矩阵论简明教程

A Concise Tutorial of Matrix Theory

◆王 钢 编著

中国工信出版集团 電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

矩阵论简明教程

王　钢　编著

電子工業出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书从求解“鸡兔同笼”问题和线性变换两个角度引出矩阵的概念，分别从矩阵化简、矩阵分解、矩阵度量和矩阵分析4个方面，介绍了线性代数基础、矩阵的初等性质、矩阵的主要分解方法、范数理论、矩阵分析、矩阵的广义逆和特征值估计等主要内容。最后，结合图像处理的一个例子，简单介绍了矩阵知识如何与实际问题相结合。

本书叙述深入浅出、思路清晰，并配有少量习题，既可作为工科院校硕士研究生的教材，又可作为高等院校数学系高年级本科生的选修课教材，还可作为工科院校有关专业教师、工程技术或研究人员的参考资料。

图书在版编目（CIP）数据

矩阵论简明教程 / 王钢编著. —北京：电子工业出版社，2018.5

ISBN 978-7-121-34110-6

I. ①矩… II. ①王… III. ①矩阵论－高等学校－教材 IV. ①O151.21

中国版本图书馆 CIP 数据核字（2018）第 082195 号

策划编辑：竺南直
责任编辑：张　京
印　　刷：三河市华成印务有限公司
装　　订：三河市华成印务有限公司
出版发行：电子工业出版社
　　　　　北京市海淀区万寿路 173 信箱　　邮编：100036
开　　本：787×1 092　1/16　印张：9.5　　字数：243.2 千字
版　　次：2018 年 5 月第 1 版
印　　次：2018 年 5 月第 1 次印刷
定　　价：29.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：davidzhu@phei.com.cn。

前　言

矩阵论作为数学的一个重要分支，不但具有丰富的内容，而且在信息科学与技术、管理科学与工程等学科中都有十分广泛的应用，因此，学习和掌握矩阵理论的基本概念和基本方法就显得十分必要。目前，高等院校许多专业都把矩阵论设置成研究生的一门必修课，而本书就是针对工科院校非数学专业研究生编写的。

本书是在近十年课堂教学经验的基础上，参考了国内其他院校的相关课程讲义编写而成的。全书共分 7 章，具体内容包括：第 1 章介绍矩阵的由来，分别从“鸡兔同笼”解线性方程组和线性变换两个角度进行叙述；第 2 章介绍矩阵的基本概念、基本性质和常见的几种矩阵；第 3 章介绍矩阵化简问题，即如何把矩阵化简成对角矩阵或分块对角化(Jordon 标准型)；第 4 章介绍矩阵分解问题，即把一个矩阵拆分成几个特殊矩阵乘积的形式，这一章的最后还介绍了矩阵的广义逆问题；第 5 章介绍矩阵度量问题，即把“向量距离”的概念推广到“矩阵范数”的概念，并介绍了如何利用范数理论估计矩阵的特征值；第 6 章介绍矩阵分析问题，即利用微积分的方法来处理矩阵；第 7 章从一个图像处理的简单例子出发，介绍了矩阵如何和实际问题相结合，并拓展介绍了非负矩阵的一些相关知识。同时，在每章章末都配备了一定数量的习题，希望这些习题能够帮助读者巩固本章的知识点。

面向工科院校硕士研究生 48 课时的授课内容，在编写过程中，本书力求兼顾基础理论和应用，培养学生逻辑思维、抽象思维及实际应用的能力。为了使读者在较短时间内尽可能多地掌握矩阵理论基础知识，本书在内容的取舍和结构编排上，还做了如下一些新的尝试。

(1) 适当压缩了广义逆的内容，只介绍广义逆的基本概念和基本求解方法，并放在矩阵分解一章中，使得学生在掌握了矩阵分解的主要方法之后，能够应用这些方法去求解广义逆。

(2) 删掉了特征值估计中一些证明较复杂的理论界结果，并把它放在范数理论一章中，让学生更好地去理解范数这一概念的内涵。

(3) 采用由简入繁的思路组织内容，通过从一些已经学过的、简单的知识中引出要介绍的一些抽象知识，如从解方程组引出矩阵、从欧氏空间的距离概念引出范数的概念，从而避免一开始就介绍过于抽象的知识而打消读者进一步学习的热情。

(4) 本书每一章都采用“基本概念—基本性质—基本定理—例题”的方式展开介绍。对于一些性质和定理的证明非常详细，同时也留出一部分性质作为练习。

本书的主要内容曾为北京航空航天大学电子信息工程学院、软件学院研究生讲授。感谢北京航空航天大学电子信息工程学院给我提供了一个良好的教学、科研平台，感谢张有光副院长、陈杰副院长、王俊副院长、孙则怡老师、张颖老师、郎荣玲老师、原仓周老师在教学过程中给我提供的帮助和支持，感谢电子工业出版社竺南直编辑和张京编辑的帮助。由于作者水平有限，难免有疏漏之处，迫切希望读者批评指正。

编著者

符 号 说 明

$R(C)$	实(复)数集合
$R^n(C^n)$	n维实(复)向量集合
$R^{m\times n}(C^{m\times n})$	$m\times n$维实(复)矩阵集合
$R_r^{m\times n}(C_r^{m\times n})$	秩为r的$m\times n$维实(复)矩阵集合
V^n	n维线性空间
$\dim V$	线性空间的维数
$\mathbf{0}$	零向量、零矩阵或线性空间零元素
$\boldsymbol{I}$	单位矩阵
$\det \boldsymbol{A}$	方阵$\boldsymbol{A}$的行列式
$\mathrm{tr}\boldsymbol{A}$	方阵$\boldsymbol{A}$的迹
$\boldsymbol{A}^{\mathrm{T}}$	矩阵$\boldsymbol{A}$的转置
$\boldsymbol{A}^{\mathrm{H}}$	矩阵$\boldsymbol{A}$的复共轭转置
$\mathrm{rank}\boldsymbol{A}$	矩阵$\boldsymbol{A}$的秩
$\rho(\boldsymbol{A})$	矩阵$\boldsymbol{A}$的谱半径
$\Vert\boldsymbol{A}\Vert$	矩阵$\boldsymbol{A}$的范数
$\boldsymbol{A}^{-1}$	矩阵$\boldsymbol{A}$的逆矩阵
$\boldsymbol{A}^{(i,j,\cdots,l)}$	矩阵$\boldsymbol{A}$的$\{i,j,\cdots,l\}$——广义逆
i	虚数单位

目　录

第 1 章　线性代数基础……1
1.1　从线性方程组谈起……1
1.2　线性空间、线性变换和矩阵……3
1.3　线性子空间基本概念……9
1.4　特殊的线性子空间……12
习题……14
第 2 章　矩阵的基本概念……16
2.1　矩阵的基本运算……16
2.2　矩阵的秩……19
2.3　矩阵的迹……22
2.4　矩阵的特征值和特征向量……22
2.5　正交矩阵和酉矩阵……27
2.5.1　Gram-Schmidt 正交化方法……27
2.5.2　Givens 变换……28
2.5.3　Householder 变换……30
2.6　正规矩阵……32
2.7　正定矩阵与半正定矩阵……34
2.8　特殊的幂矩阵……37
习题……38
第 3 章　矩阵对角化……40
3.1　矩阵的相抵……40
3.2　矩阵的相似……41
3.3　矩阵的对角化……42
3.4　矩阵的正交相似对角化……44
3.5　Jordan 标准形……49
3.5.1　Jordan 标准形的存在定理……49
3.5.2　初等因子法求 Jordan 标准形……51
3.6　Hamilton-Cayley 定理及其应用……56
习题……60
第 4 章　矩阵分解及应用……63
4.1　三角分解 LU……63
4.2　矩阵的 QR 分解……66

4.3 满秩分解……72
4.4 奇异值分解……74
4.5 矩阵的极分解……78
4.6 矩阵的谱分解……80
4.7 扩展主题——广义逆矩阵……83
习题……88
第 5 章 范数理论及其应用……90
5.1 向量范数的定义……90
5.2 三个常用的不等式……91
5.3 常见的向量范数……93
5.4 向量范数的等价性……97
5.5 矩阵范数的定义……99
5.6 常见的矩阵范数……99
5.7 矩阵范数与向量范数之间的相容性……105
5.8 扩展主题 1：矩阵的非奇异性条件……107
5.9 扩展主题 2：特征值估计……109
习题……112
第 6 章 矩阵分析及应用……114
6.1 矩阵序列及其极限……114
6.2 矩阵级数……117
6.3 矩阵幂级数……119
6.4 矩阵函数……121
6.5 函数矩阵的微分……128
6.6 函数矩阵的积分……133
习题……134
第 7 章 矩阵论的高级主题……136
7.1 线性方程组求解的问题……136
7.2 非负矩阵简介……139
7.3 低秩矩阵近似……139
参考文献……145

第1章 线性代数基础

1.1 从线性方程组谈起

中国古代重要的数学著作《孙子算经》中记载着一类著名的问题——“鸡兔同笼”。书中是这样叙述的：“今有鸡兔同笼，上有三十五头，下有九十四足，问鸡、兔各几何？”这四句话的意思是：有若干只鸡和兔同在一个笼子里，从上面数，有35个头；从下面数，有94只脚，求笼中各有几只鸡和兔？

利用初中学过的知识，通过假设有 x 只鸡，y 只兔，根据题意可以列出下面的一个二元一次方程组：

$$\begin{cases} x+y=35 \\ 2x+4y=94 \end{cases} \tag{1-1}$$

利用消元法，第二行减去第一行乘以2，我们有

$$\begin{cases} x+y=35 \\ 2y=24 \end{cases}$$

根据第2行可以得到 y=12，再代入第一个方程中，可得到 x=23，即笼中有23只鸡和12只兔。

通过以上例子，我们可以发现，在消元法求解的过程中，变量 x 和 y 没有参与运算，所以我们可以把上面的二元一次方程组写成下面这种简洁的数表形式：

$$\begin{pmatrix} 1 & 1 & 35 \\ 2 & 4 & 94 \end{pmatrix} \tag{1-2}$$

进一步，上述消元法的求解过程可以仅通过简单的行初等变换加以实现，过程如下：

$$\begin{pmatrix} 1 & 1 & 35 \\ 2 & 4 & 94 \end{pmatrix} \xrightarrow{\text{第1行}\times(-2)+\text{第2行}} \begin{pmatrix} 1 & 1 & 35 \\ 0 & 2 & 24 \end{pmatrix} \xrightarrow{\text{第2行}\div 2} \begin{pmatrix} 1 & 1 & 35 \\ 0 & 1 & 12 \end{pmatrix} \xrightarrow{\text{第2行}\times(-1)+\text{第1行}} \begin{pmatrix} 1 & 0 & 23 \\ 0 & 1 & 12 \end{pmatrix}$$

采用式(1-2)这种数表的紧凑表达形式，我们同样可以得到该方程组的解，而这个数表就是本书要研究的对象——矩阵。矩阵的正式定义如下：

定义 1.1 数域 F 上的 $m\times n$ 个数构成的数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1-3}$$

称为 F 上 m 行、n 列的矩阵，记为 $\boldsymbol{A}_{m\times n}=(a_{ij})_{m\times n}$，其中 $a_{ij}\in F, i=1,2,\cdots,m; j=1,2,\cdots,n$，称为 $\boldsymbol{A}$ 的第 i 行、第 j 列元素。这里，$m\times n$ 叫作矩阵 $\boldsymbol{A}$ 的大小或维数。为了方便，在不引起混淆的情况下，本书一般省略矩阵的维数下标，即 $\boldsymbol{A}=\boldsymbol{A}_{m\times n}$。

注：①若数集 F 含有数 1 且对四则运算封闭，则称 F 为数域，这里的数域一般是指实数域或复数域；②数域 F 上的一切 m 行、n 列的矩阵的集合记为 $F^{m\times n}$。

特别地，当一个矩阵的所有元素全为零时，称这样的矩阵为零矩阵。当行数和列数相同时，称这样的矩阵为方阵，且把行数(或列数)为 n 的方阵称为 n 阶方阵。

对于一个 n 阶方阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$，我们称 $a_{11},a_{22},\cdots,a_{nn}$ 为矩阵 $\boldsymbol{A}$ 的主对角线元素，同时称 $a_{1n},a_{2(n-1)},\cdots,a_{n1}$ 为矩阵 $\boldsymbol{A}$ 的副对角线元素。由此得到如下常用矩阵。

(1) 形如 $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$，即主对角线以下的所有元素都为 0 的矩阵称为上三角矩阵；进一步，主对角线上所有元素都是 1 的上三角矩阵称为单位上三角矩阵，即

$$\boldsymbol{R}=\begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

(2) 形如 $\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$，即主对角线以上的所有元素都为 0 的矩阵称为下三角矩阵；进一步，主对角线上所有元素都是 1 的下三角矩阵称为单位下三角矩阵，即

$$\boldsymbol{L}=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix}$$

(3) 除了主对角线元素以外，其余元素均为 0 的方阵称为对角矩阵(简称对角阵)，即

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

为了方便，有时把对角矩阵记作：$\boldsymbol{A}=\mathrm{diag}\{a_{11},a_{22},\cdots,a_{nn}\}$。

(4) 主对角线元素全为 1 的对角阵称为单位阵，简记为 $\boldsymbol{I}$，即 $\boldsymbol{I}=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$。

通过前面的介绍，我们可以直观地看出，引入矩阵的概念，可以更加简练地描述线性方程组的求解过程。当然，矩阵的引入不仅是为了简化线性方程组的求解，下面，我们将从线性空间和线性变换的角度去理解矩阵这一基本概念。

1.2　线性空间、线性变换和矩阵

回顾线性代数知识可以知道，线性空间和线性变换是线性代数中最基本的两个概念。线性空间是某类客观事物从量的方法的一个抽象，而线性变换则是研究线性空间中的元素之间的最基本联系。由于矩阵也是基于线性空间发展而来的，这一节，我们先从线性空间讲起，介绍如何通过线性变换引出矩阵这一基本概念，从而使读者可以更加深刻地理解矩阵的本质。

线性空间具体定义如下。

定义 1.2　设 V 是一个非空集合，其元素用 x,y,z 等表示；F 是一个数域，其元素用 k,l,m 等表示。如果 V 满足以下 8 条性质：

(I) 在 V 中定义一个“加法”运算，即当 $x,y\in V$ 时，有唯一的和 $x+y\in V$（封闭性），且加法运算满足下列性质：

① 结合律　$x+(y+z)=(x+y)+z$；

② 交换律　$x+y=y+x$；

③ 零元律　存在零元素 $\mathbf{0}$，使 $x+\mathbf{0}=x$；

④ 负元律　对于任一元素 $x\in V$，存在一元素 $y\in V$，使 $x+y=\mathbf{0}$，且称 y 为 x 的负元素，记为 $(-x)$，则有 $x+(-x)=\mathbf{0}$。

(II) 在 V 中定义一个“数乘”运算，即当 $x\in V,k\in F$ 时，有唯一的 $kx\in V$（封闭性），且数乘运算满足下列性质：

⑤ 分配律　$k(x+y)=kx+ky$；

⑥ 分配律　$(k+l)x=kx+lx$；

⑦ 结合律　$k(lx)=(kl)x$；

⑧ 恒等律　$1x=x$；（前面已经提及：数域中一定有元素 1）

则称 V 为数域 F 上的线性空间或向量空间。

简言之，线性空间是指一个非空集合，在其上定义了两种基本运算：加法运算和数乘运算，对这两种运算各自定义了 4 条性质，共 8 条性质。关于对线性空间这一概念的理解，需要注意以下几点：

(1) 线性空间 V 是关于某一个数域上的非空集合。同一个集合，对于不同数域，可能构成不同的线性空间，甚至对有的数域能构成线性空间，而对其他数域不能构成线性空间。当数域 F 为实数域时，V 就称为实线性空间；当 F 为复数域时，V 就称为复线性空间。

(2) 因为 V 中所定义的加法和数乘运算统称为 V 的线性运算，所以不难理解为什么称 V 为线性空间了。关于线性运算应该注意是否满足唯一性、封闭性。

例 1.1　设 $R^{+}=\{全体正实数\}$，$k\in R$，其“加法”及“数乘”运算定义为：

$$x \oplus y = xy, \quad k \circ x = x^k$$

证明：R^+ 是实数域 R 上的线性空间。

证明：首先需要证明两种运算的唯一性和封闭性，其次再证明满足 8 条性质。

(1)唯一性是显然的，对于封闭性，我们有：

若 $x>0$， $y>0$，$k \in R$，则有：

$$x \oplus y = xy \in R^+, \quad k \circ x = x^k \in R^+$$

封闭性得证。

(2)验证 8 条性质成立。

① $x \oplus (y \oplus z) = x \oplus (yz) = xyz = (xy)z = (x \oplus y) \oplus z$；

② $x \oplus y = xy = yx = y \oplus x$；

③ 1 是零元素，因为 $x \oplus 1 = x \cdot 1 = x$；

④ $\dfrac{1}{x}$ 是 x 的负元素，因为 $x \oplus \dfrac{1}{x} = x \cdot \dfrac{1}{x} = 1$；

⑤ $k \circ (x \oplus y) = (xy)^k = x^k y^k = (k \circ x) \oplus (k \circ y)$；

⑥ $(k+l) \circ x = x^{k+l} = x^k x^l = (k \circ x) \oplus (l \circ x)$；

⑦ $k \circ (l \circ x) = (x^l)^k = x^{kl} = (kl) \circ x$；

⑧ $1 \circ x = x^1 = x$。

由此可证，R^+ 是实数域 R 上的线性空间。

定理 1.1 线性空间 V 具有如下性质：

(1)零元素是唯一的，任一元素的负元素也是唯一的。

(2)如下恒等式成立：$0x = \mathbf{0}$， $(-1)x = (-x)$。

证明思路：(1)的证明采用反证法。

(2)第一个恒等式中需要注意：等号左边是数字 0，等号右边是一个零向量。

两个结论的证明非常简单，具体过程这里留作练习。

空间是现代数学的一个基本概念。引出了线性空间的概念之后，我们面临一个如何清晰地表述出线性空间的问题。首先，我们引出线性表示的定义。

定义 1.3 对于 $\boldsymbol{x} \in V$，一组向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in V$，且存在数域 F 中的一组数 $c_1, c_2 \cdots, c_m$，使得

$$\boldsymbol{x} = c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 + \cdots + c_m \boldsymbol{x}_m \tag{1-4}$$

则称 $\boldsymbol{x}$ 为向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 的线性组合，或称 $\boldsymbol{x}$ 可由向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 线性表示。

更进一步，我们有如下定义。

定义 1.4 如果存在一组不全为零的数 $c_1, c_2, \cdots, c_m \in F$，使得对于向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in V$ 的线性组合满足：

$$\sum_{i=1}^{m} c_i \boldsymbol{x}_i = \mathbf{0} \tag{1-5}$$

则称向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 线性相关，否则称其线性无关。

定义 1.5　线性空间 V 中线性无关向量组所含向量的最大个数称为 V 的维数。当线性空间 V 的维数为 n 时，记作 $\dim V=n$。

维数为 n 的线性空间 V 称为 n 维线性空间；当 $n=+\infty$ 时，称为无限维线性空间。本书仅介绍有限维线性空间。

下面，我们介绍线性代数中的两个重要概念：基和坐标。对于线性空间中的任意一个向量都可以在一组基(相当于度量单位)下表示成一组坐标(相当于具体的尺度)。这里，对于度量标准——基的选取，不但希望对线性空间中的每一个向量都适用，同时也希望尽可能简洁，这就要求基的定义满足以下两个约束：①基的组成向量线性无关；②线性空间中的任何一个向量都可以由基线性表示。

定义 1.6　设 V 是数域 F 上的线性空间，$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r\,(r\geqslant 1)$ 是属于 V 的 r 个任意向量，如果它满足：

(1) $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 线性无关；

(2) V 中任一元素 $\boldsymbol{x}$ 均可由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 线性表示，即存在一组数 $c_1,c_2,\cdots,c_r\in F$，使得

$$\boldsymbol{x}=c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_r\boldsymbol{x}_r$$

则称 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 为线性空间 V 的一个基，并称 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 为该基的基向量。

根据基的定义可知，系数 $c_1,c_2,\cdots,c_r$ 是被向量 $\boldsymbol{x}$ 和基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 唯一确定的，称之为向量 x 在基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 下的坐标，记作 $(c_1,c_2,\cdots,c_r)$。

关于对线性空间的基这一概念的理解，需要注意以下几点：

(1) 基是线性空间 V 中的最大线性无关向量组；

(2) 基中所含向量的个数也是线性空间 V 的维数；

(3) 一个线性空间的基是不唯一的，但不同的基所含向量的个数相等。

由于线性空间中基的选取不具有唯一性，那么，两个不同的基之间存在什么关系呢？为了刻画不同基之间的转换关系，得到了过渡矩阵的概念。假设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 及 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_r$ 都是线性空间 V 的基，根据基的定义可知，一组基可以由另一组基线性表示，即有如下关系：

$$\begin{cases}\boldsymbol{y}_1=c_{11}\boldsymbol{x}_1+c_{21}\boldsymbol{x}_2+\cdots+c_{r1}\boldsymbol{x}_r\\ \boldsymbol{y}_2=c_{12}\boldsymbol{x}_1+c_{22}\boldsymbol{x}_2+\cdots+c_{r2}\boldsymbol{x}_r\\ \qquad\vdots\\ \boldsymbol{y}_r=c_{1r}\boldsymbol{x}_1+c_{2r}\boldsymbol{x}_2+\cdots+c_{rr}\boldsymbol{x}_r\end{cases}$$

写成矩阵形式就有：

$$(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_r)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r)\cdot\boldsymbol{A}$$

这里，$\boldsymbol{A}=\begin{bmatrix}c_{11}&c_{12}&\cdots&c_{1r}\\c_{21}&c_{22}&\cdots&c_{2r}\\\vdots&\vdots&\ddots&\vdots\\c_{r1}&c_{r2}&\cdots&c_{rr}\end{bmatrix}$，则称矩阵 $\boldsymbol{A}$ 为从基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 到基 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_r$ 的过渡矩阵。

通过过渡矩阵可以把同一线性空间中的不同基联系起来，即不同基之间是可以相互转换的。

我们知道，线性空间描述的是一个满足某些性质的元素的集合。如果把线性空间理解成一个“静态”概念，则线性变换描述的是线性空间中一个元素“运动”到另一个元素的“动态”概念。

定义 1.7 设 V 是数域 F 上的一个线性空间，T 是 V 到自身的一个映射，使得对于 V 中的任意元素 $\boldsymbol{x}$ 均存在唯一的 $\boldsymbol{y}\in V$ 与之对应，则称 T 为 V 的一个变换或算子，记为

$$T\boldsymbol{x}=\boldsymbol{y}$$

称 $\boldsymbol{y}$ 为 $\boldsymbol{x}$ 在变换 T 下的像，$\boldsymbol{x}$ 为 $\boldsymbol{y}$ 的原像。若变换 T 同时还满足如下性质：

$$T(k\boldsymbol{x}+l\boldsymbol{y})=k(T\boldsymbol{x})+l(T\boldsymbol{y}) \qquad \forall \boldsymbol{x},\boldsymbol{y}\in V,k,l\in F$$

称 T 为线性变换或线性算子。

例 1.2 对于一个二维实向量空间 $R^2=\left\{\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\middle|\ \xi_i\in R,\ i=1,2\right\}$，我们可以证明如下结论：将其绕原点旋转 θ 角的操作就是一个线性变换。

证明：关于旋转操作，如图 1-1 所示。

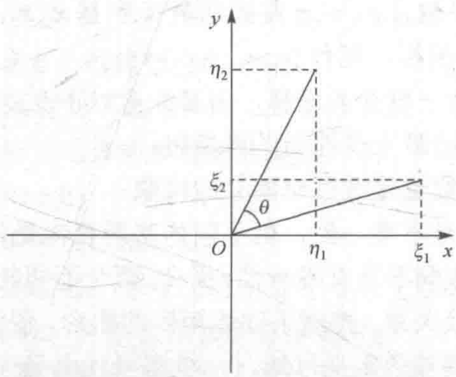

图 1-1 旋转操作

我们取 $\boldsymbol{x}=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\in R^2$，将其绕原点旋转 θ 角后得到点 $\boldsymbol{y}=T\boldsymbol{x}=\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}$，这里有 $\begin{cases}\eta_1=\xi_1\cos\theta-\xi_2\sin\theta\\ \eta_2=\xi_1\sin\theta+\xi_2\cos\theta\end{cases}$，写成矩阵的形式：

$$\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\in R^2$$

显然，这种旋转操作是一种变换，下面证明其为线性变换。

$$\forall \boldsymbol{x}=\begin{pmatrix}x_1\\ x_2\end{pmatrix},\quad \boldsymbol{z}=\begin{pmatrix}z_1\\ z_2\end{pmatrix}\in R^2,\quad k,l\in R$$

简单计算可得：

$$k\boldsymbol{x}+l\boldsymbol{z}=\begin{pmatrix}kx_1\\kx_2\end{pmatrix}+\begin{pmatrix}lz_1\\lz_2\end{pmatrix}=\begin{pmatrix}kx_1+lz_1\\kx_2+lz_2\end{pmatrix}$$

$$\begin{aligned}T(k\boldsymbol{x}+l\boldsymbol{z})&=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}kx_1+lz_1\\kx_2+lz_2\end{pmatrix}\\&=k\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}+l\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}z_1\\z_2\end{pmatrix}\\&=k(T\boldsymbol{x})+l(T\boldsymbol{z})\end{aligned}$$

所以，T 是线性变换。

证毕。

介绍了线性变换的定义之后，接下来简单介绍线性变换的几个性质。

性质 1：$T(0)=0$，$T(-\boldsymbol{x})=-T(\boldsymbol{x})$。

性质 2：如果向量组 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 线性相关，则 $T\boldsymbol{x}_1,\cdots,T\boldsymbol{x}_n$ 也线性相关。

性质 3：如果向量组 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 线性无关，且 T 是单射，则 $T\boldsymbol{x}_1,\cdots,T\boldsymbol{x}_n$ 线性无关。

证明：反证法，假设 $T\boldsymbol{x}_1,\cdots,T\boldsymbol{x}_2$ 线性相关，根据线性相关的定义，存在一组不全为 0 的数 $k_1,k_2,\cdots,k_n$，使得 $k_1T\boldsymbol{x}_1+k_2T\boldsymbol{x}_2+\cdots+k_nT\boldsymbol{x}_n=0$。根据线性变换的定义，可得

$$k_1T\boldsymbol{x}_1+k_2T\boldsymbol{x}_2+\cdots+k_nT\boldsymbol{x}_n=T\left(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_n\boldsymbol{x}_n\right)=0$$

因为线性变换 T 是单射，则有 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_n\boldsymbol{x}_n=0$，这与假设向量组 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 线性无关矛盾。

证毕。

本小节的最后，我们讨论如何将抽象的线性变换转化为具体的矩阵形式，即线性变换用矩阵表示问题。

设 T 是 n 维线性空间 V 的一个线性变换，且 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V 的一个基，$\forall\boldsymbol{x}\in V$，根据线性代数的知识，可知其存在唯一的坐标表示，即

$$\boldsymbol{x}=\xi_1\boldsymbol{x}_1+\xi_2\boldsymbol{x}_2+\cdots+\xi_n\boldsymbol{x}_n=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}$$

根据线性变换的定义，对上式两端同时作用线性变换 T，可以得到

$$\begin{aligned}T\boldsymbol{x}&=T(\xi_1\boldsymbol{x}_1+\xi_2\boldsymbol{x}_2+\cdots+\xi_n\boldsymbol{x}_n)=T(\xi_1\boldsymbol{x}_1)+T(\xi_2\boldsymbol{x}_2)+\cdots+T(\xi_n\boldsymbol{x}_n)\\&=\xi_1T\boldsymbol{x}_1+\xi_2T\boldsymbol{x}_2+\cdots+\xi_nT\boldsymbol{x}_n\\&=[T\boldsymbol{x}_1,T\boldsymbol{x}_2,\cdots,T\boldsymbol{x}_n]\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}=[T(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)]\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}\end{aligned}\tag{1-6}$$

通过式(1-6)可以看出：要想确定线性变换 T，只需确定基元素在该变换下的像就可以了。

下面，我们假设 $T\boldsymbol{x}_i=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{pmatrix}a_{i1}\\a_{i2}\\\vdots\\a_{in}\end{pmatrix}$，$i=1,2,\cdots,n$，则写成矩阵的形式为：

$$T[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\vdots&\vdots&\ddots&\vdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\boldsymbol{A} \tag{1-7}$$

这里，$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\vdots&\vdots&\ddots&\vdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}$。

下面，在相同的基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 下，对于任意向量 $\boldsymbol{x}$，变换后 $T\boldsymbol{x}$ 的坐标表示为：

$$T\boldsymbol{x}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\begin{pmatrix}\eta_1\\\eta_2\\\vdots\\\eta_n\end{pmatrix} \tag{1-8}$$

同时，根据式(1-6)和式(1-7)，我们还有：

$$T\boldsymbol{x}=[T(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)]\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]\boldsymbol{A}\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix} \tag{1-9}$$

上式与式(1-8)对比可知：

$$\begin{pmatrix}\eta_1\\\eta_2\\\vdots\\\eta_n\end{pmatrix}=\boldsymbol{A}\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix} \tag{1-10}$$

即：当 $\boldsymbol{x}\leftrightarrow\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}$ 时，$T\boldsymbol{x}\leftrightarrow\boldsymbol{A}\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_n\end{pmatrix}$。

这里，把 $\boldsymbol{A}$ 称为 T 在基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 下的矩阵。

综上所述，线性空间、线性变换和矩阵三者之间的关系可以简述为：线性空间是包

含一些元素的非空集合，基是刻画这个线性空间的一个“计量基准”；线性变换是把线性空间中的一个元素映射到另一个元素的变换，相当于定义在线性空间上的一种“运算”；矩阵就是这种线性变换在线性空间某一个基准条件下的具体体现。

总结前面介绍的知识，我们可以从以下两个方面来理解矩阵：

(1)从定义的角度来看，矩阵就是一个数表，用它可以更加简洁地处理线性方程组；

(2)基是线性空间的一个本质特征，从线性空间的某一个具体的基的角度来看，矩阵和线性变换是等价的。

1.3　线性子空间基本概念

虽然本书以介绍矩阵理论为主，但是线性子空间的相关知识在实际工程中被广泛应用，所以本节简单介绍线性子空间的相关概念。

定义 1.8　数域 F 上的线性空间 V 的一个非空子集合 W 称为 V 的一个线性子空间(或简称子空间)，如果 W 对于 V 中所定义的加法和数乘两种运算也构成数域 F 上的线性空间，即：

(1)如果 $\boldsymbol{x},\boldsymbol{y}\in W$，则 $\boldsymbol{x}+\boldsymbol{y}\in W$；

(2)如果 $\boldsymbol{x}\in W,k\in F$，则 $k\boldsymbol{x}\in W$。

关于线性子空间的概念，需要指出的是：

(1)因为 W 是 V 的一个子集，且关于线性运算封闭，所以线性子空间也是线性空间。

(2)容易看出，每个非零线性空间至少还有两个子空间，一个是它自身，另一个是仅由零向量构成的子集，称后者为零子空间。这两个特殊子空间有时被称为平凡子空间，而其他线性子空间叫作非平凡子空间。

(3)由于零子空间不含有线性无关的向量，因此没有基，规定其维数等于 0。

(4)任何一个线性子空间 W 的维数都不大于整个线性空间 V 的维数，即 $\dim W\leqslant\dim V$。

例 1.3　在线性空间 R^n 中，齐次线性方程组：

$$\begin{cases}a_{11}\boldsymbol{x}_1+a_{12}\boldsymbol{x}_2+\cdots+a_{1n}\boldsymbol{x}_n=0\\a_{21}\boldsymbol{x}_1+a_{22}\boldsymbol{x}_2+\cdots+a_{2n}\boldsymbol{x}_n=0\\\qquad\vdots\\a_{n1}\boldsymbol{x}_1+a_{n2}\boldsymbol{x}_2+\cdots+a_{nn}\boldsymbol{x}_n=0\end{cases}$$

的全部解向量组成一个子空间，这个子空间叫作齐次线性方程组的解空间。解空间的基就是方程组的基础解系。

通过举例说明了线性子空间的存在性以后，介绍线性子空间的生成问题。

定义 1.9　设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 是线性空间 V 中的一组向量，这组向量所有可能的线性组合所构成的集合 $\{c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_m\boldsymbol{x}_m\mid c_i\in F,i=1,2,\cdots,m\}$ 是非空的，而且容易验证对线性空间 V 的两种运算封闭，因而是 V 的一个子空间，这个子空间称为由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 生成的子

空间，记作 $L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m)=\{c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_m\boldsymbol{x}_m \mid c_i\in F, i=1,2,\cdots,m\}$。有时也把 $L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m)$ 记作 $\mathrm{span}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m)$。

关于线性子空间有如下几个比较重要的性质。

性质 4：两个向量组生成相同子空间的充分必要条件是这两个向量组等价。

性质 5：$L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m)$ 的维数等于生成向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的极大线性无关向量组的向量个数。

性质 6：设 W 是数域 F 上 n 维线性空间 V 的一个 $m(0\leqslant m\leqslant n)$ 维子空间，$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 是 W 的一组基，那么这组向量必定可扩充为整个线性空间 V 的一组基，即在 V 中必定可以找到 $n-m$ 个向量 $\boldsymbol{x}_{m+1},\boldsymbol{x}_{m+2},\cdots,\boldsymbol{x}_n$，使得 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V 的一组基。

证明：证明思路是对维数差 $n-m$ 做数学归纳法。设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 是 W 的一个基，当 $n-m=0$ 时，$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 已是 V 的一个基。

假定 $n-m=k$ 时定理成立，考虑 $n-m=k+1$ 的情形。因为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\in V$ 且线性无关，但又不是 V 的基，故有 $\boldsymbol{x}_{m+1}\in V$ 且 $\boldsymbol{x}_{m+1}$ 不能由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 线性表示，因而 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m,\boldsymbol{x}_{m+1}$ 是线性无关的。由于 $L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m,\boldsymbol{x}_{m+1})$ 是 V 的 $m+1$ 维子空间，且 $n-(m+1)=k$，由归纳假设知 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m,\boldsymbol{x}_{m+1}$ 可以扩充成 V 的基，故 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 可以扩充成 V 的基。

证毕。

下面介绍子空间的交、和、直和等重要概念。

定义 1.10 设 W_1、W_2 是线性空间 V 的两个子空间，称

$$W_1\cap W_2=\{\boldsymbol{\alpha}\mid\boldsymbol{\alpha}\in W_1,\ 且\boldsymbol{\alpha}\in W_2\}$$

为 W_1、W_2 的交。

两个子空间的交具有以下性质。

性质 7：设 W_1、W_2 是线性空间 V 的两个子空间，那么它们的交 $W_1\cap W_2$ 也是 V 的子空间。

性质 8：交换律，即 $W_1\cap W_2=W_2\cap W_1$。

性质 9：结合律，即 $(W_1\cap W_2)\cap W_3=W_1\cap(W_2\cap W_3)$。

以上性质的证明仅根据子空间的交的定义即可完成，具体证明留作练习。这里需要指出的是，以上 3 条性质可以推广到多个子空间的交。

定义 1.11 设 W_1、W_2 是线性空间 V 的两个子空间，称

$$W_1+W_2=\{\boldsymbol{x}+\boldsymbol{y}\mid\boldsymbol{x}\in W_1,\ \boldsymbol{y}\in W_2\}$$

为 W_1、W_2 的和。

两个子空间的和具有以下性质。

性质 10：设 W_1、W_2 是线性空间 V 的两个子空间，那么它们的和 W_1+W_2 也是 V 的子空间。

性质 11：交换律，即 $W_1+W_2=W_2+W_1$。

性质 12：结合律，即 $(W_1+W_2)+W_3=W_1+(W_2+W_3)$。

以上性质的证明仅根据子空间的和的定义即可完成，具体证明留作练习。以上性质同样可以推广到多个子空间的和。

关于子空间的交与和，其实可以看作子空间的两种运算。这里需要强调的是，两个