G·W·斯图尔特著

矩阵计算引论

JUZHEN JISUAN YINLUN

上海科学技术出版社

矩阵计算引论

G. W. 斯图尔特 著

王国荣　黄丽萍
徐锦龙　王守根　译

周彭年　徐国定　校

上海科学技术出版社

内容提要

全书共分七章，目次如下：①准备知识；②实际应用中的一些问题；③线性方程组的直接解法；④范数、极限和条件数；⑤线性最小二乘方问题；⑥特征值和特征向量；⑦ QR 算法。

本书可供从事矩阵计算的科技工作者及计算数学等有关专业的师生参考。学习本书需要微积分和线性代数方面的准备知识。

矩阵计算引论

G. W. 斯图尔特　著

王国荣　黄丽萍　徐锦龙　王守根　译

周彭年　徐国定　校

上海科学技术出版社出版

（上海瑞金二路 450 号）

新华书店上海发行所发行　上海群众印刷厂印刷

开本 850×1156 1/32　印张 12.375　字数 328,000

1980 年 1 月第 1 版　1980 年 1 月第 1 次印刷

印数 1—15,000

书号：13119·786　定价：1.40 元

译 者 序

数值线性代数是一门内容很广泛的学科，近年来发展很快. G. W. 斯图尔特所著《矩阵计算引论》(1973)是一本数值线性代数入门的较好教材.

本书取材精炼，内容比较新颖. 详细介绍了在电子计算机上使用的关于矩阵计算的一些较好的算法以及分析这些算法的理论.

主要内容有:解线性方程组、线性最小二乘方问题和特征值问题的算法以及范数理论和扰动理论. 此外，还讨论了实际应用中的一些问题,如舍入误差的影响和计算机存贮的有效应用等.

本书适应不同水平读者的需要. 它为一般读者写了一章线性代数的准备知识，同时也为具有一定水平的读者介绍了有关矩阵计算的一些较新的研究成果，这些内容大部分在有关章节后面的注释和参考文献中可以看到.

为了便于学习算法和今后用算法语言编写程序，本书还介绍了一种非形式语言 INFL.

作为一本教材,本书配有大量的习题,便于读者深入掌握有关的内容. 但本书又是一本入门的教材，故作者不得不舍掉一些还是比较有用的材料,如解大稀疏线性方程组的直接法和迭代法、非负矩阵的 Perron-Frobenious 理论等.

限于水平，译文中难免有不妥和错误之处，恳请读者批评指正，以便再版时改正。

译 者

1979 年 6 月

目　　录

第一章 预备知识

本书的主题是叙述和分析有关向量和矩阵的计算方法. 在这一章中，我们将讨论这些内容的初等基础理论. 它包括两个方面：第一部分是向量、矩阵及其运算的定义；第二部分是由向量或矩阵的思想引出的各种概念之间的抽象关系，例如线性相关、列空间等. 为了理解后面所提到的算法的描述，必须熟悉矩阵的运算，而要弄清算法的分析，又必须深入掌握矩阵的理论.

下面讨论 n 维实空间. 当 $n=2$ 时，这实际上就是欧氏平面，而当 $n=3$ 时，就是我们日常生活的 3 维空间. 因此，许多一般的定理可用平面或三维空间中的几何事实来形象化. 反过来，二维和三维空间的几何知识常常可以直接推广为 n 维空间中的一般定理. 在本章中，我们将不严格地运用这种几何观点.

一些数的矩形排列称为矩阵. 矩阵的大多数算法，是通过一连串变换将一些新的零引入矩阵，最后得到合适的简单形式. 因此，有必要用一批专门名词来描述具有各种零分布的矩阵. 同时，重要的是要知道，怎样的分布在标准的矩阵运算作用下保持不变，这些问题也将在本章中加以讨论.

1. 空间 $\mathfrak{R}^n$

n 维实空间中向量的概念，是笛卡儿坐标平面上点的表示法的一种自然推广. 在这种表示法中，选择一个特定的点作为原点，过这点作两条互相垂直的直线，称为坐标轴. 这样，平面上的每一个点 p 能表示为一个有序对 (ξ_1, ξ_2)，其中第一和第二个元素分别是由点 p 向第一、第二个坐标轴作投影而得到的（图 1）. 当然，点 p 和有序对或向量 (ξ_1, ξ_2) 是不同的对象，但是，它们之间的关系却非常密切，因此，我们常常说点 (ξ_1, ξ_2)，并且，在证明平面上

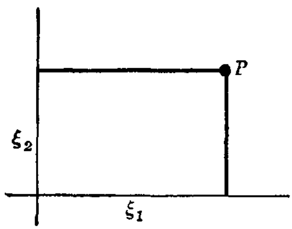

图 1

的几何定理时总是使用有序数对而不用点.

由此，我们称有序的 n 个数为 n 维向量. 并且，将这些数写成列的形式比写成行的形式更为方便，其理由以后会明白.

定义 1.1 一个 n 维向量 x 是一组按顺序排成列的实数 $\xi_1, \xi_2, \cdots, \xi_n$:

$$x=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix},$$

数 $\xi_1, \xi_2, \cdots, \xi_n$ 叫做 x 的分量.

当 n 固定时，习惯上就称向量 x. 用 $\mathfrak{R}^n$ 表示全体 n 维向量的集合，称为向量空间 $\mathfrak{R}^n$，或简称为空间 $\mathfrak{R}^n$. 实数域用 $\mathfrak{R}$ 表示，并称实数为纯量.

例题 1.2 向量空间 $\mathfrak{R}^1$ 的元素可以和域 $\mathfrak{R}$ 的元素建立自然的对应. 每一个向量 $(\alpha)\in\mathfrak{R}^1$ 与纯量 $\alpha\in\mathfrak{R}$ 对应，反过来，每个纯量 $\alpha\in\mathfrak{R}$ 与向量 $(\alpha)\in\mathfrak{R}^1$ 对应. 严格说来，空间 $\mathfrak{R}^1$ 与域 $\mathfrak{R}$ 是不同的数学对象. 但是，我们常常把它们看成是相同的，即将 $\mathfrak{R}^1$ 中的元素看作纯量，并且反过来，将纯量看作 $\mathfrak{R}^1$ 中的元素.

在全书中采用下述规定的记法：用小写希腊字母表示纯量；用小写拉丁字母表示向量；用拉丁字母所对应的希腊字母表示向量的分量. 这样，除另有说明外，纯量 α_i 就表示向量 a 的第 i 个分量. 因为希腊字母和拉丁字母之间的对应是不完全的，所以在附录 4 列出的部分对应是人为的. 特别应当注意 x 和 ξ 之间以及 y 和 η 之间的对应. 作为上述约定的一个例外，常用小写拉丁字母作为下标及求和指标.

如上所述，$\mathfrak{R}^2$ 中的一个向量 x 与平面上坐标为 ξ_1 和 ξ_2 的点相对应，同样，$\mathfrak{R}^3$ 中的向量 x 与三维空间中的坐标为 ξ_1、ξ_2 和 ξ_3

的点相对应. 通常, 用从原点到与向量对应的点画一个箭头作为这个向量的图形表示(图 2).

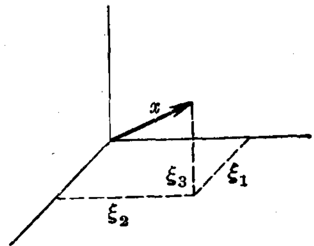

图 2

两个 n 维向量相等当且仅当它们对应的分量相等. 这样, 向量相等

$$a=b$$

等价于一组纯量的等式

$$\alpha_i=\beta_i \quad (i=1, 2, \cdots, n).$$

因此, 要证明两个向量相等, 只要证明它们的分量相等. 而要规定一个新的向量, 只要说明它的分量是怎样构成的即可.

下面讨论向量的第一个运算——两个向量的和.

定义 1.8 设 a, $b\in\mathfrak{R}^n$. a 与 b 的和记为 $a+b$, 是一个 n 维向量 c, 它的分量由下式给出:

$$\gamma_i=\alpha_i+\beta_i.$$

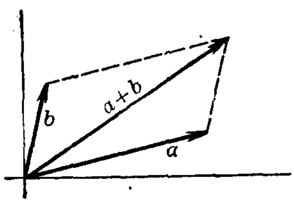

图 3

两个向量的和有如下的几何解释: 向量 a 与 b 构成一个顶点在原点的平行四边形的两邻边. 于是, a 与 b 的和就是这个平行四边形的从原点出发的对角线(图 3).

两个向量的和具有通常两个纯量和的某些性质, 如下面定理所示.

定理 1.4 设 a, b, $c\in\mathfrak{R}^n$. 则

1. $a+b=b+a$,
2. $(a+b)+c=a+(b+c)$.

证明 为了证明第一部分, 设 $x=a+b$, $y=b+a$. 于是, 由 $\mathfrak{R}$ 的性质得:

$$\xi_i=\alpha_i+\beta_i=\beta_i+\alpha_i=\eta_i.$$

因此, 由上述向量相等的意义, 得 $x=y$, 即 $a+b=b+a$.

为了证明第二部分, 设 $x=(a+b)+c$, $y=a+(b+c)$. 那末

$$\xi_i=(\alpha_i+\beta_i)+\gamma_i=\alpha_i+(\beta_i+\gamma_i)=\eta_i.$$

因此，$x=y$，即$(a+b)+c=a+(b+c)$.

由上述性质1.4中的1可推得，a_i的次序改变时，和$a_1+a_2+\cdots+a_n$是不变的，例如

$$a_1+a_2+\cdots+a_{k-1}+a_k=a_k+a_{k-1}+\cdots+a_1.$$

性质1.4中的2说明向量加法是可结合的．更一般地说，如果$a_1, a_2, \cdots, a_n\in\mathfrak{R}^n$，那末，不论和式如何组合，都与$a_1+a_2+\cdots+a_n$相同．

定义1.5 $\mathfrak{R}^n$中的零向量是指n个分量都是零的向量．

对一切n，我们也用表示纯量零的记号"0"来表示$\mathfrak{R}^n$中的零向量．零向量具有数零的某些性质．

定理1.6 设$a\in\mathfrak{R}^n$．则

1．$a+0=a$,

2．在$\mathfrak{R}^n$中有一个向量，记作$-a$，使$a+(-a)=0$.

证明 设$b=a+0$．于是，$\beta_i=\alpha_i+0=\alpha_i$，因而$b=a$，即$a+0=a$．第一部分证毕．对于第二部分，设

$$b=\begin{pmatrix}-\alpha_1\\-\alpha_2\\\vdots\\-\alpha_n\end{pmatrix},$$

并设$c=a+b$．于是

$$\gamma_i=\alpha_i+\beta_i=\alpha_i+(-\alpha_i)=0.$$

因此$c=0$，而b是第二部分中所要找的向量$-a$，即$c=a+b=a+(-a)=0$．▌

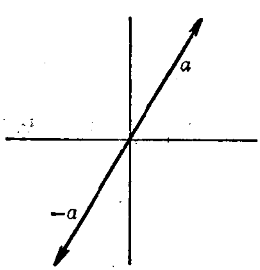

图 4

定理1.6中第二部分的向量$-a$只是由a的分量改变了符号而得到的向量．几何上看，这意味着向量a关于原点被反射(图4)．

在这里，记号"－"用于单元运算，即由已知向量a得到定理1.6中的向量$-a$．此外，还可以用这个记号表示减法的二元运算．具体地说

来，已知向量 a 和 b，可用下式定义它们的差 $a-b$：

$$a-b=a+(-b).$$

容易验证，差运算满足纯量之间减法的通常规则．例如，

$$a-b=-(b-a).$$

向量的第二个重要运算是向量与纯量相乘的运算．

定义 1.7 设 $\lambda\in\mathfrak{R}$，而 $a\in\mathfrak{R}^n$．λ 和 a 的乘积记作 $\lambda\cdot a$ 或 λa，它是一个 n 维向量 b，它的分量由

$$\beta_i=\lambda\alpha_i$$

给出．几何上看，纯量 λ 乘向量 a 的运算是用因子 $|\lambda|$ 改变 a 的长度．如果 λ 是负数，则 a 还要经过原点反射．

定理 1.8 设 λ，$\mu\in\mathfrak{R}$，而 a，$b\in\mathfrak{R}^n$．则

1．$(\lambda\mu)a=\lambda(\mu a)$，

2．$(\lambda+\mu)a=\lambda a+\mu a$，

3．$\lambda(a+b)=\lambda a+\lambda b$，

4．$1\cdot a=a$．

证明 只证明第 3 条，其余的证明留作练习．在说明定理时，按通常先乘后加的惯例．$\lambda a+\mu b$ 的意思即 $(\lambda a)+(\mu b)$．设 $x=\lambda(a+b)$，$y=\lambda a+\lambda b$．于是

$$\xi_i=\lambda(\alpha_i+\beta_i)=(\lambda\alpha_i)+(\lambda\beta_i)=\eta_i,$$

这就证明了定理的第三部分． ▌

在上面的讨论中，是用实数来定义向量及它们的运算的．因此，定理 1.4，1.6，1.8 中所列举的这些性质，都是实数性质的直接推论．也可以把这些定理的性质作为公理来描述定义在某个对象集合 $\mathscr{V}$ 上的和与积的性质．具体地说，如果

1．在 $\mathscr{V}$ 的元素之间定义和“$+$”，它满足定理 1.4 所列的性质；

2．$\mathscr{V}$ 有一个元素“0”，它满足定理 1.6 所列的性质；

3．在实数和 $\mathscr{V}$ 的元素之间定义积“$\cdot$”，它满足定理 1.8 所列的性质．

集合 $\mathscr{V}$ 就称为(在实数域上的)抽象向量空间.

抽象向量空间的元素不一定是 n 维实数组. 如下例所示:

例题 1.9 设 $\mathscr{V}$ 是定义在 $[0, 1]$ 上的全体实值函数的集合. 如果 f, $g\in\mathscr{V}$, 定义 $h=f+g$ 是一个函数, 它的值是 $h(\xi)=f(\xi)+g(\xi)$, $\xi\in[0, 1]$. 如果 $f\in\mathscr{V}$, 且 $\lambda\in\Re$, 定义 $h=\lambda f$ 是一个函数, 它的值是 $h(\xi)=\lambda f(\xi)$. 设"0"是在 $[0, 1]$ 上恒为零的函数. 根据这些定义, $\mathscr{V}$ 就是一个向量空间.

用抽象向量空间的这些性质可以证明一些并不显而易见的定理. 但是, 这种证明可能是繁琐的. 例如, 下面来证明定理: 如果 $\mathscr{V}$ 是抽象向量空间, 而 $a\in\mathscr{V}$, 则 $0\cdot a=0$. 在证明中, 保证每一个等式成立的性质列在旁边(顺便指出: 在等式 $0\cdot a=0$ 中记号"0"有两种用法, 在左边的 0 是纯量零, 在右边的 0 是零向量).

$$
\begin{aligned}
0&=a+(-a), && 1.6.2,\\
&=1\cdot a+(-a), && 1.8.4,\\
&=(0+1)a+(-a), && 1+0=1,\\
&=(0\cdot a+1\cdot a)+(-a), && 1.8.2,\\
&=0\cdot a+[1\cdot a+(-a)], && 1.4.2,\\
&=0\cdot a+[a+(-a)], && 1.8.4,\\
&=0\cdot a+0, && 1.6.2,\\
&=0\cdot a, && 1.6.1.
\end{aligned}
$$

而另一方面, 在 $\Re^n$ 中验证这个事实是容易的. 设 $b=0\cdot a$. 则 $\beta_i=0\cdot\alpha_i=0$; 因此 $b=0$. 常常发生这样情况, 涉及抽象向量空间的一些定理对 $\Re^n$ 空间来说往往是极其显然的. 因为在本章中我们只讨论 $\Re^n$, 所以不建立抽象向量空间的理论. 凡是用实数性质容易证明的事实, 在此将用而不证, 而将它的验证留作练习.

练　习

1. 完成下列计算:

(a) $\begin{pmatrix}1\\-3\end{pmatrix}+2\begin{pmatrix}-1\\1\end{pmatrix}$,　　(b) $\begin{pmatrix}\xi_1\\\xi_2\end{pmatrix}-\begin{pmatrix}\eta_1\\\eta_2\end{pmatrix}$,

(c) $\beta\begin{pmatrix}1\\ \alpha\\ \alpha^2\end{pmatrix}+\begin{pmatrix}1\\1\\1\end{pmatrix}$,　　(d) $\begin{pmatrix}0\\1\\0\end{pmatrix}+\tau\begin{pmatrix}2\\10\\1\end{pmatrix}$,

(e) $\xi_1\begin{pmatrix}1\\0\\0\end{pmatrix}+\xi_2\begin{pmatrix}0\\1\\0\end{pmatrix}+\xi_3\begin{pmatrix}0\\0\\1\end{pmatrix}$.

2. 在平面上画出下列向量轨迹的草图.

(a) $\begin{pmatrix}\xi\\ |\xi|\end{pmatrix}$, $-\infty<\xi<\infty$,

(b) $\begin{pmatrix}\xi\\ \xi^2\end{pmatrix}$, $-\infty<\xi<\infty$,

(c) $\begin{pmatrix}1\\ -1\end{pmatrix}+\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$, $0\leqslant\theta\leqslant 2\pi$,

(d) $\begin{pmatrix}\cosh\tau\\ \cosh\tau\end{pmatrix}$, $-\infty<\tau<\infty$.

3. 完成定理 1.8 的证明.

4. 在 $\mathfrak{R}^n$ 中建立下列恒等式.

$$a+(b+(c+d))=((a+b)+c)+d,$$
$$(-1)\cdot a=-a,$$
$$a+b+c+d=d+c+b+a.$$

证明在任何抽象向量空间里上述恒等式都成立.

5. 证明在 $\mathfrak{R}^n$ 中,

$$\lambda x=0 \Rightarrow \lambda=0 \text{ 或者 } x=0.$$

6. 详细验证例题 1.9 的集合 $\mathscr{V}$ 是抽象向量空间.

7. 对函数 $\phi:[0, 1]\to\mathfrak{R}$, 如果有点 $0=\xi_0<\xi_1<\xi_2<\cdots<\xi_n=1$ 和常数 η_1, η_2, $\cdots$, η_n, 使对一切 $\xi\in(\xi_{i-1}, \xi_i)$ 都有 $\phi(\xi)=\eta_i$, 则称函数 ϕ 是**阶梯函数**. 象例题 1.9 那样, 定义阶梯函数的和及其与纯量的积, 并用"0"表示恒为 0 的函数. 用这些定义, 证明在[0, 1]上的所有阶梯函数的集合是一个抽象向量空间.

2. 线性无关性、子空间和基底

给定了向量集合 $\mathscr{X}$, 通过反复应用向量的加法及其与纯量相乘的运算可以得到新的向量. 这样得到的向量可以写成下列形式

$$a_1x_1+a_2x_2+\cdots+a_mx_m, \tag{2.1}$$

其中 $x_1, x_2, \cdots, x_m \in \mathscr{X}$. 下面对形如(2.1)的向量给出有关的定义.

定义 2.1 设 $x_1, x_2, \cdots, x_m \in \mathfrak{R}^n$, 且 $\alpha_1, \alpha_2, \cdots, \alpha_m \in \mathfrak{R}$. 那末, 向量(2.1)叫做 $x_1, x_2, \cdots, x_m$ 的一个线性组合. 数 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 叫做这个线性组合的系数. 如果 $\alpha_1=\alpha_2=\cdots=\alpha_m=0$, 这个线性组合就叫做平凡的; 否则, 叫做非平凡的.

例题 2.2 设 $e_i(i=1, 2, \cdots, n)$ 表示一个 n 维向量, 它的第 i 个分量为 1, 其余分量都为 0. 于是, 对任一 n 维向量 x 有

$$x=\xi_1 e_1+\xi_2 e_2+\cdots+\xi_n e_n. \tag{2.2}$$

换句话说, 任一 n 维向量是 $e_1, e_2, \cdots, e_n$ 的一个线性组合. 这些向量以后将不断出现, 我们就用记号 $e_i(i=1, 2, \cdots, n)$ 来专门表示它们.

例题 2.3 设 x, y 是空间 $\mathfrak{R}^3$ 中不在同一直线上的两个元素, 则 x, y 的所有线性组合的集合是通过点 x, y 和原点的平面(图 1). 这个平面称为由 x, y 张成的平面.

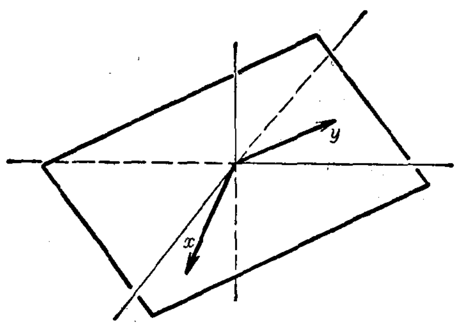

图 1

在例题 2.3 中, x, y 的所有非平凡的线性组合都不能是零向量; 因为, 若 $\alpha x+\beta y=0$, 不妨设 $\alpha\neq 0$, 则 $x=-\dfrac{\beta}{\alpha}y$, 因此 x 和 y 位于同一直线上. 另一方面, 若 z 位于 x、y 张成的平面上, 那末, 存在两个纯量 α, β 使 $\alpha x+\beta y-1\cdot z=0$. 因此, 对于象 $\{x, y\}$ 这类集合, 其中的元素是相互独立的, 以及对于象 $\{x, y, z\}$ 这类集合,

其中的元素是相互依赖的，都有必要给出适当的名称.

定义 2.4 设 $\mathscr{X}\subset\mathfrak{R}^n$. 如果 $\mathscr{X}$ 的元素的非平凡线性组合都不是零向量，则 $\mathscr{X}$ 的元素称为是**线性无关**的，否则，$\mathscr{X}$ 的元素称为是**线性相关**的.

$\mathscr{X}$ 的元素是线性无关的一个等价的说法是，如果 $x_1, x_2, \cdots, x_m\in\mathscr{X}$，且 $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m=0$，则 $\alpha_1=\alpha_2=\cdots=\alpha_m=0$. 因此，要证明 $x_1, x_2, \cdots, x_m$ 线性无关，只要假定 $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m=0$，然后证明所有的 α_i 都是零即可. 另一方面，要证明 $x_1, x_2,\cdots, x_m$ 线性相关，必须找出不全为零的纯量 $\alpha_1, \alpha_2, \cdots, \alpha_m$，使 $\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m=0$.

例题 2.5 在 $\mathfrak{R}^n$ 中，假定

$$a=\alpha_1e_1+\alpha_2e_2+\cdots+\alpha_ne_n=0.$$

那末由于 α_i 是 a 的第 i 个分量，所以它必须是零. 因而，**向量 $e_1, e_2, \cdots, e_n$ 线性无关**.

例题 2.6 包含零向量的任一向量集合都有线性相关的元素，因为线性组合 $1\cdot 0$ 是非平凡的，而且它的确是零向量.

假定 $\mathscr{X}$ 的元素线性相关，那末存在向量 $x_1, x_2, \cdots, x_m\in\mathscr{X}$ 和不全为零的纯量 $\alpha_1, \alpha_2, \cdots, \alpha_m$，使

$$\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m=0. \tag{2.3}$$

具体说来，假定 $\alpha_k\neq0$，则由(2.3)，可以解出 x_k:

$$x_k=-\frac{\alpha_1}{\alpha_k}x_1-\cdots-\frac{\alpha_{k-1}}{\alpha_k}x_{k-1}-\frac{\alpha_{k+1}}{\alpha_k}x_{k+1}-\cdots-\frac{\alpha_m}{\alpha_k}x_m. \tag{2.4}$$

这样，$\mathscr{X}$ 中向量的任何线性组合可以写成不出现向量 x_k 的形式. 就线性组合而论，x_k 并不提供额外的信息.

由线性无关的向量组成的集合有一个重要的性质，即它的向量的不同的线性组合给出不同的结果. 确切地说，有下面的定理.

定理 2.7 设 $x_1, x_2, \cdots, x_m\in\mathfrak{R}^n$ 线性无关. 假定有

$$y=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_mx_m \tag{2.5}$$

和

$$y=\beta_1x_1+\beta_2x_2+\cdots+\beta_mx_m. \tag{2.6}$$

则

$$\alpha_i=\beta_i \quad (i=1,\ 2,\ \cdots,\ m). \tag{2.7}$$

证明 从(2.5)和(2.6),可以推出

$$0=y-y=(\alpha_1-\beta_1)x_1+(\alpha_2-\beta_2)x_2+\cdots+(\alpha_m-\beta_m)x_m.$$

因为向量 x_k 线性无关,所以有

$$\alpha_i-\beta_i=0, \quad 即 \quad \alpha_i=\beta_i \quad (i=1,\ 2,\ \cdots,\ m). \quad \rbrack$$

设 z_1 和 z_2 是例题 2.3 中所述平面中的两个向量.那末存在某些纯量 α_1, α_2, β_1 和 β_2 使 $z_1=\alpha_1x+\beta_1y$ 和 $z_2=\alpha_2x+\beta_2y$. 线性组合 $\gamma_1z_1+\gamma_2z_2$ 可以写成如下形式

$$\gamma_1z_1+\gamma_2z_2=(\gamma_1\alpha_1+\gamma_2\alpha_2)x+(\gamma_1\beta_1+\gamma_2\beta_2)y,$$

它仍是 x, y 的线性组合,因此也在这个平面上. 这样,平面上的任何两个向量的线性组合仍在这个平面上.由于这个原因,这个平面就叫做 $\mathfrak{R}^3$ 的子空间. 这个想法可以推广到 $\mathfrak{R}^n$ 的任意子集.

定义 2.8 设 $\mathscr{S}$ 是 $\mathfrak{R}^n$ 的非空子集. 如果

1. $x,\ y\in\mathscr{S}\Rightarrow x+y\in\mathscr{S}$,

2. $x\in\mathscr{S}$ 且 $\alpha\in\mathfrak{R}\Rightarrow \alpha x\in\mathscr{S}$,

则称 $\mathscr{S}$ 为 $\mathfrak{R}^n$ 的子空间.

从这个定义可以推出, 如果 $\mathscr{S}$ 是 $\mathfrak{R}^n$ 的子空间, 则 $\mathscr{S}$ 中的元素的线性组合仍在 $\mathscr{S}$ 中. 反过来, $\mathfrak{R}^n$ 的任何非空子集,如果它对元素的线性组合来说是封闭的,则它是一个子空间.

例题 2.9 空间 $\mathfrak{R}^n$ 本身是 $\mathfrak{R}^n$ 的一个子空间.

只由零向量组成的集合{0}是 $\mathfrak{R}^n$ 的平凡子空间.

例题 2.3 中的平面是由向量 x 和 y 的全部线性组合形成的,这就启示了下述生成子空间的方法.

定理 2.10 设 $\mathscr{A}\subset\mathfrak{R}^n$ 是非空的, 又设 $\mathscr{S}$ 是 $\mathscr{A}$ 中元素的全部线性组合的集合,则 $\mathscr{S}$ 是 $\mathfrak{R}^n$ 的一个子空间.

证明 设 $x,\ y\in\mathscr{S}$,则对某些向量 $a_1,\ a_2,\ \cdots,\ a_m\in\mathscr{A}$ 和纯量 $\sigma_1,\ \cdots,\ \sigma_m,\ \tau_1,\ \cdots,\ \tau_m$, 有 $x=\sum_{i=1}^m\sigma_ia_i$ 和 $y=\sum_{i=1}^m\tau_ia_i$. 于是, $x+y=\sum_{i=1}^m(\sigma_i+\tau_i)a_i\in\mathscr{S}$. 同样,如果 $x=\sum_{i=1}^m\sigma_ia_i\in\mathscr{S}$,则 $\alpha x=\sum_{i=1}^m(\alpha\sigma_i)a_i$

$\in\mathscr{S}$. ▌

定理 2.10 的这个子空间 $\mathscr{S}$ 称为由 $\mathscr{A}$ 生成的子空间. 称 $\mathscr{A}$ 的所有元素张成 $\mathscr{S}$. 这样，例题 2.3 的那个平面是由集合 $\{x, y\}$ 生成的，由向量 x 和 y 所张成.

在许多情形下，用生成子空间的集合来处理问题比用子空间本身来得方便，因为前者可以只包含有限个向量，而后者如果是非平凡的，就一定包含无限多个向量. 如果生成集合的元素线性相关，那末由例题 2.6 后面的观察，可以推出这个集合的向量中至少有一个是多余的. 因而，在研究生成集合时，只要研究其元素是线性无关的那些就可以了. 这样的集合称为子空间的一个基底.

定义 2.11 设 $\mathscr{S}\subset\mathfrak{R}^n$ 是一个子空间，又设 $\mathscr{B}\subset\mathscr{S}$. 如果

1. $\mathscr{B}$ 的元素线性无关，
2. $\mathscr{B}$ 生成 $\mathscr{S}$，

则 $\mathscr{B}$ 是 $\mathscr{S}$ 的一个基底.

如果 $\mathscr{B}=\{b_1, b_2, \cdots, b_m\}$ 是 $\mathscr{S}\subset\mathfrak{R}^n$ 的一个基底，且 $x\in\mathscr{S}$，则由 2.11.2，对于某些数 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 有

$$x=\gamma_1 b_1+\gamma_2 b_2+\cdots+\gamma_m b_m.$$

由定理 2.7，数 $\gamma_1, \gamma_2, \cdots, \gamma_m$ 由向量 x 唯一确定. 它们叫做 x 关于基底 $\mathscr{B}$ 的分量.

例题 2.12 向量

$$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}$$

形成 $\mathfrak{R}^2$ 的一个基底. 向量 x 关于这个基底的分量是 $(\xi_1+\xi_2)/2$ 和 $(\xi_1-\xi_2)/2$.

例题 2.13 由例题 2.5，向量 $e_1, e_2, \cdots, e_n\in\mathfrak{R}^n$ 线性无关. 由例题 2.2，它们张成 $\mathfrak{R}^n$. 因此，集合 $\{e_1, e_2, \cdots, e_n\}$ 是 $\mathfrak{R}^n$ 的一个基底. 从等式(2.2)，我们看到向量 x 关于基底 $\{e_1, e_2, \cdots, e_n\}$ 的分量是数 $\xi_1, \xi_2, \cdots, \xi_n$，它们恰是向量 x 在通常意义下的分量.

至此，还有一个没有解决的问题，即 $\mathfrak{R}^n$ 的任意一个子空间是否都有基底. 在回答这个问题时，还要证明 $\mathfrak{R}^n$ 的一个给定子空间