JUZHEN LILUN JIQI YINGYONG

矩阵理论及其应用

黄有度 朱士信 编著

3

合肥工业大学出版社

矩阵理论及其应用

矩阵理论及其应用

黄有度　朱士信　编著

合肥工业大学出版社

图书在版编目(CIP)数据

矩阵理论及其应用/黄有度,朱士信编著.—合肥:合肥工业大学出版社,2005.8

ISBN 7-81093-287-X

Ⅰ.矩... Ⅱ.①黄...②朱... Ⅲ.矩阵—理论 Ⅳ.0151.21

中国版本图书馆 CIP 数据核字(2005)第 091454 号

矩阵理论及其应用

黄有度 朱士信 编著　　责任编辑 马国锋 孟宪余

出 版	合肥工业大学出版社	版 次	2005 年 8 月第 1 版
地 址	合肥市屯溪路 193 号	印 次	2005 年 8 月第 1 次印刷
邮 编	230009	开 本	850×1168 1/32
电 话	总编室:0551-2903038	印 张	8
	发行部:0551-2903198	字 数	193 千字
网 址	www.hfutpress.com.cn	发 行	全国新华书店
E-mail	press@hfutpress.com.cn	印 刷	合肥杏花印务股份有限公司

ISBN 7-81093-287-X/O·22　　定价:18.00 元

序

数学是在相当广泛意义下研究现实世界中数量关系和空间形式的科学。数学的特点是：内容的抽象性，应用的广泛性，推理的严谨性和结论的明确性。关于数学的作用，北京大学数学科学学院张恭庆院士将其分为三个层次：第一层次是为其他学科提供语言、概念、思想、理论和方法。自然科学和经济、管理等社会科学离开了数学便无从产生和发展。第二个层次是直接应用于工程技术、生产活动，这方面的例子是大量的。第三个层次是作为一种文化，对全社会的成员起着潜移默化的作用。一个民族数学修养的高低，对这个民族的文明有很大的影响。这也就是说数学有三个层面：作为理论思维的数学；作为技术应用的数学；作为文化修养的数学。不管从哪一个层面上讲，数学对于我们每一个人都是十分重要的。

作为数学的一个重要分支，矩阵论有一整套理论、思想和方法，它所包含的内容极为丰富。作为一种基本工具，矩阵论在自然、科学、工程技术、经济理论和管理科学中有非常广泛的应用。因此，学习和掌握矩阵的基本理论和方法，对于理、工、经、管类的研究生来说是十分必要的。矩阵论这门课程内容比较丰富，理论性比较强，概念比较抽象，思维方式也比较独特。本书的编著者不仅学术造诣深、有扎实的矩阵理论功底和宽厚的知识面，而且他们有多年从事这门课程的教学经历，有丰富的教学经验。他们在参阅了大量相关资料的基础上，结合自己科研和教学的体会，从各方面的实际情况出发，完成了这部教材。本教材内容取舍得当，结构安排合理完整，论述深入浅出、简明易懂、准确严谨，因而是一部很实用的矩阵论教材。

复旦大学李大潜院士曾说过，数学教育本质上是一种素质教育。学习数学，不仅要学到许多重要的数学概念、方法和结论，更要领会数学的精神实质和思想方法。我们相信，通过矩阵论的学习，一定会对大家在知识的积累、素质的提高和能力的培养上有所帮助。

苏仪昭

2005.8.8

前　言

矩阵理论是数学学科的重要分支，在其他科学技术领域——从自然科学、工程技术到经济管理、社会科学也有重要的应用。现代科学技术，特别是计算机和计算技术的迅猛发展，为矩阵理论的应用和研究开辟了更加广阔的前景。因此，对许多专业的研究生来说，矩阵理论是一门重要的基础课，对于其他课程的学习和将来的科研工作起着至关重要的作用。

由于矩阵理论应用领域的广泛性，本书的内容不可能全面涉及很具体的应用方法，因此往往会让学习者产生抽象、难懂的感觉。编写过程中，针对当前研究生教育的实际情况，力求做到深入浅出、简明易懂，兼顾内容的深度与广度。每一章末尾配有一定数量的习题，为便于教学，书后附有习题的参考答案，但希望学习者尽量独立完成习题的解答，否则习题的功效将会大打折扣。

矩阵理论以线性代数为基础，要求学习者对线性代数的基本概念和计算方法已牢固掌握并能熟练运用。本书分为六章，第一章是线性空间和线性变换，是和线性代数的衔接点。本章内容已在线性代数中作了简单介绍，此处再进行复习、补充和提高；第二章内积空间，在线性空间中引入内积的概念；第三章是矩阵特征值与约当标准形，引入了多项式矩阵及其标准形、一般矩阵的约当标准形的概念和计算；第四章是矩阵的范数和幂级数，引入了向量、矩阵的范数，介绍了矩阵幂级数及其收敛性的判断方法；第五章是矩阵函数及其应用，是前面各章的抽象理论和实际应用间的接口；第六章是矩阵特征值的估计与广义逆矩阵，介绍了矩阵特征值的估计方法，并引入了广义逆矩阵的概念。

本书的编写参照了国内外的一些有关教材以及合肥工业大学数学系苏家铎等编写的讲义，编写过程中，得到了合肥工业大学校领导、研究生部领导和理学院领导的大力支持和热情帮助，编著者谨表示衷心的感谢。

限于编著者的水平，书中难免不妥及谬误之处，祈望读者批评指正。

编著者

2005 年 5 月

目　录

第 1 章 线性空间与线性变换

本章将把向量空间的概念推广到线性空间，并讨论线性空间的一些性质.

§1.1 线性空间

1. 线性空间

在线性代数中已给出了向量与向量空间的概念.

n 个数构成的有序数组称为 n 维向量，由实数构成的向量称为实向量，所有的 n 维实向量的集合记为$\mathbb{R}^n$. 向量与向量之间定义了加法运算，数和向量之间定义了数乘运算. 加法运算和数乘运算满足下面将提到的一些特定规律.

若 V 是 n 维向量的非空集合，且对向量的加法和数乘运算封闭，即运算的结果仍属于该集合，则称 V 为**向量空间**.

向量空间的定义虽然只有上面一句话，但是其中有更多的内涵. 在向量空间中进行加法运算时，参加运算的是向量空间中的两个元素，即两个向量. 但是进行数乘运算时，需要有一个数来和向量进行数乘运算，这个数不是向量空间中的元素，因此向量空间必须得到一个数集的支持，以便进行数乘运算. 由于向量运算的需要，这个数集必须对加、减、乘、除这四种运算封闭(作除法时除数不为零).

对加、减、乘、除封闭的非空数集称为**数域**. 由此定义可知，任

何数域中都包含 0 和 1.

由此可见，向量空间 V 必须伴随有一个数域 P. 显然，实向量空间伴随的是实数域$\mathbb{R}$，复向量空间伴随的是复数域$\mathbb{C}$.

向量的加法和数乘满足以下八条规则（$\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$，$\forall\, \lambda, \mu \in P$）：

1）$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$；（加法的交换律）

2）$(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$；（加法的结合律）

3）V 中有一个元素 $\mathbf{0}$，对 V 中任何元素 $\boldsymbol{x}$，有 $\boldsymbol{x}+\mathbf{0}=\boldsymbol{x}$（$\mathbf{0}$ 称为 V 的**零元素**）；

4）对 V 中每个元素 $\boldsymbol{x}$，存在 V 中的元素 $\boldsymbol{y}$，使 $\boldsymbol{x}+\boldsymbol{y}=\mathbf{0}$（$\boldsymbol{y}$ 称为 $\boldsymbol{x}$ 的**负元素**，记为 $-\boldsymbol{x}$）；

5）$1\boldsymbol{x}=\boldsymbol{x}$；

6）$\lambda(\mu\boldsymbol{x})=(\lambda\mu)\boldsymbol{x}=\mu(\lambda\boldsymbol{x})$；（数乘的结合律）

7）$(\lambda+\mu)\boldsymbol{x}=\lambda\boldsymbol{x}+\mu\boldsymbol{x}$；（数乘对加法的分配律）

8）$\lambda(\boldsymbol{x}+\boldsymbol{y})=\lambda\boldsymbol{x}+\lambda\boldsymbol{y}$.（数乘对加法的分配律）

这八条规则在定义向量空间前即已给出，因此在定义向量空间时理所当然认为它们是满足的.

现在，把向量空间的概念推广到一般集合上.

定义 1 设 V 是一个非空集合，P 是一个数域. 在 V 中定义了一种二元运算，称为加法，即对 V 中任意两个元素 $\boldsymbol{x}, \boldsymbol{y}$，都有 V 中唯一的一个元素 $\boldsymbol{z}$ 与它们对应，称为 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的和，记为 $\boldsymbol{z}=\boldsymbol{x}+\boldsymbol{y}$. 在数域 P 与集合 V 的元素间定义了一种运算，称为**数乘**，即对 P 中任一数 λ 和 V 中任一元素 $\boldsymbol{x}$，都有 V 中唯一的一个元素 $\boldsymbol{y}$ 与它们对应，称为 λ 和 $\boldsymbol{x}$ 的**数积**，记为 $\boldsymbol{y}=\lambda\boldsymbol{x}$. 而且，加法和数乘满足上述八条规则，则称 V 为数域 P 上的**线性空间**.

加法运算和数乘运算称为**线性运算**.

线性空间中的元素也称为向量，当然这里的向量已不一定是

有序数组了.

易证线性空间有如下性质:

线性空间的零向量唯一;每个向量的负向量唯一;零和任何向量的数积为零向量;任何数与零向量的数积为零向量;-1 和任何向量的数积即为该向量的负向量.

线性空间的例子:

例 1 显然,向量空间为线性空间.

例 2 由数域 P 中的数为元素构成的所有 $m\times n$ 矩阵,按通常矩阵的加法和数与矩阵的乘法,构成线性空间 $P^{m\times n}$.

例 3 n 为一个正整数,P 为数域,则系数属于 P 而变量为 x 的所有次数小于 n 的多项式集合(包括零多项式),按通常多项式的加法和数与多项式的乘法,构成 P 上的线性空间,记为 $P[x]_n$. 若无"次数小于 n"的限制,也构成线性空间,记为 $P[x]$.

例 4 定义在区间$[a,b]$上的所有实值连续函数构成的集合,按函数的加法和数乘,构成实数域$\mathbb{R}$上的线性空间 $C[a,b]$.

以上四例的验证请读者完成.

例 5 V 为$\mathbb{R}^+=\{x\in\mathbb{R}\mid x>0\}$,$P$ 为$\mathbb{R}$,定义"加法$\oplus$"和"数乘$\otimes$"分别为:$x\oplus y=xy$,$\lambda\otimes x=x^\lambda$,$x,y\in\mathbb{R}^+$,$\lambda\in\mathbb{R}$. 按如此定义的加法和数乘,$V$ 为 P 上的线性空间.

此例中,V 的"零元素"为 1,x 的"负元素"为 $1/x$.

$$(\lambda+\mu)\otimes x=x^{\lambda+\mu}=x^\lambda x^\mu=(\lambda\otimes x)\oplus(\mu\otimes x),$$

$$\lambda\otimes(x\oplus y)=(xy)^\lambda=x^\lambda y^\lambda=(\lambda\otimes x)\oplus(\lambda\otimes y),$$

其余部分的验证留给读者完成.

此例说明,线性空间中的"加法"与"数乘"和一般的加法和数乘可能完全是两回事.

2. 线性空间的基和维数

与向量空间中的向量一样,线性空间中的向量也有线性相关

和线性无关等有关概念.

线性空间中若干个向量经数乘后再求和,称为这些向量的**线性组合**.一个向量等于一组向量的线性组合,则说该向量可由这组向量**线性表示**.

设 $x_1,x_2,\cdots,x_r$ 为数域 P 上的线性空间 V 中的一组向量,若有 P 中不全为零的一组数 $\lambda_1,\lambda_2,\cdots,\lambda_r$,使它们与 $x_1,x_2,\cdots,x_r$ 构成的线性组合成为零向量:

$$\lambda_1 x_1+\lambda_2 x_2+\cdots+\lambda_r x_r=0,$$

则称向量组 $x_1,x_2,\cdots,x_r$ **线性相关**.若满足上式的一组数不存在,或者说要使上式成立,必须所有的 λ_k 都为零,则称向量组 $x_1,x_2,\cdots,x_r$ **线性无关**.

定义 2 设线性空间 V 中有 n 个向量 $x_1,x_2,\cdots,x_n$ 满足:

1) $x_1,x_2,\cdots,x_n$ 线性无关;

2) V 中任何向量都可由 $x_1,x_2,\cdots,x_n$ 线性表示,则 $x_1,x_2,\cdots,x_n$ 称为 V 的一组**基**,n 称为 V 的**维数**.V 的维数记为 $\dim(V)$,这里 $\dim(V)=n$.

当 V 的维数为有限时,V 称有限维线性空间,否则称无限维线性空间.在前面所给的线性空间例子中,$P[x]$和 $C[a,b]$为无限维空间,其余的是有限维空间.本书只讨论有限维空间.

实际上,线性空间的一组基即为线性空间的一个最大线性无关组;反之,任何一个最大线性无关组都是一组基.(留作习题)

由此可见,一个线性空间可以有无穷多组基.

定义 3 设 $x_1,x_2,\cdots,x_n$ 为 V 的一组基,对 V 中任一向量 y,必存在一组数 $\lambda_1,\lambda_2,\cdots,\lambda_n$,使

$$y=\lambda_1 x_1+\lambda_2 x_2+\cdots+\lambda_n x_n=(x_1,x_2,\cdots,x_n)\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix},$$

$(\lambda_1,\lambda_2,\cdots,\lambda_n)^T$ 称为 $\boldsymbol{y}$ 在基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 下的坐标.

由基的线性无关性立即可得,任何向量在给定的基下的坐标是唯一的.

例如,

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

为 $P^{2\times2}$ 的一组基,$P^{2\times2}$ 为 4 维空间;

$1,x,x^2$ 为 $P[x]_3$ 的一组基,$P[x]_3$ 为 3 维空间.

在例 5 中所给的线性空间中,任何不等于 1 的正数都可以作为基(注意此空间的零元素为 1). 例如取 $a(a>0,a\neq1)$,则对任何 $x\in\mathbb{R}^+$,取 $\lambda=\log_a x$,有 $\lambda\otimes a=a^\lambda=a^{\log_a x}=x$,即 x 可由 a 线性表示,故 a 即为一组基,即此线性空间是 1 维的.

3. 基变换和坐标变换

由于线性空间可以有不同的基,而一个向量在不同的基下一般有不同的坐标,下面讨论同一向量在不同基下的坐标间的关系.

设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 和 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 为线性空间 V 的两组基,则后一组基的每个基向量用前一组基唯一线性表示为

$$\begin{aligned}\boldsymbol{y}_1&=p_{11}\boldsymbol{x}_1+p_{21}\boldsymbol{x}_2+\cdots+p_{n1}\boldsymbol{x}_n,\\ \boldsymbol{y}_2&=p_{12}\boldsymbol{x}_1+p_{22}\boldsymbol{x}_2+\cdots+p_{n2}\boldsymbol{x}_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ \boldsymbol{y}_n&=p_{1n}\boldsymbol{x}_1+p_{2n}\boldsymbol{x}_2+\cdots+p_{nn}\boldsymbol{x}_n.\end{aligned}\tag{1}$$

写成矩阵形式为

$$(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{pmatrix}p_{11}&p_{12}&\cdots&p_{1n}\\p_{21}&p_{22}&\cdots&p_{2n}\\ \cdots&\cdots&\cdots&\cdots\\p_{n1}&p_{n2}&\cdots&p_{nn}\end{pmatrix}.$$

令 $\boldsymbol{P}=(p_{ij})_{n\times n}$，得

$$(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{P}. \tag{2}$$

称 $\boldsymbol{P}$ 为从基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 到 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 的过渡矩阵. (1)式或(2)式称为**基变换公式**.

设一个向量在基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 和 $\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 下的坐标分别为 $(\lambda_1,\lambda_2,\cdots,\lambda_n)^T$ 和 $(\mu_1,\mu_2,\cdots,\mu_n)^T$，则有

$$\begin{aligned}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}&=(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n)\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}\\&=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{P}\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}.\end{aligned}$$

由坐标的唯一性知

$$\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}=\boldsymbol{P}\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix},\quad\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}=\boldsymbol{P}^{-1}\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}. \tag{3}$$

这就是同一向量在不同基下的坐标间的关系. (3)式称为**坐标变换公式**.

例 6 已知 $P[x]_3$ 中的两组基：$\boldsymbol{\alpha}_1=1,\boldsymbol{\alpha}_2=x,\boldsymbol{\alpha}_3=x^2$ 和 $\boldsymbol{\beta}_1=1$，$\boldsymbol{\beta}_2=x-x_0,\boldsymbol{\beta}_3=(x-x_0)^2$，求前一组基到后一组基的过渡矩阵，并求 $\boldsymbol{q}=3-x^2$ 在后一组基下的坐标.

解 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$，

$\boldsymbol{\beta}_2=-x_0\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$，

$\boldsymbol{\beta}_3=x_0{}^2\boldsymbol{\alpha}_1-2x_0\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3$，

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&-x_0&x_0^2\\0&1&-2x_0\\0&0&1\end{pmatrix},$$

过渡矩阵 $P=\begin{pmatrix}1&-x_0&x_0^2\\0&1&-2x_0\\0&0&1\end{pmatrix}$，易见 q 在前一组基下的坐标为 $\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}=\begin{pmatrix}3\\0\\-1\end{pmatrix}$，故在后一组基下的坐标为

$$\begin{pmatrix}\mu_1\\\mu_2\\\mu_3\end{pmatrix}=P^{-1}\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}=\begin{pmatrix}1&-x_0&x_0^2\\0&1&-2x_0\\0&0&1\end{pmatrix}^{-1}\begin{pmatrix}3\\0\\-1\end{pmatrix}$$
$$=\begin{pmatrix}1&x_0&x_0^2\\0&1&2x_0\\0&0&1\end{pmatrix}\begin{pmatrix}3\\0\\-1\end{pmatrix}=\begin{pmatrix}3-x_0^2\\-2x_0\\-1\end{pmatrix},$$

即 $q=3-x^2=(3-x_0^2)\beta_1-2x_0\beta_2-\beta_3$.

例 7 已知 $\mathbb{R}^4$ 中的两组基：$\alpha_1=(1,1,2,1)$，$\alpha_2=(0,2,1,2)$，$\alpha_3=(0,0,3,1)$，$\alpha_4=(0,0,0,4)$ 和 $\beta_1=(1,0,0,0)$，$\beta_2=(1,2,0,0)$，$\beta_3=(0,0,1,1)$，$\beta_4=(0,0,-1,1)$，求前一组基到后一组基的过渡矩阵.

解 这里所给的都是行向量，若一个行向量可由一组行向量线性表示，则把它们都转置为列向量后，仍保持原有的线性关系. 因此，若

$$(\beta_1,\beta_2,\beta_3,\beta_4)=(\alpha_1,\alpha_2,\alpha_3,\alpha_4)P,$$

则仍有

$$(\beta_1^T,\beta_2^T,\beta_3^T,\beta_4^T)=(\alpha_1^T,\alpha_2^T,\alpha_3^T,\alpha_4^T)P,$$
$$P=(\alpha_1^T,\alpha_2^T,\alpha_3^T,\alpha_4^T)^{-1}(\beta_1^T,\beta_2^T,\beta_3^T,\beta_4^T)$$