本书出版由上海发展汽车工业教育基金会资助

汽车优化设计理论与方法

张宝生　李　杰　林明芳　编著

机 械 工 业 出 版 社

本书主要讲述优化设计理论基础及数学模型，单变量函数的优化方法，无约束条件下、有约束条件下多变量函数的寻优方法，模糊优化设计的基本原理，内燃机工作过程及结构参数的最优化，汽车传动系参数和主要总成结构的最优化。

本书可作为高等工科院校汽车、拖拉机、特种车辆、内燃机、机械设计等专业高年级学生和研究生教材或教学参考书，亦可供从事有关设计与研究的工程技术人员参考。

图书在版编目（CIP）数据

汽车优化设计理论与方法/张宝生等编著．—北京：机械工业出版社，2000.9

ISBN 7-111-08144-7

Ⅰ．汽…　Ⅱ．张…　Ⅲ．汽车-整体设计　Ⅳ．U462.2

中国版本图书馆CIP数据核字（2000）第63758号

机械工业出版社（北京市百万庄大街22号　邮政编码100037）

责任编辑：孙慧波　孙本绪　版式设计：张世琴　责任校对：唐海燕

封面设计：方　芬　责任印制：

印刷厂印刷·新华书店北京发行所发行

2000年9月第1版·第1次印刷

787mm×1092mm1/16·9印张·217千字

0 001—3 000册

定价：16.00元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

本社购书热线电话（010）68993821、68326677－2527

前　　言

最优化技术在工程设计、计划管理、生产控制和科学实验等领域已得到了广泛应用，获得了很大的效益。优化设计是现代设计方法的主要内容，又是计算机辅助设计的核心部分，机械优化设计是以数学规划论为理论基础，运用计算机寻求机械设计最优参数的现代先进设计方法。采用这种方法可以使设计方案按预定目标达到最完善的地步，并带来显著的技术经济效益。目前，在汽车行业中进行优化设计的研究与应用已得到充分重视，为了大力推广这一先进技术，作者根据从事汽车优化设计的教学与科学研究实践，编写了此书，为学习和从事这一先进设计方法的工程设计与研究人员提供参考和帮助。本书力求做到系统性强，体现出基础与专业、理论与应用、教学与科研相结合的特色。

全书分四部分：第一部分为汽车优化设计基础理论，第二部分为汽车优化设计数学模型，第三部分为汽车优化设计的方法，第四部分为结合内燃机、汽车底盘主要总成设计的实例，介绍了典型机构的优化设计。

本书可以作为高等工科院校汽车、拖拉机、特种车辆、内燃机和普通机械类等专业高年级学生和研究生教材或教学参考书，亦可供从事有关设计与研究的工程技术人员参考。

本书吸收了许多单位在开展汽车优化设计研究工作中的成果和经验，先后得到有关教授和专家的审阅，并提出了宝贵意见，对此均表示衷心的感谢。书中难免有缺点和错误，恳请读者批评指正。

目　　录

第一章　概　　述

第一节　最优化技术及其发展

人们在做一切工作时，总希望所选用的方案是一切可能方案中最好的方案，这就是最优化问题。下面举一些例子予以说明。

安排生产计划时,如何在现有人力、物力条件下,合理安排产品生产,使总产值为最高。

确定生产工艺过程，如汽车生产，如何在保证汽车质量的前提下，选择最合理生产过程，使生产费用最省。

进行产品设计，例如设计一个变速器，如何在满足性能要求的前提下，使其体积最小、重量最轻。

配料方面，如何合理配料，在保证质量要求的前提下使成本最低。

交通运输，例如火车由甲站开往乙站，如何在保证安全行驶的条件下，使时间最省。又如汽车运输，如何选择合理的路线，使运输费用最省。

农业生产，例如利用温室生产蔬菜，应如何合理调节室内的温度、湿度，使蔬菜生产周期最短或产量最高。

国防上，例如多级火箭的发射问题，如何在规定时间内，烧完规定的燃料，使其达到的速度最大；或在规定时间内，达到某一速度，而使燃料最省。

从以上例子可以看出，在各个生产、科研领域中，普遍地存在着最优化问题。

最优化技术就是研究和解决最优化问题的一门学科。它研究和解决如何在一切可能的方案中寻求最优化的方案。换言之，最优化技术研究和解决两大类问题：如何将最优化问题表示成数学模型；如何根据数学模型尽快地求出其最优解。

按处理最优化问题的数学方法的不同,最优化技术发展分为两个阶段,第二次世界大战以前,处理最优化问题的数学方法主要是古典的微分法和变分法。第二次世界大战中,由于军事上的需要产生了运筹学,提出了大量不能用上述古典方法解决的最优化问题,从而产生了如线性规划、非线性规划、动态规划等新的方法。此后,最优化的理论和方法逐渐得到丰富和发展。特别从 60 年代以来,最优化技术发展迅速,成为一门新兴的学科,而且得到了广泛的应用。

我国开始从事这方面的研究与应用比较晚。值得注意的是，虽然在机械设计中采用最优化技术的历史很短，但其进展的速度却是十分惊人的。无论在机构综合、通用机械零部件设计，还是在各种专业机械和工艺装备的设计都由于采用了最优化技术而取得了显著成果。汽车工业的新技术不断引进和采用，特别是汽车 CAD、CAE、CAM 一体化技术的进步，为我国汽车工业形成独立自主开发能力创造了良好的条件，也大大提高了工程技术人员素质和水平，其中汽车优化设计理论与方法，得到了推广和普及。

促进最优化技术迅速发展的主要原因是：

(1) 近代科技与生产发展的需要。随着工程与技术的复杂化、大型化与精密化，新材料、新工艺、新技术的不断出现，机械产品的更新换代周期也日益缩短。与此相适应，迫

切需要提高设计质量、寻找最优设计方案、缩短设计周期，这就促使最优化技术迅速发展。

（2）电子计算机的飞速发展为最优化技术提供了有力的手段。由于最优化技术是要在一切可能的方案中寻求最优的方案，往往需要进行大量的计算。若没有电子计算机而用人工进行计算，不仅工作量很大，且有时是难以实现的。而有了计算机这一有效的计算工具，就使最优化技术得以迅速发展。

运用计算机进行机械最优化设计，对整个机械设计学科产生了十分深刻的影响，使许多过去无法解决的关键性问题获得了重大突破，并取得显著的经济效益与社会效益。机械优化设计作为一种新兴的技术，尽管目前还不很完善，有许多问题有待进一步研究探索，但可以预期，随着现代技术的迅速发展，最优化技术将在汽车行业中获得更广泛更有效的应用与发展。

第二节　优化设计在设计过瘂中的应用

设计一个好产品，应使所设计出来的产品既有良好的性能，又能满足生产的工艺性、使

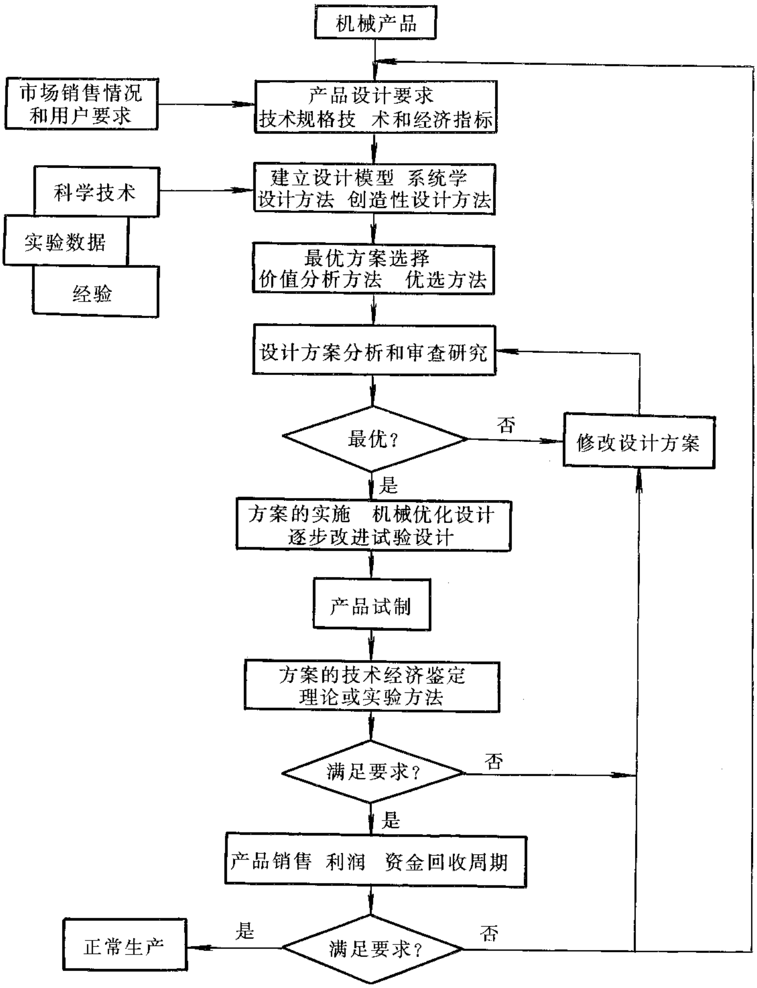

图 1-1　机械设计工作循环流程图

用的可靠性和安全性，且费用最省、消耗最低，这是一切设计活动的最终目的。要想达到此目的，其核心部分是设计过程中如何优选设计方案。图 1-1 表示机械设计工作循环流程。在传统的设计中，其全部过程皆由人来完成。

由于时间和费用的关系，所能提供的方案数非常有限，再加上分析和计算工具的限制，要想取得一个最优方案很不容易。随着设计过程的计算机化，广大设计人员迫切要求为自动选取最优设计方案建立一种迅速而有效的方法。机械优化设计就是在这种迫切要求下产生和发展起来的一种自动选优方法。这种设计方法在传统设计或计算机辅助设计中都起着重要作用，并且成为设计工作的核心部分。

计算机辅助设计　即人、机交互作用设计系统）就是根据设计题目的要求，按照某种设计思想，安排出设计模型，输入原始数据，送入中央处理机，经过分析和计算过程，以数据或图象形式输出设计方案。在此过程中，设计者还可以通过人机联系的会话系统适当修改参数，直至得到满意的设计方案，最终输出图象　自动绘图）及其他设计文件等。这种设计方法最大的优点就是能查询计算机的设计过程和由于修改参数而形成的结果　输出中间结果）。这就是说，可以根据设计者的意图利用计算机给出不同的设计方案。

由于在设计过程中需要不断地评选方案和选择最优参数，因此，在计算机辅助设计过程中引入优化设计起到了重要的作用。图 1-2 表示计算机辅助设计和优化设计的工作流程图。

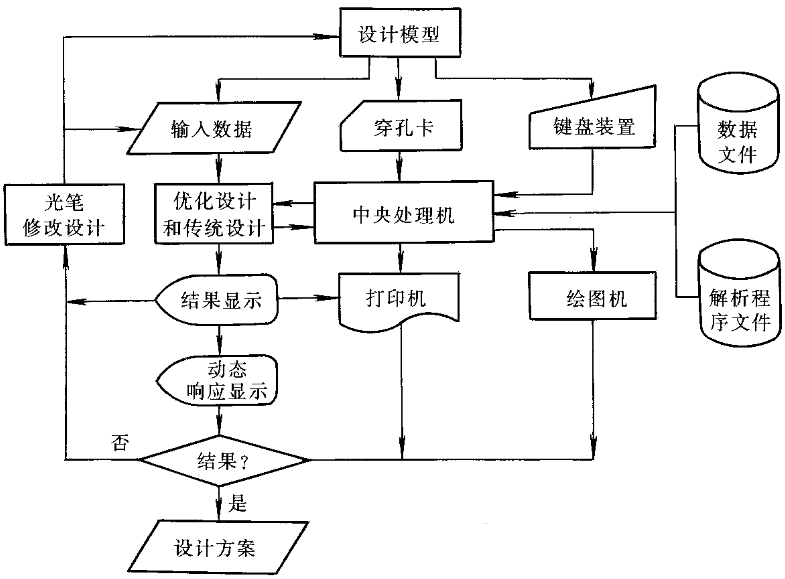

图 1-2　计算机辅助设计工作流程图

第三节　优化设计的一般过程及其技术经济效果

一、优化设计的一般过程

机械优化设计的一般过程与传统设计方法有所不同，它是以计算机自动设计选优为其基

本特征的，其过程分为四个阶段。

(1) 工程设计问题的提出　首先决定设计目标，它可以是单项设计指标，也可是多项设计指标的组合。从技术经济观点出发，机器的运动学和动力学性能、体积、重量、效率、成本、可靠性等都可以作为设计所追求的目标，然后分析设计应满足的要求。主要有以下三类：一类是某些设计参数的取值范围；二类是由某种设计性能或指标根据设计规范推导出的性能要求；三类是工艺条件对某些设计参数的限制等。

(2) 建立数学模型　将以上工程设计问题用数学方程式的形式予以全面地、准确地描述，其中包括：根据设计目标建立起评价设计方案优劣的目标函数；把设计应满足的各类要求以等式或不等式的形式建立约束方程；确定哪些参数参与选优，也就是确定设计变量。这里，一是要准，必须严格地按各种规范建立相应的数学描述；二是要全，必须把设计中应考虑的各种因素全部包括进去，这两点对于整个优化设计的效果是至关重要的。

(3) 选择优化方法　根据数学模型中函数的性质，设计的精度要求等选择适用的优化方法，并做出相应的程序设计。

(4) 得出最优设计方案　上机计算，并自动解得最优值，然后对计算结果做出分析和正确的判断，得出最优设计方案。

二、机械优化设计的技术经济效果

近十几年来机械优化设计研究的发展情况表明，优化设计已愈来愈多地应用于产品设计中，如零部件的优化设计、机构的优化设计、工艺装备基本参数的优化设计等，而且取得了显著的经济效果。

目前，对在机械设计中最优化方法的实际应用还很难做出准确的估价。但是，有很多实例可以说明，与传统的设计方法比较，采用优化设计后经济效果显著。例如，对某一大型减速器进行优化设计，使其重量比原方案减轻 12%；对一行星减速器进行优化设计，使其体积比原设计缩小 13%；对 20 台桥式起重机箱式主梁进行优化设计，使重量比原型平均减轻 14%，最多的减轻 35%。另外，对各种机构进行优化设计，不仅可以改善机构的动力学性能，还可以提高运动精度。一般来说，设计问题愈复杂，优化设计结果取得的技术经济效益也就愈显著。尽管机械优化设计方法还处在不断发展的过程中，但仅从目前已完成的一些工作实例来看，这种方法的推广和应用对机械工业必将进一步提高产品质量、降低成本、缩短生产周期。

汽车工业随着国民经济发展和交通运输体系的全面建立得到了飞速发展。目前，形成我国汽车独立自主开发能力仍是汽车工业发展中的关键之一。汽车的电子技术和计算机技术为现代汽车的科技进步，提供了两种有力的手段。汽车产品开发和科学管理中，都采用了现代的计算机辅助设计，而优化设计又是其灵魂和核心。汽车优化设计理论和方法已应用于汽车诸多领域中的很多环节，如汽车整车动力传动系统优化和匹配，汽车的发动机、底盘、车身各主要总成的优化设计、机械加工的优化设计、汽车车身 CAD/CAE/CAM 一体优化技术等，使汽车产品的性能和水平得到提高，生产的科学管理得到加强。汽车优化设计是一种有着广泛应用领域和良好效益的先进技术。

第二章　优化设计理论基础

为了讨论机械最优化设计问题，本章对于与优化方法有关的矩阵、集合、矢量、多元函数及其极值问题等数学基本知识作一简单介绍。这些概念无论在实际中还是理论上都有着重要的价值。

第一节　矩　　阵

在物理、力学、工程等各领域中，各种因素之间存在着错综复杂的联系，需要用一定的数学形式来描述和表达。其中较简单的也是较常见的表达式是线性方程组，矩阵的概念就是从线性关系中得出来的。

一、矩阵的概念

设一线性方程组

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \tag{2-1}$$

如果把上面式子中的系数按原来的顺序排列起来，记作下面的形式：

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

它就被称为矩阵。

矩阵定义：由一组数（或符号）排列成具有 m 行与 n 列的长方形“表”

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

称为 $m \times n$ 阶矩阵，简记为

$$A = (a_{ij}) \qquad \begin{aligned} i &= 1, 2, \cdots, m \\ j &= 1, 2, \cdots, n \end{aligned} \tag{2-3}$$

矩阵 A 是由 $m \times n$ 个数（或符号）组成的。其中每一个数（或符号）称为“元素”，简称“元”。用符号 a_{ij} 表示。每一横线上各元素组成一行，每一竖线上各元素组成一列，任意元 a_{ij} 的脚标 i、j 表示这个元的位置在第 i 行，第 j 列。

矩阵的符号常见的还有

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{或} \quad \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix}$$

方阵：当矩阵的行数与列数相同时，即 $m = n$，称为正方矩阵，或称方阵，如

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{1-4}$$

方阵 A 叫 n 阶方阵。矩阵 A 的元素 a_{11}、a_{22}、…、a_{nn} 称为方阵的主对角元素。但绝不可将方阵与行列式的概念相混淆，两者尽管都是一组数排列的表，但行列式可通过运算求出其值，即行列式是代表一个值的，而方阵不然，它只是一个“表”而已。

由方阵 A 的全部元素构成的行列式，称为矩阵 A 的行列式，记为 $|A|$：

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{1-5}$$

$|A|$ 有其值，若按行列式的第 i 行展开，有

$$|A| = \sum_{k=1}^{n} a_{ik} A_{jk}$$

式中 A_{jk} 为元素 a_{ik} 的代数余子式，i 可取 1，2，…，n。

例如矩阵

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$$

则 A 的行列式

$$|A| = \begin{vmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{vmatrix} = 2$$

当方阵 A 的行列式 $|A| = 0$，称 A 为奇异方阵；$|A| \neq 0$，称 A 为非奇异方阵。

行矩阵与列矩阵：仅有一行的矩阵（当 $m = 1$ 时）称为（$1 \times n$）阶行矩阵：

$$A = [a_1 \quad a_2 \quad \cdots \quad a_n] \tag{1-6a}$$

仅有一列的矩阵（当 $n = 1$ 时）称为（$m \times 1$）阶列矩阵：

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \tag{1-6b}$$

单位方阵：在 n 阶方阵中，当主对角元均为 1，其余各元素都为零，则称做单位方阵，

并用特定符号 E 表示，即

$$E=\begin{bmatrix}1 & & & \text{对称}\\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & 1\end{bmatrix} \tag{1-7}$$

在矩阵代数中，单位矩阵相当于一般代数中纯 1 的概念。

零矩阵：若矩阵的全部元素均为零，称为零矩阵，并以符号Ⓗ 表示，零矩阵可以是 $m\times n$ 阶的长方形矩阵，也可以是 n 阶方阵。

二、矩阵的转置

若将原矩阵 A 的行与列对换成列与行来写，就得到 A 的转置矩阵，用 A^T 表示，即

$$A=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}\ (m\times n\ \text{阵})$$

$$A^T=\begin{bmatrix}a_{11} & a_{21} & \cdots & a_{m1}\\ a_{12} & a_{22} & \cdots & a_{m2}\\ \vdots & \vdots & \vdots & \vdots\\ a_{1n} & a_{2n} & \cdots & a_{mn}\end{bmatrix}\ (n\times m\ \text{阶}) \tag{1-8}$$

同样，行矩阵的转置为列矩阵，列矩阵的转置为行矩阵，如

$$\begin{pmatrix}a_1 & a_2 & \cdots & a_n\end{pmatrix}^T=\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{bmatrix}$$

三、对称方阵

当方阵具有 $A=A^T$，也即各元素 $a_{ij}=a_{ji}$的性质时，称 A 为对称方阵。其全部元素沿主对角线呈对称分布，例如

$$A=\begin{bmatrix}1 & 8 & -1 & 3\\ 8 & 4 & 5 & 2\\ -1 & 5 & 7 & -3\\ 3 & 2 & -3 & 6\end{bmatrix}$$

四、矩阵的运算

1. 矩阵相等

若有矩阵 A 与B，它们的阶数相同，并且各对应元素完全相等，即 $a_{ij}=b_{ij}$，则该两矩阵称为相等，记作 $A=B$。否则，若至少有一对相应元素不相等，称作两个矩阵不相等，记作 $A\neq B$。

2. 矩阵的加减

两个同阶数的矩阵 A 与B 可以进行加减运算，其和或差 C 亦为同阶矩阵。矩阵 C 中各元素为矩阵 A、B 中各对应元素之和或差。即若有矩阵的运算关系：

$$C = A \pm B$$

则必有相对应元素的对应关系

$$c_{ij} = a_{ij} \pm b_{ij} \tag{2-9}$$

例如
$$\begin{bmatrix} 1 & 3 & 6 & 1 \\ -1 & 2 & 3 & 2 \\ 5 & -4 & 1 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 2 & -1 & 0 \\ 2 & -5 & 0 & 6 \\ 1 & -1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 5 & 5 & 1 \\ 1 & -3 & 3 & 8 \\ 6 & -5 & 3 & 6 \end{bmatrix}$$

矩阵加法满足交换律和结合律，设有同阶矩阵 A、B、C，则有

$$A + B = B + A \tag{2-10}$$

$$(A + B) + C = A + (B + C) \tag{2-11}$$

3. 矩阵的乘法

(1) 若以数 λ 乘矩阵 A，得同阶矩阵，记 $C = \lambda A$，规定 C 中各元素就是 A 中各元素乘以 λ，即 $c_{ij} = \lambda a_{ij}$。

$$C = \lambda A = \lambda \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{bmatrix} \tag{2-12}$$

(2) 若以两个矩阵 A 与 B 相乘，则必须 A 的列数等于 B 的行数时才可以进行这种运算，它的乘积仍是一个矩阵 C，C 的行数同 A，C 的列数同 B，C 的第 i 行、第 j 列的元素 c_{ij} 等于 A 中第 i 行各元素 a_{i1}，a_{i2}，…，a_{ip} 与 B 中第 j 列各元素 b_{1j}，b_{2j}，…，b_{pj} 逐对相乘之积的总和，即

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{ip} b_{pj} \tag{2-13}$$

$$C = AB = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ a_{21} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{p1} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ c_{21} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}$$

例 1
$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix}$$

例 2
$$(x_1 \quad x_2 \quad x_3) \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = (x_1 y_1 + x_2 y_2 + x_3 y_3)$$

例 3
$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} (b_1 \quad b_2 \quad b_3) = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{bmatrix}$$

关于矩阵相乘的次序问题，在一般情况下，矩阵的乘法不满足交换率，即 $AB \neq BA$，这可以举一例子予以说明。设有矩阵 A、B 为

$$A=\begin{bmatrix}1&2&1\\-1&0&2\\3&1&0\end{bmatrix}\quad B=\begin{bmatrix}1&0&-1\\2&1&3\\-1&0&4\end{bmatrix}$$

则按矩阵乘法有

$$AB=\begin{bmatrix}4&2&9\\-3&0&9\\5&1&0\end{bmatrix}\quad \text{而}\ BA=\begin{bmatrix}-2&1&1\\10&7&4\\11&2&-1\end{bmatrix}$$

故矩阵相乘的先后次序不容忽视。但在个别情况下，矩阵相乘的次序也可以互换，即 $AB=BA$，则称 A 与 B 为彼此可换。例如：

$$\begin{pmatrix}1&2\\-2&0\end{pmatrix}\begin{pmatrix}-3&2\\-2&-4\end{pmatrix}=\begin{pmatrix}-7&-6\\6&-4\end{pmatrix}$$

$$\begin{pmatrix}-3&2\\-2&-4\end{pmatrix}\begin{pmatrix}1&2\\-2&0\end{pmatrix}=\begin{pmatrix}-7&-6\\6&-4\end{pmatrix}$$

可换的情形只有在 A 与 B 均为同阶方阵时才能出现，否则两种乘积 AB 与 BA 的阶数不同，永不可能相等。综上所述，矩阵的相乘应按先后次序进行。

至于矩阵乘法的结合律则是永远成立的（证明从略）。若 A 的列数等于 B 的行数，又 B 的行数等于 C 的列数，则有

$$(AB)C=A(BC) \tag{1-14}$$

利用矩阵乘法可以简洁地写出线性方程组的矩阵表达式。已知线性方程组：

$$\left.\begin{array}{l}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\\vdots\qquad\ \vdots\qquad\ \vdots\qquad\ \vdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n\end{array}\right\}$$

写成矩阵形式

$$\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\vdots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}$$

令

$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\vdots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{bmatrix}\quad X=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\quad B=\begin{bmatrix}b_1\\b_2\\\vdots\\b_n\end{bmatrix}$$

则有

$$AX=B$$

这就是上述线性方程组的矩阵表达式。

关于矩阵乘积的某些性质

（1）当两矩阵之积为 0 时，并不意味着其中之一必为零矩阵，例如

$$\begin{bmatrix} 3 & -1 & 2 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

（2）当存在 $AB=AC$ 的关系时，$B=C$ 的关系不一定成立。

例如

$$\begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix}$$

（3）当矩阵 A 与单位方阵相乘时，其积仍为 A，即

$$EA = A \quad \text{或} \quad AE = A \tag{2-15}$$

（4）乘积的转置

$$(AB)^T = B^T A^T \tag{2-16}$$

4. 逆矩阵

对于一个 n 阶方阵 A（非奇异方阵），如果另有一个 n 阶方阵 B，能满足两者之积等于单位方阵，即 $AB=E$ 时，则 B 叫做 A 的逆矩阵，记作 $B=A^{-1}$。一个矩阵如果有逆矩阵，就叫它为可逆矩阵。逆矩阵是唯一的，显然由此推知

$$\begin{aligned} AA^{-1} &= A^{-1}A = E \\ (A^{-1})^{-1} &= A \end{aligned} \tag{2-17}$$

由此看，A 也是 A^{-1} 的逆阵。关于逆矩阵是矩阵运算中的一个重要问题，具体求法不再介绍，可参考有关数学书籍。

前面已提过数学方程组写成矩阵的形式

$$AX = B$$

若矩阵 A 是非奇异的（即 $|A| \neq 0$），则以 A^{-1} 左乘上式等号两端，所以

$$A^{-1}AX = A^{-1}B$$

因有 $AA^{-1} = E$，则

$$EX = A^{-1}B$$

$$X = A^{-1}B \tag{2-18}$$

这里，只要求出系数矩阵的逆阵 A^{-1}，再求出乘积 $A^{-1}B$，即可求出未知量 X。在普通代数中，没有表达其解的数学形式，而用矩阵方法，这点做到了，并且形式极为简洁。

5. 矩阵的正定

（1）二次型　所谓二次型是指含有 n 个变量 x_1，x_2，…，x_n 的一个二次齐次函数，形如：

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = {} & a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + \cdots + a_{nn}x_n^2 \\ & + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\ & + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_3 \\ & + \cdots + 2a_{(n-1)n}x_{n-1}x_n \end{aligned}$$

又可写成形式

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n \\ & + \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x^2 \end{aligned}$$

$$= \sum_{ij=1}^{n} a_{ij} x_i x_j \qquad (a_{ij} = a_{ji}) \tag{2-19}$$

对于这样一个一般形式的二次型，总可以通过某种线性变换 $X = PY$，其中 $Y = (y_1, y_2, \cdots, y_n)^T$，$P$ 是一个常系数矩阵，把二次型中含 X_i、X_j $(i \neq j)$ 的一切都消去，只剩下变量平方项的和，得到 $f(y_1, y_2, \cdots, y_n) = a_1 y_1^2 + a_2 y_2^2 + \cdots + a_n y_n^2$
即原二次型变成形如

$$f(x_1, x_2, \cdots, x_n) = a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2 \tag{2-20}$$

这种形式的二次型叫做二次型的标准型（或叫法式）。

一般形式的二次型即式（2-19）又可以写成如下形式

$$\begin{aligned} f(x_1 x_2 \cdots x_n) = {} & x_1 (a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n) \\ & + x_2 (a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n) + \cdots \\ & + x_n (a_{n1} x_1 + a_{n2} x_2 + \cdots + a_{nn} x_n) \end{aligned}$$

按上式规律，它可以写成矩阵乘积的形式

$$f = (x_1 \quad x_2 \quad \cdots \quad x_n) \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{2-21}$$

若记

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则式（2-21）可简记为

$$f = X^T A X \tag{2-22}$$

该式为二次型函数的矩阵表示式，式中 A 称为二次型 f 的矩阵。由于 $a_{ij} = a_{ji}$（见式（2-19）），故 A 为对称矩阵。

（2）矩阵的正定与负定　设有实二次型

$$f = X^T A X$$

或

$$f = \sum_{ij=1}^{n} a_{ij} X_i X_j \qquad (a_{ij} = a_{ji})$$

对于任意的非零 X，恒有 $X^T A X > 0$，即对于不完全为零的任何实数 x_1、x_2、$\cdots$、x_n 都有 $f > 0$，则称此二次型为正定的，而其对应的矩阵 A 也相应叫做正定矩阵。

若对于不全为零的任何实数 x_1、x_2、$\cdots$有 $f \geqslant 0$ 或 $f < 0$、$f \leqslant 0$，则称此二次型分别为半正定、负定或半负定的，而相应的矩阵 A 也分别称为半正定、负定和半负定的。

（3）判定矩阵正定的方法　在矩阵理论的专著中有一些判定矩阵正定性的方法，这里只介绍其中的一种。

定理　若对称矩阵 A 为正定，其充分必要条件为其各阶主子式（即对应的各阶行列式）

均大于零。即 n 阶方阵

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

为正定的充分必要条件为

$$a_{11}>0,\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\cdots,\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ a_{21} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}>0,\cdots,A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}>0 \tag{1-23}$$

第二节　集　　合

一、集合及其记号

集合是数学中最简单、最基本的概念之一。一般可以把集合理解为具有某种共同性质的一些对象组成的全体。例如某方程所有根的集合，全体自然数的集合，直线上所有点的集合，在平面上与一定点等距离的所有点的集合（圆周）等。

集合的元素：组成某一集合的那些对象叫做该集合的元素。因此，在上述集合的例子中，方程根、每个自然数、直线上和圆周上的各点都分别为相应的各集合的元素。

记号：习惯上用大写字母 A，B，C，R^n 等表示集合，而用小写字母 a，b，c，x，y 表示元素。若 a 是集合 A 的元素，则记为 $a\in A$，读作“a 属于 A”，否则记为 $a\overline{\in}A$，读做“a 不属于 A”。

若某一集合 A 是由某些元素构成的，则可以用括号 { } 将这些元素括起来表示这个集合。例如 A 由 1，2，3，4 四个元素组成，则表示为：$A=\{1，2，3，4\}$。又例如由全体自然数组成的集合可写作 $N=\{1，2，3，\cdots\}$。仅含有限个元素的集合叫做有限集合，例如 $B=\{b_1，b_2，b_3，b_5\}$。含有无限个元素组成的集合叫做无限集合，例如自然数的集合 $N=\{1，2，3，\cdots\}$。

如果一个集合 S 是由具有某种共同性质 P 的元素 x 组成的，把这样的集合记作 $S=\{$元素 $x\mid x$ 具有性质 $P\}$。例如，xOy 平面内直线 $x+y=3$ 上的一切点所成的集合 A 记为

$$A=\{\text{点}(x,y)\mid x+y=3\}$$

二、子集的概念

如果集合 A 的每一个元素都属于集合 B，则称 A 为 B 的子集，记为

$$A\subset B\ \text{或}\ \ B\supset A$$

读作“A 包含于 B”或“B 包含 A”。例如 n 维实欧氏空间 R^n 上有一域 D，集合 D 的每一个元素都属于集合 R^n，则称集合 D 为集合 R^n 的子集，记为

$$D\subset R^n$$

当 x 是集合D 的元素时，即 x 属于D，记为 $x \in D$。又因集合 D 是集合R^n 的子集，故记做

$$x \in D \subset R^n$$

第三节　矢量及其运算

一、矢量的表示法

二维矢量　在平面上取一固定点 O（图 2-1），从 O 点做一有向线段OA，这样的线段叫矢量，记做$\overrightarrow{OA}$。O 为矢量的始点（或原点），A 为矢量的端点。若过 O 点建立直角坐标系 xOy，则向量$\overrightarrow{OA}$的大小和方向由（x_A，y_A）所确定。称 x_A 和 y_A 为这个向量沿坐标轴方向的分量。这样，平面上任意一点 P 唯一对应一个矢量$\overrightarrow{OP}$。矢量的端点对应着一对实数（x，y），即为 P 点的坐标。从而建立了以坐标（x，y）为标记的平面上点 P 与向量$\overrightarrow{OP}$的一一对应关系。为了今后研究的方便，把P点所确定的矢量用列矩阵的形式表示出来，记作

$$\overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{2-24}$$

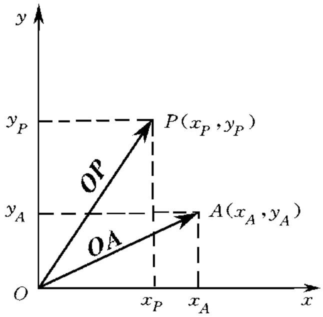

图　2-1

因此，可以把平面看成是由所有以 O 点为开始点的矢量（x　y）^{T} 的集合，它构成了一个二维矢量空间，故将平面矢量称为二维矢量。

由此推理，在直线上构成的矢量集合是一维矢量空间，而在一般三坐标空间的矢量集合则是三维矢量空间。

对于三维矢量空间能产生直观的认识，易于理解。对于更多维的空间向量就不易直观描述了。但为了说明某些问题。仍用几何描述把矢量的概念推广到多维空间中去。

一般地，以一组有序的 n 个数x_1，x_2，…，x_n 表示一个矢量，$\boldsymbol{X}$ 用列矩阵表示，记作

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = (x_1, x_2, \cdots, x_n)^T \tag{2-25}$$

它是一个 n 维矢量，x_1、x_2、…、x_n 是这个矢量的几个坐标轴分量。n 维矢量的全体组成了n 维空间，于是一切有序的几个实数组成的集合构成了 n 维矢量的欧氏空间，记作 R^n。X 是该集合中的元素，则

$$X \in R^n$$

二、矢量的相等，零矢量，单位坐标矢量

1．矢量的相等

设 $\boldsymbol{X}$，$\boldsymbol{Y}$ 是n 维空间的两个矢量，它们的分量分别是 x_1，x_2，…，x_n 和y_1，y_2，…，y_n，记作：

$$\boldsymbol{X} = (x_1, x_2, \cdots, x_n)^T \quad \boldsymbol{Y} = (y_1, y_2, \cdots, y_n)^T$$

显然，若两矢量相等，则有 $x_i = y_i$，$i = 1$，2…，n，并记作