

研究生用书

Second Edition

Basic  
Matrix  
Analysis

# 矩阵分析基础

王永茂 刘德友 李德生 编著 (第2版)

清华大学出版社



Second Edition

Basic  
Matrix  
Analysis

# 矩阵分析基础

王永茂 刘德友 李德生 编著 (第2版)

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.

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究.侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

矩阵分析基础/王永茂,刘德友,李德生编著.--2版.--北京:清华大学出版社,2016  
ISBN 978-7-302-44991-1

I. ①矩… II. ①王… ②刘… ③李 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(2016)第216152号

责任编辑:陈明

封面设计:傅瑞学

责任校对:王淑云

责任印制:王静怡

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](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

印装者:北京密云胶印厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:12.75

字 数:308千字

版 次:2012年9月第1版

2016年9月第2版

印 次:2016年9月第1次印刷

印 数:1~2000

定 价:28.00元

产品编号:068663-01

第2版

言

前

本书自2012年出版以来,经历了多所高校三年试用实践的检验.在此期间,我们走访了一些使用过该书的学生和教师,这些学生和教师对本书提出了许多宝贵的意见.依据这些意见和建议,参照理工科学生对矩阵理论教学的基本要求,展望数学教育改革的走向,作者对第1版进行了修改和补充.第2版除了对第1版的部分内容和习题进行了补充和修改外,还增设了“自学园地”和“习题答案或提示”.

第2版的三位作者都是燕山大学教师,均从事研究生教育多年,该书是作者经过多年教学实践,几经修改、补充编写而成.与其他同类书相比,除在结构、形式上有所不同外,在内容上力求更加精炼、适应面更宽、可供选择性更强,在例题的选取、内容的讲解上力求深入浅出、通俗易懂.在自学园地中,对一些典型的有代表性的例题进行了详细讲解,这样更便于学生课后复习和自学,也便于教师在讲课时进行选择.对习题中的计算题给出了参考答案,对证明题给出了提示.本书除第1、2章必学外,其他各章内容相对较独立,各类不同专业可根据需要进行选取.

本书在编写的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,此次出版得到了燕山大学研究生院领导的大力支持和燕山大学研究生课程建设经费资助,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

2016年于燕山大学

第1版

言

前

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

|                               |           |
|-------------------------------|-----------|
| <b>第 1 章 线性空间与线性变换</b> .....  | <b>1</b>  |
| 1.1 线性空间及其性质 .....            | 1         |
| 1.2 线性空间的维数、基与坐标 .....        | 3         |
| 1.3 线性映射与线性变换 .....           | 10        |
| 1.3.1 线性映射与线性变换的定义和性质 .....   | 10        |
| 1.3.2 线性变换的特征值和特征向量 .....     | 14        |
| 1.4 线性子空间 .....               | 15        |
| 1.5 自学园地 .....                | 19        |
| 习题 1 .....                    | 25        |
| <b>第 2 章 酉空间和酉变换</b> .....    | <b>29</b> |
| 2.1 酉空间和欧氏空间 .....            | 29        |
| 2.2 向量的正交与标准正交基 .....         | 32        |
| 2.3 酉(正交)变换 .....             | 36        |
| * 2.4 几种特殊的子空间 .....          | 39        |
| 2.4.1 子空间的同构 .....            | 39        |
| 2.4.2 不变子空间 .....             | 40        |
| 2.4.3 正交子空间 .....             | 41        |
| 2.5 自学园地 .....                | 43        |
| 习题 2 .....                    | 46        |
| <b>第 3 章 矩阵的分解</b> .....      | <b>49</b> |
| 3.1 若尔当(Jordan)型分解 .....      | 49        |
| 3.1.1 $\lambda$ -矩阵及其性质 ..... | 49        |
| 3.1.2 $n$ 阶方阵的若尔当标准形 .....    | 54        |
| * 3.1.3 单纯矩阵的谱分解 .....        | 61        |
| 3.2 $n$ 阶方阵的三角分解 .....        | 62        |
| 3.2.1 矩阵的三角分解 .....           | 62        |
| 3.2.2 三角分解的应用 .....           | 64        |

|              |                                            |            |
|--------------|--------------------------------------------|------------|
| 3.3          | 埃尔米特矩阵及其分解                                 | 64         |
| 3.4          | 矩阵的最大秩分解                                   | 70         |
| * 3.5        | 矩阵的奇异值分解                                   | 73         |
| 3.6          | 自学园地                                       | 75         |
|              | 习题 3                                       | 82         |
| <b>第 4 章</b> | <b>范数及其应用</b>                              | <b>85</b>  |
| 4.1          | 向量范数                                       | 85         |
| 4.2          | 矩阵范数                                       | 88         |
| 4.3          | 算子范数                                       | 90         |
| * 4.4        | 矩阵范数的推广                                    | 94         |
| 4.5          | 范数的应用                                      | 96         |
| 4.6          | 自学园地                                       | 98         |
|              | 习题 4                                       | 100        |
| <b>第 5 章</b> | <b>矩阵分析</b>                                | <b>101</b> |
| 5.1          | 矩阵级数                                       | 101        |
| 5.2          | 矩阵的微分                                      | 105        |
| 5.2.1        | 对于数量变量的微分法                                 | 105        |
| 5.2.2        | 对于向量变量的微分法                                 | 107        |
| 5.2.3        | 对于矩阵变量的微分法                                 | 112        |
| 5.2.4        | 复合函数的微分法                                   | 114        |
| 5.3          | 矩阵的积分                                      | 115        |
| * 5.4        | 微分理论的应用                                    | 116        |
| 5.4.1        | 矩阵微分方程                                     | 116        |
| 5.4.2        | 线性向量微分方程                                   | 118        |
| 5.5          | 自学园地                                       | 120        |
|              | 习题 5                                       | 123        |
| <b>第 6 章</b> | <b>矩阵函数</b>                                | <b>125</b> |
| 6.1          | 矩阵多项式                                      | 125        |
| 6.2          | 矩阵函数的定义及性质                                 | 129        |
| 6.3          | $f(\mathbf{A})$ 用若尔当标准形表示(标准形 I)           | 131        |
| 6.4          | $f(\mathbf{A})$ 用拉格朗日-西尔维斯特内插多项式表示(标准形 II) | 133        |
| 6.5          | $f(\mathbf{A})$ 用有限级数表示(标准形 III)           | 136        |
| 6.6          | 自学园地                                       | 139        |
|              | 习题 6                                       | 142        |
| <b>第 7 章</b> | <b>广义逆矩阵</b>                               | <b>144</b> |
| 7.1          | 广义逆矩阵及其性质                                  | 144        |

|              |                        |            |
|--------------|------------------------|------------|
| 7.2          | 自反广义逆矩阵 $A^-$ .....    | 148        |
| 7.3          | 伪逆矩阵 $A^+$ .....       | 150        |
| 7.4          | 广义逆矩阵的应用 .....         | 154        |
| 7.5          | 自学园地 .....             | 160        |
|              | 习题 7 .....             | 166        |
| <b>第 8 章</b> | <b>矩阵的扰动问题简介 .....</b> | <b>168</b> |
| 8.1          | 特征值问题的稳定性 .....        | 168        |
| 8.2          | 盖尔斯高林圆盘定理 .....        | 171        |
| 8.3          | 矩阵逆与线性方程组解的扰动 .....    | 175        |
|              | 8.3.1 矩阵逆的扰动界限 .....   | 176        |
|              | 8.3.2 方程组的扰动问题 .....   | 177        |
|              | 习题 8 .....             | 179        |
|              | <b>习题参考答案或提示 .....</b> | <b>181</b> |
|              | <b>参考文献 .....</b>      | <b>192</b> |



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 节~1.3 节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  $\alpha, \beta \in V$ ,存在唯一的元素  $\gamma \in V$  与之对应,称为  $\alpha$  与  $\beta$  的和,记作  $\alpha + \beta = \gamma$ ;

(2) 在  $V$  中定义了数乘,使对于任意的  $\alpha \in V$  及任意  $\lambda \in F$ ,存在唯一的元素  $\delta \in V$  与之对应,称为  $\lambda$  与  $\alpha$  的积,记作  $\delta = \lambda\alpha$ ;

(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

$$\textcircled{1} \alpha + \beta = \beta + \alpha;$$

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

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

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

$\textcircled{5}$  存在单位数  $1 \in F$ ,使得  $1\alpha = \alpha$ ;

$$\textcircled{6} \lambda(\mu\alpha) = (\lambda\mu)\alpha;$$

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

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

其中  $\alpha, \beta, \gamma \in V, \lambda, \mu \in F$ .

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** (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

$$P[x]_n = \{p_n \mid p_n = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, a_n, a_{n-1}, \cdots, a_0 \in \mathbb{R}\},$$

则  $P[x]_n$  按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,且满足定义 1.1 中 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

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

在 $\mathbb{R}^+$ 中定义加法及数乘为

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

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 $\lambda \circ a = a^\lambda \in \mathbb{R}^+$ ;

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

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

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

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

$$(5) \text{存在单位数 } 1 \in \mathbb{R}, \text{ 使 } 1 \circ a = a^1 = a;$$

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

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

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

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 $\mathbb{R}$ 上的线性空间,  $\mathbb{C}^n$  是复数域 $\mathbb{C}$ 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 $\mathbf{X}_1, \mathbf{X}_2$ 均为 $\mathbf{X} \in V(F)$ 的负元素,  $\mathbf{0}$ 为 $V(F)$ 的零元素, 则

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

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

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

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

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

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

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

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

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

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

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

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

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

这与  $\mathbf{X} \neq \mathbf{0}$  矛盾, 故  $\lambda \neq 0$  与  $\mathbf{X} \neq \mathbf{0}$  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, 定义 1.1 中规律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  $i=1, 2, \dots, n$  记为  $i \in \underline{n}$ .

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

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

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

**定义 1.4**  $V(F)$  中的向量组  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$  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

$$\sum_{i=1}^m k_i\mathbf{X}_i = \mathbf{0}; \quad (1-1)$$

否则称向量组  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$  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  $m \geq 2$  时,  $V(F)$  中的向量组  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$  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  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  称为  $V(F)$  的基, 若

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

(2)  $V(F)$  中任一向量均可由向量组  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$  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  $X_1, X_2, \dots, X_n$  为  $V_n(F)$  的基, 则对  $X \in V_n(F)$ , 其表达式

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

中的  $k_i, i \in \bar{n}$  称为向量  $X$  关于基  $X_1, X_2, \dots, X_n$  的坐标. 坐标也可以表示为向量形式  $(k_1, k_2, \dots, k_n)^T$ .

**例 1.4** 在  $\mathbb{R}^n$  中, 向量组  $e_1 = (1, 0, \dots, 0)^T, e_2 = (0, 1, 0, \dots, 0)^T, \dots, e_n = (0, \dots, 0, 1)^T$  是线性无关的, 它是  $\mathbb{R}^n$  的基, 易证向量组

$$e'_1 = (1, 1, \dots, 1)^T, e'_2 = (0, 1, \dots, 1)^T, \dots, e'_n = (0, \dots, 0, 1)^T$$

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

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

是  $P[x]_n$  的一个基, 若  $p(x) \in P[x]_n$ , 且

$$p(x) = a + a_1 x + a_2 x^2 + \dots + a_n x^n,$$

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

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

为  $P[x]_n$  的基, 其中  $a \in \mathbb{R}$  为常数, 则由泰勒公式知

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

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

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  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  与  $\epsilon'_1, \epsilon'_2, \dots, \epsilon'_n$  为  $V_n(F)$  的两个基, 且

$$\begin{cases} \epsilon'_1 = a_{11}\epsilon_1 + a_{21}\epsilon_2 + \dots + a_{n1}\epsilon_n, \\ \epsilon'_2 = a_{12}\epsilon_1 + a_{22}\epsilon_2 + \dots + a_{n2}\epsilon_n, \\ \vdots \\ \epsilon'_n = a_{1n}\epsilon_1 + a_{2n}\epsilon_2 + \dots + a_{nn}\epsilon_n, \end{cases} \quad (1-2)$$

若记

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

则式(1-2)可表示为

$$[\epsilon'_1, \epsilon'_2, \dots, \epsilon'_n] = [\epsilon_1, \epsilon_2, \dots, \epsilon_n]A, \quad (1-3)$$

或

$$\begin{bmatrix} (\epsilon'_1)^T \\ (\epsilon'_2)^T \\ \vdots \\ (\epsilon'_n)^T \end{bmatrix} = A^T \begin{bmatrix} \epsilon_1^T \\ \epsilon_2^T \\ \vdots \\ \epsilon_n^T \end{bmatrix}, \quad (1-4)$$

其中矩阵  $A$  称为由基  $\epsilon_1, \epsilon_2, \dots, \epsilon_n$  到基  $\epsilon'_1, \epsilon'_2, \dots, \epsilon'_n$  的过渡矩阵或基变换矩阵.

**定理 1.2** 设  $[\epsilon'_1, \epsilon'_2, \dots, \epsilon'_n] = [\epsilon_1, \epsilon_2, \dots, \epsilon_n]A$ , 若  $X \in V(F)$ , 有

$$X = x_1\epsilon_1 + x_2\epsilon_2 + \dots + x_n\epsilon_n = y_1\epsilon'_1 + y_2\epsilon'_2 + \dots + y_n\epsilon'_n, \quad (1-5)$$

则

$$(x_1, x_2, \dots, x_n)^T = A(y_1, y_2, \dots, y_n)^T, \quad (1-6)$$

或

$$(y_1, y_2, \dots, y_n)^T = A^{-1}(x_1, x_2, \dots, x_n)^T. \quad (1-7)$$

**证** 由  $X$  的表达式(1-5)有

$$\begin{aligned} X &= [\epsilon_1, \epsilon_2, \dots, \epsilon_n](x_1, x_2, \dots, x_n)^T = [\epsilon'_1, \epsilon'_2, \dots, \epsilon'_n](y_1, y_2, \dots, y_n)^T \\ &= [\epsilon_1, \epsilon_2, \dots, \epsilon_n]A(y_1, y_2, \dots, y_n)^T. \end{aligned}$$

再由坐标表示式的唯一性知式(1-6)成立. 又  $A$  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  $A$  已知时,  $X$  关于两个基的坐标之间的关系.

**例 1.6** 设  $e_1, e_2, \dots, e_n$  与  $e'_1, e'_2, \dots, e'_n$  为  $\mathbb{R}^n$  的两个基(见本节例 1.4), 若  $X \in \mathbb{R}^n$ , 且

$$X = x_1e_1 + x_2e_2 + \dots + x_n e_n = y_1e'_1 + y_2e'_2 + \dots + y_n e'_n,$$

求  $e_1, e_2, \dots, e_n$  到  $e'_1, e'_2, \dots, e'_n$  的过渡矩阵  $A$  及  $X$  在两个基下的坐标之间的关系.

**解** 显然有

$$[e'_1, e'_2, \dots, e'_n] = [e_1, e_2, \dots, e_n]A,$$

其中

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

为过渡矩阵, 且

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

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

**例 1.7** 在  $\mathbb{R}^4$  中, 求  $\alpha = (1, 2, 1, 1)^T$  在基  $\epsilon_1 = (1, 1, 1, 1)^T, \epsilon_2 = (1, 1, -1, -1)^T, \epsilon_3 = (1, -1, 1, -1)^T, \epsilon_4 = (1, -1, -1, 1)^T$  下的坐标.

**解** 设  $\alpha = (1, 2, 1, 1)^T$  在所给基  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$  下的坐标为  $k_1, k_2, k_3, k_4$ , 则有

$$\alpha = k_1 \epsilon_1 + k_2 \epsilon_2 + k_3 \epsilon_3 + k_4 \epsilon_4,$$

即

$$\begin{aligned} (1, 2, 1, 1)^T &= k_1(1, 1, 1, 1)^T + k_2(1, 1, -1, -1)^T \\ &\quad + k_3(1, -1, 1, -1)^T + k_4(1, -1, -1, 1)^T, \end{aligned}$$

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  $\alpha$  在所给基  $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$  下的坐标为

$$\left(\frac{5}{4}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}\right)^T.$$

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

$$A = k_1 A_1 + k_2 A_2 + k_3 A_3 + k_4 A_4,$$

即

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

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  $(1, 1, 0, -1)^T$ .

**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

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

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

**解法 1** 令

$$A = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \quad B = (\beta_1, \beta_2, \beta_3, \beta_4) = \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix},$$

设由基  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  到基  $\beta_1, \beta_2, \beta_3, \beta_4$  的过渡矩阵为  $T$ , 则应有

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

即

$$B = AT,$$

所以

$$\begin{aligned} T = A^{-1}B &= \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \\ &= \frac{1}{13} \begin{bmatrix} 3 & 2 & -6 & 5 \\ 5 & -1 & 3 & 4 \\ -2 & 3 & 4 & 1 \\ -3 & -2 & -7 & 8 \end{bmatrix} \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \end{aligned}$$

故  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  到  $\beta_1, \beta_2, \beta_3, \beta_4$  的过渡矩阵为

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

设  $\xi$  在  $\beta_1, \beta_2, \beta_3, \beta_4$  下的坐标为  $Y = (y_1, y_2, y_3, y_4)^T$ , 即

$$\xi = (\beta_1, \beta_2, \beta_3, \beta_4)Y = BY,$$

则

$$Y = B^{-1}\xi = \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{13} \begin{bmatrix} 4 & -6 & -8 & 11 \\ 2 & -3 & 9 & -1 \\ -3 & -2 & -7 & 8 \\ -1 & 8 & 2 & -6 \end{bmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{13} \begin{pmatrix} 3 \\ -5 \\ 14 \\ -17 \end{pmatrix}.$$

解法 2 设

$$\epsilon_1 = (1, 0, 0, 0)^T, \quad \epsilon_2 = (0, 1, 0, 0)^T,$$

$$\epsilon_3 = (0, 0, 1, 0)^T, \quad \epsilon_4 = (0, 0, 0, 1)^T,$$

则

$$\alpha_1 = \epsilon_1 + 2\epsilon_2 - \epsilon_3 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix},$$

$$\alpha_2 = \epsilon_1 - \epsilon_2 + \epsilon_3 + \epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix},$$

$$\alpha_3 = -\epsilon_1 + 2\epsilon_2 + \epsilon_3 + \epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} -1 \\ 2 \\ 1 \\ 1 \end{bmatrix},$$

$$\alpha_4 = -\epsilon_1 - \epsilon_2 + \epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix},$$

所以

$$(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}. \quad (1-8)$$

同理

$$\beta_1 = 2\epsilon_1 + \epsilon_2 + \epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix},$$

$$\beta_2 = \epsilon_2 + 2\epsilon_3 + 2\epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix},$$

$$\beta_3 = -2\epsilon_1 + \epsilon_2 + \epsilon_3 + 2\epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} -2 \\ 1 \\ 1 \\ 2 \end{bmatrix},$$

$$\beta_4 = \epsilon_1 + 3\epsilon_2 + \epsilon_3 + 2\epsilon_4 = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \end{bmatrix},$$

故

$$(\beta_1, \beta_2, \beta_3, \beta_4) = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix}, \quad (1-9)$$



把式(1-8)代入式(1-9)得

$$\begin{aligned}
 (\beta_1, \beta_2, \beta_3, \beta_4) &= (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \\
 &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & -1 & 2 & -1 \\ -1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \\
 &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \frac{1}{13} \begin{bmatrix} 3 & 2 & -6 & 5 \\ 5 & -1 & 3 & 4 \\ -2 & 3 & 4 & 1 \\ -3 & -2 & -7 & 8 \end{bmatrix} \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \\
 &= (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},
 \end{aligned}$$

所以 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 到 $\beta_1, \beta_2, \beta_3, \beta_4$ 的过渡矩阵为

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

设 $\xi$ 在 $\beta_1, \beta_2, \beta_3, \beta_4$ 下的坐标为 $Y = (y_1, y_2, y_3, y_4)^T$ , 即

$$\xi = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = (\beta_1, \beta_2, \beta_3, \beta_4) \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix},$$

由式(1-9)知

$$\begin{aligned}
 \xi &= (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \\
 \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} &= \begin{bmatrix} 2 & 0 & -2 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{13} \begin{bmatrix} 4 & -6 & -8 & 11 \\ 2 & -3 & 9 & -1 \\ -3 & -2 & -7 & 8 \\ -1 & 8 & 2 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \\
 &= \frac{1}{13} \begin{bmatrix} 3 \\ -5 \\ 14 \\ -17 \end{bmatrix}.
 \end{aligned}$$