矩阵理论及其应用

徐英凯 李国东 邱启荣 编著

中国电力出版社
www.cepp.com.cn

矩阵理论及其应用

徐英凯 李国东 邱启荣 编著

中国电力出版社
www.cepp.com.cn

内容提要

矩阵理论不但是学习经典数学的基础，也是研究现代工程技术的数学基础。例如系统工程、系统控制理论、信号处理、优化方法、稳定性理论、工程结构分析等，无不与矩阵理论发生紧密的结合。

全书共有8章，内容包含了矩阵理论研究和实际计算的基础知识。第1、2章是矩阵理论基础知识，主要介绍了线性空间、线性变换、矩阵特征值等基本概念、运算和基本理论，并讨论了一些特殊的常用矩阵。第3、4章的内容以实际计算中常用的和揭示矩阵性质的矩阵分解和标准形为线索，讲述了矩阵的Jordan标准形以及它的应用、矩阵的LU分解、满秩分解、奇异值分解等分解形式。第5章讲述了矩阵的度量及其应用，给出了定量刻划矩阵的度量工具——范数的定义及其性质的讨论，其中包括了一般性的矩阵范数定义以及应用广泛的特殊矩阵范数一相容矩阵范数和算子矩阵范数的定义。利用矩阵范数，对矩阵特征值、广义逆及最小二乘问题进行了简要的扰动分析。第6章的矩阵分析，主要讲述了矩阵序列和级数、矩阵的Kronecker积以及函数矩阵微积分等知识。第7章讲述了矩阵函数的定义以及各种计算方法。第8章讲述了广义逆的定义、性质和各种计算方法。

本书可作为工科研究生、管理类研究生教学的主要参考教材，也可作为工程技术人员、科技工作者和各工科专业教师更新基础知识和进修提高的参考书。

图书在版编目（CIP）数据

矩阵理论及其应用/徐英凯，李国东，邱启荣编著．北京：中国电力出版社，2003

ISBN 7-5083-1721-1

Ⅰ.矩… Ⅱ.①徐…②李…③邱… Ⅲ.矩阵-理论 Ⅳ.0151.21

中国版本图书馆CIP数据核字（2003）第073536号

中国电力出版社出版、发行

（北京三里河路6号 100044 http://www.cepp.com.cn）

北京通天印刷厂印刷

各地新华书店经售

*

2003年9月第一版 2003年9月北京第一次印刷

787毫米×1092毫米 16开本 10.5印张 231千字

印数0001—2500册 定价26.00元

前　言

矩阵理论既是学习经典数学的基础，又是一门最有实用价值的数学理论。它不仅是数学的一个重要的分支，而且业已成为现代各科技领域处理大量有限维空间形式与数量关系的强有力的工具，它不仅能使所描述的问题具有极简洁的形式，而且也是使所描述的问题得以深入系统研究的手段。特别是由于计算机和计算方法的普及和发展，不仅为矩阵理论的应用开辟了广阔的前景，也使工程技术的研究发生了新的变化，开拓了崭新的研究途径。例如系统工程、系统控制理论、信号处理、优化方法、稳定性理论、工程结构分析等，无不与矩阵理论发生紧密的结合。因此矩阵的理论和方法已经成为研究现代工程技术的数学基础。古典的线性代数理论已不能满足现代工程技术的需要。我们根据国家教委制定的工科研究生学习《矩阵论》课程的基本要求，编写此书。

本书编者结合各学科研究生的教学和科研工作的需要和多年的教学经验，力求在不失数学严谨性的同时深入浅出，并注意了与本科《线性代数》的衔接，同时也对同一问题给出了多种解法，其可读性强，便于大家学习和理解。

本书曾在华北电力大学多届研究生中试用，在广泛征求学生意见和专家意见的基础上，经过多次修改，并由徐英凯、李国东、邱启荣三位老师编写而成，其中徐英凯编写了第1~3章，李国东编写了第4、5章，邱启荣编写了第6~8章。华北电力大学校领导、研究生处、理学院领导对本书的编写与出版给予了大力支持和帮助，编者在此表示衷心的感谢。

由于编者水平有限，书中如有不妥之处，请国内同行与读者提出批评指正。

编者

2003年7月

目　录

第1章 线性空间

线性空间既是代数学的基本概念，也是矩阵论的基本概念之一，本章首先介绍这一概念。学习过这一部分内容的读者可以将本章略过，或仅作为回顾浏览。

§1.1 集合与映射

人们经常使用“一组”、“一队”、“一批”这样的词汇来描述某类事物，它们被用来表示一定事物的集体。在数学上称它们为**集合**或**集**。组成集合的东西称为集合的**元素**。

我们用大写字母 A，B，C，…表示集合，用小写字母 a，b，c，…表示集合的元素。

如果 a 是集合 A 的元素，就称 a 属于 A，记作 $a\in A$；如果 a 不是集合 A 的元素，就称 a 不属于 A，记作 $a\notin A$。

若集合 A 中仅含有有限个元素，则称其为**有限集合**，比如由 -1，1 这两个元素组成的集合就是一个有限集合。

若集合 A 中含有无限个元素，则称其为**无限集合**，比如由全体实数组成的集合就是一个无限集合。

集合可以用列出其所含有的全部元素，或者给出集合中元素所具备的特征的方式来表示。当然前一种表示方式仅对有限集合适用，而无限集合的表示要用第二种方式。

例如，由 -1，1 这两个元素组成的集合可以表示为 $A=\{-1,1\}$，而由全体实数组成的集合可以表示为 $A=\{x\mid x\in \boldsymbol{R}\}$，当然第一个集合也可以表示为 $A=\{x\mid x^2-1=0, x\in \boldsymbol{R}\}$。

不含有任何元素的集合称为**空集合**，记为 Φ。例如集合 $A=\{x\mid x^2+1=0, x\in \boldsymbol{R}\}$ 就是一个空集合。

为了方便起见，我们约定：

$\boldsymbol{N}$ 为全体自然数组成的集合；

$\boldsymbol{Z}$ 为全体整数组成的集合；

$\boldsymbol{Q}$ 为全体有理数组成的集合；

$\boldsymbol{R}$ 为全体实数组成的集合；

$\boldsymbol{C}$ 为全体复数组成的集合。

设 A，B 是两个集合，如果 A 中的每一个元素都是 B 的元素，则称 A 是 B 的子集，记作 $A\subseteq B$，或记作 $B\supseteq A$。

根据这样的定义，一个集合 A 是它自身的子集，同时规定空集合 Φ 是任意集合的子集，这两个子集称为 A 的当然子集，而 A 的其他子集称为 A 的真子集。

设 A，B 是两个集合，如果 $A\subseteq B$，并且 $B\subseteq A$，则称这两个集合相等，记作 $A=B$。

设 A，B 是两个集合，由 A 的所有元素和 B 的所有元素组成的集合称为 A 与 B 的**并集**，记作 $A\cup B$，即 $A\cup B=\{x\mid x\in A$ 或 $x\in B\}$。

显然，$A\subseteq A\cup B$，$B\subseteq A\cup B$。

由 A 和 B 的公共元素组成的集合称为 A 与 B 的**交集**，记作 $A\cap B$，即 $A\cap B=\{x\mid x\in A$ 且 $x\in B\}$。

显然 $A\cap B\subseteq A$，$A\cap B\subseteq B$。

类似地，设给定 n 个集合 A_1，A_2，$\cdots$，A_n。由 A_1，A_2，$\cdots$，A_n 的所有元素组成的集合称为 A_1，A_2，$\cdots$，A_n 的并集；而由 A_1，A_2，$\cdots$，A_n 的公共元素组成的集合称为 A_1，A_2，$\cdots$，A_n 的交集，分别记为 $A_1\cup A_2\cup\cdots\cup A_n$ 和 $A_1\cap A_2\cap\cdots\cap A_n$。

设 A，B 是两个集合，由一切属于 A 但不属于 B 的元素组成的集合称为 B 在 A 中的**余集**，或称为 A 与 B 的**差**，记作 $A-B$，即 $A-B=\{x\mid x\in A$ 但 $x\notin B\}$。

例如，$\boldsymbol{R}-\boldsymbol{Q}$ 就是全体无理数的集合。

设 A，B 是两个集合，集合 $A+B=\{a+b\mid a\in A,\ b\in B\}$ 称为 A 与 B 的**和集**。要注意和集与集合的并是两个完全不同的概念。

设 A，B 是两个集合，集合 $A\times B=\{(a,\ b)\mid a\in A,\ b\in B\}$ 称为 A 与 B 的**积**。

以上我们对集合进行了简单的讨论。

另一个常用的基本概念是**映射**，以下来讨论这个概念和它的一些性质。

定义 1.1.1 设 A，B 是两个非空集合，A 到 B 的一个映射是指一个对应法则，通过这一法则，对于集合 A 中的每一个元素 x，都有集合 B 中的一个唯一确定的元素 y 与之对应。

我们用字母 f，g，σ，μ，$\cdots$表示映射，用记号 $f: A\to B$ 表示 f 是 A 到 B 的一个映射。

如果通过映射 f，与 A 中元素 x 对应的 B 中元素是 y，则记作 $f: x\mapsto y$；或 $f(x)=y$。y 叫做元素 x 在 f 下的象，x 叫做 y 在 f 下的原象。

A 在 f 下的象的集合记作 $f(A)=\{f(x)\mid x\in A\}$。

如果对于每一个 $x\in A$，$f(x)$ 都已给出，那么映射 f 就被确定了。

【例 1-1】 $\boldsymbol{R}^{n\times n}$ 是全体 n 阶实矩阵 A 组成的集合，令 $f(A)=\det A$，则 $\forall A\in\boldsymbol{R}^{n\times n}$，通过法则 f，在 $\boldsymbol{R}$ 中有唯一确定的实数 $x=\det A$ 与之对应，所以 $f: \boldsymbol{R}^{n\times n}\to\boldsymbol{R}$ 是一个映射。

【例 1-2】 设 $\boldsymbol{R}^+$ 是一切正实数的集合，$\forall x\in\boldsymbol{R}^+$，令 $f(x)=\pm\sqrt{x}$ 与之对应。由于 $x>0$ 时，$f(x)$ 不能由 x 唯一确定，所以 f 不是 $\boldsymbol{R}^+$ 到 $\boldsymbol{R}$ 的映射。

【例 1-3】 设 $P_n[x]$ 是所有次数不大于 n 的实系数多项式的集合，$\forall f(x)\in P_n[x]$，令 $\sigma(f(x))=f'(x)$，则 σ 是 $P_n[x]$ 到自身的一个映射（其中 $f'(x)$ 表示

$f(x)$对 x 求一阶导数)。

某个集合 A 到自身的映射也称为 A 的一个**变换**。

【例 1-4】 设 A 是任意一个非空集合，$\forall x\in A$，令 $f(x)=x$ 与之对应：$f: x\mapsto x$。显然这是 A 到 A 的一个映射，这个映射称为 A 的**恒等映射**(或**恒等变换**)，记作 j_A。

关于两个集合间的映射有以下几点需要注意：

(1) A、B 可以是相同的集合，也可以是不同的集合；

(2) 对于 A 中的每一个元素 x，B 中必有一个唯一确定的元素与之对应；

(3) 一般说来，B 中的元素不一定都是 A 中元素的象；

(4) A 中不同元素的象可能相同。

定义 1.1.2 设 f，g 都是 A 到 B 的映射，如果 $\forall x\in A$，都有 $f(x)=g(x)$，则称映射 f 与 g 是相等的，记作 $f=g$。

定义 1.1.3 设 f 是 A 到 B 的一个映射，如果 $f(A)=B$，则称 f 是 A 到 B 上的一个映射，也称 f 是一个满射。

定义 1.1.4 设 f 是 A 到 B 的一个映射，如果对于 A 中的任意两个元素 x_1 和 x_2，当 $x_1\neq x_2$ 时，有 $f(x_1)\neq f(x_2)$，则称 f 是 A 到 B 的一个单射。

定义 1.1.5 设 f 是 A 到 B 的一个映射，g 是 B 到 C 的一个映射，若对于每一个 $x\in A$，有 C 中的一个唯一确定的元素 $g(f(x))$ 与之对应，则得到一个由 A 到 C 的映射 gf，称其为 f 与 g 的合成(或称为映射的积)。

事实：设给定映射 $f: A\rightarrow B$，$g: B\rightarrow C$，$h: C\rightarrow D$，则合成映射 $h(gf)$ 与 $(hg)f$ 都是 A 到 D 的映射，并且有 $h(gf)=(hg)f$。

证：由定义 1.1.5 易见 $h(gf)$ 与 $(hg)f$ 都是 A 到 D 的映射。

设 $u=gf$，$v=hg$，则 $\forall x\in A$ 有

$$hu(x)=h(u(x))=h(g(f(x))),\quad vf(x)=v(f(x))=h(g(f(x)))$$

所以 $hu=vf$，$h(gf)=(hg)f$。

一般地，映射的合成不满足交换律，即 $fg\neq gf$。

定义 1.1.6 如果映射 $f: A\rightarrow B$ 既是满射又是单射，即 f 满足以下条件：

(1) $f(A)=B$；

(2) $\forall x_1, x_2\in A$，若 $f(x_1)=f(x_2)$，则 $x_1=x_2$。

则称 f 是 A 到 B 的一个 1-1 映射。

特别地，一个有限集合 A 到自身的 1-1 映射称为 A 的一个**置换**。显然，如果 A 是有限集合，则恒等映射 j_A 是一个置换，这个置换也称为**恒等置换**。

定理 1.1.1 设 $f: A\rightarrow B$ 是一个映射，则以下两个条件等价：

(1) f 是 1-1 映射；

(2) 存在映射 $g: B\rightarrow A$，使得 $gf=j_A$，$fg=j_B$。

并且当条件(2)成立时，映射 g 由映射 f 唯一确定。

证：若条件(1)成立，因为 $f(A)=B$，所以 $\forall y\in B$，有 $x\in A$，使得 $f(x)=y$；又因为 f 是单射，所以这个 x 是由 y 唯一确定的。规定 $g: y\mapsto x$。

如果 $f(x)=y$，则 g 是 B 到 A 的一个映射。

设 $x\in A$，而 $f(x)=y$，则有 $gf(x)=g(f(x))=g(y)=x$，所以 $gf=j_A$。

设 $y\in B$，而 $f(x)=y$，那么 $g(y)=x$，所以 $fg(y)=f(g(y))=f(x)=y$，所以 $fg=j_B$。

反之，若条件（2）成立，设 $y\in B$，令 $g(y)=x\in A$。因为 $fg=j_B$，所以 $f(x)=f(g(y))=j_B(y)=y$，所以 f 是满射。

再设 x_1，$x_2\in A$，而 $f(x_1)=f(x_2)$，所以 $x_1=j_A(x_1)=gf(x_1)=gf(x_2)=j_A(x_2)=x_2$，所以 f 是单射。

所以 f 是 1-1 映射。

此外，当条件（2）成立时，令 $g:B\to A$，$h:B\to A$ 均具有性质：$gf=hf=j_A$，$fg=fh=j_B$，则有 $g=gj_B=g(fh)=(gf)h=j_Ah=h$。

所以 g 由 f 唯一确定。

定义 1.1.7 满足定理 1.1.1 条件（2）的映射 $g:B\to A$ 叫做 f 的**逆映射**，记作 f^{-1}。

显然有 $f^{-1}f=j_A$，$ff^{-1}=j_B$，$(f^{-1})^{-1}=f$。

需要注意的是：一个映射不一定有逆映射。

§1.2 线性空间及其性质

在解析几何中，我们已经学习过平面与空间的向量，线性空间是解析几何中向量概念的一般化。

定义 1.2.1 设 P 是一个数域，P 中的元素用小写拉丁字母 a，b，c，…表示，令 V 是一个非空集合，V 中的元素用小写希腊字母 α，β，γ，…表示。称 V 中的元素为向量，P 中的元素为纯量。如果下列条件被满足，则称 V 是 P 上的一个**线性空间**（或**向量空间**）：

（1）在 V 中定义了一个“加法”。对于 V 中任意两个向量 α 与 β，有 V 中一个唯一确定的向量与它们对应，这个向量叫做 α 与 β 的和，记作 $\alpha+\beta$。

（2）有一个“纯量乘法”。对于 P 中每一个纯量 a 和 V 中每一个向量 α，有 V 中一个唯一确定的向量与它们对应，这个向量叫做 a 与 α 的积，记作 $a\alpha$。

（3）向量的加法和纯量的乘法满足下列运算律：

1）$\alpha+\beta=\beta+\alpha$；

2）$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$；

3）在 V 中存在一个零向量，记作 0，对于 V 中每一个向量 α，都有 $0+\alpha=\alpha$；

4）对于 V 中每一个向量 α，在 V 中存在一个向量 α'，使得 $\alpha+\alpha'=0$，这样的 α' 叫做 α 的负向量，记作 $\alpha'=-\alpha$；

5）$a(\alpha+\beta)=a\alpha+a\beta$；

6）$(a+b)\alpha=a\alpha+b\alpha$；

7）$(ab)\alpha=a(b\alpha)$；

8）$1\cdot\alpha=\alpha$。

V 中所定义的加法与纯量乘法称为 V 中的线性运算，在不致发生混淆的情况下，将数域 P 上的线性空间 V 简称为线性空间 V。当 P 为实数域 $\boldsymbol{R}$ 或复数域 $\boldsymbol{C}$ 时，分别称 V 为实线性空间或复线性空间。

【例 1-5】 在解析几何里，平面或空间中所有向量，对于向量的加法和实数与向量的乘法这两种线性运算构成实数域 $\boldsymbol{R}$ 上的线性空间。分别记作 $\boldsymbol{R}^2$，$\boldsymbol{R}^3$。

【例 1-6】 分量属于数域 P 的全体 n 元数组 $(x_1, x_2, \cdots, x_n)^T$ 构成 P 上的一个线性空间，记作 P^n。当 $P=\boldsymbol{C}$ 时，P^n 称为 n 元复线性空间，记作 $\boldsymbol{C}^n$；当 $P=\boldsymbol{R}$ 时，P^n 称为 n 元实线性空间，记作 $\boldsymbol{R}^n$。

【例 1-7】 数域 P 上一切 $m\times n$ 矩阵所成的集合对于矩阵的加法和数与矩阵的乘法构成 P 上的一个线性空间。

当 $P=\boldsymbol{C}$ 时，则此空间记作 $\boldsymbol{C}^{m\times n}$；当 $P=\boldsymbol{R}$ 时，则此空间记作 $\boldsymbol{R}^{m\times n}$。

特别地，P 上一切 1 行 n 列矩阵所成的集合和一切 n 行 1 列矩阵所成的集合分别构成 P 上的线性空间。前者称为 P 上的行空间，后者称为 P 上的列空间（例 1-6 中的 P^n 就是一个列空间）。

【例 1-8】 给定 $A\in\boldsymbol{C}^{m\times n}$，记 $R(A)=\{y\mid y=Ax,\ x\in\boldsymbol{C}^n\}$，$N(A)=\{x\mid Ax=0,\ x\in\boldsymbol{C}^n\}$，按 $\boldsymbol{C}^n$ 中的加法和数乘运算，$R(A)$，$N(A)$ 都是 $\boldsymbol{C}$ 上的线性空间。

证：(1) 设 $y_1, y_2\in R(A)$，则存在 $x_1, x_2\in\boldsymbol{C}^n$，使得 $y_1=Ax_1$，$y_2=Ax_2$，因为 $\boldsymbol{C}^n$ 是线性空间，所以 $x_1+x_2\in\boldsymbol{C}^n$，所以 $A(x_1+x_2)\in R(A)$。又因为 $A(x_1+x_2)=Ax_1+Ax_2=y_1+y_2$，所以 $y_1+y_2\in R(A)$。

(2) 同理，设 $k\in\boldsymbol{C}$，所以 $kx_1\in\boldsymbol{C}^n$，则 $A(kx_1)\in R(A)$。又因为 $A(kx_1)=kAx_1=ky_1$，所以 $ky_1\in R(A)$。

(3) 1) $y_1+y_2=Ax_1+Ax_2=A(x_1+x_2)=A(x_2+x_1)=Ax_2+Ax_1=y_2+y_1$；

2) $(y_1+y_2)+y_3=(Ax_1+Ax_2)+Ax_3=Ax_1+(Ax_2+Ax_3)=y_1+(y_2+y_3)$；

3) 因为 $0\in\boldsymbol{C}^n$，所以 $0=A0\in R(A)$；

4) 设 $x_1\in\boldsymbol{C}^n$，所以 $-x_1\in\boldsymbol{C}^n$，所以 $0=Ax_1+A(-x_1)=A(x_1-x_1)\in R(A)$；

5) 设 $a\in\boldsymbol{C}$，$a(y_1+y_2)=a(Ax_1+Ax_2)=aAx_1+aAx_2=ay_1+ay_2$；

6) 设 $a, b\in\boldsymbol{C}$，$(a+b)y_1=(a+b)Ax_1=aAx_1+bAx_1=ay_1+by_1$；

7) $(ab)y_1=(ab)Ax_1=a(bAx_1)=a(by_1)$；

8) $1\cdot y_1=1\cdot Ax_1=Ax_1=y_1$。

所以 $R(A)$ 为 $\boldsymbol{C}$ 上的线性空间。

同理可证明 $N(A)$ 也是 $\boldsymbol{C}$ 上的线性空间。

【例 1-9】 复数域 $\boldsymbol{C}$ 可以视为实数域 $\boldsymbol{R}$ 上的线性空间。

事实上，二复数之和还是复数，一个实数与一个复数之积还是复数，同时定义 1.2.1 中的 (3) 也成立。

【例 1-10】 $\boldsymbol{R}$ 是实数域，V 为全体正实数组成的集合，定义 V 中两个元素的加法 $\oplus$ 为

$$a\oplus b=ab\ (a,\ b\in V)$$

定义 $\boldsymbol{R}$ 中元素与 V 中元素的数乘运算$\bigcirc$为

$$k\bigcirc a = a^k \quad (a\in V,\ k\in \boldsymbol{R})$$

则 V 是 $\boldsymbol{R}$ 上的一个线性空间。

证：对任意 a，b，$c\in V$，k，$l\in \boldsymbol{R}$ 有：

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

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

(3) 1 是零元素，因为 $1\oplus b = 1\cdot b = b$;

(4) a 的负元素是 a^{-1}，因为 $a\oplus a^{-1} = aa^{-1} = 1$;

(5) $k \bigcirc(a\oplus b) = (ab)^k = a^k b^k = (k\bigcirc a)\oplus(k\bigcirc b)$;

(6) $(k+l)\bigcirc a = a^{k+l} = a^k a^l = (k\bigcirc a)\oplus(k\bigcirc b)$;

(7) $(kl)a = a^{kl} = (a^l)^k = k\bigcirc(l\bigcirc a)$;

(8) $l\bigcirc a = a^1 = a$。

【例 1-11】 任意数域 P 总可以视为它自身上的线性空间。

【例 1-12】 仅由 $\boldsymbol{C}$ 上线性空间 V 中的零元素 0 构成的单元素集合 $0 = \{0 | 0\in V\}$，按 V 中的运算定义运算，则 0 是 $\boldsymbol{C}$ 上的一个线性空间，这个空间叫做零空间。

【例 1-13】 当 $b\neq 0$ 时，相容的线性方程组 $Ax = b$ 的解的全体 $S = \{x | Ax = b,\ x\in \boldsymbol{C}^n\}$ 是否构成线性空间？

证：(反证法)

若 S 是线性空间，则 $\forall x_1$，$x_2\in S$，有 $x_1 + x_2\in S$。但是 $A(x_1 + x_2) = Ax_1 + Ax_2 = 2b\neq b$，所以 $x_1 + x_2\notin S$。

所以 S 不是线性空间。

【例 1-14】 设 $\{a_n\}$，$\{b_n\}$ 是两个收敛于 0 的实数无穷序列，则 $\lim\limits_{n\to\infty}(a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n = 0$；且 $\forall a\in R$，有 $\lim\limits_{n\to\infty} aa_n = a\lim\limits_{n\to\infty} a_n = 0$。

易证定义 1.2.1 中的 (3) 也成立。

所以，一切收敛于 0 的实序列对于如上定义的加法和数与序列的乘法构成 $\boldsymbol{R}$ 上的一个线性空间。

【例 1-15】 n 阶线性齐次常微分方程

$$L[x] = \frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_1\frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}} + a_2\frac{\mathrm{d}^{n-2}x}{\mathrm{d}t^{n-2}} + \cdots + a_{n-1}\frac{\mathrm{d}x}{\mathrm{d}t} + a_n x = 0$$

的解的全体 $S = \{x(t) | L[x] = 0\}$，对于普通函数的加法、数乘法构成 $\boldsymbol{C}$ 上的线性空间。

对于零向量与负向量，我们有以下的定理：

定理 1.2.1 在线性空间 V 中，零向量是唯一的；对于 V 中的任意向量 α，其负向量也是唯一的。

证：设 0 和 $0'$ 都是 V 的零向量，则 $\forall \alpha\in V$，有 $0+\alpha = \alpha$，$\alpha + 0' = \alpha$，

所以 $0 = 0 + 0' = 0'$。

设 α'，α''都是 α 的负向量，则有 $\alpha' + \alpha = 0$，$\alpha'' + \alpha = 0$，所以 $\alpha' = \alpha' + 0 = \alpha' + (\alpha +$

$\alpha'') = (\alpha' + \alpha) + \alpha'' = 0 + \alpha'' = \alpha''$。

定义 1.2.2 在线性空间 V 中，两个向量 α 与 β 的差为 $\alpha + (-\beta)$，记作 $\alpha - \beta$。

定理 1.2.2 对任意向量 α，β，$\gamma \in V$，对任意的数 $a \in P$，有：

(1) $0\alpha = 0$，$a0 = 0$；

(2) $a(-\alpha) = (-a)\alpha = -a\alpha$；

(3) $a\alpha = 0 \Rightarrow a = 0$ 或 $\alpha = 0$；

(4) $\alpha + \beta = \gamma \Leftrightarrow \alpha = \gamma - \beta$。

证：(1)

$$\begin{aligned} 0\alpha &= 0\alpha + 0 = 0\alpha + (0\alpha - 0\alpha) = (0\alpha + 0\alpha) - 0\alpha \\ &= (0 + 0)\alpha - 0\alpha = 0\alpha - 0\alpha = 0 \end{aligned}$$

同理可证 $a0 = 0$。

(2) 因为 $a\alpha + a(-\alpha) = a(\alpha + (-\alpha)) = a0 = 0$，所以 $a(-\alpha) = -a\alpha$。

同理可证 $(-a)\alpha = -a\alpha$。

(3) 设 $a\alpha = 0$，但 $a \neq 0$，则 $\alpha = 1\alpha = \left(\frac{1}{a}a\right)\alpha = \frac{1}{a}(a\alpha) = \frac{1}{a}0 = 0$。

(4) $\alpha + \beta = \gamma \Leftrightarrow \alpha + \beta + (-\beta) = \gamma + (-\beta) \Leftrightarrow \alpha + (\beta - \beta)$

$\Leftrightarrow \alpha + 0 = \gamma - \beta \Leftrightarrow \alpha = \gamma - \beta$

§1.3 基、维数与坐标

在线性代数中，学习过向量的线性相关性，现在我们应用它来研究线性空间。

定义 1.3.1 设 V 是数域 P 上的一个线性空间，V 中满足以下两个条件的向量组 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 叫做 V 的一个基：

(1) α_1，α_2，…，α_n 线性无关；

(2) V 中的每一个向量都可以由 α_1，α_2，…，α_n 线性表示。

其中，n 称为线性空间 V 的维数，记作 $\dim V = n$。维数为 n 的线性空间称为 n 维线性空间，记作 V^n。

【例 1-16】 在线性空间 $\boldsymbol{R}^2$ 中，任意两个不共线的向量都构成 $\boldsymbol{R}^2$ 的一个基；在线性空间 $\boldsymbol{R}^3$ 中，任意三个不共面的向量都构成 $\boldsymbol{R}^3$ 的一个基。并且 $\dim\boldsymbol{R}^2 = 2$；$\dim\boldsymbol{R}^3 = 3$。

【例 1-17】 设 $\boldsymbol{C}^{m\times n}$ 是数域 $\boldsymbol{C}$ 上一切 $m \times n$ 矩阵构成的线性空间，考虑如下的 mn 个矩阵

$$E_{ij} = \begin{array}{c} \begin{bmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{bmatrix} (i) \\ (j) \end{array}$$

在 E_{ij} 中，除去第 i 行第 j 列位置上的元素是 1 外，其余的元素都是 0，$(i = 1, 2, \cdots,$

m；$j=1, 2, \cdots, n$)。由定义 1.3.1 可见 $\{E_{ij}\}$ 构成所有 $m\times n$ 矩阵组成的线性空间 $\boldsymbol{C}^{m\times n}$的一个基，并且 $\dim\boldsymbol{C}^{m\times n}=m\times n$。

【例 1-18】 零空间的维数是零。

定理 1.3.1 设 $\{\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_n\}$ 是数域 P 上的线性空间 V^n 的一个基，则 V^n 中的每一个向量可以唯一的被表示为基向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合。

证：设 $\forall\alpha\in V^n$，则

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$$

$$\alpha=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n$$

两式相减有 $(x_1-y_1)\alpha_1+(x_2-y_2)\alpha_2+\cdots+(x_n-y_n)\alpha_n=0$。

因为 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关，所以 $x_1=y_1, x_2=y_2, \cdots, x_n=y_n$。

反之，任给一个有序数组 $x_1, x_2, \cdots, x_n$，总有唯一的元素 $\alpha\in V^n$ 可以由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性表示为 $\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in V^n$（其中 $x_1, x_2, \cdots, x_n\in P$)。

由此可知，如果 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V^n 的一个基，则 V^n 中元素的全体可以表示为

$$V^n=\{\alpha\mid\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n, x_1, x_2, \cdots, x_n\in P\}$$

也就是说，V^n 中的向量 α 与有序数组 $x_1, x_2, \cdots, x_n$ 之间构成一一对应关系。

定义 1.3.2 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性空间 V^n 的一个基，$\forall\alpha\in V^n$，有且仅有一个有序数组 $x_1, x_2, \cdots, x_n$，使得 $\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$，我们称 $x_1, x_2, \cdots, x_n$ 为元素 α 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标，记作 $(x_1, x_2, \cdots, x_n)^T$。

【例 1-19】 取定 $\boldsymbol{R}^3$ 中三个不共面的向量 $\alpha_1, \alpha_2, \alpha_3$，$\forall\xi\in\boldsymbol{R}^3$，$\xi$ 可以唯一的表示为 $\xi=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3$。因此，向量 ξ 关于基 $\{\alpha_1, \alpha_2, \alpha_3\}$ 的坐标为 $(x_1, x_2, x_3)^T$。

【例 1-20】 在 $\boldsymbol{R}^n$ 中如下的 n 个向量：$\varepsilon_i=(0, \cdots, 0, \underset{(i)}{1}, 0, \cdots, 0)^T$，$i=1, 2, \cdots, n$ 作为 $\boldsymbol{R}^n$ 的一个基，此基称为 $\boldsymbol{R}^n$ 的标准基。

$\forall\alpha=(a_1, a_2, \cdots, a_n)^T\in\boldsymbol{R}^n$，有 $\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$，所以 α 关于基 $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$ 的坐标为 $(a_1, a_2, \cdots, a_n)^T$。

【例 1-21】 在 $\boldsymbol{R}^n$ 中如下的 n 个向量

$$\begin{aligned}\varepsilon'_1&=(1, 1, 1, \cdots, 1)^T\\ \varepsilon'_2&=(0, 1, 1, \cdots, 1)^T\\ \varepsilon'_3&=(0, 0, 1, \cdots, 1)^T\\ &\vdots\\ \varepsilon'_n&=(0, 0, \cdots, 0, 1)^T\end{aligned}$$

也是 $\boldsymbol{R}^n$ 的一个基，因为 $\forall\alpha=(a_1, a_2, \cdots, a_n)^T\in\boldsymbol{R}^n$，有

$$\alpha=a_1\varepsilon_1'+(a_2-a_1)\varepsilon'_2+\cdots+(a_n-a_{n-1})\varepsilon'_n$$

所以 α 关于基 $\{\varepsilon'_1, \varepsilon'_2, \cdots, \varepsilon'_n\}$ 的坐标为 $(a_1, a_2-a_1, \cdots, a_n-a_{n-1})^T$。

【例 1-22】 求线性空间 $P_n[x]$ 的基、维数以及向量 p 的坐标。

在 $P_n[x]$ 中，它的一个基为 $p_1=1$，$p_2=x$，$p_3=x^2$，$\cdots$，$p_n=x^{n-1}$，$p_{n+1}=x^n$，所

以 $\dim P_n[x]=n+1$。

任何次数不大于 n 的多项式 $p=a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}+a_nx^n$ 可以表示为

$$p=a_0p_1+a_1p_2+a_2p_3+\cdots+a_{n-1}p_n+a_np_{n+1}$$

所以 p 在基 $\{p_1, p_2, \cdots, p_{n+1}\}$ 下的坐标为 $p=(a_0, a_1, a_2, \cdots, a_n)^T$。

如果在 $P_n[x]$ 中另取一个基 $\{p'_1, p_2', \cdots, p'_{n+1}\}$，其中 $p_1'=1$，$p_2'=x-a$，$\cdots$，$p'_{n+1}=(x-a)^n$，则由 p 在 $x=a$ 点的 Taylor 多项式 $p=p(a)+p'(a)(x-a)+\cdots+\dfrac{p^{(n)}(a)}{n!}(x-a)^n$ 可知，p 在基 $\{p_1', p_2', \cdots, p'_{n+1}\}$ 下的坐标为 $p=\left(p(a), p'(a), \cdots, \dfrac{p^{(n)}(a)}{n!}\right)^T$。

【例 1-23】 线性空间 $C^{m\times n}$的一个基为

$$E_{ij}=\begin{array}{c}\begin{bmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \end{bmatrix}(i) \\ (j)\end{array}\qquad (i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$$

$\forall M\in C^{m\times n}$，有

$$M=a_{11}E_{11}+a_{12}E_{12}+\cdots+a_{1n}E_{1n}+a_{21}E_{21}+\cdots+a_{mn}E_{mn}=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij}$$

所以 $\dim C^{m\times n}=mn$；$M=(a_{ij})$ 在基 $\{E_{ij}\}_{j=1,2,\cdots,n}^{i=1,2,\cdots,m}$ 下的坐标为 $(a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, \cdots, a_{mn})^T$。

设 n 维线性空间 V^n 的向量 α，β 关于 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标分别为 $(x_1, x_2, \cdots, x_n)^T$ 和 $(y_1, y_2, \cdots, y_n)^T$，$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$，$\beta=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n$，那么，$\alpha+\beta=(x_1+y_1)\alpha_1+(x_2+y_2)\alpha_2+\cdots+(x_n+y_n)\alpha_n$。

设 $a\in P$，那么 $a\alpha=(ax_1)\alpha_1+(ax_2)\alpha_2+\cdots+(ax_n)\alpha_n$，所以有：

定理 1.3.2 设 V 是数域 P 上的一个 n 维线性空间，$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 是 V 的一个基，α，$\beta\in V$，它们关于基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标分别是 $(x_1, x_2, \cdots, x_n)^T$ 和 $(y_1, y_2, \cdots, y_n)^T$，那么，$\alpha+\beta$ 关于这个基的坐标就是 $(x_1+y_1, x_2+y_2, \cdots, x_n+y_n)^T$。

又设 $a\in P$，那么，$a\alpha$ 关于这个基的坐标就是 $(ax_1, ax_2, \cdots, ax_n)^T$。

一个向量的坐标依赖于基的选取，对于线性空间 V 的两个基来说，同一个向量的坐标一般是不相同的。

以下我们来讨论，一个向量关于不同的基的坐标之间的关系。

设 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 和 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 是 V^n 的两个基，由此可知 β_j 作为 V^n 中

的元素，可以表示为

$$\beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n$$
$$\beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n$$
$$\vdots$$
$$\beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n \tag{1-1}$$

这里 $(a_{1j}, a_{2j}, \cdots, a_{nj})^T$ 是 β_j 关于基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标 $(j = 1, 2, \cdots, n)$。以这 n 个坐标为列，作一个 n 阶矩阵 T，即

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

矩阵 T 叫做由基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的过渡矩阵。

式（1-1）可以记为

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\ T \tag{1-2}$$

设 $\alpha \in V$，α 关于基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标为 $(x_1, x_2, \cdots, x_n)^T$，关于基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的坐标为 $(y_1, y_2, \cdots, y_n)^T$，于是有

$$\alpha = \sum_{i=1}^{n} x_i\alpha_i = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1-3}$$

以及

$$\alpha = \sum_{i=1}^{n} y_i\beta_i = (\beta_1, \beta_2, \cdots, \beta_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{1-4}$$

将式（1-2）代入式（1-4）有

$$\alpha = ((\alpha_1, \alpha_2, \cdots, \alpha_n)\ T)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)\left[T\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right] \tag{1-5}$$

由式（1-5）可见，α 关于基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标为 $T\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$。

但是我们知道向量 α 关于一个基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标是唯一的，所以

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = T \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{1-6}$$

于是有：

定理 1.3.3 设 V 是数域 P 上的一个 n 维线性空间，T 是由基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的过渡矩阵，那么 V 中向量 α 关于基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的坐标 $(x_1, x_2, \cdots, x_n)^T$ 与关于基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的坐标 $(y_1, y_2, \cdots, y_n)^T$ 由式（1-6）联系着。

如果 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，$\{\beta_1, \beta_2, \cdots, \beta_n\}$，$\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$ 都是线性空间 V^n 的基；并且设由基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的过渡矩阵为 $A = (a_{ij})$；设由基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 到基 $\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$ 的过渡矩阵为 $B = (b_{ij})$。则由以上的讨论可知

$$\beta_k = \sum_{i=1}^{n} a_{ik}\alpha_i \qquad (k = 1, 2, \cdots, n) \tag{1-7}$$

$$\gamma_j = \sum_{k=1}^{n} b_{kj}\beta_k \qquad (j = 1, 2, \cdots, n) \tag{1-8}$$

将式（1-7）代入式（1-8）有

$$\gamma_j = \sum_{k=1}^{n} b_{kj}\sum_{i=1}^{n} a_{ik}\alpha_i = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)\alpha_i = \sum_{i=1}^{n} c_{ij}\alpha_i \tag{1-9}$$

其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \tag{1-10}$$

令 $C = (c_{ij})$，根据以上的讨论可知：由基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 到基 $\{\gamma_1, \gamma_2, \cdots, \gamma_n\}$ 的过渡矩阵为 $C = (c_{ij})$；同时由式（1-10）亦可看出 $C = AB$。这是因为由式（1-7）、式（1-8）有

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

$$(\gamma_1, \gamma_2, \cdots, \gamma_n) = (\beta_1, \beta_2, \cdots, \beta_n) B$$

所以

$$(\gamma_1, \gamma_2, \cdots, \gamma_n) = ((\alpha_1, \alpha_2, \cdots, \alpha_n) A) B = (\alpha_1, \alpha_2, \cdots, \alpha_n) AB$$

由此可见，我们所熟悉的矩阵乘法的定义正反映了这种线性代入过程。

过渡矩阵反映了线性空间的不同的基之间的关系，这是一个很重要的概念，现在进一步讨论过渡矩阵的一些性质。

如果由基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 的过渡矩阵为 A，则有

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n) A \tag{1-11}$$

再设基 $\{\beta_1, \beta_2, \cdots, \beta_n\}$ 到基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ 的过渡矩阵为 B，则有

$$(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\beta_1, \beta_2, \cdots, \beta_n) B \tag{1-12}$$