线性代数

邓辉文　编

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

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

清华大学出版社

线性代数

邓辉文　编

清华大学出版社

北京

内容简介

本书以线性方程组为主线、以矩阵和向量为工具，阐述线性代数的基本概念、基本理论和方法，使全书内容联系紧密，具有较强的逻辑性．全书共分5章，分别介绍线性方程组、矩阵代数、向量代数、特征值和特征向量以及二次型．对每章的学习内容简述其起源和作用．

由于线性代数概念多、结论多，内容较抽象，本书尽量从简单实例入手，力求通俗易懂、由浅入深，对重点内容提供较多的典型例题，以帮助学生更好地理解、掌握和运用线性代数的知识．每章有精选习题，有些选自历年的研究生入学考试题目，书后有习题答案．专业术语均有对应的英文．本书简单介绍了使用MATLAB求解线性代数问题的一些常见命令，希望能引起大家的学习兴趣，较早进入MATLAB世界．

本书适合于普通高等院校非数学专业各类理工科本科生特别是计算机各专业、电子信息及有关各专业、自动化专业、经济和管理学科等专业学生作为教学用书．

本书有配套的《线性代数学习指导与习题解答》辅助用书，同时由清华大学出版社出版，本书电子教案可在清华大学出版社网站下载．

图书在版编目（CIP）数据

线性代数/邓辉文编．—北京：清华大学出版社，2008．7
ISBN 978-7-302-17760-9

Ⅰ．线…　Ⅱ．邓…　Ⅲ．线性代数　Ⅳ．O151．2

中国版本图书馆CIP数据核字（2008）第077790号

责任编辑：刘　颖
责任校对：赵丽敏
责任印制：何　芊
出版发行：清华大学出版社　　**地　　址**：北京清华大学学研大厦A座
http://www.tup.com.cn　　**邮　　编**：100084
社　总　机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 刷 者：北京四季青印刷厂
装 订 者：北京国马印刷厂
经　　销：全国新华书店
开　　本：185×230　**印　张**：13　**字　数**：279千字
版　　次：2008年7月第1版　　**印　　次**：2008年7月第1次印刷
印　　数：1～4000
定　　价：19.80元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。联系电话：（010）62770177转3103　　产品编号：028937－01

前　言

什么是代数？代数(algebra)最早就是求解方程或方程组，在清代传入我国，当时将Algebra翻译成“阿尔热巴拉”，直到1859年才翻译成“代数”.根据现代数学的观点，代数就是在所考虑的对象之间规定一些运算后得到的数学结构.

什么是线性代数？线性代数(linear algebra)涉及的运算主要是称为加减和数乘的线性运算，这些线性运算须满足一定的性质进而构成线性空间.线性代数需要解决的第一个问题就是求解来源于实际应用问题的线性方程组.

线性代数的研究对象是什么？线性代数的研究对象是**线性空间**，包括其上的线性变换.它与高等代数、近世代数的研究对象略有所不同.

从广义的角度看，线性代数研究线性科学中的“线性问题”.直观地讲，对所考虑的变量来讲，和式中各项次数最高为一次的问题就是线性问题.即使是大量出现的非线性问题有时也可以转换成线性问题进行处理，如在一定条件下，曲线可用直线近似，曲面可用平面近似，函数增量可用函数的微分近似.

矩阵和向量是重要的代数工具.线性问题的讨论往往涉及矩阵和向量，它们是重要的代数工具.在一定的意义上，它们以及其上的一些运算本身就构成线性空间.因此，线性代数的主要内容分别是线性方程组、向量空间、矩阵代数，以及与线性变换密切相关的方阵的特征值和二次型这种线性空间之间特殊的双线性函数等.

线性代数的特点是什么？内容较抽象、概念和定理较多，前后联系紧密，环环相扣，相互渗透.

为何要学习线性代数？线性代数是一种数学建模方法，科研工作者必须掌握，虽然其有关内容具有一定的抽象性.前面已经提到，线性化是重要的数学方法，在高等数学特别是优化问题的讨论中会用到.在计算机程序设计语言特别是MATLAB中，矩阵是最基本的数据结构.在微积分(高等数学)、微分方程、离散数学、算法分析与设计、计算机图形图像处理及数字信号处理等课程中，矩阵、向量、线性变换是经常要用的知识.随着计算机的普及，线性代数在理论和实际应用中的重要性更加突出，这使得诸如计算机专业、电子信息专业、自动控制专业以及经济管理专业等对线性代数的内容从深度和广度方面都提出了更高的要求.

学习线性代数要达到的目的.通过线性代数的学习，一方面可以进一步培养抽象思维能力和严密的逻辑推理能力，为进一步学习和研究打下坚实的理论基础，另一方面为立志报考研究生的同学提供必要的线性代数理论知识、解题技巧和方法.

本书适用对象.本书是根据作者多年的教学经验编写的，同时也参考了国内外的线性代数教材.所选内容适合于普通高等院校非数学专业各类理工科本科学生，特别是计算机各专

业、电子信息及相关各专业、自动化专业、经济和管理学科等专业本科学生作为教学用书，也可作为理工科考研学生和有关工作者的参考书.

本书主要内容. 全书共分5章，分别介绍线性方程组、矩阵代数、向量代数、特征值与特征向量和二次型. 全书以线性方程组为主线、以矩阵和向量为工具阐述线性代数的基本概念、基本理论和方法，使全书内容联系紧密，具有较强的逻辑性. 由于线性代数概念多、结论多，内容较抽象，本书尽量从简单实例入手，力求通俗易懂、由浅入深，对重点内容提供较多的典型例题，以帮助学生更好地理解、掌握和运用线性代数的知识. 每章都有精选习题，有些选自历年的研究生入学考试线性代数题目，书后有习题答案.

MATLAB程序设计语言. 计算机科学的研究和发展，给线性代数内容注入了新的活力，出现了各种各样的数学软件，如MATLAB、Mathematic等. 本书介绍了使用MATLAB求解线性代数问题的一些常见命令，希望能引起大家的学习兴趣，较早进入MATLAB世界. 因为MATLAB强大的数值计算和符号计算功能、卓越的数据可视化能力和适用于各行各业的不同的工具箱（Toolbox），使得MATLAB成为多学科多种工作平台的程序设计语言，在欧美的几乎所有高校中，MATLAB已经成为线性代数、概率论与数理统计、自动控制理论、数字信号处理、动态系统仿真等课程的基本教学工具，是攻读学位的大学生、硕士生和博士生必须掌握的基本技能.

本书讲授约需54课时，根据教学课时数以及学生具体情况，对于第2章、第3章和第5章内容，特别是个别难度较大的例题，进行适当删减，可作为专科学生、网络学院学生、成教学生的教材. 在学习过程中，若能结合与本书配套的教学辅助用书《线性代数学习指导与习题解答》进行学习，则能起到举一反三、加深对课本内容理解的作用.

由于编者水平有限，缺点和疏漏在所难免，肯请大家不吝指正，万分感激.

编　者
2008年5月

目　录

第 1 章　线性方程组

众所周知，代数(algebra)最早就是求解(线性、非线性)方程或方程组.

历史上，线性代数(linear algebra)遇到的第一个问题是求解线性方程组，在西方它是在 17 世纪后期由 G. W. Leibniz(1646—1716)开创的.

线性方程组是线性代数的基本内容，是贯穿线性代数的一条主线. 线性代数各部分内容或多或少与线性方程组有关，如矩阵方程、向量组的线性相关性、特征值及特征向量和二次型的标准化等.

线性方程组又称为一次方程组，大家在中学就学习过较简单的二元一次方程组和三元一次方程组. 在自然科学、管理科学和工程应用中，如石油探测、管理决策、电路分析和减肥食谱研究等大量科学技术问题中，往往需要求解一个有成百上千甚至更多个未知量的大型线性方程组，求解线性方程组的重要性是显而易见的. 因此，从这个角度去看，也有必要讨论一般的线性方程组的求解问题.

一般的线性方程组的求解问题的讨论离不开矩阵. 实际上，矩阵是重要的数学工具之一，在很多的实际应用和理论研究中都经常会用到.

1.1　线性方程组与矩阵的有关概念

1.1.1　线性方程组的有关概念

在中学学习过诸如以下较简单的二元一次方程组和三元一次方程组：

$$\begin{cases} x+y=1, \\ 2x-y=5; \end{cases} \qquad \begin{cases} x+2y-5z=19, \\ 2x+8y+3z=-22, \\ x+3y+2z=-11. \end{cases}$$

它们都是本章要讨论的线性方程组.

再看一个例子.

例 1.1　求经过四个点$(-1,2)$，$(0,0.25)$，$(1,1)$，$(2,-1)$的三次多项式 ax^3+bx^2+cx+d 就需要求解以下的线性方程组：

$$\begin{cases} a\cdot(-1)^3+b\cdot(-1)^2+c\cdot(-1)+d=2, \\ a\cdot 0^3+b\cdot 0^2+c\cdot 0+d=0.25, \\ a\cdot 1^3+b\cdot 1^2+c\cdot 1+d=1, \\ a\cdot 2^3+b\cdot 2^2+c\cdot 2+d=-1, \end{cases} \text{即} \begin{cases} -a+b-c+d=2, \\ d=0.25, \\ a+b+c+d=1, \\ 8a+4b+2c+d=-1. \end{cases}$$

其中 a,b,c 和 d 看作未知量.

对所考虑的未知量来说,和式中每项次数最高为一次的方程称为**线性方程**(linear equation),否则称为**非线性方程**(nonlinear equation).

对未知量 x,y,z 来说,如 $2x+3y+4z=5$ 是线性方程,而 $3xy+y-2z=5$,$e^y+xy-e=0$ 和 $3x^2+\sin y=2z$ 是非线性方程.若 $a_{i1},a_{i2},\cdots,a_{in},b_i(i=1,2,\cdots,m)$是常数,对未知量 $x_1,x_2,\cdots,x_n$ 来说,$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$ 是线性方程.

在微积分中,对于未知函数 $y(x)$以及未知函数 $y(x)$的导数来说,和式中每项次数最高为一次的微分方程称为**线性微分方程**,如 $y''+p(x)y'+Q(x)y=f(x)$,而 $3y'y=f(x)$是非线性微分方程.

每个方程均是线性方程的方程组称为**线性方程组**(system of linear equations).具有 n 个未知量的线性方程组称为 ***n* 元线性方程组**(system of linear equations with n unknowns).

n 元线性方程组的一般形式为

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2, \\ \qquad\vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m. \end{cases} \tag{1.1}$$

(1.1)式是有 m 个方程的 n 元线性方程组,可称为 ***m*×*n* 线性方程组**,其中 a_{ij} 称为第 i 个方程第 j 个未知量 x_j 的**系数**(coefficient)($i=1,2,\cdots,m,j=1,2,\cdots,n$),$b_i$ 称为第 i 个方程的**常数项**(constant)($i=1,2,\cdots,m$),m 和 n 是任意正整数,其关系可能为下列三种情况之一:$m=n,m>n$ 及 $m<n$.

对于 n 元线性方程组,应该讨论下述问题:

(1) 解的存在性.

(2) 求出其所有解,包括解的个数.

在讨论 n 元线性方程组的有关问题时,矩阵是一个很方便的工具.

1.1.2 矩阵的有关概念

1. 矩阵

为了引入矩阵概念,先看一个三阶幻方的例子.

例 1.2(三阶幻方问题) 将 1 到 9 共 9 个正整数按一定顺序填入如下表格的空里,使其每行的 3 个数相加、每列的 3 个数相加、表格所构成的正方形的两条对角线上的 3 个数相加之和均相同.

解 容易知道，题目中所提到的 3 个数相加之和为 15，于是得到其中的一种填法为

6	1	8
7	5	3
2	9	4

将得到的数表

$$\begin{matrix} 6 & 1 & 8 \\ 7 & 5 & 3 \\ 2 & 9 & 4 \end{matrix}$$

称为矩阵，记为

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

类似地，可以考虑 $n(n\geqslant 4)$ 阶幻方问题.

矩阵就是由一些数，也可以是一些表示数的符号，按一定顺序排成若干行和若干列而构成的一个表格.

定义 1.1 由 mn 个数按一定顺序构成的有 m 行及 n 列的数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1.2}$$

称为 $\boldsymbol{m\times n}$ **矩阵**(matrix of size $m\times n$).

已经看出，为了表示元素的排列顺序，通常用一对圆括号( )或方括号[ ]将它们括起来，但不能使用{ }或| |等符号.

通常用大写的黑斜体(英文或希腊)字母来表示矩阵，如 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_3$.

在矩阵(1.2)中，对于 $i=1,2,\cdots,m$，$a_{i1},a_{i2},\cdots,a_{in}$ 称为**第 $\boldsymbol{i}$ 行元素**(i-th row)，对于 $j=1,2,\cdots,n$，$a_{1j},a_{2j},\cdots,a_{mj}$ 称为**第 $\boldsymbol{j}$ 列元素**(j-th column). 用 a_{ij} 表示第 i 行第 j 列位置的元素，称为 (i,j) 位置元素. (i,j) 位置元素 a_{ij} 是用双下标表示的，第一个下标表示该元素所在的行，第二个下标表示该元素所在的列，这种表示方法本身就有一定的创意.

例如，分别将线性方程组(1.1)的第 1 个方程到第 m 个方程中未知量 $x_1,x_2,\cdots,x_n$ 的

系数依次抽取出来，如根据第 i 个方程就得到 $a_{i1}, a_{i2}, \cdots, a_{in}(i=1,2,\cdots,m)$. 特别地，若在第 i 个方程中不含未知量 x_j，则表明 $a_{ij}=0$. 再将它们按第1个方程到第 m 个方程的顺序分别排成第1行，第2行，……，直到第 m 行，得到一个数表，这个仅由线性方程组(1.1)各方程未知量的 mn 个系数构成的矩阵

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

称为线性方程组(1.1)的**系数矩阵**(coefficient matrix)，常用黑斜体字母 $\boldsymbol{A}$ 表示.

又如，分别将线性方程组(1.1)的第1个方程到第 m 个方程中未知量 $x_1, x_2, \cdots, x_n$ 的系数以及右边的常数项依次抽取出来，如根据第 i 个方程就得到 $a_{i1}, a_{i2}, \cdots, a_{in}, b_i(i=1, 2,\cdots,m)$. 再将它们按第1个方程到第 m 个方程的顺序分别排成第1行，第2行，……，直到第 m 行，得到一个由 $m(n+1)$ 个数构成的矩阵

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

在以后的讨论中，通常用黑斜体字母 $\boldsymbol{B}$ 表示，称为线性方程组(1.1)的**增广矩阵**(augmented matrix).

事实上，增广矩阵是最早出现的矩阵."矩阵"一词来源于拉丁语，表示排数的意思，1850年英格兰数学家 J. J. Sylvester(1814—1897)首次使用矩阵术语.

在1.2节将会看到，对线性方程组进行同解变换，关键是对未知量 $x_1, x_2, \cdots, x_n$ 的系数和方程右边的常数项进行变换，即对其增广矩阵进行相应的初等行变换，因而可以略去未知量 $x_1, x_2, \cdots, x_n$，这样做不但简单而且不易出错.

说明 在以后的讨论中，将系数矩阵用 $\boldsymbol{A}$ 表示，增广矩阵用 $\boldsymbol{B}$ 表示.

例1.3 写出下列线性方程组的系数矩阵和增广矩阵：

$$\begin{cases} x_1 + 2x_2 = 3, \\ 4x_1 + 7x_2 + x_3 = 10, \\ x_2 - x_3 = 2, \\ 2x_1 + 3x_2 + x_3 = 4. \end{cases}$$

解 系数矩阵和增广矩阵分别为

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 7 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & 1 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 1 & 10 \\ 0 & 1 & -1 & 2 \\ 2 & 3 & 1 & 4 \end{pmatrix}.$$

注意 在写线性方程组的系数矩阵和增广矩阵时，一方面要按一定顺序，如 x_1,x_2,x_3 或 x,y,z 得出未知量的系数，另一方面，若有缺位，如例 1.3 中的第 1 个方程没有 x_3，就认为 x_3 的系数为 0.

为了方便，可以将矩阵(1.2)写成 $(a_{ij})_{m\times n}$，其中 a_{ij} 是矩阵的代表元素. 当仅用一个字母表示矩阵时，可以在该字母的右下角写上 $m\times n$，以表明矩阵的行数和列数，如 $\boldsymbol{A}_{m\times n}$.

矩阵 $\boldsymbol{A}_{m\times n}$ 中，$m\times n$ 为矩阵 $\boldsymbol{A}$ 的**型**(size)，这里 $m\times n$ 没有乘积之意. 有 m 行 n 列的矩阵 $\boldsymbol{A}$，就是"$m\times n$ 矩阵 $\boldsymbol{A}$"，读作"m 行 n 列矩阵 $\boldsymbol{A}$". 在 $m\times n$ 中，行数 m 写在"×"的前面，而列数 n 写在"×"的后面.

两个矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是同型的是指其行数、列数分别相同.

矩阵中的元素一般取自一些数组成的集合，在该集合上有一些满足特定性质的运算.

若矩阵中的元素全取自复数集合 $\mathbb{C}$，这样的矩阵称为**复矩阵**(complex matrix)，如 $\begin{pmatrix} 1 & 2 & \mathrm{i} \\ 1-\mathrm{i} & -1 & 0 \\ 3 & \mathrm{i} & 1+\mathrm{i} \end{pmatrix}$，其中 $\mathrm{i}=\sqrt{-1}$ 是虚数单位. 若矩阵中的元素全取自实数集合 $\mathbb{R}$，这样的矩阵称为**实矩阵**(real matrix). 若矩阵中的元素全取自整数集合，这种矩阵称为**整数矩阵**.

约定，若没有特殊说明，今后所讨论的矩阵为实矩阵.

一般的 $m\times n$ 矩阵，其元素排列起来像一个矩形，它是一个矩形阵列. 为了方便，将只有一行的矩阵 $(a_1,a_2,\cdots,a_n)$ 称为**行矩阵**，只有一列的矩阵 $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ 称为**列矩阵**.

由于 $n\times n$ 矩阵的元素排列起来像一个正方形，因此 $n\times n$ 矩阵常称为**方阵**(square)，为了将方阵的行数或列数表示出来，$n\times n$ 矩阵又称为 ***n* 阶矩阵**(matrix of order n)或 ***n* 阶方阵**(square matrix of order n). n 阶方阵 $\boldsymbol{A}$ 也可以表示为 $\boldsymbol{A}_n$ 或 $\boldsymbol{A}_{n\times n}$. 由一个数 a 构成的一阶方阵 (a) 就是元素 a 本身，即 $(a)=a$.

在 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 中，称元素 $a_{11},a_{22},\cdots,a_{nn}$ 为方阵 $\boldsymbol{A}$ 的**主对角线**(principal diagonal)，简称为 $\boldsymbol{A}$ 的**对角线**(diagonal). 有时称元素 $a_{1n},a_{2,n-1},\cdots,a_{n1}$ 为方阵 $\boldsymbol{A}$ 的**次对角线**(secondary diagonal).

实际上，矩阵是一种非常重要的数学工具之一，在处理很多问题时都会用到矩阵. 下面再举一个例子，以加深对矩阵概念的理解.

例 1.4 三个商店 S_1,S_2,S_3 进货四种产品 P_1,P_2,P_3,P_4 的数量(单位：件)如表 1.1.

表 1.1

商店 \ 产品	P_1	P_2	P_3	P_4
S_1	a_{11}	a_{12}	a_{13}	a_{14}
S_2	a_{21}	a_{22}	a_{23}	a_{24}
S_3	a_{31}	a_{32}	a_{33}	a_{34}

四种产品 P_1, P_2, P_3, P_4 的单价(单位：元)及单件重量(单位：kg)如表 1.2.

表 1.2

产品	单价	单件重量
P_1	b_{11}	b_{12}
P_2	b_{21}	b_{22}
P_3	b_{31}	b_{32}
P_4	b_{41}	b_{42}

于是,得到两个矩阵分别为

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix}.$$

2. 转置矩阵

在例 1.4 中,如果将三个商店 S_1, S_2, S_3 进货四种产品 P_1, P_2, P_3, P_4 的数量列成表 1.3 的形式.

表 1.3

产品 \ 商店	S_1	S_2	S_3
P_1	a_{11}	a_{21}	a_{31}
P_2	a_{12}	a_{22}	a_{32}
P_3	a_{13}	a_{23}	a_{33}
P_4	a_{14}	a_{24}	a_{34}

这时得到的矩阵为

$$\begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{pmatrix},$$

此矩阵称为例 1.4 中矩阵 $\boldsymbol{A}$ 的转置矩阵，它是这些元素 a_{ij} 的另外一种排列方式.

定义 1.2 设矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

称下列矩阵为 $\boldsymbol{A}$ 的**转置矩阵**(transpose of a matrix)，记为 $\boldsymbol{A}^{\mathrm{T}}$ 或 $\boldsymbol{A}'$.

$$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

一个矩阵的转置矩阵就是将其行与对应的列互换得到的矩阵，例如对于

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

则 $\boldsymbol{A}$ 的转置矩阵为

$$\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} -1 & 2 & -2 \\ 3 & 4 & -1 \\ 0 & 1 & -1 \\ 3 & 2 & 5 \end{pmatrix}.$$

显然，对于任意矩阵 $\boldsymbol{A}$，有 $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{A}$. 一个方阵的转置矩阵也可以理解为是沿主对角线的翻转.

说明 为了理解为何要讨论转置矩阵，我们是将一个矩阵的转置矩阵理解为表格的不同排列方式，它实际上也是矩阵的一种运算——转置运算.

3. 几种特殊矩阵

有了矩阵概念后，下面介绍几种常用的特殊矩阵.

(1) 单位矩阵

在平面直角坐标系中，相对于坐标原点将点 (x,y) 旋转(rotation) θ 角变换到新的点 (x',y')，见图 1.1.

这时，有

$$\begin{cases} x' = x\cos\theta - y\sin\theta, \\ y' = x\sin\theta + y\cos\theta. \end{cases}$$

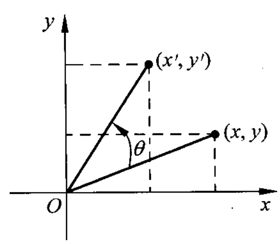

图 1.1

将两个等式右边 x 与 y 的系数用一个矩阵表示为

$$\boldsymbol{R}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},$$

它称为上述变换的旋转矩阵. 当 $\theta=0$ 时,有$\begin{cases}x'=x,\\ y'=y,\end{cases}$而 $\boldsymbol{R}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$,它称为二阶单位矩阵.

一般地,对角线元素全为 1,其他元素全为 0 的 n 阶方阵称为 **n 阶单位矩阵**(identity matrix of order n)或 **n 阶单位阵**,记为 $\boldsymbol{E}$ 或 $\boldsymbol{E}_n$,有时记为 $\boldsymbol{I}$ 或 $\boldsymbol{I}_n$,即

$$\boldsymbol{E}=\begin{bmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{bmatrix},$$

未写出的元素全为 0,也可以在是 0 的部分写一个较大的 0,今后会经常遇到这种处理方式.

将单位矩阵记为 $\boldsymbol{I}$ 在线性代数中应该更好,如产生 n 阶单位矩阵的 MATLAB 命令为 eye(n),只是由于在一般的代数结构讨论中关于乘法运算的“单位元”习惯用 $\boldsymbol{e}$ 或 $\boldsymbol{E}$,与数 1 有同样作用.

二阶单位阵和三阶单位阵分别为

$$\boldsymbol{E}_2=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\quad \boldsymbol{E}_3=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

(2) 对角矩阵

在平面直角坐标系中,为了改变一个对象的大小,可以进行如下缩放(scaling)变换

$$\begin{cases}x'=\lambda_1 x,\\ y'=\lambda_2 y.\end{cases}$$

按上面的方式,该变换的矩阵为 $\boldsymbol{S}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}$,它是上述变换的缩放矩阵,称为二阶对角矩阵.

更一般地,除对角线元素外全为 0 的 n 阶方阵称为 **n 阶对角矩阵**(diagonal matrix of order n)或 **n 阶对角阵**.

设 n 阶对角矩阵中对角线元素依次为 $\lambda_1,\lambda_2,\cdots,\lambda_n$,则该对角阵为

$$\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

也可简记为 $\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

显然,单位矩阵是对角矩阵.

对角线元素相同的 n 阶对角矩阵称为**数量矩阵**(scalar matrix),其一般形式为

$$\mathrm{diag}(k,k,\cdots,k).$$

(3) 零矩阵

元素全为 0 的 $m\times n$ 矩阵称为**零矩阵**(zero matrix),记为 $\mathbf{0}$ 或 $\mathbf{0}_{m\times n}$. n 阶零方阵记为 $\mathbf{0}_n$. 注意 $\mathbf{0}$ 是加粗的,与数 0 有区别. 有各种形式的零矩阵,如

$$\mathbf{0}_{2\times 3}=\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix},\quad \mathbf{0}_3=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}.$$

零矩阵的作用类似于数 0.

(4) 上(下)三角矩阵

线性方程组

$$\begin{cases}x_1+2x_2+3x_3+4x_4=3,\\ \quad\ 7x_2+x_3-2x_4=10,\\ \qquad\quad -x_3+2x_4=2,\\ \qquad\qquad\quad x_4=4\end{cases}$$

的系数矩阵为

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

此矩阵的非零元素都位于右上角的三角形中,称其为一个上三角矩阵.

一般地,对角线以下元素全为 0 的方阵称为**上三角矩阵**(uppertriangular matrix)或上三角阵,其一般形式为

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\ &a_{22}&\cdots&a_{2n}\\ &&\ddots&\vdots\\ &&&a_{nn}\end{pmatrix}.$$

在关于矩阵计算的内容中,还涉及下三角矩阵. 对角线以上元素全为 0 的方阵称为**下三角矩阵**(lower triangular matrix)或下三角阵,其一般形式为

$$\begin{pmatrix}a_{11}&&&\\a_{21}&a_{22}&&\\ \vdots&\vdots&\ddots&\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.$$