高等学校经济与管理类系列教材

线性代数及其MATLAB实验

编 著◇李继根

华东师范大学出版社

高等学校经济与管理类系列教材

线性代数及其MATLAB实验

编　著◇李继根

华东师范大学出版社

图书在版编目(CIP)数据

线性代数及其MATLAB实验/李继根编著．—上海：华东师范大学出版社，2016
ISBN 978-7-5675-5815-1

Ⅰ．①线… Ⅱ．①李… Ⅲ．①线性代数—计算机辅助计—Matlab软件 Ⅳ．①O151.2-39

中国版本图书馆CIP数据核字(2016)第259742号

线性代数及其MATLAB实验

编　　著　李继根
项目编辑　孙小帆
特约审读　李　帆
版式设计　卢晓红
封面设计　俞　越

出版发行　华东师范大学出版社
社　　址　上海市中山北路3663号　邮编200062
网　　址　www.ecnupress.com.cn
电　　话　021-60821666　行政传真　021-62572105
客服电话　021-62865537　门市(邮购)电话　021-62869887
地　　址　上海市中山北路3663号华东师范大学校内先锋路口
网　　店　http://hdsdcbs.tmall.com

印 刷 者　常熟高专印刷有限公司
开　　本　787×1092　16开
印　　张　21
字　　数　473千字
版　　次　2017年5月第1版
印　　次　2017年5月第1次
书　　号　ISBN 978-7-5675-5815-1/O·271
定　　价　46.00元

出 版 人　王　焰

前 言

比起微积分和概率统计，在几门大学数学基础课中，线性代数目前远未达到其应有的地位. 与此同时，学生也普遍感到这门课抽象、繁琐、枯燥和无用. 我们认为课程内容的艰深和教育生态的制约是造成这种局面的主要原因.

线性代数研究的是线性空间及其扩充——模，以及作用在线性空间上的线性变换. 事实上，线性代数是“第二代数学模型”的典范，采用公理化体系来演绎的话，只要几页篇幅就能阐述和证明线性代数的主要结论. 作为重视这种演绎体系的极端体现，法国在基础教育领域开展过“新数学运动”. 由于一味注重形式上的严格性，忽视乃至抹杀数学的直觉性，使得学生恍如杂技中表演钻火圈游戏的小白鼠，在考试的皮鞭挥舞之下拼命奔跑，完全变成了枯燥规则的奴隶.

尽管许多人一度认为线性代数已经“寿终正寝”，但计算机的横空出世却让它“枯木逢春”，成为矩阵计算(又称数值线性代数)，进而演变成科学计算中的得力工具，并进一步促进了各种化解“维数之咒”的新兴课题的蓬勃发展. 世界顶尖的数值分析学家 L・N・特雷弗腾早在 1997 年就深刻地指出：“如果除了微积分与微分方程之外，还有什么领域是数学科学的基础的话，那就是数值线性代数.” 2000 年评选出的“20 世纪 10 大算法”，其中有 3 个与矩阵计算直接相关(Krylov 子空间迭代法、矩阵计算的分解方法、QR 算法). 吴文俊先生更深刻地指出：“从数学有史料为依据的几千年发展过程来看，以公理化思想为主的演绎倾向与以机械化思想为主的算法倾向互为消长.”因此对待演绎体系与算法体系，合理的态度应该是取两者之长，兼收并蓄，而不能厚此薄彼，褒一贬一. 对算法体系的研究催生了各种程序库，进而催生了科学计算软件 MATLAB.

回顾历史，不难发现线性代数与 MATLAB 佳偶天成，在学习线性代数的同时使用 MATLAB，是天经地义的事. 事实上，尽管许多专业都有与 MATLAB 相结合的课程，但是它们都无法做到线性代数的矩阵向量思想与 MATLAB 的这种紧密结合.

有鉴于此，美国于 1990 年代实施了 ATLAST(Augment the Teaching of Linear Algebra using Software Tools，用软件工具增强线性代数教学). 20 年后，美国的线性代数教学已经普遍用计算机解决问题. 受其影响，西安电子科技大学的陈怀琛先生从 2005 年开始在工科线性代数教学中引入 MATLAB 机算，并逐步滢入线性代数改革的深水区.

正是基于上述分析，并结合我校实情，我们“抛弃各种有形和无形的思想枷锁”(钱旭红语)，从 2011 年起，以我校谢国瑞教授荣获国家优秀教材一等奖(2002 年)的《线性代数及其应用》为出发点，开设了“线性代数实验班”. 我们认为线性代数课程首先要弱化线性代数作为“第二代数学模型”的形式化特征，教学重点应放在“怎样融合线性代数的课程内容与数值线性代数为主的计算技术”，要将学生从繁冗的线性代数计算中解放出来，同时

兼顾学生继续深造(考研或出国等)的需求. 基于这种观念,同时为响应参加实验班的“小白鼠”们的强烈呼吁,我们编写了这本教材.

在本教材的体系、选材和编写中,我们力求突出以下特点:

1. 淡化形式,注重实质

代数概念是一种典型的操作性概念,反映的是一种潜在的运算过程,因为大多数代数概念既可以看作过程,也可以看作对象. 因此代数操作是一种形式的符号操作,是对形式的数学符号、概念和程序的操作,需要借助于解析式、记号、图像、示意图等外部表征,这就导致代数操作需要更高的抽象水平.

基于此,有人认为线性代数课程正是训练学生抽象思维的绝佳时机. 由此催生的教材,其典型模式多为以定义开头,继之以定理和公式,再辅以应用. 前已指出,按此风格来处理线性代数,极端者只需几页篇幅. 但这样的教材尽管精炼简洁,却经常让学生觉得从心理上难以接受,因为在他们看来,许多数学概念就像孙悟空一样,是从石头缝里蹦出来的. 关于这一点,孟岩在博文《理解矩阵》中给出了形象的描述. 更要命的是,这种干瘪枯燥急需教师的阐释发挥(注水),这自然加剧了学生对优质教育资源的争夺.

早在 1993 年,陈重穆、宋乃庆两位先生就大力呼吁要“淡化形式,注重实质”. 因为形式化是不能一蹴而就的,就像微积分,历经几百年才有了严格化的“$\varepsilon-\delta$”语言.

2. 注重启发式教学,力争将“冰冷的美丽”转变成“火热的思考”

作为启发式教学的重要辅助工具,教材必须充分反映学生的思维过程,要通过一系列启发性的问题和各种各样的尝试和想法,让学生在观察、比较和推理中形成结论. 我们在本书中,充分注意学生已有的基础和经验,注重采用多种方式自然地引入数学基本概念和基本方法. 例如,从“鸡兔同笼”问题引入矩阵概念,从解的简洁表示引入行列式,从行阶梯矩阵的秩引入任意矩阵的秩,从方程的冗余引入向量组的线性表示,从考察像与原像的最简关系引入特征对,等等.

关于范德蒙行列式的计算技巧,谢国瑞教授曾经将之比拟为“翻江倒海”,可见他的课堂必定是汪洋恣肆,脑洞大开. 数学教学是数学思维活动的教学. 教育的最终目的是点燃和激起学生火热的思考,进而对知识进行主动建构. 教学是教师展示自己对知识理解的舞台,教材亦当如此. 通过对课程内容进行启发式教学法加工,本书希望给学生展示一种更加合乎数学思维、更易接受的知识阐述方式,从而让学生不再惧怕乃至爱上线性代数这门实践性很强的理论性课程.

3. 适当渗入 MATLAB 实验和计算思维

对于教材中的许多例题,在给出手算解答的同时,也给出了 MATLAB 机算程序,目的是让学生鲜明地看到 MATLAB 机算是化繁为简的大杀器. 更进一步地,每章末都设计了一些 MATLAB 实验,这既是为了加深学生对相关知识的理解,同时也旨在培养学生进行简单程序设计的能力. 另外,许多课程内容的阐述中还渗入了计算思维的分析,比如拉普拉斯展开法和克莱姆法则何以被摒弃. 这些点到即止式的安排,除了能让学生初步掌握 MATLAB 计算工具,更能培养学生根据问题类型灵活选择恰当的计算技术(算法意识)

并编程实现机算的能力.

为方便读者查找程序，例题和实验解答的程序文件采用了近似的名称，例如，程序文件 ex1204. m 对应的是例 1. 2. 4，程序文件 sy1203. m 对应的是第 1 章实验 2 的第 3 小题. 另外，由于 MATLAB 对大小写敏感，因此自定义函数名采用首字母大写等命名方式，以区别于内置函数和 ATLAST 中的函数. 读者可向出版社或作者(jgli@ecust. edu. cn)发邮件索取本书的配套程序. 另外，有些证明的末尾添加了标记“■”，以示区隔。

4. 适度引入萌系风格的语言

要让这本书彻底做到画风清奇、骨骼清秀，臣妾真的做不到. 但纵观人类的文明史，可见语言本身就一直处在发展演变之中，权威性的《新华字典》也一直在不断地收录新词. 因此，放下高冷的身段，适度引入一些萌萌哒的语言，还是能够办到的. 当然，这种调侃乃至戏说，只是传播知识的一种手段，其终极目的，是为了拉近抽象的线性代数知识与学生之间的心理距离. 如果为了坚持维护数学这座“高堂华屋”的庄严神圣，一定要他们“沐浴更衣焚香斋戒”，不仅会事倍功半，而且早晚还会遭遇事与愿违甚至适得其反的结果. 杨过用剑，树枝即剑，关键在于观念. 我们认为，“玩数学”(陈省身语)比“学数学”更逼近数学学习的本质.

要特别感谢参加过线性代数实验班的几百只“小白鼠”，“天王盖地虎，宝塔镇河妖”曾经是大家的“接头暗号”. 没有和他们的那些鲜活的互动，没有他们的鼓励和鞭策，本书恐怕也不会面世.

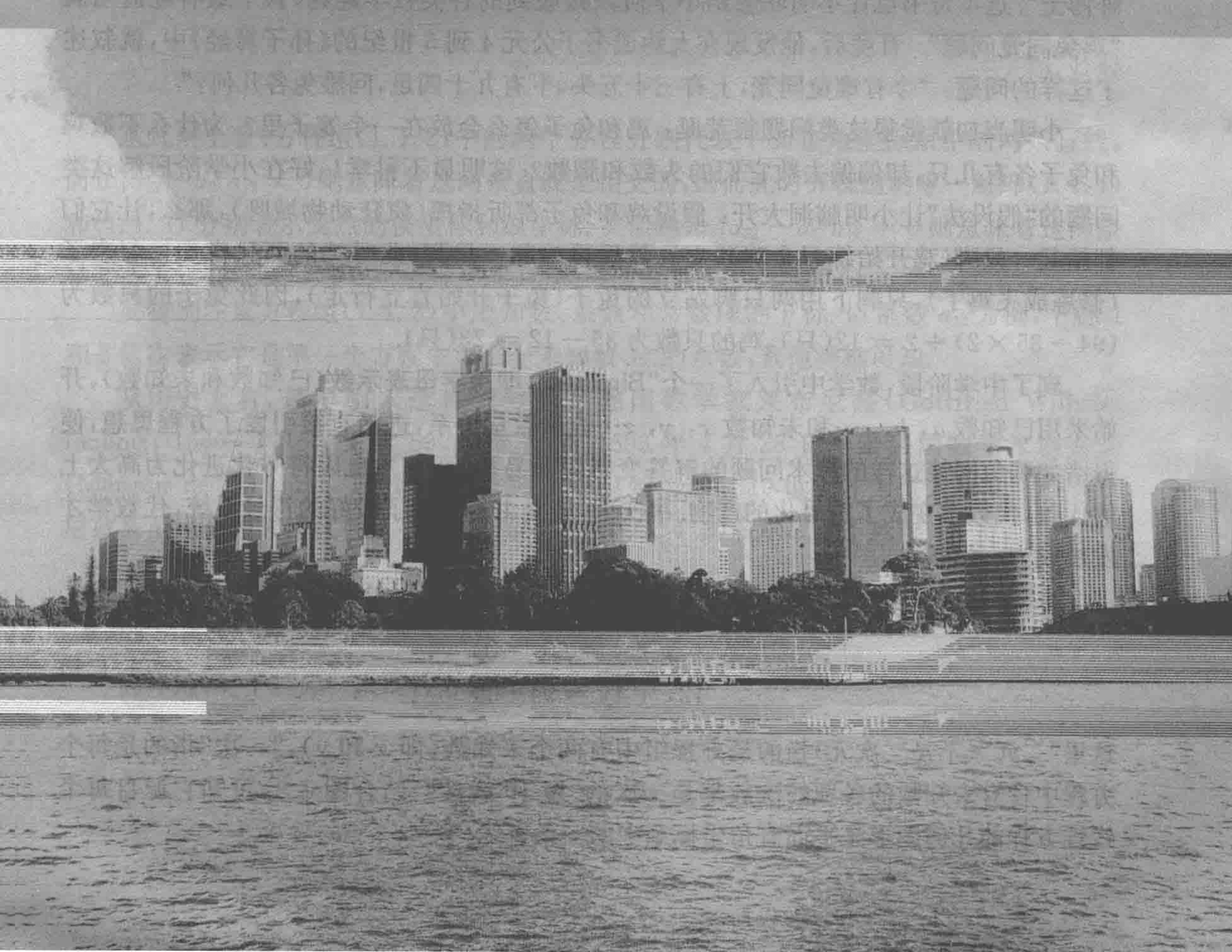
经典教材《社会心理学》一书的作者、著名心理学家戴维 · 迈尔斯写道：“我希望以一种充满热情的、富有个性的方式来讲述心理学，而不仅仅用一种严谨的科学方式”. 把心理学一词置换为线性代数，也就成了我们的热望.

李继根

于华东理工大学数学系

目 录

第 1 章 矩阵



面对求解线性方程组的千年难题，年轻的“矩阵博士”一上手就抓住了问题的核心，扔掉未知数等累赘，将之简洁地表示为增广矩阵，然后念动符号学的真诀，并使用矩阵运算等精良兵器，进一步将之转化为“符简理深”的矩阵方程，从而轻松地得到了线性方程组解的矩阵表示.

1.1 矩阵的概念与基本运算

1.1.1 从鸡兔同笼问题谈起

还记得大明湖畔的小明吗？最近他看完了一本神奇的书，即马丁·加德纳（Martin Gardner，1914—2010）的《矩阵博士的魔法数》（详见文献[67]）。这本奇书涉猎面很广，文笔隽永，诙谐有趣，可读性极强，人文根底非常深厚. 作为术数家，书中的主人公矩阵博士发现了许多离奇的数字现象与规律. 例如，布阿战争结束于 1902 年，而 $1902+1+9+0+2=1914$ 年正好是第一次世界大战开始之年. 再如，$13^2=169$，反过来 $961=31^2$，这里 31 正好也是 13 的逆序数. 另外 $1+6+9=16$，正好也是 $1+3=4$ 的平方数. 怎么会这么巧？简直太简直了！

小明被矩阵博士深深折服，开始以“矩阵博士”自诩，并把自己的各种签名改成了“矩阵博士”. 这本奇书也让小明联想到小学阶段接触到的各类数学趣题，其中最有趣的当属“鸡兔同笼问题”. 百度后，他发现在大约成书于公元 4 到 5 世纪的《孙子算经》中，就叙述了这样的问题：“今有雉兔同笼，上有三十五头，下有九十四足，问雉兔各几何？”

小明当初就觉得这类问题很荒诞：鸡和兔子怎么会放在一个笼子里？为什么不数鸡和兔子各有几只，却偏偏去数它们的头数和脚数？这明显不科学！好在小学阶段解这类问题的“假设法”让小明脑洞大开：假设鸡和兔子都听指挥（疯狂动物城呀），那么，让它们都抬起一只脚（鸡开始练习金鸡独立），然后再抬起一只脚，此时鸡的两只脚都抬起来了（修炼成飞鸡了），只剩下用两只脚站立的兔子（兔子开始直立行走），因此兔子的只数为 $(94-35\times 2)\div 2=12$（只），鸡的只数为 $35-12=23$（只）.

到了中学阶段，数学中引入了一个“Big idea”，即**用字母表示数**（已知数和未知数），开始采用已知数 a, b, c…和未知数 x, y, z…表示数量关系，进而直接引发了方程思想，使得诸如鸡兔同笼这样的算术问题的解答变得轻而易举，而算术也因符号化进化为高大上的代数，代数学变成了符号化的产物. 事实上，正是由于使用了较好的符号系统，代数学才发展成为一门学科.

小明知道，按中学的代数方法，令鸡有 x 只，兔子有 y 只，可得二元一次线性方程组

$$\begin{cases} x+y=35 & \text{①} \\ 2x+4y=94 & \text{②} \end{cases} \tag{1.1.1}$$

这里“二元”（不是二次元）指的是方程组中有两个未知数（即 x 和 y），“一次”指的是每个方程中含有未知数的各项的次数都是一次的. 至于“线性”，结合图 1.1，可知它源自每个线性方程的几何图形是平面直角坐标系中的一条直线.

采用加减消元法，②－①×2，即得 $2y=24$，故 $y=12$，代入①中，即得 $x=35-y=23$. 表现在几何图形上，就是图 1.1 中点 $P(23,12)$ 的横纵坐标分别代表鸡和兔子的只数.

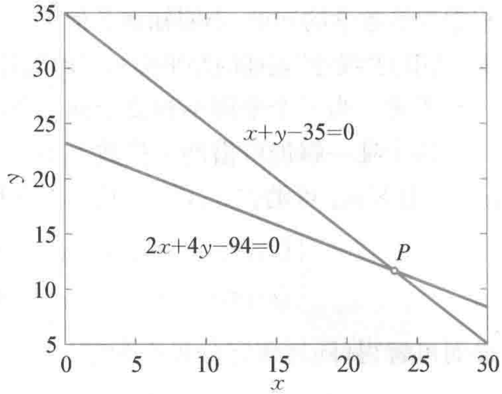


图 1.1　鸡兔同笼问题的几何意义

有了代数工具，就可以考察更一般的二元一次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2=b_1 & ① \\ a_{21}x_1+a_{22}x_2=b_2 & ② \end{cases} \tag{1.1.2}$$

其中的未知数是 x_1 和 x_2，a_{11}，a_{12}，a_{21}，a_{22} 是未知数的系数，b_1，b_2 是常数项.

仍然采用加减消元法，②×a_{11}－①×a_{21}，可消去 x_1，得：

$$(a_{11}a_{22}-a_{21}a_{12})x_2=b_2a_{11}-b_1a_{21}$$

当 $a_{11}a_{22}-a_{21}a_{12}\neq 0$ 时，即得

$$x_2=\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}} \tag{1.1.3}$$

将上式回代入方程①中，可得

$$x_1=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}} \tag{1.1.4}$$

从几何上看，方程组(1.1.2)中的两个方程分别代表平面直角坐标系中的两条直线，而 $a_{11}a_{22}-a_{21}a_{12}\neq 0$ 则意味着这两条直线是相交的，因而此时方程组的唯一解即(1.1.4)和(1.1.3)分别表示交点的横坐标和纵坐标. 类似地，$a_{11}a_{22}-a_{21}a_{12}=0$ 则意味着这两条直线是平行的(此时方程组无解)或重合的(此时方程组有无数解).

要特别注意方程组(1.1.2)中未知数、系数和常数项的下标. 以系数 a_{12} 为例，下标 1 和 2 依次表示它是第一个方程中第二个未知数 x_2 的系数，其他类推可知.

从历史上看，最早引入这种记号的是德国数学家莱布尼兹(Gottfried Wilhelm Leibniz，1646—1716). 文献[53]中指出，在 1693 年 4 月 28 日写给法国数学家洛必达(Marquis de l'Hôpital，1661—1704)的书信(这封信迟至 19 世纪才为世人所知)中，莱布尼兹甚至完全用数字代替了字母，从而将方程组(1.1.2)表示成了下列形式：

$$\begin{cases} 10+11x+12y=0 \\ 20+21x+22y=0 \end{cases}$$

同样地，对于三元一次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+a_{13}x_3=b_1 & ① \\ a_{21}x_1+a_{22}x_2+a_{23}x_3=b_2 & ② \\ a_{31}x_1+a_{32}x_2+a_{33}x_3=b_3 & ③ \end{cases} \tag{1.1.5}$$

根据高等数学的知识，小明知道几何上三个“线性”方程分别代表空间直角坐标系中的三个平面(这里的“线性”表征的是平面)，因此当且仅当它们相交于同一点时方程组(1.1.5)有唯一解.

思考 当三个平面不相交于同一点时，方程组(1.1.5)的解又是何种状况呢?

这个唯一解仍可借助于代数上的加减消元法来求出. 通过 ②$\times a_{11}-$①$\times a_{21}$ 及 ③$\times a_{11}-$①$\times a_{31}$ 可消去 x_1，将三元一次方程组(1.1.5) 化为二元一次方程组

$$\begin{cases}(a_{11}a_{22}-a_{21}a_{12})x_2+(a_{11}a_{23}-a_{21}a_{13})x_3=b_2a_{11}-b_1a_{21}\\(a_{11}a_{32}-a_{31}a_{12})x_2+(a_{11}a_{33}-a_{31}a_{13})x_3=b_3a_{11}-b_1a_{31}\end{cases}\tag{1.1.6}$$

进而可解得(同样假定分母不为零)

$$x_2=\frac{a_{11}b_2a_{33}+b_1a_{23}a_{31}+a_{13}a_{21}b_3-a_{13}b_2a_{31}-b_1a_{21}a_{33}-a_{11}a_{23}b_3}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}}\tag{1.1.7}$$

$$x_3=\frac{a_{11}a_{22}b_3+a_{12}b_2a_{31}+b_1a_{21}a_{32}-b_1a_{22}a_{31}-a_{12}a_{21}b_3-a_{11}b_2a_{32}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}}\tag{1.1.8}$$

将式(1.1.7)和(1.1.8)回代入方程组(1.1.5)中的方程①，可得

$$x_1=\frac{b_1a_{22}a_{33}+a_{12}a_{23}b_3+a_{13}b_2a_{32}-a_{13}a_{22}b_3-a_{12}b_2a_{33}-b_1a_{23}a_{32}}{a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}}\tag{1.1.9}$$

如此庞大的阵容，简直亮瞎了小明的钛合金眼. 显然上述的代数思路，尽管看上去似乎容易推广至四元、五元乃至 n 元线性方程组，但随着 n 的增大，问题马上就来了：计算太繁琐了，而且方程组的唯一解没有什么明显的规律. 这说明沿着上述思路求解线性方程组似乎困难重重. 至于前述的几何思路，三元及以下的线性方程组尚可想象其几何意义，但推广到四元、五元乃至 n 元后，高维的形象化问题将极大地考验人类的脑洞. 对于方程(组)的求解问题为何会贯穿人类的文明史，小明觉得这种困窘给出了一种解释.

1.1.2 矩阵的概念及其应用

两种思路都遇到了困难，小明只好重新回到方程组(1.1.2). 他注意到在式(1.1.3)和式(1.1.4)中，方程组的解 x_1，x_2 只与四个系数 a_{11}，a_{12}，a_{21}，a_{22} 及两个常数项 b_1，b_2 有关，如果略去未知数、加号和等号，则可将方程组(1.1.2)简化为如下的一张表

$$\begin{matrix}a_{11} & a_{12} & b_1\\ a_{21} & a_{22} & b_2\end{matrix}$$

方程组(1.1.2)显然与此表一一对应，因此对方程组(1.1.2)的研究就转化为对此表的研究.

小明马上发现这种想法具有一般性，也就是说可以推广到 n 元一次线性方程组(俗称“**正方形方程组**”，更一般地，线性代数中许多名词都涉及其几何形象)，即

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\\cdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n\end{cases}\tag{1.1.10}$$

的解也完全由未知数系数 $a_{ij}(i,j=1,2,\cdots,n)$ 和常数项 $b_i(i=1,2,\cdots,n)$ 所确定，

并且 a_{ij} 与 b_i 按原位置可排为一张表：

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{matrix}$$

因此研究线性方程组(1.1.10)就转化为研究这张表.

至此,对这种由一堆数构成的表,小明觉得有必要在数学上给出一般性的定义. 百度后,他得知这个概念就是矩阵(matrix).

定义 1.1.1(矩阵的定义)　由 mn 个数 a_{ij} $(i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$ 排成的 m 行 n 列的数表

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

称为一个 $m\times n$ **维(阶)矩阵**(matrix),其中的 a_{ij} 表示矩阵中**行号**为 i、**列号**为 j 的矩阵元素.

常用黑粗的英文大写字母 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\cdots$ 表示矩阵. 有时为了标明矩阵的行数 m、列数 n 以及矩阵元素的通项 a_{ij},也记作 $\boldsymbol{A}_{m\times n}$ 或 $(a_{ij})_{m\times n}$.

小明发现矩阵的维数或元素取特殊值时可得到大量的**特殊矩阵**.

从维数上考虑,当 $m=n$ 时,矩阵 $\boldsymbol{A}$ 称为 n **阶方阵**(square matrix),此时元素 a_{11}, a_{22}, $\cdots$, a_{nn} 称为矩阵 $\boldsymbol{A}$ 的**主对角元**(main diagonal),它们所在的对角线称为矩阵 $\boldsymbol{A}$ 的**主对角线**(main diagonal line),元素 a_{1n}, $a_{2,n-1}$, $\cdots$, a_{n1} 所在的对角线称为矩阵 $\boldsymbol{A}$ 的**次对角线**(anti-diagonal line);当 $m=1$ 时,矩阵 $\boldsymbol{A}$ 退化为 n **维行向量**(row vector);当 $n=1$ 时,矩阵 $\boldsymbol{A}$ 退化为 m **维列向量**(column vector);当 $m=n=1$ 时,矩阵 $\boldsymbol{A}$ 已退化为一个数(这一点很重要,因为矩阵可视为数的一种推广).

从元素上考虑,如果方阵 $\boldsymbol{U}$ 的主对角线以下的元素全为 0,即 $u_{ij}=0(i>j)$,则称 $\boldsymbol{U}$ 为**上三角矩阵**(upper triangular matrix),此时 $\boldsymbol{U}=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}$;类似地,如果方阵 $\boldsymbol{L}$ 满足 $l_{ij}=0(i<j)$,则称 $\boldsymbol{L}$ 为**下三角矩阵**(lower triangular matrix),此时 $\boldsymbol{L}=\begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}$;特别地,既是上三角矩阵又是下三角矩阵的方阵 $\boldsymbol{D}=$

$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$称为**对角矩阵**(diagonal matrix),也记作$\boldsymbol{D}=\mathrm{diag}(d_1, d_2, \cdots, d_n)$;如果对角矩阵$\boldsymbol{D}$的主对角元全部相等,即$\boldsymbol{D}=\mathrm{diag}(a, a, \cdots, a)$,则称为**数量阵**(scalar matrix),特别地,当$a=1$时,称对角矩阵$diag(1, 1, \cdots, 1)$为**单位矩阵**(identity matrix),记为$\boldsymbol{I}$.

另外,每个元素都为 0 的矩阵(未必是方阵)称为**零矩阵**(zero matrix),记为$\boldsymbol{O}$.

思考 为什么要引入新记号$\mathrm{diag}(d_1, d_2, \cdots, d_n)$对角矩阵$\boldsymbol{D}$来表示?提示:周星星同学的电影《美人鱼》,主题是什么?

小明很快发现了几个需要注意的问题:

(1) 在高等数学中,向量$\{a_1, a_2\}$用的是大括号,在几何上对应的是某个坐标系中以原点O为起点,以$P(a_1, a_2)$为终点的一条有向线段$\overrightarrow{OP}$,即代数向量$\{a_1, a_2\}\leftrightarrow$几何向量(即有向线段)$\overrightarrow{OP}$;起点为$P(a_1, a_2)$,终点为$Q(b_1, b_2)$的几何向量(即有向线段)$\overrightarrow{PQ}$,其代数表示则为$\{b_1-a_1, b_2-a_2\}$.

高等数学中不区分行向量与列向量,即将行列向量都统一表示为行向量. 线性代数中显然区分了行向量与列向量,而且通常用黑粗的希腊字母表示向量,例如n维行向量$\boldsymbol{\alpha}=(a_1, a_2, \cdots, a_n)$. 由于向量不再使用大括号,因此$(a_1, a_2, \cdots, a_n)$既可以表示某个点$P$的坐标(可视为点向量),也可以表示几何向量(即有向线段)$\overrightarrow{OP}$,还可视为此几何向量的代数表示. 抽取出这种代数表示的本质,即n**元有序数组**$(a_1, a_2, \cdots, a_n)$,就得到了向量(vector)这个更抽象的数学概念. 当然,作为约定俗成的习惯,除非特别指明,线性代数中的向量一般指的是列向量.

(2) 在物理中,与只有大小的标量(scalar)相对应,一般称向量为矢量(vector),泛指既有大小又有方向的物理量(如速度,位移等). 实际应用中遇到的一般是物理矢量. 几何向量可看作物理矢量的可视化,而代数向量则可视为物理矢量的运算工具.

(3) 基于维数视角,可以将序列"长方阵→方阵→行向量或列向量→数"视为矩阵维数在逐渐收缩(请脑补出动画)或者在逐步特殊化,其逆序列"数→行向量或列向量→方阵→长方阵"则可视为矩阵维数在逐渐扩张(再请脑补出动画)或者在逐步一般化. 众所周知,一般三角形的边长都满足三角不等式(任意两边之和大于第三边),因此直角三角形的边长也满足三角不等式,但直角三角形的勾股定理这个特殊性质,推广到一般三角形后,就不再满足了. 按此逻辑,对长方阵成立的性质,对方阵、向量和数应当也成立;反之对向量和方阵成立的性质,推广到长方阵,则未必成立. 在线性代数中,这种特殊化与一般化之间的逻辑关系很重要.

(4) 基于元素视角,可以将序列"方阵→上(下)三角阵→对角阵→数量阵→单位阵"视为矩阵元素取值的逐步特殊化,其逆序列"单位阵→数量阵→对角阵→上(下)三角阵→方阵"则可视为矩阵元素取值的逐步一般化. 注意前述的特殊化与一般化之间的逻辑,对这些矩阵也是成立的.

(5) 可以从三个层面来理解矩阵：从宏观的符号层面看，矩阵就是一个“超数”，一个“完全的抽象物”$\boldsymbol{A}$，一个具有某些指定运算的总体性的数学对象；从微观的元素层面看，矩阵就是“一堆数”a_{ij}，由许多个体按一定位置关系排列而成；从中观的列(行)向量层面看，矩阵就是由各列(行)组成的一组向量，即列(行)向量组. 掌握矩阵三个层面之间的关系，并能够灵活地互相转换，在线性代数中至关重要.

(6) 对于上(下)三角矩阵和对角矩阵等特殊矩阵，有时为了凸显它们的特殊结构，会略去其中的零元素，而代之以空白.

事实上，矩阵概念是线性代数中最重要的概念之一. 它最早由英国数学家凯莱(Arthur Cayley，1821—1895)提出于19世纪40年代，并由他的好基友西尔维斯特(James Joseph Sylvester，1814—1897)用matrix命名. Matrix的英文本意是“子宫、控制中心的母体、孕育生命的地方”. 在电影《黑客帝国》(The Matrix，1999)中，“母体”(Matrix)就是一套复杂的模拟系统程序，它是由具有人工智能的机器建立的，模拟了人类以前的世界，用以控制人类. 矩阵的数学定义很好地解释了母体制造世界的数学逻辑基础.

例 1.1.1(囚徒困境)　警方逮捕了甲、乙两名嫌疑犯，但却没有足够证据指控二人入罪. 于是警方分开囚禁这两名嫌疑犯，并分别和二人见面，然后向双方提供以下相同的选择：

(1) 若一人认罪并作证检控对方(称为“背叛”对方)，而对方保持沉默，此人将即时获释，至于沉默者将判监10年；

(2) 若二人都保持沉默(称为互相“合作”)，则二人都判监1年；

(3) 若二人都互相检举(称为互相“背叛”)，则二人都判监8年.

两名囚徒的上述困境，可用表1.1概述如下：

表1.1　囚徒困境

	甲沉默	甲背叛
乙沉默	甲乙各1年刑	甲释放，乙10年刑
乙背叛	甲10年刑，乙释放	甲乙各8年刑

若用矩阵分别表示甲、乙两人可能受到的处罚结果，则为$\begin{pmatrix}1 & 0\\10 & 8\end{pmatrix}$和$\begin{pmatrix}1 & 10\\0 & 8\end{pmatrix}$. 显然第一个矩阵是下三角矩阵，第二个则是上三角矩阵，它们的主对角元都是1和8.

例 1.1.2(病态矩阵)　在数值分析课程中，有一类著名的希尔伯特矩阵$\boldsymbol{H}=(h_{ij})$，是同一种典型的病态矩阵(ill-conditioned matrix). 例如三阶希尔伯特矩阵为

$$\boldsymbol{H}_3=\begin{pmatrix}1 & \frac{1}{2} & \frac{1}{3}\\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4}\\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5}\end{pmatrix}$$

仔细观察，小明发现 $\boldsymbol{H}_3$ 中的矩阵元素 h_{ij} 与其行号 i 和列号 j 都有关，即成立与数列类似的“通项公式”：$h_{ij}=\dfrac{1}{i+j-1}$.

思考 什么叫病态矩阵？为什么希尔伯特矩阵被视为病态矩阵？

温馨提示：亲爱的读者，小明不是病态，只是历经岁月沧桑，依然保有儿时的“十万个为什么”精神，请以实际行动支持他.

例 1.1.3(飞机航线问题) 某航空公司在 A，B，C，D 四城市之间开辟了若干航线，如图 1.2 所示. 如果从 A 到 B 有航班，则用箭头从 A 指向 B.

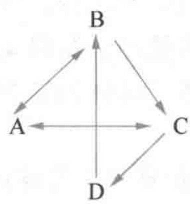


图 1.2 航班图

表 1.2 航班表

出发城市 \ 到达城市	A	B	C	D
A		√	√	
B	√		√	
C	√			√
D		√		

我们先用表 1.2 来表示航班图. 表格中√表示有航班. 显然把表中的√改成 1，空白地方填上 0，就可得到所谓的**邻接矩阵**(adjacency matrix，即元素仅取 0 或 1 的矩阵)：

$$\boldsymbol{G}=\begin{pmatrix}0&1&1&0\\1&0&1&0\\1&0&0&1\\0&1&0&0\end{pmatrix}$$

例 1.1.4(系数矩阵和增广矩阵) 按照矩阵的概念，“矩阵博士”小明马上发现由 m 个 n 元线性方程组成的线性方程组(俗称“**长方形方程组**”)

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\\cdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m\end{cases}\tag{1.1.11}$$

可以简化为 $m\times(n+1)$ 维矩阵(注意保留系数和常数的位置信息)

$$\overline{\boldsymbol{A}}=\left(\begin{array}{cccc:c}a_{11}&a_{12}&\cdots&a_{1n}&b_1\\a_{21}&a_{22}&\cdots&a_{2n}&b_2\\\vdots&\vdots&\ddots&\vdots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}&b_m\end{array}\right)$$

称为线性方程组(1.1.11)的**增广矩阵**(augmented matrix，即在系数矩阵右侧增加了一个常数列，读作“A 拔”)，其中左侧由系数构成的 $m\times n$ 维矩阵(注意保留位置信息)

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

称为线性方程组(1.1.11)的**系数矩阵**(coefficient matrix). 对于线性方程组而言,系数矩阵和增广矩阵的地位相当重要.

我们看到,矩阵博士小明甫一出场就亮瞎了很多人的双眼. 对于古老的线性方程组,他大胆地舍弃未知数等非本质因素,紧紧抓住系数、常数及其位置信息等本质因素,从而抽象出矩阵的概念. 利用矩阵的各种符号,今后讨论线性方程组时就不必再拘泥于细枝末节,即思维不再局限于元素层面,而是可以根据问题的需要在三个层面(即**符号层面、向量层面**和**元素层面**)之间灵活切换. 正如英国数学家怀特海(Alfred North Whitehead, 1861—1947)所指出的:"术语或符号的引入,往往是为了理论的易于表达和解决问题. 特别是在数学中,只要细加分析,即可发现符号化给数学理论的表述和论证带来极大的方便."事实上,矩阵符号的真正威力还没有被充分地显示出来,因为只有为矩阵附加上各种运算,才能充分领略矩阵的伟大、神秘和诱人.

1.1.3　矩阵的代数运算:线性运算

小明知道数具有加减乘除乘方开方等六则运算,向量则具有加减、数乘、数量积(点积)和向量积(叉积)等运算,而矩阵既然是数和向量的推广,因此他自然想将这些运算尽可能推广到矩阵中. 当然,按照前述的特殊化与一般化逻辑,当矩阵特殊化为向量(行数为 1 或列数为 1 的矩阵)乃至数(即一阶矩阵)时,这些矩阵运算必须与数或向量的相关运算吻合.

事实上,上述这种推广要求就是**固本原则**,这是丹齐克(Tobias Dantzig, 1884—1956)在名著《数:科学的语言》中反复提及的术语. 丹齐克指出,按照固本原则建立的代数运算规则,可以比拟为一个一心想扩张的国家所采取的政策,因为在扩张的同时,这个国家又希望永久保存那些使其强大的固有法律.

小明首先给出了矩阵相等的定义,并据此给出了矩阵的线性运算.

定义 1.1.2(同维矩阵和矩阵的相等)　行数对应相等,同时列数也对应相等的两个矩阵称为**同维矩阵**. 给定两个同维的 $m\times n$ 矩阵 $A=(a_{ij})$ 和 $B=(b_{ij})$. 如果

$$a_{ij}=b_{ij}, \ (i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$$

则称矩阵 A 与矩阵 B 相等,记作 $A=B$.

注意领略符号化的巨大威力,即 mn 个式子 $a_{ij}=b_{ij}$ 被简练地表达为一个式子 $A=B$. 显然,数(即一阶矩阵)a, b 相等的记号 $a=b$,已经被完美地推广到了矩阵相等,即 $A=B$,同时完美地遵循了固本原则.

定义 1.1.3(矩阵的加法) 两个同维的 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 的和记作 $\boldsymbol{A}+\boldsymbol{B}$,定义为

$$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})$$

定义 1.1.4(矩阵的数乘) 矩阵 $\boldsymbol{A}=(a_{ij})$ 与数 λ 的乘积记作 $\lambda\boldsymbol{A}$ 或 $\boldsymbol{A}\lambda$,定义为

$$\lambda\boldsymbol{A}=(\lambda a_{ij})$$

显然只有同维的矩阵方可相加,其结果即为它们对应位置上的元素分别相加,而且还是同维的矩阵;数乘运算则要注意"阳光普照",即矩阵的每一个元素都要乘上同一个数.

另外,对矩阵的加法运算 $\boldsymbol{A}+\boldsymbol{B}$ 和数乘运算 $\lambda\boldsymbol{A}$,当矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 特殊为向量 $\boldsymbol{a}$, $\boldsymbol{b}$ 时,两者分别特殊为向量的加法运算 $\boldsymbol{a}+\boldsymbol{b}$ 和数乘运算 $\lambda\boldsymbol{a}$;当矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 进一步特殊为数 a, b 时,两者显然就是数的加法运算 $a+b$ 和乘法运算 λa. 仍然完美地遵循了固本原则.

在矩阵的数乘运算 $\lambda\boldsymbol{A}$ 中,小明取 $\lambda=-1$,得到了矩阵 $\boldsymbol{A}=(a_{ij})$ 的**负矩阵** $-\boldsymbol{A}=(-a_{ij})$,进而定义了矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的**减法**,即

$$\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-\boldsymbol{B})=(a_{ij}-b_{ij})$$

矩阵的加(减)法运算和数乘运算统称为矩阵的**线性运算**. 小明注意到,如果更宽泛地看待"线性运算",那么高等数学中已经多次出现过线性运算,例如两可导(积)函数和的导数(积分)等于导数(积分)的和. 问题是那里的运算对象是函数,而不是这里的矩阵.

思考 什么是线性? 什么又是非线性?

在上述运算(比如加法)中,小明发现形式上无非就是把 $a+b$ 中的字母 a 和 b 分别替换成了同维矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 这两个"超数",因此他大胆地假设:字母代数的许多性质(比如运算律)都能推广到更一般的矩阵代数! 按照前述的特殊化与一般化逻辑,他明白这种类比推广是从"特殊到一般"的思维方式,不能保证结果一定准确,因此对每一个推广,都必须"小心求证"(胡适语).

根据定义,不难证明矩阵代数也满足字母代数的以下运算规律,其中 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 是同维矩阵,λ, μ 为任意实数:

(1) 交换律 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$;

(2) 结合律 $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$, $(\lambda\mu)\boldsymbol{A}=\lambda(\mu\boldsymbol{A})=\mu(\lambda\boldsymbol{A})$;

(3) 分配律 $\lambda(\boldsymbol{A}\pm\boldsymbol{B})=\lambda\boldsymbol{A}\pm\lambda\boldsymbol{B}$, $(\lambda\pm\mu)\boldsymbol{A}=\lambda\boldsymbol{A}\pm\mu\boldsymbol{A}$;

(4) $\boldsymbol{A}\pm\boldsymbol{O}=\boldsymbol{A}$, $\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{O}$;

(5) $0\cdot\boldsymbol{A}=\boldsymbol{O}$, $1\cdot\boldsymbol{A}=\boldsymbol{A}$.

例 1.1.5 设 $\boldsymbol{A}=\begin{pmatrix}1 & 3\\ -1 & 2\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}7 & -4\\ -8 & -5\end{pmatrix}$, $\boldsymbol{\Lambda}=\mathrm{diag}(1,-2)$.

求:(1) $\boldsymbol{A}-\boldsymbol{B}+3\boldsymbol{I}$; (2) $2\boldsymbol{A}-3\boldsymbol{\Lambda}$.

分析:同字母代数一样,矩阵代数的运算次序也是先乘后加. 题中 $\boldsymbol{A}$, $\boldsymbol{B}$ 是二阶方阵,