

矩阵论

尚有林 编著



科学出版社

矩 阵 论

尚有林 编著



科 学 出 版 社

北 京

内 容 简 介

本书全面系统地介绍了矩阵的主要理论、方法及应用. 全书共分 9 章, 包括线性空间与线性变换、内积空间、矩阵的标准形、矩阵的分解、特征值的估计、矩阵分析、矩阵的应用、矩阵的广义逆、非负矩阵. 本书取材广泛, 理论与应用密切结合, 参考许多矩阵理论在实际生活中的问题, 特别是工程技术中应用的文献, 帮助读者学会如何使用矩阵这一重要数学工具, 灵活解决科学和工程技术中的实际问题.

本书适合于需要矩阵知识比较多、比较深刻的理科(数学、物理、力学)和信息科学与技术(电子、通信、自动控制、计算机、系统工程、模式识别、信号处理等)各学科教师以及工程技术人员阅读、参考, 也可作为工科研究生的教材.

图书在版编目(CIP)数据

矩阵论/尚有林编著. —北京: 科学出版社, 2013

ISBN 978-7-03-038575-8

I. ①矩… II. ①尚… III. ①矩阵论 IV. ①O151.21

中国版本图书馆 CIP 数据核字(2013)第 214914 号

责任编辑: 胡海霞 昌 盛 / 责任校对: 彭 涛

责任印制: 阎 磊 / 封面设计: 迷底书装

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>

文林印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2013 年 9 月第 一 版 开本: 720×1000 B5

2013 年 9 月第一次印刷 印张: 15

字数: 302 000

定价: 34.00 元

(如有印装质量问题, 我社负责调换)

前 言

矩阵是许多理工科的重要数学工具. 就其本身的研究而言, 矩阵理论也是极富创造性的领域, 其创造性又极大地推动和丰富其他众多学科的发展: 许多新的理论、方法、技术的诞生和发展就是矩阵理论的创造性与推广的结果. 可以毫不夸张地说, 矩阵理论在数学、物理、力学、信号与信息处理、电子、通信、系统科学、控制论、模式识别、土木、电机、航空和航天等众多学科中是最富创造性和灵活性, 并起着不可替代作用的数学工具.

一方面, 作者长期从事运筹学与最优化、系统科学、微分方程等领域的科学研究, 阅读和参考了大量矩阵理论在这些学科中应用的文献, 深刻感受到矩阵理论在科学研究中所起的重要作用; 另一方面, 在长期的研究生课程教学中, 作者对工科各专业研究生的矩阵理论的不足与缺乏印象深刻. 矩阵理论本身的重要性以及作者在教学与研究中的体会激发了编著本书的意愿. 本书是作者在多年研究和教学的基础上完成的, 并参考大量科学文献, 突出理论与应用并重的特点.

本书加大了选材的广度与深度, 试图在以下三个方面形成特点.

- (1) 本书利用较多的篇幅介绍矩阵的一些应用比较广泛的理论, 如 Rayleigh 商、矩阵的微分与积分等;
- (2) 本书突出矩阵理论与各类科学与工程实际问题应用的密切结合, 书中参考了许多应用实例的文献, 并给出矩阵理论的较多应用;
- (3) 本书重点突出, 对比较基础和重要的理论及应用花费了大量的篇幅, 而对某些应用较少的理论仅给出基本介绍, 便于读者有选择性和侧重性的阅读.

本书对于长期从事理工科各学科教学的教师及工程技术人员有一定的帮助, 同时也可作为工科研究生的教材, 可根据实际需要安排 40~60 学时, 讲授的基本内容大体上就是本书前七章的内容(带 * 内容除外).

本书得到了河南科技大学学术著作出版基金、国家自然科学基金(No. 10971053)以及河南科技大学研究生教改创新基金的资助. 在编著过程中, 得到了河南科技大学数学与统计学院老师们的极大关心和支持, 尤其是李培峦博士、李小申教授、程东明教授对本书初稿提出了很多的修改建议. 在本书稿完成过程中, 河南工业大学的张宏伟教授、河南理工大学的成军祥教授、华北水利水电大学的罗党教授、刘法贵教授给予了热情指导, 并得到了河南科技大学学科建设处、研究生处等有关领导和同志们的大力支持, 在此一并感谢. 感谢使用本书的老师们的批评和

鼓励,感谢本书的责任编辑在编印本书时的出色工作.

虽然作者竭力而为,但由于水平有限,书中难免存在不足之处,请广大读者不吝赐教.

尚有林

2012 年 12 月

河南科技大学

目 录

前言

第 1 章 线性空间与线性变换	1
1.1 线性空间的概念	1
1.2 基变换与坐标变换	3
1.3 子空间与维数定理	9
1.4 线性空间的同构	15
1.5 线性变换及其矩阵	17
1.6 不变子空间	22
* 1.7 随机向量	26
习题 1	34
第 2 章 内积空间	37
2.1 欧氏空间与酉空间	37
2.2 正交变换与酉变换	43
2.3 内积空间的同构	46
2.4 点到子空间的距离与最小二乘法	47
2.5 Hermite 矩阵	48
* 2.6 Rayleigh 商	53
习题 2	57
第 3 章 矩阵的标准形	60
3.1 矩阵的相似对角形	60
3.2 矩阵的 Jordan 标准形	65
3.3 矩阵多项式与最小多项式	73
习题 3	83
第 4 章 矩阵的分解	86
4.1 矩阵的三角分解	86
4.2 矩阵的满秩分解	93
4.3 矩阵的 Schur 分解	97
4.4 矩阵的奇异值分解	100
4.5 方阵的极分解	109
4.6 矩阵的谱分解	111
习题 4	118

第 5 章 特征值的估计	121
5.1 特征值的界的估计	121
5.2 圆盘定理	124
5.3 谱半径的估计	125
习题 5	127
第 6 章 矩阵分析	129
6.1 向量范数	129
6.2 矩阵范数	131
6.3 向量序列和矩阵序列的极限	141
6.4 矩阵幂级数	144
6.5 矩阵函数	147
* 6.6 矩阵的微分	152
* 6.7 矩阵的积分	164
6.8 常用矩阵函数的性质	167
习题 6	170
第 7 章 矩阵的应用	173
* 7.1 Rayleigh 商的应用	173
7.2 矩阵奇异值分解的应用	177
7.3 矩阵函数在微分方程组中的应用	185
7.4 线性系统的能控性与能观测性	189
7.5 非经典阻尼系统的求解	192
习题 7	197
第 8 章 矩阵的广义逆	198
8.1 矩阵的左逆与右逆	198
8.2 减号广义逆与线性方程组的解	200
8.3 Moore-Penrose 广义逆	205
习题 8	211
第 9 章 非负矩阵	213
9.1 正矩阵	213
9.2 非负矩阵	217
9.3 素矩阵	221
9.4 随机矩阵	223
9.5 M 矩阵	225
参考文献	231

第 1 章 线性空间与线性变换

矩阵是处理有限维空间形式和数量关系的重要工具,线性空间与线性变换是其中的基本研究对象.本章在线性代数的基础上,将上述概念推广到向量空间 $\mathbf{R}^n$,建立线性空间的概念、定义线性变换、介绍其中的基本理论及其矩阵方法,这是学习矩阵分析及其应用的入门知识.

1.1 线性空间的概念

人们谈论问题,往往都是就一定“范围”来说的,脱离了这个“范围”,就难以讲清楚了,甚至只能在某个“范围”内才能提出或研究某种问题.明白了这一点,就比较容易理解我们引入数域及线性空间的目的.

由所有有理数组成的集合具有这样的性质:这个集合中任意两数的和、差、积、商(除数不为零)仍是该集中的数,这个集合用 $\mathbf{Q}$ 表示.类似地,由所有实数构成的集合 $\mathbf{R}$,以及由所有复数构成的集合 $\mathbf{C}$ 也都具有这一性质,这三个集合的包含关系为

$$\mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}.$$

因此,我们说“一个复数”时,自然包括这个数可能是有理数或实数这两种特殊情况在内.

在引入线性空间这一重要概念之前,首先要给出数域的概念.

如果复数的一个非空集合 P 含有非零的数,且其中任意两数的和、差、积、商(除数不为零)仍属于该集合,则称数集 P 为一个数域.于是上述集合 $\mathbf{Q}, \mathbf{R}, \mathbf{C}$ 都是数域,分别称为有理数域、实数域及复数域.又如集合

$$\mathbf{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbf{Q}\}$$

也构成一数域,请读者加以验证.但是,由所有整数组成的集合 $\mathbf{Z}$ 是不构成数域的.

数域有个简单性质,即所有的数域都包含有理数域.特别地,每个数域都包含 0 和 1.现在我们可以给出线性空间的定义了.

定义 1.1 设 V 是一个非空集合, P 是一数域.如果

(1) 在集合 V 上定义了一个二元运算(通常称为加法),即 V 中任意两个元素 α, β , 经过这个运算后所得到的结果,仍是集合 V 中的唯一确定的元素,这元素称为 α 与 β 的和,并记作 $\alpha + \beta$;

(2) 在数域 P 与集合 V 的元素之间还定义了一种运算,称为数量乘法,即对 P 中任意数 k 与 V 中任意元素 α ,经过这一运算后所得的结果仍为 V 中一个唯一确定的元素,称为 k 与 α 的数量乘积,记作 $k\alpha$;

(3) 上述两个运算满足下列八条规则:

(i) 对任意 $\alpha, \beta \in V, \alpha + \beta = \beta + \alpha$;

(ii) 对任意 $\alpha, \beta, \gamma \in V, (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$;

(iii) V 中存在一个零元素,记作 0 ,对任意 $\alpha \in V$,都有 $\alpha + 0 = \alpha$;

(iv) 对任一 $\alpha \in V$,都有 $\beta \in V$,使得 $\alpha + \beta = 0$,元素 β 称为 α 的负元素,记作 $-\alpha$;

(v) 对任一 $\alpha \in V$,都有 $1 \cdot \alpha = \alpha$;

(vi) 对任一 $\alpha \in V, k, l \in P$,都有 $k(l\alpha) = (kl)\alpha$;

(vii) 对任一 $\alpha \in V, k, l \in P$,都有 $(k+l)\alpha = k\alpha + l\alpha$;

(viii) 对任一 $k \in P, \alpha, \beta \in V$,都有 $k(\alpha + \beta) = k\alpha + k\beta$,

则集合 V 称为数域 P 上的线性空间或向量空间. V 中的元素常称为向量. V 中的零元素称为零向量. 当 P 是实数域时, V 称为实线性空间;当 P 是复数域时, V 称为复线性空间. 数域 P 上的线性空间有时简称为线性空间.

由定义 1.1 可以证明:线性空间 V 中的零向量是唯一的; V 中每个元素 α 的负元素也是唯一的;并且有

$$0 \cdot \alpha = 0, \quad k \cdot 0 = 0, \quad (-1) \cdot \alpha = -\alpha,$$

这里 $k \in P, \alpha \in V$. 又 V 中元素的减法可以定义为(对任何 $\alpha, \beta \in V$)

$$\alpha - \beta = \alpha + (-\beta).$$

下面是一些常见的线性空间的例子.

例 1.1 若 P 是数域, V 是分量属于 P 的 n 元有序数组的集合

$$V = \{(x_1 \ x_2 \ \cdots \ x_n) | x_i \in P\},$$

若对 V 中任意两元素

$$\mathbf{X} = (x_1 \ x_2 \ \cdots \ x_n), \quad \mathbf{Y} = (y_1 \ y_2 \ \cdots \ y_n)$$

及每个 $k \in P$, 定义加法及数量乘法为

$$\mathbf{X} + \mathbf{Y} = (x_1 + y_1 \ x_2 + y_2 \ \cdots \ x_n + y_n), \quad k\mathbf{X} = (kx_1 \ kx_2 \ \cdots \ kx_n),$$

则容易验证, 集合 V 构成数域 P 上的线性空间. 这个线性空间记为 P^n .

例 1.2 所有元素属于数域 P 的 $m \times n$ 矩阵组成的集合, 按通常定义的矩阵加法及数与矩阵的数量乘法, 也构成数域 P 上的一个线性空间, 并把它记为 $P^{m \times n}$.

例 1.3 若 n 为正整数, P 是数域, 则系数属于 P 而未定元为 t 的所有次数小于 n 的多项式的集合, 这个集合连同零多项式在内, 按通常多项式的加法及数与多项式的乘法构成数域 P 上的线性空间. 用 $P[t]_n$ 代表这个空间. 若把“次数小于 n 的”这一限制条件取消, 则也得到一个线性空间, 并记为 $P[t]$.

例 1.4 所有定义在区间 $[a, b] (a \leq b)$ 上的实值连续函数构成的集合, 按照函

数的加法及数与函数的乘法,显然构成实数域上的一个线性空间,记为 $\mathbf{R}[a, b]$.

在讨论线性空间的问题时,下面几个概念是必须熟知的.

定义 1.2 设 V 是数域 P 上的线性空间, $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的一组向量,如果 P 中有一组不全为零的数 $k_1, k_2, \dots, k_n$, 使得

$$k_1\alpha_1 + k_2\alpha_2 + \dots + k_n\alpha_n = \mathbf{0}, \quad (1.1)$$

则称向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性相关;若等式(1.1)当且仅当 $k_1 = k_2 = \dots = k_n = 0$ 时才成立,则称这组向量是线性无关的.

由定义 1.2 得知,如果向量 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性相关,则使得式(1.1)成立的数 $k_1, k_2, \dots, k_n$ 中至少有一个不等于零,例如, $k_1 \neq 0$, 则有

$$\alpha_1 = -\frac{k_2}{k_1}\alpha_2 - \frac{k_3}{k_1}\alpha_3 - \dots - \frac{k_n}{k_1}\alpha_n,$$

这时,向量 α_1 是向量 $\alpha_2, \alpha_3, \dots, \alpha_n$ 的线性组合. 或者说,向量 α_1 可由 $\alpha_2, \alpha_3, \dots, \alpha_n$ 线性表示.

一般地,一组向量(含有限个向量)线性相关是指其中至少有一个向量可由这组中其余向量线性表出;反过来,如果这组向量具有这一性质,则这组向量必定线性相关. 但不难推知,线性无关的一组向量,其任一向量都不可能由这组向量中其余向量线性表出.

1.2 基变换与坐标变换

定义 1.3 设 V 是线性空间,若存在一组线性无关的向量 $\alpha_1, \alpha_2, \dots, \alpha_n$, 使空间中任一向量可由它们线性表示,则称向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 为 V 的一组基,基所含向量个数为 V 的维数,记为 $\dim V = n, n < +\infty$ 或者 $n = +\infty$.

若 V 的一组基中向量个数为 n , 称 n 为 V 的维数,记为 $\dim V = n$;若基中向量个数不是有限个,称 V 是无限维向量空间. 本书主要讨论有限维线性空间.

在 n 维线性空间中,其任意的 n 个线性无关向量都构成它的一组基. 由线性空间维数定义可知,在有限维线性空间中,基是存在的,但不是唯一的. 因为当维数是 n 时,空间里的任何 n 个线性无关的向量都可以作为它的一组基.

线性空间中的基与维数是依赖于向量的线性相关与线性无关的概念来定义的. 线性空间 V 作为一个向量集合,其中向量的线性相关与线性无关的定义与线性代数中给出的定义完全类似. 因而在线性空间中,极大线性无关组、等价等概念在形式上与向量空间 $\mathbf{R}^n$ 中的定义一样,与上述概念相关的性质和结果也可平移到线性空间中,在此不再叙述.

我们再来看看前面几个例子中线性空间 $P^n, P^{m \times n}, P[t]_n$ 的维数.

首先,容易证明

$\alpha_1=(1 \ 0 \ \cdots \ 0)$, $\alpha_2=(0 \ 1 \ \cdots \ 0)$, $\cdots$, $\alpha_n=(0 \ 0 \ \cdots \ 1)$ 是线性空间 P^n 的 n 个线性无关向量, 又显然 P^n 中任一向量

$$\alpha=(k_1 \ k_2 \ \cdots \ k_n)$$

都可由这 n 个线性无关向量线性表出, 即

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n,$$

从而得知 P^n 是 n 维线性空间. 今后用得较多的是 $\mathbf{R}^n$ 及 $\mathbf{C}^n$.

再考察线性空间 $P^{m \times n}$, 若用 E_{ij} 表示第 i 行、第 j 列上的元素等于 1 而其他元素均等于零的 $m \times n$ 矩阵, 则下列的 mn 个矩阵 $E_{11}, E_{12}, \cdots, E_{ij}, \cdots, E_{nm}$ ($i=1, 2, \cdots, m; j=1, 2, \cdots, n$) 构成 $P^{m \times n}$ 的一组基, 故 $P^{m \times n}$ 是 mn 维线性空间. 今后用得较多的是 $\mathbf{R}^{m \times n}$ 及 $\mathbf{C}^{m \times n}$, 包括它们当 $m=n$ 时的特殊情况.

最后, 由于 $1, t, t^2, \cdots, t^{n-1}$ 是 $P[t]_n$ 的一组基, 故 $P[t]_n$ 是 n 维线性空间.

例 1.5 向量组 $e_1=(1 \ 0 \ 0 \ \cdots \ 0)^T, e_2=(0 \ 1 \ 0 \ \cdots \ 0)^T, \cdots, e_n=(0 \ 0 \ \cdots \ 0 \ 1)^T$ 是 F^n 的一组基, 所以 $\dim F^n=n$.

例 1.6 求矩阵空间 $\mathbf{R}^{2 \times 2}$ 的维数与一组基.

解 任取矩阵 $A=\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}$, 有

$$A=\alpha_{11}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}+\alpha_{12}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}+\alpha_{21}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}+\alpha_{22}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

因此 $\mathbf{R}^{2 \times 2}$ 中任何一个向量都可写成向量组

$$E_{11}=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21}=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22}=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

的线性组合, 又任取数 k_i , 由

$$k_1E_{11}+k_2E_{12}+k_3E_{21}+k_4E_{22}=\begin{pmatrix} k_1 & k_2 \\ k_3 & k_4 \end{pmatrix}=\mathbf{O},$$

得 $k_i=0, i=1, 2, 3, 4$, 故 $E_{11}, E_{12}, E_{21}, E_{22}$ 线性无关, 由定义 1.2 知 $E_{11}, E_{12}, E_{21}, E_{22}$ 是 $\mathbf{R}^{2 \times 2}$ 的一组基, $\dim \mathbf{R}^{2 \times 2}=4$.

类似地, $E_{ij}, i=1, 2, \cdots, m; j=1, 2, \cdots, n$ 是矩阵空间 $\mathbf{R}^{m \times n}$ 的一组基,

$$\dim \mathbf{R}^{m \times n}=m \times n.$$

例 1.7 向量组 $1, x, x^2, \cdots, x^{n-1}$ 是线性空间 $P_n[x]$ 的一组基, $\dim P_n[x]=n$.

以上例子中, 空间维数 $\dim V$ 都为有限数, 这样的空间称为有限维线性空间. 若 $\dim V$ 不是有限数, 则称 V 为无限维线性空间, 由于有限维空间与无限维空间在研究方法上的较大差异, 这里只讨论有限维空间. 约定记号 $V_n(F)$ 表示 V 是数域 F 上的 n 维线性空间. 由于基就是向量集合 V 的极大线性无关组, 从而线性空间的基也不是唯一的.

定理 1.1 n 个线性无关的向量构成的向量组都是 n 维线性空间的基.

在线性空间 $V_n(F)$ 中, 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是一组基, 则 $\alpha_1, \alpha_2, \dots, \alpha_n$ 线性无关, $\forall \beta \in V_n(F), \alpha_1, \alpha_2, \dots, \alpha_n, \beta$ 线性相关, 故 β 可由 $\alpha_1, \alpha_2, \dots, \alpha_n$ 唯一地线性表示. 因此有如下定理.

定理 1.2 设 V 是数域 P 上的 n 维线性空间, $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的一组基, 则 V 中任一向量 α 都可以表示为这组基的线性组合, 且表示式是唯一的.

证明 由定义 1.3 知

$$\alpha = k_1 \alpha_1 + k_2 \alpha_2 + \dots + k_n \alpha_n, \quad (1.2)$$

如果 α 还有另一表示式

$$\alpha = l_1 \alpha_1 + l_2 \alpha_2 + \dots + l_n \alpha_n, \quad (1.3)$$

则由式(1.2)、式(1.3)即得

$$(k_1 - l_1) \alpha_1 + (k_2 - l_2) \alpha_2 + \dots + (k_n - l_n) \alpha_n = 0.$$

因向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 V 的一组基, 故线性无关, 所以

$$k_1 - l_1 = k_2 - l_2 = \dots = k_n - l_n = 0,$$

从而有 $k_i = l_i (i=1, 2, \dots, n)$. 这就证明了表示式的唯一性.

定义 1.4 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是线性空间 $V_n(F)$ 的一组基, $\forall \beta \in V$,

$$\beta = \sum_{i=1}^n x_i \alpha_i = (\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad (1.4)$$

则称数 $x_1, x_2, \dots, x_n$ 是 β 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标, 式(1.4)中向量 $(x_1 \quad x_2 \quad \dots \quad x_n)^T$ 为 β 的坐标向量, 也简称为坐标.

式(1.4)中第二个等式是借助于矩阵的形式运算来表示的, 显然 $(\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n)$ 在一般的向量意义下不一定为矩阵, 但这种表示会给今后的矩阵处理带来很多便利.

表示式(1.2)中的数 $k_1, k_2, \dots, k_n$ 称为向量 α 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标. 定理 1.2 说明, 取定一组基后, 每个向量 α 在这组基下的坐标是唯一取定的. α 的第 i 个坐标 $k_i (i=1, 2, \dots, n)$ 也称为 α 的第 i 个分量.

例 1.8 求 $\mathbf{R}^{2 \times 2}$ 中向量 $\begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix}$ 在基 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的坐标.

解 $\begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix} = 3E_{11} + E_{12} + 4E_{21} + 5E_{22} = (E_{11} \quad E_{12} \quad E_{21} \quad E_{22}) \begin{pmatrix} 3 \\ 1 \\ 4 \\ 5 \end{pmatrix}$, 故该向量

在所给基下的坐标为 $(3 \ 1 \ 4 \ 5)^T$. 一般地, $\mathbf{R}^{2 \times 2}$ 中向量 $\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ 在基 $\{E_{ij}\}$ 下的坐标为 $(\alpha_{11} \ \alpha_{12} \ \alpha_{21} \ \alpha_{22})^T$.

例 1.9 已知 $1, x, x^2, x^3$ 和 $1, (x-1), (x-1)^2, (x-1)^3$ 是线性空间 $P_4[x]$ 的两组基, 求 $P_4[x]$ 中向量 $f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$ 在这两组基下的坐标.

解 因为 $f(x) = (1 \ x \ x^2 \ x^3) \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$, 所以 $f(x)$ 在基 $1, x, x^2, x^3$ 下的坐标

为 $(\alpha_0 \ \alpha_1 \ \alpha_2 \ \alpha_3)^T$.

又由 Taylor 公式可知

$$\begin{aligned} f(x) &= f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3 \\ &= (1 \ (x-1) \ (x-1)^2 \ (x-1)^3) \begin{pmatrix} f(1) \\ f'(1) \\ \frac{1}{2}f''(1) \\ \frac{1}{3}f'''(1) \end{pmatrix}, \end{aligned}$$

故 $f(x)$ 在基 $1, (x-1), (x-1)^2, (x-1)^3$ 下的坐标为 $(f(1) \ f'(1) \ \frac{f''(1)}{2!} \ \frac{f'''(1)}{3!})^T$.

从以上例子和式 (1.2) 可以看到, 不论 $V_n(F)$ 为何种具体的线性空间, 当在 $V_n(F)$ 中取定一组基时, $V_n(F)$ 中向量在该组基下的坐标都是线性空间 F^n 中的向量. 正是这一特点, 奠定了可以用数量矩阵和 $\mathbf{R}^n$ 中向量来研究一般的线性空间中有关问题的基础, 一般同一个向量在不同基下的坐标是不同的.

在线性空间 $V_n(F)$ 中取定一组基 $\alpha_1, \alpha_2, \dots, \alpha_n, \forall \beta \in V_n(F)$, 取坐标作为对应关系, β 唯一地对应于 F^n 中一个向量 X (β 的坐标); 反之, $\forall X \in F^n, (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)X$ 就是 X 所对应的 $V_n(F)$ 中的向量, 因此, 坐标关系建立了线性空间 $V_n(F)$ 和 F^n 的一一对应关系 σ , 显然 σ 满足

$$\begin{aligned} \sigma(\alpha + \beta) &= \sigma(\alpha) + \sigma(\beta), \\ \sigma(k\alpha) &= k\sigma(\alpha). \end{aligned}$$

在对应关系 σ 下, 数域 F 上任何一个 n 维线性空间 $V_n(F)$ 都和 n 维线性空间 F^n 同构, 1.4 节将具体讲述这一内容.

设 $V_n(F)$ 中向量 β_i 的坐标为 $X_i, i=1, 2, \dots, n$, 则 β_i 的线性组合 $\sum_{i=1}^n k_i \beta_i$ 的坐标是 $\sum_{i=1}^n k_i X_i$, 又知零向量的坐标为 0, 所以

$$\sum_{i=1}^n k_i \beta_i = 0 \Leftrightarrow \sum_{i=1}^n k_i X_i = 0.$$

该结果可叙述为下述定理.

定理 1.3 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是 n 维线性空间 $V_n(F)$ 的一组基, $V_n(F)$ 中向量 β_i 在该组基下的坐标为 $X_i, i=1, 2, 3, \dots, n$, 则 $V_n(F)$ 中向量组 $\beta_1, \beta_2, \dots, \beta_n$ 线性相关的充分必要条件是坐标向量组 $X_1, X_2, \dots, X_n$ 是 F^n 中的线性相关组.

定理 1.3 说明由坐标建立的 $V_n(F)$ 和 F^n 之间的一一对应关系保持线性关系不变. 若不计较向量的具体形式, 仅就线性关系而言, $V_n(F)$ 中有关问题都可归结为我们所熟悉的线性空间 F^n 中的相关问题, 可应用熟悉的方法和已建立的理论来解决.

例 1.10 讨论 $P_4[x]$ 中向量

$$f_1 = 1 + 2x + 4x^3, \quad f_2 = x + x^2 + 4x^3, \quad f_3 = 1 + x - 3x^2, \quad f_4 = -2x^2 + x^3$$

的线性相关性.

解 在 $P_4[x]$ 中取基 $1, x, x^2, x^3$, 则向量组对应的坐标分别为

$$X_1 = (1 \ 2 \ 0 \ 4)^T, \quad X_2 = (0 \ 1 \ 1 \ 4)^T,$$

$$X_3 = (1 \ 1 \ -3 \ 0)^T, \quad X_4 = (0 \ 0 \ -2 \ 1)^T,$$

$$A = (X_1 \ X_2 \ X_3 \ X_4) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & -3 & -2 \\ 4 & 4 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$\text{rank}(A) = 4$, 因而 X_1, X_2, X_3, X_4 线性无关, 即 f_1, f_2, f_3, f_4 线性无关.

从定理 1.1 可以知道, 线性空间 $V_n(F)$ 的基不是唯一的, 同一向量在不同基下的坐标一般也不相同, 这里讨论 $V_n(F)$ 中不同基之间的关系和同一向量在不同基下坐标的关系.

设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 和 $\beta_1, \beta_2, \dots, \beta_n$ 是 $V_n(F)$ 中的两组基, 则由基的定义

$$\beta_i = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n) \begin{pmatrix} C_{1i} \\ C_{2i} \\ \vdots \\ C_{ni} \end{pmatrix} \triangleq (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n) C_i, \quad i=1, 2, \dots, n. \quad (1.5)$$

由矩阵分块运算, 式(1.5)中 n 个式子可表示为

$$(\beta_1 \ \beta_2 \ \cdots \ \beta_n) = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n) C, \quad (1.6)$$

其中 $C = (C_1 \ C_2 \ \cdots \ C_n) \in F^{m \times n}$.

定义 1.5 设 $\alpha_1, \alpha_2, \dots, \alpha_n$ 与 $\beta_1, \beta_2, \dots, \beta_n$ 是 n 维线性空间 $V_n(F)$ 的两组基, 若有矩阵 $C \in F^{m \times n}$, 使得

$$(\beta_1 \ \beta_2 \ \cdots \ \beta_n) = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)C,$$

则称矩阵 C 是从基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 到基 $\beta_1, \beta_2, \dots, \beta_n$ 的过渡矩阵(基变换矩阵).

C 作为过渡矩阵, 它一定是可逆矩阵, C^{-1} 是从 $\beta_1, \beta_2, \dots, \beta_n$ 到 $\alpha_1, \alpha_2, \dots, \alpha_n$ 的过渡矩阵, 从式(1.5)可知, 在构成上, 矩阵 C 的第 i 列是 β_i 在基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 下的坐标 C_i .

设向量 $\alpha \in V_n(F)$, α 在两组基下坐标分别为 X 和 Y , 则有

$$\alpha = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)X, \quad (1.7)$$

$$\alpha = (\beta_1 \ \beta_2 \ \cdots \ \beta_n)Y, \quad (1.8)$$

因此

$$\alpha = (\beta_1 \ \beta_2 \ \cdots \ \beta_n)Y = (\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n)CY. \quad (1.9)$$

比较式(1.7)与式(1.9), 有

$$X = CY.$$

定理 1.4 设线性空间 $V_n(F)$ 的一组基 $\alpha_1, \alpha_2, \dots, \alpha_n$ 到另一组基 $\beta_1, \beta_2, \dots, \beta_n$ 的过渡矩阵为 C , $V_n(F)$ 中向量 α 在两组基下的坐标分别为 X, Y , 则有

$$X = CY. \quad (1.10)$$

例 1.11 设 $\mathbf{R}^3$ 的两组基分别为 $\alpha_1 = (1 \ 0 \ -1)^T, \alpha_2 = (2 \ 1 \ 1)^T, \alpha_3 = (1 \ 1 \ 1)^T$ 和 $\beta_1 = (0 \ 1 \ 1)^T, \beta_2 = (-1 \ 1 \ 0)^T, \beta_3 = (1 \ 2 \ 1)^T$.

(1) 求从基 $\alpha_1, \alpha_2, \alpha_3$ 到基 $\beta_1, \beta_2, \beta_3$ 的过渡矩阵 C .

(2) 求向量 $\alpha = \alpha_1 + 2\alpha_2 - 3\alpha_3$ 在基 $\beta_1, \beta_2, \beta_3$ 下的坐标.

解 (1) 由定义 1.5, 过渡矩阵 C 是使等式 $(\beta_1 \ \beta_2 \ \beta_3) = (\alpha_1 \ \alpha_2 \ \alpha_3)C$ 成立的矩阵, 对线性空间 $\mathbf{R}^3$ 而言, 这是矩阵等式

$$\begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} C,$$

从而可求得

$$C = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ -1 & -3 & -2 \\ 2 & 4 & 4 \end{pmatrix}.$$

(2) 由 $\alpha = \alpha_1 + 2\alpha_2 - 3\alpha_3$, 得 α 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的坐标 $X = (1 \ 2 \ -3)^T$, 由 $X = CY$ 得

$$Y = C^{-1}X = \begin{pmatrix} 0 & 1 & 1 \\ -1 & -3 & -2 \\ 2 & 4 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -7 \\ -1 \\ 3 \end{pmatrix}.$$

例 1.12 从例 1.10 可知 $f_1=1+2x+4x^3, f_2=x+x^2+4x^3, f_3=1+x-3x^2, f_4=-2x^2+x^3$ 也是线性空间 $P_4[x]$ 的一组基, 求空间的基 $1, x, x^2, x^3$ 到基 f_1, f_2, f_3, f_4 的过渡矩阵 C , 并求向量 $f=1+x+x^2+x^3$ 在基 f_1, f_2, f_3, f_4 下的坐标 Y .

解 因为易求得 f_i 在基 $1, x, x^2, x^3$ 下的坐标 C_i , 从而易得从基 $1, x, x^2, x^3$ 到基 f_1, f_2, f_3, f_4 的过渡矩阵 C , 即

$$C=(C_1 \quad C_2 \quad C_3 \quad C_4)=\begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & -3 & -2 \\ 4 & 4 & 0 & 1 \end{pmatrix},$$

所以

$$C^{-1}=\begin{pmatrix} 6 & -\frac{9}{2} & \frac{1}{2} & 1 \\ -7 & \frac{11}{2} & -\frac{1}{2} & -1 \\ -5 & \frac{9}{2} & -\frac{1}{2} & -1 \\ 4 & -4 & 0 & 1 \end{pmatrix}.$$

又

$$f=1+x+x^2+x^3=(1 \quad x \quad x^2 \quad x^3)\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}=(f_1 \quad f_2 \quad f_3 \quad f_4)C^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

所以

$$Y=C^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}=\begin{pmatrix} 3 \\ -3 \\ -2 \\ 1 \end{pmatrix}.$$

1.3 子空间与维数定理

线性空间 $V_n(F)$ 中 V 作为集合, 对于集合中子集, 会涉及交与并等运算方式, 这里讨论对于这些关系的有关性质, 我们首先给出子空间的概念.

定义 1.6 设 $V_n(F)$ 为线性空间, W 是 V 的非空子集合. 若 W 的元素关于 V 中的加法与数乘运算也构成线性空间, 则称 W 是 V 的一个子空间.

例 1.13 任何线性空间有两个平凡子空间:一个是它自身 $V \subseteq V$, 另一个是 $W = \{0\}$, 称为零元素空间, 显然 $\dim\{0\} = 0$.

子集的包含关系使得 $V_n(F)$ 的一个子集合是否为子空间的判别比较方便.

定理 1.5 设 W 是线性空间 $V_n(F)$ 的非空子集合, 则 W 是 $V_n(F)$ 的子空间的充分必要条件是:

- (1) 若 $\alpha, \beta \in W$, 则 $\alpha + \beta \in W$;
- (2) 若 $\alpha \in W, k \in F$, 则 $k\alpha \in W$.

证明 必要性是显然的, 只证充分性.

设 W 满足条件(1)与条件(2), 则只需验证定义 1.1 中八条运算法则也满足即可.

因为 $k\alpha \in W$, 取 $k=0$, 则 $0\alpha=0 \in W$, 又取 $k=-1$, $-\alpha=(-1)\alpha \in W$, 即 W 中存在零元素和一个元素的负元素. 又因为 $W \subseteq V$, 对 V 中加法与数乘, 定义 1.1 的其余六条法则对 W 中元素进行运算时必须满足, 故由定义 1.1, W 是线性空间, 从而是 $V_n(F)$ 的子空间.

例 1.14 在线性空间 $\mathbf{R}^{n \times n}$ 中取集合.

$$W_1 = \{A | A \in \mathbf{R}^{n \times n}, A^T = A\};$$

$$W_2 = \{B | B \in \mathbf{R}^{n \times n}, |B| \neq 0\}.$$

讨论 W_1 与 W_2 是否为 $\mathbf{R}^{n \times n}$ 的子空间.

解 由于 $\forall A_1, A_2 \in W_1$, 有

$$(A_1 + A_2)^T = A_1^T + A_2^T = A_1 + A_2,$$

$$(kA_1)^T = kA_1^T = kA_1,$$

所以 $A_1 + A_2, kA_1 \in W_1$, 由定理 1.5, W_1 是 $\mathbf{R}^{n \times n}$ 的子空间.

因为 $0 \notin W_2$, 所以 W_2 不是 $\mathbf{R}^{n \times n}$ 的子空间.

例 1.15 设 $V_n(F)$ 是线性空间, $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 V 中一组向量, 则由它们的一切线性组合构成的集合

$$L\{\alpha_1, \alpha_2, \dots, \alpha_m\} = \left\{ \alpha \mid \alpha = \sum_{i=1}^m k_i \alpha_i, k_i \in F \right\}$$

是 V 的一个子空间, 称为由 $\alpha_1, \alpha_2, \dots, \alpha_m$ 生成的子空间.

证明 $\forall \alpha, \beta \in L\{\alpha_1, \alpha_2, \dots, \alpha_m\}$, 即

$$\alpha = \sum_{i=1}^m k_i \alpha_i, \quad \beta = \sum_{i=1}^m l_i \alpha_i,$$

则

$$\alpha + \beta = \sum_{i=1}^m (k_i + l_i) \alpha_i \in L\{\alpha_1, \alpha_2, \dots, \alpha_m\},$$

$$k\alpha = \sum_{i=1}^m (kk_i) \alpha_i \in L\{\alpha_1, \alpha_2, \dots, \alpha_m\},$$