# 矩阵理论及其应用

$A \in C^{m \times n}$

$A: C^n \longrightarrow C^m$

$x \longmapsto Ax$

$C^n = N(A) \oplus R(A^H)$

北京[illegible]學院出版社

# 矩阵理论及其应用


蒋正新
施国梁
编著




北京航空學院出版社

## 内 容 简 介

本书是专为工科研究生和应用数学专业本科生编写的《矩阵理论及其应用》教材。在内容上除包括了常用的矩阵理论及方法外，还列举了它们在数值分析、最优化、控制论、随机过程及微分方程等方面的应用。以便为学习《泛函分析》奠定一定的基础。全书共分六章：线性代数基础、矩阵的因式分解、广义逆矩阵、矩阵分析、特征值的估计、非负矩阵及其应用。每章都附有适量的例题与习题。

本书也适合于工科学校的教师、工程师以及有关的科技工作者自学之用。


**矩阵理论及其应用**

蒋正新　施国梁　編著

責任編輯　郭維烈

北京航空学院出版社出版

新华书店北京发行所发行　各地新华书店经售

北京航空学院印刷厂印刷

850×1168　1/32　印张：13　字数：338千字

1988年3月第一版　1988年3月第一次印刷　印数：7000册

ISBN 7-81012-030-1/O·003　定价：2.50元

# 前　言

本书是根据我过去七年在北航研究生院为工科研究生讲授《矩阵论》的讲稿修改扩充而成的。由于研究生听课人数逐年增多（每届数百人），急需教材使用，便约請施国梁同志共同加工。由于讲稿以工科研究生为对象，所以不太追求数学上的系统与严谨，而注重矩阵的理论与方法在工程实际中的应用。因此，作为通用教材，一定存在着多方面的问题，甚至是错误的。敬請批评与指正。

本书初稿曾蒙北京工业学院王朝瑞、史荣昌副教授的评审，提出了很好的修改意见，并指出了个别错漏之处。在此，谨向二位表示衷心的感谢。

蒋正新

一九八七年五月

# 符号说明

| | |
|---|---|
| $A^+$ | $A$的Moore-Penrose广义逆 |
| $A^T$ | $A$的转置 |
| $A^H$ | $A$的共轭转置 |
| $A_L^{-1}$ | $A$的左逆 |
| $A_R^{-1}$ | $A$的右逆 |
| $(A)_{ij}$ | 矩阵$A$的第（$i$，$j$）元素 |
| $\sigma(A)$ | $A$的谱，$A$的所有特征值集合 |
| $\rho(A)$ | $A$的谱半径，$\max\{\|\lambda\| \mid \lambda\in\sigma(A)\}$ |
| $\gamma(A)$ | $A$的素性指数 |
| $\delta(A)$ | $\max\{\|\lambda\| \mid \lambda\in\sigma(A),\ \lambda\neq\rho(A)\}$ |
| $\mathscr{N}(A)$ | $A$的零空间，$A$的核 |
| $\mathscr{R}(A)$ | $A$的值域，$A$的列空间 |
| $\\|A\\|$ | $A$的范数， |
| $A\geqslant 0$ | $A$的每个元素都是非负的 |
| $A>0$ | $A$的每个元素都是正的 |
| $A\geqslant B$ | $A$的每个元素都不小于$B$的相应元素 |
| $\det A$ | $A$的行列式 |
| $\mathrm{adj}A$ | $A$的伴随矩阵 |
| $\mathrm{tr}A$ | $A$的迹，$A$的主对角元之和 |
| $\mathbf{R}$ | 实数的集合，实直线 |
| $\mathbf{R}^n$ | $n$维实坐标向量空间 |
| $\mathbf{R}^{m\times n}$ | $m\times n$阶实矩阵空间 |
| $\mathbf{R}_r^{m\times n}$ | 所有秩为$r$的$m\times n$阶实矩阵集合 |
| $\mathbf{C}$ | 复数集，复平面 |
| $\mathbf{C}^{m\times n}$ | $m\times n$阶复矩阵空间 |
| $\mathrm{Re}\lambda$ | 复数$\lambda$的实部 |

| | |
|---|---|
| $\operatorname{Im}\lambda$ | 复数$\lambda$的虚部 |
| $E^n$ | $n$维欧氏空间 |
| $X^c$ | 子空间$X$的代数补 |
| $X^\perp$ | 集合$X$的正交补 |
| $(x\|y)$ | 量向$x$与$y$的内积 |
| $x\perp y$ | 向量$x$与$y$垂直 |
| $X\cup Y$ | $X$与$Y$的并集 |
| $X\cap Y$ | $X$与$Y$的交集 |
| $X\oplus Y$ | 子空间$X$与$Y$的直和 |
| $A\oplus B$ | 矩阵$A$与$B$的直和，$\operatorname{diag}(A, B)$ |
| $L(X, Y)$ | $X$到$Y$的所有线性变换集合 |
| $\dim(X)$ | 空间$X$的维数 |
| $\operatorname{rank}(A)$ | 矩阵$A$的秩 |
| $\operatorname{null}(A)$ | 矩阵$A$或线性变换$A$的零度 |
| $\operatorname{span}M$ | 向量集合$M$张成的线性子空间 |
| $P_{L,M}$ | 沿着空间$M$向空间$L$的垂直投影 |
| $\bigoplus_{i=1}^{k} A_i$ | $\operatorname{diag}(A_1, A_2, \cdots, A_k)$ |
| $\min_{x\in S} f(x)$ | $f(x)$在$S$上的最小值 |
| $\max_{x\in S} f(x)$ | $f(x)$在$S$上的最大值 |
| $\inf_{x\in S} f(x)$ | $f(x)$在$S$上的下确界 |
| $\sup_{x\in S} f(x)$ | $f(x)$在$S$上的上确界 |
| $\operatorname{diag}(a_1, a_2, \cdots, a_n)$ | $n\times n$阶对角阵 |
| $\longrightarrow$ | 蕴含着，或必要条件 |
| $\Longleftrightarrow$ | 互为充分必要条件 |
| $x\longmapsto y$ | 给定$x$，有唯一的$y$与之对应 |

# 目　录

**符号说明**

**第一章　线性代数基础**

第一节　线性空间……………………………………（3）
第二节　欧氏空间和酉空间…………………………（18）
第三节　线性变换……………………………………（35）
第四节　例题与习题…………………………………（73）

**第二章　矩阵的几种重要分解**

第一节　满秩方阵的$UR$分解及对称正定阵
　　　　的$R^TR$分解 ……………………………………（87）
第二节　舒尔引理与正规矩阵………………………（93）
第三节　幂等矩阵、投影算子以及矩阵
　　　　的谱分解式…………………………………（103）
第四节　矩阵的Hermite标准形及秩分解…………（112）
第五节　矩阵的奇异值分解和极分解……………（117）
*第六节　Jordan 标准形的存在性 …………………（125）
第七节　例题与习题…………………………………（133）

**第三章　矩阵的广义逆**

第一节　Moore-Penrose 广义逆……………………（148）
第二节　M-P 逆的几种显式表示……………………（154）
第三节　矩阵方程$AXB=D$的求解与广义逆
　　　　的性质………………………………………（155）
第四节　广义逆与矛盾方程$Ax=b$的求解…………（161）

第五节 几种计算$A^+$的方法…………………………（165）
第六节 广义逆的应用举例…………………………（179）
第七节 例题与习题…………………………………（192）
**第四章 矩阵分析**
第一节 矩阵范数……………………………………（205）
第二节 矩阵序列与矩阵级数………………………（219）
第三节 幂级数及 Lagrange-Sylvester 定理 ………（228）
第四节 函数矩阵的微分与积分，矩阵的标量
函数的微分…………………………………（243）
第五节 矩阵分析在微分方程中的应用……………（254）
第六节 稳定性与 Lyapunov 定理…………………（270）
第七节 矩阵函数$e^{tA}$与Laplace 变换……………（277）
第八节 例题与习题…………………………………（289）
**第五章 特征值的估计**
第一节 特征值的界的估计…………………………（306）
第二节 特征值所在区域的确定，盖尔斯果林
的圆盘定理…………………………………（311）
第三节 圆盘定理的应用及推广……………………（315）
第四节 特殊类型矩阵的特征值估计………………（328）
第五节 例题与习题…………………………………（344）
**第六章 非负矩阵及其应用**
第一节 正矩阵的基本性质…………………………（353）
第二节 素矩阵（*primitive matrix*）…………（361）
第三节 不可约矩阵…………………………………（367）
第四节 随机矩阵与Markov过程 …………………（382）
第五节 习题…………………………………………（398）
**参考文献**……………………………………………（406）

# 第一章　线性代数基础

我们总是习惯于在通常的欧几里德空间 $E^3$ 中考虑和研究问题,你能说出其中的道理吗?不言而喻,我们置身于这一空间之中,对于在这种空间描述的问题当然易于体察。也就是说，$E^3$ 中有与我们直觉思维相合的欧氏几何，它具体、形象，能启发我们去想象。还有，我们从中学到大学所学到的初等数学和高等数学的理论与方法,都是在$E^3$ 中建立起来的。一句话，$E^3$中的几何对象我们感到生动具体。$E^3$ 中 所采用的数学工具我们感到熟悉。所以在遇到问题时,总希望能象在 $E^3$ 中一样得到几何解释,也希望能用得心应手的理论方法加以解决。不过，随着科学技术的发展，我们要研究的对象已远远超出了 $E^3$ 的范围。例如，在系统工程和现代控制理论中，需要研究状态向量的集合；在信息工程中，需要研究各种序列的集合；在微分方程、积分方程和函数逼近理论中，需要研究各类函数的集合；在数值分析和计算中，则需要研究矩阵或变换的集合；等等。这时，我们自然会提出这样的一个问题：对于一般性的抽象集合，能否也建立起类似于 $E^3$ 中的一些结构，把 $E^3$ 中的某些经典理论和方法也平行地转移过来，从而使来自不同领域的各种复杂问题能有一个统一的数学描述，能有一个简单直观的几何解释，能用与 $E^3$ 中的共同原理加以解决。事实表明，这种想法在一定程度上是可行的。这也是线性代数和泛函分析的出发点之一。但是，要达到这样一个目的，首先要认真地考察$E^3$的数学结构。

从数学上不难看出，$E^3$至少有以下几重结构：

**集合论方面的结构。** $E^3$ 中的点可以看作三维实坐标向量的集合

$$V=\{x=(\xi_1, \xi_2, \xi_3) \mid \xi_i \in \mathbf{R}\}$$

我们把它叫做空间 $E^3$ 的基集。对 $V$ 的子集引入集合论中的交、并、补（$\cap$、$\cup$、$c$）三种运算，便成为一个集合代数（Boole 代数）。

**拓扑方面的结构**。对 $V$ 中的两点

$$x=(\xi_1, \xi_2, \xi_3), \quad y=(\eta_1, \eta_2, \eta_3)$$

按照等式

$$d(x, y)=[(\xi_1-\eta_1)^2+(\xi_2-\eta_2)^2+(\xi_3-\eta_3)^2]^{1/2}$$

规定距离，也就是在基集 $V$ 上定义距离函数，便成为一个距离空间（这是一个最常用的拓扑空间）。由距离在 $V$ 中定义开集、闭集、邻域、点列的收敛性、映射的连续性等拓扑概念，从而为建立 $E^3$ 上的微分学奠定基础。

**代数方面的结构**。对 $V$ 中的向量规定加法及与实数的乘法运算：

$$x+y \triangleq (\xi_1+\eta_1, \xi_2+\eta_2, \xi_3+\eta_3)$$

$$\lambda \cdot y \triangleq (\lambda\xi_1, \lambda\xi_2, \lambda\xi_3)$$

很显然，这两种运算的结果都仍是 $V$ 中的向量，且满足一定的法则。这样便成为实数域 $\mathbf{R}$ 上的一个线性空间。据上述两种运算定义向量的相关性、子空间、基底与维数，以及基变换和坐标变换等。

**欧氏几何方面的结构**。在 $V$ 中规定向量 $x$ 和 $y$ 的点积：

$$(x \mid y) \triangleq \xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3$$

再由 $(x \mid y)=0$ 规定 $x$ 与 $y$ 正交（即垂直）。还可进一步规定向量的长度，向量间的夹角，正交变换，研究 $E^3$ 中的几何对象在正交变换群之下的不变量，即所谓的欧氏几何学。

**测度论方面的结构**。把 $V$ 中的向量视为 $E^3$ 中的点，对于 $E^3$ 中直线上、平面上、空间中的点集，分别考虑它们所构成的“长

度”、“面积”、“体积”，从而引入抽象测度的概念。这为$E^3$上的积分学奠定了基础。

上述几种结构在 $E^3$ 中不是孤立的，而是有着密切的联系。用数学的术语来讲，就是所赋予基集$V$的几种结构彼此是相容的。我们也看到空间中的某些几何观念是由拓扑结构提供的（如邻域、开集、闭集等）；有一些几何观念是由代数结构提供的（如基底，坐标，维数，基变换与坐标变换等）；还有一些几何观念，如正交性，正交变换等，是由几种结构的结合所提供的；特别是$E^3$上的微积分，更是几种结构的高度综合。

从以上的分析看，要想把 $E^3$ 的结构推广到一般的集合，则先要考虑把几种结构分别作相应的推广。本章就按这种想法，先赋予集合以代数结构，使之成为一个线性空间。然后，进一步赋予内积（即 $E^3$ 中点积的推广），并由之导出几何结构及拓扑结构，使之成为抽象的欧氏空间或酉空间。最后考虑空间上的线性变换。这样便简明扼要地给出线性代数的基础知识。学过这部分内容的读者，作为复习浏览一下就行了。而只学过线性代数初步的读者，要花相当的时间认真阅读，最好能参考更详尽的教程（如北大编《高等代数》或张远达编《线性代数原理》）。

## 第一节　线性空间

### 一、定义及例子

**定义1.1.1（线性空间）**　一个线性空间是由下面几个部分组成的：

1．一个域$F$，其元素叫做标量；

2．一个集合$V$，其元素叫做向量；

3．一个称作为向量加法的运算，记之为“+”。也就是说，对$V$中的每一对向量$x$，$y$，都有$V$中的一个向量$x+y$与之对应，并叫做$x$与$y$的和，且满足

($a$) 加法交换律，$x+y=y+x$；

($b$) 加法结合律，$x+(y+z)=(x+y)+z$；

($c$) 在$V$中存在唯一的向量0，叫做零向量，使得对一切$V$中的向量$x$有$x+0=x$；

($d$) 对$V$中的每个向量$x$，都存在唯一的$V$中的向量$-x$，使得$x+(-x)=0$；

4．一个称作为向量的标量乘法的运算。也就是说，对域$F$中的每个标量$\lambda$和$V$中的每个向量$x$，都有$V$中的一个向量$\lambda x$与之对应，并叫做$\lambda$与$x$的乘积，且满足：

($e$) 在$F$中存在标量1使得对$V$中的每个向量$x$，都有$1x=x$；

($f$) 对$F$中的一切$\lambda$，$\mu$和$V$中的任意$x$有$(\lambda\mu)x=\lambda(\mu x)$；

($g$) 对$F$中的一切$\lambda$和$V$中的一切$x$，$y$有$\lambda(x+y)=\lambda x+\lambda y$；

($h$) 对$F$中的一切$\lambda$，$\mu$和$V$中的一切$x$有$(\lambda+\mu)x=\lambda x+\mu x$。

线性空间的定义表明：它是域$F$和向量集合$V$的复合；在$V$中定义了向量的加法，在$F$与$V$之间定义了标量与向量的乘法，并且两种运算的结果仍落在$V$中，即$V$对这两种运算保持封闭；同时这两种运算还要满足八条法则。如果我们把乘法用符号"·"表示，则上述内容可以简记为（$V$，$F$，+，·，8），这时称$V$是域$F$上的一个线性空间（或向量空间）。$V$又叫做空间的基集。

由定义可知，若$\{x_1, x_2, \cdots, x_n\}$是线性空间（$V$，$F$，+，·，8）中的任一组向量，$\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$是任一组标量，则

$$y \triangleq \lambda_1 x_1+\lambda_2 x_2+\cdots+\lambda_n x_n=\sum_{i=1}^{n} \lambda_i x_i$$

也是$V$中的向量。这时，我们把$y$叫做向量组$\{x_1, x_2, \cdots, x_n\}$

的一个**线性组合**，而把$\sum_{i=1}^{n}\lambda_i x_i$叫做$y$的一个**线性表出**。

以后我们只讨论域$F$取实数域$\mathbf{R}$或复数域$\mathbf{C}$的情况，但在不强调数域是实的或复的时，仍记为$F$。并把线性空间$(V,F,+,\cdot,8)$简记为$V$。

**例1.** 取基集$V=\{x|x=(\xi_1,\xi_2,\cdots,\xi_n)^T,\ \xi_i\in\mathbf{C}\}$，取标量域$F=\mathbf{C}$。又设$y=(\eta_1,\eta_2,\cdots,\eta_n)^T\in V$，$\lambda\in\mathbf{C}$。规定：

1. $x=y\Longleftrightarrow\xi_i=\eta_i$，$i=1,2,\cdots,n$，
2. $x+y=(\xi_1+\eta_1,\ \xi_2+\eta_2,\ \cdots,\ \xi_n+\eta_n)^T$，
3. $\lambda\cdot x=(\lambda\cdot\xi_1,\ \lambda\cdot\xi_2,\ \cdots,\ \lambda\cdot\xi_n)^T$，
4. $0=(0,0,\cdots,0)^T$。

则很容易看出$V$对上述运算保持封闭，并满足八条运算法则。因此$(V,\mathbf{C},+,\cdot,8)$是一个复数域上的线性空间，叫做**$n$维复坐标向量空间**，记作$\mathbf{C}^n$。类似地，将例1中的$\mathbf{C}$全换成$\mathbf{R}$，则可得$n$维实坐标向量空间，记之为$\mathbf{R}^n$。$\mathbf{C}^n$与$\mathbf{R}^n$是最常用的两个线性空间。

值得注意的是，基集$V$仍取例1中的$V$，而域$F$取实数域$\mathbf{R}$，按例1中定义的运算，则$(V,\mathbf{R},+,\cdot,8)$是实数域$\mathbf{R}$上的一个线性空间，它与$\mathbf{C}^n$，$\mathbf{R}^n$是不同的。而反过来，若基集取，

$$V=\{x|x=(\xi_1,\xi_2,\cdots,\xi_n)^T,\ \xi_i\in\mathbf{R}\}$$

而域$F$取复数域$\mathbf{C}$，仍按例1规定运算，则不构成一个线性空间。为什么？

**例2.** 取基集$V=\{A|A=(a_{ij})_{m\times n},\ a_{ij}\in\mathbf{C}\}$，取标量域$F=\mathbf{C}$。又设$B=(b_{ij})\in V$，$\lambda\in\mathbf{C}$。规定

1. $A=B\Longleftrightarrow a_{ij}=b_{ij}$，$1\leqslant i\leqslant m$，$1\leqslant j\leqslant n$，
2. $A+B=(a_{ij}+b_{ij})$，

3．$\lambda A=(\lambda a_{ij})$，

4．$V$中的零向量为$m\times n$阶零矩阵。

则容易验证（$V$，$\mathbf{C}$，$+$，$\cdot$，8）是一个复数域上的线性空间，记之为$\mathbf{C}^{m\times n}$。类似地可定义实数域上的线性空间$\mathbf{R}^{m\times n}$。

**例3．** 取基集$V=\left\{p(x)\mid p(x)=\sum\limits_{i=0}^{n}\alpha_i x^{n-i},\ \alpha_i\in\mathbf{R}\right\}$，取域$F=\mathbf{R}$，又设$q(x)=\sum\limits_{i=0}^{n}\beta_i x^{n-i}$，$\lambda\in\mathbf{R}$。规定

1．$p(x)=q(x)\Longleftrightarrow\alpha_i=\beta_i$，$i=0,1,\cdots,n$，

2．$p(x)+q(x)=\sum\limits_{i=0}^{n}(\alpha_i+\beta_i)x^{n-i}$，

3．$\lambda p(x)=\sum\limits_{i=0}^{n}(\lambda\alpha_i)x^{n-i}$，

4．$V$中的零向量为零多项式。

则不难验证$V$是$\mathbf{R}$上的一个线性空间（$V$，$\mathbf{R}$，$+$，$\cdot$，8），并记之为$R[x]_n$。如果把$V$改为

$$V=\left\{p(x)\mid p(x)=\sum_{i=0}^{n}\alpha_i x^{n-i},\ \alpha_i\in\mathbf{R},\ \alpha_0\neq 0\right\}$$

而其它规定不变，则不再构成一个线性空间。为什么？

**例4．** 对固定的$A\in\mathbf{C}^{m\times n}$，取基集

$$V=\{x\in\mathbf{C}^n\mid Ax=0,\ 0\in\mathbf{C}^m\}$$

取域$F=\mathbf{C}$，并定义与$\mathbf{C}^n$中相同的运算。则($V$，$\mathbf{C}$，$+$，$\cdot$,8,)是一个复数域上的线性空间，叫做矩阵$A$的零空间（或核），也叫做方程$Ax=0$的解空间，记之为$\mathcal{N}(A)$。

**例5．** 对固定的$A\in\mathbf{C}^{m\times n}$，取基集

$$V=\{y\in\mathbf{C}^m\mid y=Ax,\ x\in\mathbf{C}^n\}$$

取域$F=\mathbf{C}$，并定义与$\mathbf{C}^m$中相同的运算。则（$V$，$\mathbf{C}$，$+$，$\cdot$,8）是一个复数域上的线性空间，叫做$A$的列空间，或$A$的值域，记之为$\mathcal{R}(A)$。

## 二、基底与维数

**定义1.1.2（线性相关）** 设$V$是一个线性空间，$\{x_1, x_2, \cdots, x_n\}$是$V$中的一组向量。如果在$F$中存在一组不全为零的标量$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$，使得

$$\sum_{i=1}^{n} \alpha_i x_i = 0$$

则称向量组是线性相关的。否则，便说成是线性无关的。由单个零向量构成的向量组$\{0\}$认为是相关的，而单个非零向量认为是线性无关的。

根据定义1.1.2很容易推出：

**定理1.1.3** 若向量组$\{x_1, x_2, \cdots, x_n\}$是线性相关的，则其中至少有一个向量$x_j (1 \leqslant j \leqslant n)$能用其余的向量线性表出；若该向量组是线性无关的，则其中沒有任何向量能用其余向量线性表出。

上述定理的逆也是成立的，所以也把它当作线性相关和线性无关的定义。

**定义1.1.4（空间的基底）** 设$\{x_1, x_2, \cdots, x_n\}$是线性空间$V$中的一组线性无关的向量，并且$V$中的每个向量$x$都可用之线性表出，即在$F$中存在一组标量$\{\xi_1, \xi_2, \cdots, \xi_n\}$使得

$$x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n = \sum_{i=1}^{n} \xi_i x_i$$

则称$\{x_1, x_2, \cdots, x_n\}$为$V$的一个基底，而把标量组$(\xi_1, \xi_2, \cdots, \xi_n)^T$叫做$x$关于基底$\{x_1, x_2, \cdots, x_n\}$的坐标。通常把基底排成一个形式上的矩阵（因为它不一定是矩阵）$\mathscr{B} = (x_1 \ x_2 \ \cdots \ x_n)$。而把坐标视为一个$n$维的坐标向量，当$F = \mathbf{C}$时，坐标属于$\mathbf{C}^n$，当$F = \mathbf{R}$时，坐标属于$\mathbf{R}^n$。

如果$\mathscr{B} = (x_1, x_2, \cdots, x_n)$是空间$V$的一个基底，我们把基底向量的所有可能的线性组合记为

$$S=\left\{x \mid x=\sum_{i=1}^{n} \alpha_i x_i,\ \alpha_i \in F\right\}$$

则由线性空间的定义知$S \subseteq V$；而据基底的定义又有$V \subseteq S$。因此有$V=S$。所以，一个线性空间$V$是由它的一个基底生成的，或线性张成的。一组向量$\{y_1, y_2, \cdots, y_n\}$（不一定线性无关）线性张成的集合通常有如下记法：

$$\begin{aligned} S &= L(y_1, y_2, \cdots, y_n) \\ &= \operatorname{span}(y_1, y_2, \cdots, y_n) \\ &= [y_1, y_2, \cdots, y_n] \\ &= \left\{x \mid x=\sum_{i=1}^{n} \alpha_i y_i,\ \alpha_i \in F\right\} \end{aligned}$$

因此，$V=\operatorname{span}(x_1, x_2, \cdots, x_n)=[x_1, x_2, \cdots, x_n]$。

如果线性空间$V \neq \{0\}$，则它的基底有无穷多个。但是，它所有的基底都含有相同个数（或基数）的向量。为证明这一论断，先证明一个引理。

**引理1.1.5** 设$\{x_1, x_2, \cdots, x_n\}$是空间$V$的一个基底，$\{y_1, y_2, \cdots, y_k\}$是$V$的一组线性无关的向量。则从$\{x_1, x_2, \cdots, x_n\}$中至少可以找出一个向量$x_{i_0}$ $(1 \leqslant i_0 \leqslant n)$，使得向量组$\{x_{i_0}, y_2, \cdots, y_k\}$是线性无关的。

**证明** 设结论不真，即任取$x_i$ $(1 \leqslant i \leqslant n)$，向量组$\{x_i, y_2, \cdots, y_k\}$都是线性相关的。也就是说，存在着一组不全为零的数$\{\alpha_{1i}, \alpha_{2i}, \cdots, \alpha_{ki}\}$，使得

$$\alpha_{1i} x_i + \sum_{l=2}^{k} \alpha_{li} y_l = 0$$

首先可以断定$\alpha_{1i} \neq 0$，否则与$\{y_1, y_2, \cdots, y_k\}$是线性无关的条件矛盾。从而有

$$x_i = -\sum_{l=2}^{k} \frac{\alpha_{li}}{\alpha_{1i}} y_l,\quad i=1, 2, \cdots, n$$

另一方面，由于每个$y_j$都可用$\{x_1, x_2, \cdots, x_n\}$线性表出，

不妨设$y_1=\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n$，再将前面的$x_i$的表达式代入该式中，便有

$$y_1=-\sum_{i=1}^{n}\beta_i\sum_{l=2}^{k}\frac{\alpha_{li}}{\alpha_{1i}}y_l$$

$$=\sum_{l=2}^{k}\left(-\sum_{i=1}^{n}\beta_i\frac{\alpha_{li}}{\alpha_{1i}}\right)y_l$$

这意味着$\{y_1, y_2, \cdots, y_k\}$线性相关。与原设矛盾，故引理得证。

**定理1.1.6** 线性空间$V$的任两个基底所含向量的个数相等。

**证明** 设$\{x_1, x_2, \cdots, x_n\}$和$\{y_1, y_2, \cdots, y_k\}$是$V$的两个基底，且$k>n$。由引理1.1.5知，在$\{x_1, x_2, \cdots, x_n\}$中必有向量，不妨设为$x_1$(否则,可以对向量组$\{x_1, x_2, \cdots, x_n\}$重新赋序)，使得$\{x_1, y_2, \cdots, y_k\}$线性无关。再据引理1.1.5，又可从$\{x_2, x_3, \cdots, x_n\}$中找到向量，不妨设为$x_2$，使$\{x_1, x_2, y_3, \cdots, y_k\}$线性无关。如此做到第$n$步，便得到线性无关组$\{x_1, x_2, \cdots, x_n, y_{n+1}, \cdots, y_k\}$。但是，由于每个$y_i$都能用基底$\{x_1, x_2, \cdots, x_n\}$线性表出，故向量组$\{x_1, x_2, \cdots, x_n, y_{n+1}, \cdots, y_k\}$必是线性相关的。从而导出矛盾。这表明$k>n$是不可能的。同理可证明$k<n$也是不可能的。从而只有$k=n$，即定理得证。

定理1.1.6表明，线性空间的基底所含向量个数是刻画该空间特征的一个数字，我们把它定义为线性空间的维数。

**定义1.1.7（维数）** 如果线性空间$V$的基底含有$n$个向量，则称$V$是$n$维空间，记为$\dim V=n$。顺便指出，我们这里只是就$n$为有限数的情况来定义和证明的，因为线性代数只研究有限维的线性空间。

**定理1.1.8** $n$维线性空间$V$的任意$n$个线性无关向量都可以作为$V$的一个基底。