在对矩阵作初等行变换时，如果主元选取的位置不同，会得到不同的基础解系．所以齐次线性方程组的基础解系不是惟一的，但基础解系所含解向量的个数都相同，等于自由变量的个数．由于系数矩阵的秩数等于主元变量的个数，所以自由变量的个数等于变量总数与矩阵秩数之差，即有下面的定理．

定理 3.3　如果齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 中，$\boldsymbol{A}$ 是 $m\times n$ 矩阵，$R(\boldsymbol{A})=r<n$，则该方程组的基础解系存在，而且基础解系含有 $(n-r)$ 个解向量．方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的通解由（3.4）式给出.

3.3.2　非齐次线性方程组解的结构

对于同一个系数矩阵 $\boldsymbol{A}$，方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 和 $\boldsymbol{Ax}=\boldsymbol{0}$ 虽不相同，但它们必有联系．下面讨论它们的解向量之间的关系．为方便起见，称齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 为非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的**导出组**.

如果向量 $\boldsymbol{\beta}$ 是方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个特解，即 $\boldsymbol{A\beta}=\boldsymbol{b}$，则对满足该方程组的其他解 $\boldsymbol{x}$ 来

说，因为

$$\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta})=\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{\beta}=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}\ ,$$

所以向量 $\boldsymbol{x}-\boldsymbol{\beta}$ 是导出组的解．按照公式（3.4）应该有

$$\boldsymbol{x}-\boldsymbol{\beta}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\mathrm{L}\ +c_s\boldsymbol{\alpha}_s\ ,$$

其中 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 是导出组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系，从而

$$\boldsymbol{x}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\mathrm{L}\ +c_s\boldsymbol{\alpha}_s+\boldsymbol{\beta}\ . \qquad (3.5)$$

这说明方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的任何解 $\boldsymbol{x}$ 都可由（3.5）式表示．反之将（3.5）式给出的向量代入方程组，注意到（3.3）式，得到

$$\begin{aligned}&\boldsymbol{A}(c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\mathrm{L}\ +c_s\boldsymbol{\alpha}_s+\boldsymbol{\beta})\\&=c_1\boldsymbol{A}\boldsymbol{\alpha}_1+c_2\boldsymbol{A}\boldsymbol{\alpha}_2+\mathrm{L}\ +c_s\boldsymbol{A}\boldsymbol{\alpha}_s+\boldsymbol{A}\boldsymbol{\beta}\ =\boldsymbol{0}+\boldsymbol{0}+\mathrm{L}\ +\boldsymbol{0}+\boldsymbol{b}=\boldsymbol{b}\ .\end{aligned}$$

这说明（3.5）式给出的向量一定是方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解向量．于是得到非齐次线性方程组解的结构定理如下．

定理 3.4 非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的**通解**（全部解）由（3.5）式给出，其中 $\boldsymbol{\beta}$ 是该方程组的一个特解，并且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 是导出组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系，这里 $c_1,c_2,\mathrm{L}\ ,c_s$ 是任意常数．

例 3.6 求解非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，其中 $\boldsymbol{A}=\begin{pmatrix}1&3&2&3\\1&-3&1&1\\4&-6&5&6\\3&-3&4&5\end{pmatrix}$，$\boldsymbol{b}=\begin{pmatrix}4\\1\\7\\6\end{pmatrix}$．

解 按标准程序，对增广矩阵作初等行变换：

$$(\boldsymbol{A}\,\vdots\,\boldsymbol{b})=\begin{pmatrix}\langle 1\rangle&3&2&3&4\\1&-3&1&1&1\\4&-6&5&6&7\\3&-3&4&5&6\end{pmatrix}\to\begin{pmatrix}1&3&2&3&4\\0&-6&\langle -1\rangle&-2&-3\\0&-18&-3&-6&-9\\0&-12&-2&-4&-6\end{pmatrix}\to\begin{pmatrix}\langle 1\rangle&-9&0&-1&-2\\0&6&\langle 1\rangle&2&3\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix}.$$

最后矩阵的前 4 列由系数矩阵 $\boldsymbol{A}$ 变化而来，第 2，4 列是非主元列，显示 x_2,x_4 是自由变量．令 $x_2=1$，$x_4=0$ 及 $x_2=0$，$x_4=1$，可读出导出组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系：

$$\boldsymbol{\alpha}_1=(9,\ \ 1,\ \ -6,\ \ 0)^{\mathrm{T}}\ ,\ \ \boldsymbol{\alpha}_2=(1,\ \ 0,\ \ -2,\ \ 1)^{\mathrm{T}}\ .$$

求原方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的一个特解，最简便的方法是令所有自由变量都取 0，这时，最后矩阵的右侧常数列（多余方程对应的零元素除外）恰好给出主元变量的取值（因没有移项过程，故不必变号），所以特解 $\boldsymbol{\beta}$ 也能从最后的矩阵中读出．本例中令 $x_2=x_4=0$，从右侧常数列读出 $x_1=-2$，$x_3=3$，即 $\boldsymbol{\beta}=(-2,\ \ 0,\ \ 3,\ \ 0)^{\mathrm{T}}$．原方程组的通解按（3.5）式写出为

$$\boldsymbol{x}=c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\boldsymbol{\beta}=c_1\begin{pmatrix}9\\1\\-6\\0\end{pmatrix}+c_2\begin{pmatrix}1\\0\\-2\\1\end{pmatrix}+\begin{pmatrix}-2\\0\\3\\0\end{pmatrix}.$$

其中 c_1,c_2 是任意常数．另外，从最后的矩阵中还可以读出 $R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{b})=2$．

例 3.7　解线性方程组 $\begin{cases}x_1+x_2+x_3+x_4=3\\x_1+x_2+2x_3+x_4=3\\-x_1+2x_2+5x_3-x_4=0\end{cases}$．

解　按标准程序，对增广矩阵作初等行变换：

$$\begin{pmatrix}\langle 1\rangle & 1 & 1 & 1 & 3\\1 & 1 & 2 & 1 & 3\\-1 & 2 & 5 & -1 & 0\end{pmatrix}\to\begin{pmatrix}1 & 1 & 1 & 1 & 3\\0 & 0 & \langle 1\rangle & 0 & 0\\0 & 3 & 6 & 0 & 3\end{pmatrix}$$

$$\to\begin{pmatrix}1 & 1 & 0 & 1 & 3\\0 & 0 & 1 & 0 & 0\\0 & \langle 3\rangle & 0 & 0 & 3\end{pmatrix}\to\begin{pmatrix}\langle 1\rangle & 0 & 0 & 1 & 2\\0 & \langle 1\rangle & 0 & 0 & 1\\0 & 0 & \langle 1\rangle & 0 & 0\end{pmatrix}.$$

最后一个步骤中，消元与换行同时进行．在最后的矩阵中，将选择过的主元都做了标记．前 4 列显示 x_4 是自由变量，令 $x_4=1$，读第 4 列（注意变号），得基础解系 $\boldsymbol{\alpha}=(-1,\ 0,\ 0,\ 1)^{\mathrm{T}}$；令 $x_4=0$，读第 5 列（右侧常数列，不必变号）得特解 $\boldsymbol{\beta}=(2,\ 1,\ 0,\ 0)^{\mathrm{T}}$．原方程组的通解是

$$\boldsymbol{x}=c\boldsymbol{\alpha}+\boldsymbol{\beta}=c(-1,\ 0,\ 0,\ 1)^{\mathrm{T}}+(2,\ 1,\ 0,\ 0)^{\mathrm{T}}.$$

当方程组有无穷多解时，如果主元选取的位置不同，则会得到不同的基础解系和特解，所以通解的形式不是惟一的．如例 3.4～例 3.7 的通解都还有其他形式．

在定理 3.4 的推导过程中，反复运用了“方程的解能使方程等式成立”的原理．运用类似的方法，不难得到以下关于解向量组合的结论．

（1）若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，则 $c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2$ 也是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，其中 c_1,c_2 是任意常数．

（2）若 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，则 $c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2$ 仅当 $c_1+c_2=1$ 时才是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解．

（3）若 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，则 $\boldsymbol{\beta}_1-\boldsymbol{\beta}_2$ 是导出组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解．

（4）若 $\boldsymbol{\beta}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解，$\boldsymbol{\alpha}$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，则 $\boldsymbol{\alpha}+\boldsymbol{\beta}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解．

例 3.8　已知线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有两个解向量 $\boldsymbol{\beta}_1=(0,\ 1,\ 2)^{\mathrm{T}}$，$\boldsymbol{\beta}_2=(2,\ 3,\ -1)^{\mathrm{T}}$，并且 $R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{b})=2$，求方程组的全部解．

解 解向量是 3 维的，即 $n=3$．$R(\boldsymbol{A})=r=2$，根据定理 3.3，$\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系包含 $n-r=1$ 个向量．另一方面根据上面的结论（3），$\boldsymbol{\beta}_1-\boldsymbol{\beta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解，所以它构成基础解系．于是根据定理 3.4，原方程组的全部解为

$$\boldsymbol{x}=c(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)+\boldsymbol{\beta}_1=c\left(-2,\,-2,\,3\right)^{\mathrm{T}}+\left(0,\,1,\,2\right)^{\mathrm{T}}.$$

3.4 矩阵方程的矩阵消元解法

线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，当系数矩阵 $\boldsymbol{A}$ 可逆时，有解 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$．在例 3.1 中，我们并未求逆矩阵 $\boldsymbol{A}^{-1}$，而是对增广矩阵作初等行变换，直接得到了解 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$．这种方法的原理是：逆矩阵 $\boldsymbol{A}^{-1}$ 由多个初等矩阵（初等变换）合成，在逐步获得这些初等变换的同时，立即作用于向量 $\boldsymbol{b}$，从而省略了求逆矩阵的中间步骤．这个原理可同样地用来解矩阵方程

$$\boldsymbol{AX}=\boldsymbol{B}. \tag{3.6}$$

如果矩阵 $\boldsymbol{A}$ 可逆，则方程有解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$．为了直接求得此解，可以构造增广矩阵 $(\boldsymbol{A}\vdots\boldsymbol{B})$，对其施行初等行变换，当把子块 $\boldsymbol{A}$ 变为单位矩阵时，子块 $\boldsymbol{B}$ 就变成了 $\boldsymbol{A}^{-1}\boldsymbol{B}$．

例 3.9 解矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$，其中 $\boldsymbol{A}=\begin{pmatrix}1&2\\3&4\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}5&6\\7&8\end{pmatrix}$．

解：对增广矩阵作如下初等行变换：

$$(\boldsymbol{A}\quad\boldsymbol{B})=\begin{pmatrix}\langle 1\rangle&2&5&6\\3&4&7&8\end{pmatrix}\to\begin{pmatrix}1&2&5&6\\0&\langle -2\rangle&-8&-10\end{pmatrix}\to\begin{pmatrix}1&0&-3&-4\\0&1&4&5\end{pmatrix},$$

从最后矩阵的右半部分读出解：

$$\boldsymbol{X}=\begin{pmatrix}-3&-4\\4&5\end{pmatrix}.$$

另一种矩阵方程

$$\boldsymbol{XA}=\boldsymbol{B},$$

当矩阵 $\boldsymbol{A}$ 可逆时有解 $\boldsymbol{X}=\boldsymbol{BA}^{-1}$．由于 $\boldsymbol{A}^{-1}$ 右乘于 $\boldsymbol{B}$，所以应该对 $\boldsymbol{B}$ 作相应的初等列变换（参看定理 2.4）．具体做法是将矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 上下拼接构成增广矩阵

$$\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}\end{pmatrix},$$

对其施行初等列变换，当把子块 $\boldsymbol{A}$ 变为单位矩阵时，子块 $\boldsymbol{B}$ 就变成了 $\boldsymbol{BA}^{-1}$．

例 3.10　解矩阵方程 $\boldsymbol{X}\begin{pmatrix}2&1&-1\\2&1&0\\1&-1&1\end{pmatrix}=\begin{pmatrix}4&3&2\\1&-2&5\end{pmatrix}$.

解　对增广矩阵作初等列变换．具体做法是选定主元，通过初等列变换把主元变为 1，同时把主元所在行的其他元素都变为 0：

$$\begin{pmatrix}2&1&-1\\2&1&0\\1&-1&\langle 1\rangle\\4&3&2\\1&-2&5\end{pmatrix}\to\begin{pmatrix}3&0&-1\\2&\langle 1\rangle&0\\0&0&1\\2&5&2\\-4&3&5\end{pmatrix}\to\begin{pmatrix}\langle 3\rangle&0&-1\\0&1&0\\0&0&1\\-8&5&2\\-10&3&5\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\-\frac{8}{3}&5&-\frac{2}{3}\\-\frac{10}{3}&3&\frac{5}{3}\end{pmatrix}.$$

从最后矩阵的下半部分读得解

$$\boldsymbol{X}=\frac{1}{3}\begin{pmatrix}-8&15&-2\\-10&9&5\end{pmatrix}.$$

更复杂的矩阵方程

$$\boldsymbol{AXB}=\boldsymbol{G},$$

当矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$ 均可逆时有解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{G}\boldsymbol{B}^{-1}$．这时可仿照例 3.9 先解方程 $\boldsymbol{AY}=\boldsymbol{G}$，求得矩阵 $\boldsymbol{Y}=\boldsymbol{A}^{-1}\boldsymbol{G}$，再仿照例 3.10 解方程 $\boldsymbol{XB}=\boldsymbol{Y}$，求得 $\boldsymbol{X}=\boldsymbol{Y}\boldsymbol{B}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{G}\boldsymbol{B}^{-1}$．

如果矩阵方程（3.6）中矩阵 $\boldsymbol{A}$ 不是可逆矩阵，甚至不是方阵，则方程的解可能出现多种情况．设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{B}$ 是 $m\times k$ 矩阵，则矩阵方程（3.6）相当于 k 个线性方程组．仿照定理 3.2 不难得到以下结论．

（1）方程（3.6）有解的充分必要条件是 $R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{B})$．

（2）如果 $R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{B})=r$，那么当 $r=n$ 时，方程（3.6）有惟一解；当 $r<n$ 时方程（3.6）有无穷多解．

参照定理 3.4，可以得到关于矩阵方程（3.6）的通解的结论：若

$$R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{B})=r,$$

则对增广矩阵 $(\boldsymbol{A}\,|\,\boldsymbol{B})$ 按标准程序作初等行变换（主元在子块 $\boldsymbol{A}$ 的范围内选取），可以得到最简矩阵

$$(\boldsymbol{A}\,|\,\boldsymbol{B})\to\begin{pmatrix}\boldsymbol{E}_r&\boldsymbol{Q}&\boldsymbol{G}\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

在最后矩阵的 r 个非零行中，左半部分由 r 个基本单位列（主元列）$\boldsymbol{E}_r$ 和 $(n-r)$ 个非单位列（非主元列）$\boldsymbol{Q}$ 构成（$\boldsymbol{E}_r$ 和 $\boldsymbol{Q}$ 应理解为它们的列是相互穿插、相互交错的）．右半部分是矩阵 $\boldsymbol{G}$．于是可以给出矩阵方程（3.6）的通解为

$$\boldsymbol{X}=\boldsymbol{UC}+\boldsymbol{V}\text{，}\tag{3.7}$$

其中$\boldsymbol{U}$是在$r\times(n-r)$负矩阵$-\boldsymbol{Q}$中插入$(n-r)$个基本单位行构成的矩阵，插入的位置与非主元列$\boldsymbol{Q}$所在的列序相对应，$\boldsymbol{V}$是在$r\times k$矩阵$\boldsymbol{G}$中插入$(n-r)$行零元素构成的矩阵，插入的位置与非主元列$\boldsymbol{Q}$所在的列序相对应，不妨表示为

$$\boldsymbol{U}=\begin{pmatrix}-\boldsymbol{Q}\\ \boldsymbol{E}_{n-r}\end{pmatrix},\ \boldsymbol{V}=\begin{pmatrix}\boldsymbol{G}\\ \boldsymbol{O}\end{pmatrix},$$

$\boldsymbol{C}$是$(n-r)\times k$矩阵，它的所有元素都是任意常数，即有$k(n-r)$个任意常数．

例 3.11 求矩阵方程$\begin{pmatrix}1&1&0&2&0\\0&0&1&3&0\\0&0&0&0&1\\0&0&0&0&0\end{pmatrix}\boldsymbol{X}=\begin{pmatrix}4&5\\6&7\\8&9\\0&0\end{pmatrix}$的通解．

解 增广矩阵已是最简形式，其中第 2，4 列是非主元列，所以按上面的结论由前三行（非零行）可直接读出通解为

$$\boldsymbol{X}=\begin{pmatrix}-1&-2\\1&0\\0&-3\\0&1\\0&0\end{pmatrix}\begin{pmatrix}c_{11}&c_{12}\\c_{21}&c_{22}\end{pmatrix}+\begin{pmatrix}4&5\\0&0\\6&7\\0&0\\8&9\end{pmatrix},$$

其中$c_{ij}(i,j=1,2)$是任意常数．

本例中，增广矩阵的第 1 列和第 2 列相同，因而也可将第 1，4 列作为非主元列，这样会得出不同的结果，这是允许的．当方程有无穷多解时，通解形式的多样性是很自然的．

（3.7）式中，$\boldsymbol{U}$称为**基础解系矩阵**，$\boldsymbol{V}$称为**特解矩阵**，$\boldsymbol{C}$称为**任意常数矩阵**．如果在方程（3.6）中，$\boldsymbol{B}=\boldsymbol{O}$是$m\times k$零矩阵，则必有$\boldsymbol{V}=\boldsymbol{O}$，通解（3.7）式就成为$\boldsymbol{X}=\boldsymbol{UC}$，这是定理 3.3 的推广．如果在方程（3.6）中，$R(\boldsymbol{A})=R(\boldsymbol{A}\,|\,\boldsymbol{B})=n$（$n$是矩阵$\boldsymbol{A}$的列数），则经过初等行变换得到的最简矩阵中不出现非主元列$\boldsymbol{Q}$，即初等行变换的结果为

$$(\boldsymbol{A}\quad\boldsymbol{B})\to\begin{pmatrix}\boldsymbol{E}_n&\boldsymbol{G}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}.$$

此时，（3.7）式中不出现第一项，方程有惟一解$\boldsymbol{X}=\boldsymbol{G}$，这是矩阵$\boldsymbol{A}$可逆情况的推广．

习 题 三

1．运用矩阵消元法解下列线性方程组：

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

2．确定 a, b 的值，使下列线性方程组有解，并求其一组解；同时指出方程组有惟一解还是无穷多解？并指出在原方程组中，（如果有的话）哪些方程是多余的？

(1) $\begin{cases} 2x_1 - x_2 + 3x_3 = 3 \\ 3x_1 + x_2 - 5x_3 = 0 \\ 4x_1 - x_2 + x_3 = 3 \\ x_1 + 3x_2 - 13x_3 = a \end{cases}$；　(2) $\begin{cases} x_1 + 2x_2 - 2x_3 + 2x_4 = 2 \\ x_2 - x_3 - x_4 = 1 \\ x_1 + x_2 - x_3 + 3x_4 = a \\ x_1 - x_2 + x_3 + 5x_4 = b \end{cases}$.

3．运用矩阵消元法求解下列线性方程组，指出方程组有惟一解、有无穷多解还是无解？读出系数矩阵和增广矩阵的秩，并指出在原方程组中，（如果有的话）哪一个方程是多余的？

(1) $\begin{cases} 2x_1 - x_2 + 3x_3 = 9 \\ 3x_1 - 5x_2 + x_3 = -4 \\ 4x_1 - 7x_2 + x_3 = 5 \end{cases}$；　(2) $\begin{cases} 5x_1 + 2x_2 - 7x_3 + 14x_4 = 21 \\ 5x_1 - x_2 + 8x_3 - 13x_4 + 3x_5 = 12 \\ 10x_1 + x_2 - 2x_3 + 7x_4 - x_5 = 29 \\ 15x_1 + 3x_2 + 15x_3 + 9x_4 + 7x_5 = 130 \\ 2x_1 - x_2 - 4x_3 + 5x_4 - 7x_5 = -13 \end{cases}$；

(3) $\begin{cases} x_1 + 2x_2 + 3x_3 + x_4 = 3 \\ x_1 + 4x_2 + 5x_3 + 2x_4 = 2 \\ 2x_1 + 9x_2 + 8x_3 + 3x_4 = 7 \\ 3x_1 + 7x_2 + 7x_3 + 2x_4 = 12 \end{cases}$；　(4) $\begin{cases} 4x_1 - 3x_2 + 2x_3 - x_4 = 8 \\ 3x_1 - 2x_2 + x_3 - 3x_4 = 7 \\ 2x_1 - x_2 - 5x_4 = 6 \\ 5x_1 - 3x_2 + x_3 - 8x_4 = 1 \end{cases}$.

4．求下列齐次线性方程组的基础解系以及系数矩阵的秩，并用基础解系表示方程组的通解．

(1) $\begin{cases} 3x_1 + 5x_2 + 2x_3 = 0 \\ 4x_1 + 7x_2 + 5x_3 = 0 \\ x_1 + x_2 - 4x_3 = 0 \end{cases}$；　(2) $\begin{cases} 2x_1 - 4x_2 + 5x_3 + 3x_4 = 0 \\ 3x_1 - 6x_2 + 4x_3 + 2x_4 = 0 \\ 4x_1 - 8x_2 + 17x_3 + 11x_4 = 0 \end{cases}$；

(3) $\begin{cases} x_1 + 2x_2 + 4x_3 - 3x_4 = 0 \\ 3x_1 + 5x_2 + 6x_3 - 4x_4 = 0 \\ 4x_1 + 5x_2 - 2x_3 + 3x_4 = 0 \\ 3x_1 + 8x_2 + 24x_3 - 19x_4 = 0 \end{cases}$；　(4) $\begin{cases} 2x_1 + 3x_2 - x_3 + 5x_4 = 0 \\ 3x_1 + x_2 + 2x_3 - 7x_4 = 0 \\ 4x_1 + x_2 - 3x_3 + 6x_4 = 0 \\ x_1 - 2x_2 + 4x_3 - 7x_4 = 0 \end{cases}$.

5．求解下列非齐次线性方程组，用向量形式表示它的通解，并求增广矩阵的秩．

(1) $\begin{cases} 2x_1+3x_2+x_3=4 \\ x_1-2x_2+4x_3=-5 \\ 3x_1+8x_2-2x_3=13 \\ 4x_1-x_2+9x_3=-6 \end{cases}$；　(2) $\begin{cases} 2x_1+x_2-x_3+x_4=1 \\ 3x_1-2x_2+x_3-3x_4=4 \\ x_1+4x_2-3x_3+5x_4=-2 \end{cases}$；

(3) $\begin{cases} 2x_1+7x_2+3x_3+x_4=5 \\ x_1+3x_2+5x_3-2x_4=3 \\ x_1+5x_2-9x_3+8x_4=1 \\ 5x_1+18x_2+4x_3+5x_4=12 \end{cases}$；　(4) $\begin{cases} 2x_1+3x_2-x_3+x_4=1 \\ 8x_1+12x_2-9x_3+8x_4=3 \\ 4x_1+6x_2+3x_3-2x_4=3 \\ 2x_1+3x_2+9x_3-7x_4=3 \end{cases}$.

6．讨论 p,q 为何值时，下列线性方程组有解、无解？有解时，求其通解．

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

7．已知线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有 3 个解向量

$$\boldsymbol{\beta}_1=(3,\ 1,\ 2,\ 5)^{\mathrm{T}},\ \boldsymbol{\beta}_2=(0,\ -1,\ 1,\ 1)^{\mathrm{T}},\ \boldsymbol{\beta}_3=(1,\ 2,\ 3,\ 4)^{\mathrm{T}},$$

并且系数矩阵 $\boldsymbol{A}$ 的秩为 2，求方程组的通解．

8．求解下列矩阵方程：

(1) $\begin{pmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}\boldsymbol{X}=\begin{pmatrix} 1 & -1 & 3 \\ 4 & 3 & 2 \\ 1 & -2 & 5 \end{pmatrix}$；　(2) $\boldsymbol{X}\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}=\begin{pmatrix} 4 & -6 \\ 2 & 1 \end{pmatrix}$；

(3) $\begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & -3 \\ 3 & 2 & -4 \end{pmatrix}\boldsymbol{X}\begin{pmatrix} -2 & 4 \\ 1 & -3 \end{pmatrix}=\begin{pmatrix} 3 & 1 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}$.

9．设 $\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & -1 & 2 & 0 \\ 4 & -2 & 6 & -1 \\ 2 & 4 & -2 & -3 \end{pmatrix}$，试求矩阵 $\boldsymbol{B}$，使 $\boldsymbol{AB}=\boldsymbol{O}$，且 $R(\boldsymbol{B})=2$．（提示：先解方程组 $\boldsymbol{Ax}=\boldsymbol{0}$）

10．单项选择题

(1) 运用矩阵消元法求解线性方程组，应该　（　）

A．对系数矩阵施行初等行变换；　B．对增广矩阵施行初等行变换；

C．对系数矩阵施行初等列变换；　D．对增广矩阵施行初等列变换．

(2) 对线性方程组的增广矩阵施行初等行变换，如果能将某一行的全部元素变为零，则该方程组（　）

A．有惟一解；　B．无解；　C．有无穷多解；　D．有多余方程．

(3) 下列矩阵中，满秩矩阵是　（　）

A．$\begin{pmatrix}1&1&1\\1&1&1\end{pmatrix}$；　B．$\begin{pmatrix}1&1&1\\1&2&3\\2&2&2\end{pmatrix}$；　C．$\begin{pmatrix}1&2\\2&4\\3&6\end{pmatrix}$；　D．$\begin{pmatrix}1&0&0\\0&0&1\end{pmatrix}$．

(4) 设 n 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的增广矩阵为 $(\boldsymbol{A}\vdots\boldsymbol{b})$，方程组有解的充分必要条件是　(　)

A．$R(\boldsymbol{A}\vdots\boldsymbol{b})=n$；　B．$R(\boldsymbol{A})=n$；

C．$R(\boldsymbol{A}\vdots\boldsymbol{b})=R(\boldsymbol{A})$；　D．$R(\boldsymbol{A}\vdots\boldsymbol{b})<n$ 且 $R(\boldsymbol{A})<n$．

(5) 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，则齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充分必要条件是　(　)

A．$m<n$；　B．$m>n$；　C．$m=n$；　D．$R(\boldsymbol{A})<n$．

(6) 已知 $\boldsymbol{A}$ 是 9×6 矩阵，方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有 4 个自由变量，则 $R(\boldsymbol{A})=$　(　)

A．2；　B．3；　C．4；　D．5．

(7) 如果线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 中方程的个数少于未知量的个数，则　(　)

A．$\boldsymbol{Ax}=\boldsymbol{b}$ 必无解；　B．导出组 $\boldsymbol{Ax}=\boldsymbol{0}$ 必有非零解；

C．$\boldsymbol{Ax}=\boldsymbol{b}$ 必有无穷多解；　D．导出组 $\boldsymbol{Ax}=\boldsymbol{0}$ 必无非零解．

(8) 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，$\boldsymbol{b}$ 是 m 维非零向量，则关于线性方程组的下列说法中，正确的是　(　)

A．$\boldsymbol{Ax}=\boldsymbol{0}$ 无非零解时，$\boldsymbol{Ax}=\boldsymbol{b}$ 无解；

B．$\boldsymbol{Ax}=\boldsymbol{0}$ 有无穷多解时，$\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解；

C．$\boldsymbol{Ax}=\boldsymbol{b}$ 有惟一解时，$\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解；

D．$\boldsymbol{Ax}=\boldsymbol{b}$ 无解时，$\boldsymbol{Ax}=\boldsymbol{0}$ 无非零解．

(9) 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的两个解向量，$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的两个解向量，则　(　)

A．$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解；　B．$\boldsymbol{\alpha}_1+\boldsymbol{\beta}_1$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解；

C．$\boldsymbol{\alpha}_1-\boldsymbol{\beta}_1$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解；　D．$\boldsymbol{\beta}_1-\boldsymbol{\beta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解．

(10) 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是 5 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的三个解向量，已知 $R(\boldsymbol{A})=3$，则该方程组的通解可能是　(　)

A．$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+c_3\boldsymbol{\alpha}_3$；　B．$c_1(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)+c_2(\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3)+\boldsymbol{\alpha}_3$；

C．$c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$；　D．$c(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)+\boldsymbol{\alpha}_3$．

第 4 章　向量的线性关系

矩阵的每一行（列）都构成行（列）向量．对矩阵施行初等行变换，实质是行向量的线性运算．线性方程组解的结构涉及解向量的线性组合，而解向量的读取又与矩阵的列向量密切相关．向量的线性关系还会扩展到其他领域．本章详细介绍向量的线性关系理论．

4.1　向量的线性相关性

4.1.1　两种线性关系

给定向量 $\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$，如果存在一组数 $k_1,k_2,\mathrm{L}\ ,k_s$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+k_s\boldsymbol{\alpha}_s \tag{4.1}$$

成立，则称 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 的**线性组合**，或称 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ **线性表示**．

线性表示体现了一种替代性，如果 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 线性表示，那么在向量组 $\boldsymbol{\beta},\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 中，$\boldsymbol{\beta}$ 是“多余”的，它可由其他向量通过线性组合（4.1）来替代．

例 4.1　基本单位向量组

$$\boldsymbol{e}_1=\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \boldsymbol{e}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \boldsymbol{e}_3=\begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ \boldsymbol{e}_4=\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

可以线性表示任何向量，而且线性表示的系数很容易确定．比如不难验证

$$\begin{pmatrix}-1\\2\\5\\3\end{pmatrix}=-\begin{pmatrix}1\\0\\0\\0\end{pmatrix}+2\begin{pmatrix}0\\1\\0\\0\end{pmatrix}+5\begin{pmatrix}0\\0\\1\\0\end{pmatrix}+3\begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

上式右端的线性运算（数乘与加法）就是矩阵的线性运算．上式作为线性表示式，其表示系数就是左端向量的分量．这个原理可直接推广到具有 n 个分量的 n 维向量，体现了基本单位向量组的**完备性**．

给定向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$，如果存在一组不全为零的数 $k_1,k_2,\mathrm{L}\ ,k_s$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\mathrm{L}\ +k_s\boldsymbol{\alpha}_s=\mathbf{0} \tag{4.2}$$

成立，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ **线性相关**.

反之，如果只有当 $k_1,k_2,\mathrm{L}\ ,k_s$ 全为零时，上式才成立，即由（4.2）式可以推出 $k_1=k_2=\mathrm{L}\ =k_s=0$，则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ **线性无关**.

线性相关体现了向量组内部存在某种关联性，而线性无关则表明向量组内部没有任何线性关联性.

以下是几个简单的特殊情况.

（1）单个非零向量线性无关.

若 $\boldsymbol{\alpha}_1\neq\mathbf{0}$，则 $\boldsymbol{\alpha}_1$ 至少有一个分量非零，由等式 $k_1\boldsymbol{\alpha}_1=\mathbf{0}$，比较该分量知 $k_1=0$.

（2）含有零向量的向量组线性相关.

比如对于向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathbf{0}$，等式 $0\cdot\boldsymbol{\alpha}_1+0\cdot\boldsymbol{\alpha}_2+1\cdot\mathbf{0}=\mathbf{0}$ 成立，其中系数为 $0,0,1$，不全为零.

（3）若向量组中有两个向量的对应分量成比例，则该向量组线性相关.

比如，对于 $\boldsymbol{\alpha}_1=(1,\ 1,\ 1)$，$\boldsymbol{\alpha}_2=(2,\ 2,\ 2)$，$\boldsymbol{\alpha}_3=(1,\ 2,\ 3)$，等式

$$2\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+0\cdot\boldsymbol{\alpha}_3=\mathbf{0}$$

成立.

例 4.2　基本单位向量之间线性无关. 比如在等式

$$k_1\begin{pmatrix}1\\0\\0\\0\end{pmatrix}+k_2\begin{pmatrix}0\\1\\0\\0\end{pmatrix}+k_3\begin{pmatrix}0\\0\\1\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

中，比较左、右两端的各分量，可得 $k_1=k_2=k_3=0$，从而可知基本单位向量

$$(1,\ \ 0,\ \ 0,\ \ 0)^{\mathrm{T}}\text{，}(0,\ \ 1,\ \ 0,\ \ 0)^{\mathrm{T}}\text{，}(0,\ \ 0,\ \ 1,\ \ 0)^{\mathrm{T}}$$

线性无关.

定理 4.1　向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$（$s\geq 2$）线性相关的充分必要条件是其中至少有一个向量可以由其他向量线性表示.

证　充分性：不妨设 $\boldsymbol{\alpha}_1$ 可由 $\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 线性表示，即 $\boldsymbol{\alpha}_1=k_2\boldsymbol{\alpha}_2+\mathrm{L}\ +k_s\boldsymbol{\alpha}_s$，移项后得到等式

$$(-1)\boldsymbol{\alpha}_1=k_2\boldsymbol{\alpha}_2+\mathrm{L}\ +k_s\boldsymbol{\alpha}_s=\mathbf{0}\text{，}$$

式中至少第一个系数 (-1) 非零，因此向量组线性相关.

必要性：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 线性相关，即有等式（4.2），其中系数 $k_1,k_2,\mathrm{L}\ ,k_s$ 至少有一个非

零．不妨设 $k_1 \neq 0$，则以 k_1 遍除等式，再移项得

$$\boldsymbol{\alpha}_1 = \left(-\frac{k_2}{k_1}\right)\boldsymbol{\alpha}_2 + \mathrm{L} + \left(-\frac{k_s}{k_1}\right)\boldsymbol{\alpha}_s ,$$

这说明 $\boldsymbol{\alpha}_1$ 可由 $\boldsymbol{\alpha}_2, \mathrm{L}, \boldsymbol{\alpha}_s$ 线性表示．

这是两种线性关系的互通性定理．定理 4.1 表明，线性组合的替代性直接反映了向量组内部的关联性，而向量组线性相关即意味着有某些向量是“多余”的、能够被替代的．反之，向量组线性无关即意味着无替代性，向量都是“有用”的、没有“多余”的．

第 3 章曾介绍过齐次线性方程组的基础解系（参看 3.3 节），基础解系中向量的无替代性等价于线性无关性，因此可以给出基础解系更明确的定义．

如果齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的一组解向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}, \boldsymbol{\alpha}_s$ 满足条件：

（1）方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的任一解向量都可 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}, \boldsymbol{\alpha}_s$ 线性表示；

（2）$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}, \boldsymbol{\alpha}_s$ 线性无关．

则称 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathrm{L}, \boldsymbol{\alpha}_s$ 是方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 的一个**基础解系**．

构造基础解系通常是利用基本单位向量的完备性（例 4.1）和线性无关性（例 4.2）实现的．如例 3.5 中，取自由变量 $(x_2,\ x_3,\ x_5)$ 依次为 $(1,\ 0,\ 0)$、$(0,\ 1,\ 0)$、$(0,\ 0,\ 1)$，这 3 个基本单位向量保证了由此得出的基础解系的完备性和线性无关性．当然构造基础解系也可利用其他向量，比如，取自由变量 $(x_2,\ x_3,\ x_5)$ 依次为 $(1,\ 0,\ 0)$、$(1,\ 1,\ 0)$、$(1,\ 1,\ 1)$，也可得到基础解系，只是主元变量的取值要经过一番计算，不能直接读取了．

4.1.2 线性关系和线性方程组

矩阵形式的线性方程组

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix},$$

可以写作向量形式

$$x_1\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + x_2\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + x_3\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}.$$

不难看出，这两种形式作为向量的相等关系，都相当于同样的线性方程组

$$\begin{cases} x_1 + 4x_2 + 7x_3 = 3 \\ 2x_1 + 5x_2 + 8x_3 = 5 \\ 3x_1 + 6x_2 + 9x_3 = 7 \end{cases}.$$

一般地，设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \mathrm{L}\ ,\ \boldsymbol{\alpha}_s)$，其中 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 是矩阵 $\boldsymbol{A}$ 的列向量，则方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 等价于**向量形式** $x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\mathrm{L}\ +x_s\boldsymbol{\alpha}_s=\boldsymbol{\beta}$，这表明向量 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 的线性组合．当 $\boldsymbol{\beta}=\boldsymbol{0}$ 时，又关系到 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 是否线性相关．由此看来，表明线性关系的 (4.1) 式和 (4.2) 式实际上是线性方程组的向量形式，下面的定理就是将两者联系起来的结果．

定理 4.2　列向量 $\boldsymbol{\beta}$ 能够由列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 线性表示的充分必要条件是线性方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 有解，其中 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \mathrm{L}\ ,\ \boldsymbol{\alpha}_s)$．同时，线性方程组的解 $\boldsymbol{x}=(x_1,\ x_2,\ \mathrm{L}\ ,\ x_s)^{\mathrm{T}}$ 正好给出了线性表示式（4.1）的表示系数．

定理 4.3　列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L}\ ,\boldsymbol{\alpha}_s$ 线性相关的充分必要条件是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 有非零解，其中 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \mathrm{L}\ ,\ \boldsymbol{\alpha}_s)$．同时，非零解向量 $\boldsymbol{x}=(x_1,\ x_2,\ \mathrm{L}\ ,\ x_s)^{\mathrm{T}}$ 正好给出了线性相关等式（4.2）中不全为零的系数．

如果 $\boldsymbol{A}$ 是方阵，则定理 4.3 中的充分必要条件可简化为 $|\boldsymbol{A}|=0$．

定理 4.2 和定理 4.3 给出了具体求解向量间线性关系的途径．

例 4.3　设有 4 个向量：$\boldsymbol{\beta}=(2,\ 7,\ 5)$，$\boldsymbol{\alpha}_1=(1,\ 2,\ 1)$，$\boldsymbol{\alpha}_2=(2,\ 5,\ 3)$，$\boldsymbol{\alpha}_3=(-1,\ 1,\ 2)$，问 $\boldsymbol{\beta}$ 可否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示？$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是否线性相关？

解　把所有的向量写作列向量．根据定理 4.2 和定理 4.3，应该考虑解线性方程组

$$\begin{pmatrix}1&2&1\\2&5&1\\1&3&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}2\\5\\7\end{pmatrix}.$$

用矩阵消元解法，按标准程序对增广矩阵作初等行变换：

$$\begin{pmatrix}\langle 1\rangle&2&-1&2\\2&5&1&7\\1&3&2&5\end{pmatrix}\to\begin{pmatrix}1&2&-1&2\\0&\langle 1\rangle&3&3\\0&1&3&3\end{pmatrix}\to\begin{pmatrix}\langle 1\rangle&0&-7&-4\\0&\langle 1\rangle&3&3\\0&0&0&0\end{pmatrix}.$$

因没有矛盾方程，故方程组有解，从而根据定理 4.2，$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示．最后矩阵的前 3 列中，第 3 列是非主元列，显示 x_3 是自由变量，这表明线性方程组的导出组有非零解，所以根据定理 4.3，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关．

此外，线性方程组因存在自由变量而有无穷多解，从而线性表示方式也有无穷多种．令 $x_3=0$，可读出 $x_1=-4,x_2=3$，于是得到一个线性表示式

$$\boldsymbol{\beta}=-4\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+0\cdot\boldsymbol{\alpha}_3\ .$$

从本例看出，解决线性关系问题，可先构造相应的线性方程组，进而落实到对矩阵作初等行变换，得到最简矩阵后，就能直接读取需要的结论．

4.2 极大线性无关组与向量组的秩

4.2.1 极大线性无关组的概念

向量组中的部分向量构成一个**部分组**.比如，$\alpha_1,\alpha_2,\alpha_3$ 是 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 的一个部分组.如果存在不全为零的数 k_1,k_2,k_3，使 $k_1\alpha_1+k_2\alpha_2+k_3\alpha_3=0$ 成立，那么下式也成立：

$$k_1\alpha_1+k_2\alpha_2+k_3\alpha_3+0\cdot\alpha_4=\mathbf{0},$$

这说明向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 也线性相关．由此可得到：

（1）如果某些部分向量线性相关，那么全体向量一定线性相关；

（2）如果全体向量线性无关，那么任意部分向量一定线性无关.

这两个结论互为逆否命题．应当注意它们的逆命题是不成立的，即部分向量线性无关时，全体向量未必线性无关；全体向量线性相关时，部分向量未必线性相关．也就是说，线性相关的向量组，其部分组可能线性无关.

一个线性无关的部分组用来线性表示该部分组自身的向量是显然的，比如用 α_1,α_2 来线性表示 α_1，有 $\alpha_1=1\cdot\alpha_1+0\cdot\alpha_2$，但是未必能线性表示部分组以外的其他向量．若有一个线性无关的部分组，能够线性表示整个向量组中的任何向量，则称它为整个向量组的一个**极大线性无关组**，简称**极大无关组**.

由于极大无关组可以线性表示其他任何向量，所以在极大无关组之中，再添一个向量，必然形成线性相关（参看定理 4.1），这说明在保持线性无关性的前提下，向量的个数已达最大可能，自然应冠之以“极大”.

显然，如果一个向量组线性无关，那么该向量组的极大无关组就是向量组本身.

因为极大无关组本身线性无关，即没有“多余”的向量，所以极大无关组提供了用最少数的向量来代表全体向量的依据．比如基础解系就是齐次线性方程组全体解向量的极大无关组.

4.2.2 极大线性无关组的求法

定理 4.2、定理 4.3 和例 4.3 都表明，分析向量的线性关系可归结为求解线性方程组.具体做法是将所有的向量写成列向量后拼成矩阵，对该矩阵施行初等行变换．初等行变换不仅不会改变显示线性关系的（4.1）式或（4.2）式中的系数，而且最终还能把它们求出来.因此，可以得到：

定理 4.4 对矩阵施行初等行变换，不改变矩阵列向量之间的线性关系.

例 4.4 求 5 个向量构成的向量组 $\alpha_1=(1,\ 2,\ 3,\ 2)$，$\alpha_2=(-1,\ 2,\ 1,\ -2)$，$\alpha_3=(1,\ -1,\ -1,\ 1)$，$\alpha_4=(-2,\ 1,\ -2,\ -5)$，$\alpha_5=(1,\ -3,\ -2,\ 2)$ 的极大无关组，并且用极大无关组线性表示其他向量.

解　把所有的向量写成列向量，按顺序拼成矩阵，然后作初等行变换，执行标准程序如下（所选主元都做了标记）：

$$\begin{pmatrix} 1 & -1 & \langle 1\rangle & -2 & 1 \\ 2 & 2 & -1 & 1 & -3 \\ 3 & 1 & -1 & -2 & -2 \\ 2 & -2 & 1 & -5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & -2 & 1 \\ 3 & \langle 1\rangle & 0 & -1 & -2 \\ 4 & 0 & 0 & -4 & -1 \\ 1 & -1 & 0 & -3 & 1 \end{pmatrix}$$

$$\to \begin{pmatrix} 3 & 1 & 0 & -1 & -2 \\ 4 & 0 & 1 & -3 & -1 \\ 4 & 0 & 0 & -4 & \langle -1\rangle \\ 4 & 0 & 0 & -4 & -1 \end{pmatrix} \to \begin{pmatrix} -5 & 1 & 0 & 7 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ -4 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

其中第 2 个步骤，消元与换行同时进行．在最后的矩阵中，第 2、3、5 列是主元列（含主元的列），也是基本单位列．根据基本单位向量的线性无关性和完备性（参看例 4.1 和例 4.2）可知，最后矩阵的 5 个列向量中，第 2、3、5 列向量构成极大无关组，用它线性表示第 1、4 列向量，表示系数分别为第 1、4 列向量的分量，即 $-5,\ 0,\ -4$ 和 $7,\ 1,\ 4$．根据定理 4.4，原向量组也有完全相同的线性关系，于是可直接得到本例的结论：原向量组中，$\alpha_2,\alpha_3,\alpha_5$ 是一个极大无关组，并且有线性表示式

$$\alpha_1 = -5\alpha_2 - 4\alpha_5\ ,\quad \alpha_4 = 7\alpha_2 + \alpha_3 + 4\alpha_5\ .$$

4.2.3　向量组的秩

例 4.4 在对矩阵作初等行变换时，如果主元选取的位置不同，会得到不同的极大无关组，可见极大无关组不是惟一的．因为矩阵的秩是常数 3，无论主元怎么选取，最后得到的基本单位向量必然是 3 个，所以用初等变换法求得的极大无关组尽管不惟一，但所含向量的个数都相同．然而用其他方法获得极大无关组（只要满足线性无关性和完备性即可），其所含向量的个数是否仍是一个固定的常数呢？答案是肯定的，不过推导过程并不简单，先要建立有关线性关系与向量个数的定理．

定理 4.5　当向量组中向量的个数大于向量的维数时，向量组一定线性相关．

证　设 $\alpha_1,\alpha_2,\mathrm{L}\ ,\alpha_s$ 都是 n 维向量（有 n 个分量），则矩阵 $\boldsymbol{A}=(\alpha_1,\ \alpha_2,\ \mathrm{L}\ ,\ \alpha_s)$ 是 $n\times s$ 矩阵，于是齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有 n 个方程，s 个变量．已知 $s>n$，即变量多，方程少，所以 $\boldsymbol{Ax}=\boldsymbol{0}$ 一定含有自由变量，从而必有非零解，根据定理 4.3，可知向量组线性相关．

应该指出的是，当向量的个数小于向量的维数时，不能得到线性相关还是线性无关的确定结论．因为此时齐次线性方程组的方程数虽然多于变量数，但独立方程数却不确定，从而有无非零解也不确定。

定理 4.6　设有两个向量组，向量组（Ⅰ）有 s 个向量，可以线性表示向量组（Ⅱ）中的每一个向量．如果向量组（Ⅱ）有 r 个向量，并且线性无关，那么 $r \le s$．

证 设向量组（Ⅰ）是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$，向量组（Ⅱ）是$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$，记$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s)$，根据定理 4.2，向量组（Ⅰ）可以线性表示向量组（Ⅱ），即线性方程组$\boldsymbol{Ax}=\boldsymbol{\beta}_i$（$i$=1, 2, …, r）都有解，这r个解向量记为$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$，它们都是$s$维向量且$\boldsymbol{Ax}_i=\boldsymbol{\beta}_i$，其中$i$=1, 2, …, r.

用反证法，设$r>s$，则根据定理 4.5，这些解向量线性相关，即存在不全为零的数$k_1,k_2,\cdots,k_r$，使

$$k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_r\boldsymbol{x}_r=\boldsymbol{0}$$

成立．以矩阵A左乘该式两端，可得

$$k_1\boldsymbol{\beta}_1+k_2\boldsymbol{\beta}_2+\cdots+k_r\boldsymbol{\beta}_r=\boldsymbol{0},$$

这与向量组（Ⅱ）线性无关矛盾，于是否定假设，结论$r\le s$得证．

这个定理的逆否命题是：当向量个数少的向量组可线性表示向量个数多的向量组时$(s<r)$，后者必定线性相关．

定理 4.7 如果两个线性无关的向量组能够相互线性表示，那么它们所含向量的个数必定相同．

证 设向量组（Ⅰ）有s个向量，向量组（Ⅱ）有r个向量，已知向量组（Ⅰ）可线性表示线性无关的向量组（Ⅱ），根据定理 4.6，可得$r\le s$；又已知向量组（Ⅱ）可线性表示线性无关的向量组（Ⅰ），同理可得$s\le r$，于是$r=s$．

两个向量组可以相互线性表示，称为**等价**．等价意味着两个向量组能够相互替代．等价的向量组，向量的个数未必相同，但等价的线性无关的向量组，所含向量个数一定相同．

由于极大无关组可以线性表示整个向量组中的任何向量，所以两个极大无关组可以相互线性表示，于是根据定理 4.7，它们所含向量的个数必然相同．这说明，极大无关组虽然不惟一，但它们所含向量的个数是惟一的，是一个常数，我们称之为**向量组的秩**．如例 4.4 中极大无关组含有 3 个向量，所以向量组的秩等于 3．

向量组的秩体现了整个向量组不计“多余”向量的实际规模．向量组的秩不仅是单个极大无关组中线性无关向量的个数的最大值，而且是所有线性无关部分组的向量个数的最大值，因此一个线性无关的部分组，如果所含向量个数等于向量组的秩，那么它一定是极大无关组，一定可以线性表示所有的向量．

第 3 章中通过自由变量依次取 1 来求齐次线性方程组的基础解系（参看 3.3 节），很自然地得到：不同的基础解系含有相同的向量个数．但是从理论上说，基础解系只要满足线性无关性和完备性，并不一定要通过自由变量的取值来得到，应当允许通过其他方式获得，甚至不必求出来而只就其存在性进行推断．定理 4.7 保证了基础解系作为解向量的极大无关组，无论来源如何，其所含解向量的个数必然相同，所以说定理 3.3 至此才得到了完整的证明．

对于有限个向量来说，求向量组的秩可以用例 4.4 的矩阵变换法．对矩阵$\boldsymbol{A}$作初等行变换，求得的是矩阵$\boldsymbol{A}$的列向量组的秩（简称**列秩**），并且恰好等于$R(\boldsymbol{A})$；对转置矩阵$\boldsymbol{A}^{\mathrm{T}}$

作初等行变换，求得的是 A^{T} 的列向量组亦即 A 的行向量组的秩（简称**行秩**）．因为 $R(\boldsymbol{A})=R(\boldsymbol{A}^{\mathrm{T}})$，所以有下列等式：

$$R(\boldsymbol{A})=\boldsymbol{A}\text{ 的行秩}=\boldsymbol{A}\text{ 的列秩}. \tag{4.3}$$

（4.3）式说明矩阵 A 有 3 种秩，只要求出其中之一，另两个也就得到了．（4.3）式还说明，向量组的秩与矩阵的秩没有区别，所以共用同一个名称——秩．

习 题 四

1．判断下列向量组是否线性相关．

（1）$(1,\ 2,\ 3)$，$(3,\ 6,\ 9)$；

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

（3）$(2,\ -3,\ 1)$，$(3,\ -1,\ 5)$，$(1,\ -4,\ 3)$；

（4）$(1,\ 0,\ 0)$，$(0,\ 1,\ 0)$，$(0,\ 0,\ 1)$，$(1,\ 2,\ 3)$；

（5）$(4,\ -5,\ 2,\ 6)$，$(2,\ -2,\ 1,\ 3)$，$(6,\ -3,\ 3,\ 9)$，$(4,\ -1,\ 5,\ 6)$．

2．已知向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性无关，试证向量组 $2\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2$，$\boldsymbol{\alpha}_2+4\boldsymbol{\alpha}_3$，$\boldsymbol{\alpha}_1+5\boldsymbol{\alpha}_3$ 也线性无关．

3．证明：两个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性相关的充分必要条件是它们的分量对应成比例．

4．试问下列向量组中，向量 $\boldsymbol{\beta}$ 能否由其余向量线性表示？若能，则写出线性表示式．

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

（2）$\boldsymbol{\beta}=(7,\ -2,\ 15)$，$\boldsymbol{\alpha}_1=(2,\ 3,\ 5)$，$\boldsymbol{\alpha}_2=(3,\ 7\ \ 8)$，$\boldsymbol{\alpha}_3=(1,\ -6,\ 1)$；

（3）$\boldsymbol{\beta}=(1,\ 2,\ 5)$，$\boldsymbol{\alpha}_1=(3,\ 2,\ 6)$，$\boldsymbol{\alpha}_2=(7,\ 3\ \ 9)$，$\boldsymbol{\alpha}_3=(5,\ 1,\ 3)$；

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

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

5．求下列向量组的秩与一个极大无关组，并将其余向量用此极大无关组线性表示．

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

（2）$\boldsymbol{\alpha}_1=(1,\ 1,\ 3,\ 1)$，$\boldsymbol{\alpha}_2=(-1,\ 1,\ -1,\ 3)$，$\boldsymbol{\alpha}_3=(5,\ -2,\ 8,\ -9)$，$\boldsymbol{\alpha}_4=(-1,\ 3,\ 1,\ 7)$；

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

6．单项选择题

（1）列向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$（$n\geqslant 2$）线性相关的充分必要条件是　（　）

A．其中有 1 个零向量；　　B．其中有 1 个向量可由其余向量线性表示；

C．增加 1 个向量后成为线性相关；　　D．其中有 2 个向量的对应分量成比例．

（2）设 A 是 5 阶矩阵，$R(A)=2$，则齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的所有解向量中，线性无关的解向量个数是　（　）

A．2；　　B．3；　　C．4；　　D．5．

(3) 设 A 是 $m\times n$ 矩阵，则齐次线性方程组 $Ax=0$ 仅有零解的充分必要条件是 ( )

A．A 的行向量组线性无关；　　B．A 的行向量组线性相关；

C．A 的列向量组线性无关；　　D．A 的列向量组线性相关．

(4) 列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ （$n\geq 2$）线性无关的充分必要条件是 ( )

A．所有向量非零；　　B．方程组 $(\alpha_1,\alpha_2,\cdots,\alpha_n)x=0$ 无非零解；

C．有一个部分组线性无关；　　D．α_1 不能由其余向量线性表示．

(5) 向量组 $\alpha_1=(1,\ 0,\ 0)$，$\alpha_2=(0,\ 1,\ 0)$，$\alpha_3=(1,\ 1,\ 1)$，$\alpha_4=(0,\ 1,\ 1)$ 的极大无关组可以是 ( )

A．α_1,α_2；　　B．$\alpha_1,\alpha_3,\alpha_4$；　　C．$\alpha_2,\alpha_3,\alpha_4$；　　D．$\alpha_1,\alpha_2,\alpha_3,\alpha_4$．

(6) 设 n 维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 的极大无关组含有 r 个向量，则 ( )

A．$r\leq n$；　　B．$r<n$；　　C．$r<m$；　　D．$r\geq m$．

(7) 设 $\alpha_1,\alpha_2,\cdots,\alpha_8$ 是 6 维向量组，则该向量组 ( )

A．线性相关；　　B．其中至少有 2 个向量可由其余向量线性表示；

C．其中至少有 3 个向量可由其余向量线性表示；　　D．其中只有 1 个向量可由其余向量线性表示．

(8) n 个 $n+1$ 维向量构成的向量组 ( )

A．一定线性相关；　　B．增加一个 $n+1$ 维向量后成为线性相关；

C．一定线性无关；　　D．增加两个 $n+1$ 维向量后成为线性相关．

(9) 设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_{10}$ 中，$\alpha_1,\alpha_2,\alpha_3$ 线性无关，则该向量组的秩 r 满足 ( )

A．$3\leq r\leq 10$；　　B．$3\leq r\leq 7$；　　C．$r=7$；　　D．$r=3$．

(10) 若向量组的秩为 r，则该向量组中 ( )

A．任意 r 个向量线性无关；　　B．任意 $r+1$ 个向量线性相关；

C．任意 r 个向量线性相关；　　D．任意 $r-1$ 个向量线性无关．

第 5 章　矩阵的特征值与特征向量

5.1　特征值与特征向量

设 $\boldsymbol{A}$ 是 n 阶矩阵，$\boldsymbol{x}$ 是 n 维列向量，则 $\boldsymbol{Ax}$ 仍是 n 维列向量，但通常与原来的向量 $\boldsymbol{x}$ 有很大差异．现在考虑 $\boldsymbol{Ax}$ 与 $\boldsymbol{x}$ 能否成比例（对应分量成比例）．如果存在一个特殊的数 λ 与一个特殊的非零向量 $\boldsymbol{x}$，使

$$\boldsymbol{Ax}=\lambda\boldsymbol{x}\ ,\tag{5.1}$$

那么称 λ 是矩阵 $\boldsymbol{A}$ 的**特征值**，$\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的对应于特征值 λ 的**特征向量**．

为了求矩阵的特征值与特征向量，将（5.1）式移项，并提取公因子得到

$$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}\ .\tag{5.2}$$

这是齐次线性方程组的矩阵形式，它的系数矩阵是方阵 $(\lambda\boldsymbol{E}-\boldsymbol{A})$，齐次线性方程组有非零解 $\boldsymbol{x}$ 的充分必要条件是系数行列式等于零（参看定理 1.4），即

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0\ .\tag{5.3}$$

由（5.3）式可求出特征值 λ，因此方程（5.3）称为矩阵 $\boldsymbol{A}$ 的**特征方程**．求得特征值 λ 后，就可以求解齐次线性方程组（5.2）了．具体做法是运用矩阵消元解法，对系数矩阵作初等行变换，然后读出基础解系．

例 5.1　求矩阵 $\boldsymbol{A}=\begin{pmatrix}3 & 1\\ 5 & -1\end{pmatrix}$ 的特征值和特征向量．

解　先计算行列式：

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\left|\begin{pmatrix}\lambda & 0\\ 0 & \lambda\end{pmatrix}-\begin{pmatrix}3 & 1\\ 5 & -1\end{pmatrix}\right|=\begin{vmatrix}\lambda-3 & -1\\ -5 & \lambda+1\end{vmatrix}=(\lambda-3)(\lambda+1)-5=\lambda^2-2\lambda-8\ ,$$

于是特征方程即一元二次方程 $\lambda^2-2\lambda-8=0$．求得两个特征值为 $\lambda_1=-2$，$\lambda_2=4$．

当 $\lambda_1=-2$ 时，解齐次线性方程组（5.2），即作初等行变换：

$$(\lambda_1\boldsymbol{E}-\boldsymbol{A})=\begin{pmatrix}-2-3 & -1\\ -5 & -2+1\end{pmatrix}=\begin{pmatrix}-5 & \langle-1\rangle\\ -5 & -1\end{pmatrix}\to\begin{pmatrix}5 & 1\\ 0 & 0\end{pmatrix}.$$

最后的矩阵中，第 1 列是非主元列，于是读出基础解系 $\boldsymbol{\alpha}_1=\begin{pmatrix}1\\ -5\end{pmatrix}$．对应于特征值 $\lambda_1=-2$ 的全部特征向量是 $c_1\boldsymbol{\alpha}_1$，c_1 是任意非零常数．

当 $\lambda_2=4$ 时，类似地作初等行变换：

$$(\lambda_2 \boldsymbol{E}-\boldsymbol{A})=\begin{pmatrix}1 & \langle -1\rangle\\ -5 & 5\end{pmatrix}\to\begin{pmatrix}-1 & 1\\ 0 & 0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_2=\begin{pmatrix}1\\1\end{pmatrix}$，对应于特征值 $\lambda_2=4$ 的全部特征向量是 $c_2\boldsymbol{\alpha}_2$，c_2 是任意非零常数.

当 $\boldsymbol{A}$ 是 n 阶矩阵时，$|\lambda \boldsymbol{E}-\boldsymbol{A}|$ 是 n 阶行列式，其对角线上元素都含有未知数 λ，而其他元素都是矩阵 $\boldsymbol{A}$ 的对应元素的相反数，因此它的展开式是 n 次多项式，称之为 $\boldsymbol{A}$ 的**特征多项式**. 相应地，称矩阵 $(\lambda \boldsymbol{E}-\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的**特征矩阵**. 特征方程（5.3）是 n 次代数方程，它有 n 个根（可能有重根或复数根）.

齐次线性方程组（5.2）当 λ 满足（5.3）式时一定有基础解系. 构成基础解系的特征向量称为矩阵 $\boldsymbol{A}$ 的**基础特征向量**. 如例 5.1 中的矩阵 $\boldsymbol{A}$ 有 2 个基础特征向量 $\boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_2$. 基础特征向量的线性组合给出了全部特征向量.

例 5.2 求矩阵 $\boldsymbol{A}=\begin{pmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 5\end{pmatrix}$ 的特征值和特征向量.

解 先解特征方程 $|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2 & -2 & 2\\ -2 & \lambda-5 & 4\\ 2 & 4 & \lambda-5\end{vmatrix}=0$.

为了简化计算，可先将行列式的第 3 行加到第 2 行，然后提取公因子 $(\lambda-1)$，再计算行列式，可得方程 $(\lambda-1)^2(\lambda-10)=0$，所以特征值是 $\lambda_1=\lambda_2=1$，$\lambda_3=10$.

当 $\lambda=1$ 时，对特征矩阵作初等行变换：

$$(\lambda \boldsymbol{E}-\boldsymbol{A})=\begin{pmatrix}\langle -1\rangle & -2 & 2\\ -2 & -4 & 4\\ 2 & 4 & -4\end{pmatrix}\to\begin{pmatrix}1 & 2 & -2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_1=(-2,\ 1,\ 0)^{\mathrm{T}}$，$\boldsymbol{\alpha}_2=(2,\ 0,\ 1)^{\mathrm{T}}$.

当 $\lambda=10$ 时，对特征矩阵作初等行变换：

$$(\lambda \boldsymbol{E}-\boldsymbol{A})=\begin{pmatrix}8 & -2 & \langle 2\rangle\\ -2 & 5 & 4\\ 2 & 4 & 5\end{pmatrix}\to\begin{pmatrix}4 & -1 & 1\\ -18 & \langle 9\rangle & 0\\ -18 & 9 & 0\end{pmatrix}\to\begin{pmatrix}-2 & 1 & 0\\ 2 & 0 & 1\\ 0 & 0 & 0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_3=(1,\ 2,\ -2)^{\mathrm{T}}$.

对应于 $\lambda=1$ 的全部特征向量是 $c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2$，其中 c_1,c_2 是任意不全为零的常数；对应于 $\lambda=10$ 的全部特征向量是 $c_3\boldsymbol{\alpha}_3$，其中 c_3 是任意非零常数.

应该指出的是，线性组合必须在对应于同一特征值的基础特征向量之中进行，比如本例中当 $c_1c_2c_3\neq 0$ 时，线性组合 $c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2+c_3\boldsymbol{\alpha}_3$ 不再是矩阵 $\boldsymbol{A}$ 的特征向量．

例 5.3　求矩阵 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$ 的特征值和特征向量．

解　先解特征方程

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{vmatrix}=(\lambda-2)(\lambda-1)^2=0,$$

得特征值 $\lambda_1=2$，$\lambda_2=\lambda_3=1$．

当 $\lambda=2$ 时，对特征矩阵作初等行变换：

$$\begin{pmatrix}3 & -1 & 0\\ 4 & -1 & 0\\ \langle -1\rangle & 0 & 0\end{pmatrix}\to\begin{pmatrix}0 & \langle -1\rangle & 0\\ 0 & -1 & 0\\ 1 & 0 & 0\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_1=(0,\ 0,\ 1)^{\mathrm{T}}$．

当 $\lambda=1$ 时，对特征矩阵作初等行变换：

$$\begin{pmatrix}2 & \langle -1\rangle & 0\\ 4 & -2 & 0\\ -1 & 0 & -1\end{pmatrix}\to\begin{pmatrix}-2 & 1 & 0\\ 0 & 0 & 0\\ -1 & 0 & \langle -1\rangle\end{pmatrix}\to\begin{pmatrix}-2 & 1 & 0\\ 1 & 0 & 1\\ 0 & 0 & 0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_2=(1,\ 2,\ -1)^{\mathrm{T}}$．

全部特征向量是 $c_1\boldsymbol{\alpha}_1$（对应于 $\lambda=2$）和 $c_2\boldsymbol{\alpha}_2$（对应于 $\lambda=1$），其中 c_1,c_2 是任意非零常数．

5.2　矩阵的相似与矩阵的对角化

设 $\boldsymbol{A}$，$\boldsymbol{B}$ 都是 n 阶矩阵，若存在 n 阶可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B},\tag{5.4}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似**，或称 $\boldsymbol{A}$ 经过**相似变换** $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 变换成 $\boldsymbol{B}$，记为 $\boldsymbol{A}\backsim\boldsymbol{B}$．

矩阵的相似具有以下性质：

（1）**反身性**　$\boldsymbol{A}\backsim\boldsymbol{A}$．

因为 $\boldsymbol{E}^{-1}\boldsymbol{A}\boldsymbol{E}=\boldsymbol{A}$，所以 $\boldsymbol{A}\backsim\boldsymbol{A}$．

（2）**对称性** 若 $A \backsim B$，则 $B \backsim A$．

$A \backsim B$ 意味着 $P^{-1}AP = B$，等式两端左乘 P、右乘 P^{-1}，得 $(P^{-1})^{-1}B(P^{-1}) = A$，所以 $B \backsim A$．

（3）**传递性** 若 $A \backsim B$，$B \backsim C$，则 $A \backsim C$．

由条件 $P^{-1}AP = B$，$Q^{-1}BQ = C$，以前式代入后式得：$Q^{-1}P^{-1}APQ = C$，也就是 $(PQ)^{-1}A(PQ) = C$，所以 $A \backsim C$．

要判断两个矩阵 A，B 是否相似比较困难，因为很难直接找到（5.4）式中的变换矩阵 P．不过我们可以通过相似变换把 A，B 都变为比较简单的矩阵，再通过传递性说明它们是否相似．最简单的矩阵莫过于对角矩阵．

若矩阵 A 能通过相似变换化为对角矩阵，则称矩阵 A 可以**对角化**．对角化是指存在可逆矩阵 P、对角矩阵 Λ，使 $P^{-1}AP = \Lambda$，即 $A \backsim \Lambda$．

例 5.4 将矩阵 $A = \begin{pmatrix} 3 & 1 \\ 5 & -1 \end{pmatrix}$ 对角化．

解 在例 5.1 中，已求得矩阵 A 的两个特征向量 α_1，α_2（分别对应特征值 $\lambda_1 = -2$，$\lambda_2 = 4$），按照（5.1）式，应有 $A\alpha_1 = \lambda_1\alpha_1$，$A\alpha_2 = \lambda_2\alpha_2$，拼成矩阵得到

$$A(\alpha_1,\ \alpha_2) = (A\alpha_1,\ A\alpha_2) = (\lambda_1\alpha_1,\ \lambda_2\alpha_2) = (\alpha_1,\ \alpha_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

记 $P = (\alpha_1,\ \alpha_2) = \begin{pmatrix} 1 & 1 \\ -5 & 1 \end{pmatrix}$，$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$，即有 $AP = P\Lambda$．容易看出这里的矩阵 P 可逆，所以 $P^{-1}AP = \Lambda$，矩阵 A 已经对角化．

本例的结果能够直接推广到 n 阶矩阵，设 A 是 n 阶矩阵，$\alpha_1, \alpha_2, \mathrm{L}, \alpha_n$ 是 A 的分别对应于特征值 $\lambda_1, \lambda_2, \mathrm{L}, \lambda_n$ 的 n 个特征向量，它们可以拼成 n 阶方阵．记

$$P = (\alpha_1,\ \alpha_2,\ \mathrm{L},\ \alpha_n),\quad \Lambda = \begin{pmatrix} \lambda_1 & 0 & \mathrm{L} & 0 \\ 0 & \lambda_2 & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & \mathrm{L} & \lambda_n \end{pmatrix},$$

根据（5.1）式和矩阵的乘法规则，必有

$$AP = P\Lambda \tag{5.5}$$

若矩阵 P 可逆，则 A 可对角化，即 $P^{-1}AP = \Lambda$．n 阶方阵可逆与满秩是同一概念，而满秩即意味着列向量组 $\alpha_1, \alpha_2, \mathrm{L}, \alpha_n$ 线性无关（参看（4.3）式）．然而特征向量并非总是线性无关的，不过我们从例 5.1～例 5.3 能够猜测到以下定理．

定理 5.1 方阵 A 的所有基础特征向量线性无关．

证明定理可在（4.2）式的基础上反复运用（5.1）式，因证明较长，此处从略．由于线性无关的 n 维向量个数不会超过 n 个（参看定理 4.5），所以矩阵 A 的基础特征向量的个数

不会超过 A 的阶数．根据（5.5）式及矩阵 P 可逆的条件可得以下定理．

定理 5.2　n 阶方阵 A 可以对角化的充分必要条件是它的基础特征向量的个数等于 n．

下面将矩阵对角化问题作几点归纳：

（1）若矩阵 A 的特征方程（5.3）没有重根，则 A 一定可以对角化．

这是因为每个特征值至少提供一个基础特征向量，n 个不同的特征值恰好提供 n 个基础特征向量．

（2）若矩阵 A 的特征方程（5.3）有重根，则 A 可否对角化要看齐次线性方程组（5.2）的求解情况．当基础特征向量个数等于 n 时，可以对角化，如例 5.2；当基础特征向量的个数不足时，不可以对角化，即矩阵 A 不可能与对角矩阵相似，如例 5.3．

（3）如果矩阵 A 可以对角化，那么相似变换矩阵 P 由 A 的 n 个基础特征向量拼接而成．对角矩阵 Λ 作为相似变换的结果，由对应的 n 个特征值构成，而且这些特征值的排列顺序与矩阵 P 中特征向量的排列顺序一致．

矩阵的相似变换为求矩阵的乘幂带来了方便．比如利用（5.4）式，可得：

$$\boldsymbol{B}^3=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})=\boldsymbol{P}^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{A}(\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}^3\boldsymbol{P}\text{，}$$

从而 $\boldsymbol{A}^3=\boldsymbol{P}\boldsymbol{B}^3\boldsymbol{P}^{-1}$．一般地，如果 $\boldsymbol{A}\backsim\boldsymbol{B}$，则有乘幂公式

$$\boldsymbol{A}^n=\boldsymbol{P}\boldsymbol{B}^n\boldsymbol{P}^{-1}\text{．}\tag{5.6}$$

式中 P 是相似变换矩阵．直接计算乘幂 A^n，运算量很大，而当矩阵 B 简单时，B^n 的计算却很方便．尤其是当 B 为对角矩阵时，计算 B^n 只要将主对角线上的元素乘方即可（参看（2.4）式）．

例 5.5　求例 5.2 中矩阵 $\boldsymbol{A}=\begin{pmatrix}2&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}$ 的乘幂 $\boldsymbol{A}^n$．

解　已求得矩阵 A 的特征值及 3 个基础特征向量（参看例 5.2），故令

$$\boldsymbol{P}=\begin{pmatrix}-2&2&1\\1&0&2\\0&1&-2\end{pmatrix}\text{，}\quad\boldsymbol{\Lambda}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&10\end{pmatrix}\text{，}$$

由定理 5.2 知，A 可对角化且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$．容易求得逆矩阵 $\boldsymbol{P}^{-1}=\frac{1}{9}\begin{pmatrix}-2&5&4\\2&4&5\\1&2&-2\end{pmatrix}$，应用公式（5.6）得

$$\boldsymbol{A}^n=\boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}=\frac{1}{9}\begin{pmatrix}-2&2&1\\1&0&2\\0&1&-2\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&10^n\end{pmatrix}\begin{pmatrix}-2&5&4\\2&4&5\\1&2&-2\end{pmatrix}$$

$$=\begin{pmatrix} 1+a_n & 2a_n & -2a_n \\ 2a_n & 1+4a_n & -4a_n \\ -2a_n & -4a_n & 1+4a_n \end{pmatrix},$$

其中 $a_n=\dfrac{10^n-1}{9}$．将这个结果稍作变形，可得简单的计算公式

$$\boldsymbol{A}^n=\boldsymbol{E}+a_n(\boldsymbol{A}-\boldsymbol{E}).$$

从本例能够看出，用公式（5.6）求乘幂不仅运算简单，而且便于分析计算结果．应该指出的是，由于齐次线性方程组（5.2）的基础解系不惟一，所以这里的相似变换矩阵 $\boldsymbol{P}$ 也不惟一，但乘幂计算结果是一样的．

5.3 实对称矩阵的对角化

5.3.1 向量的内积与正交矩阵

对于两个 n 维列向量

$$\boldsymbol{\alpha}=(a_1,\ a_2,\ \mathrm{L},\ a_n)^{\mathrm{T}},\quad \boldsymbol{\beta}=(b_1,\ b_2,\ \mathrm{L},\ b_n)^{\mathrm{T}},$$

称以下运算为向量 $\boldsymbol{\alpha}$，$\boldsymbol{\beta}$ 的**内积**：

$$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=a_1b_1+a_2b_2+\mathrm{L}\ +a_nb_n. \tag{5.7}$$

向量的内积表现为对应的分量“两两相乘再相加”，在第 1 章和第 2 章中，我们曾把这种运算称为组合运算．内积显然有对称性，即

$$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}.$$

内积可以看作 $1\times n$ 矩阵和 $n\times 1$ 矩阵的乘积，因此满足矩阵的运算规则：

$$(k\boldsymbol{\alpha})^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\alpha}^{\mathrm{T}}(k\boldsymbol{\beta})=k\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}\quad（k\text{ 为常量}），$$

$$(\boldsymbol{\alpha}+\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{\gamma}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\gamma}+\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\gamma}\quad（\boldsymbol{\gamma}\text{ 也是 }n\text{ 维向量}）.$$

如果向量的分量都是实数，那么

$$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}=a_1^2+a_2^2+\mathrm{L}\ +a_n^2\geqslant 0,$$

当且仅当 $\boldsymbol{\alpha}=\boldsymbol{0}$ 时等号成立．于是定义向量 $\boldsymbol{\alpha}$ 的**长度**为

$$|\boldsymbol{\alpha}|=\sqrt{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}}.$$

$|\boldsymbol{\alpha}|$ 又叫做向量 $\boldsymbol{\alpha}$ 的**模**．满足 $|\boldsymbol{\alpha}|=1$ 的向量 $\boldsymbol{\alpha}$ 称为**单位向量**．

对于一般的非零向量 $\boldsymbol{\alpha}$，$\dfrac{1}{|\boldsymbol{\alpha}|}\boldsymbol{\alpha}$ 一定是单位向量，称为把向量 $\boldsymbol{\alpha}$ **单位化**．比如，欲将向量 $\boldsymbol{\alpha}=(1,\ -2,\ 2,\ 4)^{\mathrm{T}}$ 单位化，只需以 $|\boldsymbol{\alpha}|=\sqrt{25}=5$ 遍除各分量，便得到单位向量

$$\boldsymbol{\alpha}_0=\left(\frac{1}{5},\ -\frac{2}{5},\ \frac{2}{5},\ \frac{4}{5}\right)^{\mathrm{T}}.$$

如果 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=0$，则称向量 $\boldsymbol{\alpha}$，$\boldsymbol{\beta}$ 是**正交**的.

向量的内积与正交是几何向量（三维向量）数量积、垂直等概念的直接推广．与正交概念相联系的是正交矩阵.

如果 n 阶方阵 $\boldsymbol{Q}$ 满足 $\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{E}$，则称 $\boldsymbol{Q}$ 为**正交矩阵**．例如

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\ \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

都是正交矩阵．显然，如果 $\boldsymbol{Q}$ 是正交矩阵，那么 $\boldsymbol{Q}^{-1}=\boldsymbol{Q}^{\mathrm{T}}$．可见正交矩阵的逆矩阵是很容易得到的.

因为证明逆矩阵只需做一个乘法（参看定理 2.3），所以判定 $\boldsymbol{Q}$ 是否为正交矩阵，只需检验 $\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{E}$ 或者 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{E}$．乘积 $\boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}}$，$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}$ 实际上分别是矩阵 $\boldsymbol{Q}$ 的行向量、列向量之间的内积运算，于是有以下正交矩阵的判定定理.

定理 5.3　n 阶矩阵 $\boldsymbol{Q}$ 是正交矩阵的充分必要条件是 $\boldsymbol{Q}$ 的所有行（列）向量都是单位向量，而且两两正交.

这里所谓**两两正交**指的是任意两个向量都正交．验证 n 阶矩阵的正交性，需要检验 $\dfrac{n(n-1)}{2}$ 对向量的正交性.

5.3.2　实对称矩阵的特征值与特征向量

设 n 阶矩阵 $\boldsymbol{A}$ 的所有元素都是实数，并且 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为**实对称矩阵**.

一般矩阵的相似对角化是有条件的（参看定理 5.2），而实对称矩阵有完满的结果：实对称矩阵的特征值都是实数，而且一定可以对角化，即 n 阶实对称矩阵一定有 n 个基础特征向量（证明略）.

关于实对称矩阵的对角化，还有如下进一步的结论.

定理 5.4　设 $\boldsymbol{A}$ 是实对称矩阵，则 $\boldsymbol{A}$ 的对应于不同特征值的特征向量必定正交，因而存在正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$，其中 $\boldsymbol{\Lambda}$ 是由 $\boldsymbol{A}$ 的特征值构成的对角矩阵.

证　只证前半部分，后半部分通过例子说明．设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 是对应于不同特征值 λ_1,λ_2 的特征向量，即 $\boldsymbol{A}\boldsymbol{\alpha}=\lambda_1\boldsymbol{\alpha}$，$\boldsymbol{A}\boldsymbol{\beta}=\lambda_2\boldsymbol{\beta}$，于是

$$\lambda_1\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=(\lambda_1\boldsymbol{\alpha})^{\mathrm{T}}\boldsymbol{\beta}=(\boldsymbol{A}\boldsymbol{\alpha})^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\lambda_2\boldsymbol{\beta}=\lambda_2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta},$$

移项得 $(\lambda_1-\lambda_2)\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=0$．由于 $\lambda_1\neq\lambda_2$，所以必有 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=0$，即 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 正交.

定理 5.4 表明，实对称矩阵不仅一定可以对角化，而且可以通过正交的相似变换矩阵 $\boldsymbol{Q}$ 实现对角化.

例 5.6 设 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\2&1&3\\3&3&6\end{pmatrix}$，求正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$ 是对角矩阵.

解 先解特征方程

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1&-2&-3\\-2&\lambda-1&-3\\-3&-3&\lambda-6\end{vmatrix}=0.$$

将行列式的第 1 行减去第 2 行，再计算行列式易得 $\lambda(\lambda+1)(\lambda-9)=0$，所以 $\boldsymbol{A}$ 的特征值是 $\lambda_1=0$，$\lambda_2=-1$，$\lambda_3=9$.

当 $\lambda=0$ 时，特征矩阵及其初等行变换为

$$\begin{pmatrix}\langle-1\rangle&-2&-3\\-2&-1&-3\\-3&-3&-6\end{pmatrix}\to\begin{pmatrix}1&2&3\\0&3&\langle3\rangle\\0&3&3\end{pmatrix}\to\begin{pmatrix}1&-1&0\\0&1&1\\0&0&0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_1=(1,\ \ 1,\ \ -1)^{\mathrm{T}}$.

当 $\lambda=-1$ 时，特征矩阵及其初等行变换为

$$\begin{pmatrix}\langle-2\rangle&-2&-3\\-2&-2&-3\\-3&-3&-7\end{pmatrix}\to\begin{pmatrix}1&1&1.5\\0&0&0\\0&0&\langle-2.5\rangle\end{pmatrix}\to\begin{pmatrix}1&1&0\\0&0&1\\0&0&0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_2=(1,\ \ -1,\ \ 0)^{\mathrm{T}}$.

当 $\lambda=9$ 时，特征矩阵及其初等行变换为

$$\begin{pmatrix}8&-2&-3\\-2&8&-3\\-3&-3&\langle3\rangle\end{pmatrix}\to\begin{pmatrix}\langle5\rangle&-5&0\\-5&5&0\\-1&-1&1\end{pmatrix}\to\begin{pmatrix}1&-1&0\\0&-2&1\\0&0&0\end{pmatrix},$$

读出基础解系 $\boldsymbol{\alpha}_3=(1,\ \ 1,\ \ 2)^{\mathrm{T}}$.

矩阵 $\boldsymbol{A}$ 的全部特征向量是 $c_1\boldsymbol{\alpha}_1, c_2\boldsymbol{\alpha}_2, c_3\boldsymbol{\alpha}_3$，其中 c_1, c_2, c_3 是任意非零常数. 根据基础解系及向量线性关系的理论可知，基础解系不惟一，任意确定 c_1, c_2, c_3 的值，都可得到一组（3 个）基础特征向量. 现在需要寻找一组两两正交并且长度为 1 的特征向量，以便拼成本例所求的正交矩阵 $\boldsymbol{Q}$（参看定理 5.3）.

根据定理 5.4，特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 必然两两正交（也可通过内积运算加以验证），所以只要将它们单位化即可. 得到单位化向量，再拼成正交矩阵

$$\boldsymbol{Q}=\left(\frac{\boldsymbol{\alpha}_1}{|\boldsymbol{\alpha}_1|},\ \frac{\boldsymbol{\alpha}_2}{|\boldsymbol{\alpha}_2|},\ \frac{\boldsymbol{\alpha}_3}{|\boldsymbol{\alpha}_3|}\right)=\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}}\end{pmatrix}=\frac{1}{\sqrt{6}}\begin{pmatrix}\sqrt{2} & \sqrt{3} & 1\\ \sqrt{2} & -\sqrt{3} & 1\\ -\sqrt{2} & 0 & 2\end{pmatrix},$$

即为所求，相似对角矩阵为

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 9\end{pmatrix}.$$

例 5.7　对于例 5.2 中的矩阵 $\boldsymbol{A}$，求正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$ 是对角矩阵.

解　在例 5.2 中，已求得特征值是 1，1，10，对应的特征向量是

$$\boldsymbol{\alpha}_1=(-2,\ 1,\ 0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(2,\ 0,\ 1)^{\mathrm{T}},\quad \boldsymbol{\alpha}_3=(1,\ 2,\ -2)^{\mathrm{T}},$$

其中 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 都与 $\boldsymbol{\alpha}_3$ 正交（根据定理 5.4 或直接验证），但 $\boldsymbol{\alpha}_1$ 与 $\boldsymbol{\alpha}_2$ 不正交．对应于特征值 1 的全部特征向量是 $c_1\boldsymbol{\alpha}_1+c_2\boldsymbol{\alpha}_2$（$c_1,c_2$ 是任意不全为零的常数），现在需要从中找出两个正交的特征向量 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$，这就是所谓的**正交化**，通常称作将向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 正交化．正交化过程如下：

令 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$，$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-k_1\boldsymbol{\beta}_1$，其中 k_1 是待定系数．由条件 $\boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\beta}_2=0$，即 $\boldsymbol{\beta}_1^{\mathrm{T}}(\boldsymbol{\alpha}_2-k_1\boldsymbol{\beta}_1)=0$，可确定

$$k_1=\frac{\boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\alpha}_2}{\boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\beta}_1}=-\frac{4}{5}.$$

于是得到

$$\boldsymbol{\beta}_1=\begin{pmatrix}-2\\ 1\\ 0\end{pmatrix},\quad \boldsymbol{\beta}_2=\begin{pmatrix}2\\ 0\\ 1\end{pmatrix}+\frac{4}{5}\begin{pmatrix}-2\\ 1\\ 0\end{pmatrix}=\frac{1}{5}\begin{pmatrix}2\\ 4\\ 5\end{pmatrix}.$$

最后将特征向量 $\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，$\boldsymbol{\alpha}_3$ 单位化后拼成矩阵并提取公因子，得

$$\boldsymbol{Q}=\left(\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|},\ \frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},\ \frac{\boldsymbol{\alpha}_3}{|\boldsymbol{\alpha}_3|}\right)=\frac{1}{3\sqrt{5}}\begin{pmatrix}-6 & 2 & \sqrt{5}\\ 3 & 4 & 2\sqrt{5}\\ 0 & 5 & -2\sqrt{5}\end{pmatrix},$$

即为所求的正交矩阵．相似对角矩阵为

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 10\end{pmatrix}.$$

2 个向量的正交化过程比较简单．如果是 3 个向量 $\alpha_1, \alpha_2, \alpha_3$，那么需要构造 3 个相互正交的向量 $\beta_1, \beta_2, \beta_3$，其中 β_1, β_2 的求法与例 5.7 相同，然后令 $\beta_3 = \alpha_3 - k_1\beta_1 - k_2\beta_2$，再由条件 $\beta_1^{\mathrm{T}}\beta_3 = 0$ 和 $\beta_2^{\mathrm{T}}\beta_3 = 0$ 确定待定系数 k_1, k_2．至于更多向量的正交化过程可依此类推．这种正交化方法称为**施密特正交化方法**．施密特正交化过程也可以通过矩阵的初等变换来实现．

例 5.8 将向量组 $\alpha_1 = (1,\ 2,\ 2,\ -1)^{\mathrm{T}}$，$\alpha_2 = (1,\ 1,\ -5,\ 3)^{\mathrm{T}}$，$\alpha_3 = (3,\ 2,\ 8,\ -7)^{\mathrm{T}}$ 正交化．

解 先将 3 个向量拼成矩阵 $\boldsymbol{A} = (\alpha_1,\ \alpha_2,\ \alpha_3)$，再计算内积矩阵

$$\boldsymbol{G} = (g_{ij}) = \begin{pmatrix} 10 & -10 & 30 \\ -10 & 36 & -56 \\ 30 & -56 & 126 \end{pmatrix},$$

其中 $g_{ij} = \alpha_i^{\mathrm{T}}\alpha_j$ 是内积（$i, j = 1,\ 2,\ 3$）．为了寻找 $\alpha_1, \alpha_2, \alpha_3$ 适当的线性组合，要对矩阵 $\boldsymbol{A}$ 作初等列变换，所以将 2 个矩阵上、下拼接成为 $\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{A} \end{pmatrix}$，经过初等列变换，把上半部分 $\boldsymbol{G}$ 的主对角线右上方元素变为 0 的同时，下半部分 $\boldsymbol{A}$ 的列向量便已两两正交，即

$$\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{A} \end{pmatrix} = \begin{pmatrix} \langle 10 \rangle & -10 & 30 \\ -10 & 36 & -56 \\ 30 & -56 & 126 \\ 1 & 1 & 3 \\ 2 & 1 & 2 \\ 2 & -5 & 8 \\ -1 & 3 & -7 \end{pmatrix} \to \begin{pmatrix} 10 & 0 & 0 \\ -10 & \langle 26 \rangle & -26 \\ 30 & -26 & 36 \\ 1 & 2 & 0 \\ 2 & 3 & -4 \\ 2 & -3 & 2 \\ -1 & 2 & -4 \end{pmatrix} \to \begin{pmatrix} 10 & 0 & 0 \\ -10 & 26 & 0 \\ 30 & -26 & 10 \\ 1 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -3 & -1 \\ -1 & 2 & -2 \end{pmatrix}.$$

经过以上正交化，得到的向量为

$$\beta_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ -1 \end{pmatrix},\quad \beta_2 = \begin{pmatrix} 2 \\ 3 \\ -3 \\ 2 \end{pmatrix},\quad \beta_3 = \begin{pmatrix} 2 \\ -1 \\ -1 \\ -2 \end{pmatrix}.$$

习 题 五

1．求下列矩阵的特征值和特征向量：

(1) $\begin{pmatrix} 0 & 2 \\ -3 & 5 \end{pmatrix}$；　(2) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$；　(3) $\begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$；

(4) $\begin{pmatrix} 5 & 6 & -3 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix}$；　(5) $\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix}$.

2．设 $\boldsymbol{A}=\begin{pmatrix} -1 & 4 \\ 2 & -3 \end{pmatrix}$，求 $\boldsymbol{A}^n$ ．

3．第 1 题中的各矩阵能否对角化？若能，则写出相应的变换矩阵 $\boldsymbol{P}$ 及对角矩阵 $\boldsymbol{\Lambda}$ ．

4．已知二阶矩阵 $\boldsymbol{A},\boldsymbol{B}$ 与同一个对角矩阵 $\boldsymbol{\Lambda}$ 相似，并且

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1}\boldsymbol{A}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}=\boldsymbol{\Lambda}，\quad \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}^{-1}\boldsymbol{B}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}=\boldsymbol{\Lambda}．$$

问是否必有 $\boldsymbol{A}\sim\boldsymbol{B}$ ？若是，则求出矩阵 $\boldsymbol{P}$ ，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ ．

5．判断下列矩阵是不是正交矩阵？

(1) $\begin{pmatrix} \frac{2}{7} & \frac{6}{7} & \frac{3}{7} \\ \frac{6}{7} & -\frac{3}{7} & \frac{2}{7} \\ \frac{3}{7} & \frac{2}{7} & -\frac{6}{7} \end{pmatrix}$；　(2) $\begin{pmatrix} \frac{1}{9} & -\frac{8}{9} & -\frac{4}{9} \\ -\frac{8}{9} & \frac{1}{9} & -\frac{4}{9} \\ -\frac{4}{9} & -\frac{4}{9} & \frac{7}{9} \end{pmatrix}$；　(3) $\begin{pmatrix} \frac{1}{5} & -\frac{2}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ -\frac{2}{5} & \frac{1}{5} & \frac{2}{5} \end{pmatrix}$.

6．对于下列矩阵 $\boldsymbol{A}$ ，试求正交矩阵 $\boldsymbol{Q}$ 和对角矩阵 $\boldsymbol{\Lambda}$ ，使得 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}$ ：

(1) $\boldsymbol{A}=\begin{pmatrix} 3 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 5 \end{pmatrix}$；　(2) $\boldsymbol{A}=\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$；　(3) $\boldsymbol{A}=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$.

7．单项选择题

(1) 设 $\boldsymbol{A}$ 是可逆矩阵，则在 $\boldsymbol{A}$ 的特征值中，（　）

A．都是相同的数；　B．都是非零数；　C．可能有数 0；　D．一定有数 0．

(2) 设 $\boldsymbol{A}$ 是非奇异矩阵，2 是 $\boldsymbol{A}$ 的特征值，则（　）

A．一定有非零向量 $\boldsymbol{x}$ ，使得 $\boldsymbol{A}^{-1}\boldsymbol{x}=2\boldsymbol{x}$ ；　B．一定有非零向量 $\boldsymbol{x}$ ，使得 $2\boldsymbol{A}^{-1}\boldsymbol{x}=\boldsymbol{x}$ ；

C．不存在非零向量 $\boldsymbol{x}$ ，使得 $\boldsymbol{A}^{-1}\boldsymbol{x}=2\boldsymbol{x}$ ；　D．不存在非零向量 $\boldsymbol{x}$ ，使得 $2\boldsymbol{A}^{-1}\boldsymbol{x}=\boldsymbol{x}$ ．

(3) 下列向量中，哪一个是矩阵 $\boldsymbol{A}=\begin{pmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & 3 \end{pmatrix}$ 的特征向量？（　）

A．$(2,\ 1,\ 1)^{\mathrm{T}}$ ；　B．$(2,\ 1,\ -1)^{\mathrm{T}}$　C．$(2,\ -1,\ 1)^{\mathrm{T}}$ ；　D．$(-2,\ 1,\ 1)^{\mathrm{T}}$ ．

(4) 设 λ_1,λ_2 是矩阵 $\boldsymbol{A}$ 的两个不同的特征值，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 分别是对应于 λ_1,λ_2 的特征向量．在下列何种情况下，$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$ 必定是 $\boldsymbol{A}$ 的特征向量．（　）

A．$k_1\neq 0$ 且 $k_2\neq 0$ ；　B．$k_1\neq 0$ 或 $k_2\neq 0$ ；

C．$k_1+k_2\neq 0$ 且 $k_1k_2=0$ ；　D．$k_1+k_2\neq 0$ 或 $k_1k_2=0$ ．

(5) 矩阵 A 和 B 相似的充分必要条件是 ( )

A．存在非奇异矩阵 C，使 $AC=BC$；　　B．$AB=BA$；

C．存在非奇异矩阵 C，使 $AC=CB$；　　D．以上都不对．

(6) n 阶矩阵 A 的基础特征向量的个数 ( )

A．少于 n 个；　　B．不超过 n 个；　　C．等于 n 个；　　D．超过 n 个．

(7) n 阶矩阵 A 可以对角化的充分条件是 ( )

A．A 可逆；　　B．能找出 n 个 A 的特征向量；

C．A 的特征值都非零；　　D．A 的特征值都不相同．

(8) A 是三阶矩阵，$\lambda_1,\lambda_2,\lambda_3$ 为其特征值，互不相同．在什么情况下，$\lim\limits_{n\to\infty}A^n=O$？ ( )

A．$1\leqslant|\lambda_1|\leqslant|\lambda_2|\leqslant|\lambda_3|$；　　B．$|\lambda_1|<1\leqslant|\lambda_2|\leqslant|\lambda_3|$；

C．$|\lambda_1|<|\lambda_2|<1\leqslant|\lambda_3|$；　　D．$|\lambda_1|<|\lambda_2|<|\lambda_3|<1$．

(9) 设 α 是 n 维单位列向量，即 $\alpha^{\mathrm{T}}\alpha=1$，$E$ 是 n 阶单位矩阵，则 $E-2\alpha\alpha^{\mathrm{T}}$ ( )

A．是对称矩阵，但非正交矩阵；　　B．是正交矩阵，但非对称矩阵；

C．是对称矩阵，也是正交矩阵；　　D．非对称矩阵，也非正交矩阵．

(10) 设 A 是实对称矩阵，则 ( )

A．A 的特征值各不相同；　　B．A 一定可以利用正交矩阵实现对角化；

C．A 的特征向量相互正交；　　D．A 不一定能利用正交矩阵实现对角化．

第 6 章　二 次 型

6.1　二次型及其标准形

6.1.1　二次方程与几何图形

在平面解析几何中，方程 $x^2+2y^2=1$ 表示椭圆；$x^2-2y^2=1$ 表示双曲线．这两个方程的左端都是关于 x, y 的二次齐次函数，它们都只含有平方项．含有乘积项的方程稍微复杂一些，如 $x^2+2xy+3y^2=1$．该方程通过配方可化为 $(x+y)^2+2y^2=1$，从而知道它表示椭圆．另一个方程 $x^2+4xy+3y^2=1$，经配方可知它表示双曲线．

在空间解析几何中，含有 3 个变量的二次方程

$$a_1x^2+a_2y^2+a_3z^2+a_4xy+a_5yz+a_6xz=1,$$

可能表示椭球面、双曲面或者椭圆柱面、双曲柱面等，到底表示什么曲面，与系数 a_i (i =1, 2,…,6)的构成有关．

看来，有必要对上述方程左端的二次齐次函数进行深入的研究，找出一些规律．

6.1.2　二次型的矩阵表示

研究二次函数，不能局限于 2 个或 3 个变量．当变量很多时，用带有下标的字母来表示较好．含有 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次函数的一般形式是

$$f(x_1, x_2, \cdots, x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j, \tag{6.1}$$

其中的各项 $a_{ij}x_ix_j$，当 $i=j$ 时表示**平方项**；当 $i\neq j$ 时表示**乘积项**．以后，将二次齐次函数简称为**二次型**．若所有系数 a_{ij} 都是实数，则称二次型（6.1）为**实二次型**．本章只讨论实二次型．

在(6.1)式中，每种同类的乘积项都出现了两次，比如 $i=1, j=2$ 时的 $a_{12}x_1x_2$ 和 $i=2, j=1$ 时的 $a_{21}x_2x_1$．它们作为同类项可以合并为 $(a_{12}+a_{21})x_1x_2$，当两个系数的和不变时，并不计较这两个系数各自取什么值．为确定起见，我们约定

$$a_{ij}=a_{ji}(i, j=1, 2, \cdots, n) \tag{6.2}$$

对于已经合并同类项的二次型，可按（6.2）式将乘积项拆成两个相等的项，如

$$x_1^2+3x_1x_2+2x_2^2=x_1^2+1.5x_1x_2+1.5x_2x_1+2x_2^2\text{.}$$

二次型（6.1）共有 $n\times n$ 项，这么多的项，不便于书写，也不便于研究．如果用矩阵来表示二次型，则可显得简单明了．

例 6.1　把二次型 $f(x_1,x_2,x_3)=\sum_{i=1}^{3}\sum_{j=1}^{3}a_{ij}x_ix_j$ 化为矩阵形式．

解　利用矩阵的乘法规则，可得

$$\begin{aligned}f(x_1,x_2,x_3)&=x_1(a_{11}x_1+a_{12}x_2+a_{13}x_3)+x_2(a_{21}x_1+a_{22}x_2+a_{23}x_3)+x_3(a_{31}x_1+a_{32}x_2+a_{33}x_3)\\&=\begin{pmatrix}x_1 & x_2 & x_3\end{pmatrix}\begin{pmatrix}a_{11}x_1+a_{12}x_2+a_{13}x_3\\a_{21}x_1+a_{22}x_2+a_{23}x_3\\a_{31}x_1+a_{32}x_2+a_{33}x_3\end{pmatrix}\\&=\begin{pmatrix}x_1 & x_2 & x_3\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\text{，}\end{aligned}$$

其中 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$ 是变量列，$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}$ 是二次型的系数矩阵．按照约定（6.2），矩阵 $\boldsymbol{A}$ 是对称矩阵．

显然例 6.1 可以推广到有更多变量的二次型，即任意一个二次型（6.1），都可以表示为矩阵形式如下：

$$f(x_1,x_2,\mathrm{L}\ ,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\text{，}\qquad(6.3)$$

其中 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\mathrm{M}\\x_n\end{pmatrix}$，$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\mathrm{L}&a_{1n}\\a_{21}&a_{22}&\mathrm{L}&a_{2n}\\\mathrm{M}&\mathrm{M}&&\mathrm{M}\\a_{n1}&a_{n2}&\mathrm{L}&a_{nn}\end{pmatrix}$ 是对称矩阵．

（6.3）式表明，一个二次型和一个对称矩阵是相互对应的，矩阵 $\boldsymbol{A}$ 是二次型（6.1）的系数矩阵，简称为二次型的矩阵．矩阵 $\boldsymbol{A}$ 的元素与二次型的系数相互确定，二次型系数的确定规则是：

（1）矩阵 $\boldsymbol{A}$ 主对角线上的元素等于二次型的平方项系数；

（2）矩阵 $\boldsymbol{A}$ 的其他元素等于二次型的（已合并同类项的）乘积项系数的一半；

（3）矩阵 $\boldsymbol{A}$ 各元素的行标与列标，对应二次型各项变量的下标．

比如，利用这个规则，很容易确定二次型 $x_1x_2+3x_2^2$ 的矩阵是 $\begin{pmatrix}0&0.5\\0.5&3\end{pmatrix}$；反之，根据

矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & -3\\ 2 & 0 & 4\\ -3 & 4 & -1\end{pmatrix}$，立即可写出它所对应的二次型是

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=x_1^2-x_3^2+4x_1x_2-6x_1x_3+8x_2x_3\text{．}$$

6.1.3　二次型的标准形

前面曾经提到，可以对二次型进行配方化简．比如 $f(x_1,x_2)=x_1^2+4x_1x_2+3x_2^2=(x_1+2x_2)^2-x_2^2$，若令 $y_1=x_1+2x_2$，$y_2=x_2$，则二次型等价变形为平方项的代数和（简称**平方和**）形式 $f(x_1,x_2)=y_1^2-y_2^2$．平方和的形式简单，而且能突出二次型的某些本质特性，所以化平方和成了研究二次型的首要目标．二次型的平方和形式称为二次型的**标准形**．

例 6.2　将二次型 $f=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$ 化为标准形．

解　先将含 x_1 的项配方，再对余下的含 x_2 的项配方……以此类推．配方过程如下：

$$\begin{aligned}f&=x_1^2+2x_1(x_2+x_3)+(x_2+x_3)^2-(x_2+x_3)^2+2x_2^2+5x_3^2+6x_2x_3\\&=(x_1+x_2+x_3)^2+x_2^2+4x_2x_3+4x_3^2=(x_1+x_2+x_3)^2+(x_2+2x_3)^2\text{．}\end{aligned}$$

作变量替换

$$\begin{cases}y_1=x_1+x_2+x_3\\ y_2=\quad x_2+2x_3\\ y_3=\quad\quad x_3\end{cases}\text{，}$$

则二次型已经化为标准形 $f=y_1^2+y_2^2$．

例 6.3　将二次型 $f=2x_1x_2+2x_1x_3-6x_2x_3$ 化为标准形．

解　由于没有平方项，不能直接配方，所以先要设法变造出平方项．具体的方法是让某一个乘积项成为平方差，如令

$$\begin{cases}x_1=y_1+y_2\\ x_2=y_1-y_2\\ x_3=\quad\quad y_3\end{cases}\text{，}$$

代入可得 $f=2y_1^2-2y_2^2+2(y_1+y_2)y_3-6(y_1-y_2)y_3$，化简后配方，得

$$\begin{aligned}f&=2y_1^2-4y_1y_3+2y_3^2-2y_3^2-2y_2^2+8y_2y_3\\&=2(y_1-y_3)^2-2(y_2^2-4y_2y_3+4y_3^2)+8y_3^2-2y_3^2\\&=2(y_1-y_3)^2-2(y_2-2y_3)^2+6y_3^2\end{aligned}$$

再引入一组新的变量，令

$$\begin{cases} z_1 = y_1 - y_3 \\ z_2 = y_2 - 2y_3 \\ z_3 = y_3 \end{cases},$$

则二次型化成了标准形 $f = 2z_1^2 - 2z_2^2 + 6z_3^2$.

以上两例所采用的化二次型为标准形的方法称为**配方法**. 配方法总可以进行下去：有平方项时直接配方(参照例 6.2) 没有平方项只有乘积项时先‘造出’平方项(参照例 6.3).因此任何二次型都可以化为标准形.

一个二次型的标准形不是惟一的. 比如例 6.2 中，若先对含 x_2 的项配方，则可得到

$$f = 2\left(\frac{1}{2}x_1 + x_2 + \frac{3}{2}x_3\right)^2 + \frac{1}{2}(x_1 - x_3)^2 .$$

虽然也是平方和，但平方项的前置系数已不相同，而且引入新变量的替换关系式也与先前的不一致.

配方法的原理比较简单，很容易理解，但是其中的代数运算显得比较凌乱，尤其是当变量增多时，用配方法演算起来非常烦琐. 这是因为配方法仍然停留在初等代数之中，并未利用矩阵化技术的优势.

6.2 二次型的线性变换与惯性定理

6.2.1 二次型中的线性变换

例 6.2 和例 6.3 中，变量之间的替换关系式都是线性的，因而称之为**线性变换**. 线性变换的一般形式是

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \mathrm{L} + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \mathrm{L} + c_{2n}y_n \\ \qquad \mathrm{L}\ \mathrm{L} \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \mathrm{L} + c_{nn}y_n \end{cases}. \tag{6.4}$$

这与 2.2 节中的线性方程组（2.2）十分相似，同样可用矩阵形式表示为

$$\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}, \tag{6.5}$$

其中

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \mathrm{M} \\ x_n \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \mathrm{M} \\ y_n \end{pmatrix}, \quad \boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \mathrm{L} & c_{1n} \\ c_{21} & c_{22} & \mathrm{L} & c_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ c_{n1} & c_{n2} & \mathrm{L} & c_{nn} \end{pmatrix}.$$

矩阵 $\boldsymbol{C}$ 叫做线性变换（6.4）的**变换矩阵**．选定变量 $\boldsymbol{y}$ 的一组值，可以按（6.4）式或（6.5）式确定变量 $\boldsymbol{x}$ 的一组取值．作为替换，总希望是“双向”的，即反过来选定 $\boldsymbol{x}$ 的一组值，要求能惟一确定 $\boldsymbol{y}$ 的一组取值．这个要求就是线性变换（6.4）可逆，也就是变换矩阵 $\boldsymbol{C}$ 可逆．在二次型（6.1）中作线性变换（6.4），通常都要求可逆．可逆变换的另一个好处是能保留二次型的一些固有特性而不至于发生退化，因此**可逆线性变换**又叫做**非退化线性变换**．

检验线性变换是否可逆并不难，只需检验矩阵的行列式 $|\boldsymbol{C}| \neq 0$ 即可（参看定理 2.2）．

例 6.4　将例 6.2 和例 6.3 中的线性变换写作矩阵形式（6.5）．

解　（1）例 6.2 中线性变换的矩阵形式显然是

$$\boldsymbol{y}=\begin{pmatrix}1&1&1\\0&1&2\\0&0&1\end{pmatrix}\boldsymbol{x}\ .$$

由于式中变换矩阵的行列式等于 1（三角行列式），所以矩阵可逆．用初等变换法（参看例 2.7）容易求出逆矩阵，从而得到形如（6.5）式的线性变换的矩阵形式为

$$\boldsymbol{x}=\begin{pmatrix}1&1&1\\0&1&2\\0&0&1\end{pmatrix}^{-1}\boldsymbol{y}=\begin{pmatrix}1&-1&1\\0&1&-2\\0&0&1\end{pmatrix}\boldsymbol{y}\ .$$

（2）例 6.3 中作了两次线性变换，它们分别是 $\boldsymbol{x}=\boldsymbol{C}_1\boldsymbol{y}$ 和 $\boldsymbol{z}=\boldsymbol{C}_2\boldsymbol{y}$，其中

$$\boldsymbol{C}_1=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix},\quad \boldsymbol{C}_2=\begin{pmatrix}1&0&-1\\0&1&-2\\0&0&1\end{pmatrix}.$$

不难发现，用配方法得到的线性变换都可逆，所以从后一式解出 $\boldsymbol{y}=\boldsymbol{C}_2^{-1}\boldsymbol{z}$，代入前一式即得 $\boldsymbol{x}=\boldsymbol{C}_1\boldsymbol{C}_2^{-1}\boldsymbol{z}$．求出逆矩阵，再做一次矩阵的乘法，便得

$$\boldsymbol{x}=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&-1\\0&1&-2\\0&0&1\end{pmatrix}^{-1}\boldsymbol{z}=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&2\\0&0&1\end{pmatrix}\boldsymbol{z}=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix}\boldsymbol{z}\ .$$

以上经过矩阵运算得到了所要的结果．本例的变换矩阵还可以用等式变形法得出．比如把例 6.2 中得到的变换式

$$\begin{cases}y_1=x_1+x_2+\ x_3\\ y_2=\quad\ x_2+2x_3\\ y_3=\qquad\ x_3\end{cases}$$

当作方程组来解，将 y_1,y_2,y_3 视为已知值，通过逐步代入，解出 $x_3=y_3$，$x_2=y_2-2(y_3)$，

$x_3 = y_1 - (y_2 - 2y_3) - (y_3) = y_1 - y_2 + y_3$．这些等式联立起来得到

$$\begin{cases} x_1 = y_1 - y_2 + y_3 \\ x_2 = \quad y_2 - 2y_3 \\ x_3 = \quad\quad y_3 \end{cases}.$$

由此能直接写出形如（6.5）式的矩阵形式．

等式变形法和矩阵运算法的实质是一样的．当变量很多时，用等式变形法会因大量的字母参加运算而显得不胜其烦，如果像本例的（2）那样需要经过多个层次的代入、变形，则求解过程更加复杂．

例 6.4 中特定的线性变换可以将指定的二次型化为标准形．为了寻找这种特定的线性变换，先来看看一般的线性变换（6.5）会产生怎样的矩阵结果．

在二次型 $f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 中，作线性变换 $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$，可得

$$f = (\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y}\ .$$

这说明经过线性变换，二次型变成了关于变量 $\boldsymbol{y}$ 的新的二次型．新的二次型的矩阵是 $\boldsymbol{B} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$．运用转置矩阵的性质（参看 2.2 节）不难发现，$\boldsymbol{B}$仍是对称矩阵．

如果变换后二次型的矩阵 $\boldsymbol{B}$ 是对角矩阵，即

$$\boldsymbol{B} = \begin{pmatrix} b_1 & 0 & \mathrm{L} & 0 \\ 0 & b_2 & \mathrm{L} & 0 \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ 0 & 0 & \mathrm{L} & b_n \end{pmatrix},$$

那么根据系数确定规则，已经得到了二次型的标准形．所以将二次型化为标准形，实质上是寻找合适的可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 成为对角矩阵．

6.2.2 化二次型为标准形的矩阵变换法

利用可逆矩阵 $\boldsymbol{C}$，将对称矩阵 $\boldsymbol{A}$ 变成对称矩阵 $\boldsymbol{B} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，也是一种变换．当然它不是变量间的变换，更不是线性变换，这种变换称为**合同变换**．合同变换是矩阵间的变换．

为了控制合同变换的方向，先考虑一种简单情形：设 $\boldsymbol{C}$ 是初等矩阵，则根据定理 2.4 可知，乘积 $\boldsymbol{A}\boldsymbol{C}$ 相当于对矩阵 $\boldsymbol{A}$ 施行一次初等列变换，再左乘 $\boldsymbol{C}^{\mathrm{T}}$ 即 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，相当于对矩阵 $\boldsymbol{A}\boldsymbol{C}$ 施行一次内容相同的初等行变换．可见，合同变换就是这种**配对式初等变换**．

矩阵的初等变换作为一项线性技术，具有方向明确、操作简便、有序渐进的特点，我们在前几章中已经看到了它的优越性，所以对矩阵 $\boldsymbol{A}$ 施行一系列的初等列变换，再对其施行内容相同的一系列初等行变换，可以很顺利地将矩阵 $\boldsymbol{A}$ 变成对角矩阵．这就是化二次型为标准形的**矩阵变换法**．

为了将一系列的初等变换记录下来，可以拼接一个单位矩阵充当“记录本”．由于在合

同变换 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 中，需要得到右侧的矩阵 $\boldsymbol{C}$，所以应该记录列变换，即在矩阵 $\boldsymbol{A}$ 的下方拼接一个单位矩阵 $\boldsymbol{E}$，对单位矩阵施行一系列的列变换后，可直接读出变换矩阵 $\boldsymbol{C}$（如果在 $\boldsymbol{A}$ 的右侧拼上一个单位矩阵，则记录下来的将是 $\boldsymbol{C}$ 的转置矩阵 $\boldsymbol{C}^{\mathrm{T}}$）.

例 6.5　设 $\boldsymbol{A}=\begin{pmatrix}1 & 3 & -2\\ 3 & 11 & -2\\ -2 & -2 & 9\end{pmatrix}$，求可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为对角矩阵.

解　对拼接后的 6×3 矩阵施行配对的初等变换．由于变换的目标是对角矩阵，所以主元（注意加上标记）应在矩阵块 $\boldsymbol{A}$ 的主对角线上选取：

$$\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{E}\end{pmatrix}=\begin{pmatrix}\langle 1\rangle & 3 & -2\\ 3 & 11 & -2\\ -2 & -2 & 9\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}\langle 1\rangle & 0 & 0\\ 3 & 2 & 4\\ -2 & 4 & 5\\ 1 & -3 & 2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 4\\ 0 & 4 & 5\\ 1 & -3 & 2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

以上两步变换中，第 1 步是将第 1 列乘 – 3 和 2 分别加到第 2、第 3 列上去，把主元所在行的其他元素变为 0．第 2 步作内容相同的行变换，即将第 1 行乘 – 3 和 2 分别加到第 2、第 3 行．接下去选择第 2 个主元，将第 2 列乘 – 2 加到第 3 列上去，同样把主元所在行的其他元素变为 0，再作相同的行变换，将第 2 行乘 – 2 加到第 3 行：

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \langle 2\rangle & 4\\ 0 & 4 & 5\\ 1 & -3 & 2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & \langle 2\rangle & 0\\ 0 & 4 & -3\\ 1 & -3 & 8\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -3\\ 1 & -3 & 8\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix}.$$

从最后矩阵的下半部分直接读出变换矩阵，从上半部分读出对角矩阵，即

$$\boldsymbol{C}=\begin{pmatrix}1 & -3 & 8\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix},\quad \boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -3\end{pmatrix}.$$

本例的计算表明，合同变换（配对式初等变换）虽然也是通过对矩阵施行初等变换来实现的，但它与前几章所用的初等变换的标准程序有所不同，主要有以下区别：

（1）主元必须在矩阵 $\boldsymbol{A}$ 所在子块的主对角线上选取；

（2）主元不必变为 1，只需将同行（列）的其他元素变为 0 即可；

（3）施行列变换后，必须配对地施行内容相同的行变换.

例 6.5 的初等变换是分 4 个步骤完成的．稍作观察不难发现，配对的初等变换之间相互没有干扰，可以一次性完成．事实上，用来配对的行变换不需要进行运算，只需将主元下方的元素改写为 0（对称位置上的元素值）即可．所以例 6.5 的 4 个步骤能够合并为 2 个步骤．

例 6.6 求一个非退化的线性变换，把二次型 $f=2x_1x_2+2x_1x_3-6x_2x_3$ 化为标准形．

解 此二次型的矩阵是 $\boldsymbol{A}=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix}$．

由于 $\boldsymbol{A}$ 的主对角线上没有非零元素，无法选出主元，所以先要进行变造．其方法是把具有非零元素的第 2 列加到第 1 列上去，再作配对变换，把第 2 行加到第 1 行：

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

接下去的配对变换和例 6.5 相仿．这里将配对的变换合并为一个步骤：

$$\begin{pmatrix}\langle 2\rangle&1&-2\\1&0&-3\\-2&-3&0\\1&0&0\\1&1&0\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&\langle -0.5\rangle&-2\\0&-2&-2\\1&-0.5&1\\1&0.5&1\\0&0&1\end{pmatrix}\to\begin{pmatrix}2&0&0\\0&-0.5&0\\0&0&6\\1&-0.5&3\\1&0.5&-1\\0&0&1\end{pmatrix}.$$

从最后一个矩阵读出 $\boldsymbol{C}=\begin{pmatrix}1&-0.5&3\\1&0.5&-1\\0&0&1\end{pmatrix}$，作初等变换 $\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&-0.5&3\\1&0.5&-1\\0&0&1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}$，可将原二次型化为标准形 $f=2z_1^2-0.5z_2^2+6z_3^2$．

本例的二次型与例 6.3 相同，但得到的标准形却不同，这是因为所用的方法不同．配方法往往不能直接得到变换矩阵 $\boldsymbol{C}$，需要附加矩阵的求逆运算和乘积运算．而矩阵变换法略去了变量字母，便于集中力量关注二次型的系数运算，最后能直接读出变换矩阵 $\boldsymbol{C}$．

6.2.3 二次型的惯性定理

例 6.3 和例 6.6 得到了同一个二次型的不同的标准形，不过也有共同之处：两个平方和

的项数一样，都是 3 项；两个平方和中，正项数目和负项数目相同，都分别是 2 项和 1 项．这是不是二次型的固有特性呢？

如前所述，合同变换 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 相当于对矩阵 $\boldsymbol{A}$ 施行一系列的初等列变换和一系列配对的初等行变换．因为初等变换不改变矩阵的秩（参看定理 3.1），所以标准形中的对角矩阵 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 满足 $R(\boldsymbol{B})=R(\boldsymbol{A})$．这说明矩阵 $\boldsymbol{B}$ 的对角线上非零元素个数等于原始矩阵 $\boldsymbol{A}$ 的秩，即标准形中的平方和项数就是 $R(\boldsymbol{A})$．无论得到什么样的标准形，这一点是不变的．

现在考虑正负平方项的数目．将例 6.6 的结果记为

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{z}\ ,\quad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=a_1z_1^2-a_2z_2^2+a_3z_3^2\ ,$$

式中 a_1,a_2,a_3 均为正数．假定另有一个线性变换把二次型化为 1 项正平方加上 2 项负平方，即

$$\boldsymbol{x}=\boldsymbol{B}\boldsymbol{y}\ ,\quad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=b_1y_1^2-b_2y_2^2-b_3y_3^2\ ,$$

其中 b_1,b_2,b_3 都是正数．将两个变换合成，得到

$$\boldsymbol{y}=\boldsymbol{B}^{-1}\boldsymbol{x}=\boldsymbol{B}^{-1}(\boldsymbol{C}\boldsymbol{z})=(\boldsymbol{B}^{-1}\boldsymbol{C})\boldsymbol{z}\ ,$$

此式代表两组变量 $\boldsymbol{y}$ 和 $\boldsymbol{z}$ 之间的线性变换式（共有 3 个等式），取其中之一

$$y_1=c_1z_1+c_2z_2+c_3z_3\ .$$

如果仅考虑实系数的线性变换，那么 c_1,c_2,c_3 作为矩阵 $\boldsymbol{B}^{-1}\boldsymbol{C}$ 的第一行元素，都是实数．令 $z_2=0$，$y_1=0$，代入上式得 $c_1z_1+c_3z_3=0$．这里变量多（2 个），方程少（1 个），所以可从中解出一组非零的实数解（z_1,z_3 取不全为零的实数）．在这样的取值之下，会得到相互矛盾的结果：一方面二次型等于

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=a_1z_1^2-0+a_3z_3^2>0\ ;$$

另一方面，二次型又必定等于

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=0-b_2y_2^2-b_3y_3^2\le 0\ .$$

这个矛盾说明，二次型的不同的标准形中，不可能有数目不等的正项和负项．以上讨论可归结为下面的定理．

定理 6.1　设有实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，对称矩阵 $\boldsymbol{A}$ 的秩为 r，有两个实的可逆变换 $\boldsymbol{x}=\boldsymbol{B}\boldsymbol{y}$ 和 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{z}$，分别使二次型化为

$$f=b_1y_1^2+b_2y_2^2+\mathrm{L}\ +b_ry_r^2\quad \text{和}\quad f=c_1z_1^2+c_2z_2^2+\mathrm{L}\ +c_rz_r^2\ ,$$

则 $b_1,b_2,\mathrm{L}\ ,b_r$ 中正（负）数的个数与 $c_1,c_2,\mathrm{L}\ ,c_r$ 中正（负）数的个数相等．

这个定理称为二次型的**惯性定理**．标准形中正项数目称为二次型的**正惯性指数**，负项数目称为**负惯性指数**，矩阵 $\boldsymbol{A}$ 的秩也叫做**二次型的秩**．

6.3 二次型的正交变换与有定性

6.3.1 用正交变换化二次型为标准形

在二次型中作线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$，相当于把 $\boldsymbol{x}$ 坐标系中的几何图形变换为 $\boldsymbol{y}$ 坐标系中的图形．一般的非退化线性变换虽然不改变图形的拓扑性质，但可能会改变向量的度量性质．比如由某个三元二次型构成的方程 $f(x_1,x_2,x_3)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=1$ 在 x_1,x_2,x_3 坐标系中表示椭球面，经可逆线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$，方程变为 $\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y}=1$，在 y_1,y_2,y_3 坐标系中仍表示椭球面，但椭球面的形状可能因为各个方向上拉伸、压缩的程度不同而发生改变．

为了保持度量性质不变，需要作正交变换，即在线性变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$ 中，矩阵 $\boldsymbol{Q}$ 是正交矩阵．比如将平面图形转动 θ 角的旋转变换

$$\begin{cases} x_1=y_1\cos\theta-y_2\sin\theta \\ x_2=y_1\sin\theta+y_2\cos\theta \end{cases},$$

就是一个正交变换．

用正交变换将二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 化为标准形，相当于求正交矩阵 $\boldsymbol{Q}$，使 $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}$ 成为对角矩阵．求正交矩阵的方法在 5.3 节中实际上已经作过讨论，这里将有关步骤归纳如下：

（1）解特征方程 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$，求出 n 个特征值 $\lambda_1,\lambda_2,\mathrm{L},\lambda_n$；

（2）通过特征矩阵对应的方程组 $(\lambda_i\boldsymbol{E}-\boldsymbol{A})\boldsymbol{\alpha}=\boldsymbol{0}$（$i=1,2,\mathrm{L},n$），求出 n 个基础特征向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\mathrm{L},\boldsymbol{\alpha}_n$；

（3）对应于相同特征值的基础特征向量需要进行正交化，正交化后的特征向量记为 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\mathrm{L},\boldsymbol{\beta}_n$；

（4）将所有的基础特征向量单位化后拼成正交矩阵 $\boldsymbol{Q}$，即

$$\boldsymbol{Q}=\left(\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|},\ \frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},\ \mathrm{L},\ \frac{\boldsymbol{\alpha}_n}{|\boldsymbol{\alpha}_n|}\right);$$

（5）用正交变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$ 把二次型化为标准形，其平方项的前置系数，恰好为 n 个特征值，即 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q})y=\lambda_1y_1^2+\lambda_2y_2^2+\mathrm{L}+\lambda_ny_n^2$．

例 6.7 用正交变换将下列二次型化为标准形：

$$f=2x_1^2+5x_2^2+5x_3^2+4x_1x_2-4x_1x_3-8x_2x_3.$$

解 二次型的矩阵是 $\boldsymbol{A}=\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$．在例 5.7 中，已求得该矩阵的特征值是 1，1，10，对应的特征向量是

$$\boldsymbol{\alpha}_1=(-2,\ 1,\ 0)^{\mathrm{T}},\ \boldsymbol{\alpha}_2=(2,\ 0,\ 1)^{\mathrm{T}},\ \boldsymbol{\alpha}_3=(1,\ 2,\ -2)^{\mathrm{T}};$$

将前两个特征向量正交化，得到

$$\boldsymbol{\beta}_1=(-2,\ 1,\ 0)^{\mathrm{T}},\quad \boldsymbol{\beta}_2=\frac{1}{5}(2,\ 4,\ 5)^{\mathrm{T}};$$

接下来将特征向量 $\boldsymbol{\beta}_1$，$\boldsymbol{\beta}_2$，$\boldsymbol{\alpha}_3$ 单位化后拼成矩阵

$$\boldsymbol{Q}=\left(\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|},\ \frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},\ \frac{\boldsymbol{\alpha}_3}{|\boldsymbol{\alpha}_3|}\right)=\frac{1}{3\sqrt{5}}\begin{pmatrix}-6 & 2 & \sqrt{5}\\ 3 & 4 & 2\sqrt{5}\\ 0 & 5 & -2\sqrt{5}\end{pmatrix};$$

作正交变换 $(x_1,\ x_2,\ x_3)^{\mathrm{T}}=\boldsymbol{Q}(y_1,\ y_2,\ y_3)^{\mathrm{T}}$，可将二次型化为 $f=y_1^2+y_2^2+10y_3^2$．

正交变换的优点多（既保持拓扑性质，又保持度量性质），当然所花的代价也大（求适当的正交矩阵步骤多、运算量大）．

6.3.2　二次型的有定性

二次型的**有定性**是指二次型作为二次函数，其函数值能够保持固定的符号．有定性分为正定、负定、半正定、半负定 4 种情形．

设有 n 元实二次型

$$f(x_1,x_2,\mathrm{L}\ ,x_n)=f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\ .$$

如果对任何非零向量 $\boldsymbol{x}\neq 0$，都有 $f(\boldsymbol{x})>0$（$f(\boldsymbol{x})<0$），则称 f 为**正定（负定）二次型**，并称实对称矩阵 $\boldsymbol{A}$ 为**正定（负定）矩阵**；如果对任何向量 $\boldsymbol{x}$，都有 $f(\boldsymbol{x})\geq 0$（$f(\boldsymbol{x})\leq 0$），但不能排除 $\boldsymbol{x}\neq 0$ 时会出现 $f(\boldsymbol{x})=0$，则称 f 为**半正定（半负定）二次型**，并称实对称矩阵 $\boldsymbol{A}$ 为**半正定（半负定）矩阵**．除此之外，如果二次型既可能取正值，又可能取负值，则该二次型就称为**不定的**．

判别二次型有定性的第 1 个方法是利用惯性定理 6.1．二次型的标准形是平方和，如果这些平方和的系数没有负数，则保证了二次型总是取非负值．如果进一步，平方项的系数全部为正，则保证了二次型总是取正值．由此可见，实二次型为正定（负定）的充分必要条件是它的正（负）惯性指数等于变量个数 n；实二次型为半正定（半负定）的充分必要条件是它的负（正）惯性指数等于零，而且正（负）惯性指数小于变量个数 n．如本章例 6.2 给出的二次型是半正定的，例 6.7 给出的二次型是正定的，例 6.3 和例 6.5 给出的二次型都是不定的．

根据惯性指数判别二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的有定性，需要先将二次型化为标准形．如果仅仅判别有定性，则不必记录变换矩阵 $\boldsymbol{C}$，所以在运用矩阵变换法时，无须拼接单位矩阵，直接对矩阵 $\boldsymbol{A}$ 施行配对初等变换即可．

判别二次型有定性的第 2 个方法是根据特征值的符号．因为二次型通过正交变换化成的标准形是以对称矩阵的特征值为系数的平方和，所以实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定（负定）

的充分必要条件是矩阵 $\boldsymbol{A}$ 的所有 n 个特征值都是正数（负数）；实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定（半负定）的充分必要条件是矩阵 $\boldsymbol{A}$ 的所有 n 个特征值中有数 0，但没有负数（正数）.

利用特征值判别二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的有定性，不需要将二次型化为标准形，只需从特征方程 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 中解出 n 个特征值即可．但 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$ 是关于 λ 的 n 次方程，其求解的难度往往超过将二次型化为标准形．

作为有定性的特例，判断二次型的正定性还有第 3 个方法，即考察 n 个行列式的值．设有对称矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1n} \\ a_{21} & a_{22} & \mathrm{L} & a_{2n} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{n1} & a_{n2} & \mathrm{L} & a_{nn} \end{pmatrix}.$$

处于矩阵 $\boldsymbol{A}$ 的左上角的 k 阶子块（$k\leq n$）

$$\boldsymbol{A}_k=\begin{pmatrix} a_{11} & a_{12} & \mathrm{L} & a_{1k} \\ a_{21} & a_{22} & \mathrm{L} & a_{2k} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ a_{k1} & a_{k2} & \mathrm{L} & a_{kk} \end{pmatrix}$$

称为矩阵 $\boldsymbol{A}$ 的**主子矩阵**．$\boldsymbol{A}_k$ 的行列式 $|\boldsymbol{A}_k|$（$k=1,2,\mathrm{L}\ ,n$）称为矩阵 $\boldsymbol{A}$ 的**顺序主子式**.

在二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 中，若令后 $n-k$ 个变量等于 0，则 f 成了只含有前 k 个变量的二次型 $f=\boldsymbol{x}_{(k)}^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{x}_{(k)}$，这里 $\boldsymbol{x}_{(k)}$ 表示具有 k 个分量的变量列．于是矩阵 $\boldsymbol{A}$ 的正定性与主子矩阵 $\boldsymbol{A}_k$（$k=1,2,\mathrm{L}\ ,n$）的正定性联系起来了，而且呈现出递进式（k 值逐步增大）的联系．另一方面，正定矩阵的行列式必须为正．由此得到如下定理.

定理 6.2 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定的充分必要条件是矩阵 $\boldsymbol{A}$ 的所有顺序主子式全都大于 0，即 $|\boldsymbol{A}_k|>0$（$k=1,2,\mathrm{L}\ ,n$）.

根据顺序主子式的符号判断二次型的正定性，不必化标准形，也不必解特征方程，只需求出 n 个行列式的值即可．当然计算行列式的值有一定的运算量.

判断二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的负定性，相当于判断 $-f=\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{A})\boldsymbol{x}$ 的正定性.

例 6.8 设 $\boldsymbol{A}=\begin{pmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{pmatrix}$，用 3 种方法判断二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正定性.

解 （1）**惯性指数法**．判断正定性，无须记录变换矩阵，故不拼接单位矩阵，直接对矩阵 $\boldsymbol{A}$ 施行配对式初等变换：

$$\boldsymbol{A}=\begin{pmatrix}5&2&-4\\2&\langle 1\rangle&-2\\-4&-2&5\end{pmatrix}\xrightarrow{\text{列变换}}\begin{pmatrix}1&2&0\\0&\langle 1\rangle&0\\0&-2&1\end{pmatrix}\xrightarrow{\text{行变换}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

矩阵 $\boldsymbol{A}$ 已化为对角矩阵，正惯性指数等于 3，所以 f 是正定二次型．

（2）特征值法．先化简特征多项式

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-5&\langle -2\rangle&4\\-2&\lambda-1&2\\4&2&\lambda-5\end{vmatrix}\overset{\text{行变换}}{=}\begin{vmatrix}\lambda-5&-2&4\\-2&\lambda-1&2\\\lambda-1&0&\lambda-1\end{vmatrix}$$

$$=(\lambda-1)\begin{vmatrix}\lambda-5&-2&4\\-2&\lambda-1&2\\1&0&\langle 1\rangle\end{vmatrix}\overset{\text{列变换}}{=}(\lambda-1)\begin{vmatrix}\lambda-9&-2&4\\-4&\lambda-1&2\\0&0&1\end{vmatrix}$$

$$=(\lambda-1)[(\lambda-9)(\lambda-1)-8]=(\lambda-1)(\lambda^2-10\lambda+1)\text{．}$$

解特征方程 $(\lambda-1)(\lambda^2-10\lambda+1)=0$，得到 3 个特征值 $\lambda_1=5+2\sqrt{6}$，$\lambda_2=5-2\sqrt{6}$ 和 $\lambda_3=1$，它们都是正数，所以 f 是正定二次型．

（3）顺序主子式法．计算行列式：

$$|\boldsymbol{A}_1|=|5|=5>0\text{，}\quad|\boldsymbol{A}_2|=\begin{vmatrix}5&2\\2&1\end{vmatrix}=1>0\text{，}$$

$$|\boldsymbol{A}_3|=\begin{vmatrix}5&2&-4\\2&\langle 1\rangle&-2\\-4&-2&5\end{vmatrix}\overset{\text{列变换}}{=}\begin{vmatrix}1&2&0\\0&1&0\\0&-2&1\end{vmatrix}=\begin{vmatrix}1&0\\0&1\end{vmatrix}=1>0\text{．}$$

3 个顺序主子式都大于 0，所以 f 是正定二次型．

从本例看出，顺序主子式法仍未脱离初等变换．对于本例来说，惯性指数法最简便，特征值法最麻烦．

习 题 六

1．写出下列二次型的矩阵．

（1） $f=x_1^2-2x_1x_2+3x_1x_3-2x_2^2+8x_2x_3+3x_3^2$ **；**

（2） $f=x_1^2+2x_1x_2-x_1x_3+2x_3^2$ **；**

（3） $f=x_1x_2-x_3x_4$ **．**

2．写出下列对称矩阵所对应的二次型（写成各项之和的形式）．

(1) $\boldsymbol{A}=\begin{pmatrix}1&-1&-3&1\\-1&0&-2&4\\-3&-2&3&-5\\1&4&-5&0\end{pmatrix}$；　(2) $\boldsymbol{A}=\begin{pmatrix}0&1&2&-4\\1&0&-1&-1\\2&-1&0&3\\-4&-1&3&0\end{pmatrix}$.

3．用配方法将下列二次型化为标准形，并求变换矩阵．

(1) $f=x_1^2+4x_2^2+2x_3^2+2x_1x_3$；　(2) $f=x_1^2-3x_2^2-2x_1x_2+2x_1x_3-6x_2x_3$；

(3) $f=-4x_1x_2+2x_1x_3+2x_2x_3$．

4．用矩阵变换法将下列二次型化为标准形，并求变换矩阵．

(1) $f=x_1^2+2x_2^2+7x_3^2-2x_1x_2+4x_1x_3+4x_2x_3$；　(2) $f=2x_1x_2-4x_1x_3+10x_2x_3$；

(3) $f=x_1^2-7x_3^2-2x_1x_2-6x_1x_3+2x_1x_4-4x_2x_3+8x_2x_4-10x_3x_4$．

5．对于下列矩阵 $\boldsymbol{A}$，求可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$ 为对角矩阵，并写出这个对角矩阵．

(1) $\boldsymbol{A}=\begin{pmatrix}1&2&0\\2&0&1\\0&1&3\end{pmatrix}$；　(2) $\boldsymbol{A}=\begin{pmatrix}0&1&-2\\1&0&-1\\-2&-1&0\end{pmatrix}$.

6．用正交变换法将下列二次型化为标准形，并求所用的正交矩阵．

(1) $f=2x_1^2+x_2^2+3x_3^2-4x_1x_2-4x_1x_3$；　(2) $f=4x_1^2+x_2^2+x_3^2-4x_1x_2-4x_1x_3-8x_2x_3$．

7．求下列二次型的惯性指数，并说明它的有定性．

(1) $f=x_1^2+2x_1x_2+2x_2^2+4x_2x_3+4x_3^2$；　(2) $f=x_1x_2-x_1x_3+2x_2x_3+x_4^2$．

8．用 3 种方法判断二次型 $f=2x_1^2+6x_2^2+4x_3^2-2x_1x_3-2x_2x_3$ 的正定性．

9．求 a 的值，使二次型 $f=x_1^2+x_2^2+5x_3^2+2ax_1x_2-2x_1x_3+4x_2x_3$ 为正定二次型．

10．单项选择题

(1) 下列各式中，哪一个不等于 $x_1^2+6x_1x_2+3x_2^2$．　(　　)

A．$\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}1&2\\4&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$；　B．$\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}1&3\\3&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$；

C．$\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}1&-1\\-5&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$；　D．$\begin{pmatrix}x_1&x_2\end{pmatrix}\begin{pmatrix}1&-1\\7&3\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$.

(2) 矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-1&0\\-1&3&0\\0&0&0\end{pmatrix}$ 对应的二次型是　(　　)

A．$x_1^2-2x_1x_2+3x_2^2$；　B．$x_1^2-x_1x_2+3x_2^2$；

C．$x_1^2-x_1x_2-x_1x_3+3x_2^2$；　D．$x_1^2-x_1x_2-x_2x_3+3x_2^2$．

(3) 用配方法把二次型 $f(x_1,x_2,\cdots,x_n)$ 化为平方和　(　　)

A．要求 f 中有平方项；　B．能直接得到线性变换矩阵；

C．最多需要经过 n 次配方；　D．一定能够把二次型变为标准形．

(4) 用矩阵变换法化二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$ 为标准形，在矩阵 $\boldsymbol{A}$ 的下方拼接单位矩阵，是为了　(　　)

A．得到线性变换的矩阵；　B．核对惯性指数的取值；

C．简化初等变换的运算；　　D．控制合同变换的方向．

(5) 用正交变换化二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为标准形，需要实施哪个过程．　（　）

A．对矩阵 $\boldsymbol{A}$ 施行配对的初等变换；　　B．对二次型进行一系列的配方；

C．求出矩阵 $\boldsymbol{A}$ 的特征值和特征向量；　　D．将以上 3 种过程相结合．

(6) 设 $\boldsymbol{A}=\begin{pmatrix}-2&0&0\\0&2&0\\0&0&5\end{pmatrix}$，则存在可逆矩阵 $\boldsymbol{C}$，使得 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}=$　（　）

A．$\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}$；　B．$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$；　C．$\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&3\end{pmatrix}$；　D．$\begin{pmatrix}-3&0&0\\0&-2&0\\0&0&-5\end{pmatrix}$．

(7) 如果将任意的 $x_1\neq 0, x_2\neq 0,\mathrm{L}\ , x_n\neq 0$ 代入实二次型 $f(x_1,x_2,\mathrm{L}\ ,x_n)$ 中，都有 $f>0$，则该二次型　（　）

A．正定；　　B．半正定；　　C．不是正定；　　D．未必正定．

(8) 设 $a=x_1+x_2+\mathrm{L}\ +x_n$，则 $f(x_1,x_2,\mathrm{L}\ ,x_n)=(nx_1-a)^2+(nx_2-a)^2+\mathrm{L}\ +(nx_n-a)^2$　（　）

A．是正定二次型；　B．是半正定二次型；　C．是不定的二次型；　D．不是二次型．

(9) 实二次型 $f(x_1,x_2,\mathrm{L}\ ,x_n)$ 为正定的充分必要条件是　（　）

A．f 的秩为 n；　　B．f 的正惯性指数等于 f 的秩；

C．f 的负惯性指数为零；　　D．f 的正惯性指数为 n．

(10) 实二次型 $\begin{pmatrix}x&y&z\end{pmatrix}\begin{pmatrix}2&k&0\\k&2&0\\0&0&k^2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$ 为正定的充分必要条件是　（　）

A．$k\neq 0$；　　B．$k<2$ 或 $k>-2$；

C．$0<k<2$ 或 $-2<k<0$；　　D．$-2<k<2$．

第二篇　概 率 统 计

第 7 章　随机事件及其概率

7.1　随 机 事 件

7.1.1　随机试验与样本空间

在现实生活中，我们常常会遇到许多预先不能确定结果的现象，比如投掷一枚硬币，事先不能断定正面朝上还是反面朝上．又比如在既有红球又有白球的袋中任取 5 个球，若球的大小、质感均相同，则无法断定会出现几个白球．尽管我们能罗列出所有可能发生的结果，但不能事先断定将发生哪一个特定的结果，这种现象称为**随机现象**．现实世界中随机现象无处不在，除了上面提及的投币与摸球，再如射击目标、测量潮位、交通事故、股票行情等．概率论和数理统计就是在研究随机现象的过程中发展起来的．

我们把实验、观察、研究某一随机现象叫做**随机试验**，简称**试验**．随机试验中每一个特定的结果都称为**随机事件**，简称**事件**．

例 7.1　教师上第一堂课，还不认识学生，随机地（无任何倾向性、随意地）叫一个学号，请对应的学生回答问题．站起来的可能“是女生”，可能“是戴眼镜的学生”或者“是穿红衣服的学生”，可能“是高个子学生”，也可能“是名叫张华的学生”……，这些都是随机事件．

随机事件是随机试验的某种结果，通常表现为一个明确的判断句．随机事件在试验之前无法断然确定，不过事后可以观察到其是否发生．为了便于研究，随机事件用大写字母 A，B，C……表示．

在众多的随机事件中，有的可以分解成更简单的事件，比如例 7.1 中，“站起来的是女生”可以分解成“是女生甲”、“是女生乙”、“是女生丙”，……，而有的事件不能再分解，如“站起来的是张华”．我们把可以分解的事件称为**复合事件**，而把不能再分解的“最简单”的事件称为**基本事件**，有时还把基本事件几何化，叫做“**样本点**”．显然，基本事件是随机试验中各种最基本的可能结果，例 7.1 中的基本事件是教室里的每一个学生．

全体基本事件构成的集合称为该随机试验的**样本空间**，记为 Ω．例如投币试验中以 ω_0 表示“反面朝上”，以 ω_1 表示“正面朝上”，则样本空间 $\Omega=\{\omega_0,\omega_1\}$；在例 7.1 中若将教室里的 n 个学生用 ω_i（$i=1,2,\mathrm{L},n$）表示，则样本空间 $\Omega=\{\omega_1,\omega_2,\mathrm{L},\omega_n\}$．

样本空间可以是无限集，例如观察一小时内到达某交通路口的车辆数，其样本空间是 $\Omega=\{0,1,2,\mathrm{L}\}$．样本空间还可以是连成一片的区间，例如测量长江某处的潮位，其样本空间是 $\Omega=\{x\mid x\geq 0\}$．

例 7.2　在一批灯泡中，任取一只测试它的寿命（1000～3000h）：

（1）试说出 3 个不同的事件，并估计它们发生的可能性；

（2）指出 2 个不同的样本点；

（3）表述试验的样本空间．

本例是讨论性题目，没有标准答案，所以这里不作解答．

7.1.2　随机事件与集合

从例 7.1 看出，随机事件都是基本事件的某种组合，因此随机事件都是样本空间 Ω 的子集，从而可以用集合论的观点来研究随机事件．集合之间有相等、包含、互补等关系，还有并、交、差等运算，事件之间自然也有这些关系与运算，而且运算方法与集合论完全相同．

例 7.3　一副扑克牌有 4 种花色 52 张牌，从中任取一张观察其结果．这个试验的样本空间由 52 张牌构成，即 $\Omega=\{\omega_1,\omega_2,\mathrm{L},\omega_{52}\}$．可以约定一个排序方式，比如 ω_1～ω_{13} 表示黑桃 A、黑桃 2 直至黑桃 K 等，设事件 $A=$“抽到红心”，$B=$“抽到红牌”，$C=$“抽到方块”，A 由 13 个样本点即 13 张红心牌构成，比如编号为 ω_{14}～ω_{26}，B 由 26 个样本点即 13 张红心牌与 13 张方块牌构成，比如编号为 ω_{14}～ω_{39}．

如果实际抽到了红心 2，则认为事件 A 和 B 都已发生，而事件 C 没有发生．由此可见，我们说一个事件 A 发生，指的是事件 A 中的某一个基本事件发生，即

$$\text{事件 } A \text{ 发生} \Leftrightarrow \text{基本事件}\{\omega\}\text{ 发生且 } \omega\in A\text{．}$$

特别地，样本空间 Ω 是其本身的子集（平凡子集），也是一个事件，而且因为试验结果总是 Ω 中的一个基本事件发生，所以 Ω 是必然发生的，称 Ω 为**必然事件**．必然事件的语言表述有许多形式，比如例 7.3 中

$$\Omega=\{\omega_1,\omega_2,\mathrm{L},\omega_{52}\}=\text{“抽到一张牌”}$$
$$=\text{“抽到红牌或黑牌”}=\text{“抽到 }\omega_1\text{～}\omega_{52}\text{ 中的一张”．}$$

另一个特殊事件是空集 $\varnothing$．空集 $\varnothing$ 是 Ω 的子集，当然是事件，但是试验结果出现的任何基本事件都不属于 $\varnothing$，所以 $\varnothing$ 是一定不会发生的，称 $\varnothing$ 为**不可能事件**．不可能事件的语言表述也可以有许多形式，比如在投币试验中

$$\varnothing=\text{“反面朝上而且正面朝上”}=\text{“硬币变成一只鸟飞走了”．}$$

注意这里指的不是变魔术，魔术师所做的是“让人们以为硬币变成了鸟”，这是完全可能发生的，但硬币真的变成鸟就不可能了.

7.1.3 事件的关系与运算

因为事件都是样本点构成的集合，所以事件之间的关系和运算可以用集合的语言来描述，但从上述对必然事件和不可能事件的描述可知，在概率论中，更注重对事件的事件性语言描述.

1．相等关系

两个事件 A 和 B **相等**，即 $A=B$，是指 A，B 所包含的样本点完全相同．用事件性语言表述，是指 A 和 B 同时发生且同时不发生，尽管它们可以有不同的表述形式，但实质上是两个相同的事件．如投币试验中，$A=$“正面朝上”，$B=$“反面朝下”是相同的事件，因此 $A=B$.

2．包含关系

事件 A **包含于**事件 B，即 $A\subset B$，指的是 A 所包含的样本点全部在 B 中．用事件性语言表述，是指事件 A 发生必然导致事件 B 发生．在例 7.3 中，“抽到红心”导致已经“抽到了红牌”，所以 $A\subset B$．需要说明的是，事件的内涵越多，其外延就越小，因此内涵多的事件必然被内涵少的事件所包含，如例 7.3 中，“抽到红色大牌”$\subset$“抽到红牌”.

3．事件的和（并）

作为集合，事件 A 和 B 的并记为 $A\cup B$，指的是将 A，B 所包含的样本点合并，形成一个新的事件．用事件性语言表述，$A\cup B=$“A 发生或者 B 发生”，更常用的表述是“A 和 B 至少有一个发生”．如在例 7.3 中，$B=A\cup C$.

在概率论中，习惯于将“事件的并”称为“**事件的和**”，记为

$$A\cup B=A+B.$$

这里的‘＋’号表示集合的并，不要与数量的加法相混淆，如 $A+A=A$ 而决非 $A+A=2A$．多个事件的并称为多个事件的和，记为

$$\bigcup_{i=1}^{n}A_i=\sum_{i=1}^{n}A_i=A_1+A_2+\cdots+A_n,$$

它表示 n 个事件 $A_1,A_2,\cdots,A_n$ 至少有一个发生.

4．事件的乘积（交）

作为集合，事件 A 和 B 的交记为 $A\cap B$，指的是由 A，B 所包含的样本点的公共部分所构成的新的事件．用事件性语言表述，$A\cap B=$“A 和 B 同时发生”．如在例 7.3 中，由

于 $A \subset B$，所以 $A \cap B = A$，这和集合论中的结果完全一致．

在概率论中，习惯于将“事件的交”称为“**事件的乘积**”，记为

$$A \cap B = AB .$$

这里的乘积表示集合的交，不要与数量的乘法相混淆，如 $A \cdot A = A$ 而不再记作 $A \cdot A = A^2$，更不能由此得出 $A = 1$．多个事件的交称为多个事件的乘积，记为

$$\bigcap_{i=1}^{n} A_i = \prod_{i=1}^{n} A_i = A_1 A_2 \cdots A_n ,$$

它表示 n 个事件 $A_1, A_2, \cdots, A_n$ 同时发生．

5．事件的逆

由于随机事件都是样本空间 Ω 的子集，所以 Ω 相当于全集．集合（事件）A 的补集称为事件 A 的**逆事件**，记为 $\overline{A}$．用事件性语言表述，$\overline{A}$ =“A 不发生”．A 和 $\overline{A}$ 是**互逆事件**，又叫做**对立事件**．与集合论中的结果相同，$\overline{\overline{A}} = A$，并且满足 $A + \overline{A} = \Omega$，$A\overline{A} = \varnothing$．如例 7.3 中，$\overline{B}$ =“抽到黑牌”．

6．事件的差

两个**事件的差**记为 $A - B$，指的是在事件 A 中去掉属于事件 B 的样本点，形成一个新的事件，用事件性语言表述，$A - B$ =“A 发生但 B 不发生”．如在例 7.3 中，$B - A = C$．显然，差的运算可以转换为乘积与逆的运算，即

$$A - B = A\overline{B} ,\quad \Omega - A = \overline{A} .$$

7．互不相容事件

如果两个事件 A 和 B 不相交（没有公共样本点），即 $AB = \varnothing$，则称 A，B 为**互不相容事件**，简称**不相容**．用事件性语言表述，即 A，B 不可能同时发生．显然对立事件 A 和 $\overline{A}$ 是互不相容事件，但互不相容事件未必是对立事件，如例 7.3 中的事件 A 和事件 C 互不相容，但它们并非对立事件．互不相容是研究随机事件时常见的现象，在许多情况下需要将一个事件转化为互不相容的事件之和，所以要加强对互不相容概念的理解．

既然事件就是集合，事件的关系与运算等同于集合的关系与运算，因而事件运算的规则如交换律、结合律、分配律以及德·摩根律 $\overline{A + B} = \overline{A}\,\overline{B}$，$\overline{AB} = \overline{A} + \overline{B}$ 等都与集合的相应运算规则相同，事件的运算还可借助于文氏图来进行，此处不再赘述．

例 7.4　生产加工 3 个零件，A_i $(i = 1,2,3)$ 表示“第 i 个零件是正品”．

(1) 用事件性语言表述事件 $A_1A_2A_3$ 和 $\overline{A_1A_2 + A_2A_3 + A_1A_3}$；

(2) 将事件“至少有一个是次品”和“恰有一个是次品”符号化，即用事件 A_i $(i = 1,2,3)$ 及其运算来表示．

解　(1) $A_1A_2A_3$ =“3 个零件全是正品”=“3 个零件中没有 1 个是次品”；

$\overline{A_1A_2 + A_2A_3 + A_1A_3}$ = “至少有 2 个正品的情况不发生”

= “3 个零件中至多有 1 个正品”；

(2) “至少有 1 个是次品” = $\bar{A}_1 + \bar{A}_2 + \bar{A}_3 = \overline{A_1A_2A_3}$ ；

“恰有 1 个是次品” = $\bar{A}_1A_2A_3 + A_1\bar{A}_2A_3 + A_1A_2\bar{A}_3$

$= A_1A_2 + A_2A_3 + A_1A_3 - A_1A_2A_3$.

注意：逆事件的横线记号不能随意中断或相连，如 $\overline{A_1A_2A_3}$ 表示“并非 3 个事件同时发生”，即“3 个事件至少有 1 个不发生”，而 $\bar{A}_1\bar{A}_2\bar{A}_3$ 则表示“3 个事件全都不发生”. 横线记号的相连与中断，在不经意的书写中很容易混淆，然而它们的含义却有非常大的差异.

7.2 事件的概率

7.2.1 古典概率

概率就是事件发生的可能性大小，事件 A 的概率记为 $P(A)$. 现实生活中，人们常常会去估计事件的概率值，例如在投币试验中，出现正面和反面的可能性相同，各占 50%，故认为 P (**正**)=0.5， P (**反**)=0.5. 又如在例 7.3 中，52 张牌里有 13 张红心，由于每张牌被抽到的机会均等，因此很自然地认为 $P(A)$ =13/25=0.25. 这是一种简单而又直观地计算概率的方法，当然它的前提是有限性和等可能性.

如果一个随机试验满足以下两个条件：

(1) **有限性**——样本空间中的元素（样本点）有限，即

$$\Omega = \{\omega_1, \omega_2, \text{L}, \omega_n\},$$

(2) **等可能性**——基本事件发生的可能性相同，即

$$P(\omega_1) = P(\omega_2) = \text{L} = P(\omega_n) = \frac{1}{n},$$

那么称该随机试验属于**古典概型**. 对于古典概型，事件 A 的概率可按以下公式计算：

$$P(A) = \frac{m}{n} = \frac{A\text{所包含的样本点数}}{\text{样本点总数}}. \qquad (7.1)$$

用这个公式定义的概率称为**古典概率**.

7.2.2 概率的性质

(1) **有界性** 对于任意事件 A ，$0 \le P(A) \le 1$；

(2) **完备性** 对于必然事件 Ω 和不可能事件 $\varnothing$ ， $P(\Omega) = 1$ ， $P(\varnothing) = 0$ ；

（3）**单调性**　对于任意两个事件 A 和 B，若 $A \subset B$，则 $P(A) \leq P(B)$；

（4）**互补性**　对于任意事件 A，$P(\bar{A}) = 1 - P(A)$；

（5）**可加性**　对于任意两个事件 A 和 B，都有

$$P(A+B) = P(A) + P(B) - P(AB)\text{．} \tag{7.2}$$

若 A，B 互不相容，则

$$P(A+B) = P(A) + P(B)\text{．} \tag{7.3}$$

若 n 个事件 $A_1, A_2, \cdots, A_n$ 两两互不相容，则

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)\text{．} \tag{7.4}$$

这 3 个公式统称为概率的**加法公式**．公式（7.2）表明，事件之和的概率应在概率之和的基础上予以扣减．公式（7.3）和（7.4）表明，互不相容事件之和的概率等于它们的概率之和．

以上性质（1）～（4）能直接从公式（7.1）推得．公式（7.2）可以利用文氏图说明：如图 7-1 所示，各集合包含的样本点数显然满足平衡式

$$m_{A+B} = m_A + m_B - m_{AB}\text{，}$$

上式两端除以样本点总数 n，便得公式（7.2）．如果 $AB = \varnothing$，即 A，B 不相容，则从 $P(AB) = 0$ 立得公式（7.3）．公式（7.4）可从公式（7.3）递推得到．

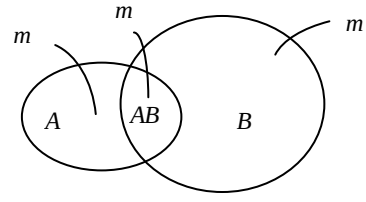

图 7-1　两个事件的样本点

概率的上述 5 条性质不仅适用于古典概型，而且适用于其他概率模型，以后运用这些性质时，不再加以说明．

例 7.5　学校开设数学类和外语类选修课．调查结果表明，有 53%的学生选修数学课，有 45%的学生选修外语课，两类课都选的学生有 20%．问参加选修课学习的学生占多少百分比？

解　设 A =“选修数学课”，B =“选修外语课”，已知 $P(A) = 0.53$，$P(B) = 0.45$，$P(AB) = 0.2$．“参加选修课学习”指的是选修数学课或者选修外语课，即求 $P(A+B)$．根据加法公式（7.2），可算得 $P(A+B) = P(A) + P(B) - P(AB)$ ＝0.53＋0.45－0.2=0.78，所以参加选修课学习的学生占 78%．

7.2.3　古典概率的计算

公式（7.1）给出了计算古典概率的非常简便有效的方法，但是实际应用起来往往有较大的难度，难度来自于对等可能样本空间的描述．

例 7.6 同时抛掷两枚硬币，求落下后出现“一正一反（朝上）”的概率.

解 1 样本空间是 $\Omega=\{$ 两正，两反，一正一反 $\}$，即 $n=3$，故 P（一正一反）$=1/3$.

解 2 将硬币编号，根据出现正反的排列不同，写出样本空间是 $\Omega=\{$〈正，正〉,〈正，反〉,〈反，正〉,〈反，反〉$\}$，即 $n=4$，故 P（一正一反）$=2/4=0.5$.

两种解法得到不同的结果是由于对样本空间理解的不一致. 事实上，每枚硬币出现正、反的机会均等，两枚硬币不同搭配的机会也均等，所以解 2 是正确的，解 1 因忽略了等可能性而导致错误.

实际上，公式（7.1）只要求给出样本空间的容量，即试验的基本结果的总可能数，并不需要详细描述样本空间的结构. 为了使公式便于操作，我们将公式（7.1）改写为

$$P(A)=\frac{m}{n}=\frac{\text{对}A\text{有利的可能数}}{\text{试验的总可能数}}. \tag{7.5}$$

运用公式（7.5）计算古典概率，可兼顾简便性与正确性.

例 7.7 12 件产品中有 9 件正品，3 件次品，从中任意抽取 5 件，设 $A=$“恰有两件次品”，$B=$“至少有一件次品”，$C=$“至少有两件次品”.

（1）5 件产品是同时取出的，求 $P(A)$，$P(B)$，$P(C)$；

（2）5 件产品是一件一件依次取出的，取出后不再放回（称为**无放回抽样**），求 $P(A)$，$P(B)$，$P(C)$.

解 （1）12 件产品中任取 5 件，总可能数是组合数 C_{12}^5；事件 A 是指 5 件产品中有 2 件次品和 3 件正品，2 件次品从 3 件次品中取出，其可能数是 C_3^2；3 件正品从 9 件正品中取出，其可能数是 C_9^3，根据乘法原理，对 A 有利的可能数是 $C_3^2C_9^3$，于是

$$P(A)=\frac{C_3^2C_9^3}{C_{12}^5}=\frac{7}{22}=0.318.$$

对 B 有利的可能数是在总可能数中扣除“5 件都是正品”的可能数 C_9^5，于是

$$P(B)=\frac{C_{12}^5-C_9^5}{C_{12}^5}=1-\frac{7}{44}=0.841.$$

求 $P(B)$ 的算式中，实际上用到了概率的互补性质. 由此可见，当可能数的构成较复杂时，逆向思维往往事半功倍. 对 C 有利的可能数应该分为“恰有 2 件次品”和“恰有 3 件次品”计算，然后相加，于是

$$P(C)=\frac{C_3^2C_9^3+C_3^3C_9^2}{C_{12}^5}=\frac{4}{11}=0.364.$$

（2）由于事件 A，B，C 都是不计次序的结果，所以同时取出 5 件和依次取出 5 件的效果是一样的，因而仍有 $P(A)=0.318$，$P(B)=0.841$，$P(C)=0.364$.

对本例作更加深入的考虑可以发现，所谓的“同时”取出，实际上总有先后次序，只

是因为时间间隔较短或者认为无须分辨先后而未加记录，才笼统地说成“同时”. 即使产品体积很小可以一把抓出，在抓的时候产品与手的接触是有先后的，5 件产品被“取出”离开整批产品，从空间距离上也有先后之分，只是其中的时间间隔非常短暂无法分清先后，只能冠之以“同时”.

由此看来，在进行无放回抽样时，无论是“同时取”还是“依次取”，对于不计次序的事件来说，概率的计算完全相同，这叫做**无放回抽样的同一性**.

7.2.4　概率的统计定义

古典概型的局限性是显而易见的. 例如测量某人的体重，样本空间是某区段内的所有实数，样本点是无限的；又如抛掷不均匀的骰子，基本事件虽然是有限的 6 个点数，但它们并非等可能. 这两个例子都不能用古典概率的公式（7.1）计算概率. 即使是古典概型，也有公式（7.1）中的 n 和 m 非常难以分析计算的时候，所以人们考虑要通过其他途径去计算概率.

“概率是事件发生的可能性大小”，虽然这样的理解很直观很合理，但再深入一步思考就发现这个认识很肤浅. 例如说某人投篮的命中率是 0.7，这个 0.7 是怎么得来的呢？仅仅是主观的粗略估计难免离客观实际很远. 最可靠的估计应该来自于以往大量的投篮实践，比如他曾经有过 100 次投篮经历，其中命中 70 次，两者的比值叫做**频率**，以频率来估计概率，其依据就比较充分，投篮次数越多，这个概率就越可靠. 可见概率的背后应该有大量的试验，这是支撑概率的条件.

一般地，随机现象在一次试验中无法断言其结果，但经过大量的试验，会呈现出某种规律性. 历史上有人做过抛掷硬币的试验，结果如表 7-1 所示，容易看出，随着抛掷次数的增加，正面朝上的频率围绕着一个确定的常数 0.5 作幅度越来越小的摆动并逐渐稳定于 0.5，这叫做**频率的稳定性**. 这种建立在大量试验基础上的规律称为**统计规律**. 概率的实质就是统计规律，可能性大小要与大量的试验相联系.

表 7-1　历次抛掷硬币试验的记录

试验者	抛掷次数 N	正面朝上的次数 M	正面朝上的频率 M/N
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5
维尼	30 000	14 994	0.499 8

以上从合理性与可靠性两方面说明能够用频率来估算概率，因而有了**概率的统计定义**：

$$P(A)=\lim_{\to\infty}\frac{M}{N} \tag{7.6}$$

式中 N 是在相同条件下重复试验的次数，M 是在 N 次试验中事件 A 出现的次数，比值 M/N 称为事件 A 出现的**频率**．频率的稳定性不仅被皮尔逊等人的实验所证明，而且后来又有数学家从理论上证明了这一事实，因此（7.6）式是计算概率的普遍适用的公式．

不过用公式（7.6）计算概率也有一定的困难，比如要做大量的试验，而且得到的仅仅是概率的近似值，又没有简便的误差估计方法．但这些困难丝毫不会影响公式（7.6）的重要意义，它的最大作用在于确立了概率的统计观点．

概率从哪里来？概率来自于统计规律，概率的取值需要有统计的支撑．如果一个从未打过枪的人，声称他的射击命中率是 0.7，那么他不是在吹牛，便是对概率一无所知，或者仅是开开玩笑而已．类似的例子很多，比如“明天下雨的可能性是 60%”，同样一句话，出自于气象部门，往往背后有大量的统计资料，而出自于其他人，则多半是一句不必负责任的空话、套话．

概率到哪里去？概率对实践的指导意义正是在于它需要面对大量的统计试验．还是以射击为例，命中率 0.7，对于一次射击意义不大，因为 0.7 不能保证命中目标，即使打不中，也不能否定这个 0.7．概率等于 0.7 的意义在于：若进行多次射击，他将有 70%左右的次数会击中目标；反之，命中频率与 70%差异很大，就有理由怀疑这个 0.7．再以投币为例，正面朝上的概率是 0.5，其依据是硬币的质地均匀（等可能性），那么均匀的依据又是什么呢？也许可以用精密的仪器来证明金属内部的均匀性，但是硬币两侧图案的不同已经造成了不均匀，等可能性该用什么来保证呢？正确的答案也许让人始料不及：硬币的均匀性应该通过大量的试验来证明，如果按公式（7.6）计算出的概率不是 0.5，就可以怀疑硬币的均匀性．在这里，因果关系颠倒过来了．

概率对于一次试验意义不大，但是当概率很大（超过 0.9）或很小（小于 0.1）时，对一次试验还是有指导意义的．可以认为小概率事件在一次试验中基本上不会发生，这就是**小概率原理**．比如飞机失事的报道很多，但是人们仍然向往坐飞机出行，而且不会有人在登机前与家人作生死话别，这是因为发生空难的概率太小了，仅为百万分之几．又比如人们在做决策时，有 90%以上的把握，都会断言“不出意外的话肯定成功”．不过应当指出的是：小概率原理不能保证没有风险，以概率的观点看问题，凡有随机因素，便不应该指望有绝对的把握，对此要有清醒的认识．此外，小概率原理面对大量的试验反而不适用，比如航空公司安排的航班数以千百计，万万不能因小概率原理而对飞行安全掉以轻心．

7.3　事件的独立性

7.3.1　条件概率

设 A，B 是两个事件，在已知事件 A 发生的条件下，求事件 B 发生的概率，称为**条件概率**，记为 $P(B|A)$，相应地把 $P(B)$ 称为**原概率**，两者的概率意义不同．

在计算条件概率 $P(B|A)$ 时，因为 A 已发生，排除了 A 之外的样本点，所以总的可能数应该是 m （参看图 7-1），而对事件 B 有利的可能数也要限制在 A 之内，应该是 m ，于是根据（7.5）式，可得

$$P(B|A)=\frac{m_{AB}}{m_A}\text{．}\tag{7.7}$$

显然这与 $p(B)=m/n$ 在数值上是不一定相等的．在（7.7）式右端分式中，分子分母同除以样本空间的容量 n，立得

$$P(B|A)=\frac{P(AB)}{P(A)}\text{．}\tag{7.8}$$

公式（7.7）和（7.8）都可以用来计算条件概率，当然要求分母非零即 $P(A)>0$．

例 7.8　一批种子的发芽率为 0.9，播种后能成长为活苗的概率为 0.72．当你看到一粒种子发了芽时，对它成活的可能性有何估算？

解　设 $A=$“种子发芽”，$B=$“种子成活”，则 $P(A)=0.9$；由于种子成为活苗必须先发芽，因此另一个概率实际上是 $P(AB)=0.72$（如果换一个角度理解为 $P(B)=0.72$，那么从 $B\subset A$ 知 $AB=B$，也可以推得 $P(AB)=0.72$）．现在已知种子发了芽，求它成活的概率，即求 $P(B|A)$，根据公式（7.8）计算

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{0.72}{0.9}=0.8\text{，}$$

你对发芽种子应抱有 80%的成活期望．

7.3.2　乘法公式

将公式（7.8）稍作变形，可得

$$P(AB)=P(A)P(B|A)\quad(P(A)>0)\text{．}\tag{7.9}$$

事件 A，B 的地位是对等的，因此将它们互换位置，即得

$$P(AB)=P(B)P(A|B)\quad(P(B)>0)\text{．}\tag{7.10}$$

公式（7.9）和（7.10）称为概率的**乘法公式**．乘法公式表明，事件乘积的概率等于各事件概率的乘积，不过后一个概率要添加条件．

乘法公式可推广到 3 个事件．先利用结合律将 3 个事件看作两个事件，得到 $P(ABC)=$

$P(A)P(BC|A)$，再对后一个概率应用乘法公式并注意添加条件，于是

$$P(ABC)=P(A)P(B|A)P(C|AB) \quad (P(AB)>0). \tag{7.11}$$

推广到更多事件的**乘法规则**是：事件乘积的概率等于各事件概率的乘积，不过后面的概率要不断地添加前款事件作为条件.

例 7.9 袋中有 52 个球，其中有 13 个红球，从中随机地依次取 3 个球（无放回）.

（1）求 3 个都是红球的概率；

（2）求前 2 个是红球，第 3 个不是红球的概率.

解 （1）设 $A_i=$“第 i 次取到红球”（$i=1,2,3$）. 3 个都是红球，意为 A_1,A_2,A_3 同时发生，即求概率 $P(A_1A_2A_3)$，利用乘法公式（7.11）得到

$$P(A_1A_2A_3)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)=\frac{13}{52}\times\frac{12}{51}\times\frac{11}{50}=0.01294.$$

这里条件概率是利用公式（7.7）计算的，如 $P(A_3|A_1A_2)$ 是在前两次都取到红球的条件下第 3 次取到红球的概率，因为第 3 次面对着 50 个球，其中只有 11 个红球，所以这时取到红球的概率当然是 11/50. 本例也可以把“依次”抽取当作“同时”抽取，按照例 7.7 的做法计算概率

$$P(A_1A_2A_3)=\frac{C_{13}^3}{C_{52}^3}=\frac{13\times12\times11}{52\times51\times50}=0.01294.$$

得到同样的结果应在意料之中，这是因为无放回抽样具有同一性，“同时”抽取和“依次”抽取并无本质的不同.

（2）所求概率是 $P(A_1A_2\bar{A}_3)$，运用乘法规则，

$$P(A_1A_2\bar{A}_3)=P(A_1)P(A_2|A_1)P(\bar{A}_3|A_1A_2)=\frac{13}{52}\times\frac{12}{51}\times\frac{39}{50}=0.04588.$$

本例的事件与次序有关，但仍然可以当作“同时”抽取来对待. 引入另一个事件 $B=$“恰有 2 次抽到红球”，则

$$P(B)=\frac{C_{13}^2C_{39}^1}{C_{52}^3}=\frac{13\times12\times39\times3}{52\times51\times50}.$$

显然 $B=\bar{A}_1A_2A_3+A_1\bar{A}_2A_3+A_1A_2\bar{A}_3$，此式右端是 3 个事件的和，这 3 个事件互不相容且机会均等，根据加法公式可知

$$P(A_1A_2\bar{A}_3)=\frac{1}{3}P(B)=\frac{13\times12\times39}{52\times51\times50}=0.04588,$$

得到同样的结果应该顺理成章. 当然，对于与次序有关的事件，利用乘法规则计算概率比较方便一些.

7.3.3 事件的独立性

如果两个事件 A 和 B，其中任何一个事件是否发生都不影响另一个事件发生的概率

（如烧香和下雨），则称事件 A，B **相互独立**．所谓不影响概率，当然指的是

$$P(B\mid A)=P(B)，\quad P(A\mid B)=P(A)，\tag{7.12}$$

即条件概率中的条件不起作用，条件概率等于原概率．在这种情况下乘法公式变成

$$P(AB)=P(A)P(B)．\tag{7.13}$$

两个事件不独立，称为**相依**．公式（7.13）称为**独立条件下的乘法公式**．不难看出，从（7.13）式出发通过公式（7.9）和（7.10）可推得（7.12）式，所以（7.13）式是事件 A，B 相互独立的充分必要条件，有的书上干脆用（7.13）式来定义独立性．

根据独立性的概念可知，事件 A，B 相互独立，等价于事件 A，$\overline{B}$ 相互独立，也等价于事件 $\overline{A}$，B 相互独立，还等价于事件 $\overline{A}$，$\overline{B}$ 相互独立．

公式（7.13）可以推广到多个事件，比如 3 个事件 A，B，C 相互独立，则乘法公式（7.11）就成了

$$P(ABC)=P(A)P(B)P(C)．\tag{7.14}$$

独立条件下的乘法规则是：相互独立事件乘积的概率等于各事件概率的乘积．独立性为计算乘积概率带来了很大的便利．比如当事件 A，B 相互独立时，加法公式（7.2）就成了

$$P(A+B)=P(A)+P(B)-P(A)P(B)．$$

例 7.10　两个射手射击同一目标，已知甲射中目标的概率为 0.9，乙射中目标的概率为 0.8，求两人各射一次而目标被击中的概率．

解　设 $A=$“甲击中目标”，$B=$“乙击中目标”，则所求的概率是 $P(A+B)$．一般情况下，两射手的射击互不影响，即 A，B 相互独立．已知 $P(A)=0.9$，$P(B)=0.8$，所以

$$P(A+B)=P(A)+P(B)-P(AB)=0.9+0.8-0.9\times0.8=0.98．$$

事件的独立性往往可以像本例那样根据经验来判断，即把不存在明显影响的若干事件看作是相互独立的．当然这在特殊情况下也许会出现偏差，比如本例中两人先后射击，后一射手心理压力很重以至不能正常发挥，这时 A，B 就不独立了．应该指出的是，在随机试验中，事件的独立性是很普遍的现象，一般情况下实际经验的判断大多是有效的；反之，若按独立性计算得到的概率与实际情况有较大的差异，则应怀疑原先的独立性判断，这正是数理统计中独立性检验的理论依据．

例 7.9 中的摸球试验是无放回抽样，即每次摸出的球不再放回．如果每次摸出一个球后放回袋中，经充分混合后再作下一次摸球，那么就成了**有放回抽样**．在无放回抽样中，因为每次摸球的结果直接影响到剩余球的构成状况，所以前后两次摸球是相依的（不独立）；而有放回抽样时，球的构成状况始终相同，因此前后两次摸球相互独立．

摸球试验是很重要的随机试验，几乎所有的随机试验都可以用摸球试验来模拟，所以对于无放回抽样和有放回抽样的独立性要有清醒的认识．在实际问题中，如果袋中球的数目相对于抽样次数而言非常庞大，那么每次无放回抽样对袋中球的构成状况影响甚微，这时无放回抽样可视作有放回抽样．比如抽样检验产品，尽管是无放回的，但计算概率时通常都应用独立性条件．

例 7.11 3 台设备同时开动，在一小时内它们不出现故障的概率分别为 0.9，0.8 和 0.8，求在一小时内至少有一台设备发生故障的概率．

解 设 A，B，C 分别表示一小时内这 3 台设备不出现故障，则 $P(A)=0.9$，$P(B)=0.8$，$P(C)=0.8$．至少有 1 台设备发生故障是 $\overline{A}+\overline{B}+\overline{C}$，该事件的对立事件是 ABC，即 3 台设备都不出现故障．3 台设备的运行显然互不影响，即 A，B，C 相互独立，所以

$$P(\overline{A}+\overline{B}+\overline{C})=1-P(ABC)=1-P(A)P(B)P(C)=1-0.9\times0.8\times0.8=0.434.$$

计算结果表明，虽然单个设备的可靠性都很好，无故障概率均不小于 80%，但整体的风险却高达 43.4%，这是在可靠性设计中具有广泛应用的概率原理．

在本例的计算中，应注意事件 $\overline{A}$，$\overline{B}$，$\overline{C}$ 并非互不相容，所以不能应用加法公式(7.4)．若将事件 $\overline{A}+\overline{B}+\overline{C}$ 分解成互不相容的事件之和，则会产生 7 项之多，由此可见，利用逆事件求概率既简便又能保证思路的清晰，当涉及的事件较多时，这种简便性尤为突出．

这里特别要强调互不相容和相互独立这两个概念的区别．不相容是指不能同时发生，简称“不相交”，而相互独立并不排除同时发生，而是指发生与否互不影响，简称“不影响”，两者不能混为一谈．一般地，当 $P(A)>0$，$P(B)>0$ 时，若 A，B 不相容则必相依（不独立），反之，若 A，B 相互独立则必相容，即 $P(AB)=0$ 和 $P(AB)=P(A)P(B)$ 不可能同时成立．

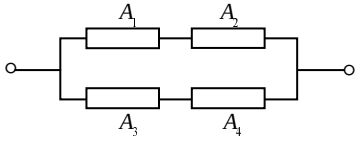

图 7-2 电器元件的串并联

例 7.12 将 4 个电器元件按图 7-2 连接，然后接入电路．每个元件的可靠度（正常工作的概率）均为 r（$0<r<1$），元件发生故障则出现断路．求图 7-2 电路能正常工作不出现断路的概率．

解 以 A_i $(i=1,2,3,4)$ 分别表示 4 个元件正常工作，则 $P(A_i)=r$ $(i=1,2,3,4)$．串联电路正常工作意味着电路中各部分同时正常，而并联电路正常工作则意味着电路中各分路至少有一个正常，所以图 7-2 电路正常工作即事件 $A_1A_2+A_3A_4$．4 个元件是否出现故障应该相互独立，因此

$$\begin{aligned}P(A_1A_2+A_3A_4)&=P(A_1A_2)+P(A_3A_4)-P(A_1A_2A_3A_4)\\&=P(A_1)P(A_2)+P(A_3)P(A_4)-P(A_1)P(A_2)P(A_3)P(A_4)\\&=r^2+r^2-r^4=r^2(2-r^2).\end{aligned}$$

本例的计算过程表明，串联电路会降低可靠性，而并联电路能提高可靠性．

7.3.4 全概率公式

1．全概率公式

如图 7-3 所示，将样本空间 Ω 分为 4 块 A_1，A_2，A_3，A_4，事件 B 也同时被分割．计算图

中各块的样本点数，不难得到

$$\begin{aligned}P(B)&=P(A_1B)+P(A_2B)+P(A_3B)+P(A_4B)\\&=P(A_1)P(B\mid A_1)+P(A_2)P(B\mid A_2)\\&\quad+P(A_3)P(B\mid A_3)+P(A_4)P(B\mid A_4)\text{．}\end{aligned}$$

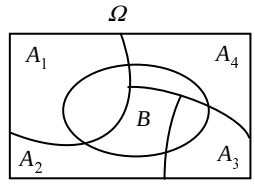

图 7-3　全概率公式示意图

类似地可以推广到一般情况：设 B 为任一事件，A_1, A_2, L, A_n 是把样本空间 Ω 分为 n 份的完备事件组，即满足 A_1, A_2, L, A_n 两两互不相容，$P(A_i)>0\ (i=1,2,\cdots,n)$，而且 $A_1+A_2+\text{L}+A_n=\Omega$，则

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B\mid A_i)\text{．}\tag{7.15}$$

公式（7.15）称为**全概率公式**．特别地，取 $n=2$，有

$$P(B)=P(A)P(B\mid A)+P(\bar{A})P(B\mid\bar{A})\text{．}\tag{7.16}$$

当不容易直接求概率 $P(B)$ 时，全概率公式往往很有效．全概率公式（7.15）、（7.16）体现了转移难点、分类计算概率的方法．

例 7.13　设仓库内有 10 箱产品，分别来自于甲厂 5 箱，乙厂 3 箱，丙厂 2 箱，而三个厂的次品概率依次为 0.1，0.07，0.05．从这 10 箱产品中任取一箱，再从此箱中取一件产品，求取得正品的概率．

解　本例分层次给出概率条件，用全概率公式最为恰当．设 B =“取得正品”，A_1 =“甲厂生产”，A_2 =“乙厂生产”，A_3 =“丙厂生产”，则 $P(A_1)=0.5$，$P(A_2)=0.3$，$P(A_3)=0.2$；$P(B\mid A_1)=1-0.1=0.9$，$P(B\mid A_2)=0.93$，$P(B\mid A_3)=0.95$．于是根据全概率公式：

$$\begin{aligned}P(B)&=P(A_1)P(B\mid A_1)+P(A_2)P(B\mid A_2)+P(A_3)P(B\mid A_3)\\&=0.5\times0.9+0.3\times0.93+0.2\times0.95=0.919\text{．}\end{aligned}$$

2．贝叶斯公式

在全概率公式中，关于事件 A_i 的信息是已知的，所以 $P(B\mid A_i)\ (i=1,2,\cdots,n)$ 称为**先验概率**．有时需反过来由结果事件 B 的信息来估计或验证原先的信息，即求概率 $P(A_i\mid B)$，这个概率称为**后验概率**．为了避免下标的混乱，将后验概率记作 $P(A_k\mid B)$．利用乘法公式（7.9）、（7.10），再结合全概率公式（7.15）可得

$$P(A_k\mid B)=\frac{P(A_kB)}{P(B)}=\frac{P(A_k)P(B\mid A_k)}{\sum_{i=1}^{n}P(A_i)P(B\mid A_i)}\ (k=1,2,\cdots,n)\text{，}\tag{7.17}$$

这个公式称为**贝叶斯公式**．在贝叶斯公式中，分母是 n 项之和，分子恰为这 n 项中的第 k 项．

例 7.14　在例 7.13 中，如果实际抽到的产品是正品，问所抽到的该产品是甲、乙、丙

厂生产的概率各为多少？

解 本例是求后验概率 $P(A_k \mid B)$ $(k=1,2,3)$，例 7.13 中计算 $P(B)$ 的 3 项之和为 $P(B)$ ＝0.45＋0.279＋0.19＝0.919，所以按贝叶斯公式，可算得：

$$P(A_1 \mid B)=\frac{0.45}{0.919}=0.4897\text{，}\quad P(A_2 \mid B)=\frac{0.279}{0.919}=0.3036\text{，}\quad P(A_3 \mid B)=\frac{0.19}{0.919}=0.2067\text{．}$$

后验概率往往被看作是在得到新的信息（B 已发生）后，对原概率 $P(A_1)=0.5$，$P(A_2)=0.3$，$P(A_3)=0.2$ 的一种修正．

习 题 七

1．设 A，B，C 表示 3 个事件，利用 A，B，C 及其运算表示下列事件．

（1）A 发生，B，C 都不发生；　（2）A，B 都发生，C 不发生；　（3）所有 3 个事件都发生；

（4）3 个事件中至少有 1 个发生；　（5）3 个事件都不发生；　（6）只有 B 发生；

（7）只有 B 不发生；　（8）不多于 1 个事件发生；　（9）不多于 2 个事件发生；

（10）3 个事件中至少有 2 个发生．

2．向指定的目标射击 3 枪，以 A_1，A_2，A_3 分别表示事件“第一、二、三枪击中目标”，试用 A_1，A_2，A_3 及其运算表示以下事件．

（1）只击中第 1 枪；　（2）只击中 1 枪；　（3）3 枪都未击中；　（4）至少击中 1 枪．

3．某村有 200 户人家，34 户没有孩子，98 户有 1 个孩子，49 户有 2 个孩子，19 户有多于 2 个孩子．从中任选 1 户人家，这户人家只有 1 个孩子的概率为多少？这户人家有至少 1 个孩子的概率为多少？

4．从一批由 37 件正品，3 件次品组成的产品中任取 3 件产品，求：

（1）3 件中恰有 1 件次品的概率；　（2）3 件全是次品的概率；　（3）3 件全是正品的概率；

（4）3 件中至少有 1 件次品的概率；　（5）3 件中至少有 2 件次品的概率．

又，如果抽取方式改为分 3 次抽取，每次无放回地取一件产品，则上述概率如何求？

5．某城市有 50%的住户订日报，有 65%的住户订晚报，有 85%的住户至少订这两种报纸中的一种．记者随机采访该城市的一家住户，发现其同时订两种报纸的概率是多少？

6．一大型超市声称，进入商店的小偷有 60%可以被电视监测器发现，有 40%被保安人员发现，有 20%被监测器和保安人员同时发现，试求小偷被发现的概率．

7．某公司有职工 210 名，对他们进行调查发现有 160 人会使用计算机，其中 78 人受过高等教育，而不会使用计算机的人中有 43 人未受过高等教育．现从所有职工中任选一人，求：

（1）他受过高等教育的概率；

（2）他不会使用计算机的概率；

（3）已知他没有受过高等教育，求他会使用计算机的概率；

（4）已知他会使用计算机，求他受过高等教育的概率；

（5）求他既会使用计算机又受过高等教育的概率；

（6）求他既不会使用计算机，又没有受过高等教育的概率．

8．对 100 家企业 2001 年、2002 年的经营情况进行调查，得到的结果是：有 55 家企业两年都盈利，有 15 家企业两年都亏损，其余的企业都为一年盈利、一年亏损，其中先盈后亏的企业有 20 家，现从中任选一家企业，求：

（1）它在 2002 年是盈利的概率；　　（2）它在 2001 年是亏损的概率；

（3）它连续两年是盈利的概率；　　（4）它连续两年是亏损的概率；

（5）已知它在 2001 年是盈利，求它在 2002 年是盈利的概率；

（6）已知它在 2001 年是亏损，求它在 2002 年是亏损的概率．

9．甲、乙两城市都位于长江下游，根据一百余年来气象的记录，知道甲、乙两城市一年中雨天占的比例分别为 20%和 18%，两地同时下雨的比例为 12%，问：

（1）乙市为雨天时，甲市为雨天的概率是多少？

（2）甲市为雨天时，乙市为雨天的概率是多少？

（3）甲、乙两城市至少有一个为雨天的概率是多少？

10．某种动物从出生活到 20 岁的概率为 0.8，活到 25 岁的概率为 0.4，问现年 20 岁的这种动物活到 25 岁的概率是多少？

11．一批零件共 100 件，其中次品 10 个，从这批零件中每次任取一个零件，取出的零件不再放回，求第 3 次才取到正品的概率．

12．两个电池 A 和 B 并联后再与电池 C 串联，构成一个复合电源接入电路，各电池是否发生故障相互独立，设电池 A，B，C 损坏的概率分别是 0.3，0.2，0.1，求电路因电池故障而无法供电的概率．

13．加工某一零件共需 4 道工序，设第 1、2、3、4 道工序的次品率分别是 2%、3%、5%、2%，假定各道工序是互不影响的，求加工出来的零件的次品率．

14．甲、乙、丙 3 人向同一飞机射击，设甲、乙、丙射中的概率分别为 0.4、0.5、0.7．

（1）求只有 1 人射中飞机的概率；

（2）求恰有 2 人射中飞机的概率；

（3）求 3 人都射中飞机的概率．

15．某人投篮的命中率仅为 0.3，有一天他去练习投篮，打算投不中接着投，投中为止．

（1）他在第 4 次才投中的概率是多少？

（2）他投篮不超过 4 次就结束练习的概率是多少？

16．有 6 个电子元件，它们工作时发生故障的概率都是 r（$0<r<1$），若发生故障则出现断路．

（1）将 6 个元件两两串联形成 3 个组合件，然后把这 3 个组合件并联起来接入电路，求电路工作时出现断路的概率；

（2）将 6 个元件 3 个一组并联形成 2 个组合件，然后把这 2 个组合件串联起来接入电路，求电路工作时不出现断路能正常工作的概率．

17．某工厂由甲、乙、丙 3 个车间生产同一种产品，每个车间的产量分别占全厂的 25%、35%、40%，各车间产品的次品率分别为 5%、4%、2%，求全厂产品的次品率．

18．上题中，如果从全厂总产品中抽取一件产品抽得的是次品，求它依次是甲、乙、丙车间生产的概

率．

19．2 台车床加工同样的零件，第 1 台加工后的废品率为 0.03，第 2 台加工后的废品率为 0.02，加工出来的零件放在一起，已知这批加工后的零件中，由第 1 台车床加工的占 2/3，由第 2 台车床加工的占 1/3，从这批零件中任取一件，求这件是合格品的概率．

20．单项选择题

(1) 设有 3 个事件 A，B，C，用 A，B，C 表示“至多有 3 个事件发生”为 （ ）

A．$\overline{A}+\overline{B}+\overline{C}$； B．$\overline{ABC}$； C．$A\overline{B}\overline{C}+\overline{A}B\overline{C}+\overline{A}\overline{B}C$； D．$\Omega$．

(2) 在某学校学生中任选一名学生，设 A =“选出的学生是男生”，B =“选出的学生是三年级学生”，C =“选出的学生是篮球运动员”，则 ABC 的含义是 （ ）

A．选出的学生是三年级男生； B．选出的学生是三年级男子篮球运动员；

C．选出的学生是男子篮球运动员； D．选出的学生是三年级篮球运动员．

(3) 掷一颗骰子的试验，观察其出现的点数，记 A =“掷出偶数点”，B =“掷出奇数点”，C =“掷出的点数小于 5”，D =“掷出 1 点”，则下述关系错误的是 （ ）

A．$B=\overline{A}$； B．A 与 D 互不相容； C．$C=\overline{D}$； D．$\Omega=A+B$．

(4) 某事件的概率为 0.2，如果试验 5 次，则该事件 （ ）

A．一定会出现 1 次；B．一定出现 5 次； C．至少出现 1 次； D．出现的次数不确定．

(5) 对一个有限总体进行有放回抽样时，各次抽样的结果 （ ）

A．相互独立； B．是相容的； C．互为逆事件； D．不相容但非逆事件．

(6) 若 $P(A)=0.5$，$P(B)=0.5$，则 $P(A+B)=$ （ ）

A．0.25； B．1； C．0.75； D．不确定．

(7) 已知 $P(A)=0.4$，$P(B)=0.3$，$P(A+B)=0.6$，则事件 A 和 B （ ）

A．相容但不独立； B．独立但不相容； C．独立且相容； D．不独立也不相容．

(8) 玩 3 种电脑游戏，都有机会获得小奖品，各种游戏是否获奖相互独立，获奖概率分别是 0.2、0.3、0.4．某人打算把 3 种游戏都玩一遍，则他获奖的概率是 （ ）

A．0.024； B．0.664； C．0.876； D．0.900．

(9) 3 人抽签决定谁可以得到惟一的一张足球票．现制作 2 张假票与真票混在一起，3 人依次抽取，则 （ ）

A．第 1 人获得足球票的机会最大； B．第 2 人获得足球票的机会最大；

C．第 3 人获得足球票的机会最大； D．3 人获得足球票的机会相同．

(10) 已知 $P(A)=0.5$，$P(B)=0.4$，$P(A+B)=0.6$，则 $P(A|B)=$ （ ）

A．0.2； B．0.45； C．0.6； D．0.75．

第 8 章　随机变量及其概率分布

8.1　离散型随机变量及其分布律

8.1.1　随机变量

随机试验的结果往往表现为数量，如抛掷骰子出现的是点数，摸球试验摸出的是红球数（参看例 7.9），射击目标的结果是击中次数，观测潮位看到的是水位高度．有些随机试验的结果并不直接表现为数量，但可以将其数量化，如抽牌试验中将牌张编号（参看例 7.3），抽出的便是编号数．又如观察婴儿的性别，用 0 表示出生女婴，用 1 表示出生男婴，则观察到的就是数 0 或数 1．

用字母 X 表示试验的数值结果，X 可能取许多不同的数值，因此 X 是变量．另一方面，试验之前无法断定 X 取什么值，试验之后可以观察到它的取值，即变量 X 具有随机性，其取值有一定的概率表现，所以将随机试验的数值结果称为**随机变量**．随机变量通常用大写字母 X, Y, Z 等表示．

例 8.1　投币试验．抛出一枚硬币，X 为“出现正面的次数”，则 X 的可能取值为 1，0，即 $(X=1)$ =“正面朝上”，$(X=0)$ =“反面朝上”，显然 X 是随机变量，并且 $P(X=1)=P(X=0)=0.5$．

例 8.2　抽牌试验．将 52 张牌按例 7.3 的方式编号，随机地抽一张牌，设 X 为“抽得牌张的编号”．显然 X 是随机变量，它有 52 个可能的取值，其中$(14\leq X\leq 26)$=“抽到红心”，并且 $P(14\leq X\leq 26)=0.25$．

从以上两例看出，随机变量的取值或取值范围表示随机事件，并且有对应的概率．随机变量本身并不是事件，记号 $P(X)$ 是没有意义的．研究随机变量最主要的是分析它的概率特征，即研究随机变量各种取值情况的概率．随机变量的全部概率特征称为随机变量的**概率分布**．

8.1.2　离散型随机变量

如果随机变量 X 的取值可以一一列出，则称 X 是**离散型随机变量**．离散型随机变量并不排除可能取无穷多个数值，只是要求这些取值可以按某种顺序一一列出而不是连成一片的．设 X 的可能取值为 x_k （$k=1,2,\mathrm{L}$），并且相应的概率

$$P(X=x_k)=p_k \quad (k=1,2,\mathrm{L}) \tag{8.1}$$

都已知道，则该随机变量的概率特征就完全搞清楚了．所以将（8.1）式称为离散型随机变量 X 的**分布律**，写成矩阵（表格）形式为

$$X \sim \begin{pmatrix} x_1 & x_2 & \mathrm{L} & x_k & \mathrm{L} \\ p_1 & p_2 & \mathrm{L} & p_k & \mathrm{L} \end{pmatrix}. \tag{8.2}$$

这里“～”号表示“服从于”的意思，矩阵的第 1 行是 X 的所有可能取值，第 2 行是这些取值所对应的概率．分布律有形象化的矩阵形式（8.2），故又称作**分布列**．

在分布律中，所有的概率均应非负，另一方面根据可加性，所有概率之和表示必然事件的概率，应该等于 1，因此分布律具有以下性质：

（1）**非负性** $p_k \geq 0$（$k=1,2,\mathrm{L}$）；

（2）**完备性** $\sum\limits_k p_k = 1$．

例 8.3 一种有奖储蓄，20 万户为一开奖组．设特等奖 20 户，每户奖金 4000 元；一等奖 120 户，每户奖金 400 元；二等奖 1200 户，每户奖金 40 元；末等奖 4 万户，每户奖金 4 元．求一个储蓄户得奖额 X 的分布律．

解 注意 X 是“得奖额”（不是“奖等”），它的所有可能取值为 4000，400，40，4，0（最后一个值 0 不可遗漏）．对应的概率是获奖名额与总户数（20 万）之比，容易求得分布律为

$$X \sim \begin{pmatrix} 4000 & 400 & 40 & 4 & 0 \\ 0.0001 & 0.0006 & 0.006 & 0.2 & 0.7933 \end{pmatrix},$$

其中最后一个概率 0.7933 利用分布律的完备性求得．可见绝大多数储户是不中奖的，这是各种摸奖、抽奖、彩票活动的共同特点．

例 8.4 10 件产品中，有 3 件次品，7 件正品．现从这 10 件产品中任取 2 件产品，记 X 是“抽得的次品数”，求 X 的分布律．

解 求分布律首先要认清 X 的确切含义，正确地分析其所有的可能取值．本例中，抽得的次品数 X 可能取值为 0，1，2．对应的概率可参照例 7.7，利用组合数求得：

$$P(X=0)=\frac{C_7^2}{C_{10}^2}=\frac{7}{15},\quad P(X=1)=\frac{C_3^1C_7^1}{C_{10}^2}=\frac{7}{15},\quad P(X=2)=\frac{C_3^2}{C_{10}^2}=\frac{1}{15},$$

于是 X 的分布律为 $X \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{7}{15} & \frac{7}{15} & \frac{1}{15} \end{pmatrix}$．这个分布律显然满足非负性和完备性的要求．

例 8.5 求 a 的值，使 X 的分布律为 $P(X=k)=a(0.4)^k$（$k=1,2,\mathrm{L}$）．

解 这里 X 的可能取值有无限多个，因而完备性表现为无穷级数之和．由

$\sum_{k=1}^{\infty} a(0.4)^k = a \times \frac{0.4}{1-0.4} = 1$，推出 $a = 1.5$．

离散型随机变量的分布律可以列表（如例 8.3，例 8.4），也可用公式（8.1）表示（如例 8.5）．这两种不同的表现形式，其本质是相同的，都给出了一个以概率为函数值，自变量取离散值的离散型函数．这个函数作为分布律刻画了随机变量的概率特征，描述出离散型随机变量的概率分布，因而不同的随机变量有不同的分布律．

以下讨论三种常见的分布：两点分布、二项分布、泊松分布．

8.1.3　两点分布

若随机变量 X 的可能取值仅两点 0 和 1，且 $P(X=1)=p$，则 X 的分布律为

$$X \sim \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}, \tag{8.3}$$

其中 $q=1-p$，X 具有这样的分布律，就称 X 服从参数为 p 的**两点分布**，两点分布又称为**0－1 分布**．两点分布的分布律可用公式表示为

$$P(X=k)=p^k(1-p)^{1-k}=p^k q^{1-k} \quad (k=0,1).$$

例 8.6　袋中有 6 只白球和 4 只红球，从中任取一只，X 为“取得的白球数”，求 X 的分布律．

解　显然 X 只能取 1（取得白球）和 0（取得红球），并且 $P(X=1)=0.6$，所以 X 服从参数为 0.6 的两点分布，其分布律为 $X \sim \begin{pmatrix} 1 & 0 \\ 0.6 & 0.4 \end{pmatrix}$．

两点分布是最简单的分布，同时又是很重要的分布，因为任何随机试验都与两点分布相联系：设 A 是试验中的某一事件，X 是“一次试验中 A 出现的次数”，显然 $(X=1)$ 表示 A 发生而 $(X=0)$ 表示 A 不发生，若 $P(A)=p$，则 X 的分布律由矩阵（8.3）给出，即 X 服从参数为 p 的两点分布．

以后还会看到，将一个复杂的随机变量分解成两点分布之和，可以在相关的概率分析中起到事半功倍的作用．

8.1.4　二项分布

1．伯努利（Bernoulli）试验

将随机试验在相同的条件下独立地重复 n 次，观察事件 A 出现的次数，称为**伯努利试验**，或 n **次重复独立试验**，事件 A 出现 k 次的概率记为 $P_n(k)$．这里的“独立”是指前后各次试验互不干扰、互不影响．

伯努利试验是很常见的试验，如投篮 n 次观察投中多少次，有放回地抽样 n 次（抽牌、摸球、检验产品等）观察抽中的次数等．

例 8.7 产品次品率为 0.2，从中有放回地抽取 5 次，每次取一个，求出现 2 次次品的概率．

解 有放回抽样具有独立性，属于伯努利试验，本例即求 $P_5(2)$．设 A_i =“第 i 次抽样中出现次品”，有放回抽样中单次概率都相同，即

$$P(A_i)=0.2,\quad P(\bar{A}_i)=0.8\quad (i=1,2,3,4,5).$$

5 次抽样情况可以是

$$A_1A_2\bar{A}_3\bar{A}_4\bar{A}_5,\quad A_1\bar{A}_2A_3\bar{A}_4\bar{A}_5,\quad \bar{A}_1\bar{A}_2A_3\bar{A}_4A_5,\ \text{L},$$

相当于在 5 个位置中安放 2 个次品位置，这样的情况共有 C_5^2 种，互不相容，其概率按独立性条件计算都是 $0.2^2\times0.8^3$，所以由加法公式得 $P_5(2)=C_5^2 0.2^2 0.8^3$．

一般地，在伯努利试验中，A 出现的概率是 p，则

$$P_n(k)=C_n^k p^k(1-p)^{n-k}\quad (k=0,1,2,\text{L},n).\tag{8.4}$$

这种概率模型称为**伯努利概型**．

2．**二项分布**

如果随机变量 X 的分布律由公式（8.4）给出，即

$$P(X=k)=C_n^k p^k(1-p)^{n-k}=C_n^k p^k q^{n-k}\quad (k=0,1,2,\text{L},n),\tag{8.5}$$

式中 $q=1-p$，则称 X 服从参数为 n 和 p 的**二项分布（或伯努利分布）**，记为 $X\sim B(n,p)$．这里的 B 是 Bernoulli（伯努利）的第一个字母．二项分布的名称来源于牛顿二项展开式

$$1=(p+q)^n=\sum_{k=0}^{n}C_n^k p^k q^{n-k},$$

公式（8.5）给出的 $n+1$ 个概率值恰是上述二项展开式的各项，二项分布的完备性也同时由此得以验证．

结合（8.4）式和（8.5）式可知，当 X 是 n 次重复独立试验中事件 A 出现的次数，且 $P(A)=p$ 时，X 服从二项分布，其中 n 是重复试验的总次数，p 是单次试验中事件 A 的概率．当 $n=1$ 时，二项分布退化为两点分布，因此两点分布可记为 $X\sim B(1,p)$．

例 8.8 某篮球运动员投篮命中率的统计值是 0.7，现投篮 10 次，问他投中 7 次的概率是多少？

解 投篮属于重复独立试验，所以投中次数 X 服从二项分布．其中参数 $n=10$ 是投篮总次数，$p=0.7$ 是单次投中的概率，所求概率按公式（8.5）计算为

$$P(X=7)=C_{10}^7 p^7 q^3=120\times0.7^7\times0.3^3=0.2668.$$

可见“命中率为 0.7”与“10 投 7 中”相去甚远，后者只有不到 27%的可能性，发生的概

率是很小的.

二项分布的概率计算有现成的公式 (8.5), 遇到实际问题要注重分析是否属于具有“重复、独立”特征的伯努利概型，以便应用二项分布的现成结论. 下面是几个二项分布的典型环境：

(1) 固定总次数的射击、投篮等独立运作；

(2) 有放回抽样问题；

(3) 总量庞大的无放回抽样问题（近似于有放回抽样）；

(4) 多台同类仪器或设备并联使用中的故障问题.

这里将设备运行问题归入伯努利概型是因为其实质上相当于总量庞大的无放回抽样问题.

例 8.9　日光灯管正常使用一个月，会有 1%的损坏率. 一个教室内安装了 16 支日光灯，问一个月内有 3 支或 3 支以上灯管不亮的概率是多少？

解　这是设备并联使用问题，可按伯努利概型来对待. 试验次数是日光灯总数 $n=16$，单独一根灯管损坏的概率是 $p=1\%=0.01$，设 X 是损坏的灯管数，则 $X \sim B(16,0.01)$. 所求的概率是 $P(X \geq 3)$，利用逆事件求概率比较简便，因为

$$\begin{aligned} P(X \leq 2) &= P(X=0)+P(X=1)+P(X=2) \\ &= C_{16}^{0}p^{0}(1-p)^{16}+C_{16}^{1}p^{1}(1-p)^{15}+C_{16}^{2}p^{2}(1-p)^{14} \\ &= 0.85156+0.13761+0.01042=0.99949, \end{aligned}$$

所以 $P(X \geq 3)=1-P(X \leq 2)=0.00051$. 计算表明，除非有人为的非正常因素，3 支以上灯管不亮几乎是不可能的.

例 8.10　一批螺丝次品率为 0.05，商店将 10 个螺丝包成一包出售，并承诺其中多于一个次品即可退货. 问商店将承受多大的退货率？

解　设 X 是一包螺丝中次品的个数，退货率即概率 $P(X>1)$. 本例是任取 10 个螺丝（一包）的无放回抽样问题，因为一批螺丝总量庞大，所以可视作伯努利概型，即 $X \sim B(10,0.05)$. 利用逆事件求概率比较简便：

$$P(X>1)=1-P(X \leq 1)=1-P_{10}(0)-P_{10}(1)=1-0.95^{10}-10\times 0.05\times 0.95^{9}=0.0861,$$

即退货率约为 8.6 %.

8.1.5　泊松（Poisson）分布

如果随机变量 X 的可能取值为 $0,1,2,\mathrm{L},k,\mathrm{L}$ （无穷），且分布率为

$$P(X=k)=\frac{\lambda^{k}}{k!}\mathrm{e}^{-\lambda} \quad (\lambda>0,\ \ k=1,2,\mathrm{L}), \tag{8.6}$$

则称 X 服从参数为 λ 的**泊松分布**，记为 $X \sim P(\lambda)$. 这里的 P 是 Poisson（泊松）的第一个字母. 验证泊松分布 (8.6) 的完备性需要用到幂级数知识，此处从略.

泊松分布来源于“排队现象”，刻画稀有事件出现的概率．如某时间段的电话呼叫、机器纺织的纱线断头、大型商店的顾客到来、交通路口的车辆通过等，都可以用泊松分布来描述．

理论推导表明，当 n 很大，p 很小而 np 大小适中时，二项分布近似于泊松分布，即有近似式

$$P_n(k)\approx\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\quad(\lambda=np)\text{，}\qquad C_n^k p^k(1-p)^{n-k}\approx\frac{(np)^k}{k!}\mathrm{e}^{-np}\text{．}$$

这个近似式为计算二项分布的概率提供了方便，尤其是当 n 很大时，计算组合数 C_n^k 有困难，而利用 $\lambda=np$ 转换成泊松公式就比较容易．

为了使计算更为方便，人们专门编制了泊松分布数值表以备查阅（见附表 2）．泊松分布数值表通常给出向后累计的概率，求单个概率或者向前累计的概率应注意还要另外做减法．

例 8.11 某厂有同类设备 80 台，发生故障的概率都是 0.01，当一台设备发生故障时需要有一个工人来维修处理．现考虑用两种方式进行管理：

（1）由 1 名工人负责维修 20 台设备；

（2）由 3 名工人负责维修 80 台设备．

试分别求出设备发生故障而需要等待维修的概率．

解 （1）设 X 是 20 台设备中发生故障的设备数，由于 X 面对一个维修工人，所以当 $X>1$ 时即出现等待维修的现象，所求的概率显然是 $P(X>1)$．设备运行问题服从二项分布，即 $X\sim B(20,0.01)$，但 $n=20$ 较大，$p=0.01$ 较小而 $\lambda=np=0.2$ 适中，所以可认为近似地有 $X\sim P(0.2)$，查附表 2 可直接得到概率为 $P(X>1)=P(X\geqslant 2)=0.0175$．

（2）设 X 是 80 台设备中发生故障的设备数，X 面对 3 个维修工人，故所求的概率是 $P(X>3)$，由于 $X\sim B(80,0.01)$ 而 $\lambda=80\times0.01=0.8$ 大小适中，所以近似地有 $X\sim P(0.8)$，查附表 2 可直接得到概率为 $P(X>3)=P(X\geqslant 4)=0.0091$．

经比较发现，后一种管理方式不仅每人平均负责维修的设备数有所增加，而且维修等待的概率大大降低，因此由多人共同看管设备的管理经济效益要好于单人分别看管设备．由此可见，用概率论知识建立模型，能够解决某些有效运用人力、物力资源的问题．

8.2 连续型随机变量及其概率密度

8.2.1 连续型随机变量

1．概率密度函数

若随机变量 X 的所有可能取值不能一一列出，而是连成一片形成区间，则 X 就可成为

连续型随机变量．如测量零件尺寸、检验电池寿命、观察降雨量等，都要归结为连续型随机变量．

由于 X 要面对一大片的取值，单点取值的概率 $P(X=x)$ 必然微乎其微，而且无法一一列出，所以不能用建立分布律的方法来对待连续型随机变量，应该转而研究随机变量取值范围的概率 $P(a<X\leq b)$ ．

离散型随机变量取值范围的概率曾在例 8.9 中涉及，那是一些概率的累计值，即

$$P(a<X\leq b)=\sum_{x_k\in(a,b]}P(X=x_k)=\sum_{x_k\in(a,b]}p_k\ ,$$

其中 $p_k=P(X=x_k)$ 是由（8.1）式确定的反映随机变量特征的离散型函数．由此引申出去，连续型随机变量取值范围的概率应当是一个“连续”的累计过程．按照微积分的观点，连续的累计过程是和的推广，含有无限分割、无穷积累的思想，需要用积分来表现，而被积函数与上面和式中的 p_k 相类似，应该是反映该随机变量特征的连续型函数，即

$$P(a<X\leq b)=\int_a^b f(x)\mathrm{d}x\ .\qquad(8.7)$$

式中体现随机变量特征的函数 $f(x)$ 称为随机变量 X 的**概率密度函数**，简称**密度函数**或**密度**．注意（8.7）式中大写字母 X 表示随机变量，小写字母 x 表示积分变量，两者切勿相混，尤其是 $(a<x\leq b)$ 不能表示随机事件，因而 $P(a<x\leq b)$ 是没有意义的．

知道概率密度函数 $f(x)$ 就可以通过积分计算有关随机变量 X 的各种概率，所以密度 $f(x)$ 全面描述了连续型随机变量 X 的概率特征，因而不同的随机变量有不同的密度函数．

2．密度函数的性质

根据概率的非负性和完备性可知，密度函数具有如下和离散型随机变量的分布律类似的性质：

（1）**非负性**　$f(x)\geq 0$；　　（2）**完备性**　$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$ ．

例 8.12　设 $f(x)=\begin{cases}k(3x-x^2), & 0\leq x\leq 3\\ 0, & \text{其他}\end{cases}$ 是随机变量 X 的概率密度函数，求 k 的值及 $P(1<X\leq 5)$，$P(X\leq 1)$ ．

解　由完备性（注意分段函数的积分处理）知，

$$1=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_0^3 k(3x-x^2)\mathrm{d}x=k\left[\frac{3}{2}x^2-\frac{1}{3}x^3\right]_0^3=\frac{9}{2}k\ ,$$

所以 $k=2/9$ ．计算连续型随机变量的概率要做积分，根据（8.7）式，

$$P(1<X\leq 5)=\int_1^5 f(x)\mathrm{d}x=\int_1^3\frac{2}{9}(3x-x^2)\mathrm{d}x=\frac{20}{27}\ ,$$

$$P(X \leq 1) = 1 - P(1 < X \leq 5) = \frac{7}{27}.$$

后一个概率也可以通过积分计算，结果是一样的.

3．概率的几何意义

定积分的几何意义是面积，而概率表现为积分，所以概率（8.7）的几何意义是密度函数曲线与坐标横轴之间、介于两直线 $x = a$ 和 $x = b$ 范围内的面积，如图 8-1 所示．简言之：**概率值等于密度曲线下方的面积**.

令图 8-1 中的两直线分别向左、右无限伸展，就得到

$$P(X \leq b) = \int_{-\infty}^{b} f(x)\mathrm{d}x, \quad P(X > a) = \int_{a}^{+\infty} f(x)\mathrm{d}x,$$

$$P(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1,$$

其中前两式分别表示密度曲线下方，直线 $x = b$ 左侧的所有面积和直线 $x = a$ 右侧的所有面积，第三式正是密度的完备性公式，即整个密度曲线下方的面积等于 1．

概率的几何意义很重要，因为许多常见随机变量的密度表达式都很复杂，积分计算难度较大，往往需要通过面积的关系来理解概率原理、把握计算方法.

4．单点概率

在图 8-1 中，设想直线 $x = a$ 和 $x = b$ 无限靠近，最后变成一条线，如图 8-2 所示，由此可知，单点取值 $(X = x)$ 的概率等于密度曲线下方线段的面积．因为线段的宽度为零，所以单点概率为零，即

$$P(X = x) = 0.$$

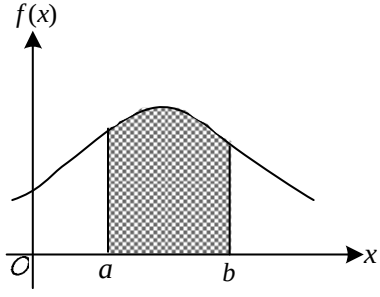

图 8-1　概率的几何意义

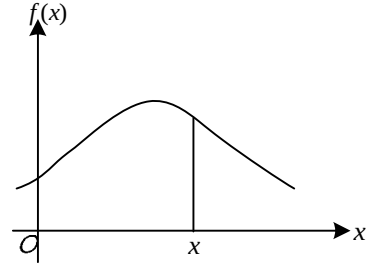

图 8-2　单点概率

这是连续型随机变量的重要特征，因此对于连续型随机变量来说，增减个别的点并不影响概率的计算，如

$$P(a \leq X < b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X \leq b)\ .$$

单点概率为零还透现出一个重要的概率原理：不可能事件的概率为零，但概率为零的事件未必是不可能事件.

利用无穷小的概念，可以更加深入地理解单点概率．在图 8-2 中，将曲线下方的线段看成具有无穷小的宽度 $\mathrm{d}x$，则单点概率为

$$P(X = x) = f(x)\mathrm{d}x\ , \tag{8.8}$$

即单点概率不再简单地等于 0，而是一个无穷小量，其内涵要丰富得多．(8.8) 式表明，密度函数值 $f(x)$ 本身并不是概率，但它表示各点概率之间的比例，密度的称呼因此而名副其实．单点概率虽然都等于 0，但它们作为无穷小可以进行比较，不同点处的密度曲线的起伏，正是反映了随机变量不同取值可能性的相对大小关系．

8.2.2　均匀分布

如果随机变量 X 的概率密度函数是

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \text{ 剟} x \quad b \\ 0, & \text{其他} \end{cases}\ , \tag{8.9}$$

则称 X 服从区间 $[a,b]$ 上的**均匀分布**，记为 $X \sim U(a,b)$．均匀分布的密度曲线如图 8-3 所示．

服从均匀分布的随机变量 X，其所有可能取值分布在区间 $[a,b]$ 上，而且在该区间上所有取值的可能性都相同（密度曲线无起伏），由此冠名为均匀分布．均匀分布的密度曲线下方的图形是一个矩形（参看图 8-3），完备性要求该矩形的面积为 1，所以矩形的高度等于矩形长度 $b-a$ 的倒数．

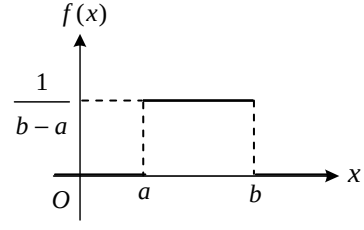

图 8-3　均匀分布密度曲线

均匀分布的概率总是等于图 8-3 中矩形的某一部分面积．由于矩形的高度恒定，所以计算均匀分布的概率时，只需考虑对应的矩形宽度是多少即可．比如当 $a < x < b$ 时，区间 $(-\infty, x]$ 在矩形中切割到的宽度是 $x-a$，因此

$$P(X \leq x) = \frac{x-a}{b-a}\ .$$

在估计误差时常常用到均匀分布．假设在数值计算中，数据都只保留到小数点后第 5 位，第 6 位以后的数字按四舍五入处理，则误差 X 是服从区间 $[-0.5\times10^{-5}, 0.5\times10^{-5}]$ 上均匀分布的随机变量．这样，对大量运算后产生的舍入误差进行分析就有章可循了．

8.2.3 指数分布

如果随机变量 X 的概率密度函数是

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0\\ 0, & x<0\end{cases}\quad(\lambda>0),\tag{8.10}$$

则称 X 服从参数为 λ 的**指数分布**，记为 $X\sim e(\lambda)$．这里的字母 e 通常被用来代表指数函数．指数分布的完备性不难通过积分予以验证，即函数（8.10）的积分为

$$\int_{-\infty}^{+\infty}f(x)dx=\int_{0}^{+\infty}\lambda e^{-\lambda x}dx=-e^{-\lambda x}\Big|_{0}^{+\infty}=1.$$

指数分布也来自于“排队现象”，往往与泊松分布紧密联系．在实际问题中，诸如动植物的寿命、随机服务系统中的服务时间等，都可用指数分布来描述.

8.3 分布函数与函数的分布

8.3.1 分布函数

设 X 是随机变量，x 是一个数，则概率 $P(X\leq x)$ 与 x 有关，随 x 的变化而变化，因而是 x 的函数．称 $F(x)=P(X\leq x)$ 为 X 的**分布函数**．与密度函数不同，分布函数 $F(x)$ 的取值本身就是一个概率值，$F(x)$ 的定义域为整个实数域．

分布函数 $F(x)$ 是在区间 $(-\infty,x]$ 上的“累积概率”，不要与单点概率相混淆．对于离散型随机变量，累积概率表现为求和，即

$$F(x)=\sum_{x_k\leq x}p_k,\tag{8.11}$$

其中 $p_k=P(X=x_k)$ 是 X 的分布律．对于连续型随机变量，累积概率表现为积分，即

$$F(x)=\int_{-\infty}^{x}f(x)dx,\tag{8.12}$$

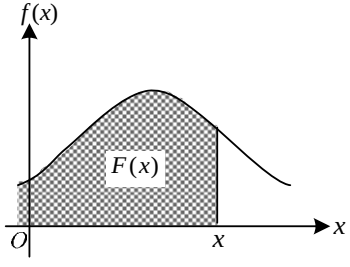

图 8-4 分布函数的几何意义

其中 $f(x)$ 是 X 的概率密度函数．利用积分上限函数的求导法则，可以得到

$$F'(x)=f(x),$$

在 $f(x)$ 的连续点成立．这表明连续型随机变量的分布函数和密度函数恰好构成一对原函数与导函数．

从（8.11）式可以看出，离散型随机变量的分布函数是具有跳跃性的阶梯函数，应用起来不甚方便，所以下面主要针对连续型随机变量研究分布函数．根据（8.12）式可知，$F(x)$ 的几何意义是密度曲线之下、x 左

方的所有面积，如图 8-4 所示．由此不难得到分布函数的如下性质：

（1）**有界性**　$0 \leq F(x) \leq 1$；

（2）**单调性**　$F(x)$ 是 x 的单调不减函数；

（3）**完备性**　$F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$，$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$；

（4）**互补性**　$P(X>x)=1-F(x)$．

此外，结合图 8-1 和图 8-4 不难发现

$$P(a<X\leq b)=F(b)-F(a)\ . \tag{8.13}$$

以上性质对离散型随机变量也适用．公式（8.13）可以用来计算概率，不需要积分，比用公式（8.7）方便得多．可惜的是在多数情况下，分布函数不易求得，即使求得也不是简单的初等形式，所以处理概率问题还是多以密度函数为主．

例 8.13　设 $X\sim U(a,b)$（均匀分布），求分布函数 $F(x)$．

解　可以按照公式（8.12）计算积分，但在上节中已提到：对于均匀分布来说，求概率不必做积分，只需求相应的矩形面积即可．当 $x<a$ 时，区间 $(-\infty,x]$ 未切割到矩形面积，当 $x\geq b$ 时，切割到整个矩形（参看图 8-3），因此 $F(x)$ 的值（面积）分别是 0 和 1；而当 $a\leq x<b$ 时，切割到宽度为 $x-a$ 的矩形．综合起来得到

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leq x<b\\ 1, & x\geq b\end{cases}\ .$$

均匀分布的分布函数 $F(x)$的图形如图 8-5 所示，$F(x)$ 的图形连续，尖点处无导数，恰为密度函数 $f(x)$ 的间断点（参看图 8-3）．

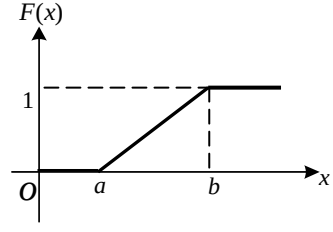

图 8-5　均匀分布函

8.3.2　函数的分布

设 X 是随机变量，则 X 的函数 $Y=g(X)$ 也是随机变量．现在已知 X 的分布，希望由此求出函数 $Y=g(X)$ 的分布．

当 X 是离散型随机变量时，求函数的分布比较容易．

例 8.14　已知 X 的分布律为 $\begin{pmatrix}-1 & 0 & 1 & 2 & 5\\ 0.1 & 0.2 & 0.3 & 0.1 & 0.3\end{pmatrix}$，求 $Y=X^2$ 的分布律．

解　$Y=X^2$ 的可能取值显然是 $(-1)^2,0^2,1^2,2^2,5^2$．由于事件 $(Y=4)$ 和事件 $(X=2)$ 相同，所以概率也相等，其他的概率情况相类似，于是 Y 的分布律为

$$Y \sim \begin{pmatrix} 1 & 0 & 1 & 4 & 25 \\ 0.1 & 0.2 & 0.3 & 0.1 & 0.3 \end{pmatrix}.$$

将取值相同的概率合并得到

$$Y \sim \begin{pmatrix} 0 & 1 & 4 & 25 \\ 0.2 & 0.4 & 0.1 & 0.3 \end{pmatrix}.$$

一般地，若 X 的分布律为 $P(X=x_k)=p_k$（$k=1,2,\mathrm{L}$），则 $Y=g(X)$ 的所有可能取值为 $y_k=g(x_k)$（$k=1,2,\mathrm{L}$），并且概率值不变，即 Y 的分布律为 $P(Y=y_k)=p_k$（$k=1,2,\mathrm{L}$）.

当 X 是连续型随机变量时，求函数的分布往往比较复杂．设 X 的概率密度是 $f_X(x)$，求 $Y=g(X)$ 的概率密度 $f_Y(y)$，需要先求出 Y 的分布函数 $F_Y(y)=P(Y\le y)$，然后求导得到 $f_Y(y)=F_Y'(y)$．在求概率 $P(Y\le y)$ 时，要通过不等式变形或者积分换元计算积分，有时还要分各种情况进行讨论，总之颇费周折，此处从略．不过应该强调的是，随机变量函数的分布，其数学表达式有一定的复杂性，绝不能简单地、想当然地随意描画．

8.4 正 态 分 布

8.4.1 正态分布的定义与性质

1．正态分布的定义

如果随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\sigma>0), \tag{8.14}$$

则称 X 服从参数为 μ 和 σ^2 的正态分布，记为 $X\sim N(\mu,\sigma^2)$．

正态分布是最重要的分布．一方面，在自然界中取值受众多微小独立因素综合影响的随机变量，一般都服从正态分布，如测量误差、质量指标、身高体重、考试成绩、用电数量、炮弹落点等．可见随机变量服从正态分布是非常普遍的现象；另一方面，许多其他分布又可以用正态分布来近似或导出．无论在理论上还是在实践中，正态分布都有着极其广泛的应用．

2．正态曲线的性质

正态密度函数（8.14）的图像是钟形的曲线，如图 8-6 所示，称为**正态曲线**，又称**高斯**（Gauss）**曲线**．正态曲线有如下性质：

（1）正态曲线关于直线 $x=\mu$ 对称；

（2）$x=\mu$ 处正态曲线达到最高点 $f(\mu)=(1/\sqrt{2\pi}\sigma)$，两侧逐渐降低，并以 x 轴为渐近线；

（3）在对应于 $x=\mu\pm\sigma$ 处，正态曲线有两个拐点；

（4）整个正态曲线下方的面积为 1，即

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t=1,$$

式中 $x=\sigma t+\mu$．这个积分称为**概率积分**，又称**高斯积分**．高斯积分的计算过程要用到较多的微积分知识，此处从略．不过高斯积分的结果作为概率常识应当牢记．

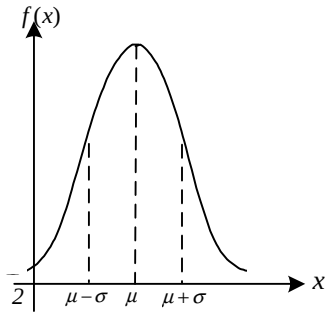

图 8-6　正态曲线

（5）若 σ 固定，μ 变动，则曲线会向左或向右平移，几何形状不变，即 μ 是曲线的位置参数；若 μ 固定，σ 变动，则因面积恒定为 1 而最高点和拐点位置发生变化，故 σ 越大（小），曲线越平坦（陡峭），即 σ 是曲线的形状参数．

3．标准正态分布

当 $\mu=0,\sigma=1$ 时，称 X 服从**标准正态分布**，记为 $X\sim N(0,1)$．标准正态分布的概率密度是

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\quad .$$

本书将 $\varphi(x)$ 作为专用记号，以后不再说明．标准正态曲线的对称性、最高点、拐点、渐近线、面积（积分）情况与一般正态分布类似，不再赘述．

8.4.2　正态分布的概率计算

1．标准正态分布的概率计算

标准正态分布的分布函数为

$$\Phi(x)=\int_{-\infty}^{x}\varphi(t)\mathrm{d}t=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}t,\qquad(8.15)$$

这里 $\Phi(x)$ 也作为专用记号，以后不再说明．积分（8.15）表示标准正态曲线之下、x 左方的面积，如图 8-7 所示．

计算积分（8.15）之值的工作量很大，为方便起见，人们编制了 $\Phi(x)$ 的数值表，称为**标准正态分布表**（见附表 1），这样，计算概率只要查表就行了．不过表中只有 $x\geq 0$ 的数值，若遇到负值，则要利用图形的对称性和完备性，按公式 $\Phi(-x)=1-\Phi(x)$ 进行非负化换算．

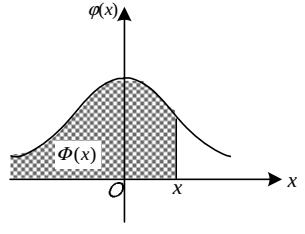

图 8-7　标准正态分布函数

例 8.15 设 $X \sim N(0,1)$，求以下概率：

（1） $P(1<X<2)$；　　（2） $P(-1<X\leq 2)$；

（3） $P(|X|<1.2)$；　　（4） $P(|X|\geq 1.8)$；

（5） $P(X<-1.3)$；　　（6）已知 $P(|X|>a)=0.05$．求 a 的值．

解 （1）利用公式（8.13）并查附表 1，得

$$P(1<X<2)=\Phi(2)-\Phi(1)=0.9772-0.8413=0.1359;$$

（2）遇到负值，要作非负化换算：

$$P(-1<X\leq 2)=\Phi(2)-\Phi(-1)=\Phi(2)-[1-\Phi(1)]$$
$$=0.9772+0.8413-1=0.8185;$$

（3）含绝对值的不等式要作等价变形：

$$P(|X|<1.2)=P(-1.2<X<1.2)=\Phi(1.2)-[1-\Phi(1.2)]$$
$$=2\times 0.8849-1=0.7698;$$

（4）利用对称性及互补性，

$$P(|X|\geq 1.8)=2P(X\geq 1.8)=2[1-\Phi(1.8)]=2(1-0.9641)=0.0718;$$

（5）这是直接的分布函数值，作非负化换算即可：

$$P(X<-1.3)=\Phi(-1.3)=1-\Phi(1.3)=1-0.9032=0.0968;$$

（6）利用对称性得到 $P(X>a)=0.025$，再从互补性求得 $\Phi(a)=0.975$，由此倒查附表 1 得 $a=1.96$．

2．一般正态分布的概率计算

对于非标准正态分布，可通过线性变换化为标准正态分布来处理．设 $X\sim N(\mu,\sigma^2)$，则通过积分换元不难证明有以下重要结论：

$$Y=\frac{X-\mu}{\sigma}\sim N(0,1)\text{．}\qquad (8.16)$$

这个结论叫做正态分布的**标准化**．对不等式作等价变形并根据结论（8.16）可以得到

$$P(X<x)=P\left(\frac{X-\mu}{\sigma}<\frac{x-\mu}{\sigma}\right)=P\left(Y<\frac{x-\mu}{\sigma}\right)=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

这说明非标准正态分布的概率计算，仍然可以直接利用标准正态分布的分布函数 $\Phi(x)$，只是在填写自变量时，先要进行标准化变换．

例 8.16 设 $X\sim N(1,4)$，求以下概率：

（1） $P(X<3)$；　　（2） $P(0<X<1.6)$；

（3） $P(X>1.7)$；　　（4） $P(|X|>1)$.

解　本例中 $\mu=1$，$\sigma=2$，于是

（1） $P(X<3)=\Phi\left(\dfrac{3-1}{2}\right)=\Phi(1)=0.8413$；

（2） $P(0<X<1.6)=\Phi\left(\dfrac{1.6-1}{2}\right)-\Phi\left(\dfrac{0-1}{2}\right)=\Phi(0.3)-\Phi(-0.5)$

$=\Phi(0.3)-[1-\Phi(0.5)]=0.6179+0.6915-1=0.3085$；

（3） $P(X>1.7)=1-\Phi\left(\dfrac{1.7-1}{2}\right)=1-\Phi(0.35)=1-0.6368=0.3632$；

（4）这里的绝对值不适用对称性，

$$P(|X|>1)=P(X>1)+P(X<-1)=\left[1-\Phi\left(\frac{1-1}{2}\right)\right]+\Phi\left(\frac{-1-1}{2}\right)$$

$$=1-\Phi(0)+[1-\Phi(1)]=0.5+1-0.8413=0.6587.$$

例 8.17　设 $X\sim N(\mu,\sigma^2)$，求概率

$P(|X-\mu|<\sigma)$，$P(|X-\mu|<2\sigma)$，$P(|X-\mu|<3\sigma)$.

解　注意到结论（8.16）及标准正态分布的对称性，可得

$$P(|X-\mu|<\sigma)=P\left(\left|\frac{X-\mu}{\sigma}\right|<1\right)=2\Phi(1)-1=0.6826;$$

$$P(|X-\mu|<2\sigma)=2\Phi(2)-1=0.9544;$$

$$P(|X-\mu|<3\sigma)=2\Phi(3)-1=0.9974.$$

计算表明，正态变量有 99.74%的机会落入以 3σ 为半径的区间 $(\mu-3\sigma,\mu+3\sigma)$ 内，因此可以说，正态变量几乎都在该区间内取值．这就是统计工作者经常使用的“3σ”规则，尤其在统计实践的快速分析中应用频繁．

习 题 八

1．将一枚硬币连续抛两次，以 X 表示所抛两次中出现正面的次数，试写出随机变量 X 的分布律．

2．若 X 服从二点分布，且 $P(X=1)=2P(X=0)$，求 X 的分布律．

3．设随机变量 X 分布律为 $P(X=k)=\dfrac{2A}{n}$（$k=1,2,\cdots,n$），试确定常数 A.

4．在 8 根灯管中混有 2 根坏灯管，现从中任取 3 根灯管，X 为取得的好灯管数，试在下列两种情况下求 X 的分布律．

（1）无放回地取 3 根灯管；　　（2）有放回地每次取一根，共 3 次取 3 根灯管．

5．某射手有 5 发子弹，射一次命中的概率为 0.8，如果命中了就停止射击，如果不命中就一直射到子弹用尽，求射击次数 X 的概率分布．

6．轮胎厂检验产品质量的方法是对每批轮胎随机抽取 10 个逐个检验，若次品在一个或一个以内就检验通过．若某批轮胎的次品率为 0.2，则该批轮胎验收通过的概率有多大?

7．根据经验数据，自学考试某课程的通过率为 60%．若 10 个人相互独立地参加考试，有 4 人通过的概率是多少？有 6 人或 6 人以上通过的概率是多少？

8．某车间有 12 台相同的车床，每台车床由于工艺上的原因，时常需要停车，设每台车床停车（或开车）是相互独立的，每台车床在任一时刻处于停车状态的概率为 1/3．计算在任一指定时刻，车间里恰有 2 台车床处于停车状态的概率．

9．电视机厂要求该厂产品次品率不超过 0.1%，在每日生产线完成的电视机中随机抽选检查，发现次品立即停产．如果该企业希望当次品率为 0.1%时停产的概率不超过 0.05，则每天应检查多少台电视机?

10．某航线的航班常常有旅客预定票后又临时取消，每班平均为 4 人．若预定票而取消的人数服从以平均人数为参数的泊松分布，求

（1）正好有 4 人取消的概率；　　（2）不超过 3 人（含 3 人）取消的概率；

（3）超过 6 人（含 6 人）取消的概率；　　（4）无人取消的概率．

11．某商店出售某种高档商品，根据以往经验，每月销售量 X 服从 $\lambda=3$ 的泊松分布，问在月初进货时要库存此商品多少件，才能以不低于 99%的概率满足顾客的需要？

12．一台仪表以 0.2 为一个刻度，读数时选取与指针靠近的刻度值．实际测量值是均匀分布的，求实际测量值（指针值）与读数之间的偏差

（1）小于 0.04 的概率；　　（2）大于 0.05 的概率．

13．某城市每天用电量不超过 100 万 kW·h，以 X 表示每天的耗电率（即用电量除以 100 万 kW·h 所得之商），它的概率密度函数为 $f(x)=\begin{cases}ax(1-x)^2, & 0<x<1\\ 0, & \text{其他}\end{cases}$．

（1）求 a 的值；

（2）若该城市发电厂每天供电量为 80 万 kW·h，求供电不能满足需要（即耗电率大于 0.8）的概率．

14．某种型号的电子管寿命 X (h)作为随机变量，其概率密度函数为 $f(x)=\begin{cases}\dfrac{100}{x^2}, & x\ldots100\\ 0, & \text{其他}\end{cases}$．

（1）求 X 的分布函数并作出其图形；

（2）若一无线电器材配有 3 个这样的电子管，试计算该器材使用 150h 内不需要更换电子管的概率．

15．设 $X\sim N(\mu,\sigma^2)$，X 的概率密度函数为 $f(x)=a\mathrm{e}^{-4(\quad)}$，试确定 a,b,μ,σ 的值．

16．设 $X\sim N(0,1)$，求概率：

$P(X<2.2)$，$P(X>1.76)$，$P(X<-0.76)$，$P(|X|<1.55)$，$P(|X|>2.5)$．

17．设 $X\sim N(3,4)$．

（1）求概率 $P(2<X\leq 5)$，$P(X>-1.5)$，$P(X<2)$，$P(|X|<1)$，$P(|X|>0.5)$；

（2）确定 c 使 $P(X>c)=P(X\leq c)$．

18．设成年男子身高 X （cm）$\sim N(170,36)$，某种公交车车门的高度是按成年男子碰头的概率在 1%以下来设计的，问车门的高度最少应为多少?

19．某班一次数学考试成绩 $X \sim N(70,100)$，若规定低于 60 分为“不及格”，高于 85 分为“优秀”，问该班级：

（1）数学成绩“优秀”的学生占总人数的百分之几？

（2）数学成绩“不及格”的学生占总人数的百分之几？

20．单项选择题

（1）随机变量的取值总是　（　）

A．正的数；　B．整数；　C．有限个数；　D．实数．

（2）下面哪一个符合概率分布的要求．　（　）

A．$P(X=x)=\dfrac{x}{6}$（$x=1,2,3$）；　B．$P(X=x)=\dfrac{x}{4}$（$x=-1,2,3$）；

C．$P(X=x)=\dfrac{x}{10}$（$x=1,2,3$）；　D．$P(X=x)=\dfrac{x^2}{15}$（$x=-1,2,3$）．

（3）离散型随机变量 X 的分布律为 $P(X=k)=ak$（$k=1,2,3,4$），则 $a=$　（　）

A．0.05；　B．0.1；　C．0.2；　D．0.25．

（4）随机猜测“选择题”的答案，每道题猜对的概率为 0.25，则 4 道选择题相互独立地猜对 2 道及 2 道以上的概率约为　（　）

A．0.1；　B．0.3；　C．0.5；　D．0.7．

（5）某厂生产的零件合格率约为 99%，零件出厂时每 200 个装一盒，设每盒中的不合格数为 X，则 X 通常服从　（　）

A．正态分布；　B．均匀分布；　C．泊松分布；　D．二项分布．

（6）已知 n 个随机变量 X_i（$i=1,2,\mathrm{L},n$）相互独立且都服从参数为 a 的两点分布，则 $X=X_1+X_2+\mathrm{L}+X_n$ 服从　（　）

A．两点分布；　B．二项分布；　C．泊松分布；　D．正态分布．

（7）设连续型随机变量 X 的分布函数是 $F(x)$，密度函数是 $f(x)$，则 $P(X=x)=$　（　）

A．$F(x)$；　B．$f(x)$；　C．0；　D．以上都不对．

（8）设连续型随机变量 X 的分布函数是 $F(x)$，密度函数是 $f(x)$，则对于一个固定的 x，下列说法正确的是　（　）

A．$f(x)$ 不是概率值，$F(x)$ 是概率值；　B．$f(x)$ 是概率值，$F(x)$ 不是概率值；

C．$f(x)$ 和 $F(x)$ 都是概率值；　D．$f(x)$ 和 $F(x)$ 都不是概率值．

（9）设 $X \sim N(10,25)$，则 $P(X<5)$ 和 $P(X>20)$ 的概率分别为　（　）

A．0.0228，0.1587；　B．0.3413，0.4772；　C．0.1587，0.0228；　D．0.4772，0.3413．

（10）随机变量 X 服从区间 $[a,b]$ 上的均匀分布是指　（　）

A．X 的取值是个常数；

B．X 取区间 $[a,b]$ 上任何值的概率都等于同一个正的常数；

C．X 落在区间 $[a,b]$ 的任何子区间内的概率都相同；

D．X 落在区间 $[a,b]$ 的任何子区间内的概率都与子区间的长度成正比．