矩阵分析基础

王永茂　刘德友　编著

清华大学出版社

矩阵分析基础

王永茂　刘德友　编著

清华大学出版社
北　京

内 容 简 介

本书系统、概括地论述了工程中常用的矩阵理论和方法，主要包括：线性空间与线性变换、酉空间和酉变换、矩阵的分解、范数及其应用、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.

本书可作为理工科硕士研究生和高年级本科生的教材，也可供高校教师、科研工作者和工程技术人员参考.

图书在版编目（CIP）数据

矩阵分析基础/王永茂，刘德友编著. —北京：清华大学出版社，2012.9

ISBN 978-7-302-29649-2

Ⅰ. ①矩…　Ⅱ. ①王…　②刘…　Ⅲ. ①矩阵分析—研究生—教材　Ⅳ. ①O151.21

中国版本图书馆 CIP 数据核字(2012)第 185118 号

责任编辑：石　磊　赵从棉
封面设计：傅瑞学
责任校对：王淑云
责任印制：宋　林

出版发行：清华大学出版社
　网　址：http://www.tup.com.cn，http://www.wqbook.com
　地　址：北京清华大学学研大厦 A 座　　**邮　编**：100084
　社 总 机：010-62770175　　**邮　购**：010-62786544
　投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
　质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：清华大学印刷厂
经　销：全国新华书店
开　本：185mm×230mm　**印　张**：11　**字　数**：238 千字
版　次：2012 年 9 月第 1 版　**印　次**：2012 年 9 月第 1 次印刷
印　数：1～3000
定　价：24.00 元

产品编号：045992-01

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．它在系统科学、计算科学、优化方法、控制论、图论、稳定性理论、保险学、风险分析等众多领域中具有广泛的应用，特别是计算机的广泛应用，更进一步促进了矩阵理论的发展．

自20世纪80年代末，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，经过近30年的发展，现在矩阵分析几乎已成为所有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．但是由于矩阵分析作为研究生的教学内容时间尚短，且由于各学校的专业性质不同、侧重点不同、讲解内容也不同，因此，矩阵分析的教学还不够规范，没有形成体系．虽然在近十几年里，出版了几部《矩阵分析》教材，但这些教材多数还局限于自编自用的阶段，因此，编写一本能够普遍适用的《矩阵分析》教材势在必行．基于此，本书在作者1991年编写的《矩阵分析》讲义和2005年出版的《矩阵分析》教材一书的基础之上，参考了大量的相关资料，精选一些主要内容，经过多年教学实践，几经修改、补充编成此书，目的是使该教材对多数理工科硕士研究生的矩阵分析教学有所帮助．

本书内容由4部分构成，即线性代数基础（第1～3章）、矩阵分析（第4～6章）、广义逆矩阵（第7章）和矩阵的扰动问题简介（第8章）．本书在选材上力求简明精练、深入浅出．每章末附有较多习题，供学生练习时使用．本书全部讲完约需60学时，教学时数在30～40之间的可选讲第1～4章和第6章，标注“*”号的章节为选学内容．各专业可视具体情况，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！

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

2012年6月

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，因此在本章的 1.1～1.3 节中仅对线性空间和线性映射的概念作简单介绍，1.4 节讲解子空间的概念.

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 $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，存在唯一的元素 $\boldsymbol{\gamma}\in V$ 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 $\lambda\in F$，存在唯一的元素 $\boldsymbol{\delta}\in V$ 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 $(\lambda+\mu)\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\alpha}$；

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

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为何，统称为向量.

注 (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

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

则 $P[x]_n$ 按着通常的多项式加法及数与多项式的乘法构成线性空间.

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 $\mathbb{R}$ 上的线性空间.

例 1.2 正实数全体记作$\mathbb{R}^+$,即

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

在$\mathbb{R}^+$中定义加法及数乘为

$$a\oplus b\overset{\text{def}}{=\!=}ab,\quad \lambda\circ a\overset{\text{def}}{=\!=}a^\lambda,$$

其中 $a,b\in\mathbb{R}^+,\lambda\in\mathbb{R}$. 证明$\mathbb{R}^+$对上述加法和数乘构成$\mathbb{R}$ 上的线性空间.

证 实际上要验证 10 条.

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

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

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

(2) $(a\oplus b)\oplus c=(ab)\oplus c=(ab)c=a(bc)=a\oplus(bc)=a\oplus(b\oplus c)$;

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

(4) 对于任何 $a\in\mathbb{R}^+$,有负元素 $a^{-1}\in\mathbb{R}^+$,使 $a\oplus a^{-1}=aa^{-1}=1$;

(5) 存在单位数 $1\in\mathbb{R}$,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 $a,b,c\in\mathbb{R}^+,\lambda,\mu\in\mathbb{R}$.

因此$\mathbb{R}^+$对所定义的加法"$\oplus$"、数乘"$\circ$"构成$\mathbb{R}$ 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 按通常的向量加法与数乘,$\mathbb{R}^n$ 是实数域上的线性空间,$\mathbb{C}^n$ 是复数域上的线性空间,但$\mathbb{R}^n$ 不是复数域$\mathbb{C}$ 上的线性空间.

线性空间具有以下简单性质.

性质 1 零元素唯一,负元素唯一.

证 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

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

由于负元素的唯一性,可记向量 $\boldsymbol{X}$ 的负元素为$-\boldsymbol{X}$.

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

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

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

(3) $\lambda\boldsymbol{0}=\boldsymbol{0}$;

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

证 因为 $\forall \boldsymbol{X}\in V(F)$,有

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

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

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

(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,则即满足规律③,④,因此有以下定理.

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 中的运算封闭.

1.2 线性空间的维数、基与坐标

在$\mathbb{R}^n$ 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$\mathbb{R}^n$ 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.

为了方便,以后将 $i=1,2,\cdots,n$ 记为 $i\in\underline{n}$.

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

与$\mathbb{R}^n$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，若

(1) $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$是线性无关向量组；

(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$n\in\mathbb{Z}^+$($\mathbb{Z}^+$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. 坐标也可以表示为向量形式$(k_1,k_2,\cdots,k_n)^{\mathrm{T}}$.

例 1.4 在$\mathbb{R}^n$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$\mathbb{R}^n$的基，易证向量组

$$\boldsymbol{e}_1'=(1,1,\cdots,1)^{\mathrm{T}},\boldsymbol{e}_2'=(0,1,\cdots,1)^{\mathrm{T}},\cdots,\boldsymbol{e}_n'=(0,\cdots,0,1)^{\mathrm{T}}$$

也是线性无关的，故它也是$\mathbb{R}^n$的基. 由此可见，线性空间的基不唯一.

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

$$p(x)=a+a_1x+a_2x^2+\cdots+a_nx^n,$$

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$\mathbb{R}$中的常数，则由泰勒公式知

$$p(x)=p(a)+p'(a)(x-a)+\frac{p''(a)}{2!}(x-a)^2+\cdots+\frac{p^{(n)}(a)}{n!}(x-a)^n,$$

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，元素的坐标由基唯一确定，当基改变时，坐标将随之改变. 那么，如果$V_n(F)$的两个基之间的关系已知，如何由此推断$V_n(F)$的元素关于这两个基的坐

标之间的关系呢?

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

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

若记

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

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$\boldsymbol{A}$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

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

再由坐标表示式的唯一性知式(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$\mathbb{R}^n$的两个基(见本节例 1.4),若$\boldsymbol{X}\in\mathbb{R}^n$,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

$$[e_1', e_2', \cdots, e_n'] = [e_1, e_2, \cdots, e_n]\boldsymbol{A},$$

其中

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & \ddots & 0 \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

为过渡矩阵，且

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix},$$

代入式(1-7)得 $y_1 = x_1, y_2 = x_2 - x_1, y_3 = x_3 - x_2, \cdots, y_n = x_n - x_{n-1}$.

例 1.7 在$\mathbb{R}^4$ 中，求$\boldsymbol{\alpha} = (1,2,1,1)^{\mathrm{T}}$ 在基$\boldsymbol{\varepsilon}_1 = (1,1,1,1)^{\mathrm{T}}, \boldsymbol{\varepsilon}_2 = (1,1,-1,-1)^{\mathrm{T}}, \boldsymbol{\varepsilon}_3 = (1,-1,1,-1)^{\mathrm{T}}, \boldsymbol{\varepsilon}_4 = (1,-1,-1,1)^{\mathrm{T}}$ 下的坐标.

解 设$\boldsymbol{\alpha} = (1,2,1,1)^{\mathrm{T}}$ 在所给基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4$下的坐标为 k_1, k_2, k_3, k_4，则有

$$\boldsymbol{\alpha} = k_1 \boldsymbol{\varepsilon}_1 + k_2 \boldsymbol{\varepsilon}_2 + k_3 \boldsymbol{\varepsilon}_3 + k_4 \boldsymbol{\varepsilon}_4,$$

即

$$\begin{aligned}(1,2,1,1)^{\mathrm{T}} &= k_1(1,1,1,1)^{\mathrm{T}} + k_2(1,1,-1,-1)^{\mathrm{T}} \\ &\quad + k_3(1,-1,1,-1)^{\mathrm{T}} + (1,-1,-1,1)^{\mathrm{T}},\end{aligned}$$

所以

$$\begin{cases} k_1 + k_2 + k_3 + k_4 = 1, \\ k_1 + k_2 - k_3 - k_4 = 2, \\ k_1 - k_2 + k_3 - k_4 = 1, \\ k_1 - k_2 - k_3 + k_4 = 1, \end{cases}$$

解得

$$k_1 = \frac{5}{4}, \quad k_2 = \frac{1}{4}, \quad k_3 = -\frac{1}{4}, \quad k_4 = -\frac{1}{4}.$$

所以$\boldsymbol{\alpha}$ 在所给基$\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3, \boldsymbol{\varepsilon}_4$下的坐标为

$$(5/4, 1/4, -1/4, -1/4)^{\mathrm{T}}.$$

例 1.8 在$\mathbb{R}^{2\times 2}$中，求 $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$在基

$$A_1=\begin{bmatrix}1&1\\1&1\end{bmatrix},\quad A_2=\begin{bmatrix}1&1\\1&0\end{bmatrix},\quad A_3=\begin{bmatrix}1&1\\0&1\end{bmatrix},\quad A_4=\begin{bmatrix}1&0\\1&1\end{bmatrix}$$

下的坐标.

解 令

$$A=k_1A_1+k_2A_2+k_3A_3+k_4A_4,$$

即

$$\begin{bmatrix}1&2\\1&0\end{bmatrix}=k_1\begin{bmatrix}1&1\\1&1\end{bmatrix}+k_2\begin{bmatrix}1&1\\1&0\end{bmatrix}+k_3\begin{bmatrix}1&1\\0&1\end{bmatrix}+k_4\begin{bmatrix}1&0\\1&1\end{bmatrix}$$
$$=\begin{bmatrix}k_1+k_2+k_3+k_4 & k_1+k_2+k_3\\ k_1+k_2+k_4 & k_1+k_3+k_4\end{bmatrix},$$

得

$$\begin{cases}k_1+k_2+k_3+k_4=1,\\k_1+k_2+k_3=2,\\k_1+k_2+k_4=1,\\k_1+k_3+k_4=0,\end{cases}$$

求得

$$k_1=1,\quad k_2=1,\quad k_3=0,\quad k_4=-1,$$

所以 A 在基 A_1,A_2,A_3,A_4 下的坐标为 $(1,1,0,-1)^{\mathrm T}$.

例 1.9 在 $\mathbb{R}^4$ 中,求由基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 到基 $\beta_1,\beta_2,\beta_3,\beta_4$ 的过渡矩阵,其中

$$\begin{cases}\alpha_1=(1,2,-1,0)^{\mathrm T},\\\alpha_2=(1,-1,1,1)^{\mathrm T},\\\alpha_3=(-1,2,1,1)^{\mathrm T},\\\alpha_4=(-1,-1,0,1)^{\mathrm T},\end{cases}\qquad\begin{cases}\beta_1=(2,1,0,1)^{\mathrm T},\\\beta_2=(0,1,2,2)^{\mathrm T},\\\beta_3=(-2,1,1,2)^{\mathrm T},\\\beta_4=(1,3,1,2)^{\mathrm T},\end{cases}$$

并求 $\xi=(1,-1,-1,1)^{\mathrm T}$ 在基 $\beta_1,\beta_2,\beta_3,\beta_4$ 下的坐标.

解 设

$$\varepsilon_1=(1,0,0,0)^{\mathrm T},\quad \varepsilon_2=(0,1,0,0)^{\mathrm T},$$
$$\varepsilon_3=(0,0,1,0)^{\mathrm T},\quad \varepsilon_4=(0,0,0,1)^{\mathrm T},$$

则

$$\alpha_1=\varepsilon_1+2\varepsilon_2-\varepsilon_3=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\begin{bmatrix}1\\2\\-1\\0\end{bmatrix},$$

$$\alpha_2=\varepsilon_1-\varepsilon_2+\varepsilon_3+\varepsilon_4=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)\begin{bmatrix}1\\-1\\1\\1\end{bmatrix},$$

$$\boldsymbol{\alpha}_3=-\boldsymbol{\varepsilon}_1+2\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3+\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}-1\\2\\1\\1\end{bmatrix},$$

$$\boldsymbol{\alpha}_4=-\boldsymbol{\varepsilon}_1-\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}-1\\-1\\0\\1\end{bmatrix},$$

所以

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{bmatrix}. \tag{1-8}$$

同理

$$\boldsymbol{\beta}_1=2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}2\\1\\0\\1\end{bmatrix},$$

$$\boldsymbol{\beta}_2=\boldsymbol{\varepsilon}_2+2\boldsymbol{\varepsilon}_3+2\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}0\\1\\2\\2\end{bmatrix},$$

$$\boldsymbol{\beta}_3=-2\boldsymbol{\varepsilon}_1+\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3+2\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}-2\\1\\1\\2\end{bmatrix},$$

$$\boldsymbol{\beta}_4=\boldsymbol{\varepsilon}_1+3\boldsymbol{\varepsilon}_2+\boldsymbol{\varepsilon}_3+2\boldsymbol{\varepsilon}_4=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}1\\3\\1\\2\end{bmatrix},$$

故

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}, \tag{1-9}$$

把式(1-8)代入式(1-9)得

$$(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4)=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}$$

$$=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{bmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{bmatrix}^{-1}\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}$$

$$=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\frac{1}{13}\begin{bmatrix}3&2&-6&5\\5&-1&3&4\\-2&3&4&1\\-3&-2&-7&8\end{bmatrix}\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}$$

$$=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)\begin{bmatrix}1&0&0&1\\1&1&0&1\\0&1&1&1\\0&0&1&0\end{bmatrix},$$

所以$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$到$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$的过渡矩阵为

$$\boldsymbol{T}=\begin{bmatrix}1&0&0&1\\1&1&0&1\\0&1&1&1\\0&0&1&0\end{bmatrix}.$$

设$\boldsymbol{\xi}$在$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4$下的坐标为$\boldsymbol{Y}=(y_1,y_2,y_3,y_4)^{\mathrm{T}}$,即

$$\boldsymbol{\xi}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3,\boldsymbol{\beta}_4)\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix},$$

由式(1-9)知

$$\boldsymbol{\xi}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\boldsymbol{\varepsilon}_3,\boldsymbol{\varepsilon}_4)\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix},$$

$$\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}=\begin{bmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{bmatrix}^{-1}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}=\frac{1}{13}\begin{bmatrix}4&-6&-8&11\\2&-3&9&-1\\-3&-2&-7&8\\-1&8&2&-6\end{bmatrix}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}$$