研究生教学用书

矩阵分析

王永茂 编著

机械工业出版社
CHINA MACHINE PRESS

研 究 生 教 学 用 书

矩 阵 分 析

王永茂　编著
王晓敏　主审

机 械 工 业 出 版 社

本书系统、概括地论述了工程中常用的矩阵理论和方法.主要内容包括:线性空间与线性变换、矩阵的分解、范数和极限、矩阵分析、矩阵函数、广义逆矩阵、矩阵的扰动问题简介.各章末配有一定数量的习题.

本书可作为工科硕士研究生的教材,也可供理工科教师、科技工作者参考.

图书在版编目(CIP)数据

矩阵分析/王永茂编著.—北京:机械工业出版社,2005.8
ISBN 7-111-17298-1

Ⅰ.矩… Ⅱ.王… Ⅲ.矩阵分析 Ⅳ.0151.21

中国版本图书馆CIP数据核字(2005)第098815号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)
责任编辑:郑 丹 版式设计:霍永明 责任校对:王 欣
封面设计:鞠 杨 责任印制:杨 曦
北京机工印刷厂印刷
2005年9月第1版第1次印刷
890mm×1240mm A5·5.75印张·156千字
定价:12.00元

前　言

矩阵理论是经典数学的基础,也是实用性最强的数学分支之一,是处理大量有限维空间形式与数学关系的强有力的工具.矩阵理论在系统科学、优化方法、控制论、图论、稳定性理论等众多领域中都有广泛的应用.计算机的普及进一步促进了矩阵理论的发展.

自20世纪80年代末期开始,一些学校开始把"矩阵分析"作为理工科部分专业研究生的教学内容,经过十几年的发展,"矩阵分析"已成为众多理工科学校硕士研究生的学位课和部分专业的博士研究生入学考试科目.

本书是在作者1991年编写的《矩阵分析》一书的基础上精选了一些主要内容,经过多年教学实践,几经修改、补充编写而成的.

本书内容分为四大部分,即线性代数基础(第1章、第2章),矩阵分析(第3章、第4章、第5章),广义逆矩阵(第6章),矩阵的扰动分析简介(第7章).前6章为必学内容,第7章为选学内容.本书选材力求简明精练、深入浅出.每章末附有较多习题,供学生练习使用.使用本书授课需用50～60学时,各专业可根据需要灵活掌握.

本书在编写过程中得到了燕山大学研究生院和理学院部分教师的帮助,在此表示衷心感谢!还要感谢我的研究生魏静和尚勤同学,她们打印了全部书稿.

由于编者水平有限,书中缺点、错误在所难免,热忱欢迎读者批评指正.

编　者

目　　录

第 1 章　线性空间与线性变换

在线性代数中我们已经学过线性空间和线性变换的概念,因此本章的 1.1～1.3 节中仅对线性空间和线性映射的概念作简单的介绍,在 1.4～1.6 节中讲解酉空间和欧氏空间的性质及正交变换,1.7 节中讲解子空间的概念.

1.1　线性空间及其性质

定义 1.1　设 V 是一个非空集合,F 是一个数域,称 V 为按所定义的运算构成 F 上的线性空间(简称 V 为 F 上的线性空间或向量空间).

如果

(1) 在 V 中定义了加法,使对于任意的 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}\in V$ 存在惟一的元素 $\boldsymbol{\gamma}\in V$ 与之对应,称为 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的和,记作 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\gamma}$.

(2) 在 V 中定义了数乘,使对于任意的 $\boldsymbol{\alpha}\in V$ 及任意 $\lambda\in F$,存在惟一的元素 $\boldsymbol{\delta}\in V$ 与之对应,称为 λ 与 $\boldsymbol{\alpha}$ 的积,记作 $\boldsymbol{\delta}=\lambda\boldsymbol{\alpha}$.

(3) 所定义的加法与数乘合起来是线性运算,即这两种运算满足运算法则

1) $\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$

2) $(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$

3) 在 V 中存在零元素,以 $\mathbf{0}$ 表示,使对于任何 $\boldsymbol{\alpha}\in V$,都有 $\boldsymbol{\alpha}+\mathbf{0}=\boldsymbol{\alpha}$.

4) 对任何 $\boldsymbol{\alpha}\in V$,都有 $\boldsymbol{\beta}\in V$,使 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\mathbf{0}$,称 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}$ 的负元素.

5) 存在单位数 $1\in F$,使得 $1\boldsymbol{\alpha}=\boldsymbol{\alpha}$.

6) $\lambda(\mu\boldsymbol{\alpha})=(\lambda\mu)\boldsymbol{\alpha}$

7) $\lambda(\boldsymbol{\alpha}+\boldsymbol{\beta})=\lambda\boldsymbol{\alpha}+\lambda\boldsymbol{\beta}$

8) $(\lambda+\mu)\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\alpha}$

式中,$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in V$; $\lambda,\mu\in F$.

数域 F 上的线性空间 V 记为 $V(F)$,V 中元素不论其本来的性质如何,统称为向量.

注意:①这里的向量不一定是有序数组;②这里的单位数 1 不一定是实数 1,它与所定义的运算有关.

例 1.1　次数不超过 n 的实系数多项式全体记为 $P[x]_n$,即

$$P[x]_n=\{p_n \mid p_n=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\ a_n,\cdots,a_0\in\mathbf{R}\}$$

则 $P[x]_n$ 按照通常的多项式加法及数与多项式的乘法构成实数域上的线性空间.

这是因为 $P[x]_n$ 对于上述两种运算显然是封闭的,且满足 8 条运算律,所以 $P[x]_n$ 是 $\mathbf{R}$ 上的线性空间.

例 1.2　正实数全体记作 $\mathbf{R}^+$,即

$$\mathbf{R}^+=\{a \mid a>0,\ a\in\mathbf{R}\}$$

在 $\mathbf{R}^+$ 中定义加法及数乘为

$$a\oplus b\triangleq ab \qquad \lambda\circ a\triangleq a^\lambda$$

其中,$a,b\in\mathbf{R}^+$, $\lambda\in\mathbf{R}$.

证明:$\mathbf{R}^+$ 对上述加法和数乘构成 $\mathbf{R}$ 上的线性空间.

证　实际上要验证以下 10 个方面:

对加法的封闭性:对于任何的 $a,b\in\mathbf{R}^+$,有 $a\oplus b=ab\in\mathbf{R}^+$.

对数乘的封闭性:对于任何 $\lambda\in\mathbf{R},a\in\mathbf{R}^+$,有 $\lambda\circ a=a^\lambda\in\mathbf{R}^+$,且

1) $a\oplus b=ab=ba=b\oplus a$

2) $$\begin{aligned}(a\oplus b)\oplus c&=(ab)\oplus c=(ab)c=a(bc)\\&=a\oplus(bc)=a\oplus(b\oplus c)\end{aligned}$$

3) 在 $\mathbf{R}^+$ 中存在零元素 1,使对任何 $a\in\mathbf{R}^+$,有 $a\oplus 1=a\cdot 1=$

a.

4）对于任何 $a\in\mathbf{R}^+$，有负元素 $a^{-1}\in\mathbf{R}^+$，使

$$a\oplus a^{-1}=aa^{-1}=1$$

5）存在单位数 $1\in\mathbf{R}$，使 $1\circ a=a^1=a$.

6）$\lambda\circ(\mu\circ a)=\lambda\circ a^\mu=(a^\mu)^\lambda=a^{\lambda\mu}=(\lambda\mu)\circ a$

7）$(\lambda+\mu)\circ a=a^{\lambda+\mu}=a^\lambda a^\mu=a^\lambda\oplus a^\mu=(\lambda\circ a)\oplus(\mu\circ a)$

8）$\lambda\circ(a\oplus b)=(ab)^\lambda=a^\lambda b^\lambda=(\lambda\circ a)\oplus(\lambda\circ b)$

其中，$a,b,c\in\mathbf{R}^+$；$\lambda,\mu\in\mathbf{R}$.

因此，$\mathbf{R}^+$ 对所定义的加法"$\oplus$"、数乘"$\circ$"构成 $\mathbf{R}$ 上的线性空间.

从这个例子可以看到线性空间中定义的加法和数乘不一定是通常意义上的加法与数乘，只是人为地把这两种运算叫做加法与数乘.

例 1.3　按通常的向量的加法与数乘，$\mathbf{R}^n$ 是实数域上的线性空间，$\mathbf{C}^n$ 是复数域上的线性空间，但 $\mathbf{R}^n$ 不是复数域 $\mathbf{C}$ 上的线性空间.

线性空间具有以下简单性质：

性质 1　零元素惟一，负元素惟一.

事实上，设 $\mathbf{0}_1,\mathbf{0}_2$ 均为 $V(F)$ 的零元素，则

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

其次，设 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 均为 $\boldsymbol{X}\in V(F)$ 的负元素，$\mathbf{0}$ 为 $V(F)$ 的零元素，则

$$\boldsymbol{X}_1=\boldsymbol{X}_1+\mathbf{0}=\boldsymbol{X}_1+(\boldsymbol{X}+\boldsymbol{X}_2)=(\boldsymbol{X}_1+\boldsymbol{X})+\boldsymbol{X}_2=\mathbf{0}+\boldsymbol{X}_2=\boldsymbol{X}_2$$

性质 2　设 $\lambda,0,-1,1\in F$；$\boldsymbol{X},-\boldsymbol{X},\mathbf{0}\in V(F)$，则

(1) $0\boldsymbol{X}=\mathbf{0}$

(2) $(-1)\boldsymbol{X}=-\boldsymbol{X}$

(3) $\lambda\mathbf{0}=\mathbf{0}$

(4) 若 $\lambda\boldsymbol{X}=\mathbf{0}$，则 $\lambda=0$ 或 $\boldsymbol{X}=\mathbf{0}$

证　$\forall\boldsymbol{X}\in V(F)$，有

(1) $\boldsymbol{X}+0\boldsymbol{X}=1\boldsymbol{X}+0\boldsymbol{X}=(1+0)\boldsymbol{X}=1\boldsymbol{X}=\boldsymbol{X}$，故 $0\boldsymbol{X}=\mathbf{0}$.

(2) $\boldsymbol{X}+(-1)\boldsymbol{X}=(1-1)\boldsymbol{X}=0\boldsymbol{X}=\mathbf{0}$，故 $(-1)\boldsymbol{X}=-\boldsymbol{X}$.

(3) $\lambda\mathbf{0}=\lambda(\boldsymbol{X}-\boldsymbol{X})=\lambda\boldsymbol{X}-\lambda\boldsymbol{X}=(\lambda-\lambda)\boldsymbol{X}=0\boldsymbol{X}=\mathbf{0}$.

(4) 若 $\lambda \neq 0$ 且 $\boldsymbol{X} \neq \boldsymbol{0}$,则

$$\boldsymbol{X} = 1\boldsymbol{X} = \left(\frac{1}{\lambda}\lambda\right)\boldsymbol{X} = \frac{1}{\lambda}(\lambda \boldsymbol{X}) = \frac{1}{\lambda}\boldsymbol{0} = \boldsymbol{0}$$

这与 $\boldsymbol{X} \neq \boldsymbol{0}$ 矛盾,故 $\lambda \neq 0$ 与 $\boldsymbol{X} \neq \boldsymbol{0}$ 不能同时成立.

定义 1.2 设 V 是一个线性空间,L 是 V 的一个非空子集,如果 L 对 V 中所定义的加法和数乘两种运算构成线性空间,则称 L 为 V 的子空间.

一个非空子集要满足什么条件才能够成为子空间呢? 因 L 是 V 的一部分,V 中的运算对 L 而言,规律 1)、2)、5)、6)、7)、8)显然是满足的,因此只要 L 对所定义的两种运算封闭且满足规律 3)、4)即可,但由线性空间的性质知,若 L 对运算封闭,即满足规律 3)、4),因此有

定理 1.1 线性空间 V 的非空子集 L 构成线性空间的充要条件是 L 对于 V 中所定义的运算封闭.

1.2 线性空间的维数、基与坐标

在 $\mathbf{R}^n$ 中有了向量的坐标表示式后,对于理论分析和实际应用都很方便,为此,本节将 $\mathbf{R}^n$ 中有关基、维数和坐标等概念推广到一般线性空间中去. 首先需要定义 $V(F)$ 中元素的线性相关性等概念.

为使用方便,以后将 $i = 1, 2, \cdots, n$ 记为 $i \in \underline{n}$.

定义 1.3 设 $\boldsymbol{X}_i \in V(F)$,$\boldsymbol{X} \in V(F)$,$k_i \in F$,$i \in \underline{m}$,若有

$$\boldsymbol{X} = k_1\boldsymbol{X}_1 + k_2\boldsymbol{X}_2 + \cdots + k_m\boldsymbol{X}_m = \sum_{i=1}^{m} k_i\boldsymbol{X}_i$$

则称 $\boldsymbol{X}$ 可由 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 线性表示,或称 $\boldsymbol{X}$ 是 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 的线性组合.

定义 1.4 $V(F)$ 中的向量组 $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$ 称为是线性相关的,系指存在 m 个不全为零的数 $k_i \in F$,$i \in \underline{m}$,使

$$\sum_{i=1}^{m} k_i\boldsymbol{X}_i = \boldsymbol{0} \tag{1.1}$$

否则,称向量组 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_m$ 是线性无关的,即仅当 $k_i=0,i\in\underline{m}$ 时,式(1.1)成立.

与 $\mathbf{R}^n$ 类似,在 $V(F)$ 中,下列命题成立.

命题 1　当 $m\geqslant 2$ 时,$V(F)$ 中的向量组 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_m$ 线性相关的充要条件为该向量组中至少有一个向量可由组中其余向量线性表示.

命题 2　若 $V(F)$ 中向量组的某一子向量组线性相关,则该向量组也线性相关.反之,若该向量组线性无关,则其任何子向量组也线性无关.

定义 1.5　线性空间 $V(F)$ 中的向量组 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 称为 $V(F)$ 的基系指

(1) $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 是线性无关组.

(2) $V(F)$ 中任一向量均可由向量组 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 线性表示.

线性空间 $V(F)$ 的基中所含向量的个数 n 称为 $V(F)$ 的维数,记为

$$\dim V(F)=n$$

并称 $V(F)$ 为 n 维线性空间,可简记为 $V_n(F)$.特别地,若对任何 $n\in\mathbf{Z}^+$($\mathbf{Z}^+$ 表示正整数集),在 $V(F)$ 中均可找到 n 个线性无关的向量,则称 $V(F)$ 为无限维线性空间.只含零向量的线性空间的维数规定为零.

定义 1.6　设 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 为 $V_n(F)$ 的基,则对 $\boldsymbol{X}\in V_n(F)$,其表达式

$$\boldsymbol{X}=\sum_{i=1}^{n}k_i\boldsymbol{X}_i,\quad k_i\in F$$

中的 $k_i,i\in\underline{n}$ 称为向量 $\boldsymbol{X}$ 关于基 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 的坐标.

例 1.4　在 $\mathbf{R}^n$ 中,向量组 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,0,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,\cdots,0,1)^{\mathrm{T}}$ 是线性无关的,它是 $\mathbf{R}^n$ 的基,易证向量组 $\boldsymbol{e}_1'=(1,1,\cdots,1)^{\mathrm{T}},\quad \boldsymbol{e}_2'=(0,1,1,\cdots,1)^{\mathrm{T}},\cdots,\boldsymbol{e}_n'=(0,\cdots,0,1)^{\mathrm{T}}$ 也是线性无关的,故它也是 $\mathbf{R}^n$ 的基.由此可见,线性空间的基不惟

一.

例 1.5 设 $P[x]_n$ 为次数不大于 n 的实系数多项式全体所组成的集合,易证

$$1, x, x^2, \cdots, x^n$$

是 $P[x]_n$ 的一个基.若 $p(x) \in P[x]_n$,且

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

则 $a_0, a_1, \cdots, a_n$ 就是 $p(x)$ 关于基 $1, x, x^2, \cdots, x^n$ 的坐标.若又取

$$1, (x-a), (x-a)^2, \cdots, (x-a)^n$$

为 $P[x]_n$ 的基,其中 a 为域 F 中的常数,则由泰勒公式知

$$\begin{aligned} p(x) = {} & p(a) + p'(a)(x-a) + \\ & \frac{p''(a)}{2!}(x-a)^2 + \cdots + \frac{p^{(n)}(a)}{n!}(x-a)^{(n)} \end{aligned}$$

故 $p(x)$ 在基 $1, (x-a), (x-a)^2, \cdots, (x-a)^n$ 下的坐标为 $p(a)$, $p'(a)$, $\frac{1}{2!}p''(a)$, $\cdots$, $\frac{1}{n!}p^{(n)}(a)$.

由此可见,在线性空间中,元素的坐标由基惟一确定,当基改变时,坐标将随之改变.

如果 $V_n(F)$ 的两个基之间的关系已知,如何推断 $V_n(F)$ 的元素关于这两个基的坐标之间的关系?

设 $(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)$ 与 $(\boldsymbol{\varepsilon}'_1, \boldsymbol{\varepsilon}'_2, \cdots, \boldsymbol{\varepsilon}'_n)$ 为 $V_n(F)$ 的两个基,且

$$\begin{cases} \boldsymbol{\varepsilon}'_1 = a_{11}\boldsymbol{\varepsilon}_1 + a_{21}\boldsymbol{\varepsilon}_2 + \cdots + a_{n1}\boldsymbol{\varepsilon}_n \\ \boldsymbol{\varepsilon}'_2 = a_{12}\boldsymbol{\varepsilon}_1 + a_{22}\boldsymbol{\varepsilon}_2 + \cdots + a_{n2}\boldsymbol{\varepsilon}_n \\ \qquad\qquad \vdots \\ \boldsymbol{\varepsilon}'_n = a_{1n}\boldsymbol{\varepsilon}_1 + a_{2n}\boldsymbol{\varepsilon}_2 + \cdots + a_{nn}\boldsymbol{\varepsilon}_n \end{cases} \tag{1.2}$$

若记 $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

则式(1.2)可表示为

$$(\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', \cdots, \boldsymbol{\varepsilon}_n') = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A} \tag{1.3}$$

或
$$\begin{bmatrix} (\boldsymbol{\varepsilon}_1')^{\mathrm{T}} \\ (\boldsymbol{\varepsilon}_2')^{\mathrm{T}} \\ \vdots \\ (\boldsymbol{\varepsilon}_n')^{\mathrm{T}} \end{bmatrix} = \boldsymbol{A}^{\mathrm{T}} \begin{bmatrix} (\boldsymbol{\varepsilon}_1)^{\mathrm{T}} \\ (\boldsymbol{\varepsilon}_2)^{\mathrm{T}} \\ \vdots \\ (\boldsymbol{\varepsilon}_n)^{\mathrm{T}} \end{bmatrix} \tag{1.4}$$

其中,矩阵 A 称为由基$(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)$到基$(\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', \cdots, \boldsymbol{\varepsilon}_n')$的过渡矩阵或基变换矩阵.

定理 1.2　设$(\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', \cdots, \boldsymbol{\varepsilon}_n') = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}$,若 $\boldsymbol{X} \in V(F)$,有

$$\boldsymbol{X} = x_1\boldsymbol{\varepsilon}_1 + x_2\boldsymbol{\varepsilon}_2 + \cdots + x_n\boldsymbol{\varepsilon}_n = y_1\boldsymbol{\varepsilon}_1' + y_2\boldsymbol{\varepsilon}_2' + \cdots + y_n\boldsymbol{\varepsilon}_n' \tag{1.5}$$

则
$$(x_1, x_2, \cdots, x_n)^{\mathrm{T}} = \boldsymbol{A}(y_1, y_2, \cdots, y_n)^{\mathrm{T}} \tag{1.6}$$

或
$$(y_1, y_2, \cdots, y_n)^{\mathrm{T}} = \boldsymbol{A}^{-1}(x_1, x_2, \cdots, x_n)^{\mathrm{T}} \tag{1.7}$$

证　由 $\boldsymbol{X}$ 的表达式(1.5)有

$$\begin{aligned}\boldsymbol{X} &= (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)(x_1, x_2, \cdots, x_n)^{\mathrm{T}} \\ &= (\boldsymbol{\varepsilon}_1', \boldsymbol{\varepsilon}_2', \cdots, \boldsymbol{\varepsilon}_n')(y_1, y_2, \cdots, y_n)^{\mathrm{T}} \\ &= (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}(y_1, y_2, \cdots, y_n)^{\mathrm{T}}\end{aligned}$$

再由坐标表示式的惟一性知式(1.6)成立.又 $\boldsymbol{A}$ 可逆,故也可得到式(1.7).

式(1.6)或式(1.7)就是当基变换矩阵 $\boldsymbol{A}$ 已知时,$\boldsymbol{X}$ 关于两个基的坐标之间的关系.

例 1.6　设$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$与$(\boldsymbol{e}_1', \boldsymbol{e}_2', \cdots, \boldsymbol{e}_n')$为 $\mathbf{R}^n$ 的两个基(见例 1.4),若 $\boldsymbol{X} \in \mathbf{R}^n$,且

$$\boldsymbol{X} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n = y_1\boldsymbol{e}_1' + y_2\boldsymbol{e}_2' + \cdots + y_n\boldsymbol{e}_n'$$

求$(\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)$到$(\boldsymbol{e}_1', \boldsymbol{e}_2', \cdots, \boldsymbol{e}_n')$的过渡矩阵 $\boldsymbol{A}$ 及 $\boldsymbol{X}$ 在两个基下的坐标之间的关系.

解　显然有

$$(\boldsymbol{e}_1', \boldsymbol{e}_2', \cdots, \boldsymbol{e}_n') = (\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n)\boldsymbol{A}$$

其中
$$A=\begin{pmatrix}1&0&0&\cdots&0\\1&1&0&\cdots&0\\\vdots&\ddots&\ddots&\ddots&\vdots\\1&1&\ddots&\ddots&0\\1&1&\cdots&1&1\end{pmatrix}$$
即为过渡矩阵，且
$$A^{-1}=\begin{pmatrix}1&0&0&\cdots&0\\-1&1&0&\cdots&0\\0&\ddots&\ddots&\ddots&\vdots\\\vdots&\ddots&\ddots&\ddots&0\\0&\cdots&0&-1&1\end{pmatrix}$$
代入式(1.7)得

$y_1=x_1,y_2=x_2-x_1,y_3=x_3-x_2,\cdots,y_n=x_x-x_{n-1}$

1.3 线性映射与线性变换

本节我们引进线性映射和线性变换的概念，它们在研究线性空间之间的关系中是很有用的.

定义1.7 设 M 与 M' 是两个集合，若对于每个 $X\in M$，通过某种法则 σ，在 M' 中有惟一确定的 Y 与之对应，则称 σ 是由 M 到 M' 的一个映射，记为
$$\sigma:M\to M' \quad 或 \quad \sigma(X)=Y$$
Y 称为 X 的像，X 称为 Y 的原像.

例1.7 设函数 $y=f(x)$ 的定义域为 $[a,b]$，值域为 $[c,d]$，则 f 即为由数集 $[a,b]$ 到数集 $[c,d]$ 的一个映射.

例1.8 对于任何 $A\in\mathbf{R}^{n\times n}$，定义 $\sigma(A)=\det A$，则 σ 是由 $\mathbf{R}^{n\times n}$ 到实数集 $\mathbf{R}$ 的一个映射.

例1.9 记 $F(x)=\{f(x)\mid f(x)可导\}$
$$F'(x)=\left\{f'(x)\mid f'(x)=\frac{df(x)}{dx},\ f(x)\in F(x)\right\}$$

则 $\sigma[f(x)] = f'(x)$ 是 $F(x)$ 到 $F'(x)$ 的一个映射.

定义 1.8　设 $V(F)$ 与 $W(F)$ 是两个线性空间，σ 是由 $V(F)$ 到 $W(F)$ 的一个映射，且

$$\sigma(\boldsymbol{X}+\boldsymbol{Y}) = \sigma(\boldsymbol{X})+\sigma(\boldsymbol{Y}) \quad \forall\, \boldsymbol{X},\boldsymbol{Y} \in V(F)$$

$$\sigma(\lambda\boldsymbol{X}) = \lambda\sigma(\boldsymbol{X}) \quad \forall\, \lambda \in F, \boldsymbol{X} \in V(F)$$

则称映射 σ 为由 $V(F)$ 到 $W(F)$ 的线性映射. 特别地，若还有 $W(F) \subset V(F)$，则称线性映射 σ 为 $V(F)$ 上的线性变换.

如例 1.8 中的 $\sigma(\boldsymbol{A}) = \det\boldsymbol{A}$ 不是线性映射. 例 1.9 中的 $\sigma[f(x)] = f'(x)$ 是线性映射，如将 $F(x)$ 换成 $P[x]$（数域 P 上的多项式全体），则 $\sigma[f(x)]$ 是 $P[x]$ 上的线性变换；当 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$，$\boldsymbol{X} \in \mathbf{R}^n$ 时，令 $\sigma(\boldsymbol{X}) = \boldsymbol{A}\boldsymbol{X}$，则 σ 是 $\mathbf{R}^n$ 上的线性变换；恒等变换 $I(\boldsymbol{X}) = \boldsymbol{X}$，零变换 $O(\boldsymbol{X}) = \boldsymbol{0}$ 都是 $\mathbf{R}^n$ 上的线性变换.

线性映射具有以下性质：

定理 1.3　设 $\sigma: V(F) \to U(F)$ 是线性映射，则

1) $\sigma(\boldsymbol{0}) = \boldsymbol{0}'$，其中 $\boldsymbol{0}$ 与 $\boldsymbol{0}'$ 分别为 $V(F)$ 与 $U(F)$ 的零元素.

2) $\sigma(-\boldsymbol{X}) = -\boldsymbol{X}, \forall\, \boldsymbol{X} \in V(F)$.

3) σ 将 $V(F)$ 中的线性相关向量组映射为 $U(F)$ 中的线性相关向量组.

4) 设 $V_1(F)$ 是 $V(F)$ 的子空间，若

$$\sigma(V_1(F)) = \{\sigma(\boldsymbol{X}) \mid \boldsymbol{X} \in V_1(F)\}$$

则 $\sigma(V_1(F))$ 是 $U(F)$ 的子空间，且

$$\dim\sigma(V_1(F)) \leqslant \dim V_1(F)$$

证　这里我们只就 3)、4) 进行证明，其他证明留给读者.

3) 设 $\boldsymbol{\alpha}_i \in V(F), i \in \underline{m}$ 是线性相关向量组，由线性相关的定义，存在 m 个不全为零的数 $k_1, k_2, \cdots, k_m$，使得

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \boldsymbol{0} \tag{1.8}$$

对式(1.8)两边取映射，有

$$k_1\sigma(\boldsymbol{\alpha}_1) + k_2\sigma(\boldsymbol{\alpha}_2) + \cdots + k_m\sigma(\boldsymbol{\alpha}_m) = \boldsymbol{0}', \boldsymbol{0}' \in U(F)$$

所以 3) 成立.

4) $\sigma(V_1(F))$是$U(F)$的子空间的证明留给读者,这里只证明

$$\dim\sigma(V_1(F)) \leqslant \dim V_1(F)$$

设$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$是$\sigma(V_1(F))$的基,且$\boldsymbol{\beta}_i = \sigma(\boldsymbol{\alpha}_i),\boldsymbol{\alpha}_i \in V_1$, $i \in \underline{m}$则由3)知,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关,故$V_1$中至少有$m$个线性无关的向量,所以4)成立.

定理1.4 设线性映射$\sigma_i:V(F) \to U(F)(i \in \underline{m})$及$\mu$: $U(F) \to W(F)$,则

$$\sigma = \sum_{i=1}^{m}\lambda_i\sigma_i,\ \lambda_i \in F, i \in \underline{m}$$

是由$V(F)$到$U(F)$的线性映射,而

$$\mu\sigma_i,\quad i \in \underline{m}$$

是由$V(F)$到$W(F)$的线性映射.(证略)

定理1.5 设$\sigma:V_n(F) \to V_n(F)$是线性变换,$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$为$V_n(F)$的基,则

1) 若$\sigma_1:V_n(F) \to V_n(F)$为另一线性变换,且$\sigma_1(\boldsymbol{\varepsilon}_i) = \sigma(\boldsymbol{\varepsilon}_i), i \in \underline{n}$,则$\sigma_1 = \sigma$.

2) 存在惟一的$\boldsymbol{A} \in \boldsymbol{F}^{n\times n}$,使

$$\sigma(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}$$

并称$\boldsymbol{A}$为线性变换σ在基$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$下的矩阵.

3) 对于任意$\boldsymbol{X} \in V_n(F)$,若有$\boldsymbol{a},\boldsymbol{b} \in \boldsymbol{F}^n$,使

$$\boldsymbol{X} = (\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{a},\quad \sigma(\boldsymbol{X}) = (\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{b}$$

则$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{a}$.

证 1) 任取$\boldsymbol{X} \in V_n(F)$,设

$$\boldsymbol{X} = k_1\boldsymbol{\varepsilon}_1 + k_2\boldsymbol{\varepsilon}_2 + \cdots + k_n\boldsymbol{\varepsilon}_n$$

则由
$$\begin{aligned}\sigma(\boldsymbol{X}) &= \sigma(k_1\boldsymbol{\varepsilon}_1 + k_2\boldsymbol{\varepsilon}_2 + \cdots + k_n\boldsymbol{\varepsilon}_n)\\ &= k_1\sigma(\boldsymbol{\varepsilon}_1) + k_2\sigma(\boldsymbol{\varepsilon}_2) + \cdots + k_n\sigma(\boldsymbol{\varepsilon}_n)\\ &= k_1\sigma_1(\boldsymbol{\varepsilon}_1) + k_2\sigma_1(\boldsymbol{\varepsilon}_2) + \cdots + k_n\sigma_1(\boldsymbol{\varepsilon}_n)\\ &= \sigma_1(k_1\boldsymbol{\varepsilon}_1 + k_2\boldsymbol{\varepsilon}_2 + \cdots + k_n\boldsymbol{\varepsilon}_n) = \sigma_1(\boldsymbol{X})\end{aligned}$$

知 $\sigma=\sigma_1$.

2) 设 $\sigma(\boldsymbol{\varepsilon}_i)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{a}_i,\boldsymbol{a}_i\in \boldsymbol{F}^n,i\in \underline{n}$,令

$$\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n)$$

则有

$$\sigma(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}$$

又由坐标的惟一性知 $\boldsymbol{A}$ 惟一.

3) 因为

$$\begin{aligned}\sigma(\boldsymbol{X})&=\sigma[(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{a}]=[\sigma(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)]\boldsymbol{a}\\&=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}\boldsymbol{a}\end{aligned}$$

又由 $\sigma(\boldsymbol{X})=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{b}$ 及坐标的惟一性,得

$$\boldsymbol{b}=\boldsymbol{A}\boldsymbol{a}$$

定理1.6　设 $\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$ 与 $\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n$ 为 $V_n(F)$ 的两个基,σ 为 $V_n(F)$ 上的线性变换,且

$$\sigma(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}$$

$$\sigma(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)=(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)\boldsymbol{B}$$

$$(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{T}$$

则
$$\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$$

即 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似.

证　因为

$$\sigma(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)=(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)\boldsymbol{B}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{T}\boldsymbol{B}$$

另一方面

$$\begin{aligned}\sigma(\boldsymbol{\varepsilon}'_1,\boldsymbol{\varepsilon}'_2,\cdots,\boldsymbol{\varepsilon}'_n)&=\sigma[(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{T}]\\&=[\sigma(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)]\boldsymbol{T}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n)\boldsymbol{A}\boldsymbol{T}\end{aligned}$$

所以
$$\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}\boldsymbol{B}$$

由于 $\boldsymbol{T}$ 可逆,故 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$.

例 1.10　设

$$\boldsymbol{\varepsilon}_1=(1,0,2,1)^{\mathrm{T}},\ \boldsymbol{\varepsilon}_2=(0,1,0,1)^{\mathrm{T}},\boldsymbol{\varepsilon}_3=(0,0,1,1)^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon}_4=(0,0,0,1)^{\mathrm{T}},\ \boldsymbol{\varepsilon}'_1=(1,0,0,0)^{\mathrm{T}},\boldsymbol{\varepsilon}'_2=(1,1,3,0)^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon}'_3=(1,0,1,0)^{\mathrm{T}},\boldsymbol{\varepsilon}'_4=(1,0,0,1)^{\mathrm{T}}$$