科学版研究生教学丛书

矩阵论

张凯院　徐仲 等编著

科学出版社

科学版研究生教学丛书

矩 阵 论

张凯院 徐 仲 等 编著

科 学 出 版 社

北 京

内 容 简 介

本书分为7章，主要介绍线性空间与线性变换、向量范数与矩阵范数、矩阵分析、矩阵分解、矩阵的特征值估计、广义逆矩阵以及特殊矩阵.各章均配有适量的习题，书后附有部分习题答案或提示.本书内容丰富，论述翔实严谨.突出线性空间的结构和线性变换，并以它们为主线将各章内容贯穿起来；安排了较多的典型例题，便于读者自学；网络教学课件（光盘）、教学辅导书等配套资源丰富.

本书可作为普通高等院校理工科研究生和数学专业高年级本科生的教材，也可供从事科学计算和工程技术的有关人员参考.

图书在版编目(CIP)数据

矩阵论/张凯院等编著．—北京：科学出版社，2013
科学版研究生教学丛书
ISBN 978-7-03-036132-5

Ⅰ.①矩…　Ⅱ.①张…Ⅲ.①矩阵论-研究生-教材
Ⅳ.①O151.21

中国版本图书馆CIP数据核字(2012)第290534号

责任编辑：姚莉丽／责任校对：包志虹
责任印制：阎　磊／封面设计：陈　敬

科学出版社出版
北京东黄城根北街16号
邮政编码：100717
http://www.sciencep.com
北京九天志诚印刷有限公司印刷
科学出版社发行　各地新华书店经销
*
2013年1月第　一　版　开本：720×1000　B5
2013年1月第一次印刷　印张：20
字数：391 000

定价：36.00元

（如有印装质量问题，我社负责调换）

前　言

矩阵论是高等学校和科研院所普遍开设的一门数学基础课.作为数学的一个重要分支,矩阵理论具有极为丰富的内容;作为一种基本工具,矩阵理论在数学学科及其他科学技术领域都有非常广泛的应用.因此,学习和掌握矩阵的基本理论与方法是十分重要的.

本书以程云鹏等编著的研究生教材《矩阵论》(西北工业大学出版社,1989 年第 1 版,1999 年第 2 版,2006 年第 3 版)为蓝本,修改、更新或增加定理与引理论证过程、定义、性质、例题、习题及习题答案或解答等近百个,修改或增加文字叙述段落 30 多处,重组或增加教学内容模块 3 个,前 6 章增加了“本章要点评述”,并对全书进行了较多的文字和叙述方式修改.

本书致力于以近代数学思想、观点和语言统一处理有关题材,并使其内容比传统工科的相应教材有较大的拓宽、充实、更新和提高.突出线性空间的结构和线性变换两大核心内容的地位及训练,并以它们为主线将各章内容贯穿起来,使之成为一个有机的整体.以期达到培养研究生抽象思维和逻辑推理的能力,也使课程体系整体优化.

本书注意揭示数学理论的相关背景,多处安排了不同领域的一些典型范例.力求遵循教育学和教学法的原理,符合教学过程中研究生的认知规律,接近我国研究生的实际水平.虽然在整体上以线性空间和线性变换为主线,但在具体题材的处理上,尽量做到由易到难、由具体到抽象、由特殊到一般.矩阵论中所使用的各种推证方法、公理化定义、抽象化思维、计算技巧及应用等都很有特色,是其他课程所无法替代的,也是提高研究生数学素质不可缺少的一环.在注重基础理论的表述及论证严谨等基础上,为了便于理解,本书对于矩阵理论中教学难度较大的内容作了颇有特色的处理.

本书力求通俗易懂,便于研究生在教师指导下自学.为了使研究生能正确理解概念,掌握运算技巧和解题方法,在书中安排了较多的例题,各章均配有适量的习题.这些题目是经过精心挑选的,其中不少具有新意,书后附有部分习题答案与提示.另外,还有《矩阵论网络教学课件》(光盘)(科学出版社)、《矩阵论辅导讲案》和《矩阵论导教导学导考》(西北工业大学出版社)等辅导资料与本书相配合,给研究生自学提供了极大的方便.

本书题材丰富,不同专业的研究生可根据需要对内容进行取舍.选用本书作教材,理科研究生约需 80 学时,工科研究生删去带“ * ”的内容后约需 60 学时.学习过工科线性代数课程的读者,均可阅读本书.

参加本书编写的有张凯院、徐仲、吕全义、陆全等任课教师，全书由张凯院统稿并负责修改定稿. 作为西北工业大学研究生高水平课程教材之一，本书在编写过程中，得到了西北工业大学理学院、研究生院、教务处、出版社等部门的协助与支持，编者在此表示衷心的感谢.

限于编者水平，书中疏漏和不妥之处在所难免，恳请同行、读者指正.

编　者

2012年3月于西北工业大学

符号说明

符号	含义
$\mathbf{N}_0$	正整数集合
$a \in S$	元素 a 属于集合 S
$a \notin S$	元素 a 不属于集合 S
$S_1 \supset S_2$	集合 S_1 包含集合 S_2
$S_1 \cap S_2$	集合 S_1 与集合 S_2 的交集
$S_1 \cup S_2$	集合 S_1 与集合 S_2 的并集
$S_1 + S_2$	集合 S_1 与集合 S_2 的和
$S_1 \oplus S_2$	集合 S_1 与集合 S_2 的直和
$\sigma: S_1 \to S_2$	σ 是集合 S_1 到集合 S_2 的映射
$\mathbf{R}(\mathbf{C})$	实(复)数集合
$\mathbf{R}^n(\mathbf{C}^n)$	n 维实(复)向量集合
$\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$	$m\times n$ 实(复)矩阵集合
$\mathbf{R}_r^{m\times n}(\mathbf{C}_r^{m\times n})$	秩为 r 的 $m\times n$ 实(复)矩阵集合
$\mathbf{R}_+^n$	n 维非负实向量集合
V^n	n 维线性空间
$\dim V$	线性空间 V 的维数
$W^\perp$	子空间 W 的正交补
$\mathbf{0}$	零向量或线性空间的零元素
$\boldsymbol{e}_i$	第 i 个分量为 1,其余分量为 0 的向量
$\boldsymbol{O}$	零矩阵
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{E}_{ij}$	第 i 行第 j 列元素为 1,其余元素为 0 的矩阵
$\det\boldsymbol{A}$	方阵 $\mathbf{A}$ 的行列式
$\mathrm{tr}\boldsymbol{A}$	方阵 $\mathbf{A}$ 的迹
$\mathrm{adj}\boldsymbol{A}$	方阵 $\mathbf{A}$ 的伴随矩阵
$\boldsymbol{A}^{\mathrm{T}}$	矩阵 $\mathbf{A}$ 的转置
$\boldsymbol{A}^{\mathrm{H}}$	矩阵 $\mathbf{A}$ 的共轭转置
$(\mathbf{A})_{ij}$	矩阵 $\mathbf{A}$ 的第 i 行第 j 列元素
$R(\mathbf{A})$	矩阵 $\mathbf{A}$ 的值域

续表

符号	含义
$N(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的零空间
$\mathrm{rank}\boldsymbol{A}$	矩阵 $\boldsymbol{A}$ 的秩，$\dim R(\boldsymbol{A})$
$n(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的零度，$\dim N(\boldsymbol{A})$
$\rho(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的谱半径
$\mathrm{cond}(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的条件数
$\boldsymbol{J}$	方阵的 Jordan 标准形
$\boldsymbol{J}_i(\lambda_i)$	方阵的第 i 个 Jordan 块
$\Vert\boldsymbol{A}\Vert$	矩阵 $\boldsymbol{A}$ 的任意范数
$\Vert\boldsymbol{A}\Vert_F$	矩阵 $\boldsymbol{A}$ 的 Frobenius 范数
$\Vert\boldsymbol{x}\Vert_p$	向量 $\boldsymbol{x}$ 的 p-范数
$\lambda_i(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的第 i 个特征值
$\sigma_i(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的第 i 个奇异值
$\boldsymbol{A}\sim\boldsymbol{B}$	方阵 $\boldsymbol{A}$ 相似于方阵 $\boldsymbol{B}$
$\boldsymbol{A}\otimes\boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的直积，Kronecker 积
$\overline{\mathrm{vec}}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 按行拉直的列向量
$\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$	以 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为对角元素的对角矩阵
$\mathrm{diag}(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n)$	以 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 为对角元素的准对角矩阵
$(\boldsymbol{x},\boldsymbol{y})$	向量 $\boldsymbol{x}$ 与向量 $\boldsymbol{y}$ 的内积
$\boldsymbol{x}\perp\boldsymbol{y}$	向量 $\boldsymbol{x}$ 与向量 $\boldsymbol{y}$ 正交
$L(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m)$ 或 $\mathrm{span}\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\}$	向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 生成的子空间
$f(\lambda)\mid g(\lambda)$	多项式 $f(\lambda)$ 整除多项式 $g(\lambda)$
$\partial f(\lambda)$	多项式 $f(\lambda)$ 的次数
$\mathrm{Re}(\lambda)$	复数 λ 的实部
$\mathrm{Im}(\lambda)$	复数 λ 的虚部
G_i	方阵 $\boldsymbol{A}$ 的第 i 个 Gerschgorin 圆
$R_i(\boldsymbol{A})$ 或 R_i	方阵 $\boldsymbol{A}$ 的第 i 个 Gerschgorin 圆的半径
$R(\boldsymbol{x})$	方阵 $\boldsymbol{A}$ 的 Rayleigh 商
T_0	零变换
T_e	单位变换
$T_{L,M}$	沿着空间 M 到空间 L 的投影变换
T_L	空间 L 上的正交投影变换
$\boldsymbol{P}_{L,M}$	沿着空间 M 到空间 L 的投影矩阵

续表

符号	含义
$\boldsymbol{P}_L$	空间 L 上的正交投影矩阵
$\boldsymbol{A}^+$	矩阵 $\boldsymbol{A}$ 的 Moore-Penrose 逆
$\boldsymbol{A}^{(i,j,\cdots,l)}$	矩阵 $\boldsymbol{A}$ 的$\{i,j,\cdots,l\}$-逆
$A\{i,j,\cdots,l\}$	矩阵 $\boldsymbol{A}$ 的$\{i,j,\cdots,l\}$-逆的集合
$\boldsymbol{A}^{(d)}$	矩阵 $\boldsymbol{A}$ 的 Drazin 逆
$\boldsymbol{A}^{\#}$	矩阵 $\boldsymbol{A}$ 的群逆
$\boldsymbol{L}_{-1}$	反周期 Jacobi 矩阵
$\boldsymbol{A}>\boldsymbol{B}$	$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 Hermite 矩阵，且 $\boldsymbol{A}-\boldsymbol{B}$ 为正定矩阵
$\boldsymbol{A}\geqslant\boldsymbol{B}$	$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 Hermite 矩阵，且 $\boldsymbol{A}-\boldsymbol{B}$ 为非负定矩阵
$\boldsymbol{A}>\boldsymbol{O}(\boldsymbol{x}>\boldsymbol{0})$	矩阵 $\boldsymbol{A}$(向量 $\boldsymbol{x}$)的每个元素都是正数
$\boldsymbol{A}\geqq\boldsymbol{O}(\boldsymbol{x}\geqq\boldsymbol{0})$	矩阵 $\boldsymbol{A}$(向量 $\boldsymbol{x}$)的每个元素都是非负实数
$\boldsymbol{A}\geqslant\boldsymbol{O}(\boldsymbol{x}\geqslant\boldsymbol{0})$	非零矩阵 $\boldsymbol{A}$(向量 $\boldsymbol{x}$)的每个元素都是非负实数
V. L. 稳定	Volterra-Lyapunov 稳定
P_0^+	主子式的值非负，且同阶主子式中至少有一个为正值的矩阵集合
命题 $A\Leftrightarrow$命题 B	命题 A 等价于命题 B
命题 $A\Rightarrow$命题 B	由命题 A 可推导出命题 B
$\forall$	任意

目　录

前言
符号说明
第 1 章　线性空间与线性变换 ······ 1
1.1　线性空间 ······ 1
1.2　线性变换及其矩阵 ······ 18
1.3　两个特殊的线性空间 ······ 54
本章要点评述 ······ 73
第 2 章　范数理论及其应用 ······ 75
2.1　向量范数及其性质 ······ 75
2.2　矩阵范数 ······ 83
2.3　范数的一些应用 ······ 90
本章要点评述 ······ 94
第 3 章　矩阵分析及其应用 ······ 95
3.1　矩阵序列 ······ 95
3.2　矩阵级数 ······ 97
3.3　矩阵函数 ······ 103
3.4　函数矩阵的微分和积分 ······ 113
3.5　矩阵函数的一些应用 ······ 119
本章要点评述 ······ 123
第 4 章　矩阵分解 ······ 125
4.1　Gauss 消去法与矩阵的三角分解 ······ 125
4.2　矩阵的 QR 分解 ······ 137
4.3　矩阵的满秩分解 ······ 153
4.4　矩阵的奇异值分解 ······ 157
本章要点评述 ······ 163
第 5 章　特征值的估计及对称矩阵的极性 ······ 164
5.1　特征值的估计 ······ 164
5.2　广义特征值问题 ······ 183
5.3　对称矩阵特征值的极性 ······ 184
5.4　矩阵的直积及其应用 ······ 192

本章要点评述 …… 201
第 6 章　广义逆矩阵 …… 203
6.1　广义逆矩阵的概念与性质 …… 203
*6.2　投影矩阵与 Moore 逆 …… 213
*6.3　广义逆矩阵的计算方法 …… 218
6.4　广义逆矩阵与线性方程组的求解 …… 233
*6.5　约束广义逆和加权广义逆 …… 241
*6.6　Drazin 广义逆 …… 245
本章要点评述 …… 252
***第 7 章　若干特殊矩阵类介绍** …… 254
7.1　正定矩阵与正稳定矩阵 …… 255
7.2　对角占优矩阵 …… 263
7.3　非负矩阵 …… 270
7.4　M 矩阵与广义 M 矩阵 …… 274
7.5　Toeplitz 矩阵及其有关矩阵 …… 282
7.6　其他特殊矩阵 …… 289
部分习题答案或提示 …… 297
参考文献 …… 310

第1章 线性空间与线性变换

线性空间与线性变换是矩阵理论中的两个重要的概念. 本章先简要地论述这两个概念及其有关理论,然后再讨论两个特殊的线性空间,这就是 Euclid 空间和酉空间. 所有论述是在假定读者已经具备了 n 维向量空间的理论、矩阵的初步运算、线性方程组的理论和二次型的有关知识的基础上进行的.

1.1 线性空间

为了统一研究在数学、力学及其他学科中遇到的规定了线性运算的不同集合的公共性质,需要引入线性空间的概念. 线性空间是对 n 维向量空间的推广.

1.1.1 集合与映射

集合是数学中最基本的概念之一. 所谓集合是指能够作为整体看待的一堆东西. 例如,由一些数(有限个或无限个)组成的集合,称为数集合或数集;一个线性代数方程组解的全体组成一个集合,称为解集合;一个已知半径和圆心的开圆内的所有点组成一个集合,称为点集合或点集;等等. 组成集合的事物称为这个集合的元素. 如果用 S 表示集合,a 表示 S 的元素,常用记号 $a\in S$ 表示 a 是 S 的元素,读为 a 属于 S;用记号 $a\notin S$ 表示 a 不是 S 的元素,读为 a 不属于 S.

因为一个集合由它的元素组成,所以给出一个集合的方式不外乎是列举出它的全部元素或者给出这个集合中元素所具有的特征性质. 例如,由数 1,2,3 组成的集合 N,可记为 $N=\{1,2,3\}$,这就是用列举出其全部元素的方式;但适合方程

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

的全部点组成的点集 P,因其元素有无穷多个,不可能全部列举出来,就用其元素所具有的特征性质的方式把它记为

$$P=\left\{(x,y)\left|\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\right.\right\}.$$

一般地,M 是具有某些性质的全部元素 a 所组成的集合时,可记为

$$M=\{a|a\text{ 具有的性质}\}.$$

全体正整数构成的集合,称为正整数集,常记为

$$\mathbf{N}_0=\{n|n\text{ 是正整数}\}.$$

不包含任何元素的集合称为**空集合**,常记为$\varnothing$. 例如,一个无解的线性代数方程

组的解集合就是一个空集合. 空集合在集合运算中所起的作用,类似于数零在数的运算中所起的作用.

如果集合 S_1 的元素全是集合 S_2 的元素,即由 $a \in S_1$ 可以推出 $a \in S_2$,那么就称 S_1 为 S_2 的子集合,记为 $S_1 \subset S_2$ 或 $S_2 \supset S_1$. 例如,全体偶数组成的集合是全体整数组成的集合(称为**整数集**)的子集合. 规定空集合是任一集合的子集合. 按定义,每个集合都是它自身的子集合. 把集合的这两个特殊子集合,统称为其当然子集合或假子集合,而把它的其余子集合统称为其非当然子集合或真子集合.

如果两个集合 S_1 与 S_2 含有完全相同的元素,即 $a \in S_1$,当且仅当 $a \in S_2$ 时,那么就称它们**相等**,记为 $S_1 = S_2$. 如果集合 S_1 与 S_2 同时满足 $S_1 \subset S_2$ 与 $S_2 \subset S_1$,那么 S_1 和 S_2 相等.

把既属于集合 S_1,又属于集合 S_2 的全体元素所组成的集合称为 S_1 与 S_2 的**交**,记为

$$S_1 \cap S_2 = \{x \mid x \in S_1 \text{ 且 } x \in S_2\}.$$

例如,圆 $(x-1)^2+y^2=1$ 包含的所有点组成的点集与圆 $(x-2)^2+y^2=1$ 包含的所有点组成的点集的交,就是两个圆公共部分所有点组成的点集. 两集合的交具有下面的关系:

$$S_1 \cap S_2 \subset S_1, \quad S_1 \cap S_2 \subset S_2.$$

属于集合 S_1,或属于集合 S_2 的全体元素组成的集合,称为 S_1 与 S_2 的**并**,记为

$$S_1 \cup S_2 = \{x \mid x \in S_1 \text{ 或 } x \in S_2\}.$$

两集合 S_1 与 S_2 的并满足关系 $S_1 \cup S_2 \supset S_1, S_1 \cup S_2 \supset S_2$.

集合 S_1 与集合 S_2 的**和集**是指如下的集合:

$$\{x+y \mid x \in S_1, y \in S_2\},$$

常用记号 S_1+S_2 来表示,于是有

$$S_1+S_2 = \{x+y \mid x \in S_1, y \in S_2\}.$$

需要指出,两集合的和集概念不同于它们的并集概念,例如

$$\{1,2,3\} \cup \{2,3,4\} = \{1,2,3,4\}, \quad \{1,2,3\}+\{2,3,4\} = \{3,4,5,6,7\}.$$

某些数集(含非零的数),如果其中任意两个数的和、差、积、商(除数不为零)仍在该数集中(即关于四则运算封闭的数集),那么称该数集为数域. 例如,实数集关于四则运算封闭,因此它形成一个数域,称其为实数域,记为 **R**. 同样,复数集也形成一个数域,称其为复数域,记为 **C**. 读者可以验证有理数集形成一个有理数域,但奇数集不能形成数域,偶数集也不能形成数域.

下面简要介绍矩阵论中关于集合间的映射这一重要概念.

设 S 与 S' 是两个集合. 所谓集合 S 到集合 S' 的一个**映射**或**映照**,是指一个法则(或规则)$\sigma: S \to S'$,它使 S 中每一个元素 a 都有 S' 中一个确定的元素 a' 与之对应(a' 未必取遍集合 S'),记为

$$\sigma(a) = a' \quad \text{或} \quad a \to a' (= \sigma(a)),$$

a'称为a在映射σ下的**象**，而a称为a'在映射σ下的一个**原象**（或**象源**）.

S到S自身的映射，有时也称为S到自身的一个**变换**. 这种特殊的映射，在矩阵论中也是经常出现的.

例如，设S是数域K[①]上全体n阶矩阵$\boldsymbol{A}$的集合($n\geqslant 2$)，定义

$$\sigma_1(\boldsymbol{A})=\det\boldsymbol{A} \quad (\boldsymbol{A}\in S),$$

则有$\sigma_1: S\to K$，即σ_1是S到K的一个映射；如果定义

$$\sigma_2(\boldsymbol{A})=(\det\boldsymbol{A})\boldsymbol{I} \quad (\boldsymbol{A}\in S),$$

这里$\boldsymbol{I}$是n阶单位矩阵，则有$\sigma_2: S\to S$，即σ_2是S到自身的一个变换.

设P_n是所有次数不超过n的实系数多项式的集合，定义

$$\sigma(f(t))=f'(t) \quad (f(t)\in P_n),$$

则σ是P_n到自身的一个变换（实为求导运算）.

又如，对于指数函数$y=\mathrm{e}^x$而言，它可视为$\mathbf{R}$到自身的映射. 一般地，定义在$\mathbf{R}$上的实函数$y=f(x)$，都可视为$\mathbf{R}$到自身的一种特殊映射.

关于映射，还可定义它的运算如下.

设σ_1与σ_2都是集合S到S'的映射，如果对于每个$a\in S$，都有$\sigma_1(a)=\sigma_2(a)$，则称映射σ_1与σ_2相等，记为$\sigma_1=\sigma_2$.

设σ,τ依次是集合S到S_1，S_1到S_2的映射，映射的乘积$\tau\sigma$定义如下：

$$(\tau\sigma)(a)=\tau(\sigma(a)) \quad (a\in S),$$

此即相继施行映射σ和τ的结果，$\tau\sigma$是S到S_2的一个映射.

设σ,τ,μ依次为集合S到S_1，S_1到S_2，S_2到S_3的映射，可以证明映射的乘积满足结合律，但不满足交换律，即分别有

$$(\mu\tau)\sigma=\mu(\tau\sigma), \quad \tau\sigma\neq\sigma\tau.$$

1.1.2 线性空间及其性质

线性空间是线性代数最基本的概念之一，也是学习矩阵论的重要基础. 线性空间的概念，是某类事物从量的方面的一个抽象. 为了便于理解这个抽象概念，先看如下几个熟知的例子.

例 1.1 所有n维实向量的集合$\mathbf{R}^n$（或复向量的集合$\mathbf{C}^n$），对n维向量的加法及数乘n维向量的运算是封闭的. 加法运算还满足交换律与结合律；数乘向量的运算满足分配律与结合律.

例 1.2 在集合P_n中，按通常意义定义多项式加法及实数与多项式乘法，若$f(t)\in P_n$，$g(t)\in P_n$，则$f(t)+g(t)\in P_n$；若$k\in\mathbf{R}$，则$kf(t)\in P_n$. 因此，P_n对这两种运算是封闭的. 容易验证，对P_n的这两种运算，如例 1.1 中所论的诸算律也成立.

---

① 数域K表示一般的数域，它可以是$\mathbf{R}$，可以是$\mathbf{C}$，也可以是其他数域.

例 1.3　常系数二阶齐次线性微分方程

$$y''-3y'+2y=0$$

解的集合 D,对于函数加法及数与函数乘法有:若 $y_1,y_2\in D$,则 $y_1+y_2\in D$;若 $k\in\mathbf{R}$,则 $ky_1\in D$. 因此,D 关于这两种运算是封闭的,且满足如例 1.1 中所论的诸算律.

例 1.4　在所有 n 阶实矩阵的集合 $\mathbf{R}^{n\times n}$(或复矩阵的集合 $\mathbf{C}^{n\times n}$)中,如果 $\boldsymbol{A},\boldsymbol{B}\in\mathbf{R}^{n\times n}$(或 $\mathbf{C}^{n\times n}$),则 $\boldsymbol{A}+\boldsymbol{B}\in\mathbf{R}^{n\times n}$(或 $\mathbf{C}^{n\times n}$);如果 $k\in\mathbf{R}$(或 $\mathbf{C}$),则 $k\boldsymbol{A}\in\mathbf{R}^{n\times n}$(或 $\mathbf{C}^{n\times n}$). 因此,集合对于这两种运算是封闭的,且满足如例 1.1 中所论的诸算律.

此外,在数学、力学及其他学科中,还有大量如例 1.1～例 1.4 这样的集合. 因此,有必要不考虑集合的对象,抽去它们的具体内容的含义,来研究这类集合的公共性质,并引进一个概括性名词,于是就有如下的线性空间的概念.

定义 1.1　设 V 是一个非空集合,它的元素用 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$ 等表示,并称之为**向量**;K 是一个数域,它的元素用 k,l,m 等表示. 如果 V 满足下列条件:

(1) 在 V 中定义一个加法运算,即当 $\boldsymbol{x},\boldsymbol{y}\in V$ 时,有唯一的和 $\boldsymbol{x}+\boldsymbol{y}\in V$,且加法运算满足下列性质:

1) 交换律　$\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$;

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

3) 存在**零元素** $\mathbf{0}$,使 $\boldsymbol{x}+\mathbf{0}=\boldsymbol{x}$;

4) 存在**负元素**,即对任一向量 $\boldsymbol{x}\in V$,存在向量 $\boldsymbol{y}\in V$,使 $\boldsymbol{x}+\boldsymbol{y}=\mathbf{0}$,则称 $\boldsymbol{y}$ 为 $\boldsymbol{x}$ 的负元素,记为 $-\boldsymbol{x}$,于是有 $\boldsymbol{x}+(-\boldsymbol{x})=\mathbf{0}$;

(2) 在 V 中定义数乘(数与向量的乘法)运算,即当 $\boldsymbol{x}\in V,k\in K$ 时,有唯一的 $k\boldsymbol{x}\in V$,且数乘运算满足下列性质:

5) 数因子分配律　$k(\boldsymbol{x}+\boldsymbol{y})=k\boldsymbol{x}+k\boldsymbol{y}$;

6) 分配律　$(k+l)\boldsymbol{x}=k\boldsymbol{x}+l\boldsymbol{x}$;

7) 结合律　$k(l\boldsymbol{x})=(kl)\boldsymbol{x}$;

8) $1\boldsymbol{x}=\boldsymbol{x}$.

则称 V 为数域 K 上的**线性空间或向量空间**.

V 中所定义的加法及数乘运算统称为 V 的线性运算. 在不致产生混淆时,将数域 K 上的线性空间简称为线性空间. 数 k 与向量 $\boldsymbol{x}$ 的乘积 $k\boldsymbol{x}$ 也可以写成 $\boldsymbol{x}k$.

需要指出,不管 V 的元素如何,当 K 为实数域 $\mathbf{R}$ 时,称 V 为**实线性空间**;当 K 为复数域 $\mathbf{C}$ 时,就称 V 为**复线性空间**.

例 1.1 中的 $\mathbf{R}^n$ 形成实线性空间,$\mathbf{C}^n$ 形成复线性空间;例 1.2 与例 1.3 中的集合,在其各自的加法和数乘运算的定义下,都形成实线性空间;例 1.4 中的 $\mathbf{R}^{n\times n}$ 形成实线性空间,$\mathbf{C}^{n\times n}$ 形成复线性空间. 特别地,将例 1.2 所给的线性空间称为**多项式空间**,将例 1.4 所给的线性空间称为**矩阵空间**.

例 1.5　设 $\mathbf{R}^+$ 为所有正实数构成的集合,其加法与数乘运算分别定义为

$$m\oplus n=mn,\quad k\circ m=m^k,$$

证明 $\mathbf{R}^{+}$ 是 $\mathbf{R}$ 上的线性空间.

证 设 $m,n\in\mathbf{R}^{+}$,$k\in\mathbf{R}$,则有

$$m\oplus n=mn\in\mathbf{R}^{+},\quad k\circ m=m^{k}\in\mathbf{R}^{+},$$

即 $\mathbf{R}^{+}$ 对所定义的加法运算“$\oplus$”与数乘运算“$\circ$”是封闭的,且有:

(1) $m\oplus n=mn=nm=n\oplus m$;

(2) $(m\oplus n)\oplus p=(mn)\oplus p=mnp=m\oplus(np)=m\oplus(n\oplus p)$;

(3) 1 是 $\mathbf{R}^{+}$ 的零元素,因为 $m\oplus 1=m\times 1=m$;

(4) $\dfrac{1}{m}$是 m 的负元素,因为 $m\oplus\dfrac{1}{m}=m\times\dfrac{1}{m}=1$;

(5) $k\circ(m\oplus n)=k\circ(mn)=(mn)^{k}=m^{k}n^{k}=(k\circ m)\oplus(k\circ n)$;

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

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

(8) $1\circ m=m^{1}=m$.

故 $\mathbf{R}^{+}$ 是实线性空间.

定理 1.1 线性空间 V 有唯一的零元素,任一元素也有唯一的负元素.

证 设 $\mathbf{0}_1,\mathbf{0}_2$ 是 V 的两个零元素,考虑和 $\mathbf{0}_1+\mathbf{0}_2$. 由于 $\mathbf{0}_1$ 是 V 的零元素,因此

$$\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_2+\mathbf{0}_1=\mathbf{0}_2;$$

又由于 $\mathbf{0}_2$ 也是 V 的零元素,因此 $\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_1$. 从而有

$$\mathbf{0}_1=\mathbf{0}_1+\mathbf{0}_2=\mathbf{0}_2.$$

这就证明了零元素的唯一性.

为了证明负元素的唯一性,设元素 $\boldsymbol{x}$ 有两个负元素 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$,即有

$$\boldsymbol{x}+\boldsymbol{x}_1=\mathbf{0},\quad \boldsymbol{x}+\boldsymbol{x}_2=\mathbf{0}.$$

于是

$$\boldsymbol{x}_1=\boldsymbol{x}_1+\mathbf{0}=\boldsymbol{x}_1+(\boldsymbol{x}+\boldsymbol{x}_2)=(\boldsymbol{x}_1+\boldsymbol{x})+\boldsymbol{x}_2=\mathbf{0}+\boldsymbol{x}_2=\boldsymbol{x}_2.$$ 证毕

利用负元素,定义 V 中向量的减法运算如下:

$$\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{x}+(-\boldsymbol{y}).$$

可以证明,若 $\boldsymbol{x}\in V,k\in K$,则 $0\boldsymbol{x}=\mathbf{0},k\mathbf{0}=\mathbf{0},(-1)\boldsymbol{x}=-\boldsymbol{x}$.

同 n 维线性空间 $\mathbf{R}^{n}$ 中向量组的线性相关性一样,如果 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 为线性空间 V 中的 m(有限正整数)个向量,$\boldsymbol{x}\in V$,且存在数域 K 中一组数 $c_1,c_2,\cdots,c_m$,使得

$$\boldsymbol{x}=c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_m\boldsymbol{x}_m.\tag{1.1.1}$$

则称 $\boldsymbol{x}$ 为向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的**线性组合**,有时也称向量 $\boldsymbol{x}$ 可由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ **线性表示**. 如果式(1.1.1)中 $c_1,c_2,\cdots,c_m$ 不全为零,且使

$$c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2+\cdots+c_m\boldsymbol{x}_m=\mathbf{0},\tag{1.1.2}$$

则称向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ **线性相关**,否则称其为**线性无关**(即式(1.1.2)仅当 $c_1,c_2,\cdots,c_m$ 全为零时才成立,称 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 线性无关). 注意到上述概念都只涉及加法与数乘运算,与向量本身的属性无直接关系,因此对于 $\mathbf{R}^{n}$ 中向量成立的相应结论可以照

搬到一般的线性空间中来. 对 V 中线性无关的向量组所含向量的最大个数作如下的定义.

定义 1.2　线性空间 V 中线性无关向量组所含向量最大个数称为 V 的**维数**. 若 n 是具有这个性质的正整数,则称 V 的维数是 n,记为 $\dim V=n$.

维数是 n 的线性空间称为数域 K 上的 n 维线性空间,记为 V^n. 当 $n=+\infty$时,称为无限维线性空间.

譬如,例 1.3 中的集合 D 仅有两个线性无关的向量 $y_1=\mathrm{e}^x, y_2=\mathrm{e}^{2x}$,且 D 中任一向量 y 都可以由 y_1 和 y_2 线性表示,即有 $y=c_1\mathrm{e}^x+c_2\mathrm{e}^{2x}$,故 $\dim D=2$,即 D 是 $\mathbf{R}$ 上的二维线性空间. 同样,例 1.4 中的 $\mathbf{R}^{n\times n}$是 $\mathbf{R}$ 上的 n^2 维线性空间. 这是因为 $\mathbf{R}^{n\times n}$ 的任一向量

$$\boldsymbol{A}=(a_{ij})_{n\times n}=\sum_{i,j=1}^{n}a_{ij}\boldsymbol{E}_{ij}\ ,$$

其中 $\boldsymbol{E}_{ij}$ 表示第 i 行第 j 列的元素为 1,其余元素都为 0 的 n 阶矩阵,且易验证 $\boldsymbol{E}_{ij}(i,j=1,2,\cdots,n)$线性无关,故 $\dim\mathbf{R}^{n\times n}=n^2$. 又如,所有实系数多项式的集合,在通常的多项式加法及数乘多项式的运算下形成的实线性空间是无限维的,这是因为对于任意正整数 N,都有 N 个线性无关的向量 $1,t,t^2,\cdots,t^{N-1}$.

容易验证,$1,t,t^2,\cdots,t^n$ 是例 1.2 中的线性空间 P_n 的一个最大线性无关组,且 $\dim P_n=n+1$.

无限维线性空间是一个专门研究的对象,它与有限维线性空间有较大的差别. 本书主要讨论有限维线性空间.

1.1.3　线性空间的基与坐标

在解析几何中,为了借助于数量运算以实现向量的运算,必须引进向量的坐标. 对有限维线性空间,坐标同样是一个有力的工具,现引入如下.

定义 1.3　设 V 是数域 K 上的线性空间,$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r(r\geqslant1)$是 V 中的任意 r 个向量,如果它满足

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

(2) V 中任一向量 $\boldsymbol{x}$ 都是 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 的线性组合.

则称 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 为 V 的一个**基**或**基底**,并称 $\boldsymbol{x}_i(i=1,2,\cdots,r)$为**基向量**.

由定义 1.3 可见,线性空间的维数是其基中所含向量的个数.

例如,e^x 与 e^{2x}就是线性空间 D 的一个基;而 $\boldsymbol{E}_{ij}(i,j=1,2,\cdots,n)$就是线性空间 $\mathbf{R}^{n\times n}$的一个基.

根据定义 1.3,容易看出:齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系中所含的向量,就是其解空间的一个基.

需要指出,一个线性空间的基不是唯一的. 例如,n 维向量组

$$\begin{cases} e_1=(1,0,\cdots,0) \\ e_2=(0,1,\cdots,0) \\ \cdots\cdots \\ e_n=(0,0,\cdots,1) \end{cases} \quad 和 \quad \begin{cases} y_1=(1,1,\cdots,1,1) \\ y_2=(0,1,\cdots,1,1) \\ \cdots\cdots \\ y_n=(0,0,\cdots,0,1) \end{cases} \tag{1.1.3}$$

都是线性空间 $\mathbf{R}^n$ 的基. 这是因为

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0 \quad 且 \quad \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix} \neq 0,$$

从而它们各自都线性无关,而且对任一向量 $\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbf{R}^n$,分别有

$$\boldsymbol{x}=\xi_1\boldsymbol{e}_1+\xi_2\boldsymbol{e}_2+\cdots+\xi_n\boldsymbol{e}_n, \quad \boldsymbol{x}=\xi_1\boldsymbol{y}_1+(\xi_2-\xi_1)\boldsymbol{y}_2+\cdots+(\xi_n-\xi_{n-1})\boldsymbol{y}_n,$$

于是由定义 1.3 知上面论断成立.

定义 1.4 称线性空间 V^n 的一个基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 为 V^n 的一个**坐标系**. 设向量 $\boldsymbol{x}\in V^n$,它在该基下的线性表示式为

$$\boldsymbol{x}=\xi_1\boldsymbol{x}_1+\xi_2\boldsymbol{x}_2+\cdots+\xi_n\boldsymbol{x}_n, \tag{1.1.4}$$

则称 $\xi_1,\xi_2,\cdots,\xi_n$ 为 $\boldsymbol{x}$ 在该坐标系中的**坐标**或**分量**,记为

$$(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}. \tag{1.1.5}$$

必须指出,在不同的坐标系(或基)中,同一向量的坐标一般是不同的. 例如,$\mathbf{R}^n$ 的任一向量$(\xi_1,\xi_2,\cdots,\xi_n)$在式(1.1.3)的第一基中的坐标是$(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$,但在第二基中的坐标却是$(\xi_1,\xi_2-\xi_1,\cdots,\xi_n-\xi_{n-1})^{\mathrm{T}}$. 然而却有下面的定理.

定理 1.2 设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V^n 的一个基,$\boldsymbol{x}\in V^n$,则 $\boldsymbol{x}$ 可唯一地表示成 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 的线性组合.

证 设 $\boldsymbol{x}$ 可由 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 线性表示为

$$\boldsymbol{x}=\xi_1\boldsymbol{x}_1+\xi_2\boldsymbol{x}_2+\cdots+\xi_n\boldsymbol{x}_n, \quad \boldsymbol{x}=\xi_1'\boldsymbol{x}_1+\xi_2'\boldsymbol{x}_2+\cdots+\xi_n'\boldsymbol{x}_n,$$

相减得

$$(\xi_1'-\xi_1)\boldsymbol{x}_1+(\xi_2'-\xi_2)\boldsymbol{x}_2+\cdots+(\xi_n'-\xi_n)\boldsymbol{x}_n=\mathbf{0}.$$

由于 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V^n 的基,从而线性无关,故有

$$\xi_i'=\xi_i \quad (i=1,2,\cdots,n).$$

证毕

需要指出,使用表达式(1.1.5)时,应该指明在 V^n 中所选用的坐标系,这是因为 V^n 的基不唯一,而向量的坐标随基的不同而不同. 若 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 为 V^n 的一个基,则在这个坐标系中,基向量 $\boldsymbol{x}_i(i=1,2,\cdots,n)$正好表示为

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

它的坐标为

$$\boldsymbol{e}_i=(0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}} \quad (i=1,2,\cdots,n).$$

由此可见,假如在定义 1.1 中没有 $1\boldsymbol{x}=\boldsymbol{x}$ 的规定,就无法把 $\boldsymbol{x}_i$ 写成 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 的线性组合,从而基、维数等概念都没有了.