最优设计的数学方法

程　极　泰　编著

国防工业出版社

最优设计的数学方法

程极泰　编著

国防工业出版社

内 容 简 介

本书介绍最优设计所涉及到的最优化数学方法，着重介绍近年来常用的单纯形直接解法和它的改型使用的方法；以及序列无约束极小化方法，也就是SUMT解法。

第一章是一些必要的数学基础知识。第二章到第六章是最优设计问题的介绍，以及和最优设计有关的最优化的数学方法。第七章是具体计算和计算程序的介绍，它们的详细计算程序放在书末。

本书可供从事工程设计技术人员，以及科研、教学人员参考。

最优设计的数学方法

程极泰　编著

*

国防工業出版社 出版

新华书店北京发行所发行　　各地新华书店经售

国防工业出版社印刷厂印装

*

$850\times1168\ 1/32$　印张 $8\ 3/4$　221千字

1981年11月第一版　1984年7月第二次印刷　印数：7,701—15,200册

统一书号：15034·2249　　定价：1.10元

前　言

近年来，工程技术中的很多设计问题，越来越多地希望通过适当的数学方法，经由电子计算机的计算，得到最优设计的效果。这种最优设计问题，通常能够表示成一种静态的数学模型，从这种模型可以实现的设计变量中，应用求极值的技巧，就可以得到最优设计解。

在最优化的设计问题中，大量地是数学规划中的非线性数学规划问题。我们曾经和上海船舶运输科学研究所的张燮元等同志一起进行过一些最优设计计算，取得较好的结果。这本《最优设计的数学方法》就是在和有关的几个单位共同实践基础上，结合国内外有关参考文献编写成的。

本书的主要内容是介绍求解非线性规划的直接法和序列无约束极小化方法，并给出这两种方法的详细计算程序。

序列无约束极小化方法要引用变尺度迭代方法，为了使读者能了解第五章介绍的无约束极小化方法的有关数学内容，我们在第一章简要的给出了一些要引用到的线性代数基础知识。为了要清楚说明序列无约束极小化方法在最优设计问题中的作用，我们在第二章中列出一些实例，在第三章中比较了各种数学规划方法，在第四章中讨论了使用这些数学方法的一些准则。第六章是关于第七章计算程序中直接法方面的一些说明。

本书第七章中的两种程序是分别由何焕熹、张克邦二同志试算和应用过的，也是分别由他们作详细的文字介绍的。本书在编写过程中曾得到许多单位的同志大力帮助，在此表示衷心的感谢。

编写这本书，虽然是基于一些实践基础之上，但总的说来，还是在匆忙中完成的，有不妥和错误之处请广大读者批评指正。

目　　录

第一章 预 备 知 识

§1 矩阵概念和矩阵的运算

1. n 维实线性空间 R^n

在一个平面上，选定一点为原点，过原点作两条互相垂直的直线构成坐标轴，在坐标轴上选定单位长以后，就可以把平面上的每一点 p 用一对实数 ξ_1、ξ_2 有序的表为序对 (ξ_1, ξ_2)，它的第一个元 ξ_1 和第二个元 ξ_2 各是 p 点在第一条（横）轴上和第二条（纵）轴上的投影（见图 1-1）。我们把序对 (ξ_1, ξ_2) 称为向量，因为它也可以被看成为从原点到 p 点这个径向向量的两个分量。当然，p 点和序对，或者向量，都是不同的对象，但是因为它们之间的关系是很明显的，我们将不加什么区别的，看我们的需要，说它是一点、一个序对或一个向量。

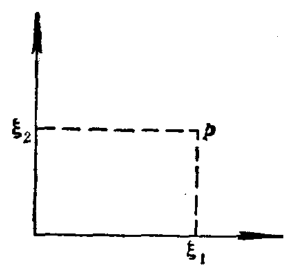

图1-1 平面上 p 点的坐标

这种平面向量的概念可以很自然地推广到一般的实 n 维空间中的向量概念，也就是用一个有序 n 数组表示出一般的向量，定义：

n 向量：顺序排成一列的 n 个实数 $\xi_1, \xi_2, \cdots, \xi_n$ 的整体

$$\mathbf{x}=\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \tag{1-1}$$

称为一个 n 向量 $\mathbf{x}$，实数 $\xi_1, \xi_2, \cdots, \xi_n$ 称为向量 $\mathbf{x}$ 的 n 个分量。

在 n 固定时，或者从上下文已经知道 n 的数值时，我们就把 n 向量 $\mathbf{x}$ 简称为 $\mathbf{x}$。我们用 R^n 表示所有 n 向量的集合，称这个集合是一个向量空间 R^n，或者简称为空间 R^n。

实数集合表为 R，它的每一个元是一个实数，是一个纯量。但是，向量空间 R^1 的各个元自然是和 R 的各个元对应的，每一个向量

$$(\alpha) \in R^1$$ ●

相应的有一个纯量

$$\alpha \in R$$

反之亦然。严格地说，空间 R^1 和实数集合 R 是不同的数学对象，但是我们在这里常把它们等价的使用。

两个 n 向量的相等：设 $\mathbf{a}$，$\mathbf{b} \in R^n$，且

$$\mathbf{a} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \text{ 和 } \mathbf{b} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \tag{1-2}$$

只有当它们的对应元都相等时，才叫做相等，也就是当且仅当纯量等式组

$$\alpha_i = \beta_i \qquad i = 1, 2, \cdots, n \tag{1-3}$$

成立时，才有向量等式

$$\mathbf{a} = \mathbf{b} \tag{1-4}$$

的成立。

两个 n 向量的和：设 $\mathbf{a}, \mathbf{b} \in R^n$，$\mathbf{a}$，$\mathbf{b}$ 的和写成 $\mathbf{a}+\mathbf{b}$，它是一个 n 向量 $\mathbf{c}$，它的各个分量是：

$$\gamma_i = \alpha_i + \beta_i \qquad i = 1, 2, \cdots, n \tag{1-5}$$

设 $\mathbf{a}$，$\mathbf{b}$，$\mathbf{c} \in R^n$，则有加法运算下的交换律和结合律：

（1）$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$

● $\in$ 表示属于的意思。

（2）$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$

因为，设 $\mathbf{x}=\mathbf{a}+\mathbf{b}$，$\mathbf{y}=\mathbf{b}+\mathbf{a}$，且

$$\mathbf{x}=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix} \quad 和 \quad \mathbf{y}=\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix}$$

则从实数 R 的性质，直接有：

$$\xi_i=\alpha_i+\beta_i=\beta_i+\alpha_i=\eta_i \qquad i=1,2,\cdots,n$$

所以 $\mathbf{x}=\mathbf{y}$，也就是（1）是成立的。再设：

$$\mathbf{x}=(\mathbf{a}+\mathbf{b})+\mathbf{c}$$
$$\mathbf{y}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$$

则有：

$$\xi_i=(\alpha_i+\beta_i)+\gamma_i=\alpha_i+(\beta_i+\gamma_i)=\eta_i$$
$$i=1,2,\cdots,n$$

所以 $\mathbf{x}=\mathbf{y}$，也就是有（2）的性质。

零向量：R^n 中的零向量是 n 个分量都是零的一个 n 向量，这里仍用“o”表示 R^n 中的这个零向量。

设 $\mathbf{a}$ 是 R^n 中的任何一个向量，则它相关于零向量有两个重要性质：

（3）$\mathbf{a}+\mathbf{o}=\mathbf{a}$

（4）在 R^n 中有 $\mathbf{a}$ 的一个逆向量，写成 $(-\mathbf{a})$，它使

$$\mathbf{a}+(-\mathbf{a})=\mathbf{o}$$

因为设 $\mathbf{b}=\mathbf{a}+\mathbf{o}$，则有：

$$\beta_i=\alpha_i+0=\alpha_i \qquad i=1,2,\cdots,n$$

（等式中的 0 是纯量的 0），所以 $\mathbf{b}=\mathbf{a}$，故有（3）。再设

$$\mathbf{b}=\begin{pmatrix}-\alpha_1\\ -\alpha_2\\ \vdots\\ -\alpha_n\end{pmatrix} \quad 且 \quad \mathbf{c}=\mathbf{a}+\mathbf{b}$$

则有：

$$\gamma_i=\alpha_i+\beta_i=\alpha_i+(-\alpha_i)=0 \qquad i=1,2,\cdots,n$$

所以 $\mathbf{c}=\mathbf{o}$，并且 $\mathbf{b}$ 就是（4）中所要找的逆向量，即

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

这里的符号“－”表示关于向量的一种一元运算。我们还用这个符号表示“减法”的二元运算，也就是，在给定向量 $\mathbf{a}$ 和 $\mathbf{b}$ 以后，我们用“－”这个符号定义它们的差是：

$$\mathbf{a}-\mathbf{b}=\mathbf{a}+(-\mathbf{b}) \tag{1-6}$$

容易证明，差运算满足纯量之间的减法规律，例如：

$$\mathbf{a}-\mathbf{b}=-(\mathbf{b}-\mathbf{a}) \tag{1-7}$$

纯量和向量的相乘：设 $\lambda\in R$，$\mathbf{a}\in R^n$，则 λ 和 $\mathbf{a}$ 的乘积写成 $\lambda\cdot\mathbf{a}$ 或 $\lambda\mathbf{a}$，它是一个 n 向量，它的分量是：

$$\beta_i=\lambda\alpha_i \qquad i=1,2,\cdots,n \tag{1-8}$$

设 λ，$\mu\in R$，$\mathbf{a}$，$\mathbf{b}\in R^n$，则纯量和向量相乘的运算有下列四个性质：

（5）$(\lambda\mu)\mathbf{a}=\lambda(\mu\mathbf{a})$

（6）$(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}$

（7）$\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$

（8）$1\cdot\mathbf{a}=\mathbf{a}$

我们有了式（1-8）的定义，就可以把任何一个 n 向量看成是 $1\cdot\mathbf{a}$ 的运算结果，同时 $\lambda\mathbf{a}+\mu\mathbf{b}$ 是表示 $(\lambda\mathbf{a})+(\mu\mathbf{b})$，即先纯量乘再做加法运算，就以（7）的等式来说，则：

设　$\mathbf{x}=\lambda(\mathbf{a}+\mathbf{b}) \qquad \mathbf{y}=\lambda\mathbf{a}+\lambda\mathbf{b}$

因

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

$$i=1,2,\cdots,n$$

所以（7）是成立的。其余证明类似。

从上面的分析看来，向量空间 R^n 在加法运算，纯量相乘运算，以及零向量和逆向量的存在，给出从（1）到（8）的线性运算性质，所以，我们常常把满足这八个性质的向量空间 R^n 称为线性空间。

一般地说，一个线性空间 S 是指一种集合，它的元不必是由一切 n 个有序实数列成的 n 向量，而可以是在实数（甚至是复数）域上的一个抽象向量空间，只要：

一是对 S 的各个元可以定义一种“＋”的加法运算，使它满足（1）（2）的关系，即当 $\mathbf{a},\mathbf{b},\mathbf{c}\in S$ 时，则有：

（1）$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$

（2）$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c})$

二是在 S 中有一个“$\mathbf{o}$”元，它使（3）（4）成立，也就是当 $\mathbf{a}\in S$ 时，则有：

（3）$\mathbf{a}+\mathbf{o}=\mathbf{a}$

（4）在 S 中有 $\mathbf{a}$ 元的一个逆元，写成 $(-\mathbf{a})$，它使

$$\mathbf{a}+(-\mathbf{a})=\mathbf{o}$$

三是对 S 的各个元可以给出和实数的一种“·”相乘运算，它使（5）（6）（7）（8）成立，也就是当 $\lambda,\mu\in R$，$\mathbf{a},\mathbf{b}\in S$ 时，则有：

（5）$(\lambda\mu)\cdot\mathbf{a}=\lambda\cdot(\mu\cdot\mathbf{a})$

（6）$(\lambda+\mu)\cdot\mathbf{a}=\lambda\cdot\mathbf{a}+\mu\cdot\mathbf{a}$

（7）$\lambda\cdot(\mathbf{a}+\mathbf{b})=\lambda\cdot\mathbf{a}+\lambda\cdot\mathbf{b}$

（8）$1\cdot\mathbf{a}=\mathbf{a}$

例如，设 S 是所有定义在（0，1）上的实值函数的集合。若 f，$g\in S$，定义 $h=f+g$ 是一个函数，它的值是：

$$h(\xi)=f(\xi)+g(\xi),\qquad \xi\in(0,1)$$

若 $f\in S$，$\lambda\in R$，定义 $h=\lambda f$ 是一个函数，它的值是：

$$h(\xi)=\lambda f(\xi),\qquad \xi\in(0,1)$$

设“$\mathbf{o}$”是在（0，1）上恒等于零的函数，则 S 将满足（1）

到（8）的性质，所以一切定义在（0，1）上的实函数集合S形成一个线性空间，或抽象向量空间。

有了线性空间S的基本性质，无论S是否是n向量空间R^n或是更一般的抽象向量空间，总可以根据线性空间的定义导出其它更进一步的性质。例如，若S是一个抽象向量空间，且$\mathbf{a}\in S$，则有：

$$0\cdot\mathbf{a}=\mathbf{o}$$

左边的零是纯量，右边的零是抽象零向量。证明它成立的步骤如下：

$$
\begin{aligned}
\mathbf{o}&=\mathbf{a}+(-\mathbf{a}) && \text{(由（4））}\\
&=1\cdot\mathbf{a}+(-\mathbf{a}) && \text{(由（8））}\\
&=(0+1)\cdot\mathbf{a}+(-\mathbf{a}) && \text{(由（3））}\\
&=(0\cdot\mathbf{a}+1\cdot\mathbf{a})+(-\mathbf{a}) && \text{(由（6））}\\
&=0\cdot\mathbf{a}+[1\cdot\mathbf{a}+(-\mathbf{a})] && \text{(由（2））}\\
&=0\cdot\mathbf{a}+[\mathbf{a}+(-\mathbf{a})] && \text{(由（8））}\\
&=0\cdot\mathbf{a}+\mathbf{o} && \text{(由（4））}\\
&=0\cdot\mathbf{a} && \text{(由（3））}
\end{aligned}
$$

当然，线性空间最常见的还是n向量空间R^n，这种线性空间又可称为n维向量空间，它之所以称为“n维”，主要的不是因为它是由n个有序实数列成的，而是从一般线性空间的“基底”概念来加以说明的。基底的概念是建立在线性空间中的线性独立的概念上的。

2．线性独立向量

设m个向量$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m\in R^n$和$m$个纯量$\alpha_1, \alpha_2, \cdots, \alpha_m\in R$经过加法运算和纯量相乘的运算，得到新的$n$向量是：

$$\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\cdots+\alpha_m\mathbf{x}_m \tag{1-9}$$

则称这个n向量是m个n向量$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m$的一个“线性组合”，其中纯量$\alpha_1, \alpha_2, \cdots, \alpha_m$称为线性组合的系数。如果一种组合是由$\alpha_1=\alpha_2=\cdots=\alpha_m=0$构成时，则称线性组合是“平凡的”，否则，

线性组合是一种非平凡的线性组合，简称为“非凡的”。

通常用 $\mathbf{e}^{(i)}(i=1,2,\cdots,n)$ 表示 n 个 n 向量，$\mathbf{e}^{(i)}$ 的第 i 个分量是 1，其余的分量都是零。这样，任何一个 n 向量 $\mathbf{x}$ 就可以写成：

$$\mathbf{x}=\xi_1\mathbf{e}^{(1)}+\xi_2\mathbf{e}^{(2)}+\cdots+\xi_n\mathbf{e}^{(n)} \tag{1-10}$$

其中

$$\mathbf{x}=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix},\quad \mathbf{e}^{(1)}=\begin{pmatrix}1\\0\\0\\ \vdots\\0\end{pmatrix},\quad \mathbf{e}^{(2)}=\begin{pmatrix}0\\1\\0\\ \vdots\\0\end{pmatrix},\cdots\mathbf{e}^{(n)}=\begin{pmatrix}0\\0\\0\\ \vdots\\1\end{pmatrix} \tag{1-11}$$

换句话说，任何一个 n 向量总是 $\mathbf{e}^{(1)},\mathbf{e}^{(2)},\cdots,\mathbf{e}^{(n)}$ 这 n 个 n 向量的一个线性组合。

设 X 是 R^n 空间中一部分 n 向量构成的一个集合，称 X 是 R^n 的一个子集，表示成：

$$X\subset R^n$$

如果在 X 中，没有任何的一些元组成的非凡线性组合能够是零向量，则称 X 的元是线性独立的。或者等价地说，若 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m\in X$，且

$$\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\cdots+\alpha_m\mathbf{x}_m=\mathbf{o}$$

则 $\alpha_1=\alpha_2=\cdots=\alpha_m=0$，那么，称 X 中的 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m$ 是“线性独立的”。反之，若有不全为零的纯量 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 使 $\alpha_1\mathbf{x}_1+\alpha_2\mathbf{x}_2+\cdots+\alpha_m\mathbf{x}_m=0$，则称 $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m$ 是“线性相关的”。

在式（1-11）中给出的 R^n 向量空间中的 n 个向量 $\mathbf{e}^{(1)},\mathbf{e}^{(2)},\cdots,\mathbf{e}^{(n)}$ 是线性独立的向量，因为如果它们的一个线性组合能表出一个零向量 $\mathbf{a}=\mathbf{o}$，即

$$\mathbf{a}=\alpha_1\mathbf{e}^{(1)}+\alpha_2\mathbf{e}^{(2)}+\cdots+\alpha_n\mathbf{e}^{(n)}=\mathbf{o}$$

由于 α_i 就是 $\mathbf{a}$ 的第 i 个分量，它必须是零，所以就有 $\alpha_1=\alpha_2=\cdots=\alpha_n=0$ 的结论。

在 R^n 中，可以由 $m<n$ 个线性独立向量张成一个子空间，也就是定义：

子空间：设 S 是 R^n 的一个非空子集，若

（1） $\mathbf{x}, \mathbf{y} \in S \Longrightarrow \mathbf{x}+\mathbf{y} \in S$

（2） $\mathbf{x} \in S,\ \alpha \in R \Longrightarrow \alpha\mathbf{x} \in S$

则称 S 是 R^n 的一个子空间。这里使用的符号“$\Rightarrow$”表示“若…则…”的语句。

基底：设 $S \subset R^n$ 也就是 S 是 R^n 的一个子空间，且设 $B \subset S$，若

（3） B 的各元是线性独立的；

（4） B 产生了 S，即 B 中各元的任一线性组合都是 S 的元，反之，S 中任何一元都是 B 中各元的一个线性组合。

则称 B 是 S 的一个“基底”。

如果 S 是 R^n 的一个非凡子空间，则 S 的基底所包含的元的个数，称为 S 这个子空间的维。由于 $\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \cdots, \mathbf{e}^{(n)}$ 是 R^n 的基底，所以我们说向量空间 R^n 是一个 n 维线性空间。

向量空间的概念是我们讨论矩阵概念的基础。

3．矩阵

矩阵产生于很多具体的有关问题，最重要的一种直接联系就是用矩阵来表示线性变换。

所谓线性变换，是指一种附合线性运算关系的函数。以从 R^2 到 R^2 的一种函数或映射的简单情况来说，我们用符号

$$f: R^2 \longrightarrow R^2$$

表示一个二自变量，值域是二维的函数或映射，f 把 R^2 中任何一个二向量映射到 R^2 中一个二向量，即

$$\mathbf{y}=f(\mathbf{x}),\quad \mathbf{x} \in R^2,\quad \mathbf{y} \in R^2$$

或者用分量写出是：

$$\eta_1=f_1(\xi_1, \xi_2)$$
$$\eta_2=f_2(\xi_1, \xi_2)$$

如果对于任何向量 $\mathbf{x}, \mathbf{y} \in R^2$ 和纯量 $\alpha, \beta \in R$，函数或映

射 f 具有下面性质：

$$f(\alpha \mathbf{x}+\beta \mathbf{y})=\alpha f(\mathbf{x})+\beta f(\mathbf{y}) \tag{1-12}$$

则称函数 f 是一个线性变换。

事实上，线性变换 f 是完全由它在基底向量 $\mathbf{e}^{(1)}$ 和 $\mathbf{e}^{(2)}$ 上的数值唯一决定的。也就是，设

$$f(\mathbf{e}^{(1)})=\mathbf{a}^{(1)}=\begin{pmatrix}\alpha_{11}\\ \alpha_{21}\end{pmatrix},\quad f(\mathbf{e}^{(2)})=\mathbf{a}^{(2)}=\begin{pmatrix}\alpha_{12}\\ \alpha_{22}\end{pmatrix} \tag{1-13}$$

则对任何的 $\mathbf{x}=\xi_1\mathbf{e}^{(1)}+\xi_2\mathbf{e}^{(2)}$，从（1-12）可以有：

$$\begin{aligned}\mathbf{y}=f(\mathbf{x})&=f(\xi_1\mathbf{e}^{(1)}+\xi_2\mathbf{e}^{(2)})=\xi_1 f(\mathbf{e}^{(1)})+\xi_2 f(\mathbf{e}^{(2)})\\&=\xi_1\mathbf{a}^{(1)}+\xi_2\mathbf{a}^{(2)}\end{aligned}$$

也就是，用分量表出为：

$$\eta_1=\alpha_{11}\xi_1+\alpha_{12}\xi_2,\quad \eta_2=\alpha_{21}\xi_1+\alpha_{22}\xi_2$$

它完全由 2×2 的阵列（即二行二列的阵列）

$$\begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{pmatrix} \tag{1-14}$$

决定。（1-14）这个阵列就称为表示线性变换 f 的矩阵。

一般地说，如果给出一种映射

$$f: R^n \longrightarrow R^m$$

对于任何向量 $\mathbf{x}$，$\mathbf{y}\in R^n$ 和纯量 α，$\beta\in R$ 也有线性变换性质（1-12），则表示线性变换 f 的对应矩阵将是由 m 行 n 列的一个 $m\times n$ 矩阵

$$\mathbf{A}=\begin{pmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n}\\ \cdots & \cdots & & \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn}\end{pmatrix} \tag{1-15}$$

表示，阵列中的 $m\times n$ 个实数（或在其它需要的情况下是复数）α_{ij}（$i=1,2,\cdots,m$，$j=1,2,\cdots,n$）称为矩阵 A 的元。

所有 $m\times n$ 矩阵的集合用 $R^{m\times n}$ 表示。

矩阵（1-15）右边的表示方法，也有用方括弧的，写成：

$$\mathbf{A}=\begin{bmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n}\\ \cdots & \cdots & & \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn}\end{bmatrix}$$

更多地是用罗马小体字母代替希腊字母写出矩阵的各元，也就是：

$$\mathbf{A}=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix} \tag{1-16}$$

并且常常简称为$m\times n$矩阵，写成$\mathbf{A}=(a_{ij})$。

我们常常把$n\times 1$矩阵

$$\mathbf{C}=\begin{pmatrix}a_{11}\\ a_{21}\\ \vdots\\ a_{n1}\end{pmatrix} \tag{1-17}$$

看成是和一个n向量的同一表达形式，从而在$R^{n\times 1}$和R^n之间就不加什么区别。

类似地，我们称$1\times n$矩阵

$$\mathbf{R}=(a_{11}, a_{12}, \cdots, a_{1n}) \tag{1-18}$$

是一个行向量，而且在式（1-18）中二元之间加上逗号主要是防止书写上混乱，并且一个行向量并不用来和n维空间的坐标点对应，我们仍然是以列向量（1-17）和点做等价对应的描述。

在$m\times n$矩阵（1-16）中，若$m=n$，则称$\mathbf{A}$是一个n阶方阵。如果一个方阵的非零元仅在对角元$a_{ii}(i=1,2,\cdots,n)$上出现，则称这个方阵是一个对角阵，写成：

$$\begin{pmatrix}\delta_1 & 0 & \cdots & 0\\ 0 & \delta_2 & \cdots & 0\\ \cdots & \cdots & & \\ 0 & 0 & \cdots & \delta_n\end{pmatrix}=\mathrm{diag}(\delta_1, \delta_2, \cdots, \delta_n) \tag{1-19}$$

如果一个 n 阶方阵 $\mathbf{A}$ 的各元有

$$a_{ij}=a_{ji} \qquad i, j=1, 2, \cdots, n \tag{1-20}$$

的性质时，则称为一个对称矩阵。

我们常常看具体的需要将一个矩阵分块为子矩阵表示出来，例如一个 4 × 4 的矩阵

$$\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

可用 2 × 2 的子矩阵写成分块矩阵的形式，即

$$\mathbf{A}=\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

其中

$$\mathbf{A}_{11}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \mathbf{A}_{12}=\begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix}$$

$$\mathbf{A}_{21}=\begin{pmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix}, \qquad \mathbf{A}_{22}=\begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix}$$

设 $\mathbf{A}\in R^{m\times n}$，$\mathbf{A}$ 的第 j 列是 m 向量的列向量，即

$$\mathbf{a}^{(j)}=\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad j=1, 2, \cdots, n$$

则 $\mathbf{A}$ 可以用各列向量分块写成：

$$\mathbf{A}=(\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \cdots, \mathbf{a}^{(n)})$$

类似地，若 $\mathbf{A}$ 的第 i 行是行向量：

$$\mathbf{b}^{(i)}=(a_{i1}, a_{i2}, \cdots, a_{in}) \qquad i=1, 2, \cdots, m$$

则 $\mathbf{A}$ 可以用各行向量分块写成：